



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2019-0065251
(43) 공개일자 2019년06월11일

- (51) 국제특허분류(Int. Cl.)
C12N 15/86 (2006.01) *A61K 35/76* (2015.01)
A61P 21/00 (2006.01) *A61P 25/28* (2006.01)
A61P 33/06 (2006.01) *A61P 35/00* (2006.01)
C12N 15/10 (2017.01) *C12N 15/11* (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 35/76 (2013.01)
- (21) 출원번호 10-2019-7007526
- (22) 출원일자(국제) 2017년08월18일
 심사청구일자 없음
- (85) 번역문제출일자 2019년03월14일
- (86) 국제출원번호 PCT/US2017/047687
- (87) 국제공개번호 WO 2018/035503
 국제공개일자 2018년02월22일
- (30) 우선권주장
 62/376,855 2016년08월18일 미국(US)
 (뒷면에 계속)

- (71) 출원인
 더 리젠츠 오브 더 유니버시티 오브 캘리포니아
 미합중국 캘리포니아 94607-5200 오클랜드 프랭클린 스트리트 1111, 5층
 카트레카 드루바
 미국, 92093 캘리포니아, 메일코드: 0910 라졸라, 길먼 드라이브 9500 유씨 샌디에고
 (뒷면에 계속)
- (72) 발명자
 말리 프라스한트
 미국, 92093 캘리포니아, 메일코드: 0910 라졸라, 길먼 드라이브 9500 유씨 샌디에고
 카트레카 드루바
 미국, 92093 캘리포니아, 메일코드: 0910 라졸라, 길먼 드라이브 9500 유씨 샌디에고
 콜라도 애나 모레노
 미국, 92093 캘리포니아, 메일코드: 0910 라졸라, 길먼 드라이브 9500 유씨 샌디에고
- (74) 대리인
 손민

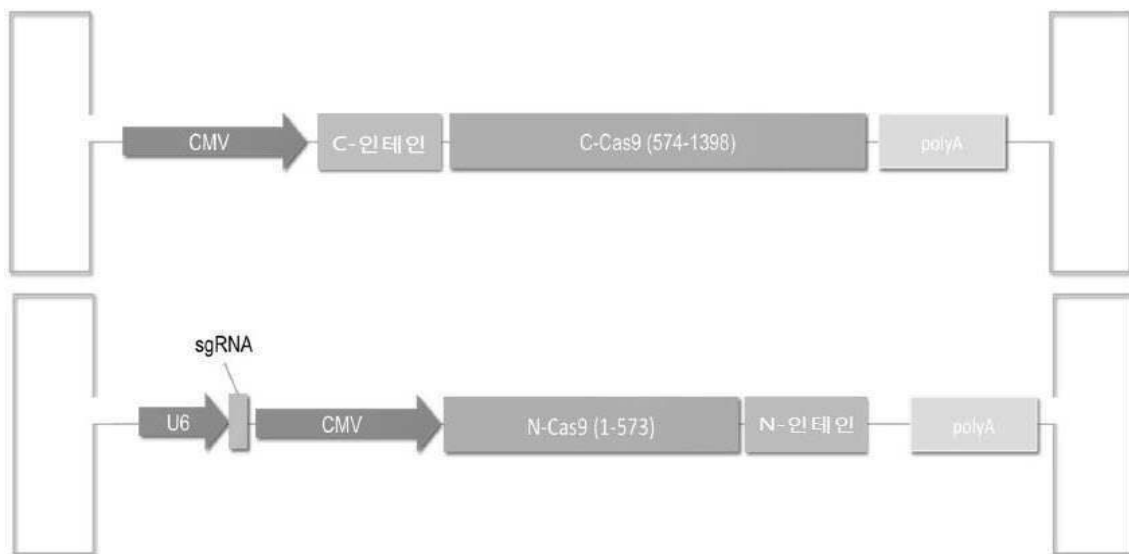
전체 청구항 수 : 총 46 항

(54) 발명의 명칭 모듈러 AAV 전달 시스템을 통한 CRISPR-Cas 개놈 가공

(57) 요약

본 개시내용은 유전자 편집의 보다 우수한 전달, 특이성 및 선택성을 허용하는 유일한 모듈러 CRISPR-Cas9 구조를 지닌 신규한 전달 시스템에 관한 것이다. 이는 이미 기술된 split-Cas9 시스템보다 유의적인 증진을 나타낸다. 모듈러 구조는 "조절가능"하다. 추가의 국면은 공간적으로 및 일시적으로 둘 다 제어됨으로써, 유도성 편집을 위한 잠재능을 생성하는 시스템에 관한 것이다. 추가의 국면은 호밍제(homing agent)에 대한 접합을 허용하는 변형된 바이러스 캡시드에 관한 것이다.

대표도 - 도2



(52) CPC특허분류

A61P 21/00 (2018.01)
A61P 25/28 (2018.01)
A61P 33/06 (2018.01)
A61P 35/00 (2018.01)
C12N 15/102 (2013.01)
C12N 15/111 (2013.01)
C12N 15/113 (2013.01)
C12N 15/90 (2013.01)
C12N 9/22 (2013.01)

(71) 출원인

콜라도 애나 모레노

미국, 92093 캘리포니아, 메일코드: 0910 라 줄라,
길먼 드라이브 9500 유씨 샌디에고

말리 프라스한트

미국, 92093 캘리포니아, 메일코드: 0910 라 줄라,
길먼 드라이브 9500 유씨 샌디에고

(30) 우선권주장

62/415,858 2016년11월01일 미국(US)
62/481,589 2017년04월04일 미국(US)

명세서

청구범위

청구항 1

(a) (i) C-인테인(C-intein)을 암호화하는 폴리뉴클레오타이드, (ii) C-Cas9를 암호화하는 폴리뉴클레오타이드, 및 (iii) 제1 벡터에 대한 프로모터 서열을 포함하는 제1 발현 벡터; 및

(b) (i) N-Cas9를 암호화하는 폴리뉴클레오타이드, (ii) N-인테인을 암호화하는 폴리뉴클레오타이드, 및 (iii) 제2 벡터에 대한 프로모터 서열을 포함하는 제2 발현 벡터를 포함하는, CRISPR-기반의 게놈 또는 에피게놈 편집(epigenome editing)을 위한 재조합 시스템(recombinant system)으로서,

여기서 임의로, 제1 및 제2 발현 벡터 둘 모두는 아데노-관련 바이러스(AAV) 또는 렌티바이러스 벡터이고,

여기서 제1 및 제2 발현 벡터의 동시-발현은 전체 Cas9 단백질의 발현을 야기하는 재조합 시스템.

청구항 2

제1항에 있어서, 제1 발현 벡터의 프로모터 서열이 CMV 프로모터를 포함하는 재조합 시스템.

청구항 3

제1항 또는 제2항에 있어서, 제2 벡터의 프로모터 서열이 gRNA 서열, 임의로 sgRNA에 작동적으로 연결된 제1 프로모터, 및 제2 프로모터를 포함하는 재조합 시스템.

청구항 4

제3항에 있어서, 제1 프로모터 서열이 U6 프로모터인 재조합 시스템.

청구항 5

제3항 또는 제4항에 있어서, 제2 프로모터 서열이 CMV 프로모터인 재조합 시스템.

청구항 6

제1항에 있어서, 제1 및 제2 발현 벡터 둘 다가 폴리-A 테일을 추가로 포함하는 재조합 시스템.

청구항 7

제1항에 있어서, 제1 발현 벡터가 테트라사이클린 반응 성분을 추가로 포함하고/하거나, 제2 발현 벡터가 테트라사이클린 조절가능한 활성화제를 추가로 포함하거나, 여기서 제1 발현 벡터가 테트라사이클린 조절가능한 활성인자를 추가로 포함하고/하거나, 제2 발현 벡터가 테트라사이클린 반응 성분을 추가로 포함하는 재조합 시스템.

청구항 8

제7항에 있어서, 테트라사이클린 반응 성분이 tetO의 하나 이상의 반복체(repeat)를 포함하는 재조합 시스템.

청구항 9

제7항에 있어서, 테트라사이클린 반응 성분이 tetO의 7개 반복체를 포함하는 재조합 시스템.

청구항 10

제7항에 있어서, 테트라사이클린 조절가능한 활성인자가 rtTa 및, 임의로, 2A를 포함하는 재조합 시스템.

청구항 11

제1항에 C-Cas9가 dC-Cas9이고 N-Cas9가 dN-Cas9인 재조합 시스템.

청구항 12

제11항에 있어서, 제1 발현 벡터 및/또는 제2 발현 벡터가 KRAB, DNMT3A, 또는 DNMT3L 중 하나 이상을 추가로 포함하는 제조합 시스템.

청구항 13

제11항에 있어서, 제1 발현 벡터 및/또는 제2 발현 벡터가 VP64, RtA, 또는 P65 중 하나 이상을 추가로 포함하는 제조합 시스템.

청구항 14

제12항에 있어서, 억제, 사일런싱(silencing), 또는 하향조절(downregulation)을 위해 표적화된 유전자에 대한 gRNA를 추가로 포함하는 제조합 시스템.

청구항 15

제13항에 있어서, 발현, 활성화, 또는 상향조절(upregulation)을 위해 표적화된 유전자에 대한 gRNA를 추가로 포함하는 제조합 시스템.

청구항 16

제15항에 있어서, 발현, 활성화, 또는 상향조절에 대해 표적화된 유전자 및, 임의로, 프로모터를 암호화하는 제3의 발현 벡터를 추가로 포함하는 제조합 시스템.

청구항 17

제1항 내지 제16항 중 어느 한 항에 있어서, 제1 발현 벡터 및/또는 제2 발현 벡터가 miRNA 회로를 추가로 포함하는 제조합 시스템.

청구항 18

제1항에 따른 제조합 발현 시스템을 포함하는 조성물로서, 여기서 제1 발현 벡터가 제1 바이러스 캡시드 속에 캡슐화(encapsulation)되며 제2 발현 벡터가 제2 바이러스 캡시드 속에 캡슐화되고, 임의로, 제1 바이러스 캡시드 및/또는 제2 바이러스 캡시드가 AAV 또는 렌티바이러스 캡시드인 조성물.

청구항 19

제18항에 있어서, AAV가 AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11, 또는 AAV-DJ 중 하나인 조성물.

청구항 20

제18항에 있어서, 제1 바이러스 캡시드 및/또는 제2 바이러스 캡시드가 비천연 아미노산, SpyTag, 또는 KTag의 그룹 중 하나 이상을 포함하도록 변형된 조성물.

청구항 21

제20항에 있어서, 비천연 아미노산이 N-엡실론-((2-아지도에톡시)카보닐)-L-라이신인 조성물.

청구항 22

제20항에 있어서, 제1 바이러스 캡시드 및/또는 제2 바이러스 캡시드가 펩타이드, 앵타머(aptamer), 올리고뉴클레오타이드, 아프이보디(affibody), DARPin, 쿠니츠 도메인(Kunitz domain), 피노머(fynomer), 비사이클릭 펩타이드, 안티칼린(anticalin), 또는 애드넥틴(adnectin) 중 하나 이상으로 슈도타이핑(pseudotyping)되는 조성물.

청구항 23

제18항에 있어서, 제1 바이러스 캡시드 및/또는 제2 바이러스 캡시드가 AAV2 캡시드인 조성물.

청구항 24

제23항에 있어서, 비천연 아미노산, SpyTag, 또는 KTag가 VP1의 아미노산 잔기 R447, S578, N587 또는 S662에서 혼입되는 조성물.

청구항 25

제18항에 있어서, 제1 바이러스 캡시드 및/또는 제2 바이러스 캡시드가 AAV-DJ 캡시드인 조성물.

청구항 26

제25항에 있어서, 비천연 아미노산, SpyTag, 또는 KTag가 VP1의 아미노산 잔기 N589에 혼입되는 조성물.

청구항 27

제18항에 있어서, 제1 바이러스 캡시드 및 제2 바이러스 캡시드가 연결되어 있는 조성물.

청구항 28

유효량의 제27항의 조성물을 이를 필요로 하는 대상체에게 투여함을 포함하는, 상기 대상체에서 통증 관리 방법으로서, 여기서 조성물은 SCN9A, SCN10A, SCN11A, SCN3A, TrpV1, SHANK3, NR2B, IL-10, PENK, POMC, 또는 MVIIA-PC 중 하나 이상을 표적화하는 gRNA를 암호화하는 벡터를 포함하는 방법.

청구항 29

유효량의 제27항의 조성물을 이를 필요로 하는 대상체에게 투여함을 포함하는, 상기 대상체에서 말라리아의 치료 또는 예방 방법으로서, 여기서 조성물은 CD81, MUC13, 또는 SR-B1 중 하나 이상을 표적화하는 gRNA를 암호화하는 벡터를 포함하는 방법.

청구항 30

유효량의 제27항의 조성물을 이를 필요로 하는 대상체에게 투여함을 포함하는, 상기 대상체에서 C형 간염의 치료 또는 예방 방법으로서, 여기서 조성물은 CD81, MUC13, SR-B1, GYPA, GYPC, PKLR, 또는 ACKR1 중 하나 이상을 표적화하는 gRNA를 암호화하는 벡터를 포함하는 방법.

청구항 31

유효량의 제27항의 조성물을 이를 필요로 하는 대상체에게 투여함을 포함하는, 상기 대상체에서 조혈 줄기 세포 요법의 면역 거부 치료 또는 예방 방법으로서, 여기서 조성물은 CCR5를 표적화하는 gRNA를 암호화하는 벡터를 포함하는 방법.

청구항 32

유효량의 제27항의 조성물을 이를 필요로 하는 대상체에게 투여함을 포함하는, 상기 대상체에서 HIV의 치료 또는 예방 방법으로서, 여기서 조성물은 CCR5를 표적화하는 gRNA를 암호화하는 벡터를 포함하는 방법.

청구항 33

유효량의 제27항의 조성물을 이를 필요로 하는 대상체에게 투여함을 포함하는, 상기 대상체에서 근육 디스트로피(muscular dystrophy)의 치료 또는 예방 방법으로서, 여기서 조성물은 디스트로핀을 표적화하는 gRNA를 암호화하는 벡터를 포함하는 방법.

청구항 34

유효량의 제27항의 조성물을 이를 필요로 하는 대상체에게 투여함을 포함하는, 상기 대상체에서 암의 치료 또는 개선 방법으로서, 여기서 조성물은 PDCD-1, NODAL, 또는 JAK-2 중 하나 이상을 표적화하는 gRNA를 암호화하는 벡터를 포함하는 방법.

청구항 35

유효량의 제27항의 조성물을 이를 필요로 하는 대상체에게 투여함을 포함하는, 상기 대상체에서 사이토크롬

p450 장애의 치료 또는 예방 방법으로서, 여기서 조성물은 CYP2D6을 표적화하는 gRNA를 암호화하는 벡터를 포함하는 방법.

청구항 36

유효량의 제27항의 조성물을 이를 필요로 하는 대상체에게 투여함을 포함하는, 상기 대상체에서 알츠하이머 질환(Alzheimer's)의 치료 또는 예방 방법으로서, 여기서 조성물은 L1rB2를 표적화하는 gRNA를 암호화하는 벡터를 포함하는 방법.

청구항 37

제29항 내지 제36항 중 어느 한 항에 있어서, 대상체가 포유동물인 방법.

청구항 38

제37항에 있어서, 포유동물이 쥐, 개, 고양이, 말, 소, 원숭이, 또는 사람 환자인 방법.

청구항 39

VP1의 아미노산 잔기 R447, S578, N587 또는 S662에서 비천연 아미노산, SpyTag, 또는 KTag를 포함하는, 변형된 AAV2 캡시드.

청구항 40

제39항에 있어서, 비천연 아미노산이 N-엡실론-((2-아지도에톡시)카보닐)-L-라이신인, 변형된 AAV2 캡시드.

청구항 41

제39항에 있어서, 변형된 AAV2 캡시드가 펩타이드, 앵타머, 올리고뉴클레오타이드, 아피보디, DARPin, 쿠니츠도메인, 피노머, 비사이클릭 펩타이드, 안티칼린, 또는 애드넥틴 중 하나 이상으로 슈도타이핑되는, 변형된 AAV2 캡시드.

청구항 42

제39항에 있어서, 리포펙타민으로 코팅된, 변형된 AAV2 캡시드.

청구항 43

VP1의 아미노산 잔기 N589에서 비천연 아미노산, SpyTag, 또는 KTag를 포함하는, 변형된 AAV-DJ 캡시드.

청구항 44

제43항에 있어서, 비천연 아미노산이 N-엡실론-((2-아지도에톡시)카보닐)-L-라이신인, 변형된 AAV2 캡시드.

청구항 45

제43항에 있어서, 변형된 AAV-DJ 캡시드가 펩타이드, 앵타머, 올리고뉴클레오타이드, 아피보디, DARPin, 쿠니츠도메인, 피노머, 비사이클릭 펩타이드, 안티칼린, 또는 애드넥틴 중 하나 이상으로 슈도타이핑되는, 변형된 AAV2 캡시드.

청구항 46

제43항에 있어서, 리포펙타민으로 코팅된, 변형된 AAV2 캡시드.

발명의 설명

기술 분야

[0001] 본 출원은 2016년 8월 18일자로 출원된 미국 특허원 일련 번호 제62/376,855호, 2016년 11월 1일자로 출원된 미국 특허원 일련 번호 제62/415,858호, 및 2017년 4월 4일자로 출원된 미국 특허원 일련 번호 제62/481,589호에 대한 35 U.S.C. 119(e) 하의 우선권을 청구하며, 이의 전문은 본원에 참고로 포함된다.

배경 기술

- [0002] 본 발명의 배경의 하기 논의는 단지 독자가 본 발명을 이해하는데 도움을 주기 위해 제공되며 본 발명에 대한 선행 기술을 설명하거나 이를 구성하기 위해 허용되는 것은 아니다.
- [0003] 크리스퍼(clustered regularly interspaced short palindromic repeats: CRISPR)-CRISPR-관련된(Cas) 시스템으로부터 기원한 RNA-안내된 효과기(effector)의 최근의 등장은 다양한 유기체의 게놈을 가공하는 능력을 변환시켜왔다.
- [0004] 현재, 아데노-관련 바이러스(Adeno-Associated Viruse: AAV)가 이들의 전체적인 안전성, 약한 면역 반응, 긴 전이유전자 발현, 고 감염 효능으로 인하여 유전자 치료요법에 광범위하게 이용되어 왔으며, 임상 시험에서 이미 사용되고 있다. 그러나, 주요 단점은 AAV가 대략 4.5 kb의 제한된 패키징 능력(packaging capacity)을 가짐으로써 이를 크기가 대략 4.2kb인, 단일 안내 RNA 백터이고, 유전자 편집에 필수적인 다른 성분을 지닌 스트렙토코쿠스 피오게네스(*Streptococcus pyogenes*) Cas9 (SpCas9)로 전달하기가 어렵다는 것이다.

발명의 내용

해결하려는 과제

- [0005] 따라서, 이러한 기술적 한계를 극복하기 위한 필요성이 당해 분야에 존재한다. 본 개시내용은 이러한 필요성을 충족시키며 또한 관련된 장점을 제공한다.
- [0006] 게놈 편집에 의해 현재 직면한 주요 과제 중 일부는: 전달, 특이성, 및 생성물 선택성이다. 본 개시내용의 국면은 이러한 챌린지를 극복하는 방법에 관한 것이다(도 1).

과제의 해결 수단

- [0007] 따라서, 일 국면에서, 본 개시내용은 CRISPR-효과기의 프로그램 가능한 혼입 및 바이러스 및 비-바이러스 전달 시도 둘 다의 장점을 통합할 목적을 지닌 손쉬운 슈도타이핑(facile pseudotyping)이 가능한 모듈러 전달 시스템(modular delivery system)에 관한 것이다.
- [0008] 질환의 유전적 및 병리학적 기반의 증가하는 지식과 결합된, CRISPR 기반 게놈 및 에피게놈 가공(epigenome engineering)을 위한 안전하고 효율적인 유전자 전달 플랫폼의 개발은 다양한 사람 질환을 표적화하고 또한 질환 내성을 가공하는 능력을 변환시킬 수 있다. 이와 관련하여, 광범위한 신규한 바이러스 및 비-바이러스 시도는 CRISPR 시약의 시험관내(*in vitro*) 및 생체내(*in vivo*) 전달을 향해 개발되어 왔다.
- [0009] 본 개시내용은 게놈 편집의 보다 우수한 전달, 특이성 및 선택성을 허용하는 유일한 모듈러 CRISPR-Cas9 구조(architecture)를 지닌 신규한 전달 시스템에 관한 것이다. 이는 기존에 기술된 split-Cas9 시스템보다 유의적인 개선을 나타낸다. 모듈러 구조는 "조절가능하다". 추가의 국면은 공간적으로 및 일시적으로 둘 다 조절될 수 있어서, 유도성 편집에 대한 잠재능을 생성할 수 있는 시스템에 관한 것이다. 추가의 국면은 호밍제(homing agent), 즉, 세포, 기관, 또는 조직에 대한 캡시드의 표적화 및/또는 국제화를 가능하도록 하는 제제에 대한 접합을 허용하는 변형된 바이러스 캡시드(viral capsid)에 관한 것이다.
- [0010] 본 개시내용의 국면은 CRISPR-기반 게놈 또는 에피게놈 편집을 위한 재조합 발현 시스템에 관한 것이다. 일부 구현예에서, 재조합 발현 시스템은: (a) (i) C-인테인을 암호화하는 폴리뉴클레오타이드, (ii) C-Cas9를 암호화하는 폴리뉴클레오타이드, 및 (iii) 제1 백터에 대한 프로모터 서열을 포함하는 제1 발현 백터; 및 (b) (i) N-Cas9를 암호화하는 폴리뉴클레오타이드, (ii) N-인테인을 암호화하는 폴리뉴클레오타이드, 및 (iii) 제2 백터에 대한 프로모터 서열을 포함하는 제2 발현 백터를 포함하거나, 대안적으로는 이로 필수적으로 이루어지거나, 여전히 추가로 이로 이루어지며, 여기서 임의로, 제1 및 제2 발현 백터는 아데노-관련 바이러스(AAV) 또는 렌티바이러스 백터이고, 여기서 제1 및 제2 발현 백터의 동시-발현은 전체 Cas9 단백질의 발현을 야기한다.
- [0011] 일부 구현예에서, 제1 발현 백터의 프로모터 서열은 CMV 프로모터를 포함하거나, 대안적으로는 이로 필수적으로 이루어지거나, 여전히 추가로 이로 이루어진다.
- [0012] 일부 구현예에서, 제2 백터의 프로모터 서열은 gRNA 서열, 임의로 sgRNA에 작동적으로 연결된 제1 프로모터, 및 제2 프로모터를 포함하거나, 대안적으로는 이로 필수적으로 이루어지거나, 여전히 추가로 이로 이루어진다. 일부 구현예에서, 제1 프로모터 서열은 U6 프로모터이다. 일부 구현예에서, 제2 프로모터 서열은 CMV 프로모터이

다.

- [0013] 일부 구현예에서, 제1 및 제2 발현 벡터 둘 다는 폴리-A 테일(poly-A tail)을 포함하거나, 대안적으로는 이로 필수적으로 이루어지거나, 여전히 추가로 이로 이루어진다.
- [0014] 일부 구현예에서, 제1 발현 벡터는 테트라사이클린 반응 성분을 포함하거나, 대안적으로는 이로 필수적으로 이루어지거나, 여전히 추가로 이로 이루어지고/지거나 제2 발현 벡터는 테트라사이클린 조절가능한 활성화제를 추가로 포함하거나, 대안적으로는 이로 필수적으로 이루어지거나, 여전히 추가로 이로 이루어지거나, 여기서 제1 발현 벡터는 테트라사이클린 조절가능한 활성인자를 추가로 포함하거나, 대안적으로는 이로 필수적으로 이루어지거나, 여전히 추가로 이로 이루어지고/지거나, 제2 발현 벡터는 테트라사이클린 반응 성분을 추가로 포함하거나, 대안적으로는 이로 필수적으로 이루어지거나, 여전히 추가로 이로 이루어진다. 일부 구현예에서, 테트라사이클린 반응 성분은 tetO의 하나 이상의 반복체, 임의로 tetO의 7개 반복체를 포함한다. 일부 구현예에서, 테트라사이클린 조절가능한 활성인자는 rtTa 및, 임의로, 2A를 포함한다.
- [0015] 일부 구현예에서, C-Cas9는 dC-Cas9이고 N-Cas9는 dN-Cas9이다. 추가의 구현예에서, 제1 발현 벡터 및/또는 제2 발현 벡터는 대안적으로는 KRAB, DNMT3A, 또는 DNMT3L 중 하나 이상을 추가로 포함하거나, 대안적으로는 이로 필수적으로 이루어지거나, 여전히 추가로 이로 이루어지고/지거나 여전히 추가로 이루어진다. 추가의 구현예에서, 재조합 발현 시스템은 억제(repression), 사일런싱(silencing) 또는 하향조절(downregulation)을 위한 표적화된 유전자에 대한 gRNA를 추가로 포함하거나, 대안적으로는 이로 필수적으로 이루어지거나, 여전히 추가로 이로 이루어진다. 다른 구현예에서, 제1 발현 벡터 및/또는 제2 발현 벡터는 VP64, RtA, 또는 P65 중 하나 이상을 추가로 포함하거나, 대안적으로는 이로 필수적으로 이루어지거나, 여전히 추가로 이로 이루어지고/지거나 여전히 추가로 이루어진다. 추가의 구현예에서, 재조합 발현 시스템은 발현, 활성화, 또는 상향조절(upregulation)을 위해 표적화된 유전자에 대한 gRNA를 추가로 포함하거나, 대안적으로는 이로 필수적으로 이루어지거나, 여전히 추가로 이로 이루어진다. 여전히 추가의 구현예에서, 재조합 발현 시스템은 발현, 활성화, 또는 상향조절을 위해 표적화된 유전자 및, 임의로, 프로모터를 암호화하는 제3의 발현 벡터를 추가로 포함하거나, 대안적으로는 이로 필수적으로 이루어지거나, 여전히 추가로 이로 이루어진다.
- [0016] 일부 구현예에서, 제1 발현 벡터 및/또는 제2 발현 벡터는 miRNA 회로를 추가로 포함하거나, 대안적으로는 이로 필수적으로 이루어지거나, 여전히 추가로 이로 이루어지고/지거나 여전히 추가로 이루어진다.
- [0017] 추가의 국면은 개시된 재조합 발현 시스템을 포함하는 조성물에 관한 것이며, 여기서 제1 발현 벡터는 제1 바이러스 캡시드 속에 캡슐화(encapsulation)되며 제2 발현 벡터는 제2 바이러스 캡시드 속에 캡슐화되고, 임의로, 제1 바이러스 캡시드 및/또는 제2 바이러스 캡시드는 AAV 또는 렌티바이러스 캡시드이다. 일부 구현예에서, AAV는 AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11, 또는 AAV-DJ 중 하나이다.
- [0018] 일부 구현예에서, 제1 바이러스 캡시드 및/또는 제2 바이러스 캡시드는 다음의 그룹 중 하나 이상을 포함하도록 변형된다: 비천연 아미노산, SpyTag, 또는 KTag. 일부 구현예에서, 비천연 아미노산은 N-엡실론-((2-아지도에톡시)카보닐)-L-라이신이다.
- [0019] 일부 구현예에서, 제1 바이러스 캡시드 및/또는 제2 바이러스 캡시드는 펩타이드, 앵타머(aptamer), 올리고뉴클레오타이드, 아프이보디(affibody), DARPin, 쿠니츠 도메인(Kunitz domain), 피노머(fynomer), 비사이클릭 펩타이드, 안티칼린(anticalin), 또는 애드넥틴(adnectin) 중 하나 이상으로 슈도타이프(pseudotype)된다.
- [0020] 일부 구현예에서, 제1 바이러스 캡시드 및/또는 제2 바이러스 캡시드는 AAV2 캡시드이다. 추가의 구현예에서, 비천연 아미노산, SpyTag, 또는 KTag는 VP1의 아미노산 잔기 R447, S578, N587 또는 S662에서 혼입된다.
- [0021] 일부 구현예에서, 제1 바이러스 캡시드 및/또는 제2 바이러스 캡시드는 AAV-DJ 캡시드이다. 추가의 구현예에서, 비천연 아미노산, SpyTag, 또는 KTag는 VP1의 아미노산 잔기 N589에 혼입된다.
- [0022] 일부 구현예에서, 제1 바이러스 캡시드 및 제2 바이러스 캡시드는 연결되어 있다.
- [0023] 본 개시내용의 일부 국면은 유효량의 개시된 조성물을 이를 필요로 하는 대상체에게 투여함을 포함하는, 상기 대상체에서의 통증 관리 방법에 관한 것이며, 여기서 조성물은 SCN9A, SCN10A, SCN11A, SCN3A, TrpV1, SHANK3, NR2B, IL-10, PENK, POMC, 또는 MVIIA-PC 중 하나 이상을 표적화하는 gRNA를 암호화하는 벡터를 포함한다.
- [0024] 본 개시내용의 일부 국면은 유효량의 개시된 조성물을 이를 필요로 하는 대상체에게 투여함을 포함하는, 상기 대상체에서 말라리아의 치료 또는 예방 방법에 관한 것이며, 여기서 조성물은 CD81, MUC13, 또는 SR-B1 중 하나

이상을 표적화하는 gRNA를 암호화하는 벡터를 포함한다.

- [0025] 본 개시내용의 일부 국면은 유효량의 개시된 조성물을 이를 필요로 하는 대상체에게 투여함을 포함하는, 상기 대상체에서 C형 간염의 치료 또는 예방 방법에 관한 것이며, 여기서 조성물은 CD81, MUC13, SR-B1, GYPA, GYPC, PKLR, 또는 ACKR1 중 하나 이상을 표적화하는 gRNA를 암호화하는 벡터를 포함한다.
- [0026] 본 개시내용의 일부 국면은 유효량의 개시된 조성물을 이를 필요로 하는 대상체에게 투여함을 포함하는, 상기 대상체에서 조혈 줄기 세포 요법의 면역 거부 치료 또는 예방 방법에 관한 것이며, 여기서 조성물은 CCR5를 표적화하는 gRNA를 암호화하는 벡터를 포함한다.
- [0027] 본 개시내용의 일부 국면은 유효량의 개시된 조성물을 이를 필요로 하는 대상체에게 투여함을 포함하는, 상기 대상체에서 HIV의 치료 또는 예방 방법에 관한 것이며, 여기서 조성물은 CCR5를 표적화하는 gRNA를 암호화하는 벡터를 포함한다.
- [0028] 본 개시내용의 일부 국면은 유효량의 개시된 조성물을 이를 필요로 하는 대상체에게 투여함을 포함하는, 상기 대상체에서 근육 위축의 치료 또는 예방 방법에 관한 것이며, 여기서 조성물은 디스트로핀을 표적화하는 gRNA를 암호화하는 벡터를 포함한다.
- [0029] 본 개시내용의 일부 국면은 유효량의 개시된 조성물을 이를 필요로 하는 대상체에게 투여함을 포함하는, 상기 대상체에서 암의 치료 또는 개선 방법에 관한 것이며, 여기서 조성물은 PDCD-1, NODAL, 또는 JAK-2 중 하나 이상을 표적화하는 gRNA를 암호화하는 벡터를 포함한다.
- [0030] 본 개시내용의 일부 국면은 유효량의 개시된 조성물을 이를 필요로 하는 대상체에게 투여함을 포함하는, 상기 대상체에서 사이토크롬 p450 장애의 치료 또는 예방 방법에 관한 것이며, 여기서 조성물은 CYP2D6을 표적화하는 gRNA를 암호화하는 벡터를 포함한다.
- [0031] 본 개시내용의 일부 국면은 유효량의 개시된 조성물을 이를 필요로 하는 대상체에게 투여함을 포함하는, 상기 대상체에서 알츠하이머 질환(Alzheimer's disease)의 치료 또는 예방 방법에 관한 것이며, 여기서 조성물은 LILRB2에서 표적화하는 gRNA를 암호화하는 벡터를 포함한다.
- [0032] 개시된 방법 중 임의의 하나 이상 중 일부 구현예에서, 대상체는 포유동물, 임의로 쥐, 개, 고양이, 말, 소, 원숭이, 또는 사람 환자이다.
- [0033] 추가의 국면은 VP1의 아미노산 잔기 R447, S578, N587 또는 S662에서 비천연 아미노산, SpyTag, 또는 KTag를 포함하는 변형된 AAV2 캡시드에 관한 것이다. 일부 구현예에서, 비천연 아미노산은 N-엡실론-((2-아지도에톡시)카보닐)-L-라이신이다. 일부 구현예에서, 변형된 AAV2 캡시드는 펩타이드, 앵타머, 올리고뉴클레오타이드, 아피보디, DARPin, 쿠니츠 도메인, 피노머, 비사이클릭 펩타이드, 안티칼린, 또는 애드벡틴 중 하나 이상으로 슈도파이프된다. 일부 구현예에서, 변형된 AAV2 캡시드는 리포펙타민으로 코팅된다.
- [0034] 추가의 국면은 VP1의 아미노산 잔기 N589에서 비천연 아미노산, SpyTag, 또는 KTag를 포함하는 변형된 AAV-DJ 캡시드에 관한 것이다. 일부 구현예에서, 비천연 아미노산은 N-엡실론-((2-아지도에톡시)카보닐)-L-라이신이다. 일부 구현예에서, 변형된 AAV-DJ 캡시드는 펩타이드, 앵타머, 올리고뉴클레오타이드, 아피보디, DARPin, 쿠니츠 도메인, 피노머, 비사이클릭 펩타이드, 안티칼린, 또는 애드벡틴 중 하나 이상으로 슈도타이프된다. 일부 구현예에서, 변형된 AAV-DJ 캡시드는 리포펙타민으로 코팅된다.

도면의 간단한 설명

- [0035] 도 1은 본원에 의해 지정된 CRISPR 전달 및 국면과 관련된 챌린지(challenge)를 묘사하는 차트이다.
- 도 2는 동시-발현시 재구성된, 각각 split-인테인, split-Cas9를 전달하는 예시적인 이중-AAV의 대략도를 나타낸다.
- 도 3은 예시적인 유도성 Split-Cas9 시스템의 개략도를 나타낸다.
- 도 4는 (A)가 KRAB 억제인자 도메인(repressor domain)을 지닌, 유전자 억제용의 예시적인 split-Cas9 시스템을 나타내고 (B)는 VP64 및 Rta 도메인을 지닌, 유전자 활성화용의 예시적인 split-Cas 시스템을 나타낸다.
- 도 5는 miRNA 회로를 지닌 이중 AAV의 예시적인 개략도를 나타낸다.
- 도 6은 바이러스-앵타머-세포 상호작용의 개략도를 나타낸다.

도 7은 (A) 예시적인 TK-GFP 벡터 개략도 및 (B) 다양한 감염 다중도(multiplicities of infection: MOI)에서 HEK293T 세포의 AAV-DJ TK-GFP 형질도입에 대한 병합된 형광성 및 상 현미경 영상을 나타낸다.

도 8은 (A) 독시사이클린이 투여되지 않은, ApoB를 표적화하는 AAV8 유도성 이중-Cas9 시스템이 투여된 3마리의 마우스, (B) APOB를 표적화하는 AAV8 유도성 이중-Cas9 시스템이 투여되고 200 mg의 독시사이클린이 주당 3회로 4주 동안 함께 투여되며, 독시사이클린과 함께 투여된 경우 1.7% 인델 형성(indel formation)을 나타내는 3마리의 마우스를 나타낸다.

도 9는 CXCR4를 표적화하는 시험관내 억제제를 나타낸다. 293T 세포는 이중-AAVDJ split-Cas9 바이러스로 형질유도되며, 세포는 3일 후에 수집하여, RNA를 추출하고 RT-qPCR을 수행하였다.

도 10은 생체내 CD81 억제를 나타내며, 생체내 억제를 위해, 3마리의 마우스에게 pAAV8_gCD81_KRAB_dCas9 벡터를 투여하였다. 간을 AAV 투여 4주 후에 수거하고, RNA를 추출하며, RT-qPCR 실험을 수행하였다. 결과는 억제 벡터를 투여한 마우스 대 야생형으로부터의 CD81 유전자의 35% 억제를 나타낸다.

도 11은 항-CD81로 염색한 간을 나타낸다. 상부로부터 하부까지: 주요 항체 대조군은 없고, 마우스에게 AAV8 gCD81 억제 split-Cas9 벡터, 야생형 대조군을 투여하였다.

도 12는 (a) AAVDJ_VR_dCas9 벡터를 사용하여 RT-qPCR 측정된 바와 같은 시험관내 RHOX 활성화의 증거를 나타내는 dC-Cas9_V를 사용한 시험관내 활성화를 나타낸다. 대조군은 AAVS1 유전자자리(locus)를 표적화하는 gRNA로 이루어지고; (b)는 AAVDJ_VR_dCas9 벡터를 사용한 RT-qPCR에 의해 측정된 바와 같은 시험관내 ASCL1 활성화의 증거를 나타낸다.

도 13은 (A) UAA 농도가 변하는 음성 대조군(UAA의 부재하에서)에 대해 GFP+ 세포 표준화된 wrt의 수를 나타내는 막대그래프 및 (B) 신타타제 농도가 변하는 음성 대조군에 대해 GFP+ 세포 표준화된 wrt의 수를 나타내는 막대그래프를 나타낸다.

도 14는 동일한 용적의 상이한 돌연변이체에 의해 형질유도된 세포%를 나타내는 막대그래프를 나타낸다.

도 15는 동일한 용적의 상이한 변이체에 의해 형질유도된 세포의 %를 나타내는 막대그래프를 나타낸다.

도 16는 모듈러 split-Cas9 이중 AAV 시스템을 통한 다목적 게놈 가공(versatile genome engineering)을 나타낸다: (a) 게놈 편집을 위한 인테그린-매개된 split-Cas9 pAAV의 예시적인 개략도, 좌측, 및 일시적인 유도성 게놈 가공, 우측. (b) 좌측으로부터 우측으로, HEK293T 세포내에서 시험관내, CD34+ 조혈 줄기 세포에서 생체외(ex vivo), 및 ApoB 유전자자리에서 생체내 AAVS1 유전자자리에서 인델(indel) 빈도. (c) 독시사이클린(dox:200 µg/ml)이 보충된 배지인, 유도성-Cas9(iCas9) AAV와 비교한 Cas9 AAV를 사용한 시험관내 AAVS1 유전자자리의 상대적인 활성. (d) Cas9 AAV와 유도성 Cas9 AAV 사이의 생체내 ApoB 편집의 상대적인 활성. 마우스는 iCas9 AAV로 형질유도되었으며, 여기서 독시사이클린의 존재 또는 부재하에 염수가 투여되었다(dox: 200 mg; 총 12회 주사; 오차 바아는 SEM이다). (e) dCas9-KRAB 억제 융합 단백질을 통한, 게놈 억제의 예시적인 개략도, 및 dCas9-VP64-RTA 융합 단백질을 통한 게놈 활성화의 개략도. (f) 2개의 명백한 스페이서를 표적화하는, HEK293T 세포내에서 시험관내 CXCR4 억제의 증거. (g) 성체 마우스 간에서 생체내 CD81 억제의 증거. (h) 이중-gRNA를 사용한 시험관내 ASCL1 활성화의 증거. (i) 성체 마우스 간에서 생체내 Afp 활성화의 증거. (j) 간 단면의 대표적인 면역형광성 염색 및 상대적인 발현 수준의 상응하는 정량적 분석이 나타나 있다: DAPI(하부 패널) 및 항-CD81(상부 패널). 좌측 패널은 음성 대조군(제2 항체 염색된 단면)이고, 중간 패널은 양성 대조군(비-표적화 AAV)이며, 우측 패널은 CD81 AAV로 형질도입된 마우스이다.(스케일 바아(scale bar): 250 µm; 오차 바아는 SEM이다).

도 17은 클릭-화학 핸들(click-chemistry handle)의 UAA 매개된 혼입을 통한 다목적의 캡시드 슈도타이핑을 나타낸다: (a) 바이러스 캡시드에 UAA를 첨가하고, UAA에 대한 효과기의 후속적인 클릭-화학 기반 화학 연결을 위한 시도의 예시적인 개략도. (b) UAA와의 교체를 위해 검정된 표면 잔기의 위치(번호매긴 VP1 잔기. (c) 2mM UAA (0.4mM 라이신)의 존재 및 부재하에서 AAV2 돌연변이체의 상대적인 역가(titer): 293T 세포를 동량의 바이러스로 형질도입하고 형광성 세포의 수를 정량화하였다; UAA의 부재하에서 바이러스 조립은 관찰되지 않았다. (d) Alexa594 DIBO 알킨을 통한 AAV의 형광단 슈도타이핑을 수행하였다: 바이러스 위로의 성공적인 연결은 293T의 형질감염 2시간 후에 바이러스의 형광성 가시화를 통해 확인하였다(스케일 바아: 250 µm). (e) 알킨-태그된 올리고뉴클레오타이드를 통한 AAV의 올리고뉴클레오타이드 슈도타이핑을 수행하였다: 상응하는 상보성 올리고뉴클레오타이드를 지닌 AAV의 DNA 배열 스폿(spot) 위에서의 선택적인 포획은 이들 스폿에 분산된 293T 세포의 특이

적인 바이러스 형질도입을 통해 입증되었다(스케일 바아: 250 μm). (f) 완전히 프로그램가능한 모듈러 AAV를 생성하기 위한 게놈 가공 효과기 및 캡시드 효과기에서 프로그램작동가능성을 조합시키는 통합된 모듈러 AAV 플랫폼의 개념. (g) mAAV 통합된 시스템이 기능성인지, 즉, UAA 변형된 AAV가 split-Cas9 기반 게놈 가공 페이로드(payload)를 혼입하여 강한 게놈 편집을 달성할 수 있는지를 확인: 인텔 신호 및 대표적인 NHEJ 프로파일이 나타나 있다.

도 18은 mAAV를 통한 생체내 및 시험관내 게놈 조절을 나타낸다: (a) 생체내 mAAV-매개된 게놈 가공을 위한 작업 흐름의 예시적인 개략도: AAV 플라스미드를 설계하고 작제한 후, 바이러스 생산 및 요덕사놀 구매를 통해 정제한다. 이후에, 마우스에게 $\sim 0.5\text{E}12\text{-}1\text{E}12$ GC를 꼬리 정맥 또는 복강내 경로를 통해 주사하고 전체 조직을 4주째에 가공을 위해 수거한다. (b) 생체내 CD81 억제: 마우스에게 복강내(IP) 주사에 의해 AAV를 표적화하지 않는 1E12 GC 또는 AAV를 표적화하는 CD81을 제공하였다. 전체 조직 수준에서 CD81의 ~ 40 내지 60% 억제를 정량적 RT-PCR을 통해 본 실험에서 관찰하였다. (c) 좌측: 2개의 명백한 스페이서 gRHOF2_1 및 gRHOF2_2 뿐만 아니라 둘 다의 조합인, 이중-gRHOF2의 표적화를 통한 293T 세포내에서의 시험관내 RHOF2 활성화. $\sim 1.25\text{-}7$ 배의 활성화가 이러한 상이한 조건 하에서 정량적 RT-PCR을 통해 관찰되었다: 간에서 생체내 Afp 활성화: 마우스에게 IP 주사에 의해 비표적화 1E12 GC 또는 Afp AAV를 제공하였다. 전체 조직 수준에서 Afp의 ~ 1.25 내지 3배 활성화가 정량적 RT-PCR을 통해 본 실험에서 관찰되었다.

도 19는 UAA 혼입의 최적화를 나타낸다: 신테타제 및 UAA 농도: (a) Y39에서 TAG 정지 부위를 지닌 GFP 리포터 서열 내로 UAA 도입: 형질감염 48시간 후 293T 세포의 형광 영상을 상이한 실험 조건 하에서 나타낸다 - 음성 대조군, wt-GFP 형질감염, 및 2mM UAA (N-엡실론-((2-아지도에톡시)카보닐)-L-라이신; 구조 나타냄)의 부재 또는 존재하에서 GFP-Y39TAG 리포터 cum tRNA-tRNA 신테타제 형질감염. 후자의 조건에서 UAA 혼입은 강한 GFP 발현을 회복한다. (b) UAA 혼입에 있어서 신테타제 양의 역할: 리포터 플라스미드(2mM UAA 하에서)에 대해 상대적인 tRNA-tRNA 신테타제 플라스미드의 최적화를 수행하였다. 5:1 비는 1:1 비에 비해서 거의 5배 더 높은 UAA 혼입을 나타내었다. (c) UAA 혼입에 있어서 UAA 농도의 최적화: 5:1 비의 tRNA-tRNA 신테타제 대 리포터 플라스미드의 존재하에서 UAA 농도의 범위를 평가하였다. 고 농도의 UAA에서 배양물 속에 보다 큰 세포 사멸이 존재하였지만, 혼입 효능에 있어서 유의적인 차이는 관찰되지 않았다.

도 20은 AAV 표면에 모이어티(moiety)의 클릭-화학 매개된 용이한 연결을 통한 다목적의 캡시드 슈도타이핑을 나타낸다. (a) 동일한 배양 조건 하에서 생산된 AAV2-N587UAA 및 AAV-DJ-N589UAA의 바이러스 역가의 비교. (b) UAA 혼입이 AAV 활성화에 영향을 미치지 않음을 확인(실험은 293T에서 수행함). (c) 항체 중화에 대한 '차폐된 AAV' 내성의 표시. (d) AAV-DJ-N589UAA 바이러스의 상대적인 활성(mCherry 발현을 통해 검정됨)을 돼지 혈청에 대한 노출 후 소 분자 및 중합체 모이어티의 범위에 대해 구속하였다.

도 21은 AAV-CRISPR 억제 및 활성화에 대한 도메인 최적화를 나타낸다: (a) AAV-CRISPR 억제에 대한 도메인 최적화: 다중 C 말단 도메인 융합체의 활성: KRAB 또는 DNA 메틸트랜스퍼라제(DNMT3A 또는 DNMT3L)를 평가하나, 일시적인 억제 검정에서 유의적인 추가의 억제는 관찰되지 않았다(오차 바아는 SEM이다; 세포: HEK293T, 유전자 자리: CXCR4) (b) AAV-CRISPR 활성화를 위한 도메인 최적화: 다수의 N 말단 도메인 융합체의 활성: VP64 및 P65를 평가하였으며, 특히 VP64 도메인의 첨가는 ~ 4 -배 더 높은 유전자 발현을 생성하였다. (오차 바아는 SEM이다; $p=0.0007$; HEK293Ts, 유전자자리: ASCL1).

도 22는 (a) 게놈 조절을 위한 인테인-매개된 split-dCas9 pAAV의 개략도. (b) KRAB-dCas9-Nr1 억제인자 융합 단백질을 통한 게놈 억제, 및 dCas9-VP64-RTA 융합 단백질을 통한 게놈 활성화를 가능하도록 하는 효과기 카세트의 모듈러 사용을 위한 시도. (c) 성체 마우스 간에서 생체내 Afp 활성화의 증거를 나타낸다. 대조군 마우스에게는 동일한 역가, $5\text{E}+11$ vg/마우스로 비-표적화 AAV8 바이러스를 제공하였다. (오차 바아는 SEM이다; $p=0.0117$). (d) 시험관내 활성화를 위한 도메인의 최적화 후(상기 새로운 도 1), VP64 활성화 도메인을 dCas9 벡터에 가하고 생체내 Afp 활성화 실험을 AAV8 $5\text{E}+11$ vg/마우스를 제공받는 마우스에서 반복하였다. 대조군 마우스에게 동일한 역가, $5\text{E}+11$ vg/마우스에서 비-표적화 AAV8 바이러스를 제공하였다. >6 배의 활성도가 추가의 VP64 도메인을 지닌 Afp에서 관찰되었다. (오차 바아는 SEM이다; $p=0.0271$).

도 23은 mdx 마우스에서 Split-Cas9 이중 AAV 시스템이 디스트로핀 발현을 구조(rescue)함을 나타낸다. (a) Mdx 마우스 모델은 엑손 23에서 예비성숙 정지 코돈(stop codon)을 갖는다. 단일 또는 이중-gRNA Cas9 시스템을 사용하여 2개의 상이한 시도를 이용하였다. 단일-gRNA를 설계하여 엑손 23의 상부 코돈을 표적화하였다. 이중-gRNA를 설계하여 엑손 23의 상부 및 하부를 표적화함으로써, 돌연변이된 엑손 23의 절개를 이끔으로써, 디스트로핀 유전자의 관독 프레임을 회수하고 단백질 발현을 회복하였다. (b) mdx 마우스에서 디스트로핀 면역형광성을

엑손 23 결실용 $1E+12$ vg/마우스 AAV8 split-Cas9 이중 gRNA 시스템으로 형질도입하였다. (디스트로핀, 상부 3개 패널; 핵, 4',6'-디아미디노-2-페닐인돌(DAPI), 하부 3개 패널; 스케일 바: $250 \mu\text{m}$), (c) Dmd 편집용 표적 서열의 목록. 엑손 23의 gRNA-L 및 gRNA-R 가공 절개부, 및 gRNA-T는 엑손 23에서 예비성숙 정지 코돈을 표적화한다; PAM은 밑줄쳐져 있고; 암호화 서열은 KD단의 경우에 존재하고 인트론 서열은 하단 경우에 존재한다. (d) 디스트로핀에 대한 웨스턴 블롯(western blot)은 디스트로핀 발현의 회복을 나타낸다. WT 마우스로부터의 단백질에 대한 비교는 회복된 디스트로핀이 이중-gRNA 및 단일-gRNA 방법 둘 모두에 대해 정상적인 양의 약 ~7 내지 10% 임을 입증한다.

도 24는 통증 관리에 관한 것이다: 마우스에게 $1E+12$ vg/마우스의 AAV5 Nav 1.7 KRAB 억제 작제물(dCas9)을 척추강내 주사하였다. 관찰한 바와 같이, 약 70% 억제가 SCN9A 유전자(Nav 1.7)에서 관찰되며, Nav 1.8이 억제 신호를 나타내지 않으므로, 특이적인 것으로 밝혀졌다. 이는 후근 신경절(DRG)을 표적화하는 작제물의 생체내 가능성을 입증한다.

도 25는 mCherry를 발현하는 $1E+12$ vg/마우스의 다양한 혈청형(AAV5, AAV1, AAV8, AAV9, AAVDJ)을 복강내 주사한 마우스에서 mCherry 발현을 나타낸다. 마우스의 그룹에 4주 동안 주당 1회 $1E+12$ vg/마우스 AAV5 mCherry(상기 AAV5_multiple)를 척추강내 주사하였다. 알 수 있는 바와 같이, AAV9 및 AAVDJ는 다른 혈청형과 비교하여 보다 높은 형질도입 효능을 나타낸다.

도 26은 SpyTag 및 KTag 또는 슈도타이핑된 하이브리드화 올리고뉴클레오타이드를 사용하여 2개의 AAV 캡시드를 연결하는 개략도이다.

도 27은 아지드-알킨 반응 또는 SpyTag 및 KTag으로 비천연 아미노산을 사용한 슈도타이핑의 일반적인 패러다임을 나타내는 개략도이다.

도 28은 (a) AAV2-N587UAA 및 AAV-DJ-N589UAA(오차 바아는 +/- SEM이다)의 바이러스 역가의 비교 및 (b) UAA 혼입이 AAV 활성화에 부정적으로 영향을 미치지 않는지의 확인(실험은 다양한 vg/세포로 HEK 293T에서 수행하였다)을 나타낸다(오차 바아는 +/- SEM이다).

도 29는 (a) AAVDJ 및 AAVDJ-N589UAA의 SDS-PAGE 분해된 캡시드 단백질의 꼬마지에 염색, (b) 알킨-올리고뉴클레오타이드(10 kDa)의 처리에 이은 AAVDJ 및 AAVDJ-N589UAA의 SDS-PAGE 분해된 캡시드 단백질의 꼬마지에 염색, 및 (c) 알킨-올리고뉴클레오타이드를 사용한 처리에 이은 변성되지 않은 AAV-DJ 및 AAV-DJN589UAA의 웨스턴 블롯, 및 상보성 올리고뉴클레오타이드-바이오틴 접합체 및 이어서 스트렙타비딘-HRP를 사용한 프로브화를 나타낸다.

도 30은 AAV 표면에 대한 효과기의 클릭-화학 매개된 연결을 통한 다목적 캡시드 슈도타이핑을 나타낸다: (a) 항체 중화에 대한 '숨겨진(cloaked) AAV' 내성의 표시. (b) HEK 293T 세포의 AAV-mCherry 기반 형질도입을 통해 검정된 돼지 혈청에 대한 노출 후 광범위한 작은 분자 및 중합체 모이어티에 대해 구속된 AAVDJ 및 AAVDJ-N589UAA 바이러스의 상대적인 활성화. (c) HEK 293T 세포의 AAV-mCherry 기반 형질도입을 통해 검정된 돼지 혈청에 대한 노출 후 광범위한 소 분자 및 중합체 모이어티에 대해 구속된 AAVDJ 및 AAVDJ-N589UAA 바이러스의 상대적 활성화. (d) HEK 293T 세포($1E+5$ vg/세포)에서 AAVDJ-N589UAA, AAVDJ-N589UAA+올리고, 및 AAVDJ-N589UAA+올리고+리포펙타민의 AAVSI 편집율(NHEJ 현상%).

도 31은 AAV 내로 UAA 혼입의 최적화: (a) UAA 혼입의 신테타제 양의 역할: 리포터 플라스미드(2mM UAA)에 대해 상대적인 tRNA 및 tRNA 신테타제 플라스미드의 양의 최적화를 수행하였다. 5:1 비는 1:1 비와 비교하여 거의 5배 더 높은 UAA 혼입을 나타내었다. (b) UAA 혼입에 있어서 UAA 농도의 최적화: 5:1 비의 tRNA 및 tRNA 신테타제 대 리포터 플라스미드의 존재하에서 UAA 농도의 범위를 평가하였다. 고 농도의 UAA에서 배양물 속에서 보다 큰 세포 사멸이 있었지만 혼입 효능에 있어서 유의적인 차이는 관찰되지 않았다. (c) eTF1-E55D의 존재하에서 UAA-AAV 역가에 있어 1.5 내지 4배 증가가 관찰되었다.

도 32는 세포주에 걸친 '숨겨진(cloaked) AAV'의 형질도입 효능을 나타낸다: 구체적으로 다양한 세포주에서 AAV-DJ-N589UAA 및 AAV-DJ-N589UAA+올리고+리포펙타민의 형질도입 효능.

도 33은 gRNA 작제물이 MCP-VP64를 보충하는 gRNA-M2M 및 Com-KRAB를 보충하는 gRNA-Com을 통한 내인성 사람 유전자에서 동시 활성화 및 억제를 매개하는 방법의 개략도를 나타낸다.

도 34는 동시 활성화 및 억제에 대한 벡터 설계(2개 벡터 시스템)를 나타낸다.

도 35는 유전자 억제 및 유전자 과발현에 대한 3개의 벡터 시스템을 나타낸다. 마우스에게 유전자 억제를 위한

본 발명자의 split-Cas9 시스템(벡터 a 및 b)(gRNA는 상이한 유전자를 표적화하기 위해 교체될 수 있다) 및 과 발현을 위한 목적한 CMV 프로모터 및 유전자(벡터 c)를 함유하는 제3 벡터로 척추강내 주사될 것이다.

도 36은 염기 편집 모델(base editing model)을 포함하는 split-Cas 시스템의 개략도를 나타낸다.

도 37a는 다음의 성분을 포함하는 이중 AAV(pX600) 시스템내 2개의 벡터 중 하나에 대한 예시적인 서열이다: CMV 프로모터, dCInteinCCas9, KRAB, 및 PolyA. 도 37b는 도 37a에서의 서열의 밑줄치고/치거나 강조된 부위 각각에 대한 주석 정보를 제공한다. 도 37c는 도 37a에 의해 암호화된 작제물의 그래프 맵(graphical map)이다.

도 38a는 다음의 성분을 포함하는 이중 AAV(pX600) 시스템에서 2개의 벡터 중 하나에 대한 예시적인 서열이다: CMV 프로모터, dCInteinCCas9, DNMT3L, 및 PolyA. 도 38b는 도 38a에서 서열의 밑줄치고/치거나 강조된 부위 각각에 대한 주석 정보를 제공한다. 도 38c는 도 38a에 의해 암호화된 작제물(construct)의 그래프 맵이다.

도 39a는 다음의 성분을 포함하는 이중 AAV(pX600) 시스템에서 2개 벡터 중 하나에 대한 예시적인 서열이다: CMV 프로모터, dCInteinCCas9, DNMT3A, 및 PolyA. 도 39b는 도 39a에서 서열의 밑줄치고/치거나 강조된 부위 각각에 대한 주석 정보를 제공한다. 도 39c는 도 39a에 의해 암호화된 작제물의 그래프 맵이다.

도 40a는 다음의 성분을 포함하는 이중 AAV(Custom) 시스템에서 2개 벡터 중 하나에 대한 예시적인 서열이다: U6 프로모터에 이은 안내 RNA 클로닝 부위, CMV 프로모터, CP64, 및 dNCas9NIntein. 도 40b는 도 40a에서 서열의 밑줄치고/치거나 강조된 부위 각각에 대한 주석 정보를 제공한다. 도 40c는 도 40a에 의해 암호화된 작제물의 그래프 맵이다.

도 41a는 다음의 성분을 포함하는 이중 AAV(Custom) 시스템에서 2개 벡터 중 하나에 대한 예시적인 서열이다: U6 프로모터에 이은 안내 RNA 클로닝 부위, CMV 프로모터, CP65, 및 dNCas9NIntein. 도 41b는 도 41a에서 서열의 밑줄치고/치거나 강조된 부위 각각에 대한 주석 정보를 제공한다. 도 41c는 도 41a에 의해 암호화된 작제물의 그래프 맵이다.

도 42a는 다음의 성분을 포함하는 이중 AAV 시스템에서 2개 벡터 중 하나에 대한 예시적인 서열이다: miRNA 인식 부위, Zac, iU6 프로모터, gSa, CMV 프로모터, 및 tTRKRAB. 도 42b는 도 42a에서 서열의 밑줄치고/치거나 강조된 부위 각각에 대한 주석 정보를 제공한다. 도 42c는 도 42a에 의해 암호화된 작제물의 그래프 맵이다.

도 43a는 다음의 성분을 포함하는 이중 AAV 시스템에서 2개 벡터 중 하나에 대한 예시적인 서열이다: tet0(Custom), U6 프로모터에 이은 안내 RNA 클로닝 부위, CMV 프로모터, NCas9NIntein, 및 M2rtTA. 도 43b는 도 43a에서 서열의 밑줄치고/치거나 강조된 부위 각각에 대한 주석 정보를 제공한다. 도 43c는 도 43a에 의해 암호화된 작제물의 그래프 맵이다.

도 44a는 다음의 성분을 포함하는 이중 AAV 시스템에서 2개 벡터 중 하나에 대한 예시적인 서열이다: tet0, CBL, 및 iCInteinCCas9. 도 44b는 도 44a에서 서열의 밑줄치고/치거나 강조된 부위 각각에 대한 주석 정보를 제공한다. 도 44c는 도 44a에 의해 암호화된 작제물의 그래프 맵이다.

도 45a는 다음 성분을 포함하는 이중 AAV(pX600) 시스템에서 2개 벡터 중 하나에 대한 예시적인 서열이다: CMV 프로모터, CIntein-CCas9, BE3C, 및 PolyA. 도 45b는 도 45a에서 서열의 밑줄치고/치거나 강조된 부위 각각에 대한 주석 정보를 제공한다. 도 45c는 도 45a에 의해 암호화된 작제물의 그래프 맵이다.

도 46a 및 도 46b는 다음 성분을 포함하는 이중 AAV(Custom) 시스템에서 2개 벡터 중 하나에 대한 예시적인 서열을 제공한다: U6 프로모터에 이은 안내 RNA 클로닝 부위, CMV 프로모터, BE3N, 및 dNCas9NIntein. 도 46c는 도 46a 및 도 46b에서 서열의 밑줄치고/치거나 강조된 부위 각각에 대한 주석 정보를 제공한다. 도 46d는 도 46a 및 도 46b에 의해 암호화된 작제물의 그래프 맵이다.

도 47a 및 도 47b 는 다음 성분을 포함하는 AAV (pX601) 벡터에 대한 예시적인 서열을 제공한다: CMV 프로모터, Cas9Sa, U6 프로모터, 및 gSa. 도 47c는 도 47a 및 도 47b.에서 서열의 밑줄치고/치거나 강조된 부위 각각에 대한 주석 정보를 제공한다. 도 47d는 도 47a 및 도 47b에 의해 암호화된 작제물의 그래프 맵이다.

[0082] 도 48a는 다음 성분을 포함하는 이중 AAV(pX600) 시스템에서 2개 벡터 중 하나에 대한 예시적인 서열이다: CMV 프로모터, dCInteinCCas9, VR, 및 PolyA. 도 48b는 도 48a에서 서열의 밑줄치고/치거나 강조된 부위 각각에 대한 주석 정보를 제공한다. 도 48c는 도 48a에 의해 암호화된 작제물의 그래프 맵이다.

도 49a는 다음의 성분을 포함하는 이중 AAV(pX600) 시스템에서 2개 벡터 중 하나에 대한 예시적인 서열이다: CMV 프로모터, dCInteinCCas9, EcoRV, 및 PolyA. 도 49b는 도 49a에서 서열의 밑줄치고/치거나 강조된 부위

각각에 대한 주석 정보를 제공한다. 도 50c는 도 49a에 의해 암호화된 작제물의 그래프 맵이다.

도 50a는 다음의 성분을 포함하는 이중 AAV(Custom) 시스템에서 2개 벡터 중 하나에 대한 예시적인 서열이다: U6 프로모터에 이은 안내 RNA 클로닝 부위, CMV 프로모터, KRAB, 및 dNCas9NIntein. 도 50b는 도 50a에서 서열의 밑줄치고/치거나 강조된 부위 각각에 대한 주석 정보를 제공한다. 도 50c는 도 50a에 의해 암호화된 작제물의 그래프 맵이다.

도 51a는 다음 성분을 포함하는 이중 AAV(Custom) 시스템에서 2개 벡터 중 하나에 대한 예시적인 서열이다: U6 프로모터에 이은 안내 RNA 클로닝 부위, CMV 프로모터, EcoRV, 및 dNCas9. 도 51b는 도 51a에서 서열의 밑줄치고/치거나 강조된 부위 각각에 대한 주석 정보를 제공한다. 도 51c는 도 51a에 의해 암호화된 작제물의 그래프 맵이다.

표의 간단한 설명

표 1은 실시예 1에서 사용된 안내 RNA 스페이서 서열을 나열한다.

표 2a는 실시예 1에 사용된 qPCR 프라이머의 올리고뉴클레오타이드 서열을 나열한다.

표 2b는 실시예 1에 사용된 NGS 프라이머의 올리고뉴클레오타이드 서열을 나타낸다.

표 2c는 실시예 1에서 사용된 AAV 구축을 위한 올리고뉴클레오타이드의 올리고뉴클레오타이드 서열을 나열한다.

발명을 실시하기 위한 구체적인 내용

[0036] 본 개시내용에 따른 구현예는 이후에 보다 충분히 기술될 것이다. 그러나, 본 개시내용의 국면은 상이한 형태로 구현될 수 있으며 본원에 제시된 구현예로 한정되는 것으로 해석되지 않아야 한다. 오히려, 이들 구현예를 제공함으로써 본 개시내용이 철저하고 완전해지도록 할 것이며, 본 발명의 영역을 당해 분야의 기술자에게 완전히 전달할 것이다. 본원의 상세한 설명에 사용된 전문용어는 특수한 구현예만을 기술할 목적이며 제한하려는 것이 아니다.

[0037] **정의**

[0038] 달리 정의하지 않는 한, 본원에 사용된 모든 용어(기술적 및 과학적 용어 포함)는 본 발명이 속한 당해 분야의 기술자가 일반적으로 이해하는 바와 동일한 의미를 갖는다. 사전에 일반적으로 사용시 규정된 것들과 같은 용어는 본원의 문맥에서 이들의 의미와 일치하는 것으로 해석되어야 하며 본원에서 이와 같이 표현해서 정의되지 않는 한 이상적이거나 너무나 형식적인 의미로 해석되어서는 안된다. 하기에 명확하게 정의하지 않는 한, 이러한 용어는 이의 일반적인 의미에 따라 해석되어야 한다.

[0039] 본원의 상세한 설명에 사용된 전문 용어는 특수한 구현예만을 기술할 목적을 위한 것이며 본 발명을 제한하는 것으로 의도되지 않는다. 본원에 언급된 모든 공보, 특허원, 특허 및 다른 참고문헌은 이의 전문이 참고로 포함된다.

[0040] 본 기술의 실시는 달리 나타내지 않는 한 조직 배양, 면역학, 분자 생물학, 미생물학, 세포 생물학, 및 재조합 DNA의 통상의 기술을 사용할 것이며, 이는 본 발명의 기술 내에 있다.

[0041] 내용이 달리 나타내지 않는 한, 본원에 기술된 본 발명의 다양한 특징은 임의의 조합으로 사용될 수 있는 것으로 구체적으로 의도된다. 더욱이, 본 개시내용은 또한 일부 구현예에서, 본원에 제시된 임의의 특징 또는 특징의 조합을 배제하거나 제외시킬 수 있는 것으로 고려된다. 설명하기 위해, 명세서가 복합체가 성분 A, B 및 C를 포함하는 것으로 기술하는 경우, 임의의 A, B 또는 C, 또는 이의 조합이 제외될 수 있고 단수로 또는 임의의 조합으로 거부될 수 있는 것으로 구체적으로 의도된다.

[0042] 달리 명확하게 나타내지 않는 한, 모든 구체적인 구현예, 특징, 및 용어는 인용된 구현예, 특징, 또는 용어 및 이의 생물학적 등가물 둘 다를 포함하는 것으로 의도된다.

[0043] 범위를 포함하는 모든 수치 정의, 예컨대, pH, 온도, 시간, 농도, 및 분자량은 적절하게는 1.0 또는 0.1의 (+) 또는 (-) 증분으로 또는 대안적으로 10%, 또는 대안적으로 5%, 또는 대안적으로 2%로 변하는 근사치이다. 항상 명확하게 기술하지 않지만, 모든 수치 정의는 용어 "약"이 선행하는 것으로 이해되어야 한다. 또한 항상 명확하게 기술되지 않지만, 본원에 기술된 시약은 단지 예시적이며 이의 등가물이 당해 분야에 공지되어 있음이 이해되어야 한다.

- [0044] 본 발명의 명세서 및 첨부된 청구범위에 사용된 것으로서, 단수형("a", "an", 및 "the")은 내용에서 명확하게 달리 나타내지 않는 한, 복수형을 또한 포함하는 것으로 의도된다.
- [0045] 본원에 사용된 바와 같은 용어 "약"은 양 또는 농도 등과 같은 측정가능한 값을 지칭하는 경우 규정된 양의 20%, 10%, 5%, 1%, 0.5%, 또는 심지어 0.1%의 변화를 포함함을 의미한다.
- [0046] 또한 본원에 사용된 바와 같은, "및/또는"은 관련된 나열된 항목 중 하나 이상의 특정 및 모든 가능한 조합 뿐만 아니라 대안("또는")으로 해석되는 경우 조합의 결여도 지칭하며 이를 포함한다.
- [0047] 본원에 사용된 바와 같은 용어 "세포"는 임의로 대상체 또는 상업적으로 이용가능한 공급원으로부터 임의로 수득된 원핵 또는 진핵 세포를 지칭할 수 있다.
- [0048] 본원에 사용된 바와 같은, 용어 "포함하는"은 조성물 및 방법이 인용된 성분들을 포함하지만, 다른 것을 배제하지는 않음을 의미하는 것으로 의도된다. 본원에 사용된 바와 같이, 전이 어구 "로 필수적으로 이루어진"(및 문법적 변이체)는 인용된 물질 또는 단계 및 인용된 구현예의 기본적 및 신규한 특성에 실질적으로 영향을 미치지 않는 것을 포함하는 것으로 해석되어야 한다. 따라서, 본원에 사용된 바와 같은 용어 "로 필수적으로 이루어진"은 "포함하는"과 등가물인 것으로 해석되어서는 안된다. "로 이루어진"은 본원에 개시된 조성물을 투여하기 위한 다른 성분 및 실질적인 방법 단계의 미량 성분 이상을 배제함을 의미할 수 있다. 이들 전이 용어 각각에 의해 정의된 국면은 본 개시내용의 영역 내에 있다.
- [0049] 용어 "암호화하는"은 이것이 핵산 서열에 적용되는 경우, 이의 천연 상태에서 폴리펩타이드를 "암호화하는" 것으로 일컬어지는 폴리뉴클레오타이드를 지칭하거나 당해 분야의 기술자에게 잘 공지된 방법에 의해 조작되는 경우, 폴리펩타이드 및/또는 이의 단편에 대한 mRNA를 생산하기 위해 전사되고/되거나 해독될 수 있음을 의미한다. 안티센스 가닥은 이러한 핵산의 상보체이며, 암호화 서열은 이로부터 유추될 수 있다.
- [0050] 용어 "등가물" 또는 "생물학적 등가물"은 특수한 분자, 생물학적, 또는 세포 물질을 지칭하는 경우 상호교환적으로 사용되며 여전히 바람직한 구조 또는 기능을 유지하면서 최소의 상동성을 갖는 것들을 의도한다.
- [0051] 본원에 사용된 바와 같은, 용어 "발현"은 이에 의해 폴리펩타이드가 mRNA로 전사되고/되거나 이의 의해 전사된 mRNA가 후속적으로 펩타이드, 폴리펩타이드, 또는 단백질로 해독되는 공정을 지칭한다. 폴리뉴클레오타이드가 게놈 DNA로부터 기원하는 경우, 발현은 진핵 세포 속에서 mRNA의 스플라이싱을 포함할 수 있다. 유전자의 발현 수준은 세포 또는 조직 샘플 속에서 mRNA 또는 단백질의 양을 측정함으로써 측정할 수 있고; 또한, 다수의 유전자의 발현 수준은 특수한 샘플에 대한 발현 프로파일을 확립하기 위해 측정할 수 있다.
- [0052] 본원에 사용된 바와 같은, 용어 "기능성"은 이것이 특수한, 규정된 효과를 달성하는 것으로 의도되는 임의의 분자, 생물학적, 또는 세포 물질을 변형시키기 위해 사용될 수 있다.
- [0053] 본원에 사용된 바와 같은, 용어 "핵산 서열", "올리고뉴클레오타이드", 및 "폴리뉴클레오타이드"는 상호교환적으로 사용되어 임의의 길이의 뉴클레오타이드의 중합체 형태, 즉, 리보뉴클레오타이드 또는 데옥시리보뉴클레오타이드를 지칭한다. 따라서, 이러한 용어는 단일-, 이중-, 또는 다중-가닥의 DNA 또는 RNA, 게놈 DNA, cDNA, DNA-RNA 하이브리드, 또는 푸린 및 피리미딘 염기 또는 다른 천연의, 화학적으로 또는 생화학적으로 변형된, 비-천연, 또는 유도체화된 뉴클레오타이드 염기를 포함하는 중합체를 포함하나, 이에 한정되지 않는다.
- [0054] 본원에 사용된 바와 같은 용어 "분리된"은 다른 물질로부터 실질적으로 유리되는 분자 또는 생물체(biologicals) 또는 세포 물질을 지칭한다.
- [0055] 본원에 사용된 바와 같은, 용어 "기관"은 개개 기관의 특수 부위인 구조이며, 여기서 개개 기관의 특정 기능 또는 기능들은 국소적으로 수행되고 이들은 형태학적으로 별개이다. 기관의 비-제한적 예는 피부, 혈관, 각막, 흉선, 신장, 심장, 간, 제대, 장, 신경, 폐, 태반, 척장, 갑상선 및 뇌를 포함한다.
- [0056] 용어 "단백질", "펩타이드" 및 "폴리펩타이드"는 상호교환적으로 및 이들의 가장 광의의 의미에서 사용되어 핵산, 아미노산 유사체 또는 펩티도미메틱(peptidomimetic) 중 2개 이상의 소단위(subunit)를 지칭한다. 소단위는 펩타이드 결합에 의해 연결될 수 있다. 다른 국면에서, 소단위는 다른 결합, 예컨대, 에스테르, 에테르 등에 의해 연결될 수 있다. 단백질 또는 펩타이드는 적어도 2개의 아미노산을 함유하여야만 하며 단백질 또는 펩타이드의 서열을 포함할 수 있는 아미노산의 최대 수에 있어서 제한은 없다. 본원에 사용된 바와 같은 용어 "아미노산"은 글리신 및 D 및 L 광학 이성체 둘 다를 포함하는 천연 및/또는 비천연 또는 합성 아미노산, 아미노산 유사체 및 펩티도미메틱을 지칭한다. 펩타이드는 이들의 구조로 정의될 수 있다. 예를 들면, "비사이클릭 펩타이드"는 임의로 항체 모사체(antibody mimetic)로서 기능하도록 가공된, 2개의 폐환된 부위를 포함하는 펩

타이드의 계열을 지칭한다.

- [0057] 용어 "조직"은 본원에서 살아있는 또는 죽은 유기체의 조직 또는 살아있거나 죽은 유기체로부터 기원하거나 이를 모사하도록 설계된 임의의 조직을 지칭한다. 조직은 건강하고/하거나, 죽어 있고/있거나 유전적 돌연변이를 가질 수 있다. 생물학적 조직은 임의의 단일 조직(예컨대, 상호연결될 수 있는 세포의 수집) 또는 기관 또는 기관 본체의 일부 또는 영역을 이루는 조직의 그룹을 포함할 수 있다. 조직은 균질한 세포 물질을 포함할 수 있거나, 이는 예를 들면 폐 조직, 골격 조직, 및/또는 근육 조직을 포함할 수 있는 흉곽을 포함하는 본체의 영역에서 발견된 것과 같은 복합 구조일 수 있다. 예시적인 조직은 이의 임의의 조합을 포함하는, 간, 폐, 갑상선, 피부, 췌장, 혈관, 방광, 신장, 뇌, 담도계, 십이지장, 복부대동맥, 엉덩 정맥, 심장 및 장의 조직을 포함하나, 이에 한정되지 않는다.
- [0058] "유효량" 또는 "효과적인 양"은 의도한 목적을 달성하기에 충분한 양이다. 일 국면에서, 유효량은 기술된 치료학적 목적을 달성하기 위해 기능하는 것, 예컨대, 치료학적 유효량이다. 본원에 상세히 기술된 바와 같이, 유효량, 또는 용량은 목적 및 구성물에 의존하며, 본 개시내용에 따라 측정될 수 있다.
- [0059] 본원에 사용된 바와 같이, 용어 "CRISPR"은 균집을 이루어 일정하게 산재된 짧은 팔린드롬 반복체 경로 (palindromic repeats pathway)에 의존한 서열 특이적인 유전자 조각의 기술을 지칭하며, 이는 RNA 간섭 (interference)과는 달리 전사 수준에서 유전자 발현을 조절한다. 본원에 사용된 바와 같은 용어 "gRNA" 또는 "안내 RNA"는 CRISPR 기술을 사용하는 교정용의 구체적인 유전자를 표적화하는데 사용된 안내 RNA 서열을 지칭한다. 표적 특이성에 대한 gRNA 및 공여체 치료학적 폴리뉴클레오타이드의 설계 기술은 당해 분야에 잘 알려져 있다. 예컨대, 본원에 참고로 혼입된, Doench et al. (2014) *Nature Biotechnol.* 32(12):1262-7 및 Graham et al. (2015) *Genome Biol.* 16: 260을 참고한다. 본원에 사용된 경우, gRNA는 이중 또는 단일 gRNA를 지칭할 수 있다. 이들 둘 다의 비-제한적인 예시적 구현예가 본원에 제공된다.
- [0060] 용어 "Cas9"는 이의 명칭(UniProtKB G3ECR1(CAS9_STRTR)) 뿐만 아니라 엔도뉴클레아제 활성을 결여한(예컨대, RuvC 및 HNH 도메인 둘 다에서 돌연변이를 지닌), 죽은 Cas9 또는 dCas9에 의해 지칭된 CRISPR 관련 엔도뉴클레아제를 지칭한다. 용어 "Cas9"는 또한 다른 큰 Cas9 단백질을 포함하나, 이에 한정되지 않는, 이에 대해 적어도 약 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 또는 99% 동일성(identity)을 갖는 언급된 Cas9의 등가물을 지칭할 수 있다.
- [0061] 용어 "인테인"은 자가 절개할 수 있고 단백질 스플라이싱(splicing)을 통해 단백질의 나머지 부위를 연결할 수 있는 단백질의 부류를 지칭한다. "split-인테인"은 2개의 유전자로부터 오는 인테인을 지칭한다. 비-제한적 예는 본원에서 split-Cas9 시스템의 일부로 개시된 엔. 폰크티포르메(*N. punctiforme*)내의 split 인테인이다. 접두어 N 및 C는 split 인테인의 맥락에서 인테인의 1/2을 암호화하는 유전자를 어느 단백질이 종료하는지를 확립하기 위해 사용될 수 있다.
- [0062] 본원에 사용된 바와 같은, 용어 "재조합 발현 시스템"은 재조합에 의해 형성된 특정 유전 물질의 발현을 위한 유전 작제물을 지칭한다.
- [0063] 본원에 사용된 바와 같은 용어 "아데노-관련 바이러스" 또는 "AAV"는 이러한 명칭과 관련되고 파르보비리다에과의 데펜도파르보바이러스(dependoparvovirus) 속(genus)에 속하는, 바이러스의 부류의 구성원을 지칭한다. 이러한 바이러스의 다수의 혈청형은 유전자 전달에 적합한 것으로 알려져 있다; 모든 공지된 혈청형은 다양한 조직 유형으로부터의 세포를 감염시킬 수 있다. 연속적으로 번호매긴, 적어도 11개가 선행 분야에 개시되어 있다. 본원에 개시된 방법에 유용한 비-제한적인 예시적인 혈청형은 11개의 혈청형 중 임의의 것, 예컨대, AAV2 및 AAV8, 또는 변이체 혈청형, 예컨대, AAV-DJ를 포함한다.
- [0064] 본원에 사용된 바와 같은 용어 "렌티바이러스"는 이러한 명칭과 관련되고 레트로비리다(Retroviridae)에과의 렌티바이러스 속에 속하는 바이러스 부류의 구성원을 지칭한다. 일부 렌티바이러스는 질환을 유발하는 것으로 알려져 있지만, 다른 렌티바이러스는 유전자 전달에 적합한 것으로 알려져 있다. 예컨대, 문헌 'Tomás et al. (2013) *Biochemistry, Genetics and Molecular Biology: "Gene Therapy - Tools and Potential Applications,"* ISBN 978-953-51-1014-9, DOI: 10.5772/52534'을 참고한다.
- [0065] 본원에 사용된 바와 같은, 용어 "백터"는 비-분열 및/또는 서서히 분열하는 세포를 감염시키고 형질도입시키며 표적 세포의 게놈내로 통합시키는 능력을 보유한 재조합체 백터를 의도한다. 백터는 야생형 바이러스로부터 유도되거나 이를 기반으로 할 수 있다. 본 개시내용의 국면은 아데노-관련 바이러스 또는 렌티바이러스 백터에 관

한 것이다.

[0066] 본원에 사용된 바와 같은 용어 "프로모터"는 유전자와 같은 암호화 서열의 발현을 조절하는 임의의 서열을 지칭한다. 프로모터는 예를 들면, 구성적, 유도성, 억제성, 또는 조직-특이성일 수 있다. "프로모터"는 전사의 개시 및 전사율이 조절되는 폴리뉴클레오타이드 서열의 영역인 조절 서열이다. 이는 RNA 폴리머라제 및 다른 전사 인자와 같은 조절 단백질 및 분자가 결합할 수 있는 유전 성분을 함유할 수 있다. 비-제한적인 예시적인 프로모터는 CMV 프로모터 및 U6 프로모터를 포함한다. 비-제한적인 예시적인 프로모터 서열은 하기 본원에 제공된다:

[0067] CMV 프로모터

```
ATACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTC
ATTAGTTCATAGCCCATATATGGAGTTCGCGTTACATAACTTACGGTAAATGGC
CCGCCTGGCTGACCGCCCAACGACCCCGCCATTGACGTCAATAATGACGTATG
TTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTT
ACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCC
CCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGA
CCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACC
ATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCAC
GGGGATTTCCAAGTCTCCACCCCATGACGTCAATGGGAGTTTGTTTTGGCACCA
AAATCAACGGGACTTTCCAAAATGTCGTAACAACCTCCGCCCCATTGACGCAAATG
GGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAAC
CGTCAGATCGCCTGGAGACGCCATCCACGCTGTTTTGACCTCCATAGAAGACACC
GGGACCGATCCAGCCTCCGGACTCTAGAGGATCGAACCCCTT
```

[0068]

[0069] 또는 이의 생물학적 증가물.

[0070] U6 프로모터

```
GAGGGCCTATTTCCCATGATTCCTTCATATTTGCATATACGATACAAGGCTGTTA
GAGAGATAATTAGAATTAATTTGACTGTAACACAAAGATATTAGTACAAAATA
CGTGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAAATTATGTTTT
AAAATGGACTATCATATGCTTACCGTAACTTGAAAGTATTTGATTTCTTGGCTTT
ATATATCTTGTGGAAAGGACGAAACACC
```

[0071]

[0072] 또는 이의 생물학적 증가물.

[0073] 다수의 효과기 성분, 예컨대, 테트라사이클린 반응 성분(예컨대, tet0), tet-조절가능한 활성인자, T2A, VP64, RtA, KRAB, 및 miRNA 센서 회로가 이러한 벡터에서 사용하기 위해 본원에 개시되어 있다. 이러한 효과기 성분의 특성 및 기능은 당해 분야에서 일반적으로 이해되어 있으며 이들 다수의 효과기 성분은 상업적으로 이용가능하다. 이의 비-제한적인 예시적인 서열은 본원에 개시되어 있으며 이의 추가의 설명은 하기 본원에서 제공된다.

[0074] 본원에 사용된 바와 같은 용어 "엡타머"는 고 친화성 및 특이성으로 하나 이상의 선택된 표적에 결합할 수 있는 단일 가닥 DNA 또는 RNA 분자를 지칭한다. 비-제한적인 예시적인 표적은 단백질 또는 펩타이드를 포함하나 이에 한정되지 않는다.

[0075] 본원에 사용된 바와 같은 용어 "아피보디(affibody)"는 고 친화성으로 다수의 표적 단백질 또는 펩타이드에 결합하도록 가공된 작은 단백질로 구성된 항체 모사체 유형을 지칭한다. 일반적인 아피보디 구조는 특이적인 표적에 결합시키기 위해 이후에 변형될 수 있는 3개의 나선-다발(helix-bundle)을 기반으로 한다.

[0076] 본원에 사용된 바와 같은 용어 "DARPin"은 표적 단백질에 대해 고 특이성 및 친화성을 지닌 가공된 항체 모사체의 유형인, 설계된 안키린 반복체 단백질을 지칭한다. 일반적으로, DARPin은 단백질 모티프(안키린)의 적어도 3개의 반복체, 임의로 4개 또는 5개의 반복체를 포함하며, 분자량이 약 14 내지 18 kDa이다.

[0077] 본원에 사용된 바와 같은 용어 "쿠니츠 도메인(Kunitz domain)"은 프로테아제 억제제로서 기능하는 단백질 내에서 발견된 디설피드 우측 알파+베타 폴드 도메인(disulfide right alpha+beta fold domain)을 지칭한다. 일반적으로, 쿠니츠 도메인은 길이가 대략 50 내지 60개 아미노산이고 분자량이 약 6 kDa이다.

[0078] 본원에 사용된 바와 같은 용어 "피노머(fynomer)"는 사람 Fyn SH3 도메인(GeneCards 참고, FYN에 기술됨)으로부터 기원한 작은 결합 단백질을 지칭하며, 이는 항체 모사체가 되도록 가공될 수 있다.

[0079] 본원에 사용된 바와 같은 용어 "안티칼린"은 항체와 구조적으로 관련되지 않은 항원에 결합할 수 있는 인공 단백질을

포함하는, Pieris Pharmaceuticals에 의해 현재 시판되는, 항체 모사체의 유형을 지칭한다. 안티칼린은 사람 립 칼린(lipcalin)으로부터 기원하며 특수한 표적에 결합하도록 변형된다.

[0080] 본원에 사용된 바와 같은 용어 "애드넥틴"은 모노바디(monobody)를 지칭하며, 이는 항체 모사체로서 제공되는 합성 결합 단백질이고, 이는 피브로넥틴 제III형 도메인(FN3)을 사용하여 작제된다.

[0081] 본 개시내용이 폴리펩타이드, 단백질, 폴리뉴클레오타이드 또는 항체에 관한 것인 경우, 이의 등가물 또는 생물학적 등가물은 본 개시내용의 범위 내에 있는 것으로 의도됨이 분명하게 인용하지 않고 달리 의도되지 않아도 부여되어야 한다. 본원에 사용된 바와 같은, 용어 "이의 생물학적 등가물은 참고 단백질, 항체, 폴리펩타이드 또는 핵산을 지칭하는 경우 목적인 구조 또는 기능성을 여전히 유지하면서 최소의 상동성을 갖는 것으로 고려된다. 본원에서 구체적으로 인용하지 않는 한, 본원에 언급된 임의의 폴리뉴클레오타이드, 폴리펩타이드 또는 단백질은 또한 이의 등가물을 포함하는 것으로 의도된다. 예를 들면, 등가물은 참고 단백질, 폴리펩타이드 또는 핵산에 대해 적어도 약 70% 상동성 또는 동일성, 또는 적어도 80% 상동성 또는 동일성 및 대안적으로, 또는 적어도 약 85%, 또는 대안적으로 적어도 약 90%, 또는 대안적으로 적어도 약 95%, 또는 대안적으로 98%의 상동성 또는 동일성을 지닐 수 있으며 실질적으로 증가의 생물학적 활성을 나타낸다. 대안적으로는, 폴리뉴클레오타이드를 지칭하는 경우, 이의 등가물은 참고 폴리뉴클레오타이드 또는 이의 상보체에 대해 엄격한 조건(stringent condition) 하에서 하이브리드화하는 폴리뉴클레오타이드이다.

[0082] 본 출원인은 본원에서 하기 기술된 유전자 및 단백질 전달 및 발현 기술에서 사용하기 위한 폴리펩타이드 및/또는 폴리뉴클레오타이드 서열을 제공한다. 항상 분명하게 기술되지는 않지만, 본원에 제공된 서열을 사용하여 발현 생성물 뿐만 아니라 동일한 생물학적 특성을 갖는 단백질을 생산하는 실질적으로 동일한 서열도 제공할 수 있음이 이해되어야 한다. "생물학적 등가물" 또는 "생물학적으로 활성인" 폴리펩타이드는 본원에 기술된 바와 같은 증가의 폴리뉴클레오타이드에 의해 암호화된다. 이들은 디폴트 조건(default condition) 하에서 수행된 서열 동일성 방법을 사용하여 비교하는 경우 참고 폴리펩타이드에 대해 적어도 60%, 또는 대안적으로, 적어도 65%, 또는 대안적으로, 적어도 70%, 또는 대안적으로, 적어도 75%, 또는 대안적으로, 적어도 80%, 또는 대안적으로 적어도 85%, 또는 대안적으로 적어도 90%, 또는 대안적으로 적어도 95% 또는 대안적으로 적어도 98% 동일한 주요 아미노산 서열을 지닐 수 있다. 구체적인 폴리펩타이드 서열이 특수한 구현예의 예로서 제공된다. 아미노산에 대한 서열은 유사한 전하를 지닌 대안의 아미노산을 지닌 아미노산으로 변형될 수 있다. 또한, 증가의 폴리뉴클레오타이드는 엄격한 조건 하에서 참고 폴리뉴클레오타이드 또는 이의 상보체 또는 폴리펩타이드에 대한 참고시, 엄격한 조건 하에서 폴리뉴클레오타이드 또는 이의 상보성 가닥을 암호화하는 참고 서열에 하이브리드화하는 폴리뉴클레오타이드에 의해 암호화된 폴리펩타이드에 하이브리드화하는 것이다. 대안적으로, 증가의 폴리펩타이드 또는 단백질은 증가의 폴리뉴클레오타이드로부터 발현된 것이다.

[0083] "하이브리드화"는 하나 이상의 폴리뉴클레오타이드가 반응하여 뉴클레오타이드 잔기의 염기들 사이에 수소 결합을 통해 안정화되는 복합체를 형성하는 반응을 지칭한다. 수소 결합은 왓슨-크릭 염기 쌍화(Watson-Crick base pairing), 후그스타인 결합(Hoogstein binding), 또는 임의의 다른 서열-특이적인 방식에 의해 발생할 수 있다. 복합체는 이분체(duplex) 구조를 형성하는 2개의 가닥, 다중 가닥 복합체를 형성하는 3개 이상의 가닥, 단일의 자가-가수분해하는 가닥, 또는 이들의 임의의 조합을 포함할 수 있다. 하이브리드화 반응은 PC 반응의 개시, 또는 리보자임에 의한 폴리뉴클레오타이드의 효소적 절단과 같은, 보다 집중적인 공정에서의 단계를 구성할 수 있다.

[0084] 엄격한 하이브리드화 조건의 예는 다음을 포함한다: 약 25°C 내지 약 37°C의 항온처리 온도; 약 6x SSC 내지 약 10x SSC의 하이브리드화 완충액 농도; 약 0% 내지 약 25%의 포름아미드 농도; 및 약 4x SSC 내지 약 8x SSC의 세척 용액. 중간 하이브리드화 조건의 예는 다음을 포함한다: 약 40°C 내지 약 50°C의 항온처리 온도; 약 9x SSC 내지 약 2x SSC의 완충제 농도; 약 30% 내지 약 50%의 포름아미드 농도; 및 약 5x SSC 내지 약 2x SSC의 세척 용액. 높은 엄격한 조건의 예는 다음을 포함한다: 약 55°C 내지 약 68°C의 항온처리 온도; 약 1x SSC 내지 약 0.1x SSC의 완충제 농도; 약 55% 내지 약 75%의 포름아미드 농도; 및 약 1x SSC, 0.1x SSC, 또는 탈이온수의 세척 용액. 일반적으로, 하이브리드화 항온처리 시간은 5분 내지 24시간이고, 1, 2회 이상의 세척 단계가 있으며, 세척 항온처리 시간은 약 1, 2, 또는 15분이다. SSC는 0.15 M NaCl 및 15 mM 시트레이트 완충액이다. 다른 완충제 시스템을 사용한 SSC의 등가물을 사용할 수 있는 것으로 이해된다.

[0085] "상동성" 또는 "동일성"은 2개의 펩타이드 사이 또는 2개의 핵산 분자 사이의 서열 유사성을 지칭한다. 상동성은 비교 목적을 위해 정렬될 수 있는 각각의 서열내 위치를 비교함으로써 측정할 수 있다. 비교된 서열내 위치가 동일한 염기 또는 아미노산에 의해 점유되는 경우, 분자는 이후에 이러한 위치에서 상동성이다. 서열 사이의

상동성 정도는 서열에 의해 공유된 매칭(matching)의 수 또는 상동성 위치의 수의 함수이다. "관련되지 않은" 또는 "비-상동성" 서열은 본 발명의 서열 중 하나와 40% 미만의 동일성, 또는 대안적으로 25% 미만의 동일성을 공유한다.

[0086] **본 개시내용의 수행 방식**

[0087] 본 개시내용은 유전자 편집의 보다 우수한 전달, 특이성 및 선택성을 허용하는 유일한 모듈러 CRISPR-Cas9 구조를 지닌 신규한 전달 시스템에 관한 것이다. 이는 앞서 기술된 split-Cas9 시스템보다 유의적인 개선을 나타낸다. 모듈러 구조는 "조절가능"하다. 추가의 국면은 공간적으로 및 일시적으로 둘 모두 조절될 수 있어서, 유도성 편집을 위한 잠재능을 생성할 수 있는 시스템에 관한 것이다. 추가의 국면은 호밍제(homing agent)에 대한 접합을 허용하는 변형된 바이러스 캡시드에 관한 것이다.

[0088] Split-Cas 시스템

[0089] 일 국면에서, 본 개시내용은 "split-Cas9"에 관한 것이며, 여기서 Cas9는 2개의 1/2 - C-Cas9 및 N-Cas9 -로 분할되며 2개의 인테인 모이어티 또는 "split 인테인"과 융합된다. 예컨대, 문헌 'Volz et al. (2015) Nat Biotechnol. 33(2):139-42; Wright et al. (2015) PNAS 112(10) 2984-89'를 참고한다. "split 인테인"은 2개의 유전자로부터 온다. "split-인테인"의 비-제한적인 예는 엔. 폰크티포르메(*N. punctiforme*)로부터 원래 기원한 C-인테인 및 N-인테인 서열이다. 비-제한적인 예시적 split-Cas9는 574 내지 1398번 잔기를 포함하는 C-Cas9 및 1 내지 573번 잔기를 포함하는 N-Cas9를 갖는다. dCas9에 대한 예시적인 split-Cas9는 dC-Cas9 및 dN-Cas9로 나타낸, dCas9의 이러한 동일한 잔기를 포함하는 2개의 도메인을 포함한다.

[0090] 이러한split-Cas9 모듈의 비-제한적인 예시적인 서열은 하기 본원에 제공된다. 아미노산 번호는 야생형 Cas9와 관련하여 제공된다.

[0091] C인테인(굵은 글씨) +CCas9(정상 글씨)(H840, 굵고 밑줄쳐진, 변형되지 않은 서열)

MIKIATRKYL GKQNVYDIGVERDHN FALKNGFIASCFDSVEISGVEDRFN
 ASLGT YHDL LKIIKDKDFLDNEENEDILEDIVL TLTLFEDREMEERLKT
 YAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGF
 ANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGL
 QTVKVVDELVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGI
 KELGSQILKEHPVENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDV
 DHIIVPQSFLKDDSDIDNKVLRSDKNRSGSDNVPSEEVVKKMKNYWRQLLN
 AKLITQRKFDNLTKAERGGSELDKAGFIKRQLVETRQITKHVAQILD SR
 MNTKYDENDKLIREVKVITLKS KLVSDFRKFDFYK VREINNYHHAHDAY
 LNAVVG TALIKKYPKLESEFVYGDYK VYDVRKMIAKSEQEIGKATAKYFF
 YSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL S
 MPQVNIIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPT
 VAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSEFEKNPIDFLEAKGYK
 EVKLDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFYLY
 ASHYEKLKGSPEQNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLD
 KVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYS
 TKEVLDATLIHQ SITGLYETRIDLSQLGGD

[0092]
 [0093] 또는 이의 생물학적 증가물.

[0094] C인테인(굵은 글씨)+dCCas9(정상)(H840A, 굵은 이탤릭체, 변형된 서열)

MIKIATRKYL GKQNVYDIGVERDHN FALKNGFIASCFDSVEISGVEDRFN
 ASLGT YHDL LKIIKDKDFLDNEENEDILEDIVL TLTLFEDREMEERLKT
 YAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGF
 ANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGL
 QTVKVVDELVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGI
 KELGSQILKEHPVENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDV
 DHIIVPQSFLKDDSDIDNKVLRSDKNRSGSDNVPSEEVVKKMKNYWRQLLN
 AKLITQRKFDNLTKAERGGSELDKAGFIKRQLVETRQITKHVAQILD SR
 MNTKYDENDKLIREVKVITLKS KLVSDFRKFDFYK VREINNYHHAHDAY
 LNAVVG TALIKKYPKLESEFVYGDYK VYDVRKMIAKSEQEIGKATAKYFF
 YSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL S
 MPQVNIIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPT
 VAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSEFEKNPIDFLEAKGYK
 EVKLDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFYLY
 ASHYEKLKGSPEQNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLD
 KVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYS
 TKEVLDATLIHQ SITGLYETRIDLSQLGGD

[0095]
 [0096] 또는 이의 생물학적 증가물.

[0097] NCas9(정상)(D10, 짧은 밑줄쳐진, 변형되지 않은 서열)+N-인테인(짧은 글씨)

MGPKKRKRVA AADYKDDDDKGIHG VPAADKKYSIGLDIGTNSVGVAVITD
EYKVP SKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYT
RRKNRICYLQEIFSNEMAKVDDSFHRL EESFLVEEDKKHERHPFGNIV
DEVAYHEKYPTIYHLRKKLV DSTDKADLR LIYLALAHMIKFRGHFLIEGD
LNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRIE
NLIAQLPGEKKNLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDD
DLDNLLAQIGDQYADFLAAKNLSDAILLSDILRVNTEITKAPLSASMIK
RYDEHHQDL TLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFY
KFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAIL
RRQEDFY PFLKDNREKIEKILTFRIPYVVGPLARGNSRF AWMTRKSEETI
TPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNE
LTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIE
CLSYE TEILTV EYGLLPIGKIVEKRIE CTVYSVDNNGNYTQPV AQWHDR
GEQEVFEYCLEDGSLIRATKDHKFMITVDGQMLPIDEIFERELDLMRVDNL
PN

[0098]

[0099] 또는 이의 생물학적 증가물.

[0100] dNCas9(정상)(D10A, 짧은 이탤릭체, 변형된 서열)+N-인테인(짧은 글씨)

MGPKKRKRVA AADYKDDDDKGIHG VPAADKKYSIGL4IGTNSVGVAVITD
EYKVP SKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYT
RRKNRICYLQEIFSNEMAKVDDSFHRL EESFLVEEDKKHERHPFGNIV
DEVAYHEKYPTIYHLRKKLV DSTDKADLR LIYLALAHMIKFRGHFLIEGD
LNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRIE
NLIAQLPGEKKNLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDD
DLDNLLAQIGDQYADFLAAKNLSDAILLSDILRVNTEITKAPLSASMIK
RYDEHHQDL TLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFY
KFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAIL
RRQEDFY PFLKDNREKIEKILTFRIPYVVGPLARGNSRF AWMTRKSEETI
TPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNE
LTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIE
CLSYE TEILTV EYGLLPIGKIVEKRIE CTVYSVDNNGNYTQPV AQWHDR
GEQEVFEYCLEDGSLIRATKDHKFMITVDGQMLPIDEIFERELDLMRVDNL
PN

[0101]

[0102] 또는 이의 생물학적 증가물.

[0103] 본 개시내용의 국면은: (a) (i) C-인테인을 암호화하는 폴리뉴클레오타이드, (ii) C-Cas9를 암호화하는 폴리뉴클레오타이드, 및 (iii) 프로모터 서열을 포함하는 제1의 발현 벡터; 및 (b)(i) N-Cas9를 암호화하는 폴리뉴클레오타이드, (ii) N-인테인을 암호화하는 폴리뉴클레오타이드, 및 (iii) 프로모터 서열을 포함하는 제2 발현 벡터를 포함하거나, 대안적으로 이로 필수적으로 이루어지거나, 여전히 추가로 이로 이루어진 CRISPR-기반 계놈 또는 에피게놈 편집용의 재조합 발현 시스템에 관한 것이며, 여기서 제1 및 제2 발현 벡터의 동시-발현은 기능성 Cas9 단백질의 발현을 야기한다.

[0104] 일부 구현예에서, 재조합 발현 시스템의 제1 및 제2 발현 벡터는 아데노-관련 바이러스(AAV) 벡터 또는 렌티바이러스 벡터이다.

[0105] 본원에 개시된 벡터에 대한 효과기 성분의 첨가는 CRISPR-기반 계놈 또는 에피게놈 편집에 있어서 특수한 사용을 위해 재조합 발현 시스템을 조정하기 위한 Cas9 발현의 조절을 허용한다. 개시된 "split-Cas9" 및/또는 재조합 발현 시스템의 맥락에서 비-제한적인 예시적인 효과기 성분 및 이의 용도는 하기 제공된다. 하기 기술된 효과기 성분 각각은 재조합 발현 시스템에서 특수한 기능의 맥락에서 기술되어 있음이 인식되어야 한다. 따라서, 이들 기능 중 하나 이상이 요구되는 경우, 이러한 효과기 성분은 재조합 발현 시스템에서 사용될 수 있다. 대조적으로, 이들 기능 중 하나 만이 요구되는 경우, 상응하는 효과기 성분 만이 재조합 발현 시스템에서 사용될 수 있다.

[0106] 일시적인 조절을 위한 효과기 성분

[0107] 일 국면에서, 재조합 발현 시스템의 제1 및/또는 제2 벡터는 유도성 발현을 허용하는 효과기 성분을 포함하거나, 또는 대안적으로 이로 필수적으로 이루어지거나, 여전히 추가로 이로 이루어지며, 여기서 특이적인 외부 제제의 도입은 벡터의 발현을 유도한다. 일반적으로, 이러한 유도는 특이적인 제제와 효과기 성분 사이의 상호작용으로 인하여 달성되어 전사 또는 해독의 완료를 허용한다.

[0108] 이러한 유도성 스위치(inducible switch)의 비-제한적 예는 본원에서 "Tet-ON" 시스템으로 지칭된 테트라사이클린 의존성 시스템이다. Tet-ON 시스템은 테트라사이클린 반응 성분("TRE")을 포함하며, 이는 TRE의 하부의 유전자의 전사 억제인자, 및 TRE에 결합하여 TRE의 하부의 유전자의 발현을 허용하는, 상응하는 테트라사이클린-조절가능한 활성인자("tet-조절가능한 활성인자"로서 작용하는, 테트라사이클린 반응 성분("TRE")을 포함한다.

tet-조절가능한 활성인자는 TRE에 결합하기 위하여 테트라사이클린 또는 이의 유도체(예컨대, 독시사이클린을 포함하나, 이에 한정되지 않음)의 존재를 필요로 한다. 따라서, Tet-ON 시스템을 사용함으로써, TRE의 하부의 유전자의 발현은, tet-조절가능한 성분이 또한 전사되는 것을 제외하고는, 테트라사이클린 또는 이의 유도체(예컨대, 독시사이클린을 포함하나 이에 한정되지 않음)의 첨가에 의해 "턴-온(turned on)"될 수 있다.

[0109] 일부 구현예에서, TRE는 TetO, 또는 임의로 이의 하나 이상의 반복 단위 또는 이의 7개의 반복 단위를 포함한다. TetO에 대한 표준 핵산 서열은: ACTCCCTATCAGTGATAGAGAA이다. TRE는 프로모터 서열을 추가로 포함할 수 있다. TetO의 7개의 반복 단위 및 최소의 CMV 프로모터를 포함하는, 이러한 TRE는 다음의 핵산 서열이다:

[0110] tet07-minCMV 프로모터

```
TTTACTCCCTATCAGTGATAGAGAACGTATGAAGAGTTTACTCCCTATCAGTGAT
AGAGAACGTATGCAGACTTTACTCCCTATCAGTGATAGAGAACGTATAAGGAGT
TTACTCCCTATCAGTGATAGAGAACGTATGACCAGTTTACTCCCTATCAGTGATA
GAGAACGTATCTACAGTTTACTCCCTATCAGTGATAGAGAACGTATATCCAGTTT
ACTCCCTATCAGTGATAGAGAACGTATAAGCTTTAGGCGTGTACGGTGGGCGCCT
ATAAAAGCAGAGCTCGTTTGTGAACCGTCAGATCGCTGGAGCAATTCCACAA
CACTTTTGTCTTATACCAACTTCCGTACCACTTCTACCTCGTAAA
```

[0111]

[0112] 또는 이의 생물학적 증가물.

[0113] 추가의 예시적인 서열은 TetO의 7개의 반복 단위를 포함한다:

[0114] tet07

```
TTTACTCCCTATCAGTGATAGAGAACGTATGAAGAGTTTACTCCCTATCAGTGAT
AGAGAACGTATGCAGACTTTACTCCCTATCAGTGATAGAGAACGTATAAGGAGT
TTACTCCCTATCAGTGATAGAGAACGTATGACCAGTTTACTCCCTATCAGTGATA
GAGAACGTATCTACAGTTTACTCCCTATCAGTGATAGAGAACGTATATCCAGTTT
ACTCCCTATCAGTGATAGAGAACGTATAA
```

[0115]

[0116] 또는 이의 생물학적 증가물.

[0117] 일부 구현예에서, tet-조절가능한 활성인자는 또한 "역 테트라사이클린-제어된 트랜스활성인자"로 또한 알려진 rtTA를 포함한다. 예컨대, 문헌 'Gossen et al. (1995) Science 268(5218):1766-1769'을 참고한다. tet-조절가능한 활성인자가 하나 이상의 유전자를 암호화하는 벡터(즉, 멀티시스트론성 벡터(multicistronic vector))로 제공되는 경우, tet-조절가능한 활성인자는 다른 벡터 생성물로부터 이의 해리를 허용하는 "자가-절단성" 펩타이드를 추가로 포함할 수 있다. 이러한 자가-절단성 펩타이드의 비-제한적 예는 2A이며, 이는 피코르나바이러스에서 처음 발견된 짧은 단백질 서열이다. 펩타이드 2A는 리보소움이 2A 성분의 C-말단에서 펩타이드 결합의 합성을 거르도록 함으로써 기능하여, 2A 서열의 말단과 이의 하부 펩타이드 사이의 분리를 야기한다. 이러한 "절단"은 C-말단에서 글리신 잔기와 프롤린 잔기 사이에서 일어난다. 2A 및 rtTA 둘 다를 포함하는 tet-조절가능한 활성인자의 비-제한된 예시적인 아미노산 서열은 하기에 제공된다:

[0118] 2A(짧은 글씨) +M2rtTA(정상 글씨)(tet 활성인자)

```
GSGATNFSLLKQAGDVEENPGPMSRLDKSKVINGALELLNGVGIEGLTTR
KLAQKLGVEQPTLYWHVKNKRALLDALPIEMLDHRHHTFCPLEGESWQDF
LRNNAKSFRCALLSHRDGAKVHLGTRPTEKQYETLENQLAFLCQQGFSLE
NALYALSAVGHFTLGCVLEEQEHQVAKEERETPTD SMPPLL RQAIELFD
RQGAEP AFLG LELIICGLEKQLKCESGGPADALDDFDLMDLPADALDDF
DLMDLPADALDDFDLMDLPG
```

[0119]

[0120] 또는 이의 생물학적 증가물.

[0121] 일부 구현예에서, Tet-ON 시스템은 본원에 개시된 재조합 발현 시스템과 같은 split Cas-9 시스템내로 통합될 수 있다.

[0122] 일부 구현예에서, 제1 벡터는 테트라사이클린 반응 성분("TRE")을 포함하며 제2 벡터는 테트라사이클린-조절가능한 활성인자("tet-조절가능한 활성인자")를 포함한다. 일부 구현예에서, 제2 벡터는 TRE를 포함하고 제1 벡터는 tet-조절가능한 활성인자를 포함한다.

[0123] 비-제한적인 예는 도에 묘사되어 있다: C-Cas9 벡터의 경우, Tet 오퍼레이터(TetO) 및 최소 CMV 프로모터를 포함하는 TRE, N-Cas9 벡터의 경우, rtTA를 포함하는 tet-조절가능한 활성인자가 임의로 첨가될 수 있다. 시스템에 대한 독시사이클린의 도입은 rtTa가 TetO에 결합하도록 하며 C-Cas9의 전사를 개시하여, 유전자 편집을 허용한다. (도 3). 본 출원인은 이러한 비-제한적 예시적인 시스템을 생체내에서 시험하였으며 편집이 DOX+ 마우스의

존재하에서 관찰되지만, DOX- 마우스에서는 관찰되지 않음을 입증하였다(도 7).

[0124] 조직 특이성에 대한 효과기 성분

[0125] 일 국면에서, 제조합체 발현 시스템의 제1 및/또는 제2 벡터는 조직 특이적인 발현을 제공하는 효과기 성분 또는 "회로"를 포함하거나, 대안적으로 이로 필수적으로 이루어지거나, 여전히 추가로 이로 이루어지는데, 즉, 벡터의 발현이 단백질, 올리고뉴클레오타이드, 또는 다른 생물학적 성분과 같은 하나 이상의 제제에 의해 유도되는 경우, 하나 이상의 특이적인 조직 속에 존재한다.

[0126] 회로와 같은 것의 비-제한적인 예는 조절가능한 microRNA("miRNA") 회로 또는 스위치이다. miRNA 스위치는 microRNA에 의해 양으로 또는 음으로 조절되도록 설계될 수 있는 유전자 발현의 억제인자 또는 활성인자이다.

[0127] MircoRNA는 표적 mRNA에 대해 쌍을 이루어 mRNA 가닥을 2개의 조각으로 1회 이상 절단, 폴리(A) 테일의 단축, 및/또는 mRNA 해독 효능의 감소를 통한 mRNA의 탈안정화함으로써 mRNA를 사일런싱하는 작은 비-암호화 RNA 분자이다. 특이적인 조직에서 발현되는 특이적인 miRNA는 다양한 데이터베이스, 예를 들면, miRmine(guanlab.ccmb.med.umich.edu/mirmine/) 및 MESAdb(konulab.fen.bilkent.edu.tr/mirna/mirna.php)에서 카탈로그화된다. 본원에서 관련될 수 있는 miRNA 및 상응하는 miRNA 표적의 비-제한적 예가 제공된다:

[0128] HeLa:

[0129] miR-21-5p: **uagcuuauacagacugauguuga**

[0130] 삽입된 표적: **TCAACATCAGTCTGATAAGCTAAGATCTA**

[0131] HUVEC:

[0132] **miR-126-3p:ucguaccgugaguaauaauugcg**

[0133] 삽입된 표적: **CGCATTATTACTCACGGTACGAAGATCAC**

[0134] 심장:

[0135] **miR-1a-3p:uggaauguaaagaaguau**

[0136] 삽입된 표적: **ATACATACTTCTTTACATTCCAAGATCAC**

[0137] 간:

[0138] **miR-122a-5p:uggagugugacaauugguuug**

[0139] 삽입된 표적: **CAAACACCATTGTCACACTCCAAGATCAC**

[0140] 또는 이의 각각의 생물학적 등가물. 조직 특이적인 miRNA를 선택하고 이러한 miRNA에 의해 표적화된 miRNA 회로를 생성함으로써, 벡터 발현을 고도로 조직 특이적으로도 조정할 수 있다.

[0141] 예를 들면, 예시적인 벡터는 이의 5' UTR에서 miRNA 표적 부위에 의해 음으로 조절된 발현의 억제인자로 구성된 miRNA 회로를 함유할 수 있다. 따라서, 벡터가 miRNA를 발현하는 표적 조직 유형으로 전달되는 경우, 억제인자는 억제되고, 상응하는 벡터가 활성화된다. 대조적으로, 벡터가 miRNA 부위를 함유하지 않는 부정확한 조직 유형으로 전달되는 경우, 벡터는 억제된다.

[0142] 일부 구현예에서, 제1 및/또는 제2 벡터는 특이적인 조직을 표적화하는 miRNA 스위치를 혼입시킨다. 이러한 혼입의 비-제한적인 예시적 개략도는 도 5에서 제공된다. 일부 구현예에서, miRNA 스위치는 miRNA 표적에 의해 음으로 조절된 발현의 억제인자를 포함한다.

[0143] 유전자 편집을 위한 효과기 성분

[0144] 본원에 개시된 제조합 발현 시스템은 활성 또는 죽은 Cas9를 사용할 수 있으므로, 다양한 임의의 효과기 성분을 혼입시켜 본원에 기술된 계통을 따라 게놈 편집을 촉진시킬 수 있다.

[0145] **녹아웃(Knock-out) 및 녹인(Knock-in):** 본원에 개시된 제조합 발현 시스템은 CRISPR-기반 게놈 또는 에피게놈 편집(epigenome editing)용으로 설계된다. 일반적으로, CRISPR-기반 게놈 또는 에피게놈 편집은 Cas9의 기능에 의존하여 gRNA와 표적 서열 사이의 쌍화를 촉진한다. gRNA는 일반적으로 특이적인 표적 유전자를 표적화하도록 설계되며 또한 CRISPR RNA(crRNA) 및 트랜스-작용성 CRISPR RNA (tracrRNA)를 추가로 포함할 수 있다. 표적 유전자에 대한 Cas9-gRNA 복합체의 쌍화시, 활성 Cas9 효소는 표적 특이적인 절단을 개시하여 유전자를 파괴하고, 임의로 유전자내에서 녹아웃 또는 녹인될 수 있다. 이는 CRISPR-Cas9 유전자 편집에 대해 취해진 전통적인 시도이며 구체적으로, 유전 질환을 사용한, 치료학적 적용에 매우 유용한 것으로 입증된다.

[0146] 대안적으로, 죽은 Cas9("dCas9")가 사용되는 경우, Cas9-gRNA 복합체는 편집; 하향조절, 억제, 또는 사일런싱(silencing); 상향조절, 과발현, 또는 활성화; 또는 표적 유전자의 메틸화의 변경을 포함하나, 이에 한정되지 않는 상이한 편집 효과를 위해 구성될 수 있다.

[0147] **염기 편집:** 일부 구현예에서, 염기 편집 시도는 dCas9를 사용하는 제조합 발현 시스템, 예컨대, split-Cas9 이중 AAV 시스템내로 혼입될 수 있다.

[0148] 예를 들면, 사이티딘의 우라실로의 전환을 지시하므로, 점-돌연변이(point-mutation)를 고정시키는데 유용한 사이티딘 데아미나제 효소는 제1 및/또는 제2 벡터내로 혼입될 수 있다. 이러한 시도는 이중-가닥 브레이크(BREAK)를 필요로 하지 않으며 전통적인 CRISPR-Cas9 유전자 편집에 의해 제안된 위험으로서, 무작위적인 인텔의 도입없이 점 돌연변이를 사용한 유전자 교정시 효과적이다. 따라서, 이러한 시스템은 오프-표적 무작위 인텔 형성을 최소화함으로써 생성물 선택성을 증가시킨다. 이러한 시도의 비-제한적 예는 제3-세대 염기 편집인자, APOBEC-XTEN-dCas9(A840H)-UGI(Komor et al. (2016) Nature 533:420-424 및 보충 소재에 개시됨)를 사용하며, 이는 편집된U의 반대쪽 G를 함유하는 편집되지 않은 가닥에 닉킹(nicking)한다. 제조합 발현 시스템, 예컨대, 본원에 개시된 split-Cas9 시스템으로 적응시킬 수 있는 코머(Komor) 등으로부터의 APOBEC1을 포함하는 Cas9에 대한 작제물은 하기에 제공된다:

[0149] BE3(**rAPOBEC1(굵은 글씨, 밑줄쳐짐)**-XTEN-Cas9n-UGI-NLS)

MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELRKETCLLYEINWGGRHSIWRHT
SQNTNKHVEVNFIEKFTTERTYFCPNTRCSITWFLSWSPCGECSRATIEFLSRYPH
VTLFIYIARLYHHADPRNROGLRDLISSGVITQIMTEOESGYCWRNFVNYSNSNE
AHWPRYPHLWVRLYVLELYLGLPPLCNILRRKOPOLTFETIALOSCHYORL
PPHILWATGLKSGSETPGTSSEATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSSK
 FKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNE
 MAKVDDSFHRLSEESFLVEEDKKHERHPHIFGNIVDEVAYHEKYPTIYHLRKKLV DST
 DKADRLIYLALAHMKFRGHFLIEGDLNPDNSVDKLFQLVQTYNQLFEENPINAS
 GVDKAILSARLSKSRLENLIAQLPGEKKNLGFNLIALSLGLTPNFKSNFDLAEDA
 KLQLSKDTHYDDDLNLLAQIGDQYADFLAAKNLSDAILSDILRVNTEITKAPLSAS
 MIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEYFYKFKI
 PILEKMDGTEELLVKNLREDLLRQRTFDNGSIPHQIHLGELHAILRRQEDFYPLKD
 NREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEVVDKGSASQSFIE
 RMTNFDKNLPNEKVLPHKSHLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIV
 DLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIKDKDFL
 DNEENEDILEDIVLTLTLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSR
 KLINGIRDKQSGKTILDFLKSDGFANRNFMLIHDDSLTFKEDIQKAQVSGQGDLSLHE
 HIANLAGSPAIKKGLQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRE
 RMKRIEEGKELGSQLKEHPVENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDY
 DVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKMKMKNYWRQLLAKLI
 TQRKFDNLTKAERGGSELKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKL
 REVKVITLKSFLVSDFRKDFQFYKREINNYHHAHDAYLNAVVGTAIIKKYPKLES
 EFVYGDYKVVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIE
 TNGETGEIVWDKGRDFATVRKVLSPQVNIKKTEVQTGGFSKESILPKRNSDKLIA
 RKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSEFEKN
 PIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNF
 LYLASHYEKLGKSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLTKVLSA
 YNKHHRDKPIREQAENIHLFTLNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSI
 TGLYETRIDLSQLGGDSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDI
 LVHTAYDESTDENVMMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSPKPKKRV

[0150] .

[0151] 추가의 예는 사람 AID(UniProt 참고 번호 Q7Z599), 사람 APOBEC3G (UniProt 참고 번호 Q9HC16), 랫트 APOBEC1(UniProt 참고 번호 P38483), 및 칠성장어 CDA1 (GenBank 참고 번호 EF094822)를 포함하나, 이에 한정되지 않는다. 염기 편집 구현예에서, 염기-편집인자는 Cas9니카제(nickase)를 이용한다. 이는 단일-가닥 브레이크

크를 생성하는 능력은 보유하면서 돌연변이되는 Cas9의 2개의 절단 도메인 중 하나 만을 생성한다. 예를 들면, 도 37에 제공된 예시적인 유전자 편집 작제물은 Cas9 절단 도메인 내에 D10A 돌연변이를 함유할 것이다. 일부 구현예에서, 이러한 시도는 생체내 설정에서 사용될 수 있다.

[0152] 일부 구현예에서, 재조합 발현 시스템에서 제1 및/또는 제2 벡터는 사이티딘의 우리딘으로의 전환을 지시하므로 접-돌연변이를 고정시키는데 유용한 사이티딘 데아미나제 효소를 암호화한다.

[0153] **억제 및 활성화:** 일부 국면은 게놈 조절을 위해 dCas9를 사용하는 재조합 발현 시스템의 용도에 관한 것이다. 전통적인 CRISPR-Cas9 모델에 따른 유전자 편집을 사용한 하나의 개념은 유전자를 영구적으로 편집한 후 발생할 수 있는 알려지지 않은 효과이다. 이는 효소와 관련된 알려지지 않은 기능 및 무차별한 활성을 지닌 유전자가 많으므로, 중요하다. 이러한 이유로, 게놈 조절은 유전자 편집으로부터 올 수 있는 가능한 결과없이 유전자 발현을 조절할 수 있도록 하므로, 매력적인 대안이다. 일부 구현예에서, 시스템은 조절된 유전자 발현을 위해 구성된다.

[0154] 일부 구현예에서, 전사 활성인자 또는 전사 억제인자는 dCas9를 사용하는 재조합 발현 시스템, 예컨대, split-Cas9 이중 AAV 시스템 내로 임의로 혼입된다 이러한 구현예에서, gRNA는 표적 유전자의 프로모터를 표적화하도록 설계된다.

[0155] 비-제한적인 예시적인 전사 억제인자는 Krueppel-관련된 박스("KRAB")이며, 이는 고등 척추동물에서 고도로 보존된 전사 억제 모듈이며, 이의 예시적인 서열은 하기 제공된다:

[0156] KRAB

DAKSLTAWSRITLVTFKDFVDFDTREEWKLDDTAQQIVYRNVMLENYKNL
VSLGYQLTKPDVILRLEKGEEP

[0157] 또는 이의 생물학적 등가물.

[0158] 비-제한적인 예시적인 전사 활성인자는 하기 제공된 예시적인 서열인, VP74, RTa, 및 p65이다:

VP64

GSGRADALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLDMLGSDALDD
FDLDMLIN

RTa

RDSREGMFLPKPEAGSAISDVFEGRVCQPKRIRPFHPPGSPWANRPLPA
SLAPTPTGPVHEPVGSLTPAPVPQPLDPAPA VTPEASHLLEDPDEETSQA
VKALREMADTVIPQKEEAICGQMDLSHPPPRGHLDELTTTLESMTEDLN

P65

SQYLPDTPDRHRIEKRKRTYETFKSIMKKSFPFGPTDPRPPRRRIA VPSRSSASVPKP
APQYPFTSSLSTINYDEFPTMVFPSGQISQASALAPAPPQVLPQAPAPAPAMVSAL
AQAPAPVPLAPGPPQAVAPPAPKPTQAGEGTLSEALLQLQFDDDELGALLGNSTDP
AVFTDLASVDNSEFQQLLNQGIPVAPHTTEPMLMEYPAITRLVTGAQRPPDPAPAPL
GAPGLPNGLLSGDEDFSSIADMDFSALL

[0159] 또는 이의 각각의 생물학적 등가물.

[0160] 일부 구현예에서, 재조합 발현 시스템에서 제1 및/또는 제2 벡터는 KRAB를 포함한다. 추가의 구현예에서, 이러한 재조합 발현 시스템은 표적 유전자를 사일런스하거나, 억제하거나, 하향조절하는데 사용된다. 여전히 추가의 구현예에서, 재조합 발현 시스템은 표적 유전자에 대해 프로모터를 표적화하는 gRNA를 포함한다.

[0161] 본 출원인은 이러한 시스템을 시험관내 및 생체내에서 시험하고, 시험관 내에서 90% 이하의 억제 및 생체내에서 35% 이하의 억제를 나타내었다(각각 도 8 및 9).

[0162] 일부 구현예에서, 재조합 발현 시스템에서 제1 및/또는 제2 벡터는 VP64, RTa, 및/또는 p65를 포함한다. 추가의 구현예에서, 이러한 재조합 발현 시스템을 사용하여 표적 유전자를 활성화시키거나, 과발현시키거나 상향조절할 수 있다. 여전히 추가의 구현예에서, 재조합 발현 시스템은 표적 유전자에 대한 프로모터를 표적화하는 gRNA를 포함한다. 표적 유전자의 활성화, 과발현, 또는 상향 조절과 관련된 구현예에서, 재조합 발현 시스템은 또한 활성화, 과발현, 또는 상향조절을 위한 표적 유전자를 암호화하는 제3 벡터를 포함할 수 있다.

[0163] 본 출원인은 40배 까지의 시험관내 상대적 발현에 있어서의 증가를 측정하였다(도 11).

[0164] **메틸화:** 일부 구현예에서, 메틸화의 조절인자는 재조합 발현 시스템내로 임의로 혼입되므로; 표적 유전자의 후생적 변형을 허용한다. 이러한 구현예에서, gRNA는 표적 유전자의 프로모터를 표적화하도록 설계될 수 있다.

[0165] 메틸화의 이러한 조절인자의 비-제한적 예는 DNMT3A 및 DNMT3L을 포함하나 이에 한정되지 않으며; 이의 예시적인 서열은 하기 제공된다:

DNMT3A

```
TYGLRRREDWPSRLQMFFANNHDQEFDPKVVYPPVPAEKRPKIRVLSLFDGIATGL
LVLKDLGIQVDRIASEVCEDESITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDL
VIGGSPCNDLSIVNPARKGLYEGTGRLLFFEFYRLLHDARPKEGDDRRPFFWLFENVVA
MGVSDKRDRISRFLESNPVMIDAKEVSAHRARYFWGNLPGMNRPLASTVNDKLELQ
ECLEHGRIAKFSKVRTITTRSNSIKQGKDQHFPVFMNEKEDILWCTEMERVFGFPVHY
TDVSNMSRLARQLLGRSWSVPVIRHLFAPLKEYFACV
```

DNMT3L

```
GSSSELSSSVSPGTGRDLIAEYVKANQRNIEDICCGSLQVHTQHPLFEGGICAPCKDK
FLDALFLYDDDDGYQSYCSICCSGETLLICGNPDCTRCYCFECVDSLVPGTSGKVHA
MSNWVCYLCLPSSRSGLLQRRRWRSQKAFYDRESENPLEMFETVPVWRRQPVR
VLSLFDIKKELTSLGFLESGSDPGQLKHVVVDVTDTVRKDVVEEWGPFDLVYGATPPL
GHTCDRPPSWYLFQFHRLQLYARPKPGSPRPFWMFVDNLVLNKEDLDVASRFLEM
EPVTIPDVHGGSLQNAVVRVWSNIPAIRSRHWALVSEEELSLAQNKQSSKLAAKWPT
KLVKNCFLPLREYFKYFSTELTSSL
```

[0166] ... 또는 이의 각각의 생물학적 증가물.

[0167] 일부 구현예에서, 재조합 발현 시스템에서 제1 및/또는 제2 벡터는 DNMT3A 및 DNMT3L 중 하나 이상을 포함한다. 추가의 구현예에서, 이러한 재조합 발현 시스템을 임의로 사용하여 이의 메틸화를 변경시킴으로써 표적 유전자를 사일런스하거나, 억제하거나, 하향조절한다. 여전히 추가의 구현예에서, 재조합 발현 시스템은 표적 유전자에 대해 프로모터를 표적화하는 gRNA를 포함한다.

[0168] 특이적인 용도를 위한 gRNA

[0169] 일부 구현예에서, 재조합 발현 시스템은 gRNA를 포함하며 본원에 사용된 gRNA를 기반으로 특수한 용도로 조절된다. 따라서, 일부 구현예에서, 재조합 발현 시스템의 제1 또는 제2 벡터는 gRNA를 암호화한다. 다른 구현예에서, 재조합 발현 시스템은 gRNA를 암호화하는 제3 벡터를 포함한다. 일부 구현예에서, gRNA는 이중 gRNA(dgRNA) 또는 단일 gRNA(sgRNA)이다.

[0170] 이에 대해 gRNA가 조절되는 비-제한적인 예시적인 방법 국면이 본원에 개시되어 있다. 예시적인 gRNA가 제공되는 경우, 대문자는 엑손 영역을 나타내고 소문자는 인트론 영역을 나타낸다.

[0171] 개시된 gRNA가 특수한 포유동물 중, 예컨대, 마우스 또는 사람에 대해 설계될 수 있지만, 이에 대한 동종 유전자 및 gRNA는 GenBank, BLAST, UniProt, SwissProt, KEGG, 및 GeneCard를 포함하나 이에 한정되지 않는 단백질 및 유전자 데이터베이스와 같은, 당해 분야에 공지된 기술 및 도구를 사용하여 찾을 수 있는 것으로 인정된다. 또한, 특수한 표적 및 종에 대한 입증된 gRNA 서열은 Cas 데이터베이스(rgenome.net/cas-database/)와 같은, 많은 gRNA 데이터베이스 중 하나 또는 AddGene (addgene.org/crispr/reference/grna-서열/) 또는 GeneScript(genscript.com/gRNA-database.html)를 통해 찾을 수 있다. gRNA 및/또는 표적 유전자는 이러한 비-제한된 예시적인 방법 및/또는 gRNA 및/또는 표적 유전자와 관련된 임의의 다른 질환 또는 장애에 대한 재조합 발현 시스템에 의해 표적화될 수 있다.

[0172] 용어 "억제하다"가 본원에 사용되는 경우 이는 KRAB; dCas9; 및 하나 이상의 개시된 gRNA와 같은, 그러나 이에 한정되지 않는 전사 억제인자를 사용하는 재조합 발현 시스템과 함께 사용하기 위한 참고를 의도하며; 이러한 용어는 이의 하향조절, 억제, 및/또는 사일런싱과 같은 이의 발현을 감소시키거나 제거하는 표적 유전자에 있어서의 효과를 의도한다. 유사하게, 용어 "활성화하다" 또는 "과발현하다"가 본원에 사용된 경우, 이는 VP64, RTa, 및 p65; dCas9; 및 하나 이상의 개시된 gRNA를 포함하나, 이에 한정되지 않은 전사 활성인자를 사용하는 재조합 발현 시스템을 의도하며; 이러한 용어는 이의 상향조절, 활성화 및/또는 과발현과 같은 이의 발현을 증가시키는 표적 유전자에 있어서의 효과를 의도한다. 보다 일반적으로, "조절"은 dCas9를 사용하는 재조합 발현 시스템과 함께 사용하기 위한 gRNA에 대한 참고시 사용할 수 있는 반면, "편집"은 활성(또는 "살아있는") Cas9를

사용하는 재조합 발현 시스템과 함께 사용하기 위한 gRNA를 참고로 사용될 수 있다.

[0173] **통증 관리:** 일부 구현예에서, gRNA는 재조합 발현 시스템에서 사용되어 통증 관리를 표적화한다. 장기간 오피오이드(opioid) 사용은 약물 중독 및 약물 남용과 연결되어 왔으며, 전세계적으로 대략 3240만 명의 사람들이 약물을 남용하고 있다. 또한, 서방 및 중유럽에서 최초 약물 재활 환자의 16%, 아시아에서 45%, 및 북아메리카에서 22%가 오피오이드 남용을 위한 치료를 추구하고 있는 중이다. 또한, 최근의 보고서는 모르핀의 사용과 만성 협착 손상의 기간의 2배를 연관시켰으며 연장된 통증이 만성 통증의 경우 오피오이드의 과다 사용의 결과임을 예측하였다. 이러한 이유로, 통증을 표적화하는 대안적인 방식은 전세계적 인구에게 크게 유리할 수 있다. 낮은 내지 높은 통증 비민감성(insensitivity)을 갖는, 오피오이드 펩타이드에 관여하는 유전자내 증가된 발현과 함께 SCN9A 유전자(전압-게이트된(gated) 나트륨 채널 Nav 1.7을 암호화함)에서 기능 돌연변이가 손실된 사람 및 마우스가 존재하는 것으로 알려져 있다. 사람 및 마우스는 SCN9A 내에 점 돌연변이를 가짐으로써, 단일 아미노산의 치환 및 하나의 프레임내 결실(in-frame deletion)을 유발하는 18개의 미스센스 돌연변이를 포함하는, 이의 표gus형을 생성한다.

[0174] 사람 SCN9a 설계

- 1: GGAAAGCCGACAGCCGCCGC
- 2: GCGCGGGCTCTCCTTCCC
- 3: GAGCACGGGCGAAAGACCGA
- 4: GTGTGCTCTTAAGGGGTGCG
- 5: GTGGCGTTGAGGCGAGCAC

[0175]

[0176] 마우스 SCN9 설계

- 1: GACCCATGTAACAACCTCCAC
- 2: GTGTATATTGTTGAACCCGT
- 3: AACAACTCCACTGGAGTAGA
- 4: CAAACTGTTAAGAAACGGGC
- 5: GGTTCCTGGCAAAAATTGCTGT

[0178]

[0179] 또는 이의 각각의 생물학적 등가물.

[0180] 이론에 얽매이지 않고, 본 출원인은 활성 Cas9를 사용하는 것이 통증에 대한 영구적인 불감성 및/또는 후각 감각의 상실을 유발할 수 있는 정도로 통증 관리에 대한 위험을 가지고 있다고 믿는다. 구체적으로, 본 출원인은 SCN9A 유전자내 돌연변이가 또한 후각 뉴우런에 있어서 기능적 NAV1.7 나트륨 채널의 손실을 유발하여 후각 감각의 손실을 야기함을 알고 있다. 따라서, 상기 제공된 예시적인 gRNA는 SCN9A의 프로모터 영역을 표적화하도록 설계되며 dCas9를 사용하는 본원에 개시된 재조합 발현 시스템의 구현예에서 사용될 수 있다. 이러한 gRNA를 사용하는 의도는 SCN9A를 사일런스하거나 하향조절하기 위한 것일 수 있다.

[0181] 예를 들면, 일 국면에서, 출원인은 dCas0를 사용하는, 개시된 재조합 발현 시스템, 예컨대, 이중 pAAV9_gSCN9a_dCas9 시스템을 (i) 수술 동안 통증의 방지를 위해(여기서 환자에게 수술 전 재조합 발현 시스템을 투여한다), 또는 (ii) 만성 통증의 사용을 위해 이용한다. 이론에 얽매이지 않고, 재조합 발현 시스템의 양은 환자가 약 1개월 동안 저하된 통증을 갖는데 효과적일 수 있다.

[0182] 통증 관리를 위해 표적화될 수 있는 추가의 유전자는 Nav 1.8(SCN10A 유전자), 1.9(SCN11A 유전자) 및 1.3(SCN3A 유전자) 뿐만 아니라 또한 캡사이신 수용체 및 바닐로이드 수용체 1으로 또한 알려진 일시적인 수용체 잠재적인 양이온 채널 서브계열 V 구성원 1(TrpV1)도 포함한다. 또한 억제되거나 활성화될 다른 목적 유전자는 다음과 같다.

유전자	재조합 발현 시스템의 효과
SHANK3(예컨대, 수탁 번호 제JX122810.1호)	억제/녹아웃
NMDA 수용체 길항제(NR2B 포함(예컨대, 수탁번호 제NM_000834.4호))	억제/녹아웃
IL-10(예컨대, 수탁번호 제NM_000572.2호)	활성화(과발현)
Penk(예컨대, 수탁번호 제NM_001135690.2호)	활성화(과발현)
Pomc(예컨대, 수탁번호 제NM_001035256.2호)	활성화(과발현)
MVIIA-PC(예컨대, 수탁번호 제FJ959111호)	활성화(과발현)

[0183]

[0184] 명명된 표적 중 일부에 대해 사용될 수 있는 gRNA의 비제한적 예는 다음을 포함한다:

[0185] 녹아웃의 경우 gRNA:

Nav 1.3: TCGTGGATTCTATCACTTT

Nav 1.8: CTTGGTAACGCTCTTCTTTG

Nav 1.9: CGATGGTTCCACGTGCAATA

TrpV1: TAAGCTGAATAACACCGTTG

[0186]

[0187] 억제제의 경우 gRNA:

Nav 1.3: CCGCTTCCTGTTCTGAGATC

Nav 1.8: GTCACGAGTCCACCCTGCC

Nav 1.9: CAGCCTGGATGGCTTACCTC

TrpV1: GGGACTTACCAGCTAGGTGC

[0188]

[0189] 또는 이의 각각의 생물학적 등가물. 여전히 추가의 예시적인 gRNA는 하기 본원에 제공된다:

사람에서, 억제용 gRNA			
sgID	유전자	전사체	광스페이저 서열
SCN3A+_166060543.23-PIP2	SCN3A	PIP2	GATCTCAGAACAGGAAGCGG
SCN3A+_166060199.23-PIP2	SCN3A	PIP2	GTGTAAATTACAGGAACCAA
SCN3A_-166060301.23-PIP2	SCN3A	PIP2	GACCTGGTAGCTAGGTTCTA
SCN3A+166060552.23-	SCN3A	PIP2	GATAGAGTGAATCTCAGAAC

[0190]

PIP2			
SCN3A_+166060129.23-PIP2	SCN3A	PIP2	GAATAGAGCCTGTCTGGAAA
SCN3A_+166060346.23-PIP2	SCN3A	PIP2	GTGTTATGCTGTAATTCATA
SCN3A_+166060119.23-PIP2	SCN3A	PIP2	GGTCTGGAAATGGTGATTIA
SCN3A_+166060135.23-PIP2	SCN3A	PIP2	GAAAGAAAATAGAGCCTGTC
SCN3A_+166060371.23-PIP2	SCN3A	PIP2	GCCTAACCATCTTGGATGCT
SCN3A_+166060281.23-PIP2	SCN3A	PIP2	GACCATAGAACCTAGCTACC
SCN9A_+167232419.23-PIP2	SCN9A	PIP2	GGCGGTCGCCAGCGCTCCAG
SCN9A_+167232052.23-PIP2	SCN9A	PIP2	GCCACCTGGAAAAGAAGAGAG
SCN9A_+167232416.23-PIP2	SCN9A	PIP2	GGTCGCCAGCGCTCCAGCGG
SCN9A_+167232010.23-PIP2	SCN9A	PIP2	GCCAGCAATGGGAGGAAGAA
SCN9A_-167232085.23-PIP2	SCN9A	PIP2	GTTCCAGGTGGCGTAATACA
SCN9A_+167232476.23-PIP2	SCN9A	PIP2	GGCGGGGCTGCTACCTCCAC
SCN9A_+167232437.23-PIP2	SCN9A	PIP2	GGGCGCAGTCTGCTGCAGG
SCN9A_+167232409.23-PIP2	SCN9A	PIP2	GGCGCTCCAGCGCGGCTGT
SCN9A_+167232021.23-PIP2	SCN9A	PIP2	GACCGGGTGGTTCAGCAAT
SCN9A_+167232018.23-PIP2	SCN9A	PIP2	GGGGTGGTTCAGCAATGGG
SCN10A_-38835462.23-ENST00000449082.2	SCN10A	ENST00000449082.2	GTGACTCCGGAGTAAAGCGA
SCN10A_-38835311.23-ENST00000449082.2	SCN10A	ENST00000449082.2	GGGAGCTCACCATAGAACTT
SCN10A_-38835269.23-ENST00000449082.2	SCN10A	ENST00000449082.2	GACGGATCTAGATCCTCCAG
SCN10A_+38835213.23-ENST00000449082.2	SCN10A	ENST00000449082.2	GCCGGGTAAAGGCTACTAGT
SCN10A_-38835251.23-ENST00000449082.2	SCN10A	ENST00000449082.2	GCCCGGTGTGTGCTGTAGAA
SCN10A_+38835434.23-ENST00000449082.2	SCN10A	ENST00000449082.2	GTTTACTCCGGAGTCACTGG
SCN10A_-38835449.23-ENST00000449082.2	SCN10A	ENST00000449082.2	GCTATCTCCACCAGTGACTC
SCN10A_-38835156.23-ENST00000449082.2	SCN10A	ENST00000449082.2	GACATCACCCAGGGCCAAGG
SCN10A_-38835491.23-ENST00000449082.2	SCN10A	ENST00000449082.2	GTAGTTTCGAGGGATCCAAT
SCN10A_+38835272.23-ENST00000449082.2	SCN10A	ENST00000449082.2	GCTCCAGCAGAACTGATCG
SCN11A_-38991624.23-ENST0000030232.8.3.ENST00000450244.1	SCN11A	ENST0000030232.8.3.ENST00000450244.1	GATGGGTCCAAGTCTTCCAG
SCN11A_+38992032.23-ENST0000030232.8.3.ENST00000450244.1	SCN11A	ENST0000030232.8.3.ENST00000450244.1	GGTTCTTGCTATACCCACAG
SCN11A_-38991801.23-	SCN11A	ENST0000030232	GCCAGAGAGTCGGAAGTGAA

[0191]

ENST00000302328.3,ENST0000450244.1		8.3.ENST00000450244.1	
SCN11A_+_38992029.23-ENST00000302328.3,ENST0000450244.1	SCN11A	ENST00000302328.3,ENST00000450244.1	GCCTGTATATACCCACAGTGG
SCN11A_+_38991609.23-ENST00000302328.3,ENST0000450244.1	SCN11A	ENST00000302328.3,ENST00000450244.1	GGGAAAGCCTCTGGAAAGACT
SCN11A_+_38992040.23-ENST00000302328.3,ENST0000450244.1	SCN11A	ENST00000302328.3,ENST00000450244.1	GGAAAGATGACCACCACTG
SCN11A_+_38991666.23-ENST00000302328.3,ENST0000450244.1	SCN11A	ENST00000302328.3,ENST00000450244.1	GGAAATGTCGCATAGAGCTT
SCN11A_+_38991618.23-ENST00000302328.3,ENST0000450244.1	SCN11A	ENST00000302328.3,ENST00000450244.1	GGAGTTCATAGGAAAGCCTC
SCN11A_+_38991924.23-ENST00000302328.3,ENST0000450244.1	SCN11A	ENST00000302328.3,ENST00000450244.1	GCITTAAGACTGGAATCCTA
SCN11A_+_38991653.23-ENST00000302328.3,ENST0000450244.1	SCN11A	ENST00000302328.3,ENST00000450244.1	GGGAAGTTGCCCAAGCTCTA
SHANK3_+_51135959.23-PIP2	SHANK3	PIP2	GGAAATTCGAATACAGCTCCT
SHANK3_+_51136404.23-PIP2	SHANK3	PIP2	GCTTCAGGCAGAGACCCCGG
SHANK3_+_51136356.23-PIP2	SHANK3	PIP2	GGAGCCTCCGTGGTGACACA
SHANK3_+_51136302.23-PIP2	SHANK3	PIP2	GCACGGCAGGAACCTCCCC
SHANK3_+_51136319.23-PIP2	SHANK3	PIP2	GAGCACCGGAGGGACCCGCA
SHANK3_+_51136333.23-PIP2	SHANK3	PIP2	GGCCCGGAACGACAGAGCAC
SHANK3_+_51136329.23-PIP2	SHANK3	PIP2	GGGAACGACAGAGCACCCGGA
SHANK3_+_51136143.23-PIP2	SHANK3	PIP2	GACcgcggcagggcctgaa
SHANK3_+_51136336.23-PIP2	SHANK3	PIP2	GCCTGCCGTGCGGGTCCCTC
SHANK3_+_51135950.23-PIP2	SHANK3	PIP2	GTACAGCTCTGGGCGCGCC
TRPV1_+_3500355.23-PIP2	TRPV1	PIP2	GAGCGACTCCTGCTAGTGCA
TRPV1_+_3500317.23-PIP2	TRPV1	PIP2	GCGGGCCCGGACCCACGG
TRPV1_+_3499964.23-PIP2	TRPV1	PIP2	GCTCCTTGAAGACCTGGG
TRPV1_+_3500391.23-PIP2	TRPV1	PIP2	GAGTCGCTGTGGACGCCCTT
TRPV1_+_3500224.23-PIP2	TRPV1	PIP2	GGGACTACCCAGCTAGACGC
TRPV1_+_3500327.23-PIP2	TRPV1	PIP2	GTGGTCTCCCCGCTCCGTG
TRPV1_+_3500298.23-PIP2	TRPV1	PIP2	GGGGAGAGCTGGGCTCGTGT
TRPV1_+_3500017.23-PIP2	TRPV1	PIP2	Gtgcctcaasggtgtctgtg
TRPV1_+_3499899.23-PIP2	TRPV1	PIP2	GCTGCATCAGCCGTCCTCGG

[0192]

TRPV1_-3500400.23-PIP2	TRPV1	PIP2	GGGACGCCCTTCGGCACTCA
GRIN2B_-14133341.23-PIP2	GRIN2B	PIP2	GGATTCGCGTGTCCCCCGGA
GRIN2B_+14132929.23-PIP2	GRIN2B	PIP2	GGATATGCAAGCGAGAAGAA
GRIN2B_-14132903.23-PIP2	GRIN2B	PIP2	GCTCTAGACGGACAGATTAA
GRIN2B_-14133316.23-PIP2	GRIN2B	PIP2	GGGGGAAAAAGAGCGGGTCA
GRIN2B_+14132924.23-PIP2	GRIN2B	PIP2	GGCAAGCGAGAAGAAGGGAC
GRIN2B_-14133295.23-PIP2	GRIN2B	PIP2	GCCAAAGCGTCCCTTCCTA
GRIN2B_-14133298.23-PIP2	GRIN2B	PIP2	GAAGCGTCCCTTCCTAAGG
GRIN2B_+14132855.23-PIP2	GRIN2B	PIP2	GGCTTCTACAAACCAAGGTA
GRIN2B_+14133247.23-PIP2	GRIN2B	PIP2	GACCATGTCCACCGAGGGA
GRIN2B_+14133252.23-PIP2	GRIN2B	PIP2	GGAATGACCATGTCCACCG
마우스에서, 억제용 gRNA			
sgID	유전자	전사체	광스캐이저 서열
Scn3a_+65567459.23-PIP2	Scn3a	PIP2	GTGAATCTCAGAACAGGAAG
Scn3a_+65567442.23-PIP2	Scn3a	PIP2	GAGCGGAGGCATAAGCAGAA
Scn3a_-65567234.23-PIP2	Scn3a	PIP2	GATCTGGTGGCTAGATTCTA
Scn3a_-65567301.23-PIP2	Scn3a	PIP2	GAGGAATCACAGCTCAACAA
Scn3a_-65567522.23-PIP2	Scn3a	PIP2	GATCAGAAAACGGCCCTGGA
Scn3a_-65567271.23-PIP2	Scn3a	PIP2	GGTTTTGICAGCTTACCTGA
Scn3a_-65567326.23-PIP2	Scn3a	PIP2	GGCATCCAAGATGGTTAGAA
Scn3a_+65567264.23-PIP2	Scn3a	PIP2	GATTCCTAAGGCTTCCATC
Scn3a_+65567031.23-PIP2	Scn3a	PIP2	GCAATACAGACTAGGAATTA
Scn9a_+66634758.23-PIP2	Scn9a	PIP2	GAGCTCAGGGAGCATCGAGG
Scn9a_-66634675.23-PIP2	Scn9a	PIP2	GAGAGTCGCAATTGGAGCGC
Scn9a_-66634637.23-PIP2	Scn9a	PIP2	GCCAGACCAGCCTGCACAGT
Scn9a_-66634689.23-PIP2	Scn9a	PIP2	GAGCGCAGGCTAGGCCTGCA
Scn9a_-66634610.23-PIP2	Scn9a	PIP2	GCTAGGAGTCCGGGATACCC
Scn9a_+66634478.23-PIP2	Scn9a	PIP2	GAATCCGCAGGTGCACTCAC
Scn9a_-66634641.23-PIP2	Scn9a	PIP2	GACCAGCCTGCACAGTGGGC
Scn9a_+66634731.23-PIP2	Scn9a	PIP2	GCGACGCGGTTGGCAGCCGA
Scn10a_+119719110.23-PIP2	Scn10a	PIP2	GGCAGGGTGGAACTCGTGAC
Scn10a_+119719123.23-PIP2	Scn10a	PIP2	GCACCATCCAGCAAGCAGGG
Scn10a_-119719078.23-PIP2	Scn10a	PIP2	GCGTCACTCAAGGATCTACA
Scn10a_+119719086.23-PIP2	Scn10a	PIP2	GATGGGAATGGCACCCACGA
Scn10a_+119718921.23-PIP2	Scn10a	PIP2	GCCTTTAGACGGAGAACAGA
Scn10a_+119719051.23-PIP2	Scn10a	PIP2	GAGATCCTTGAGTGACGGAC

[0193]

Scn10a - 119719025.23-P1P2	Scn10a	PIP2	GCGGGGCTCCTCCACGAAGG
Scn10a - 119719095.23-P1P2	Scn10a	PIP2	GCAAGGAATCACGCCTTCGT
Scn10a + 119718881.23-P1P2	Scn10a	PIP2	GGCCATGCGCGAATGCTGAG
Scn10a + 119719014.23-P1P2	Scn10a	PIP2	GGCAAGCCAGCCACCTTCG
Scn11a - 119825404.23-P1P2	Scn11a	PIP2	GAGGTAAGCCATCCAGGCTG
Scn11a - 119825450.23-P1P2	Scn11a	PIP2	GTTCTGTAGGGAGGCTCA
Scn11a - 119825400.23-P1P2	Scn11a	PIP2	GCCTGAAACGACAGAGGATG
Scn11a + 119825277.23-P1P2	Scn11a	PIP2	GTCAGAGGTGGAGCCAGGT
Scn11a - 119825394.23-P1P2	Scn11a	PIP2	GCCCCAGCCTGAAACGACAG
Scn11a - 119825463.23-P1P2	Scn11a	PIP2	GGCCAAGAGCGAGAATCTCC
Scn11a + 119825246.23-P1P2	Scn11a	PIP2	GGTCAGGTGTCAGAGCCCAT
Scn11a + 119825242.23-P1P2	Scn11a	PIP2	GGGTGTGAGAGCCATCGGT
Scn11a - 119825431.23-P1P2	Scn11a	PIP2	GTGCCCTGAGCCTCCCTAGC
Scn11a - 119825253.23-P1P2	Scn11a	PIP2	GTCGTGAGAACCCGACCGAT
Shank3 + 89499659.23-P1P2	Shank3	PIP2	GGGCCTCCGCGAGGCGAGCGG
Shank3 + 89499688.23-P1P2	Shank3	PIP2	GggggcggcgggggACAG
Shank3 + 89499943.23-P1P2	Shank3	PIP2	GCCCTAGCGGGCCACACAG
Shank3 + 89499679.23-P1P2	Shank3	PIP2	GcgggggACAGCGGCTCCGG
Shank3 + 89499612.23-P1P2	Shank3	PIP2	GCATCGGCCCCGGCTTCGAG
Shank3 + 89499924.23-P1P2	Shank3	PIP2	GGGGTACGGCGAGATCGCAA
Shank3 + 89499878.23-P1P2	Shank3	PIP2	GATGCCGACCGCACGACCA
Shank3 - 89499676.23-P1P2	Shank3	PIP2	GGCCGCCGCCGTGCGCCTG
Shank3 + 89499818.23-P1P2	Shank3	PIP2	GGGGCCCGGACTGTTCCTGG
Shank3 + 89499938.23-P1P2	Shank3	PIP2	GAGCGGGCCACACAGGGGTA
Trpv1 + 73234353.23-P1P2	Trpv1	PIP2	GGGACTTACCAGCTAGGTGC
Trpv1 - 73234330.23-P1P2	Trpv1	PIP2	GCCACAAAGAACAGCTCCA
Trpv1 - 73234384.23-P1P2	Trpv1	PIP2	GGCTGGTAAGTCTCTTCAT
Trpv1 + 73234339.23-P1P2	Trpv1	PIP2	GGGTGACGGCACACTCCAAA
Trpv1 - 73234537.23-P1P2	Trpv1	PIP2	GACTTAACCTGGCTGACTGT
Trpv1 + 73234478.23-P1P2	Trpv1	PIP2	GTCAGCCTCCAGAAAGTCCA
Trpv1 - 73234495.23-P1P2	Trpv1	PIP2	GGCTGCCTTGAATCTGCGG
Trpv1 + 73234635.23-P1P2	Trpv1	PIP2	GCCACGGAAGGCCCTCCAGAT
Trpv1 - 73234346.23-P1P2	Trpv1	PIP2	GCCAAGGCACTTGCTCCAIT
Trpv1 + 73234280.23-P1P2	Trpv1	PIP2	GGGCTGCTGTGTGGTAAGAG
Grin2b - 136172154.23-P1P2	Grin2b	PIP2	GCCAACCTGAATGGAAGAGA
Grin2b - 136172179.23-P1P2	Grin2b	PIP2	GAGGGAAGTGGAAAGCAAGG
Grin2b - 136172123.23-P1P2	Grin2b	PIP2	GTGGGACAGGCATGGATGAA
Grin2b + 136172089.23-P1P2	Grin2b	PIP2	GCCTGTCCAGGAACGGCAT
Grin2b - 136172145.23-P1P2	Grin2b	PIP2	GTGAGAAAAAGCCAACCTGAA
Grin2b - 136171934.23-P1P2	Grin2b	PIP2	GGATTCGAGTGTCTCCCGGA

[0194]

Grin2b - 136171999.23-P1P2	Grin2b	PIP2	GACCAAGTCGTTATAAGGAA
Grin2b - 136172002.23-P1P2	Grin2b	PIP2	GAAGTCGTTATAAGGAAAGG
Grin2b + 136171844.23-P1P2	Grin2b	PIP2	GGAAATGACCACGCTCCACGG
Grin2b + 136172019.23-P1P2	Grin2b	PIP2	GCCTCTGGTGTACTCTGT

[0195]

지 않고, 이러한 유전자 중 하나 이상을 표적화하는 것이 숙주를 감염시키는 이들 질환 중 하나 이상의 능력을 방해하는 것으로 여겨진다. 따라서, 이들 유전자 표적의 조절 또는 편집을 위해 조정된 gRNA를 포함하는 개시된 제조합체 발현 시스템의 사용이 이의 치료 및/또는 예방에 유용할 수 있다. 일부 구현예에서, 이는 이러한 gRNA를 포함하는 제조합 발현 시스템의 예방적 투여를 포함할 수 있다. 말라리아, C형 간염, 또는 이들 유전자가 관여된 임의의 다른 질환과 같은 그러나 이에 한정되지 않는 질환에서 사용하기 위한 gRNA의 비제한적 예는 다음을 포함한다:

CD81: CGAAATTGAAGACGAAGAGC

MUC13: GGAGACTGAGAGAGAAGC

Sf-b1: TGATGAGGGAGGGCACCATG

[0203]

또는 이의 각각의 생물학적 등가물.

[0204]

[0205]

조혈 줄기 세포 치료요법 및 HIV: 일부 구현예에서, gRNA를 설계하여 조혈 줄기 세포(HSC)의 면역 거부를 예방하고/하거나 HIV가 숙주 세포내로 도입되는 것을 예방한다. HSC 유전자 치료요법은 겸상 적혈구성 빈혈과 같은 다양한 사람 조혈 질환을 잠재적으로 치유할 수 있다. 그러나, HSC 유전자 치료요법의 현재의 공정은 매우 복잡하고 고가이다. 현재, 조혈 줄기 세포 이식 과정은 한 사람(공여자)로부터 HSC를 취하여 이들을 다른 사람(수용자) 내로 주입함을 포함한다. 이러한 방법에 대한 일부 결점은 외부체로부터 오는 세포에 기인한 면역 반응(또는 이식 거부)를 포함한다. 거부를 예방하기 위하여, 많은 환자들은 또한 화학치료요법 및/또는 방사선 치료요법을 필요로 할 수 있으며, 이는 자체적으로 환자를 약하게 한다. 다른 단점은 이식체 대 숙주 질환(GVHD)이며, 여기서 공여자로부터의 성숙한 T-세포는 수용자 조직을 외부로 인식하여 이러한 조직을 공격한다. 이러한 경우에, 수용자는 염증 및 T-세포 활성화를 억제시키는 의약을 복용하여야만 한다. 흥미롭게도, CCR5 공-수용체가 HSC 이식체의 거부 및 숙주 세포로 도입되는 HIV의 능력과 관련되어 있다. 실제로, HIV에 대해 내성인 사람은 CCR5-델타 32라고 불리는, CCR5 유전자내 돌연변이를 가지며, 이는 HIV가 세포를 감염시키지 않도록 하는 트렁크이트된 단백질(truncated protein)을 생성한다. 따라서, 적용 둘 다에서, CCR5를 표적화하는 gRNA를 사용한 제조합 발현 시스템을 이용할 수 있다. 비-제한적인 예시적인 gRNA가 제공된다:

[0206]

CCR5 gRNA: GGTCCTGCCGCTGCTTGTC

[0207]

또는 이의 생물학적 등가물.

[0208]

암 면역치료요법: 암 면역치료요법은 면역계의 구성성분을 사용하여 항체 또는 가공된 T-세포를 사용한 암 세포에 대한 신체의 자체의 면역 반응을 향상시킴으로써, 암을 박멸한다. 전형적으로, T-세포 기반 치료요법은 환자로부터 면역 세포의 추출에 이은 농축, 편집 또는 치료 후 재-주입을 포함한다. PDCD-1은 T-세포 면역 반응을 중단시키는데 중요한 역할을 하며, 이를 녹아웃시키는 것은 암 세포를 제거하는 T-세포의 능력을 증진시킬 수 있으며, 이들 가공된 면역 세포를 사용한 치료는 진전된 암을 지닌 환자에서 일부 현저한 반응을 생성하였다. 또한, 암 관련되지 않은 면역 반응은 또한 이러한 시도로 조절될 수 있다. 이러한 목적을 위해 PDCD-1을 표적화하는 gRNA를 지닌 예시적인 제조합 발현 시스템이 본원에 개시되어 있다. 비-제한적 예시적인 gRNA는 다음을 포함한다;

[0209]

PDCD-1 표적 서열:

1. AGCCGGCCAGTTCCAAACCC

2. AGGGCCCGGCGCAATGACAG

[0210]

[0211]

또는 이의 각각의 생물학적 등가물.

[0212]

시그널링 경로의 비정상적인 활성화는 암을 초래할 수 있다. 예를 들면, 노달(TGF-β 계열, 예컨대, Uniprot 참고 번호 Q96S42)의 하향조절은 전이성 흑색종과 관련된 분자의 하향조절을 유발할 수 있으며 고습도치 경로(hedgehog pathway)를 차단하는 것은 중앙 성장을 예방할 수 있음이 입증되었다. 따라서, 제조합 발현 시스템을 사용하여 이러한 경로 내에서 표적 유전자를 하향조절 할 수 있으므로 이들 표적에 대해 특이적인 gRNA를 설계함으로써 암을 치료하는데 사용할 수 있다.

[0213]

골수증식성 암의 큰 분획은 JAK-2(예컨대, Uniprot 참고 번호 O60674)에서 V617F 돌연변이를 나타낸다. 그러나, 이러한 돌연변이는 개인의 HSC 집단 속에서 너무 많은 gRNA가 지속되도록 하여 HSC 집단에서 V617F 돌연변이를

표적화하는 것은 또한 본 개시내용의 영역내에 있다.

[0214] **혈액 질환:** 말라리아의 임상 증상은 적혈구를 침입하여 이 속에 잔류하는 플라스모디움 기생충의 생명-주기의 혈액 단계 동안에 발생하여, 숙주 단백질 및 자원을 이들 자신의 요구에 대해 사용하도록 하여, 숙주 세포의 형질전환을 초래한다. Duffy, 글리코포린 A/C, 등과 같은 특성의 세포 표면 수용체는 적혈구내로 기생충의 도입에 필수적임이 밝혀졌다. 또한, 기생충은 적혈구 속의 피루베이트 키나제에 크게 의존한다. 이들 유전자의 녹아웃은 플라스모디움 침입에 대해 내성을 부여하는 것으로 여겨진다. 다음의 비-제한된 예시적인 gRNA는 이러한 목적을 위해 작제물에 대해 제공된다:

GYP A

1. TCTTCAAATAACCACTCCTG
2. TCAGCAACAATGTCAACACC

GYP C

1. GGCAATCTCCATAATGCCGT
2. TATCCACAGAGCCTAACCCA

[0215]

PKLR

1. TGTACGAAAAGCCAGTGATG
2. GGGITCACTCCAGACCTGTG

ACKR1 (Duffy)

1. AAGGICTGAGAATCGCGAAG
2. CATTCTGGCAGAGTTAGCAG

[0216]

또는 이의 각각의 생물학적 등가물.

[0217]

[0218] **근육 디스트로피(muscular dystrophy):** 비정상적인 디스트로핀은 다른 유전자들 중에서 근육 디스트로피와 관련되어 있다. 표 1에는 근육 디스트로피 및 다른 신경변성 질환에서 사용하기 위한 예시적인 gRNA가 개시되어 있다.

[0219] **자궁내 태아 특이적인 표적화:** 특이적인 gRNA를 예를 들면, 태아의 아빠로부터 담체 돌연변이에 대해 설계할 수 있으며, 이는 재조합 발현 시스템이 태아에 특이적으로 표적화하고 자궁내에서 모체는 표적화하지 않는다. 따라서, 태아가 모체에 존재하지 않는 질환이 있는 유전형과 함께 존재하는 경우, 이는 모체 게놈에 영향을 미치지 않고 자궁내에서 해결될 수 있다.

[0220] **사이토크롬 P450-기반 장애:** 사이토크롬 P450 효소 CYP2D6(예컨대, UniProt 참고 번호 P10635)은 변화된 약물 물질대사와 관련된 것으로 알려져 있다. 특정 집단(예컨대, 백인)의 퍼센트로 표현된 이러한 효소의 다형성(polymorphism)은 코데인의 통증-완화약물인 모르핀으로의 전환을 방지한다. CYP2D6의 적어도 2개의 활성 또는 기능성 카피가 코데인의 신속하고 완전한 물질대사에 요구된다. CYP2D6의 2개의 불활성 카피를 갖는 환자의 경우, 환자에서 CYP2D6의 적어도 하나의 활성 카피를 활성화시키거나 과발현하는 재조합 발현 시스템에서 gRNA를 제공하는 것은 코데인의 물질대사를 허용한다.

[0221] 특성의 기질의 존재 또는 특성의 생리학적 조건에 대한 노출시, 사이토크롬 P450(CYP)은 반응성 산화 종(ROS)을 생산할 수 있거나 정상의 물질대사를 파괴하거나 체내에서 조직을 손상시키는 물질대사를 불러일으킬 수 있다. CYP 유전자의 활성화 또는 억제될 수 있는 것은 따라서 약물-약물 상호작용으로부터 뿐 아니라 물질대사 보조인자의 비정상적인 수준을 생성하는 상태로부터도 독성을 방지할 수 있다.

[0222] 보다 일반적으로, 일관성이 없는 약물 반응은 환자에게 유리한 차세대 약물-약물 상호작용을 유발하도록 설계된, 표적화된 gRNA를 사용하여 지정할 수 있다.

[0223] **재프로그래밍 대식구:** 대식구는 케모킨 및 사이토키네 의해 양극화된 상이한 소집단을 함유하며 궁극적으로 면역 반응이 전-염증성인지 또는 전-재생성인지의 여부에 영향을 미친다. 특이적인 gRNA를 재조합 발현 시스템에

서 사용하여 대식구를 표적화하고 전-재생성 상태를 위해 M2 대식구에 대해 표현형을 구현할 수 있다.

[0224] **모기 퇴치:** 원인은 크게 알려져 있지 않은 것으로 보이지만, 모기 및 다른 곤충은 다른 사람은 여전히 피하면서 특성의 사람을 무는 것을 선호한다. 쌍생아 연구는 이러한 유인에 유전적인 구성성분이 존재하는 것으로 여겨 짐을 나타내지만, 구체적인 유전자는 알려져 있지 않다. 모기 유인에 영향을 미치는 다른 인자는 숙주가 풍기는 냄새이다. 이러한 유인을 유발하는 유전자를 변경하거나 개인이 모기를 퇴치하는 물질을 생산하도록 할 수 있는 gRNA의 선택을 통하여, 재조합 발현 시스템은 질환-운반 곤충으로 알려진 사람들이 방문하는 지역에 대해 기간 보호를 제공할 수 있다. 최종적으로 모기에 대해 방어할 수 있는, 골수에서 HSC를 표적화하는 gRNA는 또한 본 개시내용의 영역 내에 있다.

[0225] **알츠하이머:** 연구자들은 LilrB2(예컨대, UniProt 참고 번호 Q8N423)에 대한 B-아밀로이드의 결합이 알츠하이머 로 이끄는 첫번째 단계들 중 하나임을 밝혔다. 따라서, gRNA는 재조합 발현 시스템에서 사용하기 위해 본원에서 고려되고 있으며, 이는 궁극적으로 LilrB2의 D1D2 영역 내에서 점 돌연변이를 유발함으로써 이것이 B-아밀로이드가 알츠하이머의 발병을 예방하는데 영향을 미칠 수 있다. D1은 Uniprot 참고 번호 P21728과 관련되어 있다. D2는 Uniprot 참고 번호 14416과 관련되어 있다. 이의 비-제한적인 예시적인 서열은 하기 본원에 제공된다:

[0226] 도파민 수용체 D1

```

10      20      30      40      50
MRTLNTSAMD GTGLVVERDF SVRILTACFL SLLILSTLLG NTLVCAAVIR
60      70      80      90      100
FRHLRSKVTN FFVISLAVSD LLVAVLVMPW KAVAEIAGFW PFGSFCNIWV
110     120     130     140     150
AFDIMCSTAS ILNLCVISVD RYWAISSPFR YERKMTPKAA FILISVAVTL
160     170     180     190     200
SVLISFIPVQ LSWHKAKPTS PSDGNATSLA ETIDNCDSSL SRTYAISSSV
210     220     230     240     250
ISFYIPVAIM IVTYTRIYRI AQKQIRBIAA LERAAVHAKN CQITTGNGKP
260     270     280     290     300
VECSQPESSE KMSFKRETRV LKTLVIMGV FVCCWLPFFI LNCILPFCGS
310     320     330     340     350
GETQPFICDS NIFDVFVWFG WANSSLNPII YAFNADFRKA FSILLGCYRL
360     370     380     390     400
CPATINNAIET VSINNINGAAM FSSHHEPRGS ISKECNLVYL IPHAVGSSSED
410     420     430     440
LKKEEAAGIA RELEKLSPAL SVILDYDIDV SLEKIQBITQ NGQHPT
    
```

[0227]

[0228] 도파민 수용체 D2

```

10      20      30      40      50
MDPLNLSWYD DDLERQNSWR PFNGSDGKAD RPHYNYATL LILLIAVIVF
60      70      80      90      100
GNVLVCMVAVS REKALQTTIN YLIVSLAVAD LLVATLVMPW VVILEVVGWV
110     120     130     140     150
KFSRIHCDIF VILDVMCTA SILNLCAISI DRYTAVAMFM LYNTRYSSKR
160     170     180     190     200
RVIVMISIVW VLSFTISCP LFGLNADQN ECILANPAFV VYSSIVSEYV
210     220     230     240     250
PFIVTLVYI KIIYIVLRRRR KRVNTRSSR AFRAHLRAPL KGNCTHPEDM
260     270     280     290     300
KLCTVIMKSN GSFFVNRVV EAARRAQELE MEMLSSTSP ERIRYSPIPP
310     320     330     340     350
SHHQLTLPDP SHHGLHSTPD SPAKPERNGH AKDHPKIAKI FEIQTMPNGK
360     370     380     390     400
TRISLKTMSR RKLSQQKEKK ATQMLAIVLG VFIIICWLPFF ITHILNIHCD
410     420     430     440
CNIPEVLYSA FTWLVVNSA VNPFIYTIFN IEFKAFLLK LHC
    
```

[0229]

[0230] **갑상선 호르몬 생산:** 갑상선 장애(갑상선 기능항진증 및 갑상선 기능저하증 둘 모두)는 많은 세트의 사람 집단 에 영향을 미친다. gRNA는 갑상선 호르몬의 조절을 허용할 수 있고 이들 장애를 치료 또는 예방하도록 할 수 있는 재조합 발현 시스템에서 사용하기 위해 선택된다.

[0231] 효과기 성분의 주문

[0232] 본원에 개시된 효과기 성분은 본원에 개시된 재조합 발현 시스템, 예컨대, split-Cas9 시스템에서 2개의 벡터 각각에서 이용가능한 공간에 의존하여 다양화 방식으로 구성될 수 있음을 인지하여야 한다. 또한, 본원에 개시 된 효과기 성분은 완전한 Cas9 단백질을 암호화하는 하나의 벡터 및 CRISPR-기반의 게놈 또는 에피게놈 편집을 위한 필수적인 gRNA를 암호화하는 다른 벡터를 포함하는 Cas9 시스템에서 임의로 사용될 수 있다. 도 5는 이러한 방식으로 사용된 miRNA 회로의 예시적인 개략도를 제공한다. 도는 본원에 개시된 비-제한적인 예시적 개략도 및 다양한 효과기 성분의 배치를 제공한다.

- [0233] 예를 들면, 활성화(예컨대, VP64, RTA, P65), 억제(예컨대, KRAB), 및/또는 메틸화의 변경(예컨대, DNMT3A, DNMT3L)을 위해 사용된 효과기 성분은 재조합 발현 시스템, 예컨대, split-Cas9 시스템의 제1 발현 벡터 또는 제2 발현 벡터 위에 위치할 수 있다.
- [0234] TRE 및 tet-조절가능한 활성인자는 재조합 발현 시스템에서 2개의 상이한 벡터 속에 암호화되어야만 한다. 일부 구현예에서, tet-조절가능한 활성인자는 N-Cas9 암호화 벡터 내에서 암호화되며 TRE는 C-Cas9 암호화 벡터 내에서 암호화된다. 일부 구현예에서, 이는 역전될 수 있으며, 여기서 TRE는 N-Cas9 암호화 벡터 내에서 암호화되고 tet-조절가능한 성분은 C-Cas9 암호화 벡터 내에서 암호화된다.
- [0235] 프로모터 배치는 또한 개시된 구조내에서 고려된다. 일 국면에서, gRNA를 포함하는 작제물은 이의 상부를 암호화한, 프로모터, 임의로 U6 프로모터를 가질 수 있다. 유사하게, Cas9 또는 split Cas9의 1/2 2개 중 하나를 포함하는 작제물은 이의 상부를 암호화한 프로모터, 임의로 CMV 프로모터를 가질 수 있다.
- [0236] 캡시드 가공
- [0237] 본 개시내용의 국면은 하나 이상의 비천연 아미노산 및/또는 SpyTag 서열 또는 상응하는 KTag 서열의 첨가와 같은, 그러나 이에 한정되지 않는 양호한 특징을 부여하도록 가공된 바이러스 캡시드에 관한 것이다. 일부 구현예에서, 바이러스 캡시드는 AAV 캡시드 또는 렌티바이러스 캡시드이다.
- [0238] 다양한 부위가 캡시드 상에서 변형되어 하나 이상의 비천연 아미노산, SpyTag 서열, 또는 KTag 서열을 혼입할 수 있다. 일부 구현예에서, 표면 노출된 부위는 하나 이상의 비천연 아미노산, SpyTag 서열, 또는 KTag 서열의 혼입을 위한 적절한 부위로서 확인된다. AAV2 캡시드 내 이러한 부위의 비-제한적 예는 AAV2내 VP1의 447, 578, 87, 및 662번 잔기이다. 일부 구현예에서, 하나 이상의 비천연 아미노산, SpyTag 서열, 또는 KTag 서열의 혼입을 위한 부위는 AAV 기능을 절충하지 않는 것이다. AAV2와 관련하여, 특정의 표면 잔기는 조립체(assembly), 예컨대 509 내지 522번 및 561 내지 565번 잔기를 완벽하게 하는 것으로 알려져 있고, HSPG 결합, 예컨대, 586 내지 591, 484, 487, 및 K532를 부여한다. 138 및 139번 잔기는 표면 노출되어 있고 VP2의 N-말단에서 발견되며, 이는 AAV2 캡시드 속에 포함된다. 15개 이하의 아미노산이 139, 161, 459, 584, 및 587번 위치에서 삽입될 수 있다.
- [0239] 비천연 아미노산(또한 "UAA" 또는 "비 기본(non canonical) 아미노산"으로 지칭됨)은 천연적으로 발생할 수 있거나 화학적으로 합성될 수 있는 아미노산이지만 천연의 진핵 및 원핵세포 단백질 합성에 사용되는 22개의 기본 아미노산 중 하나가 아니다. 이의 비-제한적 예는 β -아미노산, 호모-아미노산, 프롤린 및 피루브산 유도체, 3-치환된 알라닌 유도체, 글리신 유도체, 환-치환된 페닐알라닌 및 타이로신 유도체, 선형의 코어 아미노산, 및 N-메틸 아미노산을 포함한다. 제한되지 않은 예시적인 비천연 아미노산은 Sigma Aldrich (sigmaaldrich.com/chemistry/chemistry-products.html?TablePage=16274965)를 통해 기술되어 있고 상업적으로 이용가능하다. 또한 제한되지 않는 예는 N-엡실론-((2-아지도에톡시)카보닐)-L-라이신, 피롤라이신, 및 다른 라이신 유도체를 포함한다.
- [0240] 일부 구현예에서, 비천연 아미노산은 아지드 또는 알킨을 포함한다. 비천연 아미노산에 포함된 기능성 그룹의 선택은 클릭 화학의 사용을 촉진하여 추가의 모이어티를 바이러스 캡시드에 가할 수 있다. 예를 들면, 아지드-알킨 첨가는 추가의 기능성 그룹을 아미노산 위에 혼입시키는 간단한 방식을 제공한다.
- [0241] 일부 구현예에서, 비천연 아미노산은 하전되거나 하전되지 않거나 극성 또는 비극성이다. 일부 구현예에서, 비천연 아미노산은 고도로 음으로 또는 양으로 하전되어 있다. 비천연 아미노산의 전하 및 극성의 선택은 바이러스 캡시드와 함께 취해질 다음 단계에 의존한다. 예를 들면, 바이러스 캡시드는 리포펙타민으로 캡슐(encapsulating)화될 것이며, 고도의 음으로 하전된 비천연 아미노산이 바람직할 수 있다.
- [0242] 단백질내로 비천연 아미노산을 혼입하는 방법은 당해 분야에 알려져 있고 재지정된 정지 코돈을 사용하는, 예컨대, 앰버 억제(amber suppression)를 하는 직교의 해독 시스템의 사용을 포함한다. 이러한 첨가를 수반하기 위한 직교의 tRNA 신테타제의 비-제한적 예는 MbPy1RS, MmPy1RS, 및 AcKRS를 포함하나, 이에 한정되지 않는다. 비천연 아미노산의 혼입은 또한 추가의 제제의 사용에 의해 향상될 수 있다. 비-제한적 예는 이의 예시적인 서열 이 하기 제공된 eTF1이다:

[0243] eTF1(정상)-E55D(짧은 이탤릭체, 변형된 서열)

MADDPSAADRNVEIWKIKKLIKSLAARGNGTSMISLIIPPKDQISRVAK
MLADDFGTASNISRVNRLSVLGAITSVQRLKLYNKVPPNGLVVYCGTI
VTEEGKEKKVNDIFEPFKPINTSLYLCDNKFHTEALTALLSDDSKFGFIV
IDGSGALFGTLQGNTRVFLHKFTVDLPKKHGRGGQSALRFARLRMEKRHN
YVRKVAETAVALFISGDKVNVAGLVLAGSADFKTELSQSDMFDORLQSKV
LKLVDISYGGENGFNQAIELSTEVLNVKFIQEKKLIGRYFDEISQDTGK
YCFGVEDTLKALEMGAVEILIVYENLDIMRYVLHCQGTREEEKILYLTP EQ
EKDKSHFTDKETGQEHLEIEMPLEWFANNYKFKGATLEIVTDKSSQEGS
QFVKGFGGIGILRYRVDFQGMFYQGGDEFFDLDDY

[0244]

[0245] 유사한 방법을 사용하여 바이러스 캡시드 상에 SpyTag 또는 KTag를 혼입할 수 있다. SpyTag는 상응하는 KTag 서열 ATHIKFSKRD와 쌍을 이루는 공지된 서열 AHIVMVDAYKPTK이며 SpyLigas - AddGene을 통해 이용가능한 시판되는 상업적으로 이용가능한 효소이고 GenBank 참고 번호 KJ401122와 관련된 SpyLigase의 존재하에서 및 일부 예에서는 자발적으로 연결된다.

[0246] AAV2 및 AAV-DJ로부터의 하기 AAV 서열은 비천연 아미노산, SpyTag, 또는 KTag 서열이 혼입될 수 있는 예시적인 위치를 제공한다.

[0247] AAV2 VP1(정상)(R447(짧은 글씨); S578(굵고 밑줄쳐진 글씨); N587(짧은 이탤릭체); S662(짧은 이중선 글씨))

MAADGYLPDWLEDTLSEGIQVWKLKPGPPPKPAERHKDDSRGLVLPGY
KYLGPFGNLDKGEVNEADAAALEHDKAYDRQLDSDGNPYLYKNHADADEF
QERLKEDTSFGGNLGRAVFAKRRVLEPLGLVEEPVKTAPGKKRPVEHSP
VEPDSSSGTGKAGQPPARKRLNFGQTGDADSVDPDQPLGQPPAAPSGLGT
NTMATGSGAPMADNNEGADGVGNSSGNWCHDSTWMDRVTITSTRTWALP
TYNNHLYKQISSQSGASNDNHYFGYSTPWGYFDNRFHCHFSPRDWQRLL
NNNWGFRPKRLNFKLFNIQVKEVTQNDGTTIANNLSTVQVFTDSEYQL
PYVLGSAHQGCLPPFPADVFMVPOYGYLTLNNGSQAVGRSSFYCLEYFPS
QMLRTGNNTFYSYTFEDVPHSSYAHQSGLDRLMNPIDQYLYLSTRNT
PSGTTTQSRQLQFSQAGASDIRDQSRNWLPGPCYRQQRVSKTSADNNSEY
SWTGATKYHLNGRDSLVPNPGPAMASHKDDDEEKFFPQSGVLFKQKQSEKT
NVDIEKVMITDEEIRTTNPVATEQYGSVSTNLQRGNRQAATADVNTQGV
LPGMVWQDRDYYLQGPWAKIPHTDGHFHPSPMLGGFGLKHPPIQILIKN
TPVPANPSTTFSAAKFASFITQYSTGQVSVEIEWELQKENSKRWNPEIQY
TSNYNKSVMVDFVDTNGVYSEPRPIGTRYLTRNL

[0248]

[0249] AAV-DJ VP1(정상)(N589(굵고 밑줄쳐진))

MAADGYLPDWLEDTLSEGIQVWKLKPGPPPKPAERHKDDSRGLVLPGYKYLGPFGN
LDKGEVNEADAAALEHDKAYDRQLDSDGNPYLYKNHADADEFQERLKEDTSFGGNL
RAVFAKRRLEPLGLVEEAAKTAPGKKRPVEHSPVEPDSSSGTGKAGQPPARKRLN
GQTGDADSVDPDQPIGPPAAPSGVGLTMAAGGAPMADNNEGADGVGNSSGNWCH
DSTWMDRVTITSTRTWALPTYNNHLYKQISNSTSGSSNDNAYFGYSTPWGYFDNRF
HCHFSPRDWQRLLNNNWGFRPKRLSFKLFNIQVKEVTQNEGKTIANNLTSTIQVFTDS
EYQLPYVLGSAHQGCLPPFPADVFMVPOYGYLTLNNGSQAVGRSSFYCLEYFPSQMLRT
GNNTFYTFEDVPHSSYAHQSGLDRLMNPIDQYLYLSTRQTTGGTNTQTLGFSQ
GGPNTMANQAKNWLPGPCYRQQRVSKTSADNNSEYSWGATKYHLNGRDSLVPNPG
PAMASHKDDDEEKFFPQSGVLFKQKQSEKTNVDIEKVMITDEEIRTTNPVATEQYGSV
TNLQRGNRQAATADVNTQGVLPGMVWQDRDYYLQGPWAKIPHTDGHFHPSPMLGGF
GLKHPPIQILIKNTPVPADPPTTFNQSKLNSFITQYSTGQVSVEIEWELQKENSKRWNPEI
QYTSNYKSTSVDFAVNTEGVYSEPRPIGTRYLTRNL

[0250]

[0251] 달리 제공하지 않는 한, AAV2 또는 AAV-DJ VP1 서열에서 아미노산 위치에 대한 참고는 상기 개시된 서열내 잔기의 위치를 기반으로 한다. 또한, 각각의 AAV의 VP1이 지칭되는 경우, 이의 생물학적 등가물도 포함하는 것으로 의도된다.

[0252] 일부 구현예에서, 캡시드 내로 혼입된 하나 이상의 비천연 아미노산, SpyTag, 또는 KTag는 캡시드의 표면에 추가의 모이어티 또는 "슈도타입(pseudotype)"을 도입하는데 사용된다. 모이어티는 펩타이드, 앵타머, 올리고뉴클레오타이드, 아포보디, DARPin, 쿠니츠 도메인(Kunitz domain), 피노머(fynomer), 비사이클릭 펩타이드, 안티칼린, 및 애드벡틴을 포함하나 이에 한정되지 않는다. 다양한 모이어티가 바이러스의 분리, 바이러스와 다른 바이러스의 연결, 및/또는 특수한 표적 세포, 기관 또는 조직으로 바이러스의 호밍(homing)을 허용하는 것을 포함하는, 다수의 기능에 유용할 수 있다.

[0253] 이러한 슈도타입은 클릭 화학을 통해 달성될 수 있다. SpyTag가 캡시드 위에 혼입되는 경우, 클릭 화학은 KTag의 슈도타입될 모이어티에 대한 접합을 포함한다. SpyTag의 KTag(예컨대, SpyLigase의 도입을 통해)에 대한

연결을 촉진하는 반응을 채택함으로써, 모이어티를 캡시드의 표면에 가한다. 이러한 슈도타이핑을 위한 서열의 비-제한적 예는 통증 관리에서 뉴우런 호밍을 위한 2개의 제제인, 물질-P 및 RVG에 접합된 KTag이다:

[0254]

KTag-물질P: **ATHIKFSKRD GSGSGS RPKPQQFFGLM**

[0255]

물질P-KTag: **RPKPQQFFGLM GSGSGS ATHIKFSKRD**

RVG-Ktag: YTIWMPENPRPGTPCDIFTNSRGKRASNG GKG GG GSGSGS ATHIKFSKRD

KTag-RVG: ATHIKFSKRD GSGSGS GKG GG YTIWMPENPRPGTPCDIFTNSRGKRASNG

[0256]

또는 이의 각각의 생물학적 등가물.

[0257]

상기 예시적인 구현예는 캡시드 상에서 SpyTag 및 모이어티 상에서 KTag의 용도를 나타내지만, 이의 역도 또한 달성될 수 있지만 KTag가 캡시드내로 혼입되고 SpyTag가 모이어티에 접합하는 것으로 인식하여야 한다. 비천연 아미노산과 관련하여, 임의로 구리에 의해 촉매된 아지드-알킨 반응을 사용하여 상응하는 기능성 그룹(예컨대, 비천연 아미노산은 아지드를 포함하고 모이어티는 알킨을 포함하거나, 이의 역도 가능하다)을 지닌 모이어티를 가할 수 있다.

[0258]

[0259]

일부 구현예에서, 가공된 캡시드를 사용하여 관절 전달용 바이러스에 연결시킬 수 있다. 이러한 연결은 본원에 개시된 재조합 발현 시스템의 전달에 특히 유용하며, 여기서 Cas9는 즉, 2개의 벡터 내에서 split-Cas9로서 암호화된다. 예를 들면, 하나의 캡시드는 SpyTag 및 다른 KTag를 포함할 수 있으며; 따라서, 바이러스는 SpyTag의 KTag에 대한 연결을 촉매함으로써 연결될 수 있다. 유사하게, 아지드-알킨 반응을 사용하여 바이러스의 연결을 촉진할 수 있으며, 여기서 하나는 아지드 함유 비천연 아미노산을 포함하고 다른 것은 알킨 함유 비천연 아미노산을 포함한다. 연결된 바이러스의 추가의 구현예를 하나 이상의 슈도타이핑된 모이어티를 사용하여 개발할 수 있으며, 여기서 2개의 바이러스는 서로 하이브리드화하는 모이어티를 발현하거나 자발적으로 또는 촉매작용을 통해 연결될 수 있다.

[0260]

추가적인 구현예에서, 캡시드는 면역 차폐(immune shielding)를 위해 가공될 수 있다. AAV와 같은 바이러스 캡시드에 대한 광범위한 노출은 대상체가 많은 천연 바이러스 혈청형에 대해 중화 항체를 지니도록 한다. 일부 구현예에서, 캡시드는 결실 또는 셔플링(shuffling)을 통해 변형되어 면역계를 회피할 수 있으며; 일부 구현예에서, 캡시드는 엑소솜과 관련될 수 있다. 일부 구현예에서, 특이적인 시약이 면역 차폐를 위해 혼입되거나 캡시드를 코팅하는데 사용된다. 예를 들면, 폴리(락틱-코-글리콜산), PEG, VSVG 코팅, 및/또는 지질/아민(예컨대, 리포펙타민) 코팅과 같은 중합체의 첨가를 사용할 수 있다.

[0261]

면역 차폐의 비-제한적인 예는 리포펙타민 코팅이다. 예를 들면, 알킨-올리고뉴클레오타이드를 캡시드를 포함하는 비천연 아미노산에 연결시킬 수 있다. 이후에, 변형된 바이러스는 리포펙타민으로 세척되며, 이는 최종적으로 코팅을 형성한다.

[0262]

추가적인 변형이 특수한 조직을 표적화할 목적으로 캡시드에 대해 이루어질 수 있다. 상기 나타낸 바와 같이, "호밍" 모이어티를 슈도타이핑에 사용하여 특수한 표적 세포, 기관, 또는 조직에 대한 캡시드의 국재화를 보증할 수 있다.

[0263]

추가적인 변형을 당해 분야에 공지된 캡시드에 대해 이루어져 이것이 다음에 기술된 것과 같은, 그러나 이에 한정되지 않는 특수한 방법 국면에 적합하도록 할 수 있음이 인식된다: 미국 특허 7,867,484; 7,892,809; 9,012,224; 8,632,764; 9,409,953; 9,402,921; 9,186,419; 8,889,641; 7,790,154; 7,465,583; 7,923,436; 7,301,898; 7,172,893; 7,071,172; 8,784,799; 7,235,235; 6,541,010; 6,531,135; 6,531,235; 5,792,462; 6,982,082; 6,008,035; 5,792,462; 9,617,561; 9,593,346; 9,587,250; 9,567,607; 9,493,788; 9,382,551; 9,359,618; 9,315,825; 9,217,159; 9,206,238; 9,198,984; 9,163,260; 9,133,483; 8,999,678; 8,962,332; 8,962,233; 8,940,290; 8,906,675; 8,846,031; 8,834,863; 8,685,387; 미국 특허 공보 2016/120960; 2017/0096646; 2017/0081392; 2017/0051259; 2017/0043035; 2017/0028082; 2017/0021037; 2017/0000904; 2016/0271192; 2016/0244783; 2916/0102295; 2016/0097040; 2016/0083748; 2016/0083749; 2016/0051603; 2016/0040137; 2016/0000887; 2015/0352203; 2015/0315612; 2015/0230430; 2015/0159173; 2014/0271550, 및 이러한 특허 및 특허 공보와 관련된 다른 패밀리 구성원 또는 이의 양수인 또는 발명자.

[0264]

조합 및 방법

[0265]

본원에 개시된 국면은 하나 이상의 비천연 아미노산 및/또는 SpyTag 서열 또는 상응하는 KTag 서열의 단독 또는

예컨대, 조성물의 형태의 다른 것과의 조합의 첨가와 같은, 그러나 이에 한정되지 않는 선호되는 특성을 부여하도록 가공된 재조합 발현 시스템(split-Cas9) 및 바이러스 캡시드의 용도에 관한 것이다.

- [0266] 예를 들면, 본원에 개시된 재조합 발현 시스템에 포함된 2개의 벡터를 하나 이상의 비천연 아미노산, SpyTag 서열 또는 KTag 서열을 혼입하도록 가공된 바이러스 캡시드 속에 포장할 수 있다. 대안적으로는, 하나 이상의 벡터를 변형되지 않은 바이러스 캡시드 속에 포장할 수 있다.
- [0267] 조합은 상기 나타난 바와 같은 장점, 특히 split-Cas9 시스템의 2개의 부위를 벡터 둘 다의 전달을 보증하도록 연결하는 능력을 제공한다. 또한, 바이러스 캡시드가 슈도타이핑된 구현예에서, 조직 특이적인 전달을 호밍 모이어티의 사용을 통해 달성할 수 있다.
- [0268] 일부 구현예에서, 재조합 발현 시스템, 본원에 개시된 바와 같이 가공된 바이러스 캡시드, 및/또는 split-Cas9 시스템을 포함하는 2개의 벡터가 본원에 개시된 바와 같이 가공된 2개의 바이러스 캡시드 속에 포함된 재조합 발현 시스템을 대상체에게 전달할 수 있다. 일부 구현예에서, 경로 및 용량은 치료되는 대상체 또는 상태를 기반으로 하여 측정될 수 있다.
- [0269] 본원에는 통증 관리, 간 질환, HSC 치료요법, HIV, 암 면역치료요법, 혈액 질환, 근육 디스트로피, 자궁내 태아 표적화, 사이토크롬 p450 기반 질환, 재프로그래밍 대식구, 모기 퇴치, 알츠하이머, 및 갑상선 호르몬 생산을 포함하나 이에 한정되지 않는 구체적인 용도로 조정된 gRNA가 본원에 개시되어 있다. 바이러스 캡시드의 슈도타이핑 뿐만 아니라 재조합체 발현 시스템에서 사용된 효과기 성분은 이들 각각의 용도를 위해 최적화될 수 있다.
- [0270] 예를 들면, 통증 관리의 경우, 상기 본원에 개시된 호밍 펩타이드는 바이러스 캡시드가 뉴우런을 표적화함으로써, 조직 특이성을 부여하는 것을 허용한다. 본원에 개시된 이러한 조직 특이성을 운반하기 위한 추가의 국면은 뉴우런에 대해 특이적인 miRNA 회로의 사용 및/또는 재조합 발현 시스템에서 구체적으로 개시된 gRNA의 사용을 포함하나 이에 한정되지 않는다.
- [0271] 암 면역치료요법에서 다른 예는 시그널링 경로(signaling pathway)의 조절이다. 소수의 경로만이 신체 전체에서 유전자 발현을 조절하므로, 본 출원에서 조직 특이성은 중요하다. miRNA 회로, 조직 특이적인 촉진, 및 바이러스 캡시드 속에서 표적 암에 대해 특이적인 호밍 펩타이드의 혼입은 이러한 치료가 목적인 표적에서 유전자에 유일하게 영향을 미칠 수 있음을 보증할 수 있다.
- [0272] HSC 치료요법 및 HSC를 포함하는 혈액 질환과 관련하여, 본 출원인은 전달 경로가 중요할 수 있다고 믿고, 따라서 개시된 재조합 발현 시스템 또는 조성물의 골수(여기서 조혈 줄기 세포(HSC)의 저장기) 또는 흉선(여기서 T-세포가 성숙한다) 내로의 직접적인 주사와 같은, 그러나 이에 한정되지 않는 반응계내(*in situ*) 또는 생체내 도입으로 바이러스의 전달을 제안한다. 유사한 골수 전달은 예컨대, PDCD-1 표적화 gRNA를 사용하여, 면역 장애를 위한 반응계내 또는 생체내 T-세포 편집 및/또는 HSC 편집 및/또는 암 치료에 사용할 수 있다. HSC 및/또는 T-세포는 조직 특이적인 gRNA 또는 다른 효과기 성분의 선택을 기반으로 구체적으로 편집함으로써; 면역 질환을 치료하고/하거나 예방할 수 있다. 반응계내 또는 생체내 시도는 생체의외(*ex vivo*) 변형 및 이식 세포(예컨대, HSC 및 T 세포)에 크게 의지하는 현재의 치료보다 더 효과적인 시도이며 HSC 이식 또는 T-세포 이식의 높은 가능성과 관련된다. 또한, 반응계내 또는 생체내 전달은 이러한 세포 치료요법의 비용을 감소시키는 큰 잠재능을 갖는다.
- [0273] 대안적으로, HSC 및/또는 T-세포와 관련된 이러한 및 암 관련된 구현예에서, 환자 HSC 및/또는 T-세포는 생체내에서 변형되어 환자(예컨대, 골수내로의 직접 주사를 통해)에게 전달될 수 있다. 변형된 세포는 이후에 생체내에서 확장될 수 있다. 일부 구현예에서, 환자에게 질환에 관여하는 세포의 기존 집단을 제거한 후 이러한 변형된 세포를 투여한다.
- [0274] 갑상선 관련된 구현예에서, 일시적으로 조절된 dCas9 시스템 및 임의로 갑상선으로의 호밍을 위해 변형된 바이러스 캡시드를 이용할 수 있다.
- [0275] 추가의 방법 국면은 재조합 발현 시스템의 전달을 포함할 수 있고/있거나 바이러스 캡시드가 하이드로겔을 사용할 수 있다. 하이드로겔은 생체내에서 약물-전달 생물물질로서 사용되어 왔다. 특정 조건에서 약물의 포획 및 방출의 최적화가 광범위하게 연구되었다. 하이드로겔 방출 특성을 조율함으로써, 재조합 발현 시스템 및/또는 바이러스 캡시드의 특이적인 전달을 별개의 pH 수준, 온도, 또는 생리학적 조건에 따라 조절할 수 있다. 예를 들면, 재조합 발현 시스템 및/또는 바이러스 캡시드를 예를 들면, 재조합 발현 시스템 및/또는 바이러스 캡시드와 보다 낮은 pH 수준에서 접촉하여 방출하도록 함으로써 염증이 있는 부위에 전달할 수 있다. 또한 및 이론에 얽매이지 않고, 최적화된 하이드로겔은 재조합 발현 시스템 및/또는 바이러스 캡시드를 제 위치에 유지하여 비-

특이적인 표적화를 방지하고 바람직하지 않은 부작용으로부터 더 많은 보호를 대상체에게 제공한다. 이러한 전달 시스템은 재조합 발현 시스템 및/또는 바이러스 캡시드의 특이성을 증가시킬 수 있다.

[0276] split-Cas9 시스템을 사용하는 방법에서, 동일한 역가의 Cas9의 1/2 둘 다는 기능성 Cas9가 전달시 생성되도록 보증하기 위해 중요하다. 이는 2개의 벡터를 포함하고/하거나 각각의 벡터에서 유일한 영역을 표적화하는데 qPCR을 이용하여 역가 대조군(예컨대, ATCC-VR-1616)에 대해 각각의 벡터의 역가를 측정하는 바이러스 캡시드의 쌍화에 의해 보증될 수 있다.

[0277] 방법 국면은 또한 본원에서 개시된 바이러스 캡시드를 사용하여 생체적합성을 시험하기 위해 고려된다. 물질의 생체적합성을 시험하기 위한 하나의 일반적인 방법은 동물 모델을 사용하고 조직학 및 면역조직화학을 수행하여 각각의 조직 속에 존재하는 세포를 특성화하는 것이다. 비용이 많이 드는 것 외에, 이는 또한 시간 및 노동 집약적이고, 정량화하기 어려울 수 있다. 하나의 가능한 대안은 TK-GFP를 목적한 부위로 패키징하는 바이러스 캡시드를 도입하는 것일 수 있다. 이후에 TK-GFP AAV를 식세포작용하는 대식구는 작열(glow)하여 리포터 유전자를 발현할 수 있다. B 및 T 세포에서 세포 표면 수용체의 장점을 이용하는 것은 또한 TK-GFP AAV에 의한 형질도입을 허용하여 생체내에서 림프구를 정량화하도록 할 수 있다. 대식세포 식세포작용을 촉진하거나 대식세포 특이적인 세포 수용체를 조작하는 것은 친밀하고/하거나 획득된 면역 반응의 정량화를 허용할 수 있다. 궁극적으로, 생물물질 시험은 보다 효율적이고 접근가능하게 될 것이다.

[0278] 본원에 용도에 적합한 용량은 임의의 적합한 경로, 예컨대, 정맥내, 경피, 비강내, 경구, 점막, 또는 다른 전달 방법을 통해, 및/또는 단일 또는 다수의 용량을 통해 전달될 수 있다. 실제 투여량은 사용된 재조합 발현 시스템(예컨대, AAV 또는 렌티바이러스), 표적 세포, 기관, 또는 조직, 대상체, 및 추구되는 효과의 정도에 따라 변할 수 있다. 조직, 기관, 및/또는 환자의 체격 및 체중은 또한 복용량에 영향을 미칠 수 있다. 용량은 또한 담체를 포함하나, 이에 한정되지 않는 추가의 제제를 포함할 수 있다. 적합한 담체의 비-제한적인 예는 당해 분야에 알려져 있다: 예를 들면, 물, 염수, 에탄올, 글리세롤, 락토즈, 슈크로즈, 텍스트란, 아가(agar), 펙틴, 식물-기원한 오일, 포스페이트-완충된 염수, 및/또는 희석제. 예를 들면, WO 2017/070605의 [00533] 단락에 개시된 추가의 물질이 본원에 개시된 조성물과 함께 사용하기에 적절할 수 있다. WO 2017/070605의 [00534] 내지 [00537] 단락은 또한 본원에 사용될 수 있는 CRISPR-Cas 시스템에 대한 복용량 편의의 비-제한적인 예를 제공한다. 일반적으로, 복용량 고려는 당해 분야의 기술자에 의해 잘 이해된다.

[0279] **실시예**

[0280] 다음의 실시예는 개시내용을 효과적으로 수행하는데 있어서 다양한 예로 사용될 수 있는 비제한적이고 설명적인 과정이다. 추가로, 하기 본원에 개시된 모든 참고 문헌은 이의 전문이 참고로 포함된다.

[0281] 실시예 1 - 예시적인 모듈러 AAV 시스템의 생성

[0282] **벡터 설계 및 작제**

[0283] 요약하면, split-Cas9 mAAV 벡터를 상응하는 유전자 블록(Integrated DNA Technologies)의 통상적으로 합성된 rAAV2 벡터 골격 내로 순차적 조립에 의해 작제하였다. UAA 실험을 위해, 4개의 유전자 블록(block)을 표면 잔기 R447, S578, N587 및 S662를 암호화하는 뉴클레오타이드 대신에 삽입된 'TAG'를 사용하여 합성하고 Gibson 조립을 사용하여 pAAV-RC2 벡터(Cell Biolabs) 내로 삽입하였다. ETF1-E55D의 경우, 단백질 서열을 암호화하는 유전자 블록을 합성하여 Gibson 조립을 통해 CAG 프로모터의 하부에 삽입하였다.

[0284] **포유동물 세포 배양**

[0285] HEK293T 세포를 10% FBS 및 1% 항생제-항진균제(ThermoFisher Scientific)가 보충된 둘베코 변형 이글 배지(Dulbecco's Modified Eagle Medium)(10%) 속에서 항온처리기 속에서 37°C 및 5% CO₂ 대기 하에서 성장시키고, AAV 형질도입을 위해 24-웰 플레이트(24-well plate) 속에 플레이트링하였다. pAAV 유도성-Cas9 벡터로 형질감염시킨 293T 세포에 200 ug/ml의 독시사이클린을 보충하였다. CD34(CD34+ 세포)를 발현하는 조혈 줄기 세포를 StemSpan^T CD34+ 확장 보충물(10X)이 보충된 혈청 유리된 StemSpan^T SFEM II(모두 StemCell Technologies로부터 구입) 속에서 성장시켰다. CD34+ 세포를 AAV 형질도입을 위해 96-웰 플레이트에 플레이트링하였다.

[0286] **AAV 바이러스의 생산**

[0287] AAV8 바이러스를 모든 생체내 연구에 이용하고, AAVDJ를 HEK293T 세포내에서 모든 시험관내 연구를 위해 이용하며, AAV6을 CD34+ 세포내에서 생체외 연구를 위해 이용하고, AAV2를 UAA 혼입 연구를 위해 이용하였다.

- [0288] 대량 생산: 바이러스를 Salk Institute of Biological Studies (캘리포니아주 라 줄라 소재)에서, 또는 연구실에서 유전자 전달, 표적화 및 치료요법(gT3) 코어에 의해 제조하였다. 요약하면, AAV2/8, AAV2/2, AAV2/6, AAV2/DJ 바이러스 입자를 PEI를 사용하여 15cm 플레이트 속에서 80 내지 90% 합치(confluency)에서 7.5 ug의 pXR-캡시드(pXR-8, pXR-2, pXR-6, pXR-DJ), 7.5 ug의 재조합체 전달 벡터, 및 22.5 ug의 pAd5 헬퍼 벡터로 형질감염된 HEK293T 세포를 사용하여 생산하였다. 바이러스를 72시간 후에 수거하고 요덕사놀 구배를 사용하여 정제하였다. 바이러스를 100kDA 여과기(Millipore)를 사용하여 ~1 mL의 최종 용적으로 농축시키고 ITR 영역에 대해 특이적인 프라이머를 사용한 qPCR에 의해 표준물(ATCC VR-1616)에 대해 정량화하였다.
- [0289] AAV-ITR-F: 5'-CGGCCTCAGTGAGCGA-3' 및
- [0290] AAV-ITR-R: 5'-GGAACCCCTAGTGATGGAGTT-3'
- [0291] UAA 혼입: 수거까지 형질감염 2시간 전부터, 293T 세포를 0.4mM 라이신(DMEM 속에 일반적으로 존재하는 0.8mM 라이신과 대체되는 것으로서)을 함유하고 10% FBS 및 2mM N-엡실론-((2-아지도에톡시)카보닐)-L-라이신으로 보충시킨 DMEM 속에서 성장시켰다. 피롤리실-tRNA 및 tRNA 신테타제를 함유하는 플라스미드 pAcBac1.tR4-MbPyl(Peter Schultz로부터의 선물, Addgene #50832)을 캡시드 벡터 pAAV-RC2(및 이의 돌연변이체), 돌연변이체 전달 벡터, 및 pAd5 헬퍼 벡터와 5:1 비의 캡시드 벡터와 함께 293T 세포 내로 동시-형질감염시켰다. 상기와 동일한 프로토콜을 바이러스의 수거, 정제 및 정량화를 위해 수행하였다. 기능적 활성을 추가로 정량화하기 위해, UAA AAV의 유동 세포분석기 분석(flow cytometry analysis)을 형질도입 48시간 후에 수행하고 20,000개의 세포를 FACScan 유동 세포분석기 및 Cell Quest 소프트웨어(둘 다 Becton Dickinson)를 사용하여 분석하였다.
- [0292] 소-규모 생산: 소-규모 AAV 프렙(preps)을 0.5 ug의 pXR-캡시드, 0.5 ug의 재조합체 전달 벡터, 및 1.5 ug의 pAd5 헬퍼 벡터를 PEI를 사용하여 동시-형질감염시킨, HEK293T 세포를 함유하는 6-웰 플레이트를 사용하여 제조하였다. 세포 및 상층액을 72시간 후에 수거하고, 조 추출물(crude extract)을 사용하여 세포를 형질도입하였다.
- [0293] 동물 실험
- [0294] AAV 주사: 모든 동물 과정을 샌 디에고 소재의 캘리포니아 대학의 시설 동물 보호 및 사용 위원회 (Institutional Animal Care and Use Committee: IACUC)에 의해 승인된 프로토콜에 따라 수행하였다. 모든 마우스는 Jackson labs.로부터 구입하였다. AAV 주사는 꼬리-정맥 주사를 통해 성체 C57BL/6J 마우스(10주)내에서 또는 IP 주사를 통한 신생아(4주)에서, 0.5E+12-1E+12를 사용하여 수행하였다. 주사 4주 후, 마우스를 CO₂에 의해 인도적으로 희생시켰다. 조직을 수거하고 RNeasy lysis buffer (Qiagen) 용액(ThermoFisher Scientific) 속에서 동결시켰다.
- [0295] 독시사이클린 투여: pAAV 유도성-Cas9 벡터로 형질도입된 마우스에게 0.4 mL의 1N HCl이 들어있는 10 mL의 0.9% NaCl 속의 200 mg의 독시사이클린의 IP 주사를 4주 동안 주당 3회 제공하였다.
- [0296] 조직학: 마우스를 CO₂에 의해 인도적으로 희생시켰다. 간을 OCT 화합물(VWR)을 함유하는 주형(mold) 속에서 동결시키고 드라이 아이스/2-메틸 부탄 슬러리 속에서 동결시켰다. 조직학을 무어스 암 센터 조직학 및 영상 코어 시설(Moores Cancer Center Histology and Imaging Core Facility)(캘리포니아주 라 줄라 소재)에 의해 수행하였다. 간 단면을 병리학을 위해 헤마톡실린 및 에오신(H&E)을 사용하여, 및 항-CD81(BD Biosciences, No. 562240)을 사용하여 염색하였다.
- [0297] 게놈 DNA 추출 및 NGS 제조
- [0298] 세포 및 조직으로부터의 gDNA를 DNeasy Blood and Tissue Kit(Qiagen)를 사용하여 제조업자의 프로토콜에 따라 추출하였다. 차세대 서열분석 라이브러리를 다음과 같이 제조하였다. 요약하면, 4 내지 10 ug의 투입 gDNA를 목적인 부위 주변에 150 bp를 증폭시키는 프라이머(표 2b)로 KAPA HiFi HotStart PCR Mix(Kapa Biosystems)를 사용하여 PCR로 증폭시켰다. PCR 생성물을 겔 정제(Qiagen 겔 추출 키트)하고, 추가로 PCR 정제하여(Qiagen PCR 정제 키트) 부산물을 제거하였다. 라이브러리 작제물을 Illumina kit(NEB)용 NEBNext Multiplex Oligos를 사용하여 수행하였다. 10 내지 25 ng의 투입 DNA를 색인 프라이머로 증폭시켰다. 이후에, 샘플을 정제하고 qPCR 라이브러리 정제 키트(Kapa Biosystems, KK4824)를 사용하여 정량화하였다. 이후에, 샘플을 혼주(pooling)시키고 Illumina Miseq(150 bp 쌍을 이룬 말단 런(run) 또는 150 단일-말단 런) 위에서 4nM 농도로 부하(loading)하였다. 데이터 분석을 CRISPR 게놈 분석기(Genome Analyzer) 44를 사용하여 수행하였다.

- [0299] 유전자 발현 분석 및 qRT-PCR
- [0300] 세포로부터의 RNA를 RNeasy 키트(Qiagen)를 사용하여, 및 RNeasy 플러스 유니버설 키트(Plus Universal Kit)(Qiagen)를 사용하여 조직으로부터 추출하였다. 1 ug의 RNA를 Protoscript II 리버스 트랜스크립타제 키트(Reverse Transcriptase Kit)(NEB)를 사용하여 수행하였다. 실시간 PCR(qPCR) 반응을 KAPA SYBR Fast qpcr 키트(Kapa Biosystems)를 사용하여, 유전자 특이적인 프라이머(표 2a)로 수행하였다. 데이터를 GAPDH 또는 B-액틴으로 표준화하였다.
- [0301] AAV 슈도타이핑
- [0302] Alexa 594 DIBO 알킨 테더링(tethering): AAV2 야생형 및 AAV2-S578UAA를 Alexa 594 DIBO 알킨과 함께 TBS(둘 다 ThermoFisher Scientific) 속에서 1시간 동안 실온에서 항온처리하였다. 과도한 표지를 PBS로 세척 제거하였다. 바이러스 입자를 293T 세포에 가하고 세포를 형질감염 2시간 후 영상화하였다.
- [0303] 올리고뉴클레오타이드 테더링 및 DNA 배열: 올리고 A' 및 B'(5uM)를 스트렙타비딘 작용화 배열(streptavidin functionalize array)(ArrayIt: SMSFM48)에 스폿팅(spotting)하고 실온에서 3분 45초 동안 항온처리하였다. 한편, 올리고 A를 AAV2-N587UAA_mCherry에 클릭 화학의 공정(Click-iT - ThermoFisher Scientific, C10276)을 통해 연결한 후 PBS로 세척하였다. 다음에, 배열을 PBS로 세척하고 변형된 AAV2-N587UAA_mCherry를 각각의 웰(well)에 가하고, 실온에서 30분 동안 항온처리한 후 PBS로 세척하였다. 최종적으로, 293T 세포를 각각의 웰에 가하였다. 세포를 mCherry 발현에 대해 형질감염 48시간 후 영상화하였다.
- [0304] 논의
- [0305] AAV는 이들의 온화한 면역 반응, 장기간 전이유전자 발현, 광범위한 세포를 감염시키는 능력, 및 양호한 안전성 프로파일로 인하여 유전자 전달에 매우 바람직하므로 예시적인 플랫폼을 아데노-관련 바이러스(AAV)를 사용하여 구축한다. 그러나, AAV는 제한된 패키징 능력(~4.7 kb)을 가지므로, 대형 Cas9-유사 효과기 단백질 및 이의 융합체, 및 또한 유효 유전자 및 안내-RNA 발현에 필수적인 구성성분을 혼입시키기 어렵게 한다. 따라서, 본 출원인은 split-Cas9 시스템을 확장하여 이러한 한계를 극복하였다. 본 출원인의 전달 양식에서 스타필로코쿠스 피오게네스 Cas9(SpCas9) 단백질은 엔. 폰크티포르메로부터 원래 기원한 split-인테인을 사용함으로써 1/2로 분할되며, 이에 의해 각각의 Cas9 1/2은 이의 상응하는 split-인테인 모이어티로 융합되며 동시-발현시 완전한 Cas9 단백질이 재구성된다. 이러한 전달 양식은 2개의 rAAV를 이용하며 상응하는 벡터를 적절히 설계함으로써 출원인은 수득되는 잔사 패키징 능력을 확장시켜 충분한 범위의 CRISPR-Cas 게놈 가공 기능이 가능하도록 하였다(도 16).
- [0306] 본 출원인은 시험관내 및 생체내 시나리오에서 광범위한 세포형 및 게놈 유전자자리에 걸쳐 표적화된 게놈 편집을 확인하였고(도 16a, 16b), 특히 또한, 사람 CD34+ 조혈 줄기 세포에서 풍부한 AAV6 매개된 편집을 입증하였다. 히트(hit) 및 런 시도는 게놈 편집에 충분하며 실제로 장기간 뉴클레아제 발현에 걸쳐 바람직하므로, 출원인은 다음에 합성 회로의 혼입을 가공하여 CRISPR-Cas 편집 활성의 소-분자 조절을 가능하도록 하였다. 여기서 하나의 rAAV 작제물을 설계하여 C-인테인-C-Cas9 융합체의 상부에 테트라사이클린 반응 성분(TRE)을 지닌 최소 CMV 프로모터를 지니도록 하고, 제2 rAAV 작제물에서 완전한 프로모터를 사용하여 N-인테인-N-Cas9 융합체 및 tet-조절가능한-활성인자(tetA)의 발현을 구동시켰다. 독시사이클린의 존재하에서, tetA는 TRE 부위에 결합하여 C-Cas9의 유도성 발현을 허용함으로써 유전자 편집의 일시적인 조절을 허용한다. 출원인은 시험관내 및 생체내 시나리오 둘 다에서 이러한 회로의 기능화를 입증하였다(도 16c). 이와 함께, 상기 시스템은 견고한 CRISPR-Cas9 기반 게놈 편집을 가능하도록 하며, tet 조절인자의 커플링은 AAV로부터 또한 지속적인 유전자 발현의 용이한 조절을 가능하도록 한다.
- [0307] 본 출원인은 다음에 죽은 split-Cas9 단백질을 이용하여 KRAB 도메인의 융합을 통한 표적화된 게놈 억제, 및 VP64 cum rTA 도메인을 통한 표적화된 게놈 활성화를 가공하였다(도 16d). 시험관내 실험은 AAVDJ를 이용하는 HEK293T에서 수행하고, 생체내 실험을 AAV8 혈청형을 사용하여 마우스당 0.5E12-1E12 AAV8 입자의 역가에서 꼬리 정맥 주사를 통해 AAV 전달로 C57BL/6J, 10주령 마우스에서 수행하였다. 마우스를 형질감염 4주 후 분석하였다. 출원인은 시험관내 및 생체내 시나리오 둘 다에서 및 다중 게놈 유전자자리를 거쳐 RNA 및 면역형광성 기반 단백질 발현을 통해 분석한 바와 같은, 표적화된 유전자 억제 및 활성화를 확인하였다(도 16e-j, 도 18). 특히, 본 출원인은 CD81 유전자자리(n=4)에서 ~80% 생체내 억제, 및 Afp 유전자자리(n=4)의 >2 배 생체내 활성화를 달성하였다. 따라서, 이러한 시스템은 유전자 발현의 미세한 조절을 위한 방식을 포장하고 생체내 게놈 가공 적용을 위한 반흔이 없는 시도(scarless approach)를 제공한다.

[0308] AAV내로 CRISPR 효과기 혼입에 있어서 프로그램가능성의 확립으로, 본 출원인은 다음에 캡시드 슈도타이핑에 있어서 용이한 프로그램가능성을 가능하게 하는 이들의 집중을 조율하였다. AAV 캡시드 단백질은 대형 펩타이드 또는 생물분자의 삽입(역가 또는 기능성의 유의적인 손실없이)에 대해 전형적으로 융통성이 없다. 따라서 출원인은 생물-직교 클릭 화학 핸들의 비천연-아미노산(UAA) 매개된 혼입에 의한 이러한 한계를 극복하여 용이한 캡시드 변형이 가능하도록 하는 신규하고 다재다능한 시도를 개발하였다. 출원인은 우선 AAV2 표면에서 계산적으로 맵핑된 접근가능한 아미노산 부위를 맵핑하고 잠재적인 후보 부위로서 R447, N587, S578 및 S662에서 이들의 평가에 집중하였다(도 17b). 목적인 UAA는 AAV VP1 단백질내 상응하는 아미노산에서 제지정된 넌센스 코돈(TAG)에 의해 유전적으로 암호화되었으며, 직교 UAA 특이적인 tRNA/아미노아실-tRNA 신테타제(tRNA/aaRS) 쌍을 사용하여 캡시드내로 동시-해독적으로 혼입되었다(도 17a, 도 19). 따라서, 본 출원인은 AAV2 캡시드 표면으로 아지드 변형된 라이신-기반의 아미노산 - N-엡실론-((2-아지도에톡시)카보닐)-L-라이신을 최고의 상대적인 생산 역가 및 바이러스 활성을 나타내는 N587 및 S578 변형과 함께 성공적으로 혼입할 수 있었다(도 17c).

[0309] 본 출원인은 다음에 2회의 독립된 슈도타이핑 실험을 통한 UAA 혼입에 의해 즉석 캡시드 가공이 가능하였음을 입증하였다: 첫째, 출원인은 클릭 화학 반응을 수행하여 형광성 분자인, Alexa 594 DIBO 알킨을 바이러스에 연결하고 세포의 형질도입을 통해 변형된 형광성 바이러스를 성공적으로 가시화하였다(도 17d); 둘째, 출원인은 알킨-태그된 올리고뉴클레오타이드를 클릭 화학을 통해 AAV 표면 위로 묶고 이들의 상부에 배양된 세포의 형질도입에 의해 입증된 바와 같은, 상응하는 상보성 올리고뉴클레오타이드를 지닌 DNA 배열 스폿에서 이들의 선택적인 포획을 입증하였다(도 17e). 최종적으로, 출원인은 UAA 변형된 AAV가 split-Cas9 기반 게놈 가공 페이로드(payload)를 혼입할 수 있으며(도 17f) 견고한 게놈 편집(도 17g)을 수행함으로써, 통합된 mAAV 전달 플랫폼을 달성함을 확인하였다.

[0310] 이와 함께, 출원인의시도는 Cas9 및 dCas9 기반 효과기를 사용한 내인성 유전자의 발현 및 또한 이들의 표면에서 UAA의 혼입을 통한 즉석 AAV 슈도타이핑을 편집하고 조절하는 용이하고 간단한 방법을 제공한다. 이러한 시스템은 split-Cas9 시스템의 활용을 포함하는, 몇가지 장점을 가지는데, 이는 AAV(~4.7kb)의 제한된 카고 능력(cargo capacity)으로 인하여, 게놈 편집 및 조절을 포함하는, 모든 목적인 게놈 가공 적용을 수행하는데 최적이다. 또한, 이러한 시스템의 다른 장점은 목적인 바람직한 보조 성분을 활용하여 페이로드의 전사를 최적화할 수 있다는 것이다. 출원인은 이들의 mAAV-Cas9 시스템을 활용하여 높은 수준의 생체내 전사 회귀(~80%)(도 16g, 16j) 및 생체내 전사 활성화(>2배 증가)(도 16i)를 달성할 수 있음을 나타낸다. 또한, 출원인은 이들의 시스템을 활용하여 HEK293Ts, CD34+ HSC 세포 내에서 실험관내에서 및 C57BL/6J 마우스에서 생체내에서 세포를 편집할 수 있음을 나타낸다(도 16b). CD34+ HSC를 표적화하는데 있어서 높은 치료학적 가치를 고려할 때, 출원인은 이들의 모든 AAV 시스템이 이들 세포에 대해 다목적의 전달체를 개발하기 위한 강력한 재원을 제공할 수 있다고 믿는다. 중요하게도, 출원인은 또한 이들의 유도성 합성 스위치를 사용하여 게놈 편집에 걸쳐 일시적인 조절을 입증하며, 이는 Cas9 뉴클레아제의 발현을 제한하므로, 높은 치료학적 가치가 있다(도 16c, 16d). 출원인은 이러한 mAAV 시스템이 또한 클릭 화학의 공정을 통하여 캡시드 표면으로 앵타머의 용이하고 신속한 첨가를 허용함을 나타낸다. 이는 캡시드 표면의 프로그램가능한 슈도타이핑의 숙주(host)에 대한 문을 열어 AAV 표적 세포형 특이성을 전신계적으로 가공할 뿐만 아니라 세포내로 AAV 형질도입의 기본 생물학을 연구한다. 출원인은 이러한 벡터가 지시된 진화, 분자 셔플링(shuffling) 및 진화 계통 분석을 기반으로 하는 것과 같은 신규한 AAV 벡터를 가공하기 위한 다른 전략을 보충하고, AAV 활성을 조절하기 위한 앵타머 및 다른 모이어티의 전신계 진화를 기반으로 모듈러 부분이 가능하도록 할 것임을 기대한다. 출원인은 또한 mAAV 시스템의 일부 잠재적인 제한을 주목한다: 첫째로, split-Cas9 시스템을 활용하는 것은 구성성분, C-Cas9 및 N-Cas9 둘 다로서 표적화 효능을 감소시킬 것이며, Cas9 활성을 회복하기 위하여 목적인 표적 세포로 동시-전달되어야 하고; 둘째로, UAA를 통한 캡시드의 변형은 1.5 내지 5배 더 낮은 바이러스 역가를 야기한다. 출원인은 국제화된 조직-특이적인 전달 및 AAV 생산 매개변수의 최적화를 위한 기술에 있어서의 개선으로, 이러한 국면이 혁신적으로 처리될 것임을 예측한다. 이와 함께, 본 출원인은 이들의 다목적 mAAV 합성 전달 플랫폼이, CRISPR 효과기 혼입 및 캡시드 슈도타이핑에 있어서 이들의 용이한 프로그램능력(programmability)을 통해, 기본 과학 및 치료학적 적용에서 광범위한 유용성을 가질 것임을 기대한다.

[0311] 실시예 2 - AAV2 캡시드로의 비천연 아미노산 첨가

[0312] 다음은 프로토콜의 개요이다:

[0313] 1. 비-기본형 아미노산 혼입의 시험

[0314] 2. 삽입된 TAG를 지닌 AAV 캡시드 작제물의 생성

- [0315] 3. 이의 캡시드 속에 비 기본형 아미노산을 함유하는 AAV의 생성
- [0316] 4. MUC-1 앵타머 및 A549 세포를 사용한 가설의 시험
- [0317] 5. MUC-1 앵타머를 함유하는 생성된 AAV2가 혼합된 세포의 집단 속에서 A549를 선택적으로 형질도입할 수 있는 지를 시험
- [0318] 6. 혼합된 세포의 집단 속에서 A549에 Cas9를 선택적으로 전달하고 유전자 편집에 대해 점검하기 위해 생성된 AAV2 용도
- [0319] 7. 생체내 실험: CRISPR-Cas9에 대한 AAV2 생성된 전달 메카니즘의 사용 및 표적 세포내에서 유전자 편집의 점검
- [0320] 본 출원인은 GFP 유전자의 중간에 TAG 정지 코돈을 함유하는 GFP 리포터 플라스미드내로 비 기본형 아미노산의 혼입을 시험함으로써 시작하였다. tRNA, tRNA 신테타제 및 비 기본형 아미노산의 존재하에서, 앵버(Amber) 억제제의 사용을 위해, GFP 발현을 회복시켰다(도 13a). 출원인은 또한 리포터 대 신테타제 비를 변화시키고(1:1, 1:2.5 및 1:5) 결과를 도 13b에 나타낸다.
- [0321] 본 출원인은 앵버(amber) 억제 방법을 사용하여 바이러스 캡시드에 비천연 아미노산을 가하였다. 출원인은 표면 잔기 R447, S578, N587 및 S662 대신에 정지 코돈 TAG를 가하여 혼입시켰다. 출원인은 바이러스가 tRNA/신테타제 쌍 및 비천연 아미노산의 존재하에서만 생산될 수 있다는 가설을 세웠다. 실험을 수행하여 지금까지 이를 정확히 출원인에게 나타내는 것으로 보였다. 비천연 아미노산의 부재하에서 바이러스 역가는 매우 낮은 반면 비천연 아미노산이 첨가된 경우 이들은 수배(200x) 더 높다. 출원인은 규정된 잔기에서 비 기본형 아미노산 N-엡실론-((2-아지도에톡시)카보닐)-L-라이신을 함유하는 4개의 상이한 바이러스를 생성하였다(도 14).
- [0322] 다음에 출원인은 알킨 그룹을 함유하는 MUC-1 앵타머를 설계하여 비 기본형 아미노산이 아지드 그룹을 함유하므로 클릭 화학을 통해 이를 비 기본형 아미노산에 첨가하려고 한다. AAV2는 A549 폐암 세포주를 매우 효율적으로 감염시키지 않는다. A549 세포는 이의 표면에서 MUC-1의 과발현을 나타내며 출원인은 AAV2에 첨가된 MUC-1 앵타머가 A549 세포에 대해 바이러스의 특이성을 증진시키는데 도움이 될 수 있는 것으로 믿는다.
- [0323] 실시예 3 - AAV2- SpyTag
- [0324] 링커 펩타이드를 지닌 SpyTag 및 SpyTag를 AAV2의 4개의 버전을 생성하는 HSPG 결합 펩타이드의 존재하 및 부재하 둘 모두에서 AAV2 캡시드의 잔기 N587에서 도입하였다(도 15).
- [0325] 실시예 4 - AAV-DJ
- [0326] 이러한 시스템의 보다 광범위한 사용을 촉진하기 위하여, 출원인은 또한 AAV-DJ 혈청형을 가공하여 UAA를 유사하게 혼입하였다. 이를 향해, 단백질 정렬을 기반으로, AAV-DJ내 N589를 AAV2내 N587에 대해 등가의 부위로 선택하였다. 출원인은 AAV-DJ-N589UAA 바이러스가 AAV2-N587UAA 바이러스보다 5 내지 15배 더 높은 역가를 가졌음을 관찰하였고(도 20a), AAV2 및 AAV-DJ 각각에서 잔기 N587 및 N589 대신에 UAA의 혼입이 바이러스의 활성에 부정적으로 영향을 미치지 않음을 확인하였다(도 20b).
- [0327] 혈청 속에서 AAV 중화 항체의 만연은 생체내 연구 및 치료학적 적용시 이들의 효과적인 사용에 대한 주요 장애물이다. 따라서 출원인은 이러한 시스템의 프로그래밍능력(programmability)을 이용하여, AAV 항체에 의한 중화까지 AAV의 차폐 정도를 가능하게 할 수 있는 AAV 캡시드에 신규 표면 특성을 부여하는 것이 가능한지를 추정하였다(도 20c). 이러한 '쉬스(stealth)' AAV의 가공에 대해 본 출원인은 다수의 작은 분자 및 중합체 모이어티를 AAV 캡시드 표면 위에 테더링하고 중화 AAV 항체⁴⁸⁻⁵⁰를 지닌 것으로 알려진 돼지 혈청(도 20d)에 대한 노출 후 수득되는 AAV 형질도입 능력을 평가함으로써 다수의 작은 분자 및 중합체 모이어티를 스크리닝하였다. 흥미롭게도, 본 출원인은 지질을 통한 차폐가 돼지 혈청-기반의 중화에 대한 AAV의 거의 완전한 내성을 생성하였음을 관찰하였다. 출원인은 올리고뉴클레오타이드를 AAV 표면 위에 테더링함을 통해 이를 달성하였으며, 이를 최종적으로 사용하여 시판되는 지질 중합체 제형 리포펙타민을 결합시켰다. 특히, 출원인은 wt AAV-DJ 및 AAV-DJ-N589 바이러스가 완전하게 중화되는 조건 하에서도 지질-코팅된 바이러스의 능력을 관찰하였다(도 20d). 출원인은 또한 이러한 가공된 바이러스가 완전한 게놈 편집 가능성을 보유하며, 특히 리포펙타민의 존재하에서 코팅이 변형되지 않은 바이러스와 비교하여 향상된 편집율을 나타내었음을 확인하였다. 따라서, 이러한 시도는 AAV 캡시드 표면 특성의 프로그래밍가능한 조절을 위한 방식을 포함함으로써 AAV 활성을 조절하기 위한 소 분자 및 중합체의 전신계적 평가를 가능하도록 한다.

- [0328] 실시예 5 - 조직 특이성에 대한 miRNA
- [0329] 본 출원인은 리포터 유전자로서 TK-GFP(티미딘 키나제 GFP 융합 단백질)을 사용함으로써 이러한 예시적인 시스템의 특이성 및 전달을 평가하였다. TK-GFP는 PET/SPECT를 사용하여 전체 동물의 실시간 생체내 영상화를 허용하며, 이는 qPCR과 같이 정량적 정보를 제공하면서 바이러스가 어느 조직을 감염시키는지에 대한 공간 정보를 제공한다.
- [0330] 실시예 6 - 통증 관리
- [0331] 본 출원인은 이용된 총 9마리의 C57BL/6J 마우스 내에서 이들의 통증 관리 시스템을 시험한다. 3마리의 마우스에게 pAAV9_gSCN9a_dCas9 시스템을 주사하고, 3마리의 마우스에게 빈 벡터(empty vector), pAAV9_gempty_dCas9를 주사하며, 3마리의 SNC9a 돌연변이체 마우스(Scn9atm1Dgen)를 양성 대조군으로 사용한다. 출원인은 또한 사람 뉴우런 세포를 이용하여 시험관내에서 사람 gRNA를 시험하였다.
- [0332] 실시예 7 - CD81 억제
- [0333] 출원인은 split-Cas9 및 split-dCas9 시스템을 설계하여 간에서 3개의 말라리아 숙주 유전자, CD81, Sr-b1, 및 MUC13를 표적화하여, 이들을 억제하고 편집하였다. 이들은 간세포의 플라스모디움 스포로조이테(plasmodium sporozoite) 감염에 필요한 숙주 인자이다. 출원인은 생체내에서 CD81의 억제를 시험하고, 35%의 억제를 검출하였다(도 8 및 9). 도 8은 AAV8_gCD81_KRAB_dCas9로 처리한 3마리의 마우스 및 6마리의 대조군 마우스에서 CD81의 상대적인 발현을 나타낸다. 도 9는 조직학 샘플의 3개 세트를 나타낸다: 첫번째는 주요 항체를 가지지 않으며, 두번째는 CD81의 비교적 높은 발현을 나타내는 양성 대조군이고, 세번째는 CD81의 감소된 발현을 나타내는, AAV8_gCD81_KRAB_dCas9이 전달된 세트이다.
- [0334] 실시예 8 - 통증 관리
- [0335] 통증에는 3개의 주요 특성이 존재한다: 기간(급성 내지 만성), 위치(예컨대, 근육, 구강안면), 및 원인(예컨대, 신경 손상, 염증). 출원인은 4가지 주요 종류의 통증 모델(화상 모델, 염증, 수술후, 및 신경병)을 이용하여 1) 본 출원인의 치료요법이 표적화하는 통증이 종류가 무엇인지 및 2) 본 출원인의 치료가 통증 관리를 위한 전통적인 방법, 예컨대, 오피오이드로부터 유사한 결과 또는 증진을 나타내는지를 추가로 이해한다. 이러한 통증 모델은 하기 표에 요약한다. 급성 통각 화상 모델의 경우, 출원인은 2개의 일반적으로 활용된 모델을 이용한다: 일반적으로 약물 또는 생리학적 조작의 진통 활성에 대해 스크리닝하기 위한 검정으로서 통각 과정을 평가하기 위해 이용된, 핫 플레이트 시험(hot plate test) 및 "하그레이브스(Hargreaves)" 시험. 제1 모델의 경우, 동물이 점핑 또는 이의 발의 핏기와 같은유독한 열적 자극 후 공지된 거동을 유발할 때까지 55°C에 둔다. 동물이 45 초 전에 반응하지 않은 경우, 이를 핫 플레이트로부터 제거하여 조직 손상을 피한다. 이후에, 기계적 임계값(mechanical threshold)을 폰 프레이 필라멘트(von Frey filament), 움추림 반응을 측정하는, 대수적으로 증가하는 강성(0.41, 0.70, 1.20, 2.00 g)을 지닌 나일론 섬유를 사용하여 측정한다. 열적 통각 반응을 이후에 하그레이브스로 알려진 상이한 실험에서 시험한다. 요약하면, 마우스를 가열된(30°C) 유리 표면 위의 플렉시글라스 좁은방(Plexiglas cubicle)에 두고, 유리 아래에 위치한 집중된 투사 전구를 통한 빛을 한쪽 뒷발의 발바닥 표면에 투사하였다. 열 움추림 반응을 손상 후 3시간 동안 30분 마다 측정한다. 발 움추림 잠복기로서 정의된, 광의 적용과 뒷발 움추림 반응 사이의 시간 간격(PWL: s)을 이후에 측정한다. 염증성 통증 모델의 경우, 출원인은 관절염 유전자이식 K/BxN 마우스로부터의 혈청을 야생형 마우스에 주사하여 2 내지 3주 동안 지속되는 관절/발 염증과 관련된 시작으로 강하고 큰 기계적 무해자극통증을 지닌 마우스를 생산하였다. 앞서 기술된 바와 같이 폰 프레이 필라멘트를 통한 기계적 역치를 또한 측정할 것이다. 다음 수술후 모델을 위해, 피부, 안구집, 및 마우스의 뒷발의 발바닥 국면의 근육을 통해 마취하에 절개한다. 움추림 반응을 수술 후 6일 동안 상처 주변의 분명한 부위에서 폰 프레이 필라멘트를 사용하여 측정한다.

통증 모델의 유형	상해	참고
급성 통각: 화상 모델		
	핫 플레이트 및 "하그레이브스"	Nozaki-Taguchi and Yaksh (1998) Neurosci. Lett. 254(1):25-8
염증성 통증 모델		
	관절염 (마우스내로 주사된 K/BxN 혈청)	Christianson et al., (2012) Methods Mol. Biol. 851:249-260
수술후 통증 모델		
	절개 모델(통각과민증)	Brennan et al. (1996) Pain 64(3):493-501
신경병적 통증 모델		
	척추 신경 연결/형단술	Kim and Chung (1992) Pain 50(3):355-363
	화학치료요법(시스플라틴)	Balayssac et al.(2009) Neurosci. Lett. 465(1):108-1112

[0336]

[0337]

마지막으로, 본 출원인은 2개의 신경병적 통증 모델을 이용할 것이다: 척추 신경 연결 및 시스플라틴을 이용한 화학치료요법. 첫번째 모델, 청 모델(Chung model)로 또한 알려진, 척추 신경 연결(SNL)에서, L5 및 L6 척추 신경을 L4 척추 신경으로부터 절개하고 후근 신경절(DRG)에서 멀리 단단하게 연결한다. 화학치료요법 모델의 경우, 마우스에게 8주 동안 주당 5mg/kg에서 시스플라틴의 투여량을 제공할 것이다. 신경병 모델은 기계적 무해 자극통증, 냉 무해자극통증, 및 열 통각과민증과 같은, 거동 변화를 가진 것으로 알려져 있다. 이러한 이유로, 방사상 열(radiant heat)의 적용으로 인한 움추림 잠복기에 대해 시험하기 위한 하그레이브스 시험 및 기계적 자극에 대해 시험하기 위한 폰 프레이 시험을 이용한다.

[0338]

AAV 혈청형이 DRG(후근 신경절)을 표적화하는데 최적임을 측정한 후(도 25), 출원인은 몇가지 유전자를 표적화하는 실험을 수행한다.

Nav 1.3 (SCN3A)	억제/KO
Nav 1.7 (SCN9A)	억제/KO
Nav 1.8 (SCN10A)	억제/KO
Nav 1.9 (SCN11A)	억제/KO
SHANK3	억제/KO
NMDA 수용체 길항제(NR2B 포함)	억제/KO
IL-10	활성화(과발현)
Penk	활성화(과발현)
Pomc	활성화(과발현)
MV11A-PC	활성화(과발현)

[0339]

[0340]

실험의 제1 라운드에서, 출원인은 먼저 SCN9A 유전자를 편집한다. 출원인은 C57BL/6J 마우스에게 SCN9A 유전자를 표적화하는 split-Cas9와 함께 ~1E11-1E12 vg/마우스의 AAV를 척추강내에 주사한다. 이후에, 출원인은 다른 마우스를 5개 그룹으로 분리하여 상이한 통증 모델을 시험하며, WT 마우스에게는 양성 대조군으로서 오피오이드를 주사하고, 마우스에게는 음성 대조군으로서 PBS를 주사하였다. 8주 말기에, 출원인은 마우스를 희생시키고, DRG로부터 gDNA를 추출하고 차세대 서열분석을 통해 목적인 표적화된 영역(절단 부위 주변의 150bp)을 서열분석한다. 통증의 영구적인 손실은 바람직하지 않을 수 있으므로, 출원인은 또한 dCas9 및 최적화된 억제 도메인을 통해 SCN9A를 표적화한다(도 33). 출원인은 다시 통증 모델과 함께 이러한 마우스의 세트를 시험한다. 또한, 본 출원인은 8주째에 마우스 DRG 뉴우런을 수거하고 RNA-서열분석을 수행하여 치료요법 후 유전자 발현에 있어서의 변화를 측정하였다. 출원인이 표적화하는 일부 추가의 유전자는 Nav 1.8(SCN10A 유전자), 1.9(SCN11A 유전자) 및 1.3(SCN3A 유전자)와 같은 다른 나트륨 채널 뿐 아니라, 또한 캡사이신 수용체 및 바닐로이드 수용체 1,

SHANK3, 및 NMDA 수용체 길항제로서 알려진, 일시적인 수용체 전위 양이온 채널 서브계열 V 구성원 1(TrpV1)을 포함한다. 유전자 억제 통증이 없는 상태를 달성하는데 충분하지 않을 수 있으므로, 출원인은 또한 유전자 활성화(또는 과발현)를 수행하였다.

[0341] 앞서의 연구는 SCN9A의 동시 억제 및 엔케팔린 전구체 Penk의 상향조절이 통증이 없는 표현형에 필수적일 수 있음을 나타내었다. 이러한 이유로, 출원인은 RNA 헤어핀(MS2, PP7, Com)을 지닌 gRNA 작제물을 이용하여 이들의 유사한 RNA-결합 단백질을 활성화/억제 도메인에 융합시킨다. Penk의 활성화를 위해, 출원인은 dN-Cas9 플라스미드 위에 gRNA-MS2 작제물을 작제하고 MS2 RNA 유사체인, MCP를 VP64 활성화 부위에 융합시킨다. 유사하게, 출원인은 SCN9A 특이적인 gRNA-Com을 dN-Cas9 및 이의 RNA 유사체에 가하고, COM을 KRAB로 융합시킨다. 따라서, 출원인은 특이적인 위치에 대한 선택적인 활성화/억제를 보충함으로써, 동시 활성화 및 억제를 허용할 gRNA에 부착된 RNA 헤어핀을 지닌 이중-AAV dCas9 시스템을 이용할 수 있다(도 33 및 34). 따라서, 출원인은 Penk를 동시에 활성화시키고 SCN9A를 억제하는 AAV를 마우스에게 주사하여, 마우스의 통증 표현형에 있어서 임의의 차이가 존재하는지 및 RNA-seq에 대해 수행하여 활성화/억제의 정도를 측정할 것이다. 억제를 위한 SCN9A 및 활성화를 위한 Penk 외에, 출원인은 동시 활성화/억제를 위해 다른 유전자를 표적화하는 중이다. 또한, CRISPR을 통한 동시 활성화 및 억제를 수행하는 것 외에, 출원인은 dCas9-KRAB-gRNA split-AAV 작제물을 통한 억제 및 유전자의 과발현을 통한 동시 활성화를 수행중에 있다(도 35).

[0342] **등가물**

[0343] 달리 정의하지 않는 한, 본원에 사용된 모든 기술적 및 과학적 용어는 본 기술이 속하는 분야의 통상의 기술자가 일반적으로 이해하는 바와 동일한 의미를 갖는다.

[0344] 본원에 설명적으로 기술된 본 기술은 본원에 구체적으로 개시되지 않은 임의의 성분 또는 성분들, 한계 또는 한계들의 부재하에서 적합하게 실시될 수 있다. 따라서, 예를 들어, 용어 "포함하는", "포괄하는", "함유하는" 등은 확장적으로 및 제한없이 관독될 것이다. 또한, 본원에 사용된 용어 및 표현은 기술의 측면에서 제한없이 사용되었으며, 이러한 용어 및 표현의 사용시 나타내고 기술된 특징의 임의의 등가물 또는 이의 일부를 배제하는 것으로 의도되지 않지만, 다양한 변형이 가능한 한 특허청구되어 있는 본 기술의 영역내에 있음이 인식된다.

[0345] 따라서, 본원에 제공된 물질, 방법, 및 실시예는 바람직한 국면을 대표하며, 예시적이고, 본 기술의 영역을 제한하는 것으로 의도되지 않음이 이해되어야 한다.

[0346] 본 기술은 본원에 광범위하게 및 일반적으로 기술되었다. 일반적인 개시내용에 속하는 보다 협소한 종 및 하위-종 그룹 각각은 또한 본 기술의 일부를 형성한다. 이는 본 기술의 일반적인 설명을 포함하며 배제된 물질이 본원에 구체적으로 인용되어 있는지의 여부에 상관없이, 속(genus)으로부터 임의의 대상 물질을 제거함을 단서조함으로 하거나 부정적으로 한정한다.

[0347] 또한, 본 기술의 특징 또는 국면은 마르쿠쉬(Markush) 그룹이 측면에서 기술되어 있으며, 당해 분야의 기술자는 본 기술이 또한 이에 의해 마르쿠쉬 그룹의 임의의 개개 구성원 또는 구성원의 하위 그룹의 측면에서 기술되어 있음을 인식할 것이다.

[0348] 본원에 언급된 모든 공보, 특허원, 특허 및 다른 참고문헌은 각각이 개별적으로 참고로 혼입되는 경우와 동일한 정도로 이의 전문이 참고로 표현하여 혼입된다. 충돌하는 경우, 정의를 포함하는, 본 명세서가 제어할 것이다.

[0349] 다른 국면은 다음의 특허청구범위 내에 제시되어 있다.

[0350] **참고문헌**

1. Charpentier, E. & Doudna, J.A. Biotechnology: Rewriting a genome. *Nature* 495, 50-51 (2013).
2. Hwang, W.Y. et al. Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat Biotechnol* 31, 227-229 (2013).
3. Li, D. et al. Heritable gene targeting in the mouse and rat using a CRISPR-Cas system. *Nat Biotechnol* 31, 681-683 (2013).
4. Mali, P., Esvelt, K.M. & Church, G.M. Cas9 as a versatile tool for engineering biology. *Nat Methods* 10, 957-963 (2013).
5. Mali, P. et al. RNA-guided human genome engineering via Cas9. *Science* 339, 823-826 (2013).
6. Nakayama, T. et al. Simple and efficient CRISPR/Cas9-mediated targeted mutagenesis in *Xenopus tropicalis*. *Genesis* 51, 835-843 (2013).
7. Shan, Q. et al. Targeted genome modification of crop plants using a CRISPR-Cas system. *Nat Biotechnol* 31, 686-688 (2013).
8. Yang, D. et al. Effective gene targeting in rabbits using RNA-guided Cas9 nucleases. *J Mol Cell Biol* 6, 97-99 (2014).
9. Yu, Z. et al. Highly efficient genome modifications mediated by CRISPR/Cas9 in *Drosophila*. *Genetics* 195, 289-291 (2013).
10. DiCarlo, J.E. et al. Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. *Nucleic Acids Res* 41, 4336-4343 (2013).
11. Wang, J. et al. Homology-driven genome editing in hematopoietic stem and progenitor cells using ZFN mRNA and AAV6 donors. *Nat Biotechnol* 33, 1256-1263 (2015).

[0351]

12. Yang, Y. et al. A dual AAV system enables the Cas9-mediated correction of a metabolic liver disease in newborn mice. *Nat Biotechnol* 34, 334-338 (2016).
13. Long, C. et al. Postnatal genome editing partially restores dystrophin expression in a mouse model of muscular dystrophy. *Science* 351, 400-403 (2016).
14. Nelson, C.E. et al. *In vivo* genome editing improves muscle function in a mouse model of Duchenne muscular dystrophy. *Science* 351, 403-407 (2016).
15. Tabebordbar, M. et al. *In vivo* gene editing in dystrophic mouse muscle and muscle stem cells. *Science* 351, 407-411 (2016).
16. Ran, F.A. et al. *In vivo* genome editing using *Staphylococcus aureus* Cas9. *Nature* 520, 186-191 (2015).
17. Zuris, J.A. et al. Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing *in vitro* and *in vivo*. *Nat Biotechnol* 33, 73-80 (2015).
18. Hsu, P.D., Lander, E.S. & Zhang, F. Development and applications of CRISPR-Cas9 for genome engineering. *Cell* 157, 1262-1278 (2014).
19. Truong, D.J. et al. Development of an intein-mediated split-Cas9 system for gene therapy. *Nucleic Acids Res* 43, 6450-6458 (2015).
20. Wright, A.V. et al. Rational design of a split-Cas9 enzyme complex. *Proc Natl Acad Sci U S A* 112, 2984-2989 (2015).
21. Zetsche, B., Volz, S.E. & Zhang, F. A split-Cas9 architecture for inducible genome editing and transcription modulation. *Nat Biotechnol* 33, 139-142 (2015).
22. Davis, K.M., Pattanayak, V., Thompson, D.B., Zuris, J.A. & Liu, D.R. Small molecule-triggered Cas9 protein with improved genome-editing specificity. *Nat Chem Biol* 11, 316-318 (2015).

[0352]

23. Chavez, A. et al. Highly efficient Cas9-mediated transcriptional programming. *Nat Methods* 12, 326-328 (2015).
24. Gilbert, L.A. et al. CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* 154, 442-451 (2013).
25. Hilton, I.B. et al. Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. *Nat Biotechnol* 33, 510-517 (2015).
26. Mali, P. et al. CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat Biotechnol* 31, 833-838 (2013).
27. Qi, L.S. et al. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* 152, 1173-1183 (2013).
28. Maeder, M.L. et al. CRISPR RNA-guided activation of endogenous human genes. *Nat Methods* 10, 977-979 (2013).
29. Aslanidi, G.V. et al. Optimization of the capsid of recombinant adeno-associated virus 2 (AAV2) vectors: the final threshold? *PLoS One* 8, e59142 (2013).
30. Ried, M.U., Girod, A., Leike, K., Buning, H. & Hallek, M. Adeno-associated virus capsids displaying immunoglobulin-binding domains permit antibody-mediated vector retargeting to specific cell surface receptors. *J Virol* 76, 4559-4566 (2002).
31. Shi, W., Arnold, G.S. & Bartlett, J.S. Insertional mutagenesis of the adeno-associated virus type 2 (AAV2) capsid gene and generation of AAV2 vectors targeted to alternative cell-surface receptors. *Hum Gene Ther* 12, 1697-1711 (2001).
32. Wu, P. et al. Mutational analysis of the adeno-associated virus type 2 (AAV2) capsid gene and construction of AAV2 vectors with altered tropism. *J Virol* 74, 8635-8647 (2000).
33. Xie, Q. et al. The atomic structure of adeno-associated virus (AAV-2), a vector for human gene therapy. *Proc Natl Acad Sci U S A* 99, 10405-10410 (2002).
34. Chatterjee, A., Xiao, H., Bollong, M., Ai, H.W. & Schultz, P.G. Efficient viral delivery system for unnatural amino acid mutagenesis in mammalian cells. *Proc Natl Acad Sci U S A* 110, 11803-11808 (2013).
35. Schmied, W.H., Elsasser, S.J., Uttamapinant, C. & Chin, J.W. Efficient multisite unnatural amino acid incorporation in mammalian cells via optimized pyrrolysyl tRNA synthetase/tRNA expression and engineered eRF1. *J Am Chem Soc* 136, 15577-15583 (2014).
36. Elsasser, S.J., Ernst, R.J., Walker, O.S. & Chin, J.W. Genetic code expansion in stable cell lines enables encoded chromatin modification. *Nat Methods* 13, 158-164 (2016).
37. Zheng, Y. et al. Broadening the versatility of lentiviral vectors as a tool in nucleic acid research via genetic code expansion. *Nucleic Acids Res* 43, e73 (2015).
38. Deverman, B.E. et al. Cre-dependent selection yields AAV variants for widespread gene transfer to the adult brain. *Nat Biotechnol* 34, 204-209 (2016).
39. Grimm, D. et al. *In vitro* and *in vivo* gene therapy vector evolution via multispecies interbreeding and retargeting of adeno-associated viruses. *J Virol* 82, 5887-5911 (2008).
40. Maheshri, N., Koerber, J.T., Kaspar, B.K. & Schaffer, D.V. Directed evolution of adeno-associated virus yields enhanced gene delivery vectors. *Nat Biotechnol* 24, 198-204 (2006).
41. Zinn, E. et al. In Silico Reconstruction of the Viral Evolutionary Lineage Yields a Potent Gene Therapy Vector. *Cell Rep* 12, 1056-1068 (2015).
42. Guenther, C.M. et al. Synthetic virology: engineering viruses for gene delivery. *Wiley Interdiscip Rev Nanomed Nanobiotechnol* 6, 548-558 (2014).
43. Endy, D. Foundations for engineering biology. *Nature* 438, 449-453 (2005).

[0353]

[0354]

44. Guell, M., Yang, L. & Church, G.M. Genome editing assessment using CRISPR Genome Analyzer (CRISPR-GA). *Bioinformatics* 30, 2968-2970 (2014).

45. Mali, P. et al. Barcoding cells using cell-surface programmable DNA-binding domains. *Nat Methods* 10, 403-406 (2013).

46. Rapti, K. et al. Neutralizing Antibodies Against AAV Serotypes 1, 2, 6, and 9 in Sera of Commonly Used Animal Models. *Mol. Ther.* 20, 73-83 (2009).

47. Lee, G. K., Maheshri, N., Kaspar, B. & Schaffer, D. V. PEG Conjugation Moderately Protects Adeno-Associated Viral Vectors Against Antibody Neutralization. (2005). doi:10.1002/bit.20562

48. Fitzpatrick, Z., Crommentuijn, M. H. W., Mu, D. & Maguire, C. A. Biomaterials Naturally enveloped AAV vectors for shielding neutralizing antibodies and robust gene delivery in vivo. *35*, 7598-7609 (2014).

49. Lerch, T. F. et al. Structure of AAV-DJ, a Retargeted Gene Therapy Vector: Cryo-Electron Microscopy at 4.5Å resolution. *NIH Public Access.* 20, 1310-1320 (2013).

50. Chew, W. L. et al. A multifunctional AAV - CRISPR - Cas9 and its host response. *13*, (2016).

51. Kelemen et al. A Precise Chemical Strategy To Alter the Receptor Specificity of the Adeno-Associated Virus. 10-15 (2016). doi:10.1002/anie.201604067.

[0355]

[0356]

[표 1]

기능	유기체	유전자	서열
편집	호모 사피엔스	AAVS1	GGGGCCACTAGGGACAGGAT
편집	호모 사피엔스	EMX1	GAGTCCGAGCAGAAGAAGAA
편집	호모 사피엔스	FANCF	GGAATCCCTTCTGCAGCACC
편집	호모 사피엔스	B2M	CAGCCCAAGATAGTTAAGTG
조절	호모 사피엔스	CXCR4_1	CGGGTGGTGGTAGTGAGTC
조절	호모 사피엔스	CXCR4_2	CAGACGCGAGGAAGGAGGGCGC
조절	호모 사피엔스	ASCL1	CGGGAGAAAGGAACGGGAGG
조절	호모 사피엔스	RHOXF2_1	GACGCGTGCTCTCCCTCATC
조절	호모 사피엔스	RHOXF2_1	GCTGTGGGTTGGGCCTGCTG
편집	무스 무스쿨루스	APOB	ACCCACCATCCATCCGCCA
조절	무스 무스쿨루스	CD81	CGAAATTGAAGACGAAGAGC
조절	무스 무스쿨루스	AFP	GGACAAAGACCCTTCAGAG
편집	무스 무스쿨루스	디스트로핀	ATTTGAGGTAAGCCGAGGTT
편집	무스 무스쿨루스	디스트로핀	ATAATTTCTATTATATTACA

[0357]

[0358]

[표 2a]

유기체	유전자	전방	역방
호모 사피엔스	CXCR4	GAAGCTGTTGGCTGAAAAGG	CTCACTGACGTTGGCAAAGA
호모 사피엔스	RHOXF2	GGAGATTTAGGAAGTATGGGGTTAGT G	AAAACCTCCTCTCTTACTTTTCT ACTTC
호모 사피엔스	ASCL1	CGCGGCCAACAAAGAAGATG	CGACGAGTAGGATGAGACCG
호모 사피엔스	GAPDH	ACAGTCAGCCGCATCTTCTT	ACGACCAAATCCGTTGACTC
호모 사피엔스	B-actin	CATGTACGTTGCTATCCAGGC	CTCCTTAATGTCACGCACGAT
무스 무스쿨루스	ApoB	GCTCAACTCAGGTTACCGTGA	AGGGTGTACTGGCAAGTTTGG
무스 무스쿨루스	Afp	CTTCCCTCATCCTCCTGCTAC	ACAAACTGGGTAAGGTGATGG
무스 무스쿨루스	CD81	GCTCTTCGTCTTCAATTTCTGCTCT	TGTTGGGTGCCGTTTGT
무스 무스쿨루스	GAPDH	TGGCCTTCCGTGTTCTTAC	GAGTTGCTGTTGAAGTCGCA
무스 무스쿨루스	B-actin	GTGACGTTGACATCCGTAAGA	GCCGGACTCATCGTACTCC

[0359]

[0360] [표 2b]

기능	유기체	유전자	프라이머	서열
편집	호모사피엔스	EMX1	전방	ACACTCTTCCCTACACGACGCTCTTCCGATCT AGTGCTGCTTGCTGCTGGCCA
			역방	GACTGGAGTTCAGACGTGTGCTCTTCCGATCT TTGCTTGCCCTCTGTCAATGGCG
편집	호모사피엔스	AAVS1	전방	ACACTCTTCCCTACACGACGCTCTTCCGATCT CGGTTAATGTGGCTCTGGTTCTGG
			역방	GACTGGAGTTCAGACGTGTGCTCTTCCGATCT GGGTTAGACCCAATATCAGGAGACTAG
편집	호모사피엔스	B2M	전방	ACACTCTTCCCTACACGACGCTCTTCCGATCT ATGAGTATGCCTGCCGTG
			역방	GACTGGAGTTCAGACGTGTGCTCTTCCGATCT GGGACTCATTGAGGGTAGT
편집	무스 무스쿨루스	CD81	전방	ACACTCTTCCCTACACGACGCTCTTCCGATCT AGGACCAATCCAAGCTCCGC
			역방	GACTGGAGTTCAGACGTGTGCTCTTCCGATCT TTGCGCTGCCCTTCTCAG
편집	무스 무스쿨루스	ApoB	전방	ACACTCTTCCCTACACGACGCTCTTCCGATCT TGTAAGCAAGCAGCAGGGGC
			역방	GACTGGAGTTCAGACGTGTGCTCTTCCGATCT GGTCCAAGAACAGTAGCAGGAAC

[0361]

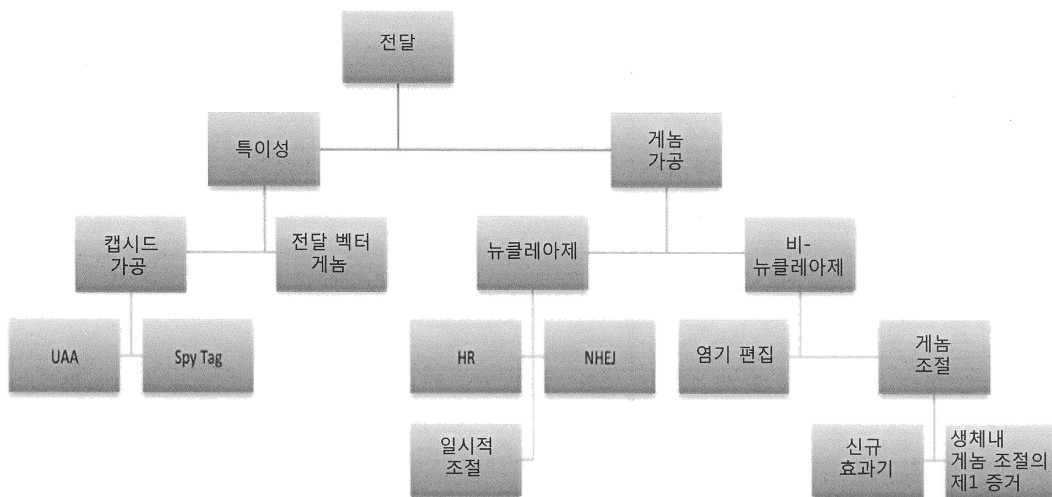
[0362] [표 2c]

명칭	서열 (5'에서 3')	5' 그룹
A	AAAACTATATTACCCTGTTAT CCCTAGCGTAACT	핵시닐
B	AAAAATATAAGCGGGAGATTC GTCCTCATA	핵시닐
A'	AGTTACGCTAGGGATAACAG GGTAATATAG	Biosg
B'	TATGAGGACGAATCTCCCGCT TATA	Biosg

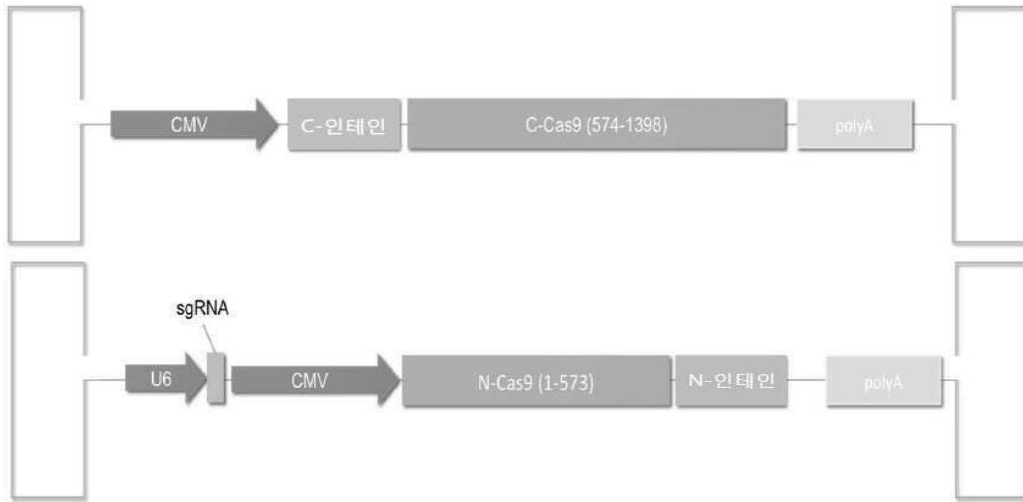
[0363]

도면

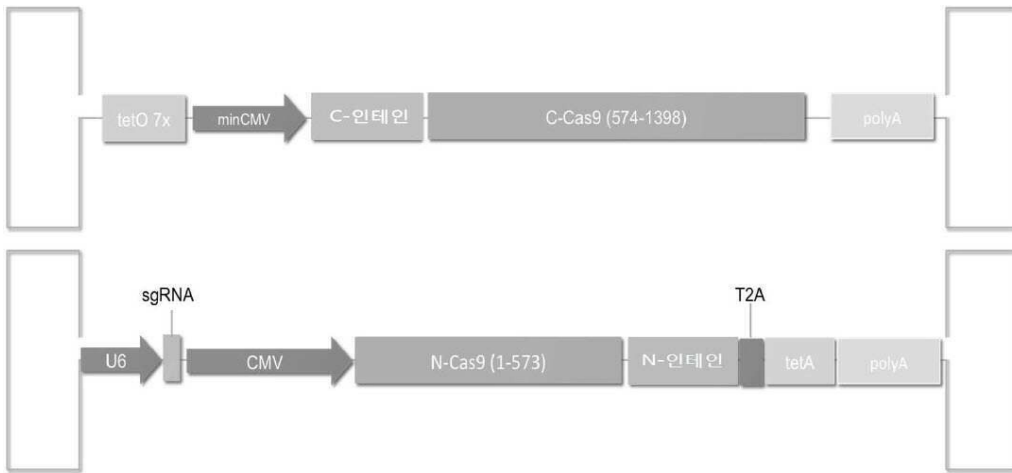
도면1



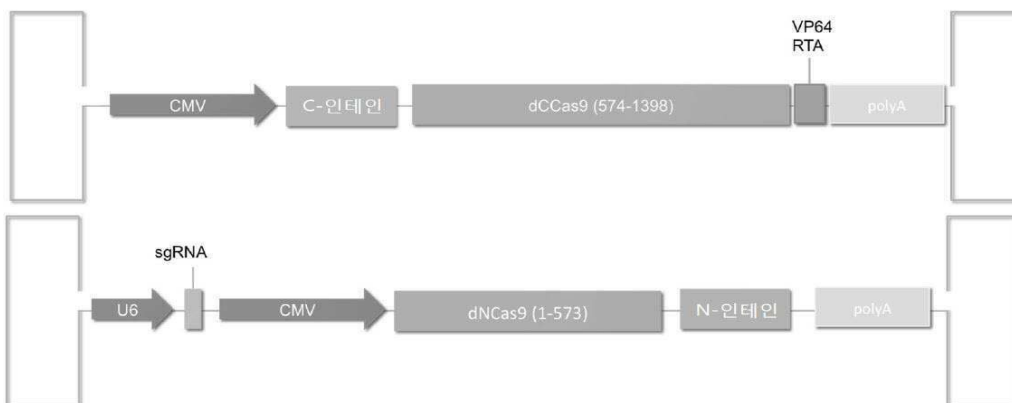
도면2



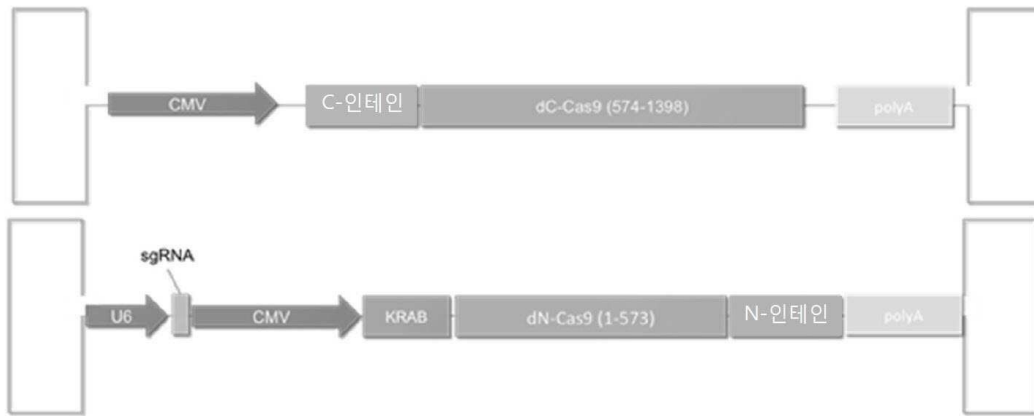
도면3



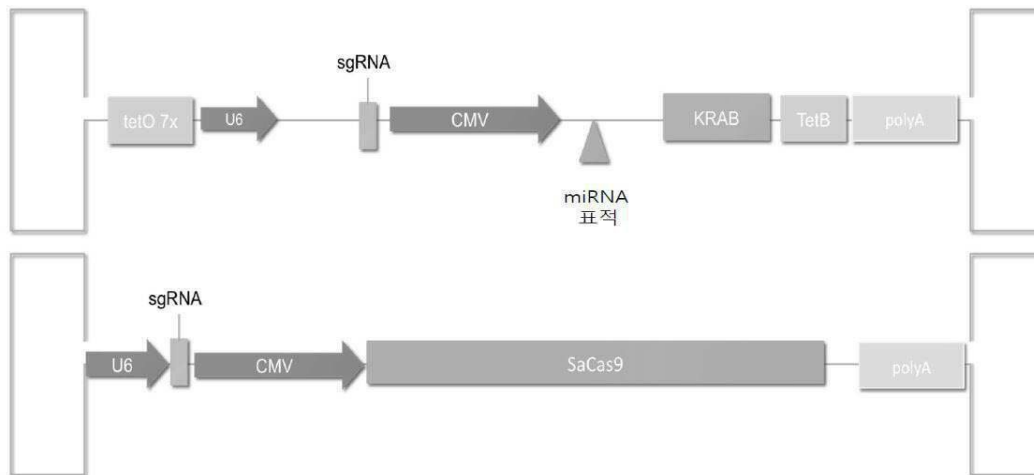
도면4a



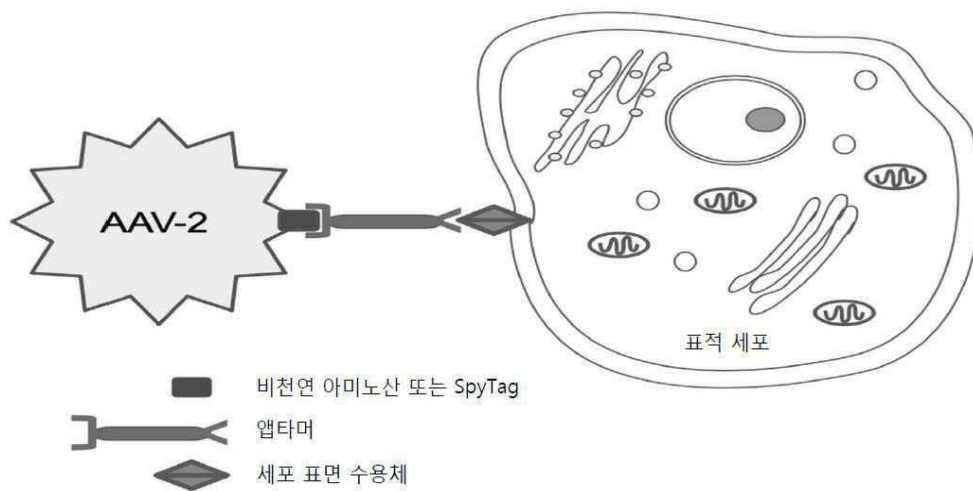
도면4b



도면5



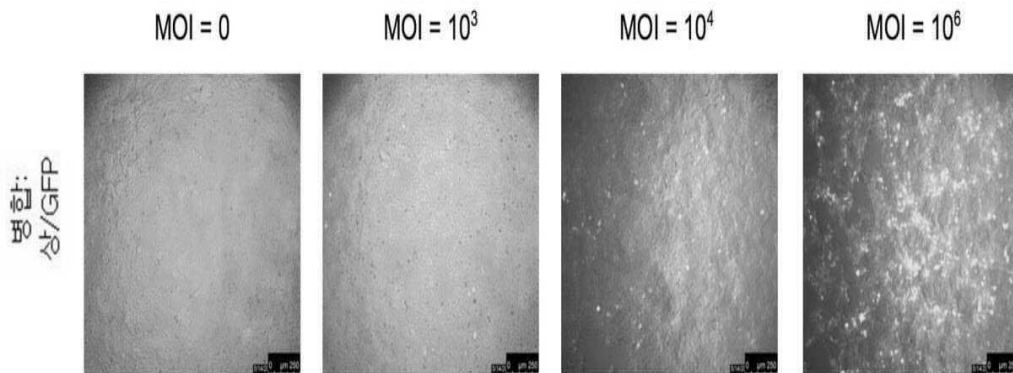
도면6



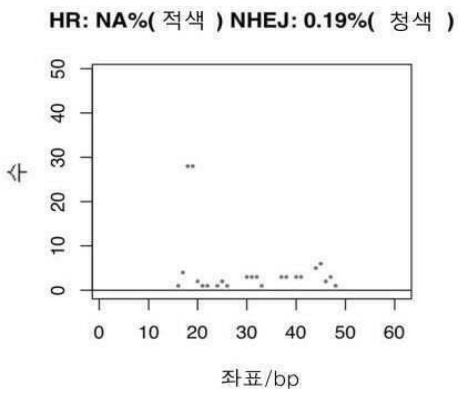
도면7a



도면7b

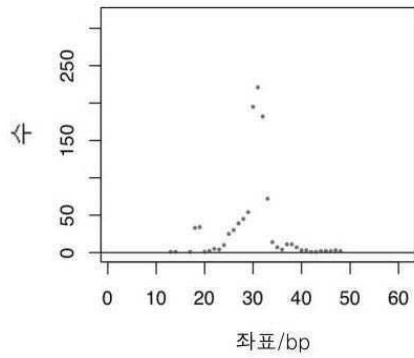


도면8a

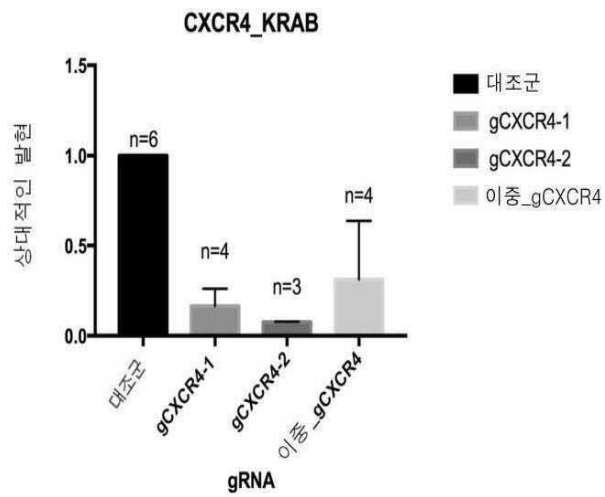


도면8b

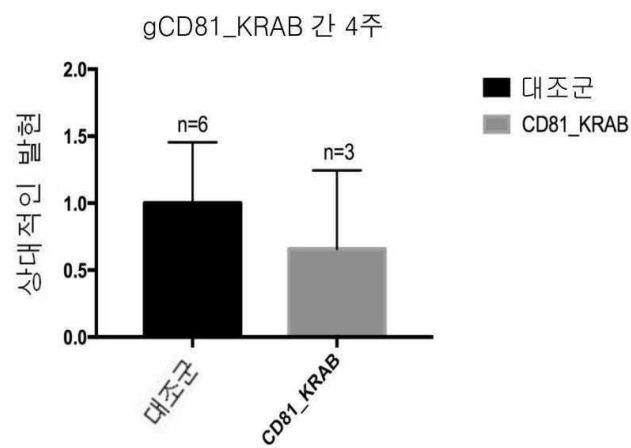
HR: NA%(적색) NHEJ: 1.7%(청색)



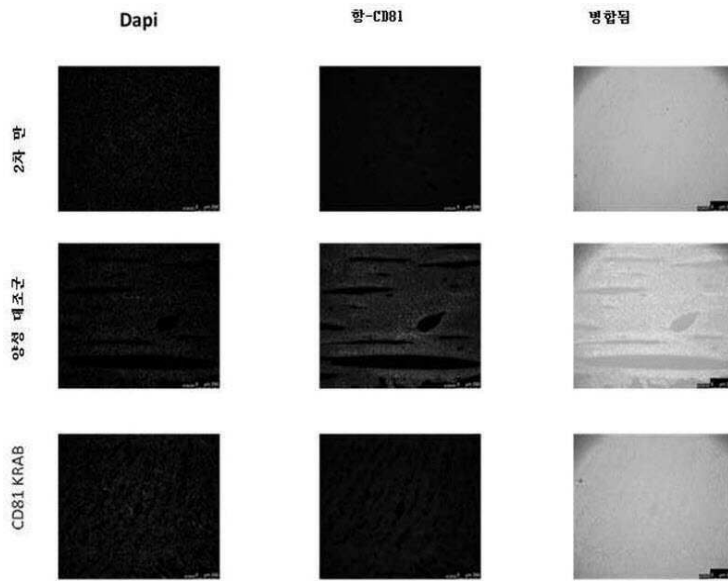
도면9



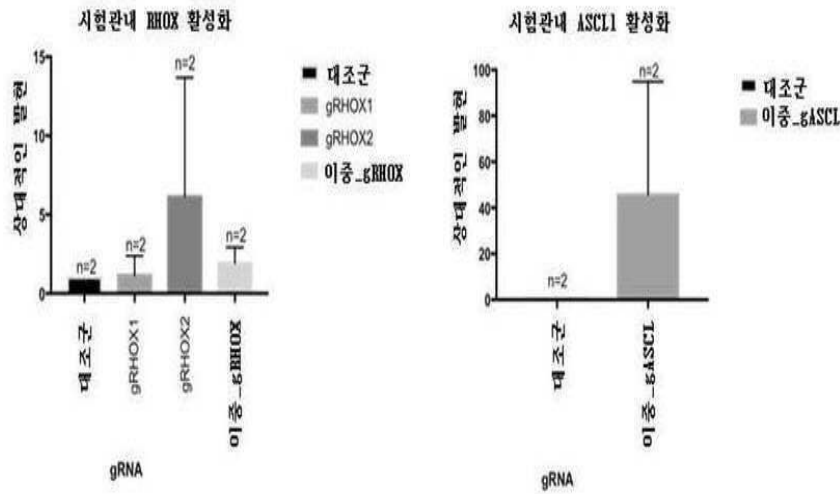
도면10



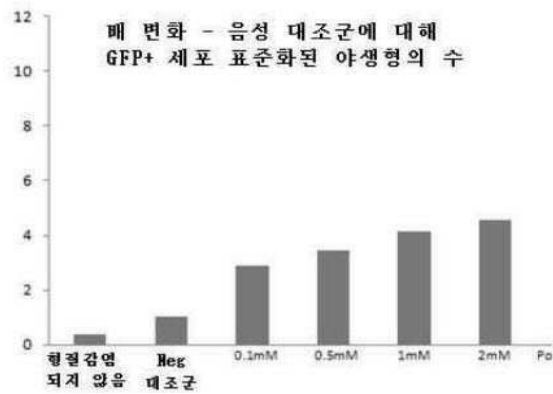
도면11



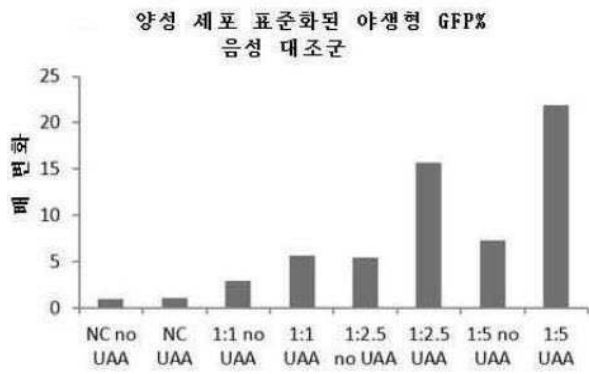
도면12



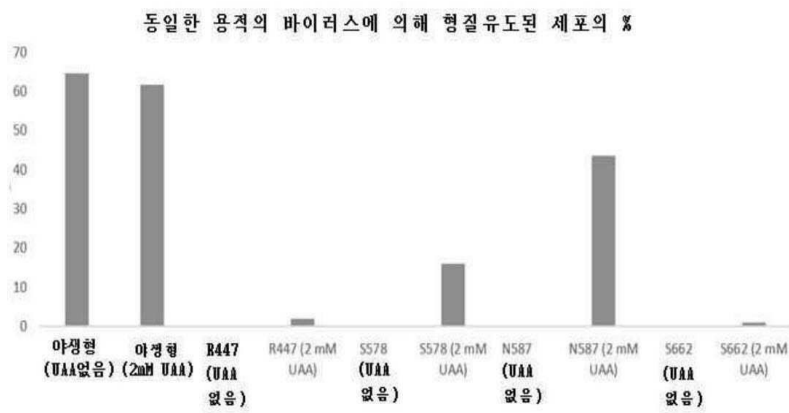
도면13a



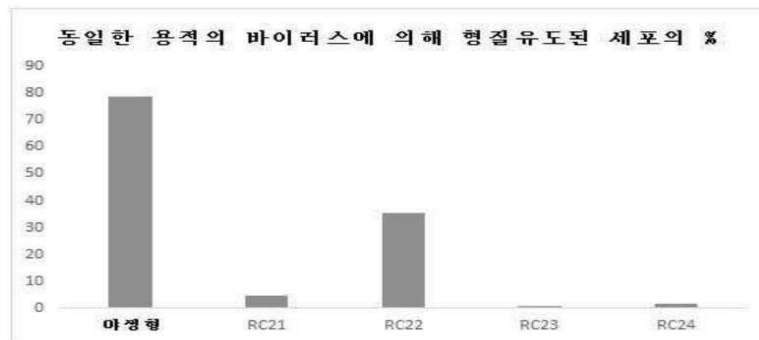
도면13b



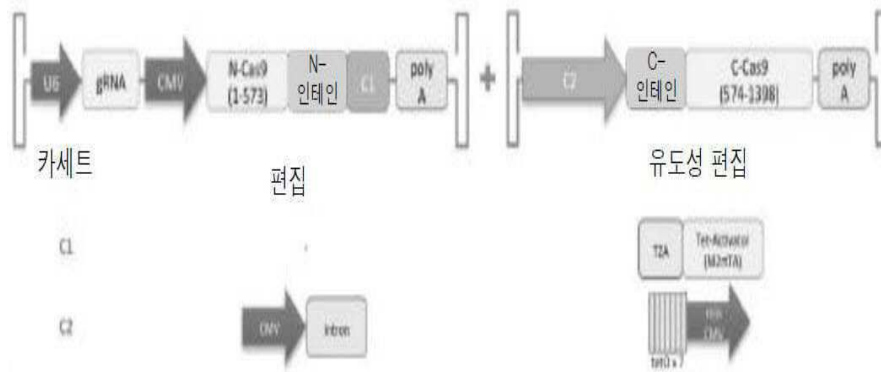
도면14



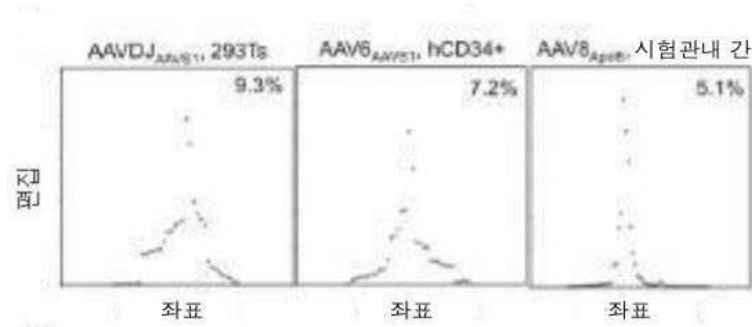
도면15



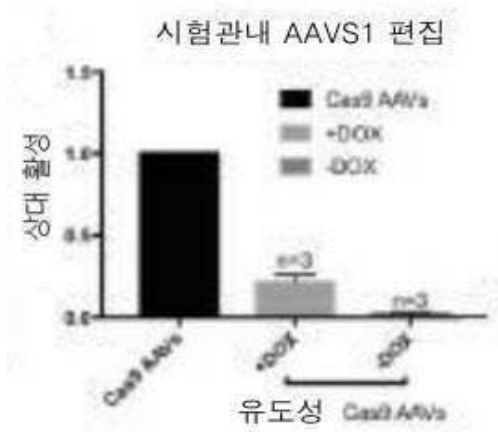
도면16a



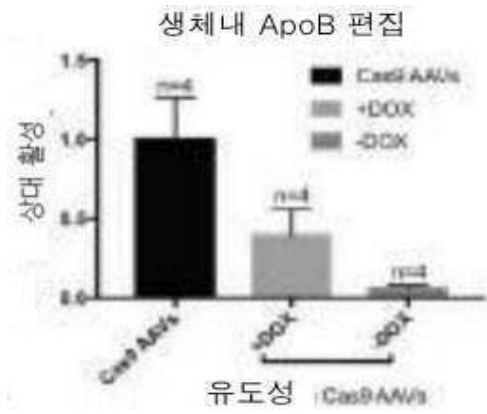
도면16b



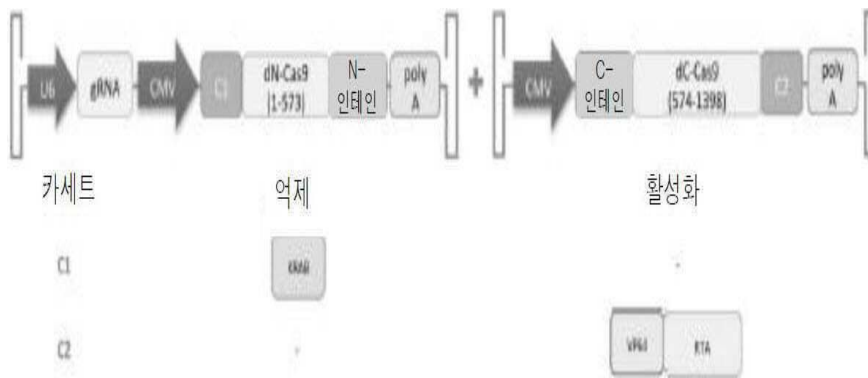
도면16c



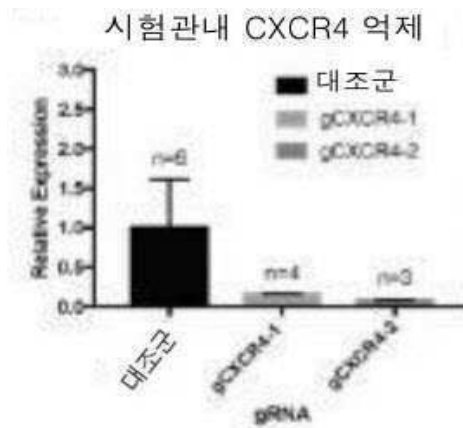
도면16d



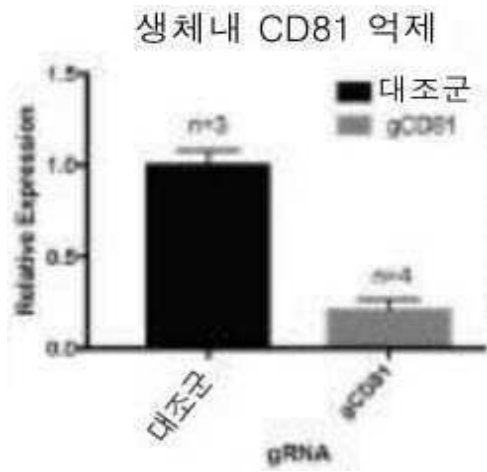
도면16e



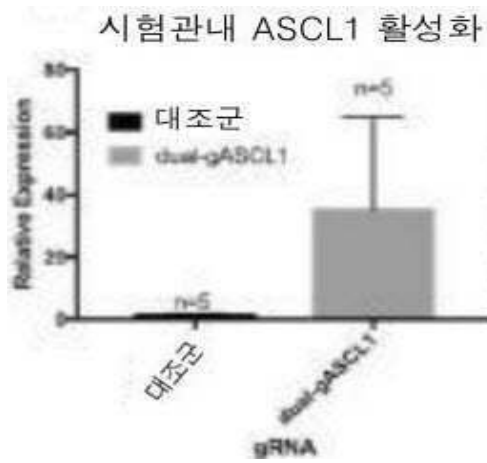
도면16f



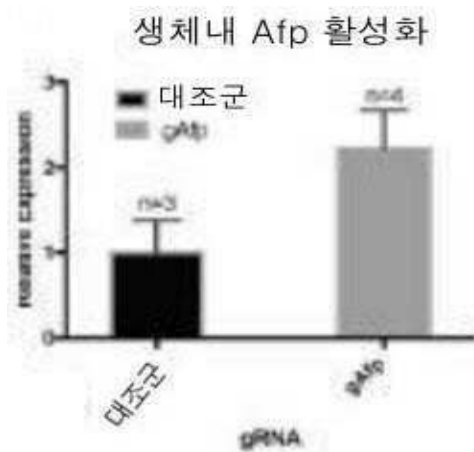
도면16g



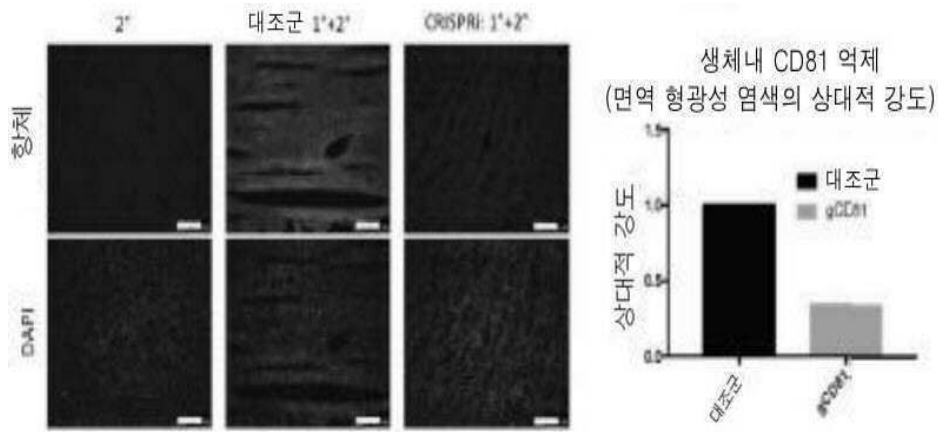
도면16h



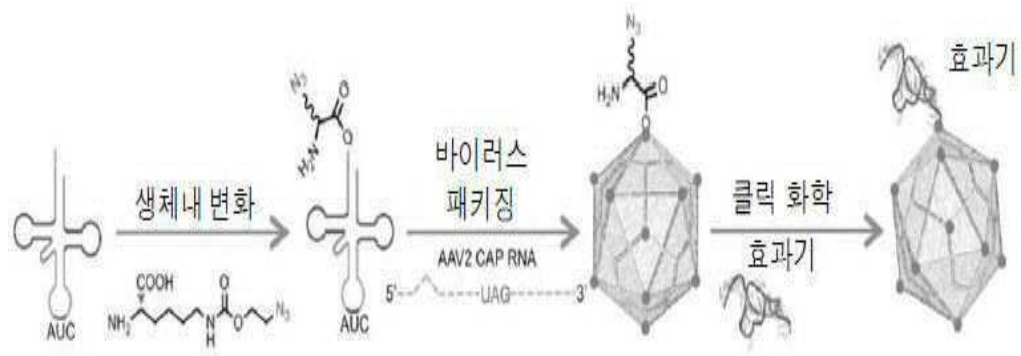
도면16i



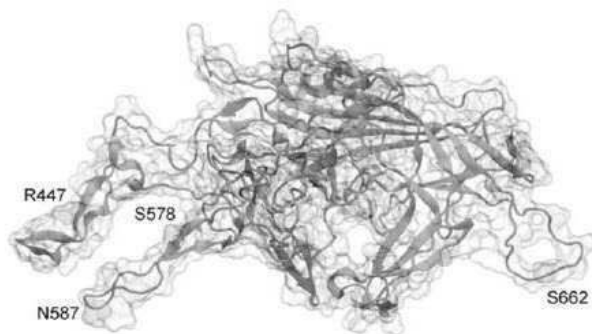
도면16j



도면17a

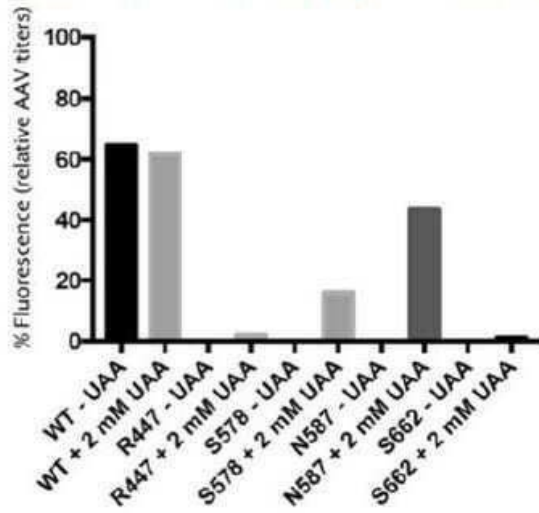


도면17b

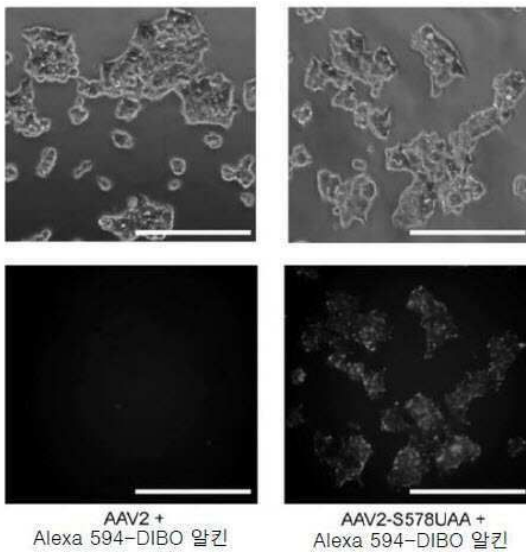


도면17c

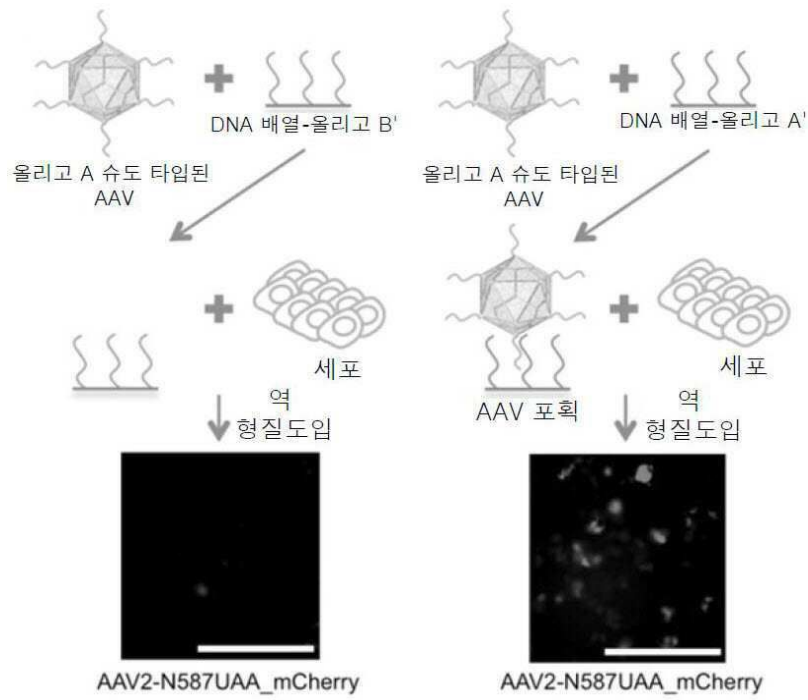
AAV 생산 동안 부위 특이적인 UAA 혼입



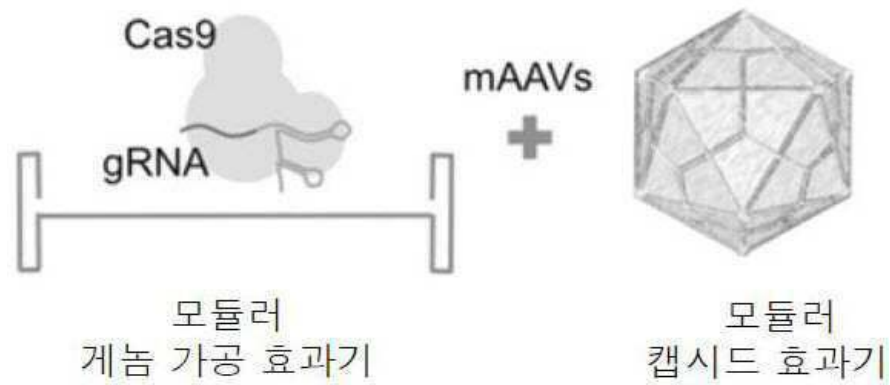
도면17d



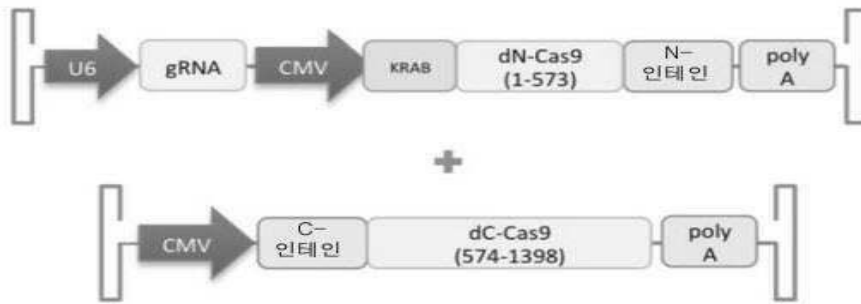
도면17e



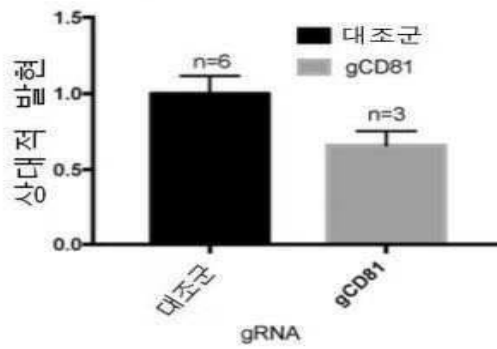
도면17f



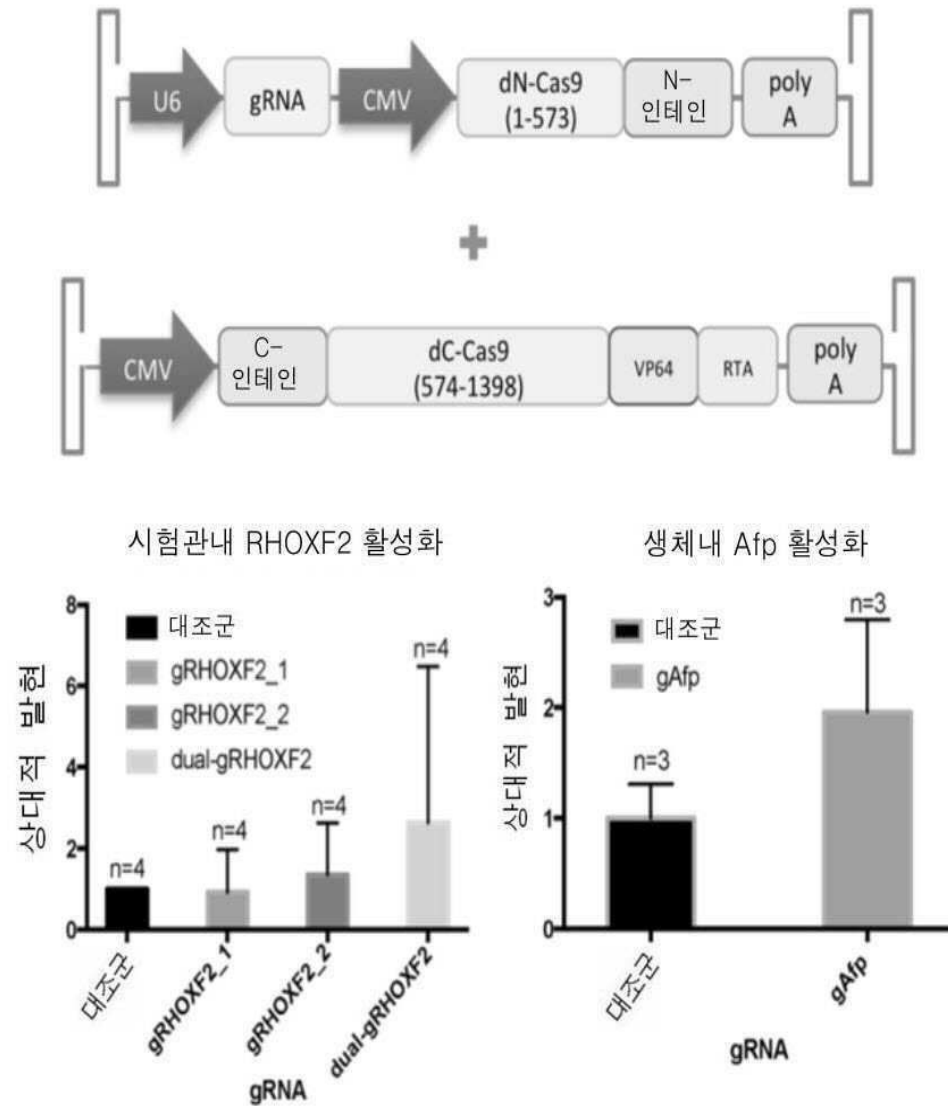
도면18b



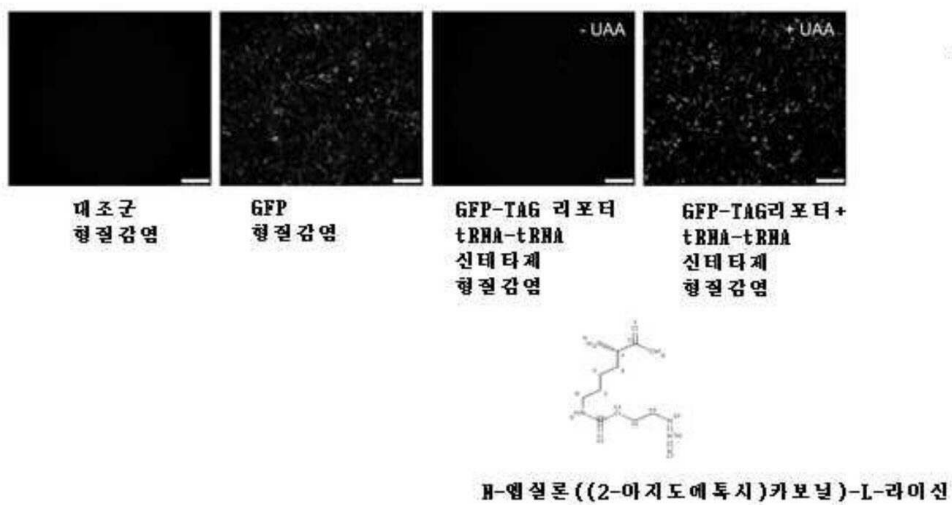
생체내 CD81 억제



도면18c

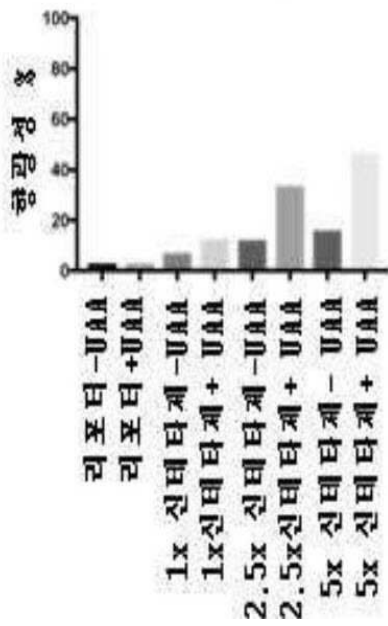


도면19a



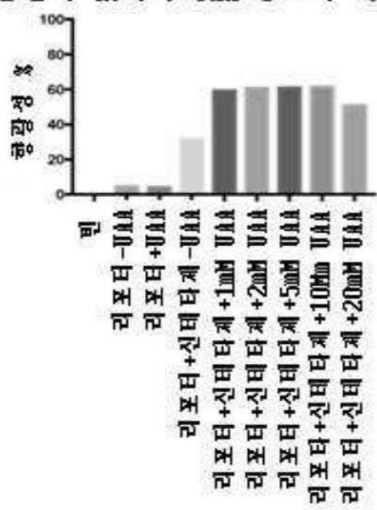
도면19b

UAA 혼입에 있어서 신테타제 양의 역할

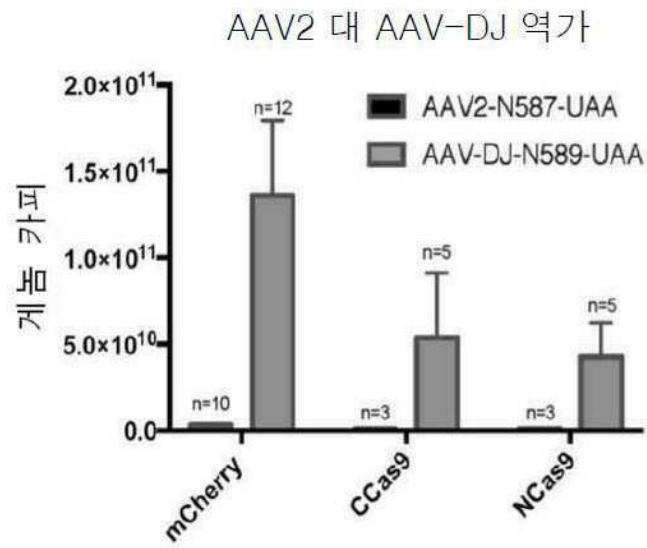


도면19c

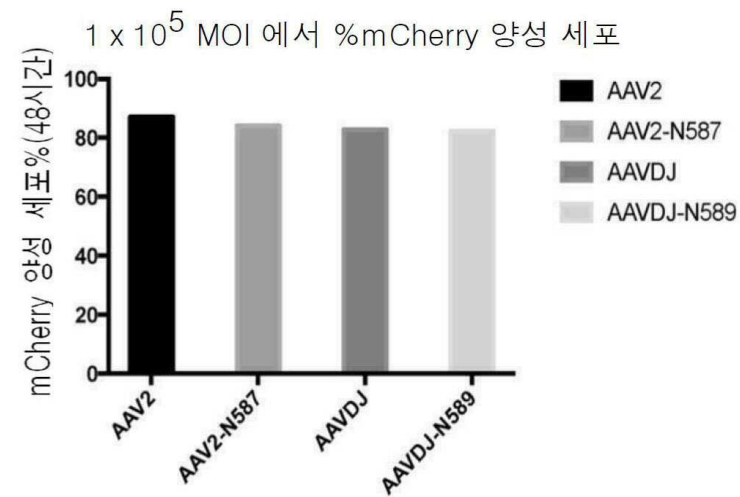
UAA 혼입에 있어서 UAA 농도의 변화



도면20a



도면20b

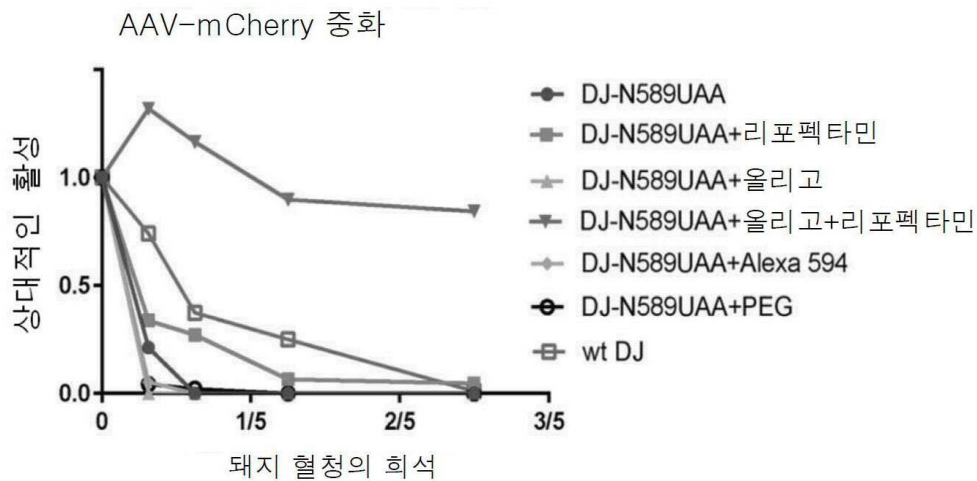


도면20c

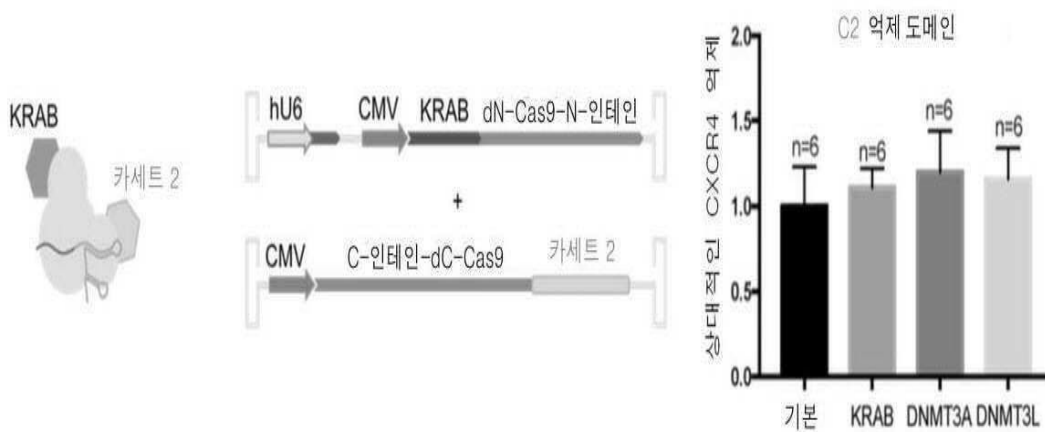


차폐 분자를 통해
항체 중화에 대한
mAAV 내성의 가공

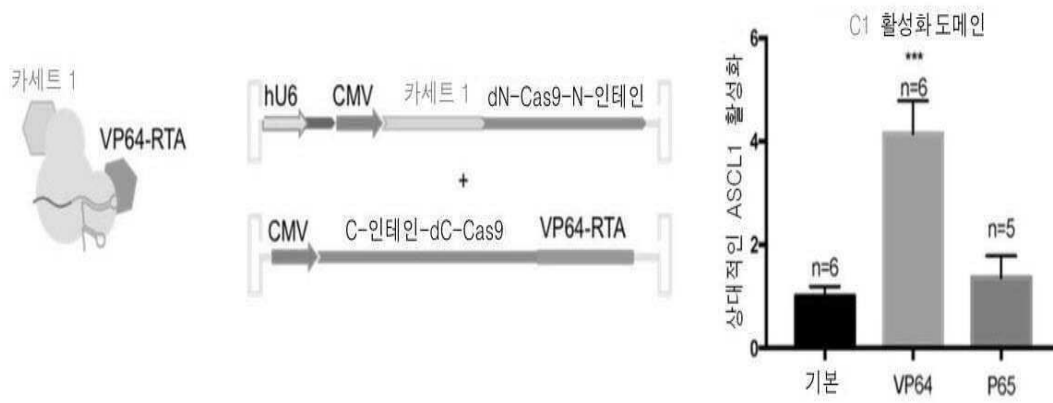
도면20d



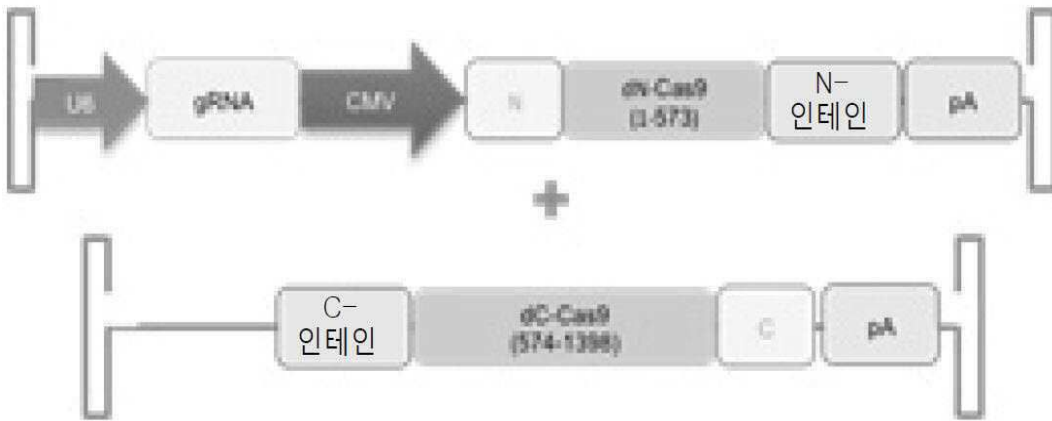
도면21a



도면21b



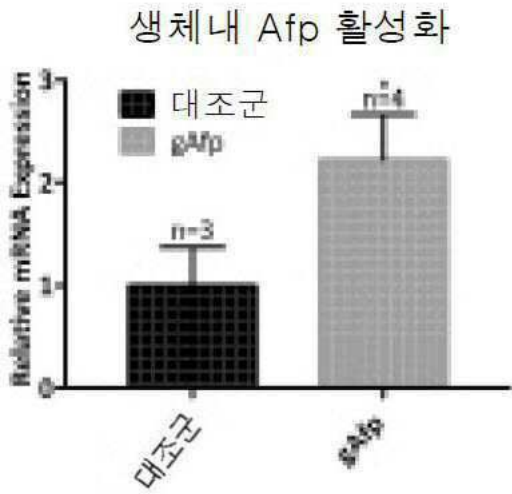
도면22a



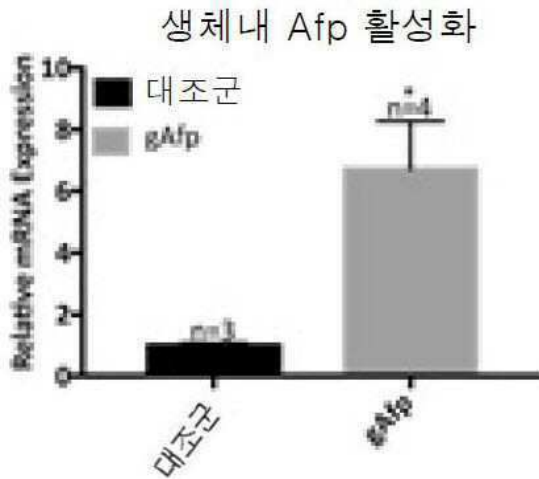
도면22b



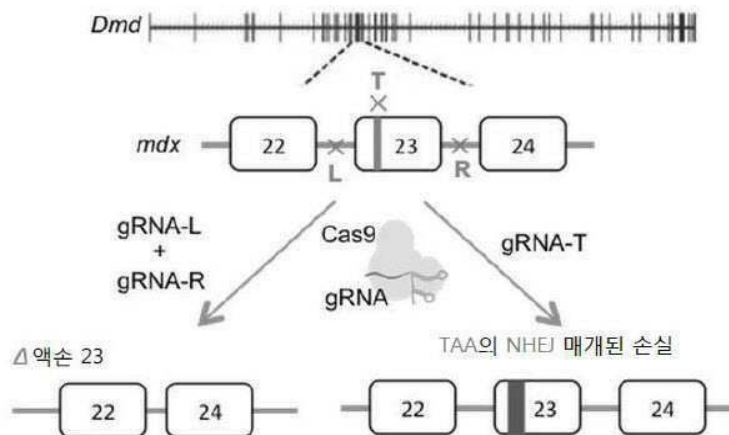
도면22c



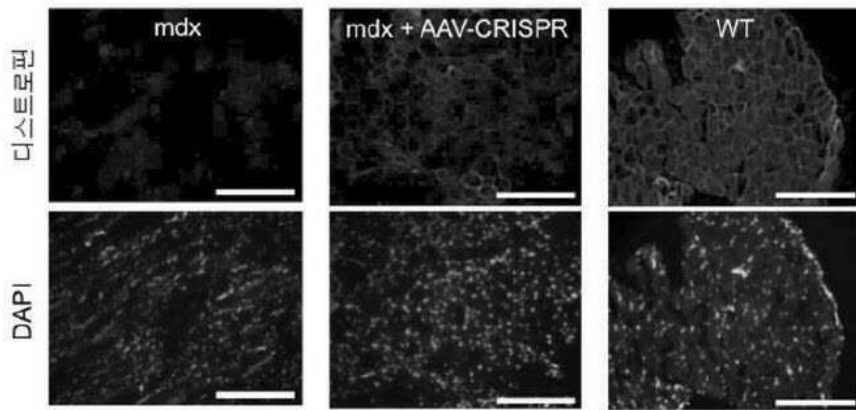
도면22d



도면23a



도면23b



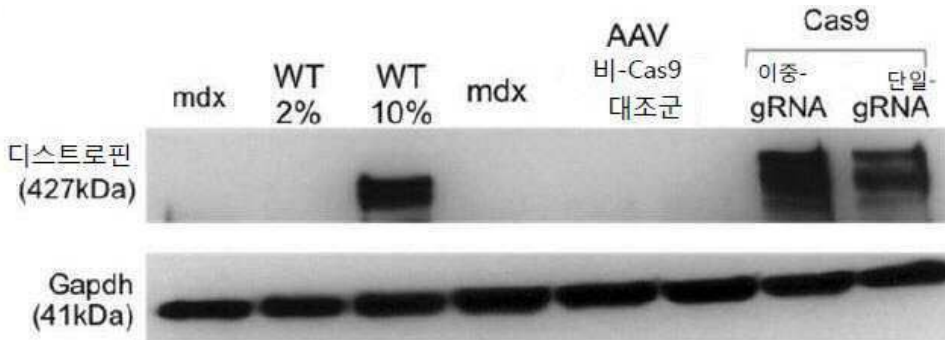
도면23c

Dmd gRNAs

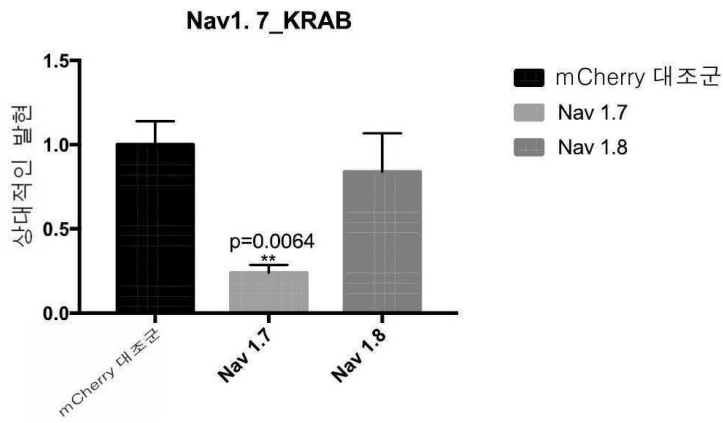
표적 서열

편집 gRNA-L	ataatttctattatattaca <u>agg</u>
편집 gRNA-R	ATTCAGgtaagccgaggtt <u>tgg</u>
편집 gRNA-T	TCTTTGAAAGAGCAATAAAA <u>TGG</u>

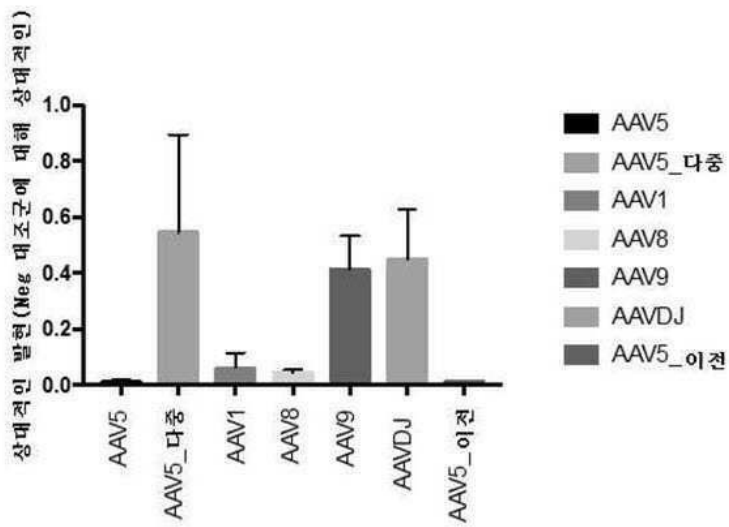
도면23d



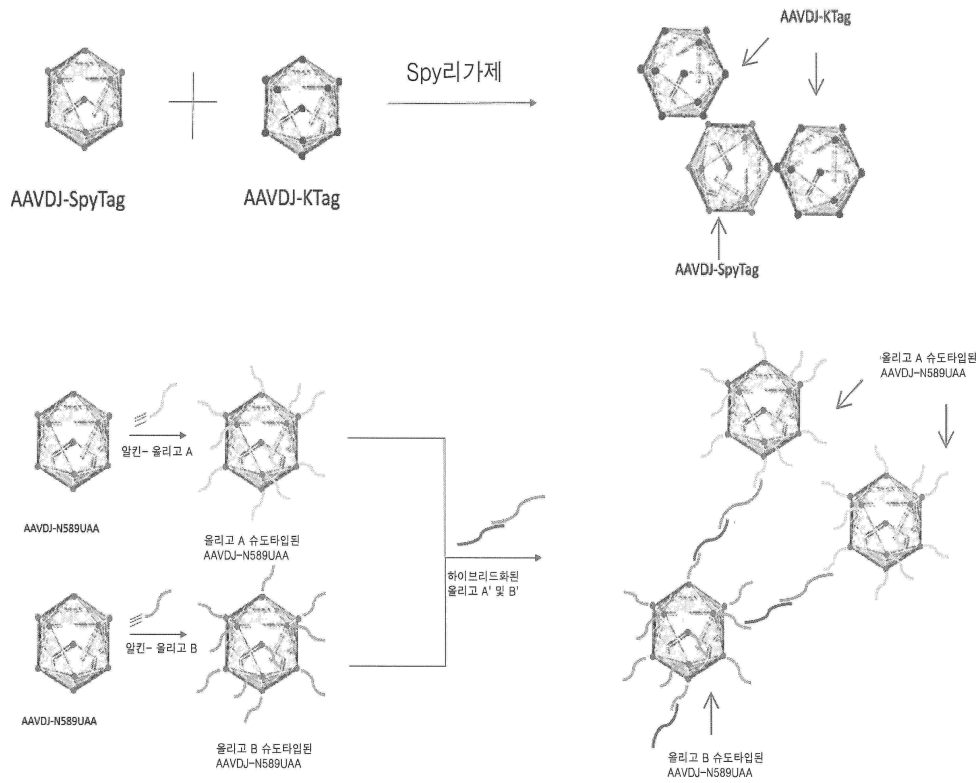
도면24



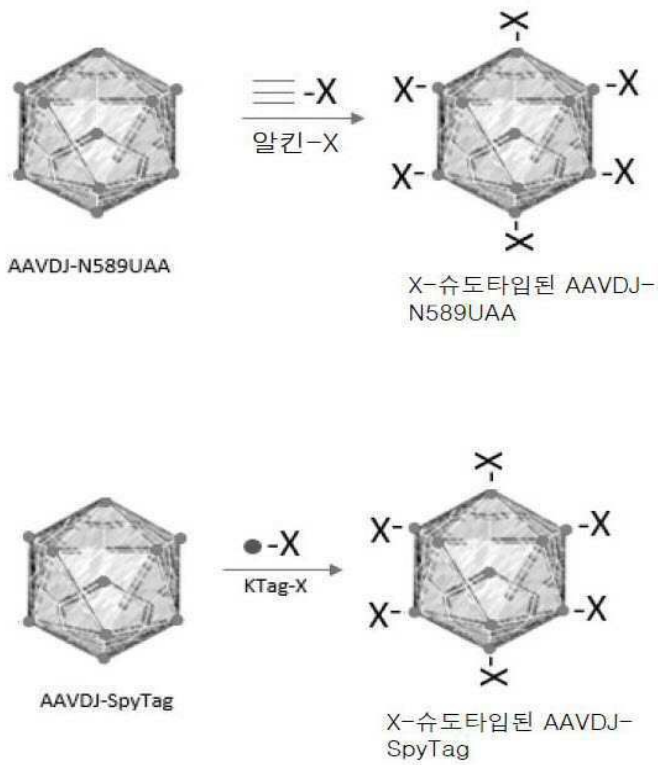
도면25



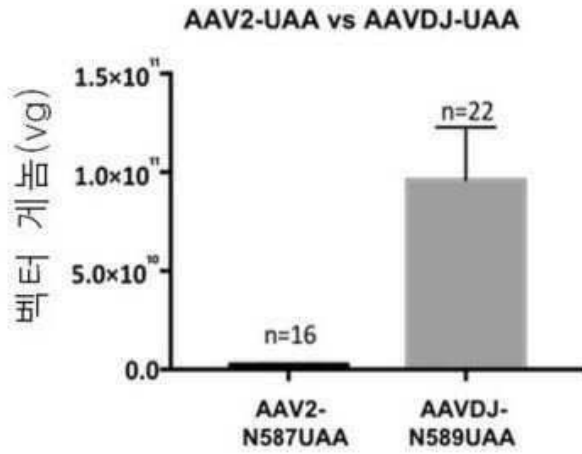
도면26



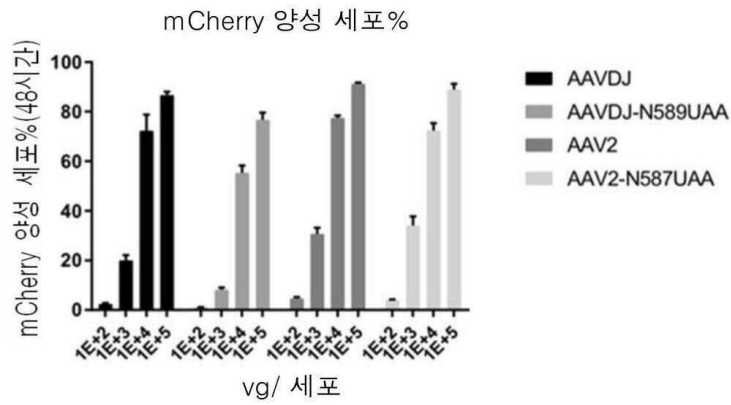
도면27



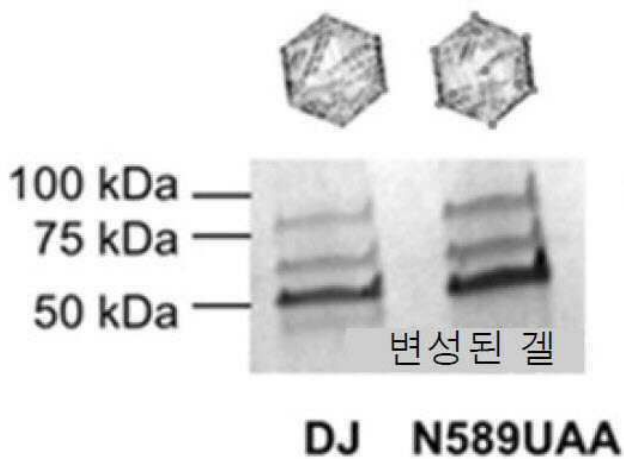
도면28a



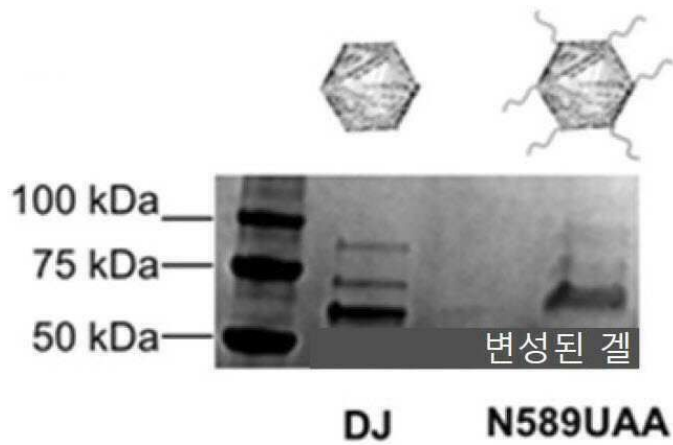
도면28b



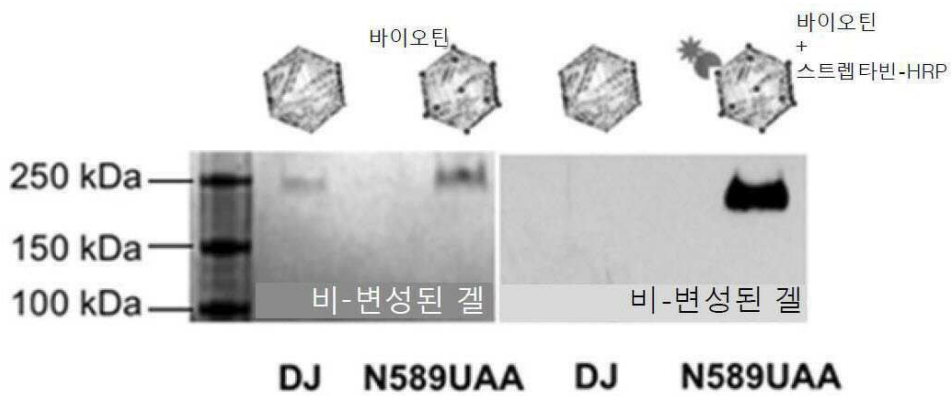
도면29a



도면29b



도면29c

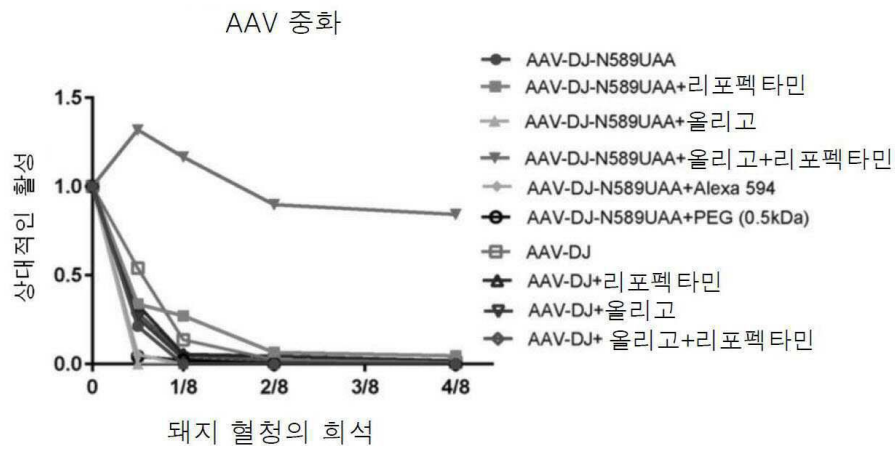


도면30a

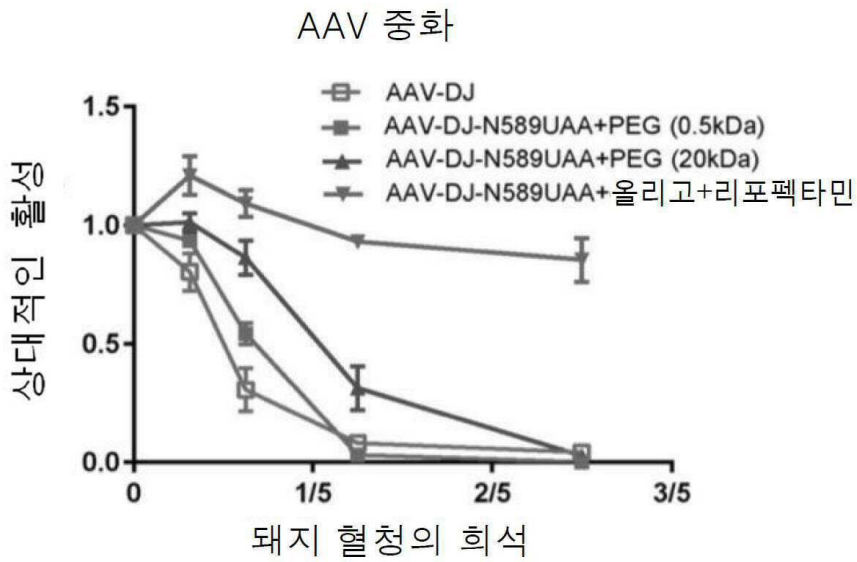


차폐 분자를 통해
항체 중화에 대한
mAAV 내성의 가공

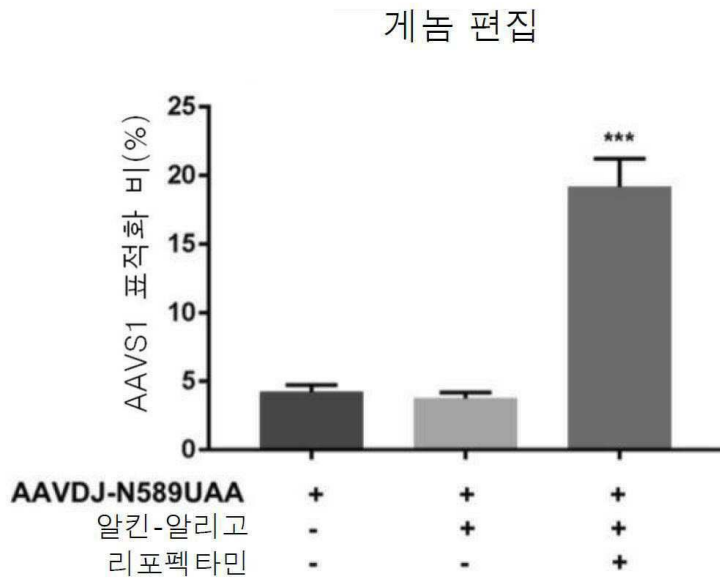
도면30b



도면30c

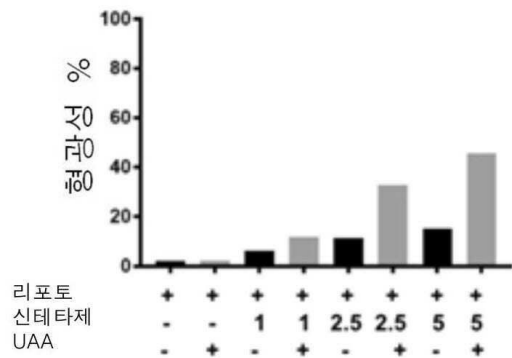


도면30d



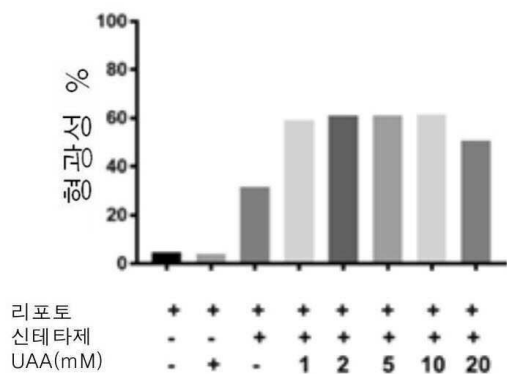
도면31a

UAA 혼입에 있어서 신테타제 양의 역할



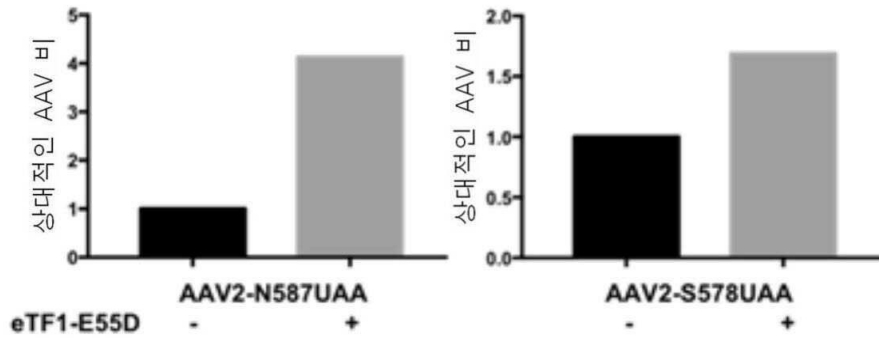
도면31b

UAA 혼입에 있어서 UAA 농도의 최적화



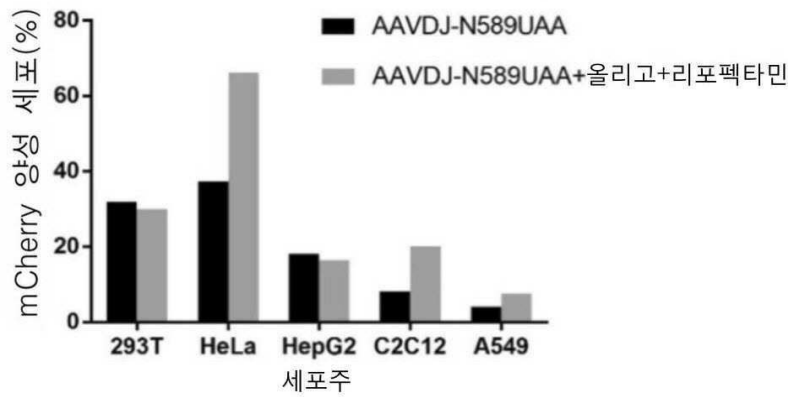
도면31c

eTF1-E55D은 UAA-AAV 생산을 향상시킨다

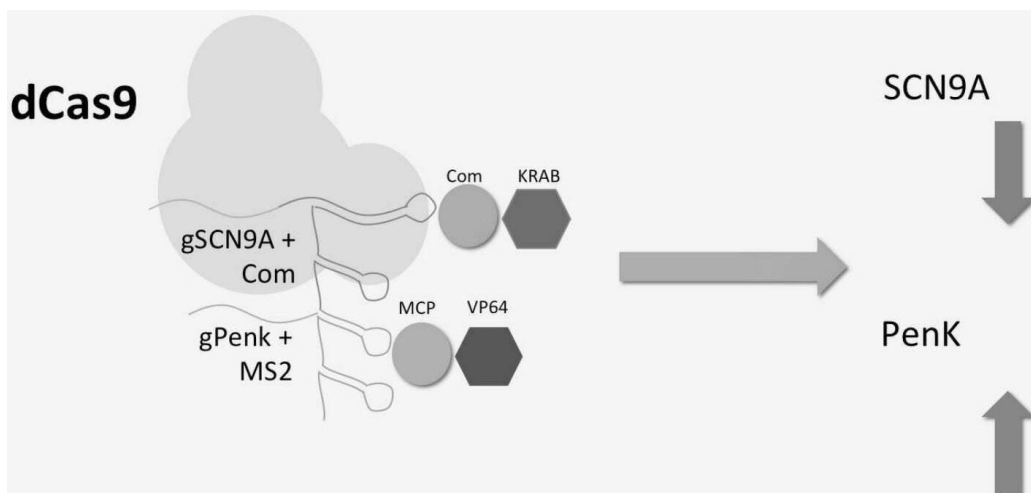


도면32

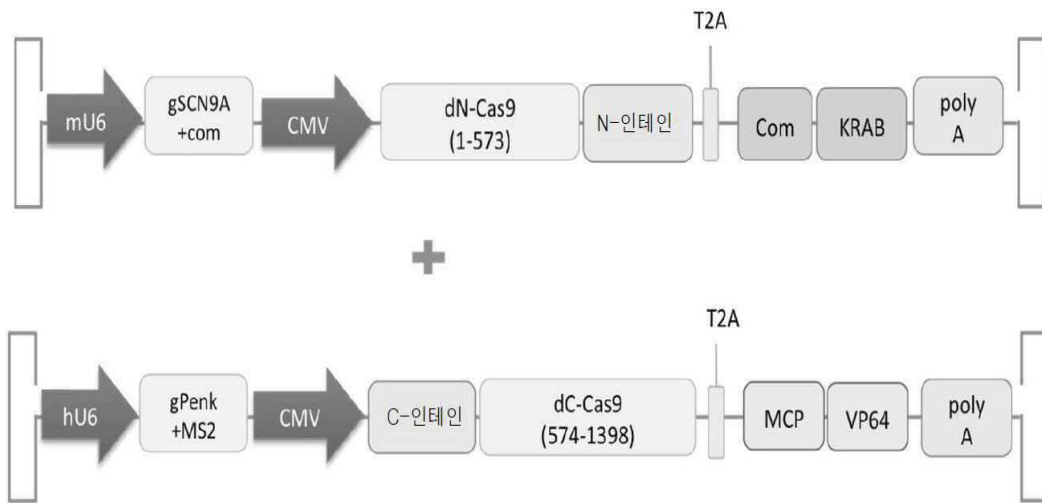
상대적인 형질유도 효능



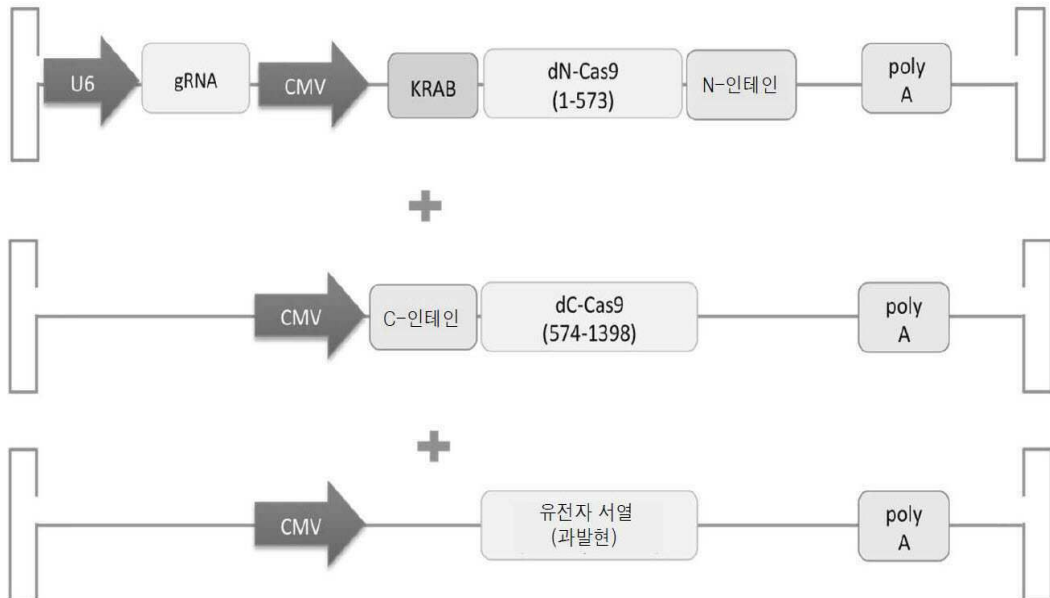
도면33



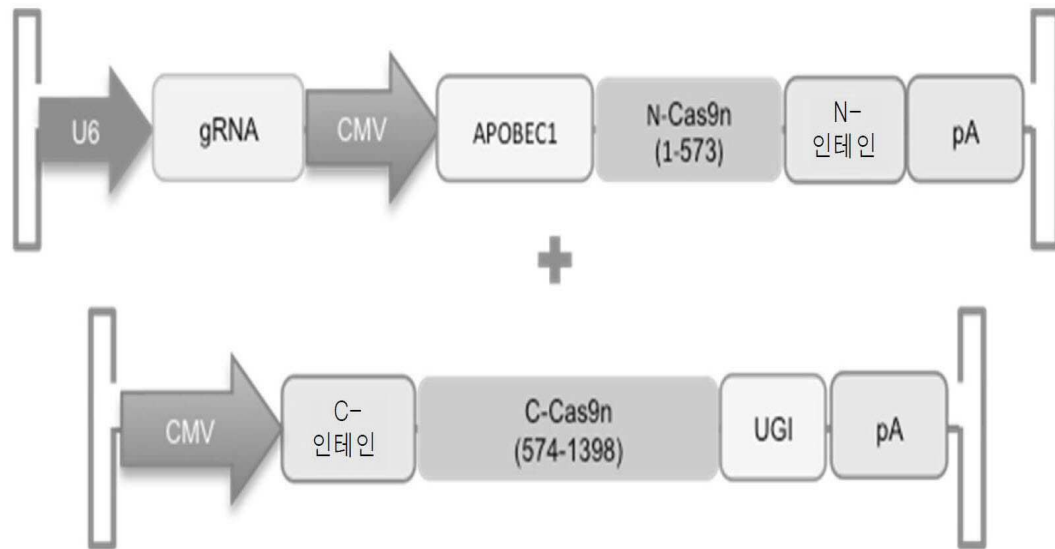
도면34



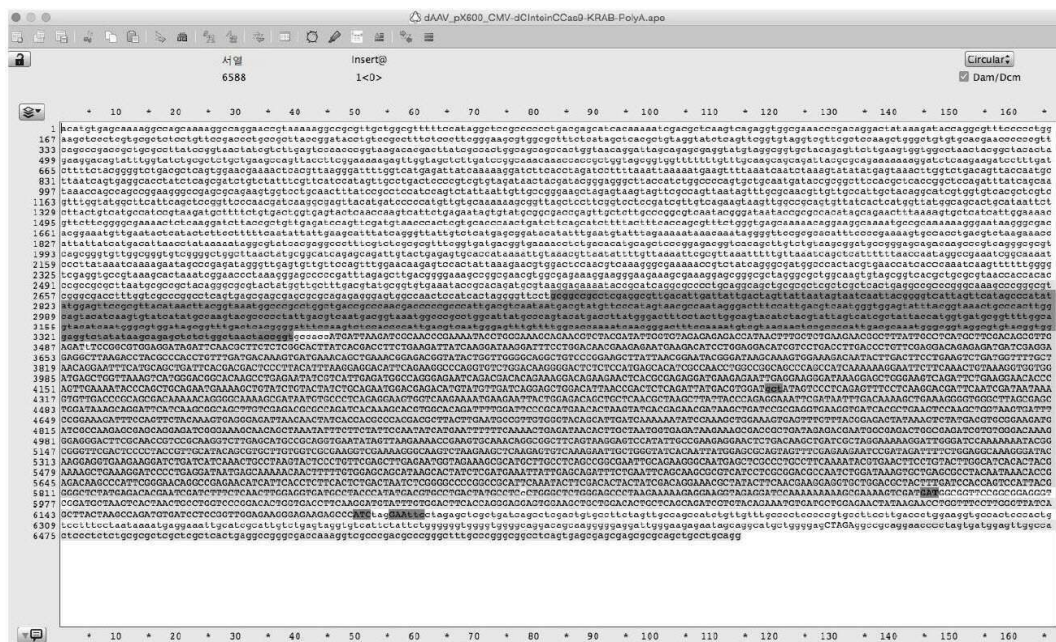
도면35



도면36



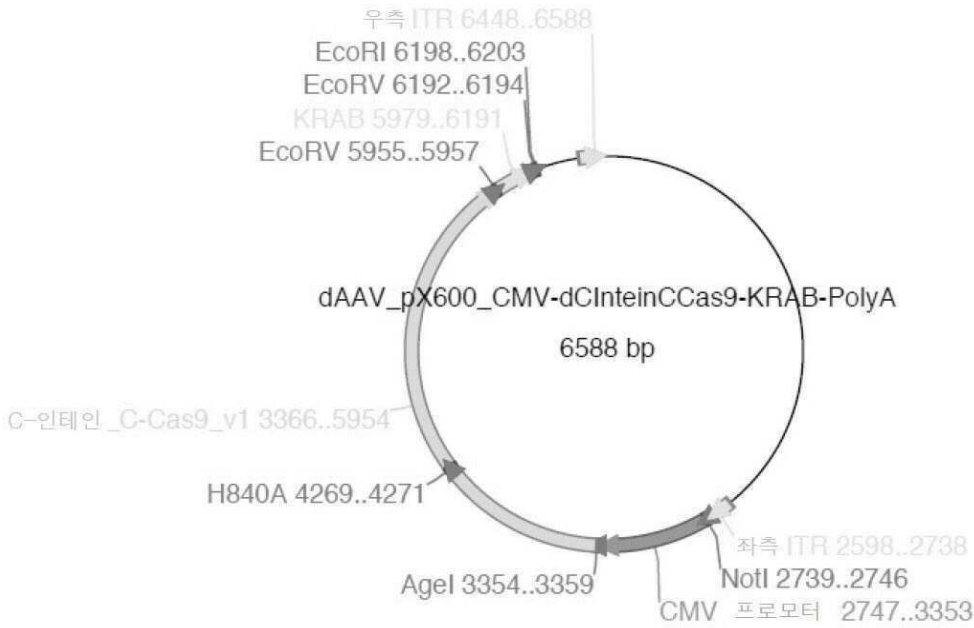
도면37a



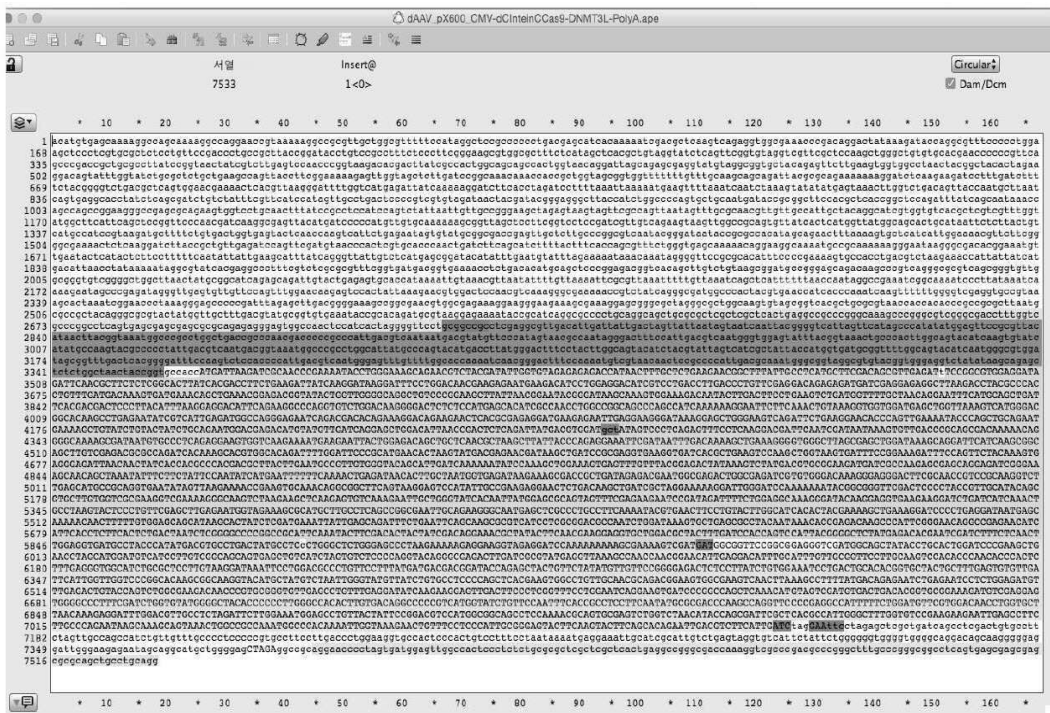
도면37b

Feature	Direction	Type	Location
작속 ITR	>>>	misc_feature	2598..2738
NotI	>>>	misc_feature	2739..2746
CMV 프로모터	>>>	misc_feature	2747..3353
Agel	>>>	misc_feature	3354..3359
C-인테인 C-Cas9_v1	>>>	misc_feature	3366..3490, 3492..4268, 4272..5887, 5889..5954
H940A	>>>	misc_feature	4269..4271
EcoRV	>>>	misc_feature	5955..5957
KRAB	>>>	misc_feature	5979..6191
EcoRV	>>>	misc_feature	6192..6194
EcoRI	>>>	misc_feature	6198..6203
우속 ITR	>>>	misc_feature	6448..6588

도면37c



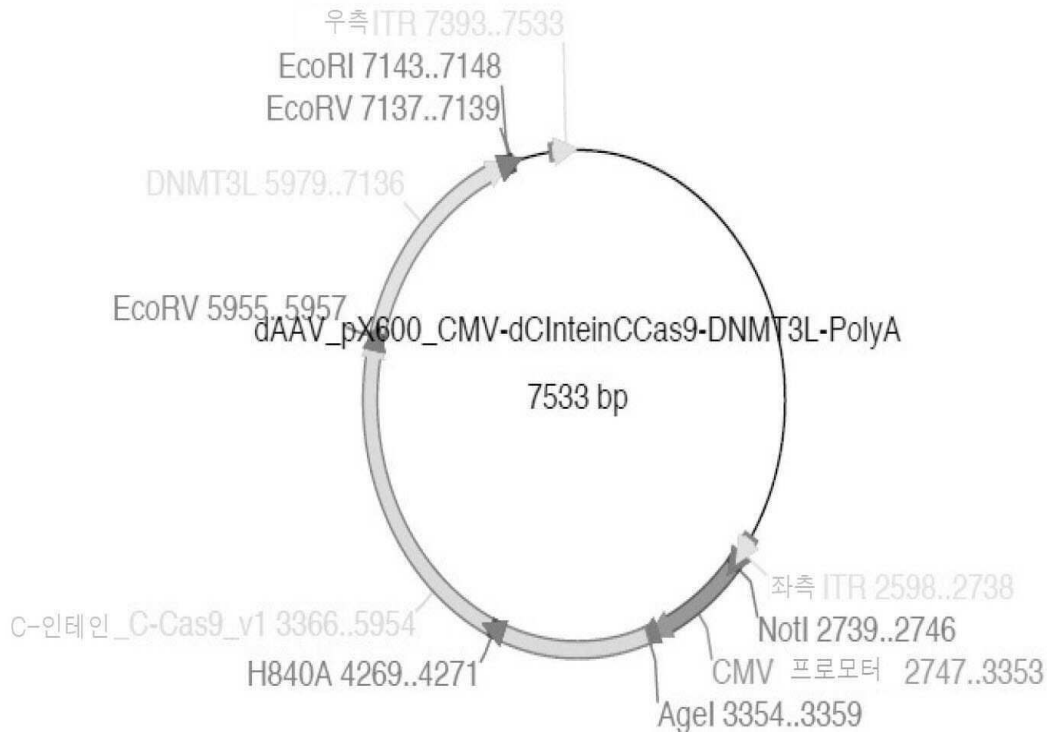
도면38a



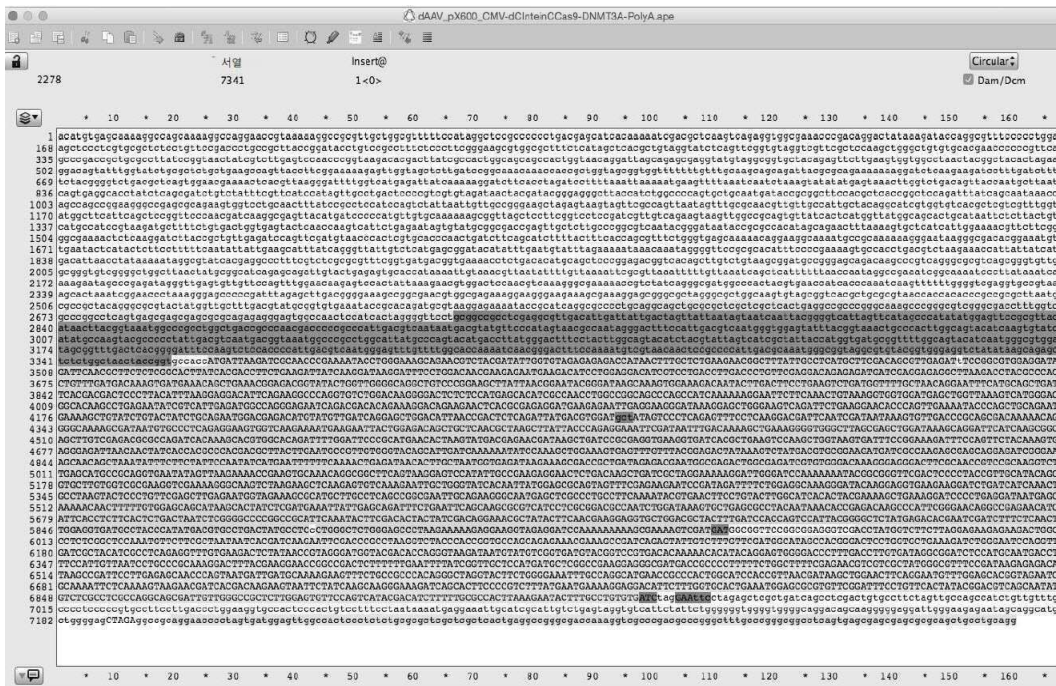
도면38b

특징	방향	유형	위치
좌측 ITR	>>>	misc_feature	2598..2738
NotI	>>>	misc_feature	2739..2746
CMV 프로모터	>>>	misc_feature	2747..3353
AgeI	>>>	misc_feature	3354..3359
C-인테인_C-Cas9_v1	>>>	misc_feature	3366..3490,3492..4268,4272..5887,5889..5954
H840A	>>>	misc_feature	4269..4271
EcoRV	>>>	misc_feature	5955..5957
DNMT3L	>>>	misc_feature	5979..7136
EcoRV	>>>	misc_feature	7137..7139
EcoRI	>>>	misc_feature	7143..7148
우측 ITR	>>>	misc_feature	7393..7533

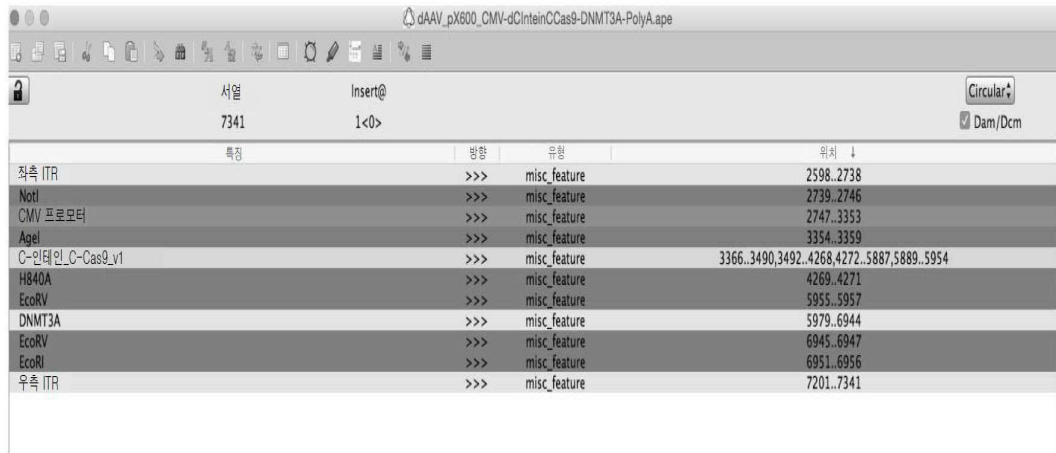
도면38c



도면39a

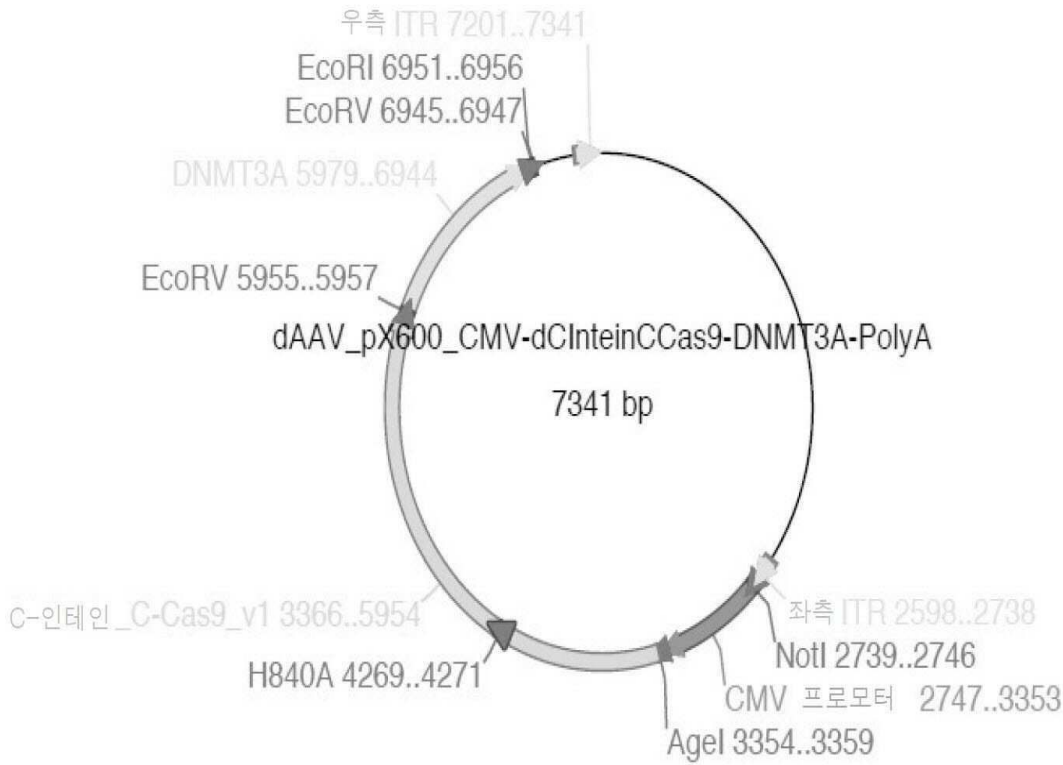


도면39b

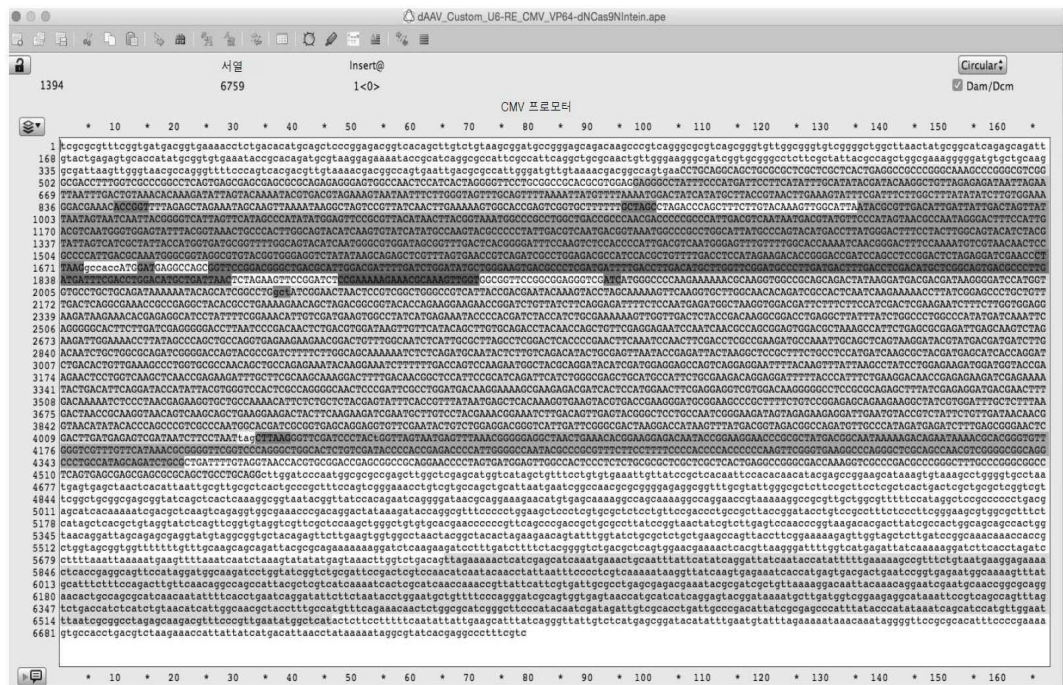


특징	방향	유형	위치 ↓
좌측 ITR	>>>	misc_feature	2598..2738
NotI	>>>	misc_feature	2739..2746
CMV 프로모터	>>>	misc_feature	2747..3353
AgeI	>>>	misc_feature	3354..3359
C-인테인_C-Cas9_v1	>>>	misc_feature	3366..3490,3492..4268,4272..5887,5889..5954
H840A	>>>	misc_feature	4269..4271
EcoRV	>>>	misc_feature	5955..5957
DNMT3A	>>>	misc_feature	5979..6944
EcoRV	>>>	misc_feature	6945..6947
EcoRI	>>>	misc_feature	6951..6956
우측 ITR	>>>	misc_feature	7201..7341

도면39c



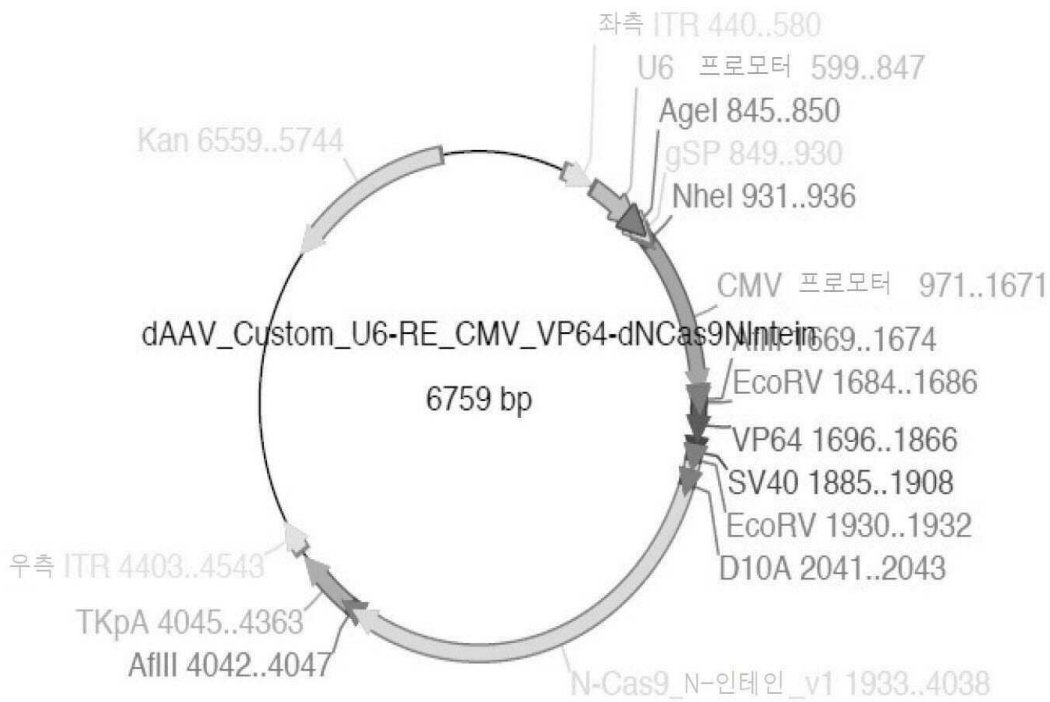
도면40a



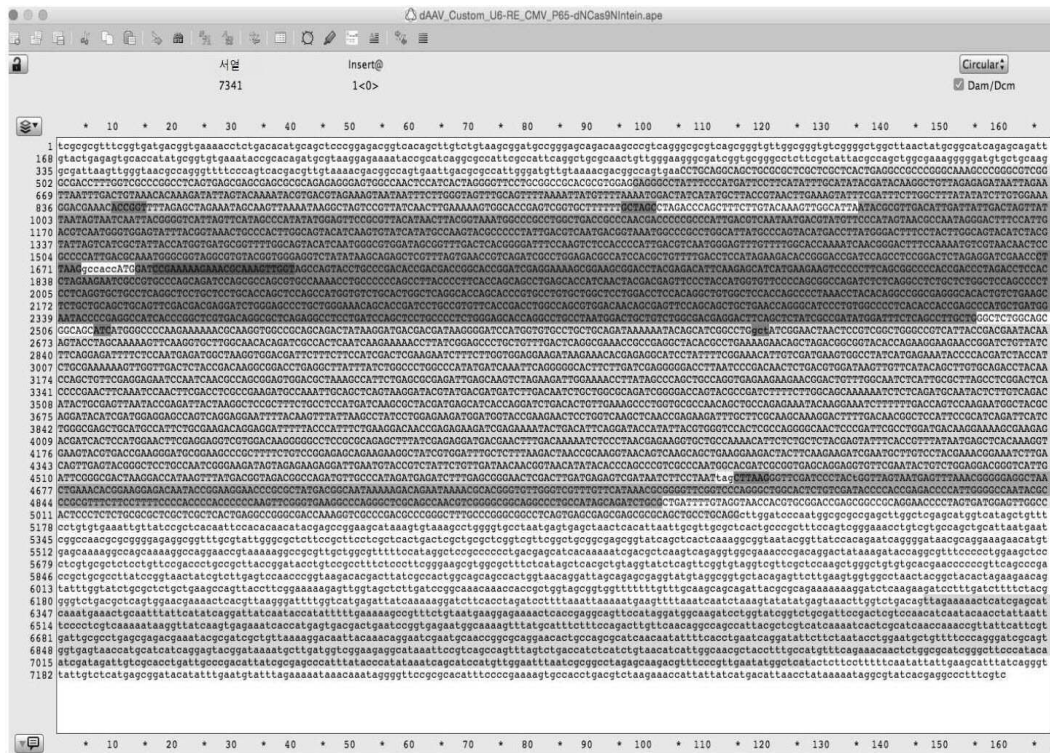
도면40b

971	서열 6759	출발 971<1>	길이 701<2>	말단 1671<2>	ORF ---	Tm ---	%GC 49%	Circular Dam/Dcm
특징	방향	유형	위치 ↓					
좌측 ITR	>>>	misc_feature	440..580					
U6 프로모터	>>>	misc_feature	599..847					
AgeI	>>>	misc_feature	845..850					
gSP	>>>	misc_feature	849..930					
NheI	>>>	misc_feature	931..936					
CMV 프로모터	>>>	misc_feature	971..1671					
AflII	>>>	misc_feature	1669..1674					
EcoRV	>>>	misc_feature	1684..1686					
VP64	>>>	misc_feature	1696..1866					
SV40	>>>	misc_feature	1885..1908					
EcoRV	>>>	misc_feature	1930..1932					
N-Cas9_N-인테인_v1	>>>	misc_feature	1933..2040,2044..4038					
D10A	>>>	misc_feature	2041..2043					
AflII	>>>	misc_feature	4042..4047					
TKpA	>>>	misc_feature	4045..4363					
우측 ITR	>>>	misc_feature	4403..4543					
Kan	<<<	misc_feature	5744..6559					

도면40c



도면41a

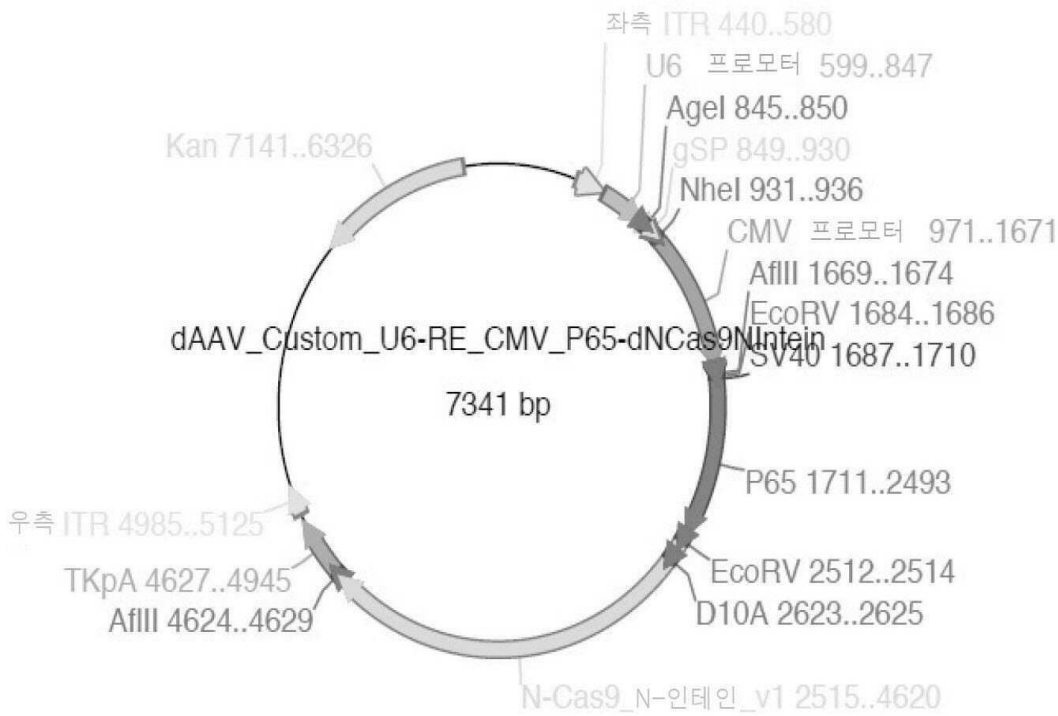


도면41b

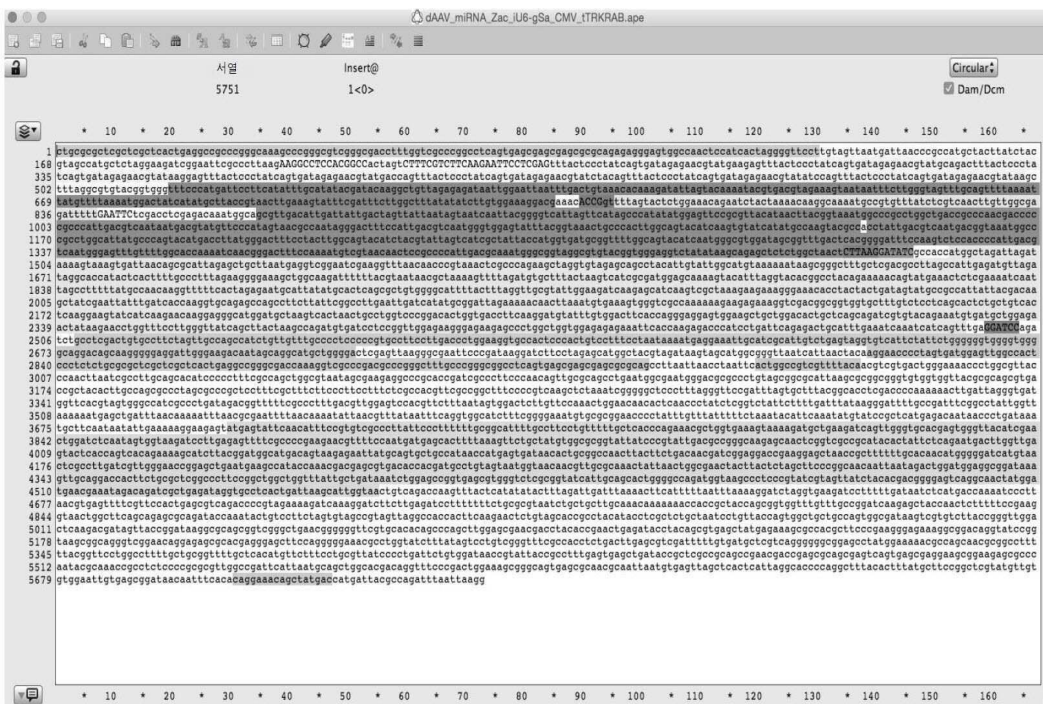
The figure shows a genomic map for the construct dAAV_Custom_U6-RE_CMV_P65-dNcas9NInfein ape, similar to 도면41a but with a focus on feature annotations. A table below the map lists various features with their start and end coordinates, orientation, and type. The features include promoters, restriction enzyme sites, and other genetic elements. The table is as follows:

특징	방향	유형	위치
좌측 ITR	>>>	misc_feature	440..580
U6 프로모터	>>>	misc_feature	599..847
Age1	>>>	misc_feature	845..850
gSP	>>>	misc_feature	849..930
NheI	>>>	misc_feature	931..936
CMV 프로모터	>>>	misc_feature	971..1671
AflII	>>>	misc_feature	1669..1674
EcoRV	>>>	misc_feature	1684..1686
SV40	>>>	misc_feature	1687..1710
P65	>>>	misc_feature	1711..2493
EcoRV	>>>	misc_feature	2512..2514
N-Cas9_N-인테인_v1	>>>	misc_feature	2515..2622.2626..4620
D10A	>>>	misc_feature	2623..2625
AflII	>>>	misc_feature	4624..4629
TKpA	>>>	misc_feature	4627..4945
우측 ITR	>>>	misc_feature	4985..5125
Kan	<<<	misc_feature	6326..7141

도면41c



도면42a



도면42b

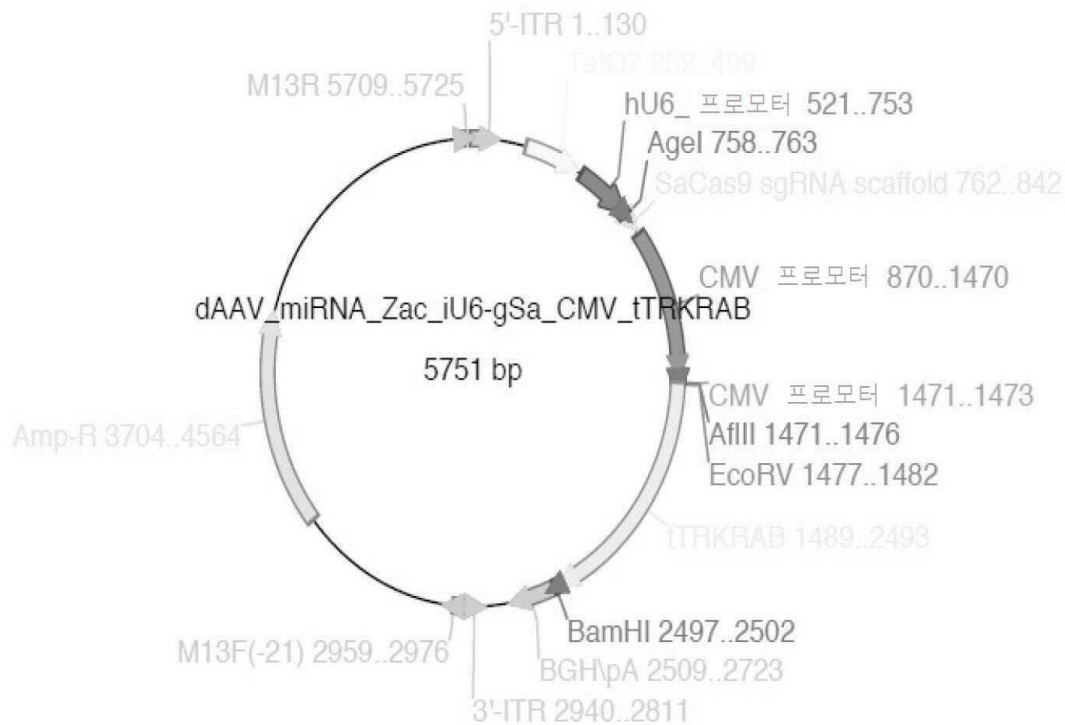
dAAV_miRNA_Zac_iU6-gSa_CMV_tTRKRAB.apc

서열: 5751, Insert@: 1<0>

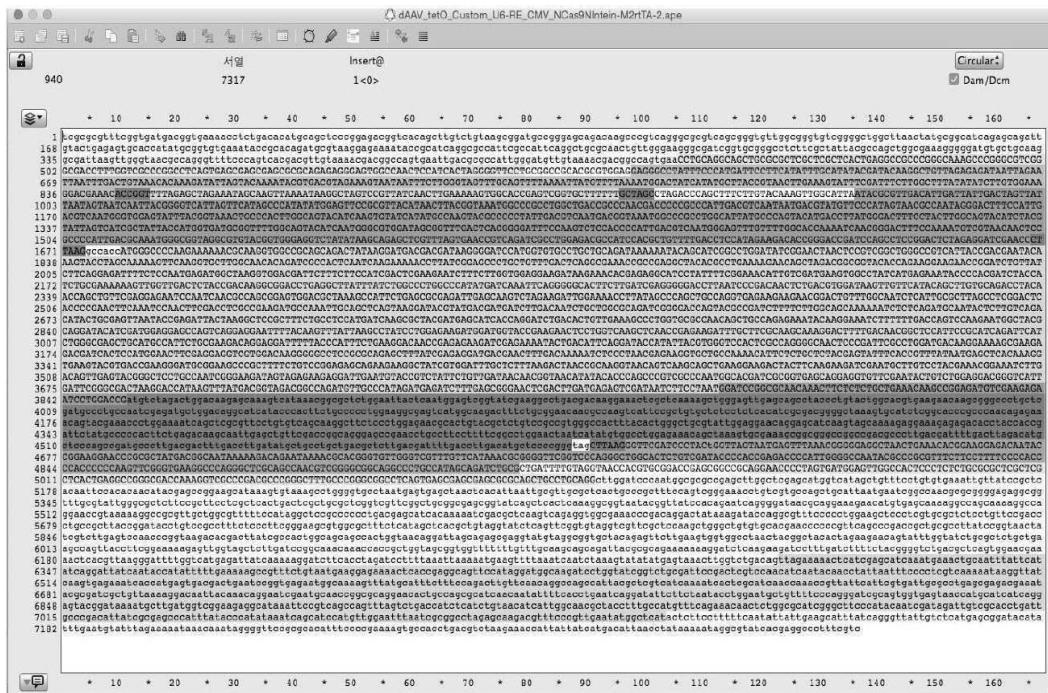
Circular: Dam/Dcm

특징	방향	유형	위치
5'-ITR	>>>	repeat_region	1..130
TetO7	>>>	promoter	252..499
hU6 프로모터	>>>	promoter	521..753
AgeI	>>>	misc_feature	758..763
SaCas9 sgRNA scaffold	>>>	misc_feature	762..842
CMV 프로모터	>>>	promoter	870..1139,1141..1470
CMV 프로모터	>>>	misc_feature	1471..1473
AflIII	>>>	misc_feature	1471..1476
EcoRV	>>>	misc_feature	1477..1482
tTRKRAB	>>>	misc_feature	1489..2493
BamHI	>>>	misc_feature	2497..2502
BGH1pA	>>>	polyA_signal	2509..2723
3'-ITR	<<<	repeat_region	2811..2940
M13F(-21)	>>>	misc_feature	2959..2976
Amp-R	>>>	CDS	3704..4564
M13R	>>>	misc_feature	5709..5725

도면42c



도면43a

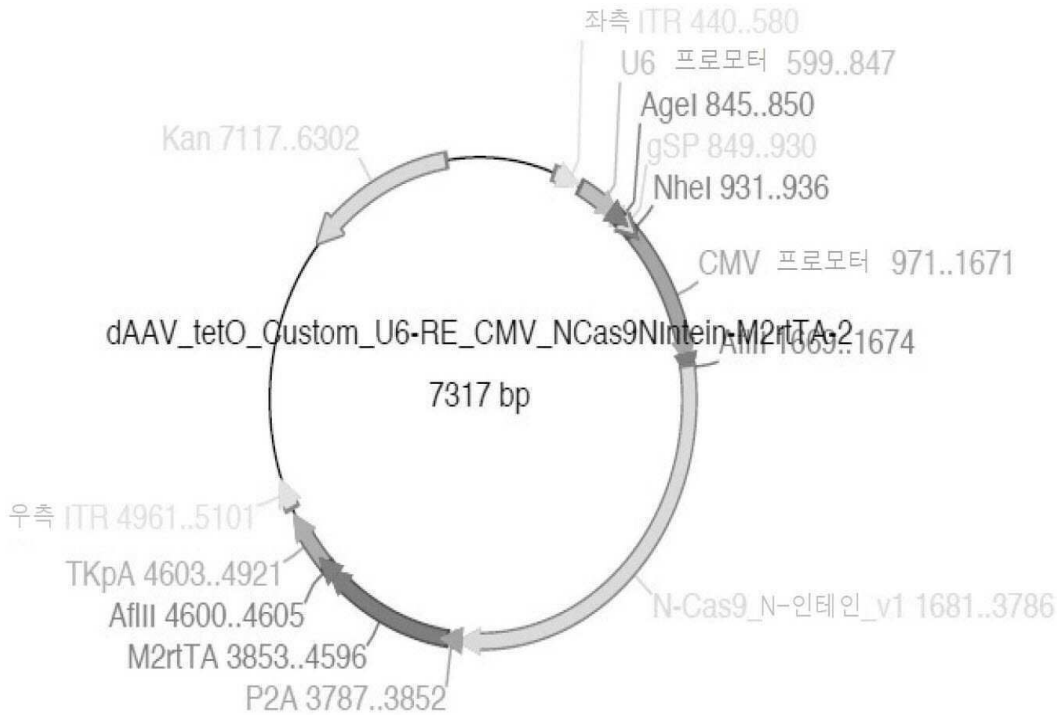


도면43b

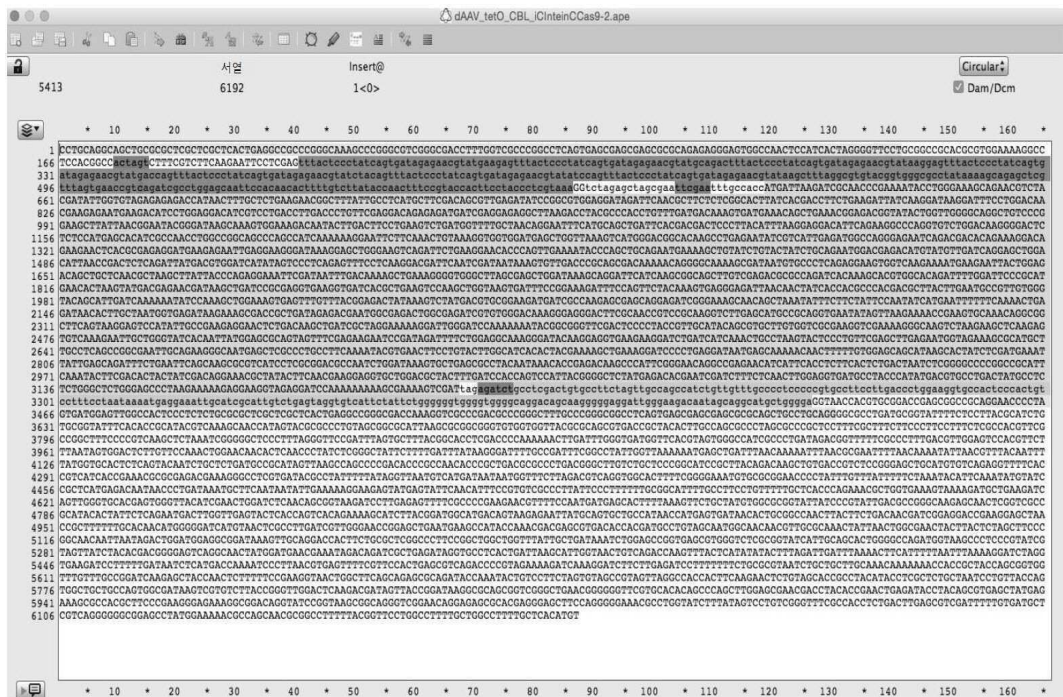
번호	서열	출발	길이	말단	ORF	Tm	%GC
971	7317	971<1>	701<2>	1671<2>	---	---	49%

특징	방향	유형	위치
지속 ITR	>>>	misc_feature	440..580
U6 프로모터	>>>	misc_feature	599..847
Agel	>>>	misc_feature	845..850
gSP	>>>	misc_feature	849..930
NheI	>>>	misc_feature	931..936
CMV 프로모터	>>>	misc_feature	971..1671
AflIII	>>>	misc_feature	1669..1674
N-Cas9_N-인덴인_v1	>>>	misc_feature	1681..3786
PZA	>>>	misc_feature	3787..3852
M2rTA	>>>	misc_feature	3853..4596
AflIII	>>>	misc_feature	4600..4605
TKpA	>>>	misc_feature	4603..4921
우측 ITR	>>>	misc_feature	4961..5101
Kan	<<<	misc_feature	6302..7117

도면43c



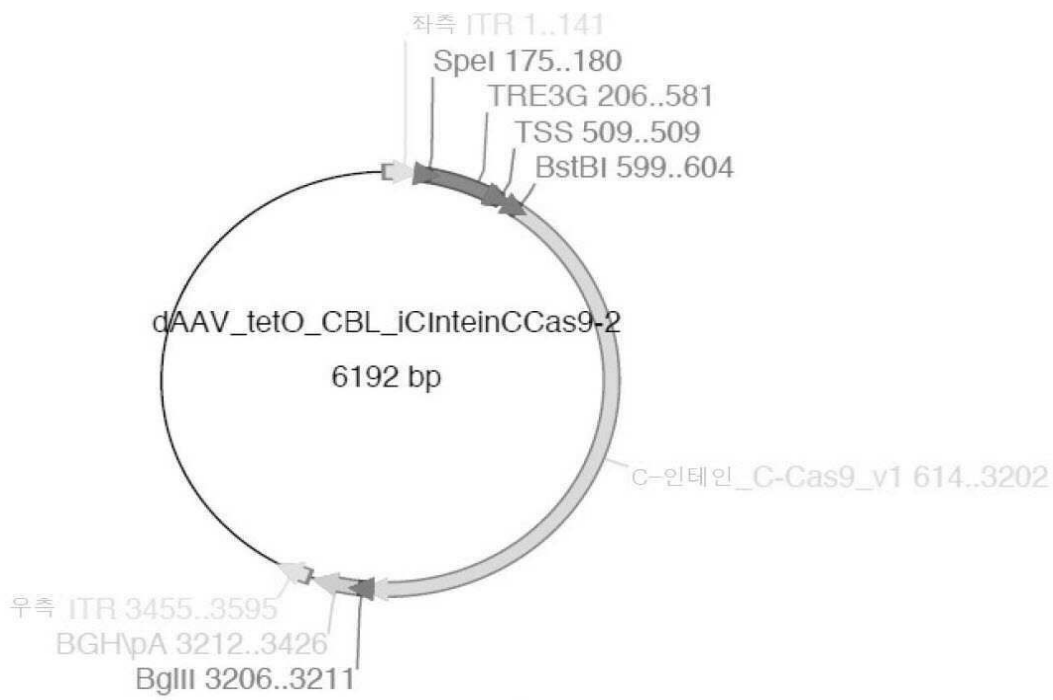
도면44a



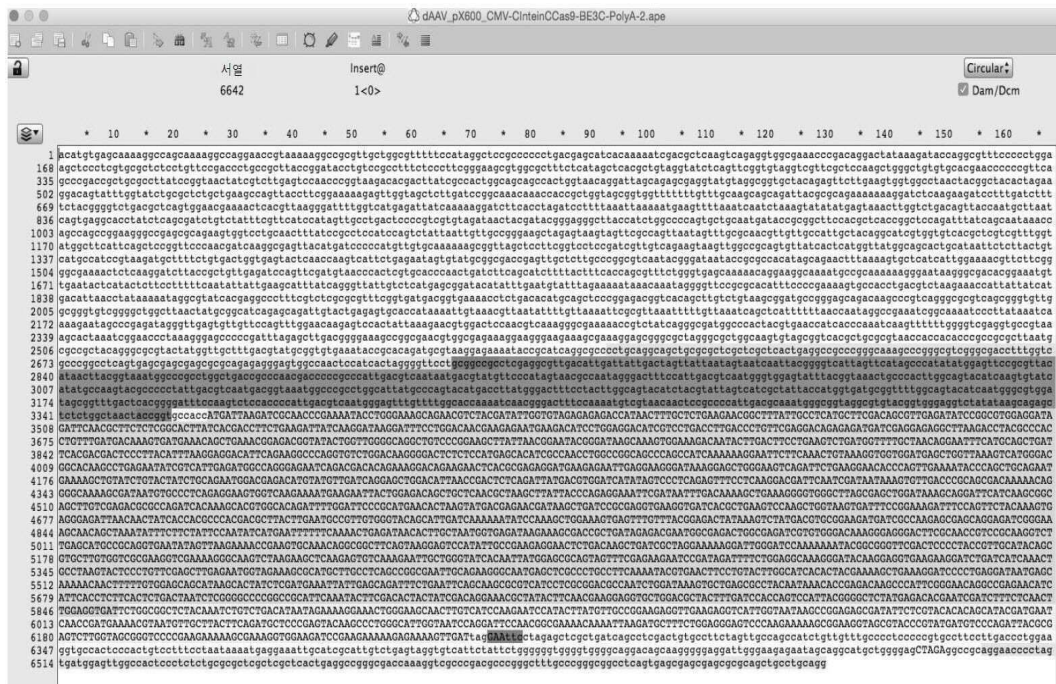
도면44b

특징	방향	유형	위치 ↓
좌측 ITR	>>>	misc_feature	1..141
SpeI	>>>	misc_feature	175..180
TRE3G	>>>	promoter	206..581
TSS	>>>	misc_feature	509..509
BstBI	>>>	misc_feature	599..604
C-인테인_C-Cas9_v1	>>>	misc_feature	614..3202
BglII	>>>	misc_feature	3206..3211
BGHpA	>>>	polyA_signal	3212..3426
우측 ITR	>>>	misc_feature	3455..3595

도면44c



도면45a

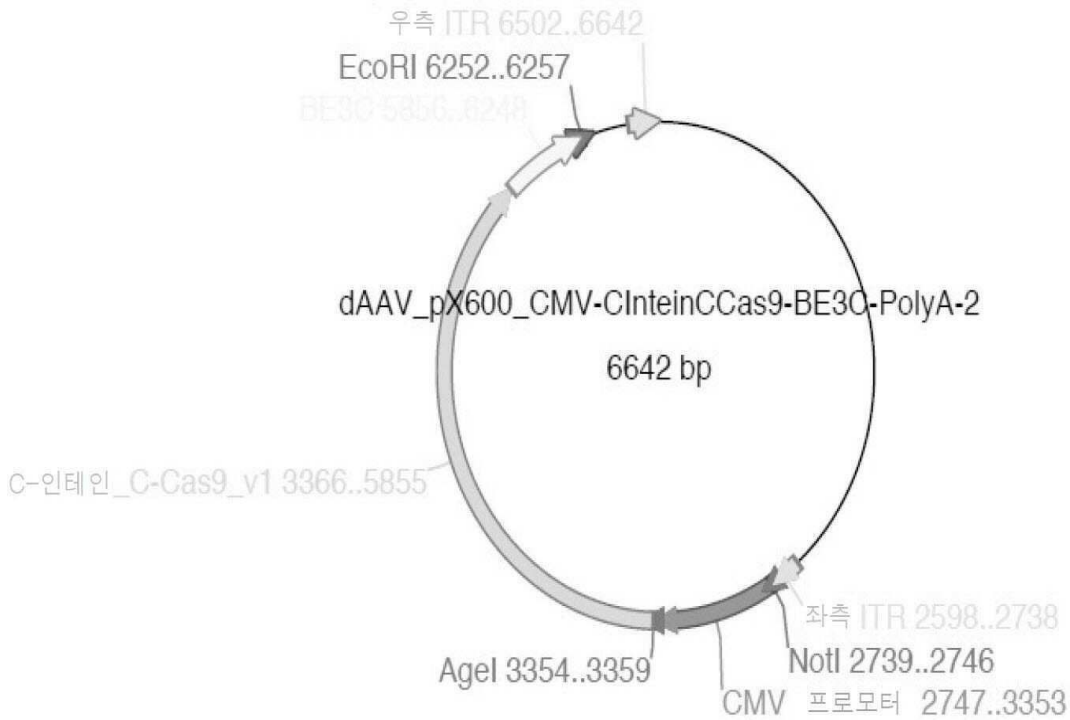


도면45b

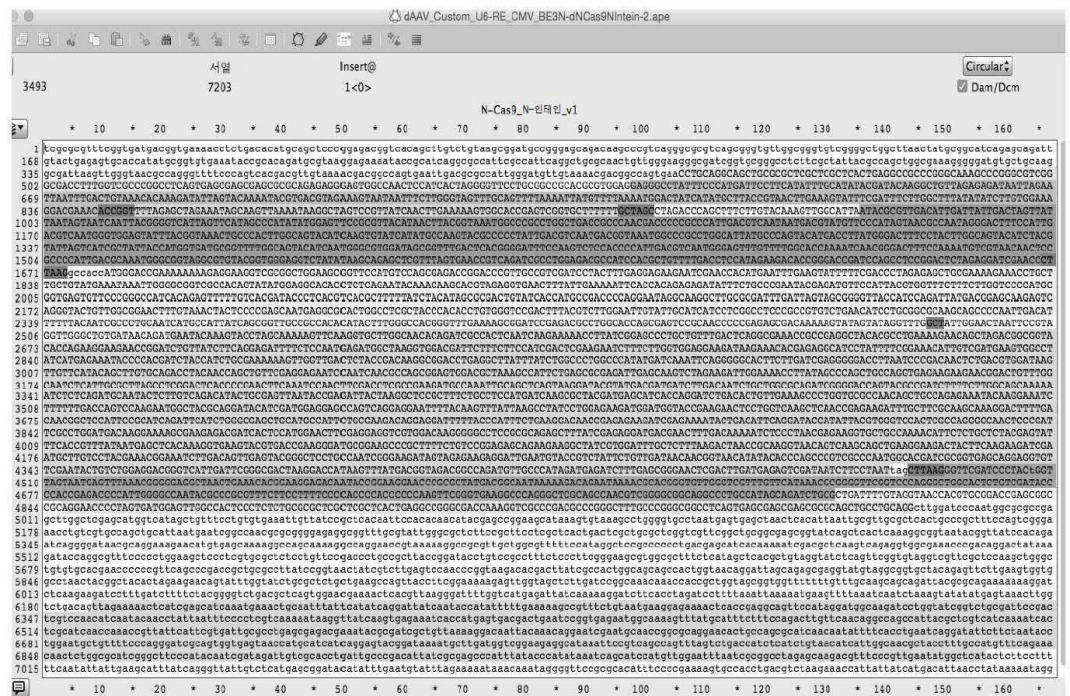
Annotation table for the sequence shown in 도면45a.

특징	방향	유형	위치
좌측 ITR	>>>	misc_feature	2598..2738
NotI	>>>	misc_feature	2739..2746
CMV 프로모터	>>>	misc_feature	2747..3353
AgeI	>>>	misc_feature	3354..3359
C-인테인 C-Cas9_v1	>>>	misc_feature	3366..5855
BE3C	>>>	misc_feature	5856..6248
EcoRI	>>>	misc_feature	6252..6257
우측 ITR	>>>	misc_feature	6502..6642

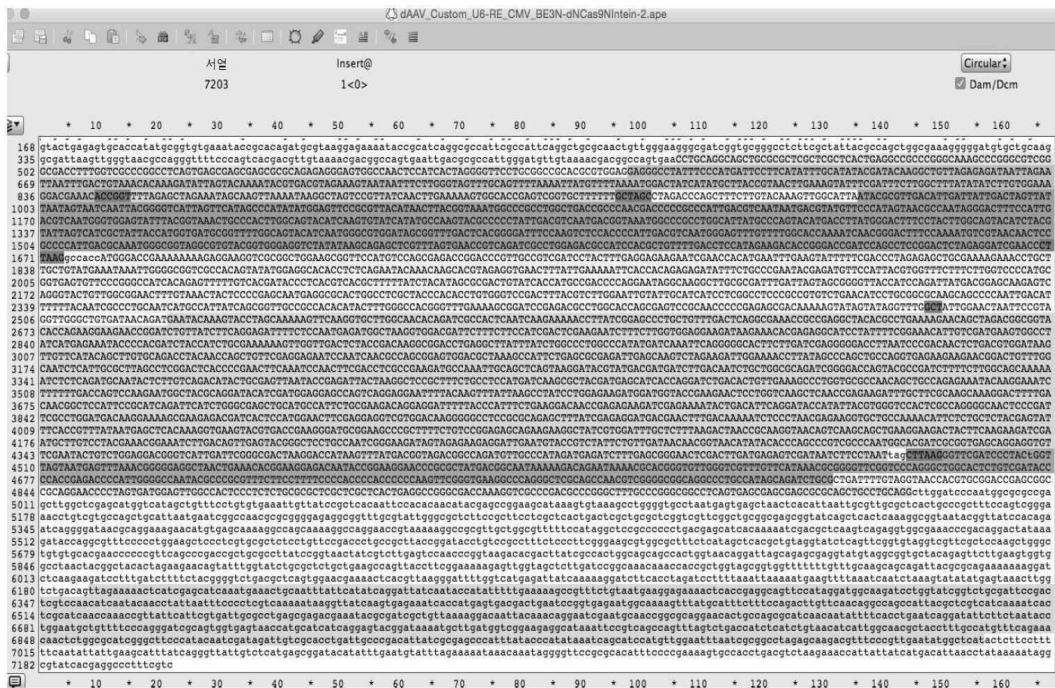
도면45c



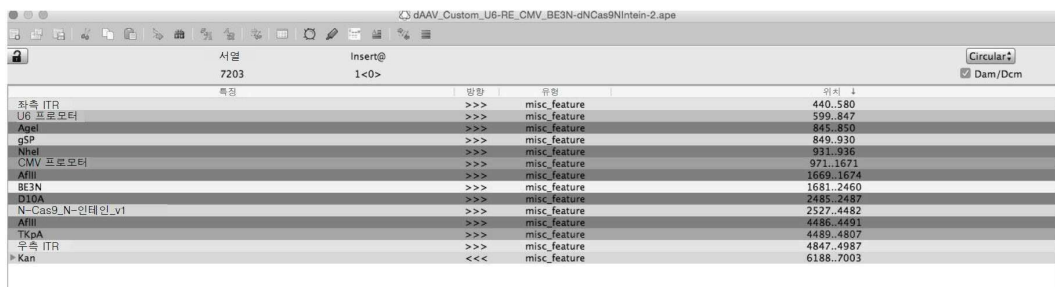
도면46a



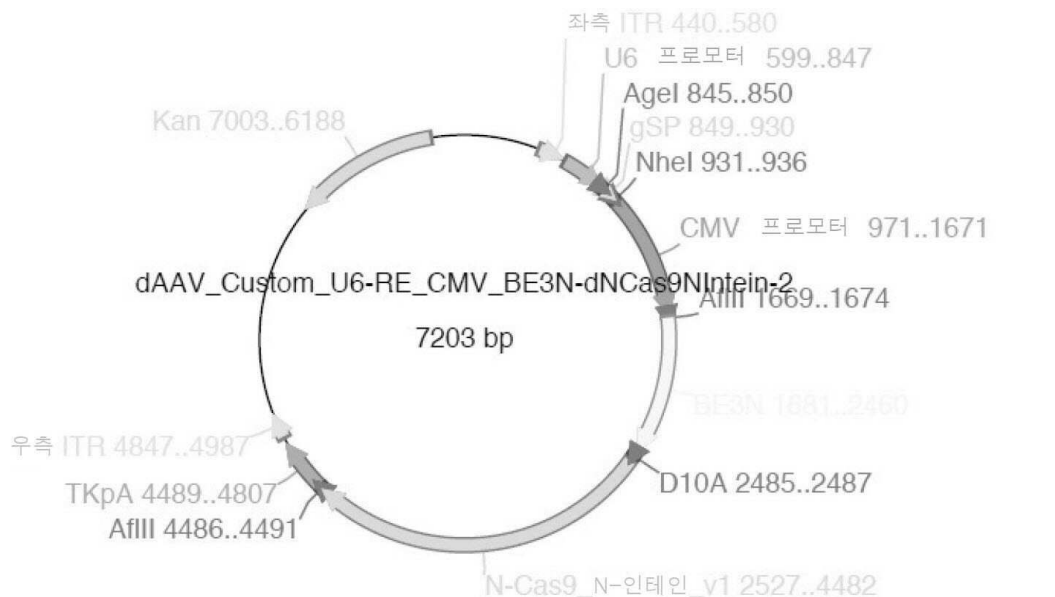
도면46b



도면46c



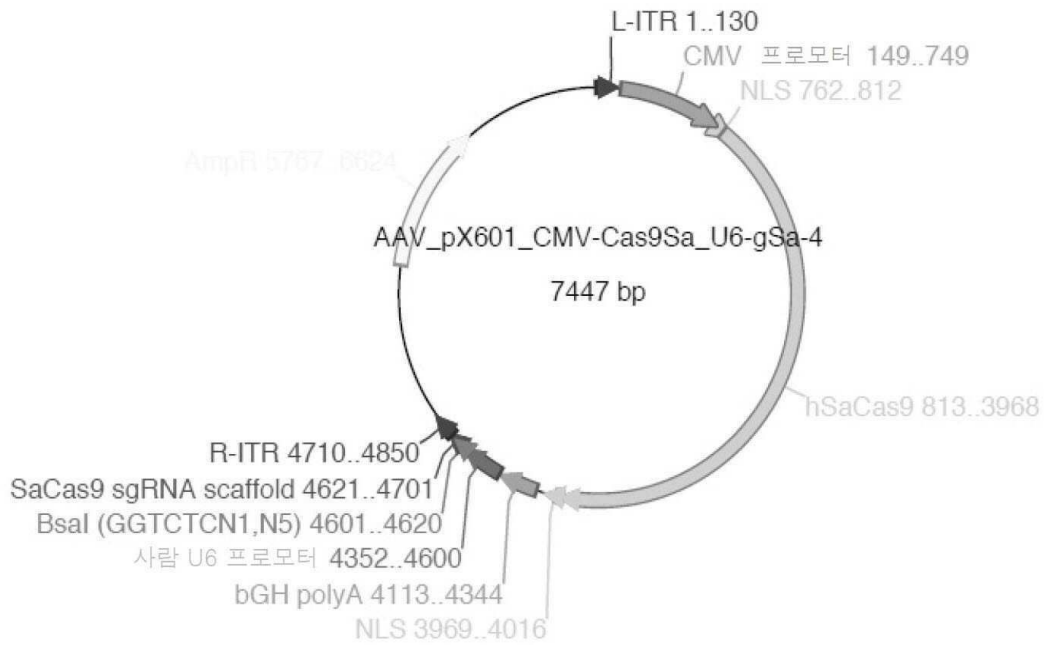
도면46d



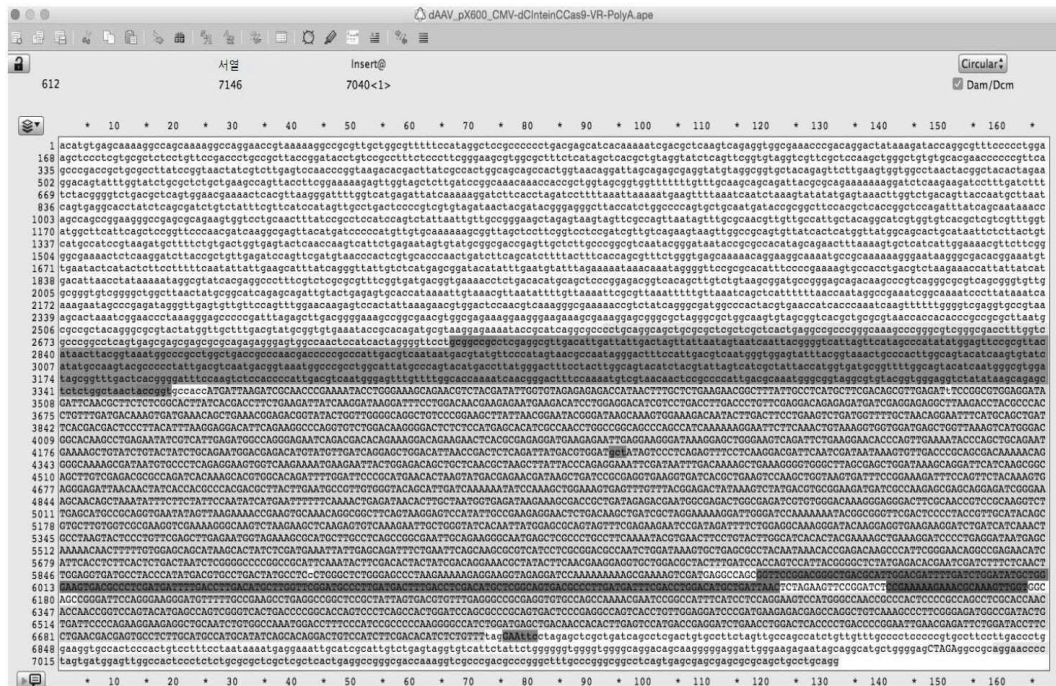
도면47c

특징	방향	유형	위치 ↓
L-ITR	>>>	misc_feature	1..130
CMV 프로모터	>>>	promoter	149..749
NLS	>>>	misc_feature	762..812
hSaCas9	>>>	exon	813..3968
NLS	>>>	misc_feature	3969..4016
bGH polyA	>>>	misc_feature	4113..4344
사람 U6 프로모터	>>>	promoter	4352..4600
Bsal (GGTCTCN1,N5)	>>>	misc_feature	4601..4620
SaCas9 sgRNA scaffold	>>>	misc_feature	4621..4701
R-ITR	>>>	misc_feature	4710..4850
AmpR	>>>	misc_feature	5767..6624

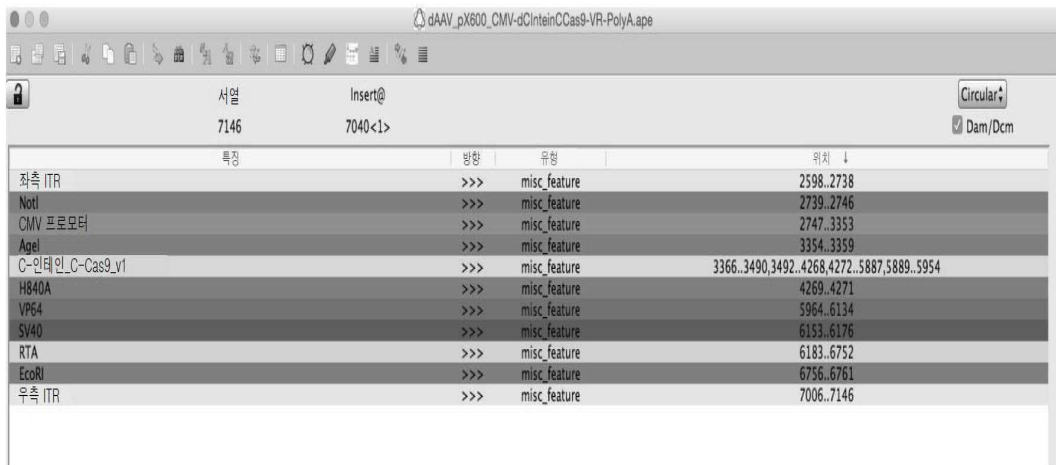
도면47d



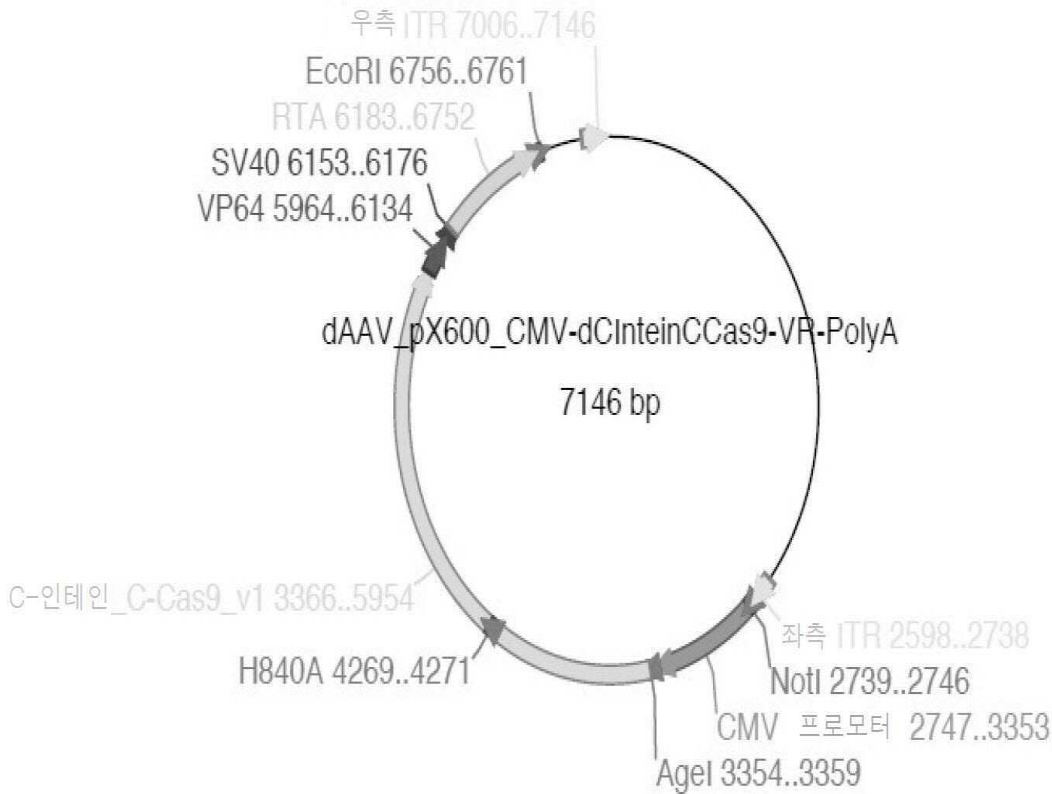
도면48a



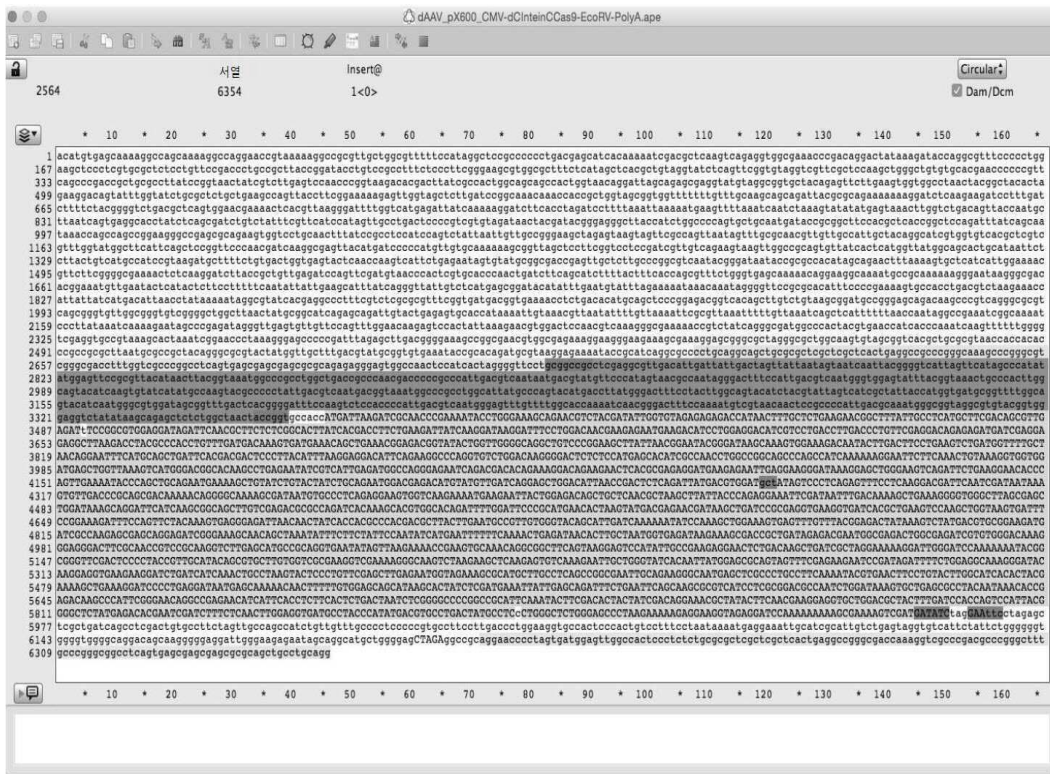
도면48b



도면48c



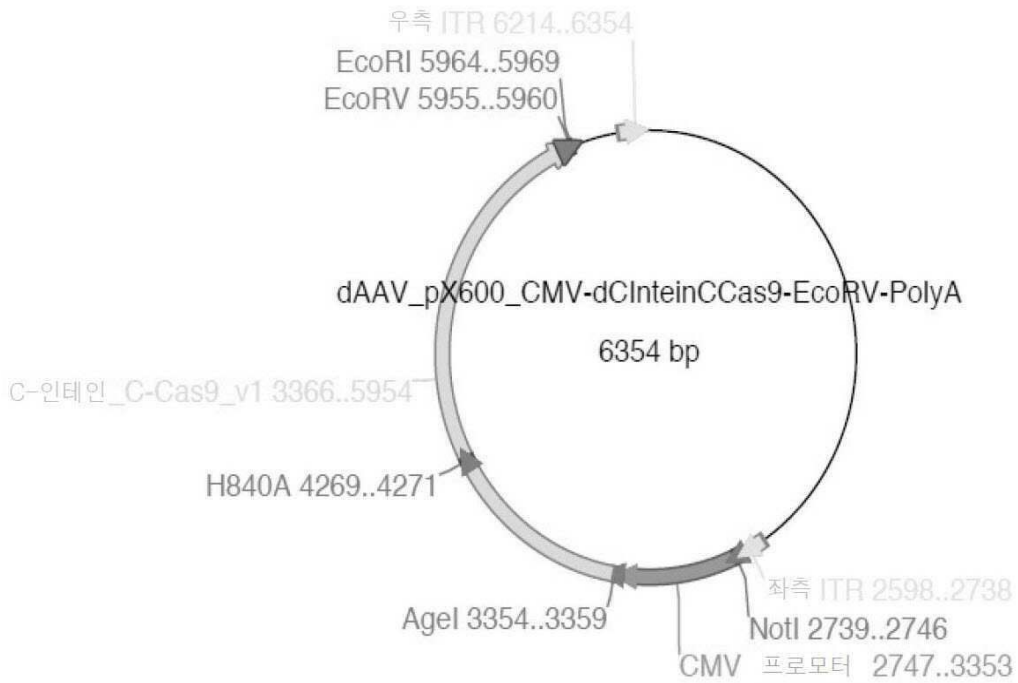
도면49a



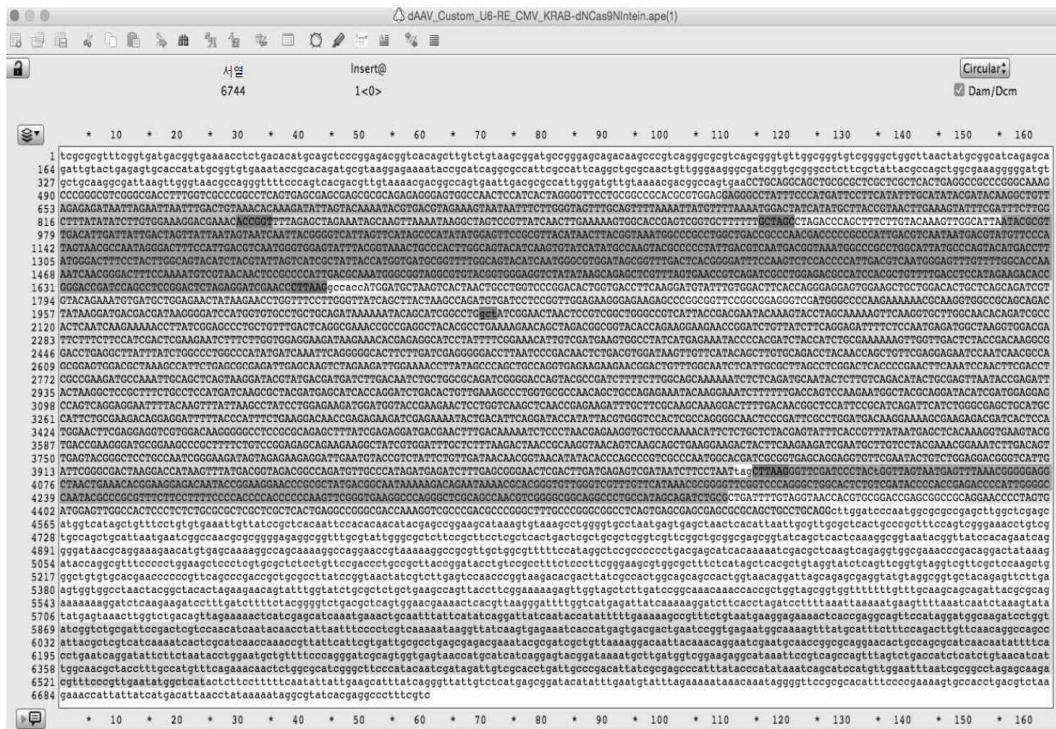
도면49b

특징	방향	유형	위치 ↓
좌측 ITR	>>>	misc_feature	2598..2738
NotI	>>>	misc_feature	2739..2746
CMV 프로모터	>>>	misc_feature	2747..3353
Agel	>>>	misc_feature	3354..3359
C-인테인_C-Cas9_v1	>>>	misc_feature	3366..3490,3492..4268,4272..5887,5889..5954
H840A	>>>	misc_feature	4269..4271
EcoRV	>>>	misc_feature	5955..5960
EcoRI	>>>	misc_feature	5964..5969
우측 ITR	>>>	misc_feature	6214..6354

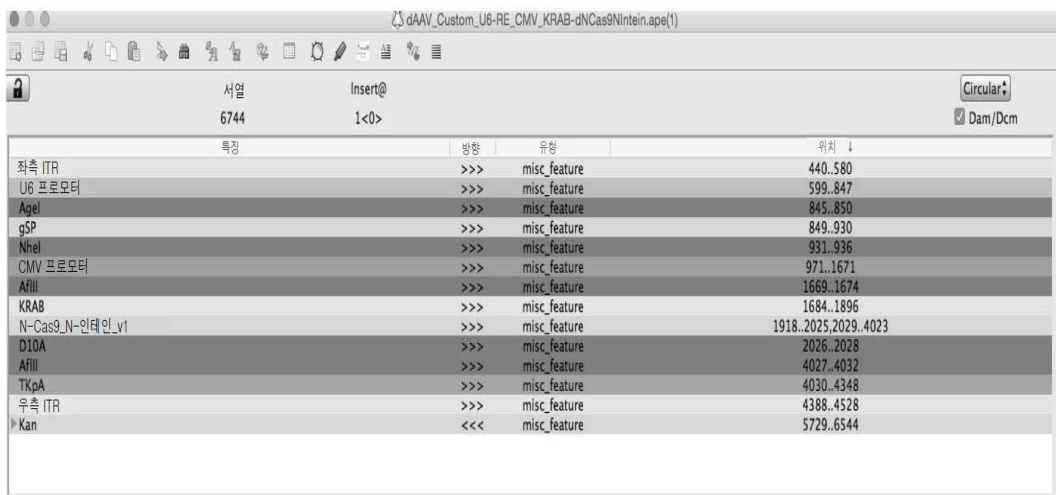
도면49c



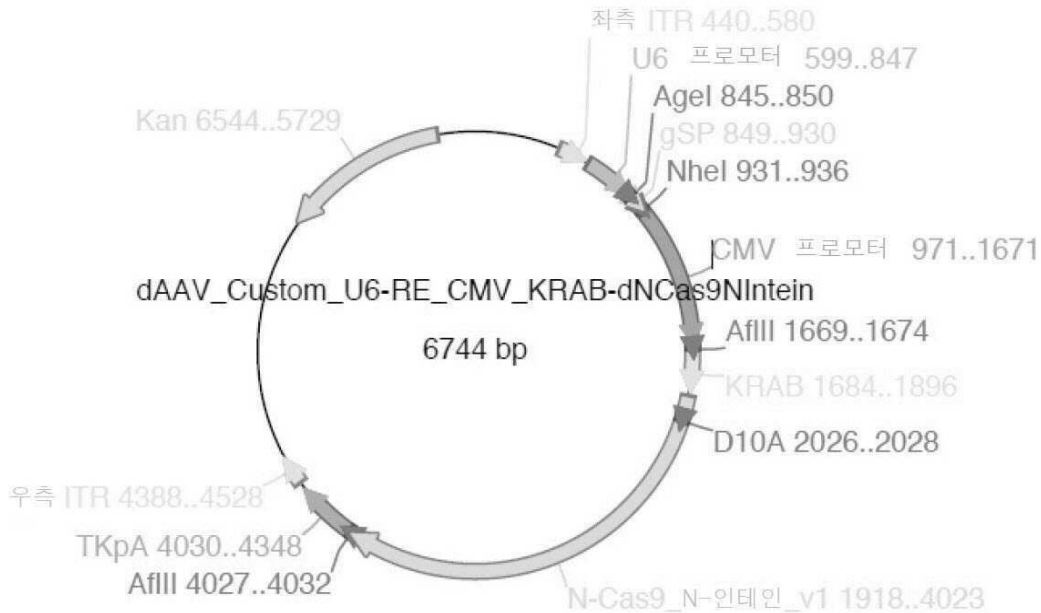
도면50a



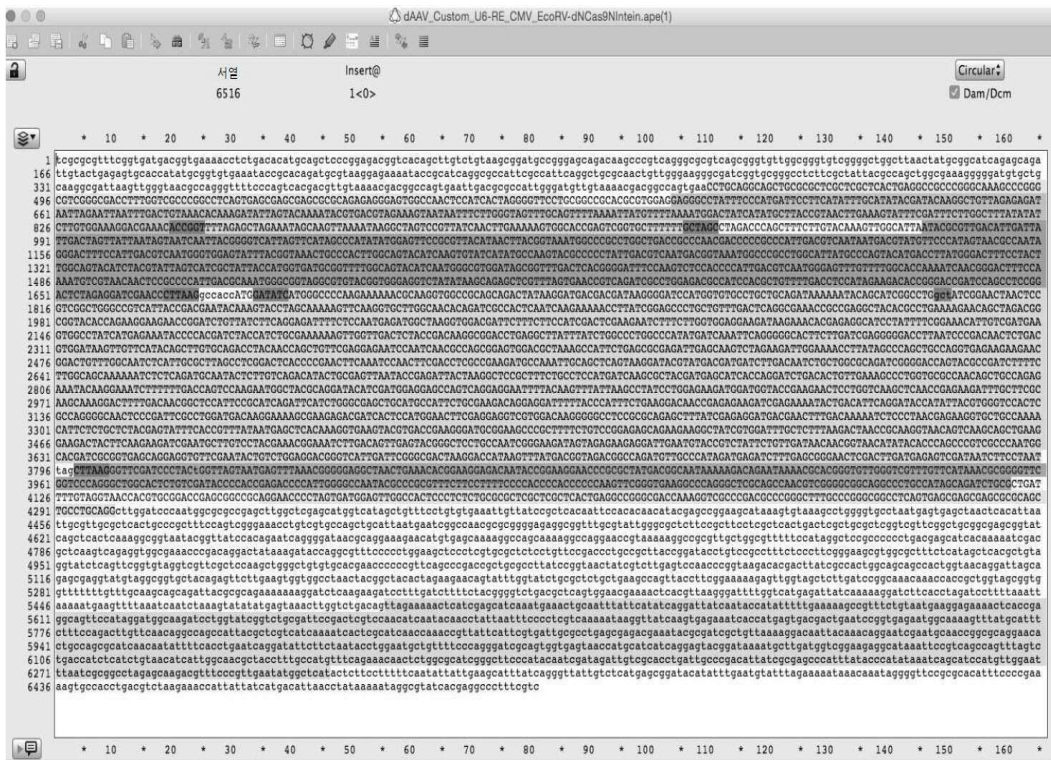
도면50b



도면50c



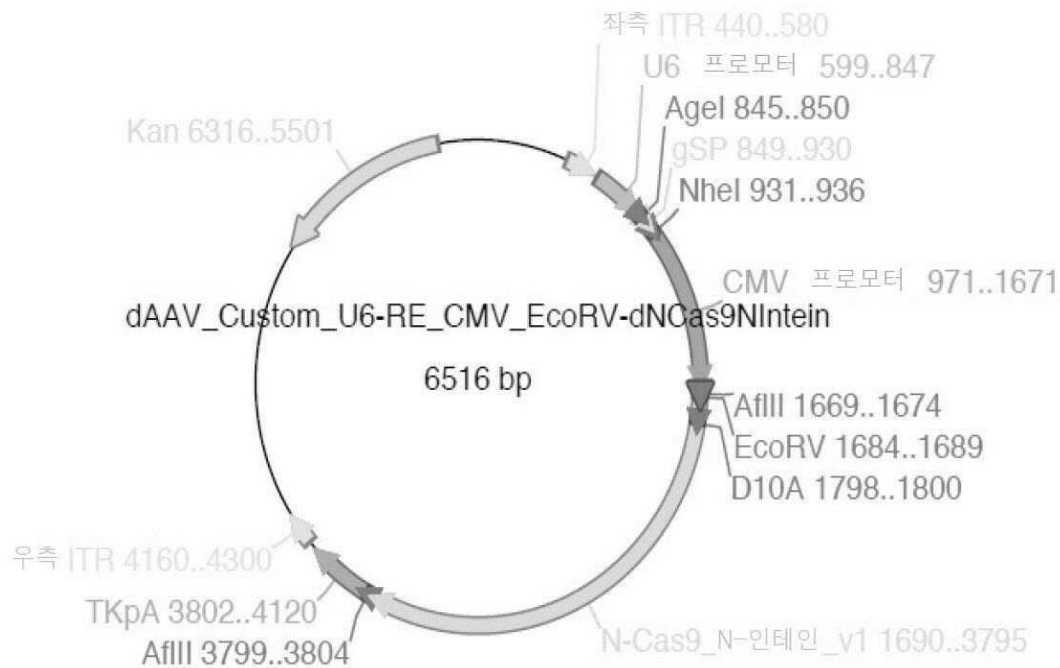
도면51a



도면51b

특징	방향	유형	위치
좌측 ITR	>>>	misc_feature	440..580
U6 프로모터	>>>	misc_feature	599..847
AgeI	>>>	misc_feature	845..850
gSP	>>>	misc_feature	849..930
NheI	>>>	misc_feature	931..936
CMV 프로모터	>>>	misc_feature	971..1671
AfIII	>>>	misc_feature	1669..1674
EcoRV	>>>	misc_feature	1684..1689
N-Cas9_N-인테인_v1	>>>	misc_feature	1690..1797,1801..3795
D10A	>>>	misc_feature	1798..1800
AfIII	>>>	misc_feature	3799..3804
TKpA	>>>	misc_feature	3802..4120
우측 ITR	>>>	misc_feature	4160..4300
Kan	<<<	misc_feature	5501..6316

도면51c



서열 목록

- <110> THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
- <120> CRISPR-CAS GENOME ENGINEERING VIA A MODULAR AAV DELIVERY SYSTEM
- <130> 114198-0121
- <140> PCT/US2017/047687
- <141> 2017-08-18
- <150> 62/481,589
- <151> 2017-04-04
- <150> 62/415,858

<151> 2016-11-01
 <150> 62/376,855
 <151> 2016-08-18
 <160> 346
 <170> PatentIn version 3.5
 <210> 1
 <211> 701
 <212> DNA
 <213> Cytomegalovirus
 <400> 1
 atacgcgttg acattgatta ttgactagtt attaatagta atcaattacg gggtcattag 60

 ttcatagccc atatatggag ttccgcgtta cataacttac ggtaaatggc cgcctggct 120
 gaccgcccac cgacccccgc ccattgacgt caataatgac gtatgttccc atagtaacgc 180
 caatagggac tticcattga cgtcaatggg tggagtatth acggtaaac gcccacttgg 240
 cagtacatca agtgtatcat atgccaagta cggcccctat tgacgtcaat gacggtaaat 300
 ggcccgcctg gcattatgcc cagtacatga ctttatggga ctttctact tggcagtaca 360
 tctacgtatt agtcatcgct attaccatgg tgatgcggtt ttggcagtac atcaatgggc 420
 gtggatagcg gtttgactca cggggatttc caagtctcca cccattgac gtcaatggga 480

 gtttgttttg gcaccaaaat caacgggact ttccaaaatg tcgtaacaac tccgcccac 540
 tgacgcaaat gggcggtagg cgtgtacggt gggaggctca tataagcaga gctcgtttag 600
 tgaaccgtea gatgcctgg agacccatc cacgctgttt tgacctccat agaagacacc 660
 gggaccgate cagcctcgg actctagagg atcgaacct t 701
 <210> 2
 <211> 249
 <212> DNA
 <213> Unknown
 <220><223> Description of Unknown: U6 promoter sequence
 <400> 2
 gagggcctat ttccatgat tccttcatat ttgcatatac gatacaaggc tgtagagag 60

 ataattagaa ttaatttgac tgtaaacaca aagatattag taaaaatac gtgacgtaga 120
 aagtaataat ttcttgggta gtttgagtt taaaattat gttttaaaat ggactatcat 180
 atgcttaccg taacttgaaa gtatttcgat ttcttggctt tatatatctt gtggaaagga 240

cgaaacacc

249

<210> 3

<211> 830

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic polypeptide

<400> 3

Met Ile Lys Ile Ala Thr Arg Lys Tyr Leu Gly Lys Gln Asn Val Tyr

1 5 10 15
Asp Ile Gly Val Glu Arg Asp His Asn Phe Ala Leu Lys Asn Gly Phe

20 25 30
Ile Ala Ser Cys Phe Asp Ser Val Glu Ile Ser Gly Val Glu Asp Arg

35 40 45
Phe Asn Ala Ser Leu Gly Thr Tyr His Asp Leu Leu Lys Ile Ile Lys

50 55 60
Asp Lys Asp Phe Leu Asp Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp

65 70 75 80
Ile Val Leu Thr Leu Thr Leu Phe Glu Asp Arg Glu Met Ile Glu Glu

85 90 95
Arg Leu Lys Thr Tyr Ala His Leu Phe Asp Asp Lys Val Met Lys Gln

100 105 110
Leu Lys Arg Arg Arg Tyr Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu

115 120 125
Ile Asn Gly Ile Arg Asp Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe

130 135 140
Leu Lys Ser Asp Gly Phe Ala Asn Arg Asn Phe Met Gln Leu Ile His

145 150 155 160
Asp Asp Ser Leu Thr Phe Lys Glu Asp Ile Gln Lys Ala Gln Val Ser

165 170 175
Gly Gln Gly Asp Ser Leu His Glu His Ile Ala Asn Leu Ala Gly Ser

180 185 190
Pro Ala Ile Lys Lys Gly Ile Leu Gln Thr Val Lys Val Val Asp Glu

Ala Tyr Leu Asn Ala Val Val Gly Thr Ala Leu Ile Lys Lys Tyr Pro
 450 455 460

Lys Leu Glu Ser Glu Phe Val Tyr Gly Asp Tyr Lys Val Tyr Asp Val
 465 470 475 480

Arg Lys Met Ile Ala Lys Ser Glu Gln Glu Ile Gly Lys Ala Thr Ala
 485 490 495

Lys Tyr Phe Phe Tyr Ser Asn Ile Met Asn Phe Phe Lys Thr Glu Ile
 500 505 510

Thr Leu Ala Asn Gly Glu Ile Arg Lys Arg Pro Leu Ile Glu Thr Asn
 515 520 525

Gly Glu Thr Gly Glu Ile Val Trp Asp Lys Gly Arg Asp Phe Ala Thr
 530 535 540

Val Arg Lys Val Leu Ser Met Pro Gln Val Asn Ile Val Lys Lys Thr
 545 550 555 560

Glu Val Gln Thr Gly Gly Phe Ser Lys Glu Ser Ile Leu Pro Lys Arg
 565 570 575

Asn Ser Asp Lys Leu Ile Ala Arg Lys Lys Asp Trp Asp Pro Lys Lys
 580 585 590

Tyr Gly Gly Phe Asp Ser Pro Thr Val Ala Tyr Ser Val Leu Val Val
 595 600 605

Ala Lys Val Glu Lys Gly Lys Ser Lys Lys Leu Lys Ser Val Lys Glu
 610 615 620

Leu Leu Gly Ile Thr Ile Met Glu Arg Ser Ser Phe Glu Lys Asn Pro
 625 630 635 640

Ile Asp Phe Leu Glu Ala Lys Gly Tyr Lys Glu Val Lys Lys Asp Leu
 645 650 655

Ile Ile Lys Leu Pro Lys Tyr Ser Leu Phe Glu Leu Glu Asn Gly Arg
 660 665 670

Lys Arg Met Leu Ala Ser Ala Gly Glu Leu Gln Lys Gly Asn Glu Leu
 675 680 685

Ala Leu Pro Ser Lys Tyr Val Asn Phe Leu Tyr Leu Ala Ser His Tyr
 690 695 700

Glu Lys Leu Lys Gly Ser Pro Glu Asp Asn Glu Gln Lys Gln Leu Phe
 705 710 715 720

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

725 730 735

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

740 745 750

Leu Ser Ala Tyr Asn Lys His Arg Asp Lys Pro Ile Arg Glu Gln Ala

755 760 765

Glu Asn Ile Ile His Leu Phe Thr Leu Thr Asn Leu Gly Ala Pro Ala

770 775 780

Ala Phe Lys Tyr Phe Asp Thr Thr Ile Asp Arg Lys Arg Tyr Thr Ser

785 790 795 800

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

805 810 815

Leu Tyr Glu Thr Arg Ile Asp Leu Ser Gln Leu Gly Gly Asp

820 825 830

<210> 4

<211> 830

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic polypeptide

<400> 4

Met Ile Lys Ile Ala Thr Arg Lys Tyr Leu Gly Lys Gln Asn Val Tyr

1 5 10 15

Asp Ile Gly Val Glu Arg Asp His Asn Phe Ala Leu Lys Asn Gly Phe

20 25 30

Ile Ala Ser Cys Phe Asp Ser Val Glu Ile Ser Gly Val Glu Asp Arg

35 40 45

Phe Asn Ala Ser Leu Gly Thr Tyr His Asp Leu Leu Lys Ile Ile Lys

50 55 60

Asp Lys Asp Phe Leu Asp Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp

65 70 75 80

Ile Val Leu Thr Leu Thr Leu Phe Glu Asp Arg Glu Met Ile Glu Glu
85 90 95
Arg Leu Lys Thr Tyr Ala His Leu Phe Asp Asp Lys Val Met Lys Gln
100 105 110
Leu Lys Arg Arg Arg Tyr Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu
115 120 125
Ile Asn Gly Ile Arg Asp Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe
130 135 140
Leu Lys Ser Asp Gly Phe Ala Asn Arg Asn Phe Met Gln Leu Ile His
145 150 155 160
Asp Asp Ser Leu Thr Phe Lys Glu Asp Ile Gln Lys Ala Gln Val Ser
165 170 175
Gly Gln Gly Asp Ser Leu His Glu His Ile Ala Asn Leu Ala Gly Ser
180 185 190
Pro Ala Ile Lys Lys Gly Ile Leu Gln Thr Val Lys Val Val Asp Glu
195 200 205
Leu Val Lys Val Met Gly Arg His Lys Pro Glu Asn Ile Val Ile Glu
210 215 220
Met Ala Arg Glu Asn Gln Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg
225 230 235 240
Glu Arg Met Lys Arg Ile Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln
245 250 255
Ile Leu Lys Glu His Pro Val Glu Asn Thr Gln Leu Gln Asn Glu Lys
260 265 270
Leu Tyr Leu Tyr Tyr Leu Gln Asn Gly Arg Asp Met Tyr Val Asp Gln
275 280 285
Glu Leu Asp Ile Asn Arg Leu Ser Asp Tyr Asp Val Asp Ala Ile Val
290 295 300
Pro Gln Ser Phe Leu Lys Asp Asp Ser Ile Asp Asn Lys Val Leu Thr
305 310 315 320
Arg Ser Asp Lys Asn Arg Gly Lys Ser Asp Asn Val Pro Ser Glu Glu

325 330 335
 Val Val Lys Lys Met Lys Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys
 340 345 350
 Leu Ile Thr Gln Arg Lys Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly
 355 360 365
 Gly Leu Ser Glu Leu Asp Lys Ala Gly Phe Ile Lys Arg Gln Leu Val

 370 375 380
 Glu Thr Arg Gln Ile Thr Lys His Val Ala Gln Ile Leu Asp Ser Arg
 385 390 395 400
 Met Asn Thr Lys Tyr Asp Glu Asn Asp Lys Leu Ile Arg Glu Val Lys
 405 410 415
 Val Ile Thr Leu Lys Ser Lys Leu Val Ser Asp Phe Arg Lys Asp Phe
 420 425 430
 Gln Phe Tyr Lys Val Arg Glu Ile Asn Asn Tyr His His Ala His Asp
 435 440 445

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

 515 520 525
 Gly Glu Thr Gly Glu Ile Val Trp Asp Lys Gly Arg Asp Phe Ala Thr
 530 535 540
 Val Arg Lys Val Leu Ser Met Pro Gln Val Asn Ile Val Lys Lys Thr
 545 550 555 560
 Glu Val Gln Thr Gly Gly Phe Ser Lys Glu Ser Ile Leu Pro Lys Arg
 565 570 575
 Asn Ser Asp Lys Leu Ile Ala Arg Lys Lys Asp Trp Asp Pro Lys Lys

<210> 5
 <211> 702
 <212> PRT
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic polypeptide
 <400> 5
 Met Gly Pro Lys Lys Lys Arg Lys Val Ala Ala Ala Asp Tyr Lys Asp
 1 5 10 15
 Asp Asp Asp Lys Gly Ile His Gly Val Pro Ala Ala Asp Lys Lys Tyr
 20 25 30
 Ser Ile Gly Leu Asp Ile Gly Thr Asn Ser Val Gly Trp Ala Val Ile
 35 40 45
 Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe Lys Val Leu Gly Asn
 50 55 60
 Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile Gly Ala Leu Leu Phe
 65 70 75 80
 Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu Lys Arg Thr Ala Arg
 85 90 95
 Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys Tyr Leu Gln Glu Ile
 100 105 110
 Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser Phe Phe His Arg Leu
 115 120 125
 Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys His Glu Arg His Pro
 130 135 140
 Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr His Glu Lys Tyr Pro
 145 150 155 160
 Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp Ser Thr Asp Lys Ala
 165 170 175
 Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His Met Ile Lys Phe Arg
 180 185 190
 Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro Asp Asn Ser Asp Val
 195 200 205

Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr Asn Gln Leu Phe Glu
 210 215 220
 Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala Lys Ala Ile Leu Ser
 225 230 235 240

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

 305 310 315 320
 Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp Ile Leu Arg Val Asn
 325 330 335
 Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser Met Ile Lys Arg Tyr
 340 345 350
 Asp Glu His His Gln Asp Leu Thr Leu Leu Lys Ala Leu Val Arg Gln
 355 360 365
 Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe Asp Gln Ser Lys Asn
 370 375 380

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

450 455 460
 Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile Pro Tyr Tyr Val Gly
 465 470 475 480
 Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp Met Thr Arg Lys Ser
 485 490 495
 Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu Val Val Asp Lys Gly
 500 505 510
 Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr Asn Phe Asp Lys Asn
 515 520 525

 Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser Leu Leu Tyr Glu Tyr
 530 535 540
 Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys Tyr Val Thr Glu Gly
 545 550 555 560
 Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln Lys Lys Ala Ile Val
 565 570 575
 Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr Val Lys Gln Leu Lys
 580 585 590
 Glu Asp Tyr Phe Lys Lys Ile Glu Cys Leu Ser Tyr Glu Thr Glu Ile

 595 600 605
 Leu Thr Val Glu Tyr Gly Leu Leu Pro Ile Gly Lys Ile Val Glu Lys
 610 615 620
 Arg Ile Glu Cys Thr Val Tyr Ser Val Asp Asn Asn Gly Asn Ile Tyr
 625 630 635 640
 Thr Gln Pro Val Ala Gln Trp His Asp Arg Gly Glu Gln Glu Val Phe
 645 650 655
 Glu Tyr Cys Leu Glu Asp Gly Ser Leu Ile Arg Ala Thr Lys Asp His
 660 665 670

 Lys Phe Met Thr Val Asp Gly Gln Met Leu Pro Ile Asp Glu Ile Phe
 675 680 685
 Glu Arg Glu Leu Asp Leu Met Arg Val Asp Asn Leu Pro Asn
 690 695 700
 <210> 6

<211> 702
 <212> PRT
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic polypeptide
 <400> 6
 Met Gly Pro Lys Lys Lys Arg Lys Val Ala Ala Ala Asp Tyr Lys Asp
 1 5 10 15
 Asp Asp Asp Lys Gly Ile His Gly Val Pro Ala Ala Asp Lys Lys Tyr
 20 25 30
 Ser Ile Gly Leu Ala Ile Gly Thr Asn Ser Val Gly Trp Ala Val Ile
 35 40 45
 Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe Lys Val Leu Gly Asn
 50 55 60
 Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile Gly Ala Leu Leu Phe
 65 70 75 80
 Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu Lys Arg Thr Ala Arg
 85 90 95
 Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys Tyr Leu Gln Glu Ile
 100 105 110
 Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser Phe Phe His Arg Leu
 115 120 125
 Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys His Glu Arg His Pro
 130 135 140
 Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr His Glu Lys Tyr Pro
 145 150 155 160
 Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp Ser Thr Asp Lys Ala
 165 170 175
 Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His Met Ile Lys Phe Arg
 180 185 190
 Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro Asp Asn Ser Asp Val
 195 200 205
 Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr Asn Gln Leu Phe Glu

Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile Pro Tyr Tyr Val Gly
 465 470 475 480
 Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp Met Thr Arg Lys Ser
 485 490 495
 Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu Val Val Asp Lys Gly
 500 505 510
 Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr Asn Phe Asp Lys Asn
 515 520 525
 Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser Leu Leu Tyr Glu Tyr
 530 535 540
 Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys Tyr Val Thr Glu Gly
 545 550 555 560
 Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln Lys Lys Ala Ile Val
 565 570 575
 Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr Val Lys Gln Leu Lys
 580 585 590

 Glu Asp Tyr Phe Lys Lys Ile Glu Cys Leu Ser Tyr Glu Thr Glu Ile
 595 600 605
 Leu Thr Val Glu Tyr Gly Leu Leu Pro Ile Gly Lys Ile Val Glu Lys
 610 615 620
 Arg Ile Glu Cys Thr Val Tyr Ser Val Asp Asn Asn Gly Asn Ile Tyr
 625 630 635 640
 Thr Gln Pro Val Ala Gln Trp His Asp Arg Gly Glu Gln Glu Val Phe
 645 650 655
 Glu Tyr Cys Leu Glu Asp Gly Ser Leu Ile Arg Ala Thr Lys Asp His
 660 665 670
 Lys Phe Met Thr Val Asp Gly Gln Met Leu Pro Ile Asp Glu Ile Phe
 675 680 685
 Glu Arg Glu Leu Asp Leu Met Arg Val Asp Asn Leu Pro Asn
 690 695 700
 <210> 7
 <211> 22

<212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 7
 actccctatc agtgatagag aa 22

<210> 8
 <211> 376
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic polynucleotide
 <400> 8
 ttactccct atcagtgata gagaacgtat gaagagtta ctcctatca gtagatagaga 60
 acgtatgcag actttactcc ctatcagtga tagagaacgt ataaggagtt tactccctat 120
 cagtgataga gaacgtatga ccagtttact ccctatcagt gatagagaac gtatctacag 180
 ttactccct atcagtgata gagaacgtat atccagtta ctcctatca gtagatagaga 240
 acgtataagc tttaggcgtg tacggtgggc gcctataaaa gcagagctcg tttagtgaac 300
 cgtcagatcg cctggagcaa ttccacaaca cttttgtctt ataccaactt tccgtaccac 360
 ttctaccct cgtaaa 376

<210> 9
 <211> 248
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic polynucleotide
 <400> 9
 ttactccct atcagtgata gagaacgtat gaagagtta ctcctatca gtagatagaga 60
 acgtatgcag actttactcc ctatcagtga tagagaacgt ataaggagtt tactccctat 120
 cagtgataga gaacgtatga ccagtttact ccctatcagt gatagagaac gtatctacag 180
 ttactccct atcagtgata gagaacgtat atccagtta ctcctatca gtagatagaga 240
 acgtataa 248

<210> 10
 <211> 270
 <212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic polypeptide

<400> 10

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

1 5 10 15

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

20 25 30

Asn Gly Ala Leu Glu Leu Leu Asn Gly Val Gly Ile Glu Gly Leu Thr

35 40 45

Thr Arg Lys Leu Ala Gln Lys Leu Gly Val Glu Gln Pro Thr Leu Tyr

50 55 60

Trp His Val Lys Asn Lys Arg Ala Leu Leu Asp Ala Leu Pro Ile Glu

65 70 75 80

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

85 90 95

Trp Gln Asp Phe Leu Arg Asn Asn Ala Lys Ser Phe Arg Cys Ala Leu

100 105 110

Leu Ser His Arg Asp Gly Ala Lys Val His Leu Gly Thr Arg Pro Thr

115 120 125

Glu Lys Gln Tyr Glu Thr Leu Glu Asn Gln Leu Ala Phe Leu Cys Gln

130 135 140

Gln Gly Phe Ser Leu Glu Asn Ala Leu Tyr Ala Leu Ser Ala Val Gly

145 150 155 160

His Phe Thr Leu Gly Cys Val Leu Glu Glu Gln Glu His Gln Val Ala

165 170 175

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

180 185 190

Arg Gln Ala Ile Glu Leu Phe Asp Arg Gln Gly Ala Glu Pro Ala Phe

195 200 205

Leu Phe Gly Leu Glu Leu Ile Ile Cys Gly Leu Glu Lys Gln Leu Lys

210 215 220

Cys Glu Ser Gly Gly Pro Ala Asp Ala Leu Asp Asp Phe Asp Leu Asp

<220><223> Description of Unknown: miR-1a-3p sequence
 <400> 15
 uggaauguaa agaaguaugu au 22

<210> 16
 <211> 29
 <212> DNA
 <213> Unknown

<220><223> Description of Unknown: Heart target sequence
 <400> 16
 atacatactt cttacattc caagatcac 29

<210> 17
 <211> 22
 <212> RNA
 <213> Unknown

<220><223> Description of Unknown: miR-122a-5p sequence
 <400> 17
 uggaguguga caaugguguu ug 22

<210> 18
 <211> 29
 <212> DNA
 <213> Unknown

<220><223> Description of Unknown: Liver target sequence

<400> 18
 caaacacat tgtcacactc caagatcac 29

<210> 19
 <211> 1710
 <212> PRT
 <213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic polypeptide
 <400> 19
 Met Ser Ser Glu Thr Gly Pro Val Ala Val Asp Pro Thr Leu Arg Arg
 1 5 10 15
 Arg Ile Glu Pro His Glu Phe Glu Val Phe Phe Asp Pro Arg Glu Leu

Ser Lys Lys Phe Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys
 275 280 285

Lys Asn Leu Ile Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu
 290 295 300

Ala Thr Arg Leu Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys
 305 310 315 320

Asn Arg Ile Cys Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys
 325 330 335

Val Asp Asp Ser Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu
 340 345 350

Glu Asp Lys Lys His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp
 355 360 365

Glu Val Ala Tyr His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys
 370 375 380

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

Ala Leu Ala His Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly
 405 410 415

Asp Leu Asn Pro Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu
 420 425 430

Val Gln Thr Tyr Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser
 435 440 445

Gly Val Asp Ala Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg
 450 455 460

Arg Leu Glu Asn Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly
 465 470 475 480

Leu Phe Gly Asn Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe
 485 490 495

Lys Ser Asn Phe Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys
 500 505 510

Asp Thr Tyr Asp Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp
 515 520 525

Gln Tyr Ala Asp Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile
 530 535 540

Leu Leu Ser Asp Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro
 545 550 555 560

Leu Ser Ala Ser Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu
 565 570 575

Thr Leu Leu Lys Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys
 580 585 590

Glu Ile Phe Phe Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp
 595 600 605

Gly Gly Ala Ser Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu
 610 615 620

Glu Lys Met Asp Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu
 625 630 635 640

Asp Leu Leu Arg Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His
 645 650 655

Gln Ile His Leu Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp
 660 665 670

Phe Tyr Pro Phe Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu
 675 680 685

Thr Phe Arg Ile Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser
 690 695 700

Arg Phe Ala Trp Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp
 705 710 715 720

Asn Phe Glu Glu Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile
 725 730 735

Glu Arg Met Thr Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu
 740 745 750

Pro Lys His Ser Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu
 755 760 765

Thr Lys Val Lys Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu

1025 1030 1035 1040
 Lys Glu His Pro Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr

 1045 1050 1055
 Leu Tyr Tyr Leu Gln Asn Gly Arg Asp Met Tyr Val Asp Gln Glu Leu
 1060 1065 1070
 Asp Ile Asn Arg Leu Ser Asp Tyr Asp Val Asp His Ile Val Pro Gln
 1075 1080 1085
 Ser Phe Leu Lys Asp Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser
 1090 1095 1100
 Asp Lys Asn Arg Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val
 1105 1110 1115 1120

 Lys Lys Met Lys Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile
 1125 1130 1135
 Thr Gln Arg Lys Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu
 1140 1145 1150
 Ser Glu Leu Asp Lys Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr
 1155 1160 1165
 Arg Gln Ile Thr Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn
 1170 1175 1180
 Thr Lys Tyr Asp Glu Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile

 1185 1190 1195 1200
 Thr Leu Lys Ser Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe
 1205 1210 1215
 Tyr Lys Val Arg Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr
 1220 1225 1230
 Leu Asn Ala Val Val Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys Leu
 1235 1240 1245
 Glu Ser Glu Phe Val Tyr Gly Asp Tyr Lys Val Tyr Asp Val Arg Lys
 1250 1255 1260

 Met Ile Ala Lys Ser Glu Gln Glu Ile Gly Lys Ala Thr Ala Lys Tyr
 1265 1270 1275 1280

Phe Phe Tyr Ser Asn Ile Met Asn Phe Phe Lys Thr Glu Ile Thr Leu
 1285 1290 1295
 Ala Asn Gly Glu Ile Arg Lys Arg Pro Leu Ile Glu Thr Asn Gly Glu
 1300 1305 1310
 Thr Gly Glu Ile Val Trp Asp Lys Gly Arg Asp Phe Ala Thr Val Arg
 1315 1320 1325
 Lys Val Leu Ser Met Pro Gln Val Asn Ile Val Lys Lys Thr Glu Val
 1330 1335 1340
 Gln Thr Gly Gly Phe Ser Lys Glu Ser Ile Leu Pro Lys Arg Asn Ser
 1345 1350 1355 1360
 Asp Lys Leu Ile Ala Arg Lys Lys Asp Trp Asp Pro Lys Lys Tyr Gly
 1365 1370 1375
 Gly Phe Asp Ser Pro Thr Val Ala Tyr Ser Val Leu Val Val Ala Lys
 1380 1385 1390
 Val Glu Lys Gly Lys Ser Lys Lys Leu Lys Ser Val Lys Glu Leu Leu
 1395 1400 1405
 Gly Ile Thr Ile Met Glu Arg Ser Ser Phe Glu Lys Asn Pro Ile Asp
 1410 1415 1420
 Phe Leu Glu Ala Lys Gly Tyr Lys Glu Val Lys Lys Asp Leu Ile Ile
 1425 1430 1435 1440
 Lys Leu Pro Lys Tyr Ser Leu Phe Glu Leu Glu Asn Gly Arg Lys Arg
 1445 1450 1455
 Met Leu Ala Ser Ala Gly Glu Leu Gln Lys Gly Asn Glu Leu Ala Leu
 1460 1465 1470
 Pro Ser Lys Tyr Val Asn Phe Leu Tyr Leu Ala Ser His Tyr Glu Lys
 1475 1480 1485
 Leu Lys Gly Ser Pro Glu Asp Asn Glu Gln Lys Gln Leu Phe Val Glu
 1490 1495 1500
 Gln His Lys His Tyr Leu Asp Glu Ile Ile Glu Gln Ile Ser Glu Phe
 1505 1510 1515 1520
 Ser Lys Arg Val Ile Leu Ala Asp Ala Asn Leu Asp Lys Val Leu Ser
 1525 1530 1535

Ala Tyr Asn Lys His Arg Asp Lys Pro Ile Arg Glu Gln Ala Glu Asn
 1540 1545 1550

Ile Ile His Leu Phe Thr Leu Thr Asn Leu Gly Ala Pro Ala Ala Phe
 1555 1560 1565

Lys Tyr Phe Asp Thr Thr Ile Asp Arg Lys Arg Tyr Thr Ser Thr Lys
 1570 1575 1580

Glu Val Leu Asp Ala Thr Leu Ile His Gln Ser Ile Thr Gly Leu Tyr
 1585 1590 1595 1600

Glu Thr Arg Ile Asp Leu Ser Gln Leu Gly Gly Asp Ser Gly Gly Ser
 1605 1610 1615

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

1620 1625 1630

Ile Gln Glu Ser Ile Leu Met Leu Pro Glu Glu Val Glu Glu Val Ile
 1635 1640 1645

Gly Asn Lys Pro Glu Ser Asp Ile Leu Val His Thr Ala Tyr Asp Glu
 1650 1655 1660

Ser Thr Asp Glu Asn Val Met Leu Leu Thr Ser Asp Ala Pro Glu Tyr
 1665 1670 1675 1680

Lys Pro Trp Ala Leu Val Ile Gln Asp Ser Asn Gly Glu Asn Lys Ile
 1685 1690 1695

Lys Met Leu Ser Gly Gly Ser Pro Lys Lys Lys Arg Lys Val
 1700 1705 1710

- <210> 20
- <211> 71
- <212> PRT
- <213> Homo sapiens
- <400> 20

Asp Ala Lys Ser Leu Thr Ala Trp Ser Arg Thr Leu Val Thr Phe Lys
 1 5 10 15

Asp Val Phe Val Asp Phe Thr Arg Glu Glu Trp Lys Leu Leu Asp Thr
 20 25 30

Ala Gln Gln Ile Val Tyr Arg Asn Val Met Leu Glu Asn Tyr Lys Asn

35 40 45

Leu Val Ser Leu Gly Tyr Gln Leu Thr Lys Pro Asp Val Ile Leu Arg

50 55 60

Leu Glu Lys Gly Glu Glu Pro

65 70

<210> 21

<211> 57

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic polypeptide

<400> 21

Gly Ser Gly Arg Ala Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu

1 5 10 15

Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp

20 25 30

Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp

35 40 45

Asp Phe Asp Leu Asp Met Leu Ile Asn

50 55

<210> 22

<211> 150

<212> PRT

<213> Unknown

<220><223> Description of Unknown: RTa sequence

<400> 22

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

1 5 10 15

Ala Ile Ser Asp Val Phe Glu Gly Arg Glu Val Cys Gln Pro Lys Arg

20 25 30

Ile Arg Pro Phe His Pro Pro Gly Ser Pro Trp Ala Asn Arg Pro Leu

35 40 45

Pro Ala Ser Leu Ala Pro Thr Pro Thr Gly Pro Val His Glu Pro Val
 50 55 60

Gly Ser Leu Thr Pro Ala Pro Val Pro Gln Pro Leu Asp Pro Ala Pro
 65 70 75 80

Ala Val Thr Pro Glu Ala Ser His Leu Leu Glu Asp Pro Asp Glu Glu
 85 90 95

Thr Ser Gln Ala Val Lys Ala Leu Arg Glu Met Ala Asp Thr Val Ile
 100 105 110

Pro Gln Lys Glu Glu Ala Ala Ile Cys Gly Gln Met Asp Leu Ser His
 115 120 125

Pro Pro Pro Arg Gly His Leu Asp Glu Leu Thr Thr Thr Leu Glu Ser
 130 135 140

Met Thr Glu Asp Leu Asn

145 150

<210> 23

<211> 261

<212> PRT

<213> Unknown

<220><223> Description of Unknown: P65 sequence

<400> 23

Ser Gln Tyr Leu Pro Asp Thr Asp Asp Arg His Arg Ile Glu Glu Lys
 1 5 10 15

Arg Lys Arg Thr Tyr Glu Thr Phe Lys Ser Ile Met Lys Lys Ser Pro
 20 25 30

Phe Ser Gly Pro Thr Asp Pro Arg Pro Pro Pro Arg Arg Ile Ala Val
 35 40 45

Pro Ser Arg Ser Ser Ala Ser Val Pro Lys Pro Ala Pro Gln Pro Tyr
 50 55 60

Pro Phe Thr Ser Ser Leu Ser Thr Ile Asn Tyr Asp Glu Phe Pro Thr
 65 70 75 80

Met Val Phe Pro Ser Gly Gln Ile Ser Gln Ala Ser Ala Leu Ala Pro
 85 90 95

Ala Pro Pro Gln Val Leu Pro Gln Ala Pro Ala Pro Ala Pro Ala Pro
 100 105 110

Ala Met Val Ser Ala Leu Ala Gln Ala Pro Ala Pro Val Pro Val Leu
 115 120 125

Ala Pro Gly Pro Pro Gln Ala Val Ala Pro Pro Ala Pro Lys Pro Thr
 130 135 140

Gln Ala Gly Glu Gly Thr Leu Ser Glu Ala Leu Leu Gln Leu Gln Phe
 145 150 155 160

Asp Asp Glu Asp Leu Gly Ala Leu Leu Gly Asn Ser Thr Asp Pro Ala
 165 170 175

Val Phe Thr Asp Leu Ala Ser Val Asp Asn Ser Glu Phe Gln Gln Leu
 180 185 190

Leu Asn Gln Gly Ile Pro Val Ala Pro His Thr Thr Glu Pro Met Leu
 195 200 205

Met Glu Tyr Pro Glu Ala Ile Thr Arg Leu Val Thr Gly Ala Gln Arg
 210 215 220

Pro Pro Asp Pro Ala Pro Ala Pro Leu Gly Ala Pro Gly Leu Pro Asn
 225 230 235 240

Gly Leu Leu Ser Gly Asp Glu Asp Phe Ser Ser Ile Ala Asp Met Asp
 245 250 255

Phe Ser Ala Leu Leu

260

<210> 24

<211> 322

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic polypeptide

<400> 24

Thr Tyr Gly Leu Leu Arg Arg Arg Glu Asp Trp Pro Ser Arg Leu Gln
 1 5 10 15

Met Phe Phe Ala Asn Asn His Asp Gln Glu Phe Asp Pro Pro Lys Val
 20 25 30

Tyr Pro Pro Val Pro Ala Glu Lys Arg Lys Pro Ile Arg Val Leu Ser
 35 40 45

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

 115 120 125
 Ala Arg Lys Gly Leu Tyr Glu Gly Thr Gly Arg Leu Phe Phe Glu Phe
 130 135 140
 Tyr Arg Leu Leu His Asp Ala Arg Pro Lys Glu Gly Asp Asp Arg Pro
 145 150 155 160
 Phe Phe Trp Leu Phe Glu Asn Val Val Ala Met Gly Val Ser Asp Lys
 165 170 175
 Arg Asp Ile Ser Arg Phe Leu Glu Ser Asn Pro Val Met Ile Asp Ala
 180 185 190

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

 260 265 270
 Glu Met Glu Arg Val Phe Gly Phe Pro Val His Tyr Thr Asp Val Ser

145 150 155 160
 Thr Val Pro Val Trp Arg Arg Gln Pro Val Arg Val Leu Ser Leu Phe
 165 170 175
 Glu Asp Ile Lys Lys Glu Leu Thr Ser Leu Gly Phe Leu Glu Ser Gly
 180 185 190
 Ser Asp Pro Gly Gln Leu Lys His Val Val Asp Val Thr Asp Thr Val
 195 200 205
 Arg Lys Asp Val Glu Glu Trp Gly Pro Phe Asp Leu Val Tyr Gly Ala
 210 215 220
 Thr Pro Pro Leu Gly His Thr Cys Asp Arg Pro Pro Ser Trp Tyr Leu
 225 230 235 240
 Phe Gln Phe His Arg Leu Leu Gln Tyr Ala Arg Pro Lys Pro Gly Ser
 245 250 255
 Pro Arg Pro Phe Phe Trp Met Phe Val Asp Asn Leu Val Leu Asn Lys
 260 265 270

 Glu Asp Leu Asp Val Ala Ser Arg Phe Leu Glu Met Glu Pro Val Thr
 275 280 285
 Ile Pro Asp Val His Gly Gly Ser Leu Gln Asn Ala Val Arg Val Trp
 290 295 300
 Ser Asn Ile Pro Ala Ile Arg Ser Arg His Trp Ala Leu Val Ser Glu
 305 310 315 320
 Glu Glu Leu Ser Leu Leu Ala Gln Asn Lys Gln Ser Ser Lys Leu Ala
 325 330 335
 Ala Lys Trp Pro Thr Lys Leu Val Lys Asn Cys Phe Leu Pro Leu Arg
 340 345 350
 Glu Tyr Phe Lys Tyr Phe Ser Thr Glu Leu Thr Ser Ser Leu
 355 360 365
 <210> 26
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 26
 ggaaagccga cagccgccgc 20

<210> 27
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 27
 ggcgcgggcc tctccttccc 20

<210> 28
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 28
 gagcacgggc gaaagaccga 20

<210> 29
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 29
 gtgtgctctt aaggggtgcg 20

<210> 30
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 30
 gtggcggttg aggcgagcac 20

<210> 31
 <211> 20
 <212> DNA

<213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 31
 gacccatgta acaactccac 20
 <210> 32
 <211> 20

<212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 32
 gtgtatattg ttgaaccggt 20
 <210> 33
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 33
 aacaactcca ctggagtaga 20
 <210> 34
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 34
 caaactgtta agaaacgggc 20
 <210> 35
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 35
 ggttctggca aaattgctgt 20
 <210> 36

<211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 36
 tcgtggattt ctatcacttt 20

<210> 37
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 37
 cttggtaacg tcttctcttg 20

<210> 38
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 38
 cgatggttcc acgtgcaata 20

<210> 39
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 39
 taagctgaat aacaccgttg 20

<210> 40
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 40

ccgcttctctg ttctgagatc 20
 <210> 41
 <211> 20

 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 41
 gtcacgagtt ccaccctgcc 20
 <210> 42
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 42
 cagcctggat ggcttacctc 20
 <210> 43
 <211> 20
 <212> DNA
 <213> Artificial Sequence

 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 43
 gggacttacc agctaggtgc 20
 <210> 44
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 44
 gatctcagaa caggaagcgg 20
 <210> 45
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 45

gtgtaaatta caggaaccaa 20

<210> 46

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 46

gacctggtag ctaggttcta 20

<210> 47

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 47

gatagagtga atctcagaac 20

<210> 48

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 48

gaatagagcc tgtctgaaa 20

<210> 49

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 49

gtggttatgct gtaattcata 20

<210> 50

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 50

ggtctggaaa tggtgattta 20

<210> 51

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 51

gaaagaaaat agagcctgtc 20

<210> 52

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 52

gctaaccat cttggatgct 20

<210> 53

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 53

gacatagaa cctagctacc 20

<210> 54

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 54
 ggcggtcgcc agcgctccag 20

<210> 55
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 55
 gccacctgga aagaagagag 20

<210> 56
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 56
 ggtcgccagc gctccagcgg 20

<210> 57
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 57
 gccagcaatg ggaggaagaa 20

<210> 58
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 58
 gttccagtg gcgtaataca 20

<210> 59
 <211> 20
 <212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 59

ggcggggctg ctacctccac 20

<210> 60

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 60

gggcgcagtc tgcttcagg 20

<210> 61

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 61

ggcgctccag cggcggctgt 20

<210> 62

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 62

gaccgggtgg ttccagcaat 20

<210> 63

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 63

ggggtggttc cagcaatggg 20

<210> 64

<211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 64
 gtgactccgg agtaaagcga 20
 <210> 65
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 65
 gggagctcac catagaactt 20

 <210> 66
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 66
 gacggatcta gatcctccag 20
 <210> 67
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 67
 gccgggtaag agctactagt 20
 <210> 68
 <211> 20

 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 68

gcccgggtgtg tgctgtagaa 20
 <210> 69
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 69
 gtttactccg gagtcactgg 20
 <210> 70
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 70
 gctatctcca ccagtgactc 20
 <210> 71
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 71
 gacatcaccc agggccaagg 20
 <210> 72
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 72
 gtagtttcga gggatccaat 20
 <210> 73
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 73
 gctcccagca gaactgatcg 20
 <210> 74
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 74
 gatgggtcca agtcttcag 20
 <210> 75
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 75
 ggttcctgct ataccacag 20
 <210> 76
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 76
 gccagagagt cggaagtgaa 20
 <210> 77
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 77
 gcctgetata cccacagtgg 20
 <210> 78
 <211> 20

<212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 78
 gggaaagcct ctggaagact 20
 <210> 79
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 79
 ggaagagatg accaccactg 20
 <210> 80
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 80
 ggaatgtcgc catagagctt 20
 <210> 81
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 81
 ggagctcata ggaaagcctc 20
 <210> 82
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 82
 gctttaagac tggaatccta 20

<210> 83
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 83
 gggaagttgc ccaagctcta 20

<210> 84
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 84
 ggaattcgaa tacagctcct 20

<210> 85
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 85
 gcttcaggca gagacccccg 20

<210> 86
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 86
 ggagcctccg tggtagacaca 20

<210> 87
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 87
gcacggcagg aaccttcccc 20

<210> 88
<211> 20
<212> DNA
<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 88
gagcaccgga gggacccgca 20

<210> 89
<211> 20
<212> DNA
<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 89
ggcccggaac gacagagcac 20

<210> 90
<211> 20
<212> DNA
<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 90
gggaacgaca gagcaccgga 20

<210> 91
<211> 20
<212> DNA
<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 91
gaccgcgcg aggccgtgaa 20

<210> 92
<211> 20
<212> DNA

<213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 92
 gctgcccgtg cgggtccctc 20

<210> 93
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 93
 gtacagctcc tgggcgcgcc 20

<210> 94
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 94
 gagcgactcc tgctagtga 20

<210> 95
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 95
 gcgggcccg gaccccacgg 20

<210> 96
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 96
 gtccttga agcacctggg 20

<210> 97

<211> 20
 <212> DNA
 <213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 97
 gagtcgctgt ggacgccctt 20

<210> 98
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 98
 gggactcacc agctagacgc 20

<210> 99
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 99
 gtggtctccc cgctccgtg 20

<210> 100
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 100
 ggggagagct ggctcgtgt 20

<210> 101
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 101

gtgcctcaaa ggtggtcgtg 20

<210> 102
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 102

gctgcatcag ccgtcctcgg 20

<210> 103
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 103

gggacgcctt tcggcactca 20

<210> 104
 <211>
 > 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 104

ggattcgcgt gtccccgga 20

<210> 105
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 105

ggatatgcaa gcgagaagaa 20

<210> 106
 <211> 20
 <212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 106

gctctagacg gacagattaa 20

<210> 107

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 107

gggggaaaaa gaggcggtca 20

<210> 108

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 108

ggcaagcgag aagaaggac 20

<210> 109

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 109

gccaaagcgt cccttccta 20

<210> 110

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 110

gaagcgtccc cttcctaagg 20

<210> 111

<211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 111
 ggcttctaca aaccaagta 20
 <210> 112
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 112
 gaccatgctc caccgagga 20

 <210> 113
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 113
 ggaatgacca tgetccaccg 20
 <210> 114
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 114
 gtgaatctca gaacaggaag 20
 <210> 115
 <211>
 > 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 115

gagcggaggc ataagcagaa 20

<210> 116

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 116

gatctggtgg ctagattcta 20

<210> 117

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 117

gaggaatcac agctcaaca 20

<210> 118

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 118

gatcagaaaa cggccctgga 20

<210> 119

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 119

ggttttgtca gcttacctga 20

<210> 120

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 120
 ggcatccaag atggttagaa 20
 <210> 121
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 121
 gattcctaag gctctccatc 20
 <210> 122
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 122
 gcaatacaga ctaggaatta 20
 <210> 123
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 123
 gagctcaggg agcatcgagg 20
 <210> 124
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 124
 gagagtcgca attggagcgc 20
 <210> 125
 <211> 20

<212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 125
 gccagaccag cctgcacagt 20
 <210> 126
 <211
 > 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 126
 gagcgcaggc taggcctgca 20
 <210> 127
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 127
 gctaggagtc cgggataccc 20
 <210> 128
 <211> 20
 <212> DNA

 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 128
 gaatccgag gtgcactcac 20
 <210> 129
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 129
 gaccagcctg cacagtgggc 20

<210> 130
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 130
 gcgacgCGGT tggcagccga 20

<210> 131
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 131
 ggCagggtgg aactcgtgac 20

<210> 132
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 132
 gcaccatcca gcaagcaggg 20

<210> 133
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 133
 gcgtcactca aggatctaca 20

<210> 134
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 134
gatgggaatg gcacccacga 20

<210> 135
<211> 20
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic oligonucleotide
<400> 135
gcctttagac ggagaacaga 20
<210> 136
<211> 20
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic oligonucleotide
<400> 136
gagatccttg agtgacggac 20
<210> 137
<211>
> 20
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic oligonucleotide
<400> 137
gcggggctcc tccacgaagg 20
<210> 138
<211> 20
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic oligonucleotide
<400> 138
gcaaggaatc acgccttcgt 20
<210> 139
<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 139

ggccatgcgc gaatgctgag 20

<210> 140

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 140

ggcaagccca gccaccttcg 20

<210> 141

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 141

gaggtaaaggc atccaggctg 20

<210> 142

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 142

gttcctgcta gggaggctca 20

<210> 143

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 143

gcctgaaacg acagaggatg 20

<210> 144

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 144

gtcagaggtg gagaccaggt 20

<210> 145

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 145

gccccagcct gaaacgacag 20

<210> 146

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 146

ggccaagagc gagaatctcc 20

<210> 147

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 147

ggtcaggtgt cagagcccat 20

<210> 148

<211>

> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 148
 gggtgtcaga gcccatcggg 20
 <210> 149
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 149
 gtgccctgag cctccctagc 20
 <210> 150
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 150
 gtctgtgaga accgaccgat 20
 <210> 151
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 151
 gggctccgca ggcgcagcgg 20
 <210> 152
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 152
 ggggccagcg cgggggacag 20
 <210> 153
 <211> 20

<212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 153
 gccgctagcg ggccacacag 20
 <210> 154
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 154
 gcgggggaca gcggtccgg 20
 <210> 155
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 155
 gcatcgccc cgcttcgag 20
 <210> 156
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 156
 ggggtacggc gagatgcaa 20
 <210> 157
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 157
 gatgccgacg cgcacgacca 20

<210> 158
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 158
 ggccgccgcc gctgcgctg 20
 <210> 159
 <211>
 > 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 159
 ggggcccgga ctgttccgg 20
 <210> 160
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 160
 gagcggcca cacagggta 20
 <210> 161
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 161
 gggacttacc agctaggtgc 20
 <210> 162
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 162
 gcccaaaag aacagctcca 20

<210> 163
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 163
 ggctggtaag tccttctcat 20

<210> 164
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 164
 ggtgcaggc aactcctaaa 20

<210> 165
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 165
 gacttaactt ggctgactgt 20

<210> 166
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 166
 gtcagcctcc cagaagtcca 20

<210> 167
 <211> 20
 <212> DNA

<213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 167
 ggctgccttg gacttctggg 20

<210> 168
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 168
 gccacggaag gcctccagat 20

<210> 169
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 169
 gccaaggcac ttgctccatt 20

<210> 170
 <211>
 > 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 170
 gggtgctgt gtgtaagag 20

<210> 171
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 171
 gccaacctga atggaagaga 20

<210> 172

<211> 20
 <212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 172

gagggaaagtg gaaagcaagg 20

<210> 173

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 173

gtgggacagg catggatgaa 20

<210> 174

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 174

gcctgtccca ggaacggcat 20

<210> 175

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 175

gtgagaaaag ccaacctgaa 20

<210> 176

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 176
 ggattcgagt gtctcccgga 20

<210> 177
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 177
 gaccaagtcg ttataaggaa 20

<210> 178
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 178
 gaagtcgtta taaggaaagg 20

<210> 179
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 179
 ggaatgacca cgtccacgg 20

<210> 180
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 180
 gcctctggtg tgtactctgt 20

<210> 181
 <211>
 > 20
 <212> DNA

<213> Mus sp.
 <400> 181
 aaagtgatag aaatccacga 20
 <210> 182
 <211> 20
 <212> DNA
 <213> Mus sp.
 <400> 182
 gtgtgtttgc aagatcaatg 20
 <210> 183
 <211> 20
 <212> DNA
 <213> Mus sp.
 <400> 183
 ctggatggga acccgctgag 20
 <210> 184
 <211> 20
 <212> DNA
 <213> Mus sp.
 <400> 184
 tatcctgacc aacacgatgg 20
 <210> 185
 <211> 20
 <212> DNA
 <213> Mus sp.
 <400> 185
 gccagttcca agggtcacgg 20
 <210> 186
 <211> 20
 <212> DNA
 <213> Mus sp.
 <400> 186
 gtgtccgtag agatttaatg 20
 <210> 187

<211> 20
 <212> DNA
 <213> Mus sp.
 <400> 187
 tatctcaaac cgtacccttg 20

<210> 188
 <211> 20
 <212> DNA
 <213> Mus sp.
 <400> 188
 ctgagtacac gagtttaggg 20

<210> 189
 <211> 20
 <212> DNA
 <213> Mus sp.
 <400> 189
 caagagaaga cgttaccaag 20

<210> 190
 <211> 20
 <212> DNA
 <213> Mus sp.
 <400> 190
 gatccattgc cacacaaca 20

<210> 191
 <211> 20
 <212> DNA
 <213> Mus sp.
 <400> 191
 ccagcaatat ggaacttcga 20

<210> 192
 <211> 20
 <212> DNA
 <213> Mus sp.

<400>	192	
	catcactgat cctaacgtgt	20
<210>	193	
<211>	20	
<212>	DNA	
<213>	Mus sp.	
<400>	193	
	tattgcacgt ggaacatcg	20
<210>	194	
<211>	20	
<212>	DNA	
<213>	Mus sp.	
<400>	194	
	gaggacgata tggaatgttg	20
<210>	195	
<211>	20	
<212>	DNA	
<213>	Mus sp.	
<400>	195	
	tttgtttgct caaggagttg	20
<210>	196	
<211>	20	
<212>	DNA	
<213>	Mus sp.	
<400>	196	
	cttaatgaga gtgtttaatg	20
<210>	197	
<211>	20	
<212>	DNA	
<213>	Mus sp.	
<400>	197	
	gaaccctctc cgacgcaccg	20
<210>	198	
<211>	20	

<212> DNA

 <213> Mus sp.
 <400> 198
 agatgcgaca gtatgacacc 20
 <210> 199
 <211> 20
 <212> DNA
 <213> Mus sp.
 <400> 199
 cgtgctcgga tcatacaggc 20
 <210> 200
 <211> 20
 <212> DNA
 <213> Mus sp.
 <400> 200
 gtacctacag atttggcgcg 20
 <210> 201
 <211> 20
 <212> DNA
 <213> Mus sp.
 <400> 201
 taagctgaat aacaccgttg 20

 <210> 202
 <211> 20
 <212> DNA
 <213> Mus sp.
 <400> 202
 aagccacata ctcttgcga 20
 <210> 203
 <211> 20
 <212> DNA
 <213> Mus sp.
 <400> 203

cctgcatca tagagccttg	20
<210> 204	
<211> 20	
<212> DNA	
<213> Mus sp.	
<400> 204	
gctccacgag aagcatgtcg	20
<210> 205	
<211> 20	
<212> DNA	
<213> Mus sp.	
<400> 205	
tatcctacgc ttgctccgaa	20
<210> 206	
<211> 20	
<212> DNA	
<213> Mus sp.	
<400> 206	
ggcaccggtt gtaaccaca	20
<210> 207	
<211> 20	
<212> DNA	
<213> Mus sp.	
<400> 207	
acatcatgga agaatacgac	20
<210> 208	
<211> 20	
<212> DNA	
<213> Mus sp.	
<400> 208	
tgactggcta cggetacaca	20
<210> 209	
<211> 30	

<212> DNA
 <213> Mus sp.
 <400> 209
 gccgaaagtg atagaaatcc acgaaggga 30
 <210> 210
 <211> 30
 <212> DNA
 <213> Mus sp.
 <400> 210
 aggagtgtgt ttgcaagatc aatgaggact 30
 <210> 211
 <211> 30
 <212> DNA
 <213> Mus sp.
 <400> 211
 ctccctggat ggaaccgc tgagcggcga 30
 <210> 212
 <211> 30
 <212> DNA

 <213> Mus sp.
 <400> 212
 ccagtatcct gaccaacacg atggaggta 30
 <210> 213
 <211> 30
 <212> DNA
 <213> Mus sp.
 <400> 213
 tccagccagt tccaagggtc acggaggaag 30
 <210> 214
 <211> 30
 <212> DNA
 <213> Mus sp.
 <400> 214
 ctcagtgtec gtagagatit aatggggcca 30

<210> 215
 <211> 30
 <212> DNA
 <213> Mus sp.
 <400> 215
 actatatctc aaaccgtacc cttgcggaga 30

<210> 216
 <211> 30
 <212> DNA
 <213> Mus sp.
 <400> 216
 gctgctgagt acacgagttt agggcggagc 30

<210> 217
 <211> 30
 <212> DNA
 <213> Mus sp.
 <400> 217
 tggccaagag aagacgttac caagcggag 30

<210> 218
 <211> 30
 <212> DNA
 <213> Mus sp.
 <400> 218
 atcagatcca ttgccacaca acaagggatc 30

<210> 219
 <211> 30
 <212> DNA
 <213> Mus sp.
 <400> 219
 ctgcccagca atatggaact tgcacggctt 30

<210> 220
 <211> 30
 <212> DNA

<213> Mus sp.
 <400> 220
 acttcatcac tgatcctaac gtgtgggtct 30
 <210> 221
 <211> 30
 <212> DNA
 <213> Mus sp.
 <400> 221
 gttttattgc acgtggaacc atcggggcag 30
 <210> 222
 <211> 30
 <212> DNA
 <213> Mus sp.
 <400> 222
 agaagaggac gatatggaat gttgtgggtga 30
 <210> 223
 <211> 30
 <212> DNA
 <213> Mus sp.
 <400> 223
 tcgttttggt tgctcaagga gttgtggctg 30
 <210> 224
 <211> 30
 <212> DNA
 <213> Mus sp.
 <400> 224
 tgatcttaat gagagtgttt aatgtgggcc 30
 <210> 225
 <211> 30
 <212> DNA
 <213> Mus sp.
 <400> 225
 acgagaacce tctccgacgc accgcgggcc 30
 <210> 226

<211> 30
 <212> DNA

 <213> Mus sp.
 <400> 226
 gtgcagatgc gacagtatga cacccggcat 30
 <210> 227
 <211> 30
 <212> DNA
 <213> Mus sp.
 <400> 227
 gaggcgtgct cggatcatac aggccggcgg 30
 <210> 228
 <211> 30
 <212> DNA
 <213> Mus sp.
 <400> 228
 agccgtacct acagatttgg tccgtggaat 30
 <210> 229
 <211> 30
 <212> DNA
 <213> Mus sp.
 <400> 229
 cctataagct gaataacacc gttggggact 30

 <210> 230
 <211> 30
 <212> DNA
 <213> Mus sp.
 <400> 230
 atggaagcca catactcctt gcgatggctg 30
 <210> 231
 <211> 30
 <212> DNA
 <213> Mus sp.

<400> 231
 tgctcctgcg atcatagagc cttggggcg 30

<210> 232
 <211> 30
 <212> DNA
 <213> Mus sp.

<400> 232
 aagggtcca cgagaagcat gtcgtggcgg 30

<210> 233
 <211> 30
 <212> DNA

<213> Mus sp.

<400> 233
 ccaatatacct acgcttgctc cgaacggcca 30

<210> 234
 <211> 30
 <212> DNA
 <213> Mus sp.

<400> 234
 gctaggcacc ggttgtaacc cacaggctg 30

<210> 235
 <211> 30
 <212> DNA
 <213> Mus sp.

<400> 235
 ctcaacatca tggaagaata cgactggatc 30

<210> 236
 <211> 30
 <212> DNA
 <213> Mus sp.

<400> 236
 gggtgactg gctacggcta cacatggatc 30

<210> 237

<211> 595
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic polynucleotide
 <400> 237

gcagagctct ctggctaact accggtgccca ccatgcctgg ctcagcactg ctatgctgcc 60
 tgctcttact gactggcatg aggatcagca ggggccagta cagccgggaa gacaataact 120
 gcacccactt cccagtcggc cagagccaca tgctcctaga gctgcggact gccttcagcc 180
 aggtgaagac tttctttcaa acaaaggacc agctggacaa catactgcta accgactcct 240
 taatgcagga cttaagggt tacttgggtt gccaaacctt atcggaaatg atccagtttt 300

acctggtaga agtgatgcc caggcagaga agcatggccc agaaatcaag gagcatttga 360
 attccctggg tgagaagctg aagaccctca ggatgcggct gaggcgctgt catcgatttc 420
 tcccctgtga aaataagagc aaggcagtgg agcaggtgaa gagtgat ttt aataagctcc 480
 aagaccaagg tgtctacaag gccatgaatg aatttgacat cttcatcaac tgcataagaag 540
 catacatgat gatcaaaatg aaaagctaag aattcctaga gctcgctgat cagcc 595

<210> 238
 <211> 865
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic polynucleotide
 <400> 238

gcagagctct ctggctaact accggtgccca ccatggcgcg gttcctgagg ctttgcacct 60
 ggctgctggc gcttgggtcc tgcctcctgg ctacagtga ggcggaatgc agccaggact 120
 gcgctaaatg cagctaccgc ctggttcgcc caggcgacat caatttctg gcgtgcacac 180
 tggaatgtga aggacagctg ctttctttca aaatctggga gacctgcaag gatctcctgc 240
 aggtgtccag gcccgagtcc ctttgggata acatcgacat gtacaaagac agcagcaaac 300
 aggatgagag ccaacttgcta gccaaagaat acggaggctt catgaaacgg tacggaggct 360
 tcatgaagaa gatggacgag ctatatccca tggagccaga agaagaagcg aacggaggag 420

agatccttgc caagaggtat ggcggcttca tgaagaagga tgcagatgag ggagacacct 480
 tggccaactc ctccgatctg ctgaaagagc tactgggaac gggagacaac cgtgcgaaag 540
 acagccacca acaagagagc accaacaatg acgaagacat gagcaagagg tatgggggct 600
 tcatgagaag cctcaaaaga agcccccaac tggaaatga agcaaaagag ctgcagaagc 660

gctacggggg cttcatgaga aggggtgggac gccccgagtg gtggatggac taccagaaga 720
 ggtatggggg cttcctgaag cgctttgctg agtctctgcc ctccgatgaa gaaggcgaag 780
 attactcgaa agaagtctct gagatagaga aaagatacgg gggctttatg cggttctgag 840

aattcctaga gctcgtgat cagcc 865

<210> 239

<211> 766

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 239

gcagagctct ctggctaact accggtgccca ccatgccgag attctgctac agtcgctcag 60
 gggccctggt gctggccctc ctgcttcaga cctccataga tgtgtggagc tggcctgg 120
 agagcagcca gtgccaggac ctaccacgg agagcaacct gctggcttgc atccgggctt 180
 gcaaactcga cctctcgtg gagacgcccg tgttctctgg caacggagat gaacagcccc 240

tgactgaaaa cccccggaag tacgtcatgg gtcacttccg ctgggaccgc ttcggcccca 300
 ggaacagcag cagtgtctggc agcgcggcgc agaggcgtgc ggaggaagag gcggtgtggg 360
 gagatggcag tcagagccc agtcacgcg agggcaagcg ctctactcc atggagcact 420
 tccgctgggg caagccggtg ggcaagaaac ggcgcccggt gaaggtgtac cccaacttg 480
 ctgagaacga gtcggcggag gcctttccc tagagttcaa gagggagctg gaaggcgagc 540
 ggccattagg cttggagcag gtcctggagt ccgacgcgga gaaggacgac gggccctacc 600
 ggggtggagca ctccgctgg agcaaccgc ccaaggacaa gcgttacggt ggcttcatga 660

cctccgagaa gagccagacg ccctgggtga cgtcttcaa gaacccatc atcaagaacg 720
 cgcaagaag gggccagtga gaattcctag agctcgtga tcagcc 766

<210> 240

<211> 1144

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 240

gcagagctct ctggctaact accggtgccca ccatgagtgc attgctcctc ctggccctgg 60
 tcggggctgc cgtggcttgt aaaggcaaag gagctaaatg cagtagactt atgtatgatt 120

gttgacaggg ttcatgtaga tcaggaagt gcatcgacta taaagacgac gatgacaaac 180

tggcagctgc cggtaacggt aatgggaatg ggaacggcaa cgggaacggt aacggagacg 240

gcacgagggt agcagtagga caggacacgc aagaggaat cgttgtaccg catagtctcc 300

ccttcaaggt agtagtgatc agtgctatac tggcgctggg ggttctcaca attattagtc 360

tgataatfff gataatgctg tggcaaaaaa agccccggag aatccgaatg gtcagtaagg 420

gtgaagaaga caatatggcc ataattaagg agttcatgcg attcaaggta catatggagg 480

gtagcgtcaa tggtcacgag ttcgaaatag aaggcgaagg cgaggggaga ccctatgaag 540

gaacacagac agctaaactt aaggtaacga aaggcggccc actccggttc gcctgggata 600

ttcttagtcc gcagttcatg tacggttcaa aggcgtatgt caaacatcca gcggacatcc 660

ccgattacct gaaattgagc ttcccagagg gatttaaatg ggagcgggtc atgaatttcg 720

aagatggggg agttgtgaca gtaactcaag actccagtct ccaggatggt gaattcatat 780

acaaagtcaa actcaggggc accaatttcc ccagcgacgg ccccgatcatg caaaagaaaa 840

ccatgggatg ggaggccagc tccgagcgca tgtatcctga ggatggagct cttaaaggag 900

agatcaaaca gcgcctgaag ttgaaggatg gaggccacta cgatgccgag gttaagacaa 960

cctataaggc caaaaagcca gtgcagcttc cgggagcgta caatgtaaac atcaagctgg 1020

atattacgag ccacaacgag gactacacga tagtagaaca gtacgagaga gcagaggac 1080

ggcactccac tggtggtatg gacgaattgt ataagtaaga attcctagag ctcgctgatc 1140

agcc 1144

<210> 241

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 241

cgaaattgaa gacgaagagc 20

<210> 242

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 242

ggagactgag agagagaagc 20

<210> 243

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 243

tgatgaggga gggcaccatg 20

<210> 244

<211> 20

<212> DNA

<213>

> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 244

ggtcctgccg ctgcttgca 20

<210> 245

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 245

agccggccag ttccaaacct 20

<210> 246

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 246

agggcccggc gcaatgacag 20

<210> 247

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 247

tcttcaaata accactcctg 20

<210> 248

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 248

tcagcaacaa tgtcaacacc 20

<210> 249

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 249

ggcaatctcc ataatgccgt 20

<210> 250

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 250

tatccacaga gcctaacca 20

<210> 251

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 251

tgtacgaaaa gccagtgatg 20

<210> 252

<211> 20

<212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 252
 gggttcactc cagacctgtg 20
 <210> 253
 <211
 > 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 253
 aaggtctgag aatcggaag 20
 <210> 254
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 254
 cattctggca gagttagcag 20
 <210> 255
 <211> 446
 <212> PRT

 <213> Homo sapiens
 <400> 255
 Met Arg Thr Leu Asn Thr Ser Ala Met Asp Gly Thr Gly Leu Val Val
 1 5 10 15
 Glu Arg Asp Phe Ser Val Arg Ile Leu Thr Ala Cys Phe Leu Ser Leu
 20 25 30
 Leu Ile Leu Ser Thr Leu Leu Gly Asn Thr Leu Val Cys Ala Ala Val
 35 40 45
 Ile Arg Phe Arg His Leu Arg Ser Lys Val Thr Asn Phe Phe Val Ile
 50 55 60

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

Cys Asn Ile Pro Pro Val Leu Tyr Ser Ala Phe Thr Trp Leu Gly Tyr
 405 410 415
 Val Asn Ser Ala Val Asn Pro Ile Ile Tyr Thr Thr Phe Asn Ile Glu
 420 425 430
 Phe Arg Lys Ala Phe Leu Lys Ile Leu His Cys
 435 440

<210> 257

<211> 437

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic polypeptide

<400> 257

Met Ala Asp Asp Pro Ser Ala Ala Asp Arg Asn Val Glu Ile Trp Lys

1 5 10 15
 Ile Lys Lys Leu Ile Lys Ser Leu Glu Ala Ala Arg Gly Asn Gly Thr
 20 25 30
 Ser Met Ile Ser Leu Ile Ile Pro Pro Lys Asp Gln Ile Ser Arg Val
 35 40 45
 Ala Lys Met Leu Ala Asp Asp Phe Gly Thr Ala Ser Asn Ile Lys Ser
 50 55 60
 Arg Val Asn Arg Leu Ser Val Leu Gly Ala Ile Thr Ser Val Gln Gln
 65 70 75 80

Arg Leu Lys Leu Tyr Asn Lys Val Pro Pro Asn Gly Leu Val Val Tyr
 85 90 95

Cys Gly Thr Ile Val Thr Glu Glu Gly Lys Glu Lys Lys Val Asn Ile
 100 105 110
 Asp Phe Glu Pro Phe Lys Pro Ile Asn Thr Ser Leu Tyr Leu Cys Asp
 115 120 125
 Asn Lys Phe His Thr Glu Ala Leu Thr Ala Leu Leu Ser Asp Asp Ser
 130 135 140
 Lys Phe Gly Phe Ile Val Ile Asp Gly Ser Gly Ala Leu Phe Gly Thr
 145 150 155 160
 Leu Gln Gly Asn Thr Arg Glu Val Leu His Lys Phe Thr Val Asp Leu
 165 170 175
 Pro Lys Lys His Gly Arg Gly Gly Gln Ser Ala Leu Arg Phe Ala Arg
 180 185 190
 Leu Arg Met Glu Lys Arg His Asn Tyr Val Arg Lys Val Ala Glu Thr
 195 200 205
 Ala Val Gln Leu Phe Ile Ser Gly Asp Lys Val Asn Val Ala Gly Leu
 210 215 220
 Val Leu Ala Gly Ser Ala Asp Phe Lys Thr Glu Leu Ser Gln Ser Asp
 225 230 235 240
 Met Phe Asp Gln Arg Leu Gln Ser Lys Val Leu Lys Leu Val Asp Ile
 245 250 255
 Ser Tyr Gly Gly Glu Asn Gly Phe Asn Gln Ala Ile Glu Leu Ser Thr
 260 265 270
 Glu Val Leu Ser Asn Val Lys Phe Ile Gln Glu Lys Lys Leu Ile Gly
 275 280 285
 Arg Tyr Phe Asp Glu Ile Ser Gln Asp Thr Gly Lys Tyr Cys Phe Gly
 290 295 300
 Val Glu Asp Thr Leu Lys Ala Leu Glu Met Gly Ala Val Glu Ile Leu
 305 310 315 320
 Ile Val Tyr Glu Asn Leu Asp Ile Met Arg Tyr Val Leu His Cys Gln
 325 330 335
 Gly Thr Glu Glu Glu Lys Ile Leu Tyr Leu Thr Pro Glu Gln Glu Lys
 340 345 350

Asp Lys Ser His Phe Thr Asp Lys Glu Thr Gly Gln Glu His Glu Leu
 355 360 365

Ile Glu Ser Met Pro Leu Leu Glu Trp Phe Ala Asn Asn Tyr Lys Lys
 370 375 380

Phe Gly Ala Thr Leu Glu Ile Val Thr Asp Lys Ser Gln Glu Gly Ser
 385 390 395 400

Gln Phe Val Lys Gly Phe Gly Gly Ile Gly Gly Ile Leu Arg Tyr Arg
 405 410 415

Val Asp Phe Gln Gly Met Glu Tyr Gln Gly Gly Asp Asp Glu Phe Phe
 420 425 430

Asp Leu Asp Asp Tyr

435

<210> 258

<211> 13

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic peptide

<400> 258

Ala His Ile Val Met Val Asp Ala Tyr Lys Pro Thr Lys

1 5 10

<210> 259

<211> 10

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic peptide

<400> 259

Ala Thr His Ile Lys Phe Ser Lys Arg Asp

1 5 10

<210> 260

<211> 735

<212> PRT

<213> Adeno-associated virus 2

<400> 260
Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Thr Leu Ser
1 5 10 15
Glu Gly Ile Arg Gln Trp Trp Lys Leu Lys Pro Gly Pro Pro Pro Pro
20 25 30
Lys Pro Ala Glu Arg His Lys Asp Asp Ser Arg Gly Leu Val Leu Pro
35 40 45
Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro
50 55 60
Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
65 70 75 80
Arg Gln Leu Asp Ser Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
85 90 95
Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly
100 105 110
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125
Leu Gly Leu Val Glu Glu Pro Val Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140
Pro Val Glu His Ser Pro Val Glu Pro Asp Ser Ser Ser Gly Thr Gly
145 150 155 160
Lys Ala Gly Gln Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175
Gly Asp Ala Asp Ser Val Pro Asp Pro Gln Pro Leu Gly Gln Pro Pro
180 185 190
Ala Ala Pro Ser Gly Leu Gly Thr Asn Thr Met Ala Thr Gly Ser Gly
195 200 205
Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser
210 215 220
Ser Gly Asn Trp His Cys Asp Ser Thr Trp Met Gly Asp Arg Val Ile
225 230 235 240
Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu

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

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

 625 630 635 640
 His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala Asn
 645 650 655
 Pro Ser Thr Thr Phe Ser Ala Ala Lys Phe Ala Ser Phe Ile Thr Gln
 660 665 670
 Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln Lys
 675 680 685
 Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn Tyr
 690 695 700

 Asn Lys Ser Val Asn Val Asp Phe Thr Val Asp Thr Asn Gly Val Tyr
 705 710 715 720
 Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu
 725 730 735

 <210> 261
 <211> 737

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic polypeptide

<400> 261

Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Thr Leu Ser

1 5 10 15

Glu Gly Ile Arg Gln Trp Trp Lys Leu Lys Pro Gly Pro Pro Pro Pro

20 25 30

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

35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro

50 55 60

Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp

65 70 75 80

Arg Gln Leu Asp Ser Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala

85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly

100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Leu Leu Glu Pro

115 120 125

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

130 135 140

Pro Val Glu His Ser Pro Val Glu Pro Asp Ser Ser Ser Gly Thr Gly

145 150 155 160

Lys Ala Gly Gln Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln Thr

165 170 175

Gly Asp Ala Asp Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro

180 185 190

Ala Ala Pro Ser Gly Val Gly Ser Leu Thr Met Ala Ala Gly Gly Gly

195 200 205

Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser

210 215 220

725

730

735

Leu

<210> 262

<211> 27

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic peptide

<400> 262

Ala Thr His Ile Lys Phe Ser Lys Arg Asp Gly Ser Gly Ser Gly Ser

1 5 10 15

Arg Pro Lys Pro Gln Gln Phe Phe Gly Leu Met

20 25

<210> 263

<211> 27

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic peptide

<400> 263

Arg Pro Lys Pro Gln Gln Phe Phe Gly Leu Met Gly Ser Gly Ser Gly

1 5 10 15

Ser Ala Thr His Ile Lys Phe Ser Lys Arg Asp

20 25

<210> 264

<211> 50

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic polypeptide

<400> 264

Tyr Thr Ile Trp Met Pro Glu Asn Pro Arg Pro Gly Thr Pro Cys Asp

1 5 10 15

Ile Phe Thr Asn Ser Arg Gly Lys Arg Ala Ser Asn Gly Gly Gly Lys

20 25 30
 Gly Gly Gly Ser Gly Ser Gly Ser Ala Thr His Ile Lys Phe Ser Lys
 35 40 45

Arg Asp
 50

<210> 265
 <211> 50
 <212> PRT
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic polypeptide
 <400> 265

Ala Thr His Ile Lys Phe Ser Lys Arg Asp Gly Ser Gly Ser Gly Ser
 1 5 10 15

Gly Gly Lys Gly Gly Tyr Thr Ile Trp Met Pro Glu Asn Pro Arg Pro
 20 25 30
 Gly Thr Pro Cys Asp Ile Phe Thr Asn Ser Arg Gly Lys Arg Ala Ser
 35 40 45

Asn Gly
 50

<210> 266
 <211> 16
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic primer
 <400> 266

cggcctcagt gagcga 16

<210> 267

<211> 21
 <212> DNA
 <213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic primer
 <400> 267

ggaacccta gtgatggagt t 21

<210> 268
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 268
 ggggccacta gggacaggat 20
 <210> 269
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 269
 gagtccgagc agaagaagaa 20
 <210> 270
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 270
 ggaatccctt ctgcagcacc 20
 <210> 271
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 271
 cagccaaga tagttaagtg 20
 <210> 272
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 272
cgggtggtcg gtagtgagtc 20

<210> 273
<211> 22
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 273
cagacgcgag gaaggagggc gc 22

<210> 274
<211> 20
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 274
cgggagaaag gaacgggagg 20

<210> 275
<211> 20
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 275
gacgcgtgct ctcctcatc 20

<210> 276
<211> 20
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 276
gctgtgggtt ggcctgctg 20

<210> 277
<211> 20
<212> DNA

<213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 277
 accccacat ccatccgcca 20
 <210> 278
 <211>
 > 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 278
 cgaaattgaa gacgaagagc 20
 <210> 279
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 279
 ggacaaagac cacttcagag 20
 <210> 280
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 280
 atttcaggtg agccgaggtt 20
 <210> 281
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 281
 ataatttcta ttatattaca 20
 <210> 282

<211> 20
 <212> DNA
 <213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic primer
 <400> 282
 gaagctggtg gctgaaaagg 20
 <210> 283
 <211> 27
 <212> DNA
 <213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic primer
 <400> 283
 ggagatttag gaagtatggg gttagtg 27
 <210> 284
 <211> 19
 <212> DNA
 <213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic primer
 <400> 284
 cgcgccaac aagaagatg 19
 <210> 285
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic primer
 <400> 285
 acagtcagcc gcatcttctt 20
 <210> 286
 <211> 21
 <212> DNA
 <213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic primer
 <400> 286

catgtacgtt gctatccagg c 21

<210> 287
 <211> 21
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic primer
 <400> 287

gctcaactca ggttaccgtg a 21

<210> 288
 <211> 21
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic primer
 <400> 288

cttcctcat cctcctgeta c 21

<210> 289
 <211> 23
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic primer
 <400> 289

gctcttcgtc ttcaatttcg tct 23

<210> 290
 <211> 19
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic primer
 <400> 290

tggccttcg tgttcctac 19

<210> 291
 <211> 22
 <212> DNA
 <213> Artificial Sequence

<220><
 223> Description of Artificial Sequence: Synthetic primer
 <400> 291
 gtgacgttga catccgtaaa ga 22
 <210> 292
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic primer
 <400> 292
 ctcaactgacg ttggcaaaga 20
 <210> 293
 <211> 28
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic primer
 <400> 293
 aaaacctcct ctcttacttt tctaacttc 28
 <210> 294
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic primer
 <400> 294
 cgacgagtag gatgagaccg 20
 <210> 295
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic primer
 <400> 295
 acgaccaaatt ccgttgactc 20
 <210> 296

<211> 21
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic primer
 <400> 296
 ctccctaatg tcacgcacga t 21
 <210> 297
 <211> 21
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic primer
 <400> 297
 aggggtgtact ggcaagtttg g 21
 <210> 298
 <211> 22
 <212> DNA

 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic primer
 <400> 298
 acaaactggg taaagtgat gg 22
 <210> 299
 <211> 19
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic primer
 <400> 299
 tgttgggtgc cggtttgtt 19
 <210> 300
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><

223> Description of Artificial Sequence: Synthetic primer
 <400> 300

gagttgctgt tgaagtcgca	20
<210> 301	
<211> 19	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic primer	
<400> 301	
gccggactca tcgtactcc	19
<210> 302	
<211> 54	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic primer	
<400> 302	
acactctttc cctacacgac gctcttccga tctagtctg cttgctgctg gcca	54
<210> 303	
<211> 56	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic primer	
<400> 303	
gactggagtt cagacgtgtg ctcttccgat cttgcttgt ccctctgtca atggcg	56
<210> 304	
<211> 57	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic primer	
<400> 304	
acactctttc cctacacgac gctcttccga tctcggtaa tgtggctctg gttctgg	57
<210> 305	
<211> 60	
<212> DNA	
<213> Artificial Sequence	

<220><223> Description of Artificial Sequence: Synthetic primer
 <400> 305
 gactggagtt cagacgtgtg ctcttccgat ctggggttag acccaatata aggagactag 60
 60

<210> 306
 <211> 51
 <212> DNA
 <213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic primer
 <400> 306
 acactctttc cctacacgac gctcttccga tctatgagta tgctgcccgt g 51

<210> 307
 <211> 51
 <212> DNA
 <213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic primer
 <400> 307
 gactggagtt cagacgtgtg ctcttccgat ctgggactca ttcagggtag t 51

<210> 308
 <211> 53
 <212> DNA
 <213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic primer
 <400> 308
 acactctttc cctacacgac gctcttccga tctaggacca atccaagctc cgc 53

<210> 309
 <211> 51
 <212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic primer
 <400> 309
 gactggagtt cagacgtgtg ctcttccgat ctttgcgctg cgccttctca g 51

<210> 310

<211> 54
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic primer
 <400> 310
 acactctttc cctacacgac gctcttccga tcttgtagag caagcagcag gggc 54
 <210> 311
 <211> 57
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic primer
 <400> 311
 gactggagtt cagacgtgtg ctcttccgat ctggtgtcca agaacagtag caggaac 57
 <210> 312
 <211> 35
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 312
 aaaaactata ttaccctggt atccctagcg taact 35
 <210> 313
 <211> 30
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 313
 aaaaatataa gcgggagatt cgtcctcata 30
 <210> 314
 <211> 30
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 314

agttacgcta gggataacag ggtaatatag 30

<210> 315

<211> 25

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 315

tatgaggacg aatctccgc ttata 25

<210> 316

<211> 50

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 316

ggggcttttc tgcaccaat cctgtcccta gtggcccccac tgtggggtgg 50

<210> 317

<211> 35

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 317

ggggcttttc tgtcagtggc cccactgtgg ggtgg 35

<210> 318

<211>

> 38

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 318

ggggcttttc tgtccctagt ggccccactg tggggtgg 38

<210> 319

<211> 38

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 319
 ggggcttttc tgtccctagt ggccccactg tggggtgg 38
 <210> 320
 <211> 53
 <212> DNA

<213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 320
 ggggcttttc tgcaccaac tgtggttgac agaaaagccc cactgtgggg tgg 53
 <210> 321
 <211> 53
 <212> DNA

<213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 321
 ggggcttttc tgcaccaat cctgctgtcc ctagtggccc cactgtgggg tgg 53
 <210> 322
 <211> 53
 <212> DNA
 <213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 322
 ggggcttttc tgcaccaat cctgctgtcc ctagtggccc cactgtgggg tgg 53
 <210> 323
 <211> 53
 <212> DNA
 <213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic oligonucleotide
 <400> 323
 ggggcttttc tgcaccaat cctagtgtcc ctagtggccc cactgtgggg tgg 53
 <210> 324
 <211> 51

<212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 324
 ggggcttttc tgcaccaat cctgtccct agtggcccca ctgtggggtg g 51

<210> 325
 <211> 51
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 325
 ggggcttttc tgcaccaat cctgtccct agtggcccca ctgtggggtg g 51

<210> 326
 <211> 49
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 326
 ggggcttttc tgcacaatc ctgtccctag tggccccact gtggggtg 49

<210> 327
 <211> 49
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 327
 ggggcttttc tgcaccaat ctgtccctag tggccccact gtggggtg 49

<210> 328
 <211> 49
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 328
 ggggcttttc tgcaccaat ctgtccctag tggccccact gtggggtg 49

<210> 329
 <211>
 > 23
 <212> DNA
 <213> Unknown
 <220><223> Description of Unknown: Target sequence
 <400> 329
 ataatttcta ttatattaca ggg 23
 <210> 330
 <211> 23
 <212> DNA
 <213> Unknown
 <220><223> Description of Unknown: Target sequence
 <400> 330
 atttcaggta agccgaggtt tgg 23
 <210> 331
 <211> 23
 <212> DNA
 <213> Unknown
 <220><223> Description of Unknown: Target sequence
 <400> 331
 tctttgaaag agcaataaaa tgg 23
 <210> 332
 <211> 6588
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic polynucleotide
 <400> 332
 acatgtgagc aaaaggccag caaaaggcca ggaaccgtaa aaaggccgcg ttgctggcgt 60
 ttttccatag gctccgcccc cctgacgagc atcacaaaaa tcgacgctca agtcagaggt 120
 ggcgaaacc gacaggacta taaagatacc aggcgtttcc ccctggaagc tcctcgtgc 180
 gctctcctgt tccgacctg ccgcttaccg gatacctgtc cgctttctc ccttcgggaa 240
 gcgtggcgt ttctcatagc tcacgctgta ggtatctcag ttcggtgtag gtcgttcgct 300

ccaagctggg ctgtgtgac gaacccccg ttcagcccga ccgctgcgcc ttatccgta 360
actatcgtct tgagtccaac ccgtaagac acgacttacc gccactggca gcagccactg 420
gtaacaggat tagcagagcg aggtatgtag gcggtgctac agagtctctg aagtgggtggc 480
ctaactacgg ctacactaga aggacagtat ttggtatctg cgctctgctg aagccagtta 540
ccttcggaaa aagagttggt agctcttgat ccggcaaaaa aaccaccgct gtagcgggtg 600
gtttttttgt ttgcaagcag cagattacgc gcagaaaaaa aggatctcaa gaagatcctt 660

tgatcttttc tacggggtct gacgctcagt ggaacgaaaa ctacagttaa gggatttttg 720
tcatgagatt atcaaaaagg atcttcacct agatcctttt aaattaaaaa tgaagtttta 780
aatcaatcta aagtatatat gagtaaacct ggtctgacag ttaccaatgc ttaatcagtg 840
aggcacctat ctacagcagc tgtctatttc gttcatccat agttgcctga ctccccgtcg 900
ttagataac tacgatacgg gagggcttac catctggccc cagtgtctga atgataccgc 960
ggcttccacg ctaccggct ccagatttat cagcaataaa ccagccagcc ggaagggccg 1020
agcgcagaag tggctctgca actttatccg cctccatcca gtctattaat tgttgccggg 1080

aagctagagt aagtagttcg ccagttaata gtttgcgcaa cgttgttgcc attgctacag 1140
gcatcgtggt gtcacgctcg tcgtttggta tggcttcatt cagctccggt tcccaacgat 1200
caaggcgagt tacatgatcc cccatgttgt gcaaaaaagc ggttagctcc ttcggtcctc 1260
cgatcgttgt cagaagtaag ttggccgcag tgttatcact catggttatg gcagcactgc 1320
ataattctct tactgtcatg ccatccgtaa gatgcttttc tgtgactggt gactactcaa 1380
ccaagtcatt ctgagaatag tgtatgcggc gaccgagttg ctcttgcccg gcgtcaatac 1440
gggataatac cgcgccacat agcagaactt taaaagtgt catcattgga aaacgttctt 1500

cggggcgaaa actctcaagg atcttaccgc tgttgagatc cagttcgatg taaccactc 1560
gtgcacccaa ctgatctca gcatctttta ctttaccag cgtttctggg tgagcaaaaa 1620
caggaaggca aaatgccga aaaaaggga taagggcgac acggaatgt tgaatactca 1680
tactttctct ttttcaatat tattgaagca tttatcaggg ttattgtctc atgagcggat 1740
acataattga atgtatttag aaaaataaac aaataggggt tccgcgcaca tttccccgaa 1800
aagtgccacc tgacgtctaa gaaaccatta ttatcatgac attaacctat aaaaataggc 1860
gtatcacgag gccctttcgt ctgcgcggtt tcggtgatga cggtgaaaac ctctgacaca 1920

tgcagctccc ggagacggtc acagcttgc tgaagcggga tgccgggagc agacaagccc 1980
gtcagggcgc gtcagcgggt gttggcgggt gtcgggctg gcttaactat gcggcatcag 2040
agcagattgt actgagagtg caccataaaa ttgtaaacgt taatattttg ttaaaattcg 2100
cgttaaattt ttgttaaact agctcatttt ttaaccaata ggccgaaatc ggcaaaatcc 2160

cttataaatc aaaagaatag cccgagatag ggttgagtgt tgttccagtt tggacaaga 2220
 gtccactatt aaagaacgtg gactccaacg tcaaagggcg aaaaaccgtc taccagggcg 2280
 atggcccact acgtgaacca taccccaat caagttttt ggggtcgagg tgcctaaag 2340

 cactaaatcg gaaccctaaa gggagcccc gatttagagc ttgacgggga aagccggcga 2400
 acgtggcgag aaaggaaggg aagaaagcga aaggagcggg cgctagggcg ctggcaagtg 2460
 tagcggtcac gctgcgcta accaccacac ccgccgcgct taatgcgccc ctacagggcg 2520
 cgtactatgg ttgctttgac gtatgcggtg tgaataccg cacagatgag taaggagaaa 2580
 ataccgcatc aggcgccctt gcaggcagct gcgcgctcgc tcgctcactg aggcgcccg 2640
 ggcaaagccc gggcgtcggg cgacctttg tccccggcc tcagtgagcg agcgagcgcg 2700
 cagagaggga gtggccaact ccatcactag gggttcctgc ggccgcctcg aggcgttgac 2760

 attgattatt gactagttat taatagtaat caattacggg gtcattagtt catagcccat 2820
 atatggagtt ccgcttaca taacttacgg taaatggccc gcctggctga ccgccaacg 2880
 acccccgcc attgacgtca ataatgacgt atgttcccat agtaacgcca atagggactt 2940
 tccattgacg tcaatgggtg gagtatttac ggtaaactgc ccaactggca gtacatcaag 3000
 tgtatcatat gccaagtacg cccctattg acgtcaatga cggtaaatgg cccgcctggc 3060
 attatgccc gtacatgacc ttatgggact ttctacttg gcagtacatc tacgtattag 3120
 tcatcgctat taccatggtg atgcggttt ggcagttacat caatggcgt ggatagcgg 3180

 ttgactcacg gggatttcca agtctccacc ccattgacgt caatgggagt ttgtttggc 3240
 accaaaaatca acgggacttt ccaaaatgtc gtaacaactc cgccccattg acgcaaatgg 3300
 gcggtaggcg tgiacggtgg gaggtctata taagcagagc tctctggcta actaccggtg 3360
 ccaccatgat taagatgca accgaaaat acctgggaaa gcagaacgtc tacgatattg 3420
 gtgtagagag agaccataac tttgctctga agaacggctt tattgcctca tgcttcgaca 3480
 gcgttgagat ttccggcgtg gaggatagat tcaacgctt tctcggcact taccagacc 3540
 ttctgaagat tatcaaggat aaggatttcc tggacaacga agagaatgaa gacatcctgg 3600

 aggacatcgt ccgaccttg accctgttcg aggacagaga gatgatcgag gagaggctta 3660
 agacctacgc ccacctgtt gatgacaaag tgatgaaca gctgaaacgg agacgtata 3720
 ctggttgggg caggctgtcc cggaagctta ttaacggaat acgggataag caaagtggaa 3780
 agacaatact tgacttctg aagtctgatg gttttgctaa caggaatttc atgcagctga 3840
 ttcacgacga ctcccttaca ttaaggagg acattcagaa ggcccaggtg tctggacaag 3900
 gggactctct ccatgagcac atcccaacc tggccggcag cccagccatc aaaaaggaa 3960

ttcttcaaac tgtaaagggtg gtggatgagc tggttaaagt catgggacgg cacaagcctg 4020

 agaatatcgt cattgagatg gccagggaga atcagacgac acagaaagga cagaagaact 4080
 cacgcgagag gatgaagaga attgaggaag ggataaagga gctgggaagt cagattctga 4140
 aggaacaccc agttgaaaat acccagctgc agaatgaaaa gctgtatctg tactatctgc 4200
 agaatggacg agacatgtat gttgatcagg agctggacat taaccgactc tcagattatg 4260
 acgtggatgc tatagtcctt cagagtttcc tcaaggacga ttcaatcgat aataaagtgt 4320
 tgaccgcgag cgacaaaaac aggggcaaaa gcgataatgt gccctcagag gaagtggcca 4380
 agaaaatgaa gaattactgg agacagctgc tcaacgctaa gcttattacc cagaggaaat 4440

 tcgataaatt gacaaaagct gaaaggggtg ggcttagcga gctggataaa gcaggattca 4500
 tcaagcggca gcttgctgag acgcgccaga tcacaaagca cgtggcacag attttgatt 4560
 cccgcatgaa cactaagatg gacgagaacg ataagctgat ccgcgaggtg aaggtgatca 4620
 cgctgaagtc caagctggta agtgatttcc ggaagattt ccagttctac aaagtgaggg 4680
 agattaacaa ctatcaccac gcccacgacg cttacttgaa tgccgtttgt ggtacagcat 4740
 tgatcaaaaa atatccaaag ctggaaagtg agtttgitta cggagactat aaagtctatg 4800
 acgtgcggaa gatgatcgcc aagagcgagc aggagatcgg gaaagcaaca gctaaatatt 4860

 tcttctattc caatatcatg aatTTTTTca aaactgagat aacacttgc t aatggtgaga 4920
 taagaaagcg accgctgata gagacgaatg gcgagactgg cgagatcgtg tgggacaaag 4980
 ggagggactt cgcaaccgct cgcaaggtct tgagcatgcc gcaggtgaat atagttaaga 5040
 aaaccgaagt gcaaacaggc ggcttcagta aggagtccat attgccgaag aggaactctg 5100
 acaagctgat cgctaggaag aaggattggg atccaaaaaa atacggcggg ttcgactccc 5160
 ctaccgttgc atacagcgtg cttgtggtcg cgaaggtcga aaagggcaag tctaagaagc 5220
 tcaagagtgt caaagaattg ctgggtatca caattatgga gcgcagtagt ttcgagaaga 5280

 atccgataga ttttctggag gcaaagggat acaaggaggt gaagaaggat ctgatcatca 5340
 aactgcctaa gtactcctg ttcgagcttg agaatggtag aaagcgcag ctgcctcag 5400
 ccggcgaatt gcagaagggc aatgagctcg cctgccttc aaaatacgtg aacttctgt 5460
 acttggcatc aactacgaa aagctgaaag gatccctga ggataatgag caaaaacaac 5520
 tttttgtgga gcagcataag cactatctcg atgaaattat tgagcagatt tctgaattca 5580
 gcaagcgcgt catcctcgcg gacccaatc tggataaagt gctgagcgc tacaataaac 5640
 accgagacaa gccattcgg gaacaggccg agaacatcat tcacctctc actctgacta 5700

atctcggggc cccggccgca ttcaaatact tcgacactac tatcgacagg aaacgtata 5760
 cttcaacgaa ggaggtgctg gacgctactt tgatccacca gtccattacg gggctctatg 5820
 agacacgaat cgatctttct caacttggag gtgatgccta cccatatgac gtgcctgact 5880
 atgcctccct gggctctggg agccctaaga aaaagaggaa ggtagaggat ccaaaaaaaaa 5940
 agcgaagaat cgatgatggc ggttccggcg gagggtcgga tgctaagtca ctaactgcct 6000
 ggtcccggac actggtgacc ttcaaggatg tatttggga cttcaccagg gaggagtgga 6060
 agctgctgga cactgctcag cagatcgtgt acagaaatgt gatgctggag aactataaga 6120

acctggtttc cttgggttat cagcttacta agccagatgt gatcctccgg ttggagaagg 6180
 gagaagagcc caictagaa ttctagagc tcgctgatca gcctcgactg tgccttctag 6240
 ttgccagcca tctgttgttt gccctcccc cgtgccttcc ttgacctgg aaggtgccac 6300
 tcccactgtc ctttctaata aaaatgagga aattgcatcg cattgtctga gtaggtgtca 6360
 ttctattctg gggggtgggg tgggacagga cagcaagggg gaggattggg aagagaatag 6420
 caggcatgct ggggagctag aggccgagg aaccctagt gatggagtgt gccactccct 6480
 ctctgcgcgc tcgctcgtc actgaggccg ggcgacaaa ggtcgcccga cgccccggct 6540

ttgccggggc ggcctcagtg agcgagcgag cgcgcagctg cctgcagg 6588

<210> 333

<211> 7533

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 333

acatgtgagc aaaaggccag caaaaggcca ggaaccgtaa aaaggccgcg ttgctggcgt 60
 ttttccatag gctccgcccc cctgacgagc atcacaaaaa tcgacgctca agtcagaggt 120
 ggcgaaacce gacaggacta taaagatacc aggcgtttcc ccctggaagc tcctcgtgc 180
 gctctcctgt tccgacctg ccgcttaccg gatacctgtc cgctttctc ctttcgggaa 240

gcgtggcgct ttctcatagc tcacgctgta ggtatctcag ttcggtgtag gtcgttcgct 300
 ccaagctggg ctgtgtgac gaacccccg ttcagcccga ccgctgcgcc ttatccgta 360
 actatcgtct tgagtccaac ccgtaagac acgacttate gccactggca gcagccactg 420
 gtaacaggat tagcagagcg aggtatgtag gcggtgttac agagtcttg aagtgtggc 480
 ctaactacgg ctacactaga aggacagtat ttggtatctg cgctctgctg aagccagtta 540
 ctttcgga aagagttggt agctcttgat ccggcaaaaca aaccaccgct ggtagcgggtg 600

gttttttgt ttgcaagcag cagattacgc gcagaaaaa aggatctcaa gaagatcctt 660

tgatcttttc tacggggtct gacgctcagt ggaacgaaaa ctcacgttaa gggattttgg 720

tcatgagatt atcaaaaagg atcttcacct agatcctttt aaattaaaaa tgaagtttta 780

aatcaatcta aagtataat gagtaaacct ggtctgacag ttaccaatgc ttaatcagtg 840

aggcacctat ctcagcgate tgtctatttc gttcatccat agttgctga ctccccgtcg 900

tgtagataac tacgatacgg gagggcttac catctggccc cagtgctgca atgataccgc 960

ggcttccacg ctcaccggct ccagatttat cagcaataaa ccagccagcc ggaagggccg 1020

agcgcagaag tggctctgca actttatccg cctccatcca gtctattaat tgttgccggg 1080

aagctagagt aagtagttcg ccagttaata gtttgcgcaa cgttgttgcc attgctacag 1140

gcatcgtggt gtcacgctcg tcgtttggta tggttcatt cagctccggt tcccaacgat 1200

caaggcgagt tacatgatcc cccatgttgt gcaaaaaagc ggttagctcc ttcggctctc 1260

cgatcgttgt cagaagtaag ttggccgcag tgttatcact catggttatg gcagactgc 1320

ataattctct tactgtcatg ccatccgtaa gatgcttttc tgtgactggt gactactcaa 1380

ccaagtcatt ctgagaatag tgtatgcggc gaccgagttg ctcttgcccc gcgtcaatac 1440

gggataatac cgcgccacat agcagaactt taaaagtct catcattgga aaacgttctt 1500

cggggcgaaa actctcaagg atcttaccgc tgttgagatc cagttcgatg taaccactc 1560

gtgcacccaa ctgatcttca gcatctttta ctttaccag cgtttctggg tgagcaaaaa 1620

caggaaggca aaatgccca aaaaaggga taaggcgac acggaatgt tgaatactca 1680

tactttctt ttttcaatat tattgaagca ttatcaggg ttattgtctc atgagcggat 1740

acatatttga atgtatttag aaaaataaac aaataggggt tccgcgcaca tttccccgaa 1800

aagtgccacc tgacgtctaa gaaaccatta ttatcatgac attaacctat aaaaatagc 1860

gtatcacgag gccctttcgt ctgcgcggt tccgtgatga cggtgaaaac ctctgacaca 1920

tgcagctccc ggagacggtc acagcttctc tgtaagcgga tgccgggagc agacaagccc 1980

gtcagggcgc gtcagcgggt gttggcgggt gtcggggctg gcttaactat gcggcatcag 2040

agcagattgt actgagagtg caccataaaa ttgtaaacgt taatattttg ttaaaattcg 2100

cgtaaattt ttgttaaatc agctcatttt ttaaccaata ggccgaaatc gcaaaaatcc 2160

cttataaatc aaaagaatag cccgagatag ggttgagtgt tgttccagtt tggacaaga 2220

gtccactatt aaagaactg gactccaacg tcaaaggcg aaaaaccgtc tatcagggcg 2280

atggcccact acgtgaacca tcacccaaat caagtttttt ggggtcgagg tgccgtaag 2340

cactaaatcg gaaccctaaa gggagccccc gatttagagc ttgacgggga aagccggcga 2400
 acgtggcgag aaaggaaggg aagaaagcga aaggagcggg cgctagggcg ctggcaagtg 2460
 tagcggtcac gctgcgcgta accaccacac ccgccgcgct taatgcgccg ctacagggcg 2520
 cgtactatgg ttgctttgac gtatgcggtg tgaatatccg cacagatgcg taaggagaaa 2580
 ataccgcatc aggcgccctt gcaggcagct gcgcgctcgc tcgctcactg aggccgcccg 2640
 ggcaaagccc gggcgtcggg cgacctttgg tcgcccggcc tcagtgagcg agcgagcgcg 2700
 cagagaggga gtggccaact ccatcactag gggttcctgc ggccgcctcg aggcgttgac 2760

 attgattatt gactagtat taatagtaat caattacggg gtcattagt catagcccat 2820
 atatggagtt ccgcgttaca taacttacgg taaatggccc gcctggctga ccgccaacg 2880
 accccgccc attgacgtca ataatgacgt atgttcccat agtaacgcca atagggactt 2940
 tccattgacg tcaatgggtg gagtatttac ggtaaactgc ccacttgca gtacatcaag 3000
 tgtatcatat gccaagtacg cccctattg acgtcaatga cggtaaatgg cccgcctggc 3060
 attatgccc gtacatgacc ttatgggact ttctacttg gcagtacatc tacgtattag 3120
 tcatcgctat taccatggtg atgcggtttt ggcagtacat caatgggcgt ggatagcgg 3180

 ttgactcacg gggatttcca agtctccacc cattgacgt caatgggagt ttgttttggc 3240
 accaaaatca acgggacttt ccaaaatgtc gtaacaactc cgccccattg acgcaaatgg 3300
 gcggtaggcg tgiacgggtg gaggctata taagcagagc tctctggcta actaccggtg 3360
 ccaccatgat taagatcgca acccgaaaat acctgggaaa gcagaacgtc tacgatattg 3420
 gtgtagagag agaccataac tttgctctga agaacggctt tattgcctca tgcttcgaca 3480
 gcgttgagat ttccggcgtg gaggatagat tcaacgctt tctcggcact tatcacgacc 3540
 ttctgaagat tatcaaggat aaggatttcc tggacaacga agagaatgaa gacatcctgg 3600

 aggacatcgt cctgacctg accctgttcg aggacagaga gatgatcgag gagaggctta 3660
 agacctacgc ccacctgtt gatgacaaag tgatgaaaca gctgaaacgg agacgtata 3720
 ctggttgggg caggctgtcc cggaagctta ttaacggaat acgggataag caaagtggaa 3780
 agacaatact tgacttctg aagtctgatg gttttgctaa caggaatttc atgcagctga 3840
 ttcacgacga ctccttaca ttttaaggagg acattcagaa ggcccagggtg tctggacaag 3900
 gggactctct ccatgagcac atcgccaacc tggccgcgag cccagccatc aaaaaggaa 3960
 ttcttcaaac tgtaaagggtg gtggatgagc tggttaaagt catgggacgg cacaagcctg 4020

 agaatatcgt cattgagatg gccagggaga atcagacgac acagaaagga cagaagaact 4080
 cacgcgagag gatgaagaga attgaggaag ggataaagga gctgggaagt cagattctga 4140
 aggaacaccc agttgaaaat acccagctgc agaatgaaa gctgtatctg tactatctgc 4200

agaatggacg agacatgtat gttgatcagg agctggacat taaccgactc tcagattatg 4260
 acgtggatgc tatagtcctc cagagtttcc tcaaggacga ttcaatcgat aataaagtgt 4320
 tgaccgcgag cgacaaaaac aggggcaaaa gcgataatgt gccctcagag gaagtggca 4380
 agaaaatgaa gaattactgg agacagctgc tcaacgctaa gcttattacc cagaggaaat 4440

 tcgataatth gacaaaagct gaaaggggtg ggcttagcga gctggataaa gcaggattca 4500
 tcaagcggca gcttgtcgag acgcgccaga tcacaaagca cgtggcacag attttggatt 4560
 cccgcatgaa cactaagtat gacgagaacg ataagctgat ccgcgaggtg aaggtgatca 4620
 cgctgaagtc caagctggta agtgatttcc ggaaagatth ccagttctac aaagtgaggg 4680
 agattaacaa ctatcaccac gcccacgacg ctactttaa tgccgtttg ggtacagcat 4740
 tgatcaaaaa atatccaaag ctggaaagtg agtttgttta cggagactat aaagtctatg 4800
 acgtgcggaa gatgatcgcc aagagcgagc aggagatcgg gaaagcaaca gctaaatatt 4860

 tcttctattc caatatcatg aatthtttca aaactgagat aaccttgct aatggtgaga 4920
 taagaaagcg accgctgata gagacgaatg gcgagactgg cgagatcgtg tgggacaaag 4980
 ggagggactt cgcaaccgct cgcaaggtct tgagcatgcc gcaggtgaat atagttaaga 5040
 aaaccgaagt gcaaacaggc ggcttcagta aggagtccat attgccgaag aggaactctg 5100
 acaagctgat cgctaggaag aaggattggg atccaaaaaa atacggcggg ttcgactccc 5160
 ctaccgttgc atacagcgtg cttgtggctc cgaaggtcga aaagggaag tctaagaagc 5220
 tcaagagtgt caagaattg ctgggtatca caattatgga gcgcagtagt ttcgagaaga 5280

 atccgataga thttctggag gcaaagggat acaaggaggt gaagaaggat ctgatcatca 5340
 aactgcctaa gtactccctg ttcgagcttg agaatggtag aaagcgcag ctgcctcag 5400
 ccggcgaatt gcagaagggc aatgagctcg cctgccttc aaaatcgtg aacttctctg 5460
 acttggcacc aactacgaa aagctgaaag gatccctga ggataatgag caaaaacaac 5520
 thtttgtgga gcagcataag cactatctcg atgaaatth tgagcagatt tctgaattca 5580
 gcaagcgcgt catcctcggg gacgccaatc tggataaagt gctgagcgc tacaataaac 5640
 accgagacaa gccattcgg gaacaggccg agaacatcat tcacctctc actctgacta 5700

 atctcggggc cccggccgca thcaaaact tcgacactac tatcgacagg aaacgtata 5760
 ctthaacgaa ggaggtgctg gacgctactt tgatccacca gtccattacg gggctctatg 5820
 agacacgaat cgatctthct caacttggag gtgatgccta cccatagac gtgcctgact 5880
 atgcctccct gggctctggg agccctaaga aaaagaggaa ggtagaggat ccaaaaaaa 5940
 agcgaagagt cgatgatggc ggttccggcg gagggtcgat ggagctata cctgcaactg 6000
 atcccgaagc tgaacctagc atggatgtca tccttgtcgg cagcagtgag ctgtcatcta 6060

gtgtctcccc aggtacaggg cgagacttga tgcgatatga ggttaaagcc aaccaacgga 6120

acattgagga catttgcatt tgttgcggtt ccttgcaagt ccacaccaa cacccactct 6180

ttgaggtgg catctgcgt ccttgaagg ataaattcct ggacgccctg ttcctttatg 6240

atgacgacgg ataccagagc tactgttcta tatgttgttc cggggagact ctcttatct 6300

gtggaaatcc tgactgcaca cgggtctact gctttgagtg tgttgattca ttggttggtc 6360

ccggcacaag cggcaaggta catgetatgt ctaattgggt atgttatctg tgcctccca 6420

gctcacgaag tggcctgtg caacgcagac ggaagtggcg aagtcaactt aaagcctttt 6480

atgacagaga atctgagaat cctctggaga tgtttgagac tgtaccagtc tggcgaagac 6540

aacccgtgcg ggigttagc ctgtttgagg atatcaagaa ggagttgact tcctcggtt 6600

tcctggaatc aggaagtgat cccggccagc tcaaacatgt agtcgatgtg actgacacgg 6660

tgccgaaaga tgcgaggag tggggccctt tcgatctggt gtatgggct acacccccct 6720

tggccacac ttgtgacagg cccccgtcat ggtatctgtt ccaatttcac cgctccttc 6780

aatatgcgcg acccaagcca ggttccccga ggccattttt ctggatgttc gtggacaacc 6840

tggtgcttaa caaagagat ttggacgttg cctctagatt cttggaaatg gacacctgta 6900

ctattccgga cgtccatggc ggcagcctcc aaaacgcagt gcgagtctgg tctaacatac 6960

cagcgattcg ctacgccaat tgggctttgg tgcggaaga agaattgagc ctctctgcc 7020

agaataagca aagcagtaaa ctggccgcca aatggccac aaaattggtg aagaactgtt 7080

tcctccatt gcgggagtac ttcaagtact tcagcacaga attgacgtct tcattgatct 7140

aggaattcct agagctcgt gatcagcctc gactgtgcct tctagttgcc agccatctgt 7200

tgtttgcccc tccccgtgc ctctctgac cctggaaggt gccactcca ctgtccttc 7260

ctaataaaat gaggaattg catgcattg tctgagtagg tgtcattcta ttctgggggg 7320

tgggtgggg caggacagca agggggagga ttgggaagag aatagcaggc atgctgggga 7380

gctagaggcc gcaggaacct ctagtgatgg agttggccac tcctctctg cgcgtcgt 7440

cgctcactga ggccgggcca ccaaaggtcg cccgacgcc gggctttgcc cggcgccct 7500

cagtgacga gcgagcgcgc agctgcctgc agg 7533

<210> 334

<211> 7341

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 334

acatgtgagc aaaaggccag caaaaggcca ggaaccgtaa aaaggccgcg ttgctggcgt 60
 ttttccatag gtcgcgcccc cctgacgagc atcacaaaaa tcgacgctca agtcagaggt 120

 ggcgaaacce gacaggacta taaagatacc aggcgtttcc ccctggaagc tecctcgtgc 180
 gctctcctgt tccgaccctg ccgcttaccg gataacctgtc cgcctttctc ccttcgggaa 240
 gcgtggcgct ttctcatagc tcacgctgta ggtatctcag ttcggtgtag gtcgttcgct 300
 ccaagctggg ctgtgtgcac gaaccccccg ttcagcccga ccgctgcgcc ttatecggta 360
 actatcgtct tgagtccaac ccgtaagac acgacttadc gccactggca gcagccactg 420
 gtaacaggat tagcagagcg aggtatgtag gcggtgctac agagtctctg aagtgggtggc 480
 ctaactacgg ctacactaga aggacagtat ttggtatctg cgctctgctg aagccagtta 540

 ccttcgaaa aagagttggt agctcttgat ccggcaaaaa aaccaccgct ggtagcggtg 600
 gtttttttgt ttgcaagcag cagattacgc gcagaaaaaa aggatctcaa gaagatcctt 660
 tgatcttttc tacggggctc gacgctcagt ggaacgaaaa ctcacgtaa gggattttgg 720
 tcatgagatt atcaaaaagg atcttcacct agatcctttt aaattaaaaa tgaagtttta 780
 aatcaatcta aagtatatat gagtaaacct ggtctgacag ttaccaatgc ttaatcagtg 840
 aggacctat ctcagcgate tgtctatttc gttcatccat agttgcctga ctccccgtcg 900
 tntagataac tacgatacgg gagggcttac catctggccc cagtctgca atgataccgc 960

 ggcttccacg ctaccggct ccagatttat cagcaataaa ccagccagcc ggaagggccg 1020
 agcgcagaag tggctctgca actttatccg cctccatcca gctctattaat tgttgccggg 1080
 aagctagagt aagtagttcg ccagttaata gtttgcgcaa cgttgttgcc attgctacag 1140
 gcatcgtggt gtcacgctcg tcgtttggta tggcttcatt cagctccggt tcccaacgat 1200
 caaggcgagt tacatgatcc cccatgttgt gcaaaaaagc ggtagctcc ttcggtcctc 1260
 cgatcgttgt cagaagtaag ttggccgcag tgttatcact catggttatg gcagcactgc 1320
 ataattctct tactgtcatg ccatccgtaa gatgcttttc tgtgactggt gactactcaa 1380

 ccaagtcatt ctgagaatag tgtatgcggc gaccgagttg ctcttgcccg gcgtcaatac 1440
 gggataatac cgcgccacat agcagaactt taaaagtct catcattgga aaacgttctt 1500
 cggggcgaaa actctcaagg atcttaccgc tgttgagatc cagttcgatg taaccactc 1560
 gtgcacccaa ctgatcttca gcacttttta ctttcaccag cgtttctggg tgagcaaaaa 1620
 caggaaggca aaatgccca aaaaaggga taaggcgac acggaatgt tgaatactca 1680
 tactcttct tttcaatat tattgaagca tttatcaggg ttattgtctc atgagcggat 1740

acatatttga atgtatttag aaaaataaac aaataggggt tccgcgcaca tttccccgaa 1800

 aagtgccacc tgacgtctaa gaaaccatta ttatcatgac attaacctat aaaaataggc 1860
 gtatcacgag gccctttcgt ctgcgcggtt tcggtgatga cggtgaaaac ctctgacaca 1920
 tgcagctccc ggagacggtc acagcttgtc tgaagcggga tgccgggagc agacaagccc 1980
 gtcagggcgc gtcagcgggt gttggcgggt gtcggggctg gcttaactat gcggcatcag 2040
 agcagattgt actgagagtg caccataaaa ttgtaaactg taatattttg ttaaaattcg 2100
 cgttaaattt ttgttaaadc agctcatttt ttaaccaata ggccgaaatc ggcaaaatcc 2160
 cttataaatc aaaagaatag cccgagatag ggttgagtgt tgttccagtt tggacaaga 2220

 gtccactatt aaagaactg gactccaacg tcaaagggcg aaaaaccgtc tatcagggcg 2280
 atggcccact acgtgaacca tcacccaaat caagtttttt ggggtcaggg tgccgtaag 2340
 cactaaatcg gaacctaaa gggagccccc gatttagagc ttgacgggga aagccggcga 2400
 acgtggcgag aaaggaaggg aagaaagcga aaggagcggg cgctagggcg ctggcaagtg 2460
 tagcggtcac gctgcgcgta accaccacac ccgccgcgct taatgcgccg ctacagggcg 2520
 cgtactatgg ttgctttgac gtatgcgggtg tgaataaccg cacagatgcg taaggagaaa 2580
 ataccgatc aggcgccctt gcaggcagct gcgcgctcgc tcgctcactg aggccgccc 2640

 ggcaaaagccc gggcgtcggg cgacctttgg tcgcccggcc tcagtgagcg agcagagcgcg 2700
 cagagagggga gtggccaact ccatcactag gggttcctgc ggccgcctcg aggcgttgac 2760
 attgattatt gactagtat taatagtaat caattacggg gtcattagtt catagccat 2820
 atatggagtt ccgcgttaca taacttacgg taaatggccc gcctggctga ccgccaacg 2880
 accccgccc attgacgtca ataatgacgt atgttccat agtaacgcca atagggactt 2940
 tccattgacg tcaatgggtg gagtatttac ggtaaactgc ccacttgca gtacatcaag 3000
 tgtatcatat gccaagtacg cccctattg acgtcaatga cggtaaatgg cccgctggc 3060

 attatgcca gtacatgacc ttatgggact ttctacttg gcagtacatc tacgtattag 3120
 tcacgctat taccatggtg atgcggtttt ggcagtacat caatggcgt ggatagcgg 3180
 ttgactcacg gggatttcca agtctccacc ccattgacgt caatgggagt ttgtttggc 3240
 accaaaatca acgggacttt ccaaaatgct gtaacaactc cgccccattg acgcaaatgg 3300
 gcggtaggcg tftacgggtg gaggctata taagcagagc tctctggcta actaccggtg 3360
 ccacatgat taagatcga acccgaaaat acctgggaaa gcagaacgct tacgatattg 3420
 gigttagagag agaccataac tttgctctga agaacggctt tattgctca tgcttcgaca 3480

gcgttgagat ttccggcgtg gaggatagat tcaacgcttc tctcggcact taccacgacc 3540
 ttctgaagat tatcaaggat aaggatttcc tggacaacga agagaatgaa gacatcctgg 3600
 aggacatcgt cctgaccttg accctgttcg aggacagaga gatgatcgag gagaggctta 3660
 agacctacgc ccacctgttt gatgacaaaag tgatgaaaca gctgaaacgg agacggtata 3720
 ctggttgggg caggctgtcc cggaagctta ttaacggaat acgggataag caaagtggaa 3780
 agacaatact tgacttcctg aagtctgatg gttttgctaa caggaatttc atgcagctga 3840
 ttcacgacga ctcccttaca ttttaaggagg acattcagaa ggcccaggtg tctggacaag 3900

 gggactctct ccatgagcac atcgccaacc tggccggcag cccagccatc aaaaaaggaa 3960
 ttcttcaaac tgtaaagggtg gtggatgagc tggttaaagt catgggacgg cacaagcctg 4020
 agaatatcgt cattgagatg gccagggaga atcagacgac acagaaagga cagaagaact 4080
 cacgcgagag gatgaagaga attgaggaag ggataaagga gctgggaagt cagattctga 4140
 aggaacaccc agttgaaaa acccagctgc agaatgaaaa gctgtatctg tactatctgc 4200
 agaatggacg agacatgatg gttgatcagg agctggacat taaccgactc tcagattatg 4260
 acgtggatgc tatagtcctt cagagtttcc tcaaggacga ttcaatcgat aataaagtgt 4320

 tgaccgcgag cgacaaaaac aggggcaaaa gcgataatgt gccctcagag gaagtgttca 4380
 agaaaatgaa gaattactgg agacagctgc tcaacgctaa gcttattacc cagaggaaat 4440
 tcgataatth gacaaaagct gaaaggggtg ggcttagcga gctggataaa gcaggattca 4500
 tcaagcggca gcttgtcgag acgcgccaga tcacaaagca cgtggcacag attttggatt 4560
 cccgcatgaa cactaagtat gacgagaacg ataagctgat ccgcgaggtg aaggtgatca 4620
 cgctgaagtc caagctggta agtgatttcc ggaagatth ccagttctac aaagtgaggg 4680
 agattaacaa ctatcaccac gccacgacg ctacttgaa tgccgttgtg ggtacagcat 4740

 tgatcaaaaa atatcceaag ctggaaagtg agtttgttta cggagactat aaagtctatg 4800
 acgtgcggaa gatgatgcc aagagcgagc aggagatcgg gaaagcaaca gctaaatatt 4860
 tcttctatth caatatcatg aatthtttca aaactgagat aacacttgct aatggtgaga 4920
 taagaaagcg accgctgata gagacgaatg gcgagactgg cgagatcgtg tgggacaaag 4980
 ggagggactt cgcaaccgtc cgcaaggtct tgagcatgcc gcaggtgaat atagttaaga 5040
 aaaccgaagt gcaaacagcg ggcttcagta aggagtccat attgccgaag aggaactctg 5100
 acaagctgat cgctagggaa aaggattggg atccccaaaa atacggcggg ttcgactccc 5160

 ctaccgttgc atacagcgtg ctttgtgtcg cgaaggtcga aaaggcaag tctaagaagc 5220
 tcaagagtgt caaagaatg ctgggtatca caattatgga gcgcagtagt ttcgagaaga 5280
 atccgataga thttctggag gcaaaaggat acaaggaggt gaagaaggat ctgatcatca 5340

aactgcctaa gtactccctg ttcgagcttg agaatggtag aaagcgcag cttgcctcag 5400
 ccggcgaatt gcagaagggc aatgagctcg cctgccttc aaaatcgtg aacttctgt 5460
 acttgccatc aactacgaa aagctgaaag gatcccctga ggataatgag caaaaacaac 5520
 tttttgtgga gcagcataag cactatctcg atgaaattat tgagcagatt tctgaattca 5580

gcaagcgcgt catcctcgcg gacgccaatc tggataaagt gctgagcgc tacaataaac 5640
 accgagacaa gccattcgg gaacaggccg agaacatcat tcacctctc actctgacta 5700
 atctcggggc cccggccgca ttcaataact tcgacactac tatcgacagg aaacgctata 5760
 cttcaacgaa ggaggtgctg gacgctactt tgatccacca gtccattacg gggctctatg 5820
 agacacgaat cgatctttct caacttggag gtgatgccta cccatgatgac gtgcctgact 5880
 atgcctccct gggctctggg agccctaaga aaaagaggaa ggtagaggat ccaaaaaaaa 5940
 agcgaagaat cgatgatggc ggttccggcg gagggctgac ctatggtctt cttaggagaa 6000

gagaagactg gccctctcgg ctccaatgt tcttcgctaa taatcacgat caagaattcg 6060
 acccgcctaa ggtctacca cccgtgccag cagagaaacg aaagccgatc agagtattgt 6120
 ctttgttcca tggcatagcc accggactcc tgggtgctgaa agatctggga atccaggttg 6180
 atcgctacat cgctcagag gtttgtgaag actctataac cgtagggatg gtacgacacc 6240
 aggtaagat aatgtatgtc ggtgatgtac ggtccgtgac acaaaaacac atacaggagt 6300
 ggggaccctt tgaccttctg ataggcggat ctccatgcaa tgaccttcc attgttaatc 6360
 ctgccgcaa aggactttac gaaggaaccg gccgactctt ttttgaattt tatcggttgc 6420

tccatgatgc tggcgcgaag gagggcagat accgcccctt tttctgctt ttcgagaacg 6480
 tcgtcgtat gggcgtttcc gataagagag acataagccg attccttgag agcaaccag 6540
 taatgattga tgcaaaagaa gtttctgccg cccacagggc taggtacttc tggggaatt 6600
 tgccaggcat gaaccgcca ctggcateca cgttaacga taagctggaa cttcaggaat 6660
 gtttgagca cggtagaatc gcaaaattct caaaagtaag aacgatcacg acaagaagta 6720
 attctatcaa gcaagggaaa gatcagcact tccccgtctt tatgaatgaa aaggaggaca 6780
 ttctttggtg cactgaaatg gagcgcgtgt tcggatttcc tgttactat acggacgtca 6840

gcaatatgtc tcgctcgcg aggcagcag tgttggccg ctcttggagt gttccagtca 6900
 tacgacatct ttttgcgcca cttaaagaat actttgctg tgtgatctag gaattcctag 6960
 agctcgtga tcagcctcga ctgtgccttc tagttgcag ccatctgtt tttgccctc 7020
 cccctgctt tccttgacc tggaggtgc cactcccact gtccttctc aataaaatga 7080
 ggaaattgca tcgacttgc tgagttagtg tcattctatt ctggggggtg ggggtgggca 7140
 ggacagcaag ggggaggatt gggaagagaa tagcaggcat gctggggagc tagaggccgc 7200

aggaaccct agtgatggag ttggccactic cctctctgcg cgctcgctcg ctcactgagg 7260

 ccgggcgacc aaaggtgcc cgacccccg gctttgccg ggcgcctca gtgagcgagc 7320
 gagcgcgag ctgcctgcag g 7341
 <210> 335
 <211> 6759
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic polynucleotide
 <400> 335
 tcgcgcttt cggtgatgac ggtgaaaacc tctgacacat gcagctccc gagacggta 60
 cagcttgtct gtaagcgat gccgggagca gacaagccc tcagggcgcg tcagcgggtg 120
 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180

 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc 240
 attgccatt caggctgcgc aactgttggg aaggcgatc ggtgcgggcc tettegctat 300
 tacgccagct ggcgaaagg ggatgtgctg caagcgatt aagtgggta acgccagggt 360
 tttccagtc acgacgttg aaaacgacgg ccagtgaatt gacgcgcat tgggatgttg 420
 taaaacgacg gccagtgaac ctgcaggeag ctgcgcgctc gctcgctcac tgaggccgcc 480
 cgggcaaagc ccgggcgtcg ggcgacctt ggtcgcccgg cctcagttag cgagcgagcg 540
 cgagagagg gagtggccaa ctccatcact aggggttctt gcggccgcac gcgtggagga 600

 ggcctatth cccatgattc cttcatatth gcatatacga tacaaggctg ttagagagat 660
 aattagaatt aattgactg taaacacaaa gatattagta caaaatacgt gacgtagaaa 720
 gtaataatth cttgggtagt ttgcagttth aaaattatgt tttaaaatgg actatcatat 780
 gcttaccgta acttgaaagt atttcgattt cttggcttta tatacttctg gaaaggacg 840
 aaacaccggt ttagagcta gaaatagcaa gttaaaataa ggctagtccg ttatcaactt 900
 gaaaaagtgg caccgagtcg gtgctttttt gctagcctag acccagctth cttgtacaaa 960
 gttggcatta atacgcgttg acattgatta ttgactagtt attaatagta atcaattacg 1020

 gggtcattag ttcatagecc atatatggag ttcgcgtta cataacttac ggtaaatggc 1080
 ccgcctggct gaccgcccac cgacccccgc ccattgacgt caataatgac gtatgttccc 1140
 atagtaacgc caataggac tttccattga cgtcaatggg tggagtatth acggtaact 1200
 gccacttgg cagtacatca agtgtatcat atgccaagta cgcccctat tgacgtcaat 1260
 gacggtaaat ggccccctg gcattatgcc cagtacatga cttatggga ctttctact 1320

tggcagtaca tctacgtatt agtcatcgct attacatgg tgatgcggtt ttggcagtac	1380
atcaatgggc gtggatagcg gtttgactca cggggatttc caagtctcca ccccatgac	1440
gtcaatggga gtttgtttg gcacaaaaat caacgggact ttccaaaatg tcgtaacaac	1500
tccgccccat tgacgcaaat gggcggtagg cgtgtacggt gggaggctta tataagcaga	1560
gctcgtttag tgaaccgtca gatcgcttgg agacgccatc cacgctgttt tgacctccat	1620
agaagacacc gggaccgatc cagcctccgg actctagagg atcgaaccct taaggccacc	1680
atggatgagg ccagcggttc cggacgggct gacgattgg acgattttga tctggatatg	1740
ctgggaagtg acgcctcga tgattttgac cttgacatgc ttggttcgga tgccttgat	1800
gactttgacc tcgacatgct cggcagtgc gcccttgatg atttcgacct ggacatgctg	1860
attaactcta gaagtccgg atctccgaaa aagaaacgca aagtgggtgg cggttccggc	1920
ggagggtcga tcatgggccc caagaaaaaa cgcaaggtgg ccgcagcaga ctataaggat	1980
gacgacgata aggggatcca tgggtgctcct gctgcagata aaaaatacag catcggcctg	2040
gctatcgga ctaactccgt cggctgggcc gtcatcacc acgaatacaa agtacctagc	2100
aaaaagtcca aggtgcttgg caacacagat cgccactcaa tcaagaaaaa cettatcgga	2160
gccctgctgt ttgactcagg cgaaaccgcc gaggctacac gcctgaaaag aacagctaga	2220
cggcgttaca ccagaaggaa gaaccggatc tgttatcttc aggagatfff ctccaatgag	2280
atggctaagg tggacgattc tttcttccat cgactcgaag aatctttctt ggtggaggaa	2340
gataagaaac acgagaggca tctatffff gaaacattg tcgatgaagt ggcctatcat	2400
gagaaatacc ccacgatcta ccatctgca aaaaagttgg ttgactctac cgacaaggcg	2460
gacctgaggc ttatttatct ggccctggcc catatgatca aattcagggg gcacttcttg	2520
atcgaggggg accttaatcc cgacaactct gacgtgata agttgttcat acagcttgtg	2580
cagacctaca accagctgtt cgaggagaat ccaatcaacg ccagcggagt ggacgctaaa	2640
gccattctga gcgcgagatt gagcaagtct agaagattgg aaaaccttat agcccagctg	2700
ccaggtgaga agaagaacgg actgtttggc aatctcattg cgcttagcct cggactcacc	2760
ccgaacttca aatccaactt cgacctgcc gaagatgcca aattcagct cagtaaggat	2820
acgtatgacg atgatcttga caatctgctg gcgcagatcg gggaccagta cgccgatctt	2880
ttcttggcag caaaaaatct ctcagatgca atactctgt cagacatact gcgagttaat	2940
accgagatta ctaaggctcc gctttctgcc tccatgatca agcgctacga tgagcatcac	3000
caggatctga cactgttgaa agccctgggt cgccaacagc tgccagagaa atacaaggaa	3060

atcttttttg accagtccaa gaatggctac gcaggataca tcgatggagg agccagtcag 3120

gaggaatfff acaagtttat taagcctatc ctggagaaga tggatggtac cgaagaactc 3180

ctggtcaage tcaaccgaga agatttgctt cgcaagcaaa ggacttttga caacggctec 3240

attccgcatc agattcatct gggcgagctg catgccattc tgcgaagaca ggaggatfff 3300

taccatttc tgaaggacaa ccgagagaag atcgagaaaa tactgacatt caggatacca 3360

tattacgtgg gtccactcgc caggggcaac tcccgatcgc cctggatgac aaggaaaagc 3420

gaagagacga tcaactcatg gaacttcgag gaggtcgtgg acaagggggc ctccgcgcag 3480

agctttatcg agaggatgac gaactttgac aaaaatctcc ctaacgagaa ggtgctgcca 3540

aaacattctc tgctctacga gtatttcacc gtttataatg agctcacaaa ggtgaagtac 3600

gtgaccgaag ggatgcggaa gcccgtfff ctgtccggag agcagaagaa ggctatcgtg 3660

gatttgctct ttaagactaa ccgcaaggta acagtcaagc agctgaagga agactacttc 3720

aagaagatcg aatgcttgtc ctacgaaacg gaaatcttga cagttgagta cgggctcctg 3780

ccaatcggga agatagtaga gaagaggatt gaatgtaccg tctattctgt tgataacaac 3840

ggtaacatat acaccagcc cgctgcccga tggcacgacg gcggtgagca ggagggttcc 3900

gaatactgtc tggaggacgg gtcattgatt cgggcgacta aggaccataa gtttatgacg 3960

gtagacggcc agatgttgcc catagatgag atctttgagc gggaactcga ctgatgaga 4020

gtcgataatc ttctaatta gcttaagggt tcgatcccta ctggttagta atgagtttaa 4080

acgggggagg ctaactgaaa cacggaagga gacaataccg gaaggaacct gcgctatgac 4140

ggcaataaaa agacagaata aaacgcacgg gtgttgggtc gtttgttcat aaacgcgggg 4200

ttcggtccca gggctggcac tctgtcgata ccccaccgag accccattgg ggccaatagc 4260

cccgcgttcc ttcttttcc ccaccccacc ccccaagttc ggggtgaagc ccagggtcgc 4320

cagccaactg cggggcgcca ggccctgcca tagcagatct gcgctgattt tntaggtaac 4380

cacgtgcgga ccgagcgccc gcaggaacct ctagtgatgg agttggccac tcctctctg 4440

cgctctcgtc cgctcactga ggccgggcca ccaaaggctc cccgacccc gggttttgcc 4500

cgggcggcct cagtgagcga gcgagcgcgc agctgcctgc aggcttgat cccaatggcg 4560

cgccgagctt ggctcgagca tggatcatagc tgtttcctgt gtgaaattgt tatccgctca 4620

caattccaca caacatacga gccggaagca taaagtgtaa agcctggggt gcctaatgag 4680

tgagctaact cacattaatt gcgttgcgct cactgcccgc tttccagtcg ggaaacctgt 4740

ctgcccagct gcattaatga atcgccaac gcgcggggag aggcggtttg cgtattgggc 4800

gctcttccgc ttctctgctc actgactcgc tgcgctcggg cgttcggctg cggcgagcgg 4860
tatcagctca ctcaaaggcg gtaatcggg tatccacaga atcaggggat aacgcaggaa 4920
agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc gcgttgctgg 4980
cgtttttcca taggctccgc cccctgacg agcatcaca aaatcgacgc tcaagtcaga 5040
ggtggcgaaa cccgacagga ctataaagat accaggcgtt tcccctgga agctccctcg 5100
tgcgctctcc tgttcgacc ctgcccgtta ccggatacct gtccgccttt ctcccttcgg 5160
gaagcgtggc gctttctcat agctcacgct gtaggtatct cagttcgggtg taggtcgttc 5220

gctccaagct gggctgtgtg cacgaacccc cgttcagcc cgaccgctgc gccttatccg 5280
gtaactatcg tcttgagtc aaccggtaa gacacgactt atcgccactg gcagcagcca 5340
ctggtaacag gattagcaga gcgaggtatg taggcggtgc tacagagttc ttgaagtgg 5400
ggcctaacta cggctacact agaagaacag ttttggtat ctgctctctg ctgaagccag 5460
ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa acaaccacc gctggtagcg 5520
gtggtttttt tgtttgcaag cagcagatta cgcgcagaaa aaaaggatct caagaagatc 5580
ctttgatctt ttctacgggg tctgacgctc agtggaaacga aaactcacgt taagggattt 5640

tggtcatgag attatcaaaa aggatcttca cctagatcct tttaaattaa aatgaagtt 5700
ttaaatcaat ctaaagtata tatgagtaaa ctgggtctga cagttagaaa aactcatcga 5760
gcatcaaatg aaactgcaat ttattcatal caggattatc aataccatat ttttgaaaa 5820
gccgtttctg taatgaagga gaaaactcac cgaggcagtt ccataggatg gcaagatcct 5880
ggtatcggtc tgcgattccg actcgtccaa catcaataca acctattaat ttcccctcgt 5940
caaaaataag gttatcaagt gagaaatcac catgagtac gactgaatcc ggtgagaatg 6000
gcaaaagttt atgcatttct ttccagactt gttcaacagg ccagccatta cgctcgtcat 6060

caaaatcact cgcatcaacc aaaccgttat tcattcgtga ttgcgcctga gcgagacgaa 6120
atacgcgac gctgttaaaa ggacaattac aaacaggaat cgaatgcaac cggcgcagga 6180
aactgccag cgcatcaaca atattttcac ctgaatcagg atattcttct aatacctgga 6240
atgctgtttt cccagggatc gcagtgggtga gtaaccatgc atcatcagga gtacggataa 6300
aatgcttgat ggtcgggaaga ggcataaatt ccgtcagcca gtttagtctg accatctcat 6360
ctgtaacatc attggcaacg ctacctttgc catgtttcag aaacaactct ggcgcatcgg 6420
gcttccata caatgatag attgtcgac ctgattgccc gacattatcg cgagccatt 6480

tataccata taaatcagca tccatgttg aatttaacg cggcctagag caagacgttt 6540
cccgttgaat atggctcata ctcttcttt ttcaatatta ttgaagcatt tatcagggtt 6600
attgtctcat gacggatc atatttgaat gtatttagaa aaataaaca ataggggttc 6660

cgcgcacatt tccccgaaaa gtgccacctg acgtctaaga aaccattatt atcatgacat 6720
 taacctataa aaataggcgt atcacgaggc cctttcgtc 6759

<210> 336
 <211> 7341
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 336

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
 cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc 240
 attcgccatt caggctgcbc aactgttggg aagggcgatc ggtgcgggcc tcttcgctat 300
 tacgccagct ggcgaaaagg ggatgtgctg caagggcatt aagtgggta acgccagggt 360
 tttcccagtc acgacgttgt aaaacgacgg ccagtgaatt gacgcgcat tgggatgttg 420

taaaacgacg gccagtgaac ctgcaggcag ctgcgcctc gctcgtcac tgaggccgcc 480
 cgggcaaage ccgggcgtcg ggcgacctt ggtcgcccgg cctcagttag cgagcgagcg 540
 cgagagagg gaggggccaa ctccatcact aggggttctt gcggccgcac gcgtggagga 600
 gggcctatth cccatgattc ctccatattt gcatatacga tacaaggctg ttagagagat 660
 aattagaatt aatttgactg taaacacaaa gatattagta caaaatcgt gacgtagaaa 720
 glaataatth ctgggtagt ttgcagtttt aaaattatgt tttaaatgg actatcatat 780
 gcttaccgta acttgaaagt atttcgattt ctggcttta tatatcttg ggaaaggacg 840

aaacaccggt ttagagcta gaaatagcaa gttaaaataa ggctagtccg ttatcaactt 900
 gaaaaagtgg caccgagtcg gtgctttttt gctagcctag accagcttt ctgtacaaa 960
 gtggcatta atacgcgttg acattgatta ttgactagtt attaatagta atcaattacg 1020
 gggtcattag tccatagccc atatatggag ttccgcgtta cataacttac ggtaaatggc 1080
 ccgcctggct gaccgcccga cgacccccgc ccattgacgt caataatgac gtatgttccc 1140
 atagtaacgc caatagggac tttccattga cgtcaatggg tggagtatth acggtaaact 1200
 gccacttgg cagtacatca agtgtatcat atccaagta cgccccctat tgacgtcaat 1260

gacggtaaat ggcccgcctg gcattatgcc cagtacatga ccttatggga ctttctact 1320
 tggcagtaca tctacgtatt agtcatcgct attaccatgg tgatgcggtt ttggcagtac 1380

atcaatgggc gtggatagcg gtttgactca cggggatttc caagtctcca cccattgac 1440
 gtcaatggga gtttgtttt gcacaaaat caacgggact ttccaaaatg tcgtaacaac 1500
 tccgccccat tgacgcaaat gggcggtagg cgtgtacggt gggaggtcta tataagcaga 1560
 gctcgtttag tgaaccgtca gatcgcttgg agacgccatc cacgctgttt tgacctccat 1620
 agaagacacc gggaccgatc cagcctccgg actctagagg atcgaacct taaggccacc 1680

 atggatccga aaaagaaacg caaagttggt agccagtacc tgcccgacac cgacgaccgg 1740
 caccggatcg aggaaaagcg gaagcggacc tacgagacat tcaagagcat catgaagaag 1800
 tcccccttca gcggccccac gcacctaga cctccaccta gaagaatcgc cgtgccccagc 1860
 agatccagcg ccagcgtgcc aaaactgcc ccccagcctt accccttcac cagcagcctg 1920
 agcaccatca actacgacga gttccctacc atggtgttcc ccagcggcca gatctctcag 1980
 gcctctgctc tggtccagc cctctctcag gtgctgctc aggctcctgc tctgcacca 2040
 gctccagcca tgggtgtctg actggctcag gcaccagcac ccgtgcctgt gctggctcct 2100

 ggacctccac aggctgtggc tccaccagcc cctaaaccta cacaggccgg cgagggcaca 2160
 ctgtctgaag ctctctgca gctgcagttc gacgacgagg atctgggagc cctgctggga 2220
 aacagcaccg atctgacct gttcaccgac ctggccagcg tggacaacag cgagttccag 2280
 cagctgctga accagggcat cctgtggcc cctcacacca ccgagcccat gctgatggaa 2340
 taccgccagg ceatcacccg gctcgtgaca ggcgctcaga ggccctctga tccagctcct 2400
 gcccctctgg gagcaccagg cctgccta at ggactgtgt ctggcgacga ggacttcagc 2460
 tctatcgccg atatggattt ctcagccttg ctgggctctg gcagcggcag catcatgggc 2520

 cccaagaaaa aacgcaaggt ggcccagca gactataagg atgacgacga taaggggatc 2580
 catggtgtgc ctgctcaga taaaaatac agcatcggcc tggtatcgg aactaactcc 2640
 gtcggtggg ccgtcattac cgacgaatac aaagtaaccta gcaaaaagtt caaggtgctt 2700
 ggcaacacag atgccactc aatcaagaaa aaccttatcg gagcctgct gtttgactca 2760
 ggcgaaaccg ccgaggctac acgctgaaa agaacagcta gacggcggta caccagaagg 2820
 aagaaccgga tctgttatct tcaggagatt ttctccaatg agatggctaa ggtggacgat 2880
 tctttcttc atcgactcga agaactttc ttggtggagg aagataagaa acacgagagg 2940

 catcctattt tcggaaacat tgtcgatgaa gtggcctatc atgagaaata ccccacgatc 3000
 taccatctgc gaaaaaagtt gttgactct accgacaagg cggacctgag gcttatttat 3060
 ctggccctgg cccatatgat caaatcagg gggcacttct tgatcgaggg ggacctaat 3120
 cccgacaact ctgacgtgga taagttgttc atacagcttg tgcagaccta caaccagctg 3180
 ttcgaggaga atccaatcaa cgccagcggg gtggacgcta aagccattct gagcgcgaga 3240

ttgagcaagt ctagaagatt ggaaaacctt atagcccagc tgccaggtga gaagaagaac 3300
 ggactgtttg gcaatctcat tgcgcttagc ctcgactca ccccgaactt caaatccaac 3360

 ttcgacctcg ccgaagatgc caaattgcag ctcagtaagg atacctatga cgatgatctt 3420
 gacaatctgc tggcgcagat cggggaccag tacgccgatc ttttcttggc agcaaaaaat 3480
 ctctcagatg caatactctt gtcagacata ctgagagtta ataccgagat tactaaggct 3540
 ccgctttctg cctccatgat caagecgtac gatgagcacc accaggatct gacactgttg 3600
 aaagccctgg tgcgccaaca gctgccagag aaatacaagg aaatcttttt tgaccagtcc 3660
 aagaatggct acgcaggata catcgatgga ggagccagtc aggaggaatt ttacaagttt 3720
 attaagccta tcttgagaaa gatggatggt accgaagaac tcttgggtcaa gctcaaccga 3780

 gaagatttgc ttcgaagca aaggactttt gacaacggct ccattccgca tcagattcat 3840
 ctgggcgagc tgcattccat tctgcgaaga caggaggatt tttaccatt tctgaaggac 3900
 aaccgagaga agatcgagaa aatactgaca ttcaggatac catattacgt gggtccactc 3960
 gccaggggca actcccatt cgcctggatg acaaggaaaa gcgaagagac gatcactcca 4020
 tggaaacttc aggaggctgt ggacaagggg gcctccgcgc agagctttat cgagaggatg 4080
 acgaactttg acaaaaatct ccctaacgag aagggtgctgc caaacattc tctgctctac 4140
 gagtatttca ccgtttataa tgagctcaca aagggtgaagt acgtgaccga agggatgcgg 4200

 aagcccgtt tctgtccgg agagcagaag aaggctatcg tggatttgc ctttaagact 4260
 aaccgcaagg taacagtcaa gcagctgaag gaagactact tcaagaagat cgaatgcttg 4320
 tcttacgaaa cggaaatctt gacagttgag tacgggctcc tgccaatcgg gaagatagta 4380
 gagaagagga ttgaatgtac cgtctattct gttgataaca acggtaacat atacaccag 4440
 cccgtcgccc aatggcacga tgcggtgag caggaggtgt tcgaatactg tctggaggac 4500
 gggtcattga ttcgggcgac taaggacat aagtttatga cggtagacgg ccagatgttg 4560
 cccatagatg agatcttga gcgggaactc gacttgatga gactcgataa tcttcctaat 4620

 tagcttaagg gttcgatccc tactggttag taatgagttt aaacggggga ggctaactga 4680
 aacacggaag gagacaatac cggaaggaac ccgcgctatg acggcaataa aaagacagaa 4740
 taaaacgcac ggggtgttggc tcgtttgttc ataaacgcgg ggttcggtcc cagggtcggc 4800
 actctgtcga taccaccgg agaccctatt ggggccaata cgcccgcgtt tcttctttt 4860
 cccacccca cccccaagt tcgggtgaag gccagggtc gcgagccaac gtcggggcgg 4920
 caggccctgc catagcagat ctgcctgat tttgtaggta accacgtgcg gaccgagcgg 4980

ccgcaggaac ccctagtgat ggagttggcc actccctctc tgcgcgctcg ctcgctcact 5040

 gaggccgggc gaccaaaggt cgccccgacgc cgggctttg cccgggcggc ctcagtgagc 5100
 gagcgagcgc gcagctgcct gcaggcttgg atcccaatgg cgcgccgagc ttggctcgag 5160
 catggtcata gctgtttcct gtgtgaaat gttatccgct cacaattcca cacaacatac 5220
 gagccggaag cataaagtgt aaagcctggg gtgcctaag agtgagctaa ctcacattaa 5280
 ttgcgttgcg ctactgccc gctttccagt cgggaaacct gtcgtgccag ctgcattaat 5340
 gaatcgcca acgcgcgggg agaggcgggt tgcgtattgg gcgctcttc gttcctcgc 5400
 tctactgactc gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg 5460

 cggtaatcgc gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaag 5520
 gccagcaaaa gcccaggaac cgtaaaaagg ccgcgttgct ggcgttttc catagctcc 5580
 gccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aaccgcagag 5640
 gactataaag ataccaggcg tttccccctg gaagctccct cgtgcgctct cctgttccga 5700
 ccctgccgct taccggatac ctgtccgct tttcccttc gggaagcgtg gcgctttctc 5760
 atagctcacg ctgtaggat ctcagttcgg ttaggtcgt tcgctccaag ctgggctgtg 5820
 tgcacgaacc cccggttcag cccgaccgct gcgccttacc cggtaactat cgtcttgagt 5880

 ccaaccgggt aagacacgac ttatgccac tggcagcagc cactggtaac aggattagca 5940
 gagcgaggta tgiaggcggg gctacagagt tcttgaagtg gtggcctaac tacggctaca 6000
 ctagaagaac agtatttggg atctgcgctc tgctgaagcc agttaccttc ggaaaaagag 6060
 ttggtagctc ttgatccggc aaacaaacca ccgctggtag cggtggtttt tttgtttgca 6120
 agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc ttttctacgg 6180
 ggtctgacgc tcagtggaac gaaaactcac gttaagggat tttggtcatg agattatcaa 6240
 aaaggatctt cacctagatc cttttaaatt aaaaatgaag ttttaaatca atctaaagta 6300

 tatatgagta aacttggctc gacagttaga aaaactcacc gagcatcaaa tgaactgca 6360
 atttattcat atcaggatta tcaataccat atttttgaaa aagccgttcc tgtaatgaag 6420
 gagaaaaactc accgaggcag ttccatagga tggcaagatc ctggtatcgg tctgcgattc 6480
 cgactcgtcc aacatcaata caacctatta atttcccctc gtcaaaaata aggttatcaa 6540
 gtgagaaatc accatgagtg acgactgaat ccggtgagaa tggcaaaagt ttatgcattt 6600
 ctttccagac ttgttcaaca ggccagccat tacgctcgtc atcaaaatca ctcgcatcaa 6660
 ccaaaccgtt attcattcgt gattgcgctc gagcgagacg aaatacgcga tcgctgttaa 6720

aaggacaatt acaaacagga atcgaatgca accggcgag gaacactgcc agcgcatcaa 6780
 caatattttc acctgaatca ggatattctt ctaatactg gaatgctgtt ttcccagga 6840
 tcgcagtggg gagtaacat gcatcatcag gactacggat aaaatgcttg atggtcggaa 6900
 gaggcataaa ttccgtcagc cagtttagtc tgaccatctc atctgtaaca tcattggcaa 6960
 cgctaccttt gccatgtttc agaaacaact ctggcgcatc gggcttcca tacaatcgt 7020
 agattgtcgc acctgattgc ccgacattat cgcgagccca tttatacca tataaatcag 7080
 catccatggt ggaatttaat cgcggcctag agcaagacgt ttcccgttga atatggctca 7140

tactcttct tttcaatat tattgaagca tttatcaggg ttattgtctc atgagcggat 7200
 acatatttga atgtatttag aaaaataaac aaataggggt tccgcgcaca tttcccga 7260
 aagtgccacc tgactctaa gaaaccatta ttatcatgac attaactat aaaaatagc 7320
 gtatcacgag gccctttcgt c 7341

<210> 337

<211> 5751

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 337

ctgcgcgctc gtcgctcac tgaggccgcc cgggcaaagc cgggctgc ggcgaccttt 60

ggtcgcccgg cctcagtgag cgagcgagcg cgcagagagg gaggggcaa ctccatcact 120

aggggttctt tgiagttaat gattaacccg ccattgctact tatctacgta gccatgctct 180

aggaagatcg gaattcgccc ttaagaaggc ctccacggcc actagtcttt cgtcttcaag 240

aattcctcga gtttactccc taccagtgat agagaacgta tgaagagttt actccctatc 300

agtgatagag aacgtatgca gactttactc cctatcagtg atagagaacg tataaggagt 360

ttactcccta tcagtgatag agaactgatg accagtttac tcctatcag tgatagagaa 420

cgatctaca gtttactccc taccagtgat agagaacgta tatccagttt actccctatc 480

agtgatagag aacgtataag ctttaggcgt gtacgggtggg tttccatga ttcttcata 540

tttgcatata cgatacaagg ctgttagaga gataattgga attaatttga ctgtaaacac 600

aaagatatta gtacaaaata cgtgacgtag aaagtaataa tttcttgggt agtttgcagt 660

tttaaaatta tgtttaaaa tggactatca tatgcttacc gtaacttgaa agtatttca 720

ttcttggct ttatatatct tgtggaaagg acgaaacacc ggttttagta ctctggaaac 780

agaatctact aaaacaaggc aaaatgccgt gtttatctcg tcaacttgtt ggcgagattt 840

ttgaattctc gacctcgaga caaatggcag cgttgacatt gattattgac tagttattaa 900

 tagtaatcaa ttacggggtc attagtccat agcccatata tggagtccg cgttacataa 960
 cttacggtaa atggcccgc tggctgaccg cccaacgacc cccgcccatt gacgtcaata 1020
 atgacgtatg ttcccatagt aacccaata gggactttcc attgacgtca atgggtggag 1080
 tatttacggt aaactgccca cttggcagta catcaagtgt atcatatgcc aagtacgcca 1140
 cctattgacg tcaatgacgg taaatggccc gcctggcatt atgccagta catgacctta 1200
 tgggactttc ctacttgga gtacatctac gtattagtca tcgctattac catggatgatg 1260
 cggttttggc agtacatcaa tgggcgtgga tagcggtttg actcacgggg atttccaagt 1320

 ctccacccca ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg ggactttcca 1380
 aaatgtcgt acaactccgc cccattgacg caaatgggcg gtaggcgtgt acgggtgggag 1440
 gtctataata gcagagctct ctggctaact cttaaggata tcgccacat ggctagatta 1500
 gataaaagta aagtattaa cagcgatta gagctgctta atgaggtcgg aatcgaaggt 1560
 ttaacaacc gtaaacctgc ccagaagcta ggtgtagagc agcctacatt gtattggcat 1620
 gtaaaaaata agcgggcttt gctcgacgc ttagccattg agatgttaga taggcacat 1680
 actcactttt gccctttaga aggggaaagc tggcaagatt ttttacgtaa taacgctaaa 1740

 agttttagat gtgctttact aagtcacgc gatggagcaa aagtacattt aggtacacgg 1800
 cctacagaaa aacagtatga aactctcga aatcaattag cctttttatg ccaacaaggt 1860
 ttttactag agaatgcat atatgcactc agcgtgtgg ggcattttac tttaggttgc 1920
 gtattggaag atcaagagca tcaagtgcct aaagaagaaa gggaaacacc tactactgat 1980
 agtatgccgc cattattacg acaagctatc gaattatttg atcaccaagg tgcagagcca 2040
 gccttcttat tggccttga attgatcata tgcggattag aaaaacaact taaatgtgaa 2100
 agtgggtcgc caaaaaagaa gagaaaggtc gacggcggtg gtgctttgtc tectcagcac 2160

 tctgctgtca ctcaaggaag tatcatcaag aacaaggagg gcatggatgc taagtcacta 2220
 actgcctggt cccggacact ggtgacctc aaggatgtat ttgtggactt caccagggag 2280
 gagtggaaagc tgctggacac tgctcagcag atcgtgtaca gaaatgtgat gctggagaac 2340
 tataagaacc tggtttcctt gggttatcag ctactaagc cagatgtgat cctccggttg 2400
 gagaaggag aagagccctg gctggtggag agagaaattc accaagagac ccatcctgat 2460
 tcagagactg catttgaat caaatcatca gtttgaggat ccagatctgc ctgactgtg 2520
 ccttctagtt gccagccatc tgttgttgc cctccccg tgccttctt gaccctggaa 2580

ggtgccactc ccaactgtcct ttcctaataa aatgaggaaa ttgcatcgca ttgtctgagt 2640
 aggtgtcatt ctattctggg gggtaggggtg gggcaggaca gcaaggggga ggattgggaa 2700
 gacaatagca ggcattgctgg ggactcgagt taagggcgaa ttcccataa ggatcttctt 2760
 agagcatggc tacgtagata agtagcatgg cgggttaatc attaactaca aggaaccctt 2820
 agtgatggag ttggccactc cctctctgcg cgctcgctcg ctactgagg cgggcgacc 2880
 aaaggtcgcc cgacgcccg gctttgcccg ggcggcctca gtgagcgagc gagcgcgag 2940
 ccttaattaa cctaattcac tggccgtcgt ttacaacgt cgtgactggg aaaaccctgg 3000

 cgttacceaa ctaaatcgcc ttgcagcaca tcccccttc gccagctggc gtaatagcga 3060
 agaggcccg accgatcgcc ctcccaaca gttgcgcagc ctgaatggcg aatgggacgc 3120
 gccctgtagc ggcgattaa gcgcggcggg tgtggtggtt acgcgcagcg tgaccgctac 3180
 acttgccagc gccctagcgc ccgctcctt cgtttcttc ccttcttctc tcgccagtt 3240
 cgccggcttt ccccgtaag ctctaaatcg ggggctcctt ttagggttcc gatttagtgc 3300
 ttacggcac ctgacccca aaaaacttga ttaggtgat ggttcacgta gtggccatc 3360
 gccctgatag acggttttcc gcccttgac gttggagtcc acgttcttta atagtggact 3420

 ctgttccaa aciggaacaa cactcaacc tatctcggtc tattcttttg atttataagg 3480
 gattttgccg atttcggcct attggttaaa aaatgagctg atttaacaaa aatttaacgc 3540
 gaattttaac aaaaatttaa cgtttataat tcagggtggc atctttcggg gaaatgtgag 3600
 cggaaaccct atttgttat ttttctaaat acattcaat atgtatccgc tcatgagaca 3660
 ataaccctga taaatgcttc aataatattg aaaaaggaag agtatgagta ttcaacattt 3720
 ccgtgtcgcc cttattcctt ttttgcggc attttgcctt cctgtttttg ctaccaccaga 3780
 aacgtgggtg aaagtaaaag atgctgaaga tcagttgggt gcacgagtgg gttacatcga 3840

 actggatctc aatagtggta agatccttga gagttttcgc cccgaagaac gttttccaat 3900
 gatgagcact tttaaagtcc tgctatgtgg cgcggtatta tcccgtattg acgccgggca 3960
 agagcaacte ggtcgccgca tacactattc tcagaatgac ttggttgagt actcaccagt 4020
 cacagaaaag catcttacgg atggcatgac agtaagagaa ttatgcagtg ctgccataac 4080
 catgagtgat aacctgcgg ccaacttact tctgacaacg atcggaggac cgaaggagct 4140
 aaccgctttt ttgcacaaca tgggggatca tgtaactcgc cttgatcgtt gggaaccgga 4200
 gctgaatgaa gccataccaa acgacgagcg tgacaccacg atgctgtag taatggtaac 4260

 aacgttgccg aactattaa ctggcgaact acttactcta gttcccggc aacaattaat 4320
 agactggatg gagggcgata aagttgcagg accactctg cgctcgcccc ttccggctgg 4380
 ctggtttatt gctgataaat ctggagccgg tgagcgtggg tctcgcgta tcattgcagc 4440

actggggcca gatggtaagc cctcccgtat cgtagttatc tacacgacgg ggagtccaggc 4500
 aactatggat gaacgaaata gacagatcgc tgagataggt gcctcactga ttaagcattg 4560
 gtaactgtca gaccaagttt actcatatat acttttagatt gatttaaaac ttcattttta 4620
 atttaaaagg atctagggtga agatcctttt tgataatctc atgaccaaaa tcccttaacg 4680

tgagttttcg ttccactgag cgtcagaccc cgtagaaaag atcaaaggat cttcttgaga 4740
 tccttttttt ctgcgcgtaa tctgetgctt gcaaacaaaa aaaccaccgc taccagcggg 4800
 ggtttgtttg ccggatcaag agctaccaac tctttttccg aaggtaactg gcttcagcag 4860
 agcgcagata ccaataactg tccttctagt gtagccgtag ttaggccacc acttcaagaa 4920
 ctctgtagca ccgcctacat acctcgctct gctaatctg ttaccagtgg ctgctgccag 4980
 tggcgataag tcgtgtctta ccgggttggg ctcaagacga tagttaccgg ataaggcgca 5040
 gcggtcgggc tgaacggggg gttcgtgcac acagcccagc ttggagcgaa cgacctacac 5100

cgaactgaga tacctacagc gtgagctatg agaaagcgc acgcttcccg aaggagaaaa 5160
 ggccggacagg tatccggtaa gcggcagggt cggaacagga gagcgcacga gggagcttcc 5220
 agggggaaaac gcttggatc tttatagtcc tgtcgggttt cgccacctct gacttgagcg 5280
 tcgatttttg tgatgctcgt caggggggcg gacccatgg aaaaacgcca gcaacgcggc 5340
 ctttttacgg ttcttggcct tttgctgcgg ttttgcctac atgttctttc ctgcgttacc 5400
 ccctgattct gtggataacc gtattaccgc ctttgagtga gctgataacc ctgcgccag 5460
 ccgaacgacc gagcgcagcg agtcagttag cgaggaagcg gaagagcgcc caatacga 5520

accgctctc cccgcgcgtt ggccgattca ttaatgcagc tggcagcaca ggtttcccga 5580
 ctggaaagcg ggcagtgagc gcaacgcaat taatgtgagt tagctcactc attaggcacc 5640
 ccaggcttta cactttatgc ttccggctcg tatgttgtgt ggaattgtga gcggataaca 5700
 atttcacaca ggaacagct atgacatga ttacccaga ttttaattaag g 5751

<210> 338
 <211> 7317
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic polynucleotide
 <400> 338

tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60

cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
 ttggcgggtg tcggggctgg cttactatg cggcatcaga gcagattgta ctgagagtgc 180

accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc 240
 attcgccatt caggctgcgc aactgttggg aagggcgatc ggtgcgggcc tcttcgctat 300
 tacgccagct ggcgaaaggg ggatgtgctg caaggcgatt aagtgggta acgccagggt 360
 tttcccagtc acgacgttgt aaaacgacgg ccagtgaatt gacgcgccat tgggatgttg 420
 taaaacgacg gccagtgaac ctgcaggcag ctgcgcgctc gctcgtcac tgaggccgcc 480

 cgggcaaage cggggcgtcg ggcgaccttt ggtcgcccgg cctcagttag cgagcgagcg 540
 cgcagagagg gagtggccaa ctccatcact aggggttctt gcggccgcac gcgtggagga 600
 gggcctatth ccgatgattc tttcatatth gcatatacga tacaaggctg ttagagagat 660
 aattagaatt aatttgactg taaacacaaa gatattagta caaatacgt gacgtagaaa 720
 gtaataatth ctgggtagt ttgcagttth aaaattatgt tttaaatgg actatcatat 780
 gcttaccgta acttgaaagt atttcgattt ctggcttta tatacttgt ggaaaggacg 840
 aaacaccggt ttagagcta gaaatagcaa gttaaaataa ggctagtccg ttatcaactt 900

 gaaaaagtgg caccgagtcg gtgctttttt gctagcctag acccagcttt ctgtacaaa 960
 gtggcatta atacggttg acattgatta ttgactagtt attaatagta atcaattacg 1020
 gggtcattag ttcatagccc atatatggag ttccgcgtta cataacttac ggtaatggc 1080
 ccgcctggct gaccgcccc aaccgcccgc ccattgacgt caataatgac gtatgttccc 1140
 atagtaacgc caatagggac tttccattga cgtcaatggg tggagtatth acggtaaac 1200
 gccacttgg cagtacatca agtgtatcat atgccaagta cggcccctat tgacgtcaat 1260
 gacggtaaat ggcccgcctg gcattatgcc cagtacatga ccttatggga ctttctact 1320

 tggcagtaca tctacgtatt agtcatcgct attaccatgg tgatcggtt ttggcagtac 1380
 atcaatgggc gtggatagcg gtttgactca cggggatttc caagctcca cccattgac 1440
 gtcaatggga gttgttttg gcacaaaaat caacgggact ttccaaaatg tcgtaacaac 1500
 tccgccccat tgacgcaaat gggcggtagg cgtgtacggt gggaggctta tataagcaga 1560
 gctcgtttag tgaaccgtca gatcgcttgg agacgccatc cacgctgttt tgacctcat 1620
 agaagacacc gggaccgata cagcctccgg actctagagg ategaacct taaggccacc 1680
 atgggcccc aaaaaaacg caaggtggcc gcagcagact ataaggatga cgacgataag 1740

 gggatccatg gtgtgcctgc tgcagataaa aaatacagca tcggcctgga tatcggaact 1800
 aactccgtcg gctgggcccgt cattaccgac gaatacaaaag tacctagcaa aaagtccaag 1860
 gtgcttggca acacagatcg ccaactcaatc aagaaaaacc ttatcggagc cctgctgttt 1920
 gactcaggcg aaaccgcccg ggctacacgc ctgaaaaaaa cagctagacg gcggtacacc 1980
 agaaggaaga accggatctg ttatcttcag gagatthtct ccaatgagat ggctaaggtg 2040

gacgattctt tcttccatcg actcgaagaa tctttcttgg tggaggaaga taagaaacac 2100
gagaggcatc ctatcttcgg aaacattgtc gatgaagtgg cctatcatga gaataacccc 2160

acgatctacc atctgcgaaa aaagtgggtt gactctaccg acaaggcggga cctgaggctt 2220
atctatctgg ccctggccca tatgatcaaa ttcagggggc acttcttgat cgagggggac 2280
cttaatcccg acaactctga cgtggataag ttgttcatac agcttgtgca gacctacaac 2340
cagctgttcg aggagaatcc aatcaacgcc agcggagtgg acgctaaagc cattctgagc 2400
gcgagattga gcaagtctag aagattggaa aaccttatag cccagctgcc aggtgagaag 2460
aagaacggac tgtttggcaa tctcattgcg cttagcctcg gactcacccc gaacttcaaa 2520
tccaacttcg acctcgccga agatgccaaa ttgcagctca gtaaggatac gtatgacgat 2580

gatcttgaca atctgctggc gcagatcggg gaccagtacg ccgatctttt ctggcagca 2640
aaaaatctct cagatgcaat actcttgtca gacatactgc gagttaatac cgagattact 2700
aaggctccgc tttctgcctc catgatcaag cgctacgatg agcatcacca ggatctgaca 2760
ctgttgaaag ccctgggtcg ccaacagctg ccagagaaat acaaggaaat cttttttgac 2820
cagtccaaga atggctacgc aggatacatc gatggaggag ccagtcagga ggaattttac 2880
aagtttatta agcctatcct ggagaagatg gatggtaccg aagaactcct ggtcaagctc 2940
aaccgagaag atttgcttcg caagcaaagg acttttgaca acggctccat tccgcatcag 3000

atctatctgg gcgagctgca tgccattctg cgaagacagg aggattttta cccatttctg 3060
aaggacaacc gagagaagat cgagaaaata ctgacattca ggataccata ttactgggtt 3120
ccactcgcca ggggcaactc ccgattcgcc tggatgacaa ggaaaagcga agagacgatc 3180
actccatgga acttcgagga ggtcgtggac aagggggcct ccgcgagag ctttatcgag 3240
aggatgacga actttgacaa aaatctcct aacgagaagg tgctgcaaaa acattctctg 3300
ctctacgagt atttcacgt ttataatgag ctcaaaaagg tgaagtacgt gaccgaaggg 3360
atgcggaagc ccgcttttct gtccggagag cagaagaagg ctatcgtgga tttgctcttt 3420

aagactaacc gcaaggtaac agtcaagcag ctgaaggaag actacttcaa gaagatcgaa 3480
tgcttgctct acgaaacgga aatcttgaca gttgagtacg ggctcctgcc aatcggaag 3540
atagtagaga agaggattga atgtaccgtc tattctgttg ataacaacgg taacataac 3600
accagcccc tcgccaatg gcacgatcgc ggtgagcagg aggtgttcga atactgtctg 3660
gaggacgggt cattgattcg ggcgactaag gaccataagt ttatgacggt agacggccag 3720
atgttgccca tagatgagat ctttgagcgg gaactcgact tgatgagagt cgataatctt 3780

cctaattggat ccggcgcaac aaacttctct ctgctgaaac aagccggaga tgtcgaagag 3840

aatcctggac cgatgtctag actggacaag agcaaagtca taaacggcgc tctggaatta 3900

ctcaatggag tcggtatcga aggctgacg acaaggaaac tcgctcaaaa getgggagtt 3960

gagcagccta ccctgtactg gcacgtgaag aacaagcggg ccctgctcga tgccttgcca 4020

atcgagatgc tggacaggca tcataccac ttctgcccc tggaaggcga gtcattggcaa 4080

gactttctgc ggaacaacgc caagtcattc cgtgtgtctc tctctcaca tcgcgacggg 4140

gctaaagtgc atctcggcac ccgccaaca gaaaacagt acgaaaccct ggaaaatcag 4200

ctcgcgttcc tgtgtcagca aggttctcc ctggagaacg cactgtacgc tctgtccgcc 4260

gtgggccact ttacactggg ctgctattg gaggaacagg agcatcaagt agcaaaagag 4320

gaaagagaga cacctaccac cgattctatg cccccacttc tgagacaagc aattgagctg 4380

ttcgaccggc agggagccga acctgccttc ctttctggcc tggaactaat catatgtggc 4440

ctggagaaac agctaaagtg cgaaagcggc gggccggccg acgcccttga cgattttgac 4500

ttagacatgc tccagccga tgccttgac gactttgacc ttgatatgct gcctgtctgac 4560

gctcttgacg attttgacct tgacatgctc cccgggtagc ttaagggttc gatccctact 4620

ggttagtaat gagtttaaac gggggaggct aactgaaaca cggaaggaga caataccgga 4680

aggaaccgcg gctatgacgg caataaaaag acagaataaa acgcacgggt gttgggtcgt 4740

ttgttcataa acgcgggggtt cgggtcccagg gctggcactc tgtcgatacc ccaccgagac 4800

cccatgggg ccaatacgcc cgcgtttctt ctttttcccc accccacccc ccaagttcgg 4860

gtgaaggccc agggctcga gccaacgtcg gggcggcagg ccctgccata gcagatctgc 4920

gctgattttg taggtaacca cgtcgggacc gagcggccgc aggaaccct agtgatggag 4980

ttggccactc cctctctcgc gcctcgtcgc ctactgagg ccgggcgacc aaaggtcgcc 5040

cgacgcccg gctttgcccg ggcggcctca gtgagcgagc gagcgcgag ctgcctgcag 5100

gcttggatcc caatggcgcg ccgagcttgg ctcgagcatg gtcatactg tttcctgtgt 5160

gaaattgtta tccgctcaca attccacaca acatacgagc cggaagcata aagtgtaaag 5220

cctgggtgct ctaatgagtg agctaactca cattaattgc gttgcgctca ctgcccgctt 5280

tccagtcggg aaacctgtcg tgccagctgc attaatgaat cggccaacgc gcggggagag 5340

gcggtttgcg tattgggcgc tcttccgctt cctcgtcac tgactcgtg cgctcggtcg 5400

ttcggctcgc gcgagcggtc tcagctcact caaaggcggg aatacggtta tccacagaat 5460

caggggataa cgcaggaag aacatgtgag caaaaggcca gcaaaaggcc aggaaccgta 5520

aaaaggccgc gttgctggcg tttttccata ggctccgccc ccctgacgag catcacaaaa 5580
 atcgacgctc aagtcagagg tggcgaaacc cgacaggact ataaagatac caggcgtttc 5640
 cccctggaag ctcctcgtg cgctctctg ttcgacacct gccgcttacc ggatacctgt 5700
 ccgcctttct cccttcggga agcgtggcgc tttctcatag ctacgctgt aggtatctca 5760
 gttcggtgta ggtcgttcgc tccaagctgg gctgtgtgca cgaaccccc gttcagcccc 5820
 accgctgcgc cttatccggt aactatcgtc ttgagtccaa cccgtaaga cacgacttat 5880
 cgccactggc agcagccact ggtaacagga ttagcagagc gaggtatgta ggcggtgcta 5940

 cagagtctt gaagtggagg cctaactacg gctacactag aagaacagta tttggtatct 6000
 gcgctctgct gaagccagtt accttcggaa aaagagtgg tagctcttga tccggcaaac 6060
 aaaccaccgc tggtagcggg ggtttttttg tttgcaagca gcagattacg cgcagaaaaa 6120
 aaggatctca agaagatcct ttgatctttt ctacggggtc tgacgctcag tggaacgaaa 6180
 actcagttta agggattttg gtcatgagat tatcaaaaag gatcttcacc tagatccttt 6240
 taaattaa atgaagtttt aatcaatct aaagtatata tgagtaaact tggctgaca 6300
 gttagaaaaa ctcatcgagc atcaaatgaa actgcaatth attcatatca ggattatcaa 6360

 taccatatth ttgaaaaagc cgtttctgta atgaaggaga aaactcaccg aggcagttcc 6420
 ataggatggc aagatcctgg tateggtctg cgattccgac tcgtccaaca tcaatacaac 6480
 ctattaatth ccctcgtca aaaataaggt tatcaagtga gaaatcacca tgagtgcga 6540
 ctgaatccgg tgagaatggc aaaagtttat gcatttcttt ccagacttgt tcaacaggcc 6600
 agccattacg ctgctcatca aaatcactcg catcaaccaa accgttatc attcgtgatt 6660
 gcgcctgagc gagacgaaat acgcgatcgc tgttaaaagg acaattacaa acaggaatcg 6720
 aatgcaaccg ggcaggaac actgccagcg catcaacaat atttcacct gaatcaggat 6780

 attcttctaa tacctggaat gctgttttcc cagggatcgc agtggtagt aacctgcat 6840
 catcaggagt acggataaaa tgcttgatgg tcggaagagg cataaattcc gtcagccagt 6900
 ttagtctgac catctcatct gtaacatcat tggcaacgct accttgcca tgtttcagaa 6960
 acaactctgg cgcategggc ttccataca atcgatagat tgtcgcacct gattgcccga 7020
 cattatcgg agccattta taccatata aatcagcatc catgttgaa ttaatcgcg 7080
 gcctagacga agacgtttcc cgttgaatat ggctcact cttcttttt caatattatt 7140
 gaagcattta tcagggttat tgtctcatga gcggatacat atttgaatgt atttagaaaa 7200

 ataaacaaat aggggttccg cgcacattc cccgaaaagt gccacctgac gtctaagaaa 7260
 ccattattat catgacatta acctataaaa ataggcgtat cacgaggccc tttcgtc 7317

<210> 339

<211> 6192

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 339

```

cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc cggggcaaag cccgggcgct      60
gggcgacctt tggtcgcccc gcctcagtga gcgagcgagc gcgcagagag ggagtggcca      120
actccatcac taggggttcc tgcggccgca cgcgtggaaa aggcctccac ggccactagt      180

ctttcgtctt caagaattcc tcgagtttac tcctatcag tgatagagaa cgtatgaaga      240
gtttactccc tatcagtgat agagaacgta tgcagacttt actccctatc agtgatagag      300
aacgtataag gaggtttact cctatcagtg atagagaacg tatgaccagt ttactcccta      360
tcagtgatag agaacgtatc tacagtttac tcctatcag tgatagagaa cgtatatcca      420
gtttactccc tatcagtgat agagaacgta taagctttag gcgtgtacgg tgggcgccta      480
taaaagcaga gctcgtttag tgaaccgtca gatcgctgg agcaattcca caacactttt      540
gtcttatacc aactttccgt accacttcct acctctgtaa aggtctagag ctagcgaatt      600

cgaatttgcc accatgatta agatcgcaac ccgaaaatac ctgggaaagc agaacgtcta      660
cgatattggt gtagagagag accataactt tgctctgaag aacggcttta ttgcctcatg      720
cttcgacagc gttgagatat ccggcgtgga ggatagattc aacgcttctc tcggcactta      780
tcacgacctt ctgaagatta tcaaggataa ggatttcctg gacaacgaag agaatgaaga      840
catcctggag gacatcgtcc tgaccttgac cctgttcgag gacagagaga tgatcgagga      900
gaggcttaag acctacgccc acctgtttga tgacaaagtg atgaaacagc tgaacggag      960
acggtatact ggttggggca ggctgtcccg gaagcttatt aacggaatac gggataagca     1020

aagtggaaag acaatacttg acttctgaa gtctgatggt tttgctaaca ggaatttcat     1080
gcagctgatt cagcagact cccttacatt taaggaggac attcagaagg cccaggtgtc     1140
tggacaaggg gactctctcc atgagcacaat cgccaacctg gccggcagcc cagccatcaa     1200
aaaaggaatt ctcaaacctg taaaggtggt ggatgagctg gttaaagtca tgggacggca     1260
caagcctgag aatatcgtca ttgagatggc cagggagaat cagacgacac agaaaggaca     1320
gaagaactca cgcgagagga tgaagagaat tgaggaaggg ataaaggagc tgggaagtca     1380
gattctgaag gaacaccag ttgaaaatac ccagctgcag aatgaaaagc tgtatctgta     1440

ctatctgcag aatggacgag acatgtatgt tgatcaggag ctggacatta accgactctc     1500
agattatgac gtggatcata tagtcctca gagtttctc aaggacgatt caatcgataa     1560

```

taaagtgttg acccgcagcg acaaaaacag gggcaaaagc gataatgtgc cctcagagga 1620
 agtgggtcaag aaaaatgaaga attactggag acagctgctc aacgctaagc ttattacca 1680
 gaggaattc gataatttga caaaagctga aaggggtggg cttagcgagc tggataaagc 1740
 aggattcatc aagcggcagc ttgtcgagac gcgccagatc acaaagcagc tggcacagat 1800
 tttggattcc cgcatgaaca ctaagtatga cgagaacgat aagctgatcc gcgaggtgaa 1860

 ggtgatcacg ctgaagtcca agctggtaag tgatttccgg aaagatttcc agttctacaa 1920
 agtgaggag attaacaact atcaccacgc ccacgacgct tacttgaatg ccgttgtggg 1980
 tacagcattg atcaaaaaat atccaaagct ggaaagtgag tttgtttacg gagactataa 2040
 agtctatgac gtgcggaaga tgatcgccaa gagcgagcag gagatcggga aagcaacagc 2100
 taaatatttc ttctattcca atatcatgaa ttttttcaa actgagataa cacttgctaa 2160
 tggtagata agaaagcagc cgctgataga gacgaatggc gagactggcg agatcgtgtg 2220
 ggacaaaggg agggacttcc caaccgtccg caaggtcttg agcatgccgc aggtgaatat 2280

 agttaagaaa accgaagtgc aaacaggcgg cttcagtaag gattccatata tgcgaagag 2340
 gaactctgac aagctgatcg ctaggaaaaa ggattgggat ccaaaaaaat acggcgggtt 2400
 cgactccct accgttgcac acagcgtgct tgtggtcgcg aaggtcgaag agggcaagtc 2460
 taagaagctc aagagtgtca aagaattgct gggtatcaca attatggagc gcagtagttt 2520
 cgagaagaat ccgatagatt ttctggagc aaagggatac aaggaggtga agaaggatct 2580
 gatcatcaaa ctgcctaagt actcctgtt cgagcttgag aatggtagaa agcgcagct 2640
 tgcctcagcc ggcgaattgc agaagggcaa tgagctcgcc ctgccttcaa aatacgtgaa 2700

 cttcctgtac ttggcatcac actacgaaaa gctgaaagga tcccctgagg ataagtagca 2760
 aaaacaactt tttgtggagc agcataagca ctatctgat gaaattattg agcagatttc 2820
 tgaattcagc aagcgcgtca tctcgcgga cgccaatctg gataaagtgc tgagcgccta 2880
 caataaacac cgagacaagc ccattcggga acaggccgag aacatcattc acctctcac 2940
 tctgactaat ctgggggccc cggccgcat caaatacttc gacactacta tcgacaggaa 3000
 acgtataact tcaacgaagg aggtgctgga cgctactttg atccaccagt ccattacggg 3060
 gctctatgag acacgaatcg atctttctca acttggaggt gatgcctacc catatgacgt 3120

 gcctgactat gcctctctgg gctctgggag ccctaagaaa aagaggaagg tagaggatcc 3180
 aaaaaaaaaag cgaaaagtcg attagagatc tgctcact gtgccttcta gttgccagcc 3240
 atctgtgtt tgccccccc ccgtgccttc cttgaccctg gaagtgcca ctcccactgt 3300
 cctttctaa taaatgagg aaattgcac gcattgtctg agtaggtgic attctattct 3360
 ggggggtggg gtggggcagg acagcaaggg ggaggattgg gaagacaata gcaggcatgc 3420

tggggaggta accacgtgcg gaccgagcgg ccgcaggaac ccctagtgat ggagttggcc	3480
actccctctc tgcgcgctcg ctcgctcact gaggccgggc gaccaaaggt cgccccacgc	3540
ccgggctttg cccgggcggc ctcaagtgagc gagcgagcgc gcagctgcct gcaggggcgc	3600
ctgatgcggt attttctct tacgcatctg tgcggtatth cacaccgcat acgtcaaagc	3660
aacctagta cgcgccctgt agcggcgcac taagcgcggc ggggtgtgtg gttacgcga	3720
gcgtgaccgc tacacttgcc agcgccttag cgcccctcc tttcgcttc tccccctct	3780
ttctcgccac gttcgccggc tttccccgtc aagctctaaa tcgggggctc cctttagggt	3840
tccgatttag tgccttacgg cacctcgacc ccaaaaaact tgatttgggt gatggttcac	3900
gtagtgggccc atcgccctga tagacggttt ttcgcccttt gacgttggag tccacgttct	3960
ttaatagtgg actcttgctc caaactggaa caacactcaa ccctatctcg ggctattctt	4020
ttgatttata agggattttg ccgatttcgg cctatttggt aaaaaatgag ctgatttaac	4080
aaaaattta cgcgaattht aacaaaatat taacgtttac aatthtatgg tgcactctca	4140
gtacaatctg ctctgatgcc gcatagttaa gccagccccg acaccgcca acaccgctg	4200
acgcgccctg acgggctttg ctgctccccg catccgetta cagacaagct gtgaccgtct	4260
ccgggagctg catgtgtcag aggttttcac cgtcatcacc gaaacgcgcg agacgaaagg	4320
gcctcgtgat acgcctatth ttataggtha atgtcatgat aataatggtt tcttagacgt	4380
caggtggcac ttttcgggga aatgtgcgcg gaaccctat ttgthtattt ttctaatac	4440
attcaaatat gtatccgctc atgagacaat aacctgata aatgcttcaa taatattgaa	4500
aaaggaagag tatgagtatt caacatthc gtgtcgcct tattccctth tttgcggcat	4560
tttgcctthc tgtthttgct caccagaaa cgttggtgaa agtaaaagat gctgaagatc	4620
agttgggtgc acgagtgggt tacatcgaac tggatctcaa cagcggtaag atccttgaga	4680
gtthtcgccc cgaagaacgt thtccaatga tgagacttht taaagthctg ctatgtggcg	4740
cggtattatc ccgtattgac gccgggcaag agcaactcgg tcgccgcata cactatthc	4800
agaatgactt ggtttagtac tcaccagtca cagaaaagca tcttacggat ggcatgacag	4860
taagagaatt atgcagtgct gccataacca tgagtgataa cactgcgcc aacttacttc	4920
tgacaacgat cggaggaccg aaggagctaa ccgctthttt gcacaacatg ggggatcatg	4980
taactcgctt tgatcgttg gaaccggagc tgaatgaagc cataccaaac gacgagcgtg	5040
acaccacgat gcctgtagca atggcaacaa cgttgcgcaa actatthact ggcgactac	5100
ttactctage thccccgcaa caatthaatag actggatgga ggcggataaa gttgcaggac	5160

cacttctgcg ctcgccctt ccggtggct ggttattgc tgataaatct ggagccggtg 5220

agcgtgggtc tcgcggtatc attgcagcac tggggccaga tggtaaagccc tcccgtatcg 5280

tagttatcta caccagggg agtcaggcaa ctatggatga acgaaataga cagatcgctg 5340

agataggtgc ctactgatt aagcattggt aactgtcaga ccaagtttac tcatatatac 5400

tttagattga tttaaaactt catttttaat ttaaaaggat ctaggatgaag atcctttttg 5460

ataatctcat gaccaaaatc ccttaacgtg agttttcggt ccaactgagcg tcagaccccg 5520

tagaaaagat caaaggatct tcttgagatc ctttttttct gcgcgtaatc tgctgcttgc 5580

aaacaaaaaa accaccgcta ccagcgggtg tttgtttgcc ggatcaagag ctaccaactc 5640

ttttccgaa ggtaactggc ttcagcagag cgcagatacc aaactactgc cttctagtgt 5700

agccgtagtt aggccaccac ttcaagaact ctgtagcacc gcctacatac ctgctctgctc 5760

taatcctggt accagtggct gctgccagtg gcgataagtc gtgtcttacc gggttggact 5820

caagacgata gttaccggat aaggcgcagc ggtcgggctg aacggggggt tcgtgcacac 5880

agcccagctt ggagcgaacg acctacaccg aactgagata cctacagcgt gagctatgag 5940

aaagcggcac gcttcccga gggagaaaagg cggacaggta tccggtaagc ggcagggtcg 6000

gaacaggaga gcgcacgagg gagcttccag ggggaaacgc ctggtatctt tatagtcctg 6060

tcgggtttcg ccacctctga cttgagcgtc gatTTTTGTG atgctctca gggggcgga 6120

gcctatggaa aaacgccagc aacgcggcct ttttacggtt cctggccttt tgctggcctt 6180

ttgctcacat gt 6192

<210> 340

<211> 6642

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 340

acatgtgagc aaaaggccag caaaaggcca ggaaccgtaa aaaggccgcg ttgctggcgt 60

ttttccatag gctccgcccc cctgacgagc atcacaaaaa tcgacgctca agtcagaggt 120

ggcgaaaacc gacaggacta taaagatacc agcgtttcc ccttgaagc tccctcgtgc 180

gctctcctgt tccgacctg ccgcttaccg gatacctgtc cgcctttctc ccttcgggaa 240

gcgtggcgct ttctcatagc tcacgtgta ggtatctcag ttcgggttag gtcgttcgct 300

ccaagctggg ctgtgtgcac gaaccccccg ttcagcccga ccgctgcgcc ttatccggtg 360

actatcgtct tgagtccaac ccggtaaagac acgacttacc gccactggca gcagccactg 420

gtaacaggat tagcagagcg aggtatgtag gcggtgctac agagtcttg aagtggggc 480
 ctaactacgg ctacactaga aggacagtat ttggtatctg cgctctgctg aagccagtta 540

 ccttcggaaa aagagttggg agctcttgat ccggcaaaaca aaccaccgct ggtagcggtg 600
 gttttttgt ttgcaagcag cagattacgc gcagaaaaaa aggatctcaa gaagatcctt 660
 tgatcttttc tacggggtct gacgctcagt ggaacgaaaa ctcacgttaa gggattttgg 720
 tcatgagatt atcaaaaagg atcttcacct agatcctttt aaattaaaaa tgaagtttta 780
 aatcaatcta aagtataat gagtaacctt ggtctgacag ttaccaatgc ttaatcagtg 840
 aggcacatct ctcagcgtc tgtctatttc gttcatccat agttgctga ctccccgtcg 900
 tgtagataac tacgatacgg gagggcttac catctggccc cagtgctgca atgataccgc 960

 ggcttccacg ctcaccggtt ccagatttat cagcaataaa ccagccagcc ggaagggccc 1020
 agcgcagaag tggctctgca actttatecg cctccatcca gtctattaat tgttgcggg 1080
 aagctagagt aagtagttcg ccagtaata gtttgcgcaa cgttgttgcc attgctacag 1140
 gcatcgtggg gtcacgctcg tcgtttggta tggcttcatt cagctccggt tcccaacgat 1200
 caaggcgagt tacatgatcc cccatgttgt gcaaaaaagc ggttagctcc ttcggctctc 1260
 cgatcgttgt cagaagtaag ttggccgag tgttatcact catggttatg gcagactgc 1320
 ataattctct tactgtcatg ccatccgtaa gatgcttttc tgtgactggt gagtactcaa 1380

 ccaagtcatt ctgagaatag tgtatgcggc gaccgagttg ctcttgccc gcgtcaatac 1440
 gggataatac cgcgccacat agcagaactt taaaagtct catcattgga aaacgttctt 1500
 cggggcgaaa actctcaagg atcttaccgc tgttgagatc cagttcgatg taaccactc 1560
 gtgcacccaa ctgatcttca gcatctttta ctttaccag cgtttctggg tgagcaaaaa 1620
 caggaaggca aaatgccga aaaaagggaa taaggcgac acggaatgt tgaatactca 1680
 tactttctt ttttcaatat tattgaagca tttatcaggg ttattgtctc atgagcggat 1740
 acatatttga atgtatttag aaaaataaac aaataggggt tccgcgcaca tttccccgaa 1800

 aagtgccacc tgacgtctaa gaaaccatta ttatcatgac attaacctat aaaaataggc 1860
 glatcacgag gccctttcgt ctcgcgctt tcggtgatga cggtgaaaac ctctgacaca 1920
 tgcagctccc ggagacggtc acagcttgc tgaagcgga tgccgggagc agacaagccc 1980
 gtcagggcgc gtcagcgggt gttggcgggt gtcgggctg gcttaactat gcggcatcag 2040
 agcagattgt actgagagtg caccataaaa ttgtaaacgt taatattttg ttaaaattcg 2100
 cgttaaattt ttgttaaatc agctcatttt ttaaccaata ggccgaaatc ggcataatcc 2160

cttataaatc aaaagaatag cccgagatag ggttgagtgt tgttccagtt tggacaaga 2220

 gtccactatt aaagaactg gactccaacg tcaaagggcg aaaaaccgtc tatcagggcg 2280
 atggcccact acgtgaacca taccceaat caagtttttt ggggtcgagg tcccgtaaag 2340
 cactaaatcg gaaccctaaa gggagcccc gatttagagc ttgacgggga aagccggcga 2400
 acgtggcgag aaaggaaggg aagaaagcga aaggagcggg cgctagggcg ctggcaagt 2460
 tagcggtcac gctgcgcgta accaccacac ccgccgcgct taatgcgccc ctacagggcg 2520
 cgtactatgg ttgctttgac gtatgcggtg tgaataccg cacagatgcg taaggagaaa 2580
 ataccgcatc aggcgccctt gcaggcagct gcgcgctcgc tcgctcactg aggcgcccg 2640

 ggcaaagccc gggcgtcggg cgacctttgg tcccccggcc tcagttagcg agcgagcgcg 2700
 cagagaggga gtggccaact ccatcactag gggttcctgc ggccgcctcg aggcgttgac 2760
 attgattatt gactagtatt taatagtaat caattacggg gtcattagtt catagcccat 2820
 atatggagtt ccgcgttaca taacttacgg taaatggccc gcctggctga ccgccaacg 2880
 acccccgcc attgacgtca ataatgacgt atgttcccat agtaacgcca atagggactt 2940
 tccattgacg tcaatgggtg gagtatttac ggtaaactgc ccaactggca gtacatcaag 3000
 tgtatcatat gccaagtacg cccctattg acgtcaatga cggtaaatgg cccgcctggc 3060

 attatgccc a gtacatgacc ttatgggact ttectacttg gcagtacatc tacgtattag 3120
 tcatcgctat taccatggtg atgcggtttt ggcagtacat caatggcgt ggatagcgg 3180
 ttgactcacg gggatttcca agtctccacc ccattgacgt caatgggagt ttgtttggc 3240
 accaaaaa ca acgggacttt ccaaaatgtc gtaacaactc cgccccattg acgcaaatgg 3300
 gcggtaggcg tgiacggtgg gaggtctata taagcagagc tctctggcta actaccggtg 3360
 ccaccatgat taagatgca accgaaaat acctgggaaa gcagaacgtc tacgatattg 3420
 gtgtagagag agaccataac tttgctctga agaacggctt tattgcctca tgcttcgaca 3480

 gcgttgagat atccggcgtg gaggatagat tcaacgcttc tctcggcact tatcacgacc 3540
 ttctgaagat tatcaaggat aaggatttcc tggacaacga agagaatgaa gacatcctgg 3600
 aggacatcgt ccigacctg accctgttcg aggacagaga gatgatcgag gagaggctta 3660
 agacctacgc ccacctgtt gatgacaaag tgatgaaaca gctgaaacgg agacgtata 3720
 ctggttgggg caggctgtcc cggaagctta ttaacggaat acgggataag caaagtggaa 3780
 agacaatact tgacttctg aagtctgatg gttttgctaa caggaatttc atgcagctga 3840
 ttcacgacga ctcccttaca ttaaggagg acattcagaa ggcccaggtg tctggacaag 3900

gggactctct ccatgagcac atcgccaacc tggccggcag cccagccatc aaaaaaggaa 3960
 ttcttcaaac tgtaaagggtg gtggatgagc tggttaaagt catgggacgg cacaagcctg 4020
 agaatatcgt cattgagatg gccagggaga atcagacgac acagaaagga cagaagaact 4080
 cacgcgagag gatgaagaga attgaggaag ggataaagga gctgggaagt cagattctga 4140
 aggaacaccc agttgaaaa atccagctgc agaataaaaa gctgtatctg tactatctgc 4200
 agaatggacg agacatgat gttgatcagg agctggacat taaccgactc tcagattatg 4260
 acgtggatca tatagtcctt cagagtttcc tcaaggacga ttcaatcgat aataaagtgt 4320

 tgaccgcgag cgacaaaaac aggggcaaaa gcgataatgt gccctcagag gaagtggcca 4380
 agaaaatgaa gaattactgg agacagctgc tcaacgctaa gcttattacc cagaggaaat 4440
 tcgataatth gacaaaagct gaaaggggtg ggcttagcga gctggataaa gcaggattca 4500
 tcaagcgcca gcttgtcgag acgcgccaga tcacaaagca cgtggcacag attttggatt 4560
 cccgcatgaa cactaagat gacgagaacg ataagctgat ccgagagggtg aaggatgatca 4620
 cgctgaagtc caagctggta agtgatttcc ggaaagattt ccagttctac aaagtgaggg 4680
 agattaacaa ctatcaccac gcccacgacg ctacttgaa tgccgttgtg ggtacagcat 4740

 tgatcaaaaa atatccaaag ctggaaagtg agtttgttta cggagactat aaagtctatg 4800
 acgtgcggaa gatgatgcc aagagcgagc aggagatcgg gaaagcaaca gctaaatatt 4860
 tcttctatc caatatcatg aatthtttca aaactgagat aaccttgct aatggtgaga 4920
 taagaaagcg accgctgata gagacgaatg gcgagactgg cgagatcgtg tgggacaaag 4980
 ggagggactt cgcaaccgtc cgcaaggtct tgagcatgcc gcaggtgaat atagttaaga 5040
 aaaccgaagt gcaaacaggc ggcttcagta aggagtccat attgccgaag aggaactctg 5100
 acaagctgat cgctaggaaa aaggattggg atccaaaaaa atacggcggg ttcgactccc 5160

 ctaccgttgc atacagcgtg cttgtggctg cgaaggctga aaaggcaag tetaagaagc 5220
 tcaagagtgt caagaattg ctgggtatca caattatgga gcgcagtagt ttcgagaaga 5280
 atccgataga ttttctggag gcaaaggat acaaggaggt gaagaaggat ctgatcatca 5340
 aactgcctaa gtactcctg ttcgagcttg agaatgtag aaagcgcag cttgcctcag 5400
 ccggcgaatt gcagaaggc aatgagctcg cctgccttc aaaatacgtg aacttctgt 5460
 acttggcacc aactacgaa aagctgaaag gatccctga ggataatgag caaaaacaac 5520
 tttttgtgga gcagcataag cactatctcg atgaaattat tgagcagatt tctgaattca 5580

 gcaagcgcgt catcctcgcg gacgccaatc tgataaagt gctgagcgc tacaataaac 5640
 accgagacaa gccattcgg gaacagggc agaacatcat tcacctctc actctgacta 5700
 atctcggggc cccggccgca ttcaataact tcgacactac tatcgacagg aaacgtata 5760

cttcaacgaa ggaggtgctg gacgctactt tgatccacca gtccattacg gggctctatg 5820
 agacacgaat cgatctttct caacttggag gtgattctgg cggctctaca aatctgtctg 5880
 acataataga aaaggaaact gggaagcaac ttgtcatcca agaatccata cttatgttgc 5940
 cggaagaggt tgaagaggtc attggttaata agccggagag cgatattctc gtacacacag 6000

 catacgatga atcaaccgat gaaaacgtaa tgttgcttac ttcagatgct cccgagtaca 6060
 agccctgggc attggtaatc caggattcca acggcgaaaa caaaattaag atgctttctg 6120
 gaggagtc caagaaaaag cggaaggtag cgtaccgta tgatgtcca gattacgca 6180
 gtcttgtag cgggtccccg aagaaaaagc gaaaggtgga agatccgaag aaaaagagaa 6240
 aagttgatta ggaattccta gagctcgctg atcagcctcg actgtgcctt ctagttgcca 6300
 gccatctgtt gtttgccct cccccgtgc ttcttgacc ctggaaggtg ccaactccac 6360
 tgtcctttcc taataaaatg aggaaattgc atcgattgt ctgagtaggt gtcattctat 6420

 tctgggggt ggggtgggc aggacagcaa gggggaggat tgggaagaga atagcaggca 6480
 tgctggggag ctagaggccg caggaacccc tagtgatgga gttggccact ccctctctgc 6540
 gcgctcgctc gtcactgag gccggcgac caaaggtcgc ccgacgcccg ggctttgccc 6600
 gggcggcctc agtgagcgag cgagcgcgca gctgcctgca gg 6642

 <210> 341
 <211> 7203
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic polynucleotide
 <400> 341

 tcgcggttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60

 cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
 ttggcgggtg tcgggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagcgcgc 240
 attcgcatt caggctgcbc aactgttggg aaggcgatc ggtgcgggcc tcttcgctat 300
 tacccagct ggcgaaaggg ggatgtgctg caaggcgatt aagttgggta acgccagggt 360
 ttcccagtc acgacgttgt aaaacgacgg ccagtgaatt gacgcgcat tgggatgttg 420
 taaaacgacg gccagtgaac ctgcaggcag ctgcgcgctc gctcgtcac tgaggccgcc 480

 cgggcaaage ccgggctcg ggcgacctt ggtcgcccgg cctcagtgag cgagcgagcg 540
 cgagagagg gaggggccaa ctccatcact aggggttctt gcggccgcac gcgtggagga 600

gggcctatTT cccatgattc cttcatatTT gcatatacga tacaaggctg ttagagagat 660
 aattagaatt aattgactg taaacacaaa gatattagta caaaatcgt gacgtagaaa 720
 gtaataatTT cttgggtagt ttgcagTTTT aaaattatgt tttaaaatgg actatcatat 780
 gcttaccgta acttgaaagt atttcgattt cttggcttta tatacttTgt ggaaaggacg 840
 aaacaccggt ttagagcta gaaatagcaa gttaaaataa ggctagtccg ttatcaactt 900

 gaaaaagtgg caccgagtcg gtgctTTTT gctagcctag acccagcttt cttgtacaaa 960
 gttggcatta atacgcgttg acattgatta ttgactagtt attaatagta atcaattacg 1020
 gggtcattag ttcatagccc atatatggag ttccgcgtta cataacttac ggtaaatggc 1080
 ccgcctggct gaccgcccac cgacccccgc ccattgacgt caataatgac gtatgttccc 1140
 atagtaacgc caataggac tttccattga cgtcaatggg tggagtatTT acggtaact 1200
 gcccacttgg cagtacatca agtgtatcat atgccaagta cggccctat tgacgtcaat 1260
 gacggtaaat ggccccctg gcattatgcc cagtacatga ccttatggga ctttctact 1320

 tggcagtaca tctacgtatt agtcatcgtt attaccatgg tgatgcggtt ttggcagtac 1380
 atcaatgggc gtggatagcg gtttgactca cggggatttc caagtctcca cccattgac 1440
 gtcaatggga gtttgtttg gcacccaaaat caacgggact ttccaaaatg tcgtaacaac 1500
 tccgccccat tgacgcaaat gggcggtagg cgtgtacggt gggaggctta tataagcaga 1560
 gctcgtttag tgaaccgtca gatcgctgg agacgccatc cacgtgttt tgacctcat 1620
 agaagacacc gggaccgatc cagcctccgg actctagagg atcgaacct taaggccacc 1680
 atgggaccga aaaaaagag gaaggtcgcg gctggaagcg gttccatgtc cagcgagacc 1740

 ggacccttg ccgtcgatcc tactttgagg agaagaatcg aaccacatga atttgaagta 1800
 tttttcgacc ctagagagct gcgaaaagaa acctgcttgc tgatgaaat aaattggggc 1860
 ggtcgccaca gtatatggag gcacacctc cagaatacaa acaagcacgt agaggtgaac 1920
 tttattgaaa aattcaccac agagagatat ttctgccga atacgagatg ttccattacg 1980
 tggttttttt cttggctccc atgcggtgag tgttcccggg ccatcacaga gttttgtca 2040
 cgataccctc acgtcacgt ttttatctac atagcgcgac tgatcacca tgccgacccc 2100
 agaataggc aaggcttgcg cgatttgatt agtagcgggg ttaccatcca gattatgacg 2160

 gagcaagagt cagggtactg ttggcggaac tttgtaaact actccccgag caatgaggcg 2220
 cactggcctc gctaccaca cctgtgggtc cgactttacg tcttgaatt gtattgcatc 2280
 atcctcggcc tcccgcgtg tctgaacatc ctgcggcgca agcagcccca attgacattt 2340
 tttacaatcg cctgcaatc atgccattat cagcggttgc cgccacacat actttgggccc 2400
 acgggtttga aaagcggatc cgagacgcct ggcaccagcg agtccgcaac ccccagagac 2460

gacaaaaagt atagtatagg tttggctatt ggaactaatt ccgtaggttg ggctgtgata 2520
acagatgaat acaaagtacc tagcaaaaag ttcaaggtgc ttggcaacac agatcgccac 2580

tcaatcaaga aaaaccttat cggagccctg ctgtttgact caggcgaaac cgccgagget 2640
acacgcctga aaagaacagc tagacggcgg tacaccagaa ggaagaaccg gatctgttat 2700
cttcaggaga ttttctccaa tgagatggct aaggtggacg attctttctt ccatcgactc 2760
gaagaatctt tcttgggtga ggaagataag aaacacgaga ggcatcctat tttcgaaac 2820
attgtcgatg aagtggccta tcatgagaaa taccacacga tctaccatct gcgaaaaaag 2880
ttggttgact ctaccgacaa ggccggacctg aggccttattt atctggccct ggcccatatg 2940
atcaaattca gggggcactt cttgatcgag ggggacctta atcccacaaa ctctgacgtg 3000

gataagtgtg tcatacagct tgtgcagacc tacaaccagc tgttcgagga gaatccaatc 3060
aacgccagcg gagtggacgc taaagccatt ctgagcgcga gattgagcaa gtctagaaga 3120
ttgghaaacc ttatagccca gctgccaggt gagaagaaga acggactgtt tggcaatctc 3180
attgcgctta gcctcggact caccocgaac ttcaaatcca acttcgacct cgccgaagat 3240
gccaaattgc agctcagtaa ggatagctat gacgatgatc ttgacaatct gctggcgcag 3300
atcggggacc agtacgcga tcttttcttg gcagcaaaaa atctctcaga tgcaatactc 3360
ttgtcagaca tactgcgagt taataccgag attactaagg ctccgctttc tgectccatg 3420

atcaagcgct acgatgagca tcaccaggat ctgacactgt tgaagccct ggtgcgccaa 3480
cagctgccag agaaatacaa ggaaatcttt tttgaccagt ccaagaatgg ctacgcagga 3540
tacatcgatg gaggagccag tcaggaggaa ttttacaagt ttattaagcc tatcctggag 3600
aagatggatg gtaccgaaga actcctggtc aagctcaacc gagaagattt gcttcgcaag 3660
caaaggactt ttgacaacgc ctccattccg catcagattc atctggcgca gctgcatgcc 3720
attctcgcaa gacaggagga tttttacceca tttctgaagg acaaccgaga gaagatcgag 3780
aaaactga cattcaggat accatattac gtgggtccac tcgccagggg caactcccga 3840

ttcgctgga tgacaaggaa aagcgaagag acgatcactc catggaactt cgaggaggtc 3900
gtggacaagg gggcctccgc gcagagcttt atcgagagga tgacgaactt tgacaaaaat 3960
ctccctaacg agaaggtgct gccaaaacat tctctgctct acgagtattt caccgtttat 4020
aatgagctca caaaggtgaa gtacgtgacc gaagggatgc ggaagcccgc ttttctgtcc 4080
ggagagcaga agaaggctat cgtggatttg ctctttaaga ctaaccgcaa ggtaacagtc 4140
aagcagctga aggaagacta cttcaagaag atcgaatgct tgtcctacga aacggaatc 4200

ttgacagttg agtaccggct cctgccaatc gggaagatag tagagaagag gattgaaatgt 4260

 accgtctatt ctgttgataa caacggtaac atatacacc agcccgtcgc ccaatggcac 4320
 gatcgcggtg agcaggaggt gttcgaatac tgtctggagg acgggtcatt gattcgggcg 4380
 actaaggacc ataagtttat gacggtagac ggccagatgt tgccataga tgagatcttt 4440
 gagcgggaac tcgacttgat gagagtcgat aatcttcta attagcttaa gggttcgatc 4500
 cctactggtt agtaatgagt ttaaaccggg gaggctaact gaaacacgga aggagacaat 4560
 accggaagga acccgcgcta tgacggcaat aaaaagacag aataaacgc acgggtgttg 4620
 ggtcgtttgt tcataaacgc ggggttcggt ccagggctg gcactctgtc gataccccac 4680

 cgagacccca ttggggccaa tacgcccgcg tttcttctt ttccccacc cccccccaa 4740
 gttcgggtga aggccaggg ctgcgacca acgtcggggc ggcaggccct gccatagcag 4800
 atctgcgctg atttttagg taaccacgtg cggaccgagc ggccgcagga acccctagtg 4860
 atggagttag ccaactccctc tctgcgctc cgctcgtca ctgaggccgg gcgaccaaag 4920
 gtgccccgac gcccggtt tgccccggcg gcctcagtga gcgagcgagc gcgcagctgc 4980
 ctgcaggctt ggatcccaat ggcgcgcca gcttggetcg agcatggtca tagctgtttc 5040
 ctgtgtgaaa ttgttatccg ctcaacaatc cacacaacat acgagccgga agcataaagt 5100

 gtaaagcctg ggtgacctaa tgagtgagct aactcacatt aattgcgttg cgctcaactgc 5160
 ccgctttcca gtcgggaaac ctgtcgtgcc agctgcatta atgaatcggc caacgcgagg 5220
 ggagaggcgg tttgcgtatt gggcgtctt ccgcttctc gctcaactgac tcgtgcgct 5280
 cggctgttcg gctgcggcga gcggtatcag ctcaactcaa ggcgtaata cggttatcca 5340
 cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa aggccagcaa aaggccagga 5400
 accgtaaaaa gccgcggtg ctggcgtttt tccataggct ccgccccct gacgagcacc 5460
 acaaaaatcg acgctcaagt cagaggtggc gaaacccgac aggactataa agataccagg 5520

 cgtttcccc tggaagctcc ctctgctcct ctctgttcc gacctgccg cttaccggat 5580
 acctgtccg ctttctccct tcgggaagcg tggcgtttc tcatagctca cgctgtaggt 5640
 atctcagttc ggigttagtc gttcgtcca agctgggctg tgtgcacgaa cccccgttc 5700
 agcccagccg ctgcgcctta tccgtaact atcgtcttga gtccaaccg gtaagacacg 5760
 acttatccg actggcagca gccactggta acaggattag cagagcgagg tatgtaggcg 5820
 gtgctacaga gttcttgaag tgggtgccta actacggcta cactagaaga acagtatttg 5880
 gtatctgcgc tctgctgaag ccagttacct tcggaaaaag agttggtagc tcttgatccg 5940

gcaaacaaac caccgctggt agcgggtggtt tttttgttg caagcagcag attacgcgca 6000

gaaaaaagg atctcaagaa gatcctttga tcttttctac ggggtctgac gctcagtgga 6060

acgaaaactc acgttaaggg attttgggtca tgagattatc aaaaaggatc ttcacctaga 6120

tccttttaaa ttaaaaatga agttttaaat caatctaaag tatatatgag taaacttgg 6180

ctgacagtta gaaaaactca tcgagcatca aafgaaactg caatttattc atatcaggat 6240

tatcaatacc atatttttga aaaagccgtt tctgtaatga aggagaaaac tcaccgaggc 6300

agttccatag gatggcaaga tcctggtatc ggtctgcat tccgactcgt ccaacatcaa 6360

tacaacctat taatttcccc tcgtcaaaaa taaggttatc aagtgagaaa tcacatgag 6420

tgacgactga atccggtgag aatggcaaaa gtttatgcat ttcittccag acttgitcaa 6480

caggccagcc attacgctcg tcatcaaaat cactcgcatc aaccaaacg ttattcattc 6540

gtgattgcgc ctgagcgaga cgaaatacgc gatcgctgtt aaaaggacaa ttacaaacag 6600

gaatcgaatg caaccggcgc aggaacactg ccagcgcac aacaatattt tcacctgaat 6660

caggatattc ttctaatacc tggaatgctg ttttccagg gatcgactg gtgagtaacc 6720

atgcatcacc aggagtacgg ataaaaatgct tgatggfctgg aagaggcata aattccgtca 6780

gccagtttag tctgaccatc tcatctgtaa catcattggc aacgctacct ttgcatggt 6840

tcagaaacaa ctctggcgca tcgggcttcc catacaatcg atagattgtc gcacctgatt 6900

gcccacatt atcgcgagcc catttatacc catataaate agcatccatg ttggaattta 6960

atcgcgccct agagcaagac gtttcccggt gaatatggct catactcttc ctttttcaat 7020

attattgaag catttatcag ggttattgtc tcatgagcgg atacatattt gaatgtattt 7080

agaaaaataa acaaataggg gttccgcgca ctttccccc aaaagtgcc cctgacgtct 7140

aagaaacct tattatcatg acattaacct ataaaaatag gcgtatcacg aggcccttc 7200

gtc 7203

<210> 342

<211> 7447

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 342

cctgcaggca gctgcgct cgctcgctca ctgaggccgc ccggcgctcg ggcgacctt 60

ggtcgcccg cctcagttag cgagcgagcg cgcagagagg gactggccaa ctccatcact 120

aggggttctt gcggcctcta gactcgaggc gttgacattg attattgact agttattaat 180

agtaatcaat tacggggica ttagttcata gcccatatat ggagttccgc gttacataac 240

 ttacggtaaa tggccccct ggctgaccgc ccaacgacc cgcgccattg acgtcaataa 300
 tgacgtatgt tcccatagta acgccaatag ggactttcca ttgacgtcaa tgggtggagt 360
 atttacggta aactgccac ttggcagtac atcaagtga tcatatgcca agtacgcccc 420
 ctattgacgt caatgacggt aaatggcccc cctggcatta tgcccagtac atgaccttat 480
 gggactttcc tacttggcag tacatctacg tattagtcac cgctattacc atgggtgatgc 540
 ggttttggca gtacatcaat gggcgtggat agcggtttga ctacgggga tttccaagtc 600
 tccccccat tgacgtcaat gggagtttgt tttggcacca aatcaacgg gactttccaa 660

 aatgtcgtaa caactccgc ccatgacgc aaatgggcgg taggcgtgta cgggtggagg 720
 tctatataag cagagctctc tggctaacta ccggtgccac catggcccca aagaagaagc 780
 ggaaggtcgg tatccacgga gtcccagcag ccaagcggaa ctacatcctg ggccctggaca 840
 tcggcatcac cagcgtgggc tacggcatca tcgactacga gacacgggac gtgatcgatg 900
 ccggcgtgcg gctgttcaaa gaggccaacg tggaaaacaa cgagggcagg cggagcaaga 960
 gaggcggcag aaggctgaag cggcggaggc ggcatagaat ccagagagtg aagaagctgc 1020
 tgttcgacta caacctgctg accgaccaca gcgagctgag cggcatcaac ccctacgagg 1080

 ccagagtga gggcctgagc cagaagctga gcgaggaaga gttctctgcc gccctgctgc 1140
 acctggccaa gagaagaggc gtgcacaacg tgaacgaggt ggaagaggac accggcaacg 1200
 agctgtccac caaagagcag atcagccgga acagcaaggc cctggaagag aaatacgtgg 1260
 ccgaactgca gctggaacgg ctgaagaaag acggcgaagt gcggggcagc atcaacagat 1320
 tcaagaccag cgactacgtg aaagaagcca aacagctgct gaaggtgcag aaggcctacc 1380
 accagctgga ccagagcttc atcgacacct acatcgacct gctggaacc cggcggacct 1440
 actatgaggg acctggcgag ggcagccct tccgctggaa ggacatcaa gaatggtacg 1500

 agatgctgat gggccactgc acctacttc ccgaggaact gcggagcgtg aagtacgct 1560
 acaacgccga cctgtacaac gccctgaacg acctgaacaa tctcgtgatc accagggacg 1620
 agaacgagaa gctggaatat tacgagaagt tccagatcat cgagaacgtg tccaagcaga 1680
 agaagaagcc caccctgaag cagatcgcca aagaaatcct cgtgaacgaa gaggatatta 1740
 agggctacag agtgaccagc accggcaagc ccgagttcac caacctgaag ggtgaccacg 1800
 acatcaagga cattaccgcc cggaaagaga ttattgagaa cgccgagctg ctggatcaga 1860
 ttgccaagat cctgaccatc taccagagca gcgaggacat ccaggaagaa ctgaccaatc 1920

tgaactccga gctgacccag gaagagatcg agcagatctc taatctgaag ggctataccg 1980
gcaccacaa cctgagcctg aaggccatca acctgatcct ggacgagctg tggcacacca 2040
acgacaacca gatcgtatc ttcaaccggc tgaagctggt gccaagaag gtggacctgt 2100
cccagcagaa agagatcccc accaccctgg tggacgactt catcctgagc cccgtcgtga 2160
agagaagctt catccagagc atcaaagtga tcaacgcat catcaagaag tacggcctgc 2220
ccaacgacat cattatcgag ctggccccgc agaagaactc caaggacgcc cagaaaatga 2280
tcaacgagat gcagaagcgg aaccggcaga ccaacgagcg gatcgaggaa atcatccgga 2340

ccaccggcaa agagaacgcc aagtacctga tcgagaagat caagctgcac gacatgcagg 2400
aaggcaagtg cctgtacagc ctggaagcca tcctctgga agatctgctg aacaaccct 2460
tcaactatga ggtggaccac atcatccca gaagcgtgtc cttcgacaac agcttcaaca 2520
acaagtgct cgtgaagcag gaagaaaaca gcaagaaggg caaccggacc ccattccagt 2580
acctgagcag cagcgacagc aagatcagct acgaaacctt caagaagcac atcctgaatc 2640
tggccaaggg caagggcaga atcagcaaga ccaagaaga gtatctgctg gaagaacggg 2700
acatcaacag gttctccgtg cagaaagact tcatcaaccg gaacctggtg gataccagat 2760

acgccaccag aggcctgatg aacctgctgc ggagctactt cagagtgaac aacctggacg 2820
tgaaagtga gtccatcaat ggcggttca ccagcttct gcgcggaag tggaaagtta 2880
agaaagagcg gaacaagggg tacaagcacc acgccgagga cgcctgatc attgccaacg 2940
ccgatttcat ctcaaagag tggagaagc tggacaaggc caaaaagtg atggaaaacc 3000
agatgttca gaaaagcag gccgagagca tgcccagat cgaaacgag caggagtaca 3060
aagagatctt catcacccc caccagatca agcacattaa ggacttcaag gactacaagt 3120
acagccaccg ggtggacaag aagcctaata gagagctgat taacgacacc ctgtactcca 3180

cccggaagga cgacaagggc aacaccctga tctgaacaa tctgaacggc ctgtacgaca 3240
aggacaatga caagctgaaa aagctgatca acaagagccc cgaaaagctg ctgatgtacc 3300
accacgacce ccagacctac cagaaactga agctgattat ggaacgtac ggcgacgaga 3360
agaatcccct gtacaagtac tacgaggaaa ccgggaacta cctgaccaag tactcaaaa 3420
aggacaacgg cccctgatc aagaagatta agtattacgg caacaaactg aacgcccac 3480
tggacatcac cgacgactac cccaacagca gaacaaggt cgtgaagctg tcctgaagc 3540
cctacagatt cgacgtgtac ctggacaatg gcgtgtacaa gttcgtgacc gtgaagaatc 3600

tggatgtgat caaaaagaa aactactacg aagtgaatag caagtgtat gaggaagcta 3660
agaagctgaa gaagatcagc aaccaggccg agtttatcgc ctcttctac aacaacgatc 3720
tgatcaagat caacggcgag ctgtatagag tgatcgcgct gaacaacgac ctgctgaacc 3780

ggatcgaagt gaacatgac gacatcacct accgcgagta cctggaaaac atgaacgaca 3840
 agaggccccc caggatcatt aagacaatcg cctccaagac ccagagcatt aagaagtaca 3900
 gcacagacat tctgggcaac ctgtatgaag tgaaatctaa gaagcacct cagatcatca 3960
 aaaagggcaa aaggccggcg gccacgaaaa aggccggcca ggcaaaaaag aaaaagggat 4020

 cctaccata cgatgttcca gattacgctt acccatacga tgttccagat tacgcttacc 4080
 catacgatgt tcagattac gcttaagaat tcctagagct cgctgatcag cctcgactgt 4140
 gccttctagt tgccagccat ctgttgtttg ccctcccc gtgccttctt tgaccctgga 4200
 aggtgccact cccactgtcc tttcctaata aatgaggaa attgcatcgc attgtctgag 4260
 taggtgtcat tctattctgg ggggtggggt ggggcaggac agcaaggggg aggattggga 4320
 agagaatagc aggcattctg gggaggtacc tgaggccta tttccatga ttccttcata 4380
 tttgcatata cgatacaagg ctgttagaga gataattgga attaatttga ctgtaaacac 4440

 aaagatatta gtacaaaata cgtgacgtag aaagtaataa tttcttgggt agtttgcagt 4500
 tttaaaatta tgttttaaaa tggactatca tatgcttacc gtaacttga agtatttca 4560
 tttcttggct ttatatact tgtgaaaagg acgaaacacc ggagaccacg gcaggtctca 4620
 gttttagtac tctgaaaca gaactacta aaacaaggca aatgccgtg tttatctcgt 4680
 caacttgttg gcgagatttt tgcggccgca ggaacccta gtgatggagt tggccactcc 4740
 ctctctgcgc gctcgtcgc tcactgagc cggcgacca aaggtcgccc gacgcccggg 4800
 ctttggccgg gcggcctcag tgagcgagcg agcgcgcagc tgcctgcagg ggcgcctgat 4860

 gcggtatfff ctcttacgc atctgtgcgg tatttcacac cgcatcgtc aaagcaacca 4920
 tagtacgcgc ccigtacgg cgcattaagc gcggcgggtg tgggtgttac gcgcagcgtg 4980
 accgtacac ttgccagcgc cctagcgcgc gctccttctg ctttcttccc ttcctttctc 5040
 gccacgttgc cggcctttcc ccgtcaagct ctaaatcggg ggctcccttt agggttccga 5100
 tttagtctt tacggcacct cgacccaaa aaacttgatt tgggtgatgg ttcacgtagt 5160
 gggccatcgc cctgatagac ggttttctgc cttttgacgt tggagtccac gttctttaat 5220
 agtggactct tgttccaaac tggaacaaca ctcaacccta tctcgggcta ttcctttgat 5280

 ttataagga ttttgccgat ttcggcctat tggtaaaaa atgagctgat ttaacaaaa 5340
 tttaacgga attttaaca aatattaacg tttacaattt tatggtgcac tctcagtaca 5400
 atctgctctg atgccgata gtttaagccag ccccgacacc cgccaacacc cgctgacgcg 5460
 ccctgacggg ctgtctgct cccgcatcc gcttacagac aagctgtgac cgtctccggg 5520
 agctgcatgt gtcagaggtt ttcaccgtca tcaccgaaac gcgcgagacg aaaggcctc 5580
 gtgatacgc ttttttata ggttaatgtc atgataataa tggtttctta gacgtcaggt 5640

ggcacttttc ggggaaatgt ggcggaacc cctatttgtt tatttttcta aatacattca 5700

aatatgtatc cgctcatgag acaataaacc tgataaatgc ttcaataata ttgaaaagg 5760

aagagtatga gtattcaaca tttccgtgtc gcccttattc ccttttttgc ggcattttgc 5820

cttcctgttt ttgctcacc agaaacgctg gtgaaagtaa aagatgctga agatcagttg 5880

ggtgcacgag tgggttacat cgaactggat ctcaacagcg gtaagatcct tgagagtttt 5940

cgccccgaag aacgttttcc aatgatgagc acttttaag ttctgctatg tggcgcggta 6000

ttatcccgta ttgacgccgg gcaagagcaa ctccggtgcc gcatacacta ttctcagaat 6060

gacttggttg agtactcacc agtcacagaa aagcatctta cggatggcat gacagtaaga 6120

gaattatgca gtgctgccat aaccatgagt gataaactg cggccaactt acttctgaca 6180

acgatcggag gaccgaagga gctaaccgct tttttgcaca acatggggga tcatgtaact 6240

cgcttgatc gttgggaacc ggagctgaat gaagccatac caaacgacga gcgtgacacc 6300

acgatgcctg tagcaatggc aacaacgttg cgcaactat taactggcga actacttact 6360

ctagcttccc ggcaacaatt aatagactgg atggagcggg ataaagtgc aggaccactt 6420

ctgcctcgg ccttccggc tggctggttt attgctgata aatctggagc cggtagcgt 6480

ggaagccgcg gtatcattgc agcactgggg ccagatggta agccctccc tatcgtagtt 6540

atctacacga cggggagtca ggcaactatg gatgaacgaa atagacagat cgctgagata 6600

ggtgcctcac tgattaagca ttggtaactg tcagaccaag tttactcata tatactttag 6660

attgatttaa aacttcattt ttaatttaa aggatctagg tgaagatcct ttttgataat 6720

ctcatgacca aaatccetta acgtgagttt tcgttccact gagcgtcaga ccccgtagaa 6780

aagatcaaag gatcttcttg agatcctttt tttctgcgcg taatctgctg cttgcaaaca 6840

aaaaaacac cgctaccagc ggtggtttgt ttgccgcatc aagagctacc aactctttt 6900

ccgaaggtaa ctggcttcag cagagcgcag ataccaaata ctgtccttct agtgtagccg 6960

tagttaggcc accacttcaa gaactctgta gcaccgcta catacctcgc tctgtaatc 7020

ctgttaccag tggtctctgc cagtggcgat aagtcgtgtc ttaccgggtt ggactcaaga 7080

cgatagttag cggataaggc gcagcggctg ggctgaacgg ggggttcgtg cacacagccc 7140

agcttggagc gaacgaccta caccgaactg agatacctac agcgtgagct atgagaaagc 7200

gccacgcttc ccgaaggagg aaaggcggac aggtatccgg taagcggcag ggtcggaaca 7260

ggagagcgca cgagggagct tccaggggga aacgcctggt atctttatag tctgtcggg 7320

tttcgccacc tctgacttga gcgtcgattt ttgtgatgct cgtcaggggg gcggagccta 7380

ccaagtcatt ctgagaatag tgtatgcggc gaccgagttg ctcttgcccc gcgtcaatac 1440

gggataatac cgcgccacat agcagaactt taaaagtgct catcattgga aaacgttctt 1500

cggggcgaaa actctcaagg atcttaccgc tgttgagatc cagttcgatg taaccctctc 1560

gtgcacccaa ctgatcttca gcatctttta ctttcaccag cgtttctggg tgagcaaaaa 1620

caggaaggca aaatgccga aaaaaggga taaggcgac acggaatgt tgaatactca 1680

tactcttctt ttttcaatat tattgaagca tttatcaggg ttattgtctc atgagcggat 1740

acataattga atgtatttag aaaaataaac aaataggggt tccgcgcaca tttccccgaa 1800

aagtgccacc tgacgtctaa gaaaccatta ttatcatgac attaacctat aaaaataggc 1860

gtatcacgag gccctttcgt ctcgcgctt tcggtgatga cggtgaaaac ctctgacaca 1920

tgcagctccc ggagacggtc acagcttctc tgtaagcgga tgccgggagc agacaagccc 1980

gtcagggcgc gtcagcgggt gttggcgggt gtcggggctg gcttaactat gcggcatcag 2040

agcagattgt actgagagtg caccataaaa ttgtaaacgt taatattttg ttaaaattcg 2100

cgttaaattt ttgttaaatc agctcatttt ttaaccaata ggccgaaatc ggcaaaatcc 2160

cttataaate aaaagaatag cccgagatag ggttgagtgt tgttccagtt tggacaaga 2220

gtccactatt aaagaacgtg gactccaacg tcaaaggcg aaaaaccgtc tatcagggcg 2280

atggcccact acgtgaacca taccceaat caagtttttt ggggtcgagg tgccgtaaag 2340

cactaaatcg gaaccctaaa gggagcccc gatttagagc ttgacgggga aagccggcga 2400

acgtggcgag aaaggaagg aagaaagcga aaggagcggg cgctagggcg ctggcaagtg 2460

tagcggtcac gctgcgcta accaccacac ccgccgcgt taatgcgccc ctacagggcg 2520

cgtactatgg ttgctttgac gtatgcggtg tgaataaccg cacagatgcg taaggagaaa 2580

ataccgcatc aggcgccct gcagcagct gcgcgctgc tcgctcactg aggccgccc 2640

ggcaaaagccc gggcgtcggg cgacctttgg tcgcccggcc tcagtgagcg agcgagcgcg 2700

cagagaggga gtggccaact ccatcactag gggttcctgc ggccgctcg aggcgttgac 2760

attgattatt gactagtat taatagtaat caattacggg gtcattagt catagcccat 2820

atatggagtt ccgcgttaca taactacgg taaatggccc gcctggctga ccgccaacg 2880

acccccccc attgacgca ataatgacgt atgttccat agtaacgcca ataggactt 2940

tccattgacg tcaatgggtg gagtatttac ggtaaactgc ccaactggca gtacatcaag 3000

tgtatcatat gccaagiag cccctattg acgtcaatga cggtaaatgg cccgcctggc 3060

attatgcca gtacatgacc ttatgggact ttctacttg gcagtacatc tacgtattag 3120

tcatcgctat taccatgggtg atgcggtttt ggcagtacat caatgggcgt ggatagcggg 3180
 ttgactcacg gggatttcca agtctccacc ccattgacgt caatgggagt ttgttttggc 3240
 accaaaaatca acgggacttt ccaaaaatgtc gtaacaactc cgccccattg acgcaaatgg 3300
 gcggtaggcg tgiacgggtg gaggctctata taagcagagc tctctggcta actaccgggtg 3360
 ccacatgat taagatcgca acccgaaaat acctgggaaa gcagaacgtc tacgatattg 3420
 gtgtagagag agaccataac tttgctctga agaacggctt tattgcctca tgcttcgaca 3480
 gcgttgagat ttccggcgtg gaggatagat tcaacgcttc tctcggcact taccacgacc 3540

 ttctgaagat tatcaaggat aaggatttcc tggacaacga agagaatgaa gacatcctgg 3600
 aggacatcgt ccigaccttg acctgttctg aggacagaga gatgatcgag gagaggctta 3660
 agacctacgc ccacctgttt gatgacaaag tgatgaaaca gctgaaacgg agacggtata 3720
 ctggttgggg caggctgtcc cggaagctta ttaacggaat acgggataag caaagtggaa 3780
 agacaatact tgacttctg aagtctgatg gttttgctaa caggaatttc atgcagctga 3840
 ttcacgacga ctccttaca ttaaggagg acattcagaa ggcccaggtg tctggacaag 3900
 gggactctct ccatgagcac atcgccaacc tggccggcag cccagccatc aaaaaaggaa 3960

 ttcttcaaac tglaaagggtg gtggatgagc tggttaaagt catgggacgg cacaagcctg 4020
 agaatatcgt cattgagatg gccagggaga atcagacgac acagaaagga cagaagaact 4080
 cacgcgagag gatgaagaga attgaggaag ggataaagga gctgggaagt cagattctga 4140
 aggaacaccc agltgaaaa atccagctgc agaataaaaa gctgtatctg tactatctgc 4200
 agaatggacg agacatgtat gttgatcagg agctggacat taaccgactc tcagattatg 4260
 acgtggatgc tatagtccct cagagtttcc tcaaggacga ttcaatcgat aataaagtgt 4320
 tgaccgcgag cgacaaaaac aggggcaaaa gcgataatgt gccctcagag gaagtgggtca 4380

 agaaaatgaa gaattactgg agacagctgc tcaacgctaa gcttattacc cagaggaaat 4440
 tcgataattt gacaaaagct gaaaggggtg ggcttagcga gctggataaa gcaggattca 4500
 tcaagcggca gcttgtcgag acgcgccaga tcacaaagca cgtggcacag attttggatt 4560
 cccgcatgaa cactaagtat gacgagaacg ataagctgat ccgagagggtg aaggtgatca 4620
 cgctgaagtc caagctggta agtgatttcc gaaaagattt ccagttctac aaagtgaggg 4680
 agattaacaa ctatcaccac gccacgacg ctactttaa tgccgttgtg ggtacagcat 4740
 tgatcaaaaa atatccaaag ctggaagtg agtttgttta cggagactat aaagtctatg 4800

 acgtgcggaa gatgatgcc aagagcgagc aggagatcgg gaaagcaaca gctaaatatt 4860
 tcttctatc caatatcatg aattttttca aaactgagat aaccttctc aatgggtgaga 4920
 taagaaagcg accgctgata gagacgaatg gcgagactgg cgagatcgtg tgggacaaag 4980

ggagggactt cgcaaccgtc cgcaaggctt tgagcatgcc gcaggtgaat atagttaaga 5040
 aaaccgaagt gcaaacaggc ggcttcagta aggagtccat attgccgaag aggaactctg 5100
 acaagctgat cgctaggaaa aaggattggg atccaaaaaa atacggcggg ttcgactccc 5160
 ctaccgttgc atacagcgtg cttgtggtcg cgaaggcga aaaggcaag tctaagaagc 5220

 tcaagagtgt caaagaattg ctgggtatca caattatgga ggcgagtagt ttcgagaaga 5280
 atccgataga ttttctggag gcaaagggat acaaggaggt gaagaaggat ctgatcatca 5340
 aactgcctaa gtactccctg ttcgagcttg agaatggtag aaagcgcagc cttgcctcag 5400
 ccggcgaatt gcagaagggc aatgagctcg cctgccttc aaaatacgtg aacttctctg 5460
 acttggcatc aactacgaa aagctgaaag gatccctga ggataatgag caaaaacaac 5520
 tttttgtgga gcagcataag cactatctcg atgaaattat tgagcagatt tctgaattca 5580
 gcaagcgcgt catcctcgcg gacgccaatc tggataaagt gctgagcgcc tacaataaac 5640

 accgagacaa gccattcgg gaacaggccg agaacatcat tcacctctc actctgacta 5700
 atctcggggc cccggccgca ttcaaatact tcgacactac tatcgacagg aaacgtata 5760
 cttaacgaa ggaggtgctg gacgetactt tgatccacca gtccattacg gggctctatg 5820
 agacacgaat cgatctttct caacttggag gtgatgccta cccatgatgac gtgcctgact 5880
 atgcctccct gggctctggg agccctaaga aaaagaggaa ggtagaggat ccaaaaaaaaa 5940
 agcgaagaat cgatgagcc agcggttccg gacgggctga cgattggac gatcttgatc 6000
 tggatatgct ggaagtgcg gccctcgatg attttgacct tgacatgctt ggttcggatg 6060

 cccttgatga cttgacctc gacatgctcg gcagtgacgc cttgatgat ttcgacctgg 6120
 acatgctgat taactctaga agttccggat ctccgaaaaa gaaacgcaa gttggtggca 6180
 gccgggattc caggaaggg atgtttttgc cgaagcctga ggccggctcc gctattagtg 6240
 acgtgtttga gggccgcgag gtgtgccagc caaacgaat ccggccattt catcctccag 6300
 gaagtccatg ggccaaccgc cactccccg ccagcctcgc accaacacca accggtccag 6360
 tacatgagcc agtcgggtca ctgaccccg caccagtccc tcagccactg gatccagcgc 6420
 ccgcagtgc tcccaggcc agtcacctgt tggaggatcc cgatgaagag acgagccagg 6480

 ctgtcaaage cttcgggag atggccgata ctgtgattcc ccagaaggaa gaggtgcaa 6540
 tctgtggcca aatggacctt tcccatccgc cccaagggg ccactctggat gagctgaaa 6600
 ccacacttga gtccatgacc gaggatctga acctggactc acccctgacc ccggaattga 6660
 acgagattct ggatacttc ctgaacgacg agtgcctctt gcatgccatg catatcagca 6720
 caggactgtc catcttcgac acatctctgt tttaggaatt cctagagctc getgatcagc 6780
 ctcgactgtg ctttctagtt gccagccatc tgttgtttgc cctcccccg tgccttctt 6840

gaccttgaa ggtgccactc ccactgtcct ttctaataa aatgagaaa ttgcatcgca 6900

ttgtctgagt aggtgtcatt ctattctggg ggttgggtg gggcaggaca gcaaggggga 6960

ggattggaa gagaatagca ggcatgctgg ggagctagag gccgcaggaa ccctagtga 7020

tggagtggc cactccctct ctgcgcgctc gctcgctcac tgaggccggg cgaccaaagg 7080

tcgcccgacg cccgggcttt gcccgggcgg cctcagttag cgagcgagcg cgcagctgcc 7140

tgcagg 7146

<210> 344

<211> 6354

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 344

acatgtgagc aaaaggccag caaaaggcca ggaaccgtaa aaaggccgag ttgctggcgt 60

ttttccatag gtcgccccc cctgacgagc atcacaaaaa tcgacgctca agtcagaggt 120

ggcgaaaccc gacaggacta taaagatacc aggcgtttcc ccctggaagc tcctcgtgc 180

gctctcctgt tccgacctg ccgcttaccg gatacctgac cgcctttctc ctttcgggaa 240

gcgtggcgct ttctcatagc tcacgctgta ggtatctcag ttcggtgtag gtcgttcgct 300

ccaagctggg ctgtgtgac gaacccccg ttcagcccga ccgctgcgcc ttatccggtg 360

actatcgtct tgagtccaac ccgtaagac acgacttacc gccactggca gcagccactg 420

gtaacaggat tagcagagcg aggtatgtag gcggtgctac agagtcttg aagtgtggc 480

ctaactacgg ctacactaga aggacagtat ttggtatctg cgctctgctg aagccagtta 540

ccttcggaaa aagagttggt agctcttgat ccggcaaaaca aaccaccgct ggtagcgggtg 600

gtttttttgt ttgcaagcag cagattacgc gcagaaaaaa aggatctcaa gaagatcctt 660

tgatcttttc tacggggtct gacgctcagt ggaacgaaaa ctcacgttaa gggatcttgg 720

tcatgagatt atcaaaaagg atcttcacct agatcctttt aaattaaaaa tgaagtttta 780

aatcaatcta aagtatatat gagtaacctt ggtctgacag ttaccaatgc ttaatcagtg 840

aggcacctat ctcagcgatc tgtctatttc gttcatccat agttgctga ctccccgtcg 900

tgtagataac tacgatacgg gagggcttac catctggccc cagtgctgca atgataccgc 960

ggcttccacg ctcaccggtc ccagatttat cagcaataaa ccagccagcc ggaagggccg 1020

agcgcagaag tggctctgca actttatecg cctccatcca gtctattaat tgttgccggg 1080

aagctagagt aagtagttcg ccagtaata gtttgcgcaa cgttgttggc attgctacag 1140

gcatacgtggt gtcacgctcg tcgtttggta tggcttcatt cagctccggg tcccaacgat 1200
caaggcgagt tacatgatcc cccatgttgt gcaaaaaagc ggtagctcc ttcggtcctc 1260

cgatacgttgt cagaagtaag ttggccgcag tgtatcact catggttatg gcagcactgc 1320
ataattctct tactgtcatg ccatccgtaa gatgcttttc tgtgactggt gactactcaa 1380
ccaagtcatg ctgagaatag tgtatcgggc gaccgagttg ctcttgcccg gcgtcaatac 1440
gggataatac cgcgccacat agcagaactt taaaagtgct catcattgga aaacgttctt 1500
cggggcgaaa actctcaagg atcttaccgc tgttgagatc cagttcgatg taaccactc 1560
gtgcacccaa ctgatcttca gcatctttta ctttaccag cgtttctggg tgagcaaaaa 1620
caggaaggca aaatgccga aaaaaggga taaggcgac acggaatgt tgaatactca 1680

tactcttct tttcaatat tattgaagca tttatcaggg ttattgtctc atgagcggat 1740
acataattga atgtattag aaaaataaac aaatagggt tccgcgcaca tttccccgaa 1800
aagtgcacc tgacgtctaa gaaaccatta ttatcatgac attaacctat aaaaataggc 1860
gtatcacgag gccctttcgt ctgcgcggt tccgtgatga cggtgaaaac ctctgacaca 1920
tgcagctccc ggagacggtc acagcttctc tgaagcggg tcccgggagc agacaagccc 1980
gtcagggcgc gtcagcgggt gttggcgggt gtcggggctg gcttaactat gcggcatcag 2040
agcagattgt actgagagtg caccataaaa ttgtaaacgt taatattttg ttaaaattcg 2100

cgttaaattt ttgttaaate agctcatttt ttaaccaata ggccgaaatc ggcaaaatcc 2160
cttataaate aaaagaatag cccgagatag ggttgagtgt tgttccagtt tggaacaaga 2220
gtccactatt aaagaactg gactccaacg tcaaagggcg aaaaaccgtc tatcagggcg 2280
atggcccact acgtgaacca tcacccaaat caagtttttt ggggtcagagg tgccgtaaag 2340
cactaaatcg gaacctaaa gggagcccc gatttagagc ttgacgggga aagccggcga 2400
acgtggcgag aaaggaaggg aagaaagcga aaggagcggg cgctagggcg ctggcaagtg 2460
tagcggtcac gctgcgcgta accaccacac ccgccgct taatgcgccg ctacagggcg 2520

cgtaactatg ttgctttgac gtatcgggtg tgaataaccg cacagatgag taaggagaaa 2580
ataccgatc aggcgccct gcaggcagct gcgcgctcgc tcgctcactg aggccccccg 2640
ggcaagccc ggcgctcggg cgacctttgg tcgccggcc tcagttagcg agcgagcgcg 2700
cagagagggg gtggccaact ccatcactag gggttctcgc ggccgctcg aggcgttgac 2760
attgattatt gactagtat taatagtaat caattacggg gtcattagtt catagccat 2820
atatggagt cgcggttaca taacttacgg taatggccc gcctggctga ccgccaacg 2880

acccccgccc attgacgtca ataatgacgt atgttcccat agtaacgcca atagggactt 2940

 tccattgacg tcaatgggtg gagtatttac ggtaaactgc ccacttgga gtacatcaag 3000
 tgtatcatat gccaagtacg cccctattg acgtcaatga cggtaaatgg cccgcctggc 3060
 attatgccca gtacatgacc ttatgggact ttctacttg gcaglacatc tacgtattag 3120
 tcatcgctat taccatggtg atgcggtttt ggcagtacat caatgggcgt ggatagcgg 3180
 ttgactcacg gggatttcca agtctccacc ccattgacgt caatgggagt ttgtttggc 3240
 accaaaatca acgggacttt ccaaaatgtc gtaacaactc cgccccattg acgcaaatgg 3300
 gcggtaggcg tgtacggtgg gaggtctata taagcagagc tctctggcta actaccggtg 3360

 ccacatgat taagatcgca acccgaaaat acctgggaaa gcagaacgtc tacgatattg 3420
 gtgtagagag agaccataac tttgctctga agaacggctt tattgcctca tgcttcgaca 3480
 gcgttgagat ttccggcgtg gaggatagat tcaacgettc tctcggcact tatcacgacc 3540
 ttctgaagat tatcaaggat aaggatttcc tggacaacga agagaatgaa gacatcctgg 3600
 aggacatcgt cctgacctg accctgttcg aggacagaga gatgatcgag gagaggctta 3660
 agacctacgc ccacctgttt gatgacaaaag tgatgaaaca gctgaaacgg agacgtata 3720
 ctggttgggg caggctgtcc cggaagctta ttaacggaat acgggataag caaagtggaa 3780

 agacaatact tgacttcctg aagtctgatg gttttgctaa caggaatttc atgcagctga 3840
 ttcacgacga ctcccttaca ttaaggagg acattcagaa ggcccagggtg tctggacaag 3900
 gggactctct ccatgagcac atcgccaacc tggccgcgag cccagccatc aaaaaaggaa 3960
 ttcttcaaac tgtaaagggtg gtggatgagc tggttaaagt catgggacgg cacaagcctg 4020
 agaatatcgt cattgagatg gccagggaga atcagacgac acagaaagga cagaagaact 4080
 cacgcgagag gatgaagaga attgaggaag ggataaagga gctgggaagt cagattctga 4140
 aggaacacce agttgaaaat acccagctgc agaatgaaaa gctgtatctg tactatctgc 4200

 agaatggacg agacatgtat gttgatcagg agctggacat taaccgactc tcagattatg 4260
 acgtggatgc tatagtccct cagagtttcc tcaaggacga ttcaatcgat aataaagtgt 4320
 tgaccgcgag cgacaaaaac aggggcaaaa gcgataatgt gccctcagag gaagtggca 4380
 agaaaatgaa gaattactgg agacagctgc tcaacgctaa gcttattacc cagaggaaat 4440
 tcgataatth gacaaaagct gaaaggggtg ggcttagcga gctggataaa gcaggattca 4500
 tcaagcggca gcttgtcgag acgcgccaga tcacaaagca cgtggcacag attttgatt 4560
 cccgatgaa cactaagtat gacgagaacg ataagctgat ccgcgaggtg aaggtgatca 4620

cgctgaagtc caagctggta agtgatttcc ggaaagattt ccagttctac aaagtgaggg 4680
 agattaacaa ctatcaccac gccacgacg cttactttaa tgccgttggt ggtacagcat 4740
 tgatcaaaaa atatccaaag ctggaaagtg agtttgttta cggagactat aaagtctatg 4800
 acgtgcggaa gatgatcgcc aagagcgagc aggagatcgg gaaagcaaca gctaaatatt 4860
 tcttctattc caatatcatg aatTTTTTca aaactgagat aaccttgct aatggtgaga 4920
 taagaaagcg accgctgata gagacgaatg gcgagactgg cgagatcgtg tgggacaaag 4980
 ggagggactt cgcaaccgtc cgcaaggtct tgagcatgcc gcaggtgaat atagttaaga 5040

aaaccgaagt gcaaacaggc ggcttcagta aggagtccat attgccgaag aggaactctg 5100
 acaagctgat cgctaggaag aaggattggg atccaaaaaa atacggcggg ttcgactccc 5160
 ctaccgttgc atacagcgtg cttgtggtcg cgaaggtcga aaagggcaag tctaagaagc 5220
 tcaagagtgt caaagaattg ctgggtatca caattatgga ggcagtagt ttcgagaaga 5280
 atccgataga ttttctggag gcaaagggat acaaggaggt gaagaaggat ctgatcatca 5340
 aactgcctaa gtactcctg ttcgagcttg agaatgtag aaagcgcag ctgcctcag 5400
 ccggcgaatt gcagaagggc aatgagctcg cctgccttc aaaatcgtg aacttctgt 5460

acttgcatc aactacgaa aagctgaaag gatcccctga ggataatgag caaaaacaac 5520
 tttttgtgga gcagcataag cactatctcg atgaaattat tgagcagatt tctgaattca 5580
 gcaagcgcgt catcctcgg gacccaatc tggataaagt gctgagcgc tacaataaac 5640
 accgagacaa gccattcgg gaacaggccg agaacatcat tcacctctc actctgacta 5700
 atctcggggc cccggccgca ttcaataact tcgacactac tatcgacagg aaacgtata 5760
 cttcaacgaa ggaggtgctg gacgctactt tgatccacca gtccattacg gggctctatg 5820
 agacacgaat cgatctttct caacttgag gtgatgccta cccatagac gtgcctgact 5880

atgcctccct gggtctctgg agccetaaga aaaagaggaa ggtagaggat ccaaaaaaaaa 5940
 agcgaagagt cgatgatatc taggaattcc tagagctcgc tgatcagcct cgactgtgcc 6000
 ttctagtgc cagccatctg ttgtttgcc cccccctg ccttcttga ccctggaagg 6060
 tgccactccc actgtcttt cctaataaaa tgaggaaatt gcatcgatt gtctgagtag 6120
 gtgtcattct attctggggg gtgggtggg gcaggacagc aagggggagg attgggaaga 6180
 gaatagcagg catgctgggg agctagaggc cgcaggaacc cctagtgat gagttggcca 6240
 ctccctctct gcgcctcgc tcgctcactg aggccgggcg accaaaggtc gcccgacgcc 6300

cgggctttgc cggggcgcc tcagtgagcg agcgagcgc cagctgcctg cagg 6354

<210> 345

<211> 6744

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 345

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca	60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg	120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc	180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagcgcgc	240
attcgccatt caggctgcgc aactgttggg aagggcgatc ggtgcgggcc tcttcgctat	300
tacgccagct ggcgaaaggg ggatgtgctg caaggcgatt aagtgggta acgccagggt	360
tttcccagtc acgacgttg taaaacgacgg ccagtgaatt gacgcgccat tgggatgttg	420
taaaacgacg gccagtgaac ctgcaggcag ctgcgcgctc gctcgtcac tgaggccgcc	480
cgggcaaagc ccgggcgtcg ggcgacctt ggtcggccg cctcagttag cgagcgagcg	540
cgcagagagg gaggggccaa ctccatcaact aggggttctt gcggccgcac gcgtggagga	600
gggcctatct cccatgattc ctccatattt gcatatacga tacaaggctg ttagagagat	660
aattagaatt aatttgactg taaacacaaa gatattagta caaaatcgt gacgtagaaa	720
gtaataatct ctgggtagt ttgcagtttt aaaattatgt tttaaaatgg actatcatat	780
gcttaccgta acttgaaagt atttcgattt ctggcctta tatacttctt ggaaaggacg	840
aaacaccggg tttagagcta gaaatagcaa gttaaaataa ggctagtccg ttatcaactt	900
gaaaaagtgg caccgagtcg gtgctttttt gctagcctag acccagcttt ctgtacaaa	960
gttggcatta atacgcgttg acattgatta ttgactagtt attaatagta atcaattacg	1020
gggtcattag ttcatagccc atatatggag ttccgcgtta cataacttac ggtaaatggc	1080
ccgcctggct gaccgcccac cgacccccgc ccattgacgt caataatgac gtatgttccc	1140
atagtaacgc caatagggac tttccattga cgtcaatggg tggagtatct acggtaaac	1200
gcccacttgg cagtacatca agtgtatcat atgccaaagta cgccccctat tgacgtcaat	1260
gacggtaaat ggccccctg gcattatgcc cagtacatga ccttatggga ctttcctact	1320
tggcagtaca tctactgatt agtcatcgct attaccatgg tgatcggtt ttggcagtac	1380
atcaatgggc gtggatagcg gtttgactca cggggatttc caagtcctca cccattgac	1440
gtcaatggga gtttgtttg gcacaaaat caacgggact ttccaaaatg tcgtaacaac	1500
tccgcccatt tgacgcaaat gggcggtagg cgtgtacggt gggaggtcta tataagcaga	1560
gctcgtttag tgaaccgtca gatcgccctg agacgccatc cacgctgttt tgacctccat	1620

agaagacacc gggaccgatc cagcctccgg actctagagg atcgaaccct taaggccacc 1680
 atggatgcta agtcactaac tgcctggctc cggacactgg tgaccttcaa ggatgtattt 1740
 gtggacttca ccagggagga gtggaagctg ctggacactg ctcagcagat cgtgtacaga 1800
 aatgtgatgc tggagaacta taagaacctg gtttccttgg gttatcagct tactaagcca 1860
 gatgtgatcc tccggttggg gaagggagaa gagcccggcg gttccggcgg agggtcgatg 1920

 ggccccaaga aaaaacgcaa ggtggccgca gcagactata aggatgacga cgataagggg 1980
 atccatggtg tgcctgctgc agataaaaa tacagcatcg gcctggctat cggaaactaac 2040
 tccgtcggct gggccgtcat taccgacgaa taaaaagtac ctagcaaaaa gttcaaggtg 2100
 cttggcaaca cagatcgcca ctcaatcaag aaaaacctta tccgagccct gctgtttgac 2160
 tcaggcgaaa ccgccgaggc tacacgcctg aaaagaacag ctagacggcg gtacaccaga 2220
 aggaagaacc ggatctgtta tcttcaggag attttctcca atgagatggc taaggtggac 2280
 gattctttct tccatcgact cgaagaatct ttcttgggtg aggaagataa gaaacacgag 2340

 aggcaccta ttttcgaaa cattgtcgat gaagtggcct atcatgagaa ataccccacg 2400
 atctaccate tgcgaaaaaa gttggttgac tctaccgaca aggcggacct gaggttatt 2460
 tatctggccc tggcccatat gatcaaattc agggggcact tcttgatcga gggggacctt 2520
 aatcccgaca actctgacgt ggataagttg ttcatacagc ttgtgcagac ctacaaccag 2580
 ctgttcgagg agaatecaat caacgccagc ggagtggacg ctaaagccat tctgagcgcg 2640
 agattgagca agtctagaag attggaaaac cttatagccc agctgccagg tgagaagaag 2700
 aacggactgt ttggcaatct cattgcgctt agcctcggac tcaccccga cttcaaatcc 2760

 aacttcgacc tcgccgaaga tgccaaattg cagctcagta aggatacgtg tgacgatgat 2820
 cttgacaate tgcctggcga gatcggggac cagtacccg atcttttctt ggcagcaaaa 2880
 aatctctcag atgcaatact cttgtcagac aactgcgag ttaataccga gattactaag 2940
 gctccgcttt ctgcctccat gatcaagcgc tacgatgagc atcaccagga tctgacactg 3000
 ttgaaagccc tggctgcgcca acagctgcca gagaaataca aggaatctt ttttgaccag 3060
 tccaagaatg gctacgcagg atacatcgat ggaggagcca gtcaggagga attttacaag 3120
 tttattaagc ctatcctgga gaagatggat ggtaccgaag aactcctggt caagctcaac 3180

 cgagaagatt tgcctcgcaa gcaaaggact tttgacaacg gctccattcc gcatcagatt 3240
 catctggcg agctgcatgc cattctgcga agacaggagg atttttacc atttctgaag 3300
 gacaaccgag agaagatcga gaaaatactg acattcagga taccatatta cgtgggtcca 3360
 ctgccaggg gcaactcccg attcgcctgg atgacaagga aaagcgaaga gacgatcact 3420
 ccatggaact tcgaggaggt cgtggacaag ggggcctccg cgcagagctt tctcgagagg 3480

atgacgaact ttgacaaaaa tctccctaac gagaaggtgc tgccaaaaca ttctctgctc 3540
 tacgagtatt tcaccgttta taatgagctc acaaaggtga agtacgtgac cgaagggatg 3600

 cggaagcccg cttttctgctc cggagagcag aagaaggcta tcgtggattt getctttaag 3660
 actaaccgca aggtaacagt caagcagctg aaggaagact acttcaagaa gatcgaatgc 3720
 ttgtcctacg aaacggaat cttgacagtt gagtacgggc tctgccaat cgggaagata 3780
 gtagagaaga ggattgaatg taccgtctat tctgttgata acaacggtaa catatacacc 3840
 cagcccgtcg cccaatggca cgatcgcggt gacgaggagg tgttcgaata ctgtctggag 3900
 gacgggtcat tgattcgggc gactaaggac cataagtta tgacggtaga cggccagatg 3960
 ttgccatag atgagatctt tgagcgggaa ctcgacttga tgagagtcga taatcttctt 4020

 aattagctta agggttcgat ccctactggt tagtaatgag tttaaaccggg ggaggctaac 4080
 tgaaacacgg aaggagacaa taccggaagg aaccgcgct atgacggcaa taaaaagaca 4140
 gaataaacg cacgggtgtt gggtcgtttg ttcataaacg cggggttcgg tcccagggtt 4200
 ggcactctgt cgatacccca ccgagacccc attggggcca atacgccgc gtttcttctt 4260
 tttcccacc ccaccccaca agttcgggtg aaggcccagg gctcgcagcc aacgtcgggg 4320
 cggcaggccc tgccatagca gatctgcgct gatttttag gtaaccacgt gcggaccgag 4380
 cggccgcagg aaccctagt gatggagttg gccactcct ctctgcgcgc tcgctcgctc 4440

 actgaggccg ggcgacaaa ggtcgcccga cggccgggt ttgcccgggc ggcctcagtg 4500
 agcgagcgag cgcgcagctg cctgcaggct tggatcccaa tggcgcgccg agcttggtc 4560
 gagcatggtc atagctgttt cctgtgtgaa attgttatcc gctcacaatt ccacacaaca 4620
 tacgagccgg aagcataaag tgtaaagcct ggggtgccta atgagtgagc taactcacat 4680
 taattgcgtt gcgctcactg cccgctttcc agtcgggaaa cctgtcgtgc cagctgcatt 4740
 aatgaatcgg ccaacgcgcg gggagaggcg gtttgcgtat tgggcgctct tccgcttctt 4800
 cgctcactga ctgctgcgc tcggtcgttc ggctgcggcg agcggtatca gctcactcaa 4860

 aggcggtaat acggttatcc acagaatcag gggataacgc aggaagaac atgtgagcaa 4920
 aaggccagca aaaggccagg aaccgtaaaa aggcccggtt gctggcgttt tccataggc 4980
 tccgcccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg cgaaaccga 5040
 caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc tctctgttc 5100
 cgacctgcc gcttaccgga tacctgtccg cctttctccc ttcgggaagc gtggcgcttt 5160
 ctcatagctc acgctgtagg tatctcagtt cgggttaggt cgttcgctcc aagctgggct 5220

gtgtgcacga accccccgtt cagcccgacc gctgcgcctt atccggtaac tatcgtcttg 5280

agtccaacce ggtaagacac gacttatcgc cactggcagc agccactggt aacaggatta 5340

gcagagcgag gtatgtaggc ggtgetacag agttcttgaa gtggtggcct aactacgget 5400

acactagaag aacagtatct ggtatctgcg cctgctgaa gccagttacc ttcgaaaaa 5460

gagttggtag ctcttgatcc ggcaaaacaa ccaccgctgg tagcgggtgt tttttgttt 5520

gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatccttg atcttttcta 5580

cggggtctga cgctcagtg aacgaaaact cacgttaagg gattttggtc atgagattat 5640

caaaaaggat cttcacctag atccttttaa attaaaaatg aagttttaa tcaatctaaa 5700

gtatatatga gtaaacttgg tctgacagt agaaaaactc atcgagcatc aaatgaaact 5760

gcaatttatt catatcagga ttatcaatac catatcttg aaaaagccgt ttctgtaatg 5820

aaggagaaaa ctaccgagg cagttccata ggatggcaag atcctggtat cggctcgcga 5880

ttccgactcg tccaacatca atacaacctt ttaatttccc ctctcaaaa ataaggttat 5940

caagtgagaa atcacatga gtgacgactg aatccggtga gaatggcaaa agtttatgca 6000

tttctttcca gacttgttca acagccagc cattacgctc gtcacaaaa tcaactcgc 6060

caaccaaacc gttattcatt cgtgattgcg cctgagcgag acgaaatagc cgatecgtgt 6120

taaaaggaca attacaaaca ggaatcgaat gcaaccggcg caggaacact gccagcgc 6180

caacaatatt ttacctgaa tcaggatatt cttctaatac ctggaatgct gttttccag 6240

ggatcgcagt ggtgagtaac catgcatcat caggagtacg gataaatgc ttgatggtcg 6300

gaagaggcat aaattccgtc agccagttta gtctgacct ctcactctga acatcattgg 6360

caacgctacc ttigccatgt ttcagaaaca actctggcgc atcgggcttc ccatacaatc 6420

gatagattgt cgcacctgat tgcccacat tatcgcgagc ccatttatac ccatataaat 6480

cagcatccat gttggaattt aatcgggcc tagagcaaga cgtttcccg tgaatatggc 6540

tcatactctt ccttttcaa tattattgaa gcatttatca gggttattgt ctcatgagcg 6600

gatacatatt tgaatgtatt tagaaaaata acaaatagg ggttccgcgc acatttccc 6660

gaaaagtgcc acctgacgtc taagaacca ttattatcat gacattaacc tataaaaata 6720

ggcgtatcac gaggcccttt cgtc 6744

<210> 346

<211> 6516

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 346

tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggcca	60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg	120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc	180
accatatgcg gtgtgaaata cgcacagat gcgtaaggag aaaataccgc atcagcgccc	240
attcgccatt caggctgcgc aactgttggg aagggcgatc ggtgcgggcc tettegctat	300
tacgccagct ggcgaaaggg ggatgtgctg caagcgcatt aagtgggta acgccagggt	360
tttcccagtc acgacgttgt aaaacgacgg ccagtgaatt gacgcgccat tgggatgttg	420
taaaacgacg gccagtgaac ctgcaggcag ctgcgcgctc gctcgctcac tgaggccgcc	480
cgggcaaagc ccgggcgtcg ggcgacctt ggtcgcccgg cctcagttag cgagcgagcg	540
cgcagagagg gagtggccaa ctccatcact aggggttctt cgggccgcac gcgtggagga	600
gggcctatth cccatgattc cttcatatth gcatatacga tacaaggctg ttagagagat	660
aattagaatt aattgactg taaacacaaa gatattagta caaaatcgt gacgtagaaa	720
gtaataatth cttgggtagt ttgcagthtt aaaattatgt tttaaatgg actatcatat	780
gcttaccgta acttgaaagt atttcgatth cttggcttta tatacttgt ggaaaggacg	840
aaacaccggt tttagagcta gaaatagcaa gttaaaataa ggctagtccg ttatcaactt	900
gaaaaagtgg caccgagtcg gtgctthttt gctagcctag acccagctth cttgtacaaa	960
gttggcatta atacgcgttg acattgatta ttgactagtt attaatagta atcaattacg	1020
gggtcattag ttcatagccc atatatggag ttccgcgtta cataacttac ggtaaatggc	1080
ccgcctggct gaccgcccac cgacccccgc ccattgacgt caataatgac gtatgttccc	1140
atagtaacgc caatagggac tttccattga cgtcaatggg tggagtatth acggtaact	1200
gcccacttgg cagtacatca agtgtatcat atgccaagta cgccccctat tgacgtcaat	1260
gacggtaaat ggccccctg gcattatgcc cagtacatga cttatggga ctttctact	1320
tggcagtaca tctacttatt agtcatcgtt attaccatgg tgatcggtt ttggcagtac	1380
atcaatgggc gtggatagcg gtttgactca cggggatttc caagtctcca cccattgac	1440
gtcaatggga gtttgtttg gcacaaaaat caacgggact ttccaaaatg tcgtaacaac	1500
tccgcccctat tgaccgcaat gggcggtagg cgtgtacggt gggaggtcta tataagcaga	1560
gctcgttttag tgaaccgtca gatcgcttgg agacgccatc cacgctgtth tgacctcat	1620
agaagacacc gggaccgatc cagcctccgg actctagagg atcgaacct taaggccacc	1680

atggatatca tgggcccaaa gaaaaaacgc aaggtggccg cagcagacta taaggatgac 1740

gacgataagg ggatccatgg tgtgcctgct gcagataaaa aatacagcat cggcctggct 1800

atcggaacta actccgtcgg ctgggccgtc attaccgacg aatacaaagt acctagcaaa 1860

aagttcaagg tgcttggcaa cacagatcgc cactcaatca agaaaaacct tatcggagcc 1920

ctgctgtttg actcaggcga aaccgccgag gctacacgcc tgaagaagac agctagacgg 1980

cggtacacca gaaggaagaa ccggatctgt tatcttcagg agattttctc caatgagatg 2040

gctaaggtgg acgattcttt cttccatcga ctggaagaat ctttcttggg ggaggaagat 2100

aagaaacacg agaggcatcc tattttcgga aacattgtcg atgaagtggc ctatcatgag 2160

aaatacccca cgatctacca tctgcgaaaa aagttggttg actctaccga caaggcggac 2220

ctgaggctta tttatctggc cctggcccat atgatcaaat tcagggggca cttcttgatc 2280

gagggggacc ttaateccga caactctgac gtggataagt tgttcataca gcttgtgcag 2340

acctacaacc agctgttcga ggagaatcca atcaacgcca gcggagtgga cgctaagcc 2400

attctgagcg cgagattgag caagtctaga agattggaaa accttatagc ccagctgcca 2460

ggtgagaaga agaacggact gtttggcaat ctattgcgc ttagcctcgg actcaccccg 2520

aacttcaaat ccaacttcca cctcggcga gatgccaaat tgcagctcag taaggatagc 2580

tatgacgatg atcttgacaa tctgctggcg cagatcgggg accagtacgc cgatcttttc 2640

ttggcagcaa aaaatctctc agatgcaata ctctgtcag acatactgcg agttaatacc 2700

gagattacta agctccgct ttctgcctcc atgatcaagc gctacgatga gcatcaccag 2760

gatctgacac tgttgaagc cctggtgcgc caacagctgc cagagaaata caaggaaatc 2820

ttttttgacc agtccaagaa tggctacgca ggatacatcg atggaggagc cagtccaggag 2880

gaattttaca agtttattaa gcctatcctg gagaagatgg atggtaccga agaactcctg 2940

gtcaagctca accgagaaga tttgcttcgc aagcaaagga cttttgacaa cggctccatt 3000

ccgcatcaga ttcatctggg cgagctgcat gccattctgc gaagacagga ggatttttac 3060

ccatttctga aggacaaccg agagaagatc gagaaaatac tgacattcag gataccatat 3120

tacgtgggtc cactcgccag gggcaactcc cgattcgcct ggatgacaag gaaaagcgaa 3180

gagacgatca ctccatggaa cttcgaggag gtcgtggaca agggggcctc cgcgcagagc 3240

tttatcgaga ggatgacgaa ctttgacaaa aatctccta acgagaaggt gctgccaaaa 3300

cattctctgc tciacgagta tttaccgtt tataatgagc tcacaaaggt gaagtacgtg 3360

accgaagga tgcggaagcc cgcttttctg tccggagagc agaagaagc tatcgtggat 3420

ttgctcttta agactaaccg caaggttaaca gtcaagcagc tgaaggaaga ctacttcaag 3480
 aagatcgaat gcttgtccta cgaaacggaa atcttgacag ttgagtacgg gtcctgccca 3540
 atcgggaaga tagtagagaa gaggattgaa tgtaccgtct attctgttga taacaacggt 3600
 aacatataca cccagcccggt cgccaatgg cacgatcgcg gtgagcagga ggtgttcgaa 3660
 tactgtctgg aggacgggtc attgattcgg gcgactaagg accataagtt tatgacggta 3720
 gacggccaga tgttgcccat agatgagatc tttgagcggg aactcgactt gatgagagtc 3780
 gataatcttc ctaattagct taagggttcg atccctactg gttagtaatg agtttaaacg 3840

 ggggaggcta actgaaacac ggaaggagac aataccggaa ggaacccgcg ctatgacggc 3900
 aataaaaaga cagaataaaa cgcacgggtg ttgggtcgtt tgttcataaa cgcggggttc 3960
 ggtcccaggg ctggcactct gtcgataccc caccgagacc ccatggggc caatacggcc 4020
 gcgtttcttc cttttccca cccaccccc caagttcggg tgaaggcca gggctcgag 4080
 ccaacgtcgg ggcggcaggc cctgccatag cagatctcgc ctgattttgt aggtaaccac 4140
 gtgcggaccg agcggccgca ggaacccta gtgatggagt tggccactcc ctctctgcgc 4200
 gctcgtcgc tcaactgagc cgggcgacca aaggtcggcc gacgcccggg ctttgcccgg 4260

 gcggcctcag tgagcgagcg agcgcgcagc tgctgcagg cttggatccc aatggcgcgc 4320
 cgagcttggc tcgagcatgg tcatagctgt ttctgtgtg aaattgttat ccgctcacia 4380
 ttccacacia catacgacc ggaagcataa agtgtaaage ctggggtgcc taatgagtga 4440
 gctaactcac attaattcgg ttgcgtcac tgcccgttt ccagtcggga aacctgtcgt 4500
 gccagctgca ttaatgaatc ggccaacgcg cggggagagg cggtttgcgt attgggcgct 4560
 ctccgcttc ctgcctcact gactcgtgc gctcggctgt tcggctcgg cgagcggtat 4620
 cagctcactc aaaggcggtg atacggttat ccacagaatc aggggataac gcaggaaaga 4680

 acatgtgagc aaaaggccag caaaaggcca ggaaccgtaa aaaggccgcg ttgctggcgt 4740
 ttttccatag gtcggcccc cctgacgagc atcacaaaa tcgacgctca agtcagaggt 4800
 ggcgaaacce gacaggacta taaagatacc agcgtttcc cctggaage tcctcgtgc 4860
 gctctcctgt tccgacctg ccgcttaccg gatacctgtc cgctttctc ccttcgggaa 4920
 gcgtggcgt ttctcatagc tcacgtgta ggtatctcag ttcggtgtag gtcgttcgct 4980
 ccaagctggg ctgtgtgac gaacccccg ttcagcccga ccgctgcgcc ttatccggtg 5040
 actatcgtct tgagtccaac ccggttaagac acgacttate gccactggca gcagccactg 5100

 gtaacaggat tagcagagcg aggtatgtag gcggtgctac agagtcttg aagtgtggc 5160
 ctaactaccg ctacactaga agaacagtat ttggtatctg cgctctgctg aagccagtta 5220
 ccttcggaaa aagagttggt agctcttgat ccggcaaca aaccaccgct gtagcggtg 5280

gttttttgt ttgcaagcag cagattacgc gcagaaaaa aggatctcaa gaagatcctt 5340
 tgatctttc tacgggtct gacgctcagt ggaacgaaa ctcacgttaa gggattttgg 5400
 tcatgagatt atcaaaaagg atcttcacct agatcctttt aaattaaaa tgaagtttta 5460
 aatcaatcta aagtatatat gagtaaactt ggtctgacag ttagaaaaac tcatcgagca 5520

 tcaaatgaaa ctgcaattta ttcatatcag gattatcaat accatatttt tgaaaaagcc 5580
 gtttctgtaa tgaaggagaa aactcacgca ggcagttcca taggatggca agatcctggt 5640
 atcggctctgc gattccgact cgtccaacat caatacaacc tattaatttc cctcgtcaa 5700
 aaataaggtt atcaagttag aaatcacat gagtgacgac tgaatccggt gagaatggca 5760
 aaagtttatg catttcttc cagacttggt caacaggcca gccattacgc tcgtcatcaa 5820
 aatcactcgc atcaacaaa ccgttattca ttcgtgattg cgcctgagcg agacgaata 5880
 cgcatcgct gttaaaagga caattacaaa caggaatcga atgcaaccgg cgcaggaaca 5940

 ctgccagcgc atcaacaata tttcacctg aatcaggata ttcttcta acctggaatg 6000
 ctgttttccc agggatcgca gtggtgagta accatgcac atcaggagta cggataaaat 6060
 gcttgatggt cggaagagc ataaattccg tcagccagtt tagtctgacc atctcatctg 6120
 taacatcatt ggcaacgcta cctttgcat gtttcagaaa caactctggc gcatcgggct 6180
 tccatacaa tcgatagatt gtcgcacctg attgcccgc attatcgca gccatttat 6240
 accatataa atcagatcc atgttggaat ttaatcgcg cctagagcaa gacgtttccc 6300
 gttgaatatg gtcatactc ttccttttc aatattatg aagcatttat cagggttatt 6360

 gtctcatgag cggatacata tttgaatgta tttagaaaa taaacaaata ggggttccgc 6420
 gcacatttcc ccgaaaagtg ccacctgacg tctaagaac cattattatc atgacattaa 6480
 cctataaaaa taggcgtatc acgaggcct ttcgtc 6516