(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2022-0038362
(43) 공개일자 2022년03월28일

(51) 국제특허분류(Int. Cl.)
*C12N 15/86* (2006.01) *A61K 48/00* (2006.01)
*C12N 15/113* (2010.01) *C12N 15/90* (2006.01)
*C12N 9/22* (2006.01)
(52) CPC특허분류
*C12N 15/86* (2013.01)
*A61K 48/0091* (2013.01)
(21) 출원번호 10-2022-7003415
(22) 출원일자(국제) 2020년07월02일
심사청구일자 없음
(85) 번역문제출일자 2022년01월28일
(86) 국제출원번호 PCT/US2020/040756
(87) 국제공개번호 WO 2021/003432
국제공개일자 2021년01월07일
(30) 우선권주장
62/869,907 2019년07월02일 미국(US)
*(뒷면에 계속)*

(71) 출원인
**프레드 헛친슨 켄서 리서치 센터**
미국 워싱턴 98109, 시애틀, 1100 페어뷰 애버뉴 엔.,
**유니버시티 오브 워싱톤**
미국 98105-4721 워싱턴주 시애틀 스위트 400 루즈벨트 웨이 노쓰이스트 4545
(72) 발명자
**키엠, 한스-피터**
미국 98109 워싱턴주 시애틀 페어뷰 애비뉴 1100
**리버, 안드레**
미국 98105 워싱턴주 시애틀 루스벨트 웨이 엔이 4545 스위트 400
*(뒷면에 계속)*
(74) 대리인
**양영준, 김영**

전체 청구항 수 : 총 58 항

(54) 발명의 명칭 **재조합 AD35 벡터 및 관련 유전자 요법 개선**

## *(57) 요 약*

본 개시내용은 무엇보다도 헬퍼-의존적 아데노바이러스 혈청형 35(Ad35) 벡터를 제공한다. 다양한 구현예에서, 헬퍼-의존적 Ad35 벡터는 치료적 페이로드를 이를 필요로 하는 대상체에게 전달하는 데 사용될 수 있다. 예시적인 페이로드는 대체 단백질, 항체, CAR, TCR, 작은 RNA, 및 게놈 편집 시스템을 암호화할 수 있다. 특정 구현예에서, 헬퍼-의존적 Ad35 벡터는 페이로드를 숙주 세포 게놈 내로 통합하기 위해 조작된다. 본 개시내용은 헬퍼-의존적 Ad35 벡터를 이를 필요로 하는 대상체에게 투여하는 것을 포함하는 유전자 요법 방법을 추가로 포함한다.

(52) CPC특허분류
***C12N 15/113*** (2013.01)
***C12N 15/907*** (2013.01)
***C12N 9/22*** (2013.01)
*C12N 2330/50* (2013.01)
*C12N 2710/10343* (2013.01)
*C12N 2710/10352* (2013.01)
*C12N 2800/50* (2013.01)
*C12N 2800/80* (2013.01)
*C12N 2800/90* (2013.01)

(72) 발명자
**리, 창**
미국 98105 워싱턴주 시애틀 루스벨트 웨이 엔이 4545 스위트 400
**왕, 홍지에**
미국 98109 워싱턴주 시애틀 루스벨트 웨이 엔이 4545 스위트 400

(30) 우선권주장
62/935,507 2019년11월14일 미국(US)
63/009,385 2020년04월13일 미국(US)

# 명 세 서

## 청구범위

### 청구항 1

하기를 포함하는, 재조합 아데노바이러스 혈청형 35(Ad35) 벡터 생산 시스템:

다음을 포함하는 재조합 Ad35 헬퍼 게놈:

Ad35 섬유 샤프트를 암호화하는 핵산 서열;

Ad35 섬유 놉을 암호화하는 핵산 서열; 및

Ad35 패키징 서열의 적어도 일부에 플랭킹하는 재조합효소 지시 반복부(DR), 및

다음을 포함하는 재조합 헬퍼 의존적 Ad35 공여자 게놈:

5' Ad35 도립된 말단 반복부(ITR);

3' Ad35 ITR;

Ad35 패키징 서열; 및

적어도 하나의 이종 발현 산물을 암호화하는 핵산 서열.

### 청구항 2

하기를 포함하는, 재조합 아데노바이러스 혈청형 35(Ad35) 헬퍼 벡터:

Ad35 섬유 샤프트;

Ad35 섬유 놉; 및

Ad35 패키징 서열의 적어도 일부에 플랭킹하는 재조합효소 지시 반복부(DR)를 포함하는 Ad35 게놈.

### 청구항 3

하기를 포함하는, 재조합 아데노바이러스 혈청형 35(Ad35) 헬퍼 게놈:

Ad35 섬유 샤프트를 암호화하는 핵산 서열;

Ad35 섬유 놉을 암호화하는 핵산 서열; 및

Ad35 패키징 서열의 적어도 일부에 플랭킹하는 재조합효소 지시 반복부(DR).

### 청구항 4

하기를 포함하는, 재조합 헬퍼 의존적 아데노바이러스 혈청형 35(Ad35) 공여자 벡터:

다음을 포함하는 핵산 서열:

5' Ad35 도립된 말단 반복부(ITR);

3' Ad35 ITR;

Ad35 패키징 서열; 및

적어도 하나의 이종 발현 산물을 암호화하는 핵산 서열로, 여기서 게놈은 Ad35 바이러스 구조적 단백질을 암호화하는 핵산 서열을 포함하지 않는 것인 핵산 서열; 및

Ad35 섬유 샤프트 및/또는 Ad35 섬유 놉.

### 청구항 5

5' Ad35 도립된 말단 반복부(ITR);

3' Ad35 ITR;

Ad35 패키징 서열; 및

적어도 하나의 이종 발현 산물을 암호화하는 핵산 서열

을 포함하는 재조합 헬퍼 의존적 아데노바이러스 혈청형 35(Ad35) 공여자 게놈이며, 여기서 Ad35 공여자 게놈은 야생형 Ad35 게놈에 의해 암호화된 발현 산물을 암호화하는 핵산 서열을 포함하지 않는 것인, 재조합 헬퍼 의존적 Ad35 공여자 게놈.

**청구항 6**

재조합 헬퍼 의존적 아데노바이러스 혈청형 35(Ad35) 공여자 벡터를 생산하는 방법으로, 상기 방법은 세포의 배양물로부터 재조합 헬퍼 의존적 Ad35 공여자 벡터를 단리하는 단계를 포함하며, 여기서 세포는 하기를 포함하는 것인, 방법:

다음을 포함하는 재조합 Ad35 헬퍼 게놈:

Ad35 섬유 샤프트를 암호화하는 핵산 서열;

Ad35 섬유 놉을 암호화하는 핵산 서열; 및

Ad35 패키징 서열의 적어도 일부에 플랭킹하는 재조합효소 지시 반복부(DR) 및

다음을 포함하는 재조합 헬퍼 의존적 Ad35 공여자 게놈:

5' Ad35 도립된 말단 반복부(ITR);

3' Ad35 ITR;

Ad35 패키징 서열; 및

적어도 하나의 이종 발현 산물을 암호화하는 핵산 서열.

**청구항 7**

하기를 포함하는, 재조합 아데노바이러스 혈청형 35(Ad35) 생산 시스템:

다음을 포함하는 재조합 Ad35 헬퍼 게놈:

Ad35 섬유 샤프트를 암호화하는 핵산 서열;

Ad35 섬유 놉을 암호화하는 핵산 서열; 및

Ad35 패키징 신호를 기능적으로 방해하지만 5' Ad35 도립된 말단 반복부(ITR)는 방해하지 않는 Ad35 게놈의 5' 단부의 550 개 뉴클레오티드 내의 재조합효소 지시 반복부(DR), 및

다음을 포함하는 재조합 Ad35 공여자 게놈:

5' Ad35 ITR;

3' Ad35 ITR;

Ad35 패키징 서열; 및

적어도 하나의 이종 발현 산물을 암호화하는 핵산 서열.

**청구항 8**

하기를 포함하는, 재조합 아데노바이러스 혈청형 35(Ad35) 헬퍼 벡터:

Ad35 섬유 샤프트;

Ad35 섬유 놉; 및

Ad35 패키징 신호를 기능적으로 방해하지만 5' Ad35 도립된 말단 반복부(ITR)는 방해하지 않는 Ad35 게놈의 5' 단부의 550 개 뉴클레오티드 내의 재조합효소 지시 반복부(DR)를 포함하는 Ad35 게놈.

**청구항 9**

하기를 포함하는, 재조합 아데노바이러스 혈청형 35(Ad35) 헬퍼 게놈:

Ad35 섬유 샤프트를 암호화하는 핵산 서열;

Ad35 섬유 놉을 암호화하는 핵산 서열; 및

Ad35 패키징 신호를 기능적으로 방해하지만 5' Ad35 도립된 말단 반복부(ITR)는 방해하지 않는 Ad35 게놈의 5' 단부의 550 개 뉴클레오티드 내의 재조합효소 지시 반복부(DR).

**청구항 10**

재조합 헬퍼 의존적 아데노바이러스 혈청형 35(Ad35) 공여자 벡터를 생산하는 방법으로, 상기 방법은 세포의 배양물로부터 재조합 헬퍼 의존적 Ad35 공여자 벡터를 단리하는 단계를 포함하며, 여기서 세포는 하기를 포함하는 것인, 방법:

다음을 포함하는 재조합 Ad35 헬퍼 게놈:

Ad35 섬유 샤프트를 암호화하는 핵산 서열;

Ad35 섬유 놉을 암호화하는 핵산 서열; 및

Ad35 패키징 신호를 기능적으로 방해하지만 5' Ad35 도립된 말단 반복부(ITR)는 방해하지 않는 Ad35 게놈의 5' 단부의 550 개 뉴클레오티드 내의 재조합효소 지시 반복부(DR), 및

다음을 포함하는 재조합 Ad35 공여자 게놈:

5' Ad35 ITR;

3' Ad35 ITR;

Ad35 패키징 서열; 및

적어도 하나의 이종 발현 산물을 암호화하는 핵산 서열.

**청구항 11**

제1항 내지 제4항 및 제6항 내지 제10항 중 어느 한 항에 있어서,

상기 Ad35 섬유 놉이 야생형 Ad35 섬유 놉이거나, 또는

상기 Ad35 섬유 놉이 조작된 Ad35 섬유 놉이며, 여기서 조작된 섬유 놉은 CD46에 대한 섬유 놉의 친화도를 증가시키는 돌연변이를 포함하는 것인, 재조합 Ad35 벡터 생산 시스템, 헬퍼 벡터, 헬퍼 게놈, 공여자 벡터, 또는 방법.

**청구항 12**

제11항에 있어서, 상기 돌연변이가

Ile192Val, Asp207Gly(또는 Glu207Gly), Asn217Asp, Thr226Ala, Thr245Ala, Thr254Pro, Ile256Leu, Ile256Val, Arg259Cys, 및 Arg279His로부터 선택된 돌연변이를 포함하거나; 또는 돌연변이 Ile192Val, Asp207Gly(또는 Glu207Gly), Asn217Asp, Thr226Ala, Thr245Ala, Thr254Pro, Ile256Leu, Ile256Val, Arg259Cys, 및 Arg279His 각각을 포함하는 것인, 재조합 Ad35 벡터 생산 시스템, 헬퍼 벡터, 헬퍼 게놈, 공여자 벡터, 또는 방법.

**청구항 13**

제1항, 제4항 내지 제7항, 및 제10항 중 어느 한 항에 있어서, 상기 이종 발현 산물이 조절 서열과 작동가능하게 연결된 치료 발현 산물을 포함하며, 임의적으로 여기서 치료 발현 산물은 하기를 포함하는 것인, 재조합 Ad35 벡터 생산 시스템, 공여자 게놈, 공여자 벡터, 또는 방법:

(a) β-글로빈 단백질 또는 γ-글로빈 단백질;

(b) 항체 또는 이의 면역글로불린 쇄로, 임의적으로 여기서 항체는 항-CD33 항체인 것;

(c) 제1 항체 또는 이의 면역글로불린 쇄 및 제2 항체 또는 이의 면역글로불린 쇄로, 임의적으로 여기서 항체는 항-CD33 항체인 것;

(d) CRISPR-연관 RNA-가이드된 엔도뉴클레아제 및/또는 가이드 RNA(gRNA)로, 임의적으로 여기서 CRISPR-연관 RNA-가이드된 엔도뉴클레아제는 Cas9 또는 cpf1을 포함하는 것;

(e) 염기 편집기 및/또는 gRNA로, 임의적으로 여기서 염기 편집기는 사이토신 염기 편집기(CBE) 또는 아데닌 염기 편집기(ABE)이며, 임의적으로 여기서 염기 편집기는 기능상실된 Cas9 및 기능상실된 cpf1로부터 선택된 촉매적으로 기능상실된 뉴클레아제를 포함하는 것;

(f) 바이러스 감염을 차단 또는 감소시키는 응고 인자 또는 단백질로, 임의적으로 여기서 치료 발현 산물은 인자 VII 대체 단백질 또는 인자 VIII 대체 단백질을 포함하는 것;

(g) 체크포인트 억제제;

(h) 키메라 항원 수용체 또는 조작된 T 세포 수용체; 또는

(i) γC, JAK3, IL7RA, RAG1, RAG2, DCLRE1C, PRKDC, LIG4, NHEJ1, CD3D, CD3E, CD3Z, CD3G, PTPRC, ZAP70, LCK, AK2, ADA, PNP, WHN, CHD7, ORAI1, STIM1, CORO1A, CIITA, RFXANK, RFX5, RFXAP, RMRP, DKC1, TERT, TINF2, DCLRE1B, SLC46A1, FancA, FancB, FancC, FancD1, FancD2, FancE, FancF, FancG, FancI, FancJ, FancL, FancM, FancN, FancO, FancP, FancQ, FancR, FancS, FancT, FancU, FancV, FancW, 가용성 CD40, CTLA, Fas L, PD-L1에 대한 항체, CD4에 대한 항체, CD5에 대한 항체, CD7에 대한 항체, CD52에 대한 항체, IL-1에 대한 항체, IL-2에 대한 항체, IL-4에 대한 항체, IL-6에 대한 항체, IL-10에 대한 항체, TNF에 대한 항체, 자가반응성 T 세포 상에 특이적으로 존재하는 TCR에 대한 항체, 글로빈 패밀리 유전자, WAS, phox, 디스트로핀, 피루베이트 키나제, CLN3, ABCD1, 아릴술파타제 A, SFTPB, SFTPC, NLX2.1, ABCA3, GATA1, 리보솜 단백질 유전자, TERT, TERC, DKC1, TINF2, CFTR, LRRK2, PARK2, PARK7, PINK1, SNCA, PSEN1, PSEN2, APP, SOD1, TDP43, FUS, 유비퀼린 2, 및/또는 C9ORF72로 이루어진 군으로부터 선택된 단백질로, 임의적으로 여기서 단백질은 FancA 단백질인 것.

**청구항 14**

제13항에 있어서, (d) 또는 (e)에서

상기 gRNA가 HBG1, HBG2, 및/또는 적혈구 인핸서 bcl11a의 표적 핵산 서열에 결합하며, 임의적으로 여기서 gRNA는 γ-글로빈의 발현을 증가시키도록 조작되는 것이거나; 또는

상기 gRNA가 CD33의 일부를 암호화하는 표적 핵산 서열에 결합하며, 임의적으로 여기서 CD33은 인간 CD33인, 재조합 Ad35 벡터 생산 시스템, 공여자 게놈, 공여자 벡터, 또는 방법.

**청구항 15**

제13항에 있어서, 상기 치료 발현 산물이 하기를 포함하는 것인, 재조합 Ad35 벡터 생산 시스템, 공여자 게놈, 공여자 벡터, 또는 방법:

β-글로빈 단백질 또는 γ-글로빈 단백질; 및

CRISPR-연관 RNA-가이드된 엔도뉴클레아제; 및 하기 중 1, 2, 또는 3 개를 포함하는 CRISPR 시스템:

HBG1의 표적 핵산 서열에 결합하는 gRNA;

HBG2의 표적 핵산 서열에 결합하는 gRNA; 및/또는

Bcl11a의 표적 핵산 서열에 결합하는 gRNA로,

임의적으로 여기서 gRNA는 γ-글로빈의 발현을 증가시키도록 조작되는 것.

**청구항 16**

제13항에 있어서, 상기 조절 서열(들)이 프로모터를 포함하며, 임의적으로 여기서 프로모터는 β-글로빈 프로모터이며, 임의적으로 여기서 β-글로빈 프로모터는 약 1.6 kb 길이이고/이거나 염색체 11의 위치 5228631-

5227023에 따른 핵산을 포함하는 것인, 재조합 Ad35 벡터 생산 시스템, 공여자 게놈, 공여자 벡터, 또는 방법.

**청구항 17**

제13항에 있어서, 상기 조절 서열(들)이 유전자좌 제어 영역(LCR)을 포함하며, 임의적으로 여기서 LCR은 β-글로빈 LCR인, 재조합 Ad35 벡터 생산 시스템, 공여자 게놈, 공여자 벡터, 또는 방법.

**청구항 18**

제13항에 있어서, 상기 β-글로빈 LCR이

HS1, HS2, HS3, 및 HS4를 포함하거나 또는 이로 이루어지는 β-글로빈 LCR DNAse I 과민성 부위(HS)를 포함하며, 임의적으로 여기서 β-글로빈 LCR은 약 4.3 kb의 길이를 갖는 것이거나;

HS1, HS2, HS3, HS4, 및 HS5를 포함하는 β-글로빈 LCR DNAse I HS를 포함하며, 임의적으로 여기서 β-글로빈 LCR은 약 21.5 kb의 길이를 갖는 것이거나; 또는

상기 β-글로빈 LCR이 염색체 11의 위치 5292319-5270789에 따른 서열을 포함하는 것인, 재조합 Ad35 벡터 생산 시스템, 공여자 게놈, 공여자 벡터, 또는 방법.

**청구항 19**

제13항 또는 제14항에 있어서, 상기 조절 서열(들)이 3'HS1을 포함하며, 임의적으로 여기서 3'HS1은 염색체 11의 위치 5206867-5203839에 따른 서열을 포함하는 것인, 재조합 Ad35 벡터 생산 시스템, 공여자 게놈, 공여자 벡터, 또는 방법.

**청구항 20**

제13항에 있어서, 상기 조절 서열(들)이 miRNA 결합 부위를 포함하며, 임의적으로 여기서

상기 miRNA 결합 부위는 관심 종에 의해 자연적으로 발현되는 miRNA에 대한 결합 부위이고/이거나;

상기 miRNA는 혈액 및 종양 미세환경 또는 표적 조직에서 차등 점유 프로파일을 입증하며, 임의적으로 여기서 점유 프로파일은 종양 미세환경 또는 표적 조직에서보다 혈액에서 더 높은 것이고/이거나;

상기 miRNA 결합 부위는 miR423-5, miR423-5p, miR42-2, miR181c, miR125a, 또는 miR15a 결합 부위를 포함하고/하거나;

상기 miRNA 결합 부위는 miR187 또는 miR218 결합 부위를 포함하는 것인, 재조합 Ad35 벡터 생산 시스템, 공여자 게놈, 공여자 벡터, 또는 방법.

**청구항 21**

제1항, 제4항 내지 제7항, 및 제10항 중 어느 한 항에 있어서, 상기 이종 발현 산물을 암호화하는 핵산이 통합 요소를 추가로 포함하는 페이로드의 부분이며, 임의적으로 여기서 통합 요소는 발현 산물을 포함하는 것인, 재조합 Ad35 벡터 생산 시스템, 공여자 게놈, 공여자 벡터, 또는 방법.

**청구항 22**

제21항에 있어서, 상기 통합 요소가 상동 재조합에 의한 표적 게놈으로의 통합을 위해 조작되고, 상기 통합 요소가 표적 게놈의 인접하게 연결된 서열에 상응하는 상동성 아암에 의해 플랭킹되며, 임의적으로 여기서

상동성 아암은 0.8 및 1.8 kb 사이인 것이고/이거나;

상동성 아암은 염색체 안전 항구 유전자좌에 플랭킹되는 표적 게놈의 핵산 서열에 상동이며, 임의적으로 여기서 안전 항구 유전자좌는 AAVS1, CCR5, HPRT, 또는 Rosa로부터 선택되는 것인, 재조합 Ad35 벡터 생산 시스템, 공여자 게놈, 공여자 벡터, 또는 방법.

**청구항 23**

제21항에 있어서, 상기 통합 요소가 전위에 의한 표적 게놈으로의 통합을 위해 조작되고, 상기 통합 요소가 트랜스포존 도립 반복부(IR)에 의해 플랭킹되며, 임의적으로 여기서 트랜스포존 IR은 재조합효소 DR에 의해 플랭

킹되는 것인, 재조합 Ad35 벡터 생산 시스템, 공여자 게놈, 공여자 벡터, 또는 방법.

**청구항 24**

제23항에 있어서,

상기 트랜스포존 IR이 잠자는 미녀(SB) IR이며, 임의적으로 여기서 SB IR은 pT4 IR이거나;

상기 트랜스포존 IR이 피기백, 마리너, 개구리 왕자, Tol2, TcBuster, 또는 spinON IR인, 재조합 Ad35 벡터 생산 시스템, 공여자 게놈, 공여자 벡터, 또는 방법.

**청구항 25**

제21항에 있어서, 트랜스포존 IR에 의해 플랭킹되는 통합 요소의 전위를 매개하는 트랜스포사제를 암호화하는 핵산을 포함하며, 임의적으로 여기서 트랜스포사제를 암호화하는 핵산은 지지 벡터 또는 지지 벡터 게놈으로 구성되는 것인, 재조합 Ad35 벡터 생산 시스템, 공여자 게놈, 공여자 벡터, 또는 방법.

**청구항 26**

제25항에 있어서, 상기 트랜스포사제가 잠자는 미녀, 피기백, 마리너, 개구리 왕자, Tol2, TcBuster, 또는 spinON 트랜스포사제이며, 임의적으로 여기서 트랜스포사제는 잠자는 미녀 100x(SB100x) 트랜스포사제인, 재조합 Ad35 벡터 생산 시스템, 공여자 게놈, 공여자 벡터, 또는 방법.

**청구항 27**

제25항 또는 제26항에 있어서, 상기 트랜스포사제를 암호화하는 핵산이 PGK 프로모터와 작동가능하게 연결되는 것인, 재조합 Ad35 벡터 생산 시스템, 공여자 게놈, 공여자 벡터, 또는 방법.

**청구항 28**

제1항 내지 제3항 및 제6항 내지 제10항 중 어느 한 항에 있어서, 상기 Ad35 패키징 서열의 적어도 일부에 플랭킹하고/하거나 Ad35 게놈의 5' 단부의 550 개 뉴클레오티드 내에 있고 Ad35 패키징 신호를 기능적으로 방해하지만 5' Ad35 ITR은 방해하지 않는 재조합효소 DR이 FRT, loxP, rox, vox, AttB, 또는 AttP 부위인, 재조합 Ad35 벡터 생산 시스템, 헬퍼 벡터, 헬퍼 게놈, 또는 방법.

**청구항 29**

제28항에 있어서, 상기 Ad35 패키징 서열의 적어도 일부의 절제를 위한 재조합효소를 암호화하는 핵산이 헬퍼 게놈을 포함하는 세포의 핵산 서열에 의해 암호화되는 것인, 재조합 Ad35 벡터 생산 시스템, 헬퍼 벡터, 헬퍼 게놈, 또는 방법.

**청구항 30**

제23항에 있어서, 상기 트랜스포존 IR에 플랭킹되는 재조합효소 DR이 FRT, loxP, rox, vox, AttB, 또는 AttP 부위인, 재조합 Ad35 벡터 생산 시스템, 헬퍼 벡터, 헬퍼 게놈, 또는 방법.

**청구항 31**

제21항에 있어서, 상기 통합 요소를 포함하는 핵산의 절제를 위한 재조합효소를 암호화하는 핵산이 지지 벡터 또는 지지 벡터 게놈으로 구성되는 것인, 재조합 Ad35 벡터 생산 시스템, 헬퍼 벡터, 헬퍼 게놈, 또는 방법.

**청구항 32**

제29항 또는 제31항에 있어서, 상기 재조합효소가 Flp, Cre, Dre, Vika, 또는 PhiC31 재조합효소인, 재조합 Ad35 벡터 생산 시스템, 헬퍼 벡터, 헬퍼 게놈, 또는 방법.

**청구항 33**

제32항에 있어서, 상기 재조합효소를 암호화하는 핵산이 EF1α 프로모터와 작동가능하게 연결되는 것인, 재조합 Ad35 벡터 생산 시스템, 헬퍼 벡터, 헬퍼 게놈, 또는 방법.

**청구항 34**

제21항에 있어서,

상기 페이로드가 이종 발현 산물을 포함하는 통합 요소를 포함하고,

상기 이종 발현 산물이 β-글로빈 프로모터 및 β-글로빈 긴 LCR와 작동가능하게 연결된 β-글로빈 단백질을 포함하고,

상기 통합 요소가 SB IR에 의해 플랭킹되고,

상기 SB IR이 재조합효소 DR에 의해 플랭킹되며, 임의적으로 여기서 재조합효소 DR은 FRT 부위인, 재조합 Ad35 벡터 생산 시스템, 헬퍼 벡터, 헬퍼 게놈, 또는 방법.

**청구항 35**

제21항에 있어서, 상기 페이로드가 하기를 포함하는 것인, 재조합 Ad35 벡터 생산 시스템, 헬퍼 벡터, 헬퍼 게놈, 또는 방법:

통합 요소, 및

발현 산물을 암호화하고, 통합 요소로 구성되지 않고, 통합 요소의 표적 게놈으로의 통합에 의해 기능적이지 않게 만들도록 배치되는, 조건부로 발현되는 핵산 서열.

**청구항 36**

제35항에 있어서, 상기 조건부로 발현되는 핵산 서열에 의해 암호화되는 발현 산물이 CRISPR 시스템 구성요소 또는 염기 편집기 시스템 구성요소를 포함하며, 임의적으로 여기서 구성요소는 CRISPR-연관 RNA-가이드된 엔도뉴클레아제, 염기 편집기 효소, 또는 gRNA인, 재조합 Ad35 벡터 생산 시스템, 헬퍼 벡터, 헬퍼 게놈, 또는 방법.

**청구항 37**

제21항에 있어서, 상기 페이로드가 선택 카세트를 포함하며, 임의적으로 여기서 선택 카세트는 통합 요소로 구성되는 것인, 재조합 Ad35 벡터 생산 시스템, 헬퍼 벡터, 헬퍼 게놈, 또는 방법.

**청구항 38**

제37항에 있어서, 상기 선택 카세트가 $mgmt^{P140K}$를 암호화하는 핵산 서열을 포함하거나 또는 상기 선택 카세트가 항-CD33 shRNA를 암호화하는 핵산 서열을 포함하는 것인, 재조합 Ad35 벡터 생산 시스템, 헬퍼 벡터, 헬퍼 게놈, 또는 방법.

**청구항 39**

제1항 내지 제3항 및 제6항 내지 제10항 중 어느 한 항에 있어서, 상기 재조합효소 DR에 의해 플랭킹된 Ad35 패키징 서열의 적어도 일부가 GenBank 수탁 번호 AX049983에 따른 Ad35 서열의 뉴클레오티드 138-481에 상응하는 것인, 재조합 Ad35 벡터 생산 시스템, 헬퍼 벡터, 헬퍼 게놈, 또는 방법.

**청구항 40**

제1항 내지 제3항 및 제6항 내지 제10항 중 어느 한 항에 있어서, 상기 재조합효소 DR에 의해 플랭킹된 Ad35 패키징 서열의 적어도 일부가 GenBank 수탁 번호 AX049983에 따른 Ad35 서열의 뉴클레오티드 179-344; 뉴클레오티드 366-481; 뉴클레오티드 155-481; 뉴클레오티드 159-480; 뉴클레오티드 159-446; 뉴클레오티드 180-480; 뉴클레오티드 207-480; 뉴클레오티드 140-446; 뉴클레오티드 159-446; 뉴클레오티드 180-446; 뉴클레오티드 202-446; 뉴클레오티드 159-481; 뉴클레오티드 180-384; 뉴클레오티드 180-481; 또는 뉴클레오티드 207-481에 상응하는 것인, 재조합 Ad35 벡터 생산 시스템, 헬퍼 벡터, 헬퍼 게놈, 또는 방법.

**청구항 41**

제1항 내지 제3항 및 제6항 내지 제10항 중 어느 한 항에 있어서, 상기 재조합효소 DR이 LoxP 부위인, 재조합

Ad35 벡터 생산 시스템, 헬퍼 벡터, 헬퍼 게놈, 또는 방법.

**청구항 42**

제2항, 제3항, 제8항 및 제9항 중 어느 한 항에 있어서, 상기 Ad35 헬퍼 게놈이 293 T 세포에서 증폭을 위해 Ad5 E4orf6을 포함하는 것인, 헬퍼 벡터 또는 헬퍼 게놈.

**청구항 43**

제2항, 제3항, 제8항, 및 제9항 중 어느 한 항에 있어서, 상기 헬퍼 게놈이 서열번호: 51-65 중 임의의 하나에 제시된 바와 같은 서열을 포함하거나 또는 생성하는 것인, 헬퍼 벡터 또는 헬퍼 게놈.

**청구항 44**

제2항 내지 제5항, 제8항, 및 제9항 중 어느 한 항의 헬퍼 벡터, 헬퍼 게놈, 공여자 벡터, 또는 공여자 게놈을 포함하는 세포로, 임의적으로 여기서 세포는 HEK293 세포인, 세포.

**청구항 45**

제1항, 제4항, 제6항, 제7항, 제10항, 제13항 내지 제27항 및 제44항 중 어느 한 항의 공여자 게놈을 포함하는 세포로, 임의적으로 여기서 세포는 적혈구이고, 임의적으로 여기서 세포는 조혈 줄기 세포, T-세포, B-세포, 또는 골수성 세포이고, 임의적으로 여기서 세포는 발현 산물을 분비하는 것인, 세포.

**청구항 46**

제6항 또는 제10항에 있어서, 상기 세포가 HEK293 세포인, 방법.

**청구항 47**

세포를 변형시키는 방법으로, 상기 방법은 세포를 제5항 및 제11항 내지 제27항 중 어느 한 항에 따른 Ad35 공여자 벡터와 접촉시키는 단계를 포함하는 것인, 방법.

**청구항 48**

대상체의 세포를 변형시키는 방법으로, 상기 방법은 제5항 및 제11항 내지 제27항 중 어느 한 항에 따른 Ad35 공여자 벡터를 대상체에게 투여하는 단계를 포함하며, 임의적으로 여기서 방법은 대상체로부터 세포의 단리를 포함하지 않는 것인, 방법.

**청구항 49**

질환 또는 병태의 치료를 필요로 하는 대상체에서 질환 또는 병태를 치료하는 방법으로, 상기 방법은 제5항 및 제11항 내지 제27항 중 어느 한 항에 따른 Ad35 공여자 벡터를 대상체에게 투여하는 단계를 포함하며, 임의적으로 여기서 투여는 정맥내 투여인, 방법.

**청구항 50**

제49항에 있어서, 상기 방법이 동원제를 대상체에게 투여하는 단계를 포함하며, 임의적으로 여기서 동원제는 과립구-콜로니 자극 인자, GM-CSF, S-CSF, CXCR4 길항제, 및 CXCR2 작용제 중 하나 이상을 포함하며, 임의적으로 여기서 CXCR4 길항제는 AMD3100이고/이거나 여기서 CXCR2 작용제는 GRO-β인, 방법.

**청구항 51**

제49항 또는 제50항에 있어서, 상기 Ad35 공여자 벡터가 선택 카세트를 포함하며, 임의적으로 여기서 방법은 선택제를 대상체에게 투여하는 단계를 추가로 포함하며, 임의적으로 여기서 선택 카세트는 $mgmt^{P140K}$를 암호화하고 선택제는 $O^6$BG/BCNU인, 방법.

**청구항 52**

제49항에 있어서, 상기 방법이 면역 억제제를 대상체에게 투여하는 단계를 추가로 포함하며, 임의적으로 여기서 면역 억제 레지멘은 스테로이드, IL-6 수용체 길항제, 및/또는 IL-1R 수용체 길항제를 포함하며, 임의적으로 여

기서 스테로이드는 글루코코르티코이드 또는 덱사메타손을 포함하는 것인, 방법.

**청구항 53**

제49항에 있어서, 상기 Ad35 공여자 벡터가 통합 요소를 포함하고 방법이 CD46을 발현하는 세포의 적어도 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 또는 95%, 조혈 줄기 세포의 적어도 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 또는 95%, 및/또는 적혈구 $Ter119^{+}$ 세포의 적어도 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 또는 95%에서 이의 통합 요소 카피의 통합 및/또는 발현을 유발하는 것인, 방법.

**청구항 54**

제49항에 있어서, 상기 방법이 통합 요소의 적어도 1 개 카피를 포함하는 표적 세포 게놈에서 통합 요소의 평균 적어도 2 개 카피 또는 적어도 2.5 개 카피의 통합을 유발하는 것인, 방법.

**청구항 55**

제49항에 있어서, 상기 방법이 페이로드 또는 이의 통합 요소에 의해 암호화된 발현 산물의 발현을 참조 수준의 적어도 약 20% 또는 참조 수준의 적어도 약 25%인 수준에서 유발하며, 임의적으로 여기서 참조는 대상체 또는 참조 집단에서 내인성 참조 단백질의 발현인, 방법.

**청구항 56**

제49항에 있어서, 상기 질환 또는 병태가 혈색소병증, 혈소판 장애, 빈혈, 면역 결핍 응고 인자 결핍, 판코니 빈혈, 알파-1 항트립신 결핍, 겸상 세포 빈혈, 지중해빈혈, 중간형 지중해빈혈, A형 혈우병, B형 혈우병, 폰빌레브란트병, 인자 V 결핍, 인자 VII 결핍, 인자 X 결핍, 인자 XI 결핍, 인자 XII 결핍, 인자 XIII 결핍, 베르나르-술리에 증후군, 그레이 혈소판 증후군, 또는 뮤코다당류축적증인, 방법.

**청구항 57**

제49항에 있어서, 상기 대상체가 암을 앓고 있는 대상체이고 상기 방법이 암을 치료, 예방, 또는 지연시키거나, 또는 암 재발을 지연시키며,

임의적으로 여기서 대상체는 암 발병과 연관된 하나 이상의 생식 계열 돌연변이의 보균자이고,

임의적으로 여기서 암은 역형성 성상세포종, 유방암, 난소암, 결장직장암, 미만성 내재성 뇌간 신경교종, 유잉 육종, 다형성 교모세포종, 악성 뇌교종, 흑색종, 전이성 악성 흑색종, 비인두암, 또는 소아암이고,

임의적으로 여기서 대상체에게 $O^{6}$BG, TMZ(테모졸로미드), 및/또는 BCNU(카르무스틴)를 제공하거나 또는 투여하는 것인, 방법.

**청구항 58**

제49항에 있어서, 상기 질환 또는 병태가 중간형 지중해빈혈이며, 임의적으로 여기서 벡터 또는 게놈은 하기로부터 선택된 하나 이상의 발현 산물을 암호화하는 핵산을 포함하는 것이며, 임의적으로 여기서 방법은 중간형 지중해빈혈의 증상을 감소시키고/시키거나 중간형 지중해빈혈을 치료하고/하거나 HbF를 증가시키는 것인, 방법:

내인성 γ-글로빈의 발현을 증가시키거나 또는 재활성화하는 발현 산물(들)로, 임의적으로 여기서 내인성 γ-글로빈의 발현을 증가시키거나 또는 재활성화하는 발현 산물(들)은 CRISPR-연관 RNA-가이드된 엔도뉴클레아제 또는 염기 편집기 및 다음 중 하나 이상을 포함하는 것:

HBG1의 핵산 서열에 결합하고 표적 핵산 서열과 작동가능하게 연결된 코딩 서열로부터 발현을 증가시키도록 조작된 gRNA;

HBG2의 핵산 서열에 결합하고 표적 핵산 서열과 작동가능하게 연결된 코딩 서열로부터 발현을 증가시키도록 조작된 gRNA; 및

적혈구 인핸서 bcl11a의 핵산 서열에 결합하고 BCL11A 발현을 감소시키도록 조작된 gRNA;

γ-글로빈; 및

β-글로빈.

## 발명의 설명

### 기 술 분 야

[0001] **관련 출원에 대한 상호 참조**

[0002] 본 출원은 2019년 7월 2일 출원된 미국 가출원 번호 제62/869,907호, 2019년 11월 14일 출원된 미국 가출원 번호 제62/935,507호, 및 2020년 4월 13일 출원된 미국 가출원 번호 제63/009,385호의 이익을 주장하며, 각각의 개시내용은 그 전문이 본원에 참조로 포함된다.

[0003] **정부 지원**

[0004] 본 발명은 국립보건원(National Institutes of Health)에 의해 수여된 보조금 번호 HL130040, HL141781, 및 CA204036 하에 정부 지원으로 이루어졌다. 정부는 본 발명의 특정 권리를 갖는다.

[0005] **서열 목록에 관한 진술**

[0006] 본 출원과 관련된 서열 목록은 종이 사본 대신에 텍스트 형식으로 제공되며 명세서에 참조로 포함된다. 서열 목록을 함유하는 텍스트 파일의 명칭은 F053-0107PCT_ST25.txt이다. 텍스트 파일은 945 KB이며, 2020년 7월 2일 생성되었고, EFS-Web을 통해 전자적으로 제출되고 있다.

### 배 경 기 술

[0007] 많은 의학적 병태는 유전적 돌연변이에 의해 유발되고/되거나, 적어도 부분적으로, 유전자 요법에 의해 치료가능하다. 이러한 병태는 예를 들어, 혈색소병증, 면역 결핍, 및 암을 포함한다. 혈색소병증으로 알려진 유전 질환은 전세계적으로 가장 널리 퍼진 유전 질환 유형 중 하나이며, 저개발 국가에서 태어난 환자 중에서 생존율이 상당히 감소된다. 혈색소병증의 예는 겸상적혈구병 및 지중해빈혈을 포함한다. 면역 결핍은 1차성 또는 2차성일 수 있다. 80 개 초과의 1차성 면역 결핍 질환이 세계보건기구(World Health Organization)에 의해 인식되고 있다. 유전적 돌연변이에 의해 유발되고/되거나, 적어도 부분적으로, 유전자 요법에 의해 치료가능한 의학적 병태에 대한 예방적 및 치료적 치료가 필요하다.

### 발명의 내용

[0008] 유전자 요법은 혈색소병증, 면역 결핍, 및 암을 포함하나 이에 제한되지 않는 유전적 구성요소가 있는 많은 병태를 치료할 수 있다. 분자 생물학은 유전적 조작을 위한 다양한 도구를 포함하지만, 예를 들어, 생체외 및 생체내에서 유전자 요법 맥락에서 이러한 도구의 적용은 유전자 요법 벡터에서 사용하기 위한 유전적 작제물의 개발, 뿐만 아니라 벡터 자체의 개발과 적어도 부분적으로 연관된 새로운 기회 및 도전을 불러일으킨다.

[0009] 본 개시내용은 무엇보다도, 표적 세포에서 염기 편집기의 발현을 위한 아데노바이러스 벡터 및 아데노바이러스 게놈(예를 들어, "재조합" 또는 "조작된" 아데노바이러스 벡터 및 아데노바이러스 게놈)을 포함한다. 본 개시내용은 무엇보다도, 표적 세포에서 CRISPR-연관 RNA-가이드된 엔도뉴클레아제 및/또는 가이드 RNA(gRNA)인 CRISPR 효소를 포함하는 CRISPR 시스템의 발현을 위한 아데노바이러스 벡터 및 아데노바이러스 게놈을 포함하며, 임의적으로 여기서 CRISPR 시스템의 적어도 하나의 구성요소의 발현은 자기-불활성화된다. 본 개시내용은 무엇보다도, 표적 세포에서 염기 편집 효소 및/또는 가이드 RNA(gRNA)를 포함하는 염기 편집 시스템의 발현을 위한 아데노바이러스 벡터 및 아데노바이러스 게놈을 포함하며, 임의적으로 여기서 염기 편집 시스템의 적어도 하나의 구성요소의 발현은 자기-불활성화된다. 본 개시내용은 무엇보다도, 표적 세포에서 발현 산물(예를 들어, 치료 발현 산물)의 발현을 지시하는 조절 서열을 포함하는 아데노바이러스 벡터 및 아데노바이러스 게놈을 포함하며, 여기서 조절 서열은 miRNA 결합 부위를 포함하거나 또는 여기서 조절 서열은 β-글로빈 유전자좌 제어 영역(LCR), 예컨대 β-글로빈 긴 LCR를 포함한다. 본 개시내용은 무엇보다도, 표적 세포에서 복수의 치료 발현 산물, 예를 들어, 질환 또는 병태의 치료에 함께 기여하는 치료 발현 산물을 발현하는 조합 아데노바이러스 벡터 및 아데노바이러스 게놈을 포함한다. 본 개시내용은 무엇보다도, β-글로빈 긴 LCR을 포함하는 페이로드(payload)의 표적 세포 게놈 내로의 통합을 위한 아데노바이러스 벡터 및 아데노바이러스 게놈을 포함한다. 본 개시내용은 무엇보다도, 특정 기존 벡터에 대한(예를 들어, Ad5 벡터에 대한) 면역원성이 감소된 아데노바이러스 벡터, 및 이의 아데노바이러스 게놈을 포함한다. 본 개시내용은 무엇보다도, Ad35 아데노바이러스 벡터, Ad35 아데노바이러스 게놈, HDAd35 아데노바이러스 벡터, HDAd35 아데노바이러스 게놈, 지지 벡터, 지지

게놈, Ad35 헬퍼 벡터, 및 ad Ad35 헬퍼 게놈을 포함하며, 여기서 HDAd35 벡터는 특정 기존 벡터에 대한(예를 들어, Ad5 벡터 또는 Ad5/35 벡터에 대한) 면역원성이 감소될 수 있다.

[0010] 본 개시내용은 무엇보다도, 조혈 줄기 세포의 생체내 유전자 편집 및 관련된 유전자 요법 개선을 위해 CD46을 표적화하는 재조합 Ad35 벡터를 기재한다. 본원에 개시된 벡터 설계의 특정 구현예에서, 모든 단백질은 혈청형 35로부터 유래된다. 본원에 기재된 Ad35 벡터의 특정 구현예에서, 바이러스 유전자는 벡터에 남아있지 않다. 특정 구현예에서, ITR 및 패키징 서열은 Ad35로부터 유래된다. 특정 구현예에서, Ad35 전달 벡터는 치료 용도와 연관된 구성요소로 대체되고 제거된 모든 바이러스 단백질 암호화 유전자가 있다.

[0011] 특정 구현예에서, Ad35 벡터는 헬퍼-의존적이며, 본 개시내용은 또한 새로-설계된 Ad35 헬퍼 벡터를 제공한다. 특정 구현예는 Ad35를 만들기 위해 헬퍼-의존적 및 이식유전자 플라스미드의 최적화된 비율을 제공한다.

[0012] 본 개시내용 내에 기재된 관련된 유전자 요법 개선은 다음 중 하나 이상에 관한 것이다: (i) CD46 결합을 증가시키는 Ad35 놉(knob) 단백질의 신규 돌연변이; (ii) 생체내 변형된 세포의 양성 선택을 허용하는 벡터 특징; (iii) 임상적으로 관련된 시간 창 내에서 치료 단백질의 발현을 조절하는 microRNA 제어 시스템; (iv) 정의된 부위에서 표적화된 게놈 삽입을 용이하게 하기 위한 상동성 아암(arm)의 사용; (v) 게놈 억제인자 영역을 불활성화시켜, 내인성 유전자의 증가된 발현을 허용하기 위한 CRISPR의 사용; (vi) 표적화된 CD46-발현 세포에 Ad35 벡터의 전달을 증가시키기 위한 동원 전략의 사용; (vii) 유전자 발현을 증가시키기 위한 미니- 또는 긴-형태의 유전자좌 제어 영역의 사용; (viii) 트랜스포사제(transposase) 시스템과 함께 삽입될 수 있는 트랜스포존의 크기를 증가시키기 위한 재조합효소 시스템의 사용; (ix) 벡터 전달 전에 스테로이드 전달(예를 들어, 글루코코르티코이드, 덱사메타손); 및 (x) 치료 단백질을 생성 및 분비하는 적혈구. 이러한 관련 유전자 요법 개선은 각각 본원에 기재된 Ad35 벡터로 실행될 수 있고 또한 다른 바이러스 벡터 전달 시스템과 함께 활용될 수 있다. 일 예로서, CD46 결합을 증가시키는 돌연변이된 Ad35 놉 단백질은 렌티바이러스 또는 거품 전달 시스템과 함께 활용될 수 있다.

[0013] 본원에 기재된 발전은 또한 다음에 관한 것이다: (i) HDAd5/35++ 벡터를 사용한 SB100x-매개 이식유전자 추가를 위한 생체내 HSC 형질도입/선택 기술; (ii) 적혈구 bcl11a-인핸서(예를 들어, BCL11A 발현 감소) 및 HBG1/2 프로모터 영역(γ-글로빈의 발현 증가)을 동시에 표적화함으로써 HbF 재활성화 증가; (iii) 생체내 CRISPR 게놈 조작; (iv) 지중해빈혈의 교정; (v) γ 유전자 추가 및 재활성화의 조합(SB100x 시스템); (vi) CRISPR/Cas9의 자기-불활성화; (vii) 자기-방출 카세트가 있는 HDAd를 공여자 벡터로 사용한 표적화된 통합; (viii) 분비된 치료 단백질의 고수준 생산을 위한 공장으로서 적혈구 세포를 사용한 생체내 HSC 유전자 요법; (ix) 암을 치료하기 위한 치료 접근법(예방적 및 치료적); 및 (x) HDAd35++ 벡터.

[0014] 특정 구현예는 CD46에 대한 표적화된 결합을 증가시켜, 치료 유전자의 보다 표적화되고 특이적인 전달을 허용하는 돌연변이된 놉 단백질에 관한 것이다.

[0015] 특정 구현예는 표적화된 게놈 삽입을 용이하게 하기 위한 상동성 아암의 사용에 관한 것이며, 이는 전형적으로 이식유전자 수준의 더 높은 발현을 허용하는 개방 염색질인 게놈 안전 항구(safe harbor) 내로의 염색체 통합을 제공하는 데 사용될 수 있다. 본원에 기재된 바와 같이, 특정 구현예에서, 1.8 b 상동성 아암은 잘 작동하며, 0.8을 하한으로 한다. 단일 뉴클레오티드 다형성은 1.8 b 초과의 상동성 아암에서 통합에 영향을 미치기 시작할 수 있다.

[0016] 특정 구현예는 조건화에 대한 필요성을 완화하기 위한 동원 레지멘의 사용에 관한 것이다.

[0017] 특정 구현예는 (i) 저용량 $O^6$-벤질구아닌 및 비스-클로로에틸니트로소우레아를 사용한 단기간 치료에 의해 치료 효과를 증가시킬 수 있는 $MGMT^{P140K}$ 시스템, (ii) SB100X 트랜스포사제-기반 통합 기구, 및 (iii) 마이크로-LCR-구동된 γ-글로빈 유전자를 사용한, Ad35 생체내 유전자 요법을 제공한다.

[0018] 특정 구현예는 (i) 내인성 유전자의 억제를 역전시키기 위해 HBG1/2 프로모터 내의 BCL11A 결합 부위를 표적화하는 CRISPR/Cas9 카세트, (ii) 5kb β-글로빈 미니-LCR에 의해 구동된 $^{γ}$-글로빈 유전자 카세트, 및 FRT 및 트랜스포존 부위에 의해 플랭킹된 후자의 2 개 카세트와 함께 형질도입된 세포의 생체내 선택을 허용하는 EF1α-$MGMT^{P140K}$ 발현 카세트를 포함하는 Ad35 아데노바이러스 벡터(HDAd-comb)를 포함한다.

[0019] 특정 구현예는 적혈구에서 태아 γ-글로빈 발현의 재활성화를 목표로 하는 성인 CD34+ 세포에서 CRISPR/Cas9-매개 게놈 편집 접근법을 기재한다. CD34+ 세포의 적혈구 분화를 수반하는 모델은 $^{γ}$-글로빈 재활성화를 평가하는

데 있어서 한계가 있기 때문에, CRISPR/Cas9를 발현하는 인간 β-글로빈 유전자좌-유전자이식, 헬퍼-의존적 인간 CD46-표적화 아데노바이러스 벡터(HDAd-HBG-CRISPR)는 $^{\gamma}$-글로빈 프로모터 내의 억제인자 결합 영역을 파괴하는 데 사용되었다.

[0020] 특정 구현예는 통합 CD46 표적화된 Ad35 벡터 시스템을 제공한다: 이식유전자는 (i) γ 글로빈 유전자의 발현을 구동하는 β-글로빈 유전자좌 제어 영역(LCR), 및 (ii) 생체내 유전자-변형된 HSC의 양성 선택을 위한 $MGMT^{P140K}$ 카세트의 발현을 구동하는 EF1-α(구성적 프로모터)를 포함하였다.

[0021] 특정 구현예는 통합 CD46 표적화된 Ad35 벡터 시스템을 제공한다: 이식유전자는 (i) γ 글로빈 유전자(임의적으로 3' UTR 포함)의 발현을 구동하는 21.5 kb(긴) 인간 β-글로빈 유전자좌 제어 영역(LCR(HS1-HS5)) 및 β-글로빈 프로모터(1.6 kb), 및 (ii) 생체내 유전자-변형된 HSC의 양성 선택을 위한 $MGMT^{P140K}$ 카세트의 발현을 구동하는 EF1- α(구성적 프로모터)를 포함하였다. 일부 구현예는 3'HS1(인간 β-글로빈 3'HS1; 3 kb, 예를 들어, 여기서 3'HS1은 염색체 11의 위치 5206867-5203839의 서열을 가짐)을 추가로 포함할 수 있다. 다양한 구현예에서, 3'HS1은 서열번호: 287에 제시된 바와 같은 하기 핵산 서열, 또는 서열번호: 287에 대해 적어도 80% 서열 동일성을 갖는 서열, 예를 들어, 서열번호: 287에 대해 적어도 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 서열을 갖는다. 이러한 구현예는 재조합효소 시스템(예를 들어, Flp/Frt; Cre/Lox)과 조합하여 과활성 트랜스포사제(예를 들어, SB100X)를 활용할 수 있다. 따라서, 일 특정 구현예에서, Ad35 벡터 시스템은 예를 들어, 긴 인간 β-글로빈 유전자좌 제어 영역(21.5 kb), 인간 β-글로빈 프로모터(1.6 kb), 3' UTR과 함께 인간 γ 글로빈 유전자(2.7 kb), 인간 β-글로빈 3' UTR, 및 3'HS1(3 kb)을 포함하는 전위성 이식유전자 삽입물을 포함할 수 있다. 전위성 이식유전자 삽입물은 $MGMT^{P140K}$의 발현을 구동하는 EF1- α(구성적 프로모터)를 추가로 포함할 수 있다. 다양한 구현예에서, Ad35 벡터 시스템은 예를 들어, 32.4 kb의 전위성 이식유전자 삽입물을 포함할 수 있다.

[0022] 특정 구현예는 치료 이식유전자의 발현을 제어하기 위해 HSPC가 종양에 모집될 때만 활성화되는 miRNA 조절 시스템을 제공한다. 본 개시내용의 이러한 특징은 항 PDL1-γ1을 이식유전자로서 사용하여 입증된다. 이러한 시스템은 종양 미세환경의 맥락에서 치료 이식유전자의 발현을 조절하는 데 사용될 수 있다.

[0023] 다양한 구현예에서, microRNA 제어 시스템은 유전자의 발현이 microRNA 부위(예를 들어, microRNA가 상호작용할 수 있는 것이 있는 핵산 서열)의 존재에 의해 조절되는 방법 또는 조성물을 지칭할 수 있으며, 예는 실시예 5에 제공되었다. 특정 구현예에서, microRNA 제어 시스템은 유전자가 표적 세포, 예컨대 HSPC 예를 들어, 종양 침윤 HSPC에서 배타적으로 발현되도록 유전자의 발현을 조절하였다. 일부 구현예에서, 관심 단백질 또는 핵산(예를 들어, 항암제 예컨대 CAR, TCR, 항체, 및/또는 체크포인트 억제제, 예를 들어, 체크포인트 억제제인 αPD-L1 항체(예를 들어, αPD-L1γ1 항체))을 암호화하는 핵산(예를 들어, 치료 유전자)은 microRNA 부위, 복수의 동일한 microRNA 부위, 또는 복수의 별개의 microRNA 부위를 포함하거나, 이와 회합하거나, 또는 이와 작동가능하게 연결된다. 당업자는 관심 유전자를 암호화하는 서열을 갖는 핵산 또는 이의 부분과 microRNA 부위를 회합하는 수단 및 기술에 친숙할 것이지만, 특정 비제한적인 예가 본원에 제공된다. 예를 들어, 관심 유전자(예를 들어, αPD-L1γ1 항체를 암호화하는 서열)는 관심 유전자의 발현이 종양-침윤 백혈구 세포가 아니지만, 종양-침윤 백혈구에서 발현을 억제하지 않는 세포에서 발현을 억제하는 하나 이상의 microRNA 부위의 존재에 의해 조절되도록 핵산에 존재할 수 있다. 특정한 특정 예에서, 관심 유전자(예를 들어, αPD-L1γ1 항체를 암호화하는 서열)는 관심 유전자의 발현이 종양-침윤 백혈구 세포가 아니지만, 종양-침윤 백혈구에서 발현을 억제하지 않는 세포에서 발현을 억제하는 하나 이상의 miR423-5p microRNA 부위의 존재에 의해 조절되도록 핵산에 존재할 수 있다. 다양한 구현예에서, microRNA 제어 시스템은 하나 이상의 microRNA 부위, 예를 들어, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 개, 또는 그 이상의 microRNA 부위를 포함하거나, 또는 여기서 관심 단백질 또는 핵산의 발현이 이에 의해 조절되는 핵산을 포함할 수 있다. 다양한 구현예에서, microRNA 제어 시스템은 하나 이상의 miR423-5p microRNA 부위, 예를 들어, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 개, 또는 그 이상의 miR423-5p microRNA 부위를 포함하거나, 또는 여기서 관심 단백질 또는 핵산의 발현이 이에 의해 조절되는 핵산을 포함할 수 있다. 일부 특정 구현예에서, microRNA 제어 시스템은 αPD-L1γ1 항체를 암호화하고 하나 이상의 miR423-5p microRNA 부위, 예를 들어, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 개, 또는 그 이상의 miR423-5p microRNA 부위, 예를 들어, miR423-5p microRNA 부위를 포함하거나, 또는 여기서 αPD-L1γ1 항체의 발현이 이에 의해 조절되는 핵산을 포함할 수 있다.

[0024] 본 개시내용은 조혈 줄기 세포의 생체내 유전자 편집 및 관련 유전자 요법 개선을 위해 CD46을 표적화하는 재조

합 Ad35 벡터를 기재한다. 특정 구현예에서, Ad35 전달 벡터는 치료 용도와 연관된 구성요소로 대체되고 제거된 모든 바이러스 단백질 암호화 유전자가 있다. 바이러스 단백질을 암호화하는 모든 유전자의 제거는 30 kb의 벡터 운반 용량을 제공하며, 다른 바이러스 벡터 전달 플랫폼으로 이용가능한 것보다 상당히 더 많은 공간이다. 특정 구현예에서, Ad35 벡터는 헬퍼-의존적이고, 본 개시내용은 또한 새로-설계된 Ad35 헬퍼 벡터를 제공한다. 의심의 여지를 피하기 위해, 본원에 사용된 바와 같은 용어 "유전자 편집"은 핵산 서열을 변형시키기 위한 벡터 또는 제제의 임의의 사용을 포함하나 이에 제한되지 않는다.

[0025] 본원에는 microRNA 제어 시스템 및 본원에 개시된 microRNA(또한 본원에서 miRNA로 지칭됨) 부위(또한 본원에서 표적 부위로 지칭됨)를 포함하는 다른 핵산을 포함하나 이에 제한되지 않는 본원에 제공된 핵산이거나 또는 이를 포함하고/하거나, 항체 예컨대 αPD-L1 항체(예를 들어, αPD-L1ɣ1 항체)를 포함하나 이에 제한되지 않는 본원에 개시된 제제를 암호화하는 벡터가 추가로 제공된다. 본 개시내용의 다양한 구현예 중 임의의 것에서, 벡터는 Ad5/35 벡터일 수 있으며, 임의적으로 여기서 Ad5/35 벡터는 헬퍼-의존적 Ad5/35(HDAd5/35)이다. 본 개시내용의 다양한 구현예 중 임의의 것에서, 벡터는 본원에 제공된 변이(예를 들어, 아미노산 돌연변이)를 포함하는 Ad5/35 벡터(예를 들어, HDAd5/35 벡터)일 수 있으며, 이들 중 특정 벡터는 Ad5/35++(예를 들어, HDAd5/35++)로 지정될 수 있다. 의심의 여지를 피하기 위해, 당업자는 Ad5/35 이외(예를 들어, Ad5/35++ 이외 또는 HDAd5/35++ 이외)의 벡터가 명시된 구현예를 포함하여, 임의의 벡터를 사용한 임의의 구현예가 관련 텍스트에 언급된 바와 같은 이러한 벡터 이외에, Ad5/35 벡터(예를 들어, HDAd5/35, Ad5/35++, 및 HDAd5/35++ 벡터 중 임의의 것 포함)인 벡터를 개시하는 것으로 구체적으로 판독해야 한다는 것을 본 개시내용으로부터 인식하도록 의도된다.

[0026] 본 개시내용의 다양한 구현예 중 임의의 것에서, 벡터는 Ad35 벡터일 수 있으며, 임의적으로 여기서 Ad35 벡터는 HDAd35이다. 본 개시내용의 다양한 구현예 중 임의의 것에서, 벡터는 본원에 제공된 변이(예를 들어, 아미노산 돌연변이)를 포함하는 Ad35 벡터(예를 들어, HDAd35 벡터) 일 수 있으며, 이들 중 특정 벡터는 Ad35++(예를 들어, HDAd35++)로 지정될 수 있다. 의심의 여지를 피하기 위해, 당업자는 Ad35 벡터 이외(예를 들어, Ad35++ 이외 또는 HDAd35++ 이외)의 벡터가 명시된 구현예를 포함하여, 임의의 벡터를 사용한 임의의 구현예가 관련 텍스트에 언급된 바와 같은 이러한 벡터 이외에, Ad35 벡터(예를 들어, HDAd35, Ad35++, 및 HDAd35++ 벡터 중 임의의 것 포함)인 벡터를 개시하는 것으로 구체적으로 판독해야야 한다는 것을 본 개시내용으로부터 인식하도록 의도된다.

[0027] 나타낸 바와 같이, 본원에 기재된 벡터는 겸상적혈구병의 치료, ɣ 글로빈 유전자 추가 및 재활성화, 및 ɣ 글로빈 재활성화를 위한 다중 표적 부위의 표적화를 포함하는 많은 용도를 갖는다. 또한, 인자 VIII(FVIII) 이외에, 개시된 접근법의 적용은 예를 들어 다음을 포함하는 다른 분비된 단백질에 사용될 수 있다: (i) 다른 응고 인자, 구체적으로 FXI, FVII, 폰빌레브란트(von Willebrand) 인자(VWF), 및 희귀 응혈 인자(즉, 인자 I, II, V, X, XI, 또는 XIII); (ii) 폼페병(산 알파(α)-글루코시다제), 고세병(글루코레세브로시다제), 파브리병(α-갈락토시다제 A), 및 I형 점액다당류증(α-L-이두로니다제)과 같은 리소좀 축적 질환에 대한 효소 대체 요법(ERT)에 현재 사용되는(교차-교정 메커니즘의 이점 취득) 효소; (iii) 면역결핍(예를 들어 SCID-ADA(아데노신 데아미나제)); (iv) 심혈관 질환, 예를 들어 가족성 아포지질단백질 E 결핍 및 아테롬성동맥경화증(ApoE); (v) HIV, 만성 HCV, 또는 HBV 감염에 대한 바이러스 유인 수용체(예를 들어 HIV- 가용성 CD4, 또는 광범위한 중화 항체(bNAb))의 발현에 의한 바이러스 감염; (vi) 암(예를 들어 단클론 항체(예를 들어 트라스투주맙) 또는 체크포인트 억제제(예를 들어 aPDL1)의 제어된 발현 또는 치료 용량의 화학요법을 허용하기 위한 HSC의 보호 및 (vii) 판코니 빈혈에 대한 FANCA 유전자; (viii) A형 혈우병, B형 혈우병, 또는 폰빌레브란트병으로부터 임의적으로 선택된 응고 인자 결핍, (ix) 혈소판 장애, (x) 빈혈, (xi) 알파-1 항트립신 결핍, 또는 (xii) 면역결핍. 다른 추가적인 용도는 본원의 다른 곳에 보다 상세하게 기재되어 있다.

[0028] 따라서, 일 구현예는 조혈 줄기 세포의 생체내 유전자 편집을 위해 CD46을 표적화하는 재조합 혈청형 35 아데노바이러스(Ad35) 벡터를 제공한다.

[0029] 또 다른 구현예는 치료 단백질을 발현하도록 유전적으로 변형된 적혈구이다. 예로서, 치료 단백질은 일부 경우에 바이러스 감염을 차단하거나 또는 감소시키는 응고 인자 또는 단백질을 포함한다. 임의적으로, 적혈구는 치료 단백질을 분비한다.

[0030] 또한 본원에 기재된 재조합 Ad35 벡터 또는 적혈구의 용도가 제공된다. 이들 용도는 적혈구 bcl11a-인핸서 및 HBG 프로모터 영역을 동시에 표적화함으로써 HbF 재활성화 증가; ɣ-글로빈 유전자 추가 및 내인성 ɣ-글로빈 유전자 재활성화의 조합; 생체내 CRISPR 게놈 조작; 치료 유전자 제공; (i) 혈색소병증, (ii) 판코니 빈혈,

(iii) A형 혈우병, B형 혈우병, 또는 폰빌레브란트병으로부터 임의적으로 선택된 응고 인자 결핍, (iv) 혈소판 장애, (v) 빈혈, (vi) 알파-1 항트립신 결핍, 또는 (v) 면역 결핍 치료; 지중해빈혈 치료; 암 치료, 암 재발 예방 또는 지연 또는 고위험 생식 계열 돌연변이의 보균자에서 암 발병 예방 또는 지연, 임의적으로 여기서 암은 유방암 또는 난소암인 것; CRISPR/Cas9의 자기-불활성화; 공여자 벡터로서 자기-방출 카세트가 있는 HDAd를 사용한 표적화된 통합을 위한 것을 포함한다. 이들 용도 중 임의의 것은 임의적으로 동원을 포함할 수 있으며, 예를 들면 여기서 동원은 Gro-베타, GM-CSF, S-CSF, 및/또는 AMD3100의 투여를 포함한다.

[0031] 또한 또 다른 용도 구현예는 스테로이드(예를 들어, 글루코코르티코이드 또는 덱사메타손), IL-6 수용체 길항제, 및/또는 IL-1R 수용체 길항제를 Ad35 벡터 및/또는 적혈구를 받는 대상체에게 투여하는 것을 포함하는 본원에 기재된 재조합 Ad35 벡터 또는 적혈구 중 임의의 것의 용도이다.

[0032] 또한 $O^6$BG 및 TMZ(테모졸로미드) 또는 BCNU(카르무스틴)를 Ad35 벡터 및/또는 적혈구를 받는 대상체에게 투여하는 것을 포함하는, 본원에 기재된 재조합 Ad35 벡터 또는 적혈구 중 임의의 것을 이용하는 용도 구현예가 제공된다. 이러한 용도 구현예의 예에 의해, 대상체는 역형성 성상세포종, 유방암, 결장직장암, 미만성 내재성 뇌간 신경교종, 유잉 육종, 다형성 교모세포종(GBM), 악성 뇌교종, 흑색종, 전이성 악성 흑색종, 비인두암, 또는 소아암에 대한 치료로서 $O^6$BG 및 TMZ 또는 BCNU를 받고 있다.

[0033] 또한 또 다른 구현예는 다음을 포함하는 재조합 아데노바이러스 혈청형 35(Ad35) 벡터 생산 시스템이다: Ad35 섬유 샤프트를 암호화하는 핵산 서열; Ad35 섬유 놉을 암호화하는 핵산 서열; 및 Ad35 패키징 서열의 적어도 일부를 플랭킹하는 재조합효소 DR을 포함하는 재조합 Ad35 헬퍼 게놈, 및 5' Ad35 ITR; 3' Ad35 ITR; Ad35 패키징 서열; 및 적어도 하나의 이종 발현 산물을 암호화하는 핵산 서열을 포함하는 재조합 헬퍼 의존적 Ad35 공여자 게놈.

[0034] 또한 Ad35 섬유 샤프트; Ad35 섬유 놉; 및 Ad35 패키징 서열의 적어도 일부를 플랭킹하는 재조합효소 DR을 포함하는 Ad35 게놈을 포함하는 재조합 아데노바이러스 혈청형 35(Ad35) 헬퍼 벡터 구현예가 제공된다.

[0035] 또한 Ad35 섬유 샤프트를 암호화하는 핵산 서열; Ad35 섬유 놉을 암호화하는 핵산 서열; 및 Ad35 패키징 서열의 적어도 일부를 플랭킹하는 재조합효소 DR을 포함하는 재조합 Ad35 헬퍼 게놈 구현예가 제공된다.

[0036] 또한 5' Ad35 ITR; 3' Ad35 ITR; Ad35 패키징 서열; 및 적어도 하나의 이종 발현 산물을 암호화하는 핵산 서열로, 여기서 게놈은 Ad35 바이러스 구조적 단백질을 암호화하는 핵산 서열을 포함하지 않는 것을 포함하는, 핵산 서열; 및 Ad35 섬유 샤프트 및/또는 Ad35 섬유 놉을 포함하는 재조합 헬퍼 의존적 Ad35 공여자 벡터 구현예가 제공된다.

[0037] 또한 5' Ad35 ITR; 3' Ad35 ITR; Ad35 패키징 서열; 및 적어도 하나의 이종 발현 산물을 암호화하는 핵산 서열을 포함하는 재조합 헬퍼 의존적 Ad35 공여자 게놈 구현예가 제공되며, 여기서 Ad35 공여자 게놈은 야생형 Ad35 게놈에 의해 암호화된 발현 산물을 암호화하는 핵산 서열을 포함하지 않는다.

[0038] 또 다른 구현예는 재조합 헬퍼 의존적 Ad35 공여자 벡터를 생산하는 방법으로, 상기 방법은 세포의 배양물로부터 재조합 헬퍼 의존적 Ad35 공여자 벡터를 단리하는 단계를 포함하며, 여기서 세포는 Ad35 섬유 샤프트를 암호화하는 핵산 서열; Ad35 섬유 놉을 암호화하는 핵산 서열; 및 Ad35 패키징 서열의 적어도 일부를 플랭킹하는 재조합효소 DR을 포함하는 재조합 Ad35 헬퍼 게놈, 및 5' Ad35 ITR; 3' Ad35 ITR; Ad35 패키징 서열; 및 적어도 하나의 이종 발현 산물을 암호화하는 핵산 서열을 포함하는 재조합 헬퍼 의존적 Ad35 공여자 게놈을 포함한다.

[0039] 또한 Ad35 섬유 샤프트를 암호화하는 핵산 서열; Ad35 섬유 놉을 암호화하는 핵산 서열; 및 Ad35 패키징 신호를 기능적으로 방해하지만 5' Ad35 ITR은 방해하지 않는 Ad35 게놈의 5' 단부의 550 개 뉴클레오티드 이내의 재조합효소 DR을 포함하는 재조합 Ad35 헬퍼 게놈, 및 5' Ad35 ITR; 3' Ad35 ITR; Ad35 패키징 서열; 및 적어도 하나의 이종 발현 산물을 암호화하는 핵산 서열을 포함하는 재조합 Ad35 공여자 게놈을 포함하는 재조합 Ad35 생산 시스템 구현예가 제공된다.

[0040] 또 다른 구현예는 Ad35 섬유 샤프트; Ad35 섬유 놉; 및 Ad35 패키징 신호를 기능적으로 방해하지만 5' Ad35 ITR은 방해하지 않는 Ad35 게놈의 5' 단부의 550 개 뉴클레오티드 이내의 재조합효소 DR을 포함하나 Ad35 게놈을 포함하는 재조합 Ad35 헬퍼 벡터이다.

[0041] 또 다른 구현예는 Ad35 섬유 샤프트를 암호화하는 핵산 서열; Ad35 섬유 놉을 암호화하는 핵산 서열; 및 Ad35 패키징 신호를 기능적으로 방해하지만 5' Ad35 ITR은 방해하지 않는 Ad35 게놈의 5' 단부의 550 개 뉴클레오티

드 이내의 DR을 포함하는 재조합 Ad35 헬퍼 게놈이다.

[0042] 또 다른 구현예는 재조합 헬퍼 의존적 Ad35 공여자 벡터를 생산하는 방법으로, 상기 방법은 세포의 배양물로부터 재조합 헬퍼 의존적 Ad35 공여자 벡터를 단리하는 단계를 포함하며, 여기서 세포는 Ad35 섬유 샤프트를 암호화하는 핵산 서열; Ad35 섬유 놉을 암호화하는 핵산 서열; 및 Ad35 패키징 신호를 기능적으로 방해하지만 5' Ad35 ITR은 방해하지 않는 Ad35 게놈의 5' 단부의 550 개 뉴클레오티드 이내의 재조합효소 DR을 포함하는 재조합 Ad35 헬퍼 게놈, 및 5' Ad35 ITR; 3' Ad35 ITR; Ad35 패키징 서열; 및 적어도 하나의 이종 발현 산물을 암호화하는 핵산 서열을 포함하는 재조합 Ad35 공여자 게놈을 포함한다.

[0043] 또한 또 다른 구현예는 본원에 기재된 바와 같은 헬퍼 벡터, 헬퍼 게놈, 공여자 벡터, 또는 공여자 게놈을 포함하는 세포이며, 임의적으로 여기서 세포는 HEK293 세포이다.

[0044] 또 다른 구현예는 본원에 기재된 구현예 중 임의의 하나의 공여자 게놈을 포함하는 세포이며, 임의적으로 여기서 세포는 적혈구이고, 임의적으로 여기서 세포는 조혈 줄기 세포, T-세포, B-세포, 또는 골수성 세포이고, 임의적으로 여기서 세포는 발현 산물을 분비한다.

[0045] 또한 세포를 변형시키는 방법이 제공되며, 상기 방법은 제공된 Ad35 공여자 벡터 구현예 중 임의의 하나에 따른 Ad35 공여자 벡터를 세포와 접촉시키는 단계를 포함한다.

[0046] 또한 대상체의 세포를 변형시키는 방법이 제공되며, 상기 방법은 Ad35 공여자 벡터 구현예 중 임의의 하나에 따른 Ad35 공여자 벡터를 대상체에게 투여하는 단계를 포함하며, 임의적으로 여기서 상기 방법은 대상체로부터 세포의 단리를 포함하지 않는다.

[0047] 또한 또 다른 구현예는 질환 또는 병태의 치료를 필요로 하는 대상체에서 질환 또는 병태를 치료하는 방법이며, 상기 방법은 본원에 제공된 Ad35 공여자 벡터 구현예 중 임의의 하나에 따른 Ad35 공여자 벡터를 대상체에게 투여하는 단계를 포함하며, 임의적으로 여기서 투여는 정맥내이다.

[0048] **정의**

[0049] ***단수형:*** 본원에 사용된 바와 같이, 단수형은 관사의 문법적 대상의 하나 또는 하나 초과(즉, 적어도 하나)를 지칭한다. 예로서, "요소"는 정확히 하나의 요소의 구현예 및 하나 초과의 요소를 포함하는 구현예를 개시한다.

[0050] ***약:*** 본원에 사용된 바와 같이, 용어 "약"은, 값과 관련하여 사용될 때, 참조된 값의 맥락에서 유사한 값을 지칭한다. 일반적으로, 문맥과 친숙한 당업자는 해당 문맥에서 "약"에 의해 포함된 적절한 변동 정도를 이해할 것이다. 예를 들어, 일부 구현예에서, 용어 "약"은 참조된 값의 25%, 20%, 19%, 18%, 17%, 16%, 15%, 14%, 13%, 12%, 11%, 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, 1% 이하 이내에서 값의 범위를 포함할 수 있다.

[0051] ***투여:*** 본원에 사용된 바와 같이, 용어 "투여"는 전형적으로 조성물이거나, 또는 조성물에 포함된 제제의 전달을 달성하기 위해 대상체 또는 시스템에 조성물을 투여하는 것을 지칭한다.

[0052] ***입양 세포 요법:*** 본원에 사용된 바와 같이, "입양 세포 요법" 또는 "ACT"는 치료 활성이 있는 세포를 대상체, 예를 들어, 병태, 장애, 또는 질환에 대한 치료를 필요로 하는 대상체 내로 전달하는 것을 수반한다. 일부 구현예에서, ACT는 세포의 생체외 및/또는 시험관내 조작 및/또는 확장 후에 세포를 대상체 내로 전달하는 것을 포함한다.

[0053] ***친화도:*** 본원에 사용된 바와 같이, "친화도"는 특정 결합제(예를 들어, 바이러스 벡터), 및/또는 이의 결합 모이어티와, 결합 표적(예를 들어, 세포) 사이의 비공유 상호작용의 총 합계의 강도를 지칭한다. 달리 나타내지 않는 한, 본원에 사용된 바와 같이, "결합 친화도"는 결합제 및 이의 결합 표적(예를 들어, 바이러스 벡터와 바이러스 벡터의 표적 세포) 사이의 1:1 상호작용을 지칭한다. 당업자는 친화도의 변화가 참조와의 비교에 의해 기재될 수 있거나(예를 들어, 참조에 대한 증가 또는 감소), 또는 수치상으로 기재될 수 있음을 이해한다. 친화도는 평형 해리 상수($K_D$) 및/또는 평형 결합 상수($K_A$)를 포함하나 이에 제한되지 않는 당업계에 알려진 다양한 방식으로 측정 및/또는 표현될 수 있다. $K_D$는 $k_{off}/k_{on}$의 몫인 반면, $K_A$는 $k_{on}/k_{off}$의 몫이며, 여기서 $k_{on}$은 예를 들어, 바이러스 벡터와 표적 세포의 결합 속도 상수를 지칭하고, $k_{off}$는 예를 들어, 표적 세포로부터 바이러스 벡터의 해리를 지칭한다. $k_{on}$ 및 $k_{off}$는 당업자에게 알려진 기술에 의해 결정될 수 있다.

[0054] ***제제:*** 본원에 사용된 바와 같이, 용어 "제제"는 원자, 분자, 화합물, 아미노산, 폴리펩티드, 뉴클레오티드, 핵산, 단백질, 단백질 복합체, 액체, 용액, 당류, 다당류, 지질, 또는 이의 조합 또는 복합체 중 하나 이상의 임

의의 것을 포함하나 이에 제한되지 않는 임의의 화학적 독립체를 지칭할 수 있다.

[0055] ***동종이계:*** 본원에 사용된 바와 같이, 용어 "동종이계"는 하나의 대상체로부터 유래된 다음 또 다른 대상체에게 도입된 임의의 물질, 예를 들어, 동종이계 T 세포 이식을 지칭한다.

[0056] ***사이에 또는 부터:*** 본원에 사용된 바와 같이, 용어 "사이에"는 경계를 포함하여, 표시된 상부 및 하부, 또는 제1 및 제2 경계 사이에 속하는 내용물을 지칭한다. 유사하게, 용어 "부터"는, 값의 범위의 맥락에서 사용될 때, 범위가 경계를 포함하여, 표시된 상부 및 하부, 또는 제1 및 제2 경계 사이에 속하는 내용물을 포함함을 나타낸다.

[0057] ***결합:*** 본원에 사용된 바와 같이, 용어 "결합"은 2 개 이상의 제제 사이 또는 중에서 비공유 회합을 지칭한다. "직접" 결합은 제제 사이의 물리적 접촉을 수반하고; 간접 결합은 하나 이상의 증간 제제와의 물리적 접촉 방식의 물리적 상호작용을 수반한다. 2 개 이상의 제제 사이의 결합은 상호작용하는 제제가 단리에서 또는 더 복잡한 시스템의 맥락에서(예를 들어, 담체 제제와 및/또는 생물학적 시스템 또는 세포에서 공유적으로 또는 달리 회합되는 동안) 연구되는 경우를 포함하여, 다양한 맥락 중 임의의 것에서 발생하고/하거나 평가할 수 있다.

[0058] ***암:*** 본원에 사용된 바와 같이, 용어 "암"은 세포가 상대적으로 비정상적, 비제어적, 및/또는 자율적 성장을 나타내어, 암 증식에 대한 제어의 상당한 손실을 특징으로 하는 비정상적으로 상승된 증식률 및/또는 비정상적인 성장 표현형을 나타내도록 하는, 병태, 장애, 또는 질환을 지칭한다. 일부 구현예에서, 암은 하나 이상의 종양을 포함할 수 있다. 일부 구현예에서, 암은 전암성(예를 들어, 양성), 악성, 전-전이성, 전이성, 및/또는 비-전이성인 세포이거나 또는 이를 포함할 수 있다. 일부 구현예에서, 암은 고형 종양이거나 또는 이를 포함할 수 있다. 일부 구현예에서, 암은 혈액 종양이거나 또는 이를 포함할 수 있다.

[0059] ***키메라 항원 수용체:*** 본원에 사용된 바와 같이, "키메라 항원 수용체" 또는 "CAR"은 (i) 표적 항원에 결합하는 모이어티를 포함하는 세포외 도메인; (ii) 막관통 도메인; 및 (iii) CAR이 세포외 결합 모이어티와 표적 항원의 결합에 의해 자극될 때 활성화 신호를 보내는 세포내 신호전달 도메인을 포함하는 조작된 단백질을 지칭한다. 키메라 항원 수용체를 발현하도록 유전적으로 조작된 T 세포는 CAR T 세포로 지칭될 수 있다. 따라서, 예를 들어, 특정 CAR이 T 세포에 의해 발현될 때, CAR 세포외 결합 모이어티와 표적 항원의 결합은 T 세포를 활성화할 수 있다. CAR은 또한 키메라 T 세포 수용체 또는 키메라 면역수용체로 알려져 있다.

[0060] ***조합 요법:*** 본원에 사용된 바와 같이, 용어 "조합 요법"은 2 개 이상의 제제 또는 레지멘이 함께 대상체의 병태, 장애, 또는 질환을 치료하도록 2 개 이상의 제제 또는 레지멘을 대상체에게 투여하는 것을 지칭한다. 일부 구현예에서, 2 개 이상의 치료제 또는 레지멘은 동시에, 순차적으로, 또는 중복 투약 레지멘으로 투여될 수 있다. 당업자 조합 요법이 2 개의 제제 또는 레지멘이 단일 조성물로 함께 투여되거나, 또는 동시에 투여되는 것을 포함하지만 요구하지 않음을 이해할 것이다.

[0061] ***발현 또는 활성 제어:*** 본원에 사용된 바와 같이, 제1 요소(예를 들어, 단백질, 예컨대 전사 인자, 또는 핵산 서열, 예컨대 프로모터)는 제2 요소의 발현 또는 활성이 적어도 하나의 조건 세트 하에 제1 요소의 상태(예를 들어, 존재, 부재, 형태, 화학적 변형, 상호작용, 또는 다른 활성)에 따라 전체적으로 또는 부분적으로 의존하는 경우 제2 요소(예를 들어, 단백질과 같은 제제를 암호화하는 핵산 또는 단백질)의 발현 또는 활성을 "제어"하거나 또는 "구동"한다. 발현 또는 활성의 제어는 예를 들어, 제1 요소의 상태 변화가, 적어도 하나의 조건 세트 하에, 참조 대조군과 비교하여 제2 요소의 발현 또는 활성의 적어도 10%(예를 들어, 적어도 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, 2-배, 3-배, 4-배, 5-배, 10-배, 20-배, 30-배, 40-배, 50-배, 100-배)의 변화를 초래한다는 점에서, 실질적인 제어 또는 활성일 수 있다.

[0062] ***에 상응하는:*** 본원에 사용된 바와 같이, 용어 "에 상응하는"은 적절한 참조 화합물 또는 조성물과의 비교를 통해 화합물 또는 조성물에서 구조적 요소의 위치/동일성을 지정하기 위해 사용될 수 있다. 예를 들어, 일부 구현예에서, 중합체의 단량체성 잔기(예를 들어, 폴리펩티드의 아미노산 잔기 또는 폴리뉴클레오티드의 핵산 잔기)는 적절한 참조 중합체의 잔기"에 상응하는" 것으로 식별될 수 있다. 예를 들어, 당업자는 제공된 폴리펩티드 또는 폴리뉴클레오티드 서열의 잔기가 종종 관련된 참조 서열의 체계에 따라 지정(예를 들어, 넘버링 또는 라벨링)됨을 이해한다(예를 들어, 이러한 지정이 제공된 서열의 문자 넘버링을 반영하지 않는 경우에도). 예시로서, 참조 서열이 위치 100-110에서 특정 아미노산 모티프를 포함하고, 두번째 관련 서열이 위치 110-120에서 동일한 모티프를 포함하는 경우, 두번째 관련 서열의 모티프 위치는 참조 서열의 위치 100-110"에 상응한다"고 할 수 있다. 당업자는 상응하는 위치가 예를 들어, 서열의 정렬에 의해 용이하게 식별될 수 있고, 이러한 정렬이 예를 들어, BLAST, CS-BLAST, CUDASW++, DIAMOND, FASTA, GGSEARCH/GLSEARCH, Genoogle, HMMER,

HHpred/HHsearch, IDF, Infernal, KLAST, USEARCH, parasail, PSI-BLAST, PSI-Search, ScalaBLAST, Sequilab, SAM, SSEARCH, SWAPHI, SWAPHI-LS, SWIMM, 또는 SWIPE와 같은 소프트웨어 프로그램을 포함하나 이에 제한되지 않는, 알려진 다양한 툴, 전략, 및/또는 알고리즘 중 임의의 것에 의해 통상적으로 달성됨을 이해한다.

[0063] ***투약 레지멘:*** 본원에 사용된 바와 같이, 용어 "투약 레지멘"은 대상체에게 투여되는 하나 이상의 동일하거나 또는 상이한 단위 용량의 세트를 지칭할 수 있으며, 전형적으로 일정 기간 동안 다른 대상체의 투여로부터 분리되는 것 중 각각의 복수의 단위 용량 투여를 포함한다. 다양한 구현예에서, 투약 레지멘의 하나 이상 또는 모든 단위 용량은 동일할 수 있거나 또는 달라질 수 있다(예를 들어, 시간 경과에 따라 증가하거나, 시간 경과에 따라 감소하거나, 또는 대상체 및/또는 의사의 결정에 따라 조정될 수 있음). 다양한 구현예에서, 각각의 용량 사이의 시간 기간 중 하나 이상 또는 모두는 동일할 수 있거나 또는 달라질 수 있다(예를 들어, 시간 경과에 따라 증가하거나, 시간 경과에 따라 감소하거나, 또는 대상체 및/또는 의사의 결정에 따라 조정될 수 있음). 일부 구현예에서, 주어진 치료제는 하나 이상의 용량을 수반할 수 있는 권고된 투약 레지멘을 갖는다. 전형적으로, 시판되는 약물의 적어도 하나의 권고된 투약 레지멘은 당업자에게 알려져 있다. 일부 구현예에서, 투약 레지멘은 관련 집단에 걸쳐 투여될 때 원하거나 또는 유리한 결과와 상관관계가 있다(즉, 치료 투약 레지멘임).

[0064] ***하류*** 및 ***상류:*** 본원에 사용된 바와 같이, 용어 "하류"는 제1 DNA 영역이, 제2 DNA 영역에 비해, 제1 DNA 영역 및 제2 DNA 영역을 포함하는 핵산의 C-말단에 더 가깝다는 것을 의미한다. 본원에 사용된 바와 같이, 용어 "상류"는 제1 DNA 영역이, 제2 DNA 영역에 비해, 제1 DNA 영역 및 제2 DNA 영역을 포함하는 핵산의 N-말단에 더 가깝다는 것을 의미한다.

[0065] ***유효량:*** "유효량"은 대상체에서 원하는 생리학적 변화를 초래하는 데 필요한 제형의 양이다. 유효량은 종종 연구 목적을 위해 투여된다.

[0066] ***조작된:*** 본원에 사용된 바와 같이, 용어 "조작된"은 인간의 손으로 조작된 측면을 지칭한다. 예를 들어, 폴리뉴클레오티드는 자연에서 그 순서로 함께 연결되지 않은 2 개 이상의 서열이 조작된 폴리뉴클레오티드에서 서로 직접적으로 연결되도록 사람의 손으로 조작될 때 "조작된" 것으로 간주된다. 당업자는 "조작된" 핵산 또는 아미노산 서열이 재조합 핵산 또는 아미노산 서열일 수 있고, "유전적으로 조작된" 것으로 지칭될 수 있음을 이해할 것이다. 일부 구현예에서, 조작된 폴리뉴클레오티드는 첫번째 서열과 작동가능하게 연결된 것으로 자연에서 발견되지만 두번째 서열과 작동가능하게 연결된 것으로 자연에서 발견되지 않는 코딩 서열 및/또는 조절 서열을 포함하며, 이는 인간의 손으로 두번째 서열에 작동가능하게 연결된 조작된 폴리펩티드에 있다. 일부 구현예에서, 세포 또는 유기체는 유전적 정보가 변경되도록 조작된 경우(예를 들어, 이전에 존재하지 않는 새로운 유전적 물질이 예를 들어 형질전환, 교배, 체세포 혼성화, 형질감염, 형질도입, 또는 다른 메커니즘에 의해 도입되었거나, 또는 이전에 존재하는 유전적 물질이 예를 들어 치환, 결실, 또는 교배에 의해 변경 또는 제거됨) "조작된" 또는 "유전적으로 조작된" 것으로 간주된다. 통상적인 관행이고 당업자에 의해 이해되는 바와 같이, 조작된 폴리뉴클레오티드 또는 세포의 완전하거나 또는 불완전한 자손 또는 카피는 직접 조작이 이전 독립체의 것이었을지라도 전형적으로 여전히 "조작된" 것으로 지칭된다.

[0067] ***부형제:*** 본원에 사용된 바와 같이, "부형제"는 예를 들어 원하는 일관성 또는 안정화 효과를 제공하거나 또는 이에 기여하기 위해 약제학적 조성물에 포함될 수 있는 비-치료제를 지칭한다. 일부 구현예에서, 적합한 약제학적 부형제는 예를 들어, 전분, 글루코스, 락토스, 수크로스, 젤라틴, 맥아, 벼, 밀가루, 백악, 실리카 겔, 나트륨 스테아레이트, 글리세롤 모노스테아레이트, 활석, 염화나트륨, 탈지분유, 글리세롤, 프로필렌, 글리콜, 물, 에탄올 등을 포함할 수 있다.

[0068] ***발현:*** 본원에 사용된 바와 같이, "발현"은 단백질과 같은 암호화된 제제의 핵산 서열로부터 생산을 초래하는 하나 이상의 생물학적 과정을 개별적으로 및/또는 점증적으로 지칭한다. 발현은 구체적으로 전사 및 번역 중 하나 또는 둘 다를 포함한다.

[0069] ***플랭크***: 본원에 사용된 바와 같이, 제2 요소 및 제3 요소와 인접한 서열에 존재하는 제1 요소(예를 들어, 핵산 서열 또는 아미노산 서열)는 제2 요소 및 제3 요소 사이의 인접한 서열에 위치하는 경우 제2 요소 및 제3 요소에 의해 "플랭킹된다". 따라서, 이러한 배열에서, 제2 요소 및 제3 요소는 제1 요소를 "플랭킹"하는 것으로 지칭될 수 있다. 플랭킹 요소는 플랭킹된 요소에 바로 인접하거나 또는 하나 이상의 관련 단위에 의해 플랭킹된 요소로부터 분리될 수 있다. 인접한 서열이 핵산 또는 아미노산 서열이고, 관련 단위가 각각 염기 또는 아미노산 잔기인 다양한 예에서, 플랭킹된 요소, 및 독립적으로, 제1 및/또는 제2 플랭킹 요소 사이에 있는 인접한 서열의 단위 수는 예를 들어, 50 개 이하 단위, 예를 들어, 50, 45, 40, 35, 30, 25, 20, 15, 10, 5, 4, 3, 2, 1, 또는 0 개 단위 이하일 수 있다.

[0070] ***단편:*** 본원에 사용된 바와 같이, "단편"은 참조 제제(때때로 "모체" 제제로 지칭됨)의 별개의 부분을 포함하고/하거나 이로 이루어진 구조를 지칭한다. 일부 구현예에서, 단편은 참조 제제에서 발견된 하나 이상의 모이어티가 결여되어 있다. 일부 구현예에서, 단편은 참조 제제에서 발견된 하나 이상의 모이어티를 포함하거나 또는 이로 이루어진다. 일부 구현예에서, 참조 제제는 폴리뉴클레오티드 또는 폴리펩티드와 같은 중합체이다. 일부 구현예에서, 중합체의 단편은 참조 중합체의 적어도 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500 개 이상의 단량체성 단위(예를 들어, 잔기)를 포함하거나 또는 이로 이루어진다. 일부 구현예에서, 중합체의 단편은 참조 중합체에서 발견된 단량체성 단위(예를 들어, 잔기)의 적어도 5%, 10%, 15%, 20%, 25%, 30%, 25%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 이상을 포함하거나 또는 이로 이루어진다. 참조 중합체의 단편은 참조 중합체의 상응하는 부분과 반드시 동일하지는 않다. 예를 들어, 참조 중합체의 단편은 참조 중합체에 대해 적어도 5%, 10%, 15%, 20%, 25%, 30%, 25%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 이상의 동일성을 갖는 잔기의 서열을 갖는 중합체일 수 있다. 단편은 참조 제제의 물리적 단편화에 의해 생성될 수 있거나, 또는 생성될 수 없다. 일부 경우에, 단편은 참조 제제의 물리적 단편화에 의해 생성된다. 일부 경우에, 단편은 참조 제제의 물리적 단편화에 의해 생성되지 않고 대신에 예를 들어, 새로운 합성 또는 다른 수단에 의해 생성될 수 있다.

[0071] ***유전자, 이식유전자:*** 본원에 사용된 바와 같이, 용어 "유전자"는 임의적으로 코딩 서열의 발현을 제어하는 조절 서열의 일부 또는 전부와 함께, 코딩 서열(즉, RNA 생성물 및/또는 폴리펩티드 생성물과 같은 발현 산물을 암호화하는 DNA 서열)이거나 또는 이를 포함하는 DNA 서열을 지칭한다. 일부 구현예에서, 유전자는 인트론과 같으나 이에 제한되지 않는 비-코딩 서열을 포함한다. 일부 구현예에서, 유전자는 코딩(예를 들어, 엑손) 및 비-코딩(예를 들어, 인트론) 서열을 둘 다 포함할 수 있다. 일부 구현예에서, 유전자는 프로모터인 조절 서열을 포함한다. 일부 구현예에서, 유전자는 (i) 소스 게놈과 같이, 참조 맥락에서 코딩 서열의 상류에 미리결정된 수의 뉴클레오티드를 확장하는 DNA 뉴클레오티드, 및 (ii) 소스 게놈과 같이, 참조 맥락에서 코딩 서열의 하류에 미리결정된 수의 뉴클레오티드를 확장하는 DNA 뉴클레오티드 중 하나 또는 둘 다를 포함한다. 다양한 구현예에서, 미리결정된 수의 뉴클레오티드는 500 bp, 1 kb, 2 kb, 3 kb, 4 kb, 5 kb, 10 kb, 20 kb, 30 kb, 40 kb, 50 kb, 75 kb, 또는 100 kb일 수 있다. 본원에 사용된 바와 같이, "이식유전자"는 유전자가 존재하거나 또는 유전자가 조작에 의해 배치될 수 있는 참조 맥락에 대해 내인성 또는 고유하지 않은 유전자를 지칭한다.

[0072] ***유전자 산물 또는 발현 산물:*** 본원에 사용된 바와 같이, 용어 "유전자 산물" 또는 "발현 산물"은 일반적으로 유전자로부터 전사된 RNA(처리전 및/또는 처리후) 또는 유전자로부터 전사된 RNA에 의해 암호화된 폴리펩티드(변형전 및/또는 변형후)를 지칭한다.

[0073] ***숙주 세포, 표적 세포:*** 본원에 사용된 바와 같이, "숙주 세포"는 이식유전자와 같은 외인성 DNA(재조합 또는 달리)가 도입된 세포를 지칭한다. 당업자는 "숙주 세포"가 외인성 DNA가 초기에 도입된 세포 및/또는 이의 완전하거나 또는 불완전한 자손 또는 카피일 수 있음을 이해한다. 일부 구현예에서, 숙주 세포는 하나 이상의 바이러스 유전자 또는 이식유전자를 포함한다. 일부 구현예에서, 의도되거나 또는 잠재적인 숙주 세포는 표적 세포로 지칭될 수 있다.

[0074] 다양한 구현예에서, 숙주 세포 또는 표적 세포는 다양한 표면 마커의 존재, 부재, 또는 발현 수준에 의해 식별된다.

[0075] 세포 또는 세포 집단이 특정 마커에 대해 "양성"이거나 또는 이를 발현한다는 진술은 특정 마커의 세포 상에서 또는 내에서 검출가능한 존재를 지칭한다. 표면 마커를 지칭할 때, 용어는 유세포 분석에 의해, 예를 들어, 마커에 특이적으로 결합하는 항체로 염색하고 상기 항체를 검출함으로써 검출되는 바와 같은 표면 발현의 존재를 지칭할 수 있으며, 여기서 염색은 달리 동일한 조건 하에 이소형-일치 대조군으로 동일한 절차를 수행하여 검출된 염색보다 실질적으로 높은 수준 및/또는 마커에 대해 양성인 것으로 알려진 세포에 대한 수준과 실질적으로 유사한 수준, 및/또는 마커에 대해 음성인 것으로 알려진 세포에 대한 수준보다 실질적으로 더 높은 수준에서 유세포 분석에 의해 검출가능하다.

[0076] 세포 또는 세포 집단이 특정 마커에 대해 "음성"이거나 또는 마커의 발현이 결여되어 있다는 진술은 특정 마커의 세포 상에서 또는 내에서 실질적으로 검출가능한 존재의 부재를 지칭한다. 표면 마커를 지칭할 때, 용어는 유세포 분석에 의해, 예를 들어, 마커에 특이적으로 결합하는 항체로 염색하고 상기 항체를 검출함으로써 검출되는 바와 같은 표면 발현의 부재를 지칭할 수 있으며, 여기서 염색은 달리 동일한 조건 하에 이소형-일치된 조

건과 동일한 절차를 수행하여 검출된 염색보다 실질적으로 높은 수준, 및/또는 마커에 대해 양성인 것으로 알려진 세포에 대한 수준보다 실질적으로 낮은 수준, 및/또는 마커에 대해 음성인 것으로 알려진 세포에 대한 수준과 비교하여 실질적으로 유사한 수준에서 유세포 분석에 의해 검출되지 않는다.

[0077] ***동일성***: 본원에 사용된 바와 같이, 용어 "동일성"은 중합체성 분자 사이, 예를 들어, 핵산 분자(예를 들어, DNA 분자 및/또는 RNA 분자) 사이 및/또는 폴리펩티드 분자 사이의 전반적인 관련성을 지칭한다. 2 개의 제공된 서열 사이의 퍼센트 동일성을 계산하는 방법은 당업계에 알려져 있다. 용어 "% 서열 동일성"은 서열을 비교함으로써 결정된 바와 같이, 2 개 이상의 서열 사이의 관계를 지칭한다. 당업계에서, "동일성"은 또한 이러한 서열의 스트링 사이의 일치에 의해 결정된 바와 같이 단백질 및 핵산 서열 사이의 서열 관련성 정도를 의미한다. "동일성"(종종 "유사성"으로 지칭됨)은 다음에 기재된 것들을 포함하여 알려진 방법에 의해 용이하게 계산될 수 있다: Computational Molecular Biology (Lesk, A. M., ed.) Oxford University Press, NY (1988); Biocomputing: Informatics and Genome Projects (Smith, D. W., ed.) Academic Press, NY (1994); Computer Analysis of Sequence Data, Part I (Griffin, A. M., and Griffin, H. G., eds.) Humana Press, NJ (1994); Sequence Analysis in Molecular Biology (Von Heijne, G., ed.) Academic Press (1987); 및 Sequence Analysis Primer (Gribskov, M. and Devereux, J., eds.) Oxford University Press, NY (1992). 동일성을 결정하기 위한 바람직한 방법은 테스트된 서열 사이의 최고 일치를 제공하도록 설계된다. 동일성 및 유사성을 결정하는 방법은 공개적으로 이용가능한 컴퓨터 프로그램에 성문화되어 있다. 예를 들면, 2 개의 핵산 또는 폴리펩티드 서열에 대한 퍼센트 동일성의 계산은, 예를 들어, 최적의 비교 목적을 위해 2 개의 서열(또는 하나 또는 두 서열의 보체)을 정렬함으로써 수행될 수 있다(예를 들어, 최적의 정렬을 위해 제1 및 제2 서열 중 하나 또는 둘 다에 갭이 도입될 수 있고 비교 목적을 위해 동일하지 않은 서열은 무시될 수 있다). 그런 다음 상응하는 위치에서 뉴클레오티드 또는 아미노산이 비교된다. 제1 서열에서 위치가 제2 서열에서 상응하는 위치와 동일한 잔기(예를 들어, 뉴클레오티드 또는 아미노산)에 의해 점유되면, 분자는 해당 위치에서 동일하다. 2 개의 서열 사이의 퍼센트 동일성은 서열에 의해 공유된 동일한 위치의 수의 함수이며, 임의적으로 갭의 수, 및 2 개의 서열의 최적의 정렬을 위해 도입될 필요가 있을 수 있는 각 갭의 길이를 설명한다. 2 개의 서열 사이의 서열 비교 및 퍼센트 동일성 결정은 컴퓨터 알고리즘, 예컨대 BLAST(기본 로컬 정렬 검색 도구)를 사용하여 달성될 수 있다. 서열 정렬 및 퍼센트 동일성 계산은 LASERGENE 생물정보학 컴퓨터 묶음의 Megalign 프로그램(DNASTAR, Inc., 위스콘신주 매디슨 소재)을 사용하여 수행될 수 있다. 서열의 다중 정렬은 또한 디폴트 매개변수(갭 패널티=10, 갭 길이 패널티=10)에 따른 Clustal 정렬 방법(Higgins and Sharp CABIOS, 5, 151-153 (1989)을 사용하여 수행될 수 있다. 관련 프로그램은 또한 GCG 프로그램 묶음(Wisconsin Package Version 9.0, Genetics Computer Group (GCG), 위스콘신주 매디슨 소재); BLASTP, BLASTN, BLASTX(Altschul 등, J. Mol. Biol. 215:403-410 (1990); DNASTAR(DNASTAR, Inc., 위스콘신주 매디슨 소재); 및 Smith-Waterman 알고리즘을 포함하는 FASTA 프로그램(Pearson, *Comput. Methods Genome Res.*, [Proc. Int. Symp.] (1994), Meeting Date 1992, 111-20. Editor(S): Suhai, Sandor. Publisher: Plenum, New York, N.Y를 포함한다. 본 개시내용의 맥락 내에서 서열 분석 소프트웨어가 분석을 위해 사용되는 경우, 분석 결과는 참조된 프로그램의 "디폴트 값"에 기반한 것으로 이해될 것이다. "디폴트 값"은 처음 초기화될 때 소프트웨어와 함께 원래 로드되는 임의의 값 또는 매개변수 세트를 의미할 것이다.

[0078] ***"개선하다," "증가하다," "억제하다," 또는 "감소하다"***: 본원에 사용된 바와 같이, 용어 "개선하다", "증가하다", "억제하다", 및 "감소하다", 및 이의 문법적 등가물은 참조와의 정성적 또는 정량적 차이를 나타낸다.

[0079] ***단리된:*** 본원에 사용된 바와 같이, "단리된"은 (1) 처음에 생산될 때(자연에서든 및/또는 실험적 설정에서든) 회합된 구성요소의 적어도 일부로부터 분리되고/되거나, (2) 인간의 손에 의해 설계, 생산, 제조, 및/또는 제작된 물질 및/또는 독립체를 지칭한다. 단리된 물질 및/또는 독립체는 처음에 회합된 다른 구성요소의 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 또는 99% 초과로부터 분리될 수 있다. 일부 구현예에서, 단리된 제제는 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 또는 99% 초과로 순수하다. 본원에 사용된 바와 같이, 물질은 다른 구성요소가 실질적으로 없는 경우 "*순수*"하다. 일부 구현예에서, 당업자에 의해 이해되는 바와 같이, 물질은 예를 들어, 하나 이상의 담체 또는 부형제(예를 들어, 완충액, 용매, 물 등)와 같은 특정 다른 구성요소와 조합된 후에도 여전히 "*단리된*" 또는 심지어 "*순수한*" 것으로 간주될 수 있으며; 이러한 구현예에서, 물질의 단리 또는 순도 퍼센트는 이러한 담체 또는 부형제를 포함하지 않고 계산된다. 예를 하나 들자면, 일부 구현예에서, 자연에서 발생하는 폴리펩티드 또는 폴리뉴클레오티드와 같은 생물학적 중합체는 a) 기원 또는 유래원에 의해 자연에서 천연 상태로 수반하는 구성요소의 일부 또는 전부와 연관되지 않고; b) 자연에서 생산하는 종과 동일한 종의 다른 폴리펩티드 또는 핵산이 실질적으로 없고; c) 자연에서 생산하는 종의 것이 아닌 세포 또는 다른 발현 시스템으로부터의 구성요소에 의

해 발현되거나 또는 달리 연관될 때 "*단리된*" 것으로 간주된다. 따라서, 예를 들면, 일부 구현예에서, 화학적으로 합성되거나 또는 자연에서 생산하는 것과 상이한 세포 시스템에서 합성된 폴리펩티드는 "*단리된*" 폴리펩티드인 것으로 간주된다. 대안적으로 또는 추가적으로, 일부 구현예에서, 하나 이상의 정제 기술에 적용된 폴리펩티드는 a) 자연에서 연관되고/되거나; b) 처음에 생산될 때 연관된 것이 있는 다른 구성요소로부터 분리된 정도로 "*단리된*" 폴리펩티드인 것으로 간주될 수 있다.

[0080] ***작동가능하게 연결된:*** 본원에 사용된 바와 같이, "작동가능하게 연결된" 또는 "작동적으로 연결된"은 구성요소 요소가 의도된 방식으로 기능하도록 허용하는 관계에 있도록 적어도 제1 요소 및 제2 요소의 회합을 지칭한다. 예를 들어, 핵산 조절 서열은 조절 서열 및 코딩 서열이 조절 서열에 의해 코딩 서열에 대한 발현의 제어를 허용하는 방식으로 회합되는 경우 핵산 코딩 서열에 "*작동가능하게 연결된*"다. 일부 구현예에서, "*작동가능하게 연결된*" 조절 서열은 (예를 들어, 단일 핵산에서) 코딩 서열과 직접적으로 또는 간접적으로 공유적으로 회합된다. 일부 구현예에서, 조절 서열은 코딩 서열의 발현을 *트랜스*로 제어하고 코딩 서열과 동일한 핵산에서 조절 서열의 포함은 작동가능한 연결의 요건이 아니다.

[0081] ***약제학적으로 허용되는:*** 본원에 사용된 바와 같이, 용어 "약제학적으로 허용되는," 본원에 개시된 바와 같은 조성물의 제형을 위한 하나 이상의, 또는 모든 구성요소(들)에 적용되는 바와 같이, 각 구성요소가 조성물의 다른 성분과 호환가능해야 하며 이의 수용자에게 해롭지 않음을 의미한다.

[0082] ***약제학적으로 허용되는 담체:*** 본원에 사용된 바와 같이, 용어 "약제학적으로 허용되는 담체"는 제제(예를 들어, 약제학적 제제)의 제형을 용이하게 하거나, 제제의 생체이용률을 변형시키거나, 또는 대상체의 하나의 기관 또는 부분에서 또 다른 기관 또는 부분까지 제제의 수송을 용이하게 하는 액체 또는 고체 충전제, 희석제, 부형제, 또는 용매 캡슐화 물질과 같은 약제학적으로-허용되는 물질, 조성물, 또는 비히클을 지칭한다. 약제학적으로-허용되는 담체로서 제공할 수 있는 물질의 일부 예는 다음을 포함한다: 당, 예컨대 락토스, 글루코스 및 수크로스; 전분, 예컨대 옥수수 전분 및 감자 전분; 셀룰로스, 및 이의 유도체, 예컨대 나트륨 카르복시메틸 셀룰로스, 에틸 셀룰로스 및 셀룰로스 아세테이트; 분말화 트라가칸트; 맥아; 젤라틴; 활석; 부형제, 예컨대 코코아 버터 및 좌제 왁스; 오일, 예컨대 땅콩유, 면실유, 홍화유, 참깨유, 올리브유, 옥수수유 및 대두유; 글리콜, 예컨대 프로필렌 글리콜; 폴리올, 예컨대 글리세린, 소르비톨, 만니톨 및 폴리에틸렌 글리콜; 에스테르, 예컨대 에틸 올레에이트 및 에틸 라우레이트; 한천; 완충제, 예컨대 수산화마그네슘 및 수산화알루미늄; 알긴산; 무발열원수; 등장성 염수; 링거 용액; 에틸 알코올; pH 완충 용액; 폴리에스테르, 폴리카르보네이트 및/또는 폴리무수물; 및 약제학적 제형에 이용되는 다른 무독성 호환성 물질.

[0083] ***약제학적 조성물:*** 본원에 사용된 바와 같이, 용어 "약제학적 조성물"은 활성제가 하나 이상의 약제학적으로 허용되는 담체와 함께 제형화된 조성물을 지칭한다.

[0084] ***프로모터:*** 본원에 사용된 바와 같이, "프로모터" 또는 "프로모터 서열"은 코딩 서열의 전사의 개시 및/또는 가공성에 직접적으로 또는 간적적으로(예를 들어, 프로모터-결합된 단백질 또는 물질을 통해) 참여하는 DNA 조절 영역일 수 있다. 프로모터는 적합한 조건 하에, 하나 이상의 전사 인자 및/또는 조절 모이어티와 프로모터의 결합 시 코딩 서열의 전사를 개시할 수 있다. 코딩 서열의 전사 개시에 참여하는 프로모터는 코딩 서열에 "작동가능하게 연결될" 수 있다. 특정 경우에, 프로모터는 전사 개시 부위(이의 3' 말단)에서 상류(5' 방향) 위치로 확장되는 DNA 조절 영역일 수 있거나 또는 이를 포함할 수 있으며, 이렇게 지정된 서열이 전사 사건을 시작하는데 필요한 최소 수의 염기 또는 요소 중 하나 또는 둘 다를 포함하도록 한다. 프로모터는 인핸서 및 억제 인자 서열과 같은 발현 제어 서열일 수 있거나, 이를 포함하거나, 또는 이와 작동가능하게 회합되거나 또는 이에 작동가능하게 연결될 수 있다. 일부 구현예에서, 프로모터는 유도성일 수 있다. 일부 구현예에서, 프로모터는 구성적 프로모터일 수 있다. 일부 구현예에서, 조건부(예를 들어, 유도성) 프로모터는 단방향성 또는 양방향성일 수 있다. 프로모터는 특정 종의 게놈에서 발생하는 것으로 알려진 서열과 동일한 서열일 수 있거나 또는 이를 포함할 수 있다. 일부 구현예에서, 프로모터는 하이브리드 프로모터일 수 있거나 또는 이를 포함할 수 있으며, 여기서 전사 조절 영역을 함유하는 서열은 하나의 공급원으로부터 수득될 수 있고 전사 개시 영역을 함유하는 서열은 제2 공급원으로부터 수득될 수 있다. 이식유전자 내의 코딩 서열에 제어 요소를 연결하기 위한 시스템은 당업계에 잘 알려져 있다(일반적인 분자 생물학적 및 재조합 DNA 기술은 Sambrook, Fritsch, 및 Maniatis, *Molecular Cloning: A Laboratory Manual*, Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989에 기재되어 있다).

[0085] ***참조:*** 본원에 사용된 바와 같이, "참조"는 비교가 수행되는 것과 관련된 표준 또는 대조군을 지칭한다. 예를 들어, 일부 구현예에서, 제제, 샘플, 서열, 대상체, 동물, 또는 개인, 또는 이의 집단, 또는 이를 나타내는 척

도 또는 특성은 참조, 제제, 샘플, 서열, 대상체, 동물, 또는 개인, 또는 이의 집단, 또는 이를 나타내는 척도 또는 특성과 비교된다. 일부 구현예에서, 참조는 측정된 값이다. 일부 구현예에서, 참조는 확립된 표준 또는 예상된 값이다. 일부 구현예에서, 참조는 역사적 참조이다. 참조는 정성적의 정량적일 수 있다. 전형적으로, 당업자에 의해 이해되는 바와 같이, 참조 및 비교되는 값은 비슷한 조건 하에 척도를 나타낸다. 당업자는 신뢰성 및/또는 비교를 정당화하기 위한 충분한 유사성이 존재하는 경우 이해할 것이다. 일부 구현예에서, 적절한 참조는 예를 들어, 하나 이상의 특정 값(예를 들어, 제제 또는 조건의 존재 또는 부재), 또는 이를 나타내는 척도 또는 특성을 평가하는 목적을 위해, 당업자가 비교할 수 있는 것으로 인식할 조건 하에, 제제, 샘플, 서열, 대상체, 동물, 또는 개인, 또는 이의 집단일 수 있다.

[0086] ***조절 서열***: 핵산 코딩 서열의 발현의 맥락에서 본원에 사용된 바와 같이, 조절 서열은 코딩 서열의 발현을 제어하는 핵산 서열이다. 일부 구현예에서, 조절 서열은 유전자 발현의 하나 이상의 측면(예를 들어, 세포-유형-특이적 발현, 유도성 발현 등)을 제어하거나 또는 영향을 미칠 수 있다.

[0087] ***대상체:*** 본원에 사용된 바와 같이, 용어 "대상체"는 유기체, 전형적으로 포유동물(예를 들어, 인간, 래트, 또는 마우스)을 지칭한다. 일부 구현예에서, 대상체는 질환, 장애 또는 병태를 앓고 있다. 일부 구현예에서, 대상체는 질환, 장애, 또는 병태에 취약하다. 일부 구현예에서, 대상체는 질환, 장애 또는 병태의 하나 이상의 증상 또는 특성을 나타낸다. 일부 구현예에서, 대상체는 질환, 장애 또는 병태를 앓고 있지 않다. 일부 구현예에서, 대상체는 질환, 장애, 또는 병태의 임의의 증상 또는 특성을 나타내지 않는다. 일부 구현예에서, 대상체는 질환, 장애, 또는 병태에 대한 감수성 또는 위험의 특성인 하나 이상의 특징을 갖는다. 일부 구현예에서, 대상체는 질환, 장애, 또는 병태에 대해 테스트된 대상체, 및/또는 요법이 투여된 대상체이다. 일부 경우에, 인간 대상체는 "환자" 또는 "개인"으로 상호교환가능하게 지칭될 수 있다.

[0088] ***치료제:*** 본원에 사용된 바와 같이, 용어 "치료제"는 대상체에게 투여될 때 원하는 약리학적 효과를 도출하는 임의의 제제를 지칭한다. 일부 구현예에서, 제제는 적절한 집단에 걸쳐 통계적으로 유의한 효과를 입증하는 경우 치료제인 것으로 간주된다. 일부 구현예에서, 적절한 집단은 모델 유기체 집단 또는 인간 집단일 수 있다. 일부 구현예에서, 적절한 집단은 특정 연령군, 성별, 유전적 배경, 기존 임상 상태 등과 같은 다양한 기준에 의해 정의될 수 있다. 일부 구현예에서, 치료제는 질환, 장애, 또는 병태의 치료에 사용될 수 있는 물질이다. 일부 구현예에서, 치료제는 인간에게 투여하기 위해 시판될 수 있기 전에 정부 기관에 의해 승인되었거나 또는 승인될 필요가 있는 제제이다. 일부 구현예에서, 치료제는 인간에게 투여하기 위해 의학적 처방이 필요한 제제이다.

[0089] ***치료 유효량:*** 본원에 사용된 바와 같이, "치료 유효량"은 투여되기 위한 원하는 효과를 생성하는 양을 지칭한다. 일부 구현예에서, 용어는 치료 투약 레지멘에 따라 질환, 장애, 및/또는 병태를 앓고 있거나 또는 이에 취약한 집단에게 투여될 때 질환, 장애, 및/또는 병태를 치료하기에 충분한 양을 지칭한다. 일부 구현예에서, 치료 유효량은 질환, 장애, 및/또는 병태의 하나 이상의 증상의 발병률 및/또는 중증도를 감소시키고/시키거나, 이의 발병을 지연시키는 것이다. 당업자는 용어 "*치료 유효량*"이 실제로 특정 개인에서 달성될 성공적인 치료를 필요로 하지 않음을 이해할 것이다. 오히려, 치료 유효량은 이러한 치료를 필요로 하는 환자에게 투여될 때 상당한 수의 대상체에서 특정한 원하는 약리학적 반응을 제공하는 양일 수 있다. 일부 구현예에서, 치료 유효량에 대한 언급은 하나 이상의 특이적 조직(예를 들어, 질환, 장애 또는 병태에 걸린 조직) 또는 체액(예를 들어, 혈액, 타액, 혈청, 땀, 눈물, 소변 등)에서 측정된 바와 같은 양에 대한 언급일 수 있다. 당업자는 일부 구현예에서, 특정 제제 또는 요법의 치료 유효량이 단일 용량으로 제형화 및/또는 투여될 수 있음을 이해할 것이다. 일부 구현예에서, 치료적으로 유효한 제제는 예를 들어, 투약 레지멘의 일부로서 복수의 용량으로 제형화 및/또는 투여될 수 있다.

[0090] ***치료***: 본원에 사용된 바와 같이, 용어 "치료"(또한 "치료하다" 또는 "치료하는")는 특정 질환, 장애, 또는 병태의 하나 이상의 증상, 특징, 및/또는 원인의 발병을 부분적으로 또는 완전히 완화하고/하거나, 개선하고/하거나, 덜어주고/주거나, 억제하고/하거나, 지연시키고/시키거나, 이의 중증도를 감소시키고/시키거나, 이의 발병률을 감소시키거나, 또는 임의의 이러한 결과를 달성하려는 목적을 위해 투여되는 요법의 투여를 지칭한다. 일부 구현예에서, 이러한 치료는 관련 질환, 장애, 또는 병태의 징후를 나타내지 않는 대상체 및/또는 질환, 장애, 또는 병태의 초기 징후만을 나타내는 대상체의 치료일 수 있다. 대안적으로 또는 추가적으로, 이러한 치료는 관련 질환, 장애 및/또는 병태의 하나 이상의 확립된 징후를 나타내는 대상체의 치료일 수 있다. 일부 구현예에서, 치료는 관련 질환, 장애, 및/또는 병태를 앓고 있는 것으로 진단된 대상체의 치료일 수 있다. 일부 구현예에서, 치료는 관련 질환, 장애, 또는 병태의 발병 위험 증가와 통계적으로 상관관계가 있는 하나 이상의 감수성 인자를 갖는 것으로 알려진 대상체의 치료일 수 있다. "예방적 치료"는 치료가 병태가 발병할 위

험을 줄이거나, 방지하거나, 또는 감소시키는 목적을 위해 투여되도록 치료될 병태의 징후 또는 증상을 나타내지 않거나 또는 치료될 병태의 초기 징후 또는 증상만을 나타내는 대상체에게 투여되는 치료를 포함한다. 따라서, 예방적 치료는 병태에 대한 예방 치료로서 기능을 한다. "치료적 치료"는 병태의 증상 또는 징후를 나타내는 대상체에게 투여되는 치료를 포함하며 병태의 중증도 또는 진행을 줄이려는 목적을 위해 대상체에게 투여된다.

[0091] ***단위 용량***: 본원에 사용된 바와 같이, 용어 "단위 용량"은 단일 용량 및/또는 약제학적 조성물의 물리적으로 별개의 단위로 투여되는 양을 지칭한다. 많은 구현예에서, 단위 용량은 미리결정된 양의 활성제, 예를 들면 미리결정된 바이러스 역가(주어진 부피에서 바이러스, 비리온, 또는 바이러스 입자의 수)를 함유한다. 일부 구현예에서, 단위 용량은 제제의 전체 단일 용량을 함유한다. 일부 구현예에서, 하나 초과의 단위 용량은 총 단일 용량을 달성하기 위해 투여된다. 일부 구현예에서, 의도된 효과를 달성하기 위해 다중 단위 용량의 투여가 필요하거나, 또는 필요할 것으로 예상된다. 단위 용량은 예를 들어, 미리결정된 양의 하나 이상의 치료 모이어티, 미리결정된 양의 고체 형태의 하나 이상의 치료 모이어티, 미리결정된 양의 하나 이상의 치료 모이어티를 함유하는 지속 방출 제형 또는 약물 전달 장치 등을 함유하는 액체(예를 들어, 허용되는 담체)의 부피일 수 있다. 단위 용량은 치료 모이어티(들) 이외에 다양한 구성요소 중 임의의 것을 포함하는 제형에 존재할 수 있음이 이해될 것이다. 예를 들어, 허용되는 담체(예를 들어, 약제학적으로 허용되는 담체), 희석제, 안정화제, 완충제, 보존제 등이 포함될 수 있다. 많은 구현예에서, 특정 치료제의 적절한 총 일일 투여량은 단위 용량의 일부, 또는 복수를 포함할 수 있고, 예를 들어, 건전한 의학적 판단의 범위 내에서 의사에 의해 결정될 수 있음이 당업자에 의해 이해될 것이다. 일부 구현예에서, 임의의 특정 대상체 또는 유기체에 대한 특이적 유효 용량 수준은 치료되는 장애 및 장애의 중증도; 이용되는 특이적 활성 화합물의 활성; 이용되는 특이적 조성물; 대상체의 연령, 체중, 일반적 건강, 성별, 및 식단; 이용되는 특이적 활성 화합물의 투여 시간, 및 배출률; 치료 지속기간; 이용되는 특이적 화합물(들)과 조합하거나 또는 동시에 사용되는 약물 및/또는 추가적인 요법을 포함하는 다양한 인자, 및 의학 분야에서 잘 알려진 유사 인자에 따라 달라질 수 있다.

## *도면의 간단한 설명*

[0092] 본원에 제출된 많은 도면은 컬러로 더 잘 이해된다. 출원인은 도면의 컬러 버전을 원본 제출의 일부로 간주하고 이후 절차에서 도면의 컬러 이미지를 제시할 권리를 보유한다.

도 1. 예시적인 벡터 도식. 예시적인 벡터 도식은 제공된 Ad35 벡터의 구현예에서 유용한 통합된 카세트 및 일시적 발현 카세트에서 구성요소의 가능한 배열을 나타낸다. 통합된 카세트는 frt 부위 사이에 트랜스포존 및 다른 구성요소를 포함한다. HDAd 벡터는 γ-글로빈, GFP, mCherry, 및 hFVIII(ET3)과 같은 발현 산물(Exp. Product); EF1α, PGK 프로모터, 또는 β 프로모터와 같은 프로모터(들); $mgmt^{P140K}$와 같은 선택 마커(들); 프로모터, polyA 꼬리, 및/또는 절연체(예컨대 cHS4)와 같은 조절 요소(Reg. Elements)를 포함할 수 있다. 일시적 발현 카세트는 유사한 구성요소, 뿐만 아니라 DNA 절단 분자(들)(예를 들어, spCas9) 또는 염기 편집기(들) 및 게놈 표적화 가이드(GTG; 예를 들어 sgRNA)를 포함한다. 트랜스포사제 벡터는 표적화 재조합효소(예를 들어, FlpE) 및 트랜스포사제(예를 들어, SB100x)를 포함한다. 벡터는 일 배향/방향으로 예시되어 있지만, 대안적으로 역방향으로 제공될 수 있다.

도 2a-2f. 혈색소병증의 HSPC 유전자 요법을 위한 통합 HDAd5/35++ 벡터. (도 2a) 벡터 구조. HDAd-γ-글로빈/mgmt에서, 11.8-kb 트랜스포존은 HDAd-SB 벡터로부터 제공된 과활성 잠자는 미녀(Sleeping Beauty) 트랜스포사제(SB100X)를 통해 통합을 위한 도립된 트랜스포존 반복부(IR) 및 FRT 부위에 의해 플랭킹된다(오른쪽 패널). γ-글로빈 발현 카세트는 4 개의 DNase 과민성(HS) 영역을 포함하는 β-글로빈 LCR의 4.3-kb 버전 및 0.7-kb β-글로빈 프로모터을 함유한다. 3'-UTR을 포함하는 76-Ile HBG1 유전자를 사용하였다(적혈구에서 mRNA 안정화를 위함). LCR/β-프로모터 및 EF1A 프로모터 사이의 간섭을 피하기 위해, 1.2-kb 닭 HS4 염색질 절연체(Ins)를 카세트 사이에 삽입하였다. HDAd-SB 벡터는 보편적으로 활성인 PGK 및 EF1A 프로모터 각각의 제어 하에 활성-향상된 SB100X 트랜스포사제 및 Flpe 재조합효소에 대한 유전자를 함유한다. (도 2b) 동원된 CD46tg 마우스의 생체내 형질도입. HSPC를 4 일 동안 인간 재조합 G-CSF의 s.c. 주사 이어서 AMD3100의 1회 s.c. 주사에 의해 동원하였다. AMD3100 주사 30 및 60 분 후, 동물에게 HDAd-γ-글로빈/mgmt 및 HDAd-SB의 1:1 혼합물을 i.v. 주사하였다(2 회 주사, 각각 4 X $10^{10}$ 개의 바이러스 입자). 마우스를 인간 γ-글로빈 및 $MGMT^{P140K}$에 대한 면역 반응을 피하기 위해 다음 4 주 동안 면역억제성(IS) 약물로 처리하였다. $O^6$-BG/BCNU 처리를 4 주차에 시작하고 2 주마다 3 회 반복하였다. 각 주기마다 BCNU 농도를 5에서 7.5 내지 10 mg/kg으로 증가

시켰다. 면역억제는 마지막 $O^6$-BG/BCNU 주사 후 2 주에 재개하였다. (도 2c) 유세포 분석에 의해 측정된 인간 γ-글로빈$^+$ 말초 RBC의 백분율. (도 2d) 말초 혈액 단핵 세포(MNC), 전체 세포, 적혈구 Ter119$^+$ 세포, 및 비적혈구 Ter119$^-$ 세포에서 인간 γ-글로빈$^+$ 세포의 백분율. (도 2e) 18 주차에 RBC에서 HPLC에 의해 측정된 성체 마우스 글로빈 쇄(α, β-주요, β-소수)와 비교한 인간 γ-글로빈 단백질의 백분율. (도 2f) 18 주차에 총 말초 혈액 세포에서 RT-qPCR에 의해 측정된 성체 마우스 β-주요 글로빈 mRNA와 비교한 인간 γ-글로빈 mRNA의 백분율. 어떠한 처리도 받지 않은 마우스를 대조군으로 사용하였다. 도 2c-2f에서, 각 기호는 개별 동물을 나타낸다.

도 3. 생체내 형질도입/선택 후 hCD46tg 대조군 마우스 및 대표적인 CD46tg 마우스로부터의 RBC에서 글로빈 쇄의 HPLC 분석. 숫자(볼트)는 피크 강도를 나타낸다. 각 그룹으로부터 총 4 마리의 마우스를 유사한 결과로 분석하였다. 데이터는 도 2e에 요약되어 있다. 도 3에서, 곡선하 면적(AUC) 값은 상응하는 피크의 왼쪽으로 상쇄된다.

도 4a-4c. 생체내 형질도입 후 18 주차에 수확된 골수 Lin- 세포로 이식받은 마우스("2차 수용자")의 분석. (도 4a) PBMC에서 인간 CD46-양성 세포의 백분율에 기반하여 표시된 시점에 혈액 샘플에서 측정된 생착. (도 4b) 20 주차에 골수, 비장, 및 PBMC에서 생착. (도 4c) RBC에서 HPLC에 의해 측정된 인간 γ- 대 마우스 α-글로빈 단백질의 비율. 각 기호는 개별 동물을 나타낸다. 통계적 분석은 비모수 Kruskal-Wallis 검정으로 수행하였다.

도 5a-5e. 20 주차에 2차 수용자의 골수 세포에서 이식유전자 통합 분석. (도 5a) 골수 세포의 마우스 염색체 상의 통합 부위의 국소화. 대표적인 마우스가 제시된다. 각 라인은 통합 부위이다. 이 샘플에서 통합 부위의 수는 2,197이다. (도 5b) 게놈 영역에서 통합 분포. 5 마리 마우스의 통합 부위 데이터를 풀링하고 사용하여 그래프를 생성하였다. (도 5c) 연속적 게놈 창 및 무작위화 마우스 게놈 창과 중첩되는 통합의 수 및 크기를 비교하였다. 풀링된 데이터를 도 5b)에서와 같이 사용하였다. 유사성에 대한 Pearson의 $\chi^2$ 검정 $P$ 값은 0.06381이며, 통합 패턴이 무작위에 가깝다는 것을 암시한다. (도 5d) 이식유전자 카피 수. 형질도입되지 않은 대조군 마우스 및 20 주차에 2차 수용자로부터의 총 골수 세포의 게놈 DNA를 인간 γ-글로빈-특이적 프라이머를 사용하여 qPCR에 적용하였다. 개별 동물에 대한 세포 당 카피 수가 제시되어 있다. 각 기호는 개별 동물을 나타낸다. (도 5e) 개별 클론 전구체 콜로니에서 이식유전자 카피 수. 골수 Lin$^-$ 세포를 메틸셀룰로스에 플레이팅하고, 개별 콜로니를 15 일 후에 선별하였다. qPCR을 게놈 DNA에 대해 수행하였다. 세포 당 이식유전자 카피 수로서 표현된 개별 콜로니에서 정규화된 qPCR 신호가 제시되어 있다($n$ = 113). 각 기호는 단일 세포로부터 유래된 개별 콜로니에서 카피 수를 나타낸다.

도 6. VCN을 측정하기 위한 단일 세포-유래 전구체 콜로니의 qPCR(도 7e 참조).

도 7a-7e. CD46tg 마우스에서 생체내 HSPC 형질도입/선택 후 혈액학적 매개변수(HDAd 주사 후 18 주). (도 7a) WBC 계수. (도 7b) 미처리 마우스 및 HDAd-γ-글로빈/mgmt 및 HDAd-SB 주사 후 18 주차 마우스의 대표적인 혈액 도말. 스케일 바: 20 μm. WBC의 핵은 보라색으로 염색된다. (도 7c) 혈액학적 매개변수. Hb, 헤모글로빈; HCT, 적혈구용적률; MCV, 평균 미립자 부피; MCH, 평균 미립자 헤모글로빈; MCHC, 평균 미립자 헤모글로빈 농도; RDW, 적혈구 분포 폭. $n \geq 3$, $^*P < 0.05$. 통계적 분석은 이원 ANOVA를 사용하여 수행하였다. (도 7d) 18 주차에 희생된 미처리(naive) 마우스(대조군) 및 처리된 마우스의 세포 골수 조성물. 계통 마커-양성 세포(Ter119+, CD3+, CD19+, 및 Gr-1+ 세포) 및 HSPC(LSK 세포)의 백분율이 제시되어 있다. (도 7e) 생체내 형질도입 후 18 주차에 수확된 골수 Lin- 세포의 콜로니-형성 가능성. 2,500 개의 Lin-세포의 플레이팅 후 형성된 콜로니의 수가 제시되어 있다. 도 7a 및 도 7c-7e에서, 각 기호는 개별 동물을 나타낸다. NE, 호중구; LY, 림프구; MO, 단핵구; BA, 호염구.

도 8. CD46++/Bhhth-3 지중해빈혈 모델의 생성. 암컷 CD46tg 마우스를 수컷 Hbbth-3 마우스와 교배시켰다. F1 하이브리드 마우스를 hCD46+/+ 마우스와 역교배시켜 hCD46+/+에 대해 동형접합성인 Hbbth-3 마우스를 생성하였다

도 9a-9c. CD46+/+/Hbbth-3 마우스 지중해빈혈 모델의 표현형. (도 9a) CD46tg($n$ = 3) 및 Hbbth-3 마우스($n$ = 3)와 비교한 CD46+/+/Hbbth-3 마우스($n$ = 7)의 혈액학적 매개변수. 각 기호는 개별 동물을 나타낸다. $^*P \leq 0.05$, $^{**}P \leq 0.0002$, $^{***}P \leq 0.00003$. 통계적 분석은 이원 ANOVA를 사용하여 수행하였다. RET, 망상적혈구. (도 9b) May-Grunwald/Giemsa로 염색한 후 대표적인 말초 혈액 도말. 스케일 바: 20 μm. (도

9c) CD46tg 마우스의 비장 및 간 절편(왼쪽 상단 2 개 패널)과 비교하여 $CD46^{+/+}$/Hbbth-3 마우스의 간 및 비장 절편(왼쪽 하단 2 개 패널)에서 H&E 염색에 의한 골수외 혈구형성. 스케일 바: 20 μm. 간에서 적혈구모세포의 클러스터는 왼쪽 하단 패널에 표시된다. 중간 하단 패널에서 원은 비장의 거핵구를 표시한다. 비장에서 Perl의 프러시안 블루 염색에 의한 철분 침착(입상 푸른빛 침전물)은 CD46tg의 경우 오른쪽 상단 패널 및 $CD46^{+/+}$/Hbbth-3 마우스의 경우 오른쪽 하단 패널에 제시되어 있다. 스케일 바: 25 μm.

도 10. "건강한" CD46tg 마우스와 비교하여 지중해빈혈 마우스(Hbbth-3 및 CD46+/+/Hbbth-3)의 백혈구 분석. WBC: 백혈구, NEU: 호중구, LY: 림프구, MONO: 단핵구. *p≤0.05, ** p≤0.0002, ***p ≤0.00003. 이들은 처리 전 마우스의 기준선 수준이다. (CD46tg의 경우 n=8, Hbbth3의 경우 n=4, CD46++/Hbbth3의 경우 n=20). 각 기호는 개별 동물을 나타낸다. 통계적 분석은 비모수 Kruskal-Wallis 검정으로 수행하였다.

도 11. CD46+/+/Hbbth-3 마우스에서 HSPC의 동원. 마지막 AMD3100 주사 1 시간 후에 말초 혈액에서 동원된 LSK(계통-/Sca-1+/c-Kit+/) 세포의 수가 제시되어 있다. n=17 동원된 마우스; n=3 미처리 마우스. 통계적 분석은 비모수 Kruskal-Wallis 검정으로 수행하였다.

도 12. 동원된 CD46+/+/Hbbth-3 마우스의 생체내 형질도입/선택. 동원된 CD46+/+/Hbbth3 마우스의 생체내 형질도입. HSPC를 6 일 동안 인간 재조합 G-CSF의 s.c. 주사(1-6 일) 이어서 AMD3100/플레릭사포르(Plerixafor)의 3 회 s.c 주사(5-7 일)에 의해 동원하였다. 플레릭사포르 주사 30 및 60 분 후, 동물에게 HDAd-γ-글로빈/mgtm + HDAd-SB의 1:1 혼합물을 정맥내로 주사하였다(2 회 주사, 각각 $4x10^{10}$ vp). 생체내 형질도입 후, 인간 γ-글로빈 및 $MGMT^{P140K}$ 단백질에 대한 면역 반응을 피하기 위해 면역-억제를 17 주 동안 투여하였다. 17 주차에, 처리된 마우스는 2차 이식을 위한 공여자로서 제공되거나 또는 $O^6$-BG/BCNU로 생체내 선택에 적용되었다. 2차 C57Bl/6 수용자를 면역억제 하에 16 주 동안 추적한 다음 희생시켰다. 생체내 선택에 적용된 마우스는 격주마다 증가하는(5, 7.5, 10, 10 mg/kg) $O^6$-BG/BCNU 처리를 받았다. 마지막 $O^6$-BG/BCNU 용량 2 주 후에 면역-억제를 재개하였다. 29 주차에, 마우스를 희생시키고, 이들의 골수를 C57Bl/6 2차 수용자 내로 이식하였다.

도 13a-13f. $O^6$BG/BCNU 처리를 받지 않은 생체내-형질도입된 $CD46^{+/+}$/Hbbth-3 마우스의 분석. (도 13a) 유세포 분석에 의해 측정된 말초 RBC에서 인간 γ-글로빈의 백분율. 실험을 3 회 수행하고, 상이한 기호 모양으로 표시하였다. (도 13b) 적혈구($Ter119^+$) 및 비적혈구($Ter119^-$) 혈액 세포에서 γ-글로빈 발현. 일원 ANOVA 검정에 의해 $***P \leq 0.00003$. (도 13c) 건강한(CD46tg) 마우스($n = 3$), 동원 및 생체내 형질도입 전의 $CD46^{+/+}$/Hbbth-3 마우스($n = 14$), 및 생체내 형질도입을 겪고 16 주차에 분석된 $CD46^{+/+}$/Hbbth-3 마우스($n = 8$)의 RBC 분석. $*P \leq 0.05$. 통계적 분석은 이원 ANOVA를 사용하여 수행하였다. (도 13d) 조직학적 표현형. 상단: 혈액 도말. 중간: 망상적혈구 검출을 위해 브릴리언트 크레실 블루(Brilliant cresyl blue)를 사용한 말초 혈액 도말의 초생체 염색. 대표적인 도말에서 양성으로 염색된 망상적혈구의 백분율은 다음과 같았다: CD46tg의 경우, 8% ± 0.8%; 형질도입 전 $CD46^{+/+}$/Hbbth-3의 경우, 39% ± 1.3%; 및 형질도입 후 16 주차 $CD46^{+/+}$/Hbbth-3의 경우, 26% ± 0.45%. 하단: 골수외 혈구형성. 스케일 바: 20 μm. (도 13e 및 도 13f) 2차 수용자의 분석. 16 주차에 생체내-형질도입된 마우스의 총 골수를 준치사량의 부술판 예조건화를 받은 C57BL/6 마우스 내로 이식하였다. 마우스는 관찰 기간 동안 면역억제를 받았다. (도 13e) 인간 $CD46^+$($hCD46^+$) PBMC의 백분율에 기반한 생착. (C57BL/6 수용자는 hCD46을 발현하지 않는다.) (도 13f) 인간 γ-글로빈$^+$ RBC의 백분율. 각 기호는 개별 동물을 나타낸다.

도 14a-14f. 생체내 선택 후 생체내-형질도입된 $CD46^{+/+}$/Hbbth-3 마우스에서 γ-글로빈 발현의 분석. (도 14a) 유세포 분석에 의해 측정된 말초 RBC에서 인간 γ-글로빈의 백분율. 화살표는 $O^6$-BG/BCNU 처리 시점을 나타낸다. 상이한 기호는 3 개의 독립적인 실험을 나타낸다. 16 주차까지의 데이터는 도 13a에서의 것들과 동일하다. (도 14b) 유세포 분석에 의해 분석된 희생 시(29 주) 조혈 조직에서 γ-글로빈-발현 세포의 백분율. $*P \leq 0.05$, $**P \leq 0.0002$, $***P \leq 0.00003$. (도 14c) MACS-정제된 Ter119 세포에서 γ-글로빈 발현. 29 주차에 1차 수용자로부터의 골수 세포를 $Ter119^+$ 세포에 대해 면역자기적으로 선택하였다. γ-글로빈 발현을 유세포 분석에 의해 $Ter119^+$ 및 $Ter119^-$ 세포에서 측정하였다. $***P \leq 0.0002$. (도 13d) 생체내 선택 전 대 후

(16 주 vs. 29 주)에 말초 혈액, 골수, 및 비장에서 γ-글로빈$^{+}$ 적혈구(Ter119$^{+}$) 및 비적혈구(Ter119$^{-}$) 세포의 배수 강화. $n$ = 5, $^{**}P \leq 0.0002$. (도 14e) RBC에서 HPLC에 의해 측정된 마우스 α-글로빈 단백질과 비교한 인간 γ-글로빈 단백질의 백분율. 통계적 분석은 비모수 Kruskal-Wallis 검정으로 수행하였다. (도 14f) 말초 혈액 세포에서 RT-qPCR에 의해 측정된 성체 마우스 β-주요 글로빈 mRNA에 대한 인간 γ-글로빈 mRNA의 수준. 미처리 CD46$^{+/+}$/Hbbth-3 마우스를 대조군으로 사용하였다. 각 기호는 개별 동물을 나타낸다.

도 15a-15d. RBC에서 글로빈 쇄의 HPLC 분석. (도 15a) 대조군 CD46tg 마우스에서 마우스 글로빈 피크의 대표적인 크로마토그램. 성체 마우스 알파(α), 베타(β)-소수, 및 β-주요 글로빈에 대한 피크가 표지된다. (도 15b-15d) CD46$^{+/+}$/Hbbth-3 마우스(#71)의 RBC 크로마토그램. 이들 마우스는 β-소수 및 β-주요 유전자 결실에 대해 이형접합성임을 유의한다. 약 29 분의 추가 피크가 이와 연관될 수 있다. (도 15d)에서, 인간 γ-글로빈에 특이적인 피크가 표지된다. 대표적인 크로마토그램이 제시되어 있다. 숫자(볼트)는 피크 강도를 나타낸다. 도 15c 및 15d에서, AUC 값은 상응하는 피크의 왼쪽으로 상쇄된다.

도 16. 29 주차에 처리된 CD46++/Hbbth-3 마우스의 DNA 분석. 골수 세포 당 이식유전자(γ-글로빈) 카피 수. 각 기호는 개별 동물을 나타낸다.

도 17a-17e. 생체내 HSPC 형질도입/선택에 의한 CD46+/+/Hbbth-3 마우스의 표현형 교정. (도 17a) 건강한(CD46tg) 마우스, 동원 및 생체내 형질도입 전의 CD46$^{+/+}$/Hbbth-3 마우스, 및 생체내 형질도입/선택을 겪은(HDAd 투입 후 29 주차에 분석됨) CD46$^{+/+}$/Hbbth-3 마우스의 RBC 분석($n$ = 5). $^{*}P \leq 0.05$, $^{**}P \leq 0.0002$, $^{***}P \leq 0.00003$. 통계적 분석은 이원 ANOVA를 사용하여 수행하였다. (도 17b) 망상적혈구 검출을 위해 브릴리언트 크레실 블루를 사용한 말초 혈액 도말의 초생체 염색. 화살표는 특징적인 잔여 RNA 및 미세-세포기관을 함유하는 망상적혈구를 나타낸다. 대표적인 도말에서 양성으로 염색된 망상적혈구의 백분율은 다음과 같았다: CD46의 경우, 7%; 처리전 CD46$^{+/+}$/Hbbth-3의 경우, 31%; 및 처리후 CD46$^{+/+}$/Hbbth-3의 경우, 12%. 스케일 바: 20 μm. (도 17c) 상단: 혈액 도말. 스케일 바: 20 μm. 중간: 골수 사이토스핀. 화살표는 상이한 성숙 단계에서 적혈구모세포 및 처리된 마우스에서 전적혈구모세포가 우세한 적혈구생성의 역행을 나타낸다. 스케일 바: 25 μm. 하단: Perl의 염색에 의한 조직 혈철소증. 철분 침착은 비장 조직 절편에서 헤모시데린의 세포질 청색 색소로 제시된다. (도 17c) 및 (도 18d)에서 대조군 마우스(CD46tg 및 CD46$^{+/+}$/Hbbth-3, 형질도입 전)에 대한 혈액 도말 이미지는 동일한 샘플로부터의 것이다. (도 17d) 1 개의 대표적인 CD46tg 및 1 개의 미처리 CD46$^{+/+}$/Hbbth-3 마우스 및 5 개의 처리된 CD46$^{+/+}$/Hbbth-3 마우스의 거시적 비장 이미지. (도 17e) 희생 시, 비장 크기를 비장 체중 대 총 체중의 비(mg/g)로 결정하였다. 각 기호는 개별 동물을 나타낸다. 데이터는 평균 ± SEM으로 제시된다. $^{*}P \leq 0.05$. 통계적 분석은 일원 ANOVA를 사용하여 수행하였다.

도 18a-18e. 처리된 CD46$^{+/+}$/Hbbth-3 마우스로부터의 골수 세포가 이식된 2차 C57BL/6 수용자의 분석. (도 18a) 부술판 조건화 또는 전신 조사(TBI) 후 PMBC에서 인간 CD46$^{+}$ (hCD46$^{+}$) 세포의 백분율에 기반한 주변부에서 측정된 생착률. (C57BL/6 수용자는 hCD46을 발현하지 않는다.) (도 18b) 인간 γ-글로빈-발현 말초 혈액 RBC의 백분율. 모든 마우스는 이식 후 4 주부터 시작하여 면역억제를 받았다. (도 18c) hCD46$^{+}$(공여자-유래) 세포에서 γ-글로빈$^{+}$ 세포의 백분율. (도 18c 및 도 18d) 이식 후 20 주차에 2차 C57BL/6 수용자에서 γ-글로빈/CD46 발현(부술판 예조건화). CD46$^{+}$ 세포를 3 개의 대표적인 2차 마우스의 키메라 골수로부터 면역자기적으로 분리하고 유세포 분석에 의해 γ-글로빈 발현에 대해 분석하였다. 특히, 인간과 달리, huCD46tg 마우스는 RBC 상에서 CD46을 발현한다. (도 18c) 희생 시 1차 및 2차 수용자의 γ-글로빈/CD46 마킹률. (도 18d) 2차 수용자의 조혈 조직으로부터의 CD46$^{+}$-선택된 세포에서 γ-글로빈 발현(20 주). 각 기호는 개별 동물을 나타낸다. (도 18e) 신규(두번째) 차례의 HSPC 동원/생체내 형질도입을 받은 2차 수용자에서 γ-글로빈 발현($n$ = 5). 2차 수용자(부술판-예조건화)를 이식 후 20 주차에 γ-글로빈 및 CD46 발현에 대해 분석하였다("생체내 형질도입 전"). 그런 다음 이들 마우스를 동원하고 HDAd-γ-글로빈 및 HDAd-SB 벡터로 생체내 형질도입하였다. 생체내 형질도입 4 주 후, 마우스를 희생시키고 분석하였다("생체내 형질도입 후 4 주"). $^{***}P \leq 0.00003$. 통계적 분석은 일원 ANOVA를 사용하여 수행하였다.

도 19a-19d. CD46$^{+/+}$/Hbbth-3 마우스 모델에서 생체내 형질도입/선택의 안전성. (도 19a) 생체내 선택 동안 및

후에 WBC 및 혈소판(PLT) 계수. $O^6$BG/BCNU 처리는 별표로 표시된다. $n \geq 3$. (도 19b) 순환 WBC 하위집단의 절대 수. $n \geq 3$. (도 19c) 29 주차에 희생된 대조군 및 처리된 마우스에서 세포 골수 조성물. 계통 마커-양성 세포($Ter119^+$, $CD3^+$, $CD19^+$, 및 $Gr\text{-}1^+$ 세포) 및 HSPC(LSK 세포)의 백분율이 제시된다. (도 19d) 29 주차에 수확된 골수 세포의 콜로니-형성 가능성. 각 기호는 개별 동물을 나타낸다. $*P \leq 0.05$, $**P \leq 0.0002$, $***P \leq 0.00003$. 통계적 분석은 이원 ANOVA를 사용하여 수행하였다. NEU: 호중구; LY: 림프구; MO: 단핵구.

도 20a-20f. 두번째 차례의 형질도입에 대한 항-HDAd5/$35^{++}$ 항체의 효과. (도 20a) CD46tg 마우스를 동원하고 HDAd-mgmt/GFP + HDAd-SB를 주사하였다. 혈청 샘플을 나타낸 바와 같이 수집하였다. (도 20b, 도 20c) 동원/형질도입 후 4 일차 및 4 주차에 PBMC의 유세포 분석 분석. (도 20d) 4 주차에 두번째 차례의 동원/형질도입 및 후속 GFP 분석. (도 20e) $OD_{450}$에 기반한 항-HDAd5/$35^{++}$ 항체 역가. $OD_{450}$=0.2 역가는 중화하는 것으로 간주된다. (도 20f) 상이한 코호트에서 측정된 GFP-양성 PBMC의 백분율(도 20b-20d 참조). Ctrl은 미처리 CD46tg 마우스이다. (도 20e) 및 (도 20f)에서 각 기호는 개별 동물을 나타낸다. 통계적 분석은 비모수 Kruskal-Wallis 검정으로 수행하였다.

도 21a-21d. HDAd 주사 후 18 주차(생체내 선택 10 주)에 벡터 DNA 생물분포(도 21a) 프라이머 설계. 연회색 프라이머는 이식유전자 카세트에 특이적이고 통합 및 에피솜 벡터 DNA를 둘 다 검출할 것이다. 진회색 프라이머는 플라스미드 pHCA로부터 유래된 벡터 스터퍼(stuffer) DNA를 검출할 것이다. SB100x-매개 통합 시, 진회색 프라이머에 대해 상응하는 표적 영역은 손실될 것이다. 따라서 진회색 프라이머를 사용하여 에피솜 벡터 카피를 측정한다. (도 21b) 통합된 이식유전자 카피 수의 표준 곡선. (도 21c) HCA(에피솜 벡터) 카피 수에 대한 표준 곡선. (도 21d) 세포 당 통합된 이식유전자 카피 수. 에피솜 벡터 카피(진회색 프라이머)를 총 벡터 카피에서 차감하였다(연회색 프라이머). 벡터-특이적 신호는 *GAPDH*로 정규화하였다. 각 기호는 개별 동물을 나타낸다.

도 22a-22c. $O^6$BG/BCNU 처리의 돌연변이원성을 평가하기 위한 시험관내 검정. (도 22a) 동결보존으로부터 밤새 회복시킨 후, $CD34^+$ 세포를 2 일 후 세포의 50%에서 GFP 발현이 매개되는 3000 vp/세포의 MOI에서 HDAd-mgmt/GFP 또는 HDAd 대조군으로 형질도입하였다. 그런 다음 세포를 10 mM $O^6$BG 이어서 25 mM BCNU(또는 DMSO 용매)로 2 시간 동안 처리하였다. 세척 후, 세포를 CFU 검정을 위해 메틸셀룰로스에 플레이팅하였다(35 mm 접시 당 3000 개 세포). 콜로니 및 풀링된 세포를 14 일 후에 계수하고 게놈 DNA를 전체 엑솜 서열분석에 적용하였다. (도 22b) 플레이트 당 풀링된 세포의 수. 각 기호는 개별 35 mm 접시에서 세포 수를 나타낸다. 통계적 분석은 비모수 Kruskal-Wallis 검정으로 수행하였다. (도 22c) HDAd-mgmt/GFP + $O^6$BG/BCNU 그룹으로부터의 대표적인 콜로니. 에피솜 바이러스 게놈의 손실로 인해 콜로니 주변부에서 GFP가 희미해지는 대부분의 세포에서 GFP 발현을 입증한다. 스케일 바는 1 mm이다.

도 23. 벡터 구조. *HDAd-짧은-LCR:* 이 벡터는 DNase 과민성 부위(HS) 1 내지 4의 코어 영역으로 이루어진 4.3 kb 미니-LCR 및 0.66 kb β-글로빈 프로모터를 함유한다. 트랜스포존의 길이는 11.8 kb이다. *HDAd-긴-LCR.* γ-글로빈 유전자는 21.5 kb β-글로빈 LCR(chr11: 5292319-5270789), 1.6 kb β -글로빈 프로모터(예를 들면, chr11: 5228631-5227023 또는 chr11: 5228631-5227018) 및 또한 β-글로빈 유전자좌로부터 유래된 3'HS1 영역(chr11: 5206867-5203839)의 제어 하에 있다. 적혈구 세포에서 RNA 안정화를 위해, γ-글로빈 유전자 UTR을 γ-글로빈 유전자의 3' 단부에 연결하였다. 벡터는 또한 mgmtP140K에 대한 발현 카세트를 함유하여 형질도입된 HSPC 및 HSPC 자손의 생체내 선택을 허용한다. γ-글로빈 및 mgmt 발현 카세트는 닭 글로빈 HS4 절연체(cHS4)에 의해 분리된다. 32.4 kb LCR- γ-글로빈/mgmt 트랜스포존은 SB100x에 의해 인식된 도립 반복부(IR) 및 Flpe 재조합효소에 의한 트랜스포존의 원형화를 허용하는 ftr 부위에 의해 플랭킹된다. *HDAd-SB:* 통합에 필요한 제2 벡터는 활성-향상된 *잠자는 미녀* SB100x 트랜스포사제 및 Flpe 재조합효소에 대한 발현 카세트를 함유한다.

도 24a-24f. HDAd-긴-LCR을 사용한 생체외 HSPC 형질도입 연구 후 32.4 kb 트랜스포존의 SB100x-매개 통합. (도 24a) 실험적 레지멘: CD46-유전자이식 마우스로부터의 골수 Lin- 세포를 HDAd-긴-LCR 및 HDAd-SB로 500 vp/세포의 총 MOI에서 형질도입하였다. 배양 1 일후, 마우스 당 1x106 개의 형질도입된 세포를 치명적으로 조사된 C57Bl/6 마우스 내로 이식하였다. 4 주차에, O6BG/BCNU 처리를 시작하고 2 주마다 4 회 반복하였다. 각 주기마다, BCNU 농도를 5 mg/kg에서 7.5 mg/kg, 내지 10 mg/kg(2 회)까지 증가시켰다. 20 주차에, 마우스를 희생시켰다. (도 24b) 유세포 분석에 의해 측정된 인간 γ-글로빈-양성 말초 적혈구(RBC)의 백분율. 각 기

호는 개별 동물이다. (도 24c) 이식 후 20 주차에 적혈구(Ter119$^{+}$) 골수 세포(하부 패널)에서 인간 γ-글로빈-발현을 나타내는 대표적인 유세포 분석 데이터. 상단 패널은 모의-형질도입된 세포로 이식된 마우스를 나타낸다. (도 24d) iPCR 분석의 도식: 5 마이크로그램의 게놈 DNA을 SacI로 소화시키고, 재결찰시키고, 표시된 프라이머를 사용하여 중첩 역PCR에 적용하였다(재료 및 방법 참조). (도 24e) 통합 접합부를 함유하는 클로닝된 플라스미드의 아가로스 겔 전기영동. 표시된 밴드를 절제하고 서열분석하였다. 통합 부위의 염색체 국소화는 겔 아래에 제시된다. (도 24f) 접합부 서열의 예: 5' 단부 벡터 서열, 잠자는 미녀 IR/DR 서열, 통합 접합부(chr15, 6805206) 서열번호: 1; 5' 단부 벡터 서열, 잠자는 미녀 IR/DR 서열, 통합 접합부(chrX, 16897322) 서열번호: 2; 3' 단부 벡터 서열, 잠자는 미녀 IR/DR 서열, 통합 접합부(chr4, 10207667) 서열번호: 3. 벡터체 및 IR/DR 서열은 각각 평문 및 밑줄로 지정된다. 염색체 서열은 굵은 글씨로 지정된다. IR 및 염색체 DNA의 접합부에서 SB100x에 의해 사용되는 TA 디뉴클레오티드는 괄호에 표시된다.

도 25a-25e. 32.4 kb 트랜스포존을 함유하는 HDAd-긴-LCR 및 11.8 kb 트랜스포존을 함유하는 HDAd-짧은-LCR을 사용한 생체내 HSPC 형질도입. (도 25a) 처리 레지멘: hCD46tg 마우스를 동원하고 HDAd-짧은-LCR + HDAd-SB 또는 HDAd-긴-LCR +HDAd-SB를 IV 주사하였다(두 바이러스의 1:1 혼합물의 각각 4x1010 vp 2 회). 5 주 후, O6BG/BCNU 처리를 시작하였다. 각 주기마다, BCNU 농도를 5 mg/kg에서 7.5 mg/kg, 및 10 mg/kg까지 증가시켰다. O6BG 농도는 4 개 처리 모두에서 30 mg/kg이었다. 분석을 위해 동물이 희생된 20 주까지 마우스를 추적하였다. 골수 Lin- 세포를 2차 수용자 내로의 이식을 위해 사용하였다. 그런 다음 2차 수용자를 16 주 동안 추적하였다. (도 25b) 유세포 분석에 의해 측정된 말초 적혈구 (RBC)에서 인간 γ-글로빈-양성 세포의 백분율. 각 기호는 개별 동물이다. 모의-형질도입된 마우스에서, 0.1% 미만의 세포는 γ -글로빈-양성이었다. (도 25c) 생체내 HSPC 형질도입 후 20 주차에 RBC에서 HPLC에 의해 측정된 γ-글로빈 단백질 쇄 수준. 마우스 α-글로빈 단백질 쇄에 대한 인간 γ-글로빈의 백분율이 제시된다. (도 25d) 생체내 HSPC 형질도입 후 20 주차에 전체 혈액에서 qRT-PCR에 의해 측정된 γ-글로빈 mRNA 수준. 마우스 α -글로빈 mRNA에 대한 인간 γ-글로빈 mRNA의 백분율이 제시된다. (도 25e) 생체내 HSPC 형질도입 후 20 주차에 수확된 골수 단핵 세포에서 세포 당 벡터 카피 수. 두 그룹 사이의 차이는 유의하지 않다. 통계적 분석은 이원 ANOVA를 사용하여 수행하였다.

도 26a-26d. 생체내 HSPC 형질도입 후 20 주차에 혈액학적 매개변수. (도 26a) 백혈구(WBC), 호중구(NE), 백혈구(LY), 단핵구(MO), 호산구(EO), 및 호염구(BA). (도 26b) 적혈구생성 매개변수. RBC: 적혈구, Hb: 헤모글로빈, MCV: 평균 미립자 부피, MCH: 평균 미립자 헤모글로빈, MCHC: 평균 미립자 헤모글로빈 농도, RDW: 적혈구 분포 폭. 세 그룹 사이의 차이는 유의하지 않았다. (도 26c) 세포 골수 조성물. (도 26d) 골수 Lin$^{-}$ 세포의 콜로니-형성 가능성. 그룹 사이의 차이는 도 26a-26d에서 유의하지 않았다.

도 27. 삽입 부위 분석의 도식. *잠자는 미녀* 도립 반복부(IR)와 관련하여 HDAd-긴-LCR 벡터에서 NheI 및 KpnI 부위의 국소화가 표시된다. 이들 효소는 근접하지만, SB IR/DR의 외부를 절단하고 통합되지 않은 벡터의 배경을 감소시키는 데 사용된다. 골수 Lin- 세포로부터의 게놈 DNA를 NheI 및 KpnI로 소화시키고, 열 불활성화 후, NlaIII으로 추가로 소화시켰다. NlaIII은 4-절단기이며 작은 DNA 단편을 생성할 것이다. 그런 다음 소화된 DNA를 서열이 알려지고 소화된 NlaIII 단편에 대한 호환가능한 단부가 있는 이중 가닥 올리고와 결찰시켰다. 열-불활성화 및 정리 후, 링커-결찰된 생성물을 선형 증폭에 사용하여, SB 왼쪽 아암으로부터 프라이밍된 단일-가닥(ss) DNA 집단을 생성하였다. 프라이머는 비오티닐화되어, ssDNA는 스트렙타비딘 비드로 수집될 수 있다. 광범위한 세척 후, ssDNA를 비드로부터 융출시키고 2 차례의 중첩 PCR에 의해 추가로 증폭시켰다. PCR 앰플리콘을 겔 정제하고, 클로닝하고, 서열분석하고 마우스 게놈 서열에 맵핑하여 통합 부위를 표시하였다.

도 28a-28d. LAM-PCR/NGS에 의한 HSPC에서 벡터 통합 부위의 분석. HDAd-긴-LCR + HDAd-SB를 사용한 생체내 형질도입 후 20 주차에 수확된 골수 세포로부터 단리된 게놈 DNA. (도 28a) 통합 부위의 염색체 분포. 통합 부위는 수직선으로 표시된다. (도 28b) 접합부 서열의 예: 잠자는 미녀 IR/DR 서열, 통합 접합부(chr7, 79796094) 서열번호: 4; 잠자는 미녀 IR/DR 서열, 통합 접합부(반복 영역) 서열번호: 5. IR/DR 서열은 밑줄 및 굵은 글씨로 지정된다. 염색체 서열은 평문으로 지정된다. IR 및 염색체 DNA의 접합부에서 SB100x에 의해 사용된 TA 디뉴클레오티드는 굵은 글씨이다. (도 28c) 통합 부위를 마우스 게놈에 맵핑하고 유전자에 대한 이의 위치를 분석하였다. 1 kb 상류 전사 시작 부위(TSS)(0.0%), 엑손의 5'UTR(0.0%), 단백질 코딩 서열(0.0%), 인트론(17.0%), 3'UTR(0.0%), 3'UTR에서 1 kb 하류(0.0%), 및 유전자간(83.0%)이 발생한 통합 사건의 백분율이 제시된다. (도 28d) 마우스 게놈 창에서의 통합 패턴. 연속적 게놈 창 및 무작위화 마우스 게놈 창과 중첩된 통합의 수 및 크기를 비교하였다. 이는 통합 패턴이 연속 및 무작위 창에서 유사함을 나타낸다. 임의의 주어진 창에서 최대 통합 수는 3 이하였으며; 창 당 하나의 통합은 발생률이 더 높다.

도 29a-29i. 2차 수용자의 분석. 생체내 형질도입된 CD46tg 마우스로부터 20 주차에 수확된 골수 Lin- 세포를 치명적으로 조사된 C57Bl/6 마우스 내로 이식하였다. 2차 수용자를 16 주 동안 추적하였다. (도 29a) 이식 후 4, 8, 12, 및 16 주차에 CD46-양성 PBMC의 백분율에 기반한 생착률. 두 그룹 사이의 차이는 유의하지 않았다. (도 29b) 유세포 분석에 의해 측정된 γ-글로빈- 발현 말초 혈액 RBC의 백분율. 두 그룹 사이의 차이는 유의하지 않다. (도 29c) 생체내 HSPC 형질도입 후 20 주차에 수확된 골수 MNC에서 세포 당 벡터 카피 수. 두 그룹 사이의 차이는 유의하지 않다. (도 29d) 2차 수용자의 RBC에서 HPLC에 의한 인간 γ-글로빈 쇄의 분석. 성체 마우스 α-글로빈에 대한 인간 γ-글로빈의 백분율이 제시된다. *** p<0.0001. (도 29e) 마우스 α-글로빈 mRNA에 비해 전체 혈액 세포에서 γ-글로빈 mRNA 수준. (도 29f) 모든 골수 MNC에서 적혈구(Ter119+ 세포)를 발현하는 γ-글로빈의 백분율. 통계적 분석은 이원 ANOVA를 사용하여 수행하였다. (도 29g) p.t. 16 주차에 γ-글로빈 mRNA 수준 골수 MNC 마우스 α 및 β-주요 글로빈 mRNA에 대한 인간 γ-글로빈 m-RNA의 백분율이 제시된다. (도 29h) 적혈구 특이성. 적혈구($Ter119^{+}$) 및 비-적혈구($Ter119^{-}$) 세포에서 γ-글로빈+ 세포의 백분율. (도 29i) 생체내 HSPC 형질도입 후 20 주차에 수확된 골수 MNC에서 세포 당 벡터 카피 수(VCN). 두 그룹 사이의 차이는 유의하지 않다.

도 30a-30d. 이식 후 16 주차에 2차 수용자에서 혈액학적 매개변수. (도 30a) 백혈구. (도 30b) 적혈구생성 매개변수. RBC: 적혈구, Hb: 헤모글로빈, MCV: 평균 미립자 부피, MCH: 평균 미립자 헤모글로빈, MCHC: 평균 미립자 헤모글로빈 농도, RDW: 적혈구 분포 폭. (도 30c) 세포 골수 조성물. (도 30d) 골수 Lin- 세포의 콜로니-형성 가능성. 그룹 사이의 차이는 도 30a - 30d에서 유의하지 않았다. 통계적 분석은 이원 ANOVA를 사용하여 수행하였다.

도 31a-31d. 인간 CD34+ 세포를 사용한 시험관내 연구. (도 31a) 실험의 도식: CD34+ 세포를 HDAd-긴-LCR + HD-SB 또는 HDAd-짧은-LCR + HDAd-SB로 형질도입하고 적혈구 분화(ED)에 적용하였다. O6BG-BCNU를 사용한 시험관내 선택을 ED의 5 일차에 시작하였다. 18 일차에 세포를 유세포 분석(도 31b) 및 HPLC(도 31c)에 의해 분석하였다. (도 31d) 18 차에 벡터 카피 수. 통계적 분석은 이원 ANOVA를 사용하여 수행하였다. * p<0.05; ** p<0.0001

도 32a-32h. HDAd-짧은-LCR 및 HDAd-긴-LCR을 사용한 $Hbb^{th3}$/CD46 마우스의 생체내 HSC 유전자 요법 후 인간 γ -글로빈 발현. (도 32a) 처리 레지멘. 도 25a-25e와 대조적으로, 도 32a-32d는 지중해빈혈 $Hbb^{th3}$/CD46 마우스 내의 결과를 나타낸다. (도 32b) 유세포 분석에 의해 측정된 말초 적혈구(RBC)에서 인간 γ -글로빈-양성 세포의 백분율. 각 기호는 개별 동물이다. (도 32c) 생체내 HSPC 형질도입 후 18 주차에 RBC에서 HPLC에 의해 측정된 γ -글로빈 단백질 쇄 수준. 마우스 α -글로빈 단백질 쇄에 대한 인간 γ-글로빈의 백분율이 제시된다. (도 32d) 미처리 $Hbb^{th3}$/CD46 마우스(왼쪽 패널) 및 처리 후 21 주차에 마우스의 대표적인 크로마토그램. 마우스 α- 및 β-쇄 뿐만 아니라 추가된 인간 γ -글로빈이 표시된다.

도 32e - 32h. HDAd-짧은-LCR 및 HDAd-긴-LCR을 사용한 Hbbth3/CD46+/+ 마우스의 생체내 HSPC 유전자 요법 후 인간 γ-글로빈 발현. (도 32e) 처리 레지멘: 도 25에 제시된 연구와 대조적으로, 이 연구는 지중해빈혈 Hbbth3/CD46 마우스를 사용하여 수행하였다. (도 32f) 유세포 분석에 의해 측정된 말초 적혈구(RBC)에서 인간 γ-글로빈-양성 세포의 백분율. 각 기호는 개별 동물이다. (도 32g) 생체내 HSPC 형질도입 후 10 내지 16 주차에 RBC에서 HPLC에 의해 측정된 γ-글로빈 단백질 쇄 수준. 마우스 α-글로빈 단백질 쇄에 대한 인간 γ-글로빈의 백분율이 제시된다. (도 32h) 미처리 Hbbth3/CD46+/+ 마우스(왼쪽 패널) 및 처리 후 16 주차에 마우스의 대표적인 크로마토그램. 마우스 α- 및 β- 쇄 뿐만 아니라 추가된 인간 γ-글로빈이 표시된다. 특히, 2개의 독립적인 연구를 Hbbth3/CD46+/+ 마우스를 사용하여 수행하였다. 첫번째 연구: HD-긴-LCR의 경우 N=6 및 HDAd-짧은-LCR의 경우 N=2를 21 주 동안 추적하였다. 두번째 연구: HD-긴-LCR의 경우 N=4 및 HDAd-짧은-LCR의 경우 N=5를 16 주 동안 추적하였다. 도 32f는 21 주까지 조합된 데이터를 나타낸다. 통계적 분석은 이원 ANOVA를 사용하여 수행하였다. * p<0.05; ** p<0.0001

도 33a, 33b. 희생 시 골수의 분석. 골수를 Hbbth3/CD46+/+ 마우스를 사용한 생체내 HSPC 형질도입 후 16 주차에 수집하였다. (도 33a) 골수 MNC에서 세포 당 벡터 카피 수. 두 그룹 사이의 차이는 유의하지 않다. (도 33b) 적혈구(Ter119+) 세포에서 γ-글로빈의 평균 형광 강도(MFI). 통계적 분석은 이원 ANOVA를 사용하여 수행하였다.

도 34. 처리 전 및 긴 LCR로 처리 후 10 주차에, C57BL6(정상 마우스) 및 Townes SCA 마우스의 정규화된 적혈

구 형태학을 나타내는 현미경 사진.

도 35. 처리 전 Townes 마우스, 및 처리 후(긴 LCR) 10 주차에 Townes 마우스에 대한 정규화된 적혈구생성(망상적혈구 계수)을 나타내는 현미경 사진.

도 36a-36c. 표현형 교정. (도 36a, 36b) 왼쪽 패널은 Griemsa 염색으로 염색된 혈액 도말을 나타내고 오른쪽 패널은 May-Grunwald 염색으로 염색된 혈액 도말을 나타내는 혈액 세포 형태학. 망상적혈구에서 잔여 핵 및 세포질은 보라색으로 염색된다. (도 36a) 처리 전 및 14 주 차의 비교. (도 36b) CD46tg, 처리 전 $Hbb^{th3}$/CD46 마우스, 18 주차에 HDAd-긴-LCR로 처리된 $Hbb^{th3}$/CD46 마우스, 및 21 주차에 HDAd-긴-LCR로 처리된 $Hbb^{th3}$/CD46 마우스에 대한 Giemsa 염색 및 망상적혈구의 비교. (도 36c) 골수 사이토스핀. 처리된 전적혈구모세포 우세와 함께 적혈구생성의 bac k-이동이 보인다. 스케일 바는 20 μm이다.

도 37a, 37b. 표현형 교정(16 주). (도 37a) 왼쪽 패널: Giemsa/May-Grünwald 염색으로 염색된 혈액 도말(5 분). 오른쪽 패널: 망상적혈구에 대해 브릴리언트 크레실 블루로 염색된 혈액 도말. 망상적혈구에서 잔여 핵 및 세포질은 보라색 염색으로 보인다. (도 37b) Giemsa/May-Grünwald 염색으로 염색된 골수 사이토스핀(15 분). (도 37a 및 37b) 상부 패널: 정상 골수 세포 분포 - 적혈구 계통은 적혈구 분화의 모든 단계에 의해 나타난다. 중간 패널: 백혈구 계통보다 적혈구 계통의 우세 - 적혈구 계통은 주로 전적혈구모세포 및 호염기성 적혈구모세포로 이루어진다. 하단 패널: 정상 골수 세포 분포 - 적혈구 계통은 주로 성숙한 다염성 및 정염성 적혈구모세포에 의해 나타난다. 스케일 바는 25 μm이다.

도 38: 1 주차(상단 패널) 및 10 주차(하단 패널)에 긴 LCR 벡터, 짧은 LCR 벡터, 및 대조군 CD46tg의 정규화된 적혈구 매개변수에 대한 그래프 묘사가 제시된다.

도 39a, 39b. Hbbth3/CD46+/+ 마우스의 생체내 HSPC 유전자 요법 전 및 후의 혈액학적 매개변수(16 주). (도 39a) 망상적혈구 계수. (도 39b) 혈액학적 매개변수. 통계적 분석은 이원 ANOVA를 사용하여 수행하였다. * $p<0.05$; ** $p<0.0001$

도 40a, 40b. 비장 및 간에서 골수외 조혈의 표현형 교정. (도 40ai) 희생 시 비장 크기(16 주). 왼쪽 패널: 대표적인 비장 이미지. 오른쪽 패널: 요약. 각 기호는 개별 동물을 나타낸다. 통계적 분석은 일원 ANOVA를 사용하여 수행하였다. ** $p<0.0001$. 두 벡터 사이의 차이는 유의하다. (도 40b) 간 및 비장 절편에서 헤마톡실린/에오신 염색에 의한 골수외 혈구형성. Hbbth3/CD46+/+ 마우스에서 간의 적혈구모세포 및 비장에서 거핵구의 클러스터는 검은색 화살표로 표시된다. 스케일 바는 20 μm이다. 대표적인 이미지가 제시된다.

도 41. 비장 및 간에서 혈철소증의 표현형 교정(16 주). 철분 침착은 비장 및 간 절편에서 헤모시데린의 세포질 청색 색소로 Perl의 염색에 의해 제시된다. 스케일 바는 20 μm이다. 대표적인 절편이 제시된다. (Exp: 2.24 ms, 이득: 4.1x, 포화: 1.50, 감마: 0.60).

도 42a-42c. 희생 시 골수의 분석(21 주). 골수를 $Hbb^{th3}$/CD46tg 마우스의 생체내 HSC 형질도입 후 21 주차에 수확하였다. (도 42a) 골수 MNC에서 세포 당 벡터 카피 수. (도 42b, 42c) γ-글로빈 발현의 적혈구 특이성. (도 42b) γ-글로빈 발현 적혈구($Ter119^{+}$) 및 비-적혈구($Ter119^{-}$) 세포의 백분율. *$p<0.05$. 통계적 분석은 이원 ANOVA를 사용하여 수행하였다.

도 43. 아데노바이러스 공여자 벡터의 투여 전에 CD46tg 및 $CD46^{+/+}$/$Hbb^{th-3}$ 마우스의 간 및 비장 절편에서 헤마톡실린/에오신 염색에 의한 골수외 혈구형성. 철분 침착은 비장에서 헤모시데린의 세포질 청색 색소로서 Perl의 염색에 의해 제시된다.

도 44a-44e. 생체내 HSPC 형질도입/선택에 의한 CD46+/+/Hbbth-3 마우스의 표현형 교정. (도 44a) 건강한(CD46tg) 마우스, 동원 및 생체내 형질도입 전의 CD46+/+/Hbbth-3 마우스, 및 생체내 형질도입/선택을 겪은 CD46+/+/Hbbth-3 마우스(HDAd 주입 후 29 주차에 분석됨)의 RBC 분석($n = 5$). *$P \leq 0.05$, **$P \leq 0.0002$, ***$P \leq 0.00003$. 통계적 분석은 이원 ANOVA를 사용하여 수행하였다. (도 44b) 망상적혈구 검출을 위해 브릴리언트 크레실 블루를 사용한 말초 혈액 도말의 초생체 염색. 화살표는 특징적 잔여 RNA 및 미세-세포기관을 함유하는 망상적혈구를 나타낸다. 대표적인 도말에서 양성으로 염색된 망상적혈구의 백분율은 다음과 같았다: CD46의 경우, 7%; 처리 전 CD46+/+/Hbbth-3의 경우, 31%; 및 처리 후 CD46+/+/Hbbth-3의 경우, 12%. 스케일 바: 20 μm. (도 44c) 상단: 혈액 도말. 스케일 바: 20 μm. 중간: 골수 사이토스핀. 화살표는 상이한 성숙

단계에서 적혈구모세포 및 처리된 마우스에서 우세한 전적혈구모세포가 적혈구 생성에서 역행을 나타낸다. 스케일 바: 25 μm. 하단: Perl의 염색에 의한 조직 혈철소증. 철분 침착은 비장 조직 절편에서 헤모시데린의 세포질 청색 색소로서 제시된다. c 및 도 5d에서 대조군 마우스(CD46tg 및 CD46+/+/Hbbth-3, 형질도입 전)에 대한 혈액 도말 이미지는 동일한 샘플로부터 유래된다. (도 44d) 1 개의 대표적인 CD46tg 및 1 개의 미처리 CD46+/+/Hbbth-3 마우스 및 5 개의 처리된 CD46+/+/Hbbth-3 마우스의 거시적 비장 이미지. (도 44e) 희생 시, 비장 크기는 총 체중에 대한 비장 중량의 비(mg/g)로 결정하였다. 각 기호는 개별 동물을 나타낸다. 데이터는 평균 Å} SEM으로 제시된다. *P ≤ 0.05. 통계적 분석은 일원 ANOVA를 사용하여 수행하였다.

도 45. 생체내 형질도입 후 16 주차에 CD46 및 처리된 Hbbth3/CD46 마우스의 세포 골수 조성물. 그룹 사이의 차이는 유의하지 않았다. 통계적 분석은 이원 ANOVA를 사용하여 수행하였다.

도 46. 인간 γ-글로빈 게이팅 전략. CD46/Hbbth3 마우스로부터의 고정 및 투과된 RBC를 적혈구 마커 Ter-119 및 세포내 γ -글로빈에 대해 염색하였다.

도 47a, 47b. CD34+ 세포의 전사체(transcriptome)에 대한 SB100x-매개 통합의 효과. (도 47a) 실험 도식. CD34+ 세포를 EF1α 프로모터 단독 또는 HDAd-SB와의 조합으로 제어 하에 GFP/mgmt 카세트를 함유하는 HDAd5/35++ 벡터로 형질감염시켰다. 형질도입된 세포를 16 일 동안 적혈구 분화 배지에서 확장시켰다. 2 차례의 O6BG/BCNU 선택(50 μM O6BG + 35 μM BCNU)을 통합된 트랜스포존이 있는 GFP- 양성 세포에 대해 강화시켰다. 16 일차에, GFP-양성 세포를 FACS 분류하였다(샘플 #6). 비교를 위해(샘플 #5), mgmt/GFP 벡터 단독으로 형질도입되고 선택에 적용된 CD34+ 세포를 사용하였다. 대조군 세포는 SB100x를 발현하지 않았기 때문에, 에피솜 mgmt/GFP 벡터가 손실되었고 따라서 GFP 음성이었다. 두 샘플로부터의 총 RNA를 Omega Bioservices에 의해 수행된 RNA-Seq에 적용하였다. (도 47b) mRNA 발현이 변경된 유전자(log2 배수 변화)를 이들의 p 값에 기반하여 순위를 매겼다.

도 48. 생체내 형질도입 후 16 주차에 골수 MNC에서 mgmt mRNA 발현 수준. 인간 $mgmt^{P140K}$ 및 마우스 mRPL10 수준을 총 골수 MNC에서 qRT-PCR에 의해 측정하였다. (mRPL10은 마우스 하우스키핑 유전자이다). 상대적 수준을 VCN에 의해 추가로 나누었다(도 33 참조). 통계적 분석은 이원 ANOVA를 사용하여 수행하였다.

도 49. 마우스에서 벡터 hCD46tg의 생체내 HSC 형질도입: "긴" 대 "짧은" 벡터 LCR. 마우스에서 벡터 $Hbb^{th3}$/CD46의 생체내 형질도입. 그룹 1은 7 마리의 마우스에서 HDAd-긴-LCR-γ-글로빈/mgmt 및 HDAd-SB/Flpe의 생체내 형질도입을 나타낸다. 그룹 2는 3 마리의 마우스에서 HDAd-짧은-LCR γ-글로빈/mgmt 및 HDAd-SB/Flpe의 생체내 형질도입을 나타낸다. $O^6$BG, BCNU에 대해 3 회 선택 주기만이 필요하였다.

도 50. Thbb 마우스 테스트(W6). 그래픽 결과는 짧은 LCR 벡터에 비해 긴 LCR 벡터로 형질도입될 때 마우스 사이에 차이가 없고 인간 γ-글로빈 발현이 거의 없음을 나타낸다.

도 51. Thbb 마우스 테스트(W8). 그래픽 결과는 짧은 LCR 벡터에 비해 긴 LCR로 형질도입될 때 마우스 사이의 차이를 나타내지만, 짧은 LCR 바이러스가 마우스에서 사멸되었는지는 불분명하다.

도 52. 마우스에서 인간 γ-글로빈 발현 RBC의 백분율을 나타내는 그래픽 묘사. 그래프는 생체내 선택의 3 주기 후에만 100% 마킹을 예시한다.

도 53. 마우스 HBA에 대한 상대적 인간 γ-글로빈을 나타내는 HPLC의 그래픽 묘사(10 주). 그래프는 짧은 LCR과 비교하여 긴 LCR에 대한 상당히 더 높은 γ-글로빈 수준을 나타낸다.

도 54. 긴 LCR 벡터를 함유하는 마우스 #57의 예시적인 10 주 혈액 HPLC의 그래픽 묘사.

도 55a-55e. HDR-매개 통합을 위한 AAVS1-특이적 CRISPR/Cas9 벡터 및 공여자 벡터의 특성화. (도 55a) HDAd-CRISPR 벡터 구조: AAVS1-특이적 sgRNA는 U6 프로모터로부터 PolIII에 의해 전사되고 spCas9 유전자는 EF1α 프로모터의 제어 하에 있다. Cas9 발현은 HDAd 생산자 116 세포에서 Cas9 발현을 억제하지만 CD34+ 세포에서 Cas9 발현에 부정적으로 영향을 미치지 않는 miR-183-5p 및 miR-218-5p에 의해 제어된다(Sayadaminova 등, *Mol Ther Methods Clin Dev*, 1, 14057, 2015). 상응하는 micro RNA 표적 부위(miR-T)를 β-글로빈 유전자의 3' 비번역 영역(3'UTR) 내로 포매시켰다. (도 55b) 2000 vp/세포의 MOI에서 HDAd-CRISPR 형질도입 3 일 후 T7E1 검정에 의해 측정된 인간 CD34+ 세포에서 표적 부위 절단 빈도. 특이적 절단 생성물은 474 bp 및 294 bp이다. 절단 효능은 겔 아래에 제시된다. (도 55c) HDAd-CRISPR-형질도입된 CD34+ 세포에서 발견된 상위 13 개의 가장 빈번한 indel(서열번호: 6-18, 위에서 아래 순서). 연회색으로 강조된 서열은 중간 회색으로 강조하

는 표시된 TAM 서열이 있는 가이드 RNA의 표적을 나타낸다. CRISPR/Cas9 절단 부위는 수직 화살표로 표시된다. 녹색은 NHEJ에 의해 유발된 삽입이다. (도 55d) AAVS1 부위 내로 통합을 위한 공여자 벡터의 구조(HDAd-GFP-공여자). mgmtP140K 유전자는 자기-절단 피코르나바이러스 2A 펩티드를 통해 GFP 유전자에 연결되어 있다. 유전자는 EF1α 프로모터의 제어 하에 있다. PA: 폴리-아데닐화 신호. 이식유전자 카세트는 이전에 공개된 연구(Lombardo 등, *Nat Methods* 8, 861-869, 2011)와 유사한 AAVS1 유전자좌에 상동성인 0.8 kb 영역에 의해 플랭킹된다. 상동성 영역의 상류 및 하류는 공여자 카세트를 방출하는 AAVS1-특이적 CRISPR/Cas9에 대한 인식 부위이다. (도 55e) 공여자 카세트의 방출. CD34+ 세포를 HDAd-GFP-공여자(1000 또는 2000 vp/세포의 MOI) 단독으로 또는 HDAd-CRISPR(MOI 1000 vp/세포)과 조합하여 형질감염시켰다. 3 일 후에 게놈 DNA를 GFP-특이적 프로브와 함께 서던 블롯에 적용하였다. (선형) 전장 HDAd-공여자-GFP 게놈을 36 kb에서 실행한다. 방출된 카세트를 4.7 kb에서 실행한다. 절단 빈도는 겔 아래에 제시된다.

도 56a-56f. HUDEP-2 세포에서 표적화된 통합 vs. SB100x-매개 통합. (도 56a) 실험 도식. HUDEP-2 세포를 각 바이러스에 대해 1000 vp/세포의 MOI에서 표시된 HDAd 벡터로 형질감염시켰다. 21 일 동안 확장 후, GFP 양성 세포를 96 웰 플레이트 내로 분류하였다. 단일 세포-유래 클론을 2 주 동안 추가 확장에 의해 수득하였다. GFP 발현을 세포 집단에서 형질도입후 2 및 21 일차, 또는 세포 클론에서 35 일차에 측정하였다. (도 56b) 2 및 21 일차에 공여자 벡터 단독 또는 표적화 vs SB100x 통합 메커니즘이 있는 벡터로 처리된 세포에서 GFP 유세포 분석. (도 56c) 표적화 vs SB100x 통합이 있는 총 $GFP^+$ 세포에서 GFP의 평균 형광 강도(21 일). 제시된 데이터(평균 ± SD)는 3 개의 독립적 실험을 나타낸다. (도 56d) 단일 클론에서 GFP의 평균 형광 강도. 각 기호는 하나의 세포 클론을 나타낸다. 제시된 데이터(평균 ± SD)는 2 개의 독립적인 실험을 나타낸다. (도 56e) 표적화 또는 SB100x-매개 통합이 있는 대표적인 세포 클론에서 GFP 발현을 나타내는 유세포 분석. (도 56f) GFP 프라이머를 사용한 qPCR에 의한 세포 클론에서 벡터 카피 수.

도 57a, 57b. 표적화된 통합 벡터로 형질도입된 HUDEP-2 클론의 통합 분석. (도 57a) 역PCR에 의한 통합 부위 분석. 상부 다이어그램은 활용된 *NcoI* 부위, 및 프라이머의 위치를 나타낸다(반 화살표. 진회색: 5'- 접합부에 대한 EF1α 프라이머; 연회색: 3' 접합부에 대한 pA 프라이머). 표적화된 통합에 대한 각 측면에서 예상된 앰플리콘 크기가 표시된다. 하부 겔 사진은 iPCR 결과를 나타낸다. 각 레인은 하나의 세포 클론을 나타낸다. New England Biolab의 1kb 래더(ladder)를 사용하였다. Ef1α 프라이머가 채택된 이후에 내인성 Ef1α의 추가 밴드가 검출되었다. 클론 #20의 경우, 앰플리콘 크기가 예측과 다를지라도, 클로닝 및 서열분석은 표적 통합이 있는 클론임을 입증하였다. (도 57b) In-Out PCR 분석. 상부 다이어그램은 프라이머의 위치를 나타낸다. 다양한 통합 패턴에 대한 예상된 생성물 크기가 나열된다. 하부 겔 사진은 대부분의 클론이 단일대립유전자성 표적화된 통합이 있음을 입증한다. (도 57a)의 결과와 관련하여, 클론 #17, #20 및 #36의 예상치 못한 앰플리콘 크기는 연쇄체(concatemeric) 통합으로부터 발생할 가능성이 있었다.

도 58a-58c. AAVS1/CD46tg 마우스에서 AAVS1 표적 부위의 절단. (도 58a) 시험관내 분석. 표시된 MOI의 시험관내 HDAd-CRISPR 형질도입 3 일 후 측정된 AAVS1/CD46tg 마우스의 골수 계통-음성 세포에서 표적 부위 절단 빈도. (도 58b) 이식 후 14 주차에 총 골수 단핵 세포로부터 DNA의 심층 서열분석에 의해 수득된 총 AAVS1 indel의 백분율. 각 기호는 개별 동물이다. (도 58c) 마우스에서 발견된 상위 29 개의 가장 빈번한 indel(서열번호: 19-23, 21, 21, 26-30, 27, 32, 28, 34-47), 위에서 아래 순서). 대표적인 데이터가 제시된다. 황색 서열은 청색으로 표시된 TAM 서열이 있는 가이드 RNA의 표적을 나타낸다. CRISPR/Cas9 절단 부위는 수직 화살표로 표시된다.

도 59a-59d. HDAd-AAVS1 및 HDAd-GFP-공여자를 사용한 AAVS1/CD46 Lin- 세포의 생체외 형질도입 및 치명적으로 조사된 수용자 내로 후속 이식. (도 59a) 실험 도식: 골수를 AAVS1/CD46tg 마우스로부터 수확하고 계통-음성 세포(Lin-)를 MACS에 의해 단리하였다. Lin- 세포를 HDAd-CRISPR 및 HDAd-GFP-공여자 단독으로 또는 500 vp/세포의 총 MOI에서 조합하여 형질도입하였다. 배양 1 일 후, 마우스 당 $1x10^6$ 개의 형질도입된 세포를 치명적으로 조사된 C57Bl/6 마우스 내로 이식하였다. 4 주차에, $O^6$BG/BCNU 처리를 시작하고 2 주마다 3 회 반복하였다. 각 주기마다, BCNU 농도를 5 mg/kg에서 7.5 mg/kg 내지 10 mg/kg까지 증가시켰다. 14 주차에, 마우스를 희생시키고 골수 Lin- 세포를 치명적으로 조사된 2차 C57Bl/6 수용자 내로 이식하기 위해 사용한 다음, 16 주 동안 추적하였다. (도 59b) 유세포 분석에 의해 측정된 말초 혈액 단핵 세포(PBMC)에서 GFP-양성 세포의 백분율. HDAd-CRISPR 단독, HDAd-GFP-공여자 단독, 및 HDAd-CRISPR + HDAd-GFP-공여자로 형질도입된 Lin- 세포를 이식한 그룹이 제시된다. 각 기호는 개별 동물을 나타낸다. (도 59c) Lin- 세포가 이식된 대표적인 마우스로부터의 PBMC에서 GFP+ 세포의 백분율. 4 주(선택 전) 및 12 주(선택 후)의 데이터가 제시된다. (도 59d) 계

통-양성 세포 CD3+(T-세포), CD19+(B-세포), Gr-1+(골수성 세포), 및 HSC(LSK 세포)에서 GFP+ 세포의 백분율.

도 60a-60e. 생체외 형질도입된 Lin- 세포의 생착 분석. (도 60a) 유세포 분석에 의해 측정된 PBMC에 대한 인간 CD46 발현에 기반한 이식된 세포의 생착. 각 기호는 개별 동물이다. 특히, 형질도입된 공여자 세포는 CD46을 발현하였지만, 수용자 C57Bl/6 마우스는 발현하지 않았다. (도 60b) 14 주차에 PBMC(혈액), 비장, 및 골수에서 CD46-양성 세포의 백분율. (도 60c) 14 주차에 PBMC, 비장 및 골수에서 GFP-양성 세포의 백분율. (도 60d) 상이한 형질도입 설정에서 LSK 및 계통-양성 세포의 백분율. 3 개의 그룹 사이의 차이는 유의하지 않다. (도 60e) GFP+ 콜로니의 분석. 14 주차 마우스의 총 골수 Lin- 세포를 플레이팅하고 콜로니에서 GFP 발현을 12 일 후에 분석하였다. 각 기호는 개별 마우스에 대한 평균 GFP+ 콜로니 수이다(왼쪽 패널). 모든 콜로니의 세포를 풀링하고 유세포 분석에 의해 분석하였다(오른쪽 패널).

도 61a-61f. 2차 수용자에서 GFP 마킹 분석. HDAd-GFP-공여자 또는 HDAd-CRISPR + HDAd-GFP-공여자 형질도입된 Lin- 세포를 이식한 반응자 마우스의 골수 세포를 이식 후 14 주차에 수확하고, 계통-양성 세포에 대해 고갈시키고, 치명적으로 조사된 C57Bl/6 마우스 내로 이식하였다. (도 61a) 4 마리의 수용자 마우스에서 PBMC의 GFP-유세포 분석. 오른쪽 패널은 전형적인 분석을 나타낸다. 수직 축은 hCD46에 대한 염색을 나타내고, 수평 축은 GFP 염색을 나타낸다. (도 61b) 16 주차에 PBMC, 비장 및 골수에서 GFP-양성 세포의 백분율. (도 61c) 이식 후 16 주차에 수용자에서 계통-양성 및 -음성 세포의 GFP 흐름 분석. (도 61d) GFP+ 콜로니의 분석. 16 주차 마우스의 총 골수 Lin- 세포를 플레이팅하고 콜로니에서 GFP 발현을 12 일 후에 분석하였다. 각 기호는 개별 마우스에 대한 평균 GFP+ 콜로니 수이다(왼쪽 패널). 모든 콜로니의 세포를 풀링하고 유세포 분석에 의해 분석하였다(오른쪽 패널). (도 61e) 유세포 분석에 의해 측정된 PBMC에 대한 인간 CD46 발현에 기반한 이식된 세포의 생착. (도 61f) 상이한 형질도입 설정에서 계통-양성 및 -음성 세포의 백분율. 두 그룹 사이의 차이는 유의하지 않다.

도 62a-62f. HDAd-AAVS1-CRISPR + HDAd-GFP-공여자를 사용한 AAVS1/CD46tg 마우스의 생체내 형질도입. (도 62a) 처리 레지멘. AAVS1/hCD46tg 마우스를 동원하고 HDAd-CRISPR + HDAd-GFP-공여자를 IV 주사하였다(두 바이러스의1:1 혼합물의 각 4x1010 vp 2 회). 4 주 후에, O6BG/BCNU 처리를 시작하였다. 각 주기마다, BCNU 농도를 2.5 mg/kg에서 7.5 mg/kg 및 10 mg/kg까지 증가시켰다. O6BG 농도는 3 가지 처리 모두에서 30 mg/kg 이었다. 마우스를 동물이 분석을 위해 희생되고 Lin- 세포가 2차 수용자 내로 이식될 때 12 주까지 추적하였다. 그런 다음 2차 수용자를 16 주 동안 추적하였다. (도 62b) 유세포 분석에 의해 측정된 말초 혈액 단핵 세포(PBMC)에서 GFP-양성 세포의 백분율. (도 62c) 14 주차에 PBMC, 비장 및 골수에서 GFP-양성 세포의 백분율. (도 62d) 계통-양성 세포 CD3+(T-세포), CD19+(B-세포), Gr-1+(골수성 세포), 및 HSC(LSK 세포)에서 GFP+ 세포의 백분율. (도 62e) GFP+ 콜로니의 분석. 14 주차 마우스의 총 골수 Lin- 세포를 플레이팅하고 콜로니에서 GFP 발현을 12 일 후에 분석하였다. 각 기호는 개별 마우스에 대한 평균 GFP+ 콜로니 수이다(왼쪽 패널). 모든 콜로니의 세포를 풀링하고 유세포 분석에 의해 분석하였다(오른쪽 패널). (도 62f) 14 주차에 계통-양성 및 -음성 세포의 백분율.

도 63a-63e. 도 59a-59d의 2차 수용자 분석. 14 주차에, 생체내 형질도입된 AAVS1/hCD46tg 마우스의 골수 Lin- 세포를 치명적으로 조사된 C57Bl/6 수용자 내로 이식하였다. (도 63a) 6 마리의 수용자 마우스에서 PBMC의 GFP-유세포 분석. (도 63b) 혈액, 비장 및 골수의 단핵 세포에서 GFP 발현. (도 63c) 이식 후 16 주차에 수용자에서 계통-양성 및 -음성 세포의 GFP 흐름 분석. (도 63d) 유세포 분석에 의해 측정된 PBMC에 대한 인간 CD46 발현에 기반한 이식된 세포의 생착. (도 63f) 16 주차에 계통-양성 및 -음성 세포의 백분율.

도 64a-64h. HDAd-AAVS1 및 HDAd-공여자-γ-글로빈 벡터를 사용한 AAVS1/CD46 Lin- 세포의 생체외 형질도입 및 치명적으로 조사된 수용자 내로의 후속 이식. (도 64a) 공여자의 구조. 전반적인 구조는 HDAds-GFP-공여자 벡터에 대한 것과 동일하다(도 55d 참조). 상동성 영역은 새로운 HDAd-글로빈-공여자 벡터에서 더 길다(1.8 kb vs 0.8 kb). γ-글로빈 발현 카세트는 4 개의 DNAse 과민성(HS) 영역 및 γ-글로빈 프로모터를 포함하는 γ-글로빈 LCR의 4.3 kb 버전을 함유한다(Lisowski 등, *Blood.* 110, 4175-4178, 1996). 3' UTR(적혈구에서 mRNA 안정화를 위함)을 포함하는 전장 γ-글로빈 cDNA를 사용하였다. mgmtP140K 유전자는 보편적으로 활성인 EF1α 프로모터의 제어 하에 있다. 양방향성 SV40 폴리-아데닐화 신호를 사용하여 전사를 종결한다. LCR/β-프로모터 및 EF1α 프로모터 사이의 간섭을 피하기 위해, 1.2 kb 닭 HS4 염색질 절연체(Emery 등, *Proc Natl Acad Sci USA*, 97, 9150-9155, 2000)를 카세트 사이에 삽입하였다. (도 64b) 처리 레지멘은 도 57a에 제시된 것과 동일하다. (도 64c) 유세포 분석에 의해 측정된 말초 적혈구(RBC)에서 인간 γ-글로빈-양성 세포의 백분율. 체내 형질도입 후 16 주차에 혈액 및 골수에서 적혈구(Ter119+) 및 비-적혈구(Ter119-) 세포의 인간 γ-글로빈-양성 세포의 (도 64d) 백분율 및 (도 64e) 평균 형광 강도. * $p<0.05$. (도 64f) HPLC에 의해 16 주차에

RBC에서 측정된 마우스 β-주요 쇄에 대한 γ -글로빈 쇄의 백분율. (도 64g) qRT-PCR에 의해 16 주차에 RBC에서 측정된 마우스 β-주요 RNA에 대한 γ -글로빈 mRNA의 백분율. (도 64h) Lin- 세포로부터 유래된 콜로니에서 세포 당 벡터 카피 수. 각 기호는 하나의 콜로니를 나타낸다. 동물 사이의 차이는 유의하지 않다.

도 65a, 65b. HDAd-CRISPR 및 HDAd-글로빈-공여자 벡터로 형질도입된 AAVS1/CD46 Lin- 세포의 생착. (도 65a) 유세포 분석에 의해 측정된 PBMC에 대한 인간 CD46 발현에 기반한 이식된 세포의 생착. (도 65b) 16 주차에 계통-양성 PBMC(혈액), 비장, 및 골수 세포 뿐만 아니라 골수 LSK 세포에서 CD46-양성 세포의 백분율.

도 66a-66c. 도 64a-64h의 2차 수용자 분석. HDAd-CRISPR + HDAd-글로빈-공여자 형질도입된 Lin- 세포가 이식된 마우스의 골수 세포를 이식 후 16 주차에 수확하고, 계통-양성 세포에 대해 고갈시키고, 치명적으로 조사된 C57Bl/6 마우스 내로 이식하였다. (도 66a) 5 마리의 수용자 마우스에서 RBC의 γ-글로빈 유세포 분석. (도 66b) 계통-양성 PBMC에서 CD46-양성 세포의 백분율. (도 66c) 2차 수용자 내로 이식 후 16 주차에 골수 조성물.

도 67a-67h. HDAd-CRISPR + HDAd-글로빈-공여자를 사용한 AAVS1/CD46tg 마우스의 생체내 형질도입. (도 67a) 처리 레지멘. (도 67b) γ-글로빈-양성 RBC의 백분율. (도 67c) 형질도입되지 않은 대조군 마우스 또는 형질도입 후 16 주차 마우스의 말초 RBC에서 γ-글로빈 발현의 백분율을 나타내는 대표적인 도트 팟(dot pot). (도 67d) 혈액 및 골수의 적혈구(Ter119+) 및 비-적혈구(Ter119-) 세포에서 γ-글로빈의 평균 형광 강도. * $p<0.05$. (도 67e) HPLC에 의해 16 주차에 RBC에서 측정된 마우스 β-주요 쇄에 대한 γ-글로빈 쇄의 백분율. * $p<0.05$. (도 67f) qRT-PCR에 의해 16 주차에 RBC에서 측정된 마우스 β-주요 RNA에 대한 γ-글로빈 mRNA의 백분율. * $p<0.05$. (도 67g) 4 마리의 반응자 마우스에서 Lin- 세포로부터 유래된 콜로니에서 세포 당 벡터 카피 수. 각 기호는 하나의 콜로니를 나타낸다. 동물 사이의 차이는 유의하지 않다. (도 67h) 생체내 형질도입 후 16 주차에 혈액, 비장 및 골수에서 계통-양성 세포 및 골수에서 LSK 세포의 조성물.

도 68a-68d. 도 67a-67h의 2차 수용자 분석. (도 68a) 유세포 분석에 의해 측정된 PBMC에 대한 인간 CD46 발현에 기반한 이식된 세포의 생착. (도 68b) RBC에서 γ-글로빈 발현. (도 68c) HPLC에 의해 16 주차에 2차 수용자의 RBC에서 측정된 마우스 β-주요 쇄에 대한 γ-글로빈 쇄의 백분율. (도 68d) 생체내 형질도입 후 16 주차에 혈액, 비장 및 골수에서 계통-양성 세포 조성물.

도 69a, 69b. AAVS1/CD46 유전자이식 마우스에서 AAVS1 유전자좌의 국소화 및 구조. (도 69a) 염색체 14 상의 불일치를 나타내는 TLA 데이터. AAVS1-특이적 프라이머 쌍을 사용하였다. 오른쪽 패널은 18 kb 갭이 보이는 염색체 14의 확대된 절편을 나타낸다. 갭은 추가된 인간 AAVS1 유전자좌에 상응한다. (도 69b)

도 70. 게놈 국소화를 나타내는 AAVS1 유전자좌의 상세한 구조. 음영 처리된 AAVS1 영역을 Sanger 서열분석에 의해 확인하였다. 빈 영역은 The Jackson Laboratory의 제한 분석 및 AAVS1 tg 마우스 유전적 배경 정보로부터 제하였다. CRISPR/Cas9 절단 부위는 가위로 표시된다. 반복부 #2 내지 #5는 완전한 8.2 kb 인간 AAVS1 EcoRI 단편이지만, 반복부 #1 및 #5만이 EcoRI 단편의 분획만을 함유한다. 특히, 반복부 #5는 완전한 5' 상동성 아암이 결여되어 있다. AAVS1tg 마우스에 존재하는 다중카피 AAVS1 유전자좌의 CRISPR/Cas9 절단에 따른 결과. 절단 위치에 관한 규칙은 다음과 같다: a) 반복부 #1 내지 #4에서 하나의 단일 절단: 바람직. b) 반복부 #5에서 하나의 단일 절단: 불완전한 왼쪽 상동성 아암으로 인한 선호도 감소. c) 2 개의 반대로 배향된 반복부(예를 들어 #1 및 #4)에서 2 개의 절단: 오른쪽 상동성 아암의 누락으로 인해 HDR-매개 표적화된 통합 없음. d) 동일한 방향을 직면한 2 개의 반복부(예를 들어 #1 및 #2)에서 2 개의 절단: 바람직. e) 2 개 초과의 절단의 경우, 각 측면에서 마우스 gDNA 서열에 대한 하나의 근위부만 고려: 규칙 c) 또는 d)에 따라서 적용. f) 반복부 #1 및 #5에서 절단 및 중심 영역의 결실. 또한, HDR-매개 표적화된 통합은 반복부 #2 내지 #4에서 발생하였고, CRISPR에 의한 플랭킹 반복부, 예를 들어 #1 및 #5에서 연속적 절단은 이미 통합된 이식유전자의 손실을 초래할 수 있다.

도 71a, 71b. HDAd-CRISPR + HDAd-GFP-공여자를 사용한 생체외 또는 생체내 HSC 형질도입 후 16 주차에 단리된 게놈 DNA의 서던에 의한 통합 부위 분석. (도 71a) AAVS1-특이적 프로브와의 혼성화. 상부 패널은 예상된 EcoRI 단편 크기 및 프로브의 국소화를 나타낸다. 하부 패널은 생체외 및 생체내 형질도입 설정에서 개별 마우스의 분석을 나타낸다. 더 큰 밴드는 비-표적화된 AAVS1 유전자좌 반복부를 나타낸다. (도 71b) *BlpI*-소화된 DNA와 GFP-특이적 프로브의 혼성화. 밴드 패턴은 다른 곳에서 논의된다.

도 72a-72c. HDAd-CRISPR + HDAd-GFP-공여자를 사용한 생체외 또는 생체내 HSC 형질도입 후 16 주차에 단리된 게놈 DNA의 역PCR(iPCR)에 의한 통합 부위 분석. (도 72a) 다이어그램은 *NcoI* 부위, 및 프라이머의 위치를 나

타낸다(반 화살: 5' 접합부에 대한 EF1α 프라이머; 연회색: 3' 접합부에 대한 pA 프라이머). 반복부 #5에서 표적화된 통합을 위한 각 측면에서 예상된 앰플리콘 크기가 표시된다. (도 72b) 총 골수 세포에서 게놈 DNA를 사용한 iPCR 결과. 각 레인은 하나의 마우스를 나타낸다. #009, #023, #943, #944 및 #946은 생체외 HSC 형질도입 후 마우스이다. #147, #304 및 #467은 생체내 형질도입된 동물이다. (도 72c) GFP-양성 콜로니의 iPCR 분석. 14 주차 마우스의 골수 Lin- 세포를 플레이팅하고, 20 일 후에 게놈 DNA를 GFP+ 콜로니에서 단리하고 iPCR에 사용하였다. 마우스 #943 및 #946을 분석하였다. 각 레인은 하나의 콜로니를 나타낸다. 연회색 화살표: 표적화된 통합; 진회색 화살표: 표적외(off-target) 통합; 중간 회색 화살표: 통합된 전체 HDAd 바이러스 게놈.

도 73a, 73b. HDAd-CRISPR + HDAd-글로빈-공여자를 사용한 생체외 또는 생체내 HSC 형질도입 후 16 주차에 단리된 게놈 DNA의 역PCR(iPCR)의 통합 부위 분석. (도 73a) 다이어그램은 *NcoI* 부위, 및 프라이머의 위치를 나타낸다(반 화살표. 검정색 5' 접합부에 대한 EF1α 프라이머; 회색: 3' 접합부에 대한 pA 프라이머). 반복부 #5에서 표적화된 통합을 위한 각 측면에서 예상된 앰플리콘의 크기가 제시된다. (도 73b) 총 골수 세포에서 게놈 DNA를 사용한 iPCR 결과. 각 레인은 하나의 마우스를 나타낸다. #321, #322, #856, #857, #858 및 #945는 생체외 형질도입된 마우스이다. #504, #816 #869 및 #898은 생체내 형질도입된 동물이다. 흰색 화살촉은 표적화된 통합을 나타내고; 회색, 점선 화살촉: 표적외 통합; 흰색 전체 화살표: 통합된 전체 HDAd 바이러스 게놈.

도 74a-74d. (도 74a) 생체내 HSPC 형질도입을 위한 HDAd5/35++ 벡터. HDAd-GFP/mgmt에서, 트랜스포존은 HDAd-SB 벡터에서 제공된 과활성 잠자는 미녀 트랜스포사제(SB100X)를 통해 통합을 위해 도립된 트랜스포존 반복부(IR) 및 frt 부위에 의해 플랭킹된다. 이식유전자 카세트는 β-글로빈 3'UTR에 연결된 PGK-프로모터 구동된 GFP 유전자 뿐만 아니라 EF1α-프로모터 구동된 mgmtP140K 카세트를 함유한다. 두 카세트는 닭 글로빈 HS4 절연체에 의해 분리된다. HSPC를 인간 재조합 G-CSF의 s.c. 주사(5 μg/마우스/일, 4 일) 이어서 마지막 G-CSF 주사 18 시간 후 AMD3100의 s.c. 주사(5 mg/kg)에 의해 neu/CD46 유전자이식 마우스에서 동원하였다. HDAd-GFP/mgmt+HDAd-SB의 총 8x1010 개의 바이러스 입자를 AMD3100 주사 1 시간 후 i.v. 주사하였다. HDAd 주사 후 전염증성 사이토카인 방출을 방지하기 위해, 동물에게 바이러스 주사 16 시간 및 2 시간 전에 덱사메타손(10 mg/kg)을 i.p.로 제공하였다. 6 주 후, 3 차례의 O6BG/BCNU(i.p.)를 적용하여 형질도입된 HSPC가 말초 혈액 순환으로 빠져나가는 것을 활성화시켰다(30 mg/kg O6BG 및 5, 7.5, 및 10 mg/kg BCNU). 생체내 형질도입 17 주 후, $1x10^6$ 개의 MMC 세포를 유선 지방 패드 내로 이식하였다. 5 주 후, 종양 및 다른 조직을 수확하고 GFP 발현에 대해 분석하였다. (도 74b) 왼쪽 패널: 생체내 형질도입 후 상이한 시점에서 GFP-발현 PBMC의 백분율. 각 기호는 개별 동물을 나타낸다. 오른쪽 패널: 골수, 비장, 혈액, 및 콜라게나제/디스파제-소화된 종양에서 범-백혈구 마커 CD45에 대해 염색된 세포에서 GFP+ 세포의 백분율. (도 74c) GFP에 대한 항체 및 세포외 기질 단백질인 라미닌에 대한 항체로 염색된 종양 절편. 스케일 바는 50 μm이다. (도 74d) 혈액에서 GFP+ PBMC 및 종양에서 GFP+ 세포의 면역표현형.

도 75. MMC 세포에서 래트 Neu 발현. 세포를 Neu-특이적 단클론 항체 7.16.4 이어서 항-마우스 Ig-FITC로 염색하였다. 배양된 MMC 세포의 대표적인 공초점 현미경 이미지가 제시된다. 새로운-특이적 신호는 흰색 색조로 보인다. 스케일 바는 20 μm이다.

도 76. 면역표현형을 위한 게이팅 전략.

도 77. 골수 및 비장에서 GFP+ 세포의 면역표현형(MMC 모델). 상세한 내용은 도 74d를 참조한다.

도 78a-78f. 생체내 HSPC 형질도입 후 종양-침윤 백혈구에서 GFP 발현(TC-1 모델). (도 78a) 실험 도식. HSPC를 인간 재조합 G-CSF의 s.c. 주사(5 mg/마우스/일, 4 일) 이어서 마지막 G-CSF 주사 18 시간 후 AMD3100의 s.c. 주사(5 mg/kg)에 의해 CD46tg 유전자이식 마우스에서 동원하였다. HDAd-GFP/mgmt+HDAd-SB의 총 8x1010 개의 바이러스 입자를 AMD3100 1 시간 후 i.v. 주사하였다. HDAd 주사 후 전염증성 사이토카인 방출을 방지하기 위해, 동물에게 바이러스 주사 16 시간 및 2 시간 전에 덱사메타손(10 mg/kg)을 i.p.로 제공하였다. 6 주 후, 3 차례의 $O^6$BG/BCNU(i.p.)를 적용하여 형질도입된 HSPC가 말초 혈액 순환으로 빠져나가는 것을 활성화시켰다(30 mg/kg O6BG 및 5, 7.5, 및 10 mg/kg BCNU. 생체내 형질도입 17 주 후, $5x10^4$ 개의 TC-1 세포를 유선 지방 패드 내로 이식하였다. 5 주 후, 종양 및 다른 조직을 수확하고 GFP 발현에 대해 분석하였다. (도 78b) 생체내 형질도입 후 상이한 시점에서 GFP-발현 PBMC의 백분율. 각 기호는 개별 동물을 나타낸다. (도 78c) 골수, 비장, 혈액, 및 콜라게나제/디스파제-소화된 종양에서 범백혈구 마커 CD45에 대해 염색된 세포에서 $GFP^+$ 세포의 백분율. (도 78d) 전체(악성 + 종양 침윤) 세포에서 GFP+ 세포 및 $GFP^+$ 양성 백혈구의 대표적인

유세포 분석 데이터. (도 78e). 대표적인 종양 절편. 왼쪽 패널: GFP 형광. 오른쪽 패널: GFP에 대한 항체(흰색) 및 세포외 기질 단백질 라미닌(회색)으로 염색. 스케일 바는 50 mm이다. (도 78f) 종양에서 GFP+ 세포 및 혈액에서 PBMC의 면역표현형. BD Biosciences의 림프구 유세포 분석 패널 8c(CD45, CD3, CD4, CD8, CD25, CD19) 및 골수성 패널 9c(CD45, CD11c, F4/80, MHCII, SiglecF-PecCP, Ly6C, CD11b, Ly6G)를 사용하였다.

도 79a-79c. 종양-침윤 백혈구 이외의 세포에서 억제하기 위한 miRNA 선택. (도 79a) 이식유전자 발현의 조직-특이성의 miRNA-기반 조절. miRNA는 전형적으로 천연 mRNA의 3' 비번역 영역(3' UTR)에 상주하는 miRNA 표적 부위(miR-T)로 지칭되는 표적 서열과의 염기 쌍형성을 통해 가이드 분자로서 기능한다. 이 상호작용은 mRNA 절단 또는 번역 억제를 매개하는 효과기 복합체를 모집한다. 이식유전자의 mRNA가 주어진 세포 유형에서 높은 수준으로 발현되는 miRNA에 대한 miR-T를 함유하는 경우, 이식유전자 발현은 이 세포 유형에서 방지될 것이다. 대조적으로, 특이적 miRNA를 발현하지 않는 세포 유형에서, 이식유전자가 발현될 것이다(Brown 등, Nat Med. 2006;12: 585- 591). (도 79b) MicroRNA-Seq를 5 마리의 마우스(neu/CD46tg-MMC 모델, 종양 접종 후 17 일)에서 풀링된 RNA에 대해 수행하였다. $GFP^+$ 종양 13 개의 샘플에 대한 비장, 골수 및 혈액의 작은 RNA 서열분석에 의해 식별된 정규화된 microRNA 판독물 계수(100만 개의 맵핑된 microRNA 당 판독물 + 1)가 제시된다. miR-423을 포함하는, 종양에 존재하지 않는 MicroRNA를 유사-계수 1로 산점도의 왼쪽에 정렬한다. miR-423-5p는 블롯에 표시된다. (도 79c) MicroRNA-Seq를 5 마리의 마우스(CD46tg/TC-1 모델, 17 일)에서 풀링된 RNA에 대해 수행하였다. 상위 10 개 miRNA의 상대적 발현 수준을 종양의 수준(1로 설정)과 비교하였다.

도 80a-80c. HSPC에 대한 miR-423-5p 표적 부위 과발현의 효과. (도 80a) 벡터 구조. HDAd-GFP-miR-423은 GFP 유전자에 연결된 3'UTR에서 4 개의 miR-423-5p 표적 부위를 함유한다. (도 80b) 마우스 HSPC(M)(CD46-유전자이식 마우스의 골수로부터의 Lin- 세포) 및 인간 HSPC(Hu)($CD34^+$ 세포)를 HDAd-GFP 또는 HDAd-GFP-miR423으로 각각 500 또는 3000 vp/세포의 MOI에서 감염시켰다. 3 일 후, 세포 용해물을 CDKN1A에 대한 웨스턴 블롯에 의해 분석하였다. 블롯을 항-β-액틴 항체로 재프로빙하여 로딩 차이를 조정하였다. 오른쪽 패널은 b-액틴 신호에 대해 정규화된 CDKN1A 신호의 정량화를 나타낸다. 상응하는 마우스 및 인간 HDAd-GFP/mgmt 샘플의 신호를 100%로 취하였다. (도 80c) 전구체 콜로니 형성에 대한 효과. HDAd 감염 1 일 후, 마우스 $Lin^-$ 세포(35 mm 접시 당 $2.5x10^3$ 개 세포) 또는 인간 $CD34^+$ 세포(접시 당 $3x10^3$ 개 세포)를 콜로니 검정을 위해 플레이팅하였다. 콜로니를 12 일 후 계수하였다. N=3. *$p<0.05$. 통계적 유의성을 양측 스튜던트(Student) t-검정(Microsoft Excel)에 의해 계산하였다. (이전 연구와 일치하여(Li 등, Mol Ther Methods Clin Dev. 2018;9: 390-401; Li 등, *Mol Ther Methods Clin Dev.* 9: 142-152, 2018), 상대적으로 높은 MOI에서 HSPC의 감염은 HSPC의 콜로니 형성 용량을 약간 감소시켰다)

도 81. 노던 블롯에 의해 miR-423-5p 발현 검증. 골수 계통-음성 세포, 비장, 전체 혈액 세포, 및 MMC-/TC-1-종양 침윤 백혈구의 총 RNA(2 μg)를 15% 변성 폴리아크릴아미드 겔에서 분리하고 블롯을 muRNA-423-5p에 특이적인 프로브 및 후속적으로 U6 RNA(로딩 대조군)에 대한 프로브와 혼성화하였다. Mir-423은 70 bp의 전구체 길이 및 23 bp의 성숙 miRNA 길이를 갖는다. miR-423-5p-특이적 신호는 혈액, 골수, 및 비장에서 보이지만, 두 종양 모델에서 종양-침윤 세포에는 없다.

도 82a, 82b. 인간에서 miRNA423-5p 발현. (도 82a) Ludwig 등, Nucleic Acids Res. 2016;44: 3865-3877에 공개된 miR-423-5p의 수준. 왼쪽에서 오른족으로, y-축 라벨은 다음을 포함한다: 지방세포, 동맥, 결장, 경막, 신장, 간, 폐, 근육, 심근, 피부, 비장, 위, 고환, 갑상선, 소장 십이지장, 소장 공장, 췌장, 신장 부신, 신피질, 신수질, 식도, 전립선, 골수, 정맥, 림프절, 흉막, 뇌하수체, 척수, 뇌시상, 뇌백질, 뇌꼬리핵, 뇌회백질, 대뇌피질 측두엽, 대뇌피질 전두엽, 대뇌피질 후두부, 및 소뇌. (도 82b) 2 명의 난소암 환자에서 플롯팅된 miRNA-Seq 데이터(풀링). CD45+ 세포를 고등급 장액성 난소의 생검으로부터 단리하였다. RNA를 종양-침윤 백혈구 및 일치하는 PBMC로부터 단리하고 LC Sciences, LLC의 miRNA-Seq에 적용하였다. miRNA-423-5p가 표시된다.

도 83a-83e. neu/MMC 모델에서 생체내 HSPC αPD-L1-ɣ1 면역-체크포인트 억제제 요법. (도 83a) MMC 종양 세포에서 PDL1 발현(흰색). 스케일 바는 20 μm이다. (도 83b) 요법 벡터의 전반적인 구조는 도 74a에 제시된 것과 동일하다. 벡터는 5' 단부 상의 HA 태그 및 분비 신호(LS) 및 3' 단부 상의 인간 IgG1의 힌지-CH2-CH3 도메인 및 myc 태그에 연결된 scFv 항-마우스 PD-L1에 대한 발현 카세트를 함유한다. miR423-5p 표적 부위를 3'UTR 내로 삽입하여 miR423-5p 조절에 의해 αPD-L1-ɣ1 발현을 종양-침윤 세포로 제한하였다. 벡터는 또한 $mgtm^{P140K}$에 대한 발현 카세트를 함유한다. (도 83c) HDAd-GFP/mgmt 및 HDAd-αPD-L1-ɣ1 생체내 형질도입된

HSPC가 있는 마우스에서 MMC 세포 접종 후(0 일) 종양 부피. HDAd-αPD-L1-γ1 그룹의 마우스에게 첫번째 종양 세포 주사 후 80 일차에 $1x10^5$ 개 MMC 세포의 피하 주사를 재접종하였다. 각 곡선은 개별 동물이다. (도 83d) 유세포 분석에 의한 T-세포 반응 분석. 미처리 neu-유전자이식 마우스 및 HDAd-αPD-L1-γ1-처리된 마우스(100 일)의 비장세포를 CD4, CD8, 및 세포내 IFNγ에 대해 유세포 분석에 의해 분석하거나 또는 Neu 사량체로 염색하였다. N=3. *p<0.05. (도 83e) Neu+ 및 Neucell로 자극 시 IFNγ 반응. 미처리 neu-유전자이식 마우스 및 HDAd-αPDL1-γ1-처리된 마우스(100 일)의 비장세포를 저지된 MMC 세포 (Neu+) 또는 neu유전자이식 마우스(Neu-)의 비장세포에 노출시키거나, 또는 PMA/이오노마이신("noAg")으로 처리하였다. 배양 상청액에서의 IFNγ 농도가 제시된다. N=3. * p<0.005.

도 84a-84c. αPD-L1-γ1 발현의 동역학. (도 84a) 항-HA 태그 항체를 사용한 αPD-L1-γ1 웨스턴 블롯. 3 마리의 동물을 17 일차에 희생시키고 조직을 웨스턴 블롯에 의해 αPD-L1-γ1 발현에 대해 분석하였다. αPD-L1-γ1 단백질은 완전히 환원되지 않아서, 2 개의 scFv 쇄(130 kDa)가 있는 완전한 αPD-L1-γ1 잔류물을 생성하였다(αPD-L1-γ1의 구조의 경우 오른쪽 패널 참조). β-액틴에 대한 염색은 로딩 대조군으로 사용하였다. 대표적인 샘플이 제시된다. 또한 웨스턴 블롯 신호의 정량화가 제시된다. N=5 마우스. (도 84b) 종양-침윤 백혈구, PBMC, 골수 세포 및 비장세포에서 αPD-L1-γ1 mRNA 발현. 마우스 PPIA mRNA를 내부 대조군으로 사용하였다. 결과를 2(-ΔΔCt) 방법에 따라 계산하고 상응하는 종양 샘플의 cDNA 수준을 100%로 설정하여 상대적 발현의 백분율로 제시하였다. (도 84c) 포획을 위한 재조합 마우스 PD-L1 및 검출을 위한 항-HA 항체-HRP 접합체를 사용하여 ELISA에 의해 측정된 혈청에서 분비된 αPD-L1-γ1의 수준. 각 기호는 개별 동물을 나타낸다. * p<0.05. 통계적 유의성은 양측 스튜던트 t-검정(Microsoft Excel)에 의해 계산하였다.

도 85a - 85f. ID8-$p53^{-/-}$$brca2^{-/-}$ 난소암 모델에서 면역-예방 연구. (도 85a) ID8-$p53^{-/-}$$brca2^{-/-}$ 종양의 분석. 총 $2x10^6$ 개의 ID8-$p53^{-/-}$$brca2^{-/-}$ 세포를 CD46-유전자이식 마우스에 복강내로 주사하였다. 복수/악액질이 6-8 주 후에 발생하였다. 그런 다음 종양을 제거하고 유세포 분석을 위해 디스파제/콜라게나제로 소화시켰다. 세포의 분획을 노던 블롯 분석을 위해 종양-연관 대식세포(TAM), 호중구(TAN), 및 T-세포(TIL)로 분류하였다. (도 76 참조). (도 85b) 종양-연관 백혈구의 면역표현형. (도 85c) miR-423-5p에 대한 노던 블롯. 레인 당 총 1 μg의 RNA를 로딩하였다. 상부 패널은 $^{32}$P-표지된 miR-423-5p 프로브로 프로빙 후 신호를 나타낸다. 블롯을 스트립(stripped)하고 U6 RNA 특이적 프로브로 재프로빙하였다(하부 패널). Ambion의 $^{32}$P-표지된 Decade 마커를 오른쪽 레인에서 실행하였다. (도 85d) 실험 체계. CD46-유전자이식 마우스를 동원하고 HDAd-αPDL1γ1miR423 + HDAd-SB, HDAd-GFP-miR423 + HDAd-SB를 주사하거나, 또는 모의-주사하였다. 4 차례의 $O^6$BG/BCNU 생체내 선택이 주어졌다. ID8-$p53^{-/-}$$brca2^{-/-}$ 세포를 마지막 $O^6$BG/BCNU 처리 2 주 후 복강내로 주사하였다. 종양 세포 주사 2, 6, 및 11 주 후, αPDL1γ1 수준을 혈청에서 분석하였다. 복수 또는 이환율/악액질의 발병을 종점으로 취하였다. (도 85e) Kaplan-Meier 생존 플롯. N=7. (도 85f) ELISA에 의해 측정된 혈청 αPDL1γ1 수준. 각 기호는 개별 동물이다. * p<0.05. 통계적 유의성은 양측 스튜던트 t-검정(Microsoft Excel)에 의해 계산하였다

도 86a- 86d. ID8-$p53^{-/-}$$brca2^{-/-}$ 난소암 모델에서 면역-요법 연구. (도 86a) 암 재발을 예방하기 위한 임상 설정. 생체내 HSC 형질도입은 외과적 종양 감량수술 후 또는, 수술이 옵션이 아닌 경우 화학요법과 함께 시작할 것이다. $O^6$BG/BCNU 생체내 선택은 화학요법과 조합될 수 있다. 생체내 HSPC 형질도입/선택의 결과로서, 보강된 HSPC는 암이 재발할 때까지 휴면기에 놓여 HSPC 분화 및 효과기 유전자 발현의 활성화를 촉발할 것이다. (도 86b) 실험 체계. CD46 유전자이식 마우스에게 $1x10^6$ 개의 ID8-$p53^{-/-}$$brca2^{-/-}$ 종양 세포를 복강내로 주사하였다. 일단 종양이 확립되면, 생체내 HSPC 형질도입 및 선택을 수행하였다. miR-423-기반 발현 시스템의 활성화를 혈청 αPDL1γ1 수준에 기반하여 모니터링하였다. (도 86c) Kaplan-Meier 생존 플롯. 대조군 설정에서, HDAd-GFP-miR423을 주사하였다. N=9. (도 86d) 혈청 αPDL1γ1 수준을 ELISA에 의해 측정하였다. 각 기호는 개별 동물이다. * p<0.05. 통계적 유의성은 양측 스튜던트 t-검정(Microsoft Excel)에 의해 계산하였다.

도 87a, 87b. 종양 성장의 역전 전에, αPD-L1-γ1의 피크에서 17 일차에 희생된 동물의 자가면역 반응. (도 87a) 처리 전 동물(왼쪽 패널)과 비교하여 처리된 동물(오른쪽 패널)에서 털 변색. (도 87b) 처리된 동물의 기관에 대한 조직학적 분석. 절편을 H&E로 염색하였다. 대표적인 영역이 제시된다. 스케일 바는 20 mm이다. 단핵 세포의 침윤에 주목한다.

도 88a-88h. MMC 종양이 있는 neu-유전자이식 마우스에서 항-PD-L1 단클론 항체 요법의 효과 및 혈구형성에 대

한 생체내 HSC 형질도입의 효과. 종양이 100 $mm^3$의 부피에 도달했을 때, 마우스에게 항-마우스 PD1-L1 단클론 항체 muDX400*(5 mg/kg i.p.)(4 일마다 4 회) 또는 이소형 대조군 항체를 복강내 주사하였다. (도 88a) 개별 마우스의 종양 부피가 제시된다. (도 88b) 항-PD-L1에 대한 더 긴 생존을 나타내는 Kaplan-Meier 생존 플롯. 부피가 1000 $mm^3$인 종양을 종점으로 취하였다. 두 그룹 사이의 차이는 유의하지 않다. (도 88c) 생체내 HSCPC 형질도입 후 2 주차에 도 85d에 제시된 hCD46-유전자이식 마우스의 혈액 세포 계수 (도 85a) 혈액학적 매개변수. RBC: 적혈구, Hb: 헤모글로빈, MCV: 평균 미립자 부피, MCH: 평균 미립자 헤모글로빈, MCHC: 평균 미립자 헤모글로빈 농도, RDW: 적혈구 분포 폭. 통계적 분석은 이원 ANOVA를 사용하여 수행하였다. 세 그룹 사이의 차이는 유의하지 않았다. (도 88e) GFP+ 세포 분획의 niRNA-Seq. (도 88f) 웨스턴 블롯, qRT-PCR, 및 혈청 ELISA에 의한 αPDL1 발현의 동역학. (도 88g) miRNA-조절된 유전자 발현. (도 88h) 개시된 면역-예방적 및 암 재발 예방의 요약된 도식.

도 89a-89h. 적혈구의 GFP 발현과 관련된 데이터.

도 90a-90i. 적혈구의 인간 인자 VIII 발현과 관련된 데이터.

도 91a-91d. 혈액학적 이상은 관찰되지 않는다.

도 92a-92g. 억제제 항체에도 불구하고 A형 혈우병의 표현형 교정.

도 93a-93e. 원숭이(엠. 파시쿨라리스(*M. fascicularis*))의 생체내 형질도입. (도 93a) 실험 타임라인; (도 93b - 93d) 말초 혈액에서 동원된 CD34+ 세포의 GFP 마킹; (도 93e) 골수(3 일).

도 94a-94m. 조합된 생체내 HSC 형질도입 선택. $mgmt^{P140K}$는 약물 내성 및 유전자-변형된 세포의 선택적 확장에 대한 메커니즘을 제공한다. (인간 O(6)-메틸구아닌-DNA-메틸트랜스퍼라제(MGMT)에 대한 P140K 돌연변이체는 벤질구아닌으로도 알려진 MGMT 억제제 O(6)-(4-브로모테닐) 구아닌(O6BG)에 대한 내성을 부여한다. (도 94a) $MGMT^{p140k}$에 대한 벡터. (도 94b) 주사에 대한 타임라인 및 투여량을 나타내는 실험 설계. (도 94c) PBMC에서 GFP+ 세포의 퍼센트를 나타내는 데이터. (도 94d) 26 주차에 골수에서 GFP+ 세포의 퍼센트를 나타내는 데이터. (도 94e) Ad5/35-GFP 벡터. (도 94f) 4 일의 동원 이어서 Ad5/35 주사를 받은 돼지꼬리 원숭이를 묘사하는 실험 프로토콜. (도 94g) 동물 ID 및 G-CSF, SCF, AMD3100, 및 Ad5/35-GFP의 용량. (도 94h) AMD3100은 총 CD34+ 줄기 세포 수준을 G-CSF/SCF 단독보다 3-배 및 기준선보다 65-배 더 우수하게 증가시켰으며; 왼쪽 패널 말초 혈액에서 CD34+ 줄기 세포의 백분율을 나타내었고; 오른쪽 패널은 CD34+ 세포 계수를 나타내었다. (도 94i) AD5/35 주사 후 동원된 세포는 계통 왜곡 없이 건강한 콜로니를 형성하며; 왼쪽 패널은 Ad5/35 주사 후 0 시간에서 6 시간까지 콜로니의 빈도 및 수를 나타내는 수치 데이터를 제공하고; 오른쪽 패널은 CD34+ 세포의 형태학에 대한 육안 검사를 제공한다. (도 94j) 상단 패널은 주사 후 0 시간에서 6 시간까지 Ad5/35-GFP 세포의 유세포 분석 데이터를 제시한다. 하단 패널은 주사 후 0, 2, 및 6 시간에 Ad5/35-GFP를 함유하는 콜로니 수의 수치 데이터를 제시한다. (도 94k) 3% 초과의 말초 CD34+ 세포는 Ad5/35 주사 후 GFP를 발현한다. 상단 패널은 Ad5/35 주사 후 0 일에서 8 일까지 단핵 세포(MNC) 층에서 추출된 C34+ 세포를 도시한다. 하단 패널은 주사 후 2 및 6 시간에 평균 $GFP^+$ 발현을 도시한다. (도 941) 다중 방법은 동원 및 Ad5/35 주사 후 순환 세포의 성공적인 형질도입을 확인한다. 왼쪽 패널은 벡터 DNA의 Taqman 검출을 도시한다. 오른쪽 패널은 GFP 발현의 유세포 분석 데이터를 도시한다. (도 94m) 골수로 돌아가는 변형된 세포. 왼쪽 패널은 Ad5/35 주사 후 3, 7, 및 73 일차에 CD34+ 및 GFP+ 세포에서의 변화를 나타내는 유세포 분석 데이터를 도시한다. 오른쪽 패널은 Ad5/35 주사 후 기준선, 및 3, 7, 및 73 일차에 GFP+, CD34+ 세포의 퍼센트를 도시한다.

도 95. 본원에 기재된 대표적인 Ad35 헬퍼 바이러스 및 벡터의 특징. 5-점 별은 다음 텍스트를 나타낸다: -SB100x에 대한 조합(추가 및 재활성화) 및 표적화됨; -CRISPR 또는 BE에 대한 다중 sgRNA; -Cas9의 miRNA(miR187/218) 조절된 발현; 및 -Cas9의 자동-불활성화.

도 96. HDAd-TI-combo 벡터의 도식. CRISPR 시스템은 2 개의 상이한 부위(HBG 프로모터 및 적혈구 bcl11a 인핸서)를 표적화하여, 증가된 감마 재활성화를 야기한다.

도 97a-97d. (도 97a). HDAd-SB 및 HDAd-combo의 공동감염 시, Flpe가 발현되고 IR-플랭킹된 트랜스포존을 방출한 다음, SB100x 트랜스포사제에 의해 게놈 내로 통합될 것이다. 동시에, HBG1 및 bcl11a-E CRISPR이 발현되고 DNA indel을 생성하여 γ-글로빈의 재활성화를 야기할 것이다. 트랜스포존의 Flp-매개 방출 시, CRISPR 카세트는 분해되어 세포독성을 회피할 것이다. CRISPR 시스템은 2 개의 상이한 부위(HBG 프로모터 및 적혈구

bcl11a 인핸서)를 표적화하여, 증가된 γ 재활성화를 야기한다. (도 97b) 표적화 전략; (도 97c) 적혈구 특이적 BCL11A 인핸서; (도 97d) HBG 프로모터에서 BCL11A 결합 부위(서열번호: 48). HDAd-SB 및 HdAd-comb-SB의 도식은 도 102에서 찾을 수 있다.

도 98a-98n. 이중 CRISPR 벡터 및 γ-글로빈 재활성화. (도 98a) HDAd-Bcl11ae-CRISPR, HDad-HBG-CRISPR, HDAd-Dual-CRISPR, 및 HDAd-스크램블에 대한 벡터 설계. (도 98b) 이중 gRNA 벡터에 대한 HD-Ad5/35++ CRISPR 벡터. (도 98c) 인간 적혈구 전구체 세포주(HUDEP-2)의 HD-Ad5/35++ CRISPR 형질도입은 분화 전 및 후에 제시된다. 타임라인은 HUDEP-2 세포 이미지 아래에 제시된다. (도 98d) HD-AD5/35++ "Dual" gRNA 벡터는 미처리(UNTR), BCL11A, 또는 HBG 벡터와 비교하여 세포 생존력에 부정적인 영향을 미치지 않는다. (도 98e) HD-AD5/35++ "Dual" gRNA 벡터는 UNTR, BCL11A, 또는 HBG 벡터와 비교하여 증식에 부정적인 영향을 미치지 않는다. (도 98f, 도 98g) 이중 벡터는 표적 유전자좌 (도 98f) Bcl11a 인핸서 및 (도 98g) HBG 프로모터에 대한 단일 gRNA 벡터로 관찰된 것과 유사한 유사한 편집 수준을 달성한다. (도 98h) HD-AD5/35++ "Dual" gRNA 벡터는 단일 gRNA 벡터로 관찰된 것과 유사한 표적 유전자좌의 편집 수준을 달성한다. (도 98i) HbF+ 세포의 상당히 더 높은 백분율은 단일 gRNA 벡터와 비교하여 HD-Ad5/35 "Dual" gRNA 벡터로 형질도입된 HUDEP-2 세포에서 유세포 분석에 의해 관찰되었다. 유세포 분석 데이터를 요약하는 막대 차트는 유세포 분석 데이터 아래에 있다. (도 98j) HPLC에 의해 측정된 전반적인 감마 글로빈 발현은 이중 표적화된 샘플에서 상당히 더 높았다. (도 98k) 단일 녹아웃(knock-out) 클론보다 이중 녹아웃 클론에서 상당히 더 높은 태아 글로빈 발현이 관찰었으며 이는 더 높은 감마 발현/세포로 이어지는 2 개의 돌연변이의 가능한 상승 효과를 암시한다. (도 98l) 도식은 말초 혈액 동원된 CD34+ 세포가 HDAd5/35++ CRISPR 벡터로 형질도입되었음을 나타낸다. CRISPR/Cas9 세포독성을 최소화하기 위해, 세포를 항-Cas9 펩티드를 발현하는 HDAd5/35++ 벡터로 후속적으로 형질도입하였다. 세포를 준치명적으로 조사된 NSG 마우스 내로 이식하고 분석하였다. (도 98m) 이식 후 10 주차에, HD-Ad5/35 "Dual" gRNA 벡터로 형질도입된 세포는 단일 gRNA 벡터로 형질도입된 세포와 유사한 생착을 나타내었다. 계통 조성물은 모든 그룹에서 유사하였다. (도 98n) 이중 gRNA 벡터에 의해 형질도입되고 편집된 CD34+ 세포는 NSG 마우스에서 효율적으로 주입되었다. 또한, 적혈구 분화 후 주입된 이중 표적화된 세포는 상대적으로 더 낮은 편집 수준에도 불구하고, 단일 표적화된 세포와 비교하여, 대조군에 대해 더 높은 수준의 감마 글로빈을 발현하였다.

도 99a-99u. 이중 편집된 정상 및 thal CD34+ 세포의 생체외 형질도입. (도 99a) 실험 설계. 정상 CD34+ 세포에 대해 15 일째에 콜로니에서 (도 99b) HBF 발현 및 (도 99c) MFI. *는 p=0.034를 나타낸다. (도 99d) 정상 CD34+ 세포에서 15 일째에 콜로니에서 HBF 발현을 설명하는 유세포 분석 데이터. 정상 CD34+ 세포에 대한 적혈구 분화(ED) 후 (도 99e) HBF 발현 및 (도 99f) MFI. *는 p=0.01을 나타낸다. 정상 CD34+ 세포에서 형질도입 후(txd) 48 시간에 (도 99g) HBG 부위에 대한 TE71 및 (도 99h) BCL11A 부위에 대한 TE71. (도 99i) EC 및 적혈구 분화에서 HBF 발현을 설명하는 유세포 분석 데이터. (도 99j-99u) Thal CD34+ 세포. (도 99j) 0 일차 세포, 형질도입되지 않은 세포 및 CRISPR-Dual으로 형질도입된 세포의 면역표현형 및 (도 99k) 11 일에 걸쳐 형질도입되지 않은 세포 및 CRISPR-Dual으로 형질도입된 세포를 비교하는 성장 곡선. 15 일째에 콜로니에서 (도 99l) HBF 발현 및 (도 99m) MFI. **는 p=0.0046을 나타낸다. (도 99n) 형질도입되지 않은 세포에 대한 CRISPR-Dual을 비교하는 적혈구 및 골수성 구획에서 HBF 발현. (도 99o) 형질도입되지 않은 세포에 대한 CRISPR-Dual A 및 B를 비교하는 적혈구 및 골수성 구획에서 HBF 발현. (도 99p) EC에서 HBF 발현 및 (도 99q) MFI. ***는 p=0.0003을 나타내고 ****는 p=0.00003을 나타낸다. (도 99r) P04 및 P18에서 HBF 발현을 설명하는 유세포 분석 데이터. (도 99s, 99t) (도 99s) p04 및 (도 99t) p18에서 HBG 부위 적혈구 분화에 대한 TE71. (도 99u) 형질도입 48 시간 후 BCL11A 부위에 대한 TE71.

도 100. γ-글로빈 유전자 추가 및 내인성 γ-글로빈 재활성화의 조합을 설명하는 그래픽 요약.

도 101. 본원에 사용되는 HDAd5/35++ 벡터. γ-글로빈 유전자 추가는 발현 카세트를 플랭킹하는 IR 및 frt 부위가 있는 트랜스포존 벡터(HDAd-combo 및 HDAd-SB-추가 참조) 및 SB100x 및 Flpe 재조합효소를 트랜스로 제공하는 제2 벡터(HDAd-SB)로 이루어진 SB100x 트랜스포사제 시스템을 통해 달성된다. 무작위 통합을 위한 트랜스포존 카세트는 인간 γ-글로빈의 적혈구 특이적 발현을 위한 미니 β-글로빈 LCR/프로모터로 이루어진다. 3'UTR은 적혈구 세포에서 mRNA 안정화를 제공한다. γ-글로빈 발현 단위는 보편적으로 활성인 PGK 프로모터로부터 $mgmt^{P140K}$ 발현을 위한 카세트에서 닭 글로빈 HS4 절연체에 의해 분리된다. HDAd-CRISPR 및 HDAd-combo 벡터에서 CRISPR/Cas9 카세트는 EF1α 프로모터 제어 하에 SpCas9인 HBG1/2 프로모터 내의 BCL11A 결합 부위에 특이적인 U6 촉진-구동된 sgRNA로 이루어진다. HDAd 생산자 세포에서 Cas9의 발현은 miRNA 조절 시스템에 의해 억제된다(Saydaminova 등, *Mol Ther Methods Clin Dev*. 2015, 1: 14057, 2015). HDAd-combo에서,

CRISPR/Cas9 카세트는 트랜스포존 외부에 배치되어 Flpe/SB100x-매개 통합 시 손실될 것이다(도 102 참조).

도 102. 제어된 Cas9 발현에 대한 도식. HDAd-combo에서, Flpe 재조합효소와 frt 부위의 상호작용은 트랜스포존의 원형화를 야기하여, CRISPR 카세트를 함유하는 벡터의 선형 단편을 남긴다. SB100x/Flpe 시스템을 사용한 이전 연구는 원형화된 트랜스포존이 SB100x에 의해 숙주 게놈 내로 통합되는 동안 이들 벡터 부분이 빠르게 손실됨을 입증하였다(Yant 등, *Nat Biotechnol.*, 20: 999-1005, 2002).

도 103a-103d. Cas9 및 γ-글로빈 발현을 분석하기 위해 HUDEP-2 세포를 사용한 시험관내 연구. (도 103a 및 103b) 웨스턴 블롯에 의한 Cas9 발현의 분석. HUDEP-2 세포를 HDAd-combo 단독으로 및 HDAd-SB와 조합하여(즉, Flpe 및 SB100x를 트랜스로 제공하는 벡터) 형질도입하였다. 시험관내 적혈구 분화를 형질도입 후 4 일에 시작하고 8 일 동안 계속하였다. (적혈구 분화는 γ-글로빈 발현을 허용한다). 오른쪽 패널: Cas9 및 β-액틴 항체를 프로브로서 사용하는 대표적인 웨스턴 블롯. 왼쪽 패널: Cas9 신호의 요약. 막대는 HDAd-SB 공동 감염이 있거나 없는 Cas9, 즉, Flpe/SB100x 메커니즘에 의한 Cas9의 감소를 비교한다. (도 103c) 유세포 분석에 의한 γ-글로빈 발현 분석. HUDEP-2 세포를 HDAd-CRISPR("cut"), HDAd-SB-add("add")+HDAd-SB, 또는 HDAd-combo("combo")+HDAd-SB로 형질도입하고 표시된 시점에 분석하였다. (도 103d) qRT-PCR에 의한 γ-글로빈 mRNA 수준. d.p.t., 형질도입 후 일. Diff, 분화. * $p<0.05$

도 104a-104i. CD46/β-YAC 마우스의 생체내 형질도입 후 γ-글로빈 발현 연구. (도 104a) 실험 도식. HSPC를 4 일 동안 인간 재조합 G-CSF의 피하(s.c.) 주사 이어서 AMD3100의 1 회 s.c. 주사에 의해 동원하였다. AMD3100 주사 후 30 및 60 분에, 동물에게 다음 HDAd 벡터의 1:1 혼합물을 정맥내로 주사하였다(2 회 주사, 각각 $4x10^{10}$ vp): HDAd-combo+HDAd-SB, HDAd-SB-add+HDAd-SB, 및 HDAd-cut. 인간 γ-글로빈 및 MGMT에 대한 면역 반응을 피하기 위해 마우스를 다음 4 주 동안 면역억제성(IS) 약물로 처리하였다. 4 주차에, $O^6$-BG/BCNU 처리를 시작하고 2 주마다 3 회 반복하였다. 각 주기마다, BCNU 농도를 5 mg/kg에서 7.5 mg/kg 내지 10 mg/kg까지 증가시켰다. 18 주차에 동물을 조직 샘플 분석 및 치명적으로 조사된 C57Bl/6 마우스 내로 2차 이식을 위한 골수 $Lin^-$ 세포 수확을 위해 희생시킨 다음, 이후 16 주 동안 추적하였다. (도 104b) "combo" 및 "cut" 그룹에 대한 유세포 분석에 의한 말초 적혈구에서 γ-글로빈 발현의 검출. (도 104c) HPLC에 의해 측정된 γ-글로빈 단백질 수준. 오른쪽 패널: 인간 β-글로빈, 재활성화된 인간 Aγ, 및 추가된 γ-글로빈 쇄가 표시된 RBC 용해물(18 주)의 크로마토그램. 왼쪽 패널: HPLC 데이터의 요약. "cut", "add", 및 "combo" 벡터로 처리된 CD46/β-YAC 마우스에 대한 인간 β -글로빈에 대한 총 γ-글로빈의 백분율이 제시된다. *: $p<0.05$, n.s.. (도 104d) 마우스 β-주요 mRNA 발현에 대한 γ-글로빈 mRNA 발현(qRT-PCR에 의해 측정). (도 104e) CRISPR/Cas9에 의한 퍼센트 표적 부위 절단. 생체내 "cut" 및 "combo" 형질도입된 마우스로부터 18 주차에 수확된 PBMC 및 골수 MNC로부터의 게놈 DNA를 T7EI 검정에 적용하였다. 도 105의 데이터에 대한 요약이 제시된다. * $p<0.05$). (도 104f) "add" 및 "combo" 벡터로 형질도입 후 18 주차에 골수 HSPC에서 측정된 통합 벡터 카피 수. 그룹 사이의 차이는 유의하지 않다. (도 104g) "combo" 벡터 처리된 마우스의 개별 CFU에서 VCN의 스펙트럼. 골수 $Lin^-$ 세포를 전구체 검정을 위해 플레이팅하고 VCN을 qPCR에 의해 개별 콜로니에서 측정하였다. 4 마리의 상이한 마우스의 데이터가 제시된다. (도 104h) HPLC에 의한 인간 γ/인간 β 글로빈 단백질. (도 104i) 마우스 β-주요 mRNA 발현에 대한 인간 γ-글로빈 mRNA 발현의 백분율.

도 105a, 105b. 표시된 인간 β- 및 γ-글로빈 피크가 있는 RBC 용해물의 크로마토그램. (도 105a) 상부 패널은 처리 전 β-YAC 마우스를 나타낸다. 중간 패널은 HDAd-CRISPR("cut") 형질도입 후 18 주를 제시한다. 왼쪽 패널은 Gγ 및 Aγ 둘 다의 재활성화를 나타낸다. 하부 패널은 HDAd-CRISPR("cut") 형질도입 후 18 주를 나타낸다. (도 105b) 피크는 마지막 하단 패널에 표지되어 있다. 각 크로마토그램은 개별 동물이다. 인간 β-글로빈은 감소하고 γ-글로빈은 증가함에 유의한다(역 글로빈 스위치).

도 106. "cut" 및 "combo" 벡터로 형질도입 후 16 주차에 혈액, 비장, 및 골수의 MNC에 대한 T7EI 검정 데이터. 특이적 CRISPR/Cas9 절단 단편(255 및 110 bp)은 화살표로 표시된다. 밴드 신호 정량화에 기반한 절단 백분율은 각 레인 아래에 제시된다.

도 107a-107f. CD46/β-YAC 형질도입된 마우스로부터 $Lin^-$ 세포의 2차 수용자 분석. (도 107a) 표시된 시점에서 인간 γ-글로빈 발현 말초 혈액 RBC의 백분율. 모든 마우스는 이식후 4 주부터 시작하여 면역억제를 받았다. (도 107b) 이식 후 16 주차에 인간 β-글로빈에 대한 γ-글로빈 단백질의 수준. (도 107c 및 107d) 마우스 $\beta_{주요}$-글로빈 및 인간 β-글로빈에 대한 γ-글로빈 단백질의 수준. (도 107e) 형질도입되지 않은 대조군

마우스와 비교하여 "combo" 벡터로 형질도입 후 16 주차에 혈액, 비장, 및 골수의 MNC에서 계통-양성 세포 조성물. 도 107f. γ-글로빈 프라이머를 사용하여 qPCR에 의해 측정된 HDAd-comb 그룹의 총 백혈구에서 세포 당 벡터 카피 수.

도 108a-108d. SCD에 대한 모델로서 삼중 유전자이식 CD46/Townes 마우스의 생성 및 특성화. (도 108a) CD46/Townes 마우스의 번식. Townes 마우스(hα/hα::$\beta^S/\beta^S$)를 3 차례에 걸쳐 CD46 유전자이식 마우스와 번식시켰다. CD46, HbS 및 HBA에 대해 동형접합성인 동물을 생체내 형질도입 연구에 사용하였다. (도 108b) 부동변형적혈구증가증, 다염적혈구증가증(검정색 화살표), 겸상 및 단편화 세포(별표가 있는 검정색 화살표)를 포함하는 인간 질환의 전형적인 특징이 있는 CD46/Townes 마우스의 말초 혈액 도말 스케일 바는 15 μm이다. (도 108c) 모체 "건강한" CD46-유전자이식 마우스와 비교하여 CD46/Townes 마우스의 말초 혈액의 혈액학적 분석. Ret: 망상적혈구; RBC: 적혈구, Hb: 헤모글로빈; HCT: 적혈구용적률; WBC: 백혈구. 모든 차이는 유의하다($p<0.05$). (도 108d) CD46/Townes 마우스에서 비장비대. CD46tg 및 CD46/Townes 마우스에서 비장 대 체중의 비율이 제시된다. N=3.

도 109a-109f. CD46/Townes 마우스의 생체내 HSPC 형질도입 후 γ-글로빈 발현. 마우스를 동원하고, HDAd-combo+HDAd-SB를 주사하고, 도 104에 대해 기재된 바와 같이 $O^6$BG/BCNU로 처리하였다. (도 109a) 유세포 분석에 의해 측정된 말초 RBC에서 γ-글로빈 마킹. 빈 사각형은 미처리 CD46/Townes 마우스의 RBC에서 마킹을 나타낸다. 수직 화살표는 생체내 선택 주기를 나타낸다. (도 109b) HPLC에 의해 13 주차에 측정된 RBC에서 γ-글로빈 수준. 왼쪽 패널: 개별 마우스에서 인간 α-글로빈 및 $\beta^s$-글로빈 쇄에 대한 총 γ-글로빈 수준의 요약. 빈 사각형은 미처리 CD46/Townes 마우스에서 RBC의 수준을 나타낸다. 오른쪽 패널: 처리 전(상부 패널) 및 HDAd-combo+HDAd-SB로 생체내 HSPC 형질도입 후 13 주차에 CD46/Townes 마우스의 대표적인 크로마토그램. 인간 β-, $\beta^s$, 재활성화된 Aγ, 및 추가된 γ-글로빈에 대한 피크가 표시된다. (도 109c) HPLC에 기반한 재활성화된 Aγ의 백분율. (도 109d) 개별 마우스에서 인간 α-글로빈 및 $\beta^s$-글로빈 mRNA에 대한 총 γ-글로빈 mRNA의 백분율. (도 109e) HDAd-combo로 형질도입 후 163 주차에 골수 HSPC에서 측정된 통합 벡터 카피 수. (도 109f) HDAd-combo의 주사 후 13 주차에 CD46/Townes 마우스의 HBG1/2 표적 부위 절단 총 골수 핵 세포, $Lin^-$ 세포, PBMC, 및 비장세포. 특이적 CRISPR/Cas9 절단 단편(255 및 110 bp)은 화살표로 표시된다. 밴드 신호 정량화에 기반한 절단 백분율은 각 레인 아래에 제시된다.

도 110a, 110b. 형질도입된 CD46/Townes 마우스의 $Lin^-$ 세포로 이식된 2차 수용자의 분석. (도 110a) 인간 γ-글로빈 발현 말초 혈액 RBC의 백분율. (도 110b) 이식 후 16 주차에 인간 α- 및 $\beta^S$ 글로빈에 대한 γ-글로빈 단백질의 수준.

도 111a-111c. 혈액에서 표현형 교정. (도 111a) 브릴리언트 크레실 블루에 의한 망상적혈구에 대해 염색된 혈액 도말. 이 염료는 핵 및 세포질 구획의 잔류물을 염색한다. (정량화는 도 109c, 막대의 첫번째 그룹에서 발견될 수 있다). 스케일 바는 20 μm이다. (도 111b) HDAd-combo 유전자 요법 후 적혈구의 정상적혈구 형태학을 나타내는 혈액 도말. (도 111c) 말초 혈액의 혈액학적 분석. "CD46" 및 "combo 후 13 주차 CD46/Townes" 사이의 차이는 유의하지 않다.

도 112a-112c. 비장 및 간에서 표현형 교정. (도 112a) 조직 조직학. 상부 패널: 비장에서 철분 침착. Perl의 프러시안 블루 염색에 의해 비장 절편에서 헤모시데린을 검출하였다. 스케일 바는 20 μm이다. 중간 및 하부 패널: 비장 및 간 절편에서 헤마톡실린/에오신 염색에 의한 골수외 혈구형성. 간에서 적혈구모세포 및 CD46/Townes 마우스의 비장에서 거핵구의 클러스터는 흰색 화살표로 표시된다. 스케일 바는 20 μm이다. 대표적인 이미지가 제시된다. (도 112b) 처리된 CD46/Townes 마우스에서 보상 혈구형성의 측정가능한 특성인 비장 크기는 모체 CD46 마우스와 비교가능하다. (도 112c) 도 112a의 간 절편 이미지의 4-배 더 큰 배율. 처리 전에 CD46/Townes 마우스의 간 동양혈관에 포획된 겸상 RBC(왼쪽 패널) 및 처리 후에 간 동양혈관에서 겸상 적혈구의 부재(오른쪽 패널).

도 113. Ad5/35 헬퍼 바이러스 게놈의 왼쪽 단부. 진회색으로 음영 처리된 서열은 천연 Ad5 서열에 상응하며, 즉, 음영 처리되지 않거나 또는 연회색으로 강조된 서열은 인공적으로 도입되었다. 연회색으로 강조된 서열은 (나란히 반복된) loxP 서열의 2 개 카피이다. "cre 재조합효소" 단백질의 존재 하에, 2 개의 loxP 서열 사이의 뉴클레오티드 서열이 결실된다(loxP의 하나의 카피만 남아 있음). loxP 부위 사이의 Ad5 서열은 아데노바이러

스 DNA를 (생산자 세포의 핵에서) 캡시드 내로 패키징하는 데 필수적이기 때문에, 이 결실은 패키징할 수 없는 헬퍼 아데노바이러스 게놈 DNA를 초래한다. 결과적으로, 결실 과정의 효율은 패키징된 헬퍼 게놈 DNA의 수준(원치않은 헬퍼 바이러스 "오염")에 직접 영향을 미친다. 상기 관점에서, Ad5 이외의 아데노바이러스 혈청형에 대해 동일한 체계를 번역하기 위해, 다음을 달성하는 것이 바람직하다: 1. loxP 서열 삽입에 의해 플랭킹되고 cre 재조합효소의 존재 하에 결실될 수 있도록 패키징에 필수적인 서열을 식별한다. 이러한 서열의 식별은 서열에서 유사성이 거의 없는 경우 간단하지 않다. 2. 천연 DNA 서열에서 loxP 서열의 삽입이 헬퍼 바이러스의 증식 및 패키징에 최소 영향을 미칠 위치를 결정한다(cre 재조합효소의 부재 하에). 3. 헬퍼-의존적 아데노바이러스의 생산 동안(즉, 116 세포주와 같은 cre 재조합효소 - 발현 세포주에서) 헬퍼 바이러스 패키징을 최소로 유지하고 패키징 서열의 효율적인 결실을 허용하기 위해 loxP 서열 사이의 간격을 결정한다.

도 114. Ad5 및 Ad35 패키징 신호의 정렬(서열번호: 49 및 50). Ad5의 왼쪽 단부 서열과 Ad35의 정렬은 패키징 신호를 식별하는 데 도움을 준다. 패키징(AI 내지 AV)에 중요한 Ad5 서열의 모티프는 상자에 있다(Schmid 등, *J Virol.*, 71(5):3375-4, 1997의 도 1b 참조). loxP 삽입 부위의 위치는 검정색 화살표로 표시된다. 삽입은 AI 내지 AIV를 플랭킹하고 AV를 방해하는 것으로 보인다. Schmid 등에 표시된 바와 같은 추가적인 패키징 신호 AVI 및 AVII은 이 벡터의 E1 결실의 일부로서 Ad5 헬퍼 바이러스에서 결실되었음에 유의한다.

도 115. pAd35GLN-5E4의 도식. 이는 재조합 기술을 사용하여 벡터화된 Ad35 게놈(ATCC의 Holden 균주)에서 유래된 1세대(E1/E3-결실) Ad35 벡터이다(PMID: 28538186). 그런 다음 이 벡터 플라스미드를 사용하여 loxP 부위를 삽입하였다.

도 116. 플라스미드 패키징 신호에 대한 정보. 패키징 부위(PS)1 LoxP 삽입 부위는 뉴클레오티드 178 및 344 뒤에 있다. 이는 AI 내지 AIV를 제거하여야 한다. AVI 및 AVII을 포함하는 나머지 패키징 신호(344 이후)는 결실되었다(E1 결실(345 내지 3113)의 일부). PS2 LoxP 삽입 부위는 뉴클레오티드 178 및 481 뒤에 있다. 추가적으로, 뉴클레오티드 179 내지 365는 결실되어, AI 내지 AV는 존재하지 않는다. 나머지 패키징 모티프 AVI 및 AVII은 HDAd 생산 동안 cre 재조합효소에 의해 제거가능하다. E1 결실은 482에서 3113까지이다. PS3 LoxP 삽입 부위는 뉴클레오티드 154 및 481 뒤에 있다. 3 개의 조작된 벡터가 구제될 수 있다. 재배열된 loxP 부위가 있는 바이러스 게놈의 백분율은 PS1, PS2, 및 PS3에 대해 각각 50, 20, 및 60%였다. lox P 부위가 바이러스 복제 및 유전자 발현에 중대하게 영향을 미칠 때 재배열이 발생한다. 재배열된 loxP 부위가 있는 벡터는 패키징될 수 있고 HDAd prep을 오염시킬 것이다. 서열번호: 286, 51, 및 52는 각각 PS1, PS2, 및 PS3으로 도해된 벡터를 예시한다.

도 117. 현재 HDAd5/35 플랫폼과 비교한 차세대 HDAd35 플랫폼. 두 벡터는 CMV-GFP 카세트를 함유한다. Ad35 벡터는 면역원성 Ad5 캡시드 단백질을 함유하지 않는다. 시험관내에서 CD34+ 세포의 비슷한 형질도입 효율을 나타낸다. 가교 연구는 시험관내에서 CD34+ 세포의 비슷한 형질도입 효율을 나타낸다. G-CSF 동원된 공여자의 말초 CD34+ 세포인 인간 HSC를 HDAd35(Ad35 헬퍼 P-2로 생성) 또는 Ad5 캡시드와 Ad35의 섬유를 함유하는 키메라 벡터로 500, 1000, 2000 vp/세포의 MOI에서 형질도입하였다. GFP-양성 세포의 백분율을 3 개의 독립적인 실험에서 바이러스 추가 후 48 시간에 측정하였다. 특히, HDAd35로 감염은 헬퍼 바이러스 오염으로 인해 48 시간에 세포변형 효과를 촉발하였다.

도 118. PS2 헬퍼 벡터를 원숭이 연구에 초점을 맞춰 새로 만들었다. 다음으로부터 학습된 조치가 이어진다: E1 영역의 결실, Loxp에 의해 플랭킹된 돌연변이체 패키징 신호, 돌연변이체 패키징 서열, E3 영역의 결실(27435→30540), Ad5E4orf6으로 대체, copGFP 카세트를 플랭킹하는 스터퍼 DNA의 삽입, 및 놉에 돌연변이를 도입하여 Ad35K++ 제조.

도 119. 돌연변이된 패키징 신호 서열이 제공된다. 잔기 1 내지 137은 Ad35 ITR이다. 굵은 글씨는 SwaI 부위이고, Loxp 부위는 이탤릭체이고, 돌연변이된 패키징 신호는 밑줄로 표시된다.

도 120a, 120b. 다양한 헬퍼 벡터 및 패키징 신호 변이체의 개략적 도면. 구현예에서, E3 영역(27388 → 30402)은 결실되고 CMV - eGFP 카세트는 E3 결실, Ad35K++ 내에 위치하고, eGFP는 copGFP 대신에 사용된다. (도 120a)에 제시된 패키징 신호 변이체를 함유하는 4 개의 헬퍼 벡터 모두 구제될 수 있다. loxP 부위는 증폭이 보다 효율적일 수 있을 때 재배열되었다. 추가적인 패키징 신호 변이체는 도 120b에 예시되어 있다.

도 121. HDAd-combo 벡터의 묘사.

도 122. 실험 프로토콜.

도 123. +58 적혈구 bcl11a 인핸서 영역 내에서 GATAA 모티프를 편집하기 위한 벡터. 벡터 구조는 상부 패널

에 제시되어 있다. 두 벡터는 GATAA 모티프를 표적화한다. 하부 패널은 HDAd-C-BE 벡터에 의해 매개된 염기 변화를 나타낸다. (서열번호: 65-68)

도 124a-124c. 인간 CD34+ 세포에 대한 벡터 분석. (도 124a) 세포를 2000 vp/세포의 MOI에서 감염시키고 1 일 후에 18 일 동안 적혈구 분화에 적용하였다. (도 124b) 상이한 시점에서 T7E1A 검정에 의한 표적 부위 절단을 위해 세포 분취액을 분석하였다. 왼쪽 막대: HDAd-wtCRISPR, 오른쪽 막대: HDAd-C-BE. (도 124c) 적혈구 분화의 종료에 γ-글로빈+ 세포의 백분율.

도 125. HDAd-wtCRISPR 및 HDAd-C-BE 형질도입된 CD34+ 세포의 생착. 형질도입의 MOI는 2000 vp/세포였다. 생착은 말초 혈액 단핵 세포에서 인간 CD45+ 세포의 백분율에 기반하여 측정하였다.

도 126. 염기 편집기 HDAd 벡터. sgRNA는 HBG1/2에서 적혈구 bcl11a 인핸서(상부 패널) 또는 BCL11a 단백질 결합 부위를 표적화한다. 중간 패널은 적혈구 전구체 세포주 HUDEP-2의 적혈구 분화 당일에 염기 전환율 %을 나타낸다. 오른쪽 패널은 γ-글로빈 재활성화 수준을 나타낸다. (서열번호: 67, 65, 및 71)

도 127a, 127b. (도 127a) 전형적인 겸상-유사 적혈구를 사용한 혈액 도말. (도 127b) 적혈구 매개변수.

도 128a-128c. (도 128a) 생체내 선택 없이 Townes/CD46 마우스의 생체내 형질도입. (도 128b) RBC에서 γ-글로빈 재활성화. (도 128c) 처리 전 및 처리 8 주차에 혈액 도말의 망상적혈구 염색.

도 129a-129d. 동원된 원숭이에서 생체내 HSC 형질도입. G-CSF, SCF, 및 AMD3100으로 동원 후, 2 마리의 수컷 원숭이에게 정맥내 주사에 의해 HDAd-GFP($1x10^{12}$vp/kg)를 제공하였다. HDAd 주사 전에, 동물을 덱사메타손을 전처리하여 가능한 사이토카인 방출을 차단하였다. (도 129a) 표시된 시점으로부터 정제된 말초 혈액 CD34+ 세포를 배양하고 유세포 분석에 의해 GFP 발현에 대해 분석하였다. 배양물에서 4 일에 걸쳐 GFP를 발현하는 세포의 평균 퍼센트가 제시된다 (도 129b) HDAd-GFP 주사 전(0 시간) 또는 후(6 시간) GFP를 발현하는 정제된 CD34+ 세포의 대표적인 흐름도. (도 129c) 콜로니 형성 검정을 말초 혈액 또는 총 PBMC로부터 정제된 CD34+ 세포로 개시하였다. 배양물에서 14 일 후, 개별 콜로니를 선별하고 PCR에 의해 GFP DNA의 존재에 대해 분석하였다. (도 129d) 골수 CD34+ 세포에서 GFP 발현의 분석. 대표적인 블롯이 제시된다. 이 연구에서, HDAd-GFP만을 주사하고 따라서 단기간 GFP 발현만을 측정하였다.

도 130. 가이드 서열의 스크리닝. HUDEP-2 세포를 표 14에 나열된 염기 편집기로 형질감염시켰다. γ-글로빈 발현을 형질감염 후 4 일차(4dpt) 및 시험관내 적혈구 분화 후 6 일차(Diff 6d)에 측정하였다. *HBG1/2* 프로모터에서 TGACCA 모티프를 표적화하는 CRISPR/Cas9 벡터를 양성 대조군(pos ctrl)으로 사용하였다. CBE 표적화 CCR5 코딩 영역을 음성 대조군(sgNeg)으로 포함하였다. 제시된 데이터(평균 ± SD)는 2 개의 독립적인 실험을 나타낸다.

도 131a, 131b. 상이한 버전의 사이티딘 염기 편집기의 비교. (도 131a) 293 세포(HEK293)를 WTCas9 또는 BE 벡터 + pSP-BE4-sgBCL11Ae1(3 + 1 μg)로 형질감염시키고 bcl11a 인핸서 표적 부위 절단 부위는 T7E1 검정에 의해 형질감염 4 일 후에 분석하였다. (도 131b) 동일한 연구를 적혈구백혈병 세포주(K562) WTCas9 또는 BE 벡터 + pSP-BE4-sgBCL11Ae1(2 + 0.66 μg)에서 수행하였다.

도 132a-132c. HDAd5/35++_BE 벡터의 설계 및 구제. (도 132a) 사이티딘 염기 편집기(CBE) 벡터 설계. 구제 가능하지만 낮은 수율. (도 132b) 아데닌 염기 편집기(ABE) 벡터 설계의 첫번째 버전. 구제가능하지 않음. (도 132c) 반복성을 감소시키기 위한 ABE 코돈 최적화. TadA(tRNA 아데노신 데아미나제 효소)의 코돈 최적화를 나타내는 서열 비교를 포함한다(서열번호: 260 및 261)

도 133a-133h. HDAd5/35++_BE 벡터의 구축 및 검증. (도 133a) HDAd_ABE 벡터 다이어그램. 2 개의 frt-IR에 의해 플랭킹된 4.2 kb MGMT/GFP 카세트는 HDAd_SB 벡터와 공동-전달될 때 통합된 발현을 허용한다. 8.0 kb 염기 편집기 구성요소를 일시적 발현을 위해 트랜스포존 외부에 설계하였다. 2 개의 TadAN 반복부는 반복 서열을 감소시키기 위해 코돈 최적화되었다(*는 촉매 반복부를 나타낸다). microRNA 반응성 요소(miR)를 3' 인간 β-글로빈 UTR에 포매시켜 116 세포에서 ABE 발현을 특이적으로 하향조절함으로서 생산자 세포에 대한 독성을 최소화하였다. PGK, 인간 PGK 프로모터. bGHpA, 소 성장 호르몬 폴리아데닐화 서열. SV40pA, 시미안 바이러스 40 폴리아데닐화 신호. ITR, 도립된 말단 반복부. Ψ, 패키징 신호. (도 133b) 생성된 바이러스 벡터의 정보. 나열된 수율은 하나의 3L 스피너로부터 기인한다. (도 133c) HUDEP-2 세포에서 바이러스 벡터의 검증. 세포를 표시된 MOI(vp/세포)에서 다양한 벡터로 형질도입하였다. γ-글로빈 발현을 형질감염 후 4 일차(4dpt) 및 시험관내 적혈구 분화 후 6 일차(Diff 6d)에 측정하였다. CCR5 코딩 영역을 표적화하는 CBE 벡터를 음성 대조군

(sgNeg)으로 포함하였다. 제시된 데이터(평균 ± SD)는 2 개의 독립적인 실험을 나타낸다. (도 133d) HDAd_sgHBG#2에 의한 표적 염기 전환. 표적화 염기를 포함하는 HBG1 또는 HBG2 게놈 분절을 증폭시키고 Sanger 서열분석에 적용하였다. 데이터를 EditR 1.0.9에 의해 분석하였다. 화살표는 표적화 염기를 나타낸다. 전환율 %은 크로마토그램 아래에 제시되었다. (도 133e) 분화 후 6 일차에 HPLC에 의해 측정된 α- 또는 β-글로빈에 대한 γ-글로빈 발현의 %. MOI=1000. 제시된 데이터(평균 ± SD)는 2 개의 독립적인 실험을 나타낸다. 도 133f-133h) HDAd_sgHBG#2로 형질도입된 HUDEP-2 세포로부터 유래된 대표적인 클론(#3). 단일대립유전자성 -116A→G 염기 전환이 HBG1 프로모터에서 검출되었으며(도 133f), 이는 유세포 분석에 의한 100% γ-글로빈+ 세포를 초래하였다(도 133g). γ-글로빈 단백질 수준은 HPLC에 의해 측정하였다(도 133h).

도 134a-134c. 도 133을 뒷받침하는 데이터. (도 134a) 도 133d의 보충. 표시된 바이러스로 처리된 HUDEP-2 세포에서 표적 염기 전환. (도 134b) 대표적인 단일 세포 HUDEP-2 클론. 도 133f의 보충. 화살표가 있는 B는 이중대립유전자성 편집을 나타내고 M 및 화살표는 단일대립유전자성 편집을 나타낸다. (도 134c) 상응하는 단일 세포 HUDEP-2 클론에서 γ-글로빈 발현은 상기 제시된다. 도 133g의 보충.

도 135a-135i. 생체내 형질도입 및 선택 후 βYAC 마우스에서 γ-글로빈의 재활성화. (도 135a) 실험 절차. β-YAC/CD46 마우스(n=9)를 G-CSF/AMD3100에 의해 동원하고 HDAd_sgHBG#2 + HDAd_SB로 생체내 형질도입하였다. $O^6$BG/BCNU에 의한 4 차례의 선택을 형질도입 후 각각 4, 6, 8 및 10 주차에 수행하였다. 마우스를 16 주차에 안락사시켰다. 계통$^-$ 세포를 단리하고 치명적으로 조사된 C57BL/6 마우스에 IV 주사하였다. 2차 이식된 마우스를 이후 16 주 동안 추적하였다. (도 135b) 형질도입 후 상이한 시점에서 PBMC의 GFP 마킹. 각 점은 1 마리의 동물을 나타낸다. (도 135c) PBMC에서 GFP 발현의 대표적인 점 플롯. (도 135d) 유세포 분석에 의해 측정된 혈액 세포에서 γ-글로빈 발현. (도 135e) 혈액 세포에서 γ-글로빈 발현의 대표적인 점 플롯. (도 135f) 1차 마우스의 종료점에서 혈액 및 골수의 Ter-119$^+$ 및 Ter-119$^-$ 세포에서 유세포 분석에 의한 γ-글로빈 발현. (도 135g) HPLC에 의해 측정된 적혈구 용해물에서 γ-글로빈 단백질 수준. 제시된 데이터는 마우스 α- 또는 β-글로빈 또는 인간 β-글로빈에 대한 백분율이다. (도 135h) RT-PCR에 의해 측정된 mRNA 수준에서 γ-글로빈 발현. 제시된 데이터는 마우스 *HBA* 또는 *HBB*, 또는 인간 *HBB* mRNA에 대한 배수 변화이다. (도 135i) 총 골수 세포에서 벡터 카피 수(세포 당 카피). MGMT에 대한 프라이머를 사용하였다.

도 136. 도 135h에 제시된 대표적인 데이터의 HPLC 플롯.

도 137a-137g. 표적 염기 전환. (도 137a) sgHBG#2 가이드 서열. 넘버링은 5' 단부에서 시작하였다. 주황색 배경으로 강조된 것은 보고된 BCL11A 결합 부위인 TGACCA 모티프이다. 모티프에서 2 개의 아데닌(A5 및 A8)은 2 개의 화살표로 표시하였다. (도 137b) 표적 염기 전환 백분율. *HBG1* 및 *HBG2* 프로모터 영역에서 두 A5 및 A8이 제시되었다. 각 점은 1 마리의 동물을 나타낸다(n=9). (도 137c) 마우스 #1108의 *HBG1* 및 *HBG2* 영역에서 표적 염기 전환을 나타내는 대표적인 크로마토그램. (도 137d) γ-글로빈 발현에 대한 평균 염기 전환 사이의 상관관계. 각 동물에서 평균 염기 전환 백분율은 *HBG1* 및 *HBG2* 프로모터 영역의 A5 및 A8에서 평균 수준이었다. 각 점은 1 마리의 동물을 나타낸다(n=9). (도 137e) A5 및 A8에서 염기 전환 비교. 각 점은 1 마리의 동물을 나타낸다(n=9). (도 137f) 표적화된 아데닌 뉴클레오티드에서 전환 백분율을 나타내는 차트. (도 137g) 특정 마우스에서 표적화 염기 전환을 나타내는 크로마토그램(서열번호: 250).

도 138a-138d. 안전성 프로파일. (도 138a) 형질도입 후 16 주차에 혈액 샘플을 사용한 HEMAVET®에 의한 혈액학 분석. 제시된 데이터는 HDAd_sgHBG#2로 형질도입된 9 마리의 마우스 및 3 마리의 형질도입되지 않은 대조군 마우스를 나타내는 평균 ± SD이다. (도 138b) 16 주차에 혈액 샘플의 망상적혈구 백분율. 샘플을 브릴리언트 크레실 블루로 염색하였다. 제시된 데이터는 HDAd_sgHBG#2로 형질도입된 4 마리의 마우스 및 3 마리의 형질도입되지 않은 대조군 마우스를 나타내는 평균 ± SD이다. (도 138c) 1차 마우스의 종료점에서 골수 MNC의 세포 조성물. 형질도입되지 않은 마우스를 대조군으로 사용하였다. 각 점은 1 마리의 동물을 나타낸다. (도 138d) 브릴리언트 크레실 블루로 염색한 대표적인 망상적혈구.

도 139a-139c. 2차 이식. (도 139a) 유세포 분석을 사용하여 PBMC에서 인간 CD46 발현에 의해 측정된 생착. (도 139b) PBMC에서 GFP 발현. (도 139c) γ 유세포 분석에 의해 검출된 말초 혈액 세포에서 γ-글로빈 발현.

도 140a, 140b. 유전자간 결실의 검출. (도 140a) 유전자간 4.9 k 결실의 검출은 이전에 기재되었다(Li 등, *Blood*, 131(26): 2915, 2018). 총 골수 MNC로부터 단리된 게놈 DNA를 주형으로 사용하였다. *HBG1* 및 *HBG2* 프로모터에서 2 개의 CRISPR 절단 부위에 걸쳐 있는 9.9 kb 게놈 영역을 PCR에 의해 증폭시켰다. 생성물에서 추가 5.0 kb 밴드는 4.9 k 결실의 발생을 나타낸다. 결실 백분율을 4.9 kb 결실의 비로 정의된 주형을 사용하

여 PCR에 의해 생성된 표준 곡선 공식에 따라 계산하였다. HBG1/2 프로모터를 표적화하는 CRISPR 벡터로 형질도입된 생체내 마우스로부터 유래된 샘플을 비교에 사용하였다. 각 레인은 1 마리의 동물을 나타낸다. (도 140b) 도 140a에서 결실 백분율의 요약. 각 점은 1 마리의 동물을 나타낸다.

도 141. BE vs CRISPR/Cas9의 세포독성. CRISPR/Cas9를 사용한 현재 게놈-편집 기술에 대한 주요 관심사는 이중-가닥 DNA 파괴(DSB)를 도입하여, 원치않은 큰 단편 결실 및 p53-의존적 DNA 손상 반응을 야기함으로써 숙주 세포에 해로울 수 있다는 점이다. 염기 편집기는 표적화된 게놈 유전자좌에서 정확한 뉴클레오티드 돌연변이를 설치할 수 있고 DSB를 피하는 이점이 있다. 이 연구는 HSC의 주요한 기능적 특징, 즉 준치명적으로 조사된 NSG 마우스에서 생착이 BE에 의해 영향을 받지 않지만 CRISPR/Cas9 발현 벡터로 인간 CD34+ 세포의 형질도입 후 극적으로 감소됨을 나타낸다.

도 142. BE4-sgBCL11AE1에 의해 매개된 예상된 편집. BCL11A 유전자좌의 편집을 나타내는 도식. GATAA 모티프(서열번호: 65) 및 염기 편집 후 방해된 GATAA 모티프(서열번호: 67)가 제시된다.

도 143. 표적에 대한 최적 위치. 표적화를 위한 예시적인 위치를 강조하는 핵산 서열의 도식. 도면은 표적 C가 프로토스페이서 내의 위치 4 내지 8에 있을 때 부분적으로 C에서 T로의 편집을 나타낸다.

도 144는 염기 편집기를 암호화하는 벡터의 도식이다.

도 145. 바이러스 gDNA의 다이어그램. 단일 인접한 작제물을 나타내지만 제시의 용이성을 위해서만 2 개의 절편으로 나눠진 바이러스 gDNA(HBG2-miR, 아데닌 편집기)의 도식.

도 146. TadA 서열. 2 개의 'TadA +32aa'의 DNA 서열(서열번호: 367 및 268)을 포함하는, TadA 및 TadA*의 서열(서열번호: 265 및 266)의 개략도.

도 147. 염기 편집. 야생형(서열번호: 269) 및 편집된 서열(서열번호: 269)의 개략도.

도 148. 염기 편집. HDAd5/35++_BE4-sgBCL11Ae1-FI-mgmtGFP(041318-1) 바이러스에 의한 염기 편집과 관련한 개략도 및 2 개의 겔.

도 149. γ-글로빈+ 세포의 퍼센트. 표시된 MOI에서 γ-글로빈+ 세포의 백분율을 나타내는 그래프.

도 150. 염기 편집에 의한 HbF의 재활성화. 벡터 및 관련 정보의 목록.

도 151. 벡터 및 관련 정보의 목록, 및 염기 편집기의 다양한 MOI에서 퍼센트 HbF+ 세포를 나타내는 그래프.

도 152. γ-글로빈 발현(HUDEP-2), 2차 시험. HUDEP-2 세포에서 2차 시험의 % HbF+를 나타내는 그래프.

도 153. γ-글로빈 발현(HUDEP-2), 단일 세포 유래 클론. 다양한 단일 세포 유래 클론에서 % HbF+를 나타내는 그래프.

도 154a-154s. 개별 단일 세포 유래 클론을 나타내는 데이터. 도 154a-154s는 각각 단일 세포 클론을 나타내는 데이터를 포함한다. (서열번호: 271, 250, 252)

도 155. 293FT 세포에서의 테스트. 293FT 세포에서 염기 편집기의 사용 결과를 나타내는 2 개의 겔.

도 156a-156d. 편집된 염기(293FT 세포)를 확인하기 위한 Sanger 서열분석. 도 156a-156d는 각각 sanger 서열분석 결과를 나타내는 크로마토그램(들)을 포함한다. (서열번호: 269, 275-278)

도 157. HUDEP-2 세포에서의 테스트. 형질감염 4 일 후 HUDEP-2 세포에서 염기 편집기의 사용 결과를 나타내는 2 개의 겔.

도 158. γ-글로빈 발현(HUDEP-2). γ-글로빈의 발현을 나타내는 그래프.

도 159a-159d. 편집된 염기(HUDEP-2 세포)를 확인하기 위한 Sanger 서열분석. 도 159a-159d는 각각 이용가능한 경우 Sanger 서열분석 결과를 나타내는 크로마토그램(들)을 포함한다. (서열번호: 269, 275-278)

도 160. (Maxi 제조 하에) HDAd 바이러스 생산을 위한 선택된 작제물. (Maxi 제조 하에) HDAd 바이러스 생산을 위한 특정 작제물의 구축된 벡터 표시 선택 목록.

도 161. huCD45+ 세포의 생착을 나타내는 차트.

도 162. HUDEP-2 세포의 일시적 형질감염(T7EI에 의해 절단). HUDEP-2 세포의 일시적 형질감염(T7EI에 의해

절단) 결과를 나타내는 겔.

도 163. 이중 염기 편집 벡터 적용. 이중 염기 편집 벡터 구현예(서열번호: 279)의 개략도.

도 164. SB100x 트랜스포사제에 의한 인간 γ-글로빈/mgmt. 유전자 추가 및 HBG 프로모터에서 적혈구 bcl11a 인핸서 및 BCL11A 결합 부위를 표적화하는 CRISPR를 사용한 레서스 γ-글로빈 재활성화를 나타내는 HDad5/35++combo 벡터의 벡터 도식.

도 165. HDAd-sgAAVS1-rm(no Cas9) 벡터 및 HDAd-Comb2를 나타내는 벡터 도식. 이 벡터의 특성은 1.8k 상동성 아암(HA), PBMC에서 형질도입을 추적하기 위한 GFP, HA 외부의 CRISPR 카세트, 및 HBG 프로모터 표적화이다.

도 166. LCR β-글로빈 프로모터 구동된 외인성 γ-글로빈 및 γ-글로빈 프로모터의 억제인자 결합 영역의 CRISPR/Cas9-매개 파괴를 통한 내인성 γ-글로빈의 재활성화를 사용하여 γ-글로빈을 발현하는 HDAd-rh-combo의 벡터 도식.

## *발명을 실시하기 위한 구체적인 내용*

[0093] 본 개시내용은 무엇보다도, 조혈 줄기 세포의 생체내 유전자 편집을 위해 CD46을 표적화하는 Ad5/35 및 Ad35 벡터와 같은 재조합 아데노바이러스 벡터를 기재한다. Ad35 벡터는 CD46 결합을 증가시키는 놉 단백질 돌연변이, 유전자의 발현을 조절하는 miRNA 제어 시스템, 내인성 유전자 발현을 활성화시키는 CRISPR 구성요소, 양성 선택 마커, 미니- 또는 긴-형태의 β-글로빈 유전자좌 제어 영역(LCR) 조절 서열, 트랜스포사제/재조합효소 시스템, 및/또는 조건화-없는 생체내 유전자 요법을 촉진하는 다수의 다른 유리한 진보를 포함하나 이에 제한되지 않는 본원에 개시된 다양한 다른 서열을 포함할 수 있다.

[0094] 유전자 요법에 대한 많은 도구의 개발에도 불구하고, 벡터 및/또는 치료적으로 유용한 페이로드의 설계는 당 분야에서 중요한 과제로 남아있다. 유전자 요법 페이로드는 바이러스 벡터 또는 비-바이러스 벡터에 의해 전달될 수 있다. 예시적인 비-바이러스 벡터는 양이온성 지질, 지질 나노 에멀젼, 고체 지질 나노입자, 펩티드, 및 중합체-기반 전달 시스템을 포함한다. 바이러스 벡터는 AAV, 단순 포진, 레트로바이러스, 렌티바이러스, 알파바이러스, 플라비바이러스, 랍도바이러스, 홍역 바이러스, 뉴캐슬병 바이러스, 폭스바이러스, 피코르나바이러스, 콕사키바이러스 벡터, 및 아데노바이러스 벡터를 포함할 수 있으며, 각각은 다양한 별개의 특성을 갖는다. 아데노바이러스 중에는, 또한 50 개 초과의 혈청형이 있다. 또한 단백질, 조절 핵산, CRISPR/Cas9 시스템, 염기 편집 시스템, 트랜스포존 시스템, 및 상동 재조합 시스템을 암호화하는 페이로드를 포함하나 이에 제한되지 않는 핵산 서열의 발현 및/또는 변형을 위한 치료적 페이로드가 존재한다. 본원에 제공된 유전자 요법을 위한 방법 및 조성물은 제한 없이 아데노바이러스 벡터 및/또는 다양한 치료적 페이로드의 활용에서 다양한 과제를 해결한다.

[0095] 본 명세서의 개시내용은 특정 맥락(예를 들어, 아데노바이러스 벡터 또는 게놈 맥락, 예를 들어, Ad5, Ad5/35, 또는 Ad35 맥락)에 있을 수 있지만, 각 구성요소는 임의의 이러한 맥락과 상관없이 추가로 개시되고 이와 같이 이러한 맥락과 상관없이 청구될 수 있다. 예시적인 개시내용은 본 개시내용의 서열 및 페이로드 작제물을 포함하며, 당업자는 임의의 특정 벡터, 혈청형, 또는 다른 맥락에 제한되지 않는 일반적인 연관성을 가질 수 있음을 이해할 것이다.

[0096] 본 개시내용의 측면은 이제 다음과 같이 추가로 상세하게 기재된다: (I) 유전자 요법 벡터; (II) 표적 세포 집단; (III) 투여량, 제형, 및 투여; (IV) 적용; (V) 예시적인 구현예; (VI) 실험적 실시예; 및 (VII) 마무리 단락.

[0097] **I. 유전자 요법 벡터**

[0098] 아데노바이러스(또는, 상호교환가능하게, "아데노바이러스") 벡터 및 게놈은 (a) 발현 작제물의 패키징을 지원하고 (b) 코딩 서열을 발현하기에 충분한 아데노바이러스 서열을 함유하는 그러한 작제물을 지칭한다. 아데노바이러스 게놈은 선형 이중-가닥 DNA 분자일 수 있다. 당업자가 이해하는 바와 같이, 아데노바이러스 게놈과 같은 선형 게놈은 예를 들어, 바이러스 생산 목적을 위해 원형 플라스미드에 존재할 수 있다.

[0099] 천연 아데노바이러스 게놈은 혈청형에 따라 26 kb 내지 45 kb 길이 범위이다.

[0100] 아데노바이러스 벡터는 프리마제-독립적 DNA 합성을 촉진하고 숙주 게놈 내로의 통합을 용이하게 하는 자기-프라이머로서 작용하는 도립된 말단 반복부(ITR)에 의해 양 말단 상에 플랭킹된 아데노바이러스 DNA를 포함한다. 아데노바이러스 게놈은 또한 적절한 바이러스 전사체 패키징을 용이하게 하고 게놈의 왼쪽 아암 상에 위치하는

패키징 서열을 함유한다. 바이러스 전사체는 초기 전사 단위, E1, E2, E3, 및 E4 및 Ad 비리온의 구조적 구성 요소를 암호화하는 후기 전사 단위를 포함하는 여러 단백질을 암호화한다(Lee 등, *Genes Dis.*, 4(2):43-63, 2017).

[0101] 아데노바이러스 벡터는 아데노바이러스 게놈을 포함한다. 재조합 아데노바이러스 벡터는 재조합 아데노바이러스 게놈을 포함하는 아데노바이러스 벡터이다. 재조합 아데노바이러스 벡터는 아데노바이러스의 유전적으로 조작된 형태를 포함한다. 당업자는 본 출원 전반에 걸쳐 아데노바이러스 벡터의 개시내용이 이의 아데노바이러스 게놈의 개시내용을 포함하고, 아데노바이러스 게놈의 개시내용이 개시된 아데노바이러스 게놈을 포함하는 아데노바이러스 벡터의 개시내용을 포함함을 이해할 것이다.

[0102] 아데노바이러스는 큰 20면체 형상의 비-외피보유 바이러스이다. 바이러스 캡시드는 섬유, 펜톤, 및 헥손 기반 단백질을 포함하는 3 가지 유형의 단백질을 포함한다. 헥손은 대부분 바이러스 캡시드로 구성되어, 20 개의 삼각형 면을 형성한다. 펜톤 염기는 캡시드의 12 개 꼭지점에 위치하고 섬유(또한 놉형 섬유로 지칭됨)은 각 펜톤 염기로부터 돌출된다. 펜톤 및 섬유인 이들 단백질은 캡시드가 숙주 세포에 부착하는 것을 용이하게 하기 때문에 수용체 결합 및 내재화에 특히 중요한 것이다(Lee 등, *Genes Dis.*, 4(2):43-63, 2017).

[0103] Ad35 섬유는 섬유 단백질 삼량체이며, 각 섬유 단백질은 오량체성 펜톤 염기와 상호작용하는 N-말단 꼬리, 숙주 세포 수용체에 대한 부착 부위로서 기능하는 C-말단 구형의 놉 도메인(섬유 놉), 및 꼬리 및 놉 도메인을 연결하는 중심 샤프트(shaft) 도메인을 포함한다(샤프트). 삼량체성 섬유의 꼬리 도메인은 5-배 축에서 오량체성 펜톤 염기에 부착된다. 다양한 구현예에서, Ad35 섬유 놉은 표준 야생형 Ad35 섬유 단백질의 아미노산 123 내지 320을 포함한다. 다양한 구현예에서, Ad35 섬유 놉은 표준 야생형 Ad35 섬유 단백질의 아미노산 123 내지 320의 상응하는 단편과 적어도 80%(예를 들어, 적어도 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 서열 동일성) 서열 동일성을 갖는 적어도 60 개의 아미노산 (예를 들어, 적어도 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 또는 198 개의 아미노산)을 포함한다. 다양한 구현예에서, 섬유 놉은 표준 야생형 Ad35 섬유 단백질을 포함하는 참조 섬유 놉, 섬유 단백질, 섬유 또는 벡터와 비교하여, 섬유 단백질, 섬유, 또는 벡터에 CD46과의 증가된 친화도에 대해 조작되고/되거나, CD46과의 증가된 친화도를 부여하도록 조작되며, 임의적으로 여기서 증가는 적어도 1.1-배, 예를 들어, 적어도 1, 2, 3, 4, 5, 10, 15, 또는 20-배의 증가이다. 중심 샤프트 도메인은 5.5 개의 β-반복부로 구성되며, 각각은 β-턴으로 연결된 2 개의 역평행 β-가닥을 코딩하는 15-20 개의 아미노산을 함유한다. β-반복부는 연결되어 매우 강성이고 안정한 3 개의 서로 엮인 나선형 가닥의 연장된 구조를 형성한다.

[0104] 아데노바이러스는 중간 크기의 게놈, 조작 용이성, 높은 역가, 넓은 표적-세포 범위 및 높은 감염성으로 인해 유전자 전달 벡터로서 사용하기에 특히 적합하다. 바이러스 게놈의 양 단부는 100-200 개의 염기쌍 ITR을 함유하며, 이는 바이러스 DNA 복제 및 패키징에 필요한 시스 요소이다. 게놈의 초기(E) 및 후기(L) 영역은 바이러스 DNA 복제 개시에 의해 구분되는 상이한 전사 단위를 함유한다. E1 영역(E1A 및 E1B)은 바이러스 게놈 및 몇몇 세포 유전자의 전사 조절을 담당하는 단백질을 암호화한다. E2 영역(E2A 및 E2B)의 발현은 바이러스 DNA 복제를 위한 단백질의 합성을 초래한다. 이들 단백질은 DNA 복제, 후기 유전자 발현 및 숙주 세포 차단(shut-off)에 수반된다. 대부분의 바이러스 캡시드 단백질를 포함하는 후기 유전자 산물은 주요 후기 프로모터(MLP)에 의해 비롯된 단일 1차 전사체의 상당한 처리 후에만 발현된다. MLP는 감염의 후기 단계 동안 특히 효율적이고, 이 프로모터로부터 비롯된 모든 mRNA는 번역에 선호되는 mNRA를 초래하는 5'-삼부 리더(TPL) 서열을 보유한다.

[0105] **I(A). 유전자 요법 벡터 혈청형**

[0106] 아데노바이러스 중에는, 또한 50 개 초과의 혈청형이 있다. 아데노바이러스 5형은 다량의 생화학적 및 유전적 정보가 알려져 있고, 역사적으로 아데노바이러스를 벡터로서 이용하는 대부분의 구축물에 사용된 인간 아데노바이러스이다. Ad5는 유전자 요법 연구에서 광범위하게 사용되었다.

[0107] 그러나, 대부분의 인간은 Ad5 캡시드 단백질에 대해 지시된 중화 혈청 항체를 가지며, 이는 HDAd5/35 벡터, 즉, Ad5 캡시드 단백질 및 키메라 Ad35 섬유를 함유하는 벡터와 같은 Ad5 캡시드를 포함하는 아데노바이러스 벡터로 생체내 형질도입을 차단할 수 있다. Ad5 캡시드 단백질에 대해 지시된 중화 혈청 항체의 존재는 Ad5 캡시드를 포함하는 아데노바이러스 벡터의 치료적 가치를 무시화하지 않는 반면, Ad5 캡시드를 포함하지 않는 아데노바이러스 벡터는 임상적으로 유의한 면역원성 반응의 일반적인 위험이 특히 Ad5 캡시드 단백질에 대해 지시된 중화 혈청 항체가 있는 대상체에서 감소될 것이라는 점에서 추가적인 이익을 제공할 것이다.

[0108] Ad35는 57 개의 알려진 인간 혈청형 중 가장 희귀한 것 중 하나이며, 혈청 유병률이 7% 미만이고 Ad5와의 교차 반응성이 없다. Ad35는 부분적으로 Ad35 섬유 놉의 T-세포 활성화의 약화로 인해 Ad5보다 면역원성이 적다. 또한, 정맥내(iv) 주사 후, 인간 CD46 유전자이식(hCD46tg) 마우스 및 비-인간 영장류에서 간을 포함하는 조직의 최소 형질도입만이 있다(PCR에 의해서만 검출가능). 1세대 Ad35 벡터는 백신접종 목적을 위해 임상적으로 사용되었다.

[0109] **I(A)(i). Ad35 유전자 요법 벡터**

[0110] 대표적인 천연 Ad35 아데노바이러스의 완전한 게놈은 알려져 있고 공개적으로 이용가능하다(예를 들어, Gao 등, 2003 *Gene Ther.* 10(23): 1941-9; Reddy 등 2003 *Virology* 311(2): 384-393; GenBank 수탁 번호 AX049983 참조). Ad5 게놈은 35,935 bp이고 G + C 함량이 55.2%인 반면, Ad35 게놈은 34,794 bp 이고 G + C 함량이 48.9%이다. Ad35의 게놈은 도립된 말단 반복부(ITR)에 의해 플랭킹된다. 다양한 구현예에서, Ad35 ITRS는 137 bp를 포함하며(예를 들어, GenBank 수탁 번호 AX049983의 뉴클레오티드 1-137 또는 4-140을 포함하는 5' Ad35 및 GenBank 수탁 번호 AX049983의 뉴클레오티드 34658-34794를 포함하는 3' ITR), Ad5(103 bp)보다 더 길다. 다양한 구현예에서, Ad35 5' ITR은 GenBank 수탁 번호 AX049983의 뉴클레오티드 1-200의 상응하는 단편과 적어도 80% 서열 동일성(예를 들어, 적어도 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 서열 동일성)을 갖는 적어도 80 개의 뉴클레오티드(예를 들어, 적어도 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 또는 200 개의 뉴클레오티드, 예를 들어, 80, 90, 100, 110, 120, 또는 130 개 뉴클레오티드의 하한 및 130, 140, 150, 160, 170, 180, 190, 또는 200 개 뉴클레오티드의 상한, 예를 들어, 137 개의 뉴클레오티드를 갖는 다수의 뉴클레오티드)를 포함하고 Ad35 3' ITR은 GenBank 수탁 번호 AX049983의 뉴클레오티드 34595-34794의 상응하는 단편과 적어도 80% 서열 동일성(예를 들어, 적어도 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 서열 동일성)을 갖는 적어도 80 개의 뉴클레오티드(예를 들어, 적어도 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 또는 200 개의 뉴클레오티드, 예를 들어, 80, 90, 100, 110, 120, 또는 130 개 뉴클레오티드의 하한 및 130, 140, 150, 160, 170, 180, 190, 또는 200 개 뉴클레오티드의 상한, 예를 들어, 137 개의 뉴클레오티드를 갖는 다수의 뉴클레오티드)를 포함한다. 다양한 구현예에서, ITR은 Ad35 캡슐화 및/또는 복제 중 하나 또는 둘 다에 충분하다. 다양한 구현예에서, Ad35 벡터에 대한 Ad35 ITR 서열은 첫번째 8 bp가 CATCATCA보다 CTATCTAT라는 점에서 상이하다(Wunderlich, *J. Gen Viro.* 95: 1574-1584, 2014).

[0111] 다양한 구현예에서, 아데노바이러스 게놈의 패키징은 ITR에 인접한 바이러스 게놈의 5' 단부에 위치한 시스-작용 패키징 서열 도메인에 의해 매개되고, 패키징은 왼쪽에서 오른쪽으로 극성 방식으로 발생한다. Ad35의 패키징 서열은 5 내지 7 개의 추정 "A" 반복부가 있는 게놈의 왼쪽 단부에 위치한다. 다양한 구현예에서, 본 개시내용은 Ad35 패키징 서열을 포함하는 재조합 Ad35 공여자 벡터 또는 게놈을 포함한다. 다양한 구현예에서, 본 개시내용은 재조합효소 부위에 의해 플랭킹된 패키징 서열을 포함하는 재조합 Ad35 헬퍼 벡터 또는 게놈을 포함한다. 다양한 구현예에서, Ad35 패키징 서열은 GenBank 수탁 번호 AX049983의 뉴클레오티드 138-481 또는 Ad35 벡터 또는 게놈의 패키징에 충분하거나 또는 필요한 이의 단편을 포함하는 핵산 서열을 지칭한다(예를 들어, 재조합효소 부위가 있는 서열의 플랭킹 및 재조합효소 부위의 재조합에 의한 절제는 패키징을 위한 벡터 또는 게놈 결핍을 예를 들어, 패키징 서열을 포함하는 참조와 비교하여 적어도 10%, 예를 들어, 적어도 10%, 20%, 30%, 40$, 50%, 60%, 70%, 80%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 또는 100%로 만들도록 하며, 임의적으로 여기서 참조는 재조합효소 부위에 의해 플랭킹된 패키징 서열을 포함한다). 다양한 구현예에서, Ad35 패키징 서열은 GenBank 수탁 번호 AX049983의 뉴클레오티드 137-481의 상응하는 단편과 적어도 80% 서열 동일성(예를 들어, 적어도 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 서열 동일성)을 갖는 적어도 80 개의 뉴클레오티드(예를 들어, 적어도 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 225, 250, 275, 또는 300 개의 뉴클레오티드, 예를 들어, 80, 90, 100, 110, 120, 130, 140, 또는 150 개 뉴클레오티드의 하한 및 150, 160, 170, 180, 190, 200, 225, 250, 275, 또는 300 개 뉴클레오티드의 상한을 갖는 다수의 뉴클레오티드)를 포함한다.

[0112] 다양한 구현예에서, Ad35 헬퍼 벡터는 패키징 서열에 플랭크하도록 삽입된 재조합효소 부위를 포함할 수 있으며, 여기서 제1 재조합효소 부위는 뉴클레오티드 130 내지 뉴클레오티드 400 사이에서 선택된 위치(예를 들어, 뉴클레오티드 138 내지 180, 138 내지 200, 138 내지 220, 138 내지 240, 138 내지 260, 138 내지 280, 138 내지 300, 138 내지 320, 138 내지 340, 138 내지 360, 138 내지 366, 138 내지 380, 또는 138 내지 400 사이)에 바로 인접하여(예를 들어, 앞에 또는 뒤에) 삽입되고 제2 재조합효소 부위는 뉴클레오티드 300 내지 뉴클레오티드 550 사이에서 선택된 위치(예를 들어, 뉴클레오티드 344 내지 360, 344 내지 380, 344 내지 400, 344 내지 420, 344 내지 440, 344 내지 460, 344 내지 480, 344 내지 481, 344 내지 500, 344 내지 520, 344

내지 540, 또는 344 내지 550 사이)에 바로 인접하여(예를 들어, 앞에 또는 뒤에) 삽입된다. 당업자는 용어 패키징 서열이 반드시 주어진 벡터 또는 게놈에 존재하는 패키징 요소를 모두 포함하지 않음을 이해할 것이다. 예를 들어, 헬퍼 게놈은 패키징 서열에 플랭킹하는 재조합효소 지시 반복부를 포함할 수 있으며, 여기서 플랭킹된 패키징 서열은 헬퍼 게놈에 존재하는 패키징 요소를 모두 포함하지 않는다. 따라서, 특정 구현예에서, 헬퍼 게놈의 1 또는 2 개의 재조합효소 지시 반복부는 더 큰 패키징 서열 내에 위치하여, 예를 들어 더 큰 패키징 서열이 1 또는 2 개의 재조합효소 지시 반복부의 도입에 의해 인접하지 않게 되도록 한다. 다양한 구현예에서, 헬퍼 게놈의 재조합효소 지시 반복부는 재조합효소 지시 반복부의 재조합에 의한 플랭킹된 패키징 서열의 절제가 헬퍼 게놈의 패키징 및/또는 헬퍼 게놈이 패키징되는 능력을 감소시키거나 또는 제거하도록(보다 일반적으로 방해하도록) 패키징 서열의 단편에 플랭킹한다. 예로서, 재조합효소 지시 반복부(DR)는 Ad35 패키징 신호를 기능적으로 방해하지만 5' Ad35 ITR은 방해하지 않기 위해 Ad35 게놈의 5' 단부의 550 개 뉴클레오티드 내에 위치한다. 다양한 구현예에서, DR은 Ad35 패키징 신호를 기능적으로 방해하지만 5' Ad35 ITR은 방해하지 않기 위해, Ad35 게놈의 5' 단부로부터 550 개 뉴클레오티드보다 더 가깝게, 예를 들면 540, 530, 520, 510, 500, 495,490, 480, 470, 450, 440, 400, 380, 360 개 뉴클레오티드 이내에, 또는 Ad35 게놈의 5' 단부의 360 개 뉴클레오티드 이내보다 더 가깝게 위치한다.

[0113] 다양한 구현예에서, 본 개시내용은 Ad35 5' ITR, Ad35 패키징 서열, 및 Ad35 3' ITR을 포함하는 재조합 Ad35 공여자 벡터 또는 게놈을 포함한다, 특정 구현예에서, Ad35 5' ITR, Ad35 패키징 서열, 및 Ad35 3' ITR은 표준 Ad35 게놈으로부터 유래되고/되거나, 이에 대해 적어도 80% 동일성을 갖는 재조합 Ad35 공여자 벡터 또는 게놈의 유일한 단편(예를 들어, 50 개 초과 또는 100 개 초과의 염기쌍에 대한 유일한 단편)이다.

[0114] Ad35 초기 영역은 E1A, E1B, E2A, E2B, E3, 및 E4를 포함한다. Ad35 중간 영역은 pIX 및 IVa2를 포함한다. Ad35의 후기 전사 단위는 16.9 맵 단위에 위치한 주요 후기 프로모터(MLP)로부터 전사된다. Ad35에서 후기 mRNA는 이러한 mRNA에 의해 사용되는 poly(A) 신호에 따라 5 개의 mRNA 패밀리(L1-L5)로 나눌 수 있다. MLP 공통 개시인자 요소, 및 스플라스 공여자 및 스플라이스 수용체 부위 서열을 기반으로, 삼부 리더(TPL)의 길이는 204 개 뉴클레오티드인 것으로 예측되었다. MLP에 인접한 TPL의 첫번째 리더는 45 개의 뉴클레오티드 길이이다. DNA 폴리머라제의 코딩 영역 내에 위치한 두번째 리더는 72 개의 뉴클레오티드 길이이다. 세번째 리더는 E2B 영역의 전구체 말단 단백질(pTP)의 코딩 영역 내에 놓여 있고 87 개의 뉴클레오티드 길이이다. Ad5는 2 개의 바이러스-연관(VA) RNA 유전자를 함유하지만, 하나의 바이러스-연관 RNA 유전자만이 Ad35의 게놈에서 발생한다. 이 VA RNA 유전자는 52/55K L1 단백질을 코딩하는 유전자 및 pTP 사이에 위치한다.

[0115] 특정 구현예에서, Ad35++ 벡터는 돌연변이체 Ad35 섬유 놉이 있는 키메라 벡터(예를 들어, 돌연변이체 Ad35 섬유 놉이 있는 재조합 Ad35 벡터 또는 돌연변이체 Ad35 섬유 놉이 있는 Ad5/35 벡터)이다. 특정 구현예에서, Ad35++ 게놈은 돌연변이체 Ad35 섬유 놉을 암호화하는 게놈(예를 들어, 돌연변이체 Ad35 섬유 놉을 암호화하는 재조합 Ad35 헬퍼 게놈 또는 돌연변이체 Ad35 섬유 놉을 암호화하는 Ad5/35 헬퍼 게놈)이다. 다양한 구현예에서, Ad35++ 돌연변이체 섬유 놉은 CD46에 대한 친화도가, 예를 들어, 25-배 증가하도록 돌연변이된 Ad35 섬유 놉이며, 예를 들어, Ad35++ 돌연변이체 섬유 놉은 예를 들어, 더 낮은 감염 다중도(MOI)에서 세포 형질도입 효율을 증가시키도록 한다(Li and Lieber, *FEBS Letters*, 593(24): 3623-3648, 2019).

[0116] 다양한 구현예에서, Ad35++ 돌연변이체 섬유 놉은 Ile192Val, Asp207Gly(또는 특정 Ad35 서열에서 Glu207Gly), Asn217Asp, Thr226Ala, Thr245Ala, Thr254Pro, Ile256Leu, Ile256Val, Arg259Cys, 및 Arg279His로부터 선택된 적어도 하나의 돌연변이를 포함한다. 다양한 구현예에서, Ad35++ 돌연변이체 섬유 놉은 다음 돌연변이 각각을 포함한다: Ile192Val, Asp207Gly(또는 특정 Ad35 서열에서 Glu207Gly), Asn217Asp, Thr226Ala, Thr245Ala, Thr254Pro, Ile256Leu, Ile256Val, Arg259Cys, 및 Arg279His. 다양한 구현예에서, Ad35 섬유의 아미노산 넘버링은 GenBank 수탁 AP_000601 또는 이에 상응하는 아미노산 서열에 따르며, 예를 들어, 여기서 위치 207은 Glu 또는 Asp이다. 다양한 구현예에서, Ad35 섬유는 GenBank 수탁 AP_000601에 따른 아미노산 서열을 갖는다. Ad35++ 섬유 놉 돌연변이에 대한 추가의 상세한 내용은 Wang 2008 *J. Virol.* 82(21): 10567-10579에 기재되어 있으며, 섬유 놉과 관련하여 그 전문이 본원에 참조로 포함된다.

[0117] **I(A)(ii). Ad5/35 유전자 요법 벡터**

[0118] 본 개시내용의 Ad5/35 벡터는 Ad5 캡시드 폴리뉴클레오티드 및 Ad35 섬유 놉을 포함하는 키메라 섬유 폴리뉴클레오티드를 포함하는 아데노바이러스 벡터를 포함하며, 키메라 섬유 폴리뉴클레오티드는 전형적으로 또한 Ad35 섬유 샤프트(예를 들어, Ad35 섬유 아미노산 44-323과 조합된 Ad5 섬유 아미노산 1-44)를 포함한다. 다양한 구현예에서, 섬유는 Ad35++ 돌연변이체 섬유 놉을 포함한다. 본 개시내용의 다양한 Ad5/35 벡터에서, 섬유 놉 도

메인 및 샤프트를 제외한 모든 단백질은 혈청형 5로부터 유래된 반면, 섬유 놉 도메인 및 샤프트는 혈청형 35로부터 유래되었고, CD46에 대한 친화도가 증가된 돌연변이는 Ad35 섬유 놉 내로 도입되었다(WO 2010/120541 A2 참조). 추가적으로, 다양한 구현예에서, Ad5/35 벡터의 ITR 및 패키징 서열은 Ad5로부터 유래된다. (예시적인 놉 돌연변이에 대해 표 1; 및 HDAd35 벡터 생산의 일반적인 도식에 대해 도 95 참조.)

[0119] **표 1: CD46에 대한 결합이 증가된 돌연변이된 Ad35 놉**

| | Kd (Olek) | |
|---|---|---|
| A1: Asn217Asp Thr245Pro Ile256Leu* | A1 4.82 nM | Asp207Gly +++<br>Thr245Ala ++<br>Ile256Leu + |
| A2: *Asp207Gly Thr245Ala** | A2 0.629 nM | |
| A3: *Asp207Gly Thr226Ala** | A3 1.407 nM | |
| A8: Ile192Val Ile256Val ? | A8 13.6850 nM | |
| B1: Asp207Gly* | B1 1.774 nM | |
| B2: wtAd35(207Asp) | B2 14.98 nM | |
| B3: Asn217Asp* | B3 16.85 nM | |
| B4: Thr245Ala* | B4 7.64 nM | |
| B5: Ile256Leu* | B5 10.96 nM | |
| B6: Ad3 | B6 결합 없음 | |
| B7: Ad11 | B7 11.22 nM | |
| M1: Arg279Cys** | M1 결합 없음 | |
| M3:Arg279His** | M3 결합 없음 | |
| wtAd35* | 13.7 nM | |
| wtAd35* | 15.36 nM | |
| AA: *Asp207Gly Thr245Ala Ile256Leu** | 0.943 nM | |

* Wang *et al. (J. Virol.*, 82(21):10567-10579, 2008)에 공개됨

** Wang *et al. (J. Virol.* 81(23):12785-12792, 2007)에 공개됨

[0120]

[0121] **I(B). 헬퍼-의존적 Ad35 및 Ad5/35 벡터**

[0122] 일반적으로, 천연 아데노바이러스 벡터에서 헬퍼-의존적 아데노바이러스 벡터로의 경로는 3 가지 세대를 포함한다. 1세대 아데노바이러스 벡터는 유전자 E1 및 E3을 제거하도록 조작된다. 이들 유전자 없이, 아데노바이러스 벡터는 자체적으로 복제할 수 없지만 HEK293 세포와 같은 E1-발현 포유류 세포주에서 생성될 수 있다. 1세대 변형만으로, 아데노바이러스 벡터 클로닝 용량은 제한되고, 벡터에 대한 숙주 면역 반응은 효과적인 페이로드 발현에 문제가 될 수 있다. 2세대 아데노바이러스 벡터는 E1/E3 제거 이외에도, 비-구조적 유전자 E2 및 E4를 제거하도록 조작되어, 용량 증가 및 면역원성 감소를 초래한다. 3세대 아데노바이러스 벡터(또한 무기력한, 고용량 아데노바이러스 벡터, 또는 헬퍼-의존적 아데노바이러스 벡터(HdAd)로도 지칭됨)는 모든 바이러스 코딩 서열을 제거하고, 게놈의 ITR 및 게놈 또는 이의 기능적 단편의 패키징 서열만을 유지하도록 추가로 조작된다. 이들 게놈은 바이러스 생산에 필요한 단백질을 암호화하지 않기 때문에, 헬퍼-의존적이다: 헬퍼-의존적 게놈은 바이러스 단백질을 트랜스로 제공하는 핵산 서열을 포함하는 세포에 존재하는 경우에만 벡터 내로 패키징될 수 있다. 이들 헬퍼-의존적 벡터는 또한 여전히 더 큰 용량 및 추가로 감소된 면역원성을 특징으로 한다. 각 바이러스 게놈의 서열은 적어도 각 혈청형에 대해 구별되기 때문에, 주어진 혈청형에 대한 헬퍼-의존적 바이러스 게놈, 및/또는 헬퍼 게놈을 생성하는 데 필요한 적절한 변형은 다른 혈청형과 관련한 이용가능한 정보로부터 예측될 수 없다.

[0123] 모든 바이러스 코딩 서열이 결여되도록 조작된 헬퍼-의존적 아데노바이러스 벡터(HDAd)는 광범위한 세포 유형을 효율적으로 형질도입할 수 있고, 무시할 수 있는 만성 독성으로 장기간 이식유전자 발현을 매개할 수 있다. 바이러스 코딩 서열을 결실시키고 게놈 복제(ITR) 및 캡슐화(Ψ)에 필요한 시스-작용 요소만을 남겨둠으로써, Ad 벡터에 대한 세포 면역 반응은 감소된다. HDAd 벡터는 최대 37 kb의 큰 클로닝 용량을 가져, 큰 페이로드의 전달을 허용한다. 이들 페이로드는 이식유전자 발현을 향상시키고, 연장하고, 조절하도록 큰 치료 유전자 또는 심지어 다중 이식유전자 및 큰 조절 구성요소를 포함할 수 있다. 다른 아데노바이러스 벡터와 마찬가지로, 전형적인 HDAd 게놈은 일반적으로 에피솜으로 남아있고 숙주 게놈과 통합하지 않는다(Rosewell 등, *J Genet Syndr Gene Ther. Suppl* 5:001, 2011, doi: 10.4172/2157-7412.s5-001).

[0124] 일부 HDAd 벡터 시스템에서, 하나의 바이러스 게놈(헬퍼 게놈)은 복제에 필요한 단백질을 모두 암호화하지만 패키징 서열에 조건부 결함이 있어서, 비리온 내로 패키징될 가능성이 적다. 상기 언급된 바와 같이, 이는 패키징 서열 또는 기능적으로 기여하는(예를 들어, 기능적으로 필요한) 이의 단편의 식별 및 다른 아데노바이러스 혈청형과 관련하여 기존 지식에서 확인될 수 없는 헬퍼 벡터의 증식을 무효화하지 않는 방식으로 대상체 게놈의 변형을 필요로 할 수 있다, 개별 공여자 바이러스 게놈은 바이러스 ITR, 페이로드(예를 들어, 치료적

페이로드), 및 기능적 패키징 서열(예를 들어, 정상 야생형 패키징 서열, 또는 이의 기능적 단편)을 포함하여(예를 들어, 유일하게 포함하여), 이 공여자 바이러스 게놈이 HDAd 바이러스 벡터 내로 선택적으로 패키징되고 생산자 세포로부터 단리되게 한다. HDAd 공여자 벡터는 물리적 수단에 의해 헬퍼 벡터로부터 추가로 정제될 수 있다. 일반적으로, HDAd 바이러스 벡터 및 HDAd 바이러스 벡터 제형에서 헬퍼 벡터 및/또는 헬퍼 게놈의 일부 오염이 발생할 수 있고 허용될 수 있다.

[0125] 일부 HDAd 벡터 시스템에서, 헬퍼 게놈은 Cre/loxP 시스템을 활용한다. 이러한 특정 HDAd 벡터 시스템에서, HDAd 공여자 게놈은 게놈 복제에 필요한 아데노바이러스 ITR, 및 게놈을 캡시드 내로 캡슐화하는 데 필요한 패키징 서열 또는 이의 기능적 단편인 Ψ를 포함하는 비코딩 아데노바이러스 DNA의 500 bp를 포함한다. 또한 HDAd 공여자 벡터 게놈은 총 길이가 27.7 kb 내지 37 kb일 때 가장 효율적으로 패키징될 수 있음이 관찰되었으며, 예를 들어, 치료적 페이로드 및/또는 "스터퍼" 서열로 구성될 수 있다. HDAd 공여자 게놈은 Cre 재조합효소를 발현하는 293 세포(HEK293)와 같은 세포로 전달될 수 있으며, 임의적으로 여기서 HDAd 공여자 게놈은 비-바이러스 벡터 형태, 예컨대 박테리아 플라스미드 형태로 세포에 전달된다(예를 들어, 여기서 HDAd 공여자 게놈은 박테리아 플라스미드(pHDAd)로 구축되고 제한 효소 소화로 유리된다). 동일한 세포는 헬퍼 게놈으로 형질도입될 수 있으며, 이는 Cre 재조합효소를 발현하는 293 세포의 감염 후, 패키징 서열 또는 기능적으로 기여하는(예를 들어, 기능적으로 필요한) 이의 단편이 loxP 부위 사이의 Cre-매개 부위-특이적 재조합에 의해 헬퍼 게놈으로부터 절제되도록 loxP 부위에 의해 플랭킹된 패키징 서열 또는 기능적으로 기여하는(예를 들어, 기능적으로 필요한) 이의 단편을 보유하는 E1-결실된 Ad 벡터를 포함한다. 따라서, HDAd 공여자 게놈은 Cre를 발현하는 293 세포(HEK293) 내로 형질감염될 수 있고 패키징 서열(Ψ) 또는 재조합효소 부위(예를 들어, loxP 부위)에 의해 플랭킹된 이의 기능적 단편을 보유하는 헬퍼 게놈으로 형질도입되어 Ψ의 상응하는 재조합효소에 의해 매개된 절제(예를 들어, Cre-매개 절제)가 헬퍼 바이러스 게놈을 패키징할 수 없게 만들지만, 여전히 HDAd의 증식을 위해 필요한 트랜스-작용 인자를 모두 제공할 수 있도록 한다. 패키징 서열 또는 기능적으로 기여하는(예를 들어, 기능적으로 필요한) 이의 단편의 절제 후, 헬퍼 게놈은 패키징할 수 없지만 여전히 DNA 복제를 겪을 수 있고 따라서 HDAd 공여자 게놈의 복제 및 캡슐화를 트랜스-보완할 수 있다. 일부 구현예에서, 293 세포(HEK293)에 존재하는 헬퍼 및 HDAd 공여자 게놈 사이의 상동 재조합의 결과로서 복제 적격 Ad(RCA; $E1^+$)의 생성을 방지하기 위해 "스터퍼" 서열은 E3 영역 내로 삽입되어 임의의 $E1^+$ 재조합체를 패키징하기에 너무 크게 만들 수 있다. 유사한 HDAd 생산 시스템은 FLP(예를 들어, FLPe)/frt 부위-특이적 재조합을 사용하여 개발되었으며, 여기서 헬퍼 게놈의 패키징 서열에 플랭킹 하는 frt 부위 사이의 FLP-매개 재조합은 FLP를 발현하는 293 세포(HEK293)에서 헬퍼 게놈의 캡슐화에 대해 선택한다. 헬퍼 벡터에 대해 선택하기 위한 대안적인 전략이 개발되었다. Ad35 헬퍼 바이러스는 전형적으로 E1 발현 산물이 생산자 세포주의 게놈으로부터 상보적 발현에 의해 공급될 수 있을 때, E1에서와 같은 것들을 제외한 모든 바이러스 유전자를 포함한다.

[0126] HDAd5/35 공여자 벡터, 공여자 게놈, 헬퍼 벡터 및 헬퍼 게놈은 본원에 제공되고 본 개시내용의 다양한 방법에 사용되는 조성물의 예시이다. HDAd5/35 벡터 또는 게놈은 Ad35 섬유 놉 및 Ad5 샤프트가 있는 헬퍼-의존적 키메라 Ad5/35 벡터 또는 게놈이다. HDAd5/35++ 벡터 또는 게놈은 돌연변이체 Ad35 섬유 놉이 있는 헬퍼-의존적 키메라 Ad5/35 벡터 또는 게놈이다. 벡터는 CD46에 대한 친화도가 예를 들어, 25-배 증가하도록 돌연변이되고 더 낮은 감염 다중도(MOI)에서 세포 형질도입 효율을 증가시킨다(Li & Lieber, *FEBS Letters*, 593(24): 3623-3648, 2019). Ad5/35 헬퍼 벡터는 조건부로 발현된(예를 들어, frt-부위 또는 loxP-부위 플랭킹된) 패키징 서열을 포함하고 공여자 게놈이 패키징될 수 있는 Ad5/35 비리온의 생산을 위해 필요한 트랜스-작용 인자를 모두 암호화하는 헬퍼 게놈을 포함하는 벡터이다.

[0127] HDAd35 공여자 벡터, 공여자 게놈, 헬퍼 벡터 및 헬퍼 게놈은 또한 본원에 제공되고 본 개시내용의 다양한 방법에 사용되는 조성물의 예시이다. HDAd35 벡터 또는 게놈은 헬퍼-의존적 Ad35 벡터 또는 게놈이다. HDAd35++ 벡터 또는 게놈은 CD46에 대한 친화도를 향상시키고 세포 형질도입 효율을 증가시키는 돌연변이체 Ad35 섬유 놉이 있는 헬퍼-의존적 Ad35 벡터 또는 게놈이다. Ad35 헬퍼 벡터는 조건부로 발현된(예를 들어, frt-부위 또는 loxP-부위 플랭킹된) 패키징 서열을 포함하고 공여자 게놈이 패키징될 수 있는 Ad35 비리온의 생산을 위해 필요한 트랜스-작용 인자를 모두 암호화하는 헬퍼 게놈을 포함하는 벡터이다. 본 개시내용은 HDAd35 공여자 게놈 및 Ad35 헬퍼 게놈을 포함하는 세포를 포함하는 HDAd35 공여자 벡터 생산 시스템을 추가로 포함한다. 이러한 특정 세포에서, 헬퍼 게놈에 의해 암호화되고 발현되는 바이러스 단백질은 HDAd35 공여자 게놈이 패키징되는 HDAd35 공여자 벡터의 생산에 활용될 수 있다. 따라서, 본 개시내용은 HDAd35 공여자 게놈 및 Ad35 헬퍼 게놈을 포함하는 세포를 배양함으로써 HDAd35 공여자 벡터를 생산하는 방법을 포함한다. 일부 구현예에서 세포는 Ad35 헬퍼 벡터의 패키징 서열에 플랭킹하는 재조합효소 지시 반복부에 상응하는 재조합효소를 암호화하고 발현

한다. 일부 구현예에서, Ad35 헬퍼 게놈의 플랭킹된 패키징 서열은 절제되었다.

[0128] 일부 구현예에서 Ad35 헬퍼 게놈은 모든 Ad35 코딩 서열을 암호화한다. 일부 구현예에서 Ad35 헬퍼 게놈은 E1 영역의 하나 이상의 코딩 서열 및/또는 E3 코딩 서열 및/또는 E4 코딩 서열을 제외한 모든 Ad35 코딩 서열을 암호화하고/하거나 발현한다. 다양한 구현예에서, Ad35 E1 유전자를 암호화하고/하거나 발현하지 않는 헬퍼 게놈은 Ad35 E4 유전자를 암호화하고/하거나 발현하지 않으며, 임의적으로 여기서 Ad35 헬퍼 게놈은 Ad5 E4orf6 코딩 서열을 포함하도록 추가로 조작된다. 다양한 구현예에서, 당업자에 의해 이해되는 바와 같이, HDAd 35 공여자 벡터의 생산을 위한 조성물 및 방법의 세포는 Ad5 E1 발현 산물을 발현하는 세포일 수 있다. 다양한 구현예에서, 당업자에 의해 이해되는 바와 같이, HDAd 35 공여자 벡터의 생산을 위한 조성물 및 방법의 세포는 293 T 세포(HEK293)일 수 있다.

[0129] 헬퍼는 야생형 또는 증식-적격 Ad5 벡터 또는 Ad35 벡터와 같은 야생형 또는 유사한 증식-적격 벡터로부터 조작될 수 있다. 당업자가 이해하는 바와 같이, 헬퍼 벡터의 조작에 사용될 수 있는 한 가지 전략은 E1 유전자 발현의 결실 또는 다른 기능적 파괴이다. 아데노바이러스 게놈의 5' 부분에 위치한 E1 영역은 초기 및 후기 유전자의 야생형 발현에 필요한 단백질을 암호화한다. E1 결실은 E1에 의해 제어된 특정 바이러스 유저자의 발현을 감소시키거나 또는 제거하고, E1-결실된 헬퍼 바이러스는 복제-결핍이다. 따라서, E1-결핍 헬퍼 바이러스는 E1을 발현하는 세포주를 사용하여 증식될 수 있다. 예를 들어, E1-결핍 Ad35 헬퍼 벡터가 Ad5 E4orf6을 암호화도록 조작되는 경우, 헬퍼 벡터는 Ad5 E1을 발현하는 세포주에서 증식될 수 있고, E1-결핍 Ad35 헬퍼 벡터가 Ad5 E4orf6을 암호화하는 경우, 헬퍼 벡터는 Ad5 E1을 발현하는 세포주에서 증식될 수 있다. HDAd35 벡터 생산을 위한 하나의 예시적인 세포 유형에서, HEK293 세포는 Ad5 E4 단백질 ORF6과 복합체를 형성하는 것으로 알려진 Ad5 E1b55k를 발현한다. 표 2는 Ad35 게놈에 의해 암호화된 발현 산물의 예시적인 요약을 제공한다(Gao, *Gene Ther.* 10:1941-1949, 2003 참조).

[0130] **표 2. Ad35 게놈의 예측된 번역 특징.**

| 특징 | | 시작 | 종료 |
|---|---|---|---|
| *E1 및 pIX 영역* | | | |
| E1A 261R | | 569 | 1148 |
| | 연결 | 1233 | 1441 |
| E1A 230R | | 569 | 1055 |
| | 연결 | 1233 | 1441 |
| E1A 58R | | 569 | 640 |
| | 연결 | 1233 | 1337 |
| E1B 214R (작은 T 항원) | | 1611 | 2153 |
| E1B 494R (큰 T 항원) | | 1916 | 3400 |
| pIX | | 3484 | 3903 |
| ORF-1 | | 2366 | 2689 |
| *E2 및 IVa2 영역 (상보적 가닥)* | | | |
| IVa2 | | 5579 | 5590 |
| | 연결 | 3966 | 5300 |
| E2B DNA pol | | 5069 | 8437 |
| E2B pTP | | 8440 | 10 356 |
| E2A DBP | | 22 414 | 23 415 |
| ORF-2 | | 5988 | 6482 |
| ORF-3 | | 7847 | 8257 |
| ORF-4 | | 15 663 | 15 971 |
| ORF-5 | | 15 743 | 16 216 |
| ORF-6 | | 16 457 | 17 041 |
| ORF-7 | | 17 543 | 17 938 |
| ORF-8 | | 17 994 | 18 713 |
| ORF-9 | | 21 858 | 22 436 |
| ORF-10 | | 22 128 | 22 502 |
| ORF-11 | | 23 027 | 23 488 |
| *E3 영역* | | | |
| E3 12.2K 단백질 | | 27 198 | 27 515 |
| E3 15.0 K 단백질 | | 27 469 | 27 864 |
| E3 18.5K 단백질 | | 27 849 | 28 349 |
| E3 20.3K 단백질 | | 28 369 | 28 914 |
| E3 20.6K 단백질 | | 28 932 | 29 495 |
| E3 15.2K 단백질 | | 29 817 | 30 221 |
| E3 15.3K 단백질 | | 30 214 | 30 621 |
| ORF-12 | | 25 693 | 26 019 |
| ORF-13 | | 27 908 | 28 240 |
| *E4 영역 (상보적 가닥)* | | | |
| E4 299R | | 32 075 | 32 974 |
| E4 145R | | 33 604 | 34 041 |
| E4 125R | | 34 038 | 34 415 |
| E4 117R | | 33 254 | 33 607 |
| E4 122R | | 32 877 | 33 245 |

[0131]

| 특징 | | 시작 | 종료 |
|---|---|---|---|
| ORF-14 | | 33 100 | 33 609 |
| *VA RNA 영역* | | 10 433 | 10 594 |
| *L 영역* | | | |
| L1 52,55K | | 10 653 | 11 819 |
| L1 IIIa | | 11 845 | 13 608 |
| L2 III (펜톤 염기) | | 13 690 | 15 375 |
| L2 pVII | | 15 383 | 15 961 |
| L2 V | | 16 004 | 17 059 |
| L3 pVI | | 17 399 | 18 139 |
| L3 II (헥손) | | 18 255 | 21 113 |
| L3 23K (프로테아제) | | 21 150 | 21 779 |
| L4 100K | | 23 446 | 25 884 |
| L4 22K | | 25 616 | 26 191 |
| L4 33K | | 25 616 | 25 934 |
| | 연결 | 26 104 | 26 465 |
| L4 pVIII | | 26 515 | 27 198 |
| L5 IV(섬유) | | 30 826 | 31 797 |

[0132]

[0133] 본 개시내용은 무엇보다도, 예를 들어, 2 개의 Ad35 ITR이 페이로드에 플랭킹되는 Ad35 ITR(예를 들어, 5' Ad35 ITR 및 3' ITR)을 포함하는 HDAd35 공여자 벡터 및 게놈을 포함한다. 본 개시내용은 무엇보다도, Ad35 패키징 서열 또는 이의 기능적 단편을 포함하는 HDAd35 공여자 벡터 및 게놈을 포함한다. 본 개시내용은 무엇보다도,

E1 또는 이의 단편이 결실된(예를 들어, E1 결실이 GenBank 수탁 번호 AX049983의 뉴클레오티드 481-3112 또는 본원에 제공된 또 다른 Ad35 벡터 서열의 상응하는 위치의 결실을 포함하는 경우) HDAd35 공여자 벡터 및 게놈을 포함한다. 본 개시내용은 무엇보다도, E3 또는 이의 단편이 결실된(예를 들어, E3 결실이 GenBank 수탁 번호 AX049983의 뉴클레오티드 27609 내지 30402 또는 27435-30542 또는 본원에 제공된 또 다른 Ad35 벡터 서열의 상응하는 위치의 결실을 포함하는 경우) HDAd35 벡터 및 게놈을 포함한다.

[0134] 본 개시내용은 무엇보다도, 패키징 서열 또는 기능적으로 기여하는(예를 들어, 기능적으로 필요한) 이의 단편에 플랭킹하는 2 개의 재조합 부위 요소를 포함하는 Ad35 헬퍼 벡터 및 게놈을 포함하며, 각각의 재조합 부위 요소는 재조합 부위를 포함하며, 여기서 2 개의 재조합 부위는 동일한 재조합효소에 대한 부위이다. 상기 언급된 바와 같이, Ad35 헬퍼 벡터의 구축은 다른 벡터와 관련하여 기존 지식으로부터 예측가능하게 조작될 수 없다. 이에 반해, Ad35의 관련 서열은 예를 들어, Ad5의 상응하는 서열과 매우 상이하다(예를 들어, Ad35 및 Ad5의 5' 600 내지 620 개의 뉴클레오티드 비교). 더욱이, 패키징 서열은 혈청형-특이적이다. Ad35 패키징 서열은 적어도 Ad5 패키징 단일 서열 AI, AII, AIII, AIV, 및 AV에 상응하는 서열을 포함한다. 따라서, Ad35 헬퍼 벡터의 생산은 다음을 포함하여 여러 예측할 수 없는 결정을 필요로 한다: (1) 서열 유사성이 제한적인 경우 간단하지 않은, 재조합효소 부위 요소를 대상체 게놈 내로 삽입함으로써 재조합효소 부위(예를 들어, loxP 부위)에 의해 플랭킹되는 Ad35 패키징 서열 또는 기능적으로 기여하는(예를 들어, 기능적으로 필요한) 이의 단편의 식별; (2) 예측될 수 없는, (패키징 서열 또는 기능적으로 기여하는(예를 들어, 기능적으로 필요한) 이의 단편이 절제되지 않은 조건 하에) 헬퍼 벡터의 증식을 무효화하지 않는 재조합효소 부위 요소 삽입의 식별; 및/또는 (3) (예를 들어, 116 세포주와 같은 cre 재조합효소-발현 세포주에서) HDAd35 공여자 벡터의 생산 동안 헬퍼 바이러스 패키징을 감소시키면서 패키징 서열 또는 기능적으로 기여하는(예를 들어, 기능적으로 필요한) 이의 단편의 효율적인 결실을 허용하는 재조합 부위 요소 사이의 간격 식별.

[0135] 본 개시내용은 (1) 적어도 플랭킹된 서열의 절제를 야기하는 loxP 부위의 재조합이 벡터의 증식을 예를 들어, 적어도 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 또는 100%로 감소(예를 들어, 벡터의 증식을 20%, 30%, 40%, 50%, 60%, 70%의 하한, 및 60%, 70%, 80%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 또는 100%의 상한을 갖는 백분율로 감소)시킨다는 점에서 Ad35 패키징 서열의 기능적으로 기여하거나 또는 기능적으로 필요한 단편에 플랭킹하는 loxP 부위를 포함하는 복수의 예시적인 Ad35 헬퍼 벡터 및 게놈을 포함하며, 임의적으로 여기서 퍼센트 증식은 비슷한 조건 하에 완전한 벡터(재조합효소 부위-플랭킹된 서열이 절제되지 않음) 또는 야생형 Ad35 벡터와 비교하여 절제된 벡터(절제된 재조합효소 부위-플랭킹된 서열)의 증식에 의해 생산된 바이러스 입자의 수로 측정된다.

[0136] 적어도 하나의 예시적인 Ad35 헬퍼 벡터에서, 재조합효소 부위 요소(예를 들어, loxP 요소)는 뉴클레오티드 178 뒤에 삽입되고 재조합효소 부위 요소(예를 들어, loxP 요소)는 뉴클레오티드 437 뒤에 삽입된다. loxP-플랭킹된 서열의 절제는 패키징 서열 서열 AI 내지 AIV를 제거한다. 이러한 특정 구현예에서, 뉴클레오티드 345-3113의 결실은 E1 유전자 뿐만 아니라 패키징 단일 서열 AVI 및 AVII을 제거한다. 따라서, 플랭킹된 패키징 서열 또는 이의 단편은 위치 179-344에 상응한다. 이 설명에 따른 벡터는 증식되는 것으로 나타났다.

[0137] 적어도 하나의 예시적인 Ad35 헬퍼 벡터에서, 재조합효소 부위 요소(예를 들어, loxP 요소)는 뉴클레오티드 178 뒤에 삽입되고 재조합효소 부위 요소(예를 들어, loxP 요소)는 뉴클레오티드 481 뒤에 삽입되며, 여기서 뉴클레오티드 179-365는 결실된다(패키징 서열 서열 AI 내지 AV를 제거하여, 나머지 서열 AVI 및 AVII이 재조합효소 부위 요소에 의해 플랭킹된 핵산 서열에 있도록 한다. 이러한 특정 구현예에서, 뉴클레오티드 482-3113의 결실은 E1 유전자를 제거한다. 따라서, 플랭킹된 패키징 서열 또는 이의 단편은 위치 366-481에 상응한다. 이 설명에 따른 벡터는 증식되는 것으로 나타났다.

[0138] 적어도 하나의 예시적인 Ad35 헬퍼 벡터에서, 재조합효소 부위 요소(예를 들어, loxP 요소)는 뉴클레오티드 154 다음에 삽입되고 재조합효소 부위 요소(예를 들어, loxP 요소)는 뉴클레오티드 481 다음에 삽입된다, 이러한 특정 구현예에서, 뉴클레오티드 482-3113의 결실은 E1 유전자를 제거한다. 따라서, 플랭킹된 패키징 서열 또는 이의 단편은 위치 155-481에 상응한다. 이 설명에 따른 벡터는 증식되는 것으로 나타났다.

[0139] 적어도 하나의 예시적인 Ad35 헬퍼 벡터에서, 재조합효소 부위 요소(예를 들어, loxP 요소)는 뉴클레오티드 158 뒤에 삽입되고 재조합효소 부위 요소(예를 들어, loxP 요소)는 뉴클레오티드 480 뒤에 삽입된다. 이 설명에 따른 벡터는 증식되는 것으로 나타났다. 이러한 특정 구현예에서, E3 영역을 포함하는 뉴클레오티드 27388-30402는 결실된다. 특정 구현예에서, 벡터는 $Ad35^{++}$ 벡터이다.

[0140] 적어도 하나의 예시적인 Ad35 헬퍼 벡터에서, 재조합효소 부위 요소(예를 들어, loxP 요소)는 뉴클레오티드 158

뒤에 삽입되고 재조합효소 부위 요소(예를 들어, loxP 요소)는 뉴클레오티드 446 뒤에 삽입된다. 이 설명에 따른 벡터는 증식되는 것으로 나타났다. 이러한 특정 구현예에서, E3 영역을 포함하는 뉴클레오티드 27388-30402는 결실된다. 특정 구현예에서, 벡터는 $Ad35^{++}$ 벡터이다.

[0141] 적어도 하나의 예시적인 Ad35 헬퍼 벡터에서, 재조합효소 부위 요소(예를 들어, loxP 요소)는 뉴클레오티드 179 뒤에 삽입되고 재조합효소 부위 요소(예를 들어, loxP 요소)는 뉴클레오티드 480 뒤에 삽입된다. 이 설명에 따른 벡터는 증식되는 것으로 나타났다. 이러한 특정 구현예에서, E3 영역을 포함하는 뉴클레오티드 27388-30402는 결실된다. 특정 구현예에서, 벡터는 $Ad35^{++}$ 벡터이다.

[0142] 적어도 하나의 예시적인 Ad35 헬퍼 벡터에서, 재조합효소 부위 요소(예를 들어, loxP 요소)는 뉴클레오티드 206 뒤에 삽입되고 재조합효소 부위 요소(예를 들어, loxP 요소)는 뉴클레오티드 480 뒤에 삽입된다. 이 설명에 따른 벡터는 증식되는 것으로 나타났다. 이러한 특정 구현예에서, E3 영역을 포함하는 뉴클레오티드 27,388-30,402는 결실된다. 특정 구현예에서, 뉴클레오티드 27,607-30,409 또는 27,609-30,402는 결실된다. 특정 구현예에서, 뉴클레오티드 27,240-27,608은 결실되지 않는다.

[0143] 적어도 하나의 예시적인 Ad35 헬퍼 벡터에서, 재조합효소 부위 요소(예를 들어, loxP 요소)는 뉴클레오티드 139 뒤에 삽입되고 재조합효소 부위 요소(예를 들어, loxP 요소)는 뉴클레오티드 446 뒤에 삽입된다. 이러한 특정 구현예에서, 뉴클레오티드 27609-30402는 결실된다.

[0144] 적어도 하나의 예시적인 Ad35 헬퍼 벡터에서, 재조합효소 부위 요소(예를 들어, loxP 요소)는 뉴클레오티드 158 뒤에 삽입되고 재조합효소 부위 요소(예를 들어, loxP 요소)는 뉴클레오티드 446 뒤에 삽입된다. 이러한 특정 구현예에서, 뉴클레오티드 27609-30402는 결실된다.

[0145] 적어도 하나의 예시적인 Ad35 헬퍼 벡터에서, 재조합효소 부위 요소(예를 들어, loxP 요소)는 뉴클레오티드 179 뒤에 삽입되고 재조합효소 부위 요소(예를 들어, loxP 요소)는 뉴클레오티드 446 뒤에 삽입된다. 이러한 특정 구현예에서, 뉴클레오티드 27609-30402는 결실된다.

[0146] 적어도 하나의 예시적인 Ad35 헬퍼 벡터에서, 재조합효소 부위 요소(예를 들어, loxP 요소)는 뉴클레오티드 201 뒤에 삽입되고 재조합효소 부위 요소(예를 들어, loxP 요소)는 뉴클레오티드 446 뒤에 삽입된다. 이러한 특정 구현예에서, 뉴클레오티드 27609-30402는 결실된다.

[0147] 적어도 하나의 예시적인 Ad35 헬퍼 벡터에서, 재조합효소 부위 요소(예를 들어, loxP 요소)는 뉴클레오티드 158 뒤에 삽입되고 재조합효소 부위 요소(예를 들어, loxP 요소)는 뉴클레오티드 481 뒤에 삽입된다. 이러한 특정 구현예에서, 뉴클레오티드 27609-30402는 결실된다.

[0148] 적어도 하나의 예시적인 Ad35 헬퍼 벡터에서, 재조합효소 부위 요소(예를 들어, loxP 요소)는 뉴클레오티드 179 뒤에 삽입되고 재조합효소 부위 요소(예를 들어, loxP 요소)는 뉴클레오티드 384 뒤에 삽입된다. 이러한 특정 구현예에서, 뉴클레오티드 27609-30402는 결실된다.

[0149] 적어도 하나의 예시적인 Ad35 헬퍼 벡터에서, 재조합효소 부위 요소(예를 들어, loxP 요소)는 뉴클레오티드 179 뒤에 삽입되고 재조합효소 부위 요소(예를 들어, loxP 요소)는 뉴클레오티드 481 뒤에 삽입된다. 이러한 특정 구현예에서, 뉴클레오티드 27609-30402는 결실된다.

[0150] 적어도 하나의 예시적인 Ad35 헬퍼 벡터에서, 재조합효소 부위 요소(예를 들어, loxP 요소)는 뉴클레오티드 206 뒤에 삽입되고 재조합효소 부위 요소(예를 들어, loxP 요소)는 뉴클레오티드 481 뒤에 삽입된다. 이러한 특정 구현예에서, 뉴클레오티드 27609-30402는 결실된다.

[0151] 추가적인 임의적 조작 고려사항은 원심분리, 예를 들어, CsCl 초원심분리에 의해 HDAd35 공여자 벡터로부터 헬퍼 벡터의 분리를 허용하는 크기를 갖는 헬퍼 게놈의 조작일 수 있다. 이 결과를 달성하기 위한 한 가지 방식은 헬퍼 게놈의 크기를 34,794 bp의 야생형 길이를 갖는 전형적인 Ad35 게놈과 비교하여 증가시키는 것이다. 특히, 아데노바이러스 게놈은 야생형 길이의 적어도 104%까지 조작함으로써 증가될 수 있다. 본 개시내용의 특정 헬퍼 벡터는 Ad35 E1 영역 및 E4 영역을 포함하고, E3 영역이 결실되고, 페이로드 및/또는 스터퍼 서열을 수용할 수 있다.

[0152] Ad35 헬퍼 벡터는 Ad35 공여자 벡터의 생산에 사용될 수 있다. HDAd35++ 벡터의 생산은 HDAd 벡터 게놈을 함유하는 플라스미드 및 구조적 및 비-구조적 바이러스 단백질을 제공하는 패키징-결함 헬퍼 바이러스의 공동-형질감염을 포함할 수 있다. 헬퍼 바이러스 게놈은 Ad35 공여자 벡터의 증식을 구제할 수 있고 Ad35 공여자 벡터는

예를 들어, 대규모로 생산되고 단리될 수 있다. 다양한 프로토콜이 당업계, 예를 들어, Palmer 등, 2009 Gene Therapy Protocols. Methods in Molecular Biology, Volume 433. Humana Press; Totowa, NJ: 2009. pp. 33-53에 알려져 있다.

[0153] 본 개시내용은 GFP를 암호화하는 페이로드 코딩 서열을 발현하는 접촉된 세포의 퍼센트로 측정된 바와 같이, 본 개시내용의 HDAd35 공여자 벡터가 인간 CD34+ 세포의 형질도입에서 HDAd5/35 공여자 벡터과 비슷하게 수행함을 입증하는 예시적인 데이터를 포함한다. 결과는 접촉된 세포 당 500 내지 2000 개의 벡터 입자 범위의 다중 MOI에서 확인되었다. 예시적인 실험은 예시적인 데이터를 생성하는 데 사용된 HDAd35 공여자 벡터를 사용하여 수행하였으며 loxP 부위가 플랭킹된 뉴클레오티드 366-481에 플랭킹된 상기 개시된 바와 같은 Ad35 헬퍼 벡터를 사용하여 생성하였다(예를 들어, 도 117 참조).

[0154] 다양한 예시적인 공여자 벡터가 본원에 제공된다. 본 개시내용은 비제한적인 예로서 표 3-6에 제시된 바와 같은 HDAd35 공여자 게놈을 제공한다.

[0155] **표 3: 서열번호: 304에 따른 예시적인 HDAd35 공여자 벡터.**

| 서열 특징 | 서열번호: 304에서의 위치 |
|---|---|
| Ad35 5' (ITR 포함, 패키징 서열) | 시작: 1 종료: 481 |
| FRT 재조합효소 지시 반복 | 시작: 14126 종료: 14159 (상보적) |
| pT4 트랜스포사제 도립 반복 | 시작: 14220 종료: 14463 |
| EF1α 프로모터 | 시작: 14491 종료: 15825 |
| $mgmt^{P140K}$ 선택 카세트 | 시작: 15843 종료: 16466 |
| polyA 서열 | 시작: 16484 종료: 16705 |
| pT4 트랜스포사제 도립 반복 | 시작: 16735 종료: 17000 |
| FRT 재조합효소 지시 반복 | 시작: 17107 종료: 17140 (상보적) |
| Ad35 3' (ITR 포함) | 시작: 28823 종료: 29230 |

[0156]

[0157] **표 4: 서열번호: 305에 따른 예시적인 HDAd35 공여자 벡터**

| 서열 특징 | 서열번호: 305에서의 위치 |
|---|---|
| Ad35 5' (ITR 포함, 패키징 서열) | 시작: 1 종료: 481 |
| FRT 재조합효소 지시 반복 | 시작: 14126 종료: 14159 (상보적) |
| pT4 트랜스포사제 도립 반복 | 시작: 14220 종료: 14463 |
| EF1α 프로모터 | 시작: 14478 종료: 15812 |
| $mgmt^{P140K}$ 선택 카세트 | 시작: 15830 종료: 16450 |
| 2A 펩티드-암호화 서열 | 시작: 16451 종료: 16522 |
| GFP-암호화 서열 | 시작: 16523 종료: 17242 |
| SV40 polyA 서열 | 시작: 17269 종료: 17390 |
| pT4 트랜스포사제 도립 반복 | 시작: 17501 종료: 17766 |
| FRT 재조합효소 지시 반복 | 시작: 17873 종료: 17906 (상보적) |
| Ad35 3' (ITR 포함) | 시작: 29589 종료: 29996 |

[0158]

[0159] 표 5: 서열번호: 288에 따른 예시적인 HDAd35 공여자 벡터.

| 서열 특징 | 서열번호: 288에서의 위치 |
|---|---|
| Ad35 5' (ITR 포함, 패키징 서열) | 시작: 1 종료: 481 |
| FRT 재조합효소 지시 반복 | 시작: 14126 종료: 14159 (상보적) |
| pT4 트랜스포사제 도립 반복 | 시작: 14220 종료: 14463 |
| EF1α 프로모터 | 시작: 14478 종료: 15812 |
| $mgmt^{P140K}$ 선택 카세트 | 시작: 15830 종료: 16450 |
| 2A 펩티드-암호화 서열 | 시작: 16451 종료: 16522 |
| mCherry-암호화 서열 | 시작: 16526 종료: 17230 |
| SV40 polyA 서열 | 시작: 17259 종료: 17380 |
| pT4 트랜스포사제 도립 반복 | 시작: 17491 종료: 17756 |
| FRT 재조합효소 지시 반복 | 시작: 17863 종료: 17896 (상보적) |
| Ad35 3' (ITR 포함) | 시작: 29579 종료: 29986 |

[0160]

[0161] 표 6: 서열번호: 289에 따른 예시적인 지지 벡터.

| 서열 특징 | 서열번호: 289에서의 위치 |
|---|---|
| Ad35 5' (ITR 포함, 패키징 서열) | 시작: 1 종료: 481 |
| PGK 프로모터 | 시작: 14103 종료: 14614 |
| SB100x 트랜스포사제-암호화 서열 | 시작: 14763 종료: 15785 |
| BGH polyA 서열 | 시작: 15811 종료: 16128 |
| B-글로빈 polyA 서열 | 시작: 16088 종료: 16376 (상보적) |
| Flpe 재조합효소-암호화 서열 | 시작: 16488 종료: 17759 (상보적) |
| EF1α 프로모터 | 시작: 17780 종료: 18895 (상보적) |
| Ad35 3' (ITR 포함) | 시작: 29751 종료: 30158 |

[0162]

[0163] 표 7: 서열번호: 286에 따른 예시적인 Ad35 헬퍼 벡터

| 서열 특징 | 서열번호: 286에서의 위치 |
|---|---|
| Ad35 5' (ITR 포함)(Ad35 nt 1-178) | 시작: 2582 종료: 2759 |
| LoxP 재조합효소 부위 | 시작: 2768 종료: 2801 |
| Ad35 패키징 서열 (Ad35 nt 179-344) | 시작: 2808 종료: 2973 |
| LoxP 재조합효소 부위 | 시작: 2974 종료: 3007 |
| Ad35 서열 (Ad35 nt 3112-27435) | 시작: 3016 종료: 27338 |
| 람다-1 서열 | 시작: 27393 종료: 29862 (상보적) |
| BGH polyA 서열 | 시작: 30176 종료: 30390 |
| CopGFP-암호화 서열 | 시작: 30415 종료: 31080 (상보적) |
| CMV 프로모터 | 시작: 31127 종료: 31779 (상보적) |
| 람다-2 서열 | 시작: 31831 종료: 33360 |
| Ad35 서열 (Ad35 nt 30544-31879) | 시작: 33421 종료: 34756 |
| Ad5 E4orf6 서열 | 시작: 34752 종료: 35866 |
| Ad35 3' (ITR 포함) (Ad35 nt 32972-34794) | 시작: 35864 종료: 37686 |

[0164]

[0165] 표 8: 서열번호: 51에 따른 예시적인 Ad35 헬퍼 벡터.

| 서열 특징 | 서열번호: 51에서의 위치 |
|---|---|
| Ad35 5' (ITR 포함) (Ad35 nt 1-178) | 시작: 2582 종료: 2759 |
| LoxP 재조합효소 부위 | 시작: 2768 종료: 2801 |
| Ad35 패키징 서열 (Ad35 nt 366-481) | 시작: 2808 종료: 2923 |
| LoxP 재조합효소 부위 | 시작: 2924 종료: 2957 |
| Ad35 서열 (Ad35 nt 3112-2743) | 시작: 2966 종료: 27288 |
| 람다-1 서열 | 시작: 27343 종료: 29812 (상보적) |
| BGH polyA 서열 | 시작: 30126 종료: 30340 |
| CopGFP-암호화 서열 | 시작: 30365 종료: 31030 (상보적) |
| CMV 프로모터 | 시작: 31077 종료: 31729 (상보적) |
| 람다-2 서열 | 시작: 31781 종료: 33310 |
| Ad35 서열 (Ad35 nt 30544-31879) | 시작: 33371 종료: 34706 |
| Ad5 E4orf6 서열 | 시작: 34702 종료: 35816 |
| Ad35 3' (ITR 포함) (Ad35 nt 32972-34794) | 시작: 35814 종료: 37636 |

[0166]

[0167] 표 9: 서열번호: 52에 따른 예시적인 Ad35 헬퍼 벡터.

| 서열 특징 | 서열번호: 52에서의 위치 |
|---|---|
| Ad35 5' (ITR 포함) (Ad35 nt 1-154) | 시작: 2582 종료: 2735 |
| LoxP 재조합효소 부위 | 시작: 2744 종료: 2777 |
| Ad35 패키징 서열 (Ad35 nt 155-481) | 시작: 2784 종료: 3110 |
| LoxP 재조합효소 부위 | 시작: 3111 종료: 3144 |
| Ad35 서열 (Ad35 nt 3112-27435) | 시작: 3153 종료: 27475 |
| 람다-1 서열 | 시작: 27530 종료: 29999 (상보적) |
| BGH polyA 서열 | 시작: 30313 종료: 30527 |
| CopGFP-암호화 서열 | 시작: 30552 종료: 31217 (상보적) |
| CMV 프로모터 | 시작: 31264 종료: 31916 (상보적) |
| 람다-2 서열 | 시작: 31968 종료: 33497 |
| Ad35 서열 (Ad35 nt 30544-31879) | 시작: 33558 종료: 34893 |
| Ad5 E4orf6 서열 | 시작: 34889 종료: 36003 |
| Ad35 3' (ITR 포함) (Ad35 nt 32972-34794) | 시작: 36001 종료: 37823 |

[0168]

[0169] I(C). 유전자 요법 벡터 페이로드

[0170] 본 개시내용의 Ad35 및 Ad5/35 공여자 벡터 및 게놈은 하나 이상의 발현 산물을 암호화하는 하나 이상의 코딩 서열, 코딩 서열에 작동가능하게 연결된 하나 이상의 조절 서열, 하나 이상의 스터퍼 서열 등 중 임의의 것을 포함할 수 있는 다양한 핵산 페이로드를 포함할 수 있다. 다양한 구현예에서, 페이로드는 숙주 세포 또는 시스템에서 치료 효과와 같은 원하는 결과, 예를 들어, 치료 관심 단백질의 발현 또는 치료 관심의 서열 변형을 생성하기 위한 유전자 편집 시스템, 예를 들어, CRISPR/Cas 시스템 또는 염기 편집 시스템의 발현을 달성하기 위해 조작된다. 일부 구현예에서, 페이로드는 유전자를 포함할 수 있다. 유전자는 코딩 서열 뿐만 아니라 조절 영역 예컨대 프로모터, 인핸서, 종결 영역, 유전자좌 제어 영역(LCR), 종결 및 폴리아데닐화 신호 요소, 스플라이싱 신호 요소 등을 포함할 수 있다. 용어는 대안적인 스플라이스 부위를 초래하는 변이체와 함께, mRNA 전사체로부터 스플라이싱된 모든 인트론 및 다른 DNA 서열을 추가로 포함할 수 있다. 서열은 또한 특이적 유기체 또는 세포 유형에 코돈 선호도를 제공하도록 도입될 수 있는 참조 서열 또는 서열들의 축퇴성 코돈을 포함할 수

있다.

[0171] 페이로드는 단일 유전자 또는 다중 유전자를 포함할 수 있다. 페이로드는 단일 조절 서열 또는 복수의 조절 서열을 포함할 수 있다. 페이로드는 단일 코딩 서열 또는 복수의 코딩 서열을 포함할 수 있다. 페이로드는 코딩 서열의 개별 발현 산물이 함께 예를 들어, 엔도뉴클레아제 및 가이드 RNA의 경우에서와 같이, 또는 독립적으로, 예를 들어 직접적으로 또는 간접적으로 결합하지 않는 2 개의 개별 단백질로서 기능하는 복수의 코딩 서열을 포함할 수 있다. 일부 경우에, 복수의 코딩 서열은 예를 들어 협력하여 기능할 수 있으며, 여기서 엔도뉴클레아제 및 가이드 RNA는 숙주 세포 또는 시스템에 내인성인 코딩 서열의 발현 증가를 야기하고 페이로드는 내인성 코딩 서열에 의해 암호화된 단백질의 것에 상응하는 적어도 하나의 생물학적 활성을 갖는 단백질을 추가로 암호화하고 발현한다. 당업자에 의해 이해되는 바와 같이, 표준 야생형 Ad35 게놈에 의해 암호화되지 않은 본원에 제공된 임의의 페이로드-암호화된 발현 산물은 본원에서 이종 발현 산물로 지칭될 수 있다.

[0172] **I(C)(i). 페이로드 발현 산물**

[0173] 본 개시내용의 아데노바이러스 공여자 벡터 또는 아데노바이러스 공여자 게놈의 페이로드는 다양한 발현 산물 중 임의의 것을 암호화하는 하나 이상의 코딩 서열을 포함할 수 있다. 예시적인 발현 산물은 참조 수준과 비교하여 생물학적으로 활성인 단백질의 낮은 발현 또는 활성을 특징으로 하는 질환 또는 병태의 치료를 위한 대체 요법 단백질을 포함하나 이에 제한되지 않는 단백질을 포함한다. 예시적인 발현 산물은 CRISPR/Cas 및 염기 편집기 시스템을 포함한다. 예시적인 발현 산물은 항체, CAR, 및 TCR를 포함한다. 예시적인 발현 산물은 작은 RNA를 포함한다. 다양한 구현예에서, 예를 들어, 의도된 또는 표적 효과가 CRISPR 시스템 또는 염기 편집기 시스템에 의한 숙주 세포 게놈의 편집을 포함하는 특정 경우에, 공여자 벡터 페이로드의 전부 또는 일부의 숙주 세포 게놈 내로의 통합은 의도된 또는 표적 효과를 생성하도록 공여자 벡터 또는 게놈의 표적 세포에 전달하기 위해 요구되지 않는다. 다양한 구현예에서, 예를 들어, 페이로드-암호화된 발현 산물의 발현이 형질도입된 표적 세포의 자손 세포에서 바람직한 경우, 공여자 벡터 페이로드의 전부 또는 일부의 통합은 의도된 또는 표적 효과를 생성하도록 공여자 벡터 또는 게놈의 표적 세포에 전달하기 위해 요구되거나 또는 바람직하다. 다양한 구현예에서, 페이로드는 예를 들어, 재조합 또는 전위에 의해 숙주 세포 게놈 내로 통합을 위해 조작된 핵산 서열("통합 요소")을 포함할 수 있다.

[0174] 하나 이상의 치료 단백질을 암호화하는 유전자 서열은 관련 아미노산 서열로부터 합성 또는 재조합 방법에 의해 용이하게 제조될 수 있다. 특정 구현예에서, 이들 서열 중 임의의 것을 암호화하는 유전자 서열은 또한 서열을 암호화하는 유전자 서열을 상이한 서열을 암호화하는 또 다른 유전자 서열로 대체하고 용이한 절제를 제공하기 위해 코딩 서열의 5' 및/또는 3' 단부에서 하나 이상의 제한 효소 부위를 가질 수 있다. 특정 구현예에서, 서열을 암호화하는 유전자 서열은 포유류 세포에서 발현을 위해 코돈 최적화될 수 있다.

[0175] 치료 유전자 및/또는 유전자 산물의 특정 예는 γ-글로빈, 인자 VIII, γC, JAK3, IL7RA, RAG1, RAG2, DCLRE1C, PRKDC, LIG4, NHEJ1, CD3D, CD3E, CD3Z, CD3G, PTPRC, ZAP70, LCK, AK2, ADA, PNP, WHN, CHD7, ORAI1, STIM1, CORO1A, CIITA, RFXANK, RFX5, RFXAP, RMRP, DKC1, TERT, TINF2, DCLRE1B, 및 SLC46A1; FancA, FancB, FancC, FancD1(BRCA2), FancD2, FancE, FancF, FancG, FancI, FancJ(BRIP1), FancL, FancM, FancN(PALB2), FancO(RAD51C), FancP(SLX4), FancQ(ERCC4), FancR(RAD51), FancS(BRCA1), FancT(UBE2T), FancU(XRCC2), FancV(MAD2L2), 및 FancW(RFWD3)를 포함하는 FANC 패밀리 유전자; 가용성 CD40; CTLA; Fas L; CD4, CD5, CD7, CD52 등에 대한 항체; IL1, IL2, IL6에 대한 항체; 자가반응성 T 세포 상에 특이적으로 존재하는 TCR에 대한 항체; IL4; IL10; IL12; IL13; IL1Ra, sIL1RI, sIL1RII; sTNFRI; sTNFRII; TNF에 대한 항체; P53, PTPN22, 및 DRB1*1501/DQB1*0602; 글로빈 패밀리 유전자; WAS; phox; 디스트로핀; 피루베이트 키나제; CLN3; ABCD1; 아릴술파타제 A; SFTPB; SFTPC; NLX2.1; ABCA3; GATA1; 리보솜 단백질 유전자; TERT; TERC; DKC1; TINF2; CFTR; LRRK2; PARK2; PARK7; PINK1; SNCA; PSEN1; PSEN2; APP; SOD1; TDP43; FUS; 유비퀼린 2; C9ORF72 및 본원에 기재된 다른 치료 유전자를 포함한다.

[0176] 치료 유전자는 적혈구 및 응혈과 관련된 질환에 대한 치료적으로 효과적인 반응을 제공하기 위해 선택될 수 있다. 특정 구현예에서, 질환은 지중해빈혈과 같은 혈색소병증, 또는 겸상적혈구병/특성이다. 치료 유전자는 예를 들어, 헤모글로빈의 생산을 유도하거나 또는 증가시키거나; β-글로빈, γ-글로빈, 또는 α-글로빈의 생산을 유도하거나 또는 증가시키거나; 또는 신체에서 세포에 대한 산소의 이용가능성을 증가시키는 유전자일 수 있다. 치료 유전자는 예를 들어, HBB 또는 CYB5R3일 수 있다. 예시적인 효과적인 치료는 예를 들어, 혈액 세포 계수를 증가시키거나, 혈액 세포 기능을 개선하거나, 또는 환자에서 세포의 산소화를 증가시킬 수 있다. 또 다른 특정 구현예에서, 질환은 혈우병이다. 치료 유전자는 예를 들어, 응고/응혈 인자 VIII 또는 응고/응혈 인자 IX

의 생산을 증가시키거나, 응고 인자 VIII 또는 응고 인자 IX의 정상 버전의 생산을 야기하는 유전자, 응고/응혈 인자 VIII 또는 응고/응혈 인자 IX에 대한 항체의 생산을 감소시키는 유전자, 또는 혈전의 적절한 형성을 야기하는 유전자일 수 있다. 예시적인 치료 유전자는 F8 및 F9를 포함한다. 예시적인 효과적인 치료는 예를 들어, 응고/응혈 인자 VIII 및 IX의 생산을 증가시키거나 또는 유도하거나; 응고/응혈 인자 VIII 및 IX의 기능화를 개선하거나, 또는 대상체에서 응혈 시간을 감소시킬 수 있다.

[0177] 본 개시내용의 다양한 구현예에서, 공여자 벡터는 글로빈 유전자를 암호화하며, 여기서 글로빈 유전자에 의해 암호화된 글로빈 단백질은 γ-글로빈, β-글로빈, 및/또는 α-글로빈으로부터 선택된다. 본 개시내용의 글로빈 유전자는 예를 들어, 글로빈 단백질을 암호화하는 핵산 서열에 작동가능하게 연결된 프로모터와 같은 하나 이상의 조절 서열을 포함할 수 있다. 당업자가 이해하는 바와 같이, γ-글로빈, β-글로빈, 및/또는 α-글로빈 각각은 태아 및/또는 성인 헤모글로빈의 구성요소이며 따라서 본원에 개시된 다양한 벡터에 유용하다.

[0178] 다양한 구현예에서, 글로빈 단백질의 발현을 증가시키는 것은 (i) 특정 서열을 갖는 글로빈 단백질의 세포 또는 시스템에서 양, 농도, 또는 발현(예를 들어, 암호화하는 핵산의 전사 또는 번역)의 증가; (ii) 서로에 대한 단백질의 서열과 관련없이 특정 유형의 글로빈 단백질의 세포 또는 시스템에서(예를 들어, 암호화하는 핵산의 전사 또는 번역) 양, 농도, 또는 발현(예를 들어, 당업자에 의해 또는 본 명세서에 제시된 바와 같은 γ-글로빈(또는 대안적으로 β-글로빈 또는 α-글로빈)으로 식별될 모든 단백질의 총량)의 증가; 및/또는 (iii) 이종 글로빈 단백질, 예를 들어, 유전자 요법 전에 숙주 세포에 의해 암호화되지 않은 글로빈 단백질의 세포 또는 시스템에서 발현 중 하나 이상 중 임의의 것을 지칭할 수 있다.

[0179] 다음 참고문헌은 기능적 글로빈 유전자의 특정 예시적인 서열을 설명한다. 참고문헌 1-4는 α-형 글로빈 서열에 관한 것이고 참고문헌 4-12는 β-형 글로빈 서열(β 및 γ 글로빈 서열)에 관한 것이며, 이의 서열은 본원에 참조로 포함된다: (1) GenBank 수탁 번호 Z84721 (Mar. 19, 1997); (2) GenBank 수탁 번호 NM_000517 (Oct. 31, 2000); (3) Hardison 등, J. Mol. Biol. (1991) 222(2):233-249; (4) A Syllabus of Human Hemoglobin Variants (1996), by Titus 등, published by The Sickle Cell Anemia Foundation in Augusta, Ga. (available online at globin.cse.psu.edu); (5) GenBank 수탁 번호 J00179 (Aug. 26, 1993); (6) Tagle 등, Genomics (1992) 13(3):741-760; (7) Grovsfeld 등, Cell (1987) 51(6):975-985; (8) Li 등, Blood (1999) 93(7):2208-2216; (9) Gorman 등, J. Biol. Chem. (2000) 275(46):35914-35919; (10) Slightom 등, Cell (1980) 21(3):627-638; (11) Fritsch 등, Cell (1980) 19(4): 959-972; (12) Marotta 등, J. Biol. Chem. (1977) 252(14):5040-5053. 글로빈을 암호화하는 유전자의 추가적인 코딩 및 비-코딩 영역에 대해 예를 들어, Marotta 등, Prog. Nucleic Acid Res. Mol. Biol. 19, 165-175, 1976, Lawn 등, Cell 21 (3), 647-651, 1980, 및 Sadelain 등, PNAS.; 92:6728-6732, 1995를 참조한다.

[0180] 헤모글로빈 서브유닛 β의 예시적인 아미노산 서열은 예를 들어, NCBI 수탁 번호 P68871에 제공된다. β-글로빈에 대한 예시적인 아미노산 서열은 예를 들어, NCBI 수탁 번호 NP_000509에 제공된다.

[0181] 치료 유전자 및/또는 유전자 산물 이외에도, 이식유전자는 또한 체크포인트 억제제 시약, 하나 이상의 암 항원에 특이적인 키메라 항원 수용체 분자, 및/또는 하나 이상의 암 항원에 특이적인 T-세포 수용체와 같은 치료 분자를 암호화할 수 있다.

[0182] 또 다른 예로서, 치료 유전자는 리소좀 축적 장애에 대한 치료적으로 효과적인 반응을 제공하기 위해 선택될 수 있다. 특정 구현예에서, 리소좀 축적 장애는 뮤코다당류축적증(MPS), I형; MPS II 또는 헌터 증후군; MPS III 또는 산필리포 증후군; MPS IV 또는 모르키오 증후군; MPS V; MPS VI 또는 마로토-라미 증후군; MPS VII 또는 슬라이 증후군; α-만노사이드축적증; β-만노사이드축적증; GSDI로도 알려져 있는 글리코겐 축적 질환 I형, 폰 기르케병, 또는 테이 삭스; 폼페병; 고세병; 파브리병이다. 치료 유전자는 예를 들어 효소의 생산을 암호화하거나 또는 유도하거나, 또는 달리 리소좀에서 뮤코다당류의 분해를 야기하는 유전자일 수 있다. 예시적인 치료 유전자는 IDUA 또는 이두로니다제, IDS, GNS, HGSNAT, SGSH, NAGLU, GUSB, GALNS, GLB1, ARSB, 및 HYAL1을 포함한다. 리소좀 축적 질환에 대한 예시적인 효과적인 유전적 요법은 예를 들어, 리소좀에서 다양한 물질의 분해를 담당하는 효소의 생산을 암호화 또는 유도하고; 머리(대두증), 간, 비장, 혀, 또는 성대를 포함하는 다양한 기관에서 부종을 감소, 제거, 방지, 또는 지연시키고; 뇌의 체액을 감소시키고; 심장 판막 이상을 감소시키고; 좁아진 기도를 방지 또는 연장시키고 감염 및 수면 무호흡과 같은 관련 상부 호흡기 병태를 예방하고; 뉴런의 파괴, 및/또는 연관 증상을 감소, 제거, 예방, 또는 지연시킬 수 있다.

[0183] 또 다른 예로서, 치료 유전자는 과증식성 질환에 대한 치료적으로 효과적인 반응을 제공하기 위해 선택될 수 있다. 특정 구현예에서, 과증식성 질환은 암이다. 치료 유전자는 예를 들어, 종양 억제인자 유전자, 세포자멸사

를 유도하는 유전자, 효소를 암호화하는 유전자, 항체를 암호화하는 유전자, 또는 호르몬을 암호화하는 유전자일 수 있다. 예시적인 치료 유전자 및 유전자 산물은 (본원의 다른 곳에 나열된 것들 이외에도) 101F6, 123F2(RASSF1), 53BP2, abl, ABLI, ADP, aFGF, APC, ApoAI, ApoAIV, ApoE, ATM, BAI-1, BDNF, 베타*(BLU), bFGF, BLC1, BLC6, BRCA1, BRCA2, CBFA1, CBL, C-CAM, CNTF, COX-1, CSFIR, CTS-1, 사이토신 데아미나제, DBCCR-1, DCC, Dp, DPC-4, E1A, E2F, EBRB2, erb, ERBA, ERBB, ETS1, ETS2, ETV6, Fab, FCC, FGF, FGR, FHIT, fms, FOX, FUS1, FYN, G-CSF, GDAIF, 유전자 21(NPRL2), 유전자 26(CACNA2D2), GM-CSF, GMF, gsp, HCR, HIC-1, HRAS, hst, IGF, IL-1, IL-2, IL-3, IL-5, IL-6, IL-7, IL-8, IL-9, IL-11, ING1, 인터페론 α, 인터페론 β, 인터페론 ɣ, IRF-1, JUN, KRAS, LUCA-1(HYAL1), LUCA-2(HYAL2), LYN, MADH4, MADR2, MCC, mda7, MDM2, MEN-I, MEN-II, MLL, MMAC1, MYB, MYC, MYCL1, MYCN, neu, NF-1, NF-2, NGF, NOEY1, NOEY2, NRAS, NT3, NT5, OVCA1, p16, p21, p27, p57, p73, p300, PGS, PIM1, PL6, PML, PTEN, raf, Rap1A, ras, Rb, RB1, RET, rks-3, ScFv, scFV ras, SEM A3, SRC, TALI, TCL3, TFPI, 트롬보스폰딘, 티미딘 키나제, TNF, TP53, trk, T-VEC, VEGF, VHL, WT1, WT-1, YES, 및 zac1을 포함한다. 예시적인 효과적인 유전적 요법은 종양을 억제 또는 제거하거나, 암 세포 수를 감소시키거나, 종양 크기를 감소시키거나, 종양 성장을 지연 또는 제거하거나, 종양에 의해 야기된 증상을 완화할 수 있다.

[0184] 또 다른 예로서, 치료 유전자는 감염성 질환에 대한 치료적으로 효과적인 반응을 제공하기 위해 선택될 수 있다. 특정 구현예에서, 감염성 질환은 인간 면역결핍 바이러스(HIV)이다. 치료 유전자는 예를 들어, 면역 세포가 HIV 감염에 대해 내성을 갖도록 하거나, 또는 면역 세포가 면역 재구축을 통해 바이러스를 효과적으로 중화할 수 있게 하는 유전자, 면역 세포에 의해 발현된 단백질을 암호화하는 유전자의 다형성, 환자에서 발현되지 않는 감염과 싸우는 데 유리한 유전자, 감염원, 수용체 또는 공수용체를 암호화하는 유전자; 수용체 또는 공수용체에 대한 리간드를 암호화하는 유전자; 특정 전사 인자의 작용을 차단하는 리보자임, 안티센스 RNA, 작은 간섭 RNA(siRNA) 또는 유인 RNA를 암호화하는 유전자; 우성 음성 바이러스 단백질, 세포내 항체, 인트라킨 및 자살 유전자를 암호화하는 유전자를 포함하는 바이러스 복제에 필수적인 바이러스 및 세포 유전자일 수 있다. 예시적인 치료 유전자 및 유전자 산물은 α2β1; αvβ3; αvβ5; αvβ63; BOB/GPR15; Bonzo/STRL-33/TYMSTR; CCR2; CCR3; CCR5; CCR8; CD4; CD46; CD55; CXCR4; 아미노펩티다제-N; HHV-7; ICAM; ICAM-1; PRR2/HveB; HveA; α-디스트로글리칸; LDLR/α2MR/LRP; PVR; PRR1/HveC; 및 라미닌 수용체를 포함한다. 예를 들어, HIV의 치료를 위한 치료 유효량은 HIV에 대한 대상체의 면역력을 증가시키거나, AIDS 또는 HIV와 연관된 증상을 개선시키거나, 또는 HIV에 대한 대상체에서 선천적 또는 적응 면역 반응을 유도할 수 있다. HIV에 대한 면역 반응은 항체 생산을 포함하고 AIDS 예방 및/또는 대상체의 AIDS 또는 HIV 감염 증상 개선을 초래하거나, 또는 HIV 감염성 및/또는 병독성을 감소 또는 제거할 수 있다.

[0185] 다양한 구현예에서, 본 개시내용의 벡터 또는 게놈, 예를 들어, Ad35 헬퍼 벡터 또는 Ad35 헬퍼 게놈은 예를 들어, CRISPR/Cas의 정상 활성을 억제하는 파지로부터 유래된 항-CRISPR (Acr) 단백질을 암호화하고/하거나 발현한다.

[0186] **I(C)(i)(a). 결합 도메인, 항체, CAR, 및 TCR 페이로드 발현 산물**

[0187] 본 개시내용은 다양한 결합 도메인을 포함한다. 항체는 결합 도메인의 일 예이며 세포 마커에 특이적으로 결합하는 전체 항체 또는 항체의 결합 단편, 예를 들어, Fv, Fab, Fab', $F(ab')_2$, 및 단일 쇄(sc) 형태 및 이의 단편(예를 들어, scFv)을 포함한다. 항체 또는 항원 결합 단편은 다클론 항체, 단클론 항체, 인간 항체, 인간화 항체, 합성 항체, 비-인간 항체, 재조합 항체, 키메라 항체, 이중특이적 항체, 미니바디(mini body), 및 선형 항체의 전부 또는 일부를 포함할 수 있다. 이의 기능적 단편은 중쇄 가변 도메인(VH), 경쇄 가변 도메인(VL) 및 낙타과 유래 나노바디(nanobody)의 가변 도메인(VHH) 등과 같은 단일-도메인 항체를 포함한다.

[0188] 일부 경우에, scFv는 당업계에 알려진 방법에 따라 제조될 수 있다(예를 들어, Bird 등, *Science* 242:423-426, 1988; 및 Huston 등, *Proc. Natl. Acad. Sci. USA* 85:5879-5883, 1988 참조). ScFv 분자는 유연한 폴리펩티드 링커를 사용하여 항체의 VL 및 VH 영역을 함께 연결함으로써 생성될 수 있다. 짧은 폴리펩티드 링커가 이용되는 경우(예를 들어, 5-10 개의 아미노산) 쇄간 접힘이 방지된다. 쇄간 접힘은 또한 기능적 에피토프 결합 부위를 형성하기 위해 2 개의 가변 영역을 함께 가져오는 데 필요하다. 링커 방향 및 크기의 예는 예를 들어, Hollinger 등 1993 Proc Natl Acad. Sci. U.S.A. 90:6444-6448, US 2005/0100543, US 2005/0175606, US 2007/0014794, WO2006/020258, 및 WO2007/024715를 참조한다.

[0189] scFv는 VL 및 VH 영역 사이에 적어도 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50 개, 또는 그 이상의 아미노산 잔기의 링커를 포함할 수 있다. 특정 구현예에서,

링커 서열은 임의의 자연 발생 아미노산을 포함할 수 있다. 일반적으로, scFv의 VL 및 VH를 연결하는 데 사용되는 링커 서열은 5 내지 35 개의 아미노산 길이이다. 특정 구현예에서, VL-VH 링커는 5 내지 35 개, 10 내지 30 개의 아미노산 또는 15 내지 25 개의 아미노산을 포함한다. 링커 길이의 변경은 활성을 유지하거나 또는 향상시켜, 활성 연구에서 우수한 효능을 일으킬 수 있다.

[0190] 일부 구현예에서, scFv의 링커 서열은 아미노산 글리신 및 세린을 포함한다. 특정 구현예에서, 링커 서열은 (GlyxSery)n의 1 내지 10 개의 반복부와 같은 글리신 및 세린 반복부의 세트를 포함하며, 여기서 x 및 y는 독립적으로 0 내지 10의 정수이며, 단, x 및 y는 둘 다 0이 아니고 여기서 n은 1, 2, 3, 4, 5, 6, 7, 8, 9 또는 10의 정수이고) 여기서 연결된 VH-VL 영역은 기능적 면역글로불린-유사 결합 도메인(예를 들어, scFv, scTCR)을 형성한다. 특정 예는 (Gly4Ser)n, (Gly3Ser)n(Gly4Ser)n, (Gly3Ser)n(Gly2Ser)n, (Gly3Ser)n(Gly4Ser)1, (Gly4Ser)1, (Gly3Ser)1, 또는 (Gly2Ser)1을 포함한다. 특정 구현예에서, 링커는 (Gly4Ser)4 또는 (Gly4Ser)3이다. 상기 scTCR에 대한 참조를 통해 나타낸 바와 같이, 이러한 링커는 또한 T 세포 수용체 Vα/β 및 Cα/β 쇄(예를 들어, Vα-Cα, Vβ-Cβ, Vα-Vβ)에 연결하는 데 사용될 수 있다.

[0191] 추가적인 예는 scFv-기반 그라바바디(grababody) 및 가용성 VH 도메인 항체를 포함한다. 이들 항체는 중쇄 가변 영역만을 사용하여 결합 영역을 형성한다. 예를 들어, Jespers 등, *Nat. Biotechnol*. 22:1161, 2004; Cortez-Retamozo 등, *Cancer Res*. 64:2853, 2004; Baral 등, *Nature Med*. 12:580, 2006; 및 Barthelemy 등, *J. Biol. Chem*. 283:3639, 2008 참조.

[0192] 일부 경우에, 결합 도메인이 궁극적으로 사용될 동일한 종으로부터 유래되는 것이 유익하다. 예를 들어, 인간에서 사용하기 위해, 항원 결합 도메인이 인간 항체, 인간화 항체, 또는 이의 단편 또는 조작된 형태를 포함하는 것이 유익할 수 있다. 인간 기원의 항체 또는 인간화 항체는 인간에서 면역원성이 낮거나 또는 전혀 없고 비-인간 항체와 비교하여 더 적은 수의 비-면역원성 에피토프를 갖는다. 항체 및 이의 조작된 단편은 일반적으로 인간 대상체에서 감소된 수준의 항원성을 갖거나 또는 전혀 없도록 선택될 것이다.

[0193] 특정 구현예에서, 결합 도메인은 인간화 항체 또는 이의 조작된 단편을 포함한다. 특정 구현예에서, 비-인간 항체는 인간화되며, 여기서 항체의 하나 이상의 아미노산 잔기는 인간 또는 이의 단편에서 자연적으로 생성된 항체에 대한 유사성을 증가시키도록 변형된다. 이들 비인간 아미노산 잔기는 종종 "유입" 잔기로 지칭되며, 전형적으로 "유입" 가변 도메인으로부터 취해진다. 본원에 제공된 바와 같이, 인간화 항체 또는 항체 단편은 비인간 면역글로불린 분자 및 프레임워크 영역으로부터 하나 이상의 CDR을 포함하며 여기서 프레임워크를 포함하는 아미노산 잔기는 인간 생식계열로부터 완전히 또는 대부분 유래된다. 일 측면에서, 항원 결합 도메인은 인간화된다. 인간화 항체는 CDR-이식(예를 들어, 유럽 특허 번호 EP 239,400; WO 91/09967; 및 US 5,225,539, US 5,530,101, 및 US 5,585,089 참조), 베니어링(veneering) 또는 재표면화(예를 들어, EP 592,106 및 EP 519,596; Padlan, 1991, Molecular Immunology, 28(4/5):489-498; Studnicka 등, 1994, Protein Engineering, 7(6):805-814; 및 Roguska 등, PNAS, 91:969-973, 1994 참조), 쇄 셔플링(shuffling)(예를 들어, US. 5,565,332 참조), 및 예를 들어, US 2005/0042664, US 2005/0048617, US 6,407,213, US 5,766,886, WO 9317105, Tan 등, *J. Immunol*., 169:1119-25, 2002, Caldas 등, *Protein Eng*., 13(5):353-60, 2000, Morea 등, *Methods*, 20(3):267-79, 2000, Baca 등, *J. Biol. Chem*., 272(16): 10678-84, 1997, Roguska 등, *Protein Eng*., 9(10):895-904, 1996, Couto 등, *Cancer Res*., 55 (23 Supp):5973s-5977s, 1995, Couto 등, *Cancer Res*., 55(8):1717-22, 1995, Sandhu, *Gene*, 150(2):409-10, 1994, 및 Pedersen 등, *J. Mol. Biol*., 235(3):959-73, 1994에 개시된 기술을 포함하여, 당업계에 알려진 다양한 기술을 사용하여 생성될 수 있다. 종종, 프레임워크 영역의 프레임워크 잔기는 CDR 공여자 항체로부터 상응하는 잔기로 치환되어 세포 마커 결합을 변경, 예를 들어 개선시킬 것이다. 이들 프레임워크 치환은 당업계에 잘 알려진 방법, 예를 들어, 세포 마커 결합에 중요한 프레임워크 잔기를 식별하기 위한 CDR 및 프레임워크 잔기의 상호작용 모델링 및 특정 위치에서 특이한 프레임워크 잔기를 식별하기 위한 서열 비교에 의해 식별된다. (예를 들어, US 5,585,089; 및 Riechmann 등, *Nature*, 332:323, 1988 참조).

[0194] 특정 세포 마커에 특이적으로 결합하는 항체 및 다른 결합 도메인은 단클론 항체를 수득하는 방법, 파지 디스플레이 방법, 인간 또는 인간화 항체를 생성하는 방법, 또는 당업자에게 알려진 바와 같이 항체를 생성하도록 조작된 유전자이식 동물 또는 식물을 사용하는 방법을 사용하여 제조될 수 있다(예를 들어, US 6,291,161 및 US 6,291,158 참조). 부분적 또는 완전 합성 항체의 파지 디스플레이 라이브러리가 이용가능하며 세포 마커에 결합할 수 있는 항체 또는 이의 단편에 대해 스크리닝될 수 있다. 예를 들어, 결합 도메인은 관심 세포 마커에 특이적으로 결합하는 Fab 단편에 대한 Fab 파지 라이브러리를 스크리닝함으로써 식별될 수 있다(Hoet 등, *Nat. Biotechnol*. 23:344, 2005 참조). 인간 항체의 파지 디스플레이 라이브러리가 또한 이용가능하다.

추가적으로, 편리한 시스템(예를 들어, 마우스, HuMAb mouse®(GenPharm Int'l. Inc., 캘리포니아주 마운틴뷰 소재), TC mouse®(Kirin Pharma Co. Ltd., 일본 도쿄 소재), KM-mouse®(Medarex, Inc., 뉴저지주 프린스턴 소재), 라마, 닭, 래트, 햄스터, 토끼 등)에서 관심 세포 마커를 면역원으로 사용한 하이브리도마 개발을 위한 전통적인 전략을 사용하여 결합 도메인을 개발할 수 있다. 특정 구현예에서, 항체는 특정 암 세포 유형에 의해 우선적으로 발현되는 세포 마커에 특이적으로 결합하고 비특이적 구성요소 또는 관련되지 않은 표적과 교차 반응하지 않는다. 일단 식별되면, 항체의 아미노산 서열 및 항체를 암호화하는 유전자 서열은 단리되고/되거나 결정될 수 있다.

[0195] 특정 구현예에서, 치료 유전자는 Fab 또는 scFv와 같은 항체 또는 항체의 결합 단편을 암호화할 수 있다. 발현될 수 있는 예시적인 항체(scFv 포함)는 WO2014/164553A1, US2017/0283504, US 7,083,785, US 10,189,906, US 10,174,095, WO2005102387, US2011/0206701A1, WO2014/179759A1, US2018/0037651A1, US2018/0118822A1, WO2008/047242A2, WO1996/016990A1, WO200/5103083A2, 및 WO1999/062526A2에 기재된 제공된 것들을 포함한다. 결합 도메인과 관련하여 상기 기재된 항체, 뿐만 아니라 아테졸리주맙, 블리나투모맙, 브렌툭시맙, 세툭시맙, 섬투주맙, 팔레투주맙, 겜투주맙, OKT3, 오레고보맙, 프로믹시맙, 펨브롤리주맙, 및 트라스투주맙이 또한 사용될 수 있다.

[0196] 면역 체크포인트 억제제가 또한 사용될 수 있다. 면역 체크포인트 억제제는 면역 억제 체크포인트 단백질의 기능을 억제하는 화합물을 지칭한다. 억제는 기능 감소 및 완전 차단을 포함한다. 바람직한 면역 체크포인트 억제제는 면역 체크포인트 단백질을 특이적으로 인식하는 항체이다. 다수의 면역 체크포인트 억제제가 알려져 있으며 이러한 알려진 면역 체크포인트 단백질 억제제의 유사성에서, 대안적인 면역 체크포인트 억제제가 (가까운) 미래에 개발될 수 있다. 면역 체크포인트 억제제는 펩티드, 항체, 핵산 분자 및 소분자를 포함한다. 특정 구현예에서, 면역 체크포인트 억제제는 대상체에서 CD8+ T 세포의 증식, 이동, 지속 및/또는 세포독성 활성 및 특히 대상체의 CD8+ T 세포의 종양-침윤을 향상시킨다. 또 다른 예시적인 면역 체크포인트 억제제는 실시예 4에 개시된 바와 같은 체크포인트 억제제를 포함한다. 따라서, 본 개시내용의 예시적인 면역 체크포인트 억제제는 αPD-L1γ1 항체(대안적으로 αPD-L1$\gamma_1$로 지칭됨)를 포함한다. αPD-L1γ1은 Engeland 등 *Mol Ther* 22(11):1949-1959, 2014에 추가로 기재되며, 특히 항-PD-L1 항체, 이를 암호화하는 핵산, 및 이의 용도와 관련하여그 전문이 본원에 참조로 포함된다.

[0197] PD-1 및 PD-L1 항체의 예는 US 7,488,802; US 7,943,743; US 8,008,449; US 8,168,757; US 8,217,149, WO03042402, WO2008156712, WO2010089411, WO2010036959, WO2011066342, WO2011159877, WO2011082400, 및 WO2011161699에 기재되어 있다. 일부 구현예에서, PD-1 차단제는 항-PD-L1 항체를 포함한다. 특정한 다른 구현예에서 PD-1 차단제는 항-PD-1 항체 및 유사한 결합 단백질 예컨대 리간드 PD-L1 및 PD-L2에 의해 PD-1에 결합하고 이의 활성화를 차단하는 완전 인간 IgG4 항체인 니볼루맙(MDX 1106, BMS 936558, ONO 4538); PD-1에 대한 인간화 단클론 IgG4 항체인 람브롤리주맙(MK-3475 또는 SCH 900475); PD-1에 결합하는 인간화 항체인 CT-011; B7-DC의 융합 단백질인 AMP-224; 항체 Fc 부분; PD-L1(B7-H1) 차단을 위한 BMS-936559(MDX-1105-01)를 포함한다.

[0198] 다른 면역-체크포인트 억제제는 가용성 Ig 융합 단백질인 IMP321과 같은 림프구 활성화 유전자-3(LAG-3) 억제제를 포함한다(Brignone 등, 2007, J. Immunol. 179:4202-4211). 다른 면역-체크포인트 억제제는 B7-H3 및 B7-H4 억제제와 같은 B7 억제제를 포함한다. 특히, 항-B7-H3 항체 MGA271(Loo 등, 2012, Clin. Cancer Res. July 15 (18) 3834). 또한 TIM3(T-세포 면역글로불린 도메인 및 뮤신 도메인 3) 억제제가 포함된다(Fourcade 등, *J. Exp. Med.* 207:2175-86, 2010 및 Sakuishi 등, *J. Exp. Med.* 207:2187-94, 2010). 본원에 사용된 바와 같이, 용어 "TIM-3"은 당업계에서 일반적인 의미를 가지며 T 세포 면역글로불린 및 뮤신 도메인-함유 분자 3을 지칭한다. TIM-3의 천연 리간드는 갈렉틴 9(Gal9)이다. 따라서, 본원에 사용된 바와 같은 용어 "TIM-3 억제제"는 TIM-3의 기능을 억제할 수 있는 화합물, 물질 또는 조성물을 지칭한다. 예를 들어, 억제제는 TIM-3의 발현 또는 활성을 억제하고/하거나, TIM-3 신호전달 경로를 조절 또는 차단하고/하거나 갈렉틴-9에 대한 TIM-3의 결합을 차단할 수 있다. TIM-3에 대한 특이성을 갖는 항체는 당업계에 잘 알려져 있으며 전형적으로 WO2011/155607, WO2013/006490 및 WO2010/117057에 기재된 것들이다.

[0199] 추가적인 특정 면역 체크포인트 억제제는 아테졸리주맙, BMS-936559, 이필리무맙, MEDI0680, MEDI4736, MSB0010718C, 펨브롤리주맙, 피딜리주맙, 및 트레멜리무맙을 포함한다. 또한 WO 1998/42752; WO 2000/37504; WO 2001/014424; WO 2004/035607; US 2005/0201994; US 2002/0039581; US 2002/086014; US 5,811,097; US 5,855,887; US 5,977,318; US 6,051,227; US 6,984,720; US 6,682,736; US 6,207,156; US 6,682,736; US

7,109,003; US 7,132,281; EP1212422B1; Hurwitz 등, Proc. Natl. Acad. Sci. USA, 95(17):10067-10071 (1998); Camacho 등, *J. Clin. Oncology*, 22(145): Abstract No. 2505, 2004(항체 CP-675206); 및 Mokyr 등, *Cancer Res*, 58:5301-5304, 1998 참조.

[0200] 본 개시내용은 CD4, CD5, CD7, CD52 등에 결합하는 항체 및 다른 결합 도메인; 항체; IL1, IL2, IL6에 대한 항체; 자가반응성 T 세포 상에 특이적으로 존재하는 TCR에 대한 항체; IL4; IL10; IL12; IL13; IL1Ra; sIL1RI; sIL1RII; TNF에 대한 항체; ABCA3; ABCD1; ADA; AK2; APP; 아르기나제; 아릴술파타제 A; A1AT; CD3D; CD3E; CD3G; CD3Z; CFTR; CHD7; 키메라 항원 수용체(CAR); CIITA; CLN3; 보체 인자, CORO1A; CTLA; C1 억제제; C9ORF72; DCLRE1B; DCLRE1C; 유인 수용체; DKC1; DRB1*1501/DQB1*0602; 디스트로핀; 효소; 인자 VIII, FANC 패밀리 유전자(FancA, FancB, FancC, FancD1(BRCA2), FancD2, FancE, FancF, FancG, FancI, FancJ(BRIP1), FancL, FancM, FancN(PALB2), FancO(RAD51C), FancP(SLX4), FancQ(ERCC4), FancR(RAD51), FancS(BRCA1), FancT(UBE2T), FancU(XRCC2), FancV(MAD2L2), 및 FancW(RFWD3)); Fas L; FUS; GATA1; 글로빈 패밀리 유전자(즉, γ-글로빈); F8; 글루타미나제; HBA1; HBA2; HBB; IL7RA; JAK3; LCK; LIG4; LRRK2; NHEJ1; NLX2.1; 중화 항체; ORAI1; PARK2; PARK7; phox; PINK1; PNP; PRKDC; PSEN1; PSEN2; PTPN22; PTPRC; P53; 피루베이트 키나제; RAG1; RAG2; RFXANK; RFXAP; RFX5; RMRP; 리보솜 단백질 유전자; SFTPB; SFTPC; SOD1; 가용성 CD40; STIM1; sTNFRI; sTNFRII; SLC46A1; SNCA; TDP43; TERT; TERC; TINF2; 유비퀼린 2; WAS; WHN; ZAP70; γC; 및 본원에 기재된 다른 치료 유전자를 추가로 포함한다.

[0201] 결합 도메인의 대안적인 공급원은 다음을 포함한다: 랜덤 펩티드 라이브러리를 암호화하는 서열 또는 scTCR과 같은 대안적인 비-항체 스캐폴드의 루프 영역에서 다양한 조작된 아미노산을 암호화하는 서열(예를 들어, Lake 등, Int. Immunol. 11:745, 1999; Maynard 등, J. Immunol. Methods 306:51, 2005; US 8,361,794 참조), 피브리노겐 도메인(예를 들어, Weisel 등, *Science* 230:1388, 1985 참조), 쿠니츠(Kunitz) 도메인(예를 들어, US 6,423,498 참조), 설계된 안키린 반복 단백질(DARPins; Binz 등, *J. Mol. Biol*. 332:489, 2003 및 Binz 등, *Nat. Biotechnol*. 22:575, 2004), 피브로넥틴 결합 도메인(아드넥틴 또는 모노바디(monobody); Richards 등, *J. Mol. Biol*. 326:1475, 2003; Parker 등, Protein Eng. Des. Selec. 18:435, 2005 and Hackel 등, J. Mol. Biol. 381:1238-1252, 2008), 시스테인-노트(knot) 미니단백질(Vita 등, 1995, *Proc. Nat'l. Acad. Sci. (USA)* 92:6404-6408; Martin 등, 2002, Nat. Biotechnol. 21:71, 2002 및 Huang 등, *Structure* 13:755, 2005), 테트라트리코펩티드 반복 도메인(Main 등, *Structure* 11:497, 2003 및 Cortajarena 등, *ACS Chem. Biol*. 3:161, 2008), 류신-풍부 반복 도메인(Stumpp 등, *J. Mol. Biol*. 332:471, 2003), 리포칼린 도메인(예를 들어, WO 2006/095164, Beste 등, *Proc. Nat'l. Acad. Sci. (USA)* 96:1898, 1999 및 Schonfeld 등, *Proc. Nat'l. Acad. Sci. (USA)* 106:8198, 2009 참조), V-유사 도메인(예를 들어, US 2007/0065431 참조), C-형 렉틴 도메인(Zelensky and Gready, *FEBS J*. 272:6179, 2005; Beavil 등, *Proc. Nat'l. Acad. Sci. (USA)* 89:753, 1992 및 Sato 등, *Proc. Nat'l. Acad. Sci. (USA)* 100:7779, 2003), 항원 결합 도메인이 있는 mAb2 또는 Fc-영역(Fcab™(F-Star Biotechnology, 영국 캠브리지 소재; 예를 들어, WO 2007/098934 및 WO 2006/072620 참조), 아르마딜로 반복 단백질(예를 들어, Madhurantakam 등, *Protein Sci*. 21: 1015, 2012; WO 2009/040338 참조), 아필린(Ebersbach 등, *J. Mol. Biol*. 372: 172, 2007), 아피바디(affibody), 아비버, 노틴(knottin), 피노머(fynomer), 아트리머(atrimer), 세포독성 T-림프구 연관 단백질-4(Weidle 등, *Cancer Gen. Proteo*. 10:155, 2013), 또는 기타(Nord 등, *Protein Eng*. 8:601, 1995; Nord 등, *Nat. Biotechnol*. 15:772, 1997; Nord 등, *Euro. J. Biochem*. 268:4269, 2001; Binz 등, *Nat. Biotechnol*. 23:1257, 2005; Boersma and Pluckthun, *Curr. Opin. Biotechnol*. 22:849, 2011).

[0202] 펩티드 압타머는 단백질 스캐폴드의 양쪽 단부에 부착된 펩티드 루프(세포 마커에 특이적)를 포함한다. 이 이중 구조적 제약은 펩티드 압타머의 결합 친화도를 항체와 비슷한 수준으로 증가시킨다. 가변 루프 길이는 전형적으로 8 내지 20 개의 아미노산이고 스캐폴드는 안정되고, 가용성이고, 작고, 무독성인 임의의 단백질일 수 있다. 펩티드 압타머 선택은 효모 2-하이브리드 시스템(예를 들어, Gal4 효모-2-하이브리드 시스템), 또는 LexA 상호작용 트랩 시스템과 같은 상이한 시스템을 사용하여 이루어질 수 있다.

[0203] 특정 구현예에서, 결합 도메인은 세포 마커 CD33에 결합한다. 특정 구현예에서, CD33에 결합하는 결합 도메인은 겜투주맙, 아클리주맙, 또는 HuM195 중 하나로부터 유래된다. 특정 구현예에서 CD33 결합 도메인은 서열번호: 91을 포함하는 CDRL1 서열, 서열번호: 92를 포함하는 CDRL2 서열, 및 서열번호: 93을 포함하는 CDRL3 서열을 포함하는 가변 경쇄, 및 서열번호: 94를 포함하는 CDRH1 서열, 서열번호: 95를 포함하는 CDRH2 서열, 및 서열번호: 96을 포함하는 CDRH3 서열을 포함하는 가변 중쇄를 포함하는 인간 또는 인간화 결합 도메인이다.

[0204] 특정 구현예에서, CD33 결합 도메인은 서열번호: 97을 포함하는 CDRL1 서열, 서열번호: 98을 포함하는 CDRL2 서

열, 및 서열번호: 99를 포함하는 CDRL3 서열을 포함하는 가변 경쇄, 및 서열번호: 100을 포함하는 CDRH1 서열, 서열번호: 101을 포함하는 CDRH2 서열, 및 서열번호: 102를 포함하는 CDRH3 서열을 포함하는 가변 중쇄를 포함하는 인간 또는 인간화 scFv이다. CD33에 결합하는 결합 도메인에 관한 더 많은 정보를 위해, 미국 특허 번호 제8759494호를 참조한다.

[0205] 특정 구현예에서, 인간 CD33에 결합하는 서열은 서열 서열번호: 103을 포함하는 가변 경쇄 영역, 및 서열 서열번호: 104를 포함하는 가변 중쇄 영역을 포함한다. 특정 구현예에서, 인간 CD33에 결합하는 서열은 서열 서열번호: 103을 포함하는 가변 경쇄 영역, 및 서열 서열번호: 106을 포함하는 가변 중쇄 영역을 포함한다.

[0206] 특정 구현예에서, 결합 도메인은 전장 CD33(CD33FL)에 결합한다. 특정 구현예에서, CD33FL에 결합하는 결합 도메인은 5D12, 8F5, 1H7, 린투주맙, 또는 겜투주맙 중 적어도 하나로부터 유래된다. 특정 구현예에서, CD33FL 결합 도메인은 인간 또는 인간화이며, 서열번호: 107을 포함하는 CDRL1 서열, 서열번호: 108을 포함하는 CDRL2 서열, 서열번호: 109를 포함하는 CDRL3 서열을 포함하는 가변 경쇄), 서열번호: 110을 포함하는 CDRH1 서열, 서열번호: 111을 포함하는 CDRH2 서열, 및 서열번호: 112를 포함하는 CDRH3 서열을 포함한다. CD33FL에 결합하는 결합 도메인에 관한 더 많은 정보를 위해, PCT/US17/42264을 참조한다.

[0207] 특정 구현예에서, 인간 CD33FL에 결합하는 결합 도메인은 서열 서열번호: 113을 포함하는 가변 경쇄 영역), 및 서열 서열번호: 114를 포함하는 가변 중쇄 영역을 포함한다.

[0208] 특정 구현예에서, 결합 도메인은 세포 마커 CD33델타E2(CD33ΔE2)에 결합한다. 특정 구현예에서, CD33ΔE2에 결합하는 결합 도메인은 12B12, 4H10, 11D5, 13E11, 11D11, 또는 1H7 중 적어도 하나로부터 유래된다. 특정 구현예에서, CD33ΔE2 결합 도메인은 인간 또는 인간화이며 서열번호: 115를 포함하는 CDRL1 서열, 서열번호: 116을 포함하는 CDRL2 서열, 서열번호: 117을 포함하는 CDRL3 서열을 포함하는 가변 경쇄, 서열번호: 118을 포함하는 CDRH1 서열, 서열번호: 11을 포함하는 CDRH2 서열), 및 서열번호: 120을 포함하는 CDRH3 서열을 포함한다. CD33ΔE2에 결합하는 결합 도메인에 관한 더 많은 정보를 위해, PCT/US17/42264를 참조한다.

[0209] 특정 구현예에서, 인간 CD33ΔE2에 결합하는 서열은 서열 서열번호: 121을 포함하는 가변 경쇄 영역, 및 서열 서열번호: 122를 포함하는 가변 중쇄 영역을 포함한다.

[0210] 특정 구현예에서, 결합 도메인은 세포 마커 Her2에 결합한다. 특정 구현예에서, HER2에 결합하는 결합 도메인은 트라스투주맙(허셉틴(Herceptin))으로부터 유래된다. 특정 구현예에서, 결합 도메인은 서열번호: 12를 포함하는 CDRL1 서열), 서열번호: 124를 포함하는 CDRL2 서열, 및 서열번호: 125를 포함하는 CDRL3 서열을 포함하는 가변 경쇄, 및 서열번호: 126을 포함하는 CDRH1 서열, 서열번호: 127을 포함하는 CDRH2 서열, 및 서열번호: 128을 포함하는 CDRH3을 포함하는 가변 중쇄를 포함한다.

[0211] 특정 구현예에서, 결합 도메인은 세포 마커 PD-L1에 결합한다. 특정 구현예에서, PD-L1에 결합하는 결합 도메인은 켐브롤리주맙 또는 FAZ053(Novartis) 중 적어도 하나로부터 유래된다. 특정 구현예에서, 결합 도메인은 서열번호: 129를 포함하는 CDRL1 서열, 서열번호: 130을 포함하는 CDRL2 서열, 및 서열번호: 131을 포함하는 CDRL3 서열을 포함하는 가변 경쇄, 및 서열번호: 132를 포함하는 CDRH1 서열, 서열번호: 133을 포함하는 CDRH2 서열, 및 서열번호: 134를 포함하는 CDRH3을 포함하는 가변 중쇄를 포함한다.

[0212] PD-L1에 대한 예시적인 결합 도메인은 아벨루맙 또는 아테졸리주맙을 포함하거나 또는 이로부터 유래될 수 있다. 특정 구현예에서, 아벨루맙의 가변 중쇄는 서열번호: 135를 포함한다. 특정 구현예에서, 아벨루맙의 가변 중쇄는 서열번호: 136을 포함한다.

[0213] 특정 구현예에서, 아벨루맙의 CDR 영역은 서열번호: 137을 포함하는 CDRH1; 서열번호: 138을 포함하는 CDRH2; 서열번호: 139를 포함하는 CDRH3; 서열번호: 140을 포함하는 CDRL1; 서열번호: 141을 포함하는 CDRL2; 및 서열번호: 142를 포함하는 CDRL3을 포함한다. 특정 구현예에서, 아테졸리주맙의 가변 중쇄는 서열번호: 143을 포함한다. 특정 구현예에서, 아테졸리주맙의 가변 경쇄는 서열번호: 144를 포함한다.

[0214] 특정 구현예에서, 아테졸리주맙의 CDR 영역은 서열번호: 145를 포함하는 CDRH; 서열번호: 146을 포함하는 CDRH2; 서열번호: 147을 포함하는 CDRH3; 서열번호: 148을 포함하는 CDRL1; 서열번호: 149를 포함하는 CDRL2; 및 서열번호: 150을 포함하는 CDRL3을 포함한다.

[0215] 특정 구현예에서, 결합 도메인은 세포 마커 PSMA에 결합한다. 특정 구현예에서, 결합 도메인은 서열번호: 151을 포함하는 CDRL1 서열, 서열번호: 152를 포함하는 CDRL2 서열, 서열번호: 153을 포함하는 CDRL3 서열을 포함하는 가변 경쇄를 포함한다. 특정 구현예에서, 결합 도메인은 서열번호: 154를 포함하는 CDRH1 서열,

서열번호: 155를 포함하는 CDRH2 서열, 및 서열번호: 156을 포함하는 CDRH3 서열을 포함하는 가변 중쇄를 포함한다.

[0216] 특정 구현예에서, 결합 도메인은 세포 마커 MUC16에 결합한다. 특정 구현예에서, 결합 도메인은 인간 또는 인간화이며 서열번호: 157을 포함하는 CDRL1 서열, GAS를 포함하는 CDRL2 서열, 서열번호: 158을 포함하는 CDRL3 서열을 포함하는 가변 경쇄를 포함한다. 특정 구현예에서, 결합 도메인은 인간 또는 인간화이며 서열번호: 159를 포함하는 CDRH1 서열, 서열번호: 160을 포함하는 CDRH2 서열, 및 서열번호: 161을 포함하는 CDRH3 서열을 포함하는 가변 중쇄를 포함한다.

[0217] 특정 구현예에서, 결합 도메인은 세포 마커 FOLR에 결합한다. 특정 구현예에서, FOLR에 결합하는 결합 도메인은 팔레투주맙으로부터 유래된다. 특정 구현예에서, 결합 도메인은 서열번호: 162를 포함하는 CDRL1 서열, 서열번호: 163을 포함하는 CDRL2 서열, 및 서열번호: 164를 포함하는 CDRL3 서열을 포함하는 가변 경쇄, 및 서열번호: 165를 포함하는 CDRH1 서열, 서열번호: 166을 포함하는 CDRH2 서열, 및 서열번호: 167을 포함하는 CDRH3 서열을 포함하는 가변 중쇄를 포함한다.

[0218] 메소텔린에 대한 예시적인 결합 도메인은 아마툭시맙을 포함하거나 또는 이로부터 유래될 수 있다. 특정 구현예에서, 아마툭시맙의 가변 중쇄는 서열번호: 168을 포함한다. 특정 구현예에서, 아마툭시맙의 가변 경쇄는 서열번호: 169를 포함한다.

[0219] 특정 구현예에서, 아마툭시맙의 CDR 영역은 서열번호: 170을 포함하는 CDRH1 서열; 서열번호: 171을 포함하는 CDRH2 서열; 서열번호: 172를 포함하는 CDRH3 서열; 서열번호: 173을 포함하는 CDRL1 서열;(서열번호: 174를 포함하는 CDRL2 서열; 및 서열번호: 175를 포함하는 CDRL3 서열을 포함한다.

[0220] 특정 구현예에서, 결합 도메인은 Vα/β 및 Cα/β 쇄(예를 들어, Vα-Cα, Vβ-Cβ, Vα-Vβ)를 포함하거나 또는 관심 세포 마커에 특이적인 Vα-Cα, Vβ-Cβ, Vα-Vβ 쌍(예를 들어, 펩티드-MHC 복합체)을 포함하는 sc T 세포 수용체(scTCR)이다.

[0221] 특정 구현예에서, 결합 도메인은 알려지거나 또는 식별된 TCR Vα, Vβ, Cα, 또는 Cβ의 아미노산 서열과 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 적어도 99.5%, 또는 100% 동일한 서열을 포함하며, 여기서 각 CDR은 표적화된 세포 마커에 특이적으로 결합하는 TCR 또는 이의 단편 또는 유도체로부터 0 개의 변화 또는 최대 1, 2, 또는 3 개의 변화를 포함한다.

[0222] 특정 구현예에서, 결합 도메인은 알려지거나 또는 식별된 TCR(예를 들어, 고-친화도 TCR)의 Vα, Vβ, Cα, 및/또는 Cβ로부터 유래되거나 또는 이에 기반한 Vα, Vβ, Cα, 및/또는 Cβ 영역을 포함하고 알려지거나 또는 식별된 TCR의 Vα, Vβ, Cα, 및/또는 Cβ와 비교할 때, 하나 이상(예를 들어, 2, 3, 4, 5, 6, 7, 8, 9, 10 개)의 삽입, 하나 이상(예를 들어, 2, 3, 4, 5, 6, 7, 8, 9, 10 개)의 결실, 하나 이상(예를 들어, 2, 3, 4, 5, 6, 7, 8, 9, 10 개)의 아미노산 치환(예를 들어, 보존적 아미노산 치환 또는 비-보존적 아미노산 치환), 또는 상기 언급된 변화의 조합을 포함한다. 삽입, 결실 또는 치환은 Vα, Vβ, Cα, 및/또는 Cβ 영역의 아미노- 또는 카르복시-말단 또는 양쪽 단부를 포함하여 이러한 영역의 어디든 있을 수 있으며, 단, 각 CDR은 0 개의 변화 또는 최대 1, 2, 또는 3 개의 변화를 포함하고 변형된 Vα, Vβ, Cα, 또는 Cβ 영역을 함유하는 표적 결합 도메인이 제공되면 여전히 야생형과 유사한 친화도 및 작용으로 표적에 특이적으로 결합할 수 있다.

[0223] 특정 구현예에서, 결합 도메인은 경쇄 가변 영역(VL) 또는 중쇄 가변 영역(VH), 또는 둘 다의 아미노산 서열과 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 적어도 99.5%, 또는 100% 동일한 서열이거나 또는 이를 포함하며, 여기서 각 CDR은 관심 세포 마커에 특이적으로 결합하는 단클론 항체 또는 이의 단편 또는 유도체로부터 0 개의 변화 또는 최대 1, 2, 또는 3 개의 변화를 포함한다.

[0224] 특정 구현예에서, 본 개시내용의 결합 도메인에서 VL 영역은 알려진 단클론 항체의 VL로부터 유래되거나 또는 이를 기반으로 하고 알려진 단클론 항체의 VL과 비교할 때, 하나 이상(예를 들어, 2, 3, 4, 5, 6, 7, 8, 9, 10 개)의 삽입, 하나 이상(예를 들어, 2, 3, 4, 5, 6, 7, 8, 9, 10 개)의 결실, 하나 이상(예를 들어, 2, 3, 4, 5, 6, 7, 8, 9, 10 개)의 아미노산 치환(예를 들어, 보존적 아미노산 치환), 또는 상기-언급된 변화의 조합을 함유한다. 삽입, 결실 또는 치환은 VL 영역의 아미노- 또는 카르복시-말단 또는 둘 다의 단부를 포함하여 이 영역의 어디든 있을 수 있으며, 단, 각 CDR은 0 개의 변화 또는 최대 1, 2, 또는 3 개의 변화를 포함하고 변형된 VL 영역을 함유하는 결합 도메인이 제공되면 여전히 야생형 결합 도메인과 유사한 친화도로 표적에 특이적으

로 결합할 수 있다.

[0225] 특정 구현예에서, 본 개시내용의 결합 도메인 VH 영역은 알려진 단클론 항체의 VH로부터 유래되거나 또는 이를 기반으로 할 수 있고 알려진 단클론 항체의 VH와 비교할 때, 하나 이상(예를 들어, 2, 3, 4, 5, 6, 7, 8, 9, 10 개)의 삽입, 하나 이상(예를 들어, 2, 3, 4, 5, 6, 7, 8, 9, 10 개)의 결실, 하나 이상(예를 들어, 2, 3, 4, 5, 6, 7, 8, 9, 10 개)의 아미노산 치환(예를 들어, 보존적 아미노산 치환 또는 비-보존적 아미노산 치환), 또는 상기-언급된 변화의 조합을 함유할 수 있다. 삽입, 결실 또는 치환은 VH 영역의 아미노- 또는 카르복시-말단 또는 둘 다의 단부를 포함하여 이 영역의 어디든 있을 수 있으며, 단, 각 CDR은 0 개의 변화 또는 최대 1, 2, 또는 3 개의 변화를 포함하고 변형된 VH 영역을 함유하는 결합 도메인이 제공되면 여전히 야생형 결합 도메인과 유사한 친화도로 표적에 특이적으로 결합할 수 있다.

[0226] 주어진 CDR 또는 FR의 정확한 아미노산 서열 경계는 Kabat 등 (1991) "Sequences of Proteins of Immunological Interest," 5th Ed. Public Health Service, National Institutes of Health, Bethesda, Md.(Kabat 넘버링 체계); Al-Lazikani 등, J Mol Biol 273: 927-948, 1997(Chothia 넘버링 체계); Maccallum 등, J Mol Biol 262: 732-745, 1996(Contact 넘버링 체계); Martin 등, Proc. Natl. Acad. Sci., 86: 9268-9272, 1989(AbM 넘버링 체계); Lefranc 등, Dev Comp Immunol 27(1): 55-77, 2003(IMGT 넘버링 체계); 및 Honegger and Pluckthun, J Mol Biol 309(3): 657-670, 2001("Aho" 넘버링 체계)에 의해 기재된 것들을 포함하여 다수의 잘 알려진 체계 중 임의의 것을 사용하여 용이하게 결정될 수 있다. 주어진 CDR 또는 FR의 경계는 식별을 위해 사용되는 체계에 따라 달라질 수 있다. 예를 들어, Kabat 체계는 구조적 정렬에 기반한 반면, Chothia 체계는 구조적 정보에 기반한다. Kabat 및 Chothia 체계 둘 다에 대한 넘버링은 가장 통상적인 항체 영역 서열 길이를 기반으로 하며, 삽입은 삽입 문자, 예를 들어, "30a"로 수용되고 결실은 일부 항체에서 보인다. 두 체계는 상이한 위치에서 특정 삽입 및 결실("indel")을 배치하여, 차등 넘버링을 초래한다. Contact 체계는 복잡한 결정 구조의 분석을 기반으로 하고 Chothia 넘버링 체계에 대한 많은 측면에서 유사하다. 특정 구현예에서, 본원에 개시된 항체 CDR 서열은 Kabat 넘버링에 따른다.

[0227] 전립선암과 연관된 특정 세포 마커는 PSMA, WT1, 전립선 줄기 세포 항원(PSCA), 및 SV40 T를 포함한다. 유방암과 연관된 특정 세포 마커는 HER2 및 ERBB2를 포함한다. 난소암과 연관된 특정 세포 마커는 L1-CAM, MUC16의 세포외 도메인(MUC-CD), 폴레이트 결합 단백질(폴레이트 수용체), Lewis Y, 메소텔린, 및 WT-1을 포함한다. 췌장암과 연관된 특정 세포 마커는 메소텔린, CEA 및 CD24를 포함한다. 다발성 골수종과 연관된 특정 세포 마커는 BCMA, GPRC5D, CD38, 및 CS-1을 포함한다. 백혈병 및/또는 림프종과 연관된 특정 마커는 CLL-1, CD123, CD33, 및 PD-L1을 포함한다.

[0228] 또한 예를 들면 감염원 항원에 대한 결합에 의해 감염성 질환 제제에 특이적인 결합 도메인이 고려된다. 이들은 예를 들면 바이러스-감염된 세포에 의해 발현되는 바이러스 항원 또는 다른 바이러스 마커를 포함한다. 예시적인 바이러스는 아데노바이러스, 아레나바이러스, 부냐바이러스, 코로나바이러스, 플라비바이러스, 한타바이러스, 헤파드나바이러스, 헤르페스바이러스, 파필로마바이러스, 파라믹소바이러스, 파보바이러스, 피코르나바이러스, 폭스바이러스, 오르토믹소바이러스, 레트로바이러스, 레오바이러스, 랍도바이러스, 로타바이러스, 스폰지형 바이러스 또는 토가바이러스를 포함한다. 추가적인 구현예에서, 바이러스 항원 마커는 CMV, 감기 바이러스, 엡스타인-바(Epstein-Barr), 독감 바이러스, A형, B형, 및 C형 간염 바이러스, 단순 포진, HIV, 인플루엔자, 일본 뇌염, 홍역, 소아마비, 광견병, 호흡기 세포융합체, 풍진, 천연두, 수두 대상포진 또는 웨스트 나일(West Nile) 바이러스에 의해 발현되는 펩티드를 포함한다.

[0229] 추가의 특정 예로서, 사이토메갈로바이러스 항원은 외피 당단백질 B 및 CMV pp65를 포함하고; 엡스타인-바 항원은 EBV EBNAI, EBV P18, 및 EBV P23을 포함하고; 간염 항원은 HBV의 S, M, 및 L 단백질, HBV의 프리-S 항원, HBCAG DELTA, HBV HBE, C형 간염 바이러스 RNA, HCV NS3 및 HCV NS4를 포함하고; 단순 포진 바이러스 항원은 급초기 단백질 및 당단백질 D를 포함하고; HIV 항원은 gag, pol, 및 env 유전자의 유전자 산물 예컨대 HIV gp32, HIV gp41, HIV gp120, HIV gp160, HIV P17/24, HIV P24, HIV P55 GAG, HIV P66 POL, HIV TAT, HIV GP36, Nef 단백질 및 역전사효소를 포함하고; 인플루엔자 항원은 헤마글루티닌 및 뉴라미니다제를 포함하고; 일본 뇌염 바이러스 항원은 단백질 E, M-E, M-E-NS1, NS1, NS1-NS2A 및 80% E를 포함하고; 홍역 항원은 홍역 바이러스 융합 단백질을 포함하고; 광견병 항원은 광견병 당단백질 및 광견병 핵단백질을 포함하고; 호흡기 세포융합 바이러스 항원은 RSV 융합 단백질 및 M2 단백질을 포함하고; 로타바이러스 항원은 VP7sc를 포함하고; 풍진 항원은 단백질 E1 및 E2를 포함하고; 수두 대상포진 바이러스 항원은 gpI 및 gpII를 포함한다.

[0230] 추가적인 특정 예시적인 바이러스 항원 서열은 다음을 포함한다: Nef(66-97)(서열번호: 176), Nef(116-145)(서

열번호: 177), Gag p17(17-35)(서열번호: 178), Gag p17-p24(253-284)(서열번호: 179), 및 Pol 325-355(RT 158-188)(서열번호: 180). 바이러스 항원의 추가적인 예에 대하여 Fundamental Virology, Second Edition, eds. Fields, B. N. and Knipe, D. M. (Raven Press, New York, 1991) 참조.

[0231] 암 세포와 같은 원치않은 세포 유형을 표적화하고 사멸하기 위해 면역계의 T 세포를 유전적으로 조작하는 데 상당한 진전이 이루어졌다. 많은 이들 T 세포는 키메라 항원 수용체(CAR) 작제물을 발현하도록 유전적으로 조작되었다. CAR은 암 세포를 인식하고 사멸시키기 위해 T 세포를 유전적으로 변형시키는 여러 별개의 하위구성요소를 포함하는 단백질이다. 하위구성요소는 적어도 세포외 구성요소 및 세포내 구성요소를 포함한다.

[0232] 세포외 구성요소는 원치않은 세포의 표면 상에 우선적으로 존재하는 마커에 특이적으로 결합하는 결합 도메인을 포함한다. 결합 도메인이 이러한 마커에 결합할 때, 세포내 구성요소는 결합된 암 세포를 파괴하도록 T 세포를 지시한다. 결합 도메인은 전형적으로 단클론 항체(mAb)로부터 유래된 단일-쇄 가변 단편(scFv)이지만, 이는 항체-유사 항원 결합 부위를 포함하는 다른 형식을 기반으로 할 수 있다.

[0233] 세포내 구성요소는 효과기 도메인의 포함을 기반으로 활성화 신호를 제공한다. 1세대 CAR은 CD3ζ의 세포질 영역을 효과기 도메인으로 활용하였다. 2세대 CAR은 분화 클러스터 28(CD28) 또는 4-1BB(CD137)와 조합하여 CD3ζ을 활용한 반면, 3세대 CAR은 세포내 효과기 도메인 내에서 CD28 및 401BB와 조합하여 CD3ζ를 활용하였다.

[0234] CAR은 일반적으로 또한 분자 내에서 다양한 목적을 위해 사용되는 하나 이상의 링커 서열을 포함한다. 예를 들어, 막관통 도메인은 CAR의 세포외 구성요소를 세포내 구성요소에 연결하기 위해 사용될 수 있다. 결합 도메인에 막-근위에 있는 스페이서 영역으로 종종 지칭되는 유연한 링커 서열은 결합 도메인 및 세포 막 사이의 추가적인 거리를 생성하기 위해 사용될 수 있다. 이는 막에 대한 근접성을 기반으로 결합에 대한 입체 장애를 감소시키는 데 유익할 수 있다. 이 목적을 위해 사용되는 통상적인 스페이서 영역은 IgG4 링커이다. 표적화된 세포 마커에 따라, 더 조밀한 스페이서 또는 더 긴 스페이서가 사용될 수 있다. 다른 잠재적인 CAR 하위구성요소는 본원의 다른 곳에 보다 상세하게 기재된다. CAR의 구성요소는 이제 다음과 같이 추가로 상세하게 기재된다: (a) 결합 도메인; (b) 세포내 신호전달 구성요소; (c) 링커; (d) 막관통 도메인; (e) 접합부 아미노산; 및 (f) 태그 카세트를 포함하는 제어 특징.

[0235] *(a) 결합 도메인.* 결합 도메인은 세포 마커에 결합하여 복합체를 형성하는 물질을 포함하며, 본원에 개시된 모든 결합 도메인 및 항체를 포함하나 이에 제한되지 않는다. 결합 도메인의 선택은 표적 세포의 표면을 정의하는 세포 마커의 유형 및 수에 따라 달라질 수 있다. 결합 도메인의 예는 세포 마커 리간드, 수용체 리간드, 항체, 펩티드, 펩티드 압타머, 수용체(예를 들어, T 세포 수용체), 또는 이의 조합 및 조작된 단편 또는 형식을 포함한다.

[0236] *(b) 세포내 신호전달 구성요소.* CAR의 세포내 또는 달리 세포질 신호전달 구성요소는 CAR이 발현되는 세포의 활성화를 담당한다. 따라서 용어 "세포내 신호전달 구성요소" 또는 "세포내 구성요소"는 활성화 신호를 형질도입하기에 충분한 세포내 도메인의 임의의 부분을 포함하는 것을 의미한다. 발현된 CAR의 세포내 구성요소는 효과기 도메인을 포함한다. 효과기 도메인은 적절한 신호를 수신할 때 세포에서 생물학적 또는 생리학적 반응을 직접적으로 또는 간접적으로 촉진할 수 있는 융합 단백질 또는 수용체의 세포내 부분이다. 특정 구현예에서, 효과기 도메인은 결합될 때 신호를 수신하는 단백질 또는 단백질 복합체의 일부이거나, 또는 효과기 도메인로부터 신호를 촉발하는 표적 분자에 직접적으로 결합된다. 효과기 도메인은 면역수용체 티로신-기반 활성화 모티프(ITAM)와 같은 하나 이상의 신호전달 도메인 또는 모티프를 함유할 때 세포 반응을 직접적으로 촉진할 수 있다. 다른 구현예에서, 효과기 도메인은 공동-자극 도메인과 같이, 세포 반응을 직접적으로 촉진하는 하나 이상의 다른 단백질과 회합함으로써 세포 반응을 간접적으로 촉진할 것이다.

[0237] 효과기 도메인은 암 세포에 의해 발현되는 세포 마커에 결합 시 변형된 세포의 적어도 하나의 기능의 활성화를 제공할 수 있다. 변형된 세포의 활성화는 분화, 증식 및/또는 활성화 또는 다른 효과기 기능 중 하나 이상을 포함할 수 있다. 특정 구현예에서, 효과기 도메인은 T 세포 수용체 및 공동-수용체 또는 공동-자극 분자로부터의 세포질 서열을 포함할 수 있는 공동-자극 도메인을 포함하는 세포내 신호전달 구성요소를 포함할 수 있다.

[0238] 효과기 도메인은 1, 2, 3 개 또는 그 이상의 수용체 신호전달 도메인, 세포내 신호전달 구성요소(예를 들어, 세포질 신호전달 서열), 공동-자극 도메인, 또는 이의 조합을 포함할 수 있다. 예시적인 효과기 도메인은 4-1BB(CD137), CARD11, CD3ɣ, CD3δ, CD3ε, CD3ζ, CD27, CD28, CD79A, CD79B, DAP10, FcRα, FcRβ(FcεR1b), FcRɣ, Fyn, HVEM(LIGHTR), ICOS, LAG3, LAT, Lck, LRP, NKG2D, NOTCH1, pTα, PTCH2, OX40, ROR2, Ryk, SLAMF1, Slp76, TCRα, TCRβ, TRIM, Wnt, Zap70, 또는 이의 임의의 조합으로부터 선택된 신호전달 및 자

극 도메인을 포함한다. 특정 구현예에서, 예시적인 효과기 도메인은 CD86, FcγRIIa, DAP12, CD30, CD40, PD-1, 림프구 기능-연관 항원-1(LFA-1), CD2, CD7, LIGHT, NKG2C, B7-H3, CD83에 특이적으로 결합하는 리간드, CDS, ICAM-1, GITR, BAFFR, SLAMF7, NKp80(KLRF1), CD127, CD160, CD19, CD4, CD8α, CD8β, IL2Rβ, IL2Rγ, IL7Rα, ITGA4, VLA1, CD49a, IA4, CD49D, ITGA6, VLA-6, CD49f, ITGAD, CD11d, ITGAE, CD103, ITGAL, CD11a, ITGAM, CD11b, ITGAX, CD11c, ITGB1, CD29, ITGB2, CD18, ITGB7, TNFR2, TRANCE/RANKL, DNAM1(CD226), SLAMF4(CD244, 2B4), CD84, CD96(Tactile), CEACAM1, CRTAM, Ly9(CD229), PSGL1, CD100(SEMA4D), CD69, SLAMF6(NTB-A, Ly108), SLAM(CD150, IPO-3), BLAME(SLAMF8), SELPLG(CD162), LTBR, GADS, PAG/Cbp, NKp44, NKp30, 또는 NKp46으로부터 선택된 신호전달 및 공동-자극 도메인을 포함한다.

[0239] 자극 방식으로 작용하는 세포내 신호전달 구성요소 서열은 iTAM을 포함할 수 있다. 1차 세포질 신호전달 서열을 포함하는 iTAM의 예는 CD3γ, CD3δ, CD3ε, CD3ζ, CD5, CD22, CD66d, CD79a, CD79b, 및 공통 FcRγ(FCER1G), FcγRIIa, FcRβ(Fcε Rib), DAP10, 및 DAP12로부터 유래된 것들을 포함한다. 특정 구현예에서,CD3ζ의 변이체는 적어도 하나, 2, 3 개, 또는 모든 ITAM 영역을 보유한다.

[0240] 특정 구현예에서, 효과기 도메인은 세포질 신호전달 단백질과 회합하는 세포질 부분을 포함하며, 여기서 세포질 신호전달 단백질은 림프구 수용체 또는 이의 신호전달 도메인, 복수의 ITAM을 포함하는 단백질, 공동-자극 도메인, 또는 이의 임의의 조합이다.

[0241] 세포내 신호전달 구성요소의 추가적인 예는 CD3ζ 쇄의 세포질 서열, 및/또는 결합 도메인 관여 후 신호 형질도입을 개시하기 위해 협력하여 작용하는 공동-수용체를 포함한다.

[0242] 공동-자극 도메인은 세포 마커 결합에 대한 효율적인 림프구 반응에 필요할 수 있는 활성화가 있는 도메인이다. 일부 분자는 세포내 신호전달 구성요소 또는 공동-자극 도메인으로 상호교환가능하다. 공동 자극 도메인의 예는 CD27, CD28, 4-1BB(CD 137), OX40, CD30, CD40, PD-1, ICOS, 림프구 기능-연관 항원-1(LFA-1), CD2, CD7, LIGHT, NKG2C, B7-H3, 및 CD83과 특이적으로 결합하는 리간드를 포함한다. 예를 들어, CD27 공동-자극은 시험관내에서 인간 CART 세포의 확장, 효과기 기능, 및 생존을 향상시키고 생체내에서 인간 T 세포 지속 및 항암 활성을 증가시키는 것으로 입증되었다(Song 등 Blood. 2012; 119(3):696-706). 이러한 공동-자극 도메인 분자의 추가의 예는 CDS, ICAM-1, GITR, BAFFR, HVEM(LIGHTR), SLAMF7, NKp80(KLRF1), NKp44, NKp30, NKp46, CD160, CD19, CD4, CD8α, CD8β, IL2Rβ, IL2Rγ, IL7Rα, ITGA4, VLA1, CD49a, ITGA4, IA4, CD49D, ITGA6, VLA-6, CD49f, ITGAD, CD11d, ITGAE, CD103, ITGAL, CD11a, ITGAM, CD1 1b, ITGAX, CD11c, ITGB1, CD29, ITGB2, CD18, ITGB7, TNFR2, TRANCE/RANKL, DNAM1(CD226), SLAMF4 (CD244, 2B4), CD84, CD96(Tactile), NKG2D, CEACAM1, CRTAM, Ly9(CD229), PSGL1, CD100(SEMA4D), CD69, SLAMF6(NTB-A, Ly108), SLAM(SLAMF1, CD150, IPO-3), BLAME(SLAMF8), SELPLG(CD162), LTBR, LAT, GADS, SLP-76, PAG/Cbp, 및 CD19a를 포함한다.

[0243] 특정 구현예에서, 세포내 신호전달 구성요소의 아미노산 서열은 CD3ζ의 변이체 및 4-1BB 세포내 신호전달 구성요소의 일부를 포함한다.

[0244] 특정 구현예에서, 세포내 신호전달 구성요소는 (i) CD3ζ의 신호전달 도메인의 전부 또는 일부, (ii) 4-1BB의 신호전달 도메인의 전부 또는 일부, 또는 (iii) CD3ζ 및 4-1BB의 신호전달 도메인의 전부 또는 일부를 포함한다.

[0245] 세포내 구성요소는 또한 Wnt 신호전달 경로(예를 들어, LRP, Ryk, 또는 ROR2), NOTCH 신호전달 경로(예를 들어, NOTCH1, NOTCH2, NOTCH3, 또는 NOTCH4), Hedgehog 신호전달 경로(예를 들어, PTCH 또는 SMO), 수용체 티로신 키나제(RTK)(예를 들어, 표피 성장 인자(EGF) 수용체 패밀리, 섬유모세포 성장 인자(FGF) 수용체 패밀리, 간세포 성장 인자(HGF) 수용체 패밀리, 인슐린 수용체(IR) 패밀리, 혈소판-유래 성장 인자(PDGF) 수용체 패밀리, 혈관 내피 성장 인자(VEGF) 수용체 패밀리, 트로포마이신 수용체 키나제(Trk) 수용체 패밀리, 에프린(Eph) 수용체 패밀리, AXL 수용체 패밀리, 백혈구 티로신 키나제(LTK) 수용체 패밀리, 면역글로불린-유사 및 EGF-유사 도메인 1을 함유하는 티로신 키나제(TIE) 수용체 패밀리, 수용체 티로신 키나제-유사 고아(ROR) 수용체 패밀리, 디스코이딘 도메인(DDR) 수용체 패밀리, 형질감염 동안 재배열된(RET) 수용체 패밀리, 티로신-단백질 키나제-유사(PTK7) 수용체 패밀리, 수용체 관련된 티로신 키나제(RYK) 수용체 패밀리, 또는 근육 특이적 키나제(MuSK) 수용체 패밀리); G-단백질-커플링된 수용체, GPCR(Frizzled 또는 Smoothened); 세린/트레오닌 키나제 수용체(BMPR 또는 TGFR); 또는 사이토카인 수용체(IL1R, IL2R, IL7R, 또는 IL15R)의 단백질 중 하나 이상을 포함할 수 있다.

[0246] *(c) 링커.* 본원에 사용된 바와 같이, 링커는 분자의 2 개의 다른 하위 구성요소를 연결하는 역할을 하는 CAR 분자의 임의의 부분일 수 있다. 일부 링커는 다른 구성요소를 연결하는 것 이외의 목적을 제공하지 않는 반면 많

은 링커는 추가적인 목적을 제공한다. scFv의 항체 유래된 결합 도메인의 VL 및 VH를 연결하는 맥락에서 링커는 상기 기재되어 있다. 링커는 또한 스페이서 영역, 및 접합부 아미노산을 포함할 수 있다.

[0247] 스페이서 영역은 다른 연결된 구성요소로부터 적절한 거리 및/또는 유연성을 생성하는 데 사용되는 링커 영역의 한 유형이다. 특정 구현예에서, 스페이서 영역의 길이는 원치않은 세포 인식 및 파괴를 최적화하기 위해 원치 않은 세포에 대한 개별 세포 마커를 맞춤화할 수 있다. 스페이서는 스페이서의 부재와 비교하여, 항원 결합 후 세포의 증가된 반응성을 제공하는 길이일 수 있다. 특정 구현예에서, 스페이서 영역 길이는 세포 마커 에피토프의 위치, 에피토프에 대한 결합 도메인의 친화도, 및/또는 세포 마커 인식에 반응하여 시험관내 및/또는 생체내에서 증식하는 분자를 발현하는 변형된 세포의 능력을 기반으로 선택될 수 있다. 스페이서 영역은 또한 변형된 세포에서 높은 발현 수준을 허용할 수 있다.

[0248] 예시적인 스페이서는 10 내지 250 개의 아미노산, 10 내지 200 개의 아미노산, 10 내지 150 개의 아미노산, 10 내지 100 개의 아미노산, 10 내지 50 개의 아미노산, 또는 10 내지 25 개의 아미노산을 갖는 것을 포함한다. 특정 구현예에서, 스페이서 영역은 12 개의 아미노산, 20 개의 아미노산, 21 개의 아미노산, 26 개의 아미노산, 27 개의 아미노산, 45 개의 아미노산, 또는 50 개의 아미노산이다.

[0249] 특정 구현예에서, 스페이서 영역은 IgG1, IgG2, lgG3, lgG4 또는 IgD 단독으로부터 힌지 영역 서열의 전부 또는 일부 또는 CH2 영역의 전부 또는 일부; CH3 영역의 전부 또는 일부; 또는 CH2 영역의 전부 또는 일부 및 CH3 영역의 전부 또는 일부와 조합하여 포함하는 군으로부터 선택된다.

[0250] 예시적인 스페이서는 IgG4 힌지 단독, CH2 및 CH3 도메인에 연결된 IgG4 힌지, 또는 CH3 도메인에 연결된 IgG4 힌지를 포함한다. 특정 구현예에서, 스페이서는 아미노산 서열 서열번호: 181의 IgG4 링커를 포함한다. 힌지 영역은 의도되지 않은 파트너와의 이량체화와 같은 바람직하지 않은 구조적 상호작용을 피하기 위해 변형될 수 있다.

[0251] 특정 구현예에서, 스페이서 영역은 II형 C-렉틴 도메인간(줄기) 영역 또는 분화 클러스터(CD) 분자 줄기 영역인 힌지 영역을 포함한다. 본원에 사용된 바와 같이, "야생형 면역글로불린 힌지 영역"은 CH1 및 CH2 도메인 사이에 개재되어 연결하거나(IgG, IgA, 및 IgD의 경우) 또는 항체의 중쇄에서 발견되는 CH1 및 CH3 도메인 사이에 개재되어 연결하는(IgE 및 IgM의 경우) 자연 발생 상부 및 중간 힌지 아미노산을 지칭한다.

[0252] II형 C-렉틴 또는 CD 분자의 "줄기 영역"은 C-형 렉틴-유사 도메인(CTLD; 예를 들어, 자연 살해자 세포 수용체의 CTLD와 유사) 및 소수성 부분(막관통 도메인) 사이에 위치한 II형 C-렉틴 또는 CD 분자의 세포외 도메인 부분을 지칭한다. 예를 들어, 인간 CD94의 세포외 도메인(GenBank 수탁 번호 AAC50291.1)은 아미노산 잔기 34-179에 상응하지만, CTLD는 아미노산 잔기 61-176에 상응하여, 인간 CD94 분자의 줄기 영역은 소수성 부분(막관통 도메인) 및 CTLD 사이에 위치한 아미노산 잔기 34-60을 포함한다(Boyington 등, *Immunity* 10:15, 1999 참조; 다른 줄기 영역의 설명을 위해, 또한 Beavil 등, *Proc. Nat'l. Acad. Sci. USA* 89:153, 1992; 및 Figdor 등, *Nat. Rev. Immunol.* 2:11, 2002 참조). 이러한 II형 C-렉틴 또는 CD 분자는 또한 줄기 영역 및 막관통 영역 또는 CTLD 사이에 접합부 아미노산(하기 기재됨)을 가질 수 있다. 또 다른 예에서, 233 개의 아미노산 인간 NKG2A 단백질(GenBank 수탁 번호 P26715.1)은 아미노산 71-93 범위의 소수성 부분(막관통 도메인) 및 아미노산 94-233 범위의 세포외 도메인을 갖는다. CTLD는 아미노산 119-231을 포함하고 줄기 영역은 아미노산 99-116을 포함하며, 이는 추가적인 접합부 아미노산에 의해 플랭킹될 수 있다. 다른 II형 C-렉틴 또는 CD 분자, 뿐만 아니라 이의 세포외 리간드-결합 도메인, 줄기 영역, 및 CTLD는 당업계에 알려져 있다(예를 들어, GenBank 수탁 번호 NP 001993.2; AAH07037.1; NP 001773.1; AAL65234.1; CAA04925.1 참조; 각각 인간 CD23, CD69, CD72, NKG2A, 및 NKG2D의 서열 및 이의 설명에 대한 것).

[0253] 예시적인 스페이서는 또한 Hudecek 등 *(Clin. Cancer Res.*, 19:3153, 2013) 또는 WO2014/031687에 기재된 것들을 포함한다. 특정 구현예에서, 스페이서 영역은 아미노산 서열 서열번호: 182의 CD28 링커일 수 있다. 특정 구현예에서, 스페이서 영역은 서열번호: 183이다. 특정 구현예에서, 스페이서 영역은 서열번호: 184이다.

[0254] 특정 구현예에서, 긴 스페이서는 119 개 초과의 아미노산(예를 들어, 229 개의 아미노산)이고, 중간 스페이서는 13-119 개의 아미노산이고, 짧은 스페이서는 12 개 이하의 아미노산이다. 중간 스페이서 영역의 예는 IgG4 힌지 영역 서열 및 CH3 영역의 전부 또는 일부를 포함한다. 긴 스페이서의 예는 IgG4 힌지 영역 서열, CH2 영역, 및 CH3 영역의 전부 또는 일부를 포함한다. 본 개시내용의 특정 구현예에서, 짧은 스페이서 서열이 바람직하다.

[0255] 스페이서 영역에 관한 추가의 설명으로서, 융합 단백질의 세포외 구성요소는 임의적으로 예를 들어, 적절한 세

포/세포 접촉, 항원 결합 및 활성화를 가능하게 하기 위해 결합 도메인이 숙주 세포(예를 들어, T 세포) 표면으로부터 멀리 떨어져 위치할 수 있는 세포외, 비-신호전달 스페이서 또는 링커 영역을 포함한다(Patel 등, Gene Therapy 6: 412-419 (1999)). 나타낸 바와 같이, 융합 결합 단백질의 세포외 스페이서 영역은 일반적으로 소수성 부분 또는 막관통 도메인 및 세포외 결합 도메인 사이에 위치하고, 스페이서 영역 길이는 선택된 표적 분자, 선택된 결합 에피토프, 또는 항원-결합 도메인 크기 및 친화도를 기반으로 항원 인식(예를 들어, 종양 인식)을 최대화하도록 달라질 수 있다(예를 들어, Guest 등, *J. Immunother*. 28:203-11, 2005; WO 2014/031687 참조). 특정 구현예에서, 스페이서 영역은 면역글로불린 힌지 영역을 포함한다. 면역글로불린 힌지 영역은 야생형 면역글로불린 힌지 영역 또는 변경된 야생형 면역글로불린 힌지 영역일 수 있다. 특정 구현예에서, 면역글로불린 힌지 영역은 인간 면역글로불린 힌지 영역이다. 면역글로불린 힌지 영역은 IgG, IgA, IgD, IgE, 또는 IgM 힌지 영역일 수 있다. IgG 힌지 영역은 IgG1, IgG2, IgG3, 또는 IgG4 힌지 영역일 수 있다. 예시적인 변경된 IgG4 힌지 영역은 PCT 공개 번호 WO 2014/031687에 기재되어 있다. 본원에 기재된 융합 결합 단백질에 사용되는 힌지 영역의 다른 예는 야생형 또는 이의 변이체일 수 있는 CD8α, CD4, CD28 및 CD7과 같은 1형 막 단백질의 세포외 영역에 존재하는 힌지 영역을 포함한다.

[0256] 특정 구현예에서, 세포외 스페이서 영역은 CH1 도메인, CH2 도메인, CH3 도메인, CH4 도메인, 또는 이의 임의의 조합으로부터 선택된 Fc 도메인의 전부 또는 일부를 포함한다(예를 들어, WO 2014/031687 참조). Fc 도메인 또는 이의 부분은 (예를 들어, 항체 효과기 기능을 감소시키기 위해) 변경된 야생형일 수 있다. 특정 구현예에서, 세포외 구성요소는 결합 도메인 및 소수성 부분 사이에 배치된 면역글로불린 힌지 영역, CH2 도메인, CH3 도메인, 또는 이의 임의의 조합을 포함한다. 특정 구현예에서, 세포외 구성요소는 IgG1 힌지 영역, IgG1 CH2 도메인, 및 IgG1 CH3 도메인을 포함한다. 추가의 구현예에서, IgG1 CH2 도메인은 (i) N297Q 돌연변이, (ii) 첫번째 6 개 아미노산(APEFLG)의 APPVA로의 치환, 또는 (i) 및 (ii) 둘 다를 포함한다. 특정 구현예에서, 면역글로불린 힌지 영역, Fc 도메인 또는 이의 부분, 또는 둘 다는 인간이다.

[0257] *(d) 막관통 도메인.* 나타낸 바와 같이, CAR 분자 내의 막관통 도메인은 종종 세포 막을 통해 세포외 구성요소 및 세포내 구성요소를 연결하는 역할을 한다. 막관통 도메인은 변형된 세포의 막에서 발현된 분자를 고정시킬 수 있다.

[0258] 막관통 도메인은 천연 및/또는 합성 공급원으로부터 유래될 수 있다. 공급원이 천연일 때, 막관통 도메인은 임의의 막-결합 또는 막관통 단백질으로부터 유래될 수 있다. 막관통 도메인은 적어도 T-세포 수용체, CD28, CD27, CD3 엡실론, CD45, CD4, CD5, CD8, CD9, CD16, CD22; CD33, CD37, CD64, CD80, CD86, CD134, CD137 및 CD154의 α, β 또는 ζ 쇄의 막관통 영역(들)을 포함할 수 있다. 특정 구현예에서, 막관통 도메인은 적어도 예를 들어, KIRDS2, OX40, CD2, CD27, LFA-1(CD 11a, CD18), ICOS(CD278), 4-1BB(CD137), GITR, CD40, BAFFR, HVEM(LIGHTR), SLAMF7, NKp80(KLRF1), NKp44, NKp30, NKp46, CD160, CD19, IL2Rβ, IL2Rγ, IL7R a, ITGA1, VLA1, CD49a, ITGA4, IA4, CD49D, ITGA6, VLA-6, CD49f, ITGAD, CD1 ld, ITGAE, CD103, ITGAL, CD1 la, ITGAM, CD1 lb, ITGAX, CD1 lc, ITGB1, CD29, ITGB2, CD18, ITGB7, TNFR2, DNAM1(CD226), SLAMF4 (CD244, 2B4), CD84, CD96(Tactile), CEACAM1, CRT AM, Ly9(CD229), PSGL1, CD100(SEMA4D), SLAMF6(NTB-A, Ly108), SLAM(SLAMF1, CD150, IPO-3), BLAME(SLAMF8), SELPLG(CD162), LTBR, PAG/Cbp, NKG2D, 또는 NKG2C의 막관통 영역(들)을 포함할 수 있다. 특정 구현예에서, 인간 Ig(면역글로불린) 힌지(예를 들어, IgG4 힌지, IgD 힌지), GS 링커(예를 들어, 본원에 기재된 GS 링커), KIR2DS2 힌지 또는 CD8a 힌지를 포함하는 다양한 인간 힌지가 또한 이용될 수 있다.

[0259] 특정 구현예에서, 막관통 도메인은 세포 막에서 열역학적으로 안정되고, 일반적으로 15 내지 30 개의 아미노산 길이 범위인 3차원 구조를 갖는다. 막관통 도메인의 구조는 α 나선, β 베럴, β 시트, β 나선, 또는 이의 임의의 조합을 포함할 수 있다.

[0260] 막관통 도메인은 막관통 영역에 인접한 하나 이상의 추가적인 아미노산, 예를 들어, CAR의 세포외 영역 내에서 하나 이상의 아미노산(예를 들어, 세포외 영역의 최대 15 개 아미노산) 및/또는 CAR의 세포내 영역 내에서 하나 이상의 추가적인 아미노산(예를 들어, 세포내 구성요소의 최대 15 개 아미노산)을 포함할 수 있다. 일 측면에서, 막관통 도메인은 신호전달 도메인, 공동-자극 도메인 또는 힌지 도메인이 유래된 것과 동일한 단백질로부터 유래된다. 또 다른 측면에서, 막관통 도메인은 CAR의 임의의 다른 도메인이 유래된 것과 동일한 단백질로부터 유래되지 않는다. 일부 경우에, 막관통 도메인은 수용체 복합체의 다른 의도되지 않은 구성원과의 상호작용을 최소화하기 위해 동일하거나 또는 상이한 표면 막 단백질의 막관통 도메인에 이러한 도메인의 결합을 피하기 위해 아미노산 치환에 의해 선택되거나 또는 변형될 수 있다. 일 측면에서, 막관통 도메인은 CAR-발현 세포의 세포 표면 상에서 또 다른 CAR과 동종이량체화할 수 있다. 상이한 측면에서, 막관통 도메인의 아미노산 서열은

동일한 CAR-발현 세포에 존재하는 천연 결합 파트너의 결합 도메인과의 상호작용을 최소화하기 위해 변형되거나 또는 치환될 수 있다. 특정 구현예에서, 막관통 도메인은 CD28 막관통 도메인의 아미노산 서열을 포함한다.

[0261] *(e) 접합부 아미노산.* 접합부 아미노산은 스페이서에 의해 제공되는 거리가 필요하지 않고/하거나 원치 않을 때 CAR 도메인의 서열을 연결하는 데 사용될 수 있는 링커일 수 있다. 접합부 아미노산 공동-자극 세포내 신호 전달 구성요소를 연결하는 데 사용될 수 있는 짧은 아미노산 서열이다. 특정 구현예에서, 접합부 아미노산은 9 개 이하의 아미노산이다.

[0262] 접합부 아미노산은 링커를 형성하는 짧은 올리고- 또는 단백질 링커, 바람직하게는 2 내지 9 개의 아미노산(예를 들어, 2, 3, 4, 5, 6, 7, 8, 또는 9 개의 아미노산) 길이일 수 있다. 특정 구현예에서, 글리신-세린 이중항이 적합한 접합부 아미노산 링커로서 사용될 수 있다. 특정 구현예에서, 단일 아미노산, 예를 들어, 알라닌, 글리신이 적합한 접합부 아미노산으로 사용될 수 있다.

[0263] *(f) 태그 카세트, 형질도입 마커, 및 자살 스위치를 포함하는 제어 특징.* 특정 구현예에서, CAR 작제물은 하나 이상의 태그 카세트, 형질도입 마커, 및/또는 자살 스위치를 포함할 수 있다. 일부 구현예에서, 형질도입 마커 및/또는 자살 스위치는 동일한 작제물 내에 있지만 세포 표면 상에서 별개의 분자로 발현된다. 태그 카세트 및 형질도입 마커는 시험관내, 생체내 및/또는 생체외에서 유전적으로 변형된 세포를 활성화, 이의 증식 촉진, 검출, 이에 대한 강화, 단리, 추적, 고갈 및/또는 제거에 사용될 수 있다. "태그 카세트"는 결합 특성이 태그된 단백질 및/또는 태그된 단백질을 발현하는 세포를 활성화, 이의 증식 촉진, 검출, 이에 대한 강화, 단리, 추적, 고갈 및/또는 제거에 사용될 수 있는 경우 동족 결합 분자(예를 들어, 리간드, 항체, 또는 다른 결합 파트너)가 특이적으로 결합할 수 있는 CAR에 부착되거나, 융합되거나, 또는 이의 일부인 고유한 합성 펩티드 서열을 지칭한다. 형질도입 마커는 동일한 목적을 제공할 수 있지만 자연 발생 분자로부터 유래되고 종종 나머지 CAR 분자로부터 형질도입 마커를 분리하는 스킵 요소를 사용하여 발현된다.

[0264] 동족 결합 분자에 결합하는 태그 카세트는 예를 들어, His 태그, Flag 태그, Xpress 태그, Avi 태그, 칼모듈린 태그, 폴리글루타메이트 태그, HA 태그, Myc 태그, Softag 1, Softag 3, 및 V5 태그를 포함한다. 특정 구현예에서, CAR은 Myc 태그를 포함한다.

[0265] 본원에 개시된 태그 카세트 서열에 특이적으로 결합하는 접합체 결합 분자는 상업적으로 이용가능하다. 예를 들어, His 태그 항체는 Life Technologies, Pierce Antibodies, 및 GenScript를 포함하는 공급처로부터 상업적으로 이용가능하다. Flag 태그 항체는 Pierce Antibodies, GenScript, 및 Sigma-Aldrich를 포함하는 공급처로부터 상업적으로 이용가능하다. Xpress 태그 항체는 Pierce Antibodies, Life Technologies 및 GenScript를 포함하는 공급처로부터 상업적으로 이용가능하다. Avi 태그 항체는 Pierce Antibodies, IsBio, 및 Genecopoeia를 포함하는 공급처로부터 상업적으로 이용가능하다. 칼모듈린 태그 항체는 Santa Cruz Biotechnology, Abcam, 및 Pierce Antibodies를 포함하는 공급처로부터 상업적으로 이용가능하다. HA 태그 항체는 Pierce Antibodies, Cell Signal 및 Abcam을 포함하는 공급처로부터 상업적으로 이용가능하다. Myc 태그 항체는 Santa Cruz Biotechnology, Abcam, 및 Cell Signal을 포함하는 공급처로부터 상업적으로 이용가능하다.

[0266] 형질도입 마커는 절두된 CD19(tCD19; Budde 등, Blood 122: 1660, 2013 참조); 절두된 인간 EGFR(tEGFR; Wang 등, Blood 118: 1255, 2011 참조); 인간 CD34의 세포외 도메인; 및/또는 CD34의 표적 에피토프와 조합하는 RQR8(Fehse 등, *Mol. Therapy* 1(5 Pt 1); 448-456, 2000 참조) 및 CD20 항원(Philip 등, *Blood* 124: 1277-1278, 2014 참조) 중 적어도 하나로부터 선택될 수 있다.

[0267] 특정 구현예에서, iCaspase9 작제물(iCasp9)을 암호화하는 폴리뉴클레오티드는 자살 스위치로서 CAR 뉴클레오티드 작제물 내로 삽입될 수 있다.

[0268] 제어 특징은 CAR의 다중 카피에 존재할 수 있거나 또는 스키핑 요소를 사용하여 별개의 분자로서 표현될 수 있다. 예를 들어, CAR은 1, 2, 3, 4 또는 5 개의 태그 카세트를 가질 수 있고/있거나 1, 2, 3, 4, 또는 5 개의 형질도입 마커가 또한 발현될 수 있다. 예를 들어, 구현예는 2 개의 Myc 태그 카세트, 또는 His 태그 및 HA 태그 카세트, 또는 HA 태그 및 Softag 1 태그 카세트, 또는 Myc 태그 및 SBP 태그 카세트가 있는 CAR 작제물을 포함할 수 있다. 특정 구현예에서, 발현 후 다량체화될 CAR은 상이한 태그 카세트를 포함한다. 특정 구현예에서, 형질도입 마커는 tEFGR을 포함한다. 예시적인 형질도입 마커 및 동족 쌍은 US 13/463,247에 기재되어 있다.

[0269] CAR에서 적어도 하나의 제어 특징을 포함하는 한 가지 이점은 대상체에게 투여된 CAR 발현 세포가 태그 카세트에 대한 동족 결합 분자를 사용하여 고갈될 수 있다는 점이다. 특정 구현예에서, 본 개시내용은 태그 카세트에

특이적인 항체를 사용하거나, 제어 특징에 특이적인 동족 결합 분자를 사용하거나, 또는 CAR을 발현하고 제어 특징에 대한 특이성을 갖는 제2 변형된 세포를 사용함으로써 CAR을 발현하는 변형된 세포를 고갈시키는 방법을 제공한다. 변형된 세포의 제거는 제어 특징에 특이적인 고갈제를 사용하여 달성될 수 있다.

[0270] 특정 구현예에서, 키메라 분자를 발현하는 변형된 세포는 제어 특징에 대한 특이성으로 결합하는 항체(예를 들어, 항-Tag 항체)를 사용하거나, 또는 제어 특징에 특이적으로 결합하는 다른 동족 결합에 의해 생체내에서 검출 또는 추적될 수 있으며, 제어 특징에 대한 결합 파트너는 형광 염료, 방사선-추적자, 산화철 나노입자 또는 X-선, CT-스캔, MRI-스캔, PET-스캔, 초음파, 유세포 분석, 근적외선 이미징 시스템, 또는 다른 이미징 양식에 의한 검출을 위해 당업계에 알려진 다른 이미징제에 접합된다(예를 들어, Yu, 등, *Theranostics* 2:3, 2012 참조).

[0271] 따라서, CAR이 있는 적어도 하나의 제어 특징을 나타내는 변형된 세포는 예를 들어, 태그 카세트가 없는 변형된 세포와 비교하여 더 용이하게 식별, 단리, 분류, 증식 유도, 추적, 및/또는 제거될 수 있다.

[0272] 본 개시내용의 방법 및 조성물에 유용한 예시적인 CAR 및 CAR 아키텍처는 WO2012/138475A1, US 9,624,306B2, US9266960B2, US2017/017477, EP2694549B1, US2017/0283504, US2017/0281766, US20170283500, US2018/0086846, US2010/0105136, US2010/0105136, WO2012/079000, WO2008045437, WO2016/139487A1, 및 WO2014/039523에 의해 제공된 것들을 포함한다.

[0273] TCR은 자연 발생 T 세포 수용체를 지칭한다. HSC는 선택된 TCR을 발현하도록 생체내에서 변형될 수 있다. CAR/TCR 하이브리드는 TCR의 요소 및 CAR의 요소를 갖는 단백질을 지칭한다. 예를 들어, CAR/TCR 하이브리드는 TCR 결합 도메인이 자연적으로 회합되지 않는 효과기 도메인과 함께 자연 발생 TCR 결합 도메인을 가질 수 있다. CAR/TCR 하이브리드는 돌연변이된 TCR 결합 도메인 및 ITAM 신호전달 도메인을 가질 수 있다. CAR/TCR 하이브리드는 삽입된 비-자연 발생 스페이서 영역 또는 막관통 도메인과 함께 자연 발생 TCR을 가질 수 있다.

[0274] 특정 CAR/TCR 하이브리드는 매사추세츠주 케임브리지 소재의 TCR2 Therapeutics의 TRuC®(T 세포 수용체 융합 작제물) 하이브리드를 포함한다. 예로서, TCR 융합 단백질의 생산은 국제 특허 공개 WO 2018/026953 및 WO 2018/067993, 및 출원 공개 US 2017/0166622에 기재되어 있다.

[0275] 특정 구현예에서, CAR/TCR 하이브리드는 "T-세포 수용체(TCR) 융합 단백질" 또는 "TFP"를 포함한다. TFP는 일반적으로 i) 표적 세포 상의 표면 항원에 결합할 수 있고 ii) 전형적으로 T-세포의 표면 내에 또는 표면 상에 공동-위치할 때, 온전한 TCR 복합체의 다른 폴리펩티드 구성요소와 상호작용할 수 있는 TCR을 포함하는 다양한 폴리펩티드로부터 유래된 재조합 폴리펩티드를 포함한다.

[0276] 특정 구현예에서, TFP는 암 항원(예를 들어, CD19, ROR1)에 결합하는 항체 단편을 포함하며 여기서 항체 단편의 서열은 TCR 서브유닛 또는 이의 부분을 암호화하는 핵산 서열과 인접하고 이와 동일한 판독 프레임 내에 있다. TFP는 기능적 TCR 복합체를 형성하기 위해 하나 이상의 내인성(또는 대안적으로, 하나 이상의 외인성, 또는 내인성 및 외인성의 조합) TCR 서브유닛과 회합할 수 있다.

[0277] **I(C)(i)(b). 유전자 편집 시스템 및 구성요소**

[0278] 다양한 구현예에서, 본 개시내용의 페이로드는 유전자 편집 시스템의 적어도 하나의 구성요소, 또는 모든 구성요소를 암호화한다. 본 개시내용의 유전자 편집 시스템은 CRISPR 시스템 및 염기 편집 시스템을 포함한다. 광범위하게, 유전자 편집 시스템은 CRISPR-연관 RNA-가이드된 엔도뉴클레아제 및 염기 편집 효소로부터 선택된 유전자 편집 효소 및 적어도 하나의 gRNA를 포함하는 복수의 구성요소를 포함할 수 있다. 따라서, 본 개시내용의 유전자 편집 시스템은 (i) CRISPR 시스템의 경우, CRISPR-연관 RNA-가이드된 엔도뉴클레아제인 CRISPR 효소 및 적어도 하나의 가이드 RNA(gRNA), 또는 (ii) 염기 편집 시스템의 경우, 염기 편집 효소 및 적어도 하나의 gRNA를 포함할 수 있다.

[0279] 본 개시내용은 자기-불활성화 유전자 편집 시스템이 본 개시내용의 벡터에 존재하고 벡터의 일부, 예를 들어, 통합 요소의 숙주 세포 게놈 내로의 절제 및/또는 통합 시 비-기능적으로 만드는 유전자 편집 시스템을 포함한다는 것을 포함한다. 다양한 구현예에서, 유전자 편집 시스템은 통합 요소의 절제 및/또는 통합 요소의 숙주 세포 게놈 내로의 통합 후 유전자 편집 시스템의 적어도 하나의 구성요소를 암호화하는 벡터 서열의 분해에 의해 비-기능적으로 만든다.

[0280] 본 개시내용은 다양한 구현예에서, CRISPR 효소 또는 염기 편집 효소가 PGK 프로모터와 작동가능하게 연결된 유전자 편집 시스템을 암호화하는 핵산 서열을 포함한다. 본 개시내용은 PGK가 공여자 벡터 생산을 위한 HEK293

세포와 같은 생산자 세포에서 더 약한 프로모터이지만(즉, 예를 들어, 생산자 세포의 Ef1α 프로모터와 비교하여 및/또는 HSC의 PGK 프로모터와 비교하여, 코딩 서열 발현의 상대적으로 낮거나 또는 감소된 수준을 구동) HSC에서 효율적인 이식유전자 발현을 구동하는(즉, 예를 들어, HSC의 Ef1α 프로모터와 비교하여 및/또는 HEK293 세포와 같은 생산자 세포의 PGK 프로모터와 비교하여, 코딩 서열 발현의 상대적으로 높거나 또는 증가된 수준을 구동) 실험적 발견을 포함한다.

[0281] 다양한 구현예에서, CRISPR 효소 또는 염기 편집 효소를 포함하는 유전자 편집 시스템을 암호화하는 핵산 서열은 예를 들어, TadA 및/또는 Tad*의 발현으로부터, 예를 들어, 생산자 세포(들)에서 유전자 편집 시스템 발현(예를 들어, 염기 편집 시스템 발현)의 잠재적인 유해 효과를 피하거나 또는 감소시키기 위해, HEK293 세포와 같은 생산자 세포에서 효소의 발현을 감소시키거나 또는 억제하는 microRNA 표적 부위를 포함한다. 다양한 구현예에서, miR 서열은 예를 들어, Saydaminova 등, *Mol. Ther. Meth. Clin. Dev.* 1: 14057, 2015; Li 등, *Mol. Ther. Meth. Clin. Dev.* 9: 390-401, 2018에 기재된 바와 같이, HDAd35 공여자 벡터 생산 동안 생산자 세포에서 염기 편집 또는 CRISPR 효소 발현을 억제하는 서열일 수 있으며, 상기 문헌은 본원에 참조로 포함된다.

[0282] 따라서, 의심의 여지를 피하기 위해, 본 개시내용은 유전자 편집 시스템을 암호화하는 핵산 서열이 (i) 임의적으로 핵산 서열이 본원에 개시된 바와 같은 변형된 TadA 및/또는 TadA*를 포함하는 경우, CRISPR 효소 또는 염기 편집 효소를 암호화하는 핵산 서열; (ii) CRISPR 효소 또는 염기 편집 효소 코딩 서열에 작동가능하게 연결된 PGK 프로모터; 및 (iii) HEK293 세포와 같은 생산자 세포에서 효소의 발현을 감소시키거나 또는 억제하는 microRNA 표적 부위의 전부 또는 임의의 것을 포함할 수 있는 구현예를 포함한다. 본 개시내용은 이러한 특징(i, ii, 및 iii)이 효과적인 유전자 요법에 개별적으로 및 상승적 조합으로 기여할 수 있음을 포함한다.

[0283] **I(C)(i)(b)(1). CRISPR 페이로드 발현 산물**

[0284] CRISPR(클러스터링된 규칙적인 간격을 갖는 짧은 회문 반복부)/Cas(CRISPR-연관 단백질) 뉴클레아제 시스템은 박테리아 시스템에 기반한 유전적 조작에 사용되는 조작된 뉴클레아제 시스템이다. 이는 부분적으로 많은 박테리아 및 고세균의 적응 면역 반응에 기반한다. 바이러스 또는 플라스미드가 박테리아를 침입하는 경우, 침입자의 DAN 분절은 박테리아의 "면역" 반응에 의해 CRISPR RNA(crRNA)로 전환된다. 그런 다음 crRNA는 부분적 상보성 영역을 통해, tracrRNA라 불리는 또 다른 유형의 RNA와 회합하여 "프로토스페이서"라 불리는 표적 DNA의 crRNA에 상동인 영역에 Cas 뉴클레아제를 가이드한다. Cas 뉴클레아제는 crRNA 전사체 내에 함유된 20-뉴클레오티드 상보적 가닥 서열에 의해 명시된 부위에서 이중-가닥 파괴 시 뭉툭한 단부를 생성하도록 DNA를 절단한다. 일부 경우에, Cas 뉴클레아제는 부위-특이적 DNA 인식 및 절단을 위해 crRNA 및 tracrRNA를 둘 다 필요로 한다.

[0285] 가이드 RNA(gRNA)는 표적화 요소의 일 예이다. 가장 간단한 형태에서, gRNA는 상보성에 기반한 게놈 내의 부위를 표적화하는 서열(예를 들어, crRNA)을 제공한다. 그러나, 하기 예시된 바와 같이, gRNA는 또한 추가적인 구성요소를 포함할 수 있다. 예를 들어, 특정 구현예에서, gRNA는 표적화 서열(예를 들어, crRNA) 및 표적화 서열을 절단 요소에 연결하는 구성요소를 포함할 수 있다. 이 연결 구성요소는 tracrRNA일 수 있다. 특정 구현예에서, 하기 기재된 바와 같이, crRNA 및 tracrRNA를 포함하는 gRNA는 단일 gRNA(sgRNA)로 지칭되는 단일 분자로 표현될 수 있다. gRNA는 또한 다른 메커니즘을 통해 예컨대 나노입자를 통해 또는 이중 또는 다중-목적 분자의 발현 또는 구축을 통해 절단 요소에 연결될 수 있다. 당업자는 예를 들어, 본 개시내용의 아데노바이러스 공여자 벡터 또는 게놈의 숙주 세포에서 선택된 핵산 서열 교정 또는 변형을 생성하기 위한 gRNA 또는 다른 표적화 요소가 예를 들어, 이용가능한 서열 정보를 기반으로 용이하게 설계 및 구현될 수 있음을 이해할 것이다.

[0286] 특정 구현예에서, 표적화 요소(예를 들어, gRNA)는 새롭거나 또는 향상된 특징(예를 들어, 개선된 안정성)이 있는 핵산을 제공하기 위해, 하나 이상의 변형(예를 들어, 염기 변형, 백본 변형)을 포함할 수 있다. 변형된 백본은 백본에서 인 원자를 유지하는 것들 및 백본에서 인 원자를 갖지 않는 것들을 포함할 수 있다. 인 원자를 함유하는 적합한 변형된 백본은 예를 들어, 포스포로티오에이트, 키랄 포스포로티오에이트, 포스포로디티오에이트, 포스포트리에스테르, 아미노알킬포스포트리에스테르, 메틸 및 다른 알킬 포스페이트 예컨대 3'-알킬렌 포스포네이트, 5'-알킬렌 포스포네이트, 키랄 포스포네이트, 포스피네이트, 3'-아미노 포스포라미데이트 및 아미노알킬포스포라미데이트를 포함하는 포스포라미데이트, 포스포로디아미데이트, 티오노포스포라미데이트, 티오노알킬포스포네이트, 티오노알킬포스포트리에스테르, 셀레노포스포네이트, 및 정상 3'-5' 연결, 2'-5' 연결된 유사체, 및 도립된 극성이 있는 것을 갖는 보라노포스페이트를 포함할 수 있으며 여기서 하나 이상의 뉴클레오티드간 연결은 3'에서 3'으로, 5'에서 5'으로 또는 2'에서 2'으로 연결이다. 도립된 극성을 갖는 적합한 표적화 요소는 3'-대부분 뉴클레오티드간 연결에서 단일 3'에서 3'으로 연결(즉, 핵염기가 누락되거나 또는 이 대신에 하

이드록실 기를 갖는 단일 도립된 뉴클레오티드 잔기)를 포함할 수 있다. 다양한 염(예를 들어, 염화칼륨 또는 염화나트륨), 혼합 염, 및 유리 산 형태가 또한 포함될 수 있다.

[0287] 표적화 요소는 하나 이상의 포스포로티오에이트 및/또는 헤테로원자 뉴클레오시드간 연결, 특히 --$CH_2$-NH-O-$CH_2$-, --$CH_2$-N($CH_3$)-O-$CH_2$-(즉, 메틸렌(메틸이미노) 또는 MMI 백본), --$CH_2$-O-N($CH_3$)-$CH_2$-, --$CH_2$-N($CH_3$)-N($CH_3$)-$CH_2$- 및 --O--N($CH_3$)-$CH_2$-$CH_2$-(여기서 천연 포스포디에스테르 뉴클레오티드간 연결은 --O--P(=O)(OH)-O-$CH_2$-로 표현됨)를 포함할 수 있다.

[0288] 특정 구현예에서, 표적화 요소는 모르폴리노 백본 구조를 포함할 수 있다. 예를 들어, 표적화 요소는 리보스 고리 대신에 6-원 모르폴리노 고리를 포함할 수 있다. 이러한 구현예 중 일부에서, 포스포로디아미데이트 또는 다른 비-포스포디에스테르 뉴클레오시드간 연결은 포스포디에스테르 연결을 대체한다.

[0289] 특정 구현예에서, 표적화 요소는 하나 이상의 치환된 당 모이어티를 포함할 수 있다. 적합한 폴리뉴클레오티드는 OH; F; O-, S-, 또는 N-알킬; O-, S-, 또는 N-알케닐; O-, S- 또는 N-알키닐; 또는 O-알킬-O-알킬로부터 선택된 당 치환기를 포함할 수 있으며, 여기서 알킬, 알케닐 및 알키닐은 치환 또는 비치환된 C1 내지 C10 알킬 또는 C2 내지 C10 알케닐 및 알키닐일 수 있다. O(($CH_2$)nO) m$CH_3$, O($CH_2$)nO$CH_3$, O($CH_2$)n$NH_2$, O($CH_2$)n$CH_3$, O($CH_2$)nONH$_2$, 및 O($CH_2$)nON(($CH_2$)n$CH_3$)$_2$가 특히 적합하며, 여기서 n 및 m은 1 내지 10이다.

[0290] 절단 요소의 예는 뉴클레아제를 포함한다. CRISPR-Cas 유전자좌는 50 개 초과의 유전자 패밀리를 갖고 엄격하게 보편적인 유전자가 없으며, 이는 유전자좌 아키텍처의 빠른 진화 및 극도의 다양성을 나타낸다. 예시적인 Cas 뉴클레아제는 Cas1, Cas1B, Cas2, Cas3, Cas4, Cas5, Cas6, Cas7, Cas8, Cas9(Csn1 및 Csx12로도 알려짐), Cas10, Cpf1, C2c3, C2c2 및 C2c1Csy1, Csy2, Csy3, Cse1, Cse2, Csc1, Csc2, Csa5, Csn2, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Cpf1, Csb1, Csb2, Csb3, Csx17, Csx14, Csx10, Csx16, CsaX, Csx3, Csx1, Csx15, Csf1, Csf2, Csf3, 및 Csf4를 포함한다.

[0291] 3가지 주요 유형의 Cas 뉴클레아제(I형, II형, 및 III형), 및 5 개의 I형, 3 개의 II형, 및 2 개의 III형 단백질을 포함하는 10 개의 하위유형이 있다(예를 들어, Hochstrasser and Doudna, Trends Biochem Sci, 2015:40(1):58-66) 참조. II형 Cas 뉴클레아제는 Cas1, Cas2, Csn2, 및 Cas9를 포함한다. 이들 Cas 뉴클레아제는 당업자에게 알려져 있다. 예를 들어, 스트렙토코쿠스 피오게네스(Streptococcus pyogenes) 야생형 Cas9 폴리펩티드의 아미노산 서열은 예를 들어, NCBI 참조 서열 번호 NP 269215에 제시되어 있고, 스트렙토코쿠스 써모필루스(Streptococcus thermophilus) 야생형 Cas9 폴리펩티드의 아미노산 서열은 예를 들어, NCBI 참조 서열 번호 WP_011681470에 제시되어 있다.

[0292] 특정 구현예에서, Cas9는 RNA-가이드된 이중-가닥 DNA-결합 뉴클레아제 단백질 또는 닉카제 단백질을 지칭한다. 야생형 Cas9 뉴클레아제는 상이한 DNA 가닥을 절단하는 2 개의 기능적 도메인, 예를 들어, RuvC 및 HNH를 갖는다. Cas9는 두 기능적 도메인이 활성일 때 게놈 DNA(표적 DNA)에서 이중 가닥 파괴를 유도할 수 있다. 일부 구현예에서, Cas9 효소는 코리네박터(Corynebacter), 수테렐라(Sutterella), 레지오넬라(Legionella), 트레포네마(Treponema), 필리프 액터(Filif actor), 유박테리움(Eubacterium), 스트렙토코쿠스(Streptococcus), 락토바실루스(Lactobacillus), 마이코플라스마(Mycoplasma), 박테로이데스(Bacteroides), 플라비볼라(Flaviivola), 플라보박테리움(Flavobacterium), 스페로케타(Sphaerochaeta), 아조스피릴룸(Azospirillum), 글루코나세토박터(Gluconacetobacter), 네이세리아(Neisseria), 로세부리아(Roseburia), 파르비바큘룸(Parvibaculum), 스타필로코쿠스(Staphylococcus), 니트라티프락터(Nitratifractor), 및 캄필로박터(Campylobacter)와 같은 박테리아로부터 유래된 Cas9 단백질의 하나 이상의 촉매적 도메인을 포함한다. 일부 구현예에서, Cas9는 융합 단백질이며, 예를 들어 2 개의 촉매적 도메인은 상이한 박테리아 종으로부터 유래된다.

[0293] 이전에 나타낸 바와 같이, CRISPR/Cas 시스템은 특정 경우에, crRNA 및 tracrRNA가 단일 gRNA(sgRNA)라고 불리는 하나의 분자 내로 조합될 수 있도록 조작되었다. 이 조작된 접근법에서, sgRNA는 Cas가 임의의 원하는 서열을 표적화하도록 가이드한다(예를 들어, Jinek 등, *Science* 337:816-821, 2012; Jinek 등, *eLife* 2:e00471, 2013; Segal, *eLife* 2:e00563, 2013 참조). 따라서, CRISPR/Cas 시스템은 세포의 게놈에서 원하는 표적에서 이중 가닥 파괴를 생성하고, 세포의 내인성 메커니즘을 활용하여 HDR, 또는 NHEJ에 의해 유도된 파괴를 복구하도록 조작될 수 있다. 본원에 기재된 특정 구현예는 정의된 통합 부위에서 HDR을 촉진하기 위해 상동성 아암을 활용한다.

[0294] Cas9 뉴클레아제의 유용한 변이체는 RuvC" 또는 HNH" 효소 또는 닉카제와 같은 단일 불활성 촉매적 도메인을 포

함한다. Cas9 닉카제는 활성 하나의 기능적 도메인만을 가지며, 일부 구현예에서, 표적 DNA의 한쪽 가닥만을 절단하여, 단일 가닥 파괴 또는 닉을 생성한다. 일부 구현예에서, 적어도 D10A 돌연변이를 갖는 돌연변이체 Cas9 뉴클레아제는 Cas9 닉카제이다. 다른 구현예에서, 적어도 H840A 돌연변이를 갖는 돌연변이체 Cas9 뉴클레아제는 Cas9 닉카제이다. Cas9 닉카제에 존재하는 돌연변이의 다른 예는 N854A 및 N863 A를 포함한다. 이중 가닥 파괴는 반대 DNA 가닥을 표적화하는 적어도 2 개의 DNA-표적화 RNA가 사용되는 경우 Cas9 닉카제를 사용하여 도입된다. 이중 닉 유도된 이중 가닥 파괴는 HDR 또는 NHEJ에 의해 복구된다. 이 유전자 편집 전략은 일반적으로 HDR을 선호하고 표적외 DNA 부위에서 indel 돌연변이의 빈도를 감소시킨다. 일부 구현예에서, Cas9 뉴클레아제 또는 닉카제는 표적 세포 또는 표적 유기체에 대해 코돈-최적화된다.

[0295] 특정 구현예는 스타필로코쿠스 아우레우스(*Staphylococcus aureus*) Cas9(SaCas9)를 활용할 수 있다. 특정 구현예는 다음 위치 중 하나 이상에서 돌연변이가 있는 SaCas9를 활용할 수 있다: E782, N968, 및/또는 R1015. 특정 구현예는 다음 위치 중 하나 이상에서 돌연변이가 있는 SaCas9를 활용할 수 있다: E735, E782, K929, N968, A1021, K1044 및/또는 R1015. 일부 구현예에서, 변이체 SaCas9 단백질은 다음 돌연변이 중 하나 이상을 포함한다: R1015Q, R1015H, E782K, N968K, E735K, K929R, A1021T, 및/또는 K1044N. 일부 구현예에서, 변이체 SaCas9 단백질은 D10A, D556A, H557A, N580A, 예를 들어, D10A/H557A 및/또는 D10A/D556A/H557A/N580A에서 돌연변이를 포함한다. 일부 구현예에서, 변이체 SaCas9 단백질은 E735, E782, K929, N968, R1015, A1021, 및/또는 K1044로부터 선택된 하나 이상의 돌연변이를 포함한다. 일부 구현예에서, SaCas9 변이체는 다음 돌연변이 세트 중 하나를 포함할 수 있다: E782K/N968K/R1015H(KKH 변이체); E782K/K929R/R1015H(KRH 변이체); 또는 E782K/K929R/N968K/R1015H(KRKH 변이체).

[0296] Cpf1에 의해 예시된 클래스 II, V형 CRISPR-Cas 클래스는 Zetsche 등 (2015) Cell 163(3): 759-771에서 식별되었다. Cpf1 뉴클레아제는 특히 프로토스페이서-인접 모티프 또는 PAM으로 알려진 짧은 3 개의 염기 쌍 인식 서열(TTN)에 의해 표적 부위 선택에 추가의 유연성을 제공할 수 있다. Cpf1의 절단 부위는 PAM 서열로부터 적어도 18 bp 떨어져 있다. 더욱이, 끈적한 단부가 있는 엇갈린 DSB는 방향-특이적 공여자 주형 삽입을 허용하며, 이는 비-분할 세포에 유리하다.

[0297] 특정 구현예는 조작된 Cpf1을 활용할 수 있다. 예를 들어, US 2018/0030425는 표적 특이성이 변경되고 개선된 라크노스피라세 박테리움(Lachnospiraceae bacterium) ND2006 및 악시다미노코쿠스 종(Acidaminococcus sp.) BV3L6으로부터 조작된 Cpf1 뉴클레아제를 기재한다. 특정 변이체는 예를 들어, 다음 위치 중 하나 이상에서 돌연변이(즉, 천연 아미노산을 상이한 아미노산, 예를 들어, 알라닌, 글리신, 또는 세린으로 대체)가 있는 아미노산 19-1246을 적어도 포함하는 라크노스피라세 박테리움 ND2006을 포함한다: S202, N274, N278, K290, K367, K532, K609, K915, Q962, K963, K966, K1002, 및/또는 S1003. 특정 Cpf1 변이체는 또한 다음 위치 중 하나 이상에서 돌연변이(즉, 천연 아미노산을 상이한 아미노산, 예를 들어, 알라닌, 글리신, 또는 세린(천연 아미노산이 세린인 경우 제외)으로 대체)가 있는 악시다미노코쿠스 종 BV3L6 Cpf1(AsCpf1)을 포함한다: N178, S186, N278, N282, R301, T315, S376, N515, K523, K524, K603, K965, Q1013, Q1014, 및/또는 K1054.

[0298] 다른 Cpf1 변이체는 Zetsche 등 (2015) Cell 163: 759-771에 개시된 Cpf1 폴리펩티드 뿐만 아니라 U.S. 2016/0208243에 개시된 Cpf1 폴리펩티드의 Cpf1 상동체 및 오솔로그(ortholog)를 포함한다. 다른 조작된 Cpf1 변이체는 당업자에게 알려져 있고 본 개시내용의 범위 내에 포함된다(예를 들어, WO/2017/184768 참조).

[0299] CRISPR-Cas 시스템 및 이의 구성요소에 관한 추가적인 정보는 다음에 기재되어 있다: US 8697359, US 8771945, US 8795965, US 8865406, US 8871445, US 8889356, US 8889418, US 8895308, US 8906616, US 8932814, US 8945839, US 8993233 및 US 8999641 및 이와 관련된 출원; 및 WO2014/018423, WO2014/093595, WO2014/093622, WO2014/093635, WO2014/093655, WO2014/093661, WO2014/093694, WO2014/093701, WO2014/093709, WO2014/093712, WO2014/093718, WO2014/145599, WO2014/204723, WO2014/204724, WO2014/204725, WO2014/204726, WO2014/204727, WO2014/204728, WO2014/204729, WO2015/065964, WO2015/089351, WO2015/089354, WO2015/089364, WO2015/089419, WO2015/089427, WO2015/089462, WO2015/089465, WO2015/089473 및 WO2015/089486, WO2016/205711, WO2017/106657, WO2017/127807 및 이와 관련된 출원.

[0300] 일부 구현예에서 CRISPR 시스템은 예를 들어, γ-글로빈의 발현을 증가시키기 위해 γ-글로빈을 암호화하는 핵산 서열을 변형시키도록 조작된다. 헤모글로빈의 주요 태아 형태인 헤모글로빈 F(HbF)는 γ-글로빈 폴리펩티드 서브유닛과 α-글로빈 폴리펩티드 서브유닛의 쌍형성에 의해 형성된다. 인간 태아 γ -글로빈 유전자(HBG1 및 HBG2; 진화적 복제에 의해 생성된 2 개의 고도로 상동성인 유전자)는 정상적으로 출생 시 침묵하는 반면, 성인 β-글로빈 유전자 발현(HBB 및 HBD)의 발현은 증가한다. 일생에 걸쳐 태아 γ-글로빈의 지속적인 발현을 야기

하거나 또는 허용하는 돌연변이는 β-글로빈 결핍의 표현형을 개선할 수 있다. 따라서, 태아 γ-글로빈 유전자의 재활성화는 특히 β-글로빈 결핍이 있는 대상체에서 치료적으로 유익할 수 있다. γ-글로빈의 발현을 증가시키는 다양한 돌연변이는 당업계에 알려져 있고 본원에 개시되어 있다(예를 들어, Wienert, *Trends in Genetics* 34(12): 927-940, 2018을 참조하며, γ-글로빈의 발현을 증가시키는 돌연변이와 관련하여 그 전문이 본원에 참조로 포함된다). 이러한 특정 돌연변이는 HBG1 프로모터 또는 HBG2 프로모터에서 발견된다.

[0301] 일부 구현예에서, 벡터 또는 게놈은 페이로드가 통합 요소를 포함하는 CRISPR 시스템을 포함하고 CRISPR 시스템의 적어도 하나의 구성요소는 페이로드에 존재하지만 통합 요소의 외부(예를 들어, 트랜스포존 도립 반복부에 의해 플랭킹된 전위성 통합 요소를 포함하는 페이로드 단편의 외부 또는 상동 통합을 위한 상동성 아암을 포함하는 페이로드 단편의 외부)에 있다. 페이로드가 전위성 통합 요소를 포함하는 특정한 특정 구현예에서, 전위성 통합 요소가 트랜스포존 도립 반복부에 의해 플랭킹된 경우, CRISPR 효소 및/또는 CRISPR 시스템의 하나 이상의 gRNA 중 하나 이상은 전위성 통합 요소의 외부 위치에서 (즉, 내부에 존재하지 않는) 페이로드에 존재한다(즉, 트랜스포존 도립 반복부에 의해 플랭킹된 핵산 서열에 존재하지 않음). 페이로드가 전위성 통합 요소를 포함하는 특정한 특정 구현예에서, 전위성 통합 요소가 상동성 아암에 의해 플랭킹된 경우, CRISPR 효소 및/또는 CRISPR 편집 시스템의 하나 이상의 gRNA 중 하나 이상은 통합 요소의 위부 위치에서 (즉, 내부에 존재하지 않는) 페이로드에 존재한다(즉, 상동성 아암에 의해 플랭킹된 핵산 서열에 존재하지 않음). 이러한 시스템에서, CRISPR 시스템의 발현 및/또는 활성은 일시적이며, 전위성 통합 요소의 전위의 경우 벡터를 파괴하고 전위성 통합 요소의 외부에 위치한 CRISPR 시스템 구성요소 중 하나 이상의 발현을 감소 또는 종결시킬 수 있다. (예를 들어, 전위 또는 상동 재조합에 의한) 통합 요소의 통합은 CRISPR 시스템의 발현 및/또는 활성을 불활성화할 수 있기 때문에 CRISPR 시스템을 포함하는 이러한 벡터는 때때로 "자기-불활성화" CRISPR 시스템 또는 벡터로 지칭될 수 있다. 다양한 구현예에서, 자기-불활성화 CRISPR 시스템은 조합 페이로드에 존재한다.

[0302] 본 발명자들은 자기-불활성화 CRISPR 시스템 페이로드를 포함하는 아데노바이러스 벡터(예를 들어, HDAd 아데노바이러스 벡터)가 예를 들어, 다른 CRISPR 시스템 페이로드와 비교하여 유전자 요법(예를 들어, 생체내 유전자 요법)에서 절단 빈도 증가 및/또는 형질도입 및/또는 편집된 표적 세포의 생존 증가(예를 들어, 형질도입 HSPC의 생존 증가)를 초래함을 관찰하였으며, 여기서 CRISPR 시스템은 통합 요소 내에 완전히 있거나 또는 CRISPR 시스템은 숙주 세포 게놈 내로 통합되지 않지만 발현은 벡터 파괴에 의해 불활성화되지 않는다. CRISPR 시스템의 자기-불활성화는 CRISPR 효소 및/또는 gRNA의 발현을 단축시키고, 편집된 세포의 생존을 증가시키고, 장기간 재증식하는 세포의 백분율을 증가시킨다, 일 예를 제공하기 위해, HBG1 및/또는 HGB2의 재활성화를 위한 자기-불활성화 CRISPR 시스템을 포함하고 γ-글로빈의 발현을 위한 핵산 서열을 추가로 포함하는 조합 페이로드를 포함하는 HDAd 벡터를 사용한 유전자 요법은 비-불활성화 CRISPR 시스템 또는 γ-글로빈 단독의 발현을 위한 핵산 서열을 포함하는 HDAd 벡터가 형질도입 수행 후 RBC에서 상당히 더 높은 γ-글로빈을 생성하였다.

[0303] 본원에는 예를 들어, 자기-불활성화 CRISPR 시스템을 포함하는 공여자 벡터가 통합 요소의 전위를 위한 트랜스포사제를 암호화하는 지지 벡터 또는 게놈과 조합하여 인간 대상체에게 투여되는 방법을 추가로 제공한다. 본 개시내용은 다양한 경우에 공여자 벡터가 지지 벡터의 투여 전에 투여되는 것을 포함하며, 여기서 공여자 벡터의 투여 및 지지 벡터의 투여 사이의 시간 기간은 CRISPR 시스템의 지속기간 및/또는 활성 수준을 조절하는 수단을 제공한다. 예를 들면, 다양한 구현예에서, 지지 벡터는 예를 들어, 공여자 벡터의 투여 후 일정 기간에 대상체에게 투여될 수 있으며 여기서 일정 기간은 적어도 1, 2, 3, 4, 5, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 30, 36, 42, 48, 54, 60, 66, 또는 72, 96, 또는 128 시간이다(예를 들어, 여기서 기간은 1, 2, 3, 4, 5, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 30, 36, 42, 48, 54, 60, 66, 또는 72 시간의 하한 및 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 30, 36, 42, 48, 54, 60, 66, 72, 96, 또는 128 시간의 상한을 갖는다).

[0304] 일부 구현예에서, CRISPR 시스템 구성요소를 암호화하는(예를 들어, CRISPR 효소를 암호화하는) 핵산 서열은 CRISPR 발현 및/또는 활성의 microRNA 조절을 위한 microRNA 표적 부위를 포함하도록 조작된다.

[0305] **I(C)(i)(b)(2). 염기 편집기 페이로드 발현 산물**

[0306] 본 개시내용은 무엇보다도, 염기 편집제 및 이를 암호화하는 핵산을 포함하며, 임의적으로 여기서 염기 편집제 또는 이를 암호화하는 핵산은 아데노바이러스 벡터 또는 게놈과 같은 벡터 또는 게놈에 존재한다. 염기 편집 시스템은 염기 편집 효소 및/또는 적어도 하나의 gRNA를 이의 구성요소로서 포함할 수 있다. 특정한 특정 구현예에서, 본 개시내용의 염기 편집제 및/또는 염기 편집 시스템은 Ad35 또는 Ad5/35 아데노바이러스 벡터에 존재한다. 그러나, 당업자는 본 개시내용의 염기 편집제 및 이를 암호화하는 핵산 서열이 임의의 맥락 또는 형태, 예를 들어, 아데노바이러스 벡터가 아닌 벡터, 예를 들어, 플라스미드로 존재할 수 있음을 이해할 것이다. 본

원에 개시된 바와 같은 뉴클레오티드 서열 암호화 염기 편집 시스템은 전형적으로 많은 제한된 용량의 벡터 시스템에 포함하기에 너무 크지만, 아데노바이러스 벡터의 큰 용량은 본 개시내용의 아데노바이러스 벡터 및 게놈에 이러한 서열의 포함을 허용한다. 실제로, 본원의 다른 곳에서 논의된 바와 같이, 아데노바이러스 벡터는 염기 편집 시스템을 암호화하고 하나 이상의 추가적인 코딩 서열을 추가로 암호화하는 페이로드를 포함할 수 있다. 본 개시내용의 염기 편집기를 암호화하는 페이로드를 사용한 유전자 요법을 위한 본원에 개시된 바와 같은 아데노바이러스 벡터 및 게놈의 추가적인 이점은 Ad35 게놈과 같은 아데노바이러스 게놈이 숙주 세포 게놈 내로 자연적으로 통합되지 않아, 염기 편집 시스템의 일시적 발현을 용이하게 한다는 점이며, 예를 들어, 면역원성 및/또는 유전독성을 피하는 것이 바람직할 수 있다.

[0307] 염기 편집은 게놈 DNA 또는 세포 RNA 내의 염기 또는 염기 쌍을 상이한 염기 또는 염기 쌍으로 전환함으로써 핵산 서열의 선택적 변형을 지칭한다(Rees & Liu, *Nature Reviews Genetics*, 19:770-788, 2018). DNA 염기 편집기에는 다음과 같은 2 가지 일반적인 클래스가 있다: (i) 구아닌-사이토신 염기 쌍을 티민-아데닌 염기 쌍으로 전환하는 사이토신 염기 편집기(CBE), 및 (ii) 아데닌-티민 염기 쌍을 구아닌 사이토신 염기 쌍으로 전환하는 아데닌 염기 편집기(ABE). 특정 구현예에서, CRISPR 시스템의 구성요소는 다른 효소 또는 이의 생물학적으로 활성인 단편과 조합하여 예를 들어, 돌연변이된 핵산에서 하나 이상의 이중-가닥 파괴를 만들거나, 야기하기나, 또는 생성하지 않고 핵산, 예를 들어, DNA 또는 RNA에 점 돌연변이와 같은 돌연변이를 직접적으로 설치하거나, 야기하거나, 또는 생성한다. 구성요소의 이러한 특정 조합은 염기 편집기로 알려져 있다.

[0308] DNA 염기 편집기는 핵염기 데아미나제 효소에 융합된 촉매적으로 기능상실된 뉴클레아제 및, 일부 경우에 DNA 글리코실라제 억제제를 포함할 수 있다. RNA 염기 편집기는 RNA를 변형시키는 염기를 구성요소로 사용하여 유사한 변화를 달성한다.

[0309] DNA에서 표적 유전자좌에 결합 시, 가이드 RNA 및 표적 DNA 가닥 사이에 염기 쌍형성은 단일-가닥 DNA의 작은 분절의 대체로 이어진다. 이 단일-가닥 DNA 버블 내의 DNA 염기는 데아미나제 효소에 의해 변형될 수 있다. 특정 구현예에서, 진핵생물 세포에서의 효율을 개선하기 위해, 촉매적으로 기능상실된 뉴클레아제는 또한 비-편집된 DNA 가닥에서 닉을 생성하여, 세포가 편집된 가닥을 주형으로 사용하여 비-편집된 가닥을 복구하도록 유도한다.

[0310] CBE의 경우, CRISPR-기반 편집기는 사이토신 데아미나제를 Cas 닉카제, 예를 들어, Cas9 닉카제(nCas9)와 연결함으로써 생성될 수 있다. 일 예를 피하기 위해, nCas9는 단일 가닥을 절단함으로써 표적 DNA에 닉을 생성하여, 이중-가닥 파괴를 필요로 하는 방법과 비교하여 해로운 indel 형성 가능성을 감소시킬 수 있다. DNA와 결합 후, CBE는 표적 사이토신(C)을 우라실(U) 염기로 탈아미노화한다. 이후에 생성된 U-G 쌍은 원래 C-G 쌍을 T-A로 전환하거나 또는 우라실 글리코실라제에 의해 매개된 염기 절제 복구에 의해 원래 C-G로 역전시키는 세포 불일치 복구 기구에 의해 복구된다. 다양한 구현예에서, 우라실 글리코실라제 억제제(UGI), 예를 들어, 페이로드에 존재하는 UGI의 발현은 2차 결과의 발생을 감소시키고 T-A 염기 쌍 형성의 생성을 증가시킨다.

[0311] 아데노신 염기 편집기(ABE)의 경우, 아데닌 염기 편집을 위해 DNA 상에서 작용할 수 있는 예시적인 아데노신 데아미나제는 DNA를 기질로서 수용하는 돌연변이체 TadA 아데노신 데아미나제(TadA*)를 포함한다. 이. 콜라이(*E. coli*) TadA는 전형적으로 동종이량체로서 작용하여 전달 RNA(tRNA)에서 아데노신을 탈아미노화한다. TadA* 데아미나제는 표적 'A'에서 'I'(이노신)으로의 전환을 촉매하며, 이는 세포 폴리머라제에 의해 'G'로 처리된다. 후속적으로, 원래 게놈 A-T 염기 쌍은 G-C 쌍으로 전환될 수 있다. 세포의 이노신 절제 복구는 우라실 절제만큼 활성이 아니므로, ABE는 CBE에서 UGI와 같은 임의의 추가적인 억제제 단백질을 필요로 하지 않는다. 일부 구현예에서, 전형적인 ABE는 염기 편집 동안 구조적 역할을 할 수 있는 야생형 이. 콜라이 tRNA-특이적 아데노신 데아미나제(TadA) 단량체, 데옥시아데노신 탈아미노화를 촉매하는 TadA* 돌연변이체 TadA 단량체, 및 Cas9(D10A)와 같은 Cas 닉카제를 포함하는 3 개의 구성요소를 포함할 수 있다. 특정 구현예에서, TadA 및 TadA* 사이에 위치한 링커가 있고, 특정 구현예에서 TadA* 및 Cas 닉카제 사이에 위치한 링커가 있다. 다양한 구현예에서, 하나 또는 두 링커는 적어도 6 개의 아미노산, 예를 들어, 적어도 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 또는 50 개의 아미노산(예를 들어, 5, 6, 7, 8, 9, 10, 또는 15 개, 아미노산의 하한 및 20, 25, 30, 35, 40, 45, 또는 50 개 아미노산의 상한을 가짐)을 포함한다. 다양한 구현예에서, 하나 또는 두 링커는 32 개의 아미노산을 포함한다. 일부 구현예에서, 하나 또는 두 링커는 $(SGGS)_2$-XTEN-$(SGGS)_2$에 따른 서열, 또는 당업자에게 달리 알려진 서열을 갖는다.

[0312] 염기 편집기는 하나의 염기 또는 염기 쌍을 또 다른 것으로 직접적으로 전환할 수 있으며, 삽입 및 결실(indel)과 같은 과도한 원치않은 편집 부산물을 생성하지 않고 비-분할 세포에서 점 돌연변이의 효율적인 설치를 가

능하게 한다. 예를 들어, 염기 편집기는 10%, 9%, 8%, 7%, 6%, 5.5%, 5%, 4.5%, 4%, 3.5%, 3%, 2.5%, 2%, 1.5%, 1%, 0.5%, 또는 0.1% 미만의 indel을 생성할 수 있다.

[0313] DNA 염기 편집기는 이러한 점 돌연변이를 이중 가닥 파괴를 생성하지 않고 비-분할 세포에 삽입할 수 있다. 이중 가닥 파괴의 결여로 인해, 염기 편집기는 삽입 및 결실(indel)과 같은 과도한 원치않은 편집 부산물을 초래하지 않는다. 예를 들어, 염기 편집기는 이중 가닥 파괴에 의존하는 기술과 비교하여 10%, 9%, 8%, 7%, 6%, 5.5%, 5%, 4.5%, 4%, 3.5%, 3%, 2.5%, 2%, 1.5%, 1%, 0.5%, 또는 0.1% 미만의 indel을 생성할 수 있다.

[0314] 대부분의 염기-편집 시스템의 구성요소는 (1) 표적화된 DNA 결합 단백질, (2) 핵염기 데아미나제 효소, 및 (3) DNA 글리코실라제 억제제를 포함한다.

[0315] CRISPR 시스템의 임의의 뉴클레아제는 기능상실되고 염기 편집 시스템 내에서 사용될 수 있다. 예시적인 Cas 뉴클레아제는 Cas1, Cas1B, Cas2, Cas3, Cas4, Cas5, Cas6, Cas7, Cas8, Cas9 (Csn1 및 Csx12로도 알려짐), Cas10, Cpf1, C2c3, C2c2 및 C2c1Csy1, Csy2, Csy3, Cse1, Cse2, Csc1, Csc2, Csa5, Csn2, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Cpf1, Csb1, Csb2, Csb3, Csx17, Csx14, Csx10, Csx16, CsaX, Csx3, Csx1, Csx15, Csf1, Csf2, Csf3, Csf4 및 이의 돌연변이를 포함한다.

[0316] 특정 구현예는 뉴클레아제-불활성 Cas9(dCas9)를 촉매적으로 기능상실된 뉴클레아제로서 활용한다. 그러나, CRISPR 시스템의 임의의 뉴클레아제(이 중 많은 것이 상기 기재됨)는 기능상실되고 염기 편집 시스템 내에서 사용될 수 있다. 특정 구현예에서, 고충실도의 Cas9 도메인이 선택되며 여기서 Cas9 도메인은 야생형 Cas9 도메인과 비교하여, Cas9 도메인 및 DNA의 당-포스페이트 백본 사이에 감소된 정전기적 상호작용을 나타낸다. 일부 구현예에서, Cas9 도메인(예를 들어, 야생형 Cas9 도메인)은 Cas9 도메인 및 DNA의 당-포스페이트 백본 사이의 회합을 감소시키는 하나 이상의 돌연변이를 포함한다. 고충실도의 Cas9 도메인은 당업자에게 알려져 있다. 예를 들어, 고충실도의 Cas9 도메인은 Kleinstiver, 등, *Nature* 529, 490-495, 2016; 및 Slaymaker 등, *Science* 351, 84-88, 2015에 기재되었다.

[0317] 다른 유전자-편집 시스템의 뉴클레아제가 또한 사용될 수 있다. 예를 들어, 염기-편집 시스템은 아연 핑거 뉴클레아제(ZFN)(Urnov 등, *Nat Rev Genet.*, 11(9):636-46, 2010) 및 전사 활성인자 유사 효과기 뉴클레아제(TALEN)(Joung 등, *Nat Rev Mol Cell Biol.* 14(1):49-55, 2013)를 활용할 수 있다. DNA-결합 뉴클레아제에 관한 추가적인 정보를 위해, US2018/0312825A1을 참조한다.

[0318] 특정 구현예에서, 핵염기 데아미나제 효소는 사이티딘 데아미나제 도메인 또는 아데닌 데아미나제 도메인을 포함한다.

[0319] 특정 구현예는 사이티딘 데아미나제 도메인을 핵염기 데아미나제 효소로서 활용한다. 특정 구현예는 아데닌 데아미나제 도메인을 핵염기 데아미나제 효소로서 활용한다. 추가로, 특정 구현예는 우라실 글리코실라제 억제제(UGI)를 글리코실라제 억제제로서 활용한다. 예를 들어, 특정 구현예에서, dCas9 또는 Cas9 닉카제는 사이티딘 데아미나제 도메인에 융합될 수 있다. 사이티딘 데아미나제 도메인에 융합된 dCas9 또는 Cas9 닉카제는 하나 이상의 UGI 도메인에 융합될 수 있다. 하나 초과의 UGI 도메인이 있는 염기 편집기는 indel을 덜 생성하고 표적 핵산을 더 효율적으로 탈아미노화할 수 있다.

[0320] 특정 구현예에서, 데아미나제 도메인(사이티딘 및/또는 아데닌)은 촉매적으로 기능상실된 뉴클레아제의 N-말단에 융합된다. 이는 Cas9의 N-말단에 융합된 사이티딘 데아미나제 도메인이 다른 구성과 비교하여 염기-편집 효율이 개선될 수 있기 때문이다. 이러한 구현예에서, 글리코실라제 억제제(예를 들어, UGI 도메인)는 촉매적으로 기능상실된 뉴클레아제의 C-말단에 융합될 수 있다. 다중 글리코실라제 억제제가 사용될 때, 각각은 촉매적으로 기능상실된 뉴클레아제의 C-말단에 융합될 수 있다.

[0321] 특정 구현예에서, 사이티딘 데아미나제 도메인을 활용하는 CBE는 우라실을 생성하기 위해 사이토신의 외환식 아민을 탈아미노화함으로써 구아닌-사이토신 염기 쌍을 티민-아데닌 염기 쌍으로 전환한다. 사이토신 데아미나제 효소의 예는 APOBEC1, APOBEC3A, APOBEC3G, CDA1, 및 AID를 포함한다. APOBEC1은 특히 단일 가닥 (ss)DNA를 기질로서 수용하지만 이중 가닥 (ds)DNA 상에서 작용할 수 없다.

[0322] 대부분의 염기-편집 시스템은 또한 의도된 염기 편집을 달리 복구할 수 있는 천연 DNA 복구 메커니즘을 중단시키는 역할을 하는 DNA 글리코실라제 억제제를 포함한다. 특정 구현예에서, DNA 글리코실라제 억제제는 Wang 등 (*Gene* 99, 31-37, 1991)에 기재된 우라실 DNA 글리코실라제 억제제 단백질 (UGI)과 같은 우라실 글리코실라제 억제제를 포함한다.

[0323] 염기 편집기의 구성요소는 직접적으로(예를 들어, 직접 공유 결합에 의해) 또는 링커를 통해 융합될 수 있다. 예를 들어, 촉매적으로 기능상실된 뉴클레아제는 링커를 통해 데아미나제 효소 및/또는 글리코실라제 억제제에 융합될 수 있다. 다중 글리코실라제 억제제는 또한 링커를 통해 융합될 수 있다. 당업자에 의해 이해되는 바와 같이, 링커는 임의의 펩티드 또는 이의 부분을 연결하는 데 사용될 수 있다.

[0324] 예시적인 링커는 중합체성 링커(예를 들어, 폴리에틸렌, 폴리에틸렌 글리콜, 폴리아미드, 폴리에스테르); 아미노산 링커; 탄소-질소 결합 아미드 링커; 환형 또는 비환형, 치환 또는 비치환, 분지형 또는 비분지형 지방족 또는 헤테로지방족 링커; 단량체성, 이량체성, 또는 중합체성 아미노알칸산 링커; 아미노알칸산(예를 들어, 글리신, 에탄산, 알라닌, β-알라닌, 3-아미노프로판산, 4-아미노부탄산, 5-펜탄산) 링커; 단량체성, 이량체성, 또는 중합체성 아미노헥산산(Ahx) 링커; 탄소환형 모이어티(예를 들어, 사이클로펜탄, 사이클로헥산) 링커; 아릴 또는 헤테로아릴 모이어티 링커; 및 페닐 고리 링커를 포함한다.

[0325] 링커는 또한 펩티드에서 링커까지 친핵체(예를 들어, 티올, 아미노)의 부착을 용이하게 하는 기능화된 모이어티를 포함할 수 있다. 임의의 친전자체가 링커의 일부로 사용될 수 있다. 예시적인 친전자체는 활성화된 에스테르, 활성화된 아미드, 마이클(Michael) 수용체, 알킬 할라이드, 아릴 할라이드, 아실 할라이드, 및 이소티오시아네이트를 포함한다.

[0326] 특정 구현예에서, 링커는 4-100 개의 아미노산 범위의 길이이다. 특정 구현예에서, 링커는 4 개의 아미노산, 9 개의 아미노산, 14 개의 아미노산, 16 개의 아미노산, 32 개의 아미노산, 또는 100 개의 아미노산이다.

[0327] 표적화된 DNA 결합 단백질을 사이티딘 데아미나제 효소 및 DNA 글리코실라제 억제제(예를 들어, UGI)와 연결함으로써 형성된 수많은 염기-편집(BE) 시스템이 기재되었다. 이러한 복합체는 예를 들어, BE1([APOBEC1-16 아미노산(aa) 링커-Sp dCas9(D10A, H840A)] Komer 등, *Nature*, 533, 420-424, 2016), BE2([APOBEC1-16aa 링커-Sp dCas9(D10A, H840A)-4aa 링커-UGI] 상기 Komer 등, 2016), BE3([APOBEC1-16aa 링커-Sp nCas9(D10A)-4aa 링커-UGI] 상기 Komer 등,), HF-BE3([APOBEC1-16aa 링커-HF nCas9(D10A)-4aa 링커-UGI] Rees 등, *Nat. Commun.* 8, 15790, 2017), BE4, BE4max([APOBEC1-32aa 링커-Sp nCas9(D10A)-9aa 링커-UGI-9aa 링커-UGI] Koblan 등, *Nat. Biotechnol* 10.1038/nbt.4172, 2018; Komer 등, *Sci. Adv.*, 3, eaao4774, 2017), BE4-GAM([Gam-16aa 링커-APOBEC1-32aa 링커-Sp nCas9(D10A)-9aa 링커-UGI-9aa 링커-UGI] 상기 Komer 등, 2017), YE1-BE3([APOBEC1(W90Y, R126E)-16aa 링커-Sp nCas9(D10A)-4aa 링커-UGI] Kim 등, *Nat. Biotechnol.* 35, 475-480, 2017), EE-BE3([APOBEC1(R126E, R132E)-16aa 링커-Sp nCas9(D10A)-4aa 링커-UGI] 상기 Kim 등, 2017), YE2-BE3([APOBEC1(W90Y, R132E)-16aa 링커-Sp nCas9(D10A)-4aa 링커-UGI] 상기 Kim 등, 2017), YEE-BE3([APOBEC1(W90Y, R126E, R132E)-16aa 링커-Sp nCas9(D10A)-4aa 링커-UGI] 상기 Kim 등, 2017), VQR-BE3([APOBEC1-16aa 링커-Sp VQR nCas9(D10A)-4aa 링커-UGI] 상기 Kim 등, 2017), VRER-BE3([APOBEC1-16aa 링커-Sp VRER nCas9(D10A)-4aa 링커-UGI] Kim 등, *Nat. Biotechnol.* 35, 475-480, 2017), Sa-BE3([APOBEC1-16aa 링커-Sa nCas9(D10A)-4aa 링커-UGI] 상기 Kim 등, 2017), SA-BE4([APOBEC1-32aa 링커-Sa nCas9(D10A)-9aa 링커-UGI-9aa 링커-UGI] 상기 Komer 등, 2017), SaBE4-Gam([Gam-16aa 링커-APOBEC1-32aa 링커-Sa nCas9(D10A)-9aa 링커-UGI-9aa 링커-UGI] 상기 Komer 등, 2017), SaKKH-BE3([APOBEC1-16aa 링커-Sa KKH nCas9(D10A)-4aa 링커-UGI] 상기 Kim 등, 2017), Cas12a-BE([APOBEC1-16aa 링커-dCas12a-14aa 링커-UGI], Li 등, *Nat. Biotechnol.* 36, 324-327, 2018), 표적-AID([Sp nCas9(D10A)-100aa 링커-CDA1-9aa 링커-UGI] Nishida 등, *Science,* 353, 10.1126/science.aaf8729, 2016), 표적-AID-NG([Sp nCas9(D10A)-NG-100aa 링커-CDA1-9aa 링커-UGI] Nishimasu 등, *Science*, 361(6408): 1259-1262, 2018), xBE3([APOBEC1-16aa 링커-xCas9(D10A)-4aa 링커-UGI] Hu 등, *Nature*, 556, 57-63, 2018), eA3A-BE3([APOBEC3A(N37G)-16aa 링커-Sp nCas9(D10A)-4aa 링커-UGI] Gerkhe 등, *Nat. Biotechnol.*, 10.1038/nbt.4199, 2018), A3A-BE3([hAPOBEC3A-16aa 링커-Sp nCas9(D10A)-4aa 링커-UGI] Wang 등, *Nat. Biotechnol.* 10.1038/nbt.4198, 2018), 및 BE-PLUS([10X GCN4-Sp nCas9(D10A) / ScFv-rAPOBEC1-UGI] Jiang 등, *Cell. Res*, 10.1038/s41422-018-0052-4, 2018)를 포함한다. 아데닌 데아미나제 염기 편집기를 포함하는 BE 복합체의 추가적인 예를 위해, Rees & Liu *Nat. Rev Genet.* 19(12): 770-788, 2018 참조.

[0328] 염기 편집기에 관한 추가적인 정보를 위해, US2018/0312825A1, WO2018/165629A, Urnov 등, *Nat Rev Genet.* 11(9):636-46, 2010; Joung 등, *Nat Rev Mol Cell Biol.* 14(1):49-55, 2013; Charpentier 등, *Nature.*; 495(7439):50-1, 2013; Seo & Kim, *Nature Medicine,* 24, 1493-1495, 2018, 및 Rees & Liu, *Nature Reviews Genetics,* 19, 770-78, 2018을 참조하며, 이들 각각은 그 전문이 참조로 포함되고 염기 편집기에 대해 특이적이다. 본 개시내용의 다양한 구현예에서 사용될 수 있는 특정 염기 편집기 작제물은 Zafra 등, *Nat Biotech*, 36(9):888-893, 2018, 및 Koblan 등, *Nat Biotech* 36(9):843-846, 2018에 기재되어 있으며, 이들 각각은 염기

편집기 작제물과 특이적으로 관련하여 그 전문이 본원에 참조로 포함된다.

[0329] 일부 구현예에서 염기 편집기 시스템은 예를 들어, γ-글로빈의 발현을 증가시키기 위해 γ-글로빈을 암호화하는 핵산 서열을 변형시키도록 조작된다. 헤모글로빈의 주요 태아 형태인 헤모글로빈 F(HbF)는 γ-글로빈 폴리펩티드와 α-글로빈 폴리펩티드의 쌍형성에 의해 형성된다. 인간 태아 γ -글로빈 유전자(HBG1 및 HBG2; 진화적 복제에 의해 생성된 2 개의 고도로 상동성인 유전자)는 정상적으로 출생 시 침묵하는 반면, 성인 β-글로빈 유전자 발현(HBB 및 HBD)의 발현은 증가하다. 일생에 걸쳐 태아 γ-글로빈의 발현을 야기하거나 또는 허용하는 돌연변이는 β-글로빈 결칩의 표현형을 개선할 수 있다. 따라서, 태아 γ-글로빈 유전자의 재활성화는 특히 β-글로빈 결핍이 있는 대상체에서 치료적으로 유익할 수 있다. γ-글로빈의 발현을 증가시키는 다양한 돌연변이는 당업계에 알려져 있고 본원에 개시되어 있다(예를 들어, Wienert *Trends in Genetics* 34(12): 927-940, 2018을 참조하며, γ-글로빈의 발현을 증가시키는 돌연변이와 관련하여 이의 전문이 본원에 참조로 포함되고). 이러한 특정 돌연변이는 HBG1 프로모터 또는 HBG2 프로모터에서 발견된다.

[0330] 일부 구현예에서, 벡터 또는 게놈은 페이로드가 통합 요소를 포함하고 염기 편집 시스템의 적어도 하나의 구성요소가 페이로드에 존재하지만 통합 요소의 외부(예를 들어, 트랜스포존 도립 반복부에 의해 플랭킹된 전위성 통합 요소를 포함하는 페이로드 단편의 외부 또는 상동 통합을 위한 상동성 아암을 포함하는 페이로드 단편의 외부)에 존재하는 염기 편집 시스템을 포함한다. 페이로드가 전위성 통합 요소를 포함하는 특정한 특정 구현예에서, 전위성 통합 요소가 트랜스포존 도립 반복부에 의해 플랭킹된 경우, 염기 편집 효소 및/또는 염기 편집 시스템의 하나 이상의 gRNA 중 하나 이상은 전위성 통합 요소의 외부의 위치(즉, 내부에 존재하지 않음)에서 페이로드에 존재한다(즉, 트랜스포존 도립 반복부에 의해 플랭킹된 핵산 서열에 존재하지 않음). 페이로드가 전위성 통합 요소를 포함하는 특정한 특정 구현예에서, 전위성 통합 요소가 상동성 아암에 의해 플랭킹된 경우, 염기 편집 효소 및/또는 염기 편집 시스템의 하나 이상의 gRNA 중 하나 이상은 통합 요소의 외부 위치(즉, 내부에 존재하지 않음)에서 페이로드에 존재한다(즉, 상동성 아암에 의해 플랭킹된 핵산 서열에 존재하지 않음). 이러한 시스템에서, 염기 편집 시스템의 발현 및/또는 활성은 일시적이며, 전위성 통합 요소의 해당 전위에서 벡터를 파괴하고 전위성 통합 요소의 외부에 위치한 염기 편집 시스템 구성요소 중 하나 이상의 발현을 감소시키거나 또는 종결시킬 수 있다. (예를 들어, 전위 또는 상동 재조합에 의한) 통합 요소의 통합은 염기 편집 시스템의 발현 및/또는 활성을 불활성화할 수 있기 때문에 염기 편집 시스템을 포함하는 이러한 벡터는 때때로 "자기-불활성화" 염기 편집 시스템 또는 벡터로 지칭될 수 있다. 다양한 구현예에서, 자기-불활성화 염기 편집 시스템은 조합 페이로드에 존재한다.

[0331] 본 개시내용은 자기-불활성화 염기 편집 시스템 페이로드를 포함하는 아데노바이러스 벡터(예를 들어, HDAd 아데노바이러스 벡터)가 다른 염기 편집 시스템 페이로드와 비교하여 예를 들어, 유전자 요법(예를 들어, 생체내 유전자 요법)에서 증가된 절단 빈도 및/또는 형질도입 및/또는 편집된 표적 세포의 증가된 생존(예를 들어, 형질도입 HSPC의 증가된 생존)을 생성할 수 있음을 포함하며, 여기서 염기 편집 시스템은 통합 요소 내에 완전히 있거나 또는 염기 편집 시스템은 숙주 세포 게놈 내로 통합되지 않지만 발현은 벡터 파괴에 의해 불활성화되지 않는다. 염기 편집 시스템의 자기-불활성화는 염기 편집기 효소 및/또는 gRNA의 발현을 단축시키고, 편집된 세포의 생존을 증가시키고, 장기간 재증식하는 세포의 백분율을 증가시킨다, 예를 들어, HBG1 및/또는 HBG2의 재활성화를 위한 자기-불활성화 염기 편집 시스템을 포함하고 γ-글로빈의 발현을 위한 핵산 서열을 추가로 포함하는 조합 페이로드를 포함하는 HDAd 벡터를 사용한 유전자 요법은 비-불활성화 염기 편집 시스템 또는 γ-글로빈 단독의 발현을 위한 핵산 서열을 포함하는 해당 HDAd 벡터를 형질도입한 후 RBC에서 상당히 더 높은 γ-글로빈을 생성할 수 있다.

[0332] 본원에는 자기-불활성화 염기 편집 시스템을 포함하는 공여자 벡터를 예를 들어, 통합 요소의 전위를 위한 트랜스포사제를 암호화하는 지지 벡터 또는 게놈과 조합하여 인간 대상체에게 투여하는 방법이 추가로 제공된다. 본 개시내용은 다양한 경우에 공여자 벡터가 지지 벡터의 투여 전에 투여되는 것을 포함하며, 여기서 공여자 벡터의 투여 및 지지 벡터의 투여 사이의 시간 기간은 염기 편집 시스템의 지속기간 및/또는 활성 수준을 조절하는 수단을 제공한다. 예를 들면, 다양한 구현예에서, 지지 벡터는 예를 들어, 공여자 벡터의 투여 후 일정 기간에 대상체에게 투여될 수 있으며 여기서 일정 기간은 적어도 1, 2, 3, 4, 5, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 30, 36, 42, 48, 54, 60, 66, 또는 72, 96, 또는 128 시간이다(예를 들어, 여기서 기간은 1, 2, 3, 4, 5, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 30, 36, 42, 48, 54, 60, 66, 또는 72 시간의 하한 및 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 30, 36, 42, 48, 54, 60, 66, 72, 96, 또는 128 시간의 상한을 갖는다).

[0333] 일부 구현예에서, 염기 편집 시스템 구성요소를 암호화하는(예를 들어, 염기 편집 효소를 암호화하는) 핵산 서

열은 염기 편집기 발현 및/또는 활성의 microRNA 조절을 위한 microRNA 표적 부위를 포함하도록 조작된다.

[0334] 본 개시내용은 ABE 시스템의 활용 시 문제를 추가로 인식하고 해결하였다. 본 개시내용은 염기 편집기 TadA 및 TadA* 서열의 반복성 및/또는 서열 유사성이 예를 들어, 생체내 유전자 요법을 위해 암호화된 염기 편집 시스템의 발현 및/또는 활성에 대한 이러한 벡터의 효능을 감소시키는 상동 재조합을 초래하 수 있다는 인식을 포함한다. 본 발명자들이 아는 한, 본 개시내용은 예를 들어, 생체내 유전자 요법에서 관찰된 바와 같이 이 문제의 최초 인식을 나타낸다. 문제를 해결하기 위해, TadA 및/또는 TadA*를 변형시켜 유사한 서열 사이에 감소된 상동성을 달성하였다. 다양한 구현예에서, TadA 및 TadA*를 암호화하는 핵산 서열의 5 개의 상응하는 코돈은 상이한 뉴클레오티드 서열을 갖도록 조작되며, 임의적으로 여기서 조작은 TadA 또는 TadA* 뉴클레오티드 서열에서 초기 코돈 서열을 관련 시스템, 예를 들어, 인간에서 코돈 용법에 따라 동일한 아미노산을 암호화하는 상이한 코돈 서열로 대체하는 것을 포함한다. 다양한 구현예에서, 적어도 5, 10, 15, 20, 25, 30, 35, 40, 45, 또는 50 개의 코돈은 TadA 및 TadA*를 각각 암호화하는 핵산 서열 사이에서 상이하도록 조작된다. 예시적인 조작된 서열은 도 132c에 제시되어 있다.

[0335] 다양한 구현예에서, ABE는 다음 TadA 및 TadA* 서열에 대해 적어도 하나의 서열 변형을 포함하는 TadA 및 TadA* 서열을 포함하며, 이는 예를 들어, ABE를 암호화하는 서열에서 링커에 의해 직접적으로 융합되거나 또는 분리될 수 있다. 다양한 구현예에서 TadA 서열은 하기 TadA 서열과 적어도 80% 동일성(예를 들어, 적어도 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일성)을 갖는 서열이며 본원에 제공된 임의의 또는 모든 TadA 변형을 포함할 수 있다. 다양한 구현예에서 TadA* 서열은 하기 TadA* 서열과 적어도 80% 동일성(예를 들어, 적어도 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일성)을 갖는 서열이며 본원에 제공된 임의의 또는 모든 TadA* 변형을 포함할 수 있다. 다양한 구현예에서 본 개시내용의 TadA 및/또는 TadA* 서열은 32 개의 아미노산 링커와 같은 링커를 포함하거나, 또는 포함하지 않을 수 있다. 하기 제공된 TadA 및/또는 TadA* 서열을 포함하는 것들을 포함하는 다양한 서열 및 구현예에서, 서열은 32 개의 아미노산 링커를 암호화하는 96 개의 뉴클레오티드의 3' 서열을 포함할 수 있다. 따라서, 다양한 구현예에서 TadA 서열은 하기 TadA 서열의 뉴클레오티드 1-498(96 개의 3' 뉴클레오티드 제외)과 적어도 80% 동일성(예를 들어, 적어도 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일성)을 갖는 서열이며 본원에 제공된 임의의 또는 모든 상응하는 TadA 변형을 포함할 수 있다. 또한 따라서, 다양한 구현예에서 TadA* 서열은 하기 TadA* 서열의 뉴클레오티드 1-498(96 개의 3' 뉴클레오티드 제외)와 적어도 80% 동일성(예를 들어, 적어도 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일성)을 갖는 서열이며 본원에 제공된 임의의 또는 모든 상응하는 TadA* 변형을 포함할 수 있다.

[0336] 다양한 구현예에서, ABE의 TadA 및/또는 TadA*의 서열은 TadA 및 TadA* 사이(또는 예를 들어, 뉴클레오티드 1 내지 579 또는 1 내지 498을 포함하는 이의 정렬된 부분)의 퍼센트 동일성을 80% 미만(예를 들어, 80%, 75%, 70%, 65%, 60%, 55%, 50%, 45%, 또는 40% 미만, 또는 60% 내지 80%, 65% 내지 80%, 70%, 내지 80%, 75% 내지 80%, 60% 내지 75%, 65% 내지 75%, 70% 내지 75%, 60% 내지 70%, 또는 65% 내지 70%인 퍼센트 동일성)으로 감소시키도록 조작된다. 다른 것에 의해 생성된 pCMV-ABEmax 플라스미드(Addgene #112095)에서, 81.6%의 동일성을 갖는 2 개의 594 bp TadA+32aa 반복부 사이에 109 bp 불일치가 있다. 다양한 본 구현예에서 TadA 및/또는 TadA* 변형을 위한 부위는 하기 서열에서 밑줄로 표시되고 하기 표에 기재된 것들을 포함한다. 다양한 구현예에서, TadA* 서열은 TadA* 변형 표(표 11)에 제시된 것들에 상응하는 하나 이상, 또는 모든 변형을 포함한다. 다양한 구현예에서, TadA 서열은 TadA 변형 표(표 10)에 제시된 하나 이상, 또는 모든 변형을 포함하고 TadA* 서열은 TadA* 변형 표(표 11)에 제시된 것들에 상응하는 하나 이상, 또는 모든 변형을 포함한다. 특정한 특정 구현예에서, TadA 서열은 TadA 변형 표(표 10; 서열번호: 280 참조)에 제시된 것들에 상응하는 0, 1, 2, 3, 4, 5, 6, 7, 8. 9. 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 또는 25 개의 변형(예를 들어, 1 내지 5, 5 내지 10, 5 내지 20, 5 내지 25, 10 내지 20, 10 내지 25, 15 내지 20, 15 내지 25, 또는 20 내지 25 개의 변형)을 포함하고 TadA* 서열은 TadA* 변형 표(표 11; 서열번호: 281 참조)에 제시된 것들에 상응하는 0, 1, 2, 3, 4, 5, 6, 7, 8. 9. 10, 11, 12, 13, 14, 15, 또는 16 개의 변형(예를 들어, 1 내지 5, 5 내지 10, 5 내지 16, 또는 10 내지 16 개의 변형)을 포함한다.

[0337] 당업자가 이해하는 바와 같이, 감소된-동일성 TadA 및 TadA* 서열은 생체내 및 생체외 유전적 조작을 포함하나 이에 제한되지 않는 유전적 조작 분야에서 일반적인 유용성을 갖는다. 감소된 동일성을 갖도록 조작된 TadA 및 TadA* 서열은 또한 예를 들어, 생체내 유전자 요법을 위해 페이로드(예를 들어, 본 개시내용의 페이로드), 예를 들어, Ad35, Ad35++, HDAd35, 또는 HDAd35++ 공여자 벡터 또는 공여자 게놈과 같은 아데노바이러스 벡터 또는 게놈에 포함된다.

[0338] 표 10: TadA 변형 표

| 위치 | 뉴클레오티드 변화 |
|---|---|
| 15 | T > C |
| 57 | G > A |
| 63 | A > C |
| 69 | T > C |
| 87 | G > C |
| 112 | A > C |
| 114 | A > C |
| 126 | G > A |
| 147 | C > A |
| 198 | C > A |
| 216 | C > T |
| 289 | A > C |
| 318 | G > A |
| 333 | C > A |
| 343 | T > A |
| 344 | C > G |
| 369 | C > A |
| 402 | A > C |
| 451 | C > A |
| 507 | A > C |
| 547 | A > T |
| 548 | G > C |
| 568 | A > T |
| 569 | G > C |
| 570 | C > T |

[0339]

[0340] 표 11: TadA* 변형 표

| 위치 | 뉴클레오티드 변화 |
|---|---|
| 321 | C > T |
| 330 | C > T |
| 345 | C > T |
| 382 | C > A |
| 384 | C > A |
| 465 | C > T |
| 498 | C > T |
| 499 | T > A |
| 500 | C > G |
| 501 | C > T |
| 504 | A > C |
| 516 | A > G |
| 537 | A > C |
| 592 | T > A |
| 593 | C > G |
| 594 | A > C |

*참조 서열:*

TadA(서열번호: 280)

TadA*(서열번호: 281)

[0341]

[0342] 당업자는 TadA 서열 및 TadA* 서열을 포함하는 ABE에 존재하는 TadA 변형 표 및/또는 TadA* 변형 표의 것들에 상응하는 다수의 변형이 적어도 TadA 및 TadA* 뉴클레오티드 서열 사이의 동일성 감소가 임의의 특정 변형을 필요로 하지 않지만 오히려 TadA 및 TadA* 서열 사이의 동일성의 전반적인 변화를 필요로 하는 식별된 문제에 대한 해결책인 한, 선택된 특정 변형의 고려 없이 유의할 수 있음을 추가로 이해할 것이다. 따라서, 본 개시내용은 예시적인 변형을 제공하지만, 임의의 특정 변형의 포함 또는 제외는 본원에 제시된 해결책에 중요하지 않다. 따라서 본 개시내용은 TadA 및 TadA* 변형 표에 제시된 하나 이상의 변형을 포함하고 80% 미만(예를 들어, 80%,

75%, 70%, 65%, 60%, 55%, 50%, 45%, 또는 40% 미만)인 TadA 및 TadA* 사이(또는 예를 들어, 뉴클레오티드 1 내지 579를 포함하는 이의 정렬된 부분)의 퍼센트 동일성을 갖는 TadA 및 TadA*의 감소된-동일성 서열을 포함한다.

[0343] 의심의 여지를 피하기 위해, 제공된 서열은 하기 TadA 및 TadA* 서열의 상응하는 뉴클레오티드 위치에 대한 비교에 의해 본원에 제공된 임의의 TadA 또는 TadA* 서열 변형을 포함하거나 또는 포함하지 않는 것으로 식별될 수 있다. 따라서, 본원에 제공된 임의의 TadA 또는 TadA* 서열 변형의 존재 또는 부재의 결정은 임의의 제공된 서열의 기원 또는 이력에 의존하지 않고 서열 그 자체로부터 단독으로 결정될 수 있다.

[0344] 당업자는 본 개시내용의 ABE 시스템, 뿐만 아니라 이의 TadA 및 TadA* 서열이 본 명세서에 제시된 본 맥락 또는 임의의 다른 맥락으로 제한되지 않는, 예를 들어, 특정 벡터, 혈청형, 또는 다른 맥락에서 사용하는 것으로 제한되지 않는 일반적인 유용성의 기여를 나타냄을 이해할 것이다. 실제로, 본 개시내용의 서열은 염기 편집 구성요소를 암호화하거나 또는 포함할 수 있는 임의의 실험 시스템에서 생체내, 시험관내, 또는 생체외에서 사용될 수 있다. 서열은 다양한 분자 생물학 적용에서 도구로서 유용하다.

[0345] **I(C)(i)(c). 작은 RNA 페이로드 발현 산물**

[0346] 작은 RNA는 유전자 발현을 조절하는 데 있어서 역할을 하는 짧은 비-코딩 RNA 분자이다. 특정 구현예에서, 작은 RNA는 200 개 미만의 뉴클레오티드 길이이다. 특정 구현예에서, 작은 RNA는 100 개 미만의 뉴클레오티드 길이이다. 특정 구현예에서, 작은 RNA는 50, 45, 40, 35, 30, 25, 또는 20 개 미만의 뉴클레오티드 길이이다. 특정 구현예에서, 작은 RNA는 20 개 미만의 뉴클레오티드 길이이다. 다양한 구현예에서 작은 RNA는 5, 10, 15, 20, 25, 또는 30 개 뉴클레오티드의 하한 및 20, 25, 30, 35, 40, 45, 50, 75, 또는 100 개 뉴클레오티드의 상한을 갖는 길이를 갖는다. 작은 RNA는 microRNA(miRNA, Piwi-상호작용 RNA(piRNA), 작은 간섭 RNA(siRNA), 작은 핵소체 RNA(snoRNA), tRNA-유래 작은 RNA(tsRNA) 작은 rDNA-유래 RNA(srRNA), 및 작은 핵 RNA를 포함하나 이에 제한되지 않는다. 작은 RNA의 추가적인 클래스는 계속 발견되고 있다.

[0347] 특정 구현예에서, 표적 mRNA에 상동성이거나 또는 간섭 RNA가 혼성화할 수 있는 간섭 RNA 분자는 RNA 간섭(RNAi)으로 지칭되는 과정인 표적 mRNA 분자의 분해 또는 표적 mRNA의 감소된 번역을 야기할 수 있다(Carthew, *Curr. Opin. Cell. Biol.* 13: 244-248, 2001). RNAi는 외래 RNA(예를 들어, 바이러스 RNA)를 제거하기 위해 자연적으로 세포에서 발생한다. 일부 경우에, 천연 RNAi는 분해 메커니즘을 다른 유사한 RNA 서열로 지시하는 유리 이중 가닥 RNA(dsRNA)로부터 절단된 단편을 통해 진행된다. 대안적으로, RNAi는 예를 들어, 표적 유전자의 발현을 침묵시키기 위해 제조될 수 있다. 예시적인 RNAi 분자는 작은 헤어핀 RNA(shRNA, 또한 짧은 헤어핀 RNA로도 지칭됨) 및 작은 간섭 RNA(siRNA)를 포함한다.

[0348] 본 개시내용을 제한하지 않고, 이론에 의해 얽매이지 않고, 자연 및/또는 일부 구현예에서 RNA 간섭은 전형적으로 2-단계 과정이다. 첫번째 단계인 개시 단계에서, 유입 dsRNA는 아마도 dsRNA-특이적 리보뉴클레아제의 리보뉴클레아제(RNase) III 패밀리의 구성원인 Dicer의 작용에 의해 21-23 개 뉴클레오티드(nt) siRNA로 소화되며, 이는 ATP-의존적 방식으로 (직접적으로 또는 이식유전자 또는 바이러스를 통해 도입된) dsRNA를 처리한다(절단한다). 성공적인 절단 사건은 RNA를 19-21 개 염기 쌍(bp) 이중체(siRNA)로 분해하며, 각각은 2-뉴클레오티드 3' 돌출부를 갖는다(Hutvagner & Zamore, *Curr. Opin. Genet. Dev.* 12: 225-232, 2002; Bernstein, *Nature* 409:363-366, 2001).

[0349] 두번째 단계인 효과기 단계에서, siRNA 이중체는 뉴클레아제 복합체에 결합되어 RNA-유도된 침묵 복합체(RISC)를 형성한다. RISC의 활성화를 위해 siRNA 이중체의 ATP-의존적 풀림이 요구된다. 그런 다음 활성 RISC는 염기 쌍형성 상호작용에 의해 상동 전사체를 표적화하고 전형적으로 mRNA를 siRNA의 3' 말단으로부터 12 개의 뉴클레오티드 단편으로 절단한다(Hutvagner & Zamore, *Curr. Opin. Genet. Dev.* 12: 225-232, 2002; Hammond 등, *Nat. Rev. Gen.* 2:110-119, 2001; Sharp, *Genes. Dev.* 15:485-490, 2001). 연구는 각 RISC가 단일 siRNA 및 RNase를 함유함을 나타낸다(Hutvagner & Zamore, *Curr. Opin. Genet. Dev.* 12: 225-232, 2002).

[0350] RNAi의 현저한 효력으로 인해, RNAi 경로 내에서 증폭 단계가 제안되었다. 증폭은 더 많은 siRNA를 생성하는 유입 dsRNA의 카피에 의해, 또는 형성된 siRNA의 복제에 의해 발생할 수 있다. 대안적으로 또는 추가적으로, 증폭은 RISC의 다중 전환 사건에 의해 실행될 수 있다(Hutvagner & Zamore, *Curr. Opin. Genet. Dev.* 12: 225-232, 2002; Hammond 등, *Nat. Rev. Gen.* 2:110-119, 2001; Sharp, Genes. *Dev.* 15:485-490, 2001). 또한 RNAi는 Tuschl(*Chem. Biochem.* 2: 239-245, 2001); Cullen(*Nat. Immunol.* 3:597-599, 2002); 및 Brantl(*Biochem. Biophys. Act.* 1575:15-25, 2002)에 기재되어 있다.

[0351] 일부 구현예에서, 본 개시내용과 함께 사용하기에 적합한 RNAi 분자의 합성은 다음과 같이 수행될 수 있다. 첫째로, mRNA 서열은 표적화된 이식유전자의 시작 코돈의 하류에서 스캐닝될 수 있다. 각 AA 및 3' 인접 19 개 뉴클레오티드의 발생은 잠재적인 siRNA 표적 부위로서 기록된다. 특정 구현예에서, siRNA 표적 부위는 비번역된 영역(UTR)이 조절 단백질 결합 부위에 더 풍부할 때, 오픈 리딩 프레임으로부터 선택될 수 있다. UTR-결합 단백질 및/또는 번역 개시 복합체는 siRNA 엔도뉴클레아제 복합체의 결합을 방해할 수 있다(Tuschl, *Chem. Biochem.* 2: 239-245, 2001). 비번역된 영역에서 지시된 siRNA는 또한 글리세르알데하이드 3-포스페이트 데하이드로게나제(GAPDH)에 대해 입증된 바와 같이, 효과적일 수 있음이 이해될 것이며, 여기서5' UTR에서 지시된 siRNA는 세포 GAPDH mRNA에서 90% 감소를 매개하고 단백질 수준을 완전히 폐지하였다. 둘째로, 잠재적인 표적 부위는 미국 국립생물공학정보센터(National Center for Biotechnology Information)(NCBI) 서버에서 이용가능한 Basic Local Alignment Search Tool(BLAST) 소프트웨어와 같은 임의의 서열 정렬 소프트웨어를 사용하여 적절한 게놈 데이터베이스와 비교될 수 있다. 다른 코딩 서열과 상당한 상동성을 나타내는 추정 표적 부위가 필터링될 수 있다.

[0352] 적격한 표적 서열이 siRNA 합성을 위한 주형으로 선택될 수 있다. 선택된 서열은 G/C 함량이 55% 초과인 것과 비교하여 유전자 침묵을 매개하는 데 더 효과적인 것으로 제시되었을 때 G/C 함량이 낮은 것들을 포함할 수 있다. 여러 표적 부위가 평가를 위한 표적 유전자의 길이에 따라 선택될 수 있다. 선택된 siRNA의 더 나은 평가를 위해, 음성 대조군이 사용될 수 있다. 음성 대조군 siRNA는 siRNA와 동일한 뉴클레오티드 조성물을 포함할 수 있지만 게놈에 대한 상동성이 상당히 결여되어 있다. 따라서, 다른 유전자와 임의의 상당한 상동성을 나타내지 않으면, siRNA의 스크램블된 뉴클레오티드 서열이 사용될 수 있다.

[0353] 센스 가닥은 선택된 부분의 서열을 기반으로 설계될 수 있다. 안티센스 가닥은 일상적으로 센스 가닥과 동일한 길이이며 상보적 뉴클레오티드를 포함한다. 특정 구현예에서, 가닥은 정렬되거나 또는 어닐링될 때 완전히 상보적이고 말단이 뭉툭해진다. 다른 구현예에서, 가닥은 1-, 2- 또는 3-뉴클레오티드 돌출부가 생성되도록, 즉, 센스 가닥의 3' 단부가 안티센스 가닥의 5' 단부보다 1, 2 또는 3 개의 뉴클레오티드를 추가로 확장하고/하거나 안티센스 가닥의 3' 단부가 센스 가닥의 5' 단부보다 1, 2 또는 3 개의 뉴클레오티드를 추가로 확장하도록 정렬 또는 어닐링된다. 돌출부는 표적 유전자 서열(또는 이의 보체)에 상응하는 뉴클레오티드를 포함할 수 있다. 대안적으로, 돌출부는 데옥시리보뉴클레오티드, 예를 들어 데옥시티민(dT), 또는 뉴클레오티드 유사체, 또는 다른 적합한 비-뉴클레오티드 물질을 포함할 수 있다.

[0354] 안티센스 가닥의 RISC로의 진입을 용이하게 하기 위해(따라서 표적 절단 및 침묵의 효율을 증가시키거나 또는 개선하기 위해), 센스 가닥의 5' 단부 및 안티센스 가닥의 3' 단부 사이의 염기 쌍 강도는 변경, 예를 들어, 줄어들거나 또는 감소될 수 있다. 특정 구현예에서, 염기 쌍 강도는 제1 또는 안티센스 가닥의 3' 단부 및 제2 또는 센스 가닥의 5' 단부 사이 보다 제1 또는 안티센스 가닥의 5' 단부 및 제2 또는 센스 가닥의 3' 단부 사이의 더 적은 G:C 염기 쌍으로 인해 더 적다. 특정 구현예에서, 염기 쌍 강도는 제1 또는 안티센스 가닥의 5' 단부 및 제2 또는 센스 가닥의 3' 단부 사이의 적어도 하나의 불일치된 염기 쌍으로 인해 더 적다. 바람직하게는, 불일치된 염기 쌍은 G:A, C:A, C:U, G:G, A:A, C:C 및 U:U를 포함하는 군으로부터 선택된다. 또 다른 구현예에서, 염기 쌍 강도는 제1 또는 안티센스 가닥의 5' 단부 및 제2 또는 센스 가닥의 3' 단부 사이의 적어도 하나의 워블(wobble) 염기 쌍, 예를 들어, G:U로 인해 더 적다. 또 다른 구현예에서, 염기 쌍 강도는 드문 뉴클레오티드, 예를 들어, 이노신(I)을 포함하는 적어도 하나의 염기 쌍으로 인해 더 적다. 특정 구현예에서, 염기 쌍은 I:A, I:U 및 I:C를 포함하는 군으로부터 선택된다. 또한 또 다른 구현예에서, 염기 쌍 강도는 변형된 뉴클레오티드를 포함하는 적어도 하나의 염기 쌍으로 인해 더 적다. 특정 구현예에서, 변형된 뉴클레오티드는 예를 들어, 2-아미노-G, 2-아미노-A, 2,6-디아미노-G, 및 2,6-디아미노-A로부터 선택된다.

[0355] ShRNA는 헤어핀 루프 구조를 갖는 단일-가닥 폴리뉴클레오티드이다. 단일-가닥 폴리뉴클레오티드는 이중-가닥 영역에서 하나의 가닥의 3' 단부 및 이중-가닥 영역에서 다른 가닥의 5' 단부를 연결하는 루프 분절을 갖는다. 이중-가닥 영역은 표적 서열에 혼성화가능한 제1 서열, 예컨대 이식유전자를 암호화하는 폴리뉴클레오티드, 및 제1 서열에 상보적인 제2 서열로부터 형성되며, 따라서 제1 및 제2 서열은 연결 서열이 단부를 연결하여 헤어핀 루프 구조를 형성하는 이중 가닥 영역을 형성한다. 제1 서열은 이식유전자를 암호화하는 폴리뉴클레오티드의 임의의 부분에 혼성화가능할 수 있다. shRNA의 이중 가닥 줄기 도메인은 제한 엔도뉴클레아제 부위를 포함할 수 있다.

[0356] shRNA의 전사는 폴리머라제 III(Pol III) 프로모터에서 개시되어 4-5-티민 전사 종결 부위의 위치 2에서 종결되는 것으로 생각된다. 발현 시, shRNA는 3' UU-돌출부가 있는 줄기-루프 구조로 접히는 것으로 생각되며; 후속적으로, 이러한 shRNA의 단부가 처리되어, shRNA를 21-23 개 뉴클레오티드의 siRNA-유사 분자로 전환시킨다

(Brummelkamp 등, *Science*. 296(5567):550-553, 2002; Lee 등, *Nature Biotechnol*. 20(5):500-505, 2002; Miyagishi & Taira, *Nature Biotechnol*. 20(5):497-500, 2002; Paddison 등, *Genes & Dev*. 16(8): 948-958, 2002; Paul 등, *Nature Biotechnol*. 20(5):505-508, 2002; Sui, *Proc. Natl. Acad. Sci. USA*. 99(6):5515-5520, 2002; Yu 등, *Proc. Natl. Acad. Sci. USA*. 99(9):6047-6052, 2002).

[0357] shRNA의 줄기-루프 구조는 임의적인 뉴클레오티드 돌출부, 예컨대 2-bp 돌출부, 예를 들어, 3' UU 돌출부를 가질 수 있다. 변경이 있을 수 있지만, 줄기는 전형적으로 15 내지 49, 15 내지 35, 19 내지 35, 21 내지 31 bp, 또는 21 내지 29 bp 범위이고, 루프는 4 내지 30 bp, 예를 들어, 4 내지 23 bp 범위일 수 있다. 특정 구현예에서, shRNA 서열은 45-65 bp; 50-60 bp; 또는 51, 52, 53, 54, 55, 56, 57, 58, 또는 59 bp를 포함한다. 특정 구현예에서, shRNA 서열은 52 또는 55 bp를 포함한다. 특정 구현예에서 siRNA는 15-25 bp를 갖는다. 특정 구현예에서 siRNA는 16, 17, 18, 19, 20, 21, 22, 23, 또는 24 bp를 갖는다. 특정 구현예에서 siRNA는 19 bp를 갖는다. 그러나, 당업자는 16 개 미만의 뉴클레오티드 또는 24 개 초과의 뉴클레오티드 길이를 갖는 siRNA가 또한 RNAi를 매개하는 기능을 할 수 있음을 이해할 것이다. 더 긴 RNAi 제제는 바람직하지 않을 수 있는 특정 포유류 세포에서 인터페론 또는 단백질 키나제 R(PKR) 반응을 도출하는 것으로 입증되었다. 바람직하게는 RNAi 제제는 PKR 반응을 도출하지 않는다(즉, 충분히 짧은 길이의 것이다). 그러나, 더 긴 RNAi 제제는 예를 들어, PKR 반응이 대체 수단에 의해 하향조절되거나 또는 약화되는 상황에서 유용할 수 있다.

[0358] 작은 RNA가 또한 유전자 발현을 활성화하는 데 사용될 수 있다.

[0359] **I(C)(i)(d). 조합 페이로드**

[0360] 본 개시내용은 복수의 발현 산물을 암호화하는 페이로드를 포함하는 아데노바이러스 벡터 및 게놈을 포함한다. 복수의 발현 산물을 암호화하는 페이로드는 조합 페이로드로 지칭될 수 있다. 다양한 구현예에서, 조합 페이로드는 제1 발현 산물을 암호화하는 제1 핵산 서열 및 제2 발현 산물을 암호화하는 제2 핵산 서열을 포함할 수 있다. 다양한 구현예에서, 제1 및 제2 발현 산물 각각은 예를 들어, 본원에 개시된 바와 같은 단백질(예를 들어, 치료 단백질, 예를 들어, 대체 효소), 결합 도메인, 항체, CAR, TCR, CRISPR 시스템, 염기 편집기 시스템, 작은 RNA, 및/또는 선택가능한 마커 중 임의의 것으로부터 독립적으로 선택될 수 있다, 예시적인 조합 페이로드가 본원에 개시되어 있다.

[0361] 당업자는 코딩 서열이 본원에 제공되거나 또는 달리 당업계에 알려진 다양한 프로모터 및/또는 다른 조절 서열 중 임의의 것과의 작동가능한 연결에 의해 제어되고/되거나 발현될 수 있음을 이해할 것이다. 당업자가 인식하고, 본 개시내용에 예시된 바와 같이, 벡터에서 코딩 서열을 제어 및/또는 발현하는 데 이용가능한 서열은 당업계에 알려져 있고 본원에 제공된 것들을 포함한다. 다양한 특정 예에서, 본 개시내용의 페이로드에 존재하는 코딩 서열은 임의적으로 프로모터, 인핸서, 종결 영역, 절연체, 미니-LCR, 종결 신호, 폴리아데닐화 신호, 스플라이싱 신호 등으로부터 선택된 하나 이상의 조절 서열과 작동가능하게 연결될 수 있다.

[0362] 일부 구현예에서, 조합 페이로드는 CRISPR-연관 RNA-가이드된 엔도뉴클레아제 및 적어도 하나의 가이드 RNA(gRNA)를 포함하는 CRISPR 시스템의 하나 이상, 또는 모든 구성요소를 암호화하며, 임의적으로 여기서 적어도 하나의 gRNA는 1, 2, 3, 4, 또는 5 개의 gRNA, 및 임의적으로 CRISPR 시스템의 부분이 아닌 하나 이상의 추가의 코딩 서열을 포함한다. 예를 들어, CRISPR 시스템의 gRNA는 HBG1 프로모터의 핵산 서열을 표적화하는 gRNA, HBG2 프로모터의 핵산 서열을 표적화하는 gRNA, 및/또는 적혈구 인핸서 bcl11a의 핵산 서열을 표적화하는 gRNA 중 하나 이상, 또는 전부를 포함할 수 있다. 다양한 구현예에서, (i) HBG1 프로모터-표적화된 gRNA는 HBG1 프로모터에서 BCL11A 억제인자 단백질 결합 부위의 불활성화에 의해 HBG1 프로모터에 작동가능하게 연결된 γ-글로빈 코딩 서열의 발현을 증가시키도록 설계되고/되거나, (ii) HBG2 프로모터-표적화된 gRNA는 HBG2 프로모터에서 BCL11A 억제인자 단백질 결합의 불활성화에 의해 HBG2 프로모터에 작동가능하게 연결된 γ-글로빈 코딩 서열의 발현을 증가시키도록 설계되고/되거나, (iii) bcl11a-표적화된 gRNA는 bcl11a 인핸서에 작동가능하게 연결된 γ-글로빈 코딩 서열의 발현을 증가시키도록 설계되며, 여기서 적혈구 bcl11a 인핸서의 변형 및/또는 불활성화는 적혈구 세포에서 감소된 BCL11A 억제인자 단백질 발현을 초래한다. 다양한 구현예에서, CRISPR 시스템을 포함하는 조합 페이로드는 치료 단백질을 암호화하는 핵산을 추가로 포함하며, 임의적으로 여기서 치료 단백질은 γ-글로빈 및 β-글로빈 중 하나 이상으로부터 선택된다. 일부 구현예에서, 치료 단백질은 β-글로빈 프로모터 및/또는 β-글로빈 LCR과 작동가능하게 연결된다.

[0363] 일부 구현예에서, 조합 페이로드는 염기 편집 효소 및 적어도 하나의 가이드 RNA(gRNA)를 포함하는 염기 편집기 시스템 중 하나 이상, 또는 모든 구성요소를 암호화하며, 임의적으로 여기서 적어도 하나의 gRNA는 1, 2, 3, 4, 또는 5 개의 gRNA, 및 임의적으로 염기 편집기 시스템의 부분이 아닌 하나 이상의 추가의 코딩 서열을

포함한다. 예를 들어, 염기 편집기 시스템의 gRNA는 HBG1 프로모터의 핵산 서열을 표적화하는 gRNA, HBG2 프로모터의 핵산 서열을 표적화하는 gRNA, 및/또는 적혈구 인핸서 bcl11a의 핵산 서열을 표적화하는 gRNA 중 하나 이상, 또는 전부를 포함할 수 있다. 다양한 구현예에서, (i) HBG1 프로모터-표적화된 gRNA는 HBG1 프로모터에서 BCL11A 억제인자 단백질 결합 부위의 불활성화에 의해 HBG1 프로모터에 작동가능하게 연결된 γ-글로빈 코딩 서열의 발현을 증가시키도록 설계되고/되거나, (ii) HBG2 프로모터-표적화된 gRNA는 HBG2 프로모터에서 BCL11A 억제인자 단백질 결합 부위의 불활성화에 의해 HBG2 프로모터에 작동가능하게 연결된 γ-글로빈 코딩 서열의 발현을 증가시키도록 설계되고/되거나, (iii) bcl11a-표적화된 gRNA는 bcl11a 인핸서에 작동가능하게 연결된 γ-글로빈 코딩 서열의 발현을 증가시키도록 설계되며, 여기서 적혈구 bcl11a 인핸서의 변형 및/또는 불활성화는 적혈구 세포에서 감소된 BCL11A 억제인자 단백질 발현을 초래한다. 다양한 구현예에서, 염기 편집기 시스템을 포함하는 조합 페이로드는 치료 단백질을 암호화하는 핵산을 추가로 포함하며, 임의적으로 여기서 치료 단백질은 γ-글로빈 및 β-글로빈 중 하나 이상으로부터 선택된다. 일부 구현예에서, 치료 단백질은 β-글로빈 프로모터 및/또는 β-글로빈 LCR과 작동가능하게 연결된다.

[0364] 일부 구현예에서, 조합 페이로드는 항체를 암호화하는 핵산 서열을 포함한다. 일부 구현예에서 조합 페이로드는 제1 항체를 암호화하는 제1 핵산 서열 및 제2 항체를 암호화하는 제2 핵산 서열을 포함한다. 일부 구현예에서, 항체(예를 들어, 제1 및/또는 제2 항체)는 scFv이다. 일부 구현예에서 항체는 면역글로불린 중쇄 및 면역글로불린 경쇄를 포함하는 항체이다.

[0365] 다양한 구현예에서, 조합 페이로드의 페이로드 핵산 서열에 의해 암호화된 적어도 하나의 발현 산물은 선택가능한 마커이다. 다양한 구현예에서, 선택가능한 마커는 $MGMT^{P140K}$이다.

[0366] 예시적인 Ad35 페이로드 및 시스템은 다음을 포함한다:

[0367] (i) 다양한 구현예에서, Ad35 페이로드는 SB100x에 의한 전위를 위해 트랜스포사제 도립 반복부에 의해 플랭킹된 통합 요소를 포함하고, 트랜스포사제 도립 반복부는 FLPe와 같은 FLP 재조합효소에 의한 재조합을 위해 frt 지시 반복부에 의해 플랭킹된다. 다양한 구현예에서, 통합 요소는 임의적으로 5'에서 3'으로, (a) β-글로빈 미니-LCR, (b) γ-글로빈 코딩 서열이 3'UTR(예를 들어, γ-글로빈 3'UTR)과 작동가능하게 연결된 인간 γ-글로빈 코딩 서열과 작동가능하게 연결된 β-글로빈 프로모터를 포함하며, 여기서 β-글로빈 미니-LCR이 또한 γ-글로빈 코딩 서열과 작동가능하게 연결되는 유전자 (c) cHS4 절연체 서열, 및 (d) $MGMT^{P140K}$ 코딩 서열, 2A 자기-절단 펩티드, GFP 형광 마커 코딩 서열, 및 폴리아데닐화 신호와 작동가능하게 연결된 PGK 프로모터와 같은 프로모터를 포함하는 유전자를 포함하며, 임의적으로 여기서 (a)-(d) 중 임의의 것은 Ad35 페이로드의 2 개의 가닥 중 어느 한 가닥 상에서 5'에서 3' 방향으로 암호화될 수 있다.

[0368] 다양한 구현예에서, Ad35 페이로드는 통합 요소의 외부 및 재조합효소 부위의 외부에, CRISPR 시스템을 암호화하는 핵산 서열을 추가로 포함한다. 특정한 특정 구현예에서, CRISPR 시스템을 암호화하는 핵산 서열은 임의적으로 5'에서 3'으로, (a) 제1 gRNA가 bcl11a 인핸서를 표적화하는 제1 gRNA-암호화 서열과 작동가능하게 연결된 제1 U6 프로모터를 포함하는 제1 gRNA 유전자, (b) 제2 gRNA가 HBG 프로모터를 표적화하는 제2 gRNA-암호화 서열과 작동가능하게 연결된 제2 U6 프로모터를 포함하는 제2 gRNA 유전자, 및 (c) CRISPR/Cas9 코딩 서열에 작동가능하게 연결된 EF1α 프로모터와 같은 프로모터를 포함하는 CRISPR 효소 유전자를 포함하며, 여기서 CRISPR/Cas9 코딩 서열은 3'UTR/miR 서열 및 폴리아데닐화 신호와 작동가능하게 연결된다. 다양한 구현예에서, CRISPR 시스템은 HBG 프로모터의 적혈구 bcl11a 인핸서 및 BCL11A 결합 부위를 표적화하며, 각각은 γ-글로빈 활성화 또는 재활성화를 야기하는 데 기여한다. 본원에 개시된 바와 같이, CRISPR 시스템은 전위에 의한 공여자 벡터의 절단이 비-통합된 공여자 벡터 핵산의 분해를 초래한다는 점에서 자기-불활성화될 수 있다. 다양한 구현예에서, miR 서열은 HDAd35 공여자 벡터 생산 동안 생산자 세포에서 Cas9 발현을 억제하는 서열일 수 있다(예를 들어, Saydaminova 등, *Mol. Ther. Meth. Clin. Dev.* 1: 14057, 2015; Li 등, *Mol. Ther. Meth. Clin. Dev.* 9: 390-401, 2018 참조).

[0369] 다양한 구현예에서, 본 개시내용의 Ad35 시스템은 Ad35 지지 벡터를 추가로 포함하며, 여기서 지지 벡터는 임의적으로 5'에서 3'으로, (a) FLPe 재조합효소 코딩 서열에 작동가능하게 연결된 EF1α 프로모터를 포함하는 재조합 유전자, 및 (b) SB100x 트랜스포사제 코딩 서열과 작동가능하게 연결된 PGK 프로모터를 포함하는 트랜스포사제 유전자를 포함한다.

[0370] 다양한 구현예에서 Ad35 페이로드는 Ad35 공여자 벡터 게놈에 존재한다. 다양한 구현예에서 Ad35 공여자 벡터 게놈에 존재하는 Ad35 페이로드는 Ad35 ITR에 의해 플랭킹된다. 다양한 구현예에서, Ad35 공여자 벡터 게놈은

Ad35 공여자 벡터에 존재한다. 다양한 구현예에서, 공여자 벡터는 Ad35 ++ 벡터이다.

[0371] 다양한 구현예에서 지지 게놈은 Ad35 ITR을 포함한다. 다양한 구현예에서, 지지 게놈은 Ad35 벡터에 존재한다. 다양한 구현예에서, 지지 벡터는 Ad35 ++ 벡터이다.

[0372] 다양한 구현예에서, Ad35 공여자 벡터는 헬퍼 의존적 공여자 벡터(HDAd35)이다. 이러한 특정 구현예에서, 본 개시내용의 시스템은 HDAd35 공여자 벡터 또는 게놈, 및 Ad35 헬퍼 벡터 또는 게놈을 포함할 수 있고, 다양한 구현예에서 Ad35 지지 벡터를 추가로 포함할 수 있다.

[0373] 특정 예시적인 구현예는 도 164에 예시되어 있다.

[0374] (ii) 다양한 구현예에서, Ad35 페이로드는 표적 세포 게놈과 적어도 80% 동일성(예를 들어, 적어도 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%< 또는 100% 동일성)을 갖는, 상동성 아암(예를 들어, 1.8 kb 상동성 아암)에 의해 플랭킹된 통합 요소를 포함한다. 다양한 구현예에서, 통합 요소는 임의적으로 5'에서 3'으로, (a) HS1, HS2, HS3, 및 HS4를 포함하나, HS5를 포함하지 않는 β-글로빈 미니-LCR, (b) γ-글로빈 3'UTR과 작동가능하게 연결된 γ-글로빈 코딩 서열과 작동가능하게 연결된 β-글로빈 프로모터를 포함하며, 여기서 β-글로빈 미니-LCR이 또한 γ-글로빈 코딩 서열에 작동가능하게 연결되는 유전자 (c) cHS4 절연체 서열, 및 (d) 폴리아데닐화 신호와 작동가능하게 연결된 $MGMT^{P140K}$ 코딩 서열과 작동가능하게 연결된 PGK 프로모터를 포함하는 유전자를 포함하며, 임의적으로 여기서 (a)-(d) 중 임의의 것은 Ad35 페이로드의 2 개의 가닥 중 어느 한 가닥 상에서 5'에서 3' 방향으로 암호화될 수 있다.

[0375] 다양한 구현예에서, Ad35 페이로드는 통합 요소의 외부 및 재조합효소 부위의 외부에, CRISPR 시스템을 암호화하는 핵산 서열을 추가로 포함한다. 특정한 특정 구현예에서, CRISPR 시스템을 암호화하는 핵산 서열은 임의적으로 5'에서 3'으로, (a) sgRNA가 HBG2 프로모터를 표적화하는 sgRNA-암호화 서열과 작동가능하게 연결된 U6 프로모터를 포함하는 sgRNA 유전자, 및 (b) miR 부위, β-글로빈 3'UTR 서열, 및 폴리아데닐화 신호와 작동가능하게 연결된 spCas9 코딩 서열과 작동가능하게 연결된 EF1α 프로모터를 포함하는 CRISPR 효소 유전자를 포함한다. 다양한 구현예에서, CRISPR 시스템은 HBG 프로모터의 BCL11A 결합 부위를 표적화하고 γ-글로빈 활성화 또는 재활성화를 야기할 수 있다. 본원에 개시된 바와 같이, CRISPR 시스템은 AAVS1 CRISPR에 의한 공여자 벡터의 절단이 비-통합된 공여자 벡터 핵산의 분해를 초래한다는 점에서 자기-불활성화될 수 있다. 다양한 구현예에서, miR 서열은 HDAd35 공여자 벡터 생산 동안 생산자 세포에서 Cas9 발현을 억제하는 서열일 수 있다(예를 들어, Saydaminova 등, *Mol. Ther. Meth. Clin. Dev.* 1: 14057, 2015; Li 등, *Mol. Ther. Meth. Clin. Dev.* 9: 390-401, 2018 참조).

[0376] 다양한 구현예에서, 본 개시내용의 Ad35 시스템은 Ad35 지지 벡터를 추가로 포함하며, 여기서 지지 벡터는 임의적으로 5'에서 3'으로, sgAAVS1-rm 코딩 서열에 작동가능하게 연결된 U6 프로모터를 포함한다.

[0377] 다양한 구현예에서 Ad35 페이로드는 Ad35 공여자 벡터 게놈에 존재한다. 다양한 구현예에서 Ad35 공여자 벡터 게놈에 존재하는 Ad35 페이로드는 Ad35 ITR에 의해 플랭킹된다. 다양한 구현예에서, Ad35 공여자 벡터 게놈은 Ad35 공여자 벡터에 존재한다. 다양한 구현예에서, 공여자 벡터는 Ad35 ++ 벡터이다.

[0378] 다양한 구현예에서 지지 게놈은 Ad35 ITR을 포함한다. 다양한 구현예에서, 지지 게놈은 Ad35 벡터에 존재한다. 다양한 구현예에서, 지지 벡터는 Ad35 ++ 벡터이다.

[0379] 다양한 구현예에서, Ad35 공여자 벡터는 헬퍼 의존적 공여자 벡터(HDAd35)이다. 이러한 특정 구현예에서, 본 개시내용의 시스템은 HDAd35 공여자 벡터 또는 게놈, 및 Ad35 헬퍼 벡터 또는 게놈을 포함할 수 있고, 다양한 구현예에서 Ad35 지지 벡터를 추가로 포함할 수 있다.

[0380] 특정 예시적인 구현예는 도 165에 예시되어 있다.

[0381] (iii) 다양한 구현예에서, Ad35 페이로드는 SB100x에 의한 전위를 위해 트랜스포사제 도립 반복부에 의해 플랭킹된 통합 요소를 포함하고, 트랜스포사제 도립 반복부는 FLPe와 같은 FLP 재조합효소에 의한 재조합을 위해 frt 지시 반복부에 의해 플랭킹된다. 다양한 구현예에서, 통합 요소는 임의적으로 5'에서 3'으로, (a) β-글로빈 미니-LCR, (b) γ-글로빈 코딩 서열이 3'UTR(예를 들어, γ-글로빈 3'UTR)과 작동가능하게 연결된 레서스 γ-글로빈 코딩 서열과 작동가능하게 연결된 β-글로빈 프로모터를 포함하며, 여기서, β-글로빈 미니-LCR이 또한 γ-글로빈 코딩 서열과 작동가능하게 연결되는 유전자 (c) cHS4 절연체 서열, 및 (d) 폴리아데닐화 신호와 작동가능하게 연결된 $MGMT^{P140K}$ 코딩 서열과 작동가능하게 연결된 PGK 프로모터를 포함하는 유전자를 포함하며, 임의

적으로 여기서 (a)-(d) 중 임의의 것은 Ad35 페이로드의 2 개의 가닥 중 어느 한 가닥 상에서 5'에서 3' 방향으로 암호화될 수 있다.

[0382] 다양한 구현예에서, Ad35 페이로드는 통합 요소의 외부 및 재조합효소 부위의 외부에, CRISPR 시스템을 암호화하는 핵산 서열을 추가로 포함한다. 특정한 특정 구현예에서, CRISPR 시스템을 암호화하는 핵산 서열은 임의적으로 5'에서 3'으로, (a) gRNA가 HBG 프로모터를 포함하는 gRNA-암호화 서열과 작동가능하게 연결된 U6 프로모터를 포함하는 gRNA 유전자, 및 (b) 3'UTR/miR 서열 및 폴리아데닐화 신호와 작동가능하게 연결된 CRISPR/Cas9 코딩 서열과 작동가능하게 연결된 EF1α 프로모터를 포함하는 CRISPR 효소 유전자를 포함한다. 다양한 구현예에서, CRISPR 시스템은 HBG 프로모터의 BCL11A 결합 부위를 표적화하여, γ-글로빈 활성화 또는 재활성화를 초래할 수 있다. 본원에 개시된 바와 같이, CRISPR 시스템은 전위에 의한 공여자 벡터의 절단이 비-통합된 공여자 벡터 핵산의 분해를 초래한다는 점에서 자기-불활성화될 수 있다. 다양한 구현예에서, miR 서열은 HDAd35 공여자 벡터 생산 동안 생산자 세포에서 Cas9 발현을 억제하는 서열일 수 있다(예를 들어, Saydaminova 등, *Mol. Ther. Meth. Clin. Dev.* 1: 14057, 2015; Li 등, *Mol. Ther. Meth. Clin. Dev.* 9: 390-401, 2018 참조).

[0383] 다양한 구현예에서, 본 개시내용의 Ad35 시스템은 Ad35 지지 벡터를 추가로 포함하며, 여기서 지지 벡터는 임의적으로 5'에서 3'으로, (a) FLPe 재조합효소 코딩 서열과 작동가능하게 연결된 EF1α 프로모터를 포함하는 재조합 유전자, 및 (b) SB100x 트랜스포사제 코딩 서열과 작동가능하게 연결된 PGK 프로모터를 포함하는 트랜스포사제 유전자를 포함한다.

[0384] 다양한 구현예에서 Ad35 페이로드는 Ad35 공여자 벡터 게놈에 존재한다. 다양한 구현예에서 Ad35 공여자 벡터 게놈에 존재하는 Ad35 페이로드는 A3d5 ITR에 의해 플랭킹된다. 다양한 구현예에서, Ad35 공여자 벡터 게놈은 Ad35 공여자 벡터에 존재한다. 다양한 구현예에서, 공여자 벡터는 Ad35 ++ 벡터이다.

[0385] 다양한 구현예에서 지지 게놈은 Ad35 ITR을 포함한다. 다양한 구현예에서, 지지 게놈은 Ad35 벡터에 존재한다. 다양한 구현예에서, 지지 벡터는 Ad35 ++ 벡터이다.

[0386] 다양한 구현예에서, Ad35 공여자 벡터는 헬퍼 의존적 공여자 벡터(HDAd35)이다. 이러한 특정 구현예에서, 본 개시내용의 시스템은 HDAd35 공여자 벡터 또는 게놈, 및 Ad35 헬퍼 벡터 또는 게놈을 포함할 수 있고, 다양한 구현예에서 Ad35 지지 벡터를 추가로 포함할 수 있다.

[0387] 특정 예시적인 구현예는 도 166에 예시되어 있다.

[0388] (iv) 다양한 구현예에서, Ad35 페이로드는 SB100x에 의한 전위를 위해 트랜스포사제 도립 반복부에 의해 플랭킹된 통합 요소를 포함하고, 트랜스포사제 도립 반복부는 FLPe와 같은 FLP 재조합효소에 의한 재조합을 위해 frt 지시 반복부에 의해 플랭킹된다. 다양한 구현예에서, 통합 요소는 임의적으로 5'에서 3'으로, (a) β-글로빈 미니-LCR, (b) γ-글로빈 코딩 서열이 3'UTR(예를 들어, γ-글로빈 3'UTR)과 작동가능하게 연결된 인간 γ-글로빈 코딩 서열과 작동가능하게 연결된 β-글로빈 프로모터를 포함하며, 여기서 β-글로빈 미니-LCR이 또한 γ-글로빈 코딩 서열과 작동가능하게 연결되는 유전자 (c) cHS4 절연체 서열, 및 (d) $MGMT^{P140K}$ 코딩 서열, 2A 자기-절단 펩티드, GFP 형광 마커 코딩 서열, 및 폴리아데닐화 신호와 작동가능하게 연결된 PGK 프로모터와 같은 프로모터를 포함하는 유전자를 포함하며, 임의적으로 여기서 (a)-(d) 중 임의의 것은 Ad35 페이로드의 2 개의 가닥 중 어느 한 가닥 상에서 5'에서 3' 방향으로 암호화될 수 있다.

[0389] 다양한 구현예에서, Ad35 페이로드는 통합 요소의 외부 및 재조합효소 부위의 외부에, 염기 편집 시스템을 암호화하는 핵산 서열을 추가로 포함한다. 특정한 특정 구현예에서, 염기 편집 시스템을 암호화하는 핵산 서열은 임의적으로 5'에서 3'으로, (a) 제1 gRNA가 bcl11a 인핸서를 표적화하는 제1 gRNA-암호화 서열과 작동가능하게 연결된 제1 U6 프로모터를 포함하는 제1 gRNA 유전자, (b) 제2 gRNA가 HBG 프로모터를 표적화하는 제2 gRNA-암호화 서열과 작동가능하게 연결된 제2 U6 프로모터를 포함하는 제2 gRNA 유전자, 및 (c) 3'UTR/miR 서열 및 폴리아데닐화 신호와 작동가능하게 연결된 염기 편집 효소 코딩 서열과 작동가능하게 연결된 EF1α 프로모터와 같은 프로모터를 포함하는 염기 편집 효소 유전자를 포함한다. 다양한 구현예에서, 염기 편집 시스템은 HBG 프로모터의 적혈구 bcl11a 인핸서 및 BCL11A 결합 부위를 표적화하며, 각각은 γ-글로빈 활성화 또는 재활성화를 야기하는 데 기여한다. 본원에 개시된 바와 같이, 염기 편집 시스템은 전위에 의한 공여자 벡터의 절단이 비-통합된 공여자 벡터 핵산의 분해를 초래한다는 점에서 자기-불활성화될 수 있다. 다양한 구현예에서, miR 서열은 HDAd35 공여자 벡터 생산 동안 생산자 세포에서 Cas9 발현을 억제하는 서열일 수 있다(예를 들어, Saydaminova 등, *Mol. Ther. Meth. Clin. Dev.* 1: 14057, 2015; Li 등, *Mol. Ther. Meth. Clin. Dev.* 9: 390-401, 2018 참조).

[0390] 다양한 구현예에서, 본 개시내용의 Ad35 시스템은 Ad35 지지 벡터를 추가로 포함하며, 여기서 지지 벡터는 임의적으로 5'에서 3'으로, (a) FLPe 재조합효소 코딩 서열과 작동가능하게 연결된 EF1α 프로모터를 포함하는 재조합 유전자, 및 (b) SB100x 트랜스포사제 코딩 서열과 작동가능하게 연결된 PGK 프로모터를 포함하는 트랜스포사제 유전자를 포함한다.

[0391] 다양한 구현예에서 Ad35 페이로드는 Ad35 공여자 벡터 게놈에 존재한다. 다양한 구현예에서 Ad35 공여자 벡터 게놈에 존재하는 Ad35 페이로드는 Ad35 ITR에 의해 플랭킹된다. 다양한 구현예에서, Ad35 공여자 벡터 게놈은 Ad35 공여자 벡터에 존재한다. 다양한 구현예에서, 공여자 벡터는 Ad35 ++ 벡터이다.

[0392] 다양한 구현예에서 지지 게놈은 Ad35 ITR을 포함한다. 다양한 구현예에서, 지지 게놈은 Ad35 벡터에 존재한다. 다양한 구현예에서, 지지 벡터는 Ad35 ++ 벡터이다.

[0393] 다양한 구현예에서, Ad35 공여자 벡터는 헬퍼 의존적 공여자 벡터(HDAd35)이다. 이러한 특정 구현예에서, 본 개시내용의 시스템은 HDAd35 공여자 벡터 또는 게놈, 및 Ad35 헬퍼 벡터 또는 게놈을 포함할 수 있고, 다양한 구현예에서 Ad35 지지 벡터를 추가로 포함할 수 있다.

[0394] (v) 다양한 구현예에서, Ad35 페이로드는 SB100x에 의한 전위를 위해 트랜스포사제 도립 반복부에 의해 플랭킹된 통합 요소를 포함하고, 트랜스포사제 도립 반복부는 FLPe와 같은 FLP 재조합효소에 의한 재조합을 위해 frt 지시 반복부에 의해 플랭킹된다. 다양한 구현예에서, 통합 요소는 임의적으로 5'에서 3'으로, (a) β-글로빈 미니-LCR, (b) γ-글로빈 코딩 서열이 3'UTR(예를 들어, γ-글로빈 3'UTR)과 작동가능하게 연결된 레서스 γ-글로빈 코딩 서열과 작동가능하게 연결된 β-글로빈 프로모터를 포함하며, 여기서 β-글로빈 미니-LCR이 또한 γ-글로빈 코딩 서열과 작동가능하게 연결되는 유전자 (c) cHS4 절연체 서열, 및 (d) 폴리아데닐화 신호와 작동가능하게 연결된 $MGMT^{P140K}$ 코딩 서열과 작동가능하게 연결된 PGK 프로모터를 포함하는 유전자를 포함하며, 임의적으로 여기서 (a)-(d) 중 임의의 것은 Ad35 페이로드의 2 개의 가닥 중 어느 한 가닥 상에서 5'에서 3' 방향으로 암호화될 수 있다.

[0395] 다양한 구현예에서, Ad35 페이로드는 통합 요소의 외부 및 재조합효소 부위의 외부에, 염기 편집 시스템을 암호화하는 핵산 서열을 추가로 포함한다. 특정한 특정 구현예에서, 염기 편집 시스템을 암호화하는 핵산 서열은 임의적으로 5'에서 3'으로, (a) gRNA가 HBG 프로모터를 표적화하는 gRNA-암호화 서열과 작동가능하게 연결된 U6 프로모터를 포함하는 gRNA 유전자, 및 (b) 3'UTR/miR 서열 및 폴리아데닐화 신호와 작동가능하게 연결된 염기 편집 효소 코딩 서열과 작동가능하게 연결된 EF1α 프로모터를 포함하는 염기 편집 효소 유전자를 포함한다. 다양한 구현예에서, 염기 편집 시스템은 HBG 프로모터의 BCL11A 결합 부위를 표적화하며, 이는 γ-글로빈 활성화 또는 재활성화를 초래할 수 있다. 본원에 개시된 바와 같이, 염기 편집 시스템은 전위에 의한 공여자 벡터의 절단이 비-통합된 공여자 벡터 핵산의 분해를 초래한다는 점에서 자기-불활성화될 수 있다. 다양한 구현예에서, miR 서열은 HDAd35 공여자 벡터 생산 동안 생산자 세포에서 Cas9 발현을 억제하는 서열일 수 있다(예를 들어, Saydaminova 등, *Mol. Ther. Meth. Clin. Dev.* 1: 14057, 2015; Li 등, *Mol. Ther. Meth. Clin. Dev.* 9: 390-401, 2018 참조).

[0396] 다양한 구현예에서, 본 개시내용의 Ad35 시스템은 Ad35 지지 벡터를 추가로 포함하며, 여기서 지지 벡터는 임의적으로 5'에서 3'으로, (a) FLPe 재조합효소 코딩 서열과 작동가능하게 연결된 EF1α 프로모터를 포함하는 재조합 유전자, 및 (b) SB100x 트랜스포사제 코딩 서열과 작동가능하게 연결된 PGK 프로모터를 포함하는 트랜스포사제 유전자를 포함한다.

[0397] 다양한 구현예에서 Ad35 페이로드는 Ad35 공여자 벡터 게놈에 존재한다. 다양한 구현예에서 Ad35 공여자 벡터 게놈에 존재하는 Ad35 페이로드는 Ad35 ITR에 의해 플랭킹된다. 다양한 구현예에서, Ad35 공여자 벡터 게놈은 Ad35 공여자 벡터에 존재한다. 다양한 구현예에서, 공여자 벡터는 Ad35++ 벡터이다.

[0398] 다양한 구현예에서 지지 게놈은 Ad35 ITR을 포함한다. 다양한 구현예에서, 지지 게놈은 Ad35 벡터에 존재한다. 다양한 구현예에서, 지지 벡터는 Ad35++ 벡터이다.

[0399] 다양한 구현예에서, Ad35 공여자 벡터는 헬퍼 의존적 공여자 벡터(HDAd35)이다. 이러한 특정 구현예에서, 본 개시내용의 시스템은 HDAd35 공여자 벡터 또는 게놈, 및 Ad35 헬퍼 벡터 또는 게놈을 포함할 수 있고, 다양한 구현예에서 Ad35 지지 벡터를 추가로 포함할 수 있다.

[0400] **I(C)(ii). 페이로드 조절 서열**

[0401] **I(C)(ii)(a). 프로모터 조절 서열**

[0402] 프로모터는 일반적으로 전사를 개시하기 전에 RNA 폴리머라제가 결합하는 관련 코딩 서열의 상류(5')에 있는 비-코딩 게놈 DNA 서열일 수 있다. 이 결합은 전사가 특이적 전사 개시 부위에서 개시하도록 RNA 폴리머라제를 정렬한다. 프로모터의 뉴클레오티드 서열은 효소의 속성 및 이에 부착되는 다른 관련 단백질 인자 및 RNA 합성률을 결정한다. RNA는 RNA 서열을 암호화된 폴리펩티드의 아미노산 서열 내로 번역하기 위한 주형으로서 역할을 하는 메신저 RNA(mRNA)를 생성하도록 처리된다. 5' 비-번역된 리더 서열은 mRNA의 개시 및 번역에서 역할을 할 수 있는 코딩 서열의 상류에 있는 mRNA 영역이다. 3' 전사 종결/폴리아데닐화 신호는 RNA 합성의 종결 및 3' 단부에 폴리아데닐레이트 뉴클레오티드의 첨가를 야기하도록 식물 세포에서 기능하는 코딩 영역의 하류에 있는 비-번역된 영역이다.

[0403] 프로모터는 일반적인 프로모터, 조직-특이적 프로모터, 세포-특이적 프로모터, 및/또는 세포질에 특이적인 프로모터를 포함할 수 있다. 프로모터는 강한 프로모터, 약한 프로모터, 구성적 발현 프로모터, 및/또는 유도성(조건부) 프로모터를 포함할 수 있다. 유도성 프로모터는 특정 조건, 신호, 또는 세포 사건에 반응하여 발현을 지시하거나 또는 제어한다. 예를 들어, 프로모터는 프로모터로부터 전사를 실행하기 위해 특정 리간드, 소분자, 전사 인자, 호르몬, 또는 호르몬 단백질을 필요로 하는 유도성 프로모터일 수 있다. 프로모터의 특정 예는 AFP(α-태아단백질) 프로모터, 아밀라제 1C 프로모터, 아쿠아포린-5(AP5) 프로모터, α1 -항트립신 프로모터, β-act 프로모터, β-글로빈 프로모터, β-Kin 프로모터, B29 프로모터, CCKAR 프로모터, CD14 프로모터, CD43 프로모터, CD45 프로모터, CD68 프로모터, CEA 프로모터, c-erbB2 프로모터, COX-2 프로모터, CXCR4 프로모터, 데스민 프로모터, E2F-1 프로모터, 인간 신장 인자 1α 프로모터(EF1α), CMV(사이토메갈로바이러스 바이러스) 프로모터, minCMV 프로모터, SV40(시미안 바이러스 40) 급초기 프로모터, EGR1 프로모터, eIF4A1 프로모터, 엘라스타제-1 프로모터, 엔도글린 프로모터, FerH 프로모터, FerL 프로모터, 피브로넥틴 프로모터, Flt-1 프로모터, GAPDH 프로모터, GFAP 프로모터, GPIIb 프로모터, GRP78 프로모터, GRP94 프로모터, HE4 프로모터, hGR1/1 프로모터, hNIS 프로모터, Hsp68 프로모터, Hsp68 최소 프로모터(proHSP68), HSP70 프로모터, HSV-1 바이러스 TK 유전자 프로모터, hTERT 프로모터, ICAM-2 프로모터, 칼리크레인 프로모터, LP 프로모터, 주요 후기 프로모터(MLP), Mb 프로모터, Rho 프로모터, MT(메탈로티오네인) 프로모터, MUC1 프로모터, NphsI 프로모터, OG-2 프로모터, PGK(포스포 글리세레이트 키나제) 프로모터, PGK-1 프로모터, 폴리머라제 III(Pol III) 프로모터, PSA 프로모터, ROSA 프로모터, SP-B 프로모터, 서비빈 프로모터, SYN1 프로모터, SYT8 유전자 프로모터, TRP1 프로모터, Tyr 프로모터, 유비퀴틴 B 프로모터, WASP 프로모터, 및 라우스 육종 바이러스(RSV) 긴-말단 반복부(LTR) 프로모터를 포함한다

[0404] 프로모터는 천연 프로모터 또는 복합 프로모터로서 수득될 수 있다. 천연 프로모터, 또는 최소 프로모터는 주어진 유전자의 5' 영역으로부터의 뉴클레오티드 서열을 포함하는 프로모터를 지칭한다. 천연 프로모터는 코어 프로모터 및 이의 천연 5'UTR을 포함한다. 특정 구현예에서, 5'UTR은 인트론을 포함한다. 복합 프로모터는 상이한 기원의 프로모터 요소를 조합하거나 또는 원위 인핸서를 동일하거나 또는 상이한 기원의 최소 프로모터와 조합함으로써 유도된 프로모터를 지칭한다.

[0405] 특정 구현예에서, SV40 프로모터는 서열번호: 80에 제시된 서열을 포함한다. 특정 구현예에서, dESV40 프로모터(인핸서 영역의 결실이 있는 SV40 프로모터)는 서열번호: 81에 제시된 서열을 포함한다. 특정 구현예에서, 인간 텔로머라제 촉매적 서브유닛(hTERT) 프로모터는 서열번호: 82에 제시된 서열을 포함한다. 특정 구현예에서, 슈미츠-루핀(Schmidt-Ruppin) A 균주로부터 유래된 RSV 프로모터는 서열번호: 83에 제시된 서열을 포함한다. 특정 구현예에서, hNIS 프로모터는 서열번호: 84에 제시된 서열을 포함한다. 특정 구현예에서, 인간 글루코코르티코이드 수용체 1A(hGR 1/Ap/e) 프로모터는 서열번호: 85에 제시된 서열을 포함한다.

[0406] 특정 구현예에서, 프로모터는 야생형 프로모터 서열 및 야생형 프로모터에 비해 특정 위치에서 임의적인 변화(삽입, 점 돌연변이 또는 결실 포함)가 있는 서열을 포함한다. 특정 구현예에서, 프로모터는 20 개 뉴클레오티드 스트레치 당 1 개 변화, 20 개 뉴클레오티드 스트레치 당 2 개 변화, 20 개 뉴클레오티드 스트레치 당 3 개 변화, 20 개 뉴클레오티드 스트레치 당 4 개 변화, 또는 20 개 뉴클레오티드 스트레치 당 5 개 변화를 가짐으로써 자연 발생 프로모터로부터 달라진다. 특정 구현예에서, 천연 서열은 1, 2, 3, 4, 5, 6, 7, 8, 9, 또는 10 개의 염기에서 변경될 것이다. 프로모터는 다른 바이러스 서열이 있거나 없이, LTR 서열의 50 개 뉴클레오티드 내지 100, 200, 250 또는 350 개 뉴클레오티드를 포함하는 길이로 달라질 수 있다.

[0407] 일부 프로모터는 조직 또는 세포에 특이적이고 일부 프로모터는 조직 또는 세포에 비-특이적이다. 포유류 세포의 각 유전자는 고유한 프로모터가 있고 일부 프로모터는 특정 세포 유형에서만 활성화될 수 있다. 비-특이적

프로모터, 또는 유비퀴터스 프로모터는 광범위한 세포, 조직 및 세포 주기에서 프로모터 서열에 작동가능하게 연결된 유전자 또는 뉴클레오티드 서열의 전사 개시를 보조한다. 특정 구현예에서, 프로모터는 비-특이적 프로모터이다. 특정 구현예에서, 비-특이적 프로모터는 CMV 프로모터, RSV 프로모터, SV40 프로모터, 포유류 신장 인자 1α(EF1α) 프로모터, β-act 프로모터, EGR1 프로모터, eIF4A1 프로모터, FerH 프로모터, FerL 프로모터, GAPDH 프로모터, GRP78 프로모터, GRP94 프로모터, HSP70 프로모터, β-Kin 프로모터, PGK-1 프로모터, ROSA 프로모터, 및/또는 유비퀴틴 B 프로모터를 포함한다.

[0408] 특이적 프로모터는 프로모터 서열에 작동가능하게 연결된 뉴클레오티드 서열의 세포 특이적 발현을 보조한다. 특정 구현예에서, 특이적 프로모터는 B 세포, 단핵구 세포, 백혈구, 대식세포, 췌장 선포 세포, 내피 세포, 성상세포, 및/또는 임의의 다른 세포 유형 또는 세포 주기에서 활성이다. 특정 구현예에서, 프로모터는 특이적 프로모터이다. 특정 구현예에서, SYT8 유전자 프로모터는 인간 섬에서 유전자 발현을 조절한다(Xu, 등, Nat Struct Mol Biol., 2011, 18: 372-378). 특정 구현예에서, 칼리크레인 프로모터 췌관 세포 특이적 타액선에서 유전자 발현을 조절한다. 특정 구현예에서, 아밀라제 1C 프로모터는 선포 세포에서 유전자 발현을 조절한다. 특정 구현예에서, 아쿠아포린-5(AP5) 프로모터는 선포 세포에서 유전자 발현을 조절한다(Zheng and Baum, *Methods Mol Biol.*, 434: 205-219, 2008). 특정 구현예에서, B29 프로모터는 B 세포에서 유전자 발현을 조절한다. 특정 구현예에서, CD14 프로모터는 단핵구 세포에서 유전자 발현을 조절한다. 특정 구현예에서, CD43 프로모터는 백혈구 및 혈소판에서 유전자 발현을 조절한다. 특정 구현예에서, CD45 프로모터는 조혈세포에서 유전자 발현을 조절한다. 특정 구현예에서, CD68 프로모터는 대식세포에서 유전자 발현을 조절한다. 특정 구현예에서, 데스민 프로모터는 근육 세포에서 유전자 발현을 조절한다. 특정 구현예에서, 엘라스타제-1 프로모터는 췌장 선포 세포에서 유전자 발현을 조절한다. 특정 구현예에서, 엔도글린 프로모터는 내피 세포에서 유전자 발현을 조절한다. 특정 구현예에서, 피브로넥틴 프로모터는 분화 세포 또는 치유 조직에서 유전자 발현을 조절한다. 특정 구현예에서, Flt-1 프로모터는 내피 세포에서 유전자 발현을 조절한다. 특정 구현예에서, GFAP 프로모터는 성상세포에서 유전자 발현을 조절한다. 특정 구현예에서, GPIIb 프로모터는 거핵구에서 유전자 발현을 조절한다. 특정 구현예에서, ICAM-2 프로모터는 내피 세포에서 유전자 발현을 조절한다. 특정 구현예에서, Mb 프로모터는 근육에서 유전자 발현을 조절한다. 특정 구현예에서, NphsI 프로모터는 족세포에서 유전자 발현을 조절한다. 특정 구현예에서, OG-2 프로모터는 골모세포, 치아모세포에서 유전자 발현을 조절한다. 특정 구현예에서, SP-B 프로모터는 폐 세포에서 유전자 발현을 조절한다. 특정 구현예에서, SYN1 프로모터는 뉴런에서 유전자 발현을 조절한다. 특정 구현예에서, WASP 프로모터는 조혈 세포에서 유전자 발현을 조절한다.

[0409] 특정 구현예에서, 프로모터는 종양-특이적 프로모터이다. 특정 구현예에서, AFP 프로모터는 간세포 암종에서 유전자 발현을 조절한다. 특정 구현예에서, CCKAR 프로모터는 췌장암에서 유전자 발현을 조절한다. 특정 구현예에서, CEA 프로모터는 상피암에서 유전자 발현을 조절한다. 특정 구현예에서, c-erbB2 프로모터는 유방암 및 췌장암에서 유전자 발현을 조절한다. 특정 구현예에서, COX-2 프로모터는 종양에서 유전자 발현을 조절한다. 특정 구현예에서, CXCR4 프로모터는 종양에서 유전자 발현을 조절한다. 특정 구현예에서, E2F-1 프로모터는 종양에서 유전자 발현을 조절한다. 특정 구현예에서, HE4 프로모터는 종양에서 유전자 발현을 조절한다. 특정 구현예에서, LP 프로모터는 종양에서 유전자 발현을 조절한다. 특정 구현예에서, MUC1 프로모터는 암종 세포에서 유전자 발현을 조절한다. 특정 구현예에서, PSA 프로모터는 전립선 및 전립선암에서 유전자 발현을 조절한다. 특정 구현예에서, 서비븐 프로모터는 종양에서 유전자 발현을 조절한다. 특정 구현예에서, TRP1 프로모터는 멜라닌세포 및 흑색종에서 유전자 발현을 조절한다. 특정 구현예에서, Tyr 프로모터는 멜라닌세포 및 흑색종에서 유전자 발현을 조절한다.

[0410] **I(C)(ii)(b). LCR 조절 서열**

[0411] 유전자좌 제어 영역은 염색질 부위에서 조직-특이적 및 카피 수-의존적 방식으로 연결된 유전자의 발현을 생리학적 수준으로 향상시키는 능력에 의해 작동적으로 정의된다. Li 등, Blood, 2002, 100(9): 3077-3086.

[0412] β-글로빈 LCR은 적어도 여러 측면에서 적어도 일부 LCR의 예시이다. 예를 들어, 많은 다른 LCR과 마찬가지로, β-글로빈 LCR은 작동가능하게 연결된 유전자 또는 이식유전자의 발현을 향상시키고(예를 들어, 전사 증가, 번역 증가, 및/또는 세포 또는 조직 특이성 증가) LCR의 발현 효과를 매개하기 위해 당업자에 의해 이해되는 DNAse 과민성(HS) 영역을 포함한다. 또한, 많은 다른 LCR과 마찬가지로, β-글로빈 LCR은 예를 들어, β-글로빈 LCR HS 영역(HS1-HS5)의 전부를 포함하거나 또는 β-글로빈 LCR HS 영역(예를 들어, HS1-HS4)의 하위세트를 포함하는 β-글로빈 LCR 서열을 포함하는 핵산에서 활용될 수 있다는 점에서 전체적으로 또는 부분적으로 활용될 수 있다.

[0413] 염색체 11 상의 호모 사피엔스(*Homo sapiens*) β-글로빈 영역에 대한 예시적인 핵산 서열은 GenBank 수탁 번호 NG_000007로 제공된다. β-글로빈 긴 LCR은 일부 경우에, 유전자좌의 첫번째(배아) 글로빈 유전자에 대해 6 내지 22 kb 5'에 위치한 서열이거나 또는 이를 포함할 수 있다. β-글로빈 긴 LCR은 5 개의 DNAse I 과민성 부위인 5'HS 1 내지 5를 포함할 수 있다. Li 등, Blood, 2002, 100(9): 3077-3086. NG_000007은 유전자좌 제어 영역 내에서 DNAse I 과민성 부위 HS1, HS2, HS3, 및 HS4를 설명하는 제한 부위의 위치(예를 들어, HS2의 SnaBI 및 BstXI 제한 부위, HS3의 HindIII 및 BamHI 제한 부위, 및 HS4의 BamHI 및 BanII 제한 부위)를 제공하고, 특히 과민성 부위 위치와 관련하여 그 전체가 본원에 참조로 포함된다. HS1의 서열 및 위치는 예를 들어, Pasceri 등, *Ann NY Acad. Sci.* 850:377-381, 1998; Pasceri 등, *Blood*. 92:653-663, 1998; 및 Milot 등, *Cell*. 87:105-114, 1996에 기재되어 있다. 특정 구현예에서, HS2 영역은 유전자좌 제어 영역의 위치 16,671에서 17,058까지 확장된다. HS2의 SnaBI 및 BstXI 제한 부위는 각각 위치 17,093 및 16,240에 위치한다. HS3 영역은 유전자좌 제어 영역의 위치 12,459에서 13,097까지 확장된다. HS3의 BamHI 및 HindIII 제한 부위는 각각 위치 12,065 및 13,360에 위치한다. HS4 영역은 유전자좌 제어 영역의 위치 9,048에서 9,713까지 확장된다. HS4의 BamHI 및 BanII 제한 부위는 각각 위치 8,496 및 9,576에 위치한다.

[0414] 본원에 개시된 특정 구현예는 β-글로빈 LCR의 미니-부분을 활용한다. 미니-부분은 LCR이 β-글로빈 LCR의 5 개 절편을 모두 포함하지 않는 한, HS1, HS2, HS3, HS4, 및/또는 HS5와 같은 5 개 미만의 HS 영역을 모두 포함한다. 본 개시내용의 실시예 1에서 활용되는 4.3 kb HS1-HS4 LCR은 미니-LCR의 일 예를 제공한다. 다른 미니-LCR은 예를 들어, HS1, HS2, 및 HS3; HS2, HS3, 및 HS4; HS3, HS4, 및 HS5; HS1, HS3, 및 HS5; HS1, HS2, 및 HS5; 및 HS1, HS4, 및 HS5를 포함할 수 있다. 미니-LCR의 추가적인 예에 대해, Sadelain 등, *Proc. Nat. Acad. Sci*. (USA) 92: 6728-6732, 1995; 및 Lebouich 등, *EMBO J.* 13: 3065-3076, 1994를 참조한다. 특정 구현예는 β-글로빈 프로모터와 조합하여 미니-β-글로빈 LCR을 활용할 수 있다. 특정 구현예에서, 이 조합은 5.9 kb LCR-프로모터 조합을 산출한다. LCR과 관련하여, "미니" 및 "마이크로"는 본원에서 상호교환가능하게 사용된다.

[0415] 본원에 개시된 특정 구현예는 유전자좌 제어 영역(LCR)의 긴 부분을 활용한다. 긴 β-글로빈 LCR은 HS1, HS2, HS3, HS4, 및 HS5를 포함할 수 있다. 특정 구현예에서, 긴 LCR은 β-글로빈 LCR의 HS1, HS2, HS3, HS4, 및 HS5를 포함하는 21.5 kb 서열을 포함한다. 긴 β-글로빈 LCR은 β-글로빈 프로모터와 커플링되어 높은 단백질 발현 수준을 구동할 수 있다.

[0416] 특정 구현예는 GRCh38에 열거된 바와 같이 인간 염색체 11(서열번호: 185)의 긴 β-글로빈 LCR 위치 5292319-5270789(21,531 bp)를 포함할 수 있다. 다양한 구현예에서, 긴 LCR은 18 kb, 18.5 kb, 19 kb, 19.5 kb, 20 kb, 20.5 kb, 21 kb, 21.5 kb, 또는 21.531 kb보다 더 크거나 또는 동일한 총 길이를 가질 수 있다. 다양한 구현예에서, 긴 LCR은 서열번호: 185의 길이의 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%보다 더 크거나 또는 동일한 총 길이를 가질 수 있다. 다양한 구현예에서, 긴 LCR은 서열번호: 185의 적어도 18 kb, 18.5 kb, 19 kb, 19.5 kb, 20 kb, 20.5 kb, 21 kb, 또는 21.5 kb를 포함할 수 있다. 본원에 제공된 다양한 구현예 중 임의의 것에서, 긴 LCR은 서열번호: 185의 상응하는 인접한 부분과 적어도 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 핵산이거나 또는 이를 포함할 수 있다. 본원에 제공된 다양한 구현예 중 임의의 것에서, 긴 LCR은 HS1, HS2, HS3, HS4, 및 HS5를 포함할 수 있다.

[0417] 다양한 구현예에서, Ad35 벡터 시스템은 예를 들어, 인간 염색체 11의 위치 5228631-5227023(1609 bp) 또는 β-글로빈 프로모터로서 GRCh38에 열거된 바와 같은 5228631-5227018(1614 bp)(서열번호: 186)을 포함하는 전위성 이식유전자 삽입물을 포함할 수 있다. 다양한 구현예에서, β-글로빈 프로모터는 예를 들어, 1.0 kb, 1.1. kb, 1.2 kb, 1.3 kb, 1.4 kb, 1.5 kb, 1.6 kb, 또는 1.609 kb보다 더 크거나 또는 동일한 총 길이를 가질 수 있다. 다양한 구현예에서, β-글로빈 프로모터는 서열번호: 186의 적어도 1.0 kb, 1.1. kb, 1.2 kb, 1.3 kb, 1.4 kb, 1.5 kb, 1.6 kb, 또는 1.609 kb를 포함할 수 있다. 다양한 구현예에서, 전위성 이식유전자 삽입물은 인간 염색체 11의 위치 5228631-5227023(1609 bp)을 포함할 수 있다. 다양한 구현예에서, β-글로빈 프로모터는 예를 들어, 엡실론(HBE1), G-감마(HBG2), A-감마(HBG1), 델타(HBD), 및 베타(HBB) 글로빈 유전자 및/또는 헤모글로빈 β 유전자좌에 존재하는 하나 이상의 유전자(11:5,225,463-5,227,070, 보체) 중 임의의 것을 포함하나 이에 제한되지 않는, β-글로빈 LCR에 의해 발현이 조절되는 유전자의 상류, 예를 들어, 이의 첫번째 코딩 뉴클레오티드의 바로 상류에 있는 핵산 서열의 100 bp, 200 bp, 300 bp, 400 bp, 500 bp, 1 kb, 1.5 kb, 2 kb, 2.5 kb, 3 kb, 4 kb, 또는 5 kb보다 더 크거나 또는 동일한 총 길이를 포함할 수 있다. 다양한 구현예에서, β-글로빈 프로모터는 예를 들어, 염색체 11 NC_000011.10 위치 5227021의 상류, 예를 들어, 바로

상류에 있는 핵산 서열의 100 bp, 200 bp, 300 bp, 400 bp, 500 bp, 1 kb, 1.5 kb, 2 kb, 2.5 kb, 3 kb, 4 kb, 또는 5 kb보다 더 크거나 또는 동일한 총 길이를 포함할 수 있다. 다양한 구현예에서, β-글로빈 프로모터는 서열번호: 186의 길이의 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%보다 더 크거나 또는 동일한 총 길이를 가질 수 있다. 본원에 제공된 다양한 구현예 중 임의의 것에서, β-글로빈 프로모터는 서열번호: 186의 상응하는 인접합 부분과 적어도 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 서열을 갖는 핵산이거나 또는 이를 포함할 수 있다.

[0418] 다양한 구현예에서, 긴 β-글로빈 LCR과 같은 β-글로빈 LCR은 적혈구에서 작동가능하게 연결된 코딩 서열의 발현을 야기한다. 다양한 구현예에서, 작동가능하게 연결된 코딩 서열은 또한 본원에 제시되거나 또는 달리 당업계에 알려진 바와 같은 β-글로빈 프로모터와 작동가능하게 연결된다.

[0419] 면역글로불린 중쇄 유전자좌 B 세포 LCR은 작동가능하게 연결된 코딩 서열의 발현을 향상시키는(예를 들어, 전사 증가, 번역 증가, 및/또는 세포 또는 조직 특이성 증가) 예시적인 LCR이다. 코딩 서열의 발현은 완전한 면역글로불린 중쇄 유전자좌 B 세포 LCR 서열을 포함하고/하거나 이의 발현-조절 단편을 포함하는 면역글로불린 중쇄 유전자좌 B 세포 LCR에 작동가능하게 연결될 때 향상될 수 있다. 면역글로불린 중쇄 유전자좌 B 세포 LCR은 면역글로불린 중쇄 유전자좌 B 세포 LCR의 발현-향상 효과의 적어도 일부를 매개하기 위해 당업자에 의해 이해되는 DNAse 과민성 부위(HS)를 포함한다. 면역글로불린 중쇄 유전자좌 B 세포 LCR은 인핸서-유전자좌 제어 영역(LCR)으로서 기능하는 면역글로불린 중쇄(IgH) 유전자좌의 3'Cα 영역에서 4 개의 DNase I-과민성 부위(HS1, HS2, HS3, 및 HS4)를 포함한다. 따라서, 면역글로불린 중쇄 유전자좌 B 세포 LCR은 HS1-HS4를 모두 포함하는 완전한 면역글로불린 중쇄 유전자좌 B 세포 LCR일 수 있거나, 또는 과민성 부위 HS1-HS4의 하위세트를 포함하는 이의 발현-조절 단편일 수 있다. 이러한 HS 부위는 IgH C 유전자의 10-30 kb에 대해 맵핑되고 일시적 형질감염 검정에서 림프계 세포-특이적 및 발달적으로 조절된 인핸서 요소를 유발할 수 있다. 이 핵산 서열은 버킷(Burkitt) 림프종 및 형질세포종 세포주에서 c-myc 유전자에 연결될 때 유사한 발현 패턴을 지시할 수 있는 것으로 관찰되었다. 버킷 림프종 및 형질세포종에서, B-세포 LCR에 의한 c-myc의 제어는 c-myc 유전자가 IgH 서열과 병치되게 하여, 비정상적인 c-myc 전사를 초래하는 특징적 염색체 전위로 인해 발생한다. B 세포 LCR의 추가적인 설명은 예를 들어, Madisen 등, *Mol Cell Biol.* 18(11):6281-92, 1998; Giannini 등, *J. Immunol.* 150:1772-1780, 1993; Madisen & Groudine, Genes *Dev.* 8:2212-2226, 1994; 및 Michaelson 등, *Nucleic Acids Res.* 23:975-981, 1995에서 찾을 수 있다.

[0420] 발현 작제물은 mRNA 전사체, 예를 들어, 절연체, 및/또는 polyA 꼬리의 안정성을 향상시키는 특징을 추가적으로 포함할 수 있다.

[0421] **I(C)(ii)(c). Micro RNA 부위 조절 서열**

[0422] 다양한 구현예에서, microRNA(또는 miRNA) 제어 시스템은 유전자의 발현이 microRNA 부위(예를 들어, microRNA가 상호작용할 수 있는 핵산 서열)의 존재에 의해 조절되는 방법 또는 조성물을 지칭할 수 있다. 다양한 구현예에서, 본 개시내용은 발현 산물의 발현이 상응하는 miRNA의 존재, 수준, 활성, 및/또는 이와의 접촉에 의해 제어되도록 발현 산물을 암호화하는 핵산 서열이 miRNA 표적 부위에 작동가능하게 연결된 페이로드를 포함하는 Ad35 공여자 벡터를 포함한다. 다양한 구현예에서, miRNA 부위는 miR423-5, miR423-5p, miR42-2, miR181c, miR125a, miR15a, miR187, 및/또는 miR218 중 임의의 것으로부터 선택된 miRNA에 대한 표적 부위이다. 의심의 여지를 피하기 위해 본 개시내용은 예를 들어, 본원에 기재된 바와 같이 miRNA 부위와 작동가능하게 연결된 핵산 서열이 예를 들어, 본원에 제공된 하나 이상의 발현 산물 중 임의의 것을 암호화하는 핵산 서열일 수 있음을 고려한다.

[0423] 특정 구현예에서, microRNA 제어 시스템은 유전자가 HSPC 예를 들어, 종양 침윤 HSPC와 같은 표적 세포에서 광범위하게 발현되도록 유전자의 발현을 조절하였다. 일부 구현예에서, 관심 단백질 또는 핵산(예를 들어, 항암제 예컨대 CAR, TCR, 항체, 및/또는 체크포인트 억제제, 예를 들어, 체크포인트 억제제인 αPD-L1 항체(예를 들어, αPD-L1ɣ1 항체))을 암호화하는 핵산(예를 들어, 치료 유전자)은 microRNA 부위, 복수의 동일한 microRNA 부위, 또는 복수의 별개의 microRNA 부위를 포함하거나, 이와 회합되거나, 또는 이와 작동가능하게 연결된다. 당업자는 관심 유전자를 암호화하는 서열을 갖는 핵산 또는 이의 부분과 microRNA 부위를 회합하는 수단 및 기술에 친숙할 것이지만, 특정 비-제한적인 예가 본원에 제공된다. 예를 들어, 관심 유전자(예를 들어, αPD-L1ɣ1 항체를 암호화하는 서열)는 관심 유전자의 발현이 종양-침윤 백혈구 세포가 아닌 세포에서 발현을 억제하지만, 종양-침윤 백혈구에서 발현을 억제하지 않는 하나 이상의 microRNA 부위의 존재에 의해 조절되도록 핵산

에 존재할 수 있다. 특정한 특정 예에서, 관심 유전자(예를 들어, αPD-L1ɣ1 항체를 암호화하는 서열)는 관심 유전자의 발현이 종양-침윤 백혈구 세포가 아닌 세포에서 발현을 억제하지만, 종양-침윤 백혈구에서 발현을 억제하지 않는 하나 이상의 miR423-5p microRNA 부위의 존재에 의해 조절되도록 핵산에 존재할 수 있다. 다양한 구현예에서, microRNA 제어 시스템은 하나 이상의 microRNA 부위, 예를 들어, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 개, 또는 그 이상의 microRNA 부위를 포함하거나, 또는 관심 단백질 또는 핵산의 발현이 이에 의해 조절되는 핵산을 포함할 수 있다. 다양한 구현예에서, microRNA 제어 시스템은 하나 이상의 miR423-5p microRNA 부위, 예를 들어, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 개, 또는 그 이상의 miR423-5p microRNA 부위를 포함하거나, 또는 관심 단백질 또는 핵산의 발현이 이에 의해 조절되는 핵산을 포함할 수 있다. 일부 특정 구현예에서, microRNA 제어 시스템은 αPD-L1ɣ1 항체를 암호화하고, 하나 이상의 miR423-5p microRNA 부위, 예를 들어, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 개, 또는 그 이상의 miR423-5p microRNA 부위, 예를 들어, miR423-5p microRNA 부위를 포함하거나, 또는 αPD-L1ɣ1 항체의 발현이 이에 의해 조절되는 핵산을 포함할 수 있다.

[0424] 다양한 구현예에서, microRNA 부위는 HDAd35 공여자 벡터 생산 동안 생산자 세포에서 작동가능하게 연결된 코딩 서열, 예를 들어, CRISPR 효소, 염기 편집 효소, 또는 gRNA를 암호화하는 코딩 서열의 발현을 억제하는 서열일 수 있다(예를 들어, Saydaminova 등, *Mol. Ther. Meth. Clin. Dev.* 1: 14057, 2015; Li 등, *Mol. Ther. Meth. Clin. Dev.* 9: 390-401, 2018 참조).

[0425] **I(C)(iii). 선택 서열**

[0426] 특정 구현예에서 벡터는 선택 카세트를 포함하는 선택 요소를 포함한다. 특정 구현예에서, 선택 카세트는 프로모터, 선택제에 내성을 추가하거나 또는 부여하는 cDNA, 및 이 독립적인 전사 요소의 전사를 중지시킬 수 있는 poly A 서열을 포함한다.

[0427] 선택 카세트는 (a) 항생제 또는 다른 독소에 내성을 부여하거나, (b) 영양요구성 결핍을 보완하거나, 또는 (c) 복합 배지로부터 이용가능하지 않은 중요한 영양소, 예를 들어, 바실리(Bacilli)에 대한 D-알라닌 라세마제를 암호화하는 유전자를 공급하는 하나 이상의 단백질을 암호화할 수 있다. 임의의 수의 선택 시스템을 사용하여 형질도입된 세포주를 회수할 수 있다. 특정 구현예에서, 양성 선택 카세트는 네오마이신, 하이그로마이신, 암피실린, 퓨로마이신, 플레오마이신, 제오마이신, 블라스티시딘, 비오마이신에 대한 내성 유전자를 포함한다. 특정 구현예에서, 양성 선택 카세트는 메토트렉세이트에 대한 내성을 제공하는 DHFR(디하이드로폴레이트 환원효소) 유전자, $O^6$BG/BCNU에 대한 내성을 담당하는 $MGMT^{P140K}$ 유전자, HAT 선택 배지에 존재하는 특이적 염기(아미노프테린, 하이포크산틴, 티미딘)의 형질전환을 담당하는 HPRT(하이포크산틴 포스포리보실 트랜스퍼라제) 유전자 및 일부 약물에 관한 해독을 위한 다른 유전자를 포함한다. 특정 구현예에서, 선택제는 네오마이신, 하이그로마이신, 퓨로마이신, 플레오마이신, 제오마이신, 블라스티시딘, 비오마이신, 암피실린, $O^6$BG/BCNU, 메토트렉세이트, 테트라사이클린, 아미노프테린, 하이포크산틴, 티미딘 키나제, DHFR, Gln 합성효소, 또는 ADA를 포함한다.

[0428] 특정 구현예에서, 음성 선택 카세트는 배양 배지에 존재하는 기질을 유전자를 발현하는 세포에 대한 독성 물질로 형질전환하기 위한 유전자를 포함한다. 이러한 분자는 간시클로비르 또는 FIAU의 존재에 민감한 헤르페스 바이러스(HSV TK)의 키나제 티미딘 유전자인 디프테리아 독소(DTA)의 해독 유전자(Yagi 등, *Anal Biochem.* 214(1):77-86, 1993; Yanagawa 등, *Transgenic Res.* 8(3):215-221, 1999)를 포함한다. HPRT 유전자는 또한 6-티오구아닌(6TG)을 배지에 첨가함으로써 음성 선택으로 사용될 수 있고. 모든 양성 및 음성 선택의 경우, poly A 전사 종결 서열은 상이한 기원으로부터 유래되며, 가장 고전적인 것은 SV40 poly A, 또는 진핵생물 유전자 poly A(소 성장 호르몬, 토끼 β-글로빈 등)으로부터 유래되고 있다.

[0429] 특정 구현예에서, 선택 카세트는 Olszko 등 (*Gene Therapy* 22: 591-595, 2015)에 기재된 바와 같은 $MGMT^{P140K}$를 포함한다. 특정 요소에서, 선택제는 $O^6$BG/BCNU를 포함한다.

[0430] 인간 알킬 구아닌 트랜스퍼라제(hAGT)를 암호화하는 약물 내성 유전자 MGMT는 니트로소우레아 및 테모졸로미드(TMZ)와 같은 알킬화제의 세포독성 효과에 내성을 부여하는 DNA 복구 단백질이다. 6-벤질구아닌(6-BG)은 니트로소우레아 독성을 강화하는 AGT의 억제제이며 이 제제의 세포독성 효과를 강화하기 위해 TMZ와 함께 공동 투여된다. AGT의 변이체를 암호화하는 MGMT의 여러 돌연변이체 형태 6-BG에 의한 불활성화에 고도로 내성이지만 DNA 손상을 복구하는 능력은 유지한다(Maze 등, *J. Pharmacol. Exp. Ther.* 290: 1467-1474, 1999). $MGMT^{P140K}$-기반 약물 내성 유전자 요법은 마우스, 개, 레서스 원숭이, 및 인간 세포, 구체적으로 조혈 세포에 화학보호를

부여하는 것으로 제시되었다(Zielske 등, *J. Clin. Invest.* 112: 1561-1570, 2003; Pollok 등, *Hum. Gene Ther.* 14: 1703-1714, 2003; Gerull 등, *Hum. Gene Ther.* 18: 451-456, 2007; Neff 등, *Blood* 105: 997-1002, 2005; Larochelle 등, *J. Clin. Invest.* 119: 1952-1963, 2009; Sawai 등, *Mol. Ther.* 3: 78-87, 2001).

[0431] 특정 구현예에서, 생체내 선택 카세트와의 조합은 유전자-보정된 세포의 선택적 이점 없이 질환에 대한 주요한 구성요소일 것이다. 예를 들어, SCID 및 일부 다른 면역결핍 및 FA에서, 보정된 세포는 이점이 있으며 치료 유전자를 "약간의" HSPC로 형질도입하는 것만이 치료 효능에 충분하다. 세포가 경쟁적 이점을 입증하지 않는 혈색소병증(즉, 겸상적혈구병 및 지중해빈혈)과 같은 다른 질환의 경우, 유전자 보정된 세포의 생체내 선택은 $MGMT^{P140K}$와 같은 생체내 선택 카세트와의 조합과 같이, 약간의 형질도입된 HSPC를 선택하여, 유전자 교정된 세포에서 증가를 허용하고 치료 효능을 달성하기 위함이다. 이 접근법은 또한 HSPC가 생체외 유전적 변형보다는 생체내 HIV에 내성을 가짐으로써 HIV에 적용될 수 있다.

[0432] **I(C)(iv). 스터퍼 서열**

[0433] 특정 구현예에서, 벡터는 스터퍼 서열을 포함한다. 특정 구현예에서, 스터퍼 서열은 게놈을 야생형 길이에 가까운 크기로 만들기 위해 추가될 수 있다. 스터퍼는 길이를 확장하기 위해 의도된 기능적으로 불활성인 서열을 정의하기 위해 의도된 당업계에서 일반적으로 인식되는 용어이다

[0434] 스터퍼 서열은 벡터의 효율적인 패키징 및 안정성을 달성하기 위해 사용된다. 특정 구현예에서, 스터퍼 서열은 게놈 크기를 야생형 바이러스 게놈 크기의 70% 내지 110%로 만드는 데 사용된다.

[0435] 스터퍼 서열은 바람직하게는 포유류 기원의 임의의 DNA일 수 있다. 본 발명의 바람직한 구현예에서, 스터퍼 서열은 포유류 기원의 비-코딩 서열, 예를 들어 인트론 단편이다.

[0436] 스터퍼 서열은 벡터의 크기를 미리 결정된 크기로 유지하는 데 사용될 때, 임의의 비-코딩 코딩 서열 또는 게놈을 분할 또는 비분할 세포에서 안정하게 유지하는 서열일 수 있다. 이들 서열은 다른 바이러스 게놈(예를 들어 엡스타인 바 바이러스) 또는 유기체(예를 들어 효모)로부터 유래될 수 있다. 예를 들어, 이들 서열은 센트로미어 및/또는 텔로미어의 기능적 부분일 수 있다.

[0437] **I(C)(v). 페이로드 통합 및 지지 벡터**

[0438] 유전자 요법은 종종 원하는 핵산 페이로드를 표적 세포의 게놈 내로 통합하는 데 필요하다. 다양한 시스템은 페이로드를 숙주 또는 표적 세포 게놈 내로 통합하기 위해 다양한 시스템이 설계 및/또는 사용될 수 있다. 이러한 다양한 시스템은 특정 페이로드 서열 특징 및 지지 벡터 및 지지 게놈(지지 게놈) 중 하나 이상을 포함할 수 있다.

[0439] 페이로드를 숙주 세포 게놈 내로 통합하는 아데노바이러스 벡터를 조작하는 한 가지 수단은 통합 바이러스 하이브리드 벡터을 생성하는 것이었다. 통합 바이러스 하이브리드 벡터는 표적 세포를 효율적으로 형질도입하는 벡터의 유전적 요소를 벡터 페이로드를 안정하게 통합하는 벡터의 유전적 요소와 조합한다. 예를 들어, 아데노바이러스 벡터와 조합하여 사용하기 위한 관심 통합 요소는 박테리오파지 인터그라제 PHiC31, 레트로트랜스포존, 레트로바이러스(예를 들어, LTR-매개 또는 레트로바이러스 통합-매개), 아연-핑거 뉴클레아제, DNA-결합 도메인-레트로바이러스 인터그라제 융합 단백질, AAV(예를 들어, AAV-ITR 또는 AAV-Rep 단백질-매개), 및 잠자는 미녀(SB) 트랜스포사제의 것들이 포함된다.

[0440] 본원에 기재된 Ad35 벡터는 임의적으로 트랜스포사제 및 트랜스포존을 포함하는 전위성 요소를 포함할 수 있다. 트랜스포사제는 레트로트랜스포존 또는 레트로바이러스 기원의 인터그라제, 뿐만 아니라 전위 가능한 기능적 핵산-단백질 복합체의 구성요소이고 전위를 매개하는 효소를 포함할 수 있다. 전위 반응은 트랜스포존 및 트랜스포사제 또는 인터그라제 효소를 포함한다. 특정 구현예에서, 통합 효율, 통합될 수 있는 DNA 서열 크기, 및 게놈 내로 통합될 수 있는 DNA 서열의 카피 수는 이러한 전위성 요소를 사용함으로써 개선될 수 있다. 트랜스포존은 DNA의 더 큰 분절의 상류 및 하류에 있는 말단 반복부 서열을 갖는 짧은 핵산 서열을 포함한다. 트랜스포사제는 말단 반복부 서열에 결합하고 트랜스포존을 게놈의 또 다른 부분으로 이동시키는 것을 촉매화한다.

[0441] 인간을 포함하는 척추동물의 게놈 내로 핵산의 삽입을 용이하게 하는 다수의 트랜스포사제가 당업계에 기재되었다. 이러한 트랜스포사제의 예는 잠자는 미녀("SB", 예를 들어, 연어과 어류의 게놈으로부터 유래); 피기백(piggyBac)(예를 들어, 인식목 세포 및/또는 마이오티스 루시푸구스(*Myotis lucifugus*)로부터 유래); 마리너(Mariner)(예를 들어, 드로소필라(Drosophila)로부터 유래); 개구리 왕자(frog prince)(예를 들어, 라나 피피엔스(*Rana pipiens*))로부터 유래; Tol1; Tol2(예를 들어, 송사리 어류로부터 유래); TcBuster(예를 들어, 붉은

밀가루갑충 트리볼리움 카스타네움(*Tribolium castaneum*)으로부터 유래), Helraiser, Himar1, Passport, Minos, Ac/Ds, PIF, Harbinger, Harbinger3-DR, HSmar1, 및 spinON을 포함한다.

[0442] 피기백(PB) 트랜스포사제는 예를 들어, Fraser 등, Insect Mol. Biol., 1996, 5, 141-51; Mitra 등, EMBO J., 2008, 27, 1097-1109; Ding 등, Cell, 2005, 122, 473-83; 및 미국 특허 번호 6,218,185; 6,551,825; 6,962,810; 7,105,343; 및 7,932,088에 기재된 조밀한 기능적 트랜스포사제 단백질이다. 과활성 피기백 트랜스포사제는 US 10,131,885에 기재되어 있다.

[0443] 특정 구현예에서, PB 트랜스포사제는 서열번호: 291(GenBank ABS12111.1)에 제시된 바와 같은 서열을 갖는다.

[0444] 특정 구현예에서, 개구리 왕자 트랜스포사제는 서열번호: 292(GenBank: AAP49009.1)에 제시된 바와 같은 서열을 갖는다. 또한 US2005/0241007을 참조한다.

[0445] 특정 구현예에서, TcBuster 트랜스포사제는 서열번호: 293(GenBank: ABF20545.1)에 제시된 바와 같은 서열을 갖는다.

[0446] 특정 구현예에서, Tol2 트랜스포사제는 서열번호: 294(GenBank: BAA87039.1)에 제시된 서열을 갖는다.

[0447] DNA 트랜스포존에 대한 추가적인 정보는 예를 들면, Munoz-Lopez & Garcia Perez, *Curr Genomics*, 11(2):115-128, 2010에서 찾을 수 있다.

[0448] 잠자는 미녀는 Ivics 등 *Cell* 91, 501-510, 1997; Izsvak 등, *J. Mol. Biol.*, 302(1):93-102, 2000; Geurts 등, *Molecular Therapy*, 8(1): 108-117, 2003; Mates 등 *Nature Genetics* 41:753-761, 2009; 및 미국 특허 번호 6,489,458; 7,148,203; 및 7,160,682; 미국 공개 번호 2011/117072; 2004/077572; 및 2006/252140에 기재되어 있다. 특정 구현예에서, 잠자는 미녀 트랜스포사제 효소는 서열 서열번호: 73을 갖는다. 특정 구현예에서, 과활성 잠자는 미녀(SB100x) 트랜스포사제 효소는 서열 서열번호: 74를 갖는다.

[0449] SB 트랜스포사제의 활성을 증가시키기 위해 체계적인 돌연변이생성 연구가 수행되었다. 예를 들어, Yant 등은 알라닌에 대한 SB 트랜스포사제의 N-말단 95 AA의 체계적인 교환을 수행하였다(*Mol. Cell Biol.* 24: 9239-9247, 2004). 이들 치환 중 10 개는 참조로서 SB10과 비교하여 200-400% 사이의 과활성을 야기하였다. Baus 등(*Mol. Therapy* 12: 1148-1156, 2005)에 기재된 SB16은 SB10과 비교하여 16-배 활성 증가를 갖는 것으로 보고되었다. 추가적인 과활성 SB 변이체는 Zayed 등 (*Molecular Therapy* 9(2):292-304, 2004) 및 US 9,840,696에 기재되어 있다.

[0450] SB 트랜스포존은 전위하기 위해 원형화될 필요가 있다(Yant 등, *Nature Biotechnology*, 20: 999-1005, 2002). 또한, 1.9 내지 7.2 kb의 트랜스포존의 경우, 트랜스포존 및 전위 빈도의 길이 사이에 역 선형 관계가 있다. 다시 말해서, SB 트랜스포사제는 더 작은 트랜스포존과 비교하여 더 큰 트랜스포존의 전달을 덜 효율적으로 매개한다(Geurts 등, *Mol Ther.*, 8(1):108-17, 2003).

[0451] SB 트랜스포사제는 SB ITR 사이에 위치한 핵산 트랜스포존 페이로드를 전위시킨다. 다양한 SB ITR이 당업계에 알려져 있다. 일부 구현예에서, SB ITR은 트랜스포사제에 대한 인식 신호로서 역할을 하는 32 bp 길이의 불완전한 지시 반복부를 포함하는 230 bp 서열이다. pT, pT2, pT3, pT2B, 및 pT4로 알려진 SB ITR을 포함하는 조작된 SB ITR이 당업계에 알려져 있다. 일부 구현예에서, pT4 ITR은 예를 들어, SB100x 트랜스포사제에 의한 전위를 위해, 예를 들어, 본 개시내용의 트랜스포존 페이로드를 플랭킹하는 데 사용된다.

[0452] 특정 구현예에서, 잠자는 미녀의 IR(도립 반복부)/DR(지시 반복부) 및 염색체 서열을 암호화하는 서열은 서열번호: 4를 포함한다. 특정 구현예에서, 잠자는 미녀의 IR/DR 및 염색체 서열을 암호화하는 서열은 서열번호: 5를 포함한다. 특정 구현예에서, 잠자는 미녀의 IR/DR 암호화 서열은 서열번호: 295를 포함한다. 특정 구현예에서, 잠자는 미녀의 IR/DR 및 염색체 서열을 암호화하는 서열은 서열번호: 296을 포함한다. 특정 구현예에서, 잠자는 미녀의 IR/DR 및 염색체 서열을 암호화하는 서열은 서열번호: 297을 포함한다. 특정 구현예에서, 잠자는 미녀의 IR/DR을 암호화하는 서열은 서열번호: 298을 포함한다. 특정 구현예에서, 잠자는 미녀의 IR/DR 및 염색체 서열을 암호화하는 서열은 서열번호: 299를 포함한다. 특정 구현예에서, 잠자는 미녀의 IR/DR을 암호화하는 서열은 서열번호: 300을 포함한다.

[0453] 다양한 구현예에서, Ad35 공여자 벡터 또는 게놈은 β-글로빈 발현 산물 또는 γ-글로빈 발현 산물을 암호화하는 적어도 하나의 코딩 서열을 포함하는 통합 요소를 플랭킹하는 SB100x 트랜스포존 도립 반복부를 포함하는 페이로드를 포함한다.

[0454] 다양한 구현예에서, 아데노바이러스 전위 시스템은 트랜스포존 도립 반복부에 의해 플랭킹된 통합 요소를 포함하는 Ad35 공여자 벡터 또는 게놈을 포함하고, 아데노바이러스 지지 벡터 또는 지지 게놈을 추가로 포함할 수 있다. 다양한 구현예에서, 지지 벡터는 (i) 아데노바이러스 캡시드; 및 (ii) 통합 요소를 플랭킹하는 도립 반복부에 상응하는 트랜스포사제를 암호화하는 핵산 서열을 포함하는 아데노바이러스 지지 게놈을 포함한다. 따라서, 다양한 구현예에서, 지지 벡터 또는 지지 게놈의 적어도 하나의 기능은 표적 세포에 투여된 공여자 벡터에 존재하는 통합 요소의 전위를 위한 트랜스포사제를 표적 세포에 암호화, 발현, 및/또는 전달할 수 있다. 예를 들면, 일부 구현예에서, Ad35 공여자 벡터 또는 게놈은 β-글로빈 발현 산물 또는 γ-글로빈 발현 산물을 암호화하는 적어도 하나의 코딩 서열을 포함하는 통합 요소를 플랭킹하는 SB100x 트랜스포존 도립 반복부를 포함하고, 지지 벡터 또는 지지 게놈은 SB100x 트랜스포사제를 암호화하는 코딩 서열을 포함한다. 특정 구현예에서, 예를 들어, 통합 요소가 트랜스포존 도립 반복부에 의해 플랭킹되고 트랜스포존 도립 반복부가 재조합효소 지시 반복부에 의해 플랭킹되는 경우 통합 요소는 재조합효소 지시 반복부에 의해 플랭킹된다. 이러한 특정 구현예에서, 지지 벡터 또는 지지 게놈의 적어도 하나의 기능은 표적 세포에 투여된 공여자 벡터에 존재하는 재조합효소 부위의 재조합을 위한 재조합효소를 표적 세포로 암호화, 발현, 및/또는 전달하는 것일 수 있다. 다양한 구현예에서, 지지 벡터 또는 지지 게놈은 표적 세포에 투여된 공여자 벡터에 존재하는 재조합효소의 부위의 재조합을 위해 재조합효소를 표적 세포로 암호화, 발현, 및/또는 전달할 수 있고 또한 표적 세포에 투여된 공여자 벡터에 존재하는 통합 요소의 전위를 위해 트랜스포사제를 표적 세포로 암호화, 발현, 및/또는 전달한다.

[0455] 본원에 개시된 특정 구현예는 또한 부위-특이적 재조합효소 시스템을 사용한다. 이러한 구현예에서, 적어도 하나의 치료 유전자 이외에도, 트랜스포사제-인식된 도립 반복부를 포함하는 트랜스포존은 또한 적어도 하나의 재조합효소-인식된 부위를 포함한다. 따라서, 특정 구현예에서, 본 개시내용은 또한 다음을 투여하는 것을 포함하는 치료 유전자를 게놈 내로 통합하는 방법을 제공한다: (a) (i) 트랜스포사제에 의해 인식된 도립 반복 서열 및 (ii) 재조합효소-인식 부위에 의해 플랭킹된 치료 유전자를 포함하는 트랜스포존; 및 b) 플라스미드, 에피솜, 또는 이식유전자로부터 치료 유전자를 절제하고 치료 유전자를 게놈 내로 통합하는 역할을 하는 트랜스포사제 및 재조합효소. 일부 구현예에서, (b)의 단백질(들)은 단백질(들)을 암호화하는 핵산으로 투여된다. 일부 구현예에서, (b)의 단백질(들)을 암호화하는 트랜스포존 및 핵산은 별개의 벡터 상에 존재한다. 일부 구현예에서, (b)의 단백질(들)을 암호화하는 트랜스포존 및 핵산은 동일한 벡터 상에 존재한다. 동일한 벡터 상에 존재할 때, (b)의 단백질(들)을 암호화하는 벡터의 부분은 (a)의 트랜스포존을 보유하는 부분의 외부에 위치한다. 다시 말해서, 트랜스포사제 및/또는 재조합효소 암호화 영역은 도립 반복부 및/또는 재조합효소-인식 부위에 의해 플랭킹된 영역 외부에 위치한다. 전술된 방법에서, 트랜스포사제 단백질은 표적 세포 게놈 내로 삽입되는 핵산과 같은 삽입된 핵산을 플랭킹하는 도립 반복부를 인식한다. 재조합효소 및 재조합효소-인식된 부위의 사용은 게놈 내로 추가로 통합될 수 있는 트랜스포존의 크기를 증가시킬 수 있다.

[0456] 재조합효소 시스템의 예는 Flp/Frt 시스템, Cre/loxP 시스템, Dre/rox 시스템, Vika/vox 시스템, 및 PhiC31 시스템을 포함한다.

[0457] Flp/Frt DNA 재조합효소 시스템은 사카로마이세스 세레비지에(Saccharomyces cerevisiae)로부터 단리되었다. Flp/Frt 시스템은 Frt 인식 부위에서 DNA-재조합을 촉매화하는 재조합효소 Flp(플리파제)를 포함한다. 특정 구현예에서, Flp(플리파제)는 서열 서열번호: 75를 포함하고 FRT 인식 부위는 서열번호: 76을 포함한다.

[0458] Flp 단백질의 변이체는 서열번호: 77(GenBank: ABD57356.1) 및 서열번호: 78(GenBank: ANW61888.1)을 포함한다.

[0459] Cre/loxP 시스템은 예를 들어, EP 02200009B1에 기재되어 있다. Cre는 박테리오파지 P1로부터 단리된 부위-특이적 DNA 재조합효소이다. 특정 구현예에서, Cre는 서열 서열번호: 79를 포함한다.

[0460] Cre 단백질의 인식 부위는 loxP 부위(서열번호: 80)인 34 개 염기 쌍의 뉴클레이다. Cre는 13 개 염기 쌍 도립 반복부에 결합하고 스페이서 영역 내에서 가닥 절단 및 재결찰을 촉매화함으로써 34 bp loxP DNA 서열을 재조합한다. 스페이서 영역에서 Cre에 의해 만들어진 엇갈린 DNA 절단은 6 개의 염기 쌍에 의해 분리되어 동일한 중첩 영역을 갖는 재조합 부위로만 재조합을 보장하도록 상동성 센서로서 작용하는 중첩 영역을 제공한다. 또한 사용될 수 있는 lox 인식 부위의 변이체는 lox2272(서열번호: 81); lox511(서열번호: 82); lox66(서열번호: 83); lox71(서열번호: 84); loxM2(서열번호: 85); 및 lox5171(서열번호: 86)을 포함한다.

[0461] VCre/VloxP 재조합효소 시스템은 비브리오 플라스미드 p0908로부터 단리되었다. 특정 구현예에서, 이 시스템의

VCre 재조합효소는 서열번호: 87을 포함하고 VloxP 인식 부위는 서열번호: 88을 포함한다.

[0462] sCre/SloxP 시스템은 WO 2010/143606에 기재되어 있다. Dre/rox 시스템은 US 7,422,889 및 US 7,915,037B2에 기재되어 있다. 일반적으로 서열 서열번호: 89 및 rox 인식 부위(서열번호: 90)를 갖는 엔테로박테리아(Enterobacteria) 파지 D6으로부터 단리된 Dre 재조합효소를 포함한다.

[0463] Vika/vox 시스템은 미국 특허 번호 10,253,332에 기재되어 있다. 추가적으로, PhiC31 재조합효소는 AttB/AttP 결합 부위를 인식한다.

[0464] 트랜스포존(도립 반복부 및/또는 재조합효소 인식 부위 포함)을 포함하는 벡터 핵산의 양, 및 많은 구현예에서 트랜스포사제 및/또는 재조합효소를 암호화하는 벡터 핵산의 양은 트랜스포존 핵산의 원하는 절제 및 삽입을 표적 세포 게놈 내로 제공하기에 충분한 세포 내로 도입된다. 이와 같이, 도입된 벡터 핵산의 양은 트랜스포사제 활성 및/또는 재조합효소 활성의 충분한 양 및 표적 세포 게놈 내로 삽입되는 것이 바람직한 트랜스포존의 충분한 카피 수를 제공하여야 한다. 특정 구현예는 트랜스포존 대 트랜스포사제/재조합효소를 1:1; 1:2; 또는 1:3 비로 포함한다.

[0465] 대상 방법은 핵산의 표적 세포 게놈 내로의 안정된 통합을 초래한다. 안정된 통합이란 핵산이 일시적 기간을 초과하여 표적 세포 게놈에 존재하고 염색체 유전적 물질의 일부를 통과시켜 표적 세포의 자손에게 주는 것을 의미한다.

[0466] 본 개시내용의 실시예 2는 과활성 잠자는 미녀 트랜스포사제가 32.4 kb 트랜스포존을 HSPC의 게놈 내로 통합하는 데 사용될 수 있다는 놀라운 결과를 기재한다. 이러한 구현예는 도 23에 도시된 바와 같이 Flp/Frt 시스템과 조합하여 SBX100의 사용을 포함한다.

[0467] 이전에 나타낸 바와 같이, 특정 구현예는 상동성 아암을 활용하여 상동성 지시된 복구를 활용하는 유전적 작제물의 표적화된 삽입을 용이하게 한다. 상동성 아암은 절단 부위 및 상동성을 보유하는 게놈 서열 사이의 HDR을 지지하기 위해, 예를 들어, 절단 부위의 50 개 미만의 염기 이내 , 예를 들어, 30 개 염기 이내, 15 개 염기 이내, 10 개 염기 이내, 5 개 염기 이내, 또는 절단 부위를 바로 플랭킹하는, 절단 부위에서 게놈 서열에 충분한 상동성, 예를 들어 절단 부위를 플랭킹하는 뉴클레오티드 서열과 70%, 80%, 85%, 90%, 95%, 또는 100% 상동성을 갖는 임의의 길이일 수 있다. 상동성 아암은 일반적으로 게놈 서열, 예를 들어, 이중 가닥 파괴(DSB)가 발생하는 게놈 영역과 동일하다. 그러나, 나타낸 바와 같이, 절대적인 동일성은 필요하지 않다.

[0468] 특정 구현예는 상동성-지시된 복구 주형 및 표적화된 게놈 서열 사이에 25, 50, 100, 또는 200 개 뉴클레오티드(nt), 또는 200 nt 초과의 서열 상동성(또는 10 내지 200 개 뉴클레오티드 사이, 또는 그 이상의 임의의 정수 값)을 갖는 상동성 아암을 활용할 수 있다. 특정 구현예에서, 상동성 아암은 40 - 1000 nt 길이이다. 특정 구현예에서, 상동성 아암 500-2500 개 염기 쌍, 700 - 2000 개 염기 쌍, 또는 800 -1800 개 염기 쌍. 특정 구현예에서, 상동성 아암은 적어도 800 개의 염기 쌍 또는 적어도 850 개의 염기 쌍을 포함한다. 상동성 아암의 길이는 또한 대칭적 또는 비대칭적일 수 있다.

[0469] 특정 구현예는 표적 게놈의 상응하는 단편과 서열 동일성 또는 상동성을 갖는, 적어도 25, 50, 100, 200, 400, 600, 800, 1,000, 1,200, 1,400, 1,600, 1,800, 2,000, 2,500, 또는 3,000 개 뉴클레오티드 또는 그 이상을 각각 포함하는 제1 및/또는 제2 상동성 아암을 활용할 수 있다. 일부 구현예에서, 제1 및/또는 제2 상동성 아암은 각각 25, 50, 100, 200, 400, 600, 800, 1,000, 1,200, 1,400, 1,600, 또는 1,800 개 뉴클레오티드의 하한 및 1,000, 1,200, 1,400, 1,600, 1,800, 2,000, 2,500, 또는 3,000 개 뉴클레오티드의 상한을 갖는 표적 게놈의 상응하는 단편과 서열 동일성 또는 상동성을 갖는 다수의 뉴클레오티드를 포함한다. 일부 구현예에서, 제1 및/또는 제2 상동성 아암은 각각 40 내지 1,000 개의 뉴클레오티드, 500 내지 2,500 개의 뉴클레오티드, 700 내지 2,000 개의 뉴클레오티드, 또는 800 내지 1800 개의 뉴클레오티드이거나, 또는 적어도 800 개 뉴클레오티드 또는 적어도 850 개 뉴클레오티드 길이를 갖는 표적 게놈의 상응하는 단편과 서열 동일성 또는 상동성을 갖는 다수의 뉴클레오티드를 포함한다. 제1 및 제2 상동성 아암은 동일한, 유사한, 또는 상이한 길이를 가질 수 있다.

[0470] 상동성 아암에 관한 추가적인 정보를 위해, Richardson 등, *Nat Biotechnol*. 34(3):339-44, 2016 참조.

[0471] 특정 구현예에서, 유전적 작제물(예를 들어, 세포 내에서 치료 산물의 발현을 야기하는 유전자)은 게놈 안전 항구 내로 정확하게 삽입된다. 게놈 안전 항구 부위는 숙주 세포에 대한 부작용 없이 새롭게 통합된 DNA의 예측 가능한 발현을 수용할 수 있는 게놈의 유전자내 또는 유전자외 영역이다. 유용한 안전 항구는 암호화된 단백질의 원하는 수준을 산출하기에 충분한 이식유전자 발현을 허용하여야 한다. 게놈 안전 항구 부위는 또한 세포

기능을 변경시키지 않아야 한다. 게놈 안전 항구 부위를 식별하는 방법은 Sadelain 등, *Nature Reviews* 12:51-58, 2012; 및 Papapetrou 등, *Nat Biotechnol*. 29(1):73-8, 2011에 기재되어 있다. 특정 구현예에서, 게놈 안전 항구 부위는 다음 기준 중 하나 이상(1, 2, 3, 4, 또는 5 개)를 충족한다: (i) 임의의 유전자의 5' 단부로부터 적어도 50 kb의 거리, (ii) 임의의 암-관련된 유전자로부터 적어도 300 kb의 거리, (iii) (천연 또는 조작된 뉴클레아제를 사용한 DNA 절단에 의해 측정된) 개방/접근가능한 염색질 구조 내, (iv) 유전자 전사 단위 외부의 위치 및 (v) 게놈의 초보존된 영역(UCR), microRNA 또는 긴 비-코딩 RNA 외부의 위치.

[0472] 특정 구현예에서, 게놈 안전 항구의 기준을 충족하기 위해, 염색질 부위는 알려진 암유전자로부터 150 kb 초과로 멀리 떨어져 있어야 하고, 알려진 전사 시작 부위로부터 30 kb 초과로 멀리 떨어져 있어야 하고; 코딩 mRNA와 중첩되지 않아야 한다. 특정 구현예에서, 게놈 안전 항구의 기준을 충족하기 위해, 염색질 부위는 알려진 암유전자로부터 200 kb 초과로 멀리 떨어져 있어야 하고, 알려진 전사 시작 부위로부터 40 kb 초과로 멀리 떨어져 있어야 하고; 코딩 mRNA와 중첩되지 않아야 한다. 특정 구현예에서, 게놈 안전 항구의 기준을 충족하기 위해, 염색질 부위는 알려진 암유전자로부터 300 kb 초과로 멀리 떨어져 있어야 하고, 알려진 전사 시작 부위로부터 50 kb 초과로 멀리 떨어져 있어야 하고; 코딩 mRNA와 중첩되지 않아야 한다. 특정 구현예에서, 게놈 안전 항구는 전술한 기준을 충족하고(알려진 전사 시작 부위로부터 150 kb 초과, 200 kb 초과 또는 300 kb 초과로 멀리 떨어져 있고; 코딩 mRNA와 중첩되지 않고, 코딩 mRNA와 중첩하지 않고 알려진 전사 시작 부위로부터 40 kb 초과, 또는 50 kb 초과로 멀리 떨어져 있음) 추가적으로 관련 발견의 신속한 임상 번역을 허용하기 위해 관련 동물 모델의 동물 및 인간 게놈 사이에 100% 상동이다.

[0473] 특정 구현예에서, 게놈 안전 항구는 본원에 기재된 기준을 충족하고 또한 유전자좌가 주변 유전적 물질에 영향을 미치지 않음을 추가로 입증하는 렌티바이러스 통합의 정방향:역방향의 1:1 비를 입증한다.

[0474] 특정 게놈 안전 항구 부위는 CCR5, HPRT, AAVS1, Rosa 및 알부민을 포함한다. 또한, 적절한 게놈 안전 항구 통합 부위에 대한 추가적인 정보 및 옵션에 대해 예를 들어, 미국 특허 번호 7,951,925 및 8,110,379; 미국 공개 번호 2008/0159996; 2010/00218264; 2012/0017290; 2011/0265198; 2013/0137104; 2013/0122591; 2013/0177983 및 2013/0177960 참조.

[0475] 당업계에 알려진 다양한 기술은 게놈 안전 항구와 같은 특이적 게놈 유전자좌에서 통합 요소의 통합을 지시하는 데 사용될 수 있다. 예를 들어 AAV-매개 유전자 표적화, 뿐만 아니라 부위-특이적 엔도뉴클레아제(아연-핑거 뉴클레아제, 메가뉴클레아제, 전사 활성인자-유사 효과기(TALE) 뉴클레아제), 및 CRISPR/Cas 시스템을 사용한 DNA 이중 가닥 파괴의 도입에 의해 향상된 상동 재조합은 게놈 안전 항구와 같은 미리 결정된 게놈 유전자좌에서 외래 DNA의 표적화된 삽입을 매개할 수 있는 모든 도구이다. 면역억제 레지멘은 예를 들어, 미국 가출원 번호 제63/009,218호에 기재되어 있으며, 이는 특히 면역억제 레지멘과 관련하여 그 전문이 본원에 참조로 포함된다.

[0476] 특정 구현예에서, 게놈 안전 항구와 같은 특이적 게놈 유전자좌에서 통합 요소의 통합은 표적 게놈의 CRISPR 효소-매개 절단을 사용한 상동성-지시된 통합을 포함할 수 있다. CRISPR 효소(예를 들어, Cas9)는 가이드 RNA(gRNA)에 의해 명시된 부위에서 이중 가닥 DNA를 절단한다. 이중 가닥 파괴는 공여자 주형(예컨대 왼쪽 및 오른쪽 상동성 아암을 포함하는 Ad35 페이로드 통합 요소)이 존재할 때 상동성-지시된 복구(HDR)에 의해 복구될 수 있다. 이러한 다양한 방법에서, 통합 요소는 왼쪽 및 오른쪽 상동성 아암(예를 들어, 500-3,000 bp)을 절단된 표적 게놈 내로 삽입하기 위해 포함한다는 점에서 "복구 주형"이다. CRISPR-매개 유전자 삽입은 DNA 주형의 자발적인 재조합과 비교하여 몇 배 더 효율적일 수 있으며, 이는 CRISPR-매개 유전자 삽입이 게놈 편집을 위한 효과적인 도구일 수 있음을 입증한다. 핵산 서열을 명시된 게놈 유전자좌 내로 상동성-지시 통합을 위한 예시적인 방법은 당업계, 예를 들어, Richardson 등 (*Nat Biotechnol*. 34(3):339-44, 2016)에 알려져 있다.

[0477] 다양한 구현예에서, 표적 세포 게놈의 게놈 안전 항구에서 삽입을 위한 통합 요소를 포함하는 아데노바이러스 공여자 벡터는 최대 15 kb의 길이를 갖는 핵산 서열의 통합을 유발할 수 있다. 다양한 구현예에서, 게놈 안전 항구에서 표적 세포 게놈 내로의 통합을 위한 통합 요소는 예를 들어, 길이가 1 kb, 2 kb, 3 kb, 4 kb, 또는 5 kb의 하한 및 10 kb, 11 kb, 12 kb, 13 kb, 14 kb, 또는 15 kb의 상한을 갖는 적어도 1 kb, 2 kb, 3 kb, 4 kb, 5 kb, 6 kb, 7 kb, 8 kb, 9 kb, 10 kb, 11 kb, 12 kb, 13 kb, 14 kb, 또는 15 kb의 길이를 가질 수 있다.

[0478] **II. 표적 세포 집단**

[0479] 다양한 구현예에서, 본 개시내용의 Ad35 공여자 벡터 및 게놈은 본원에 개시된 HSC, T 세포, B 세포, 및 종양

세포를 포함하나 이에 제한되지 않는 임의의 다양한 유형의 표적 세포를 형질도입할 수 있다.

[0480] II(A). HSC

[0481] 특정 구현예에서, 벡터-표적화된 세포 유형은 조혈 줄기 세포(HSC)를 포함한다. HSC는 CD46을 결합함으로써 생체내 유전적 변형을 위해 표적화된다. 나타낸 바와 같이, 본 개시내용 내에서, HSC는 CD46을 결합함으로써 생체내 유전적 변형을 위해 표적화된다. 벡터는 CD46 결합의 특이성 및/또는 강도를 증가시키기 위해 본원에 개시된 돌연변이를 포함할 수 있다. HSC는 또한 다음 마커 프로파일에 의해 식별될 수 있다: CD34+, Lin-CD34+CD38-CD45RA-CD90+CD49f+(HSC1) 및 CD34+CD38-CD45RA-CD90- CD49f+(HSC2). 인간 HSC1은 다음 프로파일에 의해 식별될 수 있다: CD34+/CD38-/CD45RA-/CD90+ 또는 CD34+/CD45RA-/CD90+ 및 마우스 LT-HSC는 Lin-Sca1+ckit+CD150+CD48-Flt3-CD34-에 의해 식별될 수 있다(여기서 Lin은 CD3, Cd4, CD8, CD11b, CD11c, NK1.1, Gr1, 및 TER119를 포함하는 성숙 세포의 임의의 마커의 발현 부재를 나타낸다). 특정 구현예에서, HSC는 CD164+ 프로파일에 의해 식별된다. 특정 구현예에서, HSC는 CD34+/CD164+ 프로파일에 의해 식별된다. HSC 마커 프로파일에 관한 추가적인 정보를 위해, WO2017/218948 참조.

[0482] II(B). T 세포

[0483] T-세포의 여러 상이한 하위세트가 발견되었으며, 각각은 별개의 기능이 있다. 예를 들어, 대부분의 T-세포는 여러 단백질의 복합체로서 존재하는 T-세포 수용체(TCR)가 있다. 실제 T-세포 수용체는 독립적인 T-세포 수용체 알파 및 베타(TCRα 및 TCRβ) 유전자로부터 생성되고 α- 및 β-TCR 쇄라고 불리는 2 개의 개별 펩티드 쇄로 구성된다.

[0484] γδ T-세포는 표면 상에 별개의 T-세포 수용체(TCR)를 보유하는 T-세포의 작은 하위세트를 나타낸다. γδ T-세포에서, TCR은 하나의 γ-쇄 및 하나의 δ-쇄로 구성된다. 이 T-세포 그룹은 αβ T-세포보다 훨씬 덜 일반적이다(총 T-세포의 2%).

[0485] CD3은 모든 성숙 T 세포 상에서 발현된다. 활성화된 T-세포는 4-1BB (CD137), CD69, 및 CD25를 발현한다. CD5 및 트랜스페린 수용체는 또한 T-세포 상에서 발현된다.

[0486] T-세포는 추가로 헬퍼 세포(CD4+ T-세포) 및 세포독성 T-세포(CTL, CD8+ T-세포)로 분류될 수 있으며, 이는 세포용해성 T-세포를 포함한다. T 헬퍼 세포는 다른 기능 중에서 B 세포의 형질 세포로의 성숙 및 세포독성 T-세포 및 대식세포의 활성화를 포함하는 면역학적 과정에서 다른 백혈구를 보조한다. 이들 세포는 또한 CD4+ T-세포로 알려져 있는데 표면 상에서 CD4 단백질을 발현하기 때문이다. 헬퍼 T-세포는 항원 제시 세포(APC)의 표면 상에서 발현된 MHC 클래스 II 분자에 의해 펩티드 항원과 함께 제시될 때 활성화되게 된다. 일단 활성화되면, 이들은 빠르게 분할되어 활성 면역 반응에서 조절하거나 또는 보조하는 사이토카인이라 불리는 작은 단백질을 분비한다.

[0487] 세포독성 T-세포는 바이러스로 감염된 세포 및 종양 세포를 파괴하고, 또한 이식 거부에 연루된다. 이들 세포는 또한 CD8+ T-세포로 알려져 있는데 표면 상에서 CD8 당단백질을 발현하기 때문이다. 이들 세포는 신체의 거의 모든 세포의 표면 상에 존재하는 MHC 클래스 I과 회합된 항원에 대한 결합에 의해 표적을 인식한다.

[0488] 특정 구현예에서, CAR은 세포독성 T-세포에서 발현되도록 유전적으로 변형된다.

[0489] 본원에 사용된 바와 같은 "중추 기억" T-세포(또는 "TCM")는 이의 표면 상에서 CD62L 또는 CCR7 및 CD45RO를 발현하고, 미처리 세포와 비교하여 CD45RA를 발현하지 않거나 또는 이의 발현이 감소된 항원 경험 CTL을 지칭한다. 특정 구현예에서, 중추 기억 세포는 CD62L, CCR7, CD25, CD127, CD45RO, 및 CD95의 발현에 대해 양성이고, 미처리 세포와 비교하여 CD45RA의 발현이 감소된다.

[0490] 본원에 사용된 바와 같은 "효과기 기억" T-세포(또는 "TEM")는 중추 기억 세포와 비교하여 이의 표면 상에서 CD62L을 발현하지 않거나 또는 이의 발현이 감소되고 미처리 세포와 비교하여 CD45RA를 발현하지 않거나 또는 이의 발현이 감소된 항원 경험 T-세포를 지칭한다. 특정 구현예에서, 효과기 기억 세포는 미처리 세포 또는 중추 기억 세포와 비교하여 CD62L 및 CCR7의 발현에 대해 음성이고, CD28 및 CD45RA의 가변 발현을 갖는다. 효과기 T-세포는 기억 또는 미처리 T-세포와 비교하여 그랜자임 B 및 퍼포린에 대해 양성이다.

[0491] 본원에 사용된 바와 같은 "미처리" T-세포는 중추 또는 효과기 기억 세포와 비교하여 CD62L 및 CD45RA를 발현하고 CD45RO를 발현하지 않는 비-항원 경험 T 세포를 지칭한다. 특정 구현예에서, 미처리 CD8+ T 림프구는 CD62L, CCR7, CD28, CD127, 및 CD45RA를 포함하는 미처리 T-세포의 표현형 마커의 발현을 특징으로 한다.

[0492] **II(C). B 세포**

[0493] B 세포는 체액 반응의 매개체이며 항원에 특이적인 항체의 생산 및 방출을 담당한다. 여러 유형의 B 세포가 존재하며 이는 핵심 마커를 특징으로 할 수 있다. 일반적으로, 미성숙 B 세포는 CD19, CD20, CD34, CD38, 및 CD45R을 발현하고, 이들이 성숙함에 따라 핵심 발현된 마커는 CD19 및 IgM이다.

[0494] **II(D). 종양**

[0495] 특정 구현예에서, 벡터는 종양을 표적화할 수 있다. 특정 구현예에서, 종양은 종양 세포 상에 존재하지만 건강한 세포 상에서는 존재하지 않는 수용체를 표적화함으로써 표적화된다. 종양은 αv 인테그린에 결합함으로써 생체내 유전적 변형을 위해 표적화될 수 있다. αv 인테그린은 혈관신생에 중요한 역할을 한다. αvβ3 및 αvβ5 인테그린은 정상 내피 세포에서 부재하거나 또는 낮은 수준으로 발현되지만 종양의 혈관신생 혈관계에서 유도된다(Brooks 등, *Cell*, 79: 1157-1164, 1994; Hammes 등, *Nature* Med, 2: 529-533, 1996). 아미노펩티다제 N/CD13은 최근에 NGR 모티프에 대한 혈관신생 수용체로서 식별되었다(Burg 등, *Cancer Res*, 59:2869-74, 1999). 아미노펩티다제 N/CD13은 암의 혈관신생 혈관 및 다른 혈관신생 조직에서 강하게 발현된다.

[0496] 특정 구현예에서, 벡터는 암 세포 항원 에피토프를 표적화함으로써 종양을 표적화할 수 있다. 암 세포 항원은 암 세포 또는 종양에 의해 발현된다.

[0497] 특정 구현예에서, 암 세포 항원 에피토프는 암 세포에 의해 우선적으로 발현된다. "우선적으로 발현된"은 암 세포 항원이 다른 세포 유형과 비교하여 암 세포에 대한 더 높은 수준으로 발견됨을 의미한다. 일부 경우에, 암 항원 에피토프는 표적화된 암 세포 유형에 의해서만 발현된다. 다른 경우에, 암 항원은 비-표적화된 세포보다 적어도 25%, 35%, 45%, 55%, 65%, 75%, 85%, 95%, 96%, 97%, 98%, 99%, 또는 100% 더 많은 표적화된 암 세포 유형 상에서 발현된다.

[0498] 특정 구현예에서, 암 세포 항원은 암성 및 건강한 조직 상에서 상당히 발현된다. 특정 구현예에서, 상당히 발현된은 이중특이적 항체의 사용이 표적내/암외 독성에 기반한 개발 동안 중단되었음을 의미한다. 특정 구현예에서, 상당히 발현된은 이중특이적 항체의 사용이 표적내/암외 독성에 기반한 잠재적인 부정적인 부작용에 관한 경고를 필요로 함을 의미한다. 일 예로서, 세툭시맙은 피부에서 EGFR 발현으로 인한 것으로 생각되는 여러 피부 발진과 연관된 항-EGFR 항체이다. 또 다른 예는 항-HER2 (ERBB2) 항체인 허셉틴(트라스투주맙)이다. 허셉틴은 심장에서 표적 발현으로 인한 심장독성과 연관되어 있다. 더욱이, CAR-T 세포로 Her2를 표적화하는 것은 폐에서 표적내, 암외 발현으로 인해 환자에게 치명적이었다.

[0499] 표 12는 특정 암 유형에서 공동 발현될 가능성이 높은 암 항원의 예를 제공한다.

[0500] 표 12

| 공동 발현될 가능성이 있는 암 항원 | 암 유형 |
|---|---|
| CD19, CD20, CD22, ROR1, CD33, CD56, CLL-1, WT-1, CD123, PD-L1, EFGR | 백혈병/림프종 |
| B-세포 성숙 항원 (BCMA), PD-L1, EFGR | 다발성 골수종 |
| PSMA, WT1, 전립선 줄기 세포 항원 (PSCA), SV40 T, PD-L1, EFGR | 전립선암 |
| HER2, ERBB2, ROR1, PD-L1, EFGR, MUC16, 폴레이트 수용체 (FOLR), CEA | 유방암 |
| CD133, PD-L1, EFGR | 줄기 세포암 |
| L1-CAM, MUC16, FOLR, Lewis Y, ROR1, 메소텔린, WT-1, PD-L1, EFGR, CD56 | 난소암 |
| 메소텔린, PD-L1, EFGR | 중피종 |
| 카르복시-탈수효소-IX (CAIX); PD-L1, EFGR | 신세포 암종 |
| GD2, PD-L1, EFGR | 흑색종 |
| 메소텔린, CEA, CD24, ROR1, PD-L1, EFGR, MUC16 | 췌장암 |
| ROR1, PD-L1, EFGR, 메소텔린, MUC16, FOLR, CEA, CD56 | 폐암 |
| 메소텔린, PD-L1, EFGR | 담관암 |
| MUC16, PD-L1, EFGR, | 방광암 |
| ROR1, 글리피칸-2, CD56, 디시알로강글리오시드, PD-L1, EFGR, | 신경모세포종 |
| CEA, PD-L1, EFGR, | 결장직장암 |
| CD56, PD-L1, EFGR, | 메르켈 세포 암종 |

[0501]

[0502] 보다 특정 예에서, 암 세포 항원은 다음을 포함한다: 메소텔린, MUC16, FOLR, PD-L1, ROR1, 글리피칸-2(GPC2), 디시알로강글리오시드(GD2), HER2, EGFR, EGFRvIII, CEA, CD56, CLL-1, CD19, CD20, CD123, CD30, CD33(전장), CD33(델타E2 변이체), CD33(C-말단 절두됨), BCMA, IGFR, MUC1, VEGFR, PSMA, PSCA, IL13Ra2, FAP, EpCAM, CD44, CD133, Tro-2, CD200, FLT3, GCC, 및 WT1. 당업자에 의해 이해되는 바와 같이, 표적화된 항원은 신호 펩티드가 결여될 수 있다.

[0503] 신경 세포 접착 분자 1(NCAM1)로도 알려져 있는 CD56은 세포-세포 및 세포-기질 접착에 수반되는 I형 막 당단백질이다. 이의 세포외 도메인은 N-말단에서 5 개의 IgG-유사 도메인 및 막-근위 영역에서 2 개의 피브로넥틴 III형 도메인을 갖는다.

[0504] 디시알로강글리오시드 GalAc베타1-4(NeuAc알파2-8NeuAc알파2-3)Gal베타1-4Glc베타1- 1Cer(GD2)은 신경모세포종을 포함하는 다양한 종양에서 발현된다. 디시알로강글리오시드 항원 GD2는 시알산 및 지질 잔기에 의해 플랭킹된 올리고당류의 백본을 포함한다. 예를 들어, Cheresh (*Surv. Synth. Pathol. Res.* 4:97, 1987) 및 US 5,653,977 참조.

[0505] EGFR의 종양 특이적 돌연변이체인 EGFR 변이체 III(EGFRvIII)은 종종 야생형 EGFR 유전자 증폭과 연관된 게놈 재배열의 생성물이다. EGFRvIII은 엑손 2-7의 인프레임 결실에 의해 형성되어, 접합부에서 글리신 치환에 따른 267 개 아미노산의 결실로 이어진다. 절두된 수용체는 리간드에 결합하는 능력을 상실하지만 구성적 키나제 활성을 획득한다. 흥미롭게도, EGFRvIII은 동일한 종양 세포에서 전장 야생형 EGFR과 함께 빈번하게 공동 발현된다. 더욱이, EGFRvIII 발현 세포는 증가된 증식, 침입, 혈관신생 및 세포자멸사에 대한 내성을 나타낸다.

[0506] EGFRvIII은 다형성 교모세포종(GBM)에서 가장 흔히 발견된다. GBM의 25-35%는 이 절두된 수용체를 보유하는 것으로 추정된다. 더욱이, 이의 발현은 종종 보다 공격적인 표현형 및 나쁜 예후를 반영한다. GBM 이외에도, EGFRvIII의 발현은 또한 비-소세포 폐암, 두경부암, 유방암, 난소암 및 전립선암과 같은 다른 고형 종양에서 보고되었다. 대조적으로, EGFRvIII은 건강한 조직에서 발현되지 않는다.

[0507] 특정 구현예에서, 표적화된 암 항원 에피토프는 표적화된 암 세포 또는 종양에 의한 높은 발현 또는 표적화된 암 세포 또는 종양에 의한 낮은 발현을 가질 수 있다. 특정 구현예에서, 높은 발현 및 낮은 발현은 유세포 분석 또는 형광-활성화 세포-분류(FACS)를 사용하여 결정될 수 있다. 유세포 분석의 당업자에 의해 이해되는 바

와 같이, "hi", "lo", "+" 및 "-"는 음성 또는 다른 집단에 대한 신호 강도를 지칭한다. 특정 구현예에서, 양성 발현(+)은 마커가 유세포 분석을 사용하여 세포 상에서 검출가능함을 의미한다. 특정 구현예에서, 음성 발현(-)은 마커가 유세포 분석을 사용하여 검출가능하지 않음을 의미한다. 특정 구현예에서, "hi"는 관심 마커의 양성 발현이 형광(예를 들어 FACS 사용)에 의해 측정될 때 발현에 대해 또한 양성인 다른 세포보다 더 밝음을 의미한다. 이러한 구현예에서, 당업자는 밝기가 검출 임계값을 기반으로 함을 인식한다. 일반적으로, 당업자는 먼저 음성 대조군 튜브를 분석하고, FSC 및 SSC에 의해 관심 집단 주위에 게이트(비트맵)를 설정하고 원하는 방출 파장에서 형광에 대한 광전자 증배관 전압 및 이득을 조정하여, 세포의 97%가 음성 대조군이 있는 형광 마커에 대해 염색되지 않은 것으로 보이도록 할 것이다. 일단 이러한 매개변수가 확립되면, 염색된 세포를 분석하고, 염색되지 않은 형광 세포 집단과 비교하여 형광을 기록하였다. 특정 구현예, 및 대표적인 전형적인 FACS 플롯에서, hi는 맨 오른쪽(x 라인) 또는 가장 높은 상단 라인(오른쪽 또는 왼쪽 상단)을 의미하는 반면 lo는 왼쪽 하단 사분면 내 또는 오른쪽 및 왼쪽 사분면 사이의 중간을 의미한다(그러나 음성 집단에 비해 이동됨). 특정 구현예에서, "hi"는 +의 20-배 초과, +의 30-배 초과, +의 40-배 초과, +의 50-배 초과, +의 60-배 초과, +의 70-배 초과, +의 80-배 초과, +의 90-배 초과, +의 100-배 초과, 또는 + 세포에 비해 검출가능한 형광의 증가 이상을 지칭한다. 반대로, "lo"는 "hi"로 정의된 것들의 상호간의 집단을 지칭할 수 있다.

[0508] **II(E). 다른 표적**

[0509] HSC, T 세포, B 세포, 및 종양(또는 암 세포) 외에도, 벡터는 박테리아 및 진균에 대한 다른 항원을 표적화할 수 있다.

[0510] 박테리아를 표적화하는 항원은 예를 들어, 탄저병, 그람-음성균, 클라미디아, 디프테리아, 헬리코박터 파일로리, 결핵균, 백일해 독소, 폐렴구균, 리케차, 포도상구균, 연쇄상구균 및 파상풍으로부터 유래될 수 있다.

[0511] 박테리아 항원 마커의 특정 예로서, 탄저병 항원은 탄저병 보호 항원을 포함하고; 그람-음성균 항원은 지질다당류를 포함하고; 디프테리아 항원은 디프테리아 독소를 포함하고; 결핵균 항원은 마이콜산, 열 충격 단백질 65(HSP65), 30 kDa 주요 분비 단백질 및 항원 85A를 포함하고; 백일해 독소 항원은 헤마글루티닌, 페르탁틴, FIM2, FIM3 및 아데닐레이트 사이클라제를 포함하고; 폐렴쌍구균 항원은 폐렴구균용혈소 및 폐렴쌍구균 캡슐 다당류를 포함하고; 리케차 항원은 rompA를 포함하고; 연쇄상구균 항원은 M 단백질을 포함하고; 파상풍 항원은 파상풍 독소를 포함한다.

[0512] 진균을 표적화하는 항원은 예를 들어, 칸디다, 콕시디오데스, 크립토코쿠스, 히스토플라스마, 리슈마니어, 말라리아원충, 원생동물, 기생체, 주혈흡충속, 백선, 톡소플라스마, 및 트리파노소마 크루지(*Trypanosoma cruzi*)로부터 유래될 수 있다.

[0513] 진균 항원의 특정 예로서, 콕시디오데스 항원은 소구체 항원을 포함하고; 크립토코쿠스 항원은 캡슐 다당류를 포함하고; 히스토플라스마 항원은 열 충격 단백질 60(HSP60)을 포함하고; 리슈마니어 항원은 gp63 및 지질포스포글리칸을 포함하고; 열대열원충 항원은 낭충 표면 항원, 종충 표면 항원, 포자소체 항원, 생식모세포/배우자 표면 항원, 원생동물 및 혈액-단계 항원 pf 155/RESA를 포함하는 다른 기생체 항원을 포함하고; 주혈흡충속 항원은 글루타티온-S-트랜스퍼라제 및 파라마이오신을 포함하고; 백선 진균 항원은 백선균을 포함하고; 톡소플라스마 항원은 SAG-1 및 p30을 포함하고; 트리파노소마 크루지 항원은 75-77 kDa 항원 및 56 kDa 항원을 포함한다.

[0514] **III. 투여량, 제형, 및 투여**

[0515] 벡터는 세포 또는 동물, 예를 들어, 인간에게 투여하기 위해 약제학적으로 허용되도록 제형화될 수 있다. 벡터는 시험관내, 생체외, 또는 생체내 투여될 수 있다. 본원에 기재된 Ad35 바이러스 벡터 벡터들은 대상체에게 투여하기 위해 제형화될 수 있다. 제형은 치료 유전자("활성 성분") 및 하나 이상의 약제학적으로 허용되는 담체와 연관된 Ad35 바이러스 벡터를 포함한다.

[0516] 본원에 개시된 바와 같이, 벡터는 당업계에 알려진 임의의 형태일 수 있다. 이러한 형태는 예를 들어, 액체, 반고체 및 고체 투여 형태, 예컨대 액체 용액(예를 들어, 주사가능 및 주입가능 용액), 분산액 또는 현탁액, 정제, 알약, 분말, 리포솜 및 좌제를 포함한다.

[0517] 임의의 특정 형태의 선택 또는 사용은 부분적으로 의도된 투여 방식 및 치료 적용에 따라 달라질 수 있다. 예를 들어, 전신 또는 국소 전달을 위해 의도된 조성물을 함유하는 조성물은 주사가능 또는 주입가능 용액의 형태일 수 있다. 따라서, 벡터는 비경구 방식(예를 들어, 정맥내, 피하, 복강내, 또는 근육내 주사)에 의한 투여를 위

해 제형화될 수 있다. 본원에 사용된 바와 같이, 비경구 투여는 일반적으로 주사에 의한 장관 및 국소 투여 이외의 투여 방식을 지칭하며, 정맥내, 비강내, 안내, 폐, 근육내, 동맥내, 척추강내, 피막내, 안와내, 심장내, 피내, 폐내, 복강내, 기관경유, 피하, 표피하, 관절내, 피막하, 지주막하, 척수내, 경막외, 대뇌내, 두개내, 경동맥내 및 수조내 주사 및 주입을 포함하나 이에 제한되지 않는다. 비경구 투여 경로는 예를 들어, 주사에 의한 투여, 경비 투여, 경폐 투여, 또는 경피 투여일 수 있다. 투여는 정맥내 주사, 근육내 주사, 복강내 주사, 피하 주사에 의한 전신 또는 국소 투여일 수 있다.

[0518] 다양한 구현예에서, 본 발명의 벡터는 용액, 마이크로에멀젼, 분산액, 리포솜, 또는 고농도에서 안정된 저장에 적합한 다른 정렬된 구조로 제형화될 수 있다. 멸균 주사가능 용액은 필요에 따라, 상기 열거된 성분 중 하나 또는 조합과 함께 적절한 용매에 본원에 기재된 조성물을 필요한 양으로 혼입한 후, 필터 멸균함으로써 제조될 수 있다. 일반적으로, 분산액은 본원에 기재된 조성물을 염기성 분산 매질 및 상기 열거된 것들로부터 필요한 다른 성분을 함유하는 멸균 비히클에 혼입함으로써 제조된다. 멸균 주사가능 용액의 제조를 위한 멸균 분말의 경우, 제조 방법은 본원에 기재된 조성물의 분말 및 이의 이전에 멸균 여과된 용액으로부터 임의의 추가적인 바람직한 성분(하기 참조)를 산출하는 진공 건조 및 동결 건조를 포함한다. 용액의 적절한 유동성은 예를 들어, 레시틴과 같은 코팅의 사용, 분산액의 경우 필요한 입자 크기의 유지 및 계면활성제의 사용에 의해 유지될 수 있다. 주사가능 조성물의 연장된 흡수는 흡수를 지연시키는 시약, 예를 들어, 모노스테아레이트 염, 및 젤라틴을 조성물에 포함함으로써 야기될 수 있다.

[0519] 벡터는 멸균 용액 또는 물 중 현탁액 또는 또 다른 약제학적으로 허용되는 액체를 포함하는 주사가능 제형 형태로 비경구로 투여될 수 있다. 예를 들어, 벡터는 치료 분자를 약제학적으로 허용되는 비히클 또는 매질, 예컨대 멸균수 및 생리 식염수, 식물성 오일, 유화제, 현탁제, 계면활성제, 안정화제, 향미 부형제, 희석제, 비히클, 보존제, 결합제와 적절하게 조합한 후, 일반적으로 허용되는 약제학적 관행에 필요한 단위 용량 형태로 혼합함으로써 제형화될 수 있다. 약제학적 제제에 포함되는 벡터의 양은 지정된 범위 내에서 적합한 용량이 제공되도록 한다. 유성 액체의 비제한적인 예는 참깨유 및 대두유를 포함하고, 가용화 제제로서 벤질 벤조에이트 또는 벤질 알코올과 조합될 수 있다. 포함될 수 있는 다른 항목은 포스페이트 완충제, 또는 나트륨 아세테이트 완충제와 같은 완충제, 프로카인 하이드로클로라이드와 같은 진정제, 젠질 알코올 또는 페놀과 같은 안정화제, 및 산화방지제이다. 제형화된 주사는 적합한 앰플에 패키징될 수 있다.

[0520] 다양한 구현예에서, 피하 투여는 주사기, 미리 충전된 주사기, 자동 주사기(예를 들어, 일회용 또는 재사용가능), 펜 주사기, 패치 주사기, 착용가능한 주사기, 피하 주입 세트가 장착된 보행용 주사기 주입 펌프와 같은 장치, 또는 피하 주사를 위한 다른 장치에 의해 달성될 수 있다.

[0521] 일부 구현예에서, 본원에 기재된 벡터는 국소 투여 방식에 의해 대상체에게 치료적으로 전달될 수 있다. 본원에 사용된 바와 같이, "국소 투여" 또는 "국소 전달"은 혈관계를 통해 의도된 표적 조직 또는 부위에 벡터 또는 벡터를 수송하는 것에 의존하지 않는 전달을 지칭할 수 있다. 예를 들어, 벡터는 조성물 또는 제제의 주사 또는 이식 또는 조성물 또는 제제를 함유하는 장치의 주사 또는 이식에 의해 전달될 수 있다. 특정 구현예에서, 표적 조직 또는 부위 부근에 국소 투여 후, 조성물 또는 제제, 또는 이의 하나 이상의 구성요소는 투여 부위가 아닌 의도된 표적 조직 또는 부위로 확산될 수 있다.

[0522] 일부 구현예에서, 본원에 제공된 조성물은 단위 투여 형태로 존재하며, 단위 투여 형태는 자기-투여에 적합할 수 있다. 이러한 단위 투여 형태는 용기, 전형적으로, 예를 들어, 바이알, 카트리지, 미리 충전된 주사기 또는 일회용 펜 내에 제공될 수 있다. US 6,302,855에 기재된 투약기 장치와 같은 투약기는 또한 예를 들어, 본원에 기재된 바와 같은 주사 시스템과 함께 사용될 수 있다.

[0523] 주사에 적합한 벡터 제형의 약제학적 형태는 멸균 수용액 또는 분산액을 포함할 수 있다. 제형은 멸균될 수 있고 주사기의 안팎으로 적절한 흐름을 허용하기 위해 유체여야 한다. 제형은 또한 제조 및 저장 조건 하에 안정될 수 있다. 담체는 예를 들어, 물 및 염수 또는 완충 수용액을 함유하는 용매 또는 분산 매질일 수 있다. 바람직하게는, 등장성 제제, 예를 들어, 당 또는 염화나트륨이 제형에 사용될 수 있다.

[0524] 게다가, 당업자는 또한 추가적인 전달 방법이 전기천공법, 초음파영동법, 골내 주사 방법을 통하거나 또는 유전자 총을 사용하는 것임을 고려할 수 있다. 벡터는 또한 마이크로칩, 나노-칩 또는 나노입자에 이식될 수 있다.

[0525] 본원에 기재된 벡터의 적합한 용량은 예를 들어, 치료될 대상체의 연령, 성별, 및 체중, 치료될 병태 또는 질환, 및 사용되는 특정 벡터를 포함하는 다양한 인자에 따라 달라질 수 있다. 대상체에게 투여되는 용량에 영향을 미치는 다른 인자는 예를 들어, 병태 또는 질환의 유형 또는 중증도를 포함한다. 다른 인자는 예를 들어,

대상체에게 동시에 또는 이전에 영향을 미치는 다른 의학적 장애, 대상체의 일반적인 건강, 대상체의 유전적 소인, 식습관, 투여 시간, 배출률, 약물 조합, 및 대상체에게 투여되는 임의의 다른 추가적인 치료제를 포함할 수 있다. 벡터 투여의 적합한 수단은 치료될 병태 또는 질환에 기반하고 대상체의 연령 및 병태에 따라 선택될 수 있다. 투여 용량 및 방법은 환자의 체중, 연령, 병태 등에 따라 달라질 수 있고, 당업자에 의해 필요에 따라 적합하게 선택될 수 있다. 임의의 특정 대상체에 대한 특이적 투여량 및 치료 레지멘은 의사의 판단에 따라 조정될 수 있다.

[0526] 벡터 용액은 치료 유효량의 본원에 기재된 조성물을 포함할 수 있다. 이러한 유효량은 부분적으로, 투여되는 조성물의 효과, 또는 하나 초과의 제제의 사용되는 경우 조성물 및 하나 이상의 추가적인 활성 제제의 조합 효과에 따라 당업자에 의해 용이하게 결정될 수 있다. 치료 유효량은 조성물의 임의의 독소 또는 유해 효과가 치료적으로 유익한 효과를 능가하는 더 큰 양일 수 있다.

[0527] 다양한 경우에, 벡터는 약제학적으로 허용되는 담체 또는 부형제를 포함하도록 제형화될 수 있다. 약제학적으로 허용되는 담체의 예는 생리학적으로 호환가능한 임의의 및 모든 용매, 분산 매질, 코팅, 항균제 및 항진균제, 등장성 및 흡수 지연제 등을 포함하나 이에 제한되지 않는다. 본 발명의 조성물은 약제학적으로 허용되는 염, 예를 들어, 산 부가 염 또는 염기 부가 염을 포함할 수 있다.

[0528] 예시적인 일반적으로 사용되는 약제학적으로 허용되는 담체는 임의의 및 모든 흡수 지연제, 산화방지제, 결합제, 완충제, 증강제 또는 충전제, 킬레이트화제, 코팅제, 붕해제, 분산 매질, 겔, 등장성제, 윤활제, 보존제, 염, 용매 또는 공용매, 안정화제, 계면활성제, 및/또는 전달 비히클을 포함한다.

[0529] 다양한 구현예에서, 본원에 기재된 바와 같은 벡터를 포함하는 조성물, 예를 들어, 주사용 멸균 제형은 주사용 증류수를 비히클로 사용하여 통상적인 약제학적 관행에 따라 제형화될 수 있다. 예를 들어, 생리 식염수 또는 글루코스 및 다른 보충물 예컨대 D-소르비톨, D-만노스, D-만니톨, 및 염화나트륨을 함유하는 등장성 용액은 임의적으로 적합한 가용화제, 예를 들어, 에탄올과 같은 알코올 및 프로필렌 글리콜 또는 폴리에틸렌 글리콜과 같은 폴리알코올, 및 폴리소르베이트 80™, HCO-50 등과 같은 비이온성 계면활성제와 조합하여, 주사용 수용액으로 사용될 수 있다.

[0530] 예시적인 산화방지제는 아스코르브산, 메티오닌, 및 비타민 E를 포함한다.

[0531] 예시적인 완충제는 시트레이트 완충제, 숙시네이트 완충제, 타르트레이트 완충제, 푸마레이트 완충제, 글루코네이트 완충제, 옥살레이트 완충제, 락테이트 완충제, 아세테이트 완충제, 포스페이트 완충제, 히스티딘 완충제, 및/또는 트리메틸아민 염을 포함한다.

[0532] 예시적인 킬레이트화제는 EDTA이다.

[0533] 예시적인 등장성제는 3가 또는 그 이상의 당을 포함하는 다가 당 알코올, 예컨대 글리세린, 에리트리톨, 아라비톨, 크실리톨, 소르비톨, 또는 만니톨을 포함한다.

[0534] 예시적인 보존제는 페놀, 벤질 알코올, 메타-크레졸, 메틸 파라벤, 프로필 파라벤, 옥타데실디메틸벤질 암모늄 클로라이드, 벤잘코늄 할라이드, 헥사메토늄 클로라이드, 알킬 파라벤 예컨대 메틸 또는 프로필 파라벤, 카테콜, 레소르시놀, 사이클로헥사놀, 및 3-펜타놀을 포함한다.

[0535] 안정화제는 증강제부터 활성 성분을 가용화하거나 또는 변성 또는 용기 벽에 부착을 방지하는 데 도움이 되는 첨가제까지 다양한 기능을 할 수 있는 광범위한 범주의 부형제를 지칭한다. 전형적인 안정화제는 다가 당 알코올; 아미노산, 예컨대 알라닌, 리신, 글리신, 글루타민, 아스파라긴, 히스티딘, 알라닌, 오르니틴, L-류신, 2-페닐알라닌, 글루탐산, 및 트레오닌; 유기 당 또는 당 알코올, 예컨대 락토스, 트레할로스, 스타키오스, 만니톨, 소르비톨, 크실리톨, 리비톨, 마이오이니시톨, 갈락티톨, 글리세롤, 및 사이클리톨, 예컨대 이노시톨; PEG; 아미노산 중합체; 황-함유 환원제, 예컨대 우레아, 글루타티온, 티옥트산, 나트륨 티오글리콜레이트, 티오글리세롤, α-모노티오글리세롤, 및 나트륨 티오술페이트; 저분자량 폴리펩티드(즉, 10 개 미만의 잔기); 단백질 예컨대 인간 혈청 알부민, 소 혈청 알부민, 젤라틴 또는 면역글로불린; 친수성 중합체 예컨대 폴리비닐피롤리돈; 단당류 예컨대 크실로스, 만노스, 프럭토스 및 글루코스; 이당류 예컨대 락토스, 말토스 및 수크로스; 삼당류 예컨대 라피노스, 및 다당류 예컨대 덱스트란을 포함할 수 있다. 안정화제는 전형적으로 치료 중량을 기준으로 0.1 내지 10,000 중량부의 범위로 존재한다.

[0536] 본원에 개시된 제형은 예를 들어, 주사에 의한 투여를 위해 제형화될 수 있다. 주사를 위해, 제형은 수용액으로, 예컨대 행크스 용액, 링거 용액, 또는 생리 식염수를 포함하는 완충액, 또는 이스코브의 변형된 듀벨코 배

지(IMDM)와 같은 배양 배지에서 제형화될 수 있다. 수용액은 현탁제, 안정화제, 및/또는 분산제와 같은 제형화제를 포함할 수 있다. 대안적으로, 제형은 사용 전에 적합한 비히클, 예를 들어, 멸균 무발열원수로 구성하기 위한 동결건조 및/또는 분말 형태일 수 있다.

[0537] 본원에 개시된 임의의 제형은 유리하게는 투여 이익을 능가하는 상당히 유해하거나, 알레르기성이거나, 또는 다른 불리한 반응을 생성하지 않는 것들을 포함하는 임의의 다른 약제학적으로 허용되는 담체를 포함할 수 있다. 예시적인 약제학적으로 허용되는 담체 및 제형은 Remington's Pharmaceutical Sciences, 18th Ed. Mack Printing Company, 1990에 기재되어 있다. 더욱이, 제형은 US FDA 생물학적 표준 사무국(Office of Biological Standards) 및/또는 다른 관련한 외국 규제 기관에서 요구하는 멸균성, 발열성, 일반적인 안전성, 및 순도 표준을 충족시키도록 제조될 수 있다.

[0538] 특정 구현예에서, 제형은 제형의 적어도 0.1% w/v 또는 w/w; 제형의 적어도 1% w/v 또는 w/w; 제형의 적어도 10% w/v 또는 w/w; 제형의 적어도 20% w/v 또는 w/w; 제형의 적어도 30% w/v 또는 w/w; 제형의 적어도 40% w/v 또는 w/w; 제형의 적어도 50% w/v 또는 w/w; 제형의 적어도 60% w/v 또는 w/w; 제형의 적어도 70% w/v 또는 w/w; 제형의 적어도 80% w/v 또는 w/w; 제형의 적어도 90% w/v 또는 w/w; 제형의 적어도 95% w/v 또는 w/w; 또는 제형의 적어도 99% w/v 또는 w/w의 활성 성분을 포함한다.

[0539] 특정 대상체에게 투여되는 Ad35 바이러스 벡터 및, 특정 구현예에서, Ad35 바이러스 벡터 및 동원 인자의 실제 용량 및 양 및 일치하는 동원 절차 및 일정은 예를 들어, 표적; 체중; 병태 유형; 병태 중증도; 알려진 경우 예정된 관련 사건; 이전 또는 동시 치료 개입; 대상체의 특발성; 및 투여 경로를 포함하는 물리적 및 생리학적 인자와 같은 매개변수를 고려하여 의사, 수의사, 또는 연구자에 의해 결정될 수 있다. 또한, 시험관내 및 생체내 검정은 임의적으로 최적의 투여량 범위를 식별하는 것을 돕기 위해 이용될 수 있다.

[0540] 치료 유전자와 연관된 Ad35 벡터의 치료 유효량은 예를 들어, $1 \times 10^7$ 내지 $50 \times 10^8$ 감염 단위(IU) 또는 $5 \times 10^7$ 내지 $20 \times 10^8$ IU 범위의 용량을 포함할 수 있다. 다른 예에서, 용량은 $5 \times 10^7$ IU, $6 \times 10^7$ IU, $7 \times 10^7$ IU, $8 \times 10^7$ IU, $9 \times 10^7$ IU, $1 \times 10^8$ IU, $2 \times 10^8$ IU, $3 \times 10^8$ IU, $4 \times 10^8$ IU, $5 \times 10^8$ IU, $6 \times 10^8$ IU, $7 \times 10^8$ IU, $8 \times 10^8$ IU, $9 \times 10^8$ IU, $10 \times 10^8$ IU, 또는 그 이상을 포함할 수 있다. 특정 구현예에서, 치료 유전자와 연관된 Ad35 벡터의 치료 유효량은 $4 \times 10^8$ IU를 포함한다. 특정 구현예에서, 치료 유전자와 연관된 Ad35 벡터의 치료 유효량은 피하로 또는 정맥내로 투여될 수 있다. 특정 구현예에서, 치료 유전자와 연관된 Ad35 벡터의 치료 유효량은 하나 이상의 동원 인자와 함께 투여 후 투여될 수 있다.

[0541] 본 개시내용의 다양한 구현예에서, 생체내 유전자 요법은 적어도 하나의 바이러스 유전자 요법 벡터를 적어도 하나의 면역 억제 레지멘과 조합하여 대상체에게 투여하는 것을 포함한다. 지지 벡터인 제2 벡터와 조합된 지지된 바이러스 유전자 요법 벡터인 제1 벡터와 같은 하나 초과의 벡터 종을 포함하는 생체내 유전자 요법에서, 제1 벡터 및 제2 벡터는 단일 제형 또는 투여 형태 또는 2 개의 별개의 제형 또는 투여 형태로 투여될 수 있다. 다양한 구현예에서, 제1 및 제2 벡터는 동시에 또는 상이한 시간에, 예를 들어, 동일한 1 시간 기간 동안 또는 중첩하지 않는 1 시간 기간 동안 투여될 수 있다. 다양한 구현예에서, 제1 및 제2 벡터는 동시에 또는 상이한 시간에, 예를 들어, 동일한 날 또는 상이한 날에 투여될 수 있다. 다양한 구현예에서, 제1 및 제2 벡터는 예를 들어, 투여량이 총 바이러스 입자 수 또는 대상체의 킬로그램 당 바이러스 입자의 수로 측정되는 경우 동일한 투여량 또는 상이한 투여량으로 투여될 수 있다. 다양한 구현예에서, 제1 및 제2 벡터는 미리 정의된 비율로 투여될 수 있다. 다양한 구현예에서, 비율은 2:1 내지 1:2 범위 내, 예를 들어, 1:1이다.

[0542] 다양한 구현예에서, 벡터는 하루에 단일 총 용량으로 대상체에게 투여된다. 다양한 구현예에서 벡터는 함께 총 용량을 구축하는 2, 3, 4, 또는 그 이상의 단위 용량으로 투여된다. 다양한 구현예에서, 벡터의 1 단위 용량은 1 일, 2 일, 3 일, 4 일, 또는 그 이상의 연속 일 각각에 하루에 대상체에게 투여된다. 다양한 구현예에서, 벡터의 2 단위 용량은 1 일, 2 일, 3 일, 4 일, 또는 그 이상의 연속 일 각각에 하루에 대상체에게 투여된다. 따라서, 다양한 구현예에서, 일일 용량은 하루에 걸쳐 대상체가 받는 벡터의 용량을 지칭할 수 있다. 다양한 구현예에서, 용어 일은 첫번째 달력 날짜의 자정부터 다음 달력 날짜의 자정까지 24 시간 기간과 같은 24 시 기간을 지칭한다.

[0543] 다양한 구현예에서, 바이러스 유전자 요법 벡터 또는 지지 벡터와 같은 벡터의 단위 용량, 일일 용량, 또는 총 용량, 또는 바이러스 유전자 요법 벡터 및 지지 벡터의 총 조합 용량은 킬로그램 당 적어도 1E8, 5E8, 1E9, 5E9, 1E10, 5E10, 1E11, 5E11, 1E12, 5E12, 1E13, 5E13, 1E14, 또는 1E15 바이러스 입자(vp/kg)일 수 있다.

다양한 구현예에서, 바이러스 유전자 요법 벡터 또는 지지 벡터와 같은 벡터의 단위 용량, 일일 용량, 또는 총 용량, 또는 바이러스 유전자 요법 벡터 및 지지 벡터의 총 조합 용량은 1E8, 5E8, 1E9, 5E9, 1E10, 5E10, 1E11, 5E11, 1E12, 5E12, 1E13, 5E13, 1E14, 또는 1E15 vp/kg으로부터 선택된 하한 및 1E8, 5E8, 1E9, 5E9, 1E10, 5E10, 1E11, 5E11, 1E12, 5E12, 1E13, 5E13, 1E14, 또는 1E15 vp/kg으로부터 선택된 상한을 갖는 범위 내에 속할 수 있다.

[0544] 다양한 구현예에서, 바이러스 유전자 요법 벡터는 적어도 1E10, 5E10, 1E11, 5E11, 1E12, 5E12, 1E13, 5E13, 1E14, 또는 1E15 vp/kg의 단위 용량, 일일 용량, 또는 총 용량으로 투여되고 지지 벡터는 적어도 1E8, 5E8, 1E9, 5E9, 1E10, 5E10, 1E11, 및 5E11 vp/kg의 단위 용량, 일일 용량, 또는 총 용량으로 투여되며, 임의적으로 여기서 바이러스 유전자 요법 벡터의 단위 용량, 일일 용량, 또는 총 용량은 1E10, 5E10, 1E11, 5E11, 1E12, 및 5E12, vp/kg으로부터 선택된 하한 및 1E11, 5E11, 1E12, 5E12, 1E13, 5E13, 1E14, 및 1E15 vp/kg으로부터 선택된 상한을 갖는 범위 내에 있고/있거나, 지지 벡터의 단위 용량, 일일 용량, 또는 총 용량은 1E8, 5E8, 1E9, 5E9, 1E10, 및 5E10 vp/kg으로부터 선택된 하한 및 1E9, 5E9, 1E10, 5E10, 1E11, 및 5E11 vp/kg으로부터 선택된 상한을 갖는 범위 내에 있다.

[0545] 다양한 구현예에서, 지지 벡터는 적어도 1E10, 5E10, 1E11, 5E11, 1E12, 5E12, 1E13, 5E13, 1E14, 또는 1E15 vp/kg의 단위 용량, 일일 용량, 또는 총 용량으로 투여되고 지지된 바이러스 유전자 요법 벡터는 적어도 1E8, 5E8, 1E9, 5E9, 1E10, 5E10, 1E11, 및 5E11 vp/kg의 단위 용량, 일일 용량, 또는 총 용량으로 투여되며, 임의적으로 여기서 지지 벡터의 단위 용량, 일일 용량, 또는 총 용량은 1E10, 5E10, 1E11, 5E11, 1E12, 및 5E12, vp/kg으로부터 선택된 하한 및 1E11, 5E11, 1E12, 5E12, 1E13, 5E13, 1E14, 및 1E15 vp/kg으로부터 선택된 상한을 갖는 범위 내에 있고/있거나, 지지된 바이러스 유전자 요법 벡터의 단위 용량, 일일 용량, 또는 총 용량은 1E8, 5E8, 1E9, 5E9, 1E10, 및 5E10 vp/kg으로부터 선택된 하한 및 1E9, 5E9, 1E10, 5E10, 1E11, 및 5E11 vp/kg으로부터 선택된 상한을 갖는 범위 내에 있다. 다양한 구현예에서, 지지된 바이러스 유전자 요법 벡터 및 지지 벡터는 미리 정의된 비율로 투여된다. 다양한 구현예에서, 비율은 2:1 내지 1:2 범위 내, 예를 들어, 1:1 이다.

[0546] **IV. 적용**

[0547] 본원에 제공된 방법 및 조성물은 적어도 부분적으로 생체내 유전자 요법에 사용하기 위해 개시된다. 그러나, 의심의 여지를 피하기 위해, 본 개시내용은 세포 및/또는 조직의 생체외 조작, 뿐만 아니라 연구 목적을 위한 세포 및/또는 조직의 조작을 포함하는 시험관내 사용을 위해 본원에 제공된 조성물 및 방법의 사용을 명시적으로 포함한다. 유전자 요법은 외인성 DNA를 숙주 세포(예컨대 표적 세포) 및/또는 핵산(예컨대 표적 핵산, 예컨대 표적 게놈, 예를 들어, 표적 세포의 게놈) 내로 도입하는 방법에서 본 개시내용의 벡터, 게놈, 또는 시스템의 사용을 포함한다. 본 개시내용은 당업자가 본원에 제공된 다양한 방법 및 조성물이 일반적으로 핵산 페이로드를 대상체, 예를 들어, 숙주 또는 표적 세포 내로 도입하도록 적용가능함을 이해하는 생체내, 시험관내, 및 생체외 요법과 관련한 조성물 및 방법의 설명 및 예시를 포함한다. 이러한 조성물 및 방법은 예를 들어, 유전자 요법에서 일반적으로 유용하기 때문에, 일반적으로 유전자 요법 및 본원에 제공된 것들을 포함하는 특히 다양한 특정 조건에서 모두 도구로서 유용하다.

[0548] **IV(A). 생체내 유전자 요법**

[0549] 바이러스 벡터를 환자에게 직접 전달하는 것을 포함하는 생체내 유전자 요법을 사용한 치료가 탐구되었다. 생체내 유전자 요법은 임의의 유전자독성 조건화(또는 유전자독성 조건화가 덜 요구될 수 있거나) 또는 생체외 세포 처리를 필요로 하지 않을 수 있기 때문에 매력적인 접근법이며 따라서 백신의 전달을 위해 이미 전 세계적으로 수행되는 것과 유사한 주사를 통해 요법이 투여될 수 있으므로, 개발도상국을 포함하여 전세계적으로 많은 기관에서 채택될 수 있다. 다양한 구현예에서 본 개시내용의 아데노바이러스 벡터를 사용한 생체내 유전자 요법 방법은 (i) 표적 세포 동원, (ii) 면역억제, (iii) 본원에 제공된 벡터, 게놈, 시스템 또는 제형의 투여, 및/또는 (iv) 형질도입된 세포 및/또는 아데노바이러스 벡터 또는 게놈의 페이로드의 통합요소를 통합한 세포의 선택 중 하나 이상의 단계를 포함할 수 있다.

[0550] 본원에 개시된 아데노바이러스 벡터 제형은 대상체(인간, 수의학 동물(개, 고양이, 파충류, 조류 등), 가축(말, 소, 염소, 돼지, 닭 등), 및 연구 동물(원숭이, 래트, 마우스, 어류 등)을 치료하기 위해 사용될 수 있다. 대상체를 치료하는 것은 치료 유효량의 본 개시내용의 하나 이상의 벡터, 게놈, 또는 시스템을 전달하는 것을 포함한다. 치료 유효량은 유효량, 예방적 치료, 및/또는 치료적 치료를 제공하는 것을 포함한다.

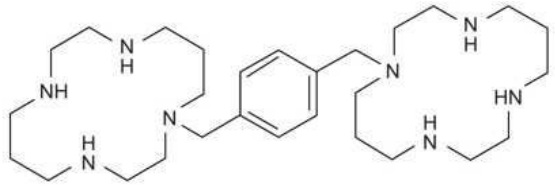
[0551] **IV(A)i. HSC의 동원**

[0552] 본원에 기재된 벡터는 동원 인자와 함께 투여될 수 있다. 특정 구현예에서, 본원에 기재된 아데노바이러스 벡터 제형은 HSPC 동원과 협력하여 투여될 수 있다. 특정 구현예에서, 아데노바이러스 공여자 벡터의 투여는 하나 이상의 동원 인자의 투여와 동시에 발생한다. 특정 구현예에서, 아데노바이러스 공여자 벡터의 투여는 하나 이상의 동원 인자의 투여 후에 수행된다. 특정 구현예에서, 아데노바이러스 공여자 벡터의 투여는 첫번째 하나 이상의 동원 인자의 투여 후에 수행되고 두번째 하나 이상의 동원 인자의 투여와 동시에 발생한다. HSPC 동원을 위한 제제는 예를 들어, 과립구-콜로니 자극 인자(G-CSF), 과립구 대식세포 콜로니 자극 인자(GM-CSF), AMD3100, SCF, S-CSF, CXCR4 길항제, CXCR2 작용제, 및 Gro-베타(GRO-β)를 포함한다. 다양한 구현예에서, CXCR4 길항제는 AMD3100이고/이거나 CXCR2 작용제는 GRO-β이다.

[0553] G-CSF는 HSPC 동원에서의 기능이 과립구 확장의 촉진 및 접착 분자의 프로테아제-의존적 및 독립적 감쇠 및 SDF-1/CXCR4 축의 파괴 둘 다를 포함할 수 있는 사이토카인이다. 특정 구현예에서, 당업자에게 알려진 임의의 상업적으로 이용가능한 형태의 G-CSF, 예를 들어, 필그라스팀(Filgrastim)(Neupogen®, Amgen Inc., 캘리포니아주 사우전드 오크스 소재) 및 PEGy길화 필그라스팀(페그필그라스팀(Pegfilgrastim), NEULASTA®, Amgen Inc., 캘리포니아주 사우전드 오크스 소재)이 본원에 개시된 바와 같은 방법 및 제형에 사용될 수 있다.

[0554] GM-CSF는 사이토카인으로 기능하며 대식세포, T 세포, 비만 세포, 자연 살해 세포, 내피 세포, 및 섬유모세포에 의해 자연적으로 분비되는 콜로니-자극 인자 2(CSF2)로도 알려진 단량체성 당단백질이다. 특정 구현예에서, 당업자에게 알려진 임의의 상업적으로 이용가능한 형태의 GM-CSF, 예를 들어, 사르그라모스팀(Sargramostim)(Leukine, Bayer Healthcare Pharmaceuticals, 워싱턴주 시애틀 소재) 및 몰그라모스팀(molgramostim)(Schering-Plough, 뉴저지주 케닐워스 소재)은 본원에 개시된 바와 같은 방법 및 제형에 사용될 수 있다.

[0555] 비사이클람 클래스의 합성 유기 분자인 AMD3100(MOZOBIL™, PLERIXAFOR™; Sanofi-Aventis, 프랑스 파리 소재)은 케모카인 수용체 길항제이며 CXCR4에 대한 SDF-1 결합을 가역적으로 억제하여, HSPC 동원을 촉진한다. AMD3100은 골수종 및 림프종 환자에서 HSPC 동원을 위해 G-CSF와 조합하여 사용되도록 승인되었다. AMD3100의 구조는 하기와 같다:

[0556] 

[0557] KIT 리간드, KL, 또는 강철 인자로도 알려져 있는 SCF는 c-kit 수용체(CD117)에 결합하는 사이토카인이다. SCF는 막관통 단백질 및 가용성 단백질 둘 다에 존재할 수 있다. 이 사이토카인은 조혈, 정자생성, 및 멜라닌생성에서 중요한 역할을 한다. 특정 구현예에서, 당업자에게 알려진 임의의 상업적으로 이용가능한 형태의 SCF, 예를 들어, 재조합 인간 SCF(안세스팀(Ancestim), STEMGEN®, Amgen Inc., 캘리포니아주 사우전드 오크스 소재)는 본원에 개시된 바와 같은 방법 및 제형에 사용될 수 있다.

[0558] 집중적인 골수억제 치료에 사용되는 화학요법은 또한 화학요법-유도된 무형성증에 따른 보상 호중구 생성의 결과로 HSPC를 말초 혈액으로 동원한다. 특정 구현예에서, HSPC의 동원을 위해 사용될 수 있는 화학치료제는 사이클로포스파미드, 에토포시드, 이포스파미드, 시스플라틴, 및 사이타라빈을 포함한다.

[0559] 세포 동원을 위해 사용될 수 있는 추가적인 제제는 다음을 포함한다: CXCL12/CXCR4 조절인자(예를 들어, CXCR4 길항제: POL6326(Polyphor, 스위스 알슈빌 소재), CXCR4를 가역적으로 억제하는 합성 환형 펩티드; BKT-140(4F-벤조일-TN14003; Biokine Therapeutics, 이스라엘 레호보트 소재); TG-0054(Taigen Biotechnology, 태국 타이베이 소재); SDF-1에 결합하여, CXCR4)에 대한 결합을 억제하는 CXCL12 중화제 NOX-A12(NOXXON Pharma, 독일 베를린 소재); 스핑고신-1-포스페이트(S1P) 작용제(예를 들어, SEW2871, Juarez 등 *Blood* 119: 707-716, 2012); 혈관 세포 접착 분자-1(VCAM) 또는 매우 후기 항원 4(VLA-4) 억제제(예를 들어, VLA-4의 α4 서브유닛에 대한 재조합 인간화 단클론 항체인 나탈리주맙(Zohren 등 *Blood* 111: 3893-3895, 2008); VLA-4의 소분자 억제제인 BIO5192(Ramirez 등 *Blood* 114: 1340-1343, 2009)); 파라갑상선 호르몬(Brunner 등 *Exp Hematol*. 36: 1157-1166, 2008); 프로테아좀 억제제(예를 들어, 보르테조밉, Ghobadi 등 *ASH Annual Meeting Abstracts*. p. 583, 2012); CXCR2 수용체에 대한 결합에 의해 호중구의 화학주성 및 활성화를 자극하는 CXC 케모카인 패밀리의

구성원인 Groβ(예를 들어, SB-251353, King 등 *Blood* 97: 1534-1542, 2001); 저산소증 유도성 인자(HIF)의 안정화(예를 들어, FG-4497, Forristal 등 ASH Annual Meeting Abstracts. p. 216, 2012); α4β1 및 α4β7 인테그린 억제제(α4β1/7)인 피라테그라스트(Kim 등 *Blood* 128: 2457-2461, 2016); α4β7 인테그린에 대한 인간화 단클론 항체인 베돌리주맙(Vedolizumab)(Rosario 등 *Clin Drug Investig* 36: 913-923, 2016); 및 인테그린 α9β1/α4β1을 표적화하는 BOP(N-(벤젠술포닐)-L-프롤릴-L-O-(1-피롤리디닐카르보닐) 티로신)(Cao 등 *Nat Commun* 7: 11007, 2016). HSPC 동원을 위해 사용될 수 있는 추가적인 제제는 예를 들어, Richter R 등 *Transfus Med Hemother* 44:151-164, 2017, Bendall & Bradstock, *Cytokine & Growth Factor Reviews* 25: 355-367, 2014, WO 2003043651, WO 2005017160, WO 2011069336, US 5,637,323, US 7,288,521, US 9,782,429, US 2002/0142462, 및 US 2010/02268에 기재되어 있다.

[0560] 특정 구현예에서, G-CSF의 치료 유효량은 0.1 μg/kg 내지 100 μg/kg을 포함한다. 특정 구현예에서, G-CSF의 치료 유효량은 0.5 μg/kg 내지 50 μg/kg을 포함한다. 특정 구현예에서, G-CSF의 치료 유효량은 0.5 μg/kg, 1 μg/kg, 2 μg/kg, 3 μg/kg, 4 μg/kg, 5 μg/kg, 6 μg/kg, 7 μg/kg, 8 μg/kg, 9 μg/kg, 10 μg/kg, 11 μg/kg, 12 μg/kg, 13 μg/kg, 14 μg/kg, 15 μg/kg, 16 μg/kg, 17 μg/kg, 18 μg/kg, 19 μg/kg, 20 μg/kg, 또는 그 이상을 포함한다. 특정 구현예에서, G-CSF의 치료 유효량은 5 μg/kg을 포함한다. 특정 구현예에서, G-CSF는 피하로 또는 정맥내로 투여될 수 있다. 특정 구현예에서, G-CSF는 1 일, 2 일 연속, 3 일 연속, 4 일 연속, 5 일 연속, 또는 그 이상 동안 투여될 수 있다. 특정 구현예에서, G-CSF는 4 일 연속 동안 투여될 수 있다. 특정 구현예에서, G-CSF는 5 일 연속 동안 투여될 수 있다. 특정 구현예에서, 단일 제제로서, G-CSF는 Ad35 전달 3, 4, 5, 6, 7, 또는 8 일 전에 시작하여, 매일 피하로 10 μg/kg의 용량으로 사용될 수 있다. 특정 구현예에서, G-CSF는 단일 제제로서 투여된 후 또 다른 동원 인자와 동시 투여될 수 있다. 특정 구현예에서, G-CSF는 단일 제제로 투여된 후 AMD3100과 동시 투여될 수 있다. 특정 구현예에서, 치료 프로토콜은 G-CSF가 1 일, 2 일, 3 일, 및 4 일 및 5 일에 투여될 수 있는 경우 5 일 치료를 포함하며, G-CSF 및 AMD3100은 Ad35 투여 6 내지 8 시간 전에 투여된다.

[0561] 투여될 GM-CSF의 치료 유효량은 예를 들어, 0.1 내지 50 μg/kg 또는 0.5 내지 30 μg/kg 범위의 용량을 포함할 수 있다. 특정 구현예에서, GM-CSF가 투여될 수 있는 용량은 0.5 μg/kg, 1 μg/kg, 2 μg/kg, 3 μg/kg, 4 μg/kg, 5 μg/kg, 6 μg/kg, 7 μg/kg, 8 μg/kg, 9 μg/kg, 10 μg/kg, 11 μg/kg, 12 μg/kg, 13 μg/kg, 14 μg/kg, 15 μg/kg, 16 μg/kg, 17 μg/kg, 18 μg/kg, 19 μg/kg, 20 μg/kg, 또는 그 이상을 포함한다. 특정 구현예에서, GM-CSF는 1 일, 2 일 연속, 3 일 연속, 4 일 연속, 5 일 연속, 또는 그 이상 동안 피하로 투여될 수 있다. 특정 구현예에서, GM-CSF는 피하로 또는 정맥내로 투여될 수 있다. 특정 구현예에서, GM-CSF는 Ad35 전달 3, 4, 5, 6, 7, 또는 8 일 전에 시작하여 매일 피하로 10 μg/kg의 용량으로 투여될 수 있다. 특정 구현예에서, GM-CSF는 단일 제제로서 투여된 후 또 다른 동원 인자와 통시 투여될 수 있다. 특정 구현예에서, GM-CSF는 단일 제제로서 투여된 후 AMD3100과 동시 투여될 수 있다. 특정 구현예에서, 치료 프로토콜은 GM-CSF가 1 일, 2 일, 3 일, 및 4 일 및 5 일에 투여될 수 있는 경우 5 일 치료를 포함하며, GM-CSF 및 AMD3100은 Ad35 투여 6 내지 8 시간 전에 투여된다. 사르그라모스팀에 대한 투약 레지멘은 200 $\mu g/m^2$, 210 $\mu g/m^2$, 220 $\mu g/m^2$, 230 $\mu g/m^2$, 240 $\mu g/m^2$, 250 $\mu g/m^2$, 260 $\mu g/m^2$, 270 $\mu g/m^2$, 280 $\mu g/m^2$, 290 $\mu g/m^2$, 300 $\mu g/m^2$, 또는 그 이상을 포함할 수 있다. 특정 구현예에서, 사르그라모스팀은 1 일, 2 일 연속, 3 일 연속, 4 일 연속, 5 일 연속, 또는 그 이상 동안 투여될 수 있다. 특정 구현예에서, 사르그라모스팀은 피하로 또는 정맥내로 투여될 수 있다. 특정 구현예에서, 사르그라모스팀에 대한 투약 레지멘은 250 $\mu g/m^2$/일 정맥내 또는 피하를 포함할 수 있고 표적화된 세포 양이 말초 혈액에서 도달할 때까지 계속될 수 있거나 또는 5 일 동안 계속될 수 있다. 특정 구현예에서, 사르그라모스팀은 단일 제제로서 투여된 후 또 다른 동원 인자와 동시 투여될 수 있다. 특정 구현예에서, 사르그라모스팀은 단일 제제로서 투여된 후 AMD3100과 동시 투여될 수 있다. 특정 구현예에서, 치료 프로토콜은 사르그라모스팀이 1 일, 2 일, 3 일, 및 4 일 및 5 일에 투여될 수 있는 경우 5 일 치료를 포험마며, 사르그라모스팀 및 AMD3100은 Ad35 투여 6 내지 8 시간 전에 투여된다.

[0562] 특정 구현예에서, AMD3100의 치료 유효량은 0.1 mg/kg 내지 100 mg/kg을 포함한다. 특정 구현예에서, AMD3100의 치료 유효량은 0.5 mg/kg 내지 50 mg/kg을 포함한다. 특정 구현예에서, AMD3100의 치료 유효량은 0.5 mg/kg, 1 mg/kg, 2 mg/kg, 3 mg/kg, 4 mg/kg, 5 mg/kg, 6 mg/kg, 7 mg/kg, 8 mg/kg, 9 mg/kg, 10 mg/kg, 11 mg/kg, 12 mg/kg, 13 mg/kg, 14 mg/kg, 15 mg/kg, 16 mg/kg, 17 mg/kg, 18 mg/kg, 19 mg/kg, 20 mg/kg, 또는 그 이상을 포함한다. 특정 구현예에서, AMD3100의 치료 유효량은 4 mg/kg을 포함한다. 특정 구현예에서, AMD3100의 치료 유효량은 5 mg/kg을 포함한다. 특정 구현예에서, AMD3100의 치료 유효량은 10 μg/kg 내지

500 μg/kg 또는 50 μg/kg 내지 400 μg/kg을 포함한다. 특정 구현예에서, AMD3100의 치료 유효량은 100 μg/kg, 150 μg/kg, 200 μg/kg, 250 μg/kg, 300 μg/kg, 350 μg/kg, 또는 그 이상을 포함한다. 특정 구현예에서, AMD3100은 피하로 또는 정맥내로 투여될 수 있다. 특정 구현예에서, AMD3100은 Ad35 전달 6 내지 11 시간 전에 160-240 μg/kg으로 피하로 투여될 수 있다. 특정 구현예에서, AMD3100의 치료 유효량은 또 다른 동원 인자의 투여와 동시에 투여될 수 있다. 특정 구현예에서, AMD3100의 치료 유효량은 또 다른 동원 인자의 투여 후에 투여될 수 있다. 특정 구현예에서, AMD3100의 치료 유효량은 G-CSF의 투여 후에 투여될 수 있다. 특정 구현예에서, 치료 프로토콜은 G-CSF가 1 일, 2 일, 3 일, 및 4 일 및 5 일에 투여되는 경우 5-일 치료를 포함하며, G-CSF 및 AMD3100은 Ad35 주사 6 내지 8 시간 전에 투여된다.

[0563] 투여될 SCF의 치료 유효량은 예를 들어, 0.1 내지 100 μg/kg/일 또는 0.5 내지 50 μg/kg/일 범위의 용량을 포함할 수 있다. 특정 구현예에서, SCF가 투여될 수 있는 용량은 0.5 μg/kg/일, 1 μg/kg/일, 2 μg/kg/일, 3 μg/kg/일, 4 μg/kg/일, 5 μg/kg/일, 6 μg/kg/일, 7 μg/kg/일, 8 μg/kg/일, 9 μg/kg/일, 10 μg/kg/일, 11 μg/kg/일, 12 μg/kg/일, 13 μg/kg/일, 14 μg/kg/일, 15 μg/kg/일, 16 μg/kg/일, 17 μg/kg/일, 18 μg/kg/일, 19 μg/kg/일, 20 μg/kg/일, 21 μg/kg/일, 22 μg/kg/일, 23 μg/kg/일, 24 μg/kg/일, 25 μg/kg/일, 26 μg/kg/일, 27 μg/kg/일, 28 μg/kg/일, 29 μg/kg/일, 30 μg/kg/일, 또는 그 이상을 포함한다. 특정 구현예에서, SCF는 1 일, 2 일 연속, 3 일 연속, 4 일 연속, 5 일 연속, 또는 그 이상 동안 투여될 수 있다. 특정 구현예에서, SCF는 피하로 또는 정맥내로 투여될 수 있다. 특정 구현예에서, SCF는 20 μg/kg/일로 피하로 주사될 수 있다. 특정 구현예에서, SCF는 단일 제제로서 투여된 후 또 다른 동원 인자와 동시 투여될 수 있다. 특정 구현예에서, SCF는 단일 제제로서 투여된 후 AMD3100과 동시 투여될 수 있다. 특정 구현예에서, 치료 프로토콜은 SCF가 1 일, 2 일, 3 일, 4 일 및 5 일에 투여될 수 있는 경우 5 일 치료를 포함하며, SCF 및 AMD3100은 Ad35 투여 6 내지 8 시간 전에 투여된다.

[0564] 특정 구현예에서, 성장 인자 GM-CSF 및 G-CSF는 골수 틈새에서 HSPC를 말초 순환 혈액에 동원하여 혈액에서 순환하는 HSPC의 분획을 증가시키기 위해 투여될 수 있다. 특정 구현예에서, 동원은 G-CSF/필그라스팀(Amgen) 및/또는 AMD3100(Sigma)의 투여로 달성될 수 있다. 특정 구현예에서, 동원은 GM-CSF/사르그라모스팀(Amgen) 및/또는 AMD3100(Sigma)의 투여로 달성될 수 있다. 특정 구현예에서, 동원은 SCF/안세스팀(Ancestim)(Amgen) 및/또는 AMD3100(Sigma)의 투여로 달성될 수 있다. 특정 구현예에서, G-CSF/필그라스팀의 투여는 AMD3100의 투여에 선행한다. 특정 구현예에서, G-CSF/필그라스팀의 투여는 AMD3100의 투여와 동시에 발생한다. 특정 구현예에서, G-CSF/필그라스팀의 투여는 AMD3100의 투여에 선행한 후, G-CSF/필그라스팀 및 AMD3100이 동시 투여된다. US 20140193376은 S1P 수용체 1(S1PR1) 조절인자 제제와 함께 CXCR4 길항제를 활용하는 동원 프로토콜을 기재한다. US 20110044997은 혈관 내피 성장 인자 수용체(VEGFR) 작용제와 함께 CXCR4 길항제를 활용하는 동원 프로토콜을 기재한다.

[0565] Ad35 바이러스 벡터는 HSPC 동원과 협력하여 투여될 수 있는 벡터의 예이다. 특정 구현예에서, Ad35 바이러스 벡터의 투여는 하나 이상의 동원 인자의 투여와 동시에 발생한다. 특정 구현예에서, Ad35 바이러스 벡터의 투여는 하나 이상의 동원 인자의 투여 후에 수행된다. 특정 구현예에서, Ad35 바이러스 벡터의 투여는 첫번째 하나 이상의 동원 인자의 투여 후에 수행되고 두번째 하나 이상의 동원 인자의 투여와 동시에 발생한다.

[0566] 특정 구현예에서, CD19 면역독소 또는 5-FU와 같은 HSC 강화제는 HSPC를 강화하기 위해 투여될 수 있다. CD19 면역독소는 골수 세포의 30%를 차지하는 모든 CD19 계통 세포를 고갈시키는 데 사용될 수 있다. 고갈은 골수로부터 빠져나가는 것을 부추긴다. HSPC가 증식하도록 강제함으로써(5-FU의 CD19 면역독소를 통해서든, 이는 분화를 자극하고 골수로부터 빠져나가고 말초 혈액 세포에서 이식유전자 마킹을 증가시킨다.

[0567] 치료 유효량은 주사, 주입, 관류에 의한 것과 같은 임의의 적절한 투여 경로를 통해, 보다 특히 골수, 정맥내, 피내, 동맥내, 결절내, 림프내, 복강내 주사, 주입, 또는 관류 중 하나 이상에 의한 투여에 의해 투여될 수 있다).

[0568] **IV(A)ii. 면역억제 레지멘**

[0569] Ad35 바이러스 벡터는 하나 이상의 스테로이드, IL-1 수용체 길항제, 및/또는 IL-6 수용체 길항제 투여를 포함할 수 있는, 하나 이상의 면역억제제 또는 면역억제 레지멘의 투여와 동시에 또는 투여 후에 투여될 수 있다. 이러한 프로토콜은 치료의 잠재적인 부작용을 완화할 수 있다.

[0570] IL-1 수용체 길항제는 알려져 있고 ADC-1001(Alligator Bioscience), FX-201(Flexion Therapeutics), Bioasis Technologies로부터 이용가능한 융합 단백질인 GQ-303(Genequine Biotherapeutics GmbH), HL-2351(Handok,

Inc.), MBIL-1RA(ProteoThera, Inc.), Anakinra(Pivor Pharmaceuticals), 인간 면역글로빈 G 또는 글로불린 S(GC Pharma)를 포함한다. IL-6 수용체 길항제는 또한 당업계에 알려져 있고 토실리주맙, BCD-089(Biocad), HS-628(Zhejiang Hisun Pharm), 및 APX-007(Apexigen)을 포함한다.

[0571] 다양한 구현예에서, 면역 억제 레지멘은 또한 적어도 하나의 바이러스 유전자 요법 벡터를 받는 대상체에게 투여되며, 여기서 면역 억제 레지멘은 (i) 대상체에게 제1 용량의 바이러스 유전자 요법 벡터를 투여하기 1 일 이상 전; (ii) 제1 용량의 바이러스 유전자 요법 벡터의 투여와 동일한 날; (iii) 하나 이상의 제2 또는 다른 후속 용량의 바이러스 유전자 요법 벡터의 투여와 동일한 날; 및/또는 (iv) 대상체에게 제1 용량의 바이러스 유전자 요법 벡터의 투여 및 제2 또는 다른 후속 용량의 바이러스 유전자 요법 벡터의 임의의 하나 이상, 또는 전부의 투여 사이에 개입하는 임의의 하나 이상, 또는 모든 일에 대상체에게 적어도 하나의 면역 억제제 투여를 포함한다.

[0572] 면역억제 레지멘은 예를 들어, 미국 가출원 번호 제63/009,218호에 추가로 기재되어 있으며, 이는 특히 면역억제 레지멘과 관련하여 그 전문이 본원에 참조로 포함된다.

[0573] **IV(A)iii. 선택**

[0574] 특정 구현예에서, 사용 방법은 교정된 세포가 비-교정된 세포에 비해 선택적 이점을 갖는 병태의 치료를 포함한다. Ad35 바이러스 벡터는 HSPC 동원과 협력하고 생체내 선택 카세트(들)와 상응하는 선택제의 투여 전에 투여될 수 있는 벡터의 예이다. 특정 구현예는 본원에 기재된 Ad35 벡터 및 $MGMT^{P140K}$ 카세트를 포함하는 Ad35의 경우 BCNU 또는 벤질구아닌 및 테모졸로미드 및/또는 항-CD33 카세트를 포함하는 Ad35 벡터의 경우 CD33-표적화 분자의 투여와 동원(예를 들어, 본원에 기재된 동원 프로토콜)을 조합한다.

[0575] 특정 구현예에서, 생체내 Ad35-매개 유전자 전달(동원이 있거나 또는 없이)은 생체내 선택 마커와 조합될 수 있다. 특정 구현예에서, 생체내 선택 마커는 Olszko 등, *Gene Therapy* 22: 591-595, 2015에 기재된 바와 같은 $MGMT^{P140K}$를 포함할 수 있다.

[0576] 인간 알킬 구아닌 트랜스퍼라제(hAGT)를 암호화하는 약물 내성 유전자 MGMT는 니트로소우레아 및 테모졸로미드(TMZ)와 같은 알킬화제의 세포독성 효과에 내성을 부여하는 DNA 복구 단백질이다. 6-벤질구아닌(6-BG)은 니트로소우레아 독성을 강화하는 AGT의 억제제이며 이 제제의 세포독성 효과를 강화하기 위해 TMZ와 공동 투여된다. AGT의 변이체를 암호화하는 MGMT의 여러 돌연변이체 형태는 6-BG에 의한 불활성화에 고도로 내성이지만, DNA 손상을 복구하는 능력을 유지한다(Maze 등 *J. Pharmacol. Exp. Ther.* 290: 1467-1474, 1999). $MGMT^{P140K}$-기반 약물 내성 유전자 요법은 마우스, 개, 레서스 원숭이, 및 인간 세포, 구체적으로 조혈 세포에 화학보호를 부여하는 것으로 나타났다(Zielske 등 *J. Clin. Invest.* 112: 1561-1570, 2003; Pollok 등 *Hum. Gene Ther.* 14: 1703-1714, 2003; Gerull 등 *Hum. Gene Ther.* 18: 451-456, 2007; Neff 등 *Blood* 105: 997-1002, 2005; Larochelle 등 *Clin. Invest.* 119: 1952-1963, 2009; Sawai 등 *Mol. Ther.* 3: 78-87, 2001).

[0577] 특정 구현예에서, 생체내 선택 마커와의 조합은 유전자-교정된 세포의 선택적 이점 없이 질환에 대한 주요한 구성요소가 될 것이다. 예를 들어, SCID 및 일부 다른 면역결핍 및 FA에서, 교정된 세포는 이점을 가지며 치료 유전자를 "약간의" HSPC 내로 형질도입하는 것만으로도 치료 효능에 충분하다. 세포가 경쟁적 이점을 입증하지 않는 혈색소병증(즉, 겸상적혈구병 및 지중해빈혈)과 같은 다른 질환의 경우, $MGMT^{P140K}$와 같은 생체내 선택 마커와 조합하는 것과 같은 유전자 교정된 세포의 생체내 선택은 약간의 형질도입된 HSPC를 선택하여, 유전자 교정된 세포에서의 증가를 허용하고 치료 효능을 달성하기 위함이다. 이 접근법은 또한 HSPC가 생체외 유전적 변형보다는 생체내에서 HIV에 내성을 가짐으로써 HIV에 적용될 수 있다.

[0578] 추가적인 접근법이 또한 사용될 수 있다. 예를 들어, 본 개시내용은 유전적으로 변형된 치료 세포에서 CD33 발현을 선택적으로 감소시키는 동시에 치료 유전자를 제공하기 위해 세포를 유전적으로 변형시키는 시스템 및 방법을 활용할 수 있다. 이 방식에서, 유전적으로 변형된 치료 세포는 환자가 받을 수 있는 동시 또는 후속 항-CD33 요법에 의해 해를 입지 않을 것이다. 그러나, 환자에서 기존 CD33-발현 세포 및/또는 유전적 변형이 결여된 투여된 세포는 보호되지 않아, 교정되지 않은 세포에 비해 유전자-교정된 세포에 대한 양성 선택을 초래할 것이다.

[0579] 특정 구현예에서, 이 접근법은 단일 세포내 전달 비히클에서 치료 유전자 및 CD33 차단 분자를 연결함으로써 달성된다. 특정 구현예에서, 단일 세포내 전달 비히클은 Ad35 바이러스 벡터이다.

[0580] 특정 구현예에서, CD33 차단 분자는 공통 Ad35 바이러스 벡터 내에 포함됨으로써 치료 유전자와 조합된 shRNA 또는 siRNA CD33 차단 분자이다. 특정 구현예에서, CD33 차단 분자는 서열번호: 187을 포함하는 shRNA 서열 또는 서열번호: 188을 포함하는 서열이다.

[0581] CD33-표적화 치료는 항-CD33 항체, 항-CD33 면역독소, 항-CD33 항체-약물 접합체, 항-CD33 항체-방사성 동위원소 접합체, 항-CD33 이중특이적 항체, 항-CD33 BiTE® 항체, 항-CD33 삼중특이적 항체, 및/또는 항-CD33 CAR을 포함한다.

[0582] **IV(B). 시험관내 및 생체외 유전자 요법**

[0583] 시험관내 유전자 요법은 외인성 DNA를 숙주 세포(예컨대 표적 세포) 및/또는 핵산(예컨대 표적 핵산, 예컨대 표적 게놈) 내로 도입하는 방법에서 본 개시내용의 벡터, 게놈, 또는 시스템의 사용을 포함하며, 여기서 숙주 세포 또는 핵산은 다세포 유기체(예를 들어, 실험실)에 존재하지 않는다. 일부 구현예에서, 표적 세포 또는 핵산은 다세포 유기체, 예컨대 포유동물(예를 들어, 마우스, 래트, 인간, 또는 비-인간 영장류)로부터 유래된다. 다세포 유기체로부터 유래된 세포의 시험관내 조작은 생체외 조작으로 지칭될 수 있고, 생체외 요법에 사용될 수 있다. 다양한 구현예에서, 본 개시내용의 방법 및 조성물은 예를 들어, 본원에 개시된 바와 같이, 제1 다세포 유기체로부터 유래된 표적 세포 또는 핵산을 변형시키기 위해 활용된 다음 조작된 표적 세포 또는 핵산은 예를 들어, 입양 세포 요법의 방법으로 제2 다세포 유기체, 예컨대 포유동물(예를 들어, 마우스, 래트, 인간, 또는 비-인간 영장류)에게 투여된다. 일부 경우에, 제1 및 제2 유기체는 동일한 단일 대상 유기체이다. 물질이 유래된 대상체에게 시험관내 조작된 물질을 되돌리는 것은 자가 요법일 수 있다. 일부 경우에, 제1 및 제2 유기체는 상이한 유기체(예를 들어, 동일한 종의 2 개의 유기체, 예를 들어, 동일한 종의 2 마리의 마우스, 2 마리의 래트, 2 명의 인간, 또는 2 마리의 비-인간 영장류)이다. 제1 대상체로부터 유래된 조작된 물질을 제2 상이한 대상체로 전달하는 것은 동종이계 요법일 수 있다.

[0584] 생체외 세포 요법은 환자 또는 정상 공여자로부터 줄기, 전구체 또는 분화된 세포의 단리, 유전적 조작이 있거나 또는 없이 생체외에서 단리된 세포의 확장 및 주입된 세포 및/또는 이의 자손의 일시적 또는 안정된 이식을 달성하기 위해 대상체에게 세포의 투여를 포함할 수 있다. 이러한 생체외 접근법은 예를 들어, 유전성, 감염성 또는 신생물성 질환을 치료하거나, 조직을 재생하거나 또는 치료제를 질환 부위에 전달하기 위해 사용될 수 있다. 다양한 생체외 요법에서 대상체를 유전자 전달 벡터에 직접 노출시키기 않고, 형질도입의 표적 세포는 임의의 유전적 조작 전 또는 후에, 효능 및 안전성을 개선하기 위해 선택, 확장, 및/또는, 분화될 수 있다.

[0585] 생체외 요법은 조혈 줄기 세포(HSC) 이식(HCT)을 포함한다. 자가 HSC 유전자 요법은 혈액 및 면역계의 여러 단일유전자성 질환 뿐만 아니라 저장 장애에 대한 치료 옵션을 나타내고, 이는 선택된 질환 상태에 대한 1차 치료 옵션이 될 수 있다. 또 다른 확립된 세포 및 유전자 요법 적용은 악성종양, 감염 및 자가면역 질환에 대해 사용하기 위한 면역 효과기 및 조절 세포의 힘을 활용하기 위해, 유전적 조작이 있거나 또는 없이 항원 특이성을 재지시하거나 또는 안정성 프로파일을 증가시키는 생체외 확장된 T 세포를 이용하는 입양 면역요법이다. 일부 경우에 유전적 조작에 수반되는 다양한 다른 유형의 체세포 줄기 세포는 표피 및 윤부 줄기 세포, 신경 줄기/전구체 세포(NSPC), 심장 줄기 세포 및 다능성 간질 세포(MSC)를 포함하는 치료 적용에 대한 가능성을 나타내고 있다.

[0586] 생체외 요법의 적용은 기능장애 세포 계통을 재구축하는 것을 포함한다. 결함성이거나 또는 부재하는 세포 계통을 특징으로 하는 선천적 질환의 경우, 계통은 정상 공여자 또는 결핍을 교정하기 위해 생체외 유전자 전달에 적용된 자가 세포로부터 유래된 기능적 전구체 세포에 의해 재생될 수 있다. 예는 SCID에 의해 제공되며, 여기서 여러 유전자 중 임의의 하나에서 결핍은 성숙 림프 세포의 발달을 차단한다. 숙주에서 다양한 계통의 공여자-유래 기능적 조혈 세포의 생성을 허용할 수 있는 조작되지 않은 정상 공여자 HSC의 이식은 SCID, 뿐만 아니라 혈액 및 면역계에 영향을 미치는 많은 다른 질환에 대한 치료 옵션을 나타낸다. 이식된 조혈 줄기/전구체 세포(HSPC)에서 결핍성 유전자의 기능적 카피를 대체하는 것을 포함할 수 있고, HCT와 유사하게, 기능적 자손의 꾸준한 공급을 제공할 수 있는 자가 HSC 유전자 요법은 이식편대숙주병(GvHD)의 위험 감소, 이식 거부의 위험 감소, 및 이식후 면역억제에 대한 필요성 감소를 포함하는 여러 이점을 가질 수 있다.

[0587] 생체외 요법의 적용은 치료 유전자 투여량 증가를 포함한다. 일부 적용에서, HSC 유전자 요법은 동종이계 HCT의 치료 효능을 증가시킬 수 있다. 치료 유전자 투여량은 이식된 세포에서 정상 수준 이상으로 조작될 수 있다.

[0588] 생체외 요법의 적용은 신규 기능을 도입하고 유전자 요법을 표적화하는 것을 포함한다. 생체외 유전자 요법은

고용향 항종양 화학요법 레지멘의 투여를 허용하는 약물 내성을 확립하거나 또는 HIV와 같은 바이러스, 또는 RNA-기반 제제(예를 들어, 리보자임, RNA 유인, 안티센스 RNA, RNA 압타머 및 작은 간섭 RNA) 및 단백질-기반 제제(예를 들어, 병원체의 게놈을 표적화하는 우성-음성 돌연변이체 바이러스 단백질, 융합 억제제 및 조작된 뉴클레아제)를 발현함으로써 다른 병원체로 사전 확립된 감염에 대한 내성을 확립하는 것과 같이, HSC 또는 이의 자손에게 신규 기능을 부여할 수 있다.

[0589] 생체외 요법의 적용은 면역 반응을 향상시키는 것을 포함한다. 신생물 질환에서, T 세포와 같은 동종이계 적응 면역 세포 유형은 암 세포를 인식하고 사멸시킬 수 있다. 불행히도, 동종반응성 림프구에 의한 건강한 조직의 인식은 또한 해로운 GvHD를 초래할 수 있다. 공여자 림프구에서 자살 유전자의 전달은 독성을 길들이면서 항-종양 잠재력을 활용할 수 있게 한다. 자가 설정에서, 형질전환 또는 감염된 세포에 대해 지시된 특이성이 있는 림프구는 환자의 조직으로부터 단리되고 생체외에서 선택적으로 확장될 수 있다 대안적으로, 이들은 형질전환 또는 감염된 세포를 접할 때 세포의 반응을 촉발하는 합성 또는 키메라 항원 수용체에 대한 유전자의 전달에 의해 생성될 수 있다. 이러한 접근법은 종양 또는 감염에 대한 기본적인 숙주 반응을 강화하거나, 또는 이를 새롭게 유도할 수 있다.

[0590] **IV(C). 유전자 요법에 의해 치료가능한 병태**

[0591] 적어도 부분적으로 본 개시내용의 아데노바이러스 벡터는 숙주 및/또는 표적 세포의 변형을 위해 생체내, 시험관내, 또는 생체외에서 사용될 수 있고, 추가로 아데노바이러스 벡터는 매우 다양한 발현 산물을 암호화하는 페이로드를 포함할 수 있기 때문에, 본원에 제공되는 다양한 기술은 광범위한 적용가능성을 갖고 매우 다양한 병태를 치료하기 위해 사용될 수 있음이 본 명세서로부터 분명할 것이다. 본 개시내용의 아데노바이러스 벡터, 게놈, 또는 시스템의 투여에 의해 치료가능한 병태의 예는 혈색소병증, 면역결핍, 점 돌연변이 병태, 암, 단백질 결핍, 감염성 질환, 및 염증성 병태를 포함하나 이에 제한되지 않는다.

[0592] 특정 구현예에서, 본원에 개시된 벡터, 게놈, 시스템 및 제형은 대상체(인간, 수의학 동물(개, 고양이, 파충류, 조류 등), 가축(말, 소, 염소, 돼지, 닭 등), 및 연구 동물(원숭이, 래트, 마우스, 어류 등)을 치료하기 위해 사용될 수 있다. 대상체를 치료하는 것은 치료 유효량을 전달하는 것을 포함한다. 치료 유효량은 유효량, 예방적 치료, 및/또는 치료적 치료를 제공하는 것을 포함한다.

[0593] 특정 구현예에서, 본원에 개시된 방법 및 제형은 혈액 장애를 치료하는 데 사용될 수 있다. 특정 구현예에서, 제형은 혈우병, 중증 β-지중해빈혈, 다이아몬드 블랙판 빈혈(DBA), 발작성 야간 혜모글로빈뇨증(PNH), 순수 적혈구 무형성증(PRCA), 불응성 빈혈, 중증 재생불량성 빈혈, 및/또는 백혈병, 림프종, 및 골수종과 같은 혈액암을 치료하기 위해 대상체에게 투여된다.

[0594] 혈색소병증은 불균형 결과를 초래하는 전 세계적인 건강 부담을 나타낸다. 헤모글로빈 단백질 또는 글로빈 유전자 발현에서의 결함은 혈색소병증이라 불리는 질환을 초래할 수 있다. 혈색소병증은 전 세계적으로 가장 흔한 유전 질환 중 하나이다.

[0595] 매년, 전 세계적으로 110만 명의 신생아가 혈색소병증에 대한 위험이 있으며, 헤모글로빈(Hb) 유전적 변이에 의해 부여된 말라리아 감염에 대한 자연적 내성으로 인해, 말라리아 열대열원충이 만연한 지리학적 영역에서 매 1,000 명마다 신생아 중 25명 정도로 영향을 미친다. 선진국에서, 환자는 만성 수혈로부터 철분 과부하의 위험이 있다. 후진 지역에서, 생존은 상당히 더 낮다. 예를 들어, 아프리카에서, 유아기 사망률은 모든 어린이에서 16%인 것과 비교하여 혈색소병증 환자에서 40%이다.

[0596] 글로빈 유전자에서 돌연변이는 겸상적혈구병(SCD) 및 헤모글로빈 C, D, 및 E 질환에서과 같이 비정상적인 형태의 헤모그로빈을 생성하거나, 또는 α 또는 β 폴리펩티드의 생산을 감소시키고 따라서 세포에서 글로빈 쇄의 불균형을 초래할 수 있다. 이러한 후자의 병태는 글로빈 쇄가 손상되는 정도에 따라 α- 또는 β-지중해빈혈이라고 불린다. 전 세계 인구의 5%는 단연코 가장 흔한(담체의 40%) b-글로빈(HBB) 유전자(글루타메이트에서 발린으로 전환; 역사적으로 E6V, 동시에 E7V)에서 겸상 세포 돌연변이가 있는 상당한 헤모글로빈 변이체를 가지고 있다. 헤모글로빈 장애의 높은 유병률 및 중증도는 일생동안 환자 관리에 비용이 들기 때문에, 이환된 환자의 삶 뿐만 아니라 건강-관리 시스템에 영향을 미치는 실질적인 부담을 나타낸다.

[0597] 2 개의 알파(α) 및 2 개의 감마(γ) 쇄를 포함하는 태아(HbF), 및 2 개의 α 및 2 개의 베타(β) 쇄를 포함하는 성체(HbA)의 2 가지 형태의 헤모글로빈이 있다. HbF에서 HbA로의 자연적 전환은 출생 직후 발생하고 마스터 조절인자인 bcl11a를 포함하는 인자에 의해 γ 글로빈 유전자의 전사 억제에 의해 조절된다. 비판적으로, 다양한 임상 관찰은 겸상적혈구병 및 β-지중해빈혈과 같은 β-혈색소병증의 중증도가 HbF의 생산 증가에 의해 개선

됨을 입증한다.

[0598] 특정 구현예에서, 치료적으로 효과적인 치료는 HbF의 발현을 유도 또는 증가시키고/시키거나, 헤모글로빈의 생산을 유도 또는 증가시키고/시키거나 β-글로빈의 생산을 유도 또는 증가시킨다. 특정 구현예에서, 치료적으로 효과적인 치료는 혈액 세포 기능을 개선하고/하거나, 세포의 산소화를 증가시킨다.

[0599] 다양한 구현예에서, 본 개시내용은 β-글로빈 긴 LCR, β-글로빈 프로모터, 및 혈액 장애의 치료를 위한 단백질 또는 제제를 암호화하는 코딩 핵산 서열을 포함하는 본 개시내용의 아데노바이러스 공여자 벡터를 사용한 혈액 장애의 치료를 포함한다. 다양한 구현예에서, 혈액 장애는 지중해빈혈이고 단백질은 β-글로빈 또는 γ-글로빈 단백질, 또는 달리 부분적으로 또는 완전히 β-글로빈 또는 γ-글기능적으로 대체하는 단백질이다. 다양한 구현예에서, 혈액 장애는 혈우병이고 단백질은 ET3 또는 달리 부분적으로 또는 완전히 인자 VIII을 기능적으로 대체하는 단백질이다. 다양한 구현예에서, 혈액 장애는 겸상 세포 빈혈과 같은 점 돌연변이 질환이고, 제제는 유전자 편집 단백질이다.

[0600] ET3은 하기 아미노산 서열을 가질 수 있다: 서열번호 301. 다양한 구현예에서, 인자 VIII 대체 단백질은 서열번호: 301에 대해 적어도 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일성을 갖는 아미노산 서열을 가질 수 있다.

[0601] β-글로빈은 하기 아미노산 서열을 가질 수 있다: 서열번호 302. 다양한 구현예에서, β-글로빈 대체 단백질은 서열번호: 302에 대해 적어도 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일성을 갖는 아미노산 서열을 가질 수 있다.

[0602] γ-글로빈은 하기 아미노산 서열을 가질 수 있다: 서열번호 303. 다양한 구현예에서, γ-글로빈 대체 단백질은 서열번호: 303에 대해 적어도 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일성을 갖는 아미노산 서열을 가질 수 있다.

[0603] 80 개 초과의 1차 면역 결핍 질환이 세계보건기구에 의해 인식되고 있다. 이러한 질환은 일부 경우에, 신체가 감염에 대한 임의의 또는 충분한 항체를 생산할 수 없는 면역계의 본질적인 결함을 특징으로 한다. 다른 경우에, 감염에 싸우는 세포 방어가 적절하게 작동하지 않는다. 전형적으로, 1차 면역 결핍은 선천적 장애이다.

[0604] 이차성, 또는 후천성 면역 결핍은 선천적 유전 이상의 결과가 아니라, 오히려 면역계가 면역계 외부에 있는 요인에 의해 손상된 개인에서 발생한다. 예는 외상, 바이러스, 화학요법, 독소, 및 오염을 포함한다. 후천성 면역결핍 증후군(AIDS)은 T 림프구의 고갈이 신체를 감염에 싸울 수 없게 만드는 인간 면역결핍 바이러스(HIV)와 같은 바이러스에 의해 유발된 2차 면역 결핍 장애의 예이다.

[0605] X-결합 중증 복합 면역결핍(SCID-X1)은 공통 감마 쇄 유전자(γC)의 돌연변이에 의해 유발된 세포성 및 체액성 면역 고갈 둘 다이며, 이는 T 및 자연 살해(NK) 림프구의 부재 및 비기능적 B 림프구의 존재를 초래한다. SCID-X1은 면역계가 예를 들어, 골수 이식(BMT) 또는 유전자 요법을 통해 재구축되지 않는 한 생후 첫 2년 내에 치명적이다.

[0606] 대부분의 개인은 BMT 또는 비-자가 유전자 요법에 대해 일치된 공여자가 부족하기 때문에, 성숙 T 세포가 고갈된 동일단배체의 모체 골수가 종종 사용되지만; 합병증은 이식편대숙주병(GVHD), 적절한 항체를 생성하지 못해 장기간 면역글로불린 대체 필요, 조혈 줄기 및 전구체 세포(HSPC)의 주입 실패로 인한 T 세포의 후기 손실, 만성 혹, 및 림프구 조절장애를 포함한다.

[0607] 판코니 빈혈(FA)은 골수 부전으로 이어지는 선천적 혈액 장애이다. 이는 부분적으로 DNA-복구 메커니즘 결핍을 특징으로 한다. FA 환자의 적어도 20%는 급성 골수성 백혈병과 같은 암, 및 피부, 간, 위장관, 및 부인과계의 암으로 발전한다. 피부 및 위장관 종양은 일반적으로 편평세포 암종이다. 암으로 발전하는 환자의 평균 연령은 백혈병의 경우 15 세, 간 종양의 경우 16 세, 및 다른 종양의 경우 23 세이다.

[0608] 치료 유전자는 특정 구현예에서, 선천적인 병태에 대해 치료적으로 효과적인 반응을 제공하도록 선택될 수 있다. 특정 구현예에서, 병태는 그레이브스병, 류마티스 관절염, 악성 빈혈, 다발성 경화증(MS), 염증성 장 질환, 전신 홍반성 루푸스(SLE), 아데노신 데아미나제 결핍증(ADA-SCID) 또는 중증 복합 면역결핍증(SCID), 비스코트-알드리치 증후군(WAS), 만성 육아종증(CGD), 판코니 빈혈(FA), 바텐병, 부신백질이영양증(ALD) 또는 이염백질이영양증(MLD), 근이영양증, 폐포단백질증(PAP), 피루베이트 키나제 결핍, 슈바크만-다이아몬드-블랙판 빈혈, 선천성 이각화증, 낭포성 섬유증, 파킨슨병, 알츠하이머병, 또는 근위축성 측색 경화증(루게릭병)일 수 있다. 특정 구현예에서, 병태에 따라, 치료 유전자는 단백질을 암호화하는 유전자 및/또는 기능이 중단된 유전자

일 수 있다.

[0609] 특정 구현예에서, 본원에 개시된 방법 및 제형은 암을 치료하는 데 사용될 수 있다. 특정 구현예에서, 제형은 급성 림프모구성 백혈병(ALL), 급성 골수형성 백혈병(AML), 만성 림프구성 백혈병(CLL), 만성 골수형성 백혈병(CML), 만성 골수단핵구 백혈병, 미만성 거대 B-세포 림프종, 여포성 림프종, 호지킨 림프종, 청소년 골수단핵구 백혈병, 다발성 골수종, 골수이형성, 및/또는 비-호지킨 림프종을 치료하기 위해 대상체에게 투여된다.

[0610] 치료될 수 있는 추가의 예시적인 암은 성상세포종, 비정형 기형 간상소체 종양, 뇌 및 중추신경계(CNS) 암, 유방암, 암육종, 연골육종, 척색종, 맥락총 암종, 맥락총 유두종, 연조직의 투명 세포 육종, 미만성 거대 B-세포 림프종, 뇌실막세포종, 상피양 육종, 생식선외 생식 세포 종양, 신장외 횡문양 종양, 유잉 육종, 위장관 간질 종양, 교모세포종, HBV-유도된 간세포 암종, 두경부암, 신장암, 폐암, 악성 간상소체 종양, 수모세포종, 흑색종, 수막종, 중피종, 다발성 골수종, 신경교 종양, 달리 명시되지 않은(NOS) 육종, 핍지교성상세포종, 희소돌기아교세포종, 골육종, 난소암, 난소 투명 세포 선암종, 난소 자궁내막 선암종, 난소 장액성 선암종, 췌장암, 췌장관 선암종, 췌장 내분비성 종양, 송과체모세포종, 전립선암, 신세포 암종, 신수질 암종, 횡문근육종, 육종, 신경집종, 피부 편평세포 암종, 및 줄기 세포암을 포함한다. 다양한 특정 구현예에서, 암은 난소암이다. 다양한 특정 구현예에서 암은 유방암이다.

[0611] 특정 구현예에서, 본원에 개시된 방법 및 제형은 점 돌연변이 병태를 치료하는 데 사용될 수 있다. 특정 구현예에서, 제형은 겸상적혈구병, 낭포성 섬유증, 테이삭스병, 및/또는 페닐케톤뇨증을 치료하기 위해 대상체에게 투여된다. 다양한 구현예에서, 본 개시내용의 트랜스포존 페이로드는 핵산 병변의 교정 편집을 위한 CRISPR-Cas를 암호화한다. 다양한 구현예에서, 본 개시내용의 트랜스포존 페이로드는 핵산 병변의 교정 편집을 위한 염기 편집기를 암호화한다.

[0612] 특정 구현예에서, 본원에 개시된 방법 및 제형은 특정 효소 결핍을 치료하는 데 사용될 수 있다. 특정 구현예에서, 제형은 헐러 증후군, 선택적 IgA 결핍, 고 IgM, IgG 하위클래스 결핍, 니만-픽병, 테이삭스병, 고세병, 파브리병, 크라베병, 포도당혈증, 단풍시럽뇨병, 페닐케톤뇨증, 글리코겐 축적 질환, 프리드라이히 운동실조, 젤웨거 증후군, 부신백질이영양증, 보체 장애, 및/또는 뮤코다당증을 치료하기 위해 대상체에게 투여된다.

[0613] 치료 유효량은 면역 및 다른 혈액 세포 및/또는 소교 세포에 기능을 제공할 수 있거나 또는 대안적으로 치료된 병태에 따라 림프구 활성화를 억제하고/하거나, 림프구에서 세포자멸사를 유도하고/하거나, 림프구의 다양한 하위세트를 제거하고/하거나, T 세포 활성화를 억제하고/하거나, 자가반응성 T 세포를 제거 또는 억제하고/하거나, Th-2 또는 Th-1 림프구 활성을 억제하고/하거나, IL-1 또는 TNF를 길항하고/하거나, 염증을 감소시키고/시키거나, 자극제에 대한 선택적 내성을 유도하고/하거나, 면역-매개 병태를 감소 또는 제거하고/하거나; 면역-매개 병태의 증상을 감소 또는 제거할 수 있다. 치료 유효량은 또한 기능적 DNA 복구 메커니즘; 계면활성제 단백질 발현; 텔로미어 유지; 리소좀 기능; 지질 또는 아밀로이드와 같은 다른 단백질의 분해를 제공하고/하거나; 리보솜 기능을 허용하고/하거나; 달리 다른 백혈구 유형의 대식세포와 같이 발달하지 않을 성숙 혈액 세포 계통의 발달을 허용할 수 있다.

[0614] 특정 구현예에서, 본 개시내용의 방법은 이를 필요로 하는 대상체에서 T-세포 매개된 면역 반응을 회복시킬 수 있다. T-세포 매개된 면역 반응의 회복은 흉선 배출 회복 및/또는 정상 T 림프구 발달 회복을 포함할 수 있다

[0615] 특정 구현예에서, 흉선 배출 회복은 말초 혈액에서 CD45RA를 발현하는 CD3+ T 세포의 빈도를 대조군 집단에서 유래된 참조 수준과 비슷한 수준으로 회복시키는 것을 포함할 수 있다. 특정 구현예에서, 흉선 배출 회복은 $10^6$ 개의 성숙 T 세포 당 T 세포 수용체 절제 원(TREC)의 수를 대조군 집단에서 유래된 참조 수준과 비슷한 수준으로 회복시키는 것을 포함할 수 있다. $10^6$ 개의 성숙 T 세포 당 TREC의 수는 Kennedy 등, *Vet Immunol Immunopathol* 142: 36-48, 2011에 기재된 바와 같이 결정될 수 있다.

[0616] 특정 구현예에서, 정상 T 림프구 발달 회복은 CD4+ 세포: CD8+ 세포의 비를 2로 회복시키는 것을 포함한다. 특정 구현예에서, 정상 T 림프구 발달 회복은 순환 T-림프구에서 αβ TCR의 존재를 검출하는 것을 포함한다. 순환 T-림프구에서 αβ TCR의 존재는 예를 들어, TCR의 α 및/또는 β 쇄에 결합하는 항체를 사용한 유세포 분석에 의해 검출될 수 있다. 특정 구현예에서, 정상 T 림프구 발달 회복은 다양한 TCR 레퍼토리의 존재를 대조군 집단에서 유래된 참조 수준과 비슷한 수준으로 검출하는 것을 포함한다. TCR 다양성은 TCRβ 유전자의 가변 영역의 유전적 재배열을 분석하는 TCRVβ 스펙트럼유형화(spectratyping)에 의해 평가될 수 있다. 강력하고 정상적인 스펙트럼유형 프로파일은 TCRVβ 분절의 17 개 패밀리에 걸쳐 크기가 부여된 단편의 가우스(Gaussian) 분포를 특징으로 할 수 있다. 특정 구현예에서, 정상 T 림프구 발달 회복은 T-세포 특이적 신호전달 경로의 회복

을 포함한다. T-세포 특이적 신호전달 경로의 회복은 T 세포 미토겐 파이토헤마글루티닌(PHA)에 노출 후 림프구 증식에 의해 평가될 수 있다. 특정 구현예에서, 정상 T 림프구 발달 회복은 백혈구 계수, 호중구 세포 계수, 단핵구 세포 계수, 림프구 세포 계수, 및/또는 혈소판 세포 계수를 대조군 집단에서 유래된 참조 수준과 비슷한 수준으로 회복시키는 것을 포함한다.

[0617] 특정 구현예에서, 본 개시내용의 방법은 이를 필요로 하는 대상체에서 림프구 재구축의 동역학 및/또는 클론 다양성을 개선할 수 있다. 특정 구현예에서, 림프구 재구축의 동역학 개선은 순환 T 림프구의 수를 대조군 집단에서 유래된 참조 수준의 범위 내로 증가시키는 것을 포함한다. 특정 구현예에서, 림프구 재구축의 동역학 개선은 절대 CD3+ 림프구 계수를 대조군 집단에서 유래된 참조 수준의 범위 내로 증가시키는 것을 포함할 수 있다. 이의 범위는 주어진 매개변수에 대해 정상(즉, 면역 손상되지 않은) 대상체에서 관찰되거나 또는 이에 의해 나타난 값의 범위일 수 있다. 특정 구현예에서, 림프구 재구축의 동역학 개선은 본원에 기재된 요법이 투여되지 않은 이를 필요로 하는 대상체와 비교하여 정상 림프구 계수에 도달하는 데 필요한 시간을 감소시키는 것을 포함할 수 있다. 특정 구현예에서, 림프구 재구축의 동역학 개선은 본원에 기재된 요법이 투여되지 않은 이를 필요로 하는 대상체와 비교하여 유전자 교정된 림프구의 빈도를 증가시키는 것을 포함할 수 있다. 특정 구현예에서, 림프구 재구축의 동역학 개선은 본원에 기재된 유전자 요법이 투여되지 않은 이를 필요로 하는 대상체와 비교하여 대상체에서 유전자 교정된 림프구의 클론 레퍼토리의 다양성을 증가시키는 것을 포함할 수 있다. 유전자 교정된 림프구의 클론 레퍼토리의 다양성 증가는 레트로바이러스 통합 부위(RIS) 분석에 의해 측정된 바와 같은 고유한 (RIS) 클론의 수를 증가시키는 것을 포함할 수 있다.

[0618] 특정 구현예에서, 본 개시내용의 방법은 이를 필요로 하는 대상체에서 골수 기능을 회복시킬 수 있다. 특정 구현예에서, 골수 기능 회복은 본원에 기재된 요법이 투여되지 않은 이를 필요로 하는 대상체와 비교하여 유전자 교정된 세포로 골수 재증식을 개선하는 것을 포함할 수 있다. 유전자 교정된 세포로 골수 재증식을 개선하는 것은 유전자 교정된 세포의 백분율을 증가시키는 것을 포함할 수 있다. 특정 구현예에서, 세포는 백혈구 및 골수 유래 세포로부터 선택된다. 특정 구현예에서, 유전자 교정된 세포의 백분율은 정량적 실시간 PCR 및 유세포 분석으로부터 선택된 검정을 사용하여 측정될 수 있다.

[0619] 특정 구현예에서, 본 개시내용의 방법은 이를 필요로 하는 대상체에서 면역화에 대한 1차 및 2차 항체 반응을 정규화할 수 있다. 면역화에 대한 1차 및 2차 항체 반응의 정규화는 항원에 대한 클래스 전환 및 기억 반응에서 기능하는 B-세포 및/또는 T-세포 사이토카인 신호전달 프로그램의 회복을 포함할 수 있다. 면역화에 대한 1차 및 2차 항체 반응의 정규화는 박테리오파지 면역화 검정에 의해 측정될 수 있다. 특정 구현예에서, B-세포 및/또는 T-세포 사이토카인 신호전달 프로그램의 회복은 T-세포 의존적 신생항원 박테리오파지 ΦX174로 면역화 후 검정될 수 있다. 특정 구현예에서, 면역화에 대한 1차 및 2차 항체 반응의 정규화는 이를 필요로 하는 대상체에서 IgA, IgM, 및/또는 IgG의 수준을 대조군 집단에서 유래된 참조 수준과 비슷한 수준으로 증가시키는 것을 포함할 수 있다. 특정 구현예에서, 면역화에 대한 1차 및 2차 항체 반응의 정규화는 이를 필요로 하는 대상체에서 IgA, IgM, 및/또는 IgG의 수준을 본원에 기재된 유전자 요법이 투여되지 않은 이를 필요로 하는 대상체보다 더 큰 수준으로 증가시키는 것을 포함할 수 있다. IgA, IgM, 및/또는 IgG의 수준은 예를 들어, 면역글로불린 테스트에 의해 측정될 수 있다. 특정 구현예에서, 면역글로불린 테스트는 항체 결합 IgG, IgA, IgM, 카파 경쇄, 람다 경쇄, 및/또는 중쇄를 포함한다. 특정 구현예에서, 면역글로불린 테스트는 혈청 단백질 전기영동법, 면역전기영동법, 방사성 면역확산, 혼탁법 및 비탁법을 포함한다. 상업적으로 이용가능한 면역글로불린 테스트 키트는 MININEPH™(Binding Site, 영국 버밍엄 소재), 및 Dako(덴마크 소재) 및 Dade Behring(독일 마르부르크 소재)의 면역글로불린 테스트 시스템을 포함한다. 특정 구현예에서, 면역글로불린 수준을 측정하는 데 사용될 수 있는 샘플은 혈액 샘플, 혈장 샘플, 뇌척수액 샘플, 및 소변 샘플을 포함한다.

[0620] 특정 구현예에서, 본 개시내용의 방법은 SCID-X1을 치료하는 데 사용될 수 있다. 특정 구현예에서, 본 개시내용의 방법은 SCID(예를 들어, JAK 3 키나제 결핍 SCID, 퓨린 뉴클레오시드 포스포릴라제(PNP) 결핍 SCID, 아데노신 데아미나제(ADA) 결핍 SCID, MHC 클래스 II 결핍 또는 재조합효소 활성화 유전자(RAG) 결핍 SCID)를 치료하는 데 사용될 수 있다. 특정 구현예에서, 치료 효능은 림프구 재구축, 클론 다양성 및 흉선형성 개선, 감염 감소, 및/또는 환자 결과 개선을 통해 관찰될 수 있다. 치료 효능은 또한 체중 증가 및 성장, 위장관 기능 개선(예를 들어, 설사 감소), 상부 호흡기 증상 감소, 구강(칸디다증)의 진균 감염 감소, 폐렴 발병률 및 중증도 감소, 수막염 및 혈류 감염 감소, 및 귀 감염 감소 중 하나 이상을 통해 관찰될 수 있다. 특정 구현예에서, 본 개시내용의 방법으로 SCIDX-1을 치료하는 것은 γC-의존적 신호전달 경로에 대한 기능성 회복을 포함한다. γC-의존적 신호전달 경로의 기능성은 각각 IL-21 및/또는 IL-2로 시험관내 자극 후 효과기 분자 STAT3 및/또는 STAT5의 티로신 인산화를 측정함으로써 검정될 수 있다. STAT3 및/또는 STAT5의 티로신 인산화는 세포내 항체

염색에 의해 측정될 수 있다.

[0621] 특정 구현예에서, 본 개시내용의 방법은 FA를 치료하는 데 사용될 수 있다. 특정 구현예에서, 치료 효능은 림프구 재구축, 클론 다양성 및 흉선형성 개선, 감염 감소, 및/또는 환자 결과 개선을 통해 관찰될 수 있다. 치료 효능은 또한 체중 증가 및 성장, 위장관 기능 개선(예를 들어, 설사 감소), 상부 호흡기 증상 감소, 구강(칸디다증)의 진균 감염 감소, 폐렴 발병률 및 중증도 감소, 수막염 및 혈류 감염 감소, 및 귀 감염 감소 중 하나 이상을 통해 관찰될 수 있다. 특정 구현예에서, 본 개시내용의 방법으로 FA를 치료하는 것은 미토마이신 C(MMC)에 대한 골수 유래 세포의 내성을 증가시키는 것을 포함한다. 특정 구현예에서, MMC에 대한 골수 유래 세포의 내성은 메틸셀룰로스 및 MMC에서 세포 생존 검정에 의해 측정될 수 있다.

[0622] 특정 구현예에서, 본 개시내용의 방법은 저감마글로불린혈증을 치료하는 데 사용될 수 있다. 저감마글로불린혈증은 B-림프구 결여에 의해 유발되고 혈액에서 낮은 수준의 항체를 특징으로 한다. 저감마글로불린혈증은 백혈병-관련 면역 기능장애 및 요법-관련 면역억제 둘 다의 결과로 만성 림프구성 백혈병(CLL), 다발성 골수종(MM), 비-호지킨 림프종(NHL) 및 다른 관련 악성종양 환자에서 발생할 수 있다. 이러한 혈액학적 악성종양에 이차적인 후천성 저감마글로불린혈증 환자, 및 HSPC 이식 후 환자는 박테리아 감염에 취약하다. 체액성 면역력의 결핍은 특히 캡슐화된 미생물에 의해 이러한 환자에서 감염-관련된 이환율 및 사망률 위험 증가에 주로 기인한다. 예를 들어, 폐렴연쇄구균, 헤모필루스 인플루엔자(Haemophilus influenzae), 및 스타필로코쿠스 아우레우스, 뿐만 아니라 레지오넬라 및 노카르디아 종(Nocardia spp.)은 CLL 환자에서 폐렴을 유발하는 빈번한 박테리아 병원체이다. 폐포자충(Pneumocystis carinii), 진균, 바이러스, 및 마이코박테리아와 같은 기회 감염이 또한 관찰되었다. 이러한 환자에서 감염의 수 및 중증도는 면역 글로불린의 투여에 의해 상당히 감소될 수 있다(Griffiths 등 *Blood* 73: 366-368, 1989; Chapel 등 *Lancet* 343: 1059-1063, 1994).

[0623] 특정 구현예에서, 제형은 급성 림프모구성 백혈병(ALL), 급성 골수형성 백혈병(AML), 부신백질이영양증, 병인불명 골수성 화생, 무거핵구성/선천성 혈소판감소증, 모세혈관 확장성운동실조, 중증성 β-지중해빈혈, 만성 육아종증, 만성 림프구성 백혈병(CLL), 만성 골수형성 백혈병(CML), 만성 골수단핵구 백혈병, 공통 가변성 면역 결핍증(CVID), 보체 장애, 선천성 무감마글로불린혈증, 다이아몬드 블랙판 증후군, 미만성 거대 B-세포 림프종, 가족성 혈구탐식성 림프조직구증, 여포성 림프종, 호지킨 림프종, 헐러 증후군, 초 IgM, IgG 하위클래스 결핍증, 청소년 골수단핵구 백혈병, 이염백질이영양증, 뮤코다당증, 다발성 골수종, 골수이형성, 비-호지킨 림프종, 발작성 야간 혜모글로빈뇨증(PNH), 항체 결핍된 1차 면역결핍증, 순수 적혈구 무형성증, 불응성 빈혈, 슈와크만-다이아몬드-블랙판(Shwachman-Diamond-Blackfan) 빈혈(DBA), 선택적 IgA 결핍증, 중증 재생불량성 빈혈, 겸상적혈구병, 특이적 항체 결핍증, 비스코트-올드리치 증후군, 및/또는 X-결합 무감마글로불린혈증(XLA)을 치료하기 위해 대상체에게 투여된다.

[0624] 치료될 수 있는 추가의 예시적인 암은 성상세포종, 비정형 기형 간상소체 종양, 뇌 및 중추신경계(CNS) 암, 유방암, 암육종, 연골육종, 척색종, 맥락총 암종, 맥락총 유두종, 연조직의 투명 세포 육종, 미만성 거대 B-세포 림프종, 뇌실막세포종, 상피양 육종, 생식선외 생식 세포 종양, 신장외 횡문양 종양, 유잉 육종, 위장관 간질 종양, 교모세포종, HBV-유도된 간세포 암종, 두경부암, 신장암, 폐암, 악성 간상소체 종양, 수모세포종, 흑색종, 수막종, 중피종, 다발성 골수종, 신경교 종양, 달리 명시되지 않은(NOS) 육종, 핍지교성상세포종, 희소돌기아교세포종, 골육종, 난소암, 난소 투명 세포 선암종, 난소 자궁내막 선암종, 난소 장액성 선암종, 췌장암, 췌장관 선암종, 췌장 내분비성 종양, 송과체모세포종, 전립선암, 신세포 암종, 신수질 암종, 횡문근육종, 육종, 신경집종, 피부 편평세포 암종, 및 줄기 세포암을 포함한다. 다양한 특정 구현예에서, 암은 난소암이다. 다양한 특정 구현예에서 암은 유방암이다.

[0625] 암의 맥락에서, 치료 유효량은 종양 세포 수를 감소시키고/시키거나, 전이 수를 감소시키고/시키거나, 종양 부피를 감소시키고/시키거나, 기대 수명을 증가시키고/시키거나, 암 세포의 세포자멸사를 유도하고/하거나, 암 세포 사멸을 유도하고/하거나, 암 세포에서 화학- 또는 방사성민감성을 유도하고/하거나, 암 세포 근처의 혈관형성을 억제하고/하거나, 암 세포 증식을 억제하고/하거나, 종양 성장을 억제하고/하거나, 전이를 예방하고/하거나, 대상체의 수명을 연장하고/하거나, 암-연관 통증을 감소시키고/시키거나, 전이 수를 감소시키고/시키거나, 치료 후 암의 재발 또는 재출현을 감소시킬 수 있다.

[0626] 특정 구현예는 외상, 바이러스, 화학요법, 독소, 및 오염에 의해 유발된 면역 결핍과 같은 이차성, 또는 후천성 면역 결핍의 치료를 포함한다. 이전에 나타낸 바와 같이, 후천성 면역결핍 증후군(AIDS)은 T 림프구의 고갈이 신체를 감염에 싸울 수 없게 만드는 인간 면역결핍 바이러스(HIV)와 같은 바이러스에 의해 유발된 2차 면역 결핍 장애의 예이다. 따라서, 또 다른 예로서, 유전자는 감염성 질환에 대한 치료적으로 효과적인 반응을 제공하

기 위해 선택될 수 있다. 특정 구현예에서, 감염성 질환은 인간 면역결핍 바이러스(HIV)이다. 치료 유전자는 예를 들어, 면역 세포가 HIV 감염에 대한 내성을 갖도록 하거나, 또는 면역 세포가 면역 재구축을 통해 바이러스를 효과적으로 중화할 수 있는 유전자, 면역 세포에 의해 발현되는 단백질을 암호화하는 유전자의 다형성, 환자에서 발현되지 않는 감염과 싸우는 데 유리한 유전자, 감염성 제제, 수용체 또는 공수용체를 암호화하는 유전자; 수용체 또는 공수용체에 대한 리간드를 암호화하는 유전자; 특정 전사 인자의 작용을 차단하기 위해 리보자임, 안티센스 RNA, 작은 간섭 RNA(siRNA) 또는 유인 RNA를 암호화하는 유전자; 우성 음성 바이러스 단백질, 세포내 항체, 인트라킨 및 자살 유전자를 암호화하는 유전자를 포함하는 바이러스 복제에 필수적인 바이러스 및 세포 유전자일 수 있다. 예시적인 치료 유전자 및 유전자 산물은 α2β1; αvβ3; αvβ5; αvβ63; BOB/GPR15; Bonzo/STRL-33/TYMSTR; CCR2; CCR3; CCR5; CCR8; CD4; CD46; CD55; CXCR4; 아미노펩티다제-N; HHV-7; ICAM; ICAM-1; PRR2/HveB; HveA; α-디스트로글리칸; LDLR/α2MR/LRP; PVR; PRR1/HveC; 및 라미닌 수용체를 포함한다. HIV의 치료를 위한 치료 유효량은 예를 들어, HIV에 대한 대상체의 면역력을 증가시키거나, AIDS 또는 HIV와 연관된 증상을 개선하거나, 또는 HIV에 대한 대상체의 선천적 또는 적응 면역 반응을 유도할 수 있다. HIV에 대한 면역 반응은 항체 생산 및 AIDS의 예방의 초래 및/또는 대상체의 AIDS 또는 HIV 감염 증상 개선, 또는 HIV 감염성 및/또는 병독성의 감소 또는 제거를 포함할 수 있다.

[0627] 특정 구현예, 제형은 암 재발을 예방 또는 지연시키거나 또는 고위험 생식 계열 돌연변이의 보균자에서 암 발병을 예방 또는 지연시키기 위해 대상체에게 투여된다. 특정 구현예에서, 제형은 더 높은 치료 용량의 테모졸로미드(TMZ) 및 벤질구아닌 또는 BCNU를 받기 위해 대상체에게 투여된다. 강한 골수억제성 표적외 효과로 인해, 유효량의 TMZ 및 벤질구아닌을 종양에 전달하는 것은 과제로 남아있다. 환자는 현재 급성 골수성 백혈병(AML), 식도암, 두경부암, 고등급 신경교종, 골수이형성 증후군, 비-소세포 폐암, NSCLC; 불응성 AML, 소세포 폐암, 역형성 성상세포종, 뇌 종양, 유방암(예를 들어, 전이성), 결장직장암(예를 들어, 전이성), 미만성 내재성 뇌간 신경교종, 유잉 육종, 다형성 교모세포종(GBM), 악성 뇌교종, 흑색종, 전이성 악성 흑색종, 재발성 악성 흑색종, 비인두암, 전이성 유방암, 및 소아암과 연관된 치료를 위해 TMZ 및 벤질구아닌을 받을 수 있다.

[0628] $MGMT^{P140K}$ 발현 종양이 있는 환자는 $MGMT^{P140K}$ 생체내 선택 카세트와 조합된 활성 성분(예컨대 CAR, TCR, 또는 체크포인트 억제제)과 함께 Ad35 바이러스 벡터의 투여로부터 이익을 얻을 것이다. 생체외 접근법은 이 접근법의 적용가능성을 제시하였다. 특정 구현예에서, TMZ 및 벤질구아닌 또는 BCNU의 치료량은 종양 부담 또는 부피를 감소시키기 위해 투여된다.

[0629] 특정 구현예에서, 치료 유효량은 면역 및 다른 혈액 세포에 기능을 제공하고/하거나, 면역-매개 병태를 감소 또는 제거하고/하거나; 면역-매개 병태의 증상을 감소 또는 제거할 수 있다.

[0630] 본원에 기재된 벡터, 동원 인자, 제형, 및 사용 방법에서, 단백질 및/또는 핵산 서열의 변이체가 또한 사용될 수 있다. 변이체는 본원에 기재되거나 또는 개시된 단백질 및 핵산 서열에 대해 적어도 70% 서열 동일성, 80% 서열 동일성, 85% 서열, 90% 서열 동일성, 95% 서열 동일성, 96% 서열 동일성, 97% 서열 동일성, 98% 서열 동일성, 또는 99% 서열 동일성을 갖는 서열을 포함하며 여기서 변이체는 실질적으로 유사하거나 또는 개선된 생물학적 기능을 나타낸다.

[0631] 본원에 기재된 생체내 유전자 요법 및/또는 HSPC 동원과 연관된 매개변수에 대해 수득된 값은 대조군 집단에서 유래된 참조 수준과 비교될 수 있고, 이 비교는 본원에 기재된 생체내 유전자 요법이 유전자 요법이 투여된 이를 필요로 하는 대상체에 효과적인지를 나타낼 수 있다. 생체내 유전자 요법 및/또는 HSPC 동원과 연관된 매개변수는 예를 들어 다음을 포함할 수 있다: 총 백혈구, 호중구, 단핵구, 림프구, 및/또는 혈소판의 수; 정상 림프구 계수에 도달하는 데 필요한 시간; 퍼센트 CD3+CD45RA+ T 세포; $10^6$ 개 세포 당 TREC의 수; CD4+인 세포의 퍼센트; CD8+인 세포의 퍼센트; CD4/CD8의 비; CD3+ T 세포에서 TCRαβ+ 세포의 퍼센트; TCR의 다양성; 유전자 교정된 림프구의 빈도; 유전자 교정된 림프구의 클론 레퍼토리의 다양성; 고유한 RIS 클론의 수; 박테리오파지 주사에 대한 1차 및 2차 항체 반응; 박테리오파지 불활성화 비; 유전자 교정된 세포의 백분율; 면역글로불린 IgA, IgM, 및/또는 IgG의 수준; 미토마이신 C에 대한 골수 유래 세포의 내성; 메틸셀룰로스 및 미토마이신 C에서 살아있는 세포의 퍼센트; γC-의존적 신호전달 경로의 기능성; 및 세포의 IL-21/미토겐 자극에 의한 STAT3의 인산화 퍼센트. 참조 수준은 대조군 집단의 하나 이상의 관련 데이터세트로부터 수득될 수 있다. 본원에 사용된 바와 같은 "데이터세트"는 원하는 조건 하에 샘플(샘플 집단)의 평가로 인한 수치 값 세트이다. 데이터세트의 값은 예를 들어, 샘플로부터 측정치를 실험적으로 수득하고 이러한 측정으로부터 데이터세트를 구축함으로써 수득될 수 있다. 당업자에 의해 이해되는 바와 같이, 참조 수준은 예를 들어, 개별 데이터포인트의 집합; 예를 들어, 평균, 중앙값, 평균의 중앙값 등으로부터 의미있는 집합적 참조 수준에 도달하기 위해 당업계에서 유용하

고 알려진 임의의 수학적 또는 통계적 공식을 기반으로 할 수 있다. 대안적으로, 참조 수준 또는 참조 수준을 생성하기 위한 데이터세트는 실험실과 같은 서비스 제공자로부터, 또는 데이터베이스 또는 데이터베이스가 저장된 서버로부터 수득될 수 있다.

[0632] 데이터세트로부터의 참조 수준은 대조군 집단에서 유래된 이전 측정치로부터 유래될 수 있다. "대조군 집단"은 유사한 명시된 특성의 대상체 또는 샘플의 임의의 그룹화이다. 그룹화는 예를 들어, 임상 매개변수, 임상 평가, 치료 레지멘, 질환 상태, 병태의 중증도 등에 따를 수 있다. 특정 구현예에서, 그룹화는 연령 범위(예를 들어, 0-2 세) 및 면역손상되지 않은 상태에 기반한다. 특정 구현예에서, 정상 대조군 집단은 테스트 대상체에 연령이 일치하고 면역손상되지 않은 개체를 포함한다. 특정 구현예에서, 연령 일치는 상황에 따라 임상적으로 관련되므로, 예를 들어, 0-6 개월; 0-1 년; 0-2 년; 0-3 년; 10-15 년을 포함한다).

[0633] 특정 구현예에서, 본원에 기재된 생체내 유전자 요법 및/또는 HSPC 동원과 연관된 특정 매개변수의 값에 대한 관련 참조 수준은 본원에 개시된 생체내 유전자 요법이 유전자 요법이 투여된 이를 필요로 하는 대상체에 대해 치료적으로 효과적인지를 결정하기 위해 대조군 집단에서 생체내 유전자 요법 및/또는 HSPC 동원과 연관된 특정 한 상응하는 매개변수의 값을 기반으로 수득된다.

[0634] 특정 구현예에서, 대조군 집단은 건강하고 면역 결핍이 없는 집단을 포함할 수 있다. 특정 구현예에서, 대조군 집단은 면역 결핍이 있고 치료 유효량의 (i) 치료 유전자와 연관된 Ad35 바이러스 벡터를 포함하는 제형; 및 (ii) 동원 인자가 투여되지 않은 집단을 포함할 수 있다. 특정 구현예에서, 대조군 집단은 면역 결핍이 있고 치료 유전자와 연관된 Ad35 바이러스 벡터를 포함하고 동원 인자를 포함하지 않는 제형이 치료 유효량으로 투여된 집단을 포함할 수 있다. 예로서, 관련 참조 수준은 대조군 대상체에서 생체내 유전자 요법 및/또는 HSPC 동원과 연관된 특정 매개변수의 값일 수 있다.

[0635] 특정 구현예에서, 결론은 샘플 값이 참조 수준과 통계적으로 유의하게 상이하거나 또는 통계적으로 유의하게 상이하지 않은지 여부에 기반하여 도출된다. 측정치는 차이가 우연에만 기반하여 발생할 것으로 예측될 수준 내에 있는 경우 통계적으로 유의하게 상이하지 않다. 대조적으로, 통계적으로 유의한 차이 또는 증가는 우연으로만 발생할 것으로 예상된 것보다 더 큰 것이다. 통계적 유의성 또는 이의 결여는 당업계에 잘 알려진 다양한 방법 중 임의의 것에 의해 결정될 수 있다. 통계적 유의성에 대해 통상적으로 사용되는 측정치의 예는 p-값이다. p-값은 특정 데이터포인트와 동등한 주어진 결과를 수득할 확률을 나타내며, 여기서 데이터포인트는 무작위 우연 단독의 결과이다. 결과는 종종 0.05 보다 작거나 또는 동일한 p-값에서 (무작위 우연이 아니라) 유의한 것으로 간주된다. 특정 구현예에서, 샘플 값은 샘플 값 및 참조 수준이 통계적으로 유의하게 상이하지 않은 경우 정상 대조군 집단에서 유래된 참조 수준과 "비슷하"다.

[0636] 특정 구현예에서, 본원에 기재된 생체내 유전자 요법 및/또는 HSPC 동원과 관련된 매개변수 및/또는 다른 데이터세트 구성요소에 대해 수득된 값은 선택된 매개변수를 사용한 분석 과정에 적용될 수 있다. 분석 과정의 매개변수는 본원에 개시된 것들 또는 본원에 기재된 지침을 사용하여 유래된 것들일 수 있다. 결과를 생성하는 데 사용되는 분석 과정은 샘플 분류에 유용한 결과를 제공할 수 있는 임의의 유형의 과정, 예를 들어, 수득된 값과 참조 수준의 비교, 선형 알고리즘, 이차 알고리즘, 결정 트리 알고리즘, 또는 투표 알고리즘일 수 있다. 분석 과정은 샘플이 주어진 클래스에 속할 확률을 결정하기 위한 임계값을 설정할 수 있다. 확률은 바람직하게는 적어도 60%, 적어도 70%, 적어도 80%, 적어도 90%, 적어도 95% 또는 그 이상이다.

[0637] 본원에 기재된 Ad35 벡터는 하기 예시적인 구현예 및 실시예에 기재된 Ad5/Ad35++ 벡터 대신에 활용될 수 있다.

[0638] 하기 예시적인 구현예 및 실시예(들)는 개시내용의 특정 구현예를 입증하기 위해 포함된다. 당업자는 많은 변화가 본원에 개시된 특이적 구현예로 이루어질 수 있고 개시내용의 사상 및 범위를 벗어나지 않고 여전히 비슷하거나 또는 유사한 결과를 수득할 수 있음을 본 개시내용에 비추어 인식하여야 한다.

[0639] **V. 예시적인 구현예**

[0640] 예시적인 구현예의 첫번째 세트는 하기를 포함할 수 있다:

[0641] 1. 하기를 포함하는, 재조합 아데노바이러스 혈청형 35(Ad35) 벡터 생산 시스템: 다음을 포함하는 재조합 Ad35 헬퍼 게놈: Ad35 섬유 샤프트를 암호화하는 핵산 서열; Ad35 섬유 놉을 암호화하는 핵산 서열; 및 Ad35 패키징 서열의 적어도 일부에 플랭킹하는 재조합효소 지시 반복부(DR), 및 다음을 포함하는 재조합 헬퍼 의존적 Ad35 공여자 게놈: 5' Ad35 도립된 말단 반복부(ITR); 3' Ad35 ITR; Ad35 패키징 서열; 및 적어도 하나의 이종 발현 산물을 암호화하는 핵산 서열.

[0642] 2. 하기를 포함하는, 재조합 아데노바이러스 혈청형 35(Ad35) 헬퍼 벡터: Ad35 섬유 샤프트; Ad35 섬유 놉; 및 Ad35 패키징 서열의 적어도 일부에 플랭킹하는 재조합효소 지시 반복부(DR)를 포함하는 Ad35 게놈.

[0643] 3. 하기를 포함하는, 재조합 아데노바이러스 혈청형 35(Ad35) 헬퍼 게놈: Ad35 섬유 샤프트를 암호화하는 핵산 서열; Ad35 섬유 놉을 암호화하는 핵산 서열; 및 Ad35 패키징 서열의 적어도 일부에 플랭킹하는 재조합효소 지시 반복부(DR).

[0644] 4. 하기를 포함하는, 재조합 헬퍼 의존적 아데노바이러스 혈청형 35(Ad35) 공여자 벡터: 다음을 포함하는 핵산 서열: 5' Ad35 도립된 말단 반복부(ITR); 3' Ad35 ITR; Ad35 패키징 서열; 및 적어도 하나의 이종 발현 산물을 암호화하는 핵산 서열, 여기서 게놈은 Ad35 바이러스 구조적 단백질을 암호화하는 핵산 서열을 포함하지 않는 것; 및 Ad35 섬유 샤프트 및/또는 Ad35 섬유 놉.

[0645] 5. 하기를 포함하는, 재조합 헬퍼 의존적 아데노바이러스 혈청형 35(Ad35) 공여자 게놈: 5' Ad35 도립된 말단 반복부(ITR); 3' Ad35 ITR; Ad35 패키징 서열; 및 적어도 하나의 이종 발현 산물을 암호화하는 핵산 서열, 여기서 Ad35 공여자 게놈은 야생형 Ad35 게놈에 의해 암호화된 발현 산물을 암호화하는 핵산 서열을 포함하지 않는 것.

[0646] 6. 재조합 헬퍼 의존적 아데노바이러스 혈청형 35(Ad35) 공여자 벡터를 생산하는 방법으로, 상기 방법은 세포의 배양물로부터 재조합 헬퍼 의존적 Ad35 공여자 벡터를 단리하는 단계를 포함하며, 여기서 세포는 Ad35 섬유 샤프트를 암호화하는 핵산 서열; Ad35 섬유 놉을 암호화하는 핵산 서열; 및 Ad35 패키징 서열의 적어도 일부에 플랭킹하는 재조합효소 지시 반복부(DR)를 포함하는 재조합 Ad35 헬퍼 게놈, 및 5' Ad35 도립된 말단 반복부(ITR); 3' Ad35 ITR; Ad35 패키징 서열; 및 적어도 하나의 이종 발현 산물을 암호화하는 핵산 서열을 포함하는 재조합 헬퍼 의존적 Ad35 공여자 게놈을 포함하는 것인, 방법.

[0647] 7. 하기를 포함하는, 재조합 아데노바이러스 혈청형 35(Ad35) 생산 시스템: 다음을 포함하는 재조합 Ad35 헬퍼 게놈: Ad35 섬유 샤프트를 암호화하는 핵산 서열; Ad35 섬유 놉을 암호화하는 핵산 서열; 및 Ad35 패키징 신호를 기능적으로 방해하지만 5' Ad35 도립된 말단 반복부(ITR)는 방해하지 않는 Ad35 게놈의 5' 단부의 550 개 뉴클레오티드 내의 재조합효소 지시 반복부(DR), 및 다음을 포함하는 재조합 Ad35 공여자 게놈:5' Ad35 ITR; 3' Ad35 ITR; Ad35 패키징 서열; 및 적어도 하나의 이종 발현 산물을 암호화하는 핵산 서열.

[0648] 8. 하기를 포함하는, 재조합 아데노바이러스 혈청형 35(Ad35) 헬퍼 벡터: Ad35 섬유 샤프트; Ad35 섬유 놉; 및 Ad35 패키징 신호를 기능적으로 방해하지만 5' Ad35 도립된 말단 반복부(ITR)는 방해하지 않는 Ad35 게놈의 5' 단부의 550 개 뉴클레오티드 내의 재조합효소 지시 반복부(DR)를 포함하는 Ad35 게놈.

[0649] 9. 하기를 포함하는, 재조합 아데노바이러스 혈청형 35(Ad35) 헬퍼 게놈: Ad35 섬유 샤프트를 암호화하는 핵산 서열; Ad35 섬유 놉을 암호화하는 핵산 서열; 및 Ad35 패키징 신호를 기능적으로 방해하지만 5' Ad35 도립된 말단 반복부(ITR)는 방해하지 않는 Ad35 게놈의 5' 단부의 550 개 뉴클레오티드 내의 재조합효소 지시 반복부(DR).

[0650] 10. 재조합 헬퍼 의존적 아데노바이러스 혈청형 35(Ad35) 공여자 벡터를 생산하는 방법으로, 상기 방법은 세포의 배양물로부터 재조합 헬퍼 의존적 Ad35 공여자 벡터를 단리하는 단계를 포함하며, 여기서 세포는 Ad35 섬유 샤프트를 암호화하는 핵산 서열; Ad35 섬유 놉을 암호화하는 핵산 서열; 및 Ad35 패키징 신호를 기능적으로 방해하지만 5' Ad35 도립된 말단 반복부(ITR)는 방해하지 않는 Ad35 게놈의 5' 단부의 550 개 뉴클레오티드 내의 재조합효소 지시 반복부(DR)를 포함하는 재조합 Ad35 헬퍼 게놈, 및 5' Ad35 ITR; 3' Ad35 ITR; Ad35 패키징 서열; 및 적어도 하나의 이종 발현 산물을 암호화하는 핵산 서열을 포함하는 재조합 Ad35 공여자 게놈을 포함하는 것인, 방법.

[0651] 11. 구현예 1-4 또는 6-10 중 어느 한 구현예에 있어서, 상기 Ad35 섬유 놉이 야생형 Ad35 섬유 놉을 포함하거나, 또는 Ad35 섬유 놉이 조작된 Ad35 섬유 놉을 포함하며, 여기서 조작된 섬유 놉은 CD46에 대한 섬유 놉의 친화도를 증가시키는 돌연변이를 포함하는 것인, 재조합 Ad35 벡터 생산 시스템, 헬퍼 벡터, 헬퍼 게놈, 공여자 벡터, 또는 방법.

[0652] 12. 구현예 11에 있어서, 상기 돌연변이가 Ile192Val, Asp207Gly(또는 Glu207Gly), Asn217Asp, Thr226Ala, Thr245Ala, Thr254Pro, Ile256Leu, Ile256Val, Arg259Cys, 및 Arg279His로부터 선택된 돌연변이를 포함하거나; 또는 돌연변이 Ile192Val, Asp207Gly(또는 Glu207Gly), Asn217Asp, Thr226Ala, Thr245Ala, Thr254Pro, Ile256Leu, Ile256Val, Arg259Cys, 및 Arg279His 각각을 포함하는 것인, 재조합 Ad35 벡터 생산 시스템, 헬퍼

벡터, 헬퍼 게놈, 공여자 벡터, 또는 방법.

[0653] 13. 구현예 1, 4-7, 또는 10-12 중 어느 한 구현예에 있어서, 상기 이종 발현 산물이 조절 서열과 작동가능하게 연결된 치료 발현 산물을 포함하며, 임의적으로 여기서 치료 발현 산물은 (a) β-글로빈 단백질 또는 γ-글로빈 단백질; (b) 항체 또는 이의 면역글로불린 쇄로, 임의적으로 여기서 항체는 항-CD33 항체를 포함하는 것; (c) 제1 항체 또는 이의 면역글로불린 쇄 및 제2 항체 또는 이의 면역글로불린 쇄로, 임의적으로 여기서 항체는 항-CD33 항체를 포함하는 것; (d) CRISPR-연관 RNA-가이드된 엔도뉴클레아제 및/또는 가이드 RNA(gRNA)로, 임의적으로 여기서 CRISPR-연관 RNA-가이드된 엔도뉴클레아제는 Cas9 또는 cpf1을 포함하는 것; (e) 염기 편집기 및/또는 gRNA로, 임의적으로 여기서 염기 편집기는 사이토신 염기 편집기(CBE) 또는 아데닌 염기 편집기(ABE)를 포함하며, 임의적으로 여기서 염기 편집기는 기능상실된 Cas9 및 기능상실된 cpf1로부터 선택된 촉매적으로 기능상실된 뉴클레아제를 포함하는 것; (f) 바이러스 감염을 차단하거나 또는 감소시키는 숙고 인자 또는 단백질로, 임의적으로 여기서 치료 발현 산물은 인자 VII 대체 단백질 또는 인자 VIII 대체 단백질을 포함하는 것; (g) 체크포인트 억제제; (h) 키메라 항원 수용체 또는 조작된 T 세포 수용체; 또는 (i) γC, JAK3, IL7RA, RAG1, RAG2, DCLRE1C, PRKDC, LIG4, NHEJ1, CD3D, CD3E, CD3Z, CD3G, PTPRC, ZAP70, LCK, AK2, ADA, PNP, WHN, CHD7, ORAI1, STIM1, CORO1A, CIITA, RFXANK, RFX5, RFXAP, RMRP, DKC1, TERT, TINF2, DCLRE1B, SLC46A1, FancA, FancB, FancC, FancD1, FancD2, FancE, FancF, FancG, FancI, FancJ, FancL, FancM, FancN, FancO, FancP, FancQ, FancR, FancS, FancT, FancU, FancV, FancW, 가용성 CD40, CTLA, Fas L, PD-L1에 대한 항체, CD4에 대한 항체, CD5에 대한 항체, CD7에 대한 항체, CD52에 대한 항체, IL-1에 대한 항체, IL-2에 대한 항체, IL-4에 대한 항체, IL-6에 대한 항체, IL-10에 대한 항체, TNF에 대한 항체, 자가반응성 T 세포 상에 특이적으로 존재하는 TCR에 대한 항체, 글로빈 패밀리 유전자, WAS, phox, 디스트로핀, 피루베이트 키나제, CLN3, ABCD1, 아릴술파타제 A, SFTPB, SFTPC, NLX2.1, ABCA3, GATA1, 리보솜 단백질 유전자, TERT, TERC, DKC1, TINF2, CFTR, LRRK2, PARK2, PARK7, PINK1, SNCA, PSEN1, PSEN2, APP, SOD1, TDP43, FUS, 유비퀼린 2, 및/또는 C9ORF72로 이루어진 군으로부터 선택된 단백질로, 임의적으로 여기서 단백질은 FancA 단백질을 포함하는 것을 포함하는, 재조합 Ad35 벡터 생산 시스템, 공여자 게놈, 공여자 벡터, 또는 방법.

[0654] 14. 구현예 13(d) 또는 13(e)에 있어서, 상기 gRNA가 HBG1, HBG2, 및/또는 적혈구 인핸서 bcl11a의 표적 핵산 서열에 결합하며, 임의적으로 여기서 gRNA는 γ-글로빈의 발현을 증가시키도록 조작되거나; 또는 상기 gRNA가 CD33의 일부를 암호화하는 표적 핵산 서열에 결합하며, 임의적으로 여기서 CD33은 인간 CD33을 포함하는 것인, 재조합 Ad35 벡터 생산 시스템, 공여자 게놈, 공여자 벡터, 또는 방법.

[0655] 15. 구현예 13에 있어서, 상기 치료 발현 산물이 β-글로빈 단백질 또는 γ-글로빈 단백질; 및 CRISPR-연관 RNA-가이드된 엔도뉴클레아제를 포함하는 CRISPR 시스템; 및 다음 중 1, 2, 또는 3 개를 포함하며: HBG1의 표적 핵산에 결합하는 gRNA; HBG2의 표적 핵산 서열에 결합하는 gRNA; 및/또는 Bcl11a의 표적 핵산 서열에 결합하는 gRNA, 임의적으로 여기서 gRNA는 γ-글로빈의 발현을 증가시키도록 조작되는 것인, 재조합 Ad35 벡터 생산 시스템, 공여자 게놈, 공여자 벡터, 또는 방법.

[0656] 16. 구현예 13에 있어서, 상기 조절 서열(들)이 프로모터를 포함하며, 임의적으로 여기서 프로모터는 β-글로빈 프로모터를 포함하며, 임의적으로 여기서 β-글로빈 프로모터는 약 1.6 kb의 길이를 갖고/갖거나 염색체 11의 위치 5228631-5227023에 따른 핵산을 포함하는 것인, 재조합 Ad35 벡터 생산 시스템, 공여자 게놈, 공여자 벡터, 또는 방법.

[0657] 17. 구현예 13에 있어서, 상기 조절 서열(들)이 유전자좌 제어 영역(LCR)을 포함하며, 임의적으로 여기서 LCR은 β-글로빈 LCR을 포함하는 것인, 재조합 Ad35 벡터 생산 시스템, 공여자 게놈, 공여자 벡터, 또는 방법

[0658] 18. 구현예 13에 있어서, 상기 β-글로빈 LCR이

[0659] HS1, HS2, HS3, 및 HS4를 포함하거나 또는 이로 이루어진 β-글로빈 LCR DNAse I 과민성 부위(HS)를 포함하며, 임의적으로 여기서 β-글로빈 LCR은 약 4.3 kb의 길이를 갖거나;

[0660] HS1, HS2, HS3, HS4, 및 HS5를 포함하는 β-글로빈 LCR DNAse I HS를 포함하며, 임의적으로 여기서 β-글로빈 LCR은 약 21.5 kb의 길이를 갖거나; 또는

[0661] 상기 β-글로빈 LCR이 염색체 11의 위치 5292319-5270789에 따른 서열을 포함하는 것인, 재조합 Ad35 벡터 생산 시스템, 공여자 게놈, 공여자 벡터, 또는 방법.

[0662] 19. 구현예 13 또는 14에 있어서, 상기 조절 서열(들)이 3'HS1을 포함하며, 임의적으로 여기서 3'HS1은 염색체 11의 위치 5206867-5203839에 따른 서열을 포함하는 것인, 재조합 Ad35 벡터 생산 시스템, 공여자 게놈, 공여자

벡터, 또는 방법.

[0663] 20. 구현예 13-19 중 어느 한 구현예에 있어서, 상기 조절 서열(들)이 miRNA 결합 부위를 포함하며, 임의적으로 여기서 miRNA 결합 부위는 관심 종에 의해 자연적으로 발현되는 miRNA에 대한 결합 부위를 포함하고/하거나; miRNA는 혈액 및 종양 미세환경 또는 표적 조직에서 차등 점유 프로파일을 입증하며, 임의적으로 여기서 점유 프로파일은 종양 미세환경 또는 표적 조직에서보다 혈액에서 더 높은 것이고/이거나; miRNA 결합 부위는 miR423-5, miR423-5p, miR42-2, miR181c, miR125a, 또는 miR15a 결합 부위를 포함하고/하거나; miRNA 결합 부위는 miR187 또는 miR218 결합 부위를 포함하는 것인, 재조합 Ad35 벡터 생산 시스템, 공여자 게놈, 공여자 벡터, 또는 방법.

[0664] 21. 구현예 1, 4-7, 또는 10-21 중 어느 한 구현예에 있어서, 상기 이종 발현 산물을 암호화하는 핵산이 통합 요소를 추가로 포함하는 페이로드의 일부이며, 임의적으로 여기서 통합 요소는 발현 산물을 포함하는 것인, 재조합 Ad35 벡터 생산 시스템, 공여자 게놈, 공여자 벡터, 또는 방법.

[0665] 22. 구현예 21에 있어서, 상기 통합 요소가 상동 재조합에 의해 표적 게놈 내로 통합을 위해 조작되고, 상기 통합 요소가 표적 게놈의 인접하게 연결된 서열에 상응하는 상동성 아암에 의해 플랭킹되며, 임의적으로 여기서 상동성 아암은 0.8 내지 1.8 kb인 것이고/이거나; 상동성 아암은 염색체 안전 항구 유전자좌에 플랭킹하는 표적 게놈의 핵산 서열에 상동성이며, 임의적으로 여기서 안전 항구 유전자좌는 AAVS1, CCR5, HPRT, 또는 Rosa로부터 선택되는 것인, 재조합 Ad35 벡터 생산 시스템, 공여자 게놈, 공여자 벡터, 또는 방법.

[0666] 23. 구현예 21에 있어서, 상기 통합 요소가 전위에 의해 표적 게놈 내로 통합을 위해 조작되고, 상기 통합 요소가 트랜스포존 도립 반복부(IR)에 의해 플랭킹되며, 임의적으로 여기서 트랜스포존 IR은 재조합효소 DR에 의해 플랭킹되는 것인, 재조합 Ad35 벡터 생산 시스템, 공여자 게놈, 공여자 벡터, 또는 방법.

[0667] 24. 구현예 23에 있어서, 상기 트랜스포존 IR이 잠자는 미녀(SB) IR이며, 임의적으로 여기서 SB IR은 pT4 IR이거나; 또는 트랜스포존 IR이 피기백, 마리너, 개구리 왕자, Tol2, TcBuster, 또는 spinON IR인, 재조합 Ad35 벡터 생산 시스템, 공여자 게놈, 공여자 벡터, 또는 방법.

[0668] 25. 구현예 21-24 중 어느 한 구현예에 있어서, 트랜스포존 IR에 의해 플랭킹된 통합 요소의 전위를 매개하는 트랜스포사제를 암호화하는 핵산을 포함하며, 임의적으로 여기서 트랜스포사제를 암호화하는 핵산은 지지 벡터 또는 지지 벡터 게놈으로 구성되는 것인, 재조합 Ad35 벡터 생산 시스템, 공여자 게놈, 공여자 벡터, 또는 방법.

[0669] 26. 구현예 25에 있어서, 상기 트랜스포사제가 잠자는 미녀, 피기백, 마리너, 개구리 왕자, Tol2, TcBuster, 또는 spinON 트랜스포사제를 포함하며, 임의적으로 여기서 트랜스포사제는 잠자는 미녀 100x(SB100x) 트랜스포사제인, 재조합 Ad35 벡터 생산 시스템, 공여자 게놈, 공여자 벡터, 또는 방법.

[0670] 27. 구현예 25 또는 26에 있어서, 상기 트랜스포사제를 암호화하는 핵산이 PGK 프로모터와 작동가능하게 연결되는 것인, 재조합 Ad35 벡터 생산 시스템, 공여자 게놈, 공여자 벡터, 또는 방법.

[0671] 28. 구현예 1-3 또는 6-27 중 어느 한 구현예에 있어서, 상기 Ad35 패키징 서열의 적어도 일부에 플랭킹되고/되거나 Ad35 게놈의 5' 단부의 550 개 뉴클레오티드 내에 있고 Ad35 패키징 신호를 기능적으로 방해하지만 5' Ad35 ITR은 방해하지 않는 재조합효소 DR이 FRT, loxP, rox, vox, AttB, 또는 AttP 부위인, 재조합 Ad35 벡터 생산 시스템, 헬퍼 벡터, 헬퍼 게놈, 또는 방법.

[0672] 29. 구현예 28에 있어서, 상기 Ad35 패키징 서열의 적어도 일부의 절제를 위한 재조합효소를 암호화하는 핵산이 헬퍼 게놈을 포함하는 세포의 핵산 서열에 의해 암호화되는 것인, 재조합 Ad35 벡터 생산 시스템, 헬퍼 벡터, 헬퍼 게놈, 또는 방법.

[0673] 30. 구현예 23-29 중 어느 한 구현예에 있어서, 상기 트랜스포존 IR에 플랭킹하는 재조합효소 DR이 FRT, loxP, rox, vox, AttB, 또는 AttP 부위인, 재조합 Ad35 벡터 생산 시스템, 헬퍼 벡터, 헬퍼 게놈, 또는 방법.

[0674] 31. 구현예 21-28 중 어느 한 구현예에 있어서, 상기 통합 요소를 포함하는 핵산의 절제를 위한 재조합효소를 암호화하는 핵산이 지지 벡터 또는 지지 벡터 게놈으로 구성되는 것인, 재조합 Ad35 벡터 생산 시스템, 헬퍼 벡터, 헬퍼 게놈, 또는 방법.

[0675] 32. 구현예 29 또는 31에 있어서, 상기 재조합효소가 Flp, Cre, Dre, Vika, 또는 PhiC31 재조합효소를 포함하는 것인, 재조합 Ad35 벡터 생산 시스템, 헬퍼 벡터, 헬퍼 게놈, 또는 방법.

[0676] 33. 구현예 32에 있어서, 상기 재조합효소를 암호화하는 핵산이 EF1α 프로모터와 작동가능하게 연결되는 것인, 재조합 Ad35 벡터 생산 시스템, 헬퍼 벡터, 헬퍼 게놈, 또는 방법.

[0677] 34. 구현예 21-33 중 어느 한 구현예에 있어서, 상기 페이로드가 이종 발현 산물을 포함하는 통합요소를 포함하며, 여기서 이종 발현 산물은 β-글로빈 프로모터 및 β-글로빈 긴 LCR과 작동가능하게 연결된 β-글로빈을 포함하고,

[0678] 여기서 통합 요소는 SB IR에 의해 플랭킹되며, 여기서 SB IR은 재조합효소 DR에 의해 플랭킹되며, 임의적으로 여기서 재조합효소 DR은 FRT 부위인, 재조합 Ad35 벡터 생산 시스템, 헬퍼 벡터, 헬퍼 게놈, 또는 방법.

[0679] 35. 구현예 21-34 중 어느 한 구현예에 있어서, 상기 페이로드가 통합 요소, 및 발현 산물을 암호화하고, 통합 요소로 구성되지 않고, 표적 게놈 내로 통합 효소의 통합에 의해 기능적이지 않게 만들도록 위치하는 조건부로 발현되는 핵산 서열을 포함하는 것인, 재조합 Ad35 벡터 생산 시스템, 헬퍼 벡터, 헬퍼 게놈, 또는 방법.

[0680] 36. 구현예 35에 있어서, 상기 조건부로 발현되는 핵산 서열에 의해 암호화되는 발현 산물이 CRISPR 시스템 구성요소 또는 염기 편집기 시스템 구성요소를 포함하며, 임의적으로 여기서 구성요소는 CRISPR-연관 RNA-가이드된 엔도뉴클레아제, 염기 편집기 효소, 또는 gRNA를 포함하는 것인, 재조합 Ad35 벡터 생산 시스템, 헬퍼 벡터, 헬퍼 게놈, 또는 방법.

[0681] 37. 구현예 21-36 중 어느 한 구현예에 있어서, 상기 페이로드가 선택 카세트를 포함하며, 임의적으로 여기서 선택 카세트는 통합 요소로 구성되는 것인, 재조합 Ad35 벡터 생산 시스템, 헬퍼 벡터, 헬퍼 게놈, 또는 방법.

[0682] 38. 구현예 37에 있어서, 상기 선택 카세트가 $mgmt^{P140K}$를 암호화하는 핵산 서열을 포함하거나 또는 상기 선택 카세트가 항-CD33 shRNA를 암호화하는 핵산 서열을 포함하는 것인, 재조합 Ad35 벡터 생산 시스템, 헬퍼 벡터, 헬퍼 게놈, 또는 방법.

[0683] 39. 구현예 1-3 또는 6-38 중 어느 한 구현예에 있어서, 상기 재조합효소 DR에 의해 플랭킹된 Ad35 패키징 서열의 적어도 일부가 GenBank 수탁 번호 AX049983에 따른 Ad35 서열의 뉴클레오티드 138-481에 상응하는 것인, 재조합 Ad35 벡터 생산 시스템, 헬퍼 벡터, 헬퍼 게놈, 또는 방법.

[0684] 40. 구현예 1-3 또는 6-38 중 어느 한 구현예에 있어서, 상기 재조합효소 DR에 의해 플랭킹된 Ad35 패키징 서열의 적어도 일부가 GenBank 수탁 번호 AX049983에 따른 Ad35 서열의 뉴클레오티드 179-344; 뉴클레오티드 366-481; 뉴클레오티드 155-481; 뉴클레오티드 159-480; 뉴클레오티드 159-446; 뉴클레오티드 180-480; 뉴클레오티드 207-480; 뉴클레오티드 140-446; 뉴클레오티드 159-446; 뉴클레오티드 180-446; 뉴클레오티드 202-446; 뉴클레오티드 159-481; 뉴클레오티드 180-384; 뉴클레오티드 180-481; 또는 뉴클레오티드 207-481에 상응하는 것인, 재조합 Ad35 벡터 생산 시스템, 헬퍼 벡터, 헬퍼 게놈, 또는 방법.

[0685] 41. 구현예 1-3 또는 6-40 중 어느 한 구현예에 있어서, 상기 재조합효소 DR이 LoxP 부위인, 재조합 Ad35 벡터 생산 시스템, 헬퍼 벡터, 헬퍼 게놈, 또는 방법.

[0686] 42. 구현예 2, 3, 8, 또는 9 중 어느 한 구현예에 있어서, 상기 Ad35 헬퍼 게놈이 293 T 세포에서 증폭을 위해 Ad5 E4orf6을 포함하는 것인, 헬퍼 벡터 또는 헬퍼 게놈.

[0687] 43. 구현예 2, 3, 8, 또는 9 중 어느 한 구현예에 있어서, 상기 헬퍼 게놈이 서열번호: 51-65 중 어느 하나에 제시된 바와 같은 서열을 포함하거나 또는 생성하는 것인, 헬퍼 벡터 또는 헬퍼 게놈.

[0688] 44. 구현예 2-5, 8, 또는 9 중 어느 한 구현예의 헬퍼 벡터, 헬퍼 게놈, 공여자 벡터, 또는 공여자 게놈을 포함하는 세포로, 임의적으로 여기서 세포는 HEK293 세포인, 세포.

[0689] 45. 구현예 1, 4, 6, 7, 10, 13-27 또는 44 중 어느 한 구현예에 있어서, 공여자 게놈을 포함하는 세포로, 임의적으로 여기서 세포는 적혈구이고, 임의적으로 여기서 세포는 조혈 줄기 세포, T-세포, B-세포, 또는 골수성 세포이고, 임의적으로 여기서 세포는 발현 산물을 분비하는 것인, 세포.

[0690] 46. 구현예 6 또는 10-41 중 어느 한 구현예에 있어서, 상기 세포가 HEK293 세포인, 방법.

[0691] 47. 세포를 변형시키는 방법으로, 상기 방법은 구현예 5 또는 11-27 중 어느 한 구현예에 따른 Ad35 공여자 벡터와 세포를 접촉시키는 단계를 포함하는 것인, 방법.

[0692] 48. 대상체의 세포를 변형시키는 방법으로, 상기 방법은 구현예 5 또는 11-27 중 어느 한 구현예에 따른 Ad35 공여자 벡터를 대상체에게 투여하는 단계를 포함하며, 임의적으로 여기서 방법은 대상체로부터 세포의 단리를

포함하지 않는 것인, 방법.

[0693] 49. 질환 또는 병태의 치료를 필요로 하는 대상체에서 질환 또는 병태를 치료하는 방법으로, 구현예 5 또는 11-27 중 어느 한 구현예에 따른 Ad35 공여자 벡터를 대상체에게 투여하는 단계를 포함하며, 임의적으로 여기서 투여는 정맥내 투여인, 방법.

[0694] 50. 구현예 49에 있어서, 상기 방법이 동원제를 대상체에게 투여하는 단계를 포함하며, 임의적으로 여기서 동원제는 과립구-콜로니 자극 인자, GM-CSF, S-CSF, CXCR4 길항제, 및 CXCR2 작용제 중 하나 이상을 포함하며, 임의적으로 여기서 CXCR4 길항제는 AMD3100을 포함하고/하거나 여기서 CXCR2 작용제는 GRO-β를 포함하는 것인, 방법.

[0695] 51. 구현예 49 또는 50에 있어서, 상기 Ad35 공여자 벡터가 선택 카세트를 포함하며, 임의적으로 여기서 방법은 대상체에게 선택제를 투여하는 단계를 추가로 포함하며, 임의적으로 여기서 선택 카세트는 $mgmt^{P140K}$를 암호화하고 선택제는 $O^6$BG/BCNU를 포함하는 것인, 방법.

[0696] 52. 구현예 49-51 중 어느 한 구현예에 있어서, 상기 방법이 면역 억제제를 대상체에게 투여하는 단계를 추가로 포함하며, 임의적으로 여기서 면역 억제 레지멘은 스테로이드, IL-6 수용체 길항제, 및/또는 IL-1R 수용체 길항제를 포함하며, 임의적으로 여기서 스테로이드는 글루코코르티코이드 또는 덱사메타손을 포함하는 것인, 방법.

[0697] 53. 구현예 49-52 중 어느 한 구현예에 있어서, 상기 Ad35 공여자 벡터가 통합 요소를 포함하고 방법이 CD46을 발현하는 세포의 적어도 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 또는 95%, 조혈 줄기 세포의 적어도 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 또는 95%, 및/또는 적혈구 $Ter119^+$ 세포의 적어도 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 또는 95%에서 이의 통합 요소의 카피의 통합 및/또는 발현을 유발하는 것인, 방법.

[0698] 54. 구현예 49-53 중 어느 한 구현예에 있어서, 상기 방법이 통합 요소의 적어도 1 개의 카피를 포함하는 표적 세포 게놈에서 통합 요소의 평균 적어도 2 개의 카피 또는 적어도 2.5 개의 카피의 통합을 유발하는 것인, 방법.

[0699] 55. 구현예 49-54 중 어느 한 구현예에 있어서, 상기 방법이 참조 수준의 적어도 약 20% 또는 참조 수준의 적어도 약 25%인 수준에서 페이로드에 의해 암호화된 발현 산물 또는 이의 통합 요소의 발현을 유발하며, 임의적으로 여기서 참조는 대상체 또는 참조 집단에서 내인성 참조 단백질의 발현인, 방법.

[0700] 56. 구현예 49-55 중 어느 한 구현예에 있어서, 상기 질환 또는 병태가 혈색소병증, 혈소판 장애, 빈혈, 면역결핍 응고 인자 결핍, 판코니 빈혈, 알파-1 항트립신 결핍, 겸상 세포 빈혈, 지중해빈혈, 중간형 지중해빈혈, A형 혈우병, B형 혈우병, 폰빌레브란트병, 인자 V 결핍, 인자 VII 결핍, 인자 X 결핍, 인자 XI 결핍, 인자 XII 결핍, 인자 XIII 결핍, 베르나르-술리에 증후군, 그레이 혈소판 증후군, 또는 뮤코다당류축적증을 포함하는 것인, 방법.

[0701] 57. 구현예 49-56 중 어느 한 구현예에 있어서, 상기 대상체가 암을 앓고 있는 대상체이고 상기 방법이 암을 치료, 예방, 또는 지연시키거나, 또는 암 재발을 지연시키며,

[0702] 임의적으로 여기서 대상체는 암의 발달과 연관된 하나 이상의 생식 계열 돌연변이의 보균자이며, 임의적으로 여기서 암은 역형성 성상세포종, 유방암, 난소암, 결장직장암, 미만성 내재성 뇌간 신경교종, 유잉 육종, 다형성 교모세포종, 악성 뇌교종, 흑색종, 전이성 악성 흑색종, 비인두암, 또는 소아암을 포함하며, 임의적으로 여기서 대상체는 $O^6$BG, TMZ(테모졸로미드), 및/또는 BCNU(카르무스틴)를 받거나 또는 투여받는 것인, 방법.

[0703] 58. 구현예 49-57 중 어느 한 구현예에 있어서, 상기 질환 또는 병태가 중간형 지중해빈혈을 포함하며, 임의적으로 여기서 벡터 또는 게놈은 내인성 γ-글로빈의 발현을 증가 또는 재활성화하는 발현 산물(들)로부터 선택된 하나 이상의 발현 산물을 암호화하는 핵산을 포함하며, 임의적으로 여기서 내인성 γ-글로빈의 발현을 증가 또는 재활성화하는 발현 산물(들)은 CRISPR-연관 RNA-가이드된 엔도뉴클레아제 또는 염기 편집기 및 HBG1의 핵산 서열에 결합하고 표적 핵산 서열과 작동가능하게 연결된 코딩 서열로부터 발현을 증가시키도록 조작된 gRNA; HBG2의 핵산 서열에 결합하고 표적 핵산 서열과 작동가능하게 연결된 코딩 서열로부터 발현을 증가시키도록 조작된 gRNA; 및 적혈구 인핸서 bcl11a의 핵산 서열에 결합하고 BCL11A의 발현을 감소시키도록 조작된 gRNA; γ-글로빈; 및 β-글로빈 중 하나 이상을 포함하며, 임의적으로 여기서 방법은 중간형 지중해빈혈의 증상을 감소시

키고/시키거나 중간형 지중해빈혈을 치료하고/하거나 HbF를 증가시키는 것인, 방법.

[0704] 예시적인 구현예의 두번째 세트는 하기를 포함할 수 있다:

[0705] 1. 조혈 줄기 세포의 생체내 유전자 편집을 위한 CD46을 표적화하는 재조합 혈청형 35 아데노바이러스(Ad35) 벡터.

[0706] 2. 구현예 1에 있어서, 상기 벡터의 섬유 놉 단백질이 CD46 결합을 증가시키는 돌연변이를 포함하는 것인, 재조합 Ad35 벡터.

[0707] 3. 구현예 2에 있어서, 상기 섬유 놉 단백질 돌연변이가 Asn217Asp, Thr254Pro, Ile256Leu, Asp207Gly(또는 Glu207Gly), Thr245Ala, Thr226Ala, Ile192Val, Ile256Val, Arg259Cys, 및 Arg279His 중 하나 이상으로부터 선택되는 것인, 재조합 Ad35 벡터.

[0708] 4. 구현예 2에 있어서, 상기 섬유 놉 단백질 돌연변이가 Asn217Asp, Thr254Pro, Ile256Leu, Asp207Gly(또는 Glu207Gly), Thr245Ala, Thr226Ala, Ile192Val, Ile256Val, Arg259Cys, 및 Arg279His를 포함하는 것인, 재조합 Ad35 벡터.

[0709] 5. 구현예 2에 있어서, 상기 섬유 놉 단백질 돌연변이가 Asn217Asp, Thr254Pro, Ile256Leu, Asp207Gly(또는 Glu207Gly), Thr245Ala, Thr226Ala, Ile192Val, Ile256Val, Arg259Cys, 및 Arg279His로 이루어지는 것인, 재조합 Ad35 벡터.

[0710] 6. 구현예 1에 있어서, 생체내 암호화된 유전자의 발현을 조절하는 miRNA 제어 시스템을 포함하는 것인, 재조합 Ad35 벡터.

[0711] 7. 구현예 6에 있어서, 상기 miRNA 제어 시스템이 혈액 및 종양 미세환경 또는 표적 조직에서 차등 점유 프로파일을 갖는 miRNA 표적 부위로 이루어지는 것인, 재조합 Ad35 벡터.

[0712] 8. 구현예 7에 있어서, 상기 점유 프로파일이 종양 미세환경 또는 표적 조직에서보다 혈액에서 더 높은 것인, 재조합 Ad35 벡터.

[0713] 9. 구현예 6에 있어서, 상기 miRNA 표적 부위가 miR423-5, miR423-5p, miR42-2, miR181c, miR125a, 및/또는 miR15a를 포함하는 것인, 재조합 Ad35 벡터.

[0714] 10. 구현예 6에 있어서, 상기 miRNA 표적 부위가 Cas9의 발현을 제어하는 것인, 재조합 Ad35 벡터.

[0715] 11. 구현예 6에 있어서, 상기 miRNA 표적 부위가 miR187, 및/또는 miR218을 포함하는 것인, 재조합 Ad35 벡터.

[0716] 12. 구현예 1에 있어서, DNA 파괴를 매개하고/하거나 내인성 유전자 발현을 활성화시키기 위해 CRISPR 구성요소를 암호화하는 뉴클레오티드를 포함하는 것인, 재조합 Ad35 벡터.

[0717] 13. 구현예 12에 있어서, 상기 CRISPR 구성요소가 뉴클레아제 및 가이드 RNA를 포함하는 것인, 재조합 Ad35 벡터.

[0718] 14. 구현예 13에 있어서, 상기 뉴클레아제가 Cas9 또는 cpf1을 포함하는 것인, 재조합 Ad35 벡터.

[0719] 15. 구현예 12에 있어서, 상기 CRISPR 구성요소가 촉매적으로 기능상실된 뉴클레아제를 포함하는 것인, 재조합 Ad35 벡터.

[0720] 16. 구현예 15에 있어서, 상기 촉매적으로 기능상실된 뉴클레아제가 기능상실된 Cas9 또는 기능상실된 cpf1을 포함하는 것인, 재조합 Ad35 벡터.

[0721] 17. 구현예 15에 있어서, 상기 촉매적으로 기능상실된 뉴클레아제가 가이드 RNA 및 사이티딘 또는 아데닌 데아미나제 또는 트랜스아미나제에 융합되는 것인, 재조합 Ad35 벡터.

[0722] 18. 구현예 13에 있어서, 상기 가이드 RNA가 HBG1 프로모터, HBG2 프로모터, 및/또는 bcl11a 인핸서에 결합하는 것인, 재조합 Ad35 벡터.

[0723] 19. 구현예 1에 있어서, 양성 선택 마커를 포함하는, 재조합 Ad35 벡터.

[0724] 20. 구현예 19에 있어서, 상기 양성 선택 마커가 항-CD33 shRNA 카세트 및/또는 $MGMT^{P140K}$ 카세트를 포함하는 것인, 재조합 Ad35 벡터.

[0725] 21. 구현예 1에 있어서, 상동성 아암을 포함하는, 재조합 Ad35 벡터.

[0726] 22. 구현예 21에 있어서, 상기 상동성 아암이 0.8 내지 1.8kb인, 재조합 Ad35 벡터.

[0727] 23. 구현예 21에 있어서, 상기 상동성 아암이 염색체 안전 항구 유전자좌에 특이적인 것인, 재조합 Ad35 벡터.

[0728] 24. 구현예 23에 있어서, 상기 염색체 안전 항구 유전자좌가 AAVS1, CCR5, HPRT, 또는 Rosa로부터 선택되는 것인, 재조합 Ad35 벡터.

[0729] 25. 구현예 1에 있어서, 트랜스포사제에 의해 인식되는 도립 반복 서열을 포함하는, 재조합 Ad35 벡터.

[0730] 26. 구현예 1에 있어서, 트랜스포사제를 암호화하는 뉴클레오티드 서열을 포함하는, 재조합 Ad35 벡터.

[0731] 27. 구현예 26에 있어서, 상기 트랜스포사제가 잠자는 미녀, 피기백, 마리너, 개구리 왕자, To12, TcBuster, 및 spinON을 포함하는 것인, 재조합 Ad35 벡터.

[0732] 28. 구현예 26에 있어서, 상기 트랜스포사제가 과활성 잠자는 미녀 트랜스포사제 또는 과활성 피기백 트랜스포사제를 포함하는 것인, 재조합 Ad35 벡터.

[0733] 29. 구현예 28에 있어서, 상기 과활성 잠자는 미녀 트랜스포사제가 SB100X를 포함하는 것인, 재조합 Ad35 벡터.

[0734] 30. 구현예 26에 있어서, 상기 트랜스포사제를 암호화하는 뉴클레오티드 서열이 PGK 프로모터의 전사 제어 하에 있는 것인, 재조합 Ad35 벡터.

[0735] 31. 구현예 1에 있어서, 재조합효소 인식 서열을 포함하는, 재조합 Ad35 벡터.

[0736] 32. 구현예 31에 있어서, 상기 재조합효소 인식 서열이 Frt, lox, rox, vox, AttB, 또는 AttP를 포함하는 것인, 재조합 Ad35 벡터.

[0737] 33. 구현예 1에 있어서, 재조합효소를 암호화하는 뉴클레오티드 서열을 포함하는, 재조합 Ad35 벡터.

[0738] 34. 구현예 33에 있어서, 상기 재조합효소가 Flp, Cre, Dre, Vika, 또는 PhiC31을 포함하는 것인, 재조합 Ad35 벡터.

[0739] 35. 구현예 33에 있어서, 상기 재조합효소를 암호화하는 뉴클레오티드 서열이 EF1α 프로모터의 전사 제어 하에 있는 것인, 재조합 Ad35 벡터.

[0740] 36. 구현예 1-35 중 어느 한 구현예에 있어서, 치료 카세트를 포함하는, 재조합 Ad35 벡터.

[0741] 37. 구현예 36에 있어서, 상기 치료 카세트가 치료 유전자를 포함하거나 또는 γC, JAK3, IL7RA, RAG1, RAG2, DCLRE1C, PRKDC, LIG4, NHEJ1, CD3D, CD3E, CD3Z, CD3G, PTPRC, ZAP70, LCK, AK2, ADA, PNP, WHN, CHD7, ORAI1, STIM1, CORO1A, CIITA, RFXANK, RFX5, RFXAP, RMRP, DKC1, TERT, TINF2, DCLRE1B, SLC46A1, FancA, FancB, FancC, FancD1(BRCA2), FancD2, FancE, FancF, FancG, FancI, FancJ(BRIP1), FancL, FancM, FancN(PALB2), FancO(RAD51C), FancP(SLX4), FancQ(ERCC4), FancR(RAD51), FancS(BRCA1), FancT(UBE2T), FancU(XRCC2), FancV(MAD2L2), FancW(RFWD3), 가용성 CD40, CTLA, Fas L, PD-L1에 대한 항체, CD4에 대한 항체, CD5에 대한 항체, CD7에 대한 항체, CD52에 대한 항체, IL-1에 대한 항체, IL-2에 대한 항체, IL-4에 대한 항체, IL-6에 대한 항체, IL-10에 대한 항체, TNF에 대한 항체, 자가반응성 T 세포 상에 특이적으로 존재하는 TCR에 대한 항체, 글로빈 패밀리 유전자, WAS, phox, 디스트로핀, 피루베이트 키나제, CLN3, ABCD1, 아릴술파타제 A, SFTPB, SFTPC, NLX2.1, ABCA3, GATA1, 리보솜 단백질 유전자, TERT, TERC, DKC1, TINF2, CFTR, LRRK2, PARK2, PARK7, PINK1, SNCA, PSEN1, PSEN2, APP, SOD1, TDP43, FUS, 유비퀼린 2, 및/또는 C9ORF72로부터 선택된 치료 유전자 산물을 암호화하는 것인, 재조합 Ad35 벡터.

[0742] 38. 구현예 36에 있어서, 상기 치료 카세트가 공통 감마(γ) 쇄, FancA, γ-글로빈, 및/또는 FVIII을 포함하거나 또는 암호화하는 치료 유전자를 포함하는 것인, 재조합 Ad35 벡터.

[0743] 39. 구현예 36에 있어서, 상기 치료 카세트가 키메라 항원 수용체, 조작된 T 세포 수용체, 및/또는 치료 항체를 암호화하는 치료 유전자를 포함하는 것인, 재조합 Ad35 벡터.

[0744] 40. 구현예 38에 있어서, 상기 치료 유전자가 β-글로빈 프로모터의 전사 제어 하에 있는 것인, 재조합 Ad35 벡터.

[0745] 41. 구현예 38에 있어서, 상기 치료 유전자가 HS1, HS2, HS3, 및 HS4로 이루어진 DNAse I 과민성 부위(HS)를

포함하는 β-글로빈 유전자좌 제어 영역(LCR)의 전사 제어 하에 있는 것인, 재조합 Ad35 벡터.

[0746] 42. 구현예 41에 있어서, 상기 β-글로빈 LCR가 대략 4.3 kb인, 재조합 Ad35 벡터.

[0747] 43. 구현예 41에 있어서, 상기 치료 유전자가 추가로 β-글로빈 프로모터의 전사 제어 하에 있는 것인, 재조합 Ad35 벡터.

[0748] 44. 구현예 43에 있어서, 상기 β-글로빈 프로모터가 대략 1.6 kb인, 재조합 Ad35 벡터.

[0749] 45. 구현예 44에 있어서, 상기 β-글로빈 프로모터가 염색체 11의 위치 5228631-5227023의 서열을 갖는 것인, 재조합 Ad35 벡터.

[0750] 46. 구현예 38에 있어서, 상기 치료 유전자가 HS1, HS2, HS3, HS4, 및 HS5를 포함하는 β-글로빈 긴 LCR의 전사 제어 하에 있는 것인, 재조합 Ad35 벡터.

[0751] 47. 구현예 46에 있어서, 상기 β-글로빈 긴 LCR이 대략 21.5 kb인, 재조합 Ad35 벡터.

[0752] 48. 구현예 47에 있어서, 상기 β-글로빈 긴 LCR이 염색체 11의 위치 5292319-5270789의 서열을 갖는 것인, 재조합 Ad35 벡터.

[0753] 49. 구현예 46에 있어서, 상기 치료 유전자가 추가로 β-글로빈 프로모터의 전사 제어 하에 있는 것인, 재조합 Ad35 벡터.

[0754] 50. 구현예 49에 있어서, 상기 β-글로빈 프로모터가 대략 1.6 kb인, 재조합 Ad35 벡터.

[0755] 51. 구현예 50에 있어서, 상기 β-글로빈 프로모터가 염색체 11의 위치 5228631-5227023의 서열을 갖는 것인, 재조합 Ad35 벡터.

[0756] 52. 구현예 46에 있어서, 3'HS1을 추가로 포함하는, 재조합 Ad35 벡터.

[0757] 53. 구현예 52에 있어서, 상기 3'HS1이 염색체 11의 위치 5206867-5203839의 서열을 갖는 것인, 재조합 Ad35 벡터.

[0758] 54. 구현예 1에 있어서, 적어도 30 kb 트랜스포존을 포함하는, 재조합 Ad35 벡터.

[0759] 55. 구현예 1에 있어서, 32.4 kb 트랜스포존을 포함하는, 재조합 Ad35 벡터.

[0760] 56. 구현예 1에 있어서, 헬퍼 바이러스를 사용하여 생성되는, 재조합 Ad35 벡터.

[0761] 57. 구현예 56에 있어서, 상기 헬퍼 바이러스가 293 T 세포에서 증폭을 위해 Ad5 E4orf6을 포함하는 것인, 재조합 Ad35 벡터.

[0762] 58. 구현예 56에 있어서, 상기 헬퍼 바이러스가 Ad35 신호전달 서열 및 패키징 서열을 포함하는 것인, 재조합 Ad35 벡터.

[0763] 59. 구현예 56에 있어서, 상기 헬퍼 바이러스가 바이러스 제조 동안 CRISPR 구성요소의 발현을 방지하기 위해 항-CRISPR(acr) 발현 카세트를 포함하는 것인, 재조합 Ad35 벡터.

[0764] 60. 구현예 56에 있어서, 상기 헬퍼 벡터가 서열번호: 51-64의 서열을 포함하거나 또는 생성하는 것인, 재조합 Ad35 벡터.

[0765] 61. 치료 단백질을 발현하도록 유전적으로 변형된 적혈구.

[0766] 62. 구현예 61에 있어서, 상기 치료 단백질이 바이러스 감염을 차단 또는 감소시키는 응고 인자 또는 단백질을 포함하는, 적혈구.

[0767] 63. 구현예 61에 있어서, 상기 적혈구가 치료 단백질을 분비하는 것인, 적혈구.

[0768] 64. 적혈구 bcl11a-인핸서 및 HBG 프로모터 영역을 동시에 표적화함으로써 HbF 재활성화를 증가시키기 위한, 구현예 1-63 중 어느 한 구현예의 재조합 Ad35 벡터 또는 적혈구의 용도.

[0769] 65. γ-글로빈 유전자 추가 및 내인성 γ-글로빈 유전자 재활성화의 조합을 위한, 구현예 1-63 중 어느 한 구현의 재조합 Ad35 벡터 또는 적혈구의 용도.

[0770] 66. 생체내 CRISPR 게놈 조작을 위한, 구현예 1-63 중 어느 한 구현예의 재조합 Ad35 벡터 또는 적혈구의 용도.

[0771] 67. 치료 유전자를 제공하기 위한, 구현예 1-63 중 어느 한 구현예의 재조합 Ad35 벡터 또는 적혈구의 용도.

[0772] 68. (i) 혈색소병증, (ii) 판코니 빈혈, (iii) 임의적으로 A형 혈우병, B형 혈우병, 또는 폰빌레브란트병으로부터 선택된 응고 인자 결핍, (iv) 혈소판 장애, (v) 빈혈, (vi) 알파-1 항트립신 결핍, 또는 (v) 면역 결핍을 치료하기 위한, 구현예 1-63 중 어느 한 구현예의 재조합 Ad35 벡터 또는 적혈구의 용도.

[0773] 69. 지중해빈혈을 치료하기 위한, 구현예 1-63 중 어느 한 구현예의 재조합 Ad35 벡터 또는 적혈구의 용도.

[0774] 70. 암을 치료하거나, 암 재발을 예방 또는 지연시키거나 또는 고위험 생식 계열 돌연변이의 보균자에서 암 발병을 예방 또는 지연시키기 위한 것이며, 임의적으로 여기서 암은 유방암 또는 난소암인, 구현예 1-63 중 어느 한 구현예의 재조합 Ad35 벡터 또는 적혈구의 용도.

[0775] 71. CRISPR/Cas9의 자기-불활성화를 위한, 구현예 1-63 중 어느 한 구현예의 재조합 Ad35 벡터 또는 적혈구의 용도.

[0776] 72. 자기-방출 카세트를 갖는 HDAd를 공여자 벡터로서 사용하여 표적화된 통합을 위한, 구현예 1-3 중 어느 한 구현예의 재조합 Ad35 벡터 또는 적혈구의 용도.

[0777] 73. 구현예 64 - 72 중 어느 한 구현예에 있어서, 동원을 포함하는, 용도.

[0778] 74. 구현예 49에 있어서, 상기 동원이 Gro-베타, GM-CSF, S-CSF, 및/또는 AMD3100의 투여를 포함하는, 용도.

[0779] 75. 구현예 64 - 72 중 어느 한 구현예에 있어서, 스테로이드, IL-6 수용체 길항제, 및/또는 IL-1R 수용체 길항제를 Ad35 벡터 및/또는 적혈구를 받는 대상체에게 투여하는 것을 포함하는, 용도.

[0780] 76. 구현예 75에 있어서, 상기 스테로이드가 글루코코르티코이드를 포함하는 것인, 용도.

[0781] 77. 구현예 75에 있어서, 상기 스테로이드가 덱사메타손을 포함하는 것인, 용도.

[0782] 78. 구현예 64-72 중 어느 한 구현예에 있어서, O6BG 및 TMZ(테모졸로미드) 또는 BCNU(카르무스틴)를 Ad35 벡터 및/또는 적혈구를 받는 대상체에게 투여하는 것을 포함하는, 용도.

[0783] 79. 구현예 78에 있어서, 상기 대상체가 역형성 성상세포종, 유방암, 결장직장암, 미만성 내재성 뇌간 신경교종, 유잉 육종, 다형성 교모세포종(GBM), 악성 뇌교종, 흑색종, 전이성 악성 흑색종, 비인두암, 또는 소아암에 대한 치료로서 O6BG 및 TMZ 또는 BCNU를 받는 것인, 용도.

[0784] 예시적인 구현예의 세번째 세트는 하기를 포함할 수 있다:

[0785] 1. 조혈 줄기 세포의 생체내 유전자 편집을 위한 CD46을 표적화하는 재조합 혈청형 35 아데노바이러스(Ad35) 벡터.

[0786] 2. 구현예 1에 있어서, 상기 벡터의 섬유 놉 단백질이 CD46 결합을 증가시키는 돌연변이를 포함하는 것인, 재조합 Ad35 벡터.

[0787] 3. 구현예 2에 있어서, 상기 섬유 놉 단백질 돌연변이가 Asn217Asp, Thr254Pro, Ile256Leu, Asp207Gly(또는 Glu207Gly), Thr245Ala, Thr226Ala, Ile192Val, Ile256Val, Arg259Cys, 및 Arg279His 중 하나 이상으로부터 선택되는 것인, 재조합 Ad35 벡터.

[0788] 4. 구현예 2에 있어서, 상기 섬유 놉 단백질 돌연변이가 Asn217Asp, Thr254Pro, Ile256Leu, Asp207Gly(또는 Glu207Gly), Thr245Ala, Thr226Ala, Ile192Val, Ile256Val, Arg259Cys, 및 Arg279His를 포함하는 것인, 재조합 Ad35 벡터.

[0789] 5. 구현예 2에 있어서, 상기 섬유 놉 단백질 돌연변이가 Asn217Asp, Thr254Pro, Ile256Leu, Asp207Gly(또는 Glu207Gly), Thr245Ala, Thr226Ala, Ile192Val, Ile256Val, Arg259Cys, 및 Arg279His로 이루어지는 것인, 재조합 Ad35 벡터.

[0790] 6. 구현예 1에 있어서, 생체내에서 암호화된 유전자의 발현을 조절하는 miRNA 제어 시스템을 포함하는, 재조합 Ad35 벡터.

[0791] 7. 구현예 6에 있어서, 상기 miRNA 제어 시스템이 혈액 및 종양 미세환경 또는 표적 조직에서 차등 점유 프로파일을 갖는 miRNA 표적 부위로 이루어지는 것인, 재조합 Ad35 벡터.

[0792] 8. 구현예 7에 있어서, 상기 점유 프로파일이 종양 미세환경 또는 표적 조직에서 보다 혈액에서 더 높은 것인,

재조합 Ad35 벡터.

[0793] 9. 구현예 6에 있어서, 상기 miRNA 표적 부위가 miR423-5, miR423-5p, miR42-2, miR181c, miR125a, 및/또는 miR15a를 포함하는 것인, 재조합 Ad35 벡터.

[0794] 10. 구현예 6에 있어서, 상기 miRNA 표적 부위가 Cas9의 발현을 제어하는 것인, 재조합 Ad35 벡터.

[0795] 11. 구현예 6에 있어서, 상기 miRNA 표적 부위가 miR187, 및/또는 miR218을 포함하는 것인, 재조합 Ad35 벡터.

[0796] 12. 구현예 1에 있어서, DNA 파괴를 매개하고/하거나 내인성 유전자 발현을 활성화하기 위해 CRISPR 구성요소를 암호화하는 뉴클레오티드를 포함하는, 재조합 Ad35 벡터.

[0797] 13. 구현예 12에 있어서, 상기 CRISPR 구성요소가 뉴클레아제 및 가이드 RNA를 포함하는 것인, 재조합 Ad35 벡터.

[0798] 14. 구현예 13에 있어서, 상기 뉴클레아제가 Cas9 또는 cpf1을 포함하는 것인, 재조합 Ad35 벡터.

[0799] 15. 구현예 12에 있어서, 상기 CRISPR 구성요소가 촉매적으로 기능상실된 뉴클레아제를 포함하는 것인, 재조합 Ad35 벡터.

[0800] 16. 구현예 15에 있어서, 상기 촉매적으로 기능상실된 뉴클레아제가 기능상실된 Cas9 또는 기능상실된 cpf1을 포함하는 것인, 재조합 Ad35 벡터.

[0801] 17. 구현예 15에 있어서, 상기 촉매적으로 기능상실된 뉴클레아제가 가이드 RNA 및 사이티딘 또는 아데닌 데아미나제 또는 트랜스아미나제에 융합되는 것인, 재조합 Ad35 벡터.

[0802] 18. 구현예 13에 있어서, 상기 가이드 RNA가 HBG1, HBG2, 및/또는 Bcl1a에 결합하는 것인, 재조합 Ad35 벡터.

[0803] 19. 구현예 1에 있어서, 양성 선택 마커를 포함하는, 재조합 Ad35 벡터.

[0804] 20. 구현예 19에 있어서, 상기 양성 선택 마커가 항-CD33 shRNA 카세트 및/또는 $MGMT^{P140k}$ 카세트를 포함하는 것인, 재조합 Ad35 벡터.

[0805] 21. 구현예 1에 있어서, 상동성 아암을 포함하는, 재조합 Ad35 벡터.

[0806] 22. 구현예 21에 있어서, 상기 상동성 아암이 0.8 내지 1.8 kb인, 재조합 Ad35 벡터.

[0807] 23. 구현예 21에 있어서, 상기 상동성 아암이 염색체 안전 항구 유전자좌에 특이적인 것인, 재조합 Ad35 벡터.

[0808] 24. 구현예 23에 있어서, 상기 염색체 안전 항구 유전자좌가 AAVS1, CCR5, HPRT, 또는 Rosa로부터 선택되는 것인, 재조합 Ad35 벡터.

[0809] 25. 구현예 1에 있어서, 트랜스포사제에 의해 인식되는 도립 반복 서열을 포함하는, 재조합 Ad35 벡터.

[0810] 26. 구현예 1에 있어서, 트랜스포사제를 암호화하는 뉴클레오티드 서열을 포함하는, 재조합 Ad35 벡터.

[0811] 27. 구현예 26에 있어서, 상기 트랜스포사제가 잠자는 미녀, 피기백, 마리너, 개구리 왕자, Tol2, TcBuster, 및 spinON을 포함하는 것인, 재조합 Ad35 벡터.

[0812] 28. 구현예 26에 있어서, 상기 트랜스포사제가 과활성 잠자는 미녀 트랜스포사제 또는 과활성 피기백 트랜스포사제를 포함하는 것인, 재조합 Ad35 벡터.

[0813] 29. 구현예 28에 있어서, 상기 과활성 잠자는 미녀 트랜스포사제가 SB100X를 포함하는 것인, 재조합 Ad35 벡터.

[0814] 30. 구현예 26에 있어서, 상기 트랜스포사제를 암호화하는 뉴클레오티드 서열이 PGK 프로모터의 전사 제어 하에 있는 것인, 재조합 Ad35 벡터.

[0815] 31. 구현예 1에 있어서, 재조합효소 인식 서열을 포함하는, 재조합 Ad35 벡터.

[0816] 32. 구현예 31에 있어서, 상기 재조합효소 인식 서열이 Frt, lox, rox, vox, AttB, 또는 AttP를 포함하는 것인, 재조합 Ad35 벡터.

[0817] 33. 구현예 1에 있어서, 재조합효소를 암호화하는 뉴클레오티드 서열을 포함하는, 재조합 Ad35 벡터.

[0818] 34. 구현예 33에 있어서, 상기 재조합효소가 Flp, Cre, Dre, Vika, 또는 PhiC31을 포함하는 것인, 재조합 Ad35

벡터.

[0819] 35. 구현예 33에 있어서, 상기 재조합효소를 암호화하는 뉴클레오티드 서열이 EF1α 프로모터의 전사 제어 하에 있는 것인, 재조합 Ad35 벡터.

[0820] 36. 구현예 1-35 중 어느 한 구현예에 있어서, 치료 카세트를 포함하는, 재조합 Ad35 벡터.

[0821] 37. 구현예 36에 있어서, 상기 치료 카세트가 치료 유전자를 포함하거나 또는 γC, JAK3, IL7RA, RAG1, RAG2, DCLRE1C, PRKDC, LIG4, NHEJ1, CD3D, CD3E, CD3Z, CD3G, PTPRC, ZAP70, LCK, AK2, ADA, PNP, WHN, CHD7, ORAI1, STIM1, CORO1A, CIITA, RFXANK, RFX5, RFXAP, RMRP, DKC1, TERT, TINF2, DCLRE1B, SLC46A1, FancA, FancB, FancC, FancD1(BRCA2), FancD2, FancE, FancF, FancG, FancI, FancJ(BRIP1), FancL, FancM, FancN(PALB2), FancO(RAD51C), FancP(SLX4), FancQ(ERCC4), FancR(RAD51), FancS(BRCA1), FancT(UBE2T), FancU(XRCC2), FancV(MAD2L2), FancW(RFWD3), 가용성 CD40, CTLA, Fas L, PD-L1에 대한 항체, CD4에 대한 항체, CD5에 대한 항체, CD7에 대한 항체, CD52에 대한 항체, IL-1에 대한 항체, IL-2에 대한 항체, IL-4에 대한 항체, IL-6에 대한 항체, IL-10에 대한 항체, TNF에 대한 항체, 자가반응성 T 세포 상에 특이적으로 존재하는 TCR에 대한 항체, 글로빈 패밀리 유전자, WAS, phox, 디스트로핀, 피루베이트 키나제, CLN3, ABCD1, 아릴술파타제 A, SFTPB, SFTPC, NLX2.1, ABCA3, GATA1, 리보솜 단백질 유전자, TERT, TERC, DKC1, TINF2, CFTR, LRRK2, PARK2, PARK7, PINK1, SNCA, PSEN1, PSEN2, APP, SOD1, TDP43, FUS, 유비퀼린 2, 및/또는 C9ORF72로부터 선택된 치료 유전자 산물을 암호화하는 것인, 재조합 Ad35 벡터.

[0822] 38. 구현예 36에 있어서, 상기 치료 카세트가 공통 감마(γ) 쇄, FancA, γ-글로빈, 및/또는 FVIII을 포함하거나 또는 암호화하는 치료 유전자를 포함하는 것인, 재조합 Ad35 벡터.

[0823] 39. 구현예 36에 있어서, 상기 치료 카세트가 키메라 항원 수용체, 조작된 T 세포 수용체, 및/또는 치료 항체를 암호화하는 치료 유전자를 포함하는 것인, 재조합 Ad35 벡터.

[0824] 40. 구현예 38에 있어서, 상기 치료 유전자가 β-글로빈 프로모터의 전사 제어 하에 있는 것인, 재조합 Ad35 벡터.

[0825] 41. 구현예 38에 있어서, 상기 치료 유전자가 HS1, HS2, HS3, 및 HS4로 이루어진 DNAse I 과민성 부위(HS)를 포함하는 β-글로빈 유전자좌 제어 영역(LCR)의 전사 제어 하에 있는 것인, 재조합 Ad35 벡터.

[0826] 42. 구현예 41에 있어서, 상기 β-글로빈 LCR이 대략 4.3 kb인, 재조합 Ad35 벡터.

[0827] 43. 구현예 41에 있어서, 상기 치료 유전자가 추가로 β-글로빈 프로모터의 전사 제어 하에 있는 것인, 재조합 Ad35 벡터.

[0828] 44. 구현예 43에 있어서, 상기 β-글로빈 프로모터가 대략 1.6 kb인, 재조합 Ad35 벡터.

[0829] 45. 구현예 44에 있어서, 상기 β-글로빈 프로모터가 염색체 11의 위치 5228631-5227023의 서열을 갖는 것인, 재조합 Ad35 벡터.

[0830] 46. 구현예 38에 있어서, 상기 치료 유전자가 HS1, HS2, HS3, HS4, 및 HS5를 포함하는 β-글로빈 긴 LCR의 전사 제어 하에 있는 것인, 재조합 Ad35 벡터.

[0831] 47. 구현예 46에 있어서, 상기 β-글로빈 긴 LCR이 대략 21.5 kb인, 재조합 Ad35 벡터.

[0832] 48. 구현예 47에 있어서, 상기 β-글로빈 긴 LCR이 염색체 11의 위치 5292319-5270789의 서열을 갖는 것인, 재조합 Ad35 벡터.

[0833] 49. 구현예 46에 있어서, 상기 치료 유전자가 추가로 β-글로빈 프로모터의 전사 제어 하에 있는 것인, 재조합 Ad35 벡터.

[0834] 50. 구현예 49에 있어서, 상기 β-글로빈 프로모터가 대략 1.6 kb인, 재조합 Ad35 벡터.

[0835] 51. 구현예 50에 있어서, 상기 β-글로빈 프로모터가 염색체 11의 위치 5228631-5227023의 서열을 갖는 것인, 재조합 Ad35 벡터.

[0836] 52. 구현예 46에 있어서, 3'HS1을 추가로 포함하는, 재조합 Ad35 벡터.

[0837] 53. 구현예 52에 있어서, 상기 3'HS1이 염색체 11의 위치 5206867-5203839의 서열을 갖는 것인, 재조합 Ad35 벡터.

[0838] 54. 구현예 1에 있어서, 적어도 30 kb 트랜스포존을 포함하는, 재조합 Ad35 벡터.

[0839] 55. 구현예 1에 있어서, 32.4 kb 트랜스포존을 포함하는, 재조합 Ad35 벡터.

[0840] 56. 구현예 1에 있어서, 헬퍼 바이러스를 사용하여 생성되는, 재조합 Ad35 벡터.

[0841] 57. 구현예 56에 있어서, 상기 헬퍼 바이러스가 293 T 세포에서 증폭을 위해 Ad5 E4orf6을 포함하는 것인, 재조합 Ad35 벡터.

[0842] 58. 구현예 56에 있어서, 상기 헬퍼 바이러스가 Ad35 신호전달 서열 및 패키징 신호를 포함하는 것인, 재조합 Ad35 벡터.

[0843] 59. 구현예 56에 있어서, 상기 헬퍼 바이러스가 바이러스 제조 동안 CRISPR 구성요소의 발현을 방지하기 위해 항-CRISPR(acr) 발현 카세트를 포함하는 것인, 재조합 Ad35 벡터.

[0844] 60. 구현예 56에 있어서, 상기 헬퍼 벡터가 서열번호: 51-65 중 임의의 하나의 서열을 포함하거나 또는 생성하는 것인, 재조합 Ad35 벡터.

[0845] 61. 치료 단백질을 발현하기 위해 유전적으로 변형된 적혈구.

[0846] 62. 구현예 61에 있어서, 상기 치료 단백질이 바이러스 감염을 차단 또는 감소시키는 응고 인자 또는 단백질을 포함하는 것인, 적혈구.

[0847] 63. 구현예 61에 있어서, 상기 적혈구가 치료 단백질을 분비하는 것인, 적혈구.

[0848] 64. 적혈구 bcl11a-인핸서 및 HBG 프로모터 영역을 동시에 표적화함으로써 HbF 재활성화를 증가시키기 위한, 구현예 1-63 중 어느 한 구현예의 재조합 Ad35 벡터 또는 적혈구의 용도.

[0849] 65. γ-글로빈 유전자 추가 및 내인성 γ-글로빈 유전자 재활성화의 조합을 위한, 구현예 1-63 중 어느 한 구현예의 재조합 Ad35 벡터 또는 적혈구의 용도.

[0850] 66. 생체내 CRISPR 게놈 조작을 위한, 구현예 1-63 중 어느 한 구현예의 재조합 Ad35 벡터 또는 적혈구의 용도.

[0851] 67. 치료 유전자를 제공하기 위한, 구현예 1-63 중 어느 한 구현예의 재조합 Ad35 벡터 또는 적혈구의 용도.

[0852] 68. (i) 혈색소병증, (ii) 판코니 빈혈, (iii) 임의적으로 A형 혈우병, B형 혈우병, 또는 폰빌레브란트병으로부터 선택된 응고 인자 결핍, (iv) 혈소판 장애, (v) 빈혈, (vi) 알파-1 항트립신 결핍, 또는 (v) 면역 결핍을 치료하기 위한, 구현예 1-63 중 어느 한 구현예의 재조합 Ad35 벡터 또는 적혈구의 용도.

[0853] 69. 지중해빈혈을 치료하기 위한, 구현예 1-63 중 어느 한 구현예의 재조합 Ad35 벡터 또는 적혈구의 용도.

[0854] 70. 암을 치료하거나, 암 재발을 예방 또는 지연시키거나 또는 고위험 생식 계열 돌연변이의 보균자에서 암 발병을 예방 또는 지연시키기 위한 것이며, 임의적으로 여기서 암은 유방암 또는 난소암인, 구현예 1-63 중 어느 한 구현예의 재조합 Ad35 벡터 또는 적혈구의 용도.

[0855] 71. CRISPR/Cas9의 자기-불활성화를 위한, 구현예 1-63 중 어느 한 구현예의 재조합 Ad35 벡터 또는 적혈구의 용도.

[0856] 72. 자기-방출 카세트가 있는 HDAd를 공여자 벡터로서 사용하여 표적화된 통합을 위한, 구현예 1-3 중 어느 한 구현예의 재조합 Ad35 벡터 또는 적혈구의 용도.

[0857] 73. 구현예 64 - 72 중 어느 한 구현예에 있어서, 동원을 포함하는, 용도.

[0858] 74. 구현예 49에 있어서, 상기 동원이 Gro-베타, GM-CSF, S-CSF, 및/또는 AMD3100의 투여를 포함하는 것인, 용도.

[0859] 75. 구현예 64 - 72 중 어느 한 구현예에 있어서, 스테로이드, IL-6 수용체 길항제, 및/또는 IL-1R 수용체 길항제를 Ad35 벡터 및/또는 적혈구를 받는 대상체에게 투여하는 것을 포함하는, 용도.

[0860] 76. 구현예 75에 있어서, 상기 스테로이드가 글루코코르티코이드를 포함하는 것인, 용도.

[0861] 77. 구현예 75에 있어서, 상기 스테로이드가 덱사메타손을 포함하는 것인, 용도.

[0862] 78. 구현예 64 - 72 중 어느 한 구현예에 있어서, $O^6$BG 및 TMZ(테모졸로미드) 또는 BCNU(카르무스틴)를 Ad35 벡

터 및/또는 적혈구를 받는 대상체에게 투여하는 것을 포함하는, 용도.

[0863] 79. 구현예 78에 있어서, 상기 대상체가 역형성 성상세포종, 유방암, 결장직장암, 미만성 내재성 뇌간 신경교종, 유잉 육종, 다형성 교모세포종(GBM), 악성 뇌교종, 흑색종, 전이성 악성 흑색종, 비인두암, 또는 소아암에 대한 치료로서 $O^6$BG 및 TMZ 또는 BCNU를 받는 것인, 용도.

[0864] **VI. 실험 실시예**

[0865] **실시예 1. 생체내 조혈 줄기 세포 유전자 요법은 뮤린 중간형 지중해빈혈을 개선시킨다.**

[0866] 이 실시예는 "건강한" 인간 CD46-유전자이식(CD46tg) 마우스에서 인간 γ-글로빈 유전자를 발현하는 통합 HDAd5/35++ 벡터를 이용하는 생체내 HSPC 유전자 요법 접근법을 예시하며; 개념 증명으로서, 이 접근법은 중간형 지중해빈혈에 대한 마우스 모델(CD46+/+/Hbbth-3 마우스)에서 예시된다. 이는 지중해빈혈에 대한 전통적인 렌티바이러스 벡터 생체외 유전자 요법에 대한 대안을 제공한다. 이 실시예에 함유된 정보 중 적어도 일부는 Wang 등, (*J Clin Invest*. 129(2):598-615, 2019; e-pub November 13, 2018)에 공개되었다.

[0867] 지중해빈혈은 전 세계적으로 인간에서 가장 흔한 선천적 질환 중 하나이며(Weatherall, *Ann N Y Acad Sci*. 1202:17-23, 2010), β-글로빈 쇄 합성이 없거나(β0/β0) 또는 결핍된다(β+/β+). 60,000 명의 어린이가 매년 주요 β-지중해빈혈을 가지고 태어난다. 치료가 없으면, 주요 지중해빈혈이 있는 어린이는 생후 10 년 내지 20 년 내에 사망한다. 헤모글로빈 사량체 형성을 위한 충분한 β-글로빈 쇄 합성이 없으면, 과도한 α-글로빈 쇄는 침전되고 골수에서 후기 적혈구모세포의 조기 사멸을 유발하거나 또는 순환 적혈구의 반감기를 감소시키는 내포물을 형성하여, β-지중해빈혈의 주요 혈액학적 특징인 비효과적인 적혈구생성 및 적혈구 사멸을 일으킨다. 생성된 빈혈은 조혈 구획의 확장을 자극하여, 적혈구 과형성 및 골수외 조혈을 생성한다.

[0868] 주요 β-지중해빈혈에 대한 주요 치료 양식은 적혈구(RBC)의 평생 수혈 및 과잉 철분을 제거하기 위한 킬레이트화를 사용한 지지적 치유; 또는 동종이계 조혈 줄기/전구체 세포(HSPC)의 이식을 사용한 치유적 치료로 이루어진다. 잘 일치하는 공여자가 없거나 또는 동종이계 HSPC 이식을 겪을 위험이 있는 환자의 경우, 렌티바이러스 벡터 야생형 β-글로빈 또는 태아 γ-글로빈 유전자 요법은 동종이계 이식의 면역학적 위험을 우회하는 치유 가능성이 있다. SIN-렌티바이러스 글로빈 벡터를 사용하여 마이크로-LCR 카세트를 혼입하는 HSPC 유전자 요법은 동물 모델 및 시험관내 환자 세포에서 β-지중해빈혈 및 겸상적혈구병(SCD) 표현형을 구제하였다(Pstaha 등, *Curr Gene Ther*. 17(5):364-378, 2017). 이에 기반하여, 지중해빈혈 및 SCD에 대한 다수의 임상 시험이 현재 유럽, 아시아, 및 미국에서 진행중이다(Pstaha 등, *Curr Gene Ther*. 17(5):364-378, 2017, Cavazanna-Calvo 등, *Nature*. 467(7313):318-322, 2010, Ferarri 등, *Hematology/Oncology Clinics of North America: Gene Therapy*. 31(5), Thompson 등, *N Engl J Med*. 378(16):1479-1493, 2018). 지금까지 이들 시험 데이터는 β+ 유전자형이 있는 대부분의 환자에 대한 장기간 수혈 독립성을 입증하지만, β0/β0 지중해빈혈의 치유는 여전히 과제로 남아있다.

[0869] 고무적인 임상 결과에도 불구하고, 현재 지중해빈혈 유전자 요법 프로토콜은 복잡하며, 백혈구분리반출술, 시험관내 배양, β- 또는 γ-글로빈 발현 카세트를 보유하는 렌티바이러스 벡터를 사용한 형질도입, 및 완전한 골수절제술로 조건화된 환자에게 재이식에 의한 공여자/환자로부터 HSPC의 수집을 수반한다. 기술적 복잡성 이외에도, 이 접근법의 다른 단점은 (a) 조혈 줄기 세포(HSC)의 다능성 및 이의 생착 가능성에 영향을 미칠 수 있는 다중 사이토카인의 존재 하에 배양의 필요성; (b) 혈색소병증 환자와 같이 만성 비악성 질환 및 기존 기관 손상이 있는 환자에서 골수절제술이 상당한 조혈 및 비조혈, 조기 또는 후기 독성과 연관된 중요한 위험 인자를 나타낼 때, 골수절제 레지멘에 대한 요건; 및 (c) 접근법에 대한 비용을 포함한다. 지중해빈혈이 자원 부족 국가에 만연하다는 사실은 더 간단하고 더 값싼 요법 접근법을 요구한다.

[0870] 백혈구분리반출술, 골수절제술, 및 HSPC 이식 없이 생체내 HSPC 유전자 전달을 위한 최소로 침습적이고 용이하게 번역가능한 접근법이 개발되었다(Richter 등, Blood. 2016;128(18):2206-2217, Richter 등, *Hematol Oncol Clin North Am*. 31(5):771-785, 2017, Ren 등, *Blood*. 128(18):2194-219, 2016). 이는 골수에서 말초 혈류 내로 HSPC를 동원하기 위한 G-CSF/AMD3100의 주사 및 통합 헬퍼-의존적 아데노바이러스(HDAd5/35++) 벡터 시스템의 정맥내 주사를 수반한다. HDAd5/35++ 벡터는 원시 HSC 상에서 발현되는 수용체인 인간 CD46을 표적화한다(Richter 등, *Blood*. 128(18):2206-2217, 2016). HDAd5/35++에서, 섬유 놉 도메인 및 샤프트를 제외한 모든 단백질은 혈청형 5로부터 유래되고; 섬유 놉 도메인 및 샤프트는 혈청형 35로부터 유래되고; CD46에 대한 친화도를 증가시키는 돌연변이는 Ad35 섬유 놉 및 ITR 내로 도입되고(WO 2010/0120541 참조) 패키징 신호는 Ad5로부터 유래된다. HdAd35++에서, 모든 단백질은 혈청형 35로부터 유래되고; CD46에 대한 친화도를 증가시키는 돌

연변이는 섬유 놉 및 ITR 내로 도입되고 패키징 신호는 Ad35로부터 유래된다.

[0871] 이식유전자 통합은 과활성 잠자는 미녀 트랜스포사제(SB100X)에 의해 무작위 패턴으로 달성된다(Mates 등, *Nat Genet*. 41(6):753-761, 2009). GFP를 리포터 유전자로 사용하는 마우스 모델에서, HSPC가 주변부에서 골수로 되돌아가 다시 형질도입되어, 생체내-형질도입된 마우스 및 2차 수용자에서 장기간 리포터를 지속하고 안정되게 발현함을 입증하였다(Richter 등, Blood. 2016;128(18):2206-2217).

[0872] 지중해빈혈을 표현형적으로 교정하기 위해 필요한 높은 수준의 이식유전자 마킹을 고려하면, 생체내 HSPC 형질도입 접근법은 $MGMT^{P140K}$ 발현 카세트를 HDAd5/35++ 벡터 내로 삽입함으로서 최적화되었다. 이는 저용량의 메틸화제, 예를 들어, $O^6$-벤질구아닌($O^6$BG) 및 비스-클로로에틸니트로소우레아(BCNU) 또는 테모졸로미드로 유전자-교정된 전구체의 생체내 선택을 허용한다(Beard 등, *J Clin Invest*. 120(7):2345-2354, 2010, Larochelle 등, *J Clin Invest*. 119(7):1952-1963, 2009, Trobridge 등, *PLoS One*. 7(9):e45173, 2012). 이전에 조합된 생체내 형질도입/선택 접근법은 안전하고 2차 수용체에서 유지되는 수준인 말초 혈액 세포의 최대 80%에서 안정된 GFP 발현을 초래하였음이 제시되었으며, 이는 자기-재생, 다중계통, 장기간 재증식 HSC의 안정된 형질도입을 나타낸다(Wang 등, *Mol Ther Methods Clin Dev*. 8:52-64, 2018).

[0873] 본원에서는 "건강한" 인간 CD46-유전자이식(CD46tg) 마우스 및, 개념 증명으로서, 중간형 지중해빈혈에 대한 마우스 모델(CD46+/+/Hbbth-3 마우스)에서 인간 γ-글로빈 유전자를 발현하는 통합 HDAd5/35++ 벡터를 사용하여 생체내 HSPC 유전자 요법 접근법을 테스트하였다.

[0874] ***재료 및 방법*. 시약**. 하기 시약을 사용하였다: G-CSF(Neupogen, Amgen), AMD3100(Sigma-Aldrich), 플레릭사포르(Mozobil, Genzyme Corp.), $O^6$-BG 및 BCNU(Sigma-Aldrich), 마이코페놀레이트 모페틸(CellCept Intravenous, Genentech), 라파마이신(Rapamune/Sirolimus, Pfizer), 및 메틸프레드니솔론 (Pfizer).

[0875] **HDAd 벡터**. 트랜스포존 벡터 HDAd-γ-글로빈/mgmt 및 SB100X-발현 인간 배아 신장-293 세포-유래 116 세포의 생성(Palmer 등, Gene Therapy Protocols. Volume 1: Production and In vivo Applications of Gene Transfer Vectors (Methods in Molecular Biology):33-53, 2009)은 이전에 기재되었다(Li 등, *Mol Ther Methods Clin Dev*. 9:142-152, 2018). 헬퍼 바이러스 오염 수준은 0.05% 미만인 것으로 밝혀졌다. 역가는 6 x $10^{12}$ 내지 12 x $10^{12}$ vp/ml였다. 이 연구에 사용되는 모든 HDAd 벡터는 Ad5 섬유 꼬리, Ad35 섬유 샤프트, 및 친화도-향상된 Ad35++ 섬유 놉으로 구성된 키메라 섬유를 함유한다(Wang 등, *J Virol*. 82(21):10567-10579, 2008). 모든 HDAd 제제는 $10^{10}$ vp에서 1 개 미만 카피의 야생형 바이러스를 가졌다(다른 곳에 기재된 프라이머를 사용하여 qPCR에 의해 측정됨; Haeussler 등, *PLoS One*. 6(8):e23160, 2011).

[0876] **인간 γ-글로빈 발현을 검출하는 세포내 유세포 분석**. FIX 및 PERM 세포 투과화 키트(Thermo Fisher Scientific)를 사용하였고, 제조업체의 프로토콜에 따랐다. 간단히 말해서, 1 X $10^6$ 개 세포를 100 μl FACS 완충액(1% FCS가 보충된 PBS)에 재현탁하고, 100 μl 시약 A(고정 배지)를 첨가하고 실온에서 2-3 분 동안 배양한 다음, 미리 냉각된 절대 메탄올 1 ml를 첨가하고, 혼합하고, 10 분 동안 암실에서 얼음 위에서 배양하였다. 그런 다음 샘플을 FACS 완충액으로 세척하고 100 μl 시약 B(투과 배지) 및 1 μg γ-글로빈 항체(Santa Cruz Biotechnology, 카탈로그 sc-21756 PE)에 재현탁하고, 실온에서 30 분 동안 배양하였다. 세척 후, 세포를 FACS 완충액에 재현탁하고 분석하였다. 적혈구 및 γ-글로빈 이중 염색을 위해, 세포를 먼저 APC 항-마우스 Ter119 항체(Ter119-APC, BioLegend, 카탈로그 116212)로 염색한 다음, 세척하고 상기 기재된 바와 같은 고정 배지로 고정하였다.

[0877] **글로빈 HPLC**. 개별 글로빈 쇄 수준을 SPD-10AV 다이오드 어레이 검출기 및 LC-10AT 바이너리 펌프가 장착된 Shimadzu Prominence 기기(Shimadzu)에서 정량화하였다. 물/아세토니트릴 중 0.1% 트리플루오로아세트산의 38%-60% 구배 혼합물을 Vydac C4 역상 칼럼(Hichrom)을 사용하여 1 ml/분의 속도로 적용하였다.

[0878] **실시간 역전사 PCR**. 총 RNA를 제조업체의 페놀-클로로포름 추출 방법에 따라 TRIzol™ 시약(Thermo Fisher Scientific, Cat. # 15596026)을 사용하여 50-100 μl 혈액으로부터 추출하였다. QuantiTect 역전사 키트(Qiagen, Cat. # 205311) 및 Power SYBR Green PCR 마스터 믹스(Thermo Fisher Scientific, Cat.# 4367659)를 사용하였다. 실시간 정량적 PCR을 StepOnePlus 실시간 PCR 시스템(Applied Biosystems)에서 수행하였다. 다음 프라이머 쌍을 이 작업에 사용하였다: 마우스 RPL10 정방향(서열번호: 189), 및 역방향(서열번호: 190); 인간 γ-글로빈 정방향(서열번호: 191), 및 역방향(서열번호: 192); 마우스 β-주요 글로빈 정방향(서열번호:

193), 및 역방향(서열번호: 194).

[0879] **자기 세포 분류**. 계통-수임 세포의 고갈을 위해, 마우스 계통 세포 고갈 키트(Miltenyi Biotec, Cat. # 130-090-858)를 제조업체의 지침에 따라 사용하였다. 1차 CD46+/+/Hbbth-3 마우스의 골수로부터의 Ter119+ 세포 또는 2차 C57BL/6 수용자의 조혈 조직으로부터의 CD46+ 세포의 선택을 위해, 인간 항-CD46-PE 1차 항체(Miltenyi Biotec, 카탈로그 130-104-508)로 염색한 후 마우스 항-Ter119 마이크로비드(Miltenyi Biotec, 카탈로그 130-049-901) 또는 항-PE 마이크로비드(Miltenyi Biotec, 카탈로그 130-048-801)를 각각 사용하였다.

[0880] **동물 연구**. 인간 CD46 게놈 유전자좌(CD46tg)에 대해 동형접합성이고 인간에서와 유사한 수준 및 패턴으로 CD46 발현을 제공하는 C57BL/6-기반 유전자이식 마우스는 이전에 기재되었다(Kemper 등, Clin Exp Immunol. 2001;124(2):180-189). CD46tg 마우스는 Roberto Cattaneo, Mayo Clinic(미국 미네소타주 로체스터 소재)에서 제공받았다. HDAd5/35++ 벡터에 의한 감염에 취약한 지중해빈혈 마우스 모델은 암컷 CD46tg 마우스와 수컷 Hbbth-3 마우스(The Jackson Laboratory)를 교배하고 F1과 CD46tg 마우스를 역교배하여, CD46+/+/Hbbth-3 마우스를 생성함으로써 수득하였다. 6 내지 10 주령의 암컷 CD46tg 및 CD46+/+/Hbbth-3 암컷을 생체내 형질도입/선택 연구에 사용하였다. 6 내지 10 주령의 암컷 C57BL/6 마우스를 2차 수용자로서 사용하였다.

[0881] **CD46tg 마우스의 동원 및 생체내 형질도입**. HSPC를 인간 재조합 G-CSF의 s.c. 주사(하루에 마우스 당 5 μg, 4 일) 이어서 5 일째에 AMD3100의 s.c. 주사(5 mg/kg)에 의해 마우스에 동원하였다. 또한, 동물은 바이러스 주사 16 시간 및 2 시간 전에 덱사메타손(10 mg/kg)을 i.p.로 투여받았다. AMD3100 후 30 및 60 분에, 동물에게 HDAd-ɣ-글로빈/mgmt 및 HDAd-SB를 주사 당 4 x $10^{10}$ vp의 용량으로 안와하정맥총을 통해 i.v. 주사하였다(총 2 회 주사, 30 분 간격). 4 주 후, 마우스에게 $O^6$-BG(15 mg/kg, i.p.)를 2 회, 30 분 간격으로 주사하였다. $O^6$-BG의 두번째 주사 1 시간 후, 마우스에게 BCNU(5 mg/kg, i.p.)를 주사하였다. BCNU 용량을 두번째 및 세번째 주기에 각각 7.5 및 10 mg/kg으로 증가시켰다.

[0882] **CD46+/+/ Hbbth-3 마우스의 동원 및 생체내 형질도입**. 이들 연구에서, 이전에 지중해빈혈 마우스 모델에 대해 기재된 바와 같이(Psatha 등, *Hum Gene Ther Methods*. 25(6):317-327, 2014), G-CSF 250 μg/kg i.p.(1-6 일) 및 플레릭사포르 5 mg/kg i.p.(이전에 AMD3100; Mozobil, Genzyme Corp.)(5-7 일)를 사용한 7-일 동원 접근법을 적용하였다. 생체내 형질도입을 상기와 같이 수행하였다. 처리 후, 조합된 면역억제를 투여하였다. 17 주차에, 마우스를 $O^6$BG(30 mg/kg, i.p.)를 사용한 4 주기의 생체내 선택에 적용하고 BCNU 용량(5, 7.5, 10, 10 mg/kg)을 용량 사이에 2-주 간격으로 증가시켰다. 면역억제는 마지막 $O^6$-BG/BCNU 용량 후 2 주에 재개하였다.

[0883] **면역억제**. 마이코페놀레이트 모페틸(20 mg/kg/d), 라파마이신(0.2 mg/kg/d), 및 메틸프레드니솔론(20 mg/kg/d)의 1 일 i.p. 주사를 수행하였다.

[0884] **2차 골수 이식**. 수용자는 Jackson Laboratory의 6-8 주령의 암컷 C57BL/6 마우스였다. 이식 당일에, 수용자 마우스를 10 Gy로 조사하였다. 생체내-형질도입된 CD46tg 마우스로부터의 골수 세포를 무균상태로 단리하고, 계통-고갈된 세포를 자기 세포 분류(MACS)를 사용하여 단리하였다. 조사 4 시간 후, 세포를 마우스 당 1 x $10^6$ 개 세포로 i.v. 주사하였다. CD46+/+/Hbbth-3 마우스 연구에서, 생체내-형질도입된 CD46+/+/Hbbth-3 마우스로부터의 2X$10^6$ 개의 전체 골수 세포를 3 개의 일일 용량 또는 치명적인 TBI(1,000 cGy)로 나눠진 100 mg/kg i.p. 부술판(Busilvex, Pierre Fabre)으로 조건화된 골수하제거된 2차 C57BL/6 수용자에 이식하였다. 20 주차에, 2차 수용자를 희생시키고, 혈액, 골수, 및 비장의 CD46+ 세포를 MACS에 의해 단리하거나, 또는 마우스를 상기 기재된 바와 같이 동원 및 생체내 형질도입에 적용하였다. 모든 2차 수용자는 4 주차에 시작하여 면역억제를 받았다.

[0885] **조직 분석**. 2.5 μm 두께의 비장 및 간 조직 절편을 적어도 24 시간 동안 4% 포름알데하이드에 고정시키고, 탈수화하고, 파라핀에 포매하였다. H&E 염색을 골수외 혈구형성의 조직학적 평가에 사용하였다. 헤모시데린을 Perl의 프러시안 블루(Prussian blue) 염색에 의해 조직 절편에서 검출하였다. 간단히 말해서, 조직 절편을 증류수 중 동일한 부피(2%)의 페로시안화칼륨 및 염산의 혼합물로 처리한 다음 뉴트럴 레드(neutral red)로 대조염색하였다. 비장 크기를 비장 중량(mg) 대 체중(g)의 비로 평가하였다.

[0886] **혈액 분석 및 골수 사이토스핀**. 혈액 샘플을 EDTA-코팅된 튜브에 수집하고, HemaVet 950FS(Drew Scientific) 또는 ProCyteDx(IDEXX) 기계에서 분석을 수행하였다. 말초 혈액 도말을 준비하고 각각 5 및 15 분 동안 May-Grunwald/Giemsa(Merck)로 염색하였다. 골수 세포의 현탁액을 사이토스핀 장치를 사용하여 슬라이드에서 원심

분리하고 May-Grunwald/Giemsa로 염색하였다.

[0887] **통계**. 데이터는 평균 ± SEM으로 제시된다. 다중 그룹의 비교를 위해, 다중 비교를 위한 본페로니 사후-테스트와 함께 일원 및 이원 ANOVA를 사용하였다. 1 개의 그룹화 변수에 대한 그룹 간 차이를 독립표본 양측 스튜던트 t 검정에 의해 결정하였다. 비모수 분석을 위해, Kruskal-Wallis 검정을 사용하였다. 통계적 분석은 GraphPad Prism 버전 6.01(GraphPad Software Inc.)을 사용하여 수행하였다. 0.05 미만의 $P$ 값이 유의한 것으로 간주되었다; $*P \leq 0.05$, $**P \leq 0.0002$, $***P \leq 0.00003$.

[0888] **동물 연구 승인**. 모든 실험은 통제 기관 심의 위원회 및 IACUC의 승인에 따라 수행되었다.

[0889] ***결과***. CD46tg 마우스에서 후속 생체내 선택과 함께 생체내 HSPC 형질도입은 대부분의 말초 RBC에서 안정된 γ-글로빈 발현을 초래한다. 치료용 HDAd5/35++ 벡터는 적혈구에서 효율적인 발현을 위한 5-kb "마이크로" β-글로빈 LCR/β-프로모터의 제어 하에 인간 γ-글로빈 유전자 뿐만 아니라 $MGMT^{P140K}$ 발현 카세트를 함유한다(도 2a, HDAd-γ-글로빈/mgmt). CD46tg 마우스는 인간과 유사한 패턴 및 수준에서 HDAd5/35++ 수용체 CD46을 발현하는 인간 CD46 유전자좌에 대해 동형접합성이고 따라서 생체내 HSPC 형질도입 연구를 위한 모델이다(Richter 등, *Blood*. 128(18):2206-2217, 2016, Kemper 등, *Clin Exp Immunol*. 124(2):180-189, 2001). "건강한" CD46tg 마우스에서 이러한 연구의 목표는 마우스 세포 상에서 인간 γ-글로빈의 수준, 동역학, 및 분포 및 접근법의 안전성을 분석하는 것이었다. 동물을 G-CSF/AMD3100으로 동원한 다음 HDAd-γ-글로빈/mgmt 및 SB100X-발현 HDAd-SB 벡터를 정맥내로 주사하였다. 3 주기의 $O^6$BG/BCNU 처리는 벡터 주사 후 4 주에 시작하였고, 마우스를 벡터 주사 후 18 주까지 추적하였다(도 2b). 먼저, RBC에서 인간 γ-글로빈 발현을 분석하였다(도 2c). 생체내 선택 시작 전(형질도입 후 4 주) 수준은 배경보다 단지 미미하게 높았다. γ-글로빈+ 세포의 백분율은 두번째 차례의 선택 후 증가하기 시작하여 세번째 차례 후 80% 초과의 수준에 도달하였다. γ-글로빈-발현 세포의 백분율은 말초 혈액 및 골수에서 비적혈구 Ter119- 세포에 비해 적혈구 Ter119+ 세포에서 7- 내지 10-배 더 높았다(도 2d). HPLC를 사용하여 성체 마우스 α- 및 β-글로빈 쇄와 비교하여 γ-글로빈 단백질의 수준을 측정하였다(도 2e 및 도 3; https://doi.org/10.1172/JCI122836DS1의 보충 자료). 18 주차에, 이들 수준은 성체 마우스 α-글로빈 및 β-주요 글로빈의 10%-15% 및 마우스 β-소수 글로빈의 25%에 도달하였다. 이를 정량적 역전사 PCR(RT-qPCR)에 의해 mRNA 수준에서 확인하였으며, 여기서 인간 γ-글로빈 mRNA는 마우스 β-주요 mRNA의 13%였다(도 2f). 원시 장기간 재증식 HSC가 형질도입되었는지 추가로 입증하기 위해, 생체내-형질도입된/선택된 마우스로부터의 계통-고갈된(Lin-) 골수 세포를 조사된 C57BL/6 마우스에 이식하였다. 말초 혈액, 골수, 및 비장에서 분석된 생착 수준은 95% 초과였고 20 주의 관찰 기간에 걸쳐 안정되었다(도 4a, 4b). 인간 γ-글로빈 수준(마우스 α-글로빈과 비교)은 ("1차") 생체내-형질도입된 마우스(형질도입 후 18 주차에 분석) 및 이식 후 14 및 20 주차에 분석된 2차 수용자에서 유사하였다(도 4c).

[0890] 생체내 HSPC 형질도입/선택 접근법은 SB100X-매개 무작위 이식유전자 통합 패턴을 바꾸지 않고 조혈을 변경하지 않는다. 하이브리드 트랜스포존/SB100X HDAd5/35++ 시스템을 사용한 생체내 형질도입은 HSPC에서 무작위 이식유전자 통합을 초래하였음이 이전에 제시되었다(Richter 등, *Blood*. 128(18):2206-2217, 2016). 생체내 선택에서 $O^6$BG/BCNU의 효과를 평가하기 위해, 골수 Lin- 세포에서 이식유전자 통합을 연구 종료 시, 즉, 2차 수용자에서 20 주차에 분석하였다. 선형 증폭-매개 PCR(LAM-PCR) 이어서 심층 서열분석은 마우스 게놈에서 통합 부위의 무작위 분포 패턴을 나타내었다(도 5a). 5 마리 마우스의 풀링된 데이터는 엑손으로 2.23% 통합, 인트론으로 31.58%, 유전자간 영역으로 65.17%, 및 비번역 영역으로 1.04%를 입증하였다(도 5b). 통합의 무작위성 수준은 전체 마우스 게놈의 임의의 주어진 창에서 우선적인 통합 없이 99%였다(도 5c). 이는 생체내 선택 및 2차 수용자에서 세포의 추가 확장이 우세한 통합 부위의 출현을 초래하지 않았음을 나타낸다(도 5d). qPCR을 사용하여 형질도입 및 비-형질도입된 세포를 둘 다 함유하는 집단에서 골수 세포 당 평균 2 개의 γ-글로빈 cDNA 카피를 측정하였다. 그런 다음 통합된 이식유전자 카피 수를 단일-세포 수준에서 정량화하였다. 이를 위해, 18 주차 마우스로부터의 골수 Lin- 세포를 메틸셀룰로스에 플레이팅하고, 개별 전구체 콜로니를 단리하고, 게놈 DNA에 대해 qPCR을 수행하였다. 이식유전자-양성 콜로니에서(n = 113), 콜로니의 86.7%는 2 또는 3 개의 통합된 카피를 가졌다(도 5e 및 도 6). 4 개의 카피를 콜로니의 6.2%에서, 8 개의 카피를 1.78%에서 발견하였다. 콜로니의 0.88%는 13, 10, 7, 6, 또는 5 개의 통합 벡터 카피를 가졌다.

[0891] 연구 종료 시(18 주) 혈액 세포 계수의 변경은 밝견되지 않았다(도 7a). RBC 매개변수의 분석은 이상을 나타내지 않았다(도 7a-7c). 골수에서 Lin+ 분획의 조성물은 처리 전 마우스 및 처리 후(18 주) 마우스에서 유사하였다(도 7d). Lin-Sca1+cKit+ (LSK) HSPC(도 7d, 마지막 레인) 및 전구체 콜로니-형성 세포(도 7e)의 수준은

또한 두 그룹에서 비슷하였다.

[0892] 인간 CD46을 발현하고 인간 중간형 지중해빈혈과 유사한 CD46+/+/Hbbth-3 마우스 모델의 생성. HDAd5/35++ 벡터는 감염을 위해 인간 CD46을 필요로 한다. 생체내 HSPC 형질도입 연구를 위한 지중해빈혈 마우스 모델을 개발하기 위해, CD46tg(CD46+/+) 마우스를 마우스 Hbb-β1 및 -β2 유전자 결실에 대해 이형접합성인 Hbbth-3 마우스와 교배하였다(Yang 등, *Proc Natl Acad Sci USA.* 92(25):11608-11612, 1995). (동형접합성 상태는 자궁내 또는 출생후 초기에 치명적이다.) Hbbth-3 마우스는 인간 중간형 지중해빈혈과 유사한 생존가능한 형태의 지중해빈혈을 나타낸다. F1 하이브리드 마우스를 CD46+/+ 마우스와 역교배하여 CD46+/+/Hbbth-3 마우스를 생성하였다(도 8). 이들 마우스는 지중해빈혈 표현형을 나타내었다. 모체 CD46tg 마우스와 비교하여, CD46+/+/Hbbth-3 마우스는 상당히 감소된 RBC 수(7.1 ± 0.1 vs. 8.63 ± 0.29 M/μl); 더 낮은 헤모글로빈(9.7 ± 0.18 vs. 13.9 ± 0.63 g/dl), 적혈구용적률(30.7% ± 0.46% vs. 41.7% ± 1.48%), 평균 미립자 헤모글로빈(13.9 ± 0.14 vs. 16.1 ± 0.23 g/dl), 및 평균 미립자 부피(43.03 ± 0.22 vs. 48.35 ± 0.9 fl)를 가졌고; RBC 분포 폭(42.9% ± 0.29% vs. 25.3% ± 0.79%)이 증가하였으며; 명백한 망상적혈구증가증(42.4% ± 1.43% vs. 11.8% ± 3.7%)을 나타내었다(도 9a). 혈액 도말에서 적혈구 형태는 Hbbth-3 마우스 혈액 도말의 형태와 유사하고 CD46tg 마우스의 정상적혈구 외형과 뚜렷하게 대조적인 혈색소 감소증, 다양한 크기 및 모양(부동변형적혈구증가증), 및 세포 단편화를 특징으로 하였다(도 9b). 마찬가지로, CD46+/+/Hbbth-3 마우스로부터의 간 및 비장의 조직학적 분석은 적혈구 전구체 또는 거핵구의 클러스터를 함유하는 골수외 혈구형성의 병소를 보인 반면(도 9c, 하단 왼쪽 및 하단 중간 패널), Perl의 염색은 모체 CD46tg 마우스로부터의 조직 절편에서 골수외 혈구형성 및 철분 축적이 부재하거나 또는 제한된 것(도 9c, 상단 패널)과 대조적으로 뚜렷한 실질 철분 침착을 입증하였다(도 9c, 하단 오른쪽 패널). CD46+/+/Hbbth-3 마우스의 이러한 특성은 인간 질환을 요약하고 후속 실험을 위한 이러한 모델의 타당성을 뒷받침한다. 특히, CD46+/+/Hbbth-3 모델에서 지중해빈혈 표현형은 또한 총 WBC의 상승된 수에 의해 나타낸 바와 같이, 적혈구 계통 이외의 계통에서 정량적 차이를 특징으로 하였다(도 10).

[0893] HDAd-γ-글로빈/mgmt 및 HDAd-SB를 사용한 HSPC 생체내 형질도입 이어서 CD46+/+/Hbbth-3 마우스에서 생체내 선택은 γ-글로빈의 높고 안정된 장기간 발현을 초래한다. 생체내 형질도입 접근법이 CD46+/+/Hbbth-3 지중해빈혈 마우스 모델의 특징적인 질환의 매개변수를 개선할 수 있는지를 결정하였다. Hbbth-3 마우스에서 이전에 검증된 변형된 G-CSF/AMD3100 동원 체계(Psatha 등, Hum Gene Ther Methods. 2014;25(6):317-327)는 마지막 플레릭사포르(AMD3100) 주사 1 시간 후(도 11), 즉, HDAd-γ-글로빈/mgmt 및 HDAd-SB가 정맥내로 주사된 시점에 말초 혈액에서 LSK 세포의 높은 수를 산출하였다. 마우스는 인간 γ-글로빈 및 MGMT 단백질에 대한 반응을 피하기 위해 면역억제를 받았다(도 12). 생체외 렌티바이러스 벡터 유전자 요법 후, 유전적으로 교정된 적혈구 모세포가 생존 이점이 있고 Hbbth-3 마우스에서 생체내 선택을 겪는다는 보고를 고려하면(Miccio 등, *Proc Natl Acad Sci USA* 105(30):10547-10552, 2008), 초기에 $O^6$BG/BCNU 처리 없이 연구를 수행하도록 계획하였다. 평균 γ-글로빈+ RBC 백분율은 CD46+/+/Hbbth-3 마우스의 생체내 형질도입 후 8 주차에 31.19% ± 2.7%에 도달하였지만 16 주차에 13.14% ± 0.4%로 감소하였다. 이 시점에, 마우스를 2 개 그룹으로 나누었다. 마우스의 절반을 혈액 및 골수 분석(그룹 1: 생체내 선택 없음) 및 2차 수용자에 대한 공여자로서 사용한 반면, $O^6$BG/BCNU 생체내 선택을 수반하는 다른 그룹(그룹 2: 생체내 선택 있음)으로 연구를 계속하였다(도 12 참조). 16 주차에, 그룹 1은 말초 RBC의 13%에서 γ-글로빈 발현을 나타내었다(도 13a, 13b). 이 수준의 γ-글로빈 마킹은 말초 혈액 망상적혈구의 백분율에서 상당한 감소를 초래하였다(도 13c, 마지막 레인). 그러나, RBC 형태학 및 골수외 혈구형성을 포함하는 다른 RBC 매개변수를 개선하기에 충분하지 않았다(도 13c, 13d). 1차 γ-글로빈 마킹 수준은 이식 전에 부술판으로 골수 조건화된 2차 C57BL/6 수용자에서 20 주에 걸쳐 유지되었다(도 13e, 13f). 이는 장기간-재증식 HSPC가 형질도입되었음을 나타낸다.

[0894] 그룹 2에서, 4 주기의 생체내 선택은 29 주차에 평균 76%에 도달하는 γ-글로빈+ RBC 백분율에서 6-배 증가를 초래하였다(도 14a). γ-글로빈 발현은 Ter119- 세포와 비교하여 유세포 분석에 의한 게이팅 또는 면역자기적으로 단리된 Ter119+ 적혈구 세포에서 γ-글로빈 발현에 의한 분석에 의해 나타낸 바와 같이 적혈구-특이적이었다(도 14b, 도 14c). 다른 연구와 일치하여(Miccio 등, *Proc Natl Acad Sci USA.* 105(30):10547-10552, 2008, Zhao 등, *Blood.* 113(23):5747-5756, 2009), 선택은 골수(또는 비장)을 빠져나가고 핵을 손실하기 전에(핵형성 및 증식 적혈구) 전구체의 수준에서 발생하였다. 이는 생체내 선택 전에 비해 선택 후 우세하게 발생한 골수 및 비장에서 γ-글로빈+Ter119+ 세포의 증가에 반영된다 (도 14d). 그러나, (제핵 RBC가 우세한) 말초 혈액의 Ter119+ 세포에서 γ-글로빈+ 마킹의 전반적인 증가(도 14b)는 아마도 지중해빈혈 배경에 의해 제공된 "천연" 생체내 선택의 부가적인 효과에 기인할 것이다. HPLC에 의해 측정된 RBC에서 마우스 α-글로빈에 대한

인간 γ-글로빈의 비는 14 주차에 거의 검출불가능한 수준에서 29 주차에 10%로 증가하였다(도 14e 및 15; 기준선(도 15b), 16 주차(도 15c) 및 29 주차(도 15d, 및 CD46tg 대조군(도 15a)에서 CD46+/+/Hbbth-3 마우스 참조). 유사하게, 처리된 마우스의 혈액 세포에서 γ-글로빈 mRNA의 수준은 증가하여, 29 주차에 마우스 β-글로빈 mRNA의 10% 인간 γ-글로빈 mRNA로 번역하였다(도 14f). 생체내 형질도입 후 29 주차에 처리된 CD46+/+/Hbbth-3 마우스에서 세포 당 1.5 개의 γ-글로빈 유전자 카피가 측정되었다(도 16).

[0895] 생체내 형질도입/선택 후 CD46+/+/Hbbth-3 마우스의 지중해빈혈 표현형의 역전. 마지막 용량의 $O^6$BG/BCNU 처리 6 주 후, CD46+/+/Hbbth-3 마우스를 희생시키고, 분석을 위해 조혈 조직을 수확하였다. 생체내 형질도입 후 29 주차에 혈액학적 매개변수는 기준선에 걸쳐 상당히 개선되었고(도 17a)(RBC: 8.53 ± 0.16 vs. 7.1 ± 0.13 M/μl, P = 0.01; 헤모글로빈: 11.27 ± 0.39 vs. 9.7 ± 0.18 g/dl, P = 0.05; 적혈구용적률: 41.37% ± 0.81% vs. 30.7% ± 0.46%, P = 0.00001; 평균 미립자 부피: 48.63 ± 0.36 vs. 43.5 ± 0.38 fl, P = 0.003; RBC 분포 폭: 39.5% ± 0.8% vs. 43% ± 0.3 %, P = 0.006; 망상적혈구: 31.13% ± 3.17% vs. 42.4% ± 1.43%, P = 0.05), 특이적 적혈구 지수(적혈구용적률 [HCT], RBC, 평균 미립자 부피)에 대해, 수준은 대조군 CD46tg 대응물과 구별하기 어려웠으며, 이는 거의 완전한 표현형 교정을 시사한다. 혈액 도말의 망상적혈구 염색은 γ-글로빈+ RBC의 가장 높은 백분율을 갖는 처리된 CD46+/+/Hbbth-3 마우스에서 망상적혈구의 인상적인 3-배 감소를 입증하였다(도 17b). 처리된 CD46+/+/Hbbth-3 마우스의 말초 혈액 도말에서 지중해빈혈 표현형의 역전을 나타내는 저변색성이고, 고도로 단편화되고 부동변형적혈구증가적인 기준선 RBC를 크기가 덜 변이되고 거의 정색성이고 잘 형상화된 RBC로 대체하였다(도 17c, 상단 패널). 대조군 및 처리된 CD46+/+/Hbbth-3 마우스의 사이토스핀에서 전적혈구모세포 및 호염기성 적혈구모세포의 유병률에 의해 나타낸 CD46+/+/Hbbth-3 마우스의 골수에서 적혈구 계통 성숙의 차단과 대조적으로, 성숙 적혈구모세포가 우세하였고 다염성 및 정염성 적혈구모세포에 의해 나타내었다(도 17c, 중간 패널). 극심한 실질 혈철소증이 미처리 CD46+/+/Hbbth-3 마우스에서 관찰된 반면, CD46tg 및 처리된 CD46+/+/Hbbth-3 마우스에서 제한된 철분 축적만이 검출될 수 있었다(도 17c, 하단 패널). 따라서, 보상 혈구형성의 측정가능한 특징인 비장 크기는 처리된 동물에서 현저하게 감소하였다(도 17d, 17e).

[0896] 조합된 생체내 형질도입/선택 접근법이 원시 HSC의 유전적 변형을 초래하였는지를 결정하기 위해, 29 주차(형질도입 후)에 수확된 처리된 CD46+/+/Hbbth-3 마우스의 골수 세포를 준치사량의 부술판 처리 또는 치명적 전신 조사(TBI) 후 C57BL/6 2차 수용자에 이식하였다(도 18a, 18b). 예상된 바와 같이, TBI를 받은 마우스에서 생착률은 부술판-처리된 동물에서보다 더 높았지만, 생착 수준으로 조정된 발현 수준은 γ-글로빈+ RBC의 상당히 상이한 빈도를 나타내지 않았다. 이식-유래된 (CD46+) 적혈구의 75% 초과가 2차 이식 후 20 주차에 γ-글로빈+였고 정상 수용자 배경(여기서 HDAd-γ-글로빈/HDAd-SB-형질도입된 CD46+/+/Hbbth-3 HSPC는 선택적 이점이 없음)에서 골수하절제 부술판 조건에 의해 생성된 경쟁 조건 하에 29 주차에 1차 처리된 마우스에서 발견된 것과 유사한 마킹률을 갖는다(도 18c, 18d)는 사실은 접근법이 장기간 재증식 HSC의 유전적 교정을 초래한다는 결론을 추가로 뒷받침한다. 더욱이, 동원 및 생체내 형질도입에 제출된 이식 후 20 주차에 2차, 부술판-조건화된 C57BL/6 수용자는 γ-글로빈-발현 세포의 현저한 강화 및 발현/MFI에서 상당한 증가를 입증하였다(도 18e).

[0897] HDAd-γ-글로빈/mgmt 및 HDAd-SB를 사용한 생체내 HSPC 형질도입 이어서 $O^6$-BG/BCNU 생체내 선택의 안정성. 마우스 연구에서, 절차는 잘 용인되었다. 명백한 혈액학적 이상은 관찰되지 않았다. 희생 시, 마지막 $O^6$-GB/BCNU 용량 후 6 주에, 모든 혈액학적 값은 정상 범위 내에 있었지만, 총 WBC 계수는 생체내 선택 전 수준과 비교하여 더 낮았으며, 이는 WBC, 특히 림프구에 대한 약물 치료의 세포감소 효과를 시사한다(도 19a, 19b). 이 효과는 또한 미처리 또는 선택전 대응물과 비교하여 골수에서 CD3+, CD19+, 및 Gr-1+ 세포의 감소된 빈도를 반영하였다(도 19c). 특히, 최하점(25-27 주; 마지막 $O^6$BG/BCNU 주사 후 2-4 주)에서 조차도, WBC 및 혈소판은 결코 무형성증 수준(즉, 호중구 1,000/μl 미만, 혈소판 20,000/μl 미만)에 도달하지 않았고, WBC는 30 주까지(마지막 $O^6$BG/BCNU 주사 후 7 주) 회복되기 시작하였다. 이는 CD46tg 모델에서, WBC 및 림프구 계수가 마지막 $O^6$BG/BCNU 주사 후 10 주에 전처리 수준으로 되돌아갔다는 관찰과 함께(도 7a), 생체내 선택 약물의 세포감소 효과가 경미하고 일시적임을 시시한다. 중요하게는, LSK 및 Ter119+ 세포의 백분율에서 골수 세포 조성, 뿐만 아니라 골수 세포의 콜로니-형성 가능성은 HSPC의 생체내 형질도입/선택에 의해 영향을 받지 않았다(도 19c, 19d).

[0898] ***논의***. 혈색소병증의 생체외 HSPC 유전자 요법에서 명백한 임상 진행에도 불구하고, 임상적으로 관련된 HSPC 생착률에 도달하기 위한 골수절제 조건화에 대한 필요성은 주요 한계이다. 또한, 기술적 복잡성은 소수의 전문화

및/또는 공인된 센터에서만 이러한 치료의 이행을 허용한다. 개발된 생체내 HSPC 유전자 요법 접근법은 골수절제술 및 HSPC 세포 이식을 필요로 하지 않고 따라서 지중해빈혈에 대한 HSPC 유전자 요법을 더 안전하고 더 접근가능하게 만든다. 접근법의 핵심 아이디어는 골수로부터 HSPC를 동원하는 것이고, 주변부에서 많은 수로 순환하는 동안, 정맥내로 주사된 HSPC-향성 HDAd5/35++ 유전자 전달 벡터 시스템으로 형질도입한다. HDAd5/35++ 벡터 시스템의 신규 특징은 (a) 정맥내 주사 후 비조혈조직의 감염을 피하면서 원시 HSC의 효율적인 형질도입을 허용하는 CD46-친화도-향상된 섬유, (b) 세포 인자와 무관하게 기능하고 유전자에 대한 선호도 없이 무작위 이식유전자 통합을 매개하는 SB100X 트랜스포사제-기반 통합 시스템, 및 (c) 저용량 $O^6$BG/BCNU의 단기간 처리에 의해 형질도입된 원시 HSC의 풀에 영향을 미치지 않고 자손 세포의 선택적 생존 및 확장을 매개하는 $MGMT^{P140K}$ 발현 카세트를 포함한다(Wang 등, *Mol Ther Methods. Clin Dev*. 8:52-64, 2018). 현재 사용되는 SIN-렌티바이러스(SIN-LV) 벡터로부터 HDAd5/35++ 벡터를 구별하는 추가적인 특징은 이 연구에서 마이크로-LCR/β-프로모터-구동된 γ-글로빈 유전자 및 11.8 kb 크기의 EF1A 프로모터-구동된 $MGMT^{P140K}$ 유전자를 혼입하는 데 사용된 큰(30 kb) 삽입 용량을 포함한다. 또한, HDAd5/35++ 벡터의 생산은 대규모 플라스미드 형질감염을 필요로 하지 않고 스피너 배양 리터 당 $3 \times 10^{12}$ 개 초과의 감염성 입자를 산출한다. 특히, 혈색소병증에 대한 임상 연구에 사용되는 SIN-LV 벡터의 수율은 적어도 2 자릿수 이하이다.

[0899] 생체내 접근법의 효능. HSPC의 1% 미만의 안정된 형질도입이 상당한 임상적 이익을 제공하는 다른 유전적 질환의 HSPC 유전자 요법과 대조적으로(즉, X-관련 SCID, Cavazzana-Calvo 등, *Science*. 288(5466):669-672, 2000; ADA-SCID, Gaspar 등, *Sci Transl Med*. 3(97):97ra80. 2011; 또는 비스코트-알드리치 증후군, Aiuti 등, *Science*. 341(6148):1233151, 2013), 환자에서 혈색소병증의 표현형 교정은 적어도 20% 교정된 적혈구 전구체를 필요로 한다(Persons 등, *Blood*. 97(10):3275-3282, 2001, Andreani 등, *Blood*.; 87(8):3494-3499, 1996, Negre 등, *Blood*. 117(20):5321-5331, 2011). 혈색소병증에 대한 뮤린 모델에서, 총 α-글로빈 mRNA의 15%에서 γ-글로빈 발현이 요법에 충분하였다(Persons 등, Blood. *2001*;97(10):3275-3282, McColl 등, Blood Med. 7:263-274, 2016, Pestina 등, *Mol Ther*. 17(2):245-252, 2009). 이 연구에서, 생체내 형질도입/선택 후, 골수 적혈구모세포의 60% 초과가 생체내 형질도입된 CD46tg 및 CD46+/+/Hbbth-3 모델에서 γ-글로빈을 발현하였다(도 2c 및 14a). 이는 40%-97% 순환 γ-글로빈-발현 RBC로 번역되었다(도 2d 및 14b). 중요하게도 또한, 두 동물 모델에서, RBC에서 지속된 γ-글로빈 마킹이 2차 수용자에서 입증되었으며, 이는 원시, 장기간 재증식 HSC가 벡터 시스템에 의해 초기에 형질도입되었음을 시사한다.

[0900] qPCR 연구는 압도적인 대부분의 골수 세포에서 세포 당 2 내지 3 개의 통합된 이식유전자 카피를 검출하였다. 이전 연구와 일치하여(Zhao 등, *Blood*. 113(23):5747-5756, 2009, Zielske 등, *Mol Ther*. 9(6):923-931, 2004), 생체내 선택이 높은 카피 수 클론에 대해 선택되었다는 것은 밝혀지지 않았다. 게놈-와이드 통합 부위 분석을 고려하여, 1,000 개의 원래 형질도입된 HSC를 예상하였다. 마우스가 10,000 내지 20,000 개의 HSC를 갖는다는 것을 고려하면(Abkowitz 등, *Blood*. 100(7):2665-2667, 2002; Chen 등, *Blood*. 107(9):3764-3771, 2006), 벡터 시스템이 HSC의 5%-10%를 표적화하였음을 의미할 것이며, 이는 생체내 선택 후 조혈의 다클론 재구축 및 장기간 치료 효과를 위한 견고한 기초가 될 것이다.

[0901] 중간형 지중해빈혈 모델에서, 거의 완전한 표현형 교정이 달성되었다. 핵심 혈액학적 매개변수(HCT, RBC, 평균 미립자 부피)는 "건강한"(모체 CD46tg) 마우스의 대응물과 구별할 수 없었다. RBC 지수의 교정 정도 및 형태는 개별 마우스에서 γ-글로빈-발현 세포의 수준과 상관관계가 있었다. 말초 RBC 및 적혈구 골수 전구체 세포는 형태학 및 성숙 과정 둘 다에서 건강한 마우스와 유사하였다. 골수외 조혈 및 실질 철분 침착은 퇴화되었고, 비장 크기는 상당히 감소하였다. CD46+/+/Hbbth-3 모델에서 지중해빈혈 표현형은 또한 백혈구증가증/림프구증가증을 특징으로 하였다(도 10). (백혈구증가증/림프구증가증은 또한 종종 비장 절제된 지중해빈혈/겸상적혈구병 환자 또는 기능적 질환 연관 무비증 환자에 존재한다; Brousse 등, *Br J Haematol*. 166(2):165-176, 2014). 흥미롭게도, CD46++/Hbbth-3 마우스의 WBC 계수는 생체내 형질도입 후 29 주차에 "건강한" CD46tg 마우스의 수준으로 되돌아갔다(도 19a). 이 효과는 접근법에 의한 지중해빈혈 표현형의 역전이 적혈구 구획을 넘어 확장되어, WBC의 정상화, 및 대부분의 경우 전반적인 비장 기능을 초래함을 시사한다.

[0902] 특히, CD46tg 마우스에서의 연구와 대조적으로, 지중해빈혈 배경의 맥락 및 $O^6$BG/BCNU 처리의 부재 하에, γ-글로빈+ RBC의 13%는 CD46+/+/Hbbth-3 마우스의 말초 혈액에서 순환하였고, 이 수준은 2차 수용자에서 장기간 유지하였다. 이는 γ-글로빈 유전자 발현이 주요 지중해빈혈의 마우스 모델에서 생체외 렌티벡터 HSPC 유전자 요법으로 보고된 것과 유사한 지중해빈혈식 유전적으로 변형된 적혈구 전구체에 생존 이점을 부여하였음을 나타낸

다(Micco 등, *Proc Natl Acad Sci USA*. 105(30):10547-10552, 2008). 그러나, 지중해빈혈 마우스 모델에서 표현형 교정은 $O^6$BG/BCNU 처리를 필요로 하였다. 이는 낮은 글로빈 마킹으로 인해 필요한 경우, 유도성 생체내 선택 시스템이 쉬운 약리학적 개입에 의한 치료 효능의 구제를 허용함을 시사한다.

[0903] 뮤린 지중해빈혈 모델에서 추가로 γ-글로빈의 수준을 증가시키기 위해, 하기 가능성이 고려될 수 있다: (a) HDAd-SB 대 HDAd-γ-글로빈/mgmt 벡터의 비를 세포 당 통합된 이식유전자 카피의 수를 증가시키기 위해 1:1에서 1:3으로 바꿀 수 있다(Zhang 등, *PLoS One*. 8(10):e75344, 2013). (b) 또한 이식유전자 통합 위치 효과를 최소화하기 위해 β-글로빈 LCR의 26.1-kb 버전을 사용하여 γ-글로빈 발현을 구동하도록 계획하였다(Wang 등, *J Virol*. 79(17):10999-11013, 2005). (c) SB100X-기반 γ-글로빈 유전자 추가 시스템 이외에도, HDAd5/35++ 벡터는 γ-글로빈 억제인자 영역을 파괴하고 내인성 γ-글로빈 유전자를 재활성화기 위해 CRISPR/Cas9를 수용할 수 있다(Li 등, *Blood*. 131(26):2915-2928, 2018).

[0904] 동원 및 발현으로부터 시간의 관계를 평가하기 위해, HDAd-mgmt/GFP 벡터 + HDAd-SB 벡터를 G-CSF 및 AMD3100으로 동원 후 hCD46tg 마우스에 투여하였다. 혈청 항-HDAd 항체를 도 20a 및 20e에 제시된 바와 같이 측정하였다. GFP를 4 일 또는 4 주 및 동원 후 4 일에 측정하였다(도 20b("B") 및 20c("C")). 동원 및 HDAd 주사의 두번째 차례(첫번째 차례 후 4 주; 도 20d). 결과는 도 20f에 제시된다. 동원의 두번째 차례(도 20d; "D")는 바이러스에 대한 중화 혈청 항체의 발달로 인해 말초 혈액 세포의 형질도입을 초래하지 않았다. 그러나, 2차 이식 수용자에서 생체내 형질도입 연구에서 나타낸 바와 같이(도 18e), 항-HDAd 항체의 발달이 약제학적으로 차단될 수 있는 경우, 두번째 처리는 γ-글로빈+ RBC 및 γ-글로빈 발현 수준/MFI의 백분율을 둘 다 증가시킬 수 있다.

[0905] 생체내 HSPC 형질도입/선택 접근법의 안전성. 이 접근법은 골수절제술/조건화 및 연관된 독성에 대한 필요성을 폐지하는 반면, 간단한 정맥내 및 피하 물질/벡터 주사에 의해 조건화되지 않은 숙주에서 HSPC를 효과적으로 표적화한다. 중요하게는, 절차는 이 연구에 수반되는 모든 동물에서 잘 용인되었다.

[0906] G-CSF/AMD3100(플레릭사포르)에 기반한 HSPC 동원과 관련하여, 접근법은 안전하고 효과적인 것으로 임상적으로 입증되었고 주요 지중해빈혈에 대한 모든 진행중인 실험에서 백혈구분리반출술에 의한 HSPC 동원 및 수집에 정기적으로 사용된다(Psatha 등, *Curr Gene Ther*. 17(5):364-378, 2017, Karponi 등, *Blood*. 126(5):616-619, 2015). 이 연구에 사용되는 동원 레지멘에 대한 대안으로서, 다른 접근법은 HSPC의 보다 효율적인 동원을 달성하기 위해 작은 합성 분자에 의한 CXCR4의 연속 차단을 수반할 수 있다(Karpova 등, *Blood*. 129(21):2939-2949, 2017).

[0907] HDAd5/35++ 벡터의 정맥내 주사는 주사 후 3 일차에 CD46tg 마우스에서 동원된 HSPC 및 PBMC 이외의 조직에서 이식유전자 발현을 초래하지 않는다(Richter 등, *Blood*. 128(18):2206-2217, 2016). 이는 1세대 CD46-표적화 Ad5/35 및 Ad5/11 벡터가 정맥내로 주사된 개코원숭이에서 초기 연구와 일치하였다(Ni 등, *Blood*. 128(18):2206-2217, 2016). 이 향성에 대한 잠재적인 설명은 CD46 수용체 밀도 및 접근가능성이 비조혈 조직에서 충분히 높지 않아 효율적인 바이러스 형질도입을 허용하지 않는다는 점이다(Richter 등, *Blood*. 128(18):2206-2217, 2016; Ong 등, *Exp Hematol*. 34(6):713-720, 2006). 여기서, 세포 당 통합된 이식유전자 카피 수는 트랜스포존 벡터를 사용하여 생체내 형질도입/선택 후 18 주차에 상이한 조직에서 측정되었다(도 21a). 카피 수에 대한 효율은 도 21b 및 21c에 제시된다. 다양한 조직에서 세포 당 통합된 트랜스포존 카피가 제시된다(도 21d). 골수, PBMC, 및 비장에서 카피 수는 2.5였다. 통합된 이식유전자를 또한 간, 폐, 및 장의 게놈 DNA에서 검출하였다. GFP 벡터 시스템을 사용한 이전 연구는 이들 기관에서의 신호가 침윤 혈액 세포 및/또는 거주 대식세포로부터 기원함을 나타내었다(Richter 등, Blood. 2016;128(18):2206-2217).

[0908] HDAd 벡터(뿐만 아니라 다른 바이러스 벡터)의 정맥내 주사는 전염증성 사이토카인의 방출과 연관되지만(Atasheva 등, *Curr Opin Virol*. 21:109-113, 2016, Grieg 등, *Mol Ther Methods Clin Dev*. 3:16079, 2016), 바이러스 주사(Seregin 등, *Mol Ther*. 17(4):685-696, 2009) 또는 벡터 용량 분획화(Illingworth 등, *Mol Ther Oncolytics*. 5:62-74, 2017) 전 날에 글루코코르티코이드를 사용한 전처리에 의해 효율적으로 차단될 수 있다. 정맥내로 주사된 종양용해성 아데노바이러스의 우수한 안전성 프로파일은 CD46-표적화 종양용해성 아데노바이러스를 사용한 시험을 포함하여 수십 개의 임상 시험에서 문서화되었다(Garcia-Carbonero 등, *J Immunother Cancer*. 5(1):71, 2017).

[0909] 생체내 선택의 안전성 및 $O^6$BG/BCNU-자극된 증식이 장기간 정지된 HSPC의 저장소를 고갈시킬 수 있다는 우려와 관련하여, 대형-동물 모델을 사용한 연구는 HSPC 소진 또는 우성 클론의 출현 없이 장기간 다중계통 선택에 대

한 증거를 제공하였다(Beard 등, *J Clin Invest*. 120(7):2345-2354, 2010, Neff 등, *J Clin Invest*. 112(10):1581-1588, 2003). 이러한 모델에서, 조혈 및 골수외 독성 프로파일은 허용가능하였다. 본 연구 및 이전 마우스 연구에서(Wang 등, *Mol Ther Methods Clin Dev*. 8:52-64, 2018, Li 등, *Blood*. 131(26):2915-2928, 2018), 생체내 선택은 골수억제 없이 잘 용인되었다. $O^6$BG/BCNU 처리 시 골수 HSPC의 빈도에서 변화는 관찰되지 않았다. WBC, 구체적으로 림프구 계수에서 경미한 감소는 일시적이었다. 생체내 선택된 HSPC의 생존을 초래하는 DNA 복구 과정의 억제제인 $O^6$BG, 및 알킬화제인 BCNU를 사용한 저용량 처리의 3 내지 4 주기는 이론적으로 돌연변이 및 종양형성을 촉발할 수 있다. 이 위험에 대한 논쟁은 이러한 치료를 받았지만 발암 징후를 시사하지 않은 원숭이 및 개의 장기간 추적 연구이다(Beard 등, *J Clin Invest*. 120(7):2345-2354, 2010, Radke 등, *Sci Transl Med*. 9(414):eaan1145, 2017, Beard 등, *Blood*. 113(21):5094-5103, 2009). HSPC에서 이 위험을 평가하려는 시도에서, 시험관내 연구를 $MGMT^{P140K}$-발현 HDAd 벡터로 형질도입된 CD34+ 세포를 사용하여 수행하고 $MGMT^{P140K}$ 발현에 의해 보호되지 않는 세포의 98%가 사멸된 용량으로 $O^6$BG/BCNU 처리에 적용하였다(도 22a-22c). 약물 노출 후 14 일차에, 미처리 CD34+ 세포 및 처리 생존된 세포의 Illumina 전체 엑솜 서열분석을 분석하였으며, 결과는 하기 표에 제시되어 있다. 약물 처리 생존된 CD34+ 세포 vs 미처리 CD34+ 세포의 전체 엑솜 서열분석. 샘플 서열을 호모 사피엔스 참조 게놈(UCSC hg19)과 비교하였다.

[0910] **샘플 #1: 미처리 $CD34^+$ 세포**

| 총 정렬된 판독물 | 퍼센트 정렬된 판독물 | 표적화된 정렬된 판독물 | 판독물 강화 | 패딩된 표적 정렬된 판독물 | 패딩된 판독물 강화 |
|---|---|---|---|---|---|
| 46,870,836 | 80.51% | 38,437,631 | 82.01% | 40,158,769 | 85.68% |
| **총 정렬된 염기** | **퍼센트 정렬된 염기** | **표적화된 정렬된 염기** | **염기 강화** | **패딩된 표적 정렬된 염기** | **패딩된 염기 강화** |
| 6,544,191,633 | 75.45% | 4,308,625,487 | 65.64% | 5,541,019,474 | 84.67% |

[0911]

[0912] **샘플 #2: 선택된 $CD34^+$ 세포**

| 총 정렬된 판독물 | 퍼센트 정렬된 판독물 | 표적화된 정렬된 판독물 | 판독물 강화 | 패딩된 표적 정렬된 판독물 | 패딩된 판독물 강화 |
|---|---|---|---|---|---|
| 47,858,908 | 81.07% | 39,945,698 | 81.38% | 40,463,838 | 84.555 |
| **총 정렬된 염기** | **퍼센트 정렬된 염기** | **표적화된 정렬된 염기** | **염기 강화** | **패딩된 표적 정렬된 염기** | **패딩된 염기 강화** |
| 6,590,512,869 | 74.93% | 4,339,416,710 | 65.84% | 5,523,089,486 | 83.80% |

[0913]

[0914] 아미노산 치환이 단백질 기능에 영향을 미치는지를 예측하는 필터로서 Tolerant의 Sorting Intolerant(SIFT; uswest.ensemble.org에서 온라인 이용가능)를 사용하여, 처리된 샘플에서 47,858,908 개의 서열분석된 염기 쌍 당 126 개의 새로운 돌연변이(염기 쌍 당 2.63 x $10^{-6}$ 개의 돌연변이)를 식별하였다. 필터로서 ClinVar을 사용하여, 잠재적인 병리학적 효과가 있는 6 개의 돌연변이가 발견되었다. 표 11은 염색체 고유 돌연변이가 발견된 것을 요약한다:

[0915] 표 13

| 고유 돌연변이의 수 | |
|---|---|
| 염색체 1 | 793 |
| 염색체 2 | 502 |
| 염색체 3 | 369 |
| 염색체 4 | 243 |
| 염색체 5 | 253 |
| 염색체 6 | 341 |
| 염색체 7 | 383 |
| 염색체 8 | 241 |
| 염색체 9 | 312 |
| 염색체 10 | 269 |
| 염색체 11 | 536 |
| 염색체 12 | 362 |
| 염색체 13 | 94 |
| 염색체 14 | 252 |
| 염색체 15 | 271 |
| 염색체 16 | 475 |
| 염색체 17 | 527 |
| 염색체 18 | 94 |
| 염색체 19 | 755 |
| 염색체 20 | 283 |
| 염색체 21 | 92 |
| 염색체 22 | 276 |
| 염색체 M | 0 |
| 염색체 X | 351 |
| 염색체 Y | 6 |
| 총 | 8080 |

[0916]

[0917] $O^6$BG/BCNU 처리가 돌연변이를 유발한다는 발견은 예상되지 못했지만; 엑솜 서열분석 데이터 결과는 불분명하다. 기능 상실 돌연변이체는 인간 집단에서 흔하다. Exome Aggregation Consortium에 의한 최근 분석은 기능 상실 돌연변이가 있는 3,230 개의 유전자를 식별하였으며, 이들 변이체 중 72%는 현재 확립된 인간 질환 표현형이 없다(Lek 등, *Nature*. 536(7616):285-291, 2016).

[0918] SB100X 트랜스포사제 및 Flpe 재조합효소 유전자를 보유하는 HDAd-SB 벡터는 통합되지 않고 세포 분열 동안 손실된다(Li 등, *Mol Ther Methods Clin Dev*. 9:142-152, 2018). 이전에 공개된 데이터와 일치하여(Li 등, *Mol Ther Methods Clin Dev*. 9:142-152, 2018), 통합 또는 에피솜 HDAd-SB 벡터는 골수 Lin-세포에서 연구 종료 시 qPCR에 의해 검출가능하지 않았다. SB100X 트랜스포사제는 유전자 내로 또는 근처에 통합에 대한 선호도 없이 무작위 이식유전자 통합을 매개한다(Richter 등, *Blood*. 128(18):2206-2217, 2016, Zhang 등, *PLoS One*. 8(10):e75344, 2013). 이 무작위 패턴은 우세한 통합 부위/클론의 출현 없이 생체내 선택 후 유지된다. 이론적으로, 무작위 통합은 활성 유전자 내로 우선적 통합 보다 상대적으로 더 안전하며, 이는 렌티바이러스 또는 AAV 벡터 형질도입 동안 발생한다(Deyle 등, *Curr Opin Mol Ther*. 11(4):442-447, 2009, Bartholomae 등, *Mol Ther*. 19(4):703-710, 2011, Schroder 등, *Cell*. 110(4):521-529, 2002). 특히, β-지중해빈혈에 대한 SIN-LV-기반 임상 시험에서, HMGA2 원종양형성 유전자의 인트론 내로 통합은 1 명의 환자에서 양성 클론 우성을 촉발하였다(Cavazzana-Calvo 등, *Nature*. 467(7313):318-322, 2010).

[0919] SB100X-매개 무작위 이식유전자 통합 및 돌연변이성 선택 약물을 사용한 치료의 조합 효과로부터 잠재적인 종양형성의 위험을 감소시키기 위해, 벡터 시스템을 제1 위험 인자를 제거하도록 설계하였다. 이는 염색체 안전 항구 부위 내로 표적화된 γ-글로빈 통합을 매개하고 마우스에서 RBC의 70% 초과에서 안정된 γ-글로빈 마킹을 초래하였다(Li 등, 21st Annual American Society of Gene and Cell Therapy Meeting. Abstract 972).

[0920] 이 접근법의 안전성은 비인간 영장류의 장기간 연구에서 처음으로 명확하게 입증될 수 있다. 이 맥락에서 원숭이 및 개코원숭이 골수 CD34+ 세포는 인간 CD34+ 세포만큼 Ad5/35 벡터에 의해 효율적으로 형질도입됨에 주목할 만하고(Tuve 등, *J Virol*. 80(24):12109-12120, 2006), 원숭이에서 GFP를 발현하는 통합 HDAd5/35++ 벡터에 의해 동원된 CD34+ 세포의 직접 생체내 도입을 입증하였다(Harworth 등, 21st Annual American Society of Gene and Cell Therapy Meeting. Abstract 995).

[0921] **접근법의 임상 번역에 대하여**. HDAd5/35++ 벡터의 생산은 일상적으로 스피너 배양물 리터 당 5 x $10^{12}$ 개의 바이러스 입자(vp)를 산출한다. Flexion의 FX201 벡터에 대한 cGMP-등급 HDAd 생산이 확립된다. 정맥내로 주사된 바이러스에 대한 선천적 면역 반응의 약리학적 제어를 위한 프로토콜은 마우스보다 인간에 대해 더 개발되었고 현재 정맥내로 주사된 고용량 rAAV 벡터를 사용한 임상 시험에서 실행되고 있다. 그러나, 대부분의 인간은 Ad5 캡시드 단백질에 대해 지시된 중화 혈청 항체를 가지며, HDAd5/35 벡터, 즉, Ad5 캡시드 단백질 및 키메라 Ad35 섬유를 함유하는 벡터로 생체내 형질도입을 차단할 것이다. 본 개시내용에 기재된 대안은 Ad35로부터 유래된 벡터를 포함한다. Ad35는 혈청 유병률이 7% 미만이고 Ad5와의 교차 반응성이 없는 57 개의 알려진 인간 혈청형 중 가장 희귀한 것 중 하나이다(Vogels 등, *J Virol.* 77(15):8263-8271, 2003, Abbink 등, *J Virol.* 81(9):4654-4663, 2007, Kostense 등, *AIDS.* 18(8):1213-1216, 2004, Flomenberg 등, *J Infect Dis.* 155(6):1127-1134, 1987, Barouch 등, *Vaccine.* 29(32):5203-5209, 2011). Ad35는 Ad5보다 면역원성이 적으며(Johnson 등, *J Immunol.* 188(12):6109-6118, 2012), 이는 부분적으로 Ad35 섬유 놉에 의한 T 세포 활성화의 감쇠에 기인한다(Adams 등, *J Gen Virol.* 93(pt 6):1339-1344, 2012. Adams 등, *Proc Natl Acad Sci USA* 108(18):7499-7504, 2011, Shoji 등, *PLoS One.* 7(1):e30302, 2012). 정맥내 주사 후, 인간 CD46-유전자이식 마우스(Sakurai 등, *Gene Ther.* 13(14):1118-1126, 2006, Sakurai 등, *Mol Ther.* 16(4):726-733, 2008) 및 비인간 영장류(Sakurai 등, *Mol Ther.* 16(4):726-733, 2008)에서 간을 포함하는 조직의 최소 형질도입만이 있다(PCR에 의해서만 검출가능). 1세대 Ad35 벡터는 백신접종 목적을 위해 임상적으로 사용되었다(Baden 등, *Ann Intern Med.* 164(5):313-322, 2016, Kazmin 등, *Proc Natl Acad Sci USA* 114(9):2425-2430, 2017). 인간에서 향후 연구를 위해, 벡터는 생체내 HSPC 유전자 요법을 위해 HDAd35++에 기반하여 생성될 것이다.

[0922] 요약하면, 이는 지중해빈혈에 대한 전통적인 렌티바이러스 벡터 생체외 유전자 요법에 대한 대안을 제공하며, 이는 요법을 단순화하고, 주요 지중해빈혈이 풍토성이고 HSPC 이식이 실행가능하지 않는 자원 부족 지역에 접근 가능하게 만들 수 있다.

[0923] **실시예 2. 29 kb β-글로빈 유전자좌 제어 영역을 사용한 뮤린 지중해빈혈의 생체내 조혈 줄기 세포 유전자 요법**

[0924] 실시예 1은 생체내 변형된 HSPC에서 γ-글로빈 유전자 발현을 구동하는 능력의 상당한 진전을 기재한다. 또한 γ-글로빈 발현 수준을 추가로 증가시키기 위해, β-글로빈 LCR의 더 긴 버전(예를 들어, 26.1 kb)을 사용하여 γ-글로빈 발현을 구동시킬 수 있음을 언급한다. 이 실시예는 추적 분석 결과를 제공한다.

[0925] 본원에 기재된 바와 같이, 조혈 줄기/전구체 세포(HSPC) 동원 이어서 통합 헬퍼-의존성 아데노바이러스 HDAd5/35++ 벡터의 정맥내 주사는 형질도입된 HSPC의 생체내 선택 후 마우스 모델에서 장기간 재증식 세포의 효율적인 형질도입 및 질환 개선을 초래하였다. HDAd5/35++ 주사와 연관된 급성 선천적 독성은 이 접근법을 임상 번역을 위해 실현가능하게 만드는 적절한 예방에 의해 제어된다. 이는 기술적으로 중요 지중해빈혈 또는 겸상 적혈구병의 유전자 요법에 대한 간단한 생체내 HSPC 형질도입 접근법으로 사용될 수 있다. 이러한 질환의 치유는 높은 발현 수준의 치료 단백질(γ- 또는 β-글로빈)을 필요로 하며, 이는 더 큰 조절 요소를 수용하는 것을 허용하지 않는 게놈 크기 제한으로 인해 렌티바이러스 벡터로 달성하기 어렵다. 이 실시예는 HDAd5/35++ 벡터의 35 kb 삽입 용량을 활용하여 총 길이가 29 kb인 β-글로빈 유전자좌의 전사 조절 영역이 HSPC 내로 효율적으로 전달될 수 있음을 입증한다. 생체내 HSPC 형질도입은 적혈구 세포에서 안정된 γ-글로빈 수준을 초래하여 뮤린 중간형 지중해빈혈의 완전한 치유를 부여하였다. 특히, 이는 최소 생체내 HSPC 선택 레지멘으로 달성되었다. 이 연구는 큰 조절 영역을 혼입하는 HDAd5/35++ 벡터가 높은 수준의 이식유전자 발현을 필요로 하는 질환의 유전자 요법에서 과제를 해결할 수 있음을 입증한다.

[0926] ***도입***. 주요 지중해빈혈 및 겸상 세포 빈혈과 같은 혈색소병증의 유전자 요법이 성공하기 위해, 전달된 유전자는 바람직하게는 통합 및 전사 침묵의 위치 효과 없이, 적혈구 세포에서 높은 수준으로 발현된다. β-글로빈 유전자좌 제어 영역(LCR)은 이러한 용도에 유익한 것으로 생각된다. 유전자 요법 적용의 경우, HS1 내지 HS5를 함유하는 β-글로빈 LCR은 유전자이식 마우스에서 시스-연결된 유전자에 대해 높은 수준의 발현을 부여하는 것으로 제시되었다(Grosveld 등, *Cell* 51: 975-985, 1987). 그러나, 이 버전의 LCR은 렌티바이러스 벡터에서 사용하기에 너무 크고(삽입 용량 8 kb), 따라서 절두된 "미니" 또는 "마이크로" LCR 버전이 개발되었다. 예를 들어, 지중해빈혈 환자에서 진행중인 임상 시험에서 2.7 kb 미니-LCR(HS2-HS4 포함) 및 266 bp β-글로빈 프로모터을 함유하는 렌티바이러스가 사용되고 있다(Negre 등, *Curr Gene Ther* 15:64-81, 2015). 실시예 1에서, CD46 유전자이식 마우스 또는 CD46/Hbb$^{th3}$ 지중해빈혈 마우스에서 γ-글로빈의 발현을 위한 HS1 내지 HS4 및 β-글로빈 프로모터를 함유하는 5.9 kb β-글로빈 LCR 버전이 이용되었다(Wang 등, *J Clin Invest* 129:598-615,

2019). 생체내 HSPC 형질도입/선택 접근법을 사용하여, 말초 혈액 적혈구의 거의 100%에서 γ-글로빈 마킹이 달성된 반면, γ-글로빈의 발현 수준은 세포 당 2-3 개 카피의 평균 통합 벡터 카피 수(VCN)를 갖는 성체 마우스 α-글로빈 수준의 10-15%였다.

[0927] $\beta_0/\beta_0$ 지중해빈혈 또는 겸상 세포 빈혈의 완전한 치유를 위해, 일반적으로 적혈구 세포에서 20%의 치료 글로빈(γ- 또는 β-글로빈) 발현 수준이 필요하다고 생각된다(Fitzhugh 등, *Blood* 130:1946-1948, 2017). 이 수준에 도달하는 한 가지 방식은 HSPC 형질도입을 개선하거나 또는 벡터 용량을 증가시킴으로써 VCN을 증가시키는 것이다. 그러나, 이러한 접근법은 역사적으로 적어도 부분적으로 활용된 벡터 시스템의 무작위 통합 패턴으로 인해, 독성 위험을 증가시키는 다른 맥락에서 관찰되었다. 이 실시예에서, 더 강한 전사 요소, 즉 더 긴 LCR 버전은 CD46-유전자이식 마우스의 생체내 HSPC 형질도입 후 RBC에서 γ-글로빈 발현을 증가시키기 위해 활용되었다.

[0928] 백혈구분리반출술, 골수절제술, 및 HSPC 이식을 필요로 하지 않는 신규 생체내 HSPC 형질도입 접근법이 제공된다(Richter 등, *Blood,* 128:2206-2217, 2016). 접근법은 생체내 HSPC 형질도입에 적합한 새로운 백터 플랫폼, 즉 헬퍼-의존적, 캡시드-변형된 아데노바이러스 벡터(HDAd5/35++)를 수반한다. 이러한 벡터의 특징은 i.v. 주사 후 비-조혈 조직의 감염을 피하면서 원시 HSC의 효율적인 형질도입을 허용하는 CD46-친화도 향상된 섬유 및 최대 30 kb의 삽입 용량을 포함한다. 제한된 접근가능성으로 인해, 골수에 위치한 HSPC는 벡터가 골수 세포 상에 존재하는 수용체를 표적화하는 경우에도, HDAd5/35++ 벡터를 포함하는 정맥내로 주사된 벡터에 의해 형질도입되지 않을 수 있다(Ni 등, *Hum Gene* Ther, 16: 664-677, 2005 및 Ni 등, *Cancer Gene Ther*, 13:1072-1081, 2006). 과립구-콜로니-자극 인자(G-CSF) 및 CXCR4 길항제 AMD3100(MOZOBIL™, PLERIXA™)의 조합은 동물 모델 및 인간에서 원시 전구체 세포를 효율적으로 동원하는 것으로 제시되었다(Fruehauf 등, *Cytotherapy*, 11: 992-1001, 2009 및 Yannaki 등, *Hum Gene Ther*, 24: 852-860, 2013). G-CSF/AMD3100을 사용하여 골수에서 말초 혈류로 HSPC를 동원한 후 HDAd5/35++ 벡터를 정맥내 주사하였다. 이는 이전에 인간 CD46 유전자이식 마우스(Richter 등, *Blood,* 128: 2206-2217, 2016; Li 등, *Mol Ther Methods Clin Dev,* 9: 390-401, 2018; Li 등, *Blood*, 131: 2915-2928. 2018; Wang 등, *J Clin Invest*, 129: 598-615. 2019; Wang 등, *Blood Adv*, 3: 2883-2894, 2019; 및 Wang 등, *Mol Ther Methods Clin Dev*, 8: 52-64, 2018), 인간화 마우스(Richter 등, *Blood,* 128: 2206-2217, 2016) 및 레서스 원숭이(Harworth 등, *ASCGT 21th Annual meeting*, 2018, DOI: 10.1016/j.ymthe.2018.05.001)에서 제시되었다. 주변부에서 되돌아가 형질도입된 HSPC는 장기간 지속되는 골수로 되돌아간다. 증식 이점 없이, 생체내 형질도입된 HSPC는 골수를 효율적으로 빠져나가지 않고 하류 분화에 기여한다. $O^6$BG/BCNU를 사용한 동물의 단기간 처리는 $mgmt^{P140K}$ 유전자-변형된 HSPC에 대한 증식 자극을 제공하고 이후에 말초 혈액 세포의 80% 초과에서 안정된 이식유전자 발현을 제공한다(Wang 등, *Mol Ther Methods Clin Dev*, 8: 52-64, 2018).

[0929] HD-Ad5/35++ 게놈은 숙주 세포 게놈 내로 통합되지 않고 세포 분열 시 손실된다. 유전자 요법 목적 및 장기간 생체내 형질도입된 HSPC를 추적하기 위해, HD-Ad5/35++ 벡터를 이식유전자 통합을 허용하도록 변형시켰다. 이는 과활성 잠자는 미녀 트랜스포사제 시스템(SB100)을 통합함으로써 수행되었다(Zhang 등, *PLoS One,* 8: e75344, 2013; Hausl 등, *Mol Ther*, 18: 1896-1906, 2010; 및 Yant 등, *Nat Biotechnol*, 20: 999-1005, 2002). 제2 벡터로부터 트랜스로 공동 발현된 트랜스포사제는 이식유전자 카세트에 플랭킹하는 특이적 DNA 서열(도립 반복부; "IR")을 인식하고 염색체 DNA의 TA 디뉴클레오티드 내로의 통합을 촉발한다. 레트로바이러스 통합과 달리, SB100x-매개 통합은 표적화된 유전자의 전사 상태에 의존하지 않는다(Yant 등, *Mol Cell Biol*, 25: 2085-2094, 2005). 여러 연구는 SB100x-매개 이식유전자 통합이 무작위적이고 원종양형성 유전자의 활성화와 연관되지 않음을 입증하였다(Richter 등, *Blood,* 128: 2206-2217, 2016; Wang 등, *Mol Ther Methods Clin Dev*, 8: 52-64, 2018; Zhang 등, *PLoS One,* 8: e75344, 2013; Hausl 등, *Mol Ther*, 18: 1896-1906, 2010; 및 Yant 등, *Nat Biotechnol*, 20: 999-1005, 2002). SB100x-기반 통합 시스템의 이점은 세포의 효율적인 상동 DNA 복구 기구에 의존하지 않는다는 점이다. 후자는 DNA 복구 및 재조합효소의 낮은 활성을 나타내는 HSPC에서 중요하다(Beerman 등, *Cell Stem Cell*, 15: 37-50, 2014). CD46-유전자이식 마우스(Richter 등, *Blood,* 128: 2206-2217, 2016; Wang 등, *J Clin Invest*, 129: 598-615. 2019; Li 등, *Mol Ther*, 27: 2195-2212, 2019; Li 등, *Mol Ther Methods Clin Dev*, 9: 142-152, 2018; 및 Wang 등, *J Virol*, 79: 10999-11013, 2005) 및 인간 CD34+ 세포(Li 등, *Mol Ther*, 27: 2195-2212, 2019)에서 HDAd35++-트랜스포존 벡터 및 SB100x/Flpe 발현 벡터를 사용한 생체내 HSC 공동-감염은 유전자에 대한 선호도 없이 세포 당 2 개 이식유전자 카피의 무작위 이식유전자 통합을 초래하였음이 입증되었다.

[0930] 인간 게놈은 일반적으로 루프 형성을 통해 조절 영역(즉, 전사 인자 결합 부위) 사이에 긴 범위 상호작용을 갖는 3-D 구조로 조직화된다. 이러한 상호작용의 대부분은 위상적으로 연관된 도메인(TAD)의 맥락에서 발생한다. TAD는 인핸서가 전사를 제어하기 위해 다른 조절 영역과 상호작용하는 염색체 조직화의 기능적 단위로 간주된다. TAD/LCR 경계 절연은 인핸서 및 프로모터의 검색 공간을 제한하고 원치않은 조절 접촉이 형성되는 것을 방지하는 것으로 생각된다. 이러한 도메인의 양 측면에서 경계는 상이한 포유류 세포 유형 사이에 그리고 심지어 종에 걸쳐 보존된다.

[0931] 현재 사용되는 렌티바이러스 및 rAAV 유전자 전달 벡터는 작은 인핸서/프로모터만을 수용하여, 종종 이식유전자 발현의 차선적 수준 및 조직 특이성, 이식유전자 침묵, 및 벡터 통합 부위 주위의 조절 영역과의 의도하지 않은 상호작용을 초래할 수 있다. 최악의 경우에 대한 시나리오에서, 후자는 원종양형성 유전자의 활성화로 이어질 수 있다.

[0932] 유전자 요법의 안전성 및 효능을 증가시키기 위해, TAD는 유전자 추가 전략에 사용되어야 한다. TAD의 중간 크기는 880 kb이다. 고처리량 염색체 형태 캡처 (3C) 검정 및 후속 4C, 5C 및 Hi-C 프로토콜 뿐만 아니라 섬유-Seq 검정의 추가 발전에 따라, 조절 게놈의 심문은 빠른 속도로 진행될 것이고, 유전자 요법 목적을 위해, 중요한 코어 요소만을 함유하는 TAD를 전달할 수 있다. β-글로빈 유전자좌 제어 영역(LCR)은 TAD의 정의 하에 속한다.

[0933] 캡시드-변형된 HDAd5/35++ 벡터는 생체내 HSPC 유전자 요법에 사용되었다(Li & Lieber, *FEBS Lett.* 593(24):3623-48, 2019; Richter 등, *Blood.* 128(18):2206-17, 2016). 접근법은 골수에서 HSPC의 동원을 수반하고, 주변부에서 많은 수로 순환하는 동안, HDAd5/35++ 벡터가 정맥내로 주사된다. 이러한 벡터는 원시 HSPC에서 발현되는 수용체인 CD46을 표적화한다(Richter 등, *Blood.* 128(18):2206-17, 2016). 형질도입된 HSPC는 장기간 지속되는 경우 골수로 되돌아간다. 무작위 통합은 활성-향상된 잠자는 미녀 트랜스포사제(SB100x)에 의해 매개된다(Boehme 등, *Mol Ther Nucleic Acids.* 5(7):e337, 2016). 표적화된 통합은 상동성 의존적 DNA 복구를 통해 달성될 수 있다(Li 등, *Mol Ther.* 27(12):2195-212, 2019). 이 접근법은 뮤린 중간형 지중해빈혈의 개선(Wang 등, *J Clin Invest.* 129(2):598-615, 2019), 뮤린 혈우병의 교정(Wang 등, *Blood Adv.* 3(19):2883-94, 2019), 및 자연암의 회귀(Li 등, *Cancer Res.* 80(3):549-560, 2019)를 초래하였다. 비-인간 영장류에서의 첫번째 데이터는 생체내 HSPC 유전자 요법 접근법이 정맥내 HDAd5/35++ 주사 후 선천적 면역 반응을 억제하기 위해 글루코코르티코이드, IL6- 및 IL1β-수용체 길항제 전처리와 조합될 때 안전함을 나타낸다(Li 등, *23rd Annual ASGCT meeting*. 2020; abstract #546). HDAd5/35++ 벡터의 정맥내 주사는 주사 후 3 일차에 CD46tg 마우스에서 동원된 HSPC 및 PBMC 이외의 조직에서 이식유전자 발현을 초래하지 않았다(Richter 등, *Blood.* 128(18):2206-17, 2016; Wang 등, *J Clin Invest.* 129(2):598-615, 2019). 이는 최근에 비-인간 영장류에서 확인되었다. 이 향성에 대한 잠재적인 설명은 CD46 수용체 밀도 및 접근가능성이 효율적인 바이러스 형질도입을 허용하기에 비-조혈조직에서 충분히 높지 않다는 것이다(Richter 등, *Blood.* 128(18):2206-17, 2016; Ni 등, *Hum Gene Ther.* 16(6):664-77, 2005).

[0934] HDAd5/35++ 벡터를 사용한 이전 연구에서, 4.3 kb HS1-HS4 미니-LCR(β-글로빈 유전자좌 제어 영역)은 생체내 HSPC 형질도입 후 인간 γ-글로빈 발현을 구동하기 위해 0.66 kb β-글로빈 프로모터와 조합하여 사용되었다(Wang 등, *J Clin Invest.* 129(2):598-615, 2019; Ong 등, *Exp Hematol.* 34(6):713-20, 2006). $Hbb^{th3}/CD46^{+/+}$ 지중해빈혈 마우스에서, 안정된(8 개월 이상) γ-글로빈 마킹은 말초 혈액 적혈구의 거의 100% 및 거의 완전한 표현형 교정에서 달성되었다(Wang 등, *J Clin Invest.* 129(2):598-615, 2019). 그러나, γ-글로빈 발현 수준은 세포 당 2 개 카피의 평균 통합 벡터 카피 수(VCN)를 갖는 성체 마우스 α-글로빈의 단지 10-15%였으며, 따라서 주요 지중해빈혈 또는 SCD에 대한 접근법의 임상 번역을 특히 어렵게 만든다. 여기서, HDAd5/35++ 벡터의 큰 용량은 γ-글로빈 발현 카세트를 포함하는 β-글로빈 TAD 코어 요소를 29 kb 길이로 통합하여 완전한 표현형 교정을 달성함으로써 활용되었다.

[0935] 이 맥락에서, 또 다른 의도는 SB100x 시스템이 32.4 kb 트랜스포존의 효율적인 통합을 매개할 수 있음을 입증하는 것이었다. 플라스미드-기반 SB 시스템을 사용한 연구로부터 SB 통합 활성은 트랜스포존의 길이와 음의 상관관계가 있는 것으로 생각되었다(Li 등, *Mol Ther Methods Clin Dev.* 9:142-52, 2018; Karsi 등, *Mar Biotechnol (NY).* 3(3):241-5, 2001). 이를 고려하면, Kay 및 Ehrhardt 그룹에 의해 개발된 첫번째 SB-기반 HDAd 벡터는 상대적으로 작은(4 kb-6 kb) 트랜스포존을 보유하였다(Turchiano 등, *PLoS One.* 9(11):e112712, 2014; Yant 등, *Nat Biotechnol.* 20(10):999-1005, 2002).

[0936] 최근에, HDAd5/35++ 벡터를 사용하여, 생체외 또는 생체내 HSPC 형질도입 후 HSPC에서 10.8 kb(Wang 등,

*Blood Adv.* 3(19):2883-94, 2019) 및 11.8 kb(Wang 등, *J Clin Invest.* 129(2):598-615, 2019; Ong 등, *Exp Hematol.* 34(6):713-20, 2006) 트랜스포존의 효율적인 SB100x-매개 통합을 입증하였다. 이 실시예는 HDAd5/35++-기반 SB100x 벡터 시스템이 32.4 kb 트랜스포존을 통합할 수 있다는 증거를 제공한다.

[0937] 전반적으로, 정상 및 지중해빈혈 마우스에서 이러한 생체내 연구 뿐만 아니라 인간 CD34+ 세포에 대한 시험관내 연구는 HDAd5/35++ 벡터를 함유하는 기재된 긴-LCR이 혈색소병증의 치료를 위한 효율적인 도구일 수 있음을 나타낸다.

[0938] ***재료 및 방법.***

[0939] **구성요소 위치**: HS5→HS1 (21.5kb): Chr11, 5292319→5270789; β-프로모터: chr11, 5228631→5227023; 및 3'HS1: Chr11, 5206867→5203839.

[0940] **HDAd 벡터**: HDAd-SB 및 HDAd-짧은-LCR 벡터의 생성은 이전에 기재되었다(Richter 등, *Blood* 128: 2206-2217, 2016; Ong 등, *Exp Hematol* 34(6):713-20, 2006). HDAd-긴-LCR 벡터의 생성을 위해, 상응하는 셔틀 플라스미드는 코스미드 벡터 pWE15를 기반으로 하였다(Stratagene, 캘리포니아주 라호이아 소재). pWE.Ad5-SB-mgmt는 Ad5 5'ITR(뉴클레오티드 1 내지 436) 및 3'ITR(뉴클레오티드 35741 내지 35938), pBS-μLCR-γ-글로빈-mgmt로부터 유래된 구 인간 EF1α 프로모터-mgmt$^{P140K}$-SV40pA-cHS4 카세트(Wang 등, *J Clin Invest* 129: 598-615, 2019), SB100x-특이적 IR/DR 부위 및 FRT 부위를 함유한다. pAd.LCR-β-GFP(21.5-kb 인간 β-글로빈 LCR 함유)에서의 GFP-BGHpA 단편(Hudecek 등, *Crit Rev Biochem Mol Biol* 52(4):355-380, 2017)은 인간 γ-글로빈 유전자 및 이의 3'UTR 영역에 의해 대체되었다(Chr 11:5,247,139 →5,249,804)(pAd-긴-LCR-β-γ-글로빈). 플라스미드 pAd-긴-LCR-β-γ-글로빈은 21.5-kb 인간 β-글로빈 LCR 및 3.0-kb 인간 β-글로빈 3'HS1을 함유한다. LCR-β-γ-글로빈-3'HS1을 함유하는 28.9-kb 단편은 EF1α-mgmt-SV40pA-cHS4의 카세트의 하류를 pWE.Ad5-SB-mgmt 내로 삽입하였다(pWE.Ad5-SB-긴-LCR-γ-글로빈/mgmt). 완전한 긴-LCR-γ-글로빈/mgmt 카세트는 SB100x-특이적 IR/DR 부위 및 FRT 부위에 의해 플랭킹되었다. 생성된 플라스미드를 Gigapack III Plus Packaging Extract(Stratagene, 캘리포니아주 라호이아 소재)를 사용하여 파지 내로 패킹하고 증식시켰다. HDAd-긴-LCR-γ-글로빈/mgmt 바이러스를 생성하기 위해, 바이러스 게놈을 116 세포에서 구제를 위한 플라스미드로부터 I-CeuI 소화에 의해 방출하였다. 단일 아미노산 변이가 있는 인간 집단에서 HBG1 유전자에 대한 2 개의 알려진 변이체가 있다(76-이소류신 또는 76-트레오닌). 빈도가 유럽인의 13%에서 동아시아인의 73%까지의 범위인 76-Ile HBG1 변이체가 사용되었다.

[0941] HDAd 바이러스를 생성하기 위해, 바이러스 게놈을 Ad5/35++-Acr 헬퍼 바이러스가 있는 116 세포에서 구제하기 위해 FseI 소화에 의해 방출하였다(Palmer 등, *Mol Ther* 8: 846-852, 2003). 이 헬퍼 바이러스는 Ad5 섬유 꼬리, Ad35 섬유 샤프트, 및 친화도-향상된 Ad35++ 섬유 놉으로 구성된 키메라 섬유를 함유하는 Ad5/35++ 헬퍼 벡터인 AdNG163-5/35++의 유도체이다(Richter 등, *Blood* 128: 2206-2217, 2016). SpCas9 활성을 억제하는 것으로 최근에 제시된 인간 코돈-최적화된 AcrIIA4-T2A-AcrIIA2 서열을 합성하고(Yang 등, *Proc Natl Acad Sci USA.* 92(25):11608-12, 1995) 셔틀 플라스미드 pBS-CMV-pA(pBS-CMV-Acr-pA)로 클로닝하였다. 이후에, 2.0-kb CMV-Acr-pA 카세트를 pBS-CMV-Acr-pA로부터 증폭시키고 In-Fusion HD 클로닝 키트(Takara)에 의해 pNG163-2-5/35++의 *SwaI* 부위 내로 삽입하였다(Richter 등, *Blood* 128: 2206-2217, 2016). 그런 다음 바이러스 게놈을 PacI 소화에 의해 방출하고 Ad5/35++-Acr 헬퍼 바이러스를 구제하고 293 세포(HEK293)에서 증식시켰다. HDAd-SB의 생성은 이전에 기재되었다(Richter 등, *Blood* 128: 2206-2217, 2016). 헬퍼 바이러스 오염 수준은 0.05% 미만이었다. 모든 제제는 박테리아 내독소가 없었다.

[0942] **CD34$^+$ 세포 배양**: G-CSF-동원된 성체 공여자로부터의 CD34$^+$ 세포를 동결된 스톡으로부터 회수하고 10% 열-불활성화 FCS, 1% BSA 0.1 mmol/l 2-머캅토에탄올, 4 mmol/l 글루타민 및 페니실린/스트렙토마이신, Flt3 리간드(Flt3L, 25 ng/ml), 인터류킨 3(10 ng/ml), 트롬보포이에틴(TPO)(2 ng/ml), 및 줄기 세포 인자(SCF)(25 ng/ml)가 보충된 Iscove의 변형된 Dulbecco 배지(IMDM)에서 밤새 배양하였다. 유세포 분석은 세포의 98% 초과가 CD34+였음을 입증하였다. 사이토카인 및 성장 인자는 Peprotech(뉴저지주 로키 힐 소재)로부터 수득하였다. CD34$^+$ 세포를 낮은 부착 12 웰 플레이트에서 바이러스로 형질도입하였다.

[0943] **적혈구 시험관내 분화**: 인간 HSPC의 적혈구 세포 내로의 분화를 Douay 등에 기재된 프로토콜에 기반하여 수행하였다(*Methods Mol Biol* 482: 127-140, 2009). 간단히 말해서, 단계 1에서, $10^4$ 개 세포/ml 밀도의 세포를 5% 인간 혈장, 2 IU/ml 헤파린, 10 μg/ml 인슐린, 330 μg/ml 트랜스페린, 1 μM 하이드로코르티손, 100

ng/ml SCF, 5 ng/ml IL-3, 3 U/ml 에리트로포이에틴(Epo), 글루타민, 및 Pen-Strep이 보충된 IMDM에서 7 일 동안 배양하였다. 단계 2에서, $1x10^5$ 개 세포/ml 밀도의 세포를 5% 인간 혈장, 2 IU/ml 헤파린, 10 μg/ml 인슐린, 330 μg/ml 트랜스페린, 100 ng/ml SCF, 3 U/ml Epo, 글루타민, 및 Pen/Strep이 보충된 IMDM에서 3 일 동안 배양하였다. 단계 3에서, $1x10^6$ 개 세포/ml 세포 밀도의 세포를 5% 인간 혈장, 2 IU/ml 헤파린, 10 μg/ml 인슐린, 330 μg/ml 트랜스페린, 3 U/ml Epo, 글루타민, 및 Pen/Strep이 보충된 IMDM에서 12 일 동안 배양하였다.

[0944] **형질도입된 CD34+ 세포의 시험관내 선택:** 형질도입된 CD34+ 세포를 시험관내 분화 프로토콜의 단계 1 에서 5 일째에 $O^6$BG/BCNU로 선택하였다. 간단히 말해서, CD34+ 세포를 50 μM $O^6$BG와 1 시간 동안 배양한 다음 35 μM BCNU와 추가 2 시간 동안 배양한 다음, 세포를 2 회 세척하고 새로운 단계 1 배지에 재현탁하였다.

[0945] **$Lin^-$ 세포 배양**: 계통 음성 세포를 Miltenyi Biotech(독일 베르기슈글라트바흐 소재)의 계통 세포 고갈 키트를 사용하여 MACS에 의해 총 마우스 골수 세포로부터 단리하였다. $Lin^-$ 세포를 10% FCS, 10% BSA, Pen-Strep, 글루타민, 10 ng/ml 인간 TPO, 20 ng/ml 마우스 SCF 및 20 ng/ml 인간 Flt-3L이 보충된 IMDM에서 배양하였다.

[0946] **글로빈 HPLC**: 개별 글로빈 쇄 수준을 SPD-10AV 다이오드 어레이 검출기 및 LC-10AT 바이너리 펌프가 장착된 Shimadzu Prominence 기기(Shimadzu, 일본 쿄토 소재)에서 정량화하였다. 물/아세토니트릴 중 0.1% 트리플루오로아세트산의 40%-60% 구배 혼합물을 Vydac C4 역상 칼럼(Hichrom, 영국 소재)을 사용하여 1 mL/분의 속도로 적용하였다.

[0947] **유세포 분석**: 세포를 1 % FCS가 보충된 PBS에 $1x10^6$ 개 세포/100 μL로 재현탁하고 얼음 위에서 10 분 동안 FcR 차단 시약(Miltenyi Biotech, 캘리포니아주 오번 소재)과 함께 배양하였다. 다음으로 염색 항체 용액을 $10^6$ 개 세포 당 100 μL에 첨가하고 암실에서 30 분 동안 얼음 위에서 배양하였다. 배양 후, 세포를 FACS 완충액(PBS, 1% FBS)에서 1 회 세척하였다. 염색 단계를 2차 염색 용액으로 반복하였다. 세척 후, 세포를 FACS 완충액에 재현탁하고 LSRII 유세포 분석기(BD Biosciences, 캘리포니아주 산호세 소재)를 사용하여 분석하였다. 파편을 전방 산란 영역 및 측방 산란 영역 게이트를 사용하여 제외하였다. 그런 다음 단일 세포를 전방 산란 높이 및 전방 산란 폭 게이트를 사용하여 게이팅하였다. 그런 다음 유세포 분석 데이터를 FlowJo(버전 10.0.8, FlowJo, LLC)를 사용하여 분석하였다. LSK 세포의 흐름 분석을 위해, 세포를 비오틴-접합된 계통 검출 칵테일(cat #: 130-092-613; Miltenyi Biotec, 캘리포니아주 샌디에이고 소재) 및 c-Kit(cat #:12-1171-83) 및 Sca-1(cat #: 25-5981-82)에 대한 항체 뿐만 아니라 APC-접합된 스트렙타비딘으로 염색하였다. eBioscience(캘리포니아주 샌디에이고 소재)로부터의 다른 항체는 항-마우스 LY-6A/E(Sca-1)-PE-Cyanine7(클론 D7), 항-마우스 CD117(c-Kit)-PE(클론 2B8), 항-마우스 CD3-APC(클론 17A2; cat #:17-0032-82), 항-마우스 CD19-PE-Cyanine7(클론 eBio1D3; cat #: 25-0193-82), 및 항-마우스 Ly-66 (Gr-1)-PE(클론 RB6-8C5; cat #: 12-5931-82)을 포함하였다. 항-마우스 Ter-119-APC(클론: Ter-119; cat #: 116211)는 Biolegend(캘리포니아주 샌디에이고 소재)로부터 수득하였다.

[0948] **인간 γ-글로빈 발현을 검출하는 세포내 유세포 분석:** FIX & PERM™(Nordic Immunological Laboratories, 네덜란드 수스테렌 소재) 세포 투과화 키트(Thermo Fisher Scientific, 매사추세츠주 월섬 소재)를 사용하고 제조업체의 프로토콜에 따랐다. 간단히 말해서, $1x10^6$ 개 세포를 100 μl FACS 완충액(1% FCS가 보충된 PBS)에 재현탁하고, 100 μl의 시약 A(고정 배지)를 첨가하고 실온에서 2-3 분 동안 배양한 다음, 1 ml 사전-냉각된 절대 메탄올을 첨가하고, 혼합하고 10 분 동안 암실에서 얼음 위에서 배양하였다. 그런 다음 샘플을 FACS 완충액으로 세척하고 100 μl 시약 B(투과 배지) 및 0.3 μg 헤모글로빈 γ 항체(Santa Cruz Biotechnology, 텍사스주 댈러스 소재, cat# sc-21756 PE)에 재현탁하고, 실온에서 30 분 동안 배양하였다. 세척 후, 세포를 FACS 완충액에 재현탁하고 분석하였다. 유세포 분석 게이팅 전략은 도 46에 제시되어 있다.

[0949] **실시간 역전사 PCR:** 총 RNA를 제조업체의 페놀-클로로포름 추출 방법에 따라 TRIzol™ 시약(Thermo Fisher Scientific)을 사용하여 50-100 μl 혈액으로부터 추출하였다. Quantitect 역전사 키트(Qiagen) 및 power SYBR™ green PCR 마스터 믹스(Thermo Fisher Scientific)를 사용하였다. 실시간 정량적 PCR을 StepOnePlus 실시간 PCR 시스템(AB Applied Biosystems)에서 수행하였다. 다음 프라이머 쌍을 사용하였다: 마우스 RPL10(하우스키핑) 정방향(서열번호: 189), 및 역방향(서열번호: 190); 인간 γ-글로빈 정방향(서열번호: 191), 및 역방향

(서열번호: 192); 마우스 β-주요 글로빈 정방향(서열번호: 193), 및 역방향(서열번호: 194), 마우스 α 글로빈 정방향(서열번호: 212), 및 역방향(서열번호: 213).

[0950] **벡터 카피 수의 측정**: 골수 세포로부터의 총 DNA를 Quick-DNA 미니프렙 키트(Zymo Research)를 사용하여 추출하였다. HDAd-짧은 LCR-γ-글로빈/mgmt 바이러스로부터 추출된 바이러스 DNA를 연속으로 희석하고 표준 곡선에 사용하였다. qPCR을 StepOnePlus 실시간 PCR 시스템(Applied Biosystems)에서 power SYBR Green PCR 마스터 믹스를 사용하여 삼중으로 수행하였다. 9.6 ng DNA(9600 pg/6 pg/세포 = 1600 개 세포)를 10 μL 반응에 사용하였다. 다음 프라이머 쌍을 사용하였다: 인간 γ-글로빈 정방향(서열번호: 195), 및 역방향(서열번호: 196).

[0951] **통합 부위 분석.** 절차를 설명하기 위해, 도 27을 참조한다. 도 28d에 대한 무작위화 데이터를 포아송(Poisson) 회귀 삽입 모델(PRIM)을 사용하여 생성하여 마우스 참조 게놈(mm9)에서 각 염색체의 길이에 따라 중첩하지 않는 20 킬로염기 창에 대한 예상된 삽입률을 계산하였다. PRIM 알고리즘은 각 창 내에서 TA 디뉴클레오티드의 수, 창이 잔류하는 염색체, 및 고유 삽입의 총 수에 기반하여 통계 모델을 생성하였다. 각 창에 대해, 예상된 삽입 수를 계산하고 관찰된 삽입 수와 비교하여 p-값을 생성하였다. 그런 다음 본페로니-교정(Bonferroni-correction)을 적용하여 삽입된 트랜스포존의 검출을 위한 강화를 나타내는 창을 식별하였다. 그런 다음 TA를 함유하는 참조 게놈으로부터의 무작위 서열을 생성하고, Bowtie2를 사용하여 맵핑하고 실제 통합 데이터에 대해 플롯팅하였다. 계산 및 플롯은 R에서 ggplot2를 사용하여 수행하였다. 도면은 HOMER 및 ChIPseeker를 사용하여 그렸다.

[0952] **통합 부위 분석(역PCR).** 총 골수 세포의 접합부를 변형되어 다른 곳에서 기재된 바와 같은 역PCR에 의해 분석하였다(Hudecek 등, *Crit Rev Biochem Mol Biol* 52(4):355-80, 2017). 간단히 말해서, 골수 세포로부터의 게놈 DNA를 제조업체의 지침에 따라 Quick-DNA 미니프렙 키트(Zymo Research)에 의해 단리하였다. 5 - 10 μg의 DNA를 *SacI*로 소화시키고 분자내 반응을 촉진하는 조건 하에 재결찰시켰다. 결찰 혼합물을 페놀/클로로포름 추출 및 에탄올 침전으로 정제한 다음 KOD Hot Start DNA 폴리머라제를 사용하여 중첩 PCR(각 30 주기)에 사용하였다. 다음 프라이머를 사용하였다: EF1α p1 정방향(서열번호: 197) 및 역방향(서열번호: 198); EF1α p2 정방향(서열번호: 199) 및 역방향(서열번호: 200); 3'HS1 p1 정방향(서열번호: 201) 및 역방향(서열번호: 202); 및 3'HS1 p2 정방향(서열번호: 203) 및 역방향(서열번호: 204). 서열번호: 197-204에서, 밑줄 표시된 염기를 하류 클로닝에 사용한다. PCR 앰플리콘을 겔 정제하고, 클로닝하고, 서열분석하고, 정렬시켜 통합 부위를 식별하였다.

[0953] **RNA-seq 분석**을 Omega Bioservices(조지아주 노르크로스 소재)에 의해 수행하였다. 데이터를 OnRamp BioInformatics, Inc.(캘리포니아주 샌디에이고 소재)에 의해 개발된 HyperScale 아키텍처를 사용하여 Rosalind(rosalind.onramp.bio/에서 온라인 이용가능)에 의해 분석하였다. 판독물을 cutadapt를 사용하여 트리밍하였다. 품질 점수를 FastQC를 사용하여 평가하였다. 개별 샘플 판독물을 HTseq4를 사용하여 정량화하고 DESeq2 R 라이브러리를 사용하여 상대 Log 발현(RLE)을 통해 정규화하였다. DEseq2를 또한 사용하여 배수 변화 및 p-값을 계산하고 임의의 공변량 교정을 수행하였다. 차등적으로 발현된 유전자의 최종 히트맵을 위한 유전자의 클러스터링을 fpc R 라이브러리를 사용한 PAM(분할적 군집 분석) 방법을 사용하여 수행하였다. Interpro9, NCBI10, MSigDB11,12, REACTOME13, WikiPathways를 포함하는 여러 데이터베이스 소스를 강화 분석을 위해 참조하였다. 강화는 실험과 관련된 배경 유전자 세트와 관련하여 계산하였다.

[0954] p-값에 대한 log-스케일 배수 변화를 플롯팅하는 맞춤형 Python 스크립트로 볼케이노 플롯(volcano plot)을 생성하였다.

[0955] **동물**:

[0956] **연구 승인:** 동물을 수반하는 모든 실험은 워싱턴 대학(University of Washington)에 의해 제시된 기관 지침에 따라 수행하였다. 워싱턴 대학은 실험실 동물 관리 국제 평가 및 인증 협회(AALAC) 공인 연구 기관이며 이 대학에서 수행된 모든 살아있는 동물 작업은 실험실 동물 복지국(OLAW) 공중 보건 보증(PHS) 정책, USDA 동물 복지법 및 규정, 실험 동물 관리 및 사용에 대한 지침 및 기관 동물 관리 및 사용 위원회(IACUC) 정책을 따른다. 연구는 워싱턴 대학 IACUC(프로토콜 번호 3108-01)에 의해 승인되었다.

[0957] 생체외 및 생체내 HSPC 형질도입 연구를 완전한 인간 CD46 유전자좌를 함유하는 C57Bl/6-기반 유전자이식 마우스 모델(hCD46tg)로 수행하였다. 이러한 마우스는 인간과 유사한 패턴 및 수준에서 hCD46을 발현한다(Wang 등, *Mol Ther Methods Clin Dev.* 8:52-64, 2018).

[0958] **$Hbb^{th3}$/CD46+/+ 마우스의 교배 및 스크리닝:** 3 차례의 역교배 후, C46에 대한 $Hbb^{th3}$ 마우스 동형접합성을 gDNA에

대한 PCR(CD46F-5'(서열번호: 205) 및 CD46R 프라이머(서열번호: 206) 사용) 뿐만 아니라 CD46 MFI 측정을 허용하는 유세포 분석에 의해 확인하였다. $Hbb^{th3}/CD46^{+/+}$ 마우스의 지중해빈혈 표현형을 하기 기재된 바와 같이, Giemsa/May-Grünwald 염색 후 말초 혈액 도말에 의해 평가하였다.

[0959] **골수 $Lin^-$ 세포 이식**: 수용자는 6 - 8 주령의 암컷 C57BL/6 마우스였다. 이식 당일에, 수용자 마우스를 1000 Rad로 조사하였다. 조사 4 시간 후 $1x10^6$ 개 $Lin^-$ 세포를 꼬리 혈관을 통해 정맥내로 주사하였다. 이 프로토콜을 생체외 형질도입 $Lin^-$ 세포의 이식 및 2차 수용자로의 이식을 위해 사용하였다.

[0960] **HSPC 동원 및 생체내 형질도입**: 이 절차는 이전에 Richter, 등, (2016) *Blood* 128: 2206-2217에 기재되었다. HSPC를 인간 재조합 G-CSF의 s.c. 주사(5 μg/마우스/일, 4 일)(Amgen 캘리포니아주 사우전드 오크스 소재) 이어서 5 일째에 AMD3100(5 mg/kg)의 s.c. 주사(Sigma-Aldrich)에 의해 마우스에서 동원하였다. 또한, 동물은 바이러스 주사 16 시간 및 2 시간 전에 덱사메타손(10 mg/kg)을 i.p.로 받았다. AMD3100 후 30 및 60 분에, 동물에게 HDAd 벡터를 주사 당 각 바이러스에 대해 $4x10^{10}$ vp의 용량으로 안와하정맥총을 통해 정맥내로 주사하였다. 4 주 후, $O^6$BG/BCNU의 생체내 선택을 시작하였다.

[0961] **2차 골수 이식**: 수용자는 Jackson Laboratory로부터 수득한 6 - 8 주령의 암컷 C57BL/6 마우스였다. 이식 당일에, 수용자 마우스를 1000 Rad로 조사하였다. 생체내 형질도입된 CD46tg 마우스로부터의 골수 세포를 무균상태로 단리하고 계통-고갈된 세포를 MACS를 사용하여 단리하였다. 조사 4 시간 후 세포를 마우스 당 $1x10^6$ 개 세포로 정맥내로 주사하였다. 20 주차에, 2차 수용자를 희생시키고 혈액, 골수 및 비장으로부터 CD46+ 세포를 MACS에 의해 단리하거나 또는 상기 기재된 바와 같이, 동원 및 생체내 형질도입에 적용하였다. 모든 2차 수용자는 4 주차에 시작하여 면역억제를 받았다.

[0962] **혈액학적 분석**: 혈액 샘플을 EDTA-코팅된 튜브에 수집하고, 분석을 HemaVet 950FS(Drew Scientific)에서 수행하였다.

[0963] **조직 분석**: 2.5 μm 두께의 비장 및 간 조직 절편을 적어도 24 시간 동안 4% 포름알데하이드에 고정시키고, 탈수하고 파라핀에 포매하였다. 골수외 혈구형성의 조직학적 평가를 위해 헤마톡실린-에오신 염색을 사용하였다. 헤모시데린은 Perl의 프러시안 블루 염색에 의해 조직 절편에서 검출되었다. 간단히 말해서, 조직 절편을 증류수 중 동일한 부피(2%)의 페로시안화칼륨 및 염산의 혼합물로 처리한 다음 뉴트럴 레드로 대조염색하였다. 세포외 혈구형성 및 혈철소증을 정량화하기 위해, 적어도 3 마리 동물로부터의 5 개의 상이한 조직 절편에서 10 개의 무작위 영역을 마우스 그룹에 대해 맹검이 된 조사자들에 의해 평가하였다. 비장 크기를 비장 중량(mg)/체중(g)의 비로 평가하였다.

[0964] **혈액 분석 및 골수 사이토스핀**: 혈액 샘플을 EDTA-코팅된 튜브에 수집하고 분석을 HemaVet 950FS(Drew Scientific, 코네티컷주 워터베리 소재)에서 수행하였다. 말초 혈액 도말 및 골수 세포 사이토스핀을 각각 5 및 15 분 동안 Giemsa/May-Grünwald/Giemsa(Merck, 독일 다름슈타트 소재)로 염색하였다. 망상적혈구를 브릴리언트 크레실 블루로 염색하였다. 혈액 도말 시 망상적혈구를 계수한 조사자들은 샘플 그룹 할당에 대해 맹검이었다. 슬라이드에는 동물 번호만이 표시되었다(동물 당 5 개 슬라이드, 5 개의 무작위 1 $cm^2$ 절편).

[0965] **통계적 분석**: 데이터는 평균 ± 평균의 표준 오차(SEM)로 제시된다. 다중 그룹의 비교를 위해, 다중 비교를 위한 본페로니 사후-테스트와 함께 일원 및 이원 분산 분석(ANOVA)을 이용하였다. 하나의 그룹화 변수에 대한 그룹 간의 차이는 독립표본 양측 스튜던트 t 검정에 의해 결정하였다. 비모수 분석을 위해 Kruskal-Wallis 검정을 사용하였다. 통계적 분석은 GraphPad Prism 버전 6.01(GraphPad Software Inc., 캘리포니아주 라호이아 소재)을 사용하여 수행하였다. *p≤0.05, ** p≤00.0001. 0.05 미만의 P 값은 유의한 것으로 간주되었다.

[0966] ***결과***.

[0967] 정맥내로 주사된 HDAd5/35++ 벡터를 사용한 생체내 형질도입 연구를 위한 모델로서, 완전한 인간 CD46 유전자좌를 함유하고 따라서 인간과 유사한 패턴 및 수준에서 hCD46을 발현하는 유전자이식 마우스를 사용하였다(hCD46tg 마우스) (Kemper, 등, (2001) *Clin Exp Immunol* 124: 180-189).

[0968] **긴 β-글로빈 LCR을 함유하는 HDAd5/35++ 벡터**. 실시예 1에 기재된 연구에서, 1.6 kb β-글로빈 프로모터에 연결된 4.3 kb 미니 LCR(HS1 내지 HS4의 코어 요소 포함(Lisowski 등, *Blood* 110: 4175-4178, 2007))의 제어 하

에 γ-글로빈을 발현하는 HDAd5/35++ 벡터(도 23, "HDAd-짧은-LCR")(Wang 등, *J Clin Invest* 129: 598-615, 2019)를 사용하였다(Wang 등, *J Clin Invest* 129: 598-615, 2019; Li, 등, () *Mol Ther Methods Clin Dev* 9: 142-152, 2018). 본 실시예에서, γ-글로빈 유전자 발현을 최대화하기 위해 하기 요소를 함유하는 HDAd5/35++ 벡터를 구축하였다: *i)* 전장 HS5 내지 HS1 영역을 포함하는 21.5 kb LCR, *ii)* 1.6 kb β-글로빈 프로모터, *iii)* γ-글로빈 mRNA를 안정화하기 위한 β-글로빈 3'UTR, 및 *iv)* 3' HS1 영역. 벡터는 HDAd-긴-LCR(도 23, "HDAd-긴-LCR")로 명명하였다. 통합을 매개하기 위해, LCR-벡터를 SB100x/Flpe 발현 HDAd 벡터(도 23, "HDAd-SB")와 조합하여 사용한다. 트랜스포존 벡터(HDAd-짧은-LCR 및 HDAd-긴-LCR)는 Flpe 재조합효소의 존재 하에 이식유전자 카세트의 원형화를 허용하는 SB100x 트랜스포사제 및 *frt* 부위에 의해 인식된 도립된/지시 반복(IR/DR) 모티프를 함유한다. HDAd-짧은-LCR 및 HDAd-긴-LCR은 둘 다 또한 보편적으로 활성인 EF1α 프로모터의 제어 하에 돌연변이체 $O^6$-메틸구아닌-DNA 메틸트랜스퍼라제($mgmt^{P140K}$)에 대한 유전자를 보유하여 저용량 $O^6$BG/BCNU 처리에 의해 안정하게 형질도입된 세포의 선택을 허용하였다(Hausl 등, *B. Mol Ther* 18(11):1896-906, 2010; Neff 등, *J Clin Invest* 112(10):1581-8, 2003).

[0969] **생체외 HSPC 형질도입/이식 연구**. 인간에서, CD46은 모든 유핵 세포에서 발현되지만, 마우스에서 상응하는 오솔로그는 고환에서만 존재한다. 정맥내로 주사된 HDAd5/35++ 벡터를 사용한 생체내 형질도입 연구를 위한 모델로서, 완전한 인간 CD46 유전자좌를 함유하고 따라서 인간과 유사한 패턴 및 수준에서 hCD46을 발현하는 유전자이식 마우스를 사용하였다(CD46tg 마우스)(Wang 등, *Mol Ther Methods Clin Dev* 8:52-64, 2018). 연역적으로, SB100x가 32.4 kb 트랜스포존의 통합을 매개할 수 있는지가 알려져 있지 않았기 때문에, HSPC 형질도입 효능이 제어될 수 있는 설정으로 생체외 HSPC 형질도입 연구를 수행하였다. HSPC에 대해 강화된 세포 분획인 CD46tg 마우스 골수 계통-음성($Lin^-$) 세포를 HDAd-긴-LCR + HDAd-SB로 생체외 형질도입하였다(도 24a). 그런 다음 생체외 형질도입된 세포를 치명적으로 조사된 C57Bl/6 마우스에 이식하였다. 4 주차에 생착률은 CD46-양성 PBMC를 기반으로 95% 초과였다. 이식 1 개월 후, 마우스를 4 차례의 $O^6$BG/BCNU 처리에 적용하여 통합된 γ-글로빈/mgmt 이식유전자를 갖는 전구체를 선택적으로 확장하였다(도 24a). 생체내 선택의 각 차례마다, γ-글로빈-양성 말초 적혈구(RBC)의 백분율을 증가시켜, 연구 종료 시 95% 초과에 도달하였다(도 24b). 20 주차에, 동물을 희생시키고 골수 단핵 세포(MNC)를 분석하였다. qPCR에 의해 측정된 평균 VCN은 세포 당 2.8 개의 카피였다. γ-글로빈 발현은 유세포 분석에 의해 적혈구 $Ter119^+$ 세포의 85.46(+/-5.9)% 및 비-적혈구($Ter119^-$) 골수 MNC의 14.54(+/-2.3)%에서 검출되었다(도 24c).

[0970] γ-글로빈 발현이 SB100x 통합된 이식유전자로부터 유래되었음을 입증하기 위해, 이식 후 20 주차에 수확된 골수 단핵 세포(MNC)로부터의 게놈 DNA에서 역PCR(iPCR) 분석을 수행하였다. iPCR 프로토콜은 게놈 DNA의 SacI로의 소화, 재결찰/원형화 단계, 중첩 PCR 및 벡터/염색체 접합부의 서열분석을 수반한다(도 24d). (도 24e)는 3 가지 대표적인 PCR 생성물 및 염색체 4, 15, 및 X 상에서 통합 부위의 국소화를 나타낸다. 생성물의 서열분석은 벡터 IR/DR-염색체 접합부에서 TA 디-뉴클레오티드를 포함하는 SB100x 매개 통합에 전형적인 벡터/염색체 접합부를 입증하였다(도 24f). 요약하면, 생체외 HSPC 형질도입 연구에서, 긴 글로빈 LCR은 SB100x 통합된 트랜스포존으로부터 기원하는 높은 수준의 γ-글로빈 발현을 부여하였다.

[0971] **짧은 LCR vs 긴 LCR를 함유하는 HDAd5/35++ 벡터를 사용한 CD46b 유전자이식 마우스에서 생체내 HSPC 형질도입**. HDAd-긴-LCR 및 미니LCR을 함유하는 실시예 1에서 이전에 사용된 벡터(Wang 등, *J Clin Invest* 129: 598-615, 2019; Li 등, *Mol Ther Methods Clin Dev* 9: 142-152, 2018)(본원에서 "HDAd-짧은-LCR"로 지칭됨)를 나란히(side-by-side) 비교하였다 (도 23). CD46-유전자이식 마우스를 G-CSF/AMD3100으로 동원하고 정맥내로 벡터를 주사하고, 5 주 후, 생체내 선택에 적용하였다(도 25a). γ-글로빈-양성 적혈구(RBC)의 백분율은 각 차례의 생체내 선택으로 증가하여 20 주차에 두 백터에 대해 95% 초과에 도달하였다(도 25b). 20 주 샘플로부터의 RBC 용해물에 대해 수행된 HPLC는 벡터 사이의 γ-글로빈/성체 마우스 α-글로빈의 백분율에서 유의한 차이를 나타내지 않았다(도 25c). 이는 또한 mRNA 수준에서 반영되었다(도 25d).

[0972] qPCR에 의해 20 주차에 측정된 골수 단핵 세포(MNC)에서 벡터 카피 수는 세포 당 2.5 개 카피였고(도 25e) 벡터 사이에 유의한 차이는 없었다. 이는 "긴" 32.4 kb 트랜스포존의 통합이 "짧은" 11.8 kb 트랜스포존의 통합만큼 효율적이었음을 나타낸다. 벡터를 사용한 생체내 HSPC 형질도입 후 32.4 kb 트랜스포존의 SB100x-매개 통합은 대부분의 적혈구 세포에서 γ-글로빈 발현에도 불구하고 혈액학적 이상(20 주)을 유발하지 않았다(도 26b). 세포 골수의 조성물(도 26c) 및 골수 $Lin^-$ 세포의 콜로니 형성-가능성(도 26d)은 그룹 사이에 유의하지 않았다.

[0973] 장기간 재증식 HSPC에서 발생된 생체내 형질도입 및 SB100x-매개 통합을 입증하기 위한 2차 이식에서, 세포 골

수의 조성물(도 26c) 및 골수 $Lin^-$ 세포의 콜로니 형성-가능성(도 26d)은 그룹 사이에 유의하지 않았다. 이식된 골수 Lin- 세포를 hCD46 이식유전자 없이 치명적으로 조사된 C57Bl/6 마우스로의 생체내 HSPC 형질도입 후 20 주차에 수확하였다). 2차 수용자에서 다중-계통 재구축을 구동하는 이식된 세포의 능력을 16 주 기간에 걸쳐 평가하였다. "1차" 생체내 HSPC 형질도입된 마우스에서와 같이, 골수의 세포 조성물 또는 말초 혈액에서 혈액학적 매개변수에 대한 높은 수준의 글로빈 발현 효과는 관찰되지 않았다.

[0974] 또한 20 주차에 수확된 골수 $Lin^-$ 세포를 사용하여 게놈-와이드 통합 부위 분석을 수행하였다. 이 검정에서, 선형 증폭-매개 PCR(LAM-PCR) 전략은 통합 접합부의 서열분석이 이어진다(도 27). 마우스 게놈에 걸친 통합 부위의 분포는 도 28a에 제시되어 있다. 통합된 이식유전자 카세트를 정확하게 처리하였고, 식별된 IR/DR 염색체 접합부는 TA 디뉴클레오티드를 함유하였다(도 28b). 대부분의 통합은 유전자간 및 인트론 영역 내에 각각 83% 및 17% 빈도로 있었다(도 28c). 통합은 전체 마우스 게놈의 임의의 주어진 창에서 우선적인 통합 없이 무작위였다(도 28d). *원종양형성 유전자* 내에서 또는 근처에서 통합은 발견되지 않았다. 이 SB100x-매개 통합 패턴은 이전 연구와 일치한다(Richter 등, *Blood* 128(18):2206-17, 2016; Neff 등, *J Clin Invest* 112(10):1581-8, 2003; Kemper 등, *Clin Exp Immunol.* 124(2):180-9, 2001; Zhang 등, *PLoS One* 8(10):e75344, 2013; Yant 등, *Nat Biotechnol* 20(10):999-1005, 2002).

[0975] **2차 수용자의 분석**. 생체내 형질도입이 장기간 재증식 HSPC에서 발생하였음을 입증하기 위해, 생체내 HSPC 형질도입 후 20 주차에 수확된 골수 $Lin^-$ 세포를 HDAd-짧은-LCR 및 HDAd-긴-LCR과 함께 치명적으로 조사된 C57Bl/6 마우스(hCD46 이식유전자 없이)에 이식하였다. 2차 수용자에서 다중-계통 재구축을 구동하는 이식된 세포의 능력을 16 주 기간에 걸쳐 평가하였다. PBMC에서 CD46 발현에 기반한 생착률은 95%였고 안정적으로 유지되었다(도 29a). 유세포 분석에 의해 측정된 RBC의 γ-글로빈 마킹은 90 내지 95% 범위에 있었고 안정되었다(도 29b). $γ$-글로빈$^+$ RBC의 백분율에서 두 벡터 사이에 유의한 차이는 없었다. 평균 통합 벡터 카피 수는 또한 장기간 재증식 세포에서 두 트랜스포존의 통합이 동일하게 효율적이었음을 나타내는 2 개의 벡터 사이에 유의하게 상이하지 않았다(도 29c). 흥미롭게도, 마우스 성체 글로빈 쇄에 대한 γ-글로빈의 백분율은 HDAd-긴-LCR 벡터가 마우스 α-글로빈의 20-25%에 도달하는 경우 시간 경과에 따라 증가하였다(도 29d 및 29e). 대조적으로, HDAd-짧은-LCR 형질도입된 골수 세포의 2차 수용자에서 γ-글로빈/마우스 α-글로빈의 백분율은 증가하지 않았다. γ-글로빈 발현 적혈구 세포의 백분율은 HDAd-긴-LCR에 대해 상당히 더 높았다(도 29f). 더 높은 γ-글로빈 발현 수준을 부여하는 것 외에도, 긴 LCR은 또한 비-적혈구 분획($Ter119^-$)에 비해 적혈구($Ter119^+$) 분획에서 γ-글로빈 발현 골수 세포의 상당히 더 높은 백분율에 의해 제시된 바와 같이 더 엄격한 적혈구-특이적 발현을 제공하였다(도 27h). 골수 MNC에서 세포 당 벡터 수 카피는 생체내 HSPC 형질도입 후 16 주차에 수확된 경우 HDAd-짧은-LCR 및 HDAd-긴-LCR 사이에서 통계적으로 유의하지 않았다(도 27i). "1차" 생체내 HSPC 형질도입된 마우스에서와 같이, 골수의 세포 조성 또는 말초 혈액에서 혈액학적 매개변수에 대한 높은 수준의 글로빈 발현 효과는 관찰되지 않았다(도 30a - 30d).

[0976] **인간 CD34+ 형질도입, 시험관내 선택, 및 적혈구 분화 후 2 개 백터의 비교.**

[0977] 마우스 적혈구 세포와 같은 이종 시스템에서 인간 β-글로빈 LCR의 분획은 LCR 내에 결합하는 전사 인자의 보존 결여로 인해 차선일 수 있다. 따라서 인간 세포에서 시험관내 연구를 수행하였다(도 31a). GCSF-동원된 건강한 공여자로부터 수득된 인간 CD34+ 세포를 4000 vp/세포의 총 MOI, 즉, 대부분의 CD34+ 세포의 형질도입을 부여하는 MOI에서 HDAd-긴-LCR + HDAd-SB 또는 HDAd-짧은-LCR + HDAd-SB로 형질도입하였다(Li 등, *Mol Ther Methods Clin Dev* 9: 390-401, 2018). 그런 다음 형질도입된 세포를 통합된 이식유전자를 갖는 세포에 대해 적혈구 분화(ED) 및 $O^6$BG/BCNU 선택에 적용하였다. 18 일에 걸쳐 형질도입된 세포의 확장 동안, 대부분의 에피솜 벡터는 손실된다. ED 종료 시, 유세포 분석에 의한 HDAd-긴-LCR + HDAd-SB 설정에 대해 상당히 더 높은 백분율의 γ-글로빈+ 유핵 세포(즉, 핵을 상실한 망상적혈구)가 발견되었다(도 31b). HPLC 분석은 또한 HDAd-긴-LCR + HDAd-SB-형질도입된 세포에서 상당히 더 높은 γ-글로빈 쇄 수준을 입증하였다(도 31c).

[0978] **중간형 지중해빈혈 γ-글로빈 수준의 마우스 모델에서 HDAd-짧은-LCR vs HDAd-긴-LCR 생체내 HSPC 형질도입 연구.** 이러한 연구를 위해(4 차례 초과) (CD46+/+) 마우스를 마우스 Hbb-베타1 및 -베타2 유전자 결실에 대해 이형접합성인 $Hbb^{th3}$ 마우스와 교배하였다(Yoshida 등, *Sci Rep* 7:43613, 2017). 생성된 $Hbb^{th3}$/$CD46^{+/+}$ 마우스는 중간형 지중해빈혈에 대한 전형적인 표현형을 갖는다(Wang 등, *J Clin Invest*, 129: 598-615. 2019).

$Hbb^{th3}/CD46^{+/+}$ 마우스를 동원하고, 정맥내로 HDAd-긴-LCR 및 HDAd-짧은-LCR 벡터 시스템을 주사하고, 4 주 후에 생체내 선택에 적용하였다(도 32a 및 32e). 중요하게는, 말초 적혈구의 γ-글로빈 마킹은 생체내 선택의 두번째 주기 후 이미 평균 40%에 도달하였고, 세번째 주기 후에 10 마리 마우스 중 9 마리에서 90% 초과에 도달하였고, HDAd-긴-LCR로 생체내 형질도입 후 12 주 차에 모든 마우스에서 거의 100%를 유지하였다(도 32b 및 32f). 대조적으로, HDAd-짧은-LCR로 형질도입된 마우스의 경우, 7 마리 마우스 중 2 마리의 RBC에서 100% γ-글로빈 마킹에 도달하기 위해 4 회 생체내 선택 주기가 필요하였고 100% 마킹은 형질도입후 16 주차에만 달성되었다. 100% 마킹률에서, 성체 마우스 α-글로빈 쇄보다 인간 γ-글로빈의 백분율(HPLC에 의해 측정)은 두 벡터에 대해 시간 경과에 따라 증가하여(대부분 질환 배경으로 인해) HDAd-긴-LCR 및 HDAd-짧은-LCR로 생체내 형질도입 후 16 주차까지 각각 평균 22%(max: 35%) 및 11%(max: 19%)에 도달하였다(도 32g 및 32h; 21 주 데이터의 경우 도 32c 및 32d). CD46tg 마우스에서 관찰된 것과 유사하게, 골수 단핵 세포의 분석은 두 벡터에 대해 비슷한 VCN을 나타내었고 HDAd-긴-LCR에 대한 적혈구 세포에서 더 높은 글로빈 발현 수준을 나타내었다(도 33). 요약하면, 이들 데이터는 *i)* 100% 마킹에 도달하기 위해 덜 강력한 생체내 선택을 필요로 하고 *ii)* RBC에서 γ-글로빈 수준을 달성함으로써 HDAd-짧은-LCR보다 HDAd-긴-LCR의 우수성이 이론적으로 SCD 및 주요 지중해빈혈 환자에서 치유력이 있어야 함을 입증한다.

[0979] **혈액학적 매개변수의 교정.** 표현형 교정은 상이한 시점에서 제시된다. 처리 전 및 긴 LCR로 처리 후 10 주차에 C57BL6 및 Townes SCA 마우스의 정규화된 적혈구 형태를 비교하는 현미경 사진(도 34) 및 처리 전 Townes 마우스, 및 긴 LCR로 처리 후 10 주차에 Townes 마우스에 대해 정규화된 적혈구생성(망상적혈구 계수)을 나타내는 현미경 사진이 제시된다(도 35). 14 주차에, Giemsa 염색 및 May-Grunwald 염색으로 염색된 혈액 세포 형태가 제시된다(도 36a). 처리 후 16 주차에, 마우스를 희생시켰다. 처리된 $Hbb^{th3}/CD46^{+/+}$ 마우스의 말초 혈액 도말에서 지중해빈형 표현형의 역전을 나타내는, 저변색성의 고도로 단편화된 부동변형적혈구증가적 기준선 RBC를 거의 정색성의 잘 형성된 RBC로 대체하였다(도 37a, 왼쪽 패널, 21 주 데이터의 경우 도 36b 참조). 말초 혈액에서 망상적혈구의 수준은 정상 CD46tg 마우스와 비슷하였다(도 37a, 오른쪽 패널, 또한 도 39 참조). 21 주에 대해 유사한 데이터는 도 36b의 오른쪽 패널에서 찾을 수 있다. 대조군 및 처리된 $Hbb^{th3}/CD46^{+/+}$ 마우스로부터의 사이토스핀에서, 호염기성 적혈구모세포의 유병률에 의해 나타낸 $Hbb^{th3}/CD46^{+/+}$ 마우스의 골수에서 적혈구 계통 성숙의 차단과 대조적으로, 골수 사이토스핀에서, 성숙 다염성 및 정염성 적혈구모세포가 우세하였다(도 37b, 21 주 데이터의 경우 도 36c 참조). 긴 LCR, 짧은 LCR, 및 대조군 CD46tg 벡터로 형질도입된 마우스의 정규화된 적혈구 매개변수가 제시된다(도 38). 생체내 형질도입 후 16 주차에 혈액학적 매개변수는 두 벡터에 대한 전처리 매개변수와 비교하여 유의하게 개선되었다(도 38, 39a). 백혈구, 적혈구, MCHC, MCV, 및 RDW-CV의 경우 이들은 CD46tg 대조군과 구별할 수 없었다(도 39a). 그러나, HDAd-짧은-LCR보다 HDAd-긴-LCR 벡터로 처리된 동물에 유리한 상당한 차이가 있으며, 구체적으로, 말초 혈액에서 망상적혈구의 백분율은 미처리, HDAd-짧은-LCR, 및 HDAd-긴-LCR-처리 $Hbb^{th3}/CD46^{+/+}$ 마우스에 대해 각각 40.9 vs 26.8 vs 9.2%였다(도 38). 또한, 헤모글로빈 수준 및 적혈구용적률은 HDAd-긴-LCR-처리 그룹에서 더 높았다.

[0980] **골수외 조혈 및 혈철소증의 교정.** 보상 혈구형성의 측정가능한 특성인 비장 크기는 두 벡터로 처리된 동물에서 정상으로 감소되었고, 이에 의해 HDAd-긴-LCR 및 HDAd-짧은-LCR 사이에 유의한 차이는 없었다(도 40a). $Hbb^{th3}/CD46^{+/+}$ 마우스와 대조적으로, 골수외 적혈구생성의 병소는 HDAd-긴-LCR로 처리 후 비장 및 간 절편에서 관찰되지 않았고 제한된 골수외 적혈구생성만이 HDAd-짧은-LCR-처리된 마우스에서 검출되었다(도 40b). 비장 및 간에서 강렬한 혈철소증은 미처리 $Hbb^{th3}/CD46^{+/+}$ 마우스에서 우세하였다(도 41, 두번째 패널. 조직의 Perl 염색 후 신호는 CD46tg(도 41, 첫번째 패널( 및 HDAd-긴-LCR 처리된 $Hbb^{th3}/CD46^{+/+}$ 마우스(도 41, 세번째 패널)에 대해 비교적 낮았던 반면, 비장 조직 $cm^2$ 당 2.7(+/-0.8)-배 더 많은 청색 스팟이 HDAd-긴-LCR-처리된 동물보다 HDAd-짧은-LCR에 대해 계수되었다(N=5).

[0981] 요약하면, HDAd-긴-LCR-처리된 동물에서 망상적혈구, 혈액 매개변수, 세포외 조혈 및 혈철소증은 대조군 CD46tg 마우스와 상당히 상이하지 않았으며, 이는 완전한 표현형 교정을 나타낸다. 또한, HDAd-긴-LCR은 여러 표현형 매개변수에서 지중해빈혈 마우스를 치유하는 데 있어서 HDAd-짧은-LCR보다 우수한 것으로 입증되었으며, 긴-LCR에서 발현된 더 높은 γ-글로빈 수준에 기인할 가능성이 더 높다.

[0982] **인간 CD34+ 형질도입 및 적혈구 분화 후 2 개 벡터의 비교.** 마우스에서 데이터를 통합하기 위해, 인간 세포에

서 시험관내 연구를 수행하였다(도 31a). GCSF-동원된 건강한 공여자에서 수득된 인간 $CD34^+$ 세포를 4000 vp/세포의 총 MOI, 즉, 대부분의 $CD34^+$ 세포의 형질도입을 부여하는 MOI에서 HDAd-긴-LCR + HDAd-SB 또는 HDAd-짧은-LCR + HDAd-SB로 형질도입하였다(Yang 등, *Proc Natl Acad Sci USA*. 92(25):11608-12, 1995). 그런 다음 형질도입된 세포를 통합된 이식유전자를 갖는 세포에 대해 적혈구 분화(ED) 및 $O^6$BG/BCNU 선택에 적용하였다. 18 일에 걸쳐 형질도입된 세포의 확장 동안, 대부분의 에피솜 벡터가 손실된다.

[0983] 골수를 $Hbb^{th3}$/CD46tg 마우스의 생체내 HSC 형질도입 후 21 주차에 수확하였다. (도 42a) 골수 MNC에서 세포 당 벡터 카피 수. 두 그룹 사이의 차이는 유의하지 않았지만 더 큰 샘플 크기로 분석하는 경우 유의하게 될 수 있다. (도 42b, 42c) γ-글로빈 발현의 적혈구 특이성. (도 42b) γ-글로빈 발현 적혈구($Ter119^+$) 및 비-적혈구($Ter119^-$) 세포의 백분율. *$p<0.05$. 통계적 분석은 이원 ANOVA를 사용하여 수행하였다.

[0984] 아데노바이러스 공여자 벡터의 투여 전 CD46tg 및 $CD46^{+/+}$/$Hbb^{th-3}$ 마우스의 간 및 비장 절편에서 헤마톡실린/에오신 염색에 의한 골수외 혈구형성(도 43). 철분 침착은 Perl 염색에 의해 비장에서 헤모시데린의 세포질 청색 색소로 제시된다.

[0985] ED 종료 시, γ-글로빈$^+$ 제핵 세포(즉, 핵이 손실된 망상적혈구)의 상당히 더 높은 백분율이 유세포 분석에 의해 발견되었고(도 31b) 또한 HDAd-짧은-LCR보다 HDAd-긴-LCR 설정에서 HPLC에 의해 상당히 더 높은 γ-글로빈쇄 수준이 발견되었다(도 31c). 18 일차에 측정된 벡터 카피 수는 두 벡터에 대해 2개였다(도 31d).

[0986] 요약하면, 마우스를 사용한 생체외 및 생체내 HSPC 형질도입 연구 뿐만 아니라 인간 HSPC를 사용한 시험관내 연구는 혈색소병증의 유전자 요법에 대한 HDAd-긴-LCR의 연관성을 뒷받침한다.

[0987] ***논의***. 이 실시예는 백혈구분리반출술, 골수절제술 및 HSPC 이식을 필요로 하지 않는 생체내 HSPC 유전자 요법 접근법의 임상 개발과 관련된 작업을 기재한다(Richter 등, *Blood*. 128(18):2206-17, 2016). 이는 특히 고령의 환자 및 동반이환 환자에서 혈색소병증의 생체외 HSPC 유전자 요법에 대한 광범위한 적용에 중요한 장애물이다. 이 접근법의 안전성 및 효능은 여러 뮤린 질환 모델(Wang 등, *J Clin Invest*. 129(2):598-615, 2019; Wang 등, *Blood Adv*. 3(19):2883-94, 2019; Li 등, *Mol Ther Methods Clin Dev*. 9:390-401, 2018) 및, 최근에 비-인간 영장류(Li 등, *23rd Annual ASGCT meeting*. 2020; abstract #546)에서 입증되었다. 두 종에서, 정맥내 HDAd5/35++ 주사와 연관된 주요 문제, 즉 급성 선천적 면역 반응은 전염증성 사이토카인을 차단하는 예방적 레지멘에 의해 해결되었다.

[0988] 생체외 HSPC 유전자 요법 설정에서 주요 지중해빈혈 및 SCD 환자에서 치유적 γ- 또는 β-글로빈 발현 수준에 도달하는 것은 여전히 과제이다. 이는 HSPC 형질도입 과정을 최적화하거나 또는 감염 다중도를 증가시킴으로써 통합된 이식유전자 카피의 수를 증가시키는 접근법을 필요로 한다. 그러나, VCN을 증가시키는 것은 유전독성을 유도하는 위험이 있다. 다른 시도는 글로빈 발현 카세트를 추가로 최적화하는 데 초점을 맞춘다(Li 등, *Cancer Res*. 80(3):549-60, 2020). 높은-페이로드 용량 HDAd 벡터를 사용하면, 렌티- 및 rAAV 벡터에 대해 설정된 게놈 크기 제한을 넘어설 기회가 있다. 본 연구는 γ-글로빈의 치유적 수준이 총 길이가 29 kb인 β-글로빈 LCR/프로모터 요소를 수용하는 통합 HDAd5/35++ 벡터를 사용한 생체내 HSPC 유전자 요법에 의해 RBC에서 달성될 수 있음을 입증한다.

[0989] 지중해빈혈 마우스에서, RBC에서 100% γ-글로빈 마킹은 HDAd-짧은-LCR 처리된 동물과 비교하여 HDAd-긴-LCR로 처리된 마우스의 초기 및 더 적은 주기의 $O^6$BG/BCNU 생체내 선택로 달성되었다. 이는 접근법의 임상 번역에 중요하다. $O^6$BG/BCNU 생체내 선택 시스템은 γ-글로빈 양성 RBC 백분율의 제어 증가를 허용하지만, 또한 일시적 백혈구감소증 및 GI-관의 부작용을 유발한다(Wang 등, *J Clin Invest*. 129(2):598-615, 2019). HDAd-긴-LCR을 사용한 덜 강렬한 생체내 선택의 요건에 대한 잠재적 설명은 긴 LCR이 $O^6$BG/BCNU에 대한 내성을 제공하는 $mgmt^{P140K}$ 유전자의 발현을 구동하는 EF1α 프로모터의 침묵을 방지한다는 것일 수 있다. 이 가설은 골수 MNC에서 mgmt mRNA 수준(VCN으로 정규화됨)이 HDAd-긴-LCR에 대해 상당히 더 높았다는 관찰에 의해 뒷받침된다(도 48).

[0990] 이 연구는 HDAd-긴-LCR을 사용한 생체내 접근법의 치료 측면에 집중하였지만, 미래에 해결해야 할 다수의 기계적 문제가 남아 있다. 이러한 개방 질문 중 하나는 긴 LCR이 떨어져 있는 이웃 유전자의 전사활성화를 방지하

는지 여부이다. 또한, mRNA 수준에서도 반영되는 HDAd-긴-LCR의 더 높은 γ-글로빈 발현 수준이 더 활성인 전사 개시 또는 통합 벡터 카피의 더 적은 침묵, 또는 둘 다에 기인하는지 여부는 완전히 명확하지 않다. HDAd-긴-LCR-처리된 $Hbb^{th3}$/CD46 마우스에서, 2차 수용자에서 CD46tg 모델에서도 볼 수 있는 현상인 마우스 성체 글로빈 쇄에 대한 γ-글로빈의 백분율이 시간 경과에 따라 증가하였다는 관찰은 구체적으로 장기간 재증식 세포에서 침묵이 시간 경과에 따라 발생하였고 긴-LCR이 이에 대해 보호되었음을 나타낼 수 있다. 통합 벡터 카피 당 더 높은 $mgmt^{P140K}$ mRNA 수준(도 48)은 또한 긴-LCR이 침묵으로부터 보호한다는 가설을 뒷받침한다. 이러한 문제를 해결하기 위해, 추가 연구는 형질도입된 $CD34^+$ 세포 클론에 초점을 맞추고 LAM-PCR/NGS(통합 부위), 염색체 형태 캡처 기술, 및 RNA-Seq를 사용한 게놈-와이드 분석을 포함할 것이다. 이러한 연구에 대한 전제조건은 SB100x 트랜스포사제-매개 이식유전자 통합 및 생체내 선택 과정이 원치않은 게놈 변경/재배열을 촉발하지 않는다는 것이다. 이를 평가하기 위한 시도에서, RNA-Seq를 시험관내에서 SB100x-매개 통합 및 $O^6$BG/BCNU 선택 후 mgtm/GFP 이식유전자를 안정하게 발현하는 인간 $CD34^+$ 세포 상에서 수행하였다(도 47a). 단지 176 개 유전자의 약간 변경된 발현은 우선적으로 히스톤 유전자를 발견하였다(도 47b). 이는 SB100x가 중요한 유전독성을 발휘하지 않음을 나타내며, 또한 통합 부위 분석에서 클론 우성의 부재 및 장기간 연구에서 혈액학적 부작용의 부재에 의해 뒷받침된다.

[0991] HDAd5/35++-기반 SB100x 시스템을 사용한 생체내 HSPC 형질도입/선택 후 16 내지 23 주차에 골수 MNC에서 분석된 통합된 이식유전자의 카피 수는 13.8(Wang 등, *J Clin Invest.* 129(2):598-615, 2019) 내지 32.4 kb 범위의 트랜스포존에 대해 세포 당 2 개 카피였다. 촉매적으로 프라이밍된 트랜스포존/트랜스포사제 복합체를 형성하기 위해, 트랜스포존의 2 개 단부는 트랜스포사제 분자에 의해 물리적으로 매우 가깝게 함께 유지되어야 한다(Uchida 등, *Nat Commun.* 10(1):4479, 2019). 이러한 제한은 frt 측면을 HDAd 벡터로 통합함으로써 해결되었고 공동 발현된 Flpe 재조합효소에 의해 인식되어 트랜스포존의 원형화로 이어진다(Turchiano 등, *PLoS One.* 9(11):e112712, 2014). 여기에 보고된 데이터는 이 과정이 HDAd5/35++ 벡터에 의해 운반되는 트랜스포존의 크기에 크게 의존하지 않고 통합을 이룰 수 있음을 시사한다.

[0992] 이 연구는 확장된 TAD/LCR 코어 요소를 사용하는 것인 치료 이식유전자의 발현 수준을 증가시킴을 입증한다. β-글로빈 LCR은 수십 년 동안 연구되었지만, 다른 유전자/클러스터에 대한 TAD 코어 요소는 덜 특성화되었다. TAD의 중앙 크기는 880 kb이다. 고처리량 염색체 형태 캡처(3C) 검정 및 후속 4C, 5C 및 Hi-C 프로토콜 뿐만 아니라 섬유-Seq 검정의 추가 발전에 따라, 조절 게놈의 심문은 빠른 속도로 진행될 것이고, 유전자 요법 목적을 위해, 중요한 코어 요소만을 함유하는 TAD를 전달할 수 있다(Liu 등, *BMC Genomics.* 20(1):217, 2019).

[0993] 요약하면, 현재 실시예는 마우스에서 HSPC의 생체내 형질도입을 위해 HDAd5/35++ 벡터의 맥락에서 더 큰 조절 요소를 이용하는 것이 주요 지중해빈혈 및 겸상 세포 빈혈에 대해 치유적인 것으로 생각된 유전자 발현 임계값을 충족하는 γ-글로빈 수준을 부여하는 벡터를 산출하였음을 나타낸다.

[0994] 인간 β-글로빈 유전자 클러스터는 염색체 11에 있고 ~100 kb에 걸쳐 있다. β-글로빈 유전자좌는 활성 염색질 hub(ACH)라고 불리는, 시스-조절 요소 및 활성 β-글로빈 유전자로 구성된 적혈구-특이적 공간 구조를 형성하는 것으로 제안되었다(Tolhius 등, *Mol Cell,* 10:1453-1465, 2002). 코어 ACH는 발달적으로 보존되고 글로빈 LCR이라고 불리는 상류 5' DNAse 과민성 영역 1 내지 5, 및 하류 3'HS1 뿐만 아니라 적혈구-특이적 트랜스작용화 인자로 이루어진다(Kim 등, *Mol Cell Biol.,* 27:4551-65, 2007). 유전자 요법 적용을 위해, HS1 내지 HS5를 함유하는 23 kb β-글로빈 LCR 및 3 kb 3'HS1 영역은 유전자이식 마우스에서 시스-연결된 유전자에에 대한 높은 수준의 적혈구-특이적 위치 독립적 발현을 부여하였음에 유의한다(Grosveld, *Cell*, 51:975-985, 1987). 이 LCR의 제어 하에 이식유전자를 전달하는 도구는 30+ kb HDAd 벡터와 함께 이용가능하다.

[0995] 많은 유전적 질환의 교정은 치료 유전자의 높은 수준 및 조직-제한된 발현을 필요로 하며, 이는 LCR을 이용함으로써 달성될 수 있다(Li 등, *Blood* 100: 3077-3086, 2002). β-주요 지중해빈혈 및 겸상 세포 빈혈의 치유를 위해, HSPC에서 약 20% 유전자 마킹 및 적혈구 세포에서 20% 치료-글로빈 쇄(β- 또는 γ-글로빈) 생산이 필요한 것으로 생각된다(Fitzhugh 등, *Blood* 130: 1946-1948, 2017). 크기 제한으로 인해, β-글로빈 LCR의 절두된 형태만이 유전자 발현 수준을 교정하기 위한 요건을 충족하기 어렵게 만드는 렌티바이러스 벡터에서 사용될 수 있다(Uchida, 등, *Nat Commun* 10: 4479, 2019). 렌티바이러스-매개 HSPC 형질도입 후 발현 수준을 증가시키는 전략은 벡터 용량 및 따라서 통합된 이식유전자 카피 수를 증가시키는 것이다. 그러나 이 접근법은 유전독성 및 종양 형성의 위험을 향상시킨다. 다른 시도는 글로빈 발현 카세트를 추가로 최적화하는 것에 초점을 맞춘다

(Uchida, 등, (2019) *Nat Commun* 10: 4479). 삽입 용량이 30 kb인 HDAd 벡터는 후자 개념을 개발하기 위한 이상적인 도구이다. 이 실시예에서, 29 kb γ-글로빈 발현 카세트를 보유하는 HDAd5/35++ 벡터를 생성하고 CD46-유전자이식 마우스에서 시험관내 및 생체내 HSPC 형질도입 후 테스트하였다.

[0996] HDAd 벡터 시스템에서, γ-글로빈 카세트의 통합은 SB100x 트랜스포사제에 의해 매개된다. SB/트랜스포존 시스템을 사용한 비-바이러스 유전자 전달은 CD19 CAR T-세포 요법(Kebriaei 등, *J Clin Invest* 126: 3363-3376, 2016), 연령 관련 황반 변성(Hudecek 등, *Crit Rev Biochem Mol Biol* 52: 355-380, 2017; Thumann 등, *Mol Ther Nucleic Acids* 6: 302-314, 2017), 및 알츠하이머병(Eyjolfsdottir 등, *Alzheimers Res Ther* 8: 30, 2016)에 대해 임상적으로 사용되고 있다. HD-Ad 매개된 SB 유전자 전달은 Kay 및 Ehrhardt 그룹에 의해 개척되었다. 이들의 연구에서, 트랜스포존은 4 kb-6 kb로 상대적으로 작았다(Hausl 등, *Mol Ther* 18: 1896-1906, 2010; Yant 등, *Nat Biotechnol* 20: 999-1005, 2002). 현재 실시예는 SB100x가 비슷한 VCN(세포 당 2-3 개 카피)에 기반하여 11.8 kb 트랜스포존과 비슷한 효능으로 32.4 kb 트랜스포존을 통합할 수 있음을 처음으로 입증한다. 그 자체로 이 발견은 SB-매개 통합의 효능이 SB 트랜스포존의 크기와 역으로 상관관계가 있다는 관찰과 모순된다(Karsi 등, *Mar Biotechnol (NY)* 3: 241-245, 2001). 시스템은 크기 제한으로부터 해제된 것으로 보인다. 먼저, 촉매적으로 프라이밍된 트랜스포존/트랜스포사제 복합체를 형성하기 위해, 트랜스포존의 2개의 단부는 트랜스포사제 분자에 의해 물리적으로 매우 가깝게 함께 유지되어야 한다(Hudecek 등, *Crit Rev Biochem Mol Biol* 52: 355-380, 2017). 이 제한은 frt 측면을 HDAd 벡터로 통합함으로써 해결되었고 공동 발현된 Flpe 재조합효소에 의해 인식되어 트랜스포존의 원형화로 이어진다(Yant 등, *Nat Biotechnol* 20: 999-1005, 2002). 큰 작제물의 전위를 제한하는 두번째 메커니즘은 자동-통합, 즉, 트랜스포존 내부의 TA 디뉴클레오티드로의 통합이라 불리는 자살 전위 메커니즘이다(Wang 등, *PLoS Genet 10*: e1004103, 2014). HDAd-짧은-LCR 및 HDAd-긴-LCR 사이의 VCN에서 보이지 않는 차이는 생체내 선택과 관련될 수 있으며, 특정 수준의 $mgmt^{P140K}$ 발현에 따라, 즉, 임계값 VCN에 도달된 세포의 경우 HSPC 및 전구체에 대해 풍부화된다.

[0997] 강력한 $O^6$BG/BCNU 생체내 선택 시스템으로 인해, 말초 혈액 적혈구의 거의 100%는 γ-글로빈을 함유하였다. 이 생체내 선택 접근법은 골수에서 세포 조성물에 영향을 미치지 않는 반면, 백혈구감소증을 초래한다. 따라서 세포독성 약물 BCNU를 수반하지 않는 대안적인 접근법에 대한 노력이 집중되고 있다. 특히, 뮤린 지중해빈혈 모델에서의 연구에 의해 뒷받침되는 바와 같이(Wang 등, *J Clin Invest* 129: 598-615, 2019), 유전자-교정된 HSPC가 교정되지 않은 세포보다 증식 이점이 있을 것이기 때문에 약제학적 생체내 선택은 혈색소병증 환자에서 필요하지 않을 수 있다(Perumbeti 등, *Blood* 114: 1174-1185, 2009).

[0998] 1차 동물 및 2차 수용자에서 HDAd-짧은-LCR 및 HDAd-긴-LCR에 대한 비슷한 VCN을 고려하면, RBC 및 골수 적혈구 전구체에서 γ-글로빈 수준(HPLC 및 qRT-PCR에 의해 측정됨)은 긴 LCR을 함유하는 벡터에 대해 상당히 더 높았다. 흥미롭게도, 2 개 벡터 사이의 차이는 2차 수용자에서 더 명백하였다. 이는 형질도입된 장기간 재증식 HSPC로부터 유래된 RBC가 더 높은 γ-글로빈 수준을 갖는다는 것을 암시한다. 또한, HDAd-긴-LCR은 더 강한 적혈구 특이성을 나타내었다. 이러한 효과는 LCR의 염색질 개방 능력으로 인해(Li 등, *Blood* 100: 3077-3086, 2002) 전사 인자에 대한 더 나은 접근을 초래하는 HDAd-긴 LCR에서 추가적인 LCR 요소, 및/또는 γ-글로빈 유전자의 전사 증가를 초래하는 추가적인 전사 인자의 결합에 기여할 수 있다. LCR의 또 다른 특징은 주목할만하며, 즉, 자율적인 조절 단위로 작용할 수 있는 능력이며, 무작위 통합 후 이웃 유전자의 적은 전사활성화를 암시한다. 이 맥락에서 보다 완전한 LCR 버전을 사용하여 접근법에 대한 잠재적인 유전독성을 감소시킨다.

[0999] **실시예 3. 내인성 태아 글로빈의 CRISPR-촉발된 재활성화 및 SB100x 트랜스포사제-매개 γ-글로빈 유전자 추가의 조합을 사용한 생체내 HSC 유전자 요법은 마우스 모델에서 겸상적혈구병을 치유한다.**

[1000] 태아 글로빈의 유전적 지속성을 갖는 환자 및, 보다 최근에, 유전자 요법 환자에서, 겸상적혈구병(SCD)의 표현형 교정 정도는 태아 γ-글로빈의 발현 수준과 상관관계가 있다. HDAd5/35++ 벡터를 사용한 생체내 조혈 줄기세포/전구체(HSPC) 형질도입 후, SB100x 트랜스포사제-매개 γ-글로빈 유전자 추가는 성체 마우스 글로빈의 10-15% γ-글로빈을 달성하여, 중간형 지중해빈혈 마우스 모델에서 유의하지만 불완전한 표현형 교정을 초래하는 것으로 최근에 보고되었다. 또한 CRISPR/Cas9에 의해 γ-글로빈 프로모터 내에서 γ-글로빈 억제인자 결합 부위의 게놈 편집은 내인성 γ-글로빈의 효율적인 재활성화를 초래하는 것으로 제시되었다. 이 실시예는 이러한 2 가지 메커니즘을 조합하여 생체내 HSPC 형질도입 후 γ-글로빈의 치유적 수준을 수득한다.

[1001] 두 모듈을 함유하는 HDAd5/35++ 아데노바이러스 벡터(HDAd-combo)를 생성하고 "건강한" CD46/β-YAC 마우스 및 SCD 마우스 모델(CD46/Townes)에서 시험관내 및 생체내 HSPC 형질도입 후 테스트하였으며, 여기서 뮤린 α- 및

β-글로빈 유전자를 인간 α-글로빈 및 인간 겸상 $\beta^S$/태아 γ-글로빈 유전자로 대체하였다. 본 HDAd-combo는 표적 부위 절단 완료 후 Cas9 발현을 감소시키는 자기-활성화 메커니즘을 함유하였다. 이는 생체내에서 상당히 더 높은 절단 빈도를 초래하였으며, 더 나은 생존 CRISPR/Cas9-편집된 HSPC 때문일 가능성이 크다. 중요하게는, γ-글로빈 추가 또는 CRISPR/Cas9 재활성화 단위 단독을 함유하는 HDAd 벡터와 비교하여, HDAd-combo로 형질도입 후 상당히 더 높은 γ-글로빈이 RBC에서 발견되었다. combo 벡터로 CD46/Townes 마우스의 생체내 HSC 형질도입 후 13 주차에, 적혈구에서 γ-글로빈 수준의 수준은 성인 인간 α- 및 $\beta^S$-쇄 수준의 30%였다. 이는 SCD의 완전한 표현형 교정을 초래하였다.

[1002] **도입:**

[1003] ***SCD 유전자 요법:*** 겸상적혈구병 및 β-지중해빈혈은 전 세계적으로 가장 흔한 단일유전자성 장애이며, 매년 317,000 명의 영향을 받은 신생아가 태어난다. SCD는 b-글로빈 유전자의 첫번째 엑손($\beta^s$ 대립유전자) 상의 단일 돌연변이에 의해 유발되어, 결함이 있는 헤모글로빈 사량체의 형성을 초래하여, 낮은 산소 농도에서 중합되어, 적혈구 파괴로 이어진다. SCD는 실질적 이환율, 낮은 삶의 질, 및 기대 수명 단축과 연관된다. SCD의 임상 과정은 태아 γ-글로빈 유전자가 HPFH 특성이 있는 환자에서 볼 수 있는 바와 같이 고도로 발현될 때 개선된다(Conley 등, *Blood* 21: 261-281, 1963; Stamatoyannopoulos 등, *Blood* 46: 683-692, 1975). SCD에서, γ-글로빈은 Hb 사량체에 혼입을 위해 겸상 β-글로빈과 경쟁하고 겸상 헤모글로빈(HbS) 중합을 억제함으로써 강력한 겸상 방지 기능을 발휘한다. HbF 수준을 증가시키는 약리학적 치료는 모든 환자에서 동일하게 효과적이지 않다. β-혈색소병증에 대한 유전자 요법의 개발은 일치하는 공여자의 제한된 이용가능성 및 가장 어린 환자에게 HSPC 이식 적용의 좁은 창에 의해 정당화되었다. 현재 SCD 유전자 요법 접근법은 HSPC의 수집, 이들의 시험관내 배양, 온전한 β-글로빈, 겸상 방지 β-글로빈 또는, 태아 γ-글로빈 발현 카세트를 보유하는 렌티바이러스 벡터를 사용한 형질도입, 및 골수-조건화된 환자로 재이식을 수반한다. γ-글로빈 유전자 추가 렌티바이러스 벡터를 사용한 I 상 유전자 요법 시험이 유망하지만, 모든 SCD 증상의 장기간 치유는 아직 달성되지 않았다(Demirci 등, *Hum Mol Genet.*, 2020. doi: 10.1093/hmg/ddaa088). 질환의 치유를 위해, RBC에서 γ-글로빈 수준은 성인 α-글로빈의 적어도 20%이어야 하고, 최적으로, $\beta^S$ 수준은 감소되어야 한다. 이는 전장 글로빈 LCR 또는 다중-양식 게놈 편집 카세트의 사용을 방지하는 삽입 크기 제한으로 인해, 렌티-바이러스 벡터로 달성하기 어렵다(Uchida 등, *Nat Commun* 10: 4479, 2019).

[1004] ***생체내 HSPC 유전자 요법 - γ-글로빈 유전자 추가:*** 생체외 HSPC 유전자 요법의 주요 위험은 이식 관련된 이환율이다(Anurathapan 등, *Biol Blood Marrow Transplant* 20: 2066-2071, 2014; Lucarelli 등, *Blood Rev* 16: 81-85, 2002; Storb 등, *Hematology Am Soc Hematol Educ Program*: 372-397, 2003). 또한, 렌티바이러스 벡터의 사용은 이식유전자 발현이 침묵되거나 또는 염색체 원종양형성 유전자가 활성화되는 위험이 있다. 중요하게는, 접근법은 복잡하고, 값이 비싸며, 따라서 SCD가 만연한 제한된 자원을 가진 나라에서 수행하기 어렵다. 간단한 생체내 HSPC 유전자 요법 접근법이 개발되었다. 이는 HSPC를 골수에서 말초 혈류로 동원하기 위한 GCSF/AMD3100의 피하 주사 및 통합 헬퍼-의존적 아데노바이러스 벡터 시스템인 HDAd5/35++ 벡터의 정맥내 주사를 수반한다. 이들 벡터는 30+kb의 삽입 용량을 가지며 원시 HSPC 상에서 발현되는 수용체인 CD46을 표적화한다(Richter 등, *Blood* 128: 2206-2217, 2016). 정맥내 HDAd5/35++ 주사와 연관된 선천적 독성은 마우스 및 비-인간 영장류에서 글루코코르티코이드, IL6- 및 IL1β-수용체 길항제 전처리에 의해 제어될 수 있다(Li 등, *23rd Annual ASGCT meeting*. 2020; abstract #546) 무작위 이식유전자 통합은 활성-향상된 잠자는 미녀 트랜스포사제(SB100x)에 의해 매개된다(Boehme 등, *Mol Ther Nucleic Acids* 5: e337, 2016). 이 시스템에서, 이식유전자 카세트는 SB100x 트랜스포사제에 의해 인식되는 도립 반복부(IR), 및 Flp 재조합효소의 존재 하에 이식유전자 카세트의 원형화를 허용하는 frt 부위에 의해 플랭킹된다. 제2 벡터인 HDAd-SB는 Flp 재조합효소 및 SB100x를 트랜스로 공급하여 GFP 카세트의 게놈 DNA의 TA 디뉴클레오티드로의 통합을 매개한다(Mates 등, *Nat Genet* 41: 753-761, 2009). HDAd5/35++ 벡터를 사용한 이전 연구에서, 4.3 kb HS1-HS4 미니-LCR(β-글로빈 유전자좌 제어 영역)을 0.66 kb β-글로빈 프로모터와 조합하여 사용하여 생체내 HSPC 형질도입 후 인간 γ-글로빈 발현을 구동하였다(Wang 등, *J Clin Invest* 129: 598-615, 2019; Li 등, *Mol Ther Methods Clin Dev* 9: 142-152, 2018). $Hbb^{th3}$/$CD46^{+/+}$ 지중해빈혈 마우스에서, 안정된(8 개월 초과) γ-글로빈 마킹은 말초 혈액 적혈구의 거의 100% 및 거의 완전한 표현형 교정에서 달성되었다(Wang 등, *J Clin Invest* 129: 598-615, 2019). 그러나, γ-글로빈 발현 수준은 세포 당 2 개 카피의 평균 통합 벡터 카피 수(VCN)를 갖는 성체 마우스 α-글로빈 수준의 단지 10-15%였으며, 따라서 특히 과제인 SCD에 대한 접근법의 임상 번역을 제공한다.

[1005] ***생체내 HSPC 유전자 요법 - 내인성 γ-글로빈의 재활성화:*** 양성 유전적 상태인 태아 헤모글로빈(HPFH)의 유전적 지속성에서, 돌연변이는 γ에서 β로의 글로빈 전환을 약화시켜, 일생에 걸쳐 높은 태아 글로빈(HbF) 수준을 유발하여 이러한 장애의 임상적 징후를 완화시킨다(Forget, *Ann N Y Acad Sci* 850: 38-44, 1998). β-글로빈 유전자좌 내에서 큰 결실을 생성하거나(Sankaran, *Hematology Am Soc Hematol Educ Program* 2011: 459-465, 2011), 또는 *HBG* 프로모터에 돌연변이를 도입함으로써 HPFH 돌연변이를 다시 일으키려 시도하는 초기 연구는 적혈구 세포에서 HbF의 수준을 증가시킬 수 있다(Wienert 등, *Nat Commun* 6: 7085, 2015; Traxler 등, *Nat Med* 22: 987-990, 2016; Lin 등, *Blood* 130: 284, 2017). 태아 글로빈 억제인자로서 BCL11A의 발견으로, 이러한 시도는 HBG 프로모터 내에서 BCL11A 결합 부위의 표적화된 파괴(Masuda 등, *Science* 351: 285-289, 2016) 또는 CRISPR/Cas9 또는, 최근에, 염기 편집기(Zeng 등, *Nat Med* 26: 535-541, 2020)에 의해 BCL11A 발현을 감소시키는 적혈구 bcl11a 인핸서의 파괴(Wu 등, *Nat Med* 25: 776-783, 2019)를 수반하는 데 더욱 집중하게 되었다. HBG1/HBG2 프로모터 표적화된 CRISPR/Cas9를 이용하여 인간 β-글로빈 유전자좌-유전자이식(β-YAC) 마우스에서 γ-글로빈을 재활성화하였다(Li 등, *Blood* 131: 2915-2928, 2018). 생체내 HSPC 형질도입 후, HSPC의 2차 이식 후 유지된 성체 마우스의 적혈구에서 인간 β-에서 γ-글로빈 발현으로의 명백한 전환을 초래하는 효율적인 표적 부위 파괴를 입증하였다. 장기간 추적 연구에서, 혈액학적 이상은 검출되지 않았으며, 이는 HBG 프로모터 편집이 조혈에 부정적으로 영향을 미치지 않았음을 나타낸다.

[1006] HDAd5/35++ 벡터로부터 CRISPR/Cas9의 발현은 줄기 세포 기능 및 형질도입된 HSPC, 구체적으로 인간 HSPC의 생존을 손상시킬 수 있는 있음이 이전에 보고되었다(Li 등, *Mol Ther Methods Clin Dev* 9: 390-401, 2018). 따라서, CRISPR/Cas9 발현을 단축하기 위한 접근법이 개발되었다(Li 등, *Mol Ther Methods Clin Dev* 9: 390-401, 2018; Li 등, *Mol Ther* 27: 2195-2212, 2019).

[1007] 여기서, 목표는 β-YAC 마우스 뿐만 아니라 Tim Townes에 의해 개발된 겸상적혈구병의 마우스 모델(hα/hα::$\beta^S$/$\beta^S$)에서 SB100x-매개 γ-글로빈 유전자 추가 및 γ-글로빈의 재활성화를 조합함으로써 생체내 HSPC 형질도입 후 γ-글로빈의 치유적 수준을 달성하는 것이었다(Wu 등, *Blood* 108: 1183-1188, 2006). 이 모델에서, 뮤린 α-글로빈 유전자를 인간 α-글로빈으로 대체하였고 뮤린 성체 β-글로빈 유전자를 함께 연결된 인간 겸상 $\beta^S$ 및 태아 γ-글로빈 유전자로 대체하였다. 이 모델은 겸상적혈구병의 주요 표현형 특징을 나타낸다.

[1008] **재료 및 방법**

[1009] **시약:** G-CSF(Neupogen™)(Amgen 캘리포니아주 사우전드 오크스 소재) 및 AMD3100(Sigma-Aldrich, 미주리주 세인트루이스 소재)을 사용하였다. $O^6$-BG 및 BCNU는 Sigma-Aldrich(미주리주 세인트루이스 소재)로부터 수득하였다.

[1010] **HDAd 벡터:** HDAd-CRISPR("cut"), HDAd-SB-추가("add") 및 HDAd-SB는 이전에 기재되었다(Li 등, *Blood* 131(26):2915-2928, 2018; Wang 등, *J Clin Invest* 129: 598-615, 2019). pHCA-combo의 클로닝은 3 단계를 수반하였다. 단계 1) *HBG1/2* 프로모터 영역에서 BCL11A 결합 부위를 표적화하는 sgHBG#2(서열번호: 258)를 합성하고, 어닐링하고 pSPgRNA(Addgene, 매사추세츠주 케임브리지 소재)의 *BbsI* 부위에 삽입하여, pSP-sgHBG#2를 생성하였다. pSP-sgHBG#2에서 0.4 kb U6-sgHBG#2 단편을 증폭시키고 pBST-sgAAVS1-miR의 *BamHI* 부위로 클로닝하여(Li 등, Mol Ther 27: 2195-2212, 2019), pBST-sgHBG#2-miR을 수득하였다. 단계 2) 1.5 kb PGK-mgmt-bGHpolyA 단편을 gBlock(IDT, 뉴저지주 뉴어크 소재)으로 합성하고 *ClaI*-소화된 pBS-LCR-글로빈-mgmt와 결찰시켜(Li 등, *Mol Ther* 27: 2195-2212, 2019), pBS-LCR-글로빈-PGK-mgmt를 수득하였다. 다음으로, pBS-Frt-IR 영역을 함유하는 4.8 kb 서열을 pBS-FRT-IR-Ef1α-mgmt로부터 증폭시키고(Li 등, *Cancer Res* 80: 549-560, 2020) *EcoRV-KpnI* 소화된 pBS-LCR-글로빈-PGK-mgmt와 결찰시켜, pBS-Frt-IR-LCR-글로빈-PGK-mgmt를 야기하였다. 이 단계에서 나중 주입 클로닝을 위한 15 bp 상동성 아암 (HA)을 함유하는 프라이머(Takara, 캘리포니아주 마운틴뷰 소재)를 사용하였다. 2 개의 Frt-IR 구성요소에 플랭킹하는 2 개의 15 bp HA는 PacI 소화시 노출되어 하기 기재된 변형된 pHCA 작제물을 사용한 재조합을 용이하게 할 수 있다. 단계 3) pHCAS1S-MCS의 5.3 kb *XbaI* 단편(Li 등, *Mol Ther* 27: 2195-2212, 2019)을 *XbaI* 제한에 의해 검출하고 재결찰시켜, pHCAS1S1-MCS를 생성하였다. U6 프로모터에서 시작하여 SV40 polyA 신호 서열까지 7.6 kb CRISPR 카세트를 pBST-sgHBG#2-miR로부터 증폭시키고 pHCAS1S1-MCS의 *NheI* 부위에 클로닝하여, pHCAS1S1-MCS-sgHBG#2를 형성하였다. 마지막으로, pBS-Frt-IR-LCR-글로빈-PGK-mgmt에서 12.0 kb HA-플랭킹된 글로빈/mgmt 카세트를 PacI 처리에 의해 방출시키고 PacI-소화된 pHCAS1S1-MCS-sgHBG#2와 재조합하여, pHCA-combo를 초래하였다. 최종 작제물을 여러 제한 효소(*HindIII*, *EcoRI* 및 *PmeI*)에 의해 스크리닝하고 이식유전자를 함유하는 전체 영역을 서열분

석함으로써 확인하였다.

[1011] HDAd5/35++ 벡터의 생산을 위해, 상응하는 플라스미드를 *PmeI*로 선형화하고 Ad5 섬유 꼬리, Ad35 섬유 샤프트, 및 친화도-향상된 Ad35++ 섬유 놉으로 구성된 키메라 섬유를 함유하는 Ad5/35++ 헬퍼 벡터인 AdNG163-5/35++(Richter 등, *Blood* 128: 2206-2217, 2016)를 사용하여 116 세포를 구제하였다(Palmer & Ng, *Mol Ther* 8: 846-852, 2003). HD-Ad5/35++ 벡터를 다른 곳에 상세하게 기재된 바와 같이 116 세포에서 증폭시켰다(Palmer & Ng, *Mol Ther* 8: 846-852, 2003). 헬퍼 바이러스 오염 수준은 0.05% 미만인 것으로 밝혀졌다. 역가는 2-5x$10^{12}$ vp/ml였다.

[1012] 본 실시예의 벡터는 도 101에 예시되어 있고, (i) 트랜스포존에 존재하는 γ-글로빈 이식유전자를 암호화하는 핵산("add") 및 (ii) 트랜스포존에 존재하지 않는, 내인성 γ-글로빈의 발현을 증가시키기 위해 HBG1/2을 표적화하는 CRISPR/Cas9 시스템을 암호화하는 핵산("CRISPR")을 둘 다 포함하는 HDAd 조합 아데노바이러스 벡터를 포함한다(둘은 함께 "조합"을 형성한다). 이중 벡터와 관련한 추가 개시내용에 대해, 또한 도 96, 102, 97a-97d, 98a-98n, 99a-99u를 참조한다).

[1013] 구체적으로, 도 96은 CRISPR 시스템이 2 개의 상이한 부위(HBG 프로모터 및 적혈구 bcl11a 인핸서)를 표적화하여, 증가된 감마 재활성화로 이어지는 HDAd-TI-combo 벡터의 개략도를 나타낸다. 도 102는 HDAd-combo에서 Flpe 재조합효소와 frt 부위의 상호작용이 어떻게 트랜스포존의 원형화로 이어져, CRISPR 카세트를 함유하는 벡터의 선형 단편을 남기는지를 나타낸다. SB100x/Flpe 시스템을 사용한 이전 연구는 원형화된 트랜스포존이 SB100x에 의해 숙주 게놈 내로 통합되는 동안 이러한 벡터 부분이 빠르게 손실된다는 것을 입증하였다(Yant 등, *Nat Biotechnol.*, 20: 999-1005, 2002). 도 97a는 HDAd-SB 및 HDAd-combo의 공동 감염시, Flpe이 어떻게 발현되고 IR-플랭킹된 트랜스포존을 방출한 다음, SB100x 트랜스포사제에 의해 게놈 내로 통합될 것인지를 나타낸다. 동시에, HBG1 및 bcl11a-E CRISPR은 발현되고 DNA indel을 생성하여 γ-글로빈의 재활성화로 이어질 것이다. 트랜스포존의 Flp-매개 방출 시, CRISPR 카세트는 분해되어, 세포독성을 피할 것이다. CRISPR 시스템은 2 개의 상이한 부위(HBG 프로모터 및 적혈구 bcl11a 인핸서)를 표적화하여, γ 재활성화 증가로 이어진다. 표적화 전략(도 97b), 적혈구 특이적 BCL11A 인핸서(도 97c), 및 HBG 프로모터서의 BCL11A 결합 부위(도 97d)가 또한 제시된다.

[1014] 이중 CRISPR 벡터 및 γ-글로빈 재활성화는 도 98a-98n에 제시된다. HDAd-Bcl11ae-CRISPR, HDad-HBG-CRISPR, HDAd-Dual-CRISPR, HDAd-스크램블(도 98a) 및 이중 gRNA 벡터에 대한 HD-Ad5/35++ CRISPR 벡터(도 98b)에 대한 벡터 설계가 제시된다. 인간 적혈구 전구체 세포주(HUDEP-2)의 HD-Ad5/35++ CRISPR 형질도입은 도 98c에서 분화 전 및 후로 제시된다. HD-AD5/35++ "Dual" gRNA 벡터는 미처리(UNTR), BCL11A, 또는 HBG 벡터와 비교하여 세포 생존력(도 98d) 또는 증식(도 98e)에 부정적으로 영향을 미치지 않는다. 이중 벡터는 표적 유전자좌(도 98f) Bcl11a 인핸서 및 (도 98g) HBG 프로모터에 대한 단일 gRNA 벡터로 관찰된 것들과 유사한 유사한 편집 수준을 달성한다. 또한, HD-AD5/35++ "Dual" gRNA 벡터는 단일 gRNA 벡터로 관찰된 것들과 유사한 표적 유전자좌의 편집 수준을 달성한다(도 98h). HbF+ 세포의 상당히 더 높은 백분율은 단일 gRNA 벡터와 비교하여 HD-Ad5/35 "Dual" gRNA 벡터로 형질도입된 HUDEP-2 세포에서 유세포 분석에 의해 관찰되었다(도 98i). HPLC에 의해 측정된 전반적인 감마 글로빈 발현은 이중 표적화된 샘플에서 상당히 더 높았다(도 98j). 단일 녹아웃 클론보다 이중 녹아웃 클론에서 상당히 더 큰 태아 글로빈 발현이 관찰되었으며 2 개 돌연변이의 가능한 상승적 효과를 암시하며, 더 높은 감마 발현/세포로 이어진다(도 98k). 도 98l은 말초 혈액 동원된 CD34+ 세포가 HDAd5/35++ CRISPR 벡터로 형질도입되었음을 나타낸다. CRISPR/Cas9 세포독성을 최소화하기 위해, 이어서 세포를 항-Cas9 펩티드를 발현하는 HDAd5/35++ 벡터로 후속적으로 형질도입하였다. 세포를 준-치명적으로 조사된 NSG 마우스 내에 이식하고 분석하였다. 이식 후 10 주차에, HD-Ad5/35 "Dual" gRNA 벡터로 형질도입된 세포는 단일 gRNA 벡터로 형질도입된 세포와 유사한 생착을 나타내었다. 계통 조성은 모든 그룹에서 유사하였다(도 98m). 이중 gRNA 벡터에 의해 형질도입되고 편집된 CD34+ 세포는 NSG 마우스에 효율적으로 주입되었다(도 98n). 또한, 적혈구 분화 후 주입된 이중 표적화된 세포는 상대적으로 더 낮은 편집 수준에도 불구하고, 단일 표적화된 세포와 비교하여, 대조군에 대해 더 높은 수준의 감마 글로빈을 발현하였다(도 9n).

[1015] 이중 편집된 정상 및 thal CD34+ 세포의 생체외 형질도입에 대한 실험 설계는 도 99a에 제시된다. 정상 CD34+ 세포에 대해 15 일째에 콜로니에서 HBF 발현(도 99b), MFI(도 99c), 및 HBF 발현을 설명하는 유세포 분석 데이터(도 99d)가 제시된다. 정상 CD34+ 세포에 대해 적혈구 분화(ED) 후 HBF 발현(도 99e) 및 MFI(도 99f)가 제시된다. HBG 부위에 대한 TE71(도 99g) 및 BCL11A 부위에 대한 TE71(도 99h)은 정상 CD34+ 세포에서 형질도입(txd) 후 48 시간에 제시된다. EC에서 HBF 발현 및 적혈구 분화를 설명하는 유세포 분석 데이터는 도 99i에서

찾을 수 있다. 도 99j-99u는 Thal CD34+ 세포에서의 결과를 제시한다. 0 일차에 세포의 면역표현형, 형질도입되지 않은 세포 및 CRISPR-Dual로 형질도입된 세포(도 99j) 및 11 일에 걸쳐 형질도입되지 않은 세포 및 CRISPR-Dual로 형질도입된 세포를 비교하는 성장 곡선(도 99k). HBF 발현(도 991) 및 MFI(도 99m)는 15 일째에 콜로니에서 제시된다. EC에서의 HBF 발현(도 99p), MFI(도 99q), 및 HBF 발현 및 P04 및 P18을 설명하는 유세포 분석 데이터(도 99r)가 또한 제시된다. p04(도 99s) 및 p18(도 99t)에서 HBG 부위 적혈구 분화에 대한 TE71이 제시되는 반면 도 99u는 형질도입 48 시간 후 BCL11A 부위에 대한 TE71을 제시한다.

[1016] **HUDEP-2 세포/적혈구 분화:** HUDEP-2 세포(Kurita 등, *PLoS One* 8: e59890, 2013)를 100 ng/ml SCF, 3 IU/ml EPO, $10^{-6}$ M 덱사메타손 및 1 μg/ml 독시사이클린(DOX)이 보충된 StemSpan SFEM 배지(STEMCELL Technologies)에서 배양하였다. 적혈구 분화를 5% 인간 AB 혈청, 100 ng/ml SCF, 3 IU/ml EPO, 10 μg/ml 인슐린, 330 μg/ml 트랜스페린, 2 U/ml 헤파린 및 1 μg/ml DOX를 함유하는 IMDM 배지에서 6 일 동안 유도하였다.

[1017] **콜로니-형성 단위(CFU) 검정:** 계통 마이너스($Lin^-$) 세포를 제조업체의 지침에 따라 마우스 계통 세포 고갈 키트(Miltenyi Biotec, 캘리포니아주 샌디에이고 소재)를 사용하여 골수 MNC에서 계통-수임 세포의 고갈에 의해 단리하였다. CFU 검정을 제조업체의 프로토콜에 따라 마우스 완전 배지와 함께 ColonyGEL(Reachbio, 워싱턴주 시애틀 소재)을 사용하여 수행하였다. 콜로니를 플레이팅 10 일 후 점수매겼다.

[1018] **T7EI 불일치 뉴클레아제 검정:** 게놈 DNA를 제공된 프로토콜에 따라 PureLink Genomic DNA 미니 키트(Life Technologies, 캘리포니아주 칼즈배드 소재)를 사용하여 단리하였다(Miller 등, *Nat Biotechnol* 25: 778-785, 2007). HBG1/2 프로모터의 표적화된 부위를 포함하는 게놈 분절을 다음 PCR 프라이머에 의해 증폭시켰다: HBG1/2 정방향(서열번호: 270), 역방향(서열번호: 271). PCR 생성물을 혼성화하고 37℃에서 20 분 동안 2.5 단위의 T7EI(NEB)로 처리하였다. 소화된 PCR 생성물을 10% TBE PAGE(Bio-Rad)로 분해하고 에티듐 브로마이드로 염색하였다. 100 bp DNA 래더(New England Biolabs)를 사용하였다. 밴드 강도를 ImageJ 소프트웨어를 사용하여 분석하였다. 절단율 % = (1-sqrt(모체 밴드/(모체 밴드 + 절단된 밴드)) x 100%.

[1019] **유세포 분석:** 세포를 1 % FCS가 보충된 PBS에 $1x10^6$ 개 세포/100 μL로 재현탁하고 얼음 위에서 10 분 동안 FcR 차단 시약(Miltenyi Biotech, 캘리포니아주 오번 소재)과 함께 배양하였다. 다음으로 염색 항체 용액을 $10^6$ 개 세포 당 100 μL로 첨가하고 암실에서 30 분 동안 얼음 위에서 배양하였다. 배양 후, 세포를 FACS 완충액(PBS, 1% FBS)에서 1 회 세척하였다. 2차 염색을 위해 염색 단계를 2차 염색 용액으로 반복하였다. 세척 후, 세포를 FACS 완충액에 재현탁하고 LSRII 유세포 분석기(BD Biosciences, 캘리포니아주 산호세 소재)를 사용하여 분석하였다. 파편을 전방 산란 영역 및 측방 산란 영역 게이트를 사용하여 제외하였다. 그런 다음 단일 세포를 전방 산란 높이 및 전방 산란 폭 게이트를 사용하여 게이팅하였다. 그런 다음 유세포 분석 데이터를 FlowJo(버전 10.0.8, FlowJo, LLC)를 사용하여 분석하였다. LSK 세포의 흐름 분석을 위해, 세포를 비오틴-접합된 계통 검출 칵테일(Miltenyi Biotec, 캘리포니아주 샌디에이고 소재)(cat #: 130-092-613) 및 c-Kit(cat #:12-1171-83) 및 Sca-1(cat #: 25-5981-82)에 대한 항체 뿐만 아니라 APC-접합된 스트렙타비딘으로 염색하였다. eBioscience(캘리포니아주 샌디에이고 소재)의 다른 항체는 항-마우스 LY-6A/E(Sca-1)-PE-Cyanine7(클론 D7), 항-마우스 CD117(c-Kit)-PE(클론 2B8), 항-마우스 CD3-APC(클론 17A2)(cat #:17-0032-82), 항-마우스 CD19-PE-Cyanine7(클론 eBio1D3)(cat #: 25-0193-82), 및 항-마우스 Ly-66(Gr-1)-PE(클론 RB6-8C5)(cat #: 12-5931-82을 포함하였다. 항-마우스 Ter-119-APC(클론: Ter-119)(cat #: 116211)는 Biolegend(캘리포니아주 샌디에이고 소재)로부터 수득하였다.

[1020] **인간 γ-글로빈 발현을 검출하는 세포내 유세포 분석:** FIX & PERM™ 세포 투과화 키트(Thermo Fisher Scientific)를 사용하고 제조업체의 프로토콜에 따랐다. 간단히 말해서, $1x10^6$ 개의 세포를 100 μl FACS 완충액(1% FCS가 보충된 PBS)에 재현탁하고, 100 μl의 시약 A(고정 배지)를 첨가하고 실온에서 2-3 분 동안 배양한 다음, 1 ml 미리 냉각된 절대 메탄올을 첨가하고, 혼합하고 10 분 동안 암실에서 얼음 위에서 배양하였다. 그런 다음 샘플을 FACS 완충액으로 세척하고 100 μl 시약 B(투과 배지) 및 1 μg 헤모글로빈 γ 항체(Santa Cruz Biotechnology, cat# sc-21756 PE)에 재현탁하고, 실온에서 30 분 동안 배양하였다. 세척 후, 세포를 FACS 완충액에 재현탁하고 분석하였다.

[1021] **글로빈 HPLC:** 개별 글로빈 쇄 수준을 SPD-10AV 다이오드 어레이 검출기 및 LC-10AT 바이너리 펌프가 장착된 Shimadzu Prominence 기기에서 정량화하였다(Shimadzu, 일본 쿄토 소재). 폴리펩티드에 대한 Vydac 214TP™ C4 역상 칼럼(214TP54 Column, C4, 300 Å, 5 μm, 4.6 mm i.d. x 250 mm)(Hichrom, 영국 소재)을 사용하였다.

물/아세토니트릴 중 0.1% 트리플루오로아세트산의 40%-60% 구배 혼합물을 1 mL/분의 속도로 적용하였다.

[1022] **벡터 카피 수의 측정:** 세포 당 아데노바이러스 게놈 카피의 절대 정량화를 위해, 게놈 DNA를 제공된 프로토콜에 따라 PureLink Genomic DNA 미니 키트(Life Technologies)를 사용하여 세포로부터 단리하고, power SYBR™ green PCR 마스터 믹스(Thermo Fisher Scientific)를 사용하여 수행된 qPCR에 대한 주형으로 사용하였다. 다음 프라이머 쌍을 사용하였다: 인간 γ-글로빈 정방향(서열번호: 195), 및 역방향(서열번호: 196); mgmt 정방향(서열번호: 220), 및 역방향(서열번호: 221).

[1023] **실시간 역전사 PCR:** 총 RNA를 5x10^6 개의 분화된 HUDEP-2 세포 또는 100 μl 혈액으로부터 제조업체의 페놀-클로로포름 추출 방법에 따라 TRIzol™ 시약(Thermo Fisher Scientific)을 사용하여 추출하였다. Quantitect 역전사 키트(Qiagen) 및 power SYBR™ green PCR 마스터 믹스(Thermo Fisher Scientific)를 사용하였다. 실시간 정량적 PCR을 StepOnePlus 실시간 PCR 시스템(AB Applied Biosystems)에서 수행하였다. 다음 프라이머 쌍을 사용하였다: 마우스 RPL10(하우스키핑) 정방향(서열번호: 189), 및 역방향(서열번호: 190); 인간 γ-글로빈 정방향(서열번호: 191), 및 역방향(서열번호: 192); 인간 β-글로빈 정방향(서열번호: 216), 및 역방향(서열번호: 217); 마우스 β-주요 글로빈 정방향(서열번호: 193), 및 역방향(서열번호: 194), 마우스 α 글로빈 정방향(서열번호: 212), 및 역방향(서열번호: 213).

[1024] **Cas9 웨스턴 블롯:** $3x10^6$ 개의 HUDEP-2 세포를 형질도입 후 다양한 시점에서 수확하고, PBS로 2 회 세척하고, 5% β-머캅토에탄올이 함유된 Laemmli 완충액으로 용해시켰다. 샘플을 95℃에서 5 분 동안 끓이고 13,000 g에서 10 분 동안 원심분리에 의해 정화시켰다. 10 μL의 용해물을 4-15% 프리캐스트 단백질 겔(Bio-Rad)을 사용하여 SDS-PAGE에 의해 분리하였다. 블롯의 Cas9 단백질을 항-Cas9-HRP(클론 7A9-3A3)에 의해 프로빙하였다(Cell Signaling Technology, 매사추세츠주 댄버스 소재). Pierce™ ECL Plus 웨스턴 블롯팅 기질(Thermo Fisher Scientific)로 처리 후 X-선 필름 상에서 화학발광 검출을 수행하였다. Cas9 검출 후, 블롯을 스트립하고 내부 대조군을 위해 Sigma-Aldrich의 항-β-액틴 항체(클론 AC-74)에 의해 다시 프로빙하였다.

[1025] **동물:** 동물을 수반하는 모든 실험은 워싱턴 대학에 의해 제시된 기관 지침에 따라 수행하였다. 워싱턴 대학은 실험실 동물 관리 국제 평가 및 인증 협회(AALAC) 공인 연구 기관이며 이 대학에서 수행된 모든 살아있는 동물 작업은 실험실 동물 복지국(OLAW) 공중 보건 보증(PHS) 정책, USDA 동물 복지법 및 규정, 실험 동물 관리 및 사용에 대한 지침 및 기관 동물 관리 및 사용 위원회(IACUC) 정책을 따른다. 연구는 워싱턴 대학 IACUC(프로토콜 번호 3108-01)에 의해 승인되었다. 인간 CD46 게놈 유전자좌를 함유하고 인간과 유사한 수준 및 패턴에서 CD46 발현을 제공하는 C57Bl/6 기반 유전자이식 마우스($hCD46^{+/+}$ 마우스)는 초기에 기재되었다(Kemper 등, *Clin Exp Immunol* 124: 180-189, 2001). 야생형 248 kb β-글로빈 유전자좌 효모 인공 염색체(β-YAC)를 보유하는 유전자이식 마우스를 사용하였다(Peterson 등, *Ann N Y Acad Sci* 850: 28-37, 1998). β-YAC 마우스를 인간 $CD46^{+/+}$ 마우스와 교배시켜 생체내 HSPC 형질도입 연구를 위한 $\beta\text{-YAC}^{+/-}/CD46^{+/+}$ 마우스를 수득하였다. 다음 프라이머를 마우스의 유전형 분석에 사용하였다: CD46 정방향(서열번호: 233), 및 역방향(서열번호: 234); β-YAC(γ-글로빈 프로모터) 정방향(서열번호: 242), 및 역방향(서열번호: 243).

[1026] *겸상적혈구병 마우스 모델:* Townes 수컷 마우스($Hbb^{tm2(HBG1,HBB*)Tow}$ 또는 $h\alpha/h\alpha::\beta^S/\beta^S$)를 Jackson Laboratory(JAX 스톡 #013071)로부터 수득하고 인간 CD46 유전자이식 암컷 마우스와 교배시켰다. 도 109a에 제시된 바와 같이, 3 차례의 교배 후, CD46, HbS 및 HBA에 대해 동형접합성인 마우스를 수득하고 실험에 사용하였다. 다음 프라이머를 유전형 분석에 사용하였다: HBB 프라이머(서열번호: 246, 251, 및 70), 및 HBA 프라이머(서열번호: 272-274); 및 상기 제시된 바와 같은 CD46 프라이머(서열번호: 233 및 234). PCR 결과는 공급처에 의해 제공된 프로토콜에 따라 해석하였다.

[1027] *HSPC 동원 및 생체내 형질도입:* HSPC를 인간 재조합 G-CSF의 s.c. 주사(5 μg/마우스/일, 4 일) 이어서 5 일째에 AMD3100의 s.c. 주사(5 mg/kg)에 의해 마우스에 동원하였다. 또한, 동물은 바이러스 주사 16 시간 및 2 시간 전에 덱사메타손(10 mg/kg)을 i.p.로 받았다. AMD3100 후 30 및 60 분에, 동물에게 주사 당 $4x10^{10}$ 개 바이러스 입자(vp)의 용량으로 안와하정맥총을 통해 바이러스 벡터를 정맥내로 주사하였다.

[1028] *생체내 선택:* 선택은 형질도입 후 1 주차(Townes 모델) 또는 4 주차(β-YAC 모델)에 시작하였다. 마우스에게 $O^6$-BG(15 mg/kg, IP)를 2 회, 30 분 간격으로 주사하였다. $O^6$-BG의 두번째 주사 1 시간 후에, 마우스에게 5

mg/kg 카르무스틴(BCNU)을 주사하였다(IP). 1 차례의 선택 후 2 및 4 주차에, 각각 7.5 및 10 mg/kg의 BCNU 용량으로 2 차례 이상 수행하였다.

[1029] *면역억제:* 마이코페놀레이트 모페틸(CellCept Intravenous)은 Genentech(오리건주 힐즈버러 소재)에서 수득하였다. 라파마이신(Rapamune/Sirolimus) 및 메틸프레드니솔론은 Pfizer(뉴욕주 뉴욕 소재)에서 수득하였다. 마이코페놀레이트 모페틸(20 mg/kg/일), 라파마이신(0.2 mg/kg/일), 메틸프레드니솔론(20 mg/kg/일)의 일일 복강내 주사를 수행하였다.

[1030] *2차 골수 이식:* 수용자는 Jackson Laboratory에서 수득한 6 - 8 주령의 암컷 C57BL/6 마우스였다. 이식 당일에, 수용자 마우스를 1000 Rad로 조사하였다. 생체내 형질도입된 CD46tg 마우스로부터 골수 세포를 단리된 무균상태로 단리하고 계통-고갈된 세포를 상기 기재된 바와 같이 MACS를 사용하여 단리하였다. 조사 6 시간 후 세포를 마우스 당 $1x10^6$ 개 세포로 정맥내로 주사하였다. 2차 수용자는 종료점 분석을 위해 이식 후 16 주 동안 유지하였다. 모든 2차 수용자는 4 주차에 시작하여 면역억제를 받았다.

[1031] **조직 분석:** 2.5 μm 두께의 비장 및 간 조직 절편을 적어도 24 시간 동안 4% 포름알데하이드에 고정시키고, 탈수하고 파라핀에 포매하였다. 헤마톡실린-에오신 염색은 골수외 혈구형성의 조직학적 평가를 위해 사용하였다. 헤모시데린을 Perl의 프러시안 블루 염색에 의해 조직 절편에서 검출하였다. 간단히 말해서, 조직 절편을 증류수 중 동일한 부피(2%)의 페로시안화칼륨 및 염산의 혼합물로 처리한 다음 뉴트럴 레드로 대조염색하였다. 비장 크기를 비장 중량(mg)/체중(g)의 비로 평가하였다.

[1032] **혈액 분석:** 혈액 샘플을 EDTA-코팅된 튜브에 수집하고 HemaVet 950FS(Drew Scientific, 코네티컷주 워터베리 소재)에서 분석을 수행하였다. 말초 혈액 도말을 각각 5 및 15 분 동안 Giemsa/May-Grünwald(Merck, 독일 다름슈타트 소재)로 염색하였다. 망상적혈구를 브릴리언트 크레실 블루로 염색하였다. 혈액 도말에 대한 망상적혈구를 계수한 조사자들은 샘플 그룹 할당에 대해 맹검이었다. 동물 번호만이 슬라이드에 보였다. (동물 당 5 개 슬라이드, 5 개의 무작위 1 $cm^2$ 절편)

[1033] **통계적 분석***:* 다중 그룹의 비교를 위해, 다중 비교를 위한 본페로니 사후-테스트와 함께 일원 및 이원 분산 분석(ANOVA)를 이용하였다. 통계적 분석은 GraphPad Prism 버전 6.01(GraphPad Software Inc., 캘리포니아주 라호이아 소재)을 사용하여 수행하였다.

[1034] ***결과 및 논의***

[1035] **γ-글로빈 유전자 추가를 위한 HDAd-combo 벡터 및 γ-글로빈 재-활성화를 위한 자기-불활성화 CRISPR/Cas9:** HDAd5/35++ 벡터의 30 kb 삽입 용량은 2 개의 치료 카세트를 1 개의 벡터로 통합하기 위해 활용되었다(도 100, 상부 패널, "HDAd-combo"): *i)* 인간 γ-글로빈 발현을 구동하기 위해 β-글로빈 프로모터와 조합된 HS1-HS4 미니-LCR로 이루어진 SB100x에 의한 γ-글로빈 유전자 추가용 카세트(Wang 등, *J Clin Invest* 129: 598-615, 2019). 이 카세트는 저용량 $O^6$BG/BCNU 처리에 의해 안정하게 형질도입된 세포의 선택을 허용하는 보편적으로 활성인 PGK 프로모터의 제어 하에 돌연변이체 $O^6$-메틸구아닌-DNA 메틸트랜스퍼라제($mgmt^{P140K}$)에 대한 유전자에 연결된다(Neff 등, *J Clin Invest* 112: 1581-1588, 2003; Wang 등, *Mol Ther Methods Clin Dev* 8: 52-64, 2018). *ii)* IR/*frt* 플랭킹된 트랜스포존 외부에 위치한 CRISPR/Cas9 발현 카세트인, γ-글로빈/$mgmt^{P140K}$ 트랜스포존 카세트는 *frt* 부위 및 IR에 의해 플랭킹된다. 이 모듈은 HBG1/2 프로모터 내에서 BCL11A 결합 부위를 표적화하는 U6-프로모터 구동된 sgRNA 및 EF1α 프로모터의 제어 하에 SpCas9로 이루어진다. HDAd combo 및 HDAd-SB의 공동 감염 및 SB100x 및 Flpe 재조합효소의 발현은 IR-플랭킹된 γ-글로빈/$mgmt^{P140K}$ 카세트의 통합을 매개하고 동시에 벡터를 파괴하고 CRISPR/Cas9 발현을 중지시킬 것이다(도 101). CRISPR/Cas9의 이러한 단축된 발현은 게놈- 편집된 세포의 생존 및 장기간 재증식 세포의 백분율을 증가시켜야 한다. 비교를 위해, HDAd5/35++ 벡터는 2 개의 상이한 모듈, HDAd-CRISPR("cut") 및 HDAd-SB-추가("add")를 별도로 함유하는 연구에 포함되었다(도 100, 중간 패널-"HDAd-cut" 및 "HDAd-SB-add").

[1036] **HUDEP-2 세포에서 벡터 검증:** 가설은 BCL11A 및 대부분의 β-글로빈 및, 낮은 수준의 γ-글로빈만을 발현하는 불멸성 인간 조혈 줄기 및 전구체 세포-유래 적혈구 전구체 세포주인 인간 제대혈 혈액-유래 적혈구 전구체(HUDEP-2) 세포에서 처음으로 테스트되었다(Kurita 등, *PLoS One* 8: e59890, 2013). HUDEP-2 세포는 γ-글로빈 재활성화 연구에 광범위하게 사용되었다(Canver 등, *Nature* 527: 192-197, 2015). HUDEP-2 세포를 대부분의 세포를 형질도입하는 MOI에서 HDAd-combo +/- HDAd-SB로 감염시키고 4 일 후, 세포를 초기에 기재된 바와 같

이 적혈구 분화 배지에서 8 일 동안 추가로 확장시켰다(Li 등, *Mol Ther* 27: 2195-2212, 2019). Cas9 웨스턴 블롯 신호는 세포가 분화/확장에 적용되면 급격하게 감소하였으며, 에피솜 HDAd-combo 벡터 카피의 손실로 인할 가능성이 크다(도 103a). HDAd-combo 벡터를 사용하여 제어된 Cas9 발현에 대한 개략도는 도 102에 제시된다. 그러나 Cas9는 12 일의 연구 기간 동안 검출가능하였다. HDAd-SB와의 공동 감염은 Cas9 발현을 35%(Diff d3) 내지 50%(Diff d8)로 감소시켰으며(도 103b), 이는 도 101에 기재된 자기-불활성화 메커니즘의 효과를 나타낸다. 유세포 분석에 의한 γ-글로빈 마킹의 분석(도 103C)은 γ-글로빈 유전자 추가 및 재활성화 모듈의 부가 효과를 시사하였다.

[1037] ***CD46/β-YAC 마우스에서 생체내 HSPC 형질도입.*** HBG1/2 프로모터를 표적화하는 HDAd5/35++ 벡터를 사용한 생체내 HSPC 형질도입 후 인간 γ-글로빈 재활성화 CD46/β-YAC 마우스가 이전에 입증되었다(Li 등, *Blood* 131: 2915-2928, 2018). 여기서, 유사한 프로토콜에 따라 새로운 HDAd-combo 벡터를 평가하였다. CD46/ β -YAC 마우스를 G-CSF/AMD3100으로 동원하고, "cut", "add" 및 "combo" 벡터를 정맥내로 주사하고, 4 주 후, 3 차례의 생체내 선택에 적용하였다(도 104a). γ-글로빈-양성 RBC의 백분율은 각 차례의 생체내 선택과 함께 증가하였으며 $O^6$BG/BCNU 주사의 마지막 차례 2 주 후 "combo" 벡터에 대해 95% 초과에 도달하였다(도 104b). "cut" 벡터로의 재활성화는 덜 효율적이고(60%) 동물 사이에 더 가변적이었다. 18 주차에, RBC 용해물을 글로빈 쇄에 대해 HPLC에 의해 분석하였다. 크로마토그램은 인간 β-글로빈, 재활성화된 인간 Gγ/Aγ (HBG1/2) 및 추가된 76-Ile G 변이체에 대한 별개의 피크를 나타낸다(Li 등, *Mol Ther Methods Clin Dev* 9: 142-152, 2018)(도 104c 왼쪽 패널, 도 105). 특히, Gγ 및 Aγ의 동시 재활성화는 "cut" 벡터로 처리된 마우스의 작은 분획에서만 보였다(도 105). "cut" 및 "combo" 벡터-처리된 마우스의 대부분은 재활성화된 Aγ만을 나타내었으며, HBG1 및 HBG2 프로모터 둘 다에서 CRISPR/Cas9의 동시 절단 결과로서 HBG2 유전자의 결실에 기인할 가능성이 크다(Li 등, *Blood* 131: 2915-2928, 2018). 도 104c(오른쪽 패널)는 인간 β-글로빈에 대한 γ-글로빈 단백질 수준을 나타낸다. "cut", "add", 및 "combo" 벡터 각각에 대해 평균, 7 %, 11% 및 17% γ-글로빈 단백질이 검출되었다. mRNA 수준에서 유사한 패턴이 보였다(도 104d). "cut" 및 "add" 벡터 사이의 차이는 유의하지 않았지만, "combo" 벡터에 대한 γ-글로빈 수준은 상당히 더 높았다. PBMC 및 골수 MNC에서 18 주차에 측정된 HBG 프로모터 표적 부위의 CRISPR/Cas9-매개 절단의 백분율은 "cut" 벡터와 비교하여 "combo" 벡터에 대해 상당히 더 높았다(도 104e, 도 106). 이는 CRISPR/Cas9 발현 감소로 이어지는 메커니즘, 및, 잠재적으로, 생체내 선택에 의해 확장된 CRISPR-편집된 HSPC의 더 나은 생존으로 인할 가능성이 더 크다. 골수 MNC에서 벡터 카피 수는 "combo" 벡터에 대한 γ-글로빈 수준 증가가 더 나은 형질도입 및 벡터 통합 때문이었다는 점을 제외하고, "add" 및 "combo" 벡터와 비슷하였다(도 104f). 상이한 마우스의 개별 전구체 콜로니에서 분석할 때, VCN은 세포 당 1 내지 6 개 카피 범위였다(도 104g). γ-글로빈 유전자 추가 및 CRISPR 절단-매개 γ-글로빈 재활성화가 장기간 재증식 HSC에서 발생하였음을 입증하기 위해, 골수 $Lin^-$ 세포를 이식하고, "cut" 및 "combo" 벡터를 갖는 β-YAC/CD46 마우스의 치명적으로 조사된 C57Bl/6 마우스로의 생체내 HSPC 형질도입 후 18 주차에 수확하였다. 2 차 수용자에서 다중-계통 재구축을 구동하는 이식된 세포의 능력을 16 주의 기간에 걸쳐 평가하였다. PBMC에서 CD46 발현에 기반한 생착률은 95%였고 안정하게 유지되었다. 유세포 분석에 의해 측정된 RBC의 γ-글로빈 마킹이 또한 "cut" 및 "combo" 벡터 각각에 대해 16 주차에 70% 및 95% 범위에서 안정하였다(도 107a). HPLC(도 107b) 또는 qRT-PCR(도 107c)에 의해 측정된 γ-글로빈 발현 수준(마우스 β-주요 기준)은 1차 마우스와 비슷하였다. 도 107b는 이식 후 16 주차에 인간 β-글로빈에 대한 γ-글로빈 단백질의 수준을 나타낸다. 도 107c 및 107d는 마우스 $\beta_{주요}$-글로빈 및 인간 β-글로빈에 대한 γ-글로빈 단백질의 수준을 나타낸다.

[1038] 혈액, 비장, 및 골수의 세포 조성물에 대한 HSPC의 유전적 조작 또는 적혈구 세포로부터의 γ-글로빈 발현에 대한 효과는 관찰되지 않았다. 도 107e 형질도입되지 않은 대조군 마우스(채워지지 않은 기호)와 비교하여 "combo" 벡터(채워진 기호)로 형질도입 후 16 주차에 혈액, 비장, 및 골수의 MNC에서 계통-양성 세포 조성. 도 107f는 혈액, 비장, 및 골수에서 세포 당 트랜스포존의 통합된 카피 수를 나타낸다.

[1039] ***SCD(Townes) 마우스에서 생체내 HSPC 형질도입 연구.*** 이 모델에서, 뮤린 α-글로빈 유전자를 인간 α-글로빈으로 대체하고 뮤린 성체 β-글로빈 유전자를 함께 연결된 인간 겸상 $\beta^S$ 및 태아 γ-글로빈 유전자로 대체하였다. β-글로빈 유전자(*HBG1*)는 CRISPR/Cas9에 의해 절단된 BCL11A 표적 부위를 함유하는 1400 bp의 5' 플랭킹 서열을 함유한다. 이는 β -글로빈 유전자의 재활성화로 이어져야 한다. Townes 모델의 게놈은 또 다른 SCD 마우스 모델인 Berkeley 모델($Hba^{0/0}$ $Hbb^{0/0}$ Tg(Hu-miniLCR α1$^G$γ$^A$γδβ$^S$)보다 더 잘 특성화되며, 이는 인간 글로빈 이식유전자의 2 개 초과의 카피를 갖는 것으로 보인다(Paszty 등, *Science* 278: 876-878, 1997).

[1040] HDAd5/35++ HSPC 유전자 요법에 적합한 Townes 모델을 만들기 위해, Townes 마우스를 인간 CD46 유전자이식 마우스와 교배하였다. 3 차례의 역교배 후, 인간 CD46 및 2 개의 인간(α, $\beta^S$/γ) 글로빈 유전자에 동형접합성인 마우스를 실험에 사용하였다(도 108a). 삼중 동형접합성 CD46/Townes 마우스는 겸상-유사 적혈구(도 108b), 중증 빈혈, 말초 혈액에서 40% 망상적혈구 뿐만 아니라 백혈구증가증 및 혈소판증가증을 나타내었다(도 108c). 후자는 조혈 장애가 적혈구 계통을 넘어 확장됨을 나타낸다. 또 다른 특징적 특성은 골수외 조혈의 결과로서 비장비대였다(도 108d).

[1041] CD46/Townes 마우스를 GCSF/AMD3100으로 동원하고 HDAd-combo+HDAd-SB 벡터를 정맥내로 주사하였다. $O^6$BG/BCNU로의 생체내 선택을 형질도입 1 주 후에 시작하고 BCNU 용량을 증가시키면서(5→7.5→10 mg/kg) 4 및 6 주차에 반복하였다. 기준선에서, 평균 5%의 RBC는 낮은 MFI로 γ-글로빈 양성이었으며, 이는 CD46/Townes 마우스에서 태아 글로빈의 불완전한 억제를 나타낸다. 3 차례의 생체내 선택 후, -글로빈-양성 RBC의 백분율은 증가하고 연구 종료까지(생체내 형질도입 후 13 주) 95% 초과에 도달하였다(도 109a). RBC 용해물의 HPLC 분석은 인간 α-글로빈 또는 $\beta^S$-글로빈의 30%인 γ-글로빈 수준을 나타내었다(도 109b 왼쪽 패널). 추가된 γ-글로빈 및 재활성화된 Aγ에 대한 피크는 명확하게 보였다(도 109b 오른쪽 패널). CD46/β-YAC 모델에서 볼 수 있는 바와 같이, 재활성화된 γ-글로빈은 총 γ-글로빈 수준에 대해 추가된 γ-글로빈보다 적게 기여하였다(도 109c). 유세포 분석에 의해 검출된 기준선 γ-글로빈의 낮은 수준은 HPLC의 검출 한계 미만이었다. RBC에서 글로빈 mRNA의 분석은 HPLC인 단백질 수준에서 보인 값을 반영하였다(도 109d). HDAd-combo 생체내 HSC 유전자 요법 후 γ-글로빈 수준은 "건강한" CD46/β-YAC 마우스에서보다 SCD CD46/Townes 모델에서 더 높았다.

[1042] 2 가지 의도된 게놈 변형을 13 주의 골수 샘플에서 검출하였다. 세포 당 평균 2.5 개의 통합된 γ-글로빈 유전자가 발견되었다(도 109e). T7EI 검정에 의해 측정된 표적 부위 절단 효율은 총 골수 MNC, $Lin^-$ 세포, PBMC, 및 비장세포에서 25-30% 범위로 비슷하였다(도 109f). CD46/Townes HSPC의 안정된 유전적 변형을 나타내기 위해, 생체내 형질도입 후 13 주차에 수확된 $Lin^-$ 세포를 2차 치명적으로 조사된 C57Bl/6 수용자에 이식하였다. RBC에서 γ-글로빈 마킹은 성인 인간 글로빈의 30% 수준에서(도 110b) 16 주에 걸쳐 안정되었다(도 110a.

[1043] ***마우스 모델에서 SCD의 표현형 교정:*** combo 벡터로 생체내 HSPC 형질도입 후 13 주차에, 겸상적혈구병의 표현형 특징을 CD46/Townes 마우스에서 분석하였다. 말초 혈액 도말에 대해 계수된 망상적혈구의 평균 백분율은 모체("건강한") CD46 유전자이식 마우스, 처리 전 CD46/Townes 마우스, 및 처리 후 13 주차의 CD46/Townes 마우스 각각에 대해 5, 39, 및 5%였다(도 111a 및 111c). 처리된 마우스에서, 혈색소감소증, 광범위하게 다양한 크기/형상(겸상 세포) 및 세포 단편화(도 108b 참조)를 특징으로 하는 CD46/Townes 마우스의 혈액 도말에서 적혈구 형태는 CD46 마우스에서 보이는 정상적혈구 적혈구 모양으로 되돌아갔다(도 111b). RBC, WBC, 및 혈소판 계수 뿐만 아니라 적혈구 특징(예를 들어 헤모글로빈 및 적혈구용적률)을 포함하는 혈액학적 매개변수는 CD46 및 처리된 CD46/Townes 마우스에서 유사하였다(도 111c). 마찬가지로, 처리된 CD46/Townes 마우스의 간 및 비장의 조직학적 분석은 실질 철분 침착 및 골수외 혈구형성의 부재를 포함하는 정규화를 나타내었다(도 112a). 처리된 CD46/Townes 마우스에서 보상 혈구형성의 측정가능한 특징인 비장 크기는 모체 CD46 마우스와 비슷하였다(도 112b).

[1044] 전반적으로, 이들 데이터는 CD46/Townes 마우스에서 겸상적혈구병의 완전한 치유를 나타낸다. 이는 SB100x 트랜스포사제-매개 γ-글로빈 유전자 추가(주요 기여) 및 내인성 γ-글로빈의 CRISPR/Cas9-촉발된 재활성화의 조합에 의해 달성된 높은 γ-글로빈 수준(20% 초과)과 직접적으로 관련된다고 가정된다. 또한, 이러한 결과는 Flpe/SB100x에 의한 Cas9 발현 감소가 HDAd-combo 게놈으로부터 CRISPR/Cas9 발현 카세트의 절제를 매개하여, CRISPR-편집된 HSPCS의 안전성 및 백분율 증가로 이어짐을 입증한다. 이 시스템이 추가 개선은 예를 들어 SCD 돌연변이를 HDAd-combo 벡터 내로 교정하는 Prime Editor를 포함함으로써, RBC에서 $\beta^S$의 양을 감소시키는 접근법을 포함할 수 있다.

[1045] **실시예 4. Ad35 벡터 생산**

[1046] 이 실시예는 Ad35 벡터 생산 및 CD34+ 세포의 형질도입에 대한 효능 입증을 기재한다. 상이한 구조를 갖는(상이한 LoxP 배치 포함) 3 개의 예시적인 Ad35 벡터를 생산하였다.

[1047] 대표적인 Ad5/35 헬퍼 바이러스 게놈의 왼쪽 단부는 도 113에 제시된다. 진회색으로 음영처리된 서열은 천연 Ad5 서열에 상응하며, 즉, 음영처리되지 않거나 또는 연회색으로 강조된 서열은 인위적으로도 도입되었다. 연회색으로 강조된 서열은 (연속 반복되는) loxP 서열의 2개 카피이다. "cre 재조합효소" 단백질의 존재 하에, 2

개의 loxP 서열 사이의 뉴클레오티드 서열이 결실된다(loxP의 1 개 카피만 남겨둠). loxP 부위 사이의 Ad5 서열은 아데노바이러스 DNA를 (생산자 세포의 핵 내의) 캡시드로 패키징하는 데 필수적이기 때문에, 이 결실은 패키징에 대한 헬퍼 아데노바이러스 게놈 DNA가 결핍되게 만든다. 결과적으로, 결실 과정의 효율은 패키징된 헬퍼 게놈 DNA(원치않은 헬퍼 바이러스 "오염")의 수준에 직접 영향을 미친다. 상기 관점에서, Ad5 이외의 아데노바이러스 혈청형에 대한 동일한 체계로 번역하기 위해, 하기를 달성하는 것이 바람직하다: 1. loxP 서열 삽입에 의해 플랭킹되고 cre 재조합효소의 존재 하에 결실될 수 있도록 패키징에 필수적인 서열을 식별한다. 이러한 서열의 식별은 서열에 유사성이 거의 없는 경우 간단하지 않다. 2. 천연 DNA 서열에서 loxP 서열의 삽입이 (cre 재조합효소의 부재 하에) 헬퍼 바이러스의 증식 및 패키징에 최소 영향을 미칠 수 있는지를 결정한다. 3. 헬퍼-의존적 아데노바이러스의 생산 동안(즉, 116 세포주와 같은 세포주를 발현하는 cre 재조합효소에서) 패키징 서열의 효율적인 결실을 허용하고 헬퍼 바이러스 패키징을 최소로 유지하는 loxP 서열 사이의 간격을 결정한다.

[1048] 도 114는 대표적인 Ad5 및 Ad35 패키징 신호(서열번호: 49 및 50)의 정렬을 나타낸다. Ad5의 왼쪽 단부 서열과 Ad35의 정렬은 패키징 신호를 식별하는 데 도움이 된다. 패키징에 중요한 Ad5 서열의 모티프(AI 내지 AV)는 선으로 표시된다(또한 Schmid 등, *J Virol.*, 71(5):3375-4, 1997의 도 1b 참조). 예시적인 loxP 삽입 부위의 위치는 검정색 화살표로 표시된다. 이러한 삽입은 AI 내지 AIV에 플랭킹하고 AV를 파괴한다. Schmid 등에 표시된 바와 같은 추가적인 패키징 신호 AVI 및 AVII은 이 벡터의 E1 결실의 일부로서 Ad5 헬퍼 바이러스에서 결실되었다.

[1049] 도 115는 Ad35 벡터 pAd35GLN-5E4의 개략도이다. 이는 재조합 기술(PMID: 28538186)을 사용하여 벡터화된 Ad35 게놈(ATCC의 Holden 균주)으로부터 유래된 1세대(E1/E3-결실) Ad35 벡터이다. 그런 다음 이 벡터 플라스미드를 사용하여 loxP 부위를 삽입하였다.

[1050] 패키징 부위(PS)1 LoxP 삽입 부위는 뉴클레오티드 178 및 344 뒤에 있으며; 이 Ad35 벡터는 서열번호: 286에 예시되어 있다. 이 LoxP 배치는 AI 내지 AIV를 제거하는 것으로 예상된다. AVI 및 AVII을 포함하는 나머지 패키징 신호(344 이후)는 결실되었다(위치 345 내지 3113에서 E1 결실의 일부로서). PS2 LoxP 삽입 부위는 뉴클레오티드 178 및 481 뒤에 있으며; 이 Ad35 벡터는 서열번호: 51에 예시되어 있다. 추가적으로, 뉴클레오티드 179 내지 365는 결실되어, AI 내지 AV가 존재하지 않는다. 나머지 패키징 모티프 AVI 및 AVII은 HDAd 생산 동안 cre 재조합효소에 의해 제거가능하다. E1 결실은 482에서 3113까지이다. PS3 LoxP 삽입 부위는 뉴클레오티드 154 및 481 뒤에 있으며; 이 Ad35 벡터는 서열번호: 52에 예시되어 있다. 이들 3 개 벡터의 패키징 신호 구조는 도 116에 제공된다.

[1051] 3 개의 조작된 벡터는 구제될 수 있다. 재배열된 loxP 부위를 갖는 바이러스 게놈의 백분율은 PS1, PS2, 및 PS3 각각에 대해 50, 20, 및 60%였다. 재배열은 lox P 부위가 바이러스 복제 및 유전자 발현에 중요하게 영향을 미칠 때 발생한다.

[1052] 현재 HDAd5/35 플랫폼과 비교하여 이 HDAd35 플랫폼은 도 117에 예시된다. 두 벡터는 CMV-GFP 카세트를 함유한다. Ad35 벡터는 면역원성 Ad5 캡시드 단백질을 함유하지 않는다. 이들 2 개의 벡터는 시험관내에서 CD34+ 세포의 비슷한 형질도입 효율을 나타내었다. 가교 연구는 시험관내에서 CD34+ 세포의 비슷한 형질도입 효율을 나타낸다. G-CSF 동원된 공여자로부터의 말초 CD34+ 세포인 인간 HSC를 HDAd35(Ad35 헬퍼 P-2로 생산) 또는 Ad35으로부터의 섬유가 있는 Ad5 캡시드를 함유하는 키메라 벡터로 MOI 500, 1000, 2000 vp/세포에서 형질도입하였다. GFP-양성 세포의 백분율을 3 개의 독립적인 실험에서 바이러스 첨가 48 시간 후에 측정하였다.

[1053] PS2 헬퍼 벡터는 원숭이 연구에 사용하기 위해 다시 제조하였다(도 118에 예시된 바와 같음). 이 버전을 만들기 위해 취해진 조치는 다음과 같다: E1 영역 결실, Loxp에 의해 플랭킹된 돌연변이체 패키징 신호, 돌연변이체 패키징 서열, E3 영역 결실(27435→30540), Ad5E4orf6으로 대체, 스터퍼 DNA 플랭킹 copGFP 카세트 삽입, 및 Ad35K++를 만들기 위해 놉에 돌연변이 도입.

[1054] 도 119는 돌연변이된 패키징 신호 서열을 나타낸다. 잔기 1 내지 137은 Ad35 ITR이다. 굵은 글씨 글자는 SwaI 부위이고, Loxp 부위는 이탤릭체로 표시되고, 돌연변이된 패키징 신호는 밑줄로 표시된다. 명확성을 위해, 이들 서열은 도 119에 개별적으로 제시된다.

[1055] 4 개의 Ad35 헬퍼 벡터 패키징 신호 변이체를 만들었다(도 120a). E3 영역(27388 → 30402)을 결실시키고 CMV - eGFP 카세트를 E3 결실, Ad35K++ 내에 위치시키고, eGFP를 copGFP 대신에 사용하였다. 이들 4 개의 패키징 신호 변이체에서 LoxP 부위는 예시된 위치에 있다(도 120a). 4 개의 헬퍼 벡터 모두를 구제할 수 있다.

[1056] 도 120b는 명시된 LoxP 부위가 있는 8 개의 추가적인 패키징 신호 변이체를 나타내는 개략도이다.

[1057] 특정 추가적인 헬퍼 벡터 및 패키징 신호 변이체에서, E3 결실의 단축(27609→30402)과 같은 변화가 도 120a의 헬퍼 벡터에서 이루어졌다.

[1058] **실시예 5. HDAd5/35++ 벡터를 사용한 생체외 및 생체내 조혈 줄기 세포 형질도입 후 AAVS1 유전자이식 마우스에서 표적화된 통합 및 높은 수준의 이식유전자 발현.**

[1059] 이 실시예에 함유된 정보의 적어도 일부는 Li 등*(Mol Ther.,* 27(12): 2195-2212, 2019; e-pub August 19, 2019)에 공개되었다.

[1060] 환자에서 현재 조혈 줄기 세포 유전자 요법은 유전자 전달을 위해 렌티바이러스 벡터를 사용한다(Nadini, *EMBO Mol Med*, 11, 2019; Wang 등, *Genome Res*, 17, 1186-1194, 2007). 렌티바이러스 벡터는 활성적으로 전사되는 유전자에 대한 강한 편향으로 인간 게놈에서 효율적으로 통합된다. 이러한 반-무작위 통합 패턴은 암-관련 유전자를 포함하는 이웃 유전자의 발현을 교란시킬 위험을 수반한다. 따라서 이 분야에서의 주요 목표는 미리 선택된 부위에 이식유전자 통합을 표적화하는 것이다. 인간 게놈 내로의 표적화된 통합을 위한 다수의 "안전 항구"가 제안되었다(예를 들어 AAVS1 및 CCR5)(Papapetrou 등, *Nat Biotechnol*, 29, 73-78, 2011). 안전 항구 부위에 대한 기준 중에는 다음이 있다: (i) 임의의 유전자의 5' 단부로부터 50 kb 초과의 거리, (ii) 암-관련 유전자로부터 300 kb 초과의 거리, (iii) 임의의 microRNA로부터 300 kb 초과의 거리, (iv) 유전자 전사 단위 외부, 및 (v) 초보존된 영역의 외부. 염색체 19에서 AAVS1 유전자좌는 AAVS1 부위 내에서 특이적 모티프(RBS)를 인식하는 바이러스-암호화된 단백질 Rep 78에 의해 매개되는 통합을 위해 야생형 AAV에 의해 사용된다(Muzyczka, *Curr Top Microbiol Immunol*, 158, 97-129, 1992, Huser 등, *PLoS Pathog*, 6, e1000985, 2010). 일부 AAV 혈청형에 대한 검출가능한 항체에 의해 입증되지만, 임의의 식별가능한 병리 없이 인간 집단의 많은 비율이 AAV에 직면하였기 때문에, AAVS1로의 통합은 안전할 수 있는 것으로 결론내렸다(Henckaerts 등, *Future Virol*, 5, 555-574, 2010). 또한, 이 유전자좌는 CD34+ 및 iPS 세포에서 개방 염색질 형태를 유지하는 DNAse I 과민성 부위 및 절연체를 함유한다(van Rensburg 등, *Gene Ther*, 20, 201-214, 2013, Lombardo 등, *Nat Methods*, 8, 861-869, 2011, Ogata 등, *J Virol*, 77, 9000-9007, 2003). 이는 게놈 편집 도구의 더 나은 접근을 허용하고 다른 한편으로는 높은 수준의 이식유전자 발현을 뒷받침한다(van Rensburg 등, *Gene Ther,* 20, 201-214, 2013, Voigt 등, *J Mol Med*, 86, 1205-1219, 2008).

[1061] 표적화된 이식유전자 통합은 상동성-지시된 복구(HDR)를 통해 달성될 수 있다(Lombardo 등, *Nat Med*, 20, 1101-1103, 2014). 조작된 부위-특이적 뉴클레아제에 의한 절단 후, DNA 이중 가닥 파괴는 전형적으로 가변 삽입 또는 결실(indel), 또는 HDR로 이어지는 오류가 발생하기 쉬운 DNA 복구 경로인 비-상동 말단 연결(NHEJ)을 통해 해결되며, 이는 상동 공여자 주형을 카피함으로써 DNA를 복구한다. 파괴 부위 주위의 게놈 서열에 상동 DNA에 의해 플랭킹된 외인성 DNA의 전달은 부위-특이적 방식으로 외인성 서열의 혼입으로 이어질 수 있다.

[1062] 표적화된 통합을 달성하기 위한 현재 접근법은 엔도뉴클레아제-암호화된 mRNA 및 공여자 플라스미드 DNA(Blair 등*, J Vis Exp*, e53583, 2016, Dreyer 등*, Biomaterials*, 69, 191-200, 2015; Kuhn 등*, Sci Rep*, 7, 15195 2017; Li 등*, Mol Med Rep*, 15, 1313-1318, 2017), 통합-결핍 렌티바이러스 벡터(IDLV)(Lombardo 등*, Nat Med*, 20, 1101-1103, 2014; Rio 등*, EMBO Mol Med*, 6, 835-848, 2014) 또는 rAAV6 벡터(De Ravin 등*, Nat Biotechnol*, 34:424-429, 2016, Hung 등*, Mol Ther*, 26, 46-467, 2018; Johnson 등*, Sci Rep*, 8:12144, 2018)를 사용하여 시험관내에서 HSC의 전기천공법에 기반한다. 헬퍼-의존적 아데노바이러스(HDAd5/35++) 벡터는 설계자 인터그라제(Li 등*, Blood*, 1431, 2915-2928, 2018, Saydaminova 등*, Mol Ther Methods Clin Dev*, 1, 14057, 2015) 및, 이 연구에서 공여자 주형을 전달하기 위해 개발되었다. HDAd5/35++ 벡터는 원시 HSC 상에서 발현되는 수용체인 인간 CD46을 표적화한다(Richter 등*, Blood*, 128, 2206-2217, 2016). 게놈을 비-분할 세포의 핵 내로 효율적으로 전달하는 HDAd5/35++ 벡터의 능력은 효율적인 표적화된 통합에 대한 전제 조건인 다량의 공여자 DNA를 허용한다. HDAd5/35++ 및 HDAd35 벡터는 최대 30 bp의 외래 DNA를 전달할 수 있기 때문에, 이들은 주어진 표적 부위에 상동인 공여자 서열의 긴 스트레치를 수용할 수 있다. 이는 상동 재조합에 의해 표적화된 유전자의 효능을 증가시켜야 하며, 상동성 영역의 길이와 직접적으로 상관관계가 있다(Balamotis 등, *Virology*, 324, 229-237, 2004, Ohbayashi 등*, Proc Natl Acad Sci USA*, 102, 13628-13633, 2005, Suzuki 등*, Proc Natl Acad Sci USA*, 105, 13781-13786, 2008). 이러한 벡터는 고수율로 생성하기 쉽고 강한 HSC 향성을 갖기 때문에, 이들은 생체내 HSC 형질도입에 이용되었다(Richter 등*, Blood*, 128, 2206-2217, 2016). 접근법의 핵심 아이디어는 G-CSF/AMD3100을 사용하여 골수로부터 HSC를 동원하고, 주변부에서 많은 수로 순환하는 동안, 정맥내로 주사된 HDAd5/35++ 벡터로 형질도입하는 것이다. 형질도입된 세포는 장기간 지속되는 경우 골수

로 되돌아간다. 접근법의 안전성 및 효능은 이전에 내인성 태아 글로빈의 CRISPR/Cas9-매개 재활성화(Li 등, *Blood*, 1431, 2915-2928, 2018) 또는 효율적인 무작위 일시적 통합을 매개하는 과활성 잠자는 미녀 트랜스포사제(SB100x)를 사용한 태아 글로빈 유전자 추가(Wang 등, *J Clin Invest*, 129, 598-615, 2019)에 의해 혈색소병증에 대한 CD46 유전자이식 마우스 모델에서 입증되었다. SB100x-매개 이식유전자 통합은 이론적으로 렌티바이러스 벡터의 준-무작위 통합보다 더 안전하지만, 여전히 이식유전자 침묵, 이웃 유전자에 대한 원치않은 효과, 및 게놈 재배열과 관련한 우려를 일으킨다. 따라서 이 연구의 목표는 AAVS1 내로 표적화된 통합을 위한 HDAd5/35++ -기반 생체내 HSC 형질도입 접근법을 변형하는 것이었다.

[1063] 인간 AAVS1 유전자좌에 대해 상동인 서열은 설치류에 부재한다(Samulski 등, *EMBO J*, 10, 3941-3950, 1991). 2 개의 유전자이식 설치류 모델이 이전에 보고되었으며, 이는 래트 또는 마우스 게놈(X-염색체)에서 AAVS1 유전자좌의 3.5-kb 단편(래트에서 머리에서 꼬리까지 7 개 카피)을 함유한다(Rizzuto 등, *J Virol*, 73, 2517-2526, 1999). 연구는 AAVS1의 개방 염색질 구조가 유전자이식 마우스에서 유지됨을 나타내었다(Young 등, *J Virol*, 74, 3953-3966, 2000). Jackson Laboratories는 AAVS1 유전자이식 마우스를 배포한다(Bakowska 등, *Gene Ther*, 10, 1691-1702, 2003). Jackson Labs의 웹사이트는 이러한 마우스가 단일 게놈 부위에 삽입된 8.2 kb 인간 AAVS1 유전자좌 단편의 5 개 카피를 함유함을 언급한다. HDAd5/35++ 벡터를 사용한 형질도입에 적합한 AAVS1 유전자이식 마우스를 만들기 위해, 인간 CD46 유전자좌에 대해 유전자이식된 마우스를 교배하였다(Kemper 등, *Clin Exp Immunol*, 124, 180-189, 2001). 모든 동물 연구는 AAVS1/CD46$^{+/+}$ 마우스로 수행하였다.

[1064] **재료 및 방법**.

[1065] ***세포***: G-CSF-동원된 성체 공여자로부터 *CD34*$^{+}$ *세포*를 수득하였다. 세포를 동결 스톡에서 회수하고 페니실린/스트렙토마이신, Flt3 리간드(Flt3L, 25 ng/ml), 인터류킨 3(10 ng/ml), 트롬보포이에틴(TPO)(2 ng/ml), 및 줄기세포 인자(SCF)(25 ng/ml)를 함유하는 StemSpan H3000(STEMCELL Technologies, 캐나다 밴쿠버 소재)에서 밤새 배양하였다. 세포를 2000 vp/세포의 MOI에서 HDAd 벡터로 형질도입하고 나타낸 바와 같이 분석하였다. *HUDEP-2 세포*. HUDEP-2 세포(Kurita 등, *PLoS One*, 8, e59890, 2013)를 또한 수득하였다. HUDEP-2 세포를 이전에 기재된 바와 같이 SCF, EPO, 독시사이클린 및 덱사메타손의 존재 하에 배양하였다(Canver 등, *Nature*, 527, 192-197, 2015). 세포를 500-1000 vp/세포의 MOI에서 HDAd 벡터로 형질도입하고 나타낸 바와 같이 분석하였다.

[1066] ***HDAd5/35++ 벡터:*** HDAd-SB, HDAd-IR-GFP/mgmt, 및 HDAd-IR-γ-글로빈/mgmt는 이전에 기재되었다(Li 등, *Mol Ther Methods Clin Dev*, 9, 142-152, 2018, Wang 등, *Mol Ther Methods Clin Dev*, 8, 52-64, 2018). HDAd-CRISPR 벡터의 클로닝을 위해, 인간 AAVS1 유전자좌를 표적화하는 sgRNA(서열번호: 207)(Mali 등, *Science*, 339, 823-826, 2013)를 합성하고, 어닐링하고, pSPgRNA(Addgene, 매사추세츠주 케임브리지 소재)의 *BbsI* 부위로 삽입하여, pSP-sgAAVS1을 생성하였다. pLentiCRISPRv2(Addgene)로부터 증폭된 Cas9 코딩 서열, pSP-sgAAVS1의 *BamHI* 소화에 의해 방출된 U6sgAAVS1 단편, 및 이전에 기재된 microRNA 표적화 영역(miR-183/218)(Saydaminova 등, *Mol Ther Methods Clin Dev*, 1, 14057, 2015)을 pBS-T-EF1α의 *EcoRV-NotI*, *BamHI* 및 *NotI* 부위에 순차적으로 클로닝하여(Saydaminova 등, *Mol Ther Methods Clin Dev*, 1, 14057, 2015), pBST-sgAAVS1-miR을 형성하였다. 재조합 아데노바이러스 플라스미드를 수득하기 위해, U6 프로모터에서 시작하여 SV40 polyA 신호 서열까지의 8 kb 카세트를 pBST-sgAAVS1-miR로부터 증폭하고 Gibson 어셈블리(New England Biolabs)에 의해 *NheI-XmaI* 소화된 pHCA와 결찰시켜(Sandig 등, *Proc Natl Acad Sci USA*, 97, 1002-1007, 2000), 상응하는 pHCA-sgAAVS1-miR 플라스미드를 생성하였다.

[1067] HDAd-GFP-공여자 벡터의 구축을 위해, AAVS1 CRISPR 절단 부위에 바로 플랭킹하는 2 개의 0.8 kb 상동성 아암(HA)을 gBlock(IDT, 캘리포니아주 산호세 소재)으로 합성하였다. PAM 서열이 있는 1 개의 23 bp sgAAVS1은 각각 5'HA의 상류 및 3'HA의 하류에 포함되어, 공여자 카세트의 방출을 매개하였다. EF1α-mgmt-2A-GFP-pA 단편을 GenScript(중국 난징 소재)에 의해 합성하고, 중첩 PCR에 의해 2 개의 5'HA와 결찰시켜, sgAAVS1-5'HA-Ef1α-mgmt-2A-GFP-pA-3'HA-sgAAVS1을 형성한 후 pHCA의 *XmaI* 부위에 삽입시켜(Sandig 등, *Proc Natl Acad Sci USA*, 97, 1002-1007, 2000), GFP 공여자 벡터 pHCA-AAVS1-GFP-mgmt를 생성하였다.

[1068] HDAd-글로빈-공여자 벡터의 클로닝은 3 단계를 수반하였다. 단계 1) 11.8 kb LCR-글로빈-mgmt 카세트를 *EcoRV-KpnI* 소화에 의해 pHM5-FR-IR-LCR-글로빈-mgmt로부터 방출하고(Li 등, *Mol Ther Methods Clin Dev*, 9, 142-152, 2018) pBS-Z로부터 증폭된 2.8 kb 플라스미드 백본과 결찰시켜(Saydaminova 등, Mol *Ther Methods Clin Dev*, 1, 14057, 2015), pBS-LCR-글로빈-mgmt를 생성하였다. AAVS1 CRISPR 절단 부위에 바로 인접한 2 개의 1.8 kb HA를 PAM 서열이 있는 23 bp sgAAVS1을 함유하는 프라이머를 사용하여 AAVS1-tg 마우스의 골수 세

포로부터 단리된 게놈 DNA로부터 PCR 증폭시켰다. 5' 및 3' 측면 HA를 pBS-LCR-글로빈-mgmt의 *EcoRV* 및 *KpnI* 부위에 각각 순차적으로 삽입하여, pBS-AAVS1-글로빈-mgmt를 생성하였다. 단계 2) pHCA의 nt1588-12121 영역을 *EcoRI* 소화에 의해 결실시키고 자기-결찰시켜, pHCAS1을 생성하였다. pHCAS1에서의 원래 *PacI* 부위를 2 개의 어닐링된 올리고 서열을 삽입하여 파괴하였다. 새로운 *PacI* 클로닝 부위를 *BstBI* 부위에 생성하여, pHCAS1-MCS를 수득하였다. 이 클로닝 부위는 2 개의 15 bp 상동 영역이 *PacI* 소화 시 노출되는 방식으로 설계하였다. pHCAS1-MCS의 크기를 1.5 kb *NheI* 단편을 제거함으로써 추가로 감소시켜, pHCAS1S-MCS를 생성하였다. 단계 3) 상기 2 개 단계로부터의 2 개의 최종 작제물을 *PacI* 소화한 후, 생성물을 Gibson 어셈블리에 의해 재조합하여, 글로빈 공여자 벡터 pHCA-AAVS1-글로빈-mgmt를 생성하였다.

[1069] HDAd5/35++ 벡터의 생산을 위해, 상응하는 플라스미드를 *PmeI*로 선형화하고 다른 곳에 상세히 기재된 바와 같이 (Palmer 등, *Mol Ther*, 8, 846-852, 2003) Ad5/35++-Acr 헬퍼 벡터(Li 등, 2018. Blood, 1431, 2915-2928)가 있는 116 세포에서 구제하였다(Palmer 등, *Mol Ther*, 8, 846-852, 2003). 헬퍼 바이러스 오염 수준은 0.05% 미만인 것으로 밝혀졌다. 역가는 6-12x$10^{12}$ vp/ml였다. 이 연구에 사용되는 모든 HDAd 벡터는 Ad5 섬유 꼬리, Ad35 섬유 샤프트, 및 친화도-향상된 Ad35++ 섬유 놉으로 구성된 키메라 섬유를 함유한다(Wang 등, *J Virol*, 82, 10567-10579, 2008).

[1070] ***불일치 민감성 뉴클레아제 검정 T7E1 검정.*** 게놈 DNA를 이전에 기재된 바와 같이 단리하였다(Miller 등, *Nat Biotechnol*, 25, 778-785, 2007). AAVS1 표적 부위를 포함하는 게놈 분절을 다음 프라이머를 사용하여 KOD Hot Start DNA 폴리머라제(MilliporeSigma, 매사추세츠주 버링턴 소재)에 의해 증폭시켰다: AAVS1 정방향(서열번호: 208); 역방향(서열번호: 209). PCR 생성물을 혼성화하고 37℃에서 20 분 동안 2.5 단위의 T7EI(NEB)로 처리하였다. 소화된 PCR 생성물을 6% TBE PAGE(Bio-Rad)에 의해 분해하고 에티듐 브로마이드로 염색하였다. 밴드 강도를 ImageJ 소프트웨어를 사용하여 분석하였다. 절단율 % = (1-sqrt(모체 밴드/(모체 밴드 + 절단된 밴드)))x100%

[1071] ***차세대 서열분석:*** 삽입/결실(indel)의 심층 서열분석을 위해, 예측된 AAVS1 절단 부위 주위의 250-bp 영역을 증폭시키고 Illumina 시스템을 사용하여 생성물을 서열분석하였다. 게놈 DNA를 이전에 기재된 바와 같이 단리하였다(Saydaminova 등, *Mol Ther Methods Clin Dev*, 1, 14057, 2015). AAVS1 표적 부위를 수반하는 249 bp 게놈 영역을 다음 프라이머를 사용하여 증폭시켰다: AAVS1 정방향(서열번호: 210); 역방향(서열번호: 211). AMPure XP Beads(Beckman Coulter, 인디애나주 인디애나폴리스 소재)를 사용하여 앰플리콘을 세정한 후, Klenow 단편을 사용하여 dA-테일링을 수행하였다. Illumina-호환성 어댑터를 T4 리가제(New England Biolabs)에 의해 생성물과 결찰시켰다. 고유한 바코드 서열을 PCR에 의해 도입하여 동일한 서열분석 수행 시 다중 샘플을 서열분석하였다. 각 단계는 AMPure XP Beads를 사용한 정제로 이어졌다. 최종 라이브러리를 Qubit(Invitrogen)에 의해 정량화하고 Agilent 2100 Bioanalyzer에서 테스트하여 앰플리콘의 평균 크기를 결정하였다. 앰플리콘을 동일한 몰 농도로 풀링하고 Illumina MiSeq 시스템에서 심층 서열분석하였다. 앰플리콘당 $10^5$ 개의 판독물을 생성하여 돌연변이 유형을 적절하게 조사하였다. 서열분석 데이터를 Cas-Analyzer 온라인 툴(available at rgenome.net/cas-analyzer/#!에서 이용가능)을 사용하여 AAVS1 참조 서열에 대해 정렬하였다(Park 등, *Bioinformatics*, 33, 286-288, 2017, NGS 데이터 분석을 위한 자바스크립트-기반 구현.

[1072] ***유세포 분석:*** 세포를 FACS 완충액(1 % 열-불활성화 FBS가 보충된 PBS)에 1x$10^6$ 개 세포/100 μL로 재현탁하고 얼음 위에서 10 분 동안 FcR 차단 시약(Miltenyi Biotech, 캘리포니아주 오번 소재)과 함께 배양하였다. 다음으로 염색 항체 용액을 $10^6$ 개 세포 당 100 μL로 첨가하고 암실에서 30 분 동안 얼음 위에서 배양하였다. 배양 후, 세포를 FACS 완충액에서 1 회 세척하였다. 2차 염색을 위해 염색 단계를 2차 염색 용액으로 반복하였다. 세척 후, 세포를 FACS 완충액에 재현탁하고 LSRII 유세포 분석기(BD Biosciences, 캘리포니아주 산호세 소재)를 사용하여 분석하였다. 파편을 전방 산란 영역 및 측방 산란 영역 게이트를 사용하여 제외하였다. 그런 다음 단일 세포를 전방 산란 높이 및 전방 산란 폭 게이트를 사용하여 게이팅하였다. 그런 다음 유세포 분석 데이터를 FlowJo(버전 10.0.8, FlowJo, LLC)를 사용하여 분석하였다. LSK 세포의 흐름 분석을 위해, 세포를 비오틴-접합된 계통 검출 칵테일(Miltenyi Biotec, 캘리포니아주 샌디에이고 소재) 및 c-Kit 및 Sca-1에 대한 항체 뿐만 아니라 APC-접합된 스트렙타비딘으로 염색하였다. eBioscience (캘리포니아주 샌디에이고 소재)로부터의 다른 항체는 항-마우스 LY-6A/E(Sca-1)-PE-Cyanine7(클론 D7), 항-마우스 CD117(c-Kit)-PE(클론 2B8), 항-마우스 CD3-APC(클론 17A2), 항-마우스 CD19-PE-Cyanine7(클론 eBio1D3), 및 항-마우스 Ly-66(Gr-1)-PE(클론 RB6-8C5)을 포함하였다. Miltenyi Biotec으로부터의 다른 항체는 항-인간 CD46-APC(클론: REA312)를 포함하였다. 항-마우스 Ter-119-APC(클론: Ter-119)는 BioLegend(캘리포니아주 샌디에이고 소재)로부터 수득하

였다.

[1073] 인간 γ-글로빈의 세포내 염색을 Santa Cruz의 PE-접합된 항-인간 γ-글로빈 항체(클론 51.7)를 사용하여 수행하였다. Invitrogen의 Fix & Perm 세포 투과화 키트를 제조업체의 지침에 따라 사용하였다.

[1074] ***실시간 역전사 PCR:*** 총 RNA를 TRIzol™ 시약(Thermo Fisher Scientific)을 사용하여 제조업체의 페놀-클로로포름 추출 방법에 따라 50-100 μL 혈액에서 추출한 다음, Qiagen의 Quantitect 역전사 키트를 사용하여 역전사시켜 cDNA를 생성하였다. 키트에 제공된 gDNA 제거 시약으로 RNA 샘플을 처리하여 잠재적인 게놈 DNA 오염을 제거하였다. 비교 실시간 PCR을 Power SYBR Green PCR 마스터 믹스(Applied Biosystems)를 사용하여 수행하고 StepOnePlus 실시간 PCR 시스템(Applied Biosystems)에서 실행하였다. 다음 프라이머 쌍을 사용하였다: 마우스 RPL10(하우스키핑) 정방향(서열번호: 189), 및 역방향(서열번호: 190); 인간 γ-글로빈 정방향(서열번호: 214), 및 역방향(서열번호: 215); 마우스 β-주요 글로빈 정방향(서열번호: 193), 및 역방향(서열번호: 217).

[1075] ***글로빈 HPLC:*** 개별 글로빈 쇄 수준을 SPD-10AV 다이오드 어레이 검출기 및 LC-10AT 바이너리 펌프가 장착된 Shimadzu Prominence 기기(Shimadzu, 일본 교토 소재)에서 정량화하였다. 물/아세토니트릴 중 0.1% 트리플루오로아세트산의 38%-58% 구배 혼합물을 Vydac C4 역상 칼럼(Hichrom, 영국 소재)을 사용하여 1 mL/분의 속도로 적용하였다.

[1076] ***콜로니 형성 단위 검정.*** 2500 개의 Lin- 세포를 ColonyGEL 1202 마우스 완전 배지(ReachBio, 워싱턴주 시애틀 소재)에 삼중으로 플레이팅하고 5 % $CO_2$ 및 최대 습도에서 37℃에서 12 일 동안 배양하였다. 콜로니를 Leica MS 5 해부 현미경(Leica Microsystems)을 사용하여 열거하였다. HDAd-GFP-공여자-형질도입된 마우스로부터 유래된 콜로니의 경우, GFP 양성 콜로니로 계수하고, 선별하고 분석하였다.

[1077] ***벡터 카피 수의 측정:*** 골수 세포 또는 단일 콜로니의 총 DNA를 PureLink Genomic DNA 미니 키트(Invitrogen)에 의해 추출하였다. HDAd-GFP-공여자 또는 HDAd-글로빈-공여자로부터 추출된 바이러스 DNA를 연속으로 희석하고 표준 곡선으로 사용하였다. qPCR을 StepOnePlus 실시간 PCR 시스템(Applied Biosystems)에서 power SYBR Green PCR 마스터 믹스를 사용하여 이중으로 수행하였다. 5 ng DNA를 10 μL 반응에 사용하였다. 다음 프라이머 쌍을 사용하였다: GFP 정방향(서열번호: 218), 및 역방향(서열번호: 219); 및 mgmt 정방향(서열번호: 220), 및 역방향(서열번호: 221). 인간 γ-글로빈 프라이머는 실시간 역전사 PCR 단락에 기재되었다.

[1078] ***AAVS1 유전자이식 마우스에서 AAVS1 유전자좌의 국소화.*** TLA 라이브러리를 이전에 기재된 바와 같이 제조하였다(de Vree 등, *Nat Biotechnol*, 32, 1019-1025, 2014). 간단히 말해서, 총 골수 세포로부터의 포름알데하이드 가교된 DNA를 *NlaIII*으로 소화시켰다. 결찰 및 역가교 후, DNA를 정제하였다. 이 생성물을 *NspI*로 추가로 소화시키고 결찰시켜 2 kb의 원형 키메라 DNA를 수득하였다. 키메라 DNA를 AAVS1 특이적 TLA 프라이머를 사용하여 PCR 증폭시켰다: 정방향(서열번호: 222), 및 역방향(서열번호: 223). PCR 증폭된 생성물로부터의 TLA 라이브러리를 Illumina Nextera XT NGS 키트를 사용하여 제조업체의 프로토콜에 따라 제조하였다. 쌍형성 단부 서열분석을 NovaSeq에서 수행하였다. TLA 프로토콜은 DNA의 재셔플링으로 이어지며, 따라서 판독물을 다음 설정을 사용한 스플릿-판독 인지 정렬기 BWA(Li 등, *Bioinformatics*, 26, 589-595, 2010)를 사용하여 정렬하였다: 이전에 제안된 바와 같은 bwasw -b 7(github.com/Cergentis/Cergentis_common에서 온라인 참조)(Vain-Hom 등, 2017. Nucleic Acids Res, 45, e62). 이러한 정렬된 bam 파일을 deepTool을 사용하여 RPKM 정규화된 bigwig 파일로 변환하였다(Ramirez 등, *Nucleic Acids Res*, 42, W187-191, 2014). 게놈 와이드 분포를 WashU 에피게놈 브라우저를 사용하여 시각화하였다(Zhou 등, *Nat Methods*, 8, 989-990, 2011).

[1079] ***서던 블롯.*** 마우스 골수의 게놈 DNA를 EcoRI 또는 Blp1로 소화시키고 Prime-It RmT 무작위 프라이머 표지화 키트(Agilent Technologies)를 사용하여 $^{32}P$로 표지된 AAVS1- 또는 GFP-특이적 프로브와 함께 서던 블롯에 적용하였다. 비-혼입된 $^{32}P$ dCTP를 MicroSpin G25 칼럼(GE Healthcare)을 통해 원심분리에 의해 제거하였다. 혼성화를 PerfectHyb Plus 혼성화 완충액(Sigma)에서 수행하였다. 블롯을 Amersham Hybond-XL 필름(GE Healthcare)에 노출시켰다.

[1080] ***역PCR:*** 총 골수 세포, 단일 콜로니, HUDEP-2 세포 혼합물 또는 클론에서의 접합부를 변형이 있는 다른 곳에 기재된 바와 같은 역PCR에 의해 분석하였다(Wang 등, *J Virol*, 79, 10999-11013, 2005). 간단히 말해서, 게놈 DNA를 진탕하면서 밤새 55℃에서 게놈 DNA 용해 완충액(100 mM Tris-Cl(pH 8.0), 50 mM EDTA, 1%(w/v) SDS, 및 400 μg/mL 프로테이나제 K)과 배양한 후, 페놀-클로로포름에 노출시키고, 이소프로판올로 침전시키고, 70% 에탄올로 세척하여 단리하였다. DNA 샘플을 10 mM Tris/HCL 완충액(pH 8.5)에 용해시켰다. 5 μg의 DNA를 5

시간 동안 37℃에서 50 μL 반응물 중 30 U *NcoI*로 소화시켰다. 열-불활성화 및 정리 후, 소화된 DNA를 분자내 결찰을 위해 밤새 16℃에서 500 μL 반응 완충액 중 2.5 μL T4 리가제(New England Biolabs, M0202L)로 처리하였다. 열-불활성화 및 정리 후, 재결찰된 생성물을 KOD Hot Start DNA 폴리머라제를 사용하여 역PCR에 사용하였다. 다음 프라이머를 사용하였다: EF1α 정방향(서열번호: 224), 및 역방향(서열번호: 225); pA 정방향(서열번호: 226), 및 역방향(서열번호: 227); HS4 정방향(서열번호: 228); 및 역방향(서열번호: 229). Ef1α 및 pA 프라이머 쌍은 GFP 공여자 벡터-처리된 샘플의 5' 및 3' 접합부를 각각 분석하기 위해 사용하였다. HS4 및 EF1α 프라이머 쌍은 글로빈 공여자 벡터-처리된 샘플의 5' 및 3' 접합부를 각각 분석하기 위해 사용하였다. PCR 앰플리콘을 겔 정제하고, 클로닝하고, 서열분석하고, 정렬하여 통합 부위를 식별하였다.

[1081] ***In-Out PCR:*** 게놈 DNA를 역PCR 섹션에 기재된 바와 같이 추출하였다. 5 ng 게놈 DNA를 25 μl의 반응물에서 KOD Hot Start DNA 폴리머라제에 의한 In-Out PCR용 주형으로 직접 사용하였다. 다음 PCR 프로그램을 사용하였다: 94℃ 2 분; 98℃ 10 초, 66℃ 30 초 및 68℃ 1.5 분으로 5 회 주기; 98℃ 10 초, 63℃ 30 초 및 68℃ 1.5 분으로 5 회 주기; 98℃ 10 초, 60℃ 30 초 및 68℃ 1.5 분으로 15 회 주기; 68℃ 5 분. 사용된 프라이머는 In-Out P1(서열번호: 230), In-Out P2(서열번호: 231), 및 In-Out P3(서열번호: 232)이다. 생성물을 1% 아가로스 겔에 용해하였다. 1 개의 단일 1.6 kb 밴드는 이중대립유전자성 표적화된 통합을 나타내고; 1 개의 1.6 kb 및 1 개의 2.0 kb 밴드는 단일대립유전자성 표적화된 통합을 나타내고; 1 개의 단일 2.0 kb 밴드는 잠재적인 표적외 통합을 나타낸다.

[1082] ***표적외 절단 부위의 인 실리코 예측:*** 인간 또는 마우스 게놈에서 AAVS1 가이드 서열의 표적외 부위를 sanger.ac.uk/htgt/wge/find_off_targets_by_seq에서 이용가능한 온라인 툴을 사용하여 예측하였다.

[1083] ***동물 연구:*** 모든 실험은 통제 기관 심의 위원회 및 IACUC의 승인을 받아 수행하였다. 마우스를 특이적 무병원체 시설에서 수용하였다. AAVS1 유전자이식 마우스(C3; B6-Tg(AAVS1)A1Xob/J)(The Jackson Laboratory)는 Bakowska 등(*Gene Ther*, 10, 1691-1702, 2003)에 기재된 바와 같이 마우스의 동결보존된 배아로부터 회수하였다. 마우스는 인간 AAVS1 유전자좌에 대해 반접합성이다. AAVS1 유전자이식 마우스를 인간 $CD46^{+/+}$ 마우스와 교배시켜 생체외 연구를 위한 $AAVS1^{+/-/}CD46^{+/-}$ 마우스 및 생체내 HSC 형질도입 연구를 위한 $AAVS1^{+/-/}CD46^{+/+}$ 마우스를 수득하였다. 다음 프라이머를 CD46 마우스의 유전형 분석에 사용하였다: 정방향(서열번호: 233), 및 역방향(서열번호: 234). CD46에 동형접합성 또는 이형접합성인 마우스는 유세포 분석에 의해 검출된 PBMC에 대한 CD46 발현의 상이한 강도에 의해 식별하였다. AAVS1 이식유전자의 유전형 분석은 Jackson Labs의 권고된 프로토콜에 따라 PCR에 의해 수행하였다.

[1084] ***골수 $Lin^-$ 세포 이식:*** 수용자는 6 - 8 주령의 암컷 C57BL/6 마우스였다. 이식 당일에, 수용자 마우스를 1000 Rad로 조사하였다. 조사 4 시간 후 $1x10^6$ 개의 $Lin^-$ 세포를 꼬리 정맥을 통해 정맥내로 주사하였다. 이 프로토콜은 생체외 형질도입 $Lin^-$ 세포의 이식 및 2차 수용자로의 이식에 사용되었다.

[1085] ***HSC 동원 및 생체내 형질도입:*** 이 절차는 이전에 기재되었다(Richter 등, *Blood*, 128, 2206-2217, 2016). 간단히 말해서, HSC를 인간 재조합 G-CSF의 s.c. 주사(5 μg/마우스/일, 4 일)(Amgen 캘리포니아주 사우전드 오크스 소재) 이어서 5 일째에 AMD3100의 s.c. 주사(5 mg/kg)(Sigma-Aldrich)에 의해 마우스에 동원하였다. 또한, 동물은 바이러스 주사 16 시간 및 2 시간 전에 덱사메타손(10 mg/kg)을 i.p.로 받았다. AMD3100 후 30 및 60 분에, 동물에게 주사 당 각 바이러스에 대해 $4x10^{10}$ vp의 용량으로 안와하정맥총을 통해 HDAd-CRISPR 및 HDAd-GFP-공여자 또는 HDAd-글로빈-공여자를 정맥내로 주사하였다. 4 주 후, 마우스에게 $O^6$-BG(15 mg/kg, IP)를 2 회, 30 분 간격으로 주사하였다. $O^6$-BG의 두번째 주사 1 시간 후, 마우스에게 BCNU(5 mg/kg, IP)를 주사하였다. BCNU 용량을 두번째 주기에 10 mg/kg로 증가시켰다. BCNU 및 $O^6$-BG는 둘 다 Sigma-Aldrich로부터 수득하였다.

[1086] ***통계적 분석:*** 다중 그룹의 비교를 위해, 다중 비교를 위한 본페로니 사후-테스트와 함께 일원 및 이원 분산 분석(ANOVA)을 이용하였다. 통계적 분석은 GraphPad Prism 버전 6.01(GraphPad Software Inc., 캘리포니아주 라호이아 소재)을 사용하여 수행하였다.

[1087] **결과**

[1088] ***HDAd-CRISPR 및 HDAd-공여자 벡터의 설계.*** CRISPR/Cas9를 발현하는 HDAd5/35++ 벡터를 생성하였다. 벡터는

AAVS1 유전자좌 내에서 ds DNA 파괴를 생성할 수 있다(도 55a). 이전 연구는 이 유전자좌로의 부위-특이적 통합이 1차 인간 세포에서 부작용 없이 강력한 이식유전자 발현을 허용하였음을 입증하였다(Lombadro 등, *Nat Methods*, 8, 861-869, 2011). 상응하는 HDAd-CRISPR 벡터의 활성을 테스트하기 위해, HSC에 대해 강화된 세포 분획인 인간 CD34+ 세포를 형질도입하였다. 42%의 빈도로 감염 후 3 일차에 AAVS1 부위-특이적 절단은 불일치 민감성 뉴클레아제 검정 T7E1 검정에 의해 입증되었다(도 55b). HDAd-CRISPR 삽입/결실(indel)의 심층 서열분석을 위해, 예측된 AAVS1 절단 부위 주위의 250-bp 영역에서 PCR 증폭을 수행하고 Illumina 시스템을 사용하여 생성물을 서열분석하였다(도 55c). indel의 80%는 1 내지 20 bp 범위의 결실이었고 10%만이 1 내지 2 bp 마이크로-삽입이었다.

[1089] HDAd5/35++ 벡터를 공여자 벡터로 사용하였다. 첫번째 HDAd-공여자 벡터는 CRISPR/Cas9 표적 부위에 바로 인접한 영역에 상동인 0.8 kb 길이 영역에 의해 양측에 플랭킹된 GFP 및 $mgmt^{P140K}$에 대한 발현 카세트를 함유하였다(도 55d). 선형 이중-가닥 아데노바이러스 게놈은 바이러스가 세포에 진입하여 핵으로 전위될 때 생성된 "말단 단백질 - TP"와 공유적으로 연결된다(Shenk, *Fields Virology*, 2:2111-2148, 1996). 이는 TP가 헬퍼 바이러스 유래인 HDAd5/35++ 게놈의 경우와 동일하다. 공여자에서 유리 DNA 단부의 부재는 HDR을 크게 감소시키는 것으로 생각된다(Cristea 등, *Biotechnol Bioeng*, 110, 871-880, 2013). AAVS1 CRISPR에 대한 sgRNA 표적 부위는 공여자 이식유전자 카세트를 플랭킹하는 공여자 벡터 내로 혼입되었다(도 55d). 따라서 HDAd-CRISPR 및 HDAd-GFP-공여자의 공동 감염은 염색체 AAVS1 표적 부위에서 dsDNA 파괴를 동시에 생성하고 핵 내부에서 유입 HDAd-공여자 게놈으로부터 공여자 카세트를 방출하여야 한다. 1000 및 2000 vp/세포의 총 MOI에서 감염 후 2 일차에 $CD34^+$ 세포에서 각각 13.2 및 18.1%의 효능을 갖는 공동 감염된 HDAd-GFP-공여자 벡터로부터 공여자 카세트의 IA HDAd-CRISPR-매개 방출이 입증되었다(도 55e). 이 발견은 또한 CRISPR/Cas9가 이중-가닥 선형 아데노바이러스 DNA를 절단할 수 있으며, 항-바이러스 요법에 영향을 미친다는 것을 나타낸다.

[1090] ***시험관내 표적화된 통합.*** 먼저, HDAd-CRISPR + HDAd-공여자 벡터 시스템을 무작위 통합을 매개하는 SB100x 벡터 시스템과 직접 비교하여 시험관내에서 표적화된 통합에 대해 테스트하였다(도 56a). 인간 적혈구 전구체 세포인 HUDEP-2 세포를 사용하였다. 이 세포주는 이배체이며 단일 콜로니 확장을 허용하여 통합 부위 분석을 용이하게 하는 것을 특징으로 한다. HUDEP-2 세포의 형질도입 후 2 일차에 수행된 GFP 유세포 분석은 유사한 형질도입률을 나타내는 SB100x-매개 및 표적화된 통합 시스템에 대한 GFP-양성 세포의 유사한 백분율을 입증하였다(도 56b, 상부 패널). 2 일차에 GFP 발현은 HDAd-GFP-공여자 단독으로의 형질도입이 유사한 GFP 마킹을 초래할 수 있기 때문에 에피솜 게놈으로부터 유래할 가능성이 있다. 21 일 동안 형질도입된 세포를 배양한 후, 세포 증식으로 인해, 에피솜 게놈은 HDAd-GFP-공여자 단독 설정에서 GFP 발현의 부재를 나타낼 때 사라졌다. 21 일차에, 세포의 4.52% 및 1.82%가 각각 SB100x-매개 및 표적화된 통합 시스템에 대해 GFP-양성이었다(도 56b, 하부 패널). 이는 SB100x 시스템이 더 높은 안정된 형질도입률을 부여함을 시사한다. 그러나, 평균 형광 강도(MFI)에 의해 반영되는 GFP 발현 수준은 HDAd-CRISPR + HDAd-GFP-공여자로 형질도입된 세포에서 21 일차에 세포 집단(도 56c) 및 단일 클론 수준(도 56d) 둘 다에서 더 높았다. 벡터 통합 분석은 단일 클론에서 수행하였다. 이식유전자 카세트에 플랭킹하는 긴 상동성 영역으로 인해, 벡터 통합 부위 분석에 통상적으로 사용되는 도구(예를 들어 LAM-PCR)를 이용하는 것은 가능하지 않았다. 벡터-세포 DNA 접합부의 존재를 입증하기 위해, 4 kb 단편으로 게놈 DNA의 엔토뉴클레아제 절단, 이의 원형화, 및 이식유전자 특이적 프라이머를 사용한 후속 PCR을 수반하는 역PCR(iPCR) 방법을 사용하였다(Wang 등, *J Virol*, 79, 10999-11013, 2005). 결과는 HDAd-CRISPR + HDAd-GFP-공여자 형질도입된 HUDEP-2 세포로부터 유래된 모든 테스트된 36 개 콜로니가 AAVS1 부위로 통합된 이식유전자를 가짐을 나타내었다(도 57a). 이는 표적화된 통합이 있는 클론에서 균질한 높은 수준의 이식유전자 발현과 일치한다. AAVS1 및 이식유전자 특이적 프라이머를 사용한 In-/out-PCR은 36 개 콜로니 중 3개에서의 통합이 두 대립유전자에서 발생하였고; 36 개 중 31 개가 단일대립유전자성 통합을 가졌고, 2 개가 분명히 연쇄체 통합체를 가짐을 입증하였다(도 57b). 대조적으로, 특이적 유전자좌의 우선적 표적화가 없는 SB100x-매개 무작위 통합(Wang 등, 2019. *J Clin Invest*, 129, 598-615, Boehme 등, *Mol Ther Nucleic Acids*, 5, e337, 2016)은 다양한 유전자 침묵 수준을 초래하였다(도 56e). 유사한 수준의 벡터 카피 수를 SB100x 및 표적화된 통합을 갖는 클론에서 검출하였다(도 56f).

[1091] 요약하면, 시험관내 연구는 HDAd-CRISPR + HDAd-GFP-공여자 시스템이 고효율로 표적화된 통합을 부여하고 SB100x 매개 시스템보다 더 높은 GFP 발현 수준을 초래하였음을 제시하였다. 안정된 통합의 효능은 표적화된 시스템에 대해 40% 더 낮았다.

[1092] ***HDAd-CRISPR + HDAd-GFP-공여자를 사용한 AAVS1/CD46 HSC의 생체외 형질도입 및 치명적으로 조사된 수용자 내***

***로 후속 이식.*** 다음으로, 표적화된 통합 시스템을 AAVS1/CD46tg 마우스의 HSC에서 테스트하였다. HSC에 대해 강화된 골수 세포 분획인 계통-음성($Lin^-$) 세포의 생체외 형질도입 후 표적 부위 절단 빈도는 1000 vp/세포의 MOI에서 HDAd-CRISPR 벡터로 형질도입 후 25%였다(도 58a). 0% 및 50% 절단에서 삽입/결실의 백분율은 도 58b에 제시된다. 예시적인 서열은 도 58c에 제시된다. HDAd-CRISPR 단독, HDAd-GFP-공여자 단독, 및 둘 다의 조합으로 생체외에서 형질도입된 AAVS1/CD46 $Lin^-$ 세포를 치명적으로 조사된 C57Bl/6 마우스로 이식한 다음, 16주 동안 추적하였다(도 59a). PBMC에 대한 인간 CD46 발현에 기반하여 이식된 세포의 생착을 표시된 시점에서 CD46+ PBMC의 퍼센트에 의해 측정하였다. 형질도입된 공여자 세포는 CD46을 발현한 반면(도 60b), 수용자 C57Bl/6 마우스는 발현하지 않았다. PBMC(혈액), 비장, 및 뼈에서 CD46+ 세포의 백분율은 도 60c 및 60d에 제시된다. GFP 마커의 발현을 또한 콜로니 및 풀링된 콜로니 세포에서 분석하였다.

[1093] 공여자 세포 생착률은 모든 3 가지 설정에 대해 비슷하였으며(도 60) 이는 HDAd-CRISPR 및 HDAd-CRISPR + HDAd-GFP-공여자 벡터에 의해 HSC에 도입된 게놈 변형이 HSC 생물학, 특히 치명적으로 조사된 수용자의 다중계통 재증식에 해로운 효과를 미치지 않았음을 시사한다. 최대 100%에 도달하는 GFP 마킹률은 이식유전자가 안정하게 발현된 HSC/전구체의 3 차례의 $O^6$BG/BCNU 선택 후 PBMC에서 보였다(도 59b, 59c). 선택 전(이식 4 주 후)에, $GFP^+$ PBMC의 백분율은 1.1%였으며, 이는 표적화된 통합이 드문 사건임을 나타낸다. $GFP^+$ PBMC는 HDAd-GFP-공여자 단독으로 형질도입된 $Lin^-$ 세포로 이식된 마우스에서 평균 0.2% 미만이었다. 이는 안정된 이식유전자 발현을 달성하기 위해 CRISPR/Cas9-매개 dsDNA 파괴의 필요성에 관해 시사한다. 이식 후 16 주차에 분석된 마우스는 골수, 비장, 및 PBMC에서 분석된 모든 계통에서 GFP 마킹을 제시하였다(도 59d). GFP 마킹률은 2차 이식 수용자에서 16 주 동안 유지되었으며, 이는 원시 HSC가 혈액, 비장, 및 골수를 포함하여(도 61b, 61c), 콜로니 및 풀링된 콜로니 세포에 대해 제시된 바와 같이(도 61d), HDAd-CRISPR + HDAd-GFP-공여자 벡터 시스템으로 유전적으로 변형됨(도 61a)을 입증한다. 인간 CD46+ 세포의 퍼센트 및 혈액, 비장, 및 골수에서의 백분율은 도 61e 및 61f에 추가로 제시된다.

[1094] ***HDAd-CRISPR + HDAd-GFP-공여자를 사용한 AAVS1/CD46tg 마우스의 생체내 HSC 형질도입.*** AAVS1/CD46-유전자이식 마우스의 생체내 HSC 형질도입을 위해, HSC를 G-CSF/AMD3100의 피하 주사에 의해 골수에서 말초 혈류로 동원하고 정맥내로 전달된 HDAd-CRISPR + HDAd-GFP-공여자 벡터에 의해 생체내에서 형질도입하였다(도 62a). 3 주기의 $O^6$BG/BCNU로 생체내 선택 후, 마우스의 60%는 개별 동물에서 35 내지 95% $GFP^+$ PBMC 범위의 PBMC에서 GFP 발현을 나타내었다(도 62b). 생체내 형질도입 후 16 주차에, 혈액, 비장 및 골수의 단핵 세포에서 유사한 마킹을 보였다(도 62c). 혈액, 비장 및 골수의 $CD3^+$, $CD19^+$, 및 $Gr\text{-}1^+$ 계통 세포에서 GFP 마킹을 보였다(도 62d). "반응자"의 골수에서, LSK 세포의 50% 초과(HSC에 대해 강화된 분획)는 GFP-양성이었다(도 62d, 마지막 그룹). 이는 또한 전구체 콜로니를 형성하는 능력인 HSC에 대한 기능적 검정에 의해 반영되었다(도 62e). 또한, 원시의 장기간 재증식 HSC의 형질도입은 2차 수용자에서 제시되었다(표시된 시험에서 GFP+ PBMC의 백분율(도 63a), 혈액, 비장, 및 골수에서 GFP+ 세포의 백분율(도 63b, 63c); 인간 CD46+ 세포의 백분율(도 63d), 및 혈액, 비장, 및 골수에서의 백분율(도 63e) 참조). 생체내 HSC 형질도입/선택 절차는 골수 세포 조성물 및 조혈에 부정적인 영향을 미치지 않았다(도 62f).

[1095] ***HDAd-CRISPR 및 HDAd-글로빈-공여자 벡터의 생체외 및 생체내 HSC 형질도입.*** HDAd-GFP-공여자 벡터를 사용한 연구는 대부분의 동물에서 안정된 HSC 형질도입을 시사하지만, 더 높은 비율의 반응자가 바람직할 것이다. 이는 HDR-매개 통합의 효능을 증가시킬 필요가 있으며, 상동성 아암의 길이를 증가시킴으로써 달성될 수 있다(Balamotis 등, *Virology*, 324, 229-237, 2004, Ohbayashi 등, *Proc Natl Acad Sci USA*, 102, 13628-13633, 2005, Suzuki 등, *Proc Natl Acad Sci USA*, 105, 13781-13786, 2008). CRISPR/Cas9 절단 부위 주위의 AAVS1 게놈 서열에 상동인 1.8 kb 영역을 갖는 새로운 HDAd-공여자 벡터를 생성하였다(도 64a). 혈색소병증의 유전자 요법에서 적용하기 위해, 미니 γ-글로빈 LCR의 제어 하에 인간 γ-글로빈 유전자(HBG1)를 사용하였다. HDAd-글로빈-공여자 벡터를 생체외 및 생체내 HSC 형질도입 프로토콜 둘 다에서 테스트하였다. 생체외 형질도입 설정에서(도 64b), 모든 마우스가 말초 적혈구(RBC)의 80%에서 γ-글로빈을 발현하는 마우스에 반응하였음을 관찰하였다(도 64c). 혈액 및 골수에서 γ-글로빈-양성 적혈구($Ter119^+$) 세포의 백분율은 비-적혈구($Ter119^-$) 세포의 백분율보다 상당히 더 높았다(도 64d). $^{\gamma}$-글로빈 MFI에 대한 경우도 동일하였다(도 64e). 이는 미니-LCR이 적혈구 세포에서 우선적 발현을 부여함을 시사한다. 16 주차에, γ-글로빈 수준은 HPLC에 의해 측정된 성체 마우스 γ-글로빈 수준의 20.52(+/-5.66)%(도 64f) 및 qRT-PCR에 의해 22.33(+/-6.21)%(도 64g)였다.

SB100x-시스템을 사용한 동일한 레지멘 하에 수행된 이전 연구에서, γ-글로빈 발현 수준은 HPLC에 의해 15.74(+/-2.69)% 및 qRT-PCR에 의해 15.40(+/-9.21)%였다(Li 등, *Mol Ther Methods Clin Dev*, 9, 142-152, 2018). 이는 γ-글로빈 발현 수준이 SB100x 시스템과 비교하여 표적화된 통합 시스템에 대해 더 높음을 의미한다. 사실, 표적화된 통합 시스템의 경우, $\beta_0/\beta_0$ 지중해빈혈 또는 겸상적혈구병 환자에 대해 성인 글로빈의 20% γ-글로빈으로 생각되는 치유 수준 범위에 있을 것이다(Wang 등,. *J Clin Invest*, 129, 598-615, 2019). 이전 연구와 일치하게(Wang 등,. *J Clin Invest*, 129, 598-615, 2019), 16 주차에 게놈 당 2 개의 통합 벡터 카피를 단일 $Lin^-$ 세포-유래 콜로니에서 평균으로 측정하였다(도 64h). $Lin^-$ 세포의 생체외 HSC 형질도입은 치명적으로 조사된 수용자에서 다중계통 생착 및 완전한 조혈 재구축에 대한 능력에 영향을 미치지 않았다(표시된 시점에서 인간 CD46+ 세포의 백분율(도 65a), 혈액, 비장, 및 골수에서의 백분율(도 65b) 참조). 2차 HSC 이식 수용자의 분석은 HDAd-CRISPR + HDAd-글로빈-공여자 벡터로 생체외 형질도입 이어서 생체내 선택이 장기간 재증식할 수 있는 HSC의 풀에 영향을 미치지 않았음을 나타내었다(RBC에서 인간 γ-글로빈+ 세포의 백분율(도 66a), 인간 CD46+ 세포의 백분율(도 66b), 및 혈액 및 골수에서의 백분율(도 66c) 참조).

[1096] HDAd-CRISPR + HDAd-글로빈-공여자 벡터를 사용한 생체내 HSC 형질도입 연구에서(도 67a), 생체내 선택 후, 5 마리 마우스 중 4 마리가 개별 마우스에서 40 내지 97% $γ\text{-글로빈}^+$ RBC 범위로 RBC에서 안정된 γ-글로빈 발현을 나타내었다(도 67b). γ-글로빈 발현은 적혈구 세포에서 우선적으로 밝혀졌다(도 67c, 67d). RBC에서 γ-글로빈 발현 수준은 성체 마우스 γ-글로빈 수준에 대해 HPLC에 의해 23.97(+/-7.22)%(도 67e, 67h) 및 qRT-PCR에 의해 24.53(+/-7.34)%(도 67f)였다. 세포 당 벡터 카피 수는 개별 마우스에서 1.5 내지 2.5 개 범위였다(도 67g). 동일한 생체내 HSC 형질도입/선택 설정에서, SB100x 기반 γ-글로빈 벡터를 사용하여, 게놈 당 평균 2 개의 통합 벡터 카피를 갖는 γ-글로빈 수준은 HPLC에 의해 10.5(+/-3.1)% 및 qRT-PCR에 의해 12.17(+/-3.38)%였다(Wang 등, *J Clin Invest*, 129, 598-615, 2019). 치명적으로 조사된 수용자 내로 HDAd-CRISPR + HDAd-글로빈-공여자를 사용한 생체내 형질도입 후 16 주차에 수확된 골수 $Lin^-$ 세포의 이식은 성체 β-글로빈의 평균 24% γ- 수준으로 16 주에 걸쳐 RBC에서 100% 생착 및 안정된 γ-글로빈 발현을 나타내었다(표시된 시점의 PBMC에서 인간 CD46+ 세포의 백분율(도 68a); 표시된 시점의 말초 혈액에서 γ-글로빈+ 세포의 백분율(도 68b); 마우스 β-주요 단백질의 백분율로서 인간 γ-글로빈(도 68c); 및 혈액, 비장, 및 골수에서 백분율(도 68d) 참조).

[1097] 요약하면, HDAd-CRISPR + HDAd-글로빈-공여자를 사용한 HSC 형질도입 연구는 SB100x-기반 시스템을 사용한 이전 연구에서 달성된 것보다 상당히 더 큰 수준으로 안정된 γ-글로빈 발현을 초래하였다.

[1098] ***AAVS 유전자이식 마우스에서 AAVS1 유전자좌의 국소화.*** 통합 부위 분석을 위한 역PCR(iPCR)은 AAVS1/CD46-유전자이식 마우스의 게놈에서 AAVS1 유전자좌 국소화에 대한 지식이 필요하다. 이를 결정하기 위해, 물리적으로 근위 서열의 가교를 수반하는 표적화된 유전자좌 증폭(TLA)/PCR 기술을 사용하였다(de Vree 등, *Nat Biotechnol*, 32, 1019-1025 2014; 재료 및 방법 참조). 그런 다음 AAVS1/CD46-tg 마우스의 골수 세포에서 수득된 TLA 데이터를 참조 마우스 게놈과 정렬하였다(도 69). TLA 결과는 위치 (Chr14:110443871-110461834)에서 염색체 14에 통합됨을 나타낸다(도 55b). 이 정보를 사용하여, 프라이머를 사용하여 유전자좌를 서열분석하였다(도 70). 왼쪽에서 오른쪽으로 및 오른쪽에서 왼쪽으로 향하는 AAVS1 유전자좌의 반복이 발견되었다. 두 말단 반복부(# 1 및 #5)를 각각 4.5 및 2.8 kb 길이로 절두하였다. 반복부 #5는 완전한 5' 상동성 영역이 결여되었다. 이러한 표적 부위 무리가 통합 부위 분석을 복잡하게 만들었다. HDAd-CRISPR + HDAd-공여자 시스템에 의한 통합에 대한 이론적 결과의 일부는 도 70에 요약되어 있다.

[1099] ***HDAd-CRISPR+HDAd-공여자를 사용한 생체외 및 생체내 HSC 형질도입 후 염색체 통합.*** 먼저 게놈 서던 블롯을 16 주차에 수확된 골수 세포의 DNA에서 수행하였다. *EcoRI*-소화된 게놈 DNA와 AAVS1 특이적 프로브의 혼성화는 모든 분석된 마우스에서 AAVS1 유전자좌의 하나(또는 그 이상)의 반복부 내로 공여자 카세트의 통합을 나타내는 3.9 kb-특이적 밴드를 나타내었다(도 71a). *Blp1*-소화된 DNA와 GFP 프로브의 혼성화는 전장 반복부 #2-4로의 통합을 나타내는 10 마리 마우스 중 5 마리에서 5.8 kb 신호를 초래하였다(도 71b). 5 및 6 kb 신호는 반복부 #1 및 5 각각으로의 통합 결과일 수 있다. 10 마리 마우스 중 2 마리는 여러 AAVS1 모티프 반복부 내로의 통합을 갖는 것으로 나타났다. 이식유전자/염색체 접합부의 존재를 입증하기 위해, iPCR을 마우스의 게놈 DNA에서 수행하였다(도 72a, 72b). 분석된 8 마리 마우스 중 6 마리는 AAVS1 부위로의 HDR-매개 통합과 일치하는 PCR 생성물을 나타내었다(도 72b). 이들 마우스 중 여러 마리는 염색체 5 상의 CRISPR/Cas9 표적외 부위 중 하나로의 통합을 초래하는 추가적인 밴드를 가졌다(도 72b). 접합부로서 ITR을 수반하는 전장 HDAd 게놈의 통합에서 유래된 밴드가 또한 발견되었다. 흥미롭게도, 이러한 통합된 전장 HDAd 게놈은 CRISPR AAVS1 표적 부위를 함유

하는 염색체인 염색체 14 상에 있었다(도 72b). 골수 세포의 풀에서 유래된 이러한 결과를 복잡하지 않게 하려는 시도에서, d GFP$^{+}$ 골수 Lin$^{-}$ 세포를 플레이팅하여 단일 세포에서 유래된 전구체 콜로니를 생성하였다(도 72c). AAVS1로의 HDR-통합에 특이적인 1 개의 밴드만 있는 마우스(예를 들어 마우스 #943)의 콜로니에 대한 분석은 모든 콜로니에서 균일한 신호를 나타낸 반면, 추가적인 표적외 통합을 갖는 마우스(예를 들어 #946)의 콜로니는 키메라 패턴을 나타내었으며: 10 개 콜로니 중 9 개는 표적내 통합만을 가졌고, 1 개 콜로니는 표적내 및 표적외 통합을 둘 다 함유하며, 이는 게놈 당 통합된 이식유전자의 평균 수가 2이기 때문에 가능하다. HDAd-CRISPR 및 HDAd-글로빈-공여자 벡터를 사용한 생체외 및 생체내 형질도입 연구에서 골수 세포의 통합 부위 분석은 유사한 결과를 보였다(도 73a 및 73b는 표적내 통합(도 73a) 및 표적내 및/또는 표적외 통합을 갖는 샘플(도 73b)을 나타낸다). HDAd-CRISPR + HDAd-글로빈-공여자를 사용한 생체외 HSC 형질도입 설정에서, HDAd-GFP-공여자 벡터를 사용한 생체내 HSC 형질도입 연구와 비교하여 표적화된 통합을 갖는 동물의 더 높은 비율이 발견되었다. 이는 더 긴 상동성 영역에 기반한 더 높은 HDR 효능 때문일 수 있다.

[1100] 전반적으로 이러한 통합 연구는 AAVS1 유전자좌로 표적화된 통합의 높은 빈도를 나타낸다. 통합 분율은 CRISPR 표적외 부위 및 아마도 표적 부위에 함유된 염색체 상의 CRISPR-촉발된 큰 결실을 수반하는 영역으로 발생하였다.

[1101] **논의**. 자기-불활성화 렌티바이러스 벡터는 감마-레트로바이러스 벡터와 대조적으로 임상 HSC 유전자 요법 시험에서 삽입 부위-연관 악성 클론 확장과 연관되지 않았다. 그러나, 이 위험은 비-인간 영장류에서 최근 연구에 나타낸 바와 같이 완전히 제외될 수는 없었다(Espinoza 등, *Mol Ther*, 6, 1074-1086, 2019). 이론적으로, SB100x에 의해 매개되는 무작위 통합 패턴 및 활성화 유전자 및 프로모터 내로 통합을 위한 선호도 결여는 더 안전해야 하지만 유전독성에 관한 우려가 남아있다. 따라서, 관련 분야의 주요 노력은 AAVS1 부위과 같은 미리 선택된 부위로 표적화된 이식 유전자 통합을 목표로 한다. 인간 HSC에서 아연 핑거 뉴클레아제 mRNA 및 AAV6-매개 공여자 주형 전달은 AAVS1 유전자좌로 50% 초과의 표적화된 통합을 초래하였다(De Ravin 등, *Nat Biotechnol*, 34, 424-429, 2016). 공여자 주형을 전달하기 위해 AAVS1-특이적 CRISPR/Cas9 RNP 및 AAV6을 이용하는 다른 연구에서, 부위-특이적 통합의 빈도는 25%였다(Johnson 등, 2018. Sci Rep, 8, 12144). CCR5로의 표적화된 통합에 대해 유사한 비율이 달성되었다(Hung 등, *Mol Ther*, 26, 456-467, 2018).

[1102] AAVS1로의 표적화된 통합을 위한 이 접근법은 다수의 새로운 측면이 있다. (*i*) 공여자 주형을 전달하기 위한 헬퍼-의존적, 캡시드-변형된 HDAd벡터의 사용. 상응하는 게놈은 두 단부 상에서 바이러스 TP 단백질에 공유적으로 연결된 이중가닥 선형 DNA이다. 단일-가닥 AAV6 공여자 벡터와 대조적으로, 이중-가닥 선형 아데노바이러스 DNA는 HDR에 대한 최적의 주형이 아닌 것으로 생각된다. 이러한 잠재적인 단점을 보완하기 위해, AAVS1 CRISPR/Cas9 절단 부위를 HDAd-공여자 벡터 내로 혼입하여 유리 "재조합원성" DNA 단부를 생성하였다. (*ii*) HDAd벡터의 삽입 용량은 30 kb이기 때문에 rAAV6 또는 IDLV 벡터의 패키징 용량을 초과하는 상동성 아암을 혼입하는 것이 가능하였다. 이전 연구(Balamotis 등, *Virology*, 324, 29-237, 2004, Ohbayashi 등, *Proc Natl Acad Sci USA*, 102, 13628-13633, 2005, 및 Suzuki 등, *Proc Natl Acad Sci USA*, 105, 13781-13786, 2008) 및 0.8 및 1.8 kb 상동성 영역을 갖는 HDAd-공여자 벡터의 비교는 상동성을 증가시키는 것이 높은 수준의 이식유전자 발현을 갖는 반응자 마우스의 수 뿐만 아니라 표적화된 통합을 갖는 마우스의 분율을 개선시켰음을 시사한다. (*iii*) 큰 HDAd5/35++ 삽입 용량은 또한 mgmt$^{P140K}$-기반 생체내 선택 카세트를 공여자 주형에 포함하는 것을 허용하여, 저용량 O$^{6}$BG/BCNU를 사용한 단기간 치료에 의해 형질도입된 원시 HSC의 풀에 영향을 미치지 않으면서 자손 세포의 선택적 생존 및 확장을 매개한다(Wang 등, *Mol Ther Methods Clin Dev*, 8, 52-64, 2018). HDR의 낮은 효능 및 결과적으로 HSC에서 표적화된 통합을 고려하면(Genovese 등, *Nature*, 510, 235-240, 2014), 생체내 HSC 선택은 말초 혈액 세포에서 높은 이식유전자 마킹 수준을 달성하는 데 중요한 것으로 보인다. (*iv*) 마지막으로, 고수율 HDAd5/35++ 벡터 생산의 용이성 및 원시 HSC에 대한 향성으로 인해, 이들은 동원된 동물에 정맥내 주사를 통한 생체내 HSC 형질도입에 사용될 수 있다. 따라서, 표적화된 이식유전자 통합으로 혈색소병증의 생체내 HSC 유전자 요법의 원리 증명을 수행하는 것이 가능하였다.

[1103] 안정된 이식유전자(GFP 또는 γ-글로빈) 발현을 달성하기 위해, HDAd-공여자 및 HDAd-CRISPR의 공동 감염이 필수적이었으며, 이는 CRISPR-매개 게놈 DNA 파괴 및, 가장 가능성 있게는 HDAd-공여자 벡터에서 공여자 주형의 방출이 통합을 크게 자극하였음을 시사한다. HDAd-공여자 + HDAd-CRISPR을 사용한 생체내 형질도입 후 HSC로 이식유전자 통합을 위한 지표는 생체내 선택 완료 후 안정된 높은 수준의 이식유전자 발현을 나타낸 마우스의 분율(즉 "반응자")이었다. HDAd-GFP-공여자 + HDAd-CRISPR의 경우 16 개 중 6 개(37.5%) 및 HDAd-글로빈-공여자 + HDAd-CRISPR의 경우 5 개 중 4 개(80%)였다. 특히, 표적화된 통합 빈도가 높은 "반응자" 비율은 생체외

형질도입 설정에서 두 벡터에 대해 100%였다. 이는 표적화된 생체내 HSC 형질도입 접근법에 대한 제한 인자가 HSC 감염의 효능임을 나타낸다. 초기 감염 단계는 이론적으로 최적화된 HSC 동원 레지멘(Psatha 등, *Hum Gene Ther Methods*, 25, 317-327, 2014) 및 하루 간격으로 2 차례의 HDAd 주사에 의해 개선될 수 있다.

[1104] 이러한 데이터는 벡터 시스템이 생체외 및 생체내 형질도입 설정에서 HSC의 표적화된 통합을 달성하기 위한 효율적인 도구임을 나타낸다. 이는 대부분 HDAd-공여자 벡터를 비-분할 세포의 핵으로 전달하는 높은 효능, 벡터 백본에서 공여자 카세트를 방출하는 능력, 및 큰 상동성 영역을 혼입하는 HDAd 벡터의 용량 때문일 수 있다.

[1105] 이 연구에서 중요한 발견은 표적화된 통합 시스템이 시험관내, 생체외, 및 생체내 형질도입 설정에서 SB100x-기반 시스템보다 더 높은 이식유전자 발현 수준을 부여하였다는 점이었다. 이는 특히 성인 글로빈 수준의 20%를 초과하는 수준에서 γ-글로빈을 필요로 하는 혈색소병증($\beta_0/\beta_0$ 지중해빈혈 및 겸상적혈구병)의 유전자 요법과 관련된다. HDAd-CRISPR + HDAd-글로빈-공여자를 사용하여 생체외 또는 생체내 형질도입된 "반응자" 마우스에서, 이러한 이론적으로 치유 수준이 달성되었다. 이는 SB100x 트랜스포사제 시스템이 γ-글로빈 유전자 추가를 위해 활용된 지중해빈혈 마우스 모델의 이전 연구에 비해 중요한 개선이다(Wang 등, *J Clin Invest*, 129, 598-615, 2019). 이식유전자 발현에 대한 에피게놈 효과는 HSC(Wang 등, *Genome Res*, 17, 1186-1194, 2007, Huser 등, *PLoS Pathog*, 6, e1000985, 2010, van Rensburg 등, *Gene Ther*, 20, 201-214, 2013) 및 AAVS1 유전자이식 마우스에서 개방 염색질 구성을 유지하는 것으로 알려진 AAVS1 유전자좌로의 통합 후 덜 명백할 수 있다. 한편으로, 무작위 SB100x-매개 통합은 침묵을 겪는 영역으로 이식유전자를 배치하는 것을 제외할 수 없다.

[1106] 통합 부위 분석은 HUDEP-2 세포의 시험관내 형질도입 후 거의 100% 표적화된 통합 효능을 시사한다. 생체외 및 생체내 HSC 형질도입 연구에서, 게놈 골수 DNA에 대한 서던 블롯 및 iPCR은 둘 다 골수 HSC에서 효율적인 표적화된 통합을 나타내었다. 예를 들어, 통합 접합부의 iPCR은 마우스의 75%에서 표적화된 통합을 입증하였으며, 이러한 마우스 중 대부분은 표적외 통합이 없었다. 이는 단일 CFU에서 유래된 콜로니의 분석에 의해 추가로 확인되었다. 낮은 빈도에서, 통합은 또한 인 실리코 예측된 CRISPR Cas9 표적외 부위 중 2 개에서 발견되었다. 또한, AAVS1 유전자좌를 보유하는 염색체인 염색체 14에서 통합된 전장 HDAd-공여자 게놈이 발견되었다. HDAd ITR은 DNA 파괴가 일어나기 쉽고 이는 DNA 파괴가 발생하는 게놈 부위로 비효율적인 통합을 초래할 수 있음이 이전에 밝혀졌다(Wang 등, *J Virol*, 79, 10999-11013, 2005, Wang 등, *J Virol*, 80, 11699-11709, 2006). 표적 부위 주위의 CRISPR/Cas9-유도된 원치않은 큰 결실/전좌(7-8 kb)에 대한 최근 연구를 고려하면(Kosicki 등, *Nat Biotechnol*, 36, 765-771, 2018), 표적 부위에서 멀리 떨어진 CRISPR-Cas9 DNA 파괴가 완전한 HDAd 게놈의 통합에 연루될 수 있음이 가능하다. 전반적으로, 큰 결실/전좌에 대한 보고서는 CRISPR/Cas9의 안정성에 문제를 제기한다. 반면에, 동물에서 CRISPR/Cas9-매개 생식계열 편집과 연관된 발달 효과는 지금까지 보고되지 않았기 때문에, 이러한 해로운 염색체 변화가 있는 세포가 발달 동안 선택될 가능성이 있다. 이 가설에 대한 뒷받침은 CRISPR Cas9-편집된 HSC가 이식되었고 HBG1/2 영역에서 9 kb 결실이 시간 경과에 따라 PBMC에서 사라졌다는 최근 NHP 연구에서 기인한다(Humbert 등, 23rd Annual Meeting of the ASGCT, abstract # 974, 2019).

[1107] 이들 연구로부터, AAVS1tg 마우스 모델은 다중 AAVS1 표적 유전자좌의 존재로 인해 CRISPR/Cas9를 수반하는 표적화된 통합 연구에 차선이라는 결론을 내릴 수 있으며, 이 중 일부는 HDAd-공여자 벡터가 있는 상동성 영역을 손실하는 정도로 절두되다. 절두된 AAVS1 유전자좌의 존재는 또한 재배열이 이전에 보고된 바와 같이 AAVS1 유전자이식 마우스에서 발생할 수 있음을 시사한다(Linden 등, *Proc Natl Acad Sci USA*, 93, 7966-7972, 1996).

[1108] **실시예 6. 면역 체크포인트 억제제를 사용한 예방적 생체내 조혈 줄기 세포 유전자 요법은 동계 마우스 종양 모델에서 종양 성장을 역전시킨다.**

[1109] 이 실시예에 함유된 정보 중 적어도 일부는 Li 등(*Cancer Res*. 80(3):549-560, 2020; 2019년 11월 14일 온라인 공개됨)에 공개되었다.

[1110] 암-연관 생식계열 돌연변이에 대한 집단-와이드 테스트는 난소 및 유방 암종의 1/5 초과가 선천적 위험과 연관되어 있음을 규명하였다. 난관난소절개술 및/또는 유방절제술은 현재 고위험 돌연변이가 있는 여성에게 제공되는 유일한 효과적인 옵션이다. 목표는 선천적 돌연변이의 보균자에 대한 면역-예방을 제공하는 오래 지속되는 접근법을 개발하는 것이다. 이 접근법은 초기 단계에서, 종양이 골수에서 조혈 줄기/전구체 세포(HSPC)를 모집하고 종양-촉진 세포로 분화한다는 사실을 활용한다. 생체내에서 HSPC를 유전적으로 변형하기 위해 기술적으로 간단한 기술이 개발되었다. 기술은 HSPC 동원 및 통합 HDAd5/35++ 벡터의 정맥내 주사를 수반한다. GFP-발현 벡터로의 생체내 HSPC 형질도입 및 후속 동계 종양 세포 이식은 종양 침윤 백혈구에서 80% 초과의 GFP-마킹을 나타내었다. 이식유전자의 발현을 제어하기 위해, HSPC가 종양에 의해 모집되고 분화되는 경우에만 활성화되는

miRNA 조절 시스템이 개발되었다. 면역 체크포인트 억제제 αPD-L1-ɣ1을 효과기 유전자로 사용하여 접근법을 테스트하였다. 이식된 마우스 유선 암종(MMC) 종양이 있는 생체내 HSPC-형질도입된 마우스에서, 초기 종양 성장 후, 종양은 퇴행하고 관찰 기간 내내 재발하지 않았다. 퇴행은 T-세포 매개되었다. 항-PD-L1 단클론 항체를 사용한 "통상적인" 치료는 유의한 항-종양 효과가 없었으며, 이는 αPD-L1-ɣ1의 조기 자기-활성화 발현이 MMC 종양에서 면역억제성 환경을 극복할 수 있음을 나타낸다. 접근법의 효능 및 안전성은 예방적 및 치료적 설정 둘 다에서 전형적인 생식 계열 돌연변이가 있는 난소암 모델(ID8 $p53^{-/-}$ $brca2^{-/-}$)에서 추가로 검증되었다.

[1111] ***재료 및 방법***.

[1112] **HDAd5/35++ 벡터:** HDAd-SB는 Richter 등, *Blood*. 128: 2206-2217, 2016에 기재되어 있다. 마우스 αPD-L1-ɣ$_1$ 이식유전자는 Engeland 등, *Mol Ther*. 22: 1949-1959, 2014)에 기재되어 있고; 116 세포에서 HDAd5/35++ 벡터의 생산은 Palmer 등, *Methods in Molecular Biology*, 33-53, 2009에 기재되어 있다. 헬퍼 바이러스 오염 수준은 0.05% 미만인 것으로 밝혀졌다. 역가는 $6\text{-}12x10^{12}$ vp/ml였다. 이 연구에 사용된 모든 HDAd 벡터는 Ad5 섬유 꼬리, Ad35 섬유 샤프트, 및 친화도-향상된 Ad35++ 섬유 놉으로 구성된 키메라 섬유를 함유한다(Wang 등, *J Virol*. 82: 10567-10579, 2008). 모든 HDAd 제제는 다른 곳에 기재된 프라이머를 사용하여 qPCR에 의해 측정된 1010 vp에서 1 개 미만의 카피 야생형 바이러스를 가졌다(Haussler 등, *PLoS One*. 6: e23160, 2011).

[1113] **HDAd-GFP/mgmt 및 HDAd-αPD-L1ɣ$_1$miR423 벡터의 구축**. 단계 1: PGK 프로모터, β-글로빈 3' UTR 및 BGH polyA 단편을 pHCA-HBG-CRISPR/mgmt로부터 PCR 증폭시킨 후(Li 등, Blood. 2018;131: 2915-2928), Gibson 어셈블리(New England Biolabs)에 의해 pBS-Z-Ef1α의 *BstBI* 부위로 삽입하여(Saydaminova 등, *Mol Ther Methods Clin Dev*. 1: 14057, 2015), pBS-PGK-3'UTR을 생성하였다. GFP 코딩 서열을 pHM5-frt-IR-EF1α-mgmt-2a-GFP로부터 PCR 증폭시키고(Wang 등, *Mol Ther Methods Clin Dev*. 8: 52-64, 2018) *EcoRI* 선형화 pBS-PGK-3'UTR과 결찰시켜, pBS-PGK-GFP를 생성하였다. 단계 2: Ef1α-$mgmt^{P140K}$-SV40pA-cHS4 절연체 카세트를 pHM5-T/μLCR-ɣ-글로빈-mgmt-FRT2로부터 증폭시키고(Li 등, *Mol Ther Methods Clin Dev* 9: 142-152, 2018) *PacI*-소화된 pHM5-T/μLCR-ɣ-글로빈-mgmt-FRT2와 결찰시켜, pHM5-FRT-IR-Ef1α-mgmt를 형성하였다. cHS4의 3' 측면에서 *BsrGI* 부위를 하류 사용을 위해 프라이머에 의해 도입하였다. pHM5-FRT-IR-Ef1α-mgmt의 박테리아 플라스미드 백본을 나중 주입 클로닝을 위한 15 bp 상동성 아암(HA)을 함유하는 프라이머(Takara, 캘리포니아주 마운틴뷰 소재)를 사용하여 pBS-Z-Ef1α로부터 백본으로 전환하여, pBS-FRT-IR-Ef1α-mgmt를 생성하였다. 2 개의 Frt-IR 구성요소에 플랭킹하는 2 개의 15 bp HA는 *PacI* 소화 시 노출되어 하기 기재된 변형된 pHCA 작제물과의 재조합을 용이하게 할 수 있다. 그런 다음, PGK-GFP-3'UTR-BGHpA 단편을 단계 1의 pBS-PGK-GFP에서 pBS-FRT-IR-Ef1α-mgmt의 *BsrGI* 부위로 이동시켜, pBS-FRT-IR-GFP/mgmt를 생성하였다. 단계 3: pHCA에서 원래 *PacI* 부위를 2 개의 어닐링된 올리고 서열을 삽입하여 파괴시켰다. 2 개의 HA와 함께 새로운 *PacI* 부위를 *BstBI* 부위에서 생성하였다. 마지막으로, pBS-FRT-IR-GFP/mgmt 및 변형된 pHCA 둘 다의 *PacI* 소화 후, 생성물을 주입 클로닝에 의해 재조합하여, 후속 바이러스 구제에 사용된 pHCA-FRT-IR-GFP/mgmt를 생성하였다. HDAd-αPD-L1ɣ$_1$은 GFP 코딩 서열 대신에, 항-PD-L1-ɣ1 이식유전자가 단계 1 에서 pBS-PGK-3'UTR의 EcoRI에 삽입된 것을 제외하고, 이 실시예에서 다른 곳에 기재된 HDAd-GFP/mgmt와 유사하게 구축하였다. microRNA 조절된 유전자 발현을 위해, 합성된 4 개의 miR423 올리고(정방향(서열번호: 24); 역방향(서열번호: 25))를 어닐링하고 pBS-PGK-3'UTR의 *AvrII-XhoI* 부위에 삽입하여, pBS-PGK-miR423-3'UTR을 생성한 다음, 항-PD-L1-ɣ1 삽입에 사용하였다.

[1114] HDAd-GFP-423은 4 개의 miR423 표적 부위를 HDAd-GFP/mgmt의 3'UTR에 삽입함으로써 유사한 방식으로 구축하였다.

[1115] **유세포 분석**: 세포를 1 % FCS가 보충된 PBS에 $1x10^6$ 개 세포/100 μL로 재현탁하고 얼음 위에서 10 분 동안 FcR 차단 시약(Miltenyi Biotech, 캘리포니아주 오번 소재)과 함께 배양하였다. 다음으로 염색 항체 용액을 $10^6$ 개 세포 당 100 μL로 첨가하고 암실에서 30 분 동안 얼음 위에서 배양하였다. 배양 후, 세포를 FACS 완충액(PBS, 1% FBS)에서 1 회 세척하였다. 2차 염색을 위해 염색 단계를 2차 염색 용액으로 반복하였다. 세척 후, 세포를 FACS 완충액에 재현탁하고 LSRII 유세포 분석기(BD Biosciences, 캘리포니아주 산호세 소재)를 사용하여 분석하였다. 파편은 전방 산란 영역 및 측방 산란 영역 게이트를 사용하여 제외하였다. 그런 다음 단일 세포를 전방 산란 높이 및 전방 산란 폭 게이트를 사용하여 게이팅하였다. 그런 다음 유세포 분석 데이터를 FlowJo(버전 10.0.8, FlowJo, LLC)를 사용하여 분석하였다. 일치된 이소형-대조군을 모든 실험에 포함하였다.

[1116] **면역표현형에 대한 유세포 분석**: 림프구 유세포 분석 패널 8c(CD45-APC/Cy7, 클론 30-F11, cat# 103116; CD3-

APC, 클론 17A2, cat# 100236; CD4-PE/Cy7, 클론 GK1.5, cat# 100422; CD8a-PE, 클론 53-6.7, cat# 100708; CD25-BV421, 클론 PC61, cat# 102043; CD19-BV510, 클론 6D5, cat# 115546; 이들 항체는 모두 BioLegend에서 수득함) 및 골수성 패널 9c(CD45-APC/Cy7, 클론 30-F11, BioLegend, cat# 103116; CD11c-APC, 클론 N418, BioLegend, cat# 117310; F4/80-PE, 클론 C1:A3-1, Cedarlane, cat# CL8940PE; MHCII-BV510, 클론 M5/114.15.2, BioLegend, cat# 107635; Siglec F-PerCP, 클론 1RNM44N, eBioscience, cat# 46-1702-82; Ly6C-BV421, 클론 AL-21, BD Biosciences, cat# 562727; CD11b-PE/Cy7, 클론 M1/70, eBioscience, cat# 25-0112-82; Ly6G-BV605, 클론 1A8, BioLegend, cat# 127639)를 사용하였다. 게이팅 전략은 도 76에 제시된다. LSK (계통$^-$ /Sca-1$^+$/c-Kit$^+$) 세포는 Richter 등, Blood. 2016;128: 2206-2217에서 이전에 특성화되었다. 다음 항체를 또한 사용하였다: 비오틴-접합된 계통 검출 칵테일(Miltenyi Biotec, 샌디에이고 소재, cat# 130-092-613); 항-마우스 LY-6A/E(Sca-1)-PE-Cyanine7(클론 D7, eBioscience, 샌디에이고 소재, cat# 25-5981-82); 항-마우스 CD117(c-Kit)-PE(클론 2B8, eBioscience, 샌디에이고 소재, cat# 12-1171-83); 항-마우스 CD3-APC(클론 17A2, Invitrogen, 매사추세츠주 윌섬 소재, cat# 17-0032-82); 항-마우스 CD19-PE-Cyanine7(클론 eBio1D3, eBioscience, 샌디에이고 소재, cat# 25-0193-82); 항-마우스 Ly-6G(Gr-1)-PE(클론 RB6-8C5, eBioscience, 캘리포니아주 샌디에이고 소재, cat #12-5931-82); 항-인간 CD46-APC(클론 E4.3, BD Pharmingen, 캘리포니아주 샌디에이고 소재, cat# 564253).

[1117] **IFNɣ-유세포 분석**: 50 mL 팔콘 튜브에 부착된 70 μm 세포 여과기를 통해 새로 수확된 비장을 통과시켜 비장세포를 단리하였다. 300 x g에서 10 분 동안 원심분리 후, 세포를 1 mL 1 x BD Pharm Lyse™ 용해 용액(BD Pharmingen, 캘리포니아주 샌디에이고 소재, cat# 555899)에 재현탁하고 30 초 동안 배양함으로써 적혈구를 제거하였다. 20 mL RPMI-1640 배지를 첨가하여 용해 반응을 중지시켰다. 원심분리 및 10% 열-불활성화 FBS, 100 단위/ml 페니실린 및 100 mg/ml 스트렙토마이신을 함유하는 RPMI-1640 배지에 재현탁 후, 수득된 비장세포를 5% $CO_2$를 함유하는 가습 인큐베이터의 96-웰 조직 배양 플레이트에 5 x $10^6$ 개 세포/ml(200 μl/웰)로 배양하였다. 1 x 세포 자극 칵테일 및 단백질 수송 억제제(eBioscience, 샌디에이고 소재, cat# 00-4975-93)가 세포 내에서 IFN-ɣ 생산의 유도 및 축적을 위해 배양 배지에 존재하였다. 12 시간 동안 자극 후, 세포를 수집하고, 먼저 상기 기재된 바와 같이 세포 표면 마커로 염색한 다음, 제조업체의 지침에 따라 IFN-ɣ (BioLegend, 캘리포니아주 샌디에이고 소재, cat# 505842)에 대한 세포내 염색에 적용하였다.

[1118] **Neu-사량체 유세포 분석**: PE-표지된 H-2Dq/RNEU420-429(H-2D(q)PDSLRDLSVF)(서열번호: 290) 사량체를 국립 알레르기 및 전염병 질환 연구소 MHC 사량체 코어 시설(조지아주 애틀랜타 소재)에서 수득하고, 제조업체의 지침에 따라 사용하였다.

[1119] **유세포 분석, FACS, 및 웨스턴 블롯을 위한 종양-침윤 백혈구의 단리**: 종양 부피가 500 $mm^3$에 도달했을 때 마우스를 희생시켰다. 종양을 수확하고, 절단하고 부드럽게 혼합하면서 37℃에서 30 분 동안 5 mL의 RPMI 1640에서 300 U/mL 콜라게나제 I(Sigma-Aldrich, 미주리주 세인트루이스 소재, cat# C0130) 및 1 mg/mL 디스파제 II(Sigma-Aldrich, cat# 4942078001)로 소화시켰다. 소화 후, 2000 U/mL DNase I(Sigma-Aldrich, cat# 260913)을 첨가하여 방출된 DNA를 제거함으로서 점도를 감소시켰다. 소화된 조직을 주사기 플런저를 사용하여 70 μm 세포 여과기를 통해 통과시켜 단일 세포 현탁액을 수득하였다. 이후에, 종양 침윤 백혈구를 마우스 CD45(TIL) 마이크로비드(Miltenyi Biotech, 캘리포니아주 오번 소재, cat# 130-110-618)를 사용하여 단일 세포 현탁액으로부터 정제하였다.

[1120] **면역형광 연구**: 종양 슬라이드를 아세톤/메탄올로 고정시키고(10 분) PBS로 2 회 세척하였다. 슬라이드를 5% 블롯팅 등급 우유(Bio-Rad, 캘리포니아주 에르쿨레스 소재)를 함유하는 PBS를 사용하여 실온에서 20 분 동안 차단한 후 PBS 중 1 차 항체와 함께 실온에서 1 시간 동안 배양하였다. 그런 다음 슬라이드를 PBS로 2 회 세척하고 2차 항체와 함께 실온에서 1 시간 동안 배양한 후 PBS로 3 회 세척하였다. 슬라이드를 PBS로 2 회 세척하고, 형광용 장착 배지(Vector Laboratories 캘리포니아주 벌링게임 소재)로 장착한 다음 형광 현미경을 사용하여 분석하였다. 항-라미닌 다클론(1차) 항체(1:200; #Z0097; Dako, 캘리포니아주 카핀테리아 소재) 및 염소 항-토끼 IgG Alexa Fluor568(2차) 항체(1:200; Molecular Probes, 캘리포니아주 칼즈배드 소재)를 사용하여 라미닌을 검출하였다.

[1121] **마우스 조직의 면역조직화학**: 조직을 10% 포르말린에 고정시키고 헤마톡실린 및 에오신 염색으로 처리하였다. 모든 샘플은 2 명의 숙련된 병리학자에 의해 맹검 방식으로 전형적인 염증 징후에 대해 검사하였다.

[1122] **T-세포 검정**: MMC 세포(Neu-양성) 및 동계 neu/CD46-유전자이식 마우스(Neu-음성)의 비장세포를 20 분 동안 50

mg/m의 최종 농도에서 미토마이신 C로 처리한 다음, 광범위하게 세척하였다. 테스트 동물(HDAd-αPD-L1-γ1 처리) 및 미처리 대조군 동물(미처리)의 비장 세포를 미토마이신 C 처리된 세포와 1:1로 혼합하고 10 U/ml IL-2의 존재 하에 1 일 동안 배양하였다. 대조군 비장세포를 또한 PMA/이오노마이신으로 처리하였다. 상청액의 IFNγ 농도를 IFNγ ELISA(InVitrogen, cat# 88-7214-22)에 의해 측정하였다

[1123] MicroRNA 어레이 분석은 UW Functional Genomics, Proteomics & Metabolomics Facility Core using Affymetrix miRNA 4.0 어레이에 의해 수행하였다

[1124] **실시간 PCR**: 총 RNA를 제조업체의 지침에 따라 TRIzol™(Invitrogen)을 사용하여 종양 침윤 백혈구, PBMC, 비장 세포 및 골수 세포로부터 추출한 다음, 역전사시켜 Qiagen의 QuantiTect 역전사 키트(cat# 205311)를 사용하여 cDNA를 생성하였다. 키트에 제공된 gDNA 제거 시약을 사용하여 잠재적인 게놈 DNA 오염을 제거하였다. Power SYBR Green PCR 마스터 믹스(Applied Biosystems)를 사용하여 비교 실시간 PCR을 수행하였다. 다음 프라이머를 사용하였다: 항-마우스 PDL1 정방향(서열번호: 238), 및 역방향(서열번호: 239); 마우스 PPIA 정방향(서열번호: 240), 및 역방향(서열번호: 241); 마우스 RPL10 정방향(서열번호: 189), 및 역방향(서열번호: 190).

[1125] 마우스 PPIA를 내부 대조군으로 사용하였다. 두번째 내부 대조군 마우스 RPL10이 또한 포함되었고 유사한 결과를 관찰하였다. 결과는 $2^{(-\Delta\Delta Ct)}$ 방법에 따라 계산하고 상응하는 종양 샘플의 cDNA 수준을 100%로 설정하여, 상대적 발현의 백분율로 표시하였다.

[1126] **계통-고갈된($Lin^-$) 골수 세포의 단리**: 계통-수임 세포의 고갈을 위해, 마우스 계통 세포 고갈 키트(Miltenyi Biotec, 캘리포니아주 샌디에이고 소재)를 제조업체의 지침에 따라 사용하였다.

[1127] **콜로니 형성 단위 검정**. 총 2500 개의 $Lin^-$ 세포를 ColonyGEL 1202 마우스 완전 배지(ReachBio, 워싱턴주 시애틀 소재)에 삼중으로 플레이팅하고 5 % $CO_2$ 및 최대 습도에서 37 ℃에서 12 일 동안 배양하였다. 콜로니를 Leica MS 5 해부 현미경(Leica Microsystems)을 사용하여 열거하였다.

[1128] ***세포:*** 마우스 유선 암종(MMC) 세포를 neu/CD46-tg 마우스의 자발적 종양으로부터 확립하였다. MMC 세포 입증을 Neu-특이적 단클론 항체 7.16.4를 사용하여 면역형광에 의해 수행하였다(Knutson 등, Cancer Res. 2004;64: 1146-1151). TC-1 세포는 American Type Culture Collection(ATCC, 버지니아주 머내서스 소재)으로부터 수득하였다. TC-1 세포는 HPV-16 E6 및 E7 단백질을 안정하게 발현하는 불멸성 뮤린 상피 세포이다. C57B1/6-유래 난소암 ID8 $p53^{-/-}$ $brca2^{-/-}$ 세포는 이전에 기재되었다. Walton 등, Cancer Res. 2016;76: 6118-6129. 이 세포주는 ID8 세포에서 *p53* 및 *brca2*의 CRISPR/Cas9 녹아웃에 의해 생성하였다. MMC 및 TC-1 세포를 10% 소태아 혈청, 1 mmol/l 나트륨 피루베이트, 10 mmol/l HEPES, 2 mmol/l L-글루타민, 100 단위/ml 페니실린 및 100 mg/ml 스트렙토마이신이 보충된 RPMI-1640에서 유지하였다. ID8 $p53^{-/-}$ $brca2^{-/-}$ 세포를 4% 소태아 혈청, 100 μg/mL 페니실린, 100 μg/mL 스트렙토마이신, 및 ITS(5 μg/mL 인슐린, 5 μg/mL 트랜스페린, 및 5 ng/mL 나트륨 셀레나이트)가 보충된 DMEM에서 배양하였다. 마이코플라즈마의 부재를 abm(캐나다 리치먼드(BC) 소재)의 PCR 마이코플라즈마 검출 키트를 사용하여 확인하였다. 증폭을 위해 동결보존된 세포를 해동시키고 4 회 계대하였다.

[1129] 난소암 생검은 환자를 식별하는 데 도움이 되는 임의의 기밀 정보 없이 대서양 난소암 연구 컨소시엄(POCRC) 시료 보관소에 의해 제공되었다(Fred Hutchinson Cancer Research Center IRB 프로토콜 # 6289). 생검의 종양 조직을 4 mm 조각으로 절개하고 Strauss 등 *(PLoS One.* 6: e16186, 2011)에 이전에 기재된 바와 같이 콜라게나제/디스파제(Roche)와 함께 37℃에서 2 시간 동안 소화시켰다. 백혈구는 인간 CD45 마이크로비드(Miltenyi Biotech, cat #130-045-801)를 사용하여 자기 활성화된 세포 분류에 의해 단리하였다. 2 개의 고등급 장액성 난소암 생검의 종양-연관 백혈구를 풀링하고 RNA를 LC Sciences, LLC(텟사스주 휴스턴 소재)에 의해 일치하는 PBMC RNA와 비교하여 miRNA-Seq에 의해 분석하였다.

[1130] ***MicroRNA 분석: miRNA-Seq:*** 작은 RNA 서열분석을 이전에 기재된 바와 같이 수행하였다(Valdmanis 등, Nat Med. 2016;22: 557-562.). RNA를 miRNeasy 미니 키트(Qiagen Cat# 1071023)를 사용하여 추출하였다. 샘플 당 1 μg의 RNA를 ATP의 부재 하에 T4 RNA 리가제 1(New England Biosciences cat# M0204)을 사용하여 3' 보편적인 miRNA 클로닝 링커(New England Biosciences cat# S1315)에 결찰시켰다. 결찰된 샘플을 15% 우레아-폴리아크릴아미드 겔 상에서 실행하였다. 작은 RNA에 상응하는 단편(17-28 nt)을 겔로부터 절단하고, 다시 T4 RNA 리가제 1을 사용하여 5' 바코드에 결찰시켰다. 그런 다음 바코드화 샘플을 다중화하고 Illumina MiSeq 기계에서

서열분석하여 Precision Medicine의 UW Center에서 50 bp 단일-단부 판독물을 수득하였다. 바코드 및 어댑터를 서열에서 프리밍한 후 Bowtie 버전 0.12.7을 사용하여 miRBase 상의 마우스 microRNA에 정렬하여, 2 개의 불일치를 허용하였다(Langmead 등, *Genome Biol*. 10: R25, 2009).

[1131] ***작은 RNA에 대한 노던 블롯.*** 이 프로토콜은 Valdmanis 등, Nat Med. 2016;22: 557-562에 기재되어 있다. 다음 $^{32}$P-γ-ATP 표지된 프로브를 사용하였다: miRNA의 경우 423-5p(서열번호: 235); U6의 경우 snRNA(서열번호: 236). 방사성 RNA 분자량 마커는 Ambion으로부터 수득하였다.

[1132] ***웨스턴 블롯:*** 조직 용해물을 SDS-PAGE에 의해 분리하고 블롯을 닭 항-HA-태그-HRP(Abcam, ab1190)와 함께 배양하였다. X-선 필름 상에서 화학발광 검출을 Pierce™ ECL Plus 웨스턴 블롯팅 기질(Thermo Fisher Scientific, cat# 34029)로 처리 후 수행하였다.

[1133] ***αPD-L1-γ$_1$ ELISA:*** 재조합 마우스 PD-L1 단백질(Sino Biological Inc, cat # 50010-M08H)을 2 μg/ml로 사용하여 ELISA 플레이트를 코팅하였다. 테스트 동물의 혈청을 1:10 희석으로 첨가하고 닭 항-HA-태그-HRP(Abcam, ab1190)를 사용하여 *αPD-L1-γ$_1$*을 측정하였다.

[1134] ***동물***: 동물을 수반하는 모든 실험은 통제 기관 심의 위원회 및 IACUC의 승인을 받아 수행되었다.

[1135] ***hCD46-유전자이식 마우스:*** 인간 CD46 게놈 유전자좌를 함유하고 인간과 유사한 수준 및 패턴으로 CD46을 발현하는 C57Bl/6 기반 유전자이식은 Kemper 등 *(Clin Exp Immunol.* 124: 180-189, 2001)에 기재되어 있다. 이들은 C57Bl/6 유래 TC-1 세포를 사용한 이식 연구에 사용되었다. *Neu 유전자이식 마우스:* Neu-tg 마우스(균주명: FVB/N-Tg(MMTVneu)202Mul)는 Jackson Laboratory(메인주 바 하버 소재)에서 수득하였다. 이러한 마우스는 마우스 유선 종양 바이러스 프로모터(게놈 당 1 개의 이식유전자 카피)의 제어 하에 돌연변이되지 않은 비활성화된 래트 neu를 보유한다. 생체내 형질도입 연구를 위해, CD46tg 및 neu-tg 마우스를 교배하여 CD46$^{+/+}$/neu$^{+}$ 마우스를 수득하였다.

[1136] ***생체내 HSPC 형질도입/선택:*** 도 74a 참조.

[1137] ***CD8 세포 고갈:*** CD8-T 세포를 200 μg 래트 항-마우스 CD8 IgG(169.4; ATCC)의 복강내 주사를 사용하여 고갈시켰다. 주사를 3 일마다 반복하여 고갈을 유지하였다.

[1138] ***통계:*** 생체내 데이터의 통계적 유의성을 Kaplan-Meier 생존 곡선 및 로그-순위 테스트(GraphPad Prism Version 4)에 의해 분석하였다. 시험관내 데이터의 통계적 유의성은 양측 스튜던트 t-검정(Microsoft Excel)에 의해 계산하였다. 0.05 초과의 P 값은 통계적으로 유의하지 않은 것(n.s.)으로 간주된다.

[1139] ***결과 및 논의.***

[1140] 50 세 이전에 유방암 진단 또는 모든 연령에서 난소암 진단을 받은 적어도 한 명의 직계 가족이 있는 여성은 이제 유전자 검사에 적용된다. 표적화된 포획 및 대량 병렬 게놈 서열분석을 사용하여, 생식 계열 돌연변이를 검출하고 암 발병 위험을 예측하는 일련의 다중 유전자 검사를 확립하였다. 이러한 검사 플랫폼 중에는 BROCA가 있다(Walsh 등, *Proc Natl Acad Sci USA*. 108: 18032-18037, 2011, Shirts 등, *Genet Med*. 18: 974-981, 2016). BROCA를 사용하여, 난소 및 유방 암종의 1/5 초과가 선천적 위험과 연관되어 있음을 규명하였다(Tung 등, *Cancer*. 121: 25-33, 2015). 문제는 고위험 보균자의 예방을 위한 현재 옵션이 지속적으로 개선되는 유전적 진단에 뒤쳐진다는 것이다. 불임, 심혈관 질환, 골다공증, 갱년기 증상, 및 심리적 효과를 포함하는 예방적 난관난소절개술 및 유방절제술의 부작용은 여성의 일생에 걸쳐 예상된다. CA125 및 HE4와 같은 혈청 마커의 사용은 난소암 사망률의 상당한 감소를 나타내지 않았다(Jacobs 등, *Lancet*. 387: 945-95, 2016). Her2/neu, HIF1α, 또는 MUC1과 같은 종양-연관 항원에 대한 예방적 백신은 모든 종양 세포 상에서 이들 항원의 존재에 의존하고, 항원-손실 돌연변이체의 발달에 의해 애를 먹고 있다(Knutson 등, *Cancer Res*. 64:1146-1151, 2004).

[1141] 목표는 종양 재발에 대한 위험이 높은 환자, 및 궁극적으로, 암에 걸리기 쉬운 선천적 돌연변이의 보균자에서 암의 면역-예방을 허용하는 오래 지속되고 기술적으로 간단한 접근법을 개발하는 것이다. 종양 진행 동안, 악성 세포는 HSPC를 활성화하고 동원하는 다수의 특이적 케모카인을 분비하여 혈액 순환에 들어가고 종양-지지 세포로 분화하는 경우 종양으로 국소화하게 한다(Hanahan 등, *Cell*. 144: 646-674, 2011, Mantovani 등, *Trends Immunol*. 23: 549- 555, 2002). HSPC-유래 골수성 및 림프성 세포는 예를 들어 장액성 관 상피내 암종(STIC)에서 암 발병 초기 단계에 존재한다(Okla 등, *Front Immunol*. 10: 691, 2019; Colvin, *Front Oncol*. 4: 137, 2014; Baert 등, *Front Immunol*. 10: 1273, 2019). Sarkar 등, Genes *Dev*. 31: 1109-1121, 2017. 이 접근법

은 조혈 줄기 세포의 유전적 변형에 기반한다. 이들 세포는 자기 재생할 수 있기 때문에, 1-회 개입은 평생 치료 효과가 있어야 한다. 백혈구분리반출술, 골수절제술 및 이식 없이 HSPC로의 생체내 유전자 전달을 가능하게 만드는 최소로 침습성이고 비용-효율적인 기술이 개발되었다(Richter 등, *Blood*. 128: 2206-2217, 2016, Wang 등, *J Clin Invest*. 129: 598-615, 2019). 이 접근법의 핵심 아이디어는 G-CSF/AMD3100을 사용하여 골수에서 HSPC를 동원하고, 주변부에서 많은 수로 순환하는 동안, 정맥내로 주사된 HSPC-향성 헬퍼-의존적 아데노바이러스 HDAd5/35++ 유전자 전달 벡터 시스템으로 형질도입하는 것이다. 이들 벡터는 원시 조혈 줄기 세포 상에서 발현되는 수용체인 CD46을 사용한다. 형질도입된 세포는 장기간 지속되는 경우 골수로 돌아간다. 이 연구에서 사용된 HDAd5/35++ 벡터 시스템의 새로운 특징은 다음을 포함한다: (*i*) i.v. 주사(간 포함) 후 비-조혈 조직의 감염을 피하면서 원시 HSPC의 효율적인 형질도입을 허용하는 CD46-친화도 향상된 섬유, (*ii*) 세포 당 1 내지 2 개의 통합 벡터 카피가 있는 유전자에 대한 선호도 없이 세포 요인과 무관하게 기능하고 무작위 이식유전자 통합을 매개하는 SB100X 트랜스포사제-기반 통합 시스템(도 74a), 및 (*iii*) 저용량 $O^6$BG/BCNU를 사용한 단기간 치료에 의해 형질도입된 원시 HSPC의 풀에 영향을 미치지 않으면서 자손 세포의 선택적 생존 및 확장을 매개하는 $MGMT^{P140K}$ 발현 카세트(Wang 등, *Mol Ther Methods Clin Dev*. 8: 52-64, 2018). 혈색소병증에 대한 마우스 모델에서 생체내 HSPC 유전자 요법 방법의 효능 및 안전성이 최근에 입증되었다(Wang 등, *J Clin Invest*. 129: 598-615, 2019, Li 등, *Blood*. 131: 2915-2928, 2018). 여기서, 이 접근법은 암 성장을 예방하는 데 사용된다.

[1142] ***생체내 HSPC 형질도입 후 종양-침윤 백혈구에서 GFP 발현.*** 동계 종양이 있는 2 개의 인간 CD46 유전자이식 마우스 모델을 이용하였다. (CD46은 HDAd5/35++ 벡터를 사용한 HSPC 형질도입이 필요하다). 첫번째 모델은 유방 조직에서 마우스 유선 종양 바이러스 프로모터로부터의 래트 neu를 과발현하는 인간 CD46/래트 neu-유전자이식 마우스를 포함하였다. Neu-tg 마우스는 Treg에 의존하고 유방암 환자에서 관찰된 것과 유사한 Neu에 대한 활성 면역 허용을 발달시킨다(Knuston 등, *J Immunol*. 177: 84-91, 2006). 마우스 유선 암종 세포(MMC)는 자발적 neu/CD46-유전자이식 마우스 종양으로부터 유래된 Neu-양성 유방암 세포주이다(도 75). HSPC를 neu/CD46 tg 마우스에 동원하고 통합 GFP-발현 HDAd5/35++ 벡터(도 74a)를 주사하였다. 이전 연구와 유사하게(Wang 등, *Mol Ther Methods Clin Dev*. 8: 52-64, 2018), $O^6$BG/BCNU를 사용한 3 차례의 저용량 처리는 PBMC의 80%에서 안정된 GFP 발현을 초래하였다(도 74). 생체내 HSPC 형질도입 후 17 주차에, 동계 MMC 세포를 유선 지방 패드에 이식하고 종양 성장을 모니터링하였다. 종양이 700 mm의 부피에 도달했을 때(Palmer 등, Methods in Molecular Biology, 2009:33-53), 동물을 희생시키고 GFP 발현을 분석하였다. 골수 세포, 비장세포, PBMC, 및 종양-침윤 백혈구의 80%는 GFP를 발현하였다(도 74b). 종양에서, $GFP^+$ 세포는 종양 기질에서 우세하게 발견되었다(도 74c). 면역표현형은 $GFP^+$ 종양-침윤 세포가 림프구(우세하게는 Treg), 호중구, DC/MDSC, 및 대식세포임을 나타내었다(도 74d, 76). 이 패턴은 말초 혈액(도 74d), 골수 및 비장(도 77)에서 $GFP^+$ 세포와 상이하였으며, 이는 종양이 HSPC를 전문화된 전종양 세포로 활성적으로 분화시킴을 나타낸다. 종양에 대한 생체내 형질도입된 HSPC의 효율적인 모집은 CD46tg 마우스 및 HPV16 E6/E7-양성 마우스 폐암 세포주인 TC-1 세포로 이루어진 두번째 모델에서 추가로 확인되었다(도 78a-78c).

[1143] ***종양-침윤 백혈구에서 miRNA-조절된 이식유전자 발현.*** 도 74b 및 도 78c는 GFP(보편적으로 활성인 EF1α 프로모터의 제어 하에)가 종양-침윤 백혈구 뿐만 아니라 골수, 비장, PBMC, 및 상주 대식세포를 포함하는 다른 조직에서도 발현됨을 예시한다. 자가-면역 반응을 최소화하기 위해, 요법 접근법은 치료 이식유전자가 (*i*) 종양에서 우선적으로 발현되고, (*ii*) 종양이 발달하기 시작할 때만 자동적으로 활성화되고, (*iii*) 종양이 사라질 때 중단되는 것을 필요로 한다. 이러한 요건은 miRNA 조절을 통해 충족될 수 있다. 조혈 동안, miRNA 프로파일은 분화 단계 및 세포 계통에 따라 변한다(Chen 등, Science. 2004;303: 83-86). 종양-연관 골수성 세포는 별개의 mRNA 및 miRNA 발현 프로파일을 갖는다(Thorsson 등, *Immunity*. 48: 812-830 e814, 2018). 마지막으로, 인간의 상이한 종양 유형에서 발견된 골수성 및 림프성 세포에서 miRNA의 보존 정도가 높다 (Thorsson 등, *Immunity*. 48: 812-830 e814, 2018). 이식유전자 발현의 miRNA 조절 원리는 도 79a에 제시된다. 생체내 HSPC-형질도입된 마우스 모델을 사용하여, 골수, 비장, PBMC, 및 종양으로부터의 GFP+/CD45+ 세포를 분류하고 (도 74b, 78c) 이들의 miRNA 발현 프로파일을 분석하였다. 목표는 골수, 혈액 및 비장 세포에서 높은 수준으로 발현되지만, 종양-연관 백혈구에 부재하는 miRNA를 찾는 것이었다. 총 RNA(5 마리의 마우스로부터 풀링됨)를 차세대 miRNA 서열분석에 적용하였다(도 79b, 79c). 상기 기준을 충족하는 일련의 miRNA를 식별하였다. neu/CD46tg-MMC(도 79b) 및 CD46tg-TC-1 모델(도 79c) 둘 다에서 목록의 상위에 있는 miRNA인 miR423-5p에 초점을 맞췄다. miR-423-5p는 인간 및 마우스 사이에 보존되어 있고 따라서 클리닉에 대한 접근법의 추가 개발에

사용될 수 있다. MMC 및 TC-1 종양이 있는 생체내 형질도입된 마우스로부터의 GFP+ 분획에서 miRNA-423-5p의 발현 프로파일을 microRNA 어레이(제시되지 않음) 및 노던 블롯 분석(도 81)에 의해 검증하였다.

[1144] miR-423-5p 조절이 또한 인간에서 사용될 수 있는지를 평가하기 위해, 일련의 인간 조직에 걸쳐 microRNA를 평가한 공개된 데이터세트에서 miR-423-5p의 수준을 조사하였다. Ludwig 등, Nucleic Acids Res. 2016;44: 3865-3877. miR-423-5p는 발현된 microRNA의 상위 20%에 있고 골수 및 비장을 포함하는 조직에 걸쳐 균일하게 분포하는 것으로 밝혀졌다(도 82a). 일치하는 PBMC 및 종양 생검을 2 명의 고등급 장액성 난소암 환자로부터 수득하였다. miRNA-Seq를 일치하는 PBMC로부터의 RNA에 비해 종양-침윤($CD45^+$) 백혈구로부터의 RNA에서 수행하였다(도 82b). 이 분석은 PBMC에서 miR423-5p의 높은 발현 수준 및 종양-침윤 백혈구에서 낮은 발현 수준을 확인하였다. 이러한 데이터는 마우스에서 관찰된 결과가 인간 연구로 번역될 강한 잠재력을 가짐을 입증한다.

[1145] **HSPC에 대한 HDAd-매개 miR-423 표적 부위 발현의 효과**. miRNA-423-5p는 모든 정상 조직에서 발현되고 따라서 유전자 발현의 조절에 수반될 가능성이 가장 높다. "mirtarbase"에서 miR-423-5p에 대한 표적 mRNA의 검색은 사이클린-의존적 키나제 억제제 1A(CDKN1A) mRNA를 1차 표적으로 식별하였다(mirtarbase.mbc.nctu.edu.tw/php/detail.php?mirtid= MIRT000589#target에서 온라인으로 이용가능). 다른 표적 mRNA는 전사 신장 인자 A 유사 1(TCEAL1), bcl2 유사 11(bcl2L11), 및 증식-연관 2G4(PA2G4)를 포함한다. HDAd 벡터로부터 추가된 miR-423-5p 표적 부위의 발현이 CDKN1A의 발현에 영향을 미치는지를 평가하기 위해, GFP 함유 mRNA에 연결된 표적 부위가 있거나 없는 2 개의 HDAd-GFP 벡터를 구축하였다(도 80a). 마우스 및 인간 HSPC, 즉, miR-423-5p 발현 수준이 높은 세포 유형을 대부분의 세포의 형질도입을 초래할 MOI에서 감염시키고(Li 등, *Mol Ther*. 27(12):2195-2212, 2019) 3 일 후 웨스턴 블롯에 의해 CDKN1A 단백질 수준을 분석하였다(도 80b). 두 세포 유형에서 2 개의 HDAd 벡터 사이의 유의한 차이는 발견되지 않았다. 또한, miR-423-5p 표적 부위 과발현의 해로운 효과가 전구체 콜로니 검정에서 관찰되지 않았다(도 80c). 본원의 다른 곳에 요약된 바와 같이, miR423-5p 표적 부위를 함유하는 요법 벡터를 사용한 생체내 HSPC 형질도입은 조혈에서 이상을 유발하지 않았다. 종합하면, 이는 개시된 miR-423-5p-기반 조절 시스템이 HSPC에서 안전함을 시사한다.

[1146] ***면역-예방 연구.*** 유전성 유방암 및 난소암에서, 유전적 변이체는 DNA 복구 메커니즘을 방해하여 더 높은 돌연변이 부담 및 신생항원 존재를 초래한다. 이는 종종 비정상적인 카피 수 및 낮은 면역원성을 특징으로 하는 비-유전성 유방암 및 난소암보다 종양을 면역요법으로 더 잘 처리할 수 있게 만든다(Thorsson 등, Immunity. 2018;48: 812-830 e814). 여기서, 체크포인트 억제제 αPD-L1-γ1은 면역치료 이식유전자로서 선택되었다. 이전에, 바이러스 유전자 전달 후 종양내 αPD-L1-γ1 발현은 종양 성장 감쇠를 초래하는 것으로 제시되었다(Engeland 등, *Mol Ther*. 22: 1949-1959, 2014, Reul 등, *Front Oncol*. 9: 52, 2019). MMC 세포 배양에서, 강한 PD-L1 발현이 관찰되었으며(도 83a), 이는 MMC 종양을 αPD-L1-γ1 요법에 민감하게 만들어야 한다. miR423-5p 표적 부위의 4 개의 카피를 αPD-L1-γ1 유전자에 연결된 글로빈 3' UTR로 통합하였다(도 83b). 실험 체계는 도 74a에 제시된 것과 동일하였다. 대조군 HDAd-GFP/mgmt 벡터로 생체내 형질도입된 마우스에서, 이식된 MMC 종양은 빠르게 성장하여 종양 세포 이식 후 35 일에 종료점 부피에 도달하였다(도 83c, 왼쪽 패널). αPD-L1-γ1 모델에서, 초기 종양 성장 후, 7 개 종양 중 6 개가 퇴행되고 관찰 기간(100 일) 내에 재발하지 않았다. 처리된 마우스는 첫번째 주사 후 11 주 차에 주어진 또 다른 접종된 MMC 세포를 거부하였다. 항-CD8 mAb 주사에 의한 CD8 세포의 고갈은 치료 효과를 무효화하였다. 항-종양 T-세포 반응을 관찰 기간의 종료 시(100 일) 측정하였다. 유세포 분석에 의한 비장 세포 분석은 상당히 더 높은 백분율의 인터페론-γ (IFNγ)-생산 CD4 및 CD8 세포 뿐만 아니라 Neu-사량체로 양성으로 염색된 더 높은 빈도의 CD8 세포를 나타내었다(도 83d). HDAd-αPD-L1-γ1-처리된 동물로부터의 비장세포는 (Neu-양성) MMC 세포로 자극 시, Neu-음성 세포와 비교하여 30-배 더 큰 IFNγ 분비를 나타내었다(도 83e). 예상된 바와 같이, 미처리 CD46/neu-tg 마우스는 Neu-특이적 T-세포를 보유하였지만, 종양에서 면역억제성 T-세포의 존재로 인해 종양 성장을 제어할 수 없었다(Knutson 등, J Immunol. 2006;177: 84-91).

[1147] **MMC/neu-유전자이식 마우스 모델에서 αPD-L1-$\gamma_1$ 발현의 동역학 및 특이성**. HDAd-αPDL1γ1miR423-처리된 동물의 별개의 그룹에서, 종양이 줄어들기 시작하기 전에 이식 후 17 일차에 수확하였다. 이들 종양(300-400 $mm^3$)에서, 웨스턴 블롯 분석 8에 의해 PBMC, 골수, 및 비장에서보다 종양에서 10-배 더 높은 수준의 αPD-L1-$\gamma_1$을 관찰하였다(도 84a). 종양-침윤 백혈구에서 *αPD-L1-*$\gamma_1$ mRNA의 우선적 발현을 qRT-PCR에 의해 확인하였다(도 84b). 이 발현 패턴은 miR-423-조절이 종양-침윤 골수성 및 림프성 세포 이외의 HSPC 자손에서 *αPD-L1-*$\gamma_1$ 발현을 억제하였음을 시사하였다. 혈청 *αPD-L1-*$\gamma_1$은 MMC 세포 주사 후 검출가능하게 되고 종양이 사라지면

감소하였으며, 이는 *αPD-L1-γ*$_1$ 발현의 기능적 자기조절을 나타내며(도 84b), 즉 이식유전자 발현은 HSPG가 종양-연관 백혈구로 분화된 후에만 시작됨을 나타낸다. MMC 세포 주사 후 2 주부터 시작하여, 털 변색 및 조직의 염증성 침윤에 의해 반영된 자가-면역 반응이 관찰되었다(도 87, 도 87a에서는 마우스를 나타내고 도 87b에서는 신장, 간, 및 폐에서의 샘플을 나타냄). 중요하게는, 종양이 사라지고 4 주 후 희생된 동물에서, 모든 기관의 조직학은 정상으로 돌아왔다. 이러한 관찰은 *αPD-L1-γ*$_1$이 발현되고 혈류로 방출되는 한, 일시적 자가-면역 반응(neu-발현 조직/세포 유형에 대한 것일 가능성이 가장 큼)이 발생할 수 있음을 나타낸다. 특히, miR-423-5p 표적 부위가 없는 HDAd *αPD-L1-γ*$_1$ 벡터를 사용한 연구는 마지막 $O^6$BG/BCNU 처리 2 주 후 처리된 동물에서 발생하는 20% 초과의 체중 손실로 인해 종료하였다. 이는 조절된 *αPD-L1-γ*$_1$ 발현에 대한 필요성을 강조한다. 관찰된 자가-면역 반응은 *αPD-L1-γ*$_1$을 종양에 물리적으로 테더링하거나 또는 세포내 면역조절 효과기(예를 들어, 종양-촉진 백혈구를 종양-사멸 세포로 재분극화하는 miRNA)를 사용함으로써 최소화될 수 있다. 추가로, 벡터는 또한 항체(Erbitux)-의존적 세포독성에 의해 형질도입된 모든 세포의 파괴를 허용하는 절두된 EGFR 수용체를 함유할 수 있다(Wang 등, Blood. 2011;118: 1255-1263).

[1148] 생체내 HSPC *αPD-L1-γ*$_1$ 유전자 요법 접근법의 효능은 neu-tg/MMC 모델에서, 다른 면역요법 접근법이 종양 재발을 예방하지 못했다는 점을 고려하면 주목할만하다(Knutson 등, *Cancer Res*. 64:1146-1151, 2004, Burgents 등, *J Immunother*. 33: 482-491, 2010). 이 맥락에서, 항-마우스 PD-L1 단클론 항체의 4 차례의 복강내 주사는 종양 성장에 유의한 영향을 미치지 않았다(도 88a, 88b). 이러한 데이터는 종양 발달 동안 초기에(HSPC 자손 세포가 종양을 침투하자마자) *αPD-L1-γ*$_1$ 1의 종양내 발현이 억제인자 및 효과기 면역 세포 사이의 균형을 종양 제거 쪽으로 기울일 수 있음을 나타낸다.

[1149] **p53 및 brca2 돌연변이가 있는 난소암 모델에서 면역예방 및 요법 연구**. C57Bl/6 유래된 뮤린 난소암 ID8 세포는 전형적인 암-연관 생식 계열 돌연변이(brca1, brca2, p53, Nf1, Rb1, Pten...)를 함유하지 않고 복강내 주사 후 종양을 잘 형성하지 않는다. Walton 등, *Cancer Res*. 76: 6118-6129, 2016. 종양-억제인자 유전자의 CRISPR/Cas9 녹아웃에 의해 생성된 더 새롭게 개선된 ID8-유래 모델에서, 이러한 결핍을 해결한다. Walton 등, Cancer Res. 2016;76: 6118-6129; Walton 등, Sci Rep. 2017;7: 16827. 이러한 모델 중에는 ID8-p53$^{-/-}$-brca2$^{-/-}$ 세포가 있다. 2x10$^6$ 개의 ID8-p53$^{-/-}$-brca2$^{-/-}$ 세포를 CD46-유전자이식 마우스로 복강내 주사하면 6-8 주 내에 종양 성장 및 복수의 발병(또는 사망)을 초래하였다(도 84c 및 85a). 복강내 종양은 장간막을 따라 광범위하게 퍼져 다른 장기(비장, 간, 림프절)을 침습하였다. 복강내 ID8-p53$^{-/-}$-brca2$^{-/-}$ 종양에서 종양-침윤 백혈구의 면역표현형은 Treg 뿐만 아니라 면역억제성 DC/MDSC 및 TAM의 의 명백한 존재를 나타내었다(도 85b). 종양 침윤 T-세포(TIL), 대식세포(TAM), 및 호중구(TAN)를 복막의 ID8 p53$^{-/-}$ brca2$^{-/-}$ 종양으로부터 단리하고 miRNA-423-5p 수준을 노던 블롯으로 분석하였다. MMC 및 TC-1 모델에서 관찰된 바와 같이, miR-423-5p는 골수 단핵 세포에서 발현되지만 TIL, TAN, 및 TAM)을 포함하는 종양-침윤 백혈구에서 검출가능하지 않았으며, 이는 모든 3 가지 세포 유형이 종양에 의해 특이적으로 재프로그램화되었음을 나타낸다(도 85c).

[1150] 먼저, ID8-Trp53$^{-/-}$-brca2$^{-/-}$ 모델을 예방적 설정에 사용하였다(도 85d). HDAd-αPDL1γ1miR423 + HDAd-SB 또는 HAd-GFP-miR423 + HDAd-SB(대조군)를 사용한 HSPC 생체내 형질도입/선택 후, ID8-p53$^{-/-}$brca2$^{-/-}$ 세포를 복강내로 주사하고 혈청 αPDL1γ1 수준 및 이환율 및 복수의 발병을 모니터링하였다. 모든 대조군 마우스는 생체내 형질도입 후 70 일에 종료점에 도달하였지만, 모니터링 기간이 종료될 때(종양 세포 접종 후 11 주) HDAd-αPDL1γ$_1$miR423 + HDAd-SB 처리된 동물의 100%가 생존하였다(도 85e). 6 주차(세포 주사 후)에 상승된 혈청 αPDL1γ1 수준은 종양이 성장하였고 혈청 αPDL1γ1 발현이 활성화되었음을 시사한다(도 85f). 11 주차에, 혈청 αPDL1γ1은 배경 수준으로 돌아갔으며 이는 종양이 제거되었음을 나타낸다. 이 연구에서, 자가-면역 반응의 징후(예를 들어, 털 변색)는 관찰되지 않았으며, 종양 및 정상 조직 사이에 공유되는 항원(예를 들어, Neu)의 부재 때문일 가능성이 가장 크다. 기재된 접근법의 안전성을 평가하는 맥락에서, 또한 HDAd-αPDL1γ$_1$miR423을 사용한 생체내 HSPC 형질도입이 조혈에서 이상을 유발하지 않았음을 제시하였다(도 88c, 88d). 동계 종양 세포가 이식된 마우스에서, PBMC에서 GFP 양성 세포의 백분율을 표시된 시점에서 측정하고, GFP 양성 세포를 miRNAseq에 대해 수확하였다(도 88e). 결과는 관심 발현 패턴을 갖는 miRNA를 식별하였다(도 58e). 종양(TIL), PBMC, 골수, 및 비장에서 PDL1에 대한 웨스턴 블롯은 도 88f에서 mRNA에 대한 발현으로 제시되고 정량화

된다. 종양 이식 전 및 이식 후 표시된 시점에서 혈청 αPDLA ELISA $OD_{450}$이 또한 도 88f에 제시된다. 개략도는 도 88g 및 88h에 제시된다.

[1151] 예방 접근법은 종양-발달의 극초기 단계에서 자동으로 시작한다는 이점을 갖지만, 고위험 돌연변이를 보유하는 건강한 여성에서 즉시 적용하는 것은 임상 번역에서 규제 장애물에 직면할 가능성이 있을 것이다. 따라서, 보다 현실적인 목표는 이 접근법을 사용하여 1차 요법 후 암 재발을 예방하는 것이다. 이 경우에, 생체내 HSPC 선택은 환자의 화학요법 치료에 직접적으로 포함될 수 있다. 도 86a는 임상 설정에서, 생체내 HSC 형질도입이 어떻게 외과적 종양 감량수술 후 시작하거나, 또는, 수술이 옵션이 아닌 경우 화학요법과 함께 시작함을 나타낸다. $O^6$BG/BCNU 생체내 선택은 화학요법과 조합될 수 있다. 생체내 HSPC 형질도입/선택의 결과로서, 무장된 HSPC는 암이 재발할 때까지 휴면기에 놓여 HSPC 분화 및 효과기 유전자 발현의 활성화를 촉발할 것이다. 이 설정은 또한 종양-특이적 신생-항원 및 종양의 면역-표현형이 외과적 생검의 분석으로부터 알려질 것이라는 이점을 가지며, 이는 적절한 면역요법 효과기 유전자를 선택하게 할 수 있다. 한편으로, "완전히 발달한" 암 특징을 갖는 암의 재발을 예방하는 것(Hanahan 등, Cell. 2011;144: 646-674)은 발달 초기 단계에서 종양을 표적화하는 것보다 더 어렵다.

[1152] 이러한 "치료적" 설정을 시뮬레이션하기 위해, CD46-유전자이식 마우스에게 먼저 ID8-$Trp53^{-/-}$-$brca2^{-/-}$ 세포를 주사한 후 2 주 후에 생체내 HSPC 형질도입/선택을 하였다(도 86b). 대조군 설정의 모든 마우스(HDAd-GFP-miR423 + HDAd-SB 형질도입된 HSPC)는 종양 세포 주사 후 12 주에 종료점에 도달하였지만, αPDL1-γ1 발현 벡터로 처리된 모든 마우스는 15 주차에 건강하였다(도 86c). 예방 연구에서와 같이, 11 주차에 상승된 혈청 αPDL1-γ1 수준은 종양이 초기에 성장하지만 자기-조절된 αPDL1-γ1 메커니즘이 활성화되면 사라졌음을 시사한다(도 86d). 이러한 데이터는 기재된 접근법이 수술/1차 화학요법 후 암 재발을 예방할 수 있음을 나타낸다.

[1153] TC-1(마우스 폐암) 종양(도 78a-81), MMC(마우스 유방암) 종양(도 79a-79c 및 81), 및 ID8-$p53^{-/-}$/$brca2^{-/-}$(마우스 난소암 종양)(도 85c)에 존재하는 종양-침윤 백혈구에 대한 mRNA 프로파일링/노던 블롯 분석을 수행하였다. miR423-5p가 검출불가능하지만, 정상 조혈 구획에서 높은 수준으로 존재하는 3 가지 종양 유형 모두에서 발견되었다. 인간 난소암 생검으로부터의 데이터와 함께(도 82a, 82b), 이는 miR423-5p-기반 시스템이 효과기 유전자 발현의 조절을 위해 종에 걸쳐 상이한 종양 유형에 광범위하게 사용될 수 있음을 나타내다.

[1154] 현재 암 발병의 고위험과 연관된 생식 계열 돌연변이가 있는 여성에게 제공되는 제한된 예방적 옵션, 및 집단-와이드 스크리닝으로 인해 이러한 보균자 수의 증가를 고려하면, 이 생체내 HSPC 유전자 요법 접근법은 주요 의학적 문제를 해결하는 유망한 접근법이다.

[1155] **실시예 7. 분비된 치료 단백질의 고수준 생산을 위한 공장으로서 적혈구 세포를 사용한 생체내 HSC 유전자 요법.**

[1156] 이 실시예는 생체내 HSC 형질도입/선택 후 적혈구 세포에서 비-적혈구 단백질의 발현 및 성숙 적혈구에서 발현된 단백질의 저장을 나타낸다. 이 시스템을 사용하여 단일 정맥내 개입 후 장기간 치료 교정을 제공할 수 있다. 이 실시예에 함유된 정보의 적어도 일부는 Wang 등 (*Blood Adv* 3(19): 2883-2894, 2019; e-pub October 4, 2019)에 공개되었다.

[1157] 인간 성인에서 초 당 240 만개의 새로운 적혈구가 생성된다. 인체 세포의 거의 1/4는 적혈구이다(Pierige 등, *Adv Drug Deliv Rev*. 60(2):286-295, 2008). 적혈구생성 과정에서 HSC는 공통 골수성 전구체 및 전적혈구모세포를 통해 정염성 적혈구모세포(Wright의 염색에 기반)로 분화한다. 이 단계에서, 핵은 배출되고, 세포는 골수에서 빠져나와 망상적혈구로서 순환계로 들어간다. 성인에서 순환하는 적혈구의 0.5% 내지 2.5%($1x10^5$/μl) 및 유아의 2% 내지 6%가 망상적혈구이다. 망상적혈구는 여전히 mRNA로부터 헤모글로빈을 생성할 수 있다. 1 내지 2 일 후, 이들은 궁극적으로 모든 세포기관을 손실하고 성숙 적혈구가 되며, 더이상 단백질 생합성을 할 수 없다. 열성적인 적혈구 전구체에서 적혈구까지 분화는 7 일이 걸린다. 적혈구의 수명은 120 일이다. 오래되고 죽어가는 적혈구는 비장의 식세포 시스템에 의해 제거된다.

[1158] HSC가 열성적인 적혈구 세포로 분화되면, 대량의 α 및 β 글로빈 쇄가 생성된 다음 이후에 적혈구에서 사량체성 헤모글로빈으로 저장된다. 건강한 개인은 혈액 100 ml 당 12 내지 20 그램의 헤모글로빈을 가지며 적혈구 중량의 95%는 헤모글로빈이다(세포 당 $270x10^6$ 개의 Hb 분자). 이러한 효율적인 생합성에 대한 기초는 높은 수준의 전사 및 효율적으로 번역되는 안정된 mRNA를 허용하는 강한 적혈구 특이적 유전자좌 제어 영역(LCR)이다.

[1159] 적혈구생성의 엄청난 속도 및 효능 및 헤모글로빈 생산의 강력한 기구를 사용하여 적혈구 전구체 세포로부터 비-적혈구 분비된 단백질을 생산하였다(전적혈구모세포에서 망상적혈구까지의 분화 단계 포함). 이식유전자는 미니-β-글로빈 LCR의 제어 하에 있고 mRNA 안정화를 위해 β-글로빈 유전자의 5'UTR 영역을 함유하였다. 치료 단백질의 장기간 평생 생산을 허용하기 위해, 유전자 전달 벡터는 원시 HSC를 표적화하였다. 생체내 HSC 형질도입 접근법은 골수에서 말초 혈류로의 HSC의 G-CSF/AMD3100-촉발된 동원 및 통합 헬퍼-의존적 아데노바이러스 벡터 시스템의 정맥내 주사를 수반한다. 이식유전자 통합은 과활성 잠자는 미녀 트랜스포사제(SB100x)를 사용하여 (무작위 패턴으로) 달성되지만, 특정 구현예에서, 상동성 지시된 복구를 통해 달성될 수 있다.

[1160] 적혈구 세포가 혈액 순환으로 분비되는 치료 단백질의 높은 수준의 생산에 사용될 수 있다는 증거 또는 원리로서, 본원에서는 응고 인자 VIII의 생체공학적 형태에 초점을 맞췄다. 연구 결과는 A형 혈우병 치료와 관련이 있다. 최근에, 임상 발전은 B형 혈우병에 대한 간-지시된 인자 IX 유전자 전달을 위한 재조합 아데노-연관 바이러스(rAAV)-기반 유전자 요법을 사용하여 이루어졌다(High 등, Methods Mol Biol. 2011;807:429-457). 전임상 연구는 또한 동물 모델에서 FVIII 발현 rAAV 벡터로 A형 혈우병을 치료하는 실현가능성을 입증하였다(Brown 등, *Mol Ther Methods Clin Dev*. 1:14036, 2014, Callan 등, *PLoS One*. 11(3):e0151800, 2016, Greig 등, *Hum Gene Ther*. 28(5):392-402, 2017). 그러나, 간-지시된 rAAV A형 혈우병 유전자 요법의 광범위한 적용은 여러 장애물에 직면할 수 있다: (*i*) 간세포에서 rAAV 게놈의 대부분의 에피솜 속성 및 구체적으로 어린이에서 세포 분열로 인한 이들의 손실. (*ii*) rAAV 벡터 생산에 대한 고비용, (*iii*) 유전자 침묵 또는 유전독성을 방지하기 위해 종종 요구되는 큰 전사적 조절 요소를 수용할 수 없는 rAAV의 제한된 패키징 용량(Grieger 등, *J Virol*. 79(15):9933-9944, 2005, Chandler 등, *J Clin Invest*. 125(2):870-880, 2015), 및 (*iv*) 구체적으로 바이러스 간염과 같은 기저 간 질환이 있는 환자, 또는 혈우병 환자의 많은 분율을 차지하는 간세포가 활성적으로 분열하는 어린이에서(Nault 등, *Mol Cell Oncol*. 3(2):e1095271, 2016, Nault 등, *Nat Genet*. 47(10):1187-1193, 2015), 원종양형성 유전자 부근의 잠재적인 rAAV 통합(Russell 등, Nat Genet. 2015;47(10):1187-1193)으로 인한 종양 형성 위험 증가.

[1161] HDAd 벡터를 사용하여 적혈구 세포로부터 FVIII을 발현하는 접근법은 이러한 문제를 해결한다. 이 연구는 미니 LCR의 제어 하에 GFP를 리포터 유전자로 사용하여 생체내 HSC 형질도입/선택 후 적혈구 세포에서 비-적혈구 단백질의 발현 및 성숙 적혈구에서 GFP의 저장을 달성할 가능성이 있음을 나타낸다(도 89a-89h 참조). 그런 다음 "건강한" hCD46 유전자이식 마우스에서 접근법이 항-FVIII 혈장 항체의 존재에도 불구하고 A형 혈우병 마우스 모델에서 생리학적 수준의 FVIII의 생체공학적 형태 및 표현형 교정을 초래함을 입증하였다.

[1162] 제안된 접근법은 단일 정맥내 개입 후 평생 치료 교정을 제공할 수 있다. 적혈구로의 분화 시 유전자 변형된 HSC의 거대한 증폭 및 이러한 세포의 고효율 단백질 합성 기구는 치유적 수준에서 FVIII 생산에 대한 기초를 생성한다. 또한, HSC의 유일한 분획의 유전적 변형은 이식유전자 산물에 대한 내성을 초래할 수 있다. 생체내 유전자를 HSC로 전달하기 위해 새로 개발된 이러한 접근법은 골수절제술 및 HSC 이식을 필요로 하지 않는다. 이는 골수에서 말초 혈류로 HSC를 동원하기 위한 G-CSF/AMD3100의 주사 및 통합 헬퍼-의존적 아데노바이러스(HDAd) 벡터 시스템의 정맥내 주사를 수반한다(도 90b). HDAd5/35++ 및 HDAd35 벡터는 원시 HSC 상에서 발현되는 수용체인 CD46을 표적화한다. 이식유전자 통합은 과활성 잠자는 미녀 트랜스포사제(SB100x)를 사용하여 (무작위 패턴으로) 달성된다(도 90a). CD46-유전자이식 마우스에서 생체내 HSC 형질도입/선택 후, 생체공학적 인간 인자 VIII 버전(ET3)의 과생리적 혈청 농도 및 활성이 입증되었다(도 90c-90i; 91a-91d; 92a-92g). ET3 유전자는 ET3 발현을 적혈구로 제한하는 미니 β-글로빈 LCR의 제어 하에 있었다. 적혈구 세포로부터 고수준 ET3 생산에도 불구하고, 조혈에 대한 효과는 관찰되지 않았다. 억제 항-ET3 항체의 초기 개발 후, 처리된 마우스의 50%에서 크게 감소된 혈청 항체 수준은 흉선에서 낮은 수준의 ET3 발현 및 내성 발달로 인할 가능성이 가장 크다. CD46-tg/A형 혈우병 마우스로부터 HSC의 생체외 및 생체내 형질도입 및 치명적으로 조사된 A형 혈우병 마우스로의 후속 이식 후, 표현형 교정은 생리학적 인자 VIII 혈청 활성, 정상 aPTT, 및 꼬리 절단 후 정상적인 출혈 시간에 기반하여 달성되었다.

[1163] **논의**. FVIII 이외에도, 다른 분비된 단백질에 대한 이 접근법의 적용은 예를 들어 다음에 사용될 수 있다: (*i*) 다른 응고 인자, 구체적으로 FXI, FVII(Binny 등, *Blood*. 119(4):957-966, 2012), 폰빌레브란트 인자(VWF)(De Meyer 등, *Arterioscler Thromb Vasc Biol*. 28(9):1621-1626, 2008), 뿐만 아니라 희귀 응혈 인자(즉, 인자 I, II, V, X, XI, 또는 XIII); (*ii*) 폼페병(산 α-글루코시다제), 고세병(글루코레세브로시다제), 파브리병(α-갈락토시다제 A), 및 I형 점액다당류증(α-L-이두로니다제)과 같은 현재 리소좀 축적 질환에 대한 효소 대체 요법(ERT)에 사용되는 효소(교차-교정 메커니즘의 이점을 취함)(Penati 등, *J Inherit Metab Dis*. 40(4):543-554, 2017); (*iii*) 면역결핍 예를 들어, SCID-ADA(Cicalese 등, *Mol. Ther*. 26(3):917-931 2018)(아데노신 데아미나

제); (*iv*) 심혈관 질환, 예를 들어, 가족성 아포지질단백질 E 결핍 및 아테롬성동맥경화증(ApoE)(Wacker 등, *Arterioscler Thromb Vasc Biol*. 38(1):206-217, 2018); (*v*) 바이러스 유인 수용체의 발현에 의한 바이러스 감염(예를 들어 HIV의 경우 - 가용성 CD4(Falkenhagen 등, *Mol Ther Nucleic Acids*. 9:132-144, 2017), 또는 HIV(Kuhlmann 등, *Mol Ther*. 27(1):164-177, 2019), 만성 HCV(Quadeer 등, *Nat Commun*. 10(1):2073, 2019), 또는 HBV(Kuciinskaite-Kodze 등, *Virus es*. 211:209-221, 2016) 감염에 대한 광범위한 중화 항체(bNAb); 및 (*vi*) 암(예를 들어 단클론 항체(예를 들어 트라스투주맙(Zafir-Laviee 등, *J Control Release*. 291:80-89, 2018) 또는 체크포인트 억제제(예를 들어 aPDL1(Engeland 등, *Mol Ther*. 22(11):1949-1959, 2014))의 제어된 발현).

[1164] **실시예 8. 생체내 HSC 형질도입 후 비-인간 영장류에서 SB100x-매개 유전자 추가 및 내인성 γ-글로빈의 BE-매개 재활성화 둘 다의 검증.**

[1165] 이 실시예는 SB100x-매개 유전자 추가 및 내인성 γ-글로빈의 BE-매개 재활성화가 둘 다 생체내 HSC 형질도입 후 비-인간 영장류에서 효과적임을 검증하는 연구를 기재한다.

[1166] **유전자 전달 벡터**: 유전자 전달 벡터인 HDAd-combo가 사용될 것이다: 벡터는 다음 이식유전자의 SB100x 트랜스포사제-매개 무작위 게놈 통합을 함유한다: *i)* 적혈구에서 효율적인 발현을 위한 미니-LCR의 제어 하에 레서스 γ-글로빈 유전자, *ii)* $O^6$BG/BCNU로 형질도입된 세포의 생체내 선택을 위한 보편적으로 활성인 EF1a 프로모터의 제어 하에 레서스 $mgmt^{P140K}$, *iii)* 말초 혈액 T-세포 형질도입 및 벡터 생물분포 연구의 분석을 위한 보편적으로 활성인 EF1a 프로모터의 제어 하에 GFP. 이는 HBG 프로모터에서 BCL11a 억제인자 단백질 결합 부위의 불활성화 및 적혈구 bcl11a 인핸서의 동시 불활성화(이는 적혈구 세포에서 BCL11a 억제인자 단백질 발현 감소를 초래함)를 통해 내인성 γ-글로빈의 재활성화를 위한 아데닌 염기 편집기를 추가로 포함할 것이다. 또한, 염기 편집기 발현 카세트는 iCas-BE의 일시적 발현만을 초래하는 트랜스포존의 Flp 재조합효소 매개된 절제 시 제거될 것이다. 마지막으로, SB100x 트랜스포사제 및 Flp 재조합효소를 함유하는 벡터는 통합되지 않고 HSC 세포 증식 동안 손실될 것이다(도 121).

[1167] **치료 프로토콜**: 이전에 테스트된 HSC 동원 및 $O^6$BG/BCNU 생체내 선택 프로토콜을 사용하여 3 마리의 히말라야 원숭이(*Macaca mulatta*)로 6-개월 연구가 수행될 것이다(도 122). 프로토콜은 1 마리의 동물을 테스트하는 것으로 시작될 것이다. 연구는 8 주(마지막 생체내 선택 주기의 종료)에 심각한 합병증이 발생하지 않으면 나머지 2 마리 동물에서 반복할 것이다.

[1168] **동원**: 피하로 주어진 GCSF 및 SCF은 5 일 동안 아침에 제공될 것이다(각각 50 μg/kg). 피하로 주어진 GCSF/SCF + AMD3100은 마지막 2 일 동안 오후에 생길 것이다(5 mg/kg).

[1169] **전처리**: 4 mg/kg으로 투여된 덱사메타손은 HDAd5/35++ 주사 16 전에 정맥내로 주어질 것이다. 20 mg/kg으로 투여된 메틸프레드니솔론 및 4 mg/kg으로 투여된 덱사메타손은 정맥내로 주어질 것이지만, 100 mg으로 투여된 아나킨라는 HDAd5/35++ 주사 30 분 전에 피하로 주어질 것이다.

[1170] **HDAd 주사**: 2 차례의 HDAd 주사는 정맥내로 1) -1 일에 저용량(2 mL/분으로 20 mL의 포스페이트 완충된 염수 중 $3x10^{11}$ vp/kg)로 주어질 것이고, 2) 2 개의 완전 용량(2 mL/분으로 20 mL의 포스페이트 완충된 염수 중 $1x10^{12}$ vp/kg)은 0 일차에 30 분 간격으로 주어질 것이다.

[1171] **일시적 면역억제**: 면역억제는 $O^6$BG/BCNU의 첫번째 용량(4 주)까지 1 일에 시작하고, 필요한 경우, $O^6$BG/BCNU의 마지막 용량 후 2 주간 계속될 것이다. 면역억제는 0.2 mg/kg/일의 라파마이신, 30 mg/kg/일의 마이코페놀레이트 모페틸, 및 0.25 mg/kg/일의 타크롤리무스를 포함하며, 모두 음식을 통해 경구로 매일 제공될 것이다.

[1172] **$O^6$BG/BCNU를 사용한 생체내 선택**: $O^6$BG: 동물은 적어도 30 분에 걸쳐 정맥내로 주입되는 200 mL의 염수 중 120 mg/$m^2$ $O^6$BG를 받을 것이다. BCNU는 $O^6$BG 주입을 시작하고 60 분 후에 투여될 것이다. 그런 다음 동물은 BCNU 투여 6 내지 8 시간 후 적어도 30 분에 걸쳐 정맥내로 200 mL의 염수 중 또 다른 용량의 $O^6$BG를 받을 것이다. 첫번째 처리는 HDAd 주사 4 주 후에 주어지고; 두번째 및 세번째 처리는 γ-글로빈 마킹 및 혈액학에 따라 2 주 간격(임의적)으로 주어질 것이다.

[1173] **수집할 데이터**: 혈액 샘플은 도 122에 나타낸 바와 같이 수집될 것이다. 매일 신체 관찰 및 매주 체중 측정이

수행될 것이다.

[1174] **혈액 샘플**: 2 시간 및 6 시간 혈액 샘플을 위해, 다음 검정이 수행될 것이다: CD34+에서 GFP+ 세포의 백분율 및 CD38-/Cd45RA에서 GFP+ 세포의 퍼센트, CD90+ 세포가 정량화되고, 콜로니 형성 단위 검정을 사용하여 %GFP+ 콜로니의 퍼센트, SDF1-a로의 이동, 및 CXCR4 및/또는 VLA-4의 발현율을 평가할 것이다(예를 들어, 도 93b-93e). 모든 다른 샘플의 경우, 혈액 세포 계수, 화학, c-반응성 단백질, 및 전염증성 사이토카인이 측정될 것이다. γ-글로빈 발현은 유세포 분석(적혈구/비-적혈구 세포)을 통해 측정되는 반면, HPLC 및 qRT-PCR은 사용하여 추가된 γ-글로빈에 대한 재활성화된 γ-글로빈 수준을 측정할 것이다. 사이토스핀은 γ-글로빈 면역형광을 평가하는 데 사용될 것이다. 벡터 카피 수 및 Cas9, SB100x, 및 Flpe mRNA 수준이 측정될 것이다. 백혈구($CD4^+$, $CD8^+$, CD25, CD45RO, CD45RA, CCR-7, CD62L, FOXP3, 인테그린 αeβ7)에서 GFP 발현이 측정될 것이다.

[1175] **골수 샘플**: 골수 샘플은 4 일째에 수집한 다음 매월 수집될 것이다(도 122 참조). 골수 샘플의 계통 조성물은 유세포 분석에 의해 평가될 것이다. CD34+ 세포의 벡터 카피 수가 또한 측정될 것이다. γ-글로빈은 Ter119+/Ter119- 마커로 분류함으로써 유세포 분석을 사용하여 평가될 것이다. HPLC 및 qRT-PCR을 사용하여 추가된 γ-글로빈에 대한 재활성화된 γ-글로빈 수준이 측정될 것이다. 이러한 분석 외에도, 부검 시, 전체 게놈 서열분석을 CD34+ 세포에서 수행하여 SB100-매개 통합 및 염기 편집기 표적외 효과를 식별할 것이다. RNA 서열분석을 또한 CD34+ 세포에서 수행하여 처리 전 및 처리 후 사이의 mRNA 및 miRNA 프로파일을 비교할 것이다.

[1176] **부검으로 수득한 조직(생식계열 조직 및 정액 포함):** 일상적인 조직학이 수행될 것이고, 벡터 카피 수가 주요 조직 그룹에서 측정될 것이다. γ-글로빈 및 GFP 면역형광이 조직 절편에서 평가될 것이다.

[1177] **결과**: 이 실험은 SB100x-매개 유전자 추가 및 내인성 γ-글로빈의 BE-매개 재활성화 둘 다가 생체내 HSC 형질도입 후 비-인간 영장류에서 효과적임을 검증할 것이다. 벡터는 SCA 환자에서 치유적일 수 있는 적혈구에서 γ-글로빈 발현 수준을 달성함이 입증될 것이다(즉 80% 초과의 $\gamma$-글로빈$^+$ RBC 및 성체 레서스 글로빈의 20% 초과의 γ-글로빈 수준). 또한 장기간 혈액학적 부작용의 부재 및 원치않은 게놈 재배열의 부재 및 HSC 전사체의 변화를 입증할 것이다. 마지막으로, 정맥내로 주사된 HDAd5/35++ 벡터가 기억 T-세포를 형질도입함을 입증할 것이다.

[1178] **실시예 9. 내인성 γ-글로빈 빈 발현의 재-활성화를 위한 염기 편집기를 발현하는 HDAd5/35++ 벡터를 사용한 인간 및 레서스 원숭이 HSC 형질도입.**

[1179] 사이티딘 또는 아데닌 데아미나제 또는 트랜스아미나제에 융합된 불활성 Cas9는 태아 글로빈을 재활성화하는 도구로서 역할을 할 수 있다. 사이티딘 염기 편집기를 발현하는 HDAd 벡터(HDAd-C-BE)를 적혈구 bcl11a 인핸서를 표적화하고 중요한 GATA 결합 모티프를 파괴하는 HDAd-CRISPR/Cas9 벡터와 비교하였다(도 123). 동일한 영역에 대해 야생형 CRISPR을 발현하는 HDAd 벡터를 구축하였다. 두 벡터를 HDAd 형질도입 후, 18 일에 걸쳐 적혈구 분화에 적용된 인간 CD34+ 세포에서 테스트하였다(도 124a). HDAd-wtCRISPR 형질도입된 세포의 경우, 편집된 표적 부위의 백분율의 점진적인 감소가 관찰되었으며, CRISPR-관련된 세포독성 때문일 가능성이 가장 크다(도 124b). 게놈 편집의 효능은 HDAd-C-BE 벡터의 경우 더 낮았지만, 편집률은 안정되게 유지되어, γ-글로빈의 비슷한 재활성화를 초래하였다(도 124c). 이식 후, HDAd-C-BE 형질도입된 CD34+ 세포의 생착은 형질도입되지 않은 대조군 세포만큼 효율적이었다(도 125). 요약하면, 이러한 데이터는 염기-편집기 벡터가 잠재적으로 wtCRISPR-발현 벡터보다 HSC에서 게놈 편집을 위한 더 나은 도구임을 나타낸다. 보다 최근에, HBG1/2 프로모터에서 3 개의 상이한 영역에 대해 아데닌 편집기를 발현하는 HDAd 벡터 시리즈가 개발되었다. γ-글로빈 재활성화는 염기 편집기 벡터로 여러 억제인자 수준을 동시에 표적화함으로서 실질적으로 증가될 수 있을 것으로 예상된다. 이 목표를 향해 적혈구 bcl11a 인핸서(도 126, 상부 패널) 또는 HBG1/2의 BCL11a 단백질 결합 부위(도 126, 하부 패널)를 표적화하는 염기 편집기를 발현하는 HDAd 벡터를 테스트하였다. 시험관내 연구에서 γ-글로빈 재활성화는 두 벡터에 대해 각각 9 및 53%였다.

[1180] **SCA 마우스 모델(Townes 모델)의 데이터:** B6;129-$Hbb^{tm2(HBG1,HBB*)Tow}$/$Hbb^{tm3(HBG1,HBB)Tow}$ $Hba^{tm1(HBA)Tow}$/J; hα/hα::$\beta^A$/$\beta^S$, hα/hα::-383 γ-$\beta^A$/-1400 γ-$\beta^S$.

[1181] 마우스는 인간 α-글로빈, γ-글로빈(프로모터를 함유하는 -383 및 -1400 영역 포함), 상응하는 마우스 유전자 대신에 $\beta^{87}$-SCA 글로빈을 함유하고 말초 혈액에서 망상적혈구의 40%, 낮은 적혈구용적률, 낮은 헤모글로빈 수준, 및 백혈구증가증(도 127b)을 갖는 심각한 SCA 표현형(도 127a)을 나타낸다. 이러한 마우스를 CD46에 대

한 동형접합성 및 3 개의 글로빈 유전자 치환을 달성하도록 번식시켰다(CD46/Townes 마우스). 이전에 개발된 HDAd-HBG-CRISPR 벡터가 CD46/Townes 마우스의 생체내 HSC 형질도입 후 γ-글로빈을 활성화시키는지 여부를 결정하기 위해 테스트하였다(도 128a). $O^6$BG/BCNU 선택 없이, RBC의 γ-글로빈 마킹은 60%에 도달하였으며, 이는 Townes 마우스의 적혈구생성의 기능적 결핍이 게놈-편집된 HSC/ 적혈구 전구체 세포에 대한 강한 증식 자극을 제공함을 나타낸다(도 128b). HDAd-HBG-CRISPR 벡터의 치료 효과는 크게 개선된 적혈구 표현형 및 말초 망상적혈구의 5-배 감소를 반영하였다(도 128c). 이는 이 모델(및 잠재적으로 SCA 환자)의 치유가 $O^6$BG/BCNU 생체내 HSC 선택에 대한 필요성 없이 달성될 수 있음을 나타낸다.

[1182] **비-인간 영장류(NHP)에서 생체내 HSC 유전자 전달:** 이들 데이터는 G-CSF, SCF, 및 AMD3100으로 동원 후 HDAd-GFP 주사를 받은 2 마리의 NHP(돼지꼬리 원숭이(*Macaca nemestrina*))로부터 수득된다(도 129a; 도 93a; 도 94e-94g). 말초 혈액 샘플을 벡터 주사 직전, 및 벡터 주사 2 및 6 시간 후에 수집하였다. 단리된 CD34+ 세포를 생체외에서 배양하고 콜로니 형성 검정에 플레이팅하였다. 벡터 투여 후 단리된 CD34+ 세포의 평균 3%는 GFP+였으며(도 129b; 도 93b, 93c; 도 94h), 이는 말초 혈액에서 동원된 CD34+ 세포가 HDAd5/35++ 벡터의 단일 정맥내 투여에 의해 형질도입될 수 있음을 시사한다. 이러한 CD34+ 세포가 콜로니-형성 가능성을 유지하는지를 테스트하기 위해, 콜로니 검정을 수행하고, PCR을 통해 GFP 이식유전자를 보유하는 콜로니의 백분율을 결정하였다. 주사 후 시점으로부터 CD34+ 세포에서 유래된 콜로니의 최대 55%가 벡터에 의해 형질도입되었다(도 129c; 도 93d; 또한 도 94i-94m 참조). 마지막으로, 말초 동원 후 골수 구획으로 되돌아가는 생체내 벡터-표적화된 세포의 능력을 테스트하기 위해, 골수 흡인물을 벡터 투여 3 일 후 동물 중 하나로부터 수집하였다. 골수-상주 CD34+ 세포의 3.7% 또는 2.9%는 GFP+였고, 생체내 전달 전 vs 전달 후 수집된 세포에서 콜로니 형성 가능성의 주목할만한 차이는 관찰되지 않았다(도 129d; 도 93e). 이러한 비-인간 영장류 연구(마우스에서보다 10배 더 낮은 벡터 용량으로 수행됨)는 기재된 생체내 전달 접근법이 검증된 전임상 모델에서 실행가능하고 안전함을 입증한다.

[1183] **실시예 10. 염기 편집기를 사용한 생체내 HSC 유전자 요법은 β-YAC 마우스에서 태아 γ-글로빈의 효율적인 재활성화를 허용한다**

[1184] 이 실시예는 생체내에서 HDAd5/35++ 벡터에 의해 전달된 염기 편집기가 예를 들어, 혈색소병증의 치료를 위해 정확한 게놈 조작을 위한 유용하고 효과적인 전략임을 입증한다.

[1185] 염기 편집기는 표적화된 게놈 유전자좌에서 정확한 뉴클레오티드 돌연변이를 설치할 수 있고 이중-가닥 DNA 파괴를 피하는 이점을 제시하다. 여기서, 중요한 모티프는 HDAd5/35++ 벡터를 통해 전달된 염기 편집기로 γ-글로빈 재활성화 조절을 표적화하였다. 최적화된 설계를 통해, *HBG1/2* 프로모터에서 *BCL11A* 인핸서를 표적화하거나 또는 자연 발생 태아 혜모글로빈의 유전적 지속성(HPFH) 돌연변이를 재생성하는 사이티딘 및 아데닌 염기 편집기(CBE 및 ABE)의 패널은 성공적으로 구제되었다. HUDEP-2 세포에서, 5 개의 테스트된 벡터는 모두 표적 염기 전환을 효율적으로 설치하고 실질적인 γ-글로빈 재활성화로 이어졌다. *HBG1/2* 프로모터에서 -113A에서 G로의 HPFH 돌연변이에 특이적인 ABE 벡터 HDAd-ABE-sgHBG#2를 사용함으로써 상당한 γ-글로빈 단백질 생산(β-글로빈에 비해 23%)이 관찰되었다. 따라서 이 벡터가 하류 동물 연구에 선택되었다. 248 kb의 인간 β-글로빈 유전자좌를 보유하는 마우스(β-YAC 마우스)를 사용하여 글로빈 전환을 정확하게 반영하였다. *FRT* 및 트랜스포존 부위에 의해 플랭킹된 EF1α-$MGMT^{P140K}$ 발현 카세트는 형질도입된 세포의 생체내 선택을 허용하는 벡터에 포함되었다. HDAd-ABE-HBG#2 + HDAd-SB로의 생체내 형질도입 및 저용량의 화학선택 후, 말초 적혈구에서 평균 40% 초과의 HbF-양성 세포가 측정되었다. 이는 인간 β-글로빈에 비해 γ-글로빈 생산의 21%에 상응하였다. 총 골수 세포에서 -113 A에서 G로의 전환은 평균 20%였다. 형질도입되지 않은 마우스와 비교하여, 처리 후 혈액학적 매개변수, 적혈구생성 및 골수 세포 조성물에서 변경은 관찰되지 않았으며, 이는 접근법의 우수한 안전성 프로파일을 입증한다. 상위 점수의 잠재적인 표적외 게놈 부위에서 검출가능한 편집은 발견되지 않았다. 골수 계통 마이너스 세포를 형질도입 후 16 주차에 1차 마우스로부터 단리하고 치명적으로 조사된 C57BL/6J 마우스에 주입하였다. HbF-양성 세포의 백분율은 16 주에 걸쳐 2차 수용자에서 유지되었으며, 이는 게놈 편집이 장기간 재증식 마우스 HSC에서 발생하였음을 나타낸다. 관찰은 HDAd5/35++ 벡터에 의해 전달된 염기 편집기가 혈색소병증의 치료를 위한 정확한 생체내 게놈 조작을 위한 유망한 전략을 나타냄을 입증한다.

[1186] CRISPR/Cas9와 같은 뉴클레아제에 기반한 게놈 조작 전략은 임상 평가 단계에 들어가는 다중 유전자 요법 연구와 함께 주목할만한 발전을 달성하였다. CRISPR/Cas9-매개 유전자 편집은 고전적인 비-상동 말단 결합(NHEJ)을 포함하는 내인성 복구 메커니즘을 촉발하는 이중-가닥 DNA 파괴(DSB)에 의존한다. 공여자 DNA 주형의 존재 하에, 상동성-지시된 복구(HDR)는 전형적으로 더 낮은 빈도로 발생할 수 있다. 최신 연구는 혈액 장애에 대한 유

전적 요법에 중요한 조혈 줄기 및 전구체 세포(HSPC)에서 관심 유전자의 매우 효율적인 파괴를 입증하였다(Martin 등, *Cell Stem Cell* 24: 821-828.e825, 2019; Wu 등, *Nature Medicine* 25: 776-783, 2019). 그러나, 연구는 뉴클레아제-유도된 DSB가 원치않은 큰 단편 결실 및 p53-의존적 DNA 손상 반응을 생성함으로써(Haapaniemi 등, *Nature Medicine*, 24(7):927-903, 2018; Ihry 등, *Nature Medicine*, 24(7):939-946, 2018; Kosicki 등, *Nature Biotechnology* 36: 765, 2018) 숙주 세포에 대한 부작용을 유발할 수 있음을 보고하였다(Haapaniemi 등, *Nature Medicine*, 24(7):927-903, 2018; Ihry 등, Nature Medicine, 24(7):939-946, 2018; Kosicki 등, *Nature Biotechnology* 36: 765, 2018).

[1187] 염기 편집기(BE)는 DSB를 생성하지 않으면서 표적화된 게놈 유전자좌에서 정확한 뉴클레오티드 치환을 설치할 수 있다. 이들은 핵염기 데아미나제 효소 및, 일부 경우에, DNA 글리코실라제 억제제에 융합된, DSB를 만들 수 없는 Cas9 닉카제(nCas9)와 같은 촉매적으로 기능상실된 뉴클레아제를 포함한다. 현재, 2 가지 주요 범주인 사이티딘 염기 편집기(CBE) 및 아데닌 염기 편집기(ABE)가 있으며, 이는 nCas9와 커플링된 단일 가이드 RNA(sgRNA)에 의해 지시된 좁은 표적가능한 창(일반적으로 약 5 개 염기 쌍)에서 각각 C > T 및 A > G 전이를 전환시킨다(Gaudelli 등, *Nature* 551: 464-471, 2017; Komor 등, *Nature* 533: 420-424, 2016; Nishida 등, Science 353, 2016). CBE 및 ABE 사이의 핵심 차이는 CBE가 사이티딘 데아미나제(예를 들어, APOBEC1)를 함유하고 ABE가 실험실에서 진화된 TadA 데옥시아데노신 데아미나제를 사용하는 데아미나제 영역에 위치한다. 다중 그룹은 다양한 진핵생물 세포에서 효율적인 염기 편집을 보고하였다(Zhang 등, *Genome Biology* 20: 101, 2019; Chadwick 등, *Arterioscler Thromb Vasc Biol* 37: 1741-1747, 2017; Zeng 등, *Nature Medicine* 26: 535-541, 2020; Lim 등, *Mol Ther*, 82(4):1177-1189, 2020; Gao 등, *Nature* 553: 217-221, 2018). 인간에서 알려진 모든 병원성 단일 뉴클레오티드 다형성(SNP)의 60%는 현재 BE에 의해 잠재적으로 역전될 수 있는 것으로 예측된다(Rees 등, *Nature Reviews Genetics* 19: 770-788, 2018).

[1188] β-혈색소병증은 주로 β-지중해빈혈 및 겸상적혈구병(SCD)를 포함하는 정상 β글로빈 생산이 부재하거나 또는 겹핍된 유전 질환의 통상적인 그룹이다. 특이적 유전적 결함에 따라, β-지중해빈혈 및 SCD 환자는 다양한 중증도의 질환 징후를 나타낸다. 신생아 스크리닝 및 치료 예방으로 SCD 어린이의 사망률이 크게 감소하였지만, 대부분의 β-주요 지중해빈혈($\beta^0$) 및 SCD 환자는 평생 급성 및 만성 합병증으로 고통을 받는다(Ware 등, *Lancet* 390: 311-323, 2017; Higgs 등, *Lancet* 379: 373-383, 2012). 그러나, 대부분의 임신 단계 동안 우세하고 일반적으로 출생 직후 침묵하는 태아 헤모글로빈(HbF) 수준이 높은 일부 성인 환자에서, 질환 증상은 현저하게 경미하다. 이러한 태아 헤모글로빈의 유전적 지속성(HPFH) 현상은 HbF의 강한 보호 효과를 입증하고 β-글로빈 장애 환자에 대한 유전적 요법 전략으로서 γ-글로빈의 재활성화에 대한 우수한 근거를 제공한다.

[1189] 다수의 HPFH 돌연변이가 보고되었다(Orkin & Bauer, *Annual Review of Medicine* 70: 257-271, 2019 및 Wienert 등, *Trends in Genetics: TIG* 34: 927-940, 2018에 의해 검토됨). *HBG1/2* 프로모터에서 약 -150, -175 및 -200 부위에 위치한 HPFH SNP의 3 가지 주요한 클러스터가 있다. 이들 부위에서 HPFH 돌연변이의 도입은 HbF 억제인자(예를 들어, BCL11A 및 ZBTB7A)의 결합 부위를 파괴하거나 또는 활성인자(예를 들어, TAL1 및 KLF1)의 기능 획득 결합 부위를 생성하여, 억제된 HbF 발현으로 이어질 수 있다(Traxler 등, *Nature Medicine* 22: 987-990, 2016; Martyn 등, *Nature Genetics* 50: 498-503, 2018). HbF 재활성화는 또한 주요한 HbF 억제인자인 BCL11A과 같은 HbF 조절인자의 발현을 조절함으로써 달성될 수 있다(Sankaran 등, *Science* 322: 1839-1842, 2008). 직접 BCL11A 녹아웃은 발달적으로 필수적인 역할로 인해 옵션이 아니었지만, 적혈구-특이적 인핸서의 편집에 의한 BCL11A의 부분적 하향조절은 동물 생존력을 유지하면서 효율적인 HbF 유도를 허용한다(Wu 등, *Nature Medicine* 25: 776-783, 2019; Canver 등, *Nature* 527: 192-197, 2015). BE:sgRNA 리보뉴클레오단백질(RNP) 전기천공법을 사용하여, 최근 연구는 +58 BCL11A 인핸서의 중요한 모티프를 염기 편집기로 파괴하는 것이 환자-유래된 $CD34^+$ HSPC에서 치료적 HbF 유도로 이어짐을 입증하였다.

[1190] 생체내 HSC 형질도입에 의한 단순화된 유전자 요법 접근법이 최근에 규명되었다. HSC 향성을 위한 키메라 섬유, 가장 통상적으로 사용되는 이식유전자를 수용하기 위한 32 kb 초과의 페이로드 등을 포함하는 다양한 유리한 특성으로 인해 헬프-의존적 HDAd5/35++ 벡터가 사용되었다. 이 연구에서, 최적화된 설계를 사용하여 *BCL11A* 인핸서 또는 *HBG1/2* 프로모터를 표적화하는 BE 벡터의 패널을 성공적으로 생성하였다. 유전자이식 마우스 모델에서, HDAD-ABE 벡터를 사용한 생체내 HSC 염기 편집은 HPFH 돌연변이를 재생성하고 효율적인 HbF 유도로 이어짐이 여기에 제시된다.

[1191] ***재료 및 방법.***

[1192] **생체내 형질도입 및 선택을 위한 시약:** G-CSF(Neupogen™)(Amgen, 캘리포니아주 사우전드 오크스 소재),

AMD3100(MilliporeSigma, 매사추세츠주 버링턴 소재) 및 덱사메타손 나트륨 포스페이트(Fresenius Kabi USA, 일리노이주 레이크 주리치 소재)를 사용하였다. $O^6$-벤질구아닌($O^6$-BG) 및 카르무스틴(BCNU)은 MilliporeSigma로부터 수득하였다.

[1193] **HDAd 벡터 생성:** 하버드의 David R. Liu 실험실에서 개발된 염기 편집 시스템을 사용하였다(Koblan 등, *Nature Biotechnology* 36: 843-846, 2018). pCMV_AncBE4max 및 pCMV_ABEmax 플라스미드는 Addegene(매사추세츠주 워터타운 소재)으로부터 구입하였다. Addgene으로부터 구입한 다음 플라스미드를 또한 사용하였다: 도 131a 및 131b의 BE4, ABE7.10, pLenti-BE3RA-PGK-Puro 및 pLenti-FNLS-PGK-Puro 및 BE3RA(Zafra 등, *Nature Biotechnology* 36: 888-893, 2018). 하기 기재된 올리고 및 gBlock은 Integrated DNA Technologies(IDT)(아이오와주 코랄빌 소재)에 의해 합성하고 표 14에 나열하였다.

[1194] **표 14: 염기 편집기를 위한 가이드 서열.**

| 편집기 | 명칭 | 서열 (5'에서 3'으로)* | 표적화 부위 / 비고 |
|---|---|---|---|
| **BCL11A** | | | |
| CBE | sgBCL#1 | TTTATCACAGGCTCCAGGAA | GATAA 모티프 |
| CBE | sgBCL#2 | TTTTATCACAGGCTCCAGGA | GATAA 모티프 |
| ABE | sgBCL#3 | TTTATCACAGGCTCCAGGAA | GATAA 모티프 |
| ABE | sgBCL#4 | TTTTATCACAGGCTCCAGGA | GATAA 모티프 |
| ABE-xCas9 | sgBCL#5 | CTGTGATAAAAGCAACTGTT | GATAA 모티프, NGC PAM |
| ABE-xCas9 | sgBCL#6 | GATAAAAGCAACTGTTAGCT | GATAA 모티프, NGC PAM |
| **HBG** | | | |
| CBE | sgHBG#1 | CTTGACCAATAGCCTTGACA | TGACCA, -114-115 CC>GG |
| ABE | sgHBG#2 | CTTGACCAATAGCCTTGACA | TGACCA, -113 A>G |
| ABE | sgHBG#3 | GCTATTGGTCAAGGCAAGGC | -111 T>C |
| ABE | sgHBG#4 | GTGGGGAAGGGGCCCCCAAG | -198 T>C |
| CBE-xCas9 | sgHBG#5 | CCTTCCCCACACTATCTCAA | -197 C>T, -196 C>T, NGC PAM |
| ABE-xCas9 | sgHBG#6 | AGATATTTGCATTGAGATAG | -175 T>C, NGT PAM |
| **HBB** | | | |
| CBE | sgHBB_STOP | CTTGCCCCACAGGGCAGTAA | HBB CDS, TGG > TAA (HBB STOP) |
| ABE | sgHBB_SKIP | AGACTCACCCTGAAGTTCTC | HBB 인트론, GT > GC (HBB 스플라이싱 SKIP) |
| **양성 대조군** | | | |
| CRISPR/Cas9 | sgHBG_CRISPR | CTTGTCAAGGCTATTGGTCA | TGACCA |
| **음성 대조군** | | | |
| CBE | sgNeg | GGTGTCGAAATGAGAAGAAG | CCR5 CDS, nt673 C>T (CCR5 STOP) |

밑줄 표시: 중요한 모티프에서 표적화 염기(들).
*상단에서 하단으로: 서열번호: 244, 245, 244, 245, 248-250, 250, 252-259.

[1195]

[1196] *CBE 및 ABE 작제물의 첫번째 버전:* 클로닝은 3 단계를 수반하였다. 단계 1) BE4에서 *BsmBI* 부위는 *EagI-NaeI* 단편을 gBlock #1로 교체함으로써 파괴하였다. pCMV_AncBE4max에서 *BsmBI* 부위는 *BsmBI-NarI* 단편을 gBlock #2로 교체함으로써 파괴하였다. *BsmBI* sgRNA 클로닝 부위가 있는 pBST-CRISPR이라는 벡터는 주입을 사용하여(Takara, 캘리포니아주 마운틴뷰 소재) 다음 4 개의 단편을 조합함으로써 생성되었다: #3FR을 사용하여 LentiCRISPRv2(Addgene)로부터 증폭된 2.3 kb U6-충전제-gRNA 스캐폴드 단편, 각각 #4FR 및 #5FR을 사용하여 pBST-sgBCL11Ae1로부터 증폭된 1.4 k b 및 1.0 kb 단편(Li 등, Blood 131: 2915-2928, 2018), 및 *BsaI-BamHI* 소화에 의해 방출된 pBST-sgBCL11Ae1의 9.6 kb 단편. 중간 플라스미드 pBS-U6-Ef1α는 주입을 사용하여 다음 3 개의 단편을 연결함으로써 구성되었다: 각각 프라이머 #6FR 및 #7FR을 사용하여 pBST-CRISPR로부터 증폭된 3.6 kb U6-충전제-스캐폴드-Ef1α 서열 및 2.9 kb 벡터 백본, 및 *BseRI* 클로닝 부위를 함유하는 0.5 kb gBlock(#8). 이 중간체를 *BseRI*로 소화시키고 *EagI-PmeI* 처리 후 BE4-ΔBsmBI의 5.5 kb 단편과 재조합하여, pBS-BE4를 생성하였다. 6.6 kb pBS 백본-U6-충전제-gRNA 스캐폴드-Ef1α 서열을 #9FR을 사용하여 pBS-BE4로부터 PCR 증폭시킨 후, *NotI-AgeI*-소화된 pCMV-ABEmax 및 pCMV_AncBE4max-ΔBsmBI를 주입하여, 각각 pBS-AncBE4max 및 pBS-ABEmax를 생성하였다. 다음으로, sgRNA 올리고를 합성하고, 어닐링하고 pBS-BE4, pBS-AncBE4max 및 pBS-ABEmax의 *BsmBI* 부위로 삽입하여, pBS-ABEmax-sgHBG#2와 같은 일체형 염기 편집 구성요소를 갖는 셔터 플라스미드를 생성하였다. 단계 2) *PacI* 클로닝 부위가 있는 21.0 kb pHCAS3-MCS 벡터를, 스터퍼 DNA를 *EcoRI* 제한에 의해 줄이고 1.8 kb *EcoRI* 단편과 재결찰시키는 것을 제외하고 이전에 기재된 것과 유사하게 생성하였다(Li 등, *Cancer Res* 80: 549-560, 2020). 2.2 kb PGK-MGMT$^{P140K}$-2A-GFP-bGHpolyA 서열을 #10FR에 의해 pHCA-이중-MGMT-GFP로부터 증폭시키고(Li 등, Blood 131: 2915-2928, 2018) *PacI*-소화된 pHM5-FRT-IR-

Ef1α-GFP와 재조합하여(Richter 등, *Blood* 128: 2206-2217, 2016), pHM5-FI-PGK-MGMT-GFP를 생성하였다. 이후에, *I-CeuI* 및 *PI-SceI* 부위 사이의 단편을 #11FR에 의해 이 작제물로부터 pHCAS3-MCS의 *PshAI* 부위로 옮기고 주입 클로닝하여, pHCAS3-FI-PGK-MGMT-GFP-MCS를 형성하였다. 단계 3) 단계 1의 셔틀 플라스미드 및 단계 2에서 생성된 벡터를 *PacI*로 처리하고 재조합하여 pHCA-ABEmax-sgHBG#2-FI-MGMT-GFP와 같은 최종 작제물을 생성하였다. 상이한 sgRNA 서열을 갖는 최종 pHCA 작제물은, 상이한 sgRNA가 단계 1에서 사용되었다는 것을 제외하고 유사하게 생성되었다.

[1197] *ABE 작제물의 두번째 버전:* ABE 작제물의 두번째 버전은 프로모터, 대안적인 코돈 용법 및 miRNA-조절된 유전자 발현에서 첫번째 버전과 상이하다. 클로닝은 또한 3 단계를 수반하였다. 단계 1) miR183/218 표적 서열이 있는 1.5 kb 3' β-글로빈 UTR을 프라이머 #12FR을 사용하여 pBST-sgHBG1-miR로부터 증폭시킨 후(Li 등, *Blood* 131: 2915-2928, 2018), pBS-ABEmax-sgHBG#2의 *NotI-HpaI* 부위에 삽입하여, pBS-ABEmax-sgHBG#2-miR을 생성하였다. ABE 작제물의 두번째 버전에 대한 셔틀 플라스미드, 예를 들어, pBS-ABEopti-sgHBG#2-miR은 주입 클로닝에 의해 *AscI-EcoRV*-소화된 pBS-ABEmax-sgHBG#2-miR과 다음 4 개의 단편을 연결함으로서 수득하였다: #13FR을 사용하여 pHM5-FI-PGK-MGMT-GFP로부터 증폭된 인간 PGK 프로모터, 서열 반복성을 감소시키는 대안적인 코돈 용법을 갖는 2 개의 *TadA* 유전자를 함유하는 2 개의 gBlock(#14 및 #15), 및 #16FR을 사용하여 pBS-ABEmax-sgHBG#2로부터 증폭된 1.9 kb 서열. 단계 2) pHM-FRT-IR-Ef1α-MGMT(P140K)-2A-GFP-pA의 *PshAI-NotI* 부위 사이의 SV40 polyA 서열을 bGH polyA 서열(gBlock #17)로 대체하여, pHM-FI-Ef1α-MGMT(P140K)-GFP-bGHpA를 수득하였다. 그런 다음, *I-CeuI* 및 *PI-SceI* 부위 사이의 전체 4.9 kb 트랜스포존을 #11FR을 사용하여 pHCAS3-MCS의 *PshAI* 부위로 전달하여, pHCAS3-FI-Ef1α-MGMT-GFP-MCS를 생성하였다. 단계 3) 단계 1 및 2로부터의 생성된 작제물을 *PacI* 처리 후 주입 클로닝에 의해 조합하여, pHCA-ABEopti-sgHBG#2-FI-MGMT-GFP를 생성하였다. 상이한 sgRNA 서열을 갖는 최종 pHCA 작제물을 유사하게 생성하였다.

[1198] Phusion Hot Start II 고충실도 DNA 폴리머라제를 클로닝에 수반되는 모든 PCR 증폭에 사용하였다. 최종 작제물을 여러 제한 효소(*HindIII*, *EcoRI* 및 *PmeI*)에 의해 스크리닝하고 이식유전자를 함유하는 전체 영역을 서열분석함으로써 확인하였다.

[1199] HDAd5/35++ 벡터의 생산을 위해, 상응하는 플라스미드를 *PmeI*로 선형화하고 Ad5 섬유 꼬리, Ad35 섬유 샤프트, 및 친화도-향상된 Ad35++ 섬유 놉으로 구성된 키메라 섬유를 함유하는 Ad5/35++ 헬퍼 벡터(Richter 등, Blood 128: 2206-2217, 2016)인 AdNG163-5/35++로 116 세포에서 구제하였다(Palmer & Ng, *Mol Ther* 8: 846-852, 2003). HD-Ad5/35++ 벡터를 다른 곳에 상세히 기재된 바와 같이 116 세포에서 증폭시켰다(Palmer & Ng, *Mol Ther* 8: 846-852, 2003). 헬퍼 바이러스 오염 수준은 0.05% 미만인 것으로 밝혀졌다. 역가는 2-5x$10^{12}$ 개 바이러스 입자(vp)/mL였다.

[1200] **세포주의 형질감염:** 293FT(Thermo Fisher Scientific) 및 K562 세포를 판매처의 지침에 따라 배양하였다. 6-웰 플레이트에 미리 시딩된 293FT 세포를 제조업체의 프로토콜에 따라 리포펙타민 3000(Thermo Fisher Scientific)을 사용하여 4 μg 플라스미드(3 μg 염기 편집기 또는 CRISPR/Cas9 + 1 μg pSP-sgBCL11AE(Li 등, *Mol Ther Methods Clin Dev* 9: 390-401, 2018))로 형질감염시켰다. K562 세포를 공급자의 프로토콜에 따라 뉴클레오펙션(카탈로그 # V4XC-2024)(Lonza, 스위스 바젤 소재)을 사용하여 2.66 μg 플라스미드(2 μg 염기 편집기 또는 CRISPR/Cas9 + 0.6 μg pSP-sgBCL11AE)로 형질감염시켰다. 게놈 DNA을 분석을 위해 형질감염 후 4 일째에 단리하였다.

[1201] **HUDEP-2 세포 및 적혈구 분화:** HUDEP-2 세포(Kurita 등, *PloS One* 8: e59890, 2013)를 100 ng/mL SCF, 3 IU/mL EPO, $10^{-6}$ M 덱사메타손 및 1 μg/mL 독시사이클린(DOX)이 공급된 StemSpan SFEM 배지(STEMCELL Technologies)에서 배양하였다. 적혈구 분화를 5% 인간 AB 혈청, 100 ng/mL SCF, 3 IU/mL EPO, 10 μg/mL 인슐린, 330 μg/mL 트랜스페린, 2 U/mL 헤파린 및 1 μg/mL DOX를 함유하는 IMDM에서 6 일 동안 유도하였다.

[1202] **콜로니-형성 단위(CFU) 검정:** 계통 마이너스($Lin^-$) 세포를 제조업체의 지침에 따라 마우스 계통 세포 고갈 키트(Miltenyi Biotec, 캘리포니아주 샌디에이고 소재)를 사용하여 골수 MNC에서 계통-수임 세포의 고갈에 의해 단리하였다. CFU 검정을 제조업체의 프로토콜에 따라 마우스 완전 배지와 함께 ColonyGEL(Reachbio, 워싱턴주 시애틀 소재)을 사용하여 수행하였다. 콜로니를 플레이팅 후 10 일에 점수매겼다.

[1203] **T7EI 불일치 뉴클레아제 검정:** 게놈 DNA를 제공된 프로토콜에 따라 PureLink Genomic DNA 미니 키트(Life Technologies, 캘리포니아주 칼즈배드 소재)를 사용하여 단리하였다(Miller 등, *Nat Biotechnol* 25: 778-785, 2007). 적혈구 *BCL11A* 인핸서의 표적 부위를 포함하는 게놈 분절을 PCR 프라이머에 의해 증폭시켰다: BCL11A

정방향(서열번호: 247) 및 역방향(서열번호: 263). PCR 생성물을 혼성화하고 2.5 단위의 T7EI(New England Biolabs)로 37℃에서 30 분 동안 처리하였다. 소화된 PCR 생성물을 10% TBE PAGE(Bio-Rad)로 분해하고 에티듐 브로마이드로 염색하였다. 100 bp DNA 래더(New England Biolabs)를 사용하였다. 밴드 강도는 ImageJ 소프트웨어를 사용하여 분석하였다. 절단율% = (1-sqrt(모체 밴드/(모체 밴드 + 절단된 밴드)) * 100%.

[1204] **유세포 분석:** 세포를 FACS 완충액(PBS, 1% FBS) 중 $1x10^6$ 개 세포/100 μL로 재현탁하고 FcR 차단 시약(Miltenyi Biotech, 캘리포니아주 오번 소재)과 함께 얼음 위에서 10 분 동안 배양하였다. 다음으로 염색 항체 용액을 $10^6$ 개 세포 당 100 μL로 첨가하고 암실에서 30 분 동안 얼음 위에서 배양하였다. 배양 후, 세포를 FACS 완충액에서 1 회 세척하였다. 2차 염색을 위해 염색 단계를 2차 염색 용액으로 반복하였다. 세척 후, 세포를 FACS 완충액에 재현탁하고 LSRII 유세포 분석기(BD Biosciences, 캘리포니아주 산호세 소재)를 사용하여 분석하였다. 파편을 전방 산란 영역 및 측방 산란 영역 게이트를 사용하여 제외하였다. 그런 다음 단일 세포를 전방 산란 높이 및 전방 산란 폭 게이트를 사용하여 게이팅하였다. 그런 다음 유세포 분석 데이터를 FlowJo(버전 10.0.8, FlowJo, LLC)를 사용하여 분석하였다. LSK 세포의 분석을 위해, 세포를 비오틴-접합된 계통 검출 칵테일(카탈로그 # 130-092-613)(Miltenyi Biotec, 캘리포니아주 샌디에이고 소재), c-Kit(클론 2B8, 카탈로그 # 12-1171-83) 및 *Sca-1*(클론 D7, 카탈로그 # 25-5981-82)에 대한 항체로 염색한 후, APC-접합된 스트렙타비딘(카탈로그 # 17-4317-82)(eBioscience, 캘리포니아주 샌디에이고 소재)으로 2차 염색하였다. eBioscience로부터의 다른 항체는 항-마우스 CD3-APC(클론 17A2)(카탈로그 # 17-0032-82), 항-마우스 CD19-PE-Cyanine7(클론 eBio1D3)(카탈로그 # 25-0193-82), 및 항-마우스 Ly-66(Gr-1)-PE(클론 RB6-8C5)(카탈로그 # 12-5931-82를 포함하였다. 항-마우스 Ter-119-APC(클론 Ter-119)(카탈로그 # 116211)는 Biolegend(캘리포니아주 샌디에이고 소재)로부터 수득하였다.

[1205] **인간 γ-글로빈 발현을 검출하는 세포내 유세포 분석:** FIX & PERM™ 세포 투과화 키트(Thermo Fisher Scientific)를 사용하고 제조업체의 프로토콜에 따랐다. 간단히 말해서, $5x10^6$ 개의 HUDEP-2 세포를 100 μL FACS 완충액에 재현탁하였다. 100 μL의 시약 A(고정 배지)를 첨가하고 실온에서 2-3 분 동안 배양하였다. 그런 다음 1 mL 미리 냉각된 절대 메탄올을 첨가하고 10 분 동안 암실에서 얼음 위에서 배양하였다. 그런 다음 샘플을 FACS 완충액으로 세척하고, 0.6 μg 헤모글로빈 γ 항체(클론 51-7, 카탈로그 # sc-21756 PE)(Santa Cruz Biotechnology, 텍사스주 댈러스 소재)를 함유하는 100 μL 시약 B(투과 배지)에 재현탁하고, 실온에서 30 분 동안 배양하였다. 세척 후, 세포를 FACS 완충액에 재현탁하고 분석하였다.

[1206] **글로빈 HPLC:** 개별 글로빈 쇄 수준을 SPD-10AV 다이오드 어레이 검출기 및 LC-10AT 바이너리 펌프가 장착된 Shimadzu Prominence 기기(Shimadzu, 일본 교토 소재)로 정량화하였다. 폴리펩티드용 Vydac 214TP™ C4 역상 칼럼(214TP54 Column, C4, 300 Å, 5 μm, 4.6 mm i.d. x 250 mm)(Hichrom, 영국 소재)을 사용하였다. 물/아세토니트릴 중 0.1% 트리플루오로아세트산의 40%-60% 구배 혼합물을 1 mL/분의 속도로 적용하였다.

[1207] **벡터 카피 수의 측정:** 세포 당 아데노바이러스 게놈 카피의 절대 정량화를 위해, 게놈 DNA를 제공된 프로토콜에 따라 PureLink Genomic DNA 미니 키트(Life Technologies)를 사용하여 세포로부터 단리하고, power SYBR™ green PCR 마스터 믹스(Thermo Fisher Scientific)를 사용하여 수행된 qPCR에 대한 주형으로 사용하였다. 다음 프라이머 쌍을 사용하였다: *MGMT* 정방향(서열번호: 220), 및 역방향(서열번호: 221).

[1208] **실시간 역전사 PCR:** 총 RNA를 TRIzol™ 시약(Thermo Fisher Scientific)을 사용하여 $5x10^6$ 개의 분화된 HUDEP-2 세포 또는 100 μL 혈액으로부터 추출한 후 페놀-클로로포름으로 추출하였다. QuantiTect 역전사 키트(Qiagen) 및 power SYBR™ green PCR 마스터 믹스(Thermo Fisher Scientific)를 사용하였다. 실시간 정량적 PCR을 StepOnePlus 실시간 PCR 시스템(AB Applied Biosystems)에서 수행하였다. 다음 프라이머 쌍을 사용하였다: 마우스 *RPL10*(하우스키핑) 정방향(서열번호: 189), 및 역방향(서열번호: 190); 인간 γ-글로빈 정방향(서열번호: 191), 및 역방향(서열번호: 192); 인간 β-글로빈 정방향(서열번호: 216), 및 역방향(서열번호: 217); 마우스 β-주요 글로빈 정방향(서열번호: 193), 및 역방향(서열번호: 194), 마우스 α 글로빈 정방향(서열번호: 212), 및 역방향(서열번호: 213).

[1209] **염기 편집의 검출:** 게놈 DNA를 상기 기재된 바와 같이 단리하였다. *BCL11A* 인핸서 및 *HBG1/2* 프로모터의 표적 부위를 포함하는 게놈 분절을 다음 프라이머를 사용하여 KOD Hot Start DNA 폴리머라제(MilliporeSigma)로 증폭시켰다: *HBG1* 정방향(서열번호: 31), 역방향(서열번호: 33); *HBG2* 정방향(서열번호: 69), 역방향(서열번호: 72); 및 상기 제시된 *BCL11A* 프라이머. 앰플리콘을 NucleoSpin Gel & PCR 정리 키트(Takara)를 사용하여 정제

하고 다음 프라이머로 서열분석하였다: *HBG1-seq*(서열번호: 105); *HBG2-seq*(서열번호: 237); 및 *BCL11A-seq*(서열번호: 247). 염기 편집 수준을 EditR 1.0.9를 사용하여 Sanger 서열분석 결과로부터 정량화하였다(Kluesner 등, *CRISPR J* 1: 239-250, 2018).

[1210] **동물 연구:** 동물을 수반하는 모든 실험은 워싱턴 대학에 의해 제시된 기관 지침에 따라 수행하였다. 워싱턴 대학은 실험실 동물 관리 국제 평가 및 인증 협회(AALAC) 공인 연구 기관이며 이 대학에서 수행된 모든 살아있는 동물 작업은 실험실 동물 복지국(OLAW) 공중 보건 보증(PHS) 정책, USDA 동물 복지법 및 규정, 실험 동물 관리 및 사용에 대한 지침 및 기관 동물 관리 및 사용 위원회(IACUC) 정책을 따른다. 연구는 워싱턴 대학 IACUC(프로토콜 번호 3108-01)에 의해 승인되었다. 인간 CD46 게놈 유전자좌를 함유하고 인간과 유사한 수준 및 패턴으로 CD46 발현을 제공하는 C57BL/6J 기반 유전자이식 마우스(hCD46$^{+/+}$ 마우스)는 이전에 기재되었다(Kemper 등, *Clin Exp Immunol* 124: 180-189, 2001). 야생형 248 kb β-글로빈 유전자좌 효모 인공 염색체(β-YAC)를 보유하는 유전자 이식 마우스를 사용하였다(Peterson 등, *Ann N Y Acad Sci* 850: 28-37, 1998). β-YAC 마우스를 인간 CD46$^{+/+}$ 마우스와 교배하여 생체내 HSPC 형질도입 연구를 위한 β-YAC$^{+/-}$/CD46$^{+/+}$ 마우스를 수득하였다. 다음 프라이머를 마우스의 유전형 분석에 사용하였다: *CD46* 정방향(서열번호: 233) 및 역방향(서열번호: 234); β-YAC(γ-글로빈 프로모터) 정방향(서열번호: 242) 및 역방향(서열번호: 243).

[1211] *HSPC 동원 및 생체내 형질도입:* HSPC를 인간 재조합 G-CSF의 피하(SC) 주사(5 μg/마우스/일, 4 일) 이어서 5 일째에 AMD3100의 SC 주사(5 mg/kg)에 의해 마우스에 동원하였다. 또한, 동물은 바이러스 주사 16 시간 및 2 시간 전에 덱사메타손(10 mg/kg, IP)을 받았다. AMD3100 후 30 및 60 분에, 동물에게 2 개 용량의 바이러스(4x10$^{10}$ vp/용량 x 2 회 용량)를 안와하정맥총을 통해 바이러스 벡터로 정맥내로 주사하였다. 염기 편집 및 SB 바이러스를 1:1 비로 공동 전달하였다.

[1212] *생체내 선택:* 선택을 형질도입 후 1 주차(Townes 모델) 또는 4 주차(β-YAC 모델)에 시작하였다. 마우스에게 O$^6$-BG(15 mg/kg, IP)를 2 회, 30 분 간격으로 주사하였다. O$^6$-BG의 두번째 주사 1 시간 후, 마우스에게 5 mg/kg BCNU를 주사하였다(IP). 선택의 첫번째 차례 후 2 및 4 주차에, 2 차례 이상을 각각 7.5 및 10 mg/kg의 BCNU 용량으로 수행하였다.

[1213] *2차 골수 이식:* 수용자는 Jackson Laboratory로부터 수득된 6 - 8 주령의 암컷 C57BL/6J 마우스였다. 이식 당일에, 수용자 마우스를 1000 Rad로 조사하였다. 생체내 형질도입된 CD46tg 마우스로부터 골수 세포를 무균상태로 단리하고 계통-고갈된 세포를 상기 기재된 바와 같이 MACS를 사용하여 단리하였다. 조사 6 시간 후 세포를 마우스 당 1x10$^6$ 개의 세포로 정맥내로 주사하였다. 2차 수용자를 종료점 분석을 위해 이식 후 16 주 동안 유지하였다.

[1214] **조직 분석:** 2.5 μm 두께의 비장 및 간 조직 절편을 4% 포름알데하이드에 적어도 24 시간 고정시키고, 탈수하고 파라핀에 포매하였다. 헤마톡실린-에오신으로 염색을 골수외 혈구형성의 조직학적 평가를 위해 사용하였다. 헤모시데린을 Perl의 프러시안 블루 염색에 의해 조직 절편에서 검출하였다. 간단히 말해서, 조직 절편을 증류수 중 동일한 부피(2%)의 페로시안화칼륨 및 염산의 혼합물로 처리한 다음 뉴트럴 레드로 대조염색하였다.

[1215] **혈액 분석:** 혈액 샘플을 EDTA-코팅된 튜브에 수집하고 분석을 HemaVet 950FS(Drew Scientific, 코네티컷주 워터베리 소재)에서 수행하였다. 말초 혈액 도말을 Giemsa/May-Grünwald(Merck, 독일 다름슈타트 소재)로 각각 5 및 15 분 동안 염색하였다. 망상적혈구를 브릴리언트 크레실 블루로 염색하였다. 혈액 도말에 대한 망상적혈구를 계수한 조사자들은 샘플 그룹 할당에 대해 맹검이었다. 슬라이드에는 동물 번호만이 보였다(동물당 5 개 슬라이드, 5 개의 무작위 1 cm$^2$ 절편).

[1216] **통계적 분석:** 다중 그룹의 비교를 위해, 다중 비교를 위한 본페로니 사후-테스트와 함께 일원 및 이원 분산 분석(ANOVA)을 이용하였다. 통계적 분석은 GraphPad Prism 버전 6.01(GraphPad Software Inc., 캘리포니아주 라호이아 소재)을 사용하여 수행하였다.

[1217] ***결과. 염기 편집기 및 가이드 RNA의 선택.*** BE4(Komo 등, *Science Advances* 3: eaao4774, 2017), AncBE4max(Koblan 등, *Nature Biotechnology* 36: 843-846, 2018), BE3RA 및 FNLS(Zafra 등, *Nature Biotechnology* 36: 888-893, 2018)를 포함하는 사이티딘 염기 편집기(CBE)의 다중 버전의 편집 활성을 비교하였다. 염기 편집기(BE)를 보편적인 EF1α 프로모터에 의해 서브클로닝하고 구동시켰다. +58 *BCL11A* 인핸서 영역에서 GATAA 모티프를 표적화하는 인간 U6 프로모터 하에 가이드 RNA를 발현하는 두번째 플라스미드(Canver

등, *Nature* 527: 192-197, 2015)를 공동 형질감염에 사용하였다. BE3RA는 293FT 세포에서 더 높은 편집을 나타내었지만(도 131a), AncBE4max 시스템은 절단 검정에 의해 측정된 K562 적혈구 세포에서 가장 높은 활성을 입증하였다(도 131b). 따라서, AncBE4max를 하류 연구에 사용하였다. 아데닌 염기 편집기(ABE)의 경우, David Liu 그룹에 의해 개발된 ABEmax 시스템을 사용하고 AncBE4max에 대한 것과 유사한 접근법을 사용하여 최적화하였다(Koblan 등, *Nature Biotechnology* 36: 843-846, 2018). 또한 xCas9(3.7)-BE4 및 xCas9(3.7)-ABE(7.10) 편집기를 광범위한 PAM 호환성으로 인해 가이드 서열 스크리닝에 사용하였다(Hu 등, *Nature* 556: 57-63, 2018).

[1218] 염기 편집기의 최적의 표적가능한 창은 5' 단부 첫번째 염기를 위치 1로서 계수하는 프로토스페이서의 위치 4-8이다. 단일 가이드 RNA(sgRNA) 서열의 패널을 +58 *BCL11A* 인핸서에서 GATAA 모티프(sgBCL#1 내지 #6)에 특이적이거나 또는 HBG1/2 프로모터에서 다양한 자연 발행 태아 헤모글로빈의 유전적 지속성(HPFH) 돌연변이(sgHBG#1 내지 #6)를 재생성하도록 설계하였다. 서열 및 이의 특이적 표적 모티프/염기는 표 14에 제시되었다. 가이드 서열을 γ-글로빈 발현을 재활성화하는 효능에 대해 적혈구 전구체 세포주 HUDEP-2 세포(Kurita 등, *PloS One* 8: e59890, 2013)에서 테스트하였다. 세포를 형질감염 후 4 일차에 적혈구 분화에 넣었다. 모든 12개의 sgRNA 서열은 CCR5 발현을 표적화하지만 헤모글로빈-관련 유전자는 표적화하지 않는 음성 CBE 대조군과 비교하여 상당한 γ-글로빈 발현을 야기하였다(도 130). sgHBG#2는 분화 후 6 일차에 41% $HbF^+$ 세포를 초래하였다. *HBG* 프로모터에서 BCL11A 결합 부위를 표적화하는 이전에 기재된 CRISPR 벡터를 양성 대조군으로 사용하고 $HbF^+$ 세포의 84%를 생성하였다(Li 등, Blood 131: 2915-2928, 2018). 따라서, sgBCL#1(CBE), sgHBG#1(CBE), sgHBG#2(ABE) 및 sgHBG#4(ABE)를 활성 뿐만 아니라 표적 부위의 다양성을 고려하여 바이러스 벡터 전달을 위해 선택하였다. 음성 대조군 벡터 sgNeg(CBE) 및 sgHBG#1 및 sgBCL#1을 둘 다 함유하는 벡터(Dual, CBE)를 또한 구축하였다.

[1219] **BE를 발현하는 헬프-의존적 아데노바이러스 벡터(HDAd)의 생성.** 다음으로, 목표는 효율적인 생체내 BE 전달을 위한 바이러스 벡터를 생성하는 것이었다. 필요한 조절 요소를 함유하는 8 kb 초과 크기의 염기 편집기로 인해, 하나의 렌티바이러스 벡터(LV) 또는 아데노 연관 벡터(AAV)에 맞추기가 어렵다. 조혈 줄기 세포(HSC)의 효율적인 형질도입을 위해 HDAd5/35++라고 불리는 변형된 섬유를 갖는 HDAd 벡터를 개발하였다(Li 등, *Mol Ther Methods Clin Dev* 9: 142-152, 2018). HDAd 벡터는 36 kb 패키징 용량을 수용하여, BE 구성요소를 위한 충분한 공간을 제공할 수 있다. 첫번째 시도에서, BE 효소(CBE의 경우 rAPOBEC1-nCas9-2xUGI 또는 ABE의 경우 2xTadA-nCas9)를 EF1α 프로모터 하에 배치하였다. 인간 U6 프로모터에 의해 구동된 sgRNA를 포함하는 전체 BE 구성요소를 HDAd 벡터 플라스미드 pHCA에 클로닝하였다. *FRT* 및 트랜스포존 부위에 의해 플랭킹된 MGMT/GFP 카세트를 또한 벡터로 클로닝하여 O6BG/BCNU 처리에 의해 형질도입된 세포의 선택을 용이하게 하였다(도 132a 및 132b). 특히, BE 구성요소를 트랜스포존의 외부에 배치하였다. 이 설계는 i) MGMT/GFP의 통합된 발현을 유지하면서 BE의 일시적 발현을 허용하고; ii) 잠자는 미녀 트랜스포사제를 발현하는 또 다른 벡터(HDAd-SB)와 공동감염 시 편집 효소의 보다 빠른 분해를 허용한다(벡터 설계의 특정 측면의 추가 논의 및/또는 추가적인 예시를 위해, 실시예 3을 또한 참조한다). 3-리터 스피너 당 수율은 상대적으로 낮았지만(평균 $1x10^{12}$ 개의 바이러스 입자 또는 vp), 4 개의 CBE 벡터를 모두 구제하였다. 이는 뉴클레아제 발현을 조절하는 메커니즘 없이 구제가 능하지 않은 HDAd-CRISPR 벡터와 대조적이다(Saydaminova 등, *Mol Ther Methods Clin Dev* 1: 14057, 2015). 결과는 DSB-무함유 BE 시스템이 CRISPR/Cas9보다 HDAd 생산자 세포에 독성이 적을 수 있음을 시사하였다. ABE 벡터의 경우, 바이러스가 재배열된 것으로 보였고 CsCl 구배를 사용한 초원심분리 후 별개의 HDAd 밴드는 관찰되지 않았다. ABE 및 CBE 벡터 사이의 주요 차이는 데아미나제 영역이므로, ABE 벡터에서 2 개의 TadA-32aa 반복부는 원인 요소였을 가능성이 있었다. 따라서, ABE 벡터의 첫번째 버전에 대해 다음 변형이 이루어졌다: i) 2 개의 TadA-32aa 반복부 사이의 서열 반복성은 대안적인 코돈 용법에 의해 추가로 감소되었다(도 132c); ii) PGK 프로모터를 사용하여 BE 효소를 구동하였다. HSC에서 구성하는 동안(Li 등, *Cancer Res* 80: 549-560, 2020), PGK 프로모터는 116 생산자 세포에서 Ef1α보다 더 낮은 유전자 발현을 구동하여(Qin 등, *PloS One* 5: e10611, 2010), 잠재적인 TadA-연관 부작용을 제거한다; iii) miR183/218-기반 유전자 조절 시스템을 활용하여 BE 발현을 추가로 제어하였다(Saydaminova 등, *Mol Ther Methods Clin Dev* 1: 14057, 2015)(도 133a). 최적화된 설계의 이러한 두번째 버전의 작제물은 정상 수율 범위 내에 있는 $3.3x10^{12}$ vp/스피너의 평균 수율로 2 개의 HDAd-ABE 바이러스를 성공적으로 구제하였다(도 133b).

[1220] 다음으로 HDAd 벡터를 HUDEP-2 세포에서 조사하였다. 테스트된 5 개의 벡터는 모두 표적 염기 전환을 효율적으로 설치하고 실질적인 γ-글로빈 재활성화로 이어졌다(도 133 및 도 134). 일시적 형질감염에 의한 스크리닝 데이터와 일치하게, HDAd-ABE-sgHBG#2 벡터는 최고 수준의 $HbF^+$ 세포를 유도하였다(MOI 1000 vp/세포에서 71%).

흥미롭게도, sgBCL#1 및 sgHBG#1 단독은 각각 17% 및 39% $HbF^+$ 세포를 매개하였지만, sgBCL#1 및 sgHBG#1을 동시에 발현하는 이중 표적화 벡터는 sgHBG#2의 수준과 비슷한 수준에서 HbF 유도를 생성하였으며(도 133c), 이는 상승 효과를 나타낸다. 음성 대조군 벡터에 대해 유의한 HbF 유도는 측정되지 않았다. HPLC에 의해 측정된 γ-글로빈 단백질 수준은 유세포 분석 데이터와 일치하였다. 인간 β-글로빈에 비해 인간 γ-글로빈의 23%가 sgHBG#2로 형질도입 후 관찰되었으며, 이는 상당한 전환을 입증한다(도 133e 및 133h). MOI 1000에서, 4 개의 sgRNA에 대한 염기 전환 빈도는 25 - 51% 범위 내에 있었다(도 133d 및 도 134a). sgHBG#2의 경우, 위치 5 및 8에서 각각 40% 및 34% A > G 전환이 검출되었다(도 133d). $A_8$에서 G로의 전환은 -113A > G HPFH 돌연변이를 시뮬레이션하였다(표 14)(Martyn 등, *Blood* 133(8):852-856, 2019). *HBG1* 및 *HBG2* 사이에 상당한 편집 차이는 발견되지 않았다. 단일 세포 유래 클론에서, $A_5$ 및 $A_8$ 부위에서 단일대립유전자성 편집은 HbF-양성 세포의 100%를 부여하였으며(도 133f 및 133g), 이는 HbF 억제를 조절하는 데 있어서 이들 부위의 중요한 역할을 확인하였다. sgHBG#1 및 sgHBG#4로부터 유래된 클론에서 유사한 결과가 제시되었다. sgBCL#1로 형질도입된 클론에서, BCL11A 인핸서의 GATAA 모티프에서 이중대립유전자성 G > A 돌연변이는 15% HbF-발현 세포를 야기하였다(도 134b 및 134c). 종합하여, 이들 데이터는 *BCL11A* 인핸서 또는 *HBG1/2* 프로모터에서 중요한 부위에 특이적인 HDAd-BE 벡터가 HbF 발현을 효율적으로 재활성화할 수 있음을 입증한다.

[1221] **염기 편집기를 사용한 생체내 형질도입 후 β-YAC 마우스에서 γ-글로빈의 재활성화.** HDAd5/35++ 벡터를 사용한 HSC의 생체내 형질도입에 의해 단순화된 유전자 요법 접근법이 달성되었다(Richter 등, Blood 128: 2206-2217, 2016). 따라서, 이러한 신규 생체내 전략법으로 염기 편집의 효능을 조사하였다. 완전한 82 kb β-글로빈 유전자좌를 포함하는 248 kb의 인간 DNA를 함유하는 β-YAC 마우스를 사용하였다(Peterson 등, *PNAS USA* 90: 7593-7597, 1993). 마우스를 HDAd5/35++ 벡터로의 형질도입을 허용하도록 인간 CD46 유전자이식 마우스와 교배하였다. HDAd-ABE-sgHBG#2는 HUDEP-2 세포에서 γ-글로빈 발현을 유도하는 가장 높은 효능을 인해 선택하였다. G-CSF/플레릭사포르로 동원 후, β-YAC/CD46 마우스에게 HDAd-ABE-sgHBG#2 및 HDAd-SB 벡터를 정맥내로 주사하였다. 형질도입 4 주 후, 마우스를 4 차례의 $O^6$BG/BCNU($O^6$-벤질구아닌/카르무스틴) 처리에 적용하여 통합된 MGMT-GFP 이식유전자를 갖는 전구체를 선택적으로 확장시켰다(도 135a). 선택 후, PBMC에서 GFP 마킹은 60%에 도달하였다(도 135b 및 135c). 특히, 말초 혈액 세포에서 γ-글로빈 발현은 형질도입 전 1%에서 형질도입 후 16 주차에 평균(n = 9) 43%까지 증가하였으며, 이는 상당한 γ-글로빈 재활성화를 입증한다(도 135d 및 135e). 상이한 마우스 사이에 존재하는 큰 변이는 아마도 MGMT의 더 낮은 발현을 초래할 수 있는 MGMT-2A-GFP의 바이시스트론 설계에 의해 유발되고 따라서 생체내 선택 효능에 영향을 미쳤다. γ-글로빈$^+$ 세포는 주로 골수 샘플의 두 혈액에서 적혈구(RBC) 분획(Ter-$119^+$)에 상주하였다(도 135f). 16 주차에 RBC 용해물에서, 인간 β-글로빈 단백질에 비해 γ-글로빈의 최대 21%가 고성능 액체 크로마토그래피(HPLC)에 의해 측정되었다(도 135g 및 도 136). γ-글로빈 mRNA 발현은 HPLC 데이터와 일치하였다(도 135h). 16 주차에 총 골수 단핵 세포에서, 통합 벡터 카피 수는 세포 당 최대 2.5 개 카피였다(평균 1.4 개)(도 135I).

[1222] *HBG1/2* 프로모터에서 염기 편집을 분석하였다. *HBG1* 및 *HBG2*의 $A_5$ 및 $A_8$ 부위에서 A > G 전환 빈도는 평균 15-30%였다(도 137a-137c). 염기 편집 빈도는 γ-글로빈 발현 수준과 밀접하게 상관관계가 있는 것으로 밝혀졌다(Pearson 검정, R = 0.92, p < 0.001) (도 137d). γ-글로빈 발현이 가장 높은 마우스에서, 82% 표적 염기 전환이 달성되었다(도 137b). 참고로, $A_5$에서의 전환율%은 *HBG1* 및 *HBG2* 영역 둘 다의 $A_8$ 부위에서보다 약간 더 높은 경향이 있었지만, 통계적 차이는 발견되지 않았다(도 137b). 일부 염기 편집기는 다중 표적이 프로토스페이서에 존재할 때 진행적 편집을 나타내는 것으로 제시되었다. 그러나, $A_9$ 부위에서 편집은 발견되지 않았다(도 137a 및 137c). 이는 위치 9가 최적 편집 창 외부에 위치하기 때문이며, 이는 편집 창의 협소함을 입증할 가능성이 있었다.

[1223] 요약하면, 이들 데이터는 *HBG1/2* 프로모터에 특이적인 염기 편집기를 사용한 생체내 형질도입 이어서 선택이 β-YAC/CD46 마우스에서 효율적인 표적 염기 전환 및 γ-글로빈 재활성화를 야기함을 입증한다.

[1224] **생체내 HSC 염기 편집 후 우수한 안전성 프로파일 및 안정된 효능.** 16 주차에, 동물을 안락사시키고 조직 샘플을 다중 혈액학 및 조직학 분석에 적용하였다. 백혈구(K/μL), 적혈구(M/μL), Hb(g/dL), MCV(fL), MCHC(g/dL), RDW(%) 및 혈소판(K/μL)을 포함하는 혈액학적 매개변수는 미처리 β-YAC/CD46 마우스의 것과 유사하였다(도 138a 및 138b). 브릴리언트 크레실 블루 염색에 의해 측정된 말초 혈액 내 망상적혈구의 백분율은 처리되지 않은 마우스와 비슷하였다(도 138d). 비장 및 간 절편에서 골수외 적혈구생성의 병소는 관찰되지 않

았다. PBMC, 비장 및 골수 단핵 세포에서 세포 조성물은 대조군 마우스와 구별할 수 없는 것으로 입증되었다(도 138c). 게다가, 이전에 보고된 다른 유전자 요법 벡터와 비교하여(Li 등, *Blood* 131: 2915-2928, 2018; Wang 등, *J Clin Invest.* 129(2): 598-615, 2018; Li 등, *Molecular Therapy* 27: 2195-2212, 2019), HDAd-ABE-sgHBG#2는 생체내 형질도입/선택 후 체중, 행동 및 외형의 명백한 변화를 유발하지 않았다.

[1225] 생체내 형질도입이 장기간 재증식 HSPC에서 발생하였음을 입증하기 위해, 16 주차에 수확된 골수 계통 마이너스($Lin^-$) 세포를 형질도입 후 치명적으로 조사된 C57BL/6J 마우스(인간 CD46 유전자 없음)에 이식하였다. 2차 수용자에서 다중-계통 재구축을 구동하는 이식된 세포의 능력을 16 주의 기간에 걸쳐 평가하였다. PBMC에서 huCD46 발현에 기반한 생착률은 95% 초과였고 안정되게 유지하였다(도 139a). PMBC의 GFP 마킹은 1차 마우스의 것과 비슷하였다(도 139b). γ-글로빈$^+$ RBC의 퍼센트는 평균 40%였고 안정되었다(도 139c).

[1226] 이러한 관찰과 함께 생체내 HSC 염기 편집은 전반적으로 안전하였음이 입증되었다. 변형된 HSPC는 장기간 지속되었고 안정된 이식유전자 발현으로 2차 수용자 마우스를 재구축할 수 있었다.

[1227] **최소 유전자간 결실 및 상위-점수 표적외 부위에서 검출가능하지 않은 편집.** DSB-의존적 유전자 편집 전략의 균형은 잠재적인 게놈 큰-단편 결실이다(Kosicki 등, *Nature Biotechnology* 36: 765, 2018). DSB-생성 뉴클레아제에 의해 *HBG1/2* 프로모터를 표적화하는 경우, 이 부작용은 *HBG1* 및 *HBG2* 영역 사이의 높은 유사성으로 인해 더 유의하게 될 수 있다. 2 개 영역 중 하나에 특이적인 가이드 서열은 또한 다른 하나를 표적화할 수 있다. CRISPR/Cas9를 사용하여 *HBG1/2* 프로모터에서 *BCL11A* 결합 부위를 표적화하는 것은 4.9 kb 유전자간 결실을 야기함이 보고되었다(Traxler 등, *Nature Medicine* 22: 987-990, 2016; Li 등, *Blood* 131: 2915-2928, 2018). 결과적으로, 전체 *HBG2* 유전자가 제거된다. 따라서, 반-정량적 PCR에 의한 게놈 결실을 조사하였다(Li 등, *Blood* 131: 2915-2928, 2018). 2 개의 표적화 부위에 플랭킹하는 한 쌍의 프라이머를 사용하여 9.9 kb 게놈 분절을 증폭시켰다. 4.9 kb 결실의 존재는 추가로 단축된 5.0 kb PCR 앰플리콘을 생성할 것이다. 결실 백분율은 표준 곡선을 확립함으로써 5.0 kb 대 9.9 kb 앰플리콘의 비와 양의 상관관계가 있었다(Li 등, *Blood* 131: 2915-2928, 2018의 도 7c 참조). 염기 편집기-처리된 마우스에서 평균 4.9 kb 결실은 1% 미만이었던 것으로 밝혀졌다(도 140). 일부 마우스에서, 거의 검출할 수 없었다. 이는 HDAd-HBG-CRISPR 벡터를 사용한 형질도입으로부터 유래된 것보다 상당히 더 낮았다(Li 등, *Blood* 131: 2915-2928, 2018).

[1228] 다음으로, 표적외 분석을 수행하여 시스템의 충실도를 조사하였다. 인 실리코 분석은 인간 및 마우스 게놈 둘 다에서 가이드 서열에 대한 2 개 이하의 염기 쌍(bp) 불일치가 있는 잠재적인 표적외 부위가 없음을 나타내었다. 인간 및 마우스에서는 3 bp 불일치가 있는 각각 10 및 2 개의 잠재적인 표적외가 있었다. 이러한 예측된 표적에서 표적외 편집의 가능성은 모든 부위가 프로토스페이서의 PAM-근위 절반에서 적어도 1 bp 불일치를 보유하기 때문에 낮았던 것으로 추측되었다. 4 bp 불일치로 인해, 인간 및 마우스에서 각각 79 및 74 개의 잠재적 표적이 반환되었다. 연구를 마우스에서 수행한 이후, 10 개의 상위-점수의 게놈 부위(2 개는 3 bp 불일치; 7 개는 4 bp 불일치)를 표적내 염기 설치가 가장 높은 마우스로부터 증폭시키고 이어서 Sanger 서열분석하였다. 이들 부위 중 어느 것도 검출가능한 편집을 나타내지 않았다.

[1229] 종합하면, 이들 데이터는 최소 유전자간 결실 및 생체내 염기 편집 시스템의 높은 충실도에 대한 증거를 제공하였다.

[1230] **실시예 11. 염기 편집기 구현예에 관한 추가 설명**

[1231] 도 141은 혈액학 분석(도 141a) 및 골수 MNC에서 세포 비교(도 141b)를 포함하는 염기 편집기의 안전성 프로파일을 제시한다. 염기 편집기 BE4-sgBCL11AE1의 활성으로 생성될 것으로 예상되는 편집의 예는 도 142에 제시되어 있다. 도 143은 C에서 T로(상단 이미지) 또는 G에서 A로(하단 이미지) 염기 변환을 수행할 때 염기 편집 효율을 최대화하기 위한 최적의 프로토스페이서 서열 배열을 나타낸다. 도 144는 표적 C가 프로토스페이서의 위치 4 내지 8에 있을 때 C에서 T로의 편집을 위한 벡터를 나타낸다. 도 145는 단일 인접한 작제물을 나타내지만 제시의 용이성만을 위해 2 개의 섹션으로 나눠진 바이러스 gDNA(HBG2-miR, 아데닌 편집기)의 다이어그램을 나타낸다. 도 146은 TadA 및 TadA*의 서열을 나타낸다. Sanger 서열분석을 수행하여 서열의 염기 편집을 확인하였다(도 147). 도 148은 HDAd5/35++_BE4-sgBCL11Ae1-FI-mgmtGFP(041318-1) 바이러스에 의한 염기 편집을 나타내고, 도 149는 표시된 MOI에서 γ-글로빈+ 세포의 백분율을 나타낸다. 도 150은 염기 편집에 의한 HbF의 재활성화를 위한 사이티딘 염기 편집기 및 아데닌 염기 편집기를 나타낸다. 도 151은 예시적인 염기 편집기 및 염기 편집기의 다양한 MOI에서 퍼센트 HbF+ 세포를 나타낸다. 도 152는 HUDEP-2 세포에서 2차 시험의 % HbF+를 나타낸다. 도 153은 단일 세포 유래 클론의 결과를 나타낸다. 도 154a-154s는 개별 단일 세포 유래 클론을 나타내

는 데이터를 제시한다. 또한 염기 편집기를 293FT 세포에서 테스트하였다(도 155). 도 156a-156d는 sanger 서열분석 결과를 나타낸다. 또한 염기 편집기를 HUDEP-2 세포에서 테스트하였다(도 157). γ-글로빈의 발현은 도 158에 제시되어 있다. 도 159a-159d는 이용가능한 경우 sanger 서열분석 결과를 나타낸다. 작제물은 도 160에 제시된 바와 같이 Maxi 제조를 위해 선택하였다.

[1232] 예를 들어, HDAd-AAVS1-CRISPR 또는 HDAd-글로빈-BE4 염기 편집기로 편집된 huCD45+ 세포의 생착은 도 161에 제시되어 있다.

[1233] HUDEP-2 세포의 일시적 형질감염(T7EI에 의한 절단)은 도 162에 제시되어 있다.

[1234] HbF에 대한 염기 편집 작제물의 비제한적인 예는 (1) pHCA-ABEmax-sgHBG2-miR-FI-mgmtGFP; (2) pHCA-ABEmax-sgHBG4-miR-FI-mgmtGFP; 또는 (3) pHCA-ABEmax-Dual-Skip-miR-FI-mgmtGFP를 포함할 수 있다.

[1235] 염기 편집기의 적어도 하나의 적용은 이중 염기 편집 벡터를 포함하며, 이 적용은 도 163에 예시되어 있다.

[1236] 단일 세포 유래 클론에서, 단일대립유전자성 또는 이중대립유전자성 표적 염기 전환은 HbF-양성 세포의 100%를 부여하였다. 혼합된 HUDEP-2 세포의 HBG1/2 프로모터에서 60% -113 A에서 G로의 HPFH 돌연변이를 ABE 벡터 HDAd-ABE-HBG#2를 사용하여 관찰하였다(도 135 참조). 이 벡터는 특정 추가 동물 연구를 위해 선택되었다. 동물 연구를 248 kb의 인간 β-글로빈 유전자좌(β-YAC 마우스)를 보유하고 따라서 글로빈 전환을 정확하게 반영하는 마우스에서 수행하였다(예를 들어, 도 137 참조). FRT 및 트랜스포존 부위에 의해 플랭킹된 EF1α-$mgmt^{P140K}$ 발현 카세트는 형질도입된 세포의 생체내 선택을 허용하기 위해 벡터에 포함되었다(예를 들어, 도 136 참조). HDAd-ABE-HBG#2 + HDAd-SB를 사용한 생체내 형질도입 및 저용량의 $O^6$BG/BCNU로의 선택 후, HbF-양성 세포의 평균 35%가 말초 적혈구에서 측정되었다(도 138). 8 마리 마우스 중 1 마리에서, 거의 완전한 -113 A에서 G로의 전환 및 HbF-양성 세포의 90%가 달성되었다. 혈액 세포 계수의 변이는 발견되지 않았다(도 141). 골수 샘플의 세포 조성물은 형질도입되지 않은 마우스의 것과 비슷하였으며, 이는 우수한 안전성 프로파일을 입증한다(도 141). 골수 계통 마이너스 세포를 형질도입 후 14 주차에 1차 마우스로부터 단리하고 치명적으로 조사된 C57BL/6J 마우스에 주입하였다. HbF-양성 세포의 백분율은 2차 수용자에서 16 주에 걸쳐 유지되었으며 이는 게놈 편집이 장기간 재증식 마우스 HSC에서 발생하였음을 나타낸다. 이러한 관찰은 생체내에서 HDAd5/35++ 벡터에 의해 전달된 염기 편집기가 예를 들어, 혈색소병증의 치료를 위해 정확한 게놈 조작을 위한 전략임을 입증한다.

[1237] **VII. 마무리 단락**

[1238] 본원에 개시되고 참조된 서열의 변이체가 또한 포함된다. 아미노산 잔기가 생물학적 활성을 제거하지 않으면서 치환, 삽입, 또는 결실될 수 있는지를 결정하는 지침은 DNASTAR™(위스콘신주 매디슨 소재) 소프트웨어와 같이 당업계에 잘 알려진 컴퓨터 프로그램을 사용하여 찾을 수 있다. 바람직하게는, 본원에 개시된 단백질 변이체에서 아미노산 변화는 보존적 아미노산 변화, 즉, 유사하게 하전 또는 비하전된 아미노산의 치환이다. 보존적 아미노산 변화는 측쇄와 관련된 아미노산 패밀리 중 하나의 치환을 수반한다.

[1239] 펩티드 또는 단백질에서, 아미노산의 적합한 보존적 치환은 당업자에게 알려져 있고 일반적으로 생성된 분자의 생물학적 활성을 변경하지 않으면서 이루어질 수 있다. 당업자는 일반적으로, 폴리펩티드의 비필수 영역에서의 단일 아미노산 치환이 생물학적 활성을 실질적으로 변경하지 않음을 인식한다(예를 들어, Watson 등, Molecular Biology of the Gene, 4th Edition, 1987, The Benjamin/Cummings Pub. Co., p. 224 참조). 자연 발생 아미노산은 일반적으로 다음과 같은 보존적 치환 패밀리로 나눠진다: 그룹 1: 알라닌(Ala), 글리신(Gly), 세린(Ser), 및 트레오닌(Thr); 그룹 2: (산성): 아스파르트산(Asp), 및 글루탐산(Glu); 그룹 3: (산성; 또한 극성, 음으로 하전된 잔기 및 이의 아미드로 분류됨): 아스파라긴(Asn), 글루타민(Gln), Asp, 및 Glu; 그룹 4: Gln 및 Asn; 그룹 5: (염기성; 또한 극성, 양으로 하전된 잔기로 분류됨): 아르기닌(Arg), 리신(Lys), 및 히스티딘(His); 그룹 6 (큰 지방족, 비극성 잔기): 이소류신(Ile), 류신(Leu), 메티오닌(Met), 발린(Val) 및 시스테인(Cys); 그룹 7(비하전된 극성): 티로신(Tyr), Gly, Asn, Gln, Cys, Ser, 및 Thr; 그룹 8(큰 방향족 잔기): 페닐알라닌(Phe), 트립토판(Trp), 및 Tyr; 그룹 9(비-극성): 프롤린(Pro), Ala, Val, Leu, Ile, Phe, Met, 및 Trp; 그룹 11(지방족): Gly, Ala, Val, Leu, 및 Ile; 그룹 10 (작은 지방족, 비극성 또는 약간 극성 잔기): Ala, Ser, Thr, Pro, 및 Gly; 및 그룹 12(황-함유): Met 및 Cys. 추가적인 정보는 Creighton (1984) Proteins, W.H. Freeman and Company에서 찾을 수 있다.

[1240] 이러한 변화를 만들기 위해, 아미노산의 소수성 지수(hydropathic index)가 고려될 수 있다. 단백질에 대한 상

호작용성 생물학적 기능을 부여하는 데 있어서 소수성 아미노산 지수의 중요성은 일반적으로 당업계에서 이해된다(Kyte & Doolittle, *J. Mol. Biol.* 157(1), 105-32, 1982). 각 아미노산은 소수성 및 전하 특성에 기반하여 소수성 지수가 할당되었다(Kyte and Doolittle, 1982). 이들 값은 다음과 같다: Ile(+4.5); Val(+4.2); Leu(+3.8); Phe(+2.8); Cys(+2.5); Met(+1.9); Ala(+1.8); Gly(-0.4); Thr(-0.7); Ser(-0.8); Trp(-0.9); Tyr(-1.3); Pro(-1.6); His(-3.2); 글루타메이트(-3.5); Gln(-3.5); 아스파르테이트(-3.5); Asn(-3.5); Lys(-3.9); 및 Arg(-4.5).

[1241] 특정 아미노산은 유사한 소수성 지수 또는 점수를 갖는 다른 아미노산으로 치환될 수 있고 여전히 유사한 생물학적 활성을 갖는 단백질을 초래하며, 즉, 여전히 생물학적 기능적으로 동등한 단백질을 수득할 수 있음이 당업계에 알려져 있다. 이러한 변화를 만드는 데 이어서, 소수성 지수가 ±2 이내인 아미노산의 치환이 바람직하고, ±1 이내의 것이 특히 바람직하고, ±0.5 이내의 것이 더욱 더 특히 바람직하다. 또한 유사 아미노산의 치환이 친수성에 기반하여 효과적으로 이루어질 수 있음이 당업계에서 이해된다.

[1242] 미국 특허 번호 제4,554,101호에 상세히 기재된 바와 같이, 다음 친수성 값이 아미노산 잔기에 할당되었다: Arg(+3.0); Lys(+3.0); 아스파르테이트(+3.0±1); 글루타메이트(+3.0±1); Ser(+0.3); Asn(+0.2); Gln(+0.2); Gly(0); Thr(-0.4); Pro(-0.5±1); Ala(-0.5); His(-0.5); Cys(-1.0); Met(-1.3); Val(-1.5); Leu(-1.8); Ile(-1.8); Tyr(-2.3); Phe(-2.5); Trp(-3.4). 아미노산은 유사한 친수성 값을 갖는 또 다른 아미노산으로 치환되고 여전히 생물학적으로 동등한 단백질, 특히 면역학적으로 동등한 단백질을 수득할 수 있음이 이해된다. 이러한 변화에서, 친수성 값이 ±2 이내인 아미노산의 치환이 바람직하고, ±1 이내의 것이 특히 바람직하고, ±0.5 이내의 것이 더욱더 특히 바람직하다.

[1243] 상기 요약된 바와 같이, 아미노산 치환은 아미노산 측쇄 치환기의 상대적 유사성, 예를 들어, 소수성, 친수성, 전하, 크기 등에 기반할 수 있다.

[1244] 다른 곳에 나타낸 바와 같이, 유전자 서열의 변이체는 코돈 최적화된 변이체, 서열 다형성, 스플라이스 변이체, 및/또는 암호화된 생성물의 기능에 통계적으로 유의한 정도로 영향을 미치지 않는 돌연변이를 포함할 수 있다.

[1245] 본원에 개시된 단백질, 핵산, 및 유전자 서열의 변이체는 또한 본원에 개시된 단백질, 핵산, 또는 유전자 서열에 대해 적어도 70% 서열 동일성, 80% 서열 동일성, 85% 서열, 90% 서열 동일성, 95% 서열 동일성, 96% 서열 동일성, 97% 서열 동일성, 98% 서열 동일성, 또는 99% 서열 동일성을 갖는 서열을 포함한다.

[1246] "% 서열 동일성"은 서열을 비교함으로써 결정된 바와 같은 2 개 이상의 서열 사이의 관계를 지칭한다. 당업계에서, "동일성"은 또한 이러한 서열의 스트링 사이의 일치에 의해 결정된 바와 같은 단백질, 핵산, 또는 유전자 서열 사이의 서열 관련성 정도를 의미한다. "동일성"(종종 "유사성"으로 언급됨)은 다음 문헌에 기재된 것들을 포함하여 알려진 방법에 의해 용이하게 계산될 수 있다: Computational Molecular Biology (Lesk, A. M., ed.) Oxford University Press, NY (1988); Biocomputing: Informatics and Genome Projects (Smith, D. W., ed.) Academic Press, NY (1994); Computer Analysis of Sequence Data, Part I (Griffin, A. M., and Griffin, H. G., eds.) Humana Press, NJ (1994); Sequence Analysis in Molecular Biology (Von Heijne, G., ed.) Academic Press (1987); 및 Sequence Analysis Primer (Gribskov, M. and Devereux, J., eds.) Oxford University Press, NY (1992). 동일성을 결정하는 바람직한 방법은 테스트된 서열 사이의 최고 일치를 제공하도록 설계된다. 동일성 및 유사성을 결정하는 방법은 공개적으로 이용가능한 컴퓨터 프로그램에 의해 성문화된다. 서열 정렬 및 퍼센트 동일성 계산은 LASERGENE 생물정보학 컴퓨터 묶음의 Megalign 프로그램(DNASTAR, Inc., 위스콘신주 매디슨 소재)을 사용하여 수행될 수 있다. 서열의 다중 정렬은 또한 디폴트 매개변수(갭 패널티=10, 갭 길이 패널티=10)에 따른 Clustal 정렬 방법(Higgins and Sharp CABIOS, 5, 151-153 (1989)을 사용하여 수행될 수 있다. 관련 프로그램은 또한 GCG 프로그램 묶음(Wisconsin Package Version 9.0, Genetics Computer Group(GCG), 위스콘신주 매디슨 소재); BLASTP, BLASTN, BLASTX(Altschul, 등, J. Mol. Biol. 215:403-410 (1990); DNASTAR(DNASTAR, Inc., 위스콘신주 매디슨 소재); 및 Smith-Waterman 알고리즘을 포함하는 FASTA 프로그램(Pearson, Comput. Methods Genome Res., [Proc. Int. Symp.] (1994), Meeting Date 1992, 111-20. Editor(s): Suhai, Sandor. Publisher: Plenum, New York, N.Y.을 포함한다. 본 개시내용의 맥락 내에서 서열 분석 소프트웨어가 분석에 사용되는 경우, 분석 결과는 참조된 프로그램의 "디폴트 값"에 기반함이 이해될 것이다. 본원에 사용되는 바와 같은 "디폴트 값"은 처음에 초기화될 때 소프트웨어와 함께 원래 로드되는 임의의 값 또는 매개변수의 세트를 의미할 것이다.

[1247] 변이체는 또한 엄격한 혼성화 조건 하에 본원에 개시된 서열에 혼성화하고 참조 서열과 동일한 기능을 제공하는 핵산 분자를 포함한다. 예시적인 엄격한 혼성화 조건은 50% 포름아미드, 5XSSC(750 mM NaCl, 75 mM 트리나트륨

시트레이트), 50 mM 나트륨 포스페이트(pH 7.6), 5XDenhardt의 용액, 10% 덱스트란 술페이트, 및 20 μg/ml 변성되고 전단된 연어 정자 DNA를 포함하는 용액에서 42 ℃에서 밤새 배양한 후, 50℃에서 0.1XSSC에서 충전제를 세척하는 것을 포함한다. 혼성화 및 신호 검출의 엄격성에서의 변화는 주로 포름아미드 농도(포름아미드의 백분율이 낮을수록 낮은 엄격도를 초래함); 염 농도, 또는 온도의 조작을 통해 달성된다. 예를 들어, 적당히 높은 엄격도 조건은 6XSSPE(20XSSPE=3M NaCl; 0.2 M $NaH_2PO_4$; 0.02 M EDTA, pH 7.4), 0.5% SDS, 30% 포름아미드, 100 μg/ml 연어 정자 차단 DNA를 포함하는 용액에서 37℃에서 밤새 배양한 후; 1XSSPE, 0.1% SDS로 50 ℃에서 세척하는 것을 포함한다. 또한, 심지어 더 낮은 엄격도를 달성하기 위해, 엄격한 혼성화 후 수행되는 세척은 더 높은 염 농도(예를 들어 5XSSC)에서 수행될 수 있다. 상기 조건에서의 변이는 혼성화 실험에서 배경을 억제하는 데 사용되는 대체 차단 시약의 포함 및/또는 치환을 통해 달성될 수 있다. 전형적인 차단 시약은 Denhardt의 시약, BLOTTO, 헤파린, 변성된 연어 정자 DNA, 및 상업적으로 이용가능한 독점 제형을 포함한다. 특이적 차단 시약의 포함은 호환성 문제로 인해, 상기 기재된 혼성화 조건의 변형을 필요로 할 수 있다.

[1248] "특이적으로 결합한다"는 (예를 들어, CAR 결합 도메인 또는 나노입자 선택된 세포 표적화 리간드의) 결합 도메인과 이의 동족 결합 분자를 $10^5$ $M^{-1}$ 이상의 친화도 또는 Ka(즉, 1/M 단위의 특정 결합 상호작용의 평형 결합 상수)로 회합하지만, 관련 환경 샘플에서 임의의 다른 분자 또는 구성요소와 유의하게 회합하지 않는 것을 지칭한다. "특이적으로 결합한다"는 또한 본원에서 "결합한다"로 지칭된다. 결합 도메인은 "높은 친화도" 또는 "낮은 친화도"로 분류될 수 있다. 특정 구현예에서, "높은 친화도" 결합 도메인은 적어도 $10^7$ $M^{-1}$, 적어도 $10^8$ $M^{-1}$, 적어도 $10^9$ $M^{-1}$, 적어도 $10^{10}$ $M^{-1}$, 적어도 $10^{11}$ $M^{-1}$, 적어도 $10^{12}$ $M^{-1}$, 또는 적어도 $10^{13}$ $M^{-1}$의 Ka를 갖는 결합 도메인을 지칭한다. 특정 구현예에서, "낮은 친화도" 결합 도메인은 최대 $10^7$ $M^{-1}$, 최대 $10^6$ M-1, 최대 $10^5$ $M^{-1}$의 Ka를 갖는 결합 도메인을 지칭한다. 대안적으로, 친화도는 M 단위(예를 들어, $10^{-5}$ M 내지 $10^{-13}$ M)의 특정 결합 상호작용의 평형 해리 상수(Kd)로서 정의될 수 있다. 특정 구현예에서, 결합 도메인은 야생형(또는 모체) 결합 도메인보다 동족 결합 분자에 더 강한 결합을 갖는 선택된 또는 조작된 결합 도메인을 지칭하는 "향상된 친화도"를 가질 수 있다. 예를 들어, 향상된 친화도는 참조 결합 도메인보다 더 높은 동족 결합 분자에 대한 Ka(평형 결합 상수) 또는 참조 결합 도메인보다 낮은 동족 결합 분자에 대한 Kd(해리 상수), 또는 참조 결합 도메인보다 낮은 동족 결합 분자에 대한 오프레이트($K_{off}$) 때문일 수 있다. 특정 동족 결합 분자에 특이적으로 결합하는 결합 도메인을 검출할 뿐만 아니라 결합 친화성을 결정하기 위해, 웨스턴 블롯, ELISA, 및 BIACORE® 분석과 같은 다양한 검정이 알려져 있다(또한, 예를 들어, Scatchard, 등, 1949, Ann. N.Y. Acad. Sci. 51:660; 및 US 5,283,173, US 5,468,614, 또는 등가물 참조).

[1249] 달리 표시되지 않는 한, 본 개시내용의 실시는 면역학, 분자 생물학, 미생물학, 세포 생물학 및 재조합 DNA의 통상적인 기술을 이용할 수 있다. 이들 방법은 다음 공개물에 기재되어 있다. 예를 들어, Sambrook, 등, Molecular Cloning: A Laboratory Manual, 2nd Edition (1989); F. M. Ausubel, 등, eds., Current Protocols in Molecular Biology, (1987); the series Methods IN Enzymology (Academic Press, Inc.); M. MacPherson, 등, PCR: A Practical Approach, IRL Press at Oxford University Press (1991); MacPherson 등, eds. PCR 2: Practical Approach, (1995); Harlow and Lane, eds. Antibodies, A Laboratory Manual, (1988); 및 R. I. Freshney, ed. Animal Cell Culture (1987) 참조.

[1250] 당업자에 의해 이해되는 바와 같이, 본원에 개시된 각 구현예는 특정 언급된 요소, 단계, 성분 또는 구성요소를 포함하거나, 이로 본질적으로 이루어지거나, 또는 이로 이루어질 수 있다. 따라서, 용어 "포함하다" 또는 "포함하는"은 "포함하거나, 이로 이루어지거나, 또는 이로 본질적으로 이루어지는"을 재인용하는 것으로 해석되어야 한다. 전환 용어 "포함하다" 또는 "포함한다"는 명시되지 않은 요소, 단계, 성분, 또는 구성요소를 포함하나 이에 제한되지 않고, 심지어 다량으로 포함하는 것을 허용함을 의미한다. 전환 어구 "로 이루어지는"은 명시되지 않은 임의의 요소, 단계, 성분 또는 구성요소를 제외한다. 전환 어구 "로 본질적으로 이루어진"은 명시된 요소, 단계, 성분 또는 구성요소에 대한 구현예의 범위 및 구현예에 실질적으로 영향을 미치지 않는 것들을 제한한다. 실질적 효과는 본 개시내용에 기재된 관련 실험 방법에 따라 청구된 효과를 수득하는 능력에서 통계적으로 유의한 감소를 야기할 것이다.

[1251] 달리 표시되지 않는 한, 명세서 및 청구범위에서 사용되는 성분의 양, 분자량과 같은 특성, 반응 조건 등을 나타내는 모든 숫자는 모든 경우에 용어 "약"에 의해 수식되는 것으로 이해되어야 한다. 따라서, 달리 반대로 표시되지 않는 한, 명세서 및 첨부된 청구범위에 제시된 수치 매개변수는 본 발명에 의해 수득하고자 하는 원하는 특성에 따라 달라질 수 있는 근사치이다. 최소한, 및 청구범위의 범위에 대한 등가 원칙의 적용을 제한하려는

시도가 아니라, 각 수치 매개변수는 적어도 보고된 유효 자릿수에 비추어 그리고 반올림 기법을 적용함으로써 해석되어야 한다. 추가로 명확성이 요구되는 경우, 용어 "약" 및 "대략"은 본원에서 상호교환가능하게 사용되며 언급된 수치 값 또는 범위와 함께 사용되는 경우 당업자에 의해 합리적으로 귀속되는 의미를 가지며, 즉, 언급된 값의 ±20%; 언급된 값의 ±19%; 언급된 값의 ±18%; 언급된 값의 ±17%; 언급된 값의 ±16%; 언급된 값의 ±15%; 언급된 값의 ±14%; 언급된 값의 ±13%; 언급된 값의 ±12%; 언급된 값의 ±11%; 언급된 값의 ±10%; 언급된 값의 ±9%; 언급된 값의 ±8%; 언급된 값의 ±7%; 언급된 값의 ±6%; 언급된 값의 ±5%; 언급된 값의 ±4%; 언급된 값의 ±3%; 언급된 값의 ±2%; 또는 언급된 값의 ±1%의 범위 내에서 언급된 값 또는 범위보다 다소 많거나 또는 다소 적은 것을 나타낸다.

[1252] 본 발명의 넓은 의미를 제시하는 수치 범위 및 매개변수가 근사치임에도 불구하고, 특정 실시예에 제시된 수치 값은 가능한한 정확하게 보고된다. 그러나, 임의의 수치 값은 본질적으로 각각의 테스트 측정에서 발견된 표준 편차에서 반드시 비롯되는 특정 오류를 함유한다.

[1253] 본원에서 값의 범위에 대한 언급은 단지 범위 내에 속하는 각각의 개별 값을 개별적으로 언급하는 약칭 방법을 제공하는 것으로 의도된다. 본원에 달리 나타낸 바와 같이, 각각의 개별 값은 본원에서 개별적으로 인용된 것처럼 명세서에 포함된다. 본원에 기재된 모든 방법은 본원에 달리 표시되지 않거나 또는 문맥상 달리 명백하게 모순되지 않는 한 임의의 적합한 순서로 수행될 수 있다. 본원에 제공된 임의의 및 모든 예, 또는 예시적인 언어(예를 들어, "예컨대")의 사용은 단지 본 발명을 더 잘 설명하기 위한 것으로 의도되고 달리 청구된 본 발명의 범위를 제한하지 않는다. 명세서의 어떤 언어도 본 발명의 실시에 필수적인 임의의 청구되지 않은 요소를 나타내는 것으로 해석되어서는 안 된다.

[1254] 본원에 개시된 본 발명의 대안적인 요소 또는 구현예의 그룹화는 제한으로 해석되어서는 안 된다. 각 그룹 구성원은 개별적으로 또는 본원에서 발견되는 그룹의 다른 구성원 또는 다른 요소와 조합하여 언급되고 청구될 수 있다. 그룹의 하나 이상의 구성원은 편의성 및/또는 특허성의 이유로 그룹에 포함되거나, 또는 그룹으로부터 삭제될 수 있는 것으로 예상된다. 임의의 이러한 포함 또는 결실이 발생할 때, 명세서는 변형된 그룹을 함유하며 따라서 첨부된 청구범위에 사용되는 모든 Markush 그룹의 서면 설명을 충족하는 것으로 간주된다.

[1255] 본 발명의 특정 구현예는 본 발명을 수행하기 위해 발명자들에게 알려진 최고의 모드를 포함하여 본원에 기재된다. 물론, 이러한 기재된 구현예에 대한 변이는 전술된 설명을 읽을 때 당업자에게 명백하게 될 것이다. 본 발명자들은 당업자가 이러한 변이를 적절하게 이용하기를 기대하고, 본 발명자들은 본 발명이 본원에 구체적으로 기재된 것보다 달리 실시되기를 의도한다. 따라서, 본 발명은 적용가능한 법에 의해 허용되는 바와 같이 여기에 첨부된 청구범위에서 인용된 주제의 모든 변형 및 등가물을 포함한다. 더욱이, 본원에 달리 표시되지 않거나 또는 문맥상 달리 명백하게 모순되지 않는 한, 모든 가능한 변이에서 상기 기재된 요소의 임의의 조합은 본 발명에 포함된다.

[1256] 또한, 본 명세서에 걸쳐 특허, 인쇄된 간행물, 저널 기사 및 다른 서면 텍스트에 대한 수많은 참조가 이루어졌다(본원에서 참조된 자료). 참조된 자료 각각은 개별적으로 참조된 교시를 위해 그 전체가 본원에 참조로 포함된다. 참조된 자료가 시간 경과에 따라 개정되는 경우(예를 들어, 서열 데이터베이스 항목 등), 해당 참조의 내용은 참조가 본 출원에 대한 우선권 주장에서 출원 시 포함되는 날짜를 기준으로 포함된다.

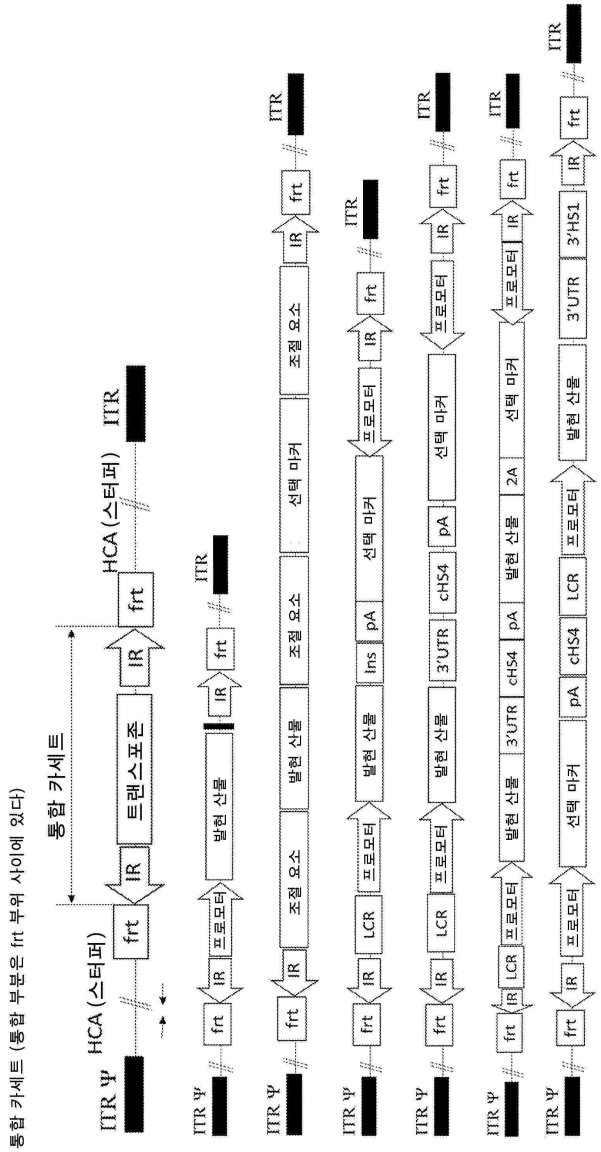
[1257] 마지막으로, 본원에 개시된 본 발명의 구현예는 본 발명의 원리를 예시하는 것으로 이해되어야 한다. 이용될 수 있는 다른 변형은 본 발명의 범위 내에 있다. 따라서, 제한되지 않는 예로서, 본 발명의 대안적인 구성이 본원의 교시에 따라 활용될 수 있다. 따라서, 본 발명은 제시되고 기재된 것으로 정확하게 제한되지 않는다.

[1258] 본원에 제시되는 세부사항은 예로서 그리고 단지 본 발명의 바람직한 구현예의 예시적인 논의를 위한 목적이며 본 발명의 다양한 구현예의 원리 및 개념적 측면의 가장 유용하고 용이하게 이해되는 설명인 것으로 여겨지는 것을 제공하기 위해 제시된다. 이와 관련하여, 본 발명의 기본적인 이해를 위해 필요한 것보다 상세하게 본 발명의 구조적 세부사항을 나타내려는 시도 없이, 설명은 본 발명의 여러 형태가 어떻게 실제로 구현될 수 있는지를 당업자에게 명백하게 하는 도면 및/또는 실시예와 함께 기재되었다.

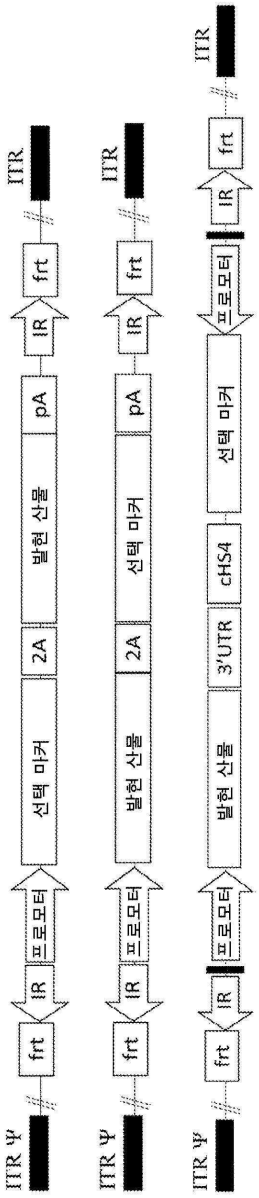
[1259] 본 개시내용에 사용되는 정의 및 설명은 실시예에서 달리 명확하고 모호하지 않게 변형되거나 또는 의미의 적용이 임의의 구축을 무의미하거나 또는 본질적으로 무의미하게 만드는 경우 외에는 임의의 미래 구축에서 제어되는 것을 의미하고 의도된다. 용어의 구축이 무의미하거나 또는 본질적으로 무의미하게 만드는 경우에, 정의는 Webster's Dictionary, 3rd Edition 또는 Oxford Dictionary of Biochemistry and Molecular Biology (Eds. Attwood T 등, Oxford University Press, Oxford, 2006)와 같이 당업자에게 알려진 사전에서 취해야 한다.

## 도면

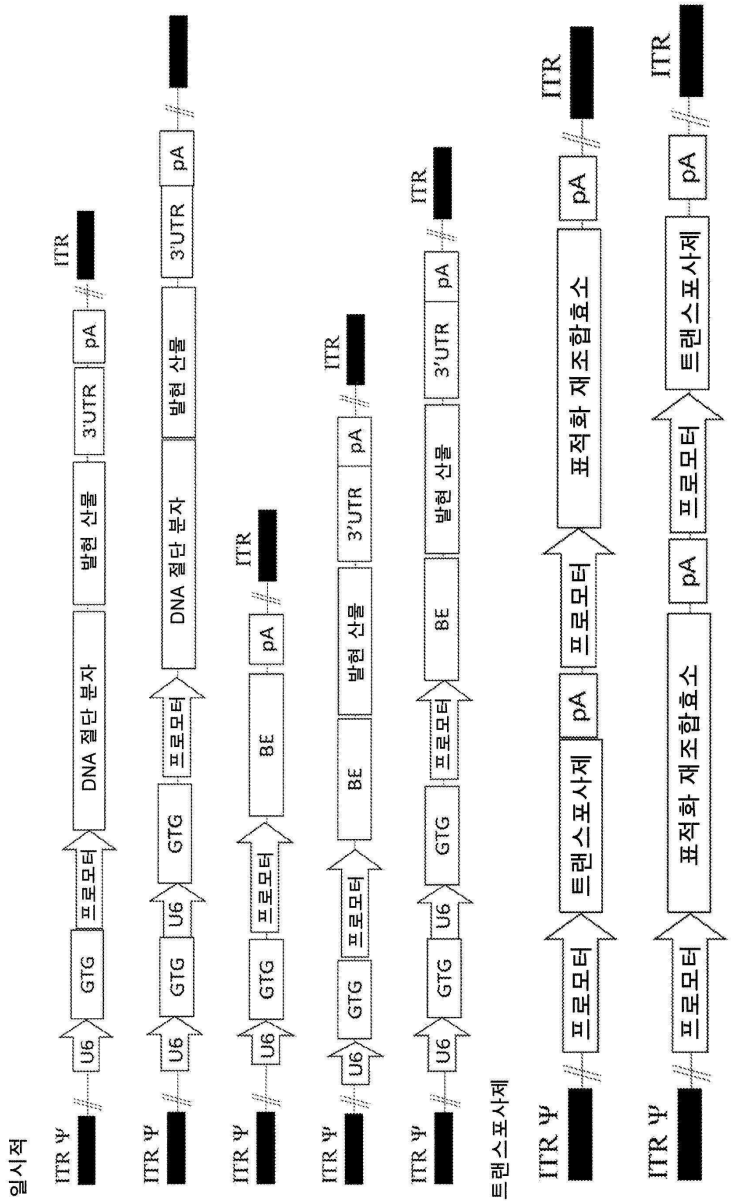
### 도면1a


통합 카세트 (통합 부분은 frt 부위 사이에 있다)
통합 카세트
ITR Ψ
HCA (스터퍼)
frt
IR
트랜스포존
IR
frt
HCA (스터퍼)
ITR
IR
프로모터
발현 산물
IR
frt
조절 요소
선택 마커
LCR
Ins
pA
3'UTR
cHS4
2A
3'HS1

도면1b

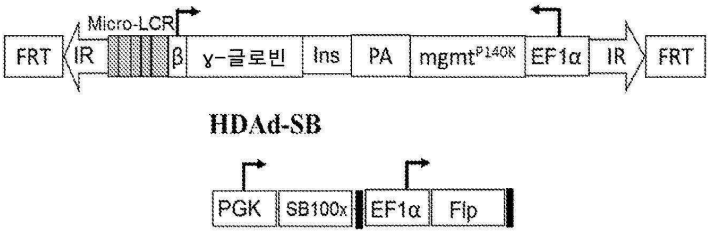

ITR
Ψ
frt
IR
프로모터
선택 마커
2A
발현 산물
pA
IR
frt
ITR
ITR
Ψ
frt
IR
프로모터
발현 산물
2A
선택 마커
pA
IR
frt
ITR
ITR
Ψ
frt
IR
프로모터
발현 산물
3'UTR
cHS4
선택 마커
프로모터
IR
frt
ITR

도면1c

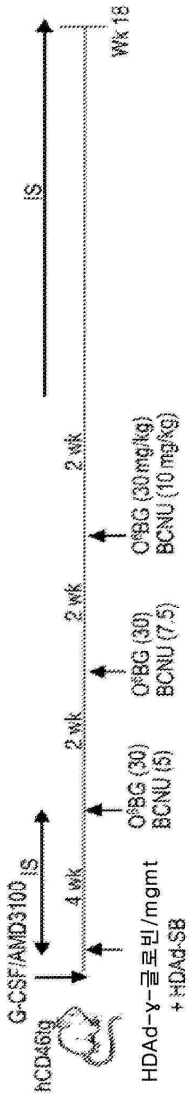



도면2a

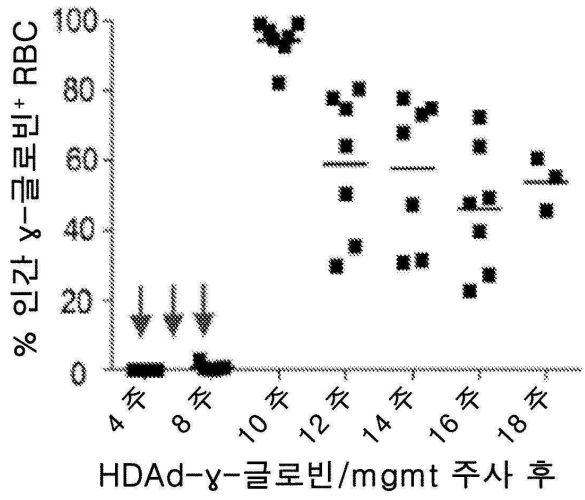
HDAd-γ-글로빈/mgmt



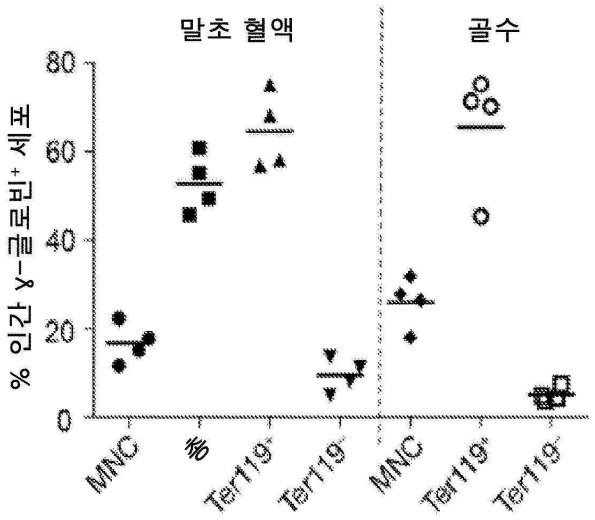
도면2b


Wk 18
IS
2 wk
O⁶BG (30 mg/kg)
BCNU (10 mg/kg)
2 wk
O⁶BG (30)
BCNU (7.5)
2 wk
O⁶BG (30)
BCNU (5)
4 wk
HDAd-γ-글로빈/mgmt
+ HDAd-SB
IS
G-CSF/AMD3100
RCD46tg

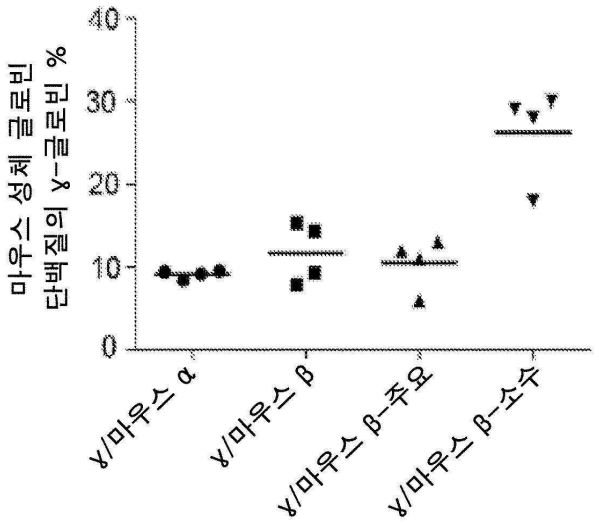
도면2c


% 인간 γ-글로빈⁺ RBC
100
80
60
40
20
0
4 주
8 주
10 주
12 주
14 주
16 주
18 주
HDAd-γ-글로빈/mgmt 주사 후

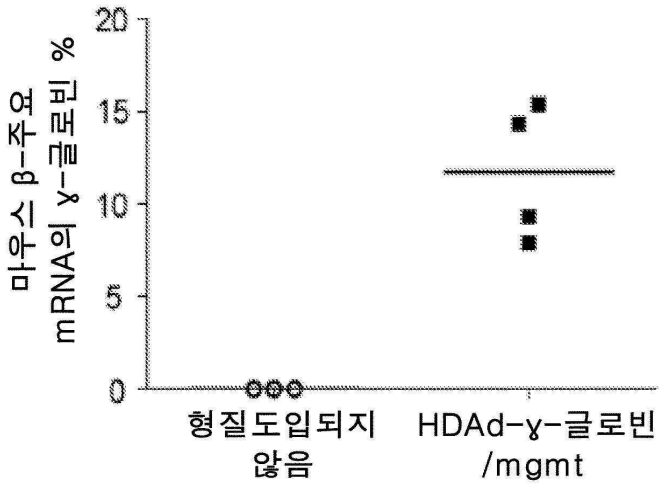
도면2d



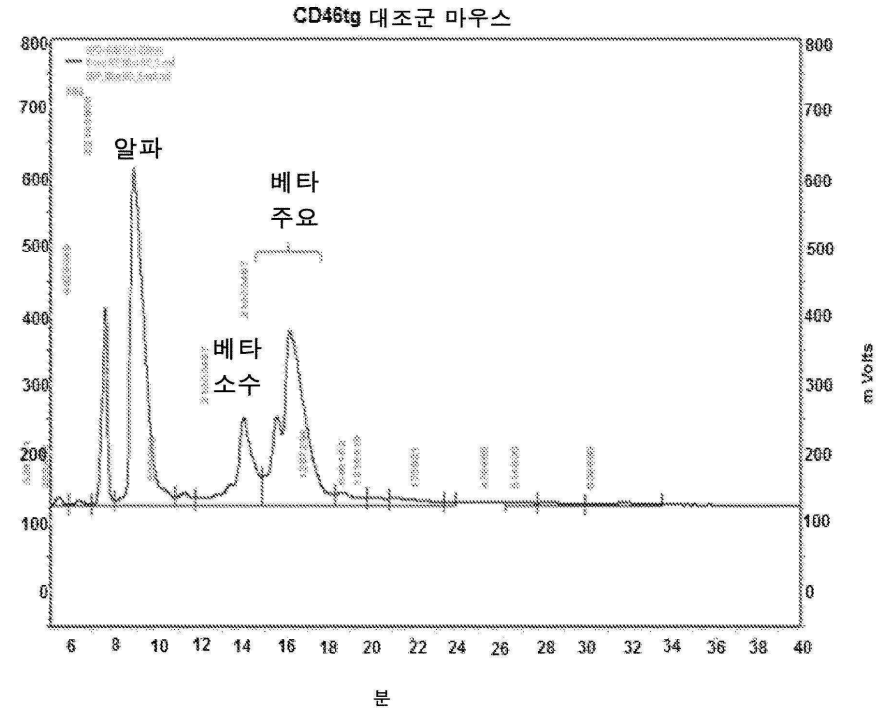

도면2e



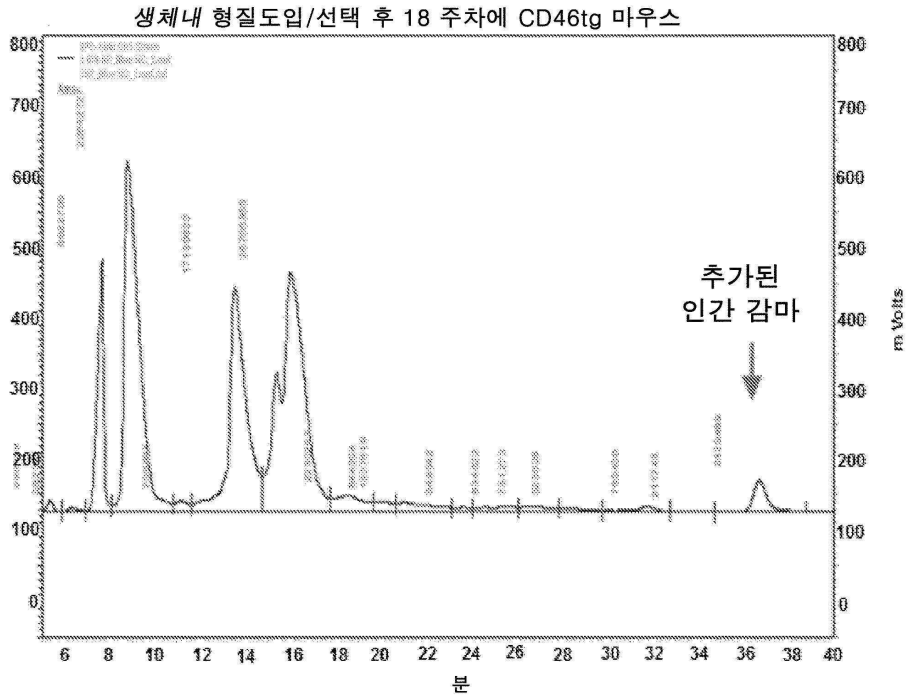

도면2f



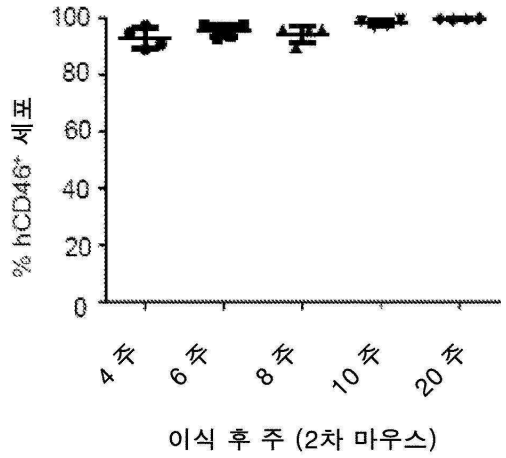
도면3a


CD46tg 대조군 마우스
알파
베타
주요
베타
소수
m Volts
분

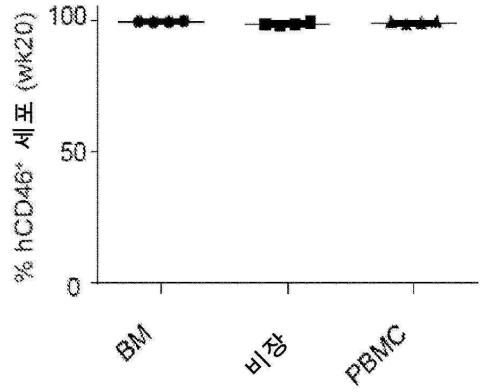
도면3b


생체내 형질도입/선택 후 18 주차에 CD46tg 마우스
추가된
인간 감마
m Volts
분

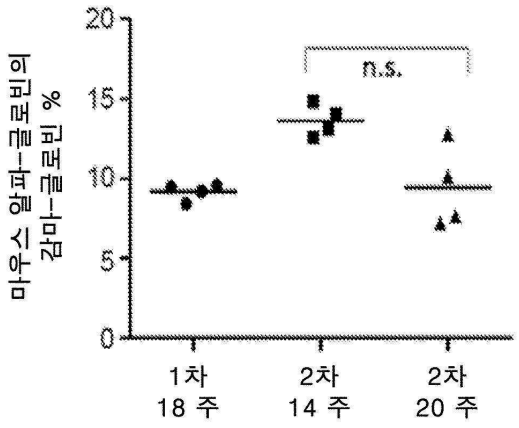
도면4a




도면4b

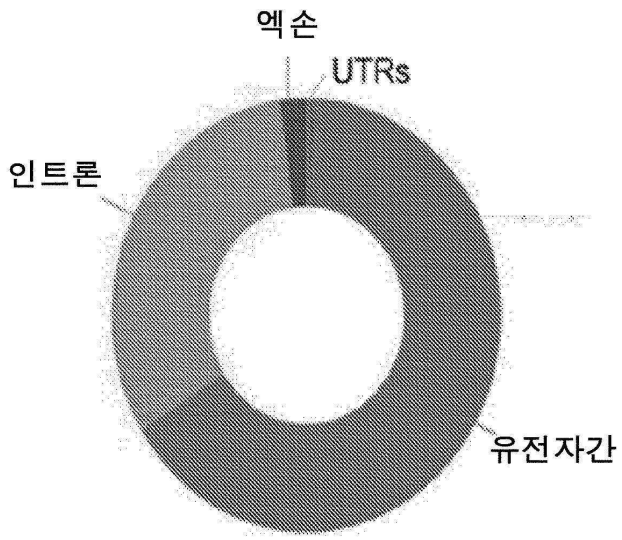



도면4c

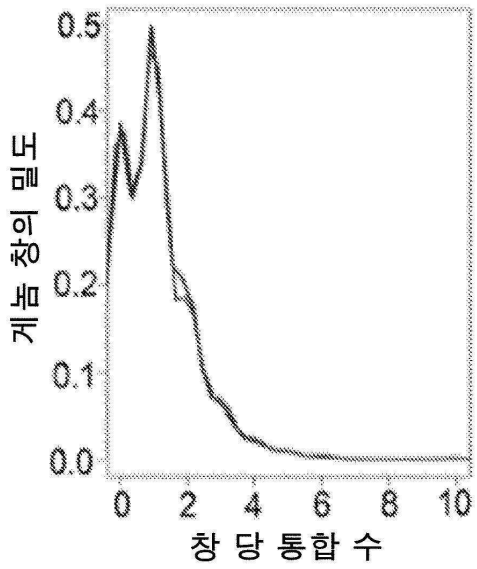


도면5a

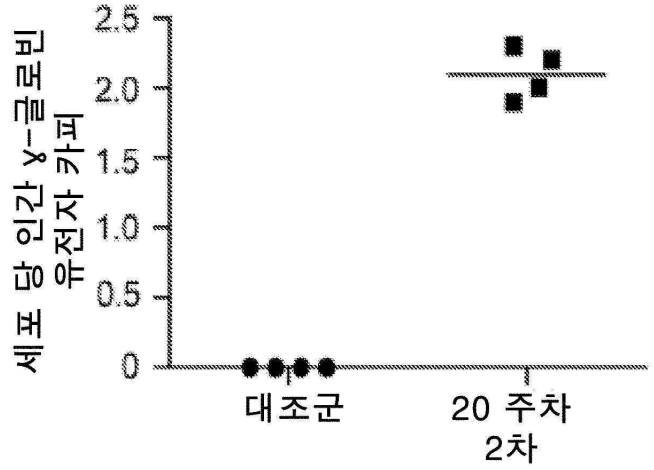
도면5b



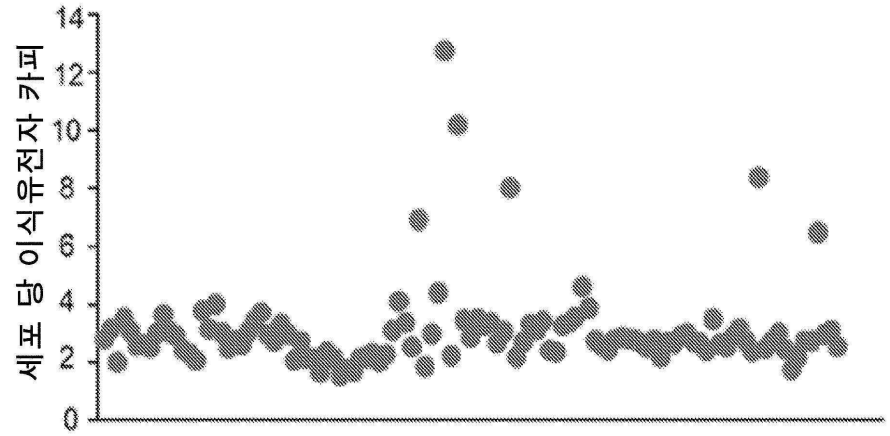
도면5c



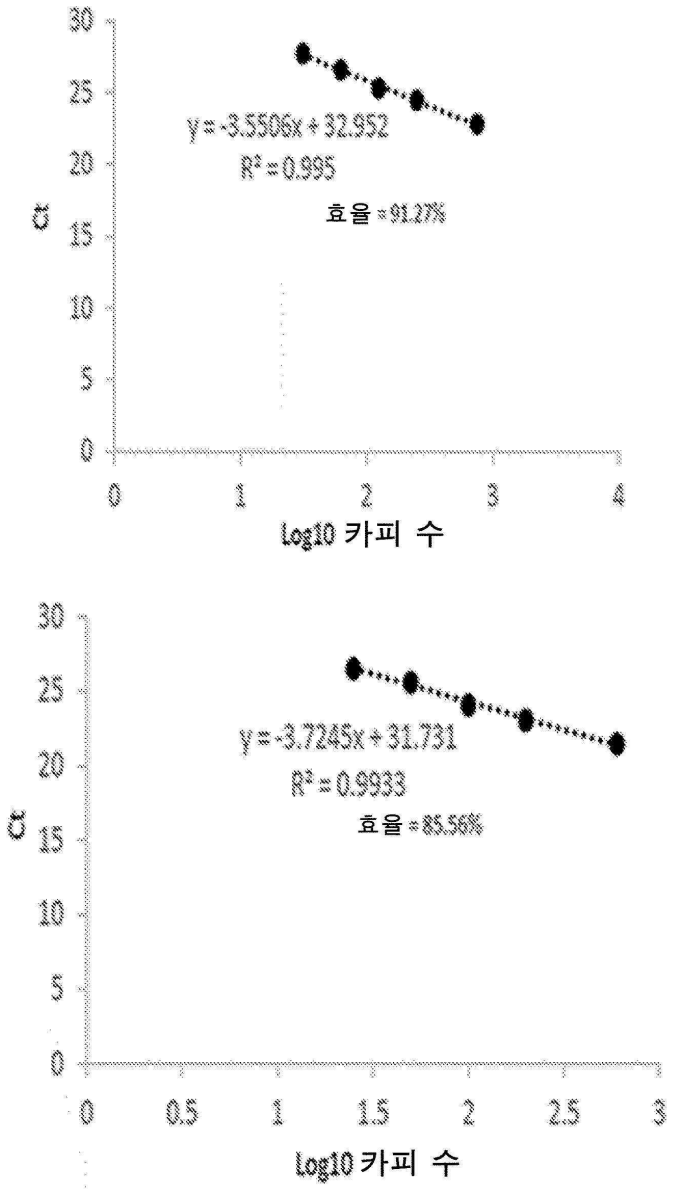

도면5d



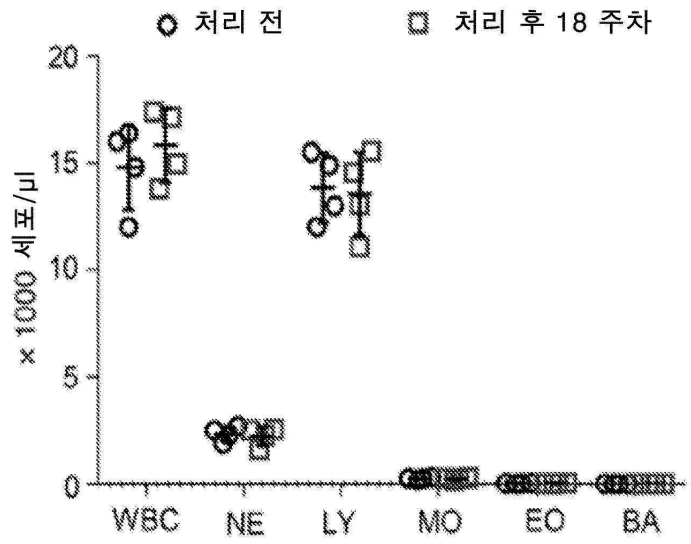

도면5e



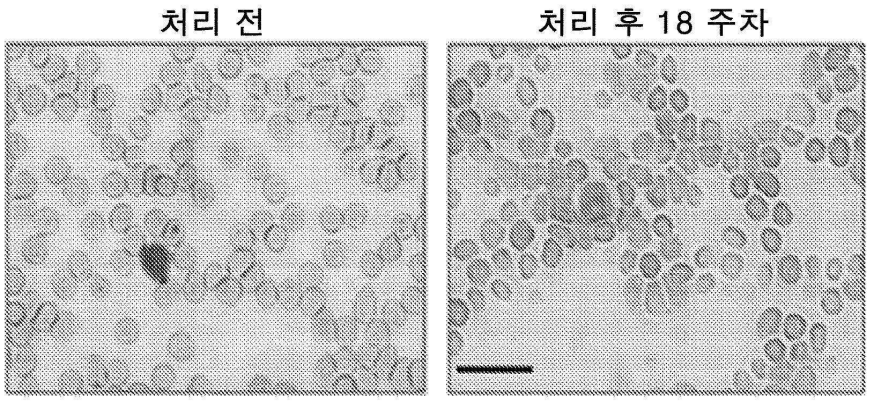
도면6


30
25
20
15
10
5
0
Ct
y = -3.5506x + 32.952
R² = 0.995
효율 = 91.27%
0
1
2
3
4
Log10 카피 수
30
25
20
15
10
5
0
Ct
y = -3.7245x + 31.731
R² = 0.9933
효율 = 85.56%
0
0.5
1
1.5
2
2.5
3
Log10 카피 수

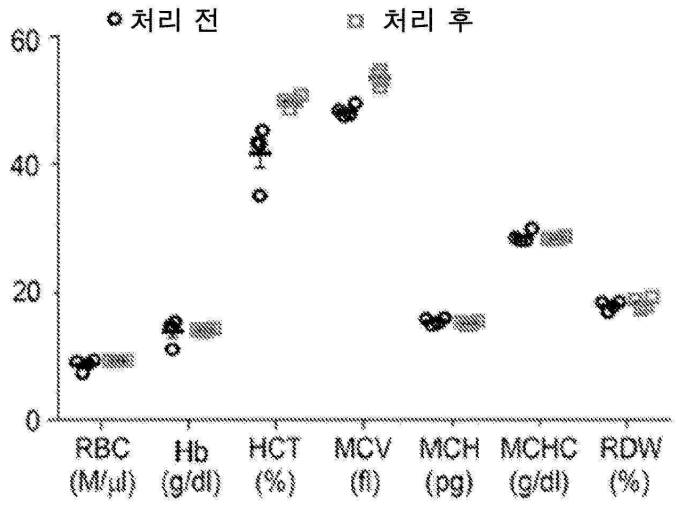
도면7a



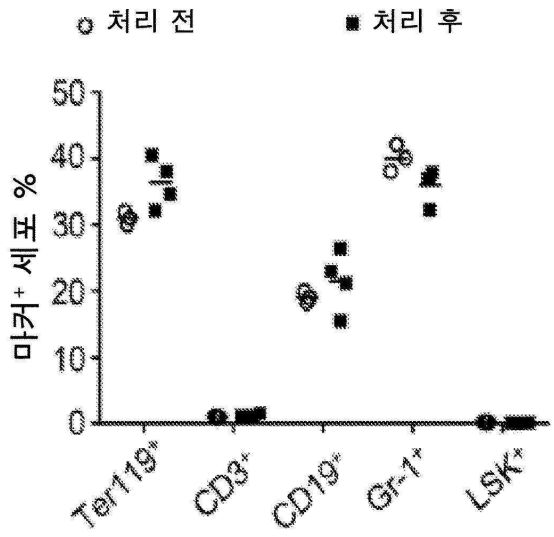

도면7b



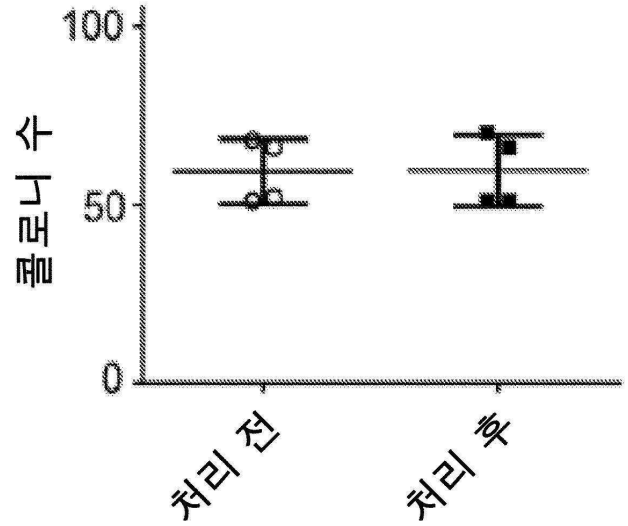

도면7c



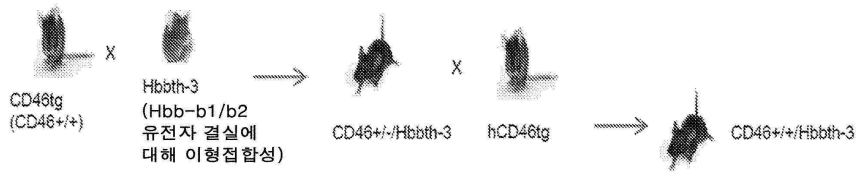
도면7d



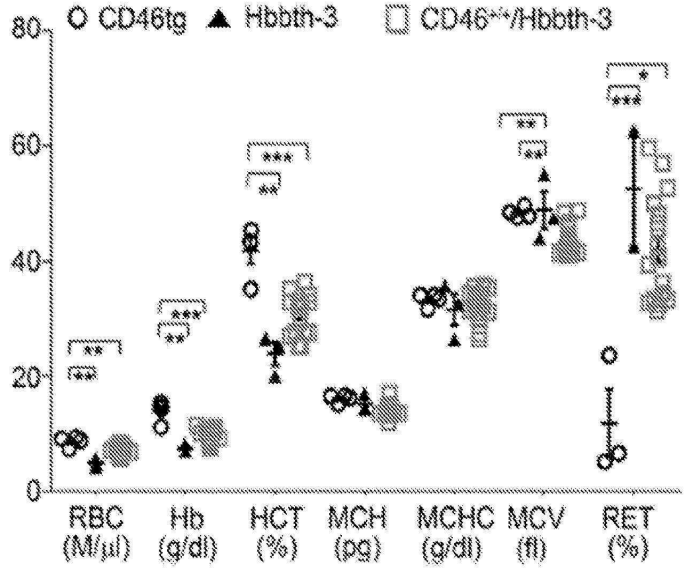

도면7e



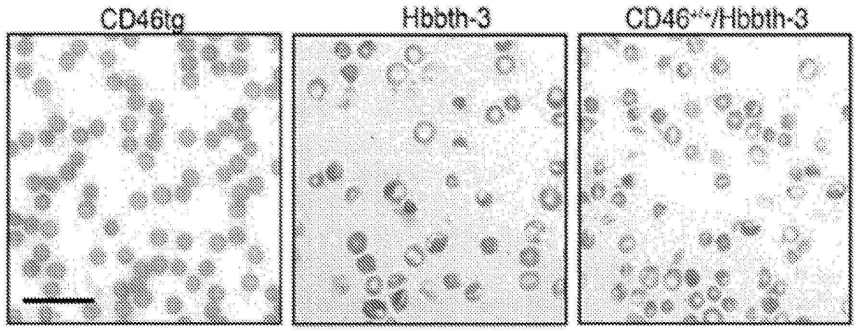

도면8



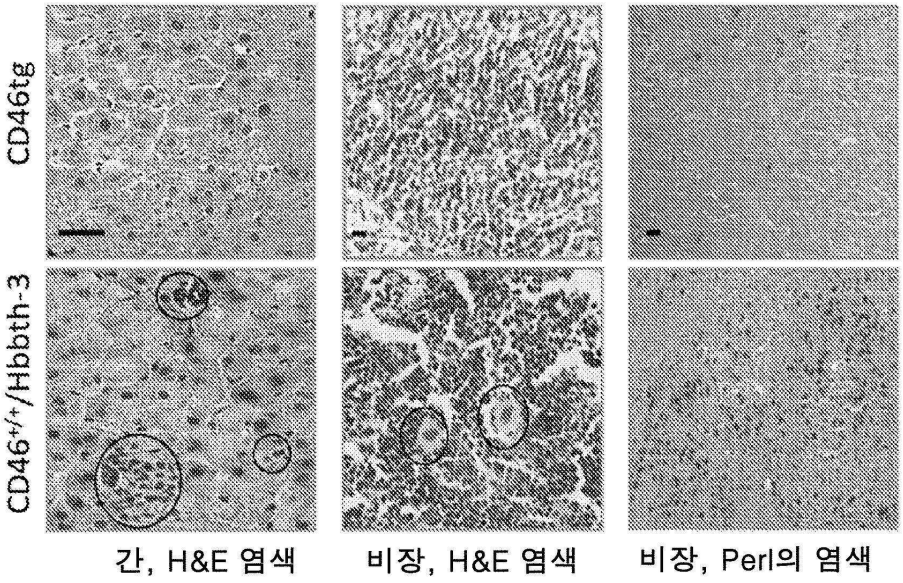
도면9a



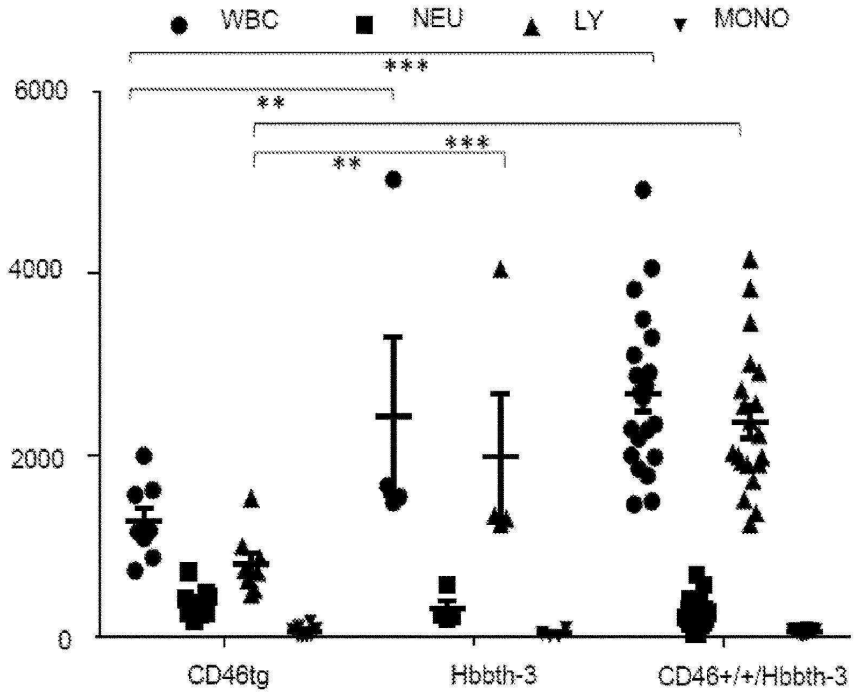

도면9b



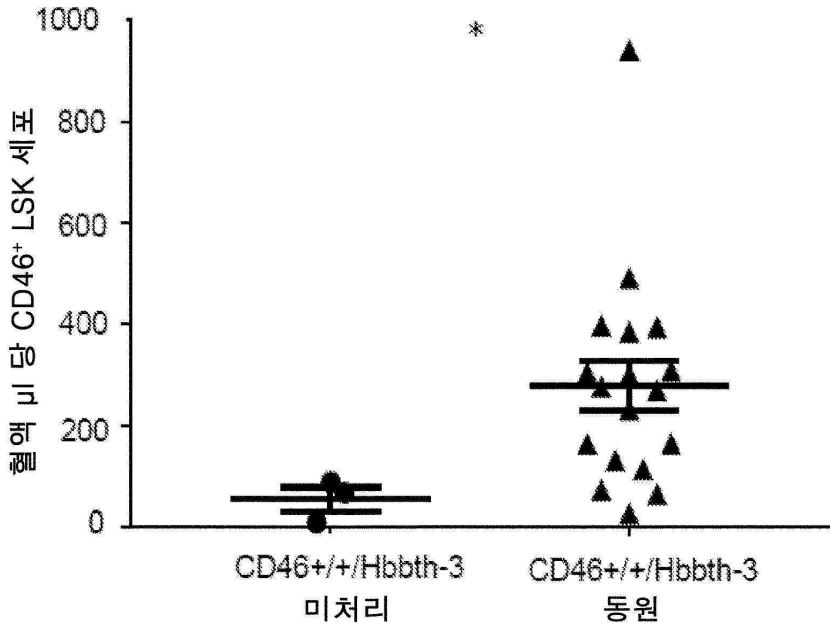

도면9c



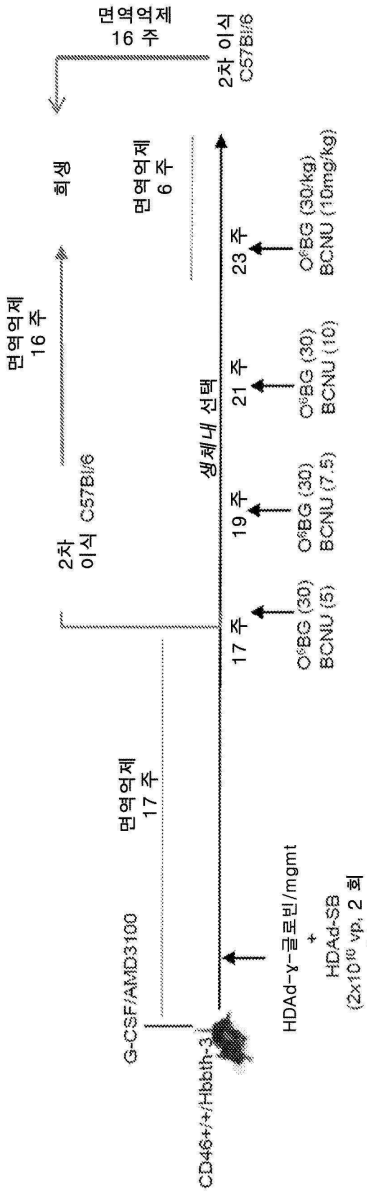
도면10


WBC
NEU
LY
MONO
***
**
***
**
6000
4000
2000
0
CD46tg
Hbbth-3
CD46+/+/Hbbth-3

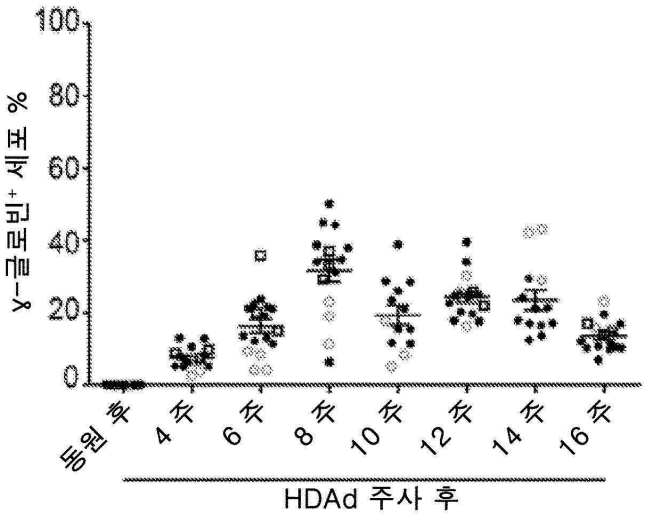
도면11


1000
800
600
400
200
0
혈액 μl 당 CD46+ LSK 세포
*
CD46+/+/Hbbth-3
미처리
CD46+/+/Hbbth-3
동원

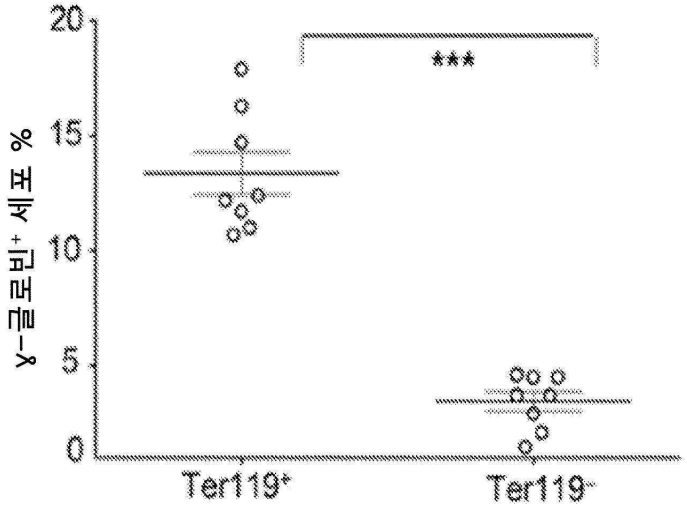
도면12


CD46+/+/Hbb$^{th-3}$
G-CSF/AMD3100
면역억제
17 주
HDAd-γ-글로빈/mgmt
+
HDAd-SB
($2x10^{10}$ vp, 2 회
IV
17 주
O$^6$BG (30)
BCNU (5)
19 주
O$^6$BG (30)
BCNU (7.5)
21 주
O$^6$BG (30)
BCNU (10)
23 주
O$^6$BG (30/kg)
BCNU (10mg/kg)
생체내 선택
면역억제
6 주
희생
2차 이식
C57Bl/6
면역억제
16 주
2차
이식 C57Bl/6
면역억제
16 주

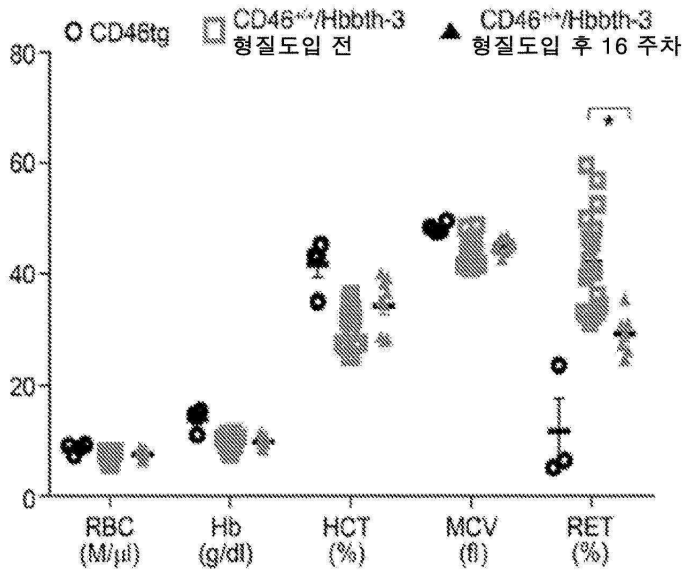
도면13a


γ-글로빈$^+$ 세포 %
100
80
60
40
20
0
동원 후
4 주
6 주
8 주
10 주
12 주
14 주
16 주
HDAd 주사 후

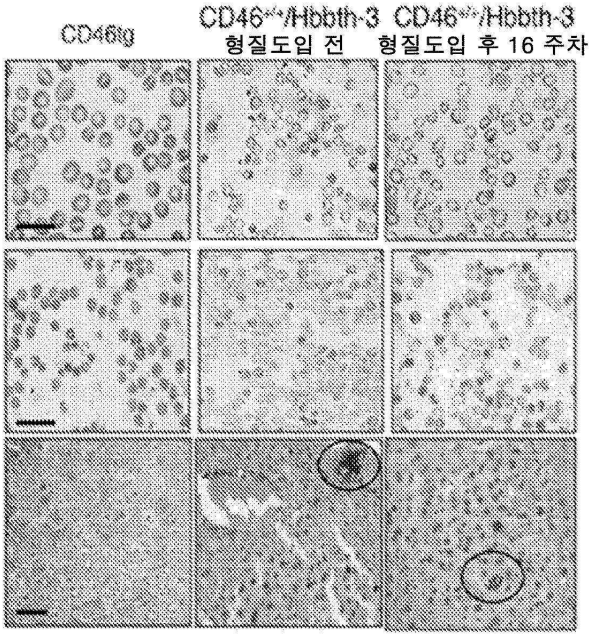
도면13b


20
15
10
5
0
γ-글로빈+ 세포 %
***
Ter119+
Ter119-

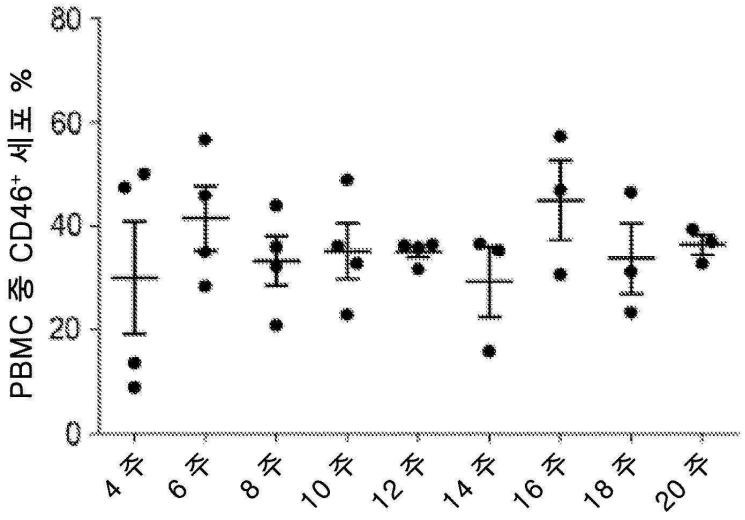
도면13c


CD46tg
CD46-/-/Hbbth-3
형질도입 전
CD46-/-/Hbbth-3
형질도입 후 16 주차
80
60
40
20
0
*
RBC
(M/μl)
Hb
(g/dl)
HCT
(%)
MCV
(fl)
RET
(%)

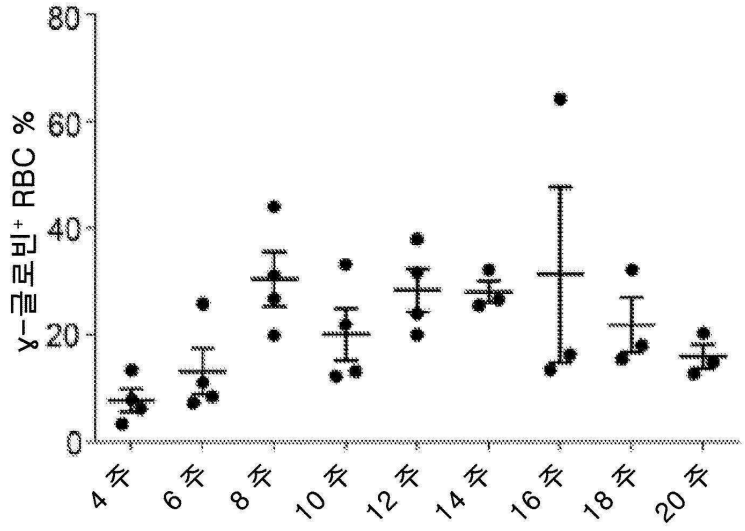
도면13d


CD46tg
CD46-/-/Hbbth-3
형질도입 전
CD46-/-/Hbbth-3
형질도입 후 16 주차

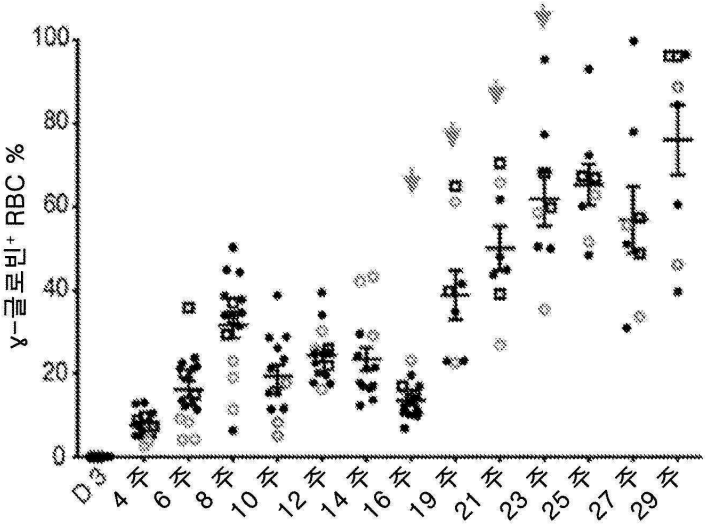
도면13e


PBMC 중 CD46+ 세포 %
80
60
40
20
0
4 주
6 주
8 주
10 주
12 주
14 주
16 주
18 주
20 주

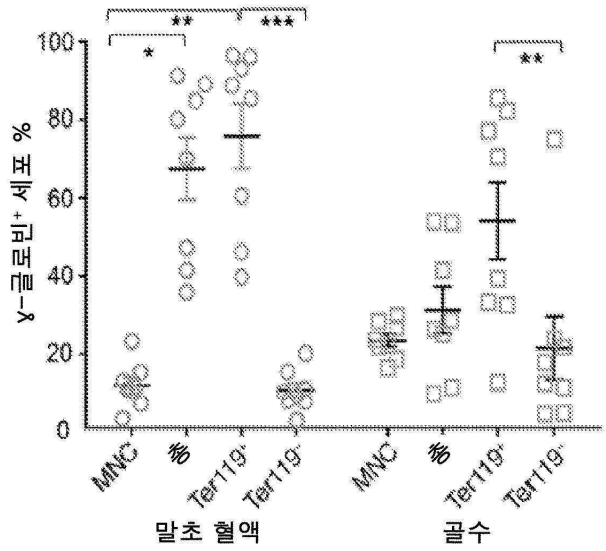
도면13f


80
60
40
20
0
γ-글로빈⁺ RBC %
4 주
6 주
8 주
10 주
12 주
14 주
16 주
18 주
20 주

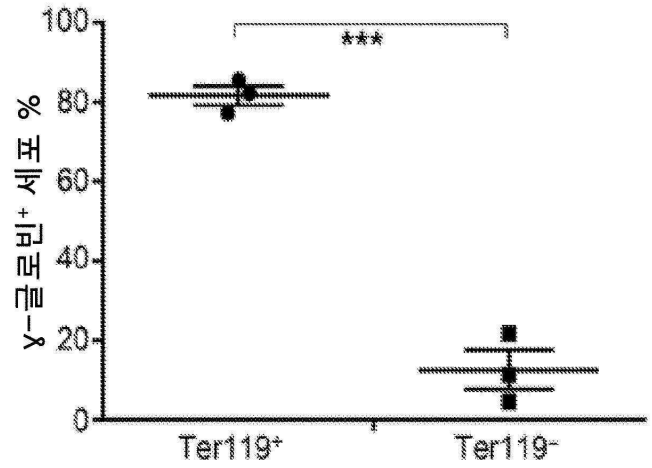
도면14a


100
80
60
40
20
0
γ-글로빈⁺ RBC %
D3
4 주
6 주
8 주
10 주
12 주
14 주
16 주
19 주
21 주
23 주
25 주
27 주
29 주

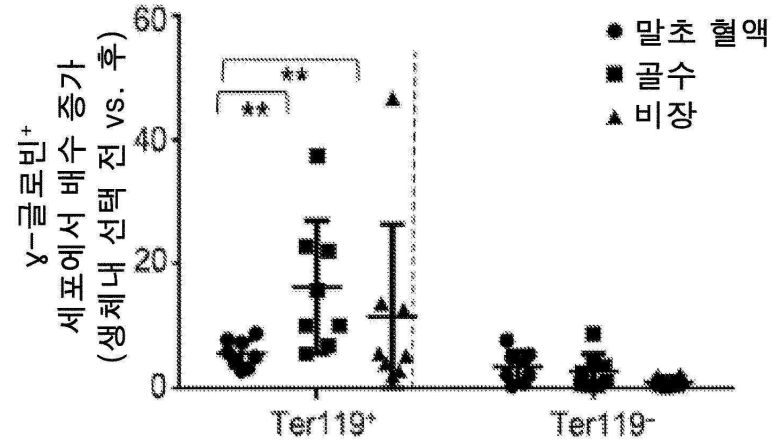
도면14b



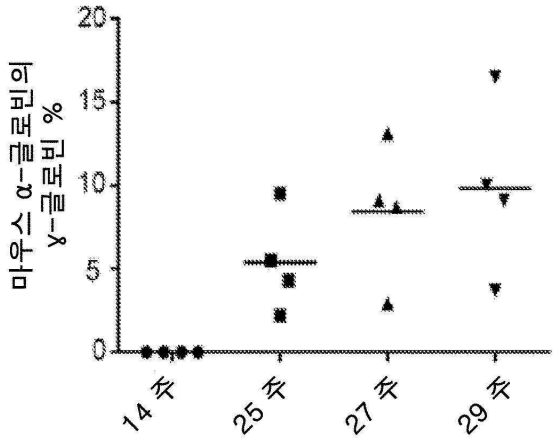

도면14c



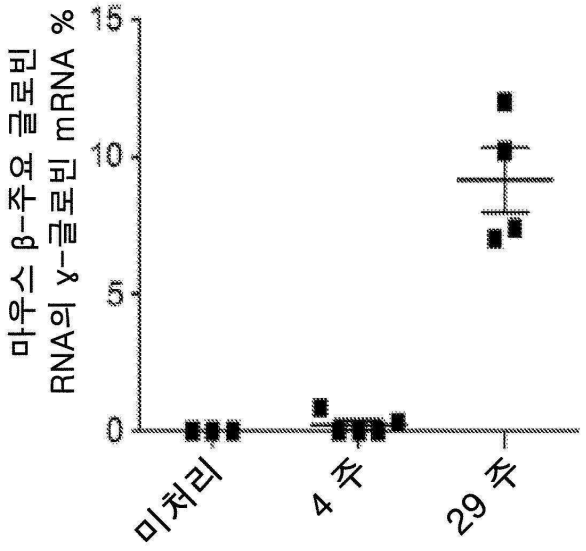

도면14d



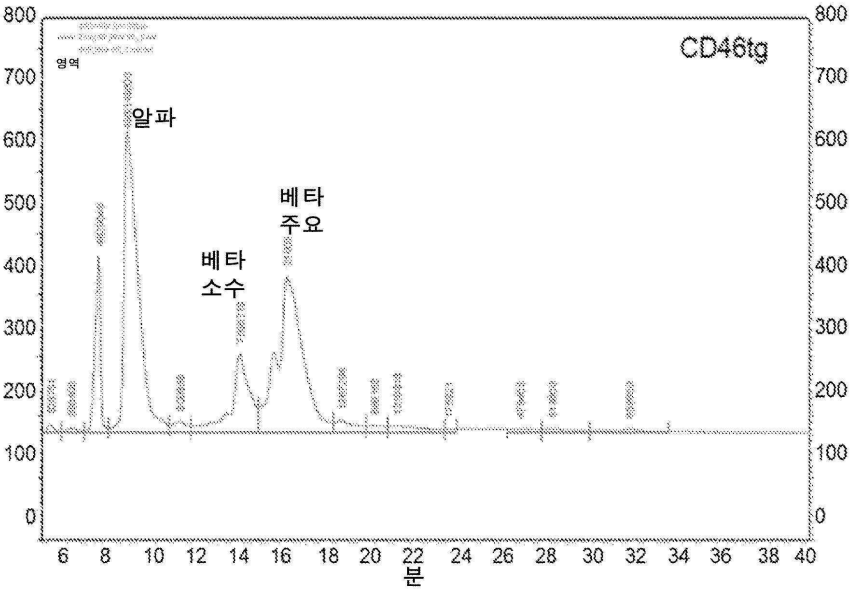
도면14e



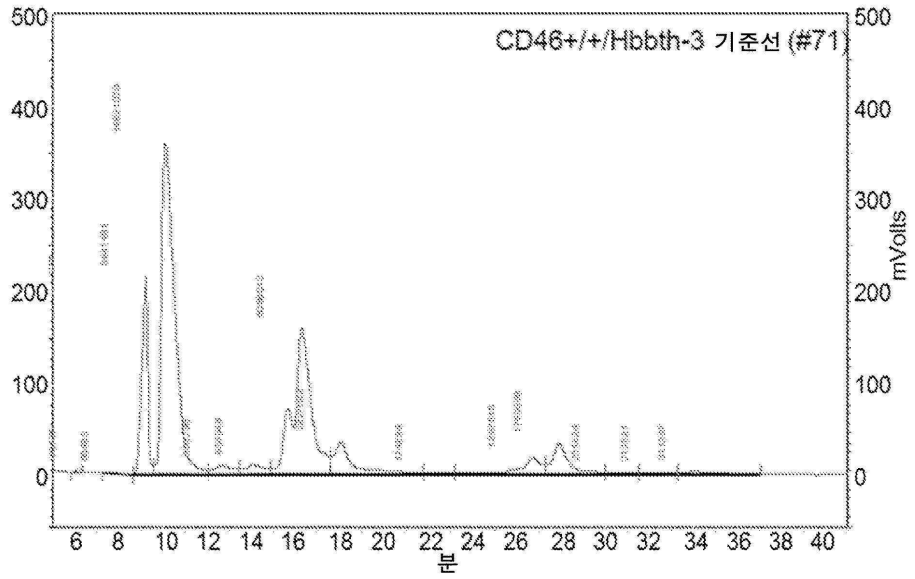

도면14f



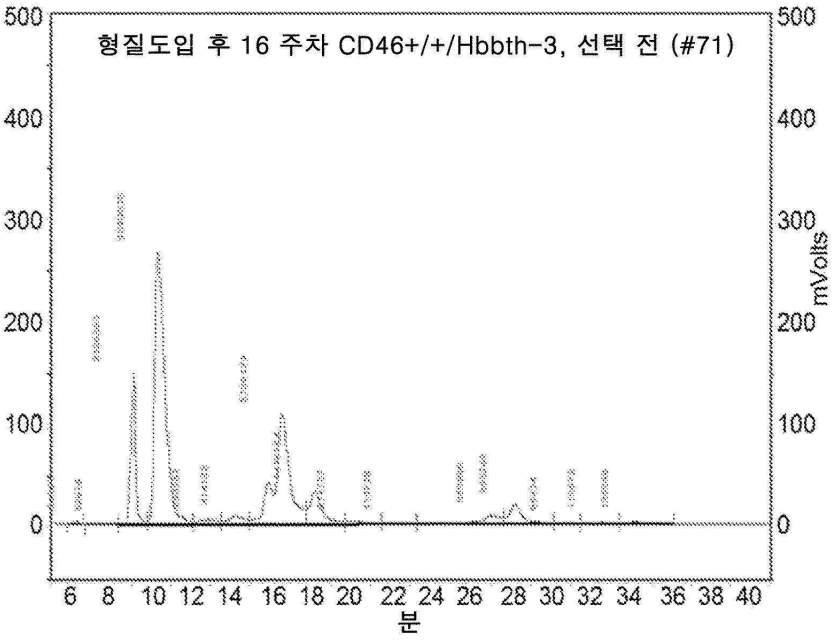

도면15a



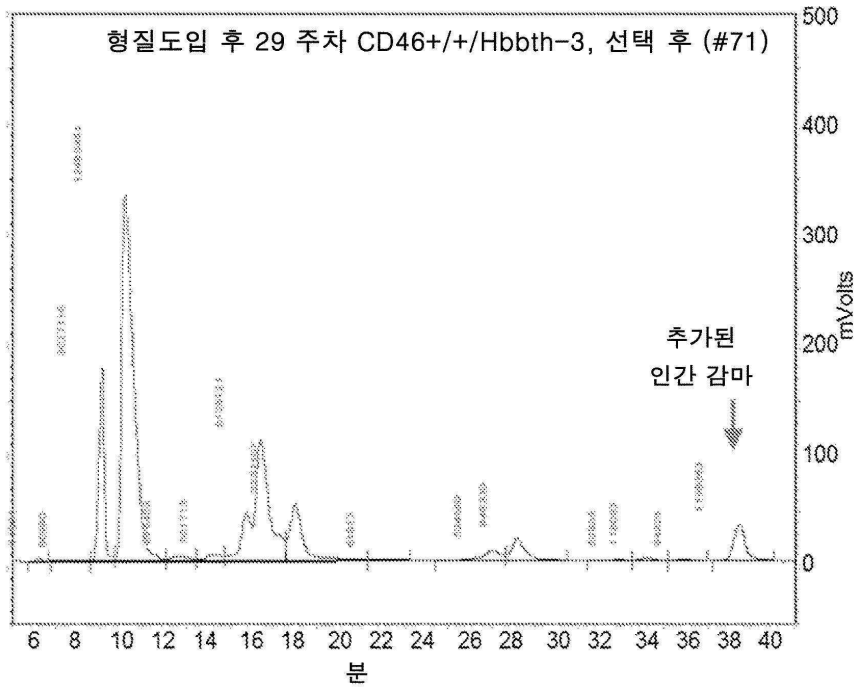
도면15b


CD46+/+/Hbbth-3 기준선 (#71)
500
400
300
200
100
0
mVolts
6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40
분

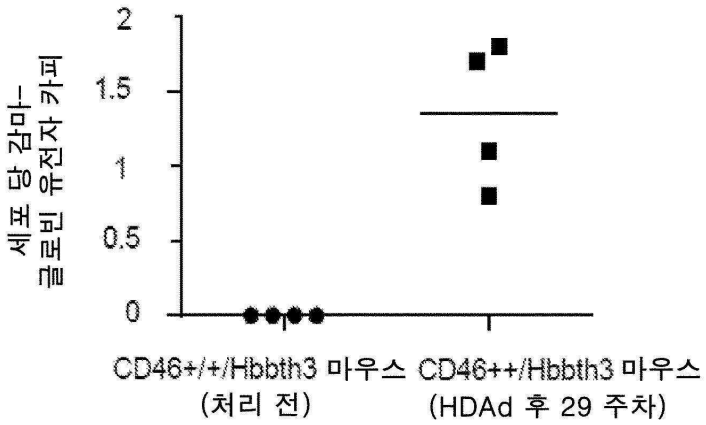
도면15c


형질도입 후 16 주차 CD46+/+/Hbbth-3, 선택 전 (#71)
500
400
300
200
100
0
mVolts
6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40
분

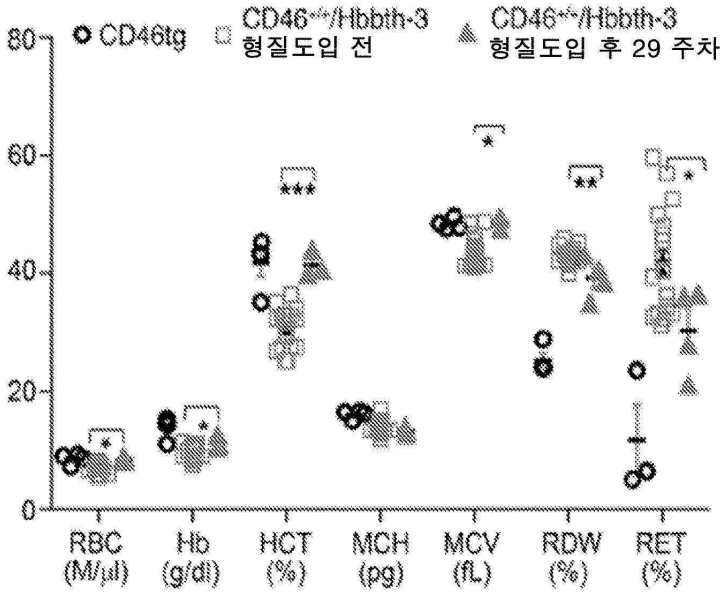
도면15d


형질도입 후 29 주차 CD46+/+/Hbbth-3, 선택 후 (#71)
추가된
인간 감마
500
400
300
200
100
0
mVolts
6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40
분

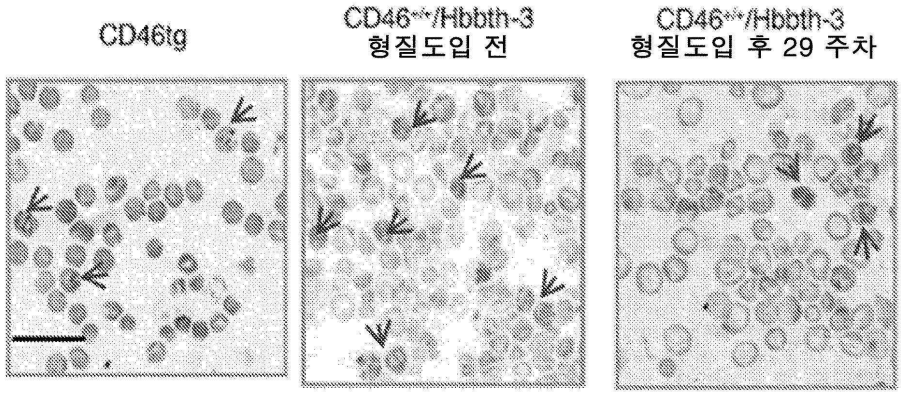
도면16


세포 당 감마-
글로빈 유전자 카피
2
1.5
1
0.5
0
CD46+/+/Hbbth3 마우스
(처리 전)
CD46++/Hbbth3 마우스
(HDAd 후 29 주차)

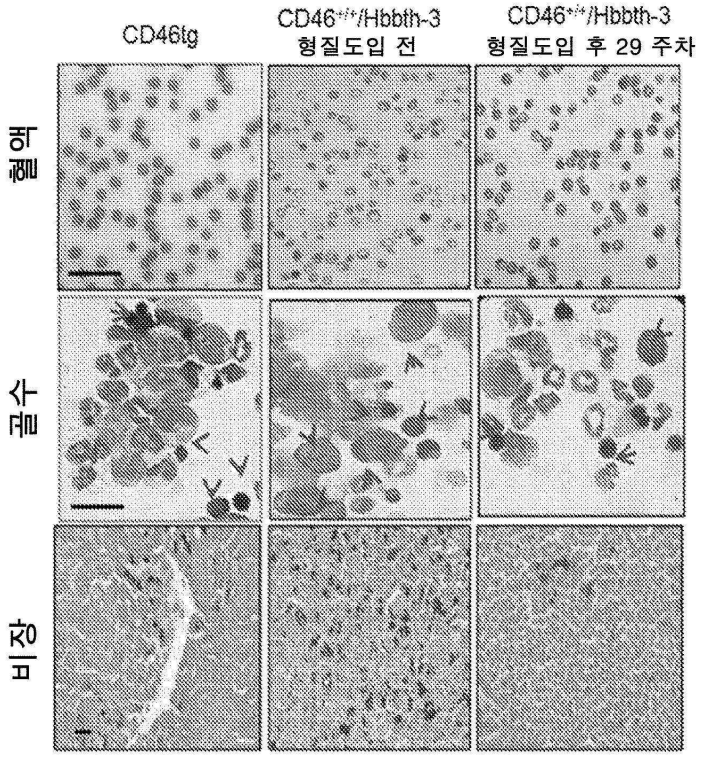
도면17a


CD46tg
CD46-/-/Hbbth-3
형질도입 전
CD46-/-/Hbbth-3
형질도입 후 29 주차
80
60
40
20
0
RBC (M/μl)
Hb (g/dl)
HCT (%)
MCH (pg)
MCV (fL)
RDW (%)
RET (%)

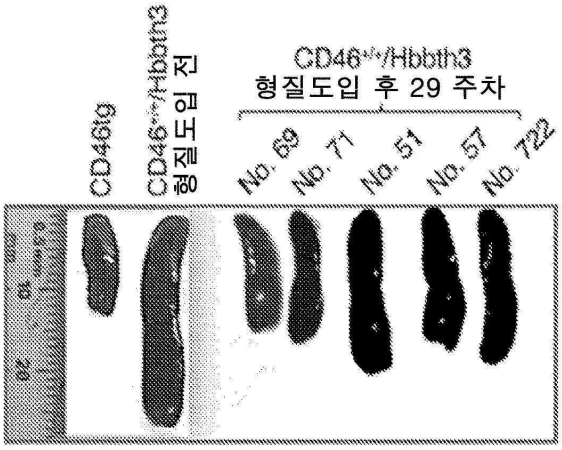
도면17b


CD46tg
CD46-/-/Hbbth-3
형질도입 전
CD46-/-/Hbbth-3
형질도입 후 29 주차

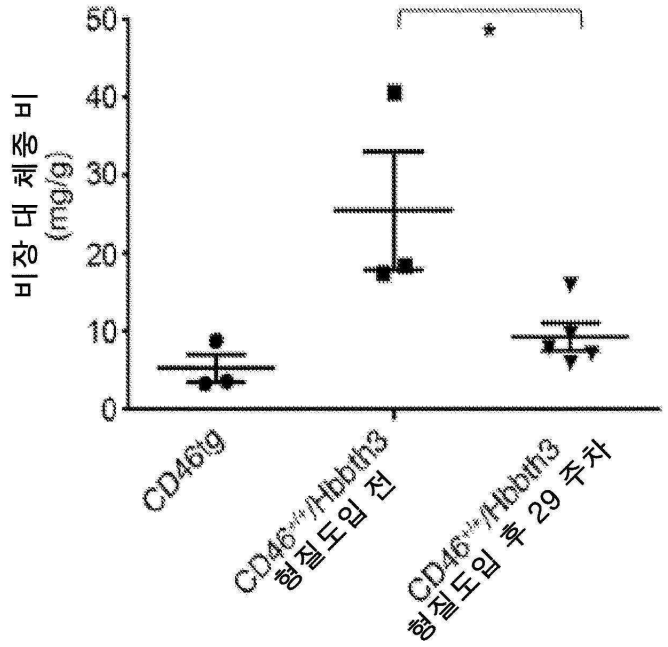
도면17c


CD46tg
CD46+/+/Hbbth-3
형질도입 전
CD46+/+/Hbbth-3
형질도입 후 29 주차
혈액
골수
비장

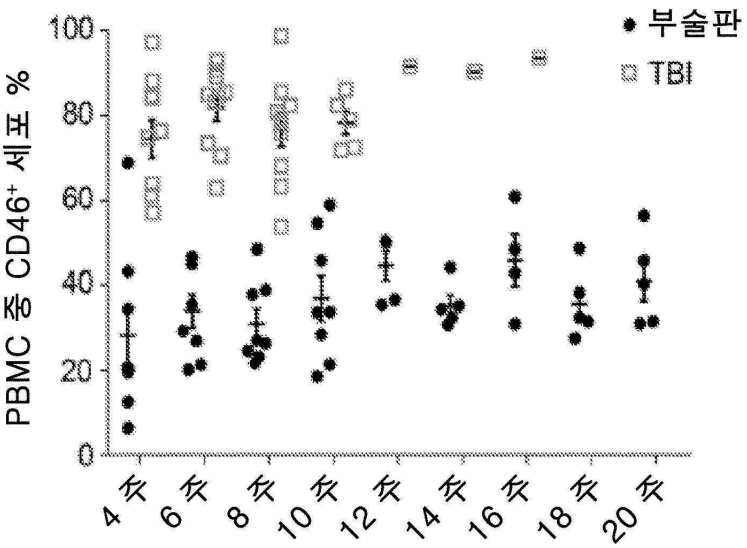
도면17d


CD46tg
CD46+/+/Hbbth3
형질도입 전
CD46+/+/Hbbth3
형질도입 후 29 주차
No. 69
No. 71
No. 51
No. 57
No. 722

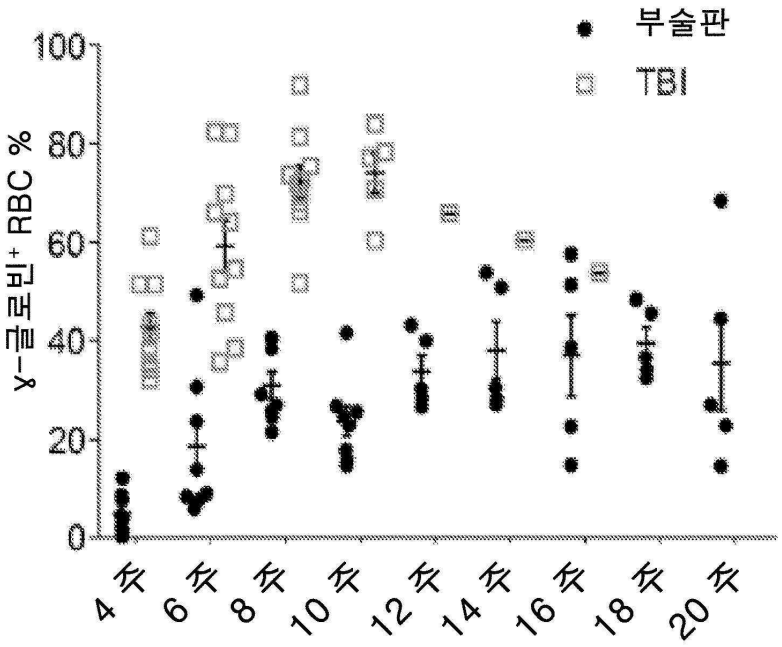
도면17e


비장 대 체중 비 (mg/g)
50
40
30
20
10
0
*
CD46tg
CD46hi/Hbbth3 형질도입 전
CD46hi/Hbbth3 형질도입 후 29 주차

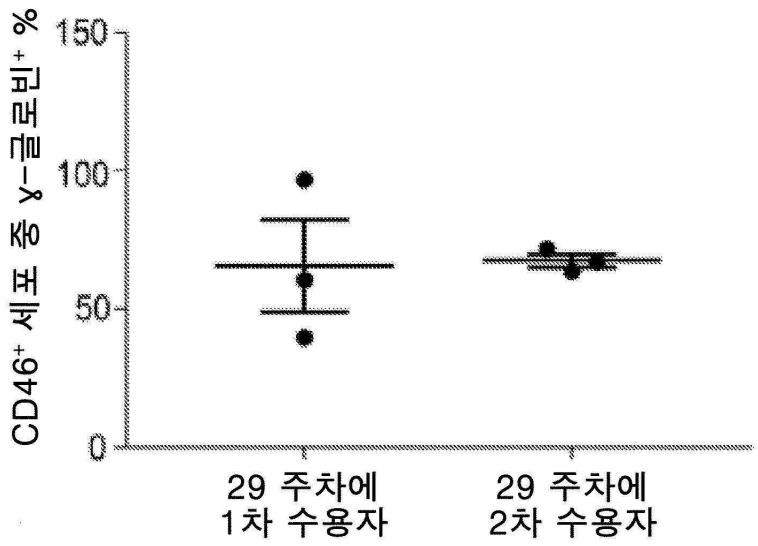
도면18a


PBMC 중 CD46+ 세포 %
100
80
60
40
20
0
부술판
TBI
4 주
6 주
8 주
10 주
12 주
14 주
16 주
18 주
20 주

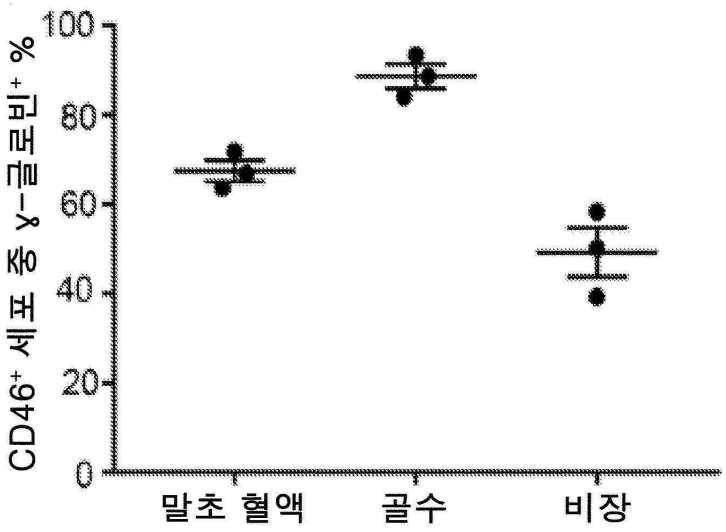
도면18b


부술판
TBI
γ-글로빈+ RBC %
100
80
60
40
20
0
4 주
6 주
8 주
10 주
12 주
14 주
16 주
18 주
20 주

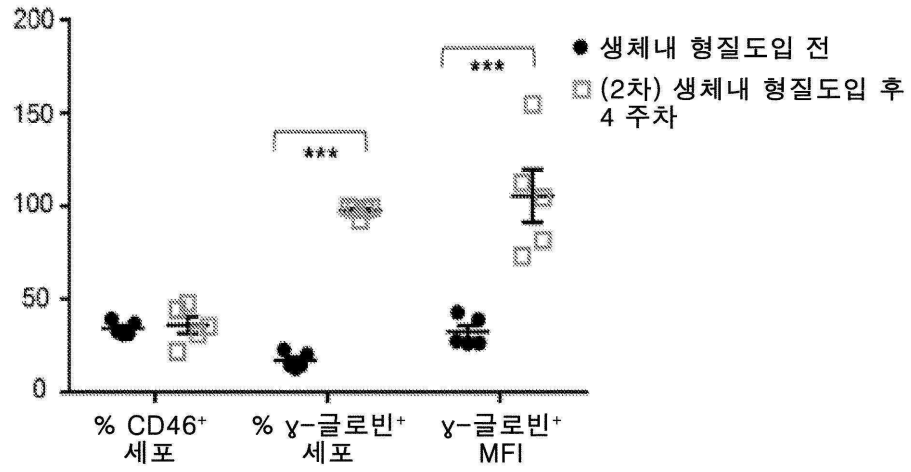
도면18c


CD46+ 세포 중 γ-글로빈+ %
150
100
50
0
29 주차에
1차 수용자
29 주차에
2차 수용자

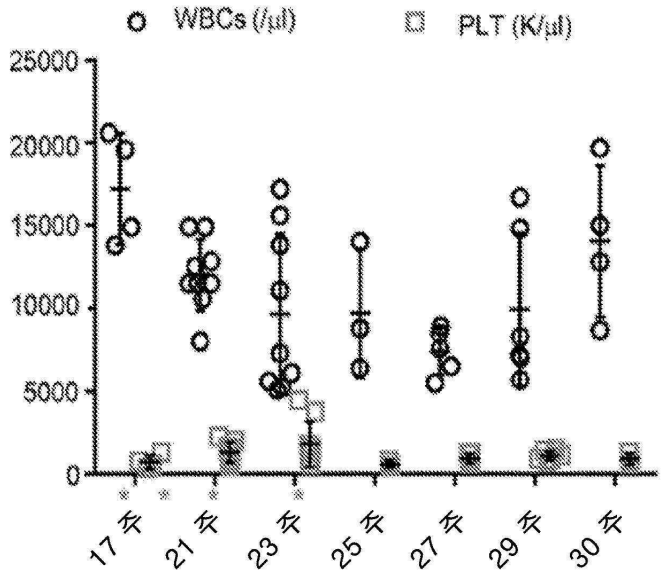
도면18d



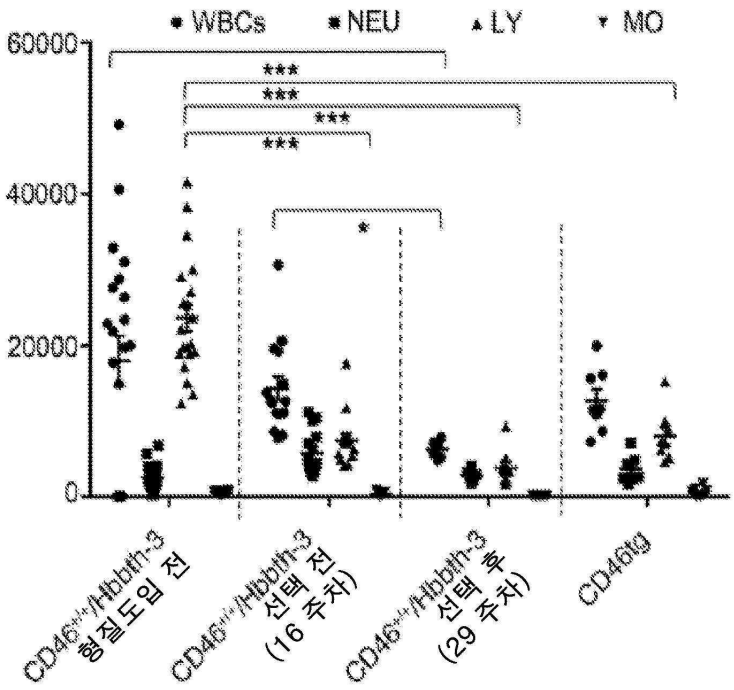

도면18e



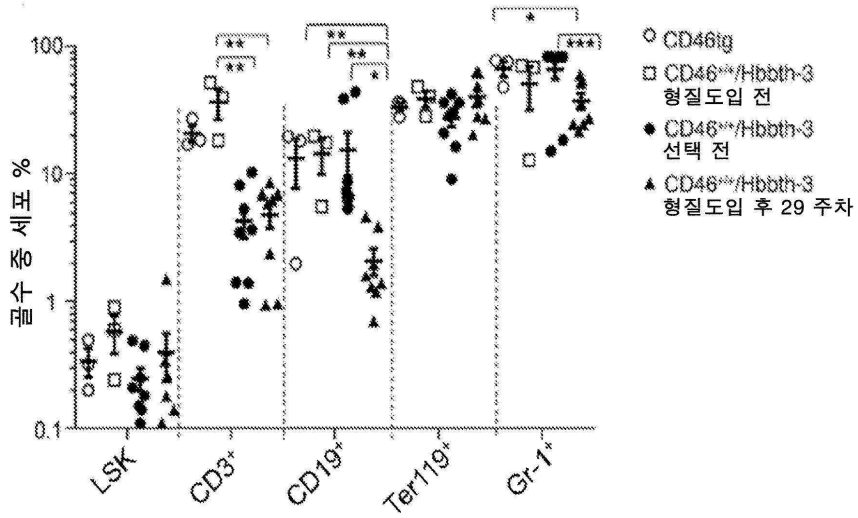

도면19a



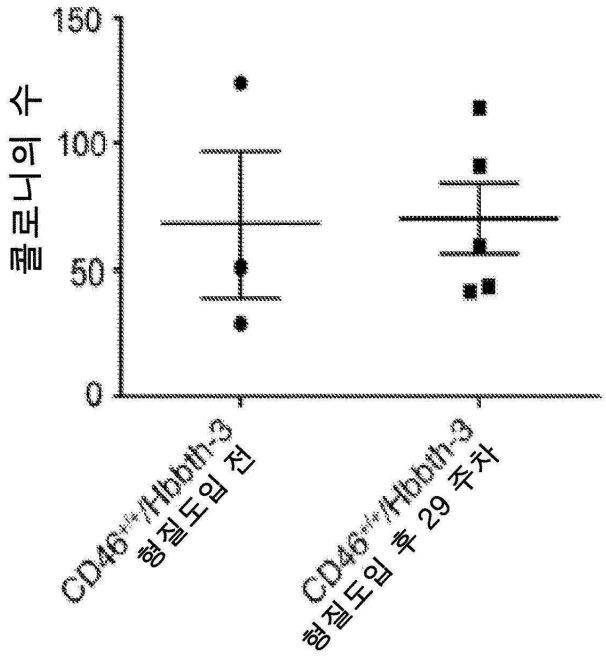
도면19b


WBCs
NEU
LY
MO
60000
40000
20000
0
***
***
***
***
*
CD46tg
CD46tg/Hbbth-3 형질도입 전
CD46tg/Hbbth-3 선택 전 (16 주차)
CD46tg/Hbbth-3 선택 후 (29 주차)
CD46tg

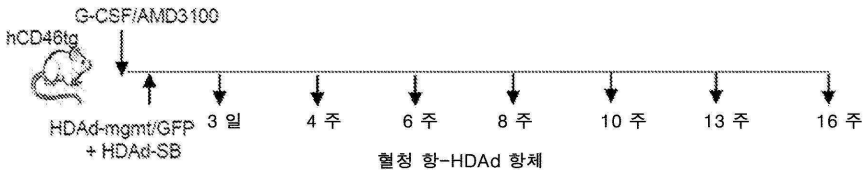
도면19c


골수 중 세포 %
100
10
1
0.1
LSK
CD3+
CD19+
Ter119+
Gr-1+
CD46tg
CD46tg/Hbbth-3 형질도입 전
CD46tg/Hbbth-3 선택 전
CD46tg/Hbbth-3 형질도입 후 29 주차

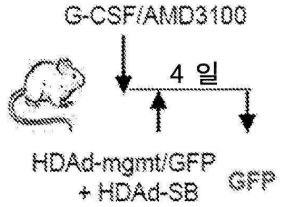
도면19d



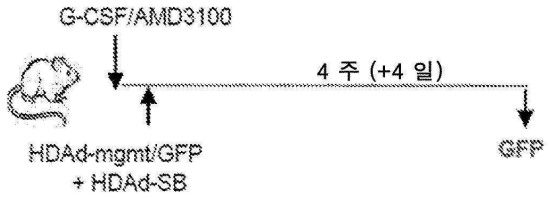

도면20a



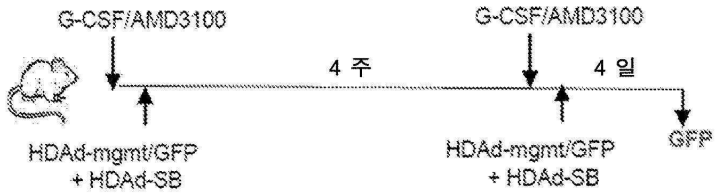

도면20b



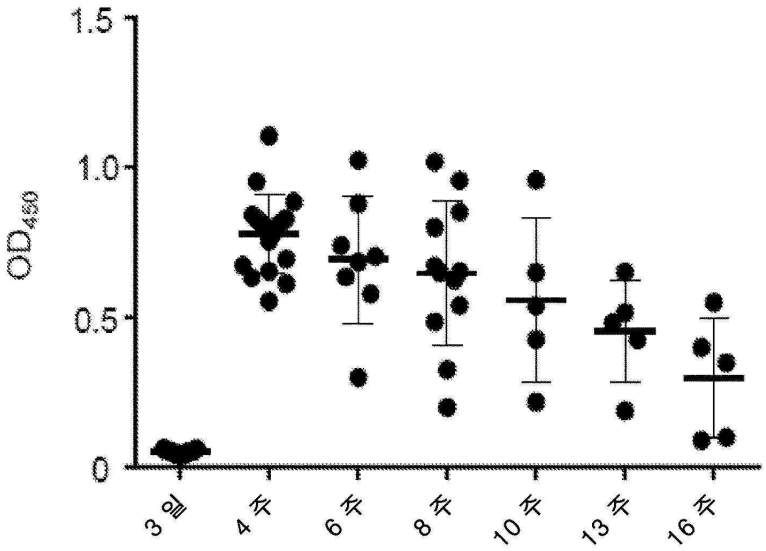

도면20c



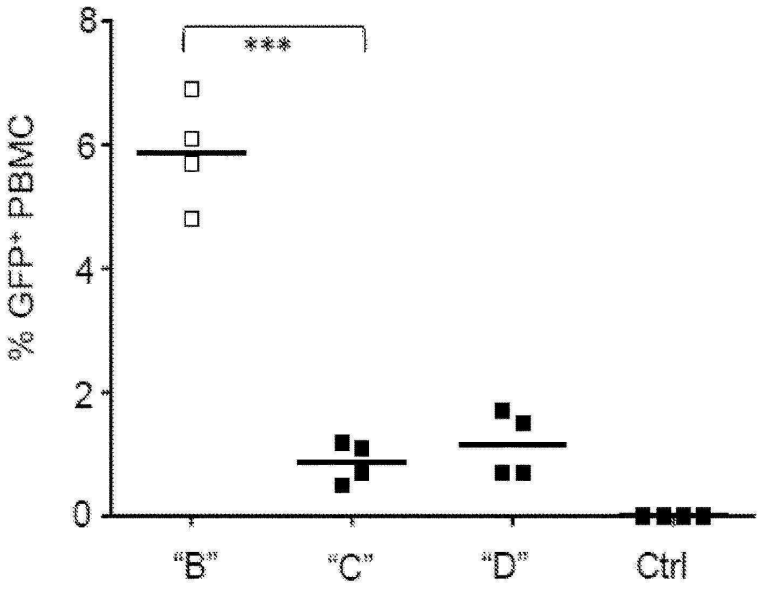
도면20d



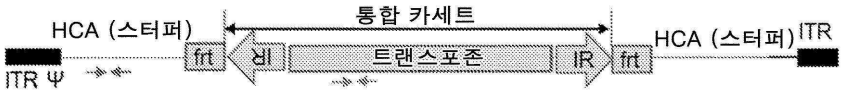

도면20e



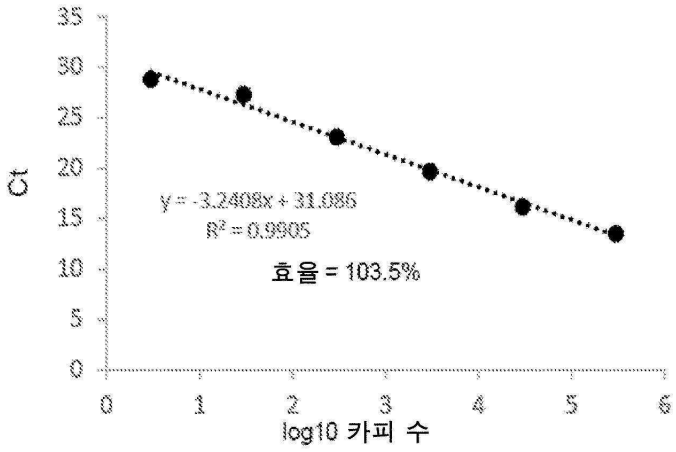

도면20f



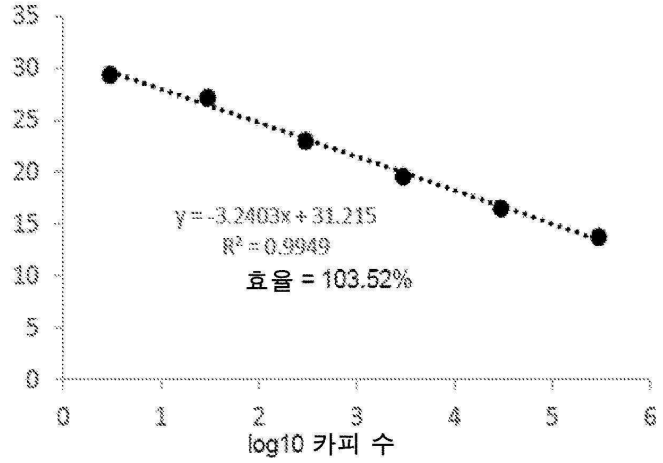

도면21a



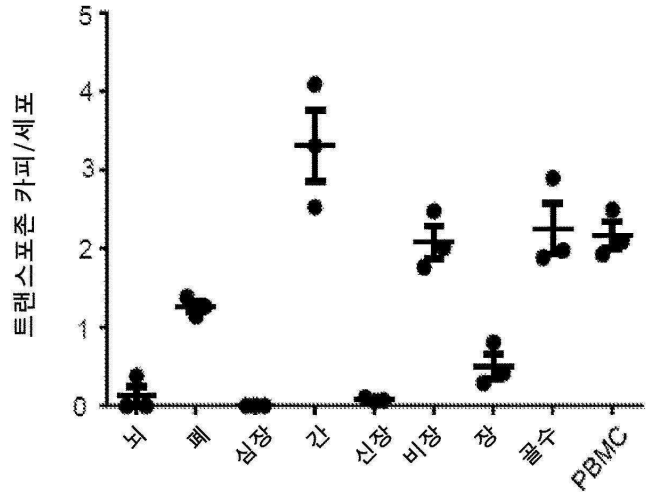
도면21b



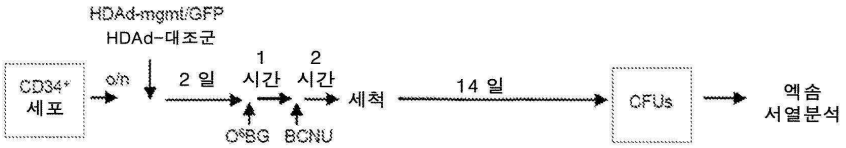

도면21c



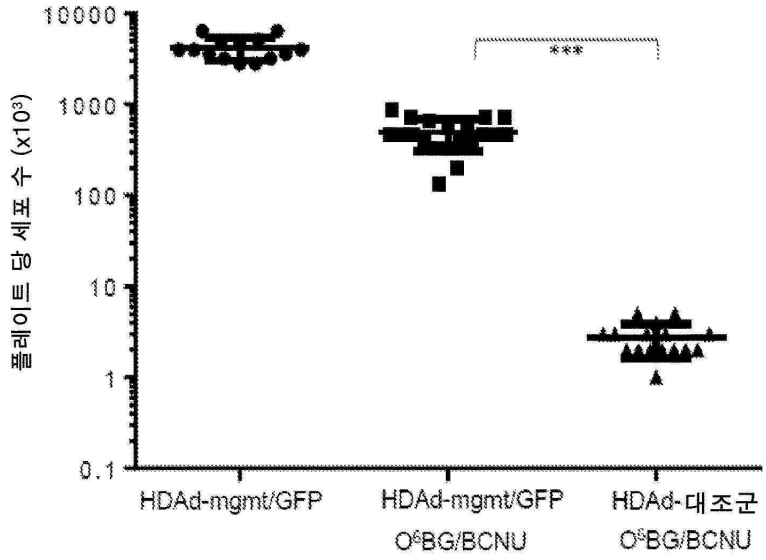

도면21d



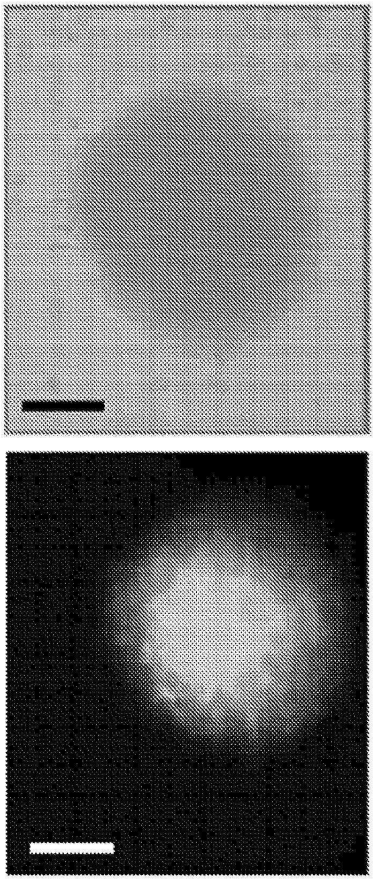
도면22a



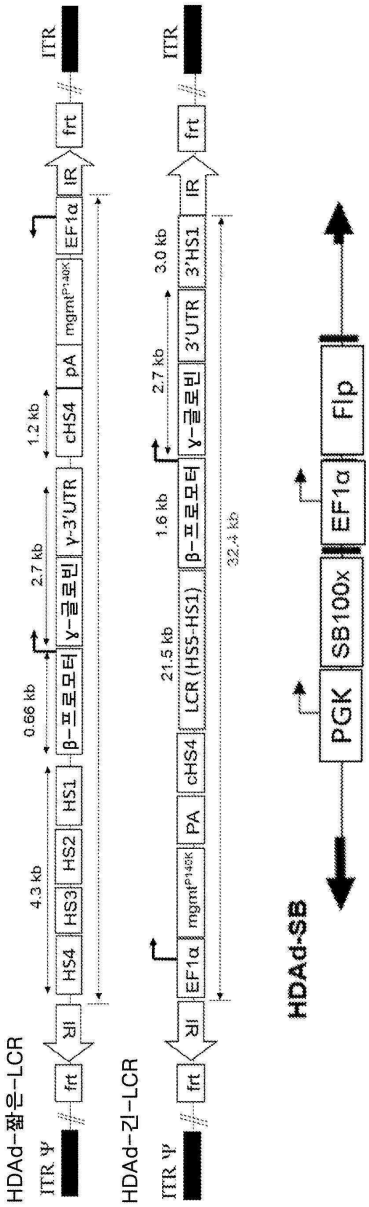

도면22b



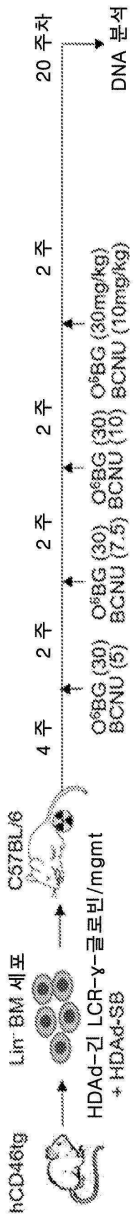

도면22c



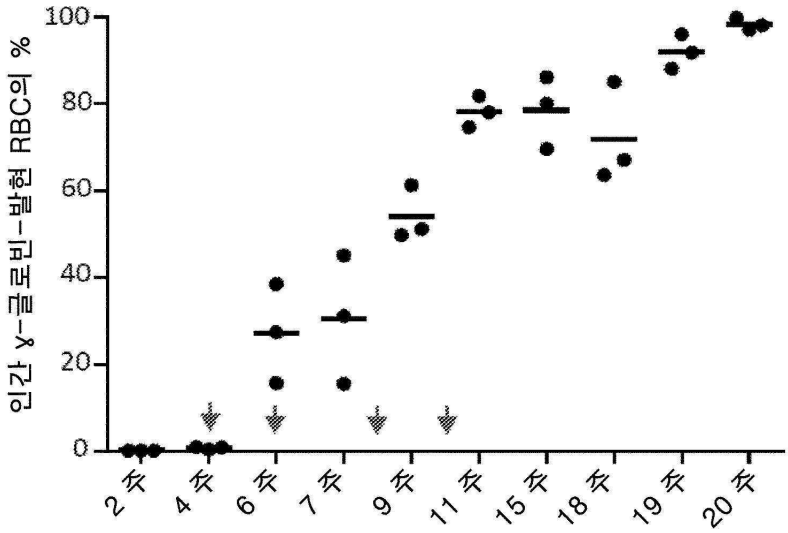
도면23


HDAd-짧은-LCR
ITR Ψ
frt
IR
HS4
HS3
HS2
HS1
β-프로모터
γ-글로빈
γ-3'UTR
cHS4
pA
mgmt
EF1α
IR
frt
ITR
4.3 kb
0.66 kb
2.7 kb
1.2 kb
HDAd-긴-LCR
ITR Ψ
frt
IR
EF1α
mgmt
PA
cHS4
LCR (HS5-HS1)
β-프로모터
γ-글로빈
3'UTR
3'HS1
IR
frt
ITR
21.5 kb
1.6 kb
2.7 kb
3.0 kb
32.4 kb
HDAd-SB
PGK
SB100x
EF1α
Flp

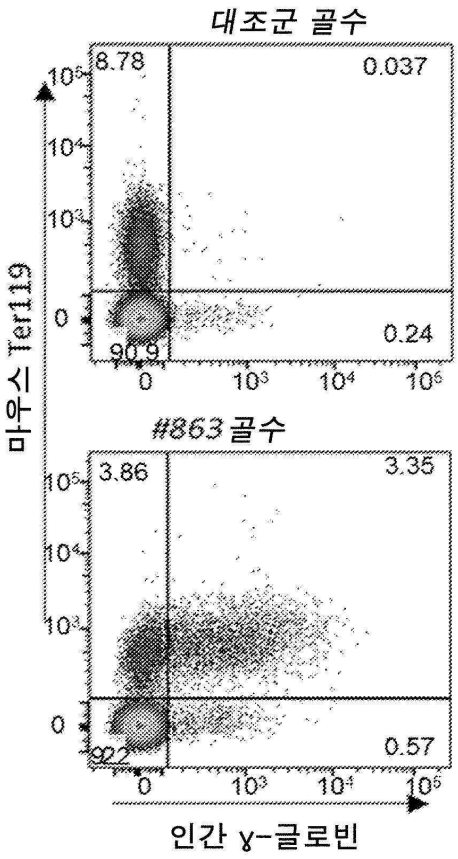
도면24a


hCD46tg
Lin⁻ BM 세포
HDAd-긴 LCR-γ-글로빈/mgmt
+ HDAd-SB
C57BL/6
4 주
O⁶BG (30)
BCNU (5)
2 주
O⁶BG (30)
BCNU (7.5)
2 주
O⁶BG (30)
BCNU (10)
2 주
O⁶BG (30mg/kg)
BCNU (10mg/kg)
2 주
20 주차
DNA 분석

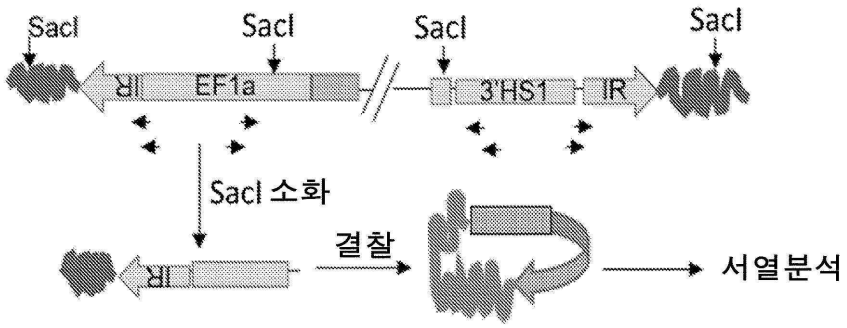
도면24b


인간 γ-글로빈-발현 RBC의 %
100
80
60
40
20
0
2 주
4 주
6 주
7 주
9 주
11 주
15 주
18 주
19 주
20 주

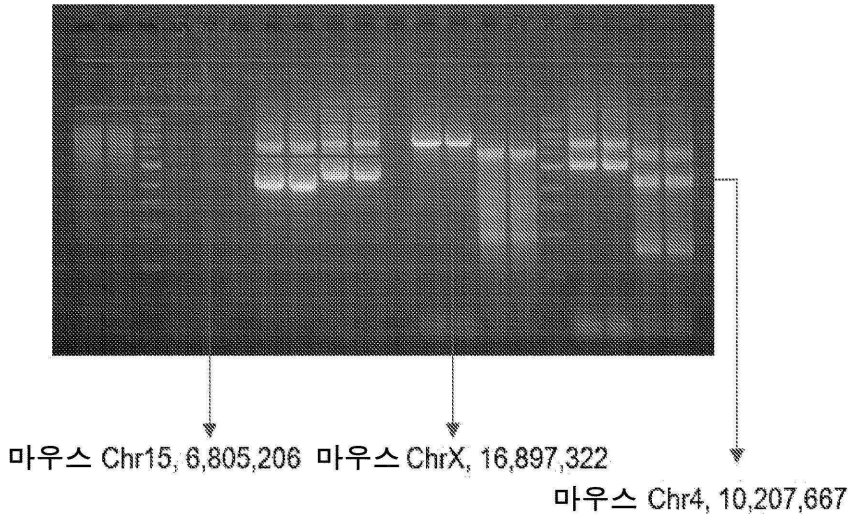
도면24c


대조군 골수
8.78
0.037
90.9
0.24
#863 골수
3.86
3.35
92
0.57
마우스 Ter119
인간 γ-글로빈

도면24d


SacI
SacI
SacI
SacI
IR
EF1a
3'HS1
IR
SacI 소화
결찰
IR
서열분석

도면24e


마우스 Chr15, 6,805,206
마우스 ChrX, 16,897,322
마우스 Chr4, 10,207,667

CCCTGGGATTCCCCAAGGCAGGGGCGAGTCCTTTTGTATGAATTACTCAAATCGATAACTAGAAACTTAATTAACAACGA
GATCTTATAATTTGCATACTTCTGCCTGCTGGGGACTTTCCACACCCTAGCTGACACAAGAATTTGAAATACATCCACAG
GTACACCTCCAATTGACTCAAATGATGTCAATTAGTCTATCATAATCTTCTAAAGCCATGACATCATTTTAACTGGAATT
TTCCAAGCTGTTTAAAGGCACAGTCAACTTAGTGTATGTAAACTTCTGACCCACTGGAATTGTGATACAGTGAATTATAA
GTGAAATAATCTGTCTGTAAACAATTGTTGGAAAAATGACTTGTGTCATGCACAAAGTAGATGTCCTAACTGACTTGCCA
AAACTATTGTTTGTTAACAAGAAATTTGTGGAGTAGTTGAAAAACGAGTTTTAATGACTCCAACTTAAGTGTATGTAAAC
TTCCGACTTCAACTG[TA]**AGAATGGCCATTCATCTATAGTAGCACACAATATTTGCATTTGTGCGACAGTATAAGGGACA**
**ATTATGCTATCAGGCATTTTTCCAAAGTGAGTAATCGAAGTTTTTATACCTTTGTGTGCCATGTTTGCTACCATGGTGGG**
**ATAATCTTACACGCGTTCTCGCGACCGGCCAGGAAAGACGCAACAAACCGGAATCTTCTGCGGCAAAAGCTTTATTGCTT**

서열번호: 1

TAGAAACTTAATTAACAACGAGATCTTATAATTTGCATACTTCTGCCTGCTGGGGACTTTCCACACCCTAGCTGACACAA
GAATTTGAAATACATCCACAGGTACACCTCCAATTGACTCAAATGATGTCAATTAGTCTATCATAATCTTCTAAAGCCAT
GACATCATTTTAACTGGAATTTTCCAAGCTGTTTAAAGGCACAGTCAACTTAGTGTATGTAAACTTCTGACCCACTGGAA
TTGTGATACAGTGAATTATAAGTGAAATAATCTGTCTGTAAACAATTGTTGGAAAAATGACTTGTGTCATGCACAAAGTA
GATGTCCTAACTGACTTGCCAAAACTATTGTTTGTTAACAAGAAATTTGTGGAGTAGTTGAAAAACGAGTTTTAATGACT
CCAACTTAAGTGTATGTAAACTTCCGACTTCAACTG[TA]**CAAGTAGACCAAATATCCATATACATAAAAGAAAAAAATAGA**
**AAAAATTTCTAGTGACAGAAAAATGACAAAGAACATACTGCTTTATTACTACTATTAAGATGTTTGCTTCCATTACACTC**
**ATATGAGTCATGATATTTTTTCTTCATTTTTTTCTANTNNCACTNGAAAT**

서열번호: 2

GTTGCTAGGAATGAGCCAAATTCATCTGTATTAAACAGTGGGAGCTTGTGGAAGGCTACTCGAAATGTTTGACCCAAGTT
AAACAATTTAAAGGCAATGCTACCAAATACTAATTGAGTGTATGTTAACTTCTGACCCACTGGGAATGTGATGAAAGAAA
TAAAAGCTGAAATGAATCATTCTCTCTACTATTATTCTGATATTTCACATTCTTAAAATAAAGTGGTGATCCTAACTGAC
CTTAAGACAGGGAATCTTTACTCGGATTAAATGTCAGGAATTGTGAAAAAGTGAGTTTAAATGTATTTGGCTAAGGTGTA
TGTAAACTTCCGACTTCAACTG[TA]**TATCCTCCCCGTTGCACCCTCTTGATGATGCTGAGATGAACACAGATGCTCACTCC**
**TTGAGGGCTCTAAGCTTATGCTGACACAGACACAGGTGCTCACTTCTATGAATGGCCTAAGATTTGAGGACATCATGAGG**
**ACAAGTGTGATAAAATCTTGGAACAACCTCCCAGAGGTCT**

서열번호: 3

5' 단부 벡터 서열
잠자는 미녀 (sleeping beauty)
IR/DR 서열
통합 접합부 **(chr15, 6805206)**

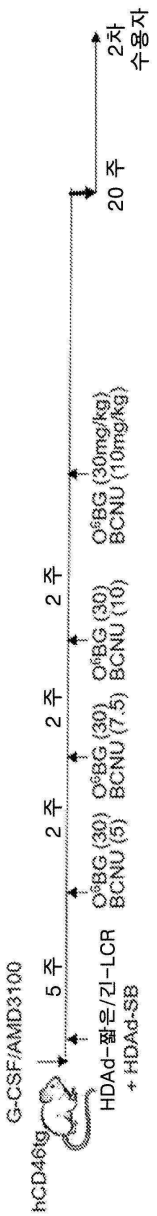
5' 단부 벡터 서열
잠자는 미녀 IR/DR 서열
통합 접합부 **(chrX, 16897322)**

3' 단부 벡터 서열
잠자는 미녀 IR/DR 서열
통합 접합부 **(chr4, 10207667)**

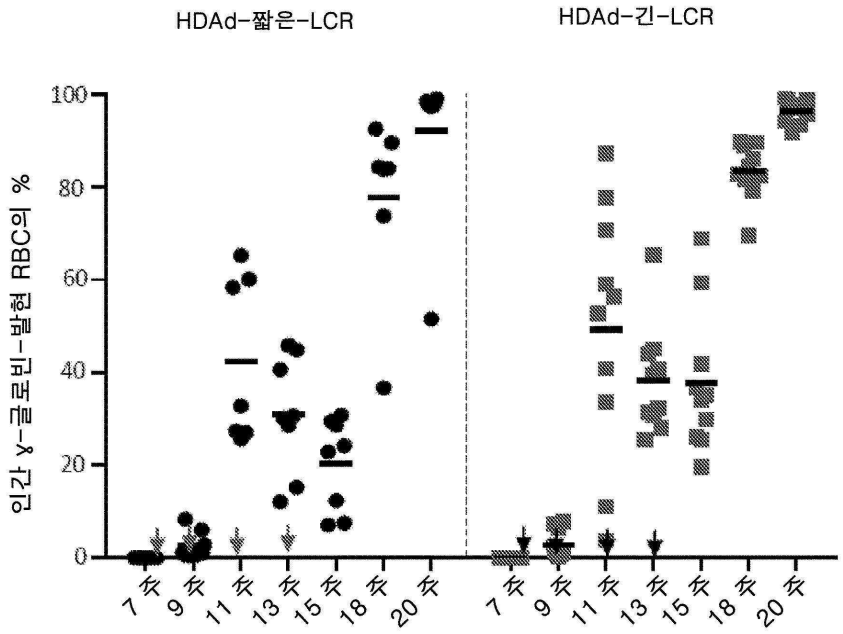


공개특허 10-2022-0038362

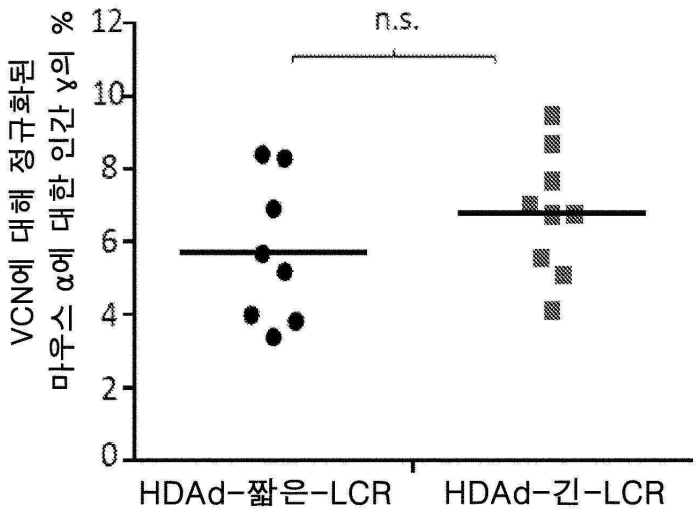
도면25a


hCD46tg
G-CSF/AMD3100
5 주
HDAd-폴은/긴-LCR
+ HDAd-SB
2 주
O6BG (30)
BCNU (5)
2 주
O6BG (30)
BCNU (7.5)
2 주
O6BG (30)
BCNU (10)
O6BG (30mg/kg)
BCNU (10mg/kg)
20 주
2차
수용자

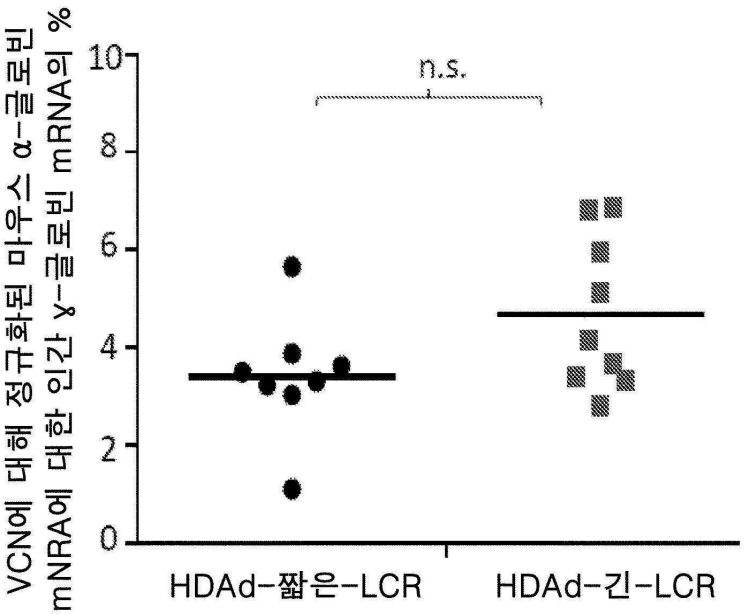
도면25b


HDAd-짧은-LCR
HDAd-긴-LCR
인간 γ-글로빈-발현 RBC의 %
100
80
60
40
20
0
7 주
9 주
11 주
13 주
15 주
18 주
20 주

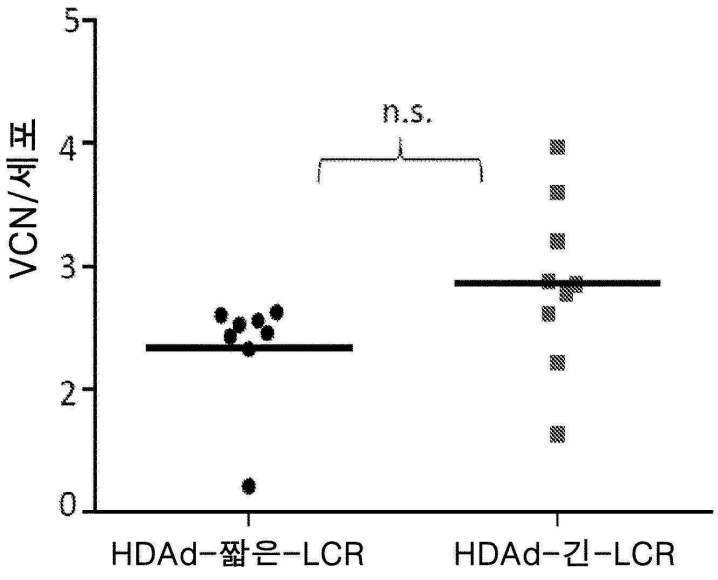
도면25c


VCN에 대해 정규화된
마우스 α에 대한 인간 γ의 %
12
10
8
6
4
2
0
n.s.
HDAd-짧은-LCR
HDAd-긴-LCR

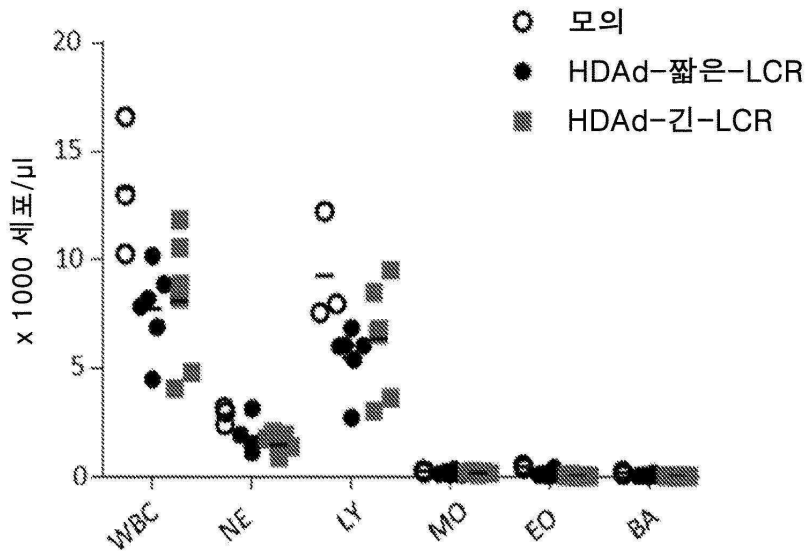
도면25d


VCN에 대해 정규화된 마우스 α-글로빈 mNRA에 대한 인간 γ-글로빈 mRNA의 %
10
8
6
4
2
0
n.s.
HDAd-짧은-LCR
HDAd-긴-LCR

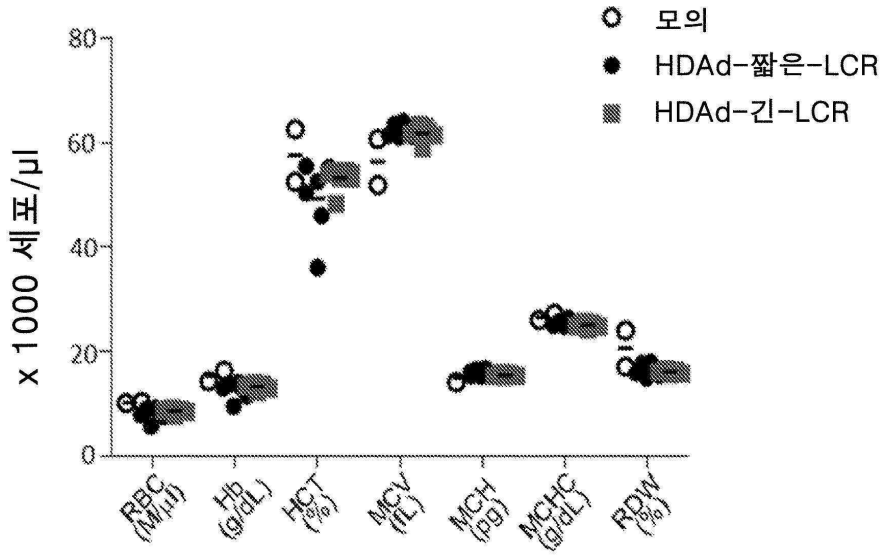
도면25e


VCN/세포
5
4
3
2
0
n.s.
HDAd-짧은-LCR
HDAd-긴-LCR

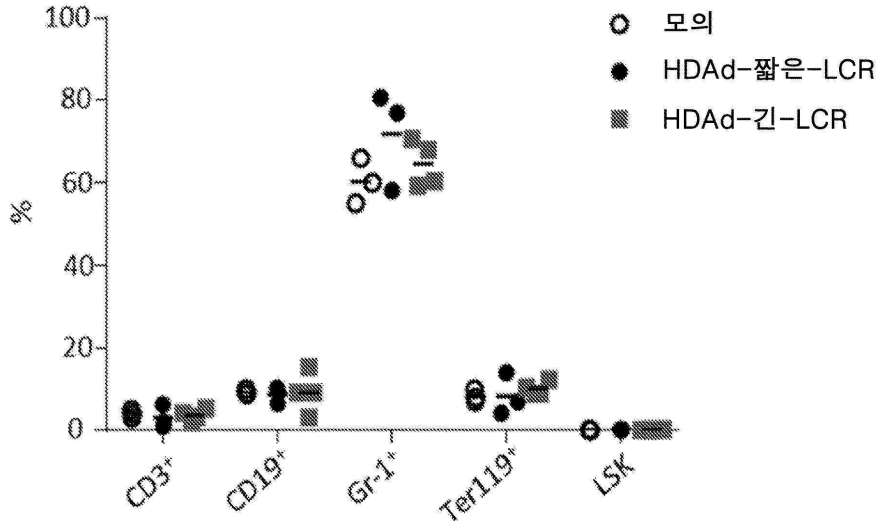
도면26a


모의
HDAd-짧은-LCR
HDAd-긴-LCR
x 1000 세포/μl
20
15
10
5
0
WBC
NE
LY
MO
EO
BA

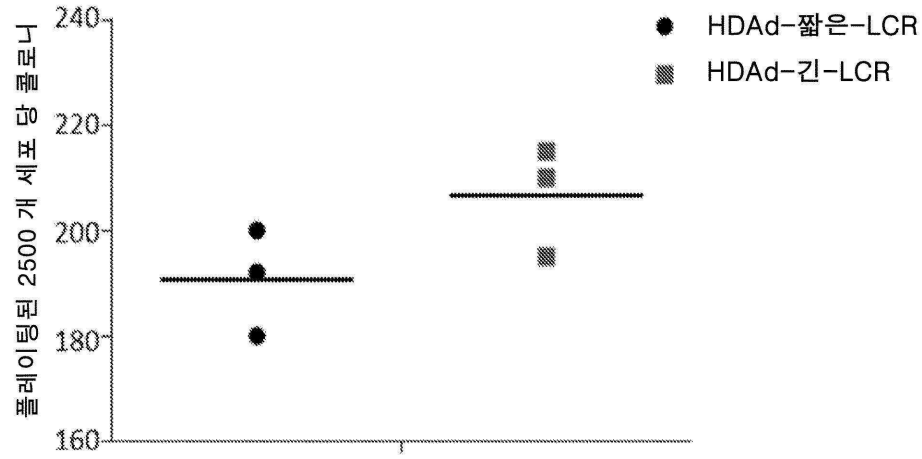
도면26b


모의
HDAd-짧은-LCR
HDAd-긴-LCR
x 1000 세포/μl
80
60
40
20
0
RBC (M/μl)
Hb (g/dL)
HCT (%)
MCV (fL)
MCH (pg)
MCHC (g/dL)
RDW (%)

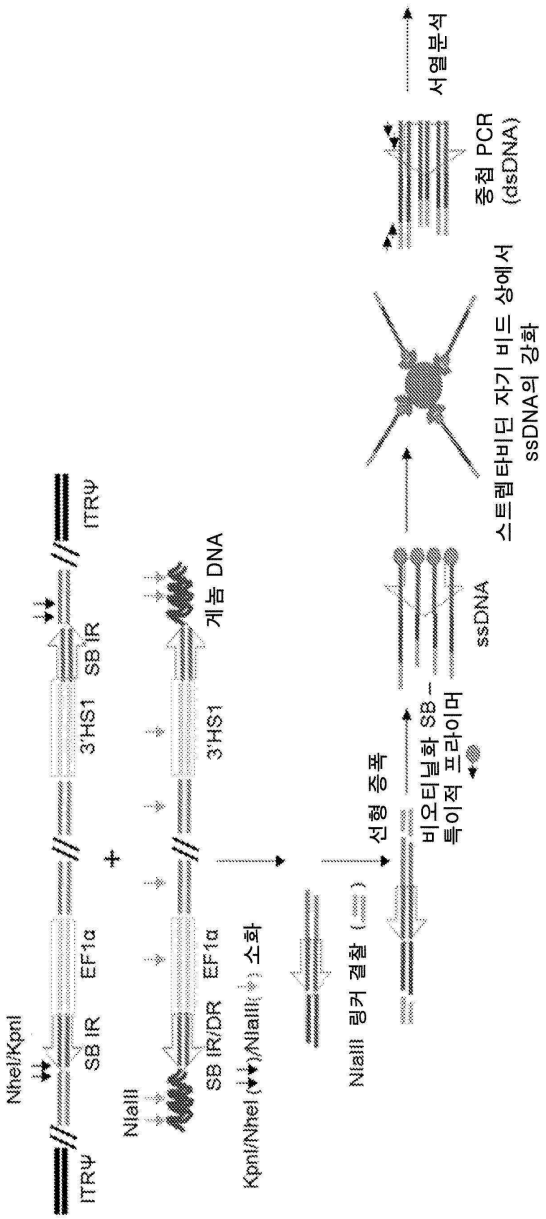
도면26c


100
80
60
40
20
0
%
모의
HDAd-짧은-LCR
HDAd-긴-LCR
CD3+
CD19+
Gr-1+
Ter119+
LSK

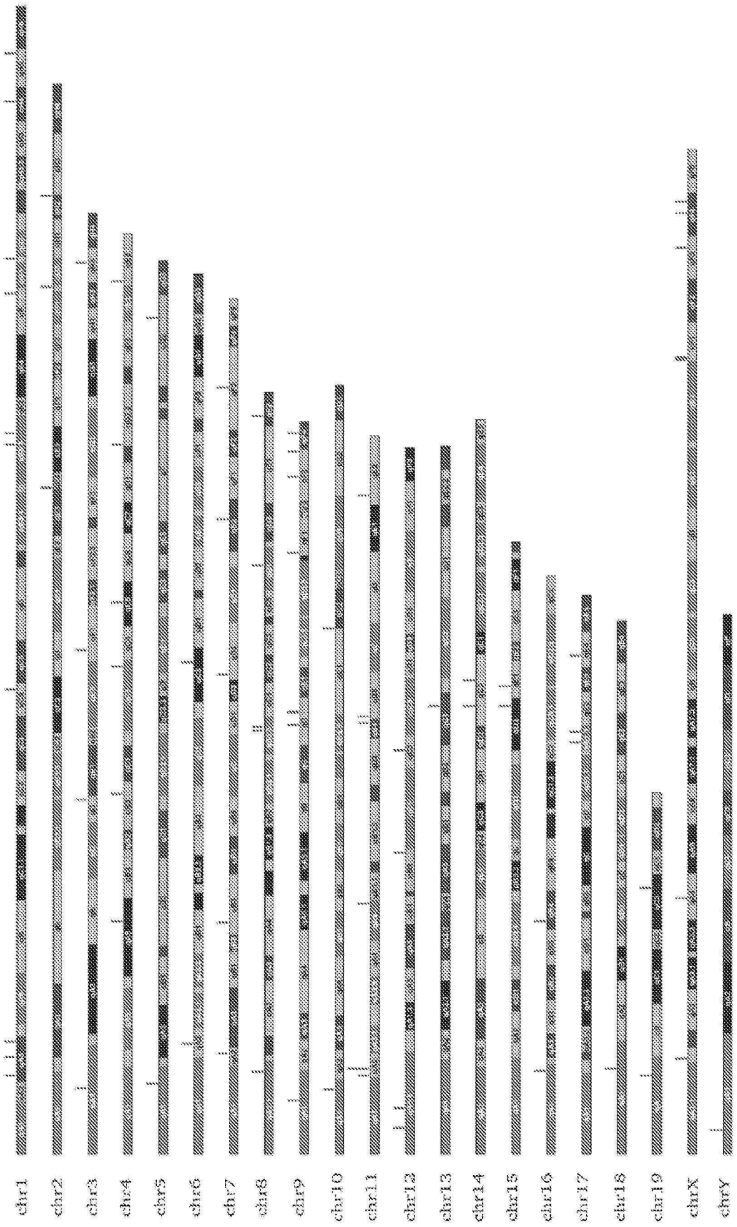
도면26d


240
220
200
180
160
폴레이팅된 2500 개 세포 당 콜로니
HDAd-짧은-LCR
HDAd-긴-LCR

도면27


NheI/KpnI
ITRΨ
SB IR
EF1α
3'HS1
SB IR
ITRΨ
NlaIII
SB IR/DR
EF1α
3'HS1
게놈 DNA
KpnI/NheI (↓↓)/NlaIII (↓) 소화
NlaIII 링커 결찰 ( )
선형 증폭
비오티닐화 SB 특이적 프라이머
ssDNA
스트렙타비딘 자기 비드 상에서 ssDNA의 강화
중첩 PCR (dsDNA)
서열분석

도면28a

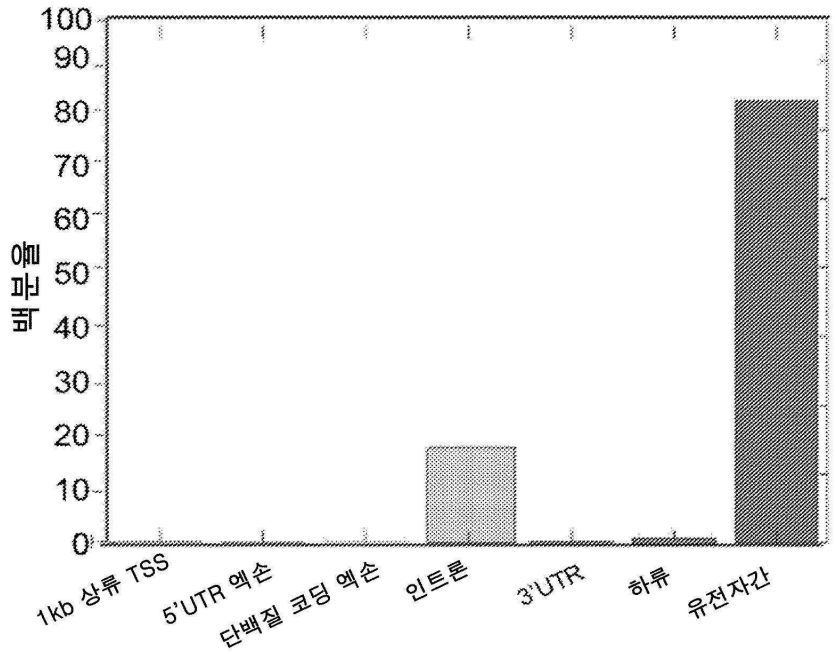

chr1
chr2
chr3
chr4
chr5
chr6
chr7
chr8
chr9
chr10
chr11
chr12
chr13
chr14
chr15
chr16
chr17
chr18
chr19
chrX
chrY

도면28b

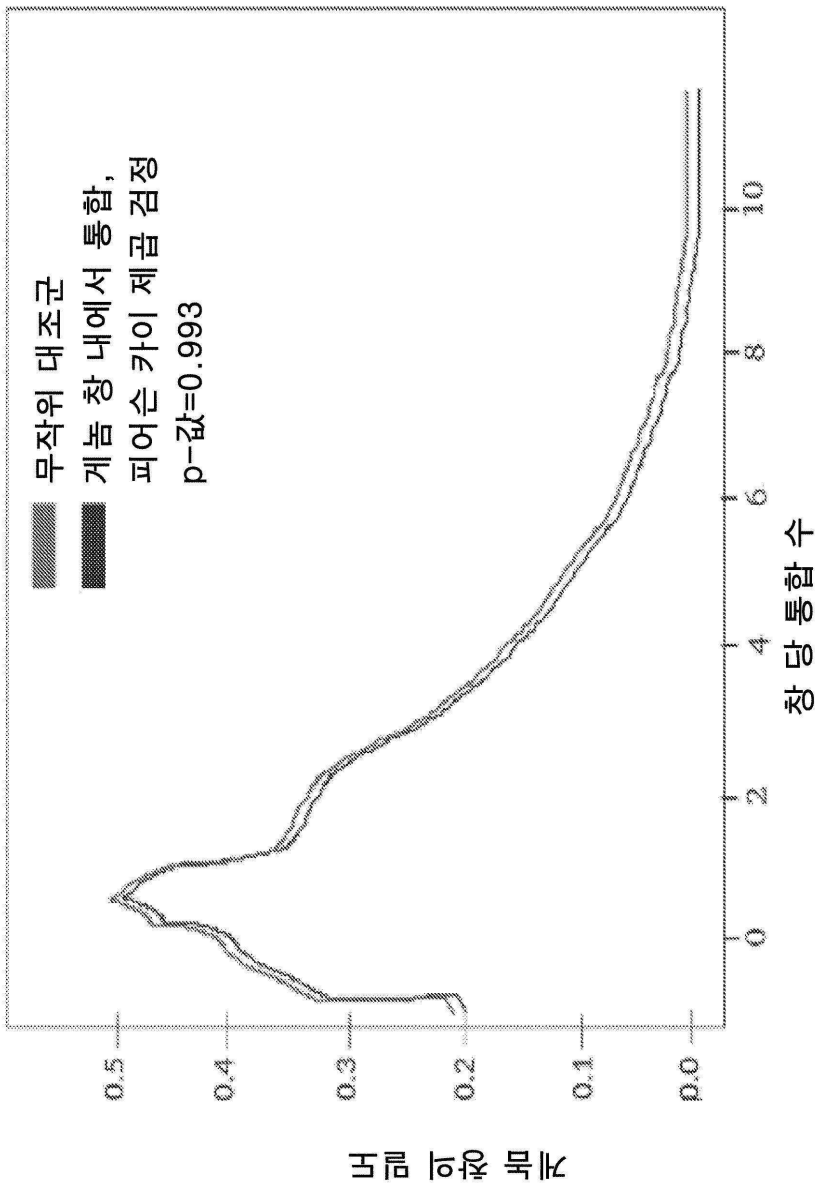
참조는 이녀 IR/DR 서열
융합 접합부 (chr7, 79796094)
ACTTAAGTGT ATGTAAACTT CCGACTTCAA CTGTAGGGTA CGTGATTCTC TGGGCATCTC TGCCCACTAC CATG
서열번호: 4

참조는 이녀 IR/DR 서열
융합 접합부(반복 영역)
ACTTAAGTGT ATGTAAACTT CCGACTTCAA CTGTAAATTT TCCACCTTTT TCAGTTTTCC TGGCCATATT TCATG
서열번호: 5

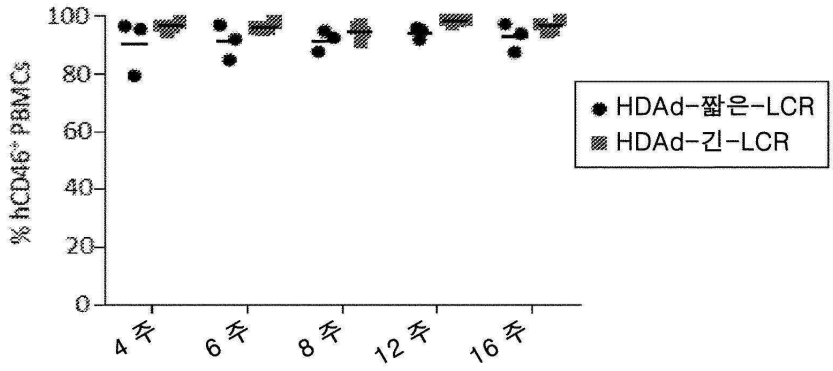
도면28c


100
90
80
70
60
50
40
30
20
10
0
백분율
1kb 상류 TSS
5'UTR 엑손
단백질 코딩 엑손
인트론
3'UTR
하류
유전자간

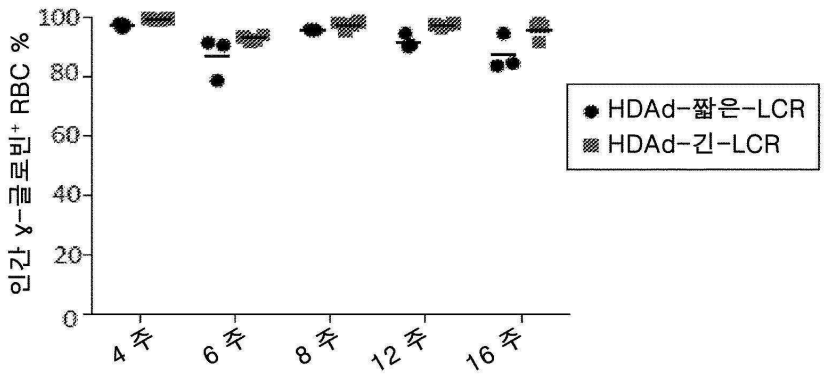
도면28d


무작위 대조군
게놈 창 내에서 통합,
피어슨 카이 제곱 검정
p-값=0.993
창 당 통합 수
0
2
4
6
8
10
0.0
0.1
0.2
0.3
0.4
0.5
게놈 창의 밀도

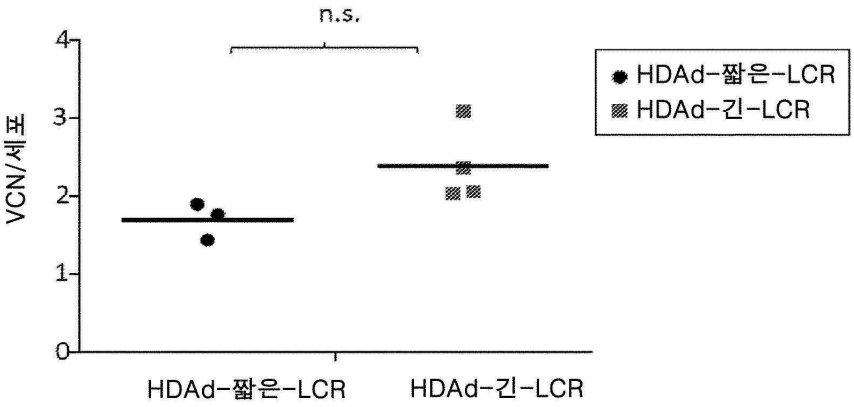
도면29a


% hCD46+ PBMCs
100
80
60
40
20
0
4 주
6 주
8 주
12 주
16 주
HDAd-짧은-LCR
HDAd-긴-LCR

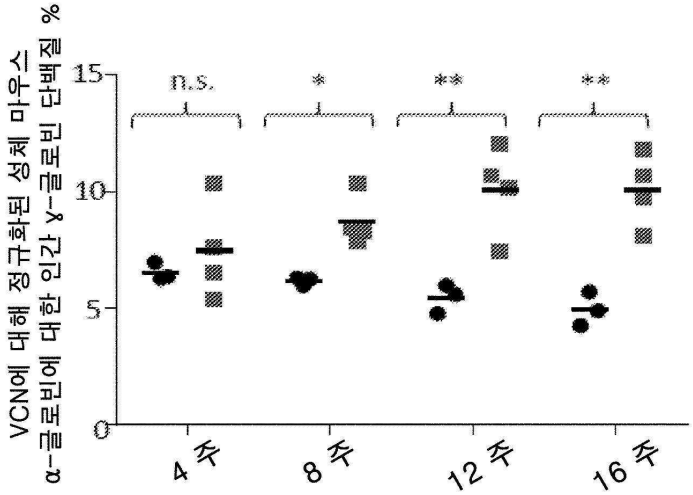
도면29b



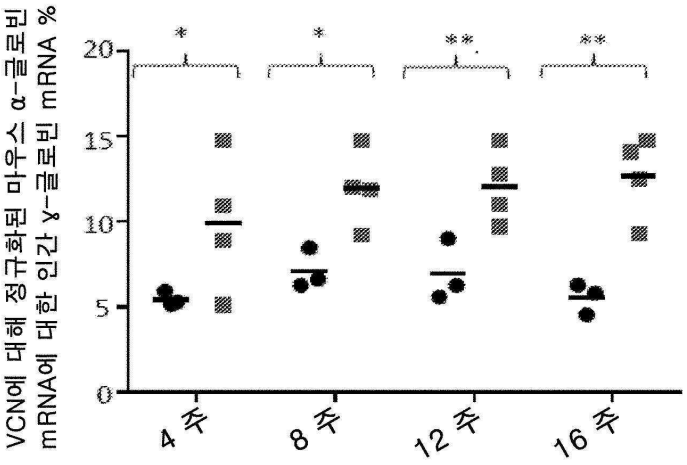

도면29c



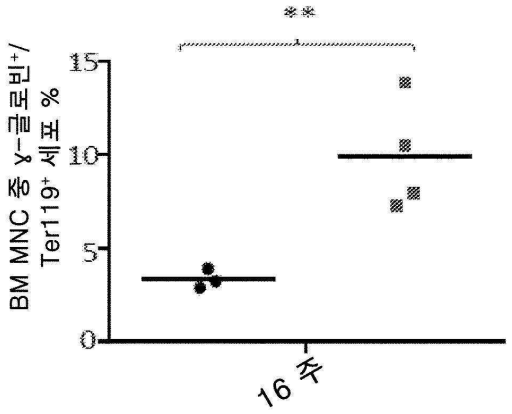

도면29d



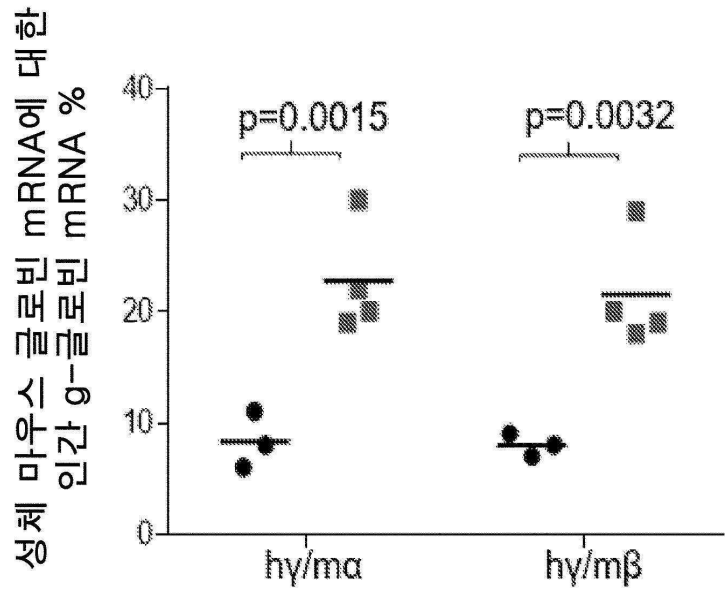
도면29e



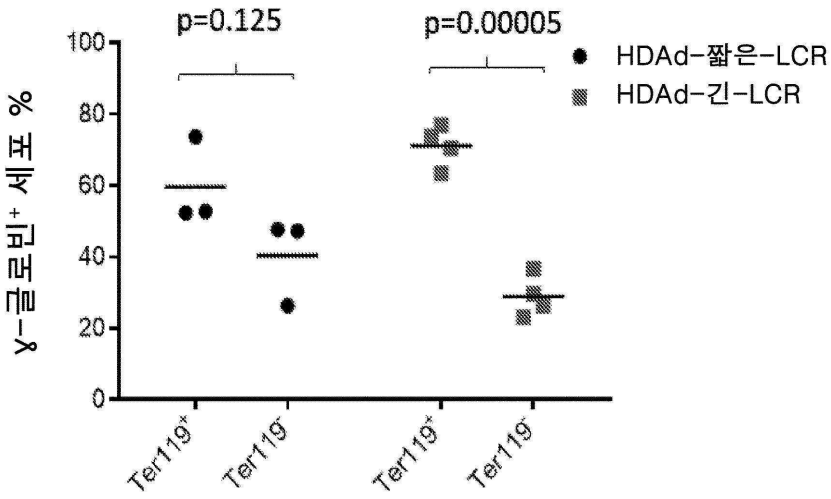

도면29f



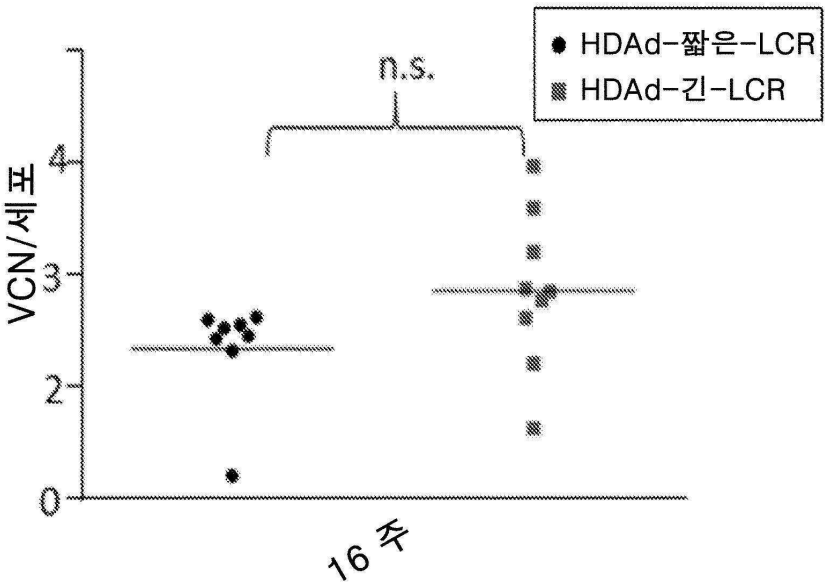

도면29g



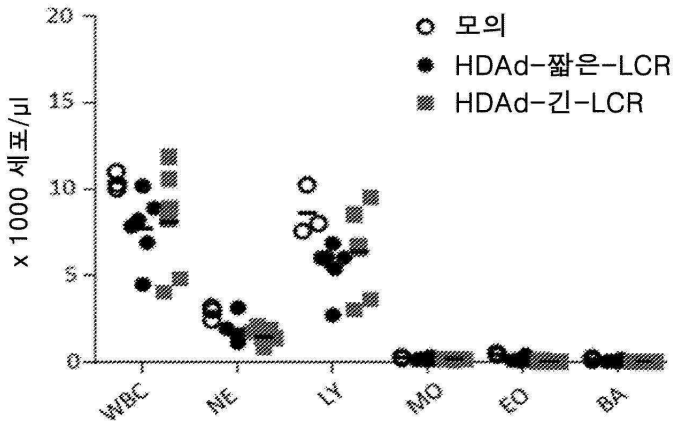
도면29h



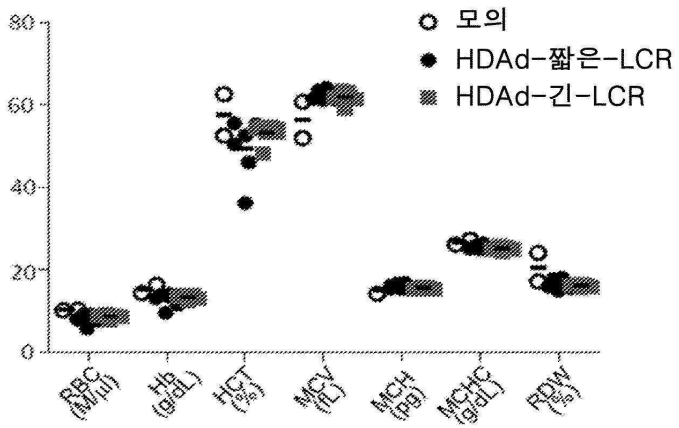

도면29i



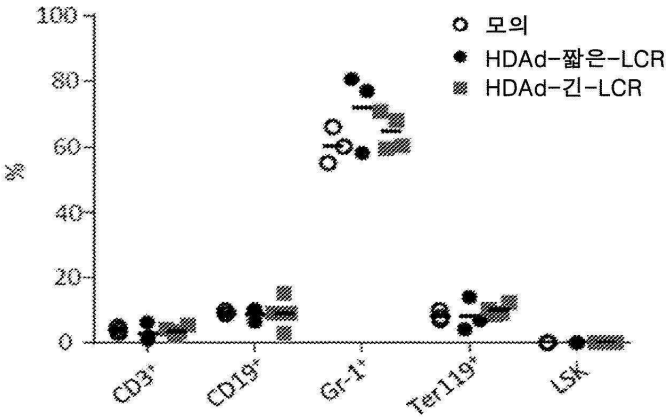

도면30a



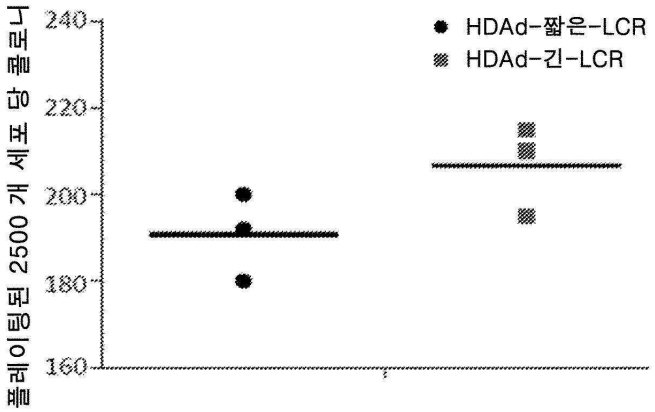
도면30b



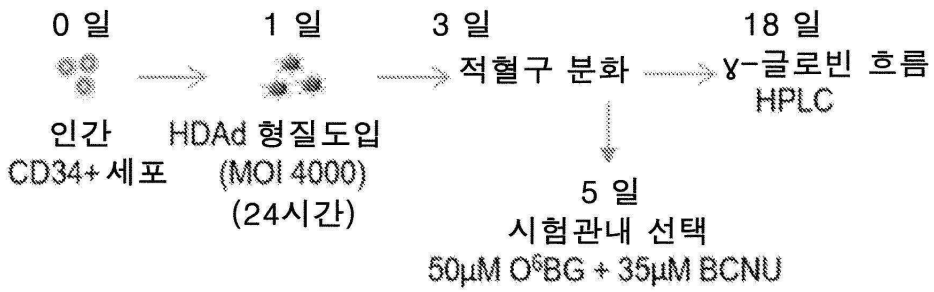

도면30c



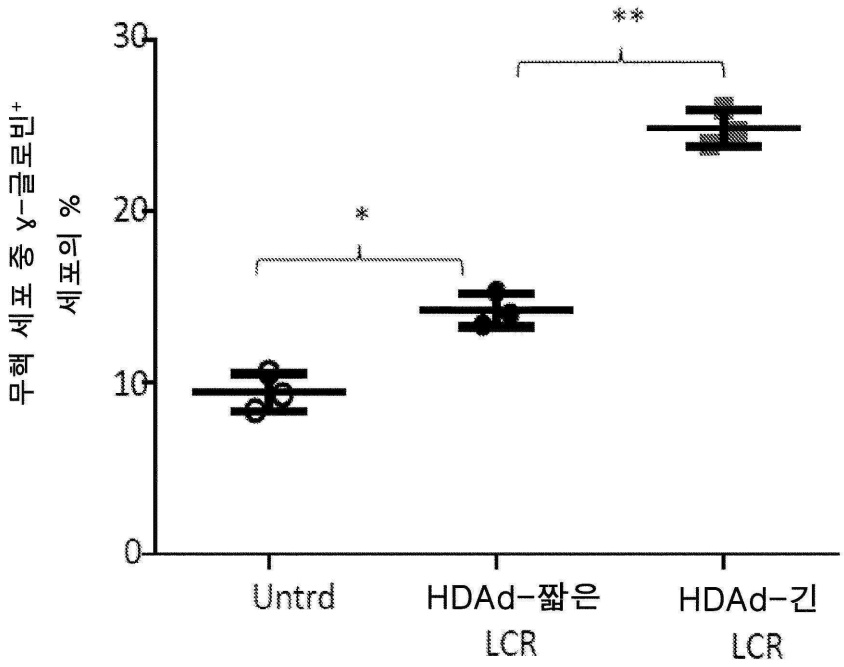

도면30d



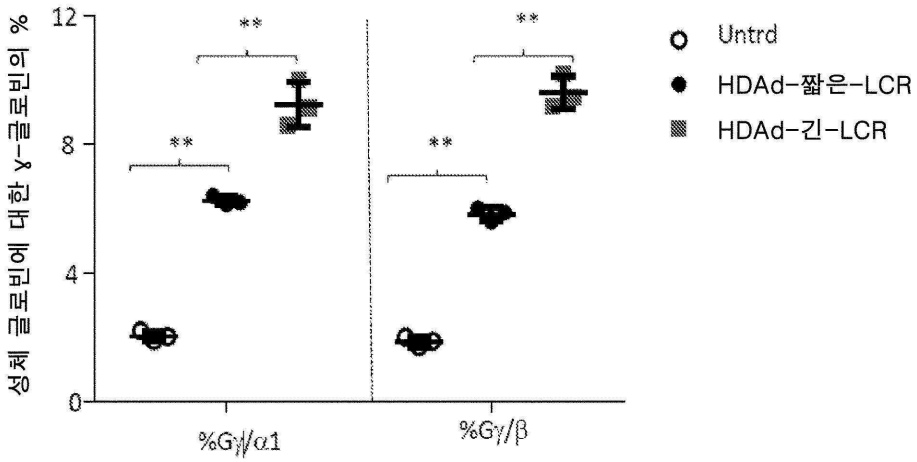
도면31a



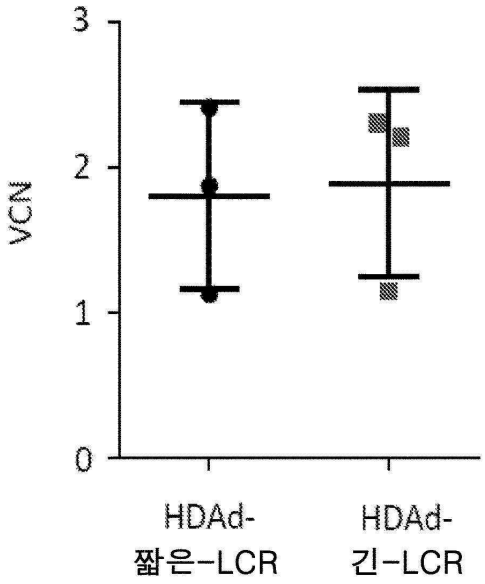

도면31b



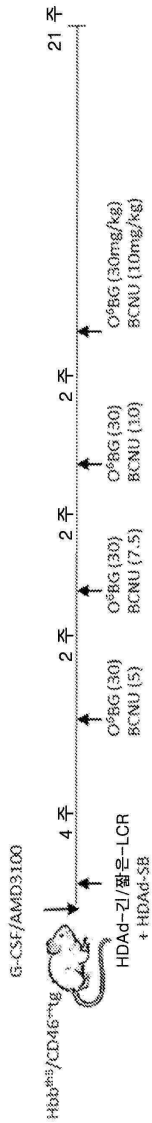

도면31c



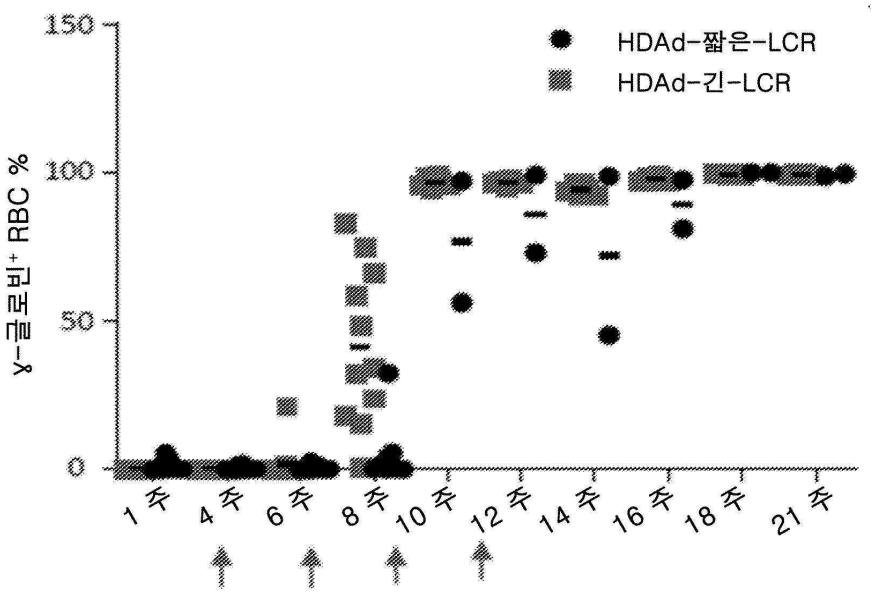
도면31d



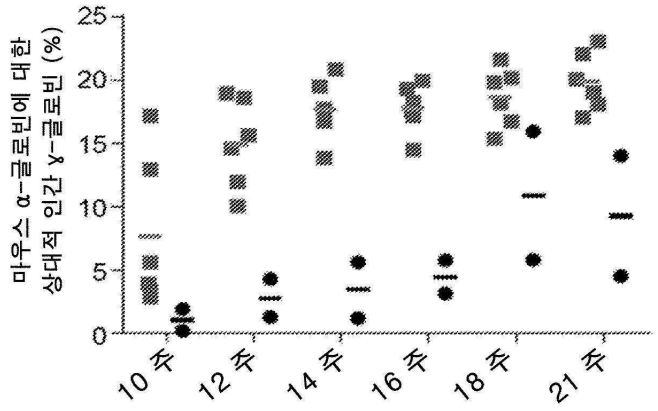
도면32a


21 주
2 주
2 주
2 주
4 주
G-CSF/AMD3100
HDAd-ZI/폴은-LCR
+ HDAd-SB
O⁶BG (30)
BCNU (5)
O⁶BG (30)
BCNU (7.5)
O⁶BG (30)
BCNU (10)
O⁶BG (30mg/kg)
BCNU (10mg/kg)

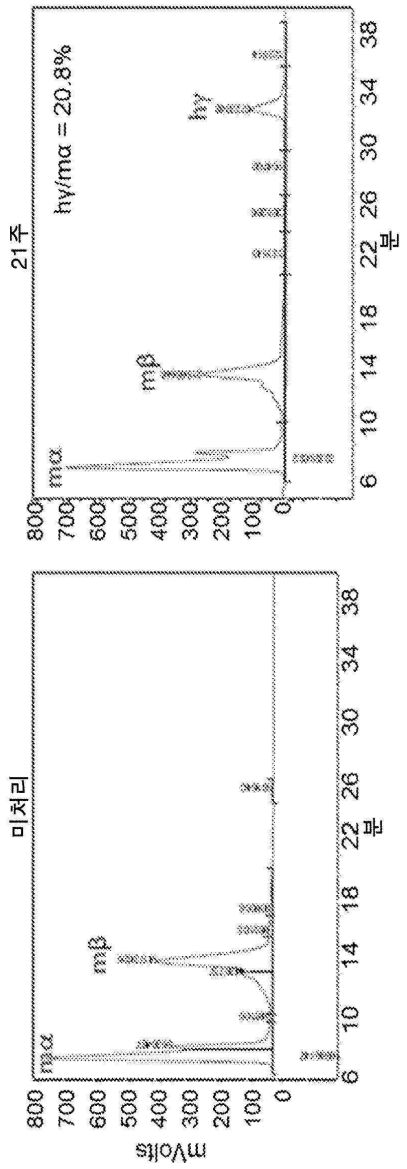
도면32b


150
100
50
0
γ-글로빈+ RBC %
HDAd-짧은-LCR
HDAd-긴-LCR
1 주
4 주
6 주
8 주
10 주
12 주
14 주
16 주
18 주
21 주

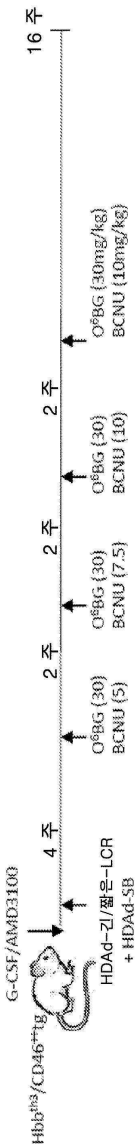
도면32c


25
20
15
10
5
0
마우스 α-글로빈에 대한
상대적 인간 γ-글로빈 (%)
10 주
12 주
14 주
16 주
18 주
21 주

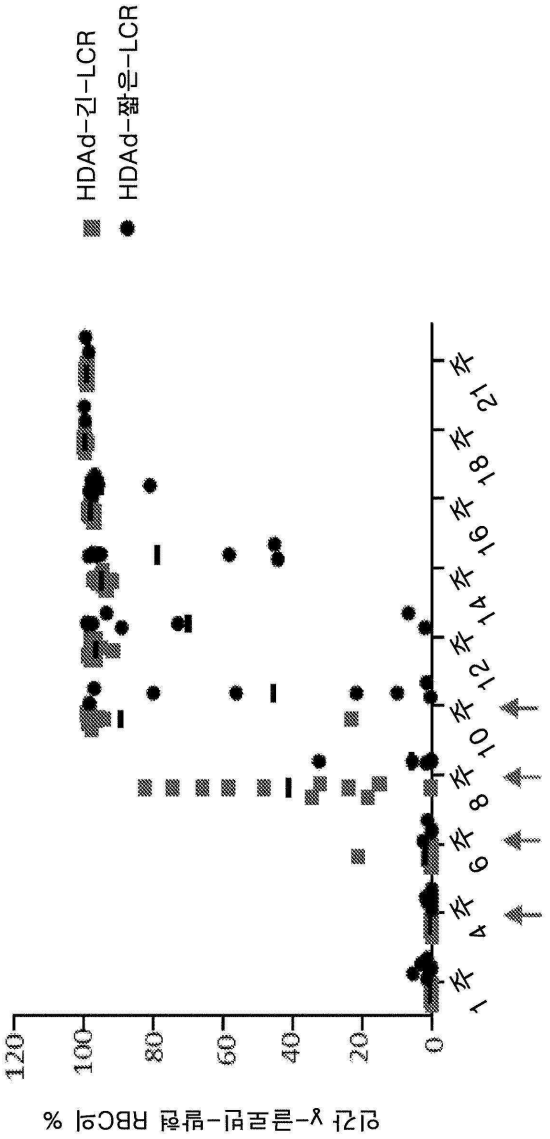
도면32d



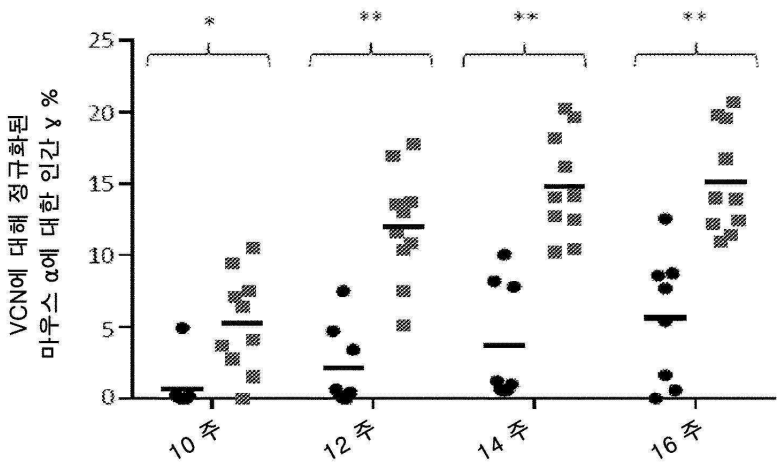
도면32e


16 주
2 주
2 주
2 주
2 주
4 주
G-CSF/AMD3100
Hbb^th3/CD46^++tg
HDAd-긴/짧은-LCR
+ HDAd-SB
O6BG (30)
BCNU (5)
O6BG (30)
BCNU (7.5)
O6BG (30)
BCNU (10)
O6BG (30mg/kg)
BCNU (10mg/kg)

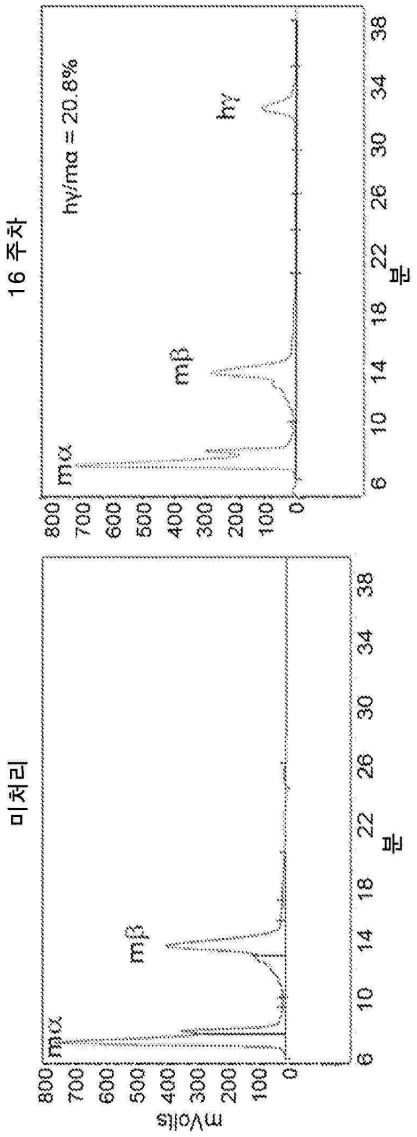
도면32f


HDAd-긴-LCR
HDAd-짧은-LCR
인간 γ-글로빈-발현 RBC의 %
120
100
80
60
40
20
0
1 주
4 주
6 주
8 주
10 주
12 주
14 주
16 주
18 주
21 주

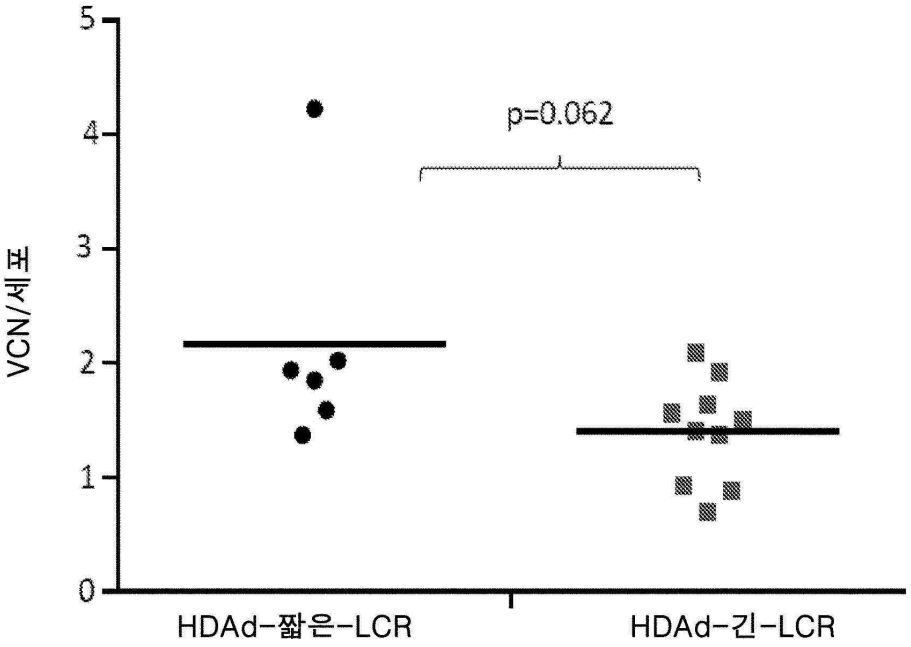
도면32g


VCN에 대해 정규화된
마우스 α에 대한 인간 γ %
25
20
15
10
5
0
*
**
**
**
10 주
12 주
14 주
16 주

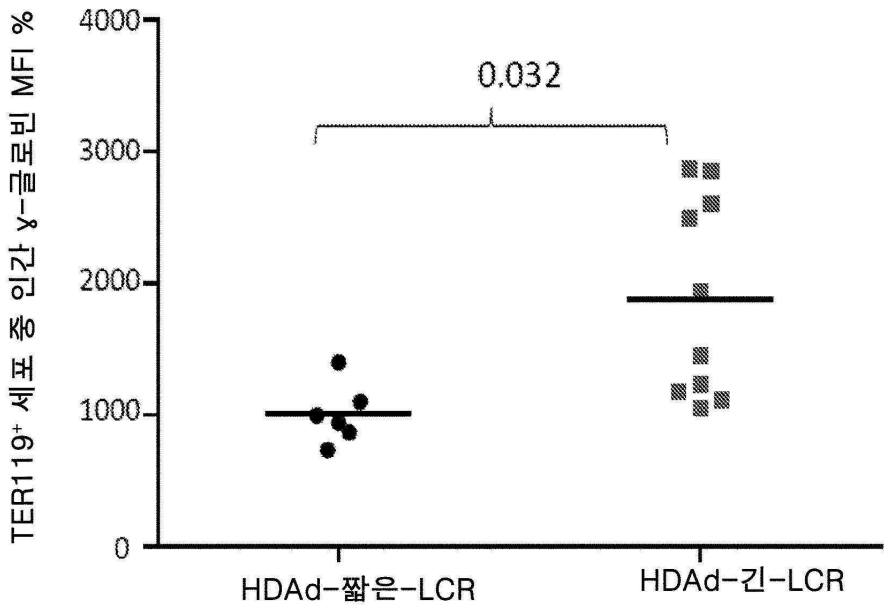
도면32h



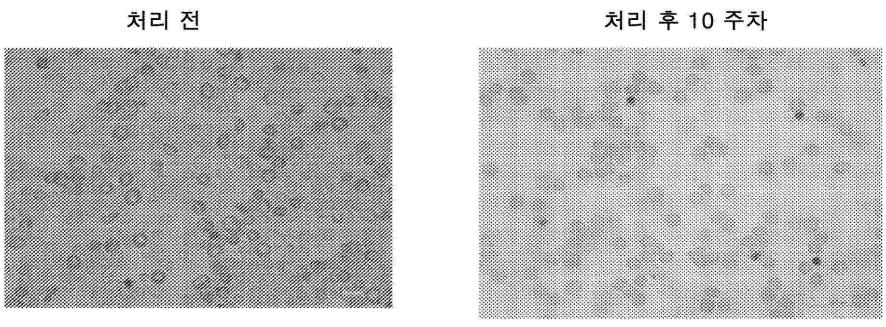
도면33a



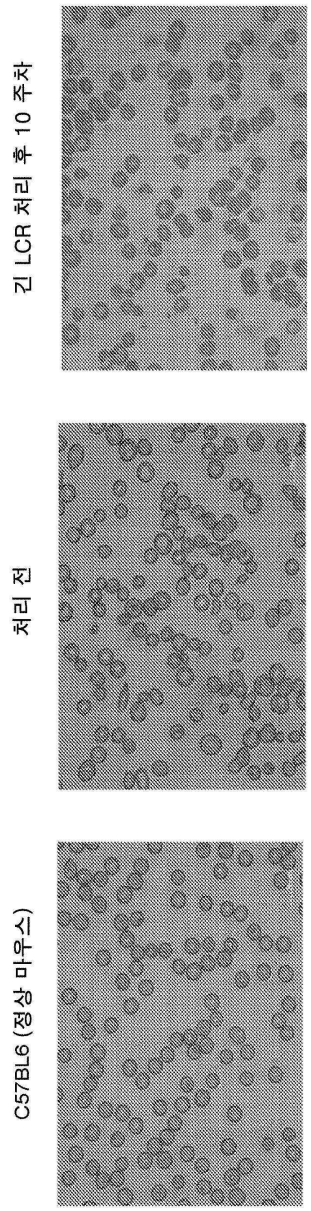

도면33b



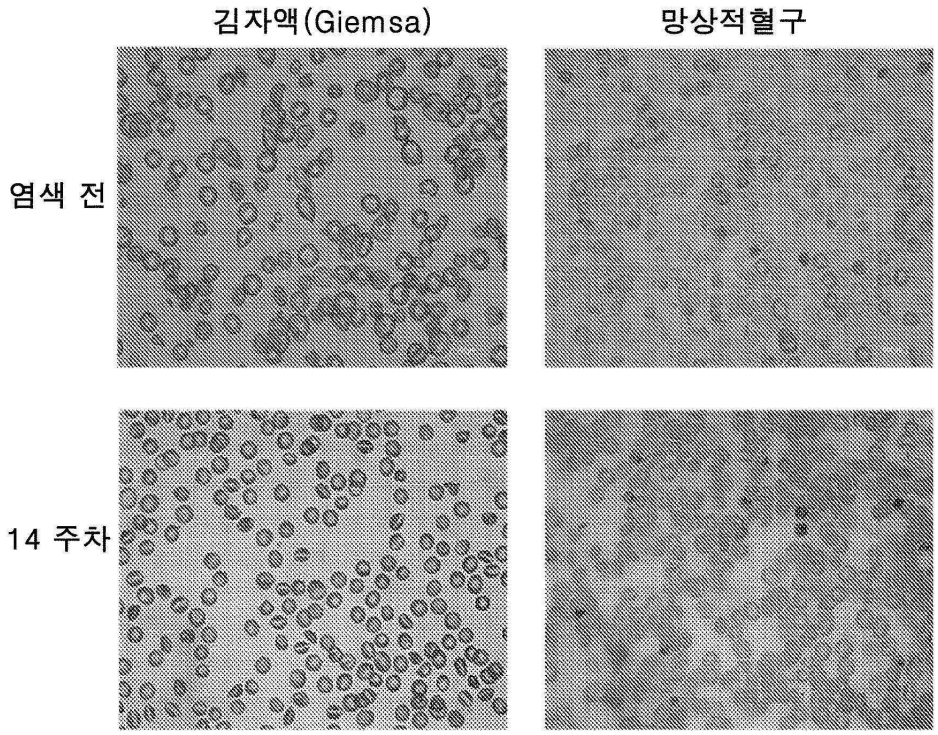

도면34



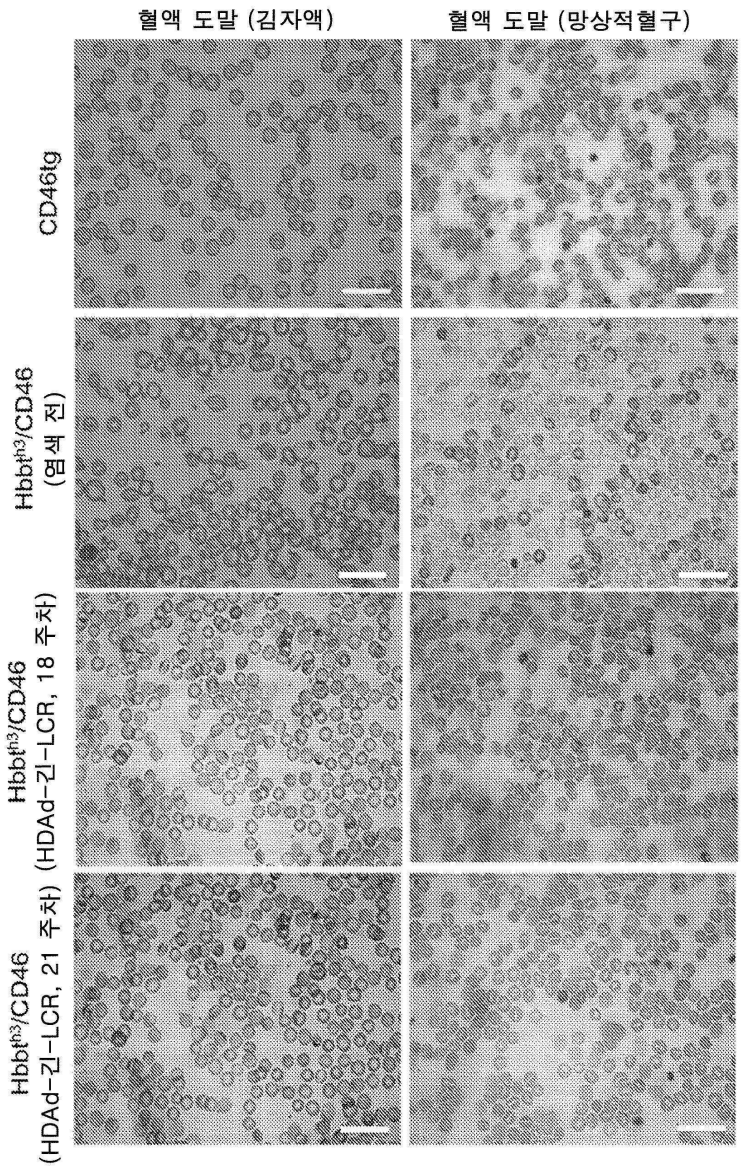
도면35



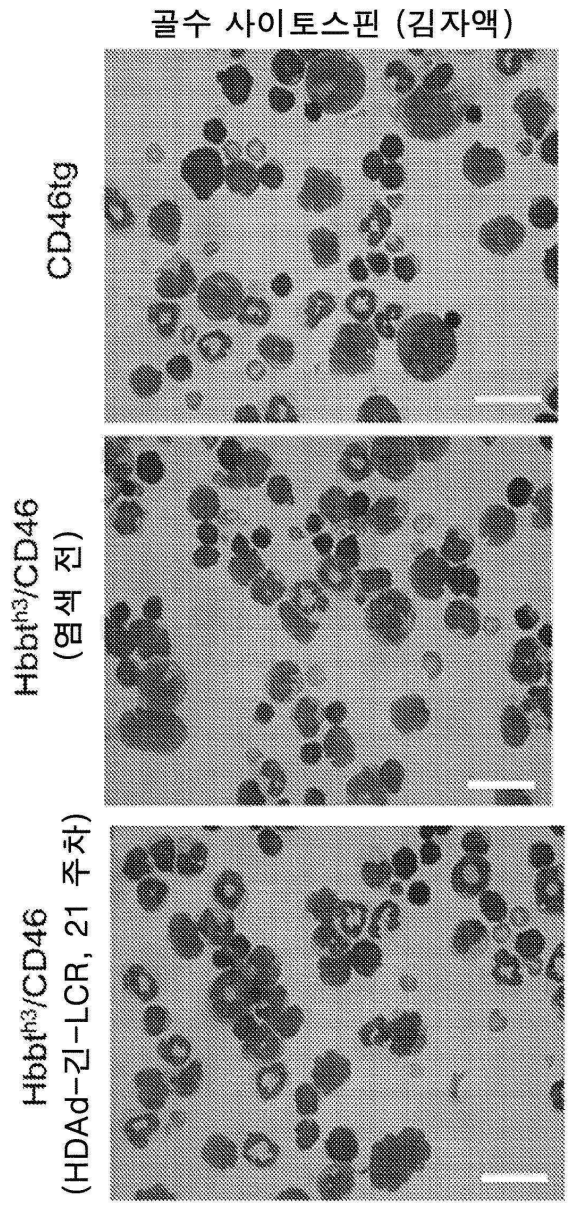
도면36a



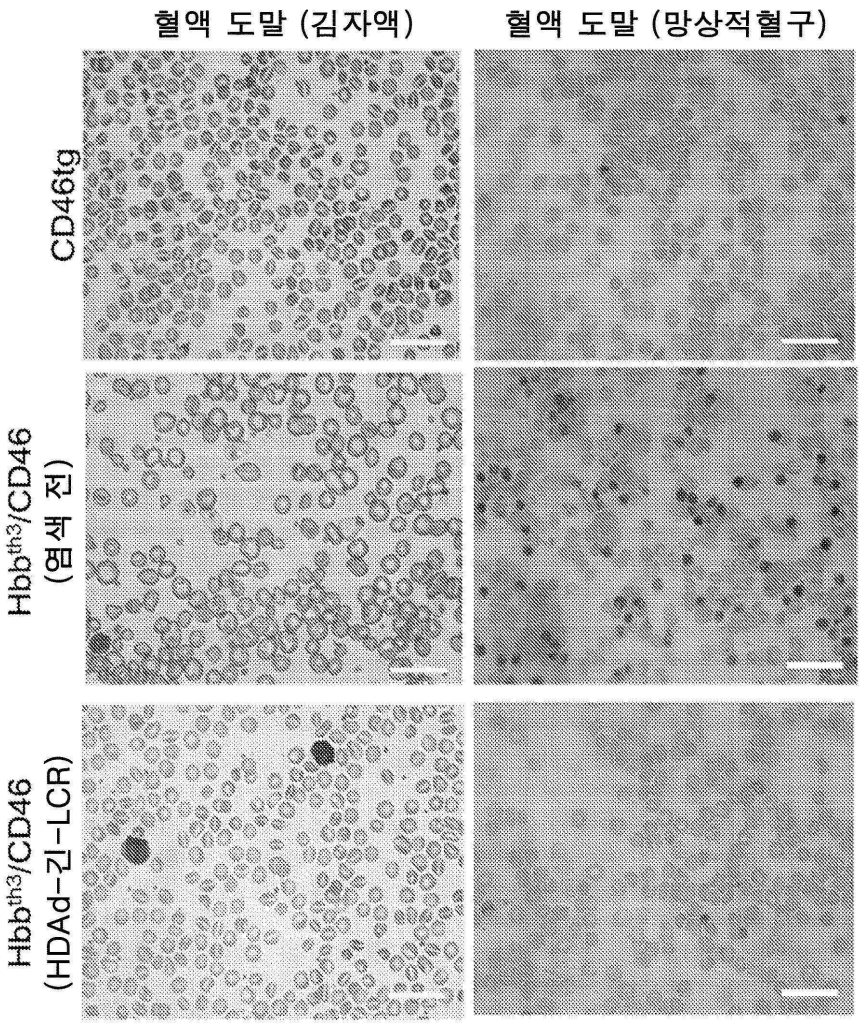
도면36b


혈액 도말 (김자액)
혈액 도말 (망상적혈구)
CD46tg
Hbbth3/CD46 (염색 전)
Hbbth3/CD46 (HDAd-긴-LCR, 18 주차)
Hbbth3/CD46 (HDAd-긴-LCR, 21 주차)

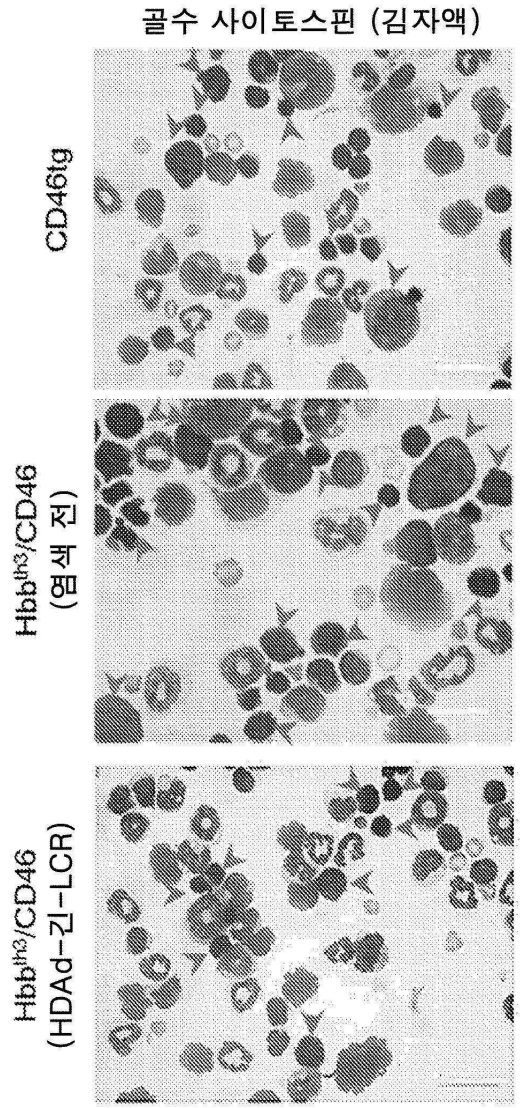
도면36c


골수 사이토스핀 (김자액)
CD46tg
Hbbth3/CD46
(염색 전)
Hbbth3/CD46
(HDAd-긴-LCR, 21 주차)

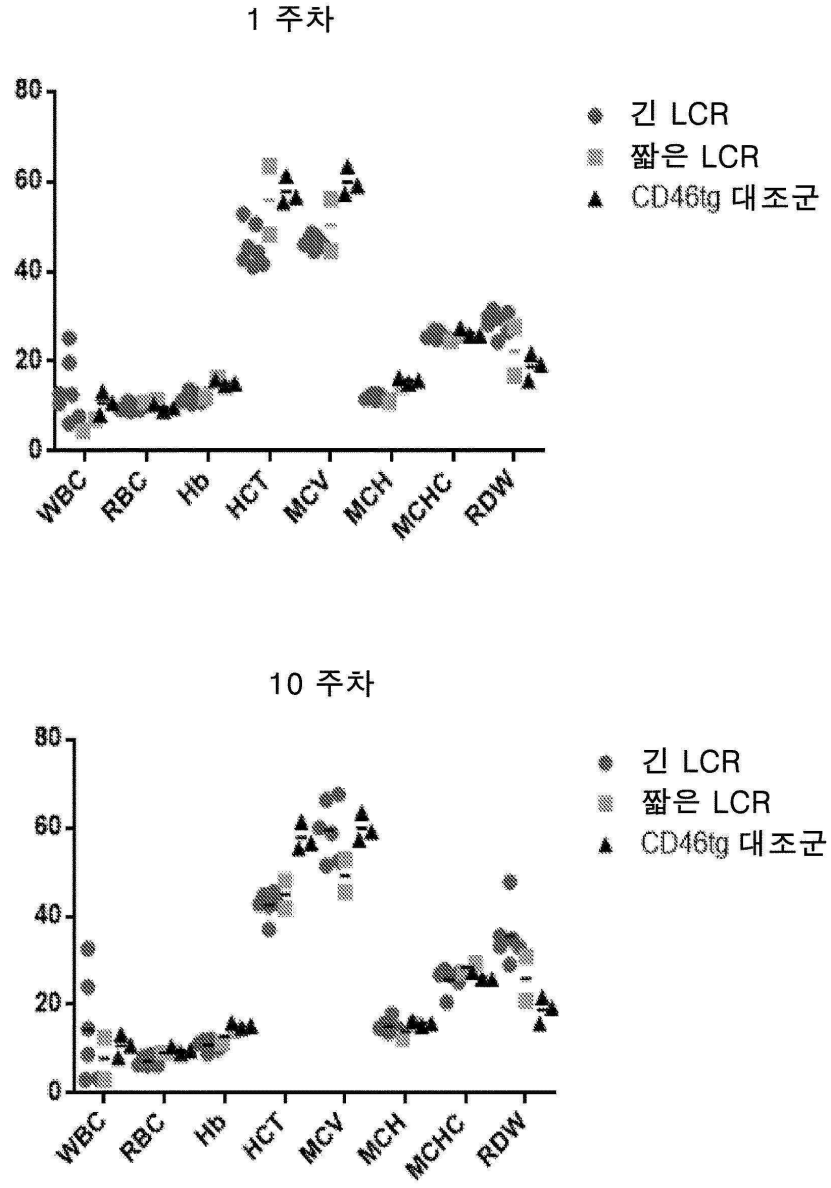
도면37a


혈액 도말 (김자액)
혈액 도말 (망상적혈구)
CD46tg
Hbb^th3/CD46 (염색 전)
Hbb^th3/CD46 (HDAd-긴-LCR)

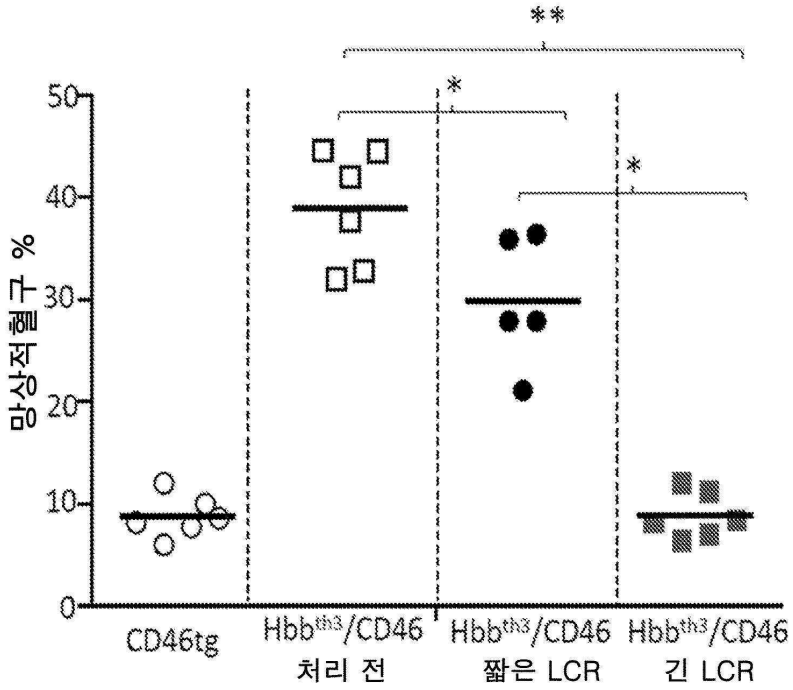
도면37b



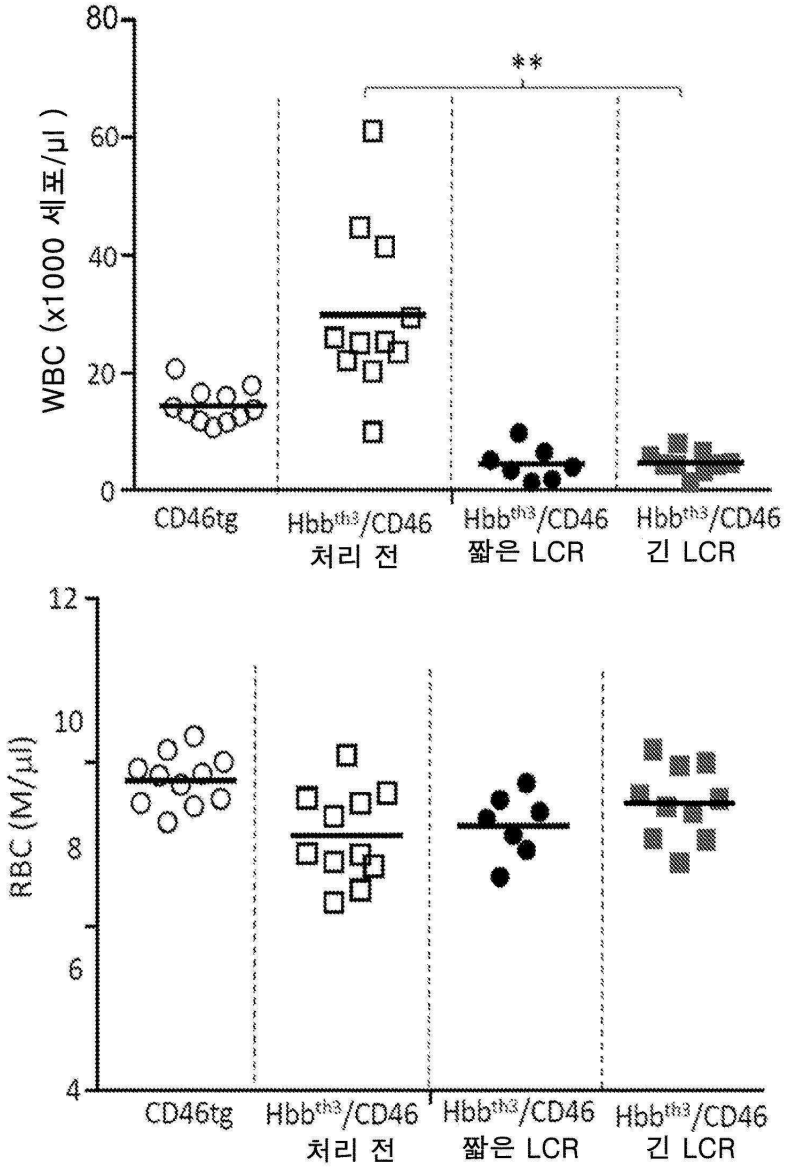
도면38


1 주차
80
60
40
20
0
WBC
RBC
Hb
HCT
MCV
MCH
MCHC
RDW
긴 LCR
짧은 LCR
CD46tg 대조군
10 주차
80
60
40
20
0
WBC
RBC
Hb
HCT
MCV
MCH
MCHC
RDW
긴 LCR
짧은 LCR
CD46tg 대조군

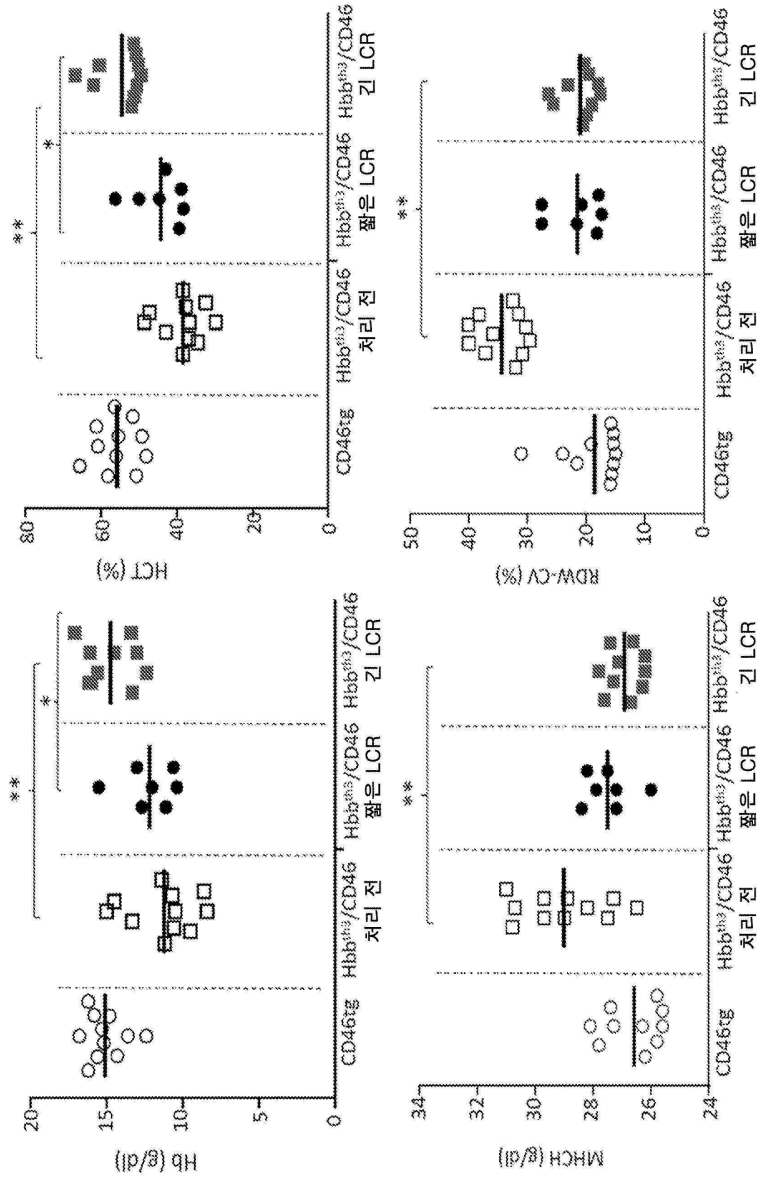
도면39a


**
*
*
50
40
30
20
10
0
망상적혈구 %
CD46tg
Hbb^th3/CD46
처리 전
Hbb^th3/CD46
짧은 LCR
Hbb^th3/CD46
긴 LCR

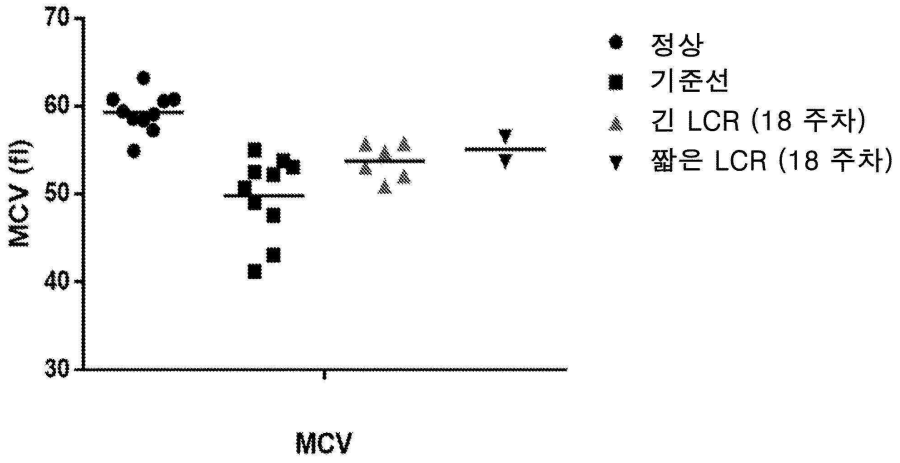
도면39ba


80
60
40
20
0
WBC (x1000 세포/μl )
**
CD46tg
Hbb^th3/CD46
처리 전
Hbb^th3/CD46
짧은 LCR
Hbb^th3/CD46
긴 LCR
12
10
8
6
4
RBC (M/μl)
CD46tg
Hbb^th3/CD46
처리 전
Hbb^th3/CD46
짧은 LCR
Hbb^th3/CD46
긴 LCR

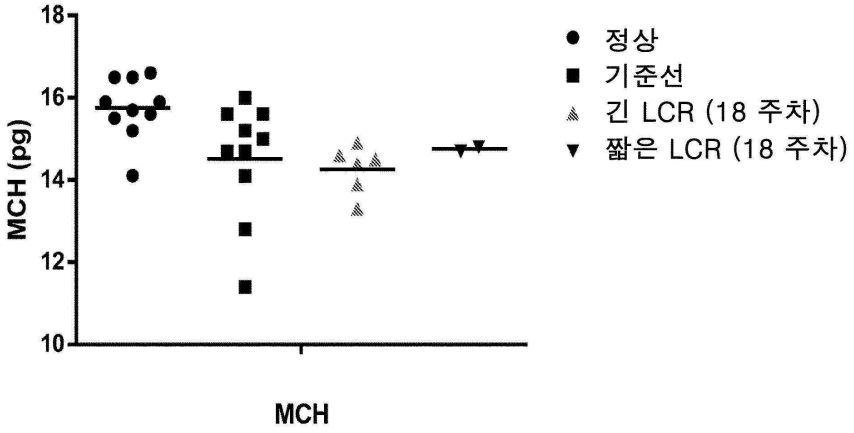
도면39bb


Hb (g/dl)
HCT (%)
RDW-CV (%)
MHCH (g/dl)
CD46tg
Hbb^th3/CD46 처리 전
Hbb^th3/CD46 짧은 LCR
Hbb^th3/CD46 긴 LCR
**
*

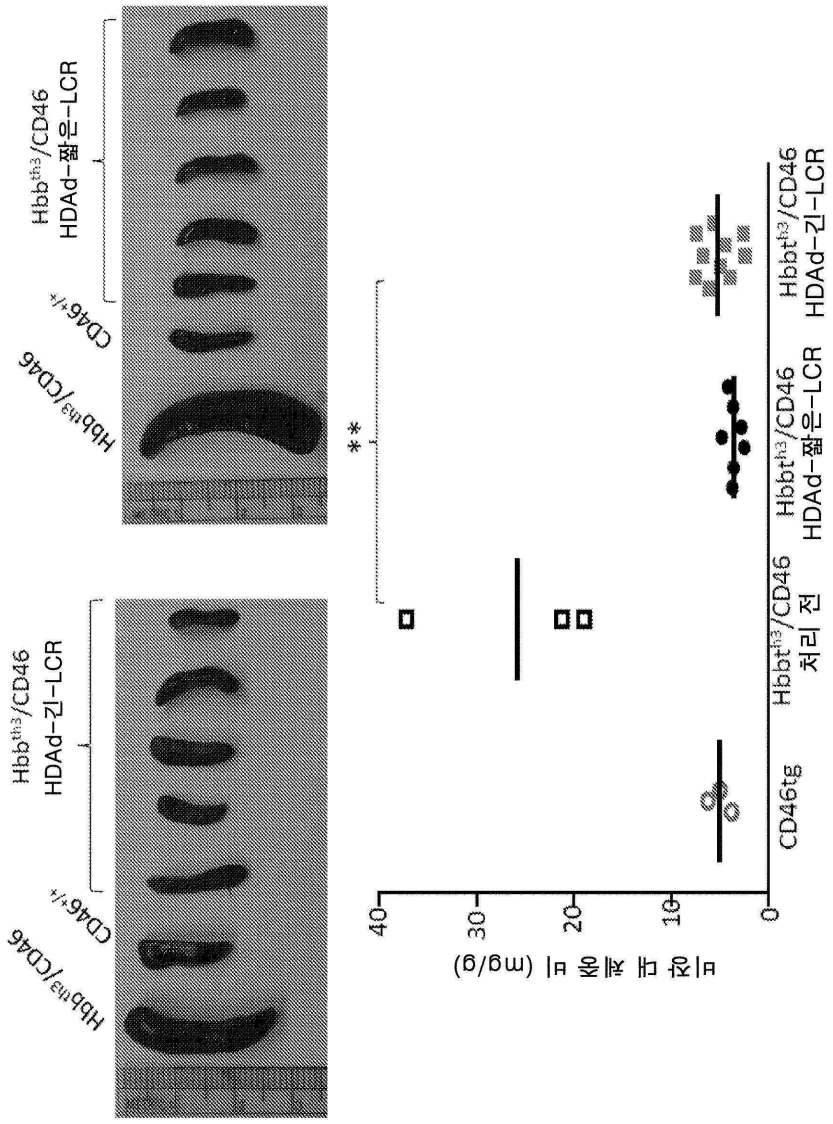
도면39bc


70
60
50
40
30
MCV (fl)
MCV
정상
기준선
긴 LCR (18 주차)
짧은 LCR (18 주차)

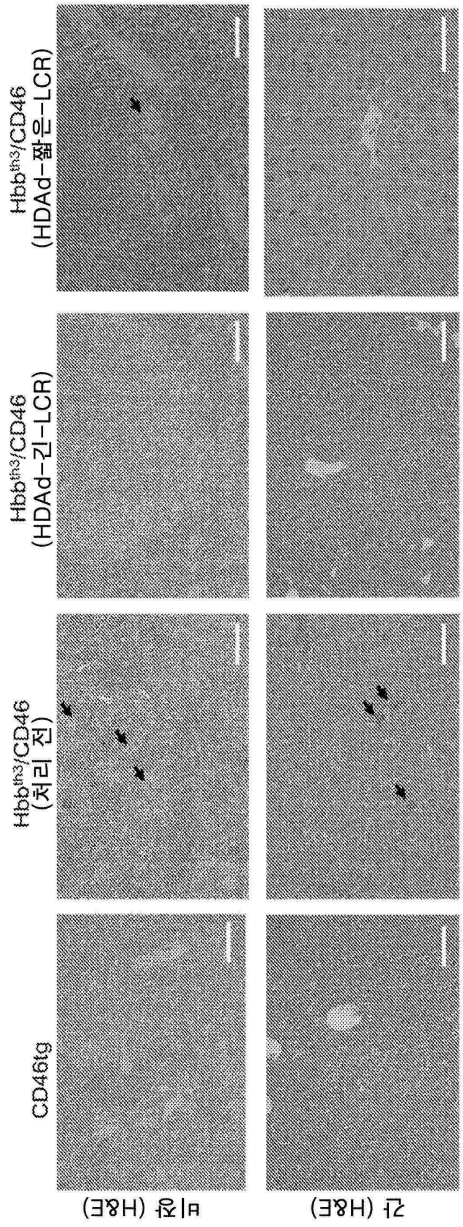
MCH


18
16
14
12
10
MCH (pg)
MCH
정상
기준선
긴 LCR (18 주차)
짧은 LCR (18 주차)

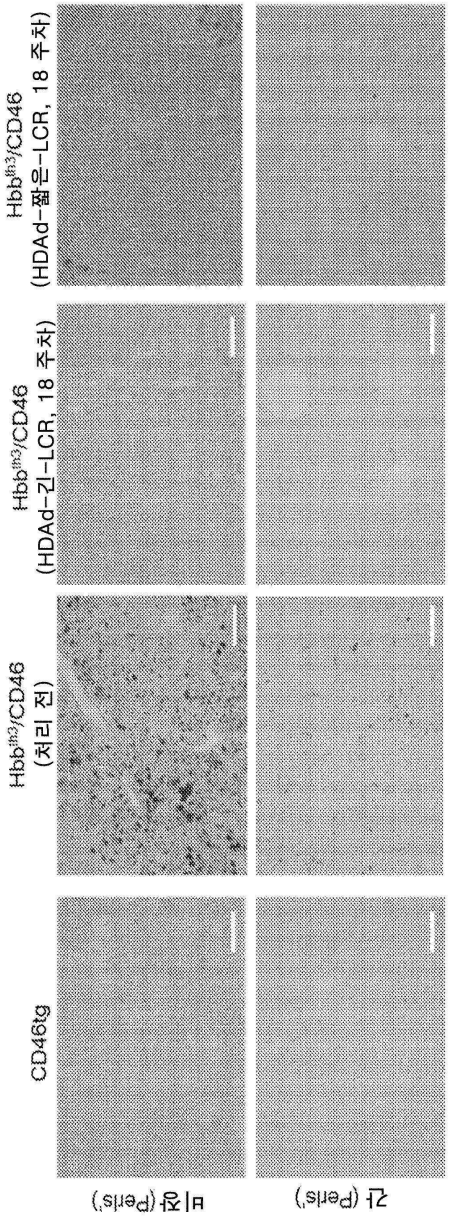
도면40a


Hbb^th3/CD46
HDAd-짧은-LCR
CD46+/+
Hbb^th3/CD46
Hbb^th3/CD46
HDAd-긴-LCR
CD46+/+
Hbb^th3/CD46
**
비장 내 전염 바이러스 (mg/g)
0
10
20
30
40
CD46tg
Hbbt^h3/CD46
처리 전
Hbbt^h3/CD46
HDAd-짧은-LCR
Hbbt^h3/CD46
HDAd-긴-LCR

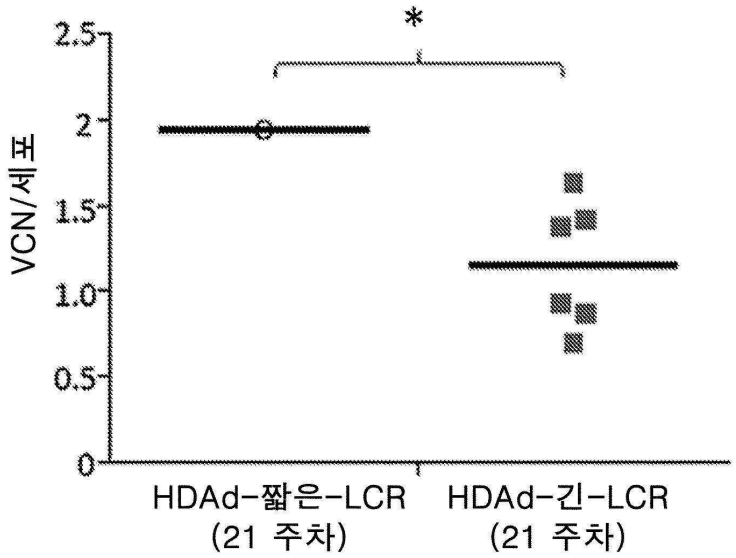
도면40b


CD46tg
Hbb^th3/CD46 (처리 전)
Hbb^th3/CD46 (HDAd-긴-LCR)
Hbb^th3/CD46 (HDAd-짧은-LCR)
비장 (H&E)
간 (H&E)

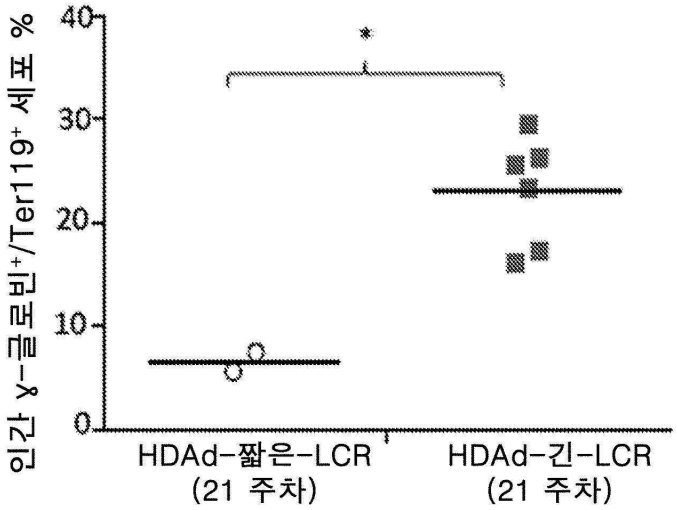
도면41


CD46tg
Hbb^th3/CD46 (처리 전)
Hbb^th3/CD46 (HDAd-긴-LCR, 18 주차)
Hbb^th3/CD46 (HDAd-짧은-LCR, 18 주차)
비장 (Peris')
간 (Peris')

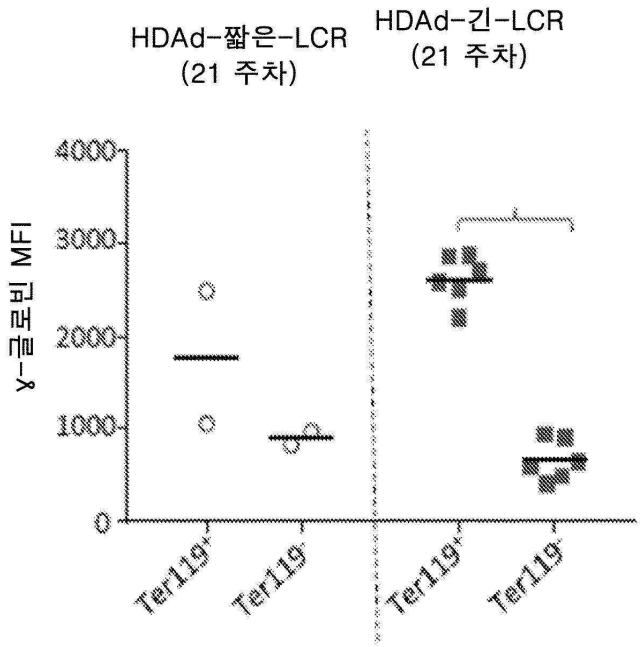
도면42a


*
2.5
2
1.5
1.0
0.5
0
VCN/세포
HDAd-짧은-LCR
(21 주차)
HDAd-긴-LCR
(21 주차)

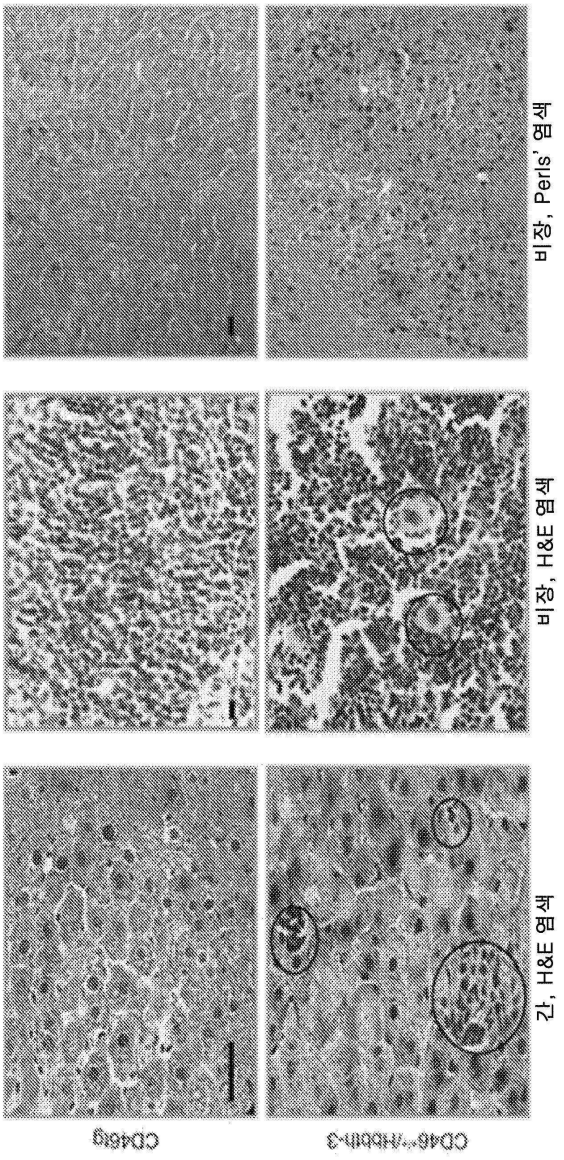
도면42b


*
40
30
20
10
0
인간 γ-글로빈+/Ter119+ 세포 %
HDAd-짧은-LCR
(21 주차)
HDAd-긴-LCR
(21 주차)

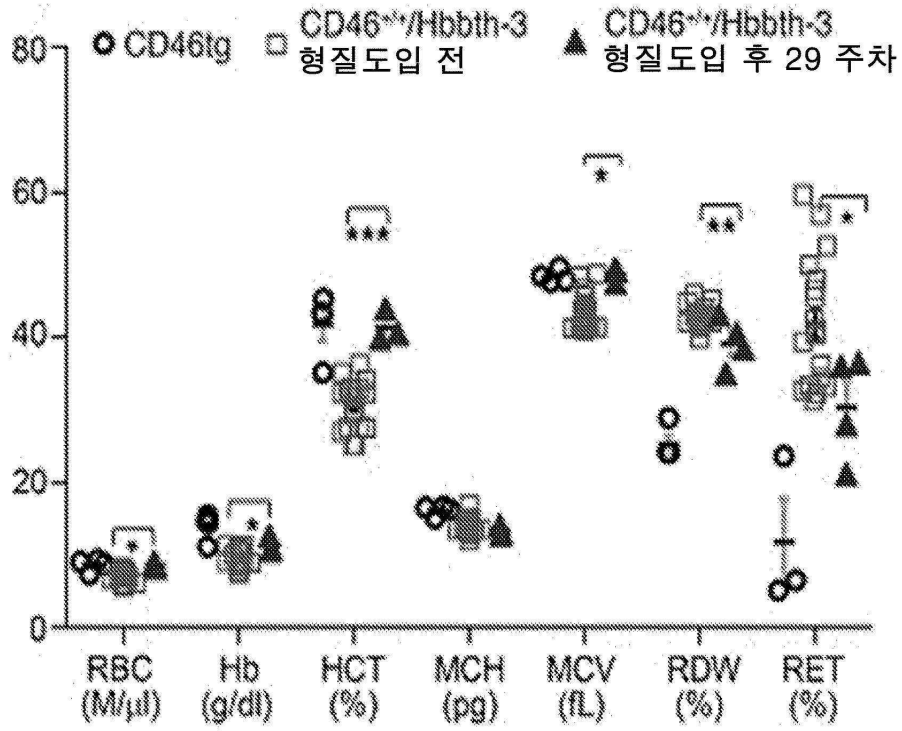
도면42c


HDAd-짧은-LCR
(21 주차)
HDAd-긴-LCR
(21 주차)
γ-글로빈 MFI
4000
3000
2000
1000
0
Ter119+
Ter119-
Ter119+
Ter119-

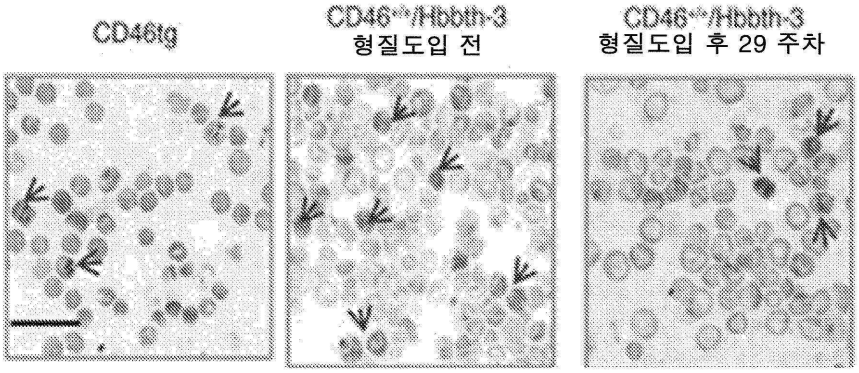
도면43



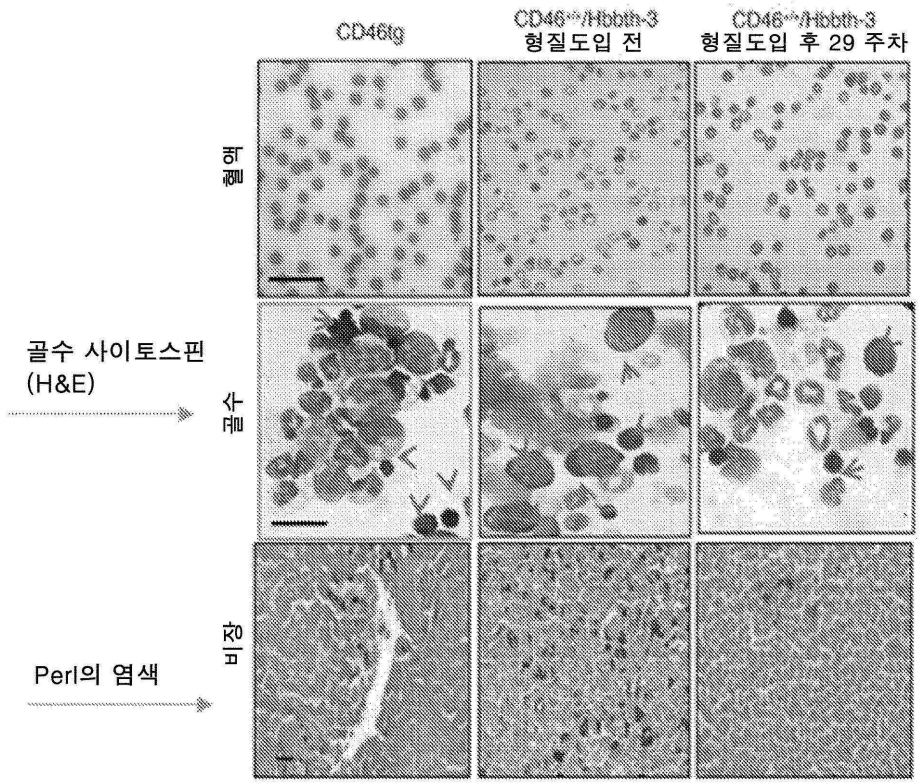
도면44a


CD46tg
CD46tm/Hbbth-3
형질도입 전
CD46tm/Hbbth-3
형질도입 후 29 주차
80
60
40
20
0
RBC (M/μl)
Hb (g/dl)
HCT (%)
MCH (pg)
MCV (fL)
RDW (%)
RET (%)

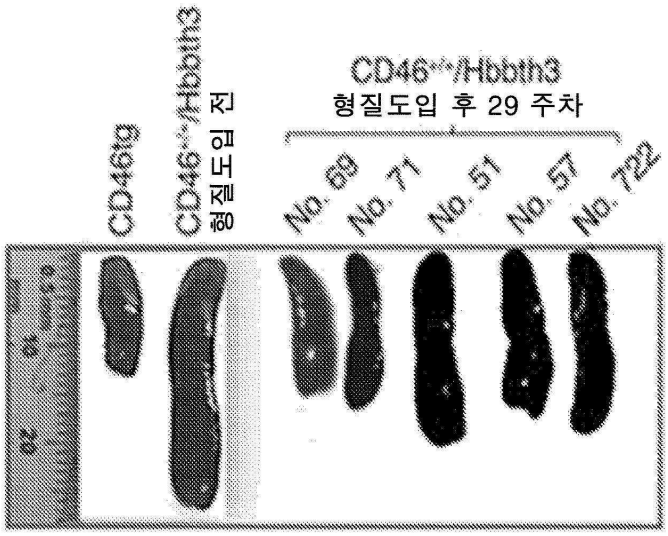
도면44b


CD46tg
CD46tm/Hbbth-3
형질도입 전
CD46tm/Hbbth-3
형질도입 후 29 주차

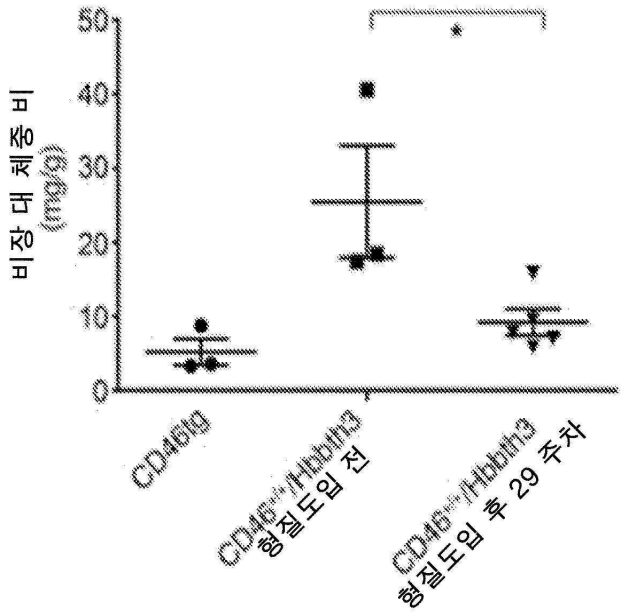
도면44c


CD46tg
CD46-/-/Hbbth-3
형질도입 전
CD46-/-/Hbbth-3
형질도입 후 29 주차
혈액
골수
비장
골수 사이토스핀
(H&E)
Perl의 염색

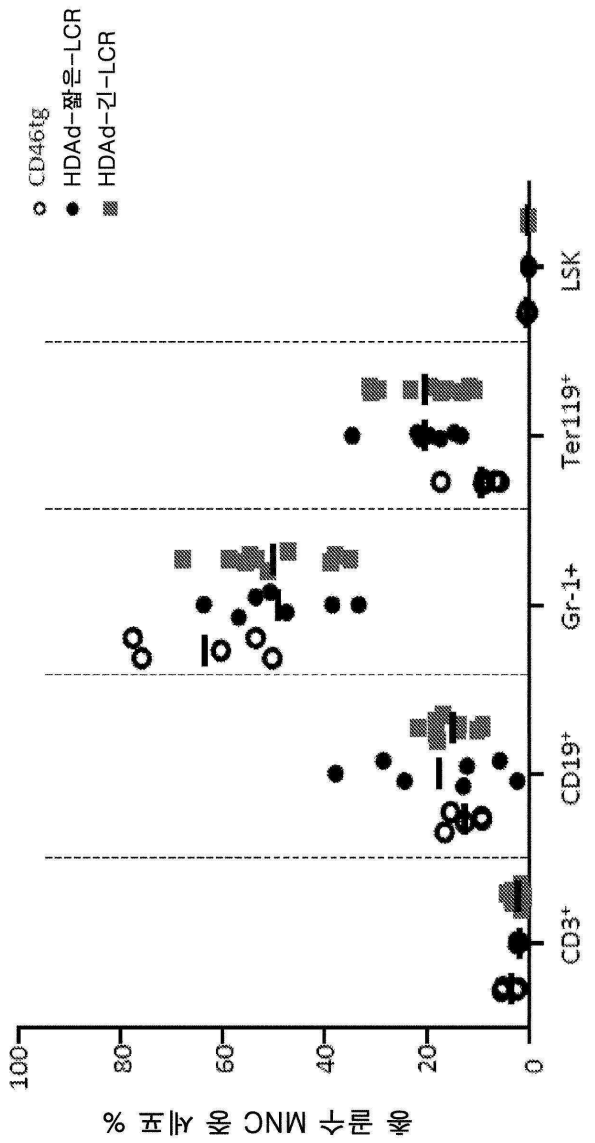
도면44d


CD46tg
CD46-/-/Hbbth3
형질도입 전
CD46-/-/Hbbth3
형질도입 후 29 주차
No. 69
No. 71
No. 51
No. 57
No. 722

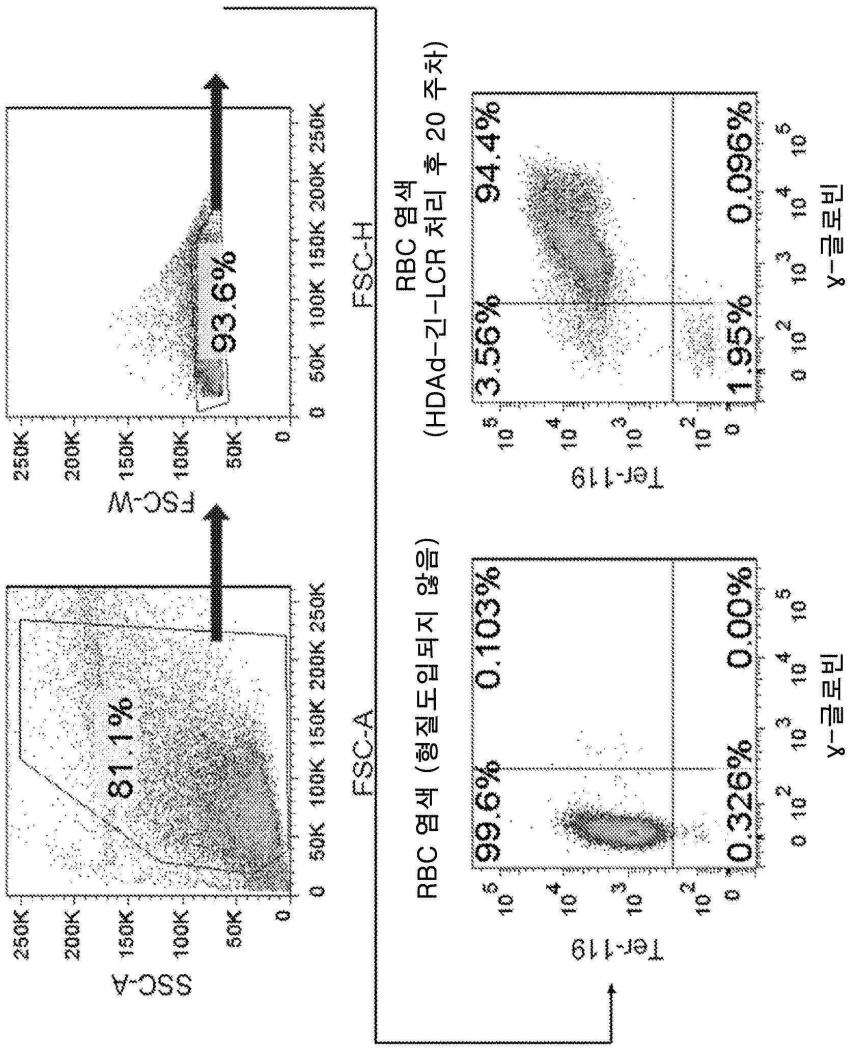
도면44e


비장 대 체중 비
(mg/g)
50
40
30
20
10
0
CD46tg
CD46hi/Hbbth3
형질도입 전
CD46hi/Hbbth3
형질도입 후 29 주차

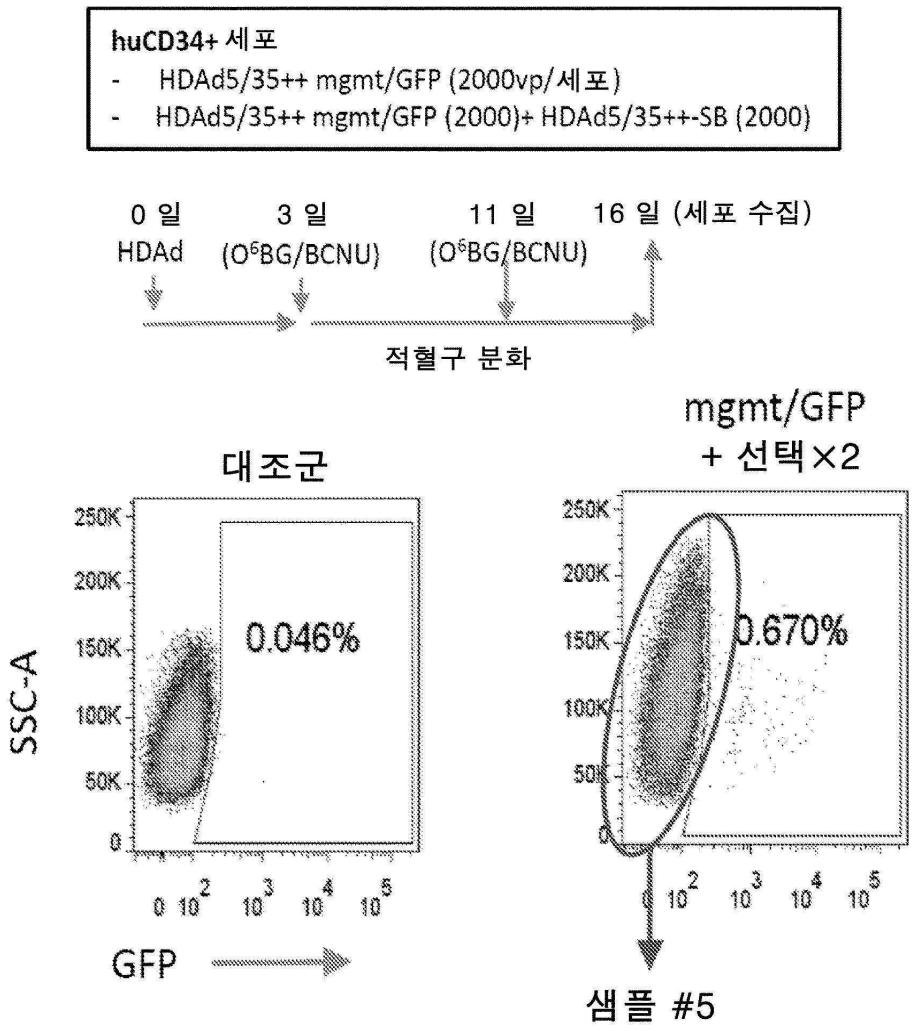
도면45


CD46tg
HDAd-짧은-LCR
HDAd-긴-LCR
총 골수 MNC 중 세포 %
0
20
40
60
80
100
CD3+
CD19+
Gr-1+
Ter119+
LSK

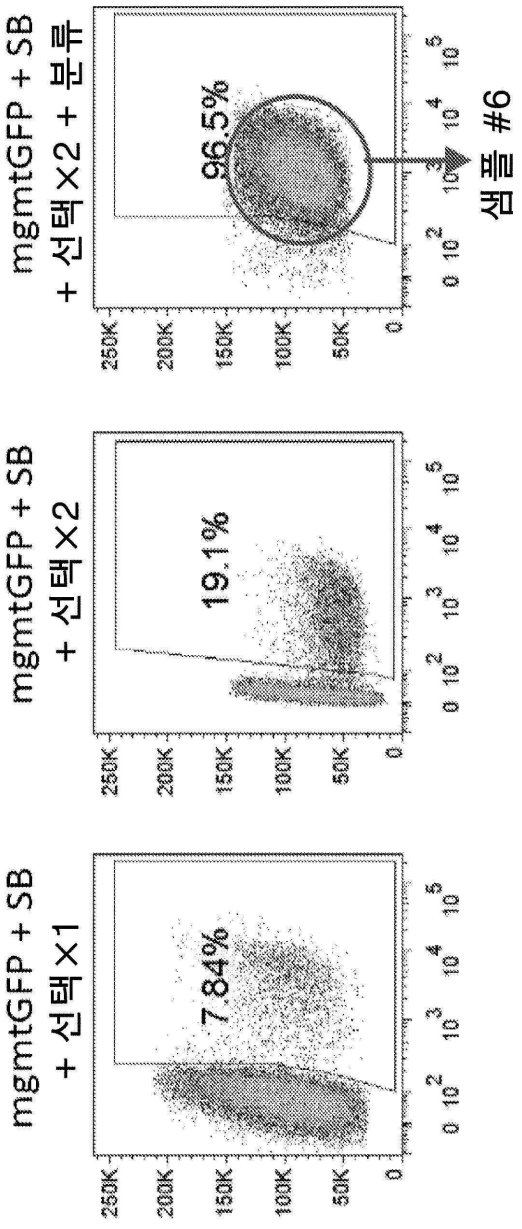
도면46


SSC-A
FSC-A
81.1%
FSC-W
FSC-H
93.6%
RBC 염색 (헝질도입되지 않음)
Ter-119
γ-글로빈
99.6%
0.103%
0.326%
0.00%
RBC 염색
(HDAd-긴-LCR 처리 후 20 주차)
3.56%
94.4%
1.95%
0.096%

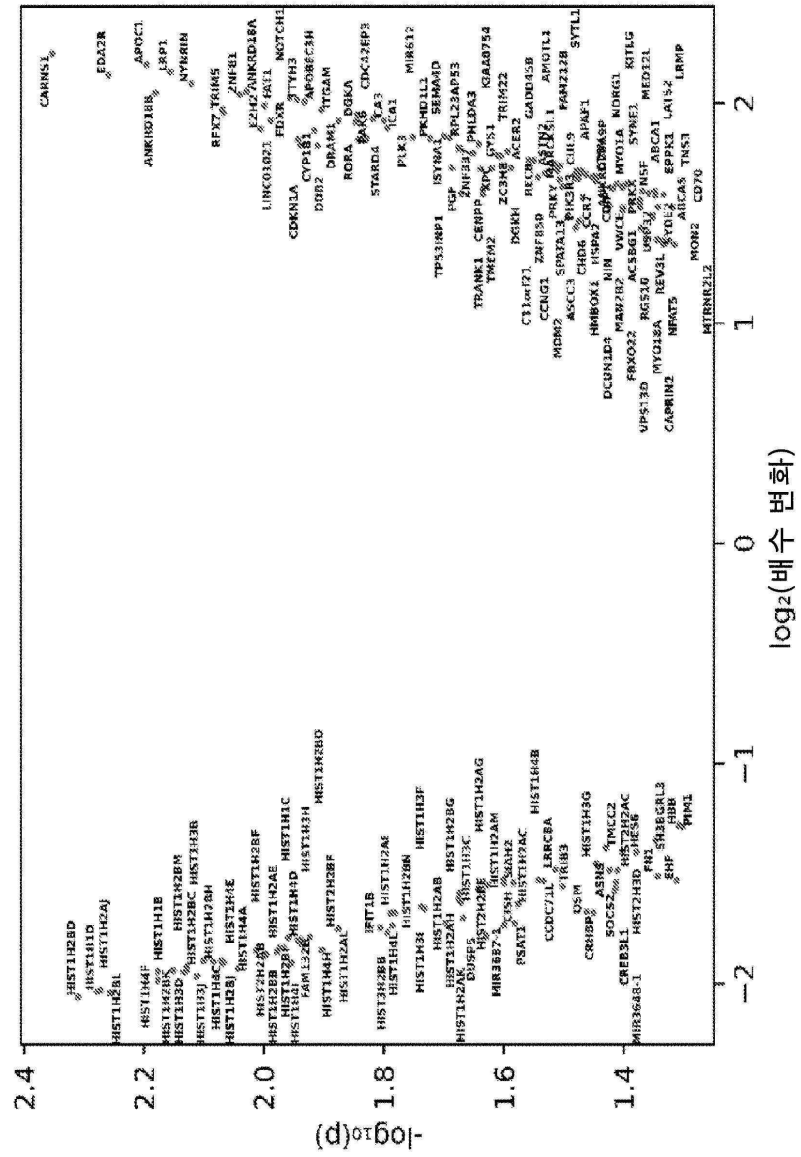
도면47aa


huCD34+ 세포
- HDAd5/35++ mgmt/GFP (2000vp/세포)
- HDAd5/35++ mgmt/GFP (2000)+ HDAd5/35++-SB (2000)
0 일
HDAd
3 일
(O6BG/BCNU)
11 일
(O6BG/BCNU)
16 일 (세포 수집)
적혈구 분화
대조군
mgmt/GFP
+ 선택×2
SSC-A
250K
200K
150K
100K
50K
0
0.046%
0.670%
0 10^2 10^3 10^4 10^5
GFP
샘플 #5

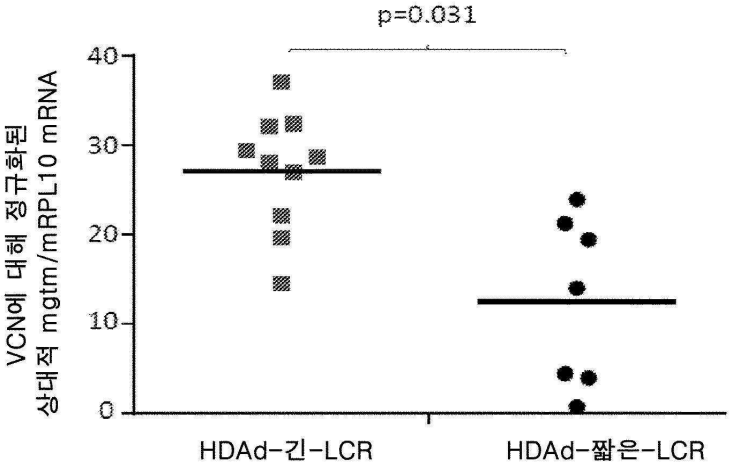
도면47ab


mgmtGFP + SB
+ 선택×1
7.84%
mgmtGFP + SB
+ 선택×2
19.1%
mgmtGFP + SB
+ 선택×2 + 분류
96.5%
샘플 #6

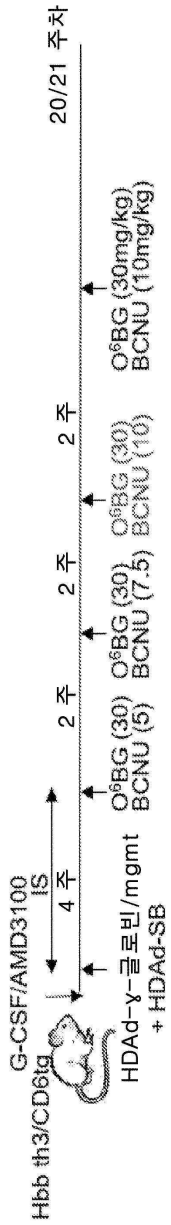
도면47b


-log10(p)
log2(배수 변화)

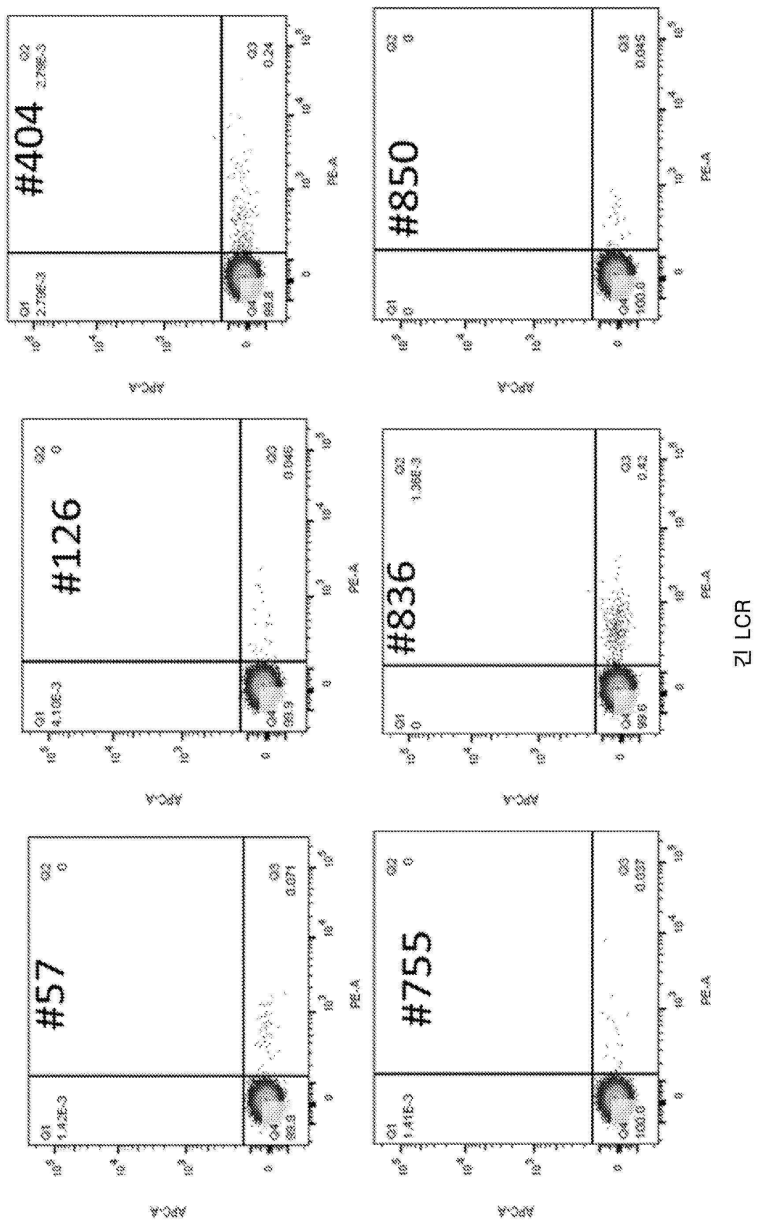
도면48


p=0.031
VCN에 대해 정규화된
상대적 mgtm/mRPL10 mRNA
HDAd-긴-LCR
HDAd-짧은-LCR

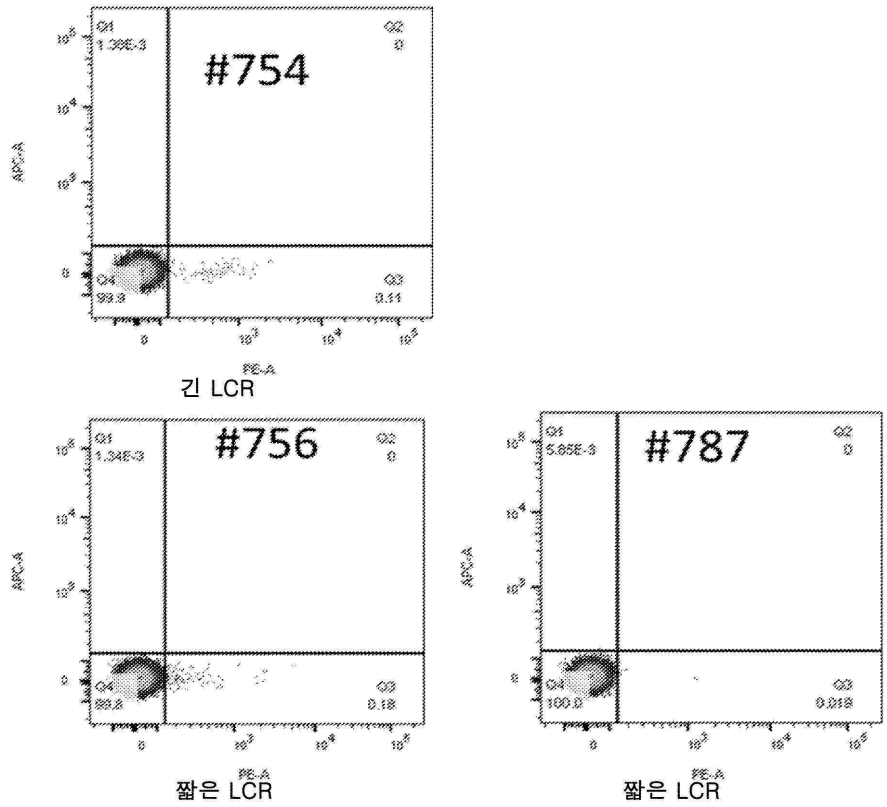
도면49


Hbb th3/CD6tg
G-CSF/AMD3100
IS
4 주
HDAd-γ-글로빈/mgmt
+ HDAd-SB
O6BG (30)
BCNU (5)
2 주
O6BG (30)
BCNU (7.5)
2 주
O6BG (30)
BCNU (10)
2 주
O6BG (30mg/kg)
BCNU (10mg/kg)
20/21 주차

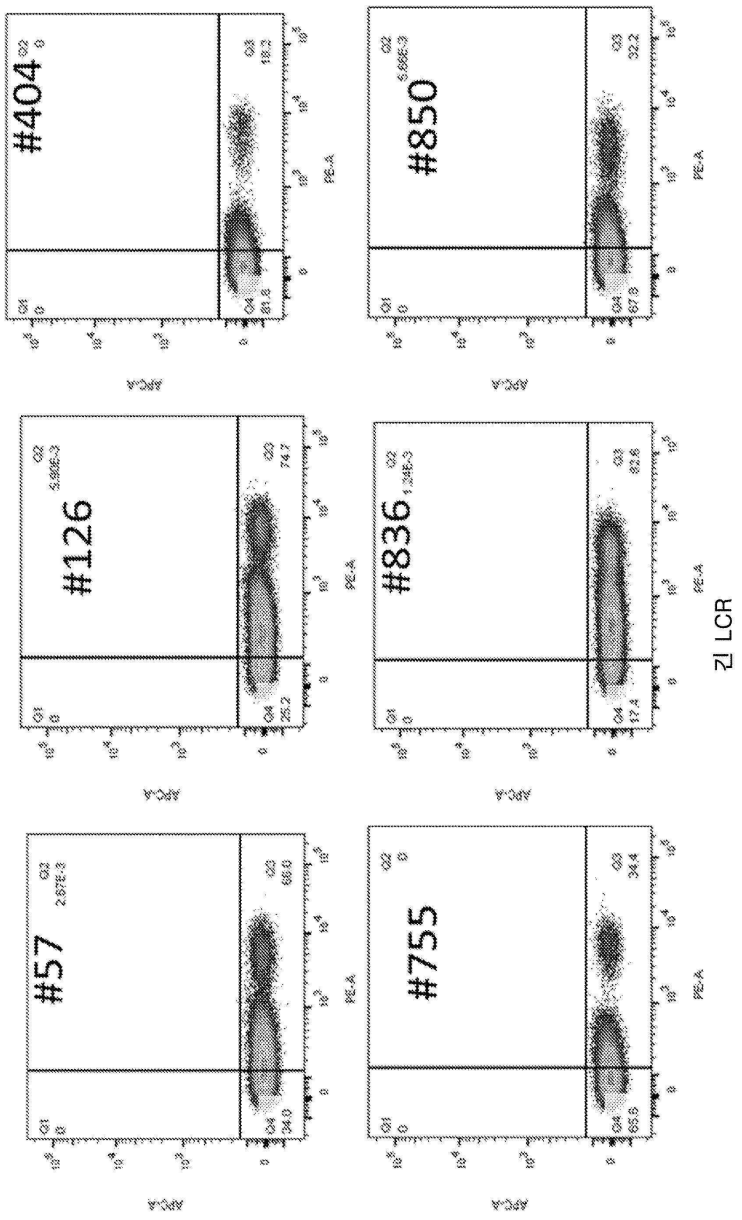
도면50a



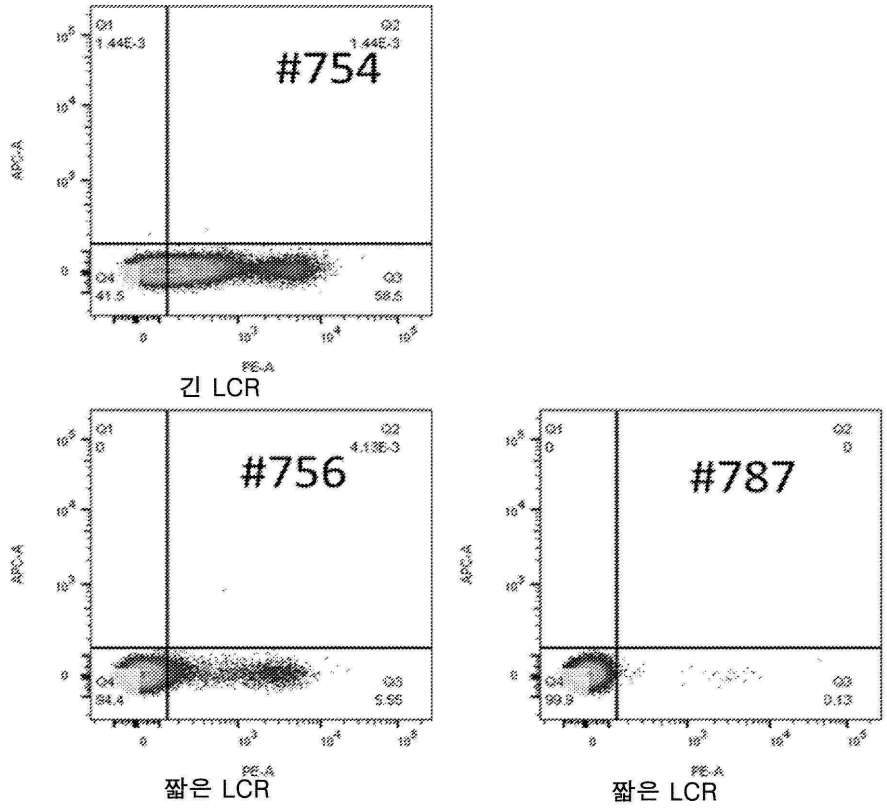
도면50b



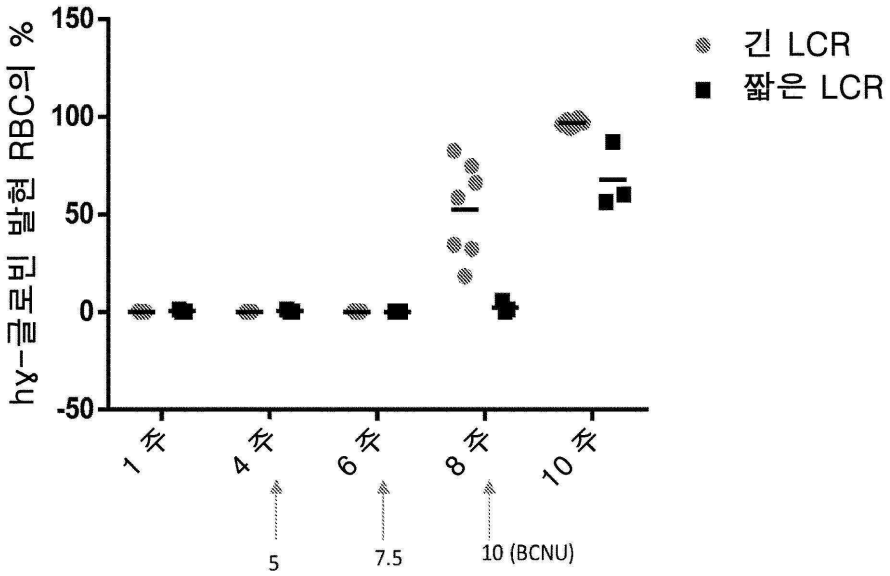
도면51a



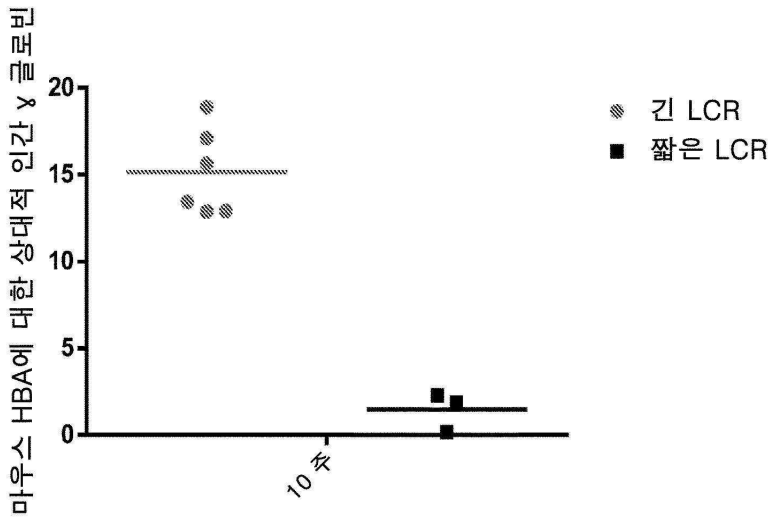
도면51b


#754
긴 LCR
#756
짧은 LCR
#787
짧은 LCR

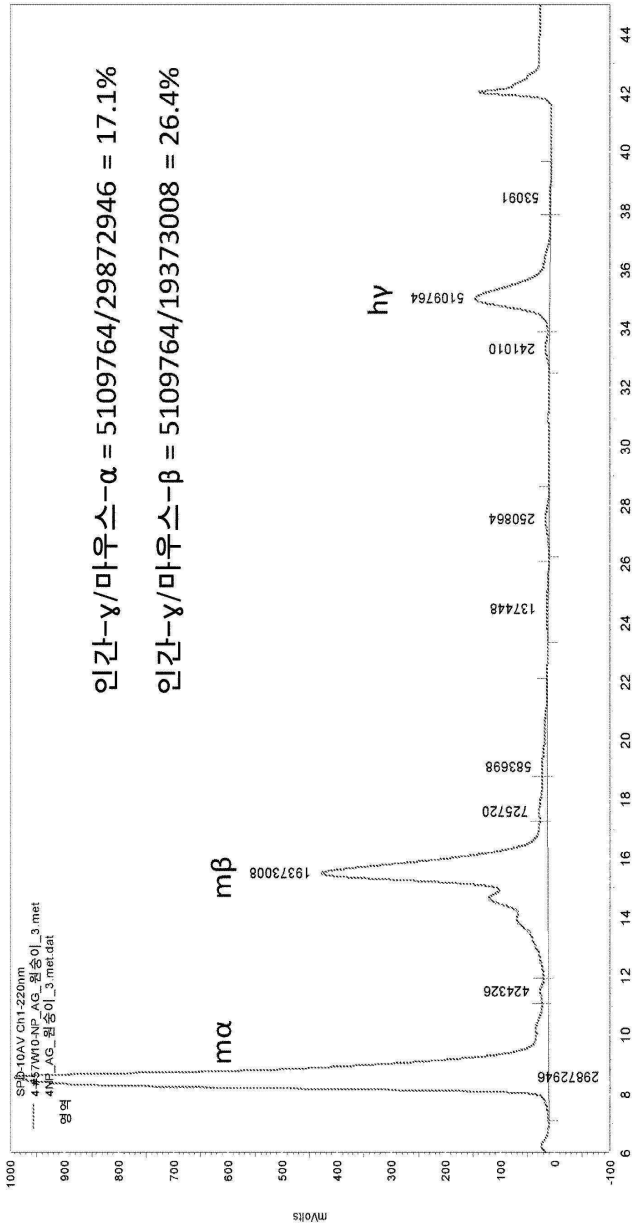
도면52


긴 LCR
짧은 LCR
hγ-글로빈 형광 RBC의 %
150
100
50
0
-50
1 주
4 주
6 주
8 주
10 주
5
7.5
10 (BCNU)

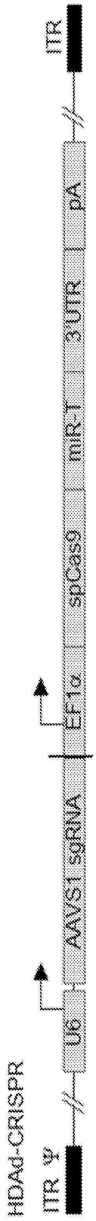
도면53



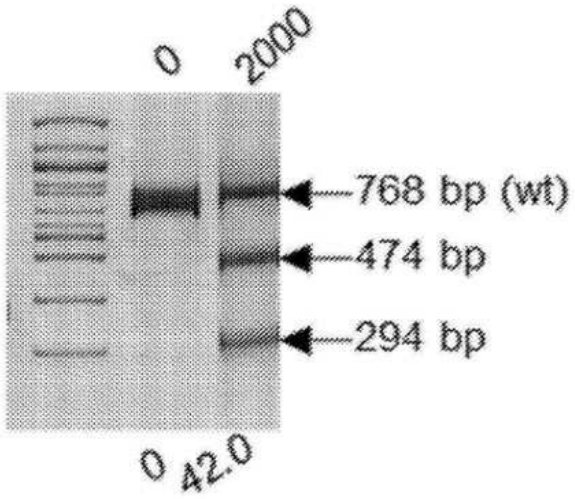
도면54


인간-γ/마우스-α = 5109764/29872946 = 17.1%
인간-γ/마우스-β = 5109764/19373008 = 26.4%
mα
mβ
hγ
29872946
424326
19373008
725720
583698
137448
250864
241010
5109764
53091
mVolts

도면55a



도면55b

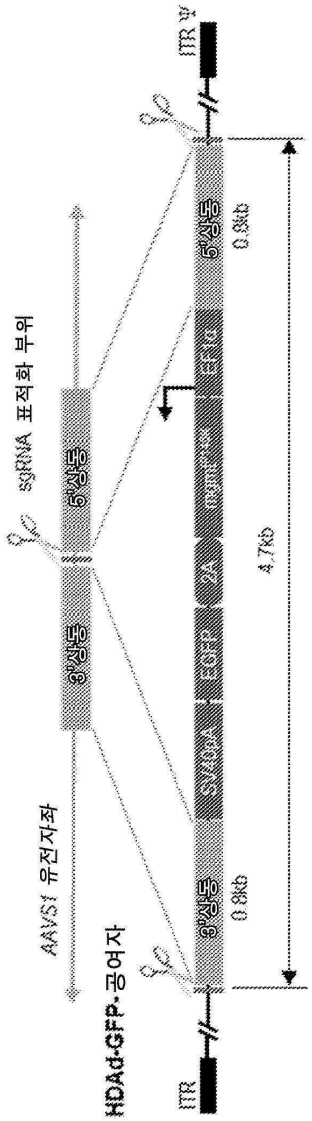



도면55c

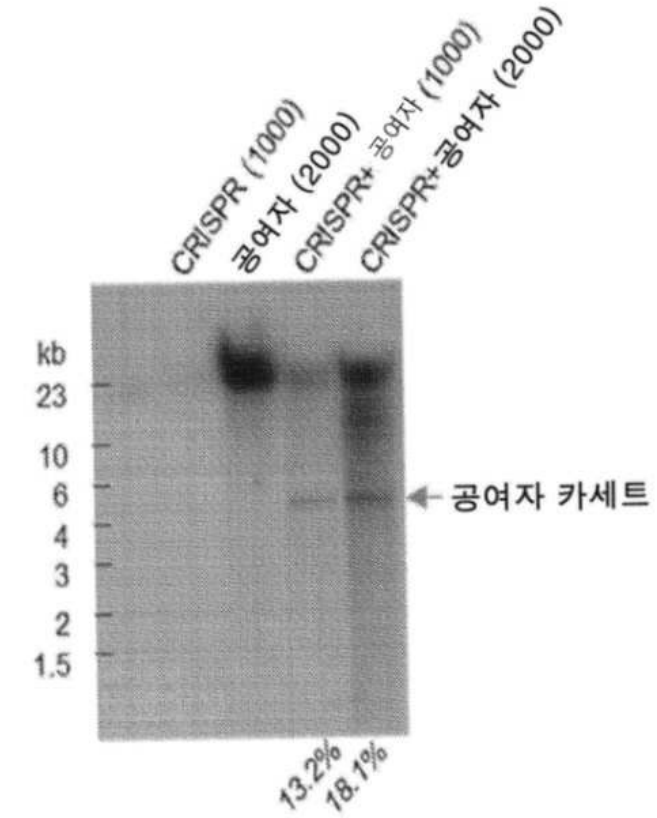
```
                                                ↓
TGTCCCCTCCACCCCACAGTGGGGCCACTAGGGACAGGATTGGTGACAGAAAAGCCCCATCCTT
TGTCCCCTCCACCCCACAGTG--------------------GGTGACAGAAAAGCCCCATCCTT
TGTCCCCTCCACCCCACA------------------------GATTGGTGACAGAAAAGCCCCATCCTT
TGTCCCCTCCACCCCACAGTGGGGCCACTAGG----------GACAGAAAAGCCCCATCCTT
TGTCCCCTCCACCCCACAGTGGGGCCACTA------------GGTGACAGAAAAGCCCCATCCTT
TGTCCCCTCCACCCCACAGTGGGGCCACTAG-----------GGTGACAGAAAAGCCCCATCCTT
TGTCCCCTCCACCCCACAGTGGGG-----------CAGGATTGGTGACAGAAAAGCCCCATCCTT
TGTCCCCTCCACCCCACAGTGGGGCCACT------AGGATTGGTGACAGAAAAGCCCCATCCTT
TGTCCCCTCCACCCCACAGTGGGGCCACTAG-----GGATTGGTGACAGAAAAGCCCCATCCTT
TGTCCCCTCCACCCCACAGTGGGGCCACTAGGGACA-GATTGGTGACAGAAAAGCCCCATCCTT
TGTCCCCTCCACCCCACAGTGGGGCCACTAGGGAC-GGATTGGTGACAGAAAAGCCCCATCCTT

TGTCCCCTCCACCCCACAGTGGGGCCACTAGGGACAGAGGATTGGTGACAGAAAAGCCCCATCC
TGTCCCCTCCACCCCACAGTGGGGCCACTAGGGACAGCAGGATTGGTGACAGAAAAGCCCCATC
```

(서열번호: 6-18)

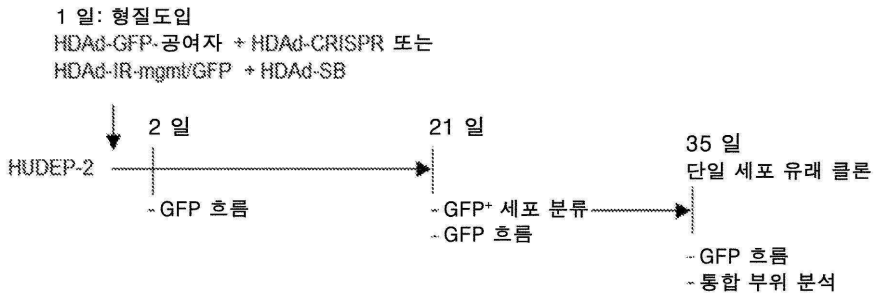
도면55d


ITR
HDAd-GFP-공여자
AAVS1 유전자좌
3'상동
5'상동
sgRNA 표적화 부위
5'상동
ITR
0.8kb
4.7kb
0.8kb

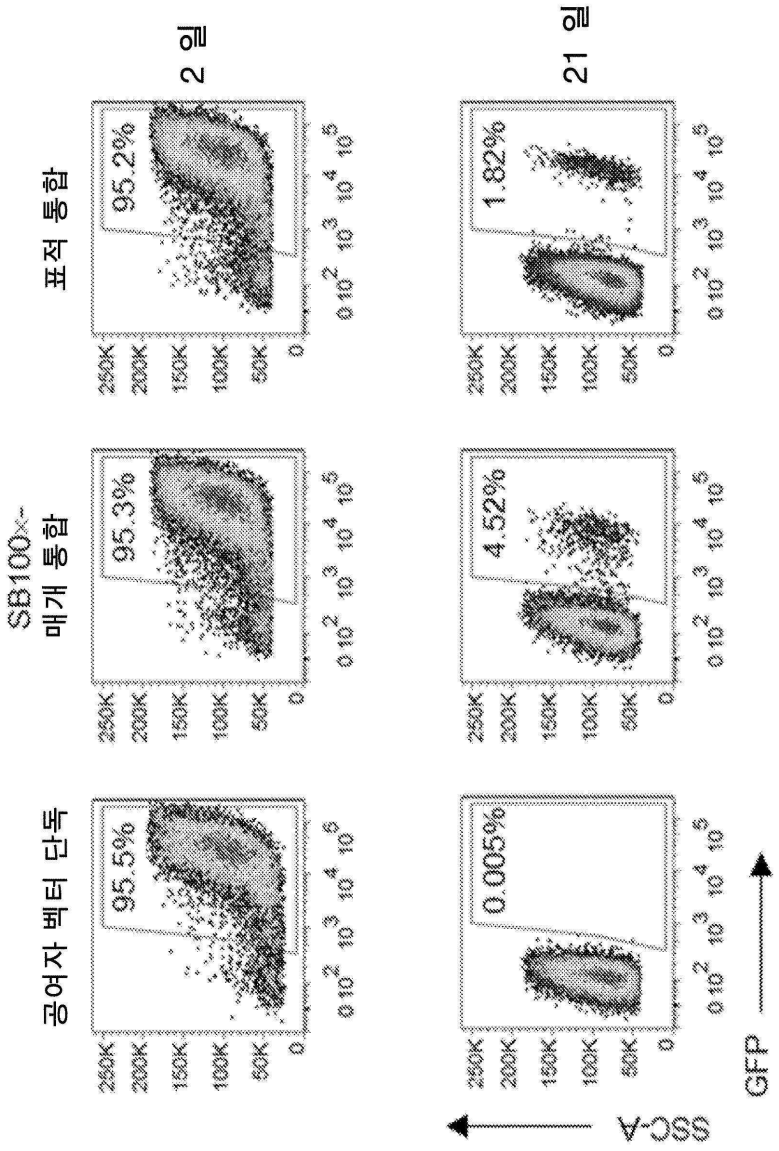
도면55e


CRISPR (1000)
공여자 (2000)
CRISPR+ 공여자 (1000)
CRISPR+공여자 (2000)
kb
23
10
6
4
3
2
1.5
공여자 카세트
13.2%
18.1%

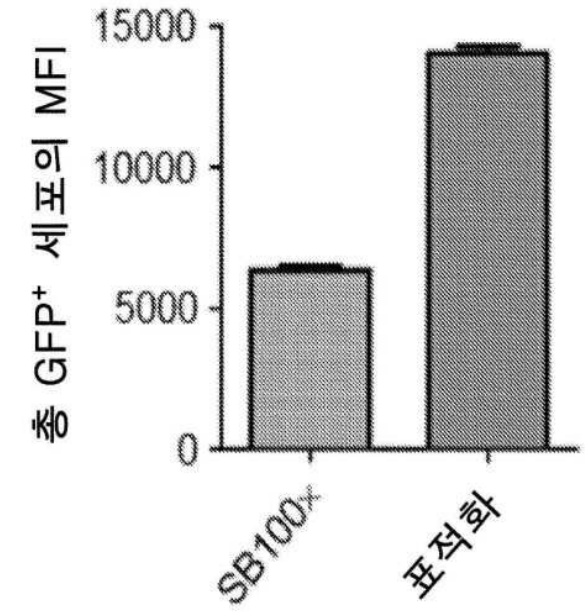
도면56a


1 일: 형질도입
HDAd-GFP-공여자 + HDAd-CRISPR 또는
HDAd-IR-mgmt/GFP + HDAd-SB
HUDEP-2
2 일
- GFP 흐름
21 일
- GFP+ 세포 분류
- GFP 흐름
35 일
단일 세포 유래 클론
- GFP 흐름
- 통합 부위 분석

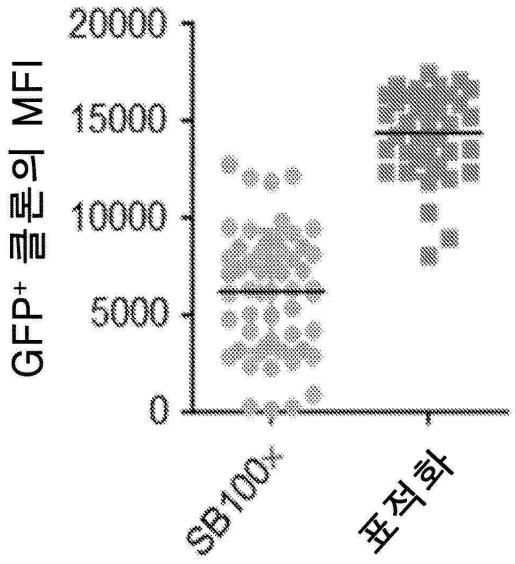
도면56b


2 일
21 일
표적 통합
SB100x-
매개 통합
공여자 벡터 단독
95.2%
95.3%
95.5%
1.82%
4.52%
0.005%
SSC-A
GFP

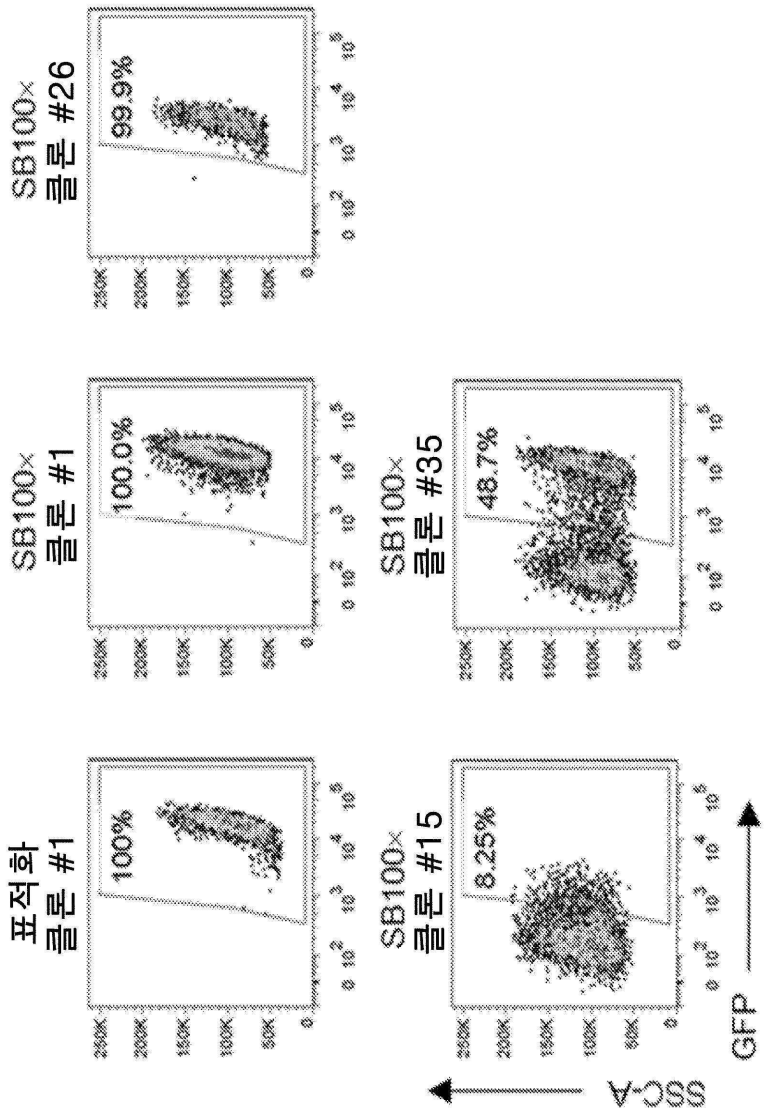
도면56c


15000
10000
5000
0
총 GFP+ 세포의 MFI
SB100x
표적화

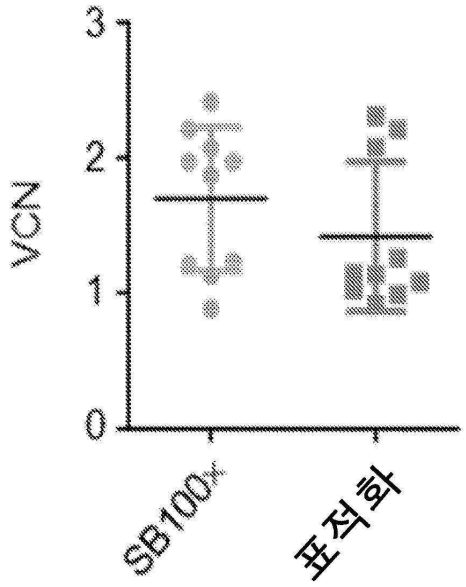
도면56d


20000
15000
10000
5000
0
GFP+ 클론의 MFI
SB100x
표적화

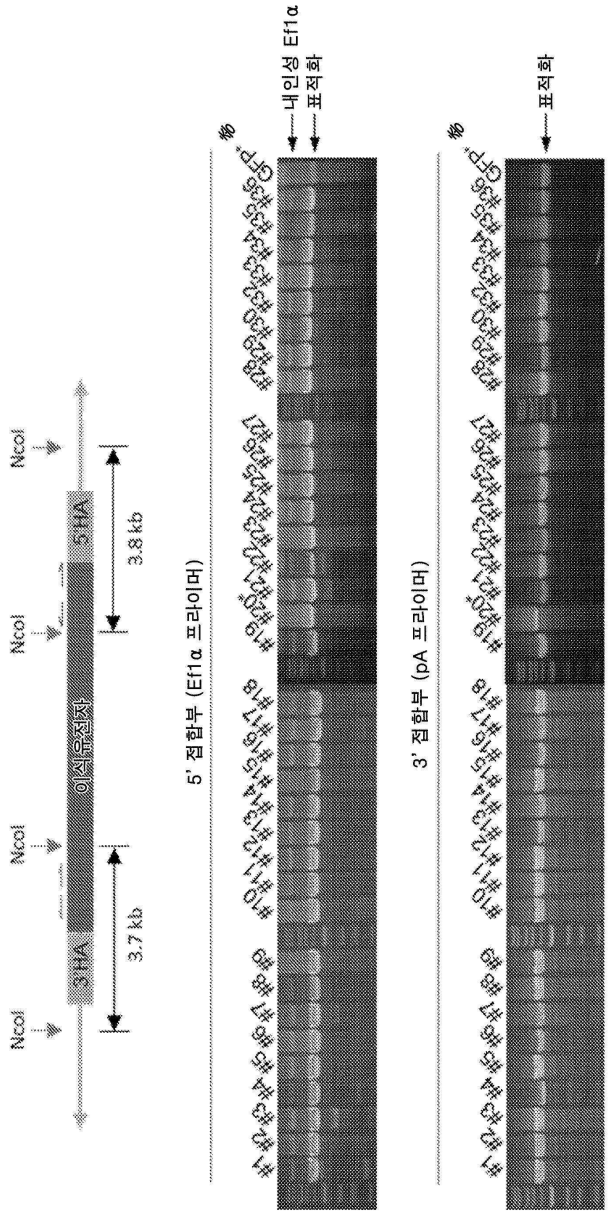
도면56e


SB100×
클론 #26
99.9%
SB100×
클론 #1
100.0%
SB100×
클론 #35
48.7%
표적화
클론 #1
100%
SB100×
클론 #15
8.25%
SSC-A
GFP

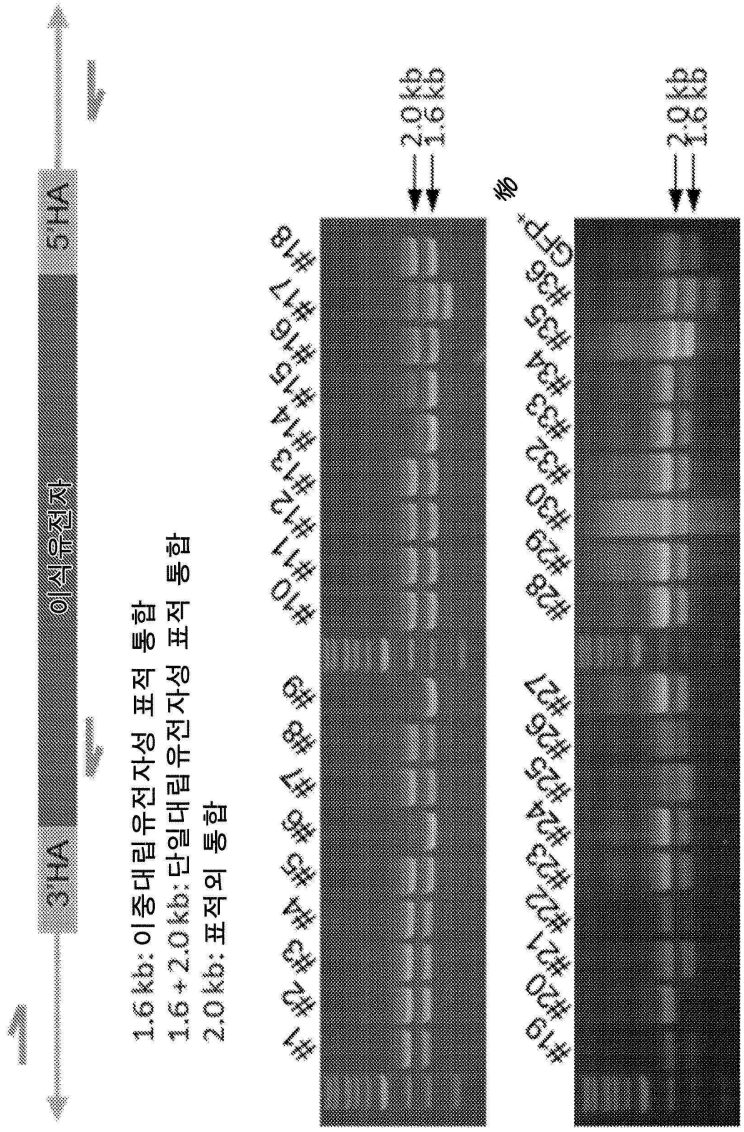
도면56f



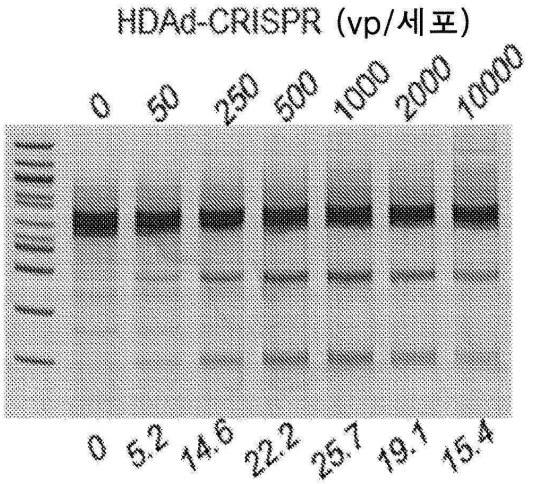
도면57a


NcoI
NcoI
NcoI
NcoI
5HA
3HA
이식유전자
3.8 kb
3.7 kb
5' 접합부 (Ef1α 프라이머)
내인성 Ef1α
표적화
3' 접합부 (pA 프라이머)
표적화

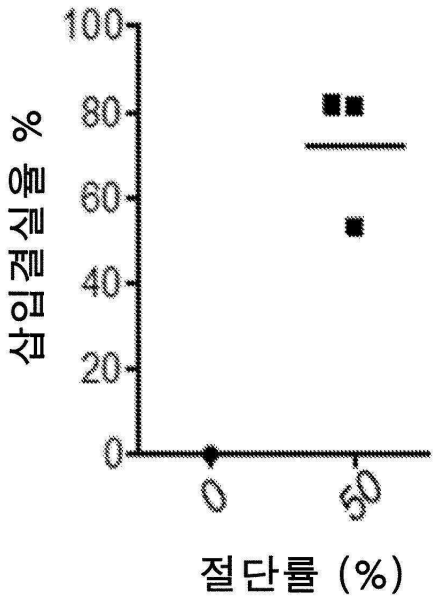
도면57b


5'HA
이식유전자
3'HA
1.6 kb: 이중대립유전자성 표적 통합
1.6 + 2.0 kb: 단일대립유전자성 표적 통합
2.0 kb: 표적외 통합
2.0 kb
1.6 kb
#1 #2 #3 #4 #5 #6 #7 #8 #9 #10 #11 #12 #13 #14 #15 #16 #17 #18
#19 #20 #21 #22 #23 #24 #25 #26 #27 #28 #29 #30 #31 #32 #33 #34 #35 #36
GFP+

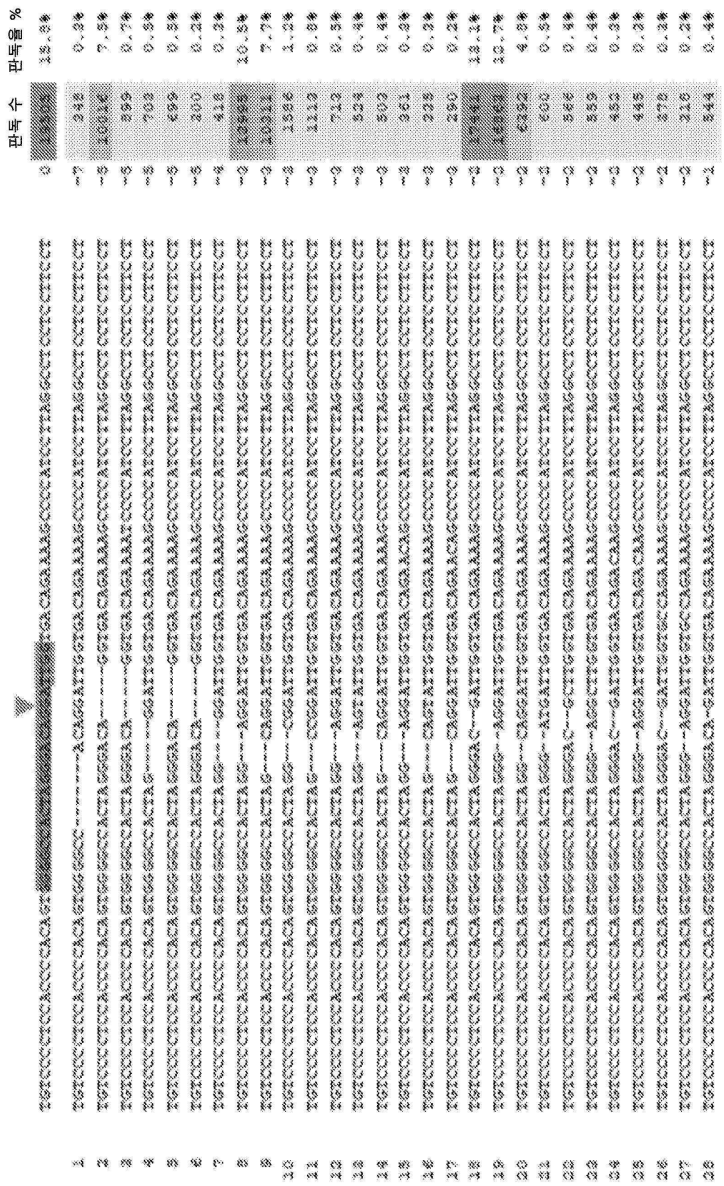
도면58a


HDAd-CRISPR (vp/세포)
0
50
250
500
1000
2000
10000
0
5.2
14.6
22.2
25.7
19.1
15.4

도면58b

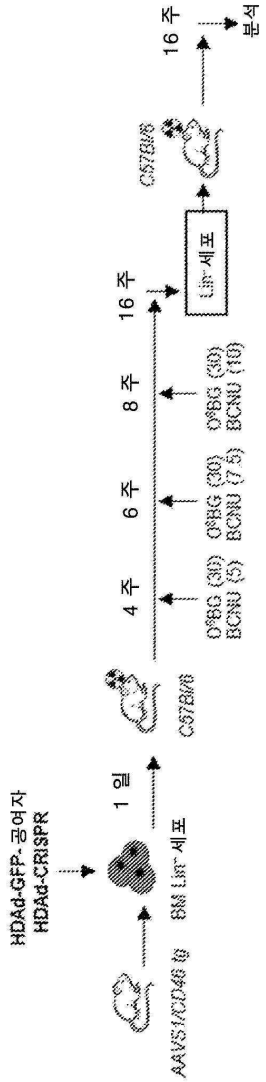


도면58c

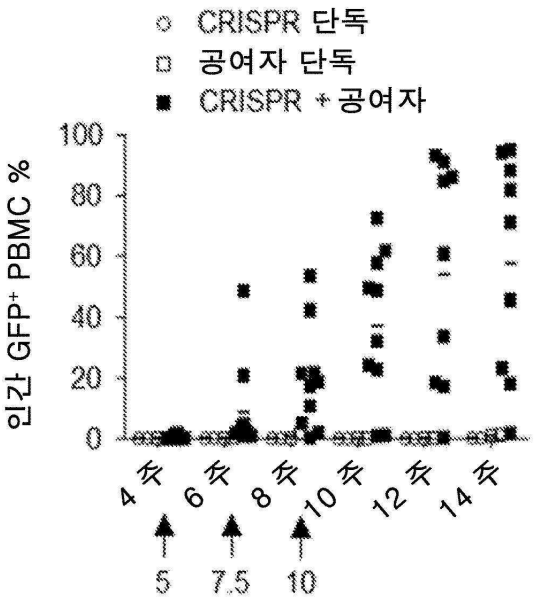


(서열번호: 19-23, 21, 21, 26-30, 27, 32, 28, 34-47)

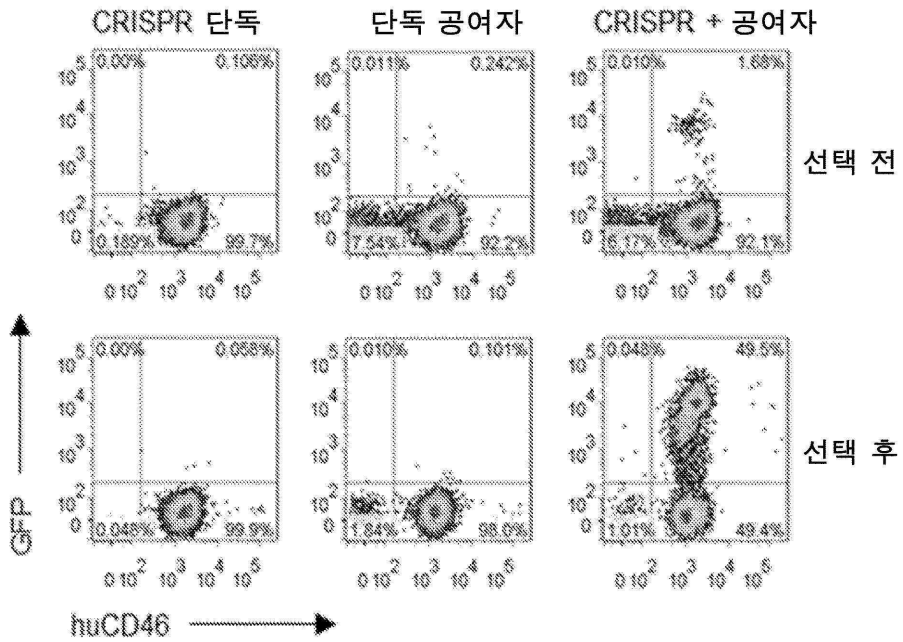
도면59a


AAVS1/CD46 Ig
6N Lin⁻ 세포
HDAd-GFP-공여자
HDAd-CRISPR
1 일
C57BL/6
4 주
6 주
8 주
16 주
OBG (30)
BCNU (5)
OBG (30)
BCNU (7.5)
OBG (30)
BCNU (10)
Lin⁻ 세포
C57BL/6
16 주
분석

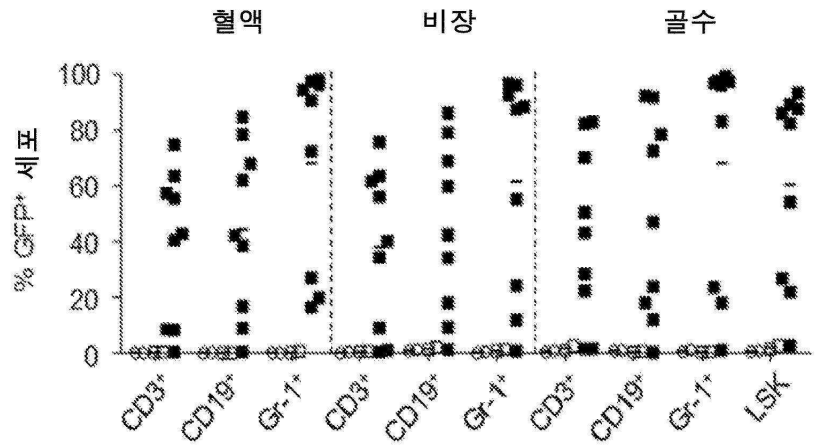
도면59b


CRISPR 단독
공여자 단독
CRISPR + 공여자
인간 GFP⁺ PBMC %
100
80
60
40
20
0
4 주
6 주
8 주
10 주
12 주
14 주
5
7.5
10

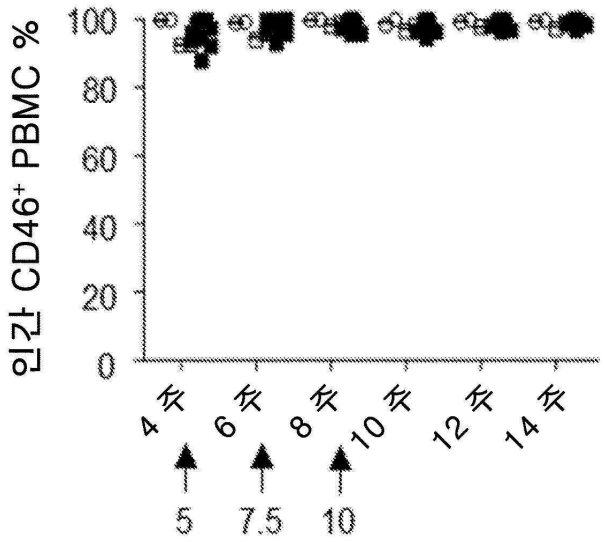
도면59c


CRISPR 단독
단독 공여자
CRISPR + 공여자
선택 전
선택 후
GFP
huCD46

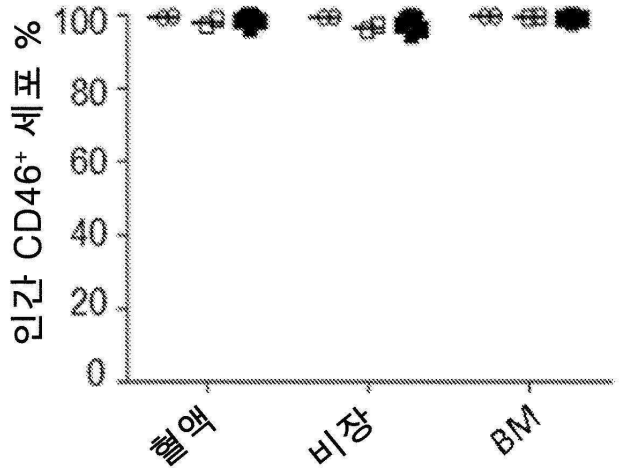
도면59d


혈액
비장
골수
% GFP+ 버퍼
CD3+
CD19+
Gr-1+
LSK

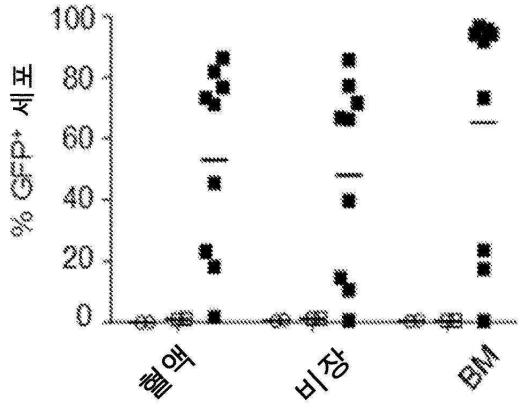
도면60a



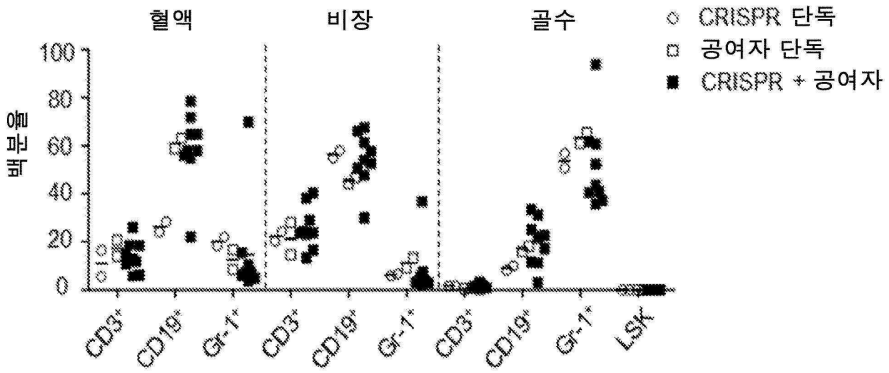

도면60b



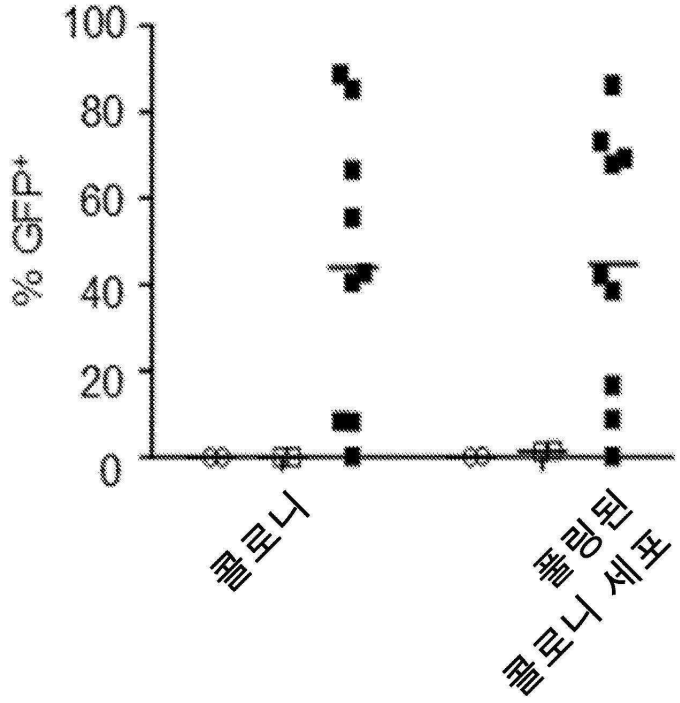

도면60c



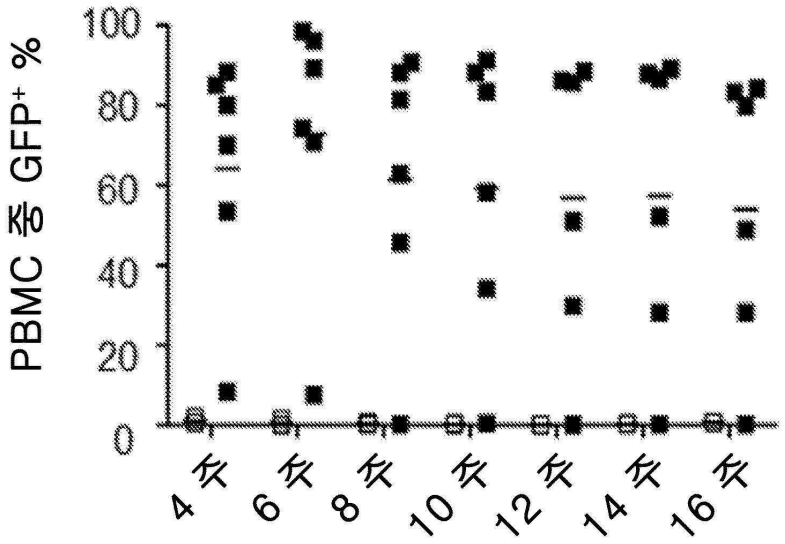
도면60d


혈액
비장
골수
CRISPR 단독
공여자 단독
CRISPR + 공여자
퍼센트
100
80
60
40
20
0
CD3+
CD19+
Gr-1+
CD3+
CD19+
Gr-1+
CD3+
CD19+
Gr-1+
LSK

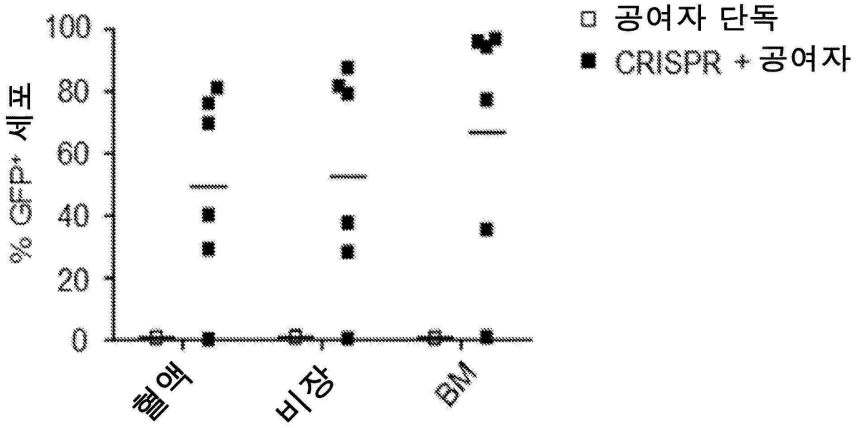
도면60e


% GFP+
100
80
60
40
20
0
콜로니
돌림된 콜로니 세포

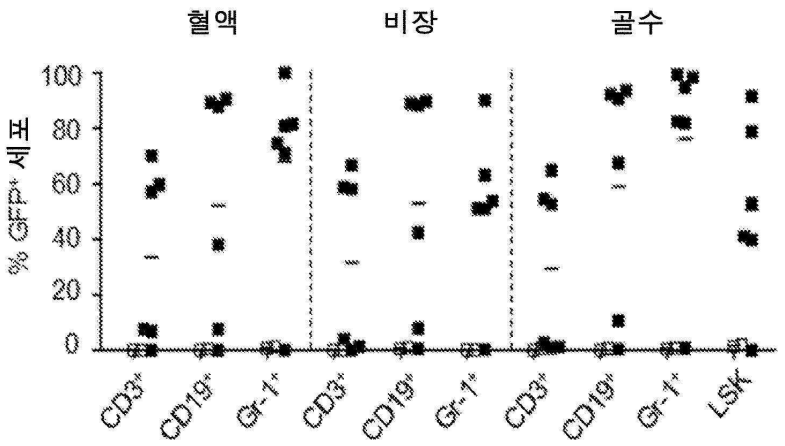
도면61a



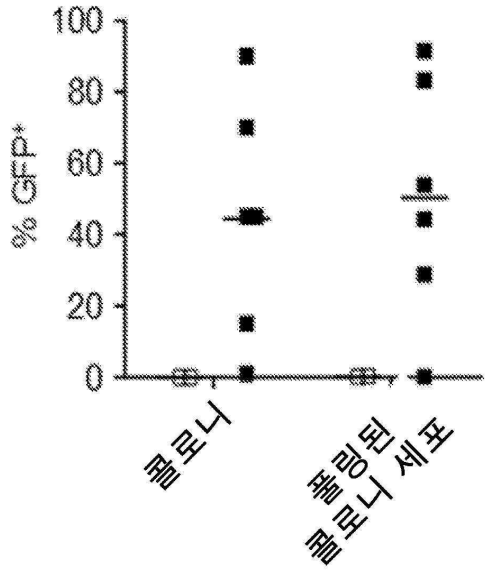

도면61b



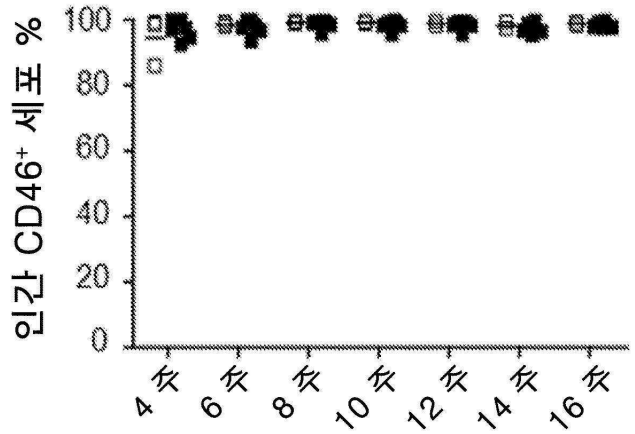

도면61c



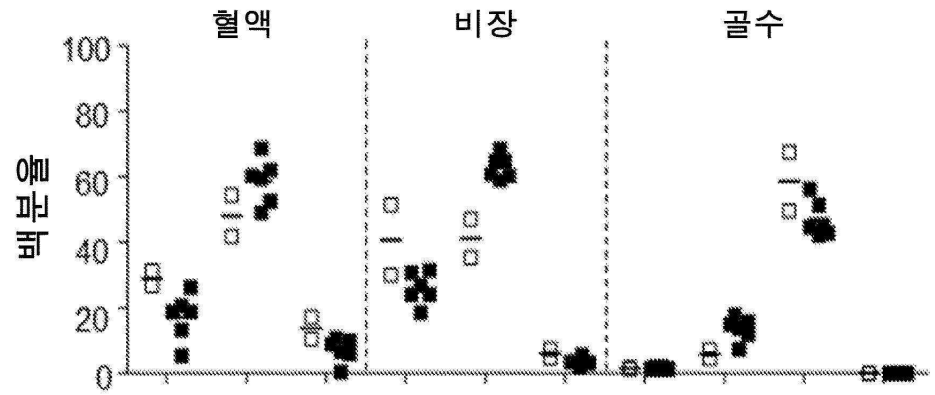
도면61d



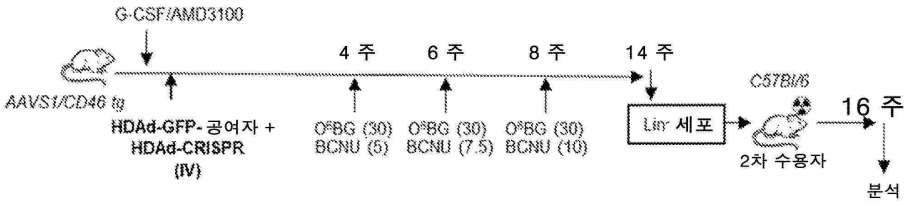

도면61e



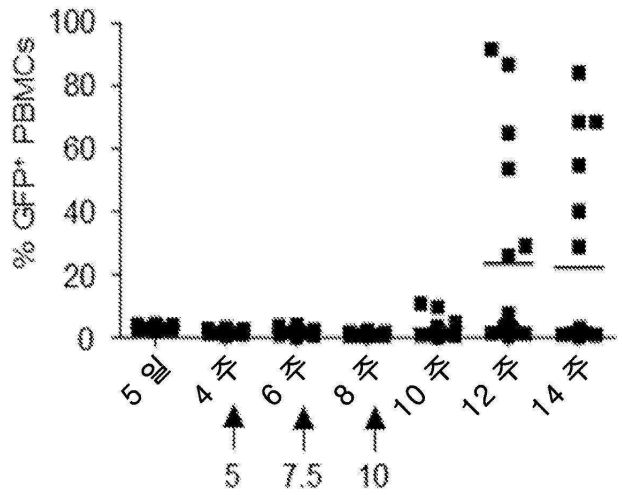

도면61f



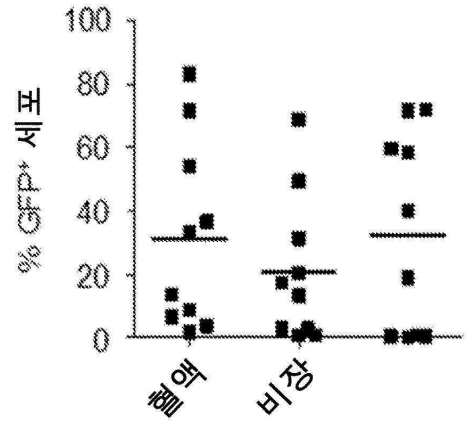
도면62a



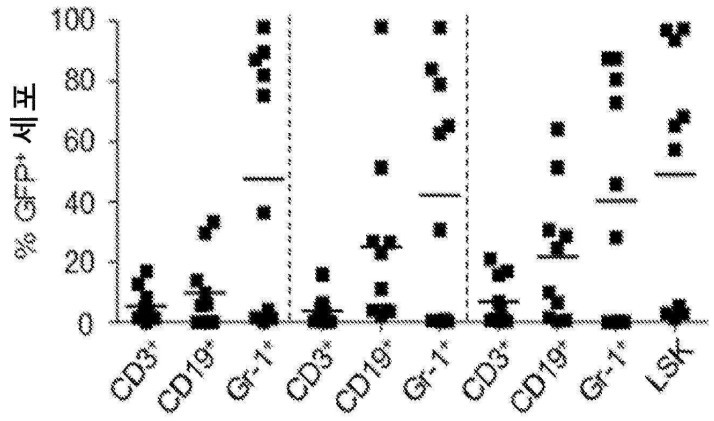

도면62b



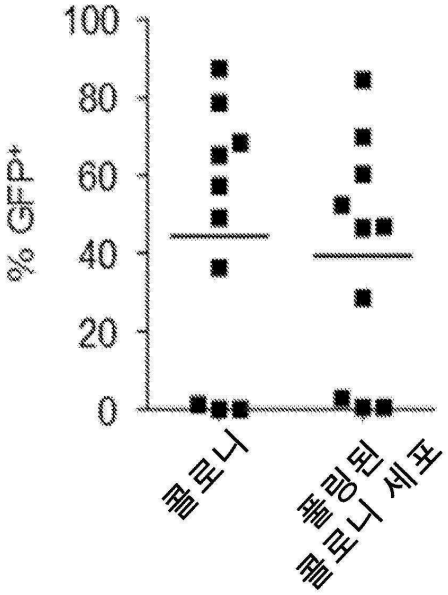

도면62c



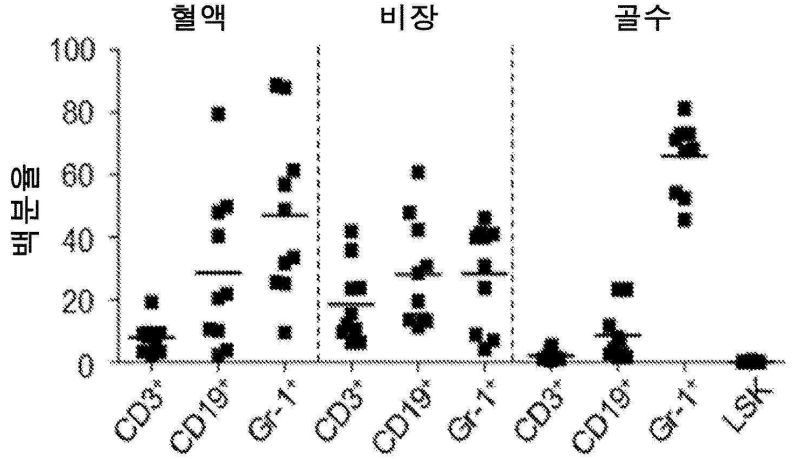
도면62d



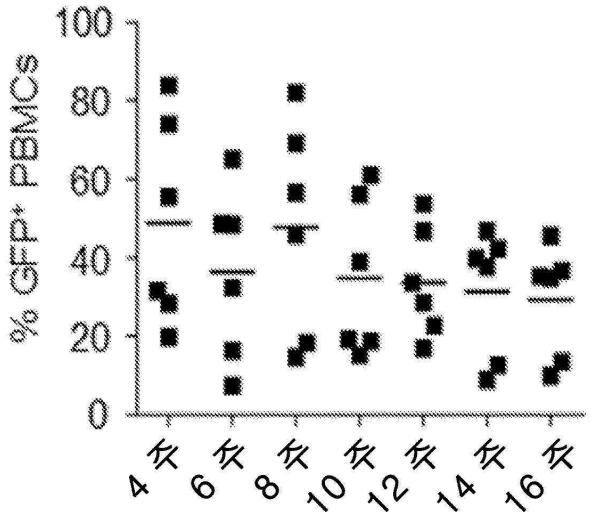

도면62e



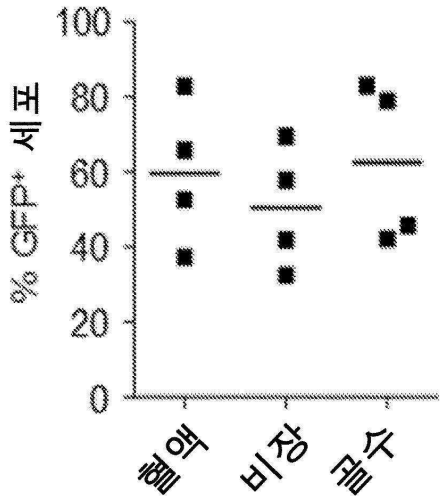

도면62f



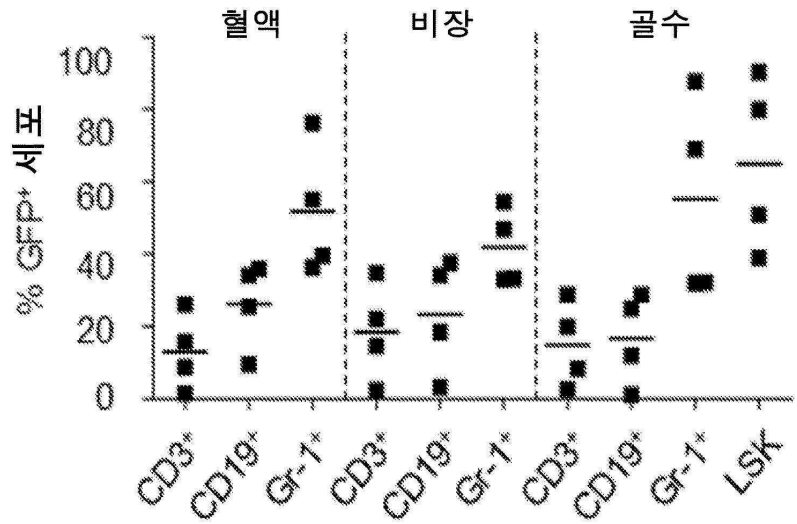
도면63a



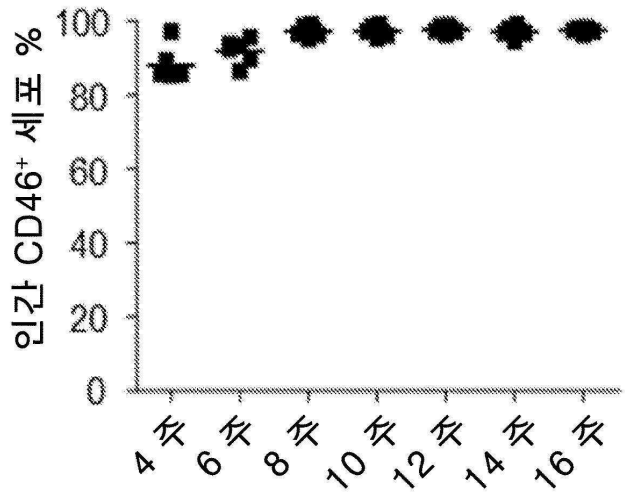

도면63b



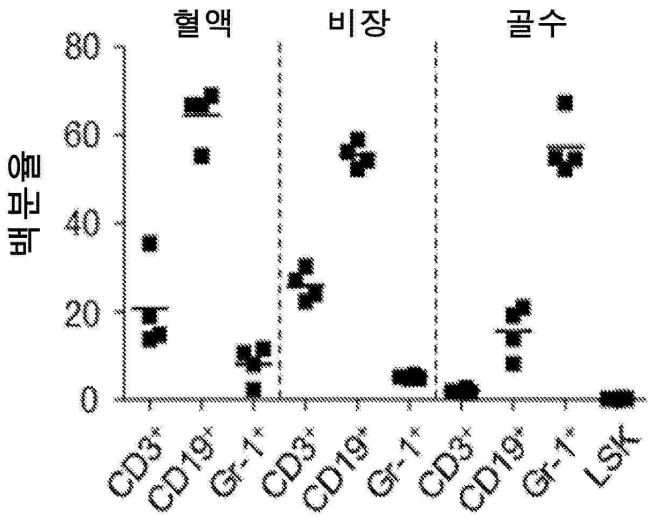

도면63c



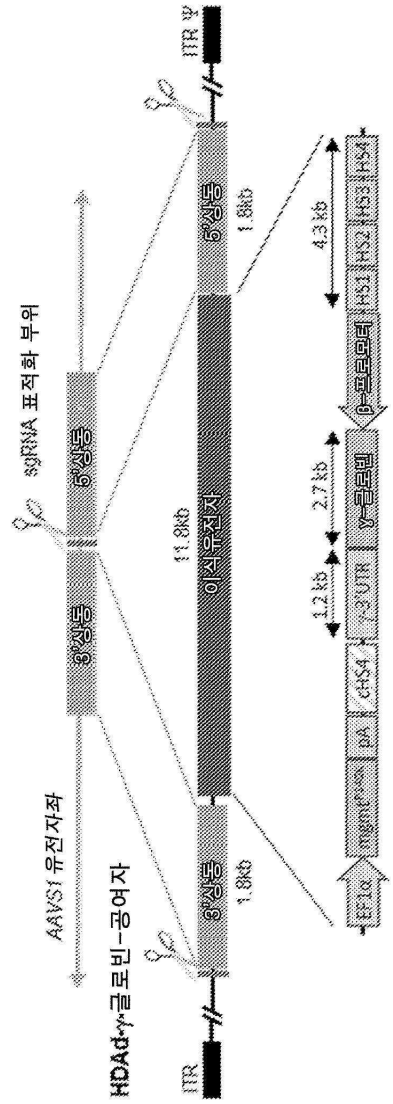
도면63d


인간 CD46+ 세포 %
100
80
60
40
20
0
4 주
6 주
8 주
10 주
12 주
14 주
16 주

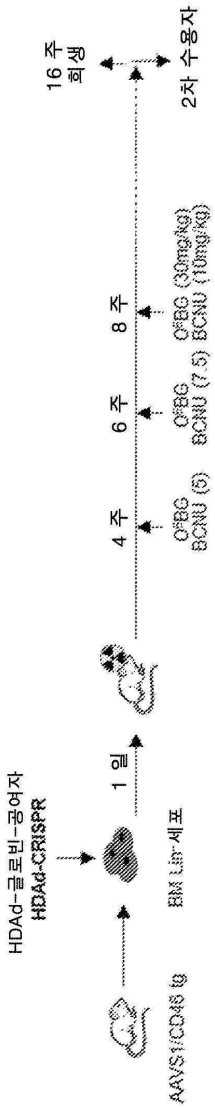
도면63e


혈액
비장
골수
백분율
80
60
40
20
0
CD3+
CD19+
Gr-1+
CD3+
CD19+
Gr-1+
CD3+
CD19+
Gr-1+
LSK

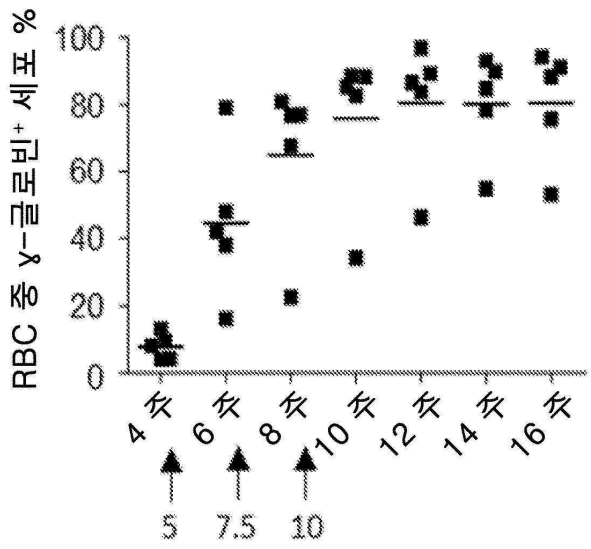
도면64a


ITR Ψ
sgRNA 표적화 부위
5'상동
3'상동
1.8kb
11.8kb
이식유전자
4.3kb
2.7kb
1.2kb
β-글로빈
HS1
HS2
HS3
HS4
γ-글로빈
3'UTR
cHS4
pA
EF1α
AAVS1 유전자좌
HDAd-γ-글로빈-공여자
ITR

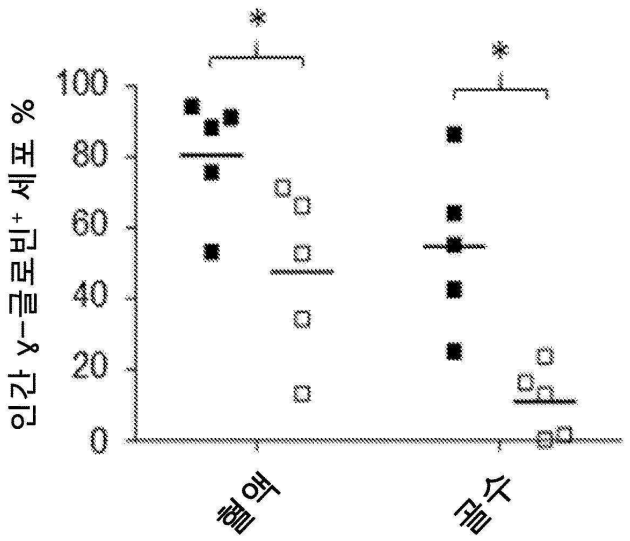
도면64b


AAVS1/CD46 tg
BM Lin⁻ 세포
HDAd-글로빈-공여자
HDAd-CRISPR
1 일
4 주
O⁶BG
BCNU (5)
6 주
O⁶BG
BCNU (7.5)
8 주
O⁶BG (30mg/kg)
BCNU (10mg/kg)
16 주
희생
2차 수용자

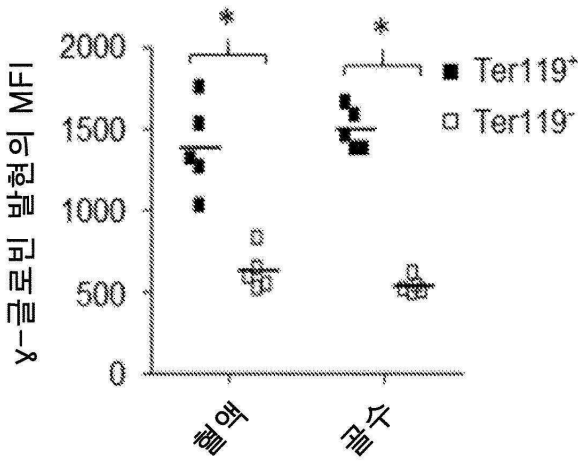
도면64c


RBC 중 γ-글로빈⁺ 세포 %
100
80
60
40
20
0
4 주
6 주
8 주
10 주
12 주
14 주
16 주
5
7.5
10

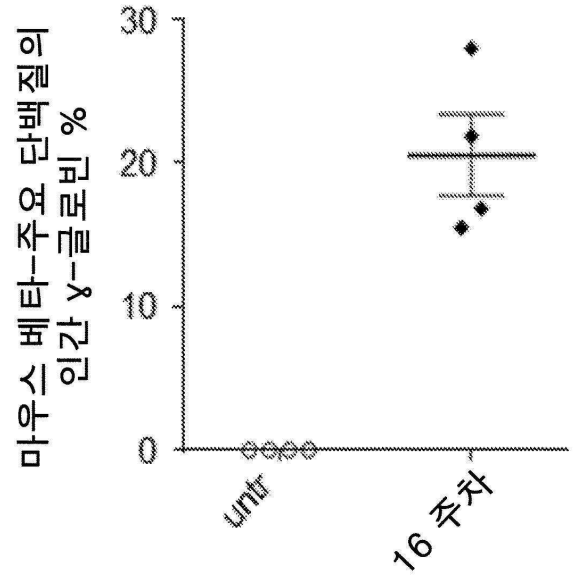
도면64d



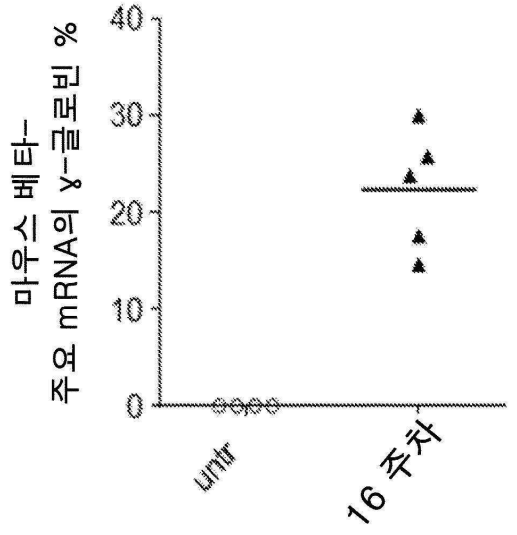

도면64e



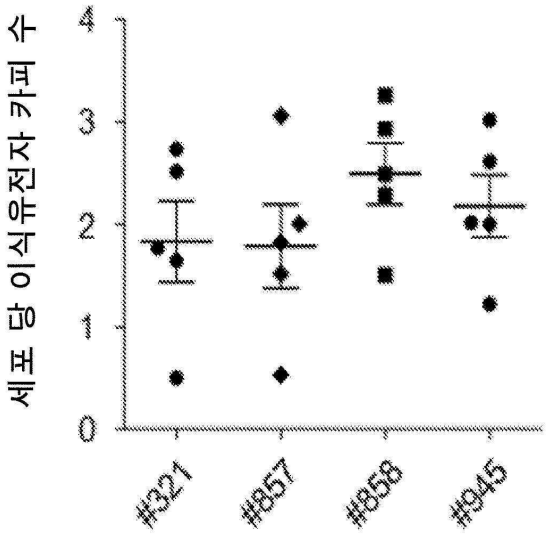

도면64f



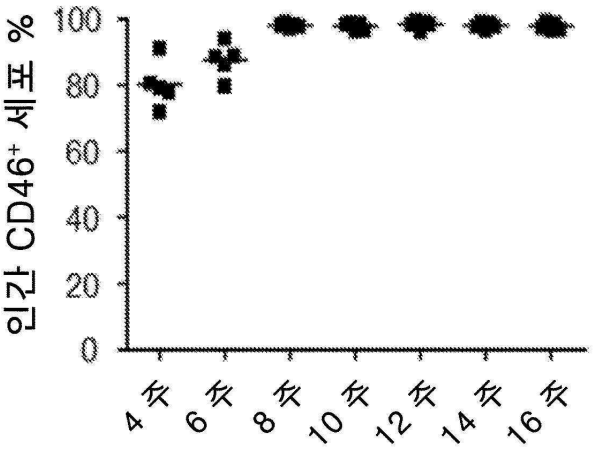
도면64g



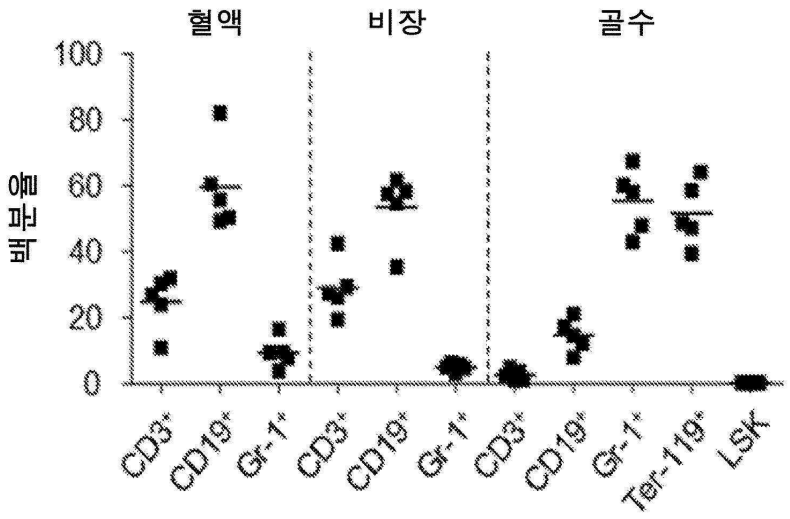

도면64h



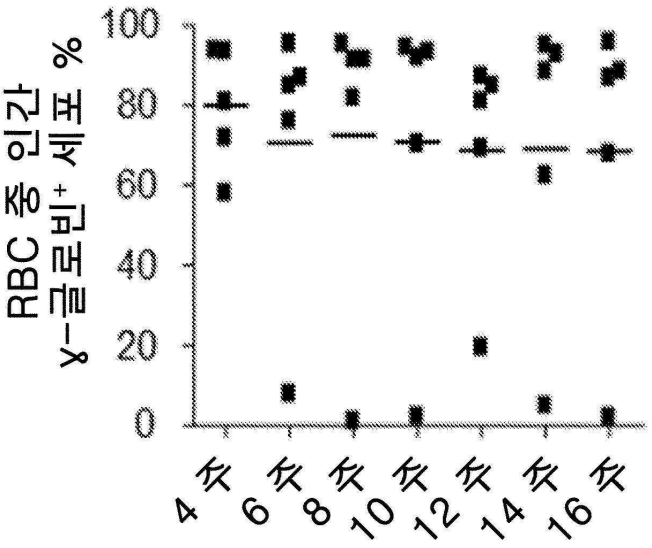

도면65a



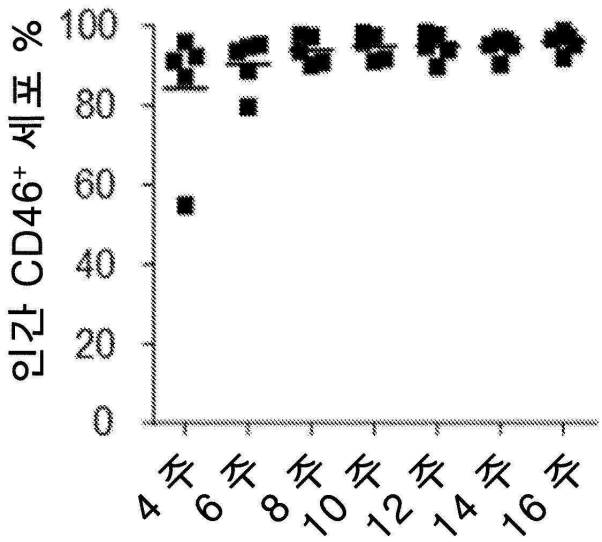
도면65b


혈액
비장
골수
백분율
100
80
60
40
20
0
CD3+
CD19+
Gr-1+
CD3+
CD19+
Gr-1+
CD3+
CD19+
Gr-1+
Ter-119+
LSK

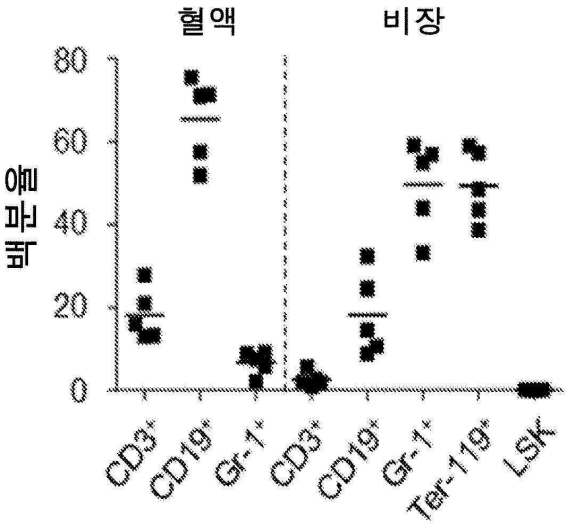
도면66a


RBC 중 인간 γ-글로빈+ 세포 %
100
80
60
40
20
0
4 주
6 주
8 주
10 주
12 주
14 주
16 주

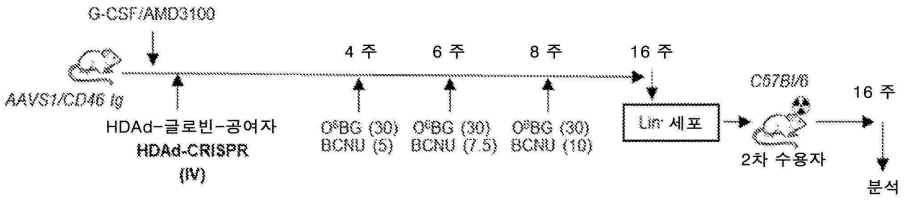
도면66b



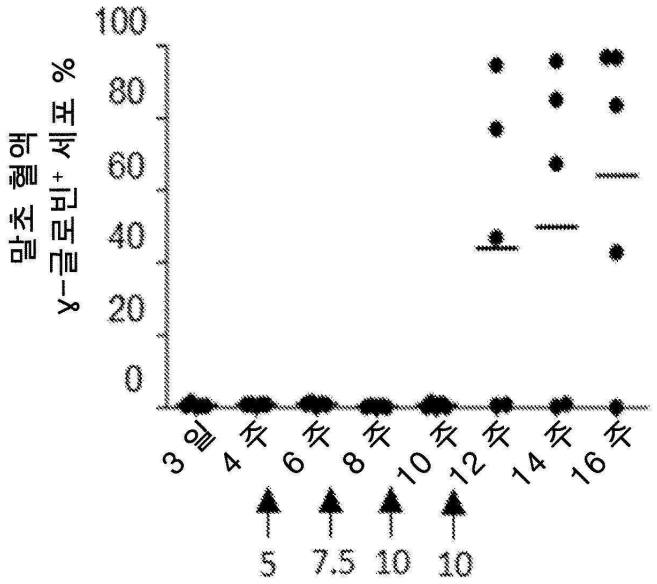

도면66c



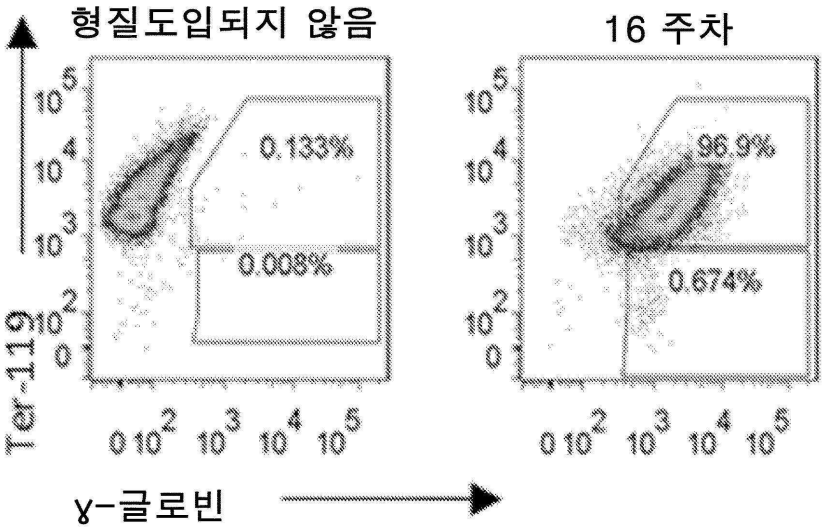

도면67a



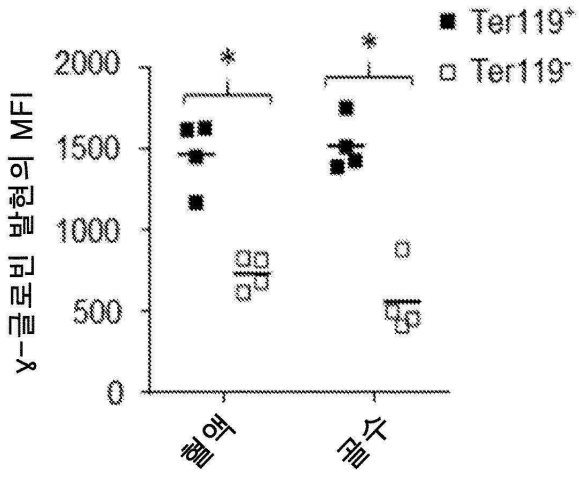
도면67b


말초 혈액 γ-글로빈+ 세포 %
100
80
60
40
20
0
3 일
4 주
6 주
8 주
10 주
12 주
14 주
16 주
5
7.5
10
10

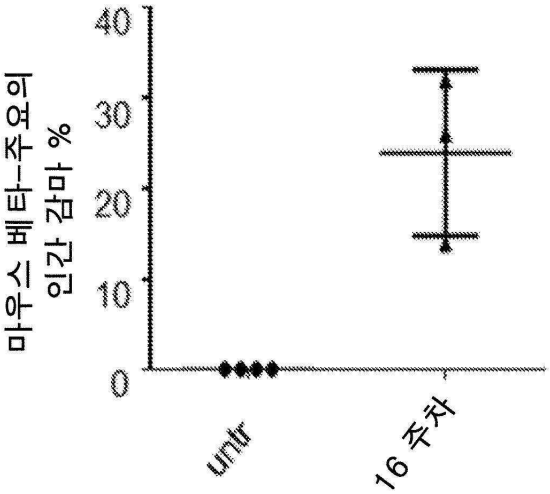
도면67c


형질도입되지 않음
16 주차
Ter-119
γ-글로빈
0.133%
0.008%
96.9%
0.674%

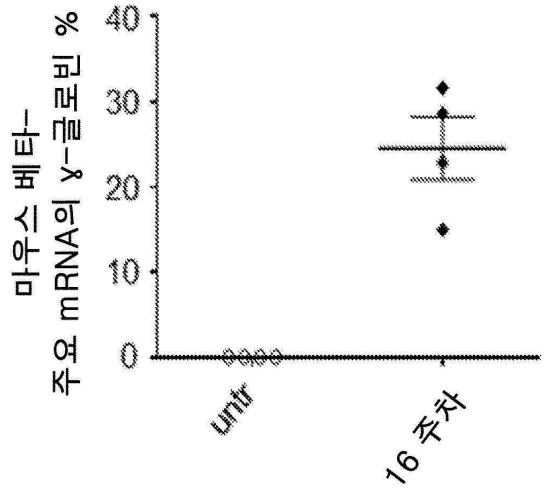
도면67d



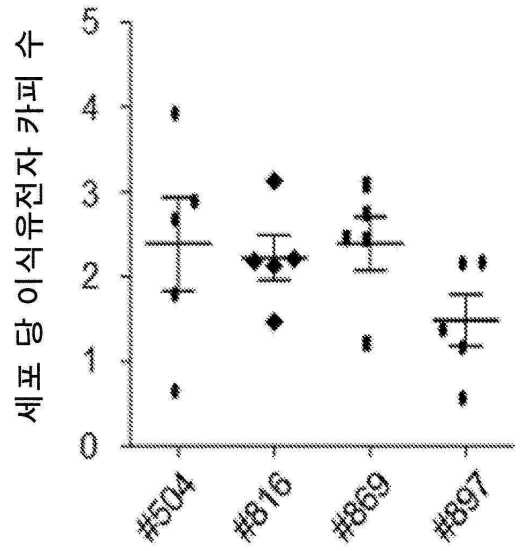

도면67e



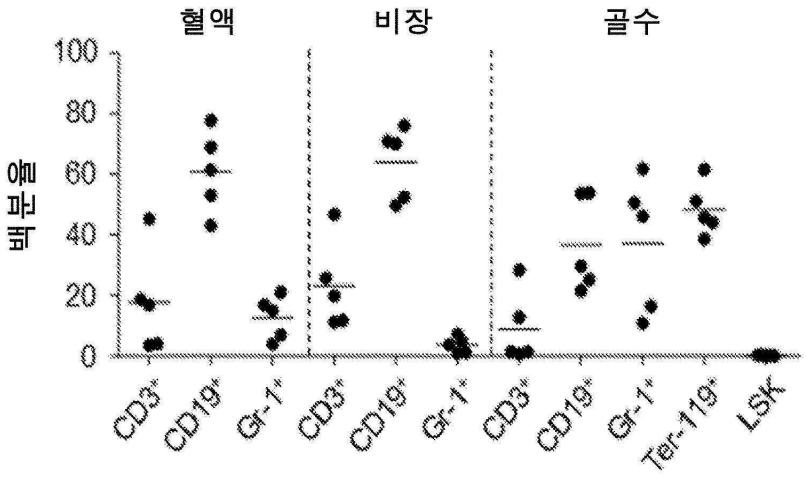

도면67f



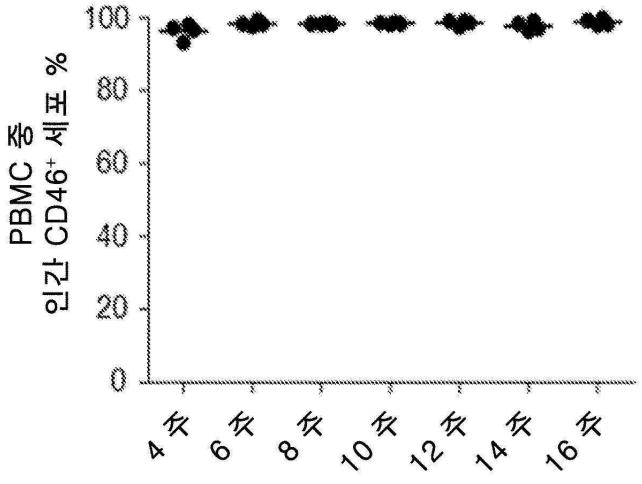
도면67g



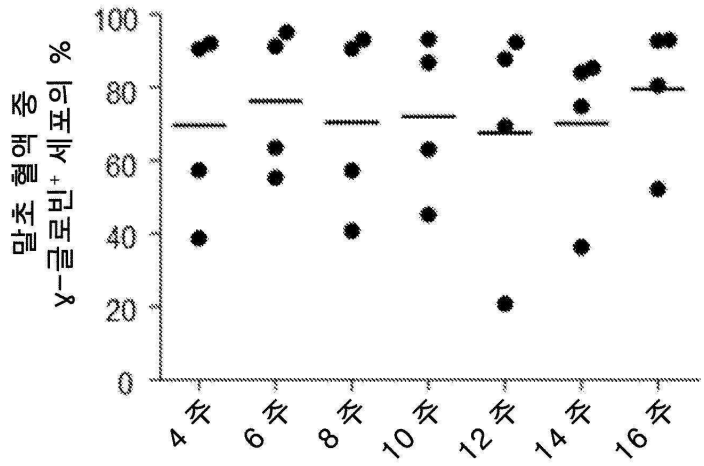

도면67h



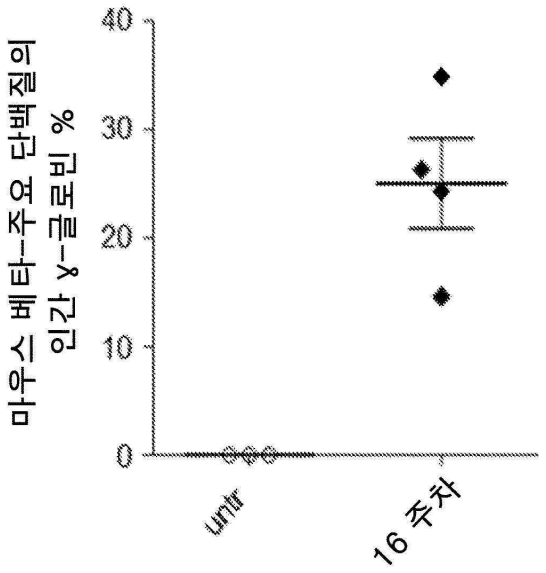

도면68a



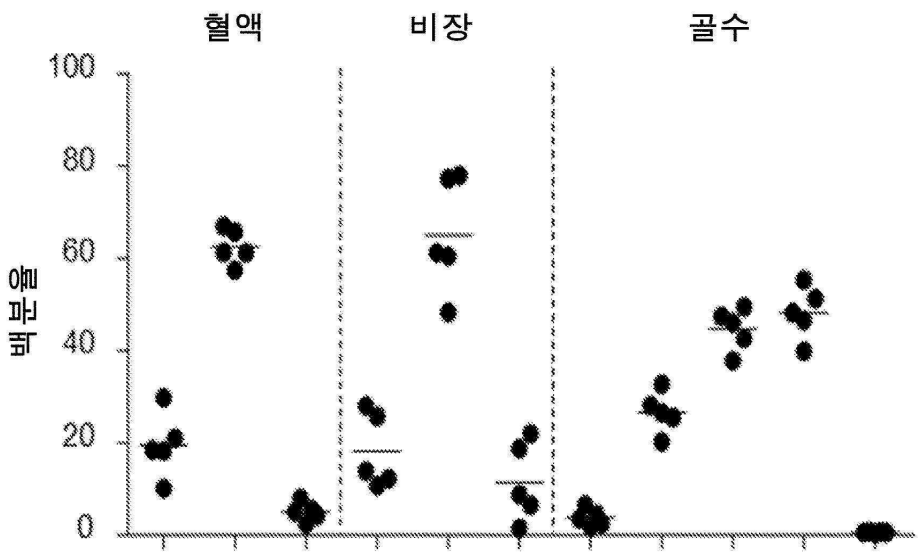
도면68b


말초 혈액 중
γ-글로빈+ 세포의 %
100
80
60
40
20
0
4 주
6 주
8 주
10 주
12 주
14 주
16 주

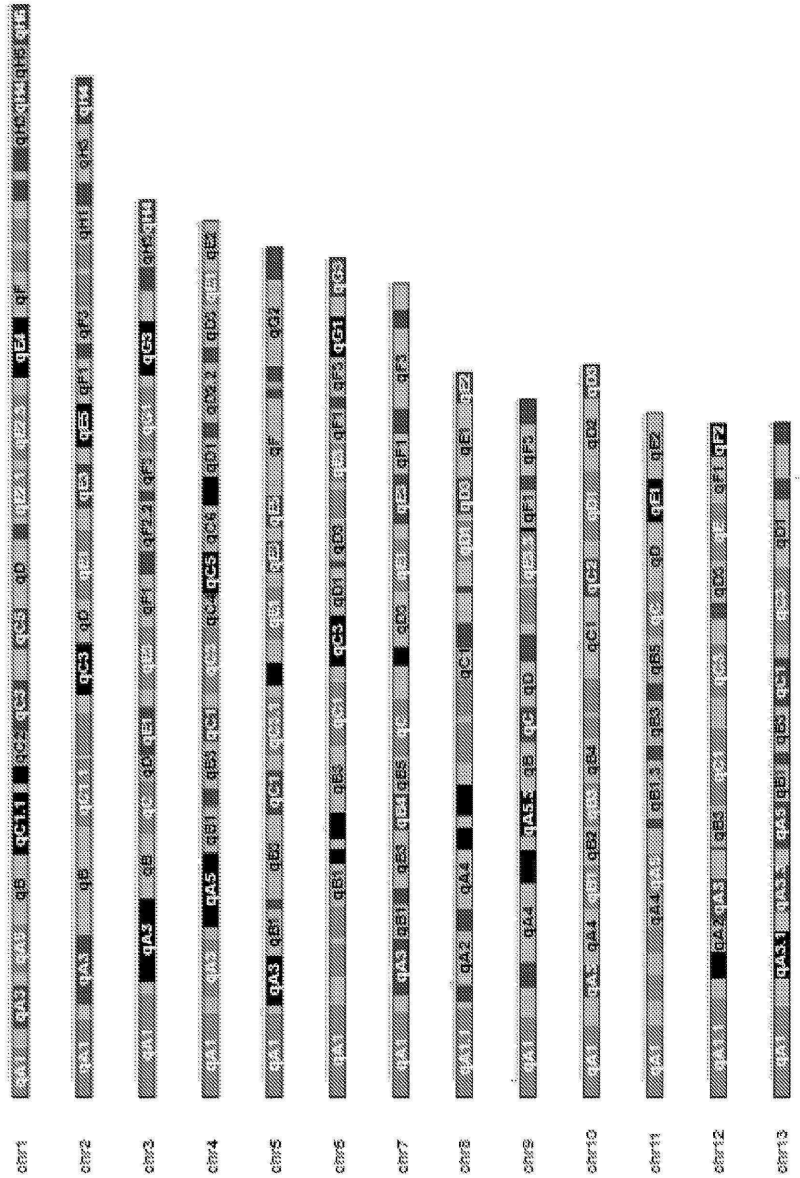
도면68c


마우스 베타-주요 단백질의
인간 γ-글로빈 %
40
30
20
10
0
unir
16 주차

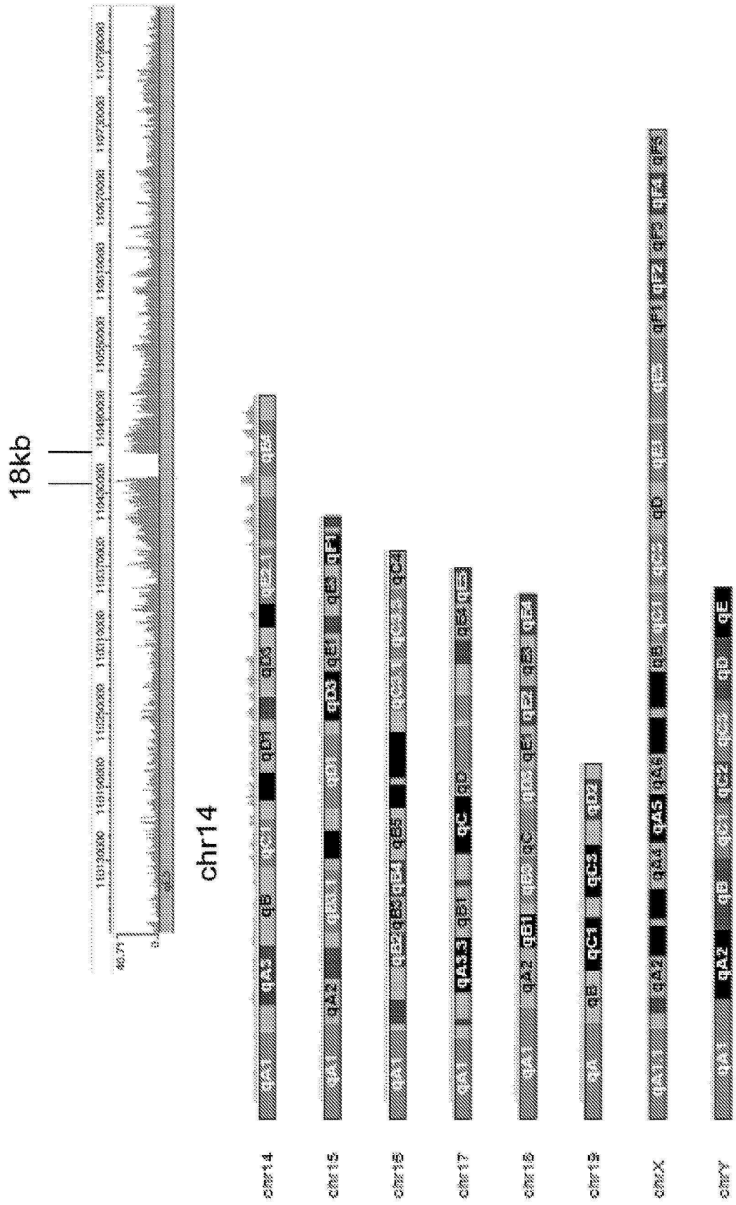
도면68d



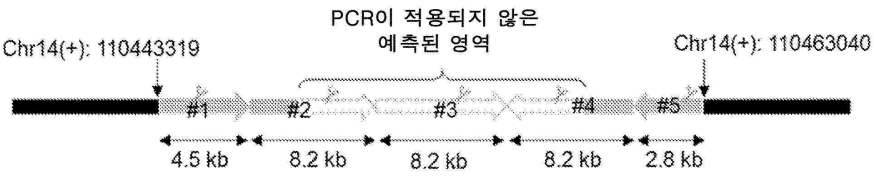
도면69a


chr1
chr2
chr3
chr4
chr5
chr6
chr7
chr8
chr9
chr10
chr11
chr12
chr13

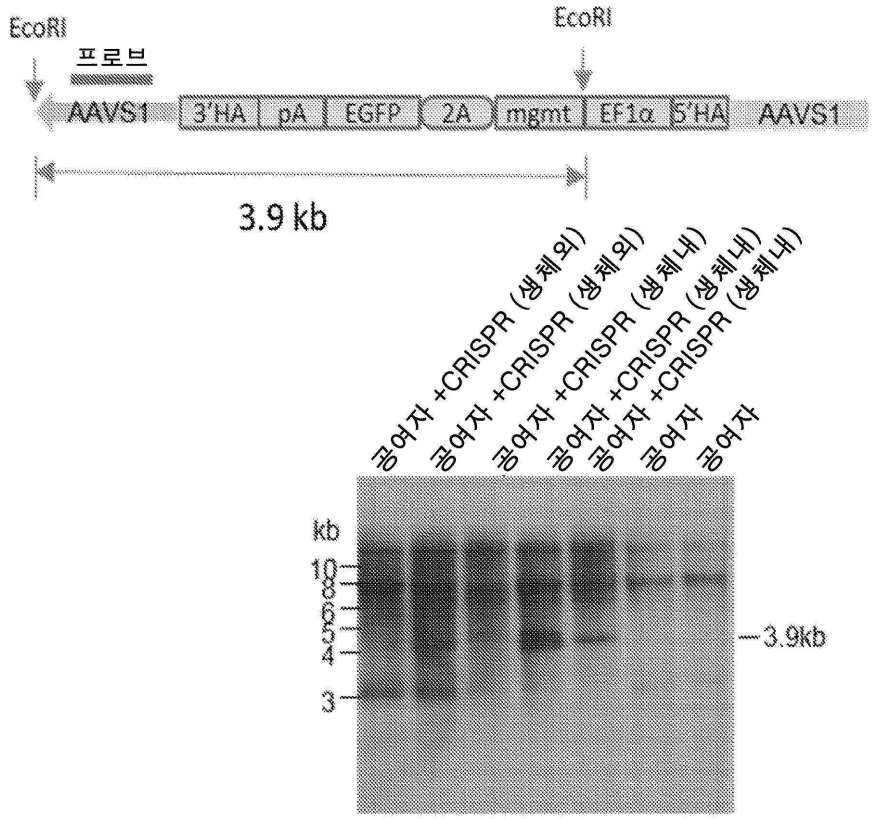
도면69b


18kb
chr14
chr14
chr15
chr16
chr17
chr18
chr19
chrX
chrY

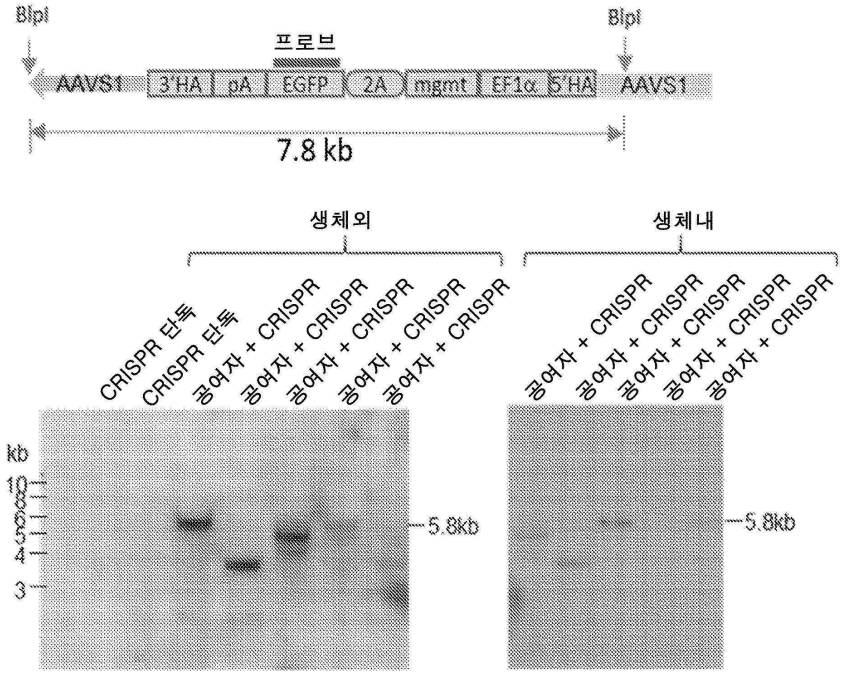
도면70


PCR이 적용되지 않은
예측된 영역
Chr14(+): 110443319
Chr14(+): 110463040
#1
#2
#3
#4
#5
4.5 kb
8.2 kb
8.2 kb
8.2 kb
2.8 kb

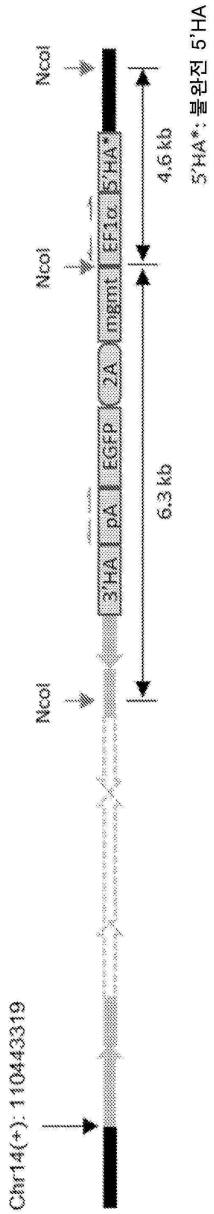
도면71a


EcoRI
프로브
EcoRI
AAVS1
3'HA
pA
EGFP
2A
mgmt
EF1α
5'HA
AAVS1
3.9 kb
공여자 +CRISPR (생체외)
공여자 +CRISPR (생체외)
공여자 +CRISPR (생체내)
공여자 +CRISPR (생체내)
공여자 +CRISPR (생체내)
공여자
공여자
kb
10
8
6
5
4
3
3.9kb

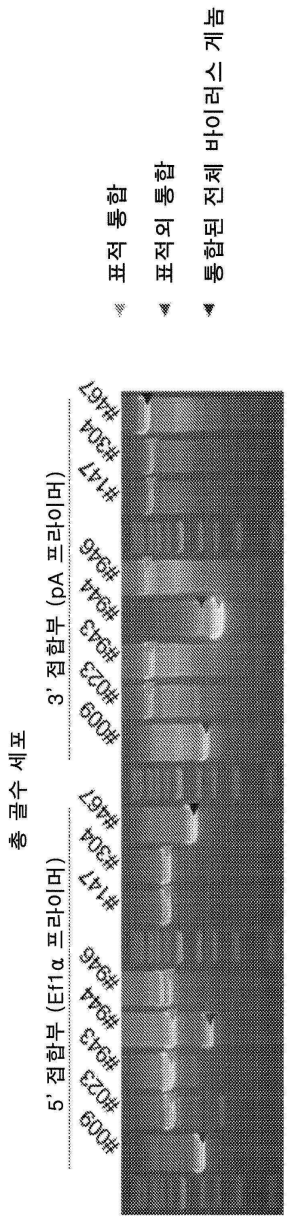
도면71b


BlpI
프로브
BlpI
AAVS1
3'HA
pA
EGFP
2A
mgmt
EF1α
5'HA
AAVS1
7.8 kb
생체외
생체내
CRISPR 단독
CRISPR 단독
공여자 + CRISPR
공여자 + CRISPR
공여자 + CRISPR
공여자 + CRISPR
공여자 + CRISPR
공여자 + CRISPR
공여자 + CRISPR
공여자 + CRISPR
공여자 + CRISPR
공여자 + CRISPR
kb
10
8
6
5
4
3
5.8kb
5.8kb

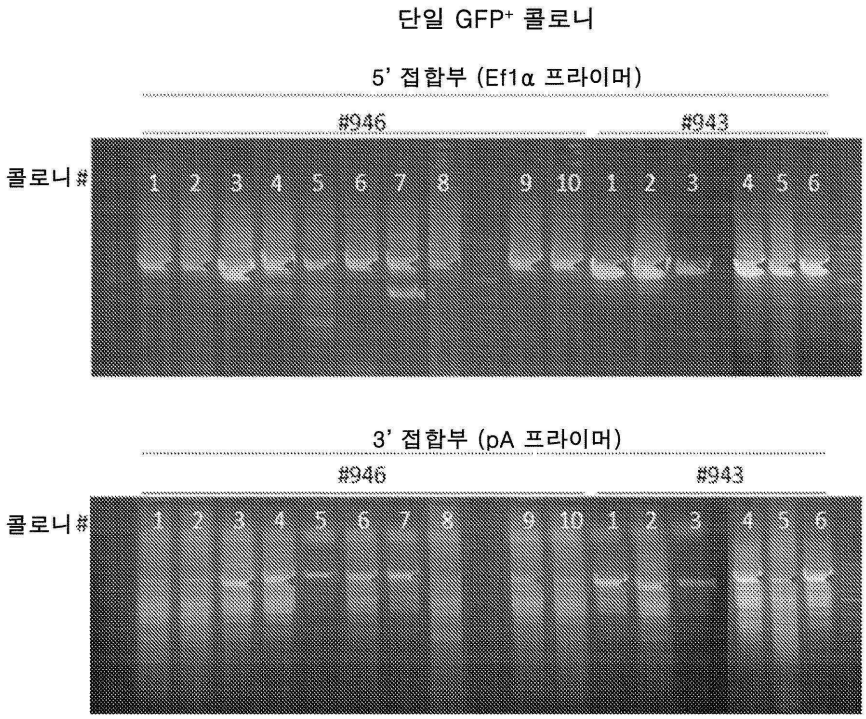
도면72a


Chr14(+): 110443319
NcoI
NcoI
NcoI
3'HA
pA
EGFP
2A
mgmt
EF1α
5'HA*
6.3 kb
4.6 kb
5'HA*: 불완전 5'HA

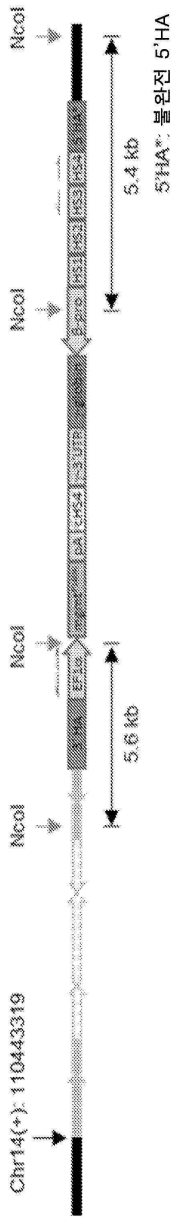
도면72b


표적 통합
표적외 통합
통합된 전체 바이러스 게놈
총 골수 세포
5' 접합부 (Ef1α 프라이머)
3' 접합부 (pA 프라이머)
#009
#023
#943
#944
#946
#147
#304
#467

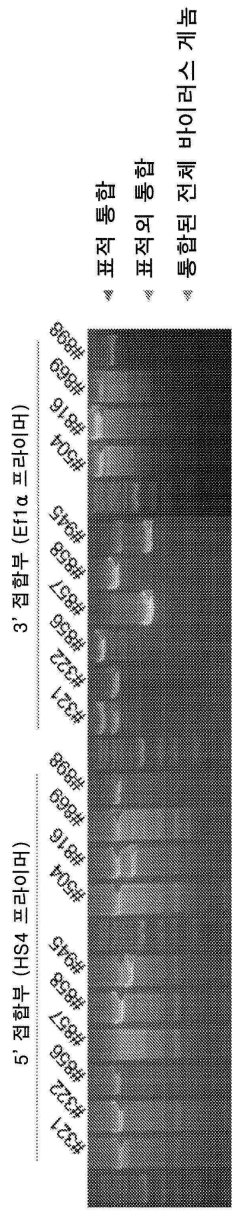
도면72c


단일 GFP+ 콜로니
5' 접합부 (Ef1α 프라이머)
#946
#943
콜로니#
1 2 3 4 5 6 7 8 9 10 1 2 3 4 5 6
3' 접합부 (pA 프라이머)
#946
#943
콜로니#
1 2 3 4 5 6 7 8 9 10 1 2 3 4 5 6

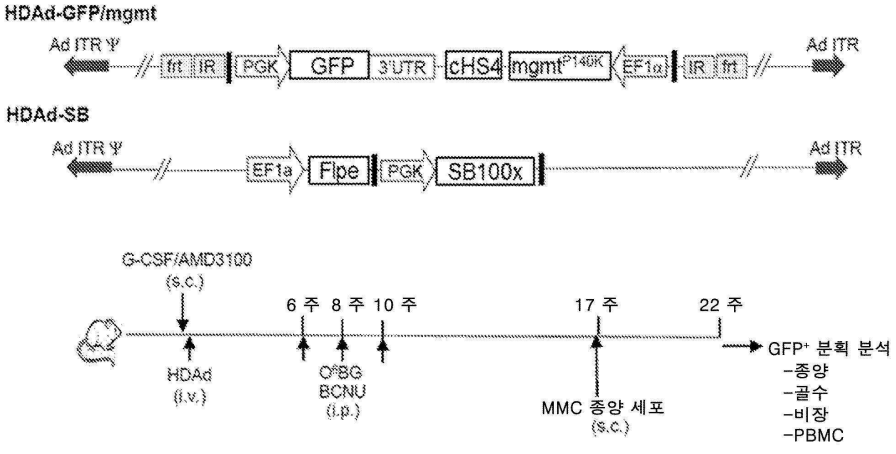
도면73a



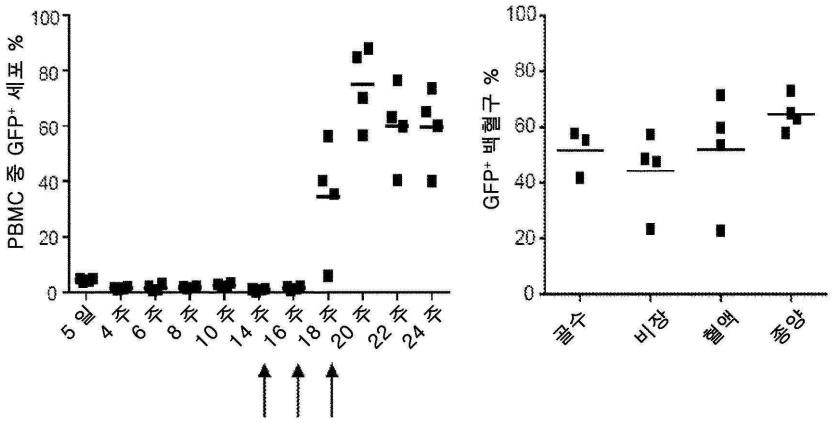
도면73b


표적 통합
표적외 통합
통합된 전체 바이러스 게놈
3' 접합부 (Ef1α 프라이머)
#898
#639
#816
#504
#945
#858
#857
#856
#322
#321
5' 접합부 (HS4 프라이머)
#898
#639
#816
#504
#945
#858
#857
#856
#322
#321

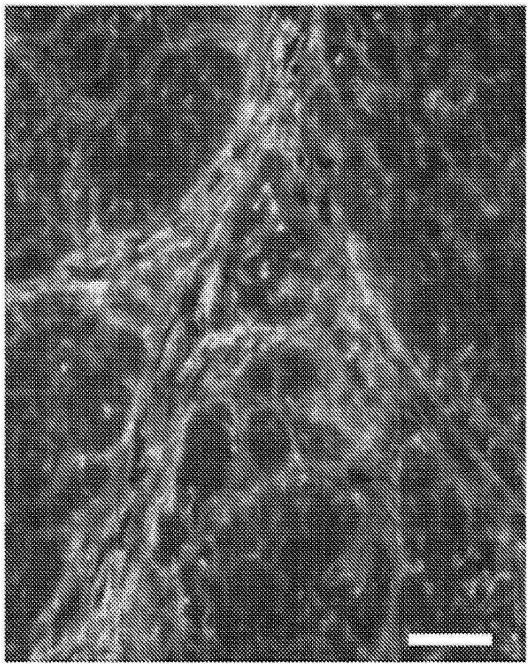
도면74a


HDAd-GFP/mgmt
Ad ITR Ψ
frt
IR
PGK
GFP
3'UTR
cHS4
mgmtP140K
EF1α
IR
frt
Ad ITR
HDAd-SB
Ad ITR Ψ
EF1a
Flpe
PGK
SB100x
Ad ITR
G-CSF/AMD3100
(s.c.)
6 주
8 주
10 주
17 주
22 주
HDAd
(i.v.)
O6BG
BCNU
(i.p.)
MMC 종양 세포
(s.c.)
GFP+ 분획 분석
-종양
-골수
-비장
-PBMC

도면74b

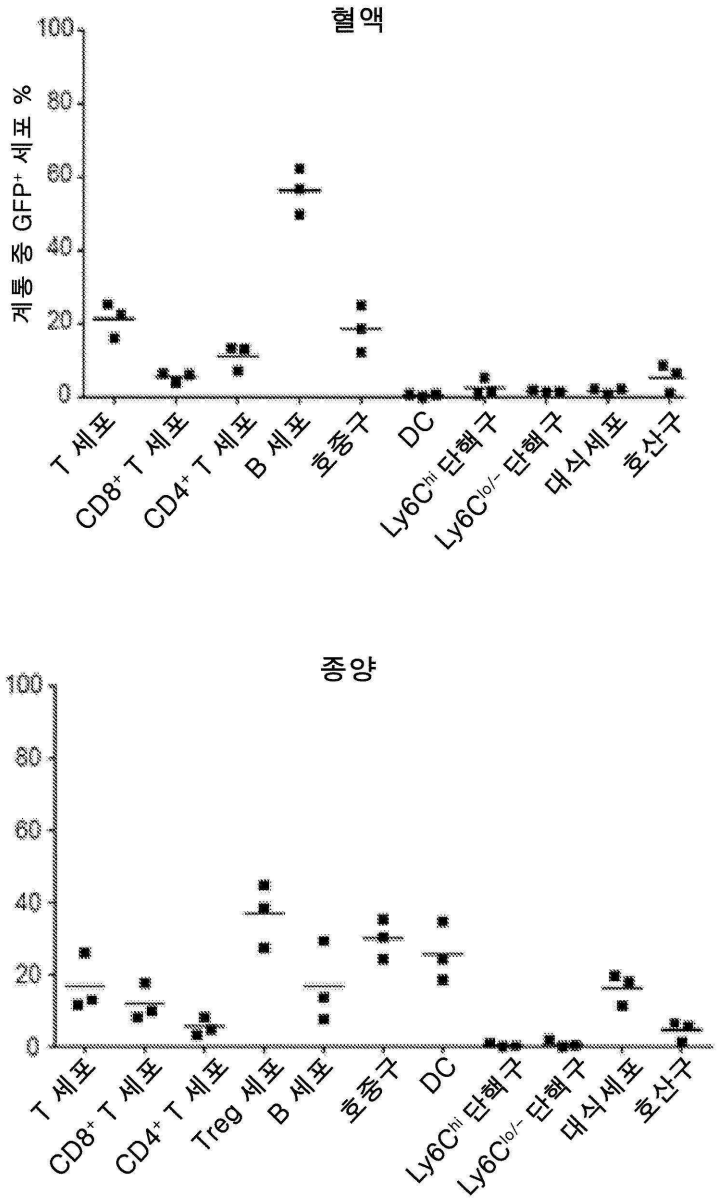



도면74c

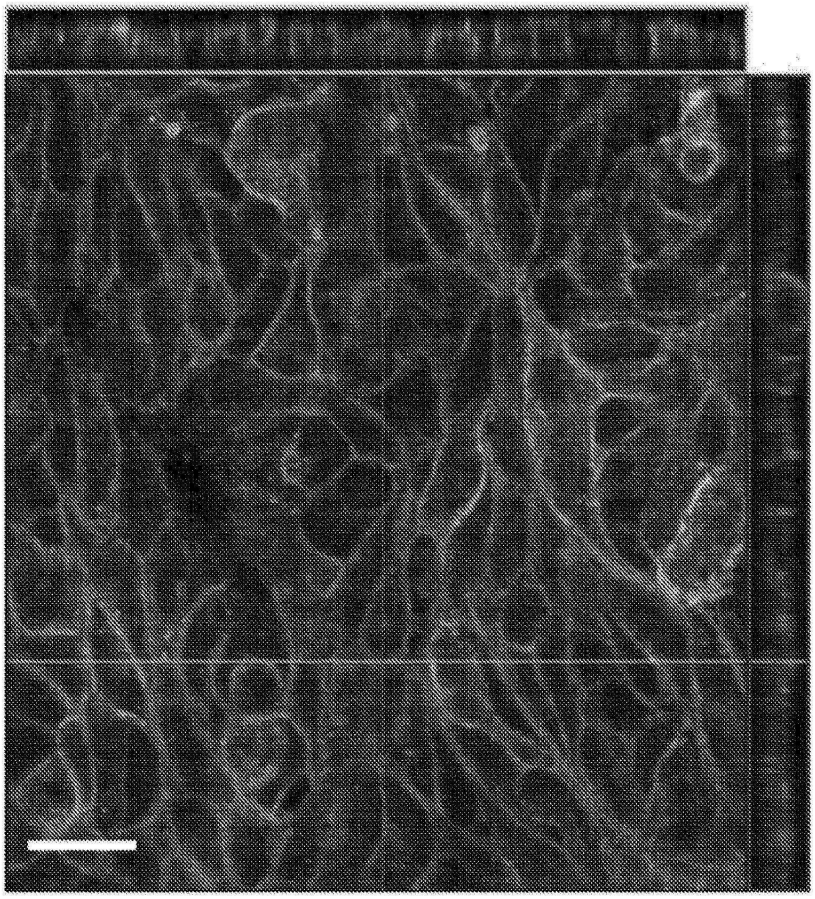


GFP/라미닌

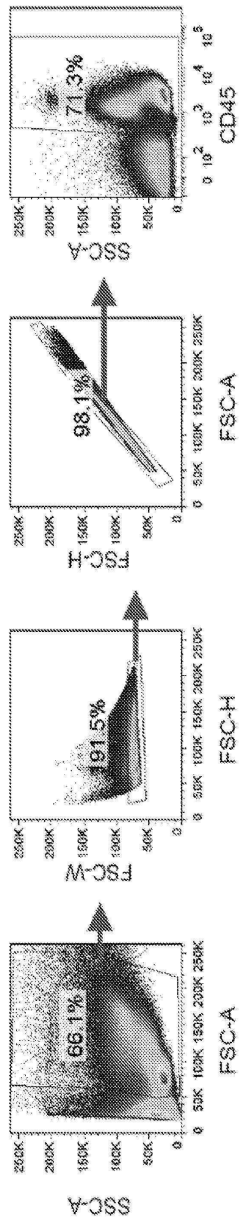
도면74d


혈액
계통 중 GFP+ 세포 %
100
80
60
40
20
0
T 세포
CD8+ T 세포
CD4+ T 세포
B 세포
호중구
DC
Ly6Chi 단핵구
Ly6Clo/- 단핵구
대식세포
호산구
종양
100
80
60
40
20
0
T 세포
CD8+ T 세포
CD4+ T 세포
Treg 세포
B 세포
호중구
DC
Ly6Chi 단핵구
Ly6Clo/- 단핵구
대식세포
호산구

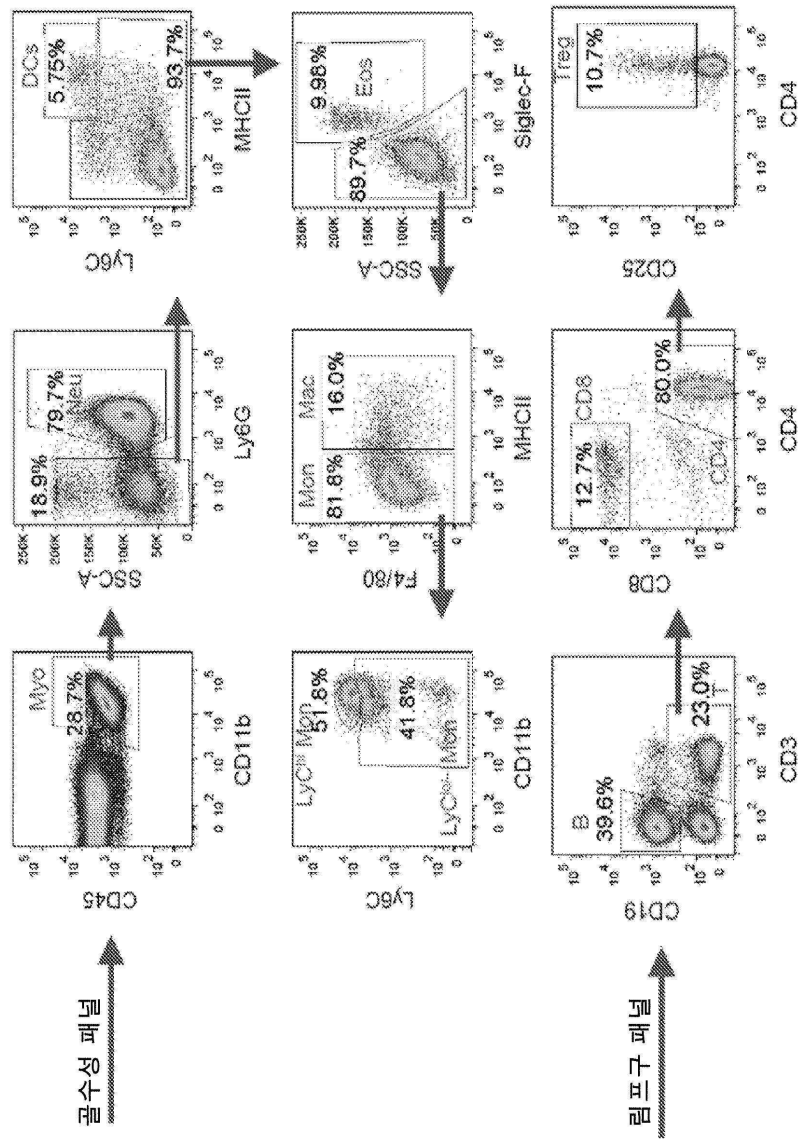
도면75



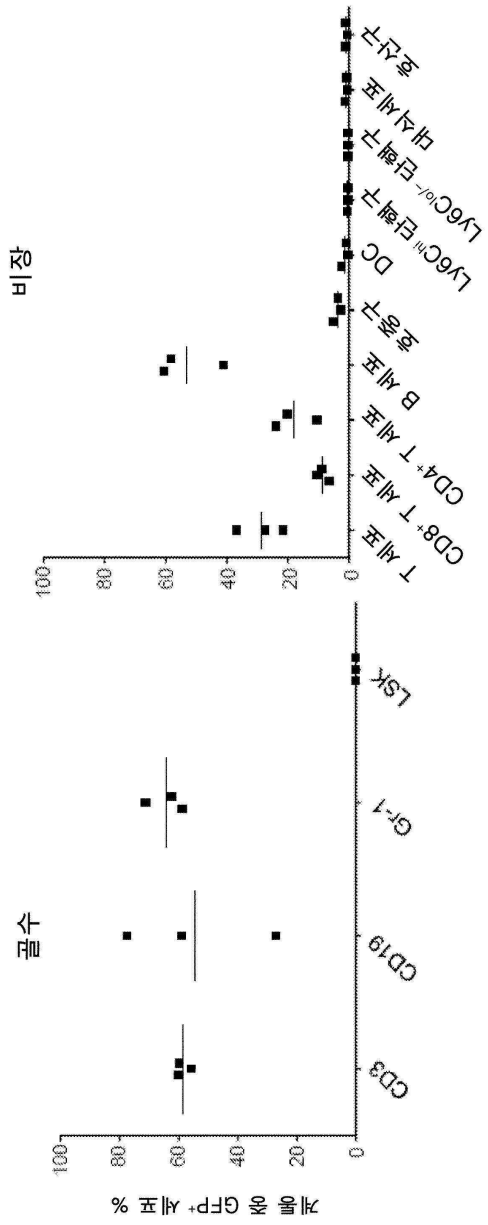
도면76a



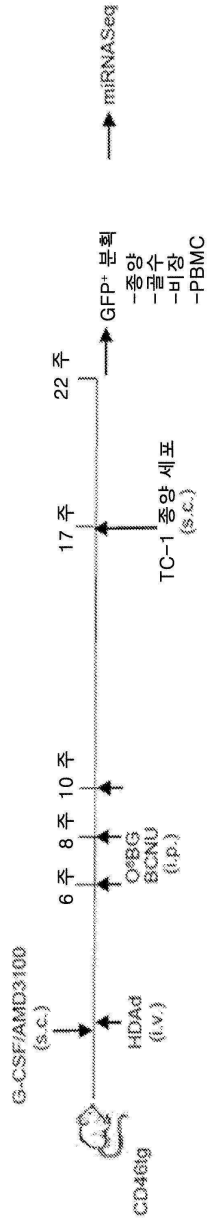
도면76b


골수성 패널
림프구 패널
CD45
CD11b
Myo
28.7%
Ly6G
SSC-A
18.9%
79.7%
Neu
Ly6C
DCs
5.75%
93.7%
MHCII
Ly6C
CD11b
Ly6C+ Mon
51.8%
41.8%
Ly6C- Mon
F4/80
SSC-A
Mon
81.8%
Mac
16.0%
MHCII
Siglec-F
9.98%
Eos
89.7%
CD19
B
39.6%
CD3
T
23.0%
CD8
CD8
12.7%
CD4
80.0%
CD4
CD25
Treg
10.7%
CD4

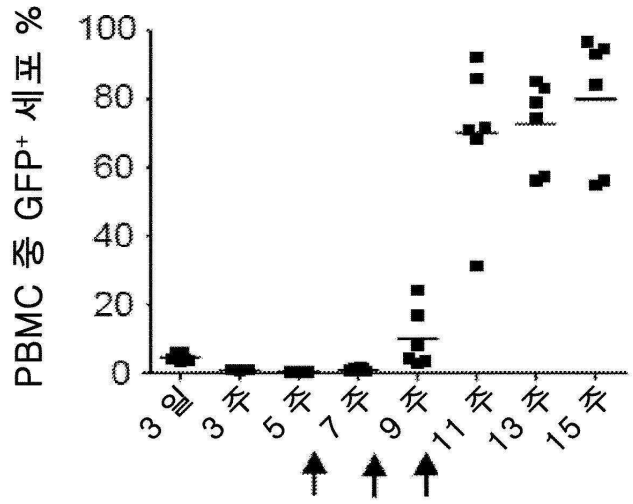
도면77


비장
골수
거동 중 GFP+ 세포 %
100
80
60
40
20
0
T 세포
CD8+ T 세포
CD4+ T 세포
B 세포
호중구
DC
Ly6C(hi) 단핵구
Ly6C(lo) 단핵구
대식세포
호산구
CD3
CD19
Gr-1
LSK

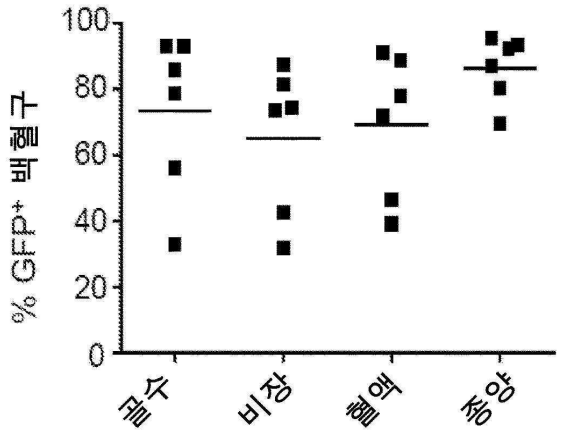
도면78a


miRNASeq
GFP+ 분획
-종양
-골수
-비장
-PBMC
22 주
17 주
TC-1 종양 세포
(s.c.)
6 주
8 주
10 주
O6BG
BCNU
(i.p.)
G-CSF/AMD3100
(s.c.)
HDAd
(i.v.)
CD46tg

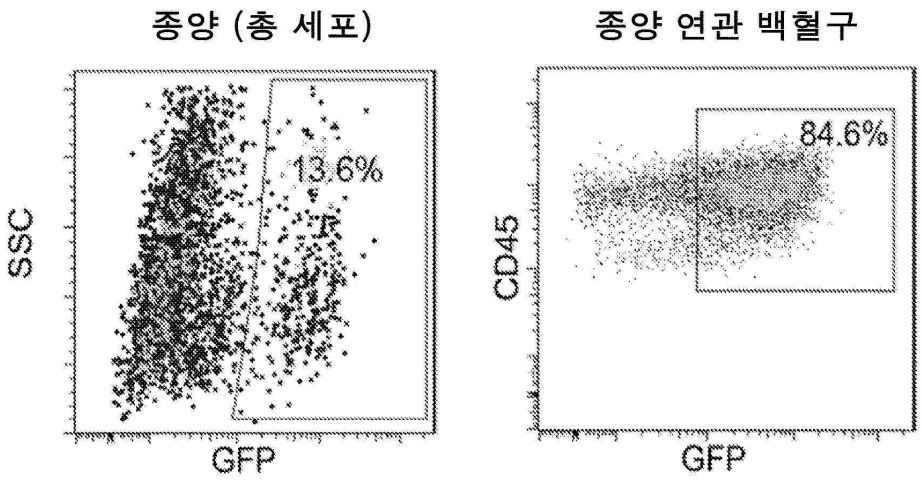
도면78b



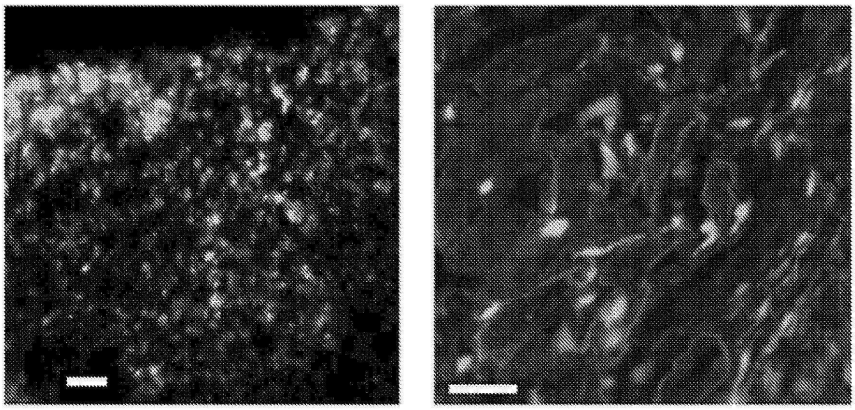

도면78c



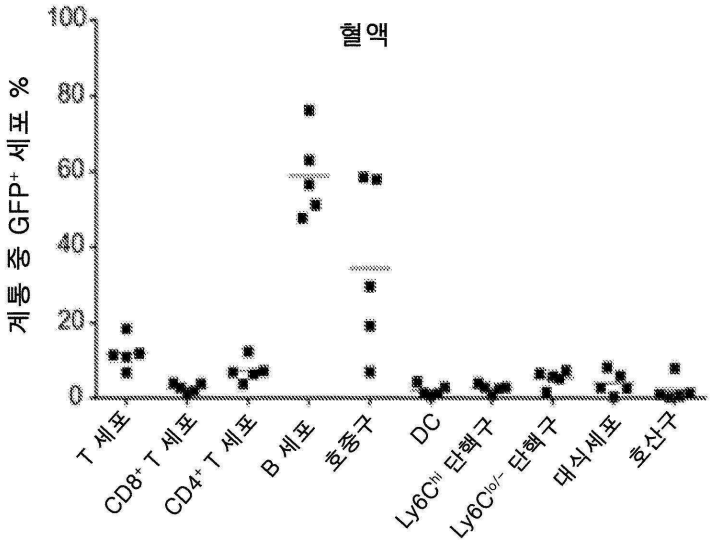

도면78d

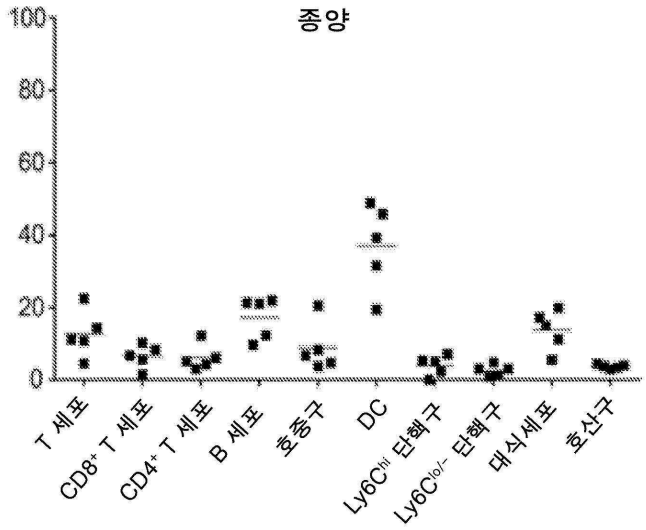


도면78e

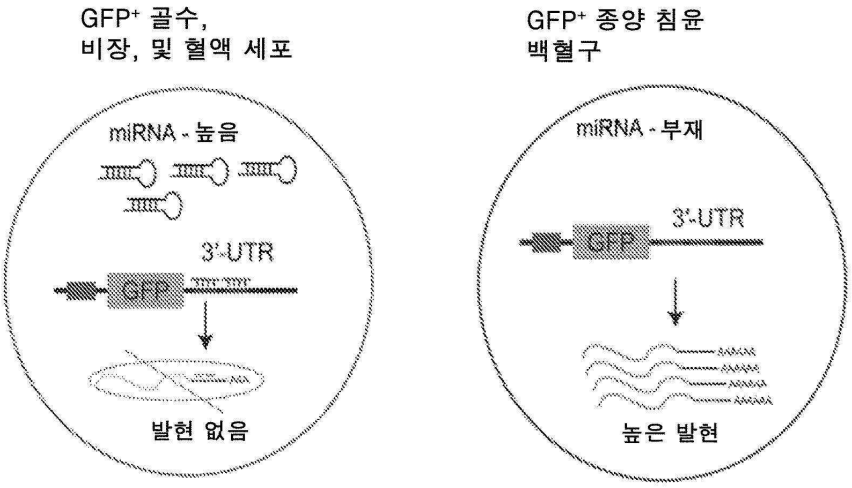

GFP
GFP/라미닌

도면78f

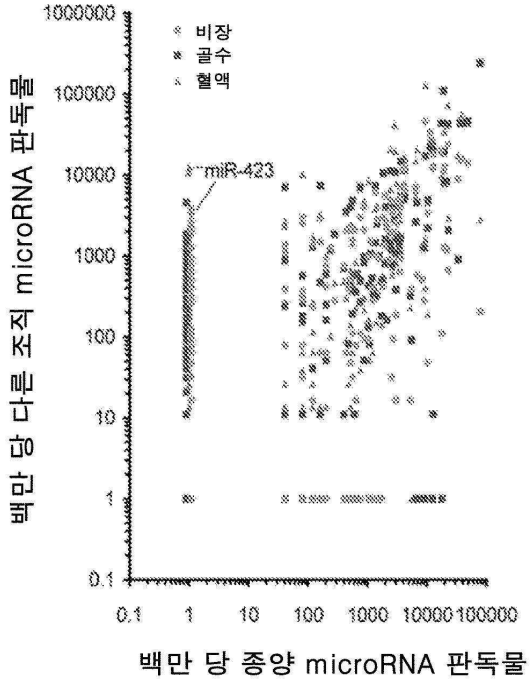

혈액
계통 중 GFP+ 세포 %
100
80
60
40
20
0
T 세포
CD8+ T 세포
CD4+ T 세포
B 세포
호중구
DC
Ly6Chi 단핵구
Ly6Clo/- 단핵구
대식세포
호산구


종양
100
80
60
40
20
0
T 세포
CD8+ T 세포
CD4+ T 세포
B 세포
호중구
DC
Ly6Chi 단핵구
Ly6Clo/- 단핵구
대식세포
호산구

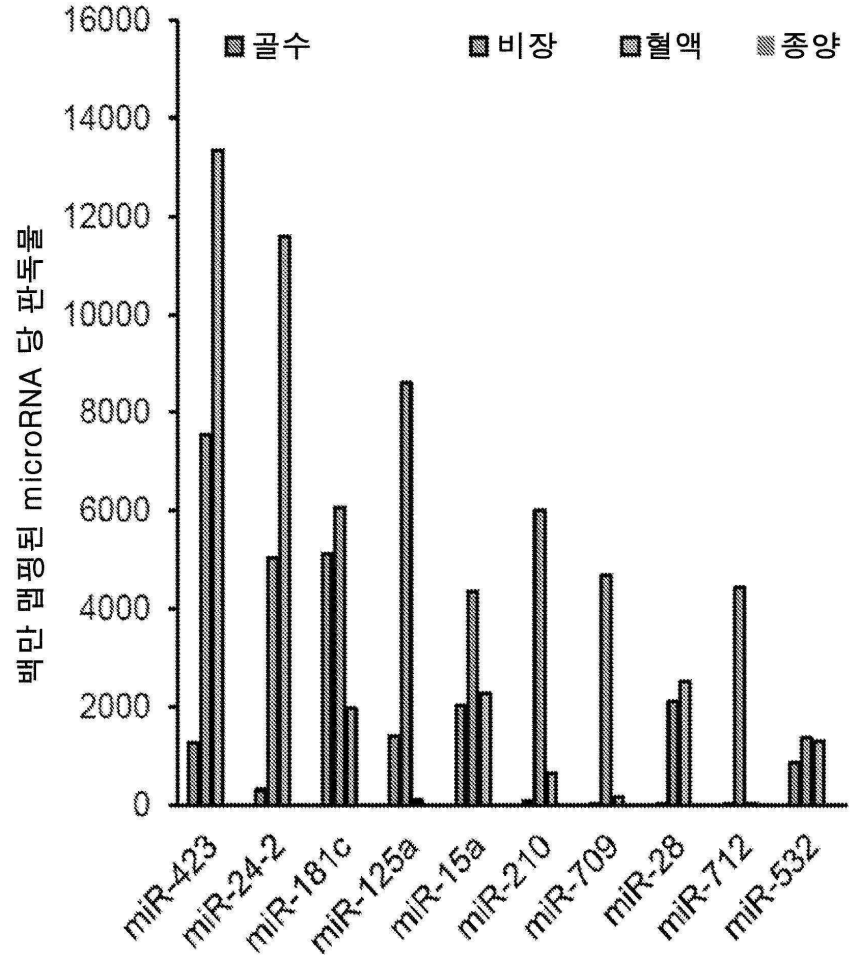
도면79a


GFP+ 골수,
비장, 및 혈액 세포
GFP+ 종양 침윤
백혈구
miRNA - 높음
3'-UTR
GFP
발현 없음
miRNA - 부재
3'-UTR
GFP
높은 발현

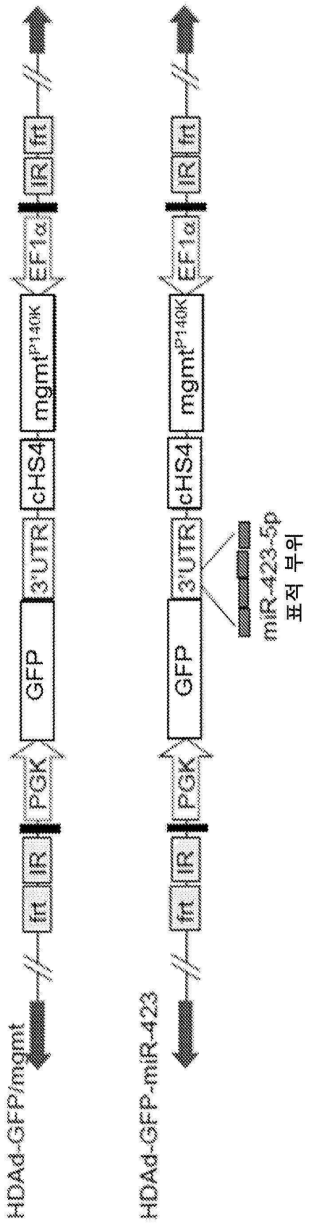
도면79b


비장
골수
혈액
miR-423
백만 당 다른 조직 microRNA 판독물
백만 당 종양 microRNA 판독물
1000000
100000
10000
1000
100
10
1
0.1
0.1
1
10
100
1000
10000
100000

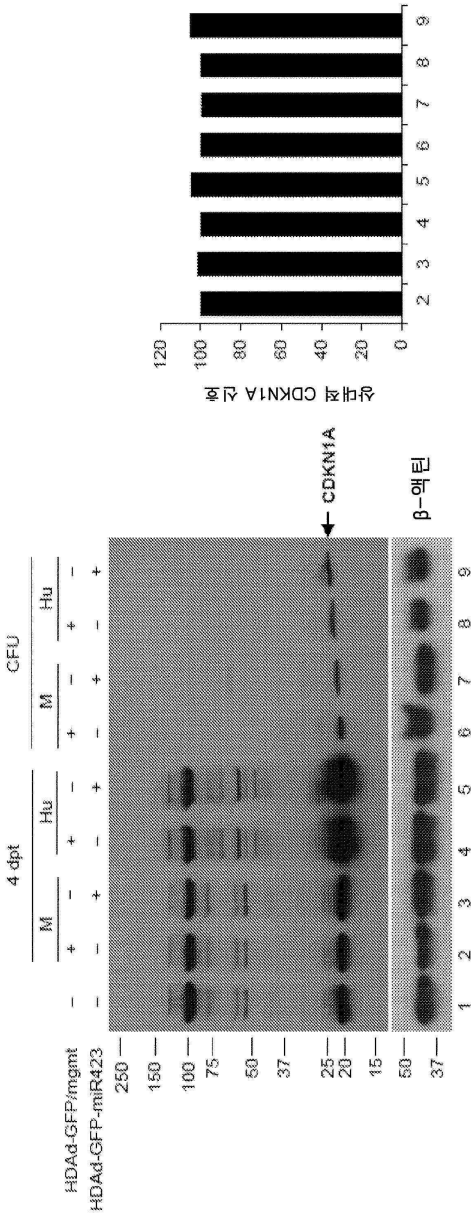
도면79c


16000
14000
12000
10000
8000
6000
4000
2000
0
백만 맵핑된 microRNA 당 판독물
골수
비장
혈액
종양
miR-423
miR-24-2
miR-181c
miR-125a
miR-15a
miR-210
miR-709
miR-28
miR-712
miR-532

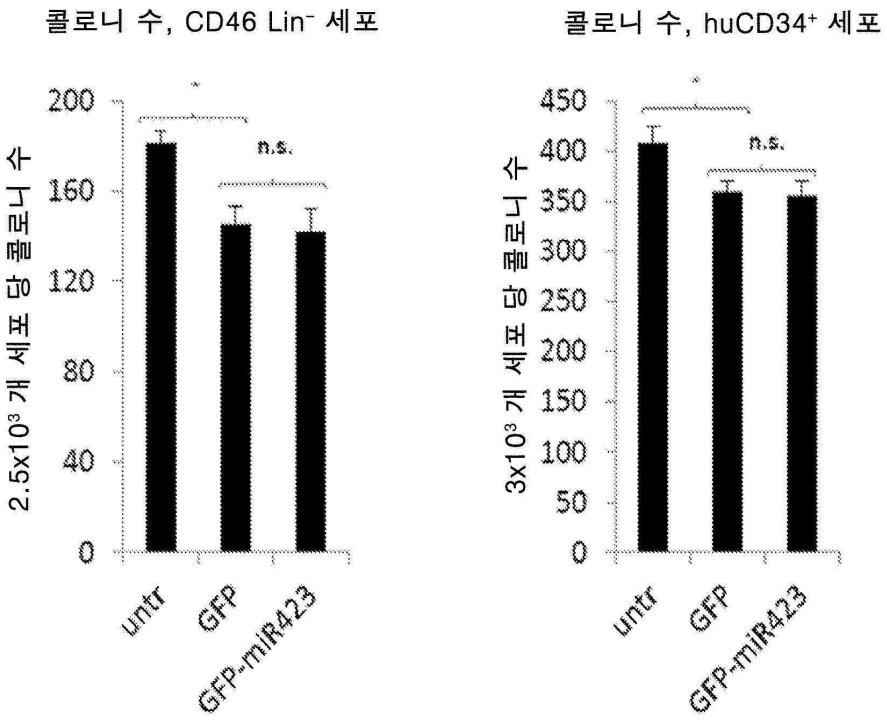
도면80a


HDAd-GFP/mgmt
frt
IR
PGK
GFP
3'UTR
cHS4
mgmt^P140K
EF1α
IR
frt
HDAd-GFP-miR-423
frt
IR
PGK
GFP
3'UTR
miR-423-5p
표적 부위
cHS4
mgmt^P140K
EF1α
IR
frt

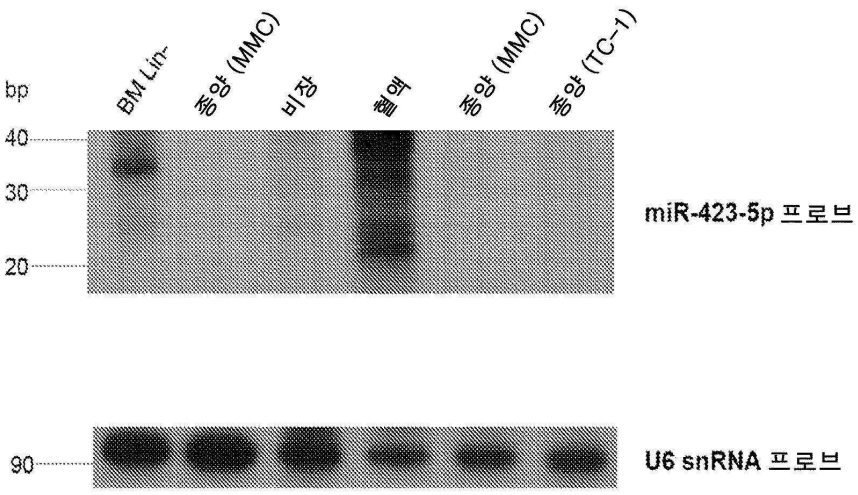
도면80b


상대적 CDKN1A 신호
120
100
80
60
40
20
0
1 2 3 4 5 6 7 8 9
CDKN1A
β-액틴
4 dpt
CFU
M
Hu
HDAd-GFP/mgmt
HDAd-GFP-miR423
250
150
100
75
50
37
25
20
15
50
37

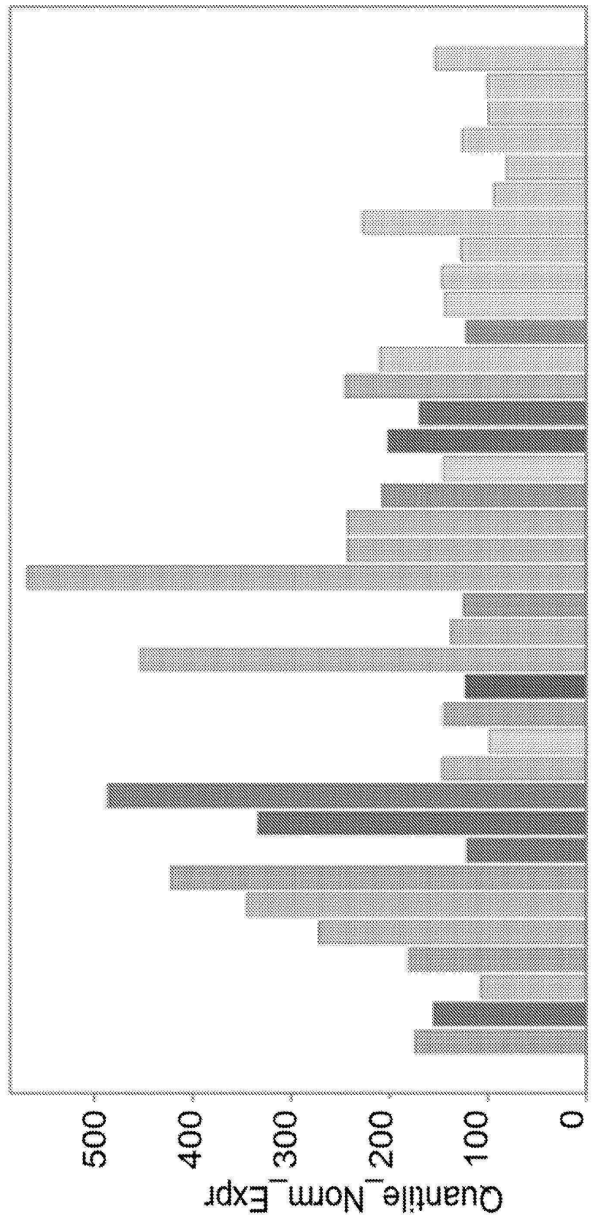
도면80c


콜로니 수, CD46 Lin⁻ 세포
2.5x10³ 개 세포 당 콜로니 수
200
160
120
80
40
0
n.s.
untr
GFP
GFP-miR423
콜로니 수, huCD34⁺ 세포
3x10³ 개 세포 당 콜로니 수
450
400
350
300
250
200
150
100
50
0
n.s.
untr
GFP
GFP-miR423

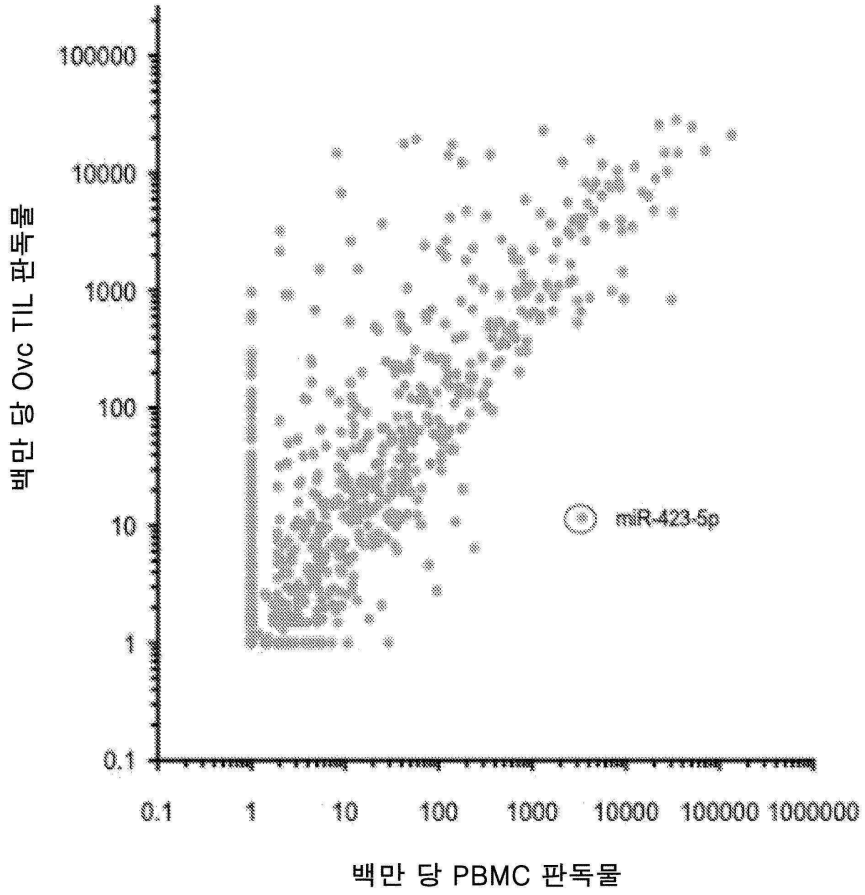
도면81


bp
BM Lin-
중앙 (MMC)
비장
혈액
중앙 (MMC)
중앙 (TC-1)
40
30
20
miR-423-5p 프로브
90
U6 snRNA 프로브

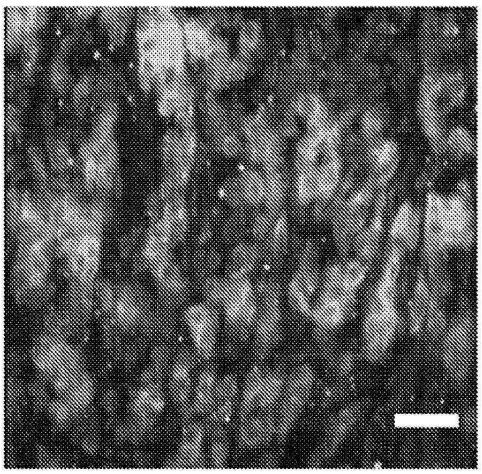
도면82a



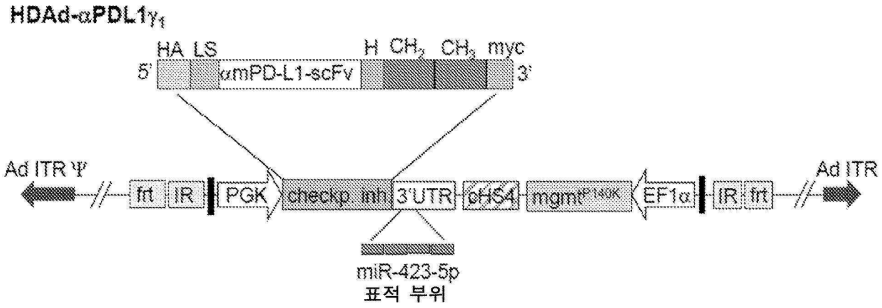
도면82b


100000
10000
1000
100
10
1
0.1
백만 당 Ovc TIL 판독물
miR-423-5p
0.1
1
10
100
1000
10000
100000
1000000
백만 당 PBMC 판독물

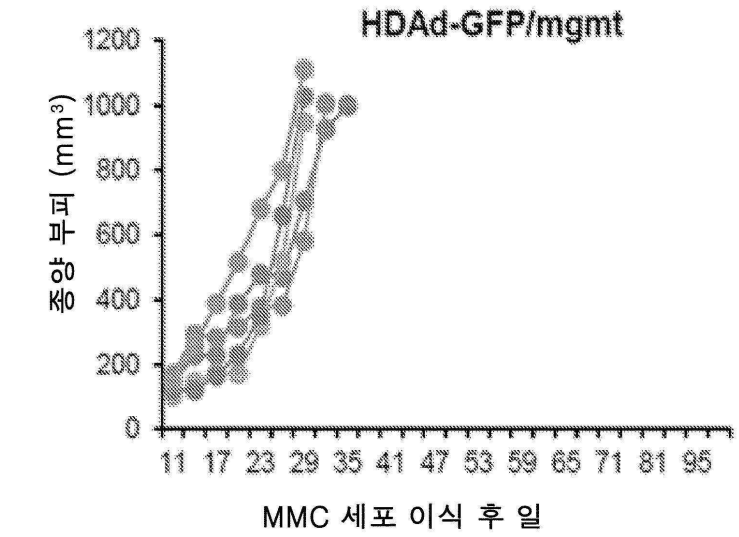
도면83a

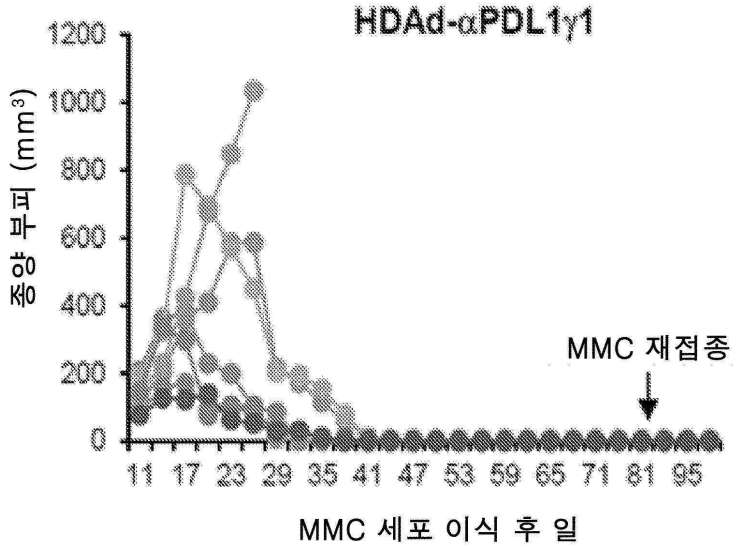


도면83b

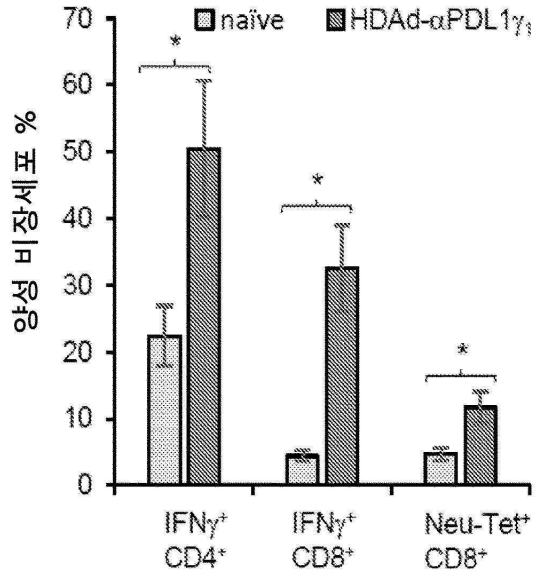

HDAd-αPDL1γ1
HA LS
H CH2 CH3 myc
5'
αmPD-L1-scFv
3'
Ad ITR Ψ
frt
IR
PGK
checkp. inh.
3'UTR
cHS4
mgmtP140K
EF1α
IR
frt
Ad ITR
miR-423-5p
표적 부위

도면83c

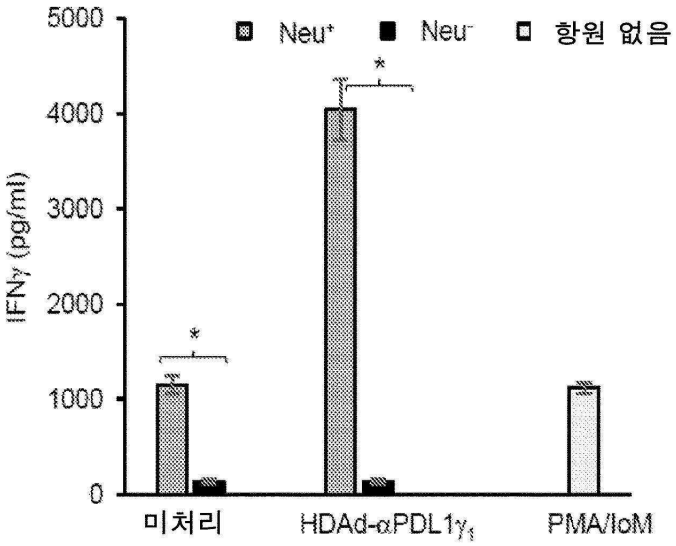

HDAd-GFP/mgmt
종양 부피 (mm3)
1200
1000
800
600
400
200
0
11 17 23 29 35 41 47 53 59 65 71 81 95
MMC 세포 이식 후 일


HDAd-αPDL1γ1
종양 부피 (mm3)
1200
1000
800
600
400
200
0
MMC 재접종
11 17 23 29 35 41 47 53 59 65 71 81 95
MMC 세포 이식 후 일

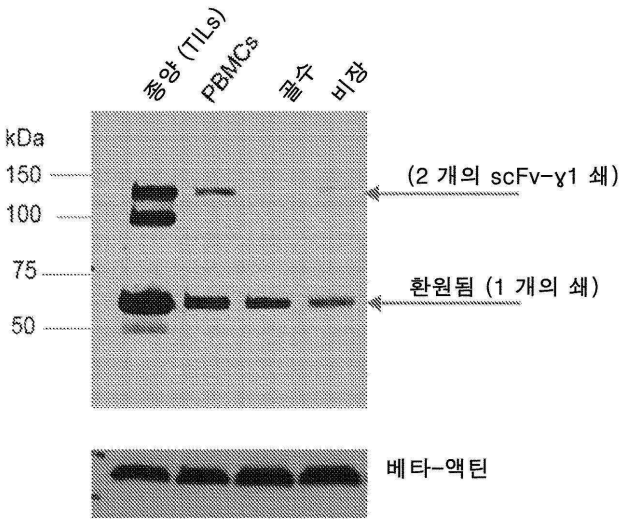
도면83d

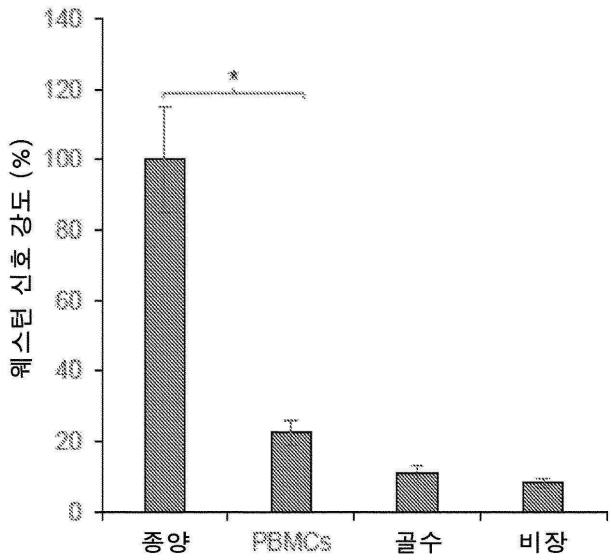

naïve
HDAd-αPDL1γ1
양성 비장세포 %
70
60
50
40
30
20
10
0
*
*
*
IFNγ+ CD4+
IFNγ+ CD8+
Neu-Tet+ CD8+

도면83e

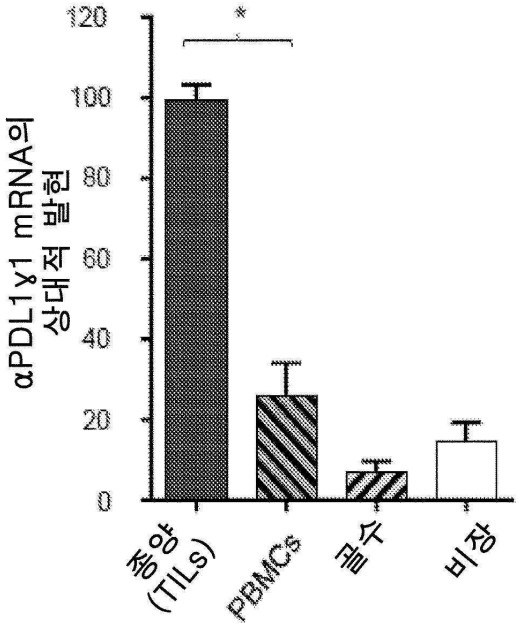

Neu+
Neu-
항원 없음
IFNγ (pg/ml)
5000
4000
3000
2000
1000
0
*
*
미처리
HDAd-αPDL1γ1
PMA/IoM

도면84a

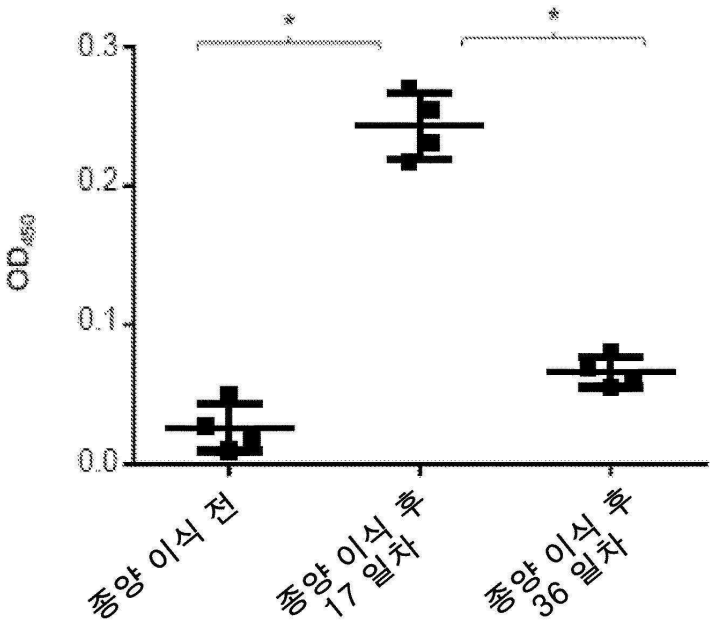

종양 (TILs)
PBMCs
골수
비장
kDa
150
100
75
50
(2 개의 scFv-γ1 쇄)
환원됨 (1 개의 쇄)
베타-액틴


헤스틴 신호 강도 (%)
140
120
100
80
60
40
20
0
*
종양
PBMCs
골수
비장

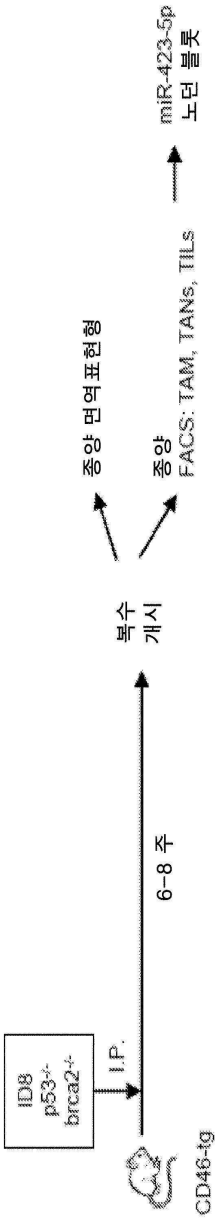
도면84b


αPDL1γ1 mRNA의 상대적 발현
120
100
80
60
40
20
0
*
종양 (TILs)
PBMCs
골수
비장

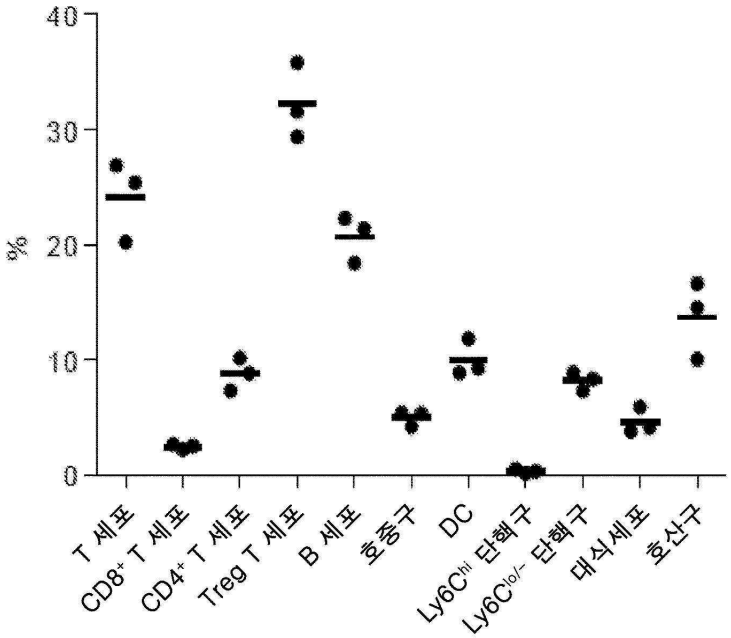
도면84c


OD450
0.3
0.2
0.1
0.0
*
*
종양 이식 전
종양 이식 후 17 일차
종양 이식 후 36 일차

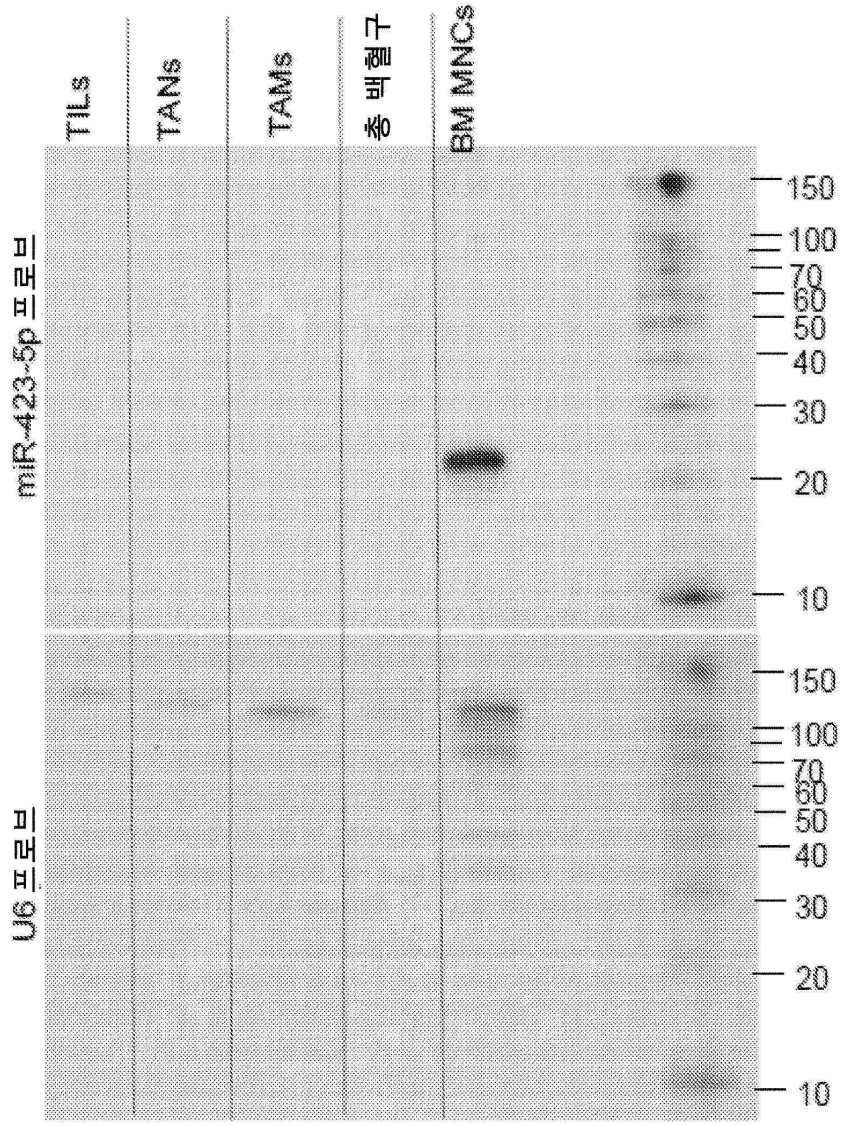
도면85a


ID8
p53-/-
brca2-/-
I.P.
6-8 주
CD46-tg
복수
게시
종양 면역표현형
종양
FACS: TAM, TANs, TILs
miR-423-5p
노던 블롯

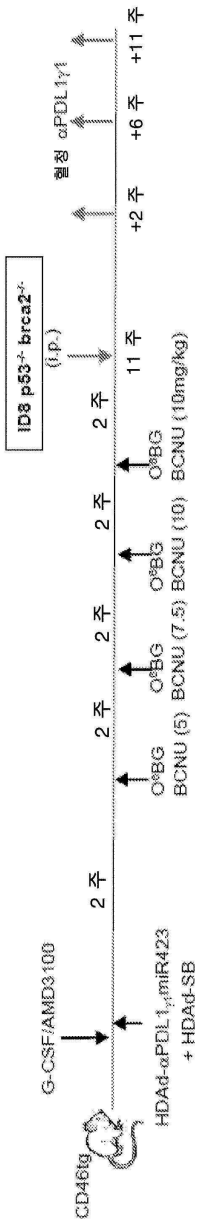
도면85b


40
30
20
10
0
%
T 세포
CD8+ T 세포
CD4+ T 세포
Treg T 세포
B 세포
호중구
DC
Ly6Chi 단핵구
Ly6Clo/- 단핵구
대식세포
호산구

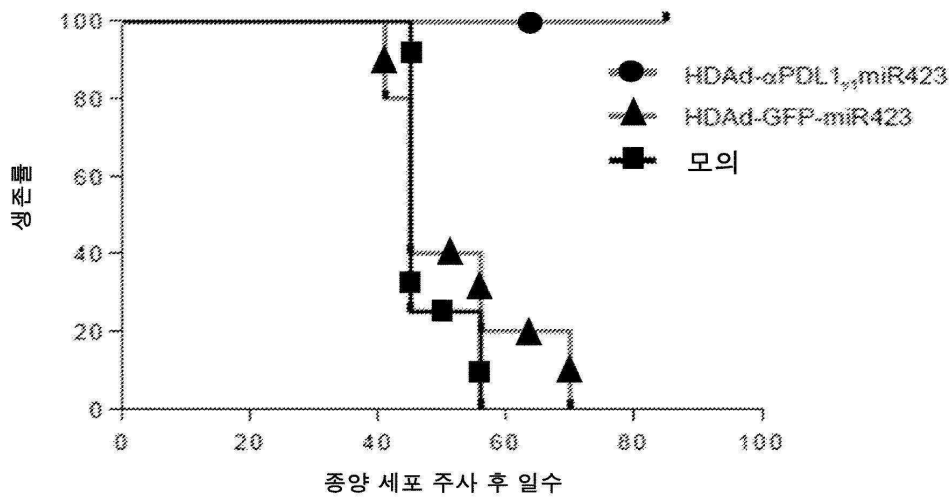
도면85c


TILs
TANs
TAMs
중 백혈구
BM MNCs
miR-423-5p 프로브
U6 프로브
150
100
70
60
50
40
30
20
10

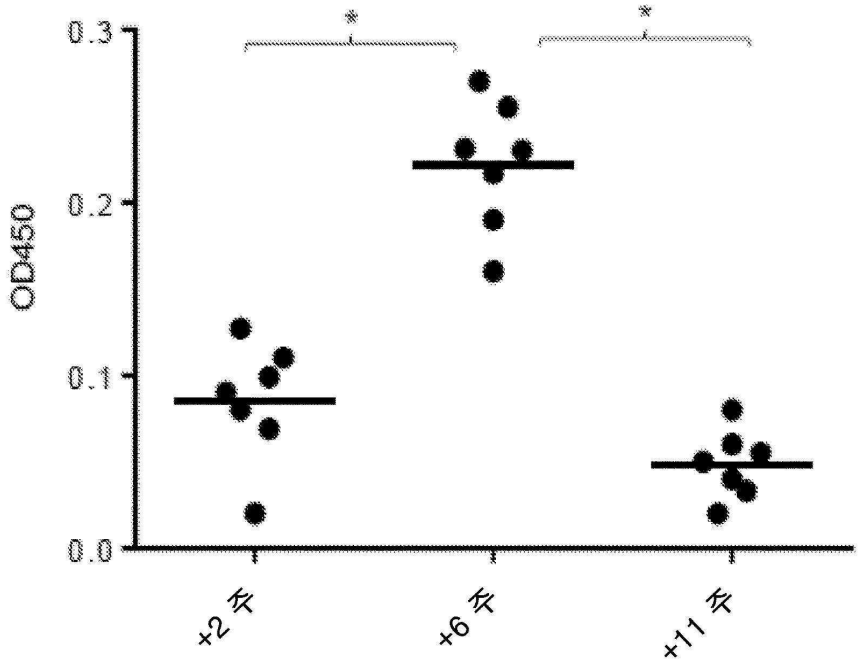
도면85d


CD46tg
G-CSF/AMD3100
HDAd-αPDL1γ1miR423
+ HDAd-SB
2 주
O⁶BG
BCNU (5)
BCNU (7.5)
BCNU (10)
BCNU (10mg/kg)
ID8 p53-/- brca2-/-
(i.p.)
11 주
혈청 αPDL1γ1
+2 주
+6 주
+11 주

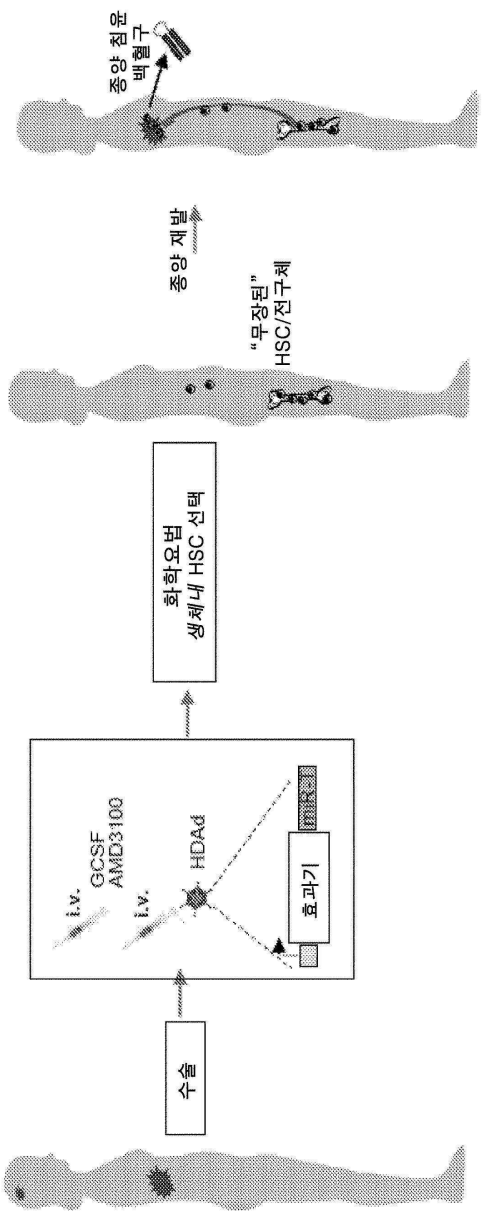
도면85e


생존률
종양 세포 주사 후 일수
HDAd-αPDL1γ1miR423
HDAd-GFP-miR423
모의
0
20
40
60
80
100

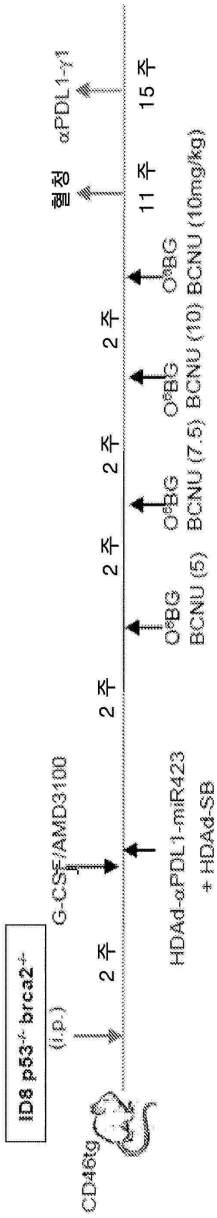
도면85f



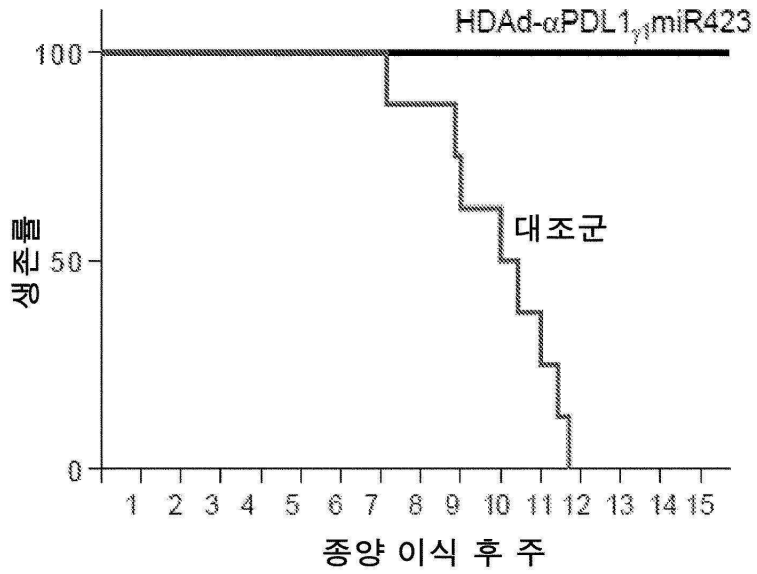
도면86a


종양 침윤
백혈구
종양 재발
"무장된"
HSC/전구체
화학요법
생체내 HSC 선택
i.v.
GCSF
AMD3100
i.v.
HDAd
miR-21
효과기
수술

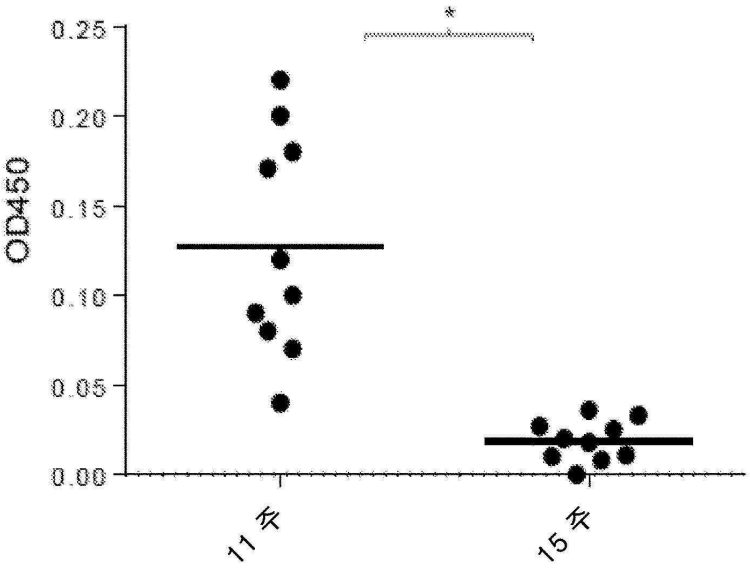
도면86b


ID8 p53-/- brca2-/-
(i.p.)
CD46tg
2 주
G-CSF/AMD3100
HDAd-αPDL1-miR423
+ HDAd-SB
2 주
O6BG
BCNU (5)
2 주
O6BG
BCNU (7.5)
2 주
O6BG
BCNU (10)
2 주
O6BG
BCNU (10mg/kg)
11 주
희생
αPDL1-γ1
15 주

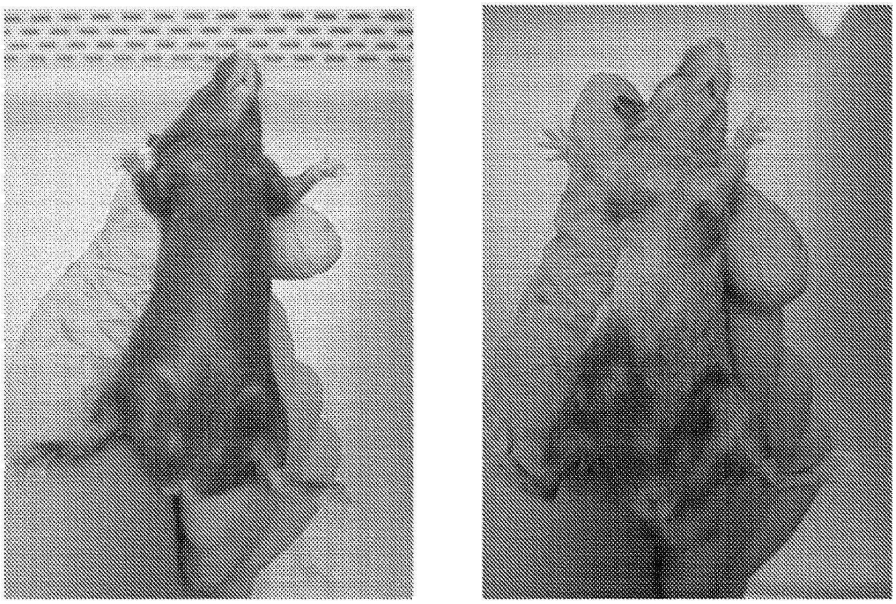
도면86c


HDAd-αPDL1γ1miR423
대조군
생존율
100
50
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
종양 이식 후 주

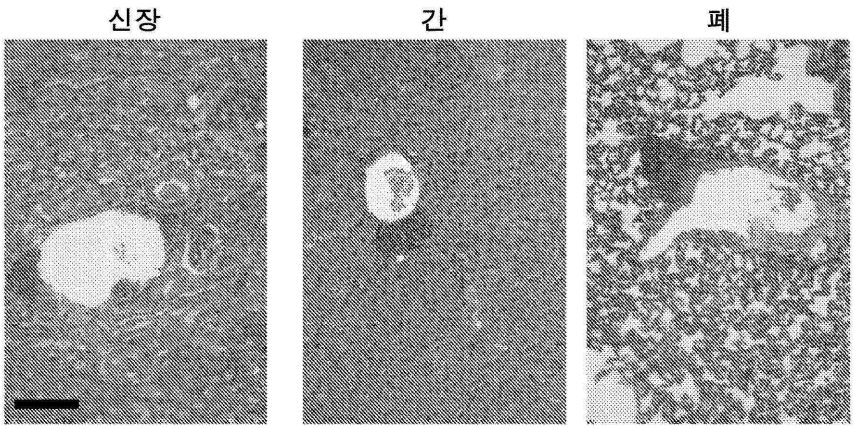
도면86d


*
OD450
0.25
0.20
0.15
0.10
0.05
0.00
11 주
15 주

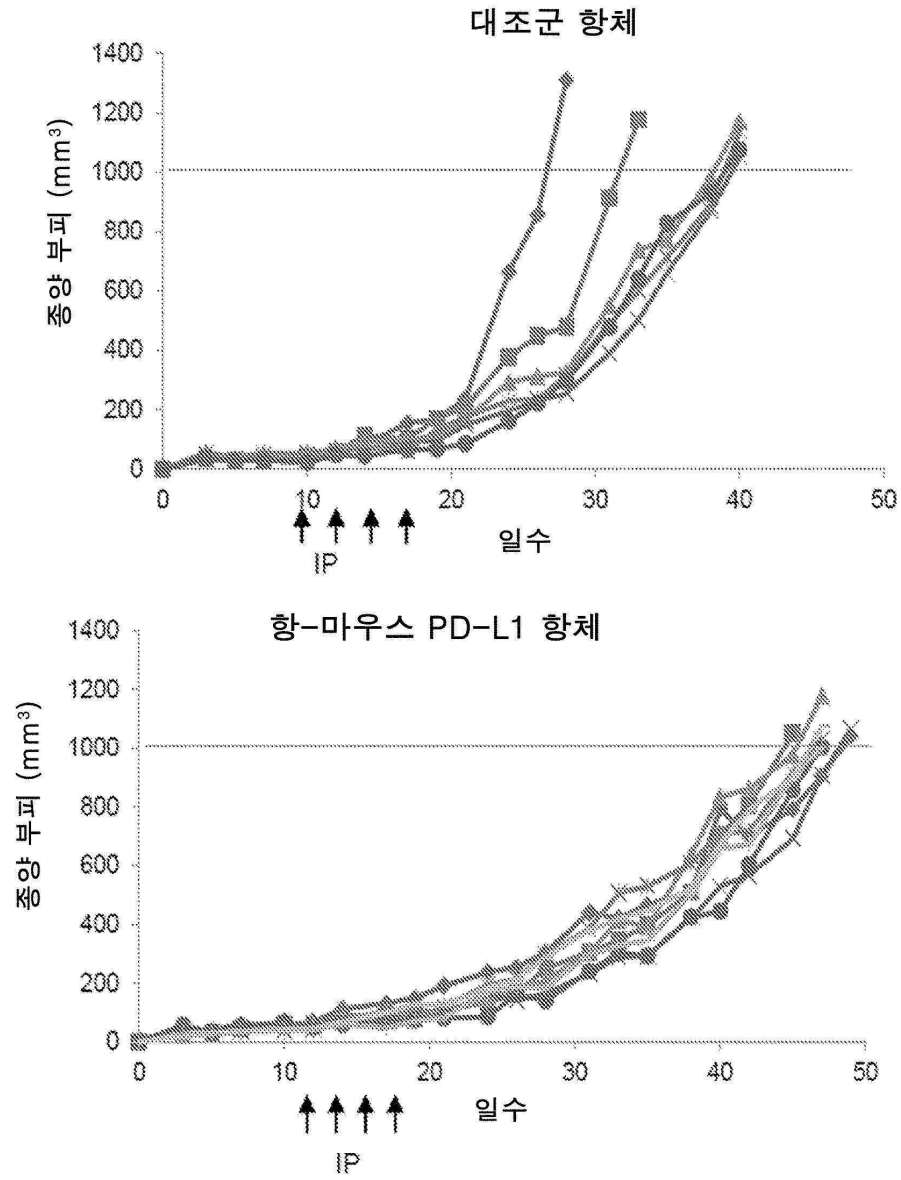
도면87a



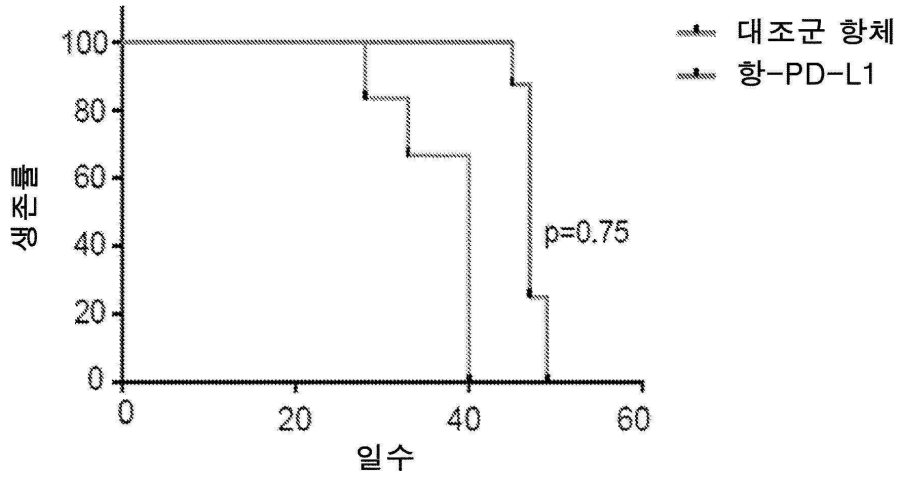
도면87b



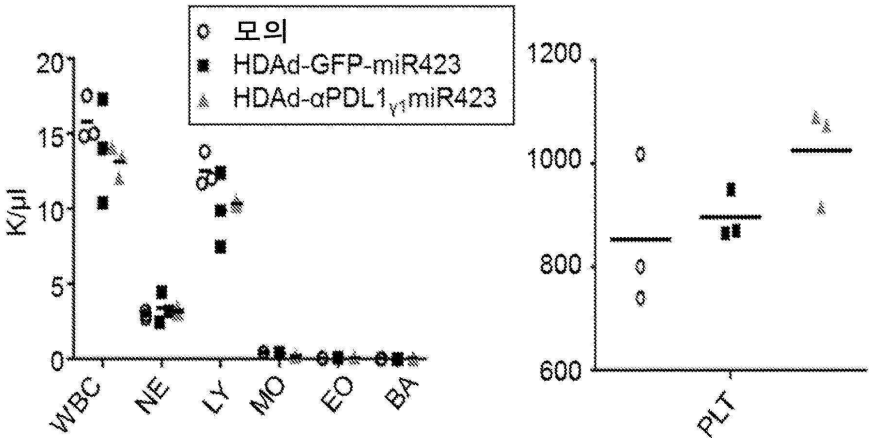
도면88a


대조군 항체
종양 부피 (mm³)
1400
1200
1000
800
600
400
200
0
0
10
20
30
40
50
일수
IP
항-마우스 PD-L1 항체
종양 부피 (mm³)
1400
1200
1000
800
600
400
200
0
0
10
20
30
40
50
일수
IP

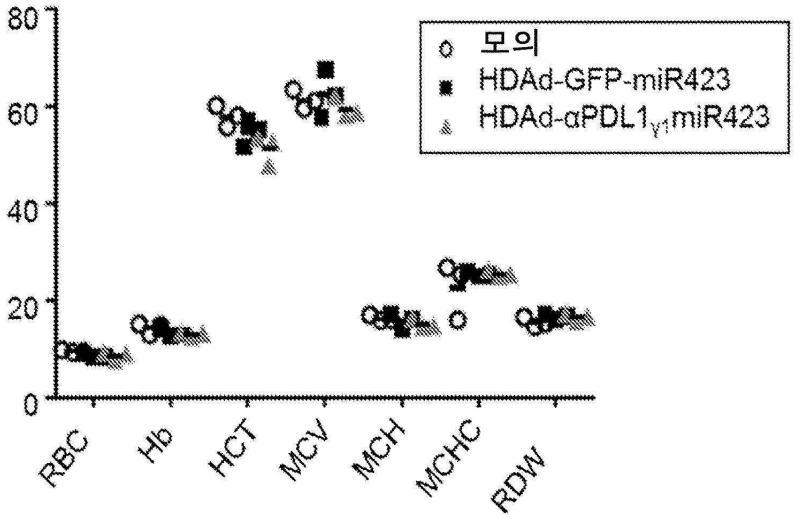
도면88b


대조군 항체
항-PD-L1
생존률
100
80
60
40
20
0
p=0.75
0
20
40
60
일수

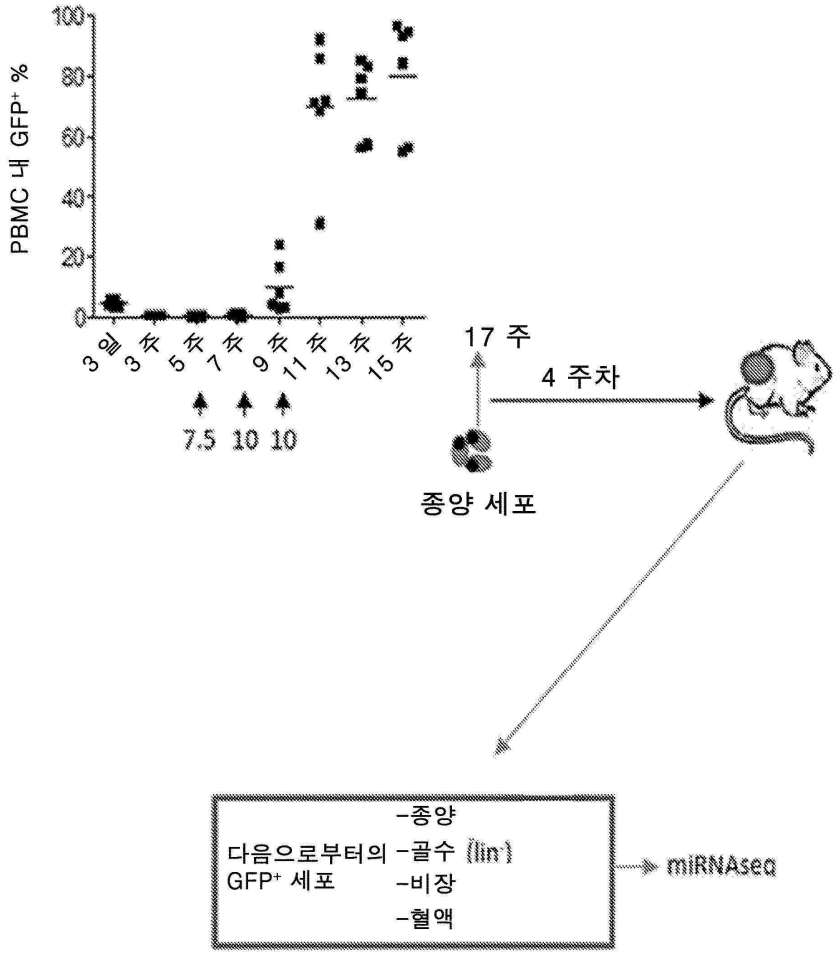
도면88c


모의
HDAd-GFP-miR423
HDAd-αPDL1$_{Y1}$miR423
K/μl
20
15
10
5
0
WBC
NE
LY
MO
EO
BA
1200
1000
800
600
PLT

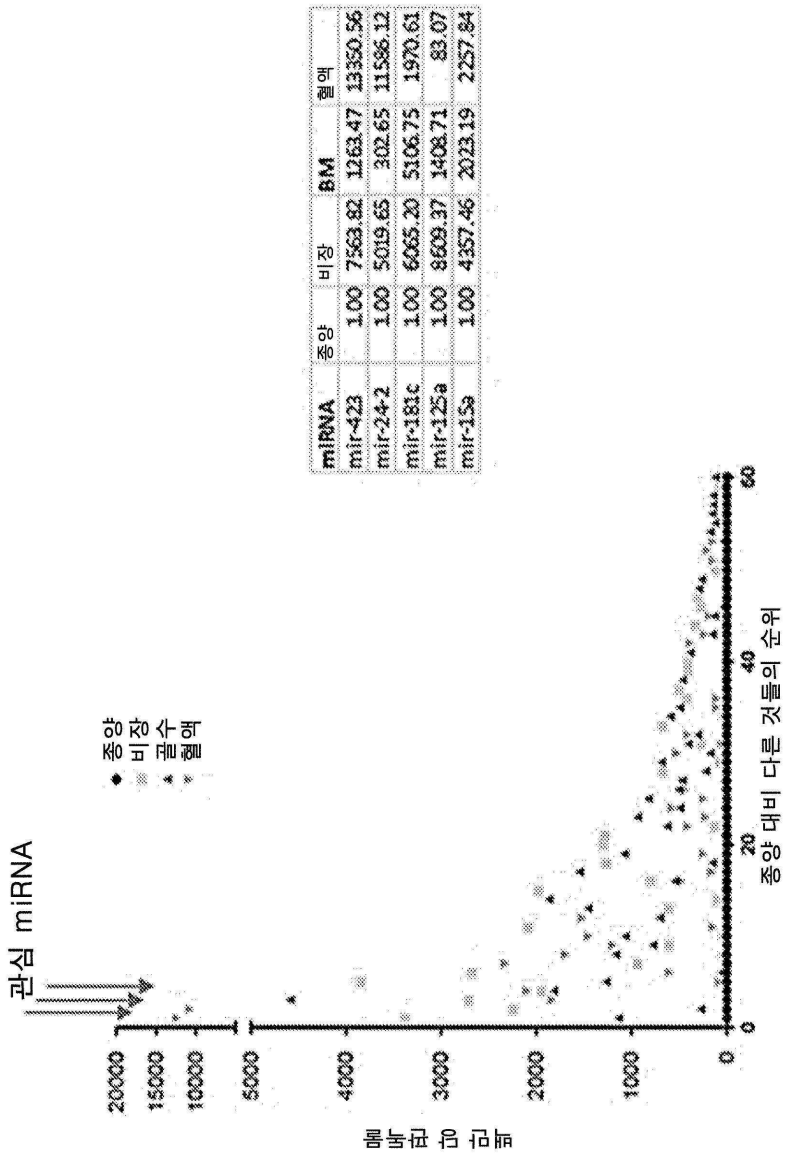
도면88d


모의
HDAd-GFP-miR423
HDAd-αPDL1$_{Y1}$miR423
80
60
40
20
0
RBC
Hb
HCT
MCV
MCH
MCHC
RDW

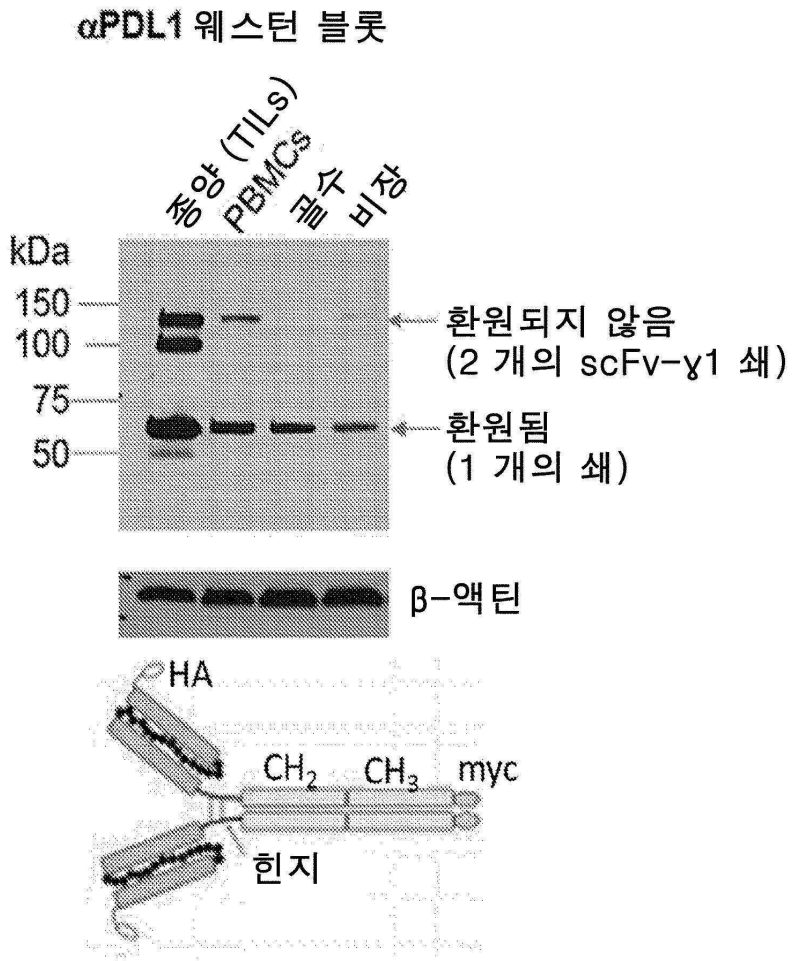
도면88ea


PBMC 내 GFP+ %
100
80
60
40
20
0
3 일
3 주
5 주
7 주
9 주
11 주
13 주
15 주
7.5
10
10
17 주
4 주차
종양 세포
다음으로부터의 GFP+ 세포
-종양
-골수 (lin-)
-비장
-혈액
miRNAseq

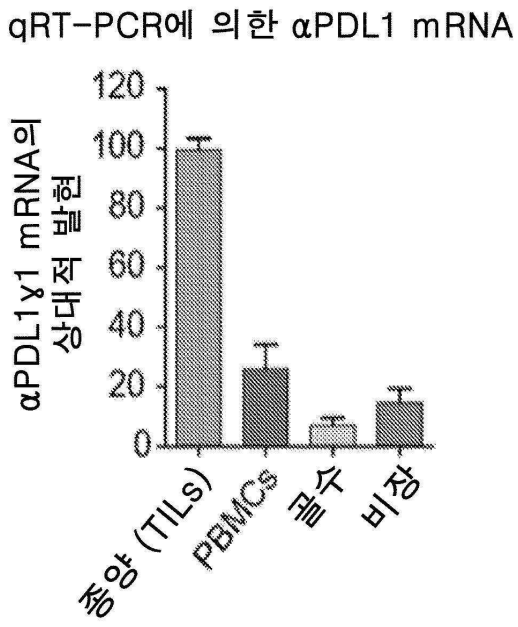
도면88eb

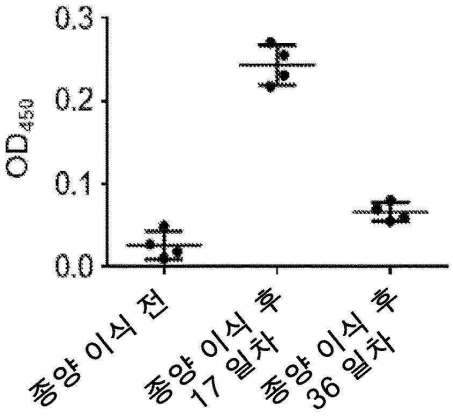

miRNA 중앙 비장 BM 혈액
mir-423 1.00 7563.82 1263.47 13350.56
mir-24-2 1.00 5019.65 302.65 11586.12
mir-181c 1.00 6065.20 5106.75 1970.61
mir-125a 1.00 8609.37 1408.71 83.07
mir-15a 1.00 4357.46 2023.19 2257.84
중앙 대비 다른 것들의 순위
관심 miRNA
중앙 비장 혈액
백만 당 판독물

도면88fa

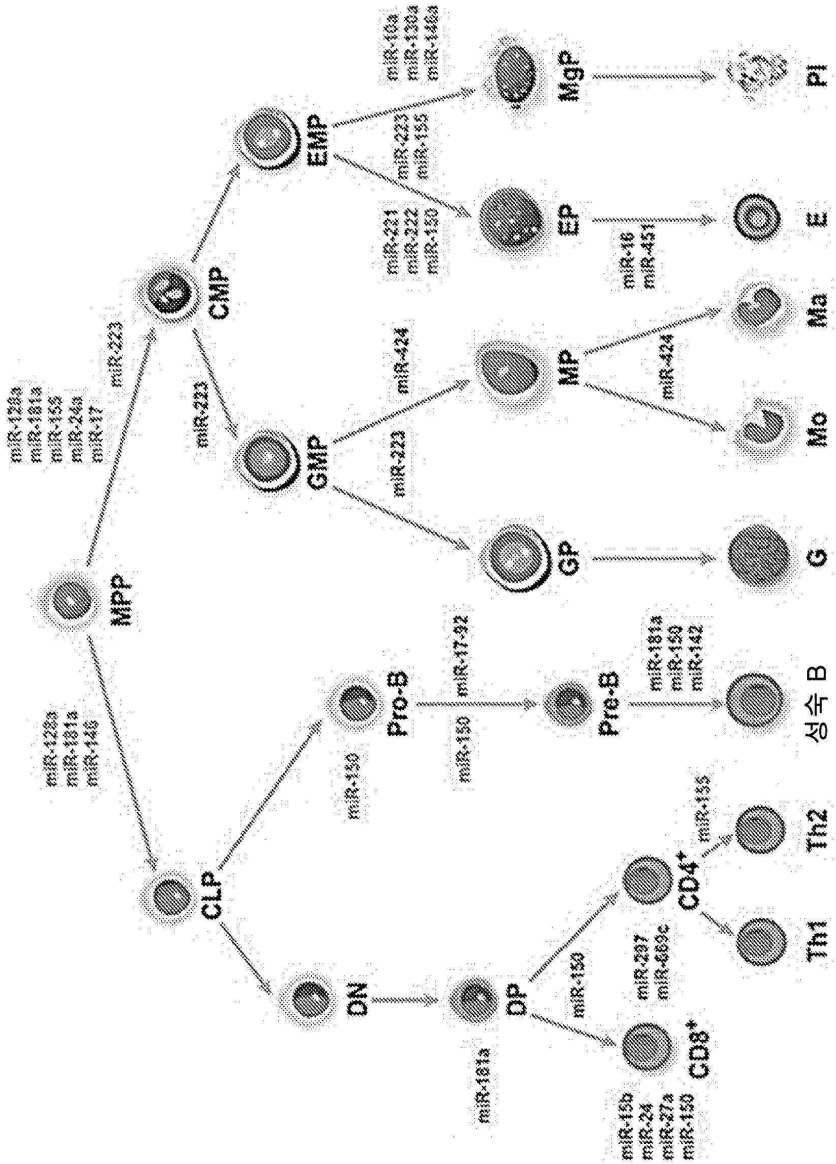

αPDL1 웨스턴 블롯
종양 (TILs)
PBMCs
골수
비장
kDa
150
100
75
50
환원되지 않음
(2 개의 scFv-γ1 쇄)
환원됨
(1 개의 쇄)
β-액틴
HA
CH2
CH3
myc
힌지

도면88fb

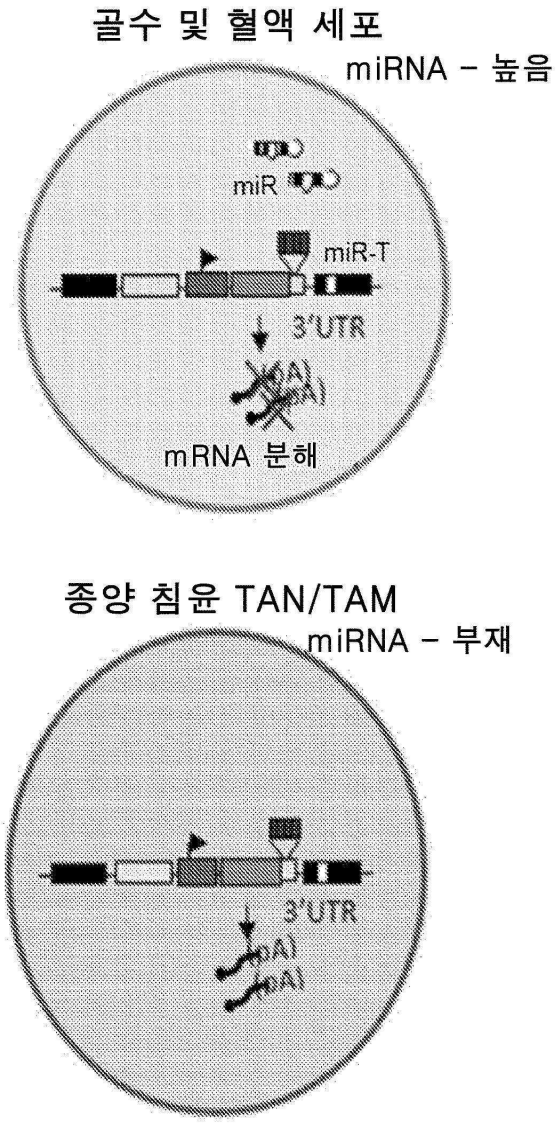

qRT-PCR에 의한 αPDL1 mRNA
αPDL1γ1 mRNA의 상대적 발현
120
100
80
60
40
20
0
종양 (TILs)
PBMCs
골수
비장


혈청 αPDL1 ELISA
OD450
0.3
0.2
0.1
0.0
종양 이식 전
종양 이식 후 17 일차
종양 이식 후 36 일차

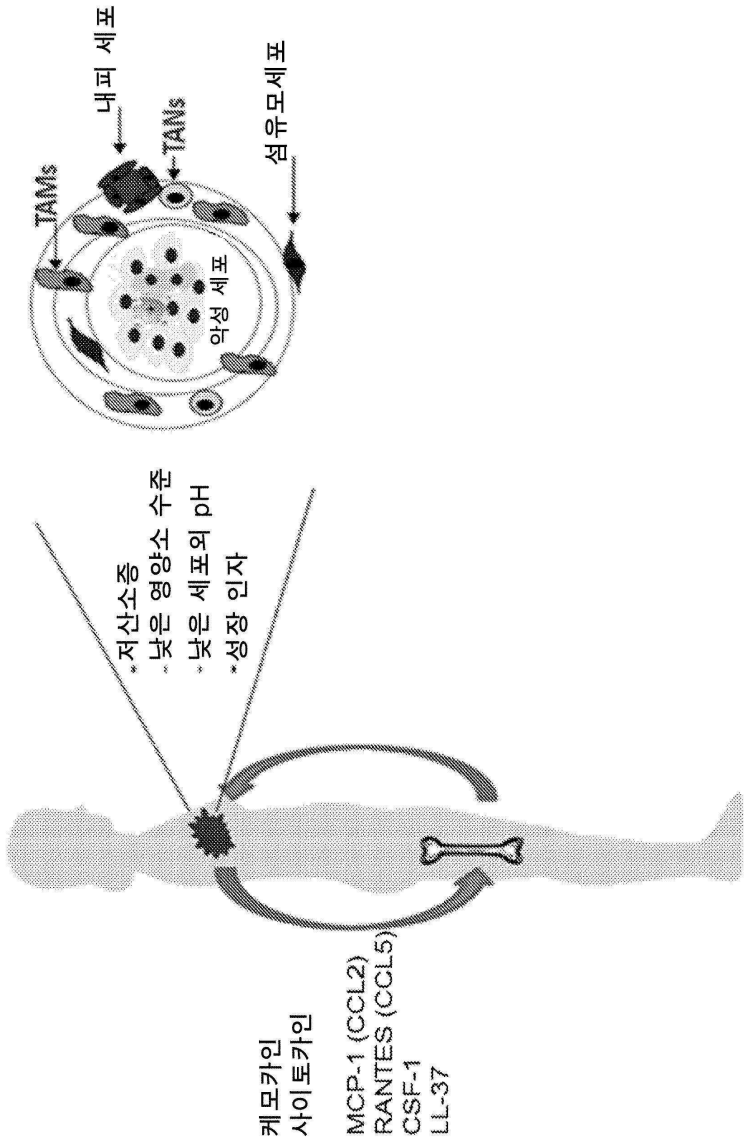
도면88ga


miR-10a
miR-130a
miR-146a
MgP
Pl
miR-223
miR-155
EMP
miR-221
miR-222
miR-150
EP
E
miR-16
miR-451
miR-128a
miR-181a
miR-155
miR-24a
miR-17
miR-223
CMP
miR-223
miR-424
MP
Ma
miR-424
GMP
miR-223
Mo
GP
G
MPP
miR-128a
miR-181a
miR-146
miR-150
Pro-B
miR-17-92
miR-150
Pre-B
miR-181a
miR-150
miR-142
성숙 B
CLP
CD4+
miR-155
Th2
Th1
miR-150
DN
DP
miR-297
miR-669c
miR-181a
CD8+
miR-15b
miR-24
miR-27a
miR-150

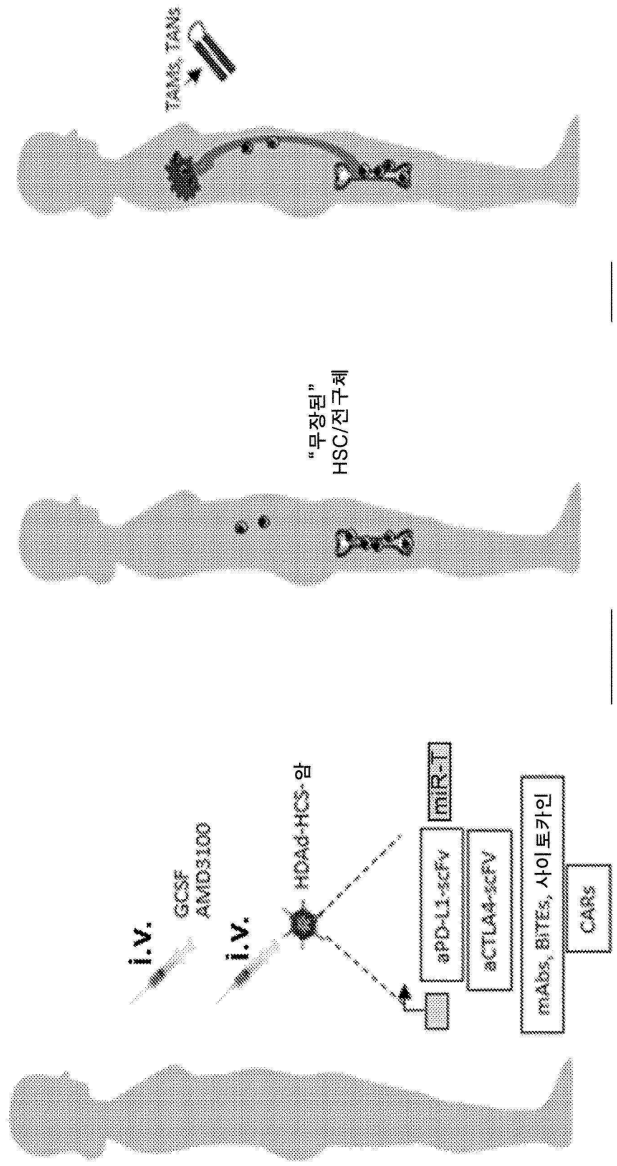
도면88gb


골수 및 혈액 세포
miRNA - 높음
miR
miR-T
3'UTR
(A)n
mRNA 분해
종양 침윤 TAN/TAM
miRNA - 부재
3'UTR
(A)n

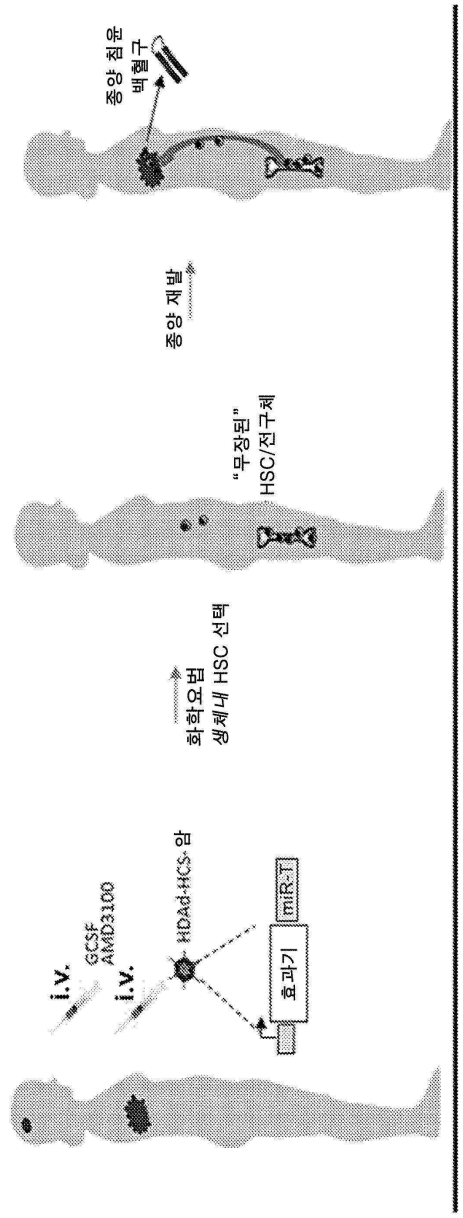
도면88ha


내피 세포
TANs
섬유모세포
TAMs
악성 세포
•저산소증
•낮은 영양소 수준
•낮은 세포외 pH
•성장 인자
케모카인
사이토카인
MCP-1 (CCL2)
RANTES (CCL5)
CSF-1
LL-37

도면88hb


TAMs, TANs
"무장된"
HSC/전구체
i.v.
GCSF
AMD3100
i.v.
HDAd-HCS-암
miR-T
aPD-L1-scFv
aCTLA4-scFv
mAbs, BiTEs, 사이토카인
CARs

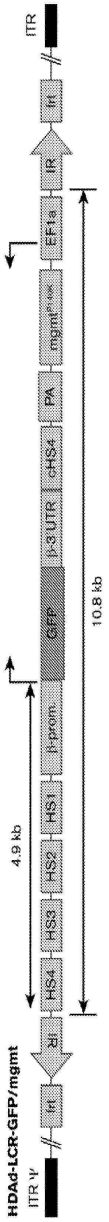
도면88hc



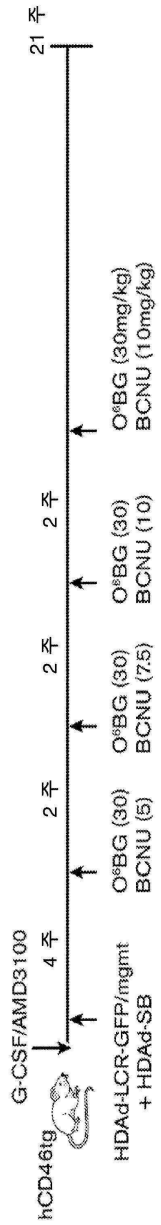

생검:
→ 알려진 면역 표현형 → 최적의 면역-효과기 유전자 선택
→ 알려진 신생-항원 → 개인맞춤형 CAR/BiTE

-본격적인 종양 치료 어려움

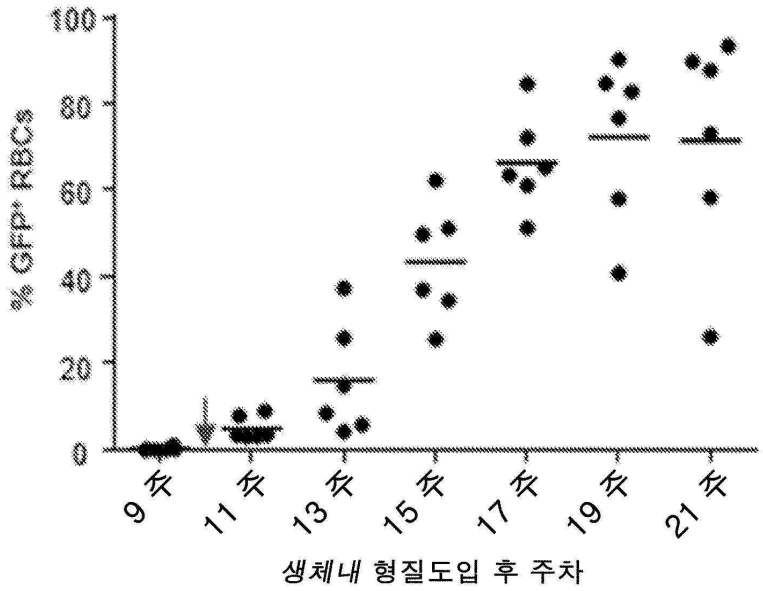
도면89a


HDAd-LCR-GFP/mgmt
ITR Ψ
Int
attI
HS4
HS3
HS2
HS1
β-prom.
GFP
β-3'UTR
cHS4
PA
mgmt$^{P140K}$
EF1α
IR
Int
ITR
4.9 kb
10.8 kb

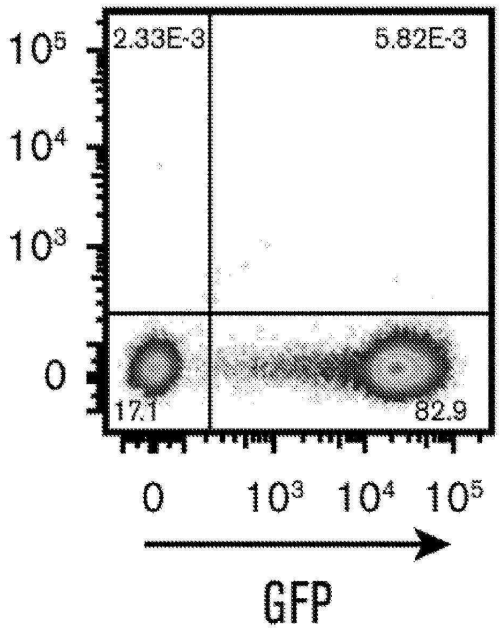
도면89b


21 주
2 주
2 주
2 주
4 주
G-CSF/AMD3100
hCD46tg
HDAd-LCR-GFP/mgmt
+ HDAd-SB
O6BG (30)
BCNU (5)
O6BG (30)
BCNU (7.5)
O6BG (30)
BCNU (10)
O6BG (30mg/kg)
BCNU (10mg/kg)

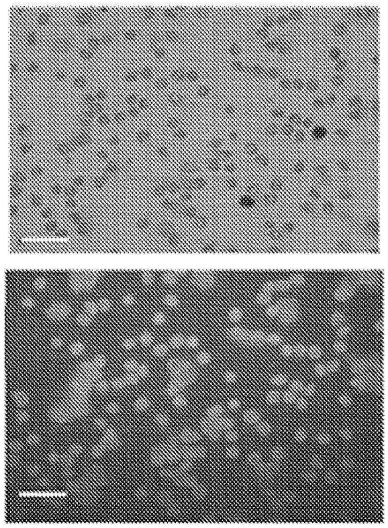
도면89c


100
80
60
40
20
0
% GFP+ RBCs
9 주
11 주
13 주
15 주
17 주
19 주
21 주
생체내 형질도입 후 주차

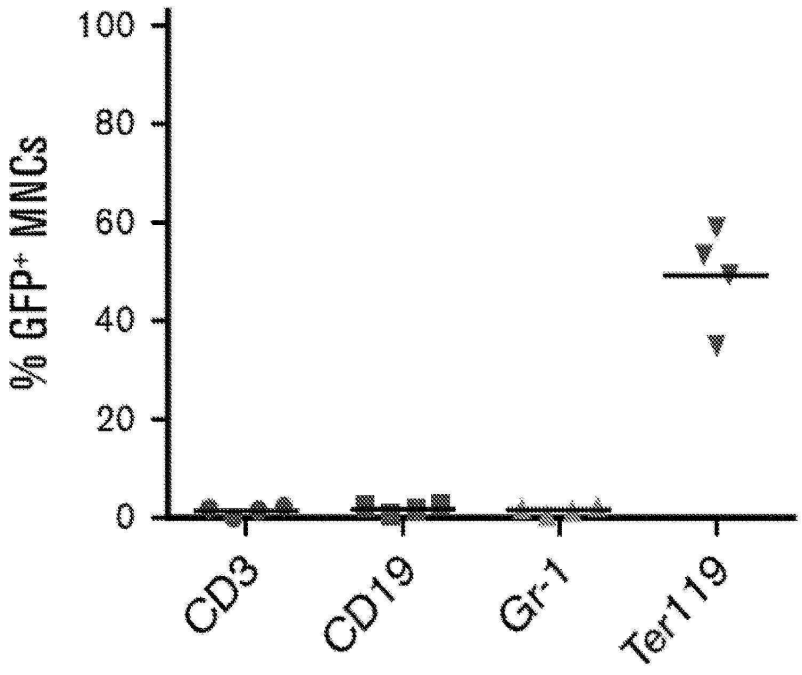
도면89d


10^5
10^4
10^3
0
2.33E-3
5.82E-3
17.1
82.9
0
10^3
10^4
10^5
GFP

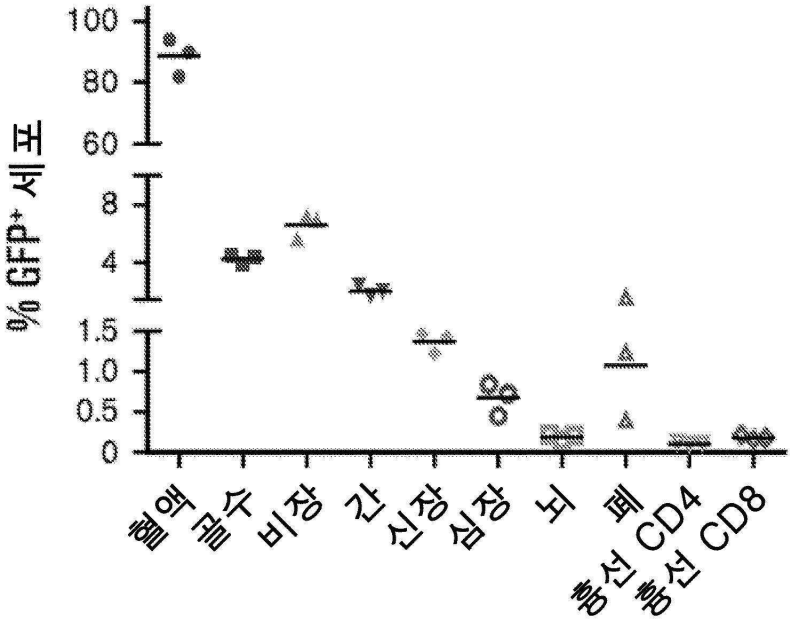
도면89e



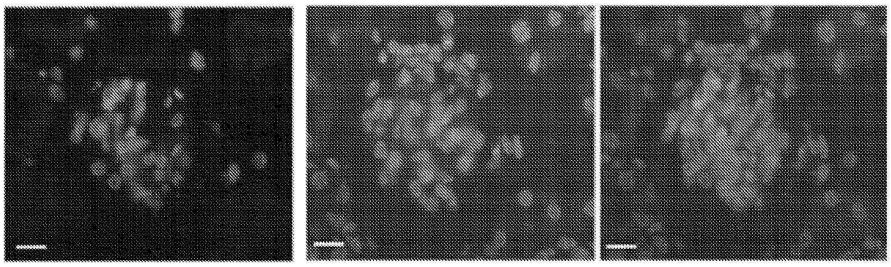
도면89f



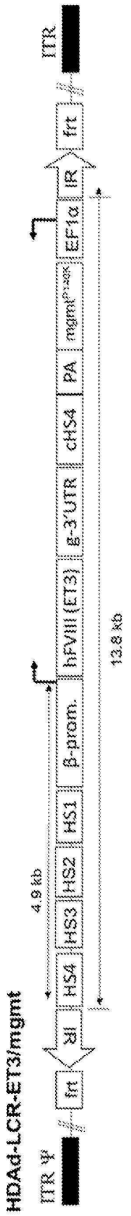
도면89g


% GFP+ 세포
100
80
60
8
4
1.5
1.0
0.5
0
혈액
골수
비장
간
신장
심장
뇌
폐
흉선 CD4
흉선 CD8

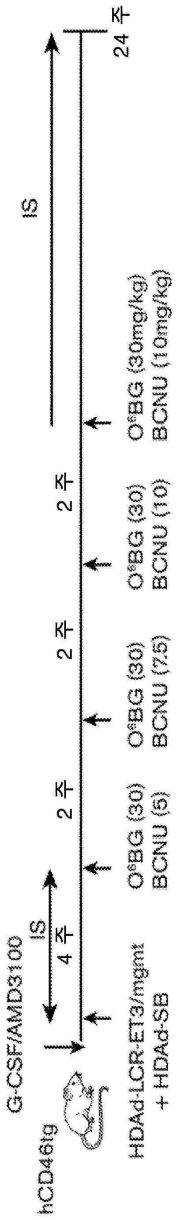
도면89h


GFP
DAPI (핵)
DAPI (병합)

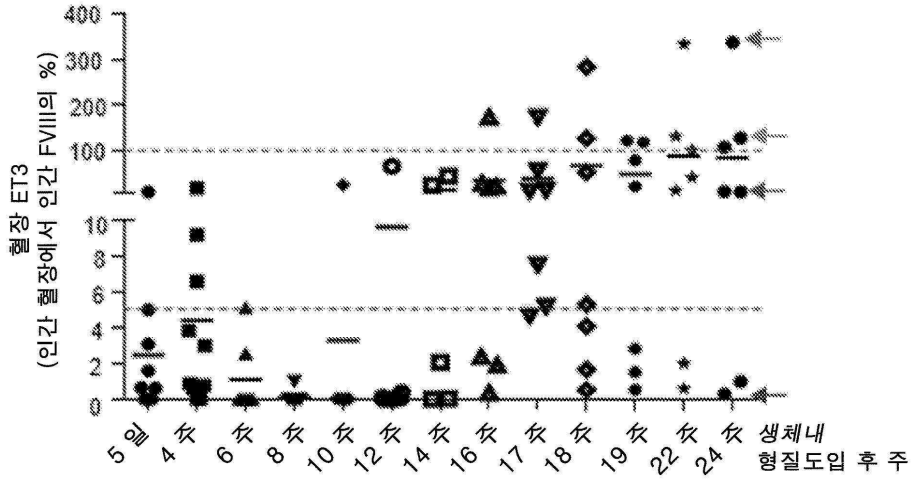
도면90a


HDAd-LCR-ET3/mgmt
ITR Ψ
frt
RI
HS4
HS3
HS2
HS1
β-prom.
hFVIII (ET3)
g-3'UTR
cHS4
PA
mgmt
EF1α
IR
frt
ITR
4.9 kb
13.8 kb

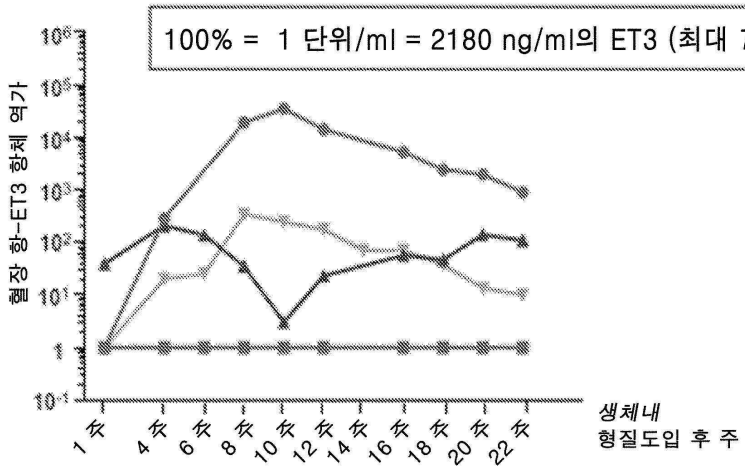
도면90b


hCD46tg
G-CSF/AMD3100
IS
4 주
2 주
2 주
2 주
IS
24 주
HDAd-LCR-ET3/mgmt
+ HDAd-SB
O⁶BG (30)
BCNU (6)
O⁶BG (30)
BCNU (7.5)
O⁶BG (30)
BCNU (10)
O⁶BG (30mg/kg)
BCNU (10mg/kg)

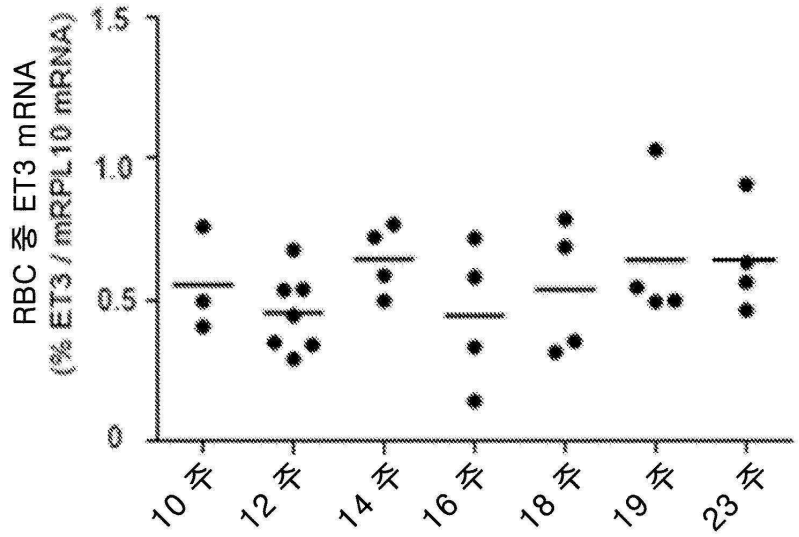
도면90c


혈장 ET3
(인간 혈장에서 인간 FVIII의 %)
400
300
200
100
10
8
6
4
2
0
5 일
4 주
6 주
8 주
10 주
12 주
14 주
16 주
17 주
18 주
19 주
22 주
24 주
생체내
형질도입 후 주

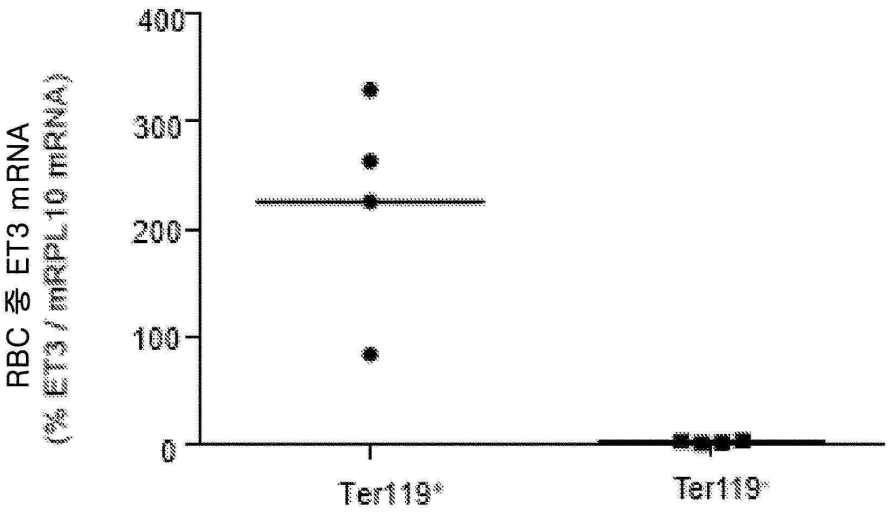
도면90d



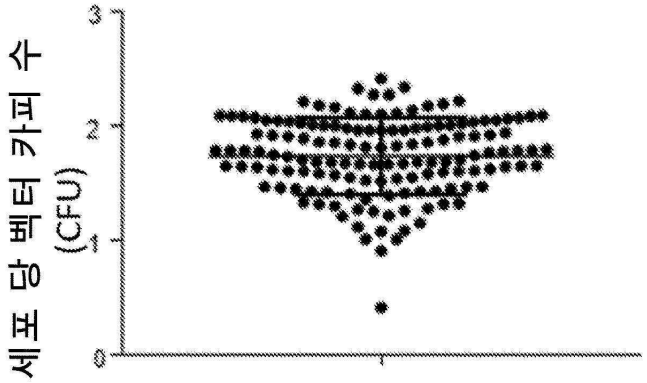

도면90e



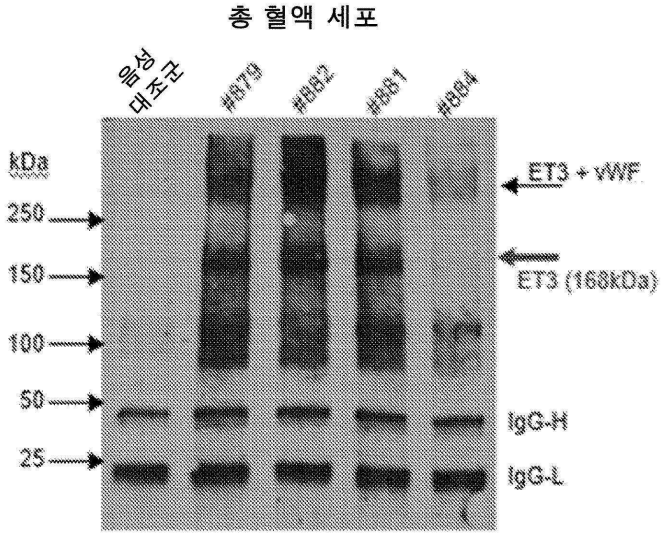

도면90f

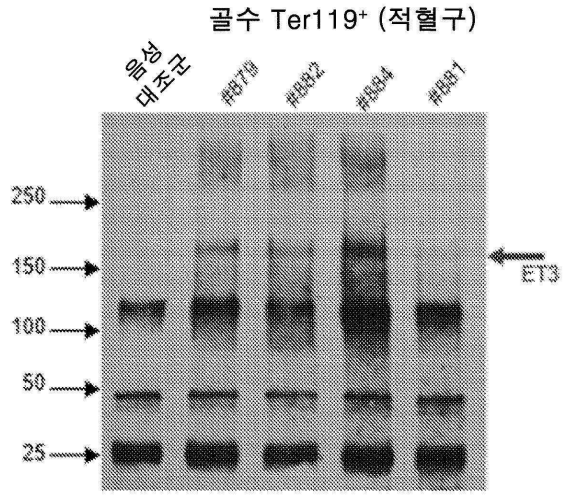


도면90g

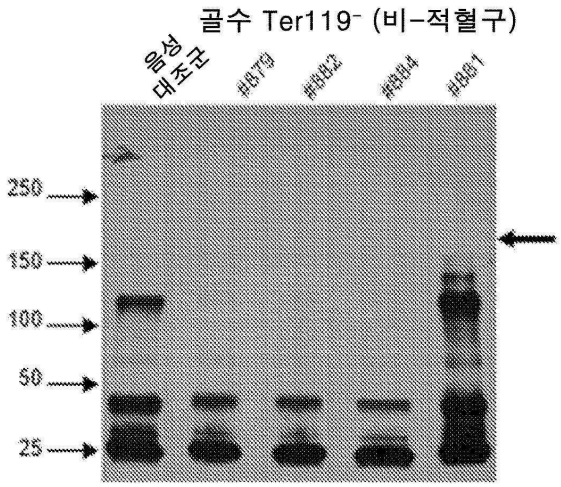



도면90ha

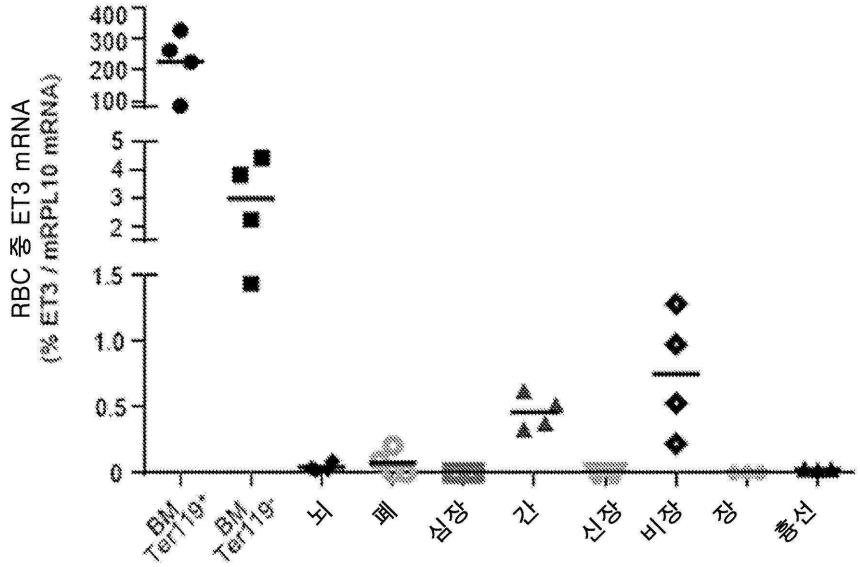


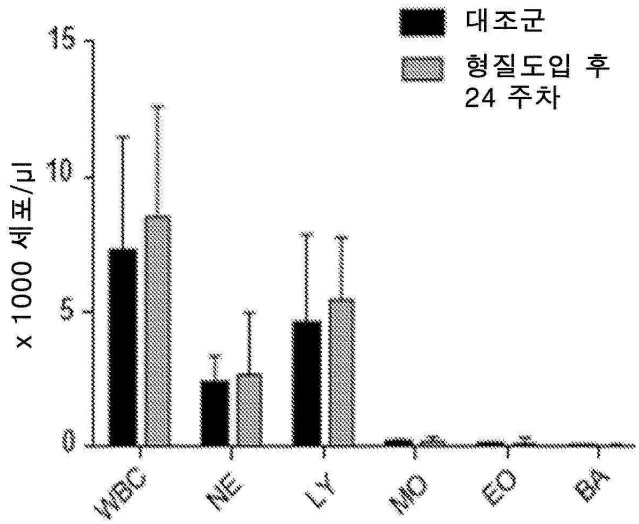
도면90hb



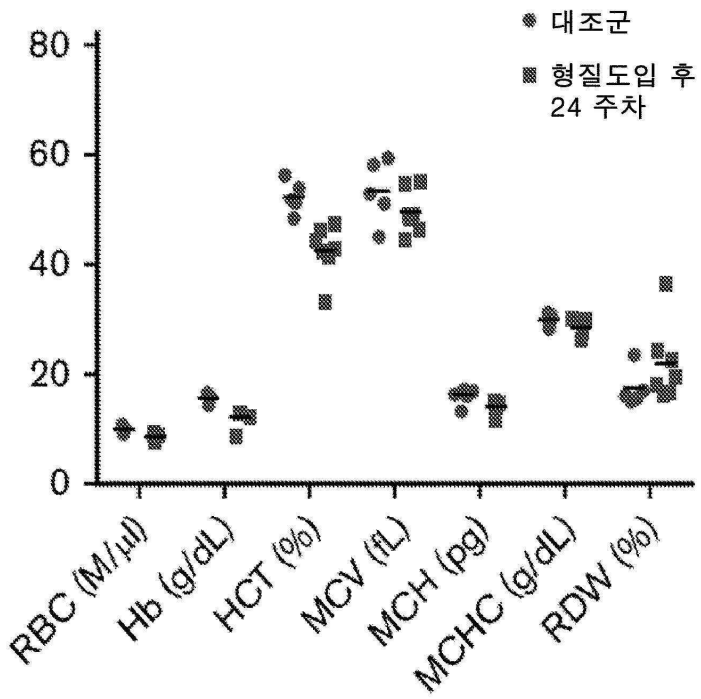

도면90i



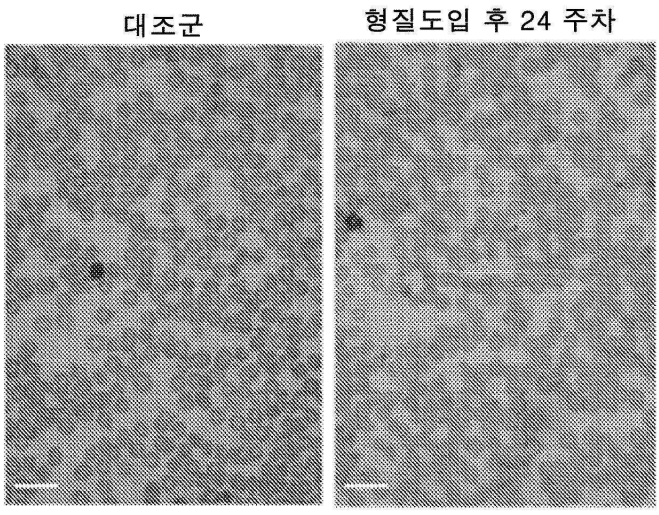

도면91a



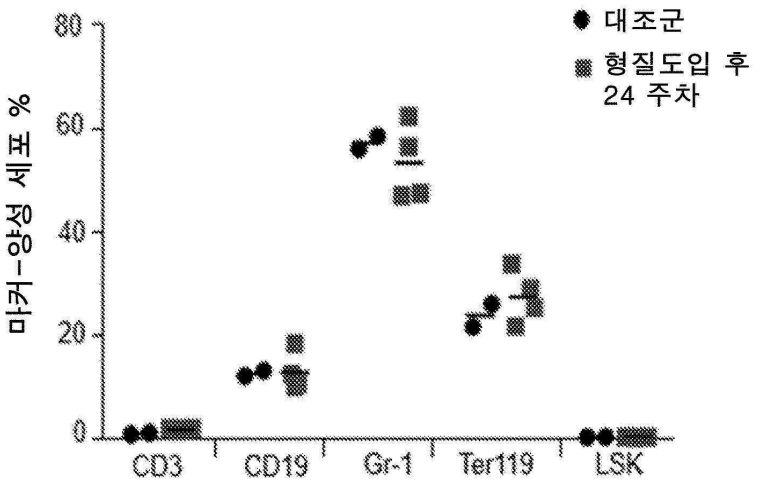
도면91b


대조군
형질도입 후 24 주차
80
60
40
20
0
RBC (M/μl)
Hb (g/dL)
HCT (%)
MCV (fL)
MCH (pg)
MCHC (g/dL)
RDW (%)

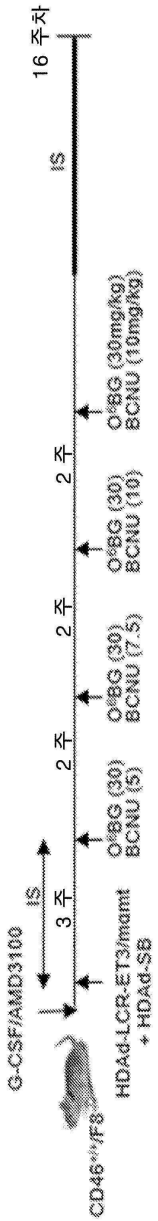
도면91c


대조군
형질도입 후 24 주차

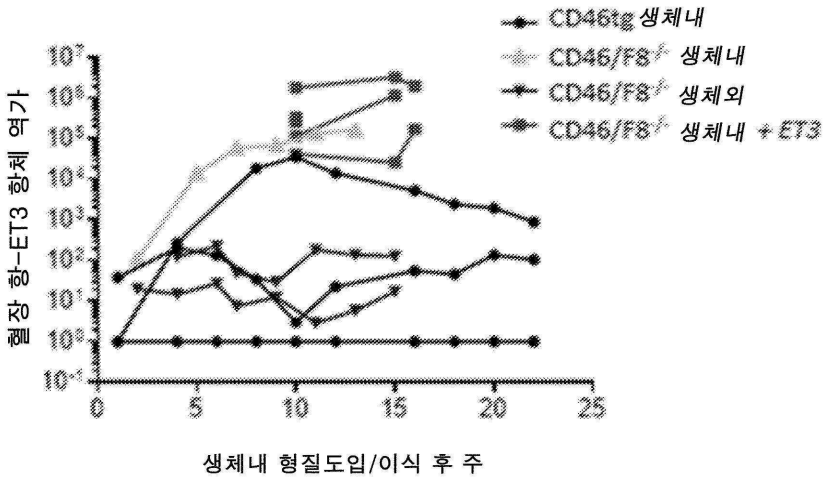
도면91d


대조군
형질도입 후
24 주차
마커-양성 세포 %
80
60
40
20
0
CD3
CD19
Gr-1
Ter119
LSK

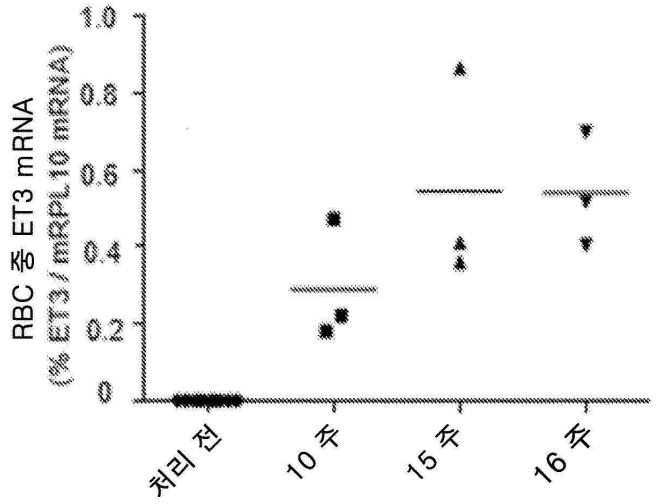
도면92a


16 주차
IS
O⁶BG (30mg/kg)
BCNU (10mg/kg)
2 주
O⁶BG (30)
BCNU (10)
2 주
O⁶BG (30)
BCNU (7.5)
2 주
O⁶BG (30)
BCNU (5)
IS
3 주
G-CSF/AMD3100
HDAd-LCR-ET3/mami
+ HDAd-SB
CD46tg/tgF8-/-

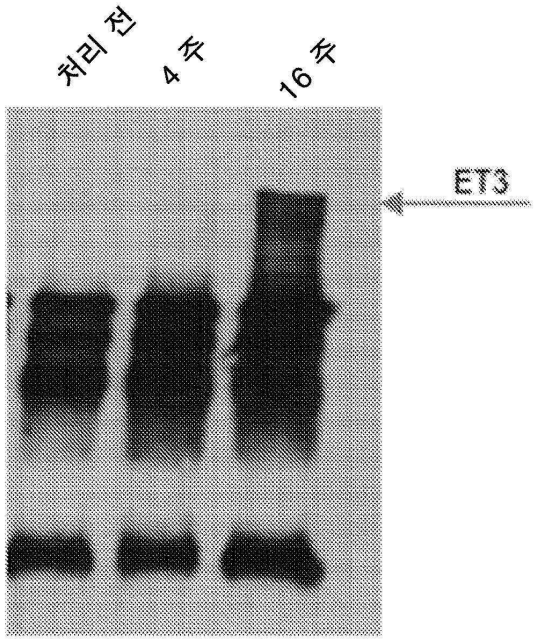
도면92b


CD46tg 생체내
CD46/F8-/- 생체내
CD46/F8-/- 생체외
CD46/F8-/- 생체내 + ET3
혈장 항-ET3 항체 역가
생체내 형질도입/이식 후 주

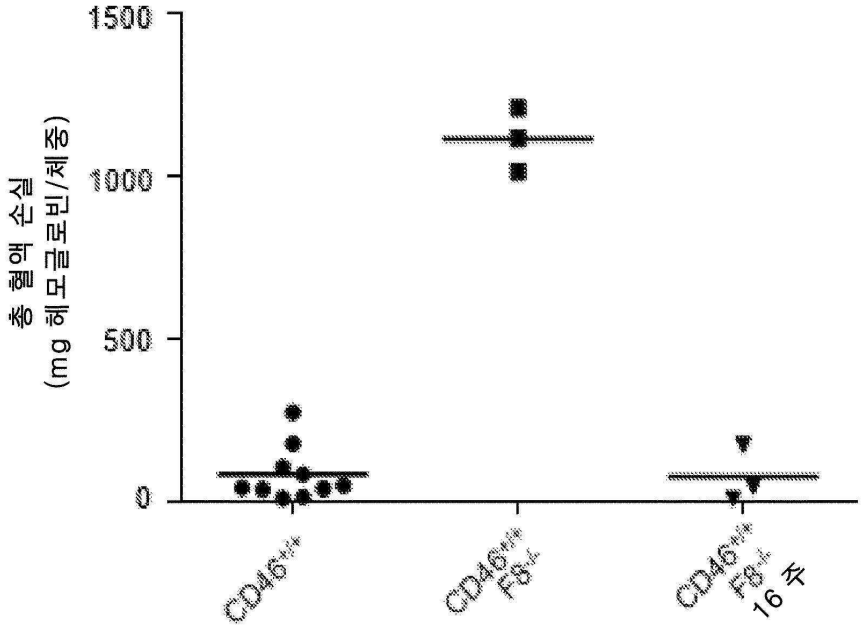
도면92c


RBC 중 ET3 mRNA
(% ET3 / mRPL10 mRNA)
처리 전
10 주
15 주
16 주

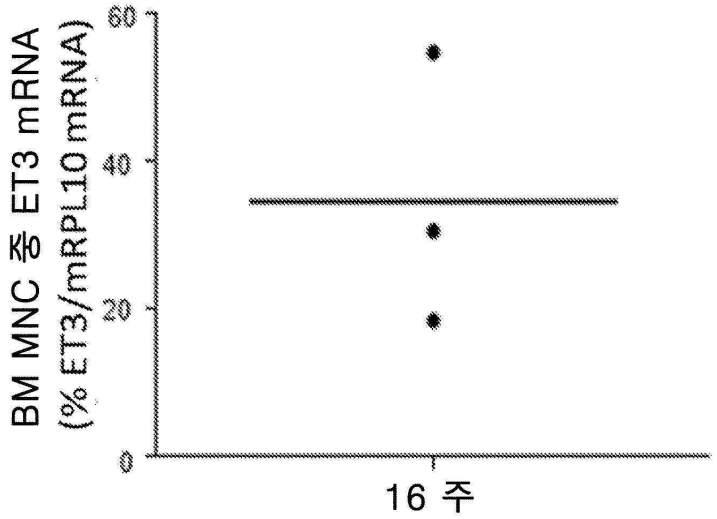
도면92d


처리 전
4 주
16 주
ET3

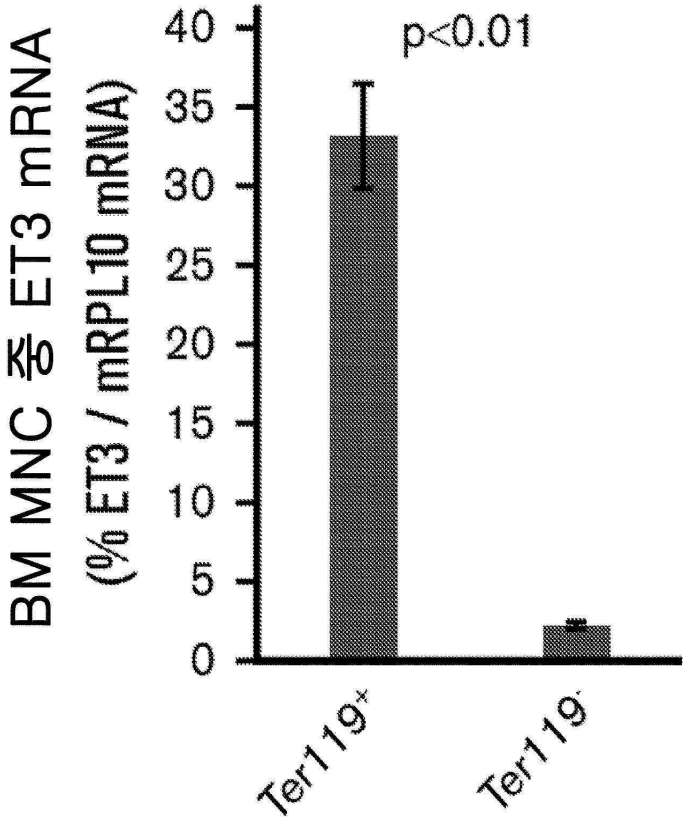
도면92e


1500
1000
500
0
총 혈액 손실
(mg 헤모글로빈/체중)
CD46+/+
CD46+/+ F8-/-
CD46+/+ F8-/-
16 주

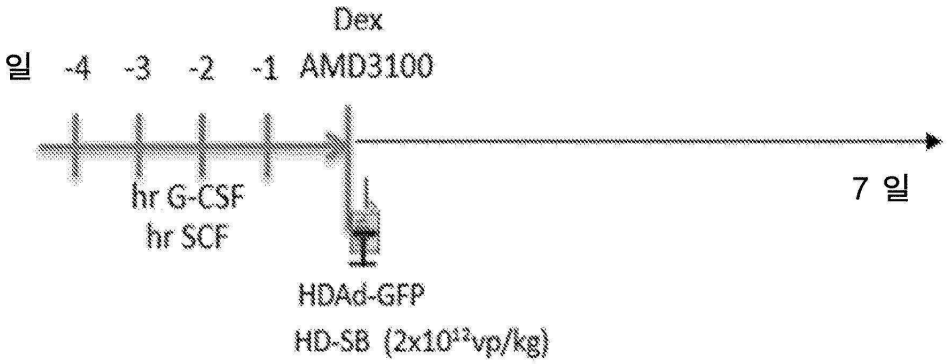
도면92f


BM MNC 중 ET3 mRNA
(% ET3/mRPL10 mRNA)
60
40
20
0
16 주

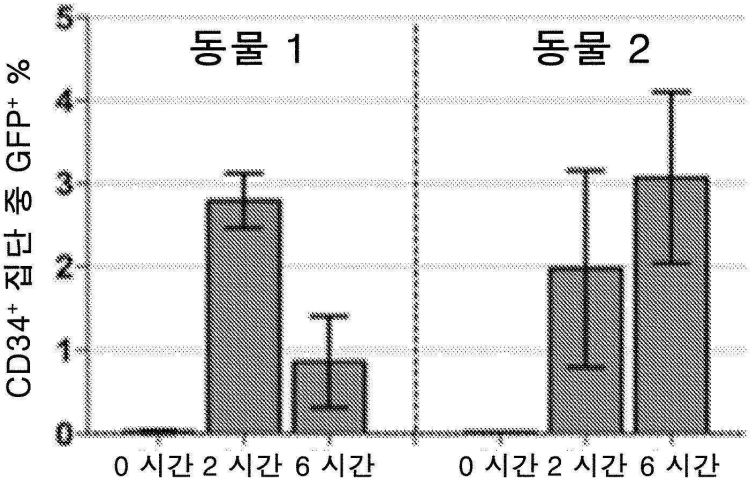
도면92g


BM MNC 중 ET3 mRNA
(% ET3 / mRPL10 mRNA)
40
35
30
25
20
15
10
5
0
p<0.01
Ter119+
Ter119-

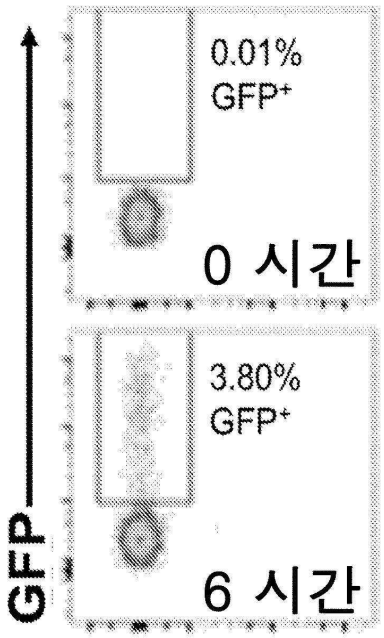
도면93a



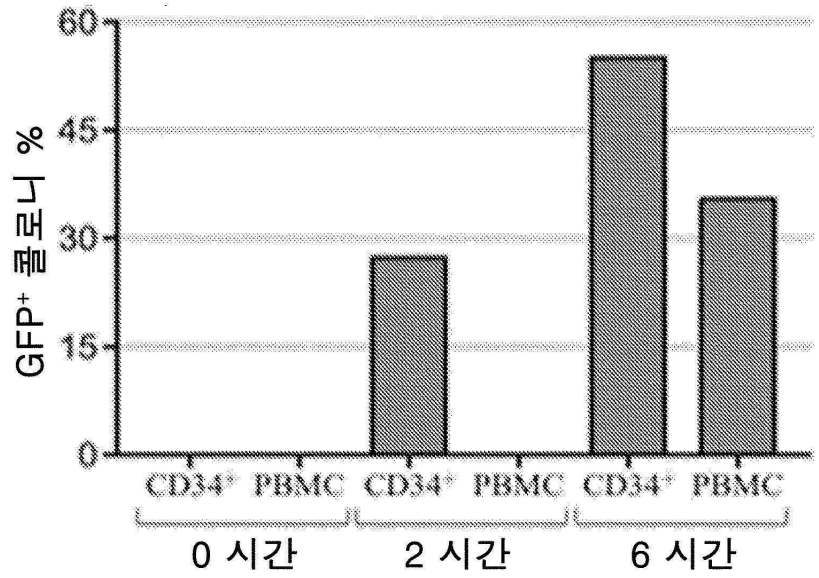

도면93b



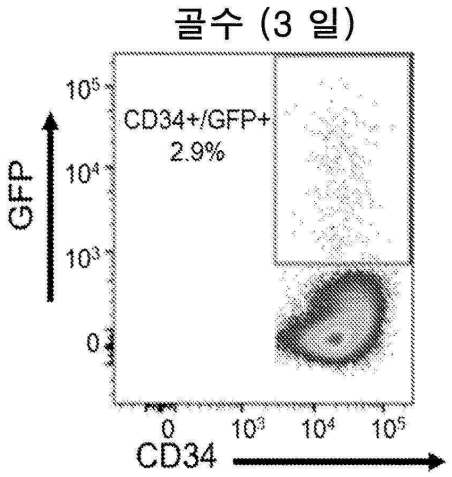

도면93c



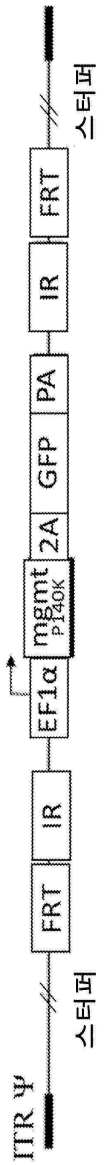
도면93d


60
45
30
15
0
GFP+ 콜로니 %
CD34+
PBMC
CD34+
PBMC
CD34+
PBMC
0 시간
2 시간
6 시간

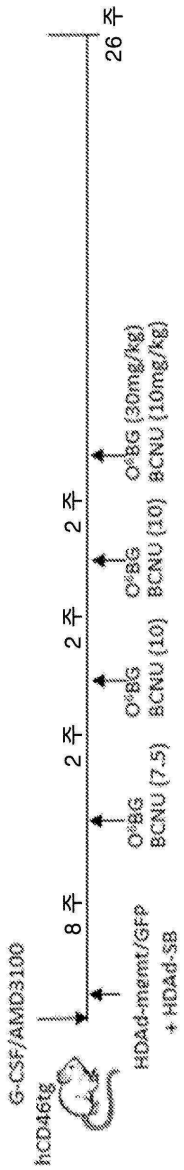
도면93e


골수 (3 일)
CD34+/GFP+
2.9%
GFP
CD34
0
10³
10⁴
10⁵

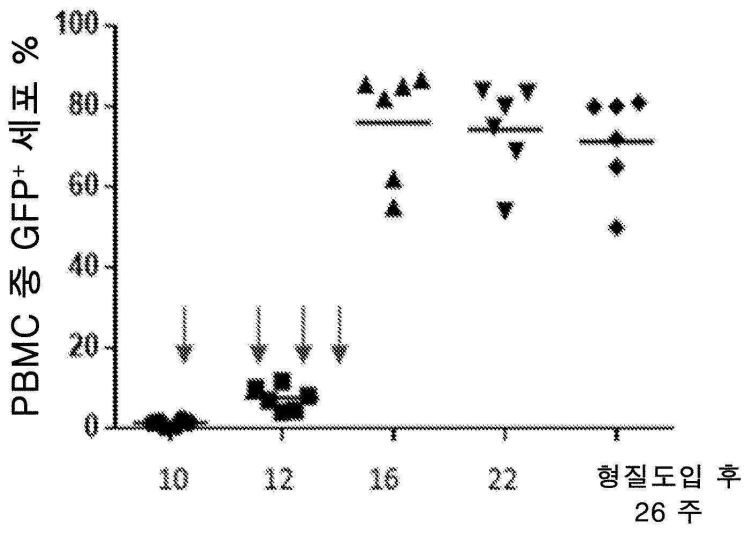
도면94a



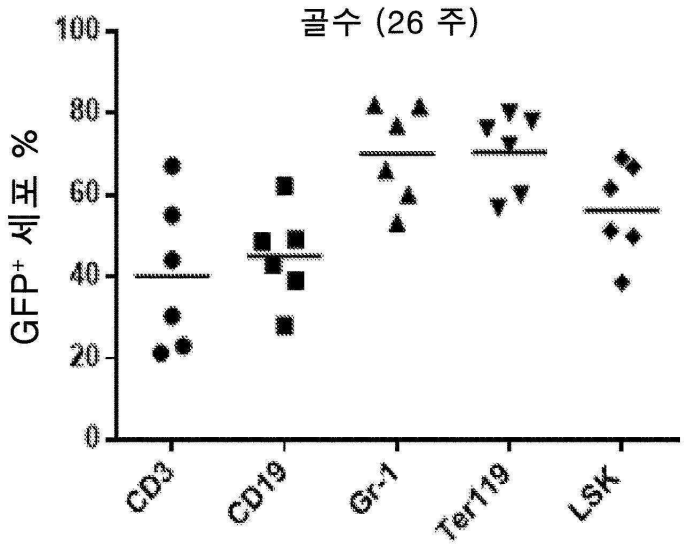
도면94b


주
26
2 주
O⁶BG (30mg/kg)
BCNU (10mg/kg)
2 주
O⁶BG
BCNU (10)
2 주
O⁶BG
BCNU (10)
2 주
O⁶BG
BCNU (7.5)
8 주
HDAd-mgmt/GFP
+ HDAd-SB
G-CSF/AMD3100
hCD46tg

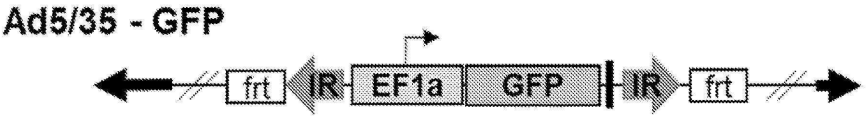
도면94c


PBMC 중 GFP⁺ 세포 %
100
80
60
40
20
0
10
12
16
22
형질도입 후
26 주

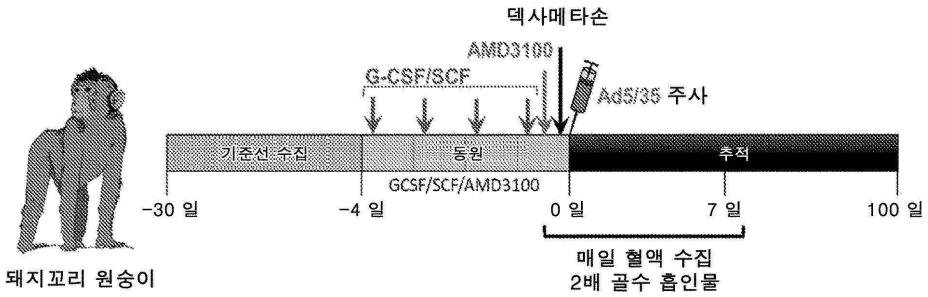
도면94d




도면94e



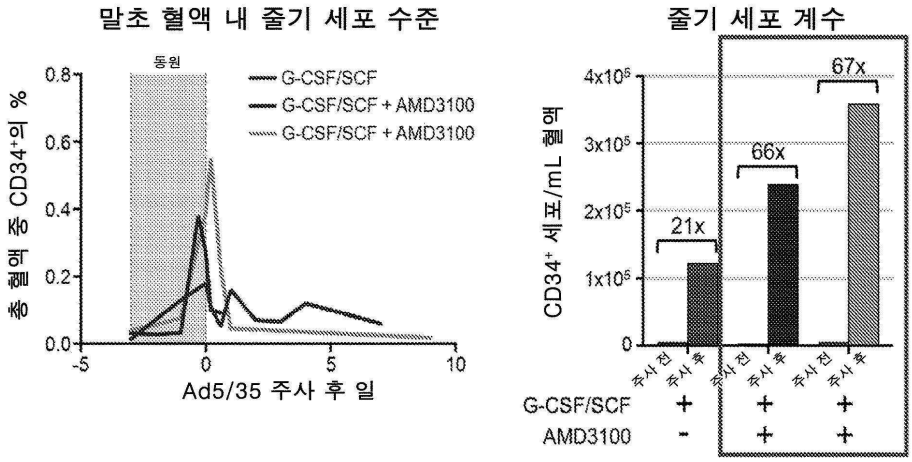

도면94f



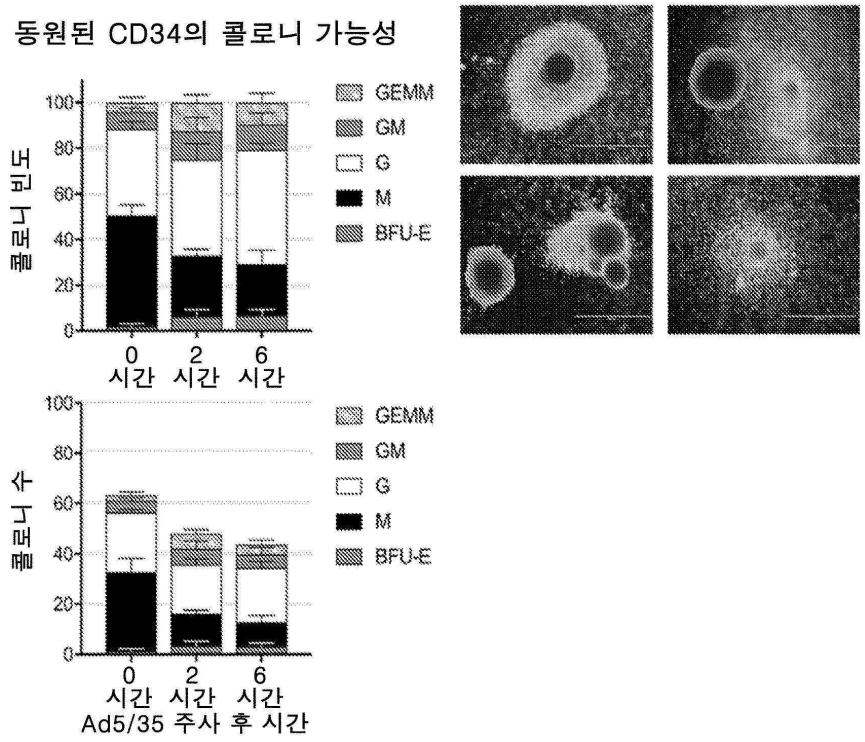

도면94g

| 동물 ID | G-CSF | SCF | AMD3100 | 용량 |
|---|---|---|---|---|
| Z12406 | 50μg/kg | 50μg/kg | - | $2x10^{12}$vp/kg |
| Z13126 | 50μg/kg | 50μg/kg | 4mg/kg | $4x10^{12}$vp/kg |
| Z15263 | 50μg/kg | 50μg/kg | 4mg/kg | $3x10^{12}$vp/kg |

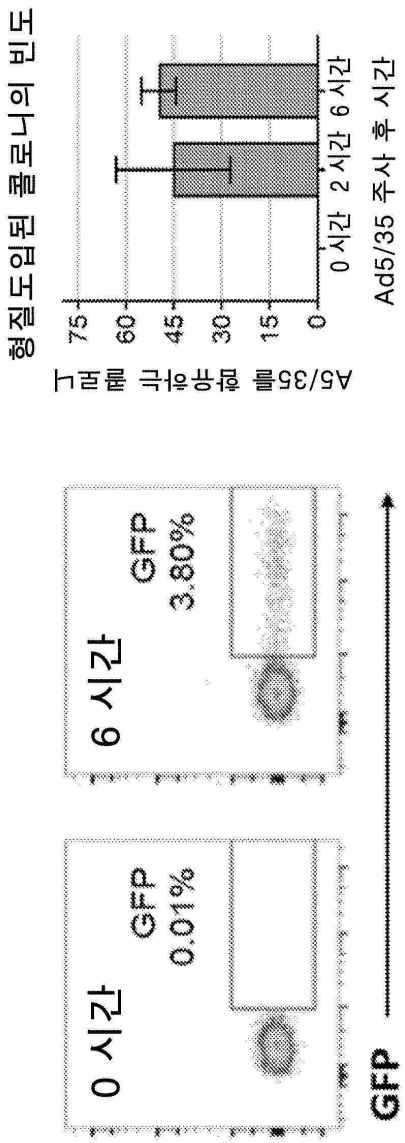
도면94h


말초 혈액 내 줄기 세포 수준
동원
G-CSF/SCF
G-CSF/SCF + AMD3100
G-CSF/SCF + AMD3100
총 혈액 중 CD34+의 %
Ad5/35 주사 후 일
줄기 세포 계수
CD34+ 세포/mL 혈액
21x
66x
67x
주사 전
주사 후
G-CSF/SCF
AMD3100

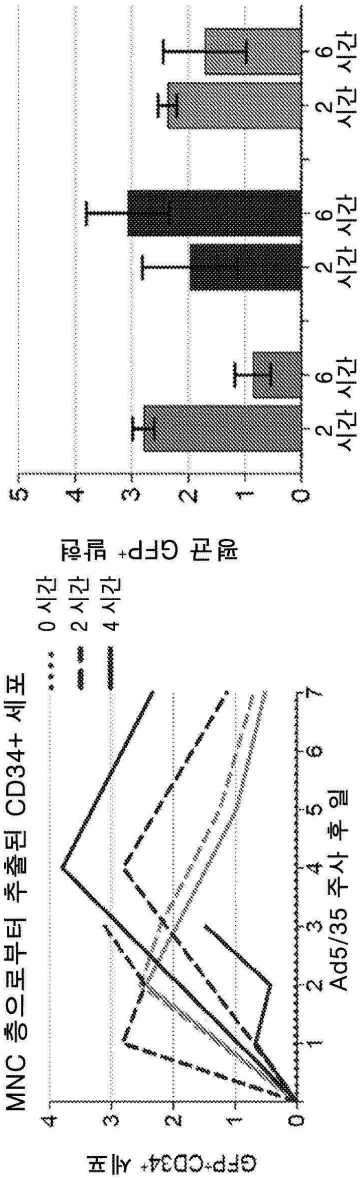
도면94i


동원된 CD34의 콜로니 가능성
GEMM
GM
G
M
BFU-E
콜로니 빈도
콜로니 수
0 시간
2 시간
6 시간
Ad5/35 주사 후 시간

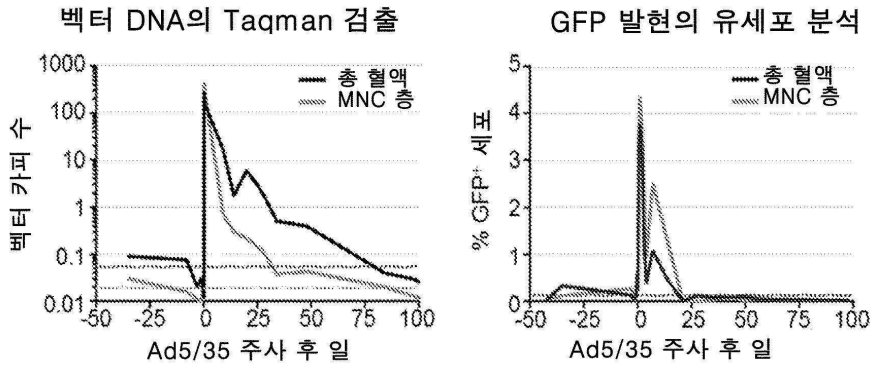
도면94j


형질도입된 콜로니의 빈도
Ad5/35를 함유하는 콜로니
75
60
45
30
15
0
0 시간
2 시간
6 시간
Ad5/35 주사 후 시간
0 시간
GFP
0.01%
6 시간
GFP
3.80%
GFP

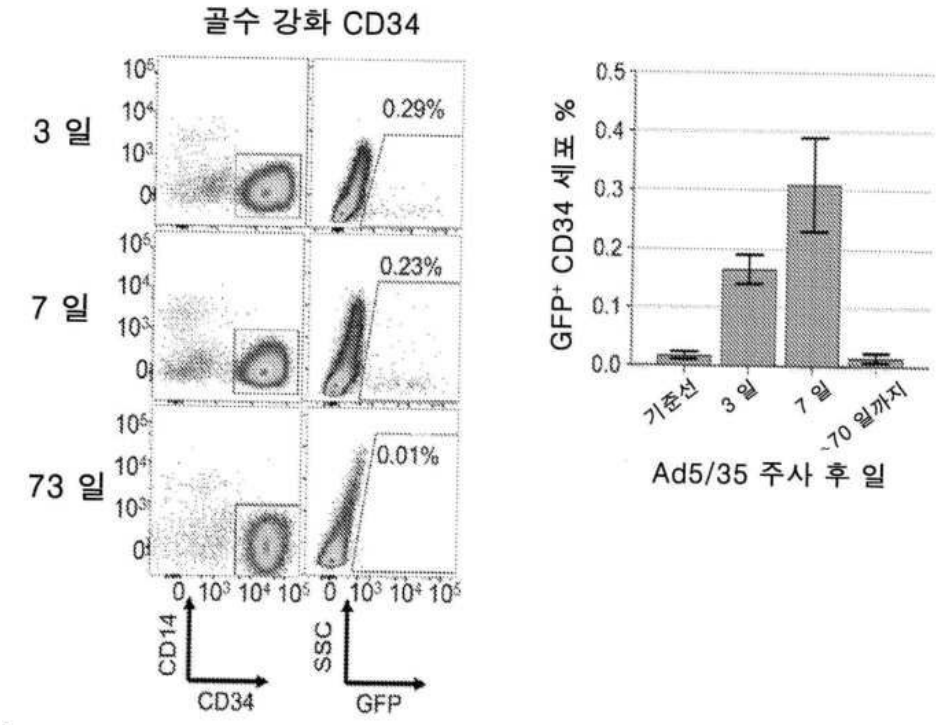
도면94k


평균 GFP+ 발현
2 시간
6 시간
2 시간
6 시간
2 시간
6 시간
GFP+CD34+ 세포
MNC 층으로부터 추출된 CD34+ 세포
0 시간
2 시간
4 시간
Ad5/35 주사 후 일

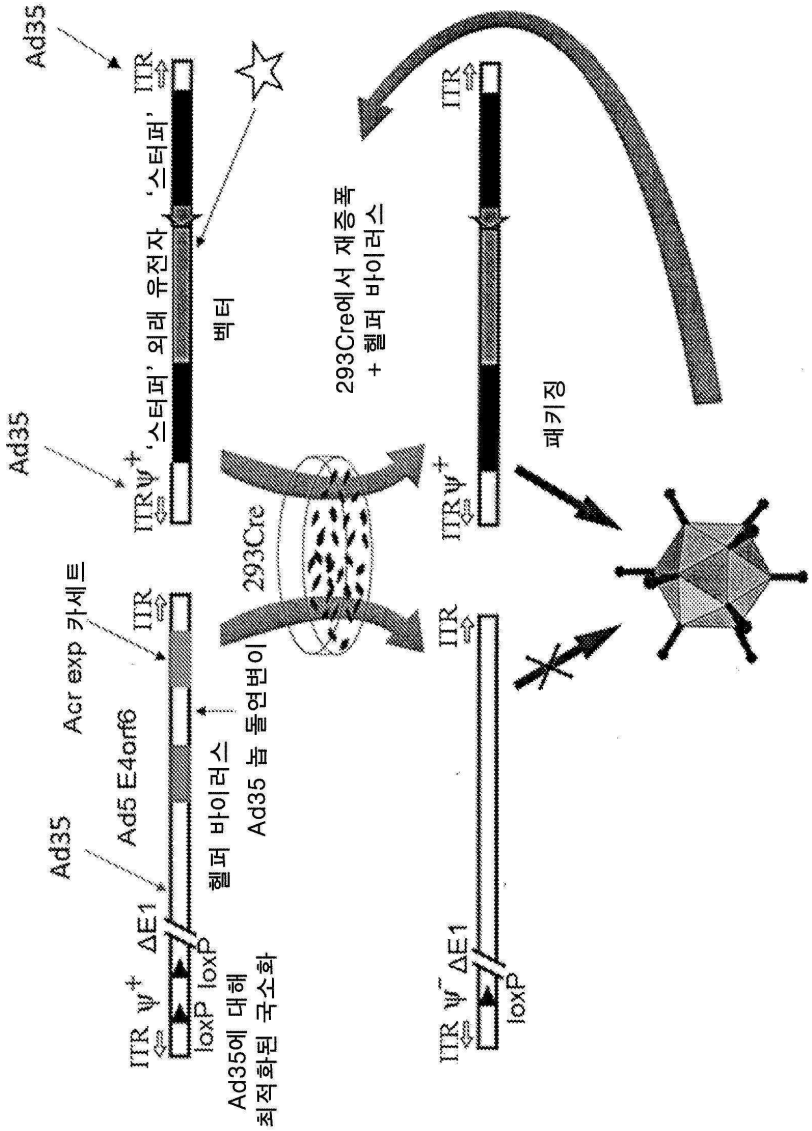
도면941


벡터 DNA의 Taqman 검출
벡터 카피 수
총 혈액
MNC 층
Ad5/35 주사 후 일
GFP 발현의 유세포 분석
% GFP+ 세포
총 혈액
MNC 층
Ad5/35 주사 후 일

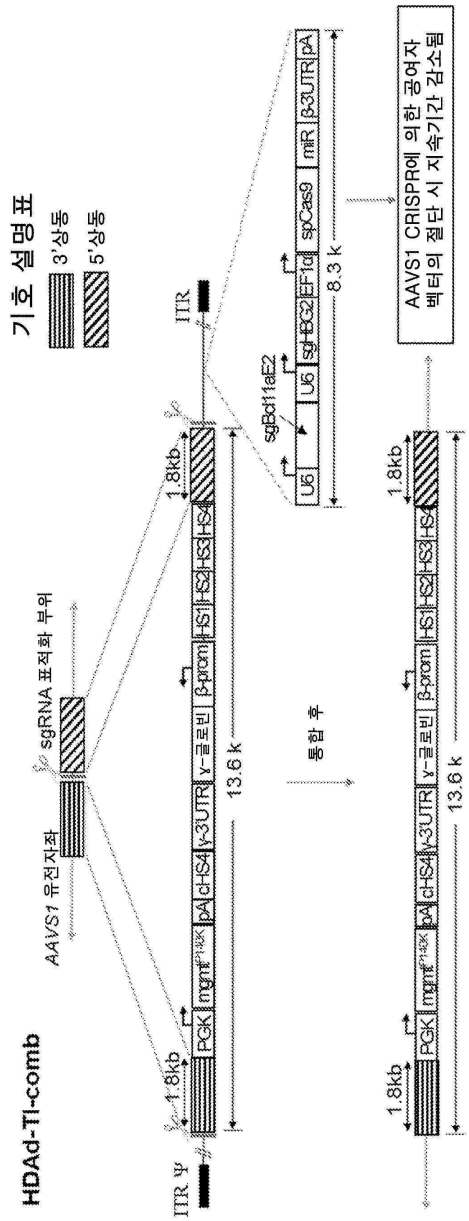
도면94m


골수 강화 CD34
3 일
7 일
73 일
0.29%
0.23%
0.01%
CD14
CD34
SSC
GFP
GFP+ CD34 세포 %
0.5
0.4
0.3
0.2
0.1
0.0
기준선
3 일
7 일
~70 일까지
Ad5/35 주사 후 일

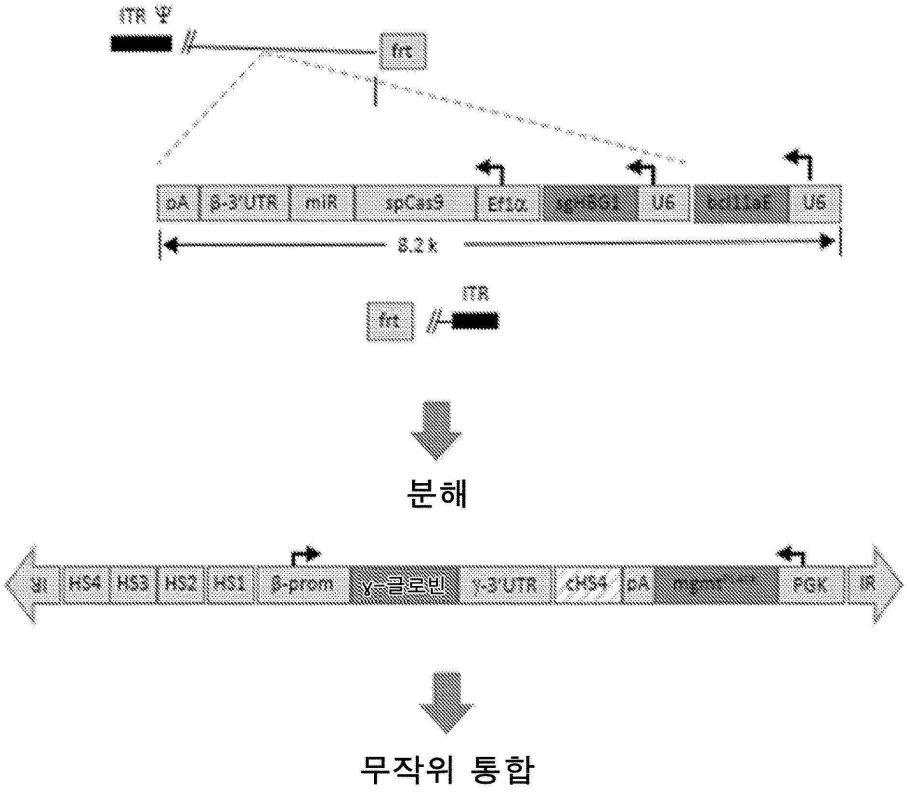
도면95


Ad35
ITR
'스터퍼'
외래 유전자
'스터퍼'
ITR
ψ+
벡터
293Cre에서 재증폭
+ 헬퍼 바이러스
패키징
293Cre
Acr exp 카세트
Ad5 E4orf6
헬퍼 바이러스
Ad35 놉 돌연변이
ΔE1
loxP
Ad35에 대해
최적화된 국소화
ψ-

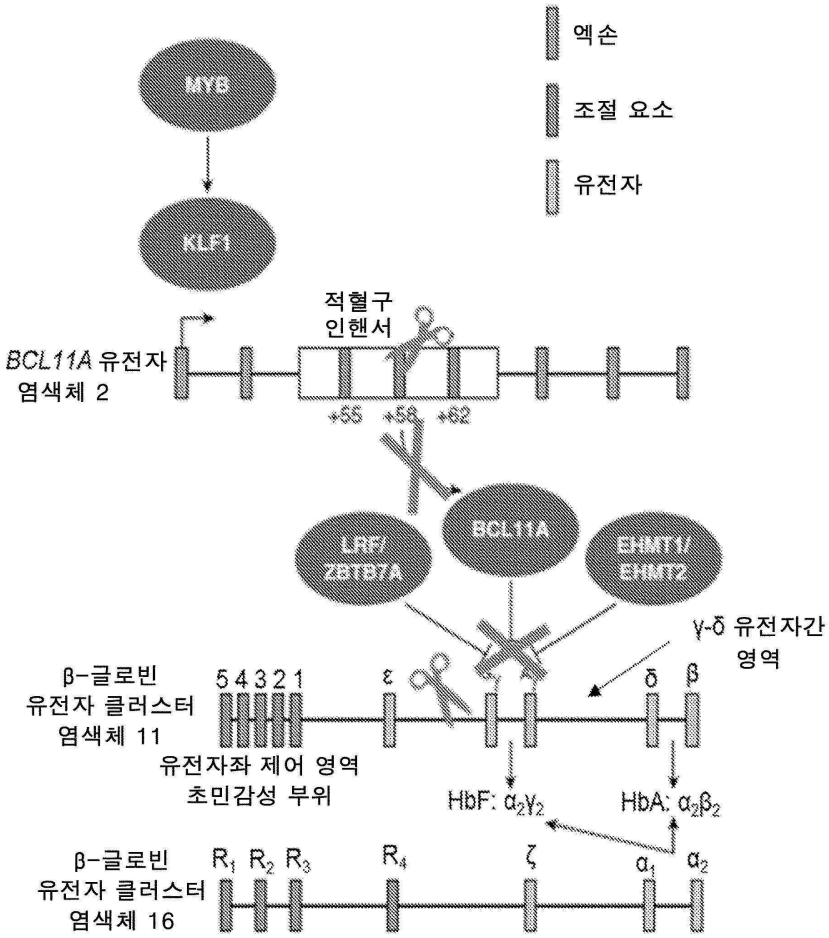
도면96


기호 설명표
3'상동
5'상동
HDAd-TI-comb
AAVS1 유전자좌
sgRNA 표적화 부위
ITR Ψ
1.8kb
PGK
mgmt
pA
cHS4
γ-3UTR
γ-글로빈
β-prom
HS1
HS2
HS3
HS4
13.6 k
ITR
sgBcl11aE2
U6
sgHBG2
EF1α
spCas9
miR
β3UTR
pA
8.3 k
통합 후
AAVS1 CRISPR에 의한 공여자
벡터의 절단 시 지속기간 감소됨

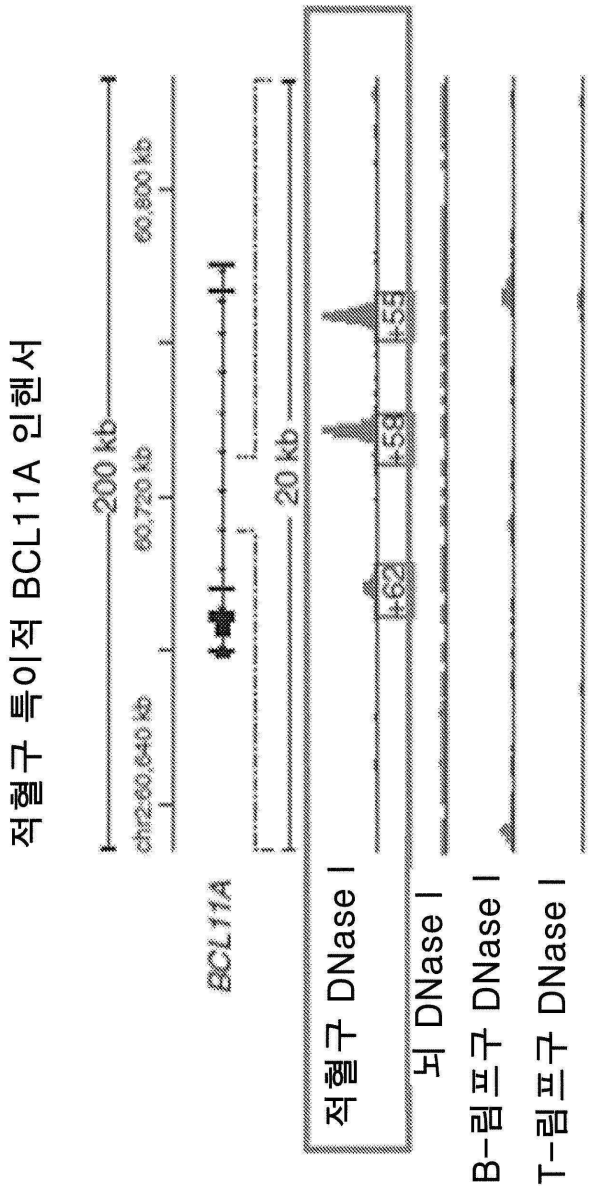
도면97a


ITR
frt
pA
β-3'UTR
miR
spCas9
Eflα
U6
8.2 k
frt
ITR
분해
β-prom
γ-글로빈
γ-3'UTR
cHS4
pA
PGK
IR
무작위 통합

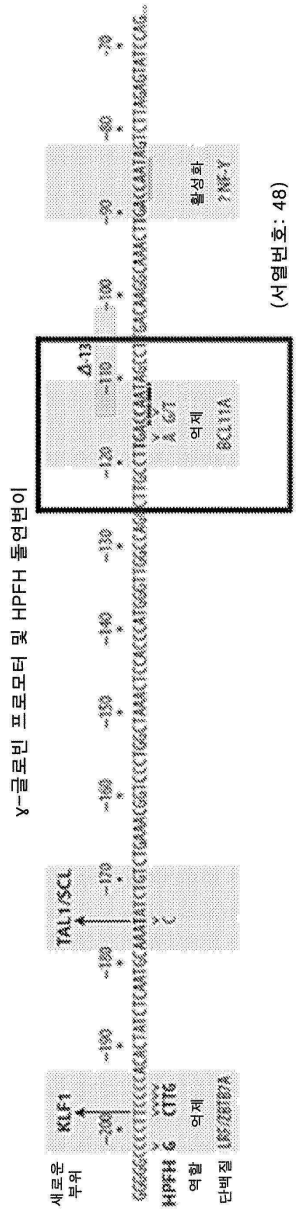
도면97b


엑손
조절 요소
유전자
MYB
KLF1
적혈구
인핸서
BCL11A 유전자
염색체 2
+55
+58
+62
LRF/
ZBTB7A
BCL11A
EHMT1/
EHMT2
γ-δ 유전자간
영역
β-글로빈
유전자 클러스터
염색체 11
5 4 3 2 1
ε
δ
β
유전자좌 제어 영역
초민감성 부위
HbF: α2γ2
HbA: α2β2
β-글로빈
유전자 클러스터
염색체 16
R1 R2 R3
R4
ζ
α1
α2

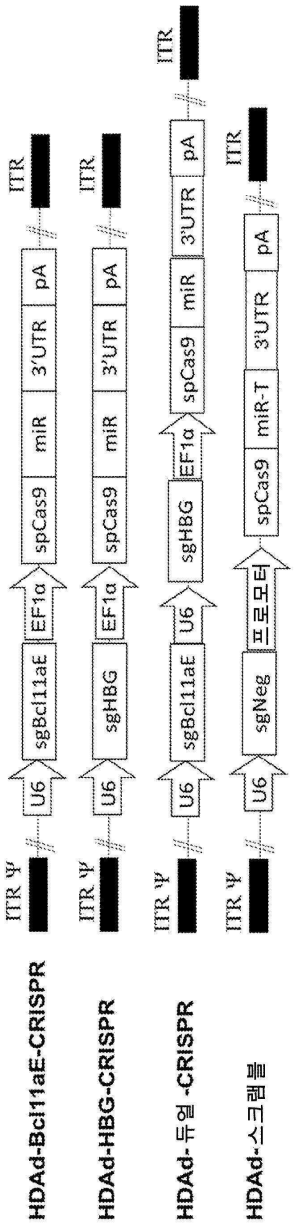
도면97c


적혈구 특이적 BCL11A 인핸서
chr2:60,640 kb
60,720 kb
60,800 kb
200 kb
BCL11A
20 kb
적혈구 DNase I
뇌 DNase I
B-림프구 DNase I
T-림프구 DNase I

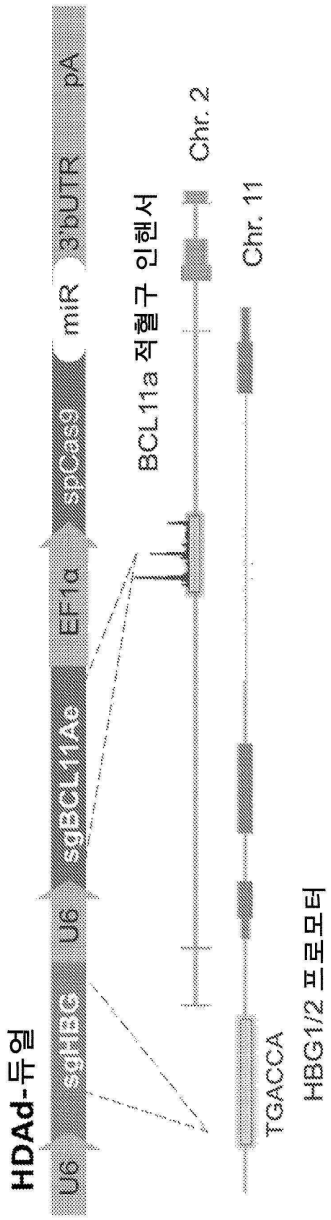
도면97d



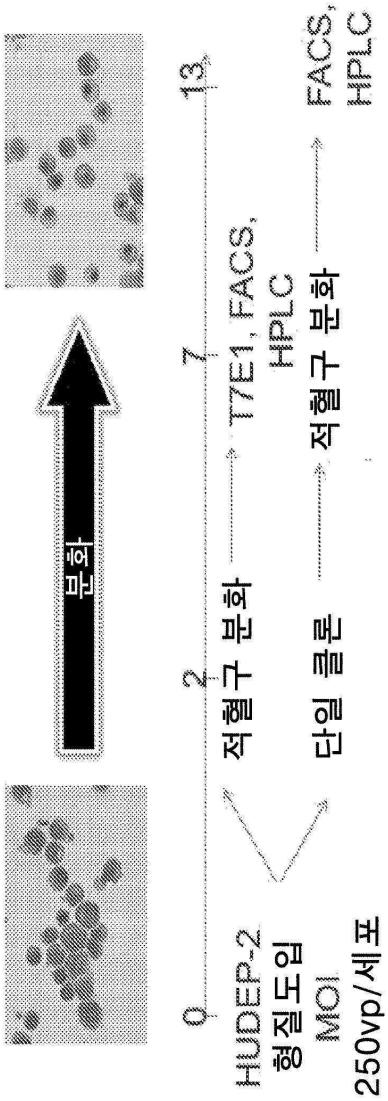
도면98a


HDAd-Bcl11aE-CRISPR
ITR Ψ
U6
sgBcl11aE
EF1α
spCas9
miR
3'UTR
pA
ITR
HDAd-HBG-CRISPR
ITR Ψ
U6
sgHBG
EF1α
spCas9
miR
3'UTR
pA
ITR
HDAd- 듀얼 -CRISPR
ITR Ψ
U6
sgBcl11aE
U6
sgHBG
EF1α
spCas9
miR
3'UTR
pA
ITR
HDAd-스크램블
ITR Ψ
U6
sgNeg
프로모터
spCas9
miR-T
3'UTR
pA
ITR

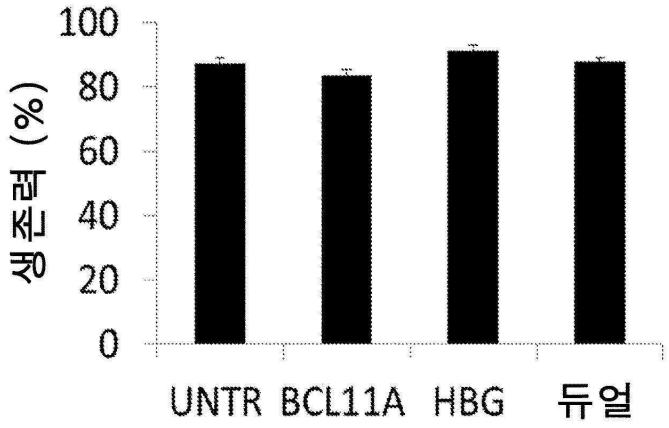
도면98b


HDAd-듀얼
U6
sgHBG
U6
sgBCL11Ae
EF1a
SpCas9
miR
3bUTR
pA
BCL11a 적혈구 인핸서
Chr. 2
Chr. 11
TGACCA
HBG1/2 프로모터

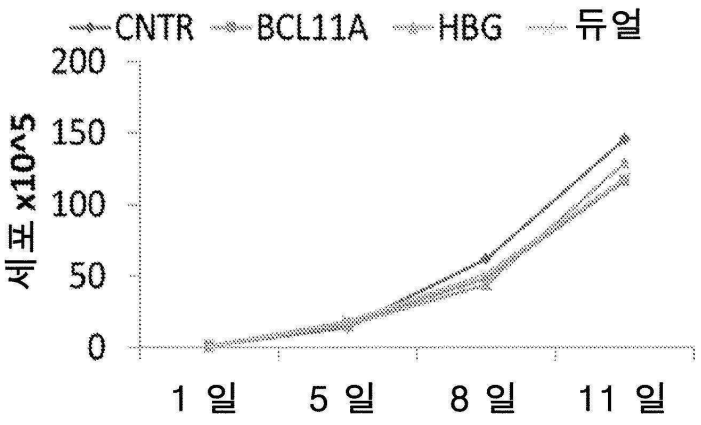
도면98c


분화
0
2
7
13
HUDEP-2
형질도입
MOI
250vp/세포
적혈구 분화
단일 클론
T7E1, FACS,
HPLC
적혈구 분화
FACS,
HPLC

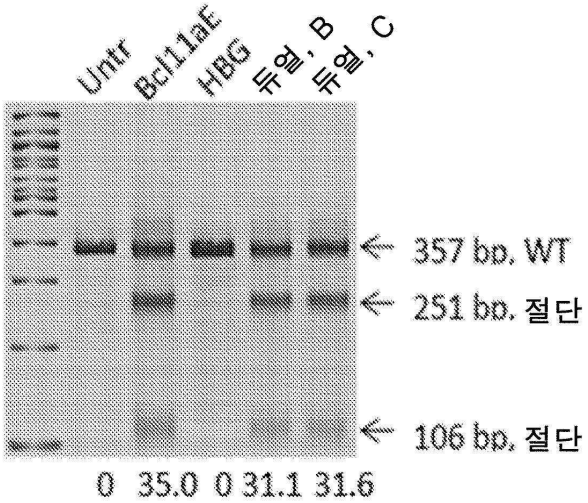
도면98d


생존률 (%)
100
80
60
40
20
0
UNTR
BCL11A
HBG
듀얼

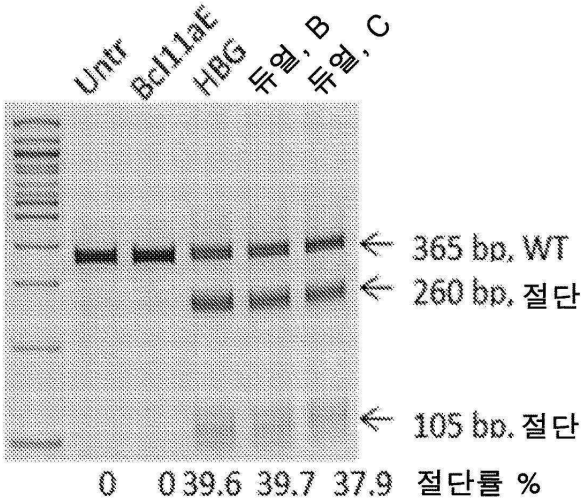
도면98e



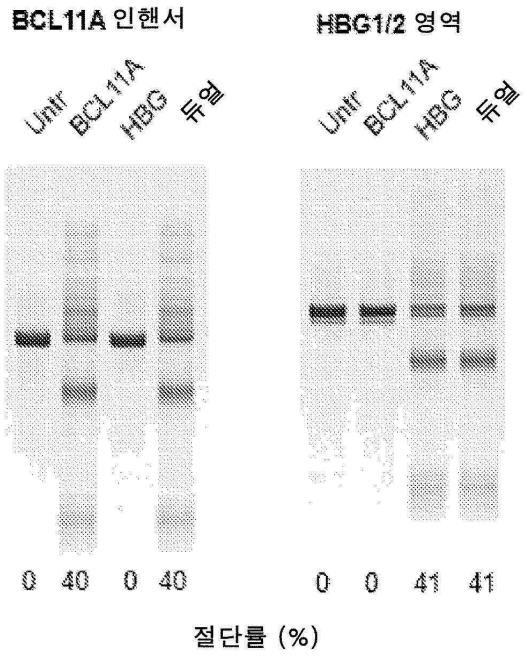

도면98f



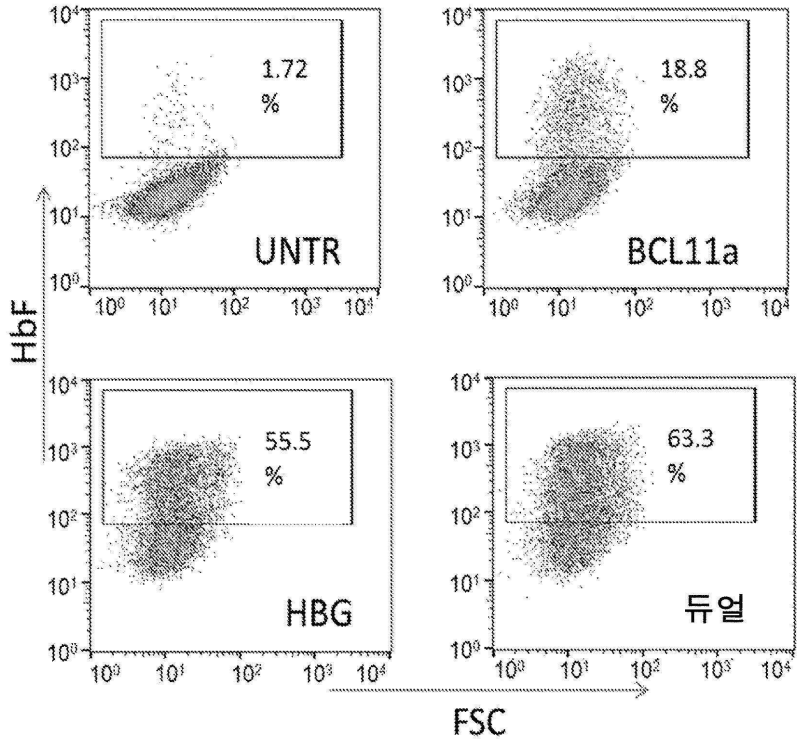

도면98g

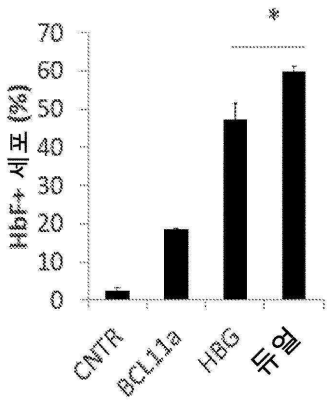


도면98h

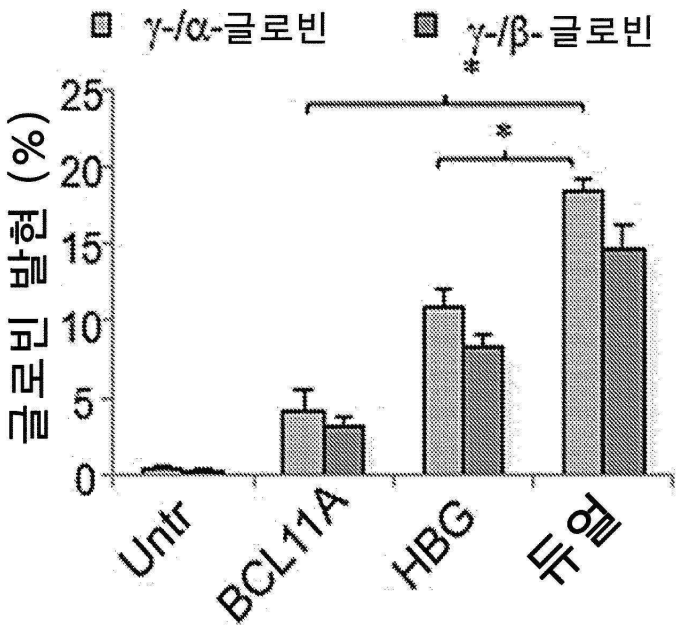


도면98i

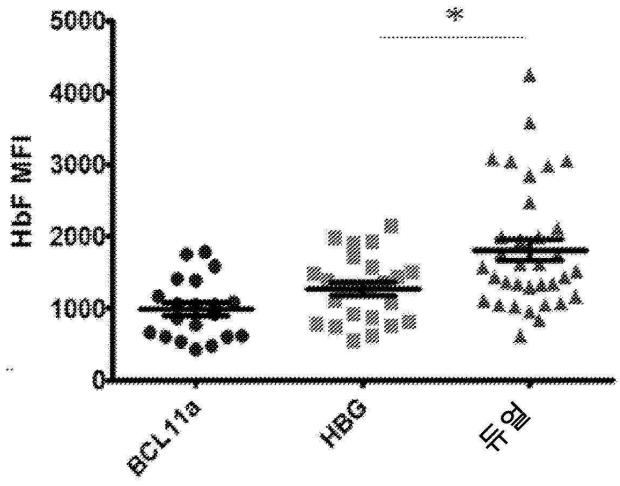

1.72
%
UNTR
18.8
%
BCL11a
55.5
%
HBG
63.3
%
듀얼
HbF
FSC


HbF+ 세포 (%)
70
60
50
40
30
20
10
0
*
CNTR
BCL11a
HBG
듀얼

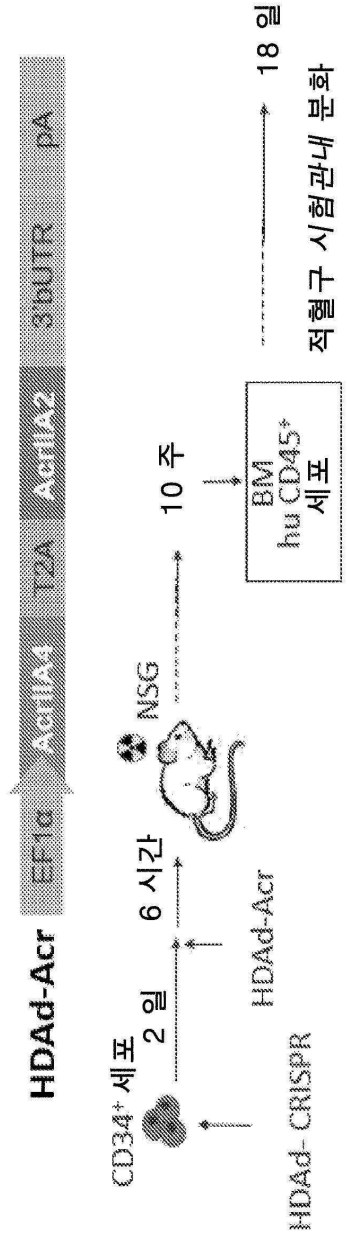
도면98j


γ-/α-글로빈
γ-/β-글로빈
글로빈 발현 (%)
25
20
15
10
5
0
Untr
BCL11A
HBG
듀엘

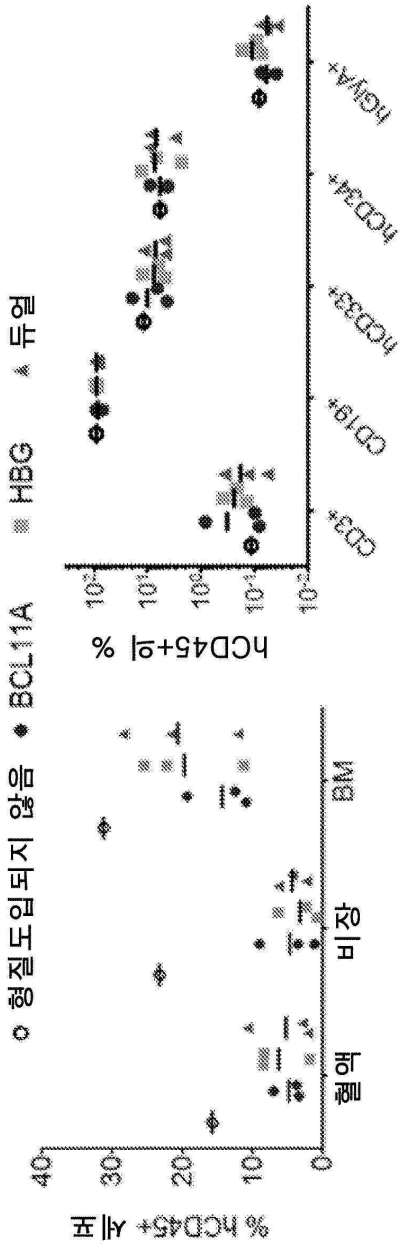
도면98k


HbF MFI
5000
4000
3000
2000
1000
0
BCL11a
HBG
듀엘

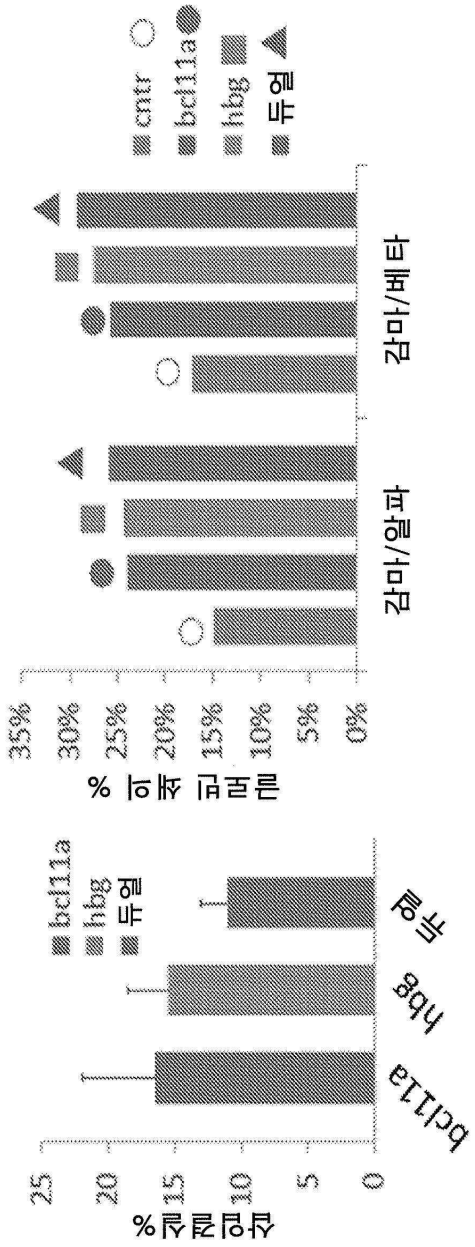
도면981


HDAd-Acr
EF1α
AcrIIA4
T2A
AcrIIA2
3bUTR
pA
CD34+ 세포
2 일
6 시간
NSG
10 주
BM
hu CD45+
세포
18 일
적혈구 시험관내 분화
HDAd-Acr
HDAd- CRISPR

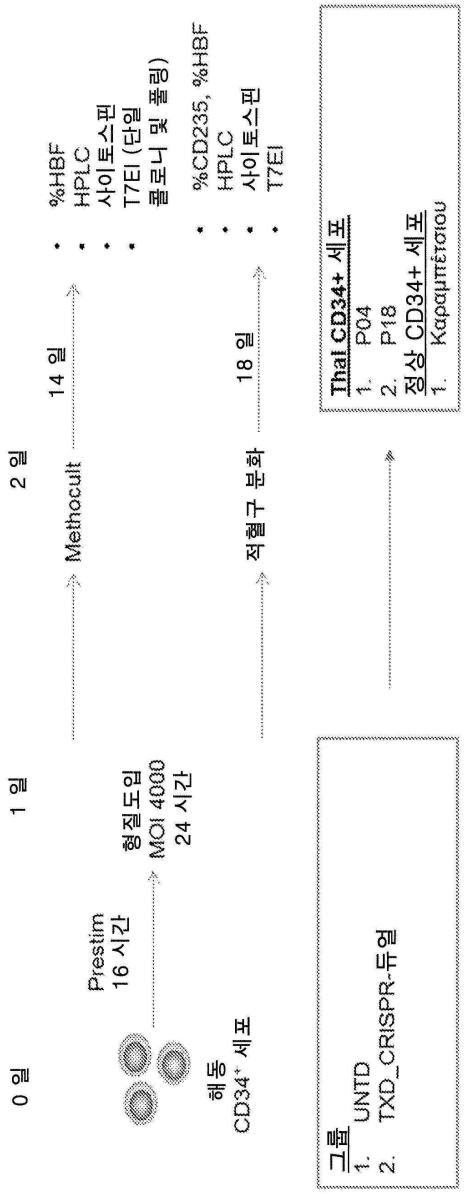
도면98m


% hCD45+ 세포
40
30
20
10
0
혈액
비장
BM
hCD45+의 %
hCD3+
hCD19+
hCD33+
hCD34+
hGlyA+
○ 형질도입되지 않음 ◆ BCL11A ■ HBG ▲ 듀얼

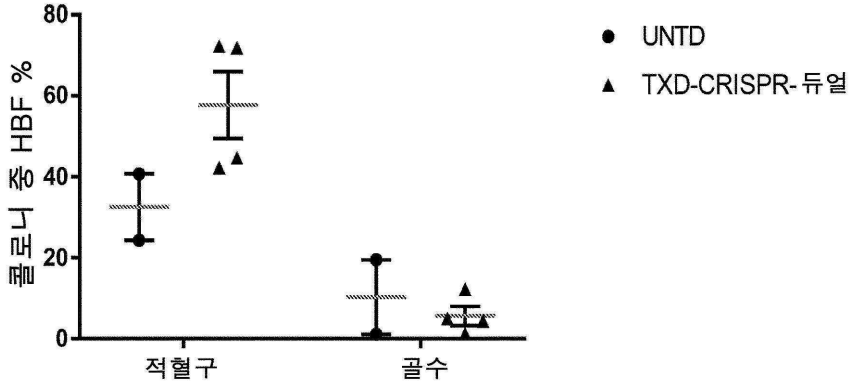
도면98n


cntr
bcl11a
hbg
듀얼
감마/베타
감마/알파
글로빈 쇄의 %
35%
30%
25%
20%
15%
10%
5%
0%
bcl11a
hbg
듀얼
삽입결실%
25
20
15
10
5
0

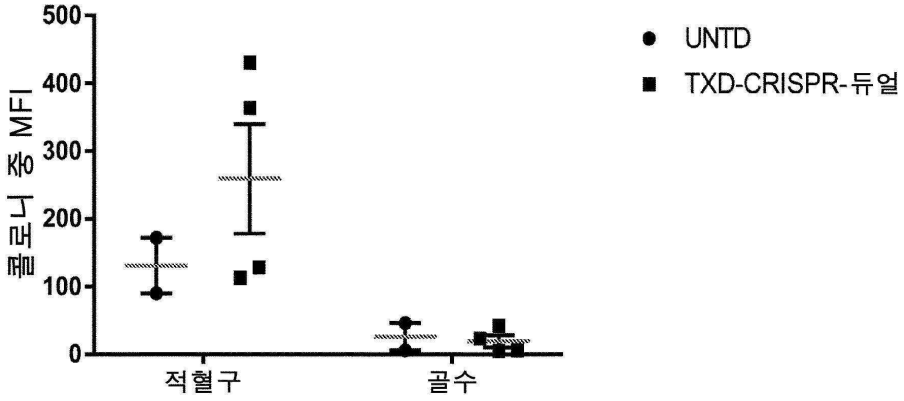
도면99a


0 일
1 일
2 일
14 일
18 일
Prestim
16 시간
형질도입
MOI 4000
24 시간
해동
CD34+ 세포
Methocult
적혈구 분화
%HBF
HPLC
사이토스핀
T7EI (단일 콜로니 및 풀링)
%CD235, %HBF
HPLC
사이토스핀
T7EI
그룹
1. UNTD
2. TXD_CRISPR-듀얼
Thal CD34+ 세포
1. P04
2. P18
정상 CD34+ 세포
1. Καραμπέτσιου

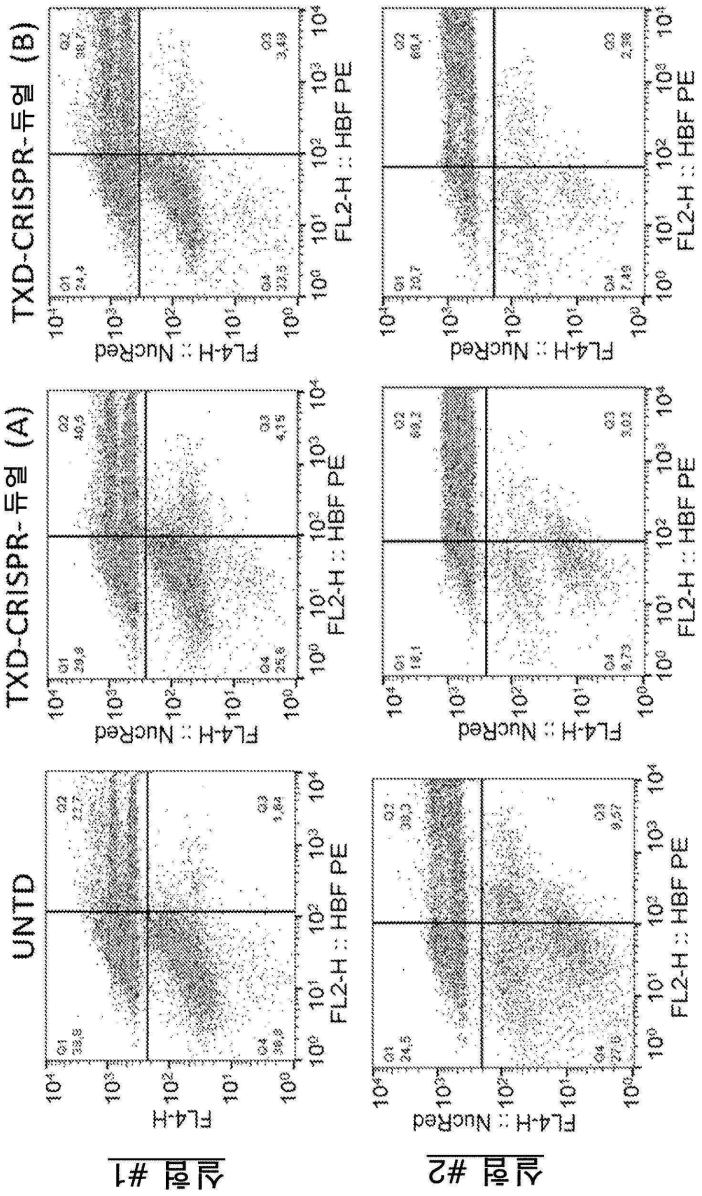
도면99b


콜로니 중 HBF %
80
60
40
20
0
적혈구
골수
UNTD
TXD-CRISPR-듀얼

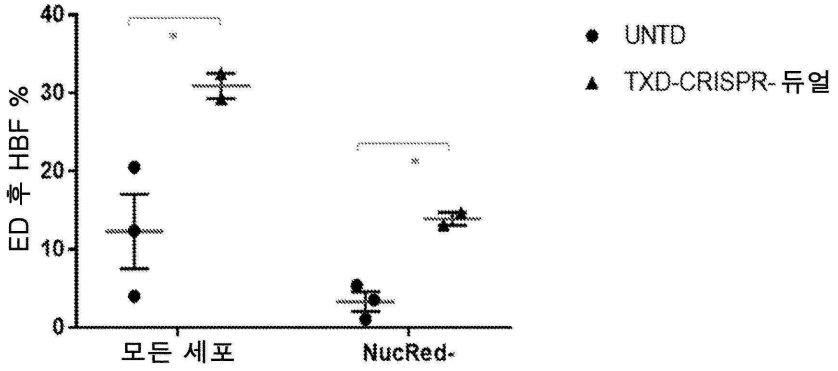
도면99c


500
400
300
200
100
0
콜로니 중 MFI
적혈구
골수
UNTD
TXD-CRISPR-듀얼

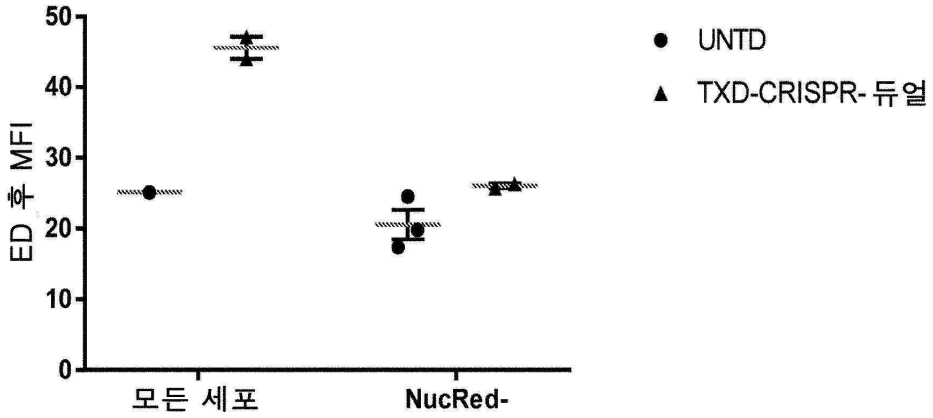
도면99d


TXD-CRISPR-듀얼 (B)
TXD-CRISPR-듀얼 (A)
UNTD
FL4-H :: NucRed
FL2-H :: HBF PE
실험 #1
실험 #2

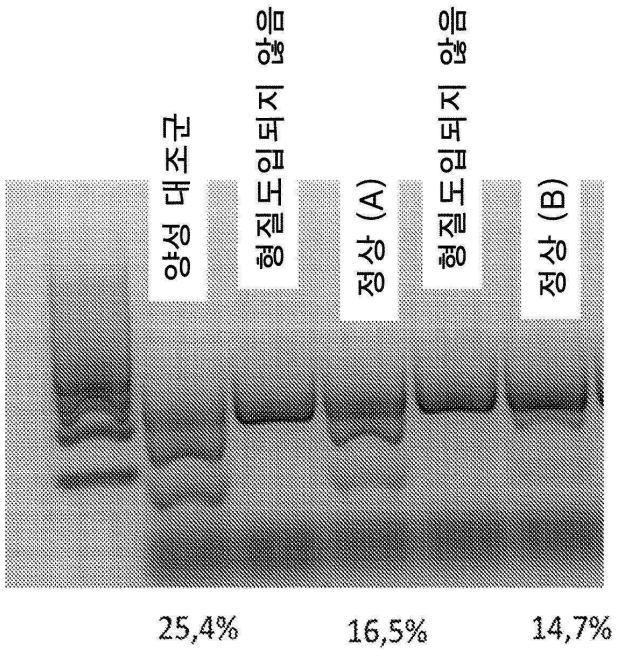
도면99e



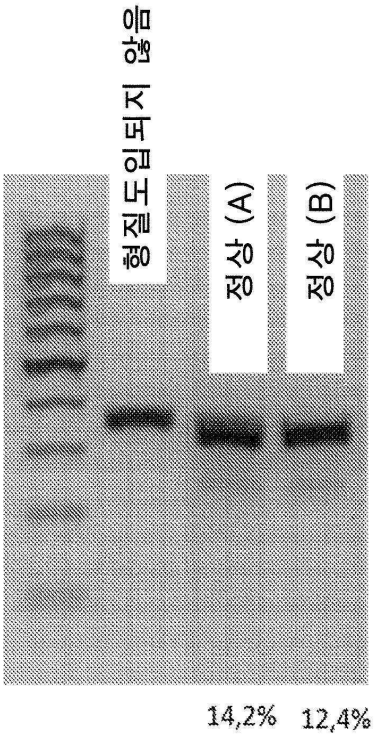

도면99f



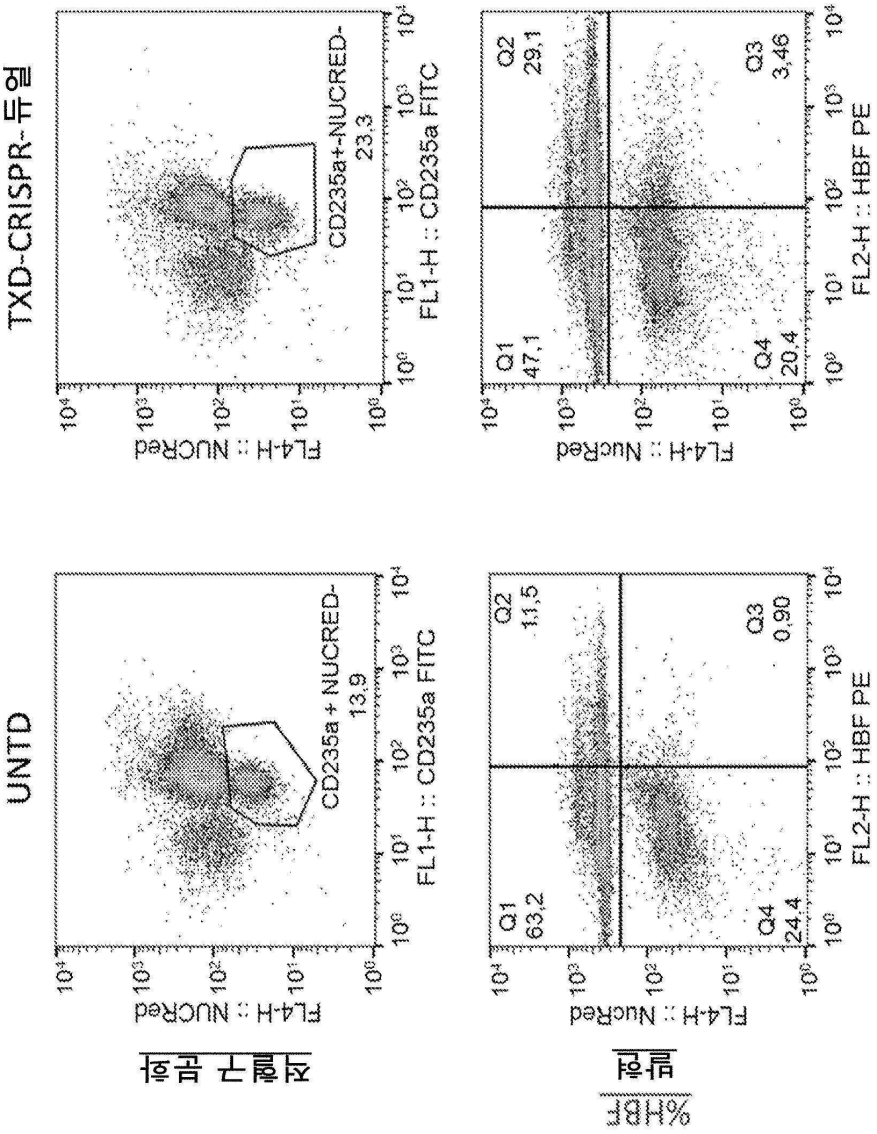

도면99g



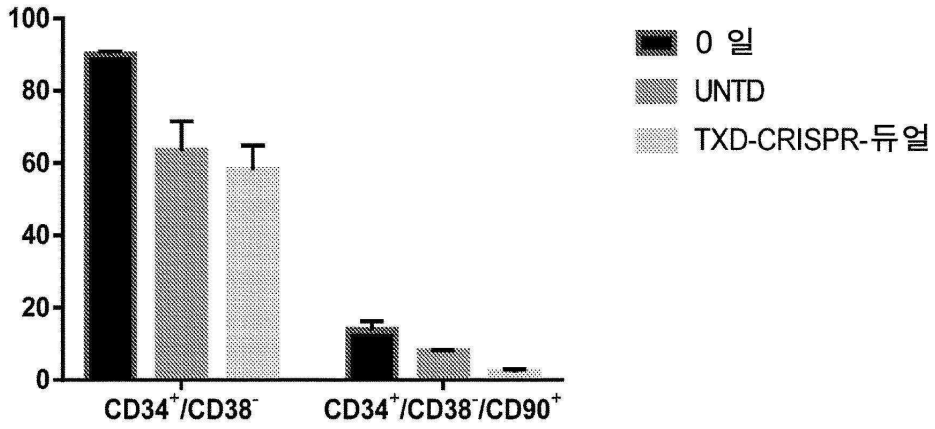
도면99h



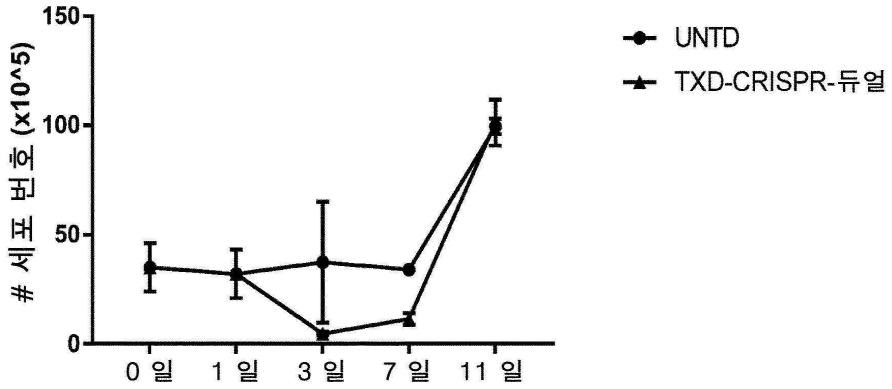
도면99i


TXD-CRISPR-듀얼
UNTD
적혈구 분화
%HBF
발현
CD235a+-NUCRED-
23.3
CD235a + NUCRED-
13.9
FL1-H :: CD235a FITC
FL4-H :: NUCRed
FL2-H :: HBF PE
Q1
47,1
Q2
29,1
Q3
3,46
Q4
20,4
Q1
63,2
Q2
11,5
Q3
0,90
Q4
24,4

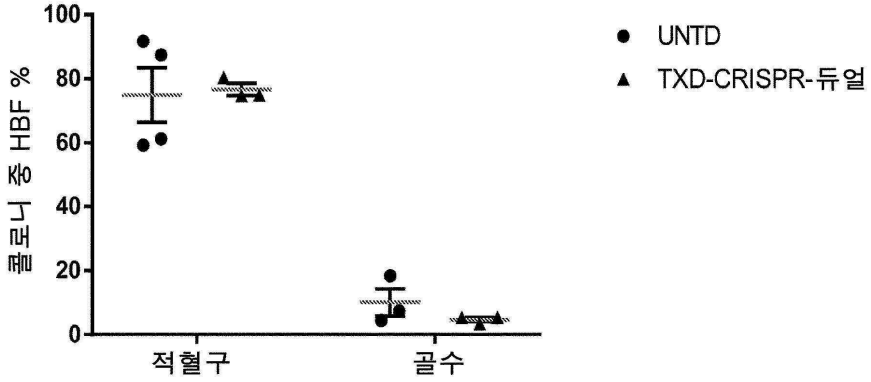
도면99j


100
80
60
40
20
0
CD34+/CD38-
CD34+/CD38-/CD90+
0 일
UNTD
TXD-CRISPR-듀얼

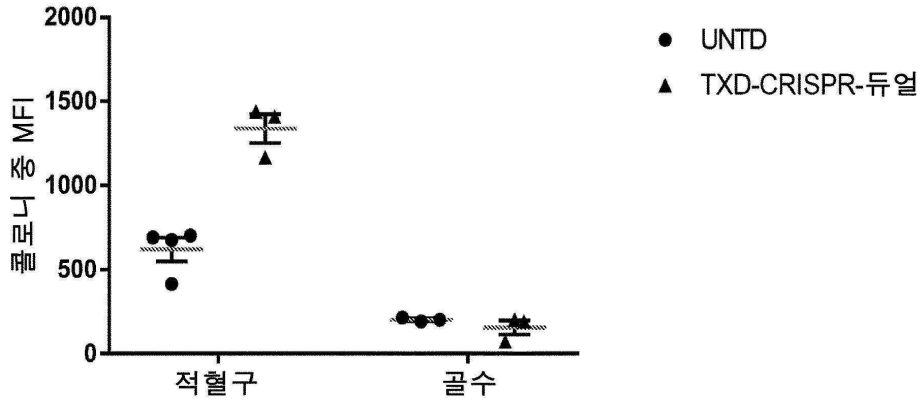
도면99k



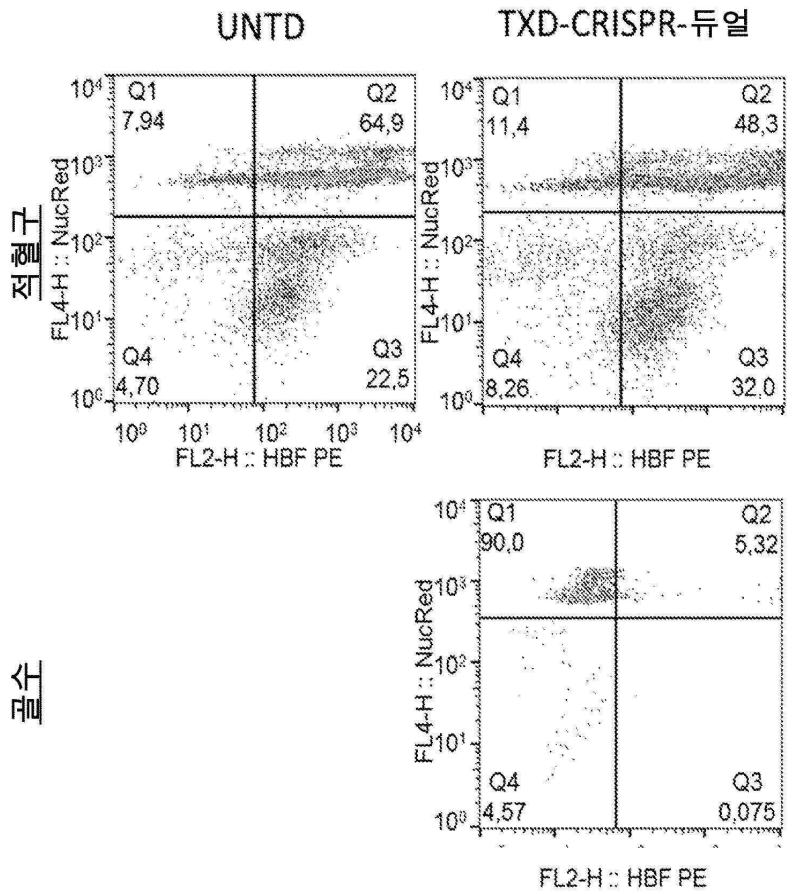

도면99l



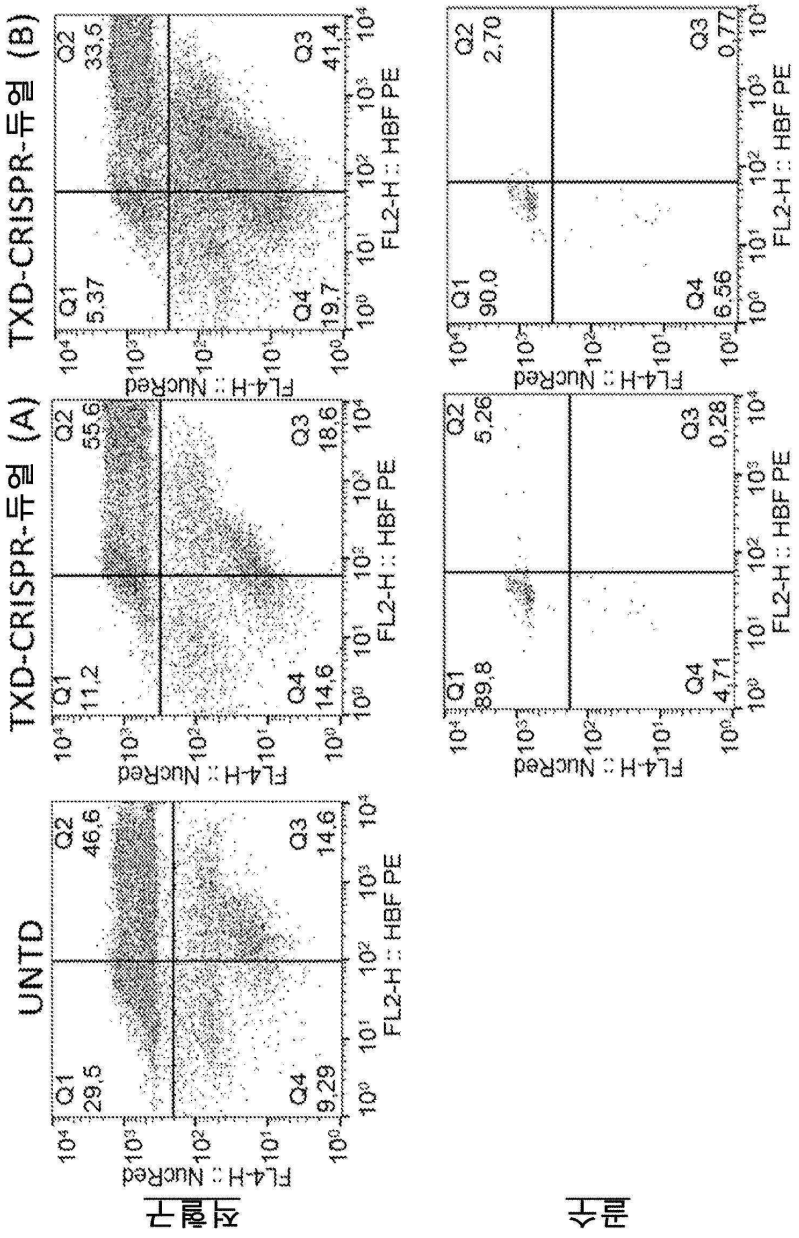

도면99m



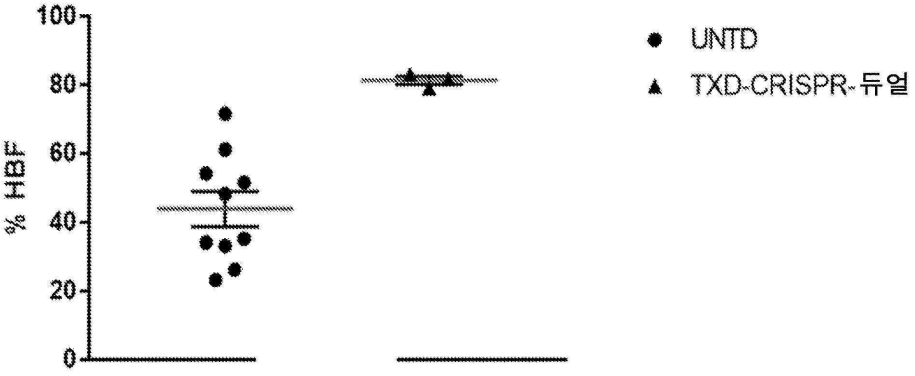
도면99n


UNTD
TXD-CRISPR-듀얼
적혈구
골수
Q1
7,94
Q2
64,9
Q4
4,70
Q3
22,5
Q1
11,4
Q2
48,3
Q4
8,26
Q3
32,0
Q1
90,0
Q2
5,32
Q4
4,57
Q3
0,075
FL4-H :: NucRed
FL2-H :: HBF PE

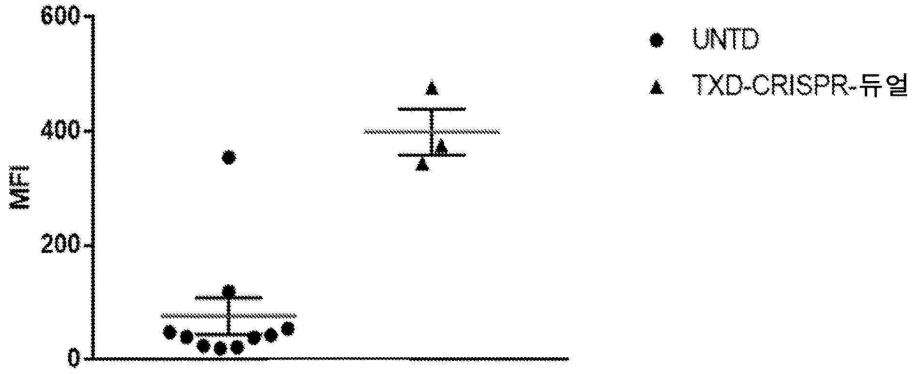
도면99o


TXD-CRISPR-듀얼 (B)
TXD-CRISPR-듀얼 (A)
UNTD
FL4-H :: NucRed
FL2-H :: HBF PE
Q1 5,37
Q2 33,5
Q3 41,4
Q4 19,7
Q1 90,0
Q2 2,70
Q3 0,77
Q4 6,56
Q1 11,2
Q2 55,6
Q3 18,6
Q4 14,6
Q1 89,8
Q2 5,26
Q3 0,28
Q4 4,71
Q1 29,5
Q2 46,6
Q3 14,6
Q4 9,29
적혈구
골수

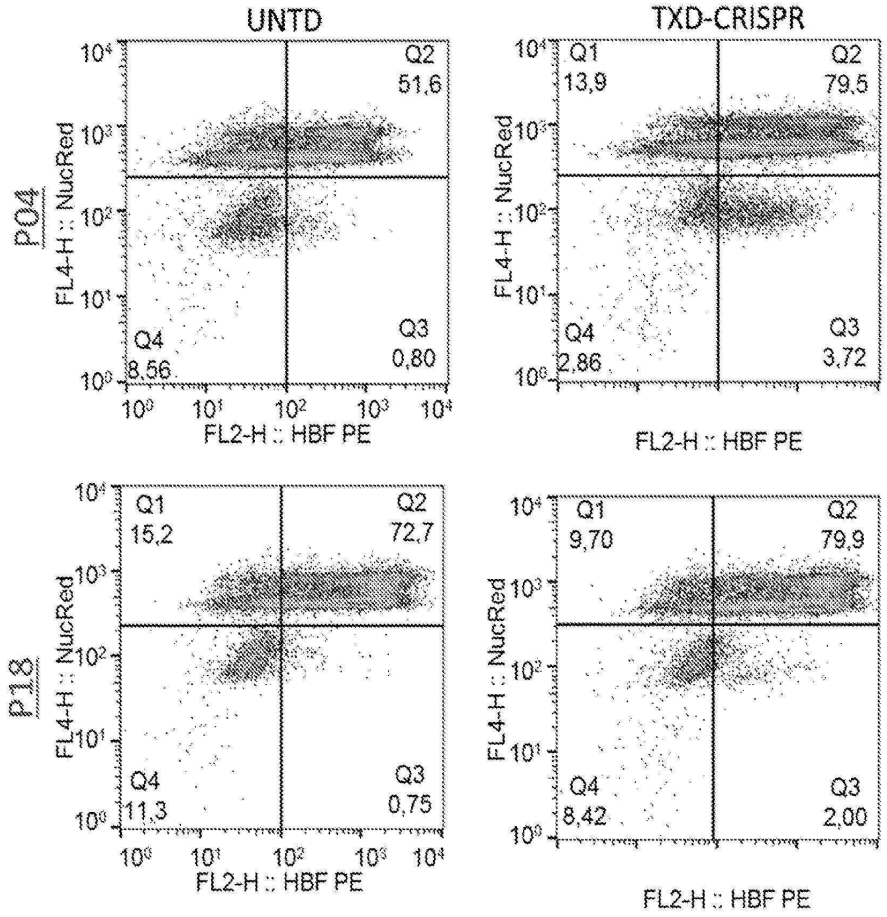
도면99p


% HBF
100
80
60
40
20
0
UNTD
TXD-CRISPR-듀얼

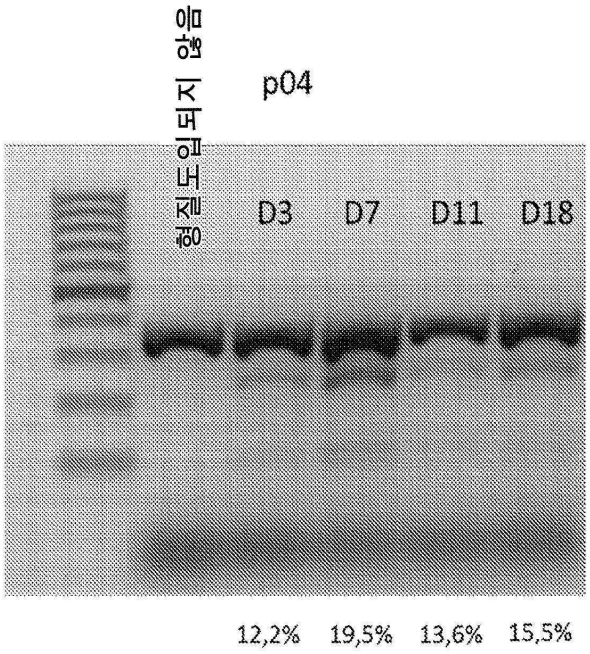
도면99q


600
400
200
0
MFI
UNTD
TXD-CRISPR-듀얼

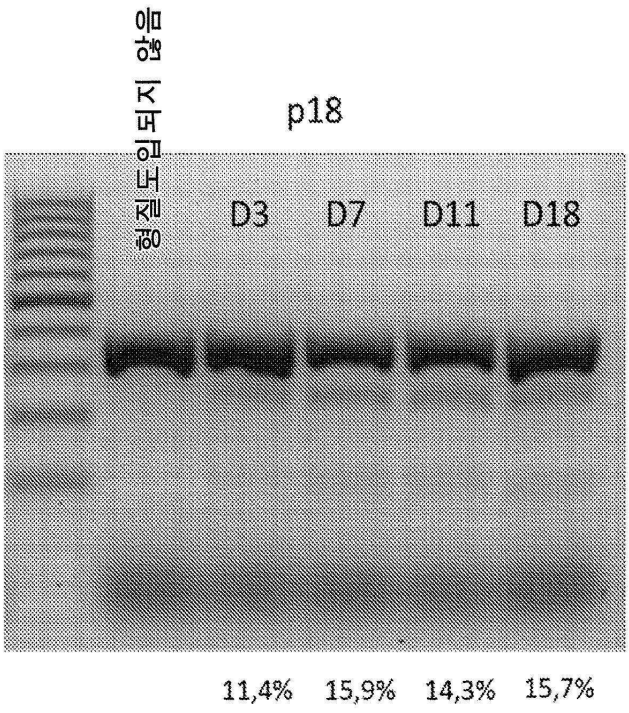
도면99r


UNTD
TXD-CRISPR
P04
P18
FL4-H :: NucRed
FL2-H :: HBF PE
Q1
13,9
Q2
51,6
79,5
Q3
0,80
3,72
Q4
8,56
2,86
15,2
72,7
11,3
0,75
9,70
79,9
8,42
2,00

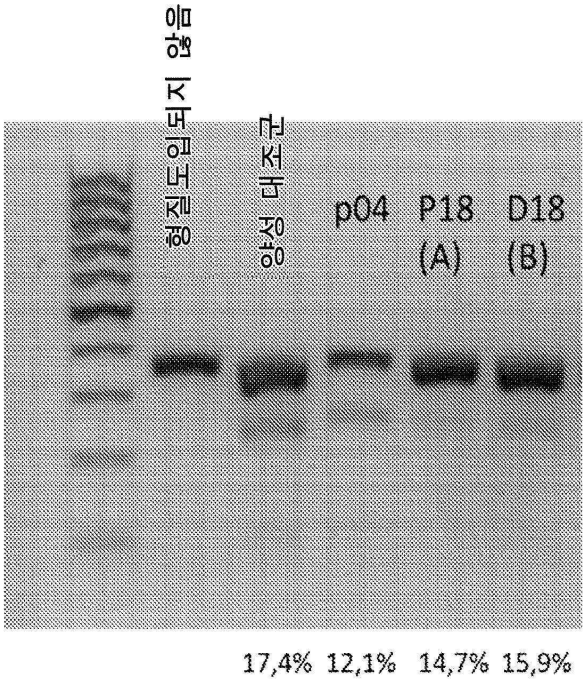
도면99s


형질도입되지 않음
p04
D3
D7
D11
D18
12,2%
19,5%
13,6%
15,5%

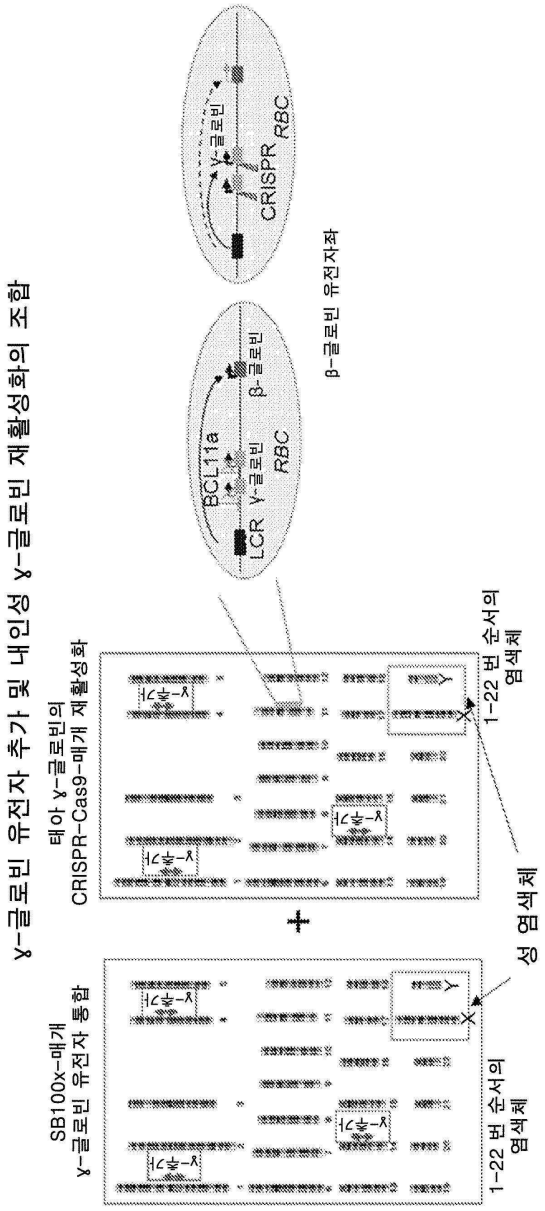
도면99t


형질도입되지 않음
p18
D3
D7
D11
D18
11,4%
15,9%
14,3%
15,7%

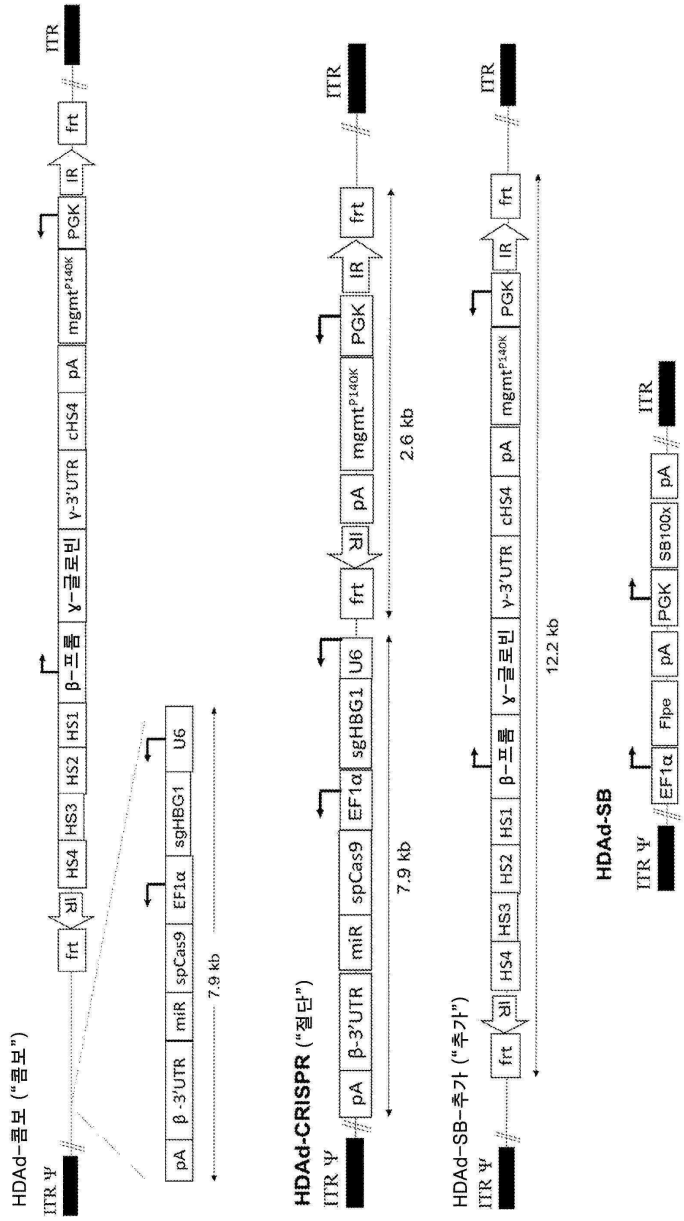
도면99u



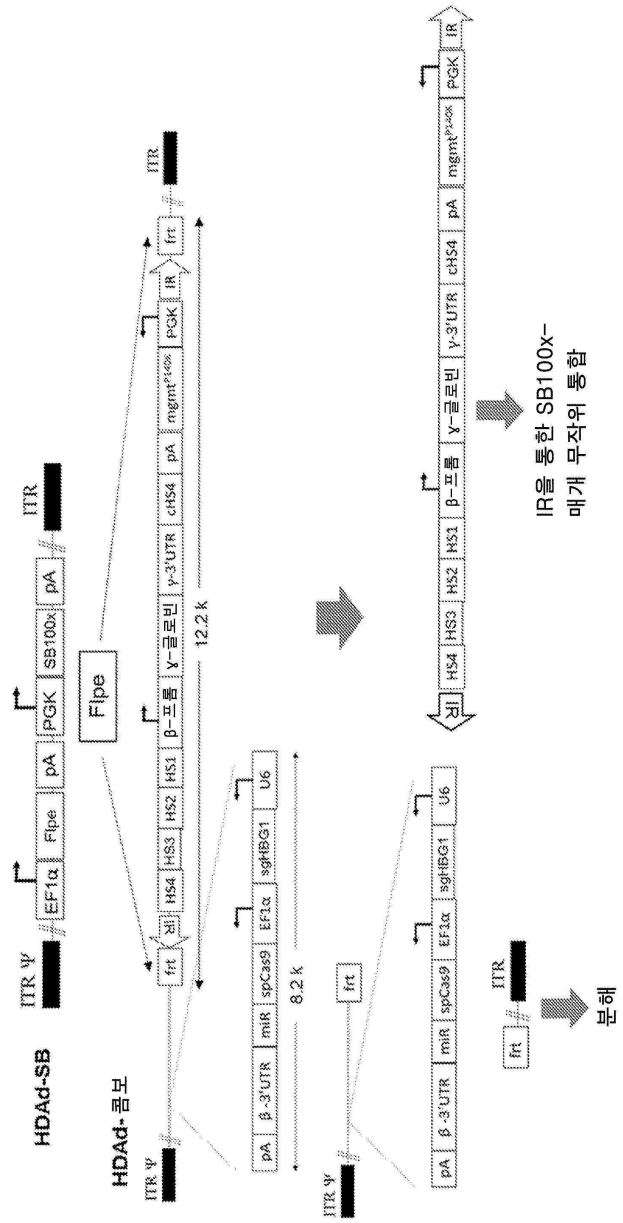
도면100


γ-글로빈 유전자 추가 및 내인성 γ-글로빈 재활성화의 조합
SB100x-매개
γ-글로빈 유전자 통합
테아 γ-글로빈의
CRISPR-Cas9-매개 재활성화
γ-글로빈
RBC
CRISPR
γ-글로빈
BCL11a
LCR
γ-글로빈
β-글로빈
RBC
β-글로빈 유전자좌
1-22 번 순서의
염색체
성 염색체
1-22 번 순서의
염색체

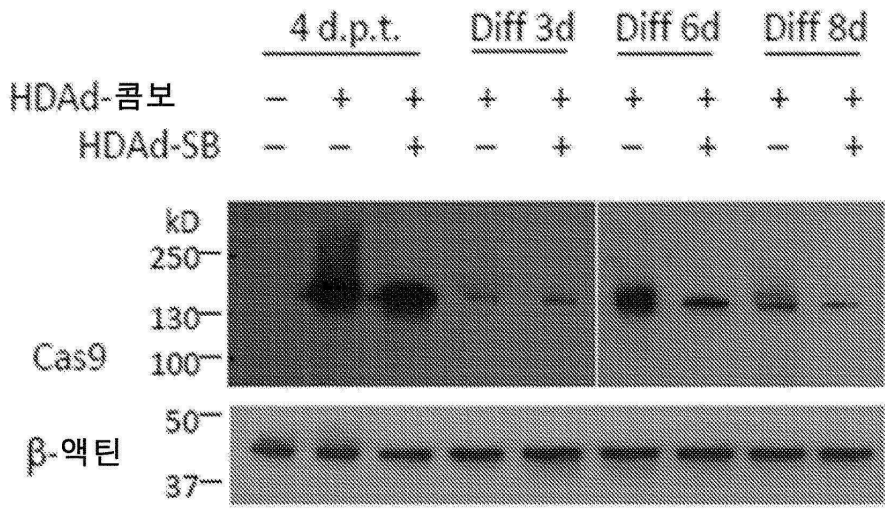
도면101


HDAd-콤보 ("콤보")
ITR Ψ
frt
처
HS4
HS3
HS2
HS1
β-프롬
γ-글로빈
γ-3'UTR
cHS4
pA
mgmt$^{P140K}$
PGK
IR
frt
ITR
pA
β-3'UTR
miR
spCas9
EF1α
sgHBG1
U6
7.9 kb
HDAd-CRISPR ("절단")
ITR Ψ
pA
β-3'UTR
miR
spCas9
EF1α
sgHBG1
U6
frt
처
pA
mgmt$^{P140K}$
PGK
IR
frt
ITR
7.9 kb
2.6 kb
HDAd-SB-추가 ("추가")
ITR Ψ
frt
처
HS4
HS3
HS2
HS1
β-프롬
γ-글로빈
γ-3'UTR
cHS4
pA
mgmt$^{P140K}$
PGK
IR
frt
ITR
12.2 kb
HDAd-SB
ITR Ψ
EF1α
Flpe
pA
PGK
SB100x
pA
ITR

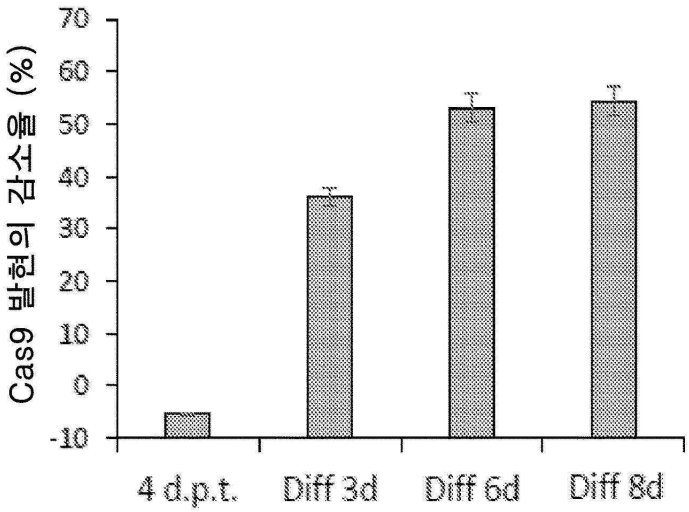
도면102


HDAd-SB
ITR Ψ
EF1α
Flpe
pA
PGK
SB100x
pA
ITR
HDAd-콤보
Flpe
frt
12.2 k
8.2 k
HS4
HS3
HS2
HS1
β-프롬
γ-글로빈
γ-3'UTR
cHS4
pA
PGK
IR
pA
β-3'UTR
miR
spCas9
EF1α
sgHBG1
U6
IR을 통한 SB100x-
매개 무작위 통합
분해

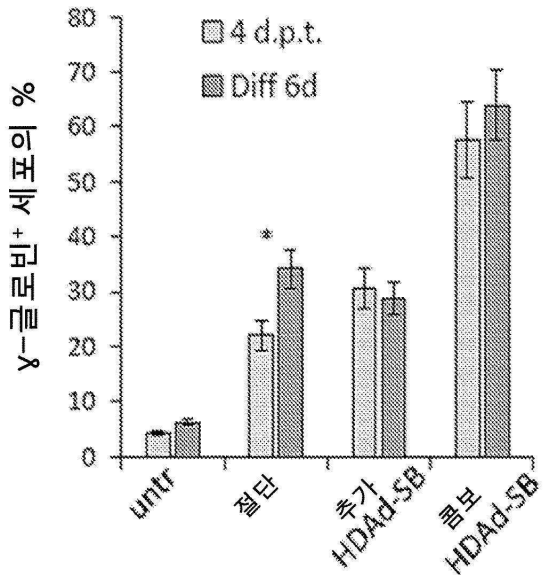
도면103a



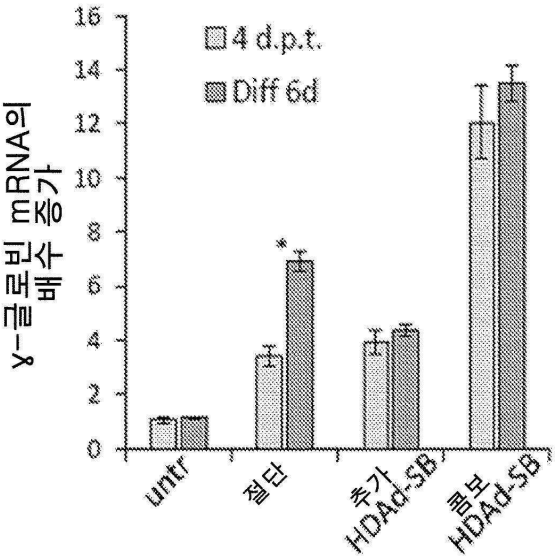

도면103b



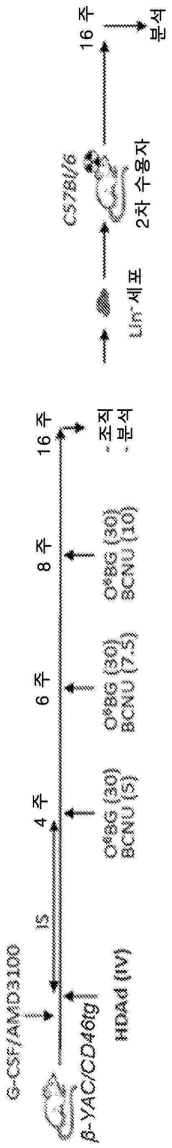
도면103c



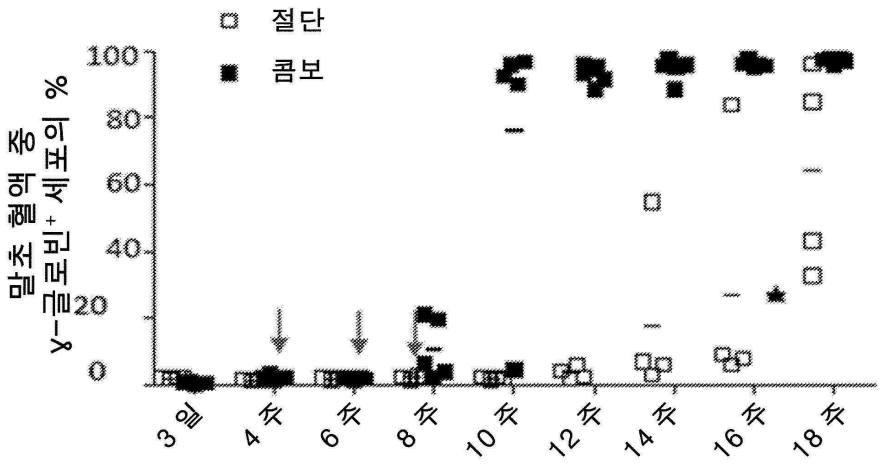

도면103d



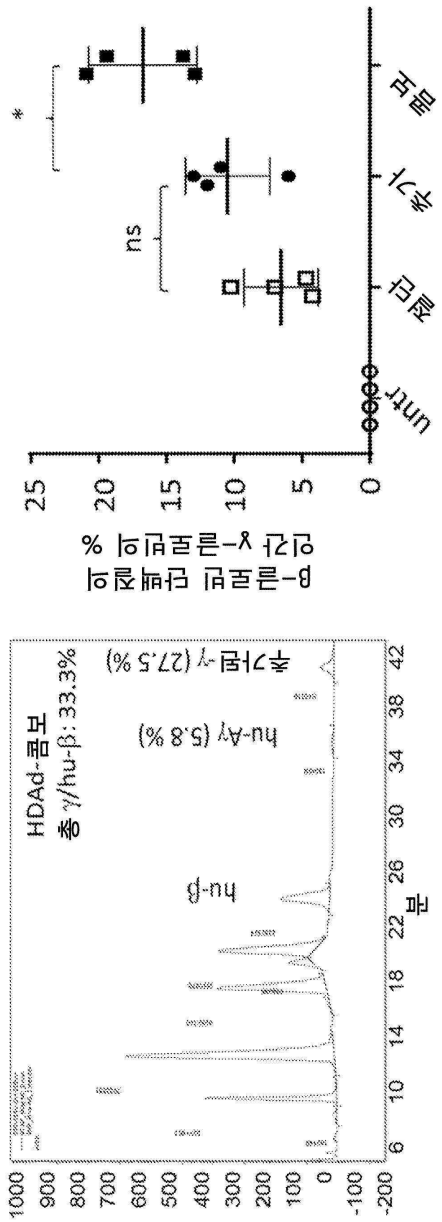
도면104a




도면104b

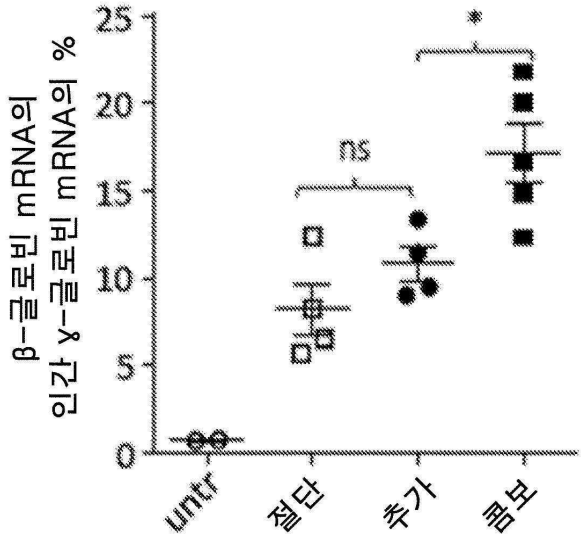


도면104c

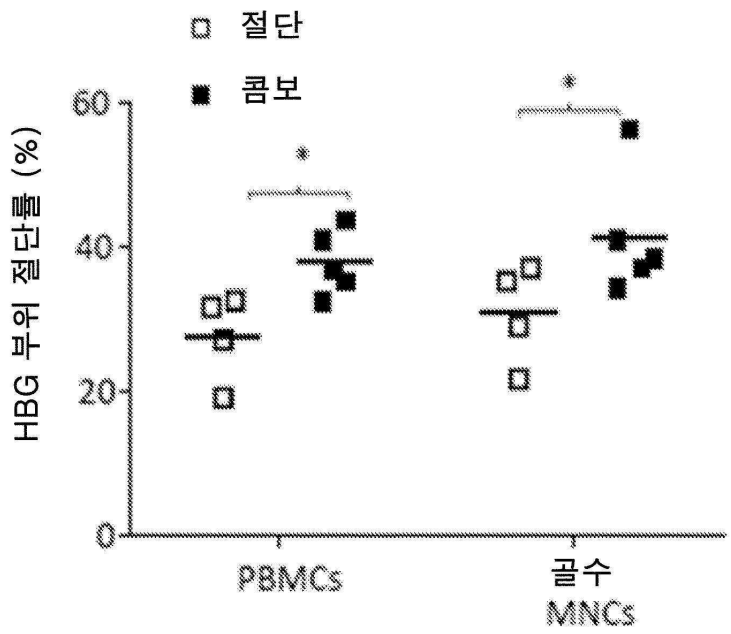

β-글로빈 단백질의
인간 γ-글로빈의 %
0
5
10
15
20
25
unTr
혈단
추가
혈모
ns
*


HDAd-콤보
총 γ/hu-β: 33.3%
hu-β
hu-Aγ (5.8 %)
추가된-γ (27.5 %)
분
6
10
14
18
22
26
30
34
38
42
1000
900
800
700
600
500
400
300
200
100
0
-100
-200

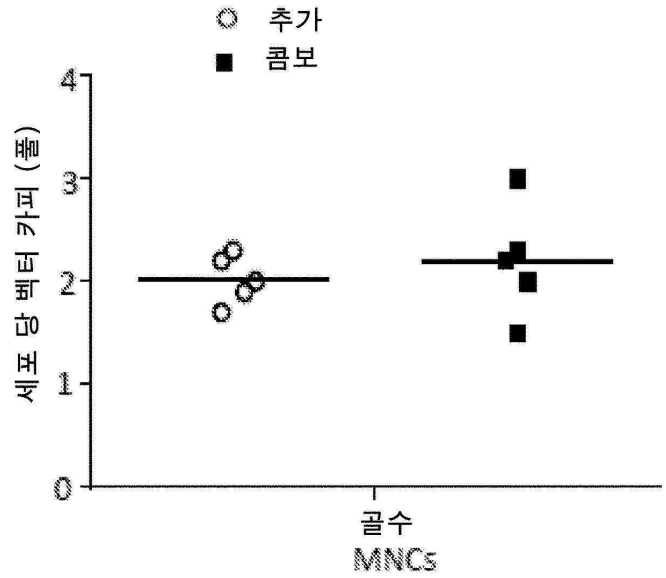
도면104d



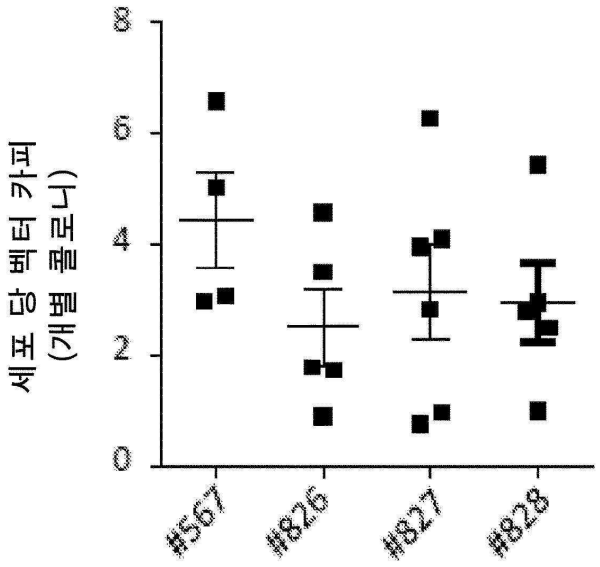

도면104e



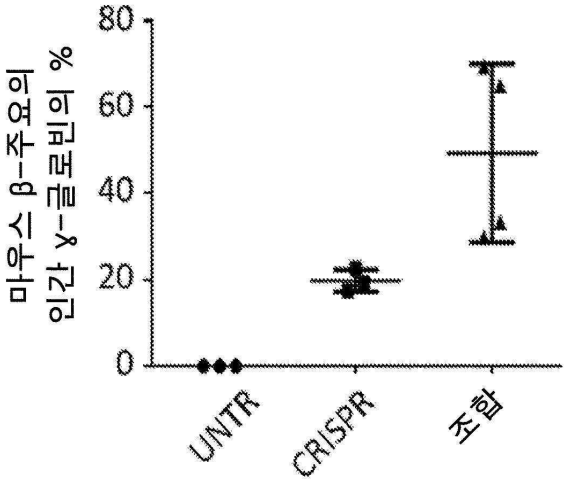
도면104f



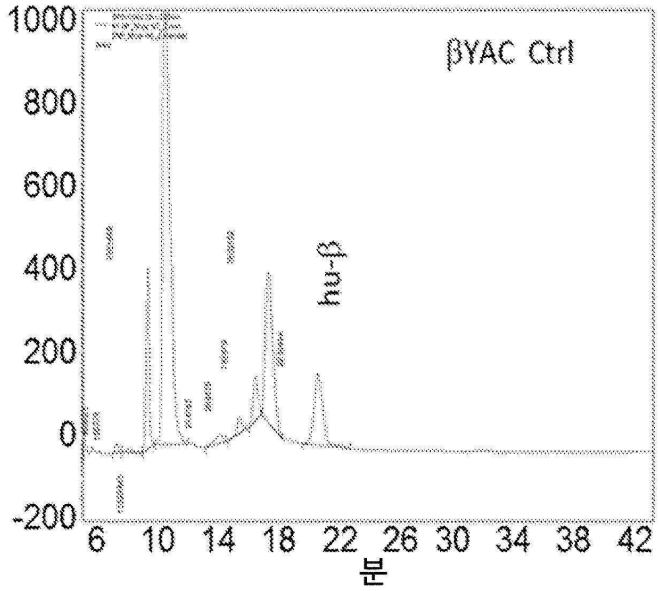

도면104g



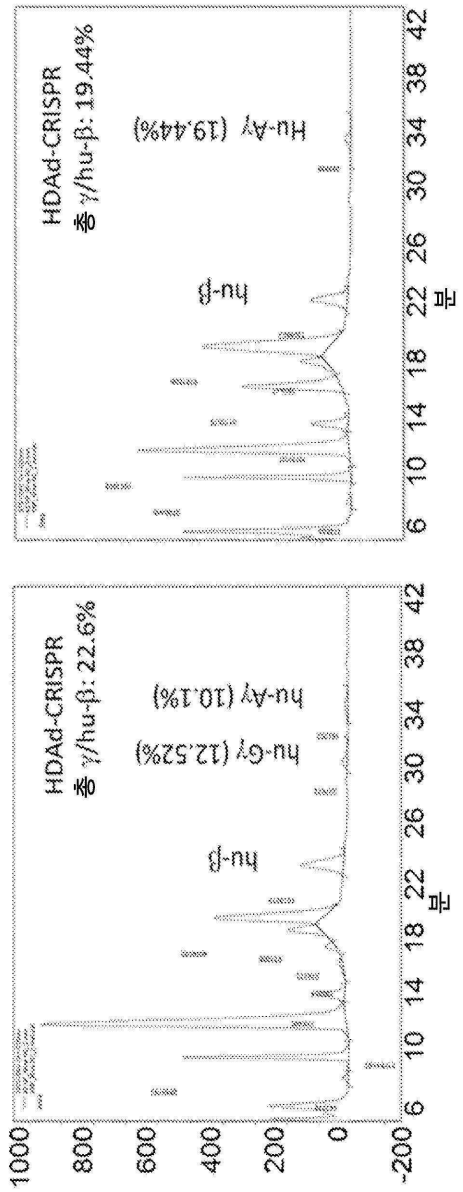
도면104h



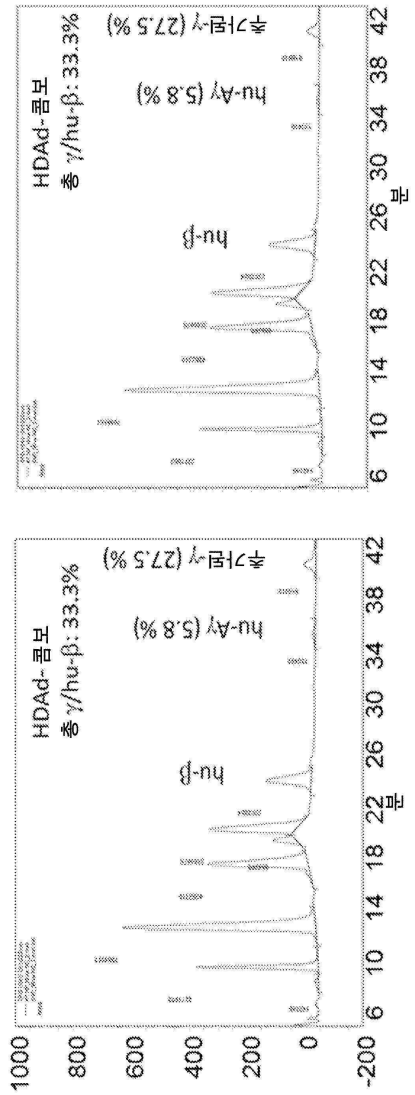

도면105a



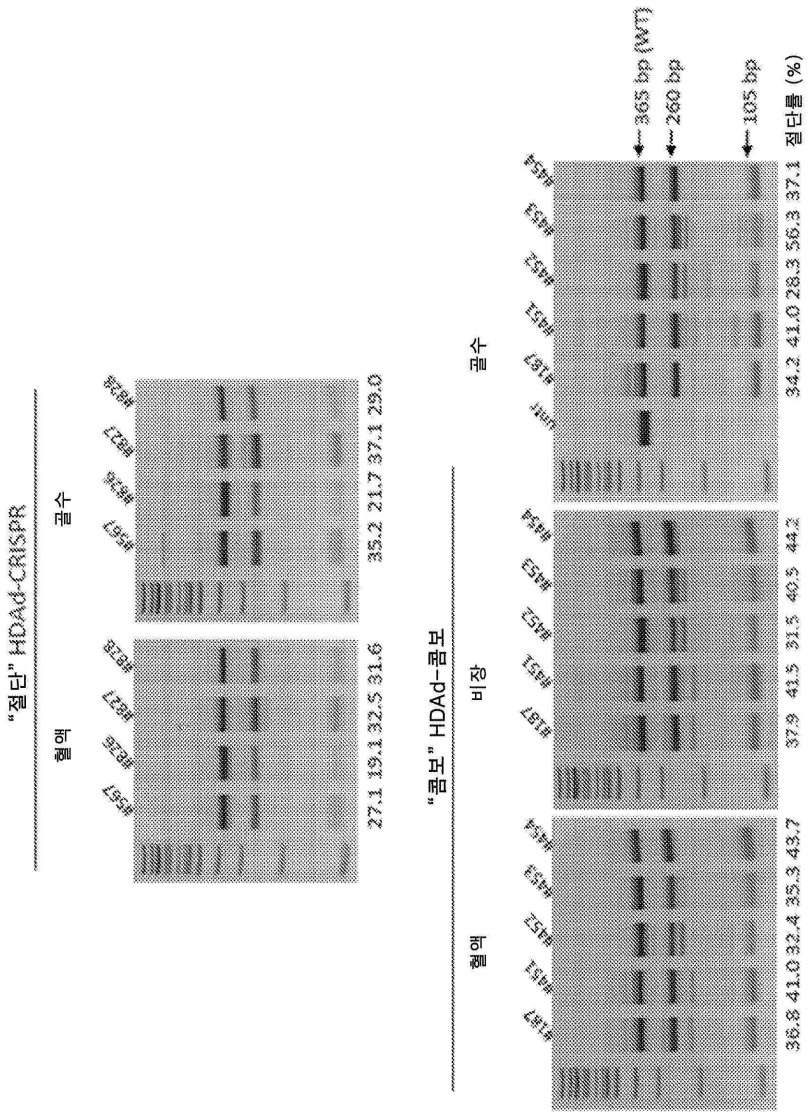
도면105b


HDAd-CRISPR
총 γ/hu-β: 19.44%
Hu-Aγ (19.44%)
hu-β
HDAd-CRISPR
총 γ/hu-β: 22.6%
hu-Gγ (12.52%)
hu-Aγ (10.1%)
hu-β
분
6
10
14
18
22
26
30
34
38
42
1000
800
600
400
200
0
-200

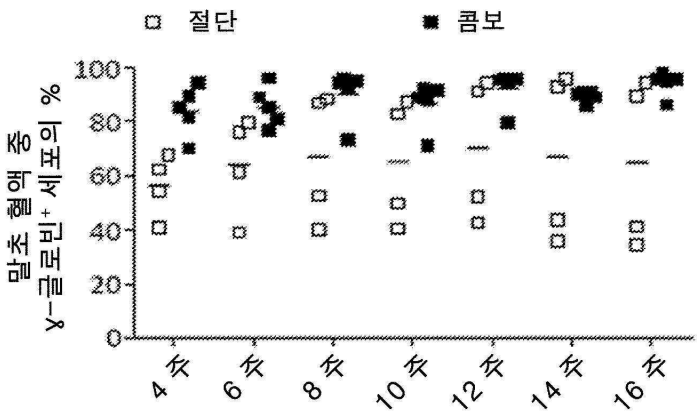
도면105c


HDAd-콤보
총 γ/hu-β: 33.3%
hu-β
hu-Aγ (5.8 %)
추가된-γ (27.5 %)
분
1000
800
600
400
200
0
-200
6
10
14
18
22
26
30
34
38
42

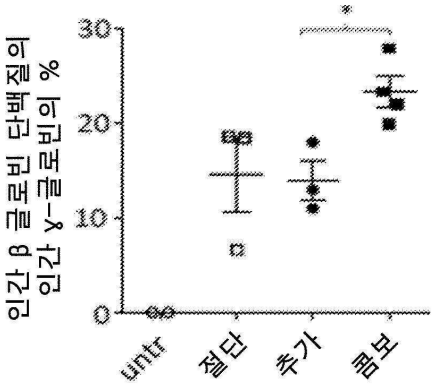
도면106



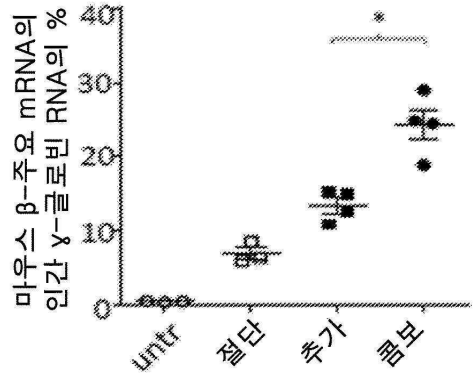

도면107a



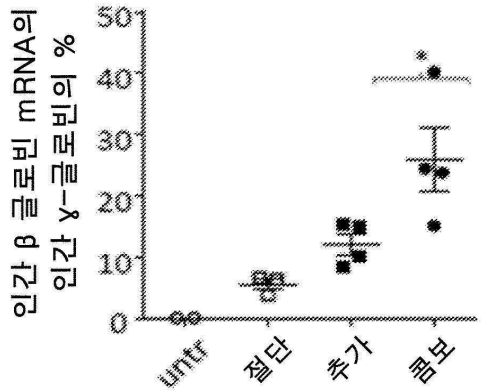
도면107b



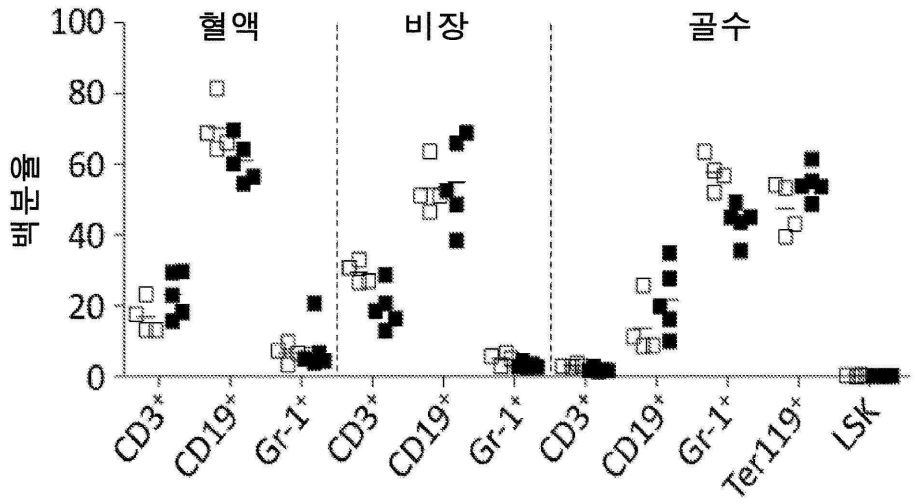

도면107c



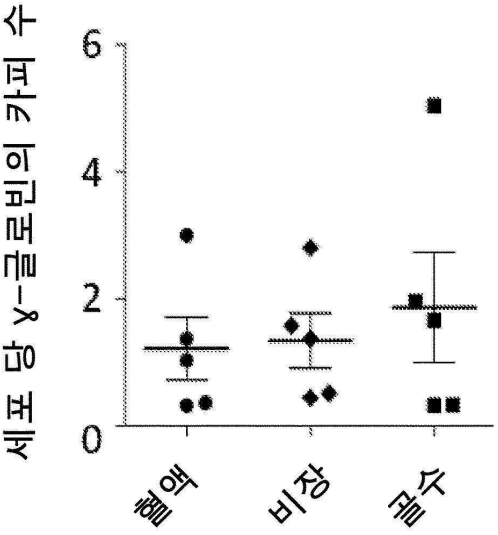

도면107d



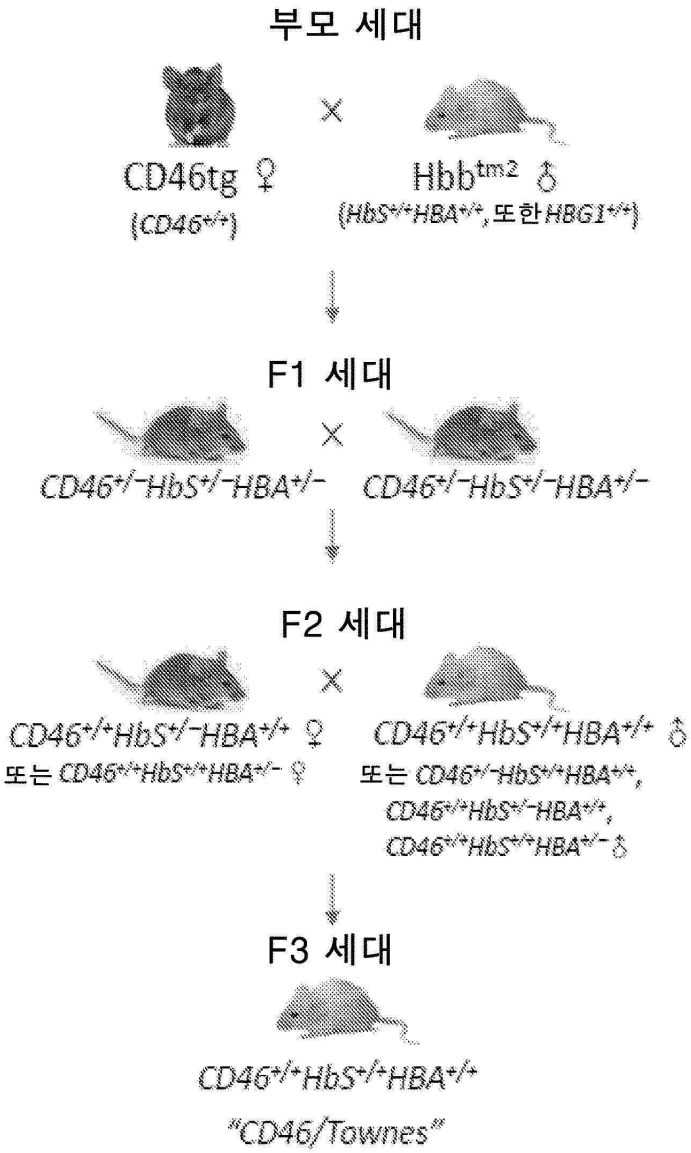
도면107e



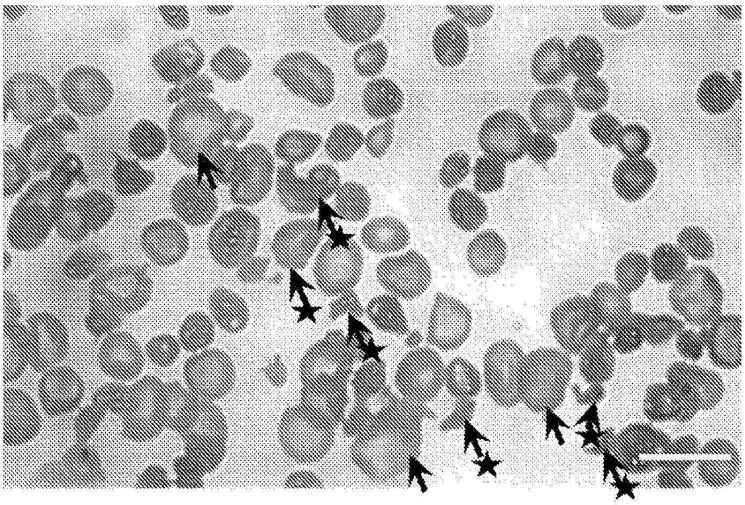

도면107f



도면108a

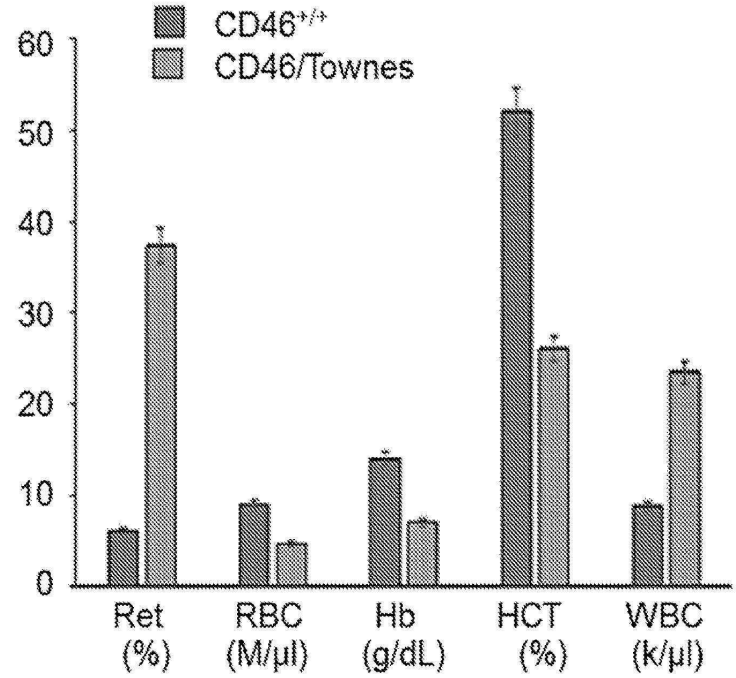



도면108b

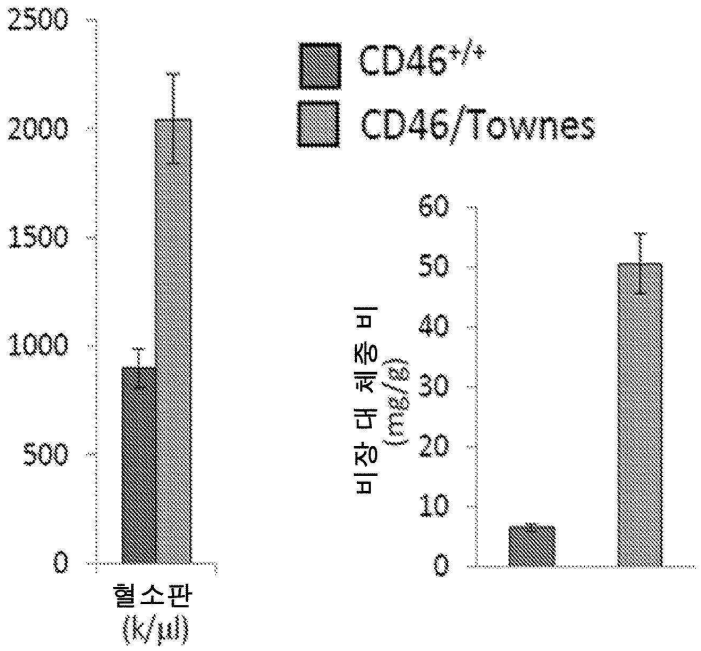


스케일 바는
15 μm이다

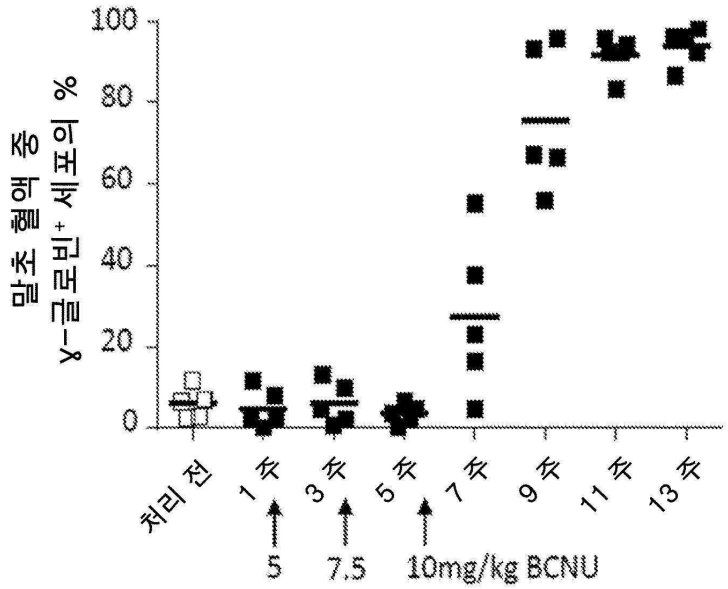
도면108c



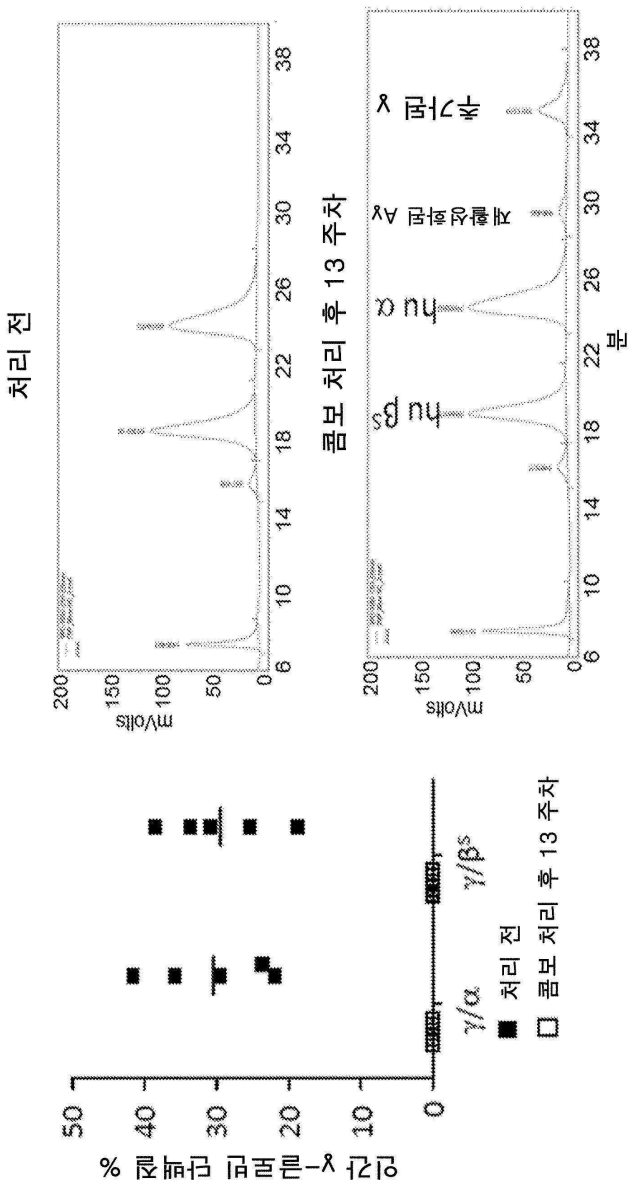

도면108d



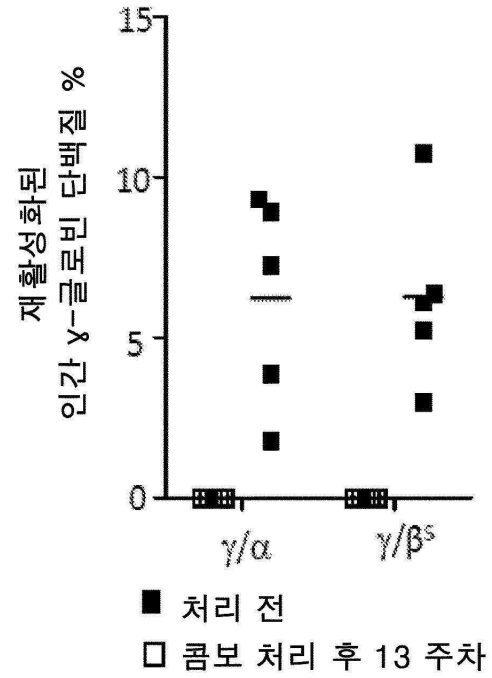
도면109a


말초 혈액 중 γ-글로빈+ 세포의 %
100
80
60
40
20
0
처리 전
1 숫
3 숫
5 숫
7 숫
9 숫
11 숫
13 숫
5
7.5
10mg/kg BCNU

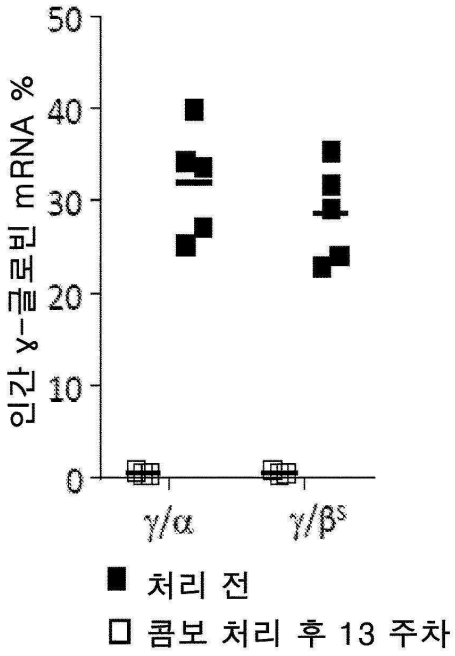
도면109b



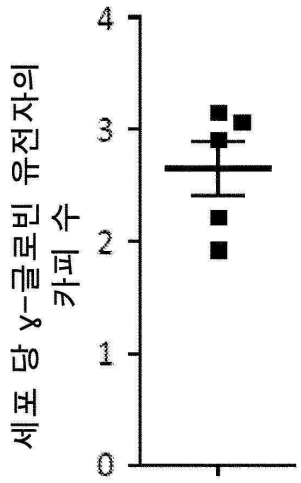
도면109c



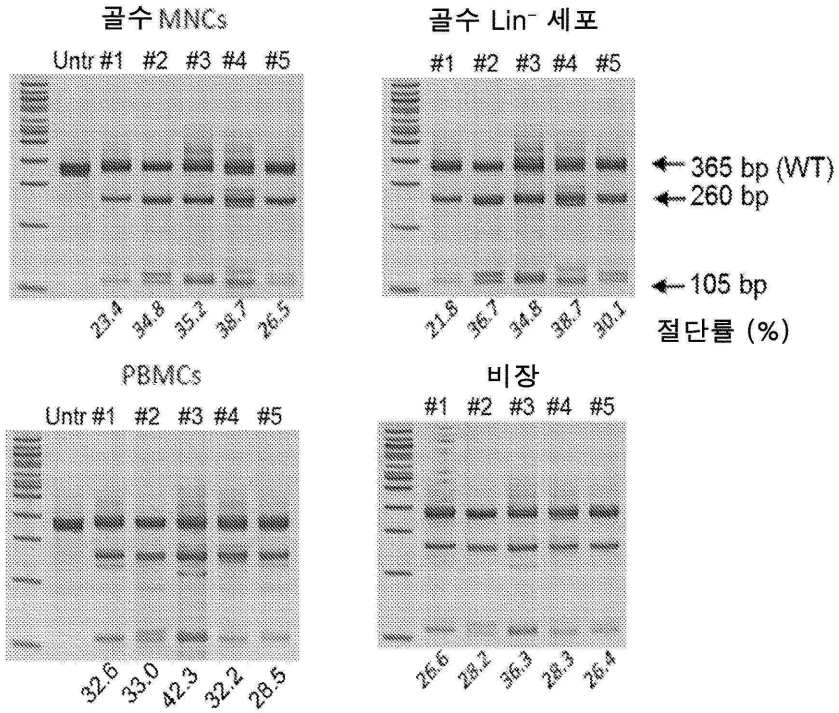

도면109d



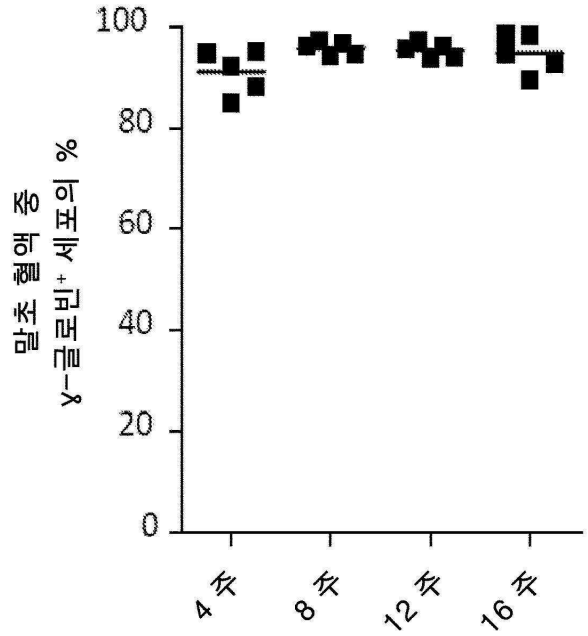
도면109e



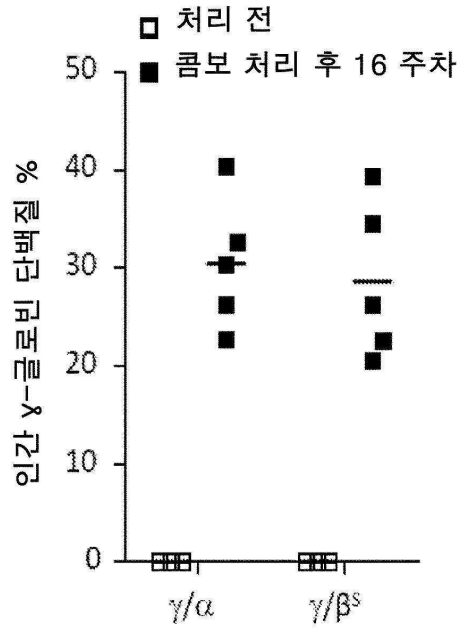

도면109f



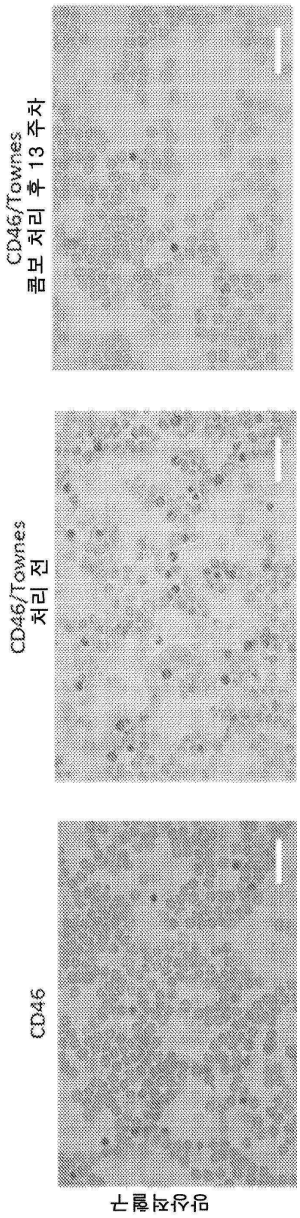
도면110a



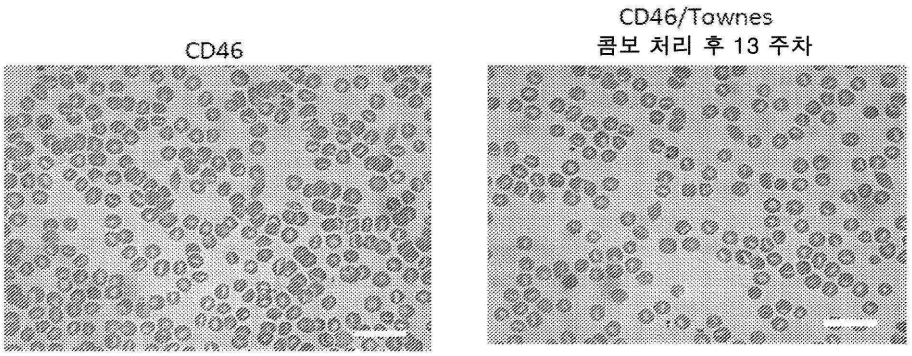

도면110b



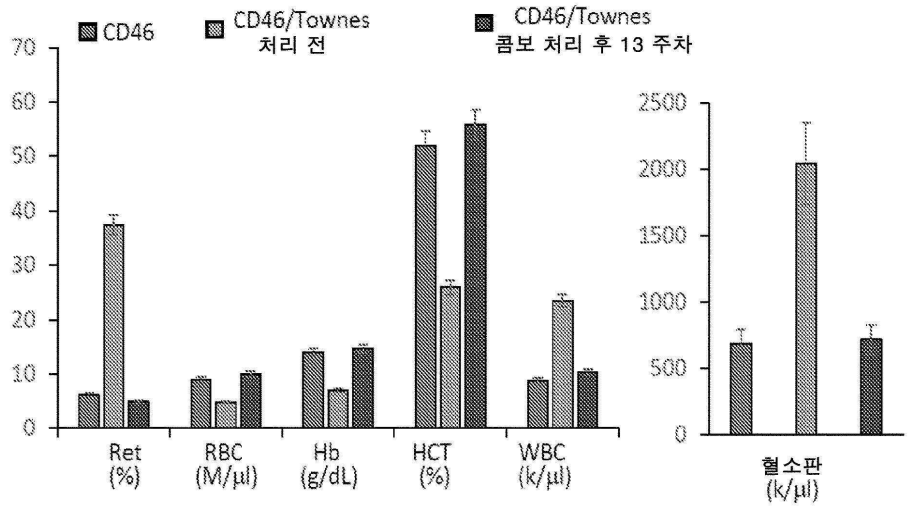
도면111a



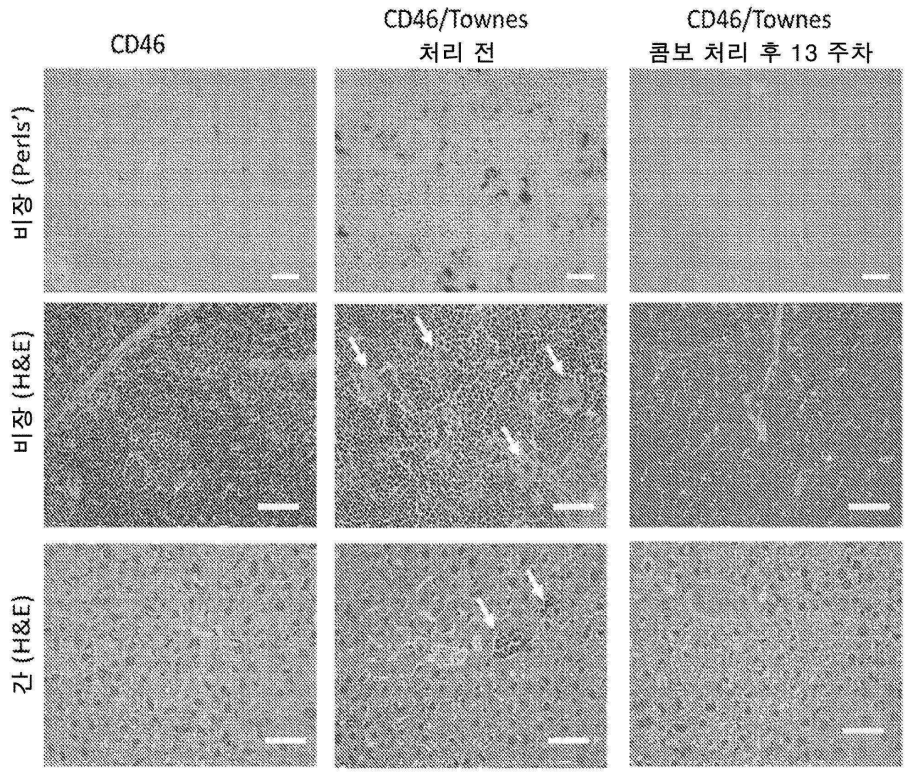

도면111b



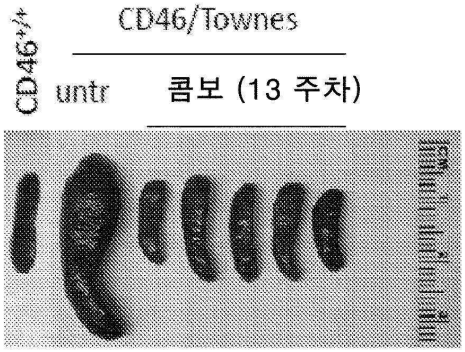
도면111c



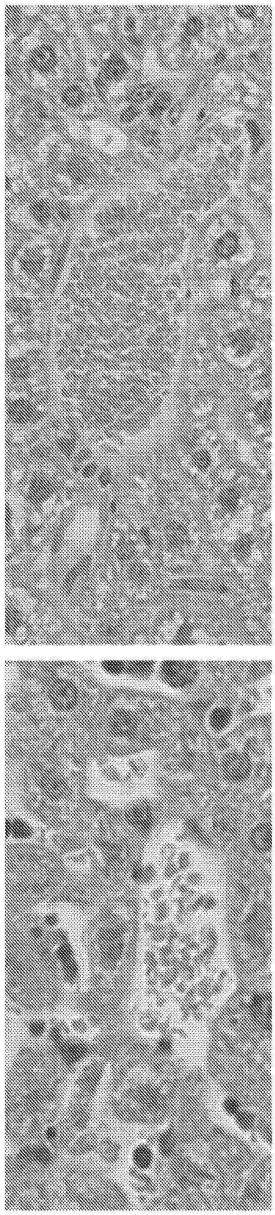

도면112a



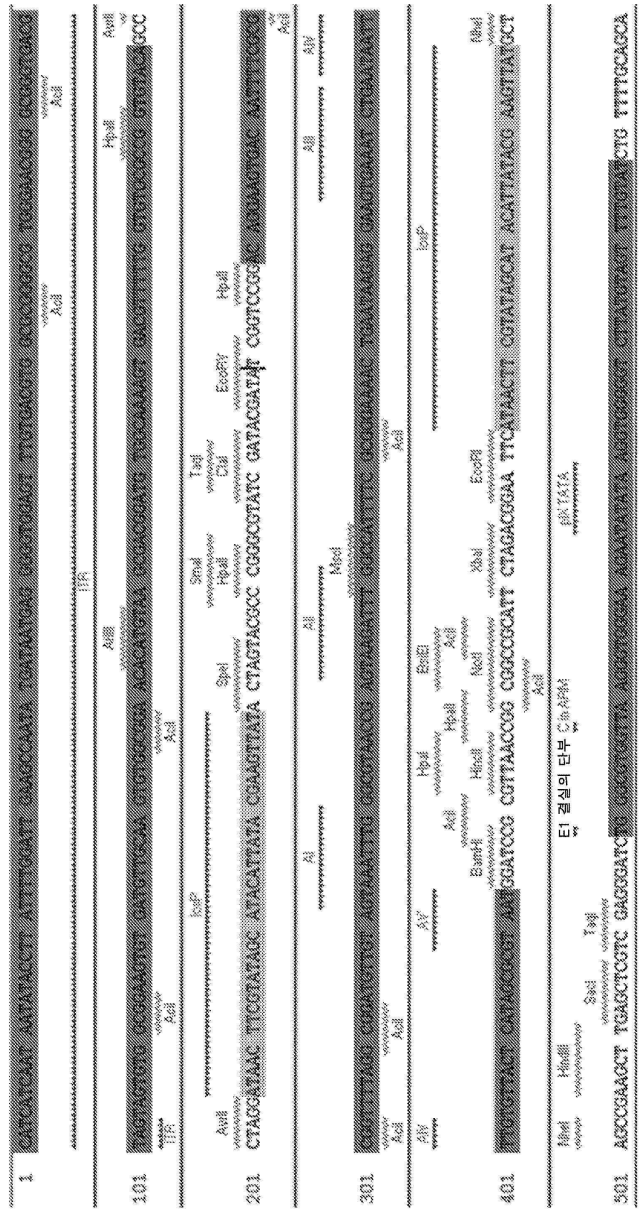
도면112b




도면112c

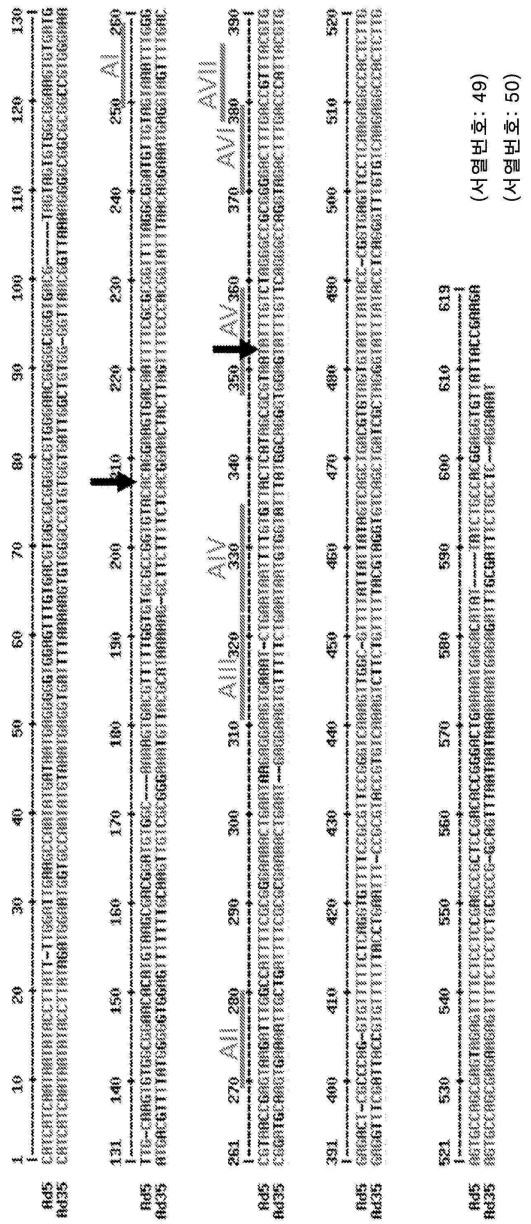


도면113



서열번호: 262

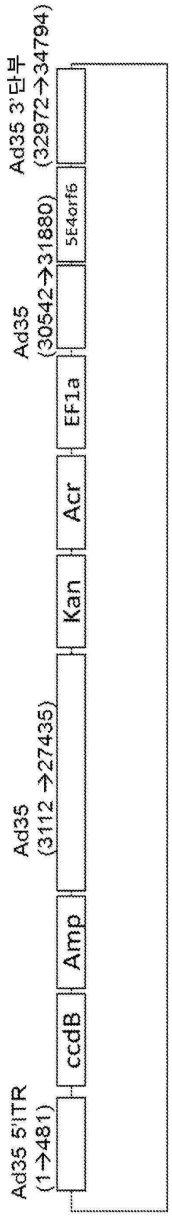
도면114



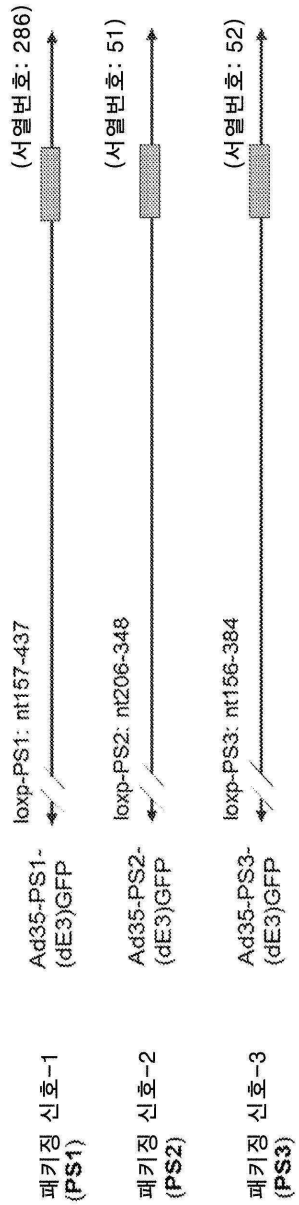
(서열번호: 49)

(서열번호: 50)

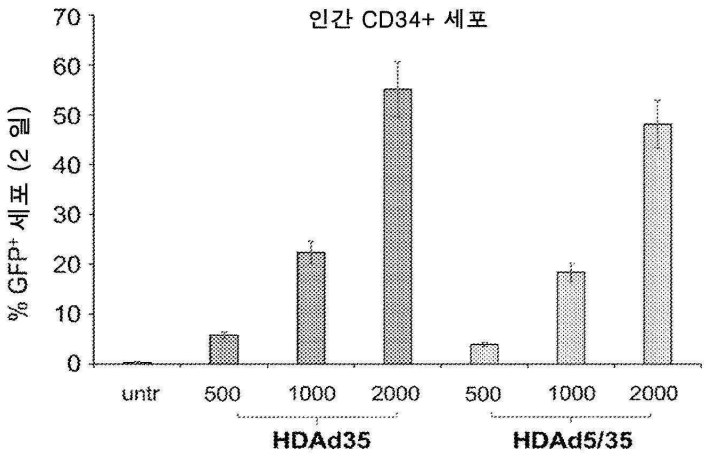
도면115


Ad35 5'ITR
(1→481)
ccdB
Amp
Ad35
(3112 →27435)
Kan
Acr
EF1a
Ad35
(30542→31880)
5E4orf6
Ad35 3'단부
(32972→34794)

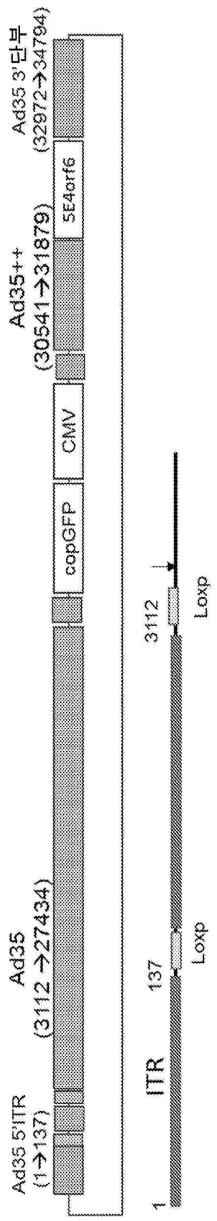
도면116


(서열번호: 286)
(서열번호: 51)
(서열번호: 52)
loxp-PS1: nt157-437
loxp-PS2: nt206-348
loxp-PS3: nt156-384
Ad35-PS1-(dE3)GFP
Ad35-PS2-(dE3)GFP
Ad35-PS3-(dE3)GFP
패키징 신호-1 (PS1)
패키징 신호-2 (PS2)
패키징 신호-3 (PS3)

도면117


인간 CD34+ 세포
% GFP+ 세포 (2 일)
70
60
50
40
30
20
10
0
untr
500
1000
2000
500
1000
2000
HDAd35
HDAd5/35

도면118


Ad35 5'ITR
(1→137)
Ad35
(3112→27434)
Ad35++
(30541→31879)
Ad35 3'단부
(32972→34794)
copGFP
CMV
SE4orf6
1
ITR
137
Loxp
3112
Loxp

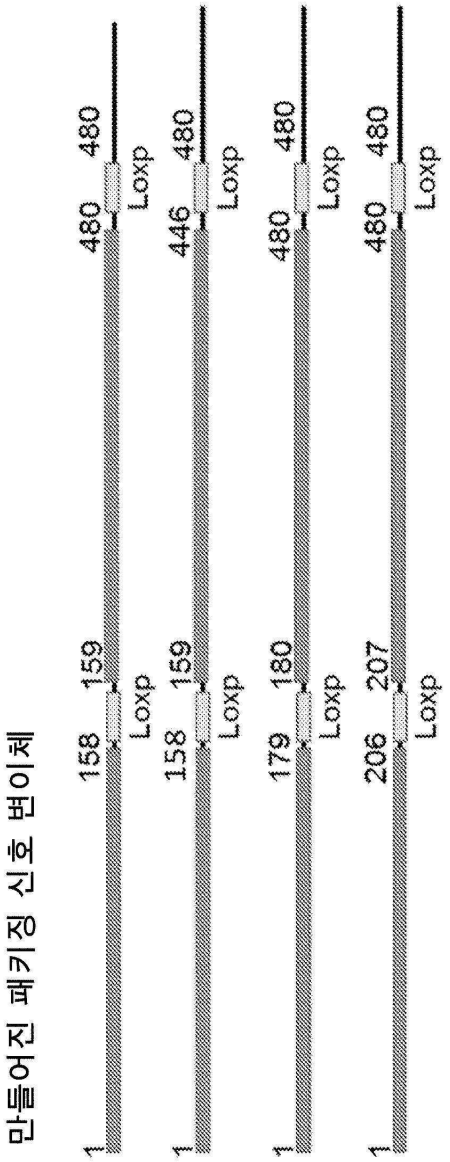
CATCATCAATAATATACCTTATAGATGGAATGGTGCCAATATGTAAATGAGGTGATTTTAAAAAGTGTGGGCCGTGTGGTGATTGG
CTGTGGGGTTAACGGTTAAAAGGGGCGGCGCGGCCGTGGGAAAATGACGTT**ATTTAAAT**ATAACTTCGTATAGCATACATTATA
CGAAGTTATGAGGTAGTTTTGTTCAGGGGCAAGTGAAAATTGACCCATTACGCGCGAAAACTGAATGAGGAAGTGTTTTTCTGAA
TAATGTGGTATTTATGGCAGGGTGGAGTATTTGACCGGATCCAGGTAGACTTTGCTGATTTTCGTGGAGGTTT*ATAACTTCGTATA*
*GCATACATTATACGAAGTTAT***ATTTAAAT** (서열번호: 264)

**Ad35 ITR 1→137bp:** (서열번호: 264의 위치 1-137)
CATCATCAATAATATACCTTATAGATGGAATGGTGCCAATATGTAAATGAGGTGATTTTAAAAAGTGTGGGCCGTGT
GGTGATTGGCTGTGGGGTTAACGGTTAAAAGGGGCGGCGCGGCCGTGGGAAAATGACGTT
SwaI 위치: (서열번호: 264의 위치 146-145 및 363-370) ATTTAAAT
Loxp 위치: (서열번호: 264의 위치 138-179 및 329-362) ATAACTTCGTATAGCATACATTATACGAAGTTAT
돌연변이된 패키징 신호: (서열번호: 264의 위치 138-145)
GAGGTAGTTTTGTTCAGGGGCAAGTGAAAATTGACCCATTACGCGCGAAAACTGAATGAGGAAGTGTTTTTCTGAAT
AATGTGGTATTTATGGCAGGGTGGAGTATTTGACCGGATCCAGGTAGACTTTGCTGATTTTCGTGGAGGTTT

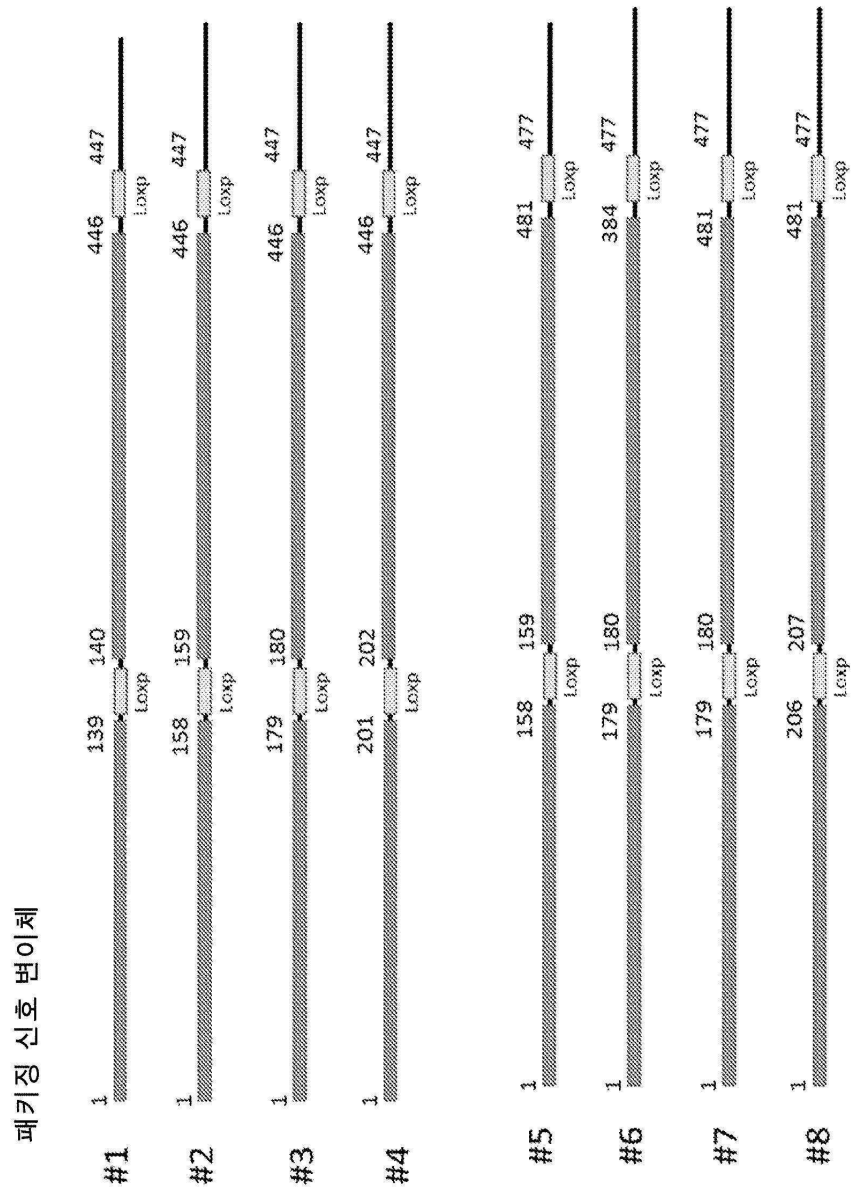


공개특허 10-2022-0038362

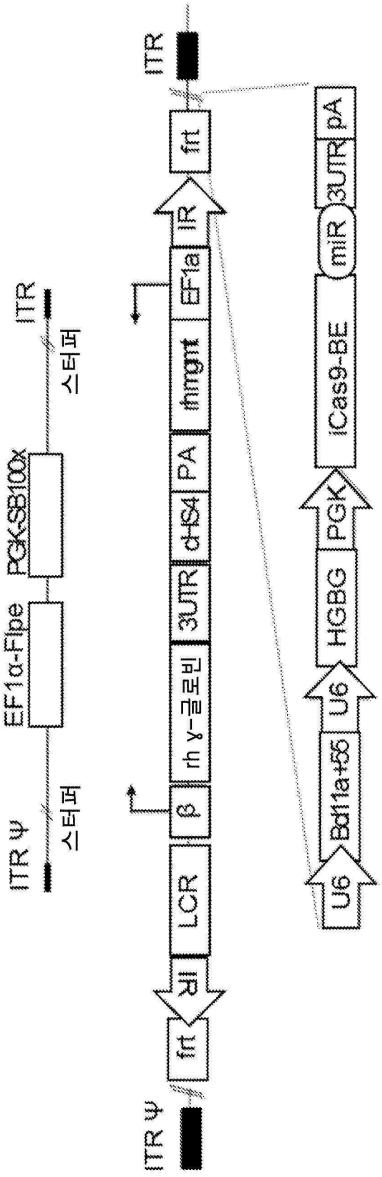
도면120a


만들어진 패키징 신호 변이체
1
158
159
Loxp
480
1
158
159
Loxp
446
480
1
179
180
Loxp
480
1
206
207
Loxp
480

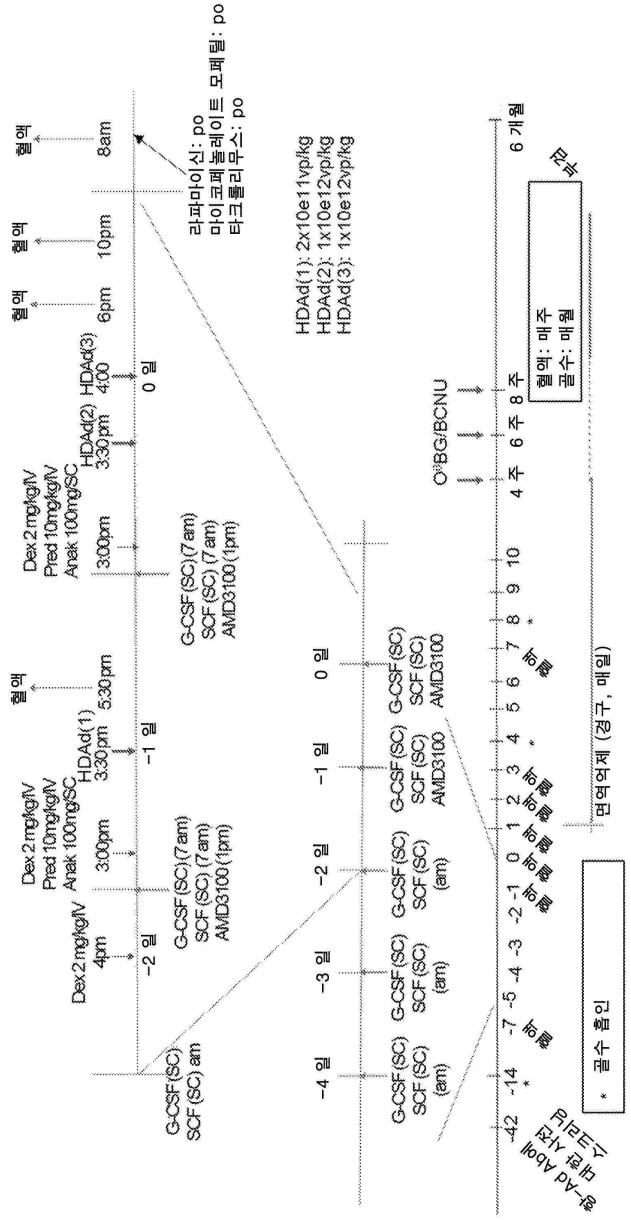
도면120b


패키징 신호 변이체
#1
#2
#3
#4
#5
#6
#7
#8
1
139
140
446
447
loxp
158
159
179
180
201
202
481
477
384
206
207

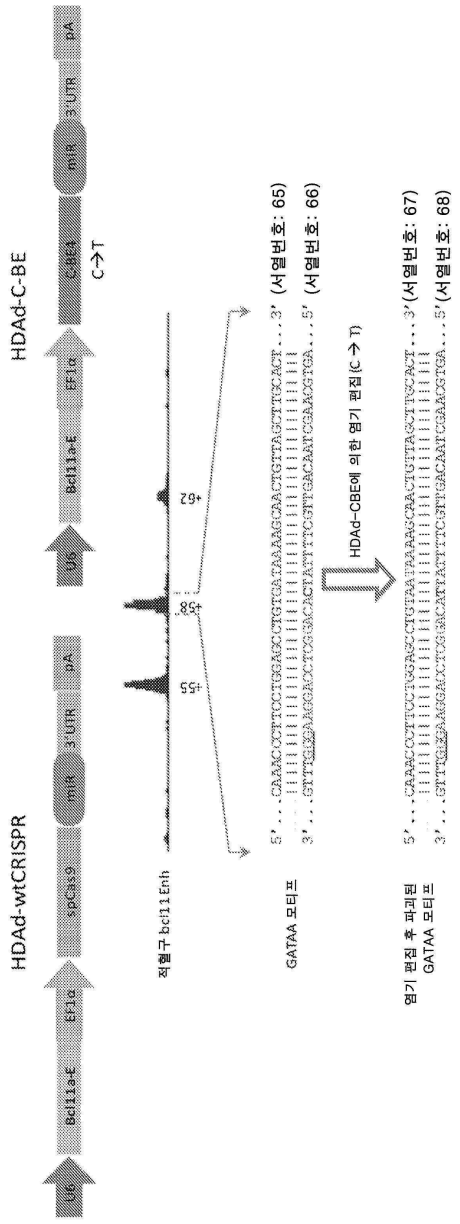
도면121


ITR Ψ
스터퍼
EF1α-Flpe
PGK-SB100x
스터퍼
ITR
ITR Ψ
frt
IR
LCR
β
rh γ-글로빈
3UTR
cHS4
PA
hmgrt
EF1a
IR
frt
ITR
U6
Bcl11a+55
U6
HGBG
PGK
iCas9-BE
miR
3UTR
pA

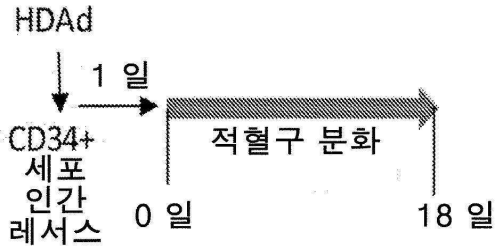
도면122


혈액
Dex 2 mg/kg/IV
Pred 10mg/kg/IV
Anak 100mg/SC
3:00pm
HDAd(1)
3:30pm
5:30pm
-1 일
-2 일
G-CSF (SC)(7 am)
SCF (SC) (7 am)
AMD3100 (1pm)
Dex 2 mg/kg/IV
4pm
G-CSF (SC)
SCF (SC) am
-4 일
-3 일
G-CSF (SC)
SCF (SC)
(am)
0 일
HDAd(2)
3:30pm
HDAd(3)
4:00
6pm
10pm
8am
라파마이신: po
마이코페놀레이트 모페틸: po
타크롤리무스: po
HDAd(1): 2x10e11vp/kg
HDAd(2): 1x10e12vp/kg
HDAd(3): 1x10e12vp/kg
O⁶BG/BCNU
-42
-14
-7
-5
-4
-3
-2
-1
0
1
2
3
4
5
6
7
8
9
10
4 주
6 주
8 주
6 개월
농출
* 골수 흡인
면역억제 (경구, 매일)
혈액: 매주
골수: 매월
끝
항-Ad Abdl
내성 시험
스크리닝

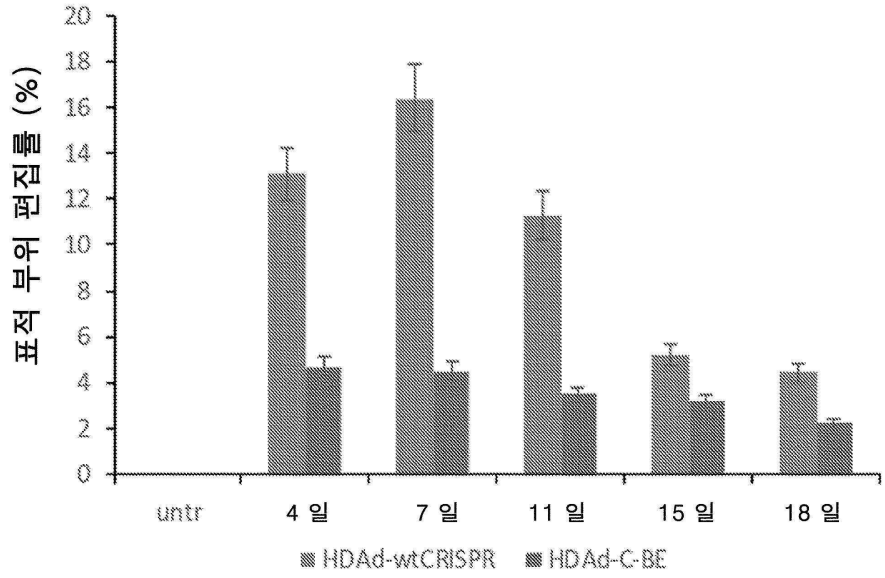
도면123



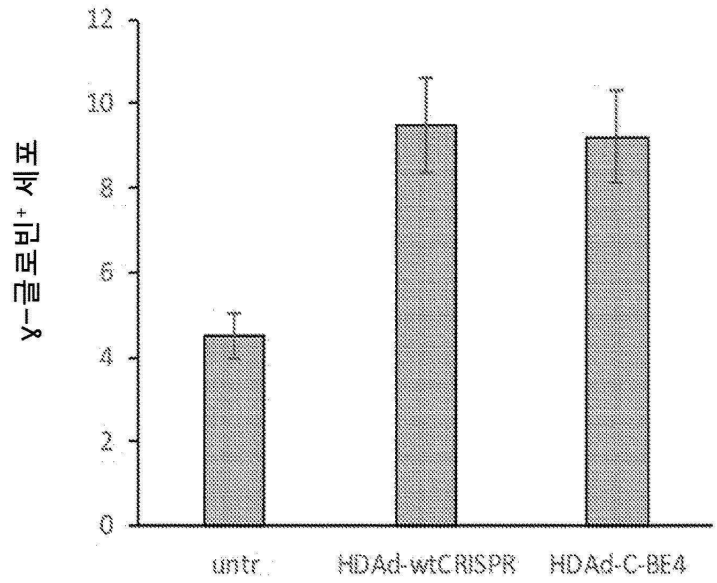

도면124a



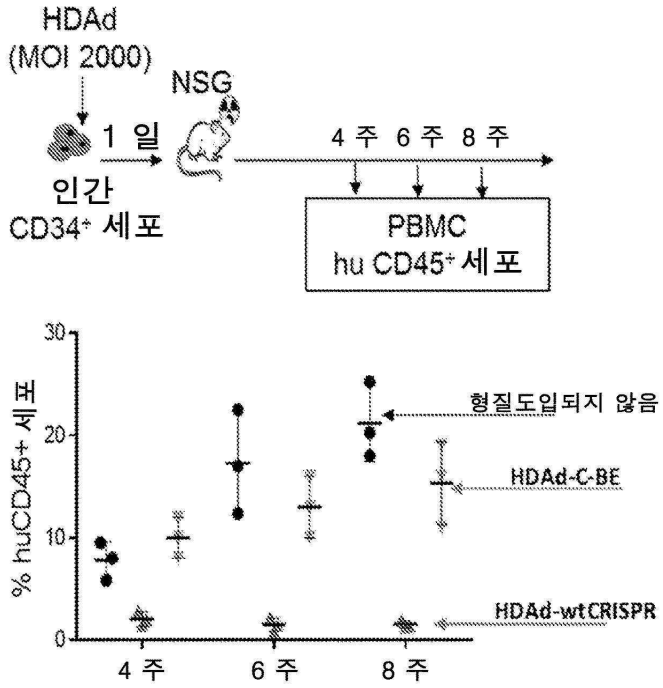
도면124b



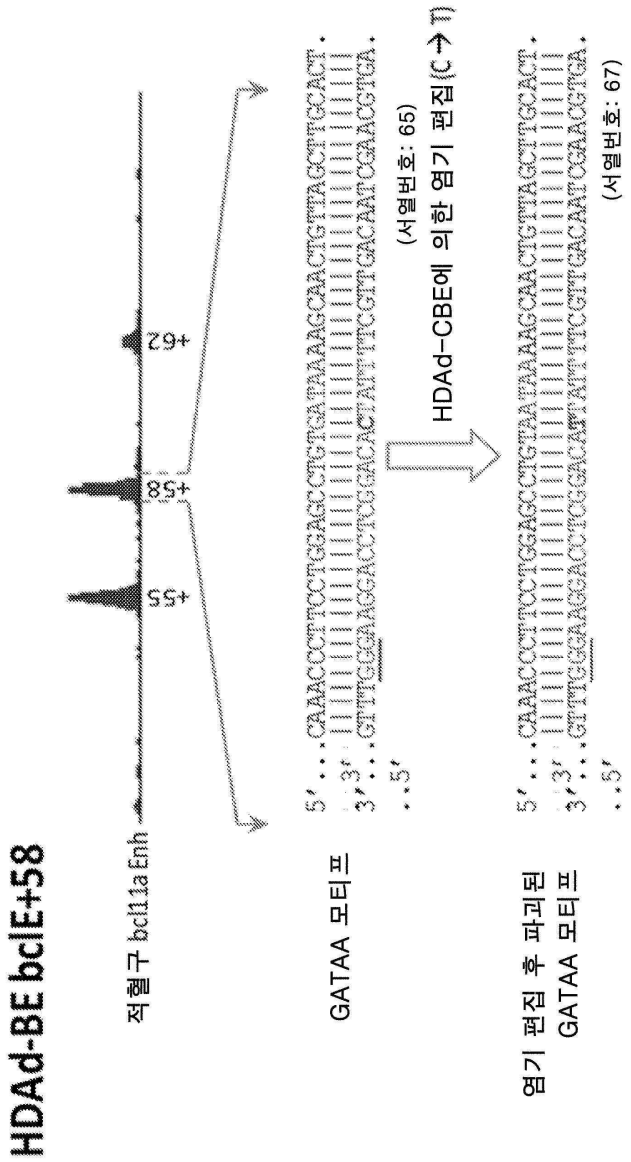

도면124c



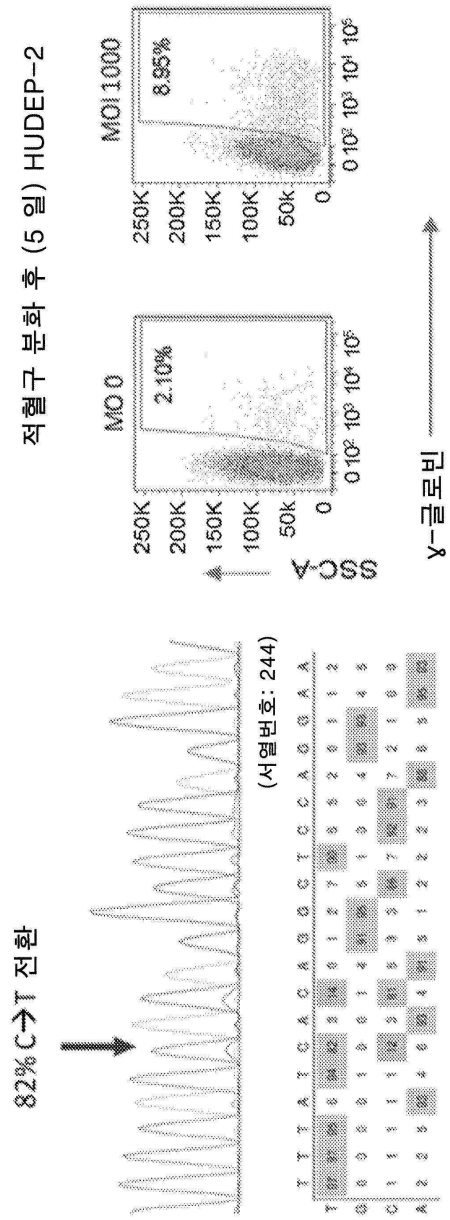
도면125


HDAd
(MOI 2000)
1 일
인간
CD34+ 세포
NSG
4 주 6 주 8 주
PBMC
hu CD45+ 세포
% huCD45+ 세포
30
20
10
0
4 주
6 주
8 주
형질도입되지 않음
HDAd-C-BE
HDAd-wtCRISPR

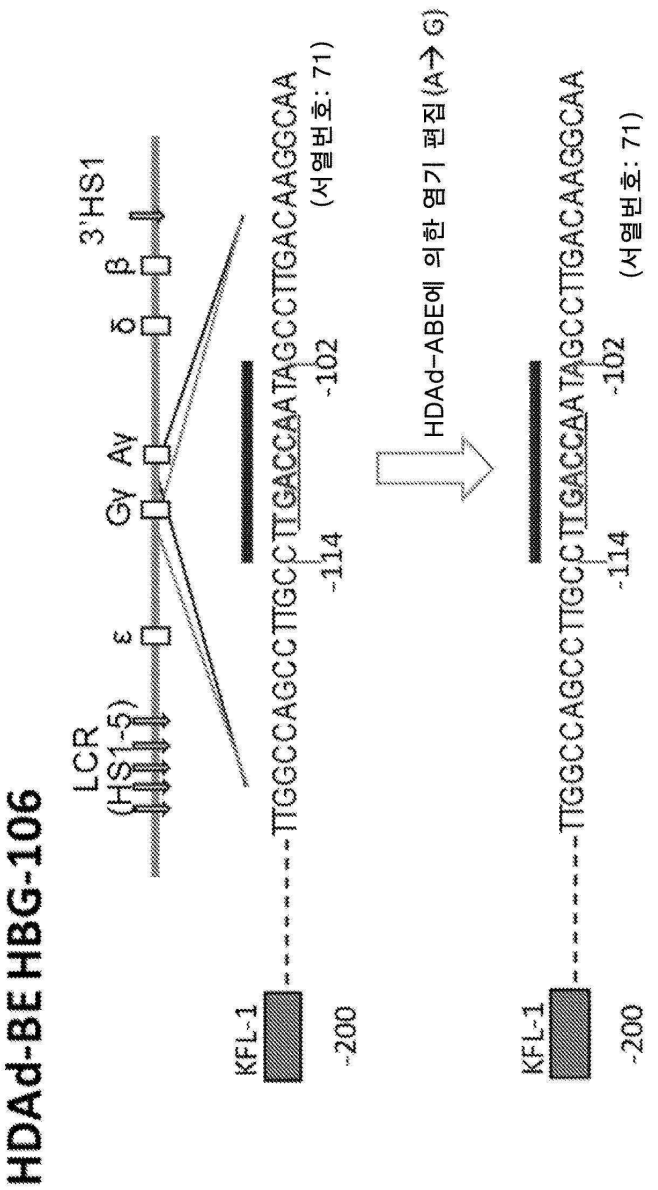
도면126a


HDAd-BE bclE+58
적혈구 bcl11a Enh
+55
+58
+62
GATAA 모티프
5'...CAAACCCTTCCTGGAGCCTGTGATAAAAGCAACTGTTAGCTTGCACT
3'...GTTTGGGAAGGACCTCGGACACTATTTTCGTTGACAATCGAACGTGA
5'
(서열번호: 65)
HDAd-CBE에 의한 염기 편집(C → T)
염기 편집 후 파괴된
GATAA 모티프
5'...CAAACCCTTCCTGGAGCCTGTGATAAAAGCAACTGTTAGCTTGCACT
3'...GTTTGGGAAGGACCTCGGACAATTATTTTCGTTGACAATCGAACGTGA
5'
(서열번호: 67)

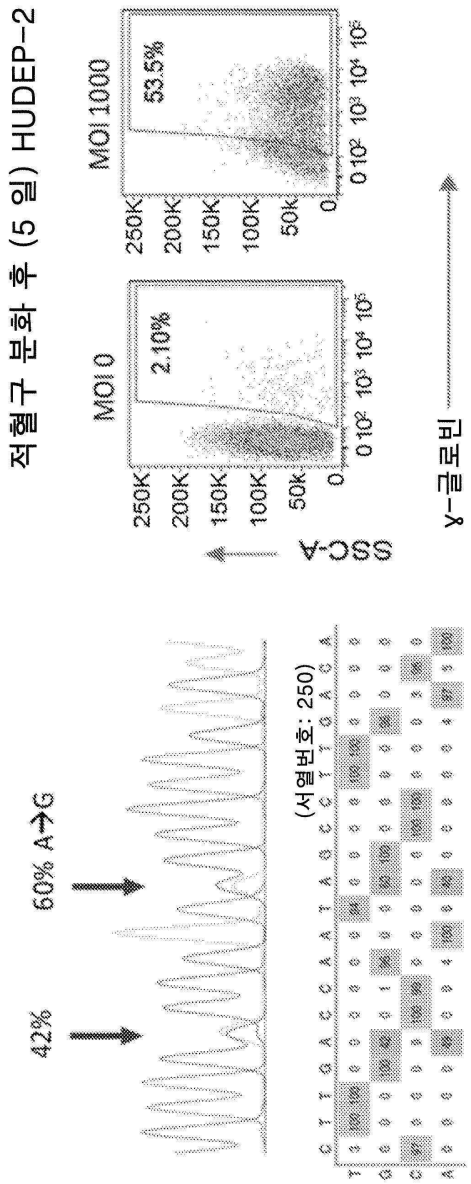
도면126b


적혈구 분화 후 (5 일) HUDEP-2
MOI 0
MOI 1000
2.10%
8.95%
SSC-A
γ-글로빈
82% C→T 전환
(서열번호: 244)

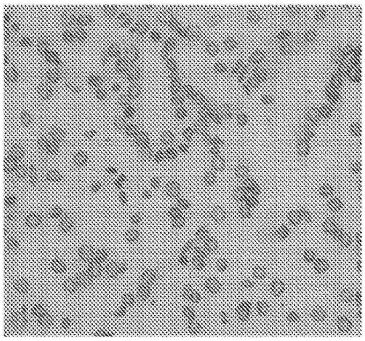
도면126c


HDAd-BE HBG-106
LCR
(HS1-5)
ε
Gγ
Aγ
δ
β
3'HS1
KFL-1
-200
-114
-102
TTGGCCAGCCTTGCCTTGACCAATAGCCTTGACAAGGCAA
(서열번호: 71)
HDAd-ABE에 의한 염기 편집(A→ G)
KFL-1
-200
-114
-102
TTGGCCAGCCTTGCCTTGACCAATAGCCTTGACAAGGCAA
(서열번호: 71)

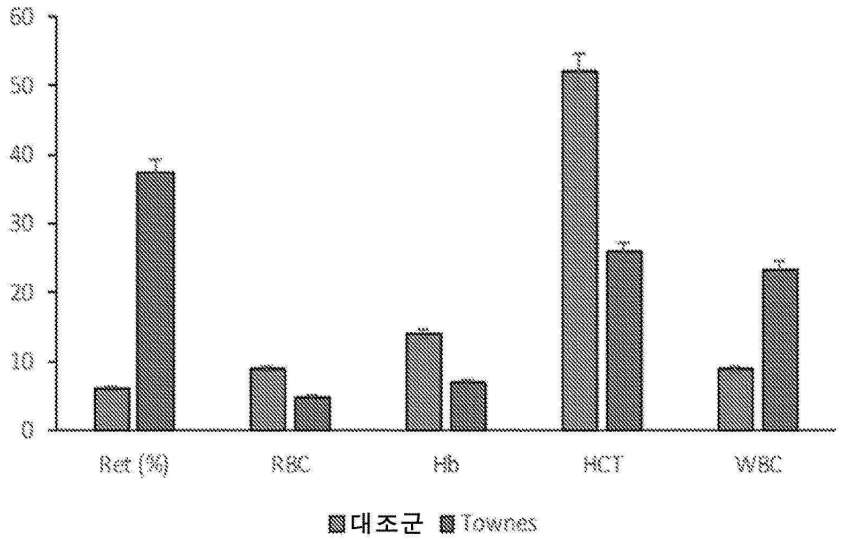
도면126d



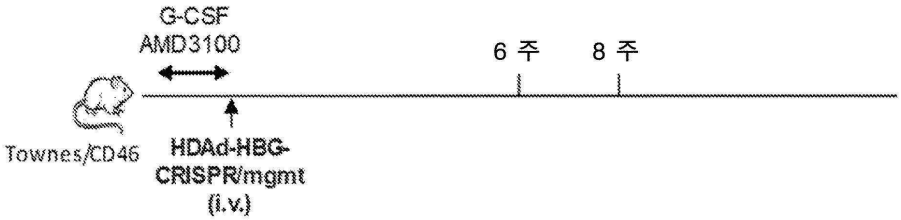

도면127a



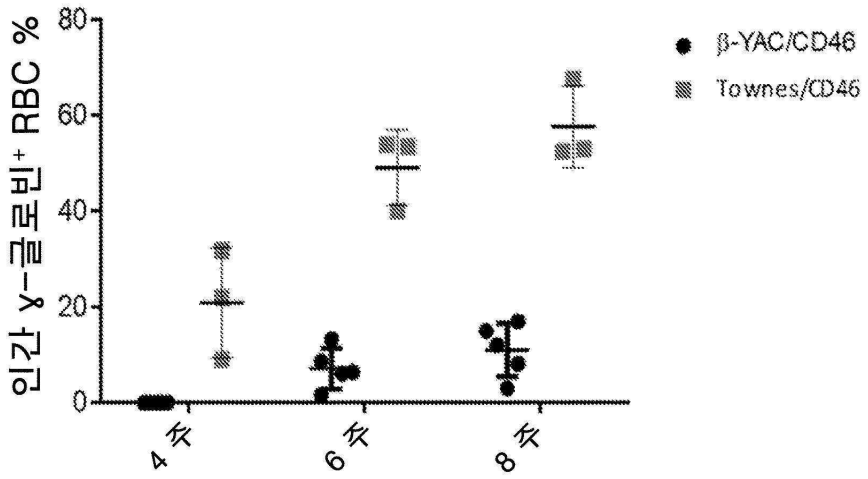
도면127b



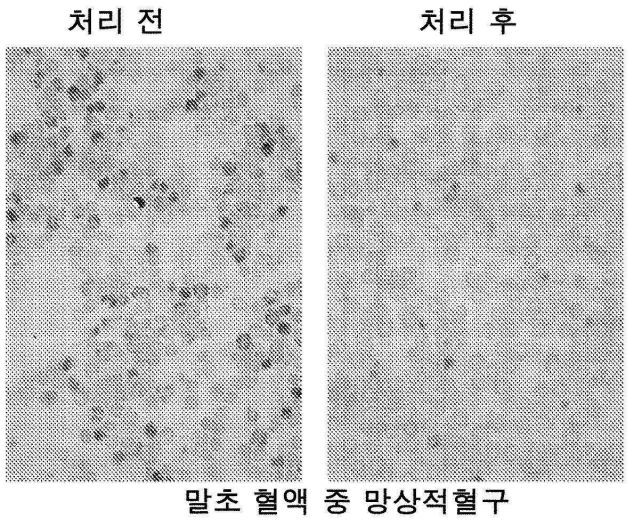

도면128a



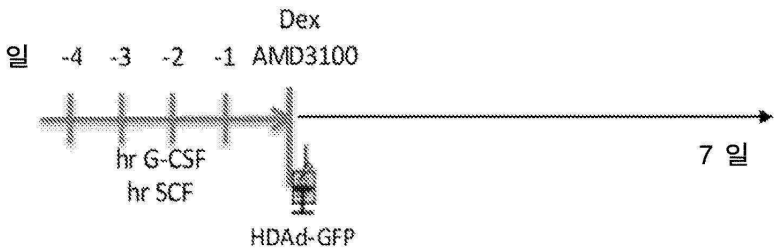

도면128b



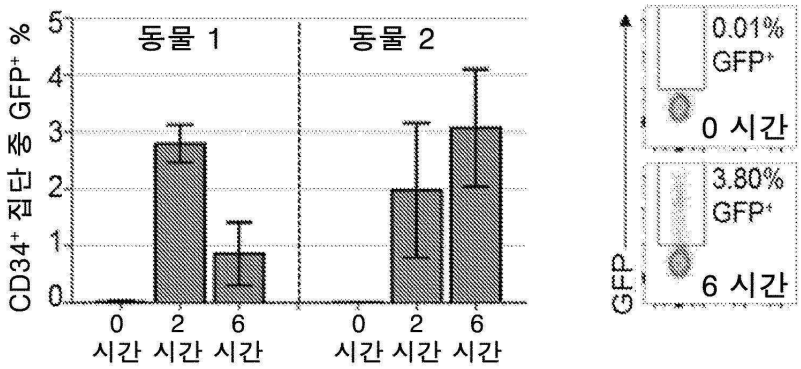
도면128c



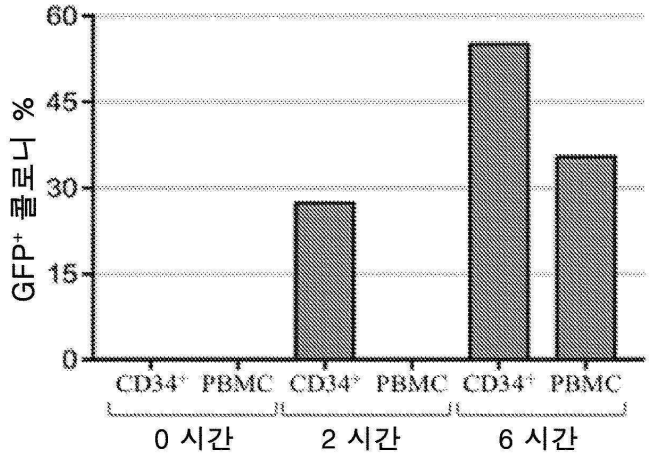

도면129a



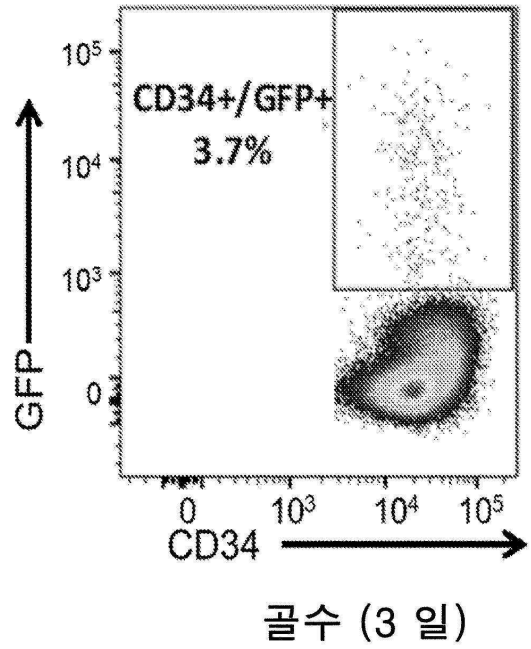

도면129b



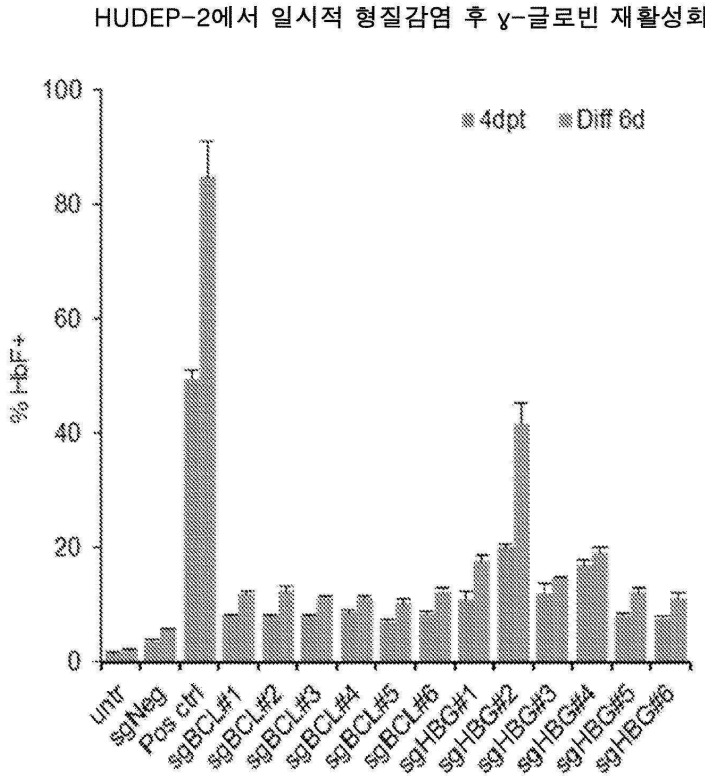
도면129c



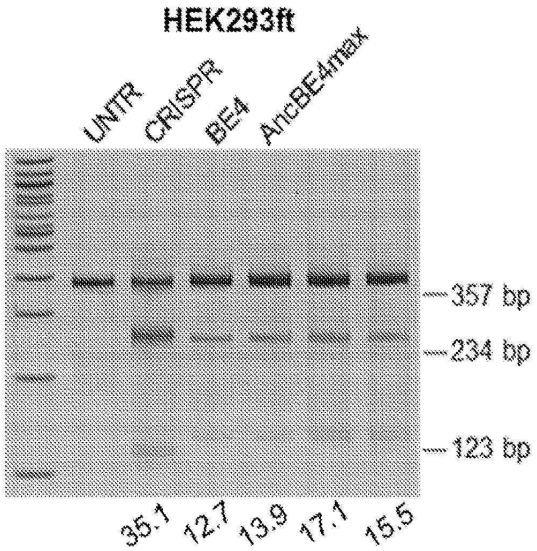

도면129d



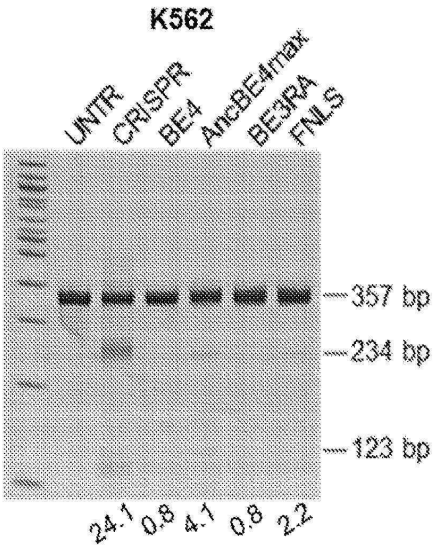
도면130



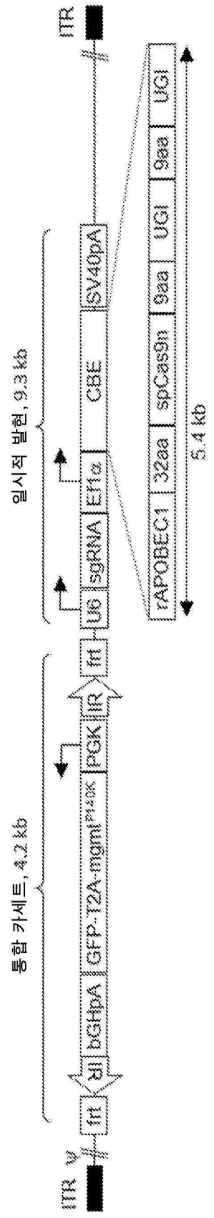

도면131a



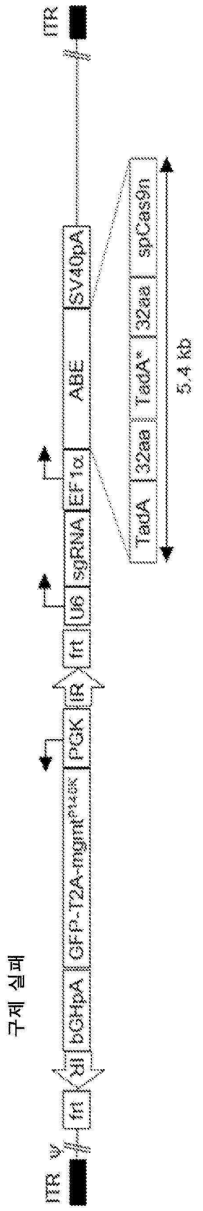
도면131b


K562
UNTR
CRISPR
BE4
AncBE4max
BE3RA
FNLS
357 bp
234 bp
123 bp
24.1
0.8
4.1
0.8
2.2

도면132a


ITR
ψ
frt
IR
bGHpA
GFP-T2A-mgmP140K
PGK
IR
frt
U6
sgRNA
Ef1α
CBE
SV40pA
통합 카세트, 4.2 kb
일시적 발현, 9.3 kb
rAPOBEC1
32aa
spCas9n
9aa
UGI
9aa
UGI
5.4 kb

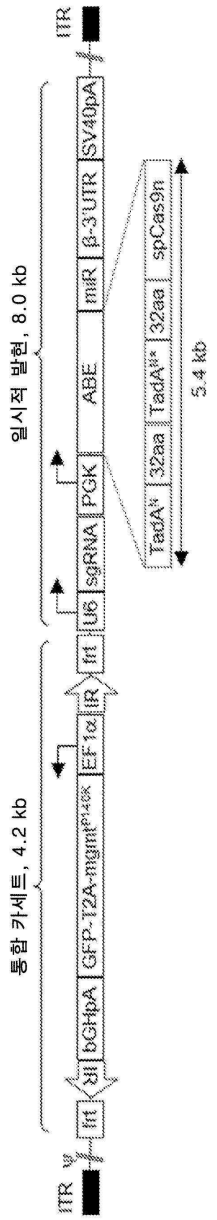
도면132b


구체 실패
ITR
ψ
frt
IR
bGHpA
GFP-T2A-mgmtP140K
PGK
IR
frt
U6
sgRNA
EF1α
ABE
SV40pA
TadA
32aa
TadA*
32aa
spCas9n
5.4 kb
ITR

도면132c

```
TadA^N+32aa    1 TCTGAAGTCGAGTTCAGCCACGAGTATTGGATGAGGCACGCACTGACCCTGGCAAAACGAGCCTGGGACG
TadA^N+32aa    1 .....G..G.....TTC.........C.........A..T..C.........C..GA.G..AC.C..T.

TadA^N+32aa   71 AAAGAGAAGTCCCCGTCGGCGCCGTGCTGGTGCACACAACAATCGCGTGATCGGAGAAGGATGGAACAGGCC
TadA^N+32aa   71 .G..G..G..G..T..G..A.....................TG.......A.A.........C..G..C.........AG.

TadA^N+32aa  141 AATCGGACGCCACGACCCTACCGCACACGCAGAGATCATGGCACTGAGGCAGGGAGGACTGGTCATGCAG
TadA^N+32aa  141 C.....C.TG.........A..A..C..T..C..A..T.....C.....A.....C..C.............

TadA^N+32aa  211 AATTATCGCCTGATCGATGCCACCCTGTATGTGACACTGGAGCCATGCGTGATGTGCGCAGGAGCAATGA
TadA^N+32aa  211 ..C..CA.A.....T..C.............C.......T.C.....T.....................C..C..C.....

TadA^N+32aa  281 TCCACAGCCGGATCGGAAGAGTGGTGTTCGGAGCACAGAGACGCCAAGACCGCAGCAGGCAGCCTGAT
TadA^N+32aa  281 .....TCTA.........CC.C.........T..C.TGA.GA.T..A..A..T..C..C.......TCT.....

TadA^N+32aa  351 GGATGTGCTGCACCACCCAGGCATGAACCACCGGGTGGAGATCACAGAGGGCATCCTGGCAGACGAGTGC
TadA^N+32aa  351 ...C.............T.....C.........T...A.A..C..A..T..C.....A.....................T..A..T

TadA^N+32aa  421 GCCCGCCCTGCTGAGCGATTTCTTTAGAATGAGGAGACAGGAGATCAAGGCCCAGAAGAAGGCACAGAGCT
TadA^N+32aa  421 ..............T..T.........C.G...CCT.........T.T.T..T..T.....................C.......

TadA^N+32aa  491 CCACCGACTCTGGAGGCTCTAGCGGAGGATCCTCTGGAAGCGAGACACCAGGCACATCCGAGTCCGCCAC
TadA^N+32aa  491 .........TAG...C..A.........G..C.........CTCT.....C..T.......AG....AG...A..

TadA^N+32aa  561 ACCAGAGTCTTCCGGCGGCTCCTCCGGAGGATCC  (서열번호: 260)
TadA^W+32aa  561 ...T..AAGCAG....G...AG.AG...G..GAG.  (서열번호: 261)
```

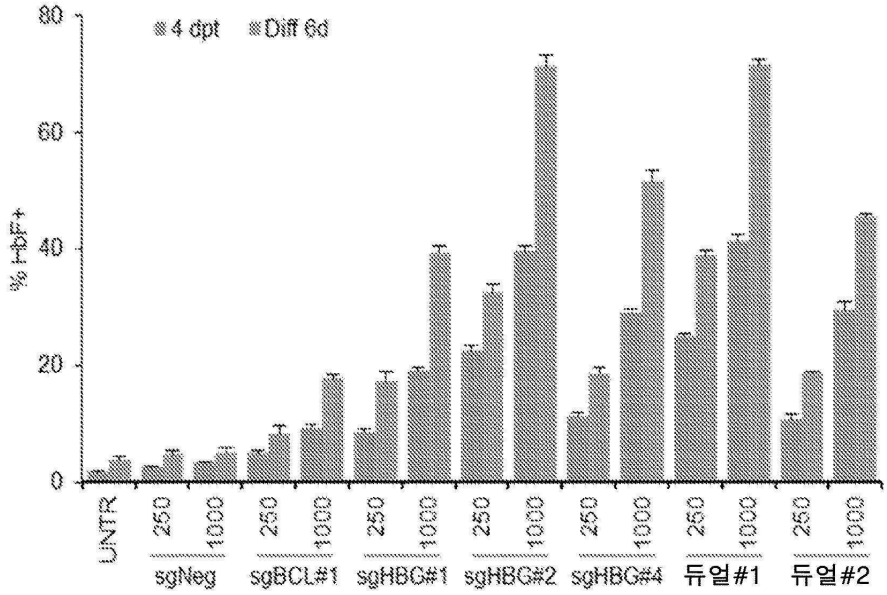
도면133a


ITR
frt
IRII
bGHpA
GFP-T2A-mgmtP140K
EF1α
IRI
frt
U6
sgRNA
PGK
ABE
miR
β-3'UTR
SV40pA
ITR
통합 카세트, 4.2 kb
일시적 발현, 8.0 kb
TadA8
32aa
TadA8*
32aa
spCas9n
5.4 kb

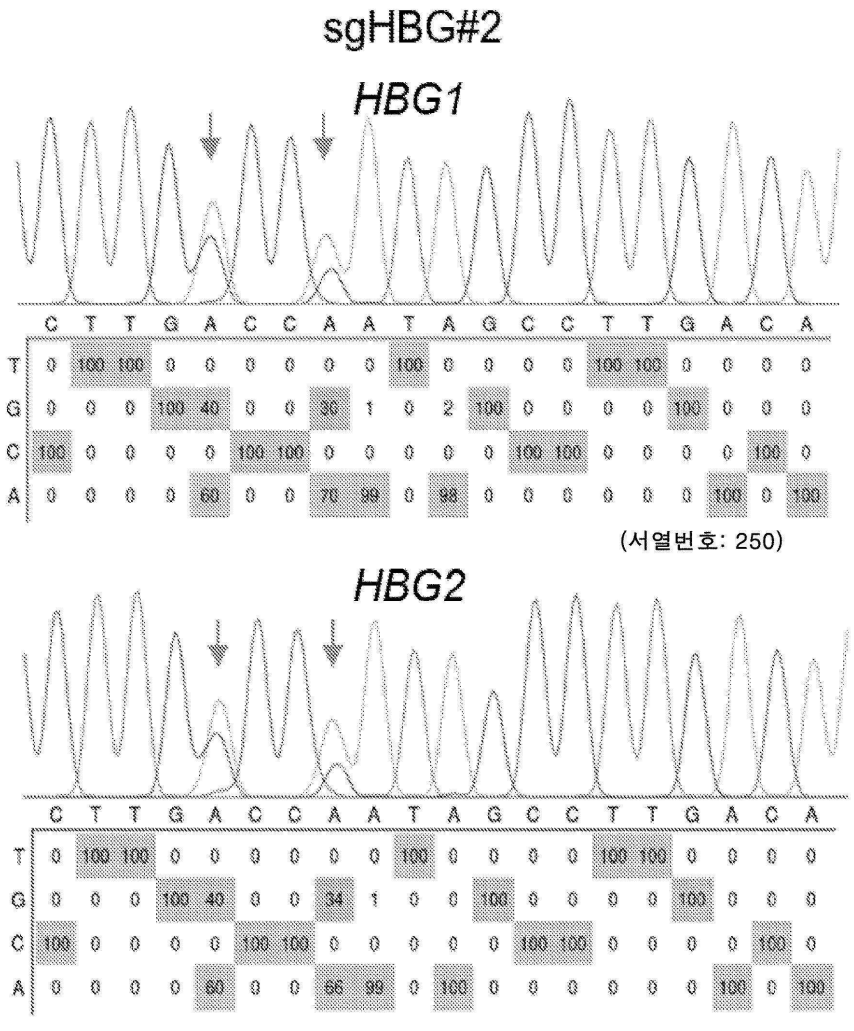
도면133b

| 명칭 | 편집기 | 가이드 RNA | 수율 ($\times 10^{12}$ vp) |
|---|---|---|---|
| HDAd5/35++_BE4max-sgBCL#1-FI-mgmt/GFP | CBE | sgBCL#1 | 1.2 |
| HDAd5/35++_BE4max-sgHBG#1-FI-mgmt/GFP | CBE | sgHBG#1 | 0.8 |
| HDAd5/35++_ABEopti-sgHBG#2-FI-mgmt/GFP | ABE | sgHBG#2 | 3.8 |
| HDAd5/35++_ABEopti-sgHBG#4-FI-mgmt/GFP | ABE | sgHBG#4 | 2.8 |
| HDAd5/35++_BE4max-Dual#1-FI-mgmt/GFP | CBE | sgBCL#1 + sgHBG#1 | 0.5 |
| HDAd5/35++_BE4max-sgNeg-FI-mgmt/GFP | CBE | sgNeg | 1.5 |

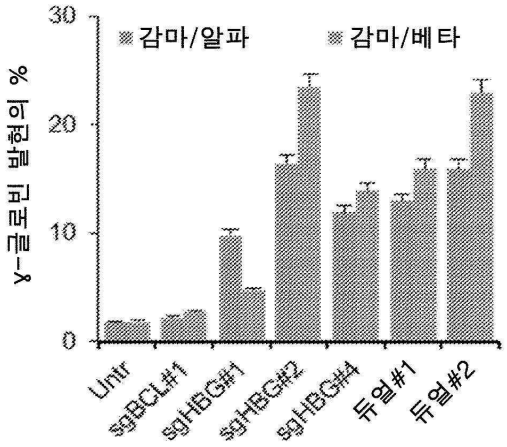
도면133c



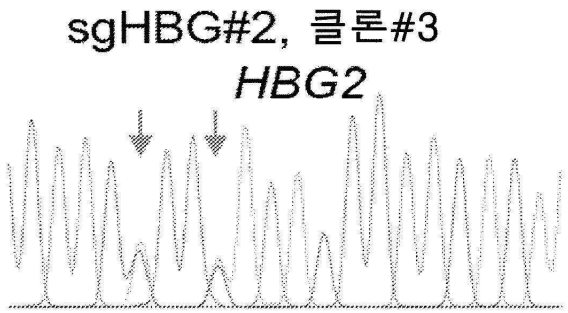

도면133d



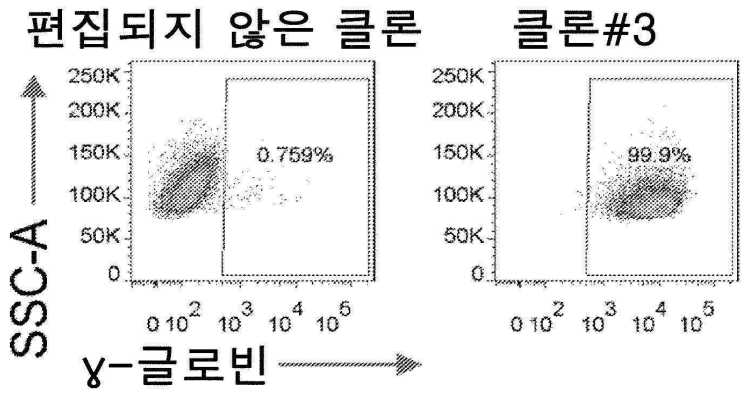
도면133e



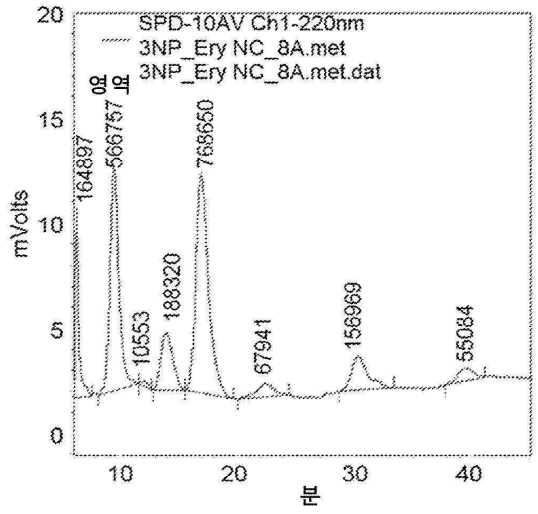

도면133f



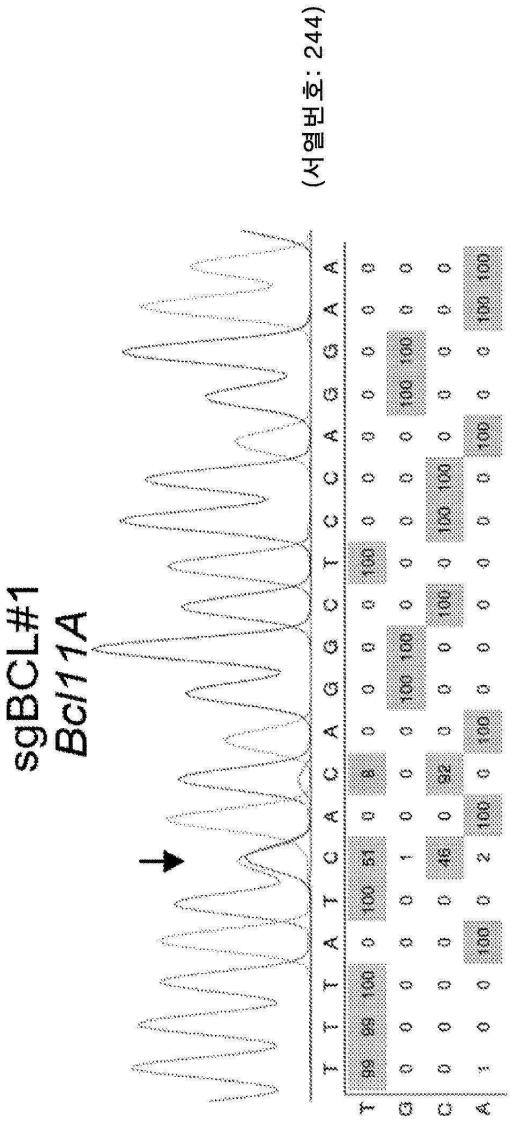

도면133g



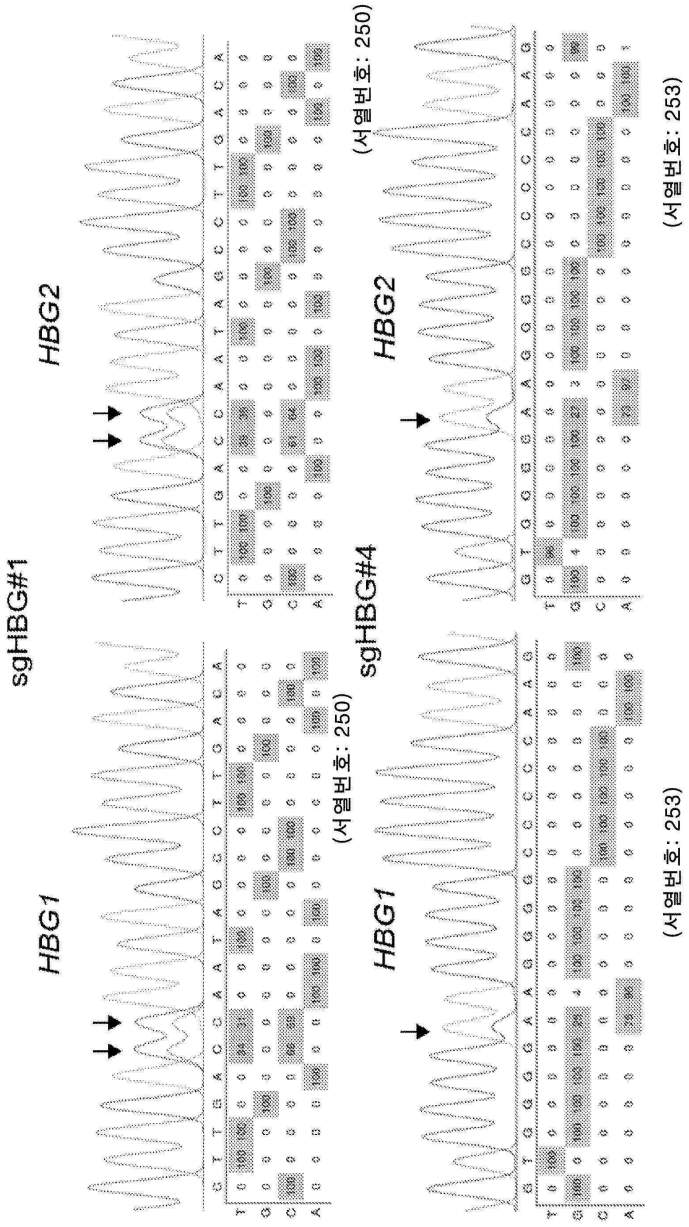
도면133h



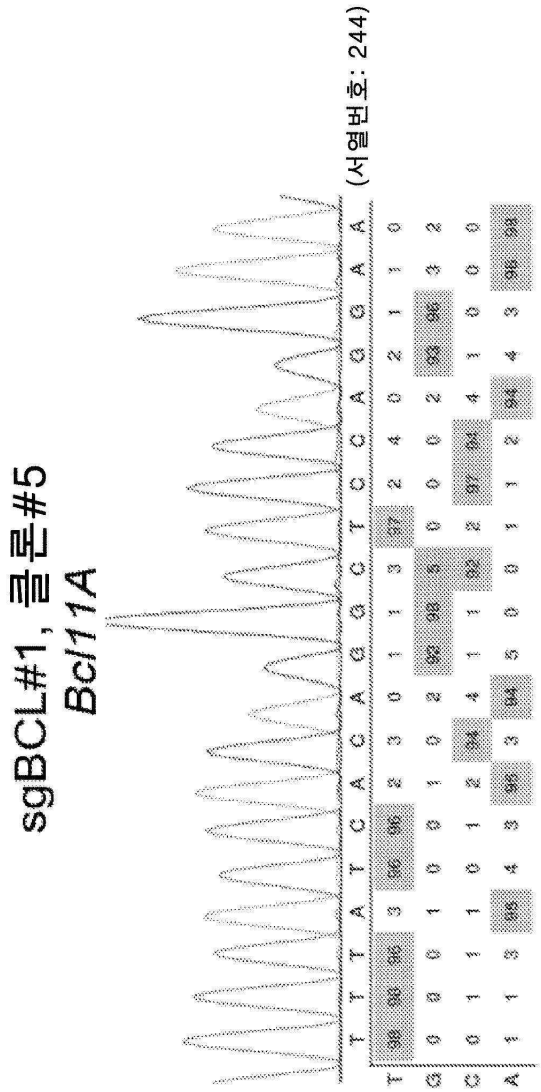

도면134aa



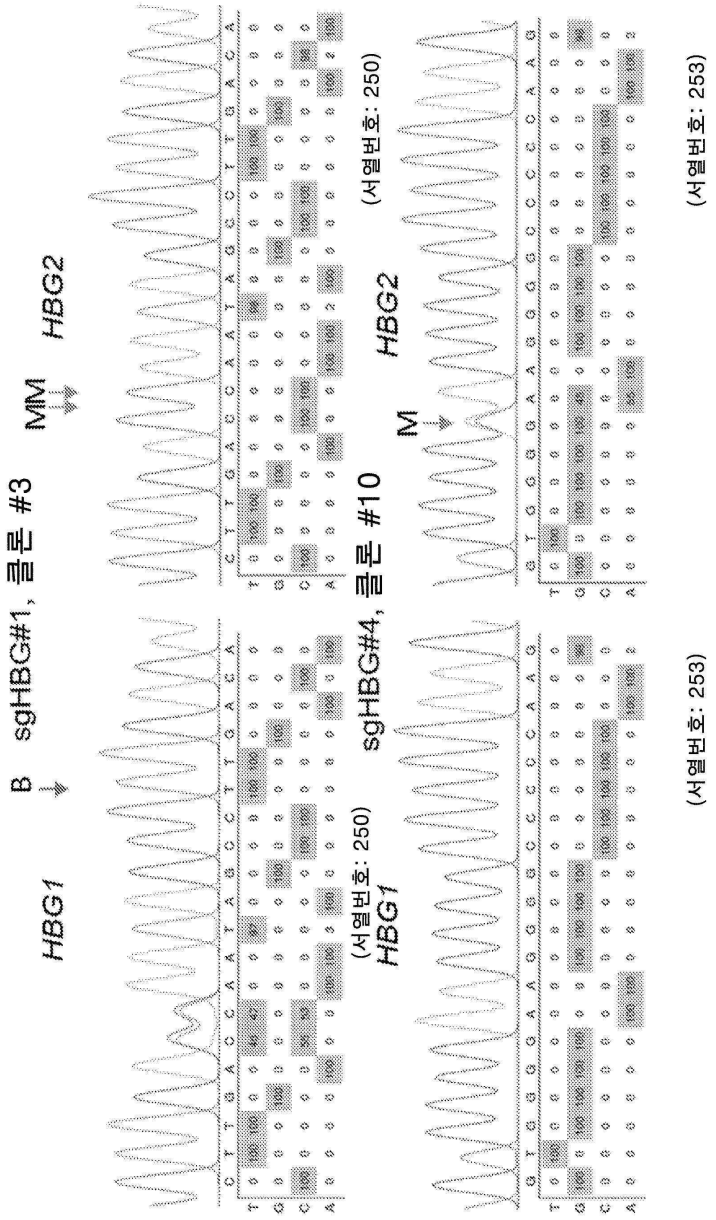
도면134ab


sgHBG#1
HBG1
HBG2
(서열번호: 250)
(서열번호: 253)
sgHBG#4
HBG1
HBG2
(서열번호: 250)
(서열번호: 253)

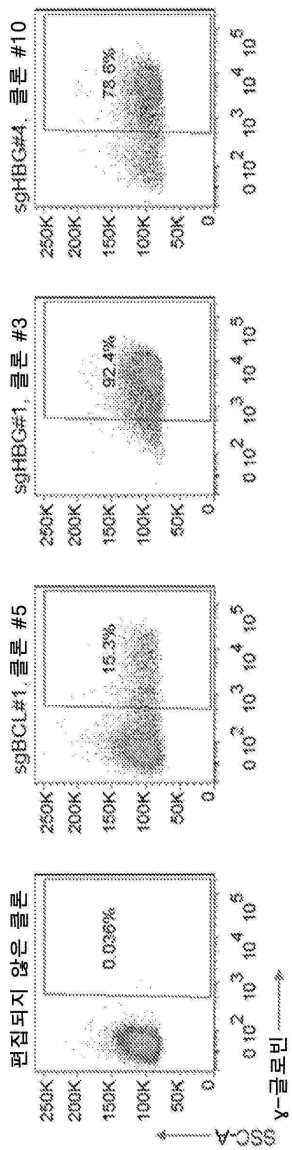
도면134ba


sgBCL#1, 클론#5
Bcl11A
(서열번호: 244)
T T T A T C A C A G G C C T C C A G G A A
T 38 92 95 3 96 96 2 3 0 1 0 2 92 98 6 2 97 2 0 0 2 93 96 3 2
G 0 0 0 1 0 0 1 0 2 94 4 1 1 2 0 0 0 1 0 2 0 1 0 0 0
C 0 1 1 1 0 1 2 94 3 94 5 0 0 1 2 94 1 0 1 2 4 1 0 0 0
A 1 1 3 95 4 3 95 3 94 5 0 0 1 2 94 4 3 96 96

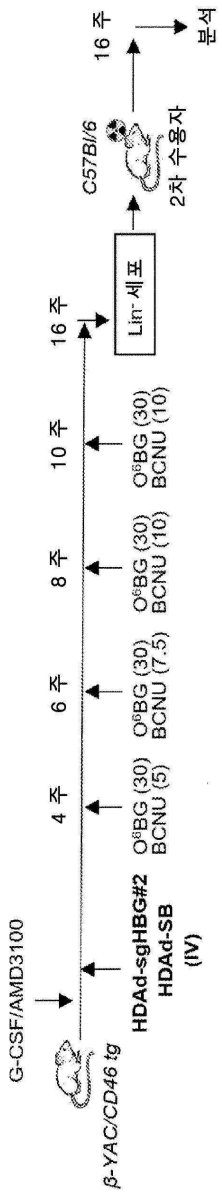
도면134bb


B
sgHBG#1, 클론 #3
MM
HBG1
HBG2
(서열번호: 250)
(서열번호: 253)
sgHBG#4, 클론 #10
M
HBG1
HBG2
(서열번호: 250)
(서열번호: 253)

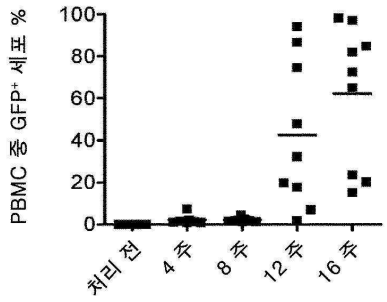
도면134c


편집되지 않은 클론
sgBCL#1, 클론 #5
sgHBG#1, 클론 #3
sgHBG#4, 클론 #10
0.036%
15.3%
92.4%
78.6%
SSC-A
γ-글로빈

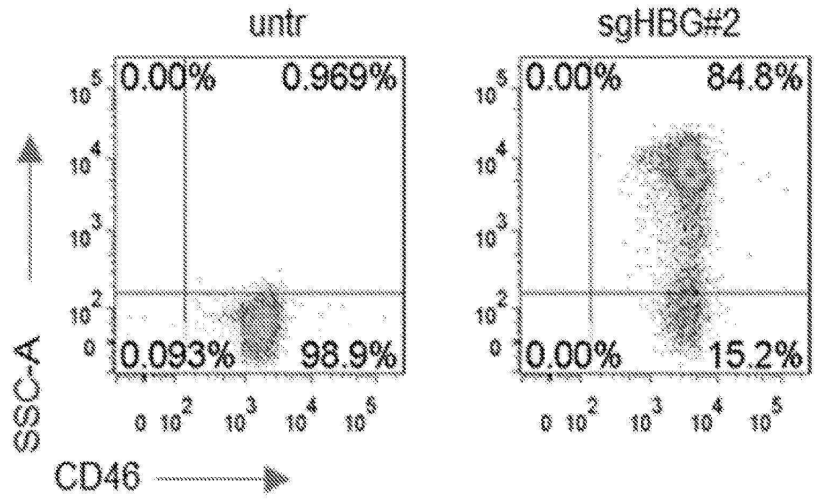
도면135a



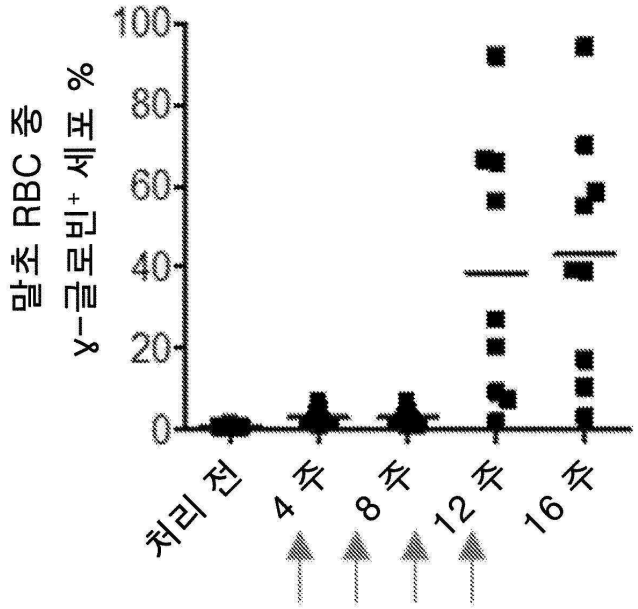

도면135b



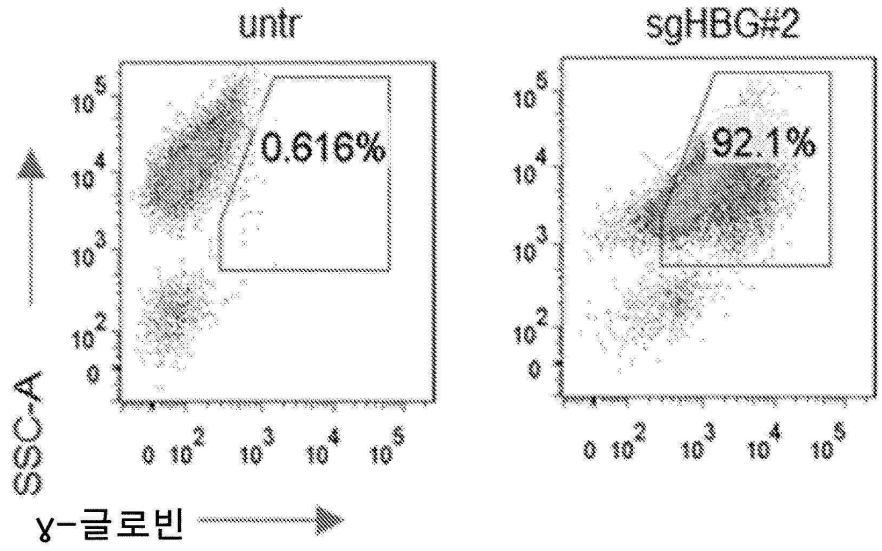
도면135c



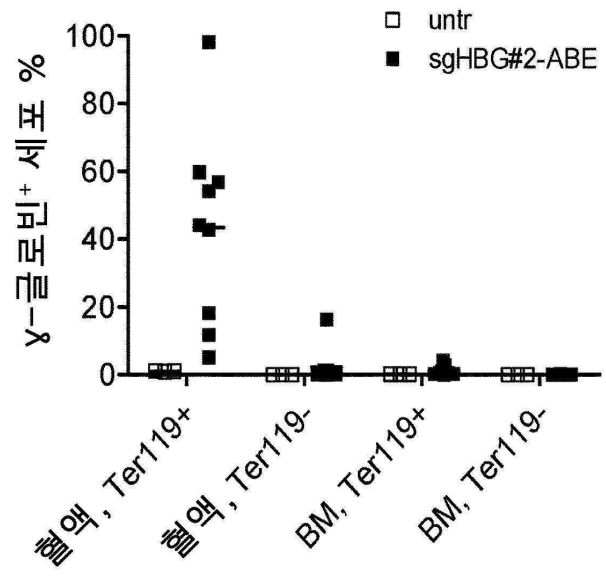

도면135d



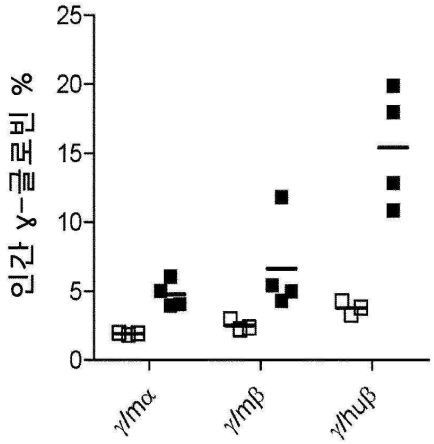
도면135e



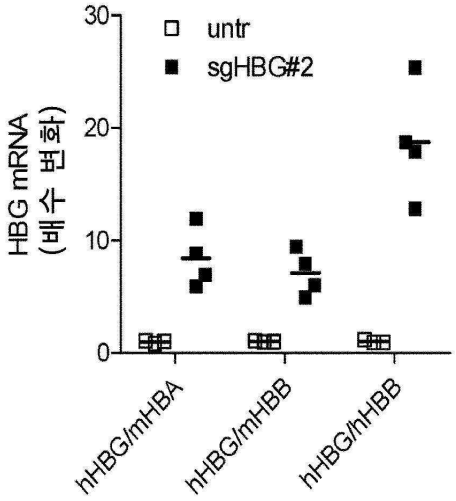

도면135f



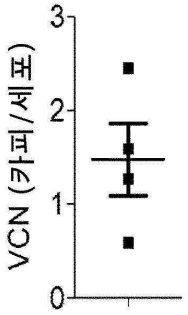

도면135g



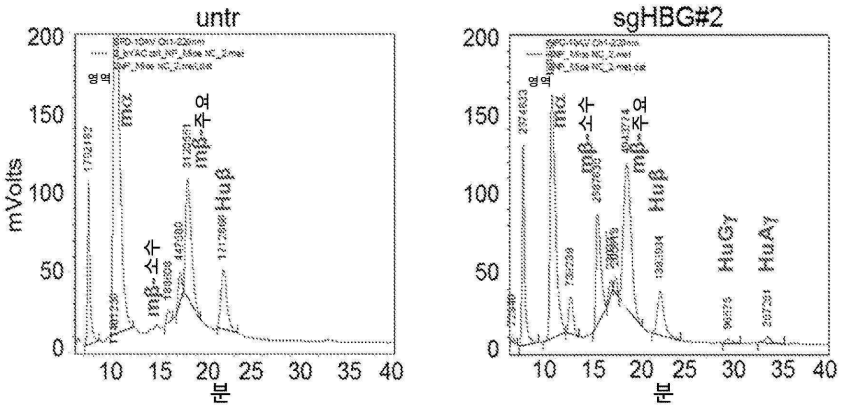
도면135h



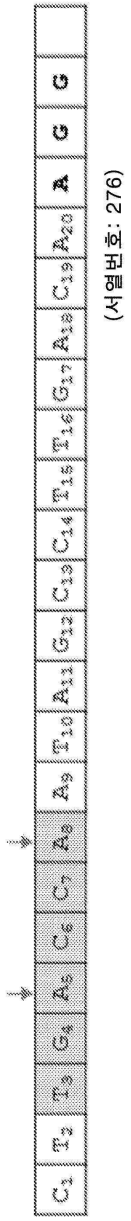

도면135i



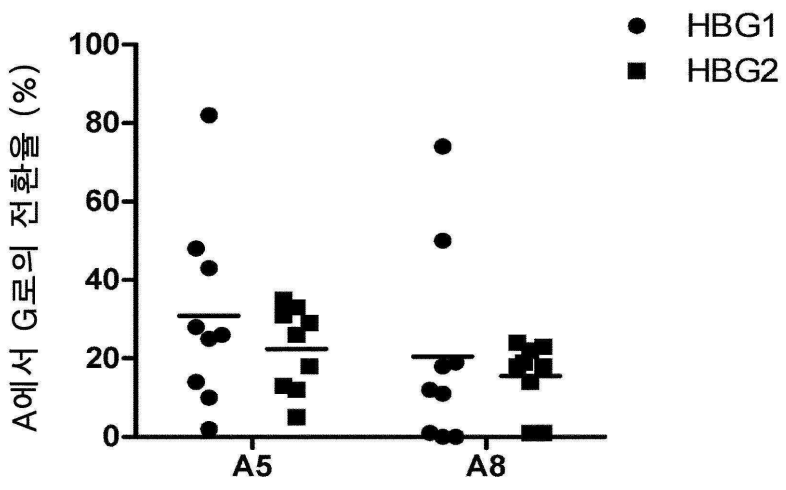

도면136



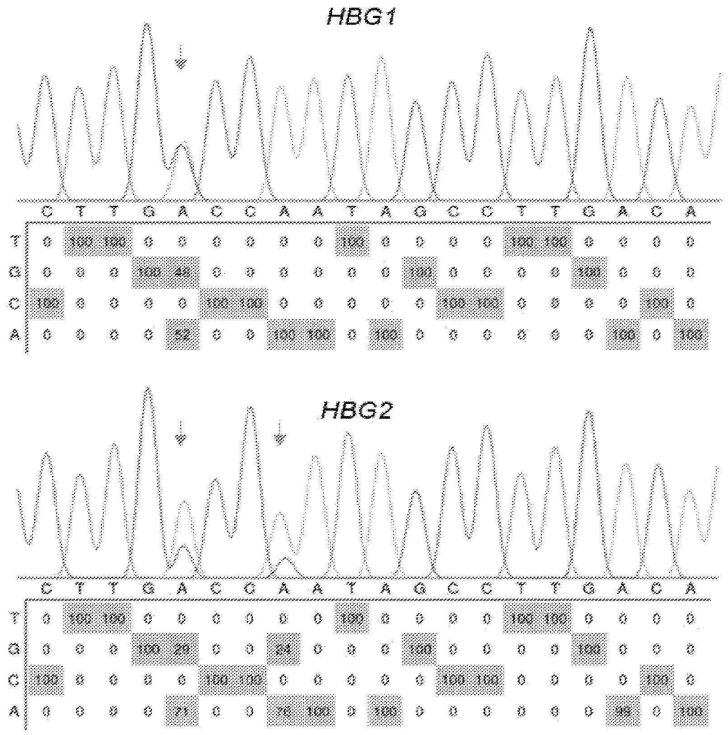
도면137a




도면137b



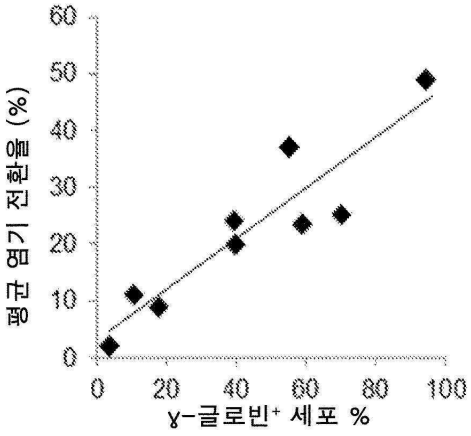
도면137c



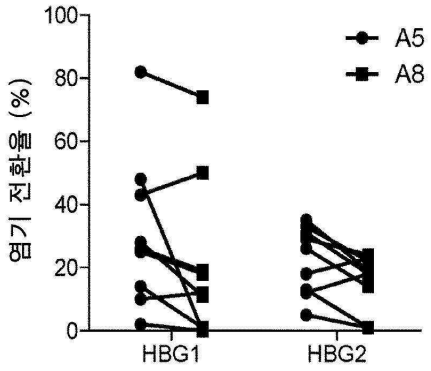

(서열번호: 250)

(서열번호: 250)

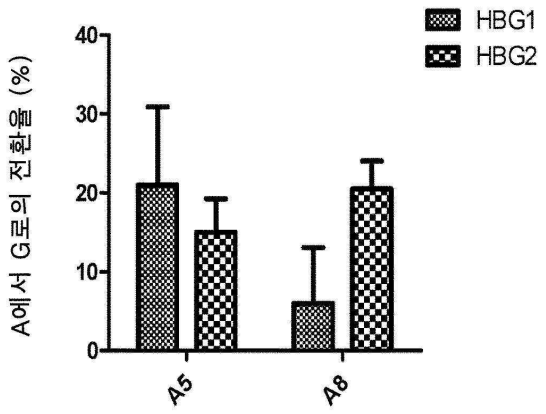
도면137d



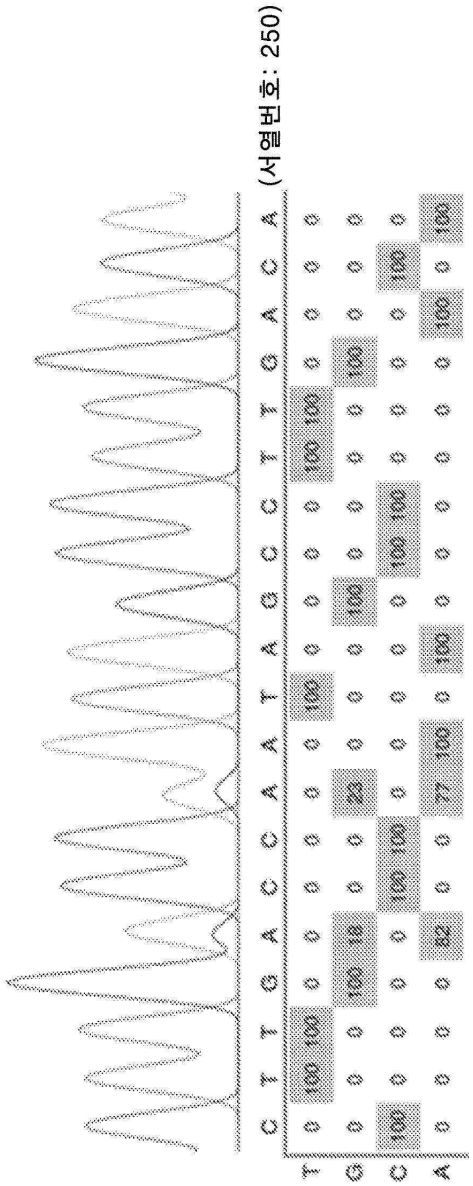

도면137e



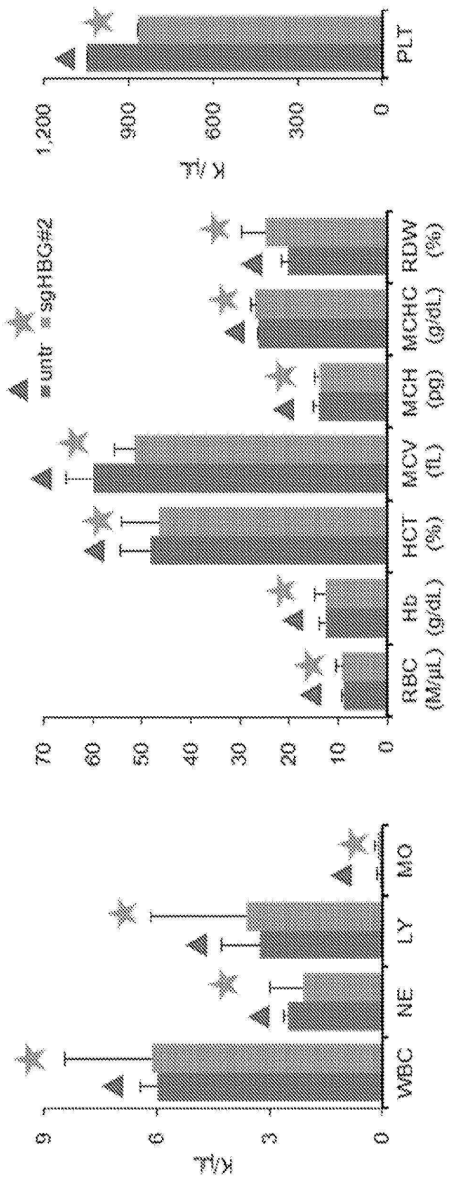
도면137f



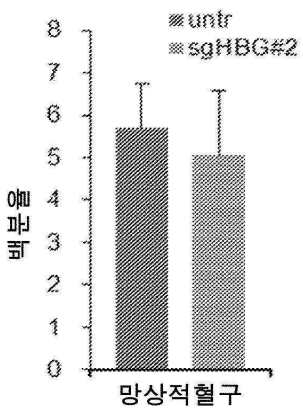

도면137g



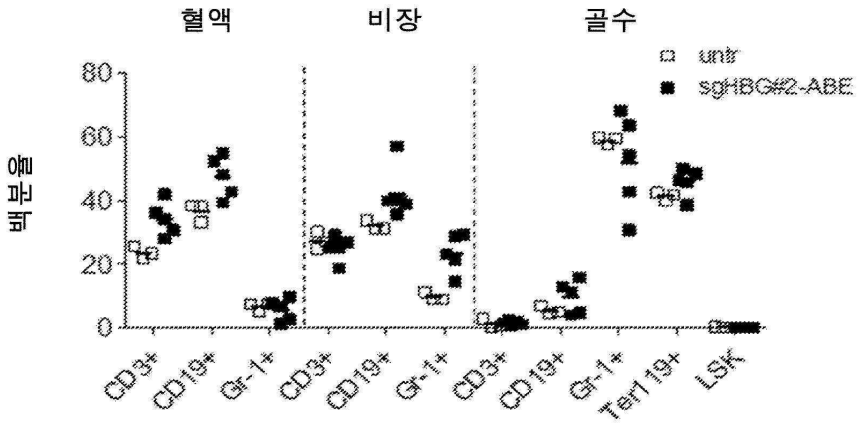
도면138a



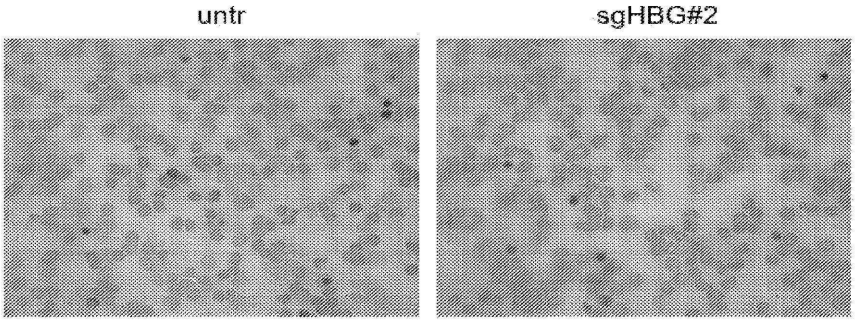

도면138b



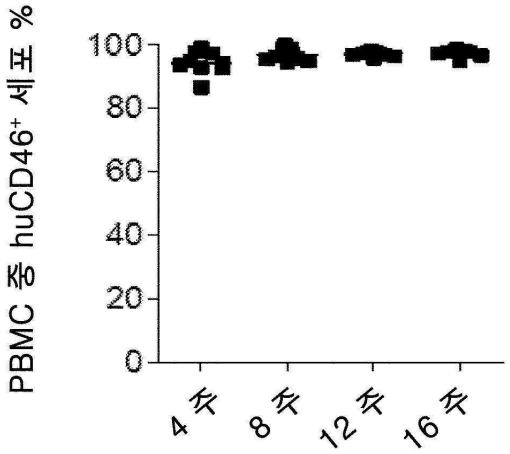
도면138c



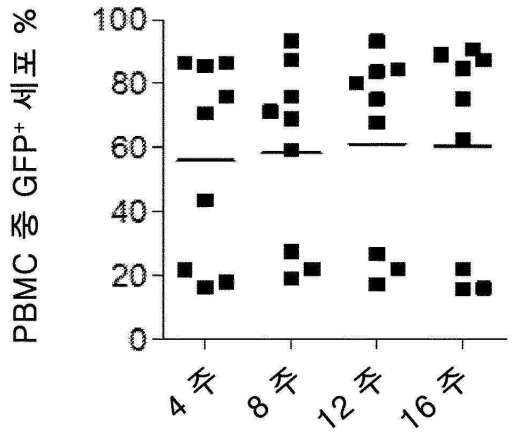

도면138d



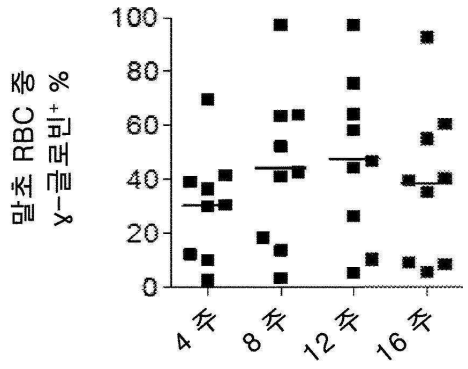

도면139a



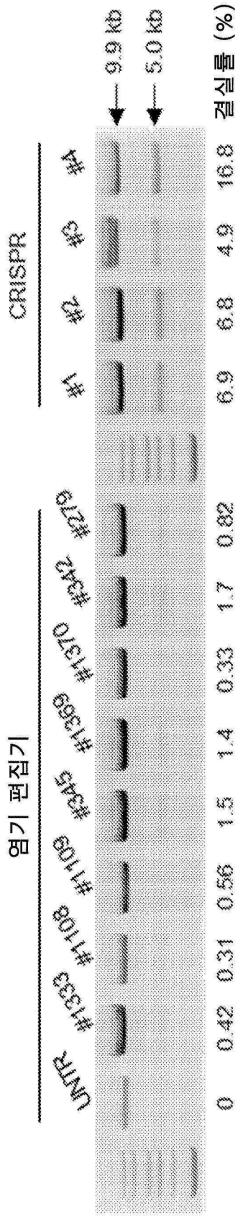
도면139b



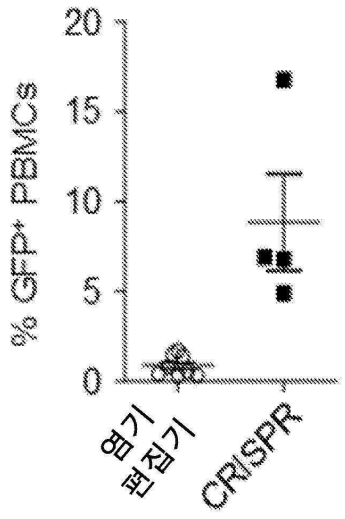

도면139c



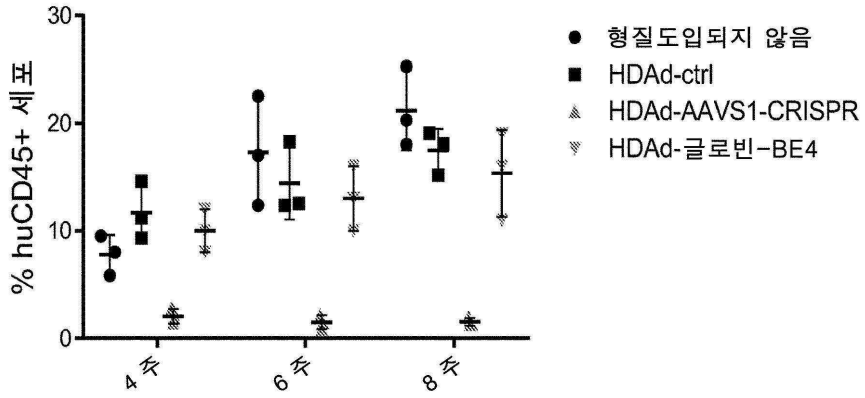
도면140a


CRISPR
#4
#3
#2
#1
6.9
6.8
4.9
16.8
결실률 (%)
6.9 kb
5.0 kb
염기 편집기
UNTR
#1333
#1108
#1109
#1345
#1369
#1370
#342
#279
0
0.42
0.31
0.56
1.5
1.4
0.33
1.7
0.82

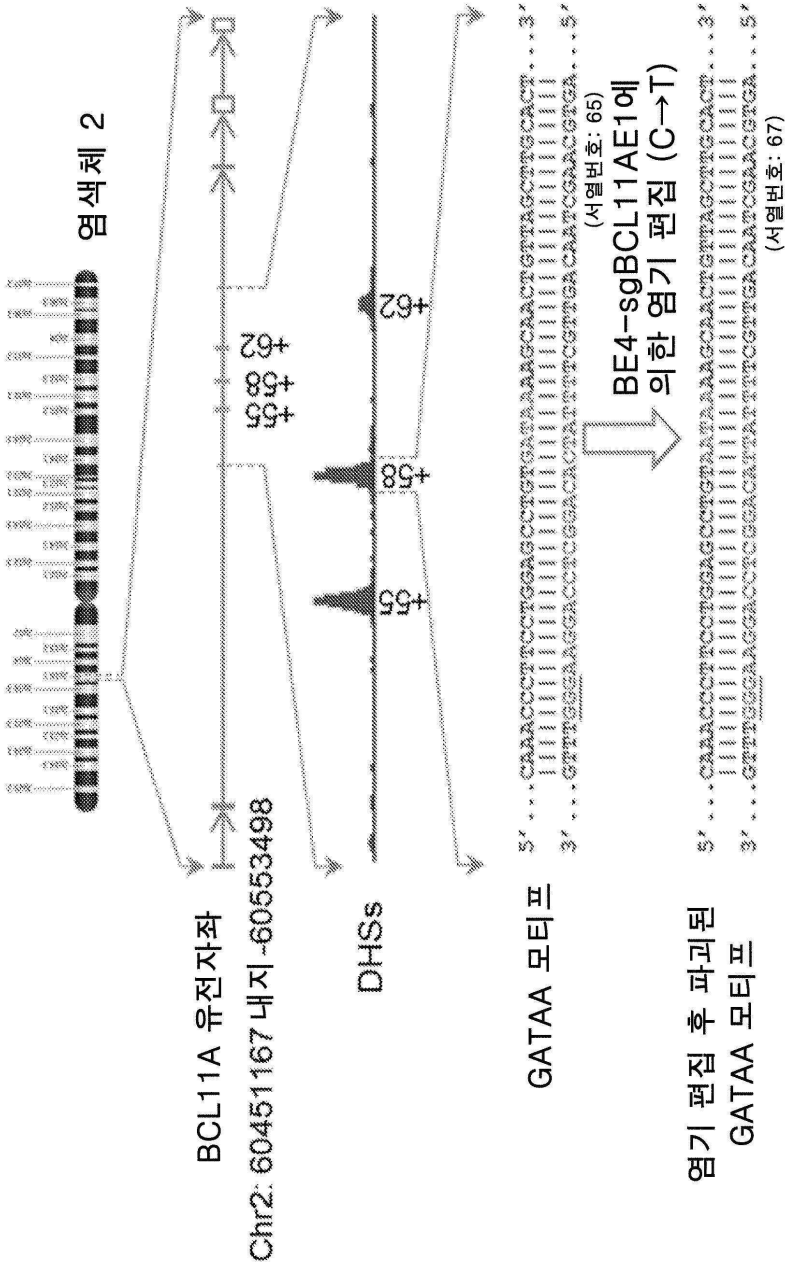
도면140b



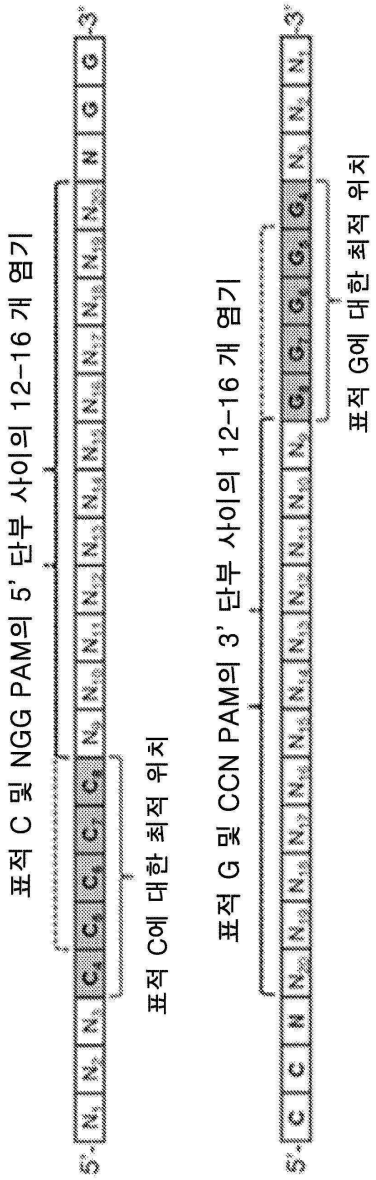

도면141



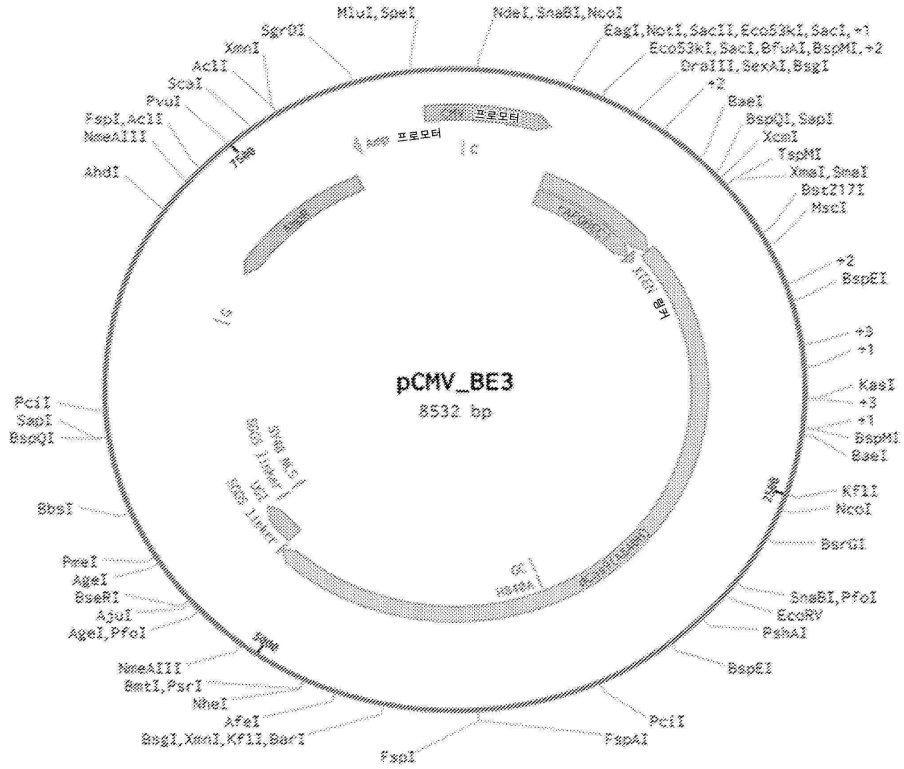
도면142


염색체 2
BCL11A 유전자좌
Chr2: 60451167 내지 60553498
+62
+55
+58
+62
DHSs
+55
+58
+62
GATAA 모티프
5′ ...CAAACCCTTCCTGGAGCCTGTGATAAAAGCAACTGTTAGCTTGCACT...3′
3′ ...GTTTGGGAAGGACCTCGGACACTATTTTCGTTGACAATCGAACGTGA...5′
(서열번호: 65)
BE4-sgBCL11AE1에
의한 염기 편집 (C→T)
염기 편집 후 파괴된
GATAA 모티프
5′ ...CAAACCCTTCCTGGAGCCTGTGATAAAAGCAACTGTTAGCTTGCACT...3′
3′ ...GTTTGGGAAGGACCTCGGACATTATTTTCGTTGACAATCGAACGTGA...5′
(서열번호: 67)

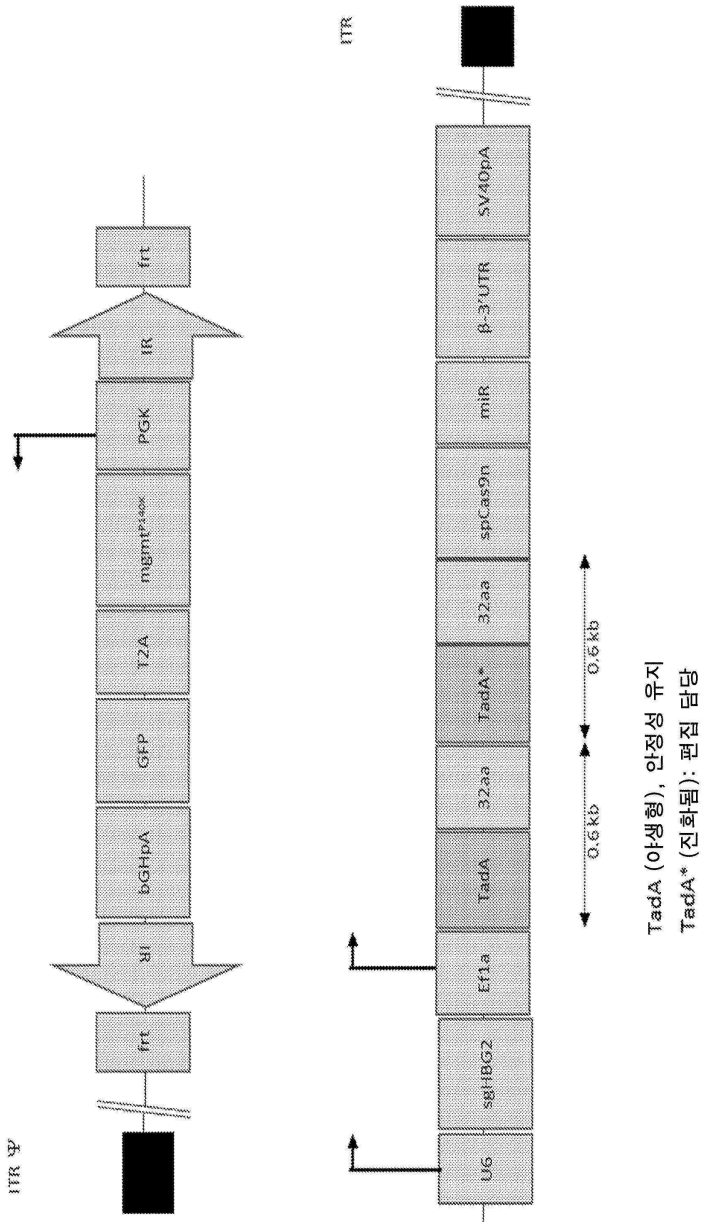
도면143


표적 C 및 NGG PAM의 5' 단부 사이의 12-16 개 염기
표적 C에 대한 최적 위치
표적 G 및 CCN PAM의 3' 단부 사이의 12-16 개 염기
표적 G에 대한 최적 위치

도면144


pCMV_BE3
8532 bp
MluI,SpeI
NdeI,SnaBI,NcoI
EagI,NotI,SacII,Eco53kI,SacI,+1
Eco53kI,SacI,BfuAI,BspMI,+2
DraIII,SexAI,BsgI
+2
BaeI
BspQI,SapI
XcmI
TspMI
XmaI,SmaI
BstZ17I
MscI
+2
BspEI
+3
+1
KasI
+3
+1
BspMI
BaeI
KflI
NcoI
BsrGI
SnaBI,PfoI
EcoRV
PshAI
BspEI
PciI
FspAI
FspI
BsgI,XmnI,KflI,BarI
AfeI
NheI
BmtI,PsrI
NmeAIII
AgeI,PfoI
AjuI
BseRI
AgeI
PmeI
BbsI
PciI
SapI
BspQI
AhdI
NmeAIII
FspI,AclI
PvuI
ScaI
AclI
XmnI
SgrDI
pCMV 프로모터
AmpR 프로모터
f1 ori
pUC ori
CMV enhancer
T7 promoter
hGH poly(A) signal
3xFlag
ECO Linker I
ECO Linker II
lacZ α
lacZ α
2000
1000
3000
4000
5000
6000
7000
8000

도면145


ITR Ψ
frt
SI
bGHpA
GFP
T2A
mgmt^P140K
PGK
IR
frt
ITR
U6
sgHBG2
Ef1a
TadA
32aa
TadA*
32aa
spCas9n
miR
β-3'UTR
SV40pA
ITR
0.6 kb
0.6 kb
TadA (야생형), 안정성 유지
TadA* (진화됨): 편집 담당

도면146



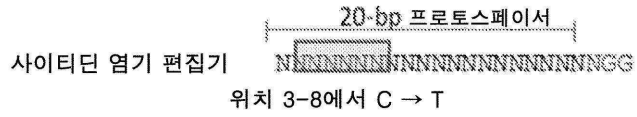
(서열번호: 265)
(서열번호: 266)

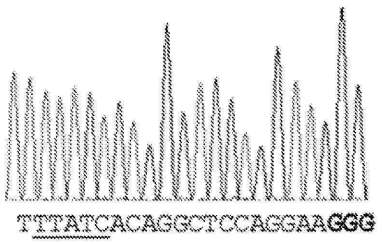
2 개의 'TadA + 32aa'의 DNA 서열



(서열번호: 267)
(서열번호: 268)

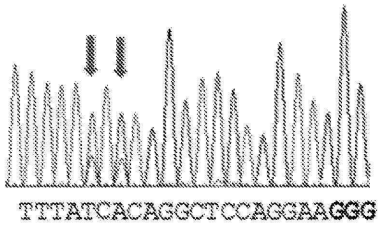
도면147




야생형 서열
밑줄 표시: GATAA 모티프
굵은 글씨: PAM 모티프

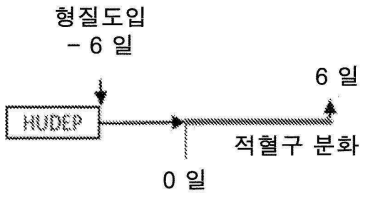
(서열번호: 269)



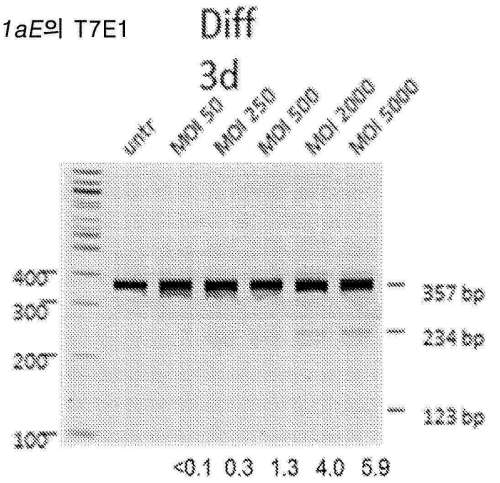

염기 편집된 서열
화살표: 편집된 염기
보라색: GATAA 모티프의 파괴된 염기

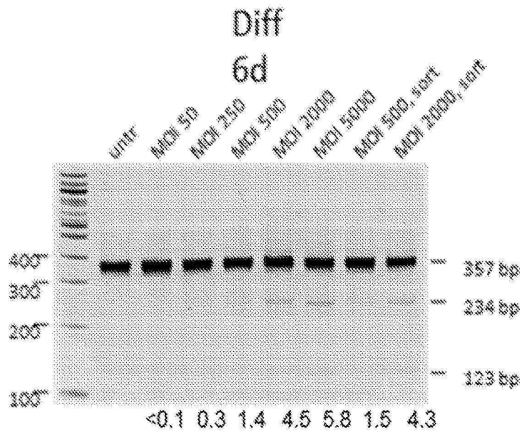
(서열번호: 269)

도면148

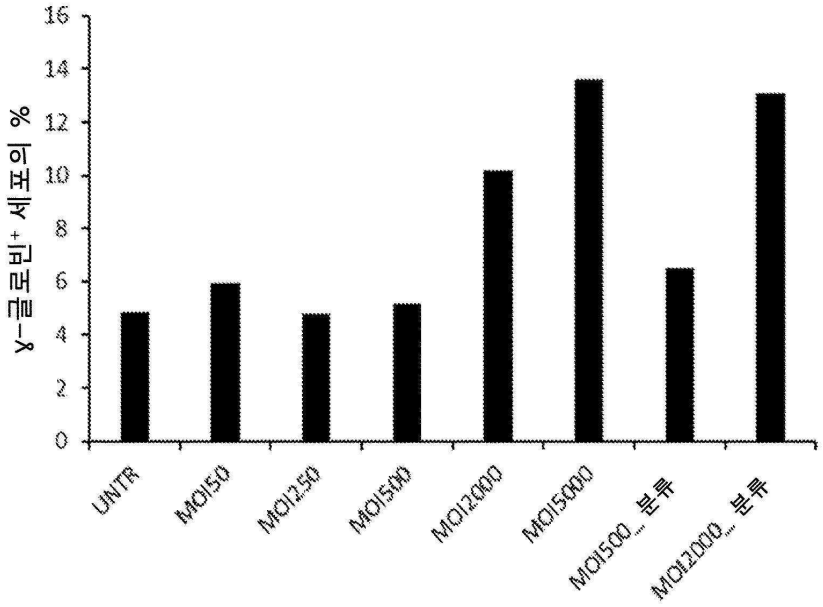



*bcl11aE*의 T7E1



도면149




도면150

염기 편집에 의한 HbF의 재활성화

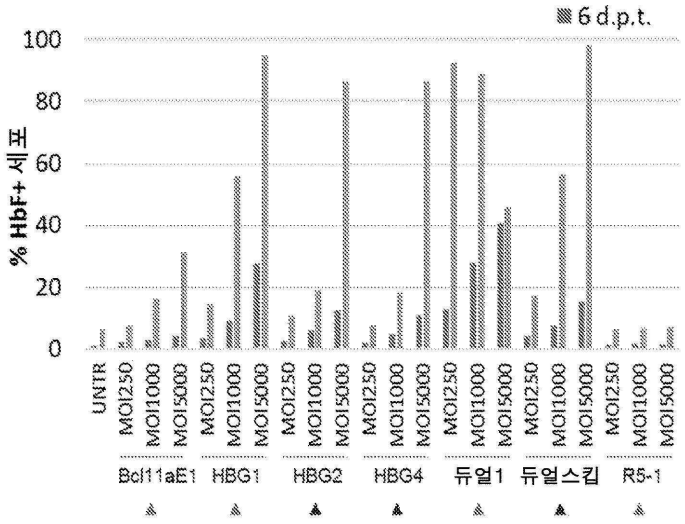
사이티딘 염기 편집기: C → T
아데닌 염기 편집기: A → G

구축된 벡터:

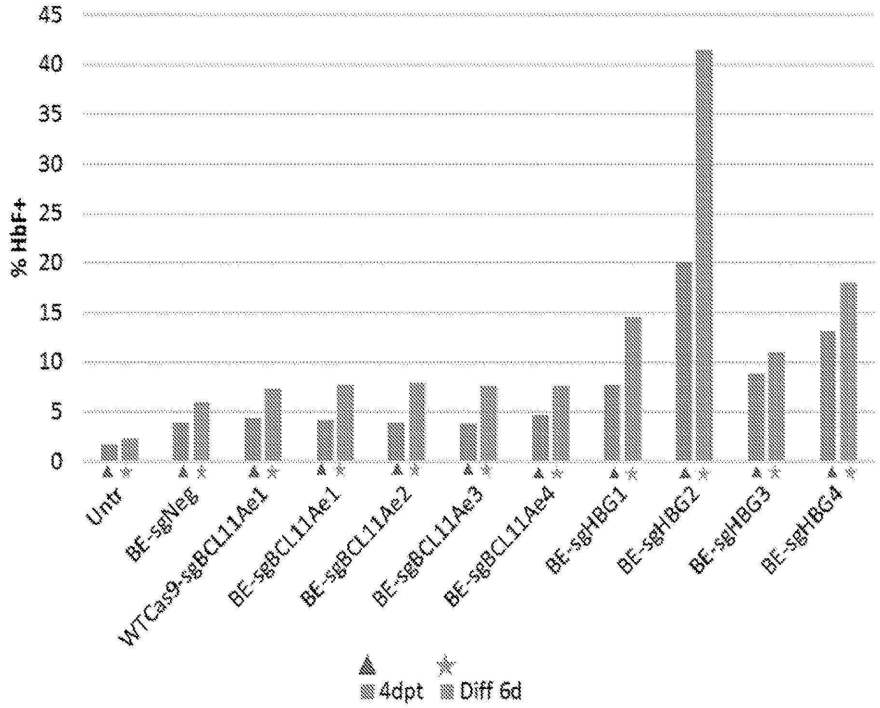
| | |
|---|---|
| pBS-AncBE4max-sgBCL11AE1 | +58 GATA 모티프, G --> A |
| pBS-AncBE4max-sgBCL11AE2 | +58 GATA 모티프, G --> A |
| pBS-ABEmax-sgBCL11AE3 | +58 GATA 모티프, T --> C |
| pBS-ABEmax-sgBCL11AE4 | +58 GATA 모티프, T --> C |
| pBS-AncBE4max-sgHBG1 | 13bp 결실, CCAAT상자, CC --> TT |
| pBS-ABEmax-sgHBG2 | 13bp 결실, CCAAT상자, A --> G |
| pBS-ABEmax-sgHBG3 | 13bp 결실, CCAAT상자, T --> C |
| pBS-ABEmax-sgHBG4 | -198 T --> C |
| pBS-AncBE4max-sgRS-1 | 대조군 |

도면151

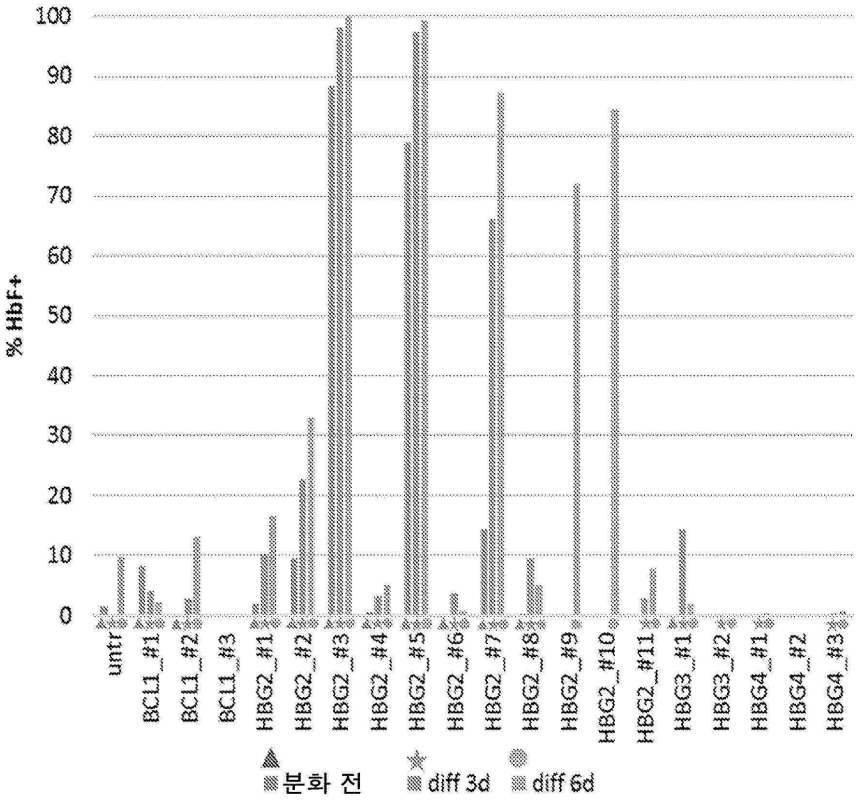
| | 날짜 | 명칭 | 역가 | 메모 |
|---|---|---|---|---|
| | 101918-3 | HDAd5/35++_BE4max-sgBCL11Ae1-FI-mgmtGFP | 0.5 x 10e12 | |
| ⟶ | 041619-2 | HDAd5/35++_BE4max-sgHBG1-FI-mgmtGFP | 매우 낮음 (0.05 x 10e12로 사용) | |
| ⟶ | 032819-1 | HDAd5/35++_ABEopti-sgHBG2-FI-mgmtGFP | 2.8 x 10e12 | |
| ⟶ | 041819-1 | HDAd5/35++_ABEopti-sgHBG4-FI-mgmtGFP | 1.52 x 10e12 | |
| ⟶ | 041619-3 | HDAd5/35++_BE4max-Dual1-FI-mgmtGFP | 매우 낮음 (0.05 x 10e12로 사용) | sgBcl1 + sgHBG1 |
| | 041619-1 | HDAd5/35++_ABEopti-Dual-skip-FI-mgmtGFP | 2.5 x 10e12 | sgHBG2 + sgHBB-Stop |
| | 041819-2 | HDAd5/35++_BE4max-sgR5-1-FI-mgmtGFP | 0.9 x 10e12 | 비-표 적화 대조군 |



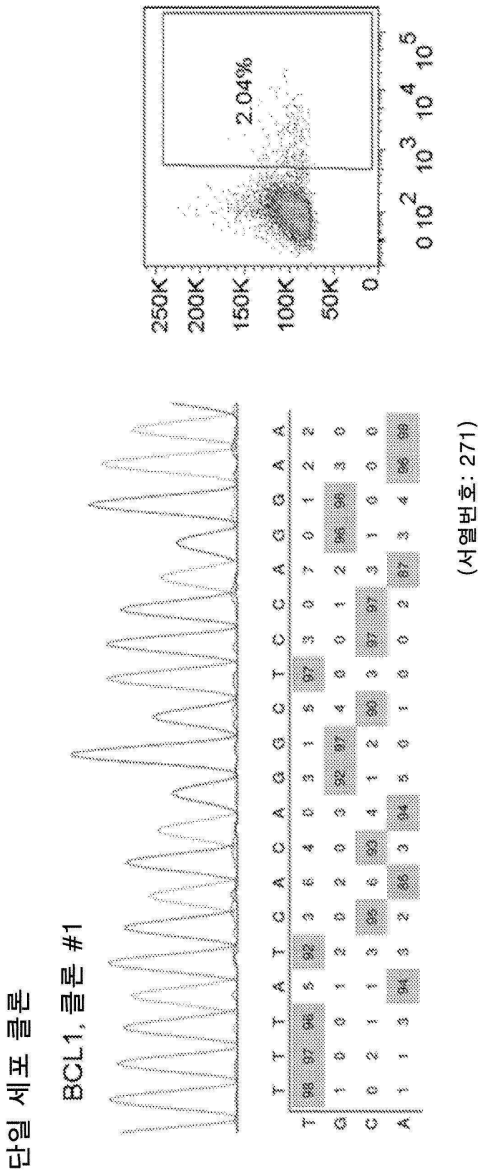
도면152


45
40
35
30
25
20
15
10
5
0
% HbF+
Untr
BE-sgNeg
WTCas9-sgBCL11Ae1
BE-sgBCL11Ae1
BE-sgBCL11Ae2
BE-sgBCL11Ae3
BE-sgBCL11Ae4
BE-sgHBG1
BE-sgHBG2
BE-sgHBG3
BE-sgHBG4
4dpt
Diff 6d

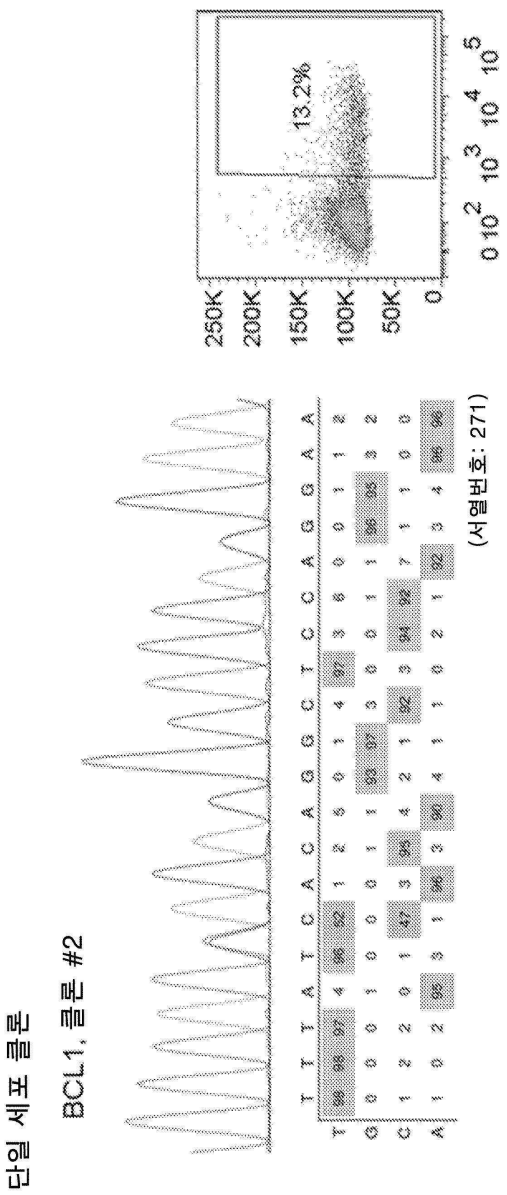
도면153


100
90
80
70
60
50
40
30
20
10
0
% HbF+
untr
BCL1_#1
BCL1_#2
BCL1_#3
HBG2_#1
HBG2_#2
HBG2_#3
HBG2_#4
HBG2_#5
HBG2_#6
HBG2_#7
HBG2_#8
HBG2_#9
HBG2_#10
HBG2_#11
HBG3_#1
HBG3_#2
HBG4_#1
HBG4_#2
HBG4_#3
분화 전
diff 3d
diff 6d

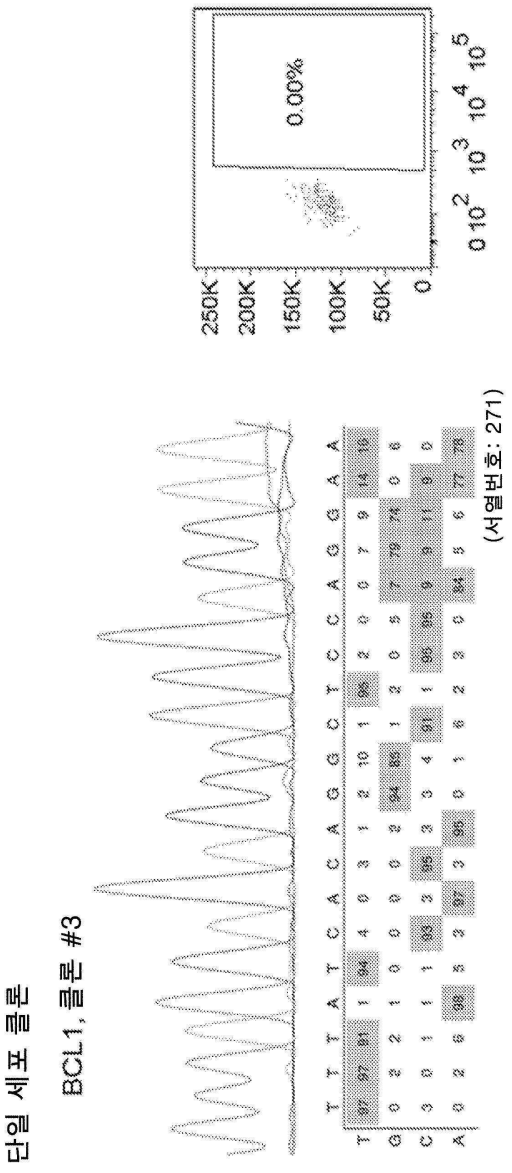
도면154a


단일 세포 클론
BCL1, 클론 #1
2.04%
250K
200K
150K
100K
50K
0
0 10^2 10^3 10^4 10^5
(서열번호: 271)

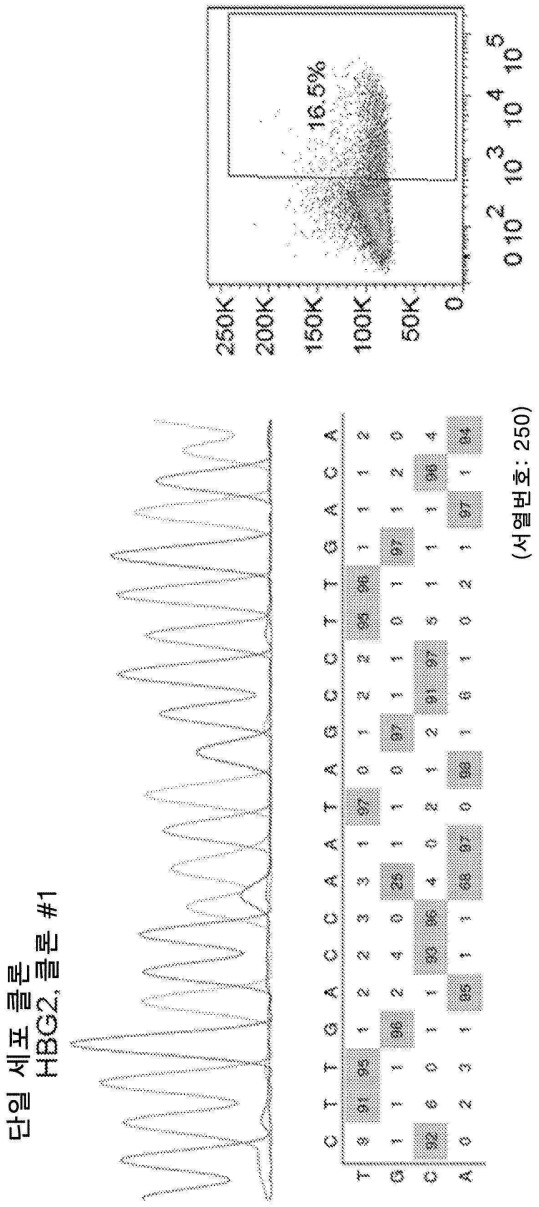
도면154b


단일 세포 클론
BCL1, 클론 #2
13.2%
250K
200K
150K
100K
50K
0
0 10^2 10^3 10^4 10^5
(서열번호: 271)

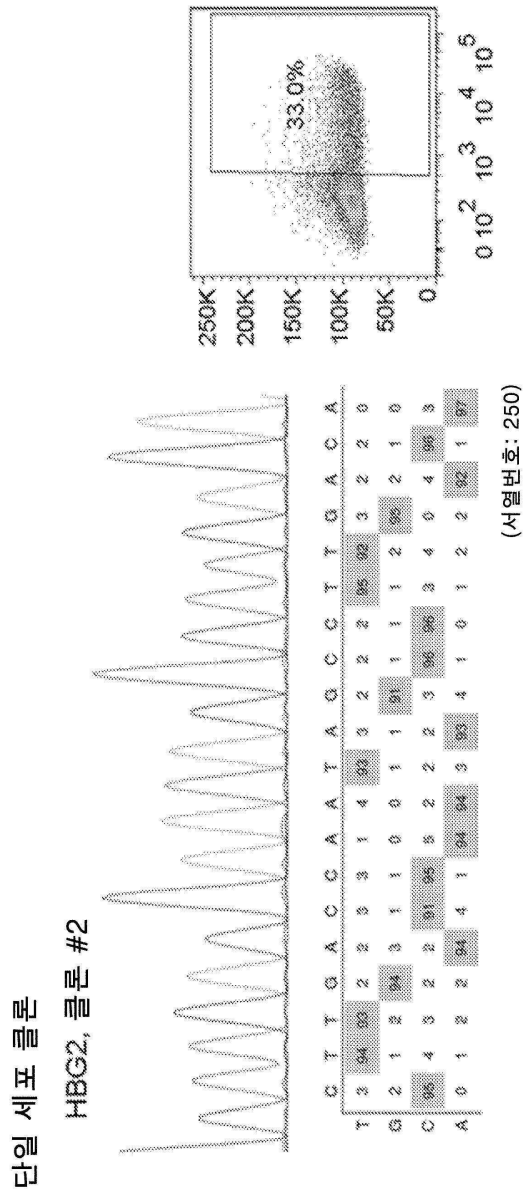
도면154c


단일 세포 클론
BCL1, 클론 #3
250K
200K
150K
100K
50K
0
0.00%
0 10^2 10^3 10^4 10^5
(서열번호: 271)

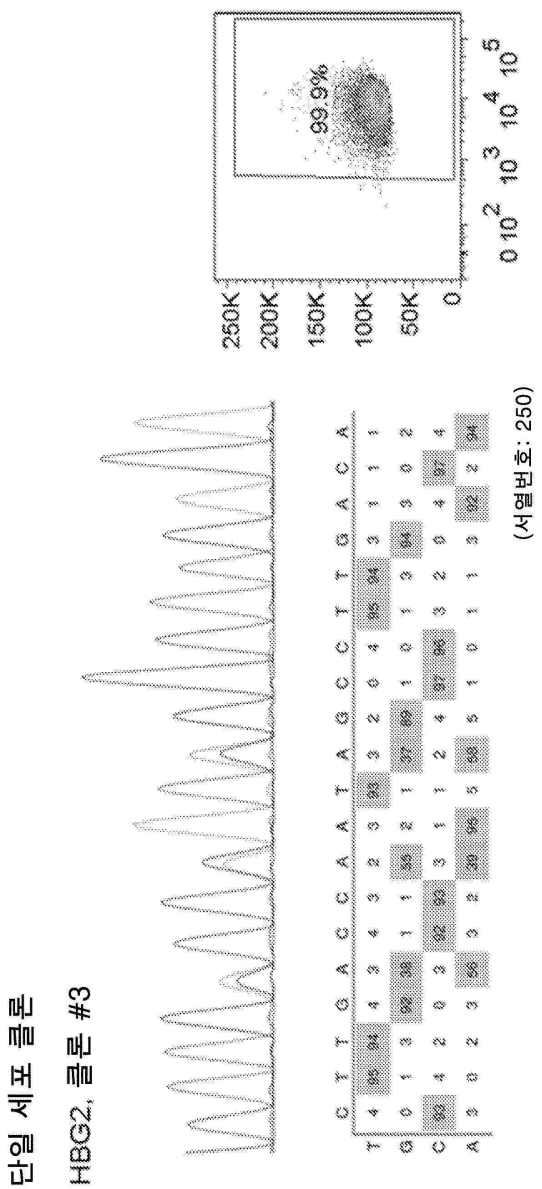
도면154d


단일 세포 클론
HBG2, 클론 #1
250K
200K
150K
100K
50K
0
16.5%
0 10^2 10^3 10^4 10^5
(서열번호: 250)

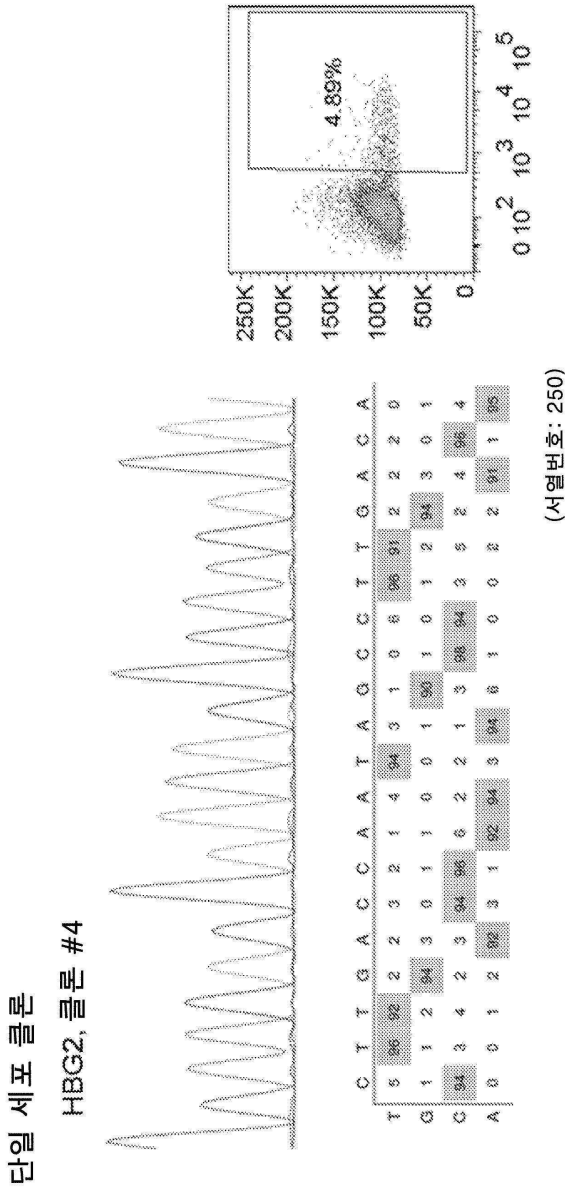
도면154e


단일 세포 클론
HBG2, 클론 #2
33.0%
250K
200K
150K
100K
50K
0
0 10^2 10^3 10^4 10^5
(서열번호: 250)

도면154f


단일 세포 클론
HBG2, 클론 #3
99.9%
250K
200K
150K
100K
50K
0
0 10^2 10^3 10^4 10^5
(서열번호: 250)

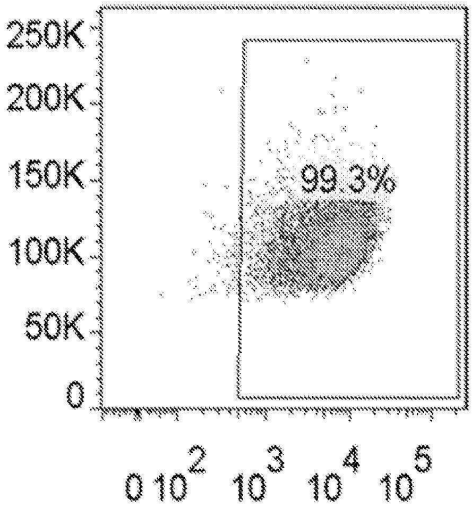
도면154g


단일 세포 클론
HBG2, 클론 #4
250K
200K
150K
100K
50K
0
4.89%
0 10^2 10^3 10^4 10^5
(서열번호: 250)

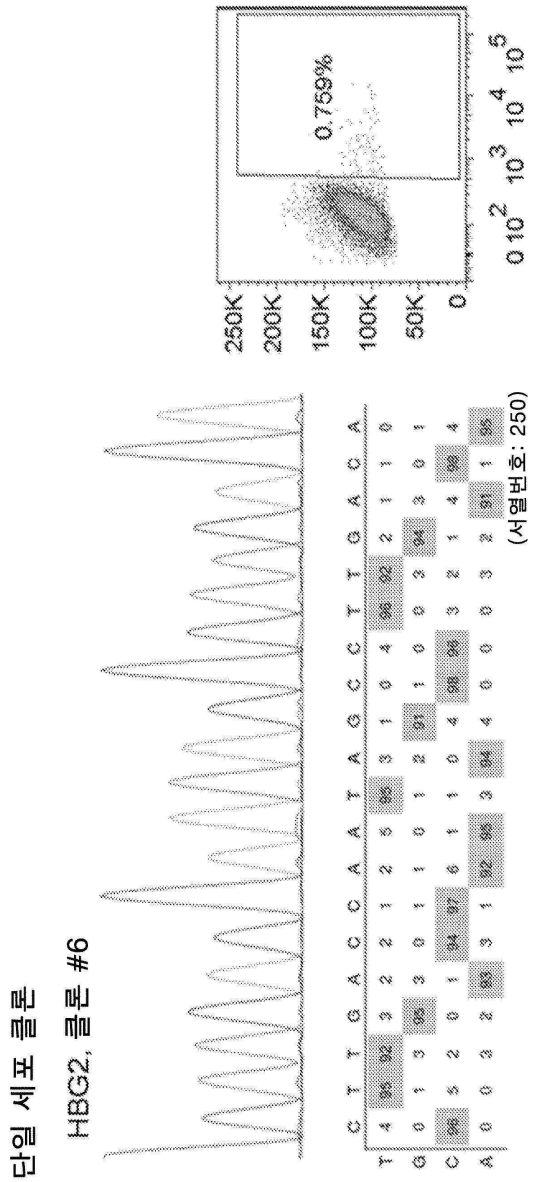
도면154h

단일 세포 클론

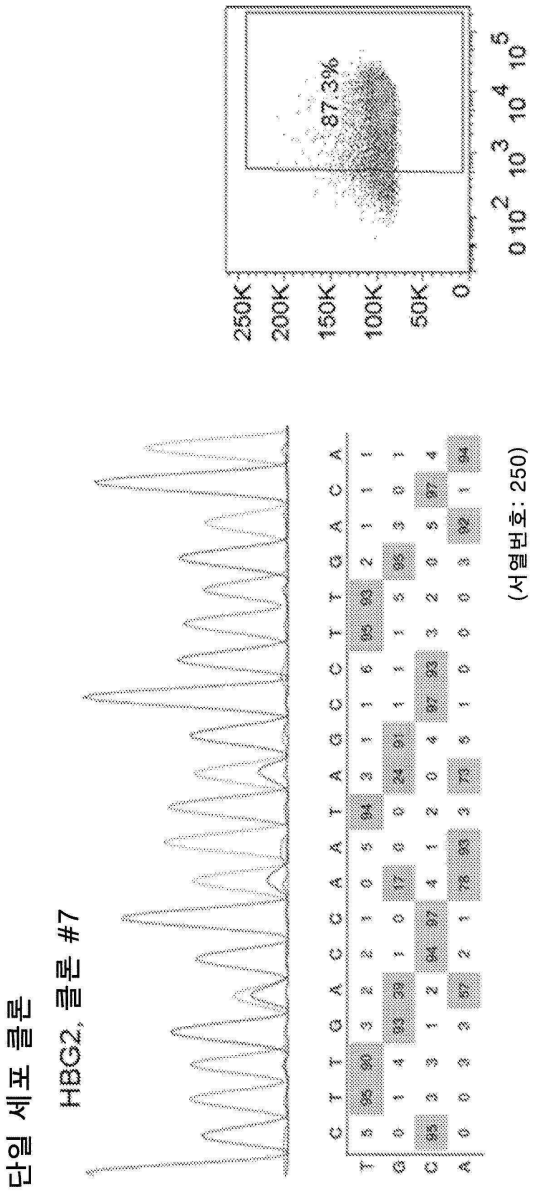
HBG2, 클론 #5



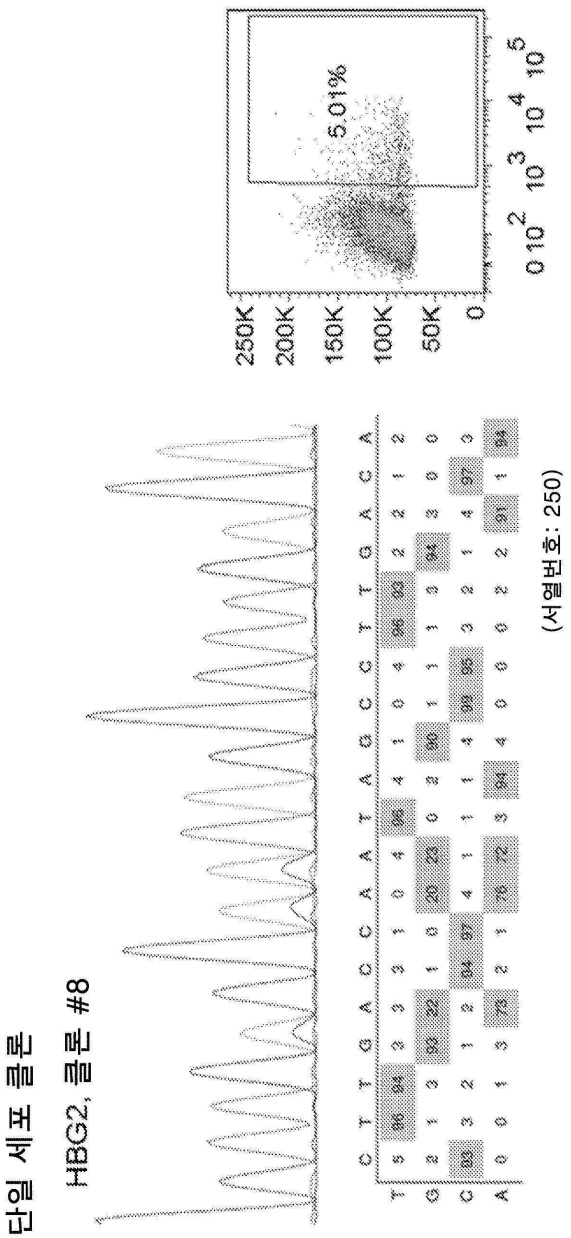
도면154i


단일 세포 클론
HBG2, 클론 #6
0.759%
250K
200K
150K
100K
50K
0
0 10^2 10^3 10^4 10^5
(서열번호: 250)

도면154j


단일 세포 클론
HBG2, 클론 #7
87.3%
250K
200K
150K
100K
50K
0
0 10^2 10^3 10^4 10^5
(서열번호: 250)

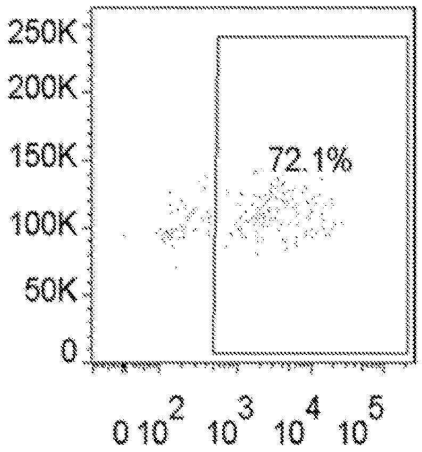
도면154k


단일 세포 클론
HBG2, 클론 #8
(서열번호: 250)
5.01%
250K
200K
150K
100K
50K
0
0 10^2 10^3 10^4 10^5

도면154l

단일 세포 클론

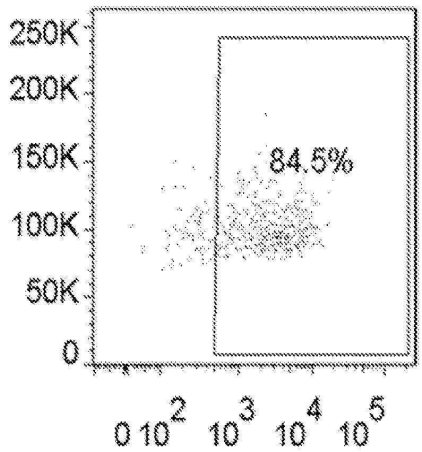
HBG2, 클론 #9




도면154m

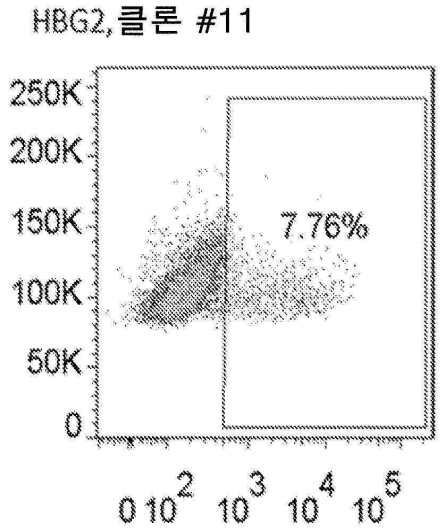
단일 세포 클론

HBG2, 클론 #10

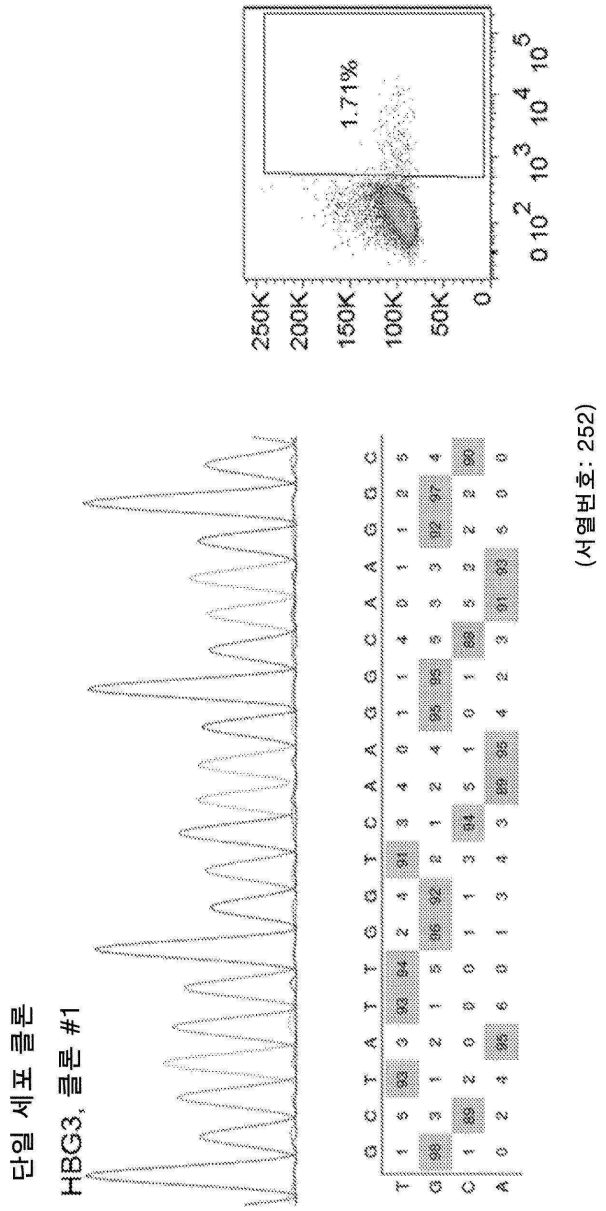


도면154n

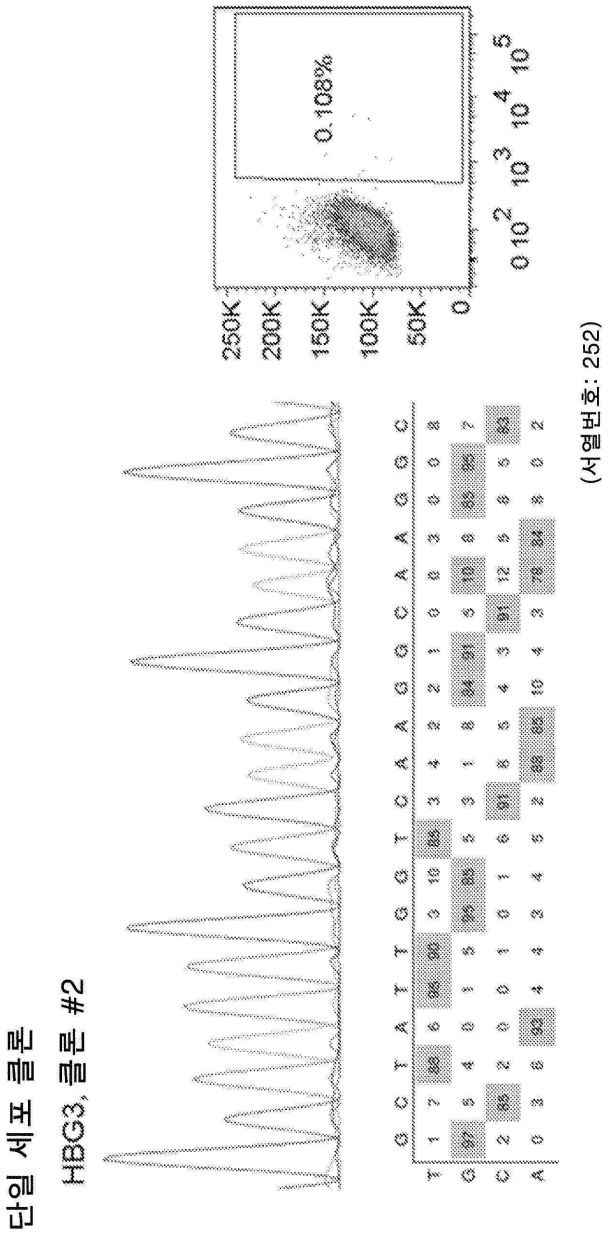
단일 세포 클론



도면154o

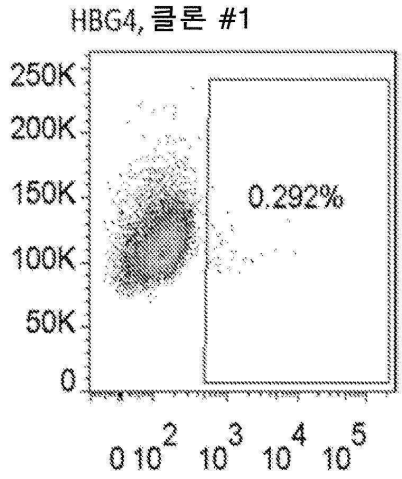

단일 세포 클론
HBG3, 클론 #1
250K
200K
150K
100K
50K
0
1.71%
0 10^2 10^3 10^4 10^5
(서열번호: 252)

도면154p


단일 세포 클론
HBG3, 클론 #2
0.108%
250K
200K
150K
100K
50K
0
0 10^2 10^3 10^4 10^5
(서열번호: 252)

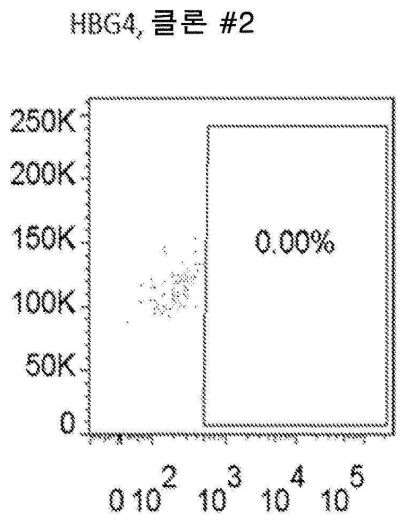
도면154q

단일 세포 클론



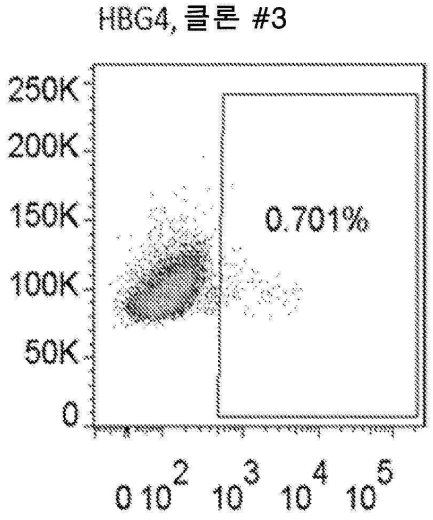

도면154r

단일 세포 클론

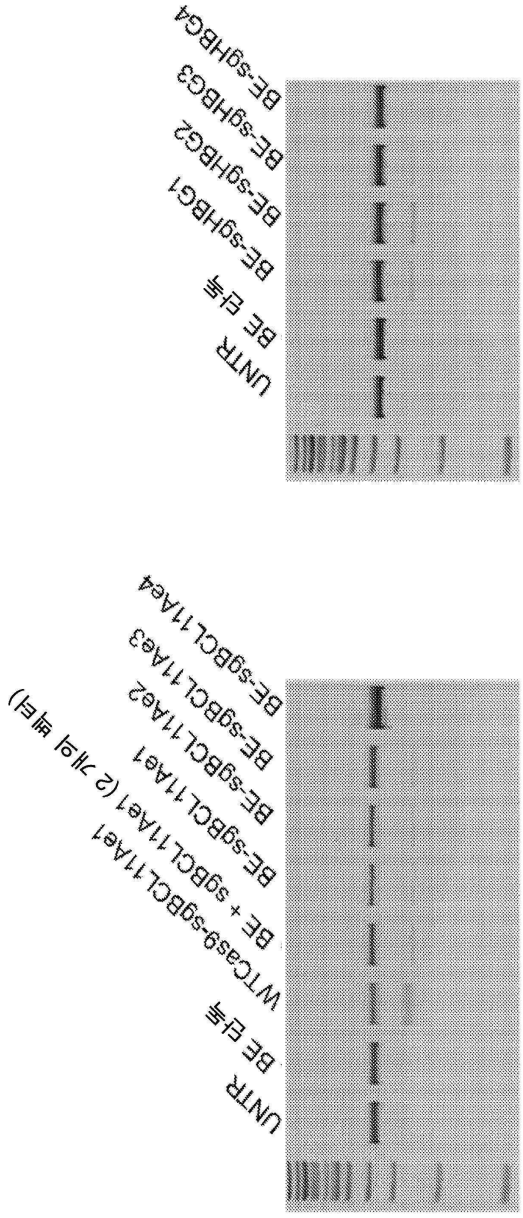


도면154s

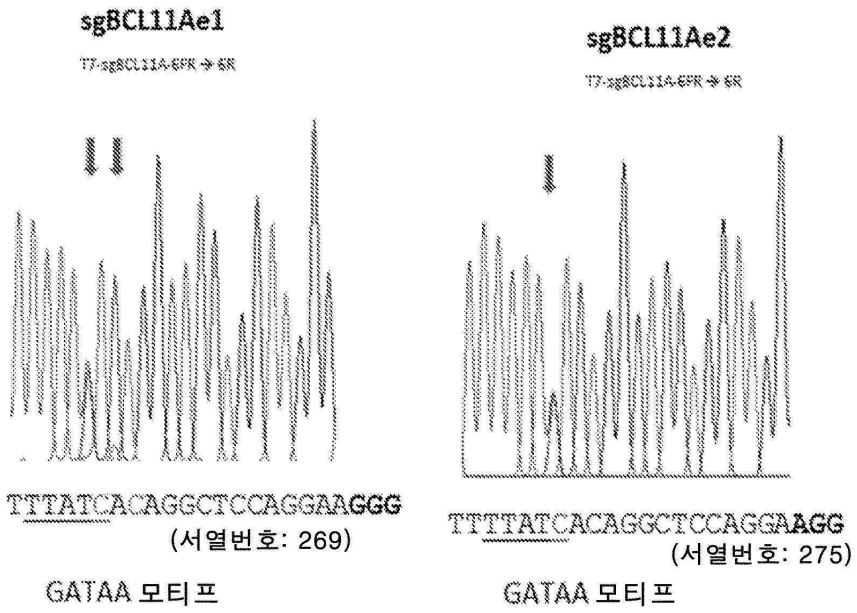
단일 세포 클론



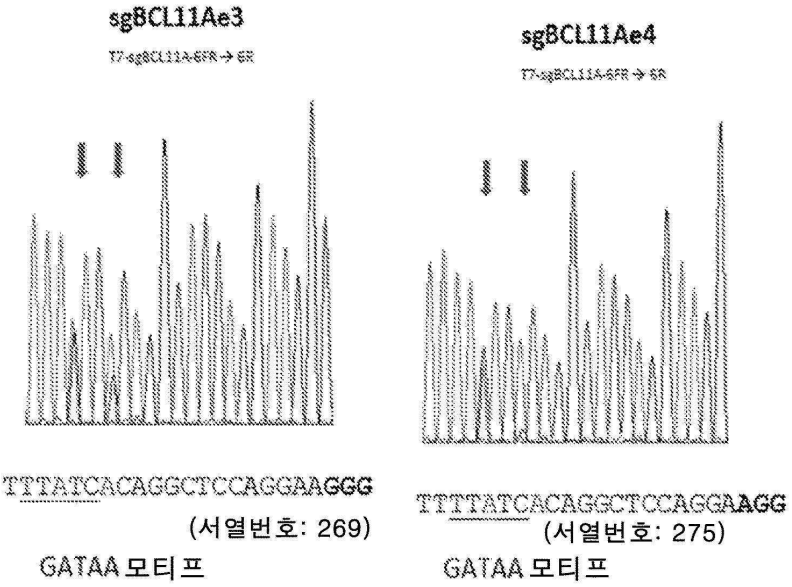
도면155


UNTR
BE 단독
BE-sgHBG1
BE-sgHBG2
BE-sgHBG3
BE-sgHBG4
UNTR
BE 단독
WTCas9-sgBCL11Ae1
BE + sgBCL11Ae1 (2 개의 벡터)
BE-sgBCL11Ae1
BE-sgBCL11Ae2
BE-sgBCL11Ae3
BE-sgBCL11Ae4

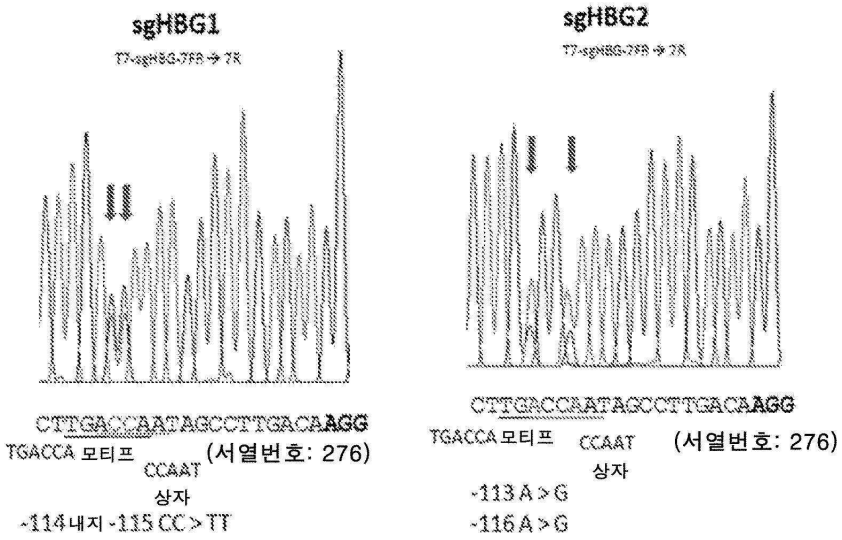
도면156a



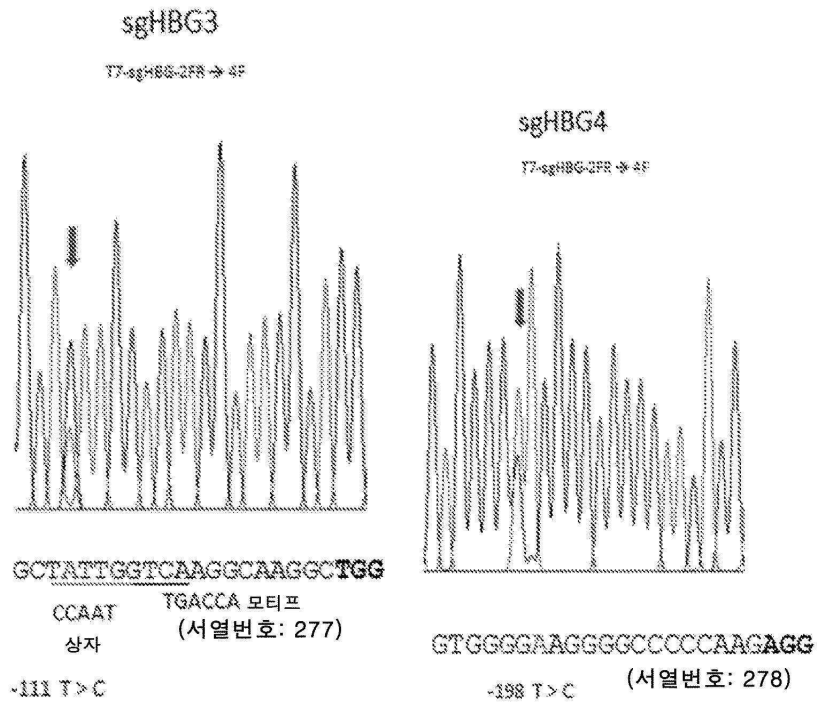

도면156b



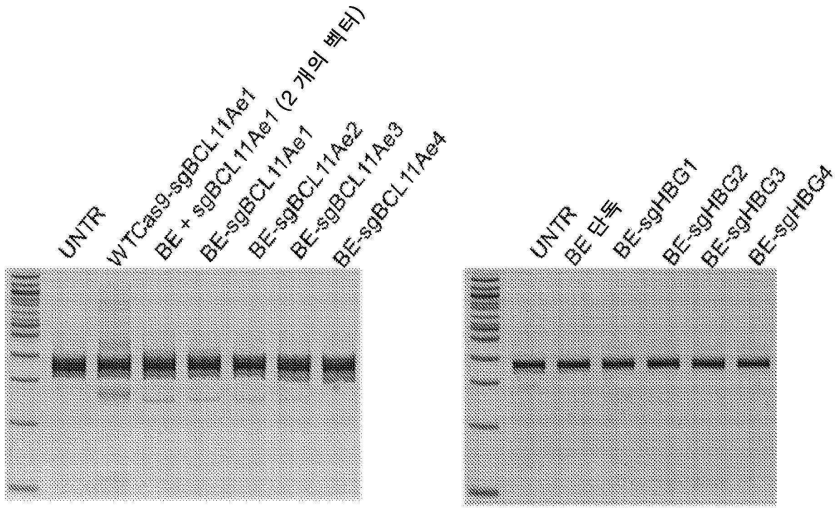
도면156c



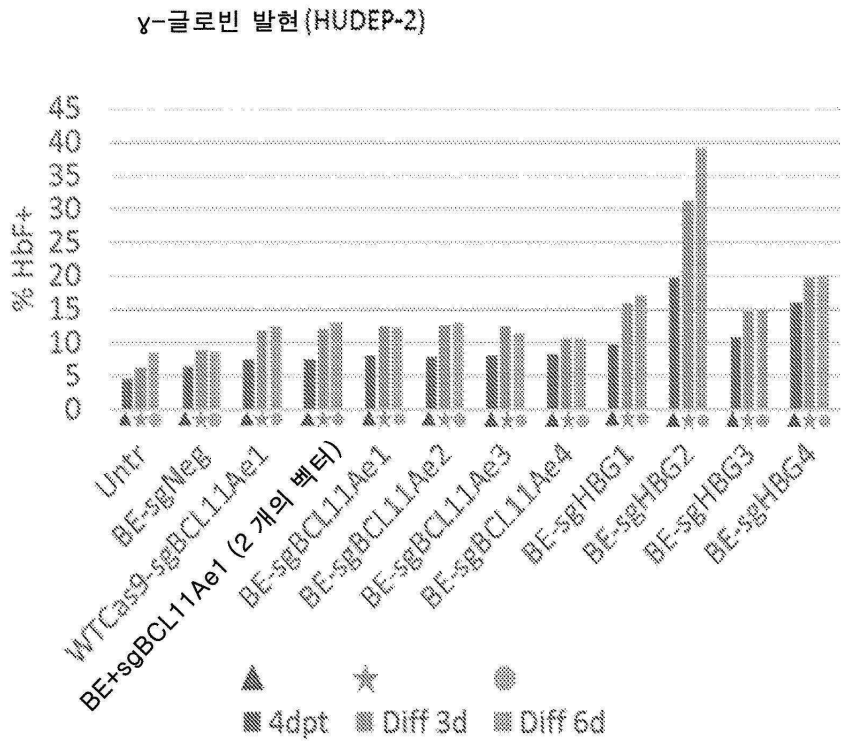

도면156d



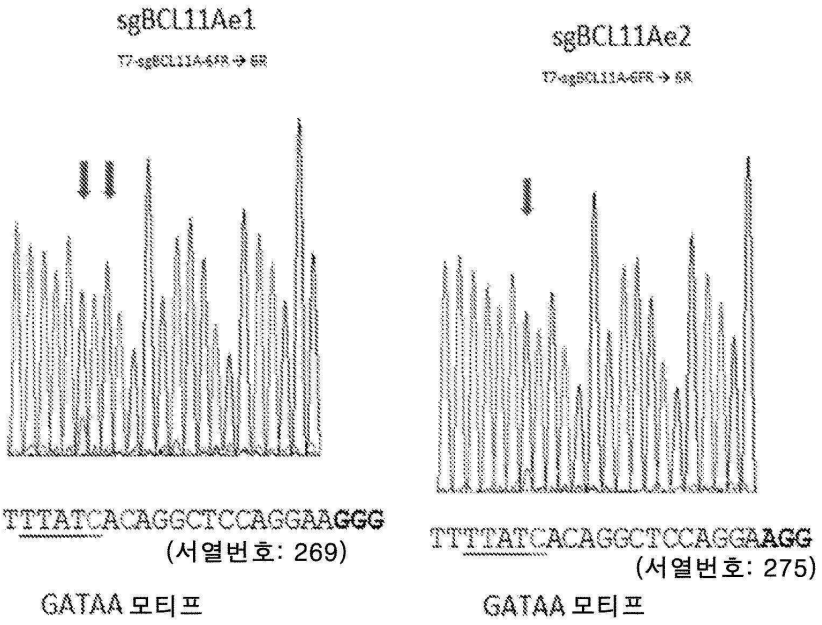
도면157


UNTR
WTCas9-sgBCL11Ae1
BE + sgBCL11Ae1 (2 개의 벡터)
BE-sgBCL11Ae1
BE-sgBCL11Ae2
BE-sgBCL11Ae3
BE-sgBCL11Ae4
UNTR
BE 단독
BE-sgHBG1
BE-sgHBG2
BE-sgHBG3
BE-sgHBG4

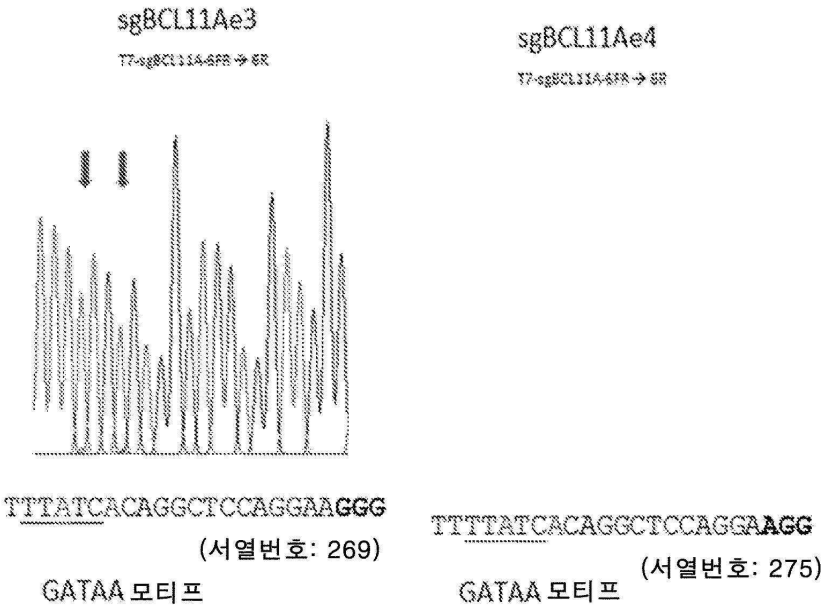
도면158


γ-글로빈 발현 (HUDEP-2)
% HbF+
45
40
35
30
25
20
15
10
5
0
Untr
BE-sgNeg
WTCas9-sgBCL11Ae1
BE+sgBCL11Ae1 (2 개의 벡터)
BE-sgBCL11Ae1
BE-sgBCL11Ae2
BE-sgBCL11Ae3
BE-sgBCL11Ae4
BE-sgHBG1
BE-sgHBG2
BE-sgHBG3
BE-sgHBG4
4dpt
Diff 3d
Diff 6d

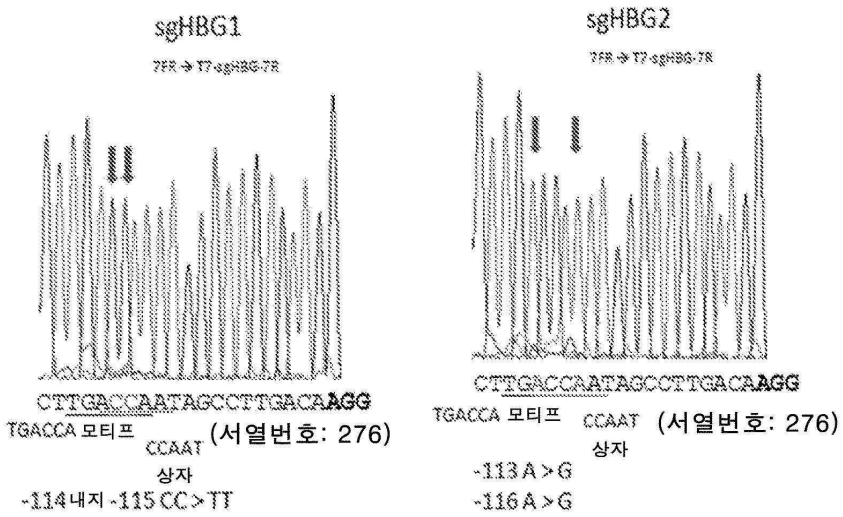
도면159a



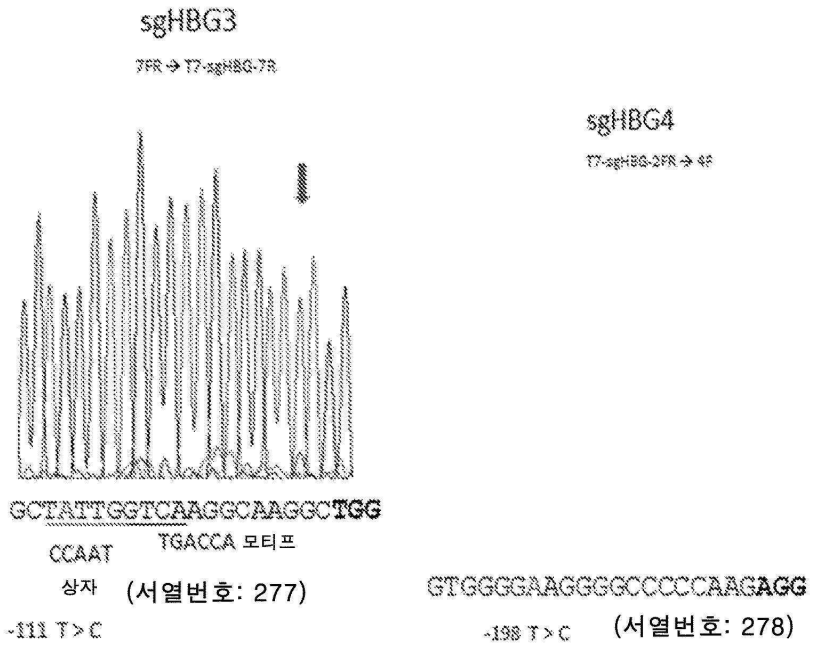

도면159b



도면159c




도면159d

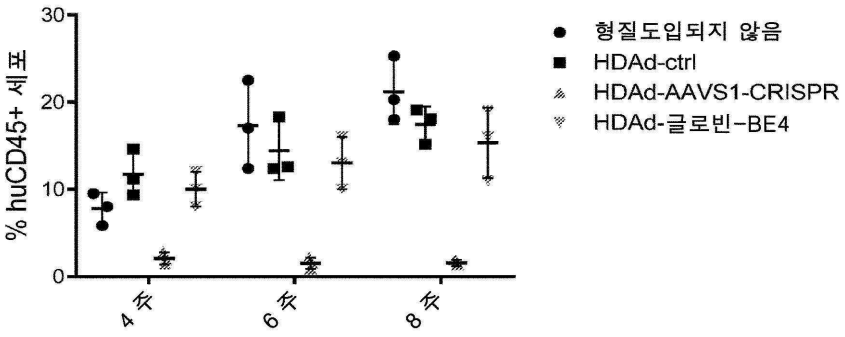


도면160

HDAd 바이러스 생산을 위해 선택된 작제물 (Maxi 제조 사용)

| 구축된 벡터: | | 상응하는 Ad 작제물: |
|---|---|---|
| pBS-AncBE4max-sgBCL11AE1 | +58 GATA 모티프, G --> A | pHCA-BE4max-sgBCL11Ae1-Fl-mgmt/GFP |
| pBS-AncBE4max-sgBCL11AE2 | +58 GATA 모티프, G --> A | |
| pBS-ABEmax-sgBCL11AE3 | +58 GATA 모티프, T --> C | |
| pBS-ABEmax-sgBCL11AE4 | +58 GATA 모티프, T --> C | |
| pBS-AncBE4max-sgHBG1-1 | 13bp 결실, CCAAT상자, CC --> TT | pHCA-ABEmax-sgHBG2-Fl-mgmt/GFP |
| pBS-ABEmax-sgHBG1-2 | 13bp 결실, CCAAT상자, A --> G | |
| pBS-ABEmax-sgHBG1-3 | 13bp 결실, CCAAT상자, T --> C | pHCA-ABEmax-sgHBG4-Fl-mgmt/GFP |
| pBS-ABEmax-sgHBG1-4 | -198 T --> C | |
| pBS-AncBE4max-sgR5-1 | 대조군 | pHCA-BE4max-sgR5-1-Fl-mgmt/GFP(Ctrl) |

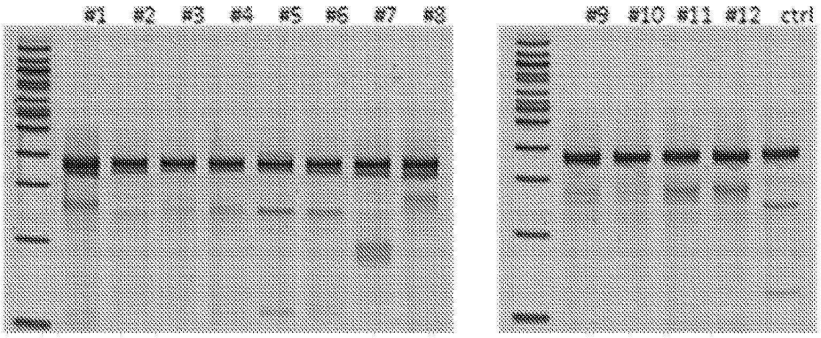
도면161



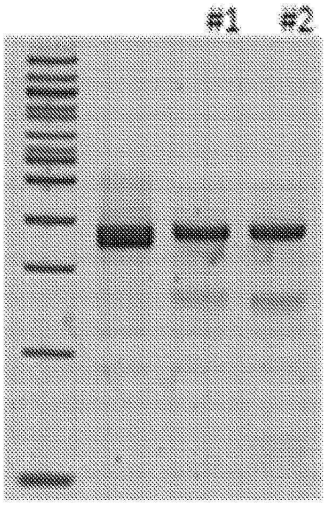
도면162

HUDEP-2 세포의 일시적 형질감염:
(T7E1에 의한 절단)

Bcl11a 인핸서



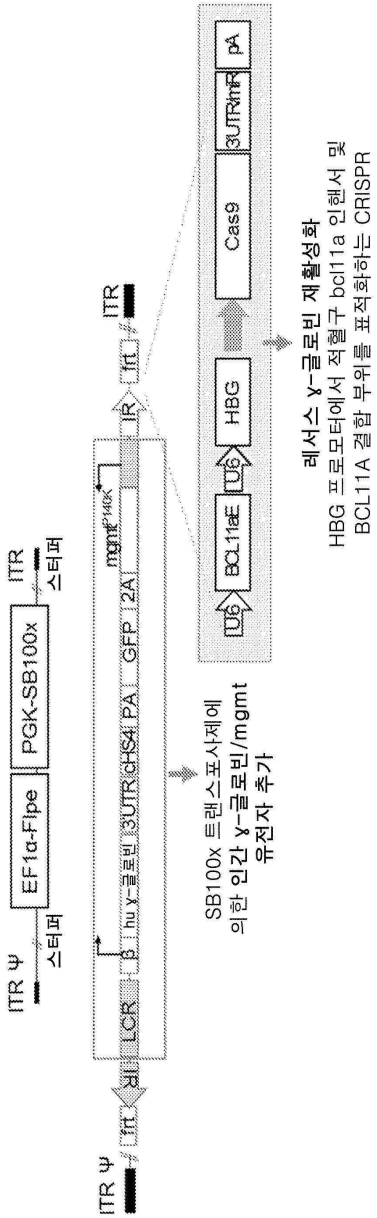

HBG 프로모터



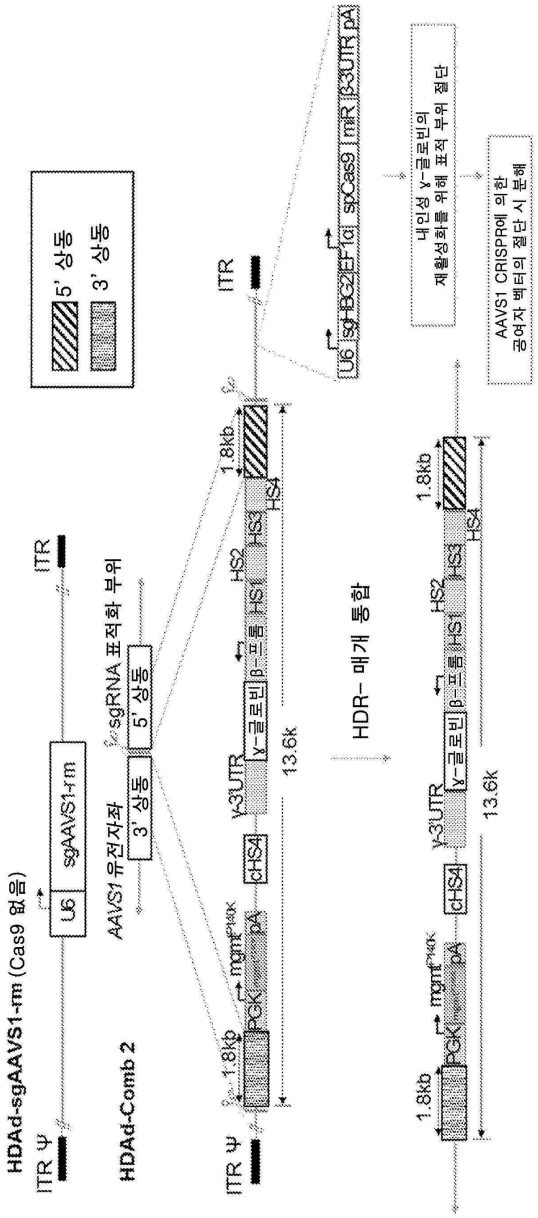
도면163


이중 염기 편집 벡터 적용
BCL11A 결합 부위를 파괴함으로써 γ-글로빈의 재활성화 + CRISPR-STOP에 의한 돌연변이체 HBB의 종결
GAG → GTG 겸상 돌연변이
사이티딘 편집기에 의한 TGG → TAA
ATGGTGCATCTGACTCCTGAGGAGAAGTCTGCCGTTACTGCCCTGTGGGGCAAGGTGA
ACGTGGATGAAGTTGGTGGTGAGGCCCTGGGCAGGCTGCTGGTGGTCTACCCTTGGAC
CCAGAGGTTCTTTGAGTCCTTTGGGGATCTGTCCACTCCTGATGCTGTTATGGGCAAC
CCTAAGGTGAAGGCTCATGGCAAGAAAGTGCTCGGTGCCTTTAGTGATGGCCTGGCTC
ACCTGGACAACCTCAAGGGCACCTTTGCCACACTGAGTGAGCTGCACTGTGACAAGCT
GCACGTGGATCCTGAGAACTTCAGGCTCCTGGGCAACGTGCTGGTCTGTGTGCTGGCC
CATCACTTTGGCAAAGAATTCACCCCACCAGTGCAGGCTGCCTATCAGAAAGTGGTGG
CTGGTGTGGCTAATGCCCTGGCCCACAAGTATCACTAA
(서열번호: 279)
사이티딘 편집기에 의한 CAG → TAG

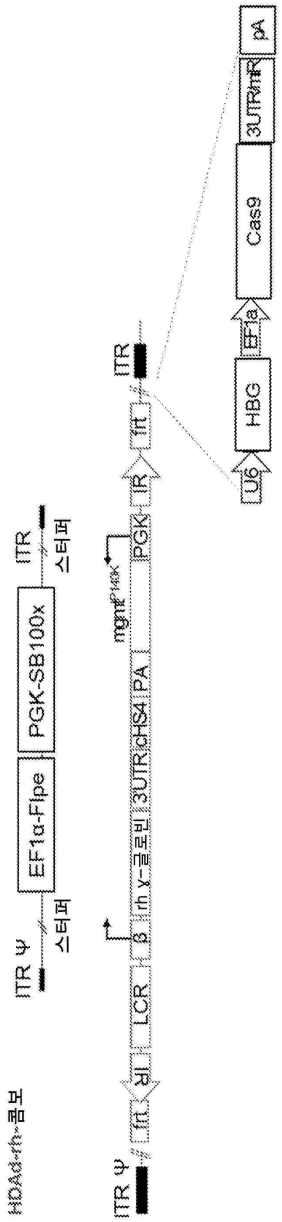
도면164


ITR Ψ
스터퍼
EF1a-Flpe
PGK-SB100x
ITR
스터퍼
ITR Ψ
frt
IR
LCR
β
hu γ-글로빈 3UTR
cHS4
PA
GFP
2A
mgmtP140K
IR
frt
ITR
SB100x 트랜스포사제에
의한 인간 γ-글로빈/mgmt
유전자 추가
U6
BCL11aE
U6
HBG
Cas9
3UTRWHP
pA
레서스 γ-글로빈 재활성화
HBG 프로모터에서 적혈구 bcl11a 인핸서 및
BCL11A 결합 부위를 표적화하는 CRISPR

도면165


5' 상동
3' 상동
HDAd-sgAAVS1-rm (Cas9 없음)
ITR Ψ
U6
sgAAVS1-rm
ITR
HDAd-Comb 2
AAVS1 유전자좌
3' 상동
sgRNA 표적화 부위
5' 상동
ITR Ψ
1.8kb
PGK
mgmt140K
pA
cHS4
γ-3'UTR
γ-글로빈
β-프롬
HS1
HS2
HS3
HS4
13.6k
HDR- 매개 통합
ITR
U6 sg-HBG2/EF1α spCas9 mR β3UTR pA
내인성 γ-글로빈의
재활성화를 위해 표적 부위 절단
AAVS1 CRISPR에 의한
공여자 벡터의 절단 시 분해

도면166




## 서 열 목 록

SEQUENCE LISTING

<110> Fred Hutchinson Cancer Research Center

University of Washington

<120> RECOMBINANT AD35 VECTORS AND RELATED GENE THERAPY IMPROVEMENTS

<130> F053-0107PCT/19-132-WO-PCT

<150> US 62/869,907

<151> 2019-07-02

<150> US 62/935,507

<151> 2019-11-14

<150> US 63/009,385

<151> 2020-04-13

<160> 325

<170> PatentIn version 3.5

<210> 1

<211> 720

<212> DNA

<213> Artificial Sequence

<220><223> 5 end vector sequence, Sleeping Beauty IR/DR sequence,

integration junction (chr15, 6805206)

<400> 1

ccctgggatt ccccaaggca ggggcgagtc cttttgtatg aattactcaa atcgataact 60

agaaacttaa ttaacaacga gatcttataa tttgcatact tctgcctgct ggggactttc 120

cacaccctag ctgacacaag aatttgaaat acatccacag gtacacctcc aattgactca 180

aatgatgtca attagtctat cataatcttc taaagccatg acatcatttt aactggaatt 240

ttccaagctg tttaaaggca cagtcaactt agtgtatgta aacttctgac ccactggaat 300

tgtgatacag tgaattataa gtgaaataat ctgtctgtaa acaattgttg gaaaaatgac 360

ttgtgtcatg cacaaagtag atgtcctaac tgacttgcca aaactattgt ttgttaacaa 420

gaaatttgtg gagtagttga aaaacgagtt ttaatgactc caacttaagt gtatgtaaac 480

ttccgacttc aactgtaaga atggcccatt catctatagt agcacacaat atttgcattt 540

gtgcgacagt ataagggaca attatgctat caggcatttt tccaaagtga gtaatcgaag 600

tttttatacc tttgtgtgcc atgtttgcta ccatggtggg ataatcttac acgcgttctc 660

gcgaccggcc aggaaagacg caacaaaccg gaatcttctg cggcaaaagc tttattgctt 720

<210> 2

<211> 607

<212> DNA

<213> Artificial Sequence

<220><223> 5 end vector sequence, Sleeping Beauty IR/DR sequence,

integration junction (chrX, 16897322)

<220><221> misc_feature

<222> (594)..(594)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (596)..(597)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (602)..(602)

<223> n is a, c, g, or t

<400> 2

tagaaactta attaacaacg agatcttata atttgcatac ttctgcctgc tggggacttt 60

ccacacccta gctgacacaa gaatttgaaa tacatccaca ggtacacctc caattgactc 120

aaatgatgtc aattagtcta tcataatctt ctaaagccat gacatcattt taactggaat 180

tttccaagct gtttaaaggc acagtcaact tagtgtatgt aaacttctga cccactggaa 240

ttgtgataca gtgaattata agtgaaataa tctgtctgta aacaattgtt ggaaaaatga 300

cttgtgtcat gcaaagtaga tgtcctaact gacttgccaa aactattgtt tgttaacaag 360

aaatttgtgg agtagttgaa aaacgagttt taatgactcc aacttaagtg tatgtaaact 420

tccgacttca actgtacaag tagaccaaat atccatatac ataaaagaaa aaaatagaaa 480

aaatttctag tgacagaaaa atgacaaaga acatactgtt tattactact attaagatgt 540

ttgcttccat tacactcata tgagtcatga tatttttct tcatttttt ctantnncac 600

tngaaat 607

<210> 3

<211> 520

<212> DNA

<213> Artificial Sequence

<220><223> 3 end vector sequence, Sleeping Beauty IR/DR sequence,
integration junction (chr4, 10207667)

<400> 3

gttgctagga atgagccaaa ttcatctgta ttaaacagtg ggagcttgtg gaaggctact 60

cgaaatgttt gacccaagtt aaacaattta aaggcaatgc taccaaatac taattgagtg 120

tatgttaact tctgacccac tgggaatgtg atgaaagaaa taaaagctga aatgaatcat 180

tctctctact attattctga tatttcacat tcttaaaata aagtggtgat cctaactgac 240

cttaagacag ggaatcttta ctcggattaa atgtcaggaa ttgtgaaaaa gtgagtttaa 300

atgtatttgg ctaaggtgta tgtaaacttc cgacttcaac tgtatatcct ccccgttgca 360

ccctcttgat gatgctgaga tgaacacaga tgctcactcc ttgagggctc taagcttatg 420

ctgacacaga cacaggtgct cacttctatg aatggcctaa gatttgagga catcatgagg 480

acaagtgtga taaaatcttg gaacaacctc ccagaggtct 520

<210> 4

<211> 74

<212> DNA

<213> Artificial Sequence

<220><223> Sleeping beauty IR/DR sequence, integration junction (chr7, 79796094)

<400> 4

acttaagtgt atgtaaactt ccgacttcaa ctgtagggta cctgattctc tgggcatctc 60

tgcccactac catg 74

<210> 5

<211> 75

<212> DNA

<213> Artificial Sequence

<220><223> Sleeping beauty IR/DR sequence, Integration junction (repeat region)

<400> 5

acttaagtgt atgtaaactt ccgacttcaa ctgtaaattt tccacctttt tcagttttcc 60

tcgccatatt tcatg 75

<210> 6

<211> 64

<212> DNA

<213> Artificial Sequence

<220><223> indels from HDAd-CRISPR-transduced CD34+ cells

<400> 6

tgtcccctcc accccacagt ggggccacta gggacaggat tggtgacaga aaagccccat 60

cctt 64

<210> 7

<211> 44

<212> DNA

<213> Artificial Sequence

<220><223> indels from HDAd-CRISPR-transduced CD34+ cells

<400> 7

tgtcccctcc accccacagt gggtgacaga aaagccccat cctt 44

<210> 8

<211> 45

<212> DNA

<213> Artificial Sequence

<220><223> indels from HDAd-CRISPR-transduced CD34+ cells

<400> 8

tgtcccctcc accccacaga ttggtgacag aaaagcccca tcctt 45

<210> 9

<211> 52

<212> DNA

<213> Artificial Sequence

<220><223> indels from HDAd-CRISPR-transduced CD34+ cells

<400> 9

tgtcccctcc accccacagt ggggccacta gggacagaaa agccccatcc tt 52

<210> 10

<211> 53

<212> DNA

<213> Artificial Sequence

<220><223> indels from HDAd-CRISPR-transduced CD34+ cells

<400> 10

tgtcccctcc accccacagt ggggccacta ggtgacagaa aagccccatc ctt 53

<210> 11

<211> 54

<212> DNA

<213> Artificial Sequence

<220><223> indels from HDAd-CRISPR-transduced CD34+ cells

<400> 11

tgtcccctcc accccacagt ggggccacta gggtgacaga aaagccccat cctt 54

<210> 12

<211> 54

<212> DNA

<213> Artificial Sequence

<220><223> indels from HDAd-CRISPR-transduced CD34+ cells

<400> 12

tgtcccctcc accccacagt ggggcaggat tggtgacaga aaagccccat cctt 54

<210> 13

<211> 58

<212> DNA

<213> Artificial Sequence

<220><223> indels from HDAd-CRISPR-transduced CD34+ cells

<400> 13

tgtcccctcc accccacagt ggggccacta ggattggtga cagaaaagcc ccatcctt 58

<210> 14

<211> 59

<212> DNA

<213> Artificial Sequence

<220><223> indels from HDAd-CRISPR-transduced CD34+ cells

<400> 14

tgtcccctcc accccacagt ggggccacta gggattggtg acagaaaagc cccatcctt 59

<210> 15

<211> 63

<212> DNA

<213> Artificial Sequence

<220><223> indels from HDAd-CRISPR-transduced CD34+ cells

<400> 15

tgtcccctcc accccacagt ggggccacta gggacagatt ggtgacagaa aagccccatc 60

ctt 63

<210> 16

<211> 63

<212> DNA

<213> Artificial Sequence

<220><223> indels from HDAd-CRISPR-transduced CD34+ cells

<400> 16

tgtcccctcc accccacagt ggggccacta gggacggatt ggtgacagaa aagccccatc 60

ctt 63

<210> 17

<211> 64

<212> DNA

<213> Artificial Sequence

<220><223> indels from HDAd-CRISPR-transduced CD34+ cells

<400> 17

tgtcccctcc accccacagt ggggccacta gggacagagg attggtgaca gaaaagcccc 60

atcc 64

<210> 18

<211> 64

<212> DNA

<213> Artificial Sequence

<220><223> indels from HDAd-CRISPR-transduced CD34+ cells

<400> 18

tgtcccctcc accccacagt ggggccacta gggacagcag gattggtgac agaaaagccc 60

catc 64

<210> 19

<211> 80

<212> DNA

<213> Mus musculus

<400> 19

tgtcccctcc accccacagt ggggccacta gggacaggat tggtgacaga aaagccccat 60

ccttaggcct cctccttcct 80

<210> 20

<211> 73

<212> DNA

<213> Mus musculus

<400> 20

tgtcccctcc accccacagt ggggccacag gattggtgac agaaaagccc catccttagg 60

cctcctcctt cct 73

<210> 21

<211> 75

<212> DNA

<213> Mus musculus

<400

> 21

tgtcccctcc accccacagt ggggccacta gggacaggtg acagaaaagc cccatcctta 60

ggcctcctcc ttcct 75

<210> 22

<211> 75

<212> DNA

<213> Mus musculus

<400> 22

tgtcccctcc accccacagt ggggccacta gggacaggtg acagaaaatc cccatcctta 60

ggcctcctcc ttcct 75

<210> 23

<211> 75

<212> DNA

<213> Mus musculus

<400> 23

tgtcccctcc accccacagt ggggccacta gggattggtg acagaaaagc cccatcctta 60

ggcctcctcc ttcct 75

<210> 24

<211> 110

<212> DNA

<213> Artificial Sequence

<220><223> microRNA regulated gene expression forward primer

<400> 24

ctaggaaagt ctcgctctct gcccctcatc acaaagtctc gctctctgcc cctcacgata 60

aagtctcgct ctctgcccct cattcaaaag tctcgctctc tgcccctcac 110

<210> 25

<211> 110

<212> DNA

<213> Artificial Sequence

<220><223> microRNA regulated gene expression reverse primer

<400> 25

tcgagtgagg ggcagagagc gagacttttg aatgaggggc agagagcgag actttatcgt 60

gaggggcaga gagcgagact ttgtgatgag gggcagagag cgagactttc 110

<210> 26

<211> 76

<212> DNA

<213> Mus musculus

<400> 26

tgtcccctcc accccacagt ggggccacta ggggattggt gacagaaaag ccccatcctt 60

aggcctcctc cttcct 76

<210> 27

<211> 77

<212> DNA

<213> Mus musculus

<400> 27

tgtcccctcc accccacagt ggggccacta ggaggattgg tgacagaaaa gccccatcct 60

taggcctcct ccttcct 77

<210> 28

<211> 77

<212> DNA

<213> Mus musculus

<400> 28

tgtcccctcc accccacagt ggggccacta gcaggattgg tgacagaaaa gccccatcct 60

taggcctcct ccttcct 77

<210> 29

<211> 77

<212> DNA

<213> Mus musculus

<400> 29

tgtcccctcc accccacagt ggggccacta ggcggattgg tgacagaaaa gccccatcct 60

taggcctcct ccttcct 77

<210> 30

<211> 77

<212> DNA

<213> Mus musculus

<400> 30

tgtcccctcc accccacagt ggggccacta gccggattgg tgacagaaaa gccccatcct 60

taggcctcct ccttcct 77

<210> 31

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> HBG1 forward primer

<400> 31

cacactccac actttttgt ttac 24

<210> 32

<211> 77

<212> DNA

<213> Mus musculus

<400> 32

tgtcccctcc accccacagt ggggccacta ggagtattgg tgacagaaaa gccccatcct 60

taggcctcct ccttcct 77

<210> 33

<211> 26

<212> DNA

<213> Artificial Sequence

<220><223> HBG1 reverse primer

<400> 33

aagtgtcttt actgctttta tttgct 26

<210> 34

<211> 77

<212> DNA

<213> Mus musculus

<400> 34

tgtcccctcc accccacagt ggggccacta ggaggattgg tgacagaaca gccccatcct 60

taggcctcct ccttcct 77

<210> 35

<211> 77

<212> DNA

<213> Mus musculus

<400> 35

tgtcccctcc accccacagt ggggccacta gcagtattgg tgacagaaaa gccccatcct 60

taggcctcct ccttcct 77

<210> 36

<211> 77

<212> DNA

<213> Mus musculus

<400> 36

tgtcccctcc accccacagt ggggccacta gcaggattgg tgacagaaca gccccatcct 60

taggcctcct ccttcct 77

<210> 37

<211> 78

<212> DNA

<213> Mus musculus

<400> 37

tgtcccctcc accccacagt ggggccacta gggacgattg gtgacagaaa agccccatcc 60

ttaggcctcc tccttcct 78

<210> 38

<211> 78

<212> DNA

<213> Mus musculus

<400> 38

tgtcccctcc accccacagt ggggccacta gggaggattg gtgacagaaa agccccatcc 60

ttaggcctcc tccttcct 78

<210> 39

<211> 78

<212> DNA

<213> Mus musculus

<400> 39

tgtcccctcc accccacagt ggggccacta ggcaggattg gtgacagaaa agccccatcc 60

ttaggcctcc tccttcct 78

<210> 40

<211> 78

<212> DNA

<213> Mus musculus

<400> 40

tgtcccctcc accccacagt ggggccacta gggatgattg gtgacagaaa agccccatcc 60

ttaggcctcc tccttcct 78

<210> 41

<211> 78

<212> DNA

<213> Mus musculus

<400> 41

tgtcccctcc accccacagt ggggccacta gggacgcttg gtgacagaaa agccccatcc 60

ttaggcctcc tccttcct 78

<210> 42

<211> 78

<212> DNA

<213> Mus musculus

<400

> 42

tgtcccctcc accccacagt ggggccacta gggaggcttg gtgacagaaa agccccatcc 60

ttaggcctcc tccttcct 78

<210> 43

<211> 78

<212> DNA

<213> Mus musculus

<400> 43

tgtcccctcc accccacagt ggggccacta gggacgattg gtgacagaca agccccatcc 60

ttaggcctcc tccttcct 78

<210> 44

<211> 78

<212> DNA

<213> Mus musculus

<400> 44

tgtcccctcc accccacagt ggggccacta gggaggattg gtgacagaca agccccatcc 60

ttaggcctcc tccttcct 78

<210> 45

<211> 78

<212> DNA

<213> Mus musculus

<400> 45

tgtcccctcc accccacagt ggggccacta gggacgattg gtgccagaaa agccccatcc 60

ttaggcctcc tccttcct 78

<210> 46

<211> 78

<212> DNA

<213> Mus musculus

<400> 46

tgtcccctcc accccacagt ggggccacta gggaggattg gtgccagaaa agccccatcc 60

ttaggcctcc tccttcct 78

<210> 47

<211> 79

<212> DNA

<213> Mus musculus

<400> 47

tgtcccctcc accccacagt ggggccacta gggacagatt ggtgacagaa aagccccatc 60

cttaggcctc ctccttcct 79

<210> 48

<211> 141

<212> DNA

<213> Artificial Sequence

<220><223> BCL11A binding site at HBG promoter

<400> 48

gggggcccct tccccacact atctcaatgc aaatatctgt ctgaaacggt ccctggctaa 60

actccaccat gggttggcca gscttgcctt gaccaatagc cttgacaagg caaacttgac 120

caatagtctt agagtatcca g 141

<210> 49

<211> 599

<212> DNA

<213> Adenovirus serotype 5

<400> 49

catcatcaat aatatacctt attttggatt gaagccaata tgataatgag ggggtggagt 60

ttgtgacgtg gcgcggggcg tgggaacggg gcgggtgacg tagtagtgtg gcggaagtgt 120

gatgttgcaa gtgtggcgga acacatgtaa gcgacggatg tggcaaaagt gacgtttttg 180

gtgtgcgccg gtgtacacag gaagtgacaa ttttcgcgcg gttttaggcg gatgttgtag 240

taaatttggg cgtaaccgag taagatttgg ccattttcgc gggaaaactg aataagagga 300

agtgaaatct gaataatttt gtgttactca tagcgcgtaa tatttgtcta gggccgcggg 360

actttgaccg tttacgtgga gactcgccca ggtgtttttc tcaggtgttt tccgcgttcc 420

gggtcaaagt tggcgtttta ttattatagt cagctgacgt gtagtgtatt tatacccggt 480

gagttcctca agaggccact cttgagtgcc agcgagtaga gttttctcct ccgagccgct 540

ccgacaccgg gactgaaaat gagacatatt atctgccacg gaggtgttat taccgaaga 599

<210> 50

<211> 600

<212> DNA

<213> Adenovirus serotype 35

<400> 50

catcatcaat aatatacctt atagatggaa tggtgccaat atgtaaatga ggtgatttta 60

aaaagtgtgg gccgtgtggt gattggctgt ggggttaacg gttaaaaggg gcggcgcggc 120

cgtgggaaaa tgacgtttta tgggggtgga gtttttttgc aagttgtcgc gggaaatgtt 180

acgcataaaa aggcttcttt tctcacggaa ctacttagtt ttcccacggt atttaacagg 240

aaatgaggta gttttgaccg gatgcaagtg aaaattgctg attttcgcgc gaaaactgaa 300

tgaggaagtg tttttctgaa taatgtggta tttatggcag ggtggagtat ttgttcaggg 360

ccaggtagac tttgacccat tacgtggagg tttcgattac cgtgtttttt acctgaattt 420

ccgcgtaccg tgtcaaagtc ttctgttttt acgtaggtgt cagctgatcg ctagggtatt 480

tatacctcag ggtttgtgtc aagaggccac tcttgagtgc cagcgagaag agttttctcc 540

tctgcgccgg cagtttaata ataaaaaaat gagagatttg cgatttctgc ctcaggaaat 600

<210> 51

<211> 32767

<212> DNA

<213> Artificial Sequence

<220><223> Ad35E4PS2/WL-ps2

<220><221> misc_feature

<222> (2582)..(2759)

<223> Ad35 1-->178

<220><221> misc_feature

<222> (2768)..(2801)

<223> loxP

<220><221> misc_feature

<222> (2808)..(2923)

<223> Ad35 366-->481

<220><221> misc_feature

<222> (2924)..(2957)

<223> loxP

<220><221> misc_feature

<222> (2966)..(27288)

<223> Ad35

3112-->27435

<220><221> misc_feature

<222> (27343)..(29812)

<223> lambda-1 (Complementary)

<220><221> misc_feature

<222> (30126)..(30340)

<223> BGH polyA

<220><221> misc_feature

<222> (30365)..(31030)

<223> copGFP (Complementary)

<220><221> misc_feature

<222> (31077)..(31729)

<223> CMV (Complementary)

<220><221> misc_feature

<222> (31781)..(33310)

<223> lambda-2

<220><221> misc_feature

<222> (33371)..(34706)

<223> Ad35 30544-->31879

<220><221> misc_feature

<222> (34702)..(35816)

<223> Ad5E4orf6

<220><221> misc_feature

<222> (35814)..(37636)

<223> Ad35 32972-->34794

<400> 51

taaacttggc gcgccctgag tgatttttct ctggtcccgc cgcatccata ccgccagttg 60

tttaccctca caacgttcca gtaaccgggc atgttcatca tcagtaaccc gtatcgtgag 120

catcctctct cgtttcatcg gtatcattac ccccatgaac agaaatcccc cttacacgga 180

ggcatcagtg accaaacagg aaaaaaccgc ccttaacatg gcccgcttta tcagaagcca 240

gacattaacg cttctggaga aactcaacga gctggacgcg gatgaacagg cagacatctg 300

tgaatcgctt cacgaccacg ctgatgagct ttaccgcagc tgcctcgcgc gtttcggtga 360

tgacggtgaa aacctctgac acatgcagct cccggagacg gtcacagctt gtctgtaagc 420

ggatgccggg agcagacaag cccgtcaggg cgcgtcagcg ggtgttggcg ggtgtcgggg 480

cgcagccatg acccagtcac gtagcgatag cggagtgtat actggcttaa ctatgcggca 540

tcagagcaga ttgtactgag agtgcaccat atgcggtgtg aaataccgca cagatgcgta 600

aggagaaaat accgcatcag gcgctcttcc gcttcctcgc tcactgactc gctgcgctcg 660

gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg cggtaatacg gttatccaca 720

gaatcagggg ataacgcagg aaagaacatg tgagcaaaag gccagcaaaa ggccaggaac 780

cgtaaaaagg ccgcgttgct ggcgtttttc cataggctcc gccccccctga cgagcatcac 840

aaaaatcgac gctcaagtca gaggtggcga aacccgacag gactataaag ataccaggcg 900

tttccccctg gaagctccct cgtgcgctct cctgttccga ccctgccgct taccggatac 960

ctgtccgcct ttctcccttc gggaagcgtg gcgctttctc atagctcacg ctgtaggtat 1020

ctcagttcgg tgtaggtcgt tcgctccaag ctgggctgtg tgcacgaacc ccccgttcag 1080

cccgaccgct gcgccttatc cggtaactat cgtcttgagt ccaacccggt aagacacgac 1140

ttatcgccac tggcagcagc cactggtaac aggattagca gagcgaggta tgtaggcggt 1200

gctacagagt tcttgaagtg gtggcctaac tacggctaca ctagaaggac agtatttggt 1260

atctgcgctc tgctgaagcc agttaccttc ggaaaaagag ttggtagctc ttgatccggc 1320

aaacaaacca ccgctggtag cggtggtttt tttgtttgca agcagcagat tacgcgcaga 1380

aaaaaaggat ctcaagaaga tcctttgatc ttttctacgg ggtctgacgc tcagtggaac 1440

gaaaactcac gttaagggat tttggtcatg agattatcaa aaaggatctt cacctagatc 1500

cttttaaatt aaaaatgaag ttttaaatca atctaaagta tatatgagta aacttggtct 1560

gacagttacc aatgcttaat cagtgaggca cctatctcag cgatctgtct atttcgttca 1620

tccatagttg cctgactccc cgtcgtgtag ataactacga tacgggaggg cttaccatct 1680

ggccccagtg ctgcaatgat accgcgagac ccacgctcac cggctccaga tttatcagca 1740

ataaaccagc cagccggaag ggccgagcgc agaagtggtc ctgcaacttt atccgcctcc 1800

atccagtcta ttaattgttg ccgggaagct agagtaagta gttcgccagt taatagtttg 1860

cgcaacgttg ttgccattgc tgcaggcatc gtggtgtcac gctcgtcgtt tggtatggct 1920

tcattcagct ccggttccca acgatcaagg cgagttacat gatcccccat gttgtgcaaa 1980

aaagcggtta gctccttcgg tcctccgatc gttgtcagaa gtaagttggc cgcagtgtta 2040

tcactcatgg ttatggcagc actgcataat tctcttactg tcatgccatc cgtaagatgc 2100

ttttctgtga ctggtgagta ctcaaccaag tcattctgag aatagtgtat gcggcgaccg 2160

agttgctctt gcccggcgtc aacacgggat aataccgcgc cacatagcag aactttaaaa 2220

gtgctcatca ttggaaaacg ttcttcgggg cgaaaactct caaggatctt accgctgttg 2280

agatccagtt cgatgtaacc cactcgtgca cccaactgat cttcagcatc ttttactttc 2340

accagcgttt ctgggtgagc aaaaacagga aggcaaaatg ccgcaaaaaa gggaataagg 2400

gcgacacgga aatgttgaat actcatactc ttccttttc aatattattg aagcatttat 2460

cagggttatt gtctcatgag cggatacata tttgaatgta tttagaaaaa taaacaaata 2520

ggggttccgc gcacatttcc ccgaaaagtg ccacctgtct agctacgata tcctgtttaa 2580

acatcatcaa taatatacct tatagatgga atggtgccaa tatgtaaatg aggtgatttt 2640

aaaaagtgtg ggccgtgtgg tgattggctg tggggttaac ggttaaaagg ggcggcgcgg 2700

ccgtgggaaa atgacgtttt atgggggtgg agttttttg caagttgtcg cgggaaatga 2760

tttaaatata acttcgtata gcatacatta tacgaagtta tggatcctag actttgaccc 2820

attacgtgga ggtttcgatt accgtgtttt ttacctgaat ttccgcgtac cgtgtcaaag 2880

tcttctgttt ttacgtaggt gtcagctgat cgctagggta tttataactt cgtatagcat 2940

acattatacg aagttatatt taaataggaa tgtttatgcc ttaccagtgt aacatgaatc 3000

atgtgaaagt gttgttggaa ccagatgcct tttccagaat gagcctaaca ggaatctttg 3060

acatgaacac gcaaatctgg aagatcctga ggtatgatga tacgagatcg agggtgcgcg 3120

catgcgaatg cggaggcaag catgccaggt tccagccggt gtgtgtagat gtgaccgaag 3180

atctcagacc ggatcatttg gttattgccc gcactggagc agagttcgga tccagtggag 3240

aagaaactga ctaaggtgag tattgggaaa actttggggt gggattttca gatggacaga 3300

ttgagtaaaa atttgttttt tctgtcttgc agctgacatg agtggaaatg cttcttttaa 3360

ggggggagtc ttcagccctt atctgacagg gcgtctccca tcctgggcag gagttcgtca 3420

gaatgttatg ggatctactg tggatggaag acccgttcaa cccgccaatt cttcaacgct 3480

gacctatgct actttaagtt cttcaccttt ggacgcagct gcagccgctg ccgccgcctc 3540

tgtcgccgct aacactgtgc ttggaatggg ttactatgga agcatcgtgg ctaattccac 3600

ttcctctaat aacccttcta cactgactca ggacaagtta cttgtccttt tggcccagct 3660

ggaggctttg acccaacgtc tgggtgaact ttctcagcag gtggccgagt tgcgagtaca 3720

aactgagtct gctgtcggca cggcaaagtc taaataaaaa aaattccaga atcaatgaat 3780

aaataaacga gcttgttgtt gatttaaaat caagtgtttt tatttcattt ttcgcgcacg 3840

gtatgccctg gaccaccgat ctcgatcatt gagaactcgg tggatttttt ccagaatcct 3900

atagaggtgg gattgaatgt ttagatacat gggcattagg ccgtctttgg ggtggagata 3960

gctccattga agggattcat gctccggggt agtgttgtaa atcacccagt cataacaagg 4020

tcgcagtgca tggtgttgca caatatcttt tagaagtagg ctgattgcca cagataagcc 4080

cttggtgtag gtgtttacaa accggttgag ctgggagggg tgcattcgag gtgaaattat 4140

gtgcattttg gattggattt ttaagttggc aatattgccg ccaagatccc gtcttgggtt 4200

catgttatga aggactacca agacggtgta tccggtacat ttaggaaatt tatcgtgcag 4260

cttggatgga aaagcgtgga aaaatttgga gacacccttg tgtcctccga gattttccat 4320

gcactcatcc atgataatag caatggggcc gtgggcagcg gcgcgggcaa acacgttccg 4380

tgggtctgac acatcatagt tatgttcctg agttaaatca tcataagcca ttttaatgaa 4440

tttggggcgg agcgtaccag attggggtat gaatgttcct tcgggccccg gagcatagtt 4500

cccctcacag atttgcattt cccaagcttt cagttctgag ggtggaatca tgtccacctg 4560

gggggctatg aagaacaccg tttcgggggc gggggtgatt agttgggatg atagcaagtt 4620

tctgagcaat tgagatttgc cacatccggt ggggccataa ataattccga ttacaggttg 4680

caggtggtag tttagggaac ggcaactgcc gtcttctcga agcaaggggg ccacctcgtt 4740

catcatttcc cttacatgca tattttcccg caccaaatcc attaggaggc gctctcctcc 4800

tagtgataga agttcttgta gtgaggaaaa gtttttcagc ggttttagac cgtcagccat 4860

gggcattttg gaaagagttt gctgcaaaag ttctagtctg ttccacagtt cagtgatgtg 4920

ttctatggca tctcgatcca gcagacctcc tcgtttcgcg ggtttggacg gctcctggag 4980
tagggtatga gacgatgggc gtccagcgct gccagggttc ggtccttcca gggtctcagt 5040
gttcgagtca gggttgtttc cgtcacagtg aaggggtgtg cgcctgcttg ggcgcttgcc 5100
agggtgcgct tcagactcat tctgctggtg gagaacttct gtcgcttggc gccctgtatg 5160
tcggccaagt agcagtttac catgagttcg tagttgagcg cctcggctgc gtggcctttg 5220
gcgcggagct tacctttgga agttttcttg cataccgggc agtataggca tttcagcgca 5280
tacagcttgg gcgcaaggaa aatggattct ggggagtatg catccgcgcc gcaggaggcg 5340

caaacagttt cacattccac cagccaggtt aaatccggtt cattggggtc aaaaacaagt 5400
tttccgccat atttttgat gcgtttctta cctttggtct ccataagttc gtgtcctcgt 5460
tgagtgacaa acaggctgtc cgtatctccg tagactgatt ttacaggcct cttctccagt 5520
ggagtgcctc ggtcttcttc gtacaggaac tctgaccact ctgatacaaa ggcgcgcgtc 5580
caggccagca caaaggaggc tatgtgggag gggtagcgat cgttgtcaac cagggggtcc 5640
acctttcca aagtatgcaa acacatgtca ccctcttcaa catccaggaa tgtgattggc 5700
ttgtaggtgt atttcacgtg acctggggtc cccgctgggg gggtataaaa gggggcggtt 5760

ctttgctctt cctcactgtc ttccggatcg ctgtccagga acgtcagctg ttggggtagg 5820
tattccctct cgaaggcggg catgacctct gcactcaggt tgtcagtttc taagaacgag 5880
gaggatttga tattgacagt gccggttgag atgcctttca tgaggttttc gtccatttgg 5940
tcagaaaaca caatttttt attgtcaagt ttggtggcaa atgatccata cagggcgttg 6000
gataaaagtt tggcaatgga tcgcatggtt tggttctttt ccttgtccgc gcgctctttg 6060
gcggcgatgt tgagttggac atactcgcgt gccaggcact tccattcggg gaagatagtt 6120
gttaattcat ctggcacgat tctcacttgc caccctcgat tatgcaaggt aattaaatcc 6180

acactggtgg ccacctcgcc tcgaaggggt tcattggtcc aacagagcct acctcctttc 6240
ctagaacaga aaggggaag tgggtctagc ataagttcat cgggagggtc tgcatccatg 6300
gtaaagattc ccggaagtaa atccttatca aaatagctga tgggagtggg gtcatctaag 6360
gccatttgcc attctcgagc tgccagtgcg cgctcatatg ggttaagggg actgccccag 6420
ggcatgggat gggtgagagc agaggcatac atgccacaga tgtcatagac gtagatggga 6480
tcctcaaaga tgcctatgta ggttggatag catcgccccc ctctgatact tgctcgcaca 6540
tagtcatata gttcatgtga tggcgctagc agccccggac ccaagttggt gcgattgggt 6600

ttttctgttc tgtagacgat ctggcgaaag atggcgtgag aattggaaga gatggtgggt 6660

ctttgaaaaa tgttgaaatg ggcatgaggt agacctacag agtctctgac aaagtgggca 6720

taagattctt gaagcttggt taccagttcg gcggtgacaa gtacgtctag ggcgcagtag 6780

tcaagtgttt cttgaatgat gtcataacct ggttggtttt tcttttccca cagttcgcgg 6840

ttgagaaggt attcttcgcg atccttccag tactcttcta gcggaaaccc gtctttgtct 6900

gcacggtaag atcctagcat gtagaactga ttaactgcct tgtaagggca gcagcccttc 6960

tctacgggta gagagtatgc ttgagcagct tttcgtagcg aagcgtgagt aagggcaaag 7020

gtgtctctga ccatgacttt gagaaattgg tatttgaagt ccatgtcgtc acaggctccc 7080

tgttcccaga gttggaagtc tacccgtttc ttgtaggcgg ggttgggcaa agcgaaagta 7140

acatcattga agagaatctt accggctctg ggcataaaat tgcgagtgat gcggaaaggc 7200

tgtggtactt ccgctcgatt gttgatcacc tgggcagcta ggacgatttc gtcgaaaccg 7260

ttgatgttgt gtcctacgat gtataattct atgaaacgcg gcgtgcctct gacgtgaggt 7320

agcttactga gctcatcaaa ggttaggtct gtggggtcag ataaggcgta gtgttcgaga 7380

gcccattcgt gcaggtgagg atttgcatgt aggaatgatg accaaagatc taccgccagt 7440

gctgtttgta actggtcccg atactgacga aaatgccggc caattgccat tttttctgga 7500

gtgacacagt agaaggttct ggggtcttgt tgccatcgat cccacttgag tttaatggct 7560

agatcgtggg ccatgttgac gagacgctct tctcctgaga gtttcatgac cagcatgaaa 7620

ggaactagtt gtttgccaaa ggatcccatc caggtgtaag tttccacatc gtaggtcagg 7680

aagagtcttt ctgtgcgagg atgagagccg atcgggaaga actggatttc ctgccaccag 7740

ttggaggatt ggctgttgat gtgatggaag tagaagtttc tgcggcgcgc cgagcattcg 7800

tgtttgtgct tgtacagacg gccgcagtag tcgcagcgtt gcacgggttg tatctcgtga 7860

atgagctgta cctggcttcc cttgacgaga aatttcagtg ggaagccgag gcctggcgat 7920

tgtatctcgt gctcttctat attcgctgta tcggcctgtt catcttctgt ttcgatggtg 7980

gtcatgctga cgagcccccg cgggaggcaa gtccagacct cggcgcggga ggggcggagc 8040

tgaaggacga gagcgcgcag gctggagctg tccagagtcc tgagacgctg cggactcagg 8100

ttagtaggta gggacagaag attaacttgc atgatctttt ccagggcgtg cgggaggttc 8160

agatggtact tgatttccac aggttcgttt gtagagacgt caatggcttg cagggttccg 8220

tgtcctttgg gcgccactac cgtacctttg ttttttcttt tgatcggtgg tggctctctt 8280

gcttcttgca tgctcagaag cggtgacggg gacgcgcgcc gggcggcagc ggttgttccg 8340

gacccggggg catggctggt agtggcacgt cggcgccgcg cacgggcagg ttctggtatt 8400

gcgctctgag aagacttgcg tgcgccacca cgcgtcgatt gacgtcttgt atctgacgtc 8460

tctgggtgaa agctaccggc cccgtgagct tgaacctgaa agagagttca acagaatcaa 8520

tttcggtatc gttaacggca gcttgtctca gtatttcttg tacgtcacca gagttgtcct 8580

ggtaggcgat ctccgccatg aactgctcga tttcttcctc ctgaagatct ccgcgacccg 8640

ctctttcgac ggtggccgcg aggtcattgg agatacggcc catgagttgg gagaatgcat 8700

tcatgcccgc ctcgttccag acgcggctgt aaaccacggc cccctcggag tctcttgcgc 8760

gcatcaccac ctgagcgagg ttaagctcca cgtgtctggt taagaccgca tagttgcata 8820

ggcgctgaaa aaggtagttg agtgtggtgg caatgtgttc ggcgacgaag aaatacatga 8880

tccatcgtct cagcggcatt tcgctaacat cgcccagagc ttccaagcgc tccatggcct 8940

cgtagaagtc cacggcaaaa ttaaaaaact gggagtttcg cgcggacacg gtcaattcct 9000

cctcgagaag acggatgagt tcggctatgg tggcccgtac ttcgcgttcg aaggctcccg 9060

ggatctcttc ttcctcttct atctcttctt ccactaacat ctcttcttcg tcttcaggcg 9120

ggggcggagg gggcacgcgg cgacgtcgac ggcgcacggg caaacggtcg atgaatcgtt 9180

caatgacctc tccgcggcgg cggcgcatgg tttcagtgac ggcgcggccg ttctcgcgcg 9240

gtcgcagagt aaaaacaccg ccgcgcatct ccttaaagtg gtgactggga ggttctccgt 9300

ttgggaggga gagggcgctg attatacatt ttattaattg gcccgtaggg actgcgcgca 9360

gagatctgat cgtgtcaaga tccacgggat ctgaaaacct ttcgacgaaa gcgtctaacc 9420

agtcacagtc acaaggtagg ctgagtacgg cttcttgtgg gcgggggtgg ttatgtgttc 9480

ggtctgggtc ttctgtttct tcttcatctc gggaaggtga gacgatgctg ctggtgatga 9540

aattaaagta ggcagttcta agacggcgga tggtggcgag gagcaccagg tctttgggtc 9600

cggcttgctg gatacgcagg cgattggcca ttccccaagc attatcctga catctagcaa 9660

gatctttgta gtagtcttgc atgagccgtt ctacgggcac ttcttcctca cccgttctgc 9720

catgcatacg tgtgagtcca aatccgcgca ttggttgtac cagtgccaag tcagctacga 9780

ctctttcggc gaggatggct tgctgtactt gggtaagggt ggcttgaaag tcatcaaaat 9840

ccacaaagcg gtggtaagcc cctgtattaa tggtgtaagc acagttggcc atgactgacc 9900

agttaactgt ctggtgacca gggcgcacga gctcggtgta tttaaggcgc gaataggcgc 9960

gggtgtcaaa gatgtaatcg ttgcaggtgc gcaccagata ctggtaccct ataagaaaat 10020

gcggcggtgg ttggcggtag agaggccatc gttctgtagc tggagcgcca ggggcgaggt 10080

cttccaacat aaggcggtga tagccgtaga tgtacctgga catccaggtg attcctgcgg 10140

cggtagtaga agcccgagga aactcgcgta cgcggttcca aatgttgcgt agcggcatga 10200

agtagttcat tgtaggcacg gtttgaccag tgaggcgcgc gcagtcattg atgctctata 10260

gacacggaga aaatgaaagc gttcagcgac tcgactccgt agcctggagg aacgtgaacg 10320

ggttgggtcg cggtgtaccc cggttcgaga cttgtactcg agccggccgg agccgcggct 10380

aacgtggtat tggcactccc gtctcgaccc agcctacaaa aatccaggat acggaatcga 10440
gtcgttttgc tggtttccga atggcaggga agtgagtcct attttttttt ttttttgccg 10500
ctcagaatgc atcccgtgct gcgacagatg cgccccccaac aacagccccc ctcgcagcag 10560
cagcagcagc aaccacaaaa ggctgtccct gcaactactg caactgccgc cgtgagcggt 10620
gcgggacagc ccgcctatga tctggacttg gaagagggcg aaggactggc acgtctaggt 10680
gcgccttcgc ccgagcggca tccgcgagtt caactgaaaa aagattctcg cgaggcgtat 10740
gtgccccaac agaacctatt tagagacaga agcggcgagg agccggagga gatgcgagct 10800

tcccgcttta acgcgggtcg tgagctgcgt cacggtttgg accgaagacg agtgttgcga 10860
gacgaggatt tcgaagttga tgaagtgaca gggatcagtc ctgccagggc acacgtggct 10920
gcagccaacc ttgtatcggc ttacgagcag acagtaaagg aagagcgtaa cttccaaaag 10980
tcttttaata atcatgtgcg aaccctgatt gcccgcgaag aagttaccct tggtttgatg 11040
catttgtggg atttgatgga agctatcatt cagaacccta ctagcaaacc tctgaccgcc 11100
cagctgtttc tggtggtgca acacagcaga gacaatgagg ctttcagaga ggcgctgctg 11160
aacatcaccg aacccgaggg gagatggttg tatgatctta tcaacattct acagagtatc 11220

atagtgcagg agcggagcct gggcctggcc gagaaggtag ctgccatcaa ttactcggtt 11280
ttgagcttgg gaaaatatta cgctcgcaaa atctacaaga ctccatacgt tcccatagac 11340
aaggaggtga agatagatgg gttctacatg cgcatgacgc tcaaggtctt gaccctgagc 11400
gatgatcttg gggtgtatcg caatgacaga atgcatcgcg cggttagcgc cagcaggagg 11460
cgcgagttaa gcgacaggga actgatgcac agtttgcaaa gagctctgac tggagctgga 11520
accgagggtg agaattactt cgacatggga gctgacttgc agtggcagcc taatcgcagg 11580
gctctgagcg ccgcgacggc aggatgtgag cttccttaca tagaagaggc ggatgaaggc 11640

gaggaggaag agggcgagta cttggaagac tgatggcaca acccgtgttt tttgctagat 11700
ggaacagcaa gcaccggatc ccgcaatgcg ggcggcgctg cagagccagc cgtccggcat 11760
taactcctcg gacgattgga cccaggccat gcaacgtatc atggcgttga cgactcgcaa 11820
ccccgaagcc tttagacagc aaccccaggc caaccgtcta tcggccatca tggaagctgt 11880
agtgccttcc cgatctaatc ccactcatga gaaggtcctg gccatcgtga acgcgttggt 11940
ggagaacaaa gctattcgtc cagatgaggc cggactggta tacaacgctc tcttagaacg 12000
cgtggctcgc tacaacagta gcaatgtgca aaccaatttg gaccgtatga taacagatgt 12060

acgcgaagcc gtgtctcagc gcgaaaggtt ccagcgtgat gccaacctgg gttcgctggt 12120

ggcgttaaat gctttcttga gtactcagcc tgctaatgtg ccgcgtggtc aacaggatta 12180

tactaacttt ttaagtgctt tgagactgat ggtatcagaa gtacctcaga gcgaagtgta 12240

tcagtccggt cctgattact tctttcagac tagcagacag ggcttgcaga cggtaaatct 12300

gagccaagct tttaaaaacc ttaaaggttt gtggggagtg catgccccgg taggagaaag 12360

agcaaccgtg tctagcttgt taactccgaa ctcccgcctg ttattactgt tggtagctcc 12420

tttcaccgac agcggtagca tcgaccgtaa ttcctatttg ggttacctac taaacctgta 12480

tcgcgaagcc atagggcaaa gtcaggtgga cgagcagacc tatcaagaaa ttacccaagt 12540

cagtcgcgct ttgggacagg aagacactgg cagtttggaa gccactctga acttcttgct 12600

taccaatcgg tctcaaaaga tccctcctca atatgctctt actgcggagg aggagaggat 12660

ccttagatat gtgcagcaga gcgtgggatt gtttctgatg caagaggggg caactccgac 12720

tgcagcactg gacatgacag cgcgaaatat ggagcccagc atgtatgcca gtaaccgacc 12780

tttcattaac aaactgctgg actacttgca cagagctgcc gctatgaact ctgattattt 12840

caccaatgcc atcttaaacc cgcactggct gcccccacct ggtttctaca cgggcgaata 12900

tgacatgccc gaccctaatg acggatttct gtgggacgac gtggacagcg atgtttttc 12960

acctctttct gatcatcgca cgtggaaaaa ggaaggcggt gatagaatgc attcttctgc 13020

atcgctgtcc ggggtcatgg gtgctaccgc ggctgagccc gagtctgcaa gtccttttcc 13080

tagtctaccc ttttctctac acagtgtacg tagcagcgaa gtgggtagaa taagtcgccc 13140

gagtttaatg ggcgaagagg agtacctaaa cgattccttg ctcagaccgg caagagaaaa 13200

aaatttccca aacaatggaa tagaaagttt ggtggataaa atgagtagat ggaagactta 13260

tgctcaggat cacagagacg agcctgggat catggggact acaagtagag cgagccgtag 13320

acgccagcgc catgacagac agaggggtct tgtgtgggac gatgaggatt cggccgatga 13380

tagcagcgtg ttggacttgg gtgggagagg aaggggcaac ccgtttgctc atttgcgccc 13440

tcgcttgggt ggtatgttgt gaaaaaaaat aaaaaagaaa aactcaccaa ggccatggcg 13500

acgagcgtac gttcgttctt ctttattatc tgtgtctagt ataatgaggc gagtcgtgct 13560

aggcggagcg gtggtgtatc cggagggtcc tcctccttcg tacgagagcg tgatgcagca 13620

gcagcaggcg acggcggtga tgcaatcccc actggaggct ccctttgtgc ctccgcgata 13680

cctggcacct acggagggca gaaacagcat tcgttactcg gaactggcac ctcagtacga 13740

taccaccagg ttgtatctgg tggacaacaa gtcggcggac attgcttctc tgaactatca 13800

gaatgaccac agcaacttct tgaccacggt ggtgcagaac aatgacttta cccctacgga 13860

agccagcacc cagaccatta actttgatga acgatcgcgg tggggcggtc agctaaagac 13920

catcatgcat actaacatgc caaacgtgaa cgagtatatg tttagtaaca agttcaaagc 13980

gcgtgtgatg gtgtccagaa aacctcccga cggtgctgca gttggggata cttatgatca 14040

caagcaggat attttggaat atgagtggtt cgagtttact ttgccagaag gcaacttttc 14100

agttactatg actattgatt tgatgaacaa tgccatcata gataattact tgaaagtggg 14160

tagacagaat ggagtgcttg aaagtgacat tggtgttaag ttcgacacca ggaacttcaa 14220

gctgggatgg gatcccgaaa ccaagttgat catgcctgga gtgtatacgt atgaagcctt 14280

ccatcctgac attgtcttac tgcctggctg cggagtggat tttaccgaga gtcgtttgag 14340

caaccttctt ggtatcagaa aaaaacagcc atttcaagag ggttttaaga ttttgtatga 14400

agatttagaa ggtggtaata ttccggccct cttggatgta gatgcctatg agaacagtaa 14460

gaaagaacaa aaagccaaaa tagaagctgc tacagctgct gcagaagcta aggcaaacat 14520

agttgccagc gactctacaa gggttgctaa cgctggagag gtcagaggag acaattttgc 14580

gccaacacct gttccgactg cagaatcatt attggccgat gtgtctgaag gaacggacgt 14640

gaaactcact attcaacctg tagaaaaaga tagtaagaat agaagctata atgtgttgga 14700

agacaaaatc aacacagcct atcgcagttg gtatctttcg tacaattatg gcgatcccga 14760

aaaaggagtg cgttcctgga cattgctcac cacctcagat gtcacctgcg gagcagagca 14820

ggtttactgg tcgcttccag acatgatgaa ggatcctgtc actttccgct ccactagaca 14880

agtcagtaac taccctgtgg tgggtgcaga gcttatgccc gtcttctcaa agagcttcta 14940

caacgaacaa gctgtgtact cccagcagct ccgccagtcc acctcgctta cgcacgtctt 15000

caaccgcttt cctgagaacc agattttaat ccgtccgccg gcgcccacca ttaccaccgt 15060

cagtgaaaac gttcctgctc tcacagatca cgggaccctg ccgttgcgca gcagtatccg 15120

gggagtccaa cgtgtgaccg ttactgacgc cagacgccgc acctgtccct acgtgtacaa 15180

ggcactgggc atagtcgcac cgcgcgtcct ttcaagccgc actttctaaa aaaaaaatgt 15240

ccattcttat ctcgcccagt aataacaccg gttggggtct gcgcgctcca agcaagatgt 15300

acggaggcgc acgcaaacgt tctacccaac atcccgtgcg tgttcgcgga cattttcgcg 15360

ctccatgggg tgccctcaag ggccgcactc gcgttcgaac caccgtcgat gatgtaatcg 15420

atcaggtggt tgccgacgcc cgtaattata ctcctactgc gcctacatct actgtggatg 15480

cagttattga cagtgtagtg gctgacgctc gcaactatgc tcgacgtaag agccggcgaa 15540

ggcgcattgc cagacgccac cgagctacca ctgccatgcg agccgcaaga gctctgctac 15600

gaagagctag acgcgtgggg cgaagagcca tgcttagggc ggccagacgt gcagcttcgg 15660

gcgccagcgc cggcaggtcc cgcaggcaag cagccgctgt cgcagcggcg actattgccg 15720

acatggccca atcgcgaaga ggcaatgtat actgggtgcg tgacgctgcc accggtcaac 15780

gtgtacccgt gcgcacccgt ccccctcgca cttagaagat actgagcagt ctccgatgtt 15840

gtgtcccagc ggcgaggatg tccaagcgca aatacaagga agaaatgctg caggttatcg 15900
cacctgaagt ctacggccaa ccgttgaagg atgaaaaaaa accccgcaaa atcaagcggg 15960
ttaaaaagga caaaaaagaa gaggaagatg gcgatgatgg gctggcggag tttgtgcgcg 16020
agtttgcccc acggcgacgc gtgcaatggc gtgggcgcaa agttcgacat gtgttgagac 16080
ctggaacttc ggtggtcttt acacccggcg agcgttcaag cgctactttt aagcgttcct 16140
atgatgaggt gtacggggat gatgatattc ttgagcaggc ggctgaccga ttaggcgagt 16200
ttgcttatgg caagcgtagt agaataactt ccaaggatga gacagtgtca ataccttgg 16260

atcatggaaa tcccacccct agtcttaaac cggtcacttt gcagcaagtg ttacccgtaa 16320
ctccgcgaac aggtgttaaa cgcgaaggtg aagatttgta tcccactatg caactgatgg 16380
tacccaaacg ccagaagttg gaggacgttt tggagaaagt aaaagtggat ccagatattc 16440
aacctgaggt taaagtgaga cccattaagc aggtagcgcc tggtctgggg gtacaaactg 16500
tagacattaa gattcccact gaaagtatgg aagtgcaaac tgaacccgca aagcctactg 16560
ccacctccac tgaagtgcaa acggatccat ggatgcccat gcctattaca actgacgccg 16620
ccggtcccac tcgaagatcc cgacgaaagt acggtccagc aagtctgttg atgcccaatt 16680

atgttgtaca cccatctatt attcctactc ctggttaccg aggcactcgc tactatcgca 16740
gccgaaacag tacctcccgc cgtcgccgca agacacctgc aaatcgcagt cgtcgccgta 16800
gacgcacaag caaaccgact cccggcgccc tggtgcggca agtgtaccgc aatggtagtg 16860
cggaaccttt gacactgccg cgtgcgcgtt accatccgag tatcatcact taatcaatgt 16920
tgccgctgcc tccttgcaga tatggccctc acttgtcgcc ttcgcgttcc catcactggt 16980
taccgaggaa gaaactcgcg ccgtagaaga gggatgttgg gacgcggaat gcgacgctac 17040
aggcgacggc gtgctatccg caagcaattg cggggtggtt ttttaccagc cttaattcca 17100

attatcgctg ctgcaattgg cgcgatacca ggcatagctt ccgtggcggt tcaggcctcg 17160
caacgacatt gacattggaa aaaaaacgta taaataaaaa aaaatacaat ggactctgac 17220
actcctggtc ctgtgactat gttttcttag agatggaaga catcaatttt tcatccttgg 17280
ctccgcgaca cggcacgaag ccgtacatgg gcacctggag cgacatcggc acgagccaac 17340
tgaacggggg cgccttcaat tggagcagta tctggagcgg gcttaaaaat tttggctcaa 17400
ccataaaaac atacgggaac aaagcttgga acagcagtac aggacaggcg cttagaaata 17460
aacttaaaga ccagaacttc caacaaaaag tagtcgatgg gatagcttcc ggcatcaatg 17520

gagtggtaga tttggctaac caggctgtgc agaaaaagat aaacagtcgt ttggacccgc 17580

cgccagcaac cccaggtgaa atgcaagtgg aggaagaaat tcctccgcca gaaaaacgag 17640

gcgacaagcg tccgcgtccc gatttggaag agacgctggt gacgcgcgta gatgaaccgc 17700

cttcttatga ggaagcaacg aagcttggaa tgcccaccac tagaccgata gccccaatgg 17760

ccaccggggt gatgaaacct tctcagttgc atcgacccgt caccttggat ttgccccctc 17820

cccctgctgc tactgctgta cccgcttcta agcctgtcgc tgccccgaaa ccagtcgccg 17880

tagccaggtc acgtcccggg ggcgctcctc gtccaaatgc gcactggcaa aatactctga 17940

acagcatcgt gggtctaggc gtgcaaagtg taaaacgccg tcgctgcttt taattaaata 18000

tggagtagcg cttaacttgc ctatctgtgt atatgtgtca ttacacgccg tcacagcagc 18060

agaggaaaaa aggaagaggt cgtgcgtcga cgctgagtta ctttcaagat ggccacccca 18120

tcgatgctgc cccaatgggc atacatgcac atcgccggac aggatgcttc ggagtacctg 18180

agtccgggtc tggtgcagtt cgcccgcgcc acagacacct acttcaatct gggaaataag 18240

tttagaaatc ccaccgtagc gccgacccac gatgtgacca ccgaccgtag ccagcggctc 18300

atgttgcgct tcgtgcccgt tgaccgggag gacaatacat actcttacaa agtgcggtac 18360

accctggccg tgggcgacaa cagagtgctg gatatggcca gcacgttctt tgacattagg 18420

ggcgtgttgg acagaggtcc cagtttcaaa ccctattctg gtacggctta caactctctg 18480

gctcctaaag gcgctccaaa tgcatctcaa tggattgcaa aaggcgtacc aactgcagca 18540

gccgcaggca atggtgaaga agaacatgaa acagaggaga aaactgctac ttacactttt 18600

gccaatgctc ctgtaaaagc cgaggctcaa attacaaaag agggcttacc aataggtttg 18660

gagatttcag ctgaaaacga atctaaaccc atctatgcag ataaacttta tcagccagaa 18720

cctcaagtgg gagatgaaac ttggactgac ctagacggaa aaaccgaaga gtatggaggc 18780

agggctctaa agcctactac taacatgaaa ccctgttacg ggtcctatgc gaagcctact 18840

aatttaaaag gtggtcaggc aaaaccgaaa aactcggaac cgtcgagtga aaaaattgaa 18900

tatgatattg acatggaatt ttttgataac tcatcgcaaa gaacaaactt cagtcctaaa 18960

attgtcatgt atgcagaaaa tgtaggtttg gaaacgccag acactcatgt agtgtacaaa 19020

cctggaacag aagacacaag ttccgaagct aatttgggac aacagtctat gcccaacaga 19080

cccaactaca ttggcttcag agataacttt attggactca tgtactataa cagtactggt 19140

aacatggggg tgctggctgg tcaagcgtct cagttaaatg cagtggttga cttgcaggac 19200

agaaacacag aactttctta ccaactcttg cttgactctc tgggcgacag aaccagatac 19260

tttagcatgt ggaatcaggc tgtggacagt tatgatcctg atgtacgtgt tattgaaaat 19320

catggtgtgg aagatgaact tcccaactat tgttttccac tggacggcat aggtgttcca 19380

acaaccagtt acaaatcaat agttccaaat ggagaagata ataataattg gaaagaacct 19440

gaagtaaatg gaacaagtga gatcggacag ggtaatttgt ttgccatgga aattaacctt 19500

caagccaatc tatggcgaag tttcctttat tccaatgtgg ctctgtatct cccagactcg 19560

tacaaataca ccccgtccaa tgtcactctt ccagaaaaca aaaacaccta cgactacatg 19620

aacgggcggg tggtgccgcc atctctagta gacacctatg tgaacattgg tgccaggtgg 19680

tctctggatg ccatggacaa tgtcaaccca ttcaaccacc accgtaacgc tggcttgcgt 19740

taccgatcta tgcttctggg taacggacgt tatgtgcctt tccacataca agtgcctcaa 19800

aaattcttcg ctgttaaaaa cctgctgctt ctcccaggct cctacactta tgagtggaac 19860

tttaggaagg atgtgaacat ggttctacag agttccctcg gtaacgacct gcgggtagat 19920

ggcgccagca tcagtttcac gagcatcaac ctctatgcta ctttttccc catggctcac 19980

aacaccgctt ccacccttga agccatgctg cggaatgaca ccaatgatca gtcattcaac 20040

gactacctat ctgcagctaa catgctctac cccattcctg ccaatgcaac caatattccc 20100

atttccattc cttctcgcaa ctgggcggct ttcagaggct ggtcatttac cagactgaaa 20160

accaaagaaa ctccctcttt ggggtctgga tttgacccct actttgtcta ttctggttct 20220

attccctacc tggatggtac cttctacctg aaccacactt ttaagaaggt ttccatcatg 20280

tttgactctt cagtgagctg gcctggaaat gacaggttac tatctcctaa cgaatttgaa 20340

ataaagcgca ctgtggatgg cgaaggctac aacgtagccc aatgcaacat gaccaaagac 20400

tggttcttgg tacagatgct cgccaactac aacatcggct atcagggctt ctacattcca 20460

gaaggataca aagatcgcat gtattcattt ttcagaaact tccagcccat gagcaggcag 20520

gtggttgatg aggtcaatta caaagacttc aaggccgtcg ccataccta ccaacacaac 20580

aactctggct ttgtgggtta catggctccg accatgcgcc aaggtcaacc ctatcccgct 20640

aactatccct atccactcat tggaacaact gccgtaaata gtgttacgca gaaaaagttc 20700

ttgtgtgaca gaaccatgtg gcgcataccg ttctcgagca acttcatgtc tatgggggcc 20760

cttacagact tgggacagaa tatgctctat gccaactcag ctcatgctct ggacatgacc 20820

tttgaggtgg atcccatgga tgagcccacc ctgctttatc ttctcttcga agttttcgac 20880

gtggtcagag tgcatcagcc acaccgcggc atcatcgagg cagtctacct gcgtacaccg 20940

ttctcggccg gtaacgctac cacgtaagaa gcttcttgct tcttgcaaat agcagctgca 21000

accatggcct gcggatccca aaacggctcc agcgagcaag agctcagagc cattgtccaa 21060

gacctgggtt gcggacccta ttttttggga acctacgata agcgcttccc ggggttcatg 21120

gccccgata agctcgcctg tgccattgta aatacggccg gacgtgagac ggggggagag 21180

cactggttgg ctttcggttg gaacccacgt tctaacacct gctacctttt tgatcctttt 21240

ggattctcgg atgatcgtct caaacagatt taccagtttg aatatgaggg tctcctgcgc 21300

cgcagcgctc ttgctaccaa ggaccgctgt attacgctgg aaaaatctac ccagaccgtg 21360
cagggccccc gttctgccgc ctgcggactt ttctgctgca tgttccttca cgcctttgtg 21420
cactggcctg accgtcccat ggacggaaac cccaccatga aattgctaac tggagtgcca 21480
aacaacatgc ttcattctcc taaagtccag cccaccctgt gtgacaatca aaaagcactc 21540
taccattttc ttaataccca ttcgccttat tttcgctctc atcgtacaca catcgaaagg 21600
gccactgcgt tcgaccgtat ggatgttcaa taatgactca tgtaaacaac gtgttcaata 21660
aacatcactt tatttttta catgtatcaa ggctctggat tacttattta tttacaagtc 21720

gaatgggttc tgacgagaat cagaatgacc cgcaggcagt gatacgttgc ggaactgata 21780
cttgggttgc cacttgaatt cgggaatcac caacttggga accggtatat cgggcaggat 21840
gtcactccac agctttctgg tcagctgcaa agctccaagc aggtcaggag ccgaaatctt 21900
gaaatcacaa ttaggaccag tgctctgagc gcgagagttg cggtacaccg gattgcagca 21960
ctgaaacacc atcagcgacg gatgtctcac gcttgccagc acggtgggat ctgcaatcat 22020
gcccacatcc agatcttcag cattggcaat gctgaacggg gtcatcttgc aggtctgcct 22080
acccatggcg ggcacccaat taggcttgtg gttgcaatcg cagtgcaggg ggatcagtat 22140

catcttggcc tgatcctgtc tgattcctgg atacacggct ctcatgaaag catcatattg 22200
cttgaaagcc tgctgggctt tactaccctc gggataaaac atcccgcagg acctgctcga 22260
aaactggtta gcctgcacag ccggcatcat tcacacagca gcgggcgtca ttgttggcta 22320
tttgcaccac acttctgccc cagcggtttt gggtgatttt ggttcgctcg ggattctcct 22380
ttaaggctcg ttgtccgttc tcgctggcca catccatctc gataatctgc tccttctgaa 22440
tcataatatt gccatgcagg cacttcagct tgccctcata atcattgcag ccatgaggcc 22500
acaacgcaca gcctgtacat tcccaattat ggtgggcgat ctgagaaaaa gaatgtatca 22560

ttccctgcag aaatcttccc atcatcgtgc tcagtgtctt gtgactagtg aaagttaact 22620
ggatgcctcg gtgctcttcg tttacgtact ggtgacagat gcgcttgtat tgttcgtgtt 22680
gctcaggcat tagtttaaaa caggttctaa gttcgttatc cagcctgtac ttctccatca 22740
gcagacacat cacttccatg cctttctccc aagcagacac caggggcaag ctaatcggat 22800
tcttaacagt gcaggcagca gctcctttag ccagagggtc atctttagcg atcttctcaa 22860
tgcttctttt gccatccttc tcaacgatgc gcacgggcgg gtagctgaaa cccactgcta 22920
caagttgcgc ctcttctctt tcttcttcgc tgtcttgact gatgtcttgc atggggatat 22980

gtttggtctt ccttggcttc tttttggggg gtatcggagg aggaggactg tcgctccgtt 23040

ccggagacag ggaggattgt gacgtttcgc tcaccattac caactgactg tcggtagaag 23100

aacctgaccc cacacggcga caggtgtttt tcttcggggg cagaggtgga ggcgattgcg 23160

aagggctgcg gtccgacctg gaaggcggat gactggcaga acccccttccg cgttcggggg 23220

tgtgctccct gtggcggtcg cttaactgat ttccttcgcg gctggccatt gtgttctcct 23280

aggcagagaa acaacagaca tggaaactca gccattgctg tcaacatcgc cacgagtgcc 23340

atcacatctc gtcctcagcg acgaggaaaa ggagcagagc ttaagcattc caccgcccag 23400

tcctgccacc acctctaccc tagaagataa ggaggtcgac gcatctcatg acatgcagaa 23460

taaaaaagcg aaagagtctg agacagacat cgagcaagac ccgggctatg tgacaccggt 23520

ggaacacgag gaagagttga aacgctttct agagagagag gatgaaaact gcccaaaaca 23580

gcgagcagat aactatcacc aagatgctgg aaatagggat cagaacaccg actacctcat 23640

agggcttgac ggggaagacg cgctccttaa acatctagca agacagtcgc tcatagtcaa 23700

ggatgcatta ttggacagaa ctgaagtgcc catcagtgtg gaagagctca gctgcgccta 23760

cgagcttaac ctttttcac ctcgtactcc ccccaaacgt cagccaaacg gcacctgcga 23820

gccaaatcct cgcttaaact tttatccagc ttttgctgtg ccagaagtac tggctaccta 23880

tcacatcttt tttaaaaatc aaaaaattcc agtctcctgc cgcgctaatc gcacccgcgc 23940

cgatgcccta ctcaatctgg gacctggttc acgcttacct gatatagctt ccttggaaga 24000

ggttccaaag atcttcgagg gtctgggcaa taatgagact cgggccgcaa atgctctgca 24060

aaagggagaa aatggcatgg atgagcatca cagcgttctg gtggaattgg aaggcgataa 24120

tgccagactc gcagtactca agcgaagcgt cgaggtcaca cacttcgcat atcccgctgt 24180

caacctgccc cctaaagtca tgacggcggt catggaccag ttactcatta agcgcgcaag 24240

tccccttca gaagacatgc atgacccaga tgcctgtgat gagggtaaac cagtggtcag 24300

tgatgagcag ctaacccgat ggctgggcac cgactctccc cgggatttgg aagagcgtcg 24360

caagcttatg atggccgtgg tgctggttac cgtagaacta gagtgtctcc gacgtttctt 24420

taccgattca gaaaccttgc gcaaactcga agagaatctg cactacactt ttagacacgg 24480

ctttgtgcgg caggcatgca agatatctaa cgtggaactc accaacctgg tttcctacat 24540

gggtattctg catgagaatc gcctaggaca aagcgtgctg cacagcaccc ttaaggggga 24600

agcccgccgt gattacatcc gcgattgtgt ctatctctac ctgtgccaca cgtggcaaac 24660

cggcatgggt gtatggcagc aatgtttaga agaacagaac ttgaaagagc ttgacaagct 24720

cttacagaaa tctcttaagg ttctgtggac agggttcgac gagcgcaccg tcgcttccga 24780

cctggcagac ctcatcttcc cagagcgtct cagggttact ttgcgaaacg gattgcctga 24840

ctttatgagc cagagcatgc ttaacaattt tcgctctttc atcctggaac gctccggtat 24900

cctgcccgcc acctgctgcg cactgccctc cgactttgtg cctctcacct accgcgagtg 24960

ccccccgccg ctatggagtc actgctacct gttccgtctg gccaactatc tctcctacca 25020

ctcggatgtg atcgaggatg tgagcggaga cggcttgctg gagtgccact gccgctgcaa 25080

tctgtgcacg ccccaccggt ccctagcttg caacccccag ttgatgagcg aaacccagat 25140

aataggcacc tttgaattgc aaggccccag cagccaaggc gatgggtctt ctcctgggca 25200

aagtttaaaa ctgaccccgg gactgtggac ctccgcctac ttgcgcaagt ttgctccgga 25260

agattaccac ccctatgaaa tcaagttcta tgaggaccaa tcacagcctc caaaggccga 25320

actttcggct tgcgtcatca cccaggggggc aattctggcc caattgcaag ccatccaaaa 25380

atcccgccaa gaatttctac tgaaaaaggg taaggggtc taccttgacc cccagaccgg 25440

cgaggaactc aacacaaggt tccctcagga tgtcccaacg acgagaaaac aagaagttga 25500

aggtgcagcc gccgccccca gaagatatgg aggaagattg ggacagtcag gcagaggagg 25560

cggaggagga cagtctggag gacagtctgg aggaagacag tttggaggag gaaaacgagg 25620

aggcagagga ggtggaagaa gtaaccgccg acaaacagtt atcctcggct gcggagacaa 25680

gcaacagcgc taccatctcc gctccgagtc gaggaacccg gcggcgtccc agcagtagat 25740

gggacgagac cggacgcttc ccgaacccaa ccagcgcttc caagaccggt aagaaggatc 25800

ggcagggata caagtcctgg cgggggcata agaatgccat catctcctgc ttgcatgagt 25860

gcgggggcaa catatccttc acgcggcgct acttgctatt ccaccatggg gtgaactttc 25920

cgcgcaatgt tttgcattac taccgtcacc tccacagccc ctactatagc cagcaaatcc 25980

cgacagtctc gacagataaa gacagcggcg gcgacctcca acagaaaacc agcagcggca 26040

gttagaaaat acacaacaag tgcagcaaca ggaggattaa agattacagc caacgagcca 26100

gcgcaaaccc gagagttaag aaatcggatc tttccaaccc tgtatgccat cttccagcag 26160

agtcggggtc aagagcagga actgaaaata aaaaaccgat ctctgcgttc gctcaccaga 26220

agttgtttgt atcacaagag cgaagatcaa cttcagcgca ctctcgagga cgccgaggct 26280

ctcttcaaca agtactgcgc gctgactctt aaagagtagg cagcgaccgc gcttattcaa 26340

aaaaggcggg aattacatca tcctcgacat gagtaaagaa attcccacgc cttacatgtg 26400

gagttatcaa ccccaaatgg gattggcagc aggcgcctcc caggactact ccacccgcat 26460

gaattggctc agcgccgggc cttctatgat ttctcgagtt aatgatatac gcgcctaccg 26520

aaaccaaata cttttggaac agtcagctct taccaccacg ccccgccaac accttaatcc 26580

cagaaattgg cccgccgccc tagtgtacca ggaaagtccc gctcccacca ctgtattact 26640

tcctcgagac gcccaggccg aagtccaaat gactaatgca ggtgcgcagt tagctggcgg 26700

ctccacccta tgtcgtcaca ggcctcggca taatataaaa cgcctgatga tcagaggccg 26760

aggtatccag ctcaacgacg agtcggtgag ctctccgctt ggtctacgac cagacggaat 26820
ctttcagatt gccggctgcg ggagatcttc cttcacccct cgtcaggctg ttctgacttt 26880
ggaaagttcg tcttcgcaac cccgctcggg cggaatcggg accgttcaat ttgtagagga 26940
gtttactccc tctgtctact tcaacccctt ctccggatct cctgggcact acccggacga 27000
gttcataccg aacttcgacg cgattagcga gtcagtggac ggctacgatt gatgtctggt 27060
gacgcggctg agctatctcg gctgcgacat ctagaccact gccgccgctt tcgctgcttt 27120
gcccgggaac ttattgagtt catctacttc gaactcccca aggatcaccc tcaaggtccg 27180

gcccacggag tgcggattac tatcgaaggc aaaatagact ctcgcctgca acgaattttc 27240
tcccagcggc ccgtgctgat cgagcgagac cagggaaaca ccacggttag taatcaatta 27300
cggggtcatt agttcatagc ccatatatgg agttgcgatc gctgcgggcc atgtcataca 27360
ccgccttcag agcagccgga cctatctgcc cgttcgtgcc gtcgttgtta atcaccacat 27420
ggttattctg ctcaaacgtc ccggacgcct gcgaccggct gtctgccatg ctgcccggtg 27480
taccgacata accgccggtg gcatagccgc gcatcagccg gtaaagattc cccacgccaa 27540
tccggctggt tgcctccttc gtgaagacaa actcaccacg gtgaacaatc cccgctggct 27600

catatttgcc gccggttccc gtaaatcctc cggttgcaaa atggaatttc gccgcagcgg 27660
cctgaatggc tgtaccgcct gacgcggatg cgccgccacc aacagccccg ccaatggcgc 27720
tgccgatact cccgacaatc cccaccattg cctgcttaag cagaatttct gtcatcatgg 27780
acagcacgga acgggtgaag ctgcgccagt tctgctcact gccggtcagc atcgccgcca 27840
tattctgtgc aataccatca aaggtctgcg tggctgcact ttttacctgc gacatactgt 27900
ccgtggcgct ctcttcccac tcactccagc cggacttcag gcctgccatc cagttcccgc 27960
gaagctggtc ttcagccgcc caggtctttt tctgctctga catgacgtta ttcagcgcca 28020

gcggattatc gccatactgt tccttcaggc gctgttccgt ggcttcccgt tctgcctgcc 28080
ggtcagtcag cccccggctt ttcgcatcaa tggcggcccg ttttgcccgt tgctgctgtg 28140
cgaatttatc cgcctgctgc gccagcgcgt tcaggcgctc ctgatacgta accttgtcgc 28200
caagtgcagc cagctggcgt ttgtactcca gcgtctcatc tttatgcgcc agcagggatt 28260
tctcctgtgc agacagctgg cgacgttgcg ccgcctcctc cagtaccgcg aactgactct 28320
ccgccttcca caaatcccgg cgctgctggc tgattttctc atttgctccg gcatgcttct 28380
ccagcgtccg gagttctgcc tgaagcgtca gcagggcagc atgagcactg tcttcctgac 28440

gatcgcccgc agacaccttc acgctggact gtttcggctt tttcagcgtc gcttcataat 28500

ccttttcgc cgccgccatc agcgtgttgt aatccgcctg caggattttc ccgtctttca 28560

gtgccttgtt cagttcttcc tgacgggcgg tatatttctc cagcggcgtc tgcagccgtt 28620

cgtaagcctt ctgcgcctct tcggtatatt tcagccgtga cgcttcggta tcgctctgct 28680

gctgcgcatt tttgtcctgt tgagtctgct gctcagcctt ctttcgggcg gcttcaagcg 28740

caagacgggc cttttcacga tcatcccagt aacgcgcccg cgcttcatcg ttaacaaaat 28800

aatcatcctt gcgcagattc cagatgtcgt ctgctttctt atacgcagcc tctgccttaa 28860

tcagcatctc ctgcgcggta tcaggacgac caatatccag caccgcatcc cacatggatt 28920

tgaatgcccg cgcagtcctg tctgcccagg tctccagcgt gcccatgttc tctttcaggc 28980

ggcgggtctg gtcatcaaac cctttcgttg cggcctcgtt cgccgcctgc aatgccccgg 29040

cttcatcgcc ggaacgctgc aactgagcaa catacgcaat ctgctccgcc gacacgttat 29100

ggaactggcg agccatcgcc gtcagccccg acgtcgggtc tgtggtcagc ttcccgaagg 29160

cttcagcgac cttgtccacc tccacgccgg atgcagagga gaaacgcgcc acactctggc 29220

tgatggacgc aatctgagcc tcaccgctta cccccgcctt aaccagtgcg ctgagtgact 29280

cgctggtctg gttaaacgtc agccctgccg cctgcccggc tctggacagg accagcatac 29340

gatctgccgt cagtcccgcc tgattgccgg aaaggaccag cgttttgttg aaatcggaca 29400

gggttgagtt gccctgatac caggcatacg ccagcgcacc ggtcgccacc gccagcgagg 29460

tggcccccac catcggcagg gtgatcgcac cggcaagccc cctgaacatg gggatcatcc 29520

cgccgaagga gtccttcacc tgccccccct gttgcagcag gatcagccac ggactttgcc 29580

cgcctgcaag ctgcgtggcc acgtcggtga actgtgcagg cagcatacgc atggcggctt 29640

tatactgccc gacggaaatc cccgctttct gtgcagccag cgcctgtcgg ctcagcgact 29700

gttcaacgac tgccgctgtt tttttcgcat cactttccgt accagaaaaa tgacgcctga 29760

ctctggccat ctgctcgtca aatctggccg catccagact caaatcaacg acgtcgacta 29820

agctctagca tttgtgaacc atcaccctaa tcaagttttt tggggtcgag gtgccgtaaa 29880

gcactaaatc ggaaccctaa agggagcccc cgatttagag cttgacgggg aaagccggcg 29940

aacgtggcga gaaaggaagg gaagaaagcg aaaggagcgg gcgctagggc gctggcaagt 30000

gtagcggtca cgctgcgcgt aaccaccaca cccgccgcgc ttaatgcgcc gctacagggc 30060

gcgtggggat acccccctaga gccccagctg gttctttccg cctcagaagc catagagccc 30120

accgcatccc cagcatgcct gctattgtct tcccaatcct ccccccttgct gtcctgcccc 30180

accccacccc ccagaataga atgacaccta ctcagacaat gcgatgcaat ttcctcattt 30240

tattaggaaa ggacagtggg agtggcacct tccagggtca aggaaggcac gggggagggg 30300

caaacaacag atggctggca actagaaggc acagtcgagg ctgatcagcg ggtttgctag 30360

cttaggcgaa ggcgatgggg gtcttgaagg cgtgctggta ctccacgatg cccagctcgg 30420

tgttgctgtg cagctcctcc acgcggcgga aggcgaacat ggggcccccg ttctgcagga 30480

tgctggggtg gatggcgctc ttgaagtgca tgtggctgtc caccacgaag ctgtagtagc 30540

cgccgtcgcg caggctgaag gtgcgggcga agctgcccac cagcacgtta tcgcccatgg 30600

ggtgcaggtg ctccacggtg gcgttgctgc ggatgatctt gtcggtgaag atcacgctgt 30660

cctcggggaa gccggtgccc accaccttga agtcgccgat cacgcggccg gcctcgtagc 30720

ggtagctgaa gctcacgtgc agcacgccgc cgtcctcgta cttctcgatg cgggtgttgg 30780

tgtagccgcc gttgttgatg gcgtgcagga aggggttctc gtagccgctg gggtaggtgc 30840

cgaagtggta gaagccgtag cccatcacgt ggctcagcag gtaggggctg aaggtcaggg 30900

cgcctttggt gctcttcatc ttgttggtca tgcggccctg ctcgggggtg ccctctccgc 30960

cgcccaccag ctcgaactcc acgccgttca gggtgccggt gatgcggcac tcgatcttca 31020

tggcgggcat ggtggctagc ctagccagct tgggtctccc tatagtgagt cgtattaatt 31080

tcgataagcc agtaagcagt gggttctcta gttagccaga gagctctgct tatatagacc 31140

tcccaccgta cacgcctacc gcccatttgc gtcaatgggg cggagttgtt acgacatttt 31200

ggaaagtccc gttgattttg gtgccaaaac aaactcccat tgacgtcaat ggggtggaga 31260

cttggaaatc cccgtgagtc aaaccgctat ccacgcccat tgatgtactg ccaaaaccgc 31320

atcaccatgg taatagcgat gactaatacg tagatgtact gccaagtagg aaagtcccat 31380

aaggtcatgt actgggcata atgccaggcg ggccatttac cgtcattgac gtcaataggg 31440

ggcgtacttg gcatatgata cacttgatgt actgccaagt gggcagttta ccgtaaatac 31500

tccacccatt gacgtcaatg gaaagtccct attggcgtta ctatgggaac atacgtcatt 31560

attgacgtca atgggcgggg gtcgttgggc ggtcagccag gcgggccatt taccgtaagt 31620

tatgtaacgc ggaactccat atatgggcta tgaactaatg accccgtaat tgattactat 31680

taataactac aataatcaat gtcaacgcgt atatctggcc cgtacatcgc gaagcagcgc 31740

aaaacgccta accctaagca gattcttcat gcaattaagc ttcgcggtgc ttcttcagta 31800

cgctacggca aatgtcatcg acgtttttat ccggaaactg ctgtctggct ttttttgatt 31860

tcagaattag cctgacgggc aatgctgcga agggcgtttt cctgctgagg tgtcattgaa 31920

caagtcccat gtcggcaagc ataagcacac agaatatgaa gcccgctgcc agaaaaatgc 31980

attccgtggt tgtcatacct ggtttctctc atctgcttct gctttcgcca ccatcatttc 32040

cagcttttgt gaaagggatg cggctaacgt atgaaattct tcgtctgttt ctactggtat 32100

tggcacaaac ctgattccaa tttgagcaag gctatgtgcc atctcgatac tcgttcttaa 32160

ctcaacagaa gatgctttgt gcatacagcc cctcgtttat tatttatctc ctcagccagc 32220

cgctgtgctt tcagtggatt tcggataaca gaaaggccgg gaaataccca gcctcgcttt 32280
gtaacggagt agacgaaagt gattgcgcct acccggatat tatcgtgagg atgcgtcatc 32340
gccattgctc cccaaataca aaaccaattt cagccagtgc ctcgtccatt ttttcgatga 32400
actccggcac gatctcgtca aaactcgcca tgtacttttc atcccgctca atcacgacat 32460
aatgcaggcc ttcacgcttc atacgcgggt catagttggc aaagtaccag gcatttttc 32520
gcgtcaccca catgctgtac tgcacctggg ccatgtaagc tgactttatg gcctcgaaac 32580
caccgagccg gaacttcatg aaatcccggg aggtaaacgg gcatttcagt tcaaggccgt 32640

tgccgtcact gcataaacca tcgggagagc aggcggtacg catactttcg tcgcgataga 32700
tgatcgggga ttcagtaaca ttcacgccgg aagtgaattc aaacagggtt ctggcgtcgt 32760
tctcgta 32767

<210> 52
<211> 32759
<212> DNA
<213> Artificial Sequence
<220><223> Ad35E4PS3/WL-ps3
<220><221> misc_feature
<222> (2582)..(2735)
<223> Ad35 1-->154
<220><221> misc_feature
<222> (2744)..(2777)
<223> loxP
<220><221> misc_feature
<222> (2784)..(3110)

<223> Ad35 155-->481
<220><221> misc_feature
<222> (3111)..(3144)
<223> loxP
<220><221> misc_feature
<222> (27530)..(29999)
<223> lambda-1 (Complementary)
<220><221> misc_feature

<222> (30313)..(30527)

<223> BGH polyA

<220><221> misc_feature

<222> (30552)..(31217)

<223> copGFP (Complementary)

<220><221> misc_feature

<222> (31264)..(31916)

<223> CMV (Complementary)

<220><221> misc_feature

<222> (31968)..(33497)

<223> lambda-2

<220><221> misc_feature

<222> (33558)..(34893)

<223> Ad35 30544-->31879

<220><221> misc_feature

<222> (34889)..(36003)

<223> Ad5E4orf6

<220><221> misc_feature

<222> (36001)..(37823)

<223> Ad35 32972-->34794

<400> 52

taaacttggc gcgccctgag tgatttttct ctggtcccgc cgcatccata ccgccagttg 60

tttaccctca caacgttcca gtaaccgggc atgttcatca tcagtaaccc gtatcgtgag 120

catcctctct cgtttcatcg gtatcattac ccccatgaac agaaatcccc cttacacgga 180

ggcatcagtg accaaacagg aaaaaaccgc ccttaacatg gcccgcttta tcagaagcca 240

gacattaacg cttctggaga aactcaacga gctggacgcg gatgaacagg cagacatctg 300

tgaatcgctt cacgaccacg ctgatgagct ttaccgcagc tgcctcgcgc gtttcggtga 360

tgacggtgaa aacctctgac acatgcagct cccggagacg gtcacagctt gtctgtaagc 420

ggatgccggg agcagacaag cccgtcaggg cgcgtcagcg ggtgttggcg ggtgtcgggg 480

cgcagccatg acccagtcac gtagcgatag cggagtgtat actggcttaa ctatgcggca 540

tcagagcaga ttgtactgag agtgcaccat atgcggtgtg aaataccgca cagatgcgta 600

aggagaaaat accgcatcag gcgctcttcc gcttcctcgc tcactgactc gctgcgctcg 660

gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg cggtaatacg gttatccaca 720

gaatcagggg ataacgcagg aaagaacatg tgagcaaaag gccagcaaaa ggccaggaac 780

cgtaaaaagg ccgcgttgct ggcgtttttc cataggctcc gccccccctga cgagcatcac 840

aaaaatcgac gctcaagtca gaggtggcga aacccgacag gactataaag ataccaggcg 900

tttccccctg gaagctccct cgtgcgctct cctgttccga ccctgccgct taccggatac 960

ctgtccgcct ttctcccttc gggaagcgtg gcgctttctc atagctcacg ctgtaggtat 1020

ctcagttcgg tgtaggtcgt tcgctccaag ctgggctgtg tgcacgaacc ccccgttcag 1080

cccgaccgct gcgccttatc cggtaactat cgtcttgagt ccaacccggt aagacacgac 1140

ttatcgccac tggcagcagc cactggtaac aggattagca gagcgaggta tgtaggcggt 1200

gctacagagt tcttgaagtg gtggcctaac tacggctaca ctagaaggac agtatttggt 1260

atctgcgctc tgctgaagcc agttaccttc ggaaaaagag ttggtagctc ttgatccggc 1320

aaacaaacca ccgctggtag cggtggtttt tttgtttgca agcagcagat tacgcgcaga 1380

aaaaaaggat ctcaagaaga tcctttgatc ttttctacgg ggtctgacgc tcagtggaac 1440

gaaaactcac gttaagggat tttggtcatg agattatcaa aaaggatctt cacctagatc 1500

cttttaaatt aaaaatgaag ttttaaatca atctaaagta tatatgagta aacttggtct 1560

gacagttacc aatgcttaat cagtgaggca cctatctcag cgatctgtct atttcgttca 1620

tccatagttg cctgactccc cgtcgtgtag ataactacga tacgggaggg cttaccatct 1680

ggccccagtg ctgcaatgat accgcgagac ccacgctcac cggctccaga tttatcagca 1740

ataaaccagc cagccggaag ggccgagcgc agaagtggtc ctgcaacttt atccgcctcc 1800

atccagtcta ttaattgttg ccgggaagct agagtaagta gttcgccagt taatagtttg 1860

cgcaacgttg ttgccattgc tgcaggcatc gtggtgtcac gctcgtcgtt tggtatggct 1920

tcattcagct ccggttccca acgatcaagg cgagttacat gatccccat gttgtgcaaa 1980

aaagcggtta gctccttcgg tcctccgatc gttgtcagaa gtaagttggc cgcagtgtta 2040

tcactcatgg ttatggcagc actgcataat tctcttactg tcatgccatc cgtaagatgc 2100

ttttctgtga ctggtgagta ctcaaccaag tcattctgag aatagtgtat gcggcgaccg 2160

agttgctctt gcccggcgtc aacacgggat aataccgcgc cacatagcag aactttaaaa 2220

gtgctcatca ttggaaaacg ttcttcgggg cgaaaactct caaggatctt accgctgttg 2280

agatccagtt cgatgtaacc cactcgtgca cccaactgat cttcagcatc ttttactttc 2340

accagcgttt ctgggtgagc aaaaacagga aggcaaaatg ccgcaaaaaa gggaataagg 2400

gcgacacgga aatgttgaat actcatactc ttccttttc aatattattg aagcatttat 2460

cagggttatt gtctcatgag cggatacata tttgaatgta tttagaaaaa taaacaaata 2520

ggggttccgc gcacatttcc ccgaaaagtg ccacctgtct agctacgata tcctgtttaa 2580

acatcatcaa taatatacct tatagatgga atggtgccaa tatgtaaatg aggtgatttt 2640

aaaaagtgtg ggccgtgtgg tgattggctg tggggttaac ggttaaaagg ggcggcgcgg 2700

ccgtgggaaa atgacgtttt atgggggtgg agtttattta aatataactt cgtatagcat 2760

acattatacg aagttatgga tccttttgca agttgtcgcg ggaaatgtta cgcataaaaa 2820

ggcttctttt ctcacggaac tacttagttt tcccacggta tttaacagga aatgaggtag 2880

ttttgaccgg atgcaagtga aaattgctga ttttcgcgcg aaaactgaat gaggaagtgt 2940

ttttctgaat aatgtggtat ttatggcagg gtggagtatt tgttcagggc caggtagact 3000

ttgacccatt acgtggaggt ttcgattacc gtgtttttta cctgaatttc cgcgtaccgt 3060

gtcaaagtct tctgttttta cgtaggtgtc agctgatcgc tagggtattt ataacttcgt 3120

atagcataca ttatacgaag ttatatttaa ataggaatgt ttatgcctta ccagtgtaac 3180

atgaatcatg tgaaagtgtt gttggaacca gatgcctttt ccagaatgag cctaacagga 3240

atctttgaca tgaacacgca aatctggaag atcctgaggt atgatgatac gagatcgagg 3300

gtgcgcgcat gcgaatgcgg aggcaagcat gccaggttcc agccggtgtg tgtagatgtg 3360

accgaagatc tcagaccgga tcatttggtt attgcccgca ctggagcaga gttcggatcc 3420

agtggagaag aaactgacta aggtgagtat tgggaaaact ttggggtggg attttcagat 3480

ggacagattg agtaaaaatt tgtttttct gtcttgcagc tgacatgagt ggaaatgctt 3540

cttttaaggg gggagtcttc agcccttatc tgacagggcg tctcccatcc tgggcaggag 3600

ttcgtcagaa tgttatggga tctactgtgg atggaagacc cgttcaaccc gccaattctt 3660

caacgctgac ctatgctact ttaagttctt cacctttgga cgcagctgca gccgctgccg 3720

ccgcctctgt cgccgctaac actgtgcttg gaatgggtta ctatggaagc atcgtggcta 3780

attccacttc ctctaataac ccttctacac tgactcagga caagttactt gtccttttgg 3840

cccagctgga ggctttgacc caacgtctgg gtgaactttc tcagcaggtg gccgagttgc 3900

gagtacaaac tgagtctgct gtcggcacgg caaagtctaa ataaaaaaaa ttccagaatc 3960

aatgaataaa taaacgagct tgttgttgat ttaaaatcaa gtgtttttat ttcatttttc 4020

gcgcacggta tgccctggac caccgatctc gatcattgag aactcggtgg attttttcca 4080

gaatcctata gaggtgggat tgaatgttta gatacatggg cattaggccg tctttggggt 4140

ggagatagct ccattgaagg gattcatgct ccggggtagt gttgtaaatc acccagtcat 4200

aacaaggtcg cagtgcatgg tgttgcacaa tatcttttag aagtaggctg attgccacag 4260

ataagccctt ggtgtaggtg tttacaaacc ggttgagctg ggaggggtgc attcgaggtg 4320

aaattatgtg cattttggat tggattttta agttggcaat attgccgcca agatcccgtc 4380

ttgggttcat gttatgaagg actaccaaga cggtgtatcc ggtacattta ggaaatttat 4440

cgtgcagctt ggatggaaaa gcgtggaaaa atttggagac acccttgtgt cctccgagat 4500
tttccatgca ctcatccatg ataatagcaa tggggccgtg ggcagcggcg cgggcaaaca 4560
cgttccgtgg gtctgacaca tcatagttat gttcctgagt taaatcatca taagccattt 4620
taatgaattt ggggcggagc gtaccagatt ggggtatgaa tgttccttcg ggccccggag 4680
catagttccc ctcacagatt tgcatttccc aagctttcag ttctgagggt ggaatcatgt 4740
ccacctgggg ggctatgaag aacaccgttt cgggggcggg ggtgattagt tgggatgata 4800
gcaagtttct gagcaattga gatttgccac atccggtggg gccataaata attccgatta 4860

caggttgcag gtggtagttt agggaacggc aactgccgtc ttctcgaagc aaggggggcca 4920
cctcgttcat catttccctt acatgcatat tttcccgcac caaatccatt aggaggcgct 4980
ctcctcctag tgatagaagt tcttgtagtg aggaaaagtt tttcagcggt tttagaccgt 5040
cagccatggg cattttggaa agagtttgct gcaaaagttc tagtctgttc cacagttcag 5100
tgatgtgttc tatggcatct cgatccagca gacctcctcg tttcgcgggt ttggacggct 5160
cctggagtag ggtatgagac gatgggcgtc cagcgctgcc agggttcggt ccttccaggg 5220
tctcagtgtt cgagtcaggg ttgtttccgt cacagtgaag gggtgtgcgc ctgcttgggc 5280

gcttgccagg gtgcgcttca gactcattct gctggtggag aacttctgtc gcttggcgcc 5340
ctgtatgtcg gccaagtagc agtttaccat gagttcgtag ttgagcgcct cggctgcgtg 5400
gcctttggcg cggagcttac ctttggaagt tttcttgcat accgggcagt ataggcattt 5460
cagcgcatac agcttgggcg caaggaaaat ggattctggg gagtatgcat ccgcgccgca 5520
ggaggcgcaa acagtttcac attccaccag ccaggttaaa tccggttcat tggggtcaaa 5580
aacaagtttt ccgccatatt ttttgatgcg tttcttacct ttggtctcca taagttcgtg 5640
tcctcgttga gtgacaaaca ggctgtccgt atctccgtag actgatttta caggcctctt 5700

ctccagtgga gtgcctcggt cttcttcgta caggaactct gaccactctg atacaaaggc 5760
gcgcgtccag gccagcacaa aggaggctat gtgggagggg tagcgatcgt tgtcaaccag 5820
ggggtccacc ttttccaaag tatgcaaaca catgtcaccc tcttcaacat ccaggaatgt 5880
gattggcttg taggtgtatt tcacgtgacc tggggtcccc gctgggggggg tataaaaggg 5940
ggcggttctt tgctcttcct cactgtcttc cggatcgctg tccaggaacg tcagctgttg 6000
gggtaggtat tccctctcga aggcgggcat gacctctgca ctcaggttgt cagtttctaa 6060
gaacgaggag gatttgatat tgacagtgcc ggttgagatg cctttcatga ggttttcgtc 6120

catttggtca gaaaacacaa ttttttttatt gtcaagtttg gtggcaaatg atccatacag 6180

ggcgttggat aaaagtttgg caatggatcg catggtttgg ttcttttcct tgtccgcgcg 6240

ctctttggcg gcgatgttga gttggacata ctcgcgtgcc aggcacttcc attcggggaa 6300

gatagttgtt aattcatctg gcacgattct cacttgccac cctcgattat gcaaggtaat 6360

taaatccaca ctggtggcca cctcgcctcg aaggggttca ttggtccaac agagcctacc 6420

tcctttccta gaacagaaag ggggaagtgg gtctagcata agttcatcgg gagggtctgc 6480

atccatggta aagattcccg gaagtaaatc cttatcaaaa tagctgatgg gagtggggtc 6540

atctaaggcc atttgccatt ctcgagctgc cagtgcgcgc tcatatgggt taaggggact 6600

gccccagggc atgggatggg tgagagcaga ggcatacatg ccacagatgt catagacgta 6660

gatgggatcc tcaaagatgc ctatgtaggt tggatagcat cgccccccctc tgatacttgc 6720

tcgcacatag tcatatagtt catgtgatgg cgctagcagc cccggaccca agttggtgcg 6780

attgggtttt tctgttctgt agacgatctg gcgaaagatg gcgtgagaat tggaagagat 6840

ggtgggtctt tgaaaaatgt tgaaatgggc atgaggtaga cctacagagt ctctgacaaa 6900

gtgggcataa gattcttgaa gcttggttac cagttcggcg gtgacaagta cgtctagggc 6960

gcagtagtca agtgtttctt gaatgatgtc ataacctggt tggtttttct tttcccacag 7020

ttcgcggttg agaaggtatt cttcgcgatc cttccagtac tcttctagcg gaaacccgtc 7080

tttgtctgca cggtaagatc ctagcatgta gaactgatta actgccttgt aagggcagca 7140

gcccttctct acgggtagag agtatgcttg agcagctttt cgtagcgaag cgtgagtaag 7200

ggcaaaggtg tctctgacca tgactttgag aaattggtat ttgaagtcca tgtcgtcaca 7260

ggctccctgt tcccagagtt ggaagtctac ccgtttcttg taggcggggt tgggcaaagc 7320

gaaagtaaca tcattgaaga gaatcttacc ggctctgggc ataaaattgc gagtgatgcg 7380

gaaaggctgt ggtacttccg ctcgattgtt gatcacctgg gcagctagga cgatttcgtc 7440

gaaaccgttg atgttgtgtc ctacgatgta taattctatg aaacgcggcg tgcctctgac 7500

gtgaggtagc ttactgagct catcaaaggt taggtctgtg gggtcagata aggcgtagtg 7560

ttcgagagcc cattcgtgca ggtgaggatt tgcatgtagg aatgatgacc aaagatctac 7620

cgccagtgct gtttgtaact ggtcccgata ctgacgaaaa tgccggccaa ttgccatttt 7680

ttctggagtg acacagtaga aggttctggg gtcttgttgc catcgatccc acttgagttt 7740

aatggctaga tcgtgggcca tgttgacgag acgctcttct cctgagagtt tcatgaccag 7800

catgaaagga actagttgtt tgccaaagga tcccatccag gtgtaagttt ccacatcgta 7860

ggtcaggaag agtctttctg tgcgaggatg agagccgatc gggaagaact ggatttcctg 7920

ccaccagttg gaggattggc tgttgatgtg atggaagtag aagtttctgc ggcgcgccga 7980

gcattcgtgt ttgtgcttgt acagacggcc gcagtagtcg cagcgttgca cgggttgtat 8040

ctcgtgaatg agctgtacct ggcttcccttgacgagaaat ttcagtggga agccgaggcc 8100

tggcgattgt atctcgtgct cttctatatt cgctgtatcg gcctgttcat cttctgtttc 8160

gatggtggtc atgctgacga gcccccgcgg gaggcaagtc cagacctcgg cgcgggaggg 8220

gcggagctga aggacgagag cgcgcaggct ggagctgtcc agagtcctga gacgctgcgg 8280

actcaggtta gtaggtaggg acagaagatt aacttgcatg atcttttcca gggcgtgcgg 8340

gaggttcaga tggtacttga tttccacagg ttcgtttgta gagacgtcaa tggcttgcag 8400

ggttccgtgt cctttgggcg ccactaccgt acctttgttt tttcttttga tcggtggtgg 8460

ctctcttgct tcttgcatgc tcagaagcgg tgacggggac gcgcgccggg cggcagcggt 8520

tgttccggac ccgggggcat ggctggtagt ggcacgtcgg cgccgcgcac gggcaggttc 8580

tggtattgcg ctctgagaag acttgcgtgc gccaccacgc gtcgattgac gtcttgtatc 8640

tgacgtctct gggtgaaagc taccggcccc gtgagcttga acctgaaaga gagttcaaca 8700

gaatcaattt cggtatcgtt aacggcagct tgtctcagta tttcttgtac gtcaccagag 8760

ttgtcctggt aggcgatctc cgccatgaac tgctcgattt cttcctcctg aagatctccg 8820

cgacccgctc tttcgacggt ggccgcgagg tcattggaga tacggcccat gagttgggag 8880

aatgcattca tgcccgcctc gttccagacg cggctgtaaa ccacggcccc ctcggagtct 8940

cttgcgcgca tcaccacctg agcgaggtta agctccacgt gtctggttaa gaccgcatag 9000

ttgcataggc gctgaaaaag gtagttgagt gtggtggcaa tgtgttcggc gacgaagaaa 9060

tacatgatcc atcgtctcag cggcatttcg ctaacatcgc ccagagcttc caagcgctcc 9120

atggcctcgt agaagtccac ggcaaaatta aaaaactggg agtttcgcgc ggacacggtc 9180

aattcctcct cgagaagacg gatgagttcg gctatggtgg cccgtacttc gcgttcgaag 9240

gctcccggga tctcttcttc ctcttctatc tcttcttcca ctaacatctc ttcttcgtct 9300

tcaggcgggg gcggaggggg cacgcggcga cgtcgacggc gcacgggcaa acggtcgatg 9360

aatcgttcaa tgacctctcc gcggcggcgg cgcatggttt cagtgacggc gcggccgttc 9420

tcgcgcggtc gcagagtaaa aacaccgccg cgcatctcct taaagtggtg actgggaggt 9480

tctccgtttg ggagggagag ggcgctgatt atacatttta ttaattggcc cgtagggact 9540

gcgcgcagag atctgatcgt gtcaagatcc acgggatctg aaaaacctttc gacgaaagcg 9600

tctaaccagt cacagtcaca aggtaggctg agtacggctt cttgtgggcg ggggtggtta 9660

tgtgttcggt ctgggtcttc tgtttcttct tcatctcggg aaggtgagac gatgctgctg 9720

gtgatgaaat taaagtaggc agttctaaga cggcggatgg tggcgaggag caccaggtct 9780

ttgggtccgg cttgctggat acgcaggcga ttggccattc cccaagcatt atcctgacat 9840

ctagcaagat ctttgtagta gtcttgcatg agccgttcta cgggcacttc ttcctcaccc 9900

gttctgccat gcatacgtgt gagtccaaat ccgcgcattg gttgtaccag tgccaagtca 9960
gctacgactc tttcggcgag gatggcttgc tgtacttggg taagggtggc ttgaaagtca 10020
tcaaaatcca caaagcggtg gtaagcccct gtattaatgg tgtaagcaca gttggccatg 10080
actgaccagt taactgtctg gtgaccaggg cgcacgagct cggtgtattt aaggcgcgaa 10140
taggcgcggg tgtcaaagat gtaatcgttg caggtgcgca ccagatactg gtaccctata 10200
agaaaatgcg gcggtggttg gcggtagaga ggccatcgtt ctgtagctgg agcgccaggg 10260
gcgaggtctt ccaacataag gcggtgatag ccgtagatgt acctggacat ccaggtgatt 10320

cctgcggcgg tagtagaagc ccgaggaaac tcgcgtacgc ggttccaaat gttgcgtagc 10380
ggcatgaagt agttcattgt aggcacggtt tgaccagtga ggcgcgcgca gtcattgatg 10440
ctctatagac acggagaaaa tgaaagcgtt cagcgactcg actccgtagc ctggaggaac 10500
gtgaacgggt tgggtcgcgg tgtaccccgg ttcgagactt gtactcgagc cggccggagc 10560
cgcggctaac gtggtattgg cactcccgtc tcgacccagc ctacaaaaat ccaggatacg 10620
gaatcgagtc gttttgctgg tttccgaatg gcagggaagt gagtcctatt ttttttttt 10680
tttgccgctc agaatgcatc ccgtgctgcg acagatgcgc ccccaacaac agccccccctc 10740

gcagcagcag cagcagcaac cacaaaaggc tgtccctgca actactgcaa ctgccgccgt 10800
gagcggtgcg ggacagcccg cctatgatct ggacttggaa gagggcgaag gactggcacg 10860
tctaggtgcg ccttcgcccg agcggcatcc gcgagttcaa ctgaaaaaag attctcgcga 10920
ggcgtatgtg ccccaacaga acctatttag agacagaagc ggcgaggagc cggaggagat 10980
gcgagcttcc cgctttaacg cgggtcgtga gctgcgtcac ggtttggacc gaagacgagt 11040
gttgcgagac gaggatttcg aagttgatga agtgacaggg atcagtcctg ccagggcaca 11100
cgtggctgca gccaaccttg tatcggctta cgagcagaca gtaaaggaag agcgtaactt 11160

ccaaaagtct tttaataatc atgtgcgaac cctgattgcc cgcgaagaag ttacccttgg 11220
tttgatgcat ttgtgggatt tgatggaagc tatcattcag aaccctacta gcaaacctct 11280
gaccgcccag ctgtttctgg tggtgcaaca cagcagagac aatgaggctt tcagagaggc 11340
gctgctgaac atcaccgaac ccgaggggag atggttgtat gatcttatca acattctaca 11400
gagtatcata gtgcaggagc ggagcctggg cctggccgag aaggtagctg ccatcaatta 11460
ctcggttttg agcttgggaa aatattacgc tcgcaaaatc tacaagactc catacgttcc 11520
catagacaag gaggtgaaga tagatgggtt ctacatgcgc atgacgctca aggtcttgac 11580

cctgagcgat gatcttgggg tgtatcgcaa tgacagaatg catcgcgcgg ttagcgccag 11640

caggaggcgc gagttaagcg acagggaact gatgcacagt ttgcaaagag ctctgactgg 11700

agctggaacc gagggtgaga attacttcga catgggagct gacttgcagt ggcagcctaa 11760

tcgcagggct ctgagcgccg cgacggcagg atgtgagctt ccttacatag aagaggcgga 11820

tgaaggcgag gaggaagagg gcgagtactt ggaagactga tggcacaacc cgtgtttttt 11880

gctagatgga acagcaagca ccggatcccg caatgcgggc ggcgctgcag agccagccgt 11940

ccggcattaa ctcctcggac gattggaccc aggccatgca acgtatcatg gcgttgacga 12000

ctcgcaaccc cgaagccttt agacagcaac cccaggccaa ccgtctatcg gccatcatgg 12060

aagctgtagt gccttcccga tctaatccca ctcatgagaa ggtcctggcc atcgtgaacg 12120

cgttggtgga gaacaaagct attcgtccag atgaggccgg actggtatac aacgctctct 12180

tagaacgcgt ggctcgctac aacagtagca atgtgcaaac caatttggac cgtatgataa 12240

cagatgtacg cgaagccgtg tctcagcgcg aaaggttcca gcgtgatgcc aacctgggtt 12300

cgctggtggc gttaaatgct ttcttgagta ctcagcctgc taatgtgccg cgtggtcaac 12360

aggattatac taacttttta agtgctttga gactgatggt atcagaagta cctcagagcg 12420

aagtgtatca gtccggtcct gattacttct ttcagactag cagacagggc ttgcagacgg 12480

taaatctgag ccaagctttt aaaaacctta aaggtttgtg gggagtgcat gccccggtag 12540

gagaaagagc aaccgtgtct agcttgttaa ctccgaactc ccgcctgtta ttactgttgg 12600

tagctccttt caccgacagc ggtagcatcg accgtaattc ctatttgggt tacctactaa 12660

acctgtatcg cgaagccata gggcaaagtc aggtggacga gcagacctat caagaaatta 12720

cccaagtcag tcgcgctttg ggacaggaag acactggcag tttggaagcc actctgaact 12780

tcttgcttac caatcggtct caaaagatcc ctcctcaata tgctcttact gcggaggagg 12840

agaggatcct tagatatgtg cagcagagcg tgggattgtt tctgatgcaa gaggggggcaa 12900

ctccgactgc agcactggac atgacagcgc gaaatatgga gcccagcatg tatgccagta 12960

accgaccttt cattaacaaa ctgctggact acttgcacag agctgccgct atgaactctg 13020

attatttcac caatgccatc ttaaacccgc actggctgcc cccacctggt ttctacacgg 13080

gcgaatatga catgcccgac cctaatgacg gatttctgtg ggacgacgtg gacagcgatg 13140

ttttttcacc tctttctgat catcgcacgt ggaaaaagga aggcggtgat agaatgcatt 13200

cttctgcatc gctgtccggg gtcatgggtg ctaccgcggc tgagcccgag tctgcaagtc 13260

cttttcctag tctacccttt tctctacaca gtgtacgtag cagcgaagtg ggtagaataa 13320

gtcgcccgag tttaatgggc gaagaggagt acctaaacga ttccttgctc agaccggcaa 13380

gagaaaaaaa tttcccaaac aatggaatag aaagtttggt ggataaaatg agtagatgga 13440

agacttatgc tcaggatcac agagacgagc ctgggatcat ggggactaca agtagagcga 13500

gccgtagacg ccagcgccat gacagacaga ggggtcttgt gtgggacgat gaggattcgg 13560

ccgatgatag cagcgtgttg gacttgggtg ggagaggaag gggcaacccg tttgctcatt 13620

tgcgccctcg cttgggtggt atgttgtgaa aaaaaataaa aaagaaaaac tcaccaaggc 13680

catggcgacg agcgtacgtt cgttcttctt tattatctgt gtctagtata atgaggcgag 13740

tcgtgctagg cggagcggtg gtgtatccgg agggtcctcc tccttcgtac gagagcgtga 13800

tgcagcagca gcaggcgacg gcggtgatgc aatccccact ggaggctccc tttgtgcctc 13860

cgcgatacct ggcacctacg gagggcagaa acagcattcg ttactcggaa ctggcacctc 13920

agtacgatac caccaggttg tatctggtgg acaacaagtc ggcggacatt gcttctctga 13980

actatcagaa tgaccacagc aacttcttga ccacggtggt gcagaacaat gactttaccc 14040

ctacggaagc cagcacccag accattaact ttgatgaacg atcgcggtgg ggcggtcagc 14100

taaagaccat catgcatact aacatgccaa acgtgaacga gtatatgttt agtaacaagt 14160

tcaaagcgcg tgtgatggtg tccagaaaac ctcccgacgg tgctgcagtt ggggatactt 14220

atgatcacaa gcaggatatt ttggaatatg agtggttcga gtttactttg ccagaaggca 14280

acttttcagt tactatgact attgatttga tgaacaatgc catcatagat aattacttga 14340

aagtgggtag acagaatgga gtgcttgaaa gtgacattgg tgttaagttc gacaccagga 14400

acttcaagct gggatgggat cccgaaacca agttgatcat gcctggagtg tatacgtatg 14460

aagccttcca tcctgacatt gtcttactgc ctggctgcgg agtggatttt accgagagtc 14520

gtttgagcaa ccttcttggt atcagaaaaa aacagccatt tcaagagggt tttaagattt 14580

tgtatgaaga tttagaaggt ggtaatattc cggccctctt ggatgtagat gcctatgaga 14640

acagtaagaa agaacaaaaa gccaaaatag aagctgctac agctgctgca gaagctaagg 14700

caaacatagt tgccagcgac tctacaaggg ttgctaacgc tggagaggtc agaggagaca 14760

attttgcgcc aacacctgtt ccgactgcag aatcattatt ggccgatgtg tctgaaggaa 14820

cggacgtgaa actcactatt caacctgtag aaaaagatag taagaataga agctataatg 14880

tgttggaaga caaaatcaac acagcctatc gcagttggta tctttcgtac aattatggcg 14940

atcccgaaaa aggagtgcgt tcctggacat tgctcaccac ctcagatgtc acctgcggag 15000

cagagcaggt ttactggtcg cttccagaca tgatgaagga tcctgtcact ttccgctcca 15060

ctagacaagt cagtaactac cctgtggtgg gtgcagagct tatgcccgtc ttctcaaaga 15120

gcttctacaa cgaacaagct gtgtactccc agcagctccg ccagtccacc tcgcttacgc 15180

acgtcttcaa ccgctttcct gagaaccaga ttttaatccg tccgccggcg cccaccatta 15240

ccaccgtcag tgaaaacgtt cctgctctca cagatcacgg gaccctgccg ttgcgcagca 15300

gtatccgggg agtccaacgt gtgaccgtta ctgacgccag acgccgcacc tgtccctacg 15360

tgtacaaggc actgggcata gtcgcaccgc gcgtcctttc aagccgcact ttctaaaaaa 15420
aaaatgtcca ttcttatctc gcccagtaat aacaccggtt ggggtctgcg cgctccaagc 15480
aagatgtacg gaggcgcacg caaacgttct acccaacatc ccgtgcgtgt tcgcggacat 15540
tttcgcgctc catggggtgc cctcaagggc cgcactcgcg ttcgaaccac cgtcgatgat 15600
gtaatcgatc aggtggttgc cgacgcccgt aattatactc ctactgcgcc tacatctact 15660
gtggatgcag ttattgacag tgtagtggct gacgctcgca actatgctcg acgtaagagc 15720
cggcgaaggc gcattgccag acgccaccga gctaccactg ccatgcgagc cgcaagagct 15780

ctgctacgaa gagctagacg cgtggggcga agagccatgc ttagggcggc cagacgtgca 15840
gcttcgggcg ccagcgccgg caggtcccgc aggcaagcag ccgctgtcgc agcggcgact 15900
attgccgaca tggcccaatc gcgaagaggc aatgtatact gggtgcgtga cgctgccacc 15960
ggtcaacgtg tacccgtgcg cacccgtccc cctcgcactt agaagatact gagcagtctc 16020
cgatgttgtg tcccagcggc gaggatgtcc aagcgcaaat acaaggaaga aatgctgcag 16080
gttatcgcac ctgaagtcta cggccaaccg ttgaaggatg aaaaaaaacc ccgcaaaatc 16140
aagcgggtta aaaaggacaa aaaagaagag gaagatggcg atgatgggct ggcggagttt 16200

gtgcgcgagt ttgccccacg gcgacgcgtg caatggcgtg ggcgcaaagt tcgacatgtg 16260
ttgagacctg gaacttcggt ggtctttaca cccggcgagc gttcaagcgc tacttttaag 16320
cgttcctatg atgaggtgta cggggatgat gatattcttg agcaggcggc tgaccgatta 16380
ggcgagtttg cttatggcaa gcgtagtaga ataacttcca aggatgagac agtgtcaata 16440
cccttggatc atggaaatcc cacccctagt cttaaaccgg tcactttgca gcaagtgtta 16500
cccgtaactc cgcgaacagg tgttaaacgc gaaggtgaag atttgtatcc cactatgcaa 16560
ctgatggtac ccaaacgcca gaagttggag gacgttttgg agaaagtaaa agtggatcca 16620

gatattcaac ctgaggttaa agtgagaccc attaagcagg tagcgcctgg tctgggggta 16680
caaactgtag acattaagat tcccactgaa agtatggaag tgcaaactga acccgcaaag 16740
cctactgcca cctccactga agtgcaaacg gatccatgga tgcccatgcc tattacaact 16800
gacgccgccg gtcccactcg aagatcccga cgaaagtacg gtccagcaag tctgttgatg 16860
cccaattatg ttgtacaccc atctattatt cctactcctg gttaccgagg cactcgctac 16920
tatcgcagcc gaaacagtac ctcccgccgt cgccgcaaga cacctgcaaa tcgcagtcgt 16980
cgccgtagac gcacaagcaa accgactccc ggcgccctgg tgcggcaagt gtaccgcaat 17040

ggtagtgcgg aacctttgac actgccgcgt gcgcgttacc atccgagtat catcacttaa 17100

tcaatgttgc cgctgcctcc ttgcagatat ggccctcact tgtcgccttc gcgttcccat 17160

cactggttac cgaggaagaa actcgcgccg tagaagaggg atgttgggac gcggaatgcg 17220

acgctacagg cgacggcgtg ctatccgcaa gcaattgcgg ggtggttttt taccagcctt 17280

aattccaatt atcgctgctg caattggcgc gataccaggc atagcttccg tggcggttca 17340

ggcctcgcaa cgacattgac attggaaaaa aaacgtataa ataaaaaaaa atacaatgga 17400

ctctgacact cctggtcctg tgactatgtt ttcttagaga tggaagacat caatttttca 17460

tccttggctc cgcgacacgg cacgaagccg tacatgggca cctggagcga catcggcacg 17520

agccaactga acgggggcgc cttcaattgg agcagtatct ggagcgggct taaaaatttt 17580

ggctcaacca taaaaacata cgggaacaaa gcttggaaca gcagtacagg acaggcgctt 17640

agaaataaac ttaaagacca gaacttccaa caaaaagtag tcgatgggat agcttccggc 17700

atcaatggag tggtagattt ggctaaccag gctgtgcaga aaaagataaa cagtcgtttg 17760

gacccgccgc cagcaacccc aggtgaaatg caagtggagg aagaaattcc tccgccagaa 17820

aaacgaggcg acaagcgtcc gcgtcccgat ttggaagaga cgctggtgac gcgcgtagat 17880

gaaccgcctt cttatgagga agcaacgaag cttggaatgc ccaccactag accgatagcc 17940

ccaatggcca ccggggtgat gaaaccttct cagttgcatc gacccgtcac cttggatttg 18000

ccccctcccc ctgctgctac tgctgtaccc gcttctaagc ctgtcgctgc cccgaaacca 18060

gtcgccgtag ccaggtcacg tcccgggggc gctcctcgtc caaatgcgca ctggcaaaat 18120

actctgaaca gcatcgtggg tctaggcgtg caaagtgtaa aacgccgtcg ctgcttttaa 18180

ttaaatatgg agtagcgctt aacttgccta tctgtgtata tgtgtcatta cacgccgtca 18240

cagcagcaga ggaaaaaagg aagaggtcgt gcgtcgacgc tgagttactt tcaagatggc 18300

caccccatcg atgctgcccc aatgggcata catgcacatc gccggacagg atgcttcgga 18360

gtacctgagt ccgggtctgg tgcagttcgc ccgcgccaca gacacctact tcaatctggg 18420

aaataagttt agaaatccca ccgtagcgcc gacccacgat gtgaccaccg accgtagcca 18480

gcggctcatg ttgcgcttcg tgcccgttga ccgggaggac aatacatact cttacaaagt 18540

gcggtacacc ctggccgtgg gcgacaacag agtgctggat atggccagca cgttctttga 18600

cattaggggc gtgttggaca gaggtcccag tttcaaaccc tattctggta cggcttacaa 18660

ctctctggct cctaaaggcg ctccaaatgc atctcaatgg attgcaaaag gcgtaccaac 18720

tgcagcagcc gcaggcaatg gtgaagaaga acatgaaaca gaggagaaaa ctgctactta 18780

cacttttgcc aatgctcctg taaaagccga ggctcaaatt acaaaagagg gcttaccaat 18840

aggtttggag atttcagctg aaaacgaatc taaacccatc tatgcagata aactttatca 18900

gccagaacct caagtgggag atgaaacttg gactgaccta gacggaaaaa ccgaagagta 18960

tggaggcagg gctctaaagc ctactactaa catgaaaccc tgttacgggt cctatgcgaa 19020

gcctactaat ttaaaaggtg gtcaggcaaa accgaaaaac tcggaaccgt cgagtgaaaa 19080

aattgaatat gatattgaca tggaattttt tgataactca tcgcaaagaa caaacttcag 19140

tcctaaaatt gtcatgtatg cagaaaatgt aggtttggaa acgccagaca ctcatgtagt 19200

gtacaaacct ggaacagaag acacaagttc cgaagctaat ttgggacaac agtctatgcc 19260

caacagaccc aactacattg gcttcagaga taactttatt ggactcatgt actataacag 19320

tactggtaac atgggggtgc tggctggtca agcgtctcag ttaaatgcag tggttgactt 19380

gcaggacaga aacacagaac tttcttacca actcttgctt gactctctgg gcgacagaac 19440

cagatacttt agcatgtgga atcaggctgt ggacagttat gatcctgatg tacgtgttat 19500

tgaaaatcat ggtgtggaag atgaacttcc caactattgt tttccactgg acggcatagg 19560

tgttccaaca accagttaca aatcaatagt tccaaatgga gaagataata ataattggaa 19620

agaacctgaa gtaaatggaa caagtgagat cggacagggt aatttgtttg ccatggaaat 19680

taaccttcaa gccaatctat ggcgaagttt cctttattcc aatgtggctc tgtatctccc 19740

agactcgtac aaatacaccc cgtccaatgt cactcttcca gaaaacaaaa acacctacga 19800

ctacatgaac gggcgggtgg tgccgccatc tctagtagac acctatgtga acattggtgc 19860

caggtggtct ctggatgcca tggacaatgt caacccattc aaccaccacc gtaacgctgg 19920

cttgcgttac cgatctatgc ttctgggtaa cggacgttat gtgcctttcc acatacaagt 19980

gcctcaaaaa ttcttcgctg ttaaaaacct gctgcttctc ccaggctcct acacttatga 20040

gtggaacttt aggaaggatg tgaacatggt tctacagagt tccctcggta acgacctgcg 20100

ggtagatggc gccagcatca gtttcacgag catcaacctc tatgctactt ttttccccat 20160

ggctcacaac accgcttcca cccttgaagc catgctgcgg aatgacacca atgatcagtc 20220

attcaacgac tacctatctg cagctaacat gctctacccc attcctgcca atgcaaccaa 20280

tattcccatt tccattcctt ctcgcaactg ggcggctttc agaggctggt catttaccag 20340

actgaaaacc aaagaaactc cctctttggg gtctggattt gacccctact ttgtctattc 20400

tggttctatt ccctacctgg atggtacctt ctacctgaac cacactttta agaaggtttc 20460

catcatgttt gactcttcag tgagctggcc tggaaatgac aggttactat ctcctaacga 20520

atttgaaata aagcgcactg tggatggcga aggctacaac gtagcccaat gcaacatgac 20580

caaagactgg ttcttggtac agatgctcgc caactacaac atcggctatc agggcttcta 20640

cattccagaa ggatacaaag atcgcatgta ttcatttttc agaaacttcc agcccatgag 20700

caggcaggtg gttgatgagg tcaattacaa agacttcaag gccgtcgcca taccctacca 20760

acacaacaac tctggctttg tgggttacat ggctccgacc atgcgccaag gtcaacccta 20820

tcccgctaac tatccctatc cactcattgg aacaactgcc gtaaatagtg ttacgcagaa 20880
aaagttcttg tgtgacagaa ccatgtggcg cataccgttc tcgagcaact tcatgtctat 20940
gggggcccct acagacttgg gacagaatat gctctatgcc aactcagctc atgctctgga 21000
catgaccttt gaggtggatc ccatggatga gcccaccctg ctttatcttc tcttcgaagt 21060
tttcgacgtg gtcagagtgc atcagccaca ccgcggcatc atcgaggcag tctacctgcg 21120
tacaccgttc tcggccggta acgctaccac gtaagaagct tcttgcttct tgcaaatagc 21180
agctgcaacc atggcctgcg gatcccaaaa cggctccagc gagcaagagc tcagagccat 21240

tgtccaagac ctgggttgcg gaccctattt tttgggaacc tacgataagc gcttcccggg 21300
gttcatggcc cccgataagc tcgcctgtgc cattgtaaat acggccggac gtgagacggg 21360
gggagagcac tggttggctt tcggttggaa cccacgttct aacacctgct acctttttga 21420
tccttttgga ttctcggatg atcgtctcaa acagatttac cagtttgaat atgagggtct 21480
cctgcgccgc agcgctcttg ctaccaagga ccgctgtatt acgctggaaa aatctaccca 21540
gaccgtgcag ggcccccgtt ctgccgcctg cggacttttc tgctgcatgt tccttcacgc 21600
ctttgtgcac tggcctgacc gtcccatgga cggaaacccc accatgaaat tgctaactgg 21660

agtgccaaac aacatgcttc attctcctaa agtccagccc accctgtgtg acaatcaaaa 21720
agcactctac cattttctta atacccattc gccttatttt cgctctcatc gtacacacat 21780
cgaaagggcc actgcgttcg accgtatgga tgttcaataa tgactcatgt aaacaacgtg 21840
ttcaataaac atcactttat ttttttacat gtatcaaggc tctggattac ttatttattt 21900
acaagtcgaa tgggttctga cgagaatcag aatgacccgc aggcagtgat acgttgcgga 21960
actgatactt gggttgccac ttgaattcgg gaatcaccaa cttgggaacc ggtatatcgg 22020
gcaggatgtc actccacagc tttctggtca gctgcaaagc tccaagcagg tcaggagccg 22080

aaatcttgaa atcacaatta ggaccagtgc tctgagcgcg agagttgcgg tacaccggat 22140
tgcagcactg aaacaccatc agcgacggat gtctcacgct tgccagcacg gtgggatctg 22200
caatcatgcc cacatccaga tcttcagcat tggcaatgct gaacggggtc atcttgcagg 22260
tctgcctacc catggcgggc acccaattag gcttgtggtt gcaatcgcag tgcaggggga 22320
tcagtatcat cttggcctga tcctgtctga ttcctggata cacggctctc atgaaagcat 22380
catattgctt gaaagcctgc tgggctttac taccctcggg ataaaacatc ccgcaggacc 22440
tgctcgaaaa ctggttagcc tgcacagccg gcatcattca cacagcagcg ggcgtcattg 22500

ttggctattt gcaccacact tctgccccag cggttttggg tgattttggt tcgctcggga 22560

ttctccttta aggctcgttg tccgttctcg ctggccacat ccatctcgat aatctgctcc 22620

ttctgaatca taatattgcc atgcaggcac ttcagcttgc cctcataatc attgcagcca 22680

tgaggccaca acgcacagcc tgtacattcc caattatggt gggcgatctg agaaaaagaa 22740

tgtatcattc cctgcagaaa tcttcccatc atcgtgctca gtgtcttgtg actagtgaaa 22800

gttaactgga tgcctcggtg ctcttcgttt acgtactggt gacagatgcg cttgtattgt 22860

tcgtgttgct caggcattag tttaaaacag gttctaagtt cgttatccag cctgtacttc 22920

tccatcagca gacacatcac ttccatgcct ttctcccaag cagacaccag gggcaagcta 22980

atcggattct taacagtgca ggcagcagct cctttagcca gagggtcatc tttagcgatc 23040

ttctcaatgc ttcttttgcc atccttctca acgatgcgca cgggcgggta gctgaaaccc 23100

actgctacaa gttgcgcctc ttctctttct tcttcgctgt cttgactgat gtcttgcatg 23160

gggatatgtt tggtcttcct tggcttcttt ttgggggggta tcggaggagg aggactgtcg 23220

ctccgttccg gagacaggga ggattgtgac gtttcgctca ccattaccaa ctgactgtcg 23280

gtagaagaac ctgaccccac acggcgacag gtgtttttct tcgggggcag aggtggaggc 23340

gattgcgaag ggctgcggtc cgacctggaa ggcggatgac tggcagaacc ccttccgcgt 23400

tcgggggtgt gctccctgtg gcggtcgctt aactgatttc cttcgcggct ggccattgtg 23460

ttctcctagg cagagaaaca acagacatgg aaactcagcc attgctgtca acatcgccac 23520

gagtgccatc acatctcgtc ctcagcgacg aggaaaagga gcagagctta agcattccac 23580

cgcccagtcc tgccaccacc tctaccctag aagataagga ggtcgacgca tctcatgaca 23640

tgcagaataa aaaagcgaaa gagtctgaga cagacatcga gcaagacccg ggctatgtga 23700

caccggtgga acacgaggaa gagttgaaac gctttctaga gagagaggat gaaaactgcc 23760

caaaacagcg agcagataac tatcaccaag atgctggaaa tagggatcag aacaccgact 23820

acctcatagg gcttgacggg gaagacgcgc tccttaaaca tctagcaaga cagtcgctca 23880

tagtcaagga tgcattattg gacagaactg aagtgcccat cagtgtggaa gagctcagct 23940

gcgcctacga gcttaacctt ttttcacctc gtactccccc caaacgtcag ccaaacggca 24000

cctgcgagcc aaatcctcgc ttaaactttt atccagcttt tgctgtgcca gaagtactgg 24060

ctacctatca catctttttt aaaaatcaaa aaattccagt ctcctgccgc gctaatcgca 24120

cccgcgccga tgccctactc aatctgggac ctggttcacg cttacctgat atagcttcct 24180

tggaagaggt tccaaagatc ttcgagggtc tgggcaataa tgagactcgg gccgcaaatg 24240

ctctgcaaaa gggagaaaat ggcatggatg agcatcacag cgttctggtg gaattggaag 24300

gcgataatgc cagactcgca gtactcaagc gaagcgtcga ggtcacacac ttcgcatatc 24360

ccgctgtcaa cctgcccccct aaagtcatga cggcggtcat ggaccagtta ctcattaagc 24420

gcgcaagtcc cctttcagaa gacatgcatg acccagatgc ctgtgatgag ggtaaaccag 24480

tggtcagtga tgagcagcta acccgatggc tgggcaccga ctctccccgg gatttggaag 24540

agcgtcgcaa gcttatgatg gccgtggtgc tggttaccgt agaactagag tgtctccgac 24600

gtttctttac cgattcagaa accttgcgca aactcgaaga gaatctgcac tacactttta 24660

gacacggctt tgtgcggcag gcatgcaaga tatctaacgt ggaactcacc aacctggttt 24720

cctacatggg tattctgcat gagaatcgcc taggacaaag cgtgctgcac agcaccctta 24780

agggggaagc ccgccgtgat tacatccgcg attgtgtcta tctctacctg tgccacacgt 24840

ggcaaaccgg catgggtgta tggcagcaat gtttagaaga acagaacttg aaagagcttg 24900

acaagctctt acagaaatct cttaaggttc tgtggacagg gttcgacgag cgcaccgtcg 24960

cttccgacct ggcagacctc atcttcccag agcgtctcag ggttactttg cgaaacggat 25020

tgcctgactt tatgagccag agcatgctta acaattttcg ctctttcatc ctggaacgct 25080

ccggtatcct gcccgccacc tgctgcgcac tgccctccga ctttgtgcct ctcacctacc 25140

gcgagtgccc cccgccgcta tggagtcact gctacctgtt ccgtctggcc aactatctct 25200

cctaccactc ggatgtgatc gaggatgtga gcggagacgg cttgctggag tgccactgcc 25260

gctgcaatct gtgcacgccc caccggtccc tagcttgcaa cccccagttg atgagcgaaa 25320

cccagataat aggcaccttt gaattgcaag gccccagcag ccaaggcgat gggtcttctc 25380

ctgggcaaag tttaaaactg accccgggac tgtggacctc cgcctacttg cgcaagtttg 25440

ctccggaaga ttaccacccc tatgaaatca agttctatga ggaccaatca cagcctccaa 25500

aggccgaact ttcggcttgc gtcatcaccc agggggcaat tctggcccaa ttgcaagcca 25560

tccaaaaatc ccgccaagaa tttctactga aaaagggtaa gggggtctac cttgaccccc 25620

agaccggcga ggaactcaac acaaggttcc ctcaggatgt cccaacgacg agaaaacaag 25680

aagttgaagg tgcagccgcc gcccccagaa gatatggagg aagattggga cagtcaggca 25740

gaggaggcgg aggaggacag tctggaggac agtctggagg aagacagttt ggaggaggaa 25800

aacgaggagg cagaggaggt ggaagaagta accgccgaca aacagttatc ctcggctgcg 25860

gagacaagca acagcgctac catctccgct ccgagtcgag gaacccggcg gcgtcccagc 25920

agtagatggg acgagaccgg acgcttcccg aacccaacca gcgcttccaa gaccggtaag 25980

aaggatcggc agggatacaa gtcctggcgg gggcataaga atgccatcat ctcctgcttg 26040

catgagtgcg ggggcaacat atccttcacg cggcgctact tgctattcca ccatggggtg 26100

aactttccgc gcaatgtttt gcattactac cgtcacctcc acagccccta ctatagccag 26160

caaatcccga cagtctcgac agataaagac agcggcggcg acctccaaca gaaaaccagc 26220

agcggcagtt agaaaataca caacaagtgc agcaacagga ggattaaaga ttacagccaa 26280

cgagccagcg caaacccgag agttaagaaa tcggatcttt ccaaccctgt atgccatctt 26340
ccagcagagt cggggtcaag agcaggaact gaaaataaaa aaccgatctc tgcgttcgct 26400
caccagaagt tgtttgtatc acaagagcga agatcaactt cagcgcactc tcgaggacgc 26460
cgaggctctc ttcaacaagt actgcgcgct gactcttaaa gagtaggcag cgaccgcgct 26520
tattcaaaaa aggcgggaat tacatcatcc tcgacatgag taaagaaatt cccacgcctt 26580
acatgtggag ttatcaaccc caaatgggat tggcagcagg cgcctcccag gactactcca 26640
cccgcatgaa ttggctcagc gccgggcctt ctatgatttc tcgagttaat gatatacgcg 26700

cctaccgaaa ccaaatactt ttggaacagt cagctcttac caccacgccc cgccaacacc 26760
ttaatcccag aaattggccc gccgccctag tgtaccagga aagtcccgct cccaccactg 26820
tattacttcc tcgagacgcc caggccgaag tccaaatgac taatgcaggt gcgcagttag 26880
ctggcggctc caccctatgt cgtcacaggc ctcggcataa tataaaacgc ctgatgatca 26940
gaggccgagg tatccagctc aacgacgagt cggtgagctc tccgcttggt ctacgaccag 27000
acggaatctt tcagattgcc ggctgcggga gatcttcctt cacccctcgt caggctgttc 27060
tgactttgga aagttcgtct tcgcaacccc gctcgggcgg aatcgggacc gttcaatttg 27120

tagaggagtt tactccctct gtctacttca accccttctc cggatctcct gggcactacc 27180
cggacgagtt cataccgaac ttcgacgcga ttagcgagtc agtggacggc tacgattgat 27240
gtctggtgac gcggctgagc tatctcggct gcgacatcta gaccactgcc gccgctttcg 27300
ctgctttgcc cgggaactta ttgagttcat ctacttcgaa ctccccaagg atcaccctca 27360
aggtccggcc cacggagtgc ggattactat cgaaggcaaa atagactctc gcctgcaacg 27420
aattttctcc cagcggcccg tgctgatcga gcgagaccag ggaaacacca cggttagtaa 27480
tcaattacgg ggtcattagt tcatagccca tatatggagt tgcgatcgct gcgggccatg 27540

tcatacaccg ccttcagagc agccggacct atctgcccgt tcgtgccgtc gttgttaatc 27600
accacatggt tattctgctc aaacgtcccg gacgcctgcg accggctgtc tgccatgctg 27660
cccggtgtac cgacataacc gccggtggca tagccgcgca tcagccggta aagattcccc 27720
acgccaatcc ggctggttgc ctccttcgtg aagacaaact caccacggtg aacaatcccc 27780
gctggctcat atttgccgcc ggttcccgta aatcctccgg ttgcaaaatg gaatttcgcc 27840
gcagcggcct gaatggctgt accgcctgac gcggatgcgc cgccaccaac agccccgcca 27900
atggcgctgc cgatactccc gacaatcccc accattgcct gcttaagcag aatttctgtc 27960

atcatggaca gcacggaacg ggtgaagctg cgccagttct gctcactgcc ggtcagcatc 28020

gccgccatat tctgtgcaat accatcaaag gtctgcgtgg ctgcactttt tacctgcgac 28080

atactgtccg tggcgctctc ttcccactca ctccagccgg acttcaggcc tgccatccag 28140

ttcccgcgaa gctggtcttc agccgcccag gtctttttct gctctgacat gacgttattc 28200

agcgccagcg gattatcgcc atactgttcc ttcaggcgct gttccgtggc ttcccgttct 28260

gcctgccggt cagtcagccc ccggcttttc gcatcaatgg cggcccgttt tgcccgttgc 28320

tgctgtgcga atttatccgc ctgctgcgcc agcgcgttca ggcgctcctg atacgtaacc 28380

ttgtcgccaa gtgcagccag ctggcgtttg tactccagcg tctcatcttt atgcgccagc 28440

agggatttct cctgtgcaga cagctggcga cgttgcgccg cctcctccag taccgcgaac 28500

tgactctccg ccttccacaa atcccggcgc tgctggctga ttttctcatt tgctccggca 28560

tgcttctcca gcgtccggag ttctgcctga agcgtcagca gggcagcatg agcactgtct 28620

tcctgacgat cgcccgcaga caccttcacg ctggactgtt tcggcttttt cagcgtcgct 28680

tcataatcct ttttcgccgc cgccatcagc gtgttgtaat ccgcctgcag gattttcccg 28740

tctttcagtg ccttgttcag ttcttcctga cgggcggtat atttctccag cggcgtctgc 28800

agccgttcgt aagccttctg cgcctcttcg gtatatttca gccgtgacgc ttcggtatcg 28860

ctctgctgct gcgcattttt gtcctgttga gtctgctgct cagccttctt tcgggcggct 28920

tcaagcgcaa gacgggcctt ttcacgatca tcccagtaac gcgcccgcgc ttcatcgtta 28980

acaaaataat catccttgcg cagattccag atgtcgtctg ctttcttata cgcagcctct 29040

gccttaatca gcatctcctg cgcggtatca ggacgaccaa tatccagcac cgcatcccac 29100

atggatttga atgcccgcgc agtcctgtct gcccaggtct ccagcgtgcc catgttctct 29160

ttcaggcggc gggtctggtc atcaaaccct ttcgttgcgg cctcgttcgc cgcctgcaat 29220

gccccggctt catcgccgga acgctgcaac tgagcaacat acgcaatctg ctccgccgac 29280

acgttatgga actggcgagc catcgccgtc agccccgacg tcgggtctgt ggtcagcttc 29340

ccgaaggctt cagcgacctt gtccacctcc acgccggatg cagaggagaa acgcgccaca 29400

ctctggctga tggacgcaat ctgagcctca ccgcttaccc ccgccttaac cagtgcgctg 29460

agtgactcgc tggtctggtt aaacgtcagc cctgccgcct gcccggctct ggacaggacc 29520

agcatacgat ctgccgtcag tcccgcctga ttgccggaaa ggaccagcgt tttgttgaaa 29580

tcggacaggg ttgagttgcc ctgataccag gcatacgcca gcgcaccggt cgccaccgcc 29640

agcgaggtgg cccccaccat cggcagggtg atcgcaccgg caagcccccc gaacatgggg 29700

atcatcccgc cgaaggagtc cttcacctgc ccccccctgtt gcagcaggat cagccacgga 29760

ctttgcccgc ctgcaagctg cgtggccacg tcggtgaact gtgcaggcag catacgcatg 29820

gcggctttat actgcccgac ggaaatcccc gctttctgtg cagccagcgc ctgtcggctc 29880

agcgactgtt caacgactgc cgctgttttt ttcgcatcac tttccgtacc agaaaaatga 29940

cgcctgactc tggccatctg ctcgtcaaat ctggccgcat ccagactcaa atcaacgacg 30000

tcgactaagc tctagcattt gtgaaccatc accctaatca agttttttgg ggtcgaggtg 30060

ccgtaaagca ctaaatcgga accctaaagg gagcccccga tttagagctt gacggggaaa 30120

gccggcgaac gtggcgagaa aggaagggaa gaaagcgaaa ggagcgggcg ctagggcgct 30180

ggcaagtgta gcggtcacgc tgcgcgtaac caccacaccc gccgcgctta atgcgccgct 30240

acagggcgcg tggggatacc ccctagagcc ccagctggtt ctttccgcct cagaagccat 30300

agagcccacc gcatccccag catgcctgct attgtcttcc caatcctccc ccttgctgtc 30360

ctgccccacc ccacccccca gaatagaatg acacctactc agacaatgcg atgcaatttc 30420

ctcattttat taggaaagga cagtgggagt ggcaccttcc agggtcaagg aaggcacggg 30480

ggaggggcaa acaacagatg gctggcaact agaaggcaca gtcgaggctg atcagcgggt 30540

ttgctagctt aggcgaaggc gatgggggtc ttgaaggcgt gctggtactc cacgatgccc 30600

agctcggtgt tgctgtgcag ctcctccacg cggcggaagg cgaacatggg gcccccgttc 30660

tgcaggatgc tggggtggat ggcgctcttg aagtgcatgt ggctgtccac cacgaagctg 30720

tagtagccgc cgtcgcgcag gctgaaggtg cgggcgaagc tgcccaccag cacgttatcg 30780

cccatggggt gcaggtgctc cacggtggcg ttgctgcgga tgatcttgtc ggtgaagatc 30840

acgctgtcct cggggaagcc ggtgcccacc accttgaagt cgccgatcac gcggccggcc 30900

tcgtagcggt agctgaagct cacgtgcagc acgccgccgt cctcgtactt ctcgatgcgg 30960

gtgttggtgt agccgccgtt gttgatggcg tgcaggaagg ggttctcgta gccgctgggg 31020

taggtgccga agtggtagaa gccgtagccc atcacgtggc tcagcaggta ggggctgaag 31080

gtcagggcgc ctttggtgct cttcatcttg ttggtcatgc ggccctgctc gggggtgccc 31140

tctccgccgc ccaccagctc gaactccacg ccgttcaggg tgccggtgat gcggcactcg 31200

atcttcatgg cgggcatggt ggctagccta gccagcttgg gtctccctat agtgagtcgt 31260

attaatttcg ataagccagt aagcagtggg ttctctagtt agccagagag ctctgcttat 31320

atagacctcc caccgtacac gcctaccgcc catttgcgtc aatggggcgg agttgttacg 31380

acattttgga aagtcccgtt gattttggtg ccaaaacaaa ctcccattga cgtcaatggg 31440

gtggagactt ggaaatcccc gtgagtcaaa ccgctatcca cgcccattga tgtactgcca 31500

aaaccgcatc accatggtaa tagcgatgac taatacgtag atgtactgcc aagtaggaaa 31560

gtcccataag gtcatgtact gggcataatg ccaggcgggc catttaccgt cattgacgtc 31620

aataggggggc gtacttggca tatgatacac ttgatgtact gccaagtggg cagtttaccg 31680

taaatactcc acccattgac gtcaatggaa agtccctatt ggcgttacta tgggaacata 31740

cgtcattatt gacgtcaatg ggcgggggtc gttgggcggt cagccaggcg ggccatttac 31800
cgtaagttat gtaacgcgga actccatata tgggctatga actaatgacc ccgtaattga 31860
ttactattaa taactacaat aatcaatgtc aacgcgtata tctggcccgt acatcgcgaa 31920
gcagcgcaaa acgcctaacc ctaagcagat tcttcatgca attaagcttc gcggtgcttc 31980
ttcagtacgc tacggcaaat gtcatcgacg tttttatccg gaaactgctg tctggctttt 32040
tttgatttca gaattagcct gacgggcaat gctgcgaagg gcgttttcct gctgaggtgt 32100
cattgaacaa gtcccatgtc ggcaagcata agcacacaga atatgaagcc cgctgccaga 32160

aaaatgcatt ccgtggttgt catacctggt ttctctcatc tgcttctgct ttcgccacca 32220
tcatttccag cttttgtgaa agggatgcgg ctaacgtatg aaattcttcg tctgtttcta 32280
ctggtattgg cacaaacctg attccaattt gagcaaggct atgtgccatc tcgatactcg 32340
ttcttaactc aacagaagat gctttgtgca tacagcccct cgtttattat ttatctcctc 32400
agccagccgc tgtgctttca gtggatttcg gataacagaa aggccgggaa atacccagcc 32460
tcgctttgta acggagtaga cgaaagtgat tgcgcctacc cggatattat cgtgaggatg 32520
cgtcatcgcc attgctcccc aaatacaaaa ccaatttcag ccagtgcctc gtccattttt 32580

tcgatgaact ccggcacgat ctcgtcaaaa ctcgccatgt acttttcatc ccgctcaatc 32640
acgacataat gcaggccttc acgcttcata cgcgggtcat agttggcaaa gtaccaggca 32700
ttttttcgcg tcacccacat gctgtactgc acctgggcca tgtaagctga ctttatggc 32759

<210> 53
<211> 32767
<212> DNA
<213> Artificial Sequence
<220><223> pAd35GFP-5E4-CCDb
<400> 53

catcatcaat aatatacctt atagatggaa tggtgccaat atgtaaatga ggtgatttta 60
aaaagtgtgg gccgtgtggt gattggctgt ggggttaacg gttaaaaggg gcggcgcggc 120
cgtgggaaaa tgacgtttta tgggggtgga gttttttgc aagttgtcgc gggaaatgtt 180

acgcataaaa aggcttcttt tctcacggaa ctacttagtt ttcccacggt atttaacagg 240
aaatgaggta gttttgaccg gatgcaagtg aaaattgctg attttcgcgc gaaaactgaa 300
tgaggaagtg tttttctgaa taatgtggta tttatggcag ggtggagtat ttgttcaggg 360
ccaggtagac tttgacccat tacgtggagg tttcgattac cgtgtttttt acctgaattt 420

ccgcgtaccg tgtcaaagtc ttctgttttt acgtaggtgt cagctgatcg ctagggtatt 480

tgttcaaaaa aaagcccgct cattaggcgg gctgggttat attccccaga acatcaggtt 540

aatggcgttt ttgatgtcat tttcgcggtg gctgagatca gccacttctt ccccgataac 600

ggagaccggc acactggcca tatcggtggt catcatgcgc cagctttcat ccccgatatg 660

caccaccggg taaagttcac gggagacttt atctgacagc agacgtgcac tggccagggg 720

gatcaccatc cgtcgcccgg gcgtgtcaat aatatcactc tgtacatcca caaacagacg 780

ataacggctc tctcttttat aggtgtaaac cttaaactgc ataatctgac ctcctggtta 840

tgtgtgggag ggctaaccat ggatccatgg ttaccaatgc ttaatcagtg aggcacctat 900

ctcagcgatc tgtctatttc gttcatccat agttgcctga ctccccgtcg tgtagataac 960

tacgatacgg gagggcttac catctggccc cagtgctgca atgataccgc gagacccacg 1020

ctcaccggct ccagatttat cagcaataaa ccagccagcc ggaagggccg agcgcagaag 1080

tggtcctgca actttatccg cctccatcca gtctattaat tgttgccggg aagctagagt 1140

aagtagttcg ccagttaata gtttgcgcaa cgttgttgcc attgctgcag gcatcgtggt 1200

gtcacgctcg tcgtttggta tggcttcatt cagctccggt tcccaacgat caaggcgagt 1260

tacatgatcc cccatgttgt gcaaaaaagc ggttagctcc ttcggtcctc cgatcgttgt 1320

cagaagtaag ttggccgcag tgttatcact catggttatg gcagcactgc ataattctct 1380

tactgtcatg ccatccgtaa gatgcttttc tgtgactggt gagtactcaa ccaagtcatt 1440

ctgagaatag tgtatgcggc gaccgagttg ctcttgcccg gcgtcaacac gggataatac 1500

cgcgccacat agcagaactt taaaagtgct catcattgga aaacgttctt cggggcgaaa 1560

actctcaagg atcttaccgc tgttgagatc cagttcgatg taacccactc gtgcacccaa 1620

ctgatcttca gcatctttta ctttcaccag cgtttctggg tgagcaaaaa caggaaggca 1680

aaatgccgca aaaaagggaa taagggcgac acggaaatgt tgaatactca tactcttcct 1740

ttttcaatat tattgaagca tttatcaggg ttattgtctc atgagcggat acatatttga 1800

atgtatttag aaaaataaac aaagctacct taagagagtg tcaggaatgt ttatgcctta 1860

ccagtgtaac atgaatcatg tgaaagtgtt gttggaacca gatgcctttt ccagaatgag 1920

cctaacagga atctttgaca tgaacacgca aatctggaag atcctgaggt atgatgatac 1980

gagatcgagg gtgcgcgcat gcgaatgcgg aggcaagcat gccaggttcc agccggtgtg 2040

tgtagatgtg accgaagatc tcagaccgga tcatttggtt attgcccgca ctggagcaga 2100

gttcggatcc agtggagaag aaactgacta aggtgagtat tgggaaaact ttggggtggg 2160

attttcagat ggacagattg agtaaaaatt tgtttttttct gtcttgcagc tgacatgagt 2220

ggaaatgctt cttttaaggg gggagtcttc agcccttatc tgacagggcg tctcccatcc 2280

tgggcaggag ttcgtcagaa tgttatggga tctactgtgg atggaagacc cgttcaaccc 2340
gccaattctt caacgctgac ctatgctact ttaagttctt cacctttgga cgcagctgca 2400
gccgctgccg ccgcctctgt cgccgctaac actgtgcttg gaatgggtta ctatggaagc 2460
atcgtggcta attccacttc ctctaataac ccttctacac tgactcagga caagttactt 2520
gtccttttgg cccagctgga ggctttgacc caacgtctgg gtgaactttc tcagcaggtg 2580
gccgagttgc gagtacaaac tgagtctgct gtcggcacgg caaagtctaa ataaaaaaaa 2640
ttccagaatc aatgaataaa taaacgagct tgttgttgat ttaaaatcaa gtgtttttat 2700

ttcatttttc gcgcacggta tgccctggac caccgatctc gatcattgag aactcggtgg 2760
atttttttcca gaatcctata gaggtgggat tgaatgttta gatacatggg cattaggccg 2820
tctttggggt ggagatagct ccattgaagg gattcatgct ccggggtagt gttgtaaatc 2880
acccagtcat aacaaggtcg cagtgcatgg tgttgcacaa tatcttttag aagtaggctg 2940
attgccacag ataagccctt ggtgtaggtg tttacaaacc ggttgagctg ggaggggtgc 3000
attcgaggtg aaattatgtg cattttggat tggattttta agttggcaat attgccgcca 3060
agatcccgtc ttgggttcat gttatgaagg actaccaaga cggtgtatcc ggtacattta 3120

ggaaatttat cgtgcagctt ggatggaaaa gcgtggaaaa atttggagac acccttgtgt 3180
cctccgagat tttccatgca ctcatccatg ataatagcaa tggggccgtg ggcagcggcg 3240
cgggcaaaca cgttccgtgg gtctgacaca tcatagttat gttcctgagt taaatcatca 3300
taagccattt taatgaattt ggggcggagc gtaccagatt ggggtatgaa tgttccttcg 3360
ggccccggag catagttccc ctcacagatt tgcatttccc aagctttcag ttctgagggt 3420
ggaatcatgt ccacctgggg ggctatgaag aacaccgttt cgggggcggg ggtgattagt 3480
tgggatgata gcaagtttct gagcaattga gatttgccac atccggtggg gccataaata 3540

attccgatta caggttgcag gtggtagttt agggaacggc aactgccgtc ttctcgaagc 3600
aaggggggcca cctcgttcat catttccctt acatgcatat tttcccgcac caaatccatt 3660
aggaggcgct ctcctcctag tgatagaagt tcttgtagtg aggaaaagtt tttcagcggt 3720
tttagaccgt cagccatggg cattttggaa agagtttgct gcaaaagttc tagtctgttc 3780
cacagttcag tgatgtgttc tatggcatct cgatccagca gacctcctcg tttcgcgggt 3840
ttggacggct cctggagtag ggtatgagac gatgggcgtc cagcgctgcc agggttcggt 3900
ccttccaggg tctcagtgtt cgagtcaggg ttgtttccgt cacagtgaag gggtgtgcgc 3960

ctgcttgggc gcttgccagg gtgcgcttca gactcattct gctggtggag aacttctgtc 4020

gcttggcgcc ctgtatgtcg gccaagtagc agtttaccat gagttcgtag ttgagcgcct 4080

cggctgcgtg gcctttggcg cggagcttac ctttggaagt tttcttgcat accgggcagt 4140

ataggcattt cagcgcatac agcttgggcg caaggaaaat ggattctggg gagtatgcat 4200

ccgcgccgca ggaggcgcaa acagtttcac attccaccag ccaggttaaa tccggttcat 4260

tggggtcaaa aacaagtttt ccgccatatt ttttgatgcg tttcttacct ttggtctcca 4320

taagttcgtg tcctcgttga gtgacaaaca ggctgtccgt atctccgtag actgatttta 4380

caggcctctt ctccagtgga gtgcctcggt cttcttcgta caggaactct gaccactctg 4440

atacaaaggc gcgcgtccag gccagcacaa aggaggctat gtgggagggg tagcgatcgt 4500

tgtcaaccag ggggtccacc ttttccaaag tatgcaaaca catgtcaccc tcttcaacat 4560

ccaggaatgt gattggcttg taggtgtatt tcacgtgacc tggggtcccc gctgggggggg 4620

tataaaaggg ggcggttctt tgctcttcct cactgtcttc cggatcgctg tccaggaacg 4680

tcagctgttg gggtaggtat tccctctcga aggcgggcat gacctctgca ctcaggttgt 4740

cagtttctaa gaacgaggag gatttgatat tgacagtgcc ggttgagatg cctttcatga 4800

ggttttcgtc catttggtca gaaaacacaa tttttttatt gtcaagtttg gtggcaaatg 4860

atccatacag ggcgttggat aaaagtttgg caatggatcg catggtttgg ttcttttcct 4920

tgtccgcgcg ctctttggcg gcgatgttga gttggacata ctcgcgtgcc aggcacttcc 4980

attcggggaa gatagttgtt aattcatctg gcacgattct cacttgccac cctcgattat 5040

gcaaggtaat taaatccaca ctggtggcca cctcgcctcg aaggggttca ttggtccaac 5100

agagcctacc tcctttccta gaacagaaag ggggaagtgg gtctagcata agttcatcgg 5160

gagggtctgc atccatggta aagattcccg gaagtaaatc cttatcaaaa tagctgatgg 5220

gagtggggtc atctaaggcc atttgccatt ctcgagctgc cagtgcgcgc tcatatgggt 5280

taaggggact gccccagggc atgggatggg tgagagcaga ggcatacatg ccacagatgt 5340

catagacgta gatgggatcc tcaaagatgc ctatgtaggt tggatagcat cgccccccctc 5400

tgatacttgc tcgcacatag tcatatagtt catgtgatgg cgctagcagc cccggaccca 5460

agttggtgcg attgggtttt tctgttctgt agacgatctg gcgaaagatg gcgtgagaat 5520

tggaagagat ggtgggtctt tgaaaaatgt tgaaatgggc atgaggtaga cctacagagt 5580

ctctgacaaa gtgggcataa gattcttgaa gcttggttac cagttcggcg gtgacaagta 5640

cgtctagggc gcagtagtca agtgtttctt gaatgatgtc ataacctggt tggtttttct 5700

tttcccacag ttcgcggttg agaaggtatt cttcgcgatc cttccagtac tcttctagcg 5760

gaaacccgtc tttgtctgca cggtaagatc ctagcatgta gaactgatta actgccttgt 5820

aagggcagca gcccttctct acgggtagag agtatgcttg agcagctttt cgtagcgaag 5880

cgtgagtaag ggcaaaggtg tctctgacca tgactttgag aaattggtat ttgaagtcca 5940

tgtcgtcaca ggctccctgt tcccagagtt ggaagtctac ccgtttcttg taggcggggt 6000

tgggcaaagc gaaagtaaca tcattgaaga gaatcttacc ggctctgggc ataaaattgc 6060

gagtgatgcg gaaaggctgt ggtacttccg ctcgattgtt gatcacctgg gcagctagga 6120

cgatttcgtc gaaaccgttg atgttgtgtc ctacgatgta taattctatg aaacgcggcg 6180

tgcctctgac gtgaggtagc ttactgagct catcaaaggt taggtctgtg gggtcagata 6240

aggcgtagtg ttcgagagcc cattcgtgca ggtgaggatt tgcatgtagg aatgatgacc 6300

aaagatctac cgccagtgct gtttgtaact ggtcccgata ctgacgaaaa tgccggccaa 6360

ttgccatttt ttctggagtg acacagtaga aggttctggg gtcttgttgc catcgatccc 6420

acttgagttt aatggctaga tcgtgggcca tgttgacgag acgctcttct cctgagagtt 6480

tcatgaccag catgaaagga actagttgtt tgccaaagga tcccatccag gtgtaagttt 6540

ccacatcgta ggtcaggaag agtctttctg tgcgaggatg agagccgatc gggaagaact 6600

ggatttcctg ccaccagttg gaggattggc tgttgatgtg atggaagtag aagtttctgc 6660

ggcgcgccga gcattcgtgt ttgtgcttgt acagacggcc gcagtagtcg cagcgttgca 6720

cgggttgtat ctcgtgaatg agctgtacct ggcttccctt gacgagaaat ttcagtggga 6780

agccgaggcc tggcgattgt atctcgtgct cttctatatt cgctgtatcg gcctgttcat 6840

cttctgtttc gatggtggtc atgctgacga gcccccgcgg gaggcaagtc cagacctcgg 6900

cgcgggaggg gcggagctga aggacgagag cgcgcaggct ggagctgtcc agagtcctga 6960

gacgctgcgg actcaggtta gtaggtaggg acagaagatt aacttgcatg atcttttcca 7020

gggcgtgcgg gaggttcaga tggtacttga tttccacagg ttcgtttgta gagacgtcaa 7080

tggcttgcag ggttccgtgt cctttgggcg ccactaccgt acctttgttt tttcttttga 7140

tcggtggtgg ctctcttgct tcttgcatgc tcagaagcgg tgacggggac gcgcgccggg 7200

cggcagcggt tgttccggac ccgggggcat ggctggtagt ggcacgtcgg cgccgcgcac 7260

gggcaggttc tggtattgcg ctctgagaag acttgcgtgc gccaccacgc gtcgattgac 7320

gtcttgtatc tgacgtctct gggtgaaagc taccggcccc gtgagcttga acctgaaaga 7380

gagttcaaca gaatcaattt cggtatcgtt aacggcagct tgtctcagta tttcttgtac 7440

gtcaccagag ttgtcctggt aggcgatctc cgccatgaac tgctcgattt cttcctcctg 7500

aagatctccg cgacccgctc tttcgacggt ggccgcgagg tcattggaga tacggcccat 7560

gagttgggag aatgcattca tgcccgcctc gttccagacg cggctgtaaa ccacggcccc 7620

ctcggagtct cttgcgcgca tcaccacctg agcgaggtta agctccacgt gtctggtgaa 7680

gaccgcatag ttgcataggc gctgaaaaag gtagttgagt gtggtggcaa tgtgttcggc 7740

gacgaagaaa tacatgatcc atcgtctcag cggcatttcg ctaacatcgc ccagagcttc 7800
caagcgctcc atggcctcgt agaagtccac ggcaaaatta aaaaactggg agtttcgcgc 7860
ggacacggtc aattcctcct cgagaagacg gatgagttcg gctatggtgg cccgtacttc 7920
gcgttcgaag gctcccggga tctcttcttc ctcttctatc tcttcttcca ctaacatctc 7980
ttcttcgtct tcaggcgggg gcggaggggg cacgcggcga cgtcgacggc gcacgggcaa 8040
acggtcgatg aatcgttcaa tgacctctcc gcggcggcgg cgcatggttt cagtgacggc 8100
gcggccgttc tcgcgcggtc gcagagtaaa aacaccgccg cgcatctcct taaagtggtg 8160

actgggaggt tctccgtttg ggagggagag ggcgctgatt atacatttta ttaattggcc 8220
cgtagggact gcgcgcagag atctgatcgt gtcaagatcc acgggatctg aaaacctttc 8280
gacgaaagcg tctaaccagt cacagtcaca aggtaggctg agtacggctt cttgtgggcg 8340
ggggtggtta tgtgttcggt ctgggtcttc tgtttcttct tcatctcggg aaggtgagac 8400
gatgctgctg gtgatgaaat taaagtaggc agttctaaga cggcggatgg tggcgaggag 8460
caccaggtct ttgggtccgg cttgctggat acgcaggcga ttggccattc cccaagcatt 8520
atcctgacat ctagcaagat ctttgtagta gtcttgcatg agccgttcta cgggcacttc 8580

ttcctcaccc gttctgccat gcatacgtgt gagtccaaat ccgcgcattg gttgtaccag 8640
tgccaagtca gctacgactc tttcggcgag gatggcttgc tgtacttggg taagggtggc 8700
ttgaaagtca tcaaaatcca caaagcggtg gtaagcccct gtattaatgg tgtaagcaca 8760
gttggccatg actgaccagt taactgtctg gtgaccaggg cgcacgagct cggtgtattt 8820
aaggcgcgaa taggcgcggg tgtcaaagat gtaatcgttg caggtgcgca ccagatactg 8880
gtaccctata agaaaatgcg gcggtggttg gcggtagaga ggccatcgtt ctgtagctgg 8940
agcgccaggg gcgaggtctt ccaacataag gcggtgatag ccgtagatgt acctggacat 9000

ccaggtgatt cctgcggcgg tagtagaagc ccgaggaaac tcgcgtacgc ggttccaaat 9060
gttgcgtagc ggcatgaagt agttcattgt aggcacggtt tgaccagtga ggcgcgcgca 9120
gtcattgatg ctctatagac acggagaaaa tgaaagcgtt cagcgactcg actccgtagc 9180
ctggaggaac gtgaacgggt tgggtcgcgg tgtaccccgg ttcgagactt gtactcgagc 9240
cggccggagc cgcggctaac gtggtattgg cactcccgtc tcgacccagc ctacaaaaat 9300
ccaggatacg gaatcgagtc gttttgctgg tttccgaatg gcagggaagt gagtcctatt 9360
ttttttttt ttgccgctca gatgcatccc gtgctgcgac agatgcgccc ccaacaacag 9420

ccccctcgc agcagcagca gcagcaacca caaaaggctg tccctgcaac tactgcaact 9480

gccgccgtga gcggtgcggg acagcccgcc tatgatctgg acttggaaga gggcgaagga 9540

ctggcacgtc taggtgcgcc ttcgcccgag cggcatccgc gagttcaact gaaaaaagat 9600

tctcgcgagg cgtatgtgcc ccaacagaac ctatttagag acagaagcgg cgaggagccg 9660

gaggagatgc gagcttcccg ctttaacgcg ggtcgtgagc tgcgtcacgg tttggaccga 9720

agacgagtgt tgcgagacga ggatttcgaa gttgatgaag tgacagggat cagtcctgcc 9780

agggcacacg tggctgcagc caaccttgta tcggcttacg agcagacagt aaaggaagag 9840

cgtaacttcc aaaagtcttt taataatcat gtgcgaaccc tgattgcccg cgaagaagtt 9900

acccttggtt tgatgcattt gtgggatttg atggaagcta tcattcagaa ccctactagc 9960

aaacctctga ccgcccagct gtttctggtg gtgcaacaca gcagagacaa tgaggctttc 10020

agagaggcgc tgctgaacat caccgaaccc gaggggagat ggttgtatga tcttatcaac 10080

attctacaga gtatcatagt gcaggagcgg agcctgggcc tggccgagaa ggtagctgcc 10140

atcaattact cggttttgag cttgggaaaa tattacgctc gcaaaatcta caagactcca 10200

tacgttccca tagacaagga ggtgaagata gatgggttct acatgcgcat gacgctcaag 10260

gtcttgaccc tgagcgatga tcttggggtg tatcgcaatg acagaatgca tcgcgcggtt 10320

agcgccagca ggaggcgcga gttaagcgac agggaactga tgcacagttt gcaaagagct 10380

ctgactggag ctggaaccga gggtgagaat tacttcgaca tgggagctga cttgcagtgg 10440

cagcctagtc gcagggctct gagcgccgcg acggcaggat gtgagcttcc ttacatagaa 10500

gaggcggatg aaggcgagga ggaagagggc gagtacttgg aagactgatg gcacaacccg 10560

tgttttttgc tagatggaac agcaagcacc ggatcccgca atgcgggcgg cgctgcagag 10620

ccagccgtcc ggcattaact cctcggacga ttggacccag gccatgcaac gtatcatggc 10680

gttgacgact cgcaaccccg aagcctttag acagcaaccc caggccaacc gtctatcggc 10740

catcatggaa gctgtagtgc cttcccgatc taatcccact catgagaagg tcctggccat 10800

cgtgaacgcg ttggtggaga acaaagctat tcgtccagat gaggccggac tggtatacaa 10860

cgctctctta gaacgcgtgg ctcgctacaa cagtagcaat gtgcaaacca atttggaccg 10920

tatgataaca gatgtacgcg aagccgtgtc tcagcgcgaa aggttccagc gtgatgccaa 10980

cctgggttcg ctggtggcgt taaatgcttt cttgagtact cagcctgcta atgtgccgcg 11040

tggtcaacag gattatacta acttttaag tgctttgaga ctgatggtat cagaagtacc 11100

tcagagcgaa gtgtatcagt ccggtcctga ttacttcttt cagactagca gacagggctt 11160

gcagacggta aatctgagcc aagcttttaa aaaccttaaa ggtttgtggg gagtgcatgc 11220

cccggtagga gaaagagcaa ccgtgtctag cttgttaact ccgaactccc gcctgttatt 11280

actgttggta gctcctttca ccgacagcgg tagcatcgac cgtaattcct atttgggtta 11340

cctactaaac ctgtatcgcg aagccatagg gcaaagtcag gtggacgagc agacctatca 11400

agaaattacc caagtcagtc gcgctttggg acaggaagac actggcagtt tggaagccac 11460

tctgaacttc ttgcttacca atcggtctca aaagatccct cctcaatatg ctcttactgc 11520

ggaggaggag aggatcctta gatatgtgca gcagagcgtg ggattgtttc tgatgcaaga 11580

gggggcaact ccgactgcag cactggacat gacagcgcga aatatggagc ccagcatgta 11640

tgccagtaac cgacctttca ttaacaaact gctggactac ttgcacagag ctgccgctat 11700

gaactctgat tatttcacca atgccatctt aaacccgcac tggctgcccc cacctggttt 11760

ctacacgggc gaatatgaca tgcccgaccc taatgacgga tttctgtggg acgacgtgga 11820

cagcgatgtt ttttcacctc tttctgatca tcgcacgtgg aaaaaggaag gcggtgatag 11880

aatgcattct tctgcatcgc tgtccggggt catgggtgct accgcggctg agcccgagtc 11940

tgcaagtcct tttcctagtc taccctttc tctacacagt gtacgtagca gcgaagtggg 12000

tagaataagt cgcccgagtt taatgggcga agaggagtac ctaaacgatt ccttgctcag 12060

accggcaaga gaaaaaaatt tcccaaacaa tggaatagaa agtttggtgg ataaaatgag 12120

tagatggaag acttatgctc aggatcacag agacgagcct gggatcatgg ggactacaag 12180

tagagcgagc cgtagacgcc agcgccatga cagacagagg ggtcttgtgt gggacgatga 12240

ggattcggcc gatgatagca gcgtgttgga cttgggtggg agaggaaggg gcaacccgtt 12300

tgctcatttg cgccctcgct tgggtggtat gttgtgaaaa aaaataaaaa agaaaaactc 12360

accaaggcca tggcgacgag cgtacgttcg ttcttcttta ttatctgtgt ctagtataat 12420

gaggcgagtc gtgctaggcg gagcggtggt gtatccggag ggtcctcctc cttcgtacga 12480

gagcgtgatg cagcagcagc aggcgacggc ggtgatgcaa tccccactgg aggctccctt 12540

tgtgcctccg cgatacctgg cacctacgga gggcagaaac agcattcgtt actcggaact 12600

ggcacctcag tacgatacca ccaggttgta tctggtggac aacaagtcgg cggacattgc 12660

ttctctgaac tatcagaatg accacagcaa cttcttgacc acggtggtgc agaacaatga 12720

ctttacccct acggaagcca gcacccagac cattaacttt gatgaacgat cgcggtgggg 12780

cggtcagcta aagaccatca tgcatactaa catgccaaac gtgaacgagt atatgtttag 12840

taacaagttc aaagcgcgtg tgatggtgtc cagaaaacct cccgacggtg ctgcagttgg 12900

ggatacttat gatcacaagc aggatatttt ggaatatgag tggttcgagt ttactttgcc 12960

agaaggcaac ttttcagtta ctatgactat tgatttgatg aacaatgcca tcatagataa 13020

ttacttgaaa gtgggtagac agaatggagt gcttgaaagt gacattggtg ttaagttcga 13080

caccaggaac ttcaagctgg gatgggatcc cgaaaccaag ttgatcatgc ctggagtgta 13140

tacgtatgaa gccttccatc ctgacattgt cttactgcct ggctgcggag tggattttac 13200

cgagagtcgt ttgagcaacc ttcttggtat cagaaaaaaa cagccatttc aagagggttt 13260
taagattttg tatgaagatt tagaaggtgg taatattccg gccctcttgg atgtagatgc 13320
ctatgagaac agtaagaaag aacaaaaagc caaaatagaa gctgctacag ctgctgcaga 13380
agctaaggca aacatagttg ccagcgactc tacaagggtt gctaacgctg gagaggtcag 13440
aggagacaat tttgcgccaa cacctgttcc gactgcagaa tcattattgg ccgatgtgtc 13500
tgaaggaacg gacgtgaaac tcactattca acctgtagaa aaagatagta agaatagaag 13560
ctataatgtg ttggaagaca aaatcaacac agcctatcgc agttggtatc tttcgtacaa 13620

ttatggcgat cccgaaaaag gagtgcgttc ctggacattg ctcaccacct cagatgtcac 13680
ctgcggagca gagcaggttt actggtcgct tccagacatg atgaaggatc ctgtcacttt 13740
ccgctccact agacaagtca gtaactaccc tgtggtgggt gcagagctta tgcccgtctt 13800
ctcaaagagc ttctacaacg aacaagctgt gtactcccag cagctccgcc agtccacctc 13860
gcttacgcac gtcttcaacc gctttcctga gaaccagatt ttaatccgtc cgccggcgcc 13920
caccattacc accgtcagtg aaaacgttcc tgctctcaca gatcacggga ccctgccgtt 13980
gcgcagcagt atccggggag tccaacgtgt gaccgttact gacgccagac gccgcacctg 14040

tccctacgtg tacaaggcac tgggcatagt cgcaccgcgc gtcctttcaa gccgcacttt 14100
ctaaaaaaaa aatgtccatt cttatctcgc ccagtaataa caccggttgg ggtctgcgcg 14160
ctccaagcaa gatgtacgga ggcgcacgca aacgttctac ccaacatccc gtgcgtgttc 14220
gcggacattt tcgcgctcca tggggtgccc tcaagggccg cactcgcgtt cgaaccaccg 14280
tcgatgatgt aatcgatcag gtggttgccg acgcccgtaa ttatactcct actgcgccta 14340
catctactgt ggatgcagtt attgacagtg tagtggctga cgctcgcaac tatgctcgac 14400
gtaagagccg gcgaaggcgc attgccagac gccaccgagc taccactgcc atgcgagccg 14460

caagagctct gctacgaaga gctagacgcg tggggcgaag agccatgctt agggcggcca 14520
gacgtgcagc ttcgggcgcc agcgccggca ggtcccgcag gcaagcagcc gctgtcgcag 14580
cggcgactat tgccgacatg gcccaatcgc gaagaggcaa tgtatactgg gtgcgtgacg 14640
ctgccaccgg tcaacgtgta cccgtgcgca cccgtccccc tcgcacttag aagatactga 14700
gcagtctccg atgttgtgtc ccagcggcga ggatgtccaa gcgcaaatac aaggaagaaa 14760
tgctgcaggt tatcgcacct gaagtctacg gccaaccgtt gaaggatgaa aaaaaacccc 14820
gcaaaatcaa gcgggttaaa aaggacaaaa aagaagagga agatggcgat gatgggctgg 14880

cggagtttgt gcgcgagttt gccccacggc gacgcgtgca atggcgtggg cgcaaagttc 14940

gacatgtgtt gagacctgga acttcggtgg tctttacacc cggcgagcgt tcaagcgcta 15000

cttttaagcg ttcctatgat gaggtgtacg gggatgatga tattcttgag caggcggctg 15060

accgattagg cgagtttgct tatggcaagc gtagtagaat aacttccaag gatgagacag 15120

tgtcaatacc cttggatcat ggaaatccca cccctagtct taaaccggtc actttgcagc 15180

aagtgttacc cgtaactccg cgaacaggtg ttaaacgcga aggtgaagat ttgtatccca 15240

ctatgcaact gatggtaccc aaacgccaga agttggagga cgttttggag aaagtaaaag 15300

tggatccaga tattcaacct gaggttaaag tgagacccat taagcaggta gcgcctggtc 15360

tgggggtaca aactgtagac attaagattc ccactgaaag tatggaagtg caaactgaac 15420

ccgcaaagcc tactgccacc tccactgaag tgcaaacgga tccatggatg cccatgccta 15480

ttacaactga cgccgccggt cccactcgaa gatcccgacg aaagtacggt ccagcaagtc 15540

tgttgatgcc caattatgtt gtacacccat ctattattcc tactcctggt taccgaggca 15600

ctcgctacta tcgcagccga aacagtacct cccgccgtcg ccgcaagaca cctgcaaatc 15660

gcagtcgtcg ccgtagacgc acaagcaaac cgactcccgg cgccctggtg cggcaagtgt 15720

accgcaatgg tagtgcggaa cctttgacac tgccgcgtgc gcgttaccat ccgagtatca 15780

tcacttaatc aatgttgccg ctgcctcctt gcagatatgg ccctcacttg tcgccttcgc 15840

gttcccatca ctggttaccg aggaagaaac tcgcgccgta gaagagggat gttgggacgc 15900

ggaatgcgac gctacaggcg acggcgtgct atccgcaagc aattgcgggg tggtttttta 15960

ccagccttaa ttccaattat cgctgctgca attggcgcga taccaggcat agcttccgtg 16020

gcggttcagg cctcgcaacg acattgacat tggaaaaaaa acgtataaat aaaaaaaaat 16080

acaatggact ctgacactcc tggtcctgtg actatgtttt cttagagatg gaagacatca 16140

atttttcatc cttggctccg cgacacggca cgaagccgta catgggcacc tggagcgaca 16200

tcggcacgag ccaactgaac gggggcgcct tcaattggag cagtatctgg agcgggctta 16260

aaaattttgg ctcaaccata aaaacatacg ggaacaaagc ttggaacagc agtacaggac 16320

aggcgcttag aaataaactt aaagaccaga acttccaaca aaaagtagtc gatgggatag 16380

cttccggcat caatggagtg gtagatttgg ctaaccaggc tgtgcagaaa aagataaaca 16440

gtcgtttgga cccgccgcca gcaaccccag gtgaaatgca agtggaggaa gaaattcctc 16500

cgccagaaaa acgaggcgac aagcgtccgc gtcccgattt ggaagagacg ctggtgacgc 16560

gcgtagatga accgccttct tatgaggaag caacgaagct tggaatgccc accactagac 16620

cgatagcccc aatggccacc ggggtgatga aaccttctca gttgcatcga cccgtcacct 16680

tggatttgcc ccctcccccc gctgctactg ctgtacccgc ttctaagcct gtcgctgccc 16740

cgaaaccagt cgccgtagcc aggtcacgtc ccgggggcgc tcctcgtcca aatgcgcact 16800

ggcaaaatac tctgaacagc atcgtgggtc taggcgtgca aagtgtaaaa cgccgtcgct 16860

gcttttaatt aaatatggag tagcgcttaa cttgcctatc tgtgtatatg tgtcattaca 16920

cgccgtcaca gcagcagagg aaaaaaggaa gaggtcgtgc gtcgacgctg agttactttc 16980

aagatggcca ccccatcgat gctgccccaa tgggcataca tgcacatcgc cggacaggat 17040

gcttcggagt acctgagtcc gggtctggtg cagttcgccc gcgccacaga cacctacttc 17100

aatctgggaa ataagtttag aaatcccacc gtagcgccga cccacgatgt gaccaccgac 17160

cgtagccagc ggctcatgtt gcgcttcgtg cccgttgacc gggaggacaa tacatactct 17220

tacaaagtgc ggtacaccct ggccgtgggc gacaacagag tgctggatat ggccagcacg 17280

ttctttgaca ttaggggcgt gttggacaga ggtcccagtt tcaaacccta ttctggtacg 17340

gcttacaact ctctggctcc taaaggcgct ccaaatgcat ctcaatggat tgcaaaaggc 17400

gtaccaactg cagcagccgc aggcaatggt gaagaagaac atgaaacaga ggagaaaact 17460

gctacttaca cttttgccaa tgctcctgta aaagccgagg ctcaaattac aaaagagggc 17520

ttaccaatag gtttggagat ttcagctgaa aacgaatcta aacccatcta tgcagataaa 17580

ctttatcagc cagaacctca agtgggagat gaaacttgga ctgacctaga cggaaaaacc 17640

gaagagtatg gaggcagggc tctaaagcct actactaaca tgaaaccctg ttacgggtcc 17700

tatgcgaagc ctactaattt aaaaggtggt caggcaaaac cgaaaaactc ggaaccgtcg 17760

agtgaaaaaa ttgaatatga tattgacatg gaatttttg ataactcatc gcaaagaaca 17820

aacttcagtc ctaaaattgt catgtatgca gaaaatgtag gtttggaaac gccagacact 17880

catgtagtgt acaaacctgg aacagaagac acaagttccg aagctaattt gggacaacag 17940

tctatgccca acagacccaa ctacattggc ttcagagata actttattgg actcatgtac 18000

tataacagta ctggtaacat gggggtgctg gctggtcaag cgtctcagtt aaatgcagtg 18060

gttgacttgc aggacagaaa cacagaactt tcttaccaac tcttgcttga ctctctgggc 18120

gacagaacca gatactttag catgtggaat caggctgtgg acagttatga tcctgatgta 18180

cgtgttattg aaaatcatgg tgtggaagat gaacttccca actattgttt tccactggac 18240

ggcataggtg ttccaacaac cagttacaaa tcaatagttc caaatggaga agataataat 18300

aattggaaag aacctgaagt aaatggaaca agtgagatcg gacagggtaa tttgtttgcc 18360

atggaaatta accttcaagc caatctatgg cgaagtttcc tttattccaa tgtggctctg 18420

tatctcccag actcgtacaa atacaccccg tccaatgtca ctcttccaga aaacaaaaac 18480

acctacgact acatgaacgg gcgggtggtg ccgccatctc tagtagacac ctatgtgaac 18540

attggtgcca ggtggtctct ggatgccatg gacaatgtca acccattcaa ccaccaccgt 18600

aacgctggct tgcgttaccg atctatgctt ctgggtaacg gacgttatgt gcctttccac 18660

atacaagtgc ctcaaaaatt cttcgctgtt aaaaacctgc tgcttctccc aggctcctac 18720
acttatgagt ggaactttag gaaggatgtg aacatggttc tacagagttc cctcggtaac 18780
gacctgcggg tagatggcgc cagcatcagt ttcacgagca tcaacctcta tgctactttt 18840
ttccccatgg ctcacaacac cgcttccacc cttgaagcca tgctgcggaa tgacaccaat 18900
gatcagtcat tcaacgacta cctatctgca gctaacatgc tctaccccat tcctgccaat 18960
gcaaccaata ttcccatttc cattccttct cgcaactggg cggctttcag aggctggtca 19020
tttaccagac tgaaaaccaa agaaactccc tctttggggt ctggatttga cccctacttt 19080

gtctattctg gttctattcc ctacctggat ggtaccttct acctgaacca cacttttaag 19140
aaggtttcca tcatgtttga ctcttcagtg agctggcctg gaaatgacag gttactatct 19200
cctaacgaat ttgaaataaa gcgcactgtg gatggcgaag gctacaacgt agcccaatgc 19260
aacatgacca aagactggtt cttggtacag atgctcgcca actacaacat cggctatcag 19320
ggcttctaca ttccagaagg atacaaagat cgcatgtatt catttttcag aaacttccag 19380
cccatgagca ggcaggtggt tgatgaggtc aattacaaag acttcaaggc cgtcgccata 19440
ccctaccaac acaacaactc tggctttgtg ggttacatgg ctccgaccat gcgccaaggt 19500

caaccctatc ccgctaacta tccctatcca ctcattggaa caactgccgt aaatagtgtt 19560
acgcagaaaa agttcttgtg tgacagaacc atgtggcgca taccgttctc gagcaacttc 19620
atgtctatgg gggcccttac agacttggga cagaatatgc tctatgccaa ctcagctcat 19680
gctctggaca tgacctttga ggtggatccc atggatgagc ccaccctgct ttatcttctc 19740
ttcgaagttt tcgacgtggt cagagtgcat cagccacacc gcggcatcat cgaggcagtc 19800
tacctgcgta caccgttctc ggccggtaac gctaccacgt aagaagcttc ttgcttcttg 19860
caaatagcag ctgcaaccat ggcctgcgga tcccaaaacg gctccagcga gcaagagctc 19920

agagccattg tccaagacct gggttgcgga ccctattttt tgggaaccta cgataagcgc 19980
ttcccggggt tcatggcccc cgataagctc gcctgtgcca ttgtaaatac ggccggacgt 20040
gagacggggg gagagcactg gttggctttc ggttggaacc cacgttctaa cacctgctac 20100
ctttttgatc cttttggatt ctcggatgat cgtctcaaac agatttacca gtttgaatat 20160
gagggtctcc tgcgccgcag cgctcttgct accaaggacc gctgtattac gctggaaaaa 20220
tctacccaga ccgtgcaggg cccccgttct gccgcctgcg gacttttctg ctgcatgttc 20280
cttcacgcct ttgtgcactg gcctgaccgt cccatggacg gaaaccccac catgaaattg 20340

ctaactggag tgccaaacaa catgcttcat tctcctaaag tccagcccac cctgtgtgac 20400

aatcaaaaag cactctacca ttttcttaat acccattcgc cttattttcg ctctcatcgt 20460

acacacatcg aaagggccac tgcgttcgac cgtatggatg ttcaataatg actcatgtaa 20520

acaacgtgtt caataaacat cactttattt ttttacatgt atcaaggctc tggattactt 20580

atttatttac aagtcgaatg ggttctgacg agaatcagaa tgacccgcag gcagtgatac 20640

gttgcggaac tgatacttgg gttgccactt gaattcggga atcaccaact tgggaaccgg 20700

tatatcgggc aggatgtcac tccacagctt tctggtcagc tgcaaagctc caagcaggtc 20760

aggagccgaa atcttgaaat cacaattagg accagtgctc tgagcgcgag agttgcggta 20820

caccggattg cagcactgaa acaccatcag cgacggatgt ctcacgcttg ccagcacggt 20880

gggatctgca atcatgccca catccagatc ttcagcattg gcaatgctga acggggtcat 20940

cttgcaggtc tgcctaccca tggcgggcac ccaattaggc ttgtggttgc aatcgcagtg 21000

caggggggatc agtatcatct tggcctgatc ctgtctgatt cctggataca cggctctcat 21060

gaaagcatca tattgcttga aagcctgctg ggctttacta ccctcggtat aaaacatccc 21120

gcaggacctg ctcgaaaact ggttagctgc acagccggca tcattcacac agcagcgggc 21180

gtcattgttg gctatttgca ccacacttct gccccagcgg ttttgggtga ttttggttcg 21240

ctcgggattc tcctttaagg ctcgttgtcc gttctcgctg gccacatcca tctcgataat 21300

ctgctccttc tgaatcataa tattgccatg caggcacttc agcttgccct cataatcatt 21360

gcagccatga ggccacaacg cacagcctgt acattcccaa ttatggtggg cgatctgaga 21420

aaaagaatgt atcattccct gcagaaatct tcccatcatc gtgctcagtg tcttgtgact 21480

agtgaaagtt aactggatgc ctcggtgctc ttcgtttacg tactggtgac agatgcgctt 21540

gtattgttcg tgttgctcag gcattagttt aaaacaggtt ctaagttcgt tatccagcct 21600

gtacttctcc atcagcagac acatcacttc catgcctttc tcccaagcag acaccagggg 21660

caagctaatc ggattcttaa cagtgcaggc agcagctcct ttagccagag ggtcatcttt 21720

agcgatcttc tcaatgcttc ttttgccatc cttctcaacg atgcgcacgg gcgggtagct 21780

gaaacccact gctacaagtt gcgcctcttc tctttcttct tcgctgtctt gactgatgtc 21840

ttgcatgggg atatgtttgg tcttccttgg cttctttttg gggggtatcg gaggaggagg 21900

actgtcgctc cgttccggag acagggagga ttgtgacgtt tcgctcacca ttaccaactg 21960

actgtcggta gaagaacctg accccacacg gcgacaggtg tttttcttcg ggggcagagg 22020

tggaggcgat tgcgaagggc tgcggtccga cctggaaggc ggatgactgg cagaacccct 22080

tccgcgttcg ggggtgtgct ccctgtggcg gtcgcttaac tgatttcctt cgcggctggc 22140

cattgtgttc tcctaggcag agaaacaaca gacatggaaa ctcagccatt gctgtcaaca 22200

tcgccacgag tgccatcaca tctcgtcctc agcgacgagg aaaaggagca gagcttaagc 22260

attccaccgc ccagtcctgc caccacctct accctagaag ataaggaggt cgacgcatct 22320

catgacatgc agaataaaaa agcgaaagag tctgagacag acatcgagca agacccgggc 22380

tatgtgacac cggtggaaca cgaggaagag ttgaaacgct ttctagagag agaggatgaa 22440

aactgcccaa aacagcgagc agataactat caccaagatg ctggaaatag ggatcagaac 22500

accgactacc tcatagggct tgacggggaa gacgcgctcc ttaaacatct agcaagacag 22560

tcgctcatag tcaaggatgc attattggac agaactgaag tgcccatcag tgtggaagag 22620

ctcagctgcg cctacgagct taacctttt tcacctcgta ctccccccaa acgtcagcca 22680

aacggcacct gcgagccaaa tcctcgctta aacttttatc cagcttttgc tgtgccagaa 22740

gtactggcta cctatcacat ctttttaaa aatcaaaaaa ttccagtctc ctgccgcgct 22800

aatcgcaccc gcgccgatgc cctactcaat ctgggacctg gttcacgctt acctgatata 22860

gcttccttgg aagaggttcc aaagatcttc gagggtctgg gcaataatga gactcgggcc 22920

gcaaatgctc tgcaaaaggg agaaaatggc atggatgagc atcacagcgt tctggtggaa 22980

ttggaaggcg ataatgccag actcgcagta ctcaagcgaa gcgtcgaggt cacacacttc 23040

gcatatcccg ctgtcaacct gccccctaaa gtcatgacgg cggtcatgga ccagttactc 23100

attaagcgcg caagtcccct ttcagaagac atgcatgacc cagatgcctg tgatgagggt 23160

aaaccagtgg tcagtgatga gcagctaacc cgatggctgg gcaccgactc tccccgggat 23220

ttggaagagc gtcgcaagct tatgatggcc gtggtgctgg ttaccgtaga actagagtgt 23280

ctccgacgtt tctttaccga ttcagaaacc ttgcgcaaac tcgaagagaa tctgcactac 23340

acttttagac acggctttgt gcggcaggca tgcaagatat ctaacgtgga actcaccaac 23400

ctggtttcct acatgggtat tctgcatgag aatcgcctag gacaaagcgt gctgcacagc 23460

acccttaagg gggaagcccg ccgtgattac atccgcgatt gtgtctatct ctacctgtgc 23520

cacacgtggc aaaccggcat gggtgtatgg cagcaatgtt tagaagaaca gaacttgaaa 23580

gagcttgaca agctcttaca gaaatctctt aaggttctgt ggacagggtt cgacgagcgc 23640

accgtcgctt ccgacctggc agacctcatc ttcccagagc gtctcagggt tactttgcga 23700

aacggattgc ctgactttat gagccagagc atgcttaaca attttcgctc tttcatcctg 23760

gaacgctccg gtatcctgcc cgccacctgc tgcgcactgc cctccgactt tgtgcctctc 23820

acctaccgcg agtgcccccc gccgctatgg agtcactgct acctgttccg tctggccaac 23880

tatctctcct accactcgga tgtgatcgag gatgtgagcg gagacggctt gctggagtgc 23940

cactgccgct gcaatctgtg cacgccccac cggtccctag cttgcaaccc ccagttgatg 24000

agcgaaaccc agataatagg caccttttgaa ttgcaaggcc ccagcagcca aggcgatggg 24060

tcttctcctg ggcaaagttt aaaactgacc ccgggactgt ggacctccgc ctacttgcgc 24120

aagtttgctc cggaagatta ccacccctat gaaatcaagt tctatgagga ccaatcacag 24180
cctccaaagg ccgaactttc ggcttgcgtc atcacccagg gggcaattct ggcccaattg 24240
caagccatcc aaaaatcccg ccaagaattt ctactgaaaa agggtaaggg ggtctacctt 24300
gacccccaga ccggcgagga actcaacaca aggttccctc aggatgtccc aacgacgaga 24360
aaacaagaag ttgaaggtgc agccgccgcc cccagaagat atggaggaag attgggacag 24420
tcaggcagag gaggcggagg aggacagtct ggaggacagt ctggaggaag acagtttgga 24480
ggaggaaaac gaggaggcag aggaggtgga agaagtaacc gccgacaaac agttatcctc 24540

ggctgcggag acaagcaaca gcgctaccat ctccgctccg agtcgaggaa cccggcggcg 24600
tcccagcagt agatgggacg agaccggacg cttcccgaac ccaaccagcg cttccaagac 24660
cggtaagaag gatcggcagg gatacaagtc ctggcggggg cataagaatg ccatcatctc 24720
ctgcttgcat gagtgcgggg gcaacatatc cttcacgcgg cgctacttgc tattccacca 24780
tggggtgaac tttccgcgca atgttttgca ttactaccgt cacctccaca gcccctacta 24840
tagccagcaa atcccgacag tctcgacaga taaagacagc ggcggcgacc tccaacagaa 24900
aaccagcagc ggcagttaga aaatacacaa caagtgcagc aacaggagga ttaaagatta 24960

cagccaacga gccagcgcaa acccgagagt taagaaatcg gatctttcca accctgtatg 25020
ccatcttcca gcagagtcgg ggtcaagagc aggaactgaa aataaaaaac cgatctctgc 25080
gttcgctcac cagaagttgt ttgtatcaca agagcgaaga tcaacttcag cgcactctcg 25140
aggacgccga ggctctcttc aacaagtact gcgcgctgac tcttaaagag taggcagcga 25200
ccgcgcttat tcaaaaaagg cgggaattac atcatcctcg acatgagtaa agaaattccc 25260
acgccttaca tgtggagtta tcaaccccaa atgggattgg cagcaggcgc ctcccaggac 25320
tactccaccc gcatgaattg gctcagcgcc gggccttcta tgatttctcg agttaatgat 25380

atacgcgcct accgaaacca aatacttttg gaacagtcag ctcttaccac cacgccccgc 25440
caacacctta atcccagaaa ttggcccgcc gccctagtgt accaggaaag tcccgctccc 25500
accactgtat tacttcctcg agacgcccag gccgaagtcc aaatgactaa tgcaggtgcg 25560
cagttagctg gcggctccac cctatgtcgt cacaggcctc ggcataatat aaaacgcctg 25620
atgatcagag gccgaggtat ccagctcaac gacgagtcgg tgagctctcc gcttggtcta 25680
cgaccagacg gaatctttca gattgccggc tgcgggagat cttccttcac ccctcgtcag 25740
gctgttctga ctttggaaag ttcgtcttcg caaccccgct cgggcggaat cgggaccgtt 25800

caatttgtag aggagtttac tccctctgtc tacttcaacc ccttctccgg atctcctggg 25860

cactacccgg acgagttcat accgaacttc gacgcgatta gcgagtcagt ggacggctac 25920

gattgatgtc tggtgacgcg gctgagctat ctcggctgcg acatctagac cactgccgcc 25980

gctttcgctg ctttgcccgg gaacttattg agttcatcta cttcgaactc cccaaggatc 26040

accctcaagg tccggcccac ggagtgcgga ttactatcga aggcaaaata gactctcgcc 26100

tgcaacgaat tttctcccag cggcccgtgc tgatcgagcg agaccaggga aacaccacgg 26160

ttagtaatca attacggggt cattagttca tagcccatat atggagttaa cttgtttatt 26220

gcagcttata atggttacaa ataaagcaat agcatcacaa atttcacaaa taaagcattt 26280

ttttcactgc attctagttg tggtttgtcc aaactcatca atgtatctta tcatgtctgc 26340

tcgaagcggc cggccgcccc gactctagag tcgcggcctc attaggaagt tcctatactt 26400

tctagagaat aggaacttct cagaagaact cgtcaagaag gcgatagaag gcgatgcgct 26460

gcgaatcggg agcggcgata ccgtaaagca cgaggaagcg gtcagcccat tcgccgccaa 26520

gctcttcagc aatatcacgg gtagccaacg ctatgtcctg atagcggtcc gccacaccca 26580

gccggccaca gtcgatgaat ccagaaaagc ggccattttc caccatgata ttcggcaagc 26640

aggcatcgcc atgggtcacg acgagatcct cgccgtcggg catgcgcgcc ttgagcctgg 26700

cgaacagttc ggctggcgcg agcccctgat gctcttcgtc cagatcatcc tgatcgacaa 26760

gaccggcttc catccgagta cgtgctcgct cgatgcgatg tttcgcttgg tggtcgaatg 26820

ggcaggtagc cggatcaagc gtatgcagcc gccgcattgc atcagccatg atggatactt 26880

tctcggcagg agcaaggtga gatgacagga gatcctgccc cggcacttcg cccaatagca 26940

gccagtccct tcccgcttca gtgacaacgt cgagcacagc tgcgcaagga acgcccgtcg 27000

tggccagcca cgatagccgc gctgcctcgt cctgcagttc attcagggca ccggacaggt 27060

cggtcttgac aaaaagaacc gggcgcccct gcgctgacag ccggaacacg gcggcatcag 27120

agcagccgat tgcctgttgt gcccagtcat agccgaatag cctctccacc caagcggccg 27180

gagaacctgc gtgcaatcca tcttgttcaa tggccgatcc cataacaccc cttgtattac 27240

tgtttatgta agcagacagt tttactgttc gtgatgatat atttttatct tgtgcaatgt 27300

aacaggttgt ggccatagcg ggcccgggat tttcctccac gtccccgcat gttagaagac 27360

ttcccctgcc ctcggctctg gaagttccta tactttctag agaataggaa cttcccgcca 27420

gaatgcgttc gcacagccgc cagccggtca ctccgttgat ggttactcgg aacagcaggg 27480

agccgtcggg gttgatcagg cgctcgtcga taattttgtt gccgttccac agggtccctg 27540

ttacagtgat ctttttgccg tcgaacacgg cgatgccttc atacggccgt ccgaaatagt 27600

cgatcatgtt cggcgtaacc ccgtcgatta ccagtgtgcc atagtgcagg atcaccttaa 27660

agtgatgatc atccacaggg tacaccacct taaaaatttt ttcgatctgg cccatttggt 27720

cgccgctcag accttcatac gggatgatga catggatgtc gatcttcagc ccattttcac 27780

cgctcaggac aatcctttgg atcggagtta cggacacccc gagattctga aacaaactgg 27840

acacacctcc ctgttcaagg acttggtcca ggttgtagcc ggctgtctgt cgccagtccc 27900

caacgaaatc ttcgagtgtg aagaccatgg atccgggccc ggggttttct tcaacgtctc 27960

cagcctgctt cagcaggctg aagttagtag ctccgcttcc tcgagctcga gatctggcga 28020

aggcgatggg ggtcttgaag gcgtgctggt actccacgat gcccagctcg gtgttgctgt 28080

gcagctcctc cacgcggcgg aaggcgaaca tggggcccccc gttctgcagg atgctggggt 28140

ggatggcgct cttgaagtgc atgtggctgt ccaccacgaa gctgtagtag ccgccgtcgc 28200

gcaggctgaa ggtgcgggcg aagctgccca ccagcacgtt atcgcccatg gggtgcaggt 28260

gctccacggt ggcgttgctg cggatgatct tgtcggtgaa gatcacgctg tcctcgggga 28320

agccggtgcc caccaccttg aagtcgccga tcacgcggcc ggcctcgtag cggtagctga 28380

agctcacgtg cagcacgccg ccgtcctcgt acttctcgat gcgggtgttg gtgtagccgc 28440

cgttgttgat ggcgtgcagg aaggggttct cgtagccgct ggggtaggtg ccgaagtggt 28500

agaagccgta gcccatcacg tggctcagca ggtaggggct gaaggtcagg gcgcctttgg 28560

tgctcttcat cttgttggtc atgcggccct gctcgggggt gccctctccg ccgcccacca 28620

gctcgaactc cacgccgttc agggtgccgg tgatgcggca ctcgatcttc atggcgggca 28680

tggtggcgac cggtagcgct agcggcttcg gtaccacgcg ttcgctcgaa ttaatcaatt 28740

ctttgccaaa atgatgagac agcacaataa ccagcacgtt gcccaggagc tgtaggaaaa 28800

agaagaaggc atgaacatgg ttagcagagg ctctagagcc gccggtcaca cgccagaagc 28860

cgaaccccgc cctgccccgt ccccccgaa ggcagccgtc cccccgcgga cagccccgag 28920

gctggagagg gagaagggga cggcggcgcg gcgacgcacg aaggccctcc ccgcccattt 28980

ccttcctgcc ggggccctcc cggagcccct caaggctttc acgcagccac agaaaagaaa 29040

caagccgtca ttaaaccaag cgctaattac agcccggagg agaagggccg tcccgcccgc 29100

tcacctgtgg gagtaacgcg gtcagtcaga gccggggcgg gcggcgcgag gcggcgcgga 29160

gcggggcacg gggcgaaggc aacgcagcga ctcccgcccg ccgcgcgctt cgctttttat 29220

agggccgccg ccgccgccgc ctcgccataa aaggaaactt tcggagcgcg ccgctctgat 29280

tggctgccgc cgcacctctc cgcctcgccc cgccccgccc ctcgccccgc cccgccccgc 29340

ctggcgcgcg ccccccccccc ccccccgccc ccatcgctgc acaaaataat taaaaaataa 29400

ataaatacaa aattgggggt ggggaggggg gggagatggg gagagtgaag cagaacgtgg 29460

ggctcacctc gaccatggta atagcgatga ctaatacgta gatgtactgc caagtaggaa 29520

agtcccataa ggtcatgtac tgggcataat gccaggcggg ccatttaccg tcattgacgt 29580

caataggggg cgtacttggc atatgataca cttgatgtac tgccaagtgg gcagtttacc 29640
gtaaatactc cacccattga cgtcaatgga aagtccctat tggcgttact atgggaacat 29700
acgtcattat tgacgtcaat gggcgggggt cgttgggcgg tcagccaggc gggccattta 29760
ccgtaagtta tgtaacgcgg aacaacaaca attgcattca ttttatgttt caggttcagg 29820
gggaggtgtg gtcctgcgat tccatcgagt gcacctacac cctgctgaag accctatgcg 29880
gcctaagaga cctgctacca atgaattaaa aaaaaatgat taataaaaaa tcacttactt 29940
gaaatcagca ataaggtctc tgttgaaatt ttctcccagc agcacctcac ttccctcttc 30000

ccaactctgg tattctaaac cccgttcagc ggcatacttt ctccatactt taaaggggat 30060
gtcaaatttt agctcctctc ctgtacccac aatcttcatg tctttcttcc cagatgacca 30120
agagagtccg gctcagtgac tccttcaacc ctgtctaccc ctatgaagat gaaagcacct 30180
cccaacaccc ctttataaac ccagggttta tttccccaaa tggcttcaca caaagcccag 30240
acggagttct tactttaaaa tgtttaaccc cactaacaac cacaggcgga tctctacagc 30300
taaaagtggg agggggactt acagtggatg acactgatgg taccttacaa gaaaacatac 30360
gtgctacagc acccattact aaaaataatc actctgtaga actatccatt ggaaatggat 30420

tagaaactca aaacaataaa ctatgtgcca aattgggaaa tgggttaaaa tttaacaacg 30480
gtgacatttg tataaaggat agtattaaca ccttatggac tggaataaac cctccaccta 30540
actgtcaaat tgtggaaaac actaatacaa atgatggcaa acttacttta gtattagtaa 30600
aaaatggagg gcttgttaat ggctacgtgt ctctagttgg tgtatcagac actgtgaacc 30660
aaatgttcac acaaaagaca gcaaacatcc aattaagatt atattttgac tcttctggaa 30720
atctattaac tgaggaatca gacttaaaaa ttccacttaa aaataaatct tctacagcga 30780
ccagtgaaac tgtagccagc agcaaagcct ttatgccaag tactacagct tatcccttca 30840

acaccactac tagggatagt gaaaactaca ttcatggaat atgttactac atgactagtt 30900
atgatagaag tctatttccc ttgaacattt ctataatgct aaacagccgt atgatttctt 30960
ccaatgttgc ctatgccata caatttgaat ggaatctaaa tgcaagtgaa tctccagaaa 31020
gcaacatagc tacgctgacc acatcccccc tttttctttc ttacattaca gaagacgaca 31080
actaaaataa agtttaagtg tttttattta aaatcacaaa attcgagtag ttattttgcc 31140
tccaccttcc catttgacag aatacacagt cctttctccc cggctggcct taaaaagcat 31200
catatcatgg gtaacagaca tattcttagg tgttatattc cacacggttt cctgtcgagc 31260

caaacgctca tcagtgatat taataaactc cccgggcagc tcacttaagt tcatgtcgct 31320

gtccagctgc tgagccacag gctgctgtcc aacttgcggt tgcttaacgg gcggcgaagg 31380

agaagtccac gcctacatgg gggtagagtc ataatcgtgc atcaggatag ggcggtggtg 31440

ctgcagcagc gcgcgaataa actgctgccg ccgccgctcc gtcctgcagg aatacaacat 31500

ggcagtggtc tcctcagcga tgattcgcac cgcccgcagc ataaggcgcc ttgtcctccg 31560

ggcacagcag cgcaccctga tctcacttaa atcagcacag taactgcagc acagcaccac 31620

aatattgttc aaaatcccac agtgcaaggc gctgtatcca aagctcatgg cggggaccac 31680

agaacccacg tggccatcat accacaagcg caggtagatt aagtggcgac ccctcataaa 31740

cacgctggac ataaacatta cctcttttgg catgttgtaa ttcaccacct cccggtacca 31800

tataaacctc tgattaaaca tggcgccatc caccaccatc ctaaaccagc tggccaaaac 31860

ctgcccgccg gctatacact gcagggaacc gggactggaa caatgacagt ggagagccca 31920

ggactcgtaa ccatggatca tcatgctcgt catgatatca atgttggcac aacacaggca 31980

cacgtgcata cacttcctca ggattacaag ctcctcccgc gttagaacca tatcccaggg 32040

aacaacccat tcctgaatca gcgtaaatcc cacactgcag ggaagacctc gcacgtaact 32100

cacgttgtgc attgtcaaag tgttacattc gggcagcagc ggatgatcct ccagtatggt 32160

agcgcgggtt tctgtctcaa aaggaggtag acgatcccta ctgtacggag tgcgccgaga 32220

caaccgagat cgtgttggtc gtagtgtcat gccaaatgga acgccggacg tagtcattct 32280

cgtattttgt atagcaaaac gcggccctgg cagaacacac tcttcttcgc cttctatcct 32340

gccgcttagc gtgttccgtg tgatagttca agtacagcca cactcttaag ttggtcaaaa 32400

gaatgctggc ttcagttgta atcaaaactc catcgcatct aattgttctg aggaaatcat 32460

ccacggtagc atatgcaaat cccaaccaag caatgcaact ggattgcgtt tcaagcagga 32520

gaggagaggg aagagacgga agaaccatgt taatttttat tccaaacgat ctcgcagtac 32580

ttcaaattgt agatcgcgca gatggcatct ctcgccccca ctgtgttggt gaaaaagcac 32640

agctaaatca aaagaaatgc gattttcaag gtgctcaacg gtggcttcca acaaagcctc 32700

cacgcgcaca tccaagaaca aaagaatacc aaaagaagga gcattttcta actcctcaat 32760

catcata 32767

<210> 54

<211> 30493

<212> DNA

<213> Artificial Sequence

<220><223> pHCA(Ad35)

<400> 54

ggccgaagga ttacatgagc ttagaaatgt aattagcata gtgtgtggca tagtgtagat 60

accaaataaa tatgatctct ccttctactc ttgaaaatgc aaacacattc ttggtggtcc 120

taaaatagcc tgtaacatgg tttactcagc agcatttgct attcaaggca gatctgcctt 180

tagtcattgg ctgcgctcct gaacagctgt gtgaaaggct aacttttgta aaccaaatca 240

aaataaaatg cagcaaaaat ttgtcactga aaggaaatcc tcagtatatc cttttatgaa 300

atgaaagatc cctcatccaa acttaacttt tttaaaagtg cgcatttgga gatatagccc 360

tttcttatga atcctaattc aattttggcc ataaacacac gttgatgttc cccaccccaa 420

agcacatagc aacaagagta ggttctatat tgaaaataat gacaatttaa aaacatgtac 480

ttatttcact gtatgtggac agtgtctatg attgcatcat gaagtgtcat ataaccatgt 540

acgtgtacat gagagagaga tagagagaga agtggtaggg tggtggtggt agaggggatg 600

gcgatagtaa tcatggtaat ggtagaggtg atggaggtgg taatgacgga ggtaagggtg 660

gtagtgatga tggtggtggt ggtaatggtg gtggatgtgg tggtggcaat tgggatggtg 720

ggatggtggt agccatggtg atggtggtaa tggtgttgat ttaaagggtg gtggtagtga 780

aggtgagggt agtggtggtg gaggtggtgg tgctggtagc aatagtgatg gtggtgatgg 840

tgttgatgag ggtgttggga tcagggtgag ttcccacagt atatttcatt cttgttgtac 900

cactctgtca acagcaccac tgactgggac agaggaagaa ggcacactct gaatgtgtta 960

ttaacagaaa cctcaaaaca gtctgtctcc ttgtagtcat tcaaaattat ctttttctta 1020

cctggaaaac tgaaactgaa ttaccgggaa aaacacagga gatttttgtt tgttaatatg 1080

ctgccaataa agtaatttta tgtcaaattt aactacagga aagggcaagg cattttctaa 1140

gttccttaga tgtcatgtgg ctaaaaaaaa caaaaggatg gacagcagtt agatactgta 1200

cacttagctg tttgaagcca tatattcaga aagcagatgt tgggagttgg tgtttgagga 1260

ctgatttcct ggaggtattt tatataggcc aagttcattg ttctaaactc taagggcttg 1320

acttgaggga ggaaaagagg caagaacatg tttagttttg ctgacagcat cacatgggca 1380

gccctaaggc tagacaactt tagggcctga agcttattct aggaaagaag cacctacaga 1440

gtggcactgg gctcccctcc actatagaga tgaagtcata tgacagtaaa gggcaggcag 1500

ggctgcctag ggggcccaga actgacactt ccattagaat gagcacaggc cagggagaga 1560

agtggggaac cagagagaag gagctggaat tctagtagga caaacggtaa gtgaacaaca 1620

agaacaagtt aagagtgtgt gcagtattct ttcaaagact gaaaaaatag tgatgtgata 1680

gaatggcagg tggctctgag caggccagga gaaggactgg gggcagagca tcccaggcag 1740

gagggcagca agtgggaagg ccctggggtg gggcttttgg actgttccag tgacgggcag 1800

gcagccagtg tgcctgtcac acaatgcacc agggaagtag tcgtgaattt gcagagggtc 1860

ttgcaggcta tgggaaaggg attggattgt attttgtttg tagggaagcc atcgggggac 1920

ttaagcagag gaaggattgg cttcatctct ttgaaaaagt tctctctgga tgctgatggg 1980

aggagaaatg gaaggaaaag aaacactttt aggggcaaga acttttgaga agggtggaat 2040

tgggagtgtg gagttggggc cagctttggc acaggagggg aagctaaaca cgtggccgca 2100

tgagggcctg taattctacc tgaaatgggt accatttgtt agggtaaaca aatgaaccaa 2160

atgcccagtg atacagacca agtgttggca aacttcttct gtgatggccc aggtagtaaa 2220

tgtctcaggc ttcgcaggcc atgtggtctc tgttgaagct ctgtgtagta gacaatatgt 2280

taatgactgg gcgtgactgt gtgctaataa aagtttattt acaaaaacag cccgtgggct 2340

ggatttagct cacaggctgt agtttgccaa cctctgacct agagcatgaa ctgagcatct 2400

tcttggaggg aaataagttc tttccaagtt gccctcctca cattgcaggg ggccatgtag 2460

gcccattatt cacagaagag tgggtgggca acctttctgg agcagaaaaa cgtaaagatt 2520

tcttccgtag tgcaagtaag gtgaccattt ctaaaccgtg caagtgatcc agcagtccca 2580

aaagttgttt cacttctcat tgtgcgcccg ttctcaggtg ctccgaagct tccagtcctt 2640

tgtagggaca tggatgaaat tggaaatcat cattctcagt aaactatcgc aagaacaaaa 2700

aaccaaacac cgcatattct cactcatagg tgggaattga acaatgagat cacatggaca 2760

cgggaagggg aacatcacat tctggggact gttgtggggt ggggggaggg gggagggata 2820

gcattgggag atatacctaa tgctagatga caagttagtg ggtgcagcgc accagtgtgg 2880

cacatgtata catatgtaac taacctgcac aatgtgcaca tgtaccctaa aacttaaagt 2940

ataataataa aagaaaaaaa aaaagagagg agagaaacat catcccctcc aggataccct 3000

tgggccttgt tcttatagtc ttgtacattg ttgaacaatt tgcatgggct agtggattaa 3060

agcacaccct ccaccctcag gccctcaagg gtctctatga taatacagtc tcaccttcta 3120

ccctttccat caccatccta ggtgctatgg ccaaccttga ggctgccatg ttaggtctat 3180

gcatttccca cctccaccac ataactctct gaaggccagg tagtttccta ttcatcttgg 3240

taaccccaaa gcctcgtgac agggctcagc tggcatctgc ggatgtgaat gaaccattgg 3300

agaaaatggt actctgcaaa taactctgtt attttcccat ttcctgtgta aggcctagag 3360

acaatgactt tttaattgca ccccttcccc tctgtatgac actggccttc tcttgtgtcc 3420

agcaatgtgg gtggcctaga tgatttctaa gggacttctg gccaagatga acagcagctg 3480

catcttactg agcatttact atgtgccata tactcagcca cagctctagg ggcatagaag 3540

caggagctct cagggtcagg gcagtgagtg agcaagcgag cacctatgcc agccctgcct 3600

ctggatgggg acttgagagg gtgatggaag cctgcagcac tggagggagg cagacaaaga 3660

caggcctgtg ctgaggggggc ccggagcaag agagagggag gcaatgacag cagagacatg 3720

cctgcgcctt gggtttgagt gcccagtggt caaatccact tccctgtggc tgatgcttgc 3780

ctttctaact ttggaattta ggggttggag atctggtgag aaggtaggag ggagatgagg 3840

aggagaaggg aaaggcagga aggaaggggga gggaaaggaa aagcaaaagg ggaggaggaa 3900

ggtttccaac aaattattct atatcaactg cggaaatcaa aatttgttgc ccaaatctta 3960

gaagctcatg tccctcctcc ccagaagtct ggaatgcagc actccagggg tagcttataa 4020

cccaaatatc tatctgtaaa aagagaaaca ttgggctttc gagctgtgga ttctcagtaa 4080

aagcaagagg cctcagccta cacaggccag cccagagttt gaggaacccc aggcccacac 4140

ccacagggct ggcccctggg tctgcatact ccctagaaat gtgcacactt ctgagcctca 4200

actctgtcct ggagtctaac agcatccctc tccttcctgg ggcagttcca cctccagaaa 4260

cctgttacct tgggccttat gtcaaggaaa ctgtgggaaa gagctaggca ggaatgcaga 4320

tgaggccagc atgggctcct aaaagtttag aaataggcag tgtcatgctc ccaggtgcct 4380

gcataaacca gctgaaaaat ggagctcccc tcaccagcac tctcccttca aacagactgt 4440

gatttgcagg tcactggttt accaagccag gctacccagg caggacccag atgccaagcc 4500

cagtggtgtc ctgcaagctg agcagtgctc agttcttgca aaaaaaggtc tgtgtgaagg 4560

caaggcctct gcctggcttc tcaccccagt tgggtgtctg gaacaggaag gagcccttac 4620

tgcagaaaaa ggaggaggga gcaaagggag cgaacagctg cgtgctccat ggggaggatc 4680

cccaaagtag aaaggcgcat acacactgca gcccttgacc cagaatgctc acagctacat 4740

tacagattca ggtctcctca gtgtagtggg gctgctgatg agactgtggc atcctcaggg 4800

gtcaggacac acattttcca tcactcttct gatggcaaaa aacctctgag ccaatgccaa 4860

cctctgatca ttaaaaaaaa gtgctcacag cagtgtgtgg tttaggatca tgccctgtgt 4920

ggtttggaac acgtgcacaa ccacaccttg ttcatcacca tcccagaaac cctgacgcag 4980

gcaaagagca gagttattaa ccctacttta ctgatgtgga tactgaggcc cagaggctca 5040

tgcaagttat caataagtgg cagggacagt tgcctctaga ttaactagcc cctaggatca 5100

cctgggtctt ggaaggggac ccataaacat gagctcccct ctcttggggc cagatttgca 5160

cctgtgccgc gccttcagcc tgcatgaagt aggggctgct ggcaaagact caaagctgta 5220

aatctgggtt ttctcttgag gcttctaagg gagctgtttc gacaactcac tctgttccca 5280

gctggctgcc cctgcatagg gttttaaagc agcctagctt tctgccaggc ttggcagtgg 5340

acaacgctgg tcagaacatc ccagagagct accagaatga agtaagtttg cttctactct 5400

ttacctgttt atgggctgtc tctgccactg gaatgaaagg cactgagaac agtgcctggc 5460

ctgcagaagg ccctggaaat acctgagctc ctaatctggg aataggagta ggaagagctt 5520

tggaggcagg gcacctgagt ttgagatcta caacttcctg cctgtgtgac attgggaaag 5580
tctccatcct ttctgagcct cagtctccac cctggggaag tggaaatatc aatctctgtg 5640
acacagaagc aaatgagcga atgtgcacaa agtaccttgc acaagagaga cgctcaaaca 5700
cttgcctcca ggtttcaccg agaactacag agtaagatag atttgttccc agtggaggaa 5760
gcctgggaat aatttgcccc tagactatga attcctgggg ctcaagatcg agcacagggc 5820
caggcacaca gaagggaccc tggaaatgtg gcaggaggcc agagatagac aggcccttag 5880
agctcatacc catgccctct gacctcaaga agaaagaaac ctgctcaaaa tctcacaaag 5940

agcttgttcc aaccctgaat cgagtctgag gactccttcc tgagtccagc acttttctg 6000
caagaagtat atgcctccaa agctgatggg cgcaaatctt gaaccccgtc acataaacac 6060
aaagggagga ggtgactaga gctcctccta ctggatatgt ctaaggtcac cagtctaaag 6120
aaaagggatg gatagaatga ggccagtatt tttgcagcca tccaaatgtc cacatacgct 6180
gttacactga gggctcctct ctcccccgtc ttcagcccta cttgcattta gaggtgagaa 6240
agatatgggc tgaggggttg tttttcatcg tattgtagat ggaaagcaca ctgcccttgg 6300
ggccatccaa atgtggacct tgatgtagca ccccaccttc tggatggcca tccttctgaa 6360

agtcactgaa tttctcagac tttattctct ttatccataa agaaggagaa taataataat 6420
ccccccaccc tgcccaacca ctgactggtt gggaagctca gaagaaatac tgggcacggc 6480
atcccattgt aatctataga gtgagtcgct tcttaatatt aaatggctga acacagaaga 6540
tgtgcaaaaa gtactgtgtc cccttcctcc tccaactgaa catttcatgc cctttgcacc 6600
ctcattttgt ctaggagctg ccttatgaag ggaataggta cctgctccga gctggaggaa 6660
tctttgccac ttatggtggg gtatggactg agacagagat ggcatgtgac atgcgcactg 6720
agtctcaact ccatgcaggc tctggagcac tctcaaattg gagtactaat gccttttaaa 6780

ttctcacact agcaatcctt tgacctactg atctagggat ctagggaaag aatcgtgatc 6840
ttaacttcaa agggaaggac aaaatgttct gcctcctgtt aaaactccat acactaagtg 6900
cagagactgg atgccttatt aaccttgggt agatgcccaa atgttcaaaa ggtcaaactc 6960
ttctgttccc cagatcgcca gagtcattaa ccagtcacac tattaaatga atgaacagat 7020
gctgaaaagg tacttgcatt actgagattt cttatggtga tggcccctgc ctgatatgta 7080
ttcagcattt tgtagttttc aatgtgcatt agagtatagt ggtgatgaca ttggcctctg 7140
agtttgccac ttcttatatc tgtgactttg gtcaaattgc ttaatctctc tgagtctcgg 7200

tttcctggag ataataatag cttcttcttc ccagggttat catgaggatt acaggagata 7260

atgccccaaa aatgcttagt aaagtgccta gcacctagtc aatgctgaat taaaggtggt 7320

tattcttact tttcgttcat ttgaactttg ttctcaggga gggcaaagga tagacaaagc 7380

cccatagcta gtgaggagta gctgcaagac tagaacccag gtgttctgag ccctagtctt 7440

aggccaagaa caactgttac gtgagatgca cgttttcctt caagggagct cacaattatt 7500

tccatgtaaa ttcaaggact gctaaaagag aactctcctc tgggactgat atcattttat 7560

ttcaagattg atttgaaaca tgtttttgt ttgtttgttt gttttctagg aaagaacaag 7620

agaaccagtc aagctgaatg cctgaagcaa atctctgtta gcgatgtttt caggatgagg 7680

gagagtggtg caagaaacgt gcttccagat gcacatggtt tcctgggact agggttcagg 7740

gtgtcatccc tgggtgttat taagtgtcag aaggagagca aacaagggaa acatctgaga 7800

tccagctaag gctacaccct ggaaatgcaa gcccagctct tgcaaaggac ctcctttggc 7860

cactcacctt ccaggcctta caataacttg tttggactgc aggtttcttg gtggactcac 7920

aggccattct gcttttattt ggtcaacctc agttcacaag cacccagatg ctgagatcct 7980

cagcatgtgc agcagagttt catattagca ctgggtacct ttctgaggct acagggatac 8040

cgtacagcag cacctgtcac gtccagccaa aggagtgggc tctctcaatg tcatccaatg 8100

ctgtttcaac tgtgaagaag accatctgag agagttgctt ttggaggctg aggcaaattt 8160

ttaaaattct ttgttctcct caactggggt gaattcttgg tcttctagga cagcttgaag 8220

ttttagaaag agtcaagcca ctcagaacca acagagaact ctttcagaga acaaggtgtg 8280

gcatagagga ggcagagggc tgatcttgat caaatccaaa gtgtgactct aaagcaatga 8340

atgtgaattt ttggcaaagc ttacaaaggg ctctaaaggc catctgcaaa gagaagccaa 8400

gcctgatcga tgaatcacta gtgcggccgg atatcgatcg gcacgctgtt gattttctca 8460

tagtaaggaa cagtgggccc tttcagtccc acttctgtag tctgtggtac tacaaatggt 8520

gagcccatga tgttgccatt catagggtta ttctccagca gtaatgactg gccagccact 8580

cccatagccg cagggctagg atttattgtc aatggaggga cctgcagttc tgcacaagca 8640

gtactaggat gagcacctgg gcccattgca agggtgacat cttcaaggca aggcctctta 8700

attttattag ggtagccccc atcagccatg tctggaaact ggaagtggtc ttcttcttgt 8760

ctcctcttaa cagttccctg tgaatggaag agaagagaag aggagaagag aggagaggag 8820

aagggaagag aggtgacaca cacacacaca cacacacaca cacacacaga gagagagaga 8880

gagagacaga gagaaagaga gagagagaga gagaggaatt tttataaagg tttggcacat 8940

taaagctaat gaacaggaaa tgtgcatgat aaaacagacc tctcagttta aagacttata 9000

gttgtgaaaa ctataaaata cagcctgtct ttggaaccat agtgcttatt tattcattat 9060

tatgtttcat ctaaactgtc taattacatt tcaaataagg cattatgttg tctgtatact 9120

aaaacgggat agaacgttat tcaaagggta atctgcccac ttcaaggaga gttcaacaaa 9180

actatgcaga agtcactaaa tgaaccatgc tgccaaaggc aggcattgga gagaaaacta 9240

gaagtagcta aatagtttta attctttcct gtctacagac acatagattt taacgaagga 9300

ataccatagt atagaattga acttttaggc tgccttctag tcttggttaa atgcatcagg 9360

ctgcagtggt aaaattgaat acaacagagc ccttacagga aagaagtaga tctggatgtg 9420

ttttcttggg gagctgttta aaatactgtt tttgggaaag cacaagtttc agaacagtca 9480

ttgtaggcat cgtattcatt gttccattta ttttttacaca cacacacaca cacacacaca 9540

cacacactct cacacattgc tatgtgtaca caaaaataat ttggaagaac ctatacccaa 9600

caatttggag tggtcattta tttgggatga ctggcaattc cctttctatt ctcttcattt 9660

ctgcttgttt gtctttaacg agaacgactc ataatccaaa aatttaaaaa agtataaagt 9720

tatctaaata agaaattttc ctctgaagat gcatcctcag gttggggaga tattaaacaa 9780

tgagaaaagg ccccaatctg ggatctgaac cttgggggag ctgcccatca tttatagaag 9840

cacagccttt gggaacaaag caaagtcact agcaatgtga gacttcctac tcttcatggc 9900

ttcatacagt catccatcgc tgttgtgtta atgaccatga cctgtatgtt agcaggtaaa 9960

tgggaaagga agtgggggca aaggagtatg tgcaggaatg atcaaaataa ggaaaggaag 10020

agagggatct ggaaatcacc tgaatgccga taggtgaaca ggtagaattc ttttaaagct 10080

tcccccaccc ggtacccccc aaataacccc tttccagctt tggaagtttc actaggacat 10140

acagtgctca tcctctgatg tcaccttaag tttggctctt ctggtttgat gagcttgtag 10200

cccactagga gctcaaggca tgcatggggc cacttgccag cacgatgagg ggcatgactg 10260

tcatggccaa gtgaacatca aagcagatcc ccagggctgt atgtctcagg ccttggtgca 10320

catcagaatc acttagaaac atccacattc ctgggccctc ccaccacaaa ctgacagctt 10380

catccagggt gtggcccagg catcgggagt ttttccaaca gctccatggc tgattctcaa 10440

cagaaaacca ctggcccaga gcaagggtgg aggcagcgtg gcatagggct ctgaccttgg 10500

ccttgccact gaacctctca gagccccagt ttctttatgt gtaaaatgag tgtaattata 10560

gttcttttct catgaaggtg ctctgactat taagtgaaac ggggcacatt gtatgacacc 10620

taatagctcc tcactaactg gtacccggca ttataaaggg caggtatgga agggttctgg 10680

gagtccaata cccttcttaa agacagagag gtctctgaga cccagagagg ggcaggcctt 10740

acccagagtt gctcagccag agggcaacaa ggcccaggtc agatgcaggg cccctccacc 10800

accactcagc tgcctccaga cccactgcct tcgccatgtt gttggtagga cactgcatcg 10860

cccccacaga aggggcttgc caacttgagt gagaggactt gcacacttct ttgacttttc 10920

ttttgagatg cccacaatct gaacaagggc acttcaaggg acagctctgt caccaaactc 10980

atctgaggcc tgaataccat gggtcaggca ggaatgggtt ggagaggtgt agagcaggca 11040
caataagagg gctgaggccc atgcagtcat cagtgcccac tttcccagga gtctgactgg 11100
gcacagcacc catagtgtcc ctgagctggt ccatggagca gctcactaac tgtttggccc 11160
acagcaggtg ctcagtaaat ggcagttgaa cgaatcaatg gacaaaggaa cataaattac 11220
ccaacacaca gggagctcag ccatttactc aatccattat ggagtaacct acaaacaagc 11280
cactgggtcc caaactgaaa ttgtgtctct tctacattct cccaaagaat ccaataggtt 11340
aaaaatagaa atgtatgaaa tagatcaatc agggatgatt gcatgtggat ttgacataag 11400

gatcccctgc agggagtctg agctggcaac agtcaggccc aaagtgctgt ccatgatgtc 11460
tcgaactgca agacagtttt aacaatggcg aagcaatgca gaaccaggca ggccaaggag 11520
ggggtggggg ttggggaaag gaagggaggg aaggggctgt gaggggcaat ggtctggcat 11580
ccctgccacg tgagcctctg aaatttgctg gcagcttcta tgggctccca gagctttcac 11640
ttaattgttg gtctgccact aacctgctgg gagtaaggtg cagggatgga ggaggcaggg 11700
catgaccacc agacactaaa ggtaccagct ggggccactg gcaaagggaa ggaggctgca 11760
cctctcctac atgagagccc gtatacacac acctttcca gcactcatca actgcatccc 11820

aagcaaatgg tccctgatca attccaattc tagaaaccaa ctgactactc aataacaaag 11880
tagatcccag caggccgcca ctgctggagc ggatgccact tttgctatgc caagtctgtg 11940
gctggacagc tgctggcatg tacactcact gactttcata aggatgccta ataaaggggg 12000
caggctcacc tggcttttct caggggtggg gtttggggtg ccgatagagg ctgctgtttt 12060
ggcagagtgg caagctgcaa gcctcttctg agctttcatt tttcaatgga cttcagtgag 12120
aattcacttt gtcagaggcc atgcagctcc atgttttgga tttcatggaa tgagctttca 12180
acagtgagcc tgaagtgccc tggctgaaca gcaagaacac cagccaaccc taaacaaggc 12240

cgaggagagg cggctgtgtt tacacggaag gctcagcctt gctgtaatag cgtctgcctt 12300
caccagacat cagtgaggcg tggaaatcta ttatccagtt aattttgccc ctagataaag 12360
acttgctttc gtgtcttctc tttcacagtc ccatgatctg ttactcatct caactgcgag 12420
aagttggctg ggctttcccc tgtgcccagt gccacactcg tgccttcact gggtcacctg 12480
tgcctgtggc tgatgccgct gaggttttgc ctgcccagac tgggtgtttc tgactaaatc 12540
ccacagccac cattttagat caagggcagg agatagctca ctgctccgga atgacctccc 12600
ctcccagaat cctggtaggg gcggaaggtc cccaaccaag ctcccagccc tttctaaatg 12660

aatctccctg cttcacccat gtgcttttct ccagtctctg cggtcttgat gacagcaggg 12720

tattagtcct agctgtccca cagctcctac ttctttcagg cctctccctg tgacaatcag 12780

tagccactgg caggatttcc tcagagcata tctcgatttg ctttcagaca attagttaaa 12840

aggacactgg accccagacg tcccaactcc cagccagagc cctcacaggc ccggcctttg 12900

gtggtgagga agggggaggg agtgagtgac agtgccctgg catcttttag aaacgaattc 12960

ctttctctcc atacataaat gcctgcagag tcccatttca gaatccggca gacaaagcca 13020

ccaatgtgat ccccatgacc ttataaacat tcattaaaat gcatttcaag gcatgtgatg 13080

gcctccccac cccctagata atgagaaaac aaaggtttct cttctgatag agacaagttc 13140

agctctgaag tcaacattat ttctggttct gtctgaacaa tgacatatgg caactcttcc 13200

ctttctatag ttctagtcca gaatgacaaa aaaggggaaa aatttcttag agaaggtaga 13260

gattatacga atacagtcca tgaaatgagc ataaggagaa taaagaatat aacttatcca 13320

aagaagtctg gcaggctgtt ataaatgctt gattttggac actgtagttg gaggtttaac 13380

atggacacca ataaaaaggt cagcaaaggg tatgcactgt tcctattggg caagaagata 13440

ggaggtcaaa ggtaaccagg aaagataaac tcagggagac ttattttccc tccagagggc 13500

actgggcttg taggccctgg gcaaaattgt caaaaaggtg aaaatcgcct gtggtttatt 13560

tagtctgctc tttcttcact agtgcctcac cagttcagtt caggccaatt tgctagccag 13620

ttctcctagt ctcttccctt acatccatac tccatgaaca actttattca atattgttac 13680

attttccccc tatgaatttt gaggcacctg atgttggcca actaaattag acttcaagcc 13740

ctctgaggac aaggatggcc cccagtagac tatctcagaa gaaagtaatg aagatgtcgt 13800

ggggcatgat ggaggtagct gtagaataga taagaaaata atgtatcagg aatgggtttg 13860

aagagacaaa gagatggggg atggagagaa agagagagag tgagagcgag gaggagagaa 13920

agacagaatg aatatgtttg catttgcaca tgggatcaaa ggaatatccg tgtgtatagt 13980

atcaatgcct gtatttgggg atgtgttatt tacatgttta ttttcagcag aatgtcccag 14040

caaactgaaa ctgatctttc tgtagaacaa acaggatttt gagcatttcc agaattgcaa 14100

gacatgaggc tctggagaat accagggtgg ttgggaatgg cagggaaccc tgaatgctaa 14160

tgctgattgg cttcaagagg agtcagttat gtctcaggag caaatctctg ataatgggtg 14220

gtttggagca agaacctcga cctacaaatc ggggatttat tggtgtccct gattcgccct 14280

gaaccagcac tcgatgcagg cagatgcagg ctgcggtgga ggaaggggcc ctggaatgca 14340

catgagaaga cctgcagtta agtcctagct cctccgtttg ttgtgtggct gtgggcaagc 14400

cactcccctc cctgaagctc aatttcttct tctgtaaaat gggtacttgg cccaccatgt 14460

gtgcttcaca gcattagtga aaggttccac tgtactcact gaggagagca cttggggagc 14520

agtaatcact actgatatca ttctcaaggc cctgtgtctt tgtgttgggc ctctgttgct 14580

ttcttacaaa taaggatgca ttcactagga tagggggatg gagtggaagt agggggaggc 14640

aagcagctta gtcctgacag cactgacctg acagcactga cctgacagcc agggatgctg 14700

tcttggtggc tgagcctgga aactcctgcc tcctgctgct tgccaccacc tgccactctg 14760

aaggctgcca cagtcatcag cctttagcca ccctgacctg agggctgggg cttcaacagg 14820

cttctgccag aactatctga gcagcagcct ggagtggctg gaagtgcact gacaaaggca 14880

gagatgagca aacctgggtg ggcagtgttc tagaactact ttctcctttt ctgtagtgta 14940

gacatgccag ttggaagaac tgacagctct aaggatgatc ctcaacctat agatgtatat 15000

gaatttaatg agctcagaat ttgtattgta ccatctcctg agaagaatca aggtattctg 15060

ccattgaaca agcatttgtt gagagctagc ctctaactcc tagaccgtca gaactgctgg 15120

gcccttcaag acgggctgct cacacccact catgttaagc ctggtgaggc ctgtactctg 15180

ttttcacagg aagaaatcct cacccagtct tccccaaaca cattcccagg ttctgtcatt 15240

agtgggatag agatgattac tgtggggaga agagaaacat ctggatggat ttggtgaggt 15300

tgatctatag aggaagtagg tgctgcctga ggtagctgta atagaagcta aaggtcaaag 15360

gagagggccc tgtcccaatc cagatgactc cacttctgct ggacccaggt tcacaagctt 15420

aatctacatt tcacctaaat ttggctaaca agcccaaaat cacacaggca aagggagaag 15480

tggaggcaga accgaggttg gaggccacca gggccaccgg gcagagatca tttaagccca 15540

accttctcac ttctccctgg gctctgcctc tcttaaagga ccttgtggtg tgacctcttg 15600

taggtccctt tcacactcgg ggcctcagtt tccccactgt aaagtgaatg ggtcccagct 15660

ttggtaagct tatgcttacc tgatgctttc ttcctgggct gctcttgtag agaaaagata 15720

aatcttcttc ctccatccac gagggcttct ttccctgggg gtgagagtag gctgaggaga 15780

gccacttgca cacactctta aagaaagtat tacctgcacc agctcagtga gaggcacaga 15840

tcagactgtt acttgaatca aattatgagc ctccccaaat atatctatga catttaaata 15900

ggggattact tgaacataga ctttgggatc cggtgtggag tgcaggagac tagcaaagtg 15960

aatcctgaga gtagcaggtc tgcacctgtt ggatcgagaa aggcggccta caattctggt 16020

caaatgagct gtgcttattg acatattcta ttagagagta ctaccaggtc accagtcacc 16080

agaaaggctg ccagctctcc aaccacctcc agggaactat cctgaatggg gccttaacaa 16140

gcctaagaga gggttggttt gggtcccaag ccaatatttg ctctgcttta tgtcagtcat 16200

atggaaccca aaccaaccct ctcctatgtg cctcaccagt cggtgcaggg atcccaattt 16260

caagtttggt tttttatggt caaagtccag catagattaa atgaaggggt gtgatgatgg 16320

tgttaaaaga gaactccaga ccagtttaac tcttggacac acatcccatc tcaccatggt 16380

gcttccaacc ttccagagat gatgggctcc tattttctga tgacaaagcc ctccacagga 16440

ttgctgcctg gccatcaggg agtgcctctg taactgaggc tgagatccca ctttcagtcc 16500
tccagctgtg gcccatccct gctccgccca ccgggtatgg cctgtcctag gctcttaggt 16560
atggctgcat tgtgaaatga tggctacaga gctggcatct cctgtagtct ggttcatcta 16620
gtgcactacc tcatagttaa aagaaatctg tttaagccac tgagggtggc tcctagtgcc 16680
aactccaaga acaggaagct tccctttttt gggaggaggg gcagatggta acatggatcg 16740
tccaggtcaa tgggagcagg gcaaccacag taagtactgg acaacaacac aaaactccat 16800
gtgtggcttc catcgagtcc ctctccaatt ggtttggtct tctccgtccc atgcagcact 16860

ttagcaaggg gcctggctga aggctatgaa ttgtgtggag cctcctcatt gcagtctcca 16920
accatctgat gctgggaaaa tgtcaccagg atgcagccat gccgtgtggc caatgaaccg 16980
agaaaacacc ccttttctag aatgctctaa agaggcagaa taatccagag gtgaggaagg 17040
aaatactcca ccagagaccc aggcagttcc tacaaaagcc agactttcct tcacctaggg 17100
agtgacaaga ccagtggaaa acactctcaa gcagtaaccc ccaaatgctc tgcaagccag 17160
tggcgtccag ataccgcaca agcgagtggg ctgtctaatc ccatcatcat gatgtaaata 17220
tctctaggct gccctgggct gtgcctgacc ctgtcttcag ctttccacac ctccacctac 17280

agcccatgca cagaaggacc acccaggaat gctgcaagtg tggcacctcc agggccaccc 17340
agggagaagg agggcagcta tgctggtggc tccaggccca tttggcgggt ggtaccttca 17400
caccacaaag cccaaactga ggccccagat ttggctgatg agggcatatt ggacaggggt 17460
cacttatgct cttccccatt gccacctggc ctctggctac ctggacttgg ctacctgtgg 17520
atcctctcac aggtgccacc atcttggctg agtctccaga tgcgaggtcc ctgaggcagt 17580
ggcaggcttc tcgctaatgc tgatgggatt aggaatggga taggtgggga gggccctgga 17640
ctgggccctg atgagccaag tgggttttta gaggggctac tggtacattt cagggacagg 17700

acatctggta gagctaagct ggggcaataa ggagccactg ctaatctgag agctagaaac 17760
aatcagcttc tgggtcatta ttaattaggg tagtttgggc tgtgtggaag tcacgtacta 17820
tatggggtag ccacagctct ctctacagat aatctctaag acttctgatt gggaccgtgt 17880
gaatgcagta gcaatatctc ttcttactgc caggccctgc cagtcctgcc tccacgccct 17940
ggctggcccc ccttatgatc tgacccatgc caggctgcca tagtatgtta cttctgcatt 18000
agcactcctt gggacctgcc tctccactgt ccctcagact ttaaagaact atacaaaccc 18060
aaggggctct tcccaagaga attgatatga cttgaggtga ttccatttct ggaagtagtc 18120

actccatttt ctgcctcact ctttcagtgc ttcacagagc aggttcgaac gaaggagcca 18180

tccaactaac cgtcatgttc gggcaaccga agaagggagt ggcaggattt cctttggaga 18240

cttctggaat tagacagcag tttaatgcaa gcatctaaat tctcttcctc ccagagtctc 18300

attaaaacta cagtaagagt ttgtgttttg ttttgttttt aaagacaaaa tcccaccagg 18360

atagagagaa taggagagga gataacagca tcataattta tgaaactaaa atgcagatag 18420

accaggatta actgactaca cagcaccaag gaagctgaat cacaagacag cagaggagaa 18480

aactggaaag gatcgtggtc tatacggcag aatcttccca agcctcagga ggaggagctc 18540

tagatgttcc cagatctggg aggtaaagtg gaatgggggg acatggtcag cgtaatgggg 18600

ttgggctgga agctggttaa ggagcaggca gatctctgaa tcccctctct gactctgtgt 18660

ccccaggcat ctgcctgtcc cccaccctgg aagaggtctg gcttgaccct ttgtctggtg 18720

aatttcctgc tctgctttcc tggtcctgct ggccagatca gtggaggcca ctcacttcac 18780

cccacaggga tgttctgtgt tgccctacac ctgggaactg gaggtactgg aggcaggctg 18840

tggtgagctt gaaagcaaaa cacagagggc agtccaatct ctttggccat atttcttctg 18900

catatccaat accatgtcca caactctgct agtgtcctga tggtggtggg ctctacacat 18960

tcccgggaag ctgaaggcag ataatgacca ggacaggtca acctctcttc ttctgaaagc 19020

cttcatctac taatggcctg ggactcttcc cttaaatgct tagattgtgt cttccactaa 19080

ggtttttttgc tgttgctgtt gtttgtttgt ttgtttgttt gtttgtttgt tttgagacgg 19140

aatctcactc tgtcgcccag gctggagtgt agtggcacaa tctcagctca ccacaacctt 19200

cacctcctag gttgaaggga ttctcctgcc tcagcctcct gagtagctag gattacaggc 19260

acatgccacc atgcctggct aatttttgta tttttagtag agacaggatt tcgccatgtt 19320

ggccaggctg gtcttgaact cctgacctca ggtgatctgc ctaccttggt ctcccaaagt 19380

gctgggatta ctggtgtgag ccaccacacc cggccaaggt ttttgtttgt ttgtttgttt 19440

gtttgttttg tattgaggca gggtatcact ctggtcaccc aggctggagt gcagtagtgc 19500

aatcacggct cactgaaacc tccacctccc tggcgggctc aggtgatcct gccacctcag 19560

cttcccaggt agctgggact acaggcttgt accaccactc ccagctaatt tttgcgtttt 19620

tagtagagac agggtttccc catgttgccc aggttggtct caaactctgg gctcaagcga 19680

tctgcctgcc tcagcctccc aaagtgctgg gattacaggt gtaagccacc gtacccggcc 19740

ccgccactaa ggttttgaaa atgaagcaat tacaagttta agtctattaa taagtgatga 19800

agctatgtag aaaagcagaa taattatctt ggatcaggaa ggtcacatga ggatctactt 19860

gggggttgtc aatattctat ttcttgacct gatcagtgtt gacagcaggt tttaattttt 19920

tacttctttt tgtttgtttg tttttgagac ggagtcttgc tctgtctccc aggctggagt 19980

gcagtggtat gatctcggct cactgcaacc tccgcctcct gggttcaagc tgttctcctg 20040

cctcagcctc cccagtagct gggattacag gcaggcacca ccacgaccag ctaatttttg 20100

tatttttagt agagactggg tttcaccatc ttggccaggc tggtctcgaa cttctgatct 20160

cgtgatccgc cctccttggc ctcccaaagt gctgggatta caggcttgag ccagcgtgcc 20220

cggcccattt tttacttcct tattaaactg tacatatagg ccttgcacac ttttctgcat 20280

caatgttata ttccacaata aagggaaaag gtatatacac aacttgatac cagtaatgtg 20340

aaacatatat ttctacatag aaaaaaaaat gactgaaata ctgcactcca atgtgttcac 20400

acagtagttg tttctggatt atttatatat taaatgttta tatattgtat tatgccatga 20460

ggtttgtgtt ttctctccac ttttctgcat tttccaagtt tactacaaag agcacatatt 20520

actcttataa tcagaaagtc ataaaatata tttaaaaaga caaaattgaa actaataagg 20580

atcaacacaa aacagatgag ccatctgtgg aaatccgcac agaatactac ctaaagagat 20640

tggtgacgtg catgatctca ctaggatgag cacaaagctt gccagagcct agggtctatt 20700

tctagggttg gctcttggaa gccaggatag ttgttatctc tgggaagagg gaggggcaca 20760

caaggggctt ctaaaacatt ctgaatgttc tatttctgaa cctggttggt gggtacatga 20820

ctgttggttt tattattata tgttttatat actcttccgt atgtatggtg tggattccaa 20880

aaaaagattt cctttagaga aaaccagaat cacataagta gaaaatatgg tgctatgttg 20940

aaggaacaac tcaagtttat ataaaatcat catcatttat aggcttaaaa agttgctttg 21000

gaattttggt ctaactgact tgtcttttct gcagcaaacc acgctccttc tggacgtgct 21060

ccaggcagag gggattaggg tgggttcaag gctgcaagta cctagctcag cacactctct 21120

tcaggggact tagagtttgt ctggtgttgg ctctctgagc tcttgtcagg aatgccgacc 21180

cttccgaggt tcaggatttg aagcctgcct tcccacccca gatttggtcc acacagacac 21240

tcaagtatgt atttcaacta caaatgacct gtactttcct attactcctc tctttcatgg 21300

taacctttct ggtatccttc cttccctaca tttatgggag ggggacatca ttctctgctc 21360

tcctgtcact gaaggctcca ccttctgtct tcttctgacc catctggttt tcctggggcc 21420

acctcctctc cttaccaccc taacgctttt gtaacttgag gagaaatgag agatcaccta 21480

gtcaggtcat cattctctgt agatgaagag gcccaatggt ttgctcaaga attgccaagc 21540

gagttaaaga cagagagtat gagagtcagc aagacctaca gaaagcatct atctgcactg 21600

ttttgcaggg acttagcctt tgtgtgtgga ctcctggaat gccacccact aagaaacatt 21660

gtctgacacc aactccccac ttggtaggtg gggacactga aactcatggc aggaaagggc 21720

cttgccccaa gccagggcag agtgtcactc atcactctca attttcagtc cagggcacct 21780

tgttgtgact atcccaaagg cagccacttt ccctggtctg aaagacctga agagagaaga 21840

gaagagaagg atggaaggca gagtatgcgg ctttgattca tttcctggtg aaaacagatc 21900

tatacgagaa gcaaatttca cgaaagggaa gagaagaaag tgtcccatac gttgctggcc 21960
tgtttcaacc ttgctttgat tcttgctgaa aagggtaccg tgtatttctg agttcaacat 22020
gcagaccagt gttaggaaag ccactgcacc tccactttag cctccagggc tgtgccctgc 22080
aaatggcctg cagccttggt gcctcgctct ccagactgca ttttggaaga tgggacagag 22140
gcttatggaa gcccacatta gaacgggggа gcagaatggg tgagatgagg gatccttgat 22200
agtgaaccag atgaaggaat ggtagccaaa tgccaggcct cctttgtggc ttcaatccaa 22260
aggctctgga gcccttccag ggcagaacat caggcatgtt tacccccact gtcctcaaca 22320

gtgacagagg tgcaatcttg ggcagctggc cattttgaaa gcaacctcct taatctcaac 22380
tgggaaggct ccctagcagg acccctgtgt tgcacacctg gaggaagcta gactaaccag 22440
aagctcagca cggttccatc tgggatgccc aggtctgaga cgaaaaaggt aactctcttt 22500
tctgggtcct ggcccagttg tgtctctctc cacctcattc tctgagatgc ctgtctcccc 22560
ttttttgtcc catcaggagg caagagctat cactgggcca gactccacca gaagccaagc 22620
cagcttgtta cccagcttct cagggagcaa agaacagcct tgtttctatc ttatccccac 22680
tgtcccctgc ccctgcccca cctcccagcc attcagcttc tggcttcccc agagctgcct 22740

gcttctttgt ggtcctccat tccttgaaaa gaccttctag tcattagtgt atataaatgg 22800
ccacttagcc cagattacag tgaggtcaac agctggggct ctgagaattg tcacacactg 22860
gcacaggaga ggaggctatt cttccagaga atttggaggg cactcccatc cacttacaac 22920
aaaaagccca tccactgtgc ttggcagtag gtgatctgag aaccaatgga accaggttaa 22980
tcctgtggca ctgttgagtg aggagagcag tggcgggcac tggaaaatat cagagacaag 23040
gcaggagacc tgaaatctag gcttagctcc tcatatactt ggcagctgta tgacctcaga 23100
caaccagtgt tacctctcta agcctcagtt tcctcatgca aaaggagggg gaataacaac 23160

agagcccact gcttgggggt gttgtgagga caggatgaaa aaacaaacag aaatccctca 23220
gtacaggatt cagtgcagtg gacagtcttg caaggtctgg ttcagccctc cacccctacc 23280
ctcaccagta taaagaactc tggcctacaa gtcagatgac ctgagtttta atctcagctt 23340
tgccattagc cgtgtgaact tgagaaagtc cctttccttt ttacatctat tgggatgatc 23400
atgcattttt tgtcctttat tctgttaata tagtgtgtta cattgattgc ttttcataga 23460
ctgaaccagc cttgtattcc agggataaat ctcacttggt catggtgtat aatcctttat 23520
acaaatgttg ctgggttgag tttgctagta ttttgttgaa gatttttatg tcttgattca 23580

taaggaatat tggtgtacct tccccttta tggccacagt ttccctacaa tgatgtagtc 23640

gaactagaca acctccaata tctttcagta ttcatgtcct ctgattctgt gaaactaaga 23700

aaattaagaa atagtgattc ataggcacaa ggcaggcaaa acttagactc cttgtagaat 23760

aattaggaag ccaaatattc agtgtgctta tttctcaaat aaccttagtt tctccagtct 23820

gccccaactc cgaggcctga atatctctag atgcttatga tggcaactaa agcctaaaag 23880

ctaattcatt ttaaagttct tccaaatgca tagggtttta tttttccaga cctgggttca 23940

gatggggaat ttgacaaaca atggaaaggg ggaaaaacaa caatctaaac actgagtgac 24000

aaagtaacaa agaaatagtc tagctatcag ccagtcaagc cagccttggc tttgctatcc 24060

aaagtagtca gtctaattct accaccagtt tctgttcctg tagctgtcta ctgcctgcca 24120

gggactctgc cttcccaccc acaactacca atggaaggat gtggtgacca taccagtggc 24180

tgctgacatc tcctgccatg ggaagcataa ttgcctccag cagcctcccc cttagatcca 24240

tcatttttgt tgcacttggc ctgggctgta ctcccggcca atgactgaac atggtgagca 24300

tagtaatgca ggcccatttc tgtgaggagc aggactcctc cagtaggtga ctttggctca 24360

aggactctct attggcctgg ttgaactttt cctgaactgt gctactgtct gagactcttc 24420

ttacccaatc ctctttctcg ccccaattgt cacagaccac ctgcattgtg gtctgagtct 24480

ctccccacct tctcttgctc ttccctgttt atctttcaca ggcatttccc ccagtacatt 24540

ccttgaatgt ctaacccgat acgggtgcct gacttttggc agacctaagc agacaaaaag 24600

gagtacttgg ttacctagct cttctttcta ccacaaacat cgagggaacc ctttttccct 24660

caccctctg ccacaccccc actgccccag tgaacaacca cagagagagc tgtggtataa 24720

tattaggctg gtgcaaaagt aattgcggtt tttgccatta cttttaatgg taaaaaccgc 24780

aattactttt gcacctacct agtatttgtg tccccccaaa ttcatatgtt gaaacctaac 24840

ccacaatatg atgtcattag gaggcaagac cttgaggagg tgattagatg atggggtgga 24900

gctctcctga atgagattag tgcccttata agaagaagcc caaggaagct accttgactc 24960

ttccatcaca tgagaatgca gcaagaaggc accatctact aatcaggaag agagctctca 25020

ccagacactg aatctgccag tgtcttgatc ttgaagttcc cagcctccag aactatgcat 25080

aatgcatttc cattgtctct aagccaccca gcctatggta ttttgtcata gcagcctgaa 25140

ctgactaaga cagtgagcca catgagaagt gccccaaccc ctcccttaag cacttggctc 25200

acagatcagt gggttcattt ctgcctgagt tttattgtta ttctgtagat ttcttgggct 25260

agatatattt ttctgttatt ttccttcttc acctcagtca tgaattggtt gttttaaaaa 25320

agacaatgta agtcatgggg aaactcctga caactctact ctcctagggt tcctgataaa 25380

aggggattca gttgagtcct ctgatggtct ctacctgcca aagtccagca gcccttagca 25440

aacatgctgc tcgtttctgt agagaaggtg ctggtgtccc accatacttc tctctccctc 25500

atgaagggct tgcgacccag caaatgggtg gcttatatgg gtctgtttca aaggaagagc 25560

cagctctggg aagaaaaacg atgagcataa gcataaccta ccactgtgcc tgggaaagca 25620

gacaactttt ttgatgtgtg aatatctaat gagaatggaa tccatcaatt accttaaact 25680

taggcacagt cttcaaattc aatatatgtg ggatatactt ttagtcagtt tgtagacgtt 25740

atttgtaata aataatctgg cttctctaaa gaaattattt taagtgtttg gtttggtttg 25800

atttaatggt aaaattatat ttagtggcag agaattatag caatggtgat aaactataga 25860

gtgtcataag ttcatatctt attctcacat ttgaagctgc ctgcagatgc attcaagatg 25920

cagccagaag tcaggagact caggctgtta tttggagctc atcattttac agccttgctg 25980

gactcccact ttctcagggg aaaaatgtgg tgttgaccca gattagctct ccaggccctg 26040

ctgagttggg cactctgtaa gctggagggt cttctattgt cttcacctaa gtgtcaatca 26100

acaacccaaa tgggcatggg ggaagaggga gctgggccaa tgcccagggt gcctggtaga 26160

gagatacctt gggcactgga aggcaccagc ttcccagaga gaagggggag ggccatgaaa 26220

aagttggctg tagatgccag ggacactggg actctccagc tgtgtgtttg tgtcttctga 26280

agacttatgt ttcattcctt tggagcatgc ataatcatac actgtgggat gtgttatata 26340

gattgcttga tagttcacca ctgtaataaa atactgtgac tggaatctgc tcccagtctg 26400

cctttgatag cacttgtgca acacacattt actgagcatt tacagtgatc caggacctgt 26460

gttgtgaaaa cattgatgga caaggcagat ggtggagcac gtcagtgagg atttttaaca 26520

aaggctggta agtgctataa aggaacattg taggacacta gagaacaaag aacaggagaa 26580

cctgacttag gctggggtgg ggcgttggtt agaggaggct ccttggagga catgaggttt 26640

aagctgtgac ctgaggatga atagatgttg gccaggtgag gttaattaag cttgtcagcc 26700

ttaccagtaa aaaagaaaac ctattaaaaa aaggcttacc agagaaagta cagcgggcac 26760

acaaaccaca agctctaaag tcactctcca acctctccac aatatatata cacaagccct 26820

aaactgacgt aatgggacta aagtgtaaaa aatcccgcca aacccaacac acaccccgaa 26880

actgcgtcac cagggaaaag tacagtttca cttccgcaat cccaacaagc gtcacttcct 26940

ctttctcacg gtacgtcaca tcccattaac ttacaacgtc attttcccac ggccgcgccg 27000

ccccttttaa ccgttaaccc cacagccaat caccacacgg cccacacttt ttaaaatcac 27060

ctcatttaca tattggcacc attccatcta taaggtatat tattgatgat gtttaaacta 27120

cggcccggta cccagctttt gttcccttta gtgagggtta attgcgcgct tggcgtaatc 27180

atggtcatag ctgtttcctg tgtgaaattg ttatccgctc acaattccac acaacatacg 27240

agccggaagc ataaagtgta aagcctgggg tgcctaatga gtgagctaac tcacattaat 27300

tgcgttgcgc tcactgcccg ctttccagtc gggaaacctg tcgtgccagc tgcattaatg 27360

aatcggccaa cgcgcgggga gaggcggttt gcgtattggg cgctcttccg cttcctcgct 27420
cactgactcg ctgcgctcgg tcgttcggct gcggcgagcg gtatcagctc actcaaaggc 27480
ggtaatacgg ttatccacag aatcagggga taacgcagga aagaacatgt gagcaaaagg 27540
ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg gcgtttttcc ataggctccg 27600
ccccccctgac gagcatcaca aaaatcgacg ctcaagtcag aggtggcgaa acccgacagg 27660
actataaaga taccaggcgt ttcccccctgg aagctccctc gtgcgctctc ctgttccgac 27720
cctgccgctt accggatacc tgtccgcctt tctcccttcg ggaagcgtgg cgctttctca 27780

tagctcacgc tgtaggtatc tcagttcggt gtaggtcgtt cgctccaagc tgggctgtgt 27840
gcacgaaccc cccgttcagc ccgaccgctg cgccttatcc ggtaactatc gtcttgagtc 27900
caacccggta agacacgact tatcgccact ggcagcagcc actggtaaca ggattagcag 27960
agcgaggtat gtaggcggtg ctacagagtt cttgaagtgg tggcctaact acggctacac 28020
tagaaggaca gtatttggta tctgcgctct gctgaagcca gttaccttcg gaaaaagagt 28080
tggtagctct tgatccggca aacaaaccac cgctggtagc ggtggttttt ttgtttgcaa 28140
gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat cctttgatct tttctacggg 28200

gtctgacgct cagtggaacg aaaactcacg ttaagggatt ttggtcatga gattatcaaa 28260
aaggatcttc acctagatcc ttttaaatta aaaatgaagt tttaaatcaa tctaaagtat 28320
atatgagtaa acttggtctg acagttacca atgcttaatc agtgaggcac ctatctcagc 28380
gatctgtcta tttcgttcat ccatagttgc ctgactcccc gtcgtgtaga taactacgat 28440
acgggagggc ttaccatctg gccccagtgc tgcaatgata ccgcgagacc cacgctcacc 28500
ggctccagat ttatcagcaa taaaccagcc agccggaagg gccgagcgca gaagtggtcc 28560
tgcaacttta tccgcctcca tccagtctat taattgttgc cgggaagcta gagtaagtag 28620

ttcgccagtt aatagtttgc gcaacgttgt tgccattgct acaggcatcg tggtgtcacg 28680
ctcgtcgttt ggtatggctt cattcagctc cggttcccaa cgatcaaggc gagttacatg 28740
atcccccatg ttgtgcaaaa aagcggttag ctccttcggt cctccgatcg ttgtcagaag 28800
taagttggcc gcagtgttat cactcatggt tatggcagca ctgcataatt ctcttactgt 28860
catgccatcc gtaagatgct tttctgtgac tggtgagtac tcaaccaagt cattctgaga 28920
atagtgtatg cggcgaccga gttgctcttg cccggcgtca atacgggata ataccgcgcc 28980
acatagcaga actttaaaag tgctcatcat tggaaaacgt tcttcggggc gaaaactctc 29040

aaggatctta ccgctgttga gatccagttc gatgtaaccc actcgtgcac ccaactgatc 29100

ttcagcatct tttactttca ccagcgtttc tgggtgagca aaaacaggaa ggcaaaatgc 29160

cgcaaaaaag ggaataaggg cgacacggaa atgttgaata ctcatactct tccttttca 29220

atattattga agcatttatc agggttattg tctcatgagc ggatacatat ttgaatgtat 29280

ttagaaaaat aaacaaatag gggttccgcg cacatttccc cgaaaagtgc cacctgacgc 29340

gccctgtagc ggcgcattaa gcgcggcggg tgtggtggtt acgcgcagcg tgaccgctac 29400

acttgccagc gccctagcgc ccgctccttt cgctttcttc ccttcctttc tcgccacgtt 29460

cgccggcttt ccccgtcaag ctctaaatcg ggggctccct ttagggttcc gatttagtgc 29520

tttacggcac ctcgacccca aaaaacttga ttagggtgat ggttcacgta gtgggccatc 29580

gccctgatag acggtttttc gccctttgac gttggagtcc acgttcttta atagtggact 29640

cttgttccaa actggaacaa cactcaaccc tatctcggtc tattcttttg atttataagg 29700

gattttgccg atttcggcct attggttaaa aaatgagctg atttaacaaa aatttaacgc 29760

gaattttaac aaaatattaa cgcttacaat ttccattcgc cattcaggct gcgcaactgt 29820

tgggaagggc gatcggtgcg ggcctcttcg ctattacgcc agctggcgaa aggggatgt 29880

gctgcaaggc gattaagttg ggtaacgcca gggttttccc agtcacgacg ttgtaaaacg 29940

acggccagtg agcgcgcgta atacgactca ctatagggcg aattggagct ccactacgta 30000

gtttaaacat catcaataat ataccttata gatggaatgg tgccaatatg taaatgaggt 30060

gattttaaaa agtgtgggcc gtgtggtgat tggctgtggg gttaacggtt aaaaggggcg 30120

gcgcggccgt gggaaaatga cgttttatgg gggtggagtt tttttgcaag ttgtcgcggg 30180

aaatgttacg cataaaaagg cttcttttct cacggaacta cttagttttc ccacggtatt 30240

taacaggaaa tgaggtagtt ttgaccggat gcaagtgaaa attgctgatt ttcgcgcgaa 30300

aactgaatga ggaagtgttt ttctgaataa tgtggtattt atggcagggt ggagtatttg 30360

ttcagggcca ggtagacttt gacccattac gtggaggttt cgattaccgt gttttttacc 30420

tgaatttccg cgtaccgtgt caaagtcttc tgtttttacg taggtgtcag ctgatcgcta 30480

gggtatttga ttc 30493

<210> 55

<211> 14414

<212> DNA

<213> Artificial Sequence

<220><223> pWE15-A2

<400> 55

ccacgacggg cgttccttgc gcagctgtgc tcgacgttgt cactgaagcg ggaagggact 60

ggctgctatt gggcgaagtg ccggggcagg atctcctgtc atctcacctt gctcctgccg 120

agaaagtatc catcatggct gatgcaatgc ggcggctgca tacgcttgat ccggctacct 180
gcccattcga ccaccaagcg aaacatcgca tcgagcgagc acgtactcgg atggaagccg 240
gtcttgtcga tcaggatgat ctggacgaag agcatcaggg gctcgcgcca gccgaactgt 300
tcgccaggct caaggcgcgc atgcccgacg gcgaggatct cgtcgtgacc catggcgatg 360
cctgcttgcc gaatatcatg gtggaaaatg gccgcttttc tggattcatc gactgtggcc 420
ggctgggtgt ggcggaccgc tatcaggaca tagcgttggc tacccgtgat attgctgaag 480
agcttggcgg cgaatgggct gaccgcttcc tcgtgcttta cggtatcgcc gctcccgatt 540

cgcagcgcat cgccttctat cgccttcttg acgagttctt ctgagcggga ctctggggtt 600
cgaaatgacc gaccaagcga cgcccaacct gccatcacga gatttcgatt ccaccgccgc 660
cttctatgaa aggttgggct tcggaatcgt tttccgggac gccggctgga tgatcctcca 720
gcgcggggat ctcatgctgg agttcttcgc ccaccccggg ctcgatcccc tcgcgagttg 780
gttcagctgc tgcctgaggc tggacgacct cgcggagttc taccggcagt gcaaatccgt 840
cggcatccag gaaaccagca gcggctatcc gcgcatccat gcccccgaac tgcaggagtg 900
gggaggcacg atggccgctt tggtccggat ctttgtgaag gaaccttact tctgtggtgt 960

gacataattg gacaaactac ctacagagat ttaaagctct aaggtaaata taaaatttt 1020
aagtgtataa tgtgttaaac tactgattct aattgtttgt gtattttaga ttccaaccta 1080
tggaactgat gaatgggagc agtggtggaa tgcctttaat gaggaaaacc tgttttgctc 1140
agaagaaatg ccatctagtg atgatgaggc tactgctgac tctcaacatt ctactcctcc 1200
aaaaaagaag agaaaggtag aagaccccaa ggactttcct tcagaattgc taagtttttt 1260
gagtcatgct gtgtttagta atagaactct tgcttgcttt gctatttaca ccacaaagga 1320
aaaagctgca ctgctataca agaaaattat ggaaaaatat tctgtaacct ttataagtag 1380

gcataacagt tataatcata acatactgtt ttttcttact ccacacaggc atagagtgtc 1440
tgctattaat aactatgctc aaaaattgtg tacctttagc tttttaattt gtaaaggggt 1500
taataaggaa tatttgatgt atagtgcctt gactagagat cataatcagc cataccacat 1560
ttgtagaggt tttacttgct ttaaaaaacc tcccacacct ccccctgaac ctgaaacata 1620
aaatgaatgc aattgttgtt gttaacttgt ttattgcagc ttataatggt tacaaataaa 1680
gcaatagcat cacaaatttc acaaataaag catttttttc actgcattct agttgtggtt 1740
tgtccaaact catcaatgta tcttatcatg tctggatctg acgggtgcgc atgatcgtgc 1800

tcctgtcgtt gaggacccgg ctaggctggc ggggttgcct tactggttag cagaatgaat 1860

caccgatacg cgagcgaacg tgaagcgact gctgctgcaa aacgtctgcg acctgagcaa 1920

caacatgaat ggtcttcggt ttccgtgttt cgtaaagtct ggaaacgcgg aagtcagcgc 1980

tcttccgctt cctcgctcac tgactcgctg cgctcggtcg ttcggctgcg gcgagcggta 2040

tcagctcact caaaggcggt aatacggtta tccacagaat caggggataa cgcaggaaag 2100

aacatgtgag caaaaggcca gcaaaaggcc agcaaaaggc caggaaccgt aaaaaggccg 2160

cgttgctggc gtttttccat aggctccgcc cccctgacga gcatcacaaa aatcgacgct 2220

caagtcagag gtggcgaaac ccgacaggac tataaagata ccaggcgttt cccccctggaa 2280

gctccctcgt gcgctctcct gttccgaccc tgccgcttac cggatacctg tccgcctttc 2340

tcccttcggg aagcgtggcg ctttctcata gctcacgctg taggtatctc agttcggtgt 2400

aggtcgttcg ctccaagctg ggctgtgtgc acgaaccccc cgttcagccc gaccgctgcg 2460

ccttatccgg taactatcgt cttgagtcca acccggtaag acacgactta tcgccactgg 2520

cagcagccac tggtaacagg attagcagag cgaggtatgt aggcggtgct acagagttct 2580

tgaagtggtg gcctaactac ggctacacta gaaggacagt atttggtatc tgcgctctgc 2640

tgaagccagt taccttcgga aaaagagttg gtagctcttg atccggcaaa caaaccaccg 2700

ctggtagcgg tggttttttt gtttgcaagc agcagattac gcgcagaaaa aaaggatctc 2760

aagaagatcc tttgatcttt tctacggggt ctgacgctca gtggaacgaa aactcacgtt 2820

aagggatttt ggtcatgaga ttatcaaaaa ggatcttcac ctagatcctt ttaaattaaa 2880

aatgaagttt taaatcaatc taaagtatat atgagtaaac ttggtctgac agttaccaat 2940

gcttaatcag tgaggcacct atctcagcga tctgtctatt tcgttcatcc atagttgcct 3000

gactccccgt cgtgtagata actacgatac gggagggctt accatctggc cccagtgctg 3060

caatgatacc gcgagaccca cgctcaccgg ctccagattt atcagcaata aaccagccag 3120

ccggaagggc cgagcgcaga agtggtcctg caactttatc cgcctccatc cagtctatta 3180

attgttgccg ggaagctaga gtaagtagtt cgccagttaa tagtttgcgc aacgttgttg 3240

ccattgctgc aggcatcgtg gtgtcacgct cgtcgtttgg tatggcttca ttcagctccg 3300

gttcccaacg atcaaggcga gttacatgat cccccatgtt gtgcaaaaaa gcggttagct 3360

ccttcggtcc tccgatcgtt gtcagaagta agttggccgc agtgttatca ctcatggtta 3420

tggcagcact gcataattct cttactgtca tgccatccgt aagatgcttt tctgtgactg 3480

gtgagtactc aaccaagtca ttctgagaat agtgtatgcg gcgaccgagt tgctcttgcc 3540

cggcgtcaac acgggataat accgcgccac atagcagaac tttaaaagtg ctcatcattg 3600

gaaaacgttc ttcggggcga aaactctcaa ggatcttacc gctgttgaga tccagttcga 3660

tgtaacccac tcgtgcaccc aactgatctt cagcatcttt tactttcacc agcgtttctg 3720

ggtgagcaaa aacaggaagg caaaatgccg caaaaaaggg aataagggcg acacggaaat 3780

gttgaatact catactcttc ctttttcaat attattgaag catttatcag ggttattgtc 3840

tcatgagcgg atacatattt gaatgtattt agaaaaataa acaaataggg gttccgcgca 3900

catttccccg aaaagtgcca cctgacgtct aagaaaccat tattatcatg acattaacct 3960

ataaaaatag gcgtatcacg aggccctttc gtcttcaaga attcggccgg ccgtttaaac 4020

atctagatgg tacctcatgt ttgaccgctt atcatcgcga taagctttaa tgcggtagtt 4080

tatcacagtt aaattgctaa cgcagtcagg caccgtgtat gaaatctaac aatgcgctca 4140

tcgtcatcct cggcaccgtc accctggatg ctgtaggcat aggcttggtt atgccggtac 4200

tgccgggcct cttgcgggat atcgtccatt ccgacagcat cgccagtcac tatggcgtgc 4260

tactagcgct atatgcgttg atgcaatttc tatgcgcacc cgttctcgga gcactgtccg 4320

accgctttgg ccgccgccca gtcctgctcg cttcgctact tggagccact atcgactacg 4380

cgatcatggc gaccacaccc gtcctgtgga tctgcctcgc tggcctgccg cagttcttca 4440

acctcccggc gcagcttttc gttctcaatt tcagcatccc tttcggcata ccattttatg 4500

acggcggcag agtcataaag cacctcatta cccttgccac cgcctcgcag aacgggcatt 4560

ccctgttcct gccagttctg aatggtacgg atactcgcac cgaaaatgtc agccagctgc 4620

tttttgttga cttccattgt tcattccacg gacaaaaaca gagaaaggaa acgacagagg 4680

ccaaaaagct cgctttcagc acctgtcgtt tcctttcttt tcagagggta ttttaaataa 4740

aaacattaag ttatgacgaa gaagaacgga aacgccttaa accggaaaat tttcataaat 4800

agcgaaaacc cgcgaggtcg ccgccccgta acaaggcgga tcgccggaaa ggacccgcaa 4860

atgataataa ttatcaattg catactatcg acggcactgc tgccagataa caccaccggg 4920

gaaacattcc atcatgatgg ccgtgcggac ataggaagcc agttcatcca tcgctttctt 4980

gtctgctgcc atttgctttg tgacatccag cgccgcacat tcagcagcgt ttttcagcgc 5040

gttttcgatc aacgtttcaa tgttggtatc aacaccaggt ttaactttga acttatcggc 5100

actgacggtt accttgttct gcgctggctc atcacgctgg ataccaaggc tgatgttgta 5160

gatattggtc accggctgag gtgtttcgat tgccgctgcg tggatagcac catttgcgat 5220

agcggcgtcc ttgatgaatg acactccatt gcgaataagt tcgaaggaga cggtgtcacg 5280

aatgcgctgg tccagctcgt cgattgcctt ttgtgcagca gaggtatcaa tctcaacgcc 5340

aagcgtcatc gaagcgcaat attgctgctc accaaaacgc gtattgacca ggtgttcaac 5400

ggcaaatttc tgcccttctg atgtcagaaa ggtaaagtga ttttctttct ggtattcagt 5460

tgctgtgtgt ctggtttcag caaaaccaag ctcgcgcaat tcggctgtgc cagatttaga 5520

aggcagatca ccagacagca acgcgccacg gaaaaacagc gcatacagaa catccgtcgc 5580

cgcgccggac aacgtgataa ttttatgacc catgatttat ttccttttag acgtgagcct 5640
gtcgcacagc aaagccgccg aaagttaacg gtttgcccag gctcacaact gaaagacttt 5700
ctacggtgtg cgcgtgcgat gcgcgtagaa gactgattta tcaacctgtc tttatatcag 5760
gattcattac ctgactattt gtgggtaaag ttcgtagtgc gctgatcgtg caaaatgatt 5820
ttagttggga acagttcgca actctgtccc ataaaaatca gcatattccc atctatccca 5880
tatccagcgc attgaccatc gggatactga agggagattc catcatctct tagaaagatc 5940
accatctctt ttgtttcaat ttgcatatag ctacctggag gatttatgaa tacaaggatt 6000

ttcatggact attaccatga gattgatttt ccatctttat tcgcgagagc agtggaaagc 6060
gatgacgatg tgggtactac attgcgcatt cacctacttt gtgagcgcat ggtcgaagca 6120
tggatatgcg catgctgtga ctgccaagat cctctacgcc ggacgcatcg tggccggcat 6180
caccggcgcc acaggtgcgg ttgctggcgc ctatatcgcc gacatcaccg atggggaaga 6240
tcgggctcgc cacttcgggc tcatgagcgc ttgtttcggc gtgggtatgg tggcaggccc 6300
cgtggccggg ggactgttgg gcgccatctc cttgcatgca ccattccttg cggcggcggt 6360
gctcaacggc ctcaacctac tactgggctg cttcctaatg caggagtcgc ataagggaga 6420

gcgtcgaccg atgcccttga gagccttcaa cccagtcagc tccttccggt gggcgcgggg 6480
catgactatc gtcgccgcac ttatgactgt cttctttatc atgcaactcg taggacaggt 6540
gccggcagcg ctctgggtca ttttcggcga ggaccgcttt cgctggagcg cgacgatgat 6600
cggcctgtcg cttgcggtat tcggaatctt gcacgccctc gctcaagcct tcgtcactgg 6660
tcccgccacc aaacgtttcg gcgagaagca ggccattatc gccggcatgg cggccgacgc 6720
gctgggctac gtcttgctgg cgttcgcgac gcgaggctgg atggccttcc ccattatgat 6780
tcttctcgct tccggcggca tcgggatgcc cgcgttgcag gccatgctgt ccaggcaggt 6840

agatgacgac catcagggac agcttcaagg atcgctcgcg gctcttacca gcctaacttc 6900
gatcattgga ccgctgatcg tcacggcgat ttatgccgcc tcggcgagca catggaacgg 6960
gttggcatgg attgtaggcg ccgccctata ccttgtctgc ctccccgcgt tgcgtcgcgg 7020
tgcatggagc cgggccacct cgacctgaat ggaagccggc ggcacctcgc taacggattc 7080
accactccaa gaattggagc caatcaattc ttgcggagaa ctgtgaatgc gcaaaccaac 7140
ccttggcaga acatatccat cgcgtccgcc atctccagca gccgcacgcg gcgcatctcg 7200
ggcagcgttg ggtcctgcag atccggctgt ggaatgtgtg tcagttaggg tgtggaaagt 7260

ccccaggctc cccagcaggc agaagtatgc aaagcatgca tctcaattag tcagcaacca 7320

ggtgtggaaa gtccccaggc tccccagcag gcagaagtat gcaaagcatg catctcaatt 7380

agtcagcaac catagtcccg cccctaactc cgcccatccc gcccctaact ccgcccagtt 7440

ccgcccattc tccgccccat ggctgactaa ttttttttat ttatgcagag gccgaggccg 7500

cctcggcctc tgagctattc cagaagtagt gaggaggctt ttttggaggc ctaggctttt 7560

gcaaaaagct tcacgctgcc gcaagcactc agggcgcaag ggctgctaaa ggaagcggaa 7620

cacgtagaaa gccagtccgc agaaacggtg ctgaccccgg atgaatgtca gctactgggc 7680

tatctggaca agggaaaacg caagcgcaaa gagaaagcag gtagcttgca gtgggcttac 7740

atggcgatag ctagactggg cggttttatg gacagcaagc gaaccggaat tgccagctgg 7800

ggcgccctct ggtaaggttg ggaagccctg caaagtaaac tggatggctt tcttgccgcc 7860

aaggatctga tggcgcaggg gatcaagatc tgatcaagag acaggatgag gatcgtttcg 7920

catgattgaa caagatggat tgcacgcagg ttctccggcc gcttgggtgg agaggctatt 7980

cggctatgac tgggcacaac agacaatcgg ctgctctgat gccgccgtgt tccggctgtc 8040

agcgcagggg cgcccggttc tttttgtcaa gaccgacctg tccggtgccc tgaatgaact 8100

gcaggacgag gcagcgcggc tatcgtggct ggccaattgc aagccatcaa caaagcccgc 8160

caagagtttc tgctacgaaa gggacggggg gtttacttgg acccccagtc cggcgaggag 8220

ctcaacccaa tccccccgcc gccgcagccc tatcagcagc agccgcgggc ccttgcttcc 8280

caggatggca cccaaaaaga agctgcagct gccgccgcca cccacggacg aggaggaata 8340

ctgggacagt caggcagagg aggttttgga cgaggaggag gaggacatga tggaagactg 8400

ggagagccta gacgaggaag cttccgaggt cgaagaggtg tcagacgaaa caccgtcacc 8460

ctcggtcgca ttcccctcgc cggcgcccca gaaatcggca accggttcca gcatggctac 8520

aacctccgct cctcaggcgc cgccggcact gcccgttcgc cgacccaacc gtagatggga 8580

caccactgga accagggccg gtaagtccaa gcagccgccg ccgttagccc aagagcaaca 8640

acagcgccaa ggctaccgct catggcgcgg gcacaagaac gccatagttg cttgcttgca 8700

agactgtggg ggcaacatct ccttcgcccg ccgctttctt ctctaccatc acggcgtggc 8760

cttcccccgt aacatcctgc attactaccg tcatctctac agcccatact gcaccggcgg 8820

cagcggcagc ggcagcaaca gcagcggcca cacagaagca aaggcgaccg gatagcaaga 8880

ctctgacaaa gcccaagaaa tccacagcgg cggcagcagc aggaggagga gcgctgcgtc 8940

tggcgcccaa cgaacccgta tcgacccgcg agcttagaaa caggattttt cccactctgt 9000

atgctatatt tcaacagagc aggggccaag aacaagagct gaaaataaaa aacaggtctc 9060

tgcgatccct cacccgcagc tgcctgtatc acaaaagcga agatcagctt cggcgcacgc 9120

tggaagacgc ggaggctctc ttcagtaaat actgcgcgct gactcttaag gactagtttc 9180

gcgccctttc tcaaatttaa gcgcgaaaac tacgtcatct ccagcggcca cacccggcgc 9240

cagcacctgt cgtcagcgcc attatgagca aggaaattcc cacgccctac atgtggagtt 9300

accagccaca aatgggactt gcggctggag ctgcccaaga ctactcaacc cgaataaact 9360

acatgagcgc gggaccccac atgatatccc gggtcaacgg aatccgcgcc caccgaaacc 9420

gaattctctt ggaacaggcg gctattacca ccacacctcg taataacctt aatccccgta 9480

gttggcccgc tgccctggtg taccaggaaa gtcccgctcc caccactgtg gtacttccca 9540

gagacgccca ggccgaagtt cagatgacta actcaggggc gcagcttgcg ggcggctttc 9600

gtcacagggt gcggtcgccc gggcagggta taactcacct gacaatcaga gggcgaggta 9660

ttcagctcaa cgacgagtcg gtgagctcct cgcttggtct ccgtccggac gggacatttc 9720

agatcggcgg cgccggccgt ccttcattca cgcctcgtca ggcaatccta actctgcaga 9780

cctcgtcctc tgagccgcgc tctggaggca ttggaactct gcaatttatt gaggagtttg 9840

tgccatcggt ctactttaac cccttctcgg gacctcccgg ccactatccg gatcaattta 9900

ttcctaactt tgacgcggta aaggactcgg cggacggcta cgactgaatg ttaagtggag 9960

aggcagagca actgcgcctg aaacacctgg tccactgtcg ccgccacaag tgctttgccc 10020

gcgactccgg tgagttttgc tactttgaat tgcccgagga tcatatcgag ggcccggcgc 10080

acggcgtccg gcttaccgcc cagggagagc ttgcccgtag cctgattcgg gagtttaccc 10140

agcgcccccct gctagttgag cgggacaggg gaccctgtgt tctcactgtg atttgcaact 10200

gtcctaaccc tggattacat caagatcctc tagttaatgg atccaatcga tatcgcgaag 10260

tcgatcgacg gatcgggaga tctcccgatc ccctatggtg cactctcagt acaatctgct 10320

ctgatgccgc atagttaagc cagtatctgc tccctgcttg tgtgttggag gtcgctgagt 10380

agtgcgcgag caaaatttaa gctacaacaa ggcaaggctt gaccgacaat tgcatgaaga 10440

atctgcttag ggttaggcgt tttgcgctgc ttcgcgatgt acgggccaga tatacgcgtt 10500

gacattgatt attgactagt tattaatagt aatcaattac ggggtcatta gttcatagcc 10560

catatatgga gttccgcgtt acataactta cggtaaatgg cccgcctggc tgaccgccca 10620

acgacccccg cccattgacg tcaataatga cgtatgttcc catagtaacg ccaataggga 10680

ctttccattg acgtcaatgg gtggagtatt tacggtaaac tgcccacttg gcagtacatc 10740

aagtgtatca tatgccaagt acgcccccta ttgacgtcaa tgacggtaaa tggcccgcct 10800

ggcattatgc ccagtacatg accttatggg actttcctac ttggcagtac atctacgtat 10860

tagtcatcgc tattaccatg gtgatgcggt tttggcagta catcaatggg cgtggatagc 10920

ggtttgactc acggggattt ccaagtctcc accccattga cgtcaatggg agtttgtttt 10980

ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa ctccgcccca ttgacgcaaa 11040

tgggcggtag gcgtgtacgg tgggaggtct atataagcag agctctctgg ctaactagag 11100
aacccactgc ttactggctt atcgaaatta atacgactca ctatagggag acccaagctg 11160
gctagcgttt aaacttaagc ttcgagctcg gatccaccgg tcgccaccat ggtgagcaag 11220
ggcgaggagc tgttcaccgg ggtggtgccc atcctggtcg agctggacgg cgacgtaaac 11280
ggccacaagt tcagcgtgtc cggcgagggc gagggcgatg ccacctacgg caagctgacc 11340
ctgaagttca tctgcaccac cggcaagctg cccgtgccct ggcccaccct cgtgaccacc 11400
ctgacctacg gcgtgcagtg cttcagccgc taccccgacc acatgaagca gcacgacttc 11460

ttcaagtccg ccatgcccga aggctacgtc caggagcgca ccatcttctt caaggacgac 11520
ggcaactaca agacccgcgc cgaggtgaag ttcgagggcg acaccctggt gaaccgcatc 11580
gagctgaagg gcatcgactt caaggaggac ggcaacatcc tggggcacaa gctggagtac 11640
aactacaaca gccacaacgt ctatatcatg gccgacaagc agaagaacgg catcaaggtg 11700
aacttcaaga tccgccacaa catcgaggac ggcagcgtgc agctcgccga ccactaccag 11760
cagaacaccc ccatcggcga cggccccgtg ctgctgcccg acaaccacta cctgagcacc 11820
cagtccgccc tgagcaaaga ccccaacgag aagcgcgatc acatggtcct gctggagttc 11880

gtgaccgccg ccgggatcac tctcggcatg gacgagctgt acaagtaaag cggccgctcg 11940
agtctagagg gcccgtttaa acccgctgat cagcctcgac tgtgccttct agttgccagc 12000
catctgttgt ttgcccctcc cccgtgcctt ccttgaccct ggaaggtgcc actcccactg 12060
tcctttccta ataaaatgag gaaattgcat cgcattgtct gagtaggtgt cattctattc 12120
tgggggggtgg ggtggggcag gacagcaagg gggaggattg ggaagacaat agcaggcatg 12180
ctggggatgc ggtgggctct atggcttctg aggcggaaag aaccagctgg ggctctaggg 12240
ggtatcccca cgcgccctgt agcggcgcat taagcgcggc gggtgtggtg gttacgcgca 12300

gcgtgaccgc tacacttgcc agcgccctag cgcccgctcc tttcgctttc ttcccttcct 12360
ttctcgccac gttcgccggc tttccccgtc aagctctaaa tcgggggctc cctttagggt 12420
tccgatttag tgctttacgg cacctcgacc ccaaaaaact tgattagggt gatggttcac 12480
atcgaattcc tgcagcccgg gggatccatt aactaataaa aaaaaataat aaagcatcac 12540
ttacttaaaa tcagttagca aatttctgtc cagtttattc agcagcacct ccttgccctc 12600
ctcccagctc tggtattgca gcttcctcct ggctgcaaac tttctccaca atctaaatgg 12660
aatgtcagtt tcctcctgtt cctgtccatc cgcacccact atcttcatgt tgttgcagat 12720

gaagcgcgca agaccgtctg aagatacctt caaccccgtg tatccatatg acacggaaac 12780

cggtcctcca actgtgcctt ttcttactcc tccctttgta tcccccaatg ggtttcaaga 12840

gagtcccct ggagttctta ctttaaaatg tttaacccca ctaacaacca caggcggatc 12900

tctacagcta aaagtgggag ggggacttac agtggatgac actgatggta ccttacaaga 12960

aaacatacgt gctacagcac ccattactaa aaataatcac tctgtagaac tatccattgg 13020

aaatggatta gaaactcaaa acaataaact atgtgccaaa ttgggaaatg ggttaaaatt 13080

taacaacggt gacatttgta taaaggatag tattaacacc ttatggactg gaataaaccc 13140

tccacctaac tgtcaaattg tggaaaacac taatacaaat gatggcaaac ttactttagt 13200

attagtaaaa aacggagggc ttgttaatgg ctacgtgtct ctagttggtg tatcagacac 13260

tgtgaaccaa atgttcacac aaaagacagc aaacatccaa ttaagattat attttgactc 13320

ttctggaaat ctattaactg gagaatcaga cttaaaaatt ccacttaaaa ataaatcttc 13380

tacagcgacc agtgaaactg tagccagcag caaagccttt atgccaagta ctacagctta 13440

tcccttcaac accgctacta gggatagtga aaactacatt catggaatat gttactacat 13500

gactagttat gatagaagtc tatttccctt gaacatttct ataatgctaa acagccgtat 13560

gatttcttcc aatgttgcct atgccataca atttgaatgg aatctaaatg caagtgaatc 13620

tccagaaagc aacatagcta cgctgaccac atcccccttt ttcttttctt acattacaga 13680

agacgacaac taaaatgccc aagaataaag aaagctttgt gttatgtttc aacgtgttta 13740

tttttcaatt gcagaaaatt tcaagtcatt tttcattcag tagtatagcc ccaccaccac 13800

atagcttata cagatcaccg taccttaatc aaactcacag aaccctagta ttcaacctgc 13860

cacctccctc ccaacacaca gagtacacag tcctttctcc ccggctggcc ttaaaaagca 13920

tcatatcatg ggtaacagac atattcttag gtgttatatt ccacacggtt tcctgtcgag 13980

ccaaacgctc atcagtgata ttaataaact ccccgggcag ctcacttaag ttcatgtcgc 14040

tgtccagctg ctgagccaca ggctgctgtc caacttgcgg ttgcttaacg ggcggcgaag 14100

gagaagtcca cgcctacatg ggggtagagt cataatcgtg catcaggata gggcggtggt 14160

gctgcagcag cgcgcgaata aactgctgcc gccgccgctc cgtcctgcag gaatacaaca 14220

tggcagtggt ctcctcagcg atgattcgca ccgcccgcag cataaggcgc cttgtcctcc 14280

gggcacagca gcgcaccctg atctcactta aatcagcaca gtaactgcag cacagcacca 14340

caatattgtt caaaatccca cagtgcaagg cgctgtatcc aaagctcatg gcggggacca 14400

cagaacccac gtgg 14414

<210> 56

<211> 32724

<212> DNA

<213> Artificial Sequence

<220><223> PS1

<400> 56

catcatcaat aatatacctt atagatggaa tggtgccaat atgtaaatga ggtgatttta 60

aaaagtgtgg gccgtgtggt gattggctgt ggggttaacg gttaaaaggg gcggcgcggc 120

cgtgggaaaa tgacgtttta tgggggtgga gttttttgc aagttgtcgc gggaaatgat 180

aacttcgtat agcatacatt atacgaagtt atttacgcat aaaaaggctt cttttctcac 240

ggaactactt agttttccca cggtatttaa caggaaatga ggtagttttg accggatgca 300

agtgaaaatt gctgattttc gcgcgaaaac tgaatgagga agtgtttttc tgaataatgt 360

ggtatttatg gcagggtgat aacttcgtat agcatacatt atacgaagtt atggaatgtt 420

tatgccttac cagtgtaaca tgaatcatgt gaaagtgttg ttggaaccag atgccttttc 480

cagaatgagc ctaacaggaa tctttgacat gaacacgcaa atctggaaga tcctgaggta 540

tgatgatacg agatcgaggg tgcgcgcatg cgaatgcgga ggcaagcatg ccaggttcca 600

gccggtgtgt gtagatgtga ccgaagatct cagaccggat catttggtta ttgcccgcac 660

tggagcagag ttcggatcca gtggagaaga aactgactaa ggtgagtatt gggaaaactt 720

tggggtggga ttttcagatg gacagattga gtaaaaattt gtttttttctg tcttgcagct 780

gacatgagtg gaaatgcttc ttttaagggg ggagtcttca gcccttatct gacagggcgt 840

ctcccatcct gggcaggagt tcgtcagaat gttatgggat ctactgtgga tggaagaccc 900

gttcaacccg ccaattcttc aacgctgacc tatgctactt taagttcttc acctttggac 960

gcagctgcag ccgctgccgc cgcctctgtc gccgctaaca ctgtgcttgg aatgggttac 1020

tatggaagca tcgtggctaa ttccacttcc tctaataacc cttctacact gactcaggac 1080

aagttacttg tccttttggc ccagctggag gctttgaccc aacgtctggg tgaactttct 1140

cagcaggtgg ccgagttgcg agtacaaact gagtctgctg tcggcacggc aaagtctaaa 1200

taaaaaaaat tccagaatca atgaataaat aaacgagctt gttgttgatt taaaatcaag 1260

tgtttttatt tcatttttcg cgcacggtat gccctggacc accgatctcg atcattgaga 1320

actcggtgga ttttttccag aatcctatag aggtgggatt gaatgtttag atacatgggc 1380

attaggccgt ctttggggtg gagatagctc cattgaaggg attcatgctc cggggtagtg 1440

ttgtaaatca cccagtcata acaaggtcgc agtgcatggt gttgcacaat atcttttaga 1500

agtaggctga ttgccacaga taagcccttg gtgtaggtgt ttacaaaccg gttgagctgg 1560

gaggggtgca ttcgaggtga aattatgtgc attttggatt ggatttttaa gttggcaata 1620

ttgccgccaa gatcccgtct tgggttcatg ttatgaagga ctaccaagac ggtgtatccg 1680

gtacatttag gaaatttatc gtgcagcttg gatggaaaag cgtggaaaaa tttggagaca 1740

cccttgtgtc ctccgagatt ttccatgcac tcatccatga taatagcaat ggggccgtgg 1800

gcagcggcgc gggcaaacac gttccgtggg tctgacacat catagttatg ttcctgagtt 1860

aaatcatcat aagccatttt aatgaatttg gggcggagcg taccagattg gggtatgaat 1920

gttccttcgg gccccggagc atagttcccc tcacagattt gcatttccca agctttcagt 1980

tctgagggtg gaatcatgtc cacctggggg gctatgaaga acaccgtttc gggggcgggg 2040

gtgattagtt gggatgatag caagtttctg agcaattgag atttgccaca tccggtgggg 2100

ccataaataa ttccgattac aggttgcagg tggtagttta gggaacggca actgccgtct 2160

tctcgaagca agggggccac ctcgttcatc atttccctta catgcatatt ttcccgcacc 2220

aaatccatta ggaggcgctc tcctcctagt gatagaagtt cttgtagtga ggaaaagttt 2280

ttcagcggtt ttagaccgtc agccatgggc attttggaaa gagtttgctg caaaagttct 2340

agtctgttcc acagttcagt gatgtgttct atggcatctc gatccagcag acctcctcgt 2400

ttcgcgggtt tggacggctc ctggagtagg gtatgagacg atgggcgtcc agcgctgcca 2460

gggttcggtc cttccagggt ctcagtgttc gagtcagggt tgtttccgtc acagtgaagg 2520

ggtgtgcgcc tgcttgggcg cttgccaggg tgcgcttcag actcattctg ctggtggaga 2580

acttctgtcg cttggcgccc tgtatgtcgg ccaagtagca gtttaccatg agttcgtagt 2640

tgagcgcctc ggctgcgtgg cctttggcgc ggagcttacc tttggaagtt ttcttgcata 2700

ccgggcagta taggcatttc agcgcataca gcttgggcgc aaggaaaatg gattctgggg 2760

agtatgcatc cgcgccgcag gaggcgcaaa cagtttcaca ttccaccagc caggttaaat 2820

ccggttcatt ggggtcaaaa acaagttttc cgccatattt tttgatgcgt ttcttacctt 2880

tggtctccat aagttcgtgt cctcgttgag tgacaaacag gctgtccgta tctccgtaga 2940

ctgattttac aggcctcttc tccagtggag tgcctcggtc ttcttcgtac aggaactctg 3000

accactctga tacaaaggcg cgcgtccagg ccagcacaaa ggaggctatg tgggaggggt 3060

agcgatcgtt gtcaaccagg gggtccacct tttccaaagt atgcaaacac atgtcaccct 3120

cttcaacatc caggaatgtg attggcttgt aggtgtattt cacgtgacct ggggtccccg 3180

ctgggggggt ataaaagggg gcggttcttt gctcttcctc actgtcttcc ggatcgctgt 3240

ccaggaacgt cagctgttgg ggtaggtatt ccctctcgaa ggcgggcatg acctctgcac 3300

tcaggttgtc agtttctaag aacgaggagg atttgatatt gacagtgccg gttgagatgc 3360

ctttcatgag gttttcgtcc atttggtcag aaaacacaat ttttttattg tcaagtttgg 3420

tggcaaatga tccatacagg gcgttggata aaagtttggc aatggatcgc atggtttggt 3480

tcttttcctt gtccgcgcgc tctttggcgg cgatgttgag ttggacatac tcgcgtgcca 3540

ggcacttcca ttcggggaag atagttgtta attcatctgg cacgattctc acttgccacc 3600

ctcgattatg caaggtaatt aaatccacac tggtggccac ctcgcctcga aggggttcat 3660

tggtccaaca gagcctacct cctttcctag aacagaaagg gggaagtggg tctagcataa 3720

gttcatcggg agggtctgca tccatggtaa agattcccgg aagtaaatcc ttatcaaaat 3780

agctgatggg agtggggtca tctaaggcca tttgccattc tcgagctgcc agtgcgcgct 3840

catatgggtt aaggggactg ccccagggca tgggatgggt gagagcagag gcatacatgc 3900

cacagatgtc atagacgtag atgggatcct caaagatgcc tatgtaggtt ggatagcatc 3960

gccccctct gatacttgct cgcacatagt catatagttc atgtgatggc gctagcagcc 4020

ccggacccaa gttggtgcga ttgggttttt ctgttctgta gacgatctgg cgaaagatgg 4080

cgtgagaatt ggaagagatg gtgggtcttt gaaaaatgtt gaaatgggca tgaggtagac 4140

ctacagagtc tctgacaaag tgggcataag attcttgaag cttggttacc agttcggcgg 4200

tgacaagtac gtctagggcg cagtagtcaa gtgtttcttg aatgatgtca taacctggtt 4260

ggtttttctt ttcccacagt tcgcggttga gaaggtattc ttcgcgatcc ttccagtact 4320

cttctagcgg aaacccgtct ttgtctgcac ggtaagatcc tagcatgtag aactgattaa 4380

ctgccttgta agggcagcag cccttctcta cgggtagaga gtatgcttga gcagcttttc 4440

gtagcgaagc gtgagtaagg gcaaaggtgt ctctgaccat gactttgaga aattggtatt 4500

tgaagtccat gtcgtcacag gctccctgtt cccagagttg gaagtctacc cgtttcttgt 4560

aggcggggtt gggcaaagcg aaagtaacat cattgaagag aatcttaccg gctctgggca 4620

taaaattgcg agtgatgcgg aaaggctgtg gtacttccgc tcgattgttg atcacctggg 4680

cagctaggac gatttcgtcg aaaccgttga tgttgtgtcc tacgatgtat aattctatga 4740

aacgcggcgt gcctctgacg tgaggtagct tactgagctc atcaaaggtt aggtctgtgg 4800

ggtcagataa ggcgtagtgt tcgagagccc attcgtgcag gtgaggattt gcatgtagga 4860

atgatgacca aagatctacc gccagtgctg tttgtaactg gtcccgatac tgacgaaaat 4920

gccggccaat tgccattttt tctggagtga cacagtagaa ggttctgggg tcttgttgcc 4980

atcgatccca cttgagttta atggctagat cgtgggccat gttgacgaga cgctcttctc 5040

ctgagagttt catgaccagc atgaaaggaa ctagttgttt gccaaaggat cccatccagg 5100

tgtaagtttc cacatcgtag gtcaggaaga gtctttctgt gcgaggatga gagccgatcg 5160

ggaagaactg gatttcctgc caccagttgg aggattggct gttgatgtga tggaagtaga 5220

agtttctgcg gcgcgccgag cattcgtgtt tgtgcttgta cagacggccg cagtagtcgc 5280

agcgttgcac gggttgtatc tcgtgaatga gctgtacctg gcttcccttg acgagaaatt 5340

tcagtgggaa gccgaggcct ggcgattgta tctcgtgctc ttctatattc gctgtatcgg 5400
cctgttcatc ttctgtttcg atggtggtca tgctgacgag cccccgcggg aggcaagtcc 5460
agacctcggc gcgggagggg cggagctgaa ggacgagagc gcgcaggctg gagctgtcca 5520
gagtcctgag acgctgcgga ctcaggttag taggtaggga cagaagatta acttgcatga 5580
tcttttccag ggcgtgcggg aggttcagat ggtacttgat ttccacaggt tcgtttgtag 5640
agacgtcaat ggcttgcagg gttccgtgtc ctttgggcgc cactaccgta cctttgtttt 5700
ttcttttgat cggtggtggc tctcttgctt cttgcatgct cagaagcggt gacggggacg 5760

cgcgccgggc ggcagcggtt gttccggacc cgggggcatg gctggtagtg gcacgtcggc 5820
gccgcgcacg ggcaggttct ggtattgcgc tctgagaaga cttgcgtgcg ccaccacgcg 5880
tcgattgacg tcttgtatct gacgtctctg ggtgaaagct accggccccg tgagcttgaa 5940
cctgaaagag agttcaacag aatcaatttc ggtatcgtta acggcagctt gtctcagtat 6000
ttcttgtacg tcaccagagt tgtcctggta ggcgatctcc gccatgaact gctcgatttc 6060
ttcctcctga agatctccgc gacccgctct ttcgacggtg gccgcgaggt cattggagat 6120
acggcccatg agttgggaga atgcattcat gcccgcctcg ttccagacgc ggctgtaaac 6180

cacggccccc tcggagtctc ttgcgcgcat caccacctga gcgaggttaa gctccacgtg 6240
tctggtgaag accgcatagt tgcataggcg ctgaaaaagg tagttgagtg tggtggcaat 6300
gtgttcggcg acgaagaaat acatgatcca tcgtctcagc ggcatttcgc taacatcgcc 6360
cagagcttcc aagcgctcca tggcctcgta gaagtccacg gcaaaattaa aaaactggga 6420
gtttcgcgcg gacacggtca attcctcctc gagaagacgg atgagttcgg ctatggtggc 6480
ccgtacttcg cgttcgaagg ctcccgggat ctcttcttcc tcttctatct cttcttccac 6540
taacatctct tcttcgtctt caggcggggg cggaggggcc acgcggcgac gtcgacggcg 6600

cacgggcaaa cggtcgatga atcgttcaat gacctctccg cggcggcggc gcatggtttc 6660
agtgacggcg cggccgttct cgcgcggtcg cagagtaaaa acaccgccgc gcatctcctt 6720
aaagtggtga ctgggaggtt ctccgtttgg gagggagagg gcgctgatta tacattttat 6780
taattggccc gtagggactg cgcgcagaga tctgatcgtg tcaagatcca cgggatctga 6840
aaacctttcg acgaaagcgt ctaaccagtc acagtcacaa ggtaggctga gtacggcttc 6900
ttgtgggcgg gggtggttat gtgttcggtc tgggtcttct gtttcttctt catctcggga 6960
aggtgagacg atgctgctgg tgatgaaatt aaagtaggca gttctaagac ggcggatggt 7020

ggcgaggagc accaggtctt tgggtccggc ttgctggata cgcaggcgat tggccattcc 7080

ccaagcatta tcctgacatc tagcaagatc tttgtagtag tcttgcatga gccgttctac 7140

gggcacttct tcctcacccg ttctgccatg catacgtgtg agtccaaatc cgcgcattgg 7200

ttgtaccagt gccaagtcag ctacgactct ttcggcgagg atggcttgct gtacttgggt 7260

aagggtggct tgaaagtcat caaaatccac aaagcggtgg taagcccctg tattaatggt 7320

gtaagcacag ttggccatga ctgaccagtt aactgtctgg tgaccagggc gcacgagctc 7380

ggtgtattta aggcgcgaat aggcgcgggt gtcaaagatg taatcgttgc aggtgcgcac 7440

cagatactgg taccctataa gaaaatgcgg cggtggttgg cggtagagag gccatcgttc 7500

tgtagctgga gcgccagggg cgaggtcttc caacataagg cggtgatagc cgtagatgta 7560

cctggacatc caggtgattc ctgcggcggt agtagaagcc cgaggaaact cgcgtacgcg 7620

gttccaaatg ttgcgtagcg gcatgaagta gttcattgta ggcacggttt gaccagtgag 7680

gcgcgcgcag tcattgatgc tctatagaca cggagaaaat gaaagcgttc agcgactcga 7740

ctccgtagcc tggaggaacg tgaacgggtt gggtcgcggt gtaccccggt tcgagacttg 7800

tactcgagcc ggccggagcc gcggctaacg tggtattggc actcccgtct cgacccagcc 7860

tacaaaaatc caggatacgg aatcgagtcg ttttgctggt ttccgaatgg cagggaagtg 7920

agtcctattt ttttttttt tgccgctcag atgcatcccg tgctgcgaca gatgcgcccc 7980

caacaacagc ccccctcgca gcagcagcag cagcaaccac aaaaggctgt ccctgcaact 8040

actgcaactg ccgccgtgag cggtgcggga cagcccgcct atgatctgga cttggaagag 8100

ggcgaaggac tggcacgtct aggtgcgcct tcgcccgagc ggcatccgcg agttcaactg 8160

aaaaaagatt ctcgcgaggc gtatgtgccc caacagaacc tatttagaga cagaagcggc 8220

gaggagccgg aggagatgcg agcttcccgc tttaacgcgg gtcgtgagct gcgtcacggt 8280

ttggaccgaa gacgagtgtt gcgagacgag gatttcgaag ttgatgaagt gacagggatc 8340

agtcctgcca gggcacacgt ggctgcagcc aaccttgtat cggcttacga gcagacagta 8400

aaggaagagc gtaacttcca aaagtctttt aataatcatg tgcgaaccct gattgcccgc 8460

gaagaagtta cccttggttt gatgcatttg tgggatttga tggaagctat cattcagaac 8520

cctactagca aacctctgac cgcccagctg tttctggtgg tgcaacacag cagagacaat 8580

gaggctttca gagaggcgct gctgaacatc accgaacccg aggggagatg gttgtatgat 8640

cttatcaaca ttctacagag tatcatagtg caggagcgga gcctgggcct ggccgagaag 8700

gtagctgcca tcaattactc ggttttgagc ttgggaaaat attacgctcg caaaatctac 8760

aagactccat acgttcccat agacaaggag gtgaagatag atgggttcta catgcgcatg 8820

acgctcaagg tcttgaccct gagcgatgat cttggggtgt atcgcaatga cagaatgcat 8880

cgcgcggtta gcgccagcag gaggcgcgag ttaagcgaca gggaactgat gcacagtttg 8940

caaagagctc tgactggagc tggaaccgag ggtgagaatt acttcgacat gggagctgac 9000

ttgcagtggc agcctagtcg cagggctctg agcgccgcga cggcaggatg tgagcttcct 9060

tacatagaag aggcggatga aggcgaggag gaagagggcg agtacttgga agactgatgg 9120

cacaacccgt gttttttgct agatggaaca gcaagcaccg gatcccgcaa tgcgggcggc 9180

gctgcagagc cagccgtccg gcattaactc ctcggacgat tggacccagg ccatgcaacg 9240

tatcatggcg ttgacgactc gcaaccccga agcctttaga cagcaacccc aggccaaccg 9300

tctatcggcc atcatggaag ctgtagtgcc ttcccgatct aatcccactc atgagaaggt 9360

cctggccatc gtgaacgcgt tggtggagaa caaagctatt cgtccagatg aggccggact 9420

ggtatacaac gctctcttag aacgcgtggc tcgctacaac agtagcaatg tgcaaaccaa 9480

tttggaccgt atgataacag atgtacgcga agccgtgtct cagcgcgaaa ggttccagcg 9540

tgatgccaac ctgggttcgc tggtggcgtt aaatgctttc ttgagtactc agcctgctaa 9600

tgtgccgcgt ggtcaacagg attatactaa ctttttaagt gctttgagac tgatggtatc 9660

agaagtacct cagagcgaag tgtatcagtc cggtcctgat tacttctttc agactagcag 9720

acagggcttg cagacggtaa atctgagcca agcttttaaa aaccttaaag gtttgtgggg 9780

agtgcatgcc ccggtaggag aaagagcaac cgtgtctagc ttgttaactc cgaactcccg 9840

cctgttatta ctgttggtag ctcctttcac cgacagcggt agcatcgacc gtaattccta 9900

tttgggttac ctactaaacc tgtatcgcga agccataggg caaagtcagg tggacgagca 9960

gacctatcaa gaaattaccc aagtcagtcg cgctttggga caggaagaca ctggcagttt 10020

ggaagccact ctgaacttct tgcttaccaa tcggtctcaa aagatccctc ctcaatatgc 10080

tcttactgcg gaggaggaga ggatccttag atatgtgcag cagagcgtgg gattgtttct 10140

gatgcaagag ggggcaactc cgactgcagc actggacatg acagcgcgaa atatggagcc 10200

cagcatgtat gccagtaacc gacctttcat taacaaactg ctggactact tgcacagagc 10260

tgccgctatg aactctgatt atttcaccaa tgccatctta aacccgcact ggctgccccc 10320

acctggtttc tacacgggcg aatatgacat gcccgaccct aatgacggat ttctgtggga 10380

cgacgtggac agcgatgttt tttcacctct ttctgatcat cgcacgtgga aaaaggaagg 10440

cggtgataga atgcattctt ctgcatcgct gtccggggtc atgggtgcta ccgcggctga 10500

gcccgagtct gcaagtcctt ttcctagtct acccttttct ctacacagtg tacgtagcag 10560

cgaagtgggt agaataagtc gcccgagttt aatgggcgaa gaggagtacc taaacgattc 10620

cttgctcaga ccggcaagag aaaaaaattt cccaaacaat ggaatagaaa gtttggtgga 10680

taaaatgagt agatggaaga cttatgctca ggatcacaga gacgagcctg ggatcatggg 10740

gactacaagt agagcgagcc gtagacgcca gcgccatgac agacagaggg gtcttgtgtg 10800

ggacgatgag gattcggccg atgatagcag cgtgttggac ttgggtggga gaggaagggg 10860

caacccgttt gctcatttgc gccctcgctt gggtggtatg ttgtgaaaaa aaataaaaaa 10920

gaaaaactca ccaaggccat ggcgacgagc gtacgttcgt tcttctttat tatctgtgtc 10980

tagtataatg aggcgagtcg tgctaggcgg agcggtggtg tatccggagg gtcctcctcc 11040

ttcgtacgag agcgtgatgc agcagcagca ggcgacggcg gtgatgcaat ccccactgga 11100

ggctcccttt gtgcctccgc gatacctggc acctacggag ggcagaaaca gcattcgtta 11160

ctcggaactg gcacctcagt acgataccac caggttgtat ctggtggaca acaagtcggc 11220

ggacattgct tctctgaact atcagaatga ccacagcaac ttcttgacca cggtggtgca 11280

gaacaatgac tttaccccta cggaagccag cacccagacc attaactttg atgaacgatc 11340

gcggtggggc ggtcagctaa agaccatcat gcatactaac atgccaaacg tgaacgagta 11400

tatgtttagt aacaagttca aagcgcgtgt gatggtgtcc agaaaacctc ccgacggtgc 11460

tgcagttggg gatacttatg atcacaagca ggatattttg gaatatgagt ggttcgagtt 11520

tactttgcca gaaggcaact tttcagttac tatgactatt gatttgatga acaatgccat 11580

catagataat tacttgaaag tgggtagaca gaatggagtg cttgaaagtg acattggtgt 11640

taagttcgac accaggaact tcaagctggg atgggatccc gaaaccaagt tgatcatgcc 11700

tggagtgtat acgtatgaag ccttccatcc tgacattgtc ttactgcctg gctgcggagt 11760

ggattttacc gagagtcgtt tgagcaacct tcttggtatc agaaaaaaac agccatttca 11820

agagggtttt aagattttgt atgaagattt agaaggtggt aatattccgg ccctcttgga 11880

tgtagatgcc tatgagaaca gtaagaaaga acaaaaagcc aaaatagaag ctgctacagc 11940

tgctgcagaa gctaaggcaa acatagttgc cagcgactct acaagggttg ctaacgctgg 12000

agaggtcaga ggagacaatt ttgcgccaac acctgttccg actgcagaat cattattggc 12060

cgatgtgtct gaaggaacgg acgtgaaact cactattcaa cctgtagaaa aagatagtaa 12120

gaatagaagc tataatgtgt tggaagacaa aatcaacaca gcctatcgca gttggtatct 12180

ttcgtacaat tatggcgatc ccgaaaaagg agtgcgttcc tggacattgc tcaccacctc 12240

agatgtcacc tgcggagcag agcaggttta ctggtcgctt ccagacatga tgaaggatcc 12300

tgtcactttc cgctccacta gacaagtcag taactaccct gtggtgggtg cagagcttat 12360

gcccgtcttc tcaaagagct tctacaacga acaagctgtg tactcccagc agctccgcca 12420

gtccacctcg cttacgcacg tcttcaaccg ctttcctgag aaccagattt taatccgtcc 12480

gccggcgccc accattacca ccgtcagtga aaacgttcct gctctcacag atcacgggac 12540

cctgccgttg cgcagcagta tccggggagt ccaacgtgtg accgttactg acgccagacg 12600

ccgcacctgt ccctacgtgt acaaggcact gggcatagtc gcaccgcgcg tcctttcaag 12660

ccgcactttc taaaaaaaaa atgtccattc ttatctcgcc cagtaataac accggttggg 12720

gtctgcgcgc tccaagcaag atgtacggag gcgcacgcaa acgttctacc caacatcccg 12780

tgcgtgttcg cggacatttt cgcgctccat ggggtgccct caagggccgc actcgcgttc 12840

gaaccaccgt cgatgatgta atcgatcagg tggttgccga cgcccgtaat tatactccta 12900

ctgcgcctac atctactgtg gatgcagtta ttgacagtgt agtggctgac gctcgcaact 12960

atgctcgacg taagagccgg cgaaggcgca ttgccagacg ccaccgagct accactgcca 13020

tgcgagccgc aagagctctg ctacgaagag ctagacgcgt ggggcgaaga gccatgctta 13080

gggcggccag acgtgcagct tcgggcgcca gcgccggcag gtcccgcagg caagcagccg 13140

ctgtcgcagc ggcgactatt gccgacatgg cccaatcgcg aagaggcaat gtatactggg 13200

tgcgtgacgc tgccaccggt caacgtgtac ccgtgcgcac ccgtccccct cgcacttaga 13260

agatactgag cagtctccga tgttgtgtcc cagcggcgag gatgtccaag cgcaaataca 13320

aggaagaaat gctgcaggtt atcgcacctg aagtctacgg ccaaccgttg aaggatgaaa 13380

aaaaaccccg caaaatcaag cgggttaaaa aggacaaaaa agaagaggaa gatggcgatg 13440

atgggctggc ggagtttgtg cgcgagtttg ccccacggcg acgcgtgcaa tggcgtgggc 13500

gcaaagttcg acatgtgttg agacctggaa cttcggtggt ctttacaccc ggcgagcgtt 13560

caagcgctac ttttaagcgt tcctatgatg aggtgtacgg ggatgatgat attcttgagc 13620

aggcggctga ccgattaggc gagtttgctt atggcaagcg tagtagaata acttccaagg 13680

atgagacagt gtcaataccc ttggatcatg gaaatcccac ccctagtctt aaaccggtca 13740

ctttgcagca agtgttaccc gtaactccgc gaacaggtgt taaacgcgaa ggtgaagatt 13800

tgtatcccac tatgcaactg atggtaccca aacgccagaa gttggaggac gttttggaga 13860

aagtaaaagt ggatccagat attcaacctg aggttaaagt gagacccatt aagcaggtag 13920

cgcctggtct gggggtacaa actgtagaca ttaagattcc cactgaaagt atggaagtgc 13980

aaactgaacc cgcaaagcct actgccacct ccactgaagt gcaaacggat ccatggatgc 14040

ccatgcctat tacaactgac gccgccggtc ccactcgaag atcccgacga aagtacggtc 14100

cagcaagtct gttgatgccc aattatgttg tacacccatc tattattcct actcctggtt 14160

accgaggcac tcgctactat cgcagccgaa acagtacctc ccgccgtcgc cgcaagacac 14220

ctgcaaatcg cagtcgtcgc cgtagacgca caagcaaacc gactcccggc gccctggtgc 14280

ggcaagtgta ccgcaatggt agtgcggaac ctttgacact gccgcgtgcg cgttaccatc 14340

cgagtatcat cacttaatca atgttgccgc tgcctccttg cagatatggc cctcacttgt 14400

cgccttcgcg ttcccatcac tggttaccga ggaagaaact cgcgccgtag aagagggatg 14460

ttgggacgcg gaatgcgacg ctacaggcga cggcgtgcta tccgcaagca attgcggggt 14520

ggttttttac cagccttaat tccaattatc gctgctgcaa ttggcgcgat accaggcata 14580

gcttccgtgg cggttcaggc ctcgcaacga cattgacatt ggaaaaaaaa cgtataaata 14640

aaaaaaaata caatggactc tgacactcct ggtcctgtga ctatgttttc ttagagatgg 14700

aagacatcaa tttttcatcc ttggctccgc gacacggcac gaagccgtac atgggcacct 14760

ggagcgacat cggcacgagc caactgaacg ggggcgcctt caattggagc agtatctgga 14820

gcgggcttaa aaattttggc tcaaccataa aaacatacgg gaacaaagct tggaacagca 14880

gtacaggaca ggcgcttaga aataaactta aagaccagaa cttccaacaa aaagtagtcg 14940

atgggatagc ttccggcatc aatggagtgg tagatttggc taaccaggct gtgcagaaaa 15000

agataaacag tcgtttggac ccgccgccag caaccccagg tgaaatgcaa gtggaggaag 15060

aaattcctcc gccagaaaaa cgaggcgaca agcgtccgcg tcccgatttg gaagagacgc 15120

tggtgacgcg cgtagatgaa ccgccttctt atgaggaagc aacgaagctt ggaatgccca 15180

ccactagacc gatagcccca atggccaccg gggtgatgaa accttctcag ttgcatcgac 15240

ccgtcacctt ggatttgccc cctccccctg ctgctactgc tgtacccgct tctaagcctg 15300

tcgctgcccc gaaaccagtc gccgtagcca ggtcacgtcc cgggggcgct cctcgtccaa 15360

atgcgcactg gcaaaatact ctgaacagca tcgtgggtct aggcgtgcaa agtgtaaaac 15420

gccgtcgctg cttttaatta aatatggagt agcgcttaac ttgcctatct gtgtatatgt 15480

gtcattacac gccgtcacag cagcagagga aaaaaggaag aggtcgtgcg tcgacgctga 15540

gttactttca agatggccac cccatcgatg ctgccccaat gggcatacat gcacatcgcc 15600

ggacaggatg cttcggagta cctgagtccg ggtctggtgc agttcgcccg cgccacagac 15660

acctacttca atctgggaaa taagtttaga aatcccaccg tagcgccgac ccacgatgtg 15720

accaccgacc gtagccagcg gctcatgttg cgcttcgtgc ccgttgaccg ggaggacaat 15780

acatactctt acaaagtgcg gtacaccctg gccgtgggcg acaacagagt gctggatatg 15840

gccagcacgt tctttgacat taggggcgtg ttggacagag gtcccagttt caaaccctat 15900

tctggtacgg cttacaactc tctggctcct aaaggcgctc caaatgcatc tcaatggatt 15960

gcaaaaggcg taccaactgc agcagccgca ggcaatggtg aagaagaaca tgaaacagag 16020

gagaaaactg ctacttacac ttttgccaat gctcctgtaa aagccgaggc tcaaattaca 16080

aaagagggct taccaatagg tttggagatt tcagctgaaa acgaatctaa acccatctat 16140

gcagataaac tttatcagcc agaacctcaa gtgggagatg aaacttggac tgacctagac 16200

ggaaaaaccg aagagtatgg aggcagggct ctaaagccta ctactaacat gaaaccctgt 16260

tacgggtcct atgcgaagcc tactaattta aaaggtggtc aggcaaaacc gaaaaactcg 16320
gaaccgtcga gtgaaaaaat tgaatatgat attgacatgg aatttttga taactcatcg 16380
caaagaacaa acttcagtcc taaaattgtc atgtatgcag aaaatgtagg tttggaaacg 16440
ccagacactc atgtagtgta caaacctgga acagaagaca caagttccga agctaatttg 16500
ggacaacagt ctatgcccaa cagacccaac tacattggct tcagagataa ctttattgga 16560
ctcatgtact ataacagtac tggtaacatg ggggtgctgg ctggtcaagc gtctcagtta 16620
aatgcagtgg ttgacttgca ggacagaaac acagaacttt cttaccaact cttgcttgac 16680

tctctgggcg acagaaccag atactttagc atgtggaatc aggctgtgga cagttatgat 16740
cctgatgtac gtgttattga aaatcatggt gtggaagatg aacttcccaa ctattgtttt 16800
ccactggacg gcataggtgt tccaacaacc agttacaaat caatagttcc aaatggagaa 16860
gataataata attggaaaga acctgaagta aatggaacaa gtgagatcgg acagggtaat 16920
ttgtttgcca tggaaattaa ccttcaagcc aatctatggc gaagtttcct ttattccaat 16980
gtggctctgt atctcccaga ctcgtacaaa tacaccccgt ccaatgtcac tcttccagaa 17040
aacaaaaaca cctacgacta catgaacggg cgggtggtgc cgccatctct agtagacacc 17100

tatgtgaaca ttggtgccag gtggtctctg gatgccatgg acaatgtcaa cccattcaac 17160
caccaccgta acgctggctt gcgttaccga tctatgcttc tgggtaacgg acgttatgtg 17220
cctttccaca tacaagtgcc tcaaaaattc ttcgctgtta aaaacctgct gcttctccca 17280
ggctcctaca cttatgagtg gaactttagg aaggatgtga acatggttct acagagttcc 17340
ctcggtaacg acctgcgggt agatggcgcc agcatcagtt tcacgagcat caacctctat 17400
gctacttttt tccccatggc tcacaacacc gcttccaccc ttgaagccat gctgcggaat 17460
gacaccaatg atcagtcatt caacgactac ctatctgcag ctaacatgct ctaccccatt 17520

cctgccaatg caaccaatat tcccatttcc attccttctc gcaactgggc ggctttcaga 17580
ggctggtcat ttaccagact gaaaaccaaa gaaactccct ctttggggtc tggatttgac 17640
ccctactttg tctattctgg ttctattccc tacctggatg gtaccttcta cctgaaccac 17700
acttttaaga aggtttccat catgtttgac tcttcagtga gctggcctgg aaatgacagg 17760
ttactatctc ctaacgaatt tgaaataaag cgcactgtgg atggcgaagg ctacaacgta 17820
gcccaatgca acatgaccaa agactggttc ttggtacaga tgctcgccaa ctacaacatc 17880
ggctatcagg gcttctacat tccagaagga tacaaagatc gcatgtattc atttttcaga 17940

aacttccagc ccatgagcag gcaggtggtt gatgaggtca attacaaaga cttcaaggcc 18000

gtcgccatac cctaccaaca caacaactct ggctttgtgg gttacatggc tccgaccatg 18060

cgccaaggtc aaccctatcc cgctaactat ccctatccac tcattggaac aactgccgta 18120

aatagtgtta cgcagaaaaa gttcttgtgt gacagaacca tgtggcgcat accgttctcg 18180

agcaacttca tgtctatggg ggcccttaca gacttgggac agaatatgct ctatgccaac 18240

tcagctcatg ctctggacat gacctttgag gtggatccca tggatgagcc caccctgctt 18300

tatcttctct tcgaagtttt cgacgtggtc agagtgcatc agccacaccg cggcatcatc 18360

gaggcagtct acctgcgtac accgttctcg gccggtaacg ctaccacgta agaagcttct 18420

tgcttcttgc aaatagcagc tgcaaccatg gcctgcggat cccaaaacgg ctccagcgag 18480

caagagctca gagccattgt ccaagacctg ggttgcggac cctatttttt gggaacctac 18540

gataagcgct tcccggggtt catggccccc gataagctcg cctgtgccat tgtaaatacg 18600

gccggacgtg agacgggggg agagcactgg ttggctttcg gttggaaccc acgttctaac 18660

acctgctacc tttttgatcc ttttggattc tcggatgatc gtctcaaaca gatttaccag 18720

tttgaatatg agggtctcct gcgccgcagc gctcttgcta ccaaggaccg ctgtattacg 18780

ctggaaaaat ctacccagac cgtgcagggc ccccgttctg ccgcctgcgg acttttctgc 18840

tgcatgttcc ttcacgcctt tgtgcactgg cctgaccgtc ccatggacgg aaaccccacc 18900

atgaaattgc taactggagt gccaaacaac atgcttcatt ctcctaaagt ccagcccacc 18960

ctgtgtgaca atcaaaaagc actctaccat tttcttaata cccattcgcc ttattttcgc 19020

tctcatcgta cacacatcga aagggccact gcgttcgacc gtatggatgt tcaataatga 19080

ctcatgtaaa caacgtgttc aataaacatc actttatttt tttacatgta tcaaggctct 19140

ggattactta tttatttaca agtcgaatgg gttctgacga gaatcagaat gacccgcagg 19200

cagtgatacg ttgcggaact gatacttggg ttgccacttg aattcgggaa tcaccaactt 19260

gggaaccggt atatcgggca ggatgtcact ccacagcttt ctggtcagct gcaaagctcc 19320

aagcaggtca ggagccgaaa tcttgaaatc acaattagga ccagtgctct gagcgcgaga 19380

gttgcggtac accggattgc agcactgaaa caccatcagc gacggatgtc tcacgcttgc 19440

cagcacggtg ggatctgcaa tcatgcccac atccagatct tcagcattgg caatgctgaa 19500

cggggtcatc ttgcaggtct gcctacccat ggcgggcacc caattaggct tgtggttgca 19560

atcgcagtgc agggggatca gtatcatctt ggcctgatcc tgtctgattc ctggatacac 19620

ggctctcatg aaagcatcat attgcttgaa agcctgctgg gctttactac cctcggtata 19680

aaacatcccg caggacctgc tcgaaaactg gttagctgca cagccggcat cattcacaca 19740

gcagcgggcg tcattgttgg ctatttgcac cacacttctg ccccagcggt tttgggtgat 19800

tttggttcgc tcgggattct cctttaaggc tcgttgtccg ttctcgctgg ccacatccat 19860

ctcgataatc tgctccttct gaatcataat attgccatgc aggcacttca gcttgccctc 19920

ataatcattg cagccatgag gccacaacgc acagcctgta cattcccaat tatggtgggc 19980

gatctgagaa aaagaatgta tcattccctg cagaaatctt cccatcatcg tgctcagtgt 20040

cttgtgacta gtgaaagtta actggatgcc tcggtgctct tcgtttacgt actggtgaca 20100

gatgcgcttg tattgttcgt gttgctcagg cattagttta aaacaggttc taagttcgtt 20160

atccagcctg tacttctcca tcagcagaca catcacttcc atgcctttct cccaagcaga 20220

caccaggggc aagctaatcg gattcttaac agtgcaggca gcagctcctt tagccagagg 20280

gtcatcttta gcgatcttct caatgcttct tttgccatcc ttctcaacga tgcgcacggg 20340

cgggtagctg aaacccactg ctacaagttg cgcctcttct ctttcttctt cgctgtcttg 20400

actgatgtct tgcatgggga tatgtttggt cttccttggc ttctttttgg ggggtatcgg 20460

aggaggagga ctgtcgctcc gttccggaga cagggaggat tgtgacgttt cgctcaccat 20520

taccaactga ctgtcggtag aagaacctga ccccacacgg cgacaggtgt ttttcttcgg 20580

gggcagaggt ggaggcgatt gcgaagggct gcggtccgac ctggaaggcg gatgactggc 20640

agaaacccctt ccgcgttcgg gggtgtgctc cctgtggcgg tcgcttaact gatttccttc 20700

gcggctggcc attgtgttct cctaggcaga gaaacaacag acatggaaac tcagccattg 20760

ctgtcaacat cgccacgagt gccatcacat ctcgtcctca gcgacgagga aaaggagcag 20820

agcttaagca ttccaccgcc cagtcctgcc accacctcta ccctagaaga taaggaggtc 20880

gacgcatctc atgacatgca gaataaaaaa gcgaaagagt ctgagacaga catcgagcaa 20940

gacccgggct atgtgacacc ggtggaacac gaggaagagt tgaaacgctt tctagagaga 21000

gaggatgaaa actgcccaaa acagcgagca gataactatc accaagatgc tggaaatagg 21060

gatcagaaca ccgactacct catagggctt gacggggaag acgcgctcct taaacatcta 21120

gcaagacagt cgctcatagt caaggatgca ttattggaca gaactgaagt gcccatcagt 21180

gtggaagagc tcagctgcgc ctacgagctt aacctttttt cacctcgtac tccccccaaa 21240

cgtcagccaa acggcacctg cgagccaaat cctcgcttaa acttttatcc agcttttgct 21300

gtgccagaag tactggctac ctatcacatc ttttttaaaa atcaaaaaat tccagtctcc 21360

tgccgcgcta atcgcacccg cgccgatgcc ctactcaatc tgggacctgg ttcacgctta 21420

cctgatatag cttccttgga agaggttcca aagatcttcg agggtctggg caataatgag 21480

actcgggccg caaatgctct gcaaaaggga gaaaatggca tggatgagca tcacagcgtt 21540

ctggtggaat tggaaggcga taatgccaga ctcgcagtac tcaagcgaag cgtcgaggtc 21600

acacacttcg catatcccgc tgtcaacctg ccccctaaag tcatgacggc ggtcatggac 21660

cagttactca ttaagcgcgc aagtcccctt tcagaagaca tgcatgaccc agatgcctgt 21720

gatgagggta aaccagtggt cagtgatgag cagctaaccc gatggctggg caccgactct 21780
ccccgggatt tggaagagcg tcgcaagctt atgatggccg tggtgctggt taccgtagaa 21840
ctagagtgtc tccgacgttt ctttaccgat tcagaaacct tgcgcaaact cgaagagaat 21900
ctgcactaca cttttagaca cggctttgtg cggcaggcat gcaagatatc taacgtggaa 21960
ctcaccaacc tggtttccta catgggtatt ctgcatgaga atcgcctagg acaaagcgtg 22020
ctgcacagca cccttaaggg ggaagcccgc cgtgattaca tccgcgattg tgtctatctc 22080
tacctgtgcc acacgtggca aaccggcatg ggtgtatggc agcaatgttt agaagaacag 22140

aacttgaaag agcttgacaa gctcttacag aaatctctta aggttctgtg gacagggttc 22200
gacgagcgca ccgtcgcttc cgacctggca gacctcatct tcccagagcg tctcagggtt 22260
actttgcgaa acggattgcc tgactttatg agccagagca tgcttaacaa ttttcgctct 22320
ttcatcctgg aacgctccgg tatcctgccc gccacctgct gcgcactgcc ctccgacttt 22380
gtgcctctca cctaccgcga gtgccccccg ccgctatgga gtcactgcta cctgttccgt 22440
ctggccaact atctctccta ccactcggat gtgatcgagg atgtgagcgg agacggcttg 22500
ctggagtgcc actgccgctg caatctgtgc acgccccacc ggtccctagc ttgcaacccc 22560

cagttgatga gcgaaaccca gataataggc acctttgaat tgcaaggccc cagcagccaa 22620
ggcgatgggt cttctcctgg gcaaagttta aaactgaccc cgggactgtg gacctccgcc 22680
tacttgcgca agtttgctcc ggaagattac cacccctatg aaatcaagtt ctatgaggac 22740
caatcacagc ctccaaaggc cgaactttcg gcttgcgtca tcacccaggg ggcaattctg 22800
gcccaattgc aagccatcca aaaatcccgc caagaatttc tactgaaaaa gggtaagggg 22860
gtctaccttg acccccagac cggcgaggaa ctcaacacaa ggttccctca ggatgtccca 22920
acgacgagaa aacaagaagt tgaaggtgca gccgccgccc ccagaagata tggaggaaga 22980

ttgggacagt caggcagagg aggcggagga ggacagtctg gaggacagtc tggaggaaga 23040
cagtttggag gaggaaaacg aggaggcaga ggaggtggaa gaagtaaccg ccgacaaaca 23100
gttatcctcg gctgcggaga caagcaacag cgctaccatc tccgctccga gtcgaggaac 23160
ccggcggcgt cccagcagta gatgggacga gaccggacgc ttcccgaacc caaccagcgc 23220
ttccaagacc ggtaagaagg atcggcaggg atacaagtcc tggcgggggc ataagaatgc 23280
catcatctcc tgcttgcatg agtgcggggg caacatatcc ttcacgcggc gctacttgct 23340
attccaccat ggggtgaact ttccgcgcaa tgttttgcat tactaccgtc acctccacag 23400

cccctactat agccagcaaa tcccgacagt ctcgacagat aaagacagcg gcggcgacct 23460

ccaacagaaa accagcagcg gcagttagaa aatacacaac aagtgcagca acaggaggat 23520

taaagattac agccaacgag ccagcgcaaa cccgagagtt aagaaatcgg atctttccaa 23580

ccctgtatgc catcttccag cagagtcggg gtcaagagca ggaactgaaa ataaaaaacc 23640

gatctctgcg ttcgctcacc agaagttgtt tgtatcacaa gagcgaagat caacttcagc 23700

gcactctcga ggacgccgag gctctcttca acaagtactg cgcgctgact cttaaagagt 23760

aggcagcgac cgcgcttatt caaaaaaggc gggaattaca tcatcctcga catgagtaaa 23820

gaaattccca cgccttacat gtggagttat caaccccaaa tgggattggc agcaggcgcc 23880

tcccaggact actccacccg catgaattgg ctcagcgccg ggccttctat gatttctcga 23940

gttaatgata tacgcgccta ccgaaaccaa atacttttgg aacagtcagc tcttaccacc 24000

acgccccgcc aacaccttaa tcccagaaat tggcccgccg ccctagtgta ccaggaaagt 24060

cccgctccca ccactgtatt acttcctcga gacgcccagg ccgaagtcca aatgactaat 24120

gcaggtgcgc agttagctgg cggctccacc ctatgtcgtc acaggcctcg gcataatata 24180

aaacgcctga tgatcagagg ccgaggtatc cagctcaacg acgagtcggt gagctctccg 24240

cttggtctac gaccagacgg aatctttcag attgccggct gcgggagatc ttccttcacc 24300

cctcgtcagg ctgttctgac tttggaaagt tcgtcttcgc aaccccgctc gggcggaatc 24360

gggaccgttc aatttgtaga ggagtttact ccctctgtct acttcaaccc cttctccgga 24420

tctcctgggc actacccgga cgagttcata ccgaacttcg acgcgattag cgagtcagtg 24480

gacggctacg attgatgtct ggtgacgcgg ctgagctatc tcggctgcga catctagacc 24540

actgccgccg ctttcgctgc tttgcccggg aacttattga gttcatctac ttcgaactcc 24600

ccaaggatca ccctcaaggt ccggcccacg gagtgcggat tactatcgaa ggcaaaatag 24660

actctcgcct gcaacgaatt ttctcccagc ggcccgtgct gatcgagcga gaccagggaa 24720

acaccacggt tagtaatcaa ttacggggtc attagttcat agcccatata tggagttaac 24780

ttgtttattg cagcttataa tggttacaaa taaagcaata gcatcacaaa tttcacaaat 24840

aaagcatttt tttcactgca ttctagttgt ggtttgtcca aactcatcaa tgtatcttat 24900

catgtctgct cgaagcggcc ggccgccccg actctagagt cgcggcctca ttaggaagtt 24960

cctatacttt ctagagaata ggaacttctc agaagaactc gtcaagaagg cgatagaagg 25020

cgatgcgctg cgaatcggga gcggcgatac cgtaaagcac gaggaagcgg tcagcccatt 25080

cgccgccaag ctcttcagca atatcacggg tagccaacgc tatgtcctga tagcggtccg 25140

ccacacccag ccggccacag tcgatgaatc cagaaaagcg gccattttcc accatgatat 25200

tcggcaagca ggcatcgcca tgggtcacga cgagatcctc gccgtcgggc atgcgcgcct 25260

tgagcctggc gaacagttcg gctggcgcga gcccctgatg ctcttcgtcc agatcatcct 25320

gatcgacaag accggcttcc atccgagtac gtgctcgctc gatgcgatgt ttcgcttggt 25380

ggtcgaatgg gcaggtagcc ggatcaagcg tatgcagccg ccgcattgca tcagccatga 25440

tggatacttt ctcggcagga gcaaggtgag atgacaggag atcctgcccc ggcacttcgc 25500

ccaatagcag ccagtccctt cccgcttcag tgacaacgtc gagcacagct gcgcaaggaa 25560

cgcccgtcgt ggccagccac gatagccgcg ctgcctcgtc ctgcagttca ttcagggcac 25620

cggacaggtc ggtcttgaca aaaagaaccg ggcgcccctg cgctgacagc cggaacacgg 25680

cggcatcaga gcagccgatt gcctgttgtg cccagtcata gccgaatagc ctctccaccc 25740

aagcggccgg agaacctgcg tgcaatccat cttgttcaat ggccgatccc ataacacccc 25800

ttgtattact gtttatgtaa gcagacagtt ttactgttcg tgatgatata tttttatctt 25860

gtgcaatgta acaggttgtg gccatagcgg gcccgggatt ttcctccacg tccccgcatg 25920

ttagaagact tcccctgccc tcggctctgg aagttcctat actttctaga gaataggaac 25980

ttcccgccag aatgcgttcg cacagccgcc agccggtcac tccgttgatg gttactcgga 26040

acagcaggga gccgtcgggg ttgatcaggc gctcgtcgat aattttgttg ccgttccaca 26100

gggtccctgt tacagtgatc tttttgccgt cgaacacggc gatgccttca tacggccgtc 26160

cgaaatagtc gatcatgttc ggcgtaaccc cgtcgattac cagtgtgcca tagtgcagga 26220

tcaccttaaa gtgatgatca tccacagggt acaccacctt aaaaattttt tcgatctggc 26280

ccatttggtc gccgctcaga ccttcatacg ggatgatgac atggatgtcg atcttcagcc 26340

cattttcacc gctcaggaca atcctttgga tcggagttac ggacaccccg agattctgaa 26400

acaaactgga cacacctccc tgttcaagga cttggtccag gttgtagccg gctgtctgtc 26460

gccagtcccc aacgaaatct tcgagtgtga agaccatgga tccgggcccg gggttttctt 26520

caacgtctcc agcctgcttc agcaggctga agttagtagc tccgcttcct cgagctcgag 26580

atctggcgaa ggcgatgggg gtcttgaagg cgtgctggta ctccacgatg cccagctcgg 26640

tgttgctgtg cagctcctcc acgcggcgga aggcgaacat ggggcccccg ttctgcagga 26700

tgctggggtg gatggcgctc ttgaagtgca tgtggctgtc caccacgaag ctgtagtagc 26760

cgccgtcgcg caggctgaag gtgcgggcga agctgcccac cagcacgtta tcgcccatgg 26820

ggtgcaggtg ctccacggtg gcgttgctgc ggatgatctt gtcggtgaag atcacgctgt 26880

cctcggggaa gccggtgccc accaccttga agtcgccgat cacgcggccg gcctcgtagc 26940

ggtagctgaa gctcacgtgc agcacgccgc cgtcctcgta cttctcgatg cgggtgttgg 27000

tgtagccgcc gttgttgatg gcgtgcagga aggggttctc gtagccgctg gggtaggtgc 27060

cgaagtggta gaagccgtag cccatcacgt ggctcagcag gtaggggctg aaggtcaggg 27120

cgcctttggt gctcttcatc ttgttggtca tgcggccctg ctcgggggtg ccctctccgc 27180

cgcccaccag ctcgaactcc acgccgttca gggtgccggt gatgcggcac tcgatcttca 27240
tggcgggcat ggtggcgacc ggtagcgcta gcggcttcgg taccacgcgt tcgctcgaat 27300
taatcaattc tttgccaaaa tgatgagaca gcacaataac cagcacgttg cccaggagct 27360
gtaggaaaaa gaagaaggca tgaacatggt tagcagaggc tctagagccg ccggtcacac 27420
gccagaagcc gaaccccgcc ctgccccgtc ccccccgaag gcagccgtcc ccccgcggac 27480
agccccgagg ctggagaggg agaaggggac ggcggcgcgg cgacgcacga aggccctccc 27540
cgcccatttc cttcctgccg gggccctccc ggagcccctc aaggctttca cgcagccaca 27600

gaaaagaaac aagccgtcat taaaccaagc gctaattaca gcccggagga gaagggccgt 27660
cccgcccgct cacctgtggg agtaacgcgg tcagtcagag ccggggcggg cggcgcgagg 27720
cggcgcggag cggggcacgg ggcgaaggca acgcagcgac tcccgcccgc cgcgcgcttc 27780
gctttttata gggccgccgc cgccgccgcc tcgccataaa aggaaacttt cggagcgcgc 27840
cgctctgatt ggctgccgcc gcacctctcc gcctcgcccc gccccgcccc tcgccccgcc 27900
ccgccccgcc tggcgcgcgc ccccccccccc cccccgcccc catcgctgca caaaataatt 27960
aaaaaataaa taaatacaaa attgggggtg gggaggggggg ggagatgggg agagtgaagc 28020

agaacgtggg gctcacctcg accatggtaa tagcgatgac taatacgtag atgtactgcc 28080
aagtaggaaa gtcccataag gtcatgtact gggcataatg ccaggcgggc catttaccgt 28140
cattgacgtc aatagggggc gtacttggca tatgatacac ttgatgtact gccaagtggg 28200
cagtttaccg taaatactcc acccattgac gtcaatggaa agtccctatt ggcgttacta 28260
tgggaacata cgtcattatt gacgtcaatg ggcgggggtc gttgggcggt cagccaggcg 28320
ggccatttac cgtaagttat gtaacgcgga acaacaacaa ttgcattcat tttatgtttc 28380
aggttcaggg ggaggtgtgg tcctgcgatt ccatcgagtg cacctacacc ctgctgaaga 28440

ccctatgcgg cctaagagac ctgctaccaa tgaattaaaa aaaaatgatt aataaaaaat 28500
cacttacttg aaatcagcaa taaggtctct gttgaaattt tctcccagca gcacctcact 28560
tccctcttcc caactctggt attctaaacc ccgttcagcg gcatactttc tccatacttt 28620
aaaggggatg tcaaatttta gctcctctcc tgtacccaca atcttcatgt ctttcttccc 28680
agatgaccaa gagagtccgg ctcagtgact ccttcaaccc tgtctacccc tatgaagatg 28740
aaagcacctc ccaacacccc tttataaacc cagggtttat ttccccaaat ggcttcacac 28800
aaagcccaga cggagttctt actttaaaat gtttaacccc actaacaacc acaggcggat 28860

ctctacagct aaaagtggga gggggactta cagtggatga cactgatggt accttacaag 28920

aaaacatacg tgctacagca cccattacta aaaataatca ctctgtagaa ctatccattg 28980

gaaatggatt agaaactcaa aacaataaac tatgtgccaa attgggaaat gggttaaaat 29040

ttaacaacgg tgacatttgt ataaaggata gtattaacac cttatggact ggaataaacc 29100

ctccacctaa ctgtcaaatt gtggaaaaca ctaatacaaa tgatggcaaa cttactttag 29160

tattagtaaa aaatggaggg cttgttaatg gctacgtgtc tctagttggt gtatcagaca 29220

ctgtgaacca aatgttcaca caaaagacag caaacatcca attaagatta tattttgact 29280

cttctggaaa tctattaact gaggaatcag acttaaaaat tccacttaaa aataaatctt 29340

ctacagcgac cagtgaaact gtagccagca gcaaagcctt tatgccaagt actacagctt 29400

atcccttcaa caccactact agggatagtg aaaactacat tcatggaata tgttactaca 29460

tgactagtta tgatagaagt ctatttccct tgaacatttc tataatgcta aacagccgta 29520

tgatttcttc caatgttgcc tatgccatac aatttgaatg gaatctaaat gcaagtgaat 29580

ctccagaaag caacatagct acgctgacca catccccctt tttcttttct tacattacag 29640

aagacgacaa ctaaaataaa gtttaagtgt ttttatttaa aatcacaaaa ttcgagtagt 29700

tattttgcct ccaccttccc atttgacaga atacacagtc ctttctcccc ggctggcctt 29760

aaaaagcatc atatcatggg taacagacat attcttaggt gttatattcc acacggtttc 29820

ctgtcgagcc aaacgctcat cagtgatatt aataaactcc ccgggcagct cacttaagtt 29880

catgtcgctg tccagctgct gagccacagg ctgctgtcca acttgcggtt gcttaacggg 29940

cggcgaagga gaagtccacg cctacatggg ggtagagtca taatcgtgca tcaggatagg 30000

gcggtggtgc tgcagcagcg cgcgaataaa ctgctgccgc cgccgctccg tcctgcagga 30060

atacaacatg gcagtggtct cctcagcgat gattcgcacc gcccgcagca taaggcgcct 30120

tgtcctccgg gcacagcagc gcaccctgat ctcacttaaa tcagcacagt aactgcagca 30180

cagcaccaca atattgttca aaatcccaca gtgcaaggcg ctgtatccaa agctcatggc 30240

ggggaccaca gaacccacgt ggccatcata ccacaagcgc aggtagatta agtggcgacc 30300

cctcataaac acgctggaca taaacattac ctcttttggc atgttgtaat tcaccacctc 30360

ccggtaccat ataaacctct gattaaacat ggcgccatcc accaccatcc taaaccagct 30420

ggccaaaacc tgcccgccgg ctatacactg cagggaaccg ggactggaac aatgacagtg 30480

gagagcccag gactcgtaac catggatcat catgctcgtc atgatatcaa tgttggcaca 30540

acacaggcac acgtgcatac acttcctcag gattacaagc tcctcccgcg ttagaaccat 30600

atcccaggga acaacccatt cctgaatcag cgtaaatccc acactgcagg gaagacctcg 30660

cacgtaactc acgttgtgca ttgtcaaagt gttacattcg ggcagcagcg gatgatcctc 30720

cagtatggta gcgcgggttt ctgtctcaaa aggaggtaga cgatccctac tgtacggagt 30780

gcgccgagac aaccgagatc gtgttggtcg tagtgtcatg ccaaatggaa cgccggacgt 30840

agtcattctc gtattttgta tagcaaaacg cggccctggc agaacacact cttcttcgcc 30900

ttctatcctg ccgcttagcg tgttccgtgt gatagttcaa gtacagccac actcttaagt 30960

tggtcaaaag aatgctggct tcagttgtaa tcaaaactcc atcgcatcta attgttctga 31020

ggaaatcatc cacggtagca tatgcaaatc ccaaccaagc aatgcaactg gattgcgttt 31080

caagcaggag aggagaggga agagacggaa gaaccatgtt aatttttatt ccaaacgatc 31140

tcgcagtact tcaaattgta gatcgcgcag atggcatctc tcgcccccac tgtgttggtg 31200

aaaaagcaca gctaaatcaa aagaaatgcg attttcaagg tgctcaacgg tggcttccaa 31260

caaagcctcc acgcgcacat ccaagaacaa aagaatacca aaagaaggag cattttctaa 31320

ctcctcaatc atcatattac attcctgcac cattcccaga taattttcag ctttccagcc 31380

ttgaattatt cgtgtcagtt cttgtggtaa atccaatcca cacattacaa acaggtcccg 31440

gagggcgccc tccaccacca ttcttaaaca caccctcata atgacaaaat atcttgctcc 31500

tgtgtcacct gtagcgaatt gagaatggca acatcaattg acatgccctt ggctctaagt 31560

tcttctttaa gttctagttg taaaaactct ctcatattat caccaaactg cttagccaga 31620

agccccccgg gaacaagagc aggggacgct acagtgcagt acaagcgcag acctccccaa 31680

ttggctccag caaaaacaag attggaataa gcatattggg aaccaccagt aatatcatcg 31740

aagttgctgg aaatataatc aggcagagtt tcttgtagaa attgaataaa agaaaaattt 31800

gccaaaaaaa cattcaaaac ctctgggatg caaatgcaat aggttaccgc gctgcgctcc 31860

aacattgtta gttttgaatt agtctgcaaa aataaaaaaa aaacaagcgt catatcatag 31920

tagcctgacg aacaggtgga taaatcagtc tttccatcac aagacaagcc acagggtctc 31980

cagctcgacc ctcgtaaaac ctgtcatcgt gattaaacaa cagcaccgaa agttcctcgc 32040

ggtgaccagc atgaataagt cttgatgaag catacaatcc agacatgtta gcatcagtta 32100

aggagaaaaa acagccaaca tagcctttgg gtataattat gcttaatcgt aagtatagca 32160

aagccacccc tcgcggatac aaagtaaaag gcacaggaga ataaaaaata taattatttc 32220

tctgctgctg tttaggcaac gtcgcccccg gtccctctaa atacacatac aaagcctcat 32280

cagccatggc ttaccagaga aagtacagcg ggcacacaaa ccacaagctc taaagtcact 32340

ctccaacctc tccacaatat atatacacaa gccctaaact gacgtaatgg gactaaagtg 32400

taaaaaatcc cgccaaaccc aacacacacc ccgaaactgc gtcaccaggg aaaagtacag 32460

tttcacttcc gcaatcccaa caagcgtcac ttcctctttc tcacggtacg tcacatccca 32520

ttaacttaca acgtcatttt cccacggccg cgccgcccct tttaaccgtt aaccccacag 32580

ccaatcacca cacggcccac acttttttaaa atcacctcat ttacatattg gcaccattcc 32640

atctataagg tatattattg atgatggcca agctatttag gtgacactat agaatactca 32700

agctatgcat caagcttggt accg 32724

<210> 57

<211> 32724

<212> DNA

<213> Artificial Sequence

<220><223> PS2

<400> 57

catcatcaat aatatacctt atagatggaa tggtgccaat atgtaaatga ggtgatttta 60

aaaagtgtgg gccgtgtggt gattggctgt ggggttaacg gttaaaaggg gcggcgcggc 120

cgtgggaaaa tgacgtttta tgggggtgga gttttttgc aagttgtcgc gggaaatgat 180

aacttcgtat agcatacatt atacgaagtt attagacttt gacccattac gtggaggttt 240

cgattaccgt gtttttacc tgaatttccg cgtaccgtgt caaagtcttc tgtttttacg 300

taggtgtcag ctgatcgcta gggtatttat aacttcgtat agcatacatt atacgaagtt 360

atggaatgtt tatgccttac cagtgtaaca tgaatcatgt gaaagtgttg ttggaaccag 420

atgccttttc cagaatgagc ctaacaggaa tctttgacat gaacacgcaa atctggaaga 480

tcctgaggta tgatgatacg agatcgaggg tgcgcgcatg cgaatgcgga ggcaagcatg 540

ccaggttcca gccggtgtgt gtagatgtga ccgaagatct cagaccggat catttggtta 600

ttgcccgcac tggagcagag ttcggatcca gtggagaaga aactgactaa ggtgagtatt 660

gggaaaactt tggggtggga ttttcagatg gacagattga gtaaaaattt gttttttctg 720

tcttgcagct gacatgagtg gaaatgcttc ttttaagggg ggagtcttca gcccttatct 780

gacagggcgt ctcccatcct gggcaggagt tcgtcagaat gttatgggat ctactgtgga 840

tggaagaccc gttcaacccg ccaattcttc aacgctgacc tatgctactt taagttcttc 900

acctttggac gcagctgcag ccgctgccgc cgcctctgtc gccgctaaca ctgtgcttgg 960

aatgggttac tatggaagca tcgtggctaa ttccacttcc tctaataacc cttctacact 1020

gactcaggac aagttacttg tccttttggc ccagctggag gctttgaccc aacgtctggg 1080

tgaactttct cagcaggtgg ccgagttgcg agtacaaact gagtctgctg tcggcacggc 1140

aaagtctaaa taaaaaaaat tccagaatca atgaataaat aaacgagctt gttgttgatt 1200

taaaatcaag tgtttttatt tcatttttcg cgcacggtat gccctggacc accgatctcg 1260

atcattgaga actcggtgga ttttttccag aatcctatag aggtgggatt gaatgtttag 1320

atacatgggc attaggccgt ctttggggtg gagatagctc cattgaaggg attcatgctc 1380

cggggtagtg ttgtaaatca cccagtcata acaaggtcgc agtgcatggt gttgcacaat 1440

atcttttaga agtaggctga ttgccacaga taagcccttg gtgtaggtgt ttacaaaccg 1500

gttgagctgg gaggggtgca ttcgaggtga aattatgtgc attttggatt ggatttttaa 1560

gttggcaata ttgccgccaa gatcccgtct tgggttcatg ttatgaagga ctaccaagac 1620

ggtgtatccg gtacatttag gaaatttatc gtgcagcttg gatggaaaag cgtggaaaaa 1680

tttggagaca cccttgtgtc ctccgagatt ttccatgcac tcatccatga taatagcaat 1740

ggggccgtgg gcagcggcgc gggcaaacac gttccgtggg tctgacacat catagttatg 1800

ttcctgagtt aaatcatcat aagccatttt aatgaatttg gggcggagcg taccagattg 1860

gggtatgaat gttccttcgg gccccggagc atagttcccc tcacagattt gcatttccca 1920

agctttcagt tctgagggtg gaatcatgtc cacctggggg gctatgaaga acaccgtttc 1980

gggggcgggg gtgattagtt gggatgatag caagtttctg agcaattgag atttgccaca 2040

tccggtgggg ccataaataa ttccgattac aggttgcagg tggtagttta gggaacggca 2100

actgccgtct tctcgaagca agggggccac ctcgttcatc atttccctta catgcatatt 2160

ttcccgcacc aaatccatta ggaggcgctc tcctcctagt gatagaagtt cttgtagtga 2220

ggaaaagttt ttcagcggtt ttagaccgtc agccatgggc attttggaaa gagtttgctg 2280

caaaagttct agtctgttcc acagttcagt gatgtgttct atggcatctc gatccagcag 2340

acctcctcgt ttcgcgggtt tggacggctc ctggagtagg gtatgagacg atgggcgtcc 2400

agcgctgcca gggttcggtc cttccagggt ctcagtgttc gagtcagggt tgtttccgtc 2460

acagtgaagg ggtgtgcgcc tgcttgggcg cttgccaggg tgcgcttcag actcattctg 2520

ctggtggaga acttctgtcg cttggcgccc tgtatgtcgg ccaagtagca gtttaccatg 2580

agttcgtagt tgagcgcctc ggctgcgtgg cctttggcgc ggagcttacc tttggaagtt 2640

ttcttgcata ccgggcagta taggcatttc agcgcataca gcttgggcgc aaggaaaatg 2700

gattctgggg agtatgcatc cgcgccgcag gaggcgcaaa cagtttcaca ttccaccagc 2760

caggttaaat ccggttcatt ggggtcaaaa acaagttttc cgccatattt tttgatgcgt 2820

ttcttacctt tggtctccat aagttcgtgt cctcgttgag tgacaaacag gctgtccgta 2880

tctccgtaga ctgattttac aggcctcttc tccagtggag tgcctcggtc ttcttcgtac 2940

aggaactctg accactctga tacaaaggcg cgcgtccagg ccagcacaaa ggaggctatg 3000

tgggaggggt agcgatcgtt gtcaaccagg gggtccacct tttccaaagt atgcaaacac 3060

atgtcaccct cttcaacatc caggaatgtg attggcttgt aggtgtattt cacgtgacct 3120

ggggtccccg ctgggggggt ataaaagggg gcggttcttt gctcttcctc actgtcttcc 3180

ggatcgctgt ccaggaacgt cagctgttgg ggtaggtatt ccctctcgaa ggcgggcatg 3240
acctctgcac tcaggttgtc agtttctaag aacgaggagg atttgatatt gacagtgccg 3300
gttgagatgc ctttcatgag gttttcgtcc atttggtcag aaaacacaat ttttttattg 3360
tcaagtttgg tggcaaatga tccatacagg gcgttggata aaagtttggc aatggatcgc 3420
atggtttggt tcttttcctt gtccgcgcgc tctttggcgg cgatgttgag ttggacatac 3480
tcgcgtgcca ggcacttcca ttcggggaag atagttgtta attcatctgg cacgattctc 3540
acttgccacc ctcgattatg caaggtaatt aaatccacac tggtggccac ctcgcctcga 3600

aggggttcat tggtccaaca gagcctacct cctttcctag aacagaaagg gggaagtggg 3660
tctagcataa gttcatcggg agggtctgca tccatggtaa agattcccgg aagtaaatcc 3720
ttatcaaaat agctgatggg agtggggtca tctaaggcca tttgccattc tcgagctgcc 3780
agtgcgcgct catatgggtt aaggggactg ccccagggca tgggatgggt gagagcagag 3840
gcatacatgc cacagatgtc atagacgtag atgggatcct caaagatgcc tatgtaggtt 3900
ggatagcatc gccccctct gatacttgct cgcacatagt catatagttc atgtgatggc 3960
gctagcagcc ccggacccaa gttggtgcga ttgggttttt ctgttctgta gacgatctgg 4020

cgaaagatgg cgtgagaatt ggaagagatg gtgggtcttt gaaaaatgtt gaaatgggca 4080
tgaggtagac ctacagagtc tctgacaaag tgggcataag attcttgaag cttggttacc 4140
agttcggcgg tgacaagtac gtctagggcg cagtagtcaa gtgtttcttg aatgatgtca 4200
taacctggtt ggtttttctt ttcccacagt tcgcggttga gaaggtattc ttcgcgatcc 4260
ttccagtact cttctagcgg aaacccgtct ttgtctgcac ggtaagatcc tagcatgtag 4320
aactgattaa ctgccttgta agggcagcag cccttctcta cgggtagaga gtatgcttga 4380
gcagcttttc gtagcgaagc gtgagtaagg gcaaaggtgt ctctgaccat gactttgaga 4440

aattggtatt tgaagtccat gtcgtcacag gctccctgtt cccagagttg gaagtctacc 4500
cgtttcttgt aggcggggtt gggcaaagcg aaagtaacat cattgaagag aatcttaccg 4560
gctctgggca taaaattgcg agtgatgcgg aaaggctgtg gtacttccgc tcgattgttg 4620
atcacctggg cagctaggac gatttcgtcg aaaccgttga tgttgtgtcc tacgatgtat 4680
aattctatga aacgcggcgt gcctctgacg tgaggtagct tactgagctc atcaaaggtt 4740
aggtctgtgg ggtcagataa ggcgtagtgt tcgagagccc attcgtgcag gtgaggattt 4800
gcatgtagga atgatgacca aagatctacc gccagtgctg tttgtaactg gtcccgatac 4860

tgacgaaaat gccggccaat tgccattttt tctggagtga cacagtagaa ggttctgggg 4920

tcttgttgcc atcgatccca cttgagttta atggctagat cgtgggccat gttgacgaga 4980

cgctcttctc ctgagagttt catgaccagc atgaaaggaa ctagttgttt gccaaaggat 5040

cccatccagg tgtaagtttc cacatcgtag gtcaggaaga gtctttctgt gcgaggatga 5100

gagccgatcg ggaagaactg gatttcctgc caccagttgg aggattggct gttgatgtga 5160

tggaagtaga agtttctgcg gcgcgccgag cattcgtgtt tgtgcttgta cagacggccg 5220

cagtagtcgc agcgttgcac gggttgtatc tcgtgaatga gctgtacctg gcttcccttg 5280

acgagaaatt tcagtgggaa gccgaggcct ggcgattgta tctcgtgctc ttctatattc 5340

gctgtatcgg cctgttcatc ttctgtttcg atggtggtca tgctgacgag cccccgcggg 5400

aggcaagtcc agacctcggc gcgggagggg cggagctgaa ggacgagagc gcgcaggctg 5460

gagctgtcca gagtcctgag acgctgcgga ctcaggttag taggtaggga cagaagatta 5520

acttgcatga tcttttccag ggcgtgcggg aggttcagat ggtacttgat ttccacaggt 5580

tcgtttgtag agacgtcaat ggcttgcagg gttccgtgtc ctttgggcgc cactaccgta 5640

cctttgtttt ttcttttgat cggtggtggc tctcttgctt cttgcatgct cagaagcggt 5700

gacggggacg cgcgccgggc ggcagcggtt gttccggacc cgggggcatg gctggtagtg 5760

gcacgtcggc gccgcgcacg ggcaggttct ggtattgcgc tctgagaaga cttgcgtgcg 5820

ccaccacgcg tcgattgacg tcttgtatct gacgtctctg ggtgaaagct accggccccg 5880

tgagcttgaa cctgaaagag agttcaacag aatcaatttc ggtatcgtta acggcagctt 5940

gtctcagtat ttcttgtacg tcaccagagt tgtcctggta ggcgatctcc gccatgaact 6000

gctcgatttc ttcctcctga agatctccgc gacccgctct ttcgacggtg gccgcgaggt 6060

cattggagat acggcccatg agttgggaga atgcattcat gcccgcctcg ttccagacgc 6120

ggctgtaaac cacggccccc tcggagtctc ttgcgcgcat caccacctga gcgaggttaa 6180

gctccacgtg tctggtgaag accgcatagt tgcataggcg ctgaaaaagg tagttgagtg 6240

tggtggcaat gtgttcggcg acgaagaaat acatgatcca tcgtctcagc ggcatttcgc 6300

taacatcgcc cagagcttcc aagcgctcca tggcctcgta gaagtccacg gcaaaattaa 6360

aaaactggga gtttcgcgcg gacacggtca attcctcctc gagaagacgg atgagttcgg 6420

ctatggtggc ccgtacttcg cgttcgaagg ctcccgggat ctcttcttcc tcttctatct 6480

cttcttccac taacatctct tcttcgtctt caggcggggg cggaggggcc acgcggcgac 6540

gtcgacggcg cacgggcaaa cggtcgatga atcgttcaat gacctctccg cggcggcggc 6600

gcatggtttc agtgacggcg cggccgttct cgcgcggtcg cagagtaaaa acaccgccgc 6660

gcatctcctt aaagtggtga ctgggaggtt ctccgtttgg gagggagagg gcgctgatta 6720

tacattttat taattggccc gtagggactg cgcgcagaga tctgatcgtg tcaagatcca 6780

cgggatctga aaacctttcg acgaaagcgt ctaaccagtc acagtcacaa ggtaggctga 6840

gtacggcttc ttgtgggcgg gggtggttat gtgttcggtc tgggtcttct gtttcttctt 6900

catctcggga aggtgagacg atgctgctgg tgatgaaatt aaagtaggca gttctaagac 6960

ggcggatggt ggcgaggagc accaggtctt tgggtccggc ttgctggata cgcaggcgat 7020

tggccattcc ccaagcatta tcctgacatc tagcaagatc tttgtagtag tcttgcatga 7080

gccgttctac gggcacttct tcctcacccg ttctgccatg catacgtgtg agtccaaatc 7140

cgcgcattgg ttgtaccagt gccaagtcag ctacgactct ttcggcgagg atggcttgct 7200

gtacttgggt aagggtggct tgaaagtcat caaaatccac aaagcggtgg taagcccctg 7260

tattaatggt gtaagcacag ttggccatga ctgaccagtt aactgtctgg tgaccagggc 7320

gcacgagctc ggtgtattta aggcgcgaat aggcgcgggt gtcaaagatg taatcgttgc 7380

aggtgcgcac cagatactgg taccctataa gaaaatgcgg cggtggttgg cggtagagag 7440

gccatcgttc tgtagctgga gcgccagggg cgaggtcttc caacataagg cggtgatagc 7500

cgtagatgta cctggacatc caggtgattc ctgcggcggt agtagaagcc cgaggaaact 7560

cgcgtacgcg gttccaaatg ttgcgtagcg gcatgaagta gttcattgta ggcacggttt 7620

gaccagtgag gcgcgcgcag tcattgatgc tctatagaca cggagaaaat gaaagcgttc 7680

agcgactcga ctccgtagcc tggaggaacg tgaacgggtt gggtcgcggt gtaccccggt 7740

tcgagacttg tactcgagcc ggccggagcc gcggctaacg tggtattggc actcccgtct 7800

cgacccagcc tacaaaaatc caggatacgg aatcgagtcg ttttgctggt ttccgaatgg 7860

cagggaagtg agtcctattt ttttttttt tgccgctcag atgcatcccg tgctgcgaca 7920

gatgcgcccc caacaacagc ccccctcgca gcagcagcag cagcaaccac aaaaggctgt 7980

ccctgcaact actgcaactg ccgccgtgag cggtgcggga cagcccgcct atgatctgga 8040

cttggaagag ggcgaaggac tggcacgtct aggtgcgcct tcgcccgagc ggcatccgcg 8100

agttcaactg aaaaaagatt ctcgcgaggc gtatgtgccc caacagaacc tatttagaga 8160

cagaagcggc gaggagccgg aggagatgcg agcttcccgc tttaacgcgg gtcgtgagct 8220

gcgtcacggt ttggaccgaa gacgagtgtt gcgagacgag gatttcgaag ttgatgaagt 8280

gacagggatc agtcctgcca gggcacacgt ggctgcagcc aaccttgtat cggcttacga 8340

gcagacagta aaggaagagc gtaacttcca aaagtctttt aataatcatg tgcgaaccct 8400

gattgcccgc gaagaagtta cccttggttt gatgcatttg tgggatttga tggaagctat 8460

cattcagaac cctactagca aacctctgac cgcccagctg tttctggtgg tgcaacacag 8520

cagagacaat gaggctttca gagaggcgct gctgaacatc accgaacccg aggggagatg 8580

gttgtatgat cttatcaaca ttctacagag tatcatagtg caggagcgga gcctgggcct 8640

ggccgagaag gtagctgcca tcaattactc ggttttgagc ttgggaaaat attacgctcg 8700
caaaatctac aagactccat acgttcccat agacaaggag gtgaagatag atgggttcta 8760
catgcgcatg acgctcaagg tcttgaccct gagcgatgat cttggggtgt atcgcaatga 8820
cagaatgcat cgcgcggtta gcgccagcag gaggcgcgag ttaagcgaca gggaactgat 8880
gcacagtttg caaagagctc tgactggagc tggaaccgag ggtgagaatt acttcgacat 8940
gggagctgac ttgcagtggc agcctagtcg cagggctctg agcgccgcga cggcaggatg 9000
tgagcttcct tacatagaag aggcggatga aggcgaggag gaagagggcg agtacttgga 9060

agactgatgg cacaacccgt gttttttgct agatggaaca gcaagcaccg gatcccgcaa 9120
tgcgggcggc gctgcagagc cagccgtccg gcattaactc ctcggacgat tggacccagg 9180
ccatgcaacg tatcatggcg ttgacgactc gcaaccccga agcctttaga cagcaacccc 9240
aggccaaccg tctatcggcc atcatggaag ctgtagtgcc ttcccgatct aatcccactc 9300
atgagaaggt cctggccatc gtgaacgcgt tggtggagaa caaagctatt cgtccagatg 9360
aggccggact ggtatacaac gctctcttag aacgcgtggc tcgctacaac agtagcaatg 9420
tgcaaaccaa tttggaccgt atgataacag atgtacgcga agccgtgtct cagcgcgaaa 9480

ggttccagcg tgatgccaac ctgggttcgc tggtggcgtt aaatgctttc ttgagtactc 9540
agcctgctaa tgtgccgcgt ggtcaacagg attatactaa ctttttaagt gctttgagac 9600
tgatggtatc agaagtacct cagagcgaag tgtatcagtc cggtcctgat tacttctttc 9660
agactagcag acagggcttg cagacggtaa atctgagcca agcttttaaa aaccttaaag 9720
gtttgtgggg agtgcatgcc ccggtaggag aaagagcaac cgtgtctagc ttgttaactc 9780
cgaactcccg cctgttatta ctgttggtag ctcctttcac cgacagcggt agcatcgacc 9840
gtaattccta tttgggttac ctactaaacc tgtatcgcga agccataggg caaagtcagg 9900

tggacgagca gacctatcaa gaaattaccc aagtcagtcg cgctttggga caggaagaca 9960
ctggcagttt ggaagccact ctgaacttct tgcttaccaa tcggtctcaa aagatccctc 10020
ctcaatatgc tcttactgcg gaggaggaga ggatccttag atatgtgcag cagagcgtgg 10080
gattgtttct gatgcaagag ggggcaactc cgactgcagc actggacatg acagcgcgaa 10140
atatggagcc cagcatgtat gccagtaacc gacctttcat taacaaactg ctggactact 10200
tgcacagagc tgccgctatg aactctgatt atttcaccaa tgccatctta aacccgcact 10260
ggctgccccc acctggtttc tacacgggcg aatatgacat gcccgaccct aatgacggat 10320

ttctgtggga cgacgtggac agcgatgttt tttcacctct ttctgatcat cgcacgtgga 10380

aaaaggaagg cggtgataga atgcattctt ctgcatcgct gtccggggtc atgggtgcta 10440

ccgcggctga gcccgagtct gcaagtcctt ttcctagtct acccttttct ctacacagtg 10500

tacgtagcag cgaagtgggt agaataagtc gcccgagttt aatgggcgaa gaggagtacc 10560

taaacgattc cttgctcaga ccggcaagag aaaaaaattt cccaaacaat ggaatagaaa 10620

gtttggtgga taaaatgagt agatggaaga cttatgctca ggatcacaga gacgagcctg 10680

ggatcatggg gactacaagt agagcgagcc gtagacgcca gcgccatgac agacagaggg 10740

gtcttgtgtg ggacgatgag gattcggccg atgatagcag cgtgttggac ttgggtggga 10800

gaggaagggg caacccgttt gctcatttgc gccctcgctt gggtggtatg ttgtgaaaaa 10860

aaataaaaaa gaaaaactca ccaaggccat ggcgacgagc gtacgttcgt tcttctttat 10920

tatctgtgtc tagtataatg aggcgagtcg tgctaggcgg agcggtggtg tatccggagg 10980

gtcctcctcc ttcgtacgag agcgtgatgc agcagcagca ggcgacggcg gtgatgcaat 11040

ccccactgga ggctcccttt gtgcctccgc gatacctggc acctacggag ggcagaaaca 11100

gcattcgtta ctcggaactg gcacctcagt acgataccac caggttgtat ctggtggaca 11160

acaagtcggc ggacattgct tctctgaact atcagaatga ccacagcaac ttcttgacca 11220

cggtggtgca gaacaatgac tttaccccta cggaagccag cacccagacc attaactttg 11280

atgaacgatc gcggtggggc ggtcagctaa agaccatcat gcatactaac atgccaaacg 11340

tgaacgagta tatgtttagt aacaagttca aagcgcgtgt gatggtgtcc agaaaacctc 11400

ccgacggtgc tgcagttggg gatacttatg atcacaagca ggatattttg gaatatgagt 11460

ggttcgagtt tactttgcca gaaggcaact tttcagttac tatgactatt gatttgatga 11520

acaatgccat catagataat tacttgaaag tgggtagaca gaatggagtg cttgaaagtg 11580

acattggtgt taagttcgac accaggaact tcaagctggg atgggatccc gaaaccaagt 11640

tgatcatgcc tggagtgtat acgtatgaag ccttccatcc tgacattgtc ttactgcctg 11700

gctgcggagt ggattttacc gagagtcgtt tgagcaacct tcttggtatc agaaaaaaac 11760

agccatttca agagggtttt aagattttgt atgaagattt agaaggtggt aatattccgg 11820

ccctcttgga tgtagatgcc tatgagaaca gtaagaaaga acaaaaagcc aaaatagaag 11880

ctgctacagc tgctgcagaa gctaaggcaa acatagttgc cagcgactct acaagggttg 11940

ctaacgctgg agaggtcaga ggagacaatt ttgcgccaac acctgttccg actgcagaat 12000

cattattggc cgatgtgtct gaaggaacgg acgtgaaact cactattcaa cctgtagaaa 12060

aagatagtaa gaatagaagc tataatgtgt tggaagacaa aatcaacaca gcctatcgca 12120

gttggtatct ttcgtacaat tatggcgatc ccgaaaaagg agtgcgttcc tggacattgc 12180

tcaccacctc agatgtcacc tgcggagcag agcaggttta ctggtcgctt ccagacatga 12240

tgaaggatcc tgtcactttc cgctccacta gacaagtcag taactaccct gtggtgggtg 12300

cagagcttat gcccgtcttc tcaaagagct tctacaacga acaagctgtg tactcccagc 12360

agctccgcca gtccacctcg cttacgcacg tcttcaaccg ctttcctgag aaccagattt 12420

taatccgtcc gccggcgccc accattacca ccgtcagtga aaacgttcct gctctcacag 12480

atcacgggac cctgccgttg cgcagcagta tccggggagt ccaacgtgtg accgttactg 12540

acgccagacg ccgcacctgt ccctacgtgt acaaggcact gggcatagtc gcaccgcgcg 12600

tcctttcaag ccgcactttc taaaaaaaaa atgtccattc ttatctcgcc cagtaataac 12660

accggttggg gtctgcgcgc tccaagcaag atgtacggag gcgcacgcaa acgttctacc 12720

caacatcccg tgcgtgttcg cggacatttt cgcgctccat ggggtgccct caagggccgc 12780

actcgcgttc gaaccaccgt cgatgatgta atcgatcagg tggttgccga cgcccgtaat 12840

tatactccta ctgcgcctac atctactgtg gatgcagtta ttgacagtgt agtggctgac 12900

gctcgcaact atgctcgacg taagagccgg cgaaggcgca ttgccagacg ccaccgagct 12960

accactgcca tgcgagccgc aagagctctg ctacgaagag ctagacgcgt ggggcgaaga 13020

gccatgctta gggcggccag acgtgcagct tcgggcgcca gcgccggcag gtcccgcagg 13080

caagcagccg ctgtcgcagc ggcgactatt gccgacatgg cccaatcgcg aagaggcaat 13140

gtatactggg tgcgtgacgc tgccaccggt caacgtgtac ccgtgcgcac ccgtcccgct 13200

cgcacttaga agatactgag cagtctccga tgttgtgtcc cagcggcgag gatgtccaag 13260

cgcaaataca aggaagaaat gctgcaggtt atcgcacctg aagtctacgg ccaaccgttg 13320

aaggatgaaa aaaaaccccg caaaatcaag cgggttaaaa aggacaaaaa agaagaggaa 13380

gatggcgatg atgggctggc ggagtttgtg cgcgagtttg ccccacggcg acgcgtgcaa 13440

tggcgtgggc gcaaagttcg acatgtgttg agacctggaa cttcggtggt ctttacaccc 13500

ggcgagcgtt caagcgctac ttttaagcgt tcctatgatg aggtgtacgg ggatgatgat 13560

attcttgagc aggcggctga ccgattaggc gagtttgctt atggcaagcg tagtagaata 13620

acttccaagg atgagacagt gtcaataccc ttggatcatg gaaatcccac ccctagtctt 13680

aaaccggtca ctttgcagca agtgttaccc gtaactccgc gaacaggtgt taaacgcgaa 13740

ggtgaagatt tgtatcccac tatgcaactg atggtaccca aacgccagaa gttggaggac 13800

gttttggaga aagtaaaagt ggatccagat attcaacctg aggttaaagt gagacccatt 13860

aagcaggtag cgcctggtct gggggtacaa actgtagaca ttaagattcc cactgaaagt 13920

atggaagtgc aaactgaacc cgcaaagcct actgccacct ccactgaagt gcaaacggat 13980

ccatggatgc ccatgcctat tacaactgac gccgccggtc ccactcgaag atcccgacga 14040

aagtacggtc cagcaagtct gttgatgccc aattatgttg tacacccatc tattattcct 14100

actcctggtt accgaggcac tcgctactat cgcagccgaa acagtacctc ccgccgtcgc 14160
cgcaagacac ctgcaaatcg cagtcgtcgc cgtagacgca caagcaaacc gactcccggc 14220
gccctggtgc ggcaagtgta ccgcaatggt agtgcggaac ctttgacact gccgcgtgcg 14280
cgttaccatc cgagtatcat cacttaatca atgttgccgc tgcctccttg cagatatggc 14340
cctcacttgt cgccttcgcg ttcccatcac tggttaccga ggaagaaact cgcgccgtag 14400
aagagggatg ttgggacgcg gaatgcgacg ctacaggcga cggcgtgcta tccgcaagca 14460
attgcggggt ggttttttac cagccttaat tccaattatc gctgctgcaa ttggcgcgat 14520

accaggcata gcttccgtgg cggttcaggc ctcgcaacga cattgacatt ggaaaaaaaa 14580
cgtataaata aaaaaaaata caatggactc tgacactcct ggtcctgtga ctatgttttc 14640
ttagagatgg aagacatcaa tttttcatcc ttggctccgc gacacggcac gaagccgtac 14700
atgggcacct ggagcgacat cggcacgagc caactgaacg ggggcgcctt caattggagc 14760
agtatctgga gcgggcttaa aaattttggc tcaaccataa aaacatacgg gaacaaagct 14820
tggaacagca gtacaggaca ggcgcttaga aataaactta aagaccagaa cttccaacaa 14880
aaagtagtcg atgggatagc ttccggcatc aatggagtgg tagatttggc taaccaggct 14940

gtgcagaaaa agataaacag tcgtttggac ccgccgccag caaccccagg tgaaatgcaa 15000
gtggaggaag aaattcctcc gccagaaaaa cgaggcgaca agcgtccgcg tcccgatttg 15060
gaagagacgc tggtgacgcg cgtagatgaa ccgccttctt atgaggaagc aacgaagctt 15120
ggaatgccca ccactagacc gatagcccca atggccaccg gggtgatgaa accttctcag 15180
ttgcatcgac ccgtcacctt ggatttgccc cctccccctg ctgctactgc tgtacccgct 15240
tctaagcctg tcgctgcccc gaaaccagtc gccgtagcca ggtcacgtcc cgggggcgct 15300
cctcgtccaa atgcgcactg gcaaaatact ctgaacagca tcgtgggtct aggcgtgcaa 15360

agtgtaaaac gccgtcgctg cttttaatta aatatggagt agcgcttaac ttgcctatct 15420
gtgtatatgt gtcattacac gccgtcacag cagcagagga aaaaaggaag aggtcgtgcg 15480
tcgacgctga gttactttca agatggccac cccatcgatg ctgccccaat gggcatacat 15540
gcacatcgcc ggacaggatg cttcggagta cctgagtccg ggtctggtgc agttcgcccg 15600
cgccacagac acctacttca atctgggaaa taagtttaga aatcccaccg tagcgccgac 15660
ccacgatgtg accaccgacc gtagccagcg gctcatgttg cgcttcgtgc ccgttgaccg 15720
ggaggacaat acatactctt acaaagtgcg gtacaccctg gccgtgggcg acaacagagt 15780

gctggatatg gccagcacgt tctttgacat taggggcgtg ttggacagag gtcccagttt 15840

caaaccctat tctggtacgg cttacaactc tctggctcct aaaggcgctc caaatgcatc 15900

tcaatggatt gcaaaaggcg taccaactgc agcagccgca ggcaatggtg aagaagaaca 15960

tgaaacagag gagaaaactg ctacttacac ttttgccaat gctcctgtaa aagccgaggc 16020

tcaaattaca aaagagggct taccaatagg tttggagatt tcagctgaaa acgaatctaa 16080

acccatctat gcagataaac tttatcagcc agaacctcaa gtgggagatg aaacttggac 16140

tgacctagac ggaaaaaccg aagagtatgg aggcagggct ctaaagccta ctactaacat 16200

gaaaccctgt tacgggtcct atgcgaagcc tactaattta aaaggtggtc aggcaaaacc 16260

gaaaaactcg gaaccgtcga gtgaaaaaat tgaatatgat attgacatgg aatttttga 16320

taactcatcg caaagaacaa acttcagtcc taaaattgtc atgtatgcag aaaatgtagg 16380

tttggaaacg ccagacactc atgtagtgta caaacctgga acagaagaca caagttccga 16440

agctaatttg ggacaacagt ctatgcccaa cagacccaac tacattggct tcagagataa 16500

ctttattgga ctcatgtact ataacagtac tggtaacatg ggggtgctgg ctggtcaagc 16560

gtctcagtta aatgcagtgg ttgacttgca ggacagaaac acagaacttt cttaccaact 16620

cttgcttgac tctctgggcg acagaaccag atactttagc atgtggaatc aggctgtgga 16680

cagttatgat cctgatgtac gtgttattga aaatcatggt gtggaagatg aacttcccaa 16740

ctattgtttt ccactggacg gcataggtgt tccaacaacc agttacaaat caatagttcc 16800

aaatggagaa gataataata attggaaaga acctgaagta aatggaacaa gtgagatcgg 16860

acagggtaat ttgtttgcca tggaaattaa ccttcaagcc aatctatggc gaagtttcct 16920

ttattccaat gtggctctgt atctcccaga ctcgtacaaa tacaccccgt ccaatgtcac 16980

tcttccagaa aacaaaaaca cctacgacta catgaacggg cgggtggtgc cgccatctct 17040

agtagacacc tatgtgaaca ttggtgccag gtggtctctg gatgccatgg acaatgtcaa 17100

cccattcaac caccaccgta acgctggctt gcgttaccga tctatgcttc tgggtaacgg 17160

acgttatgtg cctttccaca tacaagtgcc tcaaaaattc ttcgctgtta aaaacctgct 17220

gcttctccca ggctcctaca cttatgagtg gaactttagg aaggatgtga acatggttct 17280

acagagttcc ctcggtaacg acctgcgggt agatggcgcc agcatcagtt tcacgagcat 17340

caacctctat gctacttttt tccccatggc tcacaacacc gcttccaccc ttgaagccat 17400

gctgcggaat gacaccaatg atcagtcatt caacgactac ctatctgcag ctaacatgct 17460

ctaccccatt cctgccaatg caaccaatat tcccatttcc attccttctc gcaactgggc 17520

ggctttcaga ggctggtcat ttaccagact gaaaaccaaa gaaactccct ctttggggtc 17580

tggatttgac ccctactttg tctattctgg ttctattccc tacctggatg gtaccttcta 17640

cctgaaccac acttttaaga aggtttccat catgtttgac tcttcagtga gctggcctgg 17700

aaatgacagg ttactatctc ctaacgaatt tgaaataaag cgcactgtgg atggcgaagg 17760

ctacaacgta gcccaatgca acatgaccaa agactggttc ttggtacaga tgctcgccaa 17820

ctacaacatc ggctatcagg gcttctacat tccagaagga tacaaagatc gcatgtattc 17880

atttttcaga aacttccagc ccatgagcag gcaggtggtt gatgaggtca attacaaaga 17940

cttcaaggcc gtcgccatac cctaccaaca caacaactct ggctttgtgg gttacatggc 18000

tccgaccatg cgccaaggtc aaccctatcc cgctaactat ccctatccac tcattggaac 18060

aactgccgta aatagtgtta cgcagaaaaa gttcttgtgt gacagaacca tgtggcgcat 18120

accgttctcg agcaacttca tgtctatggg ggcccttaca gacttgggac agaatatgct 18180

ctatgccaac tcagctcatg ctctggacat gacctttgag gtggatccca tggatgagcc 18240

caccctgctt tatcttctct tcgaagtttt cgacgtggtc agagtgcatc agccacaccg 18300

cggcatcatc gaggcagtct acctgcgtac accgttctcg gccggtaacg ctaccacgta 18360

agaagcttct tgcttcttgc aaatagcagc tgcaaccatg gcctgcggat cccaaaacgg 18420

ctccagcgag caagagctca gagccattgt ccaagacctg ggttgcggac cctatttttt 18480

gggaacctac gataagcgct tcccggggtt catggccccc gataagctcg cctgtgccat 18540

tgtaaatacg gccggacgtg agacgggggg agagcactgg ttggctttcg gttggaaccc 18600

acgttctaac acctgctacc tttttgatcc ttttggattc tcggatgatc gtctcaaaca 18660

gatttaccag tttgaatatg agggtctcct gcgccgcagc gctcttgcta ccaaggaccg 18720

ctgtattacg ctggaaaaat ctacccagac cgtgcagggc ccccgttctg ccgcctgcgg 18780

acttttctgc tgcatgttcc ttcacgcctt tgtgcactgg cctgaccgtc ccatggacgg 18840

aaaccccacc atgaaattgc taactggagt gccaaacaac atgcttcatt ctcctaaagt 18900

ccagcccacc ctgtgtgaca atcaaaaagc actctaccat tttcttaata cccattcgcc 18960

ttattttcgc tctcatcgta cacacatcga aagggccact gcgttcgacc gtatggatgt 19020

tcaataatga ctcatgtaaa caacgtgttc aataaacatc actttatttt tttacatgta 19080

tcaaggctct ggattactta tttatttaca agtcgaatgg gttctgacga gaatcagaat 19140

gacccgcagg cagtgatacg ttgcggaact gatacttggg ttgccacttg aattcgggaa 19200

tcaccaactt gggaaccggt atatcgggca ggatgtcact ccacagcttt ctggtcagct 19260

gcaaagctcc aagcaggtca ggagccgaaa tcttgaaatc acaattagga ccagtgctct 19320

gagcgcgaga gttgcggtac accggattgc agcactgaaa caccatcagc gacggatgtc 19380

tcacgcttgc cagcacggtg ggatctgcaa tcatgcccac atccagatct tcagcattgg 19440

caatgctgaa cggggtcatc ttgcaggtct gcctacccat ggcgggcacc caattaggct 19500

tgtggttgca atcgcagtgc agggggatca gtatcatctt ggcctgatcc tgtctgattc 19560

ctggatacac ggctctcatg aaagcatcat attgcttgaa agcctgctgg gctttactac 19620
cctcggtata aaacatcccg caggacctgc tcgaaaactg gttagctgca cagccggcat 19680
cattcacaca gcagcgggcg tcattgttgg ctatttgcac cacacttctg ccccagcggt 19740
tttgggtgat tttggttcgc tcgggattct cctttaaggc tcgttgtccg ttctcgctgg 19800
ccacatccat ctcgataatc tgctccttct gaatcataat attgccatgc aggcacttca 19860
gcttgccctc ataatcattg cagccatgag gccacaacgc acagcctgta cattcccaat 19920
tatggtgggc gatctgagaa aaagaatgta tcattccctg cagaaatctt cccatcatcg 19980

tgctcagtgt cttgtgacta gtgaaagtta actggatgcc tcggtgctct tcgtttacgt 20040
actggtgaca gatgcgcttg tattgttcgt gttgctcagg cattagttta aaacaggttc 20100
taagttcgtt atccagcctg tacttctcca tcagcagaca catcacttcc atgcctttct 20160
cccaagcaga caccaggggc aagctaatcg gattcttaac agtgcaggca gcagctcctt 20220
tagccagagg gtcatcttta gcgatcttct caatgcttct tttgccatcc ttctcaacga 20280
tgcgcacggg cgggtagctg aaacccactg ctacaagttg cgcctcttct ctttcttctt 20340
cgctgtcttg actgatgtct tgcatgggga tatgtttggt cttccttggc ttctttttgg 20400

ggggtatcgg aggaggagga ctgtcgctcc gttccggaga cagggaggat tgtgacgttt 20460
cgctcaccat taccaactga ctgtcggtag aagaacctga ccccacacgg cgacaggtgt 20520
ttttcttcgg gggcagaggt ggaggcgatt gcgaagggct gcggtccgac ctggaaggcg 20580
gatgactggc agaacccctt ccgcgttcgg gggtgtgctc cctgtggcgg tcgcttaact 20640
gatttccttc gcggctggcc attgtgttct cctaggcaga gaaacaacag acatggaaac 20700
tcagccattg ctgtcaacat cgccacgagt gccatcacat ctcgtcctca gcgacgagga 20760
aaaggagcag agcttaagca ttccaccgcc cagtcctgcc accacctcta ccctagaaga 20820

taaggaggtc gacgcatctc atgacatgca gaataaaaaa gcgaaagagt ctgagacaga 20880
catcgagcaa gacccgggct atgtgacacc ggtggaacac gaggaagagt tgaaacgctt 20940
tctagagaga gaggatgaaa actgcccaaa acagcgagca gataactatc accaagatgc 21000
tggaaatagg gatcagaaca ccgactacct catagggctt gacggggaag acgcgctcct 21060
taaacatcta gcaagacagt cgctcatagt caaggatgca ttattggaca gaactgaagt 21120
gcccatcagt gtggaagagc tcagctgcgc ctacgagctt aacctttttt cacctcgtac 21180
tccccccaaa cgtcagccaa acggcacctg cgagccaaat cctcgcttaa acttttatcc 21240

agcttttgct gtgccagaag tactggctac ctatcacatc ttttttaaaa atcaaaaaat 21300

tccagtctcc tgccgcgcta atcgcacccg cgccgatgcc ctactcaatc tgggacctgg 21360

ttcacgctta cctgatatag cttccttgga agaggttcca aagatcttcg agggtctggg 21420

caataatgag actcgggccg caaatgctct gcaaaaggga gaaaatggca tggatgagca 21480

tcacagcgtt ctggtggaat tggaaggcga taatgccaga ctcgcagtac tcaagcgaag 21540

cgtcgaggtc acacacttcg catatcccgc tgtcaacctg ccccctaaag tcatgacggc 21600

ggtcatggac cagttactca ttaagcgcgc aagtcccctt tcagaagaca tgcatgaccc 21660

agatgcctgt gatgagggta aaccagtggt cagtgatgag cagctaaccc gatggctggg 21720

caccgactct ccccgggatt tggaagagcg tcgcaagctt atgatggccg tggtgctggt 21780

taccgtagaa ctagagtgtc tccgacgttt ctttaccgat tcagaaacct tgcgcaaact 21840

cgaagagaat ctgcactaca cttttagaca cggctttgtg cggcaggcat gcaagatatc 21900

taacgtggaa ctcaccaacc tggtttccta catgggtatt ctgcatgaga atcgcctagg 21960

acaaagcgtg ctgcacagca cccttaaggg ggaagcccgc cgtgattaca tccgcgattg 22020

tgtctatctc tacctgtgcc acacgtggca aaccggcatg ggtgtatggc agcaatgttt 22080

agaagaacag aacttgaaag agcttgacaa gctcttacag aaatctctta aggttctgtg 22140

gacagggttc gacgagcgca ccgtcgcttc cgacctggca gacctcatct tcccagagcg 22200

tctcagggtt actttgcgaa acggattgcc tgactttatg agccagagca tgcttaacaa 22260

ttttcgctct ttcatcctgg aacgctccgg tatcctgccc gccacctgct gcgcactgcc 22320

ctccgacttt gtgcctctca cctaccgcga gtgccccccg ccgctatgga gtcactgcta 22380

cctgttccgt ctggccaact atctctccta ccactcggat gtgatcgagg atgtgagcgg 22440

agacggcttg ctggagtgcc actgccgctg caatctgtgc acgccccacc ggtccctagc 22500

ttgcaacccc cagttgatga gcgaaaccca gataataggc acctttgaat tgcaaggccc 22560

cagcagccaa ggcgatgggt cttctcctgg gcaaagttta aaactgaccc cgggactgtg 22620

gacctccgcc tacttgcgca agtttgctcc ggaagattac cacccctatg aaatcaagtt 22680

ctatgaggac caatcacagc ctccaaaggc cgaactttcg gcttgcgtca tcacccaggg 22740

ggcaattctg gcccaattgc aagccatcca aaaatcccgc caagaatttc tactgaaaaa 22800

gggtaagggg gtctaccttg acccccagac cggcgaggaa ctcaacacaa ggttccctca 22860

ggatgtccca acgacgagaa aacaagaagt tgaaggtgca gccgccgccc ccagaagata 22920

tggaggaaga ttgggacagt caggcagagg aggcggagga ggacagtctg gaggacagtc 22980

tggaggaaga cagtttggag gaggaaaacg aggaggcaga ggaggtggaa gaagtaaccg 23040

ccgacaaaca gttatcctcg gctgcggaga caagcaacag cgctaccatc tccgctccga 23100

gtcgaggaac ccggcggcgt cccagcagta gatgggacga gaccggacgc ttcccgaacc 23160

caaccagcgc ttccaagacc ggtaagaagg atcggcaggg atacaagtcc tggcgggggc 23220

ataagaatgc catcatctcc tgcttgcatg agtgcggggg caacatatcc ttcacgcggc 23280

gctacttgct attccaccat ggggtgaact ttccgcgcaa tgttttgcat tactaccgtc 23340

acctccacag cccctactat agccagcaaa tcccgacagt ctcgacagat aaagacagcg 23400

gcggcgacct ccaacagaaa accagcagcg gcagttagaa aatacacaac aagtgcagca 23460

acaggaggat taaagattac agccaacgag ccagcgcaaa cccgagagtt aagaaatcgg 23520

atctttccaa ccctgtatgc catcttccag cagagtcggg gtcaagagca ggaactgaaa 23580

ataaaaaacc gatctctgcg ttcgctcacc agaagttgtt tgtatcacaa gagcgaagat 23640

caacttcagc gcactctcga ggacgccgag gctctcttca acaagtactg cgcgctgact 23700

cttaaagagt aggcagcgac cgcgcttatt caaaaaaggc gggaattaca tcatcctcga 23760

catgagtaaa gaaattccca cgccttacat gtggagttat caaccccaaa tgggattggc 23820

agcaggcgcc tcccaggact actccacccg catgaattgg ctcagcgccg ggccttctat 23880

gatttctcga gttaatgata tacgcgccta ccgaaaccaa atacttttgg aacagtcagc 23940

tcttaccacc acgccccgcc aacaccttaa tcccagaaat tggcccgccg ccctagtgta 24000

ccaggaaagt cccgctccca ccactgtatt acttcctcga gacgcccagg ccgaagtcca 24060

aatgactaat gcaggtgcgc agttagctgg cggctccacc ctatgtcgtc acaggcctcg 24120

gcataatata aaacgcctga tgatcagagg ccgaggtatc cagctcaacg acgagtcggt 24180

gagctctccg cttggtctac gaccagacgg aatctttcag attgccggct gcgggagatc 24240

ttccttcacc cctcgtcagg ctgttctgac tttggaaagt tcgtcttcgc aaccccgctc 24300

gggcggaatc gggaccgttc aatttgtaga ggagtttact ccctctgtct acttcaaccc 24360

cttctccgga tctcctgggc actacccgga cgagttcata ccgaacttcg acgcgattag 24420

cgagtcagtg gacggctacg attgatgtct ggtgacgcgg ctgagctatc tcggctgcga 24480

catctagacc actgccgccg ctttcgctgc tttgcccggg aacttattga gttcatctac 24540

ttcgaactcc ccaaggatca ccctcaaggt ccggcccacg gagtgcggat tactatcgaa 24600

ggcaaaatag actctcgcct gcaacgaatt ttctcccagc ggcccgtgct gatcgagcga 24660

gaccagggaa acaccacggt tagtaatcaa ttacggggtc attagttcat agcccatata 24720

tggagttaac ttgtttattg cagcttataa tggttacaaa taaagcaata gcatcacaaa 24780

tttcacaaat aaagcatttt tttcactgca ttctagttgt ggtttgtcca aactcatcaa 24840

tgtatcttat catgtctgct cgaagcggcc ggccgccccg actctagagt cgcggcctca 24900

ttaggaagtt cctatacttt ctagagaata ggaacttctc agaagaactc gtcaagaagg 24960

cgatagaagg cgatgcgctg cgaatcggga gcggcgatac cgtaaagcac gaggaagcgg 25020

tcagcccatt cgccgccaag ctcttcagca atatcacggg tagccaacgc tatgtcctga 25080
tagcggtccg ccacacccag ccggccacag tcgatgaatc cagaaaagcg gccattttcc 25140
accatgatat tcggcaagca ggcatcgcca tgggtcacga cgagatcctc gccgtcgggc 25200
atgcgcgcct tgagcctggc gaacagttcg gctggcgcga gcccctgatg ctcttcgtcc 25260
agatcatcct gatcgacaag accggcttcc atccgagtac gtgctcgctc gatgcgatgt 25320
ttcgcttggt ggtcgaatgg gcaggtagcc ggatcaagcg tatgcagccg ccgcattgca 25380
tcagccatga tggatacttt ctcggcagga gcaaggtgag atgacaggag atcctgcccc 25440

ggcacttcgc ccaatagcag ccagtccctt cccgcttcag tgacaacgtc gagcacagct 25500
gcgcaaggaa cgcccgtcgt ggccagccac gatagccgcg ctgcctcgtc ctgcagttca 25560
ttcagggcac cggacaggtc ggtcttgaca aaaagaaccg ggcgcccctg cgctgacagc 25620
cggaacacgg cggcatcaga gcagccgatt gcctgttgtg cccagtcata gccgaatagc 25680
ctctccaccc aagcggccgg agaacctgcg tgcaatccat cttgttcaat ggccgatccc 25740
ataacacccc ttgtattact gtttatgtaa gcagacagtt ttactgttcg tgatgatata 25800
tttttatctt gtgcaatgta acaggttgtg gccatagcgg gcccgggatt ttcctccacg 25860

tccccgcatg ttagaagact tcccctgccc tcggctctgg aagttcctat actttctaga 25920
gaataggaac ttcccgccag aatgcgttcg cacagccgcc agccggtcac tccgttgatg 25980
gttactcgga acagcaggga gccgtcgggg ttgatcaggc gctcgtcgat aattttgttg 26040
ccgttccaca gggtccctgt tacagtgatc tttttgccgt cgaacacggc gatgccttca 26100
tacggccgtc cgaaatagtc gatcatgttc ggcgtaaccc cgtcgattac cagtgtgcca 26160
tagtgcagga tcaccttaaa gtgatgatca tccacagggt acaccacctt aaaaattttt 26220
tcgatctggc ccatttggtc gccgctcaga ccttcatacg ggatgatgac atggatgtcg 26280

atcttcagcc cattttcacc gctcaggaca atcctttgga tcggagttac ggacaccccg 26340
agattctgaa acaaactgga cacacctccc tgttcaagga cttggtccag gttgtagccg 26400
gctgtctgtc gccagtcccc aacgaaatct tcgagtgtga agaccatgga tccgggcccg 26460
gggttttctt caacgtctcc agcctgcttc agcaggctga agttagtagc tccgcttcct 26520
cgagctcgag atctggcgaa ggcgatgggg gtcttgaagg cgtgctggta ctccacgatg 26580
cccagctcgg tgttgctgtg cagctcctcc acgcggcgga aggcgaacat ggggcccccg 26640
ttctgcagga tgctggggtg gatggcgctc ttgaagtgca tgtggctgtc caccacgaag 26700

ctgtagtagc cgccgtcgcg caggctgaag gtgcgggcga agctgcccac cagcacgtta 26760

tcgcccatgg ggtgcaggtg ctccacggtg gcgttgctgc ggatgatctt gtcggtgaag 26820

atcacgctgt cctcggggaa gccggtgccc accaccttga agtcgccgat cacgcggccg 26880

gcctcgtagc ggtagctgaa gctcacgtgc agcacgccgc cgtcctcgta cttctcgatg 26940

cgggtgttgg tgtagccgcc gttgttgatg gcgtgcagga aggggttctc gtagccgctg 27000

gggtaggtgc cgaagtggta gaagccgtag cccatcacgt ggctcagcag gtaggggctg 27060

aaggtcaggg cgcctttggt gctcttcatc ttgttggtca tgcggccctg ctcgggggtg 27120

ccctctccgc cgcccaccag ctcgaactcc acgccgttca gggtgccggt gatgcggcac 27180

tcgatcttca tggcgggcat ggtggcgacc ggtagcgcta gcggcttcgg taccacgcgt 27240

tcgctcgaat taatcaattc tttgccaaaa tgatgagaca gcacaataac cagcacgttg 27300

cccaggagct gtaggaaaaa gaagaaggca tgaacatggt tagcagaggc tctagagccg 27360

ccggtcacac gccagaagcc gaaccccgcc ctgccccgtc ccccccgaag gcagccgtcc 27420

ccccgcggac agccccgagg ctggagaggg agaaggggac ggcggcgcgg cgacgcacga 27480

aggccctccc cgcccatttc cttcctgccg gggccctccc ggagcccctc aaggctttca 27540

cgcagccaca gaaaagaaac aagccgtcat taaaccaagc gctaattaca gcccggagga 27600

gaagggccgt cccgcccgct cacctgtggg agtaacgcgg tcagtcagag ccggggcggg 27660

cggcgcgagg cggcgcggag cggggcacgg ggcgaaggca acgcagcgac tcccgcccgc 27720

cgcgcgcttc gctttttata gggccgccgc cgccgccgcc tcgccataaa aggaaacttt 27780

cggagcgcgc cgctctgatt ggctgccgcc gcacctctcc gcctcgcccc gccccgcccc 27840

tcgccccgcc ccgccccgcc tggcgcgcgc cccccccccc cccccgcccc catcgctgca 27900

caaaataatt aaaaaataaa taaatacaaa attgggggtg gggaggggggg ggagatgggg 27960

agagtgaagc agaacgtggg gctcacctcg accatggtaa tagcgatgac taatacgtag 28020

atgtactgcc aagtaggaaa gtcccataag gtcatgtact gggcataatg ccaggcgggc 28080

catttaccgt cattgacgtc aataggggggc gtacttggca tatgatacac ttgatgtact 28140

gccaagtggg cagtttaccg taaatactcc acccattgac gtcaatggaa agtccctatt 28200

ggcgttacta tgggaacata cgtcattatt gacgtcaatg ggcgggggtc gttgggcggt 28260

cagccaggcg ggccatttac cgtaagttat gtaacgcgga acaacaacaa ttgcattcat 28320

tttatgtttc aggttcaggg ggaggtgtgg tcctgcgatt ccatcgagtg cacctacacc 28380

ctgctgaaga ccctatgcgg cctaagagac ctgctaccaa tgaattaaaa aaaaatgatt 28440

aataaaaaat cacttacttg aaatcagcaa taaggtctct gttgaaattt tctcccagca 28500

gcacctcact tccctcttcc caactctggt attctaaacc ccgttcagcg gcatactttc 28560

tccatacttt aaaggggatg tcaaatttta gctcctctcc tgtacccaca atcttcatgt 28620

ctttcttccc agatgaccaa gagagtccgg ctcagtgact ccttcaaccc tgtctacccc 28680

tatgaagatg aaagcacctc ccaacacccc tttataaacc cagggtttat ttccccaaat 28740

ggcttcacac aaagcccaga cggagttctt actttaaaat gtttaacccc actaacaacc 28800

acaggcggat ctctacagct aaaagtggga gggggactta cagtggatga cactgatggt 28860

accttacaag aaaacatacg tgctacagca cccattacta aaaataatca ctctgtagaa 28920

ctatccattg gaaatggatt agaaactcaa aacaataaac tatgtgccaa attgggaaat 28980

gggttaaaat ttaacaacgg tgacatttgt ataaaggata gtattaacac cttatggact 29040

ggaataaacc ctccacctaa ctgtcaaatt gtggaaaaca ctaatacaaa tgatggcaaa 29100

cttactttag tattagtaaa aaatggaggg cttgttaatg gctacgtgtc tctagttggt 29160

gtatcagaca ctgtgaacca aatgttcaca caaaagacag caaacatcca attaagatta 29220

tattttgact cttctggaaa tctattaact gaggaatcag acttaaaaat tccacttaaa 29280

aataaatctt ctacagcgac cagtgaaact gtagccagca gcaaagcctt tatgccaagt 29340

actacagctt atcccttcaa caccactact agggatagtg aaaactacat tcatggaata 29400

tgttactaca tgactagtta tgatagaagt ctatttccct tgaacatttc tataatgcta 29460

aacagccgta tgatttcttc caatgttgcc tatgccatac aatttgaatg gaatctaaat 29520

gcaagtgaat ctccagaaag caacatagct acgctgacca catccccctt tttcttttct 29580

tacattacag aagacgacaa ctaaaataaa gtttaagtgt ttttatttaa aatcacaaaa 29640

ttcgagtagt tattttgcct ccaccttccc atttgacaga atacacagtc ctttctcccc 29700

ggctggcctt aaaaagcatc atatcatggg taacagacat attcttaggt gttatattcc 29760

acacggtttc ctgtcgagcc aaacgctcat cagtgatatt aataaactcc ccgggcagct 29820

cacttaagtt catgtcgctg tccagctgct gagccacagg ctgctgtcca acttgcggtt 29880

gcttaacggg cggcgaagga gaagtccacg cctacatggg ggtagagtca taatcgtgca 29940

tcaggatagg gcggtggtgc tgcagcagcg cgcgaataaa ctgctgccgc cgccgctccg 30000

tcctgcagga atacaacatg gcagtggtct cctcagcgat gattcgcacc gcccgcagca 30060

taaggcgcct tgtcctccgg gcacagcagc gcaccctgat ctcacttaaa tcagcacagt 30120

aactgcagca cagcaccaca atattgttca aaatcccaca gtgcaaggcg ctgtatccaa 30180

agctcatggc ggggaccaca gaacccacgt ggccatcata ccacaagcgc aggtagatta 30240

agtggcgacc cctcataaac acgctggaca taaacattac ctcttttggc atgttgtaat 30300

tcaccacctc ccggtaccat ataaacctct gattaaacat ggcgccatcc accaccatcc 30360

taaaccagct ggccaaaacc tgcccgccgg ctatacactg cagggaaccg ggactggaac 30420

aatgacagtg gagagcccag gactcgtaac catggatcat catgctcgtc atgatatcaa 30480

tgttggcaca acacaggcac acgtgcatac acttcctcag gattacaagc tcctcccgcg 30540
ttagaaccat atcccaggga acaacccatt cctgaatcag cgtaaatccc acactgcagg 30600
gaagacctcg cacgtaactc acgttgtgca ttgtcaaagt gttacattcg ggcagcagcg 30660
gatgatcctc cagtatggta gcgcgggttt ctgtctcaaa aggaggtaga cgatccctac 30720
tgtacggagt gcgccgagac aaccgagatc gtgttggtcg tagtgtcatg ccaaatggaa 30780
cgccggacgt agtcattctc gtattttgta tagcaaaacg cggccctggc agaacacact 30840
cttcttcgcc ttctatcctg ccgcttagcg tgttccgtgt gatagttcaa gtacagccac 30900

actcttaagt tggtcaaaag aatgctggct tcagttgtaa tcaaaactcc atcgcatcta 30960
attgttctga ggaaatcatc cacggtagca tatgcaaatc ccaaccaagc aatgcaactg 31020
gattgcgttt caagcaggag aggagaggga agagacggaa gaaccatgtt aatttttatt 31080
ccaaacgatc tcgcagtact tcaaattgta gatcgcgcag atggcatctc tcgcccccac 31140
tgtgttggtg aaaaagcaca gctaaatcaa aagaaatgcg attttcaagg tgctcaacgg 31200
tggcttccaa caaagcctcc acgcgcacat ccaagaacaa aagaatacca aaagaaggag 31260
cattttctaa ctcctcaatc atcatattac attcctgcac cattcccaga taattttcag 31320

ctttccagcc ttgaattatt cgtgtcagtt cttgtggtaa atccaatcca cacattacaa 31380
acaggtcccg gagggcgccc tccaccacca ttcttaaaca caccctcata atgacaaaat 31440
atcttgctcc tgtgtcacct gtagcgaatt gagaatggca acatcaattg acatgccctt 31500
ggctctaagt tcttctttaa gttctagttg taaaaactct ctcatattat caccaaactg 31560
cttagccaga agcccccggg gaacaagagc aggggacgct acagtgcagt acaagcgcag 31620
acctccccaa ttggctccag caaaaacaag attggaataa gcatattggg aaccaccagt 31680
aatatcatcg aagttgctgg aaatataatc aggcagagtt tcttgtagaa attgaataaa 31740

agaaaaattt gccaaaaaaa cattcaaaac ctctgggatg caaatgcaat aggttaccgc 31800
gctgcgctcc aacattgtta gttttgaatt agtctgcaaa aataaaaaaa aaacaagcgt 31860
catatcatag tagcctgacg aacaggtgga taaatcagtc tttccatcac aagacaagcc 31920
acagggtctc cagctcgacc ctcgtaaaac ctgtcatcgt gattaaacaa cagcaccgaa 31980
agttcctcgc ggtgaccagc atgaataagt cttgatgaag catacaatcc agacatgtta 32040
gcatcagtta aggagaaaaa acagccaaca tagcctttgg gtataattat gcttaatcgt 32100
aagtatagca aagccacccc tcgcggatac aaagtaaaag gcacaggaga ataaaaaata 32160

taattatttc tctgctgctg tttaggcaac gtcgcccccg gtccctctaa atacacatac 32220

aaagcctcat cagccatggc ttaccagaga aagtacagcg ggcacacaaa ccacaagctc 32280

taaagtcact ctccaacctc tccacaatat atatacacaa gccctaaact gacgtaatgg 32340

gactaaagtg taaaaaatcc cgccaaaccc aacacacacc ccgaaactgc gtcaccaggg 32400

aaaagtacag tttcacttcc gcaatcccaa caagcgtcac ttcctctttc tcacggtacg 32460

tcacatccca ttaacttaca acgtcatttt cccacggccg cgccgcccct tttaaccgtt 32520

aaccccacag ccaatcacca cacggcccac actttttaaa atcacctcat ttacatattg 32580

gcaccattcc atctataagg tatattattg atgatggcca agctatttag gtgacactat 32640

agaatactca agctatgcat caagcttggt accgagctcg gatccactag taacggccgc 32700

cagtgtgctg gaattcgccc ttgt 32724

<210> 58

<211> 32724

<212> DNA

<213> Artificial Sequence

<220><223> PS3

<400> 58

catcatcaat aatatacctt atagatggaa tggtgccaat atgtaaatga ggtgatttta 60

aaaagtgtgg gccgtgtggt gattggctgt ggggttaacg gttaaaaggg gcggcgcggc 120

cgtgggaaaa tgacgtttta tgggggtgga gtttataact tcgtatagca tacattatac 180

gaagttattt ttgcaagttg tcgcgggaaa tgttacgcat aaaaaggctt cttttctcac 240

ggaactactt agttttccca cggtatttaa caggaaatga ggtagttttg accggatgca 300

agtgaaaatt gctgattttc gcgcgaaaac tgaatgagga agtgtttttc tgaataatgt 360

ggtatttatg gcagggtgga gtatttgttc agggccaggt agactttgac ccattacgtg 420

gaggtttcga ttaccgtgtt ttttacctga atttccgcgt accgtgtcaa agtcttctgt 480

ttttacgtag gtgtcagctg atcgctaggg tatttataac ttcgtatagc atacattata 540

cgaagttatg gaatgtttat gccttaccag tgtaacatga atcatgtgaa agtgttgttg 600

gaaccagatg ccttttccag aatgagccta acaggaatct ttgacatgaa cacgcaaatc 660

tggaagatcc tgaggtatga tgatacgaga tcgagggtgc gcgcatgcga atgcggaggc 720

aagcatgcca ggttccagcc ggtgtgtgta gatgtgaccg aagatctcag accggatcat 780

ttggttattg cccgcactgg agcagagttc ggatccagtg gagaagaaac tgactaaggt 840

gagtattggg aaaactttgg ggtgggattt tcagatggac agattgagta aaaatttgtt 900

ttttctgtct tgcagctgac atgagtggaa atgcttcttt taaggggga gtcttcagcc 960

cttatctgac agggcgtctc ccatcctggg caggagttcg tcagaatgtt atgggatcta 1020

ctgtggatgg aagacccgtt caacccgcca attcttcaac gctgacctat gctactttaa 1080
gttcttcacc tttggacgca gctgcagccg ctgccgccgc ctctgtcgcc gctaacactg 1140
tgcttggaat gggttactat ggaagcatcg tggctaattc cacttcctct aataaccctt 1200
ctacactgac tcaggacaag ttacttgtcc ttttggccca gctggaggct ttgacccaac 1260
gtctgggtga actttctcag caggtggccg agttgcgagt acaaactgag tctgctgtcg 1320
gcacggcaaa gtctaaataa aaaaaattcc agaatcaatg aataaataaa cgagcttgtt 1380
gttgatttaa aatcaagtgt ttttatttca ttttcgcgcc acggtatgcc ctggaccacc 1440

gatctcgatc attgagaact cggtggattt tttccagaat cctatagagg tgggattgaa 1500
tgtttagata catgggcatt aggccgtctt tggggtggag atagctccat tgaagggatt 1560
catgctccgg ggtagtgttg taaatcaccc agtcataaca aggtcgcagt gcatggtgtt 1620
gcacaatatc ttttagaagt aggctgattg ccacagataa gcccttggtg taggtgttta 1680
caaaccggtt gagctgggag gggtgcattc gaggtgaaat tatgtgcatt ttggattgga 1740
tttttaagtt ggcaatattg ccgccaagat cccgtcttgg gttcatgtta tgaaggacta 1800
ccaagacggt gtatccggta catttaggaa atttatcgtg cagcttggat ggaaaagcgt 1860

ggaaaaattt ggagacaccc ttgtgtcctc cgagattttc catgcactca tccatgataa 1920
tagcaatggg gccgtgggca gcggcgcggg caaacacgtt ccgtgggtct gacacatcat 1980
agttatgttc ctgagttaaa tcatcataag ccattttaat gaatttgggg cggagcgtac 2040
cagattgggg tatgaatgtt ccttcgggcc ccggagcata gttcccctca cagatttgca 2100
tttcccaagc tttcagttct gagggtggaa tcatgtccac ctgggggggct atgaagaaca 2160
ccgtttcggg ggcgggggtg attagttggg atgatagcaa gtttctgagc aattgagatt 2220
tgccacatcc ggtggggcca taaataattc cgattacagg ttgcaggtgg tagtttaggg 2280

aacggcaact gccgtcttct cgaagcaagg gggccacctc gttcatcatt tcccttacat 2340
gcatattttc ccgcaccaaa tccattagga ggcgctctcc tcctagtgat agaagttctt 2400
gtagtgagga aaagtttttc agcggtttta gaccgtcagc catgggcatt ttggaaagag 2460
tttgctgcaa aagttctagt ctgttccaca gttcagtgat gtgttctatg gcatctcgat 2520
ccagcagacc tcctcgtttc gcgggtttgg acggctcctg gagtagggta tgagacgatg 2580
ggcgtccagc gctgccaggg ttcggtcctt ccagggtctc agtgttcgag tcagggttgt 2640
ttccgtcaca gtgaaggggt gtgcgcctgc ttgggcgctt gccagggtgc gcttcagact 2700

cattctgctg gtggagaact tctgtcgctt ggcgccctgt atgtcggcca agtagcagtt 2760

taccatgagt tcgtagttga gcgcctcggc tgcgtggcct ttggcgcgga gcttaccttt 2820

ggaagttttc ttgcataccg ggcagtatag gcatttcagc gcatacagct tgggcgcaag 2880

gaaaatggat tctggggagt atgcatccgc gccgcaggag gcgcaaacag tttcacattc 2940

caccagccag gttaaatccg gttcattggg gtcaaaaaca agttttccgc catatttttt 3000

gatgcgtttc ttacctttgg tctccataag ttcgtgtcct cgttgagtga caaacaggct 3060

gtccgtatct ccgtagactg attttacagg cctcttctcc agtggagtgc ctcggtcttc 3120

ttcgtacagg aactctgacc actctgatac aaaggcgcgc gtccaggcca gcacaaagga 3180

ggctatgtgg gaggggtagc gatcgttgtc aaccaggggg tccacctttt ccaaagtatg 3240

caaacacatg tcaccctctt caacatccag gaatgtgatt ggcttgtagg tgtatttcac 3300

gtgacctggg gtccccgctg gggggggtata aaaggggggcg gttctttgct cttcctcact 3360

gtcttccgga tcgctgtcca ggaacgtcag ctgttggggt aggtattccc tctcgaaggc 3420

gggcatgacc tctgcactca ggttgtcagt ttctaagaac gaggaggatt tgatattgac 3480

agtgccggtt gagatgcctt tcatgaggtt ttcgtccatt tggtcagaaa acacaatttt 3540

tttattgtca agtttggtgg caaatgatcc atacagggcg ttggataaaa gtttggcaat 3600

ggatcgcatg gtttggttct tttccttgtc cgcgcgctct ttggcggcga tgttgagttg 3660

gacatactcg cgtgccaggc acttccattc ggggaagata gttgttaatt catctggcac 3720

gattctcact tgccaccctc gattatgcaa ggtaattaaa tccacactgg tggccacctc 3780

gcctcgaagg ggttcattgg tccaacagag cctacctcct ttcctagaac agaaaggggg 3840

aagtgggtct agcataagtt catcgggagg gtctgcatcc atggtaaaga ttcccggaag 3900

taaatcctta tcaaaatagc tgatgggagt ggggtcatct aaggccattt gccattctcg 3960

agctgccagt gcgcgctcat atgggttaag gggactgccc cagggcatgg gatgggtgag 4020

agcagaggca tacatgccac agatgtcata gacgtagatg ggatcctcaa agatgcctat 4080

gtaggttgga tagcatcgcc cccctctgat acttgctcgc acatagtcat atagttcatg 4140

tgatggcgct agcagccccg gacccaagtt ggtgcgattg ggtttttctg ttctgtagac 4200

gatctggcga aagatggcgt gagaattgga agagatggtg ggtctttgaa aaatgttgaa 4260

atgggcatga ggtagaccta cagagtctct gacaaagtgg gcataagatt cttgaagctt 4320

ggttaccagt tcggcggtga caagtacgtc tagggcgcag tagtcaagtg tttcttgaat 4380

gatgtcataa cctggttggt ttttcttttc ccacagttcg cggttgagaa ggtattcttc 4440

gcgatccttc cagtactctt ctagcggaaa cccgtctttg tctgcacggt aagatcctag 4500

catgtagaac tgattaactg ccttgtaagg gcagcagccc ttctctacgg gtagagagta 4560

tgcttgagca gcttttcgta gcgaagcgtg agtaagggca aaggtgtctc tgaccatgac 4620

tttgagaaat tggtatttga agtccatgtc gtcacaggct ccctgttccc agagttggaa 4680

gtctacccgt ttcttgtagg cggggttggg caaagcgaaa gtaacatcat tgaagagaat 4740

cttaccggct ctgggcataa aattgcgagt gatgcggaaa ggctgtggta cttccgctcg 4800

attgttgatc acctgggcag ctaggacgat ttcgtcgaaa ccgttgatgt tgtgtcctac 4860

gatgtataat tctatgaaac gcggcgtgcc tctgacgtga ggtagcttac tgagctcatc 4920

aaaggttagg tctgtggggt cagataaggc gtagtgttcg agagcccatt cgtgcaggtg 4980

aggatttgca tgtaggaatg atgaccaaag atctaccgcc agtgctgttt gtaactggtc 5040

ccgatactga cgaaaatgcc ggccaattgc catttttct ggagtgacac agtagaaggt 5100

tctggggtct tgttgccatc gatcccactt gagtttaatg gctagatcgt gggccatgtt 5160

gacgagacgc tcttctcctg agagtttcat gaccagcatg aaaggaacta gttgtttgcc 5220

aaaggatccc atccaggtgt aagtttccac atcgtaggtc aggaagagtc tttctgtgcg 5280

aggatgagag ccgatcggga agaactggat ttcctgccac cagttggagg attggctgtt 5340

gatgtgatgg aagtagaagt ttctgcggcg cgccgagcat tcgtgtttgt gcttgtacag 5400

acggccgcag tagtcgcagc gttgcacggg ttgtatctcg tgaatgagct gtacctggct 5460

tcccttgacg agaaatttca gtgggaagcc gaggcctggc gattgtatct cgtgctcttc 5520

tatattcgct gtatcggcct gttcatcttc tgtttcgatg gtggtcatgc tgacgagccc 5580

ccgcgggagg caagtccaga cctcggcgcg ggaggggcgg agctgaagga cgagagcgcg 5640

caggctggag ctgtccagag tcctgagacg ctgcggactc aggttagtag gtagggacag 5700

aagattaact tgcatgatct tttccagggc gtgcgggagg ttcagatggt acttgatttc 5760

cacaggttcg tttgtagaga cgtcaatggc ttgcagggtt ccgtgtcctt tgggcgccac 5820

taccgtacct ttgtttttc ttttgatcgg tggtggctct cttgcttctt gcatgctcag 5880

aagcggtgac ggggacgcgc gccgggcggc agcggttgtt ccggacccgg gggcatggct 5940

ggtagtggca cgtcggcgcc gcgcacgggc aggttctggt attgcgctct gagaagactt 6000

gcgtgcgcca ccacgcgtcg attgacgtct tgtatctgac gtctctgggt gaaagctacc 6060

ggccccgtga gcttgaacct gaaagagagt tcaacagaat caatttcggt atcgttaacg 6120

gcagcttgtc tcagtatttc ttgtacgtca ccagagttgt cctggtaggc gatctccgcc 6180

atgaactgct cgatttcttc ctcctgaaga tctccgcgac ccgctctttc gacggtggcc 6240

gcgaggtcat tggagatacg gcccatgagt tgggagaatg cattcatgcc cgcctcgttc 6300

cagacgcggc tgtaaaccac ggcccccctcg gagtctcttg cgcgcatcac cacctgagcg 6360

aggttaagct ccacgtgtct ggtgaagacc gcatagttgc ataggcgctg aaaaaggtag 6420

ttgagtgtgg tggcaatgtg ttcggcgacg aagaaataca tgatccatcg tctcagcggc 6480

atttcgctaa catcgcccag agcttccaag cgctccatgg cctcgtagaa gtccacggca 6540
aaattaaaaa actgggagtt tcgcgcggac acggtcaatt cctcctcgag aagacggatg 6600
agttcggcta tggtggcccg tacttcgcgt tcgaaggctc ccgggatctc ttcttcctct 6660
tctatctctt cttccactaa catctcttct tcgtcttcag gcgggggcgg agggggcacg 6720
cggcgacgtc gacggcgcac gggcaaacgg tcgatgaatc gttcaatgac ctctccgcgg 6780
cggcggcgca tggtttcagt gacggcgcgg ccgttctcgc gcggtcgcag agtaaaaaca 6840
ccgccgcgca tctccttaaa gtggtgactg ggaggttctc cgtttgggag ggagagggcg 6900

ctgattatac attttattaa ttggcccgta gggactgcgc gcagagatct gatcgtgtca 6960
agatccacgg gatctgaaaa cctttcgacg aaagcgtcta accagtcaca gtcacaaggt 7020
aggctgagta cggcttcttg tgggcggggg tggttatgtg ttcggtctgg gtcttctgtt 7080
tcttcttcat ctcgggaagg tgagacgatg ctgctggtga tgaaattaaa gtaggcagtt 7140
ctaagacggc ggatggtggc gaggagcacc aggtctttgg gtccggcttg ctggatacgc 7200
aggcgattgg ccattcccca agcattatcc tgacatctag caagatcttt gtagtagtct 7260
tgcatgagcc gttctacggg cacttcttcc tcacccgttc tgccatgcat acgtgtgagt 7320

ccaaatccgc gcattggttg taccagtgcc aagtcagcta cgactctttc ggcgaggatg 7380
gcttgctgta cttgggtaag ggtggcttga aagtcatcaa aatccacaaa gcggtggtaa 7440
gcccctgtat taatggtgta agcacagttg gccatgactg accagttaac tgtctggtga 7500
ccagggcgca cgagctcggt gtatttaagg cgcgaatagg cgcgggtgtc aaagatgtaa 7560
tcgttgcagg tgcgcaccag atactggtac cctataagaa aatgcggcgg tggttggcgg 7620
tagagaggcc atcgttctgt agctggagcg ccaggggcga ggtcttccaa cataaggcgg 7680
tgatagccgt agatgtacct ggacatccag gtgattcctg cggcggtagt agaagcccga 7740

ggaaactcgc gtacgcggtt ccaaatgttg cgtagcggca tgaagtagtt cattgtaggc 7800
acggtttgac cagtgaggcg cgcgcagtca ttgatgctct atagacacgg agaaaatgaa 7860
agcgttcagc gactcgactc cgtagcctgg aggaacgtga acgggttggg tcgcggtgta 7920
ccccggttcg agacttgtac tcgagccggc cggagccgcg gctaacgtgg tattggcact 7980
cccgtctcga cccagcctac aaaaatccag gatacggaat cgagtcgttt tgctggtttc 8040
cgaatggcag ggaagtgagt cctatttttt ttttttttgc cgctcagatg catcccgtgc 8100
tgcgacagat gcgcccccaa caacagcccc cctcgcagca gcagcagcag caaccacaaa 8160

aggctgtccc tgcaactact gcaactgccg ccgtgagcgg tgcgggacag cccgcctatg 8220

atctggactt ggaagagggc gaaggactgg cacgtctagg tgcgccttcg cccgagcggc 8280

atccgcgagt tcaactgaaa aaagattctc gcgaggcgta tgtgccccaa cagaacctat 8340

ttagagacag aagcggcgag gagccggagg agatgcgagc ttcccgcttt aacgcgggtc 8400

gtgagctgcg tcacggtttg gaccgaagac gagtgttgcg agacgaggat ttcgaagttg 8460

atgaagtgac agggatcagt cctgccaggg cacacgtggc tgcagccaac cttgtatcgg 8520

cttacgagca gacagtaaag gaagagcgta acttccaaaa gtcttttaat aatcatgtgc 8580

gaaccctgat tgcccgcgaa gaagttaccc ttggtttgat gcatttgtgg gatttgatgg 8640

aagctatcat tcagaaccct actagcaaac ctctgaccgc ccagctgttt ctggtggtgc 8700

aacacagcag agacaatgag gctttcagag aggcgctgct gaacatcacc gaacccgagg 8760

ggagatggtt gtatgatctt atcaacattc tacagagtat catagtgcag gagcggagcc 8820

tgggcctggc cgagaaggta gctgccatca attactcggt tttgagcttg ggaaaatatt 8880

acgctcgcaa aatctacaag actccatacg ttcccataga caaggaggtg aagatagatg 8940

ggttctacat gcgcatgacg ctcaaggtct tgaccctgag cgatgatctt ggggtgtatc 9000

gcaatgacag aatgcatcgc gcggttagcg ccagcaggag gcgcgagtta agcgacaggg 9060

aactgatgca cagtttgcaa agagctctga ctggagctgg aaccgagggt gagaattact 9120

tcgacatggg agctgacttg cagtggcagc ctagtcgcag ggctctgagc gccgcgacgg 9180

caggatgtga gcttccttac atagaagagg cggatgaagg cgaggaggaa gagggcgagt 9240

acttggaaga ctgatggcac aacccgtgtt ttttgctaga tggaacagca agcaccggat 9300

cccgcaatgc gggcggcgct gcagagccag ccgtccggca ttaactcctc ggacgattgg 9360

acccaggcca tgcaacgtat catggcgttg acgactcgca accccgaagc ctttagacag 9420

caaccccagg ccaaccgtct atcggccatc atggaagctg tagtgccttc ccgatctaat 9480

cccactcatg agaaggtcct ggccatcgtg aacgcgttgg tggagaacaa agctattcgt 9540

ccagatgagg ccggactggt atacaacgct ctcttagaac gcgtggctcg ctacaacagt 9600

agcaatgtgc aaaccaattt ggaccgtatg ataacagatg tacgcgaagc cgtgtctcag 9660

cgcgaaaggt tccagcgtga tgccaacctg ggttcgctgg tggcgttaaa tgctttcttg 9720

agtactcagc ctgctaatgt gccgcgtggt caacaggatt atactaactt tttaagtgct 9780

ttgagactga tggtatcaga agtacctcag agcgaagtgt atcagtccgg tcctgattac 9840

ttctttcaga ctagcagaca gggcttgcag acggtaaatc tgagccaagc ttttaaaaac 9900

cttaaaggtt tgtggggagt gcatgccccg gtaggagaaa gagcaaccgt gtctagcttg 9960

ttaactccga actcccgcct gttattactg ttggtagctc ctttcaccga cagcggtagc 10020

atcgaccgta attcctattt gggttaccta ctaaacctgt atcgcgaagc catagggcaa 10080

agtcaggtgg acgagcagac ctatcaagaa attacccaag tcagtcgcgc tttgggacag 10140

gaagacactg gcagtttgga agccactctg aacttcttgc ttaccaatcg gtctcaaaag 10200

atccctcctc aatatgctct tactgcggag gaggagagga tccttagata tgtgcagcag 10260

agcgtgggat tgtttctgat gcaagagggg gcaactccga ctgcagcact ggacatgaca 10320

gcgcgaaata tggagcccag catgtatgcc agtaaccgac ctttcattaa caaactgctg 10380

gactacttgc acagagctgc cgctatgaac tctgattatt tcaccaatgc catcttaaac 10440

ccgcactggc tgcccccacc tggtttctac acgggcgaat atgacatgcc cgaccctaat 10500

gacggatttc tgtgggacga cgtggacagc gatgtttttt cacctctttc tgatcatcgc 10560

acgtggaaaa aggaaggcgg tgatagaatg cattcttctg catcgctgtc cggggtcatg 10620

ggtgctaccg cggctgagcc cgagtctgca agtccttttc ctagtctacc cttttctcta 10680

cacagtgtac gtagcagcga agtgggtaga ataagtcgcc cgagtttaat gggcgaagag 10740

gagtacctaa acgattcctt gctcagaccg gcaagagaaa aaaatttccc aaacaatgga 10800

atagaaagtt tggtggataa aatgagtaga tggaagactt atgctcagga tcacagagac 10860

gagcctggga tcatggggac tacaagtaga gcgagccgta gacgccagcg ccatgacaga 10920

cagaggggtc ttgtgtggga cgatgaggat tcggccgatg atagcagcgt gttggacttg 10980

ggtgggagag gaaggggcaa cccgtttgct catttgcgcc ctcgcttggg tggtatgttg 11040

tgaaaaaaaa taaaaaagaa aaactcacca aggccatggc gacgagcgta cgttcgttct 11100

tctttattat ctgtgtctag tataatgagg cgagtcgtgc taggcggagc ggtggtgtat 11160

ccggagggtc ctcctccttc gtacgagagc gtgatgcagc agcagcaggc gacggcggtg 11220

atgcaatccc cactggaggc tccctttgtg cctccgcgat acctggcacc tacggagggc 11280

agaaacagca ttcgttactc ggaactggca cctcagtacg ataccaccag gttgtatctg 11340

gtggacaaca agtcggcgga cattgcttct ctgaactatc agaatgacca cagcaacttc 11400

ttgaccacgg tggtgcagaa caatgacttt acccctacgg aagccagcac ccagaccatt 11460

aactttgatg aacgatcgcg gtggggcggt cagctaaaga ccatcatgca tactaacatg 11520

ccaaacgtga acgagtatat gtttagtaac aagttcaaag cgcgtgtgat ggtgtccaga 11580

aaacctcccg acggtgctgc agttggggat acttatgatc acaagcagga tattttggaa 11640

tatgagtggt tcgagtttac tttgccagaa ggcaactttt cagttactat gactattgat 11700

ttgatgaaca atgccatcat agataattac ttgaaagtgg gtagacagaa tggagtgctt 11760

gaaagtgaca ttggtgttaa gttcgacacc aggaacttca agctgggatg ggatcccgaa 11820

accaagttga tcatgcctgg agtgtatacg tatgaagcct tccatcctga cattgtctta 11880

ctgcctggct gcggagtgga ttttaccgag agtcgtttga gcaaccttct tggtatcaga 11940

aaaaaacagc catttcaaga gggttttaag attttgtatg aagatttaga aggtggtaat 12000
attccggccc tcttggatgt agatgcctat gagaacagta agaaagaaca aaaagccaaa 12060
atagaagctg ctacagctgc tgcagaagct aaggcaaaca tagttgccag cgactctaca 12120
agggttgcta acgctggaga ggtcagagga gacaattttg cgccaacacc tgttccgact 12180
gcagaatcat tattggccga tgtgtctgaa ggaacggacg tgaaactcac tattcaacct 12240
gtagaaaaag atagtaagaa tagaagctat aatgtgttgg aagacaaaat caacacagcc 12300
tatcgcagtt ggtatctttc gtacaattat ggcgatcccg aaaaaggagt gcgttcctgg 12360

acattgctca ccacctcaga tgtcacctgc ggagcagagc aggtttactg gtcgcttcca 12420
gacatgatga aggatcctgt cactttccgc tccactagac aagtcagtaa ctaccctgtg 12480
gtgggtgcag agcttatgcc cgtcttctca aagagcttct acaacgaaca agctgtgtac 12540
tcccagcagc tccgccagtc cacctcgctt acgcacgtct tcaaccgctt tcctgagaac 12600
cagattttaa tccgtccgcc ggcgcccacc attaccaccg tcagtgaaaa cgttcctgct 12660
ctcacagatc acgggaccct gccgttgcgc agcagtatcc ggggagtcca acgtgtgacc 12720
gttactgacg ccagacgccg cacctgtccc tacgtgtaca aggcactggg catagtcgca 12780

ccgcgcgtcc tttcaagccg cactttctaa aaaaaaaatg tccattctta tctcgcccag 12840
taataacacc ggttggggtc tgcgcgctcc aagcaagatg tacggaggcg cacgcaaacg 12900
ttctacccaa catcccgtgc gtgttcgcgg acattttcgc gctccatggg gtgccctcaa 12960
gggccgcact cgcgttcgaa ccaccgtcga tgatgtaatc gatcaggtgg ttgccgacgc 13020
ccgtaattat actcctactg cgcctacatc tactgtggat gcagttattg acagtgtagt 13080
ggctgacgct cgcaactatg ctcgacgtaa gagccggcga aggcgcattg ccagacgcca 13140
ccgagctacc actgccatgc gagccgcaag agctctgcta cgaagagcta gacgcgtggg 13200

gcgaagagcc atgcttaggg cggccagacg tgcagcttcg ggcgccagcg ccggcaggtc 13260
ccgcaggcaa gcagccgctg tcgcagcggc gactattgcc gacatggccc aatcgcgaag 13320
aggcaatgta tactgggtgc gtgacgctgc caccggtcaa cgtgtacccg tgcgcacccg 13380
tccccctcgc acttagaaga tactgagcag tctccgatgt tgtgtcccag cggcgaggat 13440
gtccaagcgc aaatacaagg aagaaatgct gcaggttatc gcacctgaag tctacggcca 13500
accgttgaag gatgaaaaaa aaccccgcaa aatcaagcgg gttaaaaagg acaaaaaaga 13560
agaggaagat ggcgatgatg ggctggcgga gtttgtgcgc gagtttgccc cacggcgacg 13620

cgtgcaatgg cgtgggcgca aagttcgaca tgtgttgaga cctggaactt cggtggtctt 13680

tacacccggc gagcgttcaa gcgctacttt taagcgttcc tatgatgagg tgtacgggga 13740

tgatgatatt cttgagcagg cggctgaccg attaggcgag tttgcttatg gcaagcgtag 13800

tagaataact tccaaggatg agacagtgtc aataccctg gatcatggaa atcccacccc 13860

tagtcttaaa ccggtcactt tgcagcaagt gttacccgta actccgcgaa caggtgttaa 13920

acgcgaaggt gaagatttgt atcccactat gcaactgatg gtacccaaac gccagaagtt 13980

ggaggacgtt ttggagaaag taaaagtgga tccagatatt caacctgagg ttaaagtgag 14040

acccattaag caggtagcgc ctggtctggg ggtacaaact gtagacatta agattcccac 14100

tgaaagtatg gaagtgcaaa ctgaacccgc aaagcctact gccacctcca ctgaagtgca 14160

aacggatcca tggatgccca tgcctattac aactgacgcc gccggtccca ctcgaagatc 14220

ccgacgaaag tacggtccag caagtctgtt gatgcccaat tatgttgtac acccatctat 14280

tattcctact cctggttacc gaggcactcg ctactatcgc agccgaaaca gtacctcccg 14340

ccgtcgccgc aagacacctg caaatcgcag tcgtcgccgt agacgcacaa gcaaaccgac 14400

tcccggcgcc ctggtgcggc aagtgtaccg caatggtagt gcggaacctt tgacactgcc 14460

gcgtgcgcgt taccatccga gtatcatcac ttaatcaatg ttgccgctgc ctccttgcag 14520

atatggccct cacttgtcgc cttcgcgttc ccatcactgg ttaccgagga agaaactcgc 14580

gccgtagaag agggatgttg ggacgcggaa tgcgacgcta caggcgacgg cgtgctatcc 14640

gcaagcaatt gcggggtggt tttttaccag ccttaattcc aattatcgct gctgcaattg 14700

gcgcgatacc aggcatagct tccgtggcgg ttcaggcctc gcaacgacat tgacattgga 14760

aaaaaaacgt ataaataaaa aaaaatacaa tggactctga cactcctggt cctgtgacta 14820

tgttttctta gagatggaag acatcaattt ttcatccttg gctccgcgac acggcacgaa 14880

gccgtacatg ggcacctgga gcgacatcgg cacgagccaa ctgaacgggg gcgccttcaa 14940

ttggagcagt atctggagcg ggcttaaaaa ttttggctca accataaaaa catacgggaa 15000

caaagcttgg aacagcagta caggacaggc gcttagaaat aaacttaaag accagaactt 15060

ccaacaaaaa gtagtcgatg ggatagcttc cggcatcaat ggagtggtag atttggctaa 15120

ccaggctgtg cagaaaaaga taaacagtcg tttggacccg ccgccagcaa ccccaggtga 15180

aatgcaagtg gaggaagaaa ttcctccgcc agaaaaacga ggcgacaagc gtccgcgtcc 15240

cgatttggaa gagacgctgg tgacgcgcgt agatgaaccg ccttcttatg aggaagcaac 15300

gaagcttgga atgcccacca ctagaccgat agccccaatg gccaccgggg tgatgaaacc 15360

ttctcagttg catcgacccg tcaccttgga tttgcccccc ccccctgctg ctactgctgt 15420

acccgcttct aagcctgtcg ctgccccgaa accagtcgcc gtagccaggt cacgtcccgg 15480

gggcgctcct cgtccaaatg cgcactggca aaatactctg aacagcatcg tgggtctagg 15540

cgtgcaaagt gtaaaacgcc gtcgctgctt ttaattaaat atggagtagc gcttaacttg 15600

cctatctgtg tatatgtgtc attacacgcc gtcacagcag cagaggaaaa aaggaagagg 15660

tcgtgcgtcg acgctgagtt actttcaaga tggccacccc atcgatgctg ccccaatggg 15720

catacatgca catcgccgga caggatgctt cggagtacct gagtccgggt ctggtgcagt 15780

tcgcccgcgc cacagacacc tacttcaatc tgggaaataa gtttagaaat cccaccgtag 15840

cgccgaccca cgatgtgacc accgaccgta gccagcggct catgttgcgc ttcgtgcccg 15900

ttgaccggga ggacaataca tactcttaca aagtgcggta caccctggcc gtgggcgaca 15960

acagagtgct ggatatggcc agcacgttct ttgacattag gggcgtgttg gacagaggtc 16020

ccagtttcaa accctattct ggtacggctt acaactctct ggctcctaaa ggcgctccaa 16080

atgcatctca atggattgca aaaggcgtac caactgcagc agccgcaggc aatggtgaag 16140

aagaacatga aacagaggag aaaactgcta cttacacttt tgccaatgct cctgtaaaag 16200

ccgaggctca aattacaaaa gagggcttac caataggttt ggagatttca gctgaaaacg 16260

aatctaaacc catctatgca gataaacttt atcagccaga acctcaagtg ggagatgaaa 16320

cttggactga cctagacgga aaaaccgaag agtatggagg cagggctcta aagcctacta 16380

ctaacatgaa accctgttac gggtcctatg cgaagcctac taatttaaaa ggtggtcagg 16440

caaaaccgaa aaactcggaa ccgtcgagtg aaaaaattga atatgatatt gacatggaat 16500

tttttgataa ctcatcgcaa agaacaaact tcagtcctaa aattgtcatg tatgcagaaa 16560

atgtaggttt ggaaacgcca gacactcatg tagtgtacaa acctggaaca gaagacacaa 16620

gttccgaagc taatttggga caacagtcta tgcccaacag acccaactac attggcttca 16680

gagataactt tattggactc atgtactata acagtactgg taacatgggg gtgctggctg 16740

gtcaagcgtc tcagttaaat gcagtggttg acttgcagga cagaaacaca gaactttctt 16800

accaactctt gcttgactct ctgggcgaca gaaccagata ctttagcatg tggaatcagg 16860

ctgtggacag ttatgatcct gatgtacgtg ttattgaaaa tcatggtgtg gaagatgaac 16920

ttcccaacta ttgttttcca ctggacggca taggtgttcc aacaaccagt tacaaatcaa 16980

tagttccaaa tggagaagat aataataatt ggaaagaacc tgaagtaaat ggaacaagtg 17040

agatcggaca gggtaatttg tttgccatgg aaattaacct tcaagccaat ctatggcgaa 17100

gtttccttta ttccaatgtg gctctgtatc tcccagactc gtacaaatac accccgtcca 17160

atgtcactct tccagaaaac aaaaacacct acgactacat gaacgggcgg gtggtgccgc 17220

catctctagt agacacctat gtgaacattg gtgccaggtg gtctctggat gccatggaca 17280

atgtcaaccc attcaaccac caccgtaacg ctggcttgcg ttaccgatct atgcttctgg 17340

gtaacggacg ttatgtgcct ttccacatac aagtgcctca aaaattcttc gctgttaaaa 17400

acctgctgct tctcccaggc tcctacactt atgagtggaa ctttaggaag gatgtgaaca 17460
tggttctaca gagttccctc ggtaacgacc tgcgggtaga tggcgccagc atcagtttca 17520
cgagcatcaa cctctatgct actttttcc ccatggctca caacaccgct tccacccttg 17580
aagccatgct gcggaatgac accaatgatc agtcattcaa cgactaccta tctgcagcta 17640
acatgctcta ccccattcct gccaatgcaa ccaatattcc catttccatt ccttctcgca 17700
actgggcggc tttcagaggc tggtcattta ccagactgaa aaccaaagaa actccctctt 17760
tggggtctgg atttgacccc tactttgtct attctggttc tattccctac ctggatggta 17820

ccttctacct gaaccacact tttaagaagg tttccatcat gtttgactct tcagtgagct 17880
ggcctggaaa tgacaggtta ctatctccta acgaatttga aataaagcgc actgtggatg 17940
gcgaaggcta caacgtagcc caatgcaaca tgaccaaaga ctggttcttg gtacagatgc 18000
tcgccaacta caacatcggc tatcagggct tctacattcc agaaggatac aaagatcgca 18060
tgtattcatt tttcagaaac ttccagccca tgagcaggca ggtggttgat gaggtcaatt 18120
acaaagactt caaggccgtc gccataccct accaacacaa caactctggc tttgtgggtt 18180
acatggctcc gaccatgcgc caaggtcaac cctatcccgc taactatccc tatccactca 18240

ttggaacaac tgccgtaaat agtgttacgc agaaaaagtt cttgtgtgac agaaccatgt 18300
ggcgcatacc gttctcgagc aacttcatgt ctatgggggc ccttacagac ttgggacaga 18360
atatgctcta tgccaactca gctcatgctc tggacatgac ctttgaggtg gatcccatgg 18420
atgagcccac cctgctttat cttctcttcg aagttttcga cgtggtcaga gtgcatcagc 18480
cacaccgcgg catcatcgag gcagtctacc tgcgtacacc gttctcggcc ggtaacgcta 18540
ccacgtaaga agcttcttgc ttcttgcaaa tagcagctgc aaccatggcc tgcggatccc 18600
aaaacggctc cagcgagcaa gagctcagag ccattgtcca agacctgggt tgcggaccct 18660

atttttggg aacctacgat aagcgcttcc cggggttcat ggcccccgat aagctcgcct 18720
gtgccattgt aaatacggcc ggacgtgaga cgggggggaga gcactggttg gctttcggtt 18780
ggaacccacg ttctaacacc tgctaccttt ttgatccttt tggattctcg gatgatcgtc 18840
tcaaacagat ttaccagttt gaatatgagg gtctcctgcg ccgcagcgct cttgctacca 18900
aggaccgctg tattacgctg gaaaaatcta cccagaccgt gcagggcccc cgttctgccg 18960
cctgcggact tttctgctgc atgttccttc acgcctttgt gcactggcct gaccgtccca 19020
tggacggaaa ccccaccatg aaattgctaa ctggagtgcc aaacaacatg cttcattctc 19080

ctaaagtcca gcccaccctg tgtgacaatc aaaaagcact ctaccatttt cttaataccc 19140

attcgcctta ttttcgctct catcgtacac acatcgaaag ggccactgcg ttcgaccgta 19200

tggatgttca ataatgactc atgtaaacaa cgtgttcaat aaacatcact ttatttttt 19260

acatgtatca aggctctgga ttacttattt atttacaagt cgaatgggtt ctgacgagaa 19320

tcagaatgac ccgcaggcag tgatacgttg cggaactgat acttgggttg ccacttgaat 19380

tcgggaatca ccaacttggg aaccggtata tcgggcagga tgtcactcca cagctttctg 19440

gtcagctgca aagctccaag caggtcagga gccgaaatct tgaaatcaca attaggacca 19500

gtgctctgag cgcgagagtt gcggtacacc ggattgcagc actgaaacac catcagcgac 19560

ggatgtctca cgcttgccag cacggtggga tctgcaatca tgcccacatc cagatcttca 19620

gcattggcaa tgctgaacgg ggtcatcttg caggtctgcc tacccatggc gggcacccaa 19680

ttaggcttgt ggttgcaatc gcagtgcagg gggatcagta tcatcttggc ctgatcctgt 19740

ctgattcctg gatacacggc tctcatgaaa gcatcatatt gcttgaaagc ctgctgggct 19800

ttactaccct cggtataaaa catcccgcag gacctgctcg aaaactggtt agctgcacag 19860

ccggcatcat tcacacagca gcgggcgtca ttgttggcta tttgcaccac acttctgccc 19920

cagcggtttt gggtgatttt ggttcgctcg ggattctcct ttaaggctcg ttgtccgttc 19980

tcgctggcca catccatctc gataatctgc tccttctgaa tcataatatt gccatgcagg 20040

cacttcagct tgccctcata atcattgcag ccatgaggcc acaacgcaca gcctgtacat 20100

tcccaattat ggtgggcgat ctgagaaaaa gaatgtatca ttccctgcag aaatcttccc 20160

atcatcgtgc tcagtgtctt gtgactagtg aaagttaact ggatgcctcg gtgctcttcg 20220

tttacgtact ggtgacagat gcgcttgtat tgttcgtgtt gctcaggcat tagtttaaaa 20280

caggttctaa gttcgttatc cagcctgtac ttctccatca gcagacacat cacttccatg 20340

cctttctccc aagcagacac caggggcaag ctaatcggat tcttaacagt gcaggcagca 20400

gctcctttag ccagagggtc atctttagcg atcttctcaa tgcttctttt gccatccttc 20460

tcaacgatgc gcacgggcgg gtagctgaaa cccactgcta caagttgcgc ctcttctctt 20520

tcttcttcgc tgtcttgact gatgtcttgc atggggatat gtttggtctt ccttggcttc 20580

tttttggggg gtatcggagg aggaggactg tcgctccgtt ccggagacag ggaggattgt 20640

gacgtttcgc tcaccattac caactgactg tcggtagaag aacctgaccc cacacggcga 20700

caggtgtttt tcttcggggg cagaggtgga ggcgattgcg aagggctgcg gtccgacctg 20760

gaaggcggat gactggcaga accccttccg cgttcggggg tgtgctccct gtggcggtcg 20820

cttaactgat ttccttcgcg gctggccatt gtgttctcct aggcagagaa acaacagaca 20880

tggaaactca gccattgctg tcaacatcgc cacgagtgcc atcacatctc gtcctcagcg 20940

acgaggaaaa ggagcagagc ttaagcattc caccgcccag tcctgccacc acctctaccc 21000

tagaagataa ggaggtcgac gcatctcatg acatgcagaa taaaaaagcg aaagagtctg 21060

agacagacat cgagcaagac ccgggctatg tgacaccggt ggaacacgag gaagagttga 21120

aacgctttct agagagagag gatgaaaact gcccaaaaca gcgagcagat aactatcacc 21180

aagatgctgg aaatagggat cagaacaccg actacctcat agggcttgac ggggaagacg 21240

cgctccttaa acatctagca agacagtcgc tcatagtcaa ggatgcatta ttggacagaa 21300

ctgaagtgcc catcagtgtg gaagagctca gctgcgccta cgagcttaac ctttttcac 21360

ctcgtactcc ccccaaacgt cagccaaacg gcacctgcga gccaaatcct cgcttaaact 21420

tttatccagc ttttgctgtg ccagaagtac tggctaccta tcacatcttt tttaaaaatc 21480

aaaaaattcc agtctcctgc cgcgctaatc gcacccgcgc cgatgcccta ctcaatctgg 21540

gacctggttc acgcttacct gatatagctt ccttggaaga ggttccaaag atcttcgagg 21600

gtctgggcaa taatgagact cgggccgcaa atgctctgca aaagggagaa aatggcatgg 21660

atgagcatca cagcgttctg gtggaattgg aaggcgataa tgccagactc gcagtactca 21720

agcgaagcgt cgaggtcaca cacttcgcat atcccgctgt caacctgccc cctaaagtca 21780

tgacggcggt catggaccag ttactcatta agcgcgcaag tccccttca gaagacatgc 21840

atgacccaga tgcctgtgat gagggtaaac cagtggtcag tgatgagcag ctaacccgat 21900

ggctgggcac cgactctccc cgggatttgg aagagcgtcg caagcttatg atggccgtgg 21960

tgctggttac cgtagaacta gagtgtctcc gacgtttctt taccgattca gaaaccttgc 22020

gcaaactcga agagaatctg cactacactt ttagacacgg ctttgtgcgg caggcatgca 22080

agatatctaa cgtggaactc accaacctgg tttcctacat gggtattctg catgagaatc 22140

gcctaggaca aagcgtgctg cacagcaccc ttaaggggga agcccgccgt gattacatcc 22200

gcgattgtgt ctatctctac ctgtgccaca cgtggcaaac cggcatgggt gtatggcagc 22260

aatgtttaga agaacagaac ttgaaagagc ttgacaagct cttacagaaa tctcttaagg 22320

ttctgtggac agggttcgac gagcgcaccg tcgcttccga cctggcagac ctcatcttcc 22380

cagagcgtct cagggttact ttgcgaaacg gattgcctga ctttatgagc cagagcatgc 22440

ttaacaattt tcgctctttc atcctggaac gctccggtat cctgcccgcc acctgctgcg 22500

cactgccctc cgactttgtg cctctcacct accgcgagtg ccccccgccg ctatggagtc 22560

actgctacct gttccgtctg gccaactatc tctcctacca ctcggatgtg atcgaggatg 22620

tgagcggaga cggcttgctg gagtgccact gccgctgcaa tctgtgcacg ccccaccggt 22680

ccctagcttg caacccccag ttgatgagcg aaacccagat aataggcacc tttgaattgc 22740

aaggccccag cagccaaggc gatgggtctt ctcctgggca aagtttaaaa ctgaccccgg 22800

gactgtggac ctccgcctac ttgcgcaagt ttgctccgga agattaccac ccctatgaaa 22860

tcaagttcta tgaggaccaa tcacagcctc caaaggccga actttcggct tgcgtcatca 22920
cccaggggggc aattctggcc caattgcaag ccatccaaaa atcccgccaa gaatttctac 22980
tgaaaaaggg taaggggtc taccttgacc cccagaccgg cgaggaactc aacacaaggt 23040
tccctcagga tgtcccaacg acgagaaaac aagaagttga aggtgcagcc gccgccccca 23100
gaagatatgg aggaagattg ggacagtcag gcagaggagg cggaggagga cagtctggag 23160
gacagtctgg aggaagacag tttggaggag gaaaacgagg aggcagagga ggtggaagaa 23220
gtaaccgccg acaaacagtt atcctcggct gcggagacaa gcaacagcgc taccatctcc 23280

gctccgagtc gaggaacccg gcggcgtccc agcagtagat gggacgagac cggacgcttc 23340
ccgaacccaa ccagcgcttc caagaccggt aagaaggatc ggcagggata caagtcctgg 23400
cgggggcata agaatgccat catctcctgc ttgcatgagt gcgggggcaa catatccttc 23460
acgcggcgct acttgctatt ccaccatggg gtgaactttc cgcgcaatgt tttgcattac 23520
taccgtcacc tccacagccc ctactatagc cagcaaatcc cgacagtctc gacagataaa 23580
gacagcggcg gcgacctcca acagaaaacc agcagcggca gttagaaaat acacaacaag 23640
tgcagcaaca ggaggattaa agattacagc caacgagcca gcgcaaaccc gagagttaag 23700

aaatcggatc tttccaaccc tgtatgccat cttccagcag agtcggggtc aagagcagga 23760
actgaaaata aaaaaccgat ctctgcgttc gctcaccaga agttgtttgt atcacaagag 23820
cgaagatcaa cttcagcgca ctctcgagga cgccgaggct ctcttcaaca agtactgcgc 23880
gctgactctt aaagagtagg cagcgaccgc gcttattcaa aaaaggcggg aattacatca 23940
tcctcgacat gagtaaagaa attcccacgc cttacatgtg gagttatcaa ccccaaatgg 24000
gattggcagc aggcgcctcc caggactact ccacccgcat gaattggctc agcgccgggc 24060
cttctatgat ttctcgagtt aatgatatac gcgcctaccg aaaccaaata cttttggaac 24120

agtcagctct taccaccacg ccccgccaac accttaatcc cagaaattgg cccgccgccc 24180
tagtgtacca ggaaagtccc gctcccacca ctgtattact tcctcgagac gcccaggccg 24240
aagtccaaat gactaatgca ggtgcgcagt tagctggcgg ctccacccta tgtcgtcaca 24300
ggcctcggca taatataaaa cgcctgatga tcagaggccg aggtatccag ctcaacgacg 24360
agtcggtgag ctctccgctt ggtctacgac cagacggaat ctttcagatt gccggctgcg 24420
ggagatcttc cttcacccct cgtcaggctg ttctgacttt ggaaagttcg tcttcgcaac 24480
cccgctcggg cggaatcggg accgttcaat ttgtagagga gtttactccc tctgtctact 24540

tcaacccctt ctccggatct cctgggcact acccggacga gttcataccg aacttcgacg 24600

cgattagcga gtcagtggac ggctacgatt gatgtctggt gacgcggctg agctatctcg 24660

gctgcgacat ctagaccact gccgccgctt tcgctgcttt gcccgggaac ttattgagtt 24720

catctacttc gaactcccca aggatcaccc tcaaggtccg gcccacggag tgcggattac 24780

tatcgaaggc aaaatagact ctcgcctgca acgaattttc tcccagcggc ccgtgctgat 24840

cgagcgagac cagggaaaca ccacggttag taatcaatta cggggtcatt agttcatagc 24900

ccatatatgg agttaacttg tttattgcag cttataatgg ttacaaataa agcaatagca 24960

tcacaaattt cacaaataaa gcatttttt cactgcattc tagttgtggt ttgtccaaac 25020

tcatcaatgt atcttatcat gtctgctcga agcggccggc cgccccgact ctagagtcgc 25080

ggcctcatta ggaagttcct atactttcta gagaatagga acttctcaga agaactcgtc 25140

aagaaggcga tagaaggcga tgcgctgcga atcgggagcg gcgataccgt aaagcacgag 25200

gaagcggtca gcccattcgc cgccaagctc ttcagcaata tcacgggtag ccaacgctat 25260

gtcctgatag cggtccgcca cacccagccg gccacagtcg atgaatccag aaaagcggcc 25320

attttccacc atgatattcg gcaagcaggc atcgccatgg gtcacgacga gatcctcgcc 25380

gtcgggcatg cgcgccttga gcctggcgaa cagttcggct ggcgcgagcc cctgatgctc 25440

ttcgtccaga tcatcctgat cgacaagacc ggcttccatc cgagtacgtg ctcgctcgat 25500

gcgatgtttc gcttggtggt cgaatgggca ggtagccgga tcaagcgtat gcagccgccg 25560

cattgcatca gccatgatgg atactttctc ggcaggagca aggtgagatg acaggagatc 25620

ctgccccggc acttcgccca atagcagcca gtcccttccc gcttcagtga caacgtcgag 25680

cacagctgcg caaggaacgc ccgtcgtggc cagccacgat agccgcgctg cctcgtcctg 25740

cagttcattc agggcaccgg acaggtcggt cttgacaaaa agaaccgggc gcccctgcgc 25800

tgacagccgg aacacggcgg catcagagca gccgattgcc tgttgtgccc agtcatagcc 25860

gaatagcctc tccacccaag cggccggaga acctgcgtgc aatccatctt gttcaatggc 25920

cgatcccata acacccctg tattactgtt tatgtaagca gacagtttta ctgttcgtga 25980

tgatatattt ttatcttgtg caatgtaaca ggttgtggcc atagcgggcc cgggattttc 26040

ctccacgtcc ccgcatgtta gaagacttcc cctgccctcg gctctggaag ttcctatact 26100

ttctagagaa taggaacttc ccgccagaat gcgttcgcac agccgccagc cggtcactcc 26160

gttgatggtt actcggaaca gcagggagcc gtcggggttg atcaggcgct cgtcgataat 26220

tttgttgccg ttccacaggg tccctgttac agtgatcttt ttgccgtcga acacggcgat 26280

gccttcatac ggccgtccga aatagtcgat catgttcggc gtaaccccgt cgattaccag 26340

tgtgccatag tgcaggatca ccttaaagtg atgatcatcc acagggtaca ccaccttaaa 26400

aatttttcg atctggccca tttggtcgcc gctcagacct tcatacggga tgatgacatg 26460

gatgtcgatc ttcagcccat tttcaccgct caggacaatc ctttggatcg gagttacgga 26520

caccccgaga ttctgaaaca aactggacac acctccctgt tcaaggactt ggtccaggtt 26580

gtagccggct gtctgtcgcc agtccccaac gaaatcttcg agtgtgaaga ccatggatcc 26640

gggcccgggg ttttcttcaa cgtctccagc ctgcttcagc aggctgaagt tagtagctcc 26700

gcttcctcga gctcgagatc tggcgaaggc gatgggggtc ttgaaggcgt gctggtactc 26760

cacgatgccc agctcggtgt tgctgtgcag ctcctccacg cggcggaagg cgaacatggg 26820

gcccccgttc tgcaggatgc tggggtggat ggcgctcttg aagtgcatgt ggctgtccac 26880

cacgaagctg tagtagccgc cgtcgcgcag gctgaaggtg cgggcgaagc tgcccaccag 26940

cacgttatcg cccatggggt gcaggtgctc cacggtggcg ttgctgcgga tgatcttgtc 27000

ggtgaagatc acgctgtcct cggggaagcc ggtgcccacc accttgaagt cgccgatcac 27060

gcggccggcc tcgtagcggt agctgaagct cacgtgcagc acgccgccgt cctcgtactt 27120

ctcgatgcgg gtgttggtgt agccgccgtt gttgatggcg tgcaggaagg ggttctcgta 27180

gccgctgggg taggtgccga agtggtagaa gccgtagccc atcacgtggc tcagcaggta 27240

ggggctgaag gtcagggcgc ctttggtgct cttcatcttg ttggtcatgc ggccctgctc 27300

gggggtgccc tctccgccgc ccaccagctc gaactccacg ccgttcaggg tgccggtgat 27360

gcggcactcg atcttcatgg cgggcatggt ggcgaccggt agcgctagcg gcttcggtac 27420

cacgcgttcg ctcgaattaa tcaattcttt gccaaaatga tgagacagca caataaccag 27480

cacgttgccc aggagctgta ggaaaaagaa gaaggcatga acatggttag cagaggctct 27540

agagccgccg gtcacacgcc agaagccgaa ccccgccctg ccccgtcccc cccgaaggca 27600

gccgtccccc cgcggacagc cccgaggctg gagagggaga aggggacggc ggcgcggcga 27660

cgcacgaagg ccctccccgc ccatttcctt cctgccgggg ccctcccgga gcccctcaag 27720

gctttcacgc agccacagaa aagaaacaag ccgtcattaa accaagcgct aattacagcc 27780

cggaggagaa gggccgtccc gcccgctcac ctgtgggagt aacgcggtca gtcagagccg 27840

gggcgggcgg cgcgaggcgg cgcggagcgg ggcacggggc gaaggcaacg cagcgactcc 27900

cgcccgccgc gcgcttcgct ttttataggg ccgccgccgc cgccgcctcg ccataaaagg 27960

aaactttcgg agcgcgccgc tctgattggc tgccgccgca cctctccgcc tcgccccgcc 28020

ccgcccctcg ccccgccccg ccccgcctgg cgcgcgcccc ccccccccca ccgcccccat 28080

cgctgcacaa aataattaaa aaataaataa atacaaaatt gggggtgggg agggggggga 28140

gatggggaga gtgaagcaga acgtggggct cacctcgacc atggtaatag cgatgactaa 28200

tacgtagatg tactgccaag taggaaagtc ccataaggtc atgtactggg cataatgcca 28260

ggcgggccat ttaccgtcat tgacgtcaat agggggcgta cttggcatat gatacacttg 28320

atgtactgcc aagtgggcag tttaccgtaa atactccacc cattgacgtc aatggaaagt 28380
ccctattggc gttactatgg gaacatacgt cattattgac gtcaatgggc gggggtcgtt 28440
gggcggtcag ccaggcgggc catttaccgt aagttatgta acgcggaaca acaacaattg 28500
cattcatttt atgtttcagg ttcaggggga ggtgtggtcc tgcgattcca tcgagtgcac 28560
ctacaccctg ctgaagaccc tatgcggcct aagagacctg ctaccaatga attaaaaaaa 28620
aatgattaat aaaaaatcac ttacttgaaa tcagcaataa ggtctctgtt gaaattttct 28680
cccagcagca cctcacttcc ctcttcccaa ctctggtatt ctaaaccccg ttcagcggca 28740

tactttctcc atactttaaa ggggatgtca aattttagct cctctcctgt acccacaatc 28800
ttcatgtctt tcttcccaga tgaccaagag agtccggctc agtgactcct tcaaccctgt 28860
ctacccctat gaagatgaaa gcacctccca acaccccttt ataaacccag ggtttatttc 28920
cccaaatggc ttcacacaaa gcccagacgg agttcttact ttaaaatgtt taaccccact 28980
aacaaccaca ggcggatctc tacagctaaa agtgggaggg ggacttacag tggatgacac 29040
tgatggtacc ttacaagaaa acatacgtgc tacagcaccc attactaaaa ataatcactc 29100
tgtagaacta tccattggaa atggattaga aactcaaaac aataaactat gtgccaaatt 29160

gggaaatggg ttaaaattta acaacggtga catttgtata aaggatagta ttaacacctt 29220
atggactgga ataaaccctc cacctaactg tcaaattgtg gaaaacacta atacaaatga 29280
tggcaaactt actttagtat tagtaaaaaa tggagggctt gttaatggct acgtgtctct 29340
agttggtgta tcagacactg tgaaccaaat gttcacacaa aagacagcaa acatccaatt 29400
aagattatat tttgactctt ctggaaatct attaactgag gaatcagact taaaaattcc 29460
acttaaaaat aaatcttcta cagcgaccag tgaaactgta gccagcagca aagcctttat 29520
gccaagtact acagcttatc ccttcaacac cactactagg gatagtgaaa actacattca 29580

tggaatatgt tactacatga ctagttatga tagaagtcta tttcccttga acatttctat 29640
aatgctaaac agccgtatga tttcttccaa tgttgcctat gccatacaat ttgaatggaa 29700
tctaaatgca agtgaatctc cagaaagcaa catagctacg ctgaccacat ccccctttt 29760
cttttcttac attacagaag acgacaacta aaataaagtt taagtgtttt tatttaaaat 29820
cacaaaattc gagtagttat tttgcctcca ccttcccatt tgacagaata cacagtcctt 29880
tctccccggc tggccttaaa aagcatcata tcatgggtaa cagacatatt cttaggtgtt 29940
atattccaca cggtttcctg tcgagccaaa cgctcatcag tgatattaat aaactccccg 30000

ggcagctcac ttaagttcat gtcgctgtcc agctgctgag ccacaggctg ctgtccaact 30060

tgcggttgct taacgggcgg cgaaggagaa gtccacgcct acatgggggt agagtcataa 30120

tcgtgcatca ggatagggcg gtggtgctgc agcagcgcgc gaataaactg ctgccgccgc 30180

cgctccgtcc tgcaggaata caacatggca gtggtctcct cagcgatgat tcgcaccgcc 30240

cgcagcataa ggcgccttgt cctccgggca cagcagcgca ccctgatctc acttaaatca 30300

gcacagtaac tgcagcacag caccacaata ttgttcaaaa tcccacagtg caaggcgctg 30360

tatccaaagc tcatggcggg gaccacagaa cccacgtggc catcatacca caagcgcagg 30420

tagattaagt ggcgacccct cataaacacg ctggacataa acattacctc ttttggcatg 30480

ttgtaattca ccacctcccg gtaccatata aacctctgat taaacatggc gccatccacc 30540

accatcctaa accagctggc caaaacctgc ccgccggcta tacactgcag ggaaccggga 30600

ctggaacaat gacagtggag agcccaggac tcgtaaccat ggatcatcat gctcgtcatg 30660

atatcaatgt tggcacaaca caggcacacg tgcatacact tcctcaggat tacaagctcc 30720

tcccgcgtta gaaccatatc ccagggaaca acccattcct gaatcagcgt aaatcccaca 30780

ctgcagggaa gacctcgcac gtaactcacg ttgtgcattg tcaaagtgtt acattcgggc 30840

agcagcggat gatcctccag tatggtagcg cgggtttctg tctcaaaagg aggtagacga 30900

tccctactgt acggagtgcg ccgagacaac cgagatcgtg ttggtcgtag tgtcatgcca 30960

aatggaacgc cggacgtagt cattctcgta ttttgtatag caaaacgcgg ccctggcaga 31020

acacactctt cttcgccttc tatcctgccg cttagcgtgt tccgtgtgat agttcaagta 31080

cagccacact cttaagttgg tcaaaagaat gctggcttca gttgtaatca aaactccatc 31140

gcatctaatt gttctgagga aatcatccac ggtagcatat gcaaatccca accaagcaat 31200

gcaactggat tgcgtttcaa gcaggagagg agagggaaga gacggaagaa ccatgttaat 31260

ttttattcca aacgatctcg cagtacttca aattgtagat cgcgcagatg gcatctctcg 31320

cccccactgt gttggtgaaa aagcacagct aaatcaaaag aaatgcgatt ttcaaggtgc 31380

tcaacggtgg cttccaacaa agcctccacg cgcacatcca agaacaaaag aataccaaaa 31440

gaaggagcat tttctaactc ctcaatcatc atattacatt cctgcaccat tcccagataa 31500

ttttcagctt tccagccttg aattattcgt gtcagttctt gtggtaaatc caatccacac 31560

attacaaaca ggtcccggag ggcgccctcc accaccattc ttaaacacac cctcataatg 31620

acaaaatatc ttgctcctgt gtcacctgta gcgaattgag aatggcaaca tcaattgaca 31680

tgcccttggc tctaagttct tctttaagtt ctagttgtaa aaactctctc atattatcac 31740

caaactgctt agccagaagc cccccgggaa caagagcagg ggacgctaca gtgcagtaca 31800

agcgcagacc tccccaattg gctccagcaa aaacaagatt ggaataagca tattgggaac 31860

caccagtaat atcatcgaag ttgctggaaa tataatcagg cagagtttct tgtagaaatt 31920

gaataaaaga aaaatttgcc aaaaaaacat tcaaaacctc tgggatgcaa atgcaatagg 31980

ttaccgcgct gcgctccaac attgttagtt ttgaattagt ctgcaaaaat aaaaaaaaaa 32040

caagcgtcat atcatagtag cctgacgaac aggtggataa atcagtcttt ccatcacaag 32100

acaagccaca gggtctccag ctcgaccctc gtaaaacctg tcatcgtgat taaacaacag 32160

caccgaaagt tcctcgcggt gaccagcatg aataagtctt gatgaagcat acaatccaga 32220

catgttagca tcagttaagg agaaaaaaca gccaacatag cctttgggta taattatgct 32280

taatcgtaag tatagcaaag ccacccctcg cggatacaaa gtaaaaggca caggagaata 32340

aaaaatataa ttatttctct gctgctgttt aggcaacgtc gcccccggtc cctctaaata 32400

cacatacaaa gcctcatcag ccatggctta ccagagaaag tacagcgggc acacaaacca 32460

caagctctaa agtcactctc caacctctcc acaatatata tacacaagcc ctaaactgac 32520

gtaatgggac taaagtgtaa aaaatcccgc caaacccaac acacaccccg aaactgcgtc 32580

accagggaaa agtacagttt cacttccgca atcccaacaa gcgtcacttc ctctttctca 32640

cggtacgtca catcccatta acttacaacg tcattttccc acggccgcgc cgcccctttt 32700

aaccgttaac cccacagcca atca 32724

<210> 59

<211> 32538

<212> DNA

<213> Artificial Sequence

<220><223> pHCA-AAVS1-globin-mgmt

<400> 59

catcatcaat aatatacctt attttggatt gaagccaata tgataatgag ggggtggagt 60

ttgtgacgtg gcgcggggcg tgggaacggg gcgggtgacg tagtagtgtg gcggaagtgt 120

gatgttgcaa gtgtggcgga acacatgtaa gcgacggatg tggcaaaagt gacgtttttg 180

gtgtgcgccg gtgtacacag gaagtgacaa ttttcgcgcg gttttaggcg gatgttgtag 240

taaatttggg cgtaaccgag taagatttgg ccattttcgc gggaaaactg aataagagga 300

agtgaaatct gaataatttt gtgttactca tagcgcgtaa tatttgtcta gggccgcggg 360

gactttgacc gtttacgtgg agactcgccc aggtgttttt ctcaggtgtt ttccgcgttc 420

cgggtcaaag ttggcgtttt gattcggccg aaggattaca tgagcttaga aatgtaatta 480

gcatagtgtg tggcatagtg tagataccaa ataaatatga tctctccttc tactcttgaa 540

aatgcaaaca cattcttggt ggtcctaaaa tagcctgtaa catggtttac tcagcagcat 600

ttgctattca aggcagatct gcctttagtc attggctgcg ctcctgaaca gctgtgtgaa 660

aggctaactt ttgtaaacca aatcaaaata aaatgcagca aaaatttgtc actgaaagga 720

aatcctcagt atatccttt atgaaatgaa agatccctca tccaaactta actttttaa 780

aagtgcgcat ttggagatat agccctttct tatgaatcct aattcaattt tggccataaa 840

cacacgttga tgttccccac cccaaagcac atagcaacaa gagtaggttc tatattgaaa 900

ataatgacaa tttaaaaaca tgtacttatt tcactgtatg tggacagtgt ctatgattgc 960

atcatgaagt gtcatataac catgtacgtg tacatgagag agagatagag agagaagtgg 1020

tagggtggtg gtggtagagg ggatggcgat agtaatcatg gtaatggtag aggtgatgga 1080

ggtggtaatg acggaggtaa gggtggtagt gatgatggtg gtggtggtaa tggtggtgga 1140

tgtggtggtg gcaattggga tggtgggatg gtggtagcca tggtgatggt ggtaatggtg 1200

ttgatttaaa gggtggtggt agtgaaggtg agggtagtgg tggtggaggt ggtggtgctg 1260

gtagcaatag tgatggtggt gatggtgttg atgagggtgt tgggatcagg gtgagttccc 1320

acagtatatt tcattcttgt tgtaccactc tgtcaacagc accactgact gggacagagg 1380

aagaaggcac actctgaatg tgttattaac agaaacctca aaacagtctg tctccttgta 1440

gtcattcaaa attatctttt tcttacctgg aaaactgaaa ctgaattacc gggaaaaaca 1500

caggagattt ttgtttgtta atatgctgcc aataaagtaa ttttatgtca aatttaacta 1560

caggaaaggg caaggcattt tctaagttcc ttagatgtca tgtggctaaa aaaaacaaaa 1620

ggatggacag cagttagata ctgtacactt agctgtttga agccatatat tcagaaagca 1680

gatgttggga gttggtgttt gaggactgat ttcctggagg tattttatat aggccaagtt 1740

cattgttcta aactctaagg gcttgacttg agggaggaaa agaggcaaga acatgtttag 1800

ttttgctgac agcatcacat gggcagccct aaggctagac aactttaggg cctgaagctt 1860

attctaggaa agaagcacct acagagtggc actgggctcc cctccactat agagatgaag 1920

tcatatgaca gtaaagggca ggcagggctg cctaggggggc ccagaactga cacttccatt 1980

agaatgagca caggccaggg agagaagtgg ggaaccagag agaaggagct ggaattcctt 2040

tctctccata cataaatgcc tgcagagtcc catttcagaa tccggcagac aaagccacca 2100

atgtgatccc catgacctta taaacattca ttaaaatgca tttcaaggca tgtgatggcc 2160

tccccacccc ctagataatg agaaaacaaa ggtttctctt ctgatagaga caagttcagc 2220

tctgaagtca acattatttc tggttctgtc tgaacaatga catatggcaa ctcttccctt 2280

tctatagttc tagtccagaa tgacaaaaaa ggggaaaaat ttcttagaga aggtagagat 2340

tatacgaata cagtccatga aatgagcata aggagaataa agaatataac ttatccaaag 2400

aagtctggca ggctgttata aatgcttgat tttggacact gtagttggag gtttaacatg 2460

gacaccaata aaaaggtcag caaagggtat gcactgttcc tattgggcaa gaagatagga 2520

ggtcaaaggt aaccaggaaa gataaactca gggagactta ttttccctcc agagggcact 2580

gggcttgtag gccctgggca aaattgtcaa aaaggtgaaa atcgcctgtg gtttatttag 2640

tctgctcttt cttcactagt gcctcaccag ttcagttcag gccaatttgc tagcctctaa 2700

ctcctagacc gtcagaactg ctgggcccttcaagacgggc tgctcacacc cactcatgtt 2760

aagcctggtg aggcctgtac tctgttttca caggaagaaa tcctcaccca gtcttcccca 2820

aacacattcc caggttctgt cattagtggg atagagatga ttactgtggg gagaagagaa 2880

acatctggat ggatttggtg aggttgatct atagaggaag taggtgctgc ctgaggtagc 2940

tgtaatagaa gctaaaggtc aaaggagagg gccctgtccc aatccagatg actccacttc 3000

tgctggaccc aggttcacaa gcttaatcta catttcacct aaatttggct aacaagccca 3060

aaatcacaca ggcaaaggga gaagtggagg cagaaccgag gttggaggcc accagggcca 3120

ccgggcagag atcatttaag cccaaccttc tcacttctcc ctgggctctg cctctcttaa 3180

aggaccttgt ggtgtgacct cttgtaggtc cctttcacac tcggggcctc agtttcccca 3240

ctgtaaagtg aatgggtccc agctttggta agcttatgct tacctgatgc tttcttcctg 3300

ggctgctctt gtagagaaaa gataaatctt cttcctccat ccacgagggc ttctttccct 3360

gggggtgaga gtaggctgag gagagccact tgcacacact cttaaagaaa gtattacctg 3420

caccagctca gtgagaggca cagatcagac tgttacttga atcaaattat gagcctcccc 3480

aaatatatct atgacattta aataggggat tacttgaaca tagactttgg gatccggtgt 3540

ggagtgcagg agactagcaa agtgaatcct gagagtagca ggtctgcacc tgttggatcg 3600

agaaaggcgg cctacaattc tggtcaaatg agctgtgctt attgacatat tctattagag 3660

agtactacca ggtcaccagt caccagaaag gctgccagct ctccaaccac ctccagggaa 3720

ctatcctgaa tggggcctta acaagcctaa gagagggttg gtttgggtcc caagccaata 3780

tttgctctgc tttatgtcag tcatatggaa cccaaaccaa ccctctccta tgtgcctcac 3840

cagtcggtgc agggatccca atttcaagtt tggtttttta tggtcaaagt ccagcataga 3900

ttaaatgaag gggtgtgatg atggtgttaa aagagaactc cagaccagtt taactcttgg 3960

acacacatcc catctcacca tggtgcttcc aaccttccag agatgatggg ctcctatttt 4020

ctgatgacaa agccctccac aggattgctg cctggccatc agggagtgcc tctgtaactg 4080

aggctgagat cccactttca gtcctccagc tgtggcccat ccctgctccg cccaccgggt 4140

atggcctgtc ctaggctctt aggtatggct gcattgtgaa atgatggcta cagagctggc 4200

atctcctgta gtctggttca tctagtgcac tacctcatag ttaaaagaaa tctgtttaag 4260

ccactgaggg tggctcctag tgccaactcc aagaacagga agcttccctt ttttgggagg 4320

aggggcagat ggtaacatgg atcgtccagg tcaatgggag cagggcaacc acagtaagta 4380
ctggacaaca acacaaaact ccatgtgtgg cttccatcga gtccctctcc aattggtttg 4440
gtcttctccg tcccatgcag cactttagca aggggcctgg ctgaaggcta tgaattgtgt 4500
ggagcctcct cattgcagtc tccaaccatc tgatgctggg aaaatgtcac caggatgcag 4560
ccatgccgtg tggccaatga accgagaaaa caccccttt ctagaatgct ctaaagaggc 4620
agaataatcc agaggtgagg aaggaaatac tccaccagag acccaggcag ttcctacaaa 4680
agccagactt tccttcacct agggagtgac aagaccagtg gaaaacactc tcaagcagta 4740

acccccaaat gctctgcaag ccagtggcgt ccagataccg cacaagcgag tgggctgtct 4800
aatcccatca tcatgatgta aatatctcta ggctgccccg ggctgtgcct gaccctgtct 4860
tcagctttcc acacctccac ctacagccca tgcacagaag gaccacccag gaatgctgca 4920
agtgtggcac ctccagggcc acccagggag aaggagggca gctatgctgg tggctccagg 4980
cccatttggc gggtggtacc ttcacaccac aaagcccaaa ctgaggcccc agatttggct 5040
gatgagggca tattggacag gggtcactta tgctcttccc cattgccacc tggcctctgg 5100
ctacctggac ttggctacct gtggatcctc tcacaggtgc caccatcttg gctgagtctc 5160

cagatgcgag gtccctgagg cagtggcagg cttctcgcta atgctgatgg gattaggaat 5220
gggataggtg gggagggccc tggactgggc cctgatgagc caagtgggtt tttagagggg 5280
ctactggtac atttcaggga caggacatct ggtagagcta agctggggca ataaggagcc 5340
actgctaatc tgagagctag aaacaatcag cttctgggtc attattaatt agggtagttt 5400
gggctgtgtg gaagtcacgt actatatggg gtagccacag ctctctctac agataatctc 5460
taagacttct gattgggacc gtgtgaatgc agtagcaata tctcttctta ctgccaggcc 5520
ctgccagtcc tgcctccacg ccctggctgg ccccccttat gatctgaccc atgccaggct 5580

gccatagtat gttacttctg cattagcact ccttgggacc tgcctctcca ctgtccctca 5640
gactttaaag aactatacaa acccaagggg ctcttcccaa gagaattgat atgacttgag 5700
gtgattccat ttctggaagt agtcactcca ttttctgcct cactctttca gtgcttcaca 5760
gagcaggttc gattaagcac acagattaat ggtaccgggg ccactaggga caggattggg 5820
caccagggac tccagagtcc aggcccaacc cctccccatt caacccagga ggccaggccc 5880
cagcccttcc gccctcagat gaaggagtcc aggccccag cctctcccca ttcagaccca 5940
ggggtccagg cccagccccg cctccctaag acccagaagt ccaggccccc agcccctcct 6000

ccctcagacc cacgagtcca ggccccagcc cctcctccct cggacccagg agtccaggcc 6060

cccagtccct ccaccctcag acccaggagt ccaggcccca gcccctcctc cctcggaccc 6120

aggagtccag gccccagccc ctcctctctc aaacccagga gcccaggccc ccagctcttc 6180

tctgttcagc cctaagaatc ctggctccag cccctcctac tctagccccc aaccccctag 6240

ccactaaggc aattggggtg caggaatggg ggcagggtac cagcctcacc aagtggttga 6300

taaacccacg tggggtaccc taagaacttg ggaacagcca cagcaggggg gcgatgcttg 6360

gggacctgcc tggagaagga tgcaggacga gaaacacagc cccaggtgga gaaactggcc 6420

gggaatcaag agtcacccag agacagtgac caaccatccc tgttttccta ggactgaggg 6480

tttcagtgct aaaactaggc tgtcctgggc aaacagcata agctggtcac cccacaccca 6540

gacctgaccc aaacccagct cccctgcttc ttggccacgt aacctgagaa gggaatccct 6600

cctctctgaa ccccagccca ccccaatgct ccaggcctcc tgggataccc cgaagagtga 6660

gtttgccaag cagtcacccc acagttggag gagaatccac ccaaaaggca gcctggtaga 6720

cagggctggg gtggcctctc gtggggtcca ggccaagtag gtggcctggg gcctctgggg 6780

gatgcagggg aagggggatg caggggaacg gggatgcagg ggaacggggc tcagtctgaa 6840

gagcagagcc aggaacccct gtagggaagg ggcaggagag ccaggggcat gagatggtgg 6900

acgaggaagg gggacaggga agcctgagcg cctctcctgg gcttgccaag gactcaaacc 6960

cagaagccca gagcagggcc ttagggaagc gggaccctgc tctgggcgga ggaatatgtc 7020

ccagatagca ctggggactc tttaaggaaa gaaggatgga gaaagagaaa gggagtagag 7080

gcggccacga cctggtgaac acctaggacg caccattctc acaaagggag ttttccacac 7140

ggacaccccc ctcctcacca cagccctgcc aggacggggc tggctactgg ccttatctca 7200

caggtaaaac tgacgcacgg aggaacaata taaattgggg actagaaagg tgaagagcca 7260

aagttagaac tcaggaccaa cttattctga ttttgttttt ccaaactgct tctcctcttg 7320

ggaagtgtaa ggaagctgca gcaccaggat cagtgaaacg caccagacgg ccgcgtcaga 7380

gcagctcagg ttctgggaga gggtagcgca gggtggccac tgagaaccgg gcaggtcacg 7440

catccccccc ttccctccca ccccctgcca agctctccct cccaggatcc tctctggctc 7500

catcgtaagc aaaccttaga ggttctggca aggagagaga tggctccagg aaatgggggt 7560

gtgtcaccag ataaggaatc tgcctaacag gaggtggggg ttagacccaa tatcaggaga 7620

ctaggaagga ggaggcctaa ggatggggct tttctgtcag gtaccgtcaa ggctgcagtg 7680

agacatgatc ttgccactgc actccagcct ggacagcaga gtgaaacctt gcctcacgaa 7740

acagaataca aaaacaaaca aacaaaaaac tgctccgcaa tgcgcttcct tgatgctcta 7800

ccacataggt ctgggtactt tgtacacatt atctcattgc tgttcataat tgttagatta 7860

attttgtaat attgatatta ttcctagaaa gctgaggcct caagatgata acttttattt 7920

tctggacttg taatagcttt ctcttgtatt caccatgttg taactttctt agagtagtaa 7980

caatataaag ttattgtgag tttttgcaaa cacagcaaac acaacgaccc atatagacat 8040

tgatgtgaaa ttgtctattg tcaatttatg ggaaaacaag tatgtacttt ttctactaag 8100

ccattgaaac aggaataaca gaacaagatt gaaagaatac attttccgaa attacttgag 8160

tattatacaa agacaagcac gtggacctgg gaggagggtt attgtccatg actggtgtgt 8220

ggagacaaat gcaggtttat aatagatggg atggcatcta gcgcaatgac tttgccatca 8280

cttttagaga gctcttgggg accccagtac acaagagggg acgcagggta tatgtagaca 8340

tctcattctt tttcttagtg tgagaataag aatagccatg acctgagttt atagacaatg 8400

agccctttc tctctcccac tcagcagcta tgagatggct tgccctgcct ctctactagg 8460

ctgactcact ccaaggccca gcaatgggca gggctctgtc agggctttga tagcactatc 8520

tgcagagcca gggccgagaa ggggtggact ccagagactc tccctcccat tcccgagcag 8580

ggtttgctta tttatgcatt taaatgatat atttatttta aaagaaataa caggagactg 8640

cccagccctg gctgtgacat ggaaactatg tagaatattt tgggttccat tttttttcc 8700

ttctttcagt tagaggaaaa ggggctcagg atccacttgc ccagtgttct tccttagttc 8760

ctaccttcga ccttgatcct cctttatctt cctgaaccct gctgagatga tctatgtggg 8820

gagaatggct tctttgagaa acatcttctt cgttagtggc ctgcccctca ttcccacttt 8880

aatatccaga atcactataa gaagaatata ataagaggaa taactcttat tataggtaag 8940

ggaaaattaa gaggcatacg tgatgggatg agtaagagag gagagggaag gattaatgga 9000

cgataaaatc tactactatt tgttgagacc ttttatagtc taatcaattt tgctattgtt 9060

ttccatcctc acgctaactc cataaaaaaa cactattatt atctttattt tgccatgaca 9120

agactgagct cagaagagtc aagcatttgc ctaaggtcgg acatgtcaga ggcagtgcca 9180

gacctatgtg agactctgca gctactgctc atgggccctg tgctgcactg atgaggagga 9240

tcagatggat ggggcaatga agcaaaggaa tcattctgtg gataaaggag acagccatga 9300

agaagtctat gactgtaaat ttgggagcag gagtctctaa ggacttggat ttcaaggaat 9360

tttgactcag caaacacaag accctcacgg tgactttgcg agctggtgtg ccagatgtgt 9420

ctatcagagg ttccagggag ggtggggtgg ggtcagggct ggccaccagc tatcagggcc 9480

cagatgggtt ataggctggc aggctcagat aggtggttag gtcaggttgg tggtgctggg 9540

tggagtccat gactcccagg agccaggaga gatagaccat gagtagaggg cagacatggg 9600

aaaggtgggg gaggcacagc atagcagcat ttttcattct actactacat gggactgctc 9660

ccctataccc ccagctaggg gcaagtgcct tgactcctat gttttcagga tcatcatcta 9720

taaagtaaga gtaataattg tgtctatctc atagggttat tatgaggatc aaaggagatg 9780

cacactctct ggaccagtgg cctaacagtt caggacagag ctatgggctt cctatgtatg 9840
ggtcagtggt ctcaatgtag caggcaagtt ccagaagata gcatcaacca ctgttagaga 9900
tatactgcca gtctcagagc ctgatgttaa tttagcaatg ggctgggacc ctcctccagt 9960
agaaccttct aaccagctgc tgcagtcaaa gtcgaatgca gctggttaga ctttttttaa 10020
tgaaagcttg catgcagcac tttgggaggc tgaggtgggt ggactgcttg gagctcagga 10080
gttcaagacc atcttggaca acatggtgat accctgcctc tacaaaaagt acaaaaatta 10140
gcctggcatg gtggtgtgca cctgtaatcc cagctattag ggtggctgag gcaggagaat 10200

tgcttgaacc cgggaggcgg aggttgcagt gagctgagat cgtgccactg cactccagcc 10260
tgggggacag agcacattat aattaactgt tatttttac ttggactctt gtggggaata 10320
agatacatgt tttattctta tttatgattc aagcactgaa aatagtgttt agcatccagc 10380
aggtgcttca aaaccatttg ctgaatgatt actatacttt ttacaagctc agctccctct 10440
atcccttcca gcatcctcat ctctgattaa ataagcttca gtttttcctt agttcctgtt 10500
acatttctgt gtgtctccat tagtgacctc ccatagtcca agcatgagca gttctggcca 10560
ggcccctgtc ggggtcagtg ccccaccccc gccttctggt tctgtgtaac cttctaagca 10620

aaccttctgg ctcaagcaca gcaatgctga gtcatgatga gtcatgctga ggcttagggt 10680
gtgtgcccag atgttctcag cctagagtga tgactcctat ctgggtcccc agcaggatgc 10740
ttacagggca gatggcaaaa aaaaggagaa gctgaccacc tgactaaaac tccacctcaa 10800
acggcatcat aaagaaaatg gatgcctgag acagaatgtg acatattcta gaatatatta 10860
tttcctgaat atatatatat atatacacat ataccatatg aaacacctct aggctataag 10920
gcaacagagc tccttttttt tttttctgtg ctttcctggc tgtccaaatc tctaatgata 10980
agcatacttc tattcaatga gaatattctg taagattata gttaagaatt gtgggagcca 11040

ttccgtctct tatagttaaa tttgagcttc ttttatgatc actgtttttt taatatgctt 11100
taagttctgg ggtacatgtg ccatggtggt ttgctgcacc catcaacccg tcatctacat 11160
taggtatttc tcctaatgct atccttcccc tagcccccca cccccaacag gccccagtgt 11220
gtgatgttcc cctccctgtg tccatggatc actggttttt ttttgttttt ttttttttt 11280
taaagtctca gttaaatttt tggaatgtaa tttattttcc tggtatccta aggacttgca 11340
agttatctgg tcactttagc cctcacgttt tgatgataat cacatatttg taaacacaac 11400
acacacacac acacacacac acatatatat atatataaaa catatatata cataaacaca 11460

cataacatat ttatcgggca tttctgagca actaatcatg caggactctc aaacactaac 11520

ctatagcctt ttctatgtat ctacttgtgt agaaaccaag cgtggggact gagaaggcaa 11580

tagcaggagc attctgactc tcactgcctt tagctaggcc cctccctcat cacagctcag 11640

catagtcctg agctcttatc tatatccaca cacagtttct gacgctgccc agctatcacc 11700

atcccaagtc taaagaaaaa aataatgggt ttgcccatct ctgttgatta gaaaacaaaa 11760

caaaataaaa taagccccta agctcccaga aaacatgact aaaccagcaa gaagaagaaa 11820

atacaatagg tatatgagga gactggtgac actaagtgtc tgaatgaggc ttgagtacag 11880

aaaagaggct ctagcagcat agtggtttag aggagatgtt tctttccttc acagatgcct 11940

tagcctcaat aagcttgcgg ttgtggaagt ttacttgttt atcaccggtg acgtccatga 12000

gcaaattaag aaaaacaaca acaaatgaat gcatatatat gtatatgtat gtgtgtatat 12060

atacacatat atatatatat ttttttttctt ttcttaccag aaggttttaa tccaaataag 12120

gagaagatat gcttagaact gaggtagagt tttcatccat tctgtcctgt aagtattttg 12180

catattctgg agacgcagga agagatccat ctacatatcc caaagctgaa ttatggtaga 12240

caaaactctt ccacttttag tgcatcaatt tcttatttgt gtaataagaa aattgggaaa 12300

acgatcttca atatgcttac caagctgtga ttccaaatat tacgtaaata cacttgcaaa 12360

ggaggatgtt tttagtagca atttgtactg atggtatggg gccaagagat atatcttaga 12420

gggagggctg agggtttgaa gtccaactcc taagccagtg ccagaagagc caaggacagg 12480

tacggctgtc atcacttaga cctcaccctg tggagccaca ccctagggtt ggccaatcta 12540

ctcccaggag cagggagggc aggagccagg gctgggcata aaagtcaggg cagagccatc 12600

tattgcttac atttgcttct gacacaactg tgttcactag caacctcaaa cagacaccat 12660

gggtcatttc acagaggagg acaaggctac tatcacaagc ctgtggggca aggtgaatgt 12720

ggaagatgct ggaggagaaa ccctgggaag gtaggctctg gtgaccagga caagggaggg 12780

aaggaaggac cctgtgcctg gcaaaagtcc aggtcgcttc tcaggatttg tggcaccttc 12840

tgactgtcaa actgttcttg tcaatctcac aggctcctgg ttgtctaccc atggacccag 12900

aggttctttg acagctttgg caacctgtcc tctgcctctg ccatcatggg caaccccaaa 12960

gtcaaggcac atggcaagaa ggtgctgact tccttgggag atgccataaa gcacctggat 13020

gatctcaagg gcacctttgc ccagctgagt gaactgcact gtgacaagct gcatgtggat 13080

cctgagaact tcaaggtgag tccaggagat gtttcagccc tgttgccttt agtctcgagg 13140

caacttagac aactgagtat tgatctgagc acagcagggt gtgagctgtt tgaagatact 13200

ggggttgggg gtgaagaaac tgcagaggac taactgggct gagacccagt ggtaatgttt 13260

tagggcctaa ggagcgcctc taaaaatcta gatggacaat tttgactttg agaaaagaga 13320

ggtggaaatg aggaaaatga cttttattag attccagtag aaagaacttt catctttccc 13380

tcatttttgt tcgttttaaa acatctatct ggaggcagga caagtatggt cgttaaaaag 13440

atgcaggcag aaggcatata ttggctcagt caaagtgggg aactttggtg gccaaacata 13500

cattgctaag gctattccta tatcagctgg acacatataa aatgctgcta atgcttcatt 13560

acaaacttat atcctttaat tccagatggg ggcaaagtat gtccaggggt gaggaacaat 13620

tgaaacattt gggctggagt agattttgaa agtcagctct gtgtgtgtgt gtgtgtgcgc 13680

gcgcgcgtgt gtgtgtgtgt gtgtcaacgt gtgtttcttt taacgtcttc agcctacaac 13740

atacagggtt catggtggca agaagatagc aagatttaaa ttatggccag tgactagtgc 13800

ttgaaggggа acaactacct gcatttaatg ggaaggcaaa atctcaggct ttgagggaag 13860

ttaacatagg cttgattctg ggtagaagct gggtgtgtag ttatctggag gccaggctgg 13920

agctctcagc tcactatggg ttcatcttta ttgtctcctt tcatctcaac agctcctggg 13980

aaatgtgctg gtgaccgttt tggcaatcca tttcggcaaa gaattcaccc ctgaggtgca 14040

ggcttcctgg cagaagatgg tgactgcagt ggccagtgcc ctgtcctcca gataccactg 14100

agcctcttgc ccatgattca gagctttcaa ggataggctt tattctgcaa gcaattcaaa 14160

taataaatct attctgctga gagatcacac atgattttct tcagctcttt tttttacatc 14220

tttttaaata tatgagccac aaagggttta tattgaggga agtgtgtatg tgtatttctg 14280

catgcctgtt tgtgtttgtg gtgtgtgcat gctcctcatt tatttttata tgagatgtgc 14340

attttgttga gcaaataaaa gcagtaaaga cacttgtaca cgggagttct gcaagtggga 14400

gtaaatggtg taggagaaat ccggtgggaa gaaagacctc tataggacag gacttctcag 14460

aaacagatgt tttggaagag atgggaaaag gttcagtgaa gacctggggg ctggattgat 14520

tgcagctgag tagcaaggat ggttcttaat gaagggaaag tgttccaagc tttaggaatt 14580

caaggtttag tcaggtgtag caattctatt ttattaggag gaatactatt tctaatggca 14640

cttagctttt cacagccctt gtggatgcct aagaaagtga aattaatccc atgccctcaa 14700

gtgtgcagat tggtcacagc atttcaaggg agagacctca ttgtaagact ctgggggagg 14760

tggggactta ggtgtaagaa atgaatcagc agaggctcac aagtcagcat gagcatgtta 14820

tgtctgagaa acagaccagc actgtgagat caaaatgtag tgggaagaat ttgtacaaca 14880

ttaattggaa ggcttactta atggaatttt tgtatagttg gatgttagtg catctctata 14940

agtaagagtt taatatgatg gtgttacgga cctaatgttt gtgtctcctc aaaattcaca 15000

tgctgaatcc ccaactccca actgacctta tctgtggggg aggcttttga aaagtaatta 15060

ggtttagatg agctcataag agcagatccc catcataaaa ttattttcct tatcagaagc 15120

agagagacaa gccatttctc tttcctcccg gtgaggacac agtgagaagt ccgccatctg 15180

caatccagga agagaaccct gaccacgagt cagccttcag aaatgtgaga aaaaactctg 15240

ttgttgaagc cacccagtct tttgtatttt gttatagcac cttgcactga gtaaggcaga 15300
tgaagaagga gaaaaaaata agcttatcga aacgcgtccc catcctcact gactccgtcc 15360
tggagttgga tgagagataa tggccttacg ttgtgccagg ggagggtcgg gctggattta 15420
gcaagattta ccttctccaa agagcggtgc tgcagtggca cagctgccca cggaggtggg 15480
ggggtcaccg tccctggagg tgatgaagaa ctgtggggat gtggcactga gggacatggc 15540
cagtgggcac ggtgggtggg ttggggttgg tcttggggat cttggagggc ttttccagcc 15600
ttcatgattt gacgattgta tgaacatcta catggcaatt ctccagctgc ctgtcccagt 15660

cctactgacc cagctgtatc tctccaggca agctcttcca ccccttctgc ttgcatccag 15720
acaccatcaa acatgcaggc tcagacacag ggaccagcag tgtctgtggc ctttttgtgc 15780
tcctctccat gctgggtttt aacttgctct ttgtccttct atcctatctt cttatcctta 15840
aggctgttct gaacgctgtg acttggagag tgtcccagag ccctcaacac ctgcatgtcc 15900
cacgtccatg ctgtcctgca cttccttatc cccaagatct gcctctccgt gatgcactga 15960
attggcaaac atgtgtcacc ccagaccaac aatgtcacag caaactcccc cttgatagga 16020
caaggggaa tggctttaca ctgagacagg ggaggtttgg gttggatatg aggaggcagt 16080

ttttccccca gagggtggtg acgcactgaa caggttgccc aaggaggctg tggatgcccc 16140
atccctgcag gcattcaagg ccaggctgga tgtggctctg ggcagcctgg gctgctggtt 16200
gatgaccctg cacatagcag ggggttggat ctggatgagc actgtgctcc tttgcaaccc 16260
aggccgttct atgattctgt cattctaaat ctctctttca gcctaaagct ttttccccgt 16320
atccccccag gtgtctgcag gctcaaagag cagcgagaag cgttcagagg aaagcgatcc 16380
cgtgccacct tccccgtgcc cgggctgtcc ccgcacgctg ccggctcggg gatgcggggg 16440
gagcgccgga ccggagcgga gccccgggcg gctcgctgct gccccctagc gggggaggga 16500

cgtaattaca tccctggggg ctttgggggg gggctgtccc tatcgattac tagttttacc 16560
acatttgtag aggttttact tgctttaaaa aacctcccac acctcccccc gaacctgaaa 16620
cataaaatga atgcaattgt tgttgttaac ttgtttattg cagcttataa tggttacaaa 16680
taaagcaata gcatcacaaa tttcacaaat aaagcatttt tttcactgca ttctagttgt 16740
ggtttgtcca aactcatcaa tgtatcttat catgtctgct cgaagagcgg ccgcttcagt 16800
ttcggccagc aggcggggag cccgaggtag ctcccgctcc cttgagccag gcccctgcca 16860
gacctgagct ccctcccaag cctggcttcc ccaaccggtg gccttcatgg gccagaagcc 16920

attccttcac ggctagccct ccggagtagt tgcccacggc tccgctgctg cagaccactc 16980

tgtggcacgg gatgaggatc ttgacaggat tgcctctcat ggcgcctccc actgctcgcg 17040

cggctttggg gttgccggcc agggcggcca attgctggta agaaatcact tctccgaatt 17100

tcacaacctt aagcagcttc cataacacct gacgcgtgaa cgactcttgc tggaaaacgg 17160

gatggtgaag cgctggcacg gggaactctt cgatagcctc gggctggtgg aaataggcat 17220

tcagccaggc tgtgcactgc atcaggggct ccggacctcc gagaaccgca gcgggggctg 17280

ggacctccac ggcatcagct gcagacgtcc ccttgcccag gagctttatt tcgtgcagac 17340

cctgctcaca accagacagc tccagcttcc ccaaagggct gtccagtgtg gtgcgtttca 17400

tttcacaatc cttgtccatg gtggcgaccg tctagctcac gacacctgaa atggaagaaa 17460

aaaactttga accactgtct gaggcttgag aatgaaccaa gatccaaact caaaaagggc 17520

aaattccaag gagaattaca tcaagtgcca agctggccta acttcagtct ccacccactc 17580

agtgtgggga aactccatcg cataaaaccc ctccccccaa cctaaagacg acgtactcca 17640

aaagctcgag aactaatcga ggtgcctgga cggcgcccgg tactccgtgg agtcacatga 17700

agcgacggct gaggacggaa aggccctttt cctttgtgtg ggtgactcac ccgcccgctc 17760

tcccgagcgc cgcgtcctcc attttgagct ccctgcagca gggccgggaa gcggccatct 17820

ttccgctcac gcaactggtg ccgaccgggc cagccttgcc gcccagggcg gggcgataca 17880

cggcggcgcg aggccaggca ccagagcagg ccggcgagct tgagactacc cccgtccgat 17940

tctcggtggc cgcgctcgca ggccccgcct cgccgaacat gtgcgctggg acgcacgggc 18000

cccgtcgccg cccgcggccc caaaaaccga aataccagtg tgcagatctt ggcccgcatt 18060

tacaagacta tcttgccaga aaaaaagcgt cgcagcaggt catcaaaaat tttaaatggc 18120

tagagactta tcgaaagcag cgagacaggc gcgaaggtgc caccagattc gcacgcggcg 18180

gccccagcgc ccaggccagg cctcaactca agcacgaggc gaaggggctc cttaagcgca 18240

aggcctcgaa ctctcccacc cacttccaac ccgaagctcg ggatcaagaa tcacgtactg 18300

cagccagggg cgtggaagta attcaaggca cgcaagggcc ataacccgta aagaggccag 18360

gcccgcggga accacacacg gcacttacct gtgttctggc ggcaaacccg ttgcgaaaaa 18420

gaacgttcac ggcgactact gcacttatat acggttctcc cccaccctcg ggaaaaaggc 18480

ggagccagta cacgacatca ctttcccagt ttaccccgcg ccaccttctc taggcaccgg 18540

ttcaattgcc gacccctccc cccaacttct cggggactgt gggcgatgtg cgctctgccc 18600

actgacgggc accggagcct cacgcatgct cttctccacc tcagtgatga cgagagcggg 18660

cgggtgaggg ggcgggaacg cagcgatctc tgggttctac gttagtggga gtttaacgac 18720

ggtccctggg attccccaag gcaggggcga gtccttttgt atgaattact caaatcgata 18780

actagtgata tcctagtggc cccactgtgg ggtggagggg acagataaaa gtacccagaa 18840

ccagagccac attaaccggc cctgggaata taaggtggtc ccagctcggg gacacaggat 18900

ccctggaggc agcaaacatg ctgtcctgaa gtggacatag gggcccgggt tggaggaaga 18960

agactagctg agctctcgga cccctggaag atgccatgac agggggctgg aagagctagc 19020

acagactaga gaggtaaggg gggtagggga gctgcccaaa tgaaaggagt gagaggtgac 19080

ccgaatccac aggagaacgg ggtgtccagg caaagaaagc aagaggatgg agaggtggct 19140

aaagccaggg agacggggta ctttggggtt gtccagaaaa acggtgatga tgcaggccta 19200

caagaagggg aggcgggacg caagggagac atccgtcgga gaaggccatc ctaagaaacg 19260

agagatggca caggccccag aaggagaagg aaaagggaac ccagcgagtg aagacggcat 19320

ggggttgggt gagggaggag agatgcccgg agaggaccca gacacgggga ggatccgctc 19380

agaggacatc acgtggtgca gcgccgagaa ggaagtgctc cggaaagagc atccttgggc 19440

agcaacacag cagagagcaa ggggaagagg gagtggagga agacggaacc tgaaggaggc 19500

ggcagggaag gatctgggcc agccgtagag gtgacccagg ccacaagctg cagacagaaa 19560

gcggcacagg cccaggggag agaatgcagg tcagagaaag caggacctgc ctgggaaggg 19620

gaaacagtgg gccagaggcg gcgcagaagc cagtagagct caaagtggtc cggactcagg 19680

agagagacgg cagcgttaga gggcagagtt ccggcggcac agcaagggca ctcgggggcg 19740

agaggagggc agcgcaaagt gacaatggcc agggccaggc agatagacca gactgagcta 19800

tgggagctgg ctcaggttca ggagagggca gggcagggaa ggagacaaag tccaggaccg 19860

gctggagggg ctcaacatcg gaagagggga agtcgaggga gggatggtaa ggaggactgc 19920

atgggtcagc acaggctgcc aaagccaggg ccagttaaag cgactccaat gcggaagaga 19980

gtaggtcgaa ggggaatggt aaggaggcct ggggcagagt ggtcagcaca gagtggctaa 20040

gcccagggcc agttgaagcg gctccaattc ggaagtgggg tggtcgaagg ggaatggtaa 20100

gggggactgg gacggggtgt cagcataggg tggcaaagcc cagggccagg aacgacgggg 20160

cggatcgaga ctggcaacgg ggaaggagga tgccccaggt ggcgcagcag agggtggacc 20220

tggccccggg agacgccggg cgggggcgc tgacctggtg cagggcgctg ataccgtcgg 20280

cgttggtgga gtccagcacg gcgcgggcgg gcggcggcgc ggcggggtcg agctcggcgc 20340

cggggccagg gtcggcggcg cgcagcatca gacgcgcctc gtccaggtcg ccgcccgcac 20400

aggccgccag gaactcggcg gcgcgctcga agcggacggt gcgggcgcgg cgctctccgg 20460

ggccaggctc ggcgcccgcc cgcgcccccc actgccgcag ctgctcccgt cgccgctccc 20520

gggcagccgc cgccgccgcc cccgggccag ccgccgggcc atcctctccg gacatcgcac 20580

cgcccgcccg cccagcgagc gagcgagcgc cgagccccaa ccgccgccac caccccaatc 20640

ctgtccctag tggccccgat atcattaagt ggcaggatta atcttcgaac gaaggagcca 20700

tccaactaac cgtcatgttc gggcaaccga agaagggagt ggcaggattt cctttggaga 20760
cttctggaat tagacagcag tttaatgcaa gcatctaaat tctcttcctc ccagagtctc 20820
attaaaacta cagtaagagt ttgtgttttg ttttgttttt aaagacaaaa tcccaccagg 20880
atagagagaa taggagagga gataacagca tcataattta tgaaactaaa atgcagatag 20940
accaggatta actgactaca cagcaccaag gaagctgaat cacaagacag cagaggagaa 21000
aactggaaag gatcgtggtc tatacggcag aatcttccca agcctcagga ggaggagctc 21060
tagatgttcc cagatctggg aggtaaagtg gaatgggggg acatggtcag cgtaatgggg 21120

ttgggctgga agctggttaa ggagcaggca gatctctgaa tcccctctct gactctgtgt 21180
ccccaggcat ctgcctgtcc cccaccctgg aagaggtctg gcttgaccct ttgtctggtg 21240
aatttcctgc tctgctttcc tggtcctgct ggccagatca gtggaggcca ctcacttcac 21300
cccacaggga tgttctgtgt tgccctacac ctgggaactg gaggtactgg aggcaggctg 21360
tggtgagctt gaaagcaaaa cacagagggc agtccaatct ctttggccat atttcttctg 21420
catatccaat accatgtcca caactctgct agtgtcctga tggtggtggg ctctacacat 21480
tcccgggaag ctgaaggcag ataatgacca ggacaggtca acctctcttc ttctgaaagc 21540

cttcatctac taatggcctg ggactcttcc cttaaatgct tagattgtgt cttccactaa 21600
ggtttttttgc tgttgctgtt gtttgtttgt ttgtttgttt gtttgtttgt tttgagacgg 21660
aatctcactc tgtcgcccag gctggagtgt agtggcacaa tctcagctca ccacaacctt 21720
cacctcctag gttgaaggga ttctcctgcc tcagcctcct gagtagctag gattacaggc 21780
acatgccacc atgcctggct aatttttgta tttttagtag agacaggatt tcgccatgtt 21840
ggccaggctg gtcttgaact cctgacctca ggtgatctgc ctaccttggt ctcccaaagt 21900
gctgggatta ctggtgtgag ccaccacacc cggccaaggt ttttgtttgt ttgtttgttt 21960

gtttgttttg tattgaggca gggtatcact ctggtcaccc aggctggagt gcagtagtgc 22020
aatcacggct cactgaaacc tccacctccc tggcgggctc aggtgatcct gccacctcag 22080
cttcccaggt agctgggact acaggcttgt accaccactc ccagctaatt tttgcgtttt 22140
tagtagagac agggtttccc catgttgccc aggttggtct caaactctgg gctcaagcga 22200
tctgcctgcc tcagcctccc aaagtgctgg gattacaggt gtaagccacc gtacccggcc 22260
ccgccactaa ggttttgaaa atgaagcaat tacaagttta agtctattaa taagtgatga 22320
agctatgtag aaaagcagaa taattatctt ggatcaggaa ggtcacatga ggatctactt 22380

ggggggttgtc aatattctat ttcttgacct gatcagtgtt gacagcaggt tttaattttt 22440

tacttctttt tgtttgtttg tttttgagac ggagtcttgc tctgtctccc aggctggagt 22500

gcagtggtat gatctcggct cactgcaacc tccgcctcct gggttcaagc tgttctcctg 22560

cctcagcctc cccagtagct gggattacag gcaggcacca ccacgaccag ctaatttttg 22620

tatttttagt agagactggg tttcaccatc ttggccaggc tggtctcgaa cttctgatct 22680

cgtgatccgc cctccttggc ctcccaaagt gctgggatta caggcttgag ccagcgtgcc 22740

cggcccattt tttacttcct tattaaactg tacatatagg ccttgcacac ttttctgcat 22800

caatgttata ttccacaata aagggaaaag gtatatacac aacttgatac cagtaatgtg 22860

aaacatatat ttctacatag aaaaaaaaat gactgaaata ctgcactcca atgtgttcac 22920

acagtagttg tttctggatt atttatatat taaatgttta tatattgtat tatgccatga 22980

ggtttgtgtt ttctctccac ttttctgcat tttccaagtt tactacaaag agcacatatt 23040

actcttataa tcagaaagtc ataaaatata tttaaaaaga caaaattgaa actaataagg 23100

atcaacacaa aacagatgag ccatctgtgg aaatccgcac agaatactac ctaaagagat 23160

tggtgacgtg catgatctca ctaggatgag cacaaagctt gccagagcct agggtctatt 23220

tctagggttg gctcttggaa gccaggatag ttgttatctc tgggaagagg gaggggcaca 23280

caaggggctt ctaaaacatt ctgaatgttc tatttctgaa cctggttggt gggtacatga 23340

ctgttggttt tattattata tgttttatat actcttccgt atgtatggtg tggattccaa 23400

aaaaagattt cctttagaga aaaccagaat cacataagta gaaaatatgg tgctatgttg 23460

aaggaacaac tcaagtttat ataaaatcat catcatttat aggcttaaaa agttgctttg 23520

gaattttggt ctaactgact tgtcttttct gcagcaaacc acgctccttc tggacgtgct 23580

ccaggcagag gggattaggg tgggttcaag gctgcaagta cctagctcag cacactctct 23640

tcaggggact tagagtttgt ctggtgttgg ctctctgagc tcttgtcagg aatgccgacc 23700

cttccgaggt tcaggatttg aagcctgcct tcccacccca gatttggtcc acacagacac 23760

tcaagtatgt atttcaacta caaatgacct gtactttcct attactcctc tctttcatgg 23820

taacctttct ggtatccttc cttccctaca tttatgggag ggggacatca ttctctgctc 23880

tcctgtcact gaaggctcca ccttctgtct tcttctgacc catctggttt tcctggggcc 23940

acctcctctc cttaccaccc taacgctttt gtaacttgag gagaaatgag agatcaccta 24000

gtcaggtcat cattctctgt agatgaagag gcccaatggt ttgctcaaga attgccaagc 24060

gagttaaaga cagagagtat gagagtcagc aagacctaca gaaagcatct atctgcactg 24120

ttttgcaggg acttagcctt tgtgtgtgga ctcctggaat gccacccact aagaaacatt 24180

gtctgacacc aactccccac ttggtaggtg gggacactga aactcatggc aggaaagggc 24240

cttgccccaa gccagggcag agtgtcactc atcactctca attttcagtc cagggcacct 24300

tgttgtgact atcccaaagg cagccacttt ccctggtctg aaagacctga agagagaaga 24360

gaagagaagg atggaaggca gagtatgcgg ctttgattca tttcctggtg aaaacagatc 24420

tatacgagaa gcaaatttca cgaaagggaa gagaagaaag tgtcccatac gttgctggcc 24480

tgtttcaacc ttgctttgat tcttgctgaa aagggtaccg tgtatttctg agttcaacat 24540

gcagaccagt gttaggaaag ccactgcacc tccactttag cctccagggc tgtgccctgc 24600

aaatggcctg cagccttggt gcctcgctct ccagactgca ttttggaaga tgggacagag 24660

gcttatggaa gcccacatta gaacgggggа gcagaatggg tgagatgagg gatccttgat 24720

agtgaaccag atgaaggaat ggtagccaaa tgccaggcct cctttgtggc ttcaatccaa 24780

aggctctgga gcccttccag ggcagaacat caggcatgtt taccccccact gtcctcaaca 24840

gtgacagagg tgcaatcttg ggcagctggc cattttgaaa gcaacctcct taatctcaac 24900

tgggaaggct ccctagcagg acccctgtgt tgcacacctg gaggaagcta gactaaccag 24960

aagctcagca cggttccatc tgggatgccc aggtctgaga cgaaaaaggt aactctcttt 25020

tctgggtcct ggcccagttg tgtctctctc cacctcattc tctgagatgc ctgtctcccc 25080

ttttttgtcc catcaggagg caagagctat cactgggcca gactccacca gaagccaagc 25140

cagcttgtta cccagcttct cagggagcaa agaacagcct tgtttctatc ttatccccac 25200

tgtcccctgc ccctgcccca cctcccagcc attcagcttc tggcttcccc agagctgcct 25260

gcttctttgt ggtcctccat tccttgaaaa gaccttctag tcattagtgt atataaatgg 25320

ccacttagcc cagattacag tgaggtcaac agctggggct ctgagaattg tcacacactg 25380

gcacaggaga ggaggctatt cttccagaga atttggaggg cactcccatc cacttacaac 25440

aaaaagccca tccactgtgc ttggcagtag gtgatctgag aaccaatgga accaggttaa 25500

tcctgtggca ctgttgagtg aggagagcag tggcgggcac tggaaaatat cagagacaag 25560

gcaggagacc tgaaatctag gcttagctcc tcatatactt ggcagctgta tgacctcaga 25620

caaccagtgt tacctctcta agcctcagtt tcctcatgca aaaggagggg gaataacaac 25680

agagcccact gcttgggggt gttgtgagga caggatgaaa aaacaaacag aaatccctca 25740

gtacaggatt cagtgcagtg gacagtcttg caaggtctgg ttcagccctc cacccctacc 25800

ctcaccagta taaagaactc tggcctacaa gtcagatgac ctgagtttta atctcagctt 25860

tgccattagc cgtgtgaact tgagaaagtc cctttccttt ttacatctat tgggatgatc 25920

atgcattttt tgtcctttat tctgttaata tagtgtgtta cattgattgc ttttcataga 25980

ctgaaccagc cttgtattcc agggataaat ctcacttggt catggtgtat aatcctttat 26040

acaaatgttg ctgggttgag tttgctagta ttttgttgaa gatttttatg tcttgattca 26100

taaggaatat tggtgtacct tccccttta tggccacagt ttccctacaa tgatgtagtc 26160

gaactagaca acctccaata tctttcagta ttcatgtcct ctgattctgt gaaactaaga 26220
aaattaagaa atagtgattc ataggcacaa ggcaggcaaa acttagactc cttgtagaat 26280
aattaggaag ccaaatattc agtgtgctta tttctcaaat aaccttagtt tctccagtct 26340
gccccaactc cgaggcctga atatctctag atgcttatga tggcaactaa agcctaaaag 26400
ctaattcatt ttaaagttct tccaaatgca tagggtttta tttttccaga cctgggttca 26460
gatggggaat ttgacaaaca atggaaaggg ggaaaaacaa caatctaaac actgagtgac 26520
aaagtaacaa agaaatagtc tagctatcag ccagtcaagc cagccttggc tttgctatcc 26580

aaagtagtca gtctaattct accaccagtt tctgttcctg tagctgtcta ctgcctgcca 26640
gggactctgc cttcccaccc acaactacca atggaaggat gtggtgacca taccagtggc 26700
tgctgacatc tcctgccatg ggaagcataa ttgcctccag cagcctcccc cttagatcca 26760
tcatttttgt tgcacttggc ctgggctgta ctcccggcca atgactgaac atggtgagca 26820
tagtaatgca ggcccatttc tgtgaggagc aggactcctc cagtaggtga ctttggctca 26880
aggactctct attggcctgg ttgaactttt cctgaactgt gctactgtct gagactcttc 26940
ttacccaatc ctctttctcg ccccaattgt cacagaccac ctgcattgtg gtctgagtct 27000

ctccccacct tctcttgctc ttccctgttt atctttcaca ggcatttccc ccagtacatt 27060
ccttgaatgt ctaacccgat acgggtgcct gacttttggc agacctaagc agacaaaaag 27120
gagtacttgg ttacctagct cttctttcta ccacaaacat cgagggaacc ctttttccct 27180
caccctctg ccacaccccc actgccccag tgaacaacca cagagagagc tgtggtataa 27240
tattaggctg gtgcaaaagt aattgcggtt tttgccatta cttttaatgg taaaaaccgc 27300
aattactttt gcacctacct agtatttgtg tccccccaaa ttcatatgtt gaaacctaac 27360
ccacaatatg atgtcattag gaggcaagac cttgaggagg tgattagatg atggggtgga 27420

gctctcctga atgagattag tgcccttata agaagaagcc caaggaagct accttgactc 27480
ttccatcaca tgagaatgca gcaagaaggc accatctact aatcaggaag agagctctca 27540
ccagacactg aatctgccag tgtcttgatc ttgaagttcc cagcctccag aactatgcat 27600
aatgcatttc cattgtctct aagccaccca gcctatggta ttttgtcata gcagcctgaa 27660
ctgactaaga cagtgagcca catgagaagt gccccaaccc ctcccttaag cacttggctc 27720
acagatcagt gggttcattt ctgcctgagt tttattgtta ttctgtagat ttcttgggct 27780
agatatattt ttctgttatt ttccttcttc acctcagtca tgaattggtt gttttaaaaa 27840

agacaatgta agtcatgggg aaactcctga caactctact ctcctagggt tcctgataaa 27900

aggggattca gttgagtcct ctgatggtct ctacctgcca aagtccagca gcccttagca 27960

aacatgctgc tcgtttctgt agagaaggtg ctggtgtccc accatacttc tctctccctc 28020

atgaagggct tgcgacccag caaatgggtg gcttatatgg gtctgtttca aaggaagagc 28080

cagctctggg aagaaaaacg atgagcataa gcataaccta ccactgtgcc tgggaaagca 28140

gacaactttt ttgatgtgtg aatatctaat gagaatggaa tccatcaatt accttaaact 28200

taggcacagt cttcaaattc aatatatgtg ggatatactt ttagtcagtt tgtagacgtt 28260

atttgtaata aataatctgg cttctctaaa gaaattattt taagtgtttg gtttggtttg 28320

atttaatggt aaaattatat ttagtggcag agaattatag caatggtgat aaactataga 28380

gtgtcataag ttcatatctt attctcacat ttgaagctgc ctgcagatgc attcaagatg 28440

cagccagaag tcaggagact caggctgtta tttggagctc atcattttac agccttgctg 28500

gactcccact ttctcagggg aaaaatgtgg tgttgaccca gattagctct ccaggccctg 28560

ctgagttggg cactctgtaa gctggagggt cttctattgt cttcacctaa gtgtcaatca 28620

acaacccaaa tgggcatggg ggaagaggga gctgggccaa tgcccagggt gcctggtaga 28680

gagatacctt gggcactgga aggcaccagc ttcccagaga gaaggggggag ggccatgaaa 28740

aagttggctg tagatgccag ggacactggg actctccagc tgtgtgtttg tgtcttctga 28800

agacttatgt ttcattcctt tggagcatgc ataatcatac actgtgggat gtgttatata 28860

gattgcttga tagttcacca ctgtaataaa atactgtgac tggaatctgc tcccagtctg 28920

cctttgatag cacttgtgca acacacattt actgagcatt tacagtgatc caggacctgt 28980

gttgtgaaaa cattgatgga caaggcagat ggtggagcac gtcagtgagg atttttaaca 29040

aaggctggta agtgctataa aggaacattg taggacacta gagaacaaag aacaggagaa 29100

cctgacttag gctggggtgg ggcgttggtt agaggaggct ccttggagga catgaggttt 29160

aagctgtgac ctgaggatga atagatgttg gccaggtgag gttaatctgg ggctcaagat 29220

cgagcattaa gcttgtcagc cttaccagta aaaaagaaaa cctattaaaa aaacaccact 29280

cgacacggca ccagctcaat cagtcacagt gtaaaaaagg gccaagtgca gagcgagtat 29340

atataggact aaaaaatgac gtaacggtta aagtccacaa aaaacaccca gaaaaccgca 29400

cgcgaaccta cgcccagaaa cgaaagccaa aaaacccaca acttcctcaa atcgtcactt 29460

ccgttttccc acgttacgta acttcccatt ttaagaaaac tacaattccc aacacataca 29520

agttactccg ccctaaaacc tacgtcaccc gccccgttcc cacgccccgc gccacgtcac 29580

aaactccacc ccctcattat catattggct tcaatccaaa ataaggtata ttattgatga 29640

tgtttaaact acggcccggt acccagcttt tgttcccttt agtgagggtt aattgcgcgc 29700

ttggcgtaat catggtcata gctgtttcct gtgtgaaatt gttatccgct cacaattcca 29760

cacaacatac gagccggaag cataaagtgt aaagcctggg gtgcctaatg agtgagctaa 29820

ctcacattaa ttgcgttgcg ctcactgccc gctttccagt cgggaaacct gtcgtgccag 29880

ctgcattaat gaatcggcca acgcgcgggg agaggcggtt tgcgtattgg gcgctcttcc 29940

gcttcctcgc tcactgactc gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct 30000

cactcaaagg cggtaatacg gttatccaca gaatcagggg ataacgcagg aaagaacatg 30060

tgagcaaaag gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc 30120

cataggctcc gccccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga 30180

aacccgacag gactataaag ataccaggcg tttcccccctg gaagctccct cgtgcgctct 30240

cctgttccga ccctgccgct taccggatac ctgtccgcct ttctcccttc gggaagcgtg 30300

gcgctttctc atagctcacg ctgtaggtat ctcagttcgg tgtaggtcgt tcgctccaag 30360

ctgggctgtg tgcacgaacc ccccgttcag cccgaccgct gcgccttatc cggtaactat 30420

cgtcttgagt ccaacccggt aagacacgac ttatcgccac tggcagcagc cactggtaac 30480

aggattagca gagcgaggta tgtaggcggt gctacagagt tcttgaagtg gtggcctaac 30540

tacggctaca ctagaaggac agtatttggt atctgcgctc tgctgaagcc agttaccttc 30600

ggaaaaagag ttggtagctc ttgatccggc aaacaaacca ccgctggtag cggtggtttt 30660

tttgtttgca agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc 30720

ttttctacgg ggtctgacgc tcagtggaac gaaaactcac gttaagggat tttggtcatg 30780

agattatcaa aaaggatctt cacctagatc cttttaaatt aaaaatgaag ttttaaatca 30840

atctaaagta tatatgagta aacttggtct gacagttacc aatgcttaat cagtgaggca 30900

cctatctcag cgatctgtct atttcgttca tccatagttg cctgactccc cgtcgtgtag 30960

ataactacga tacgggaggg cttaccatct ggccccagtg ctgcaatgat accgcgagac 31020

ccacgctcac cggctccaga tttatcagca ataaaccagc cagccggaag ggccgagcgc 31080

agaagtggtc ctgcaacttt atccgcctcc atccagtcta ttaattgttg ccgggaagct 31140

agagtaagta gttcgccagt taatagtttg cgcaacgttg ttgccattgc tacaggcatc 31200

gtggtgtcac gctcgtcgtt tggtatggct tcattcagct ccggttccca acgatcaagg 31260

cgagttacat gatcccccat gttgtgcaaa aaagcggtta gctccttcgg tcctccgatc 31320

gttgtcagaa gtaagttggc cgcagtgtta tcactcatgg ttatggcagc actgcataat 31380

tctcttactg tcatgccatc cgtaagatgc ttttctgtga ctggtgagta ctcaaccaag 31440

tcattctgag aatagtgtat gcggcgaccg agttgctctt gcccggcgtc aatacgggat 31500

aataccgcgc cacatagcag aactttaaaa gtgctcatca ttggaaaacg ttcttcgggg 31560

cgaaaactct caaggatctt accgctgttg agatccagtt cgatgtaacc cactcgtgca 31620

cccaactgat cttcagcatc ttttactttc accagcgttt ctgggtgagc aaaaacagga 31680
aggcaaaatg ccgcaaaaaa gggaataagg gcgacacgga aatgttgaat actcatactc 31740
ttccttttttc aatattattg aagcatttat cagggttatt gtctcatgag cggatacata 31800
tttgaatgta tttagaaaaa taaacaaata ggggttccgc gcacatttcc ccgaaaagtg 31860
ccacctgacg cgccctgtag cggcgcatta agcgcggcgg gtgtggtggt tacgcgcagc 31920
gtgaccgcta cacttgccag cgccctagcg cccgctcctt tcgctttctt cccttccttt 31980
ctcgccacgt tcgccggctt tccccgtcaa gctctaaatc gggggctccc tttagggttc 32040

cgatttagtg ctttacggca cctcgacccc aaaaaacttg attagggtga tggttcacgt 32100
agtgggccat cgccctgata gacggttttt cgcccctttga cgttggagtc cacgttcttt 32160
aatagtggac tcttgttcca aactggaaca acactcaacc ctatctcggt ctattctttt 32220
gatttataag ggattttgcc gatttcggcc tattggttaa aaaatgagct gatttaacaa 32280
aaatttaacg cgaattttaa caaaatatta acgcttacaa tttccattcg ccattcaggc 32340
tgcgcaactg ttgggaaggg cgatcggtgc gggcctcttc gctattacgc cagctggcga 32400
aaggggatg tgctgcaagg cgattaagtt gggtaacgcc agggttttcc cagtcacgac 32460

gttgtaaaac gacggccagt gagcgcgcgt aatacgactc actatagggc gaattggagc 32520
tccactacgt agtttaaa 32538
<210> 60
<211> 32699
<212> DNA
<213> Artificial Sequence
<220><223> pHCA-Dual2-FI-globin-mgmtGFP
<400> 60
ggccgaagga ttacatgagc ttagaaatgt aattagcata gtgtgtggca tagtgtagat 60
accaaataaa tatgatctct ccttctactc ttgaaaatgc aaacacattc ttggtggtcc 120
taaaatagcc tgtaacatgg tttactcagc agcatttgct attcaaggca gatctgcctt 180
tagtcattgg ctgcgctcct gaacagctgt gtgaaaggct aacttttgta aaccaaatca 240

aaataaaatg cagcaaaaat ttgtcactga aaggaaatcc tcagtatatc cttttatgaa 300
atgaaagatc cctcatccaa acttaacttt tttaaaagtg cgcatttgga gatatagccc 360
tttcttatga atcctaattc aattttggcc ataaacacac gttgatgttc cccaccccaa 420
agcacatagc aacaagagta ggttctatat tgaaaataat gacaatttaa aaacatgtac 480

ttatttcact gtatgtggac agtgtctatg attgcatcat gaagtgtcat ataaccatgt 540

acgtgtacat gagagagaga tagagagaga agtggtaggg tggtggtggt agaggggatg 600

gcgatagtaa tcatggtaat ggtagaggtg atggaggtgg taatgacgga ggtaagggtg 660

gtagtgatga tggtggtggt ggtaatggtg gtggatgtgg tggtggcaat tgggatggtg 720

ggatggtggt agccatggtg atggtggtaa tggtgttgat ttaaagggtg gtggtagtga 780

aggtgagggt agtggtggtg gaggtggtgg tgctggtagc aatagtgatg gtggtgatgg 840

tgttgatgag ggtgttggga tcagggtgag ttcccacagt atatttcatt cttgttgtac 900

cactctgtca acagcaccac tgactgggac agaggaagaa ggcacactct gaatgtgtta 960

ttaacagaaa cctcaaaaca gtctgtctcc ttgtagtcat tcaaaattat ctttttctta 1020

cctggaaaac tgaaactgaa ttaccgggaa aaacacagga gatttttgtt tgttaatatg 1080

ctgccaataa agtaatttta tgtcaaattt aactacagga aagggcaagg cattttctaa 1140

gttccttaga tgtcatgtgg ctaaaaaaaa caaaaggatg gacagcagtt agatactgta 1200

cacttagctg tttgaagcca tatattcaga aagcagatgt tgggagttgg tgtttgagga 1260

ctgatttcct ggaggtattt tatataggcc aagttcattg ttctaaactc taagggcttg 1320

acttgaggga ggaaaagagg caagaacatg tttagttttg ctgacagcat cacatgggca 1380

gccctaaggc tagacaactt tagggcctga agcttattct aggaaagaag cacctacaga 1440

gtggcactgg gctcccctcc actatagaga tgaagtcata tgacagtaaa gggcaggcag 1500

ggctgcctag ggggcccaga actgacactt ccattagaat gagcacaggc cagggagaga 1560

agtggggaac cagagagaag gagctggaat tcctttctct ccatacataa atgcctgcag 1620

agtcccattt cagaatccgg cagacaaagc caccaatgtg atccccatga ccttataaac 1680

attcattaaa atgcatttca aggcatgtga tggcctcccc acccccctaga taatgagaaa 1740

acaaaggttt ctcttctgat agagacaagt tcagctctga agtcaacatt atttctggtt 1800

ctgtctgaac aatgacatat ggcaactctt ccctttctat agttctagtc cagaatgaca 1860

aaaaagggga aaaatttctt agagaaggta gagattatac gaatacagtc catgaaatga 1920

gcataaggag aataaagaat ataacttatc caaagaagtc tggcaggctg ttataaatgc 1980

ttgattttgg acactgtagt tggaggttta acatggacac caataaaaag gtcagcaaag 2040

ggtatgcact gttcctattg ggcaagaaga taggaggtca aaggtaacca ggaaagataa 2100

actcagggag acttattttc cctccagagg gcactgggct tgtaggccct gggcaaaatt 2160

gtcaaaaagg tgaaaatcgc ctgtggttta tttagtctgc tctttcttca ctagtgcctc 2220

accagttcag ttcaggccaa tttgctagct accacatttg tagaggtttt acttgcttta 2280

aaaaacctcc cacacctccc cctgaacctg aaacataaaa tgaatgcaat tgttgttgtt 2340

aacttgttta ttgcagctta taatggttac aaataaagca atagcatcac aaatttcaca 2400
aataaagcat ttttttcact gcattctagt tgtggtttgt ccaaactcat caatgtatct 2460
tatcatgtct gaagctttct aggtattgaa taagaaaaat gaagttaagg tggttgatgg 2520
taacactatg ctaataactg cagagccaga agcaccataa gggacatgat aagggagcca 2580
gcagacctct gatctcttcc tgaatgctaa tcttaaacat cctgaggaag aatgggactt 2640
ccatttgggg tgggcctatg atagggtaat aagacagtag tgaatatcaa gctacaaaaa 2700
gccccctttc aaattcttct cagtcctaac ttttcatact aagcccagtc cttccaaagc 2760

agactgtgaa agagtgatag ttccgggaga ctagcactgc agattccggg tcactgtgag 2820
tgggggaggc agggaagaag ggctcacagg acagtcaaac catgcccccc gtttttcctt 2880
cttcaagtag acctctataa gacaacagag acaactaagg ctgagtggcc aggcgaggag 2940
aaaccatctc gccgtaaaac atggaaggaa cacttcaggg gaaaggtggt atctctaagc 3000
aagagaactg agtggagtca aggctgagag atgcaggata agcaaatggg tagtgaaaag 3060
acattcatga ggacagctaa aacaataagt aatgtaaaat acagcatagc aaaactttaa 3120
cctccaaatc aagcctctac ttgaatcctt ttctgaggga tgaataaggc ataggcatca 3180

ggggctgttg ccaatgtgca ttagctgttt gcagcctcac cttctttcat ggagtttaag 3240
atatagtgta ttttcccaag gtttgaacta gctcttcatt tctttatgtt ttaaatgcac 3300
tgacctccca cattcccttt ttagtaaaat attcagaaat aatttaaata catcattgca 3360
atgaaaataa atgtttttta ttaggcagaa tccagatgct caaggccctt cataatatcc 3420
cccagtttag tagttggact tagggaacaa aggaaccttt aatagaaatt ggacagcaag 3480
aaagcgagct tagtgatact tgtgggccag ggcattagcc acaccagcca ccactttctg 3540
ataggcagcc tgcactggtg gggtggcggc cgccctagga ttatggcact ggtagaattc 3600

actacttatg gcactggtag aattcactac ttatggcact ggtagaattc actacttatg 3660
gcactggtag aattcactat cgttgtgctt gatctaacca tgtttcattg tgcttgatct 3720
aaccatgttt cattgtgctt gatctaacca tgtttcattg tgcttgatct aaccatgtat 3780
cgcccggggg cggccgcaca caaaaaacca acacacagat ctaatgaaaa taaagatctt 3840
ttattgaatt cttagctggc ctccaccttt ctcttcttct tggggctgtc gcctcccagc 3900
tgagacaggt cgatccgtgt ctcgtacagg ccggtgatgc tctggtggat cagggtggcg 3960
tccagcacct ctttggtgct ggtgtacctc ttccggtcga tggtggtgtc aaagtacttg 4020

aaggcggcag gggctcccag attggtcagg gtaaacaggt ggatgatatt ctcggcctgc 4080

tctctgatgg gcttatcccg gtgcttgttg taggcggaca gcactttgtc cagattagcg 4140

tcggccagga tcactctctt ggagaactcg ctgatctgct cgatgatctc gtccaggtag 4200

tgcttgtgct gttccacaaa cagctgtttc tgctcattat cctcgggggа gcccttcagc 4260

ttctcatagt ggctggccag gtacaggaag ttcacatatt tggagggcag ggccagttcg 4320

tttcccttct gcagttcgcc ggcagaggcc agcattctct tccggccgtt ttccagctcg 4380

aacagggagt acttaggcag cttgatgatc aggtcctttt tcacttcttt gtagcccttg 4440

gcttccagaa agtcgatggg attcttctcg aagctgcttc tttccatgat ggtgatcccc 4500

agcagctctt tcacactctt cagtttcttg gacttgccct tttccacttt ggccaccacc 4560

agcacagaat aggccacggt ggggctgtcg aagccgccgt acttcttagg gtcccagtcc 4620

ttctttctgg cgatcagctt atcgctgttc ctcttgggca ggatagactc tttgctgaag 4680

ccgcctgtct gcacctcggt ctttttcacg atattcactt ggggcatgct cagcactttc 4740

cgcacggtgg caaaatcccg gcccttatcc cacacgatct ccccggtttc gccgtttgtc 4800

tcgatcagag gccgcttccg gatctcgccg ttggccaggg taatctcggt cttgaaaaag 4860

ttcatgatgt tgctgtagaa gaagtacttg gcggtagcct tgccgatttc ctgctcgctc 4920

ttggcgatca tcttccgcac gtcgtacacc ttgtagtcgc cgtacacgaa ctcgctttcc 4980

agcttagggt acttttgat cagggcggtt cccacgacgg cgttcaggta ggcgtcgtgg 5040

gcgtggtggt agttgttgat ctcgcgcact ttgtaaaact ggaaatcctt ccggaaatcg 5100

gacaccagct tggacttcag ggtgatcact ttcacttccc ggatcagctt gtcattctcg 5160

tcgtacttag tgttcatccg ggagtccagg atctgtgcca cgtgctttgt gatctgccgg 5220

gtttccacca gctgtctctt gatgaagccg gccttatcca gttcgctcag gccgcctctc 5280

tcggccttgg tcagattgtc gaactttctc tgggtaatca gcttggcgtt cagcagctgc 5340

cgccagtagt tcttcatctt cttcacgacc tcttcggagg gcacgttgtc gctcttgccc 5400

cggttcttgt cgcttctggt cagcaccttg ttgtcgatgg agtcgtcctt cagaaagctc 5460

tgaggcacga tatggtccac atcgtagtcg gacagccggt tgatgtccag ttcctggtcc 5520

acgtacatat cccgcccatt ctgcaggtag tacaggtaca gcttctcgtt ctgcagctgg 5580

gtgttttcca cggggtgttc tttcaggatc tggctgccca gctctttgat gccctcttcg 5640

atccgcttca ttctctcgcg gctgttcttc tgtcccttct gggtggtctg gttctctctg 5700

gccatttcga tcacgatgtt ctcgggcttg tgccggccca tcactttcac gagctcgtcc 5760

accaccttca ctgtctgcag gatgcccttc ttaatggcgg ggctgccggc cagattggca 5820

atgtgctcgt gcaggctatc gccctggccg gacacctggg ctttctggat gtcctcttta 5880

aaggtcaggc tgtcgtcgtg gatcagctgc atgaagtttc tgttggcgaa gccgtcggac 5940

ttcaggaaat ccaggattgt cttgccggac tgcttgtccc ggatgccgtt gatcagcttc 6000

cggctcagcc tgccccagcc ggtgtatctc cgccgcttca gctgcttcat cactttgtcg 6060

tcgaacaggt gggcataggt tttcagccgt tcctcgatca tctctctgtc ctcaaacagt 6120

gtcagggtca gcacgatatc ttccagaatg tcctcgtttt cctcattgtc caggaagtcc 6180

ttgtccttga taattttcag cagatcgtgg tatgtgccca gggaggcgtt gaaccgatct 6240

tccacgccgg agatttccac ggagtcgaag cactcgattt tcttgaagta gtcctctttc 6300

agctgcttca cggtcacttt ccggttggtc ttgaacagca ggtccacgat ggccttttc 6360

tgctcgccgc tcaggaaggc gggctttctc attccctcgg tcacgtattt cactttggtc 6420

agctcgttat acacggtgaa gtactcgtac agcaggctgt gcttgggcag caccttctcg 6480

ttgggcaggt tcttatcgaa gttggtcatc cgctcgatga agctctgggc ggaagcgccc 6540

ttgtccacca cttcctcgaa gttccagggg gtgatggttt cctcgctctt tctggtcatc 6600

caggcgaatc tgctgtttcc cctggccaga gggcccacgt agtaggggat gcggaaggtc 6660

aggatcttct cgatcttttc ccggttgtcc ttcaggaatg ggtaaaaatc ttcctgccgc 6720

cgcagaatgg cgtgcagctc tcccaggtgg atctggtggg ggatgctgcc gttgtcgaag 6780

gtccgctgct tccgcagcag gtcctctctg ttcagcttca cgagcagttc ctcggtgccg 6840

tccatctttt ccaggatggg cttgatgaac ttgtagaact cttcctggct ggctccgccg 6900

tcaatgtagc cggcgtagcc gttcttgctc tggtcgaaga aaatctcttt gtacttctca 6960

ggcagctgct gccgcacgag agctttcagc agggtcaggt cctggtggtg ctcgtcgtat 7020

ctcttgatca tagaggcgct caggggggcc ttggtgatct cggtgttcac tctcaggatg 7080

tcgctcagca ggatggcgtc ggacaggttc ttggcggcca gaaacaggtc ggcgtactgg 7140

tcgccgatct gggccagcag gttgtccagg tcgtcgtcgt aggtgtcctt gctcagctgc 7200

agtttggcat cctcggccag gtcgaagttg ctcttgaagt tgggggtcag gcccaggctc 7260

agggcaatca ggttgccgaa caggccattc ttcttctcgc cgggcagctg ggcgatcaga 7320

ttttccagcc gtctgctctt gctcagtctg gcagacagga tggccttggc gtccacgccg 7380

ctggcgttga tggggttttc ctcgaacagc tggttgtagg tctgcaccag ctggatgaac 7440

agcttgtcca cgtcgctgtt gtcggggttc aggtcgccct cgatcaggaa gtggccccgg 7500

aacttgatca tgtgggccag ggccagatag atcagccgca ggtcggcctt gtcggtgctg 7560

tccaccagtt tctttctcag gtggtagatg gtggggtact tctcgtggta ggccacctcg 7620

tccacgatgt tgccgaagat ggggtgccgc tcgtgcttct tatcctcttc caccaggaag 7680

gactcttcca gtctgtggaa gaagctgtcg tccaccttgg ccatctcgtt gctgaagatc 7740

tcttgcagat agcagatccg gttcttccgt ctggtgtatc ttcttctggc ggttctcttc 7800

agccgggtgg cctcggctgt ttcgccgctg tcgaacagca gggctccgat caggttcttc 7860
ttgatgctgt gccggtcggt gttgcccagc accttgaatt tcttgctggg caccttgtac 7920
tcgtcggtga tcacggccca gcccacagag ttggtgccga tgtccaggcc gatgctgtac 7980
ttcttgtcgg acgcttcgac cttgcgcttt ttcttcggcg aagcgtaatc tggaacatcg 8040
tatgggtaca tggtggccct aggatcgatt actagctcac gacacctgaa atggaagaaa 8100
aaaactttga accactgtct gaggcttgag aatgaaccaa gatccaaact caaaaagggc 8160
aaattccaag gagaattaca tcaagtgcca agctggccta acttcagtct ccacccactc 8220

agtgtgggga aactccatcg cataaaaccc ctccccccaa cctaaagacg acgtactcca 8280
aaagctcgag aactaatcga ggtgcctgga cggcgcccgg tactccgtgg agtcacatga 8340
agcgacggct gaggacggaa aggccctttt cctttgtgtg ggtgactcac ccgcccgctc 8400
tcccgagcgc cgcgtcctcc attttgagct ccctgcagca gggccgggaa gcggccatct 8460
ttccgctcac gcaactggtg ccgaccgggc cagccttgcc gcccagggcg gggcgataca 8520
cggcggcgcg aggccaggca ccagagcagg ccggcgagct tgagactacc cccgtccgat 8580
tctcggtggc cgcgctcgca ggccccgcct cgccgaacat gtgcgctggg acgcacgggc 8640

cccgtcgccg cccgcggccc caaaaaccga aataccagtg tgcagatctt ggcccgcatt 8700
tacaagacta tcttgccaga aaaaaagcgt cgcagcaggt catcaaaaat tttaaatggc 8760
tagagactta tcgaaagcag cgagacaggc gcgaaggtgc caccagattc gcacgcggcg 8820
gccccagcgc ccaggccagg cctcaactca agcacgaggc gaaggggctc cttaagcgca 8880
aggcctcgaa ctctcccacc cacttccaac ccgaagctcg ggatcaagaa tcacgtactg 8940
cagccagggg cgtggaagta attcaaggca cgcaagggcc ataacccgta aagaggccag 9000
gcccgcggga accacacacg gcacttacct gtgttctggc ggcaaacccg ttgcgaaaaa 9060

gaacgttcac ggcgactact gcacttatat acggttctcc cccaccctcg ggaaaaaggc 9120
ggagccagta cacgacatca ctttcccagt ttaccccgcg ccaccttctc taggcaccgg 9180
ttcaattgcc gacccctccc cccaacttct cggggactgt gggcgatgtg cgctctgccc 9240
actgacgggc accggagcct cacgcatgct cttctccacc tcagtgatga cgagagcggg 9300
cgggtgaggg ggcgggaacg cagcgatctc tgggttctac gttagtggga gtttaacgac 9360
ggtccctggg attccccaag gcaggggcga gtccttttgt atgaattact caaatcgatt 9420
aggatccggc gcgcccaccg cggaaaaaaa gcaccgactc ggtgccactt tttcaagttg 9480

ataacggact agccttattt taacttgcta tttctagctc taaaacgtga taaaagcaac 9540

tgttagcggt gtttcgtcct ttccacaaga tatataaagc caagaaatcg aaatactttc 9600

aagttacgat aagcatatga tagtccattt taaaacataa ttttaaaact gcaaactacc 9660

caagaaatta ttactttcta cgtcacgtat tttgtactaa tatctttgtg tttacagtca 9720

aattaattcc aattatctct ctaacagcct tgtatcgtat atgcaaatat gaaggaatca 9780

tgggaaatag gccctcggcg cgcccaccgc ggaaaaaaag caccgactcg gtgccacttt 9840

ttcaagttga taacggacta gccttatttt aacttgctat ttctagctct aaaactgacc 9900

aatagccttg acaagcggtg tttcgtcctt tccacaagat atataaagcc aagaaatcga 9960

aatactttca agttacggta agcatatgat agtccatttt aaaacataat tttaaaactg 10020

caaactaccc aagaaattat tactttctac gtcacgtatt ttgtactaat atctttgtgt 10080

ttacagtcaa attaattcta attatctctc taacagcctt gtatcgtata tgcaaatatg 10140

aaggaatcat gggaaatagg ccctcggatc cttaattgct agcgaggcct gtactctgtt 10200

ttcacaggaa gaaatcctca cccagtcttc cccaaacaca ttcccaggtt ctgtcattag 10260

tgggatagag atgattactg tggggagaag agaaacatct ggatggattt ggtgaggttg 10320

atctatagag gaagtaggtg ctgcctgagg tagctgtaat agaagctaaa ggtcaaagga 10380

gagggccctg tcccaatcca gatgactcca cttctgctgg acccaggttc acaagcttaa 10440

tctacatttc acctaaattt ggctaacaag cccaaaatca cacaggcaaa gggagaagtg 10500

gaggcagaac cgaggttgga ggccaccagg gccaccgggc agagatcatt taagcccaac 10560

cttctcactt ctccctgggc tctgcctctc ttaaaggacc ttgtggtgtg acctcttgta 10620

ggtccctttc acactcgggg cctcagtttc cccactgtaa agtgaatggg tcccagcttt 10680

ggtaagctta tgcttacctg atgctttctt cctgggctgc tcttgtagag aaaagataaa 10740

tcttcttcct ccatccacga gggcttcttt ccctgggggt gagagtaggc tgaggagagc 10800

cacttgcaca cactcttaaa gaaagtatta cctgcaccag ctcagtgaga ggcacagatc 10860

agactgttac ttgaatcaaa ttatgagcct ccccaaatat atctatgaca tttaaatagg 10920

ggattacttg aacatagact ttgggatccg gtgtggagtg caggagacta gcaaagtgaa 10980

tcctgagagt agcaggtctg cacctgttgg atcgagaaag gcggcctaca attctggtca 11040

aatgagctgt gcttattgac atattctatt agagagtact accaggtcac cagtcaccag 11100

aaaggctgcc agctctccaa ccacctccag ggaactatcc tgaatggggc cttaacaagc 11160

ctaagagagg gttggtttgg gtcccaagcc aatatttgct ctgctttatg tcagtcatat 11220

ggaacccaaa ccaaccctct cctatgtgcc tcaccagtcg gtgcagggat cccaatttca 11280

agtttggttt tttatggtca aagtccagca tagattaaat gaaggggtgt gatgatggtg 11340

ttaaaagaga actccagacc agtttaactc ttggacacac atcccatctc accatggtgc 11400

ttccaacctt ccagagatga tgggctccta ttttctgatg acaaagccct ccacaggatt 11460

gctgcctggc catcagggag tgcctctgta actgaggctg agatcccact ttcagtcctc 11520

cagctgtggc ccatccctgc tccgcccacc gggtatggcc tgtcctaggc tcttaggtat 11580

ggctgcattg tgaaatgatg gctacagagc tggcatctcc tgtagtctgg ttcatctagt 11640

gcactacctc atagttaaaa gaaatctgtt taagccactg agggtggctc ctagtgccaa 11700

ctccaagaac aggaagcttc ccttttttgg gaggaggggc agatggtaac atggatcgtc 11760

caggtcaatg ggagcagggc aaccacagta agtactggac aacaacacaa aactccatgt 11820

gtggcttcca tcgagtccct ctccaattgg tttggtcttc tccgtcccat gcagcacttt 11880

agcaaggggc ctggctgaag gctatgaatt gtgtggagcc tcctcattgc agtctccaac 11940

catctgatgc tgggaaaatg tcaccaggat gcagccatgc cgtgtggcca atgaaccgag 12000

aaaacacccc ttttctagaa tgctctaaag aggcagaata atccagaggt gaggaaggaa 12060

atactccacc agagacccag gcagttccta caaaagccag actttccttc acctagggag 12120

tgacaagacc agtggaaaac actctcaagc agtaaccccc aaatgctctg caagccagtg 12180

gcgtccagat accgcacaag cgagtgggct gtctaatccc atcatcatga tgtaaatatc 12240

tctaggctgc cccgggctgt gcctgaccct gtcttcagct ttccacacct ccacctacag 12300

cccatgcaca gaaggaccac ccaggaatgc tgcaagtgtg gcacctccag ggccacccag 12360

ggagaaggag ggcagctatg ctggtggctc caggcccatt tggcgggtgg taccttcaca 12420

ccacaaagcc caaactgagg ccccagattt ggctgatgag ggcatattgg acaggggtca 12480

cttatgctct tccccattgc cacctggcct ctggctacct ggacttggct acctgtggat 12540

cctctcacag gtgccaccat cttggctgag tctccagatg cgaggtccct gaggcagtgg 12600

caggcttctc gctaatgctg atgggattag gaatgggata ggtggggagg gccctggact 12660

gggccctgat gagccaagtg ggtttttaga ggggctactg gtacatttca gggacaggac 12720

atctggtaga gctaagctgg ggcaataagg agccactgct aatctgagag ctagaaacaa 12780

tcagcttctg ggtcattatt aattagggta gtttgggctg tgtggaagtc acgtactata 12840

tggggtagcc acagctctct ctacagataa tctctaagac ttctgattgg gaccgtgtga 12900

atgcagtagc aatatctctt cttactgcca ggccctgcca gtcctgcctc cacgccctgg 12960

ctggcccccc ttatgatctg acccatgcca ggctgccata gtatgttact tctgcattag 13020

cactccttgg gacctgcctc tccactgtcc ctcagacttt aaagaactat acaaacccaa 13080

ggggctcttc ccaagagaat tgatatgact tgaggtgatt ccatttctgg aagtagtcac 13140

tccattttct gcctcactct ttcagtgctt cacagagcag gttcgattaa gcacacagat 13200

taattaagca cacagattaa tcgtaactat aacggtccta aggtagcgaa agcttgcatg 13260

cctgcaggtc gactctagag gatcaaacct agccaccgcg gtggcggccg gctagccggc 13320
tagccggcta gccctagaac tagtaacggc cgccagtgtg ctggaattcg gcttgtaagg 13380
taccggtgaa gttcctatac tttctagaga ataggaactt cggaatagga acttctacct 13440
agatgcatgc tcgagcggcc cctacagttg aagtcggaag tttacataca cttaagttgg 13500
agtcattaaa actcgttttt caactactcc acaaatttct tgttaacaaa caatagtttt 13560
ggcaagtcag ttaggacatc tactttgtgc atgacacaag tcatttttcc aacaattgtt 13620
tacagacaga ttatttcact tataattcac tgtatcacaa ttccagtggg tcagaagttt 13680

acatacacta agttgactgt gcctttaaac agcttggaaa attccagaaa atgatgtcat 13740
ggctttagaa gcttctgata gactaattga catcatttga gtcaattgga ggtgtacctg 13800
tggatgtatt tcaaggaatt ctgtggaatg tgtgtcagtt agggtgtgga aagtccccag 13860
gctccccagc aggcagaagt atgcaaagca tgcatatcga tactagttta taagatctcg 13920
agctagggta ccgtcaaggc tgcagtgaga catgatcttg ccactgcact ccagcctgga 13980
cagcagagtg aaaccttgcc tcacgaaaca gaatacaaaa acaaacaaac aaaaaactgc 14040
tccgcaatgc gcttccttga tgctctacca cataggtctg ggtactttgt acacattatc 14100

tcattgctgt tcataattgt tagattaatt ttgtaatatt gatattattc ctagaaagct 14160
gaggcctcaa gatgataact tttattttct ggacttgtaa tagctttctc ttgtattcac 14220
catgttgtaa ctttcttaga gtagtaacaa tataaagtta ttgtgagttt ttgcaaacac 14280
agcaaacaca acgacccata tagacattga tgtgaaattg tctattgtca atttatggga 14340
aaacaagtat gtacttttтc tactaagcca ttgaaacagg aataacagaa caagattgaa 14400
agaatacatt ttccgaaatt acttgagtat tatacaaaga caagcacgtg gacctgggag 14460
gagggttatt gtccatgact ggtgtgtgga gacaaatgca ggtttataat agatgggatg 14520

gcatctagcg caatgacttt gccatcactt ttagagagct cttggggacc ccagtacaca 14580
agaggggacg cagggtatat gtagacatct cattcttttt cttagtgtga gaataagaat 14640
agccatgacc tgagtttata gacaatgagc ccttttctct ctcccactca gcagctatga 14700
gatggcttgc cctgcctctc tactaggctg actcactcca aggcccagca atgggcaggg 14760
ctctgtcagg gctttgatag cactatctgc agagccaggg ccgagaaggg gtggactcca 14820
gagactctcc ctcccattcc cgagcagggt ttgcttattt atgcatttaa atgatatatt 14880
tattttaaaa gaaataacag gagactgccc agccctggct gtgacatgga aactatgtag 14940

aatattttgg gttccatttt tttttccttc tttcagttag aggaaaaggg gctcaggatc 15000

cacttgccca gtgttcttcc ttagttccta ccttcgacct tgatcctcct ttatcttcct 15060

gaaccctgct gagatgatct atgtggggag aatggcttct ttgagaaaca tcttcttcgt 15120

tagtggcctg cccctcattc ccactttaat atccagaatc actataagaa gaatataata 15180

agaggaataa ctcttattat aggtaaggga aaattaagag gcatacgtga tgggatgagt 15240

aagagaggag agggaaggat taatggacga taaaatctac tactatttgt tgagaccttt 15300

tatagtctaa tcaattttgc tattgttttc catcctcacg ctaactccat aaaaaaacac 15360

tattattatc tttattttgc catgacaaga ctgagctcag aagagtcaag catttgccta 15420

aggtcggaca tgtcagaggc agtgccagac ctatgtgaga ctctgcagct actgctcatg 15480

ggccctgtgc tgcactgatg aggaggatca gatggatggg gcaatgaagc aaaggaatca 15540

ttctgtggat aaaggagaca gccatgaaga agtctatgac tgtaaatttg ggagcaggag 15600

tctctaagga cttggatttc aaggaatttt gactcagcaa acacaagacc ctcacggtga 15660

ctttgcgagc tggtgtgcca gatgtgtcta tcagaggttc cagggagggt ggggtggggt 15720

cagggctggc caccagctat cagggcccag atgggttata ggctggcagg ctcagatagg 15780

tggttaggtc aggttggtgg tgctgggtgg agtccatgac tcccaggagc caggagagat 15840

agaccatgag tagagggcag acatgggaaa ggtgggggag gcacagcata gcagcatttt 15900

tcattctact actacatggg actgctcccc tataccccca gctaggggca agtgccttga 15960

ctcctatgtt ttcaggatca tcatctataa agtaagagta ataattgtgt ctatctcata 16020

gggttattat gaggatcaaa ggagatgcac actctctgga ccagtggcct aacagttcag 16080

gacagagcta tgggcttcct atgtatgggt cagtggtctc aatgtagcag gcaagttcca 16140

gaagatagca tcaaccactg ttagagatat actgccagtc tcagagcctg atgttaattt 16200

agcaatgggc tgggaccctc ctccagtaga accttctaac cagctgctgc agtcaaagtc 16260

gaatgcagct ggttagactt tttttaatga aagcttgcat gcagcacttt gggaggctga 16320

ggtgggtgga ctgcttggag ctcaggagtt caagaccatc ttggacaaca tggtgatacc 16380

ctgcctctac aaaaagtaca aaaattagcc tggcatggtg gtgtgcacct gtaatcccag 16440

ctattagggt ggctgaggca ggagaattgc ttgaacccgg gaggcggagg ttgcagtgag 16500

ctgagatcgt gccactgcac tccagcctgg gggacagagc acattataat taactgttat 16560

tttttacttg gactcttgtg gggaataaga tacatgtttt attcttattt atgattcaag 16620

cactgaaaat agtgtttagc atccagcagg tgcttcaaaa ccatttgctg aatgattact 16680

atacttttta caagctcagc tccctctatc ccttccagca tcctcatctc tgattaaata 16740

agcttcagtt tttccttagt tcctgttaca tttctgtgtg tctccattag tgacctccca 16800

tagtccaagc atgagcagtt ctggccaggc ccctgtcggg gtcagtgccc cacccccgcc 16860

ttctggttct gtgtaacctt ctaagcaaac cttctggctc aagcacagca atgctgagtc 16920

atgatgagtc atgctgaggc ttagggtgtg tgcccagatg ttctcagcct agagtgatga 16980

ctcctatctg ggtccccagc aggatgctta cagggcagat ggcaaaaaaa aggagaagct 17040

gaccacctga ctaaaactcc acctcaaacg gcatcataaa gaaaatggat gcctgagaca 17100

gaatgtgaca tattctagaa tatattattt cctgaatata tatatatata tacacatata 17160

ccatatgaaa cacctctagg ctataaggca acagagctcc ttttttttt ttctgtgctt 17220

tcctggctgt ccaaatctct aatgataagc atacttctat tcaatgagaa tattctgtaa 17280

gattatagtt aagaattgtg ggagccattc cgtctcttat agttaaattt gagcttcttt 17340

tatgatcact gttttttaa tatgctttaa gttctggggt acatgtgcca tggtggtttg 17400

ctgcacccat caacccgtca tctacattag gtatttctcc taatgctatc cttcccctag 17460

ccccccaccc ccaacaggcc ccagtgtgtg atgttcccct ccctgtgtcc atggatcact 17520

ggtttttttt tgttttttt ttttttttaa agtctcagtt aaatttttgg aatgtaattt 17580

attttcctgg tatcctaagg acttgcaagt tatctggtca ctttagccct cacgttttga 17640

tgataatcac atatttgtaa acacaacaca cacacacaca cacacacaca tatatatata 17700

tataaaacat atatatacat aaacacacat aacatattta tcgggcattt ctgagcaact 17760

aatcatgcag gactctcaaa cactaaccta tagccttttc tatgtatcta cttgtgtaga 17820

aaccaagcgt ggggactgag aaggcaatag caggagcatt ctgactctca ctgcctttag 17880

ctaggcccct ccctcatcac agctcagcat agtcctgagc tcttatctat atccacacac 17940

agtttctgac gctgcccagc tatcaccatc ccaagtctaa agaaaaaaat aatgggtttg 18000

cccatctctg ttgattagaa aacaaaacaa aataaaataa gcccctaagc tcccagaaaa 18060

catgactaaa ccagcaagaa gaagaaaata caataggtat atgaggagac tggtgacact 18120

aagtgtctga atgaggcttg agtacagaaa agaggctcta gcagcatagt ggtttagagg 18180

agatgtttct ttccttcaca gatgccttag cctcaataag cttgcggttg tggaagttta 18240

cttgtttatc accggtgacg tccatgagca aattaagaaa aacaacaaca aatgaatgca 18300

tatatatgta tatgtatgtg tgtatatata cacatatata tatatatttt ttttcttttc 18360

ttaccagaag gttttaatcc aaataaggag aagatatgct tagaactgag gtagagtttt 18420

catccattct gtcctgtaag tattttgcat attctggaga cgcaggaaga gatccatcta 18480

catatcccaa agctgaatta tggtagacaa aactcttcca cttttagtgc atcaatttct 18540

tatttgtgta ataagaaaat tgggaaaacg atcttcaata tgcttaccaa gctgtgattc 18600

caaatattac gtaaatacac ttgcaaagga ggatgttttt agtagcaatt tgtactgatg 18660

gtatggggcc aagagatata tcttagaggg agggctgagg gtttgaagtc caactcctaa 18720

gccagtgcca gaagagccaa ggacaggtac ggctgtcatc acttagacct caccctgtgg 18780
agccacaccc tagggttggc caatctactc ccaggagcag ggagggcagg agccagggct 18840
gggcataaaa gtcagggcag agccatctat tgcttacatt tgcttctgac acaactgtgt 18900
tcactagcaa cctcaaacag acaccatggg tcatttcaca gaggaggaca aggctactat 18960
cacaagcctg tggggcaagg tgaatgtgga agatgctgga ggagaaaccc tgggaaggta 19020
ggctctggtg accaggacaa gggagggaag gaaggaccct gtgcctggca aaagtccagg 19080
tcgcttctca ggatttgtgg caccttctga ctgtcaaact gttcttgtca atctcacagg 19140

ctcctggttg tctacccatg gacccagagg ttctttgaca gctttggcaa cctgtcctct 19200
gcctctgcca tcatgggcaa ccccaaagtc aaggcacatg gcaagaaggt gctgacttcc 19260
ttgggagatg ccataaagca cctggatgat ctcaagggca cctttgccca gctgagtgaa 19320
ctgcactgtg acaagctgca tgtggatcct gagaacttca aggtgagtcc aggagatgtt 19380
tcagccctgt tgcctttagt ctcgaggcaa cttagacaac tgagtattga tctgagcaca 19440
gcagggtgtg agctgtttga agatactggg gttgggggtg aagaaactgc agaggactaa 19500
ctgggctgag acccagtggt aatgttttag ggcctaagga gcgcctctaa aaatctagat 19560

ggacaatttt gactttgaga aaagagaggt ggaaatgagg aaaatgactt ttattagatt 19620
ccagtagaaa gaactttcat ctttccctca tttttgttcg ttttaaaaca tctatctgga 19680
ggcaggacaa gtatggtcgt taaaaagatg caggcagaag gcatatattg gctcagtcaa 19740
agtggggaac tttggtggcc aaacatacat tgctaaggct attcctatat cagctggaca 19800
catataaaat gctgctaatg cttcattaca aacttatatc ctttaattcc agatgggggc 19860
aaagtatgtc caggggtgag gaacaattga aacatttggg ctggagtaga ttttgaaagt 19920
cagctctgtg tgtgtgtgtg tgtgcgcgcg cgcgtgtgtg tgtgtgtgtg tcaacgtgtg 19980

tttcttttaa cgtcttcagc ctacaacata cagggttcat ggtggcaaga agatagcaag 20040
atttaaatta tggccagtga ctagtgcttg aaggggaaca actacctgca tttaatggga 20100
aggcaaaatc tcaggctttg agggaagtta acataggctt gattctgggt agaagctggg 20160
tgtgtagtta tctggaggcc aggctggagc tctcagctca ctatgggttc atctttattg 20220
tctcctttca tctcaacagc tcctgggaaa tgtgctggtg accgttttgg caatccattt 20280
cggcaaagaa ttcacccctg aggtgcaggc ttcctggcag aagatggtga ctgcagtggc 20340
cagtgccctg tcctccagat accactgagc ctcttgccca tgattcagag ctttcaagga 20400

taggctttat tctgcaagca attcaaataa taaatctatt ctgctgagag atcacacatg 20460

attttcttca gctctttttt ttacatcttt ttaaatatat gagccacaaa gggtttatat 20520

tgagggaagt gtgtatgtgt atttctgcat gcctgtttgt gtttgtggtg tgtgcatgct 20580

cctcatttat ttttatatga gatgtgcatt ttgttgagca aataaaagca gtaaagacac 20640

ttgtacacgg gagttctgca agtgggagta aatggtgtag gagaaatccg gtgggaagaa 20700

agacctctat aggacaggac ttctcagaaa cagatgtttt ggaagagatg ggaaaaggtt 20760

cagtgaagac ctgggggctg gattgattgc agctgagtag caaggatggt tcttaatgaa 20820

gggaaagtgt tccaagcttt aggaattcaa ggtttagtca ggtgtagcaa ttctatttta 20880

ttaggaggaa tactatttct aatggcactt agcttttcac agcccttgtg gatgcctaag 20940

aaagtgaaat taatcccatg ccctcaagtg tgcagattgg tcacagcatt tcaagggaga 21000

gacctcattg taagactctg ggggaggtgg ggacttaggt gtaagaaatg aatcagcaga 21060

ggctcacaag tcagcatgag catgttatgt ctgagaaaca gaccagcact gtgagatcaa 21120

aatgtagtgg gaagaatttg tacaacatta attggaaggc ttacttaatg gaatttttgt 21180

atagttggat gttagtgcat ctctataagt aagagtttaa tatgatggtg ttacggacct 21240

aatgtttgtg tctcctcaaa attcacatgc tgaatcccca actcccaact gaccttatct 21300

gtgggggagg cttttgaaaa gtaattaggt ttagatgagc tcataagagc agatccccat 21360

cataaaatta ttttccttat cagaagcaga gagacaagcc atttctcttt cctcccggtg 21420

aggacacagt gagaagtccg ccatctgcaa tccaggaaga gaaccctgac cacgagtcag 21480

ccttcagaaa tgtgagaaaa aactctgttg ttgaagccac ccagtctttt gtattttgtt 21540

atagcacctt gcactgagta aggcagatga agaaggagaa aaaaataagc ttatcgaaac 21600

gcgtccccat cctcactgac tccgtcctgg agttggatga gagataatgg ccttacgttg 21660

tgccagggga gggtcgggct ggatttagca agatttacct tctccaaaga gcggtgctgc 21720

agtggcacag ctgcccacgg aggtgggggg gtcaccgtcc ctggaggtga tgaagaactg 21780

tggggatgtg gcactgaggg acatggccag tgggcacggt gggtgggttg gggttggtct 21840

tggggatctt ggagggcttt tccagccttc atgatttgac gattgtatga acatctacat 21900

ggcaattctc cagctgcctg tcccagtcct actgacccag ctgtatctct ccaggcaagc 21960

tcttccaccc cttctgcttg catccagaca ccatcaaaca tgcaggctca gacacaggga 22020

ccagcagtgt ctgtggcctt tttgtgctcc tctccatgct gggttttaac ttgctctttg 22080

tccttctatc ctatcttctt atccttaagg ctgttctgaa cgctgtgact tggagagtgt 22140

cccagagccc tcaacacctg catgtcccac gtccatgctg tcctgcactt ccttatcccc 22200

aagatctgcc tctccgtgat gcactgaatt ggcaaacatg tgtcacccca gaccaacaat 22260

gtcacagcaa actccccctt gataggacaa gggggaatgg ctttacactg agacagggga 22320

ggtttgggtt ggatatgagg aggcagtttt tcccccagag ggtggtgacg cactgaacag 22380

gttgcccaag gaggctgtgg atgccccatc cctgcaggca ttcaaggcca ggctggatgt 22440

ggctctgggc agcctgggct gctggttgat gaccctgcac atagcagggg gttggatctg 22500

gatgagcact gtgctccttt gcaacccagg ccgttctatg attctgtcat tctaaatctc 22560

tctttcagcc taaagctttt tccccgtatc cccccaggtg tctgcaggct caaagagcag 22620

cgagaagcgt tcagaggaaa gcgatcccgt gccaccttcc ccgtgcccgg gctgtccccg 22680

cacgctgccg gctcggggat gcgggggggag cgccggaccg gagcggagcc ccgggcggct 22740

cgctgctgcc ccctagcggg ggagggacgt aattacatcc ctgggggctt tgggggggggg 22800

ctgtccctat cgattactag tttagccata gagcccaccg catccccagc atgcctgcta 22860

ttgtcttccc aatcctcccc cttgctgtcc tgccccaccc caccccccag aatagaatga 22920

cacctactca gacaatgcga tgcaatttcc tcattttatt aggaaaggac agtgggagtg 22980

gcaccttcca gggtcaagga aggcacgggg gaggggcaaa caacagatgg ctggcaacta 23040

gaaggcacag tcgctcgaag agcggccgct cgctttactt gtacagctcg tccatgccga 23100

gagtgatccc ggcggcggtc acgaactcca gcaggaccat gtgatcgcgc ttctcgttgg 23160

ggtctttgct cagggcggac tgggtgctca ggtagtggtt gtcgggcagc agcacggggc 23220

cgtcgccgat gggggtgttc tgctggtagt ggtcggcgag ctgcacgctg ccgtcctcga 23280

tgttgtggcg gatcttgaag ttcaccttga tgccgttctt ctgcttgtcg gccatgatat 23340

agacgttgtg gctgttgtag ttgtactcca gcttgtgccc caggatgttg ccgtcctcct 23400

tgaagtcgat gcccttcagc tcgatgcggt tcaccagggt gtcgccctcg aacttcacct 23460

cggcgcgggt cttgtagttg ccgtcgtcct tgaagaagat ggtgcgctcc tggacgtagc 23520

cttcgggcat ggcggacttg aagaagtcgt gctgcttcat gtggtcgggg tagcggctga 23580

agcactgcac gccgtaggtc agggtggtca cgagggtggg ccagggcacg ggcagcttgc 23640

cggtggtgca gatgaacttc agggtcagct tgccgtaggt ggcatcgccc tcgccctcgc 23700

cggacacgct gaacttgtgg ccgtttacgt cgccgtccag ctcgaccagg atgggcacca 23760

ccccggtgaa cagctcctcg cccttgctca ccatgggccc tgggttggac tccacgtctc 23820

ccgccaactt gagaaggtca aaattcaaag tctgtttcac ctcgaggttt cggccagcag 23880

gcggggagcc cgaggtagct cccgctccct tgagccaggc ccctgccaga cctgagctcc 23940

ctcccaagcc tggcttcccc aaccggtggc cttcatgggc cagaagccat tccttcacgg 24000

ctagccctcc ggagtagttg cccacggctc cgctgctgca gaccactctg tggcacggga 24060

tgaggatctt gacaggattg cctctcatgg cgcctcccac tgctcgcgcg gctttggggt 24120

tgccggccag ggcggccaat tgctggtaag aaatcacttc tccgaatttc acaaccttaa 24180

gcagcttcca taacacctga cgcgtgaacg actcttgctg gaaaacggga tggtgaagcg 24240
ctggcacggg gaactcttcg atagcctcgg gctggtggaa ataggcattc agccaggctg 24300
tgcactgcat caggggctcc ggacctccga gaaccgcagc gggggctggg acctccacgg 24360
catcagctgc agacgtcccc ttgcccagga gctttatttc gtgcagaccc tgctcacaac 24420
cagacagctc cagcttcccc aaagggctgt ccagtgtggt gcgtttcatt tcacaatcct 24480
tgtccatggt ggccctaggc cctggggaga gaggtcggtg attcggtcaa cgagggagcc 24540
gactgccgac gtgcgctccg gaggcttgca gaatgcggaa caccgcgcgg gcaggaacag 24600

ggcccacact accgccccac accccgcctc ccgcaccgcc ccttcccggc cgctgctctc 24660
ggcgcgcccc gctgagcagc cgctattggc cacagcccat cgcggtcggc gcgctgccat 24720
tgctccctgg cgctgtccgt ctgcgagggt actagtgaga cgtgcggctt ccgtttgtca 24780
cgtccggcac gccgcgaacc gcaaggaacc ttcccgactt aggggcggag caggaagcgt 24840
cgccgggggg cccacaaggg tagcggcgaa gatccgggtg acgctgcgaa cggacgtgaa 24900
gaatgtgcga gacccagggt cggcgccgct gcgtttcccg gaaccacgcc cagagcagcc 24960
gcgtccctgc gcaaacccag ggctgccttg gaaaaggcgc aaccccaacc ccgtggaaat 25020

aaatcgataa ctagtgatat catcatgtct ggatcccatc acaaagctct gacctcaatc 25080
ctatagaaag gaggaatgag ccaaaattca cccaacttat tgtgggaagc ttgtggaagg 25140
ctactcgaaa tgtttgaccc aagttaaaca atttaaaggc aatgctacca aatactaatt 25200
gagtgtatgt taacttctga cccactggga atgtgatgaa agaaataaaa gctgaaatga 25260
atcattctct ctactattat tctgatattt cacattctta aaataaagtg gtgatcctaa 25320
ctgaccttaa gacagggaat ctttactcgg attaaatgtc aggaattgtg aaaaaagtga 25380
gtttaaatgt atttggctaa ggtgtatgta aacttccgac ttcaactgta ggggatcctc 25440

tagggccgcc agtgtgatgg atatctgcag aattcggctt caggtaccgt cgacgatgta 25500
ggtcacggtc tcgaagccgc ggtgcgggtg ccagggcgtg cccttgggct ccccgggcgc 25560
gtactccacc tcacccatct ggtccatcat gatgaacggg tcgaggtggc ggtagttgat 25620
cccggcgaac gcgcggcgca ccgggaagcc ctcgccctcg aaaccgctgg gcgcggtggt 25680
cacggtgagc acgggacgtg cgacggcgtc ggcgggtgcg gatacgcggg gcagcgtcag 25740
cgggttctcg acggtcacgg cgggcatgtc gacaagccga attccagcac actggcggcc 25800
gttactaggt agctagctcg agccttcgaa gatctcctag ggaagttcct atactttcta 25860

gagaatagga acttcggaat aggaacttca ccggtgggtg aaaagccgaa ttctgcagat 25920

atcaagctta tcgataccgt cgacctcgag gggggcccg gttagatccc cgggtaccga 25980

gctcgaatat ctatgtcggg tgcggagaaa gaggtaatga aatggattaa gtggcaggat 26040

taattaagtg gcaggattaa tcttcgaacg aaggagccat ccaactaacc gtcatgttcg 26100

ggcaaccgaa gaagggagtg gcaggatttc ctttggagac ttctggaatt agacagcagt 26160

ttaatgcaag catctaaatt ctcttcctcc cagagtctca ttaaaactac agtaagagtt 26220

tgtgttttgt tttgttttta aagacaaaat cccaccagga tagagagaat aggagaggag 26280

ataacagcat cataatttat gaaactaaaa tgcagataga ccaggattaa ctgactacac 26340

agcaccaagg aagctgaatc acaagacagc agaggagaaa actggaaagg atcgtggtct 26400

atacggcaga atcttcccaa gcctcaggag gaggagctct agatgcttat gatggcaact 26460

aaagcctaaa agctaattca ttttaaagtt cttccaaatg catagggttt tatttttcca 26520

gacctgggtt cagatgggga atttgacaaa caatggaaag ggggaaaaac aacaatctaa 26580

acactgagtg acaaagtaac aaagaaatag tctagctatc agccagtcaa gccagccttg 26640

gctttgctat ccaaagtagt cagtctaatt ctaccaccag tttctgttcc tgtagctgtc 26700

tactgcctgc cagggactct gccttcccac ccacaactac caatggaagg atgtggtgac 26760

cataccagtg gctgctgaca tctcctgcca tgggaagcat aattgcctcc agcagcctcc 26820

cccttagatc catcattttt gttgcacttg gcctgggctg tactcccggc caatgactga 26880

acatggtgag catagtaatg caggcccatt tctgtgagga gcaggactcc tccagtaggt 26940

gactttggct caaggactct ctattggcct ggttgaactt ttcctgaact gtgctactgt 27000

ctgagactct tcttacccaa tcctctttct cgccccaatt gtcacagacc acctgcattg 27060

tggtctgagt ctctccccac cttctcttgc tcttccctgt ttatctttca caggcatttc 27120

ccccagtaca ttccttgaat gtctaacccg atacgggtgc ctgacttttg gcagacctaa 27180

gcagacaaaa aggagtactt ggttacctag ctcttctttc taccacaaac atcgagggaa 27240

ccctttttcc ctcacccctc tgccacaccc ccactgcccc agtgaacaac cacagagaga 27300

gctgtggtat aatattaggc tggtgcaaaa gtaattgcgg tttttgccat tacttttaat 27360

ggtaaaaacc gcaattactt ttgcacctac ctagtatttg tgtcccccca aattcatatg 27420

ttgaaaccta acccacaata tgatgtcatt aggaggcaag accttgagga ggtgattaga 27480

tgatggggtg gagctctcct gaatgagatt agtgccctta taagaagaag cccaaggaag 27540

ctaccttgac tcttccatca catgagaatg cagcaagaag gcaccatcta ctaatcagga 27600

agagagctct caccagacac tgaatctgcc agtgtcttga tcttgaagtt cccagcctcc 27660

agaactatgc ataatgcatt tccattgtct ctaagccacc cagcctatgg tattttgtca 27720

tagcagcctg aactgactaa gacagtgagc cacatgagaa gtgccccaac ccctccctta 27780

agcacttggc tcacagatca gtgggttcat ttctgcctga gttttattgt tattctgtag 27840

atttcttggg ctagatatat ttttctgtta ttttccttct tcacctcagt catgaattgg 27900

ttgttttaaa aaagacaatg taagtcatgg ggaaactcct gacaactcta ctctcctagg 27960

gttcctgata aaaggggatt cagttgagtc ctctgatggt ctctacctgc caaagtccag 28020

cagcccttag caaacatgct gctcgtttct gtagagaagg tgctggtgtc ccaccatact 28080

tctctctccc tcatgaaggg cttgcgaccc agcaaatggg tggcttatat gggtctgttt 28140

caaaggaaga gccagctctg ggaagaaaaa cgatgagcat aagcataacc taccactgtg 28200

cctgggaaag cagacaactt ttttgatgtg tgaatatcta atgagaatgg aatccatcaa 28260

ttaccttaaa cttaggcaca gtcttcaaat tcaatatatg tgggatatac ttttagtcag 28320

tttgtagacg ttatttgtaa taaataatct ggcttctcta aagaaattat tttaagtgtt 28380

tggtttggtt tgatttaatg gtaaaattat atttagtggc agagaattat agcaatggtg 28440

ataaactata gagtgtcata agttcatatc ttattctcac atttgaagct gcctgcagat 28500

gcattcaaga tgcagccaga agtcaggaga ctcaggctgt tatttggagc tcatcatttt 28560

acagccttgc tggactccca ctttctcagg ggaaaaatgt ggtgttgacc cagattagct 28620

ctccaggccc tgctgagttg ggcactctgt aagctggagg gtcttctatt gtcttcacct 28680

aagtgtcaat caacaaccca aatgggcatg ggggaagagg gagctgggcc aatgcccagg 28740

gtgcctggta gagagatacc ttgggcactg gaaggcacca gcttcccaga gagaaggggg 28800

agggccatga aaaagttggc tgtagatgcc agggacactg ggactctcca gctgtgtgtt 28860

tgtgtcttct gaagacttat gtttcattcc tttggagcat gcataatcat acactgtggg 28920

atgtgttata tagattgctt gatagttcac cactgtaata aaatactgtg actggaatct 28980

gctcccagtc tgcctttgat agcacttgtg caacacacat ttactgagca tttacagtga 29040

tccaggacct gtgttgtgaa aacattgatg gacaaggcag atggtggagc acgtcagtga 29100

ggatttttaa caaaggctgg taagtgctat aaaggaacat tgtaggacac tagagaacaa 29160

agaacaggag aacctgactt aggctggggt ggggcgttgg ttagaggagg ctccttggag 29220

gacatgaggt ttaagctgtg acctgaggat gaatagatgt tggccaggtg aggttaatct 29280

ggggctcaag atcgagcatt aagcttgtca gccttaccag taaaaaagaa aacctattaa 29340

aaaaacacca ctcgacacgg caccagctca atcagtcaca gtgtaaaaaa gggccaagtg 29400

cagagcgagt atatatagga ctaaaaaatg acgtaacggt taaagtccac aaaaaacacc 29460

cagaaaaccg cacgcgaacc tacgcccaga aacgaaagcc aaaaaaccca caacttcctc 29520

aaatcgtcac ttccgttttc ccacgttacg taacttccca ttttaagaaa actacaattc 29580

ccaacacata caagttactc cgccctaaaa cctacgtcac ccgccccgtt cccacgcccc 29640

gcgccacgtc acaaactcca ccccctcatt atcatattgg cttcaatcca aaataaggta 29700
tattattgat gatgtttaaa ctacggcccg gtacccagct tttgttccct ttagtgaggg 29760
ttaattgcgc gcttggcgta atcatggtca tagctgtttc ctgtgtgaaa ttgttatccg 29820
ctcacaattc cacacaacat acgagccgga agcataaagt gtaaagcctg gggtgcctaa 29880
tgagtgagct aactcacatt aattgcgttg cgctcactgc ccgctttcca gtcgggaaac 29940
ctgtcgtgcc agctgcatta atgaatcggc caacgcgcgg ggagaggcgg tttgcgtatt 30000
gggcgctctt ccgcttcctc gctcactgac tcgctgcgct cggtcgttcg gctgcggcga 30060

gcggtatcag ctcactcaaa ggcggtaata cggttatcca cagaatcagg ggataacgca 30120
ggaaagaaca tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa ggccgcgttg 30180
ctggcgtttt tccataggct ccgccccccct gacgagcatc acaaaaatcg acgctcaagt 30240
cagaggtggc gaaacccgac aggactataa agataccagg cgtttccccc tggaagctcc 30300
ctcgtgcgct ctcctgttcc gaccctgccg cttaccggat acctgtccgc ctttctccct 30360
tcgggaagcg tggcgctttc tcatagctca cgctgtaggt atctcagttc ggtgtaggtc 30420
gttcgctcca agctgggctg tgtgcacgaa ccccccgttc agcccgaccg ctgcgcctta 30480

tccggtaact atcgtcttga gtccaacccg gtaagacacg acttatcgcc actggcagca 30540
gccactggta acaggattag cagagcgagg tatgtaggcg gtgctacaga gttcttgaag 30600
tggtggccta actacggcta cactagaagg acagtatttg gtatctgcgc tctgctgaag 30660
ccagttacct tcggaaaaag agttggtagc tcttgatccg gcaaacaaac caccgctggt 30720
agcggtggtt tttttgtttg caagcagcag attacgcgca gaaaaaaagg atctcaagaa 30780
gatcctttga tcttttctac ggggtctgac gctcagtgga acgaaaactc acgttaaggg 30840
attttggtca tgagattatc aaaaaggatc ttcacctaga tccttttaaa ttaaaaatga 30900

agttttaaat caatctaaag tatatatgag taaacttggt ctgacagtta ccaatgctta 30960
atcagtgagg cacctatctc agcgatctgt ctatttcgtt catccatagt tgcctgactc 31020
cccgtcgtgt agataactac gatacgggag ggcttaccat ctggccccag tgctgcaatg 31080
ataccgcgag acccacgctc accggctcca gatttatcag caataaacca gccagccgga 31140
agggccgagc gcagaagtgg tcctgcaact ttatccgcct ccatccagtc tattaattgt 31200
tgccgggaag ctagagtaag tagttcgcca gttaatagtt tgcgcaacgt tgttgccatt 31260
gctacaggca tcgtggtgtc acgctcgtcg tttggtatgg cttcattcag ctccggttcc 31320

caacgatcaa ggcgagttac atgatccccc atgttgtgca aaaaagcggt tagctccttc 31380

ggtcctccga tcgttgtcag aagtaagttg gccgcagtgt tatcactcat ggttatggca 31440

gcactgcata attctcttac tgtcatgcca tccgtaagat gcttttctgt gactggtgag 31500

tactcaacca agtcattctg agaatagtgt atgcggcgac cgagttgctc ttgcccggcg 31560

tcaatacggg ataataccgc gccacatagc agaactttaa aagtgctcat cattggaaaa 31620

cgttcttcgg ggcgaaaact ctcaaggatc ttaccgctgt tgagatccag ttcgatgtaa 31680

cccactcgtg cacccaactg atcttcagca tcttttactt tcaccagcgt ttctgggtga 31740

gcaaaaacag gaaggcaaaa tgccgcaaaa aagggaataa gggcgacacg gaaatgttga 31800

atactcatac tcttcctttt tcaatattat tgaagcattt atcagggtta ttgtctcatg 31860

agcggataca tatttgaatg tatttagaaa aataaacaaa taggggttcc gcgcacattt 31920

ccccgaaaag tgccacctga cgcgccctgt agcggcgcat taagcgcggc gggtgtggtg 31980

gttacgcgca gcgtgaccgc tacacttgcc agcgccctag cgcccgctcc tttcgctttc 32040

ttcccttcct ttctcgccac gttcgccggc tttccccgtc aagctctaaa tcgggggctc 32100

cctttagggt tccgatttag tgctttacgg cacctcgacc ccaaaaaact tgattagggt 32160

gatggttcac gtagtgggcc atcgccctga tagacggttt ttcgcccttt gacgttggag 32220

tccacgttct ttaatagtgg actcttgttc caaactggaa caacactcaa ccctatctcg 32280

gtctattctt ttgatttata agggattttg ccgatttcgg cctattggtt aaaaaatgag 32340

ctgatttaac aaaaatttaa cgcgaatttt aacaaaatat taacgcttac aatttccatt 32400

cgccattcag gctgcgcaac tgttgggaag ggcgatcggt gcgggcctct tcgctattac 32460

gccagctggc gaaaggggga tgtgctgcaa ggcgattaag ttgggtaacg ccagggtttt 32520

cccagtcacg acgttgtaaa acgacggcca gtgagcgcgc gtaatacgac tcactatagg 32580

gcgaattgga gctccactac gtagtttaaa catcatcaat aatatacctt attttggatt 32640

gaagccaata tgataatgag ggggtggagt ttgtgacgtg gcgcggggcg tgggaacgg 32699

<210> 61

<211> 32312

<212> DNA

<213> Artificial Sequence

<220><223> pHCA-Dual-FI-globin-mgmt

<400> 61

ggccgaagga ttacatgagc ttagaaatgt aattagcata gtgtgtggca tagtgtagat 60

accaaataaa tatgatctct ccttctactc ttgaaaatgc aaacacattc ttggtggtcc 120

taaaatagcc tgtaacatgg tttactcagc agcatttgct attcaaggca gatctgcctt 180

tagtcattgg ctgcgctcct gaacagctgt gtgaaaggct aacttttgta aaccaaatca 240

aaataaaatg cagcaaaaat ttgtcactga aaggaaatcc tcagtatatc cttttatgaa 300
atgaaagatc cctcatccaa acttaacttt tttaaaagtg cgcatttgga gatatagccc 360
tttcttatga atcctaattc aattttggcc ataaacacac gttgatgttc cccaccccaa 420
agcacatagc aacaagagta ggttctatat tgaaaataat gacaatttaa aaacatgtac 480
ttatttcact gtatgtggac agtgtctatg attgcatcat gaagtgtcat ataaccatgt 540
acgtgtacat gagagagaga tagagagaga agtggtaggg tggtggtggt agaggggatg 600
gcgatagtaa tcatggtaat ggtagaggtg atggaggtgg taatgacgga ggtaagggtg 660

gtagtgatga tggtggtggt ggtaatggtg gtggatgtgg tggtggcaat tgggatggtg 720
ggatggtggt agccatggtg atggtggtaa tggtgttgat ttaaagggtg gtggtagtga 780
aggtgagggt agtggtggtg gaggtggtgg tgctggtagc aatagtgatg gtggtgatgg 840
tgttgatgag ggtgttggga tcagggtgag ttcccacagt atatttcatt cttgttgtac 900
cactctgtca acagcaccac tgactgggac agaggaagaa ggcacactct gaatgtgtta 960
ttaacagaaa cctcaaaaca gtctgtctcc ttgtagtcat tcaaaattat ctttttctta 1020
cctggaaaac tgaaactgaa ttaccgggaa aaacacagga gatttttgtt tgttaatatg 1080

ctgccaataa agtaatttta tgtcaaattt aactacagga aagggcaagg cattttctaa 1140
gttccttaga tgtcatgtgg ctaaaaaaaa caaaaggatg gacagcagtt agatactgta 1200
cacttagctg tttgaagcca tatattcaga aagcagatgt tgggagttgg tgtttgagga 1260
ctgatttcct ggaggtattt tatataggcc aagttcattg ttctaaactc taagggcttg 1320
acttgaggga ggaaaagagg caagaacatg tttagttttg ctgacagcat cacatgggca 1380
gccctaaggc tagacaactt tagggcctga agcttattct aggaaagaag cacctacaga 1440
gtggcactgg gctcccctcc actatagaga tgaagtcata tgacagtaaa gggcaggcag 1500

ggctgcctag ggggcccaga actgacactt ccattagaat gagcacaggc cagggagaga 1560
agtggggaac cagagagaag gagctggaat tcctttctct ccatacataa atgcctgcag 1620
agtcccattt cagaatccgg cagacaaagc caccaatgtg atccccatga ccttataaac 1680
attcattaaa atgcatttca aggcatgtga tggcctcccc acccctaga taatgagaaa 1740
acaaaggttt ctcttctgat agagacaagt tcagctctga agtcaacatt atttctggtt 1800
ctgtctgaac aatgacatat ggcaactctt ccctttctat agttctagtc cagaatgaca 1860
aaaaagggga aaaatttctt agagaaggta gagattatac gaatacagtc catgaaatga 1920

gcataaggag aataaagaat ataacttatc caaagaagtc tggcaggctg ttataaatgc 1980

ttgattttgg acactgtagt tggaggttta acatggacac caataaaaag gtcagcaaag 2040

ggtatgcact gttcctattg ggcaagaaga taggaggtca aaggtaacca ggaaagataa 2100

actcagggag acttattttc cctccagagg gcactgggct tgtaggccct gggcaaaatt 2160

gtcaaaaagg tgaaaatcgc ctgtggttta tttagtctgc tctttcttca ctagtgcctc 2220

accagttcag ttcaggccaa tttgctagct accacatttg tagaggtttt acttgcttta 2280

aaaaacctcc cacacctccc cctgaacctg aaacataaaa tgaatgcaat tgttgttgtt 2340

aacttgttta ttgcagctta taatggttac aaataaagca atagcatcac aaatttcaca 2400

aataaagcat ttttttcact gcattctagt tgtggtttgt ccaaactcat caatgtatct 2460

tatcatgtct gaagctttct aggtattgaa taagaaaaat gaagttaagg tggttgatgg 2520

taacactatg ctaataactg cagagccaga agcaccataa gggacatgat aagggagcca 2580

gcagacctct gatctcttcc tgaatgctaa tcttaaacat cctgaggaag aatgggactt 2640

ccatttgggg tgggcctatg atagggtaat aagacagtag tgaatatcaa gctacaaaaa 2700

gccccctttc aaattcttct cagtcctaac ttttcatact aagcccagtc cttccaaagc 2760

agactgtgaa agagtgatag ttccgggaga ctagcactgc agattccggg tcactgtgag 2820

tgggggaggc agggaagaag ggctcacagg acagtcaaac catgcccccc gtttttcctt 2880

cttcaagtag acctctataa gacaacagag acaactaagg ctgagtggcc aggcgaggag 2940

aaaccatctc gccgtaaaac atggaaggaa cacttcaggg gaaaggtggt atctctaagc 3000

aagagaactg agtggagtca aggctgagag atgcaggata agcaaatggg tagtgaaaag 3060

acattcatga ggacagctaa aacaataagt aatgtaaaat acagcatagc aaaactttaa 3120

cctccaaatc aagcctctac ttgaatcctt ttctgaggga tgaataaggc ataggcatca 3180

ggggctgttg ccaatgtgca ttagctgttt gcagcctcac cttctttcat ggagtttaag 3240

atatagtgta ttttcccaag gtttgaacta gctcttcatt tctttatgtt ttaaatgcac 3300

tgacctccca cattcccttt ttagtaaaat attcagaaat aatttaaata catcattgca 3360

atgaaaataa atgtttttta ttaggcagaa tccagatgct caaggccctt cataatatcc 3420

cccagtttag tagttggact tagggaacaa aggaaccttt aatagaaatt ggacagcaag 3480

aaagcgagct tagtgatact tgtgggccag ggcattagcc acaccagcca ccactttctg 3540

ataggcagcc tgcactggtg gggtggcggc cgccctagga ttatggcact ggtagaattc 3600

actacttatg gcactggtag aattcactac ttatggcact ggtagaattc actacttatg 3660

gcactggtag aattcactat cgttgtgctt gatctaacca tgtttcattg tgcttgatct 3720

aaccatgttt cattgtgctt gatctaacca tgtttcattg tgcttgatct aaccatgtat 3780

cgcccggggg cggccgcaca caaaaaacca acacacagat ctaatgaaaa taaagatctt 3840

ttattgaatt cttagctggc ctccaccttt ctcttcttct tggggctgtc gcctcccagc 3900

tgagacaggt cgatccgtgt ctcgtacagg ccggtgatgc tctggtggat cagggtggcg 3960

tccagcacct ctttggtgct ggtgtacctc ttccggtcga tggtggtgtc aaagtacttg 4020

aaggcggcag gggctcccag attggtcagg gtaaacaggt ggatgatatt ctcggcctgc 4080

tctctgatgg gcttatcccg gtgcttgttg taggcggaca gcactttgtc cagattagcg 4140

tcggccagga tcactctctt ggagaactcg ctgatctgct cgatgatctc gtccaggtag 4200

tgcttgtgct gttccacaaa cagctgtttc tgctcattat cctcgggggа gcccttcagc 4260

ttctcatagt ggctggccag gtacaggaag ttcacatatt tggagggcag ggccagttcg 4320

tttcccttct gcagttcgcc ggcagaggcc agcattctct tccggccgtt ttccagctcg 4380

aacagggagt acttaggcag cttgatgatc aggtcctttt tcacttcttt gtagcccttg 4440

gcttccagaa agtcgatggg attcttctcg aagctgcttc tttccatgat ggtgatcccc 4500

agcagctctt tcacactctt cagtttcttg gacttgccct tttccacttt ggccaccacc 4560

agcacagaat aggccacggt ggggctgtcg aagccgccgt acttcttagg gtcccagtcc 4620

ttctttctgg cgatcagctt atcgctgttc ctcttgggca ggatagactc tttgctgaag 4680

ccgcctgtct gcacctcggt ctttttcacg atattcactt ggggcatgct cagcactttc 4740

cgcacggtgg caaaatcccg gcccttatcc cacacgatct ccccggtttc gccgtttgtc 4800

tcgatcagag gccgcttccg gatctcgccg ttggccaggg taatctcggt cttgaaaaag 4860

ttcatgatgt tgctgtagaa gaagtacttg gcggtagcct tgccgatttc ctgctcgctc 4920

ttggcgatca tcttccgcac gtcgtacacc ttgtagtcgc cgtacacgaa ctcgctttcc 4980

agcttagggt acttttgat cagggcggtt cccacgacgg cgttcaggta ggcgtcgtgg 5040

gcgtggtggt agttgttgat ctcgcgcact ttgtaaaact ggaaatcctt ccggaaatcg 5100

gacaccagct tggacttcag ggtgatcact ttcacttccc ggatcagctt gtcattctcg 5160

tcgtacttag tgttcatccg ggagtccagg atctgtgcca cgtgctttgt gatctgccgg 5220

gtttccacca gctgtctctt gatgaagccg gccttatcca gttcgctcag gccgcctctc 5280

tcggccttgg tcagattgtc gaactttctc tgggtaatca gcttggcgtt cagcagctgc 5340

cgccagtagt tcttcatctt cttcacgacc tcttcggagg gcacgttgtc gctcttgccc 5400

cggttcttgt cgcttctggt cagcaccttg ttgtcgatgg agtcgtcctt cagaaagctc 5460

tgaggcacga tatggtccac atcgtagtcg gacagccggt tgatgtccag ttcctggtcc 5520

acgtacatat cccgcccatt ctgcaggtag tacaggtaca gcttctcgtt ctgcagctgg 5580

gtgttttcca cggggtgttc tttcaggatc tggctgccca gctctttgat gccctcttcg 5640

atccgcttca ttctctcgcg gctgttcttc tgtcccttct gggtggtctg gttctctctg 5700

gccatttcga tcacgatgtt ctcgggcttg tgccggccca tcactttcac gagctcgtcc 5760
accaccttca ctgtctgcag gatgcccttc ttaatggcgg ggctgccggc cagattggca 5820
atgtgctcgt gcaggctatc gccctggccg gacacctggg ctttctggat gtcctcttta 5880
aaggtcaggc tgtcgtcgtg gatcagctgc atgaagtttc tgttggcgaa gccgtcggac 5940
ttcaggaaat ccaggattgt cttgccggac tgcttgtccc ggatgccgtt gatcagcttc 6000
cggctcagcc tgccccagcc ggtgtatctc cgccgcttca gctgcttcat cactttgtcg 6060
tcgaacaggt gggcataggt tttcagccgt tcctcgatca tctctctgtc ctcaaacagt 6120

gtcagggtca gcacgatatc ttccagaatg tcctcgtttt cctcattgtc caggaagtcc 6180
ttgtccttga taattttcag cagatcgtgg tatgtgccca gggaggcgtt gaaccgatct 6240
tccacgccgg agatttccac ggagtcgaag cactcgattt tcttgaagta gtcctctttc 6300
agctgcttca cggtcacttt ccggttggtc ttgaacagca ggtccacgat ggccttttc 6360
tgctcgccgc tcaggaaggc gggctttctc attccctcgg tcacgtattt cactttggtc 6420
agctcgttat acacggtgaa gtactcgtac agcaggctgt gcttgggcag caccttctcg 6480
ttgggcaggt tcttatcgaa gttggtcatc cgctcgatga agctctgggc ggaagcgccc 6540

ttgtccacca cttcctcgaa gttccagggg gtgatggttt cctcgctctt tctggtcatc 6600
caggcgaatc tgctgtttcc cctggccaga gggcccacgt agtaggggat gcggaaggtc 6660
aggatcttct cgatcttttc ccggttgtcc ttcaggaatg ggtaaaaatc ttcctgccgc 6720
cgcagaatgg cgtgcagctc tcccaggtgg atctggtggg ggatgctgcc gttgtcgaag 6780
gtccgctgct tccgcagcag gtcctctctg ttcagcttca cgagcagttc ctcggtgccg 6840
tccatctttt ccaggatggg cttgatgaac ttgtagaact cttcctggct ggctccgccg 6900
tcaatgtagc cggcgtagcc gttcttgctc tggtcgaaga aaatctcttt gtacttctca 6960

ggcagctgct gccgcacgag agctttcagc agggtcaggt cctggtggtg ctcgtcgtat 7020
ctcttgatca tagaggcgct caggggggcc ttggtgatct cggtgttcac tctcaggatg 7080
tcgctcagca ggatggcgtc ggacaggttc ttggcggcca gaaacaggtc ggcgtactgg 7140
tcgccgatct gggccagcag gttgtccagg tcgtcgtcgt aggtgtcctt gctcagctgc 7200
agtttggcat cctcggccag gtcgaagttg ctcttgaagt tgggggtcag gcccaggctc 7260
agggcaatca ggttgccgaa caggccattc ttcttctcgc cgggcagctg ggcgatcaga 7320
ttttccagcc gtctgctctt gctcagtctg gcagacagga tggccttggc gtccacgccg 7380

ctggcgttga tggggttttc ctcgaacagc tggttgtagg tctgcaccag ctggatgaac 7440

agcttgtcca cgtcgctgtt gtcggggttc aggtcgccct cgatcaggaa gtggccccgg 7500

aacttgatca tgtgggccag ggccagatag atcagccgca ggtcggcctt gtcggtgctg 7560

tccaccagtt tctttctcag gtggtagatg gtggggtact tctcgtggta ggccacctcg 7620

tccacgatgt tgccgaagat ggggtgccgc tcgtgcttct tatcctcttc caccaggaag 7680

gactcttcca gtctgtggaa gaagctgtcg tccaccttgg ccatctcgtt gctgaagatc 7740

tcttgcagat agcagatccg gttcttccgt ctggtgtatc ttcttctggc ggttctcttc 7800

agccgggtgg cctcggctgt ttcgccgctg tcgaacagca gggctccgat caggttcttc 7860

ttgatgctgt gccggtcggt gttgcccagc accttgaatt tcttgctggg caccttgtac 7920

tcgtcggtga tcacggccca gcccacagag ttggtgccga tgtccaggcc gatgctgtac 7980

ttcttgtcgg acgcttcgac cttgcgcttt ttcttcggcg aagcgtaatc tggaacatcg 8040

tatgggtaca tggtggccct aggatcgatt actagctcac gacacctgaa atggaagaaa 8100

aaaactttga accactgtct gaggcttgag aatgaaccaa gatccaaact caaaaagggc 8160

aaattccaag gagaattaca tcaagtgcca agctggccta acttcagtct ccacccactc 8220

agtgtgggga aactccatcg cataaaaccc ctccccccaa cctaaagacg acgtactcca 8280

aaagctcgag aactaatcga ggtgcctgga cggcgcccgg tactccgtgg agtcacatga 8340

agcgacggct gaggacggaa aggccctttt cctttgtgtg ggtgactcac ccgcccgctc 8400

tcccgagcgc cgcgtcctcc attttgagct ccctgcagca gggccgggaa gcggccatct 8460

ttccgctcac gcaactggtg ccgaccgggc cagccttgcc gcccagggcg gggcgataca 8520

cggcggcgcg aggccaggca ccagagcagg ccggcgagct tgagactacc cccgtccgat 8580

tctcggtggc cgcgctcgca ggccccgcct cgccgaacat gtgcgctggg acgcacgggc 8640

cccgtcgccg cccgcggccc caaaaaccga aataccagtg tgcagatctt ggcccgcatt 8700

tacaagacta tcttgccaga aaaaaagcgt cgcagcaggt catcaaaaat tttaaatggc 8760

tagagactta tcgaaagcag cgagacaggc gcgaaggtgc caccagattc gcacgcggcg 8820

gccccagcgc ccaggccagg cctcaactca agcacgaggc gaaggggctc cttaagcgca 8880

aggcctcgaa ctctcccacc cacttccaac ccgaagctcg ggatcaagaa tcacgtactg 8940

cagccagggg cgtggaagta attcaaggca cgcaagggcc ataacccgta aagaggccag 9000

gcccgcggga accacacacg gcacttacct gtgttctggc ggcaaacccg ttgcgaaaaa 9060

gaacgttcac ggcgactact gcacttatat acggttctcc cccaccctcg ggaaaaaggc 9120

ggagccagta cacgacatca ctttcccagt ttaccccgcg ccaccttctc taggcaccgg 9180

ttcaattgcc gacccctccc cccaacttct cggggactgt gggcgatgtg cgctctgccc 9240

actgacgggc accggagcct cacgcatgct cttctccacc tcagtgatga cgagagcggg 9300

cgggtgaggg ggcgggaacg cagcgatctc tgggttctac gttagtggga gtttaacgac 9360

ggtccctggg attccccaag gcaggggcga gtccttttgt atgaattact caaatcgatt 9420

aggatccggc gcgcccaccg cggaaaaaaa gcaccgactc ggtgccactt tttcaagttg 9480

ataacggact agccttattt taacttgcta tttctagctc taaaacaacc cttcctggag 9540

cctgtgcggt gtttcgtcct ttccacaaga tatataaagc caagaaatcg aaatactttc 9600

aagttacgat aagcatatga tagtccattt taaaacataa ttttaaaact gcaaactacc 9660

caagaaatta ttactttcta cgtcacgtat tttgtactaa tatctttgtg tttacagtca 9720

aattaattcc aattatctct ctaacagcct tgtatcgtat atgcaaatat gaaggaatca 9780

tgggaaatag gccctcggcg cgcccaccgc ggaaaaaaag caccgactcg gtgccacttt 9840

ttcaagttga taacggacta gccttatttt aacttgctat ttctagctct aaaactgcct 9900

tgaccaatag ccttgcggtg tttcgtcctt tccacaagat atataaagcc aagaaatcga 9960

aatactttca agttacggta agcatatgat agtccatttt aaaacataat tttaaaactg 10020

caaactaccc aagaaattat tactttctac gtcacgtatt ttgtactaat atctttgtgt 10080

ttacagtcaa attaattcta attatctctc taacagcctt gtatcgtata tgcaaatatg 10140

aaggaatcat gggaaatagg ccctcggatc cttaattgct agcctctaac tcctagaccg 10200

tcagaactgc tgggcccttc aagacgggct gctcacaccc actcatgtta agcctggtga 10260

ggcctgtact ctgttttcac aggaagaaat cctcacccag tcttccccaa acacattccc 10320

aggttctgtc attagtggga tagagatgat tactgtgggg agaagagaaa catctggatg 10380

gatttggtga ggttgatcta tagaggaagt aggtgctgcc tgaggtagct gtaatagaag 10440

ctaaaggtca aaggagaggg ccctgtccca atccagatga ctccacttct gctggaccca 10500

ggttcacaag cttaatctac atttcaccta aatttggcta acaagcccaa aatcacacag 10560

gcaaagggag aagtggaggc agaaccgagg ttggaggcca ccagggccac cgggcagaga 10620

tcatttaagc ccaaccttct cacttctccc tgggctctgc ctctcttaaa ggaccttgtg 10680

gtgtgacctc ttgtaggtcc ctttcacact cggggcctca gtttccccac tgtaaagtga 10740

atgggtccca gctttggtaa gcttatgctt acctgatgct ttcttcctgg gctgctcttg 10800

tagagaaaag ataaatcttc ttcctccatc cacgagggct tctttccctg ggggtgagag 10860

taggctgagg agagccactt gcacacactc ttaaagaaag tattacctgc accagctcag 10920

tgagaggcac agatcagact gttacttgaa tcaaattatg agcctcccca aatatatcta 10980

tgacatttaa ataggggatt acttgaacat agactttggg atccggtgtg gagtgcagga 11040

gactagcaaa gtgaatcctg agagtagcag gtctgcacct gttggatcga gaaaggcggc 11100

ctacaattct ggtcaaatga gctgtgctta ttgacatatt ctattagaga gtactaccag 11160

gtcaccagtc accagaaagg ctgccagctc tccaaccacc tccagggaac tatcctgaat 11220

ggggccttaa caagcctaag agagggttgg tttgggtccc aagccaatat ttgctctgct 11280

ttatgtcagt catatggaac ccaaaccaac cctctcctat gtgcctcacc agtcggtgca 11340

gggatcccaa tttcaagttt ggtttttat ggtcaaagtc cagcatagat taaatgaagg 11400

ggtgtgatga tggtgttaaa agagaactcc agaccagttt aactcttgga cacacatccc 11460

atctcaccat ggtgcttcca accttccaga gatgatgggc tcctattttc tgatgacaaa 11520

gccctccaca ggattgctgc ctggccatca gggagtgcct ctgtaactga ggctgagatc 11580

ccactttcag tcctccagct gtggcccatc cctgctccgc ccaccgggta tggcctgtcc 11640

taggctctta ggtatggctg cattgtgaaa tgatggctac agagctggca tctcctgtag 11700

tctggttcat ctagtgcact acctcatagt taaaagaaat ctgtttaagc cactgagggt 11760

ggctcctagt gccaactcca agaacaggaa gcttcccttt tttgggagga ggggcagatg 11820

gtaacatgga tcgtccaggt caatgggagc agggcaacca cagtaagtac tggacaacaa 11880

cacaaaactc catgtgtggc ttccatcgag tccctctcca attggtttgg tcttctccgt 11940

cccatgcagc actttagcaa ggggcctggc tgaaggctat gaattgtgtg gagcctcctc 12000

attgcagtct ccaaccatct gatgctggga aaatgtcacc aggatgcagc catgccgtgt 12060

ggccaatgaa ccgagaaaac acccctttc tagaatgctc taaagaggca gaataatcca 12120

gaggtgagga aggaaatact ccaccagaga cccaggcagt tcctacaaaa gccagacttt 12180

ccttcaccta gggagtgaca agaccagtgg aaaacactct caagcagtaa cccccaaatg 12240

ctctgcaagc cagtggcgtc cagataccgc acaagcgagt gggctgtcta atcccatcat 12300

catgatgtaa atatctctag gctgccccgg gctgtgcctg accctgtctt cagctttcca 12360

cacctccacc tacagcccat gcacagaagg accacccagg aatgctgcaa gtgtggcacc 12420

tccagggcca cccagggaga aggagggcag ctatgctggt ggctccaggc ccatttggcg 12480

ggtggtacct tcacaccaca aagcccaaac tgaggcccca gatttggctg atgagggcat 12540

attggacagg ggtcacttat gctcttcccc attgccacct ggcctctggc tacctggact 12600

tggctacctg tggatcctct cacaggtgcc accatcttgg ctgagtctcc agatgcgagg 12660

tccctgaggc agtggcaggc ttctcgctaa tgctgatggg attaggaatg ggataggtgg 12720

ggagggccct ggactgggcc ctgatgagcc aagtgggttt ttagaggggc tactggtaca 12780

tttcagggac aggacatctg gtagagctaa gctggggcaa taaggagcca ctgctaatct 12840

gagagctaga aacaatcagc ttctgggtca ttattaatta gggtagtttg ggctgtgtgg 12900

aagtcacgta ctatatgggg tagccacagc tctctctaca gataatctct aagacttctg 12960

attgggaccg tgtgaatgca gtagcaatat ctcttcttac tgccaggccc tgccagtcct 13020

gcctccacgc cctggctggc ccccttatg atctgaccca tgccaggctg ccatagtatg 13080

ttacttctgc attagcactc cttgggacct gcctctccac tgtccctcag actttaaaga 13140

actatacaaa cccaaggggc tcttcccaag agaattgata tgacttgagg tgattccatt 13200

tctggaagta gtcactccat tttctgcctc actctttcag tgcttcacag agcaggttcg 13260

attaagcaca cagattaatc gtaactataa cggtcctaag gtagcgaaag cttgcatgcc 13320

tgcaggtcga ctctagagga tcaaacctag ccaccgcggt ggcggccggc tagccggcta 13380

gccggctagc cctagaacta gtaacggccg ccagtgtgct ggaattcggc ttgtaaggta 13440

ccggtgaagt tcctatactt tctagagaat aggaacttcg gaataggaac ttctacctag 13500

atgcatgctc gagcggcccc tacagttgaa gtcggaagtt tacatacact taagttggag 13560

tcattaaaac tcgtttttca actactccac aaatttcttg ttaacaaaca atagttttgg 13620

caagtcagtt aggacatcta ctttgtgcat gacacaagtc atttttccaa caattgttta 13680

cagacagatt atttcactta taattcactg tatcacaatt ccagtgggtc agaagtttac 13740

atacactaag ttgactgtgc ctttaaacag cttggaaaat tccagaaaat gatgtcatgg 13800

ctttagaagc ttctgataga ctaattgaca tcatttgagt caattggagg tgtacctgtg 13860

gatgtatttc aaggaattct gtggaatgtg tgtcagttag ggtgtggaaa gtccccaggc 13920

tccccagcag gcagaagtat gcaaagcatg catatcgata ctagtttata agatctcgag 13980

ctagggtacc gtcaaggctg cagtgagaca tgatcttgcc actgcactcc agcctggaca 14040

gcagagtgaa accttgcctc acgaaacaga atacaaaaac aaacaaacaa aaaactgctc 14100

cgcaatgcgc ttccttgatg ctctaccaca taggtctggg tactttgtac acattatctc 14160

attgctgttc ataattgtta gattaatttt gtaatattga tattattcct agaaagctga 14220

ggcctcaaga tgataacttt tattttctgg acttgtaata gctttctctt gtattcacca 14280

tgttgtaact ttcttagagt agtaacaata taaagttatt gtgagttttt gcaaacacag 14340

caaacacaac gacccatata gacattgatg tgaaattgtc tattgtcaat ttatgggaaa 14400

acaagtatgt actttttcta ctaagccatt gaaacaggaa taacagaaca agattgaaag 14460

aatacatttt ccgaaattac ttgagtatta tacaaagaca agcacgtgga cctgggagga 14520

gggttattgt ccatgactgg tgtgtggaga caaatgcagg tttataatag atgggatggc 14580

atctagcgca atgactttgc catcactttt agagagctct tggggacccc agtacacaag 14640

aggggacgca gggtatatgt agacatctca ttctttttct tagtgtgaga ataagaatag 14700

ccatgacctg agtttataga caatgagccc ttttctctct cccactcagc agctatgaga 14760

tggcttgccc tgcctctcta ctaggctgac tcactccaag gcccagcaat gggcagggct 14820

ctgtcagggc tttgatagca ctatctgcag agccagggcc gagaaggggt ggactccaga 14880

gactctccct cccattcccg agcagggttt gcttatttat gcatttaaat gatatattta 14940

ttttaaaaga aataacagga gactgcccag ccctggctgt gacatggaaa ctatgtagaa 15000

tattttgggt tccatttttt tttccttctt tcagttagag gaaaaggggc tcaggatcca 15060

cttgcccagt gttcttcctt agttcctacc ttcgaccttg atcctccttt atcttcctga 15120

accctgctga gatgatctat gtggggagaa tggcttcttt gagaaacatc ttcttcgtta 15180

gtggcctgcc cctcattccc actttaatat ccagaatcac tataagaaga atataataag 15240

aggaataact cttattatag gtaagggaaa attaagaggc atacgtgatg ggatgagtaa 15300

gagaggagag ggaaggatta atggacgata aaatctacta ctatttgttg agacctttta 15360

tagtctaatc aattttgcta ttgttttcca tcctcacgct aactccataa aaaaacacta 15420

ttattatctt tattttgcca tgacaagact gagctcagaa gagtcaagca tttgcctaag 15480

gtcggacatg tcagaggcag tgccagacct atgtgagact ctgcagctac tgctcatggg 15540

ccctgtgctg cactgatgag gaggatcaga tggatggggc aatgaagcaa aggaatcatt 15600

ctgtggataa aggagacagc catgaagaag tctatgactg taaatttggg agcaggagtc 15660

tctaaggact tggatttcaa ggaattttga ctcagcaaac acaagaccct cacggtgact 15720

ttgcgagctg gtgtgccaga tgtgtctatc agaggttcca gggagggtgg ggtggggtca 15780

gggctggcca ccagctatca gggcccagat gggttatagg ctggcaggct cagataggtg 15840

gttaggtcag gttggtggtg ctgggtggag tccatgactc ccaggagcca ggagagatag 15900

accatgagta gagggcagac atgggaaagg tgggggaggc acagcatagc agcatttttc 15960

attctactac tacatgggac tgctccccta tacccccagc taggggcaag tgccttgact 16020

cctatgtttt caggatcatc atctataaag taagagtaat aattgtgtct atctcatagg 16080

gttattatga ggatcaaagg agatgcacac tctctggacc agtggcctaa cagttcagga 16140

cagagctatg ggcttcctat gtatgggtca gtggtctcaa tgtagcaggc aagttccaga 16200

agatagcatc aaccactgtt agagatatac tgccagtctc agagcctgat gttaatttag 16260

caatgggctg ggaccctcct ccagtagaac cttctaacca gctgctgcag tcaaagtcga 16320

atgcagctgg ttagactttt tttaatgaaa gcttgcatgc agcactttgg gaggctgagg 16380

tgggtggact gcttggagct caggagttca agaccatctt ggacaacatg gtgatаccct 16440

gcctctacaa aaagtacaaa aattagcctg gcatggtggt gtgcacctgt aatcccagct 16500

attagggtgg ctgaggcagg agaattgctt gaacccggga ggcggaggtt gcagtgagct 16560

gagatcgtgc cactgcactc cagcctgggg gacagagcac attataatta actgttattt 16620

tttacttgga ctcttgtggg gaataagata catgttttat tcttatttat gattcaagca 16680
ctgaaaatag tgtttagcat ccagcaggtg cttcaaaacc atttgctgaa tgattactat 16740
actttttaca agctcagctc cctctatccc ttccagcatc ctcatctctg attaaataag 16800
cttcagtttt tccttagttc ctgttacatt tctgtgtgtc tccattagtg acctcccata 16860
gtccaagcat gagcagttct ggccaggccc ctgtcggggt cagtgcccca cccccgcctt 16920
ctggttctgt gtaaccttct aagcaaacct tctggctcaa gcacagcaat gctgagtcat 16980
gatgagtcat gctgaggctt agggtgtgtg cccagatgtt ctcagcctag agtgatgact 17040

cctatctggg tccccagcag gatgcttaca gggcagatgg caaaaaaaag gagaagctga 17100
ccacctgact aaaactccac ctcaaacggc atcataaaga aaatggatgc ctgagacaga 17160
atgtgacata ttctagaata tattatttcc tgaatatata tatatatata cacatatacc 17220
atatgaaaca cctctaggct ataaggcaac agagctcctt ttttttttt ctgtgctttc 17280
ctggctgtcc aaatctctaa tgataagcat acttctattc aatgagaata ttctgtaaga 17340
ttatagttaa gaattgtggg agccattccg tctcttatag ttaaatttga gcttctttta 17400
tgatcactgt ttttttaata tgctttaagt tctggggtac atgtgccatg gtggtttgct 17460

gcacccatca acccgtcatc tacattaggt atttctccta atgctatcct tcccctagcc 17520
ccccaccccc aacaggcccc agtgtgtgat gttcccctcc ctgtgtccat ggatcactgg 17580
ttttttttg ttttttttt ttttttaaag tctcagttaa atttttggaa tgtaatttat 17640
tttcctggta tcctaaggac ttgcaagtta tctggtcact ttagccctca cgttttgatg 17700
ataatcacat atttgtaaac acaacacaca cacacacaca cacacacata tatatatata 17760
taaaacatat atatacataa acacacataa catatttatc gggcatttct gagcaactaa 17820
tcatgcagga ctctcaaaca ctaacctata gccttttcta tgtatctact tgtgtagaaa 17880

ccaagcgtgg ggactgagaa ggcaatagca ggagcattct gactctcact gcctttagct 17940
aggcccctcc ctcatcacag ctcagcatag tcctgagctc ttatctatat ccacacacag 18000
tttctgacgc tgcccagcta tcaccatccc aagtctaaag aaaaaaataa tgggtttgcc 18060
catctctgtt gattagaaaa caaaacaaaa taaaataagc ccctaagctc ccagaaaaca 18120
tgactaaacc agcaagaaga agaaaataca ataggtatat gaggagactg gtgacactaa 18180
gtgtctgaat gaggcttgag tacagaaaag aggctctagc agcatagtgg tttagaggag 18240
atgtttcttt ccttcacaga tgccttagcc tcaataagct tgcggttgtg gaagtttact 18300

tgtttatcac cggtgacgtc catgagcaaa ttaagaaaaa caacaacaaa tgaatgcata 18360

tatatgtata tgtatgtgtg tatatataca catatatata tatatttttt ttcttttctt 18420

accagaaggt tttaatccaa ataaggagaa gatatgctta gaactgaggt agagttttca 18480

tccattctgt cctgtaagta ttttgcatat tctggagacg caggaagaga tccatctaca 18540

tatcccaaag ctgaattatg gtagacaaaa ctcttccact tttagtgcat caatttctta 18600

tttgtgtaat aagaaaattg ggaaaacgat cttcaatatg cttaccaagc tgtgattcca 18660

aatattacgt aaatacactt gcaaaggagg atgtttttag tagcaatttg tactgatggt 18720

atggggccaa gagatatatc ttagagggag ggctgagggt ttgaagtcca actcctaagc 18780

cagtgccaga agagccaagg acaggtacgg ctgtcatcac ttagacctca ccctgtggag 18840

ccacacccta gggttggcca atctactccc aggagcaggg agggcaggag ccagggctgg 18900

gcataaaagt cagggcagag ccatctattg cttacatttg cttctgacac aactgtgttc 18960

actagcaacc tcaaacagac accatgggtc atttcacaga ggaggacaag gctactatca 19020

caagcctgtg gggcaaggtg aatgtggaag atgctggagg agaaaccctg ggaaggtagg 19080

ctctggtgac caggacaagg gagggaagga aggaccctgt gcctggcaaa agtccaggtc 19140

gcttctcagg atttgtggca ccttctgact gtcaaactgt tcttgtcaat ctcacaggct 19200

cctggttgtc tacccatgga cccagaggtt ctttgacagc tttggcaacc tgtcctctgc 19260

ctctgccatc atgggcaacc ccaaagtcaa ggcacatggc aagaaggtgc tgacttcctt 19320

gggagatgcc ataaagcacc tggatgatct caagggcacc tttgcccagc tgagtgaact 19380

gcactgtgac aagctgcatg tggatcctga gaacttcaag gtgagtccag gagatgtttc 19440

agccctgttg cctttagtct cgaggcaact tagacaactg agtattgatc tgagcacagc 19500

agggtgtgag ctgtttgaag atactggggt tgggggtgaa gaaactgcag aggactaact 19560

gggctgagac ccagtggtaa tgttttaggg cctaaggagc gcctctaaaa atctagatgg 19620

acaattttga ctttgagaaa agagaggtgg aaatgaggaa aatgactttt attagattcc 19680

agtagaaaga actttcatct ttccctcatt tttgttcgtt ttaaaacatc tatctggagg 19740

caggacaagt atggtcgtta aaaagatgca ggcagaaggc atatattggc tcagtcaaag 19800

tggggaactt tggtggccaa acatacattg ctaaggctat tcctatatca gctggacaca 19860

tataaaatgc tgctaatgct tcattacaaa cttatatcct ttaattccag atgggggcaa 19920

agtatgtcca ggggtgagga acaattgaaa catttgggct ggagtagatt ttgaaagtca 19980

gctctgtgtg tgtgtgtgtg tgcgcgcgcg cgtgtgtgtg tgtgtgtgtc aacgtgtgtt 20040

tcttttaacg tcttcagcct acaacataca gggttcatgg tggcaagaag atagcaagat 20100

ttaaattatg gccagtgact agtgcttgaa ggggaacaac tacctgcatt taatgggaag 20160

gcaaaatctc aggctttgag ggaagttaac ataggcttga ttctgggtag aagctgggtg 20220

tgtagttatc tggaggccag gctggagctc tcagctcact atgggttcat ctttattgtc 20280

tcctttcatc tcaacagctc ctgggaaatg tgctggtgac cgttttggca atccatttcg 20340

gcaaagaatt cacccctgag gtgcaggctt cctggcagaa gatggtgact gcagtggcca 20400

gtgccctgtc ctccagatac cactgagcct cttgcccatg attcagagct ttcaaggata 20460

ggctttattc tgcaagcaat tcaaataata aatctattct gctgagagat cacacatgat 20520

tttcttcagc tctttttttt acatcttttt aaatatatga gccacaaagg gtttatattg 20580

agggaagtgt gtatgtgtat ttctgcatgc ctgtttgtgt ttgtggtgtg tgcatgctcc 20640

tcatttattt ttatatgaga tgtgcatttt gttgagcaaa taaaagcagt aaagacactt 20700

gtacacggga gttctgcaag tgggagtaaa tggtgtagga gaaatccggt gggaagaaag 20760

acctctatag gacaggactt ctcagaaaca gatgttttgg aagagatggg aaaaggttca 20820

gtgaagacct gggggctgga ttgattgcag ctgagtagca aggatggttc ttaatgaagg 20880

gaaagtgttc caagctttag gaattcaagg tttagtcagg tgtagcaatt ctattttatt 20940

aggaggaata ctatttctaa tggcacttag cttttcacag cccttgtgga tgcctaagaa 21000

agtgaaatta atcccatgcc ctcaagtgtg cagattggtc acagcatttc aagggagaga 21060

cctcattgta agactctggg ggaggtgggg acttaggtgt aagaaatgaa tcagcagagg 21120

ctcacaagtc agcatgagca tgttatgtct gagaaacaga ccagcactgt gagatcaaaa 21180

tgtagtggga agaatttgta caacattaat tggaaggctt acttaatgga atttttgtat 21240

agttggatgt tagtgcatct ctataagtaa gagtttaata tgatggtgtt acggacctaa 21300

tgtttgtgtc tcctcaaaat tcacatgctg aatccccaac tcccaactga ccttatctgt 21360

gggggaggct tttgaaaagt aattaggttt agatgagctc ataagagcag atccccatca 21420

taaaattatt ttccttatca gaagcagaga gacaagccat ttctctttcc tcccggtgag 21480

gacacagtga gaagtccgcc atctgcaatc caggaagaga accctgacca cgagtcagcc 21540

ttcagaaatg tgagaaaaaa ctctgttgtt gaagccaccc agtcttttgt attttgttat 21600

agcaccttgc actgagtaag gcagatgaag aaggagaaaa aaataagctt atcgaaacgc 21660

gtccccatcc tcactgactc cgtcctggag ttggatgaga gataatggcc ttacgttgtg 21720

ccaggggagg gtcgggctgg atttagcaag atttaccttc tccaaagagc ggtgctgcag 21780

tggcacagct gcccacggag gtgggggggt caccgtccct ggaggtgatg aagaactgtg 21840

gggatgtggc actgagggac atggccagtg ggcacggtgg gtgggttggg gttggtcttg 21900

gggatcttgg agggcttttc cagccttcat gatttgacga ttgtatgaac atctacatgg 21960

caattctcca gctgcctgtc ccagtcctac tgacccagct gtatctctcc aggcaagctc 22020

ttccacccct tctgcttgca tccagacacc atcaaacatg caggctcaga cacagggacc 22080

agcagtgtct gtggcctttt tgtgctcctc tccatgctgg gttttaactt gctctttgtc 22140
cttctatcct atcttcttat ccttaaggct gttctgaacg ctgtgacttg gagagtgtcc 22200
cagagccctc aacacctgca tgtcccacgt ccatgctgtc ctgcacttcc ttatccccaa 22260
gatctgcctc tccgtgatgc actgaattgg caaacatgtg tcaccccaga ccaacaatgt 22320
cacagcaaac tccccttga taggacaagg gggaatggct ttacactgag acaggggagg 22380
tttgggttgg atatgaggag gcagtttttc ccccagaggg tggtgacgca ctgaacaggt 22440
tgcccaagga ggctgtggat gccccatccc tgcaggcatt caaggccagg ctggatgtgg 22500

ctctgggcag cctgggctgc tggttgatga ccctgcacat agcaggggggt tggatctgga 22560
tgagcactgt gctcctttgc aacccaggcc gttctatgat tctgtcattc taaatctctc 22620
tttcagccta aagctttttc cccgtatccc cccaggtgtc tgcaggctca aagagcagcg 22680
agaagcgttc agaggaaagc gatcccgtgc caccttcccc gtgcccgggc tgtccccgca 22740
cgctgccggc tcggggatgc gggggggagcg ccggaccgga gcggagcccc gggcggctcg 22800
ctgctgcccc ctagcggggg agggacgtaa ttacatccct gggggctttg gggggggggct 22860
gtccctatcg attactagtt tagccataga gcccaccgca tccccagcat gcctgctatt 22920

gtcttcccaa tcctcccccct tgctgtcctg ccccacccca ccccccagaa tagaatgaca 22980
cctactcaga caatgcgatg caatttcctc attttattag gaaaggacag tgggagtggc 23040
accttccagg gtcaaggaag gcacgggggga ggggcaaaca acagatggct ggcaactaga 23100
aggcacagtc gctcgaagag cggccgctcg cttcagtttc ggccagcagg cggggagccc 23160
gaggtagctc ccgctccctt gagccaggcc cctgccagac ctgagctccc tcccaagcct 23220
ggcttcccca accggtggcc ttcatgggcc agaagccatt ccttcacggc tagccctccg 23280
gagtagttgc ccacggctcc gctgctgcag accactctgt ggcacgggat gaggatcttg 23340

acaggattgc ctctcatggc gcctcccact gctcgcgcgg ctttggggtt gccggccagg 23400
gcggccaatt gctggtaaga aatcacttct ccgaatttca caaccttaag cagcttccat 23460
aacacctgac gcgtgaacga ctcttgctgg aaaacgggat ggtgaagcgc tggcacgggg 23520
aactcttcga tagcctcggg ctggtggaaa taggcattca gccaggctgt gcactgcatc 23580
aggggctccg gacctccgag aaccgcagcg ggggctggga cctccacggc atcagctgca 23640
gacgtcccct tgcccaggag ctttatttcg tgcagaccct gctcacaacc agacagctcc 23700
agcttcccca aagggctgtc cagtgtggtg cgtttcattt cacaatcctt gtccatggtg 23760

gcgaccgtct agcccctggg gagagaggtc ggtgattcgg tcaacgaggg agccgactgc 23820

cgacgtgcgc tccggaggct tgcagaatgc ggaacaccgc gcgggcagga acagggccca 23880

cactaccgcc ccacaccccg cctcccgcac cgcccccttcc cggccgctgc tctcggcgcg 23940

ccccgctgag cagccgctat tggccacagc ccatcgcggt cggcgcgctg ccattgctcc 24000

ctggcgctgt ccgtctgcga gggtactagt gagacgtgcg gcttccgttt gtcacgtccg 24060

gcacgccgcg aaccgcaagg aaccttcccg acttaggggc ggagcaggaa gcgtcgccgg 24120

ggggcccaca agggtagcgg cgaagatccg ggtgacgctg cgaacggacg tgaagaatgt 24180

gcgagaccca gggtcggcgc cgctgcgttt cccggaacca cgcccagagc agccgcgtcc 24240

ctgcgcaaac ccagggctgc cttggaaaag gcgcaacccc aaccccgtgg aaataaatcg 24300

ataactagtg atatcatcat gtctggatcc catcacaaag ctctgacctc aatcctatag 24360

aaaggaggaa tgagccaaaa ttcacccaac ttattgtggg aagcttgtgg aaggctactc 24420

gaaatgtttg acccaagtta aacaatttaa aggcaatgct accaaatact aattgagtgt 24480

atgttaactt ctgacccact gggaatgtga tgaaagaaat aaaagctgaa atgaatcatt 24540

ctctctacta ttattctgat atttcacatt cttaaaataa agtggtgatc ctaactgacc 24600

ttaagacagg gaatctttac tcggattaaa tgtcaggaat tgtgaaaaaa gtgagtttaa 24660

atgtatttgg ctaaggtgta tgtaaacttc cgacttcaac tgtaggggat cctctagggc 24720

cgccagtgtg atggatatct gcagaattcg gcttcaggta ccgtcgacga tgtaggtcac 24780

ggtctcgaag ccgcggtgcg ggtgccaggg cgtgcccttg ggctccccgg gcgcgtactc 24840

cacctcaccc atctggtcca tcatgatgaa cgggtcgagg tggcggtagt tgatcccggc 24900

gaacgcgcgg cgcaccggga agccctcgcc ctcgaaaccg ctgggcgcgg tggtcacggt 24960

gagcacggga cgtgcgacgg cgtcggcggg tgcggatacg cggggcagcg tcagcgggtt 25020

ctcgacggtc acggcgggca tgtcgacaag ccgaattcca gcacactggc ggccgttact 25080

aggtagctag ctcgagcctt cgaagatctc ctagggaagt tcctatactt tctagagaat 25140

aggaacttcg gaataggaac ttcaccggtg ggtgaaaagc cgaattctgc agatatcaag 25200

cttatcgata ccgtcgacct cgaggggggg cccggttaga tccccgggta ccgagctcga 25260

atatctatgt cgggtgcgga gaaagaggta atgaaatgga ttaagtggca ggattaatct 25320

tcgaacgaag gagccatcca actaaccgtc atgttcgggc aaccgaagaa gggagtggca 25380

ggatttcctt tggagacttc tggaattaga cagcagttta atgcaagcat ctaaattctc 25440

ttcctcccag agtctcatta aaactacagt aagagtttgt gttttgtttt gtttttaaag 25500

acaaaatccc accaggatag agagaatagg agaggagata acagcatcat aatttatgaa 25560

actaaaatgc agatagacca ggattaactg actacacagc accaaggaag ctgaatcaca 25620

agacagcaga ggagaaaact ggaaaggatc gtggtctata cggcagaatc ttcccaagcc 25680

tcaggaggag gagctctaga tgcttatgat ggcaactaaa gcctaaaagc taattcattt 25740

taaagttctt ccaaatgcat agggttttat ttttccagac ctgggttcag atggggaatt 25800

tgacaaacaa tggaaagggg gaaaaacaac aatctaaaca ctgagtgaca aagtaacaaa 25860

gaaatagtct agctatcagc cagtcaagcc agccttggct ttgctatcca aagtagtcag 25920

tctaattcta ccaccagttt ctgttcctgt agctgtctac tgcctgccag ggactctgcc 25980

ttcccaccca caactaccaa tggaaggatg tggtgaccat accagtggct gctgacatct 26040

cctgccatgg gaagcataat tgcctccagc agcctccccc ttagatccat cattttgtt 26100

gcacttggcc tgggctgtac tcccggccaa tgactgaaca tggtgagcat agtaatgcag 26160

gcccatttct gtgaggagca ggactcctcc agtaggtgac tttggctcaa ggactctcta 26220

ttggcctggt tgaacttttc ctgaactgtg ctactgtctg agactcttct tacccaatcc 26280

tctttctcgc cccaattgtc acagaccacc tgcattgtgg tctgagtctc tccccacctt 26340

ctcttgctct tccctgttta tctttcacag gcatttcccc cagtacattc cttgaatgtc 26400

taacccgata cgggtgcctg acttttggca gacctaagca gacaaaaagg agtacttggt 26460

tacctagctc ttctttctac cacaaacatc gagggaaccc tttttccctc acccctctgc 26520

cacaccccca ctgccccagt gaacaaccac agagagagct gtggtataat attaggctgg 26580

tgcaaaagta attgcggttt ttgccattac ttttaatggt aaaaaccgca attacttttg 26640

cacctaccta gtatttgtgt ccccccaaat tcatatgttg aaacctaacc cacaatatga 26700

tgtcattagg aggcaagacc ttgaggaggt gattagatga tggggtggag ctctcctgaa 26760

tgagattagt gcccttataa gaagaagccc aaggaagcta ccttgactct tccatcacat 26820

gagaatgcag caagaaggca ccatctacta atcaggaaga gagctctcac cagacactga 26880

atctgccagt gtcttgatct tgaagttccc agcctccaga actatgcata atgcatttcc 26940

attgtctcta agccacccag cctatggtat tttgtcatag cagcctgaac tgactaagac 27000

agtgagccac atgagaagtg ccccaacccc tcccttaagc acttggctca cagatcagtg 27060

ggttcatttc tgcctgagtt ttattgttat tctgtagatt tcttgggcta gatatatttt 27120

tctgttattt tccttcttca cctcagtcat gaattggttg ttttaaaaaa gacaatgtaa 27180

gtcatgggga aactcctgac aactctactc tcctagggtt cctgataaaa ggggattcag 27240

ttgagtcctc tgatggtctc tacctgccaa agtccagcag cccttagcaa acatgctgct 27300

cgtttctgta gagaaggtgc tggtgtccca ccatacttct ctctccctca tgaagggctt 27360

gcgacccagc aaatgggtgg cttatatggg tctgtttcaa aggaagagcc agctctggga 27420

agaaaaacga tgagcataag cataacctac cactgtgcct gggaaagcag acaacttttt 27480

tgatgtgtga atatctaatg agaatggaat ccatcaatta ccttaaactt aggcacagtc 27540

ttcaaattca atatatgtgg gatatacttt tagtcagttt gtagacgtta tttgtaataa 27600
ataatctggc ttctctaaag aaattatttt aagtgtttgg tttggtttga tttaatggta 27660
aaattatatt tagtggcaga gaattatagc aatggtgata aactatagag tgtcataagt 27720
tcatatctta ttctcacatt tgaagctgcc tgcagatgca ttcaagatgc agccagaagt 27780
caggagactc aggctgttat ttggagctca tcattttaca gccttgctgg actcccactt 27840
tctcagggga aaaatgtggt gttgacccag attagctctc caggccctgc tgagttgggc 27900
actctgtaag ctggagggtc ttctattgtc ttcacctaag tgtcaatcaa caacccaaat 27960

gggcatgggg gaagagggag ctgggccaat gcccagggtg cctggtagag agataccttg 28020
ggcactggaa ggcaccagct tcccagagag aaggggggagg gccatgaaaa agttggctgt 28080
agatgccagg gacactggga ctctccagct gtgtgtttgt gtcttctgaa gacttatgtt 28140
tcattccttt ggagcatgca taatcataca ctgtgggatg tgttatatag attgcttgat 28200
agttcaccac tgtaataaaa tactgtgact ggaatctgct cccagtctgc ctttgatagc 28260
acttgtgcaa cacacattta ctgagcattt acagtgatcc aggacctgtg ttgtgaaaac 28320
attgatggac aaggcagatg gtggagcacg tcagtgagga tttttaacaa aggctggtaa 28380

gtgctataaa ggaacattgt aggacactag agaacaaaga acaggagaac ctgacttagg 28440
ctggggtggg gcgttggtta gaggaggctc cttggaggac atgaggttta agctgtgacc 28500
tgaggatgaa tagatgttgg ccaggtgagg ttaatctggg gctcaagatc gagcattaag 28560
cttgtcagcc ttaccagtaa aaaagaaaac ctattaaaaa aacaccactc gacacggcac 28620
cagctcaatc agtcacagtg taaaaaaggg ccaagtgcag agcgagtata tataggacta 28680
aaaaatgacg taacggttaa agtccacaaa aaacacccag aaaaccgcac gcgaacctac 28740
gcccagaaac gaaagccaaa aaacccacaa cttcctcaaa tcgtcacttc cgttttccca 28800

cgttacgtaa cttcccattt taagaaaact acaattccca acacatacaa gttactccgc 28860
cctaaaacct acgtcacccg ccccgttccc acgccccgcg ccacgtcaca aactccaccc 28920
cctcattatc atattggctt caatccaaaa taaggtatat tattgatgat gtttaaacta 28980
cggcccggta cccagctttt gttcccttta gtgagggtta attgcgcgct tggcgtaatc 29040
atggtcatag ctgtttcctg tgtgaaattg ttatccgctc acaattccac acaacatacg 29100
agccggaagc ataaagtgta aagcctgggg tgcctaatga gtgagctaac tcacattaat 29160
tgcgttgcgc tcactgcccg ctttccagtc gggaaacctg tcgtgccagc tgcattaatg 29220

aatcggccaa cgcgcgggga gaggcggttt gcgtattggg cgctcttccg cttcctcgct 29280

cactgactcg ctgcgctcgg tcgttcggct gcggcgagcg gtatcagctc actcaaaggc 29340

ggtaatacgg ttatccacag aatcagggga taacgcagga aagaacatgt gagcaaaagg 29400

ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg gcgtttttcc ataggctccg 29460

ccccccctgac gagcatcaca aaaatcgacg ctcaagtcag aggtggcgaa acccgacagg 29520

actataaaga taccaggcgt ttcccccctgg aagctccctc gtgcgctctc ctgttccgac 29580

cctgccgctt accggatacc tgtccgcctt tctcccttcg ggaagcgtgg cgctttctca 29640

tagctcacgc tgtaggtatc tcagttcggt gtaggtcgtt cgctccaagc tgggctgtgt 29700

gcacgaaccc cccgttcagc ccgaccgctg cgccttatcc ggtaactatc gtcttgagtc 29760

caacccggta agacacgact tatcgccact ggcagcagcc actggtaaca ggattagcag 29820

agcgaggtat gtaggcggtg ctacagagtt cttgaagtgg tggcctaact acggctacac 29880

tagaaggaca gtatttggta tctgcgctct gctgaagcca gttaccttcg gaaaaagagt 29940

tggtagctct tgatccggca aacaaaccac cgctggtagc ggtggttttt ttgtttgcaa 30000

gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat cctttgatct tttctacggg 30060

gtctgacgct cagtggaacg aaaactcacg ttaagggatt ttggtcatga gattatcaaa 30120

aaggatcttc acctagatcc ttttaaatta aaaatgaagt tttaaatcaa tctaaagtat 30180

atatgagtaa acttggtctg acagttacca atgcttaatc agtgaggcac ctatctcagc 30240

gatctgtcta tttcgttcat ccatagttgc ctgactcccc gtcgtgtaga taactacgat 30300

acgggagggc ttaccatctg gccccagtgc tgcaatgata ccgcgagacc cacgctcacc 30360

ggctccagat ttatcagcaa taaaccagcc agccggaagg gccgagcgca gaagtggtcc 30420

tgcaacttta tccgcctcca tccagtctat taattgttgc cgggaagcta gagtaagtag 30480

ttcgccagtt aatagtttgc gcaacgttgt tgccattgct acaggcatcg tggtgtcacg 30540

ctcgtcgttt ggtatggctt cattcagctc cggttcccaa cgatcaaggc gagttacatg 30600

atcccccatg ttgtgcaaaa aagcggttag ctccttcggt cctccgatcg ttgtcagaag 30660

taagttggcc gcagtgttat cactcatggt tatggcagca ctgcataatt ctcttactgt 30720

catgccatcc gtaagatgct tttctgtgac tggtgagtac tcaaccaagt cattctgaga 30780

atagtgtatg cggcgaccga gttgctcttg cccggcgtca atacgggata ataccgcgcc 30840

acatagcaga actttaaaag tgctcatcat tggaaaacgt tcttcggggc gaaaactctc 30900

aaggatctta ccgctgttga gatccagttc gatgtaaccc actcgtgcac ccaactgatc 30960

ttcagcatct tttactttca ccagcgtttc tgggtgagca aaaacaggaa ggcaaaatgc 31020

cgcaaaaaag ggaataaggg cgacacggaa atgttgaata ctcatactct tccttttca 31080

atattattga agcatttatc agggttattg tctcatgagc ggatacatat ttgaatgtat 31140

ttagaaaaat aaacaaatag gggttccgcg cacatttccc cgaaaagtgc cacctgacgc 31200

gccctgtagc ggcgcattaa gcgcggcggg tgtggtggtt acgcgcagcg tgaccgctac 31260

acttgccagc gccctagcgc ccgctccttt cgctttcttc ccttcctttc tcgccacgtt 31320

cgccggcttt ccccgtcaag ctctaaatcg ggggctccct ttagggttcc gatttagtgc 31380

tttacggcac ctcgacccca aaaaacttga ttagggtgat ggttcacgta gtgggccatc 31440

gccctgatag acggtttttc gccctttgac gttggagtcc acgttcttta atagtggact 31500

cttgttccaa actggaacaa cactcaaccc tatctcggtc tattcttttg atttataagg 31560

gattttgccg atttcggcct attggttaaa aaatgagctg atttaacaaa aatttaacgc 31620

gaattttaac aaaatattaa cgcttacaat ttccattcgc cattcaggct gcgcaactgt 31680

tgggaagggc gatcggtgcg ggcctcttcg ctattacgcc agctggcgaa agggggatgt 31740

gctgcaaggc gattaagttg ggtaacgcca gggttttccc agtcacgacg ttgtaaaacg 31800

acggccagtg agcgcgcgta atacgactca ctatagggcg aattggagct ccactacgta 31860

gtttaaacat catcaataat ataccttatt ttggattgaa gccaatatga taatgagggg 31920

gtggagtttg tgacgtggcg cggggcgtgg gaacggggcg ggtgacgtag tagtgtggcg 31980

gaagtgtgat gttgcaagtg tggcggaaca catgtaagcg acggatgtgg caaaagtgac 32040

gtttttggtg tgcgccggtg tacacaggaa gtgacaattt tcgcgcggtt ttaggcggat 32100

gttgtagtaa atttgggcgt aaccgagtaa gatttggcca ttttcgcggg aaaactgaat 32160

aagaggaagt gaaatctgaa taattttgtg ttactcatag cgcgtaatat ttgtctaggg 32220

ccgcggggac tttgaccgtt tacgtggaga ctcgcccagg tgtttttctc aggtgttttc 32280

cgcgttccgg gtcaaagttg gcgttttgat tc 32312

<210> 62

<211> 32670

<212> DNA

<213> Artificial Sequence

<220><223> pHCA-sgHBG1-miR

<400> 62

ggccgaagga ttacatgagc ttagaaatgt aattagcata gtgtgtggca tagtgtagat 60

accaaataaa tatgatctct ccttctactc ttgaaaatgc aaacacattc ttggtggtcc 120

taaaatagcc tgtaacatgg tttactcagc agcatttgct attcaaggca gatctgcctt 180

tagtcattgg ctgcgctcct gaacagctgt gtgaaaggct aacttttgta aaccaaatca 240

aaataaaatg cagcaaaaat ttgtcactga aaggaaatcc tcagtatatc cttttatgaa 300

atgaaagatc cctcatccaa acttaacttt tttaaaagtg cgcatttgga gatatagccc 360

tttcttatga atcctaattc aattttggcc ataaacacac gttgatgttc cccaccccaa 420

agcacatagc aacaagagta ggttctatat tgaaaataat gacaatttaa aaacatgtac 480

ttatttcact gtatgtggac agtgtctatg attgcatcat gaagtgtcat ataaccatgt 540

acgtgtacat gagagagaga tagagagaga agtggtaggg tggtggtggt agaggggatg 600

gcgatagtaa tcatggtaat ggtagaggtg atggaggtgg taatgacgga ggtaagggtg 660

gtagtgatga tggtggtggt ggtaatggtg gtggatgtgg tggtggcaat tgggatggtg 720

ggatggtggt agccatggtg atggtggtaa tggtgttgat ttaaagggtg gtggtagtga 780

aggtgagggt agtggtggtg gaggtggtgg tgctggtagc aatagtgatg gtggtgatgg 840

tgttgatgag ggtgttggga tcagggtgag ttcccacagt atatttcatt cttgttgtac 900

cactctgtca acagcaccac tgactgggac agaggaagaa ggcacactct gaatgtgtta 960

ttaacagaaa cctcaaaaca gtctgtctcc ttgtagtcat tcaaaattat ctttttctta 1020

cctggaaaac tgaaactgaa ttaccgggaa aaacacagga gatttttgtt tgttaatatg 1080

ctgccaataa agtaatttta tgtcaaattt aactacagga aagggcaagg cattttctaa 1140

gttccttaga tgtcatgtgg ctaaaaaaaa caaaaggatg gacagcagtt agatactgta 1200

cacttagctg tttgaagcca tatattcaga aagcagatgt tgggagttgg tgtttgagga 1260

ctgatttcct ggaggtattt tatataggcc aagttcattg ttctaaactc taagggcttg 1320

acttgaggga ggaaaagagg caagaacatg tttagttttg ctgacagcat cacatgggca 1380

gccctaaggc tagacaactt tagggcctga agcttattct aggaaagaag cacctacaga 1440

gtggcactgg gctcccctcc actatagaga tgaagtcata tgacagtaaa gggcaggcag 1500

ggctgcctag ggggcccaga actgacactt ccattagaat gagcacaggc cagggagaga 1560

agtggggaac cagagagaag gagctggaat tctagtagga caaacggtaa gtgaacaaca 1620

agaacaagtt aagagtgtgt gcagtattct ttcaaagact gaaaaaatag tgatgtgata 1680

gaatggcagg tggctctgag caggccagga gaaggactgg gggcagagca tcccaggcag 1740

gagggcagca agtgggaagg ccctggggtg gggcttttgg actgttccag tgacgggcag 1800

gcagccagtg tgcctgtcac acaatgcacc agggaagtag tcgtgaattt gcagagggtc 1860

ttgcaggcta tgggaaaggg attggattgt attttgtttg tagggaagcc atcgggggac 1920

ttaagcagag gaaggattgg cttcatctct ttgaaaaagt tctctctgga tgctgatggg 1980

aggagaaatg gaaggaaaag aaacactttt aggggcaaga acttttgaga agggtggaat 2040

tgggagtgtg gagttggggc cagctttggc acaggagggg aagctaaaca cgtggccgca 2100

tgagggcctg taattctacc tgaaatgggt accatttgtt agggtaaaca aatgaaccaa 2160

atgcccagtg atacagacca agtgttggca aacttcttct gtgatggccc aggtagtaaa 2220

tgtctcaggc ttcgcaggcc atgtggtctc tgttgaagct ctgtgtagta gacaatatgt 2280

taatgactgg gcgtgactgt gtgctaataa aagtttattt acaaaaacag cccgtgggct 2340

ggatttagct cacaggctgt agtttgccaa cctctgacct agagcatgaa ctgagcatct 2400

tcttggaggg aaataagttc tttccaagtt gccctcctca cattgcaggg ggccatgtag 2460

gcccattatt cacagaagag tgggtgggca acctttctgg agcagaaaaa cgtaaagatt 2520

tcttccgtag tgcaagtaag gtgaccattt ctaaaccgtg caagtgatcc agcagtccca 2580

aaagttgttt cacttctcat tgtgcgcccg ttctcaggtg ctccgaagct tccagtcctt 2640

tgtagggaca tggatgaaat tggaaatcat cattctcagt aaactatcgc aagaacaaaa 2700

aaccaaacac cgcatattct cactcatagg tgggaattga acaatgagat cacatggaca 2760

cgggaagggg aacatcacat tctggggact gttgtggggt gggggggaggg gggagggata 2820

gcattgggag atatacctaa tgctagatga caagttagtg ggtgcagcgc accagtgtgg 2880

cacatgtata catatgtaac taacctgcac aatgtgcaca tgtaccctaa aacttaaagt 2940

ataataataa aagaaaaaaa aaaagagagg agagaaacat catcccctcc aggataccct 3000

tgggccttgt tcttatagtc ttgtacattg ttgaacaatt tgcatgggct agtggattaa 3060

agcacaccct ccaccctcag gccctcaagg gtctctatga taatacagtc tcaccttcta 3120

ccctttccat caccatccta ggtgctatgg ccaaccttga ggctgccatg ttaggtctat 3180

gcatttccca cctccaccac ataactctct gaaggccagg tagtttccta ttcatcttgg 3240

taaccccaaa gcctcgtgac agggctcagc tggcatctgc ggatgtgaat gaaccattgg 3300

agaaaatggt actctgcaaa taactctgtt attttcccat ttcctgtgta aggcctagag 3360

acaatgactt tttaattgca ccccttcccc tctgtatgac actggccttc tcttgtgtcc 3420

agcaatgtgg gtggcctaga tgatttctaa gggacttctg gccaagatga acagcagctg 3480

catcttactg agcatttact atgtgccata tactcagcca cagctctagg ggcatagaag 3540

caggagctct cagggtcagg gcagtgagtg agcaagcgag cacctatgcc agccctgcct 3600

ctggatgggg acttgagagg gtgatggaag cctgcagcac tggagggagg cagacaaaga 3660

caggcctgtg ctgaggggggc ccggagcaag agagagggag gcaatgacag cagagacatg 3720

cctgcgcctt gggtttgagt gcccagtggt caaatccact tccctgtggc tgatgcttgc 3780

ctttctaact ttggaattta ggggttggag atctggtgag aaggtaggag ggagatgagg 3840

aggagaaggg aaaggcagga aggaagggga gggaaaggaa aagcaaaagg ggaggaggaa 3900

ggtttccaac aaattattct atatcaactg cggaaatcaa aatttgttgc ccaaatctta 3960

gaagctcatg tccctcctcc ccagaagtct ggaatgcagc actccagggg tagcttataa 4020
cccaaatatc tatctgtaaa aagagaaaca ttgggctttc gagctgtgga ttctcagtaa 4080
aagcaagagg cctcagccta cacaggccag cccagagttt gaggaacccc aggcccacac 4140
ccacagggct ggcccctggg tctgcatact ccctagaaat gtgcacactt ctgagcctca 4200
actctgtcct ggagtctaac agcatccctc tccttcctgg ggcagttcca cctccagaaa 4260
cctgttacct tgggccttat gtcaaggaaa ctgtgggaaa gagctaggca ggaatgcaga 4320
tgaggccagc atgggctcct aaaagtttag aaataggcag tgtcatgctc ccaggtgcct 4380

gcataaacca gctgaaaaat ggagctcccc tcaccagcac tctcccttca aacagactgt 4440
gatttgcagg tcactggttt accaagccag gctacccagg caggacccag atgccaagcc 4500
cagtggtgtc ctgcaagctg agcagtgctc agttcttgca aaaaaaggtc tgtgtgaagg 4560
caaggcctct gcctggcttc tcaccccagt tgggtgtctg gaacaggaag gagcccttac 4620
tgcagaaaaa ggaggaggga gcaaagggag cgaacagctg cgtgctccat ggggaggatc 4680
cccaaagtag aaaggcgcat acacactgca gcccttgacc cagaatgctc acagctacat 4740
tacagattca ggtctcctca gtgtagtggg gctgctgatg agactgtggc atcctcaggg 4800

gtcaggacac acattttcca tcactcttct gatggcaaaa aacctctgag ccaatgccaa 4860
cctctgatca ttaaaaaaaa gtgctcacag cagtgtgtgg tttaggatca tgccctgtgt 4920
ggtttggaac acgtgcacaa ccacaccttg ttcatcacca tcccagaaac cctgacgcag 4980
gcaaagagca gagttattaa ccctacttta ctgatgtgga tactgaggcc cagaggctca 5040
tgcaagttat caataagtgg cagggacagt tgcctctaga ttaactagcc cctaggatca 5100
cctgggtctt ggaaggggac ccataaacat gagctcccct ctcttggggc cagatttgca 5160
cctgtgccgc gccttcagcc tgcatgaagt aggggctgct ggcaaagact caaagctgta 5220

aatctgggtt ttctcttgag gcttctaagg gagctgtttc gacaactcac tctgttccca 5280
gctggctgcc cctgcatagg gttttaaagc agcctagctt tctgccaggc ttggcagtgg 5340
acaacgctgg tcagaacatc ccagagagct accagaatga agtaagtttg cttctactct 5400
ttacctgttt atgggctgtc tctgccactg gaatgaaagg cactgagaac agtgcctggc 5460
ctgcagaagg ccctggaaat acctgagctc ctaatctggg aataggagta ggaagagctt 5520
tggaggcagg gcacctgagt ttgagatcta caacttcctg cctgtgtgac attgggaaag 5580
tctccatcct ttctgagcct cagtctccac cctggggaag tggaaatatc aatctctgtg 5640

acacagaagc aaatgagcga atgtgcacaa agtaccttgc acaagagaga cgctcaaaca 5700

cttgcctcca ggtttcaccg agaactacag agtaagatag atttgttccc agtggaggaa 5760

gcctgggaat aatttgcccc tagactatga attcctgggg ctcaagatcg agcacagggc 5820

caggcacaca gaagggaccc tggaaatgtg gcaggaggcc agagatagac aggcccttag 5880

agctcatacc catgccctct gacctcaaga agaaagaaac ctgctcaaaa tctcacaaag 5940

agcttgttcc aaccctgaat cgagtctgag gactccttcc tgagtccagc actttttctg 6000

caagaagtat atgcctccaa agctgatggg cgcaaatctt gaaccccgtc acataaacac 6060

aaagggagga ggtgactaga gctcctccta ctggatatgt ctaaggtcac cagtctaaag 6120

aaaagggatg gatagaatga ggccagtatt tttgcagcca tccaaatgtc cacatacgct 6180

gttacactga gggctcctct ctcccccgtc ttcagcccta cttgcattta gaggtgagaa 6240

agatatgggc tgaggggttg tttttcatcg tattgtagat ggaaagcaca ctgcccttgg 6300

ggccatccaa atgtggacct tgatgtagca ccccaccttc tggatggcca tccttctgaa 6360

agtcactgaa tttctcagac tttattctct ttatccataa agaaggagaa taataataat 6420

ccccccaccc tgcccaacca ctgactggtt gggaagctca gaagaaatac tgggcacggc 6480

atcccattgt aatctataga gtgagtcgct tcttaatatt aaatggctga acacagaaga 6540

tgtgcaaaaa gtactgtgtc cccttcctcc tccaactgaa catttcatgc cctttgcacc 6600

ctcattttgt ctaggagctg ccttatgaag ggaataggta cctgctccga gctggaggaa 6660

tctttgccac ttatggtggg gtatggactg agacagagat ggcatgtgac atgcgcactg 6720

agtctcaact ccatgcaggc tctggagcac tctcaaattg gagtactaat gccttttaaa 6780

ttctcacact agcaatcctt tgacctactg atctagggat ctagggaaag aatcgtgatc 6840

ttaacttcaa agggaaggac aaaatgttct gcctcctgtt aaaactccat acactaagtg 6900

cagagactgg atgccttatt aaccttgggt agatgcccaa atgttcaaaa ggtcaaactc 6960

ttctgttccc cagatcgcca gagtcattaa ccagtcacac tattaaatga atgaacagat 7020

gctgaaaagg tacttgcatt actgagattt cttatggtga tggcccctgc ctgatatgta 7080

ttcagcattt tgtagttttc aatgtgcatt agagtatagt ggtgatgaca ttggcctctg 7140

agtttgccac ttcttatatc tgtgactttg gtcaaattgc ttaatctctc tgagtctcgg 7200

tttcctggag ataataatag cttcttcttc ccagggttat catgaggatt acaggagata 7260

atgccccaaa aatgcttagt aaagtgccta gcacctagtc aatgctgaat taaaggtggt 7320

tattcttact tttcgttcat ttgaactttg ttctcaggga gggcaaagga tagacaaagc 7380

cccatagcta gtgaggagta gctgcaagac tagaacccag gtgttctgag ccctagtctt 7440

aggccaagaa caactgttac gtgagatgca cgttttcctt caagggagct cacaattatt 7500

tccatgtaaa ttcaaggact gctaaaagag aactctcctc tgggactgat atcattttat 7560

ttcaagattg atttgaaaca tgtttttgt ttgtttgttt gttttctagg aaagaacaag 7620

agaaccagtc aagctgaatg cctgaagcaa atctctgtta gcgatgtttt caggatgagg 7680

gagagtggtg caagaaacgt gcttccagat gcacatggtt tcctgggact agggttcagg 7740

gtgtcatccc tgggtgttat taagtgtcag aaggagagca aacaagggaa acatctgaga 7800

tccagctaag gctacaccct ggaaatgcaa gcccagctct tgcaaaggac ctcctttggc 7860

cactcacctt ccaggcctta caataacttg tttggactgc aggtttcttg gtggactcac 7920

aggccattct gcttttattt ggtcaacctc agttcacaag cacccagatg ctgagatcct 7980

cagcatgtgc agcagagttt catattagca ctgggtacct ttctgaggct acagggatac 8040

cgtacagcag cacctgtcac gtccagccaa aggagtgggc tctctcaatg tcatccaatg 8100

ctgtttcaac tgtgaagaag accatctgag agagttgctt ttggaggctg aggcaaattt 8160

ttaaaattct ttgttctcct caactggggt gaattcttgg tcttctagga cagcttgaag 8220

ttttagaaag agtcaagcca ctcagaacca acagagaact ctttcagaga acaaggtgtg 8280

gcatagagga ggcagagggc tgatcttgat caaatccaaa gtgtgactct aaagcaatga 8340

atgtgaattt ttggcaaagc ttacaaaggg ctctaaaggc catctgcaaa gagaagccaa 8400

gcctgatcga tgaatcacta gtgcggccgg atatcgatcg gcacgctgtt gattttctca 8460

tagtaaggaa cagtgggccc tttcagtccc acttctgtag tctgtggtac tacaaatggt 8520

gagcccatga tgttgccatt catagggtta ttctccagca gtaatgactg gccagccact 8580

cccatagccg cagggctagg atttattgtc aatggaggga cctgcagttc tgcacaagca 8640

gtactaggat gagcacctgg gcccattgca agggtgacat cttcaaggca aggcctctta 8700

attttattag ggtagccccc atcagccatg tctggaaact ggaagtggtc ttcttcttgt 8760

ctcctcttaa cagttccctg tgaatggaag agaagagaag aggagaagag aggagaggag 8820

aagggaagag aggtgacaca cacacacaca cacacacaca cacacacaga gagagagaga 8880

gagagacaga gagaaagaga gagagagaga gagaggaatt tttataaagg tttggcacat 8940

taaagctaat gaacaggaaa tgtgcatgat aaaacagacc tctcagttta aagacttata 9000

gttgtgaaaa ctataaaata cagcctgtct ttggaaccat agtgcttatt tattcattat 9060

tatgtttcat ctaaactgtc taattacatt tcaaataagg cattatgttg tctgtatact 9120

aaaacgggat agaacgttat tcaaagggta atctgcccac ttcaaggaga gttcaacaaa 9180

actatgcaga agtcactaaa tgaaccatgc tgccaaaggc aggcattgga gagaaaacta 9240

gaagtagcta aatagtttta attctttcct gtctacagac acatagattt taacgaagga 9300

ataccatagt atagaattga acttttaggc tgccttctag tcttggttaa atgcatcagg 9360

ctgcagtggt aaaattgaat acaacagagc ccttacagga aagaagtaga tctggatgtg 9420

ttttcttggg gagctgttta aaatactgtt tttgggaaag cacaagtttc agaacagtca 9480

ttgtaggcat cgtattcatt gttccattta tttttacaca cacacacaca cacacacaca 9540

cacacactct cacacattgc tatgtgtaca caaaaataat ttggaagaac ctatacccaa 9600

caatttggag tggtcattta tttgggatga ctggcaattc cctttctatt ctcttcattt 9660

ctgcttgttt gtctttaacg agaacgactc ataatccaaa aatttaaaaa agtataaagt 9720

tatctaaata agaaattttc ctctgaagat gcatcctcag gttggggaga tattaaacaa 9780

tgagaaaagg ccccaatctg ggatctgaac cttgggggag ctgcccatca tttatagaag 9840

cacagccttt gggaacaaag caaagtcact agcaatgtga gacttcctac tcttcatggc 9900

ttcatacagt catccatcgc tgttgtgtta atgaccatga cctgtatgtt agcaggtaaa 9960

tgggaaagga agtgggggca aaggagtatg tgcaggaatg atcaaaataa ggaaaggaag 10020

agagggatct ggaaatcacc tgaatgccga taggtgaaca ggtagaattc ttttaaagct 10080

tcccccaccc ggtacccccc aaataacccc tttccagctt tggaagtttc actaggacat 10140

acagtgctca tcctctgatg tcaccttaag tttggctctt ctggtttgat gagcttgtag 10200

cccactagga gctcaaggca tgcatggggc cacttgccag cacgatgagg ggcatgactg 10260

tcatggccaa gtgaacatca aagcagatcc ccagggctgt atgtctcagg ccttggtgca 10320

catcagaatc acttagaaac atccacattc ctgggccctc ccaccacaaa ctgacagctt 10380

catccagggt gtggcccagg catcgggagt ttttccaaca gctccatggc tgattctcaa 10440

cagaaaacca ctggcccaga gcaagggtgg aggcagcgtg gcatagggct ctgaccttgg 10500

ccttgccact gaacctctca gagccccagt ttctttatgt gtaaaatgag tgtaattata 10560

gttcttttct catgaaggtg ctctgactat taagtgaaac ggggcacatt gtatgacacc 10620

taatagctcc tcactaactg gtacccggca ttataaaggg caggtatgga agggttctgg 10680

gagtccaata cccttcttaa agacagagag gtctctgaga cccagagagg ggcaggcctt 10740

acccagagtt gctcagccag agggcaacaa ggcccaggtc agatgcaggg cccctccacc 10800

accactcagc tgcctccaga cccactgcct tcgccatgtt gttggtagga cactgcatcg 10860

cccccacaga aggggcttgc caacttgagt gagaggactt gcacacttct ttgacttttc 10920

ttttgagatg cccacaatct gaacaagggc acttcaaggg acagctctgt caccaaactc 10980

atctgaggcc tgaataccat gggtcaggca ggaatgggtt ggagaggtgt agagcaggca 11040

caataagagg gctgaggccc atgcagtcat cagtgcccac tttcccagga gtctgactgg 11100

gcacagcacc catagtgtcc ctgagctggt ccatggagca gctcactaac tgtttggccc 11160

acagcaggtg ctcagtaaat ggcagttgaa cgaatcaatg gacaaaggaa cataaattac 11220

ccaacacaca gggagctcag ccatttactc aatccattat ggagtaacct acaaacaagc 11280

cactgggtcc caaactgaaa ttgtgtctct tctacattct cccaaagaat ccaataggtt 11340

aaaaatagaa atgtatgaaa tagatcaatc agggatgatt gcatgtggat ttgacataag 11400

gatcccctgc agggagtctg agctggcaac agtcaggccc aaagtgctgt ccatgatgtc 11460

tcgaactgca agacagtttt aacaatggcg aagcaatgca gaaccaggca ggccaaggag 11520

ggggtggggg ttggggaaag gaagggaggg aaggggctgt gaggggcaat ggtctggcat 11580

ccctgccacg tgagcctctg aaatttgctg gcagcttcta tgggctccca gagctttcac 11640

ttaattgttg gtctgccact aacctgctgg gagtaaggtg cagggatgga ggaggcaggg 11700

catgaccacc agacactaaa ggtaccagct ggggccactg gcaaagggaa ggaggctgca 11760

cctctcctac atgagagccc gtatacacac acctttttcca gcactcatca actgcatccc 11820

aagcaaatgg tccctgatca attccaattc tagaaaccaa ctgactactc aataacaaag 11880

tagatcccag caggccgcca ctgctggagc ggatgccact tttgctatgc caagtctgtg 11940

gctggacagc tgctggcatg tacactcact gactttcata aggatgccta ataaaggggg 12000

caggctcacc tggcttttct caggggtggg gtttggggtg ccgatagagg ctgctgtttt 12060

ggcagagtgg caagctgcaa gcctcttctg agctttcatt tttcaatgga cttcagtgag 12120

aattcacttt gtcagaggcc atgcagctcc atgttttgga tttcatggaa tgagctttca 12180

acagtgagcc tgaagtgccc tggctgaaca gcaagaacac cagccaaccc taaacaaggc 12240

cgaggagagg cggctgtgtt tacacggaag gctcagcctt gctgtaatag cgtctgcctt 12300

caccagacat cagtgaggcg tggaaatcta ttatccagtt aattttgccc ctagataaag 12360

acttgctttc gtgtcttctc tttcacagtc ccatgatctg ttactcatct caactgcgag 12420

aagttggctg ggctttcccc tgtgcccagt gccacactcg tgccttcact gggtcacctg 12480

tgcctgtggc tgatgccgct gaggttttgc ctgcccagac tgggtgtttc tgactaaatc 12540

ccacagccac cattttagat caagggcagg agatagctca ctgctccgga atgacctccc 12600

ctcccagaat cctggtaggg gcggaaggtc cccaaccaag ctcccagccc tttctaaatg 12660

aatctccctg cttcacccat gtgcttttct ccagtctctg cggtcttgat gacagcaggg 12720

tattagtcct agctgtccca cagctcctac ttctttcagg cctctccctg tgacaatcag 12780

tagccactgg caggatttcc tcagagcata tctcgatttg ctttcagaca attagttaaa 12840

aggacactgg accccagacg tcccaactcc cagccagagc cctcacaggc ccggcctttg 12900

gtggtgagga agggggaggg agtgagtgac agtgccctgg catcttttag aaacgaattc 12960

ctttctctcc atacataaat gcctgcagag tcccatttca gaatccggca gacaaagcca 13020

ccaatgtgat ccccatgacc ttataaacat tcattaaaat gcatttcaag gcatgtgatg 13080

gcctccccac cccctagata atgagaaaac aaaggtttct cttctgatag agacaagttc 13140

agctctgaag tcaacattat ttctggttct gtctgaacaa tgacatatgg caactcttcc 13200

ctttctatag ttctagtcca gaatgacaaa aaaggggaaa aatttcttag agaaggtaga 13260

gattatacga atacagtcca tgaaatgagc ataaggagaa taaagaatat aacttatcca 13320

aagaagtctg gcaggctgtt ataaatgctt gattttggac actgtagttg gaggtttaac 13380

atggacacca ataaaaaggt cagcaaaggg tatgcactgt tcctattggg caagaagata 13440

ggaggtcaaa ggtaaccagg aaagataaac tcagggagac ttattttccc tccagagggc 13500

actgggcttg taggccctgg gcaaaattgt caaaaaggtg aaaatcgcct gtggtttatt 13560

tagtctgctc tttcttcact agtgcctcac cagttcagtt caggccaatt tgctagctgg 13620

ctccttcgtt aattaaggat ccgagggcct atttcccatg attccttcat atttgcatat 13680

acgatacaag gctgttagag agataattag aattaatttg actgtaaaca caaagatatt 13740

agtacaaaat acgtgacgta gaaagtaata atttcttggg tagtttgcag ttttaaaatt 13800

atgttttaaa atggactatc atatgcttac cgtaacttga aagtatttcg atttcttggc 13860

tttatatatc ttgtggaaag gacgaaacac cgcaaggcta ttggtcaagg cagttttaga 13920

gctagaaata gcaagttaaa ataaggctag tccgttatca acttgaaaaa gtggcaccga 13980

gtcggtgctt tttttccgcg gtgggcgcgc cggatcctaa tcgatttgag taattcatac 14040

aaaaggactc gcccctgcct tggggaatcc cagggaccgt cgttaaactc ccactaacgt 14100

agaacccaga gatcgctgcg ttcccgcccc ctcacccgcc cgctctcgtc atcactgagg 14160

tggagaagag catgcgtgag gctccggtgc ccgtcagtgg gcagagcgca catcgcccac 14220

agtccccgag aagttggggg gaggggtcgg caattgaacc ggtgcctaga gaaggtggcg 14280

cggggtaaac tgggaaagtg atgtcgtgta ctggctccgc ctttttcccg agggtggggg 14340

agaaccgtat ataagtgcag tagtcgccgt gaacgttctt tttcgcaacg ggtttgccgc 14400

cagaacacag gtaagtgccg tgtgtggttc ccgcgggcct ggcctcttta cgggttatgg 14460

cccttgcgtg ccttgaatta cttccacgcc cctggctgca gtacgtgatt cttgatcccg 14520

agcttcgggt tggaagtggg tgggagagtt cgaggccttg cgcttaagga gcccccttcgc 14580

ctcgtgcttg agttgaggcc tggcctgggc gctggggccg ccgcgtgcga atctggtggc 14640

accttcgcgc ctgtctcgct gctttcgata agtctctagc catttaaaat ttttgatgac 14700

ctgctgcgac gctttttttc tggcaagata gtcttgtaaa tgcgggccaa gatctgcaca 14760

ctggtatttc ggtttttggg gccgcgggcg gcgacggggc ccgtgcgtcc cagcgcacat 14820

gttcggcgag gcggggcctg cgagcgcggc caccgagaat cggacggggg tagtctcaag 14880

ctcgccggcc tgctctggtg cctggcctcg cgccgccgtg tatcgccccg ccctgggcgg 14940
caaggctggc ccggtcggca ccagttgcgt gagcggaaag atggccgctt cccggccctg 15000
ctgcagggag ctcaaaatgg aggacgcggc gctcgggaga gcgggcgggt gagtcaccca 15060
cacaaaggaa aagggccttt ccgtcctcag ccgtcgcttc atgtgactcc acggagtacc 15120
gggcgccgtc caggcacctc gattagttct cgagcttttg gagtacgtcg tctttaggtt 15180
gggggggaggg gttttatgcg atggagtttc cccacactga gtgggtggag actgaagtta 15240
ggccagcttg gcacttgatg taattctcct tggaatttgc ccttttgag tttggatctt 15300

ggttcattct caagcctcag acagtggttc aaagtttttt tcttccattt caggtgtcgt 15360
gagctagtaa tcgatcctag ggccaccatg tacccatacg atgttccaga ttacgcttcg 15420
ccgaagaaaa agcgcaaggt cgaagcgtcc gacaagaagt acagcatcgg cctggacatc 15480
ggcaccaact ctgtgggctg ggccgtgatc accgacgagt acaaggtgcc cagcaagaaa 15540
ttcaaggtgc tgggcaacac cgaccggcac agcatcaaga agaacctgat cggagccctg 15600
ctgttcgaca gcggcgaaac agccgaggcc acccggctga agagaaccgc cagaagaaga 15660
tacaccagac ggaagaaccg gatctgctat ctgcaagaga tcttcagcaa cgagatggcc 15720

aaggtggacg acagcttctt ccacagactg gaagagtcct tcctggtgga agaggataag 15780
aagcacgagc ggcaccccat cttcggcaac atcgtggacg aggtggccta ccacgagaag 15840
taccccacca tctaccacct gagaaagaaa ctggtggaca gcaccgacaa ggccgacctg 15900
cggctgatct atctggccct ggcccacatg atcaagttcc ggggccactt cctgatcgag 15960
ggcgacctga accccgacaa cagcgacgtg gacaagctgt tcatccagct ggtgcagacc 16020
tacaaccagc tgttcgagga aaaccccatc aacgccagcg gcgtggacgc caaggccatc 16080
ctgtctgcca gactgagcaa gagcagacgg ctggaaaatc tgatcgccca gctgcccggc 16140

gagaagaaga atggcctgtt cggcaacctg attgccctga gcctgggcct gaccccaac 16200
ttcaagagca acttcgacct ggccgaggat gccaaactgc agctgagcaa ggacacctac 16260
gacgacgacc tggacaacct gctggcccag atcggcgacc agtacgccga cctgtttctg 16320
gccgccaaga acctgtccga cgccatcctg ctgagcgaca tcctgagagt gaacaccgag 16380
atcaccaagg ccccccctgag cgcctctatg atcaagagat acgacgagca ccaccaggac 16440
ctgaccctgc tgaaagctct cgtgcggcag cagctgcctg agaagtacaa agagattttc 16500
ttcgaccaga gcaagaacgg ctacgccggc tacattgacg gcggagccag ccaggaagag 16560

ttctacaagt tcatcaagcc catcctggaa aagatggacg gcaccgagga actgctcgtg 16620

aagctgaaca gagaggacct gctgcggaag cagcggacct tcgacaacgg cagcatcccc 16680

caccagatcc acctgggaga gctgcacgcc attctgcggc ggcaggaaga tttttaccca 16740

ttcctgaagg acaaccggga aaagatcgag aagatcctga ccttccgcat cccctactac 16800

gtgggccctc tggccagggg aaacagcaga ttcgcctgga tgaccagaaa gagcgaggaa 16860

accatcaccc cctggaactt cgaggaagtg gtggacaagg gcgcttccgc ccagagcttc 16920

atcgagcgga tgaccaactt cgataagaac ctgcccaacg agaaggtgct gcccaagcac 16980

agcctgctgt acgagtactt caccgtgtat aacgagctga ccaaagtgaa atacgtgacc 17040

gagggaatga gaaagcccgc cttcctgagc ggcgagcaga aaaaggccat cgtggacctg 17100

ctgttcaaga ccaaccggaa agtgaccgtg aagcagctga aagaggacta cttcaagaaa 17160

atcgagtgct tcgactccgt ggaaatctcc ggcgtggaag atcggttcaa cgcctccctg 17220

ggcacatacc acgatctgct gaaaattatc aaggacaagg acttcctgga caatgaggaa 17280

aacgaggaca ttctggaaga tatcgtgctg accctgacac tgtttgagga cagagagatg 17340

atcgaggaac ggctgaaaac ctatgcccac ctgttcgacg acaaagtgat gaagcagctg 17400

aagcggcgga gatacaccgg ctggggcagg ctgagccgga agctgatcaa cggcatccgg 17460

gacaagcagt ccggcaagac aatcctggat ttcctgaagt ccgacggctt cgccaacaga 17520

aacttcatgc agctgatcca cgacgacagc ctgaccttta aagaggacat ccagaaagcc 17580

caggtgtccg gccagggcga tagcctgcac gagcacattg ccaatctggc cggcagcccc 17640

gccattaaga agggcatcct gcagacagtg aaggtggtgg acgagctcgt gaaagtgatg 17700

ggccggcaca agcccgagaa catcgtgatc gaaatggcca gagagaacca gaccacccag 17760

aagggacaga agaacagccg cgagagaatg aagcggatcg aagagggcat caaagagctg 17820

ggcagccaga tcctgaaaga acaccccgtg gaaaacaccc agctgcagaa cgagaagctg 17880

tacctgtact acctgcagaa tgggcgggat atgtacgtgg accaggaact ggacatcaac 17940

cggctgtccg actacgatgt ggaccatatc gtgcctcaga gctttctgaa ggacgactcc 18000

atcgacaaca aggtgctgac cagaagcgac aagaaccggg gcaagagcga caacgtgccc 18060

tccgaagagg tcgtgaagaa gatgaagaac tactggcggc agctgctgaa cgccaagctg 18120

attacccaga gaaagttcga caatctgacc aaggccgaga gaggcggcct gagcgaactg 18180

gataaggccg gcttcatcaa gagacagctg gtggaaaccc ggcagatcac aaagcacgtg 18240

gcacagatcc tggactcccg gatgaacact aagtacgacg agaatgacaa gctgatccgg 18300

gaagtgaaag tgatcaccct gaagtccaag ctggtgtccg atttccggaa ggatttccag 18360

ttttacaaag tgcgcgagat caacaactac caccacgccc acgacgccta cctgaacgcc 18420

gtcgtgggaa ccgccctgat caaaaagtac cctaagctgg aaagcgagtt cgtgtacggc 18480

gactacaagg tgtacgacgt gcggaagatg atcgccaaga gcgagcagga aatcggcaag 18540

gctaccgcca agtacttctt ctacagcaac atcatgaact ttttcaagac cgagattacc 18600

ctggccaacg gcgagatccg gaagcggcct ctgatcgaga caaacggcga aaccggggag 18660

atcgtgtggg ataagggccg ggattttgcc accgtgcgga aagtgctgag catgccccaa 18720

gtgaatatcg tgaaaaagac cgaggtgcag acaggcggct tcagcaaaga gtctatcctg 18780

cccaagagga acagcgataa gctgatcgcc agaaagaagg actgggaccc taagaagtac 18840

ggcggcttcg acagccccac cgtggcctat tctgtgctgg tggtggccaa agtggaaaag 18900

ggcaagtcca agaaactgaa gagtgtgaaa gagctgctgg ggatcaccat catggaaaga 18960

agcagcttcg agaagaatcc catcgacttt ctggaagcca agggctacaa agaagtgaaa 19020

aaggacctga tcatcaagct gcctaagtac tccctgttcg agctggaaaa cggccggaag 19080

agaatgctgg cctctgccgg cgaactgcag aagggaaacg aactggccct gccctccaaa 19140

tatgtgaact tcctgtacct ggccagccac tatgagaagc tgaagggctc ccccgaggat 19200

aatgagcaga aacagctgtt tgtggaacag cacaagcact acctggacga gatcatcgag 19260

cagatcagcg agttctccaa gagagtgatc ctggccgacg ctaatctgga caaagtgctg 19320

tccgcctaca acaagcaccg ggataagccc atcagagagc aggccgagaa tatcatccac 19380

ctgtttaccc tgaccaatct gggagcccct gccgccttca agtactttga caccaccatc 19440

gaccggaaga ggtacaccag caccaaagag gtgctggacg ccaccctgat ccaccagagc 19500

atcaccggcc tgtacgagac acggatcgac ctgtctcagc tgggaggcga cagccccaag 19560

aagaagagaa aggtggaggc cagctaagaa ttcaataaaa gatctttatt ttcattagat 19620

ctgtgtgttg gtttttgtg tgcggccgcc cccgggcgat acatggttag atcaagcaca 19680

atgaaacatg gttagatcaa gcacaatgaa acatggttag atcaagcaca atgaaacatg 19740

gttagatcaa gcacaacgat agtgaattct accagtgcca taagtagtga attctaccag 19800

tgccataagt agtgaattct accagtgcca taagtagtga attctaccag tgccataatc 19860

ctagggcggc cgccacccca ccagtgcagg ctgcctatca gaaagtggtg gctggtgtgg 19920

ctaatgccct ggcccacaag tatcactaag ctcgctttct tgctgtccaa tttctattaa 19980

aggttccttt gttccctaag tccaactact aaactggggg atattatgaa gggccttgag 20040

catctggatt ctgcctaata aaaaacattt attttcattg caatgatgta tttaaattat 20100

ttctgaatat tttactaaaa agggaatgtg ggaggtcagt gcatttaaaa cataaagaaa 20160

tgaagagcta gttcaaacct tgggaaaata cactatatct taaactccat gaaagaaggt 20220

gaggctgcaa acagctaatg cacattggca acagcccctg atgcctatgc cttattcatc 20280

cctcagaaaa ggattcaagt agaggcttga tttggaggtt aaagttttgc tatgctgtat 20340

tttacattac ttattgtttt agctgtcctc atgaatgtct tttcactacc catttgctta 20400
tcctgcatct ctcagccttg actccactca gttctcttgc ttagagatac caccttccc 20460
ctgaagtgtt ccttccatgt tttacggcga gatggtttct cctcgcctgg ccactcagcc 20520
ttagttgtct ctgttgtctt atagaggtct acttgaagaa ggaaaaacag ggggcatggt 20580
ttgactgtcc tgtgagccct tcttccctgc ctcccccact cacagtgacc cggaatctgc 20640
agtgctagtc tcccggaact atcactcttt cacagtctgc tttggaagga ctgggcttag 20700
tatgaaaagt taggactgag aagaatttga aaggggggctt tttgtagctt gatattcact 20760

actgtcttat taccctatca taggcccacc ccaaatggaa gtcccattct tcctcaggat 20820
gtttaagatt agcattcagg aagagatcag aggtctgctg gctcccttat catgtccctt 20880
atggtgcttc tggctctgca gttattagca tagtgttacc atcaaccacc ttaacttcat 20940
ttttcttatt caatacctag aaagcttcag acatgataag atacattgat gagtttggac 21000
aaaccacaac tagaatgcag tgaaaaaaat gctttatttg tgaaatttgt gatgctattg 21060
ctttatttgt aaccattata agctgcaata aacaagttaa caacaacaat tgcattcatt 21120
ttatgtttca ggttcagggg gaggtgtggg aggtttttta aagcaagtaa aacctctaca 21180

aatgtggtac ccgggaagct gaaggcagat aatgaccagg acaggtcaac ctctcttctt 21240
ctgaaagcct tcatctacta atggcctggg actcttccct taaatgctta gattgtgtct 21300
tccactaagg ttttttgctg ttgctgttgt ttgtttgttt gtttgtttgt ttgtttgttt 21360
tgagacggaa tctcactctg tcgcccaggc tggagtgtag tggcacaatc tcagctcacc 21420
acaaccttca cctcctaggt tgaagggatt ctcctgcctc agcctcctga gtagctagga 21480
ttacaggcac atgccaccat gcctggctaa tttttgtatt tttagtagag acaggatttc 21540
gccatgttgg ccaggctggt cttgaactcc tgacctcagg tgatctgcct accttggtct 21600

cccaaagtgc tgggattact ggtgtgagcc accacacccg gccaaggttt ttgtttgttt 21660
gtttgtttgt ttgttttgta ttgaggcagg gtatcactct ggtcacccag gctggagtgc 21720
agtagtgcaa tcacggctca ctgaaacctc cacctccctg gcgggctcag gtgatcctgc 21780
cacctcagct tcccaggtag ctgggactac aggcttgtac caccactccc agctaatttt 21840
tgcgttttta gtagagacag ggtttcccca tgttgcccag gttggtctca aactctgggc 21900
tcaagcgatc tgcctgcctc agcctcccaa agtgctggga ttacaggtgt aagccaccgt 21960
acccggcccc gccactaagg ttttgaaaat gaagcaatta caagtttaag tctattaata 22020

agtgatgaag ctatgtagaa aagcagaata attatcttgg atcaggaagg tcacatgagg 22080

atctacttgg gggttgtcaa tattctattt cttgacctga tcagtgttga cagcaggttt 22140

taatttttta cttctttttg tttgtttgtt tttgagacgg agtcttgctc tgtctcccag 22200

gctggagtgc agtggtatga tctcggctca ctgcaacctc cgcctcctgg gttcaagctg 22260

ttctcctgcc tcagcctccc cagtagctgg gattacaggc aggcaccacc acgaccagct 22320

aatttttgta tttttagtag agactgggtt tcaccatctt ggccaggctg gtctcgaact 22380

tctgatctcg tgatccgccc tccttggcct cccaaagtgc tgggattaca ggcttgagcc 22440

agcgtgcccg gcccattttt tacttcctta ttaaactgta catataggcc ttgcacactt 22500

ttctgcatca atgttatatt ccacaataaa gggaaaaggt atatacacaa cttgatacca 22560

gtaatgtgaa acatatattt ctacatagaa aaaaaaatga ctgaaatact gcactccaat 22620

gtgttcacac agtagttgtt tctggattat ttatatatta aatgtttata tattgtatta 22680

tgccatgagg tttgtgtttt ctctccactt ttctgcattt tccaagttta ctacaaagag 22740

cacatattac tcttataatc agaaagtcat aaaatatatt taaaaagaca aaattgaaac 22800

taataaggat caacacaaaa cagatgagcc atctgtggaa atccgcacag aatactacct 22860

aaagagattg gtgacgtgca tgatctcact aggatgagca caaagcttgc cagagcctag 22920

ggtctatttc tagggttggc tcttggaagc caggatagtt gttatctctg ggaagaggga 22980

ggggcacaca aggggcttct aaaacattct gaatgttcta tttctgaacc tggttggtgg 23040

gtacatgact gttggtttta ttattatatg ttttatatac tcttccgtat gtatggtgtg 23100

gattccaaaa aaagatttcc tttagagaaa accagaatca cataagtaga aaatatggtg 23160

ctatgttgaa ggaacaactc aagtttatat aaaatcatca tcatttatag gcttaaaaag 23220

ttgctttgga attttggtct aactgacttg tcttttctgc agcaaaccac gctccttctg 23280

gacgtgctcc aggcagaggg gattagggtg ggttcaaggc tgcaagtacc tagctcagca 23340

cactctcttc aggggactta gagtttgtct ggtgttggct ctctgagctc ttgtcaggaa 23400

tgccgaccct tccgaggttc aggatttgaa gcctgccttc ccaccccaga tttggtccac 23460

acagacactc aagtatgtat ttcaactaca aatgacctgt actttcctat tactcctctc 23520

tttcatggta acctttctgg tatccttcct tccctacatt tatgggaggg ggacatcatt 23580

ctctgctctc ctgtcactga aggctccacc ttctgtcttc ttctgaccca tctggttttc 23640

ctggggccac ctcctctcct taccacccta acgcttttgt aacttgagga gaaatgagag 23700

atcacctagt caggtcatca ttctctgtag atgaagaggc ccaatggttt gctcaagaat 23760

tgccaagcga gttaaagaca gagagtatga gagtcagcaa gacctacaga aagcatctat 23820

ctgcactgtt ttgcagggac ttagcctttg tgtgtggact cctggaatgc cacccactaa 23880

gaaacattgt ctgacaccaa ctccccactt ggtaggtggg gacactgaaa ctcatggcag 23940

gaaagggcct tgccccaagc cagggcagag tgtcactcat cactctcaat tttcagtcca 24000

gggcaccttg ttgtgactat cccaaaggca gccactttcc ctggtctgaa agacctgaag 24060

agagaagaga agagaaggat ggaaggcaga gtatgcggct ttgattcatt tcctggtgaa 24120

aacagatcta tacgagaagc aaatttcacg aaagggaaga gaagaaagtg tcccatacgt 24180

tgctggcctg tttcaacctt gctttgattc ttgctgaaaa gggtaccgtg tatttctgag 24240

ttcaacatgc agaccagtgt taggaaagcc actgcacctc cactttagcc tccagggctg 24300

tgccctgcaa atggcctgca gccttggtgc ctcgctctcc agactgcatt ttggaagatg 24360

ggacagaggc ttatggaagc ccacattaga acgggggagc agaatgggtg agatgaggga 24420

tccttgatag tgaaccagat gaaggaatgg tagccaaatg ccaggcctcc tttgtggctt 24480

caatccaaag gctctggagc ccttccaggg cagaacatca ggcatgttta cccccactgt 24540

cctcaacagt gacagaggtg caatcttggg cagctggcca ttttgaaagc aacctcctta 24600

atctcaactg ggaaggctcc ctagcaggac ccctgtgttg cacacctgga ggaagctaga 24660

ctaaccagaa gctcagcacg gttccatctg ggatgcccag gtctgagacg aaaaaggtaa 24720

ctctcttttc tgggtcctgg cccagttgtg tctctctcca cctcattctc tgagatgcct 24780

gtctcccctt ttttgtccca tcaggaggca agagctatca ctgggccaga ctccaccaga 24840

agccaagcca gcttgttacc cagcttctca gggagcaaag aacagccttg tttctatctt 24900

atccccactg tcccctgccc ctgccccacc tcccagccat tcagcttctg gcttccccag 24960

agctgcctgc ttctttgtgg tcctccattc cttgaaaaga ccttctagtc attagtgtat 25020

ataaatggcc acttagccca gattacagtg aggtcaacag ctggggctct gagaattgtc 25080

acacactggc acaggagagg aggctattct tccagagaat ttggagggca ctcccatcca 25140

cttacaacaa aaagcccatc cactgtgctt ggcagtaggt gatctgagaa ccaatggaac 25200

caggttaatc ctgtggcact gttgagtgag gagagcagtg gcgggcactg gaaaatatca 25260

gagacaaggc aggagacctg aaatctaggc ttagctcctc atatacttgg cagctgtatg 25320

acctcagaca accagtgtta cctctctaag cctcagtttc ctcatgcaaa aggaggggga 25380

ataacaacag agcccactgc ttgggggtgt tgtgaggaca ggatgaaaaa acaaacagaa 25440

atccctcagt acaggattca gtgcagtgga cagtcttgca aggtctggtt cagccctcca 25500

cccctaccct caccagtata aagaactctg gcctacaagt cagatgacct gagttttaat 25560

ctcagctttg ccattagccg tgtgaacttg agaaagtccc tttccttttt acatctattg 25620

ggatgatcat gcatttttg tcctttattc tgttaatata gtgtgttaca ttgattgctt 25680

ttcatagact gaaccagcct tgtattccag ggataaatct cacttggtca tggtgtataa 25740

tcctttatac aaatgttgct gggttgagtt tgctagtatt ttgttgaaga tttttatgtc 25800

ttgattcata aggaatattg gtgtaccttc ccctttttatg gccacagttt ccctacaatg 25860
atgtagtcga actagacaac ctccaatatc tttcagtatt catgtcctct gattctgtga 25920
aactaagaaa attaagaaat agtgattcat aggcacaagg caggcaaaac ttagactcct 25980
tgtagaataa ttaggaagcc aaatattcag tgtgcttatt tctcaaataa ccttagtttc 26040
tccagtctgc cccaactccg aggcctgaat atctctagat gcttatgatg gcaactaaag 26100
cctaaaagct aattcatttt aaagttcttc caaatgcata gggttttatt tttccagacc 26160
tgggttcaga tggggaattt gacaaacaat ggaaaggggg aaaaacaaca atctaaacac 26220

tgagtgacaa agtaacaaag aaatagtcta gctatcagcc agtcaagcca gccttggctt 26280
tgctatccaa agtagtcagt ctaattctac caccagtttc tgttcctgta gctgtctact 26340
gcctgccagg gactctgcct tcccacccac aactaccaat ggaaggatgt ggtgaccata 26400
ccagtggctg ctgacatctc ctgccatggg aagcataatt gcctccagca gcctcccct 26460
tagatccatc atttttgttg cacttggcct gggctgtact cccggccaat gactgaacat 26520
ggtgagcata gtaatgcagg cccatttctg tgaggagcag gactcctcca gtaggtgact 26580
ttggctcaag gactctctat tggcctggtt gaacttttcc tgaactgtgc tactgtctga 26640

gactcttctt acccaatcct ctttctcgcc ccaattgtca cagaccacct gcattgtggt 26700
ctgagtctct ccccaccttc tcttgctctt ccctgtttat ctttcacagg catttccccc 26760
agtacattcc ttgaatgtct aacccgatac gggtgcctga cttttggcag acctaagcag 26820
acaaaaagga gtacttggtt acctagctct tctttctacc acaaacatcg agggaaccct 26880
ttttccctca cccctctgcc acacccccac tgccccagtg aacaaccaca gagagagctg 26940
tggtataata ttaggctggt gcaaaagtaa ttgcggtttt tgccattact tttaatggta 27000
aaaaccgcaa ttacttttgc acctacctag tatttgtgtc cccccaaatt catatgttga 27060

aacctaaccc acaatatgat gtcattagga ggcaagacct tgaggaggtg attagatgat 27120
ggggtggagc tctcctgaat gagattagtg cccttataag aagaagccca aggaagctac 27180
cttgactctt ccatcacatg agaatgcagc aagaaggcac catctactaa tcaggaagag 27240
agctctcacc agacactgaa tctgccagtg tcttgatctt gaagttccca gcctccagaa 27300
ctatgcataa tgcatttcca ttgtctctaa gccacccagc ctatggtatt ttgtcatagc 27360
agcctgaact gactaagaca gtgagccaca tgagaagtgc cccaacccct cccttaagca 27420
cttggctcac agatcagtgg gttcatttct gcctgagttt tattgttatt ctgtagattt 27480

cttgggctag atatattttt ctgttatttt ccttcttcac ctcagtcatg aattggttgt 27540

tttaaaaaag acaatgtaag tcatggggaa actcctgaca actctactct cctagggttc 27600

ctgataaaag gggattcagt tgagtcctct gatggtctct acctgccaaa gtccagcagc 27660

ccttagcaaa catgctgctc gtttctgtag agaaggtgct ggtgtcccac catacttctc 27720

tctccctcat gaagggcttg cgacccagca aatgggtggc ttatatgggt ctgtttcaaa 27780

ggaagagcca gctctgggaa gaaaaacgat gagcataagc ataacctacc actgtgcctg 27840

ggaaagcaga caacttttt gatgtgtgaa tatctaatga gaatggaatc catcaattac 27900

cttaaactta ggcacagtct tcaaattcaa tatatgtggg atatactttt agtcagtttg 27960

tagacgttat ttgtaataaa taatctggct tctctaaaga aattatttta agtgtttggt 28020

ttggtttgat ttaatggtaa aattatattt agtggcagag aattatagca atggtgataa 28080

actatagagt gtcataagtt catatcttat tctcacattt gaagctgcct gcagatgcat 28140

tcaagatgca gccagaagtc aggagactca ggctgttatt tggagctcat cattttacag 28200

ccttgctgga ctcccacttt ctcaggggaa aaatgtggtg ttgacccaga ttagctctcc 28260

aggccctgct gagttgggca ctctgtaagc tggagggtct tctattgtct tcacctaagt 28320

gtcaatcaac aacccaaatg ggcatggggg aagagggagc tgggccaatg cccagggtgc 28380

ctggtagaga gataccttgg gcactggaag gcaccagctt cccagagaga agggggaggg 28440

ccatgaaaaa gttggctgta gatgccaggg acactgggac tctccagctg tgtgtttgtg 28500

tcttctgaag acttatgttt cattcctttg gagcatgcat aatcatacac tgtgggatgt 28560

gttatataga ttgcttgata gttcaccact gtaataaaat actgtgactg gaatctgctc 28620

ccagtctgcc tttgatagca cttgtgcaac acacatttac tgagcattta cagtgatcca 28680

ggacctgtgt tgtgaaaaca ttgatggaca aggcagatgg tggagcacgt cagtgaggat 28740

ttttaacaaa ggctggtaag tgctataaag gaacattgta ggacactaga gaacaaagaa 28800

caggagaacc tgacttaggc tggggtgggg cgttggttag aggaggctcc ttggaggaca 28860

tgaggtttaa gctgtgacct gaggatgaat agatgttggc caggtgaggt taattaagct 28920

tgtcagcctt accagtaaaa aagaaaacct attaaaaaaa caccactcga cacggcacca 28980

gctcaatcag tcacagtgta aaaaagggcc aagtgcagag cgagtatata taggactaaa 29040

aaatgacgta acggttaaag tccacaaaaa acacccagaa aaccgcacgc gaacctacgc 29100

ccagaaacga aagccaaaaa acccacaact tcctcaaatc gtcacttccg ttttcccacg 29160

ttacgtaact tcccatttta agaaaactac aattcccaac acatacaagt tactccgccc 29220

taaaacctac gtcacccgcc ccgttcccac gccccgcgcc acgtcacaaa ctccacccca 29280

tcattatcat attggcttca atccaaaata aggtatatta ttgatgatgt ttaaactacg 29340

gcccggtacc cagcttttgt tccctttagt gagggttaat tgcgcgcttg gcgtaatcat 29400

ggtcatagct gtttcctgtg tgaaattgtt atccgctcac aattccacac aacatacgag 29460

ccggaagcat aaagtgtaaa gcctggggtg cctaatgagt gagctaactc acattaattg 29520

cgttgcgctc actgcccgct ttccagtcgg gaaacctgtc gtgccagctg cattaatgaa 29580

tcggccaacg cgcggggaga ggcggtttgc gtattgggcg ctcttccgct tcctcgctca 29640

ctgactcgct gcgctcggtc gttcggctgc ggcgagcggt atcagctcac tcaaaggcgg 29700

taatacggtt atccacagaa tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc 29760

agcaaaaggc caggaaccgt aaaaaggccg cgttgctggc gtttttccat aggctccgcc 29820

cccctgacga gcatcacaaa aatcgacgct caagtcagag gtggcgaaac ccgacaggac 29880

tataaagata ccaggcgttt ccccctggaa gctccctcgt gcgctctcct gttccgaccc 29940

tgccgcttac cggatacctg tccgcctttc tcccttcggg aagcgtggcg ctttctcata 30000

gctcacgctg taggtatctc agttcggtgt aggtcgttcg ctccaagctg ggctgtgtgc 30060

acgaaccccc cgttcagccc gaccgctgcg ccttatccgg taactatcgt cttgagtcca 30120

acccggtaag acacgactta tcgccactgg cagcagccac tggtaacagg attagcagag 30180

cgaggtatgt aggcggtgct acagagttct tgaagtggtg gcctaactac ggctacacta 30240

gaaggacagt atttggtatc tgcgctctgc tgaagccagt taccttcgga aaaagagttg 30300

gtagctcttg atccggcaaa caaaccaccg ctggtagcgg tggttttttt gtttgcaagc 30360

agcagattac gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt tctacggggt 30420

ctgacgctca gtggaacgaa aactcacgtt aagggatttt ggtcatgaga ttatcaaaaa 30480

ggatcttcac ctagatcctt ttaaattaaa aatgaagttt taaatcaatc taaagtatat 30540

atgagtaaac ttggtctgac agttaccaat gcttaatcag tgaggcacct atctcagcga 30600

tctgtctatt tcgttcatcc atagttgcct gactccccgt cgtgtagata actacgatac 30660

gggagggctt accatctggc cccagtgctg caatgatacc gcgagaccca cgctcaccgg 30720

ctccagattt atcagcaata aaccagccag ccggaagggc cgagcgcaga agtggtcctg 30780

caactttatc cgcctccatc cagtctatta attgttgccg ggaagctaga gtaagtagtt 30840

cgccagttaa tagtttgcgc aacgttgttg ccattgctac aggcatcgtg gtgtcacgct 30900

cgtcgtttgg tatggcttca ttcagctccg gttcccaacg atcaaggcga gttacatgat 30960

cccccatgtt gtgcaaaaaa gcggttagct ccttcggtcc tccgatcgtt gtcagaagta 31020

agttggccgc agtgttatca ctcatggtta tggcagcact gcataattct cttactgtca 31080

tgccatccgt aagatgcttt tctgtgactg gtgagtactc aaccaagtca ttctgagaat 31140

agtgtatgcg gcgaccgagt tgctcttgcc cggcgtcaat acgggataat accgcgccac 31200

atagcagaac tttaaaagtg ctcatcattg gaaaacgttc ttcggggcga aaactctcaa 31260

ggatcttacc gctgttgaga tccagttcga tgtaacccac tcgtgcaccc aactgatctt 31320

cagcatcttt tactttcacc agcgtttctg ggtgagcaaa aacaggaagg caaaatgccg 31380

caaaaaaggg aataagggcg acacggaaat gttgaatact catactcttc ctttttcaat 31440

attattgaag catttatcag ggttattgtc tcatgagcgg atacatattt gaatgtattt 31500

agaaaaataa acaaataggg gttccgcgca catttccccg aaaagtgcca cctgacgcgc 31560

cctgtagcgg cgcattaagc gcggcgggtg tggtggttac gcgcagcgtg accgctacac 31620

ttgccagcgc cctagcgccc gctcctttcg ctttcttccc ttcctttctc gccacgttcg 31680

ccggctttcc ccgtcaagct ctaaatcggg ggctcccttt agggttccga tttagtgctt 31740

tacggcacct cgaccccaaa aaacttgatt agggtgatgg ttcacgtagt gggccatcgc 31800

cctgatagac ggtttttcgc cctttgacgt tggagtccac gttctttaat agtggactct 31860

tgttccaaac tggaacaaca ctcaacccta tctcggtcta ttcttttgat ttataaggga 31920

ttttgccgat ttcggcctat tggttaaaaa atgagctgat ttaacaaaaa tttaacgcga 31980

attttaacaa aatattaacg cttacaattt ccattcgcca ttcaggctgc gcaactgttg 32040

ggaagggcga tcggtgcggg cctcttcgct attacgccag ctggcgaaag ggggatgtgc 32100

tgcaaggcga ttaagttggg taacgccagg gttttcccag tcacgacgtt gtaaaacgac 32160

ggccagtgag cgcgcgtaat acgactcact atagggcgaa ttggagctcc actacgtagt 32220

ttaaacatca tcaataatat accttatttt ggattgaagc caatatgata atgaggggt 32280

ggagtttgtg acgtggcgcg gggcgtggga acggggcggg tgacgtagta gtgtggcgga 32340

agtgtgatgt tgcaagtgtg gcggaacaca tgtaagcgac ggatgtggca aaagtgacgt 32400

ttttggtgtg cgccggtgta cacaggaagt gacaattttc gcgcggtttt aggcggatgt 32460

tgtagtaaat ttgggcgtaa ccgagtaaga tttggccatt ttcgcgggaa aactgaataa 32520

gaggaagtga aatctgaata attttgtgtt actcatagcg cgtaatattt gtctagggcc 32580

gcggggactt tgaccgttta cgtggagact cgcccaggtg tttttctcag gtgttttccg 32640

cgttccgggt caaagttggc gttttgattc 32670

<210> 63

<211> 30655

<212> DNA

<213> Artificial Sequence

<220><223> LCR globin vector (Sleeping beauty IR, Beta promoter, human gamma globin, mgmt, Efla promoter and Ad5)

<400> 63

catcatcaat aatatacctt attttggatt gaagccaata tgataatgag ggggtggagt 60

ttgtgacgtg gcgcggggcg tgggaacggg gcgggtgacg tagtagtgtg gcggaagtgt 120

gatgttgcaa gtgtggcgga acacatgtaa gcgacggatg tggcaaaagt gacgtttttg 180

gtgtgcgccg gtgtacacag gaagtgacaa ttttcgcgcg gttttaggcg gatgttgtag 240

taaatttggg cgtaaccgag taagatttgg ccattttcgc gggaaaactg aataagagga 300

agtgaaatct gaataatttt gtgttactca tagcgcgtaa tatttgtcta gggccgcggg 360

gactttgacc gtttacgtgg agactcgccc aggtgttttt ctcaggtgtt ttccgcgttc 420

cgggtcaaag ttggcgtttt gattcggccg aaggattaca tgagcttaga aatgtaatta 480

gcatagtgtg tggcatagtg tagataccaa ataaatatga tctctccttc tactcttgaa 540

aatgcaaaca cattcttggt ggtcctaaaa tagcctgtaa catggtttac tcagcagcat 600

ttgctattca aggcagatct gcctttagtc attggctgcg ctcctgaaca gctgtgtgaa 660

aggctaactt ttgtaaacca aatcaaaata aaatgcagca aaaatttgtc actgaaagga 720

aatcctcagt atatcctttt atgaaatgaa agatccctca tccaaactta acttttttaa 780

aagtgcgcat ttggagatat agccctttct tatgaatcct aattcaattt tggccataaa 840

cacacgttga tgttccccac cccaaagcac atagcaacaa gagtaggttc tatattgaaa 900

ataatgacaa tttaaaaaca tgtacttatt tcactgtatg tggacagtgt ctatgattgc 960

atcatgaagt gtcatataac catgtacgtg tacatgagag agagatagag agagaagtgg 1020

tagggtggtg gtggtagagg ggatggcgat agtaatcatg gtaatggtag aggtgatgga 1080

ggtggtaatg acggaggtaa gggtggtagt gatgatggtg gtggtggtaa tggtggtgga 1140

tgtggtggtg gcaattggga tggtgggatg gtggtagcca tggtgatggt ggtaatggtg 1200

ttgatttaaa gggtggtggt agtgaaggtg agggtagtgg tggtggaggt ggtggtgctg 1260

gtagcaatag tgatggtggt gatggtgttg atgagggtgt tgggatcagg gtgagttccc 1320

acagtatatt tcattcttgt tgtaccactc tgtcaacagc accactgact gggacagagg 1380

aagaaggcac actctgaatg tgttattaac agaaacctca aaacagtctg tctccttgta 1440

gtcattcaaa attatctttt tcttacctgg aaaactgaaa ctgaattacc gggaaaaaca 1500

caggagattt ttgtttgtta atatgctgcc aataaagtaa ttttatgtca aatttaacta 1560

caggaaaggg caaggcattt tctaagttcc ttagatgtca tgtggctaaa aaaaacaaaa 1620

ggatggacag cagttagata ctgtacactt agctgtttga agccatatat tcagaaagca 1680

gatgttggga gttggtgttt gaggactgat ttcctggagg tattttatat aggccaagtt 1740

cattgttcta aactctaagg gcttgacttg agggaggaaa agaggcaaga acatgtttag 1800
ttttgctgac agcatcacat gggcagccct aaggctagac aactttaggg cctgaagctt 1860
attctaggaa agaagcacct acagagtggc actgggctcc cctccactat agagatgaag 1920
tcatatgaca gtaaagggca ggcagggctg cctaggggggc ccattgaaat tgcggccgca 1980
aataatgggc ccggagcaag agagagggag gcaatgacag cagagacatg cctgcgcctt 2040
gggtttgagt gcccagtggt caaatccact tccctgtggc tgatgcttgc ctttctaact 2100
ttggaattta ggggttggag atctggtgag aaggtaggag ggagatgagg aggagaaggg 2160

aaaggcagga aggaagggga gggaaaggaa aagcaaaagg ggaggaggaa ggtttccaac 2220
aaattattct atatcaactg cggaaatcaa aatttgttgc ccaaatctta gaagctcatg 2280
tccctcctcc ccagaagtct ggaatgcagc actccagggg tagcttataa cccaaatatc 2340
tatctgtaaa aagagaaaca ttgggctttc gagctgtgga ttctcagtaa aagcaagagg 2400
cctcagccta cacaggccag cccagagttt gaggaacccc aggcccacac ccacagggct 2460
ggcccctggg tctgcatact ccctagaaat gtgcacactt ctgagcctca actctgtcct 2520
ggagtctaac agcatccctc tccttcctgg ggcagttcca cctccagaaa cctgttacct 2580

tgggccttat gtcaaggaaa ctgtgggaaa gagctaggca ggaatgcaga tgaggccagc 2640
atgggctcct aaaagtttag aaataggcag tgtcatgctc ccaggtgcct gcataaacca 2700
gctgaaaaat ggagctcccc tcaccagcac tctcccttca aacagactgt gatttgcagg 2760
tcactggttt accaagccag gctacccagg caggacccag atgccaagcc cagtggtgtc 2820
ctgcaagctg agcagtgctc agttcttgca aaaaaaggtc tgtgtgaagg caaggcctct 2880
gcctggcttc tcaccccagt tgggtgtctg gaacaggaag gagcccttac tgcagaaaaa 2940
ggaggaggga gcaaagggag cgaacagctg cgtgctccat ggggaggatc cccaaagtag 3000

aaaggcgcat acacactgca gcccttgacc cagaatgctc acagctacat tacagattca 3060
ggtctcctca gtgtagtggg gctgctgatg agactgtggc atcctcaggg gtcaggacac 3120
acattttcca tcactcttct gatggcaaaa aacctctgag ccaatgccaa cctctgatca 3180
ttaaaaaaaa gtgctcacag cagtgtgtgg tttaggatca tgccctgtgt ggtttggaac 3240
acgtgcacaa ccacaccttg ttcatcacca tcccagaaac cctgacgcag gcaaagagca 3300
gagttattaa ccctacttta ctgatgtgga tactgaggcc cagaggctca tgcaagttat 3360
caataagtgg cagggacagt tgcctctaga ttaactagcc cctaggatca cctgggtctt 3420

ggaaggggac ccataaacat gagctcccct ctcttggggc cagatttgca cctgtgccgc 3480

gccttcagcc tgcatgaagt aggggctgct ggcaaagact caaagctgta aatctgggtt 3540

ttctcttgag gcttctaagg gagctgtttc gacaactcac tctgttccca gctggctgcc 3600

cctgcatagg gttttaaagc agcctagctt tctgccaggc ttggcagtgg acaacgctgg 3660

tcagaacatc ccagagagct accagaatga agtaagtttg cttctactct ttacctgttt 3720

atgggctgtc tctgccactg gaatgaaagg cactgagaac agtgcctggc ctgcagaagg 3780

ccctggaaat acctgagctc ctaatctggg aataggagta ggaagagctt tggaggcagg 3840

gcacctgagt ttgagatcta caacttcctg cctgtgtgac attgggaaag tctccatcct 3900

ttctgagcct cagtctccac cctggggaag tggaaatatc aatctctgtg acacagaagc 3960

aaatgagcga atgtgcacaa agtaccttgc acaagagaga cgctcaaaca cttgcctcca 4020

ggtttcaccg agaactacag agtaagatag atttgttccc agtggaggaa gcctgggaat 4080

aatttgcccc tagactatga attcctgggg ctcaagatcg agcacagggc caggcacaca 4140

gaagggaccc tggaaatgtg gcaggaggcc agagatagac aggcccttag agctcatacc 4200

catgccctct gacctcaaga agaaagaaac ctgctcaaaa tctcacaaag agcttgttcc 4260

aaccctgaat cgagtctgag gactccttcc tgagtccagc acttttttctg caagaagtat 4320

atgcctccaa agctgatggg cgcaaatctt gaaccccgtc acataaacac aaagggagga 4380

ggtgactaga gctcctccta ctggatatgt ctaaggtcac cagtctaaag aaaagggatg 4440

gatagaatga ggccagtatt tttgcagcca tccaaatgtc cacatacgct gttacactga 4500

gggctcctct ctcccccgtc ttcagcccta cttgcattta gaggtgagaa agatatgggc 4560

tgaggggttg tttttcatcg tattgtagat ggaaagcaca ctgcccttgg ggccatccaa 4620

atgtggacct tgatgtagca ccccaccttc tggatggcca tccttctgaa agtcactgaa 4680

tttctcagac tttattctct ttatccataa agaaggagaa taataataat ccccccaccc 4740

tgcccaacca ctgactggtt gggaagctca gaagaaatac tgggcacggc atcccattgt 4800

aatctataga gtgagtcgct tcttaatatt aaatggctga acacagaaga tgtgcaaaaa 4860

gtactgtgtc cccttcctcc tccaactgaa catttcatgc cctttgcacc ctcattttgt 4920

ctaggagctg ccttatgaag ggaataggta cctgctccga gctggaggaa tctttgccac 4980

ttatggtggg gtatggactg agacagagat ggcatgtgac atgcgcactg agtctcaact 5040

ccatgcaggc tctggagcac tctcaaattg gagtactaat gccttttaaa ttctcacact 5100

agcaatcctt tgacctactg atctagggat ctagggaaag aatcgtgatc ttaacttcaa 5160

agggaaggac aaaatgttct gcctcctgtt aaaactccat acactaagtg cagagactgg 5220

atgccttatt aaccttgggt agatgcccaa atgttcaaaa ggtcaaactc ttctgttccc 5280

cagatcgcca gagtcattaa ccagtcacac tattaaatga atgaacagat gctgaaaagg 5340

tacttgcatt actgagattt cttatggtga tggcccctgc ctgatatgta ttcagcattt 5400

tgtagttttc aatgtgcatt agagtatagt ggtgatgaca ttggcctctg agtttgccac 5460

ttcttatatc tgtgactttg gtcaaattgc ttaatctctc tgagtctcgg tttcctggag 5520

ataataatag cttcttcttc ccagggttat catgaggatt acaggagata atgccccaaa 5580

aatgcttagt aaagtgccta gcacctagtc aatgctgaat taaaggtggt tattcttact 5640

tttcgttcat ttgaactttg ttctcaggga gggcaaagga tagacaaagc cccatagcta 5700

gtgaggagta gctgcaagac tagaacccag gtgttctgag ccctagtctt aggccaagaa 5760

caactgttac gtgagatgca cgttttcctt caagggagct cacaattatt tccatgtaaa 5820

ttcaaggact gctaaaagag aactctcctc tgggactgat aaacatctag tcgagtatcg 5880

acaaataggg gttccgcgca catttccccg aaaagtgcca cctgacgtaa ctataacggt 5940

cctaaggtag cgaaagcttg catgcctgca ggtcgactct agaggatcaa acctagccac 6000

cgcggtggcg gccggctagc cggctagccg gctagcccta gaactagtaa cggccgccag 6060

tgtgctggaa ttcggcttgt aaggtaccgg tgaagttcct atactttcta gagaatagga 6120

acttcggaat aggaacttct acctagatgc atgctcgagc ggcccctaca gttgaagtcg 6180

gaagtttaca tacacttaag ttggagtcat taaaactcgt ttttcaacta ctccacaaat 6240

ttcttgttaa caaacaatag ttttggcaag tcagttagga catctacttt gtgcatgaca 6300

caagtcattt ttccaacaat tgtttacaga cagattattt cacttataat tcactgtatc 6360

acaattccag tgggtcagaa gtttacatac actaagttga ctgtgccttt aaacagcttg 6420

gaaaattcca gaaaatgatg tcatggcttt agaagcttct gatagactaa ttgacatcat 6480

ttgagtcaat tggaggtgta cctgtggatg tatttcaagg aattctgtgg aatgtgtgtc 6540

agttagggtg tggaaagtcc ccaggctccc cagcaggcag aagtatgcaa agcatgcata 6600

tcgatactag tttataagat ctcgagctag cggccgtttg ttaattaagt cgacggtacc 6660

gtcaaggctg cagtgagaca tgatcttgcc actgcactcc agcctggaca gcagagtgaa 6720

accttgcctc acgaaacaga atacaaaaac aaacaaacaa aaaactgctc cgcaatgcgc 6780

ttccttgatg ctctaccaca taggtctggg tactttgtac acattatctc attgctgttc 6840

ataattgtta gattaatttt gtaatattga tattattcct agaaagctga ggcctcaaga 6900

tgataacttt tattttctgg acttgtaata gctttctctt gtattcacca tgttgtaact 6960

ttcttagagt agtaacaata taaagttatt gtgagttttt gcaaacacag caaacacaac 7020

gacccatata gacattgatg tgaaattgtc tattgtcaat ttatgggaaa acaagtatgt 7080

acttttttcta ctaagccatt gaaacaggaa taacagaaca agattgaaag aatacatttt 7140

ccgaaattac ttgagtatta tacaaagaca agcacgtgga cctgggagga gggttattgt 7200

ccatgactgg tgtgtggaga caaatgcagg tttataatag atgggatggc atctagcgca 7260
atgactttgc catcactttt agagagctct tggggacccc agtacacaag aggggacgca 7320
gggtatatgt agacatctca ttctttttct tagtgtgaga ataagaatag ccatgacctg 7380
agtttataga caatgagccc ttttctctct cccactcagc agctatgaga tggcttgccc 7440
tgcctctcta ctaggctgac tcactccaag gcccagcaat gggcagggct ctgtcagggc 7500
tttgatagca ctatctgcag agccagggcc gagaaggggt ggactccaga gactctccct 7560
cccattcccg agcagggttt gcttatttat gcatttaaat gatatattta ttttaaaaga 7620

aataacagga gactgcccag ccctggctgt gacatggaaa ctatgtagaa tattttgggt 7680
tccatttttt tttccttctt tcagttagag gaaaaggggc tcaggatcca cttgcccagt 7740
gttcttcctt agttcctacc ttcgaccttg atcctccttt atcttcctga accctgctga 7800
gatgatctat gtggggagaa tggcttcttt gagaaacatc ttcttcgtta gtggcctgcc 7860
cctcattccc actttaatat ccagaatcac tataagaaga atataataag aggaataact 7920
cttattatag gtaagggaaa attaagaggc atacgtgatg ggatgagtaa gagaggagag 7980
ggaaggatta atggacgata aaatctacta ctatttgttg agacctttta tagtctaatc 8040

aattttgcta ttgttttcca tcctcacgct aactccataa aaaaacacta ttattatctt 8100
tattttgcca tgacaagact gagctcagaa gagtcaagca tttgcctaag gtcggacatg 8160
tcagaggcag tgccagacct atgtgagact ctgcagctac tgctcatggg ccctgtgctg 8220
cactgatgag gaggatcaga tggatggggc aatgaagcaa aggaatcatt ctgtggataa 8280
aggagacagc catgaagaag tctatgactg taaatttggg agcaggagtc tctaaggact 8340
tggatttcaa ggaattttga ctcagcaaac acaagaccct cacggtgact ttgcgagctg 8400
gtgtgccaga tgtgtctatc agaggttcca gggagggtgg ggtggggtca gggctggcca 8460

ccagctatca gggcccagat gggttatagg ctggcaggct cagataggtg gttaggtcag 8520
gttggtggtg ctgggtggag tccatgactc ccaggagcca ggagagatag accatgagta 8580
gagggcagac atgggaaagg tgggggaggc acagcatagc agcatttttc attctactac 8640
tacatgggac tgctccccta tacccccagc taggggcaag tgccttgact cctatgtttt 8700
caggatcatc atctataaag taagagtaat aattgtgtct atctcatagg gttattatga 8760
ggatcaaagg agatgcacac tctctggacc agtggcctaa cagttcagga cagagctatg 8820
ggcttcctat gtatgggtca gtggtctcaa tgtagcaggc aagttccaga agatagcatc 8880

aaccactgtt agagatatac tgccagtctc agagcctgat gttaatttag caatgggctg 8940

ggaccctcct ccagtagaac cttctaacca gctgctgcag tcaaagtcga atgcagctgg 9000

ttagactttt tttaatgaaa gcttgcatgc agcactttgg gaggctgagg tgggtggact 9060

gcttggagct caggagttca agaccatctt ggacaacatg gtgataccct gcctctacaa 9120

aaagtacaaa aattagcctg gcatggtggt gtgcacctgt aatcccagct attagggtgg 9180

ctgaggcagg agaattgctt gaacccggga ggcggaggtt gcagtgagct gagatcgtgc 9240

cactgcactc cagcctgggg gacagagcac attataatta actgttattt tttacttgga 9300

ctcttgtggg gaataagata catgttttat tcttatttat gattcaagca ctgaaaatag 9360

tgtttagcat ccagcaggtg cttcaaaacc atttgctgaa tgattactat acttttttaca 9420

agctcagctc cctctatccc ttccagcatc ctcatctctg attaaataag cttcagtttt 9480

tccttagttc ctgttacatt tctgtgtgtc tccattagtg acctcccata gtccaagcat 9540

gagcagttct ggccaggccc ctgtcggggt cagtgcccca cccccgcctt ctggttctgt 9600

gtaaccttct aagcaaacct tctggctcaa gcacagcaat gctgagtcat gatgagtcat 9660

gctgaggctt agggtgtgtg cccagatgtt ctcagcctag agtgatgact cctatctggg 9720

tccccagcag gatgcttaca gggcagatgg caaaaaaaag gagaagctga ccacctgact 9780

aaaactccac ctcaaacggc atcataaaga aaatggatgc ctgagacaga atgtgacata 9840

ttctagaata tattatttcc tgaatatata tatatatata cacatatacc atatgaaaca 9900

cctctaggct ataaggcaac agagctcctt ttttttttt ctgtgctttc ctggctgtcc 9960

aaatctctaa tgataagcat acttctattc aatgagaata ttctgtaaga ttatagttaa 10020

gaattgtggg agccattccg tctcttatag ttaaatttga gcttctttta tgatcactgt 10080

ttttttaata tgctttaagt tctggggtac atgtgccatg gtggtttgct gcacccatca 10140

acccgtcatc tacattaggt atttctccta atgctatcct tcccctagcc ccccaccccc 10200

aacaggcccc agtgtgtgat gttcccctcc ctgtgtccat ggatcactgg ttttttttg 10260

ttttttttt ttttttaaag tctcagttaa atttttggaa tgtaatttat tttcctggta 10320

tcctaaggac ttgcaagtta tctggtcact ttagccctca cgttttgatg ataatcacat 10380

atttgtaaac acaacacaca cacacacaca cacacacata tatatatata taaaacatat 10440

atatacataa acacacataa catatttatc gggcatttct gagcaactaa tcatgcagga 10500

ctctcaaaca ctaacctata gccttttcta tgtatctact tgtgtagaaa ccaagcgtgg 10560

ggactgagaa ggcaatagca ggagcattct gactctcact gcctttagct aggcccctcc 10620

ctcatcacag ctcagcatag tcctgagctc ttatctatat ccacacacag tttctgacgc 10680

tgcccagcta tcaccatccc aagtctaaag aaaaaaataa tgggtttgcc catctctgtt 10740

gattagaaaa caaaacaaaa taaaataagc ccctaagctc ccagaaaaca tgactaaacc 10800

agcaagaaga agaaaataca ataggtatat gaggagactg gtgacactaa gtgtctgaat 10860

gaggcttgag tacagaaaag aggctctagc agcatagtgg tttagaggag atgtttcttt 10920

ccttcacaga tgccttagcc tcaataagct tgcggttgtg gaagtttact tgtttatcac 10980

cggtgacgtc catgagcaaa ttaagaaaaa caacaacaaa tgaatgcata tatatgtata 11040

tgtatgtgtg tatatataca catatatata tatatttttt ttcttttctt accagaaggt 11100

tttaatccaa ataaggagaa gatatgctta gaactgaggt agagttttca tccattctgt 11160

cctgtaagta ttttgcatat tctggagacg caggaagaga tccatctaca tatcccaaag 11220

ctgaattatg gtagacaaaa ctcttccact tttagtgcat caatttctta tttgtgtaat 11280

aagaaaattg ggaaaacgat cttcaatatg cttaccaagc tgtgattcca aatattacgt 11340

aaatacactt gcaaaggagg atgtttttag tagcaatttg tactgatggt atggggccaa 11400

gagatatatc ttagagggag ggctgagggt ttgaagtcca actcctaagc cagtgccaga 11460

agagccaagg acaggtacgg ctgtcatcac ttagacctca ccctgtggag ccacacccta 11520

gggttggcca atctactccc aggagcaggg agggcaggag ccagggctgg gcataaaagt 11580

cagggcagag ccatctattg cttacatttg cttctgacac aactgtgttc actagcaacc 11640

tcaaacagac accatgggtc atttcacaga ggaggacaag gctactatca caagcctgtg 11700

gggcaaggtg aatgtggaag atgctggagg agaaaccctg ggaaggtagg ctctggtgac 11760

caggacaagg gagggaagga aggaccctgt gcctggcaaa agtccaggtc gcttctcagg 11820

atttgtggca ccttctgact gtcaaactgt tcttgtcaat ctcacaggct cctggttgtc 11880

tacccatgga cccagaggtt ctttgacagc tttggcaacc tgtcctctgc ctctgccatc 11940

atgggcaacc ccaaagtcaa ggcacatggc aagaaggtgc tgacttcctt gggagatgcc 12000

ataaagcacc tggatgatct caagggcacc tttgcccagc tgagtgaact gcactgtgac 12060

aagctgcatg tggatcctga gaacttcaag gtgagtccag gagatgtttc agccctgttg 12120

cctttagtct cgaggcaact tagacaactg agtattgatc tgagcacagc agggtgtgag 12180

ctgtttgaag atactggggt tgggggtgaa gaaactgcag aggactaact gggctgagac 12240

ccagtggtaa tgttttaggg cctaaggagc gcctctaaaa atctagatgg acaattttga 12300

ctttgagaaa agagaggtgg aaatgaggaa aatgacttttt attagattcc agtagaaaga 12360

actttcatct ttccctcatt tttgttcgtt ttaaaacatc tatctggagg caggacaagt 12420

atggtcgtta aaaagatgca ggcagaaggc atatattggc tcagtcaaag tggggaactt 12480

tggtggccaa acatacattg ctaaggctat tcctatatca gctggacaca tataaaatgc 12540

tgctaatgct tcattacaaa cttatatcct ttaattccag atgggggcaa agtatgtcca 12600

ggggtgagga acaattgaaa catttgggct ggagtagatt ttgaaagtca gctctgtgtg 12660

tgtgtgtgtg tgcgcgcgcg cgtgtgtgtg tgtgtgtgtc aacgtgtgtt tcttttaacg 12720
tcttcagcct acaacataca gggttcatgg tggcaagaag atagcaagat ttaaattatg 12780
gccagtgact agtgcttgaa ggggaacaac tacctgcatt taatgggaag gcaaaatctc 12840
aggctttgag ggaagttaac ataggcttga ttctgggtag aagctgggtg tgtagttatc 12900
tggaggccag gctggagctc tcagctcact atgggttcat ctttattgtc tcctttcatc 12960
tcaacagctc ctgggaaatg tgctggtgac cgttttggca atccatttcg gcaaagaatt 13020
caccctgag gtgcaggctt cctggcagaa gatggtgact gcagtggcca gtgccctgtc 13080

ctccagatac cactgagcct cttgcccatg attcagagct ttcaaggata ggctttattc 13140
tgcaagcaat tcaaataata aatctattct gctgagagat cacacatgat tttcttcagc 13200
tctttttttt acatcttttt aaatatatga gccacaaagg gtttatattg agggaagtgt 13260
gtatgtgtat ttctgcatgc ctgtttgtgt ttgtggtgtg tgcatgctcc tcatttattt 13320
ttatatgaga tgtgcatttt gttgagcaaa taaaagcagt aaagacactt gtacacggga 13380
gttctgcaag tgggagtaaa tggtgtagga gaaatccggt gggaagaaag acctctatag 13440
gacaggactt ctcagaaaca gatgttttgg aagagatggg aaaaggttca gtgaagacct 13500

gggggctgga ttgattgcag ctgagtagca aggatggttc ttaatgaagg gaaagtgttc 13560
caagctttag gaattcaagg tttagtcagg tgtagcaatt ctattttatt aggaggaata 13620
ctatttctaa tggcacttag cttttcacag cccttgtgga tgcctaagaa agtgaaatta 13680
atcccatgcc ctcaagtgtg cagattggtc acagcatttc aagggagaga cctcattgta 13740
agactctggg ggaggtgggg acttaggtgt aagaaatgaa tcagcagagg ctcacaagtc 13800
agcatgagca tgttatgtct gagaaacaga ccagcactgt gagatcaaaa tgtagtggga 13860
agaatttgta caacattaat tggaaggctt acttaatgga atttttgtat agttggatgt 13920

tagtgcatct ctataagtaa gagtttaata tgatggtgtt acggacctaa tgtttgtgtc 13980
tcctcaaaat tcacatgctg aatccccaac tcccaactga ccttatctgt gggggaggct 14040
tttgaaaagt aattaggttt agatgagctc ataagagcag atccccatca taaaattatt 14100
ttccttatca gaagcagaga gacaagccat ttctctttcc tcccggtgag gacacagtga 14160
gaagtccgcc atctgcaatc caggaagaga accctgacca cgagtcagcc ttcagaaatg 14220
tgagaaaaaa ctctgttgtt gaagccaccc agtcttttgt attttgttat agcaccttgc 14280
actgagtaag gcagatgaag aaggagaaaa aaataagctt atcgaaacgc gtccccatcc 14340

tcactgactc cgtcctggag ttggatgaga gataatggcc ttacgttgtg ccaggggagg 14400

gtcgggctgg atttagcaag atttaccttc tccaaagagc ggtgctgcag tggcacagct 14460

gcccacggag gtgggggggt caccgtccct ggaggtgatg aagaactgtg gggatgtggc 14520

actgagggac atggccagtg ggcacggtgg gtgggttggg gttggtcttg gggatcttgg 14580

agggcttttc cagccttcat gatttgacga ttgtatgaac atctacatgg caattctcca 14640

gctgcctgtc ccagtcctac tgacccagct gtatctctcc aggcaagctc ttccaccccct 14700

tctgcttgca tccagacacc atcaaacatg caggctcaga cacagggacc agcagtgtct 14760

gtggcctttt tgtgctcctc tccatgctgg gttttaactt gctctttgtc cttctatcct 14820

atcttcttat ccttaaggct gttctgaacg ctgtgacttg gagagtgtcc cagagccctc 14880

aacacctgca tgtcccacgt ccatgctgtc ctgcacttcc ttatccccaa gatctgcctc 14940

tccgtgatgc actgaattgg caaacatgtg tcaccccaga ccaacaatgt cacagcaaac 15000

tccccttga taggacaagg gggaatggct ttacactgag acaggggagg tttgggttgg 15060

atatgaggag gcagtttttc ccccagaggg tggtgacgca ctgaacaggt tgcccaagga 15120

ggctgtggat gccccatccc tgcaggcatt caaggccagg ctggatgtgg ctctgggcag 15180

cctgggctgc tggttgatga ccctgcacat agcaggggggt tggatctgga tgagcactgt 15240

gctcctttgc aacccaggcc gttctatgat tctgtcattc taaatctctc tttcagccta 15300

aagctttttc cccgtatccc cccaggtgtc tgcaggctca aagagcagcg agaagcgttc 15360

agaggaaagc gatcccgtgc caccttcccc gtgcccgggc tgtccccgca cgctgccggc 15420

tcggggatgc gggggggagcg ccggaccgga gcggagcccc gggcggctcg ctgctgcccc 15480

ctagcggggg agggacgtaa ttacatccct gggggctttg gggggggggct gtccctatcg 15540

attactagtt ttaccacatt tgtagaggtt ttacttgctt taaaaaacct cccacacctc 15600

cccctgaacc tgaaacataa aatgaatgca attgttgttg ttaacttgtt tattgcagct 15660

tataatggtt acaaataaag caatagcatc acaaatttca caaataaagc attttttttca 15720

ctgcattcta gttgtggttt gtccaaactc atcaatgtat cttatcatgt ctgctcgaag 15780

agcggccgct tcagtttcgg ccagcaggcg gggagcccga ggtagctccc gctcccttga 15840

gccaggcccc tgccagacct gagctccctc ccaagcctgg cttccccaac cggtggcctt 15900

catgggccag aagccattcc ttcacggcta gccctccgga gtagttgccc acggctccgc 15960

tgctgcagac cactctgtgg cacgggatga ggatcttgac aggattgcct ctcatggcgc 16020

ctcccactgc tcgcgcggct ttggggttgc cggccagggc ggccaattgc tggtaagaaa 16080

tcacttctcc gaatttcaca accttaagca gcttccataa cacctgacgc gtgaacgact 16140

cttgctggaa aacgggatgg tgaagcgctg gcacggggaa ctcttcgata gcctcgggct 16200

ggtggaaata ggcattcagc caggctgtgc actgcatcag gggctccgga cctccgagaa 16260

ccgcagcggg ggctgggacc tccacggcat cagctgcaga cgtccccttg cccaggagct 16320

ttatttcgtg cagaccctgc tcacaaccag acagctccag cttccccaaa gggctgtcca 16380

gtgtggtgcg tttcatttca caatccttgt ccatggtggc gaccgtctag ctcacgacac 16440

ctgaaatgga agaaaaaaac tttgaaccac tgtctgaggc ttgagaatga accaagatcc 16500

aaactcaaaa agggcaaatt ccaaggagaa ttacatcaag tgccaagctg gcctaacttc 16560

agtctccacc cactcagtgt ggggaaactc catcgcataa aacccctccc cccaacctaa 16620

agacgacgta ctccaaaagc tcgagaacta atcgaggtgc ctggacggcg cccggtactc 16680

cgtggagtca catgaagcga cggctgagga cggaaaggcc cttttccttt gtgtgggtga 16740

ctcacccgcc cgctctcccg agcgccgcgt cctccatttt gagctccctg cagcagggcc 16800

gggaagcggc catctttccg ctcacgcaac tggtgccgac cgggccagcc ttgccgccca 16860

gggcggggcg atacacggcg gcgcgaggcc aggcaccaga gcaggccggc gagcttgaga 16920

ctacccccgt ccgattctcg gtggccgcgc tcgcaggccc cgcctcgccg aacatgtgcg 16980

ctgggacgca cgggccccgt cgccgcccgc ggccccaaaa accgaaatac cagtgtgcag 17040

atcttggccc gcatttacaa gactatcttg ccagaaaaaa agcgtcgcag caggtcatca 17100

aaaattttaa atggctagag acttatcgaa agcagcgaga caggcgcgaa ggtgccacca 17160

gattcgcacg cggcggcccc agcgcccagg ccaggcctca actcaagcac gaggcgaagg 17220

ggctccttaa gcgcaaggcc tcgaactctc ccacccactt ccaacccgaa gctcgggatc 17280

aagaatcacg tactgcagcc aggggcgtgg aagtaattca aggcacgcaa gggccataac 17340

ccgtaaagag gccaggcccg cgggaaccac acacggcact tacctgtgtt ctggcggcaa 17400

acccgttgcg aaaaagaacg ttcacggcga ctactgcact tatatacggt tctcccccac 17460

cctcgggaaa aaggcggagc cagtacacga catcactttc ccagtttacc ccgcgccacc 17520

ttctctaggc accggttcaa ttgccgaccc ctccccccaa cttctcgggg actgtgggcg 17580

atgtgcgctc tgcccactga cgggcaccgg agcctcacgc atgctcttct ccacctcagt 17640

gatgacgaga gcgggcgggt gaggggggcgg gaacgcagcg atctctgggt tctacgttag 17700

tgggagttta acgacggtcc ctgggattcc ccaaggcagg ggcgagtcct tttgtatgaa 17760

ttactcaaat cgataactag tgatatctta attaacaaac ggccgctgca ttctagttgt 17820

ggtttgtcca aactcatcaa tgtatcttat catgtctgga tcccatcaca aagctctgac 17880

ctcaatccta tagaaaggag gaatgagcca aaattcaccc aacttattgt gggaagcttg 17940

tggaaggcta ctcgaaatgt ttgacccaag ttaaacaatt taaaggcaat gctaccaaat 18000

actaattgag tgtatgttaa cttctgaccc actgggaatg tgatgaaaga aataaaagct 18060

gaaatgaatc attctctcta ctattattct gatatttcac attcttaaaa taaagtggtg 18120

atcctaactg accttaagac agggaatctt tactcggatt aaatgtcagg aattgtgaaa 18180
aagtgagttt aaatgtattt ggctaaggtg tatgtaaact tccgacttca actgtagggg 18240
atcctctagg gccgccagtg tgatggatat ctgcagaatt cggcttcagg taccgtcgac 18300
gatgtaggtc acggtctcga agccgcggtg cgggtgccag ggcgtgccct tgggctcccc 18360
gggcgcgtac tccacctcac ccatctggtc catcatgatg aacgggtcga ggtggcggta 18420
gttgatcccg gcgaacgcgc ggcgcaccgg gaagccctcg ccctcgaaac cgctgggcgc 18480
ggtggtcacg gtgagcacgg gacgtgcgac ggcgtcggcg ggtgcggata cgcggggcag 18540

cgtcagcggg ttctcgacgg tcacggcggg catgtcgaca agccgaattc cagcacactg 18600
gcggccgtta ctaggtagct agctcgagcc ttcgaagatc tcctagggaa gttcctatac 18660
tttctagaga ataggaactt cggaatagga acttcaccgg tgggtgaaaa gccgaattct 18720
gcagatatca agcttatcga taccgtcgac ctcgaggggg ggcccggtta gatccccggg 18780
taccgagctc gaatatctat gtcgggtgcg gagaaagagg taatgaaatg gcattatggg 18840
tattatgggt ctgcattaac gatagctagc ctctaactcc tagaccgtca gaactgctgg 18900
gcccttcaag acgggctgct cacacccact catgttaagc ctggtgaggc ctgtactctg 18960

ttttcacagg aagaaatcct cacccagtct tccccaaaca cattcccagg ttgtgtcatt 19020
agtgggatag agatgattat tgtggggaga agagaaacat ctggatggat ttggttaggt 19080
tgatctatag aggaagtagg tgctgcctga ggtagctgta atagaagcta aaggtcaaag 19140
gagagggccc tgtcccaatc cagatgactc cacttctgct ggacccaggt tcacaagctt 19200
aatctacatt tcacctaaat ttggctaaca agcccaaaat cacacaggca aagggagaag 19260
tggaggcaga accgaggttg gaggccacca gggccaccgg gcagagatca tttaagccca 19320
accttctcac ttctccctgg gctctgcctc tcttaaagga ccttgtggtg tgacctcttg 19380

taggtccctt tcacactcgg ggcctcagtt tccccactgt aaagtgaatg ggtcccagct 19440
ttggtaagct tatgcttacc tgatgctttc ttcctgggct gctcttgtag agaaaagata 19500
aatcttcttc ctccatccac gagggcttct ttccctgggg gtgagagtag gctgaggaga 19560
gccacttgca cacactctta aagaaagtat tacctgcacc agctcagtga gaggcacaga 19620
tcagactgtt acttgaatca aattatgagc ctccccaaat atatctatga catttaaata 19680
ggggattact tgaacataga ctttgggatc cggtgtggag tgcaggagac tagcaaagtg 19740
aatcctgaga gtagcaggtc tgcacctgtt ggatcgagaa aggcggccta caattctggt 19800

caaatgagct gtgcttattg acatattcta ttagagagta ctaccaggtc accagtcacc 19860

agaaaggctg ccagctctcc aaccacctcc agggaactat cctgaatggg gccttaacaa 19920

gcctaagaga gggttggttt gggtcccaag ccaatatttg ctctgcttta tgtcagtcat 19980

atggaaccca aaccaaccct ctcctatgtg cctcaccagt cggtgcaggg atcccaattt 20040

caagtttggt tttttatggt caaagtccag catagattaa atgaaggggt gtgatgatgg 20100

tgttaaaaga gaactccaga ccagtttaac tcttggacac acatcccatc tcaccatggt 20160

gcttccaacc ttccagagat gatgggctcc tattttctga tgacaaagcc ctccacagga 20220

ttgctgcctg gccatcaggg agtgcctctg taactgaggc tgagatccca ctttcagtcc 20280

tccagctgtg gcccatccct gctccgccca ccgggtatgg cctgtcctag gctcttaggt 20340

atggctgcat tgtgaaatga tggctacaga gctggcatct cctgtagtct ggttcatcta 20400

gtgcactacc tcatagttaa aagaaatctg tttaagccac tgagggtggc tcctagtgcc 20460

aactccaaga acaggaagct tccctttttt gggaggaggg gcagatggta acatggatcg 20520

tccaggtcaa tgggagcagg gcaaccacag taagtactgg acaacaacac aaaactccat 20580

gtgtggcttc catcgagtcc ctctccaatt ggtttggtct tctccgtccc atgcagcact 20640

ttagcaaggg gcctggctga aggctatgaa ttgtgtggag cctcctcatt gcagtctcca 20700

accatctgat gctgggaaaa tgtcaccagg atgcagccat gccgtgtggc caatgaaccg 20760

agaaaacacc ccttttctag aatgctctaa agaggcagaa taatccagag gtgaggaagg 20820

aaatactcca ccagagaccc aggcagttcc tacaaaagcc agactttcct tcacctaggg 20880

agtgacaaga ccagtggaaa acactctcaa gcagtaaccc ccaaatgctc tgcaagccag 20940

tggcgtccag ataccgcaca agcgagtggg ctgtctaatc ccatcatcat gatgtaaata 21000

tctctaggct gccctgggct gtgcctgacc ctgtcttcag ctttccacac ctccacctac 21060

agcccatgca cagaaggacc acccaggaat gctgcaagtg tggcacctcc agggccaccc 21120

agggagaagg agggcagcta tgctggtggc tccaggggtg gtaccttcac accacaaagc 21180

ccaaactgag gccccagatt tggctgatga gggcatattg gacaggggtc acttatgctc 21240

ttccccattg ccacctggcc tctggctacc tggacttggc tacctgtgga tcctctcaca 21300

ggtgccacca tcttggctga gtctccagat gcgaggtccc tgaggcagtg gcaggcttct 21360

cgctaatgct gatgggatta ggaatgggat aggtggggag ggccctggac tgggccctga 21420

tgagccaagt gggtttttag aggggctact ggtacatttc agggacagga catctggtag 21480

agctaagctg gggcaataag gagccactgc taatctgaga gctagaaaca atcagcttct 21540

gggtcattat taattagggt agtttgggct gtgtggaagt cacgtactat atggggtagc 21600

cacagctctc tctacagata atctctaaga cttctgattg ggaccgtgtg aatgcagtag 21660

caatatctct tcttactgcc aggccctgcc agtcctgcct ccacgccctg gctggccccc 21720

cttatgatct gacccatgcc aggctgccat agtatgttac ttctgcatta gcactccttg 21780

ggacctgcct ctccactgtc cctcagactt taaagaacta tacaaaccca aggggctctt 21840

cccaagagaa ttgatatgac ttgaggtgat tccatttctg gaagtagtca ctccattttc 21900

tgcctcactc tttcagtgct tcacagagca ggttcgaacg aaggagccat ccaactaacc 21960

gtcatgttcg ggcaaccgaa gaagggagtg gcaggatttc ctttggagac ttctggaatt 22020

agacagcagt ttaatgcaag catctaaatt ctcttcctcc cagagtctca ttaaaactac 22080

agtaagagtt tgtgttttgt tttgttttta aagacaaaat cccaccagga tagagagaat 22140

aggagaggag ataacagcat cataatttat gaaactaaaa tgcagataga ccaggattaa 22200

ctgactacac agcaccaagg aagctgaatc acaagacagc agaggagaaa actggaaagg 22260

atcgtggtct atacggcaga atcttcccaa gcctcaggag gaggagctct agatgttccc 22320

agatctggga ggtaaagtgg aatgggggga catggtcagc gtaatggggt tgggctggaa 22380

gctggttaag gagcaggcag atctctgaat ccctctctg actctgtgtc cccaggcatc 22440

tgcctgtccc ccaccctgga agaggtctgg cttgaccctt tgtctggtga atttcctgct 22500

ctgctttcct ggtcctgctg gccagatcag tggaggccac tcacttcacc ccacagggat 22560

gttctgtgtt gccctacacc tgggaactgg aggtactgga ggcaggctgt ggtgagcttg 22620

aaagcaaaac acagagggca gtccaatctc tttggccata tttcttctgc atatccaata 22680

ccatgtccac aactctgcta gtgtcctgat ggtggtgggc tctacacatt cccgggaagc 22740

tgaaggcaga taatgaccag gacaggtcaa cctctcttct tctgaaagcc ttcatctact 22800

aatggcctgg gactcttccc ttaaatgctt agattgtgtc ttccactaag gttttttgct 22860

gttgctgttg tttgtttgtt tgtttgtttg tttgtttgtt ttgagacgga atctcactct 22920

gtcgcccagg ctggagtgta gtggcacaat ctcagctcac cacaaccttc acctcctagg 22980

ttgaagggat tctcctgcct cagcctcctg agtagctagg attacaggca catgccacca 23040

tgcctggcta atttttgtat ttttagtaga gacaggattt cgccatgttg gccaggctgg 23100

tcttgaactc ctgacctcag gtgatctgcc taccttggtc tcccaaagtg ctgggattac 23160

tggtgtgagc caccacaccc ggccaaggtt tttgtttgtt tgtttgtttg tttgttttgt 23220

attgaggcag ggtatcactc tggtcaccca ggctggagtg cagtagtgca atcacggctc 23280

actgaaacct ccacctccct ggcgggctca ggtgatcctg ccacctcagc ttcccaggta 23340

gctgggacta caggcttgta ccaccactcc cagctaattt ttgcgttttt agtagagaca 23400

gggtttcccc atgttgccca ggttggtctc aaactctggg ctcaagcgat ctgcctgcct 23460

cagcctccca aagtgctggg attacaggtg taagccaccg tacccggccc cgccactaag 23520

gttttgaaaa tgaagcaatt acaagtttaa gtctattaat aagtgatgaa gctatgtaga 23580

aaagcagaat aattatcttg gatcaggaag gtcacatgag gatctacttg ggggttgtca 23640
atattctatt tcttgacctg atcagtgttg acagcaggtt ttaatttttt acttcttttt 23700
gtttgtttgt ttttgagacg gagtcttgct ctgtctccca ggctggagtg cagtggtatg 23760
atctcggctc actgcaacct ccgcctcctg ggttcaagct gttctcctgc ctcagcctcc 23820
ccagtagctg ggattacagg caggcaccac cacgaccagc taatttttgt atttttagta 23880
gagactgggt ttcaccatct tggccaggct ggtctcgaac ttctgatctc gtgatccgcc 23940
ctccttggcc tcccaaagtg ctgggattac aggcttgagc cagcgtgccc ggcccatttt 24000

ttacttcctt attaaactgt acatataggc cttgcacact tttctgcatc aatgttatat 24060
tccacaataa agggaaaagg tatatacaca acttgatacc agtaatgtga aacatatatt 24120
tctacataga aaaaaaaatg actgaaatac tgcactccaa tgtgttcaca cagtagttgt 24180
ttctggatta tttatatatt aaatgtttat atattgtatt atgccatgag gtttgtgttt 24240
tctctccact tttctgcatt ttccaagttt actacaaaga gcacatatta ctcttataat 24300
cagaaagtca taaaatatat ttaaaaagac aaaattgaaa ctaataagga tcaacacaaa 24360
acagatgagc catctgtgga aatccgcaca gaatactacc taaagagatt ggtgacgtgc 24420

atgatctcac taggatgagc acaaagcttg ccagagccta gggtctattt ctagggttgg 24480
ctcttggaag ccaggatagt tgttatctct gggaagaggg aggggcacac aaggggcttc 24540
taaaacattc tgaatgttct atttctgaac ctggttggtg ggtacatgac tgttggtttt 24600
attattatat gttttatata ctcttccgta tgtatggtgt ggattccaaa aaaagatttc 24660
ctttagagaa aaccagaatc acataagtag aaaatatggt gctatgttga aggaacaact 24720
caagtttata taaaatcatc atcatttata ggcttaaaaa gttgctttgg aattttggtc 24780
taactgactt gtcttttctg cagcaaacca cgctccttct ggacgtgctc caggcagagg 24840

ggattagggt gggttcaagg ctgcaagtac ctagctcagc acactctctt caggggactt 24900
agagtttgtc tggtgttggc tctctgagct cttgtcagga atgccgaccc ttccgaggtt 24960
caggatttga agcctgcctt cccaccccag atttggtcca cacagacact caagtatgta 25020
tttcaactac aaatgacctg tactttccta ttactcctct ctttcatggt aacctttctg 25080
gtatccttcc ttccctacat ttatgggagg gggacatcat tctctgctct cctgtcactg 25140
aaggctccac cttctgtctt cttctgaccc atctggtttt cctggggcca cctcctctcc 25200
ttaccaccct aacgcttttg taacttgagg agaaatgaga gatcacctag tcaggtcatc 25260

attctctgta gatgaagagg cccaatggtt tgctcaagaa ttgccaagcg agttaaagac 25320

agagagtatg agagtcagca agacctacag aaagcatcta tctgcactgt tttgcaggga 25380

cttagccttt gtgtgtggac tcctggaatg ccacccacta agaaacattg tctgacacca 25440

actccccact tggtaggtgg ggacactgaa actcatggca ggaaagggcc ttgccccaag 25500

ccagggcaga gtgtcactca tcactctcaa ttttcagtcc agggcacctt gttgtgacta 25560

tcccaaaggc agccactttc cctggtctga aagacctgaa gagagaagag aagagaagga 25620

tggaaggcag agtatgcggc tttgattcat ttcctggtga aaacagatct atacgagaag 25680

caaatttcac gaaagggaag agaagaaagt gtcccatacg ttgctggcct gtttcaacct 25740

tgctttgatt cttgctgaaa agggtaccgt gtatttctga gttcaacatg cagaccagtg 25800

ttaggaaagc cactgcacct ccactttagc ctccagggct gtgccctgca aatggcctgc 25860

agccttggtg cctcgctctc cagactgcat tttggaagat gggacagagg cttatggaag 25920

cccacattag aacgggggag cagaatgggt gagatgaggg atccttgata gtgaaccaga 25980

tgaaggaatg gtagccaaat gccaggcctc ctttgtggct tcaatccaaa ggctctggag 26040

cccttccagg gcagaacatc aggcatgttt acccccactg tcctcaacag tgacagaggt 26100

gcaatcttgg gcagctggcc attttgaaag caacctcctt aatctcaact gggaaggctc 26160

cctagcagga cccctgtgtt gcacacctgg aggaagctag actaaccaga agctcagcac 26220

ggttccatct gggatgccca ggtctgagac gaaaaaggta actctctttt ctgggtcctg 26280

gcccagttgt gtctctctcc acctcattct ctgagatgcc tgtctccccc tttttgtccc 26340

atcaggaggc aagagctatc actgggccag actccaccag aagccaagcc agcttgttac 26400

ccagcttctc agggagcaaa gaacagcctt gtttctatct tatccccact gtcccctgcc 26460

cctgccccac ctcccagcca ttcagcttct ggcttcccca gagctgcctg cttctttgtg 26520

gtcctccatt ccttgaaaag accttctagt cattagtgta tataaatggc cacttagccc 26580

agattacagt gaggtcaaca gctggggctc tgagaattgt cacacactgg cacaggagag 26640

gaggctattc ttccagagaa tttggagggc actcccatcc acttacaaca aaaagcccat 26700

ccactgtgct tggcagtagg tgatctgaga accaatggaa ccaggttaat cctgtggcac 26760

tgttgagtga ggagagcagt ggcgggcact ggaaaatatc agagacaagg caggagacct 26820

gaaatctagg cttagctcct catatacttg gcagctgtat gacctcagac aaccagtgtt 26880

acctctctaa gcctcagttt cctcatgcaa aaggaggggg aataacaaca gagcccactg 26940

cttgggggtg ttgtgaggac aggatgaaaa aacaaacaga aatccctcag tacaggattc 27000

agtgcagtgg acagtcttgc aaggtctggt tcagccctcc acccctaccc tcaccagtat 27060

aaagaactct ggcctacaag tcagatgacc tgagttttaa tctcagcttt gccattagcc 27120

gtgtgaactt gagaaagtcc ctttcctttt tacatctatt gggatgatca tgcatttttt 27180

gtcctttatt ctgttaatat agtgtgttac attgattgct tttcatagac tgaaccagcc 27240

ttgtattcca gggataaatc tcacttggtc atggtgtata atcctttata caaatgttgc 27300

tgggttgagt ttgctagtat tttgttgaag atttttatgt cttgattcat aaggaatatt 27360

ggtgtacctt ccccttttat ggccacagtt tccctacaat gatgtagtcg aactagacaa 27420

cctccaatat ctttcagtat tcatgtcctc tgattctgtg aaactaagaa aattaagaaa 27480

tagtgattca taggcacaag gcaggcaaaa cttagactcc ttgtagaata attaggaagc 27540

caaatattca gtgtgcttat ttctcaaata accttagttt ctccagtctg ccccaactcc 27600

gaggcctgaa tatctctaga tgcttatgat ggcaactaaa gcctaaaagc taattcattt 27660

taaagttctt ccaaatgcat agggttttat ttttccagac ctgggttcag atggggaatt 27720

tgacaaacaa tggaaagggg gaaaaacaac aatctaaaca ctgagtgaca aagtaacaaa 27780

gaaatagtct agctatcagc cagtcaagcc agccttggct ttgctatcca aagtagtcag 27840

tctaattcta ccaccagttt ctgttcctgt agctgtctac tgcctgccag ggactctgcc 27900

ttcccaccca caactaccaa tggaaggatg tggtgaccat accagtggct gctgacatct 27960

cctgccatgg gaagcataat tgcctccagc agcctccccc ttagatccat catttttgtt 28020

gcacttggcc tgggctgtac tcccggccaa tgactgaaca tggtgagcat agtaatgcag 28080

gcccatttct gtgaggagca ggactcctcc agtaggtgac tttggctcaa ggactctcta 28140

ttggcctggt tgaacttttc ctgaactgtg ctactgtctg agactcttct tacccaatcc 28200

tctttctcgc cccaattgtc acagaccacc tgcattgtgg tctgagtctc tccccacctt 28260

ctcttgctct tccctgttta tctttcacag gcatttcccc cagtacattc cttgaatgtc 28320

taacccgata cgggtgcctg acttttggca gacctaagca gacaaaaagg agtacttggt 28380

tacctagctc ttctttctac cacaaacatc gagggaaccc tttttccctc acccctctgc 28440

cacaccccca ctgccccagt gaacaaccac agagagagct gtggtataat attaggctgg 28500

tgcaaaagta attgcggttt ttgccattac ttttaatggt aaaaaccgca attacttttg 28560

cacctaccta gtatttgtgt ccccccaaat tcatatgttg aaacctaacc cacaatatga 28620

tgtcattagg aggcaagacc ttgaggaggt gattagatga tggggtggag ctctcctgaa 28680

tgagattagt gcccttataa gaagaagccc aaggaagcta ccttgactct tccatcacat 28740

gagaatgcag caagaaggca ccatctacta atcaggaaga gagctctcac cagacactga 28800

atctgccagt gtcttgatct tgaagttccc agcctccaga actatgcata atgcatttcc 28860

attgtctcta agccacccag cctatggtat tttgtcatag cagcctgaac tgactaagac 28920

agtgagccac atgagaagtg ccccaacccc tcccttaagc acttggctca cagatcagtg 28980

ggttcatttc tgcctgagtt ttattgttat tctgtagatt tcttgggcta gatatatttt 29040

tctgttattt tccttcttca cctcagtcat gaattggttg ttttaaaaaa gacaatgtaa 29100
gtcatgggga aactcctgac aactctactc tcctagggtt cctgataaaa ggggattcag 29160
ttgagtcctc tgatggtctc tacctgccaa agtccagcag cccttagcaa acatgctgct 29220
cgtttctgta gagaaggtgc tggtgtccca ccatacttct ctctccctca tgaagggctt 29280
gcgacccagc aaatgggtgg cttatatggg tctgtttcaa aggaagagcc agctctggga 29340
agaaaaacga tgagcataag cataacctac cactgtgcct gggaaagcag acaacttttt 29400
tgatgtgtga atatctaatg agaatggaat ccatcaatta ccttaaactt aggcacagtc 29460

ttcaaattca atatatgtgg gatatacttt tagtcagttt gtagacgtta tttgtaataa 29520
ataatctggc ttctctaaag aaattatttt aagtgtttgg tttggtttga tttaatggta 29580
aaattatatt tagtggcaga gaattatagc aatggtgata aactatagag tgtcataagt 29640
tcatatctta ttctcacatt tgaagctgcc tgcagatgca ttcaagatgc agccagaagt 29700
caggagactc aggctgttat ttggagctca tcattttaca gccttgctgg actcccactt 29760
tctcagggga aaaatgtggt gttgacccag attagctctc caggccctgc tgagttgggc 29820
actctgtaag ctggagggtc ttctattgtc ttcacctaag tgtcaatcaa caacccaaat 29880

gggcatgggg gaagagggag ctgggccaat gcccagggtg cctggtagag agataccttg 29940
ggcactggaa ggcaccagct tcccagagag aaggggggagg gccatgaaaa agttggctgt 30000
agatgccagg gacactggga ctctccagct gtgtgtttgt gccttctgaa gacttatgtt 30060
tcattccttt ggagcatgca taatcataca ctgtgggatg tgttatatag attgcttgat 30120
agttcaccac tgtaataaaa tactgtgact ggaatctgct cccagtctgc ctttgatagc 30180
acttgtgcaa cacacattta ctgagcattt acagtgatcc aggacctgtg ttgtgaaaac 30240
attgatggac aaggcagatg gtggagcacg tcagtgagga tttttaacaa aggctggtaa 30300

gtgctataaa ggaacattgt aggacactag agaacaaaga acaggagaac ctgacttagg 30360
ctggggtggg gcgttggtta gaggaggctc cttggaggac atgaggttta agctgtgacc 30420
tgaggatgaa tagatgttgg ccagtgaggt taattaacct caaatcgtca cttccgtttt 30480
cccacgttac gtcacttccc attttaagaa aactacaatt cccaacacat acaagttact 30540
ccgccctaaa acctacgtca cccgccccgt tcccacgccc cgcgccacgt cacaaactcc 30600
acccccctcat tatcatattg gcttcaatcc aaaataaggt atattattga tgatg 30655

<210> 64

<211> 32501

<212> DNA

<213> Artificial Sequence

<220><223> LCR-F8 (Ad5, Sleeping beauty IR, HS4 to HS1, beta promoter, ET3, mgmt, EF1a promoter, and Ad5)

<400> 64

catcatcaat aatatacctt attttggatt gaagccaata tgataatgag ggggtggagt 60

ttgtgacgtg gcgcggggcg tgggaacggg gcgggtgacg tagtagtgtg gcggaagtgt 120

gatgttgcaa gtgtggcgga acacatgtaa gcgacggatg tggcaaaagt gacgtttttg 180

gtgtgcgccg gtgtacacag gaagtgacaa ttttcgcgcg gttttaggcg gatgttgtag 240

taaatttggg cgtaaccgag taagatttgg ccattttcgc gggaaaactg aataagagga 300

agtgaaatct gaataatttt gtgttactca tagcgcgtaa tatttgtcta gggccgcggg 360

gactttgacc gtttacgtgg agactcgccc aggtgttttt ctcaggtgtt ttccgcgttc 420

cgggtcaaag ttggcgtttt gattcggccg aaggattaca tgagcttaga aatgtaatta 480

gcatagtgtg tggcatagtg tagataccaa ataaatatga tctctccttc tactcttgaa 540

aatgcaaaca cattcttggt ggtcctaaaa tagcctgtaa catggtttac tcagcagcat 600

ttgctattca aggcagatct gcctttagtc attggctgcg ctcctgaaca gctgtgtgaa 660

aggctaactt ttgtaaacca aatcaaaata aaatgcagca aaaatttgtc actgaaagga 720

aatcctcagt atatcctttt atgaaatgaa agatccctca tccaaactta acttttttaa 780

aagtgcgcat ttggagatat agcccttttct tatgaatcct aattcaattt tggccataaa 840

cacacgttga tgttccccac cccaaagcac atagcaacaa gagtaggttc tatattgaaa 900

ataatgacaa tttaaaaaca tgtacttatt tcactgtatg tggacagtgt ctatgattgc 960

atcatgaagt gtcatataac catgtacgtg tacatgagag agagatagag agagaagtgg 1020

tagggtggtg gtggtagagg ggatggcgat agtaatcatg gtaatggtag aggtgatgga 1080

ggtggtaatg acggaggtaa gggtggtagt gatgatggtg gtggtggtaa tggtggtgga 1140

tgtggtggtg gcaattggga tggtgggatg gtggtagcca tggtgatggt ggtaatggtg 1200

ttgatttaaa gggtggtggt agtgaaggtg agggtagtgg tggtggaggt ggtggtgctg 1260

gtagcaatag tgatggtggt gatggtgttg atgagggtgt tgggatcagg gtgagttccc 1320

acagtatatt tcattcttgt tgtaccactc tgtcaacagc accactgact gggacagagg 1380

aagaaggcac actctgaatg tgttattaac agaaacctca aaacagtctg tctccttgta 1440

gtcattcaaa attatctttt tcttacctgg aaaactgaaa ctgaattacc gggaaaaaca 1500

caggagattt ttgtttgtta atatgctgcc aataaagtaa ttttatgtca aatttaacta 1560

caggaaaggg caaggcattt tctaagttcc ttagatgtca tgtggctaaa aaaaacaaaa 1620
ggatggacag cagttagata ctgtacactt agctgtttga agccatatat tcagaaagca 1680
gatgttggga gttggtgttt gaggactgat ttcctggagg tattttatat aggccaagtt 1740
cattgttcta aactctaagg gcttgacttg agggaggaaa agaggcaaga acatgtttag 1800
ttttgctgac agcatcacat gggcagccct aaggctagac aactttaggg cctgaagctt 1860
attctaggaa agaagcacct acagagtggc actgggctcc cctccactat agagatgaag 1920
tcatatgaca gtaaagggca ggcagggctg cctaggggggc ccattgaaat tgcggccgca 1980

aataatgggc ccggagcaag agagagggag gcaatgacag cagagacatg cctgcgcctt 2040
gggtttgagt gcccagtggt caaatccact tccctgtggc tgatgcttgc ctttctaact 2100
ttggaattta ggggttggag atctggtgag aaggtaggag ggagatgagg aggagaaggg 2160
aaaggcagga aggaagggga gggaaaggaa aagcaaaagg ggaggaggaa ggtttccaac 2220
aaattattct atatcaactg cggaaatcaa aatttgttgc ccaaatctta gaagctcatg 2280
tccctcctcc ccagaagtct ggaatgcagc actccagggg tagcttataa cccaaatatc 2340
tatctgtaaa aagagaaaca ttgggctttc gagctgtgga ttctcagtaa aagcaagagg 2400

cctcagccta cacaggccag cccagagttt gaggaacccc aggcccacac ccacagggct 2460
ggcccctggg tctgcatact ccctagaaat gtgcacactt ctgagcctca actctgtcct 2520
ggagtctaac agcatccctc tccttcctgg ggcagttcca cctccagaaa cctgttacct 2580
tgggccttat gtcaaggaaa ctgtgggaaa gagctaggca ggaatgcaga tgaggccagc 2640
atgggctcct aaaagtttag aaataggcag tgtcatgctc ccaggtgcct gcataaacca 2700
gctgaaaaat ggagctcccc tcaccagcac tctcccttca aacagactgt gatttgcagg 2760
tcactggttt accaagccag gctacccagg caggacccag atgccaagcc cagtggtgtc 2820

ctgcaagctg agcagtgctc agttcttgca aaaaaaggtc tgtgtgaagg caaggcctct 2880
gcctggcttc tcaccccagt tgggtgtctg gaacaggaag gagcccttac tgcagaaaaa 2940
ggaggaggga gcaaagggag cgaacagctg cgtgctccat ggggaggatc cccaaagtag 3000
aaaggcgcat acacactgca gcccttgacc cagaatgctc acagctacat tacagattca 3060
ggtctcctca gtgtagtggg gctgctgatg agactgtggc atcctcaggg gtcaggacac 3120
acattttcca tcactcttct gatggcaaaa aacctctgag ccaatgccaa cctctgatca 3180
ttaaaaaaaa gtgctcacag cagtgtgtgg tttaggatca tgccctgtgt ggtttggaac 3240

acgtgcacaa ccacaccttg ttcatcacca tcccagaaac cctgacgcag gcaaagagca 3300

gagttattaa ccctacttta ctgatgtgga tactgaggcc cagaggctca tgcaagttat 3360

caataagtgg cagggacagt tgcctctaga ttaactagcc cctaggatca cctgggtctt 3420

ggaaggggac ccataaacat gagctcccct ctcttggggc cagatttgca cctgtgccgc 3480

gccttcagcc tgcatgaagt aggggctgct ggcaaagact caaagctgta aatctgggtt 3540

ttctcttgag gcttctaagg gagctgtttc gacaactcac tctgttccca gctggctgcc 3600

cctgcatagg gttttaaagc agcctagctt tctgccaggc ttggcagtgg acaacgctgg 3660

tcagaacatc ccagagagct accagaatga agtaagtttg cttctactct ttacctgttt 3720

atgggctgtc tctgccactg gaatgaaagg cactgagaac agtgcctggc ctgcagaagg 3780

ccctggaaat acctgagctc ctaatctggg aataggagta ggaagagctt tggaggcagg 3840

gcacctgagt ttgagatcta caacttcctg cctgtgtgac attgggaaag tctccatcct 3900

ttctgagcct cagtctccac cctggggaag tggaaatatc aatctctgtg acacagaagc 3960

aaatgagcga atgtgcacaa agtaccttgc acaagagaga cgctcaaaca cttgcctcca 4020

ggtttcaccg agaactacag agtaagatag atttgttccc agtggaggaa gcctgggaat 4080

aatttgcccc tagactatga attcctgggg ctcaagatcg agcacagggc caggcacaca 4140

gaagggaccc tggaaatgtg gcaggaggcc agagatagac aggcccttag agctcatacc 4200

catgccctct gacctcaaga agaaagaaac ctgctcaaaa tctcacaaag agcttgttcc 4260

aaccctgaat cgagtctgag gactccttcc tgagtccagc acttttctg caagaagtat 4320

atgcctccaa agctgatggg cgcaaatctt gaaccccgtc acataaacac aaagggagga 4380

ggtgactaga gctcctccta ctggatatgt ctaaggtcac cagtctaaag aaaagggatg 4440

gatagaatga ggccagtatt tttgcagcca tccaaatgtc cacatacgct gttacactga 4500

gggctcctct ctcccccgtc ttcagcccta cttgcattta gaggtgagaa agatatgggc 4560

tgaggggttg tttttcatcg tattgtagat ggaaagcaca ctgcccttgg ggccatccaa 4620

atgtggacct tgatgtagca ccccaccttc tggatggcca tccttctgaa agtcactgaa 4680

tttctcagac tttattctct ttatccataa agaaggagaa taataataat ccccccaccc 4740

tgcccaacca ctgactggtt gggaagctca gaagaaatac tgggcacggc atcccattgt 4800

aatctataga gtgagtcgct tcttaatatt aaatggctga acacagaaga tgtgcaaaaa 4860

gtactgtgtc cccttcctcc tccaactgaa catttcatgc cctttgcacc ctcattttgt 4920

ctaggagctg ccttatgaag ggaataggta cctgctccga gctggaggaa tctttgccac 4980

ttatggtggg gtatggactg agacagagat ggcatgtgac atgcgcactg agtctcaact 5040

ccatgcaggc tctggagcac tctcaaattg gagtactaat gccttttaaa ttctcacact 5100

agcaatcctt tgacctactg atctagggat ctagggaaag aatcgtgatc ttaacttcaa 5160

agggaaggac aaaatgttct gcctcctgtt aaaactccat acactaagtg cagagactgg 5220

atgccttatt aaccttgggt agatgcccaa atgttcaaaa ggtcaaactc ttctgttccc 5280

cagatcgcca gagtcattaa ccagtcacac tattaaatga atgaacagat gctgaaaagg 5340

tacttgcatt actgagattt cttatggtga tggcccctgc ctgatatgta ttcagcattt 5400

tgtagttttc aatgtgcatt agagtatagt ggtgatgaca ttggcctctg agtttgccac 5460

ttcttatatc tgtgactttg gtcaaattgc ttaatctctc tgagtctcgg tttcctggag 5520

ataataatag cttcttcttc ccagggttat catgaggatt acaggagata atgccccaaa 5580

aatgcttagt aaagtgccta gcacctagtc aatgctgaat taaaggtggt tattcttact 5640

tttcgttcat ttgaactttg ttctcaggga gggcaaagga tagacaaagc cccatagcta 5700

gtgaggagta gctgcaagac tagaacccag gtgttctgag ccctagtctt aggccaagaa 5760

caactgttac gtgagatgca cgttttcctt caagggagct cacaattatt tccatgtaaa 5820

ttcaaggact gctaaaagag aactctcctc tgggactgat aaacatctag tcgagtatcg 5880

acaaataggg gttccgcgca catttccccg aaaagtgcca cctgacgtaa ctataacggt 5940

cctaaggtag cgaaagcttg catgcctgca ggtcgactct agaggatcaa acctagccac 6000

cgcggtggcg gccggctagc cggctagccg gctagcccta gaactagtaa cggccgccag 6060

tgtgctggaa ttcggcttgt aaggtaccgg tgaagttcct atactttcta gagaatagga 6120

acttcggaat aggaacttct acctagatgc atgctcgagc ggcccctaca gttgaagtcg 6180

gaagtttaca tacacttaag ttggagtcat taaaactcgt ttttcaacta ctccacaaat 6240

ttcttgttaa caaacaatag ttttggcaag tcagttagga catctacttt gtgcatgaca 6300

caagtcattt ttccaacaat tgtttacaga cagattattt cacttataat tcactgtatc 6360

acaattccag tgggtcagaa gtttacatac actaagttga ctgtgccttt aaacagcttg 6420

gaaaattcca gaaaatgatg tcatggcttt agaagcttct gatagactaa ttgacatcat 6480

ttgagtcaat tggaggtgta cctgtggatg tatttcaagg aattctgtgg aatgtgtgtc 6540

agttagggtg tggaaagtcc ccaggctccc cagcaggcag aagtatgcaa agcatgcata 6600

tcgatactag tttataagat ctcgagctag cggccgtttg ttaattaagt cgacggtacc 6660

gtcaaggctg cagtgagaca tgatcttgcc actgcactcc agcctggaca gcagagtgaa 6720

accttgcctc acgaaacaga atacaaaaac aaacaaacaa aaaactgctc cgcaatgcgc 6780

ttccttgatg ctctaccaca taggtctggg tactttgtac acattatctc attgctgttc 6840

ataattgtta gattaatttt gtaatattga tattattcct agaaagctga ggcctcaaga 6900

tgataacttt tattttctgg acttgtaata gctttctctt gtattcacca tgttgtaact 6960

ttcttagagt agtaacaata taaagttatt gtgagttttt gcaaacacag caaacacaac 7020

gacccatata gacattgatg tgaaattgtc tattgtcaat ttatgggaaa acaagtatgt 7080
actttttcta ctaagccatt gaaacaggaa taacagaaca agattgaaag aatacatttt 7140
ccgaaattac ttgagtatta tacaaagaca agcacgtgga cctgggagga gggttattgt 7200
ccatgactgg tgtgtggaga caaatgcagg tttataatag atgggatggc atctagcgca 7260
atgactttgc catcactttt agagagctct tggggacccc agtacacaag aggggacgca 7320
gggtatatgt agacatctca ttctttttct tagtgtgaga ataagaatag ccatgacctg 7380
agtttataga caatgagccc ttttctctct cccactcagc agctatgaga tggcttgccc 7440

tgcctctcta ctaggctgac tcactccaag gcccagcaat gggcagggct ctgtcagggc 7500
tttgatagca ctatctgcag agccagggcc gagaaggggt ggactccaga gactctccct 7560
cccattcccg agcagggttt gcttatttat gcatttaaat gatatattta ttttaaaaga 7620
aataacagga gactgcccag ccctggctgt gacatggaaa ctatgtagaa tattttgggt 7680
tccatttttt tttccttctt tcagttagag gaaaaggggc tcaggatcca cttgcccagt 7740
gttcttcctt agttcctacc ttcgaccttg atcctccttt atcttcctga accctgctga 7800
gatgatctat gtggggagaa tggcttcttt gagaaacatc ttcttcgtta gtggcctgcc 7860

cctcattccc actttaatat ccagaatcac tataagaaga atataataag aggaataact 7920
cttattatag gtaagggaaa attaagaggc atacgtgatg ggatgagtaa gagaggagag 7980
ggaaggatta atggacgata aaatctacta ctatttgttg agacctttta tagtctaatc 8040
aattttgcta ttgttttcca tcctcacgct aactccataa aaaaacacta ttattatctt 8100
tattttgcca tgacaagact gagctcagaa gagtcaagca tttgcctaag gtcggacatg 8160
tcagaggcag tgccagacct atgtgagact ctgcagctac tgctcatggg ccctgtgctg 8220
cactgatgag gaggatcaga tggatggggc aatgaagcaa aggaatcatt ctgtggataa 8280

aggagacagc catgaagaag tctatgactg taaatttggg agcaggagtc tctaaggact 8340
tggatttcaa ggaattttga ctcagcaaac acaagaccct cacggtgact ttgcgagctg 8400
gtgtgccaga tgtgtctatc agaggttcca gggagggtgg ggtggggtca gggctggcca 8460
ccagctatca gggcccagat gggttatagg ctggcaggct cagataggtg gttaggtcag 8520
gttggtggtg ctgggtggag tccatgactc ccaggagcca ggagagatag accatgagta 8580
gagggcagac atgggaaagg tgggggaggc acagcatagc agcatttttc attctactac 8640
tacatgggac tgctccccta taccccagc taggggcaag tgccttgact cctatgtttt 8700

caggatcatc atctataaag taagagtaat aattgtgtct atctcatagg gttattatga 8760

ggatcaaagg agatgcacac tctctggacc agtggcctaa cagttcagga cagagctatg 8820

ggcttcctat gtatgggtca gtggtctcaa tgtagcaggc aagttccaga agatagcatc 8880

aaccactgtt agagatatac tgccagtctc agagcctgat gttaatttag caatgggctg 8940

ggaccctcct ccagtagaac cttctaacca gctgctgcag tcaaagtcga atgcagctgg 9000

ttagactttt tttaatgaaa gcttgcatgc agcactttgg gaggctgagg tgggtggact 9060

gcttggagct caggagttca agaccatctt ggacaacatg gtgataccct gcctctacaa 9120

aaagtacaaa aattagcctg gcatggtggt gtgcacctgt aatcccagct attagggtgg 9180

ctgaggcagg agaattgctt gaacccggga ggcggaggtt gcagtgagct gagatcgtgc 9240

cactgcactc cagcctgggg gacagagcac attataatta actgttattt tttacttgga 9300

ctcttgtggg gaataagata catgttttat tcttatttat gattcaagca ctgaaaatag 9360

tgtttagcat ccagcaggtg cttcaaaacc atttgctgaa tgattactat acttttttaca 9420

agctcagctc cctctatccc ttccagcatc ctcatctctg attaaataag cttcagtttt 9480

tccttagttc ctgttacatt tctgtgtgtc tccattagtg acctcccata gtccaagcat 9540

gagcagttct ggccaggccc ctgtcggggt cagtgcccca cccccgcctt ctggttctgt 9600

gtaaccttct aagcaaacct tctggctcaa gcacagcaat gctgagtcat gatgagtcat 9660

gctgaggctt agggtgtgtg cccagatgtt ctcagcctag agtgatgact cctatctggg 9720

tccccagcag gatgcttaca gggcagatgg caaaaaaaag gagaagctga ccacctgact 9780

aaaactccac ctcaaacggc atcataaaga aaatggatgc ctgagacaga atgtgacata 9840

ttctagaata tattatttcc tgaatatata tatatatata cacatatacc atatgaaaca 9900

cctctaggct ataaggcaac agagctcctt ttttttttt ctgtgctttc ctggctgtcc 9960

aaatctctaa tgataagcat acttctattc aatgagaata ttctgtaaga ttatagttaa 10020

gaattgtggg agccattccg tctcttatag ttaaatttga gcttctttta tgatcactgt 10080

ttttttaata tgctttaagt tctggggtac atgtgccatg gtggtttgct gcacccatca 10140

acccgtcatc tacattaggt atttctccta atgctatcct tcccctagcc ccccaccccc 10200

aacaggcccc agtgtgtgat gttcccctcc ctgtgtccat ggatcactgg ttttttttttg 10260

ttttttttttt ttttttaaag tctcagttaa atttttggaa tgtaatttat tttcctggta 10320

tcctaaggac ttgcaagtta tctggtcact ttagccctca cgttttgatg ataatcacat 10380

atttgtaaac acaacacaca cacacacaca cacacacata tatatatata taaaacatat 10440

atatacataa acacacataa catatttatc gggcatttct gagcaactaa tcatgcagga 10500

ctctcaaaca ctaacctata gccttttcta tgtatctact tgtgtagaaa ccaagcgtgg 10560

ggactgagaa ggcaatagca ggagcattct gactctcact gcctttagct aggcccctcc 10620

ctcatcacag ctcagcatag tcctgagctc ttatctatat ccacacacag tttctgacgc 10680

tgcccagcta tcaccatccc aagtctaaag aaaaaaataa tgggtttgcc catctctgtt 10740

gattagaaaa caaaacaaaa taaaataagc ccctaagctc ccagaaaaca tgactaaacc 10800

agcaagaaga agaaaataca ataggtatat gaggagactg gtgacactaa gtgtctgaat 10860

gaggcttgag tacagaaaag aggctctagc agcatagtgg tttagaggag atgtttcttt 10920

ccttcacaga tgccttagcc tcaataagct tgcggttgtg gaagtttact tgtttatcac 10980

cggtgacgtc catgagcaaa ttaagaaaaa caacaacaaa tgaatgcata tatatgtata 11040

tgtatgtgtg tatatataca catatatata tatatttttt ttcttttctt accagaaggt 11100

tttaatccaa ataaggagaa gatatgctta gaactgaggt agagttttca tccattctgt 11160

cctgtaagta ttttgcatat tctggagacg caggaagaga tccatctaca tatcccaaag 11220

ctgaattatg gtagacaaaa ctcttccact tttagtgcat caatttctta tttgtgtaat 11280

aagaaaattg ggaaaacgat cttcaatatg cttaccaagc tgtgattcca aatattacgt 11340

aaatacactt gcaaaggagg atgtttttag tagcaatttg tactgatggt atggggccaa 11400

gagatatatc ttagagggag ggctgagggt ttgaagtcca actcctaagc cagtgccaga 11460

agagccaagg acaggtacgg ctgtcatcac ttagacctca ccctgtggag ccacacccta 11520

gggttggcca atctactccc aggagcaggg agggcaggag ccagggctgg gcataaaagt 11580

cagggcagag ccatctattg cttacatttg cttctgacac aactgtgttc actagcaacc 11640

tcaaacagac acctcgagtt tcgattgcta gcttcgaact cgagccacca tgcagctaga 11700

gctctccacc tgtgtctttc tgtgtctctt gccactcggc tttagtgcca tcaggagata 11760

ctacctgggc gcagtggaac tgtcctggga ctaccggcaa agtgaactcc tccgtgagct 11820

gcacgtggac accagatttc ctgctacagc gccaggagct cttccgttgg gcccgtcagt 11880

cctgtacaaa aagactgtgt tcgtagagtt cacggatcaa cttttcagcg ttgccaggcc 11940

caggccacca tggatgggtc tgctgggtcc taccatccag gctgaggttt acgacacggt 12000

ggtcgttacc ctgaagaaca tggcttctca tcccgttagt cttcacgctg tcggcgtctc 12060

cttctggaaa tcttccgaag gcgctgaata tgaggatcac accagccaaa gggagaagga 12120

agacgataaa gtccttcccg gtaaaagcca aacctacgtc tggcaggtcc tgaaagaaaa 12180

tggtccaaca gcctctgacc caccatgtct tacctactca tacctgtctc acgtggacct 12240

ggtgaaagac ctgaattcgg gcctcattgg agccctgctg gtttgtagag aagggagtct 12300

gaccagagaa aggacccaga acctgcacga atttgtacta ctttttgctg tctttgatga 12360

agggaaaagt tggcactcag caagaaatga ctcctggaca cgggccatgg atcccgcacc 12420

tgccagggcc cagcctgcaa tgcacacagt caatggctat gtcaacaggt ctctgccagg 12480

tctgatcgga tgtcataaga aatcagtcta ctggcacgtg attggaatgg gcaccagccc 12540
ggaagtgcac tccatttttc ttgaaggcca cacgtttctc gtgaggcacc atcgccaggc 12600
ttccttggag atctcgccac taactttcct cactgctcag acattcctga tggaccttgg 12660
ccagttccta ctgttttgtc atatctcttc ccaccaccat ggtggcatgg aggctcacgt 12720
cagagtagaa agctgcgccg aggagcccca gctgcggagg aaagctgatg aagaggaaga 12780
ttatgatgac aatttgtacg actcggacat ggacgtggtc cggctcgatg gtgacgacgt 12840
gtctcccttt atccaaatcc gctcagttgc caagaagcat cctaaaactt gggtacatta 12900

cattgctgct gaagaggagg actgggacta tgctccctta gtcctcgccc ccgatgacag 12960
aagttataaa agtcaatatt tgaacaatgg ccctcagcgg attggtagga agtacaaaaa 13020
agtccgattt atggcataca cagatgaaac ctttaagacg cgtgaagcta ttcagcatga 13080
atcaggaatc ttgggacctt tactttatgg ggaagttgga gacacactgt tgattatatt 13140
taagaatcaa gcaagcagac catataacat ctaccctcac ggaatcactg atgtccgtcc 13200
tttgtattca aggagattac caaaaggtgt aaaacatttg aaggattttc caattctgcc 13260
aggagaaata ttcaaatata aatggacagt gactgtagaa gatgggccaa ctaaatcaga 13320

tccgcggtgc ctgacccgct attactctag tttcgttaat atggagagag atctagcttc 13380
aggactcatt ggccctctcc tcatctgcta caaagaatct gtagatcaaa gaggaaacca 13440
gataatgtca gacaagagga atgtcatcct gttttctgta tttgatgaga accgaagctg 13500
gtacctcaca gagaatatac aacgctttct ccccaatcca gctggagtgc agcttgagga 13560
tccagagttc caagcctcca acatcatgca cagcatcaat ggctatgttt ttgatagttt 13620
gcagttgtca gtttgtttgc atgaggtggc atactggtac attctaagca ttggagcaca 13680
gactgacttc ctttctgtct tcttctctgg atataccttc aaacacaaaa tggtctatga 13740

agacacactc accctattcc cattctcagg agaaactgtc ttcatgtcga tggaaaaccc 13800
aggtctatgg attctggggt gccacaactc agactttcgg aacagaggca tgaccgcctt 13860
actgaaggtt tctagttgtg acaagaacac tggtgattat tacgaggaca gttatgaaga 13920
tatttcagca tacttgctga gtaaaaacaa tgccattgaa cctaggagct ttgcccagaa 13980
ttcaagaccc cctagtgcga gcgctccaaa gcctccggtc ctgcgacggc atcagaggga 14040
cataagcctt cctacttttc agccggagga agacaaaatg gactatgatg atatcttctc 14100
aactgaaacg aagggagaag attttgacat ttacggtgag gatgaaaatc aggaccctcg 14160

cagctttcag aagagaaccc gacactattt cattgctgcg gtggagcagc tctgggatta 14220

cgggatgagc gaatcccccc gggcgctaag aaacagggct cagaacggag aggtgcctcg 14280

gttcaagaag gtggtcttcc gggaatttgc tgacggctcc ttcacgcagc cgtcgtaccg 14340

cggggaactc aacaaacact tggggctctt gggaccctac atcagagcgg aagttgaaga 14400

caacatcatg gtaactttca aaaaccaggc gtctcgtccc tattccttct actcgagcct 14460

tatttcttat ccggatgatc aggagcaagg ggcagaacct cgacacaact tcgtccagcc 14520

aaatgaaacc agaacttact tttggaaagt gcagcatcac atggcaccca cagaagacga 14580

gtttgactgc aaagcctggg cctacttttc tgatgttgac ctggaaaaag atgtgcactc 14640

aggcttgatc ggcccccttc tgatctgccg cgccaacacc ctgaacgctg ctcacggtag 14700

acaagtgacc gtgcaagaat ttgctctgtt tttcactatt tttgatgaga caaagagctg 14760

gtacttcact gaaaatgtgg aaaggaactg ccgggccccc tgccatctgc agatggagga 14820

ccccactctg aaagaaaact atcgcttcca tgcaatcaat ggctatgtga tggatacact 14880

ccctggctta gtaatggctc agaatcaaag gatccgatgg tatctgctca gcatgggcag 14940

caatgaaaat atccattcga ttcattttag cggacacgtg ttcagtgtac ggaaaaagga 15000

ggagtataaa atggccgtgt acaatctcta tccgggtgtc tttgagacag tggaaatgct 15060

accgtccaaa gttggaattt ggcgaataga atgcctgatt ggcgagcacc tgcaagctgg 15120

gatgagcacg actttcctgg tgtacagcaa gaagtgtcag actcccctgg gaatggcttc 15180

tggacacatt agagattttc agattacagc ttcaggacaa tatggacagt gggccccaaa 15240

gctggccaga cttcattatt ccggatcaat caatgcctgg agcaccaagg agccctttc 15300

ttggatcaag gtggatctgt tggcaccaat gattattcac ggcatcaaga cccagggtgc 15360

ccgtcagaag ttctccagcc tctacatctc tcagtttatc atcatgtata gtcttgatgg 15420

gaagaagtgg cagacttatc gaggaaattc cactggaacc ttaatggtct tctttggcaa 15480

tgtggattca tctgggataa aacacaatat ttttaaccct ccaattattg ctcgatacat 15540

ccgtttgcac ccaactcatt atagcattcg cagcactctt cgcatggagt tgatgggctg 15600

tgatttaaat agttgcagca tgccattggg aatggagagt aaagcaatat cagatgcaca 15660

gattactgct tcatcctact ttaccaatat gtttgccacc tggtctcctt caaaagctcg 15720

acttcacctc caagggagga gtaatgcctg gagacctcag gtgaataatc caaaagagtg 15780

gctgcaagtg gacttccaga agacaatgaa agtcacagga gtaactactc agggagtaaa 15840

atctctgctt accagcatgt atgtgaagga gttcctcatc tccagcagtc aagatggcca 15900

tcagtggact ctctttttc agaatggcaa agtaaaggtt tttcagggaa atcaagactc 15960

cttcacacct gtggtgaact ctctagaccc accgttactg actcgctacc ttcgaattca 16020

ccccagagt tgggtgcacc agattgccct gaggatggag gttctgggct gcgaggcaca 16080

ggacctctac tgagggcggc cgccacccca ccagtgcagg ctgcctatca gaaagtggtg 16140

gctggtgtgg ctaatgccct ggcccacaag tatcactaag ctcgctttct tgctgtccaa 16200

tttctattaa aggttccttt gttccctaag tccaactact aaactggggg atattatgaa 16260

gggccttgag catctggatt ctgcctaata aaaaacattt attttcattg caatgatgta 16320

tttaaattat ttctgaatat tttactaaaa agggaatgtg ggaggtcagt gcatttaaaa 16380

cataaagaaa tgaagagcta gttcaaacct tgggaaaata cactatatct taaactccat 16440

gaaagaaggt gaggctgcaa acagctaatg cacattggca acagcccctg atgcctatgc 16500

cttattcatc cctcagaaaa ggattcaagt agaggcttga tttggaggtt aaagttttgc 16560

tatgctgtat tttacattac ttattgtttt agctgtcctc atgaatgtct tttcactacc 16620

catttgctta tcctgcatct ctcagccttg actccactca gttctcttgc ttagagatac 16680

caccttccc ctgaagtgtt ccttccatgt tttacggcga gatggtttct cctcgcctgg 16740

ccactcagcc ttagttgtct ctgttgtctt atagaggtct acttgaagaa ggaaaaacag 16800

ggggcatggt ttgactgtcc tgtgagccct tcttccctgc ctcccccact cacagtgacc 16860

cggaatctgc agtgctagtc tcccggaact atcactcttt cacagtctgc tttggaagga 16920

ctgggcttag tatgaaaagt taggactgag aagaatttga aaggggggctt tttgtagctt 16980

gatattcact actgtcttat taccctatca taggcccacc ccaaatggaa gtcccattct 17040

tcctcaggat gtttaagatt agcattcagg aagagatcag aggtctgctg gctcccttat 17100

catgtccctt atggtgcttc tggctctgca gttattagca tagtgttacc atcaaccacc 17160

ttaacttcat ttttcttatt caatacctag aaagcttatc gaccccatcc tcactgactc 17220

cgtcctggag ttggatgaga gataatggcc ttacgttgtg ccaggggagg gtcgggctgg 17280

atttagcaag atttaccttc tccaaagagc ggtgctgcag tggcacagct gcccacggag 17340

gtgggggggt caccgtccct ggaggtgatg aagaactgtg gggatgtggc actgagggac 17400

atggccagtg ggcacggtgg gtgggttggg gttggtcttg gggatcttgg agggcttttc 17460

cagccttcat gatttgacga ttgtatgaac atctacatgg caattctcca gctgcctgtc 17520

ccagtcctac tgacccagct gtatctctcc aggcaagctc ttccaccct tctgcttgca 17580

tccagacacc atcaaacatg caggctcaga cacagggacc agcagtgtct gtggcctttt 17640

tgtgctcctc tccatgctgg gttttaactt gctctttgtc cttctatcct atcttcttat 17700

ccttaaggct gttctgaacg ctgtgacttg gagagtgtcc cagagccctc aacacctgca 17760

tgtcccacgt ccatgctgtc ctgcacttcc ttatccccaa gatctgcctc tccgtgatgc 17820

actgaattgg caaacatgtg tcaccccaga ccaacaatgt cacagcaaac tccccttga 17880

taggacaagg gggaatggct ttacactgag acaggggagg tttgggttgg atatgaggag 17940

gcagtttttc ccccagaggg tggtgacgca ctgaacaggt tgcccaagga ggctgtggat 18000
gccccatccc tgcaggcatt caaggccagg ctggatgtgg ctctgggcag cctgggctgc 18060
tggttgatga ccctgcacat agcaggggggt tggatctgga tgagcactgt gctcctttgc 18120
aacccaggcc gttctatgat tctgtcattc taaatctctc tttcagccta aagctttttc 18180
cccgtatccc cccaggtgtc tgcaggctca aagagcagcg agaagcgttc agaggaaagc 18240
gatcccgtgc caccttcccc gtgcccgggc tgtccccgca cgctgccggc tcggggatgc 18300
ggggggagcg ccggaccgga gcggagcccc gggcggctcg ctgctgcccc ctagcggggg 18360

agggacgtaa ttacatccct gggggctttg gggggggggct gtccctatcg attactagtt 18420
ttaccacatt tgtagaggtt ttacttgctt taaaaaacct cccacacctc cccctgaacc 18480
tgaaacataa aatgaatgca attgttgttg ttaacttgtt tattgcagct tataatggtt 18540
acaaataaag caatagcatc acaaatttca caaataaagc attttttttca ctgcattcta 18600
gttgtggttt gtccaaactc atcaatgtat cttatcatgt ctgctcgaag agcggccgct 18660
tcagtttcgg ccagcaggcg gggagcccga ggtagctccc gctcccttga gccaggcccc 18720
tgccagacct gagctccctc ccaagcctgg cttccccaac cggtggcctt catgggccag 18780

aagccattcc ttcacggcta gccctccgga gtagttgccc acggctccgc tgctgcagac 18840
cactctgtgg cacgggatga ggatcttgac aggattgcct ctcatggcgc ctcccactgc 18900
tcgcgcggct ttggggttgc cggccagggc ggccaattgc tggtaagaaa tcacttctcc 18960
gaatttcaca accttaagca gcttccataa cacctgacgc gtgaacgact cttgctggaa 19020
aacgggatgg tgaagcgctg gcacggggaa ctcttcgata gcctcgggct ggtggaaata 19080
ggcattcagc caggctgtgc actgcatcag gggctccgga cctccgagaa ccgcagcggg 19140
ggctgggacc tccacggcat cagctgcaga cgtccccttg cccaggagct ttatttcgtg 19200

cagaccctgc tcacaaccag acagctccag cttccccaaa gggctgtcca gtgtggtgcg 19260
tttcatttca caatccttgt ccatggtggc gaccgtctag ctcacgacac ctgaaatgga 19320
agaaaaaaac tttgaaccac tgtctgaggc ttgagaatga accaagatcc aaactcaaaa 19380
agggcaaatt ccaaggagaa ttacatcaag tgccaagctg gcctaacttc agtctccacc 19440
cactcagtgt ggggaaactc catcgcataa aacccctccc cccaacctaa agacgacgta 19500
ctccaaaagc tcgagaacta atcgaggtgc ctggacggcg cccggtactc cgtggagtca 19560
catgaagcga cggctgagga cggaaaggcc cttttccttt gtgtgggtga ctcacccgcc 19620

cgctctcccg agcgccgcgt cctccatttt gagctccctg cagcagggcc gggaagcggc 19680

catctttccg ctcacgcaac tggtgccgac cgggccagcc ttgccgccca gggcggggcg 19740

atacacggcg gcgcgaggcc aggcaccaga gcaggccggc gagcttgaga ctacccccgt 19800

ccgattctcg gtggccgcgc tcgcaggccc cgcctcgccg aacatgtgcg ctgggacgca 19860

cgggccccgt cgccgcccgc ggccccaaaa accgaaatac cagtgtgcag atcttggccc 19920

gcatttacaa gactatcttg ccagaaaaaa agcgtcgcag caggtcatca aaaattttaa 19980

atggctagag acttatcgaa agcagcgaga caggcgcgaa ggtgccacca gattcgcacg 20040

cggcggcccc agcgcccagg ccaggcctca actcaagcac gaggcgaagg ggctccttaa 20100

gcgcaaggcc tcgaactctc ccacccactt ccaacccgaa gctcgggatc aagaatcacg 20160

tactgcagcc aggggcgtgg aagtaattca aggcacgcaa gggccataac ccgtaaagag 20220

gccaggcccg cgggaaccac acacggcact tacctgtgtt ctggcggcaa acccgttgcg 20280

aaaaagaacg ttcacggcga ctactgcact tatatacggt tctcccccac cctcgggaaa 20340

aaggcggagc cagtacacga catcactttc ccagtttacc ccgcgccacc ttctctaggc 20400

accggttcaa ttgccgaccc ctccccccaa cttctcgggg actgtgggcg atgtgcgctc 20460

tgcccactga cgggcaccgg agcctcacgc atgctcttct ccacctcagt gatgacgaga 20520

gcgggcgggt gaggggggcgg gaacgcagcg atctctgggt tctacgttag tgggagttta 20580

acgacggtcc ctgggattcc ccaaggcagg ggcgagtcct tttgtatgaa ttactcaact 20640

agtgatatct taattaacaa acggccgctg cattctagtt gtggtttgtc caaactcatc 20700

aatgtatctt atcatgtctg gatcccatca caaagctctg acctcaatcc tatagaaagg 20760

aggaatgagc caaaattcac ccaacttatt gtgggaagct tgtggaaggc tactcgaaat 20820

gtttgaccca agttaaacaa tttaaaggca atgctaccaa atactaattg agtgtatgtt 20880

aacttctgac ccactgggaa tgtgatgaaa gaaataaaag ctgaaatgaa tcattctctc 20940

tactattatt ctgatatttc acattcttaa aataaagtgg tgatcctaac tgaccttaag 21000

acagggaatc tttactcgga ttaaatgtca ggaattgtga aaaagtgagt ttaaatgtat 21060

ttggctaagg tgtatgtaaa cttccgactt caactgtagg ggatcctcta gggccgccag 21120

tgtgatggat atctgcagaa ttcggcttca ggtaccgtcg acgatgtagg tcacggtctc 21180

gaagccgcgg tgcgggtgcc agggcgtgcc cttgggctcc ccgggcgcgt actccacctc 21240

acccatctgg tccatcatga tgaacgggtc gaggtggcgg tagttgatcc cggcgaacgc 21300

gcggcgcacc gggaagccct cgccctcgaa accgctgggc gcggtggtca cggtgagcac 21360

gggacgtgcg acggcgtcgg cgggtgcgga tacgcggggc agcgtcagcg ggttctcgac 21420

ggtcacggcg ggcatgtcga caagccgaat tccagcacac tggcggccgt tactaggtag 21480

ctagctcgag ccttcgaaga tctcctaggg aagttcctat actttctaga gaataggaac 21540

ttcggaatag gaacttcacc ggtgggtgaa aagccgaatt ctgcagatat caagcttatc 21600

gataccgtcg acctcgaggg ggggcccggt tagatccccg ggtaccgagc tcgaattcat 21660

ctatgtcggg tgcggagaaa gaggtaatga aatggcatta tgggtattat gggtctgcat 21720

taacgatagc tagcctctaa ctcctagacc gtcagaactg ctgggccctt caagacgggc 21780

tgctcacacc cactcatgtt aagcctggtg aggcctgtac tctgttttca caggaagaaa 21840

tcctcaccca gtcttcccca aacacattcc caggttgtgt cattagtggg atagagatga 21900

ttattgtggg gagaagagaa acatctggat ggatttggtt aggttgatct atagaggaag 21960

taggtgctgc ctgaggtagc tgtaatagaa gctaaaggtc aaaggagagg gccctgtccc 22020

aatccagatg actccacttc tgctggaccc aggttcacaa gcttaatcta catttcacct 22080

aaatttggct aacaagccca aaatcacaca ggcaaaggga gaagtggagg cagaaccgag 22140

gttggaggcc accagggcca ccgggcagag atcatttaag cccaaccttc tcacttctcc 22200

ctgggctctg cctctcttaa aggaccttgt ggtgtgacct cttgtaggtc cctttcacac 22260

tcggggcctc agtttcccca ctgtaaagtg aatgggtccc agctttggta agcttatgct 22320

tacctgatgc tttcttcctg ggctgctctt gtagagaaaa gataaatctt cttcctccat 22380

ccacgagggc ttctttccct gggggtgaga gtaggctgag gagagccact tgcacacact 22440

cttaaagaaa gtattacctg caccagctca gtgagaggca cagatcagac tgttacttga 22500

atcaaattat gagcctcccc aaatatatct atgacattta aataggggat tacttgaaca 22560

tagactttgg gatccggtgt ggagtgcagg agactagcaa agtgaatcct gagagtagca 22620

ggtctgcacc tgttggatcg agaaaggcgg cctacaattc tggtcaaatg agctgtgctt 22680

attgacatat tctattagag agtactacca ggtcaccagt caccagaaag gctgccagct 22740

ctccaaccac ctccagggaa ctatcctgaa tggggcctta acaagcctaa gagagggttg 22800

gtttgggtcc caagccaata tttgctctgc tttatgtcag tcatatggaa cccaaaccaa 22860

ccctctccta tgtgcctcac cagtcggtgc agggatccca atttcaagtt tggtttttta 22920

tggtcaaagt ccagcataga ttaaatgaag gggtgtgatg atggtgttaa aagagaactc 22980

cagaccagtt taactcttgg acacacatcc catctcacca tggtgcttcc aaccttccag 23040

agatgatggg ctcctatttt ctgatgacaa agccctccac aggattgctg cctggccatc 23100

agggagtgcc tctgtaactg aggctgagat cccactttca gtcctccagc tgtggcccat 23160

ccctgctccg cccaccgggt atggcctgtc ctaggctctt aggtatggct gcattgtgaa 23220

atgatggcta cagagctggc atctcctgta gtctggttca tctagtgcac tacctcatag 23280

ttaaaagaaa tctgtttaag ccactgaggg tggctcctag tgccaactcc aagaacagga 23340

agcttccctt ttttgggagg aggggcagat ggtaacatgg atcgtccagg tcaatgggag 23400

cagggcaacc acagtaagta ctggacaaca acacaaaact ccatgtgtgg cttccatcga 23460
gtccctctcc aattggtttg gtcttctccg tcccatgcag cactttagca aggggcctgg 23520
ctgaaggcta tgaattgtgt ggagcctcct cattgcagtc tccaaccatc tgatgctggg 23580
aaaatgtcac caggatgcag ccatgccgtg tggccaatga accgagaaaa caccccttt 23640
ctagaatgct ctaaagaggc agaataatcc agaggtgagg aaggaaatac tccaccagag 23700
acccaggcag ttcctacaaa agccagactt tccttcacct agggagtgac aagaccagtg 23760
gaaaacactc tcaagcagta acccccaaat gctctgcaag ccagtggcgt ccagataccg 23820

cacaagcgag tgggctgtct aatcccatca tcatgatgta aatatctcta ggctgccctg 23880
ggctgtgcct gaccctgtct tcagctttcc acacctccac ctacagccca tgcacagaag 23940
gaccacccag gaatgctgca agtgtggcac ctccagggcc acccagggag aaggagggca 24000
gctatgctgg tggctccagg ggtggtacct tcacaccaca aagcccaaac tgaggcccca 24060
gatttggctg atgagggcat attggacagg ggtcacttat gctcttcccc attgccacct 24120
ggcctctggc tacctggact tggctacctg tggatcctct cacaggtgcc accatcttgg 24180
ctgagtctcc agatgcgagg tccctgaggc agtggcaggc ttctcgctaa tgctgatggg 24240

attaggaatg ggataggtgg ggagggccct ggactgggcc ctgatgagcc aagtgggttt 24300
ttagaggggc tactggtaca tttcagggac aggacatctg gtagagctaa gctggggcaa 24360
taaggagcca ctgctaatct gagagctaga aacaatcagc ttctgggtca ttattaatta 24420
gggtagtttg ggctgtgtgg aagtcacgta ctatatgggg tagccacagc tctctctaca 24480
gataatctct aagacttctg attgggaccg tgtgaatgca gtagcaatat ctcttcttac 24540
tgccaggccc tgccagtcct gcctccacgc cctggctggc ccccttatg atctgaccca 24600
tgccaggctg ccatagtatg ttacttctgc attagcactc cttgggacct gcctctccac 24660

tgtccctcag actttaaaga actatacaaa cccaaggggc tcttcccaag agaattgata 24720
tgacttgagg tgattccatt tctggaagta gtcactccat tttctgcctc actctttcag 24780
tgcttcacag agcaggttcg aacgaaggag ccatccaact aaccgtcatg ttcgggcaac 24840
cgaagaaggg agtggcagga tttcctttgg agacttctgg aattagacag cagtttaatg 24900
caagcatcta aattctcttc ctcccagagt ctcattaaaa ctacagtaag agtttgtgtt 24960
ttgttttgtt tttaaagaca aaatcccacc aggatagaga gaataggaga ggagataaca 25020
gcatcataat ttatgaaact aaaatgcaga tagaccagga ttaactgact acacagcacc 25080

aaggaagctg aatcacaaga cagcagagga gaaaactgga aaggatcgtg gtctatacgg 25140

cagaatcttc ccaagcctca ggaggaggag ctctagatgt tcccagatct gggaggtaaa 25200

gtggaatggg gggacatggt cagcgtaatg gggttgggct ggaagctggt taaggagcag 25260

gcagatctct gaatcccctc tctgactctg tgtccccagg catctgcctg tcccccaccc 25320

tggaagaggt ctggcttgac cctttgtctg gtgaatttcc tgctctgctt tcctggtcct 25380

gctggccaga tcagtggagg ccactcactt caccccacag ggatgttctg tgttgcccta 25440

cacctgggaa ctggaggtac tggaggcagg ctgtggtgag cttgaaagca aaacacagag 25500

ggcagtccaa tctctttggc catatttctt ctgcatatcc aataccatgt ccacaactct 25560

gctagtgtcc tgatggtggt gggctctaca cattcccggg aagctgaagg cagataatga 25620

ccaggacagg tcaacctctc ttcttctgaa agccttcatc tactaatggc ctgggactct 25680

tcccttaaat gcttagattg tgtcttccac taaggttttt tgctgttgct gttgtttgtt 25740

tgtttgtttg tttgtttgtt tgttttgaga cggaatctca ctctgtcgcc caggctggag 25800

tgtagtggca caatctcagc tcaccacaac cttcacctcc taggttgaag ggattctcct 25860

gcctcagcct cctgagtagc taggattaca ggcacatgcc accatgcctg gctaattttt 25920

gtatttttag tagagacagg atttcgccat gttggccagg ctggtcttga actcctgacc 25980

tcaggtgatc tgcctacctt ggtctcccaa agtgctggga ttactggtgt gagccaccac 26040

acccggccaa ggtttttgtt tgtttgtttg tttgtttgtt ttgtattgag gcagggtatc 26100

actctggtca cccaggctgg agtgcagtag tgcaatcacg gctcactgaa acctccacct 26160

ccctggcggg ctcaggtgat cctgccacct cagcttccca ggtagctggg actacaggct 26220

tgtaccacca ctcccagcta atttttgcgt ttttagtaga gacagggttt ccccatgttg 26280

cccaggttgg tctcaaactc tgggctcaag cgatctgcct gcctcagcct cccaaagtgc 26340

tgggattaca ggtgtaagcc accgtacccg gccccgccac taaggttttg aaaatgaagc 26400

aattacaagt ttaagtctat taataagtga tgaagctatg tagaaaagca gaataattat 26460

cttggatcag gaaggtcaca tgaggatcta cttgggggtt gtcaatattc tatttcttga 26520

cctgatcagt gttgacagca ggttttaatt ttttacttct ttttgtttgt ttgtttttga 26580

gacggagtct tgctctgtct cccaggctgg agtgcagtgg tatgatctcg gctcactgca 26640

acctccgcct cctgggttca agctgttctc ctgcctcagc ctccccagta gctgggatta 26700

caggcaggca ccaccacgac cagctaattt ttgtattttt agtagagact gggtttcacc 26760

atcttggcca ggctggtctc gaacttctga tctcgtgatc cgccctcctt ggcctcccaa 26820

agtgctggga ttacaggctt gagccagcgt gcccggccca ttttttactt ccttattaaa 26880

ctgtacatat aggccttgca cacttttctg catcaatgtt atattccaca ataaagggaa 26940

aaggtatata cacaacttga taccagtaat gtgaaacata tatttctaca tagaaaaaaa 27000

aatgactgaa atactgcact ccaatgtgtt cacacagtag ttgtttctgg attatttata 27060

tattaaatgt ttatatattg tattatgcca tgaggtttgt gttttctctc cacttttctg 27120

cattttccaa gtttactaca aagagcacat attactctta taatcagaaa gtcataaaat 27180

atatttaaaa agacaaaatt gaaactaata aggatcaaca caaaacagat gagccatctg 27240

tggaaatccg cacagaatac tacctaaaga gattggtgac gtgcatgatc tcactaggat 27300

gagcacaaag cttgccagag cctagggtct atttctaggg ttggctcttg gaagccagga 27360

tagttgttat ctctgggaag agggaggggc acacaagggg cttctaaaac attctgaatg 27420

ttctatttct gaacctggtt ggtgggtaca tgactgttgg ttttattatt atatgtttta 27480

tatactcttc cgtatgtatg gtgtggattc caaaaaaaga tttcctttag agaaaaccag 27540

aatcacataa gtagaaaata tggtgctatg ttgaaggaac aactcaagtt tatataaaat 27600

catcatcatt tataggctta aaaagttgct ttggaatttt ggtctaactg acttgtcttt 27660

tctgcagcaa accacgctcc ttctggacgt gctccaggca gaggggatta gggtgggttc 27720

aaggctgcaa gtacctagct cagcacactc tcttcagggg acttagagtt tgtctggtgt 27780

tggctctctg agctcttgtc aggaatgccg acccttccga ggttcaggat ttgaagcctg 27840

ccttcccacc ccagatttgg tccacacaga cactcaagta tgtatttcaa ctacaaatga 27900

cctgtacttt cctattactc ctctctttca tggtaacctt tctggtatcc ttccttccct 27960

acatttatgg gagggggaca tcattctctg ctctcctgtc actgaaggct ccaccttctg 28020

tcttcttctg acccatctgg ttttcctggg gccacctcct ctccttacca ccctaacgct 28080

tttgtaactt gaggagaaat gagagatcac ctagtcaggt catcattctc tgtagatgaa 28140

gaggcccaat ggtttgctca agaattgcca agcgagttaa agacagagag tatgagagtc 28200

agcaagacct acagaaagca tctatctgca ctgttttgca gggacttagc ctttgtgtgt 28260

ggactcctgg aatgccaccc actaagaaac attgtctgac accaactccc cacttggtag 28320

gtggggacac tgaaactcat ggcaggaaag ggccttgccc caagccaggg cagagtgtca 28380

ctcatcactc tcaattttca gtccagggca ccttgttgtg actatcccaa aggcagccac 28440

tttccctggt ctgaaagacc tgaagagaga agagaagaga aggatggaag gcagagtatg 28500

cggctttgat tcatttcctg gtgaaaacag atctatacga gaagcaaatt tcacgaaagg 28560

gaagagaaga aagtgtccca tacgttgctg gcctgtttca accttgcttt gattcttgct 28620

gaaaagggta ccgtgtattt ctgagttcaa catgcagacc agtgttagga aagccactgc 28680

acctccactt tagcctccag ggctgtgccc tgcaaatggc ctgcagcctt ggtgcctcgc 28740

tctccagact gcattttgga agatgggaca gaggcttatg gaagcccaca ttagaacggg 28800

ggagcagaat gggtgagatg agggatcctt gatagtgaac cagatgaagg aatggtagcc 28860

aaatgccagg cctcctttgt ggcttcaatc caaaggctct ggagcccttc cagggcagaa 28920
catcaggcat gtttaccccc actgtcctca acagtgacag aggtgcaatc ttgggcagct 28980
ggccattttg aaagcaacct ccttaatctc aactgggaag gctccctagc aggacccctg 29040
tgttgcacac ctggaggaag ctagactaac cagaagctca gcacggttcc atctgggatg 29100
cccaggtctg agacgaaaaa ggtaactctc ttttctgggt cctggcccag ttgtgtctct 29160
ctccacctca ttctctgaga tgcctgtctc ccctttttg tcccatcagg aggcaagagc 29220
tatcactggg ccagactcca ccagaagcca agccagcttg ttacccagct tctcagggag 29280

caaagaacag ccttgtttct atcttatccc cactgtcccc tgcccctgcc ccacctccca 29340
gccattcagc ttctggcttc cccagagctg cctgcttctt tgtggtcctc cattccttga 29400
aaagaccttc tagtcattag tgtatataaa tggccactta gcccagatta cagtgaggtc 29460
aacagctggg gctctgagaa ttgtcacaca ctggcacagg agaggaggct attcttccag 29520
agaatttgga gggcactccc atccacttac aacaaaaagc ccatccactg tgcttggcag 29580
taggtgatct gagaaccaat ggaaccaggt taatcctgtg gcactgttga gtgaggagag 29640
cagtggcggg cactggaaaa tatcagagac aaggcaggag acctgaaatc taggcttagc 29700

tcctcatata cttggcagct gtatgacctc agacaaccag tgttacctct ctaagcctca 29760
gtttcctcat gcaaaaggag ggggaataac aacagagccc actgcttggg ggtgttgtga 29820
ggacaggatg aaaaaacaaa cagaaatccc tcagtacagg attcagtgca gtggacagtc 29880
ttgcaaggtc tggttcagcc ctccaccccct accctcacca gtataaagaa ctctggccta 29940
caagtcagat gacctgagtt ttaatctcag ctttgccatt agccgtgtga acttgagaaa 30000
gtcccctttcc ttttttacatc tattgggatg atcatgcatt ttttgtcctt tattctgtta 30060
atatagtgtg ttacattgat tgcttttcat agactgaacc agccttgtat tccagggata 30120

aatctcactt ggtcatggtg tataatcctt tatacaaatg ttgctgggtt gagtttgcta 30180
gtattttgtt gaagattttt atgtcttgat tcataaggaa tattggtgta ccttcccctt 30240
ttatggccac agtttcccta caatgatgta gtcgaactag acaacctcca atatctttca 30300
gtattcatgt cctctgattc tgtgaaacta agaaaattaa gaaatagtga ttcataggca 30360
caaggcaggc aaaacttaga ctccttgtag aataattagg aagccaaata ttcagtgtgc 30420
ttatttctca aataacctta gtttctccag tctgccccaa ctccgaggcc tgaatatctc 30480
tagatgctta tgatggcaac taaagcctaa aagctaattc attttaaagt tcttccaaat 30540

gcatagggtt ttatttttcc agacctgggt tcagatgggg aatttgacaa acaatggaaa 30600

gggggaaaaa caacaatcta aacactgagt gacaaagtaa caaagaaata gtctagctat 30660

cagccagtca agccagcctt ggctttgcta tccaaagtag tcagtctaat tctaccacca 30720

gtttctgttc ctgtagctgt ctactgcctg ccagggactc tgccttccca cccacaacta 30780

ccaatggaag gatgtggtga ccataccagt ggctgctgac atctcctgcc atgggaagca 30840

taattgcctc cagcagcctc ccccttagat ccatcatttt tgttgcactt ggcctgggct 30900

gtactcccgg ccaatgactg aacatggtga gcatagtaat gcaggcccat ttctgtgagg 30960

agcaggactc ctccagtagg tgactttggc tcaaggactc tctattggcc tggttgaact 31020

tttcctgaac tgtgctactg tctgagactc ttcttaccca atcctctttc tcgccccaat 31080

tgtcacagac cacctgcatt gtggtctgag tctctcccca ccttctcttg ctcttccctg 31140

tttatctttc acaggcattt cccccagtac attccttgaa tgtctaaccc gatacgggtg 31200

cctgactttt ggcagaccta agcagacaaa aaggagtact tggttaccta gctcttcttt 31260

ctaccacaaa catcgaggga acccttttc cctcacccct ctgccacacc cccactgccc 31320

cagtgaacaa ccacagagag agctgtggta taatattagg ctggtgcaaa agtaattgcg 31380

gtttttgcca ttacttttaa tggtaaaaac cgcaattact tttgcaccta cctagtattt 31440

gtgtcccccc aaattcatat gttgaaacct aacccacaat atgatgtcat taggaggcaa 31500

gaccttgagg aggtgattag atgatggggt ggagctctcc tgaatgagat tagtgccctt 31560

ataagaagaa gcccaaggaa gctaccttga ctcttccatc acatgagaat gcagcaagaa 31620

ggcaccatct actaatcagg aagagagctc tcaccagaca ctgaatctgc cagtgtcttg 31680

atcttgaagt tcccagcctc cagaactatg cataatgcat ttccattgtc tctaagccac 31740

ccagcctatg gtattttgtc atagcagcct gaactgacta agacagtgag ccacatgaga 31800

agtgccccaa cccctccctt aagcacttgg ctcacagatc agtgggttca tttctgcctg 31860

agttttattg ttattctgta gatttcttgg gctagatata tttttctgtt attttccttc 31920

ttcacctcag tcatgaattg gttgttttaa aaaagacaat gtaagtcatg gggaaactcc 31980

tgacaactct actctcctag ggttcctgat aaaaggggat tcagttgagt cctctgatgg 32040

tctctacctg ccaaagtcca gcagccctta gcaaacatgc tgctcgtttc tgtagagaag 32100

gtgctggtgt cccaccatac ttctctctcc ctcatgaagg gcttgcgacc cagcaaatgg 32160

gtggcttata tgggtctgtt tcaaaggaag agccagctct gggaagaaaa acgatgagca 32220

taagcataac ctaccactgt gcctgggaaa gcagacaact tttttgatgt gtgaatatct 32280

aatgagaatg gaatccatca attaccttaa acttaggcac agtcttcaaa ttcaatatat 32340

gtgggatata cttttagtca gtttgtagac gttatttgta ataaataatc tggcttctct 32400

aaagaaatta ttttaagtgt ttggtttggt ttgatttaat ggtaaaatta tatttagtgg 32460

cagagaatta tagcaatggt gataaactat agagtgtcat a 32501

<210> 65

<211> 47

<212> DNA

<213> Artificial Sequence

<220><223> top strand of GATAA motif

<400> 65

caaacccttc ctggagcctg tgataaaagc aactgttagc ttgcact 47

<210> 66

<211> 47

<212> DNA

<213> Artificial Sequence

<220><223> bottom strand of GATAA motif

<400> 66

agtgcaagct aacagttgct tttatcacag gctccaggaa gggtttg 47

<210> 67

<211> 47

<212> DNA

<213> Artificial Sequence

<220><223> top strand of disrupted GATAA motif

<400> 67

caaacccttc ctggagcctg taataaaagc aactgttagc ttgcact 47

<210> 68

<211> 47

<212> DNA

<213> Artificial Sequence

<220><223> bottom strand of distrupted GATAA motif

<400> 68

agtgcaagct aacagttgct tttattacag gctccaggaa gggtttg 47

<210> 69

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> HBG2 forward primer

<400> 69

tccttctgtc attttgcctc tgtt 24

<210> 70

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> HBB primer

<400> 70

ttgagcaatg tggacagaga agg 23

<210> 71

<211> 40

<212> DNA

<213> Artificial Sequence

<220><223> HDAd-BEHBG-106

<400> 71

ttggccagcc ttgccttgac caatagcctt gacaaggcaa 40

<210> 72

<211> 27

<212> DNA

<213> Artificial Sequence

<220><223> HBG2 reverse primer

<400> 72

cacttcattg tagttaccgt ggaaaga 27

<210> 73

<211> 340

<212> PRT

<213> Artificial Sequence

<220><223> Sleeping Beauty transposase enzyme

<400> 73

Met Gly Lys Ser Lys Glu Ile Ser Gln Asp Leu Arg Lys Lys Ile Val

1 5 10 15

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

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

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

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

260 265 270

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

Ala Thr Lys Tyr
340

<210> 74
<211> 340
<212> PRT
<213> Artificial Sequence
<220><223> Hyperactive Sleeping Beauty is SB100X
<400> 74
Met Gly Lys Ser Lys Glu Ile Ser Gln Asp Leu Arg Lys Arg Ile Val
1 5 10 15
Asp Leu His Lys Ser Gly Ser Ser Leu Gly Ala Ile Ser Lys Arg Leu
20 25 30
Ala Val Pro Arg Ser Ser Val Gln Thr Ile Val Arg Lys Tyr Lys His
35 40 45

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

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

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

Phe Gln His Asp Asn Asp Pro Lys His Thr Ser Lys Val Val Ala Lys
245 250 255
Trp Leu Lys Asp Asn Lys Val Lys Val Leu Glu Trp Pro Ser Gln Ser
260 265 270
Pro Asp Leu Asn Pro Ile Glu Asn Leu Trp Ala Glu Leu Lys Lys Arg
275 280 285
Val Arg Ala Arg Arg Pro Thr Asn Leu Thr Gln Leu His Gln Leu Cys
290 295 300

Gln Glu Glu Trp Ala Lys Ile His Pro Asn Tyr Cys Gly Lys Leu Val
305 310 315 320
Glu Gly Tyr Pro Lys Arg Leu Thr Gln Val Lys Gln Phe Lys Gly Asn
325 330 335
Ala Thr Lys Tyr
340

<210> 75

<211> 423

<212> PRT

<213> Artificial Sequence

<220><223> Flp (flippase)

<400> 75

Met Pro Gln Phe Gly Ile Leu Cys Lys Thr Pro Pro Lys Val Leu Val
1 5 10 15

Arg Gln Phe Val Glu Arg Phe Glu Arg Pro Ser Gly Glu Lys Ile Ala
20 25 30

Leu Cys Ala Ala Glu Leu Thr Tyr Leu Cys Trp Met Ile Thr His Asn
35 40 45

Gly Thr Ala Ile Lys Arg Ala Thr Phe Met Ser Tyr Asn Thr Ile Ile
50 55 60

Ser Asn Ser Leu Ser Phe Asp Ile Val Asn Lys Ser Leu Gln Phe Lys
65 70 75 80

Tyr Lys Thr Gln Lys Ala Thr Ile Leu Glu Ala Ser Leu Lys Lys Leu
85 90 95

Ile Pro Ala Trp Glu Phe Thr Ile Ile Pro Tyr Tyr Gly Gln Lys His
100 105 110

Gln Ser Asp Ile Thr Asp Ile Val Ser Ser Leu Gln Leu Gln Phe Glu
115 120 125

Ser Ser Glu Glu Ala Asp Lys Gly Asn Ser His Ser Lys Lys Met Leu
130 135 140

Lys Ala Leu Leu Ser Glu Gly Glu Ser Ile Trp Glu Ile Thr Glu Lys
145 150 155 160

Ile Leu Asn Ser Phe Glu Tyr Thr Ser Arg Phe Thr Lys Thr Lys Thr
165 170 175

Leu Tyr Gln Phe Leu Phe Leu Ala Thr Phe Ile Asn Cys Gly Arg Phe
180 185 190

Ser Asp Ile Lys Asn Val Asp Pro Lys Ser Phe Lys Leu Val Gln Asn
195 200 205

Lys Tyr Leu Gly Val Ile Ile Gln Cys Leu Val Thr Glu Thr Lys Thr

210 215 220

Ser Val Ser Arg His Ile Tyr Phe Phe Ser Ala Arg Gly Arg Ile Asp

225 230 235 240

Pro Leu Val Tyr Leu Asp Glu Phe Leu Arg Asn Ser Glu Pro Val Leu

245 250 255

Lys Arg Val Asn Arg Thr Gly Asn Ser Ser Ser Asn Lys Gln Glu Tyr

260 265 270

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

275 280 285

Lys Asn Ala Pro Tyr Ser Ile Phe Ala Ile Lys Asn Gly Pro Lys Ser

290 295 300

His Ile Gly Arg His Leu Met Thr Ser Phe Leu Ser Met Lys Gly Leu

305 310 315 320

Thr Glu Leu Thr Asn Val Val Gly Asn Trp Ser Asp Lys Arg Ala Ser

325 330 335

Ala Val Ala Arg Thr Thr Tyr Thr His Gln Ile Thr Ala Ile Pro Asp

340 345 350

His Tyr Phe Ala Leu Val Ser Arg Tyr Tyr Ala Tyr Asp Pro Ile Ser

355 360 365

Lys Glu Met Ile Ala Leu Lys Asp Glu Thr Asn Pro Ile Glu Glu Trp

370 375 380

Gln His Ile Glu Gln Leu Lys Gly Ser Ala Glu Gly Ser Ile Arg Tyr

385 390 395 400

Pro Ala Trp Asn Gly Ile Ile Ser Gln Glu Val Leu Asp Tyr Leu Ser

405 410 415

Ser Tyr Ile Asn Arg Arg Ile

420

<210> 76

<211> 34

<212> DNA

<213> Artificial Sequence

<220><223> FRT recognition site

<400> 76

gaagttccta tactttctag agaataggaa cttc 34

<210> 77

<211> 374

<212> PRT

<213> Artificial Sequence

<220><223> variant of the Flp protein

<400> 77

Met Asn Glu Gln Ala Lys Gln Leu Val Ala Val Thr Arg Gln Pro Ala

1 5 10 15

Leu Asn Ala Gly Val Gly Leu Val Leu Ala Gln Leu Ala Glu Val Glu

20 25 30

Ala Arg Gln Ile Pro Gly Ser Leu Ala Glu Ala Arg Ala His Cys Leu

35 40 45

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

50 55 60

Thr Leu Ala Ala Asp Leu Ala Ala Leu Ala Glu Cys Cys Pro Pro Gln

65 70 75 80

Met Arg Leu Val Leu Leu Gly Glu Arg Gly Asp Val Thr Leu Phe Arg

85 90 95

Trp Leu Ile Ser Val Gly Val Asp Asp Tyr Tyr Pro Ala Pro Leu Asp

100 105 110

Pro Asp Ala Leu Arg Thr Gly Leu Leu Arg Leu Leu Gly Val Pro Leu

115 120 125

Val Thr Ser Leu Arg Lys Gly Arg Val Ile Cys Val Val Gly Ala Ala

130 135 140

Gly Gly Val Gly Thr Ser Thr Val Ala Ala Asn Leu Ala Met Ala Leu

145 150 155 160

Ala Asp Gln His His Arg Gln Val Ala Leu Leu Asp Leu Asn Leu His

165 170 175

His Ser Arg His Pro Ile Leu Leu Gly Ser Asp Tyr Ala Pro Pro Gly

180 185 190

Glu Gln Trp Trp Gln Ala Thr Asp Arg Leu Asp Gly Thr Leu Leu Ala

195 200 205

His Thr Ala His Gln Leu Gly Pro Arg Leu Phe Leu Phe Tyr Asp Glu

210 215 220

Gly Gln Glu Leu Val Leu Gly Ala Glu Gln Leu Val Ala Ala Val Asn

225 230 235 240

Val Met Ala Glu His Tyr Ser Thr Leu Ile Ile Asp Val Pro Asp Leu

245 250 255

Arg Thr His Gly Leu Arg Ala Leu Leu Gln Glu Ala Asp Val Val Leu

260 265 270

Trp Leu His Asp Phe Ser Leu Gly Ala Leu Arg Leu Leu Gly Gln Cys

275 280 285

Pro Gln Gly Gly Gln Ala Gln Arg Arg Leu Leu Val Gly Asn His Cys

290 295 300

Arg Gly Lys Glu Gly Arg Val Pro Ala Gln Glu Leu Glu Arg Val Cys

305 310 315 320

Gly Gln Pro His Ala Ala Val Leu Pro Tyr Asp His Gly Val Phe Val

325 330 335

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

340 345 350

Ala Arg Ala Leu Thr Leu Leu Ala Gly Glu Leu Val Gly Ala Gln Val

355 360 365

Thr Gly Arg Gly Arg Arg

370

<210> 78

<211> 432

<212> PRT

<213> Artificial Sequence

<220><223> variant of the Flp protein

<400> 78

Met Ala Pro Lys Lys Lys Arg Lys Val Met Ser Gln Phe Asp Ile Leu

1 5 10 15

Cys Lys Thr Pro Pro Lys Val Leu Val Arg Gln Phe Val Glu Arg Phe

20 25 30

Glu Arg Pro Ser Gly Glu Lys Ile Ala Ser Cys Ala Ala Glu Leu Thr

35 40 45

Tyr Leu Cys Trp Met Ile Thr His Asn Gly Thr Ala Ile Lys Arg Ala

50 55 60

Thr Phe Met Ser Tyr Asn Thr Ile Ile Ser Asn Ser Leu Ser Phe Asp

65 70 75 80

Ile Val Asn Lys Ser Leu Gln Phe Lys Tyr Lys Thr Gln Lys Ala Thr

85 90 95

Ile Leu Glu Ala Ser Leu Lys Lys Leu Ile Pro Ala Trp Glu Phe Thr

100 105 110

Ile Ile Pro Tyr Asn Gly Gln Lys His Gln Ser Asp Ile Thr Asp Ile

115 120 125

Val Ser Ser Leu Gln Leu Gln Phe Glu Ser Ser Glu Glu Ala Asp Lys

130 135 140

Gly Asn Ser His Ser Lys Lys Met Leu Lys Ala Leu Leu Ser Glu Gly

145 150 155 160

Glu Ser Ile Trp Glu Ile Thr Glu Lys Ile Leu Asn Ser Phe Glu Tyr

165 170 175

Thr Ser Arg Phe Thr Lys Thr Lys Thr Leu Tyr Gln Phe Leu Phe Leu

180 185 190

Ala Thr Phe Ile Asn Cys Gly Arg Phe Ser Asp Ile Lys Asn Val Asp

195 200 205

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

210 215 220

Gln Cys Leu Val Thr Glu Thr Lys Thr Ser Val Ser Arg His Ile Tyr

225 230 235 240

Phe Phe Ser Ala Arg Gly Arg Ile Asp Pro Leu Val Tyr Leu Asp Glu

245 250 255

Phe Leu Arg Asn Ser Glu Pro Val Leu Lys Arg Val Asn Arg Thr Gly

260 265 270

Asn Ser Ser Ser Asn Lys Gln Glu Tyr Gln Leu Leu Lys Asp Asn Leu

275 280 285

Val Arg Ser Tyr Asn Lys Ala Leu Lys Lys Asn Ala Pro Tyr Pro Ile

290 295 300

Phe Ala Ile Lys Asn Gly Pro Lys Ser His Ile Gly Arg His Leu Met

305 310 315 320

Thr Ser Phe Leu Ser Met Lys Gly Leu Thr Glu Leu Thr Asn Val Val

325 330 335

Gly Asn Trp Ser Asp Lys Arg Ala Ser Ala Val Ala Arg Thr Thr Tyr

340 345 350

Thr His Gln Ile Thr Ala Ile Pro Asp His Tyr Phe Ala Leu Val Ser

355 360 365

Arg Tyr Tyr Ala Tyr Asp Pro Ile Ser Lys Glu Met Ile Ala Leu Lys

370 375 380

Asp Glu Thr Asn Pro Ile Glu Glu Trp Gln His Ile Glu Gln Leu Lys

385 390 395 400

Gly Ser Ala Glu Gly Ser Ile Arg Tyr Pro Ala Trp Asn Gly Ile Ile

405 410 415

Ser Gln Glu Val Leu Asp Tyr Leu Ser Ser Tyr Ile Asn Arg Arg Ile

420 425 430

<210> 79

<211> 343

<212> PRT

<213> Artificial Sequence

<220><223> Cre

<400> 79

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

1 5 10 15

Asp Ala Thr Ser Asp Glu Val Arg Lys Asn Leu Met Asp Met Phe Arg

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

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

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

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

Tyr Gly Ala Lys Asp Asp Ser Gly Gln Arg Tyr Leu Ala Trp Ser Gly
275 280 285

His Ser Ala Arg Val Gly Ala Ala Arg Asp Met Ala Arg Ala Gly Val
290 295 300

Ser Ile Pro Glu Ile Met Gln Ala Gly Gly Trp Thr Asn Val Asn Ile
305 310 315 320

Val Met Asn Tyr Ile Arg Asn Leu Asp Ser Glu Thr Gly Ala Met Val
325 330 335

Arg Leu Leu Glu Asp Gly Asp
340

<210> 80

<211> 34

<212> DNA

<213> Artificial Sequence

<220><223> loxP site

<400> 80

ataacttcgt atagcataca ttatacgaag ttat 34

<210> 81

<211> 34

<212> DNA

<213> Artificial Sequence

<220><223> lox recognition site (lox2272)

<400> 81

ataacttcgt ataaagtatc ctatacgaag ttat 34

<210> 82

<211> 34

<212> DNA

<213> Artificial Sequence

<220><223

> lox recognition site (lox511)

<400> 82

ataacttcgt ataatgtatg ctatacgaag ttat 34

<210> 83

<211> 34

<212> DNA

<213> Artificial Sequence

<220><223> lox recognition site (lox66)

<400> 83

ataacttcgt atagcataca ttatacgaac ggta 34

<210> 84

<211> 34

<212> DNA

<213> Artificial Sequence

<220><223> lox recognition site (lox71

<400> 84

taccgttcgt atagcataca ttatacgaag ttat 34

<210> 85

<211> 34

<212> DNA

<213> Artificial Sequence

<220><223> lox recognition site (loxM2)

<400> 85

ataacttcgt ataagaaacc atatacgaag ttat 34

<210> 86

<211> 34

<212> DNA

<213> Artificial Sequence

<220><223> lox recognition site (lox5171)

<400> 86

ataacttcgt ataatgtgta ctatacgaag ttat 34

<210> 87

<211> 380

<212> PRT

<213> Vibrio plasmid p0908

<400> 87

Met Ile Glu Asn Gln Leu Ser Leu Leu Gly Asp Phe Ser Gly Val Arg

1 5 10 15

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

65 70 75 80
His Arg His Ser Leu Pro Ala Thr Pro Asp Thr Val Glu Ala Phe Phe
85 90 95
Ile Glu Arg Ala Glu Glu Leu His Arg Asn Thr Leu Ser Val Tyr Arg
100 105 110
Trp Ala Ile Ser Arg Val His Arg Val Ala Gly Cys Pro Asp Pro Cys
115 120 125
Leu Asp Ile Tyr Val Glu Asp Arg Leu Lys Ala Ile Ala Arg Lys Lys

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

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

260 265 270

Lys Asp Lys Gln Thr Gly Glu Val Leu His Lys Pro Ile Thr Thr Lys

275 280 285

Thr Val Glu Gly Val Phe Tyr Ser Ala Trp Glu Thr Leu Asp Leu Gly

290 295 300

Arg Gln Gly Val Lys Pro Phe Thr Ala His Ser Ala Arg Val Gly Ala

305 310 315 320

Ala Gln Asp Leu Leu Lys Lys Gly Tyr Asn Thr Leu Gln Ile Gln Gln

325 330 335

Ser Gly Arg Trp Ser Ser Gly Ala Met Val Ala Arg Tyr Gly Arg Ala

340 345 350

Ile Leu Ala Arg Asp Gly Ala Met Ala His Ser Arg Val Lys Thr Arg

355 360 365

Ser Ala Pro Met Gln Trp Gly Lys Asp Glu Lys Asp

370 375 380

<210> 88

<211> 34

<212> DNA

<213> Artificial Sequence

<220><223> VloxP recognition site

<400> 88

tcaatttccg agaatgacag ttctcagaaa ttga 34

<210> 89

<211> 342

<212> PRT

<213> Enterobacteria phage D6

<400> 89

Met Ser Glu Leu Ile Ile Ser Gly Ser Ser Gly Gly Phe Leu Arg Asn

1 5 10 15

Ile Gly Lys Glu Tyr Gln Glu Ala Ala Glu Asn Phe Met Arg Phe Met

20 25 30

Asn Asp Gln Gly Ala Tyr Ala Pro Asn Thr Leu Arg Asp Leu Arg Leu

35 40 45

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

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

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

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

Gly His Ser Ala Arg Val Gly Ala Ala Ile Asp Met Ala Glu Lys Gln
290 295 300

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

<210> 90
<211> 32
<212> DNA
<213> Artificial Sequence
<220><223> rox recognition site
<400> 90
taactttaaa taattggcat tatttaaagt ta 32

<210>
91
<211> 10
<212> PRT
<213> Artificial Sequence
<220><223> CD33-specific CDRL1
<400> 91
Thr Ala Ser Ser Ser Val Asn Tyr Ile His
1 5 10

<210> 92
<211> 6
<212> PRT
<213> Artificial Sequence
<220><223> CD33-specific CDRL2
<400> 92
Thr Ser Lys Val Ala Ser
1 5

<210> 93
<211> 9

<212> PRT

<213> Artificial Sequence

<220><223> CD33-specific CDRL3

<400> 93

Gln Gln Trp Arg Ser Tyr Pro Leu Thr

1 5

<210> 94

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> CD33-specific CDRH1

<400> 94

Asp Tyr Val Val His

1 5

<210> 95

<211> 17

<212> PRT

<213> Artificial Sequence

<220><223> CD33-specific CDRH2

<400> 95

Tyr Ile Asn Pro Tyr Asn Asp Gly Thr Lys Tyr Asn Glu Lys Phe Lys

1 5 10 15

Gly

<210> 96

<211> 11

<212> PRT

<213> Artificial Sequence

<220><223> CD33-specific CDRH3

<400> 96

Asp Tyr Arg Tyr Glu Val Tyr Gly Met Asp Tyr

1 5 10

<210> 97

<211> 14

<212> PRT

<213> Artificial Sequence

<220><223> CD33-specific CDRL1

<400> 97

Arg Ala Ser Glu Val Asp Asn Tyr Gly Ile Ser Phe Met Asn

1 5 10

<210> 98

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> CD33-specific CDRL2

<400> 98

Ala Ala Ser Asn Gln Gly Ser

1 5

<210> 99

<211> 8

<212> PRT

<213> Artificial Sequence

<220><223> CD33-specific CDRL3

<400> 99

Gln Gln Ser Lys Glu Val Pro Trp

1 5

<210> 100

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> CD33-specific CDRH1

<400> 100

Asp Tyr Asn Met His

1 5

<210> 101

<211> 17

<212> PRT

<213> Artificial Sequence

<220><223> CD33-specific CDRH2

<400> 101

Tyr Ile Tyr Pro Tyr Asn Gly Gly Thr Gly Tyr Asn Gln Lys Phe Lys

1 5 10 15

Ser

<210> 102

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> CD33-specific CDRH3

<400> 102

Gly Arg Pro Ala Met Asp Tyr

1 5

<210> 103

<211> 115

<212> PRT

<213> Artificial Sequence

<220><223> CD33-specific variable light chain

<400> 103

Asp Ile Val Leu Thr Gln Ser Pro Thr Ile Met Ser Ala Ser Pro Gly

1 5 10 15

Glu Arg Val Thr Met Thr Cys Thr Ala Ser Ser Ser Val Asn Tyr Ile

20 25 30

His Trp Tyr Gln Gln Lys Ser Gly Asp Ser Pro Lys Arg Trp Ile Phe

35 40 45

Asp Thr Ser Lys Val Ala Ser Gly Val Pro Ala Arg Phe Ser Gly Ser

50 55 60

Gly Ser Gly Thr Ser Tyr Ser Leu Thr Ile Ser Thr Met Glu Ala Glu

65 70 75 80

Asp Ala Ala Thr Tyr Tyr Cys Gln Gln Trp Arg Ser Tyr Pro Leu Thr

85 90 95

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

100 105 110

Thr Val Ser

115

<210> 104

<211> 120

<212> PRT

<213> Artificial Sequence

<220><223> CD33-specific variable heavy chain

<400> 104

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

1 5 10 15

Ser Val Lys Met Ser Cys Lys Ala Ser Gly Tyr Lys Phe Thr Asp Tyr

20 25 30

Val Val His Trp Leu Lys Gln Lys Pro Gly Gln Gly Leu Glu Trp Ile

35 40 45

Gly Tyr Ile Asn Pro Tyr Asn Asp Gly Thr Lys Tyr Asn Glu Lys Phe

50 55 60

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

65 70 75 80

Met Glu Val Ser Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr Tyr Cys

85 90 95

Ala Arg Asp Tyr Arg Tyr Glu Val Tyr Gly Met Asp Tyr Trp Gly Gln

100 105 110

Gly Thr Ser Val Thr Val Ser Ser

115 120

<210> 105

<211> 25

<212> DNA

<213> Artificial Sequence

<220><223> HBG1-seq primer

<400> 105

tttccttaga aaccactgct aactg 25

<210> 106

<211> 115

<212> PRT

<213> Artificial Sequence

<220><223> human CD33-specific variable heavy chain

<400> 106

Asp Ile Val Leu Thr Gln Ser Pro Ala Ile Met Ser Ala Ser Pro Gly
1 5 10 15

Glu Lys Val Thr Met Thr Cys Ser Ala Asn Ser Ser Val Ser Tyr Ile
20 25 30

His Trp Tyr Gln Gln Lys Ser Gly Thr Ser Pro Lys Arg Trp Ile Phe
35 40 45

Asp Thr Ser Lys Leu Ala Ser Gly Val Pro Ala Arg Phe Ser Gly Ser

50 55 60

Gly Ser Gly Thr Ser Tyr Ser Leu Thr Ile Ser Thr Met Glu Ala Glu
65 70 75 80

Asp Ala Ala Thr Tyr Tyr Cys Gln Gln Trp Thr Ser His Pro Leu Thr
85 90 95

Phe Gly Thr Gly Thr Lys Leu Gln Leu Lys Arg Ala Asp Ala Ala Pro
100 105 110

Thr Val Ser
115

<210> 107

<211> 10

<212> PRT

<213> Artificial Sequence

<220><223> CD33FL-specific CDRL1

<400> 107

Gly Tyr Ala Phe Ser Asn Tyr Trp Met Asn
1 5 10

<210> 108

<211> 17

<212> PRT

<213> Artificial Sequence

<220><223> CD33FL-specific CDRL2

<400> 108

Gln Ile Tyr Pro Gly Asn Phe Asn Thr Asp Tyr Asn Gly Gln Phe Lys

1 5 10 15

Gly

<210> 109

<211> 12

<212> PRT

<213> Artificial Sequence

<220><223

> CD33FL-specific CDRL3

<400> 109

Phe Phe Asp Phe Gly Ala Tyr Phe Thr Leu Asp Tyr

1 5 10

<210> 110

<211> 11

<212> PRT

<213> Artificial Sequence

<220><223> CD33FL-specific CDRH1

<400> 110

Lys Ala Ser Gln Asp Ile Lys Ser Tyr Leu Ser

1 5 10

<210> 111

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> CD33FL-specific CDRH2

<400> 111

Tyr Ala Thr Thr Leu Ala Asp

1 5

<210> 112

<211> 9

<212> PRT

<213> Artificial Sequence

<220><223> CD33FL-specific CDRH3

<400> 112

Leu His His Gly Glu Ser Pro Trp Thr

1 5

<210> 113

<211> 234

<212> PRT

<213> Artificial Sequence

<220><223> human CD33FL-specific variable light chain

<400> 113

Met Glu Thr Asp Thr Leu Leu Leu Trp Val Leu Leu Leu Trp Val Pro

1 5 10 15

Gly Ser Thr Gly Asp Ile Lys Met Thr Gln Ser Pro Ser Ser Ile Tyr

20 25 30

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

35 40 45

Ile Lys Ser Tyr Leu Ser Trp Tyr Gln Gln Lys Pro Trp Lys Ser Pro

50 55 60

Lys Thr Leu Ile Tyr Tyr Ala Thr Thr Leu Ala Asp Gly Val Pro Ser

65 70 75 80

Arg Phe Ser Gly Ser Gly Ser Gly Gln Asp Tyr Ser Leu Thr Ile Ser

85 90 95

Ser Leu Glu Ser Asp Asp Thr Ala Thr Tyr Tyr Cys Leu His His Gly

100 105 110

Glu Ser Pro Trp Thr Phe Gly Glu Gly Thr Lys Leu Glu Ile Lys Arg

115 120 125

Ala Asp Ala Ala Pro Thr Val Ser Ile Phe Pro Pro Ser Ser Glu Gln

130 135 140

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

145 150 155 160

Pro Lys Asp Ile Asn Val Lys Trp Lys Ile Asp Gly Ser Glu Arg Gln

165 170 175

Asn Gly Val Leu Asn Ser Trp Thr Asp Gln Asp Ser Lys Asp Ser Thr

180 185 190

Tyr Ser Met Ser Ser Thr Leu Thr Leu Thr Lys Asp Glu Tyr Glu Arg

195 200 205

His Asn Ser Tyr Thr Cys Glu Ala Thr His Lys Thr Ser Thr Ser Pro

210 215 220

Ile Val Lys Ser Phe Asn Arg Asn Glu Cys

225 230

<210> 114

<211> 471

<212> PRT

<213> Artificial Sequence

<220><223> human CD33FL-specific variable heavy chain

<400> 114

Met Glu Thr Asp Thr Leu Leu Leu Trp Val Leu Leu Leu Trp Val Pro

1 5 10 15

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

20 25 30

Lys Pro Gly Ala Ser Val Lys Ile Ser Cys Arg Ala Ser Gly Tyr Ala

35 40 45

Phe Ser Asn Tyr Trp Met Asn Trp Val Lys Gln Arg Pro Gly Lys Gly

50 55 60

Leu Glu Trp Ile Gly Gln Ile Tyr Pro Gly Asn Phe Asn Thr Asp Tyr

65 70 75 80

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

85 90 95

Asn Thr Ala Tyr Met Gln Leu Ser Ser Leu Thr Ser Glu Asp Ser Ala

100 105 110

Val Tyr Phe Cys Ala Arg Phe Phe Asp Phe Gly Ala Tyr Phe Thr Leu

115 120 125

Asp Tyr Trp Gly Gln Gly Thr Ser Val Thr Val Ser Ser Ala Lys Thr

130 135 140

Thr Ala Pro Ser Val Tyr Pro Leu Ala Pro Val Cys Gly Asp Thr Thr

145 150 155 160

Gly Ser Ser Val Thr Leu Gly Cys Leu Val Lys Gly Tyr Phe Pro Glu

165 170 175

Pro Val Thr Leu Thr Trp Asn Ser Gly Ser Leu Ser Ser Gly Val His

180 185 190

Thr Phe Pro Ala Val Leu Gln Ser Asp Leu Tyr Thr Leu Ser Ser Ser

195 200 205

Val Thr Val Thr Ser Ser Thr Trp Pro Ser Gln Ser Ile Thr Cys Asn

210 215 220

Val Ala His Pro Ala Ser Ser Thr Lys Val Asp Lys Lys Ile Glu Pro

225 230 235 240

Arg Gly Pro Thr Ile Lys Pro Cys Pro Pro Cys Lys Cys Pro Ala Pro

245 250 255

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

260 265 270

Asp Val Leu Met Ile Ser Leu Ser Pro Ile Val Thr Cys Val Val Val

275 280 285

Asp Val Ser Glu Asp Asp Pro Asp Val Gln Ile Ser Trp Phe Val Asn

290 295 300

Asn Val Glu Val His Thr Ala Gln Thr Gln Thr His Arg Glu Asp Tyr

305 310 315 320

Asn Ser Thr Leu Arg Val Val Ser Ala Leu Pro Ile Gln His Gln Asp

325 330 335

Trp Met Ser Gly Lys Glu Phe Lys Cys Lys Val Asn Asn Lys Asp Leu

340 345 350

Pro Ala Pro Ile Glu Arg Thr Ile Ser Lys Pro Lys Gly Ser Val Arg
355 360 365

Ala Pro Gln Val Tyr Val Leu Pro Pro Pro Glu Glu Glu Met Thr Lys
370 375 380

Lys Gln Val Thr Leu Thr Cys Met Val Thr Asp Phe Met Pro Glu Asp
385 390 395 400

Ile Tyr Val Glu Trp Thr Asn Asn Gly Lys Thr Glu Leu Asn Tyr Lys
405 410 415

Asn Thr Glu Pro Val Leu Asp Ser Asp Gly Ser Tyr Phe Met Tyr Ser
420 425 430

Lys Leu Arg Val Glu Lys Lys Asn Trp Val Glu Arg Asn Ser Tyr Ser
435 440 445

Cys Ser Val Val His Glu Gly Leu His Asn His His Thr Thr Lys Ser
450 455 460

Phe Ser Arg Thr Pro Gly Lys
465 470

<210> 115

<211> 10

<212> PRT

<213> Artificial Sequence

<220><223> CD33 deltaE2-specific CDRL1

<400> 115

Gly Tyr Thr Phe Thr Thr Tyr Trp Met His
1 5 10

<210> 116

<211> 13

<212> PRT

<213> Artificial Sequence

<220><223> CD33 deltaE2-specific CDRL2

<400> 116

Ala Ile Tyr Pro Gly Asn Ser Asp Thr Ser Tyr Asn Gln
1 5 10

<210> 117

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> CD33 deltaE2-specific CDRL3

<400> 117

Tyr Asp Gly Tyr His Phe Ile

1 5

<210> 118

<211> 16

<212> PRT

<213> Artificial Sequence

<220><223> CD33 deltaE2-specific CDRH1

<400> 118

Arg Ser Ser Gln Ser Leu Leu His Ser Asn Gly Ile Thr Tyr Leu Tyr

1 5 10 15

<210> 119

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> CD33 deltaE2-specific CDRH2

<400> 119

Gln Met Ser Asn Leu Ala Ser

1 5

<210> 120

<211> 9

<212> PRT

<213> Artificial Sequence

<220><223> CD33 deltaE2-specific CDRH3

<400> 120

Ala Gln Asn Leu Glu Leu Pro Pro Thr

1 5

<210> 121

<211> 239

<212> PRT

<213> Artificial Sequence

<220><223> human CD33 deltaE2-specific variable light chain

<400> 121

Met Glu Thr Asp Thr Leu Leu Leu Trp Val Leu Leu Leu Trp Val Pro
1 5 10 15

Gly Ser Thr Gly Asp Ile Val Met Thr Gln Ala Ala Phe Ser Asn Pro
20 25 30

Val Thr Leu Gly Thr Ser Ala Ser Ile Ser Cys Arg Ser Ser Gln Ser
35 40 45

Leu Leu His Ser Asn Gly Ile Thr Tyr Leu Tyr Trp Tyr Leu Gln Lys
50 55 60

Pro Gly Gln Ser Pro Gln Leu Leu Ile Tyr Gln Met Ser Asn Leu Ala
65 70 75 80

Ser Gly Val Pro Asp Arg Phe Ser Ser Ser Gly Ser Gly Thr Asp Phe
85 90 95

Thr Leu Arg Ile Ser Arg Val Glu Ala Glu Asp Val Gly Val Tyr Tyr
100 105 110

Cys Ala Gln Asn Leu Glu Leu Pro Pro Thr Phe Gly Gly Gly Thr Lys
115 120 125

Leu Glu Ile Lys Arg Ala Asp Ala Ala Pro Thr Val Ser Ile Phe Pro
130 135 140

Pro Ser Ser Glu Gln Leu Thr Ser Gly Gly Ala Ser Val Val Cys Phe
145 150 155 160

Leu Asn Asn Phe Tyr Pro Lys Asp Ile Asn Val Lys Trp Lys Ile Asp
165 170 175

Gly Ser Glu Arg Gln Asn Gly Val Leu Asn Ser Trp Thr Asp Gln Asp
180 185 190

Ser Lys Asp Ser Thr Tyr Ser Met Ser Ser Thr Leu Thr Leu Thr Lys
195 200 205

Asp Glu Tyr Glu Arg His Asn Ser Tyr Thr Cys Glu Ala Thr His Lys

210 215 220

Thr Ser Thr Ser Pro Ile Val Lys Ser Phe Asn Arg Asn Glu Cys
225 230 235

<210> 122

<211> 473

<212> PRT

<213> Artificial Sequence

<220><223> human CD33 deltaE2-specific variable heavy chain

<400> 122

Met Glu Thr Asp Thr Leu Leu Leu Trp Val Leu Leu Leu Trp Val Pro
1 5 10 15

Gly Ser Thr Gly Glu Val Gln Leu Gln Gln Ser Gly Thr Val Leu Ala
20 25 30

Arg Pro Gly Ala Ser Val Lys Met Ser Cys Lys Ala Ser Gly Tyr Thr
35 40 45

Phe Thr Thr Tyr Trp Met His Trp Ile Lys Gln Ser Pro Gly Gln Gly
50 55 60

Leu Glu Trp Ile Gly Ala Ile Tyr Pro Gly Asn Ser Asp Thr Ser Tyr
65 70 75 80

Asn Gln Lys Phe Lys Gly Lys Ala Lys Leu Thr Ala Val Thr Ser Ala
85 90 95

Ser Thr Ala Tyr Met Glu Leu Ser Ser Leu Thr Asn Glu Asp Ser Ala
100 105 110

Val Tyr Tyr Cys Glu Ile Tyr Asp Gly Tyr His Phe Ile Tyr Trp Gly
115 120 125

Gln Gly Thr Thr Leu Thr Val Ser Ser Ala Lys Thr Thr Pro Pro Ser
130 135 140

Val Tyr Pro Leu Ala Pro Gly Cys Gly Asp Thr Thr Gly Ser Ser Val
145 150 155 160

Thr Leu Gly Cys Leu Val Lys Gly Tyr Phe Pro Glu Ser Val Thr Val
165 170 175

Thr Trp Asn Ser Gly Ser Leu Ser Ser Ser Val His Thr Phe Pro Ala
180 185 190
Leu Leu Gln Ser Gly Leu Tyr Thr Met Ser Ser Ser Val Thr Val Pro
195 200 205
Ser Ser Thr Trp Pro Ser Gln Thr Val Thr Cys Ser Val Ala His Pro
210 215 220

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

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

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

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

420 425 430

Tyr Ser Lys Leu Asn Met Lys Thr Ser Lys Trp Glu Lys Thr Asp Ser

435 440 445

Phe Ser Cys Asn Val Arg His Glu Gly Leu Lys Asn Tyr Tyr Leu Lys

450 455 460

Lys Thr Ile Ser Arg Ser Pro Gly Lys

465 470

<210> 123

<211> 11

<212> PRT

<213> Artificial Sequence

<220><223> HER2-specific CDRL1

<400> 123

Lys Ala Ser Gln Asp Val Ser Ile Gly Val Ala

1 5 10

<210> 124

<211> 6

<212> PRT

<213> Artificial Sequence

<220><223> HER2-specific CDRL2

<400> 124

Ala Ser Tyr Arg Tyr Thr

1 5

<210> 125

<211> 9

<212> PRT

<213> Artificial Sequence

<220><223> HER2-specific CDRL3

<400> 125

Gln Gln Tyr Tyr Ile Tyr Pro Tyr Thr

1 5

<210> 126

<211> 10

<212> PRT

<213> Artificial Sequence

<220><223> HER2-specific CDRH1

<400> 126

Gly Phe Thr Phe Thr Asp Tyr Thr Met Asp

1 5 10

<210> 127

<211> 16

<212> PRT

<213> Artificial Sequence

<220><223> HER2-specific CDRH2

<400> 127

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

1 5 10 15

<210> 128

<211> 9

<212> PRT

<213> Artificial Sequence

<

220><223> HER2-specific CDRH3

<400> 128

Leu Gly Pro Ser Phe Tyr Phe Asp Tyr

1 5

<210> 129

<211> 15

<212> PRT

<213> Artificial Sequence

<220><223> PD-L1-specific CDRL1

<400> 129

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

1 5 10 15

<210> 130

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> PD-L1-specific CDRL2

<400> 130

Leu Ala Ser Tyr Leu Glu Ser

1 5

<210> 131

<211> 9

<212> PRT

<213> Artificial Sequence

<220><223> PD-L1-specific CDRL3

<400> 131

Gln His Ser Arg Asp Leu Pro Leu Thr

1 5

<210> 132

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> PD-L1-specific CDRH1

<400> 132

Asn Tyr Tyr Met Tyr

1 5

<210> 133

<211> 17

<212> PRT

<213> Artificial Sequence

<220><223> PD-L1-specific CDRH2

<400> 133

Gly Ile Asn Pro Ser Asn Gly Gly Thr Asn Phe Asn Glu Lys Phe Lys

1 5 10 15

Asn

<210> 134

<211> 11

<212> PRT

<213> Artificial Sequence

<220><223> PD-L1-specific CDRH3

<400> 134

Arg Asp Tyr Arg Phe Asp Met Gly Phe Asp Tyr
1 5 10

<210> 135

<211> 120

<212> PRT

<213> Artificial Sequence

<220><223> Avelumab variable heavy chain

<400> 135

Glu Val Gln Leu Leu Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Gly
1 5 10 15

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

Ile Met Met Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val
35 40 45

Ser Ser Ile Tyr Pro Ser Gly Gly Ile Thr Phe Tyr Ala Asp Thr Val
50 55 60

Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ser Lys Asn Thr Leu Tyr
65 70 75 80

Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys
85 90 95

Ala Arg Ile Lys Leu Gly Thr Val Thr Thr Val Asp Tyr Trp Gly Gln
100 105 110

Gly Thr Leu Val Thr Val Ser Ser
115 120

<210> 136

<211> 110

<212> PRT

<213> Artificial Sequence

<220><223> Avelumab variable light chain

<400> 136

Gln Ser Ala Leu Thr Gln Pro Ala Ser Val Ser Gly Ser Pro Gly Gln

1 5 10 15

Ser Ile Thr Ile Ser Cys Thr Gly Thr Ser Ser Asp Val Gly Gly Tyr

20 25 30

Asn Tyr Val Ser Trp Tyr Gln Gln His Pro Gly Lys Ala Pro Lys Leu

35 40 45

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

50 55 60

Ser Gly Ser Lys Ser Gly Asn Thr Ala Ser Leu Thr Ile Ser Gly Leu

65 70 75 80

Gln Ala Glu Asp Glu Ala Asp Tyr Tyr Cys Ser Ser Tyr Thr Ser Ser

85 90 95

Ser Thr Arg Val Phe Gly Thr Gly Thr Lys Val Thr Val Leu

100 105 110

<210> 137

<211> 11

<212> PRT

<213> Artificial Sequence

<220><223> Avelumab-specific CDRH1

<400> 137

Ser Gly Phe Thr Phe Ser Ser Tyr Ile Met Met

1 5 10

<210> 138

<211> 17

<212> PRT

<213> Artificial Sequence

<220><223> Avelumab-specific CDRH2

<400> 138

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

1 5 10 15

Gly

<210> 139

<211> 11

<212> PRT

<213> Artificial Sequence

<220><223> Avelumab-specific CDRH3

<400> 139

Ile Lys Leu Gly Thr Val Thr Thr Val Asp Tyr

1 5 10

<210> 140

<211> 14

<212> PRT

<213> Artificial Sequence

<220><223> Avelumab-specific CDRL1

<400> 140

Thr Gly Thr Ser Ser Asp Val Gly Gly Tyr Asn Tyr Val Ser

1 5 10

<210> 141

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> Avelumab-specific CDRL2

<400> 141

Asp Val Ser Asn Arg Pro Ser

1 5

<210> 142

<211> 10

<212> PRT

<213> Artificial Sequence

<220><223> Avelumab-specific CDRL3

<400> 142

Ser Ser Tyr Thr Ser Ser Ser Thr Arg Val

1 5 10

<210> 143

<211> 118

<212> PRT

<213> Artificial Sequence

<220><223> Atezolizumab-specific variable heavy chain

<400> 143

Glu Val Gln Leu Val Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Gly
1 5 10 15

Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Ser Asp Ser
20 25 30

Trp Ile His Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val
35 40 45

Ala Trp Ile Ser Pro Tyr Gly Gly Ser Thr Tyr Tyr Ala Asp Ser Val
50 55 60

Lys Gly Arg Phe Thr Ile Ser Ala Asp Thr Ser Lys Asn Thr Ala Tyr
65 70 75 80

Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys
85 90 95

Ala Arg Arg His Trp Pro Gly Gly Phe Asp Tyr Trp Gly Gln Gly Thr
100 105 110

Leu Val Thr Val Ser Ser
115

<210> 144

<211> 107

<212> PRT

<213> Artificial Sequence

<220><223> Atezolizumab-specific variable light chain

<400> 144

Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly
1 5 10 15

Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Asp Val Ser Thr Ala
20 25 30

Val Ala Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile

35 40 45

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

50 55 60

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

65 70 75 80

Glu Asp Phe Ala Thr Tyr Tyr Cys Gln Gln Tyr Leu Tyr His Pro Ala

85 90 95

Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys

100 105

<210> 145

<211> 11

<212> PRT

<213> Artificial Sequence

<220><223> Atezolizumab-specific CDRH1

<400> 145

Ser Gly Phe Thr Phe Ser Asp Ser Trp Ile His

1 5 10

<210> 146

<211> 17

<212> PRT

<213> Artificial Sequence

<220><223> Atezolizumab-specific CDRH2

<400> 146

Trp Ile Ser Pro Tyr Gly Gly Ser Thr Tyr Tyr Ala Asp Ser Val Lys

1 5 10 15

Gly

<210> 147

<211> 9

<212> PRT

<213> Artificial Sequence

<220><223> Atezolizumab-specific CDRH3

<400> 147

Arg His Trp Pro Gly Gly Phe Asp Tyr

1 5

<210> 148

<211> 11

<212> PRT

<213> Artificial Sequence

<220><223> Atezolizumab-specific CDRL1

<400> 148

Arg Ala Ser Gln Asp Val Ser Thr Ala Val Ala

1 5 10

<210> 149

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> Atezolizumab-specific CDRL2

<400> 149

Ser Ala Ser Phe Leu Tyr Ser

1 5

<210> 150

<211> 9

<212> PRT

<213> Artificial Sequence

<220><223> Atezolizumab-specific CDRL3

<400> 150

Gln Gln Tyr Leu Tyr His Pro Ala Thr

1 5

<210> 151

<211> 11

<212> PRT

<213> Artificial Sequence

<220><223> PSMA-specific CDRL1

<400> 151

Lys Ala Ser Gln Asp Val Gly Thr Ala Val Asp

1 5 10

<210> 152

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> PSMA-specific CDRL2

<400> 152

Trp Ala Ser Thr Arg His Thr

1 5

<210> 153

<211> 9

<212> PRT

<213> Artificial Sequence

<220><223> PSMA-specific CDRL3

<400> 153

Gln Gln Tyr Asn Ser Tyr Pro Leu Thr

1 5

<210> 154

<211> 10

<212> PRT

<213> Artificial Sequence

<220><223> PSMA-specific CDRH1

<400> 154

Gly Tyr Thr Phe Thr Glu Tyr Thr Ile His

1 5 10

<210> 155

<211> 17

<212> PRT

<213> Artificial Sequence

<220><223> PSMA-specific CDRH2

<400> 155

Asn Ile Asn Pro Asn Asn Gly Gly Thr Thr Tyr Asn Gln Lys Phe Glu

1 5 10 15

Asp

<210> 156

<211> 6

<212> PRT

<213> Artificial Sequence

<220><223> PSMA-specific CDRH3

<400> 156

Gly Trp Asn Phe Asp Tyr

1 5

<210> 157

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> MUC16-specific CDRL1

<400> 157

Ser Glu Asp Ile Tyr Ser Gly

1 5

<210> 158

<211> 8

<212> PRT

<213> Artificial Sequence

<220><223> MUC16-specific CDRL3

<400> 158

Gly Tyr Ser Tyr Ser Ser Thr Leu

1 5

<210> 159

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> MUC16-specific CDRH1

<400> 159

Thr Leu Gly Met Gly Val Gly

1 5

<210> 160

<211> 16

<212> PRT

<213> Artificial Sequence

<220><223> MUC16-specific CDRH2

<400> 160

His Ile Trp Trp Asp Asp Asp Lys Tyr Tyr Asn Pro Ala Leu Lys Ser

1 5 10 15

<210> 161

<211> 12

<212> PRT

<213> Artificial Sequence

<220><223> MUC16-specific CDRH3

<400> 161

Ile Gly Thr Ala Gln Ala Thr Asp Ala Leu Asp Tyr

1 5 10

<210> 162

<211> 15

<212> PRT

<213> Artificial Sequence

<220><223> FOLR-specific CDRL1

<400> 162

Lys Ala Ser Gln Ser Val Ser Phe Ala Gly Thr Ser Leu Met His

1 5 10 15

<210> 163

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> FOLR-specific CDRL2

<400> 163

Arg Ala Ser Asn Leu Glu Ala

1 5

<210> 164

<211> 9

<212> PRT

<213> Artificial Sequence

<220><223> FOLR-specific CDRL3

<400> 164

Gln Gln Ser Arg Glu Tyr Pro Tyr Thr

1 5

<210> 165

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> FOLR-specific CDRH1

<400> 165

Gly Tyr Phe Met Asn

1 5

<210> 166

<211> 17

<212> PRT

<213> Artificial Sequence

<220><223> FOLR-specific CDRH2

<400> 166

Arg Ile His Pro Tyr Asp Gly Asp Thr Phe Tyr Asn Gln Lys Phe Gln

1 5 10 15

Gly

<210> 167

<211> 9

<212> PRT

<213> Artificial Sequence

<220><223> FOLR-specific CDRH3

<400> 167

Tyr Asp Gly Ser Arg Ala Met Asp Tyr

1 5

<210>

168

<211> 119

<212> PRT

<213> Artificial Sequence

<220><223> Amatuximab-specific variable heavy chain

<400> 168

Gln Val Gln Leu Gln Gln Ser Gly Pro Glu Leu Glu Lys Pro Gly Ala
1 5 10 15

Ser Val Lys Ile Ser Cys Lys Ala Ser Gly Tyr Ser Phe Thr Gly Tyr
20 25 30

Thr Met Asn Trp Val Lys Gln Ser His Gly Lys Ser Leu Glu Trp Ile
35 40 45

Gly Leu Ile Thr Pro Tyr Asn Gly Ala Ser Ser Tyr Asn Gln Lys Phe
50 55 60

Arg Gly Lys Ala Thr Leu Thr Val Asp Lys Ser Ser Ser Thr Ala Tyr
65 70 75 80

Met Asp Leu Leu Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr Phe Cys
85 90 95

Ala Arg Gly Gly Tyr Asp Gly Arg Gly Phe Asp Tyr Trp Gly Ser Gly
100 105 110

Thr Pro Val Thr Val Ser Ser
115

<210> 169

<211> 106

<212> PRT

<213> Artificial Sequence

<220><223> Amatuximab-specific variable light chain

<400> 169

Asp Ile Glu Leu Thr Gln Ser Pro Ala Ile Met Ser Ala Ser Pro Gly
1 5 10 15

Glu Lys Val Thr Met Thr Cys Ser Ala Ser Ser Ser Val Ser Tyr Met
20 25 30

His Trp Tyr Gln Gln Lys Ser Gly Thr Ser Pro Lys Arg Trp Ile Tyr

35 40 45

Asp Thr Ser Lys Leu Ala Ser Gly Val Pro Gly Arg Phe Ser Gly Ser

50 55 60

Gly Ser Gly Asn Ser Tyr Ser Leu Thr Ile Ser Ser Val Glu Ala Glu

65 70 75 80

Asp Asp Ala Thr Tyr Tyr Cys Gln Gln Trp Ser Lys His Pro Leu Thr

85 90 95

Phe Gly Ser Gly Thr Lys Val Glu Ile Lys

100 105

<210> 170

<211> 10

<212> PRT

<213> Artificial Sequence

<220><223> Amatuximab-specific CDRH1

<400> 170

Gly Tyr Ser Phe Thr Gly Tyr Thr Met Asn

1 5 10

<210> 171

<211> 13

<212> PRT

<213> Artificial Sequence

<220><223> Amatuximab-specific CDRH2

<400> 171

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

1 5 10

<210> 172

<211> 10

<212> PRT

<213> Artificial Sequence

<220><223> Amatuximab-specific CDRH3

<400> 172

Gly Gly Tyr Asp Gly Arg Gly Phe Asp Tyr

1 5 10

<210> 173

<211> 10

<212> PRT

<213> Artificial Sequence

<220><223> Amatuximab-specific CDRL1

<400> 173

Ser Ala Ser Ser Ser Val Ser Tyr Met His

1 5 10

<210> 174

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> Amatuximab-specific CDRL2

<400> 174

Asp Thr Ser Lys Leu Ala Ser

1 5

<210> 175

<211> 9

<212> PRT

<213> Artificial Sequence

<220><223> Amatuximab-specific CDRL3

<400> 175

Gln Gln Trp Ser Lys His Pro Leu Thr

1 5

<210> 176

<211> 32

<212> PRT

<213> Artificial Sequence

<220><223> Nef (66-97)

<400> 176

Val Gly Phe Pro Val Thr Pro Gln Val Pro Leu Arg Pro Met Thr Tyr

1 5 10 15

Lys Ala Ala Val Asp Leu Ser His Phe Leu Lys Glu Lys Gly Gly Leu

20 25 30

<210> 177

<211> 30

<212> PRT

<213> Artificial Sequence

<220><223> Nef (116-145)

<400> 177

His Thr Gln Gly Tyr Phe Pro Asp Trp Gln Asn Tyr Thr Pro Gly Pro

1 5 10 15

Gly Val Arg Tyr Pro Leu Thr Phe Gly Trp Leu Tyr Lys Leu

20 25 30

<210> 178

<211> 19

<212> PRT

<213> Artificial Sequence

<220><223> Gag p17 (17-35)

<400> 178

Glu Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys Lys Tyr Lys Leu Lys

1 5 10 15

His Ile Val

<210> 179

<211> 32

<212> PRT

<213> Artificial Sequence

<220><223> Gag p17-p24 (253-284)

<400> 179

Asn Pro Pro Ile Pro Val Gly Glu Ile Tyr Lys Arg Trp Ile Ile Leu

1 5 10 15

Gly Leu Asn Lys Ile Val Arg Met Tyr Ser Pro Thr Ser Ile Leu Asp

20 25 30

<210> 180

<211> 31

<212> PRT

<213> Artificial Sequence

<220><223> Pol 325-355 (RT 158-188)

<400> 180

Ala Ile Phe Gln Ser Ser Met Thr Lys Ile Leu Glu Pro Phe Arg Lys
1 5 10 15

Gln Asn Pro Asp Ile Val Ile Tyr Gln Tyr Met Asp Asp Leu Tyr
20 25 30

<210> 181

<211> 11

<212> PRT

<213> Artificial Sequence

<220><223> IgG4 linker

<400> 181

Glu Ser Lys Tyr Gly Pro Pro Cys Pro Pro Cys
1 5 10

<210> 182

<211> 11

<212> PRT

<213> Artificial Sequence

<220><223> CD28 linker

<400> 182

Pro Ser Pro Leu Phe Pro Gly Pro Ser Lys Pro
1 5 10

<210> 183

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> spacer

<220><221> MISC_FEATURE

<222> (1)..(5)

<223> Residues (GGGGS) can be repeated n times, wherein n is an integer
including, 1, 2, 3, 4, 5, 6, 7, 8, 9, or more

<400> 183

Gly Gly Gly Gly Ser

1 5

<210> 184

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> spacer

<220><221> MISC_FEATURE

<222> (1)..(5)

<223> (EAAAK) can be repeated n times, wherein n is an integer including 1, 2, 3, 4, 5, 6, 7, 8, 9, or more

<400> 184

Glu Ala Ala Ala Lys

1 5

<210> 185

<211> 21531

<212> DNA

<213> Homo sapiens

<400> 185

gatctctatc ccctcctgtt ttctctacgt tatttatatg ggtatcatca ccatcctgga 60

caacatcagg acagatatcc ctcaccaagc caatgttcct ctctatgttg gctcaaatgt 120

ccttgaactt tcctttcacc acccctttcca cagtcaaaag gatattgtag tttaatgcct 180

cagagttcag cttttaagct tctgacaaat tattcttcct ctttaggttc tcctttatgg 240

aatcttctgt actgatggcc atgtccttta actactatgt agatatctgc tactacctgt 300

attatgcctc tacctttatt agcagagtta tctgtactgt tggcatgaca atcatttgtt 360

aatatgactt gcctttcctt tttctgctat tcttgatcaa atggctcctc tttcttgctc 420

ctctcatttc tcctgccttc acttggacgt gcttcacgta gtctgtgctt atgactggat 480

taaaaattga tatggactta tcctaatgtt gttcgtcata atatgggttt tatggtccat 540

tattatttcc tatgcattga tctggagaag gcttcaatcc ttttactctt tgtggaaaat 600

atctgtaaac cttctggttc actctgctat agcaatttca gtttaggcta gtaagcatga 660

ggatgcctcc ttctctgatt tttcccacag tctgttggtc acagaataac ctgagtgatt 720

actgatgaaa gagtgagaat gttattgata gtcacaatga caaaaaacaa acaactacag 780

tcaaaatgtt tctcttttta ttagtggatt atatttcctg acctatatct ggcaggactc 840

tttagagagg tagctgaagc tgctgttatg accactagag ggaagaagat acctgtggag 900

ctaatggtcc aagatggtgg agccccaagc aaggaagttg ttaaggagcc cttttgattg 960

aaggtgggtg cccccacctt acagggacag gacatctgga tactcctccc agtttctcca 1020

gtttcccttt ttcctaatat atctcctgat aaaatgtcta tactcacttc cccatttcta 1080

ataataaagc aaaggctagt tagtaagaca tcaccttgca ttttgaaaat gccatagact 1140

ttcaaaatta tttcatacat cggtctttct ttatttcaag agtccagaaa tggcaacatt 1200

acctttgatt caatgtaatg gaaagagctc tttcaagaga cagagaaaag aataatttaa 1260

tttctttccc cacacctcct tccctgtctc ttaccctatc ttccttcctt ctaccctccc 1320

catttctctc tctcatttct cagaagtata ttttgaaagg attcatagca gacagctaag 1380

gctggttttt tctaagtgaa gaagtgatat tgagaaggta gggttgcatg agccctttca 1440

gtttttttagt ttatatacat ctgtattgtt agaatgtttt ataatataaa taaaattatt 1500

tctcagttat atactagcta tgtaacctgt ggatatttcc ttaagtatta caagctatac 1560

ttaactcact tggaaaactc aaataaatac ctgcttcata gttattaata aggattaagt 1620

gagataatgc ccataagatt cctattaata acagataaat acatacacac acacacacat 1680

tgaaaggatt cttactttgt gctaggaact ataataagtt cattgatgca ttatatcatt 1740

aagttctaat ttcaacacta gaaggcaggt attatctaaa tttcatactg gatacctcca 1800

aactcataaa gataattaaa ttgccttttg tcatatattt attcaaaagg gtaaactcaa 1860

actatggctt gtctaatttt atatatcacc ctactgaaca tgaccctatt gtgatatttt 1920

ataaaattat tctcaagtta ttatgaggat gttgaaagac agagaggatg gggtgctatg 1980

ccccaaatca gcctcacaat taagctaagc agctaagagt cttgcagggt agtgtaggga 2040

ccacagggtt aaggggggcag tagaattata ctcccacttt agtttcattt caaacaatcc 2100

atacacacac agccctgagc acttacaaat tatactacgc tctatacttt ttgtttaaat 2160

gtataaataa gtggatgaaa gaatagatag atagatagac agatagatga tagatagaat 2220

aaatgcttgc cttcatagct gtctccctac cttgttcaaa atgttcctgt ccagaccaaa 2280

gtaccttgcc ttcacttaag taatcaattc ctaggttata ttctgatgtc aaaggaagtc 2340

aaaagatgtg aaaaacaatt tctgacccac aactcatgct ttgtagatga ctagatcaaa 2400

aaatttcagc catatcttaa cagtgagtga acaggaaatc tcctcttttc cctacatctg 2460

agatcccagc ttctaagacc ttcaattctc actcttgatg caacagacct tggaagcata 2520

caggagagct gaacttggtc aacaaaggag aaaagtttgt tggcctccaa aggcacagct 2580

caaacttttc aagccttctc taatcttaaa ggtaaacaag ggtctcattt ctttgagaac 2640

ttcagggaaa atagacaagg acttgcctgg tgcttttggt aggggagctt gcactttccc 2700
cctttctgga ggaaatattt atccccaggt agttcccttt ttgcaccagt ggttctttga 2760
agagacttcc acctgggaac agttaaacag caactacagg gccttgaact gcacactttc 2820
agtccggtcc tcacagttga aaagacctaa gcttgtgcct gatttaagcc tttttggtca 2880
taaaacattg aattctaatc tccctctcaa ccctacagtc acccatttgg tatattaaag 2940
atgtgttgtc tactgtctag tatccctcaa gtagtgtcag gaattagtca tttaaatagt 3000
ctgcaagcca ggagtggtgg ctcatgtctg taattccagc acttgagagg tagaagtggg 3060

aggactgctt gagctcaaga gtttgatatt atcctggaca acatagcaag acctcgtctc 3120
tacttaaaaa aaaaaaaaaa attagccagg catgtgatgt acacctgtag tcccagctac 3180
tcaggaggcc gaaatgggag gatcccttga gctcaggagg tcaaggctgc agtgagacat 3240
gatcttgcca ctgcactcca gcctggacag cagagtgaaa ccttgcctca cgaaacagaa 3300
tacaaaaaca aacaaacaaa aaactgctcc gcaatgcgct tccttgatgc tctaccacat 3360
aggtctgggt actttgtaca cattatctca ttgctgttca taattgttag attaattttg 3420
taatattgat attattccta gaaagctgag gcctcaagat gataactttt attttctgga 3480

cttgtaatag ctttctcttg tattcaccat gttgtaactt tcttagagta gtaacaatat 3540
aaagttattg tgagtttttg caaacacagc aaacacaacg acccatatag acattgatgt 3600
gaaattgtct attgtcaatt tatgggaaaa caagtatgta ctttttctac taagccattg 3660
aaacaggaat aacagaacaa gattgaaaga atacattttc cgaaattact tgagtattat 3720
acaaagacaa gcacgtggac ctgggaggag ggttattgtc catgactggt gtgtggagac 3780
aaatgcaggt ttataataga tgggatggca tctagcgcaa tgactttgcc atcactttta 3840
gagagctctt ggggacccca gtacacaaga ggggacgcag ggtatatgta gacatctcat 3900

tctttttctt agtgtgagaa taagaatagc catgacctga gtttatagac aatgagccct 3960
tttctctctc ccactcagca gctatgagat ggcttgccct gcctctctac taggctgact 4020
cactccaagg cccagcaatg ggcagggctc tgtcagggct ttgatagcac tatctgcaga 4080
gccagggccg agaaggggtg gactccagag actctccctc ccattcccga gcagggtttg 4140
cttatttatg catttaaatg atatatttat tttaaaagaa ataacaggag actgcccagc 4200
cctggctgtg acatggaaac tatgtagaat attttgggtt ccatttttt ttccttcttt 4260
cagttagagg aaaaggggct cactgcacat acactagaca gaaagtcagg agctttgaat 4320

ccaagcctga tcatttccat gtcatactga gaaagtcccc acccttctct gagcctcagt 4380

ttctcttttt ataagtagga gtctggagta aatgatttcc aatggctctc atttcaatac 4440

aaaatttccg tttattaaat gcatgagctt ctgttactcc aagactgaga aggaaattga 4500

acctgagact cattgactgg caagatgtcc ccagaggctc tcattcagca ataaaattct 4560

caccttcacc caggcccact gagtgtcaga tttgcatgca ctagttcacg tgtgtaaaaa 4620

ggaggatgct tctttccttt gtattctcac atacctttag gaaagaactt agcacccttc 4680

ccacacagcc atcccaataa ctcatttcag tgactcaacc cttgacttta taaaagtctt 4740

gggcagtata gagcagagat taagagtaca gatgctggag ccagaccacc tgagtgatta 4800

gtgactcagt ttctcttagt agttgtatga ctcagtttct tcatctgtaa aatggagggt 4860

tttttaatta gtttgttttt gagaaagggt ctcactctgt cacccaaatg ggagtgtagt 4920

ggcaaaatct cggctcactg caacttgcac ttcccaggct caagcggtcc tcccacctca 4980

acatcctgag tagctggaac cacaggtaca caccaccata cctcgctaat tttttgtatt 5040

tttggtagag atggggtttc acatgttaca caggatggtc tcagactccg gagctcaagc 5100

aatctgccca cctcagcctt ccaaagtgct gggattataa gcatgattac aggagtttta 5160

acaggctcat aagattgttc tgcagcccga gtgagttaat acatgcaaag agtttaaagc 5220

agtgacttat aaatgctaac tactctagaa atgtttgcta gtatttttg tttaactgca 5280

atcattcttg ctgcaggtga aaactagtgt tctgtacttt atgcccattc atctttaact 5340

gtaataataa aaataactga catttattga aggctatcag agactgtaat tagtgctttg 5400

cataattaat catatttaat actcttggat tctttcaggt agatactatt attatcccca 5460

ttttactaca gttaaaaaaa ctacctctca acttgctcaa gcatacactc tcacacacac 5520

aaacataaac tactagcaaa tagtagaatt gagatttggt cctaattatg tctttgctca 5580

ctatccaata aatatttatt gacatgtact tcttggcagt ctgtatgctg gatgctgggg 5640

atacaaagat gtttaaattt aagctccagt ctctgcttcc aaaggcctcc caggccaagt 5700

tatccattca gaaagcattt tttactcttt gcattccact gtttttccta agtgactaaa 5760

aaattacact ttattcgtct gtgtcctgct ctgggatgat agtctgactt tcctaacctg 5820

agcctaacat ccctgacatc aggaaagact acaccatgtg gagaaggggt ggtggttttg 5880

attgctgctg tcttcagtta gatggttaac tttgtgaagt tgaaaactgt ggctctctgg 5940

ttgactgtta gagttctggc acttgtcact atgcctatta tttaacaaat gcatgaatgc 6000

ttcagaatat gggaatatta tcttctggaa tagggaatca agttatatta tgtaacccag 6060

gattagaaga ttcttctgtg tgtaagaatt tcataaacat taagctgtct agcaaaagca 6120

agggcttgga aaatctgtga gctcctcacc atatagaaag cttttaaccc atcattgaat 6180

aaatccctat aggggatttc taccctgagc aaaaggctgg tcttgattaa ttcccaaact 6240

catatagctc tgagaaagtc tatgctgtta acgttttctt gtctgctacc ccatcatatg 6300

cacaacaata aatgcaggcc taggcatgac tgaaggctct ctcataattc ttggttgcat 6360

gaatcagatt atcaacagaa atgttgagac aaactatggg gaagcagggt atgaaagagc 6420

tctgaatgaa atggaaaccg caatgcttcc tgcccattca gggctccagc atgtagaaat 6480

ctggggcttt gtgaagactg gcttaaaatc agaagcccca ttggataaga gtagggaaga 6540

acctagagcc tacgctgagc aggtttcctt catgtgacag ggagcctcct gccccgaact 6600

tccagggatc ctctcttaag tgtttcctgc tggaatctcc tcacttctat ctggaaatgg 6660

tttctccaca gtccagcccc tggctagttg aaagagttac ccatgcagag gccctcctag 6720

catccagaga ctagtgctta gattcctact ttcagcgttg gacaacctgg atccacttgc 6780

ccagtgttct tccttagttc ctaccttcga ccttgatcct cctttatctt cctgaaccct 6840

gctgagatga tctatgtggg gagaatggct tctttgagaa acatcttctt cgttagtggc 6900

ctgcccctca ttcccacttt aatatccaga atcactataa gaagaatata ataagaggaa 6960

taactcttat tataggtaag ggaaaattaa gaggcatacg tgatgggatg agtaagagag 7020

gagagggaag gattaatgga cgataaaatc tactactatt tgttgagacc ttttatagtc 7080

taatcaattt tgctattgtt ttccatcctc acgctaactc cataaaaaaa cactattatt 7140

atctttattt tgccatgaca agactgagct cagaagagtc aagcatttgc ctaaggtcgg 7200

acatgtcaga ggcagtgcca gacctatgtg agactctgca gctactgctc atgggccctg 7260

tgctgcactg atgaggagga tcagatggat ggggcaatga agcaaaggaa tcattctgtg 7320

gataaaggag acagccatga agaagtctat gactgtaaat ttgggagcag gagtctctaa 7380

ggacttggat ttcaaggaat tttgactcag caaacacaag accctcacgg tgactttgcg 7440

agctggtgtg ccagatgtgt ctatcagagg ttccagggag ggtggggtgg ggtcagggct 7500

ggccaccagc tatcagggcc cagatgggtt ataggctggc aggctcagat aggtggttag 7560

gtcaggttgg tggtgctggg tggagtccat gactcccagg agccaggaga gatagaccat 7620

gagtagaggg cagacatggg aaaggtgggg gaggcacagc atagcagcat ttttcattct 7680

actactacat gggactgctc ccctataccc ccagctaggg gcaagtgcct tgactcctat 7740

gttttcagga tcatcatcta taaagtaaga gtaataattg tgtctatctc atagggttat 7800

tatgaggatc aaaggagatg cacactctct ggaccagtgg cctaacagtt caggacagag 7860

ctatgggctt cctatgtatg ggtcagtggt ctcaatgtag caggcaagtt ccagaagata 7920

gcatcaacca ctgttagaga tatactgcca gtctcagagc ctgatgttaa tttagcaatg 7980

ggctgggacc ctcctccagt agaaccttct aaccagctgc tgcagtcaaa gtcgaatgca 8040

gctggttaga ctttttttaa tgaaagctta gctttcatta aagattaagc tcctaagcag 8100

ggcacagatg aaattgtcta acagcaactt tgccatctaa aaaaatctga cttcactgga 8160
aacatggaag cccaaggttc tgaacatgag aaatttttag gaatctgcac aggagttgag 8220
agggaaacaa gatggtgaag ggactagaaa ccacatgaga gacacgagga aatagtgtag 8280
atttaggctg gaggtaaatg aaagagaagt gggaattaat acttactgaa atctttctat 8340
atgtcaggtg ccattttatg atatttaata atctcattac atatggtaat tctgtgagat 8400
atgtattatt gaacatacta taattaatac taatgataag taacacctct tgagtactta 8460
gtatatgcta gaatcaaatt taagtttatc atatgaggcc gggcacggtg gctcatatat 8520

gggattacat gcctgtaatc ccagcacttt gggaggccaa ggcaattgga tcacctgagg 8580
tcaggagttc cagaccagcc tggccaacat ggtgaaaccc cttctctact aaaaaataca 8640
aaaaatcagc caggtgtggt ggcacgcgtc tataatccca gctactcagg aggctgaggc 8700
aggagaatca cttgaaccca ggaggtggag gttgcagtga gctaagattg caccactgca 8760
ctccagccta ggcgacagag tgagactcca tctcaaaaaa aaaaaaagaa gtttattata 8820
tgaattaact tagttttact cacaccaata ctcagaagta gattattacc tcatttattg 8880
atgaggagcc caatgtactt gtagtgtaga tcaacttatt gaaagcacaa gctaataagt 8940

agacaattag taattagaag tcagatggtc tgagctctcc tactgtctac attacatgag 9000
ctcttattaa ctggggactc gaaaatcaaa gacatgaaat aatttgtcca agcttacaga 9060
accaccaagt agtaaggcta ggatgtagac ccagttctgc tacctctgaa gacagtgttt 9120
tttccacagc aaaacacaaa ctcagatatt gtggatgcga gaaattagaa gtagatattc 9180
ctgccctgtg gcccttgctt cttacttta cttcttgtcg attggaagtt gtggtccaag 9240
ccacagttgc agaccatact tcctcaacca taattgcatt tcttcaggaa agtttgaggg 9300
agaaaaaggt aaagaaaaat ttagaaacaa cttcagaata aagagatttt ctcttgggtt 9360

acagagattg tcatatgaca aattataagc agacacttga gaaaactgaa ggcccatgcc 9420
tgcccaaatt acccttgac cccttggtca agctgcaact ttggttaaag ggagtgttta 9480
tgtgttatag tgttcattta ctcttctggt ctaacccatt ggctccgtct tcatcctgca 9540
gtgacctcag tgcctcagaa acatacatat gtttgtctag tttaagtttg tgtgaaattc 9600
taactagcgt caagaactga gggccctaaa ctatgctagg aatagtgctg tggtgctgtg 9660
ataggtacac aagaaatgag aagaaactgc agattctctg catctccctt tgccgggtct 9720
gacaacaaag tttccccaaa ttttaccaat gcaagccatt tctccatatg ctaactactt 9780

taaaatcatt tggggcttca cattgtcttt ctcatctgta aaaagaatgg aagaactcat 9840

tcctacagaa ctccctatgt cttccctgat gggctagagt tcctctttct caaaaattag 9900

ccattattgt atttccttct aagccaaagc tcagaggtct tgtattgccc agtgacatgc 9960

acactggtca aaagtaggct aagtagaagg gtactttcac aggaacagag agcaaaagag 10020

gtgggtgaat gagagggtaa gtgagaaaag acaaatgaga agttacaaca tgatggcttg 10080

ttgtctaaat atctcctagg gaattattgt gagaggtctg aatagtgttg taaaataagc 10140

tgaatctgct gccaacatta acagtcaaga aatacctccg aataactgta cctccaatta 10200

ttctttaagg tagcatgcaa ctgtaatagt tgcatgtata tatttatcat aatactgtaa 10260

cagaaaacac ttactgaata tatactgtgt ccctagttct ttacacaata aactaatctc 10320

atcctcataa ttctattagc taatacatat tatcatccta tatttcagag acttcaagaa 10380

gttaagcaac ttgctcaaga tcatctaaga agtaggtggt atttctgggc tcatttggcc 10440

cctcctaatc tctcatggca acatggctgc ctaaagtgtt gattgcctta attcatcagg 10500

gatgggctca tactcactgc agaccttaac tggcatcctc ttttcttatg tgatctgcct 10560

gaccctagta gacttatgaa atttctgatg agaaaggaga gaggagaaag gcagagctga 10620

ctgtgatgag tgatgaaggt gccttctcat ctgggtacca gtggggcctc taagactaag 10680

tcactctgtc tcactgtgtc ttagccagtt ccttacagct tgccctgatg ggagatagag 10740

aatgggtatc ctccaacaaa aaaataaatt ttcatttctc aaggtccaac ttatgttttc 10800

ttaatttta aaaaaatctt gaccattctc cactctctaa aataatccac agtgagagaa 10860

acattctttt cccccatccc ataaatacct ctattaaata tggaaaatct gggcatggtg 10920

tctcacacct gtaatcccag cactttggga ggctgaggtg ggtggactgc ttggagctca 10980

ggagttcaag accatcttgg acaacatggt gataccctgc ctctacaaaa agtacaaaaa 11040

ttagcctggc atggtggtgt gcacctgtaa tcccagctat tagggtggct gaggcaggag 11100

aattgcttga acccgggagg cggaggttgc agtgagctga gatcgtgcca ctgcactcca 11160

gcctgggggа cagagcacat tataattaac tgttattttt tacttggact cttgtgggga 11220

ataagataca tgttttattc ttatttatga ttcaagcact gaaaatagtg tttagcatcc 11280

agcaggtgct tcaaaaccat ttgctgaatg attactatac tttttacaag ctcagctccc 11340

tctatccctt ccagcatcct catctctgat taaataagct tcagtttttc cttagttcct 11400

gttacatttc tgtgtgtctc cattagtgac ctcccatagt ccaagcatga gcagttctgg 11460

ccaggcccct gtcggggtca gtgccccacc cccgccttct ggttctgtgt aaccttctaa 11520

gcaaaccttc tggctcaagc acagcaatgc tgagtcatga tgagtcatgc tgaggcttag 11580

ggtgtgtgcc cagatgttct cagcctagag tgatgactcc tatctgggtc cccagcagga 11640

tgcttacagg gcagatggca aaaaaaagga gaagctgacc acctgactaa aactccacct 11700

caaacggcat cataaagaaa atggatgcct gagacagaat gtgacatatt ctagaatata 11760

ttatttcctg aatatatata tatatataca catatacgta tatatatata tatatatata 11820

tttgttgtta tcaattgcca tagaatgatt agttattgtg aatcaaatat ttatcttgca 11880

ggtggcctct atacctagaa gcggcagaat caggctttat taatacatgt gtatagattt 11940

ttaggatcta tacacatgta ttaatatgaa acaaggatat ggaagaggaa ggcatgaaaa 12000

caggaaaaga aaacaaacct tgtttgccat tttaaggcac ccctggacag ctaggtggca 12060

aaaggcctgt gctgttagag gacacatgct cacatacggg gtcagatctg acttggggtg 12120

ctactgggaa gctctcatct taaggataca tctcaggcca gtcttggtgc attaggaaga 12180

tgtaggcaac tctgatcctg agaggaaaga aacattcctc caggagagct aaaagggttc 12240

acctgtgtgg gtaactgtga aggactacaa gaggatgaaa aacaatgaca gacagacata 12300

atgcttgtgg gagaaaaaac aggaggtcaa ggggatagag aaggcttcca gaagaatggc 12360

tttgaagctg gcttctgtag gagttcacag tggcaaagat gtttcagaaa tgtgacatga 12420

cttaaggaac tatacaaaaa ggaacaaatt taaggagagg cagataaatt agttcaacag 12480

acatgcaagg aattttcaga tgaatgttat gtctccactg agcttcttga ggttagcagc 12540

tgtgagggtt ttgcaggccc aggacccatt acaggacctc acgtatactt gacactgttt 12600

tttgtattca tttgtgaatg aatgacctct tgtcagtcta ctcggtttcg ctgtgaatga 12660

atgatgtctt gtcagcctac ttggtttcgc taagagcaca gagagaagat ttagtgatgc 12720

tatgtaaaaa cttccttttt ggttcaagtg tatgtttgtg atagaaatga agacaggcta 12780

catgatgcat atctaacata aacacaaaca ttaagaaagg aaatcaacct gaagagtatt 12840

tatacagata acaaaataca gagagtgagt taaatgtgta ataactgtgg cacaggctgg 12900

aatatgagcc atttaaatca caaattaatt agaaaaaaaa cagtggggaa aaaattccat 12960

ggatgggtct agaaagacta gcattgtttt aggttgagtg gcagtgttta aagggtgata 13020

tcagactaaa cttgaaatat gtggctaaat aactagaata ctctttattt tttcgtatca 13080

tgaatagcag atatagcttg atggccccat gcttggttta acatccttgc tgttcctgac 13140

atgaaatcct taatttttga caaaggggct attcattttc attttatatt gggcctagaa 13200

attatgtaga tggtcctgag gaaaagttta tagcttgtct atttctctct ctaacatagt 13260

tgtcagcaca atgcctaggc tataggaagt actcaaagct tgttaaattg aattctatcc 13320

ttcttattca attctacaca tggaggaaaa actcatcagg gatggaggca cgcctctaag 13380

gaaggcaggt gtggctctgc agtgtgattg ggtacttgca ggacgaaggg tggggtggga 13440

gtggctaacc ttccattcct agtgcagagg tcacagccta aacatcaaat tccttgaggt 13500

gcggtggctc actcctgtaa tcacagcagt ttgggacgcc aaggtgggca gatcacttga 13560

ggtcaggagt tggacaccag cccagccaac atagtgaaac ctggtctctg cttaaaaata 13620
taaaaattag ctggacgtgg tgacgggagc ctgtaatcca actacttggg aggctgaggc 13680
aggagaatcg cttgaaccgg ggaggtggag tttgcactga gcagagatca tgccattgca 13740
ctccagcctc cagagcgaga ctctgtctaa agaaaaacga aaacaaacaa acaaacaaac 13800
aaacaaaacc catcaaattc cctgaccgaa cagaattctg tctgattgtt ctctgactta 13860
tctaccattt tccctcctta aagaaactgt gaacttcctt cagctagagg ggcctggctc 13920
agaagcctct ggtcagcatc caagaaatac ttgatgtcac tttggctaaa ggtatgatgt 13980

gtagacaagc tccagagatg gtttctcatt tccatatcca cccacccagc tttccaattt 14040
taaagccaat tctgaggtag agactgtgat gaacaaacac cttgacaaaa ttcaacccaa 14100
agactcactt tgcctagctt caaaatcctt actctgacat atactcacag ccagaaatta 14160
gcatgcacta gagtgtgcat gagtgcaaca cacacacaca ccaattccat attctctgtc 14220
agaaaatcct gttggttttt cgtgaaagga tgttttcaga ggctgacccc ttgccttcac 14280
ctccaatgct accactctgg tctaagtcac tgtcaccacc acctaaatta tagctgttga 14340
ctcataacaa tcttcctgct tctaccactg ccccactaca atttcttccc aatatactat 14400

ccaaattagt cttttcaaaa tgtaagtcat atatggtcac ctctttgttc aaagtcttct 14460
gatagtttcc tatatcattt ataataaaac caaatcctta caattctcta caatagttgt 14520
tcatgcatat attatgttta ttacagatac atatatatag ctctcatata aataaatata 14580
tatatttatg tgtatgtgtg tagagtgttt tttcttacaa ctctatgatg taggtattat 14640
tagtgtccca aattttataa tttaggactt ctatgatctc atcttttatt ctccccttca 14700
ccgaatctca tcctacattg gccttattga tattccttga aaattctaag catcttacat 14760
ctttagggta tttacatttg ccattcccta tgccctaaat atttaatcat agtttcatat 14820

aaatgggttc ctcatcatct atgggtactc tctcaggtgt taactttata gtgaggactt 14880
tcctgccata ctacttaaag tagcgatacc ctttcaccct gtcctaatca cactctggcc 14940
ttcatttcag ttttttttttt ttctccatag cacctaatct cattggtata taacatgttt 15000
catttgctta tttaatgtca agctctttcc actatcaagt ccatgaaaac aggaacttta 15060
ttcctctatt ctgtttttgt gctgtattct tagcaatttt acaattttga atgaatgaat 15120
gagcagtcaa acacatatac aactataatt aaaaggatgt atgctgacac atccactgct 15180
atgcacacac aaagaaatca gtggagtaga gctggaagtg ctaagcctgc atagagctag 15240

ttagccctcc gcaggcagag ccttgatggg attactgagt tctagaattg gactcatttg 15300

ttttgtaggc tgagatttgc tcttgaaaac ttgttctgac caaaataaaa ggctcaaaag 15360

atgaatatcg aaaccagggt gttttttaca ctggaattta taactagagc actcatgttt 15420

atgtaagcaa ttaattgttt catcagtcag gtaaaagtaa agaaaaactg tgccaaggca 15480

ggtagcctaa tgcaatatgc cactaaagta aacattattt cataggtgtc agatatggct 15540

tattcatcca tcttcatggg aaggatggcc ttggcctgga catcagtgtt atgtgaggtt 15600

caaaacacct ctaggctata aggcaacaga gctccttttt ttttttctg tgctttcctg 15660

gctgtccaaa tctctaatga taagcatact tctattcaat gagaatattc tgtaagatta 15720

tagttaagaa ttgtgggagc cattccgtct cttatagtta aatttgagct tcttttatga 15780

tcactgtttt tttaatatgc tttaagttct ggggtacatg tgccatggtg gtttgctgca 15840

cccatcaacc cgtcatctac attaggtatt tctcctaatg ctatccttcc cctagccccc 15900

caccccaac aggccccagt gtgtgatgtt cccctccctg tgtccatgga tcactggttt 15960

ttttttgttt ttttttttt tttaaagtct cagttaaatt tttggaatgt aatttatttt 16020

cctggtatcc taggacttgc aagttatctg gtcactttag ccctcacgtt ttgatgataa 16080

tcacatattt gtaaacacaa cacacacaca cacacacaca cacatatata tatatataaa 16140

acatatatat acataaacac acataacata tttatcgggc atttctgagc aactaatcat 16200

gcaggactct caaacactaa cctatagcct tttctatgta tctacttgtg tagaaaccaa 16260

gcgtggggac tgagaaggca atagcaggag cattctgact ctcactgcct ttagctaggc 16320

ccctccctca tcacagctca gcatagtcct gagctcttat ctatatccac acacagtttc 16380

tgacgctgcc cagctatcac catcccaagt ctaaagaaaa aaataatggg tttgcccatc 16440

tctgttgatt agaaaacaaa acaaaataaa ataagcccct aagctcccag aaaacatgac 16500

taaaccagca agaagaagaa aatacaatag gtatatgagg agactggtga cactagtgtc 16560

tgaatgaggc ttgagtacag aaaagaggct ctagcagcat agtggtttag aggagatgtt 16620

tctttccttc acagatgcct tagcctcaat aagcttgcgg ttgtggaagt ttactttcag 16680

aacaaactcc tgtggggcta gaattattga tggctaaaag aagcccgggg gagggaaaaa 16740

tcattcagca tcctcaccct tagtgacaca aaacagaggg ggcctggttt tccatatttc 16800

ctcatgatgg atgatctcgt taatgaaggt ggtctgacga gatcattgct tcttccattt 16860

aagccttgct cacttgccaa tcctcagttt taaccttctc cagagaaata cacatttttt 16920

attcaggaaa catactatgt tatagtttca atactaaata atcaaagtac tgaagatagc 16980

atgcataggc aagaaaaagt ccttagcttt atgttgctgt tgtttcagaa tttaaaaaag 17040

atcaccaagt caaggacttc tcagttctag cactagaggt ggaatcttag catataatca 17100

gaggtttttc aaaatttcta gacataagat tcaaagccct gcacttaaaa tagtctcatt 17160

tgaattaact ctttatataa attgaaagca cattctgaac tacttcagag tattgtttta 17220

tttctatgtt cttagttcat aaatacatta ggcaatgcaa tttaattaaa aaaacccaag 17280

aatttcttag aattttaatc atgaaaataa atgaaggcat ctttacttac tcaaggtccc 17340

aaaaggtcaa agaaaccagg aaagtaaagc tatatttcag cggaaaatgg gatatttatg 17400

agttttctaa gttgacagac tcaagtttta accttcagtg cccatcatgt aggaaagtgt 17460

ggcataactg gctgattctg gctttctact ccttttttccc attaaagatc cctcctgctt 17520

aattaacatt cacaagtaac tctggttgta ctttaggcac agtggctccc gaggtcagtc 17580

acacaatagg atgtctgtgc tccaagttgc cagagagaga gattactctt gagaatgagc 17640

ctcagccctg gctcaaactc acctgcaaac ttcgtgagag atgaggcaga ggtacactac 17700

gaaagcaaca gttagaagct aaatgatgag aacacatgga ctcatagagg gaaacaacgc 17760

atactggggc ctatcagagg gtggagggtg agagaaggag aggatcagga aaaatcacta 17820

atggatgcta agcgtaatac ctgagtgatg agatcatcta tacaacaaac ccccttgaca 17880

ttcatttatc tatgtaacaa acctgcacat cctgtacatg tacccctgaa cttaaaataa 17940

aagttgaaaa caagaaagca acagtttgaa cacttgttat ggtctattct ctcattcttt 18000

acaattacac tagaaaatag ccacaggctt cctgcaaggc agccacagaa tttatgactt 18060

gtgatatcca agtcattcct ggataatgca aaatctaaca caaaatctag tagaatcatt 18120

tgcttacatc tatttttgtt ctgagaatat agatttagat acataatgga agcagaataa 18180

tttaaaatct ggctaattta gaatcctaag cagctctttt cctatcagtg gtttacaagc 18240

cttgtttata tttttcctat tttaaaaata aaaataaagt aagttatttg tggtaaagaa 18300

tattcattaa agtatttatt tcttagataa taccatgaaa aacattcagt gaagtgaagg 18360

gcctacttta cttaacaaga atctaattta tataattttt catactaata gcatctaaga 18420

acagtacaat atttgactct tcaggttaaa catatgtcat aaattagcca gaaagattta 18480

agaaaatatt ggatgtttcc ttgtttaaat taggcatctt acagttttta gaatcctgca 18540

tagaacttaa gaaattacaa atgctaaagc aaacccaaac aggcaggaat taatcttcat 18600

cgaatttggg tgtttctttc taaaagtcct ttatacttaa atgtcttaag acatacatag 18660

attttatttt actaatttta attatataga caataaatga atattcttac tgattacttt 18720

ttctgactgt ctaatctttc tgatctatcc tggatggcca taacacttat ctctctgaac 18780

tttgggcttt taatatagga aagaaaagca ataatccatt tttcatggta tctcatatga 18840

taaacaaata aaatgcttaa aaatgagcag gtgaagcaat ttatcttgaa ccaacaagca 18900

tcgaagcaat aatgagactg cccgcagcct acctgacttc tgagtcagga tttataagcc 18960

ttgttactga gacacaaacc tgggcctttc aatgctataa cctttcttga agctcctccc 19020

taccaccttt agccataagg aaacatggaa tgggtcagat ccctggatgc aagccaggtc 19080

tggaaccata ggcagtaagg agagaagaaa atgtgggctc tgcaactggc tccgagggag 19140

caggagagga tcaaccccat actctgaatc taagagaaga ctggtgtcca tactctgaat 19200

gggaagaatg atgggattac ccatagggct tgttttaggg agaaacctgt tctccaaact 19260

cttggccttg agatacctgg tccttattcc ttggactttg gcaatgtctg accctcacat 19320

tcaagttctg aggaagggcc actgccttca tactgtggat ctgtagcaaa ttccccctga 19380

aaacccagag ctgtatctta attggttaaa aaaaattata ttatctcaac gactgttctt 19440

ctctgagtag ccaagctcag cttggttcaa gctacaagca gctgagctgc tttttgtcta 19500

gtcattgttc ttttatttca gtggatcaaa tacgttcttt ccaaacctag gatcttgtct 19560

tcctaggcta tatattttgt cccaggaagt cttaatctgg ggtccacaga acactagggg 19620

gctggtgaag tttatagaaa aaaaatctgt atttttactt acatgtaact gaaatttagc 19680

atttcttct actttgaatg caaaggacaa actagaatga catcatcagt acctattgca 19740

tagttataaa gagaaaccac agatattttc atactacacc ataggtattg cagatctttt 19800

tgtttttgtt tttgtttgag atggagtttc gctcttattg cccaggctgg agtgcagtgg 19860

catgatttcg gctcactgca acctcccctt cctgcattca agcaattctc ctgccttggc 19920

ctcctgagta gctggggatt acaggcacct gccaccatgc cagtctaatt tttgtatttt 19980

tagtagagat ggggtttcgc catgttggcc aggctggtct tgaactcctg acctcagatg 20040

atctgcccgc cttggcctcc tgaagtgctg ggattatagg tgtgagccac cacgcctggc 20100

ccattgcaga tatttttaat tcacatttat ctgcatcact acttggatct taaggtagct 20160

gtagacccaa tcctagatct aatgctttca taaagaagca aatataataa atactatacc 20220

acaaatgtaa tgtttgatgt ctgataatga tatttcagtg taattaaact tagcactcct 20280

atgtatatta tttgatgcaa taaaaacata ttttttttagc acttacagtc tgccaaactg 20340

gcctgtgaca caaaaaaagt ttaggaattc ctggttttgt ctgtgttagc caatggttag 20400

aatatatgct cagaaagata ccattggtta atagctaaaa gaaaatggag tagaaattca 20460

gtggcctgga ataataacaa tttgggcagt cattaagtca ggtgaagact tctggaatca 20520

tgggagaaaa gcaagggaga cattcttact tgccacaagt gttttttttt ttttttttt 20580

ttatcacaaa cataagaaaa tataataaat aacaaagtca ggttatagaa gagagaaacg 20640

ctcttagtaa acttggaata tggaatcccc aaaggcactt gacttgggag acaggagcca 20700

tactgctaag tgaaaaagac gaagaacctc tagggcctga acatacagga aattgtagga 20760

acagaaattc ctagatctgg tggggcaagg ggagccatag gagaaagaaa tggtagaaat 20820

ggatggagac ggaggcagag gtgggcagat catgaggtca agagatcgag accatcctgg 20880

caaacatggt gaaatcccgt ctctactaaa aataaaaaaa ttagctgggc atggtggcat 20940

gcgcctgtag tcccagctgc tcgggaggct gaggcaggag aatcgtttga acccaggagg 21000

cgaaggttgc agtgagctga gatagtgcca ttgcactcca gtctggcaac agagtgagac 21060

tccgtctcaa aaaaaaaaaa aaaagaaaga aagaaaagaa aaagaaaaaa gaaaaaataa 21120

atggatgtag aacaagccag aaggaggaac tgggctgggg caatgagatt atggtgatgt 21180

aagggacttt tatagaatta acaatgctgg aatttgtgga actctgcttc tattattccc 21240

ccaatcatta cttctgtcac attgatagtt aaataatttc tgtgaattta ttccttgatt 21300

ctaaaatatg aggataatga caatggtatt ataagggcag attaagtgat atagcatgag 21360

caatattctt caggcacatg gatcgaattg aatacactgt aaatcccaac ttccagtttc 21420

agctctacca agtaaagagc tagcaagtca tcaaaatggg gacatacaga aaaaaaaaag 21480

gacactagag gaataatata ccctgactcc tagcctgatt aatatatcga t 21531

<210> 186

<211> 1614

<212> DNA

<213> Homo sapiens

<400> 186

gatctctatt tatttagcaa taatagagaa agcatttaag agaataaagc aatggaaata 60

agaaatttgt aaatttcctt ctgataacta gaaatagagg atccagtttc ttttggttaa 120

cctaaatttt atttcatttt attgttttat tttattttat tttattttat tttgtgtaat 180

cgtagtttca gagtgttaga gctgaaagga agaagtagga gaaacatgca aagtaaaagt 240

ataacacttt ccttactaaa ccgacatggg tttccaggta ggggcaggat tcaggatgac 300

tgacagggcc cttagggaac actgagaccc tacgctgacc tcataaatgc ttgctacctt 360

tgctgtttta attacatctt ttaatagcag gaagcagaac tctgcacttc aaaagttttt 420

cctcacctga ggagttaatt tagtacaagg ggaaaaagta caggggggatg ggagaaaggc 480

gatcacgttg ggaagctata gagaaagaag agtaaatttt agtaaaggag gtttaaacaa 540

acaaaatata aagagaaata ggaacttgaa tcaaggaaat gattttaaaa cgcagtattc 600

ttagtggact agaggaaaaa aataatctga gccaagtaga agacctttc ccctcctacc 660

cctactttct aagtcacaga ggctttttgt tcccccagac actcttgcag attagtccag 720

gcagaaacag ttagatgtcc ccagttaacc tcctatttga caccactgat taccccattg 780

atagtcacac tttgggttgt aagtgacttt ttatttattt gtatttttga ctgcattaag 840

aggtctctag ttttttatct cttgtttccc aaaacctaat aagtaactaa tgcacagagc 900

acattgattt gtatttattc tatttttaga cataatttat tagcatgcat gagcaaatta 960

agaaaaacaa caacaaatga atgcatatat atgtatatgt atgtgtgtat atatacacac 1020

atatatatat atatttttc ttttcttacc agaaggtttt aatccaaata aggagaagat 1080

atgcttagaa ccgaggtaga gttttcatcc attctgtcct gtaagtattt tgcatattct 1140

ggagacgcag gaagagatcc atctacatat cccaaagctg aattatggta gacaaaactc 1200

ttccactttt agtgcatcaa cttcttattt gtgtaataag aaaattggga aaacgatctt 1260

caatatgctt accaagctgt gattccaaat attacgtaaa tacacttgca aaggaggatg 1320

tttttagtag caatttgtac tgatggtatg gggccaagag atatatctta gagggagggc 1380

tgagggtttg aagtccaact cctaagccag tgccagaaga gccaaggaca ggtacggctg 1440

tcatcactta gacctcaccc tgtggagcca caccctaggg ttggccaatc tactcccagg 1500

agcagggagg gcaggagcca gggctgggca taaaagtcag ggcagagcca tctattgctt 1560

acatttgctt ctgacacaac tgtgttcact agcaacctca aacagacacc atgg 1614

<210> 187

<211> 55

<212> DNA

<213> Artificial Sequence

<220><223> CD33 shRNA

<400> 187

tgttcatact tctttcggat ttcaagagaa tccgaaagaa gtatgaactt ttttc 55

<210> 188

<211> 55

<212> DNA

<213> Artificial Sequence

<220><223> CD33 shRNA

<400> 188

tggagagagg aagtaccaaa ttcaagagat ttggtacttc ctctctcctt ttttc 55

<210> 189

<211> 22

<212> DNA

<213> Artificial Sequence

<220><223> mouse RPL10 forward primer

<400> 189

tgaagacatg gttgctgaga ag 22

<210> 190

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> mouse RPL10 reverse primer

<400> 190

gaacgatttg gtagggtata ggag 24

<210> 191

<211> 22

<212> DNA

<213> Artificial Sequence

<220><223> human gamma-globin forward primer

<400> 191

gtggaagatg ctggaggaga aa 22

<210> 192

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> human gamma-globin reverse primer

<400> 192

tgccatgtgc cttgactttg 20

<210> 193

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> mouse beta-major globin forward primer

<400> 193

atgccaaagt gaaggcccat 20

<210> 194

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> mouse beta-major globin reverse primer

<400> 194

cccagcacaa tcacgatcat 20

<210

> 195

<211> 22

<212> DNA

<213> Artificial Sequence

<220><223> human gamma-globin forward primer

<400> 195

gtgcttgaag gggaacaact ac 22

<210> 196

<211> 22

<212> DNA

<213> Artificial Sequence

<220><223> human gamma-globin reverse primer

<400> 196

cctggcctcc agataactac ac 22

<210> 197

<211> 39

<212> DNA

<213> Artificial Sequence

<220><223> EFlalpha p1 forward primer

<400> 197

ccccctcgag gtcgacatgg ctagagactt atcgaaagc 39

<210> 198

<211> 38

<212> DNA

<213> Artificial Sequence

<220><223> EFlalpha p1 reverse primer

<400> 198

attcgatatc aagctccaag atctgcacac tggtattt 38

<210> 199

<211> 39

<212> DNA

<213> Artificial Sequence

<220><223> EF1alpha p2 forward primer

<400> 199

ccccctcgag gtcgacgtac acgacatcac tttcccagt 39

<210> 200

<211> 37

<

212> DNA

<213> Artificial Sequence

<220><223> EF1alpha p2 reverse primer

<400> 200

attcgatatc aagctcacac tggtatttcg gtttttg 37

<210> 201

<211> 40

<212> DNA

<213> Artificial Sequence

<220><223> 3'HS1 p1 forward primer

<400> 201

ccccctcgag gtcgacctac actctcagtc agcctatgga 40

<210> 202

<211> 39

<212> DNA

<213> Artificial Sequence

<220><223> 3'HS1 p1 reverse primer

<400> 202

attcgatatc aagcttaatc ccaaaaggct gatagtctc 39

<210> 203

<211> 40

<212> DNA

<213> Artificial Sequence

<220><223> 3'HS1 p2 forward primer

<400> 203

ccccctcgag gtcgacacat ctctcacttt ctcatcacca 40

<210> 204

<211> 39

<212> DNA

<213> Artificial Sequence

<220><223> 3'HS1 p2 reverse primer

<400> 204

attcgatatc aagctaagta actgggatta caggagcac 39

<210> 205

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> CD46F primer

<400> 205

aaagggcaaa taccttaagg ggtg 24

<210> 206

<211> 25

<212> DNA

<213> Artificial Sequence

<220><223> CD46R primer

<400> 206

agcacttcga cctaaaaata gagat 25

<210> 207

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> sgRNA

<400> 207

ggggccacta gggacaggat 20

<210> 208

<211> 33

<212> DNA

<213> Artificial Sequence

<220><223> AAVS1 forward primer

<400> 208

atctcacagg taaaactgac gcacggagga aca 33

<210> 209

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> AAVS1 reverse primer

<400> 209

cgggtcacct ctcactcctt tcatttgggc 30

<210> 210

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> AAVS1 forward primer

<400> 210

cggttaatgt ggctctggtt 20

<210> 211

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> AAVS1 reverse primer

<400> 211

cctctctggc tccatcgtaa 20

<210> 212

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> mouse alpha globin forward primer

<400> 212

ctggggaaga caaaagcaac 20

<210> 213

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> mouse alpha globin reverse primer

<400> 213

gccgtggctt acatcaaagt 20

<210> 214

<211> 22

<212> DNA

<213> Artificial Sequence

<220><223> human g-globin forward primer

<400> 214

ggggcaaggt gaatgtggaa ga 22

<210> 215

<211> 22

<212> DNA

<213> Artificial Sequence

<220><223> human g-globin reverse primer

<400> 215

catgatggca gaggcagagg ac 22

<210> 216

<211> 21

<212> DNA

<213> Artificial Sequence

<220><223> human beta-globin forward primer

<400> 216

ctcatggcaa gaaagtgctc g 21

<210> 217

<211>

22

<212> DNA

<213> Artificial Sequence

<220><223> human beta-globin reverse primer

<400> 217

aattctttgc caaagtgatg gg 22

<210> 218

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> GFP forward primer

<400> 218

tcgtgaccac cctgacctac 20

<210> 219

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> GFP reverse primer

<400> 219

ggtcttgtag ttgccgtcgt 20

<210> 220

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> mgmt forward primer

<400> 220

gctgtctggt tgtgagcagg gtct 24

<210> 221

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> mgmt reverse primer

<400> 221

cgggctggtg gaaataggca ttc 23

<210> 222

<211> 22

<212> DNA

<213> Artificial Sequence

<220><223> AAVS1 specific TLA forward primer

<400> 222

ggttgtccag aaaaacggtg at 22

<210> 223

<211> 22

<212> DNA

<213> Artificial Sequence

<220><223> AAVS1 specific TLA reverse primer

<400> 223

cctctcactc ctttcatttg gg 22

<210> 224

<211> 42

<212> DNA

<213> Artificial Sequence

<220><223> EF1a forward primer

<400> 224

aacaaaagct ggttaattaa atcggacggg ggtagtctca ag 42

<210> 225

<211> 41

<212> DNA

<213> Artificial Sequence

<220><223> EF1a reverse primer

<400> 225

tattgtacca tcttaattaa aggggcgagt ccttttgtat g 41

<210> 226

<211> 46

<212> DNA

<213> Artificial Sequence

<220><223> pA forward primer

<400> 226

aacaaaagct ggttaattaa tcattttatg tttcaggttc aggggg 46

<210> 227

<211> 48

<212> DNA

<213> Artificial Sequence

<220><223> pA reverse primer

<400> 227

tattgtacca tcttaattaa atggttacaa ataaagcaat agcatcac 48

<210> 228

<211> 46

<212> DNA

<213> Artificial Sequence

<220><223> HS4 forward primer

<400> 228

aacaaaagct ggttaattaa gtttttgtat tctgtttcgt gaggca 46

<210> 229

<211> 42

<212> DNA

<213> Artificial Sequence

<220><223> HS4 reverse primer

<400> 229

tattgtacca tcttaattaa gcatttgcct aaggtcggac at 42

<210> 230

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> In-Out P1 primer

<400> 230

ccacacccag acctgaccca aacc 24

<210> 231

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> In-Out P2 primer

<400> 231

cgggaaccac acacggcact tacc 24

<210> 232

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> In-Out P3 primer

<400> 232

tctaacgctg ccgtctctct cctg 24

<210> 233

<211> 26

<212> DNA

<213> Artificial Sequence

<220><223> forward primer for genotyping CD46 mice

<400> 233

gccagttcat cttttgactc tattaa 26

<210> 234

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> reverse primer for genotyping CD46 mice

<400> 234

aatcacagca atgacccaaa 20

<210> 235

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> 32P-gamma-ATP labeled probes for miRNA 423-5p

<400> 235

aaagtctcgc tctctgcccc tca 23

<210> 236

<211> 37

<212> DNA

<213> Artificial Sequence

<220><223> 32P-gamma-ATP labeled probes for U6 snRNA

<400> 236

gaatttgcgt gtcatccttg cgcaggggcc atgctaa 37

<210> 237

<211> 25

<212> DNA

<213> Artificial Sequence

<220><223> HBG2-seq primer

<400> 237

cttatttgga aaccaatgct tacta 25

<210> 238

<211> 22

<212> DNA

<213> Artificial Sequence

<220><223> anti-mouse PDL1 forward primer

<400> 238

ggttcagcct ggtgggtctt tg 22

<210> 239

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> anti-mouse PDL1 reverse primer

<400> 239

atctggaggt atgcggtgtt cttc 24

<210> 240

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> mouse PPIA forward primer

<400> 240

gcatacaggt cctggcatct 20

<210> 241

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> mouse PPIA reverse primer

<400> 241

atccagccat tcagtcttgg 20

<210> 242

<211> 20

<212> DNA

<213

> Artificial Sequence

<220><223> forward primer for mouse beta-YAC (gamma-globin promoter)

<400> 242

aaacggtccc tggctaaact 20

<210> 243

<211> 21

<212> DNA

<213> Artificial Sequence

<220><223> reverse primer for mouse beta-YAC (gamma-globin promoter)

<400> 243

gctgaagggt gcttccttttt t 21

<210> 244

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> sgBCL

<400> 244

tttatcacag gctccaggaa 20

<210> 245

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> sgBCL

<400> 245

ttttatcaca ggctccagga 20

<210> 246

<211> 25

<212> DNA

<213> Artificial Sequence

<220><223> HBB primer

<400> 246

atgtcagaag caaatgtgag gagca 25

<210> 247

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223>

forward BCL11A-seq primer

<400> 247

agagagcctt ccgaaagagg 20

<210> 248

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> sgBCL

<400> 248

ctgtgataaa agcaactgtt 20

<210> 249

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> sgBCL

<400> 249

gataaaagca actgttagct 20

<210> 250

<211> 20

<212> DNA

<

213> Artificial Sequence

<220><223> sgHBG

<400> 250

cttgaccaat agccttgaca 20

<210> 251

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> HBB primer

<400> 251

aattctggct tatcggaggc aag 23

<210> 252

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> sgHBG

<400> 252

gctattggtc aaggcaaggc 20

<210> 253

<211

> 20

<212> DNA

<213> Artificial Sequence

<220><223> sgHBG

<400> 253

gtggggaagg ggcccccaag 20

<210> 254

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> sgHBG

<400> 254

ccttccccac actatctcaa 20

<210> 255

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> sgHBG

<400> 255

agatatttgc attgagatag 20

<210> 256

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> sgHBB_STOP

<400> 256

cttgccccac agggcagtaa 20

<210> 257

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> sgHBB_SKIP

<400> 257

agactcaccc tgaagttctc 20

<210> 258

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> sgHBB_CRISPR

<400> 258

cttgtcaagg ctattggtca 20

<210> 259

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> sgNEG

<400> 259

ggtgtcgaaa tgagaagaag 20

<210> 260

<211> 594

<212> DNA

<213> Artificial Sequence

<220><223> TadAN +32aa

<400> 260

tctgaagtcg agttcagcca cgagtattgg atgaggcacg cactgaccct ggcaaaacga 60

gcctgggacg aaagagaagt ccccgtcggc gccgtgctgg tgcacaacaa tcgcgtgatc 120

ggagaaggat ggaacaggcc aatcggacgc cacgacccta ccgcacacgc agagatcatg 180

gcactgaggc agggaggact ggtcatgcag aattatcgcc tgatcgatgc caccctgtat 240

gtgacactgg agccatgcgt gatgtgcgca ggagcaatga tccacagccg gatcggaaga 300

gtggtgttcg gagcacgaga cgccaagacc ggagcagcag gcagcctgat ggatgtgctg 360

caccacccag gcatgaacca ccgggtggag atcacagagg gcatcctggc agacgagtgc 420

gccgccctgc tgagcgattt ctttagaatg aggagacagg agatcaaggc ccagaagaag 480

gcacagagct ccaccgactc tggaggctct agcggaggat cctctggaag cgagacacca 540

ggcacatccg agtccgccac accagagtct tccggcggct cctccggagg atcc 594

<210> 261

<211> 594

<212> DNA

<213> Artificial Sequence

<220><223> TadAN+ +32aa

<400> 261

tctgaggtgg agttttccca cgagtactgg atgagacatg ccctgaccct ggccaagagg 60

gcacgcgatg agagggaggt gcctgtggga gccgtgctgg tgctgaacaa tagagtgatc 120

ggcgagggct ggaacagagc catcggcctg cacgacccaa cagcccatgc cgaaattatg 180

gccctgagac agggcggcct ggtcatgcag aactacagac tgattgacgc caccctgtac 240

gtgacattcg agccttgcgt gatgtgcgcc ggcgccatga tccactctag gatcggccgc 300

gtggtgtttg gcgtgaggaa tgcaaaaact ggcgccgcag gctctctgat ggacgtgctg 360

cactaccccg gcatgaatca cagagtcgaa attaccgagg gaatcctggc agatgaatgt 420

gccgccctgc tgtgctattt ctttcggatg cctagacagg tgtttaatgc tcagaagaag 480

gcccagagct ccaccgatag tggcggatct agcgggggct cctctggctc tgagacccct 540

ggcacaagcg agagcgcaac acctgaaagc agcgggggca gcagcggggg gagc 594

<210> 262

<211> 600

<212> DNA

<213> Artificial Sequence

<220><223> vector

<400> 262

catcatcaat aatatacctt attttggatt gaagccaata tgataatgag ggggtggagt 60

ttgtgacgtg gcgcggggcg tgggaacggg gcgggtgacg tagtagtgtg gcggaagtgt 120

gatgttgcaa gtgtggcgga acacatgtaa gcgacggatg tggcaaaagt gacgtttttg 180

gtgtgcgccg gtgtacagcc ctaggataac ttcgtatagc atacattata cgaagttata 240

ctagtacgcc cgggcgtatc gatacgatat cggtccggac aggaagtgac aattttcgcg 300

cggttttagg cggatgttgt agtaaatttg ggcgtaaccg agtaagattt ggccattttc 360

gcgggaaaac tgaataagag gaagtgaaat ctgaataatt ttgtgttact catagcgcgt 420

aatggatccg cgttaaccgg cggccgcatt ctagacggaa ttcataactt cgtatagcat 480

acattatacg aagttatgct agccgaagct tgagctcgtc gagggatctg ggcgtggtta 540

agggtgggaa agaatatata aggtgggggt cttatgtagt tttgtatctg ttttgcagca 600

<210> 263

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> BCL11A-seq reverse primer

<400> 263

ggcagctaga caggacttgg 20

<210> 264

<211> 370

<212> DNA

<213> Adenovirus serotype 35

<400> 264

catcatcaat aatatacctt atagatggaa tggtgccaat atgtaaatga ggtgatttta 60

aaaagtgtgg gccgtgtggt gattggctgt ggggttaacg gttaaaaggg gcggcgcggc 120

cgtgggaaaa tgacgttatt taaatataac ttcgtatagc atacattata cgaagttatg 180

aggtagtttt gttcaggggc aagtgaaaat tgacccatta cgcgcgaaaa ctgaatgagg 240

aagtgttttt ctgaataatg tggtatttat ggcagggtgg agtatttgac cggatccagg 300

tagactttgc tgattttcgt ggaggtttat aacttcgtat agcatacatt atacgaagtt 360

atatttaaat 370

<210> 265

<211> 198

<212> PRT

<213> Artificial Sequence

<220><223> TadA

<400> 265

Ser Glu Val Glu Phe Ser His Glu Tyr Trp Met Arg His Ala Leu Thr
1 5 10 15

Leu Ala Lys Arg Ala Trp Asp Glu Arg Glu Val Pro Val Gly Ala Val
20 25 30

Leu Val His Asn Asn Arg Val Ile Gly Glu Gly Trp Asn Arg Pro Ile
35 40 45

Gly Arg His Asp Pro Thr Ala His Ala Glu Ile Met Ala Leu Arg Gln
50 55 60

Gly Gly Leu Val Met Gln Asn Tyr Arg Leu Ile Asp Ala Thr Leu Tyr
65 70 75 80

Val Thr Leu Glu Pro Cys Val Met Cys Ala Gly Ala Met Ile His Ser
85 90 95

Arg Ile Gly Arg Val Val Phe Gly Ala Arg Asp Ala Lys Thr Gly Ala
100 105 110

Ala Gly Ser Leu Met Asp Val Leu His His Pro Gly Met Asn His Arg
115 120 125

Val Glu Ile Thr Glu Gly Ile Leu Ala Asp Glu Cys Ala Ala Leu Leu
130 135 140

Ser Asp Phe Phe Arg Met Arg Arg Gln Glu Ile Lys Ala Gln Lys Lys
145 150 155 160

Ala Gln Ser Ser Thr Asp Ser Gly Gly Ser Ser Gly Gly Ser Ser Gly
165 170 175

Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Ser Gly
180 185 190

Gly Ser Ser Gly Gly Ser
195

<210> 266

<211> 198

<212> PRT

<213> Artificial Sequence

<220><223> mutant TadA

<400> 266

Ser Glu Val Glu Phe Ser His Glu Tyr Trp Met Arg His Ala Leu Thr
1 5 10 15

Leu Ala Lys Arg Ala Arg Asp Glu Arg Glu Val Pro Val Gly Ala Val
20 25 30

Leu Val Leu Asn Asn Arg Val Ile Gly Glu Gly Trp Asn Arg Ala Ile
35 40 45

Gly Leu His Asp Pro Thr Ala His Ala Glu Ile Met Ala Leu Arg Gln
50 55 60

Gly Gly Leu Val Met Gln Asn Tyr Arg Leu Ile Asp Ala Thr Leu Tyr
65 70 75 80

Val Thr Phe Glu Pro Cys Val Met Cys Ala Gly Ala Met Ile His Ser
85 90 95

Arg Ile Gly Arg Val Val Phe Gly Val Arg Asn Ala Lys Thr Gly Ala
100 105 110

Ala Gly Ser Leu Met Asp Val Leu His Tyr Pro Gly Met Asn His Arg
115 120 125

Val Glu Ile Thr Glu Gly Ile Leu Ala Asp Glu Cys Ala Ala Leu Leu
130 135 140

Cys Tyr Phe Phe Arg Met Pro Arg Gln Val Phe Asn Ala Gln Lys Lys
145 150 155 160

Ala Gln Ser Ser Thr Asp Ser Gly Gly Ser Ser Gly Gly Ser Ser Gly
165 170 175

Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Ser Gly
180 185 190

Gly Ser Ser Gly Gly Ser
195

<210> 267

<211> 594

<212> DNA

<213> Artificial Sequence

<220><223> TadA

<400> 267

tctgaagtcg agtttagcca cgagtattgg atgaggcacg cactgaccct ggcaaagcga 60

gcatgggatg aaagagaagt ccccgtgggc gccgtgctgg tgcacaacaa tagagtgatc 120

ggagagggat ggaacaggcc aatcggccgc cacgacccta ccgcacacgc agagatcatg 180

gcactgaggc agggaggcct ggtcatgcag aattaccgcc tgatcgatgc caccctgtat 240

gtgacactgg agccatgcgt gatgtgcgca ggagcaatga tccacagcag gatcggaaga 300

gtggtgttcg gagcacggga cgccaagacc ggcgcagcag gctccctgat ggatgtgctg 360

caccaccccg gcatgaacca ccgggtggag atcacagagg gaatcctggc agacgagtgc 420

gccgccctgc tgagcgattt ctttagaatg cggagacagg agatcaaggc ccagaagaag 480

gcacagagct ccaccgactc tggaggatct agcggaggat cctctggaag cgagacacca 540

ggcacaagcg agtccgccac accagagagc tccggcggct cctccggagg atcc 594

<210> 268

<211> 594

<212> DNA

<213> Artificial Sequence

<220><223> mutant TadA

<400> 268

tctgaggtgg agttttccca cgagtactgg atgagacatg ccctgaccct ggccaagagg 60

gcacgcgatg agagggaggt gcctgtggga gccgtgctgg tgctgaacaa tagagtgatc 120

ggcgagggct ggaacagagc catcggcctg cacgacccaa cagcccatgc cgaaattatg 180

gccctgagga agggaggcct ggtcatgcag aactacagac tgattgacgc caccctgtac 240

gtgacattcg agccttgcgt gatgtgcgcc ggcgccatga tccactctag gatcggccgc 300

gtggtgtttg gcgtgaggaa cgcaaaaacc ggcgccgcag gctccctgat ggacgtgctg 360

cactaccccg gcatgaatca ccgcgtcgaa attaccgagg gaatcctggc agatgaatgt 420

gccgccctgc tgtgctattt ctttcggatg cctagacagg tgttcaatgc tcagaagaag 480

gcccagagct ccaccgactc cggaggatct agcggcggat cctctggctc tgagacacct 540

ggcacaagcg agagcgcaac acctgaaagc agcgggggca gcagcggggg gtca 594

<210> 269

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> wildtype and edited sequence

<400> 269

tttatcacag gctccaggaa ggg 23

<210> 270

<211> 29

<212> DNA

<213> Artificial Sequence

<220><223> HBG1/2 forward primer

<400> 270

cagggtttct cctccagcat cttccacat 29

<210> 271

<211> 21

<212> DNA

<213> Artificial Sequence

<220><223> HBG1/2 reverse primer

<400> 271

agcagcagta tcctcttggg g 21

<210> 272

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> HBA primer

<400> 272

tcctgcaggg tgaggaagga agg 23

<210> 273

<211> 25

<212> DNA

<213> Artificial Sequence

<220><223> HBA primer

<400> 273

tctatgcaca tcaattagca gaggc 25

<210> 274

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> HBA primer

<400> 274

ccccaaggca ctccagggac atag 24

<210> 275

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> sgBCL11Ae

<400> 275

ttttatcaca ggctccagga agg 23

<210> 276

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> sgHBG

<400> 276

cttgaccaat agccttgaca agg 23

<210> 277

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> sgHBG

<400> 277

gctattggtc aaggcaaggc tgg 23

<210> 278

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> sgHBG

<400> 278

gtggggaagg ggcccccaag agg 23

<210> 279

<211> 444

<212> DNA

<213> Artificial Sequence

<220><223> dual base editing vector

<400> 279

atggtgcatc tgactcctga ggagaagtct gccgttactg ccctgtgggg caaggtgaac 60

gtggatgaag ttggtggtga ggccctgggc aggctgctgg tggtctaccc ttggacccag 120

aggttctttg agtcctttgg ggatctgtcc actcctgatg ctgttatggg caaccctaag 180

gtgaaggctc atggcaagaa agtgctcggt gcctttagtg atggcctggc tcacctggac 240

aacctcaagg gcacctttgc cacactgagt gagctgcact gtgacaagct gcacgtggat 300

cctgagaact tcaggctcct gggcaacgtg ctggtctgtg tgctggccca tcactttggc 360

aaagaattca ccccaccagt gcaggctgcc tatcagaaag tggtggctgg tgtggctaat 420

gccctggccc acaagtatca ctaa 444

<210> 280

<211> 594

<212> DNA

<213> Artificial Sequence

<220><223> TadA

<220><221> misc_feature

<222> (15)..(15)

<223> site for modification

<220><221> misc_feature

<222> (57)..(57)

<223> site for modification

<220><221> misc_feature

<222> (63)..(63)

<223> site for modification

<220><221> misc_feature

<222> (69)..(69)

<223> site for modification

<220><221> misc_feature

<222> (87)..(87)

<223> site for modification

<220><221> misc_feature

<222> (112)..(112)

<223> site for modification

<220><221> misc_feature

<222> (114)..(114)

<223> site for modification

<220><221> misc_feature

<222> (126)..(126)

<223> site for modification

<220><221> misc_feature

<222> (147)..(147)

<223> site for modification

<220><221> misc_feature

<222> (198)..(198)

<223> site for modification

<220><221> misc_feature

<222> (216)..(216)

<223> site for modification

<220><221> misc_feature

<222> (289)..(289)

<223> site for modification

<220

><221> misc_feature

<222> (318)..(318)

<223> site for modification

<220><221> misc_feature

<222> (333)..(333)

<223> site for modification

<220><221> misc_feature

<222> (343)..(343)

<223> site for modification

<220><221> misc_feature

<222> (344)..(344)

<223> site for modification

<220><221> misc_feature

<222> (369)..(369)

<223> site for modification

<220><221> misc_feature

<222> (402)..(402)

<223> site for modification

<220><221> misc_feature

<222> (451)..(451)

<223> site for modification

<220><221> misc_feature

<222> (507)..(507)

<223> site for modification

<220><221> misc_feature

<222> (547)..(547)

<223> site for modification

<220><221> misc_feature

<222> (548)..(548)

<223> site for modification

<220><221> misc_feature

<222> (568)..(568)

<223> site for modification

<220><221> misc_feature

<222> (569)..(569)

<223> site for modification

<220><221> misc_feature

<222> (570)..(570)

<223> site for modification

<400

> 280

tctgaagtcg agtttagcca cgagtattgg atgaggcacg cactgaccct ggcaaagcga 60

gcatgggatg aaagagaagt ccccgtgggc gccgtgctgg tgcacaacaa tagagtgatc 120

ggagagggat ggaacaggcc aatcggccgc cacgacccta ccgcacacgc agagatcatg 180

gcactgaggc agggaggcct ggtcatgcag aattaccgcc tgatcgatgc caccctgtat 240

gtgacactgg agccatgcgt gatgtgcgca ggagcaatga tccacagcag gatcggaaga 300

gtggtgttcg gagcacggga cgccaagacc ggcgcagcag gctccctgat ggatgtgctg 360

caccaccccg gcatgaacca ccgggtggag atcacagagg gaatcctggc agacgagtgc 420

gccgccctgc tgagcgattt ctttagaatg cggagacagg agatcaaggc ccagaagaag 480

gcacagagct ccaccgactc tggaggatct agcggaggat cctctggaag cgagacacca 540

ggcacaagcg agtccgccac accagagagc tccggcggct cctccggagg atcc 594

<210> 281

<211> 594

<212> DNA

<213> Artificial Sequence

<220><223> mutant TadA

<220><221> misc_feature

<222> (321)..(321)

<223> site for modification

<220><221> misc_feature

<222> (330)..(330)

<223> site for modification

<220><221> misc_feature

<

222> (345)..(345)

<223> site for modification

<220><221> misc_feature

<222> (382)..(382)

<223> site for modification

<220><221> misc_feature

<222> (384)..(384)

<223> site for modification

<220><221> misc_feature

<222> (465)..(465)

<223> site for modification

<220><221> misc_feature

<222> (498)..(498)

<223> site for modification

<220><221> misc_feature

<222> (499)..(499)

<223> site for modification

<220><221> misc_feature

<222> (500)..(500)

<223> site for modification

<220><221> misc_feature

<222> (501)..(501)

<223> site for modification

<220><221> misc_feature

<222> (504)..(504)

<223> site for modification

<220><221> misc_feature

<222> (516)..(516)

<223> site for modification

<220><221> misc_feature

<222> (537)..(537)

<223> site for modification

<220><221> misc_feature

<222> (592)..(592)

<223> site for modification

<220><221> misc_feature

<222> (593)..(593)

<223> site for modification

<220><221> misc_feature

<222> (594)..(594)

<223> site for modification

<400> 281

tctgaggtgg agttttccca cgagtactgg atgagacatg ccctgaccct ggccaagagg 60

gcacgcgatg agagggaggt gcctgtggga gccgtgctgg tgctgaacaa tagagtgatc 120

ggcgagggct ggaacagagc catcggcctg cacgacccaa cagcccatgc cgcaattatg 180

gccctgagac agggcggcct ggtcatgcag aactacagac tgattgacgc caccctgtac 240

gtgacattcg agccttgcgt gatgtgcgcc ggcgccatga tccactctag gatcggccgc 300

gtggtgtttg gcgtgaggaa cgcaaaaacc ggcgccgcag gctccctgat ggacgtgctg 360

cactaccccg gcatgaatca ccgcgtcgaa attaccgagg gaatcctggc agatgaatgt 420

gccgccctgc tgtgctattt ctttcggatg cctagacagg tgttcaatgc tcagaagaag 480

gcccagagct ccaccgactc cggaggatct agcggaggct cctctggctc tgagacacct 540

ggcacaagcg agagcgcaac acctgaaagc agcgggggca gcagcggggg gtca 594

<210> 282

<400> 282

000

<210> 283

<400> 283

000

<210> 284

<400> 284

000

<210> 285

<400> 285

000

<210> 286

<211> 32767

<212> DNA

<213> Artificial Sequence

<220><223> WL-PS1

<220><221> misc_feature

<222> (2582)..(2759)

<223> 1-->178

<220><221> misc_feature

<222> (2768)..(2801)

<223> loxP

<220><221> misc_feature

<222> (2808)..(2973)

<223> 179-->344

<220><221> misc_feature

<222> (2974)..(3007)

<223> loxP

<220><221> misc_feature

<222> (3016)..(27338)

<223> 3112-->27435

<220><221> misc_feature

<222> (27393)..(29862)

<223> lambda-1 (Complementary)

<220><221> misc_feature

<222> (30176)..(30390)

<223> BGH polyA

<220><221> misc_feature

<222> (30415)..(31080)

<223> copGFP (Complementary)

<220><221> misc_feature

<222> (31127)..(31779)

<223> CMV (Complementary)

<220><221> misc_feature

<222> (31831)..(33360)

<223> lambda-2

<220><221> misc_feature

<222> (34752)..(35866)

<223> Ad5E4orf6

<220><221> misc_feature

<222> (35864)..(37686)

<223> 32972-->34794

<220><221> misc_feature

<222> (334221)..(34756)

<223> 30544-->31879

<400> 286

taaacttggc gcgccctgag tgatttttct ctggtcccgc cgcatccata ccgccagttg 60

tttaccctca caacgttcca gtaaccgggc atgttcatca tcagtaaccc gtatcgtgag 120

catcctctct cgtttcatcg gtatcattac ccccatgaac agaaatcccc cttacacgga 180

ggcatcagtg accaaacagg aaaaaaccgc ccttaacatg gcccgcttta tcagaagcca 240

gacattaacg cttctggaga aactcaacga gctggacgcg gatgaacagg cagacatctg 300

tgaatcgctt cacgaccacg ctgatgagct ttaccgcagc tgcctcgcgc gtttcggtga 360

tgacggtgaa aacctctgac acatgcagct cccggagacg gtcacagctt gtctgtaagc 420

ggatgccggg agcagacaag cccgtcaggg cgcgtcagcg ggtgttggcg ggtgtcgggg 480

cgcagccatg acccagtcac gtagcgatag cggagtgtat actggcttaa ctatgcggca 540

tcagagcaga ttgtactgag agtgcaccat atgcggtgtg aaataccgca cagatgcgta 600

aggagaaaat accgcatcag gcgctcttcc gcttcctcgc tcactgactc gctgcgctcg 660

gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg cggtaatacg gttatccaca 720

gaatcagggg ataacgcagg aaagaacatg tgagcaaaag gccagcaaaa ggccaggaac 780

cgtaaaaagg ccgcgttgct ggcgtttttc cataggctcc gccccctga cgagcatcac 840

aaaaatcgac gctcaagtca gaggtggcga aacccgacag gactataaag ataccaggcg 900

tttccccctg gaagctccct cgtgcgctct cctgttccga ccctgccgct taccggatac 960

ctgtccgcct ttctcccttc gggaagcgtg gcgctttctc atagctcacg ctgtaggtat 1020

ctcagttcgg tgtaggtcgt tcgctccaag ctgggctgtg tgcacgaacc ccccgttcag 1080

cccgaccgct gcgccttatc cggtaactat cgtcttgagt ccaacccggt aagacacgac 1140

ttatcgccac tggcagcagc cactggtaac aggattagca gagcgaggta tgtaggcggt 1200

gctacagagt tcttgaagtg gtggcctaac tacggctaca ctagaaggac agtatttggt 1260

atctgcgctc tgctgaagcc agttaccttc ggaaaaagag ttggtagctc ttgatccggc 1320

aaacaaacca ccgctggtag cggtggtttt tttgtttgca agcagcagat tacgcgcaga 1380

aaaaaaggat ctcaagaaga tcctttgatc ttttctacgg ggtctgacgc tcagtggaac 1440

gaaaactcac gttaagggat tttggtcatg agattatcaa aaaggatctt cacctagatc 1500

cttttaaatt aaaaatgaag ttttaaatca atctaaagta tatatgagta aacttggtct 1560

gacagttacc aatgcttaat cagtgaggca cctatctcag cgatctgtct atttcgttca 1620

tccatagttg cctgactccc cgtcgtgtag ataactacga tacgggaggg cttaccatct 1680

ggccccagtg ctgcaatgat accgcgagac ccacgctcac cggctccaga tttatcagca 1740

ataaaccagc cagccggaag ggccgagcgc agaagtggtc ctgcaacttt atccgcctcc 1800

atccagtcta ttaattgttg ccgggaagct agagtaagta gttcgccagt taatagtttg 1860

cgcaacgttg ttgccattgc tgcaggcatc gtggtgtcac gctcgtcgtt tggtatggct 1920

tcattcagct ccggttccca acgatcaagg cgagttacat gatcccccat gttgtgcaaa 1980

aaagcggtta gctccttcgg tcctccgatc gttgtcagaa gtaagttggc cgcagtgtta 2040

tcactcatgg ttatggcagc actgcataat tctcttactg tcatgccatc cgtaagatgc 2100

ttttctgtga ctggtgagta ctcaaccaag tcattctgag aatagtgtat gcggcgaccg 2160

agttgctctt gcccggcgtc aacacgggat aataccgcgc cacatagcag aactttaaaa 2220

gtgctcatca ttggaaaacg ttcttcgggg cgaaaactct caaggatctt accgctgttg 2280

agatccagtt cgatgtaacc cactcgtgca cccaactgat cttcagcatc ttttactttc 2340

accagcgttt ctgggtgagc aaaaacagga aggcaaaatg ccgcaaaaaa gggaataagg 2400

gcgacacgga aatgttgaat actcatactc ttccttttc aatattattg aagcatttat 2460

cagggttatt gtctcatgag cggatacata tttgaatgta tttagaaaaa taaacaaata 2520

ggggttccgc gcacatttcc ccgaaaagtg ccacctgtct agctacgata tcctgtttaa 2580

acatcatcaa taatatacct tatagatgga atggtgccaa tatgtaaatg aggtgatttt 2640

aaaaagtgtg ggccgtgtgg tgattggctg tggggttaac ggttaaaagg ggcggcgcgg 2700

ccgtgggaaa atgacgtttt atgggggtgg agttttttg caagttgtcg cgggaaatga 2760

tttaaatata acttcgtata gcatacatta tacgaagtta tggatcctta cgcataaaaa 2820

ggcttctttt ctcacggaac tacttagttt tcccacggta tttaacagga aatgaggtag 2880

ttttgaccgg atgcaagtga aaattgctga ttttcgcgcg aaaactgaat gaggaagtgt 2940

ttttctgaat aatgtggtat ttatggcagg gtgataactt cgtatagcat acattatacg 3000

aagttatatt taaataggaa tgtttatgcc ttaccagtgt aacatgaatc atgtgaaagt 3060

gttgttggaa ccagatgcct tttccagaat gagcctaaca ggaatctttg acatgaacac 3120

gcaaatctgg aagatcctga ggtatgatga tacgagatcg agggtgcgcg catgcgaatg 3180

cggaggcaag catgccaggt tccagccggt gtgtgtagat gtgaccgaag atctcagacc 3240

ggatcatttg gttattgccc gcactggagc agagttcgga tccagtggag aagaaactga 3300

ctaaggtgag tattgggaaa actttggggt gggattttca gatggacaga ttgagtaaaa 3360

atttgttttt tctgtcttgc agctgacatg agtggaaatg cttcttttaa ggggggagtc 3420

ttcagccctt atctgacagg gcgtctccca tcctgggcag gagttcgtca gaatgttatg 3480

ggatctactg tggatggaag acccgttcaa cccgccaatt cttcaacgct gacctatgct 3540

actttaagtt cttcaccttt ggacgcagct gcagccgctg ccgccgcctc tgtcgccgct 3600

aacactgtgc ttggaatggg ttactatgga agcatcgtgg ctaattccac ttcctctaat 3660

aacccttcta cactgactca ggacaagtta cttgtccttt tggcccagct ggaggctttg 3720

acccaacgtc tgggtgaact ttctcagcag gtggccgagt tgcgagtaca aactgagtct 3780

gctgtcggca cggcaaagtc taaataaaaa aaattccaga atcaatgaat aaataaacga 3840

gcttgttgtt gatttaaaat caagtgtttt tatttcattt ttcgcgcacg gtatgccctg 3900

gaccaccgat ctcgatcatt gagaactcgg tggatttttt ccagaatcct atagaggtgg 3960
gattgaatgt ttagatacat gggcattagg ccgtctttgg ggtggagata gctccattga 4020
agggattcat gctccggggt agtgttgtaa atcacccagt cataacaagg tcgcagtgca 4080
tggtgttgca caatatcttt tagaagtagg ctgattgcca cagataagcc cttggtgtag 4140
gtgtttacaa accggttgag ctgggagggg tgcattcgag gtgaaattat gtgcattttg 4200
gattggattt ttaagttggc aatattgccg ccaagatccc gtcttgggtt catgttatga 4260
aggactacca agacggtgta tccggtacat ttaggaaatt tatcgtgcag cttggatgga 4320

aaagcgtgga aaaatttgga gacacccttg tgtcctccga gattttccat gcactcatcc 4380
atgataatag caatggggcc gtgggcagcg gcgcgggcaa acacgttccg tgggtctgac 4440
acatcatagt tatgttcctg agttaaatca tcataagcca ttttaatgaa tttggggcgg 4500
agcgtaccag attggggtat gaatgttcct tcgggccccg gagcatagtt cccctcacag 4560
atttgcattt cccaagcttt cagttctgag ggtggaatca tgtccacctg gggggctatg 4620
aagaacaccg tttcgggggc gggggtgatt agttgggatg atagcaagtt tctgagcaat 4680
tgagatttgc cacatccggt ggggccataa ataattccga ttacaggttg caggtggtag 4740

tttagggaac ggcaactgcc gtcttctcga agcaaggggg ccacctcgtt catcatttcc 4800
cttacatgca tattttcccg caccaaatcc attaggaggc gctctcctcc tagtgataga 4860
agttcttgta gtgaggaaaa gtttttcagc ggttttagac cgtcagccat gggcattttg 4920
gaaagagttt gctgcaaaag ttctagtctg ttccacagtt cagtgatgtg ttctatggca 4980
tctcgatcca gcagacctcc tcgtttcgcg ggtttggacg gctcctggag tagggtatga 5040
gacgatgggc gtccagcgct gccagggttc ggtccttcca gggtctcagt gttcgagtca 5100
gggttgtttc cgtcacagtg aaggggtgtg cgcctgcttg ggcgcttgcc agggtgcgct 5160

tcagactcat tctgctggtg gagaacttct gtcgcttggc gccctgtatg tcggccaagt 5220
agcagtttac catgagttcg tagttgagcg cctcggctgc gtggcctttg gcgcggagct 5280
taccttgga agttttcttg cataccgggc agtataggca tttcagcgca tacagcttgg 5340
gcgcaaggaa aatggattct ggggagtatg catccgcgcc gcaggaggcg caaacagttt 5400
cacattccac cagccaggtt aaatccggtt cattggggtc aaaaacaagt tttccgccat 5460
atttttgat gcgtttctta cctttggtct ccataagttc gtgtcctcgt tgagtgacaa 5520
acaggctgtc cgtatctccg tagactgatt ttacaggcct cttctccagt ggagtgcctc 5580

ggtcttcttc gtacaggaac tctgaccact ctgatacaaa ggcgcgcgtc caggccagca 5640

caaaggaggc tatgtgggag gggtagcgat cgttgtcaac caggggtcc acctttcca 5700

aagtatgcaa acacatgtca ccctcttcaa catccaggaa tgtgattggc ttgtaggtgt 5760

atttcacgtg acctggggtc cccgctgggg gggtataaaa gggggcggtt ctttgctctt 5820

cctcactgtc ttccggatcg ctgtccagga acgtcagctg ttggggtagg tattccctct 5880

cgaaggcggg catgacctct gcactcaggt tgtcagtttc taagaacgag gaggatttga 5940

tattgacagt gccggttgag atgcctttca tgaggttttc gtccatttgg tcagaaaaca 6000

caatttttt attgtcaagt ttggtggcaa atgatccata cagggcgttg gataaaagtt 6060

tggcaatgga tcgcatggtt tggttctttt ccttgtccgc gcgctctttg gcggcgatgt 6120

tgagttggac atactcgcgt gccaggcact tccattcggg gaagatagtt gttaattcat 6180

ctggcacgat tctcacttgc caccctcgat tatgcaaggt aattaaatcc acactggtgg 6240

ccacctcgcc tcgaaggggt tcattggtcc aacagagcct acctcctttc ctagaacaga 6300

aaggggggaag tgggtctagc ataagttcat cgggagggtc tgcatccatg gtaaagattc 6360

ccggaagtaa atccttatca aaatagctga tgggagtggg gtcatctaag gccatttgcc 6420

attctcgagc tgccagtgcg cgctcatatg ggttaagggg actgccccag ggcatgggat 6480

gggtgagagc agaggcatac atgccacaga tgtcatagac gtagatggga tcctcaaaga 6540

tgcctatgta ggttggatag catcgccccc ctctgatact tgctcgcaca tagtcatata 6600

gttcatgtga tggcgctagc agccccggac ccaagttggt gcgattgggt ttttctgttc 6660

tgtagacgat ctggcgaaag atggcgtgag aattggaaga gatggtgggt ctttgaaaaa 6720

tgttgaaatg ggcatgaggt agacctacag agtctctgac aaagtgggca taagattctt 6780

gaagcttggt taccagttcg gcggtgacaa gtacgtctag ggcgcagtag tcaagtgttt 6840

cttgaatgat gtcataacct ggttggtttt tcttttccca cagttcgcgg ttgagaaggt 6900

attcttcgcg atccttccag tactcttcta gcggaaaccc gtctttgtct gcacggtaag 6960

atcctagcat gtagaactga ttaactgcct tgtaagggca gcagcccttc tctacgggta 7020

gagagtatgc ttgagcagct tttcgtagcg aagcgtgagt aagggcaaag gtgtctctga 7080

ccatgacttt gagaaattgg tatttgaagt ccatgtcgtc acaggctccc tgttcccaga 7140

gttggaagtc tacccgtttc ttgtaggcgg ggttgggcaa agcgaaagta acatcattga 7200

agagaatctt accggctctg ggcataaaat tgcgagtgat gcggaaaggc tgtggtactt 7260

ccgctcgatt gttgatcacc tgggcagcta ggacgatttc gtcgaaaccg ttgatgttgt 7320

gtcctacgat gtataattct atgaaacgcg gcgtgcctct gacgtgaggt agcttactga 7380

gctcatcaaa ggttaggtct gtggggtcag ataaggcgta gtgttcgaga gcccattcgt 7440

gcaggtgagg atttgcatgt aggaatgatg accaaagatc taccgccagt gctgtttgta 7500

actggtcccg atactgacga aaatgccggc caattgccat tttttctgga gtgacacagt 7560

agaaggttct ggggtcttgt tgccatcgat cccacttgag tttaatggct agatcgtggg 7620

ccatgttgac gagacgctct tctcctgaga gtttcatgac cagcatgaaa ggaactagtt 7680

gtttgccaaa ggatcccatc caggtgtaag tttccacatc gtaggtcagg aagagtcttt 7740

ctgtgcgagg atgagagccg atcgggaaga actggatttc ctgccaccag ttggaggatt 7800

ggctgttgat gtgatggaag tagaagtttc tgcggcgcgc cgagcattcg tgtttgtgct 7860

tgtacagacg gccgcagtag tcgcagcgtt gcacgggttg tatctcgtga atgagctgta 7920

cctggcttcc cttgacgaga aatttcagtg ggaagccgag gcctggcgat tgtatctcgt 7980

gctcttctat attcgctgta tcggcctgtt catcttctgt ttcgatggtg gtcatgctga 8040

cgagcccccg cgggaggcaa gtccagacct cggcgcggga ggggcggagc tgaaggacga 8100

gagcgcgcag gctggagctg tccagagtcc tgagacgctg cggactcagg ttagtaggta 8160

gggacagaag attaacttgc atgatctttt ccagggcgtg cgggaggttc agatggtact 8220

tgatttccac aggttcgttt gtagagacgt caatggcttg cagggttccg tgtcctttgg 8280

gcgccactac cgtacctttg ttttttcttt tgatcggtgg tggctctctt gcttcttgca 8340

tgctcagaag cggtgacggg gacgcgcgcc gggcggcagc ggttgttccg gacccggggg 8400

catggctggt agtggcacgt cggcgccgcg cacgggcagg ttctggtatt gcgctctgag 8460

aagacttgcg tgcgccacca cgcgtcgatt gacgtcttgt atctgacgtc tctgggtgaa 8520

agctaccggc cccgtgagct tgaacctgaa agagagttca acagaatcaa tttcggtatc 8580

gttaacggca gcttgtctca gtatttcttg tacgtcacca gagttgtcct ggtaggcgat 8640

ctccgccatg aactgctcga tttcttcctc ctgaagatct ccgcgacccg ctctttcgac 8700

ggtggccgcg aggtcattgg agatacggcc catgagttgg gagaatgcat tcatgcccgc 8760

ctcgttccag acgcggctgt aaaccacggc cccctcggag tctcttgcgc gcatcaccac 8820

ctgagcgagg ttaagctcca cgtgtctggt taagaccgca tagttgcata ggcgctgaaa 8880

aaggtagttg agtgtggtgg caatgtgttc ggcgacgaag aaatacatga tccatcgtct 8940

cagcggcatt tcgctaacat cgcccagagc ttccaagcgc tccatggcct cgtagaagtc 9000

cacggcaaaa ttaaaaaact gggagtttcg cgcggacacg gtcaattcct cctcgagaag 9060

acggatgagt tcggctatgg tggcccgtac ttcgcgttcg aaggctcccg ggatctcttc 9120

ttcctcttct atctcttctt ccactaacat ctcttcttcg tcttcaggcg ggggcggagg 9180

gggcacgcgg cgacgtcgac ggcgcacggg caaacggtcg atgaatcgtt caatgacctc 9240

tccgcggcgg cggcgcatgg tttcagtgac ggcgcggccg ttctcgcgcg gtcgcagagt 9300

aaaaacaccg ccgcgcatct ccttaaagtg gtgactggga ggttctccgt ttgggaggga 9360

gagggcgctg attatacatt ttattaattg gcccgtaggg actgcgcgca gagatctgat 9420
cgtgtcaaga tccacgggat ctgaaaacct ttcgacgaaa gcgtctaacc agtcacagtc 9480
acaaggtagg ctgagtacgg cttcttgtgg gcgggggtgg ttatgtgttc ggtctgggtc 9540
ttctgtttct tcttcatctc gggaaggtga gacgatgctg ctggtgatga aattaaagta 9600
ggcagttcta agacggcgga tggtggcgag gagcaccagg tctttgggtc cggcttgctg 9660
gatacgcagg cgattggcca ttccccaagc attatcctga catctagcaa gatctttgta 9720
gtagtcttgc atgagccgtt ctacgggcac ttcttcctca cccgttctgc catgcatacg 9780

tgtgagtcca aatccgcgca ttggttgtac cagtgccaag tcagctacga ctctttcggc 9840
gaggatggct tgctgtactt gggtaagggt ggcttgaaag tcatcaaaat ccacaaagcg 9900
gtggtaagcc cctgtattaa tggtgtaagc acagttggcc atgactgacc agttaactgt 9960
ctggtgacca gggcgcacga gctcggtgta tttaaggcgc gaataggcgc gggtgtcaaa 10020
gatgtaatcg ttgcaggtgc gcaccagata ctggtaccct ataagaaaat gcggcggtgg 10080
ttggcggtag agaggccatc gttctgtagc tggagcgcca ggggcgaggt cttccaacat 10140
aaggcggtga tagccgtaga tgtacctgga catccaggtg attcctgcgg cggtagtaga 10200

agcccgagga aactcgcgta cgcggttcca aatgttgcgt agcggcatga agtagttcat 10260
tgtaggcacg gtttgaccag tgaggcgcgc gcagtcattg atgctctata gacacggaga 10320
aaatgaaagc gttcagcgac tcgactccgt agcctggagg aacgtgaacg ggttgggtcg 10380
cggtgtaccc cggttcgaga cttgtactcg agccggccgg agccgcggct aacgtggtat 10440
tggcactccc gtctcgaccc agcctacaaa aatccaggat acggaatcga gtcgttttgc 10500
tggtttccga atggcaggga agtgagtcct attttttttt ttttttgccg ctcagaatgc 10560
atcccgtgct gcgacagatg cgcccccaac aacagccccc ctcgcagcag cagcagcagc 10620

aaccacaaaa ggctgtccct gcaactactg caactgccgc cgtgagcggt gcgggacagc 10680
ccgcctatga tctggacttg gaagagggcg aaggactggc acgtctaggt gcgccttcgc 10740
ccgagcggca tccgcgagtt caactgaaaa aagattctcg cgaggcgtat gtgccccaac 10800
agaacctatt tagagacaga agcggcgagg agccggagga gatgcgagct tcccgcttta 10860
acgcgggtcg tgagctgcgt cacggtttgg accgaagacg agtgttgcga gacgaggatt 10920
tcgaagttga tgaagtgaca gggatcagtc ctgccagggc acacgtggct gcagccaacc 10980
ttgtatcggc ttacgagcag acagtaaagg aagagcgtaa cttccaaaag tcttttaata 11040

atcatgtgcg aaccctgatt gcccgcgaag aagttaccct tggtttgatg catttgtggg 11100

atttgatgga agctatcatt cagaacccta ctagcaaacc tctgaccgcc cagctgtttc 11160

tggtggtgca acacagcaga gacaatgagg ctttcagaga ggcgctgctg aacatcaccg 11220

aacccgaggg gagatggttg tatgatctta tcaacattct acagagtatc atagtgcagg 11280

agcggagcct gggcctggcc gagaaggtag ctgccatcaa ttactcggtt ttgagcttgg 11340

gaaaatatta cgctcgcaaa atctacaaga ctccatacgt tcccatagac aaggaggtga 11400

agatagatgg gttctacatg cgcatgacgc tcaaggtctt gaccctgagc gatgatcttg 11460

gggtgtatcg caatgacaga atgcatcgcg cggttagcgc cagcaggagg cgcgagttaa 11520

gcgacaggga actgatgcac agtttgcaaa gagctctgac tggagctgga accgagggtg 11580

agaattactt cgacatggga gctgacttgc agtggcagcc taatcgcagg gctctgagcg 11640

ccgcgacggc aggatgtgag cttccttaca tagaagaggc ggatgaaggc gaggaggaag 11700

agggcgagta cttggaagac tgatggcaca acccgtgttt tttgctagat ggaacagcaa 11760

gcaccggatc ccgcaatgcg ggcggcgctg cagagccagc cgtccggcat taactcctcg 11820

gacgattgga cccaggccat gcaacgtatc atggcgttga cgactcgcaa ccccgaagcc 11880

tttagacagc aaccccaggc caaccgtcta tcggccatca tggaagctgt agtgccttcc 11940

cgatctaatc ccactcatga gaaggtcctg gccatcgtga acgcgttggt ggagaacaaa 12000

gctattcgtc cagatgaggc cggactggta tacaacgctc tcttagaacg cgtggctcgc 12060

tacaacagta gcaatgtgca aaccaatttg gaccgtatga taacagatgt acgcgaagcc 12120

gtgtctcagc gcgaaaggtt ccagcgtgat gccaacctgg gttcgctggt ggcgttaaat 12180

gctttcttga gtactcagcc tgctaatgtg ccgcgtggtc aacaggatta tactaacttt 12240

ttaagtgctt tgagactgat ggtatcagaa gtacctcaga gcgaagtgta tcagtccggt 12300

cctgattact tctttcagac tagcagacag ggcttgcaga cggtaaatct gagccaagct 12360

tttaaaaacc ttaaaggttt gtggggagtg catgccccgg taggagaaag agcaaccgtg 12420

tctagcttgt taactccgaa ctcccgcctg ttattactgt tggtagctcc tttcaccgac 12480

agcggtagca tcgaccgtaa ttcctatttg ggttacctac taaacctgta tcgcgaagcc 12540

atagggcaaa gtcaggtgga cgagcagacc tatcaagaaa ttacccaagt cagtcgcgct 12600

ttgggacagg aagacactgg cagtttggaa gccactctga acttcttgct taccaatcgg 12660

tctcaaaaga tccctcctca atatgctctt actgcggagg aggagaggat ccttagatat 12720

gtgcagcaga gcgtgggatt gtttctgatg caagaggggg caactccgac tgcagcactg 12780

gacatgacag cgcgaaatat ggagcccagc atgtatgcca gtaaccgacc tttcattaac 12840

aaactgctgg actacttgca cagagctgcc gctatgaact ctgattattt caccaatgcc 12900

atcttaaacc cgcactggct gcccccacct ggtttctaca cgggcgaata tgacatgccc 12960

gaccctaatg acggatttct gtgggacgac gtggacagcg atgtttttc acctctttct 13020

gatcatcgca cgtggaaaaa ggaaggcggt gatagaatgc attcttctgc atcgctgtcc 13080

ggggtcatgg gtgctaccgc ggctgagccc gagtctgcaa gtccttttcc tagtctaccc 13140

ttttctctac acagtgtacg tagcagcgaa gtgggtagaa taagtcgccc gagtttaatg 13200

ggcgaagagg agtacctaaa cgattccttg ctcagaccgg caagagaaaa aaatttccca 13260

aacaatggaa tagaaagttt ggtggataaa atgagtagat ggaagactta tgctcaggat 13320

cacagagacg agcctgggat catggggact acaagtagag cgagccgtag acgccagcgc 13380

catgacagac agaggggtct tgtgtgggac gatgaggatt cggccgatga tagcagcgtg 13440

ttggacttgg gtgggagagg aaggggcaac ccgtttgctc atttgcgccc tcgcttgggt 13500

ggtatgttgt gaaaaaaaat aaaaaagaaa aactcaccaa ggccatggcg acgagcgtac 13560

gttcgttctt ctttattatc tgtgtctagt ataatgaggc gagtcgtgct aggcggagcg 13620

gtggtgtatc cggagggtcc tcctccttcg tacgagagcg tgatgcagca gcagcaggcg 13680

acggcggtga tgcaatcccc actggaggct ccctttgtgc ctccgcgata cctggcacct 13740

acggagggca gaaacagcat tcgttactcg gaactggcac ctcagtacga taccaccagg 13800

ttgtatctgg tggacaacaa gtcggcggac attgcttctc tgaactatca gaatgaccac 13860

agcaacttct tgaccacggt ggtgcagaac aatgacttta cccctacgga agccagcacc 13920

cagaccatta actttgatga acgatcgcgg tggggcggtc agctaaagac catcatgcat 13980

actaacatgc caaacgtgaa cgagtatatg tttagtaaca agttcaaagc gcgtgtgatg 14040

gtgtccagaa aacctcccga cggtgctgca gttggggata cttatgatca caagcaggat 14100

attttggaat atgagtggtt cgagtttact ttgccagaag gcaacttttc agttactatg 14160

actattgatt tgatgaacaa tgccatcata gataattact tgaaagtggg tagacagaat 14220

ggagtgcttg aaagtgacat tggtgttaag ttcgacacca ggaacttcaa gctgggatgg 14280

gatcccgaaa ccaagttgat catgcctgga gtgtatacgt atgaagcctt ccatcctgac 14340

attgtcttac tgcctggctg cggagtggat tttaccgaga gtcgtttgag caaccttctt 14400

ggtatcagaa aaaaacagcc atttcaagag ggttttaaga ttttgtatga agatttagaa 14460

ggtggtaata ttccggccct cttggatgta gatgcctatg agaacagtaa gaaagaacaa 14520

aaagccaaaa tagaagctgc tacagctgct gcagaagcta aggcaaacat agttgccagc 14580

gactctacaa gggttgctaa cgctggagag gtcagaggag acaattttgc gccaacacct 14640

gttccgactg cagaatcatt attggccgat gtgtctgaag gaacggacgt gaaactcact 14700

attcaacctg tagaaaaaga tagtaagaat agaagctata atgtgttgga agacaaaatc 14760

aacacagcct atcgcagttg gtatctttcg tacaattatg gcgatcccga aaaaggagtg 14820

cgttcctgga cattgctcac cacctcagat gtcacctgcg gagcagagca ggtttactgg 14880
tcgcttccag acatgatgaa ggatcctgtc actttccgct ccactagaca agtcagtaac 14940
taccctgtgg tgggtgcaga gcttatgccc gtcttctcaa agagcttcta caacgaacaa 15000
gctgtgtact cccagcagct ccgccagtcc acctcgctta cgcacgtctt caaccgcttt 15060
cctgagaacc agattttaat ccgtccgccg gcgcccacca ttaccaccgt cagtgaaaac 15120
gttcctgctc tcacagatca cgggaccctg ccgttgcgca gcagtatccg gggagtccaa 15180
cgtgtgaccg ttactgacgc cagacgccgc acctgtccct acgtgtacaa ggcactgggc 15240

atagtcgcac cgcgcgtcct ttcaagccgc actttctaaa aaaaaaatgt ccattcttat 15300
ctcgcccagt aataacaccg gttggggtct gcgcgctcca agcaagatgt acggaggcgc 15360
acgcaaacgt tctacccaac atcccgtgcg tgttcgcgga cattttcgcg ctccatgggg 15420
tgccctcaag ggccgcactc gcgttcgaac caccgtcgat gatgtaatcg atcaggtggt 15480
tgccgacgcc cgtaattata ctcctactgc gcctacatct actgtggatg cagttattga 15540
cagtgtagtg gctgacgctc gcaactatgc tcgacgtaag agccggcgaa ggcgcattgc 15600
cagacgccac cgagctacca ctgccatgcg agccgcaaga gctctgctac gaagagctag 15660

acgcgtgggg cgaagagcca tgcttagggc ggccagacgt gcagcttcgg gcgccagcgc 15720
cggcaggtcc cgcaggcaag cagccgctgt cgcagcggcg actattgccg acatggccca 15780
atcgcgaaga ggcaatgtat actgggtgcg tgacgctgcc accggtcaac gtgtacccgt 15840
gcgcacccgt ccccctcgca cttagaagat actgagcagt ctccgatgtt gtgtcccagc 15900
ggcgaggatg tccaagcgca aatacaagga agaaatgctg caggttatcg cacctgaagt 15960
ctacggccaa ccgttgaagg atgaaaaaaa accccgcaaa atcaagcggg ttaaaaagga 16020
caaaaaagaa gaggaagatg gcgatgatgg gctggcggag tttgtgcgcg agtttgcccc 16080

acggcgacgc gtgcaatggc gtgggcgcaa agttcgacat gtgttgagac ctggaacttc 16140
ggtggtcttt acacccggcg agcgttcaag cgctactttt aagcgttcct atgatgaggt 16200
gtacggggat gatgatattc ttgagcaggc ggctgaccga ttaggcgagt ttgcttatgg 16260
caagcgtagt agaataactt ccaaggatga gacagtgtca ataccttgg atcatggaaa 16320
tcccacccct agtcttaaac cggtcacttt gcagcaagtg ttacccgtaa ctccgcgaac 16380
aggtgttaaa cgcgaaggtg aagatttgta tcccactatg caactgatgg tacccaaacg 16440
ccagaagttg gaggacgttt tggagaaagt aaaagtggat ccagatattc aacctgaggt 16500

taaagtgaga cccattaagc aggtagcgcc tggtctgggg gtacaaactg tagacattaa 16560

gattcccact gaaagtatgg aagtgcaaac tgaacccgca aagcctactg ccacctccac 16620

tgaagtgcaa acggatccat ggatgcccat gcctattaca actgacgccg ccggtcccac 16680

tcgaagatcc cgacgaaagt acggtccagc aagtctgttg atgcccaatt atgttgtaca 16740

cccatctatt attcctactc ctggttaccg aggcactcgc tactatcgca gccgaaacag 16800

tacctcccgc cgtcgccgca agacacctgc aaatcgcagt cgtcgccgta gacgcacaag 16860

caaaccgact cccggcgccc tggtgcggca agtgtaccgc aatggtagtg cggaaccttt 16920

gacactgccg cgtgcgcgtt accatccgag tatcatcact taatcaatgt tgccgctgcc 16980

tccttgcaga tatggccctc acttgtcgcc ttcgcgttcc catcactggt taccgaggaa 17040

gaaactcgcg ccgtagaaga gggatgttgg gacgcggaat gcgacgctac aggcgacggc 17100

gtgctatccg caagcaattg cggggtggtt ttttaccagc cttaattcca attatcgctg 17160

ctgcaattgg cgcgatacca ggcatagctt ccgtggcggt tcaggcctcg caacgacatt 17220

gacattggaa aaaaaacgta taaataaaaa aaaatacaat ggactctgac actcctggtc 17280

ctgtgactat gttttcttag agatggaaga catcaatttt tcatccttgg ctccgcgaca 17340

cggcacgaag ccgtacatgg gcacctggag cgacatcggc acgagccaac tgaacggggg 17400

cgccttcaat tggagcagta tctggagcgg gcttaaaaat tttggctcaa ccataaaaac 17460

atacgggaac aaagcttgga acagcagtac aggacaggcg cttagaaata aacttaaaga 17520

ccagaacttc caacaaaaag tagtcgatgg gatagcttcc ggcatcaatg gagtggtaga 17580

tttggctaac caggctgtgc agaaaaagat aaacagtcgt ttggacccgc cgccagcaac 17640

cccaggtgaa atgcaagtgg aggaagaaat tcctccgcca gaaaaacgag gcgacaagcg 17700

tccgcgtccc gatttggaag agacgctggt gacgcgcgta gatgaaccgc cttcttatga 17760

ggaagcaacg aagcttggaa tgcccaccac tagaccgata gccccaatgg ccaccggggt 17820

gatgaaacct tctcagttgc atcgacccgt caccttggat ttgccccctc cccctgctgc 17880

tactgctgta cccgcttcta agcctgtcgc tgccccgaaa ccagtcgccg tagccaggtc 17940

acgtcccggg ggcgctcctc gtccaaatgc gcactggcaa aatactctga acagcatcgt 18000

gggtctaggc gtgcaaagtg taaaacgccg tcgctgcttt taattaaata tggagtagcg 18060

cttaacttgc ctatctgtgt atatgtgtca ttacacgccg tcacagcagc agaggaaaaa 18120

aggaagaggt cgtgcgtcga cgctgagtta ctttcaagat ggccacccca tcgatgctgc 18180

cccaatgggc atacatgcac atcgccggac aggatgcttc ggagtacctg agtccgggtc 18240

tggtgcagtt cgcccgcgcc acagacacct acttcaatct gggaaataag tttagaaatc 18300

ccaccgtagc gccgacccac gatgtgacca ccgaccgtag ccagcggctc atgttgcgct 18360

tcgtgcccgt tgaccgggag gacaatacat actcttacaa agtgcggtac accctggccg 18420

tgggcgacaa cagagtgctg gatatggcca gcacgttctt tgacattagg ggcgtgttgg 18480

acagaggtcc cagtttcaaa ccctattctg gtacggctta caactctctg gctcctaaag 18540

gcgctccaaa tgcatctcaa tggattgcaa aaggcgtacc aactgcagca gccgcaggca 18600

atggtgaaga agaacatgaa acagaggaga aaactgctac ttacactttt gccaatgctc 18660

ctgtaaaagc cgaggctcaa attacaaaag agggcttacc aataggtttg gagatttcag 18720

ctgaaaacga atctaaaccc atctatgcag ataaacttta tcagccagaa cctcaagtgg 18780

gagatgaaac ttggactgac ctagacggaa aaaccgaaga gtatggaggc agggctctaa 18840

agcctactac taacatgaaa ccctgttacg ggtcctatgc gaagcctact aatttaaaag 18900

gtggtcaggc aaaaccgaaa aactcggaac cgtcgagtga aaaaattgaa tatgatattg 18960

acatggaatt ttttgataac tcatcgcaaa gaacaaactt cagtcctaaa attgtcatgt 19020

atgcagaaaa tgtaggtttg gaaacgccag acactcatgt agtgtacaaa cctggaacag 19080

aagacacaag ttccgaagct aatttgggac aacagtctat gcccaacaga cccaactaca 19140

ttggcttcag agataacttt attggactca tgtactataa cagtactggt aacatggggg 19200

tgctggctgg tcaagcgtct cagttaaatg cagtggttga cttgcaggac agaaacacag 19260

aactttctta ccaactcttg cttgactctc tgggcgacag aaccagatac tttagcatgt 19320

ggaatcaggc tgtggacagt tatgatcctg atgtacgtgt tattgaaaat catggtgtgg 19380

aagatgaact tcccaactat tgttttccac tggacggcat aggtgttcca acaaccagtt 19440

acaaatcaat agttccaaat ggagaagata ataataattg gaaagaacct gaagtaaatg 19500

gaacaagtga gatcggacag ggtaatttgt ttgccatgga aattaacctt caagccaatc 19560

tatggcgaag tttcctttat tccaatgtgg ctctgtatct cccagactcg tacaaataca 19620

ccccgtccaa tgtcactctt ccagaaaaca aaaacaccta cgactacatg aacgggcggg 19680

tggtgccgcc atctctagta gacacctatg tgaacattgg tgccaggtgg tctctggatg 19740

ccatggacaa tgtcaaccca ttcaaccacc accgtaacgc tggcttgcgt taccgatcta 19800

tgcttctggg taacggacgt tatgtgcctt tccacataca agtgcctcaa aaattcttcg 19860

ctgttaaaaa cctgctgctt ctcccaggct cctacactta tgagtggaac tttaggaagg 19920

atgtgaacat ggttctacag agttccctcg gtaacgacct gcgggtagat ggcgccagca 19980

tcagtttcac gagcatcaac ctctatgcta ctttttttccc catggctcac aacaccgctt 20040

ccacccttga agccatgctg cggaatgaca ccaatgatca gtcattcaac gactacctat 20100

ctgcagctaa catgctctac cccattcctg ccaatgcaac caatattccc atttccattc 20160

cttctcgcaa ctgggcggct ttcagaggct ggtcatttac cagactgaaa accaaagaaa 20220

ctccctcttt ggggtctgga tttgacccct actttgtcta ttctggttct attccctacc 20280

tggatggtac cttctacctg aaccacactt ttaagaaggt ttccatcatg tttgactctt 20340
cagtgagctg gcctggaaat gacaggttac tatctcctaa cgaatttgaa ataaagcgca 20400
ctgtggatgg cgaaggctac aacgtagccc aatgcaacat gaccaaagac tggttcttgg 20460
tacagatgct cgccaactac aacatcggct atcagggctt ctacattcca gaaggataca 20520
aagatcgcat gtattcattt ttcagaaact tccagcccat gagcaggcag gtggttgatg 20580
aggtcaatta caaagacttc aaggccgtcg ccatacccta ccaacacaac aactctggct 20640
ttgtgggtta catggctccg accatgcgcc aaggtcaacc ctatcccgct aactatccct 20700

atccactcat tggaacaact gccgtaaata gtgttacgca gaaaaagttc ttgtgtgaca 20760
gaaccatgtg gcgcataccg ttctcgagca acttcatgtc tatgggggcc cttacagact 20820
tgggacagaa tatgctctat gccaactcag ctcatgctct ggacatgacc tttgaggtgg 20880
atcccatgga tgagcccacc ctgctttatc ttctcttcga agttttcgac gtggtcagag 20940
tgcatcagcc acaccgcggc atcatcgagg cagtctacct gcgtacaccg ttctcggccg 21000
gtaacgctac cacgtaagaa gcttcttgct tcttgcaaat agcagctgca accatggcct 21060
gcggatccca aaacggctcc agcgagcaag agctcagagc cattgtccaa gacctgggtt 21120

gcggacccta ttttttggga acctacgata agcgcttccc ggggttcatg gcccccgata 21180
agctcgcctg tgccattgta aatacggccg gacgtgagac ggggggagag cactggttgg 21240
ctttcggttg gaacccacgt tctaacacct gctacctttt tgatcctttt ggattctcgg 21300
atgatcgtct caaacagatt taccagtttg aatatgaggg tctcctgcgc cgcagcgctc 21360
ttgctaccaa ggaccgctgt attacgctgg aaaaatctac ccagaccgtg cagggccccc 21420
gttctgccgc ctgcggactt ttctgctgca tgttccttca cgcctttgtg cactggcctg 21480
accgtcccat ggacggaaac cccaccatga aattgctaac tggagtgcca aacaacatgc 21540

ttcattctcc taaagtccag cccaccctgt gtgacaatca aaaagcactc taccattttc 21600
ttaataccca ttcgccttat tttcgctctc atcgtacaca catcgaaagg gccactgcgt 21660
tcgaccgtat ggatgttcaa taatgactca tgtaaacaac gtgttcaata aacatcactt 21720
tatttttta catgtatcaa ggctctggat tacttattta tttacaagtc gaatgggttc 21780
tgacgagaat cagaatgacc cgcaggcagt gatacgttgc ggaactgata cttgggttgc 21840
cacttgaatt cgggaatcac caacttggga accggtatat cgggcaggat gtcactccac 21900
agctttctgg tcagctgcaa agctccaagc aggtcaggag ccgaaatctt gaaatcacaa 21960

ttaggaccag tgctctgagc gcgagagttg cggtacaccg gattgcagca ctgaaacacc 22020

atcagcgacg gatgtctcac gcttgccagc acggtgggat ctgcaatcat gcccacatcc 22080

agatcttcag cattggcaat gctgaacggg gtcatcttgc aggtctgcct acccatggcg 22140

ggcacccaat taggcttgtg gttgcaatcg cagtgcaggg ggatcagtat catcttggcc 22200

tgatcctgtc tgattcctgg atacacggct ctcatgaaag catcatattg cttgaaagcc 22260

tgctgggctt tactaccctc gggataaaac atcccgcagg acctgctcga aaactggtta 22320

gcctgcacag ccggcatcat tcacacagca gcgggcgtca ttgttggcta tttgcaccac 22380

acttctgccc cagcggtttt gggtgatttt ggttcgctcg ggattctcct ttaaggctcg 22440

ttgtccgttc tcgctggcca catccatctc gataatctgc tccttctgaa tcataatatt 22500

gccatgcagg cacttcagct tgccctcata atcattgcag ccatgaggcc acaacgcaca 22560

gcctgtacat tcccaattat ggtgggcgat ctgagaaaaa gaatgtatca ttccctgcag 22620

aaatcttccc atcatcgtgc tcagtgtctt gtgactagtg aaagttaact ggatgcctcg 22680

gtgctcttcg tttacgtact ggtgacagat gcgcttgtat tgttcgtgtt gctcaggcat 22740

tagtttaaaa caggttctaa gttcgttatc cagcctgtac ttctccatca gcagacacat 22800

cacttccatg cctttctccc aagcagacac caggggcaag ctaatcggat tcttaacagt 22860

gcaggcagca gctcctttag ccagagggtc atctttagcg atcttctcaa tgcttctttt 22920

gccatccttc tcaacgatgc gcacgggcgg gtagctgaaa cccactgcta caagttgcgc 22980

ctcttctctt tcttcttcgc tgtcttgact gatgtcttgc atggggatat gtttggtctt 23040

ccttggcttc tttttggggg gtatcggagg aggaggactg tcgctccgtt ccggagacag 23100

ggaggattgt gacgtttcgc tcaccattac caactgactg tcggtagaag aacctgaccc 23160

cacacggcga caggtgtttt tcttcggggg cagaggtgga ggcgattgcg aagggctgcg 23220

gtccgacctg gaaggcggat gactggcaga accccttccg cgttcggggg tgtgctccct 23280

gtggcggtcg cttaactgat ttccttcgcg gctggccatt gtgttctcct aggcagagaa 23340

acaacagaca tggaaactca gccattgctg tcaacatcgc cacgagtgcc atcacatctc 23400

gtcctcagcg acgaggaaaa ggagcagagc ttaagcattc caccgcccag tcctgccacc 23460

acctctaccc tagaagataa ggaggtcgac gcatctcatg acatgcagaa taaaaaagcg 23520

aaagagtctg agacagacat cgagcaagac ccgggctatg tgacaccggt ggaacacgag 23580

gaagagttga aacgctttct agagagagag gatgaaaact gcccaaaaca gcgagcagat 23640

aactatcacc aagatgctgg aaatagggat cagaacaccg actacctcat agggcttgac 23700

ggggaagacg cgctccttaa acatctagca agacagtcgc tcatagtcaa ggatgcatta 23760

ttggacagaa ctgaagtgcc catcagtgtg gaagagctca gctgcgccta cgagcttaac 23820

ctttttcac ctcgtactcc ccccaaacgt cagccaaacg gcacctgcga gccaaatcct 23880

cgcttaaact tttatccagc ttttgctgtg ccagaagtac tggctaccta tcacatcttt 23940

tttaaaaatc aaaaaattcc agtctcctgc cgcgctaatc gcacccgcgc cgatgcccta 24000

ctcaatctgg gacctggttc acgcttacct gatatagctt ccttggaaga ggttccaaag 24060

atcttcgagg gtctgggcaa taatgagact cgggccgcaa atgctctgca aaagggagaa 24120

aatggcatgg atgagcatca cagcgttctg gtggaattgg aaggcgataa tgccagactc 24180

gcagtactca agcgaagcgt cgaggtcaca cacttcgcat atcccgctgt caacctgccc 24240

cctaaagtca tgacggcggt catggaccag ttactcatta agcgcgcaag tccccctttca 24300

gaagacatgc atgacccaga tgcctgtgat gagggtaaac cagtggtcag tgatgagcag 24360

ctaacccgat ggctgggcac cgactctccc cgggatttgg aagagcgtcg caagcttatg 24420

atggccgtgg tgctggttac cgtagaacta gagtgtctcc gacgtttctt taccgattca 24480

gaaaccttgc gcaaactcga agagaatctg cactacactt ttagacacgg ctttgtgcgg 24540

caggcatgca agatatctaa cgtggaactc accaacctgg tttcctacat gggtattctg 24600

catgagaatc gcctaggaca aagcgtgctg cacagcaccc ttaaggggga agcccgccgt 24660

gattacatcc gcgattgtgt ctatctctac ctgtgccaca cgtggcaaac cggcatgggt 24720

gtatggcagc aatgtttaga agaacagaac ttgaaagagc ttgacaagct cttacagaaa 24780

tctcttaagg ttctgtggac agggttcgac gagcgcaccg tcgcttccga cctggcagac 24840

ctcatcttcc cagagcgtct cagggttact ttgcgaaacg gattgcctga ctttatgagc 24900

cagagcatgc ttaacaattt tcgctctttc atcctggaac gctccggtat cctgcccgcc 24960

acctgctgcg cactgccctc cgactttgtg cctctcacct accgcgagtg ccccccgccg 25020

ctatggagtc actgctacct gttccgtctg gccaactatc tctcctacca ctcggatgtg 25080

atcgaggatg tgagcggaga cggcttgctg gagtgccact gccgctgcaa tctgtgcacg 25140

ccccaccggt ccctagcttg caacccccag ttgatgagcg aaacccagat aataggcacc 25200

tttgaattgc aaggccccag cagccaaggc gatgggtctt ctcctgggca aagtttaaaa 25260

ctgaccccgg gactgtggac ctccgcctac ttgcgcaagt ttgctccgga agattaccac 25320

ccctatgaaa tcaagttcta tgaggaccaa tcacagcctc caaaggccga actttcggct 25380

tgcgtcatca cccaggggggc aattctggcc caattgcaag ccatccaaaa atcccgccaa 25440

gaatttctac tgaaaaaggg taaggggggtc taccttgacc cccagaccgg cgaggaactc 25500

aacacaaggt tccctcagga tgtcccaacg acgagaaaac aagaagttga aggtgcagcc 25560

gccgcccccca gaagatatgg aggaagattg ggacagtcag gcagaggagg cggaggagga 25620

cagtctggag gacagtctgg aggaagacag tttggaggag gaaaacgagg aggcagagga 25680

ggtggaagaa gtaaccgccg acaaacagtt atcctcggct gcggagacaa gcaacagcgc 25740

taccatctcc gctccgagtc gaggaacccg gcggcgtccc agcagtagat gggacgagac 25800
cggacgcttc ccgaacccaa ccagcgcttc caagaccggt aagaaggatc ggcagggata 25860
caagtcctgg cgggggcata agaatgccat catctcctgc ttgcatgagt gcgggggcaa 25920
catatccttc acgcggcgct acttgctatt ccaccatggg gtgaactttc cgcgcaatgt 25980
tttgcattac taccgtcacc tccacagccc ctactatagc cagcaaatcc cgacagtctc 26040
gacagataaa gacagcggcg gcgacctcca acagaaaacc agcagcggca gttagaaaat 26100
acacaacaag tgcagcaaca ggaggattaa agattacagc caacgagcca gcgcaaaccc 26160

gagagttaag aaatcggatc tttccaaccc tgtatgccat cttccagcag agtcggggtc 26220
aagagcagga actgaaaata aaaaaccgat ctctgcgttc gctcaccaga agttgtttgt 26280
atcacaagag cgaagatcaa cttcagcgca ctctcgagga cgccgaggct ctcttcaaca 26340
agtactgcgc gctgactctt aaagagtagg cagcgaccgc gcttattcaa aaaaggcggg 26400
aattacatca tcctcgacat gagtaaagaa attcccacgc cttacatgtg gagttatcaa 26460
ccccaaatgg gattggcagc aggcgcctcc caggactact ccacccgcat gaattggctc 26520
agcgccgggc cttctatgat ttctcgagtt aatgatatac gcgcctaccg aaaccaaata 26580

cttttggaac agtcagctct taccaccacg ccccgccaac accttaatcc cagaaattgg 26640
cccgccgccc tagtgtacca ggaaagtccc gctcccacca ctgtattact tcctcgagac 26700
gcccaggccg aagtccaaat gactaatgca ggtgcgcagt tagctggcgg ctccacccta 26760
tgtcgtcaca ggcctcggca taatataaaa cgcctgatga tcagaggccg aggtatccag 26820
ctcaacgacg agtcggtgag ctctccgctt ggtctacgac cagacggaat ctttcagatt 26880
gccggctgcg ggagatcttc cttcacccct cgtcaggctg ttctgacttt ggaaagttcg 26940
tcttcgcaac cccgctcggg cggaatcggg accgttcaat ttgtagagga gtttactccc 27000

tctgtctact tcaacccctt ctccggatct cctgggcact acccggacga gttcataccg 27060
aacttcgacg cgattagcga gtcagtggac ggctacgatt gatgtctggt gacgcggctg 27120
agctatctcg gctgcgacat ctagaccact gccgccgctt tcgctgcttt gcccgggaac 27180
ttattgagtt catctacttc gaactcccca aggatcaccc tcaaggtccg gcccacggag 27240
tgcggattac tatcgaaggc aaaatagact ctcgcctgca acgaattttc tcccagcggc 27300
ccgtgctgat cgagcgagac cagggaaaca ccacggttag taatcaatta cggggtcatt 27360
agttcatagc ccatatatgg agttgcgatc gctgcgggcc atgtcataca ccgccttcag 27420

agcagccgga cctatctgcc cgttcgtgcc gtcgttgtta atcaccacat ggttattctg 27480

ctcaaacgtc ccggacgcct gcgaccggct gtctgccatg ctgcccggtg taccgacata 27540

accgccggtg gcatagccgc gcatcagccg gtaaagattc cccacgccaa tccggctggt 27600

tgcctccttc gtgaagacaa actcaccacg gtgaacaatc cccgctggct catatttgcc 27660

gccggttccc gtaaatcctc cggttgcaaa atggaatttc gccgcagcgg cctgaatggc 27720

tgtaccgcct gacgcggatg cgccgccacc aacagccccg ccaatggcgc tgccgatact 27780

cccgacaatc cccaccattg cctgcttaag cagaatttct gtcatcatgg acagcacgga 27840

acgggtgaag ctgcgccagt tctgctcact gccggtcagc atcgccgcca tattctgtgc 27900

aataccatca aaggtctgcg tggctgcact ttttacctgc gacatactgt ccgtggcgct 27960

ctcttcccac tcactccagc cggacttcag gcctgccatc cagttcccgc gaagctggtc 28020

ttcagccgcc caggtctttt tctgctctga catgacgtta ttcagcgcca gcggattatc 28080

gccatactgt tccttcaggc gctgttccgt ggcttcccgt tctgcctgcc ggtcagtcag 28140

cccccggctt ttcgcatcaa tggcggcccg ttttgcccgt tgctgctgtg cgaatttatc 28200

cgcctgctgc gccagcgcgt tcaggcgctc ctgatacgta accttgtcgc caagtgcagc 28260

cagctggcgt ttgtactcca gcgtctcatc tttatgcgcc agcagggatt tctcctgtgc 28320

agacagctgg cgacgttgcg ccgcctcctc cagtaccgcg aactgactct ccgccttcca 28380

caaatcccgg cgctgctggc tgattttctc atttgctccg gcatgcttct ccagcgtccg 28440

gagttctgcc tgaagcgtca gcagggcagc atgagcactg tcttcctgac gatcgcccgc 28500

agacaccttc acgctggact gtttcggctt tttcagcgtc gcttcataat ccttttcgc 28560

cgccgccatc agcgtgttgt aatccgcctg caggattttc ccgtctttca gtgccttgtt 28620

cagttcttcc tgacgggcgg tatatttctc cagcggcgtc tgcagccgtt cgtaagcctt 28680

ctgcgcctct tcggtatatt tcagccgtga cgcttcggta tcgctctgct gctgcgcatt 28740

tttgtcctgt tgagtctgct gctcagcctt ctttcgggcg gcttcaagcg caagacgggc 28800

cttttcacga tcatcccagt aacgcgcccg cgcttcatcg ttaacaaaat aatcatcctt 28860

gcgcagattc cagatgtcgt ctgctttctt atacgcagcc tctgccttaa tcagcatctc 28920

ctgcgcggta tcaggacgac caatatccag caccgcatcc cacatggatt tgaatgcccg 28980

cgcagtcctg tctgcccagg tctccagcgt gcccatgttc tctttcaggc ggcgggtctg 29040

gtcatcaaac cctttcgttg cggcctcgtt cgccgcctgc aatgccccgg cttcatcgcc 29100

ggaacgctgc aactgagcaa catacgcaat ctgctccgcc gacacgttat ggaactggcg 29160

agccatcgcc gtcagccccg acgtcgggtc tgtggtcagc ttcccgaagg cttcagcgac 29220

cttgtccacc tccacgccgg atgcagagga gaaacgcgcc acactctggc tgatggacgc 29280

aatctgagcc tcaccgctta cccccgcctt aaccagtgcg ctgagtgact cgctggtctg 29340

gttaaacgtc agccctgccg cctgcccggc tctggacagg accagcatac gatctgccgt 29400

cagtcccgcc tgattgccgg aaaggaccag cgttttgttg aaatcggaca gggttgagtt 29460

gccctgatac caggcatacg ccagcgcacc ggtcgccacc gccagcgagg tggcccccac 29520

catcggcagg gtgatcgcac cggcaagccc cctgaacatg gggatcatcc cgccgaagga 29580

gtccttcacc tgcccccccc gttgcagcag gatcagccac ggactttgcc cgcctgcaag 29640

ctgcgtggcc acgtcggtga actgtgcagg cagcatacgc atggcggctt tatactgccc 29700

gacggaaatc cccgctttct gtgcagccag cgcctgtcgg ctcagcgact gttcaacgac 29760

tgccgctgtt tttttcgcat cactttccgt accagaaaaa tgacgcctga ctctggccat 29820

ctgctcgtca aatctggccg catccagact caaatcaacg acgtcgacta agctctagca 29880

tttgtgaacc atcaccctaa tcaagttttt tggggtcgag gtgccgtaaa gcactaaatc 29940

ggaaccctaa agggagcccc cgatttagag cttgacgggg aaagccggcg aacgtggcga 30000

gaaaggaagg gaagaaagcg aaaggagcgg gcgctagggc gctggcaagt gtagcggtca 30060

cgctgcgcgt aaccaccaca cccgccgcgc ttaatgcgcc gctacagggc gcgtggggat 30120

acccccctaga gccccagctg gttctttccg cctcagaagc catagagccc accgcatccc 30180

cagcatgcct gctattgtct tcccaatcct ccccccttgct gtcctgcccc accccacccc 30240

ccagaataga atgacaccta ctcagacaat gcgatgcaat ttcctcattt tattaggaaa 30300

ggacagtggg agtggcacct tccagggtca aggaaggcac gggggagggg caaacaacag 30360

atggctggca actagaaggc acagtcgagg ctgatcagcg ggtttgctag cttaggcgaa 30420

ggcgatgggg gtcttgaagg cgtgctggta ctccacgatg cccagctcgg tgttgctgtg 30480

cagctcctcc acgcggcgga aggcgaacat ggggcccccg ttctgcagga tgctggggtg 30540

gatggcgctc ttgaagtgca tgtggctgtc caccacgaag ctgtagtagc cgccgtcgcg 30600

caggctgaag gtgcgggcga agctgcccac cagcacgtta tcgcccatgg ggtgcaggtg 30660

ctccacggtg gcgttgctgc ggatgatctt gtcggtgaag atcacgctgt cctcggggaa 30720

gccggtgccc accaccttga agtcgccgat cacgcggccg gcctcgtagc ggtagctgaa 30780

gctcacgtgc agcacgccgc cgtcctcgta cttctcgatg cgggtgttgg tgtagccgcc 30840

gttgttgatg gcgtgcagga aggggttctc gtagccgctg gggtaggtgc cgaagtggta 30900

gaagccgtag cccatcacgt ggctcagcag gtaggggctg aaggtcaggg cgcctttggt 30960

gctcttcatc ttgttggtca tgcggccctg ctcgggggtg ccctctccgc cgcccaccag 31020

ctcgaactcc acgccgttca gggtgccggt gatgcggcac tcgatcttca tggcgggcat 31080

ggtggctagc ctagccagct tgggtctccc tatagtgagt cgtattaatt tcgataagcc 31140

agtaagcagt gggttctcta gttagccaga gagctctgct tatatagacc tcccaccgta 31200

cacgcctacc gcccatttgc gtcaatgggg cggagttgtt acgacatttt ggaaagtccc 31260
gttgattttg gtgccaaaac aaactcccat tgacgtcaat ggggtggaga cttggaaatc 31320
cccgtgagtc aaaccgctat ccacgcccat tgatgtactg ccaaaaccgc atcaccatgg 31380
taatagcgat gactaatacg tagatgtact gccaagtagg aaagtcccat aaggtcatgt 31440
actgggcata atgccaggcg ggccatttac cgtcattgac gtcaataggg ggcgtacttg 31500
gcatatgata cacttgatgt actgccaagt gggcagttta ccgtaaatac tccacccatt 31560
gacgtcaatg gaaagtccct attggcgtta ctatgggaac atacgtcatt attgacgtca 31620

atgggcgggg gtcgttgggc ggtcagccag gcgggccatt taccgtaagt tatgtaacgc 31680
ggaactccat atatgggcta tgaactaatg accccgtaat tgattactat taataactac 31740
aataatcaat gtcaacgcgt atatctggcc cgtacatcgc gaagcagcgc aaaacgccta 31800
accctaagca gattcttcat gcaattaagc ttcgcggtgc ttcttcagta cgctacggca 31860
aatgtcatcg acgtttttat ccggaaactg ctgtctggct ttttttgatt tcagaattag 31920
cctgacgggc aatgctgcga agggcgtttt cctgctgagg tgtcattgaa caagtcccat 31980
gtcggcaagc ataagcacac agaatatgaa gcccgctgcc agaaaaatgc attccgtggt 32040

tgtcatacct ggtttctctc atctgcttct gctttcgcca ccatcatttc cagcttttgt 32100
gaaagggatg cggctaacgt atgaaattct tcgtctgttt ctactggtat tggcacaaac 32160
ctgattccaa tttgagcaag gctatgtgcc atctcgatac tcgttcttaa ctcaacagaa 32220
gatgctttgt gcatacagcc cctcgtttat tatttatctc ctcagccagc cgctgtgctt 32280
tcagtggatt tcggataaca gaaaggccgg gaaataccca gcctcgcttt gtaacggagt 32340
agacgaaagt gattgcgcct acccggatat tatcgtgagg atgcgtcatc gccattgctc 32400
cccaaataca aaaccaattt cagccagtgc ctcgtccatt ttttcgatga actccggcac 32460

gatctcgtca aaactcgcca tgtacttttc atcccgctca atcacgacat aatgcaggcc 32520
ttcacgcttc atacgcgggt catagttggc aaagtaccag gcatttttc gcgtcaccca 32580
catgctgtac tgcacctggg ccatgtaagc tgactttatg gcctcgaaac caccgagccg 32640
gaacttcatg aaatcccggg aggtaaacgg gcatttcagt tcaaggccgt tgccgtcact 32700
gcataaacca tcgggagagc aggcggtacg catactttcg tcgcgataga tgatcgggga 32760
ttcagta 32767

<210> 287

<211> 3030

<212> DNA

<213> Homo sapiens

<400> 287

ccaggctcca ttattgatat agtcatgatc tcctctgttg gggatgaagt aggcaaattt 60

gaggcactaa tttacttctc acattctttt cttgaacaga aagatagaac tggaaattaa 120

tagtagtata taaattcaaa attttagctt taataacatt taatcagaca taaataatta 180

tggtaatgtg aatttcaata ataaaatttt agttctaata taagtgtaac tgtgtaatat 240

tcatactttt tctgaaggct ttactaattt gatatggcat tacttttta ttgctgccaa 300

aactattctt attccactgt gtggtgatga gaaagtgaga gatgttctgg agatggtgat 360

tatagatagc ttccctgaag ccatagtaac cccctggaga aaaattggac ctggagtcta 420

gcagcctagg tatgggtact cgatttctta gaaagccttt acaatttcct ttatcttaaa 480

aataagggta ttgaagtaga attctagaat tttcagagga caacttaaaa tatgtgtaat 540

agttttaatt atttatcctc ataaatttaa ctgttcattt taatatattt aaggatgaat 600

tttttaaaaa gttgatttca taaaaacggg aatagaaaga tggttccata ggctgactga 660

gagtgtagag gagggatggg aagggaaaga agttgatctt cagttagact agaggaataa 720

gttttagtga tctctcacac tgcatagtga acacagttaa taatatatta tgtatttaaa 780

ttaaaaattg ctaaaaaata aatattttat gttctcacca caaaaaaagt tggaaggtga 840

ttcatatgct aattagcttg atagactctc tctacaatgt atatatagat caaacatcac 900

attgtatccc ataacatatt atatatatta tatatttata ttatatatta ttattgtatc 960

cattaatata tgcacttatt atttgccagg caaataaaaa atgtttttaa aatataaatt 1020

tatttgtaac ctccttttac ttttctgctt ggttttcttc tttcattcag tgtttaccag 1080

tttcttatag ttaattttat tttaagctgt ctcacatttt ctgaagaaaa gggaacatat 1140

taaagccaac aaaacaaata cactatcttg catgagatga tttatgtcat ggtacaatca 1200

aatgctataa atcttataaa aacttctcaa atggttagat ggctacagtt gaacagatgg 1260

accatgtcat atattttta taatgcttct aaggtatggc taatttttaa aaaatatttt 1320

agtaatgatg ggaatattat ttatagaaat cttataaaat atataatgaa atatgtaata 1380

aagtctagat aaatgtgtat atacataata tatatttatt acataatata taatatataa 1440

tgtatattta tatattacat gcattatata ttaaatataa tacattttat atattatata 1500

ttaaaatatg taataatatg ttattaaata tatacaataa tctattacat tttatgctta 1560

tataatatat aataaatata tagtatataa taaatataca ctatatattt gtatctatat 1620

atgtttataa agtcattcct ctaattaggt cataaccatt caggtaaact ggaaatttaa 1680

gcctacttca ggtttgtggt aaatagattc tctctgaact agcatattca gaatcattaa 1740

acagtcagtt ctttggacaa gtcttataga atgttcttac ctcttcagcc atcccaagac 1800

tcttgagggc ctgacctcgc ttacactaaa gcagatctgc cttatgcatc actgaagtag 1860

ggagggaaga aagtttgatg aactacttct gacccctagt ggtgtccaga aaagaccatt 1920

aaaggaatga cctttaaagg atggacatac aatttttgt ccaaggcagg acatgtgtgg 1980

gtgtctttca gtaattatgt tctaagaaca gcaaaaactc cactgccttg gcaaatagga 2040

atgttttagt tctatagaat tataaagaag ctgtctttta aacacaatat actttctcta 2100

tgtctttgga acaatgacta ttggtcatta ccctatttta aagtaagcaa gtaatcacac 2160

agggaattat tctgaaaaga cagaaaaaaa aaaaaaacca agagatttct gcatatgtag 2220

gtcagtttta atcagagggc atcagaaaag actcctgaaa gaatgacctg gttattataa 2280

tcacagattt gctttccaag tcaacattcc agacagtgct cagaggggat acgaaaaccc 2340

ttttatttct ccagactcaa attcactgct atttgtcttc tctatttatt ttattatagg 2400

cattgttctg gttgctggga actcagactg agataccata cactgactct cagatagcat 2460

aacacaacat gatgtcttgg aaaactgtaa atctttttgt tttttaaata caggtggagc 2520

atctggcaca cctgacatat tgatcttgtt tttctttaaa tcttcattta tttaccttat 2580

caaaactatg ctctttcatc ctacctttca aaacatattt taaaaaatcc tccaacatgt 2640

attttgctct ggtaatccca aaaggctgat agtctctatg gtggcaacat ggataatact 2700

gttccccatc tagatggtct catttcttct gtatctagtc tgaagaagcc tgaatgaaag 2760

tagattttta agctttgtag ctagtctgaa gcctttgtag tcagtctgaa gaaacctgca 2820

tgaaaataga ttttttttttt cctttgggac agagtcttgc tctgtcgccc agactggagt 2880

gcaatggcgc gatctcggct cactgcaact tccacctccc aggatcaagc aattctcctg 2940

cctcagtctc ccaagtaact gggattacag gagcacactg ccatgcccag ctaattattt 3000

tttgtgtttt agtagagaca gggtttcacc 3030

<210> 288

<211> 29986

<212> DNA

<213> Artificial Sequence

<220><223> HDAd35-T4-Ef1a-mgmt-mCherry

<220><221> misc_feature

<222> (1)..(481)

<223> Ad35 5'end

<220><221> misc_feature

<222> (14126)..(14159)

<223> FRT (Complementary)

<220><221> misc_feature

<222> (14220)..(14463)

<223> pT4 LIR

<220><221> misc_feature

<

222> (14478)..(15812)

<223> Efla promoter

<220><221> misc_feature

<222> (15830)..(16450)

<223> mgmt

<220><221> misc_feature

<222> (16451)..(16522)

<223> 2A

<220><221> misc_feature

<222> (16526)..(17230)

<223> mCherry

<220><221> misc_feature

<222> (17259)..(17380)

<223> SV40 pA

<220><221> misc_feature

<222> (17491)..(17756)

<223> pT4 RIR

<220><221> misc_feature

<222> (17863)..(17896)

<223> FRT (Complementary)

<220><221> misc_feature

<222> (29579)..(29986)

<223> Ad35 3'end

<400> 288

catcatcaat aatatacctt atagatggaa tggtgccaat atgtaaatga ggtgatttta 60

aaaagtgtgg gccgtgtggt gattggctgt ggggttaacg gttaaaaggg gcggcgcggc 120

cgtgggaaaa tgacgtttta tgggggtgga gttttttttgc aagttgtcgc gggaaatgtt 180

acgcataaaa aggcttcttt tctcacggaa ctacttagtt ttcccacggt atttaacagg 240

aaatgaggta gttttgaccg gatgcaagtg aaaattgctg attttcgcgc gaaaactgaa 300

tgaggaagtg tttttctgaa taatgtggta tttatggcag ggtggagtat ttgttcaggg 360

ccaggtagac tttgacccat tacgtggagg tttcgattac cgtgtttttt acctgaattt 420

ccgcgtaccg tgtcaaagtc ttctgttttt acgtaggtgt cagctgatcg ctagggtatt 480

taccggtatt caaggattac atgagcttag aaatgtaatt agcatagtgt gtggcatagt 540

gtagatacca aataaatatg atctctcctt ctactcttga aaatgcaaac acattcttgg 600

tggtcctaaa atagcctgta acatggttta ctcagcagca tttgctattc aaggcagatc 660

tgcctttagt cattggctgc gctcctgaac agctgtgtga aaggctaact tttgtaaacc 720

aaatcaaaat aaaatgcagc aaaaatttgt cactgaaagg aaatcctcag tatatccttt 780

tatgaaatga aagatccctc atccaaactt aacttttta aaagtgcgca tttggagata 840

tagccctttc ttatgaatcc taattcaatt ttggccataa acacacgttg atgttcccca 900

ccccaaagca catagcaaca agagtaggtt ctatattgaa aataatgaca atttaaaaac 960

atgtacttat ttcactgtat gtggacagtg tctatgattg catcatgaag tgtcatataa 1020

ccatgtacgt gtacatgaga gagagataga gagagaagtg gtagggtggt ggtggtagag 1080

gggatggcga tagtaatcat ggtaatggta gaggtgatgg aggtggtaat gacggaggta 1140

agggtggtag tgatgatggt ggtggtggta atggtggtgg atgtggtggt ggcaattggg 1200

atggtgggat ggtggtagcc atggtgatgg tggtaatggt gttgatttaa agggtggtgg 1260

tagtgaaggt gagggtagtg gtggtggagg tggtggtgct ggtagcaata gtgatggtgg 1320

tgatggtgtt gatgagggtg ttgggatcag ggtgagttcc cacagtatat ttcattcttg 1380

ttgtaccact ctgtcaacag caccactgac tgggacagag gaagaaggca cactctgaat 1440

gtgttattaa cagaaacctc aaaacagtct gtctccttgt agtcattcaa aattatcttt 1500

ttcttacctg gaaaactgaa actgaattac cgggaaaaac acaggagatt tttgtttgtt 1560

aatatgctgc caataaagta attttatgtc aaatttaact acaggaaagg gcaaggcatt 1620

ttctaagttc cttagatgtc atgtggctaa aaaaaacaaa aggatggaca gcagttagat 1680

actgtacact tagctgtttg aagccatata ttcagaaagc agatgttggg agttggtgtt 1740

tgaggactga tttcctggag gtattttata taggccaagt tcattgttct aaactctaag 1800

ggcttgactt gagggaggaa aagaggcaag aacatgttta gttttgctga cagcatcaca 1860

tgggcagccc taaggctaga caactttagg gcctgaagct tattctagga aagaagcacc 1920

tacagagtgg cactgggctc ccctccacta tagagatgaa gtcatatgac agtaaagggc 1980

aggcagggct gcctaggggg cccagaactg acacttccat tagaatgagc acaggccagg 2040

gagagaagtg gggaaccaga gagaaggagc tggaattcta gtaggacaaa cggtaagtga 2100

acaacaagaa caagttaaga gtgtgtgcag tattctttca aagactgaaa aaatagtgat 2160

gtgatagaat ggcaggtggc tctgagcagg ccaggagaag gactgggggc agagcatccc 2220

aggcaggagg gcagcaagtg ggaaggccct ggggtggggc ttttggactg ttccagtgac 2280

gggcaggcag ccagtgtgcc tgtcacacaa tgcaccaggg aagtagtcgt gaatttgcag 2340

agggtcttgc aggctatggg aaagggattg gattgtattt tgtttgtagg gaagccatcg 2400

ggggacttaa gcagaggaag gattggcttc atctctttga aaaagttctc tctggatgct 2460

gatgggagga gaaatggaag gaaaagaaac acttttaggg gcaagaactt ttgagaaggg 2520

tggaattggg agtgtggagt tggggccagc tttggcacag gaggggaagc taaacacgtg 2580

gccgcatgag ggcctgtaat tctacctgaa atgggtacca tttgttaggg taaacaaatg 2640

aaccaaatgc ccagtgatac agaccaagtg ttggcaaact tcttctgtga tggcccaggt 2700

agtaaatgtc tcaggcttcg caggccatgt ggtctctgtt gaagctctgt gtagtagaca 2760

atatgttaat gactgggcgt gactgtgtgc taataaaagt ttatttacaa aaacagcccg 2820

tgggctggat ttagctcaca ggctgtagtt tgccaacctc tgacctagag catgaactga 2880

gcatcttctt ggagggaaat aagttctttc caagttgccc tcctcacatt gcagggggcc 2940

atgtaggccc attattcaca gaagagtggg tgggcaacct ttctggagca gaaaaacgta 3000

aagatttctt ccgtagtgca agtaaggtga ccatttctaa accgtgcaag tgatccagca 3060

gtcccaaaag ttgtttcact tctcattgtg cgcccgttct caggtgctcc gaagcttcca 3120

gtcctttgta gggacatgga tgaaattgga aatcatcatt ctcagtaaac tatcgcaaga 3180

acaaaaaacc aaacaccgca tattctcact cataggtggg aattgaacaa tgagatcaca 3240

tggacacggg aaggggaaca tcacattctg gggactgttg tggggtgggg ggaggggtga 3300

gggatagcat tgggagatat acctaatgct agatgacaag ttagtgggtg cagcgcacca 3360

gtgtggcaca tgtatacata tgtaactaac ctgcacaatg tgcacatgta ccctaaaact 3420

taaagtataa taataaaaga aaaaaaaaaa gagaggagag aaacatcatc ccctccagga 3480

taccccttggg ccttgttctt atagtcttgt acattgttga acaatttgca tgggctagtg 3540

gattaaagca caccctccac cctcaggccc tcaagggtct ctatgataat acagtctcac 3600

cttctaccct ttccatcacc atcctaggtg ctatggccaa ccttgaggct gccatgttag 3660

gtctatgcat ttcccacctc caccacataa ctctctgaag gccaggtagt ttcctattca 3720

tcttggtaac cccaaagcct cgtgacaggg ctcagctggc atctgcggat gtgaatgaac 3780

cattggagaa aatggtactc tgcaaataac tctgttattt tcccatttcc tgtgtaaggc 3840
ctagagacaa tgacttttta attgcacccc ttcccctctg tatgacactg gccttctctt 3900
gtgtccagca atgtgggtgg cctagatgat ttctaaggga cttctggcca agatgaacag 3960
cagctgcatc ttactgagca tttactatgt gccatatact cagccacagc tctaggggca 4020
tagaagcagg agctctcagg gtcagggcag tgagtgagca agcgagcacc tatgccagcc 4080
ctgcctctgg atggggactt gagagggtga tggaagcctg cagcactgga gggaggcaga 4140
caaagacagg cctgtgctga gggggcccgg agcaagagag agggaggcaa tgacagcaga 4200

gacatgcctg cgccttgggt ttgagtgccc agtggtcaaa tccacttccc tgtggctgat 4260
gcttgccttt ctaactttgg aatttagggg ttggagatct ggtgagaagg taggagggag 4320
atgaggagga gaagggaaag gcaggaagga aggggaggga aaggaaaagc aaaaggggag 4380
gaggaaggtt tccaacaaat tattctatat caactgcgga aatcaaaatt tgttgcccaa 4440
atcttagaag ctcatgtccc tcctccccag aagtctggaa tgcagcactc caggggtagc 4500
ttataaccca aatatctatc tgtaaaaaga gaaacattgg gctttcgagc tgtggattct 4560
cagtaaaagc aagaggcctc agcctacaca ggccagccca gagtttgagg aaccccaggc 4620

ccacacccac agggctggcc cctgggtctg catactccct agaaatgtgc acacttctga 4680
gcctcaactc tgtcctggag tctaacagca tccctctcct tcctggggca gttccacctc 4740
cagaaacctg ttaccttggg ccttatgtca aggaaactgt gggaaagagc taggcaggaa 4800
tgcagatgag gccagcatgg gctcctaaaa gtttagaaat aggcagtgtc atgctcccag 4860
gtgcctgcat aaaccagctg aaaaatggag ctcccctcac cagcactctc ccttcaaaca 4920
gactgtgatt tgcaggtcac tggtttacca agccaggcta cccaggcagg acccagatgc 4980
caagcccagt ggtgtcctgc aagctgagca gtgctcagtt cttgcaaaaa aaggtctgtg 5040

tgaaggcaag gcctctgcct ggcttctcac cccagttggg tgtctggaac aggaaggagc 5100
ccttactgca gaaaaaggag gagggagcaa agggagcgaa cagctgcgtg ctccatgggg 5160
aggatcccca aagtagaaag gcgcatacac actgcagccc ttgacccaga atgctcacag 5220
ctacattaca gattcaggtc tcctcagtgt agtggggctg ctgatgagac tgtggcatcc 5280
tcaggggtca ggacacacat tttccatcac tcttctgatg gcaaaaaacc tctgagccaa 5340
tgccaacctc tgatcattaa aaaaaagtgc tcacagcagt gtgtggttta ggatcatgcc 5400
ctgtgtggtt tggaacacgt gcacaaccac accttgttca tcaccatccc agaaaccctg 5460

acgcaggcaa agagcagagt tattaaccct actttactga tgtggatact gaggcccaga 5520

ggctcatgca agttatcaat aagtggcagg gacagttgcc tctagattaa ctagccccta 5580

ggatcacctg ggtcttggaa ggggacccat aaacatgagc tcccctctct tggggccaga 5640

tttgcacctg tgccgcgcct tcagcctgca tgaagtaggg gctgctggca aagactcaaa 5700

gctgtaaatc tgggttttct cttgaggctt ctaagggagc tgtttcgaca actcactctg 5760

ttcccagctg gctgcccctg catagggttt taaagcagcc tagctttctg ccaggcttgg 5820

cagtggacaa cgctggtcag aacatcccag agagctacca gaatgaagta agtttgcttc 5880

tactctttac ctgtttatgg gctgtctctg ccactggaat gaaaggcact gagaacagtg 5940

cctggcctgc agaaggccct ggaaatacct gagctcctaa tctgggaata ggagtaggaa 6000

gagctttgga ggcagggcac ctgagtttga gatctacaac ttcctgcctg tgtgacattg 6060

ggaaagtctc catcctttct gagcctcagt ctccaccctg gggaagtgga aatatcaatc 6120

tctgtgacac agaagcaaat gagcgaatgt gcacaaagta ccttgcacaa gagagacgct 6180

caaacacttg cctccaggtt tcaccgagaa ctacagagta agatagattt gttcccagtg 6240

gaggaagcct gggaataatt tgcccctaga ctatgaattc ctggggctca agatcgagca 6300

cagggccagg cacacagaag ggaccctgga aatgtggcag gaggccagag atagacaggc 6360

ccttagagct catacccatg ccctctgacc tcaagaagaa agaaacctgc tcaaaatctc 6420

acaaagagct tgttccaacc ctgaatcgag tctgaggact ccttcctgag tccagcactt 6480

tttctgcaag aagtatatgc ctccaaagct gatgggcgca aatcttgaac cccgtcacat 6540

aaacacaaag ggaggaggtg actagagctc ctcctactgg atatgtctaa ggtcaccagt 6600

ctaaagaaaa gggatggata gaatgaggcc agtatttttg cagccatcca aatgtccaca 6660

tacgctgtta cactgagggc tcctctctcc cccgtcttca gccctacttg catttagagg 6720

tgagaaagat atgggctgag gggttgtttt tcatcgtatt gtagatggaa agcacactgc 6780

ccttggggcc atccaaatgt ggaccttgat gtagcacccc accttctgga tggccatcct 6840

tctgaaagtc actgaatttc tcagacttta ttctctttat ccataaagaa ggagaataat 6900

aataatcccc ccaccctgcc caaccactga ctggttggga agctcagaag aaatactggg 6960

cacggcatcc cattgtaatc tatagagtga gtcgcttctt aatattaaat ggctgaacac 7020

agaagatgtg caaaaagtac tgtgtcccct tcctcctcca actgaacatt tcatgccctt 7080

tgcaccctca ttttgtctag gagctgcctt atgaagggaa taggtacctg ctccgagctg 7140

gaggaatctt tgccacttat ggtggggtat ggactgagac agagatggca tgtgacatgc 7200

gcactgagtc tcaactccat gcaggctctg gagcactctc aaattggagt actaatgcct 7260

tttaaattct cacactagca atcctttgac ctactgatct agggatctag ggaaagaatc 7320

gtgatcttaa cttcaaaggg aaggacaaaa tgttctgcct cctgttaaaa ctccatacac 7380

taagtgcaga gactggatgc cttattaacc ttgggtagat gcccaaatgt tcaaaaggtc 7440

aaactcttct gttccccaga tcgccagagt cattaaccag tcacactatt aaatgaatga 7500

acagatgctg aaaaggtact tgcattactg agatttctta tggtgatggc ccctgcctga 7560

tatgtattca gcattttgta gttttcaatg tgcattagag tatagtggtg atgacattgg 7620

cctctgagtt tgccacttct tatatctgtg actttggtca aattgcttaa tctctctgag 7680

tctcggtttc ctggagataa taatagcttc ttcttcccag ggttatcatg aggattacag 7740

gagataatgc cccaaaaatg cttagtaaag tgcctagcac ctagtcaatg ctgaattaaa 7800

ggtggttatt cttacttttc gttcatttga actttgttct cagggagggc aaaggataga 7860

caaagcccca tagctagtga ggagtagctg caagactaga acccaggtgt tctgagccct 7920

agtcttaggc caagaacaac tgttacgtga gatgcacgtt ttccttcaag ggagctcaca 7980

attatttcca tgtaaattca aggactgcta aaagagaact ctcctctggg actgatatca 8040

ttttatttca agattgattt gaaacatgtt ttttgtttgt ttgtttgttt tctaggaaag 8100

aacaagagaa ccagttaagc tgaatgcctg aagcaaatcc ctgttagcga tgttttcagg 8160

atgagggaga gtggtgcaag aaacgtgctt ccagatgcac atggtttcct gggactaggg 8220

ttcagggtgt catccctggg tgttattaag tgtcagaagg agagcaaaca agggaaacat 8280

ctgagatcca gctaaggcta caccctggaa atgcaagccc agctcttgca aaggacctcc 8340

tttggccact caccttccag gccttacaat aacttgtttg gactgcaggt ttcttggtgg 8400

actcacaggc cattctgctt ttatttggtc aacctcagtt cacaagcacc cagatgctga 8460

gatcctcagc atgtgcagca gagtttcata ttagcactgg gtacctttct gaggctacag 8520

ggataccgta cagcagcacc tgtcacgtcc agccaaagga gtgggctctc tcaatgtcat 8580

ccaatgctgt ttcaactgtg aagaagacca tctgagagag ttgcttttgg aggctgaggc 8640

aaatttttaa aattctttgt tctcctcaac tggggtgaat tcttggtctt ctaggacagc 8700

ttgaagtttt agaaagagtc aagccactca gaaccaacag agaactcttt cagagaacaa 8760

ggtgtggcat agaggaggca gagggctgat cttgatcaaa tccaaagtgt gactctaaag 8820

caatgaatgt gaatttttgg caaagcttac aaagggctct aaaggccatc tgcaaagaga 8880

agccaagcct gatcgatgaa tcactagtgc ggccggatat cgatcggcac gctgttgatt 8940

ttctcatagt aaggaacagt gggccctttc agtcccactt ctgtagtctg tggtactaca 9000

aatggtgagc ccatgatgtt gccattcata gggttattct ccagcagtaa tgactggcca 9060

gccactccca tagccgcggg gctaggattt attgtcaatg gagggacctg cagttctgca 9120

caagcagtac taggatgagc acctgggccc attgcaaggg tgacatcttc aaggcaaggc 9180

ctcttaattt tattagggta gcccccatca gccatgtctg gaaactggaa gtggtcttct 9240

tcttgtctcc tcttaacagt tccctgtgaa tggaagagaa gagaggagga gaagagagga 9300
gaggagaagg gaagagaggt gacacacaca cacacacaca cacacacaca cacagagaga 9360
gagagagaga gacagagaga aagagagaga gagagagaga ggaattttta taaaggtttg 9420
gcacattaaa gctaatgaac aggaaatgtg catgataaaa cagacctctc agtttaaaga 9480
cttatagttg tgaaaactat aaaatacagc ctgtctttgg aaccatagtg cttatttatt 9540
cattattatg tttcatctaa actgtctaat tacatttcaa ataaggcatt atgttgtctg 9600
tatactaaaa cgggatagaa cgttattcaa agggtaatct gcccacttca aggagagttc 9660

aacaaaacta tgcagaagtc actaaatgaa ccatgctgcc aaaggcaggc attggagaga 9720
aaactagaag tagctaaata gttttaattc tttcctgtct acagacacat agattttaac 9780
gaaggaatac catagtatag aattgaactt ttaggctgcc ttctagtctt ggttaaatgc 9840
atcaggctgc agtggtaaaa ttgaatacaa cagagccctt acaggaaaga agtagatctg 9900
gatgtgtttt cttggggagc tgtttaaaat actgtttttg ggaaagcaca agtttcagaa 9960
cagtcattgt aggcatcgta ttcattgttc catttatttt tacacacaca cacacacaca 10020
cacacacaca cactctcaca cattgctatg tgtacacaaa aataatttgg aagaacctat 10080

acccaacaat ttggagtggt catttatttg ggatgactgg caattccctt tctattctct 10140
tcatttctgc ttgtttgtct ttaacgagaa cgactcataa tccaaaaatt taaaaaagta 10200
taaagttatc taaataagaa attttcctct gaagatgcat cctcaggttg gggagatatt 10260
aaacaatgag aaaaggcccc aatctgggat ctgaaccttg ggggagctgc ccatcattta 10320
tagaagcaca gcctttggga acaaagcaaa gtcactagca atgtgagact tcctactctt 10380
catggcttca tacagtcatc catcgctgtt gtgttaatga ccatgacctg tatgttagca 10440
ggtaaatggg aaaggaagtg ggggcaaagg agtatgtgca ggaatgatca aaataaggaa 10500

aggaagagag ggatctggaa atcacctgaa tgccgatagg tgaacaggta gaattctttt 10560
aaagcttccc ccacccggta ccccccaaat aaccccttc cagctttgga agtttcacta 10620
ggacatacag tgctcatcct ctgatgtcac cttaagtttg gctcttctgg tttgatgagc 10680
ttgtagccca ctaggagctc aaggcatgca tggggccact tgccagcacg atgaggggca 10740
tgactgtcat ggccaagtga acatcaaagc agatccccag ggctgtatgt ctcaggcctt 10800
ggtgcacatc agaatcactt agaaacatcc acattcctgg gccctcccac cacaaactga 10860
cagcttcatc cagggtgtgg cccaggcatc gggagttttt ccaacagctc catggctgat 10920

tctcaacaga aaaccactgg cccagagcaa gggtggaggc agcgtggcat agggctctga 10980

ccttggcctt gccactgaac ctctcagagc cccagtttct ttatgtgtaa aatgagtgta 11040

attatagttc ttttctcatg aaggtgctct gactattaag tgaaacgggg cacattgtat 11100

gacacctaat agctcctcac taactggtac ccggcattat aaagggcagg tatggaaggg 11160

ttctgggagt ccaataccct tcttaaagac agagaggtct ctgagaccca gagaggggca 11220

ggccttaccc agagttgctc agccagaggg caacaaggcc caggtcagat gcagggcccc 11280

tccaccacca ctcagctgcc tccagaccca ctgccttcgc catgttgttg gtaggacact 11340

gcatcgcccc cacagaaggg gcttgccaac ttgagtgaga ggacttgcac acttctttga 11400

cttttctttt gagatgccca caatctgaac aagggcactt caagggacag ctctgtcacc 11460

aaactcatct gaggcctgaa taccatgggt caggcaggaa tgggttggag aggtgtagag 11520

caggcacaat aagagggctg aggcccatgc agtcatcagt gcccactttc ccaggagtct 11580

gactgggcac agcacccata gtgtccctga gctggtccat ggagcagctc actaactgtt 11640

tggcccacag caggtgctca gtaaatggca gttgaacgaa tcaatggaca aaggaacata 11700

aattacccaa cacacaggga gctcagccat ttactcaatc cattatggag taacctacaa 11760

acaagccact gggtcccaaa ctgaaattgt gtctcttcta cattctccca aagaatccaa 11820

taggttaaaa atagaaatgt atgaaataga tcaatcaggg atgattgcat gtggatttga 11880

cataaggatc ccctgcaggg agtctgagct ggcaacagtc aggcccaaag tgctgtccat 11940

gatgtctcga actgcaagac agttttaaca atggcgaagc aatgcagaac caggcaggcc 12000

aaggaggggg tgggggttgg ggaaaggaag ggagggaagg ggctgtgagg ggcaatggtc 12060

tggcatccct gccacgtgag cctctgaaat ttgctggcag cttctatggg ctcccagagc 12120

tttcacttaa ttgttggtct gccactaacc tgctgggagt aaggtgcagg gatggaggag 12180

gcagggcatg accaccagac actaaaggta ccagctgggg ccactggcaa agggaaggag 12240

gctgcacctc tcctacatga gagcccgtat acacacacct tttccagcac tcatcaactg 12300

catcccaagc aaatggtccc tgatcaattc caattctaga aaccaactga ctactcaata 12360

acaaagtaga tcccagcagg ccgccactgc tggagcggat gccacttttg ctatgccaag 12420

tctgtggctg gacagctgct ggcatgtaca ctcactgact ttcataagga tgcctaataa 12480

agggggcagg ctcacctggc ttttctcagg ggtggggttt ggggtgccga tagaggctgc 12540

tgttttggca gagtggcaag ctgcaagcct cttctgagct ttcatttttc aatggacttc 12600

agtgagaatt cactttgtca gaggccatgc agctccatgt tttggatttc atggaatgag 12660

ctttcaacag tgagcctgaa gtgccctggc tgaacagcaa gaacaccagc caaccctaaa 12720

caaggccgag gagaggcggc tgtgtttaca cggaaggctc agccttgctg taatagcgtc 12780

tgccttcacc agacatcagt gaggcgtgga aatctattat ccagttaatt ttgcccctag 12840

ataaagactt gctttcgtgt cttctctttc acagtcccat gatctgttac tcatctcaac 12900

tgcgagaagt tggctgggct ttcccctgtg cccagtgcca cactcgtgcc ttcactgggt 12960

cacctgtgcc tgtggctgat gccgctgagg ttttgcctgc ccagactggg tgtttctgac 13020

taaatcccac agccaccatt ttagatcaag ggcaggagat agctcactgc tccggaatga 13080

cctcccctcc cagaatcctg gtaggggcgg aaggtcccca accaagctcc cagccctttc 13140

taaatgaatc tccctgcttc acccatgtgc ttttctccag tctctgcggt cttgatgaca 13200

gcagggtatt agtcctagct gtcccacagc tcctacttct ttcaggcctc tccctgtgac 13260

aatcagtagc cactggcagg atttcctcag agcatatctc gatttgcttt cagacaatta 13320

gttaaaagga cactggaccc cagacgtccc aactcccagc cagagccctc acaggcccgg 13380

cctttggtgg tgaggaaggg ggagggagtg agtgacagtg ccctggcatc ttttagaaac 13440

gaattccttt ctctccatac ataaatgcct gcagagtccc atttcagaat ccggcagaca 13500

aagccaccaa tgtgatcccc atgaccttat aaacattcat taaaatgcat ttcaaggcat 13560

gtgatggcct ccccaccccc tagataatga gaaaacaaag gtttctcttc tgatagagac 13620

aagttcagct ctgaagtcaa cattatttct ggttctgtct gaacaatgac atatggcaac 13680

tcttcccttt ctatagttct agtccagaat gacaaaaaag gggaaaaatt tcttagagaa 13740

ggtagagatt atacgaatac agtccatgaa atgagcataa ggagaataaa gaatataact 13800

tatccaaaga agtctggcag gctgttataa atgcttgatt ttggacactg tagttggagg 13860

tttaacatgg acaccaataa aaaggtcagc aaagggtatg cactgttcct attgggcaag 13920

aagataggag gtcaaaggta accaggaaag ataaactcag ggagacttat tttccctcca 13980

gagggcactg ggcttgtagg ccctgggcaa aattgtcaaa aaggtgaaaa tcgcctgtgg 14040

tttatttagt ctgctctttc ttcactagtg cctcaccagt tcagttcagg ccaatttgct 14100

agaaggtagc gaacgatcga ccggtgaagt tcctatactt tctagagaat aggaacttcg 14160

gaataggaac ttctacctag atgcatgctc agagcggccc ctagctagcg tttaaaacct 14220

acagttgaag tcggaagttt acatacactt aagttggagt cattaaaact cgtttttcaa 14280

ctactccaca aatttcttgt taacaaacaa tagttttggc aagtcagtta ggacatctac 14340

tttgtgcatg acacaagtca tttttccaac aattgtttac agacagatta tttcacttat 14400

aattcactgt atcacaattc cagtgggtca gaagtgtaca tacacgcgct tgactgtgcc 14460

tttaagcttt taattaagag taattcatac aaaaggactc gcccctgcct tggggaatcc 14520

cagggaccgt cgttaaactc ccactaacgt agaacccaga gatcgctgcg ttcccgcccc 14580

ctcacccgcc cgctctcgtc atcactgagg tggagaagag catgcgtgag gctccggtgc 14640

ccgtcagtgg gcagagcgca catcgcccac agtccccgag aagttggggg gaggggtcgg 14700

caattgaacc ggtgcctaga gaaggtggcg cggggtaaac tgggaaagtg atgtcgtgta 14760
ctggctccgc ctttttcccg agggtggggg agaaccgtat ataagtgcag tagtcgccgt 14820
gaacgttctt tttcgcaacg ggtttgccgc cagaacacag gtaagtgccg tgtgtggttc 14880
ccgcgggcct ggcctcttta cgggttatgg cccttgcgtg ccttgaatta cttccacgcc 14940
cctggctgca gtacgtgatt cttgatcccg agcttcgggt tggaagtggg tgggagagtt 15000
cgaggccttg cgcttaagga gccccttcgc ctcgtgcttg agttgaggcc tggcctgggc 15060
gctggggccg ccgcgtgcga atctggtggc accttcgcgc ctgtctcgct gctttcgata 15120

agtctctagc catttaaaat ttttgatgac ctgctgcgac gctttttttc tggcaagata 15180
gtcttgtaaa tgcgggccaa gatctgcaca ctggtatttc ggtttttggg gccgcgggcg 15240
gcgacggggc ccgtgcgtcc cagcgcacat gttcggcgag gcggggcctg cgagcgcggc 15300
caccgagaat cggacggggg tagtctcaag ctcgccggcc tgctctggtg cctggcctcg 15360
cgccgccgtg tatcgccccg ccctgggcgg caaggctggc ccggtcggca ccagttgcgt 15420
gagcggaaag atggccgctt cccggccctg ctgcagggag ctcaaaatgg aggacgcggc 15480
gctcgggaga gcgggcgggt gagtcaccca cacaaaggaa aagggccttt ccgtcctcag 15540

ccgtcgcttc atgtgactcc acggagtacc gggcgccgtc caggcacctc gattagttct 15600
cgagcttttg gagtacgtcg tctttaggtt ggggggaggg gttttatgcg atggagtttc 15660
cccacactga gtgggtggag actgaagtta ggccagcttg gcacttgatg taattctcct 15720
tggaatttgc ccttttgag tttggatctt ggttcattct caagcctcag acagtggttc 15780
aaagtttttt tcttccattt caggtgtcgt gagaattcga tatcccacca tggacaagga 15840
ttgtgaaatg aaacgcacca cactggacag ccctttgggg aagctggagc tgtctggttg 15900
tgagcagggt ctgcacgaaa taaagctcct gggcaagggg acgtctgcag ctgatgccgt 15960

ggaggtccca gcccccgctg cggttctcgg aggtccggag cccctgatgc agtgcacagc 16020
ctggctgaat gcctatttcc accagcccga ggctatcgaa gagttccccg tgccagcgct 16080
tcaccatccc gttttccagc aagagtcgtt cacgcgtcag gtgttatgga agctgcttaa 16140
ggttgtgaaa ttcggagaag tgatttctta ccagcaattg gccgccctgg ccggcaaccc 16200
caaagccgcg cgagcagtgg gaggcgccat gagaggcaat cctgtcaaga tcctcatccc 16260
gtgccacaga gtggtctgca gcagcggagc cgtgggcaac tactccggag ggctagccgt 16320
gaaggaatgg cttctggccc atgaaggcca ccggttgggg aagccaggct tgggagggag 16380

ctcaggtctg gcaggggcct ggctcaaggg agcgggagct acctcgggct ccccgcctgc 16440

tggccgaaac ctcgaggtga aacagacttt gaattttgac cttctcaagt tggcgggaga 16500

cgtggagtcc aacccagggc ccatggtgag caagggcgag gaggataaca tggccatcat 16560

caaggagttc atgcgcttca aggtgcacat ggagggctcc gtgaacggcc acgagttcga 16620

gatcgagggc gagggcgagg gccgccccta cgagggcacc cagaccgcca agctgaaggt 16680

gaccaagggt ggcccccctgc ccttcgcctg ggacatcctg tcccctcagt tcatgtacgg 16740

ctccaaggcc tacgtgaagc accccgccga catccccgac tacttgaagc tgtccttccc 16800

cgagggcttc aagtgggagc gcgtgatgaa cttcgaggac ggcggcgtgg tgaccgtgac 16860

ccaggactcc tccctgcagg acggcgagtt catctacaag gtgaagctgc gcggcaccaa 16920

cttcccctcc gacggccccg taatgcagaa gaagaccatg ggctgggagg cctcctccga 16980

gcggatgtac cccgaggacg gcgccctgaa gggcgagatc aagcagaggc tgaagctgaa 17040

ggacggcggc cactacgacg ctgaggtcaa gaccacctac aaggccaaga agcccgtgca 17100

gctgcccggc gcctacaacg tcaacatcaa gttggacatc acctcccaca acgaggacta 17160

caccatcgtg gaacagtacg aacgcgccga gggccgccac tccaccggcg gcatggacga 17220

gctgtacaag taggcggccg ctcttcgagc agatatcata agatacattg atgagtttgg 17280

acaaaccaca actagaatgc agtgaaaaaa atgctttatt tgtgaaattt gtgatgctat 17340

tgctttattt gtaaccatta taagctgcaa taaacaagtt acgttaacaa caacaattgc 17400

attcatttta tgtttcaggt tcaggggggag gtgtgggagg ttttttaaag caagtaaaac 17460

ctctacaaat gtggtattaa ttaaagatct ttaaacaatt taaaggcaat gctaccaaat 17520

actaagcgcg tgtatgtaca cttctgaccc actgggaatg tgatgaaaga aataaaagct 17580

gaaatgaatc attctctcta ctattattct gatatttcac attcttaaaa taaagtggtg 17640

atcctaactg accttaagac agggaatctt tactcggatt aaatgtcagg aattgtgaaa 17700

aagtgagttt aaatgtattt ggctaaggtg tatgtaaact tccgacttca actgtagttt 17760

aaaacgggcc cgtagtctag ggccgccagt gtgatggagt tcggcttcag gtacagcaca 17820

ctggcggccg ttactaggta gctagagcct tcagactcta gggaagttcc tatactttct 17880

agagaatagg aacttcggaa taggaacttc acccatggcg atcgctagcc tctaactcct 17940

agaccgtcag aactgctggg cccttcaaga cgggctgctc acacccactc atgttaagcc 18000

tggtgaggcc tgtactctgt tttcacagga agaaatcctc acccagtctt ccccaaacac 18060

attcccaggt tgtgtcatta gtgggataga gatgattatt gtggggagaa gagaaacatc 18120

tggatggatt tggtgaggtt gatctataga ggaagtaggt gctgcctgag gtagctgtaa 18180

tagaagctaa aggtcaaagg agagggccct gtcccaatcc agatgactcc acttctgctg 18240

gacccaggtt cacaagctta atctacattt cacctaaatt tggctaacaa gcccaaaatc 18300

acacaggcaa agggagaagt ggaggcagaa ccgaggttgg aggccaccag ggccaccggg 18360

cagagatcat ttaagcccaa ccttctcact tctccctggg ctctgcctct cttaaaggac 18420

cttgtggtgt gacctcttgt aggtcccttt cacactcggg gcctcagttt ccccactgta 18480

aagtgaatgg gtcccagctt tggtaagctt atgcttacct gatgctttct tcctgggctg 18540

ctcttgtaga gaaaagataa atcttcttcc tccatccacg agggctcctt tccctggggg 18600

tgagagtagg ctgaggagag ccacttgcac acacccttaa agaaagtatt acctgcacca 18660

gctcagtgag aggcacagat cagactgtta cttgaatcaa attatgagcc tccccaaata 18720

tatctatgac atttaaatag gggattactt gaacatagac tgtgggatcc ggtgtggagt 18780

gcgggagact agcaaagtga atcctgagag tagcaggtct gcacctgttg gatcgagaaa 18840

ggcggcctac aattctggtc aaatgagctg tgcttattga catattctat tagagagtac 18900

taccaggtca ccagtcacca gaaaggctgc cagctctcca accacctcca gggaactatc 18960

ctgaatgggg ccttaacaag tctaagagag ggttggtttg ggtcccaagc caatatttgc 19020

tctgctttat gtcagtcata tggaacccaa accaaccctc tcctatgtgc ctcaccagtc 19080

ggtgcaggga tcccaatttc aagtttggtt ttttatggtc aaagtccagc atagattaaa 19140

tgaaggggtg tgatgatggt gttaaaagag aactccagac cagtttaact cttggacaca 19200

catcccatct caccatggtg cttccaacct tccagagatg atgggctcct attttctgat 19260

gacaaagccc tccacaggat tgctgcctgg ccatcaggga gtgcctctgt aactgaggct 19320

gagatcccac tttcagtcct ccagctgtgg cccatccctg ctccgcccac cgggtatggc 19380

ctgtcctagg ctcttaggta tggctgcatt gtgaaatgat ggctacagag ctggcatctc 19440

ctgtagtctg gttcatctag tgcactacct catagttaaa agaaatctgt ttaagccact 19500

gagggtggct cctagtgcca actccaagaa caggaagctt ccctttttg ggaggagggg 19560

cagatggtaa catggatcgt ccaggtcaat gggagcaggg caaccacagt aagtactgga 19620

caacaacaca aaactccatg tgtggcttcc atcgagtccc tctccaattg gtttggtctt 19680

ctccgtccca tgcagcactt tagcaagggg cctggctgaa ggctatgaat tgtgtggagc 19740

ctcctcattg cagtctccaa ccatctgatg ctgggaaaat gtcaccagga tgcagccatg 19800

ccgtgtggcc aatgaaccga gaaaacaccc cttttctaga atgctctaaa gaggcagaat 19860

aatccagagg tgaggaagga aatactccac cagagaccca ggcagttcct acaaaagcca 19920

gactttcctt cacctaggga gtgacaagac cagtggaaaa cactctcaag cagtaacccc 19980

caaatgctct gcaagccagt ggcgtccaga taccgcacaa gcgagtgggc tgtctaatcc 20040

catcatcatg atgtaaatat ctctaggctg ccctgggctg tgcctgaccc tgtcttcagc 20100

tttccacacc tccacctaca gcccatgcac agaaggacca cccaggaatg ctgcaagtgt 20160

ggcacctcca gggccaccca gggagaagga gggcagctat gctggtggct ccaggcccat 20220
ttggcgggtg gtaccttcac accacaaagc ccaaactgag gccccagatt tggctgatga 20280
gggcatattg gacaggggtc acttatgctc ttccccattg ccacctggcc tctggctacc 20340
tggacttggc tacctgtgga tcctctcaca ggtgccacca tcttggctga gtctccagat 20400
gcgaggtccc tgaggcagtg gcgggcttct cgctaatgct gatgggatta ggaatgggat 20460
aggtggggag ggccctggac tgggccctga tgagccaagt gggtttttag aggggctact 20520
ggtacatttc agggacagga catctggtag agctaagctg gggcaataag gagccactgc 20580

taatctgaga gctagaaaca atcagcttct gggtcattat taattagggt agtttgggct 20640
gtgtggaagt cacgtactat atggggtagc cacagctctc tctacagata atctctaaga 20700
cttctgattg ggactgtgtg aatgcagtag caatatctct tcttactgcc aggccctgcc 20760
agtcctgcct ccacgccctg gctggccccc cttatgatct gacccatgcc aggctgccat 20820
agtatgttac ttctgcatta gcactccttg ggacctgcct ctccactgtc cctcagactt 20880
taaagaacta tacaaaccca aggggctctt cccaagagaa ttgatatgac ttgaggtgat 20940
tccatttctg gaagtagtca ctccattttc tgcctcactc tttcagtgct tcacagagca 21000

ggttcgaacg aaggagccat ccaactaacc gtcatgttcg ggcaaccgaa gaagggagtg 21060
gcaggatttc ctttggagac ttctggaatt agacagcagt ttaatgcaag catctaaatt 21120
ctctccctcc cagagtctca ttaaaactac agtaagagtt tgtgttttgt tttgttttta 21180
aagacaaaat cccaccagga tagagagaat aggagaggag ataacagcat cataatttat 21240
gaaactaaaa tgcagataga ccaggattaa ctgactacac agcaccaagg aagctgaatc 21300
acaagacagc agaggagaaa actggaaagg atcgtggtct atacggcaga atcttcccaa 21360
gcctcaggag gaggagctct agatgttccc agatctggga ggtaaagtgg aatgggggga 21420

catggtcagc gtaatggggt tgggctggaa gcaggttaag gagcaggcag atctctgaat 21480
cccctctctg actctgtgtc cccaggcatc tgcctgtccc ccaccctgga agaggtctgg 21540
cttgaccctt tgtctggtga atttcctgct ctgctttcct ggtcctgctg gccggatcag 21600
tggaggccac tcacttcacc ccacagggat gttctgtgtt gccctacacc tgggaactgg 21660
aggtactgga ggcaggctgt ggtgagcttg aaagcaaaac acagagggca gtccaatctc 21720
tttggccata tttcttctgc atatccaata ccatgtccac aactctgcta gtgtcctgat 21780
ggtggtgggc tctacacatt cccgggaagc tgaaggcaga taatgaccag gacaggtcaa 21840

cctctcttct tctgaaagcc ttcatctact aatggcctgg gactcttccc ttaaatgctt 21900

agattgtgtc ttccactaag gttttttgct gttgctgttg tttgtttgtt tgtttgtttg 21960

tttgtttgtt tgttttgaga cggaatctca ctctgtcgcc caggctggag tgtagtggca 22020

caatctcagc tcaccacaac cttcacctcc taggttgaag gggttctcct gcctcagcct 22080

cctgtgtagc taggattaca ggcacatgcc accatgcctg gctaattttt gtatttttgg 22140

tagagacagg atttcgccat gttggccagg ctggtcttga actcctgacc tcaggtgatc 22200

tgcctacctt ggtctcccaa agtgctggga ttacaggtgt gagccaccac acccggccaa 22260

ggtttttgtt tgtttgtttg tttgtttgtt tgttttgtat tgaggcaggg tatcactctg 22320

gtcacccagg ctggagtgca gtagtgcaat cacggctcac tgaaacctcc acctccctgg 22380

cgggctcagg tgatcctgcc acctcagctt cccaggtagc tgggactaca ggcttgtacc 22440

accactccca gctaattttt gcgtttttag tagagacagg gtttccccat gttgcccagg 22500

ttggtctcaa actctgggct caagcgatct gcctgcctca gcctcccaaa gtgctgggat 22560

tacaggtgta agccaccgta cccggccccg ccactaaggt tttgaaaatg aagcaattac 22620

aagtttaagt ctattaataa gtgatgaagc catgtagaaa agcagaataa ttatcttgga 22680

tcaggaaggt cacatgagga tctacttggg ggttgtcaat attctatttc ttgacctgat 22740

cagtgttgac agcaggtttt aatttttac ttctttttgt ttgtttgttt ttgagacgga 22800

gtcttgctct gtctcccagg ctggagtgca gtggtatgat ctcggctcac tgcaacctcc 22860

gcctcctggg ttcaagctgt tctcctgcct cagcctcccc agtagctggg attacaggca 22920

ggcaccacca cgaccagcta atttttgtat ttttagtaga gactgggttt caccatcttg 22980

gccaggctgg tctcgaactt ctgatctcgt gatccgccct ccttggcctc ccaaagtgct 23040

gggattacag gcttgagcca gcgtgcccgg cccatttttt acttccttat taaactgtac 23100

atataggcct tgcacacttt tctgcatcaa tgttatattc cacaataaag ggaaaaggta 23160

tatacacaac ttgataccag taatgtgaaa catatatttc tacatagaaa aaaaaatgac 23220

tgaaatactg cactccaatg tgttcacaca gtagttgttt ctggattatt tatatattaa 23280

atgtttatat attgtattat gccatgaggt ttgtgttttc tctccacttt tctgcatttt 23340

ccaagtttac tacaaagagc acatattact cttataatca gaaagtcata aaatatattt 23400

aaaaagacaa aattgaaact aataaggatc aacacaaaac agatgagcca tctgtggaaa 23460

tccgcacaga atactaccta aagagattgg tgacgtgcat gatctcacta ggatgagcac 23520

aaagcttgcc agagcctagg gtctatttct agggttggct cttggaagcc aggatagttg 23580

ttatctctgg gaagagggag gggcacacaa ggggcttcta aaacattctg aatgttctat 23640

ttctgaacct ggttggtggg tacatgactg ttggttttat tattatatgt tttatatact 23700

cttccgtatg tatggtgtgg attccaaaaa aagatttcct ttagagaaaa ccagaatcac 23760

ataagtagaa aatatggtgc tatgttgaag gaacaactca agtttatata aaatcatcat 23820

catttatagg cttaaaaagt tgctttggaa ttttggtcta actgacttgt cttttctgca 23880

gcaaaccacg ctccttctgg acgtgctcca ggcagagggg attagggtgg gttcaaggct 23940

gcaagtacct agctcagcac actctcttca ggggacttag agtttgtctg gtgttggctc 24000

tctgagctct tgtcaggaat gccgaccctt ccgaggttca ggatttgaag cctgccttcc 24060

caccccagat ttggtccaca cagacactca agtatgtatt tcaactacaa atgacctgta 24120

ctttcctatt actcctctct ttcatggtaa cctttctggt atccttcctt ccctacattt 24180

atgggagggg gacatcattc tctgctctcc tgtcactgaa ggctccacct tctgtcttct 24240

tctgacccat ctggttttcc tggggccacc tcctctcctt accaccctaa cgcttttgta 24300

acttgaggag aaatgagaga tcacctagtc aggtcatcat tctctgtaga tgaagaggcc 24360

caatggtttg ctcaagaatt gccaagcgag ttaaagacag agagtatgag agtcagcaag 24420

acctacagaa agcatctatc tgcactgttt tgcagggact tagcctttgt gtgtggactc 24480

ctggaatgcc acccactaag aaacattgtc tgacaccaac tccccacttg gtaggtgggg 24540

acactgaaac tcatggcagg aaagggcctt gccccaagcc agggcagagt gtcactcatc 24600

actctcaatt ttcagtccag ggcaccttgt tgtgactatc ccaaaggcag ccactttccc 24660

tggtctgaaa gacctgaaga gagaagagaa gagaaggatg gaaggcagag tatgcggctt 24720

tgattcattt cctggtgaaa acagatctat acgagaagca aatttcacga aagggaagag 24780

aagaaagtgt cccatacgtt gctggcctgt ttcaaccttg ctttgattct tgctgaaaag 24840

ggtaccgtgt atttctgagt tcaacatgca gaccagtgtt aggaaagcca ctgcacctcc 24900

actttagcct ccagggctgt gccctgcaaa tggcctgcag ccttggtgcc tcgctctcca 24960

gactgcattt tggaagatgg gacagaggct tatggaagcc cacattagaa cgggggagca 25020

gaatgggtga gatgagggat ccttgatagt gaaccagatg aaggaatggt agccaaatgc 25080

caggcctcct ttgtggcttc aatccaaagg ctctggagcc cttccagggc agaacatcag 25140

gcatgtttac ccccactgtc ctcaacagtg acagaggtgc aatcttgggc agctggccat 25200

tttgaaagca acctccttaa tctcaactgg gaaggctccc tagcaggacc cctgtgttgc 25260

acacctggag gaagctagac taaccagaag ctcagcacgg ttccatctgg gatgcccagg 25320

tctgagacga aaaaggtaac tctcttttct gggtcctggc ccagttgtgt ctctctccac 25380

ctcattctct gagatgcctg tctccccttt tttgtcccat caggaggcaa gagctatcac 25440

tgggccagac tccaccagaa gccaagccag cttgttaccc agcttctcag ggagcaaaga 25500

acagccttgt ttctatctta tccccactgt cccctgcccc tgccccacct cccagccatt 25560

cagcttctgg cttccccaga gctgcctgct tctttgtggt cctccattcc ttgaaaagac 25620

cttctagtca ttagtgtata taaatggcca cttagcccag attacagtga ggtcaacagc 25680
tggggctctg agaattgtca cacactggca caggagagga ggctattctt ccagagaatt 25740
tggagggcac tcccatccac ttacaacaaa aagcccatcc actgtgcttg gcagtaggtg 25800
atctgagaac caatggaacc aggttaatcc tgtggcactg ttgagtgagg agagcagtgg 25860
cgggcactgg aaaatatcag agacaaggca ggagacctga aatctaggct tagctcctca 25920
tatacttggc agctgtatga cctcagacaa ccagtgttac ctctctaagc ctcagtttcc 25980
tcatgcaaaa ggaggggggaa taacaacaga gcccactgct tgggggtgtt gtgaggacag 26040

gatgaaaaaa caaacagaaa tccctcagta caggattcag tgcagtggac agtcttgcaa 26100
ggtctggttc agccctccac ccctaccctc accagtataa agaactctgg cctacaagtc 26160
agatgacctg agttttaatc tcagctttgc cattagccgt gtgaacttga gaaagtccct 26220
ttccttttta catctattgg gatgatcatg catttttgt cctttattct gttaatatag 26280
tgtgttacat tgattgcttt tcatagactg aaccagcctt gtattccagg gataaatctc 26340
acttggtcat ggtgtataat cctttataca aatgttgctg ggttgagttt gctagtattt 26400
tgttgaagat ttttatgtct tgattcataa ggaatattgg tgtaccttcc ccttttatgg 26460

ccacagtttc cctacaatga tgtagtcgaa ctagacaacc tccaatatct ttcagtattc 26520
atgtcctctg attctgtgaa actaagaaaa ttaagaaata gtgattcata ggcacaaggc 26580
aggcaaaact tagactcctt gtagaataat taggaagcca aatattcagt gtgcttattt 26640
ctcaaataac cttagtttct ccagtctgcc ccaactccga ggcctgaata tctctagatg 26700
cttatgatgg caactaaagc ctaaaagcta attcatttta aagttcttcc aaatgcatag 26760
ggttttattt ttccagacct gggttcagat ggggaatttg acaaacaatg gaaaggggga 26820
aaaacaacaa tctaaacact gagtgacaaa gtaacaaaga aatagtctag ctatcagcca 26880

gtcaagccag ccttggcttt gctatccaaa gtagtcagtc taattctacc accagtttct 26940
gttcctgtag ctgtctactg cctgccaggg actctgcctt cccacccaca actaccaatg 27000
gaaggatgtg gtgaccatac cagtggctgc tgacatctcc tgccatggga agcataattg 27060
cctccagcag cctccccctt agatccatca tttttgttgc acttggcctg ggctgtactc 27120
ccggccaatg actgaacatg gtgagcatag taatgcaggc ccatttctgt gaggagcagg 27180
actcctccag taggtgactt tggctcaagg actctctatt ggcctggttg aacttttcct 27240
gaactgtgct actgtctgag actcttctta cccaatcctc tttctcgccc caattgtcac 27300

agaccacctg cattgtggtc tgagtctctc cccaccttct cttgctcttc cctgtttatc 27360

tttcacaggc atttccccca gtacattcct tgaatgtcta acccgatacg ggtgcctgac 27420

ttttggcaga cctaagcaga caaaaaggag tacttggtta cctagctctt ctttctacca 27480

caaacatcga gggaaccctt tttccctcac ccctctgcca cacccccact gccccagtga 27540

acaaccacag agagagctgt ggtataatat taggctggtg caaaagtaat tgcggttttt 27600

gccattactt ttaatggtaa aaaccgcaat tacttttgca cctacctagt atttgtgtcc 27660

ccccaaattc atatgttgaa acctaaccca caatatgatg tcattaggag gcaagacctt 27720

gaggaggtga ttagatgatg gggtggagct ctcctgaatg agattagtgc ccttataaga 27780

agaagcccaa ggaagctacc ttgactcttc catcacatga gaatgcagca agaaggcacc 27840

atctactaat caggaagaga gctctcacca gacactgaat ctgccagtgt cttgatcttg 27900

aagttcccag cctccagaac tatgcataat gcatttccat tgtctctaag ccacccagcc 27960

tatggtattt tgtcatagca gcctgaactg actaagacag tgagccacat gagaagtgcc 28020

ccaacccctc ccttaagcac ttggctcaca gatcagtggg ttcatttctg cctgagtttt 28080

attgttattc tgtagatttc ttgggctaga tatatttttc tgttattttc cttcttcacc 28140

tcagtcatga attggttgtt ttaaaaaaga caatgtaagt catggggaaa ctcctgacaa 28200

ctctactctc ctagggttcc tgataaaagg ggattcagtt gagtcctctg atggtctcta 28260

cctgccaaag tccagcagcc cttagcaaac atgctgctcg tttctgtaga gaaggtgctg 28320

gtgtcccacc atacttctct ctccctcatg aagggcttgc gacccagcaa atgggtggct 28380

tatatgggtc tgtttcaaag gaagagccag ctctgggaag aaaaacgatg agcataagca 28440

taacctacca ctgtgcctgg gaaagcagac aacttttttg atgtgtgaat atctaatgag 28500

aatggaatcc atcaattacc ttaaacttag gcacagtctt caaattcaat atatgtggga 28560

tatactttta gtcagtttgt agacgttatt tgtaataaat aatctggctt ctctaaagaa 28620

attattttaa gtgtttggtt tggtttgatt taatggtaaa attatattta gtggcagaga 28680

attatagcaa tggtgataaa ctatagagtg tcataagttc atatcttatt ctcacatttg 28740

aagctgcctg cagatgcatt caagatgcag ccagaagtca ggagactcag gctgttattt 28800

ggagctcatc attttacagc cttgctggac tcccactttc tcaggggaaa aatgtggtgt 28860

tgacccagat tagctctcca ggccctgctg agttgggcac tctgtaagct ggagggtctt 28920

ctattgtctt cacctaagtg tcaatcaaca acccaaatgg gcatgggggа agagggagct 28980

gggccaatgc ccagggtgcc tggtagagag ataccttggg cactggaagg caccagcttc 29040

ccagagagaa gggggagggc catgaaaaag ttggctgtag atgccaggga cactgggact 29100

ctccagctgt gtgtttgtgt cttctgaaga cttatgtttc attcctttgg agcatgcata 29160

atcatacact gtgggatgtg ttatatagat tgcttgatag ttcaccactg taataaaata 29220

ctgtgactgg aatctgctcc cagtctgcct ttgatagcac ttgtgcaaca cacatttact 29280

gagcatttac agtgatccag gacctgtgtt gtgaaaacat tgatggacaa ggcagatggt 29340

ggagcacgtc agtgaggatt tttaacaaag gctggtaagt gctataaagg aacattgtag 29400

gacactagag aacaaagaac aggagaacct gacttaggct ggggtggggc gttggttaga 29460

ggaggctcct tggaggacat gaggtttaag ctgtgacctg aggatgaata gatgttggcc 29520

aggtgaggta ccggtatttg tcagccttac cagtaaaaaa gaaaacctat taaaaaaaaa 29580

atacacatac aaagcctcat cagccatggc ttaccagaga aagtacagcg ggcacacaaa 29640

ccacaagctc taaagtcact ctccaacctc tccacaatat atatacacaa gccctaaact 29700

gacgtaatgg gactaaagtg taaaaaatcc cgccaaaccc aacacacacc ccgaaactgc 29760

gtcaccaggg aaaagtacag tttcacttcc gcaatcccaa caagcgtcac ttcctctttc 29820

tcacggtacg tcacatccca ttaacttaca acgtcatttt cccacggccg cgccgcccct 29880

tttaaccgtt aaccccacag ccaatcacca cacggcccac acttttttaaa atcacctcat 29940

ttacatattg gcaccattcc atctataagg tatattattg atgatg 29986

<210> 289

<211> 30158

<212> DNA

<213> Artificial Sequence

<220><223> HDAd35-T4-Ef1a-SB100-Flpe

<220><221> misc_feature

<222> (1)..(481)

<223> Ad35 5'end

<220><221> misc_feature

<222> (14103)..(14614)

<223> pgk

<220><221> misc_feature

<222> (14763)..(15785)

<223> SB100

<220><221> misc_feature

<222> (15811)..(16128)

<223> BGH pA

<220><221> misc_feature

<222> (16088)..(16376)

<223> beta globin pA (Complementary)

<220><221> misc_feature

<222> (16488)..(17759)

<223> Flpe (Complementary)

<220><221> misc_feature

<222> (17780)..(18895)

<223> EF1a (Complementary)

<220><221> misc_feature

<222> (29751)..(30158)

<223> Ad35 3'end

<400> 289

catcatcaat aatatacctt atagatggaa tggtgccaat atgtaaatga ggtgatttta 60

aaaagtgtgg gccgtgtggt gattggctgt ggggttaacg gttaaaaggg gcggcgcggc 120

cgtgggaaaa tgacgtttta tgggggtgga gttttttgc aagttgtcgc gggaaatgtt 180

acgcataaaa aggcttcttt tctcacggaa ctacttagtt ttcccacggt atttaacagg 240

aaatgaggta gttttgaccg gatgcaagtg aaaattgctg attttcgcgc gaaaactgaa 300

tgaggaagtg ttttttctgaa taatgtggta tttatggcag ggtggagtat ttgttcaggg 360

ccaggtagac tttgacccat tacgtggagg tttcgattac cgtgtttttt acctgaattt 420

ccgcgtaccg tgtcaaagtc ttctgttttt acgtaggtgt cagctgatcg ctagggtatt 480

taccggtatt caaggattac atgagcttag aaatgtaatt agcatagtgt gtggcatagt 540

gtagatacca aataaatatg atctctcctt ctactcttga aaatgcaaac acattcttgg 600

tggtcctaaa atagcctgta acatggttta ctcagcagca tttgctattc aaggcagatc 660

tgcctttagt cattggctgc gctcctgaac agctgtgtga aaggctaact tttgtaaacc 720

aaatcaaaat aaaatgcagc aaaaatttgt cactgaaagg aaatcctcag tatatccttt 780

tatgaaatga aagatccctc atccaaactt aacttttta aaagtgcgca tttggagata 840

tagccctttc ttatgaatcc taattcaatt ttggccataa acacacgttg atgttcccca 900

ccccaaagca catagcaaca agagtaggtt ctatattgaa aataatgaca atttaaaaac 960

atgtacttat ttcactgtat gtggacagtg tctatgattg catcatgaag tgtcatataa 1020

ccatgtacgt gtacatgaga gagagataga gagagaagtg gtagggtggt ggtggtagag 1080

gggatggcga tagtaatcat ggtaatggta gaggtgatgg aggtggtaat gacggaggta 1140

agggtggtag tgatgatggt ggtggtggta atggtggtgg atgtggtggt ggcaattggg 1200

atggtgggat ggtggtagcc atggtgatgg tggtaatggt gttgatttaa agggtggtgg 1260

tagtgaaggt gagggtagtg gtggtggagg tggtggtgct ggtagcaata gtgatggtgg 1320

tgatggtgtt gatgagggtg ttgggatcag ggtgagttcc cacagtatat ttcattcttg 1380

ttgtaccact ctgtcaacag caccactgac tgggacagag gaagaaggca cactctgaat 1440

gtgttattaa cagaaacctc aaaacagtct gtctccttgt agtcattcaa aattatcttt 1500

ttcttacctg gaaaactgaa actgaattac cgggaaaaac acaggagatt tttgtttgtt 1560

aatatgctgc caataaagta attttatgtc aaatttaact acaggaaagg gcaaggcatt 1620

ttctaagttc cttagatgtc atgtggctaa aaaaaacaaa aggatggaca gcagttagat 1680

actgtacact tagctgtttg aagccatata ttcagaaagc agatgttggg agttggtgtt 1740

tgaggactga tttcctggag gtattttata taggccaagt tcattgttct aaactctaag 1800

ggcttgactt gagggaggaa aagaggcaag aacatgttta gttttgctga cagcatcaca 1860

tgggcagccc taaggctaga caactttagg gcctgaagct tattctagga aagaagcacc 1920

tacagagtgg cactgggctc ccctccacta tagagatgaa gtcatatgac agtaaagggc 1980

aggcagggct gcctaggggg cccagaactg acacttccat tagaatgagc acaggccagg 2040

gagagaagtg gggaaccaga gagaaggagc tggaattcta gtaggacaaa cggtaagtga 2100

acaacaagaa caagttaaga gtgtgtgcag tattctttca aagactgaaa aaatagtgat 2160

gtgatagaat ggcaggtggc tctgagcagg ccaggagaag gactgggggc agagcatccc 2220

aggcaggagg gcagcaagtg ggaaggccct ggggtggggc ttttggactg ttccagtgac 2280

gggcaggcag ccagtgtgcc tgtcacacaa tgcaccaggg aagtagtcgt gaatttgcag 2340

agggtcttgc aggctatggg aaagggattg gattgtattt tgtttgtagg gaagccatcg 2400

ggggacttaa gcagaggaag gattggcttc atctctttga aaaagttctc tctggatgct 2460

gatgggagga gaaatggaag gaaaagaaac acttttaggg gcaagaactt ttgagaaggg 2520

tggaattggg agtgtggagt tggggccagc tttggcacag gaggggaagc taaacacgtg 2580

gccgcatgag ggcctgtaat tctacctgaa atgggtacca tttgttaggg taaacaaatg 2640

aaccaaatgc ccagtgatac agaccaagtg ttggcaaact tcttctgtga tggcccaggt 2700

agtaaatgtc tcaggcttcg caggccatgt ggtctctgtt gaagctctgt gtagtagaca 2760

atatgttaat gactgggcgt gactgtgtgc taataaaagt ttatttacaa aaacagcccg 2820

tgggctggat ttagctcaca ggctgtagtt tgccaacctc tgacctagag catgaactga 2880

gcatcttctt ggagggaaat aagttctttc caagttgccc tcctcacatt gcagggggcc 2940

atgtaggccc attattcaca gaagagtggg tgggcaacct ttctggagca gaaaaacgta 3000

aagatttctt ccgtagtgca agtaaggtga ccatttctaa accgtgcaag tgatccagca 3060

gtcccaaaag ttgtttcact tctcattgtg cgcccgttct caggtgctcc gaagcttcca 3120
gtcctttgta gggacatgga tgaaattgga aatcatcatt ctcagtaaac tatcgcaaga 3180
acaaaaaacc aaacaccgca tattctcact cataggtggg aattgaacaa tgagatcaca 3240
tggacacggg aaggggaaca tcacattctg gggactgttg tggggtgggg ggaggggga 3300
gggatagcat tgggagatat acctaatgct agatgacaag ttagtgggtg cagcgcacca 3360
gtgtggcaca tgtatacata tgtaactaac ctgcacaatg tgcacatgta ccctaaaact 3420
taaagtataa taataaaaga aaaaaaaaaa gagaggagag aaacatcatc ccctccagga 3480

tacccttggg ccttgttctt atagtcttgt acattgttga acaatttgca tgggctagtg 3540
gattaaagca caccctccac cctcaggccc tcaagggtct ctatgataat acagtctcac 3600
cttctaccct ttccatcacc atcctaggtg ctatggccaa ccttgaggct gccatgttag 3660
gtctatgcat ttcccacctc caccacataa ctctctgaag gccaggtagt ttcctattca 3720
tcttggtaac cccaaagcct cgtgacaggg ctcagctggc atctgcggat gtgaatgaac 3780
cattggagaa aatggtactc tgcaaataac tctgttattt tcccatttcc tgtgtaaggc 3840
ctagagacaa tgacttttta attgcacccc ttcccctctg tatgacactg gccttctctt 3900

gtgtccagca atgtgggtgg cctagatgat ttctaaggga cttctggcca agatgaacag 3960
cagctgcatc ttactgagca tttactatgt gccatatact cagccacagc tctaggggca 4020
tagaagcagg agctctcagg gtcagggcag tgagtgagca agcgagcacc tatgccagcc 4080
ctgcctctgg atggggactt gagagggtga tggaagcctg cagcactgga gggaggcaga 4140
caaagacagg cctgtgctga gggggcccgg agcaagagag agggaggcaa tgacagcaga 4200
gacatgcctg cgccttgggt ttgagtgccc agtggtcaaa tccacttccc tgtggctgat 4260
gcttgccttt ctaactttgg aatttagggg ttggagatct ggtgagaagg taggagggag 4320

atgaggagga gaagggaaag gcaggaagga aggggaggga aaggaaaagc aaaaggggag 4380
gaggaaggtt tccaacaaat tattctatat caactgcgga aatcaaaatt tgttgcccaa 4440
atcttagaag ctcatgtccc tcctccccag aagtctggaa tgcagcactc caggggtagc 4500
ttataaccca aatatctatc tgtaaaaaga gaaacattgg gctttcgagc tgtggattct 4560
cagtaaaagc aagaggcctc agcctacaca ggccagccca gagtttgagg aaccccaggc 4620
ccacacccac agggctggcc cctgggtctg catactccct agaaatgtgc acacttctga 4680
gcctcaactc tgtcctggag tctaacagca tccctctcct tcctggggca gttccacctc 4740

cagaaacctg ttaccttggg ccttatgtca aggaaactgt gggaaagagc taggcaggaa 4800

tgcagatgag gccagcatgg gctcctaaaa gtttagaaat aggcagtgtc atgctcccag 4860

gtgcctgcat aaaccagctg aaaaatggag ctcccctcac cagcactctc ccttcaaaca 4920

gactgtgatt tgcaggtcac tggtttacca agccaggcta cccaggcagg acccagatgc 4980

caagcccagt ggtgtcctgc aagctgagca gtgctcagtt cttgcaaaaa aaggtctgtg 5040

tgaaggcaag gcctctgcct ggcttctcac cccagttggg tgtctggaac aggaaggagc 5100

ccttactgca gaaaaaggag gagggagcaa agggagcgaa cagctgcgtg ctccatgggg 5160

aggatcccca aagtagaaag gcgcatacac actgcagccc ttgacccaga atgctcacag 5220

ctacattaca gattcaggtc tcctcagtgt agtggggctg ctgatgagac tgtggcatcc 5280

tcaggggtca ggacacacat tttccatcac tcttctgatg gcaaaaaacc tctgagccaa 5340

tgccaacctc tgatcattaa aaaaaagtgc tcacagcagt gtgtggttta ggatcatgcc 5400

ctgtgtggtt tggaacacgt gcacaaccac accttgttca tcaccatccc agaaaccctg 5460

acgcaggcaa agagcagagt tattaaccct actttactga tgtggatact gaggcccaga 5520

ggctcatgca agttatcaat aagtggcagg gacagttgcc tctagattaa ctagccccta 5580

ggatcacctg ggtcttggaa ggggacccat aaacatgagc tcccctctct tggggccaga 5640

tttgcacctg tgccgcgcct tcagcctgca tgaagtaggg gctgctggca aagactcaaa 5700

gctgtaaatc tgggttttct cttgaggctt ctaagggagc tgtttcgaca actcactctg 5760

ttcccagctg gctgcccctg catagggttt taaagcagcc tagctttctg ccaggcttgg 5820

cagtggacaa cgctggtcag aacatcccag agagctacca gaatgaagta agtttgcttc 5880

tactctttac ctgtttatgg gctgtctctg ccactggaat gaaaggcact gagaacagtg 5940

cctggcctgc agaaggccct ggaaatacct gagctcctaa tctgggaata ggagtaggaa 6000

gagctttgga ggcagggcac ctgagtttga gatctacaac ttcctgcctg tgtgacattg 6060

ggaaagtctc catcctttct gagcctcagt ctccaccctg gggaagtgga aatatcaatc 6120

tctgtgacac agaagcaaat gagcgaatgt gcacaaagta ccttgcacaa gagagacgct 6180

caaacacttg cctccaggtt tcaccgagaa ctacagagta agatagattt gttcccagtg 6240

gaggaagcct gggaataatt tgcccctaga ctatgaattc ctggggctca agatcgagca 6300

cagggccagg cacacagaag ggaccctgga aatgtggcag gaggccagag atagacaggc 6360

ccttagagct catacccatg ccctctgacc tcaagaagaa agaaacctgc tcaaaatctc 6420

acaaagagct tgttccaacc ctgaatcgag tctgaggact ccttcctgag tccagcactt 6480

tttctgcaag aagtatatgc ctccaaagct gatgggcgca aatcttgaac cccgtcacat 6540

aaacacaaag ggaggaggtg actagagctc ctcctactgg atatgtctaa ggtcaccagt 6600

ctaaagaaaa gggatggata gaatgaggcc agtatttttg cagccatcca aatgtccaca 6660

tacgctgtta cactgagggc tcctctctcc cccgtcttca gccctacttg catttagagg 6720

tgagaaagat atgggctgag gggttgtttt tcatcgtatt gtagatggaa agcacactgc 6780

ccttggggcc atccaaatgt ggaccttgat gtagcacccc accttctgga tggccatcct 6840

tctgaaagtc actgaatttc tcagacttta ttctctttat ccataaagaa ggagaataat 6900

aataatcccc ccaccctgcc caaccactga ctggttggga agctcagaag aaatactggg 6960

cacggcatcc cattgtaatc tatagagtga gtcgcttctt aatattaaat ggctgaacac 7020

agaagatgtg caaaaagtac tgtgtcccct tcctcctcca actgaacatt tcatgccctt 7080

tgcaccctca ttttgtctag gagctgcctt atgaagggaa taggtacctg ctccgagctg 7140

gaggaatctt tgccacttat ggtggggtat ggactgagac agagatggca tgtgacatgc 7200

gcactgagtc tcaactccat gcaggctctg gagcactctc aaattggagt actaatgcct 7260

tttaaattct cacactagca atcctttgac ctactgatct agggatctag ggaaagaatc 7320

gtgatcttaa cttcaaaggg aaggacaaaa tgttctgcct cctgttaaaa ctccatacac 7380

taagtgcaga gactggatgc cttattaacc ttgggtagat gcccaaatgt tcaaaaggtc 7440

aaactcttct gttccccaga tcgccagagt cattaaccag tcacactatt aaatgaatga 7500

acagatgctg aaaaggtact tgcattactg agatttctta tggtgatggc ccctgcctga 7560

tatgtattca gcattttgta gttttcaatg tgcattagag tatagtggtg atgacattgg 7620

cctctgagtt tgccacttct tatatctgtg actttggtca aattgcttaa tctctctgag 7680

tctcggtttc ctggagataa taatagcttc ttcttcccag ggttatcatg aggattacag 7740

gagataatgc cccaaaaatg cttagtaaag tgcctagcac ctagtcaatg ctgaattaaa 7800

ggtggttatt cttacttttc gttcatttga actttgttct cagggagggc aaaggataga 7860

caaagcccca tagctagtga ggagtagctg caagactaga acccaggtgt tctgagccct 7920

agtcttaggc caagaacaac tgttacgtga gatgcacgtt ttccttcaag ggagctcaca 7980

attatttcca tgtaaattca aggactgcta aaagagaact ctcctctggg actgatatca 8040

ttttatttca agattgattt gaaacatgtt ttttgtttgt ttgtttgttt tctaggaaag 8100

aacaagagaa ccagttaagc tgaatgcctg aagcaaatcc ctgttagcga tgttttcagg 8160

atgagggaga gtggtgcaag aaacgtgctt ccagatgcac atggtttcct gggactaggg 8220

ttcagggtgt catccctggg tgttattaag tgtcagaagg agagcaaaca agggaaacat 8280

ctgagatcca gctaaggcta caccctggaa atgcaagccc agctcttgca aaggacctcc 8340

tttggccact caccttccag gccttacaat aacttgtttg gactgcaggt ttcttggtgg 8400

actcacaggc cattctgctt ttatttggtc aacctcagtt cacaagcacc cagatgctga 8460

gatcctcagc atgtgcagca gagtttcata ttagcactgg gtacctttct gaggctacag 8520

ggataccgta cagcagcacc tgtcacgtcc agccaaagga gtgggctctc tcaatgtcat 8580
ccaatgctgt ttcaactgtg aagaagacca tctgagagag ttgcttttgg aggctgaggc 8640
aaatttttaa aattctttgt tctcctcaac tggggtgaat tcttggtctt ctaggacagc 8700
ttgaagtttt agaaagagtc aagccactca gaaccaacag agaactcttt cagagaacaa 8760
ggtgtggcat agaggaggca gagggctgat cttgatcaaa tccaaagtgt gactctaaag 8820
caatgaatgt gaatttttgg caaagcttac aaagggctct aaaggccatc tgcaaagaga 8880
agccaagcct gatcgatgaa tcactagtgc ggccggatat cgatcggcac gctgttgatt 8940

ttctcatagt aaggaacagt gggccctttc agtcccactt ctgtagtctg tggtactaca 9000
aatggtgagc ccatgatgtt gccattcata gggttattct ccagcagtaa tgactggcca 9060
gccactccca tagccgcggg gctaggattt attgtcaatg gagggacctg cagttctgca 9120
caagcagtac taggatgagc acctgggccc attgcaaggg tgacatcttc aaggcaaggc 9180
ctcttaattt tattagggta gcccccatca gccatgtctg gaaactggaa gtggtcttct 9240
tcttgtctcc tcttaacagt tccctgtgaa tggaagagaa gagaggagga gaagagagga 9300
gaggagaagg gaagagaggt gacacacaca cacacacaca cacacacaca cacagagaga 9360

gagagagaga gacagagaga aagagagaga gagagagaga ggaattttta taaaggtttg 9420
gcacattaaa gctaatgaac aggaaatgtg catgataaaa cagacctctc agtttaaaga 9480
cttatagttg tgaaaactat aaaatacagc ctgtctttgg aaccatagtg cttatttatt 9540
cattattatg tttcatctaa actgtctaat tacatttcaa ataaggcatt atgttgtctg 9600
tatactaaaa cgggatagaa cgttattcaa agggtaatct gcccacttca aggagagttc 9660
aacaaaacta tgcagaagtc actaaatgaa ccatgctgcc aaaggcaggc attggagaga 9720
aaactagaag tagctaaata gttttaattc tttcctgtct acagacacat agattttaac 9780

gaaggaatac catagtatag aattgaactt ttaggctgcc ttctagtctt ggttaaatgc 9840
atcaggctgc agtggtaaaa ttgaatacaa cagagccctt acaggaaaga agtagatctg 9900
gatgtgtttt cttggggagc tgtttaaaat actgtttttg ggaaagcaca agtttcagaa 9960
cagtcattgt aggcatcgta ttcattgttc catttatttt tacacacaca cacacacaca 10020
cacacacaca cactctcaca cattgctatg tgtacacaaa aataatttgg aagaacctat 10080
acccaacaat ttggagtggt catttatttg ggatgactgg caattccctt tctattctct 10140
tcatttctgc ttgtttgtct ttaacgagaa cgactcataa tccaaaaatt taaaaaagta 10200

taaagttatc taaataagaa attttcctct gaagatgcat cctcaggttg gggagatatt 10260

aaacaatgag aaaaggcccc aatctgggat ctgaaccttg ggggagctgc ccatcattta 10320

tagaagcaca gcctttggga acaaagcaaa gtcactagca atgtgagact tcctactctt 10380

catggcttca tacagtcatc catcgctgtt gtgttaatga ccatgacctg tatgttagca 10440

ggtaaatggg aaaggaagtg ggggcaaagg agtatgtgca ggaatgatca aaataaggaa 10500

aggaagagag ggatctggaa atcacctgaa tgccgatagg tgaacaggta gaattctttt 10560

aaagcttccc ccacccggta ccccccaaat aacccctttc cagctttgga agtttcacta 10620

ggacatacag tgctcatcct ctgatgtcac cttaagtttg gctcttctgg tttgatgagc 10680

ttgtagccca ctaggagctc aaggcatgca tggggccact tgccagcacg atgaggggca 10740

tgactgtcat ggccaagtga acatcaaagc agatccccag ggctgtatgt ctcaggcctt 10800

ggtgcacatc agaatcactt agaaacatcc acattcctgg gccctcccac cacaaactga 10860

cagcttcatc cagggtgtgg cccaggcatc gggagttttt ccaacagctc catggctgat 10920

tctcaacaga aaaccactgg cccagagcaa gggtggaggc agcgtggcat agggctctga 10980

ccttggcctt gccactgaac ctctcagagc cccagtttct ttatgtgtaa aatgagtgta 11040

attatagttc ttttctcatg aaggtgctct gactattaag tgaaacgggg cacattgtat 11100

gacacctaat agctcctcac taactggtac ccggcattat aaagggcagg tatggaaggg 11160

ttctgggagt ccaataccct tcttaaagac agagaggtct ctgagaccca gagaggggca 11220

ggccttaccc agagttgctc agccagaggg caacaaggcc caggtcagat gcagggcccc 11280

tccaccacca ctcagctgcc tccagaccca ctgccttcgc catgttgttg gtaggacact 11340

gcatcgcccc cacagaaggg gcttgccaac ttgagtgaga ggacttgcac acttctttga 11400

cttttctttt gagatgccca caatctgaac aagggcactt caagggacag ctctgtcacc 11460

aaactcatct gaggcctgaa taccatgggt caggcaggaa tgggttggag aggtgtagag 11520

caggcacaat aagagggctg aggcccatgc agtcatcagt gcccactttc ccaggagtct 11580

gactgggcac agcacccata gtgtccctga gctggtccat ggagcagctc actaactgtt 11640

tggcccacag caggtgctca gtaaatggca gttgaacgaa tcaatggaca aaggaacata 11700

aattacccaa cacacaggga gctcagccat ttactcaatc cattatggag taacctacaa 11760

acaagccact gggtcccaaa ctgaaattgt gtctcttcta cattctccca aagaatccaa 11820

taggttaaaa atagaaatgt atgaaataga tcaatcaggg atgattgcat gtggatttga 11880

cataaggatc ccctgcaggg agtctgagct ggcaacagtc aggcccaaag tgctgtccat 11940

gatgtctcga actgcaagac agttttaaca atggcgaagc aatgcagaac caggcaggcc 12000

aaggaggggg tgggggttgg ggaaaggaag ggagggaagg ggctgtgagg ggcaatggtc 12060

tggcatccct gccacgtgag cctctgaaat ttgctggcag cttctatggg ctcccagagc 12120

tttcacttaa ttgttggtct gccactaacc tgctgggagt aaggtgcagg gatggaggag 12180

gcagggcatg accaccagac actaaaggta ccagctgggg ccactggcaa agggaaggag 12240

gctgcacctc tcctacatga gagcccgtat acacacacct tttccagcac tcatcaactg 12300

catcccaagc aaatggtccc tgatcaattc caattctaga aaccaactga ctactcaata 12360

acaaagtaga tcccagcagg ccgccactgc tggagcggat gccacttttg ctatgccaag 12420

tctgtggctg gacagctgct ggcatgtaca ctcactgact ttcataagga tgcctaataa 12480

agggggcagg ctcacctggc ttttctcagg ggtggggttt ggggtgccga tagaggctgc 12540

tgttttggca gagtggcaag ctgcaagcct cttctgagct ttcatttttc aatggacttc 12600

agtgagaatt cactttgtca gaggccatgc agctccatgt tttggatttc atggaatgag 12660

ctttcaacag tgagcctgaa gtgccctggc tgaacagcaa gaacaccagc caaccctaaa 12720

caaggccgag gagaggcggc tgtgtttaca cggaaggctc agccttgctg taatagcgtc 12780

tgccttcacc agacatcagt gaggcgtgga aatctattat ccagttaatt ttgcccctag 12840

ataaagactt gctttcgtgt cttctctttc acagtcccat gatctgttac tcatctcaac 12900

tgcgagaagt tggctgggct ttcccctgtg cccagtgcca cactcgtgcc ttcactgggt 12960

cacctgtgcc tgtggctgat gccgctgagg ttttgcctgc ccagactggg tgtttctgac 13020

taaatcccac agccaccatt ttagatcaag ggcaggagat agctcactgc tccggaatga 13080

cctcccctcc cagaatcctg gtaggggcgg aaggtcccca accaagctcc cagccctttc 13140

taaatgaatc tccctgcttc acccatgtgc ttttctccag tctctgcggt cttgatgaca 13200

gcagggtatt agtcctagct gtcccacagc tcctacttct ttcaggcctc tccctgtgac 13260

aatcagtagc cactggcagg atttcctcag agcatatctc gatttgcttt cagacaatta 13320

gttaaaagga cactggaccc cagacgtccc aactcccagc cagagccctc acaggcccgg 13380

cctttggtgg tgaggaaggg ggagggagtg agtgacagtg ccctggcatc ttttagaaac 13440

gaattccttt ctctccatac ataaatgcct gcagagtccc atttcagaat ccggcagaca 13500

aagccaccaa tgtgatcccc atgaccttat aaacattcat taaaatgcat ttcaaggcat 13560

gtgatggcct ccccaccccc tagataatga gaaaacaaag gtttctcttc tgatagagac 13620

aagttcagct ctgaagtcaa cattatttct ggttctgtct gaacaatgac atatggcaac 13680

tcttcccttt ctatagttct agtccagaat gacaaaaaag gggaaaaatt tcttagagaa 13740

ggtagagatt atacgaatac agtccatgaa atgagcataa ggagaataaa gaatataact 13800

tatccaaaga agtctggcag gctgttataa atgcttgatt ttggacactg tagttggagg 13860

tttaacatgg acaccaataa aaaggtcagc aaagggtatg cactgttcct attgggcaag 13920

aagataggag gtcaaaggta accaggaaag ataaactcag ggagacttat tttccctcca 13980

gagggcactg ggcttgtagg ccctgggcaa aattgtcaaa aaggtgaaaa tcgcctgtgg 14040
tttatttagt ctgctctttc ttcactagtg cctcaccagt tcagttcagg ccaatttgct 14100
aggaattcta ccgggtaggg gaggcgcttt cccaaggcag tctggagcat gcgctttagc 14160
agccccgctg ggcacttggc gctacacaag tggcctctgg cctcgcacac attccacatc 14220
caccggtagg cgccaaccgg ctccgttctt tggtggcccc ttcgcgccac cttctactcc 14280
tcccctagtc aggaagttcc cccccgcccc gcagctcgcg tcgtgcagga cgtgacaaat 14340
ggaagtagca cgtctcacta gtctcgtgca gatggacagc accgctgagc aatggaagcg 14400

ggtaggcctt tggggcagcg gccaatagca gctttgctcc ttcgctttct gggctcagag 14460
gctgggaagg ggtgggtccg ggggcgggct caggggcggg ctcaggggcg gggcgggcgc 14520
ccgaaggtcc tccggaggcc cggcattctg cacgcttcaa aagcgcacgt ctgccgcgct 14580
gttctcctct tcctcatctc cgggcctttc gaccgttgat ccggtggtgg tgcaaatcaa 14640
agaactgctc ctcagtggat gttgccttta cttctaggcc tgtacggaag tgttacttct 14700
gctctaaaag ctgcggaatt gtacccgcgg aattaatacg actcactata gggactagta 14760
ccatgggaaa atcaaaagaa atcagccaag acctcagaaa aagaattgta gacctccaca 14820

agtctggttc atccttggga gcaatttcca aacgcctggc ggtaccacgt tcatctgtac 14880
aaacaatagt acgcaagtat aaacaccatg ggaccacgca gccgtcatac cgctcaggaa 14940
ggagacgcgt tctgtctcct agagatgaac gtactttggt gcgaaaagtg caaatcaatc 15000
ccagaacaac agcaaaggac cttgtgaaga tgctggagga aacaggtaca aaagtatcta 15060
tatccacagt aaaacgagtc ctatatcgac ataacctgaa aggccactca gcaaggaaga 15120
agccactgct ccaaaaccga cataagaaag ccagactacg gtttgcaact gcacatgggg 15180
acaaagatcg tacttttgg agaaatgtcc tctggtctga tgaaacaaaa atagaactgt 15240

ttggccataa tgaccatcgt tatgtttgga ggaagaaggg ggaggcttgc aagccgaaga 15300
acaccatccc aaccgtgaag cacgggggtg gcagcatcat gttgtggggg tgctttgctg 15360
caggagggac tggtgcactt cacaaaatag atggcatcat ggacgcggtg cagtatgtgg 15420
atatattgaa gcaacatctc aagacatcag tcaggaagtt aaagcttggt cgcaaatggg 15480
tcttccaaca cgacaatgac cccaagcata cttccaaagt tgtggcaaaa tggcttaagg 15540
acaacaaagt caaggtattg gagtggccat cacaaagccc tgacctcaat cctatagaaa 15600
atttgtgggc agaactgaaa aagcgtgtgc gagcaaggag gcctacaaac ctgactcagt 15660

tacaccagct ctgtcaggag gaatgggcca aaattcaccc aaattattgt gggaagcttg 15720

tggaaggcta cccgaaacgt ttgacccaag ttaaacaatt taaaggcaat gctaccaaat 15780

actaggggcc ctaaccgcgg ggatcaacgc ctagagctcg ctgatcagcc tcgactgtgc 15840

cttctagttg ccagccatct gttgtttgcc cctcccccgt gccttccttg accctggaag 15900

gtgccactcc cactgtcctt tcctaataaa atgaggaaat tgcatcgcat tgtctgagta 15960

ggtgtcattc tattctgggg ggtggggtgg ggcaggacag caagggggag gattgggaag 16020

acaatagcag gcatgctggg gatgcggtgg gctctatggc ttctgaggcg gaaagaacca 16080

gctgggggtt acctactttc tttatgtttt aaatgcactg acctcccaca ttccctttt 16140

agtaaaatat tcagaaataa tttaaataca tcattgcaat gaaaataaat gttttttatt 16200

aggcagaatc cagatgctca aggcccttca taatatcccc cagtttagta gttggactta 16260

gggaacaaag gaacctttaa tagaaattgg acagcaagaa agcgagatta gatccctatt 16320

atttttgaca ccagacaagt tggtaatggt agcgaccggc gctcagttgg aattcgctag 16380

ccgcacatac agctcactgt tcacgtcgca cctatatctg cgtgttgcct gtatatatat 16440

atacatgaga agaacggcat agtgcgtgtt tatgcttaaa tgcgtactta tatgcgtcta 16500

tttatgtagg atgaaaggta gtctagtacc tcctgtgata ttatcccatt ccatgcgggg 16560

tatcgtatgc ttccttcagc actacccttt agctgttcta tatgctgcca ctcctcaatt 16620

ggattagtct catccttcaa tgctatcatt tcctttgata ttggatcata tgcatagtac 16680

cgagaaacta gtgcgaagta gtgatcaggt attgctgtta tctgatgagt atacgttgtc 16740

ctggccacgg cagaagcacg cttatcgctc caatttccca caacattagt caactccgtt 16800

aggcccttca ttgacagaaa tgaggtcatc aaatgtcttc caatgtgaga ttttgggcca 16860

ttctttatag caaagattgg ataaggcgca tttttcttca aagccttgtt gtacgatctg 16920

actaagttat cttttaataa ttggtattcc tgtttgttgc ttgaagaatt gccggtccta 16980

tttactcgtt ttaggactgg ttcagaattc ctcaaaaatt catccaaata tacaagtgga 17040

tcgatcctac cccttgcgct aaagaagtat atgtgcctac taacgcttgt ctttgtctct 17100

gtcactaaac actggattat tactcccaga tacttatttt ggactaattt aaatgatttc 17160

ggatcaacgt tcttaatatc gctgaatctt ccacaattga tgaaagtagc taggaagagg 17220

aattggtata aagtttttgt ttttgtaaat ctcgaggtat actcaaacga atttagtatt 17280

ttctcagtga tctcccagat gctttcaccc tcacttagaa gtgctttaag catttttta 17340

ctgtggctat ttcccttatc tgcttcttcc gatgattcga actgtaattg caaactactt 17400

acaatatcag tgatatcaga ttgatgtttt tgtccattgt aaggaataat tgtaaattcc 17460

caagcaggaa ttaatttctt taatgaggct tccagaattg ttgctttttg cgtcttgtat 17520

ttaaactgga gtgatttgtt gacaatatcg aaactcagcg aattgcttat gatagtatta 17580

tagctcatga atgtggctct cttgattgct gttccgttat gagtaatcat ccaacataaa 17640

taggttagtt cagcagcaca tgatgctatt ttttcccctg aaggtctttc aaacctttcc 17700

acaaactgac gaaccaggac cttaggtggt gttttacata atatatcaaa ttgtggcatg 17760

gtggaagctt ggcatgggaa ttcctcacga cacctgaaat ggaagaaaaa aactttgaac 17820

cactgtctga ggcttgagaa tgaaccaaga tccaaactca aaaagggcaa attccaagga 17880

gaattacatc aagtgccaag ctggcctaac ttcagtctcc acccactcag tgtggggaaa 17940

ctccatcgca taaaacccct ccccccaacc taaagacgac gtactccaaa agctcgagaa 18000

ctaatcgagg tgcctggacg gcgcccggta ctccgtggag tcacatgaag cgacggctga 18060

ggacggaaag gccctttcc tttgtgtggg tgactcaccc gcccgctctc ccgagcgccg 18120

cgtcctccat tttgagctcc ctgcagcagg gccgggaagc ggccatcttt ccgctcacgc 18180

aactggtgcc gaccgggcca gccttgccgc ccagggcggg gcgatacacg gcggcgcgag 18240

gccaggcacc agagcaggcc ggccagcttg agactacccc cgtccgattc tcggtggccg 18300

cgctcgcagg ccccgcctcg ccgaacatgt gcgctgggac gcacgggccc cgtcgccgcc 18360

cgcggcccca aaaaccgaaa taccagtgtg cagatcttgg cccgcattta caagactatc 18420

ttgccagaaa aaaagcgtcg cagcaggtca tcaaaaattt taaatggcta gagacttatc 18480

gaaagcagcg agacaggcgc gaaggtgcca ccagattcgc acgcggcggc cccagcgccc 18540

aggccaggcc tcaactcaag cacgaggcga aggggctcct taagcgcaag gcctcgaact 18600

ctcccaccca cttccaaccc gaagctcggg atcaagaatc acgtactgca gccaggggcg 18660

tggaagtaat tcaaggcacg caagggccat aacccgtaaa gaggccaggc ccgcgggaac 18720

cacacacggc acttacctgt gttctggcgg caaacccgtt gcgaaaaaga acgttcacgg 18780

cgactactgc acttatatac ggttctcccc caccctcggg aaaaaggcgg agccagtaca 18840

cgacatcact ttcccagttt accccgcgcc accttctcta ggcaccggtt caattctagt 18900

atcgataaat aggggattac ttgaacatag actgtgggat ccggtgtgga gtgcgggaga 18960

ctagcaaagt gaatcctgag agtagcaggt ctgcacctgt tggatcgaga aaggcggcct 19020

acaattctgg tcaaatgagc tgtgcttatt gacatattct attagagagt actaccaggt 19080

caccagtcac cagaaaggct gccagctctc caaccacctc cagggaacta tcctgaatgg 19140

ggccttaaca agtctaagag agggttggtt tgggtcccaa gccaatattt gctctgcttt 19200

atgtcagtca tatggaaccc aaaccaaccc tctcctatgt gcctcaccag tcggtgcagg 19260

gatcccaatt tcaagtttgg ttttttatgg tcaaagtcca gcatagatta aatgaagggg 19320

tgtgatgatg gtgttaaaag agaactccag accagtttaa ctcttggaca cacatcccat 19380

ctcaccatgg tgcttccaac cttccagaga tgatgggctc ctattttctg atgacaaagc 19440

cctccacagg attgctgcct ggccatcagg gagtgcctct gtaactgagg ctgagatccc 19500
actttcagtc ctccagctgt ggcccatccc tgctccgccc accgggtatg gcctgtccta 19560
ggctcttagg tatggctgca ttgtgaaatg atggctacag agctggcatc tcctgtagtc 19620
tggttcatct agtgcactac ctcatagtta aaagaaatct gtttaagcca ctgagggtgg 19680
ctcctagtgc caactccaag aacaggaagc ttccctttt tgggaggagg ggcagatggt 19740
aacatggatc gtccaggtca atgggagcag ggcaaccaca gtaagtactg gacaacaaca 19800
caaaactcca tgtgtggctt ccatcgagtc cctctccaat tggtttggtc ttctccgtcc 19860

catgcagcac tttagcaagg ggcctggctg aaggctatga attgtgtgga gcctcctcat 19920
tgcagtctcc aaccatctga tgctgggaaa atgtcaccag gatgcagcca tgccgtgtgg 19980
ccaatgaacc gagaaaacac ccctttttcta gaatgctcta aagaggcaga ataatccaga 20040
ggtgaggaag gaaatactcc accagagacc caggcagttc ctacaaaagc cagactttcc 20100
ttcacctagg gagtgacaag accagtggaa aacactctca agcagtaacc cccaaatgct 20160
ctgcaagcca gtggcgtcca gataccgcac aagcgagtgg gctgtctaat cccatcatca 20220
tgatgtaaat atctctaggc tgccctgggc tgtgcctgac cctgtcttca gctttccaca 20280

cctccaccta cagcccatgc acagaaggac cacccaggaa tgctgcaagt gtggcacctc 20340
cagggccacc cagggagaag gagggcagct atgctggtgg ctccaggccc atttggcggg 20400
tggtaccttc acaccacaaa gcccaaactg aggccccaga tttggctgat gagggcatat 20460
tggacagggg tcacttatgc tcttccccat tgccacctgg cctctggcta cctggacttg 20520
gctacctgtg gatcctctca caggtgccac catcttggct gagtctccag atgcgaggtc 20580
cctgaggcag tggcgggctt ctcgctaatg ctgatgggat taggaatggg ataggtgggg 20640
agggccctgg actgggccct gatgagccaa gtgggttttt agaggggcta ctggtacatt 20700

tcagggacag gacatctggt agagctaagc tggggcaata aggagccact gctaatctga 20760
gagctagaaa caatcagctt ctgggtcatt attaattagg gtagtttggg ctgtgtggaa 20820
gtcacgtact atatggggta gccacagctc tctctacaga taatctctaa gacttctgat 20880
tgggactgtg tgaatgcagt agcaatatct cttcttactg ccaggccctg ccagtcctgc 20940
ctccacgccc tggctggccc cccttatgat ctgacccatg ccaggctgcc atagtatgtt 21000
acttctgcat tagcactcct tgggacctgc ctctccactg tccctcagac tttaaagaac 21060
tatacaaacc caaggggctc ttcccaagag aattgatatg acttgaggtg attccatttc 21120

tggaagtagt cactccattt tctgcctcac tctttcagtg cttcacagag caggttcgaa 21180

cgaaggagcc atccaactaa ccgtcatgtt cgggcaaccg aagaagggag tggcaggatt 21240

tcctttggag acttctggaa ttagacagca gtttaatgca agcatctaaa ttctctccct 21300

cccagagtct cattaaaact acagtaagag tttgtgtttt gttttgtttt taaagacaaa 21360

atcccaccag gatagagaga ataggagagg agataacagc atcataattt atgaaactaa 21420

aatgcagata gaccaggatt aactgactac acagcaccaa ggaagctgaa tcacaagaca 21480

gcagaggaga aaactggaaa ggatcgtggt ctatacggca gaatcttccc aagcctcagg 21540

aggaggagct ctagatgttc ccagatctgg gaggtaaagt ggaatggggg gacatggtca 21600

gcgtaatggg gttgggctgg aagcaggtta aggagcaggc agatctctga atcccctctc 21660

tgactctgtg tccccaggca tctgcctgtc ccccaccctg gaagaggtct ggcttgaccc 21720

tttgtctggt gaatttcctg ctctgctttc ctggtcctgc tggccggatc agtggaggcc 21780

actcacttca ccccacaggg atgttctgtg ttgccctaca cctgggaact ggaggtactg 21840

gaggcaggct gtggtgagct tgaaagcaaa acacagaggg cagtccaatc tctttggcca 21900

tatttcttct gcatatccaa taccatgtcc acaactctgc tagtgtcctg atggtggtgg 21960

gctctacaca ttcccgggaa gctgaaggca gataatgacc aggacaggtc aacctctctt 22020

cttctgaaag ccttcatcta ctaatggcct gggactcttc ccttaaatgc ttagattgtg 22080

tcttccacta aggttttttg ctgttgctgt tgtttgtttg tttgtttgtt tgtttgtttg 22140

tttgttttga gacggaatct cactctgtcg cccaggctgg agtgtagtgg cacaatctca 22200

gctcaccaca accttcacct cctaggttga aggggttctc ctgcctcagc ctcctgtgta 22260

gctaggatta caggcacatg ccaccatgcc tggctaattt ttgtattttt ggtagagaca 22320

ggatttcgcc atgttggcca ggctggtctt gaactcctga cctcaggtga tctgcctacc 22380

ttggtctccc aaagtgctgg gattacaggt gtgagccacc acacccggcc aaggtttttg 22440

tttgtttgtt tgtttgtttg tttgttttgt attgaggcag ggtatcactc tggtcaccca 22500

ggctggagtg cagtagtgca atcacggctc actgaaacct ccacctccct ggcgggctca 22560

ggtgatcctg ccacctcagc ttcccaggta gctgggacta caggcttgta ccaccactcc 22620

cagctaattt ttgcgttttt agtagagaca gggtttcccc atgttgccca ggttggtctc 22680

aaactctggg ctcaagcgat ctgcctgcct cagcctccca aagtgctggg attacaggtg 22740

taagccaccg tacccggccc cgccactaag gttttgaaaa tgaagcaatt acaagtttaa 22800

gtctattaat aagtgatgaa gccatgtaga aaagcagaat aattatcttg gatcaggaag 22860

gtcacatgag gatctacttg ggggttgtca atattctatt tcttgacctg atcagtgttg 22920

acagcaggtt ttaatttttt acttcttttt gtttgtttgt ttttgagacg gagtcttgct 22980

ctgtctccca ggctggagtg cagtggtatg atctcggctc actgcaacct ccgcctcctg 23040

ggttcaagct gttctcctgc ctcagcctcc ccagtagctg ggattacagg caggcaccac 23100

cacgaccagc taatttttgt atttttagta gagactgggt ttcaccatct tggccaggct 23160

ggtctcgaac ttctgatctc gtgatccgcc ctccttggcc tcccaaagtg ctgggattac 23220

aggcttgagc cagcgtgccc ggcccatttt ttacttcctt attaaactgt acatataggc 23280

cttgcacact tttctgcatc aatgttatat tccacaataa agggaaaagg tatatacaca 23340

acttgatacc agtaatgtga aacatatatt tctacataga aaaaaaaatg actgaaatac 23400

tgcactccaa tgtgttcaca cagtagttgt ttctggatta tttatatatt aaatgtttat 23460

atattgtatt atgccatgag gtttgtgttt tctctccact tttctgcatt ttccaagttt 23520

actacaaaga gcacatatta ctcttataat cagaaagtca taaaatatat ttaaaaagac 23580

aaaattgaaa ctaataagga tcaacacaaa acagatgagc catctgtgga aatccgcaca 23640

gaatactacc taaagagatt ggtgacgtgc atgatctcac taggatgagc acaaagcttg 23700

ccagagccta gggtctattt ctagggttgg ctcttggaag ccaggatagt tgttatctct 23760

gggaagaggg aggggcacac aaggggcttc taaaacattc tgaatgttct atttctgaac 23820

ctggttggtg ggtacatgac tgttggtttt attattatat gttttatata ctcttccgta 23880

tgtatggtgt ggattccaaa aaaagatttc ctttagagaa aaccagaatc acataagtag 23940

aaaatatggt gctatgttga aggaacaact caagtttata taaaatcatc atcatttata 24000

ggcttaaaaa gttgctttgg aattttggtc taactgactt gtcttttctg cagcaaacca 24060

cgctccttct ggacgtgctc caggcagagg ggattagggt gggttcaagg ctgcaagtac 24120

ctagctcagc acactctctt caggggactt agagtttgtc tggtgttggc tctctgagct 24180

cttgtcagga atgccgaccc ttccgaggtt caggatttga agcctgcctt cccaccccag 24240

atttggtcca cacagacact caagtatgta tttcaactac aaatgacctg tactttccta 24300

ttactcctct ctttcatggt aacctttctg gtatccttcc ttccctacat ttatgggagg 24360

gggacatcat tctctgctct cctgtcactg aaggctccac cttctgtctt cttctgaccc 24420

atctggtttt cctggggcca cctcctctcc ttaccaccct aacgcttttg taacttgagg 24480

agaaatgaga gatcacctag tcaggtcatc attctctgta gatgaagagg cccaatggtt 24540

tgctcaagaa ttgccaagcg agttaaagac agagagtatg agagtcagca agacctacag 24600

aaagcatcta tctgcactgt tttgcaggga cttagccttt gtgtgtggac tcctggaatg 24660

ccacccacta agaaacattg tctgacacca actccccact tggtaggtgg ggacactgaa 24720

actcatggca ggaaagggcc ttgccccaag ccagggcaga gtgtcactca tcactctcaa 24780

ttttcagtcc agggcacctt gttgtgacta tcccaaaggc agccactttc cctggtctga 24840

aagacctgaa gagagaagag aagagaagga tggaaggcag agtatgcggc tttgattcat 24900

ttcctggtga aaacagatct atacgagaag caaatttcac gaaagggaag agaagaaagt 24960
gtcccatacg ttgctggcct gtttcaacct tgctttgatt cttgctgaaa agggtaccgt 25020
gtatttctga gttcaacatg cagaccagtg ttaggaaagc cactgcacct ccactttagc 25080
ctccagggct gtgccctgca aatggcctgc agccttggtg cctcgctctc cagactgcat 25140
tttggaagat gggacagagg cttatggaag cccacattag aacgggggag cagaatgggt 25200
gagatgaggg atccttgata gtgaaccaga tgaaggaatg gtagccaaat gccaggcctc 25260
ctttgtggct tcaatccaaa ggctctggag cccttccagg gcagaacatc aggcatgttt 25320

acccccactg tcctcaacag tgacagaggt gcaatcttgg gcagctggcc attttgaaag 25380
caacctcctt aatctcaact gggaaggctc cctagcagga cccctgtgtt gcacacctgg 25440
aggaagctag actaaccaga agctcagcac ggttccatct gggatgccca ggtctgagac 25500
gaaaaaggta actctctttt ctgggtcctg gcccagttgt gtctctctcc acctcattct 25560
ctgagatgcc tgtctccccc tttttgtccc atcaggaggc aagagctatc actgggccag 25620
actccaccag aagccaagcc agcttgttac ccagcttctc agggagcaaa gaacagcctt 25680
gtttctatct tatccccact gtcccctgcc cctgccccac ctcccagcca ttcagcttct 25740

ggcttcccca gagctgcctg cttctttgtg gtcctccatt ccttgaaaag accttctagt 25800
cattagtgta tataaatggc cacttagccc agattacagt gaggtcaaca gctggggctc 25860
tgagaattgt cacacactgg cacaggagag gaggctattc ttccagagaa tttggagggc 25920
actcccatcc acttacaaca aaaagcccat ccactgtgct tggcagtagg tgatctgaga 25980
accaatggaa ccaggttaat cctgtggcac tgttgagtga ggagagcagt ggcgggcact 26040
ggaaaatatc agagacaagg caggagacct gaaatctagg cttagctcct catatacttg 26100
gcagctgtat gacctcagac aaccagtgtt acctctctaa gcctcagttt cctcatgcaa 26160

aaggaggggg aataacaaca gagcccactg cttgggggtg ttgtgaggac aggatgaaaa 26220
aacaaacaga aatccctcag tacaggattc agtgcagtgg acagtcttgc aaggtctggt 26280
tcagccctcc acccctaccc tcaccagtat aaagaactct ggcctacaag tcagatgacc 26340
tgagttttaa tctcagcttt gccattagcc gtgtgaactt gagaaagtcc ctttcctttt 26400
tacatctatt gggatgatca tgcatttttt gtcctttatt ctgttaatat agtgtgttac 26460
attgattgct tttcatagac tgaaccagcc ttgtattcca gggataaatc tcacttggtc 26520
atggtgtata atcctttata caaatgttgc tgggttgagt ttgctagtat tttgttgaag 26580

attttatgt cttgattcat aaggaatatt ggtgtacctt ccccttttat ggccacagtt 26640

tccctacaat gatgtagtcg aactagacaa cctccaatat ctttcagtat tcatgtcctc 26700

tgattctgtg aaactaagaa aattaagaaa tagtgattca taggcacaag gcaggcaaaa 26760

cttagactcc ttgtagaata attaggaagc caaatattca gtgtgcttat ttctcaaata 26820

accttagttt ctccagtctg ccccaactcc gaggcctgaa tatctctaga tgcttatgat 26880

ggcaactaaa gcctaaaagc taattcattt taaagttctt ccaaatgcat agggttttat 26940

ttttccagac ctgggttcag atggggaatt tgacaaacaa tggaaagggg gaaaaacaac 27000

aatctaaaca ctgagtgaca aagtaacaaa gaaatagtct agctatcagc cagtcaagcc 27060

agccttggct ttgctatcca aagtagtcag tctaattcta ccaccagttt ctgttcctgt 27120

agctgtctac tgcctgccag ggactctgcc ttcccaccca caactaccaa tggaaggatg 27180

tggtgaccat accagtggct gctgacatct cctgccatgg gaagcataat tgcctccagc 27240

agcctccccc ttagatccat cattttgtt gcacttggcc tgggctgtac tcccggccaa 27300

tgactgaaca tggtgagcat agtaatgcag gcccatttct gtgaggagca ggactcctcc 27360

agtaggtgac tttggctcaa ggactctcta ttggcctggt tgaacttttc ctgaactgtg 27420

ctactgtctg agactcttct tacccaatcc tctttctcgc cccaattgtc acagaccacc 27480

tgcattgtgg tctgagtctc tccccacctt ctcttgctct tccctgttta tctttcacag 27540

gcatttcccc cagtacattc cttgaatgtc taacccgata cgggtgcctg acttttggca 27600

gacctaagca gacaaaaagg agtacttggt tacctagctc ttctttctac cacaaacatc 27660

gagggaaccc tttttccctc acccctctgc cacaccccca ctgccccagt gaacaaccac 27720

agagagagct gtggtataat attaggctgg tgcaaaagta attgcggttt ttgccattac 27780

ttttaatggt aaaaaccgca attacttttg cacctaccta gtatttgtgt ccccccaaat 27840

tcatatgttg aaacctaacc cacaatatga tgtcattagg aggcaagacc ttgaggaggt 27900

gattagatga tggggtggag ctctcctgaa tgagattagt gcccttataa gaagaagccc 27960

aaggaagcta ccttgactct tccatcacat gagaatgcag caagaaggca ccatctacta 28020

atcaggaaga gagctctcac cagacactga atctgccagt gtcttgatct tgaagttccc 28080

agcctccaga actatgcata atgcatttcc attgtctcta agccacccag cctatggtat 28140

tttgtcatag cagcctgaac tgactaagac agtgagccac atgagaagtg ccccaacccc 28200

tcccttaagc acttggctca cagatcagtg ggttcatttc tgcctgagtt ttattgttat 28260

tctgtagatt tcttgggcta gatatatttt tctgttattt tccttcttca cctcagtcat 28320

gaattggttg ttttaaaaaa gacaatgtaa gtcatgggga aactcctgac aactctactc 28380

tcctagggtt cctgataaaa ggggattcag ttgagtcctc tgatggtctc tacctgccaa 28440

agtccagcag cccttagcaa acatgctgct cgtttctgta gagaaggtgc tggtgtccca 28500

ccatacttct ctctccctca tgaagggctt gcgacccagc aaatgggtgg cttatatggg 28560

tctgtttcaa aggaagagcc agctctggga agaaaaacga tgagcataag cataacctac 28620

cactgtgcct gggaaagcag acaacttttt tgatgtgtga atatctaatg agaatggaat 28680

ccatcaatta ccttaaactt aggcacagtc ttcaaattca atatatgtgg gatatacttt 28740

tagtcagttt gtagacgtta tttgtaataa ataatctggc ttctctaaag aaattatttt 28800

aagtgtttgg tttggtttga tttaatggta aaattatatt tagtggcaga gaattatagc 28860

aatggtgata aactatagag tgtcataagt tcatatctta ttctcacatt tgaagctgcc 28920

tgcagatgca ttcaagatgc agccagaagt caggagactc aggctgttat ttggagctca 28980

tcattttaca gccttgctgg actcccactt tctcagggga aaaatgtggt gttgacccag 29040

attagctctc caggccctgc tgagttgggc actctgtaag ctggagggtc ttctattgtc 29100

ttcacctaag tgtcaatcaa caacccaaat gggcatgggg gaagagggag ctgggccaat 29160

gcccagggtg cctggtagag agataccttg ggcactggaa ggcaccagct tcccagagag 29220

aaggggggagg gccatgaaaa agttggctgt agatgccagg gacactggga ctctccagct 29280

gtgtgtttgt gtcttctgaa gacttatgtt tcattccttt ggagcatgca taatcataca 29340

ctgtgggatg tgttatatag attgcttgat agttcaccac tgtaataaaa tactgtgact 29400

ggaatctgct cccagtctgc ctttgatagc acttgtgcaa cacacattta ctgagcattt 29460

acagtgatcc aggacctgtg ttgtgaaaac attgatggac aaggcagatg gtggagcacg 29520

tcagtgagga tttttaacaa aggctggtaa gtgctataaa ggaacattgt aggacactag 29580

agaacaaaga acaggagaac ctgacttagg ctggggtggg gcgttggtta gaggaggctc 29640

cttggaggac atgaggttta agctgtgacc tgaggatgaa tagatgttgg ccaggtgagg 29700

taccggtatt tgtcagcctt accagtaaaa aagaaaacct attaaaaaaa aaatacacat 29760

acaaagcctc atcagccatg gcttaccaga gaaagtacag cgggcacaca aaccacaagc 29820

tctaaagtca ctctccaacc tctccacaat atatatacac aagccctaaa ctgacgtaat 29880

gggactaaag tgtaaaaaat cccgccaaac ccaacacaca ccccgaaact gcgtcaccag 29940

ggaaaagtac agtttcactt ccgcaatccc aacaagcgtc acttcctctt tctcacggta 30000

cgtcacatcc cattaactta caacgtcatt ttcccacggc cgcgccgccc cttttaaccg 30060

ttaaccccac agccaatcac cacacggccc acacttttta aaatcacctc atttacatat 30120

tggcaccatt ccatctataa ggtatattat tgatgatg 30158

<210> 290

<211> 10

<212> PRT

<213> Artificial Sequence

<220><223> PE-labeled H-2Dq/RNEU420429 (H-2D(q)

<400> 290

Pro Asp Ser Leu Arg Asp Leu Ser Val Phe

1 5 10

<210> 291

<211> 594

<212> PRT

<213> Trichoplusia ni

<400> 291

Met Gly Ser Ser Leu Asp Asp Glu His Ile Leu Ser Ala Leu Leu Gln

1 5 10 15

Ser Asp Asp Glu Leu Val Gly Glu Asp Ser Asp Ser Glu Ile Ser Asp

20 25 30

His Val Ser Glu Asp Asp Val Gln Ser Asp Thr Glu Glu Ala Phe Ile

35 40 45

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

50 55 60

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

65 70 75 80

Arg Ile Leu Thr Leu Pro Gln Arg Thr Ile Arg Gly Lys Asn Lys His

85 90 95

Cys Trp Ser Thr Ser Lys Ser Thr Arg Arg Ser Arg Val Ser Ala Leu

100 105 110

Asn Ile Val Arg Ser Gln Arg Gly Pro Thr Arg Met Cys Arg Asn Ile

115 120 125

Tyr Asp Pro Leu Leu Cys Phe Lys Leu Phe Phe Thr Asp Glu Ile Ile

130 135 140

Ser Glu Ile Val Lys Trp Thr Asn Ala Glu Ile Ser Leu Lys Arg Arg

145 150 155 160

Glu Ser Met Thr Gly Ala Thr Phe Arg Asp Thr Asn Glu Asp Glu Ile

165 170 175

Tyr Ala Phe Phe Gly Ile Leu Val Met Thr Ala Val Arg Lys Asp Asn

180 185 190

His Met Ser Thr Asp Asp Leu Phe Asp Arg Ser Leu Ser Met Val Tyr

195 200 205

Val Ser Val Met Ser Arg Asp Arg Phe Asp Phe Leu Ile Arg Cys Leu

210 215 220

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

225 230 235 240

Phe Thr Pro Val Arg Lys Ile Trp Asp Leu Phe Ile His Gln Cys Ile

245 250 255

Gln Asn Tyr Thr Pro Gly Ala His Leu Thr Ile Asp Glu Gln Leu Leu

260 265 270

Gly Phe Arg Gly Arg Cys Pro Phe Arg Met Tyr Ile Pro Asn Lys Pro

275 280 285

Ser Lys Tyr Gly Ile Lys Ile Leu Met Met Cys Asp Ser Gly Thr Lys

290 295 300

Tyr Met Ile Asn Gly Met Pro Tyr Leu Gly Arg Gly Thr Gln Thr Asn

305 310 315 320

Gly Val Pro Leu Gly Glu Tyr Tyr Val Lys Glu Leu Ser Lys Pro Val

325 330 335

His Gly Ser Cys Arg Asn Ile Thr Cys Asp Asn Trp Phe Thr Ser Ile

340 345 350

Pro Leu Ala Lys Asn Leu Leu Gln Glu Pro Tyr Lys Leu Thr Ile Val

355 360 365

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

370 375 380

Ser Arg Ser Arg Pro Val Gly Thr Ser Met Phe Cys Phe Asp Gly Pro

385 390 395 400

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

405 410 415

Leu Ser Ser Cys Asp Glu Asp Ala Ser Ile Asn Glu Ser Thr Gly Lys

420 425 430

Pro Gln Met Val Met Tyr Tyr Asn Gln Thr Lys Gly Gly Val Asp Thr

435 440 445

Leu Asp Gln Met Cys Ser Val Met Thr Cys Ser Arg Lys Thr Asn Arg

450 455 460

Trp Pro Met Ala Leu Leu Tyr Gly Met Ile Asn Ile Ala Cys Ile Asn

465 470 475 480

Ser Phe Ile Ile Tyr Ser His Asn Val Ser Ser Lys Gly Glu Lys Val

485 490 495

Gln Ser Arg Lys Lys Phe Met Arg Asn Leu Tyr Met Ser Leu Thr Ser

500 505 510

Ser Phe Met Arg Lys Arg Leu Glu Ala Pro Thr Leu Lys Arg Tyr Leu

515 520 525

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

530 535 540

Asp Asp Ser Thr Glu Glu Pro Val Met Lys Lys Arg Thr Tyr Cys Thr

545 550 555 560

Tyr Cys Pro Ser Lys Ile Arg Arg Lys Ala Asn Ala Ser Cys Lys Lys

565 570 575

Cys Lys Lys Val Ile Cys Arg Glu His Asn Ile Asp Met Cys Gln Ser

580 585 590

Cys Phe

<210> 292

<211> 340

<212> PRT

<213> Rana pipiens

<400> 292

Met Pro Arg Pro Lys Glu Ile Gln Glu Gln Leu Arg Lys Lys Val Ile

1 5 10 15

Glu Ile Tyr Gln Ser Gly Lys Gly Tyr Lys Ala Ile Ser Lys Ala Leu

20 25 30

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

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

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

Asn Val Arg Pro Ser Val Arg Val Leu Lys Leu Lys Arg Thr Trp Val
225 230 235 240
Leu Gln Gln Asp Asn Asp Pro Lys His Thr Ser Lys Ser Thr Thr Glu
245 250 255
Trp Leu Lys Lys Asn Lys Met Lys Thr Leu Glu Trp Pro Ser Gln Ser
260 265 270

Pro Asp Leu Asn Pro Ile Glu Met Leu Trp Tyr Asp Leu Lys Lys Ala
275 280 285

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

<210> 293
<211> 636
<212> PRT
<213> Tribolium castaneum
<400> 293

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

Thr Trp Thr Gly Asp Lys Asp Glu Pro Asn Gly Leu Cys Val Ile Cys
65 70 75 80
Glu Gln Val Val Asn Asn Ser Ser Leu Asn Pro Ala Lys Leu Lys Arg
85 90 95
His Leu Asp Thr Lys His Pro Thr Leu Lys Gly Lys Ser Glu Tyr Phe
100 105 110
Lys Arg Lys Cys Asn Glu Leu Asn Gln Lys Lys His Thr Phe Glu Arg
115 120 125

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

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

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

Arg His Ala Leu Ala Val Lys Arg Ile Pro Asn Ala Leu His Glu Val
325 330 335
Leu Asn Asp Ala Val Lys Met Ile Asn Phe Ile Lys Ser Arg Pro Leu
340 345 350
Asn Ala Arg Val Phe Ala Leu Leu Cys Asp Asp Leu Gly Ser Leu His
355 360 365
Lys Asn Leu Leu Leu His Thr Glu Val Arg Trp Leu Ser Arg Gly Lys

370 375 380

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

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

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

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

Asn Leu Cys Asn Asn Lys Gln Ala Gln Lys Ser His

625 630 635

<210> 294

<211> 649

<212> PRT

<213> Oryzias latipes

<400> 294

Met Glu Glu Val Cys Asp Ser Ser Ala Ala Ala Ser Ser Thr Val Gln

1 5 10 15

Asn Gln Pro Gln Asp Gln Glu His Pro Trp Pro Tyr Leu Arg Glu Phe

20 25 30

Phe Ser Leu Ser Gly Val Asn Lys Asp Ser Phe Lys Met Lys Cys Val

35 40 45

Leu Cys Leu Pro Leu Asn Lys Glu Ile Ser Ala Phe Lys Ser Ser Pro

50 55 60

Ser Asn Leu Arg Lys His Ile Glu Arg Met His Pro Asn Tyr Leu Lys

65 70 75 80

Asn Tyr Ser Lys Leu Thr Ala Gln Lys Arg Lys Ile Gly Thr Ser Thr

85 90 95

His Ala Ser Ser Ser Lys Gln Leu Lys Val Asp Ser Val Phe Pro Val

100 105 110

Lys His Val Ser Pro Val Thr Val Asn Lys Ala Ile Leu Arg Tyr Ile

115 120 125

Ile Gln Gly Leu His Pro Phe Ser Thr Val Asp Leu Pro Ser Phe Lys

130 135 140

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

145 150 155 160

Thr Leu Arg Ser Lys Ile Ala Glu Ala Ala Leu Ile Met Lys Gln Lys

165 170 175

Val Thr Ala Ala Met Ser Glu Val Glu Trp Ile Ala Thr Thr Thr Asp

180 185 190

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

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

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

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

435 440 445

Asp Ile Leu Gln Ala Glu Thr Asn Thr Gln Leu Gly Trp Leu Leu Pro

450 455 460

Ser Val His Gln Leu Ser Leu Lys Leu Gln Arg Leu His His Ser Leu

465 470 475 480

Arg Tyr Cys Asp Pro Leu Val Asp Ala Leu Gln Gln Gly Ile Gln Thr

485 490 495

Arg Phe Lys His Met Phe Glu Asp Pro Glu Ile Ile Ala Ala Ala Ile

500 505 510

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

515 520 525

Lys Arg Gly Met Asp Tyr Ile Arg Val His Leu Glu Pro Leu Asp His

530 535 540

Lys Lys Glu Leu Ala Asn Ser Ser Ser Asp Asp Glu Asp Phe Phe Ala

545 550 555 560

Ser Leu Lys Pro Thr Thr His Glu Ala Ser Lys Glu Leu Asp Gly Tyr

565 570 575

Leu Ala Cys Val Ser Asp Thr Arg Glu Ser Leu Leu Thr Phe Pro Ala

580 585 590

Ile Cys Ser Leu Ser Ile Lys Thr Asn Thr Pro Leu Pro Ala Ser Ala

595 600 605

Ala Cys Glu Arg Leu Phe Ser Thr Ala Gly Leu Leu Phe Ser Pro Lys

610 615 620

Arg Ala Arg Leu Asp Thr Asn Asn Phe Glu Asn Gln Leu Leu Leu Lys

625 630 635 640

Leu Asn Leu Arg Phe Tyr Asn Phe Glu

645

<210> 295

<211> 33

<212> DNA

<213> Artificial Sequence

<220><223> IR/DR encoding sequence Sleeping Beauty

<400> 295

acttaagtgt atgtaaactt ccgacttcaa ctg 33

<210> 296

<211> 370

<212> DNA

<213> Artificial Sequence

<220><223> IR/DR and chromosomal sequence Sleeping Beauty

<400> 296

cagtcaactt agtgtatgta aacttctgac ccactggaat tgtgatacag tgaattataa 60

gtgaaataat ctgtctgtaa acaattgttg gaaaaatgac ttgtgtcatg cacaaagtag 120

atgtcctaac tgacttgcca aaactattgt ttgttaacaa gaaatttgtg gagtagttga 180

aaaacgagtt ttaatgactc caacttaagt gtatgtaaac ttccgacttc aactgtaaga 240

atggcccatt catctatagt agcacacaat atttgcattt gtgcgacagt ataagggaca 300

attatgctat caggcatttt tccaaagtga gtaatcgaag tttttatacc tttgtgtgcc 360

atgtttgcta 370

<210> 297

<211> 369

<212> DNA

<213> Artificial Sequence

<220><223> IR/DR and chromosomal sequence Sleeping Beauty

<400> 297

cagtcaactt agtgtatgta aacttctgac ccactggaat tgtgatacag tgaattataa 60

gtgaaataat ctgtctgtaa acaattgttg gaaaaatgac ttgtgtcatg cacaaagtag 120

atgtcctaac tgacttgcca aaactattgt ttgttaacaa gaaatttgtg gagtagttga 180

aaaacgagtt ttaatgactc caacttaagt gtatgtaaac ttccgacttc aactgtacaa 240

gtagaccaaa tatccatata cataaaagaa aaaaatagaa aaaatttcta gtgacagaaa 300

aatgacaaag aacatactgc tttattacta ctattaagat gtttgcttcc attacactca 360

tatgagtca 369

<210> 298

<211> 227

<212> DNA

<213> Artificial Sequence

<220><223> IR/DR of Sleeping Beauty

<400> 298

ttagtgtatg taaacttctg acccactgga attgtgatac agtgaattat aagtgaaata 60

atctgtctgt aaacaattgt tggaaaaatg acttgtgtca tgcacaaagt agatgtccta 120

actgacttgc caaaactatt gtttgttaac aagaaatttg tggagtagtt gaaaaacgag 180

ttttaatgac tccaacttaa gtgtatgtaa acttccgact tcaactg 227

<210> 299

<211> 371

<212> DNA

<213> Artificial Sequence

<220><223> IR/DR and chromosomal sequence Sleeping Beauty

<400> 299

caacttgagt gtatgttaac ttctgaccca ctgggaatgt gatgaaagaa ataaaagctg 60

aaatgaatca ttctctctac tattattctg atatttcaca ttcttaaaat aaagtggtga 120

tcctaactga ccttaagaca gggaatcttt actcggatta aatgtcagga attgtgaaaa 180

agtgagttta aatgtatttg gctaaggtgt atgtaaactt ccgacttcaa ctgtatatcc 240

tccccgttgc accctcttga tgatgctgag atgaacacag atgctcactc cttgagggct 300

ctaagcttat gctgacacag acacaggtgc tcacttctat gaatggccta agatttgagg 360

acatcatgag g 371

<210> 300

<211> 229

<212> DNA

<213> Artificial Sequence

<220><223> IR/DR of Sleeping Beauty

<400> 300

ttgagtgtat gttaacttct gacccactgg gaatgtgatg aaagaaataa aagctgaaat 60

gaatcattct ctctactatt attctgatat ttcacattct taaaataaag tggtgatcct 120

aactgacctt aagacaggga atctttactc ggattaaatg tcaggaattg tgaaaaagtg 180

agtttaaatg tatttggcta aggtgtatgt aaacttccga cttcaactg 229

<210> 301

<211> 1468

<212> PRT

<213> Artificial Sequence

<220><223> ET3

<400> 301

Met Gln Leu Glu Leu Ser Thr Cys Val Phe Leu Cys Leu Leu Pro Leu
1 5 10 15

Gly Phe Ser Ala Ile Arg Arg Tyr Tyr Leu Gly Ala Val Glu Leu Ser
20 25 30

Trp Asp Tyr Arg Gln Ser Glu Leu Leu Arg Glu Leu His Val Asp Thr
35 40 45

Arg Phe Pro Ala Thr Ala Pro Gly Ala Leu Pro Leu Gly Pro Ser Val
50 55 60

Leu Tyr Lys Lys Thr Val Phe Val Glu Phe Thr Asp Gln Leu Phe Ser
65 70 75 80

Val Ala Arg Pro Arg Pro Pro Trp Met Gly Leu Leu Gly Pro Thr Ile
85 90 95

Gln Ala Glu Val Tyr Asp Thr Val Val Val Thr Leu Lys Asn Met Ala
100 105 110

Ser His Pro Val Ser Leu His Ala Val Gly Val Ser Phe Trp Lys Ser
115 120 125

Ser Glu Gly Ala Glu Tyr Glu Asp His Thr Ser Gln Arg Glu Lys Glu
130 135 140

Asp Asp Lys Val Leu Pro Gly Lys Ser Gln Thr Tyr Val Trp Gln Val
145 150 155 160

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

Ser Tyr Leu Ser His Val Asp Leu Val Lys Asp Leu Asn Ser Gly Leu
180 185 190

Ile Gly Ala Leu Leu Val Cys Arg Glu Gly Ser Leu Thr Arg Glu Arg
195 200 205

Thr Gln Asn Leu His Glu Phe Val Leu Leu Phe Ala Val Phe Asp Glu
210 215 220

Gly Lys Ser Trp His Ser Ala Arg Asn Asp Ser Trp Thr Arg Ala Met

225 230 235 240

Asp Pro Ala Pro Ala Arg Ala Gln Pro Ala Met His Thr Val Asn Gly

245 250 255

Tyr Val Asn Arg Ser Leu Pro Gly Leu Ile Gly Cys His Lys Lys Ser

260 265 270

Val Tyr Trp His Val Ile Gly Met Gly Thr Ser Pro Glu Val His Ser

275 280 285

Ile Phe Leu Glu Gly His Thr Phe Leu Val Arg His His Arg Gln Ala

290 295 300

Ser Leu Glu Ile Ser Pro Leu Thr Phe Leu Thr Ala Gln Thr Phe Leu

305 310 315 320

Met Asp Leu Gly Gln Phe Leu Leu Phe Cys His Ile Ser Ser His His

325 330 335

His Gly Gly Met Glu Ala His Val Arg Val Glu Ser Cys Ala Glu Glu

340 345 350

Pro Gln Leu Arg Arg Lys Ala Asp Glu Glu Glu Asp Tyr Asp Asp Asn

355 360 365

Leu Tyr Asp Ser Asp Met Asp Val Val Arg Leu Asp Gly Asp Asp Val

370 375 380

Ser Pro Phe Ile Gln Ile Arg Ser Val Ala Lys Lys His Pro Lys Thr

385 390 395 400

Trp Val His Tyr Ile Ala Ala Glu Glu Glu Asp Trp Asp Tyr Ala Pro

405 410 415

Leu Val Leu Ala Pro Asp Asp Arg Ser Tyr Lys Ser Gln Tyr Leu Asn

420 425 430

Asn Gly Pro Gln Arg Ile Gly Arg Lys Tyr Lys Lys Val Arg Phe Met

435 440 445

Ala Tyr Thr Asp Glu Thr Phe Lys Thr Arg Glu Ala Ile Gln His Glu

450 455 460

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

465 470 475 480

Leu Ile Ile Phe Lys Asn Gln Ala Ser Arg Pro Tyr Asn Ile Tyr Pro
485 490 495
His Gly Ile Thr Asp Val Arg Pro Leu Tyr Ser Arg Arg Leu Pro Lys

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

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

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

690 695 700
Gly Leu Trp Ile Leu Gly Cys His Asn Ser Asp Phe Arg Asn Arg Gly
705 710 715 720
Met Thr Ala Leu Leu Lys Val Ser Ser Cys Asp Lys Asn Thr Gly Asp

725 730 735

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

740 745 750

Asn Asn Ala Ile Glu Pro Arg Ser Phe Ala Gln Asn Ser Arg Pro Pro

755 760 765

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

770 775 780

Ile Ser Leu Pro Thr Phe Gln Pro Glu Glu Asp Lys Met Asp Tyr Asp

785 790 795 800

Asp Ile Phe Ser Thr Glu Thr Lys Gly Glu Asp Phe Asp Ile Tyr Gly

805 810 815

Glu Asp Glu Asn Gln Asp Pro Arg Ser Phe Gln Lys Arg Thr Arg His

820 825 830

Tyr Phe Ile Ala Ala Val Glu Gln Leu Trp Asp Tyr Gly Met Ser Glu

835 840 845

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

850 855 860

Phe Lys Lys Val Val Phe Arg Glu Phe Ala Asp Gly Ser Phe Thr Gln

865 870 875 880

Pro Ser Tyr Arg Gly Glu Leu Asn Lys His Leu Gly Leu Leu Gly Pro

885 890 895

Tyr Ile Arg Ala Glu Val Glu Asp Asn Ile Met Val Thr Phe Lys Asn

900 905 910

Gln Ala Ser Arg Pro Tyr Ser Phe Tyr Ser Ser Leu Ile Ser Tyr Pro

915 920 925

Asp Asp Gln Glu Gln Gly Ala Glu Pro Arg His Asn Phe Val Gln Pro

930 935 940

Asn Glu Thr Arg Thr Tyr Phe Trp Lys Val Gln His His Met Ala Pro

945 950 955 960

Thr Glu Asp Glu Phe Asp Cys Lys Ala Trp Ala Tyr Phe Ser Asp Val

965 970 975

Asp Leu Glu Lys Asp Val His Ser Gly Leu Ile Gly Pro Leu Leu Ile

980 985 990

Cys Arg Ala Asn Thr Leu Asn Ala Ala His Gly Arg Gln Val Thr Val

995 1000 1005

Gln Glu Phe Ala Leu Phe Phe Thr Ile Phe Asp Glu Thr Lys Ser

1010 1015 1020

Trp Tyr Phe Thr Glu Asn Val Glu Arg Asn Cys Arg Ala Pro Cys

1025 1030 1035

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

1040 1045 1050

His Ala Ile Asn Gly Tyr Val Met Asp Thr Leu Pro Gly Leu Val

1055 1060 1065

Met Ala Gln Asn Gln Arg Ile Arg Trp Tyr Leu Leu Ser Met Gly

1070 1075 1080

Ser Asn Glu Asn Ile His Ser Ile His Phe Ser Gly His Val Phe

1085 1090 1095

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

1100 1105 1110

Tyr Pro Gly Val Phe Glu Thr Val Glu Met Leu Pro Ser Lys Val

1115 1120 1125

Gly Ile Trp Arg Ile Glu Cys Leu Ile Gly Glu His Leu Gln Ala

1130 1135 1140

Gly Met Ser Thr Thr Phe Leu Val Tyr Ser Lys Lys Cys Gln Thr

1145 1150 1155

Pro Leu Gly Met Ala Ser Gly His Ile Arg Asp Phe Gln Ile Thr

1160 1165 1170

Ala Ser Gly Gln Tyr Gly Gln Trp Ala Pro Lys Leu Ala Arg Leu

1175 1180 1185

His Tyr Ser Gly Ser Ile Asn Ala Trp Ser Thr Lys Glu Pro Phe

1190 1195 1200

Ser Trp Ile Lys Val Asp Leu Leu Ala Pro Met Ile Ile His Gly

1205 1210 1215

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

1220 1225 1230

Ser Gln Phe Ile Ile Met Tyr Ser Leu Asp Gly Lys Lys Trp Gln

1235 1240 1245

Thr Tyr Arg Gly Asn Ser Thr Gly Thr Leu Met Val Phe Phe Gly

1250 1255 1260

Asn Val Asp Ser Ser Gly Ile Lys His Asn Ile Phe Asn Pro Pro

1265 1270 1275

Ile Ile Ala Arg Tyr Ile Arg Leu His Pro Thr His Tyr Ser Ile

1280 1285 1290

Arg Ser Thr Leu Arg Met Glu Leu Met Gly Cys Asp Leu Asn Ser

1295 1300 1305

Cys Ser Met Pro Leu Gly Met Glu Ser Lys Ala Ile Ser Asp Ala

1310 1315 1320

Gln Ile Thr Ala Ser Ser Tyr Phe Thr Asn Met Phe Ala Thr Trp

1325 1330 1335

Ser Pro Ser Lys Ala Arg Leu His Leu Gln Gly Arg Ser Asn Ala

1340 1345 1350

Trp Arg Pro Gln Val Asn Asn Pro Lys Glu Trp Leu Gln Val Asp

1355 1360 1365

Phe Gln Lys Thr Met Lys Val Thr Gly Val Thr Thr Gln Gly Val

1370 1375 1380

Lys Ser Leu Leu Thr Ser Met Tyr Val Lys Glu Phe Leu Ile Ser

1385 1390 1395

Ser Ser Gln Asp Gly His Gln Trp Thr Leu Phe Phe Gln Asn Gly

1400 1405 1410

Lys Val Lys Val Phe Gln Gly Asn Gln Asp Ser Phe Thr Pro Val

1415 1420 1425

Val Asn Ser Leu Asp Pro Pro Leu Leu Thr Arg Tyr Leu Arg Ile

1430 1435 1440

His Pro Gln Ser Trp Val His Gln Ile Ala Leu Arg Met Glu Val

1445 1450 1455

Leu Gly Cys Glu Ala Gln Asp Leu Tyr Val

1460 1465

<210> 302

<211> 147

<212> PRT

<213> Homo sapiens

<400> 302

Met Val His Leu Thr Pro Glu Glu Lys Ser Ala Val Thr Ala Leu Trp

1 5 10 15

Gly Lys Val Asn Val Asp Glu Val Gly Gly Glu Ala Leu Gly Arg Leu

20 25 30

Leu Val Val Tyr Pro Trp Thr Gln Arg Phe Phe Glu Ser Phe Gly Asp

35 40 45

Leu Ser Thr Pro Asp Ala Val Met Gly Asn Pro Lys Val Lys Ala His

50 55 60

Gly Lys Lys Val Leu Gly Ala Phe Ser Asp Gly Leu Ala His Leu Asp

65 70 75 80

Asn Leu Lys Gly Thr Phe Ala Thr Leu Ser Glu Leu His Cys Asp Lys

85 90 95

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

100 105 110

Cys Val Leu Ala His His Phe Gly Lys Glu Phe Thr Pro Pro Val Gln

115 120 125

Ala Ala Tyr Gln Lys Val Val Ala Gly Val Ala Asn Ala Leu Ala His

130 135 140

Lys Tyr His

145

<210> 303

<211>

147

<212> PRT

<213> Homo sapiens

<400> 303

Met Gly His Phe Thr Glu Glu Asp Lys Ala Thr Ile Thr Ser Leu Trp
1 5 10 15

Gly Lys Val Asn Val Glu Asp Ala Gly Gly Glu Thr Leu Gly Arg Leu
20 25 30

Leu Val Val Tyr Pro Trp Thr Gln Arg Phe Phe Asp Ser Phe Gly Asn
35 40 45

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

50 55 60

Gly Lys Lys Val Leu Thr Ser Leu Gly Asp Ala Thr Lys His Leu Asp
65 70 75 80

Asp Leu Lys Gly Thr Phe Ala Gln Leu Ser Glu Leu His Cys Asp Lys
85 90 95

Leu His Val Asp Pro Glu Asn Phe Lys Leu Leu Gly Asn Val Leu Val
100 105 110

Thr Val Leu Ala Ile His Phe Gly Lys Glu Phe Thr Pro Glu Val Gln

115 120 125

Ala Ser Trp Gln Lys Met Val Thr Ala Val Ala Ser Ala Leu Ser Ser
130 135 140

Arg Tyr His
145

<210> 304

<211> 29230

<212> DNA

<213> Artificial Sequence

<220><223> HDAd35-T4-Ef1a-mgmt

<220><221> misc_feature

<222> (1)..(481)

<223> Ad35 5'end

<220><221> misc_feature

<222> (14126)..(14159)

<223> FRT (Complementary)

<220><221> misc_feature

<222

> (14220)..(14463)

<223> pT4 LIR

<220><221> misc_feature

<222> (14491)..(15825)

<223> EF1a

<220><221> misc_feature

<222> (15843)..(16466)

<223> mgmt

<220><221> misc_feature

<222> (16484)..(16705)

<223> pA

<220><221> misc_feature

<222> (16735)..(17000)

<223> pT4 RIR

<220><221> misc_feature

<222> (17107)..(17140)

<223> FRT (Complementary)

<220><221> misc_feature

<222> (28823)..(29230)

<223> Ad35 3'end

<400> 304

catcatcaat aatatacctt atagatggaa tggtgccaat atgtaaatga ggtgatttta 60

aaaagtgtgg gccgtgtggt gattggctgt ggggttaacg gttaaaaggg gcggcgcggc 120

cgtgggaaaa tgacgtttta tgggggtgga gttttttttgc aagttgtcgc gggaaatgtt 180

acgcataaaa aggcttcttt tctcacggaa ctacttagtt ttcccacggt atttaacagg 240

aaatgaggta gttttgaccg gatgcaagtg aaaattgctg attttcgcgc gaaaactgaa 300

tgaggaagtg ttttttctgaa taatgtggta tttatggcag ggtggagtat ttgttcaggg 360

ccaggtagac tttgacccat tacgtggagg tttcgattac cgtgtttttt acctgaattt 420

ccgcgtaccg tgtcaaagtc ttctgttttt acgtaggtgt cagctgatcg ctagggtatt 480

taccggtatt caaggattac atgagcttag aaatgtaatt agcatagtgt gtggcatagt 540

gtagatacca aataaatatg atctctcctt ctactcttga aaatgcaaac acattcttgg 600

tggtcctaaa atagcctgta acatggttta ctcagcagca tttgctattc aaggcagatc 660

tgcctttagt cattggctgc gctcctgaac agctgtgtga aaggctaact tttgtaaacc 720

aaatcaaaat aaaatgcagc aaaaatttgt cactgaaagg aaatcctcag tatatccttt 780

tatgaaatga aagatccctc atccaaactt aacttttta aaagtgcgca tttggagata 840

tagccctttc ttatgaatcc taattcaatt ttggccataa acacacgttg atgttcccca 900

ccccaaagca catagcaaca agagtaggtt ctatattgaa aataatgaca atttaaaaac 960

atgtacttat ttcactgtat gtggacagtg tctatgattg catcatgaag tgtcatataa 1020

ccatgtacgt gtacatgaga gagagataga gagagaagtg gtagggtggt ggtggtagag 1080

gggatggcga tagtaatcat ggtaatggta gaggtgatgg aggtggtaat gacggaggta 1140

agggtggtag tgatgatggt ggtggtggta atggtggtgg atgtggtggt ggcaattggg 1200

atggtgggat ggtggtagcc atggtgatgg tggtaatggt gttgatttaa agggtggtgg 1260

tagtgaaggt gagggtagtg gtggtggagg tggtggtgct ggtagcaata gtgatggtgg 1320

tgatggtgtt gatgagggtg ttgggatcag ggtgagttcc cacagtatat ttcattcttg 1380

ttgtaccact ctgtcaacag caccactgac tgggacagag gaagaaggca cactctgaat 1440

gtgttattaa cagaaacctc aaaacagtct gtctccttgt agtcattcaa aattatcttt 1500

ttcttacctg gaaaactgaa actgaattac cgggaaaaac acaggagatt tttgtttgtt 1560

aatatgctgc caataaagta attttatgtc aaatttaact acaggaaagg gcaaggcatt 1620

ttctaagttc cttagatgtc atgtggctaa aaaaaacaaa aggatggaca gcagttagat 1680

actgtacact tagctgtttg aagccatata ttcagaaagc agatgttggg agttggtgtt 1740

tgaggactga tttcctggag gtattttata taggccaagt tcattgttct aaactctaag 1800

ggcttgactt gagggaggaa aagaggcaag aacatgttta gttttgctga cagcatcaca 1860

tgggcagccc taaggctaga caactttagg gcctgaagct tattctagga aagaagcacc 1920

tacagagtgg cactgggctc ccctccacta tagagatgaa gtcatatgac agtaaagggc 1980

aggcagggct gcctaggggg cccagaactg acacttccat tagaatgagc acaggccagg 2040

gagagaagtg gggaaccaga gagaaggagc tggaattcta gtaggacaaa cggtaagtga 2100

acaacaagaa caagttaaga gtgtgtgcag tattctttca aagactgaaa aaatagtgat 2160

gtgatagaat ggcaggtggc tctgagcagg ccaggagaag gactgggggc agagcatccc 2220

aggcaggagg gcagcaagtg ggaaggccct ggggtggggc ttttggactg ttccagtgac 2280

gggcaggcag ccagtgtgcc tgtcacacaa tgcaccaggg aagtagtcgt gaatttgcag 2340

agggtcttgc aggctatggg aaagggattg gattgtattt tgtttgtagg gaagccatcg 2400

ggggacttaa gcagaggaag gattggcttc atctctttga aaaagttctc tctggatgct 2460

gatgggagga gaaatggaag gaaaagaaac acttttaggg gcaagaactt ttgagaaggg 2520

tggaattggg agtgtggagt tggggccagc tttggcacag gaggggaagc taaacacgtg 2580

gccgcatgag ggcctgtaat tctacctgaa atgggtacca tttgttaggg taaacaaatg 2640

aaccaaatgc ccagtgatac agaccaagtg ttggcaaact tcttctgtga tggcccaggt 2700

agtaaatgtc tcaggcttcg caggccatgt ggtctctgtt gaagctctgt gtagtagaca 2760

atatgttaat gactgggcgt gactgtgtgc taataaaagt ttatttacaa aaacagcccg 2820

tgggctggat ttagctcaca ggctgtagtt tgccaacctc tgacctagag catgaactga 2880

gcatcttctt ggagggaaat aagttctttc caagttgccc tcctcacatt gcagggggcc 2940

atgtaggccc attattcaca gaagagtggg tgggcaacct ttctggagca gaaaaacgta 3000

aagatttctt ccgtagtgca agtaaggtga ccatttctaa accgtgcaag tgatccagca 3060

gtcccaaaag ttgtttcact tctcattgtg cgcccgttct caggtgctcc gaagcttcca 3120

gtcctttgta gggacatgga tgaaattgga aatcatcatt ctcagtaaac tatcgcaaga 3180

acaaaaaacc aaacaccgca tattctcact cataggtggg aattgaacaa tgagatcaca 3240

tggacacggg aaggggaaca tcacattctg gggactgttg tggggtgggg ggagggggga 3300

gggatagcat tgggagatat acctaatgct agatgacaag ttagtgggtg cagcgcacca 3360

gtgtggcaca tgtatacata tgtaactaac ctgcacaatg tgcacatgta ccctaaaact 3420

taaagtataa taataaaaga aaaaaaaaaa gagaggagag aaacatcatc ccctccagga 3480

tacccttggg ccttgttctt atagtcttgt acattgttga acaatttgca tgggctagtg 3540

gattaaagca caccctccac cctcaggccc tcaagggtct ctatgataat acagtctcac 3600

cttctaccct ttccatcacc atcctaggtg ctatggccaa ccttgaggct gccatgttag 3660

gtctatgcat ttcccacctc caccacataa ctctctgaag gccaggtagt ttcctattca 3720

tcttggtaac cccaaagcct cgtgacaggg ctcagctggc atctgcggat gtgaatgaac 3780

cattggagaa aatggtactc tgcaaataac tctgttattt tcccatttcc tgtgtaaggc 3840

ctagagacaa tgacttttta attgcacccc ttcccctctg tatgacactg gccttctctt 3900

gtgtccagca atgtgggtgg cctagatgat ttctaaggga cttctggcca agatgaacag 3960

cagctgcatc ttactgagca tttactatgt gccatatact cagccacagc tctaggggca 4020

tagaagcagg agctctcagg gtcagggcag tgagtgagca agcgagcacc tatgccagcc 4080

ctgcctctgg atggggactt gagagggtga tggaagcctg cagcactgga gggaggcaga 4140

caaagacagg cctgtgctga gggggcccgg agcaagagag agggaggcaa tgacagcaga 4200

gacatgcctg cgccttgggt ttgagtgccc agtggtcaaa tccacttccc tgtggctgat 4260

gcttgccttt ctaactttgg aatttagggg ttggagatct ggtgagaagg taggagggag 4320
atgaggagga gaagggaaag gcaggaagga aggggaggga aaggaaaagc aaaaggggag 4380
gaggaaggtt tccaacaaat tattctatat caactgcgga aatcaaaatt tgttgcccaa 4440
atcttagaag ctcatgtccc tcctccccag aagtctggaa tgcagcactc caggggtagc 4500
ttataaccca aatatctatc tgtaaaaaga gaaacattgg gctttcgagc tgtggattct 4560
cagtaaaagc aagaggcctc agcctacaca ggccagccca gagtttgagg aaccccaggc 4620
ccacacccac agggctggcc cctgggtctg catactccct agaaatgtgc acacttctga 4680

gcctcaactc tgtcctggag tctaacagca tccctctcct tcctggggca gttccacctc 4740
cagaaacctg ttaccttggg ccttatgtca aggaaactgt gggaaagagc taggcaggaa 4800
tgcagatgag gccagcatgg gctcctaaaa gtttagaaat aggcagtgtc atgctcccag 4860
gtgcctgcat aaaccagctg aaaaatggag ctcccctcac cagcactctc ccttcaaaca 4920
gactgtgatt tgcaggtcac tggtttacca agccaggcta cccaggcagg acccagatgc 4980
caagcccagt ggtgtcctgc aagctgagca gtgctcagtt cttgcaaaaa aaggtctgtg 5040
tgaaggcaag gcctctgcct ggcttctcac cccagttggg tgtctggaac aggaaggagc 5100

ccttactgca gaaaaaggag gagggagcaa agggagcgaa cagctgcgtg ctccatgggg 5160
aggatcccca aagtagaaag gcgcatacac actgcagccc ttgacccaga atgctcacag 5220
ctacattaca gattcaggtc tcctcagtgt agtggggctg ctgatgagac tgtggcatcc 5280
tcaggggtca ggacacacat tttccatcac tcttctgatg gcaaaaaacc tctgagccaa 5340
tgccaacctc tgatcattaa aaaaaagtgc tcacagcagt gtgtggttta ggatcatgcc 5400
ctgtgtggtt tggaacacgt gcacaaccac accttgttca tcaccatccc agaaaccctg 5460
acgcaggcaa agagcagagt tattaaccct actttactga tgtggatact gaggcccaga 5520

ggctcatgca agttatcaat aagtggcagg gacagttgcc tctagattaa ctagccccta 5580
ggatcacctg ggtcttggaa ggggacccat aaacatgagc tcccctctct tggggccaga 5640
tttgcacctg tgccgcgcct tcagcctgca tgaagtaggg gctgctggca aagactcaaa 5700
gctgtaaatc tgggttttct cttgaggctt ctaagggagc tgtttcgaca actcactctg 5760
ttcccagctg gctgcccctg catagggttt taaagcagcc tagctttctg ccaggcttgg 5820
cagtggacaa cgctggtcag aacatcccag agagctacca gaatgaagta agtttgcttc 5880
tactctttac ctgtttatgg gctgtctctg ccactggaat gaaaggcact gagaacagtg 5940

cctggcctgc agaaggccct ggaaatacct gagctcctaa tctgggaata ggagtaggaa 6000

gagctttgga ggcagggcac ctgagtttga gatctacaac ttcctgcctg tgtgacattg 6060

ggaaagtctc catcctttct gagcctcagt ctccaccctg gggaagtgga aatatcaatc 6120

tctgtgacac agaagcaaat gagcgaatgt gcacaaagta ccttgcacaa gagagacgct 6180

caaacacttg cctccaggtt tcaccgagaa ctacagagta agatagattt gttcccagtg 6240

gaggaagcct gggaataatt tgcccctaga ctatgaattc ctggggctca agatcgagca 6300

cagggccagg cacacagaag ggaccctgga aatgtggcag gaggccagag atagacaggc 6360

ccttagagct catacccatg ccctctgacc tcaagaagaa agaaacctgc tcaaaatctc 6420

acaaagagct tgttccaacc ctgaatcgag tctgaggact ccttcctgag tccagcactt 6480

tttctgcaag aagtatatgc ctccaaagct gatgggcgca aatcttgaac cccgtcacat 6540

aaacacaaag ggaggaggtg actagagctc ctcctactgg atatgtctaa ggtcaccagt 6600

ctaaagaaaa gggatggata gaatgaggcc agtatttttg cagccatcca aatgtccaca 6660

tacgctgtta cactgagggc tcctctctcc cccgtcttca gccctacttg catttagagg 6720

tgagaaagat atgggctgag gggttgtttt tcatcgtatt gtagatggaa agcacactgc 6780

ccttggggcc atccaaatgt ggaccttgat gtagcacccc accttctgga tggccatcct 6840

tctgaaagtc actgaatttc tcagacttta ttctctttat ccataaagaa ggagaataat 6900

aataatcccc ccaccctgcc caaccactga ctggttggga agctcagaag aaatactggg 6960

cacggcatcc cattgtaatc tatagagtga gtcgcttctt aatattaaat ggctgaacac 7020

agaagatgtg caaaaagtac tgtgtcccct tcctcctcca actgaacatt tcatgccctt 7080

tgcaccctca ttttgtctag gagctgcctt atgaagggaa taggtacctg ctccgagctg 7140

gaggaatctt tgccacttat ggtggggtat ggactgagac agagatggca tgtgacatgc 7200

gcactgagtc tcaactccat gcaggctctg gagcactctc aaattggagt actaatgcct 7260

tttaaattct cacactagca atcctttgac ctactgatct agggatctag ggaaagaatc 7320

gtgatcttaa cttcaaaggg aaggacaaaa tgttctgcct cctgttaaaa ctccatacac 7380

taagtgcaga gactggatgc cttattaacc ttgggtagat gcccaaatgt tcaaaaggtc 7440

aaactcttct gttccccaga tcgccagagt cattaaccag tcacactatt aaatgaatga 7500

acagatgctg aaaaggtact tgcattactg agatttctta tggtgatggc ccctgcctga 7560

tatgtattca gcattttgta gttttcaatg tgcattagag tatagtggtg atgacattgg 7620

cctctgagtt tgccacttct tatatctgtg actttggtca aattgcttaa tctctctgag 7680

tctcggtttc ctggagataa taatagcttc ttcttcccag ggttatcatg aggattacag 7740

gagataatgc cccaaaaatg cttagtaaag tgcctagcac ctagtcaatg ctgaattaaa 7800

ggtggttatt cttacttttc gttcatttga actttgttct cagggagggc aaaggataga 7860

caaagcccca tagctagtga ggagtagctg caagactaga acccaggtgt tctgagccct 7920

agtcttaggc caagaacaac tgttacgtga gatgcacgtt ttccttcaag ggagctcaca 7980

attatttcca tgtaaattca aggactgcta aaagagaact ctcctctggg actgatatca 8040

ttttatttca agattgattt gaaacatgtt ttttgtttgt ttgtttgttt tctaggaaag 8100

aacaagagaa ccagttaagc tgaatgcctg aagcaaatcc ctgttagcga tgttttcagg 8160

atgagggaga gtggtgcaag aaacgtgctt ccagatgcac atggtttcct gggactaggg 8220

ttcagggtgt catccctggg tgttattaag tgtcagaagg agagcaaaca agggaaacat 8280

ctgagatcca gctaaggcta caccctggaa atgcaagccc agctcttgca aaggacctcc 8340

tttggccact caccttccag gccttacaat aacttgtttg gactgcaggt ttcttggtgg 8400

actcacaggc cattctgctt ttatttggtc aacctcagtt cacaagcacc cagatgctga 8460

gatcctcagc atgtgcagca gagtttcata ttagcactgg gtacctttct gaggctacag 8520

ggataccgta cagcagcacc tgtcacgtcc agccaaagga gtgggctctc tcaatgtcat 8580

ccaatgctgt ttcaactgtg aagaagacca tctgagagag ttgcttttgg aggctgaggc 8640

aaatttttaa aattctttgt tctcctcaac tggggtgaat tcttggtctt ctaggacagc 8700

ttgaagtttt agaaagagtc aagccactca gaaccaacag agaactcttt cagagaacaa 8760

ggtgtggcat agaggaggca gagggctgat cttgatcaaa tccaaagtgt gactctaaag 8820

caatgaatgt gaatttttgg caaagcttac aaagggctct aaaggccatc tgcaaagaga 8880

agccaagcct gatcgatgaa tcactagtgc ggccggatat cgatcggcac gctgttgatt 8940

ttctcatagt aaggaacagt gggccctttc agtcccactt ctgtagtctg tggtactaca 9000

aatggtgagc ccatgatgtt gccattcata gggttattct ccagcagtaa tgactggcca 9060

gccactccca tagccgcggg gctaggattt attgtcaatg gagggacctg cagttctgca 9120

caagcagtac taggatgagc acctgggccc attgcaaggg tgacatcttc aaggcaaggc 9180

ctcttaattt tattagggta gcccccatca gccatgtctg gaaactggaa gtggtcttct 9240

tcttgtctcc tcttaacagt tccctgtgaa tggaagagaa gagaggagga gaagagagga 9300

gaggagaagg gaagagaggt gacacacaca cacacacaca cacacacaca cacagagaga 9360

gagagagaga gacagagaga aagagagaga gagagagaga ggaatttta taaaggtttg 9420

gcacattaaa gctaatgaac aggaaatgtg catgataaaa cagacctctc agtttaaaga 9480

cttatagttg tgaaaactat aaaatacagc ctgtctttgg aaccatagtg cttatttatt 9540

cattattatg tttcatctaa actgtctaat tacatttcaa ataaggcatt atgttgtctg 9600

tatactaaaa cgggatagaa cgttattcaa agggtaatct gcccacttca aggagagttc 9660

aacaaaacta tgcagaagtc actaaatgaa ccatgctgcc aaaggcaggc attggagaga 9720

aaactagaag tagctaaata gttttaattc tttcctgtct acagacacat agattttaac 9780
gaaggaatac catagtatag aattgaactt ttaggctgcc ttctagtctt ggttaaatgc 9840
atcaggctgc agtggtaaaa ttgaatacaa cagagccctt acaggaaaga agtagatctg 9900
gatgtgtttt cttggggagc tgtttaaaat actgtttttg ggaaagcaca agtttcagaa 9960
cagtcattgt aggcatcgta ttcattgttc catttatttt tacacacaca cacacacaca 10020
cacacacaca cactctcaca cattgctatg tgtacacaaa aataatttgg aagaacctat 10080
acccaacaat ttggagtggt catttatttg ggatgactgg caattccctt tctattctct 10140

tcatttctgc ttgtttgtct ttaacgagaa cgactcataa tccaaaaatt taaaaaagta 10200
taaagttatc taaataagaa attttcctct gaagatgcat cctcaggttg gggagatatt 10260
aaacaatgag aaaaggcccc aatctgggat ctgaaccttg ggggagctgc ccatcattta 10320
tagaagcaca gcctttggga acaaagcaaa gtcactagca atgtgagact tcctactctt 10380
catggcttca tacagtcatc catcgctgtt gtgttaatga ccatgacctg tatgttagca 10440
ggtaaatggg aaaggaagtg ggggcaaagg agtatgtgca ggaatgatca aaataaggaa 10500
aggaagagag ggatctggaa atcacctgaa tgccgatagg tgaacaggta gaattctttt 10560

aaagcttccc ccacccggta ccccccaaat aacccctttc cagctttgga agtttcacta 10620
ggacatacag tgctcatcct ctgatgtcac cttaagtttg gctcttctgg tttgatgagc 10680
ttgtagccca ctaggagctc aaggcatgca tggggccact tgccagcacg atgaggggca 10740
tgactgtcat ggccaagtga acatcaaagc agatccccag ggctgtatgt ctcaggcctt 10800
ggtgcacatc agaatcactt agaaacatcc acattcctgg gccctcccac cacaaactga 10860
cagcttcatc cagggtgtgg cccaggcatc gggagttttt ccaacagctc catggctgat 10920
tctcaacaga aaaccactgg cccagagcaa gggtggaggc agcgtggcat agggctctga 10980

ccttggcctt gccactgaac ctctcagagc cccagtttct ttatgtgtaa aatgagtgta 11040
attatagttc ttttctcatg aaggtgctct gactattaag tgaaacgggg cacattgtat 11100
gacacctaat agctcctcac taactggtac ccggcattat aaagggcagg tatggaaggg 11160
ttctgggagt ccaataccct tcttaaagac agagaggtct ctgagaccca gagaggggca 11220
ggccttaccc agagttgctc agccagaggg caacaaggcc caggtcagat gcagggcccc 11280
tccaccacca ctcagctgcc tccagaccca ctgccttcgc catgttgttg gtaggacact 11340
gcatcgcccc cacagaaggg gcttgccaac ttgagtgaga ggacttgcac acttctttga 11400

ctttttcttt gagatgccca caatctgaac aagggcactt caagggacag ctctgtcacc 11460

aaactcatct gaggcctgaa taccatgggt caggcaggaa tgggttggag aggtgtagag 11520

caggcacaat aagagggctg aggcccatgc agtcatcagt gcccactttc ccaggagtct 11580

gactgggcac agcacccata gtgtccctga gctggtccat ggagcagctc actaactgtt 11640

tggcccacag caggtgctca gtaaatggca gttgaacgaa tcaatggaca aaggaacata 11700

aattacccaa cacacaggga gctcagccat ttactcaatc cattatggag taacctacaa 11760

acaagccact gggtcccaaa ctgaaattgt gtctcttcta cattctccca aagaatccaa 11820

taggttaaaa atagaaatgt atgaaataga tcaatcaggg atgattgcat gtggatttga 11880

cataaggatc ccctgcaggg agtctgagct ggcaacagtc aggcccaaag tgctgtccat 11940

gatgtctcga actgcaagac agttttaaca atggcgaagc aatgcagaac caggcaggcc 12000

aaggaggggg tgggggttgg ggaaaggaag ggagggaagg ggctgtgagg ggcaatggtc 12060

tggcatccct gccacgtgag cctctgaaat ttgctggcag cttctatggg ctcccagagc 12120

tttcacttaa ttgttggtct gccactaacc tgctgggagt aaggtgcagg gatggaggag 12180

gcagggcatg accaccagac actaaaggta ccagctgggg ccactggcaa agggaaggag 12240

gctgcacctc tcctacatga gagcccgtat acacacacct tttccagcac tcatcaactg 12300

catcccaagc aaatggtccc tgatcaattc caattctaga aaccaactga ctactcaata 12360

acaaagtaga tcccagcagg ccgccactgc tggagcggat gccacttttg ctatgccaag 12420

tctgtggctg gacagctgct ggcatgtaca ctcactgact ttcataagga tgcctaataa 12480

agggggcagg ctcacctggc ttttctcagg ggtggggttt ggggtgccga tagaggctgc 12540

tgttttggca gagtggcaag ctgcaagcct cttctgagct ttcatttttc aatggacttc 12600

agtgagaatt cactttgtca gaggccatgc agctccatgt tttggatttc atggaatgag 12660

ctttcaacag tgagcctgaa gtgccctggc tgaacagcaa gaacaccagc caaccctaaa 12720

caaggccgag gagaggcggc tgtgtttaca cggaaggctc agccttgctg taatagcgtc 12780

tgccttcacc agacatcagt gaggcgtgga aatctattat ccagttaatt ttgcccctag 12840

ataaagactt gctttcgtgt cttctctttc acagtcccat gatctgttac tcatctcaac 12900

tgcgagaagt tggctgggct ttcccctgtg cccagtgcca cactcgtgcc ttcactgggt 12960

cacctgtgcc tgtggctgat gccgctgagg ttttgcctgc ccagactggg tgtttctgac 13020

taaatcccac agccaccatt ttagatcaag ggcaggagat agctcactgc tccggaatga 13080

cctcccctcc cagaatcctg gtaggggcgg aaggtcccca accaagctcc cagccctttc 13140

taaatgaatc tccctgcttc acccatgtgc ttttctccag tctctgcggt cttgatgaca 13200

gcagggtatt agtcctagct gtcccacagc tcctacttct ttcaggcctc tccctgtgac 13260

aatcagtagc cactggcagg atttcctcag agcatatctc gatttgcttt cagacaatta 13320

gttaaaagga cactggaccc cagacgtccc aactcccagc cagagccctc acaggcccgg 13380

cctttggtgg tgaggaaggg ggagggagtg agtgacagtg ccctggcatc ttttagaaac 13440

gaattccttt ctctccatac ataaatgcct gcagagtccc atttcagaat ccggcagaca 13500

aagccaccaa tgtgatcccc atgaccttat aaacattcat taaaatgcat ttcaaggcat 13560

gtgatggcct ccccaccccc tagataatga gaaaacaaag gtttctcttc tgatagagac 13620

aagttcagct ctgaagtcaa cattatttct ggttctgtct gaacaatgac atatggcaac 13680

tcttcccttt ctatagttct agtccagaat gacaaaaaag gggaaaaatt tcttagagaa 13740

ggtagagatt atacgaatac agtccatgaa atgagcataa ggagaataaa gaatataact 13800

tatccaaaga agtctggcag gctgttataa atgcttgatt ttggacactg tagttggagg 13860

tttaacatgg acaccaataa aaaggtcagc aaagggtatg cactgttcct attgggcaag 13920

aagataggag gtcaaaggta accaggaaag ataaactcag ggagacttat tttccctcca 13980

gagggcactg ggcttgtagg ccctgggcaa aattgtcaaa aaggtgaaaa tcgcctgtgg 14040

tttatttagt ctgctctttc ttcactagtg cctcaccagt tcagttcagg ccaatttgct 14100

agaaggtagc gaacgatcga ccggtgaagt tcctatactt tctagagaat aggaacttcg 14160

gaataggaac ttctacctag atgcatgctc agagcggccc ctagctagcg tttaaaacct 14220

acagttgaag tcggaagttt acatacactt aagttggagt cattaaaact cgtttttcaa 14280

ctactccaca aatttcttgt taacaaacaa tagttttggc aagtcagtta ggacatctac 14340

tttgtgcatg acacaagtca tttttccaac aattgtttac agacagatta tttcacttat 14400

aattcactgt atcacaattc cagtgggtca gaagtgtaca tacacgcgct tgactgtgcc 14460

tttaagcttt taattaatgg atcactagtt gagtaattca tacaaaagga ctcgcccctg 14520

ccttggggaa tcccagggac cgtcgttaaa ctcccactaa cgtagaaccc agagatcgct 14580

gcgttcccgc cccctcaccc gcccgctctc gtcatcactg aggtggagaa gagcatgcgt 14640

gaggctccgg tgcccgtcag tgggcagagc gcacatcgcc cacagtcccc gagaagttgg 14700

ggggaggggt cggcaattga accggtgcct agagaaggtg gcgcggggta aactgggaaa 14760

gtgatgtcgt gtactggctc cgccttttтc ccgagggtgg gggagaaccg tatataagtg 14820

cagtagtcgc cgtgaacgtt ctttttcgca acgggtttgc cgccagaaca caggtaagtg 14880

ccgtgtgtgg ttcccgcggg cctggcctct ttacgggtta tggcccttgc gtgccttgaa 14940

ttacttccac gcccctggct gcagtacgtg attcttgatc ccgagcttcg ggttggaagt 15000

gggtgggaga gttcgaggcc ttgcgcttaa ggagcccctt cgcctcgtgc ttgagttgag 15060

gcctggcctg ggcgctgggg ccgccgcgtg cgaatctggt ggcaccttcg cgcctgtctc 15120

gctgctttcg ataagtctct agccatttaa aatttttgat gacctgctgc gacgcttttt 15180

ttctggcaag atagtcttgt aaatgcgggc caagatctgc acactggtat ttcggttttt 15240
ggggccgcgg gcggcgacgg ggcccgtgcg tcccagcgca catgttcggc gaggcggggc 15300
ctgcgagcgc ggccaccgag aatcggacgg gggtagtctc aagctcgccg gcctgctctg 15360
gtgcctggcc tcgcgccgcc gtgtatcgcc ccgccctggg cggcaaggct ggcccggtcg 15420
gcaccagttg cgtgagcgga aagatggccg cttcccggcc ctgctgcagg gagctcaaaa 15480
tggaggacgc ggcgctcggg agagcgggcg ggtgagtcac ccacacaaag gaaaagggcc 15540
tttccgtcct cagccgtcgc ttcatgtgac tccacggagt accgggcgcc gtccaggcac 15600

ctcgattagt tctcgagctt ttggagtacg tcgtctttag gttgggggga ggggttttat 15660
gcgatggagt ttccccacac tgagtgggtg gagactgaag ttaggccagc ttggcacttg 15720
atgtaattct ccttggaatt tgccctttt gagtttggat cttggttcat tctcaagcct 15780
cagacagtgg ttcaaagttt ttttcttcca tttcaggtgt cgtgagctag acggtcgcca 15840
ccatggacaa ggattgtgaa atgaaacgca ccacactgga cagccctttg gggaagctgg 15900
agctgtctgg ttgtgagcag ggtctgcacg aaataaagct cctgggcaag gggacgtctg 15960
cagctgatgc cgtggaggtc ccagcccccg ctgcggttct cggaggtccg gagcccctga 16020

tgcagtgcac agcctggctg aatgcctatt tccaccagcc cgaggctatc gaagagttcc 16080
ccgtgccagc gcttcaccat cccgttttcc agcaagagtc gttcacgcgt caggtgttat 16140
ggaagctgct taaggttgtg aaattcggag aagtgatttc ttaccagcaa ttggccgccc 16200
tggccggcaa ccccaaagcc gcgcgagcag tgggaggcgc catgagaggc aatcctgtca 16260
agatcctcat cccgtgccac agagtggtct gcagcagcgg agccgtgggc aactactccg 16320
gagggctagc cgtgaaggaa tggcttctgg cccatgaagg ccaccggttg gggaagccag 16380
gcttgggagg gagctcaggt ctggcagggg cctggctcaa gggagcggga gctacctcgg 16440

gctccccgcc tgctggccga aactgaagcg gccgctcttc gagcagacat gataagatac 16500
attgatgagt ttggacaaac cacaactaga atgcagtgaa aaaaatgctt tatttgtgaa 16560
atttgtgatg ctattgcttt atttgtaacc attataagct gcaataaaca agttaacaac 16620
aacaattgca ttcattttat gtttcaggtt caggggggagg tgtgggaggt tttttaaagc 16680
aagtaaaacc tctacaaatg tggtaaaact agtaatcgat ttaattaaag atctttaaac 16740
aatttaaagg caatgctacc aaatactaag cgcgtgtatg tacacttctg acccactggg 16800
aatgtgatga aagaaataaa agctgaaatg aatcattctc tctactatta ttctgatatt 16860

tcacattctt aaaataaagt ggtgatccta actgacctta agacagggaa tctttactcg 16920

gattaaatgt caggaattgt gaaaaagtga gtttaaatgt atttggctaa ggtgtatgta 16980

aacttccgac ttcaactgta gtttaaaacg ggcccgtagt ctagggccgc cagtgtgatg 17040

gagttcggct tcaggtacag cacactggcg gccgttacta ggtagctaga gccttcagac 17100

tctagggaag ttcctatact ttctagagaa taggaacttc ggaataggaa cttcacccat 17160

ggcgatcgct agcctctaac tcctagaccg tcagaactgc tgggcccttc aagacgggct 17220

gctcacaccc actcatgtta agcctggtga ggcctgtact ctgttttcac aggaagaaat 17280

cctcacccag tcttccccaa acacattccc aggttgtgtc attagtggga tagagatgat 17340

tattgtgggg agaagagaaa catctggatg gatttggtga ggttgatcta tagaggaagt 17400

aggtgctgcc tgaggtagct gtaatagaag ctaaaggtca aaggagaggg ccctgtccca 17460

atccagatga ctccacttct gctggaccca ggttcacaag cttaatctac atttcaccta 17520

aatttggcta acaagcccaa aatcacacag gcaaagggag aagtggaggc agaaccgagg 17580

ttggaggcca ccagggccac cgggcagaga tcatttaagc ccaaccttct cacttctccc 17640

tgggctctgc ctctcttaaa ggaccttgtg gtgtgacctc ttgtaggtcc ctttcacact 17700

cggggcctca gtttccccac tgtaaagtga atgggtccca gctttggtaa gcttatgctt 17760

acctgatgct ttcttcctgg gctgctcttg tagagaaaag ataaatcttc ttcctccatc 17820

cacgagggct cctttccctg ggggtgagag taggctgagg agagccactt gcacacaccc 17880

ttaaagaaag tattacctgc accagctcag tgagaggcac agatcagact gttacttgaa 17940

tcaaattatg agcctcccca aatatatcta tgacatttaa ataggggatt acttgaacat 18000

agactgtggg atccggtgtg gagtgcggga gactagcaaa gtgaatcctg agagtagcag 18060

gtctgcacct gttggatcga gaaaggcggc ctacaattct ggtcaaatga gctgtgctta 18120

ttgacatatt ctattagaga gtactaccag gtcaccagtc accagaaagg ctgccagctc 18180

tccaaccacc tccagggaac tatcctgaat ggggccttaa caagtctaag agagggttgg 18240

tttgggtccc aagccaatat ttgctctgct ttatgtcagt catatggaac ccaaaccaac 18300

cctctcctat gtgcctcacc agtcggtgca gggatcccaa tttcaagttt ggtttttat 18360

ggtcaaagtc cagcatagat taaatgaagg ggtgtgatga tggtgttaaa agagaactcc 18420

agaccagttt aactcttgga cacacatccc atctcaccat ggtgcttcca accttccaga 18480

gatgatgggc tcctattttc tgatgacaaa gccctccaca ggattgctgc ctggccatca 18540

gggagtgcct ctgtaactga ggctgagatc ccactttcag tcctccagct gtggcccatc 18600

cctgctccgc ccaccgggta tggcctgtcc taggctctta ggtatggctg cattgtgaaa 18660

tgatggctac agagctggca tctcctgtag tctggttcat ctagtgcact acctcatagt 18720

taaaagaaat ctgtttaagc cactgagggt ggctcctagt gccaactcca agaacaggaa 18780

gcttcccttt tttgggagga ggggcagatg gtaacatgga tcgtccaggt caatgggagc 18840

agggcaacca cagtaagtac tggacaacaa cacaaaactc catgtgtggc ttccatcgag 18900

tccctctcca attggtttgg tcttctccgt cccatgcagc actttagcaa ggggcctggc 18960

tgaaggctat gaattgtgtg gagcctcctc attgcagtct ccaaccatct gatgctggga 19020

aaatgtcacc aggatgcagc catgccgtgt ggccaatgaa ccgagaaaac acccctttc 19080

tagaatgctc taaagaggca gaataatcca gaggtgagga aggaaatact ccaccagaga 19140

cccaggcagt tcctacaaaa gccagacttt ccttcaccta gggagtgaca agaccagtgg 19200

aaaacactct caagcagtaa cccccaaatg ctctgcaagc cagtggcgtc cagataccgc 19260

acaagcgagt gggctgtcta atcccatcat catgatgtaa atatctctag gctgccctgg 19320

gctgtgcctg accctgtctt cagctttcca cacctccacc tacagcccat gcacagaagg 19380

accacccagg aatgctgcaa gtgtggcacc tccagggcca cccagggaga aggagggcag 19440

ctatgctggt ggctccaggc ccatttggcg ggtggtacct tcacaccaca aagcccaaac 19500

tgaggcccca gatttggctg atgagggcat attggacagg ggtcacttat gctcttcccc 19560

attgccacct ggcctctggc tacctggact tggctacctg tggatcctct cacaggtgcc 19620

accatcttgg ctgagtctcc agatgcgagg tccctgaggc agtggcgggc ttctcgctaa 19680

tgctgatggg attaggaatg ggataggtgg ggagggccct ggactgggcc ctgatgagcc 19740

aagtgggttt ttagaggggc tactggtaca tttcagggac aggacatctg gtagagctaa 19800

gctggggcaa taaggagcca ctgctaatct gagagctaga aacaatcagc ttctgggtca 19860

ttattaatta gggtagtttg ggctgtgtgg aagtcacgta ctatatgggg tagccacagc 19920

tctctctaca gataatctct aagacttctg attgggactg tgtgaatgca gtagcaatat 19980

ctcttcttac tgccaggccc tgccagtcct gcctccacgc cctggctggc ccccttatg 20040

atctgaccca tgccaggctg ccatagtatg ttacttctgc attagcactc cttgggacct 20100

gcctctccac tgtccctcag actttaaaga actatacaaa cccaaggggc tcttcccaag 20160

agaattgata tgacttgagg tgattccatt tctggaagta gtcactccat tttctgcctc 20220

actctttcag tgcttcacag agcaggttcg aacgaaggag ccatccaact aaccgtcatg 20280

ttcgggcaac cgaagaaggg agtggcagga tttcctttgg agacttctgg aattagacag 20340

cagtttaatg caagcatcta aattctctcc ctcccagagt ctcattaaaa ctacagtaag 20400

agtttgtgtt ttgttttgtt tttaaagaca aaatcccacc aggatagaga gaataggaga 20460

ggagataaca gcatcataat ttatgaaact aaaatgcaga tagaccagga ttaactgact 20520

acacagcacc aaggaagctg aatcacaaga cagcagagga gaaaactgga aaggatcgtg 20580

gtctatacgg cagaatcttc ccaagcctca ggaggaggag ctctagatgt tcccagatct 20640

gggaggtaaa gtggaatggg gggacatggt cagcgtaatg gggttgggct ggaagcaggt 20700
taaggagcag gcagatctct gaatcccctc tctgactctg tgtccccagg catctgcctg 20760
tcccccaccc tggaagaggt ctggcttgac cctttgtctg gtgaatttcc tgctctgctt 20820
tcctggtcct gctggccgga tcagtggagg ccactcactt caccccacag ggatgttctg 20880
tgttgcccta cacctgggaa ctggaggtac tggaggcagg ctgtggtgag cttgaaagca 20940
aaacacagag ggcagtccaa tctctttggc catatttctt ctgcatatcc aataccatgt 21000
ccacaactct gctagtgtcc tgatggtggt gggctctaca cattcccggg aagctgaagg 21060

cagataatga ccaggacagg tcaacctctc ttcttctgaa agccttcatc tactaatggc 21120
ctgggactct tcccttaaat gcttagattg tgtcttccac taaggttttt tgctgttgct 21180
gttgtttgtt tgtttgtttg tttgtttgtt tgtttgtttt gagacggaat ctcactctgt 21240
cgcccaggct ggagtgtagt ggcacaatct cagctcacca caaccttcac ctcctaggtt 21300
gaaggggttc tcctgcctca gcctcctgtg tagctaggat tacaggcaca tgccaccatg 21360
cctggctaat ttttgtattt ttggtagaga caggatttcg ccatgttggc caggctggtc 21420
ttgaactcct gacctcaggt gatctgccta ccttggtctc ccaaagtgct gggattacag 21480

gtgtgagcca ccacacccgg ccaaggtttt tgtttgtttg tttgtttgtt tgtttgtttt 21540
gtattgaggc agggtatcac tctggtcacc caggctggag tgcagtagtg caatcacggc 21600
tcactgaaac ctccacctcc ctggcgggct caggtgatcc tgccacctca gcttcccagg 21660
tagctgggac tacaggcttg taccaccact cccagctaat ttttgcgttt ttagtagaga 21720
cagggtttcc ccatgttgcc caggttggtc tcaaactctg ggctcaagcg atctgcctgc 21780
ctcagcctcc caaagtgctg ggattacagg tgtaagccac cgtacccggc cccgccacta 21840
aggttttgaa aatgaagcaa ttacaagttt aagtctatta ataagtgatg aagccatgta 21900

gaaaagcaga ataattatct tggatcagga aggtcacatg aggatctact tgggggttgt 21960
caatattcta tttcttgacc tgatcagtgt tgacagcagg ttttaatttt ttacttcttt 22020
ttgtttgttt gtttttgaga cggagtcttg ctctgtctcc caggctggag tgcagtggta 22080
tgatctcggc tcactgcaac ctccgcctcc tgggttcaag ctgttctcct gcctcagcct 22140
ccccagtagc tgggattaca ggcaggcacc accacgacca gctaattttt gtatttttag 22200
tagagactgg gtttcaccat cttggccagg ctggtctcga acttctgatc tcgtgatccg 22260
ccctccttgg cctcccaaag tgctgggatt acaggcttga gccagcgtgc ccggcccatt 22320

ttttacttcc ttattaaact gtacatatag gccttgcaca cttttctgca tcaatgttat 22380

attccacaat aaagggaaaa ggtatataca caacttgata ccagtaatgt gaaacatata 22440

tttctacata gaaaaaaaaa tgactgaaat actgcactcc aatgtgttca cacagtagtt 22500

gtttctggat tatttatata ttaaatgttt atatattgta ttatgccatg aggtttgtgt 22560

tttctctcca cttttctgca ttttccaagt ttactacaaa gagcacatat tactcttata 22620

atcagaaagt cataaaatat atttaaaaag acaaaattga aactaataag gatcaacaca 22680

aaacagatga gccatctgtg gaaatccgca cagaatacta cctaaagaga ttggtgacgt 22740

gcatgatctc actaggatga gcacaaagct tgccagagcc tagggtctat ttctagggtt 22800

ggctcttgga agccaggata gttgttatct ctgggaagag ggaggggcac acaaggggct 22860

tctaaaacat tctgaatgtt ctatttctga acctggttgg tgggtacatg actgttggtt 22920

ttattattat atgttttata tactcttccg tatgtatggt gtggattcca aaaaaagatt 22980

tcctttagag aaaaccagaa tcacataagt agaaaatatg gtgctatgtt gaaggaacaa 23040

ctcaagttta tataaaatca tcatcattta taggcttaaa aagttgcttt ggaattttgg 23100

tctaactgac ttgtcttttc tgcagcaaac cacgctcctt ctggacgtgc tccaggcaga 23160

ggggattagg gtgggttcaa ggctgcaagt acctagctca gcacactctc ttcaggggac 23220

ttagagtttg tctggtgttg gctctctgag ctcttgtcag gaatgccgac ccttccgagg 23280

ttcaggattt gaagcctgcc ttcccacccc agatttggtc cacacagaca ctcaagtatg 23340

tatttcaact acaaatgacc tgtactttcc tattactcct ctctttcatg gtaacctttc 23400

tggtatcctt ccttccctac atttatggga gggggacatc attctctgct ctcctgtcac 23460

tgaaggctcc accttctgtc ttcttctgac ccatctggtt ttcctggggc cacctcctct 23520

ccttaccacc ctaacgcttt tgtaacttga ggagaaatga gagatcacct agtcaggtca 23580

tcattctctg tagatgaaga ggcccaatgg tttgctcaag aattgccaag cgagttaaag 23640

acagagagta tgagagtcag caagacctac agaaagcatc tatctgcact gttttgcagg 23700

gacttagcct ttgtgtgtgg actcctggaa tgccacccac taagaaacat tgtctgacac 23760

caactcccca cttggtaggt ggggacactg aaactcatgg caggaaaggg ccttgcccca 23820

agccagggca gagtgtcact catcactctc aattttcagt ccagggcacc ttgttgtgac 23880

tatcccaaag gcagccactt tccctggtct gaaagacctg aagagagaag agaagagaag 23940

gatggaaggc agagtatgcg gctttgattc atttcctggt gaaaacagat ctatacgaga 24000

agcaaatttc acgaaaggga agagaagaaa gtgtcccata cgttgctggc ctgtttcaac 24060

cttgctttga ttcttgctga aaagggtacc gtgtatttct gagttcaaca tgcagaccag 24120

tgttaggaaa gccactgcac ctccacttta gcctccaggg ctgtgccctg caaatggcct 24180

gcagccttgg tgcctcgctc tccagactgc attttggaag atgggacaga ggcttatgga 24240

agcccacatt agaacggggg agcagaatgg gtgagatgag ggatccttga tagtgaacca 24300

gatgaaggaa tggtagccaa atgccaggcc tcctttgtgg cttcaatcca aaggctctgg 24360

agcccttcca gggcagaaca tcaggcatgt ttacccccac tgtcctcaac agtgacagag 24420

gtgcaatctt gggcagctgg ccattttgaa agcaacctcc ttaatctcaa ctgggaaggc 24480

tccctagcag gacccctgtg ttgcacacct ggaggaagct agactaacca gaagctcagc 24540

acggttccat ctgggatgcc caggtctgag acgaaaaagg taactctctt ttctgggtcc 24600

tggcccagtt gtgtctctct ccacctcatt ctctgagatg cctgtctccc ctttttgtc 24660

ccatcaggag gcaagagcta tcactgggcc agactccacc agaagccaag ccagcttgtt 24720

acccagcttc tcagggagca aagaacagcc ttgtttctat cttatcccca ctgtcccctg 24780

cccctgcccc acctcccagc cattcagctt ctggcttccc cagagctgcc tgcttctttg 24840

tggtcctcca ttccttgaaa agaccttcta gtcattagtg tatataaatg gccacttagc 24900

ccagattaca gtgaggtcaa cagctggggc tctgagaatt gtcacacact ggcacaggag 24960

aggaggctat tcttccagag aatttggagg gcactcccat ccacttacaa caaaaagccc 25020

atccactgtg cttggcagta ggtgatctga gaaccaatgg aaccaggtta atcctgtggc 25080

actgttgagt gaggagagca gtggcgggca ctggaaaata tcagagacaa ggcaggagac 25140

ctgaaatcta ggcttagctc ctcatatact tggcagctgt atgacctcag acaaccagtg 25200

ttacctctct aagcctcagt ttcctcatgc aaaaggaggg ggaataacaa cagagcccac 25260

tgcttggggg tgttgtgagg acaggatgaa aaaacaaaca gaaatccctc agtacaggat 25320

tcagtgcagt ggacagtctt gcaaggtctg gttcagccct ccacccctac cctcaccagt 25380

ataaagaact ctggcctaca agtcagatga cctgagtttt aatctcagct ttgccattag 25440

ccgtgtgaac ttgagaaagt ccctttcctt tttacatcta ttgggatgat catgcatttt 25500

ttgtccttta ttctgttaat atagtgtgtt acattgattg cttttcatag actgaaccag 25560

ccttgtattc cagggataaa tctcacttgg tcatggtgta taatccttta tacaaatgtt 25620

gctgggttga gtttgctagt attttgttga agatttttat gtcttgattc ataaggaata 25680

ttggtgtacc ttccccttttt atggccacag tttccctaca atgatgtagt cgaactagac 25740

aacctccaat atctttcagt attcatgtcc tctgattctg tgaaactaag aaaattaaga 25800

aatagtgatt cataggcaca aggcaggcaa aacttagact ccttgtagaa taattaggaa 25860

gccaaatatt cagtgtgctt atttctcaaa taaccttagt ttctccagtc tgccccaact 25920

ccgaggcctg aatatctcta gatgcttatg atggcaacta aagcctaaaa gctaattcat 25980

tttaaagttc ttccaaatgc atagggtttt atttttccag acctgggttc agatggggaa 26040

tttgacaaac aatggaaagg gggaaaaaca acaatctaaa cactgagtga caaagtaaca 26100

aagaaatagt ctagctatca gccagtcaag ccagccttgg ctttgctatc caaagtagtc 26160
agtctaattc taccaccagt ttctgttcct gtagctgtct actgcctgcc agggactctg 26220
ccttcccacc cacaactacc aatggaagga tgtggtgacc ataccagtgg ctgctgacat 26280
ctcctgccat gggaagcata attgcctcca gcagcctccc ccttagatcc atcatttttg 26340
ttgcacttgg cctgggctgt actcccggcc aatgactgaa catggtgagc atagtaatgc 26400
aggcccattt ctgtgaggag caggactcct ccagtaggtg actttggctc aaggactctc 26460
tattggcctg gttgaacttt tcctgaactg tgctactgtc tgagactctt cttacccaat 26520

cctctttctc gccccaattg tcacagacca cctgcattgt ggtctgagtc tctccccacc 26580
ttctcttgct cttccctgtt tatctttcac aggcatttcc cccagtacat tccttgaatg 26640
tctaacccga tacgggtgcc tgacttttgg cagacctaag cagacaaaaa ggagtacttg 26700
gttacctagc tcttctttct accacaaaca tcgagggaac ccttttttccc tcacccctct 26760
gccacacccc cactgcccca gtgaacaacc acagagagag ctgtggtata atattaggct 26820
ggtgcaaaag taattgcggt ttttgccatt acttttaatg gtaaaaaccg caattacttt 26880
tgcacctacc tagtatttgt gtccccccaa attcatatgt tgaaacctaa cccacaatat 26940

gatgtcatta ggaggcaaga ccttgaggag gtgattagat gatggggtgg agctctcctg 27000
aatgagatta gtgcccttat aagaagaagc ccaaggaagc taccttgact cttccatcac 27060
atgagaatgc agcaagaagg caccatctac taatcaggaa gagagctctc accagacact 27120
gaatctgcca gtgtcttgat cttgaagttc ccagcctcca gaactatgca taatgcattt 27180
ccattgtctc taagccaccc agcctatggt attttgtcat agcagcctga actgactaag 27240
acagtgagcc acatgagaag tgccccaacc cctcccttaa gcacttggct cacagatcag 27300
tgggttcatt tctgcctgag ttttattgtt attctgtaga tttcttgggc tagatatatt 27360

tttctgttat tttccttctt cacctcagtc atgaattggt tgttttaaaa aagacaatgt 27420
aagtcatggg gaaactcctg acaactctac tctcctaggg ttcctgataa aaggggattc 27480
agttgagtcc tctgatggtc tctacctgcc aaagtccagc agcccttagc aaacatgctg 27540
ctcgtttctg tagagaaggt gctggtgtcc caccatactt ctctctccct catgaagggc 27600
ttgcgaccca gcaaatgggt ggcttatatg ggtctgtttc aaaggaagag ccagctctgg 27660
gaagaaaaac gatgagcata agcataacct accactgtgc ctgggaaagc agacaacttt 27720
tttgatgtgt gaatatctaa tgagaatgga atccatcaat taccttaaac ttaggcacag 27780

tcttcaaatt caatatatgt gggatatact tttagtcagt ttgtagacgt tatttgtaat 27840

aaataatctg gcttctctaa agaaattatt ttaagtgttt ggtttggttt gatttaatgg 27900

taaaattata tttagtggca gagaattata gcaatggtga taaactatag agtgtcataa 27960

gttcatatct tattctcaca tttgaagctg cctgcagatg cattcaagat gcagccagaa 28020

gtcaggagac tcaggctgtt atttggagct catcatttta cagccttgct ggactcccac 28080

tttctcaggg gaaaaatgtg gtgttgaccc agattagctc tccaggccct gctgagttgg 28140

gcactctgta agctggaggg tcttctattg tcttcaccta agtgtcaatc aacaacccaa 28200

atgggcatgg gggaagaggg agctgggcca atgcccaggg tgcctggtag agagatacct 28260

tgggcactgg aaggcaccag cttcccagag agaaggggga gggccatgaa aaagttggct 28320

gtagatgcca gggacactgg gactctccag ctgtgtgttt gtgtcttctg aagacttatg 28380

tttcattcct ttggagcatg cataatcata cactgtggga tgtgttatat agattgcttg 28440

atagttcacc actgtaataa aatactgtga ctggaatctg ctcccagtct gcctttgata 28500

gcacttgtgc aacacacatt tactgagcat ttacagtgat ccaggacctg tgttgtgaaa 28560

acattgatgg acaaggcaga tggtggagca cgtcagtgag gatttttaac aaaggctggt 28620

aagtgctata aaggaacatt gtaggacact agagaacaaa gaacaggaga acctgactta 28680

ggctggggtg gggcgttggt tagaggaggc tccttggagg acatgaggtt taagctgtga 28740

cctgaggatg aatagatgtt ggccaggtga ggtaccggta tttgtcagcc ttaccagtaa 28800

aaaagaaaac ctattaaaaa aaaaatacac atacaaagcc tcatcagcca tggcttacca 28860

gagaaagtac agcgggcaca caaaccacaa gctctaaagt cactctccaa cctctccaca 28920

atatatatac acaagcccta aactgacgta atgggactaa agtgtaaaaa atcccgccaa 28980

acccaacaca caccccgaaa ctgcgtcacc agggaaaagt acagtttcac ttccgcaatc 29040

ccaacaagcg tcacttcctc tttctcacgg tacgtcacat cccattaact tacaacgtca 29100

ttttcccacg gccgcgccgc ccctttaac cgttaacccc acagccaatc accacacggc 29160

ccacactttt taaaatcacc tcatttacat attggcacca ttccatctat aaggtatatt 29220

attgatgatg 29230

<210> 305

<211> 29996

<212> DNA

<213> Artificial Sequence

<220><223> HDAd35-T4-Ef1a-mgmt-GFP

<220><221> misc_feature

<222> (1)..(481)

<223> Ad35 5'end

<220><221> misc_feature

<222

> (14126)..(14159)

<223> FRT (Complementary)

<220><221> misc_feature

<222> (14220)..(14463)

<223> pT4 LIR

<220><221> misc_feature

<222> (14478)..(15812)

<223> Efla promoter

<220><221> misc_feature

<222> (15830)..(16450)

<223> mgmt

<220><221> misc_feature

<222> (16451)..(16522)

<223> 2A

<220><221> misc_feature

<222> (16523)..(17242)

<223> GFP

<220><221> misc_feature

<222> (17269)..(17390)

<223> SV40pA

<220><221> misc_feature

<222> (17501)..(17766)

<223> pT4 RIR

<220><221> misc_feature

<222> (17873)..(17906)

<223> FRT (Complementary)

<220><221> misc_feature

<222> (29589)..(29996)

<223> Ad35 3'end

<400> 305

catcatcaat aatatacctt atagatggaa tggtgccaat atgtaaatga ggtgatttta 60

aaaagtgtgg gccgtgtggt gattggctgt ggggttaacg gttaaaaggg gcggcgcggc 120

cgtgggaaaa tgacgtttta tgggggtgga gttttttttgc aagttgtcgc gggaaatgtt 180

acgcataaaa aggcttcttt tctcacggaa ctacttagtt ttcccacggt atttaacagg 240

aaatgaggta gttttgaccg gatgcaagtg aaaattgctg attttcgcgc gaaaactgaa 300

tgaggaagtg ttttttctgaa taatgtggta tttatggcag ggtggagtat ttgttcaggg 360

ccaggtagac tttgacccat tacgtggagg tttcgattac cgtgtttttt acctgaattt 420

ccgcgtaccg tgtcaaagtc ttctgttttt acgtaggtgt cagctgatcg ctagggtatt 480

taccggtatt caaggattac atgagcttag aaatgtaatt agcatagtgt gtggcatagt 540

gtagatacca aataaatatg atctctcctt ctactcttga aaatgcaaac acattcttgg 600

tggtcctaaa atagcctgta acatggttta ctcagcagca tttgctattc aaggcagatc 660

tgcctttagt cattggctgc gctcctgaac agctgtgtga aaggctaact tttgtaaacc 720

aaatcaaaat aaaatgcagc aaaaatttgt cactgaaagg aaatcctcag tatatccttt 780

tatgaaatga aagatccctc atccaaactt aacttttta aaagtgcgca tttggagata 840

tagccctttc ttatgaatcc taattcaatt ttggccataa acacacgttg atgttcccca 900

ccccaaagca catagcaaca agagtaggtt ctatattgaa aataatgaca atttaaaaac 960

atgtacttat ttcactgtat gtggacagtg tctatgattg catcatgaag tgtcatataa 1020

ccatgtacgt gtacatgaga gagagataga gagagaagtg gtagggtggt ggtggtagag 1080

gggatggcga tagtaatcat ggtaatggta gaggtgatgg aggtggtaat gacggaggta 1140

agggtggtag tgatgatggt ggtggtggta atggtggtgg atgtggtggt ggcaattggg 1200

atggtgggat ggtggtagcc atggtgatgg tggtaatggt gttgatttaa agggtggtgg 1260

tagtgaaggt gagggtagtg gtggtggagg tggtggtgct ggtagcaata gtgatggtgg 1320

tgatggtgtt gatgagggtg ttgggatcag ggtgagttcc cacagtatat ttcattcttg 1380

ttgtaccact ctgtcaacag caccactgac tgggacagag gaagaaggca cactctgaat 1440

gtgttattaa cagaaacctc aaaacagtct gtctccttgt agtcattcaa aattatcttt 1500

ttcttacctg gaaaactgaa actgaattac cgggaaaaac acaggagatt tttgtttgtt 1560

aatatgctgc caataaagta attttatgtc aaatttaact acaggaaagg gcaaggcatt 1620

ttctaagttc cttagatgtc atgtggctaa aaaaaacaaa aggatggaca gcagttagat 1680

actgtacact tagctgtttg aagccatata ttcagaaagc agatgttggg agttggtgtt 1740

tgaggactga tttcctggag gtattttata taggccaagt tcattgttct aaactctaag 1800

ggcttgactt gagggaggaa aagaggcaag aacatgttta gttttgctga cagcatcaca 1860

tgggcagccc taaggctaga caactttagg gcctgaagct tattctagga aagaagcacc 1920

tacagagtgg cactgggctc ccctccacta tagagatgaa gtcatatgac agtaaagggc 1980

aggcagggct gcctaggggg cccagaactg acacttccat tagaatgagc acaggccagg 2040
gagagaagtg gggaaccaga gagaaggagc tggaattcta gtaggacaaa cggtaagtga 2100
acaacaagaa caagttaaga gtgtgtgcag tattctttca aagactgaaa aaatagtgat 2160
gtgatagaat ggcaggtggc tctgagcagg ccaggagaag gactgggggc agagcatccc 2220
aggcaggagg gcagcaagtg ggaaggccct ggggtggggc ttttggactg ttccagtgac 2280
gggcaggcag ccagtgtgcc tgtcacacaa tgcaccaggg aagtagtcgt gaatttgcag 2340
agggtcttgc aggctatggg aaagggattg gattgtattt tgtttgtagg gaagccatcg 2400

ggggacttaa gcagaggaag gattggcttc atctctttga aaaagttctc tctggatgct 2460
gatgggagga gaaatggaag gaaaagaaac acttttaggg gcaagaactt ttgagaaggg 2520
tggaattggg agtgtggagt tggggccagc tttggcacag gaggggaagc taaacacgtg 2580
gccgcatgag ggcctgtaat tctacctgaa atgggtacca tttgttaggg taaacaaatg 2640
aaccaaatgc ccagtgatac agaccaagtg ttggcaaact tcttctgtga tggcccaggt 2700
agtaaatgtc tcaggcttcg caggccatgt ggtctctgtt gaagctctgt gtagtagaca 2760
atatgttaat gactgggcgt gactgtgtgc taataaaagt ttatttacaa aaacagcccg 2820

tgggctggat ttagctcaca ggctgtagtt tgccaacctc tgacctagag catgaactga 2880
gcatcttctt ggagggaaat aagttctttc caagttgccc tcctcacatt gcagggggcc 2940
atgtaggccc attattcaca gaagagtggg tgggcaacct ttctggagca gaaaaacgta 3000
aagatttctt ccgtagtgca agtaaggtga ccatttctaa accgtgcaag tgatccagca 3060
gtcccaaaag ttgtttcact tctcattgtg cgcccgttct caggtgctcc gaagcttcca 3120
gtcctttgta gggacatgga tgaaattgga aatcatcatt ctcagtaaac tatcgcaaga 3180
acaaaaaacc aaacaccgca tattctcact cataggtggg aattgaacaa tgagatcaca 3240

tggacacggg aaggggaaca tcacattctg gggactgttg tggggtgggg ggaggggga 3300
gggatagcat tgggagatat acctaatgct agatgacaag ttagtgggtg cagcgcacca 3360
gtgtggcaca tgtatacata tgtaactaac ctgcacaatg tgcacatgta ccctaaaact 3420
taaagtataa taataaaaga aaaaaaaaaa gagaggagag aaacatcatc ccctccagga 3480
taccttggg ccttgttctt atagtcttgt acattgttga acaatttgca tgggctagtg 3540
gattaaagca caccctccac cctcaggccc tcaagggtct ctatgataat acagtctcac 3600
cttctaccct ttccatcacc atcctaggtg ctatggccaa ccttgaggct gccatgttag 3660

gtctatgcat ttcccacctc caccacataa ctctctgaag gccaggtagt ttcctattca 3720

tcttggtaac cccaaagcct cgtgacaggg ctcagctggc atctgcggat gtgaatgaac 3780

cattggagaa aatggtactc tgcaaataac tctgttattt tcccatttcc tgtgtaaggc 3840

ctagagacaa tgacttttta attgcacccc ttcccctctg tatgacactg gccttctctt 3900

gtgtccagca atgtgggtgg cctagatgat ttctaaggga cttctggcca agatgaacag 3960

cagctgcatc ttactgagca tttactatgt gccatatact cagccacagc tctaggggca 4020

tagaagcagg agctctcagg gtcagggcag tgagtgagca agcgagcacc tatgccagcc 4080

ctgcctctgg atggggactt gagagggtga tggaagcctg cagcactgga gggaggcaga 4140

caaagacagg cctgtgctga gggggcccgg agcaagagag agggaggcaa tgacagcaga 4200

gacatgcctg cgccttgggt ttgagtgccc agtggtcaaa tccacttccc tgtggctgat 4260

gcttgccttt ctaactttgg aatttagggg ttggagatct ggtgagaagg taggagggag 4320

atgaggagga gaagggaaag gcaggaagga aggggaggga aaggaaaagc aaaaggggag 4380

gaggaaggtt tccaacaaat tattctatat caactgcgga aatcaaaatt tgttgcccaa 4440

atcttagaag ctcatgtccc tcctccccag aagtctggaa tgcagcactc caggggtagc 4500

ttataaccca aatatctatc tgtaaaaaga gaaacattgg gctttcgagc tgtggattct 4560

cagtaaaagc aagaggcctc agcctacaca ggccagccca gagtttgagg aaccccaggc 4620

ccacacccac agggctggcc cctgggtctg catactccct agaaatgtgc acacttctga 4680

gcctcaactc tgtcctggag tctaacagca tccctctcct tcctggggca gttccacctc 4740

cagaaacctg ttaccttggg ccttatgtca aggaaactgt gggaaagagc taggcaggaa 4800

tgcagatgag gccagcatgg gctcctaaaa gtttagaaat aggcagtgtc atgctcccag 4860

gtgcctgcat aaaccagctg aaaaatggag ctcccctcac cagcactctc ccttcaaaca 4920

gactgtgatt tgcaggtcac tggtttacca agccaggcta cccaggcagg acccagatgc 4980

caagcccagt ggtgtcctgc aagctgagca gtgctcagtt cttgcaaaaa aaggtctgtg 5040

tgaaggcaag gcctctgcct ggcttctcac cccagttggg tgtctggaac aggaaggagc 5100

ccttactgca gaaaaaggag gagggagcaa agggagcgaa cagctgcgtg ctccatgggg 5160

aggatcccca aagtagaaag gcgcatacac actgcagccc ttgacccaga atgctcacag 5220

ctacattaca gattcaggtc tcctcagtgt agtggggctg ctgatgagac tgtggcatcc 5280

tcaggggtca ggacacacat tttccatcac tcttctgatg gcaaaaaacc tctgagccaa 5340

tgccaacctc tgatcattaa aaaaaagtgc tcacagcagt gtgtggttta ggatcatgcc 5400

ctgtgtggtt tggaacacgt gcacaaccac accttgttca tcaccatccc agaaaccctg 5460

acgcaggcaa agagcagagt tattaaccct actttactga tgtggatact gaggcccaga 5520

ggctcatgca agttatcaat aagtggcagg gacagttgcc tctagattaa ctagccccta 5580

ggatcacctg ggtcttggaa ggggacccat aaacatgagc tcccctctct tggggccaga 5640

tttgcacctg tgccgcgcct tcagcctgca tgaagtaggg gctgctggca aagactcaaa 5700

gctgtaaatc tgggttttct cttgaggctt ctaagggagc tgtttcgaca actcactctg 5760

ttcccagctg gctgcccctg catagggttt taaagcagcc tagctttctg ccaggcttgg 5820

cagtggacaa cgctggtcag aacatcccag agagctacca gaatgaagta agtttgcttc 5880

tactctttac ctgtttatgg gctgtctctg ccactggaat gaaaggcact gagaacagtg 5940

cctggcctgc agaaggccct ggaaatacct gagctcctaa tctgggaata ggagtaggaa 6000

gagctttgga ggcagggcac ctgagtttga gatctacaac ttcctgcctg tgtgacattg 6060

ggaaagtctc catcctttct gagcctcagt ctccaccctg gggaagtgga aatatcaatc 6120

tctgtgacac agaagcaaat gagcgaatgt gcacaaagta ccttgcacaa gagagacgct 6180

caaacacttg cctccaggtt tcaccgagaa ctacagagta agatagattt gttcccagtg 6240

gaggaagcct gggaataatt tgcccctaga ctatgaattc ctggggctca agatcgagca 6300

cagggccagg cacacagaag ggaccctgga aatgtggcag gaggccagag atagacaggc 6360

ccttagagct catacccatg ccctctgacc tcaagaagaa agaaacctgc tcaaaatctc 6420

acaaagagct tgttccaacc ctgaatcgag tctgaggact ccttcctgag tccagcactt 6480

tttctgcaag aagtatatgc ctccaaagct gatgggcgca aatcttgaac cccgtcacat 6540

aaacacaaag ggaggaggtg actagagctc ctcctactgg atatgtctaa ggtcaccagt 6600

ctaaagaaaa gggatggata gaatgaggcc agtatttttg cagccatcca aatgtccaca 6660

tacgctgtta cactgagggc tcctctctcc cccgtcttca gccctacttg catttagagg 6720

tgagaaagat atgggctgag gggttgtttt tcatcgtatt gtagatggaa agcacactgc 6780

ccttggggcc atccaaatgt ggaccttgat gtagcacccc accttctgga tggccatcct 6840

tctgaaagtc actgaatttc tcagacttta ttctctttat ccataaagaa ggagaataat 6900

aataatcccc ccaccctgcc caaccactga ctggttggga agctcagaag aaatactggg 6960

cacggcatcc cattgtaatc tatagagtga gtcgcttctt aatattaaat ggctgaacac 7020

agaagatgtg caaaaagtac tgtgtcccct tcctcctcca actgaacatt tcatgcccct 7080

tgcaccctca ttttgtctag gagctgcctt atgaagggaa taggtacctg ctccgagctg 7140

gaggaatctt tgccacttat ggtggggtat ggactgagac agagatggca tgtgacatgc 7200

gcactgagtc tcaactccat gcaggctctg gagcactctc aaattggagt actaatgcct 7260

tttaaattct cacactagca atcctttgac ctactgatct agggatctag ggaaagaatc 7320

gtgatcttaa cttcaaaggg aaggacaaaa tgttctgcct cctgttaaaa ctccatacac 7380

taagtgcaga gactggatgc cttattaacc ttgggtagat gcccaaatgt tcaaaaggtc 7440

aaactcttct gttccccaga tcgccagagt cattaaccag tcacactatt aaatgaatga 7500
acagatgctg aaaaggtact tgcattactg agatttctta tggtgatggc ccctgcctga 7560
tatgtattca gcattttgta gttttcaatg tgcattagag tatagtggtg atgacattgg 7620
cctctgagtt tgccacttct tatatctgtg actttggtca aattgcttaa tctctctgag 7680
tctcggtttc ctggagataa taatagcttc ttcttcccag ggttatcatg aggattacag 7740
gagataatgc cccaaaaatg cttagtaaag tgcctagcac ctagtcaatg ctgaattaaa 7800
ggtggttatt cttacttttc gttcatttga actttgttct cagggagggc aaaggataga 7860

caaagcccca tagctagtga ggagtagctg caagactaga acccaggtgt tctgagccct 7920
agtcttaggc caagaacaac tgttacgtga gatgcacgtt ttccttcaag ggagctcaca 7980
attatttcca tgtaaattca aggactgcta aaagagaact ctcctctggg actgatatca 8040
ttttatttca agattgattt gaaacatgtt ttttgtttgt ttgtttgttt tctaggaaag 8100
aacaagagaa ccagttaagc tgaatgcctg aagcaaatcc ctgttagcga tgttttcagg 8160
atgagggaga gtggtgcaag aaacgtgctt ccagatgcac atggtttcct gggactaggg 8220
ttcagggtgt catccctggg tgttattaag tgtcagaagg agagcaaaca agggaaacat 8280

ctgagatcca gctaaggcta caccctggaa atgcaagccc agctcttgca aaggacctcc 8340
tttggccact caccttccag gccttacaat aacttgtttg gactgcaggt ttcttggtgg 8400
actcacaggc cattctgctt ttatttggtc aacctcagtt cacaagcacc cagatgctga 8460
gatcctcagc atgtgcagca gagtttcata ttagcactgg gtacctttct gaggctacag 8520
ggataccgta cagcagcacc tgtcacgtcc agccaaagga gtgggctctc tcaatgtcat 8580
ccaatgctgt ttcaactgtg aagaagacca tctgagagag ttgcttttgg aggctgaggc 8640
aaatttttaa aattctttgt tctcctcaac tggggtgaat tcttggtctt ctaggacagc 8700

ttgaagtttt agaaagagtc aagccactca gaaccaacag agaactcttt cagagaacaa 8760
ggtgtggcat agaggaggca gagggctgat cttgatcaaa tccaaagtgt gactctaaag 8820
caatgaatgt gaatttttgg caaagcttac aaagggctct aaaggccatc tgcaaagaga 8880
agccaagcct gatcgatgaa tcactagtgc ggccggatat cgatcggcac gctgttgatt 8940
ttctcatagt aaggaacagt gggccctttc agtcccactt ctgtagtctg tggtactaca 9000
aatggtgagc ccatgatgtt gccattcata gggttattct ccagcagtaa tgactggcca 9060
gccactccca tagccgcggg gctaggattt attgtcaatg gagggacctg cagttctgca 9120

caagcagtac taggatgagc acctgggccc attgcaaggg tgacatcttc aaggcaaggc 9180

ctcttaattt tattagggta gcccccatca gccatgtctg gaaactggaa gtggtcttct 9240

tcttgtctcc tcttaacagt tccctgtgaa tggaagagaa gagaggagga gaagagagga 9300

gaggagaagg gaagagaggt gacacacaca cacacacaca cacacacaca cacagagaga 9360

gagagagaga gacagagaga aagagagaga gagagagaga ggaattttta taaaggtttg 9420

gcacattaaa gctaatgaac aggaaatgtg catgataaaa cagacctctc agtttaaaga 9480

cttatagttg tgaaaactat aaaatacagc ctgtctttgg aaccatagtg cttatttatt 9540

cattattatg tttcatctaa actgtctaat tacatttcaa ataaggcatt atgttgtctg 9600

tatactaaaa cgggatagaa cgttattcaa agggtaatct gcccacttca aggagagttc 9660

aacaaaacta tgcagaagtc actaaatgaa ccatgctgcc aaaggcaggc attggagaga 9720

aaactagaag tagctaaata gttttaattc tttcctgtct acagacacat agattttaac 9780

gaaggaatac catagtatag aattgaactt ttaggctgcc ttctagtctt ggttaaatgc 9840

atcaggctgc agtggtaaaa ttgaatacaa cagagccctt acaggaaaga agtagatctg 9900

gatgtgtttt cttggggagc tgtttaaaat actgtttttg ggaaagcaca agtttcagaa 9960

cagtcattgt aggcatcgta ttcattgttc catttatttt tacacacaca cacacacaca 10020

cacacacaca cactctcaca cattgctatg tgtacacaaa aataatttgg aagaacctat 10080

acccaacaat ttggagtggt catttatttg ggatgactgg caattccctt tctattctct 10140

tcatttctgc ttgtttgtct ttaacgagaa cgactcataa tccaaaaatt taaaaaagta 10200

taaagttatc taaataagaa attttcctct gaagatgcat cctcaggttg gggagatatt 10260

aaacaatgag aaaaggcccc aatctgggat ctgaaccttg ggggagctgc ccatcattta 10320

tagaagcaca gcctttggga acaaagcaaa gtcactagca atgtgagact tcctactctt 10380

catggcttca tacagtcatc catcgctgtt gtgttaatga ccatgacctg tatgttagca 10440

ggtaaatggg aaaggaagtg ggggcaaagg agtatgtgca ggaatgatca aaataaggaa 10500

aggaagagag ggatctggaa atcacctgaa tgccgatagg tgaacaggta gaattctttt 10560

aaagcttccc ccacccggta ccccccaaat aacccctttc cagctttgga agtttcacta 10620

ggacatacag tgctcatcct ctgatgtcac cttaagtttg gctcttctgg tttgatgagc 10680

ttgtagccca ctaggagctc aaggcatgca tggggccact tgccagcacg atgaggggca 10740

tgactgtcat ggccaagtga acatcaaagc agatccccag ggctgtatgt ctcaggcctt 10800

ggtgcacatc agaatcactt agaaacatcc acattcctgg gccctcccac cacaaactga 10860

cagcttcatc cagggtgtgg cccaggcatc gggagttttt ccaacagctc catggctgat 10920

tctcaacaga aaaccactgg cccagagcaa gggtggaggc agcgtggcat agggctctga 10980

ccttggcctt gccactgaac ctctcagagc cccagtttct ttatgtgtaa aatgagtgta 11040

attatagttc ttttctcatg aaggtgctct gactattaag tgaaacgggg cacattgtat 11100

gacacctaat agctcctcac taactggtac ccggcattat aaagggcagg tatggaaggg 11160

ttctgggagt ccaataccct tcttaaagac agagaggtct ctgagaccca gagaggggca 11220

ggccttaccc agagttgctc agccagaggg caacaaggcc caggtcagat gcagggcccc 11280

tccaccacca ctcagctgcc tccagaccca ctgccttcgc catgttgttg gtaggacact 11340

gcatcgcccc cacagaaggg gcttgccaac ttgagtgaga ggacttgcac acttctttga 11400

cttttctttt gagatgccca caatctgaac aagggcactt caagggacag ctctgtcacc 11460

aaactcatct gaggcctgaa taccatgggt caggcaggaa tgggttggag aggtgtagag 11520

caggcacaat aagagggctg aggcccatgc agtcatcagt gcccactttc ccaggagtct 11580

gactgggcac agcacccata gtgtccctga gctggtccat ggagcagctc actaactgtt 11640

tggcccacag caggtgctca gtaaatggca gttgaacgaa tcaatggaca aaggaacata 11700

aattacccaa cacacaggga gctcagccat ttactcaatc cattatggag taacctacaa 11760

acaagccact gggtcccaaa ctgaaattgt gtctcttcta cattctccca aagaatccaa 11820

taggttaaaa atagaaatgt atgaaataga tcaatcaggg atgattgcat gtggatttga 11880

cataaggatc ccctgcaggg agtctgagct ggcaacagtc aggcccaaag tgctgtccat 11940

gatgtctcga actgcaagac agttttaaca atggcgaagc aatgcagaac caggcaggcc 12000

aaggaggggg tgggggttgg ggaaaggaag ggagggaagg ggctgtgagg ggcaatggtc 12060

tggcatccct gccacgtgag cctctgaaat ttgctggcag cttctatggg ctcccagagc 12120

tttcacttaa ttgttggtct gccactaacc tgctgggagt aaggtgcagg gatggaggag 12180

gcagggcatg accaccagac actaaaggta ccagctgggg ccactggcaa agggaaggag 12240

gctgcacctc tcctacatga gagcccgtat acacacacct tttccagcac tcatcaactg 12300

catcccaagc aaatggtccc tgatcaattc caattctaga aaccaactga ctactcaata 12360

acaaagtaga tcccagcagg ccgccactgc tggagcggat gccacttttg ctatgccaag 12420

tctgtggctg gacagctgct ggcatgtaca ctcactgact ttcataagga tgcctaataa 12480

agggggcagg ctcacctggc ttttctcagg ggtggggttt ggggtgccga tagaggctgc 12540

tgttttggca gagtggcaag ctgcaagcct cttctgagct ttcatttttc aatggacttc 12600

agtgagaatt cactttgtca gaggccatgc agctccatgt tttggatttc atggaatgag 12660

ctttcaacag tgagcctgaa gtgccctggc tgaacagcaa gaacaccagc caaccctaaa 12720

caaggccgag gagaggcggc tgtgtttaca cggaaggctc agccttgctg taatagcgtc 12780

tgccttcacc agacatcagt gaggcgtgga aatctattat ccagttaatt ttgcccctag 12840

ataaagactt gctttcgtgt cttctctttc acagtcccat gatctgttac tcatctcaac 12900

tgcgagaagt tggctgggct ttcccctgtg cccagtgcca cactcgtgcc ttcactgggt 12960
cacctgtgcc tgtggctgat gccgctgagg ttttgcctgc ccagactggg tgtttctgac 13020
taaatcccac agccaccatt ttagatcaag ggcaggagat agctcactgc tccggaatga 13080
cctcccctcc cagaatcctg gtaggggcgg aaggtcccca accaagctcc cagccctttc 13140
taaatgaatc tccctgcttc acccatgtgc ttttctccag tctctgcggt cttgatgaca 13200
gcagggtatt agtcctagct gtcccacagc tcctacttct ttcaggcctc tccctgtgac 13260
aatcagtagc cactggcagg atttcctcag agcatatctc gatttgcttt cagacaatta 13320

gttaaaagga cactggaccc cagacgtccc aactcccagc cagagccctc acaggcccgg 13380
cctttggtgg tgaggaaggg ggagggagtg agtgacagtg ccctggcatc ttttagaaac 13440
gaattccttt ctctccatac ataaatgcct gcagagtccc atttcagaat ccggcagaca 13500
aagccaccaa tgtgatcccc atgaccttat aaacattcat taaaatgcat ttcaaggcat 13560
gtgatggcct ccccaccccc tagataatga gaaaacaaag gtttctcttc tgatagagac 13620
aagttcagct ctgaagtcaa cattatttct ggttctgtct gaacaatgac atatggcaac 13680
tcttcccttt ctatagttct agtccagaat gacaaaaaag gggaaaaatt tcttagagaa 13740

ggtagagatt atacgaatac agtccatgaa atgagcataa ggagaataaa gaatataact 13800
tatccaaaga agtctggcag gctgttataa atgcttgatt ttggacactg tagttggagg 13860
tttaacatgg acaccaataa aaaggtcagc aaagggtatg cactgttcct attgggcaag 13920
aagataggag gtcaaaggta accaggaaag ataaactcag ggagacttat tttccctcca 13980
gagggcactg ggcttgtagg ccctgggcaa aattgtcaaa aaggtgaaaa tcgcctgtgg 14040
tttatttagt ctgctctttc ttcactagtg cctcaccagt tcagttcagg ccaatttgct 14100
agaaggtagc gaacgatcga ccggtgaagt tcctatactt tctagagaat aggaacttcg 14160

gaataggaac ttctacctag atgcatgctc agagcggccc ctagctagcg tttaaaacct 14220
acagttgaag tcggaagttt acatacactt aagttggagt cattaaaact cgtttttcaa 14280
ctactccaca aatttcttgt taacaaacaa tagttttggc aagtcagtta ggacatctac 14340
tttgtgcatg acacaagtca tttttccaac aattgtttac agacagatta tttcacttat 14400
aattcactgt atcacaattc cagtgggtca gaagtgtaca tacacgcgct tgactgtgcc 14460
tttaagcttt taattaagag taattcatac aaaaggactc gcccctgcct tggggaatcc 14520
cagggaccgt cgttaaactc ccactaacgt agaacccaga gatcgctgcg ttcccgcccc 14580

ctcacccgcc cgctctcgtc atcactgagg tggagaagag catgcgtgag gctccggtgc 14640

ccgtcagtgg gcagagcgca catcgcccac agtccccgag aagttggggg gaggggtcgg 14700

caattgaacc ggtgcctaga gaaggtggcg cggggtaaac tgggaaagtg atgtcgtgta 14760

ctggctccgc ctttttcccg agggtggggg agaaccgtat ataagtgcag tagtcgccgt 14820

gaacgttctt tttcgcaacg ggtttgccgc cagaacacag gtaagtgccg tgtgtggttc 14880

ccgcgggcct ggcctcttta cgggttatgg cccttgcgtg ccttgaatta cttccacgcc 14940

cctggctgca gtacgtgatt cttgatcccg agcttcgggt tggaagtggg tgggagagtt 15000

cgaggccttg cgcttaagga gccccttcgc ctcgtgcttg agttgaggcc tggcctgggc 15060

gctggggccg ccgcgtgcga atctggtggc accttcgcgc ctgtctcgct gctttcgata 15120

agtctctagc catttaaaat ttttgatgac ctgctgcgac gctttttttc tggcaagata 15180

gtcttgtaaa tgcgggccaa gatctgcaca ctggtatttc ggtttttggg gccgcgggcg 15240

gcgacggggc ccgtgcgtcc cagcgcacat gttcggcgag gcggggcctg cgagcgcggc 15300

caccgagaat cggacggggg tagtctcaag ctcgccggcc tgctctggtg cctggcctcg 15360

cgccgccgtg tatcgccccg ccctgggcgg caaggctggc ccggtcggca ccagttgcgt 15420

gagcggaaag atggccgctt cccggccctg ctgcagggag ctcaaaatgg aggacgcggc 15480

gctcgggaga gcgggcgggt gagtcaccca cacaaaggaa aagggccttt ccgtcctcag 15540

ccgtcgcttc atgtgactcc acggagtacc gggcgccgtc caggcacctc gattagttct 15600

cgagcttttg gagtacgtcg tctttaggtt ggggggaggg gttttatgcg atggagtttc 15660

cccacactga gtgggtggag actgaagtta ggccagcttg gcacttgatg taattctcct 15720

tggaatttgc ccttttgag tttggatctt ggttcattct caagcctcag acagtggttc 15780

aaagtttttt tcttccattt caggtgtcgt gagaattcga tatcccacca tggacaagga 15840

ttgtgaaatg aaacgcacca cactggacag ccctttgggg aagctggagc tgtctggttg 15900

tgagcagggt ctgcacgaaa taaagctcct gggcaagggg acgtctgcag ctgatgccgt 15960

ggaggtccca gcccccgctg cggttctcgg aggtccggag cccctgatgc agtgcacagc 16020

ctggctgaat gcctatttcc accagcccga ggctatcgaa gagttccccg tgccagcgct 16080

tcaccatccc gttttccagc aagagtcgtt cacgcgtcag gtgttatgga agctgcttaa 16140

ggttgtgaaa ttcggagaag tgatttctta ccagcaattg gccgccctgg ccggcaaccc 16200

caaagccgcg cgagcagtgg gaggcgccat gagaggcaat cctgtcaaga tcctcatccc 16260

gtgccacaga gtggtctgca gcagcggagc cgtgggcaac tactccggag ggctagccgt 16320

gaaggaatgg cttctggccc atgaaggcca ccggttgggg aagccaggct tgggagggag 16380

ctcaggtctg gcaggggcct ggctcaaggg agcgggagct acctcgggct ccccgcctgc 16440

tggccgaaac ctcgaggtga aacagacttt gaattttgac cttctcaagt tggcgggaga 16500

cgtggagtcc aacccagggc ccatggtgag caagggcgag gagctgttca ccggggtggt 16560

gcccatcctg gtcgagctgg acggcgacgt aaacggccac aagttcagcg tgtccggcga 16620

gggcgagggc gatgccacct acggcaagct gaccctgaag ttcatctgca ccaccggcaa 16680

gctgcccgtg ccctggccca ccctcgtgac caccctgacc tacggcgtgc agtgcttcag 16740

ccgctacccc gaccacatga agcagcacga cttcttcaag tccgccatgc ccgaaggcta 16800

cgtccaggag cgcaccatct tcttcaagga cgacggcaac tacaagaccc gcgccgaggt 16860

gaagttcgag ggcgacaccc tggtgaaccg catcgagctg aagggcatcg acttcaagga 16920

ggacggcaac atcctggggc acaagctgga gtacaactac aacagccaca acgtctatat 16980

catggccgac aagcagaaga acggcatcaa ggtgaacttc aagatccgcc acaacatcga 17040

ggacggcagc gtgcagctcg ccgaccacta ccagcagaac acccccatcg gcgacggccc 17100

cgtgctgctg cccgacaacc actacctgag cacccagtcc gccctgagca aagaccccaa 17160

cgagaagcgc gatcacatgg tcctgctgga gttcgtgacc gccgccggga tcactctcgg 17220

catggacgag ctgtacaagt aaagcggccg ctcttcgagc agatatcata agatacattg 17280

atgagtttgg acaaaccaca actagaatgc agtgaaaaaa atgctttatt tgtgaaattt 17340

gtgatgctat tgctttattt gtaaccatta taagctgcaa taaacaagtt acgttaacaa 17400

caacaattgc attcatttta tgtttcaggt tcaggggggag gtgtgggagg ttttttaaag 17460

caagtaaaac ctctacaaat gtggtattaa ttaaagatct ttaaacaatt taaaggcaat 17520

gctaccaaat actaagcgcg tgtatgtaca cttctgaccc actgggaatg tgatgaaaga 17580

aataaaagct gaaatgaatc attctctcta ctattattct gatatttcac attcttaaaa 17640

taaagtggtg atcctaactg accttaagac agggaatctt tactcggatt aaatgtcagg 17700

aattgtgaaa aagtgagttt aaatgtattt ggctaaggtg tatgtaaact tccgacttca 17760

actgtagttt aaaacgggcc cgtagtctag ggccgccagt gtgatggagt tcggcttcag 17820

gtacagcaca ctggcggccg ttactaggta gctagagcct tcagactcta gggaagttcc 17880

tatactttct agagaatagg aacttcggaa taggaacttc acccatggcg atcgctagcc 17940

tctaactcct agaccgtcag aactgctggg cccttcaaga cgggctgctc acacccactc 18000

atgttaagcc tggtgaggcc tgtactctgt tttcacagga agaaatcctc acccagtctt 18060

ccccaaacac attcccaggt tgtgtcatta gtgggataga gatgattatt gtggggagaa 18120

gagaaacatc tggatggatt tggtgaggtt gatctataga ggaagtaggt gctgcctgag 18180

gtagctgtaa tagaagctaa aggtcaaagg agagggccct gtcccaatcc agatgactcc 18240

acttctgctg gacccaggtt cacaagctta atctacattt cacctaaatt tggctaacaa 18300

gcccaaaatc acacaggcaa agggagaagt ggaggcagaa ccgaggttgg aggccaccag 18360

ggccaccggg cagagatcat ttaagcccaa ccttctcact tctccctggg ctctgcctct 18420
cttaaaggac cttgtggtgt gacctcttgt aggtcccttt cacactcggg gcctcagttt 18480
ccccactgta aagtgaatgg gtcccagctt tggtaagctt atgcttacct gatgctttct 18540
tcctgggctg ctcttgtaga gaaaagataa atcttcttcc tccatccacg agggctcctt 18600
tccctggggg tgagagtagg ctgaggagag ccacttgcac acacccttaa agaaagtatt 18660
acctgcacca gctcagtgag aggcacagat cagactgtta cttgaatcaa attatgagcc 18720
tccccaaata tatctatgac atttaaatag gggattactt gaacatagac tgtgggatcc 18780

ggtgtggagt gcgggagact agcaaagtga atcctgagag tagcaggtct gcacctgttg 18840
gatcgagaaa ggcggcctac aattctggtc aaatgagctg tgcttattga catattctat 18900
tagagagtac taccaggtca ccagtcacca gaaaggctgc cagctctcca accacctcca 18960
gggaactatc ctgaatgggg ccttaacaag tctaagagag ggttggtttg ggtcccaagc 19020
caatatttgc tctgctttat gtcagtcata tggaacccaa accaaccctc tcctatgtgc 19080
ctcaccagtc ggtgcaggga tcccaatttc aagtttggtt ttttatggtc aaagtccagc 19140
atagattaaa tgaaggggtg tgatgatggt gttaaaagag aactccagac cagtttaact 19200

cttggacaca catcccatct caccatggtg cttccaacct tccagagatg atgggctcct 19260
attttctgat gacaaagccc tccacaggat tgctgcctgg ccatcaggga gtgcctctgt 19320
aactgaggct gagatcccac tttcagtcct ccagctgtgg cccatccctg ctccgcccac 19380
cgggtatggc ctgtcctagg ctcttaggta tggctgcatt gtgaaatgat ggctacagag 19440
ctggcatctc ctgtagtctg gttcatctag tgcactacct catagttaaa agaaatctgt 19500
ttaagccact gagggtggct cctagtgcca actccaagaa caggaagctt ccctttttg 19560
ggaggagggg cagatggtaa catggatcgt ccaggtcaat gggagcaggg caaccacagt 19620

aagtactgga caacaacaca aaactccatg tgtggcttcc atcgagtccc tctccaattg 19680
gtttggtctt ctccgtccca tgcagcactt tagcaagggg cctggctgaa ggctatgaat 19740
tgtgtggagc ctcctcattg cagtctccaa ccatctgatg ctgggaaaat gtcaccagga 19800
tgcagccatg ccgtgtggcc aatgaaccga gaaaacaccc cttttctaga atgctctaaa 19860
gaggcagaat aatccagagg tgaggaagga aatactccac cagagaccca ggcagttcct 19920
acaaaagcca gactttcctt cacctaggga gtgacaagac cagtggaaaa cactctcaag 19980
cagtaacccc caaatgctct gcaagccagt ggcgtccaga taccgcacaa gcgagtgggc 20040

tgtctaatcc catcatcatg atgtaaatat ctctaggctg ccctgggctg tgcctgaccc 20100

tgtcttcagc tttccacacc tccacctaca gcccatgcac agaaggacca cccaggaatg 20160

ctgcaagtgt ggcacctcca gggccaccca gggagaagga gggcagctat gctggtggct 20220

ccaggcccat ttggcgggtg gtaccttcac accacaaagc ccaaactgag gccccagatt 20280

tggctgatga gggcatattg gacaggggtc acttatgctc ttccccattg ccacctggcc 20340

tctggctacc tggacttggc tacctgtgga tcctctcaca ggtgccacca tcttggctga 20400

gtctccagat gcgaggtccc tgaggcagtg gcgggcttct cgctaatgct gatgggatta 20460

ggaatgggat aggtggggag ggccctggac tgggccctga tgagccaagt gggtttttag 20520

aggggctact ggtacatttc agggacagga catctggtag agctaagctg gggcaataag 20580

gagccactgc taatctgaga gctagaaaca atcagcttct gggtcattat taattagggt 20640

agtttgggct gtgtggaagt cacgtactat atggggtagc cacagctctc tctacagata 20700

atctctaaga cttctgattg ggactgtgtg aatgcagtag caatatctct tcttactgcc 20760

aggccctgcc agtcctgcct ccacgccctg gctggccccc cttatgatct gacccatgcc 20820

aggctgccat agtatgttac ttctgcatta gcactccttg ggacctgcct ctccactgtc 20880

cctcagactt taaagaacta tacaaaccca aggggctctt cccaagagaa ttgatatgac 20940

ttgaggtgat tccatttctg gaagtagtca ctccattttc tgcctcactc tttcagtgct 21000

tcacagagca ggttcgaacg aaggagccat ccaactaacc gtcatgttcg ggcaaccgaa 21060

gaagggagtg gcaggatttc ctttggagac ttctggaatt agacagcagt ttaatgcaag 21120

catctaaatt ctctccctcc cagagtctca ttaaaactac agtaagagtt tgtgttttgt 21180

tttgttttta aagacaaaat cccaccagga tagagagaat aggagaggag ataacagcat 21240

cataatttat gaaactaaaa tgcagataga ccaggattaa ctgactacac agcaccaagg 21300

aagctgaatc acaagacagc agaggagaaa actggaaagg atcgtggtct atacggcaga 21360

atcttcccaa gcctcaggag gaggagctct agatgttccc agatctggga ggtaaagtgg 21420

aatgggggga catggtcagc gtaatggggt tgggctggaa gcaggttaag gagcaggcag 21480

atctctgaat cccctctctg actctgtgtc cccaggcatc tgcctgtccc ccaccctgga 21540

agaggtctgg cttgaccctt tgtctggtga atttcctgct ctgctttcct ggtcctgctg 21600

gccggatcag tggaggccac tcacttcacc ccacagggat gttctgtgtt gccctacacc 21660

tgggaactgg aggtactgga ggcaggctgt ggtgagcttg aaagcaaaac acagagggca 21720

gtccaatctc tttggccata tttcttctgc atatccaata ccatgtccac aactctgcta 21780

gtgtcctgat ggtggtgggc tctacacatt cccgggaagc tgaaggcaga taatgaccag 21840

gacaggtcaa cctctcttct tctgaaagcc ttcatctact aatggcctgg gactcttccc 21900

ttaaatgctt agattgtgtc ttccactaag gtttttttgct gttgctgttg tttgtttgtt 21960

tgtttgtttg tttgtttgtt tgttttgaga cggaatctca ctctgtcgcc caggctggag 22020

tgtagtggca caatctcagc tcaccacaac cttcacctcc taggttgaag gggttctcct 22080

gcctcagcct cctgtgtagc taggattaca ggcacatgcc accatgcctg gctaattttt 22140

gtatttttgg tagagacagg atttcgccat gttggccagg ctggtcttga actcctgacc 22200

tcaggtgatc tgcctacctt ggtctcccaa agtgctggga ttacaggtgt gagccaccac 22260

acccggccaa ggtttttgtt tgtttgtttg tttgtttgtt tgttttgtat tgaggcaggg 22320

tatcactctg gtcacccagg ctggagtgca gtagtgcaat cacggctcac tgaaacctcc 22380

acctccctgg cgggctcagg tgatcctgcc acctcagctt cccaggtagc tgggactaca 22440

ggcttgtacc accactccca gctaattttt gcgtttttag tagagacagg gtttccccat 22500

gttgcccagg ttggtctcaa actctgggct caagcgatct gcctgcctca gcctcccaaa 22560

gtgctgggat tacaggtgta agccaccgta cccggccccg ccactaaggt tttgaaaatg 22620

aagcaattac aagtttaagt ctattaataa gtgatgaagc catgtagaaa agcagaataa 22680

ttatcttgga tcaggaaggt cacatgagga tctacttggg ggttgtcaat attctatttc 22740

ttgacctgat cagtgttgac agcaggtttt aatttttttac ttctttttgt ttgtttgttt 22800

ttgagacgga gtcttgctct gtctcccagg ctggagtgca gtggtatgat ctcggctcac 22860

tgcaacctcc gcctcctggg ttcaagctgt tctcctgcct cagcctcccc agtagctggg 22920

attacaggca ggcaccacca cgaccagcta atttttgtat ttttagtaga gactgggttt 22980

caccatcttg gccaggctgg tctcgaactt ctgatctcgt gatccgccct ccttggcctc 23040

ccaaagtgct gggattacag gcttgagcca gcgtgcccgg cccatttttt acttccttat 23100

taaactgtac atataggcct tgcacacttt tctgcatcaa tgttatattc cacaataaag 23160

ggaaaaggta tatacacaac ttgataccag taatgtgaaa catatatttc tacatagaaa 23220

aaaaaatgac tgaaatactg cactccaatg tgttcacaca gtagttgttt ctggattatt 23280

tatatattaa atgtttatat attgtattat gccatgaggt ttgtgttttc tctccacttt 23340

tctgcatttt ccaagtttac tacaaagagc acatattact cttataatca gaaagtcata 23400

aaatatattt aaaaagacaa aattgaaact aataaggatc aacacaaaac agatgagcca 23460

tctgtggaaa tccgcacaga atactaccta aagagattgg tgacgtgcat gatctcacta 23520

ggatgagcac aaagcttgcc agagcctagg gtctatttct agggttggct cttggaagcc 23580

aggatagttg ttatctctgg gaagagggag gggcacacaa ggggcttcta aaacattctg 23640

aatgttctat ttctgaacct ggttggtggg tacatgactg ttggttttat tattatatgt 23700

tttatatact cttccgtatg tatggtgtgg attccaaaaa aagatttcct ttagagaaaa 23760

ccagaatcac ataagtagaa aatatggtgc tatgttgaag gaacaactca agtttatata 23820

aaatcatcat catttatagg cttaaaaagt tgctttggaa ttttggtcta actgacttgt 23880

cttttctgca gcaaaccacg ctccttctgg acgtgctcca ggcagagggg attagggtgg 23940

gttcaaggct gcaagtacct agctcagcac actctcttca ggggacttag agtttgtctg 24000

gtgttggctc tctgagctct tgtcaggaat gccgaccctt ccgaggttca ggatttgaag 24060

cctgccttcc caccccagat ttggtccaca cagacactca agtatgtatt tcaactacaa 24120

atgacctgta ctttcctatt actcctctct ttcatggtaa cctttctggt atccttcctt 24180

ccctacattt atgggagggg gacatcattc tctgctctcc tgtcactgaa ggctccacct 24240

tctgtcttct tctgacccat ctggttttcc tggggccacc tcctctcctt accaccctaa 24300

cgcttttgta acttgaggag aaatgagaga tcacctagtc aggtcatcat tctctgtaga 24360

tgaagaggcc caatggtttg ctcaagaatt gccaagcgag ttaaagacag agagtatgag 24420

agtcagcaag acctacagaa agcatctatc tgcactgttt tgcagggact tagcctttgt 24480

gtgtggactc ctggaatgcc acccactaag aaacattgtc tgacaccaac tccccacttg 24540

gtaggtgggg acactgaaac tcatggcagg aaagggcctt gccccaagcc agggcagagt 24600

gtcactcatc actctcaatt ttcagtccag ggcaccttgt tgtgactatc ccaaaggcag 24660

ccactttccc tggtctgaaa gacctgaaga gagaagagaa gagaaggatg gaaggcagag 24720

tatgcggctt tgattcattt cctggtgaaa acagatctat acgagaagca aatttcacga 24780

aagggaagag aagaaagtgt cccatacgtt gctggcctgt ttcaaccttg ctttgattct 24840

tgctgaaaag ggtaccgtgt atttctgagt tcaacatgca gaccagtgtt aggaaagcca 24900

ctgcacctcc actttagcct ccagggctgt gccctgcaaa tggcctgcag ccttggtgcc 24960

tcgctctcca gactgcattt tggaagatgg gacagaggct tatggaagcc cacattagaa 25020

cgggggagca gaatgggtga gatgagggat ccttgatagt gaaccagatg aaggaatggt 25080

agccaaatgc caggcctcct ttgtggcttc aatccaaagg ctctggagcc cttccagggc 25140

agaacatcag gcatgtttac ccccactgtc ctcaacagtg acagaggtgc aatcttgggc 25200

agctggccat tttgaaagca acctccttaa tctcaactgg gaaggctccc tagcaggacc 25260

cctgtgttgc acacctggag gaagctagac taaccagaag ctcagcacgg ttccatctgg 25320

gatgcccagg tctgagacga aaaaggtaac tctcttttct gggtcctggc ccagttgtgt 25380

ctctctccac ctcattctct gagatgcctg tctccccttt tttgtcccat caggaggcaa 25440

gagctatcac tgggccagac tccaccagaa gccaagccag cttgttaccc agcttctcag 25500

ggagcaaaga acagccttgt ttctatctta tccccactgt cccctgcccc tgccccacct 25560

cccagccatt cagcttctgg cttccccaga gctgcctgct tctttgtggt cctccattcc 25620

ttgaaaagac cttctagtca ttagtgtata taaatggcca cttagcccag attacagtga 25680

ggtcaacagc tggggctctg agaattgtca cacactggca caggagagga ggctattctt 25740

ccagagaatt tggagggcac tcccatccac ttacaacaaa aagcccatcc actgtgcttg 25800

gcagtaggtg atctgagaac caatggaacc aggttaatcc tgtggcactg ttgagtgagg 25860

agagcagtgg cgggcactgg aaaatatcag agacaaggca ggagacctga aatctaggct 25920

tagctcctca tatacttggc agctgtatga cctcagacaa ccagtgttac ctctctaagc 25980

ctcagtttcc tcatgcaaaa ggaggggggaa taacaacaga gcccactgct tgggggtgtt 26040

gtgaggacag gatgaaaaaa caaacagaaa tccctcagta caggattcag tgcagtggac 26100

agtcttgcaa ggtctggttc agccctccac ccctaccctc accagtataa agaactctgg 26160

cctacaagtc agatgacctg agttttaatc tcagctttgc cattagccgt gtgaacttga 26220

gaaagtccct ttccttttta catctattgg gatgatcatg catttttgt cctttattct 26280

gttaatatag tgtgttacat tgattgcttt tcatagactg aaccagcctt gtattccagg 26340

gataaatctc acttggtcat ggtgtataat cctttataca aatgttgctg ggttgagttt 26400

gctagtattt tgttgaagat ttttatgtct tgattcataa ggaatattgg tgtaccttcc 26460

ccttttatgg ccacagtttc cctacaatga tgtagtcgaa ctagacaacc tccaatatct 26520

ttcagtattc atgtcctctg attctgtgaa actaagaaaa ttaagaaata gtgattcata 26580

ggcacaaggc aggcaaaact tagactcctt gtagaataat taggaagcca aatattcagt 26640

gtgcttattt ctcaaataac cttagtttct ccagtctgcc ccaactccga ggcctgaata 26700

tctctagatg cttatgatgg caactaaagc ctaaaagcta attcatttta aagttcttcc 26760

aaatgcatag ggttttattt ttccagacct gggttcagat ggggaatttg acaaacaatg 26820

gaaaggggga aaaacaacaa tctaaacact gagtgacaaa gtaacaaaga aatagtctag 26880

ctatcagcca gtcaagccag ccttggcttt gctatccaaa gtagtcagtc taattctacc 26940

accagtttct gttcctgtag ctgtctactg cctgccaggg actctgcctt cccacccaca 27000

actaccaatg gaaggatgtg gtgaccatac cagtggctgc tgacatctcc tgccatggga 27060

agcataattg cctccagcag cctccccctt agatccatca tttttgttgc acttggcctg 27120

ggctgtactc ccggccaatg actgaacatg gtgagcatag taatgcaggc ccatttctgt 27180

gaggagcagg actcctccag taggtgactt tggctcaagg actctctatt ggcctggttg 27240

aacttttcct gaactgtgct actgtctgag actcttctta cccaatcctc tttctcgccc 27300

caattgtcac agaccacctg cattgtggtc tgagtctctc cccaccttct cttgctcttc 27360

cctgtttatc tttcacaggc atttccccca gtacattcct tgaatgtcta acccgatacg 27420

ggtgcctgac ttttggcaga cctaagcaga caaaaaggag tacttggtta cctagctctt 27480

ctttctacca caaacatcga gggaaccctt tttccctcac ccctctgcca cacccccact 27540

gccccagtga acaaccacag agagagctgt ggtataatat taggctggtg caaaagtaat 27600

tgcggttttt gccattactt ttaatggtaa aaaccgcaat tacttttgca cctacctagt 27660

atttgtgtcc ccccaaattc atatgttgaa acctaaccca caatatgatg tcattaggag 27720

gcaagacctt gaggaggtga ttagatgatg gggtggagct ctcctgaatg agattagtgc 27780

ccttataaga agaagcccaa ggaagctacc ttgactcttc catcacatga gaatgcagca 27840

agaaggcacc atctactaat caggaagaga gctctcacca gacactgaat ctgccagtgt 27900

cttgatcttg aagttcccag cctccagaac tatgcataat gcatttccat tgtctctaag 27960

ccacccagcc tatggtattt tgtcatagca gcctgaactg actaagacag tgagccacat 28020

gagaagtgcc ccaacccctc ccttaagcac ttggctcaca gatcagtggg ttcatttctg 28080

cctgagtttt attgttattc tgtagatttc ttgggctaga tatatttttc tgttattttc 28140

cttcttcacc tcagtcatga attggttgtt ttaaaaaaga caatgtaagt catggggaaa 28200

ctcctgacaa ctctactctc ctagggttcc tgataaaagg ggattcagtt gagtcctctg 28260

atggtctcta cctgccaaag tccagcagcc cttagcaaac atgctgctcg tttctgtaga 28320

gaaggtgctg gtgtcccacc atacttctct ctccctcatg aagggcttgc gacccagcaa 28380

atgggtggct tatatgggtc tgtttcaaag gaagagccag ctctgggaag aaaaacgatg 28440

agcataagca taacctacca ctgtgcctgg gaaagcagac aacttttttg atgtgtgaat 28500

atctaatgag aatggaatcc atcaattacc ttaaacttag gcacagtctt caaattcaat 28560

atatgtggga tatactttta gtcagtttgt agacgttatt tgtaataaat aatctggctt 28620

ctctaaagaa attattttaa gtgtttggtt tggtttgatt taatggtaaa attatattta 28680

gtggcagaga attatagcaa tggtgataaa ctatagagtg tcataagttc atatcttatt 28740

ctcacatttg aagctgcctg cagatgcatt caagatgcag ccagaagtca ggagactcag 28800

gctgttattt ggagctcatc attttacagc cttgctggac tcccactttc tcaggggaaa 28860

aatgtggtgt tgacccagat tagctctcca ggccctgctg agttgggcac tctgtaagct 28920

ggagggtctt ctattgtctt cacctaagtg tcaatcaaca acccaaatgg gcatggggga 28980

agagggagct gggccaatgc ccagggtgcc tggtagagag ataccttggg cactggaagg 29040

caccagcttc ccagagagaa gggggagggc catgaaaaag ttggctgtag atgccaggga 29100

cactgggact ctccagctgt gtgtttgtgt cttctgaaga cttatgtttc attcctttgg 29160

agcatgcata atcatacact gtgggatgtg ttatatagat tgcttgatag ttcaccactg 29220

taataaaata ctgtgactgg aatctgctcc cagtctgcct ttgatagcac ttgtgcaaca 29280

cacatttact gagcatttac agtgatccag gacctgtgtt gtgaaaacat tgatggacaa 29340

ggcagatggt ggagcacgtc agtgaggatt tttaacaaag gctggtaagt gctataaagg 29400

aacattgtag gacactagag aacaaagaac aggagaacct gacttaggct ggggtggggc 29460

gttggttaga ggaggctcct tggaggacat gaggtttaag ctgtgacctg aggatgaata 29520

gatgttggcc aggtgaggta ccggtatttg tcagccttac cagtaaaaaa gaaaacctat 29580

taaaaaaaaa atacacatac aaagcctcat cagccatggc ttaccagaga aagtacagcg 29640

ggcacacaaa ccacaagctc taaagtcact ctccaacctc tccacaatat atatacacaa 29700

gccctaaact gacgtaatgg gactaaagtg taaaaaatcc cgccaaaccc aacacacacc 29760

ccgaaactgc gtcaccaggg aaaagtacag tttcacttcc gcaatcccaa caagcgtcac 29820

ttcctctttc tcacggtacg tcacatccca ttaacttaca acgtcatttt cccacggccg 29880

cgccgcccct tttaaccgtt aaccccacag ccaatcacca cacggcccac acttttaaa 29940

atcacctcat ttacatattg gcaccattcc atctataagg tatattattg atgatg 29996

<210> 306

<400> 306

000

<210> 307

<400> 307

000

<210> 308

<400> 308

000

<210> 309

<400> 309

000

<210> 310

<400> 310

000

<210> 311

<400> 311

000

<210> 312

<400> 312

000

<210> 313

<400> 313

000

<210> 314

<400> 314

000

<210> 315

<400> 315

000

<210> 316

<400> 316

000

<210> 317

<400> 317

000

<210> 318

<400> 318

000

<210> 319

<400> 319

000

<210> 320

<211> 348

<212> DNA

<213> Artificial Sequence

<220><223> SV40 promoter

<400> 320

ggtgtggaaa gtccccaggc tccccagcag gcagaagtat gcaaagcatg catctcaatt 60

agtcagcaac caggtgtgga aagtccccag gctccccagc aggcagaagt atgcaaagca 120

tgcatctcaa ttagtcagca accatagtcc cgcccctaac tccgcccatc ccgcccctaa 180

ctccgcccag ttccgcccat tctccgcccc atggctgact aatttttttt atttatgcag 240

aggccgaggc cgcctctgcc tctgagctat tccagaagta gtgaggaggc ttttttggag 300

gcctaggctt ttgcaaaaag ctcccgggag cttgtatatc cattttcg 348

<210> 321

<211> 206

<212> DNA

<213> Artificial Sequence

<220><223> dESV40 promoter

<400> 321

gcatgcatct caattagtca gcaaccatag tcccgcccct aactccgccc atcccgcccc 60

taactccgcc cagttccgcc cattctccgc cccatggctg actaattttt tttatttatg 120

cagaggccga ggccgcctcg gcctctgagc tattccagaa gtagtgagga ggctttttttg 180

gaggcctagg cttttgcaaa aagctt 206

<210> 322

<211> 463

<212> DNA

<213> Homo sapiens

<400> 322

ttggcccctc cctcgggtta ccccacagcc taggccgatt cgacctctct ccgctggggc 60

cctcgctggc gtccctgcac cctgggagcg cgagcggcgc gcgggcgggg aagcgcggcc 120

cagacccccg ggtccgcccg gagcagctgc gctgtcgggg ccaggccggg ctcccagtgg 180

attcgcgggc acagacgccc aggaccgcgc tccccacgtg gcggagggac tggggacccg 240

ggcacccgtc ctgccccttc accttccagc tccgcctcct ccgcgcggac cccgccccgt 300

cccgacccct cccgggtccc cggcccagcc ccctccgggc cctcccagcc cctccccttc 360

ctttaccgcg gccccgccct ctcctcgcgg cgcgagtttc aggcagcgct gcgtcctgct 420

gcgcacgtgg gaagccctgg ccccggccac ccccgccaga tct 463

<210> 323

<211> 648

<212> DNA

<213> Schmidt-Ruppin A strain

<400> 323

acgcgtcatg tttgacagct tatcatcgca gatccgtatg gtgcactctc agtacaatct 60

gctctgatgc cgcatagtta agccagtatc tgctccctgc ttgtgtgttg gaggtcgctg 120

agtagtgcgc gagcaaaatt taagctacaa caaggcaagg cttgaccgac aattgcatga 180

agaatctgct tagggttagg cgttttgcgc tgcttcgcga tgtacgggcc agatattcgc 240

gtatctgagg ggactagggt gtgtttaggc gaaaagcggg gcttcggttg tacgcggtta 300

ggagtcccct caggatatag tagtttcgct tttgcatagg gaggggggaaa tgtagtctta 360

tgcaatactc ttgtagtctt gcaacatggt aacgatgagt tagcaacatg ccttacaagg 420

agagaaaaag caccgtgcat gccgattggt ggaagtaagg tggtacgatc gtgccttatt 480

aggaaggcaa cagacgggtc tgacatggat tggacgaacc actaaattcc gcattgcaga 540

gatattgtat ttaagtgcct agctcgatac aataaacgcc atttgaccat tcaccacatt 600

ggtgtgcacc tccaagctgg gtaccagctg ctagcaagct tgagatct 648

<210> 324

<211> 1740

<212> DNA

<213> Homo sapiens

<400> 324

gagtagctgg gattacaggc atgtgccacc acgcctcgct aatattagta tttttcatac 60

agacaagatc tcactatgtt gctcagggta gtctcgaatt ctgggactca aatgatcctc 120

ccacttcagc ctcccaaagt gctgggatta caggcataag ccatcatgcc cggcctctga 180

cgctgtttct ttcaaccccc aggatttcag attccaccag cttatggaga agggaaccaa 240

gttcgagatg cgtgattgcc cagaaagttg gaggctgagc tgagacttga acccagagac 300

cagaacctcc agaggtcaaa gtcctcctcc tgggtccccc agagaagggc cctgagatga 360

cagctcgttg gtcctcatgg aagcgtgacc cccccagtag actttctccc acacccaacc 420

ttggtttcct catctatatg atagggacaa gccagactct acctccctgg tggtcatggt 480

ctccgcttat tcgggttcat aaccttaaag gcccctcgca ccacctcagt gagccattta 540

tgcctggcac agggccaact ctcagtgcat atctgcaaag gaaccaatga atgagtgaat 600

gaagtgacaa atgaataaag gaataaatga atgaggcact tatcatgtac caggctttcg 660

ttaccacgtc ccatttattc ctctgaggca gggtctattt tatccttgtt acagatgggg 720

aaactaaggc ccagggagga gcaaagtctt ccccaagtat gtacccactc agaacttgag 780

ctctgaatgt ctcccaccca gcttagccca agagcggggt tcagtgatgc ccacccccta 840

aggctctaga gaaaggggggt aggcccacat gccagtttgg gggtggtaaa gccaggtaag 900

ttttctttat gggtcccctg aaaccctgaa agtgaacccc agtcctgcat gaaagtgagc 960

tccccatagc tcaaggtatt caagcacaat acggctttga gtgctgaagc aggctgtgca 1020

ggcttggata gtgacatgcc ctctctgagc ctcaatttcc ccacctgtca acagcagaca 1080

gtgacagctg tgatcagggg atcacagtgc atggggatgg gtgggtgcat ggggatggag 1140

gggcatttgg gagccctccc cgataccacc ccctgcagcc acccagatag cctgtcctgg 1200

cctgtctgtc ccagtccagg gctgaaaggg tgcgggtcct gcccgcccct aggtctggag 1260

gcggagtcgc ggtgacccgg gagcccaata aatctgcaac ccacaatcac gagctgctcc 1320

cgtaagcccc aaggcgacct ccagctgtca gcgctgagca cagcgcccag ggagagggac 1380

agacagccgg ctgcatggga cagcggaacc cagagtgaga ggggaggtgg caggacagac 1440

agacagcagg ggcggacgca gagacagaca gcggggacag ggaggccgac acggacatcg 1500

acagcccata gattcctaac ccagggagcc ccggcccctc tcgccgcttc ccaccccaga 1560

cggagcgggg acaggctgcc gagcatcctc ccacccgccc tccccgtcct gcctcctcgg 1620

cccctgccag cttcccccgc ttgagcacgc agggcgtccg aggacgcgct gggcctccgc 1680

acccgccctc atggaggccg tggagaccgg ggaacggccc accttcggag cctgggacta 1740

<210> 325

<211> 1075

<212> DNA

<213> Homo sapiens

<400> 325

attagagatt gtaaattggg ctctgagctt cctaccaaca aaagcacaaa ggaaaatatg 60

atcactggta ttaaaaaaaa acacctatgg tttccaaaag attaaaacaa accagcagtt 120

ttatagaagc taacactaaa atctaaagga actacgttct atggagccac ttaatatgga 180

taaacacttt gacaatattc tttcaacaac tacagtaaca agtttcttag agtccatttc 240

tttttacatc cataatgaat tgtaaatctt ttctacttct taagtaaaac atcaccactt 300

aattctggta acttttccat attaactttt tagaacaatt gcaaacgtac cataaatgat 360

tgttgtcaca gtggtaacta tttgaccctg actgttattt tgtatatagc agcttttaaa 420

ataaaaaggc aacaagtttc taggcgtaat ttccacagat cttttatgta aaacaatgac 480

atcctttgca acttctgcca tttaatctat ctcaagcaag ctctctggaa acaaatctat 540

ttgaaagatt ctattgtaat tagaaatcag ggtaactgaa tgcactagat gaaaaccttc 600

tgactggggc caatgaagtc aataaagtca aaactgctgt gaatgctcaa ctgtctgcag 660

atcagatgtc ttgggatgga atccgttctc gaggccacca tcattaatat caatttggcc 720

atgtaataca agcctcactt gttccactgt tacaaatgtg cttaaaactg agctcattta 780

caatccaaat acatatgtag gatggtaacc aaggcatcac actaatttag gtattatgtt 840

ttaggggggaa caaaaggtat gttaatattt tattcatctc caaattaact ataaattgtg 900

cattcttgca tagatcctcc ttgggaatga gaaattagga aaatccagtt gttaaaatga 960

atgcctaaaa tcaaaataaa atttgttttt ctggcacctg cttgatgaca cagactaata 1020

accaatgaca aaattccctt gaacccaagt tttcatttcc tcctattgtg tggtc 1075