(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2023-0036110
(43) 공개일자 2023년03월14일

(51) 국제특허분류(Int. Cl.)
C12N 15/86 (2006.01) *C07K 14/005* (2006.01)
(52) CPC특허분류
C12N 15/86 (2013.01)
C07K 14/005 (2013.01)
(21) 출원번호 10-2023-7001349
(22) 출원일자(국제) 2021년06월11일
심사청구일자 없음
(85) 번역문제출일자 2023년01월12일
(86) 국제출원번호 PCT/US2021/037091
(87) 국제공개번호 WO 2021/252955
국제공개일자 2021년12월16일
(30) 우선권주장
63/038,483 2020년06월12일 미국(US)
63/146,963 2021년02월08일 미국(US)

(71) 출원인
플래그쉽 파이어니어링 이노베이션스 브이, 인크.
미국 02142 매사추세츠주 캠브리지 8쓰 플로어 캠브리지 파크웨이 55
(72) 발명자
델라그레이브, 사이먼
미국 02142 매사추세츠 캠브리지 캠브리지 파크웨이 55 8쓰 플로어 스위트 800이
레보, 케빈, 제임스
미국 02142 매사추세츠 캠브리지 캠브리지 파크웨이 55 8쓰 플로어 스위트 800이
(뒷면에 계속)
(74) 대리인
양영준, 이상남

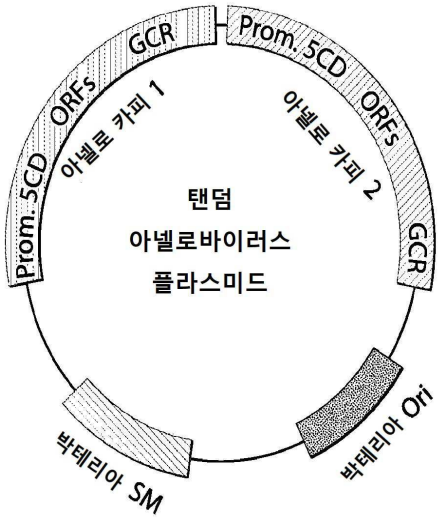
전체 청구항 수 : 총 10 항

(54) 발명의 명칭 **탠덤 아넬로바이러스 구축물**

(57) 요 약

본 발명은 아넬로벡터를 제조하는 조성물 및 이의 용도에 관한 것이다. 예를 들어, 본 명세서의 방법은 외인성 이펙터를 인코딩하는 제1 아넬로바이러스 게놈 및 나란히 배열된 제2 아넬로바이러스 게놈 또는 이의 단편을 포함하는 핵산 작제물을 제공하는 단계를 포함할 수 있다. 일부 구현예에서, 핵산 작제물은 단백질성 외부에 봉입된, 외인성 이펙터를 인코딩하는 아넬로바이러스 유전 요소를 포함하는 아넬로벡터의 생성을 초래한다.

대 표 도 - 도1a



(52) CPC특허분류
C12N 2750/00021 (2013.01)
C12N 2750/00043 (2013.01)
C12N 2750/00044 (2013.01)
(72) 발명자
나완다르, 다난제이, 마니클랄
미국 02142 매사추세츠 캠브리지 캠브리지 파크웨이 55 8쓰 플로어 스위트 800이
팀포나, 조셉, 루이스
미국 02142 매사추세츠 캠브리지 캠브리지 파크웨이 55 8쓰 플로어 스위트 800이

명 세 서

청구범위

청구항 1

a) 외인성 이펙터를 인코딩하는 서열을 포함하는 제1 아넬로바이러스 게놈;

b) 제1 아넬로바이러스 게놈과 나란히 배치된, 제2 아넬로바이러스 게놈 또는 이의 단편; 및

c) 선택적으로, (a)와 (b) 사이에 위치하는 스페이서 서열

을 포함하는 핵산(예를 들어, DNA) 작제물.

청구항 2

제1항에 있어서, 상기 제2 아넬로바이러스 게놈 또는 이의 단편은 2800, 2700, 2600, 2500, 2000, 1500, 1000, 900, 800, 700, 600, 또는 500개 미만 뉴클레오티드의 길이를 갖는, 핵산 작제물.

청구항 3

제1항 또는 제2항에 있어서, 상기 제2 아넬로바이러스 게놈 또는 이의 단편은 제1 아넬로바이러스 게놈에 대해 3'에 배치되는, 핵산 작제물.

청구항 4

제1항 내지 제3항 중 어느 한 항에 있어서, 상기 제2 아넬로바이러스 게놈 또는 이의 단편은 제1 아넬로바이러스 게놈에 대해 5'에 배치되는, 핵산 작제물.

청구항 5

제1항 내지 제4항 중 어느 한 항에 있어서, 상기 핵산 작제물은 스페이서 서열을 포함하는, 핵산 작제물.

청구항 6

제5항에 있어서, 상기 스페이서 서열은 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20개, 또는 그 이상의 아미노산의 길이, 또는 1 내지 5, 5 내지 10, 10 내지 15, 또는 15 내지 20개 아미노산의 길이를 갖는, 핵산 작제물.

청구항 7

제1항 내지 제6항 중 어느 한 항에 있어서, 상기 핵산 작제물은 스페이서 서열을 포함하지 않는, 핵산 작제물.

청구항 8

단백질성 외부에 봉입된 유전 요소를 포함하는 아넬로벡터의 제조 방법으로서,

a) 상기 구현예들 중 어느 하나의 핵산 작제물 및 아넬로바이러스 유전 요소(예를 들어, 여기서 아넬로바이러스 유전 요소는 핵산 작제물로부터 증폭된 것임)의 하나 이상의 카피를 포함하는 세포(예를 들어, 포유동물 숙주 세포)를 제공하는 단계;

b) 아넬로바이러스 유전 요소가 세포의 단백질성 외부에 봉입될 수 있도록 하는 조건 하에서 세포를 인큐베이션하는 단계

를 포함하여, 이에 의해 아넬로벡터를 제조하는 방법.

청구항 9

제1항 내지 제7항 중 어느 한 항의 핵산 작제물을 포함하는 세포.

청구항 10

외인성 이펙터를 세포로 운반하는 방법으로서, 제8항의 방법에 의해 제조된 아넬로벡터를 세포 내로 도입하는 단계 및 외인성 이펙터의 발현에 적합한 조건 하에서 세포를 인큐베이션하는 단계를 포함하는 방법.

발명의 설명

기 술 분 야

[0001] **관련 출원에 대한 상호 참조**

[0002] 본 출원은 2020년 6월 12일에 출원된 미국 가출원 63/038,483 및 2021년 2월 8일에 출원된 63/146,963의 이익을 주장한다. 전술한 출원의 내용은 본 명세서에 전문이 참조로 포함된다.

배 경 기 술

[0003] 환자에게 치료용 이펙터를 전달하기에 적합한 벡터를 제조하기 위한 조성물 및 방법의 개발에 대한 요구가 계속 존재한다. 바이러스 벡터의 생성은 일반적으로 단백질성 외부에 유전 요소를 봉입하는 것을 포함한다. 유전 요소의 생성은, 예를 들어, 백본 및 원하는 유전 요소 서열을 포함하는 플라스미드를 먼저 생성함으로써 달성될 수 있다. 그 다음, 백본 서열은 시험관 내 환화(IVC) 공정에서, 엔도뉴클레아제 절단 후 결찰을 수행함으로써 제거될 수 있다. 그러나, IVC는 때때로 대규모 공정에 적합하지 않다. 단백질성 외부에 봉입하기 위한 유전 요소를 생성하는 새로운 방법에 대한 요구가 당업계에 존재한다.

발명의 내용

[0004] 본 개시내용은 진핵 세포(예를 들어, 인간 세포 또는 인간 조직)에, 예를 들어, 유전 물질을 운반하기 위한, 이펙터, 예를 들어, 페이로드(payload)를 운반하기 위한, 또는 치료제 또는 치료용 이펙터를 운반하기 위한 운반 비히클로서 사용될 수 있는 아넬로벡터(예를 들어, 합성 아넬로벡터)를 생성하는 핵산 작제물을 제공한다. 일반적으로, 본 명세서에 기재된 핵산 작제물은 나란히 배열된 제1 카피의 유전 요소 서열(예를 들어, 돌연변이체 아넬로바이러스 게놈) 및 제2 카피의 유전 요소 서열(예를 들어, 아넬로바이러스 게놈 또는 이의 단편)의 적어도 일부를 포함한다. 이러한 구조를 갖는 핵산 작제물은 일반적으로 본 명세서에서 탠덤 작제물로 지칭된다. 이러한 탠덤 작제물은 아넬로벡터 유전 요소를 생성하는 데 사용된다. 제1 카피의 유전 요소 서열 및 제2 카피의 유전 요소 서열은, 일부 예에서 핵산 작제물에 서로 바로 인접할 수 있다. 다른 예에서, 제1 카피의 유전 요소 서열 및 제2 카피의 유전 요소 서열은, 예를 들어, 스페이서 서열에 의해 분리될 수 있다. 일부 구현예에서, 제2 카피의 유전 요소 또는 이의 일부는, 예를 들어, 본 명세서에 기재된 바와 같은, 상류 복제-촉진 서열(uRFS)을 포함한다. 일부 구현예에서, 제2 카피의 유전 요소 서열 또는 이의 일부는, 예를 들어, 본 명세서에 기재된 바와 같은, 하류 복제-촉진 서열(dRFS)을 포함한다. 일부 구현예에서, uRFS 및/또는 dRFS는 복제 원점(ORI)(예를 들어, 포유동물 ORI 또는 곤충 ORI) 또는 이의 일부를 포함한다. 일부 구현예에서, uRFS 및/또는 dRFS는 복제 원점을 포함하지 않는다. 일부 구현예에서, uRFS 및/또는 dRFS는 (예를 들어, 5' UTR에) 헤어핀 루프를 포함한다. 일부 구현예에서, 탠덤 작제물은 제2 카피의 유전 요소 또는 이의 일부가 결여된 다른 유사한 작제물보다 더 높은 수준의 유전 요소를 생성한다. 이론에 결부시키고자 하지 않고, 본 명세서에 기재된 탠덤 작제물은 회전환 복제(rolling circle replication)에 의해 복제될 수 있다. 일부 구현예에서, 탠덤 작제물은 하나 이상의 코돈-최적화된 오픈-리딩 프레임(예를 들어, 아넬로바이러스 ORF1, ORF1/1, ORF1/2, ORF2, ORF2/2, ORF2/3, 및/또는 ORF2t/3을 인코딩하며, 예를 들어, 포유동물에서의 발현을 위해 코돈 최적화된 서열)을 포함한다.

[0005] 아넬로벡터(예를 들어, 본 명세서에 기재된 바와 같은 탠덤 작제물을 사용하여 생성됨)는 일반적으로 유전 요소를 세포(예를 들어, 포유동물 세포, 예를 들어, 인간 세포) 내로 도입할 수 있는 단백질성 외부(예를 들어, 아넬로바이러스 캡시드 단백질, 예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어, 아넬로바이러스 ORF1 단백질 또는 아넬로바이러스 ORF1 핵산에 의해 인코딩된 폴리펩티드를 포함하는 단백질성 외부)에 캡슐화된 유전 요소(예를 들어, 이펙터, 예를 들어, 외인성 또는 내인성 이펙터, 예를 들어, 치료용 이펙터를 포함하거나 인코딩하는 유전 요소)를 포함한다. 일부 구현예에서, 아넬로벡터는 아넬로바이러스 ORF1 핵산(예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어, 알파토르크바이러스(*Alphatorquevirus*), 베타토르크바이러스(*Betatorquevirus*) 또는 감마토르크바이러스(*Gammatorquevirus*)의 ORF1 핵산, 예를 들어, 알파토르크바이러스 클레이드 1, 알파토르크바이러스 클레이드 2, 알파토르크바이러스 클레이드 3, 알파토르크바이러스 클레이드 4, 알파토르크바이러스 클레이드 5, 알파토르크바이러스 클레이드 6 또는 알파토르크바이러스 클레이드 7의 ORF1)

에 의해 인코딩된 폴리펩티드를 포함하는 단백질성 외부를 포함하는 입자 또는 감염성 비히클이다. 본 개시내용의 아넬로벡터의 유전 요소는 통상적으로 환형 및/또는 단일-가닥 DNA 분자(예를 들어, 환형 및 단일 가닥)이며, 일반적으로, 이를 둘러싸는 단백질성 외부 또는 이에 부착된 폴리펩티드에 결합하는 단백질 결합 서열을 포함하며, 이는 단백질성 외부 내의 유전 요소의 봉입 및/또는 단백질성 외부 내에서 다른 핵산에 비한 유전 요소의 농축을 용이하게 할 수 있다. 일부 구현예에서, 아넬로벡터의 유전 요소는, 예를 들어, 본 명세서에 기재된 바와 같은 탠덤 작제물을 사용하여 생성된다.

[0006] 일부 예에서, 유전 요소 작제물(예를 들어, 본 명세서에 기재된 바와 같은 탠덤 작제물)을 사용하여 유전 요소가 제공된다. 탠덤 작제물은, 일부 예에서, 제1 카피의 유전 요소 서열 및 제2 카피의 유전 요소 서열, 또는 이의 일부(예를 들어, uRFS 또는 dRFS)를 포함한다. 제2 카피는 제1 카피 또는 이의 일부의 동일한 카피일 수 있거나, 하나 이상의 서열 차이, 예를 들어, 치환을 포함할 수 있음이 이해된다. 일부 예에서, 제2 카피의 유전 요소 서열 또는 이의 일부(예를 들어, uRFS)는 제1 카피의 유전 요소 서열에 대해 5'에 위치한다. 일부 예에서, 제2 카피의 유전 요소 서열 또는 이의 일부(예를 들어, dRFS)는 제1 카피의 유전 요소 서열에 대해 3'에 위치한다. 일부 예에서, 제2 카피의 유전 요소 서열 또는 이의 일부 및 제1 카피의 유전 요소 서열은 탠덤 작제물에서 서로 인접해 있다. 일부 예에서, 제2 카피의 유전 요소 서열 또는 이의 일부 및 제1 카피의 유전 요소 서열은, 예를 들어, 스페이서 서열에 의해 분리되어 있다. 일부 예에서, 유전 요소 작제물은 환형 또는 선형이다. 일부 예에서, 유전 요소는 환형이다. 일부 예에서, 유전 요소는 단일-가닥이다. 일부 예에서, 유전 요소는 DNA이다. 일부 구현예에서, 단백질성 외부에 봉입하기에 적합한 유전 요소는 제1 카피의 유전 요소 서열의 회전환 복제를 통해 생성될 수 있다.

[0007] 일부 예에서, 유전 요소는, 예를 들어 세포 내에서 발현될 수 있는 이펙터(예를 들어, 핵산 이펙터, 예컨대 비-코딩 RNA 또는 폴리펩티드 이펙터, 예를 들어, 단백질)를 포함하거나 인코딩한다. 일부 구현예에서, 이펙터는, 예를 들어 본 명세서에 기재된 바와 같은, 치료제 또는 치료용 이펙터이다. 일부 구현예에서, 이펙터는, 예를 들어 야생형 아넬로바이러스 또는 표적 세포에 대하여, 내인성 이펙터 또는 외인성 이펙터이다. 일부 구현예에서, 이펙터는 야생형 아넬로바이러스 또는 표적 세포에 대하여 외인성이다. 일부 구현예에서, 아넬로벡터는 세포와 접촉시키고, 이펙터를 인코딩하는 유전 요소를 세포 내로 도입하여, 이펙터가 세포에 의해 제조되거나 발현되게 함으로써 이펙터를 세포 내로 운반할 수 있다. 특정 예에서, 이펙터는 내인성 이펙터이다(예를 들어, 표적 세포에 대하여 내인성이지만, 예를 들어, 아넬로벡터에 의해 증가된 양으로 제공됨). 다른 예에서, 이펙터는 외인성 이펙터이다. 이펙터는 일부 예에서, 세포의 기능을 조절하거나, 세포 내의 표적 분자의 활성 또는 수준을 조절할 수 있다. 예를 들어, 이펙터는 (예를 들어, PCT/US19/65995의 실시예 3 및 4에 기재된 바와 같이) 세포 내의 표적 단백질의 수준을 감소시킬 수 있다. 또 다른 예에서, 아넬로벡터는 (예를 들어, 실시예 15 및 19에 기재된 바와 같이) 생체 내에서 이펙터, 예를 들어, 외인성 단백질을 운반하고 발현할 수 있다. 아넬로벡터는 예를 들어, 유전 물질을 표적 세포, 조직 또는 대상체에 운반하거나; 이펙터를 표적 세포, 조직 또는 대상체에 운반하거나; 또는 예를 들어, 원하는 세포, 조직 또는 대상체에 대하여 치료제로서 작동할 수 있는 이펙터를 운반함으로써 질환 및 장애를 치료하기 위해 사용될 수 있다.

[0008] 일부 구현예에서, 본 명세서에 기재된 탠덤 작제물은, 예를 들어, 숙주 세포에서 합성 아넬로벡터의 유전 요소를 생성하는 데 사용될 수 있다. 합성 아넬로벡터는 야생형 바이러스(예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어, 야생형 아넬로바이러스)와 비교하여 적어도 하나의 구조적 차이, 예를 들어, 야생형 바이러스에 비하여 결실, 삽입, 치환, 변형(예를 들어, 효소적 변형)을 갖는다. 일반적으로, 합성 아넬로벡터는 단백질성 외부 내에 봉입된 외인성 유전 요소를 포함하며, 이는 유전 요소 또는 그 내에 인코딩된 이펙터(예를 들어, 외인성 이펙터 또는 내인성 이펙터)(예를 들어, 폴리펩티드 또는 핵산 이펙터)를 진핵(예를 들어, 인간) 세포 내로 운반하기 위해 사용될 수 있다. 구현예에서, 아넬로벡터는 검출 가능하고/가능하거나 원하지 않는 면역 또는 염증 반응을 야기하지 않으며, 예를 들어, 염증의 분자 마커(들), 예를 들어, TNF-알파, IL-6, IL-12, IFN, 및 B-세포 반응, 예를 들어, 반응성 또는 중화 항체의 1%, 5%, 10%, 15% 초과의 증가를 야기하지 않으며, 예를 들어, 아넬로벡터는 표적 세포, 조직 또는 대상체에 대하여 실질적으로 비-면역원성일 수 있다.

[0009] 일부 구현예에서, 본 명세서에 기재된 탠덤 작제물은 (i) 프로모터 요소 및 이펙터(예를 들어, 내인성 또는 외인성 이펙터)를 인코딩하는 서열, 및 단백질 결합 서열(예를 들어, 외부 단백질 결합 서열, 예를 들어 패키징 신호)을 포함하는 유전 요소; 및 (ii) 단백질성 외부를 포함하는 아넬로벡터의 유전 요소를 생성하는 데 사용될 수 있으며; 유전 요소는 단백질성 외부(예를 들어, 캡시드) 내에 봉입되고; 아넬로벡터는 유전 요소를 진핵(예를 들어, 포유동물, 예를 들어, 인간) 세포 내로 운반할 수 있다. 일부 구현예에서, 유전 요소는 단일-가닥 및/또는 환형 DNA이다. 대안적으로 또는 조합하여, 유전 요소는 하기의 특성 중 1개, 2개, 3개, 또는 전부를 갖는

다: 환형이고/이거나, 단일-가닥이고/이거나, 세포에 유입되는 유전 요소의 약 0.0001%, 0.001%, 0.005%, 0.01%, 0.05%, 0.1%, 0.5%, 1%, 1.5% 또는 2% 미만의 빈도로 세포의 게놈 내로 통합되고/되거나, 게놈당 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 또는 30개 카피 미만으로 표적 세포의 게놈 내로 통합됨. 일부 구현예에서, 통합 빈도는, 예를 들어 문헌[Wang et al., 2004, *Gene Therapy* 11: 711-721](본 명세서에 전문이 참조로 포함됨)에 기재된 바와 같이, 자유 벡터에서 분리된 게놈 DNA의 정량적 겔 정제 분석에 의해 결정된다. 일부 구현예에서, 유전 요소는 단백질성 외부 내에 봉입된다. 일부 구현예에서, 아넬로벡터는 유전 요소를 진핵 세포 내로 운반할 수 있다. 일부 구현예에서, 유전 요소는 야생형 아넬로바이러스의 서열(예를 들어, 야생형 토르크 테노 바이러스(TTV), 토르크 테노 미니 바이러스(TTMV) 또는 TTMDV 서열, 예를 들어, 본 명세서에 기재된 바와 같은 야생형 아넬로바이러스 서열)에 대하여 적어도 75%(예를 들어, 적어도 75, 76, 77, 78, 79, 80, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 또는 100%) 서열 동일성을 갖는 핵산 서열(예를 들어, 300 내지 4000개 뉴클레오티드, 예를 들어, 300 내지 3500개 뉴클레오티드, 300 내지 3000개 뉴클레오티드, 300 내지 2500개 뉴클레오티드, 300 내지 2000개 뉴클레오티드, 300 내지 1500개 뉴클레오티드의 핵산 서열)을 포함한다. 일부 구현예에서, 유전 요소는 야생형 아넬로바이러스의 서열(예를 들어, 본 명세서에 기재된 바와 같은, 야생형 아넬로바이러스 서열)에 대하여 적어도 75%(예를 들어, 적어도 75, 76, 77, 78, 79, 80, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 또는 100%) 서열 동일성을 갖는 핵산 서열(예를 들어, 적어도 300개 뉴클레오티드, 500개 뉴클레오티드, 1000개 뉴클레오티드, 1500개 뉴클레오티드, 2000개 뉴클레오티드, 2500개 뉴클레오티드, 3000개 뉴클레오티드 또는 그 이상의 핵산 서열)을 포함한다. 일부 구현예에서, 핵산 서열은 예를 들어, 포유동물(예를 들어, 인간) 세포에서의 발현을 위하여 코돈-최적화된다. 일부 구현예에서, 핵산 서열 내의 적어도 50%, 60%, 70%, 80%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%의 코돈은 예를 들어, 포유동물(예를 들어, 인간) 세포에서의 발현을 위하여 코돈-최적화된다.

[0010] 일부 구현예에서, 본 명세서에 기재된 탠덤 작제물은 (예를 들어, 인간 세포에 대한) 감염성 아넬로벡터, 비히클, 또는 캡시드(예를 들어, 아넬로바이러스 ORF, 예를 들어, ORF1, 폴리펩티드를 포함하는 캡시드)를 포함하는 입자의 유전 요소를 생성하여 캡시드에 결합하는 단백질 결합 서열 및 치료용 이펙터를 인코딩하는 (아넬로바이러스에 대한) 이종 서열을 포함하는 유전 요소를 캡슐화하는 데 사용될 수 있다. 구현예에서, 아넬로벡터는 유전 요소를 포유동물, 예를 들어, 인간, 세포 내로 운반할 수 있다. 일부 구현예에서, 유전 요소는 야생형 아넬로바이러스 게놈 서열에 대하여 약 6% 미만(예를 들어, 10%, 9%5, 8%, 7%, 6%, 5.5%, 5%, 4.5%, 4%, 3.5%, 3%, 2.5%, 2%, 1.5% 미만 또는 그 이하)의 동일성을 갖는다. 일부 구현예에서, 유전 요소는 야생형 아넬로바이러스 게놈 서열에 대하여 1.5%, 2%, 2.5%, 3%, 3.5%, 4%, 4.5%, 5%, 5.5% 또는 6% 이하의 동일성을 갖는다. 일부 구현예에서, 유전 요소는 야생형 아넬로바이러스에 대하여 적어도 약 2% 내지 적어도 약 5.5%(예를 들어, 2 내지 5%, 3% 내지 5%, 4% 내지 5%)의 동일성을 갖는다. 일부 구현예에서, 유전 요소는 약 2000, 3000, 4000, 4500 또는 5000개 초과의 비-바이러스 서열(예를 들어, 비 아넬로바이러스 게놈 서열)의 뉴클레오티드를 갖는다. 일부 구현예에서, 유전 요소는 약 2000 내지 5000개, 2500 내지 4500개, 3000 내지 4500개, 2500 내지 4500개, 3500 또는 4000개, 4500개(예를 들어, 약 3000 내지 4500개) 초과의 비-바이러스 서열(예를 들어, 비 아넬로바이러스 게놈 서열)의 뉴클레오티드를 갖는다. 일부 구현예에서, 유전 요소는 단일-가닥, 환형 DNA이다. 대안적으로 또는 조합하여, 유전 요소는 하기의 특성 중 1개, 2개 또는 3개를 갖는다: 환형이거나, 단일 가닥이거나, 세포에 유입되는 유전 요소의 약 0.001%, 0.005%, 0.01%, 0.05%, 0.1%, 0.5%, 1%, 1.5% 또는 2% 미만의 빈도로 세포의 게놈 내로 통합되거나, 게놈당 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 또는 30개 카피 미만으로 표적 세포의 게놈 내로 통합되거나, (예를 들어, 세포 용해물로부터의 유전 요소 서열에 대한 게놈 DNA 내로의 통합 빈도를 비교함으로써) 세포에 유입되는 유전 요소의 약 0.0001%, 0.001%, 0.005%, 0.01%, 0.05%, 0.1%, 0.5%, 1%, 1.5% 또는 2% 미만의 빈도로 통합됨. 일부 구현예에서, 통합 빈도는, 예를 들어, 문헌[Wang et al., 2004, *Gene Therapy* 11: 711-721](본 명세서에 전문이 참조로 포함됨)에 기재된 바와 같이, 자유 벡터에서 분리된 게놈 DNA의 정량적 겔 정제 분석에 의해 결정된다.

[0011] 일부 구현예에서, 본 명세서에 기재된 바와 같이, 아넬로바이러스 또는 아넬로벡터는 작용제, 예컨대 본 명세서에 기재된 이펙터를 표적 세포, 예를 들어, 치료적으로 또는 예방적으로 처치할 대상체 내의 표적 세포로 도입하기 위한 유효한 운반 비히클로서 사용될 수 있다.

[0012] 일부 구현예에서, 본 명세서에 기재된 탠덤 작제물은 하기를 (예를 들어, 연속하여) 포함하는 폴리펩티드(예를 들어, 합성 폴리펩티드, 예를 들어, ORF1 분자)를 포함하는 단백질성 외부를 포함하는 아넬로벡터의 유전 요소를 생성하는 데 사용될 수 있다:

[0013] (i) 아르기닌-풍부 영역, 예를 들어, 적어도 60%, 70% 또는 80%의 염기성 잔기(예를 들어, 아르기닌, 라이신 또

는 이의 조합)를 포함하는 적어도 약 40개의 아미노산의 서열을 포함하는 제1 영역,

[0014] (ii) 젤리-롤(jelly-roll) 도메인, 예를 들어, 적어도 6개의 베타 가닥을 포함하는 서열을 포함하는 제2 영역,

[0015] (iii) 본 명세서에 기재된 N22 도메인 서열을 포함하는 제3 영역,

[0016] (iv) 본 명세서에 기재된 아넬로바이러스 ORF1 C-말단 도메인(CTD) 서열을 포함하는 제4 영역, 및

[0017] (v) 선택적으로 폴리펩티드는, 예를 들어, 본 명세서에 기재된 바와 같은, 야생형 아넬로바이러스 ORF1 단백질에 대하여 100%, 99%, 98%, 95%, 90%, 85%, 80% 미만의 서열 동일성을 갖는 아미노산 서열을 가짐.

[0018] 일 양태에서, 본 발명은 이펙터, 예를 들어, 페이로드를 인코딩하는 서열에 작동 가능하게 연결된 프로모터 요소, 및 외부 단백질 결합 서열을 포함하는 유전 요소의 서열을 포함하는 단리된 핵산 분자(예를 들어, 핵산 작제물, 예를 들어, 탠덤 작제물)를 특징으로 한다. 일부 구현예에서, 외부 단백질 결합 서열은, 예를 들어, 본 명세서에 개시된 바와 같은, 아넬로바이러스의 5'UTR 서열에 대하여 적어도 75%(적어도 80%, 85%, 90%, 95%, 97%, 100%) 동일한 서열을 포함한다. 구현예에서, 유전 요소는 단일-가닥 DNA이고/이거나, 환형이고/이거나, 세포에 유입되는 유전 요소의 약 0.001%, 0.005%, 0.01%, 0.05%, 0.1%, 0.5%, 1%, 1.5% 또는 2% 미만의 빈도로 통합되고/되거나, 게놈당 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 또는 30개 카피 미만으로 표적 세포의 게놈 내로 통합되거나, 세포에 유입되는 유전 요소의 약 0.001%, 0.005%, 0.01%, 0.05%, 0.1%, 0.5%, 1%, 1.5% 또는 2% 미만의 빈도로 통합된다. 일부 구현예에서, 통합 빈도는, 예를 들어, 문헌[Wang et al., 2004, *Gene Therapy* 11: 711-721](본 명세서에 전문이 참조로 포함됨)에 기재된 바와 같이, 자유 벡터에서 분리된 게놈 DNA의 정량적 겔 정제 분석에 의해 결정된다. 구현예에서, 이펙터는 TTV로부터 기원하지 않으며, SV40-miR-S1이 아니다. 구현예에서, 핵산 분자는 TTMV-LY2의 폴리뉴클레오티드 서열을 포함하지 않는다. 구현예에서, 프로모터 요소는 진핵(예를 들어, 포유동물, 예를 들어, 인간) 세포에서 이펙터의 발현을 지시할 수 있다.

[0019] 일부 구현예에서, 핵산 분자는 환형이다. 일부 구현예에서, 핵산 분자는 선형이다. 일부 구현예에서, 본 명세서에 기재된 핵산 분자는 하나 이상의 변형된 뉴클레오티드(예를 들어, 염기 변형, 당 변형 또는 백본 변형)를 포함한다.

[0020] 일부 구현예에서, 핵산 분자는 ORF1 분자(예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어, 아넬로바이러스 ORF1 단백질)를 인코딩하는 서열을 포함한다. 일부 구현예에서, 핵산 분자는 ORF2 분자(예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어, 아넬로바이러스 ORF2 단백질)를 인코딩하는 서열을 포함한다. 일부 구현예에서, 핵산 분자는 ORF3 분자(예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어, 아넬로바이러스 ORF3 단백질)를 인코딩하는 서열을 포함한다. 일 양태에서, 본 발명은 (i) 프로모터 요소 및 이펙터, 예를 들어, 외인성 또는 내인성 이펙터를 인코딩하는 서열; (ii) 야생형 아넬로바이러스 서열에 대하여 적어도 75%(예를 들어, 적어도 75, 76, 77, 78, 79, 80, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 또는 100%) 서열 동일성을 갖는 적어도 72개의 인접 뉴클레오티드(예를 들어, 적어도 72, 73, 74, 75, 76, 77, 78, 79, 80, 90, 100 또는 150개 뉴클레오티드); 또는 야생형 아넬로바이러스 서열에 대하여 적어도 72%(예를 들어, 적어도 72, 73, 74, 75, 76, 77, 78, 79, 80, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 또는 100%) 서열 동일성을 갖는 적어도 100개(예를 들어, 적어도 300, 500, 1000, 1500개)의 인접 뉴클레오티드; 및 (iii) 단백질 결합 서열, 예를 들어, 외부 단백질 결합 서열 중 1개, 2개 또는 3개를 포함하는 유전 요소를 특징으로 하며, 핵산 작제물은 단일-가닥 DNA이며; 핵산 작제물은 환형이고/환형이거나, 세포에 유입되는 유전 요소의 약 0.001%, 0.005%, 0.01%, 0.05%, 0.1%, 0.5%, 1%, 1.5% 또는 2% 미만의 빈도로 통합되고/통합되거나, 게놈당 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 또는 30개 카피 미만으로 표적 세포의 게놈 내로 통합된다. 일부 구현예에서, 이펙터(예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어, 외인성 또는 내인성 이펙터)를 인코딩하는 유전 요소는 코돈 최적화된다. 일부 구현예에서, 유전 요소는 환형이다. 일부 구현예에서, 유전 요소는 선형이다. 일부 구현예에서, 본 명세서에 기재된 유전 요소는 하나 이상의 변형된 뉴클레오티드(예를 들어, 염기 변형, 당 변형 또는 백본 변형)를 포함한다. 일부 구현예에서, 유전 요소는 ORF1 분자(예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어, 아넬로바이러스 ORF1 단백질)를 인코딩하는 서열을 포함한다. 일부 구현예에서, 유전 요소는 ORF2 분자(예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어, 아넬로바이러스 ORF2 단백질)를 인코딩하는 서열을 포함한다. 일부 구현예에서, 유전 요소는 ORF3 분자(예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어, 아넬로바이러스 ORF3 단백질)를 인코딩하는 서열을 포함한다.

[0021] 일 양태에서, 본 발명은 본 명세서에 기재된 바와 같은 탠덤 작제물을 포함하는 숙주 세포를 특징으로 한다. 일부 구현예에서, 숙주 세포는 (a) ORF1 분자, ORF2 분자, 또는 ORF3 분자 중 하나 이상을 인코딩하는 서열(예를

들어, 본 명세서에 기재된 아넬로바이러스 ORF1 폴리펩티드를 인코딩하는 서열)을 포함하는 핵산(예를 들어, 핵산 분자는 플라스미드이거나, 바이러스 핵산이거나, 염색체에 통합됨); 및 (b) 유전 요소를 포함하며, 여기서 유전 요소는 (i) 이펙터(예를 들어, 외인성 이펙터 또는 내인성 이펙터)를 인코딩하는 핵산 서열(예를 들어, DNA 서열)에 작동 가능하게 연결된 프로모터 요소 및 (ii) (a)의 폴리펩티드에 결합하는 단백질 결합 서열을 포함하고, 선택적으로 유전 요소는 ORF1 폴리펩티드(예를 들어, ORF1 단백질)를 인코딩하지 않는다. 예를 들어, 숙주 세포는 (a) 및 (b)를 시스(동일한 핵산 분자의 두 부분) 또는 트랜스(상이한 핵산 분자의 각각의 부분)로 포함한다. 구현예에서, (b)의 유전 요소는 환형, 단일-가닥 DNA이다. 일부 구현예에서, 숙주 세포는, 예를 들어 본 명세서에 기재된 바와 같은 제조(manufacturing) 세포주이다. 일부 구현예에서, 숙주 세포는 부착성이거나, 현탁액 중에 존재하거나, 둘 다이다. 일부 구현예에서, 숙주 세포 또는 헬퍼 세포는 마이크로캐리어에서 성장된다. 일부 구현예에서, 숙주 세포 또는 헬퍼 세포는 cGMP 제조 실무와 양립 가능하다. 일부 구현예에서, 숙주 세포 또는 헬퍼 세포는 세포 성장을 촉진하기에 적합한 배지 중에 성장된다. 특정 구현예에서, 일단 숙주 세포 또는 헬퍼 세포가 (예를 들어, 적절한 세포 밀도로) 충분히 성장하면, 배지를 숙주 세포 또는 헬퍼 세포에 의한 아넬로벡터의 생성에 적합한 배지로 교체할 수 있다.

[0022] 일 양태에서, 본 발명은 본 명세서에 기재된 바와 같은 아넬로벡터(예를 들어, 합성 아넬로벡터)을 포함하는 약제학적 조성물을 특징으로 한다. 구현예에서, 약제학적 조성물은 약제학적으로 허용 가능한 담체 또는 부형제를 추가로 포함한다. 구현예에서, 약제학적 조성물은 표적 대상체 킬로그램당 약 10^5 내지 10^{14}개(예를 들어, 약 10^6 내지 10^{13}, 10^7 내지 10^{12}, 10^8 내지 10^{11}, 또는 10^9 내지 10^{10}개)의 게놈 당량의 아넬로벡터를 포함하는 단위 용량을 포함한다. 일부 구현예에서, 제제를 포함하는 약제학적 조성물은 허용 가능한 기간 및 온도에 걸쳐 안정할 것이고/이거나 요망되는 투여 경로 및/또는 투여 경로가 필요로 할 임의의 디바이스, 예를 들어, 니들(needle) 또는 주사기와 양립 가능할 것이다. 일부 구현예에서, 약제학적 조성물은 단회 용량 또는 다회 용량으로서의 투여를 위해 제형화된다. 일부 구현예에서, 약제학적 조성물은 투여 장소에서, 예를 들어, 의료 전문가(healthcare professional)에 의해 제형화된다. 일부 구현예에서, 약제학적 조성물은 (예를 들어, 부피당 게놈의 수에 의해 정의된 바와 같은) 요망되는 농도의 아넬로벡터 게놈 또는 게놈 당량을 포함한다.

[0023] 일 양태에서, 본 발명은 대상체에서의 질환 또는 장애의 치료 방법을 특징으로 하며, 방법은 예를 들어, 본 명세서에 기재된 바와 같은, 아넬로벡터, 예를 들어, 합성 아넬로벡터를 대상체에게 투여하는 단계를 포함한다.

[0024] 일 양태에서, 본 발명은 이펙터 또는 페이로드(예를 들어, 내인성 또는 외인성 이펙터)를 세포, 조직 또는 대상체에 운반하는 방법을 특징으로 하며, 방법은 예를 들어, 본 명세서에 기재된 바와 같은, 아넬로벡터, 예를 들어, 합성 아넬로벡터를 대상체에게 투여하는 단계를 포함하며, 아넬로벡터는 이펙터를 인코딩하는 핵산 서열을 포함한다. 구현예에서, 페이로드는 핵산이다. 구현예에서, 페이로드는 폴리펩티드이다.

[0025] 일 양태에서, 본 발명은 예를 들어, 본 명세서에 기재된 바와 같은 아넬로벡터, 예를 들어, 합성 아넬로벡터를 예를 들어, 생체 내에서 또는 생체 외에서, 세포, 예를 들어, 진핵 세포, 예를 들어, 포유동물 세포와 접촉시키는 단계를 포함하는 세포로의 아넬로벡터의 운반 방법을 특징으로 한다.

[0026] 일 양태에서, 본 발명은 아넬로벡터, 예를 들어, 합성 아넬로벡터의 제조 방법을 특징으로 한다. 방법은

[0027] (a) (i) 예를 들어, 본 명세서에 기재된 바와 같은 아넬로벡터의 유전 요소의 제1 카피의 핵산 서열, 및 아넬로벡터의 유전 요소의 제2 카피의 핵산 서열을 포함하는 제1 핵산 분자, 또는 이의 일부(예를 들어, uRFS 또는 dRFS); 및

[0028] (ii) 예를 들어, 본 명세서에 기재된 바와 같은, 아넬로바이러스 ORF1 폴리펩티드, 또는 ORF1, ORF2, ORF2/2, ORF2/3, ORF1/1, 또는 ORF1/2로부터 선택되는 하나 이상의 아미노산 서열, 또는 이에 대하여 적어도 70%(예를 들어, 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100%) 서열 동일성을 갖는 아미노산 서열을 인코딩하는 제2 핵산 분자

[0029] 를 포함하는 숙주 세포를 제공하는 단계; 및

[0030] (b) 유전 요소의 제1 카피의 핵산 서열의 복제(예를 들어, 회전환 복제)에 적합한 조건 하에서 숙주 세포를 인큐베이션하여, 유전 요소를 생성하는 단계; 및

[0031] 선택적으로 (c) 단백질성 외부(예를 들어, 제2 핵산 분자에 의해 인코딩되는 폴리펩티드를 포함함)에서 유전 요소의 봉입에 적합한 조건 하에서 숙주 세포를 인큐베이션하는 단계

[0032] 를 포함한다.

[0033] 일부 구현예에서, 제1 핵산 분자 및 제2 핵산 분자는 (예를 들어 본 명세서에 기재된 핵산 작제물에서, 예를 들어 시스로) 서로에 대해 부착된다. 일부 구현예에서, 제1 핵산 분자 및 제2 핵산 분자는 (예를 들어, 트랜스로) 분리되어 있다. 일부 구현예에서, 제1 핵산 분자는 플라스미드, 코스미드, 박미드, 미니서클, 또는 인공 염색체이다. 일부 구현예에서, 제2 핵산 분자는 플라스미드, 코스미드, 박미드, 미니서클, 또는 인공 염색체이다. 일부 구현예에서, 제2 핵산 분자는 숙주 세포의 게놈 내로 통합된다.

[0034] 일부 구현예에서, 방법은, 단계 (a) 이전에, 제1 핵산 분자 및/또는 제2 핵산 분자를 숙주 세포 내로 도입하는 단계를 추가로 포함한다. 일부 구현예에서, 제2 핵산 분자는 제1 핵산 분자 이전에, 동시에, 또는 이후에 숙주 세포 내로 도입된다. 다른 구현예에서, 제2 핵산 분자는 숙주 세포의 게놈에 통합된다. 일부 구현예에서, 제2 핵산 분자는 헬퍼 작제물, 헬퍼 바이러스 또는 다른 헬퍼 벡터이거나 이를 포함하거나 이의 일부이다.

[0035] 또 다른 양태에서, 본 발명은 아넬로벡터 조성물의 제조 방법을 특징으로 하며,

[0036] a) 예를 들어, 본 명세서에 기재된 바와 같은 아넬로벡터, 예를 들어, 합성 아넬로벡터의, 하나 이상의 구성성분(예를 들어 모든 구성성분)을 포함하는, 예를 들어, 이를 발현하는 숙주 세포를 제공하는 단계;

[0037] b) 숙주 세포로부터 아넬로벡터의 제제를 생성하기에 적합한 조건 하에서 숙주 세포를 배양하며, 여기서 제제의 아넬로벡터는 유전 요소(예를 들어, 본 명세서에 기재된 바와 같음)를 캡슐화하는 단백질성 외부(예를 들어, 아넬로벡터 ORF1 폴리펩티드를 포함함)를 포함하여, 아젤로벡터의 제제를 제조하는, 단계; 및

[0038] 선택적으로, c) 예를 들어, 대상체에게 투여하기에 적합한 약제학적 조성물로서, 아넬로벡터의 제제를 제형화하는 단계

[0039] 중 하나 이상(예를 들어, (a), (b), 및 (c) 전부)를 포함한다.

[0040] 예를 들어, 본 제조 방법에서 제공되는 숙주 세포는 (a) 본 명세서에 기재된 아넬로바이러스 ORF1 폴리펩티드를 인코딩하는 서열을 포함하는 핵산(여기서, 핵산은 플라스미드이거나, 바이러스 핵산 또는 게놈이거나, 헬퍼 세포 염색체에 통합됨); 및 (b) 유전 요소를 생성할 수 있는 탠덤 작제물을 포함하며, 여기서 유전 요소는 (i) 이펙터(예를 들어, 외인성 이펙터 또는 내인성 이펙터)를 인코딩하는 핵산 서열(예를 들어, DNA 서열)에 작동 가능하게 연결된 프로모터 요소 및 (i) (a)의 폴리펩티드에 결합하는 단백질 결합 서열(예를 들어, 패키징 서열)을 포함하고, 숙주 세포는 (a) 및 (b)를 시스 또는 트랜스로 포함한다. 구현예에서, (b)의 유전 요소는 환형, 단일-가닥 DNA이다. 일부 구현예에서, 숙주 세포는 제조 세포주이다.

[0041] 일부 구현예에서, 아넬로벡터의 구성성분은 생성 시에 (예를 들어, 일시적 형질감염에 의해) 숙주 세포 내로 도입된다. 일부 구현예에서, 숙주 세포는 아넬로벡터의 구성성분을 안정하게 발현한다(예를 들어, 아넬로벡터의 구성성분을 인코딩하는 하나 이상의 핵산은, 예를 들어, 안정한 형질감염에 의해 숙주 세포 또는 이의 자손 내로 도입됨).

[0042] 일 양태에서, 본 발명은 a) 본 명세서에 기재된 복수의 아넬로벡터 또는 본 명세서에 기재된 아넬로벡터의 제제를 제공하는 단계; 및 b) 아넬로벡터 또는 이의 제제를, 예를 들어, 대상체로의 투여에 적합한 약제학적 조성물로서 제형화하는 단계를 포함하는, 아넬로벡터 조성물의 제조 방법을 특징으로 한다.

[0043] 일 양태에서, 본 발명은 아넬로벡터를 포함하는 숙주 세포, 예를 들어, 제1 숙주 세포 또는 생성자 세포(예를 들어, PCT/US19/65995의 도 12에 나타낸 바와 같음), 예를 들어, 제1 숙주 세포의 집단의 제조 방법을 특징으로 하며, 방법은 예를 들어, 본 명세서에 기재된 바와 같은 유전 요소를 생성할 수 있는 탠덤 작제물을 숙주 세포 내로 도입하는 단계, 및 숙주 세포를 아넬로벡터의 생성에 적합한 조건 하에서 배양하는 단계를 포함한다. 구현예에서, 방법은 헬퍼, 예를 들어, 헬퍼 바이러스를, 숙주 세포 내로 도입하는 단계를 추가로 포함한다. 구현예에서, 도입하는 단계는 아넬로벡터를 이용한 숙주 세포의 형질감염(예를 들어, 화학적 형질감염) 또는 전기천공법을 포함한다.

[0044] 일 양태에서, 본 발명은, 예를 들어, 본 명세서에 기재된 바와 같은, 아넬로벡터를 포함하는 숙주 세포, 예를 들어, 제1 숙주 세포 또는 생성자 세포(예를 들어, PCT/US19/65995의 도 12에 나타낸 바와 같음)를 제공하는 단계, 및 숙주 세포로부터 아넬로벡터를 정제하는 단계를 포함하는 아넬로벡터의 제조 방법을 특징으로 한다. 일부 구현예에서, 방법은 제공하는 단계 이전에, 숙주 세포를, 예를 들어, 본 명세서에 기재된 바와 같은, 탠덤 작제물 또는 아넬로벡터와 접촉시키는 단계, 및 숙주 세포를 아넬로벡터의 생성에 적합한 조건 하에서 인큐베이션하는 단계를 추가로 포함한다. 구현예에서, 숙주 세포는 상기 숙주 세포의 제조 방법에 기재된 제1 숙주 세포

또는 생성자 세포이다. 구현예에서, 숙주 세포로부터 아넬로벡터를 정제하는 단계는 숙주 세포를 용해시키는 것을 포함한다.

[0045] 일부 구현예에서, 방법은 제1 숙주 세포 또는 생성자 세포에 의해 생성되는 아넬로벡터를 제2 숙주 세포, 예를 들어, 허용 세포(예를 들어, PCT/US19/65995의 도 12에 나타낸 바와 같음), 예를 들어, 제2 숙주 세포의 집단과 접촉시키는 제2 단계를 추가로 포함한다. 일부 구현예에서, 방법은 제2 숙주 세포를 아넬로벡터의 생성에 적합한 조건 하에서 인큐베이션하는 단계를 추가로 포함한다. 일부 구현예에서, 방법은 제2 숙주 세포로부터 아넬로벡터를 정제함으로써, 예를 들어, 아넬로벡터 시드 집단을 생성하는 단계를 추가로 포함한다. 구현예에서, 제1 숙주 세포의 집단에 비해 제2 숙주 세포의 집단으로부터 적어도 약 2 내지 100배 더 많은 아넬로벡터가 생성된다. 구현예에서, 제2 숙주 세포로부터 아넬로벡터를 정제하는 단계는 제2 숙주 세포를 용해시키는 것을 포함한다. 일부 구현예에서, 방법은 제2 숙주 세포에 의해 생성되는 아넬로벡터를 제3 숙주 세포, 예를 들어, 허용 세포(예를 들어, PCT/US19/65995의 도 12에 나타낸 바와 같음), 예를 들어, 제3 숙주 세포의 집단과 접촉시키는 제2 단계를 추가로 포함한다. 일부 구현예에서, 방법은 제3 숙주 세포를 아넬로벡터의 생성에 적합한 조건 하에서 인큐베이션하는 단계를 추가로 포함한다. 일부 구현예에서, 방법은 제3 숙주 세포로부터 아넬로벡터를 정제함으로써, 예를 들어, 아넬로벡터 스톡(stock) 집단을 생성하는 단계를 추가로 포함한다. 구현예에서, 제3 숙주 세포로부터 아넬로벡터를 정제하는 단계는 제3 숙주 세포를 용해시키는 것을 포함한다. 구현예에서, 제2 숙주 세포의 집단에 비해 제3 숙주 세포의 집단으로부터 적어도 약 2 내지 100배 더 많은 아넬로벡터가 생성된다.

[0046] 일부 구현예에서, 숙주 세포는 세포 성장을 촉진하기에 적합한 배지에서 성장된다. 특정 구현예에서, 일단 숙주 세포가 (예를 들어, 적절한 세포 밀도로) 충분하게 성장되면, 배지를 숙주 세포에 의한 아넬로벡터의 생성에 적합한 배지로 교체할 수 있다. 일부 구현예에서, 숙주 세포에 의해 생성되는 아넬로벡터는 제2 숙주 세포와의 접촉 이전에 (예를 들어, 숙주 세포를 용해시킴으로써) 숙주 세포로부터 분리된다. 일부 구현예에서, 숙주 세포에 의해 생성되는 아넬로벡터는 개재되는 정제 단계 없이 제2 숙주 세포와 접촉된다.

[0047] 일 양태에서, 본 발명은 약제학적 아넬로벡터 제제의 제조 방법을 특징으로 한다. 방법은 (a) 본 명세서에 기재된 바와 같은 아넬로벡터 제제를 제조하는 단계, (b) 제제(예를 들어, 약제학적 아넬로벡터 제제, 아넬로벡터 시드 집단 또는 아넬로벡터 스톡 집단)를 하나 이상의 약제학적 품질 관리 파라미터, 예를 들어, 아이덴티티, 순도, 역가, 효력(예를 들어, 아넬로벡터 입자당 게놈 당량) 및/또는 예를 들어, 아넬로벡터에 포함되는 유전 요소로부터의 핵산 서열에 대하여 평가하는 단계, 및 (c) 평가가 사전결정된 기준을 만족시키는, 예를 들어, 약제학적 규격을 만족시키는 약제학적 이용을 위한 제제를 제형화하는 단계를 포함한다. 일부 구현예에서, 아이덴티티를 평가하는 것은 아넬로벡터의 유전 요소의 서열, 예를 들어, 이펙터를 인코딩하는 서열을 평가하는 것(예를 들어, 확인하는 것)을 포함한다. 일부 구현예에서, 순도를 평가하는 것은 불순물, 예를 들어, 마이코플라스마, 내독소, 숙주 세포 핵산(예를 들어, 숙주 세포 DNA 및/또는 숙주 세포 RNA), 동물-유래 과정 불순물(예를 들어, 혈청 알부민 또는 트립신), 복제-적격 작용제(RCA), 예를 들어, 복제-적격 바이러스 또는 원치 않는 아넬로벡터(예를 들어, 요망되는 아넬로벡터, 예를 들어, 본 명세서에 기재된 바와 같은 합성 아넬로벡터 이외의 아넬로벡터), 유리 바이러스 캡시드 단백질, 외래성 물질(adventitious agent) 및 응집물의 양을 평가하는 것을 포함한다. 일부 구현예에서, 역가를 평가하는 것은 (예를 들어, HPLC에 의해 평가되는 바와 같은) 제제 내의 기능성 대 비-기능성(예를 들어, 감염성 대 비-감염성) 아넬로벡터의 비를 평가하는 것을 포함한다. 일부 구현예에서, 효력을 평가하는 것은 제제에서 검출 가능한 아넬로벡터 기능(예를 들어, 그 내에 인코딩된 이펙터의 발현 및/또는 기능 또는 게놈 당량)의 수준을 평가하는 것을 포함한다.

[0048] 구현예에서, 제형화된 제제는 병원체, 숙주 세포 오염물질 또는 불순물이 실질적으로 없거나; 사전결정된 비-감염성 입자의 수준 또는 사전결정된 입자:감염성 유닛의 비(예를 들어, 300:1 미만, 200:1 미만, 100:1 미만 또는 50:1 미만)를 갖는다. 일부 구현예에서, 다중 아넬로벡터가 단일의 배치(batch)에서 생성될 수 있다. 구현예에서, 배치에서 생성되는 아넬로벡터의 수준이 (예를 들어, 개별적으로 또는 함께) 평가될 수 있다.

[0049] 일 양태에서, 본 발명은 하기를 포함하는 숙주 세포를 특징으로 한다:

[0050] (i) 본 명세서에 기재된 바와 같은 탠덤 작제물을 포함하는 제1 핵산 분자, 및

[0051] (ii) 선택적으로, 예를 들어 본 명세서에 기재된 바와 같은 ORF1, ORF2, ORF2/2, ORF2/3, ORF1/1 또는 ORF1/2로부터 선택되는 아미노산 서열 또는 이에 대하여 적어도 약 70%(예를 들어, 적어도 약 70, 80, 90, 95, 96, 97, 98, 99 또는 100%) 서열 동일성을 갖는 아미노산 서열 중 하나 이상을 인코딩하는 제2 핵산 분자.

[0052] 일 양태에서, 본 발명은 본 명세서에 기재된 아넬로벡터 및 헬퍼 바이러스를 포함하는 반응 혼합물을 특징으로

하며, 헬퍼 바이러스는 외부 단백질(예를 들어, 외부 단백질 결합 서열에 결합할 수 있는 외부 단백질 및 선택적으로 지질 외피)을 인코딩하는, 폴리뉴클레오티드, 복제 단백질(예를 들어, 중합효소)을 인코딩하는 폴리뉴클레오티드 또는 이의 임의의 조합을 포함한다.

[0053] 일부 구현예에서, 아넬로벡터(예를 들어, 합성 아넬로벡터)는 단리되며, 예를 들어, 숙주 세포로부터 단리되고/되거나 용액(예를 들어, 상청액) 중 다른 구성요소로부터 단리된다. 일부 구현예에서, 아넬로벡터(예를 들어, 합성 아넬로벡터)는 예를 들어, 용액(예를 들어, 상청액)으로부터 정제된다. 일부 구현예에서, 아넬로벡터는 용액 중 다른 구성요소에 비하여 용액 중에서 농축된다.

[0054] 상기 아넬로벡터, 조성물 또는 방법 중 임의의 것의 일부 구현예에서, 아넬로벡터를 제공하는 것은, 예를 들어 본 명세서에 기재된 바와 같은, 아넬로벡터-생성 세포를 포함하는 조성물로부터 아넬로벡터를 분리하는 것(예를 들어, 수확하는 것)을 포함한다. 다른 구현예에서, 아넬로벡터를 제공하는 것은, 예를 들어 제3자로부터 아넬로벡터 또는 이의 제제를 수득하는 것을 포함한다.

[0055] 상기 아넬로벡터, 조성물 또는 방법 중 임의의 것의 일부 구현예에서, 유전 요소는, 예를 들어, 본 명세서에 기재된 방법에 따라 확인되는 바와 같은, 아넬로벡터 게놈을 포함한다. 구현예에서, 아넬로벡터 게놈은 TTV-tth8 핵산 서열, 예를 들어, TTV-tth8 핵산 서열의 뉴클레오티드 3436 내지 3707의 적어도 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 99% 또는 100%의 결실을 갖는, TTV-tth8 핵산 서열을 포함한다. 구현예에서, 아넬로벡터 게놈은 TTMV-LY2 핵산 서열, 예를 들어, TTMV-LY2 핵산 서열의 뉴클레오티드 574 내지 1371, 1432 내지 2210, 574 내지 2210, 및/또는 2610 내지 2809의 적어도 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 99% 또는 100%의 결실을 갖는, 예를 들어, TTMV-LY2 핵산 서열을 포함한다. 구현예에서, 유전 요소는 자가-복제할 수 있고/있거나 자가-증폭할 수 있다. 구현예에서, 유전 요소는 자가-복제할 수 없고/없거나 자가-증폭할 수 없다. 구현예에서, 유전 요소는, 예를 들어, 헬퍼, 예를 들어, 헬퍼 바이러스의 존재 하에 트랜스로 복제할 수 있고/있거나 증폭할 수 있다.

[0056] 상기 아넬로벡터, 조성물 또는 방법 중 임의의 것의 추가의 특징은 하기 열거된 구현예 중 하나 이상을 포함한다.

[0057] 당업자는 단지 일상적인 실험을 사용하여, 본 명세서에 기재된 본 발명의 특정 구현예의 다수의 등가물을 인식하거나 확인할 수 있을 것이다. 그러한 등가물은 하기 열거된 구현예에 의해 포함되는 것으로 의도된다.

[0058] **열거된 구현예**

[0059] 1. 핵산(예를 들어, DNA) 작제물로서,

[0060] a) 외인성 이펙터를 인코딩하는 서열을 포함하는, 선택적으로 돌연변이체인, 제1 아넬로바이러스 게놈;

[0061] b) 제1 아넬로바이러스 게놈과 나란히 배치된, 제2 아넬로바이러스 게놈 또는 이의 단편(예를 들어, 이의 약 10 내지 20, 20 내지 30, 30 내지 40, 40 내지 50, 50 내지 60, 60 내지 70, 70 내지 80, 80 내지 90, 90 내지 100, 100 내지 125, 125 내지 150, 150 내지 175, 175 내지 200, 200 내지 250, 250 내지 300, 300 내지 400, 400 내지 500, 500 내지 600, 600 내지 700, 700 내지 800, 800 내지 900, 900 내지 1000, 1000 내지 1200, 1200 내지 1400, 1400 내지 1600, 1600 내지 1800, 1800 내지 2000, 2000 내지 2200, 2200 내지 2400, 2400 내지 2600, 2600 내지 2700, 2700 내지 2800, 2800 내지 2900, 또는 2900 내지 3000개의 인접 뉴클레오티드를 포함함); 및

[0062] c) 선택적으로, (a)와 (b) 사이에 위치하는 스페이서 서열

[0063] 을 포함하고;

[0064] 선택적으로, (i) 제1 아넬로바이러스 게놈은 제2 아넬로바이러스 게놈 또는 이의 단편에 대해 5'에 배치되거나, (ii) 제1 아넬로바이러스 게놈은 제2 아넬로바이러스 게놈 또는 이의 단편에 대해 3'에 배치되는 핵산 작제물.

[0065] 2. 핵산(예를 들어, DNA) 작제물로서,

[0066] a) i) 제1 아넬로바이러스 상류 복제-촉진 서열(uRFS), 예를 들어, 5' UTR;

[0067] ii) 외인성 이펙터를 인코딩하는 서열에 작동 가능하게 연결된 프로모터;

[0068] iii) 제1 아넬로바이러스 하류 복제-촉진 서열(dRFS), 예를 들어, 3' UTR

[0069] 을 포함하는 아넬로바이러스 유전 요소 영역; 및

[0070] b) 제2 아넬로바이러스 uRFS(예를 들어, 5' UTR) 또는 제2 아넬로바이러스 dRFS(예를 들어, 3' UTR); 및

[0071] c) 선택적으로, 아넬로바이러스 유전 요소 영역과 (b) 사이에 위치하는 스페이서 서열

[0072] 을 포함하는 핵산 작제물.

[0073] 3. 구현예 1 또는 2에 있어서, d) 핵산 작제물의 복제에 적합한 백본 영역, 예를 들어, 플라스미드 백본 또는 박미드 백본을 추가로 포함하는, 핵산 작제물.

[0074] 4. 구현예 2에 있어서, uRFS는 아넬로바이러스 Rep 단백질에 결합하는, 핵산 작제물.

[0075] 5. 구현예 2에 있어서, uRFS는 복제 원점(ORI)을 포함하는, 핵산 작제물.

[0076] 6. 구현예 2에 있어서, uRFS는 복제 원점을 포함하지 않는, 핵산 작제물.

[0077] 7. 구현예 2에 있어서, uRFS는 (예를 들어, 5' UTR에) 헤어핀 루프를 포함하는, 핵산 작제물.

[0078] 8. 구현예 2에 있어서, uRFS 및 dRFS는 함께 아넬로바이러스 Rep 변위 부위를 포함하는, 핵산 작제물.

[0079] 9. 구현예 2에 있어서, uRFS 및 dRFS는 함께 아넬로바이러스 Rep 결합 부위를 포함하는, 핵산 작제물.

[0080] 10. 구현예 2에 있어서, uRFS 및 dRFS는 함께 아넬로바이러스 Rep 복제 개시 부위를 포함하는, 핵산 작제물.

[0081] 11. 구현예 1 또는 2에 있어서, 핵산 작제물이 복제를 가능하게 하는 조건 하에서 숙주 세포 내로 도입될 때, 예를 들어, 적어도 3:1, 4:1, 5:1, 6:1, 7:1, 8:1, 9:1, 10:1, 20:1, 30:1, 40:1, 50:1, 100:1, 500:1. 1000:1, 5000:1, 10,000:1, 100,000:1, 또는 1,000,000:1의 비율로, 백본보다 더 높은 수준의 유전 요소가 관찰되는, 핵산 작제물.

[0082] 12. 상기 구현예들 중 어느 하나에 있어서, 제1 uRFS는 제2 uRFS와 동일한 서열을 갖는, 핵산 작제물.

[0083] 13. 구현예 1 내지 11 중 어느 하나에 있어서, 제1 uRFS는 제2 uRFS와 상이한 서열을 갖는, 핵산 작제물.

[0084] 14. 상기 구현예들 중 어느 하나에 있어서, 핵산 작제물은 하나 이하의 dRFS를 포함하는, 핵산 작제물.

[0085] 15. 구현예 1 내지 13 중 어느 하나에 있어서, 제1 dRFS는 제2 dRFS와 동일한 서열을 갖는, 핵산 작제물.

[0086] 16. 상기 구현예들 중 어느 하나에 있어서, 핵산 작제물은 하나 이상의 uRFS를 포함하는, 핵산 작제물.

[0087] 17. 상기 구현예들 중 어느 하나에 있어서, uRFS 또는 dRFS는 전장 아넬로바이러스 유전 요소를 포함하는, 핵산.

[0088] 18. 상기 구현예들 중 어느 하나에 있어서, uRFS 또는 dRFS는 부분적인 아넬로바이러스 유전 요소를 포함하는, 핵산.

[0089] 19. 상기 구현예들 중 어느 하나에 있어서, dRFS는 5' UTR(예를 들어, 헤어핀 루프를 포함함), ORF2, 인트론, ORF2-2, ORF3, 3' UTR, 및 GC-풍부 영역 서열을 포함하는, 핵산 작제물.

[0090] 20. 상기 구현예들 중 어느 하나에 있어서, dRFS는 5' UTR(예를 들어, 헤어핀 루프를 포함함), ORF2, 인트론, ORF2-2, ORF3, 및 3' UTR 서열을 포함하는, 핵산 작제물.

[0091] 21. 상기 구현예들 중 어느 하나에 있어서, dRFS는 5' UTR(예를 들어, 헤어핀 루프를 포함함), ORF2, 인트론, ORF2-2, 및 ORF3 서열을 포함하는, 핵산 작제물.

[0092] 22. 상기 구현예들 중 어느 하나에 있어서, dRFS는 5' UTR(예를 들어, 헤어핀 루프를 포함함), ORF2, 인트론, 및 ORF2-2 서열을 포함하는, 핵산 작제물.

[0093] 23. 상기 구현예들 중 어느 하나에 있어서, dRFS는 5' UTR(예를 들어, 헤어핀 루프를 포함함), ORF2, 및 인트론 서열을 포함하는, 핵산 작제물.

[0094] 24. 상기 구현예들 중 어느 하나에 있어서, dRFS는 5' UTR(예를 들어, 헤어핀 루프를 포함함) 및 ORF2 서열을 포함하는, 핵산 작제물.

[0095] 25. 상기 구현예들 중 어느 하나에 있어서, dRFS는 5' UTR(예를 들어, 헤어핀 루프를 포함함)을 포함하는, 핵산

작제물.

[0096] 26. 구현예 19 내지 25 중 어느 하나에 있어서, dRFS는 복제 원점을 포함하는 5' UTR을 포함하는, 핵산.

[0097] 27. 상기 구현예들 중 어느 하나에 있어서, dRFS는 GC-풍부 영역 서열을 포함하지 않는, 핵산 작제물.

[0098] 28. 상기 구현예들 중 어느 하나에 있어서, dRFS는 3' UTR 및 GC-풍부 영역 서열을 포함하지 않는, 핵산 작제물.

[0099] 29. 상기 구현예들 중 어느 하나에 있어서, dRFS는 ORF3, 3' UTR, 및 GC-풍부 영역 서열을 포함하지 않는, 핵산 작제물.

[0100] 30. 상기 구현예들 중 어느 하나에 있어서, dRFS는 ORF2-2, ORF3, 3' UTR, 및 GC-풍부 영역 서열을 포함하지 않는, 핵산 작제물.

[0101] 31. 상기 구현예들 중 어느 하나에 있어서, dRFS는 인트론, ORF2-2, ORF3, 3' UTR, 및 GC-풍부 영역 서열을 포함하지 않는, 핵산 작제물.

[0102] 32. 상기 구현예들 중 어느 하나에 있어서, dRFS는 ORF2, 인트론, ORF2-2, ORF3, 3' UTR, 및 GC-풍부 영역 서열을 포함하지 않는, 핵산 작제물.

[0103] 33. 상기 구현예들 중 어느 하나에 있어서, dRFS는 5' UTR(예를 들어, 헤어핀 루프 및/또는 복제 원점을 포함함), ORF2, 인트론, ORF2-2, ORF3, 3' UTR, 및 GC-풍부 영역 서열을 포함하지 않는, 핵산 작제물.

[0104] 34. 상기 구현예들 중 어느 하나에 있어서, uRFS는 5' UTR(예를 들어, 헤어핀 루프 및/또는 복제 원점을 포함함), ORF2, 인트론, ORF2-2, ORF3, 3' UTR, 및 GC-풍부 영역 서열을 포함하는, 핵산 작제물.

[0105] 35. 상기 구현예들 중 어느 하나에 있어서, uRFS는 ORF2, 인트론, ORF2-2, ORF3, 3' UTR, 및 GC-풍부 영역 서열을 포함하는, 핵산 작제물.

[0106] 36. 상기 구현예들 중 어느 하나에 있어서, uRFS는 인트론, ORF2-2, ORF3, 3' UTR, 및 GC-풍부 영역 서열을 포함하는, 핵산 작제물.

[0107] 37. 상기 구현예들 중 어느 하나에 있어서, uRFS는 ORF2-2, ORF3, 3' UTR, 및 GC-풍부 영역 서열을 포함하는, 핵산 작제물.

[0108] 38. 상기 구현예들 중 어느 하나에 있어서, uRFS는 ORF3, 3' UTR, 및 GC-풍부 영역 서열을 포함하는, 핵산 작제물.

[0109] 39. 상기 구현예들 중 어느 하나에 있어서, uRFS는 3' UTR, 및 GC-풍부 영역 서열을 포함하는, 핵산 작제물.

[0110] 40. 상기 구현예들 중 어느 하나에 있어서, uRFS는 GC-풍부 영역 서열을 포함하는, 핵산 작제물.

[0111] 41. 상기 구현예들 중 어느 하나에 있어서, uRFS는 5' UTR(예를 들어, 헤어핀 루프를 포함함)을 포함하지 않는, 핵산 작제물.

[0112] 42. 상기 구현예들 중 어느 하나에 있어서, uRFS는 5' UTR(예를 들어, 헤어핀 루프를 포함함) 및 ORF2 서열을 포함하지 않는, 핵산 작제물.

[0113] 43. 상기 구현예들 중 어느 하나에 있어서, uRFS는 5' UTR(예를 들어, 헤어핀 루프를 포함함), ORF2, 및 인트론 서열을 포함하지 않는, 핵산 작제물.

[0114] 44. 상기 구현예들 중 어느 하나에 있어서, uRFS는 5' UTR(예를 들어, 헤어핀 루프를 포함함), ORF2, 인트론, 및 ORF2-2 서열을 포함하지 않는, 핵산 작제물.

[0115] 45. 상기 구현예들 중 어느 하나에 있어서, uRFS는 5' UTR(예를 들어, 헤어핀 루프를 포함함), ORF2, 인트론, ORF2-2, 및 ORF3 서열을 포함하지 않는, 핵산 작제물.

[0116] 46. 상기 구현예들 중 어느 하나에 있어서, uRFS는 5' UTR(예를 들어, 헤어핀 루프를 포함함), ORF2, 인트론, ORF2-2, ORF3, 및 3' UTR 서열을 포함하지 않는, 핵산 작제물.

[0117] 47. 상기 구현예들 중 어느 하나에 있어서, uRFS는 5' UTR(예를 들어, 헤어핀 루프를 포함함), ORF2, 인트론, ORF2-2, ORF3, 3' UTR, 및 GC-풍부 영역 서열을 포함하고, dRFS는 5' UTR(예를 들어, 헤어핀 루프를 포함함)을

포함하는, 핵산 작제물.

[0118] 48. 상기 구현예들 중 어느 하나에 있어서, uRFS는 5' UTR(예를 들어, 헤어핀 루프를 포함함), ORF2, 인트론, ORF2-2, ORF3, 3' UTR, 및 GC-풍부 영역 서열을 포함하고, dRFS는 5' UTR(예를 들어, 헤어핀 루프를 포함함) 및 ORF2 서열을 포함하는, 핵산 작제물.

[0119] 49. 상기 구현예들 중 어느 하나에 있어서, uRFS는 5' UTR(예를 들어, 헤어핀 루프를 포함함), ORF2, 인트론, ORF2-2, ORF3, 3' UTR, 및 GC-풍부 영역 서열을 포함하고, dRFS는 5' UTR(예를 들어, 헤어핀 루프를 포함함), ORF2, 및 인트론 서열을 포함하는, 핵산 작제물.

[0120] 50. 상기 구현예들 중 어느 하나에 있어서, uRFS는 5' UTR(예를 들어, 헤어핀 루프를 포함함), ORF2, 인트론, ORF2-2, ORF3, 3' UTR, 및 GC-풍부 영역 서열을 포함하고, dRFS는 5' UTR(예를 들어, 헤어핀 루프를 포함함), ORF2, 인트론, 및 ORF2-2 서열을 포함하는, 핵산 작제물.

[0121] 51. 상기 구현예들 중 어느 하나에 있어서, uRFS는 5' UTR(예를 들어, 헤어핀 루프를 포함함), ORF2, 인트론, ORF2-2, ORF3, 3' UTR, 및 GC-풍부 영역 서열을 포함하고, dRFS는 5' UTR(예를 들어, 헤어핀 루프를 포함함), ORF2, 인트론, ORF2-2, 및 ORF3 서열을 포함하는, 핵산 작제물.

[0122] 52. 상기 구현예들 중 어느 하나에 있어서, uRFS는 5' UTR(예를 들어, 헤어핀 루프를 포함함), ORF2, 인트론, ORF2-2, ORF3, 3' UTR, 및 GC-풍부 영역 서열을 포함하고, dRFS는 5' UTR(예를 들어, 헤어핀 루프를 포함함), ORF2, 인트론, ORF2-2, ORF3, 및 3' UTR 서열을 포함하는, 핵산 작제물.

[0123] 53. 구현예 41 내지 52 중 어느 하나에 있어서, uRFS의 5' UTR은 복제 원점을 포함하는, 핵산 작제물.

[0124] 54. 구현예 41 내지 52 중 어느 하나에 있어서, uRFS의 5' UTR은 복제 원점을 포함하지 않는, 핵산 작제물.

[0125] 55. 구현예 41 내지 54 중 어느 하나에 있어서, dRFS의 5' UTR은 복제 원점을 포함하는, 핵산 작제물.

[0126] 56. 구현예 41 내지 54 중 어느 하나에 있어서, dRFS의 5' UTR은 복제 원점을 포함하지 않는, 핵산 작제물.

[0127] 57. 상기 구현예들 중 어느 하나에 있어서, uRFS는, 100 내지 200, 200 내지 300, 300 내지 400, 400 내지 500, 500 내지 600, 600 내지 700, 700 내지 800, 800 내지 900, 900 내지 1000, 1000 내지 1100, 1100 내지 1200, 1200 내지 1300, 1300 내지 1400, 1400 내지 1500, 1500 내지 1600, 1600 내지 1700, 1700 내지 1800, 1800 내지 1900, 1900 내지 2000, 2000 내지 2200, 2200 내지 2400, 2400 내지 2500, 2500 내지 2600, 2600 내지 2800, 또는 2800 내지 3000 kb의, 예를 들어, 본 명세서에 기재된 바와 같은 유전 요소 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 서열 동일성을 갖는 서열을 포함하는, 핵산 작제물.

[0128] 58. 상기 구현예들 중 어느 하나에 있어서, dRFS는, 100 내지 200, 200 내지 300, 300 내지 400, 400 내지 500, 500 내지 600, 600 내지 700, 700 내지 800, 800 내지 900, 900 내지 1000, 1000 내지 1100, 1100 내지 1200, 1200 내지 1300, 1300 내지 1400, 1400 내지 1500, 1500 내지 1600, 1600 내지 1700, 1700 내지 1800, 1800 내지 1900, 1900 내지 2000, 2000 내지 2200, 2200 내지 2400, 2400 내지 2500, 2500 내지 2600, 2600 내지 2800, 또는 2800 내지 3000 kb의, 예를 들어, 본 명세서에 기재된 바와 같은 유전 요소 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 서열 동일성을 갖는 서열을 포함하는, 핵산 작제물.

[0129] 59. 핵산(예를 들어, DNA) 작제물로서.

[0130] a) 외인성 이펙터를 인코딩하는 서열을 포함하는 제1 아넬로바이러스 유전 요소 영역;

[0131] b) 제2 아넬로바이러스 유전 요소 영역 또는 이의 단편; 및

[0132] c) 선택적으로, (a) 및 (b) 사이에 위치하는 스페이서 서열

[0133] 을 포함하는 핵산 작제물.

[0134] 60. 구현예 59에 있어서, d) 백본 영역을 추가로 포함하며, 예를 들어, 백본 영역은 박테리아, 포유동물 세포, 또는 곤충 세포에서 DNA 작제물의 복제에 적합한 것인, 핵산 작제물.

[0135] 61. 구현에 59에 있어서,

[0136] 아넬로바이러스 유전 요소의 서열을 포함하는 유전 요소 영역,

[0137] 선택적으로, 스페이서 서열;

[0138] 아넬로벡터 탠덤 영역; 및

[0139] 백본 영역

[0140] 을 상기 순서로 포함하는, 핵산 작제물.

[0141] 62. 구현예 59에 있어서,

[0142] 아넬로벡터 탠덤 영역; 및

[0143] 선택적으로, 스페이서 서열;

[0144] 아넬로벡터 유전 요소의 서열을 포함하는 유전 요소 영역,

[0145] 백본 영역

[0146] 을 상기 순서로 포함하는, 핵산 작제물.

[0147] 63. 상기 구현예들 중 어느 하나에 있어서, 외인성 이펙터를 인코딩하는 서열의 하나 이하의 카피를 포함하는, 핵산 작제물.

[0148] 64. 상기 구현예들 중 어느 하나에 있어서, 작제물은 백본 영역의 길이를 제외하고, 약 10 kb 미만(예를 들어, 약 9 kb, 8 kb, 7 kb, 6 kb, 5 kb, 4 kb, 또는 3 kb 미만)의 길이를 갖는, 핵산 작제물.

[0149] 65. 상기 구현예들 중 어느 하나에 있어서, 핵산 작제물은 유전 요소 영역의 하나 이하의 전장 카피를 포함하는, 핵산 작제물.

[0150] 66. 상기 구현예들 중 어느 하나에 있어서, 탠덤 영역은 2800, 2700, 2600, 2500, 2000, 1500, 1000, 900, 800, 700, 600, 또는 500개 미만 뉴클레오티드의 길이를 갖는, 핵산 작제물.

[0151] 67. 상기 구현예들 중 어느 하나에 있어서, 유전 요소는, 예를 들어 표 A1, B1, 또는 C1 중 임의의 것의 아넬로바이러스 유래의 서열(예를 들어, 5' UTR, 3' UTR, 또는 GC-풍부 영역 서열), 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 서열 동일성을 갖는 서열을 포함하는, 핵산 작제물.

[0152] 68. 구현예 67에 있어서, 아넬로바이러스는 인간 아넬로바이러스인, 핵산 작제물.

[0153] 69. 상기 구현예들 중 어느 하나에 있어서, 유전 요소는 인간 세포에서 복제될 수 있는, 핵산 작제물.

[0154] 70. 상기 구현예들 중 어느 하나에 있어서, 유전 요소는 (예를 들어, ORF1, ORF2, ORF2/2, ORF2/3, ORF1/1, 및 ORF1/2의 전부) 중 하나 이상을 인코딩하지 않는, 핵산 작제물.

[0155] 71. 상기 구현예들 중 어느 하나에 있어서, 핵산 작제물은 이중 가닥 DNA인, 핵산 작제물.

[0156] 72. 상기 구현예들 중 어느 하나에 있어서, 핵산 작제물은 단일 가닥 DNA인, 핵산 작제물.

[0157] 73. 상기 구현예들 중 어느 하나에 있어서, 핵산 작제물은 환형인, 핵산 작제물.

[0158] 74. 상기 구현예들 중 어느 하나에 있어서, 핵산 작제물은 플라스미드 또는 바이러스 벡터, 예를 들어, 바큘로바이러스 벡터인, 핵산 작제물.

[0159] 75. 상기 구현예들 중 어느 하나에 있어서, 백본 영역은 박테리아, 포유동물 세포, 또는 곤충 세포에서 핵산 작제물의 복제에 충분한, 핵산 작제물.

[0160] 76. 상기 구현예들 중 어느 하나에 있어서, 백본 영역은 5' UTR(예를 들어, 헤어핀 루프 및/또는 복제 원점(ORI)을 포함함) 및 선택 마커(예를 들어, 양성 선택 마커(예를 들어, 내성 마커), 음성 선택 마커, 또는 형광 마커) 중 하나 또는 둘 다를 포함하는, 핵산 작제물.

[0161] 77. 구현예 76에 있어서, ORI는 박테리아 ORI, 바이러스 ORI, 포유동물 ORI, 또는 곤충 ORI인, 핵산 작제물.

[0162] 78. 상기 구현예들 중 어느 하나에 있어서, 스페이서 서열은 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20개, 또는 그 이상의 아미노산의 길이, 또는 1 내지 5, 5 내지 10, 10 내지 15, 또는 15 내지 20개 아미노산의 길이를 갖는, 핵산 작제물.

[0163] 79. 상기 구현예들 중 어느 하나에 있어서, 제1 유전 요소 영역은 본 명세서에 기재된 아넬로바이러스 유래의

TATA 박스, 개시인자 요소, cap 부위, 전사 시작 부위, 5' UTR 보존된 도메인, ORF1-인코딩 서열, ORF1/1-인코딩 서열, ORF1/2-인코딩 서열, ORF2-인코딩 서열, ORF2/2-인코딩 서열, ORF2/3-인코딩 서열, ORF2/3t-인코딩 서열, 3개의 오픈-리딩 프레임 영역, 폴리(A) 신호, 및/또는 GC-풍부 영역(예를 들어, 표 A1 내지 M1 중 임의의 것에 열거된 바와 같음) 중 하나 이상, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 서열을 포함하는, 핵산 작제물.

[0164] 80. 상기 구현예들 중 어느 하나에 있어서, 제2 유전 요소 영역 또는 이의 단편, 또는 탠덤 영역은 본 명세서에 기재된 아넬로바이러스 유래의 TATA 박스, 개시인자 요소, cap 부위, 전사 시작 부위, 5' UTR 보존된 도메인, ORF1-인코딩 서열, ORF1/1-인코딩 서열, ORF1/2-인코딩 서열, ORF2-인코딩 서열, ORF2/2-인코딩 서열, ORF2/3-인코딩 서열, ORF2/3t-인코딩 서열, 3개의 오픈-리딩 프레임 영역, 폴리(A) 신호, 및/또는 GC-풍부 영역(예를 들어, 표 A1 내지 M1 중 임의의 것에 열거된 바와 같음) 중 하나 이상, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 서열을 포함하는, 핵산 작제물.

[0165] 81. 상기 구현예들 중 어느 하나에 있어서, 제1 유전 요소 영역은 아넬로바이러스 게놈 서열(예를 들어, 본 명세서에 기재된 바와 같음, 예를 들어, 표 A1 내지 M1 중 임의의 것에 열거된 바와 같음), 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 서열을 포함하는, 핵산 작제물.

[0166] 82. 구현예 81에 있어서, 아넬로바이러스 게놈 서열 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 서열의 적어도 하나의 추가의 카피(예를 들어, 총 2, 3, 4, 5, 또는 6개의 카피)를 추가로 포함하는, 핵산 작제물.

[0167] 83. 구현예 81에 있어서, 상이한 아넬로바이러스 게놈 서열(예를 들어, 본 명세서에 기재된 바와 같음, 예를 들어, 표 A1 내지 M1 중 임의의 것에 열거된 바와 같음), 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 서열의 적어도 하나의 카피(예를 들어, 총 1, 2, 3, 4, 5, 또는 6개의 카피)를 추가로 포함하는, 핵산 작제물.

[0168] 84. 상기 구현예들 중 어느 하나에 있어서, 제1 유전 요소 영역 및/또는 제2 유전 요소 영역 또는 이의 단편, 또는 탠덤 영역은 본 명세서에 기재된 아넬로바이러스 유래의 5' UTR 뉴클레오티드 서열(예를 들어, 표 A1 내지 M1 중 임의의 것에 열거된 바와 같음)에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 서열을 포함하는, 핵산 작제물.

[0169] 85. 상기 구현예들 중 어느 하나에 있어서, 제1 유전 요소 영역 및/또는 제2 유전 요소 영역 또는 이의 단편, 또는 탠덤 영역은 하기 핵산 서열의 적어도 10, 15, 20, 25, 30, 31, 32, 33, 34, 35, 또는 36개의 연속적인 뉴클레오티드를 포함하는, 핵산 작제물:

[0170] (i) CGCGCTGCGCGCGCCGCCCAGTAGGGGGAGCCATGC(서열번호 160),

[0171] (ii) GCGCTX$_1$CGCGCGCGCGCCGGGGGGCTGCGCCCCCCC(서열번호 164)(여기서, X$_1$은 T, G, 또는 A로부터 선택됨);

[0172] (iii) GCGCTTCGCGCGCCGCCCACTAGGGGGCGTTGCGCG(서열번호 165);

[0173] (iv) GCGCTGCGCGCGCCGCCCAGTAGGGGGCGCAATGCG(서열번호 166);

[0174] (v) GCGCTGCGCGCGCGGCCCCCGGGGGAGGCATTGCCT(서열번호 167);

[0175] (vi) GCGCTGCGCGCGCGCGCCGGGGGGGCGCCAGCGCCC(서열번호 168);

[0176] (vii) GCGCTTCGCGCGCGCGCCGGGGGGCTCCGCCCCCCC(서열번호 169);

[0177] (viii) GCGCTTCGCGCGCGCGCCGGGGGGCTGCGCCCCCCC(서열번호 170);

[0178] (ix) GCGCTACGCGCGCGCGCCGGGGGGCTGCGCCCCCCC(서열번호 171); 또는

[0179] (x) GCGCTACGCGCGCGCGCCGGGGGGCTCTGCCCCCCC(서열번호 172);

[0180] 또는 이에 대하여 적어도 75, 76, 77, 78, 79, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 또는 100% 서열 동일성을 갖는 핵산 서열.

[0181] 86. 상기 구현예들 중 어느 하나에 있어서, 제1 유전 요소 영역 및/또는 제2 유전 요소 영역 또는 이의 단편, 또는 탠덤 영역은 GC 함량이 적어도 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 또는 80.6%인

적어도 20, 25, 30, 31, 32, 33, 34, 35, 또는 36개의 연속적인 뉴클레오티드를 포함하는, 핵산 작제물.

[0182] 87. 핵산(예를 들어, DNA) 작제물로서,

[0183] a) (선택적으로 외인성 이펙터를 인코딩하는 서열을 포함하는) 돌연변이체 또는 야생형 아넬로바이러스 게놈;

[0184] b) (a)의 아넬로바이러스 게놈과 나란히 배치된, 아넬로바이러스 게놈의 단편; 및

[0185] c) 선택적으로, (a)와 (b) 사이에 위치하는 스페이서 서열

[0186] 을 포함하는 핵산 작제물.

[0187] 88. 핵산(예를 들어, DNA) 작제물로서,

[0188] a) (선택적으로 외인성 이펙터를 인코딩하는 서열을 포함하는) 아넬로바이러스 유전 요소 영역;

[0189] b) 아넬로바이러스 유전 요소 영역의 단편; 및

[0190] c) 선택적으로, (a)와 (b) 사이에 위치하는 스페이서 서열

[0191] 을 포함하는 핵산 작제물.

[0192] 89. 핵산(예를 들어, DNA) 작제물로서,

[0193] a) i) 제1 아넬로바이러스 상류 복제-촉진 서열(uRFS), 예를 들어, 5' UTR;

[0194] ii) 선택적으로, 외인성 이펙터를 인코딩하는 서열에 작동 가능하게 연결된 프로모터;

[0195] iii) 제1 아넬로바이러스 하류 복제-촉진 서열(dRFS), 예를 들어, 3' UTR

[0196] 을 포함하는 아넬로바이러스 유전 요소 영역; 및

[0197] b) 제2 아넬로바이러스 uRFS(예를 들어, 5' UTR) 또는 제2 아넬로바이러스 dRFS(예를 들어, 3' UTR); 및

[0198] c) 선택적으로, 아넬로바이러스 유전 요소 영역과 (b) 사이에 위치하는 스페이서 서열

[0199] 을 포함하고,

[0200] 아넬로바이러스 유전 요소 영역의 2개의 전장 카피를 포함하지 않는 핵산 작제물.

[0201] 90. 핵산(예를 들어, DNA) 작제물로서,

[0202] a) i) 제1 아넬로바이러스 상류 복제-촉진 서열(uRFS), 예를 들어, 5' UTR;

[0203] ii) 선택적으로, 외인성 이펙터를 인코딩하는 서열에 작동 가능하게 연결된 프로모터;

[0204] iii) 제1 아넬로바이러스 하류 복제-촉진 서열(dRFS), 예를 들어, 3' UTR

[0205] 을 포함하는 아넬로바이러스 유전 요소 영역; 및

[0206] b) 제2 아넬로바이러스 uRFS(예를 들어, 5' UTR) 또는 제2 아넬로바이러스 dRFS(예를 들어, 3' UTR)(여기서, 제2 아넬로바이러스 uRFS는 아넬로바이러스 유전 요소 영역의 전장 카피의 일부가 아님); 및

[0207] c) 선택적으로, 아넬로바이러스 유전 요소 영역과 (b) 사이에 위치하는 스페이서 서열

[0208] 을 포함하는 핵산 작제물.

[0209] 91. 상기 구현예들 중 어느 하나의 핵산 작제물을 포함하는 세포.

[0210] 92. 구현예 91에 있어서, 박테리아 세포, 포유동물 세포 또는 곤충 세포인, 세포.

[0211] 93. 상기 구현예들 중 어느 하나의 핵산 작제물 및 세포를 포함하는 반응 혼합물.

[0212] 94. 세포 및 핵산(예를 들어, DNA) 작제물을 포함하는 반응 혼합물로서, 작제물은

[0213] a) 외인성 이펙터를 인코딩하는 서열을 포함하는, 선택적으로 돌연변이체인 제1 아넬로바이러스 게놈;

[0214] b) 제1 아넬로바이러스 게놈과 나란히 배치된, 제2 아넬로바이러스 게놈 또는 이의 단편; 및

[0215] c) 선택적으로, (a)와 (b) 사이에 위치하는 스페이서 서열

[0216] 을 포함하는 반응 혼합물.

[0217] 95. 세포 및 핵산(예를 들어, DNA) 작제물을 포함하는 반응 혼합물로서, 작제물은

[0218] a) 외인성 이펙터를 인코딩하는 서열을 포함하는 제1 아넬로바이러스 유전 요소 영역;

[0219] b) 제2 아넬로바이러스 유전 요소 영역 또는 이의 단편; 및

[0220] c) 선택적으로, (a) 및 (b) 사이에 위치하는 스페이서 서열

[0221] 을 포함하는 반응 혼합물.

[0222] 96. 세포 및 핵산(예를 들어, DNA) 작제물을 포함하는 반응 혼합물로서, 작제물은

[0223] a) i) 제1 아넬로바이러스 상류 복제-촉진 서열(uRFS), 예를 들어, 5' UTR;

[0224] ii) 외인성 이펙터를 인코딩하는 서열에 작동 가능하게 연결된 프로모터;

[0225] iii) 제1 아넬로바이러스 하류 복제-촉진 서열(dRFS), 예를 들어, 3' UTR

[0226] 을 포함하는 아넬로바이러스 유전 요소 영역; 및

[0227] b) 제2 아넬로바이러스 uRFS(예를 들어, 5' UTR) 또는 제2 아넬로바이러스 dRFS(예를 들어, 3' UTR); 및

[0228] c) 선택적으로, 아넬로바이러스 유전 요소 영역과 (b) 사이에 위치하는 스페이서 서열

[0229] 을 포함하는 반응 혼합물.

[0230] 97. 구현예 93 내지 96 중 어느 하나에 있어서, 세포는 박테리아 세포, 포유동물 세포 또는 곤충 세포인, 반응 혼합물.

[0231] 98. 구현예 93 내지 97 중 어느 하나에 있어서, 세포는 핵산 작제물을 포함하는, 반응 혼합물.

[0232] 99. 상기 구현예들 중 어느 하나의 핵산 작제물을 포함하는 조성물의 제조 방법으로서,

[0233] a) 상기 구현예들 중 어느 하나의 핵산 작제물을 제공하는 단계; 및

[0234] b) 세포(예를 들어, 박테리아 세포)를 핵산 작제물이 세포에서 복제될 수 있도록 하는 조건 하에서 핵산 작제물과 접촉시켜, 복수 카피의 핵산 작제물을 생성하는 단계

[0235] 를 포함하여, 이에 의해 핵산 작제물을 포함하는 조성물을 제조하는 방법.

[0236] 100. 구현예 99에 있어서,

[0237] c) 세포로부터 복수 카피의 핵산 작제물을 단리하는 단계

[0238] 를 추가로 포함하는, 방법.

[0239] 101. 아넬로벡터 유전 요소의 제조 방법으로서,

[0240] a) 상기 구현예들 중 어느 하나의 핵산 작제물을 제공하는 단계; 및

[0241] b) 세포(예를 들어, 포유동물 숙주 세포)를 핵산 작제물의 아넬로바이러스 유전 요소가 복제되거나 증폭될 수 있도록 하는 조건 하에서 핵산 작제물과 접촉시키는 단계

[0242] 를 포함하여, 이에 의해 아넬로벡터 유전 요소를 제조하는 방법.

[0243] 102. 구현예 101에 있어서,

[0244] c) 증폭된 아넬로바이러스 유전 요소가 세포의 단백질성 외부에 봉입될 수 있도록 하는 조건 하에서 세포를 인큐베이션하는 단계

[0245] 를 추가로 포함하는, 방법.

[0246] 103. 단백질성 외부에 봉입된 유전 요소를 포함하는 아넬로벡터의 제조 방법으로서,

[0247] a) 상기 구현예들 중 어느 하나의 핵산 작제물 및 아넬로바이러스 유전 요소(예를 들어, 여기서 아넬로바이러스 유전 요소는 핵산 작제물로부터 증폭된 것임)의 하나 이상의 카피를 포함하는 세포(예를 들어, 포유동물 숙주

세포)를 제공하는 단계;

[0248] b) 아넬로바이러스 유전 요소가 세포의 단백질성 외부에 봉입될 수 있도록 하는 조건 하에서 세포를 인큐베이션하는 단계

[0249] 를 포함하여, 이에 의해 아넬로벡터를 제조하는 방법.

[0250] 104. 구현예 102 또는 103에 있어서, 단백질성 외부는 시스 또는 트랜스로 제공되는, 방법.

[0251] 105. 구현예 102 내지 104 중 어느 하나에 있어서, 단백질성 외부에 봉입된 아넬로바이러스 유전 요소는 감염성 입자, 예를 들어, 바이러스 입자를 형성하는, 방법.

[0252] 106. 단백질성 외부에 봉입된 유전 요소를 포함하는 아넬로벡터의 제조 방법으로서,

[0253] a) 아넬로바이러스 유전 요소를 포함하는 MOLT-4 세포를 제공하는 단계;

[0254] b) 아넬로바이러스 유전 요소가 세포의 단백질성 외부에 봉입될 수 있도록 하는 조건 하에서 세포를 인큐베이션하는 단계;

[0255] 를 포함하여, 이에 의해 아넬로벡터를 제조하는 방법.

[0256] 107. 복수의 핵산(예를 들어, DNA) 작제물을 포함하는 조성물로서, 핵산 작제물 각각은

[0257] a) 외인성 이펙터를 인코딩하는 서열을 포함하는, 선택적으로 돌연변이체인, 제1 아넬로바이러스 게놈;

[0258] b) 제1 아넬로바이러스 게놈과 나란히 배치된, 제2 아넬로바이러스 게놈 또는 이의 단편; 및

[0259] c) 선택적으로, (a)와 (b) 사이에 위치하는 스페이서 서열

[0260] 을 포함하고,

[0261] 조성물은

[0262] i) 상기 구현예들 중 어느 하나의 핵산 작제물을 제공하는 단계; 및

[0263] ii) 핵산 작제물이 세포에서 복제될 수 있도록 하는 조건 하에서 세포(예를 들어, 박테리아 세포)를 핵산 작제물과 접촉시키는 단계

[0264] 를 포함하는 방법에 의해 생성되는 조성물.

[0265] 108. 복수의 핵산(예를 들어, DNA) 작제물을 포함하는 조성물로서, 핵산 작제물 각각은

[0266] a) 외인성 이펙터를 인코딩하는 서열을 포함하는 제1 아넬로바이러스 유전 요소 영역;

[0267] b) 제2 아넬로바이러스 유전 요소 영역 또는 이의 단편; 및

[0268] c) 선택적으로, (a)와 (b) 사이에 위치하는 스페이서 서열

[0269] 을 포함하고,

[0270] 조성물은

[0271] i) 상기 구현예들 중 어느 하나의 핵산 작제물을 제공하는 단계; 및

[0272] ii) 핵산 작제물이 세포에서 복제될 수 있도록 하는 조건 하에서 세포(예를 들어, 박테리아 세포)를 핵산 작제물과 접촉시키는 단계

[0273] 를 포함하는 방법에 의해 생성되는 조성물.

[0274] 109. 복수의 핵산(예를 들어, DNA)을 포함하는 조성물로서, 핵산 작제물 각각은

[0275] a) i) 제1 아넬로바이러스 상류 복제-촉진 서열(uRFS), 예를 들어, 5' UTR;

[0276] ii) 외인성 이펙터를 인코딩하는 서열에 작동 가능하게 연결된 프로모터;

[0277] iii) 제1 아넬로바이러스 하류 복제-촉진 서열(dRFS), 예를 들어, 3' UTR

[0278] 을 포함하는 아넬로바이러스 유전 요소 영역; 및

[0279] b) 제2 아넬로바이러스 uRFS(예를 들어, 5' UTR) 또는 제2 아넬로바이러스 dRFS(예를 들어, 3' UTR); 및

[0280] c) 선택적으로, 아넬로바이러스 유전 요소 영역과 (b) 사이에 위치하는 스페이서 서열

[0281] 을 포함하고,

[0282] 조성물은

[0283] i) 상기 구현예들 중 어느 하나의 핵산 작제물을 제공하는 단계; 및

[0284] ii) 핵산 작제물이 세포에서 복제될 수 있도록 하는 조건 하에서 세포(예를 들어, 박테리아 세포)를 핵산 작제물과 접촉시키는 단계

[0285] 를 포함하는 방법에 의해 생성되는 조성물.

[0286] 110. 외인성 이펙터를 세포로 운반하는 방법으로서,

[0287] a) 외인성 이펙터를 인코딩하는 서열을 포함하는, 선택적으로 돌연변이체인, 제1 아넬로바이러스 게놈;

[0288] b) 제1 아넬로바이러스 게놈과 나란히 배치된, 제2 아넬로바이러스 게놈 또는 이의 단편; 및

[0289] c) 선택적으로, (a)와 (b) 사이에 위치하는 스페이서 서열

[0290] 을 포함하는 핵산 작제물을 세포 내로 도입하는 단계; 및

[0291] 외인성 이펙터의 발현에 적합한 조건 하에서 세포를 인큐베이션하는 단계

[0292] 를 포함하고,

[0293] 핵산 작제물은

[0294] i) 상기 구현예들 중 어느 하나의 핵산 작제물을 제공하는 단계; 및

[0295] ii) 핵산 작제물이 세포에서 복제될 수 있도록 하는 조건 하에서 세포(예를 들어, 박테리아 세포)를 핵산 작제물과 접촉시키는 단계

[0296] 를 포함하는 방법에 의해 생성되는 방법.

[0297] 111. 외인성 이펙터를 세포로 운반하는 방법으로서,

[0298] a) 외인성 이펙터를 인코딩하는 서열을 포함하는 제1 아넬로바이러스 유전 요소 영역;

[0299] b) 제2 아넬로바이러스 유전 요소 영역 또는 이의 단편; 및

[0300] c) 선택적으로, (a)와 (b) 사이에 위치하는 스페이서 서열

[0301] 을 포함하는 핵산 작제물을 세포 내로 도입하는 단계; 및

[0302] 외인성 이펙터의 발현에 적합한 조건 하에서 세포를 인큐베이션하는 단계

[0303] 를 포함하고,

[0304] 조성물은

[0305] i) 상기 구현예들 중 어느 하나의 핵산 작제물을 제공하는 단계; 및

[0306] ii) 핵산 작제물이 세포에서 복제될 수 있도록 하는 조건 하에서 세포(예를 들어, 박테리아 세포)를 핵산 작제물과 접촉시키는 단계

[0307] 를 포함하는 방법에 의해 생성되는 방법.

[0308] 112. 외인성 이펙터를 세포로 운반하는 방법으로서,

[0309] a) i) 제1 아넬로바이러스 상류 복제-촉진 서열(uRFS), 예를 들어, 5' UTR;

[0310] ii) 외인성 이펙터를 인코딩하는 서열에 작동 가능하게 연결된 프로모터;

[0311] iii) 제1 아넬로바이러스 하류 복제-촉진 서열(dRFS), 예를 들어, 3' UTR

[0312] 을 포함하는 아넬로바이러스 유전 요소 영역; 및

[0313] b) 제2 아넬로바이러스 uRFS(예를 들어, 5' UTR) 또는 제2 아넬로바이러스 dRFS(예를 들어, 3' UTR); 및

[0314] c) 선택적으로, 아넬로바이러스 유전 요소 영역과 (b) 사이에 위치하는 스페이서 서열

[0315] 을 포함하는 핵산 작제물을 세포 내로 도입하는 단계; 및

[0316] 외인성 이펙터의 발현에 적합한 조건 하에서 세포를 인큐베이션하는 단계

[0317] 를 포함하고,

[0318] 조성물은

[0319] i) 상기 구현예들 중 어느 하나의 핵산 작제물을 제공하는 단계; 및

[0320] ii) 핵산 작제물이 세포에서 복제될 수 있도록 하는 조건 하에서 세포(예를 들어, 박테리아 세포)를 핵산 작제물과 접촉시키는 단계

[0321] 를 포함하는 방법에 의해 생성되는 방법.

[0322] 113. 아넬로바이러스 유전 요소 영역을 세포의 게놈 내로 통합시키는 방법으로서,

[0323] (a) 세포를 상기 구현예들 중 어느 하나의 핵산 작제물과 접촉시키되,

[0324] 핵산 작제물은 5' 상동성 영역 및 3' 상동성 영역이 측접한 아넬로바이러스 유전 요소 영역을 포함하고,

[0325] 5' 상동성 영역 및 3' 상동성 영역은 세포의 게놈에서 적어도 9개 뉴클레오티드의 핵산 서열에 대하여 적어도 90% 서열 동일성(예를 들어, 적어도 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성)을 갖는, 단계; 및

[0326] (b) 세포의 게놈 내로 아넬로바이러스 유전 요소 영역의 통합에 적합한 조건 하에서 세포를 인큐베이션하는 단계

[0327] 를 포함하는 방법.

[0328] 114. 핵산(예를 들어, DNA) 작제물로서,

[0329] a) 선택적으로 야생형 아넬로바이러스 게놈 서열에 대해 돌연변이를 포함하는, 제1 아넬로바이러스 게놈,

[0330] b) 제1 아넬로바이러스 게놈과 나란히 배치되고, 선택적으로 야생형 아넬로바이러스 게놈 서열에 대해 돌연변이를 포함하는, 제2 아넬로바이러스 게놈; 및

[0331] c) 선택적으로, (a)와 (b) 사이에 위치하는 스페이서 서열

[0332] 을 포함하고,

[0333] 제1 또는 제2 아넬로바이러스 게놈은 외인성 이펙터를 인코딩하는 서열을 포함하는 핵산 작제물.

[0334] 115. 구현예 114에 있어서, 제1 아넬로바이러스 게놈은 아넬로바이러스 ORF1, ORF2, 및/또는 ORF2/3, 및/또는 GC-풍부 영역을 인코딩하는 하나 이상의 서열을 포함하는, 핵산 작제물.

[0335] 116. 구현예 114에 있어서, 제2 아넬로바이러스 게놈은 ORF1, ORF2, 및/또는 ORF2/3, 및/또는 GC-풍부 영역을 인코딩하는 하나 이상의 서열을 포함하는, 핵산 작제물.

[0336] 117. 핵산(예를 들어, DNA) 작제물로서, 5'에서 3' 순서로

[0337] a) 외인성 이펙터를 인코딩하는 서열을 포함하되, 선택적으로 야생형 아넬로바이러스 게놈 서열에 대해 돌연변이를 포함하는, 제1 아넬로바이러스 게놈,

[0338] b) 선택적으로, 스페이서 서열, 및

[0339] c) 제1 아넬로바이러스 게놈과 나란히 배치되고, 선택적으로 야생형 아넬로바이러스 게놈 서열에 대해 돌연변이를 포함하는, 제2 아넬로바이러스 게놈

[0340] 을 포함하고,

[0341] 선택적으로 외인성 이펙터를 포함하는 서열은 길이가 약 50 내지 100, 100 내지 200, 200 내지 300, 300 내지

400, 400 내지 500, 500 내지 600, 600 내지 700, 700 내지 800, 800 내지 900, 900 내지 950, 950 내지 1000, 1000 내지 1100, 1100 내지 1200, 1200 내지 1400, 1400 내지 1600, 1600 내지 1800, 1800 내지 2000, 2000 내지 2200, 또는 2200 내지 2400개 뉴클레오티드(예를 들어, 길이가 약 986개 뉴클레오티드)인 핵산 작제물.

[0342] 118. 구현예 117에 있어서, 제1 아넬로바이러스 게놈은

[0343] (i) 아넬로바이러스 ORF2 유전자를 포함하지 않고/않거나,

[0344] (ii) 아넬로바이러스 ORF2/3 유전자의 제1 엑손을 포함하지 않고/않거나,

[0345] (iii) 절단된 아넬로바이러스 ORF1 유전자를 포함하는,

[0346] 핵산 작제물.

[0347] 119. 구현예 117에 있어서, 제1 아넬로바이러스 게놈은

[0348] (i) 절단된 아넬로바이러스 ORF2 유전자를 포함하고/하거나,

[0349] (ii) 아넬로바이러스 ORF2/3 유전자의 절단된 제1 엑손을 포함하고/하거나,

[0350] (iii) 절단된 아넬로바이러스 ORF1 유전자를 포함하는,

[0351] 핵산 작제물.

[0352] 120. 구현예 118 또는 119에 있어서, 절단은 아넬로바이러스 ORF1 유전자의 5' 말단에서 발생하는(예를 들어, 아넬로바이러스 ORF1 유전자의 약 5 내지 10, 10 내지 20, 20 내지 30, 30 내지 40, 40 내지 50, 50 내지 60, 60 내지 70, 70 내지 80, 80 내지 90, 90 내지 100, 100 내지 150, 150 내지 200, 200 내지 300, 300 내지 400, 400 내지 500, 500 내지 600, 600 내지 700, 700 내지 800, 800 내지 900, 900 내지 1000, 1000 내지 1200, 1200 내지 1400, 1400 내지 1600, 1600 내지 1800, 또는 1800 내지 2000개가 절단되는), 핵산 작제물.

[0353] 121. 구현예 118 또는 119에 있어서, 절단은 아넬로바이러스 ORF1 유전자 내에서 발생하는(예를 들어, 아넬로바이러스 ORF1 유전자의 약 5 내지 10, 10 내지 20, 20 내지 30, 30 내지 40, 40 내지 50, 50 내지 60, 60 내지 70, 70 내지 80, 80 내지 90, 90 내지 100, 100 내지 150, 150 내지 200, 200 내지 300, 300 내지 400, 400 내지 500, 500 내지 600, 600 내지 700, 700 내지 800, 800 내지 900, 900 내지 1000, 1000 내지 1200, 1200 내지 1400, 1400 내지 1600, 1600 내지 1800, 또는 1800 내지 2000개가 절단되는), 핵산 작제물.

[0354] 122. 구현예 117에 있어서, 제1 아넬로바이러스 게놈은

[0355] (i) 아넬로바이러스 ORF2/3 유전자의 제2 엑손을 포함하지 않거나 아넬로바이러스 ORF2/3 유전자의 절단된 제2 엑손을 포함하고/하거나,

[0356] (ii) GC-풍부 영역을 포함하지 않거나 절단된 GC-풍부 영역을 포함하고/하거나,

[0357] (iii) 절단된 아넬로바이러스 ORF1 유전자를 포함하는,

[0358] 핵산 작제물.

[0359] 123. 구현예 122에 있어서, 절단은 아넬로바이러스 ORF1 유전자 내에서 발생하는(예를 들어, 아넬로바이러스 ORF1 유전자의 약 5 내지 10, 10 내지 20, 20 내지 30, 30 내지 40, 40 내지 50, 50 내지 60, 60 내지 70, 70 내지 80, 80 내지 90, 90 내지 100, 100 내지 150, 150 내지 200, 200 내지 300, 300 내지 400, 400 내지 500, 500 내지 600, 600 내지 700, 700 내지 800, 800 내지 900, 900 내지 1000, 1000 내지 1200, 1200 내지 1400, 1400 내지 1600, 1600 내지 1800, 또는 1800 내지 2000개가 절단되는), 핵산 작제물.

[0360] 124. 구현예 122에 있어서, 절단은 아넬로바이러스 ORF1 유전자의 3' 말단에서 발생하는(예를 들어, 아넬로바이러스 ORF1 유전자의 약 5 내지 10, 10 내지 20, 20 내지 30, 30 내지 40, 40 내지 50, 50 내지 60, 60 내지 70, 70 내지 80, 80 내지 90, 90 내지 100, 100 내지 150, 150 내지 200, 200 내지 300, 300 내지 400, 400 내지 500, 500 내지 600, 600 내지 700, 700 내지 800, 800 내지 900, 900 내지 1000, 1000 내지 1200, 1200 내지 1400, 1400 내지 1600, 1600 내지 1800, 또는 1800 내지 2000개가 절단되는), 핵산 작제물.

[0361] 125. 구현예 117 내지 124 중 어느 하나에 있어서, 제1 아넬로바이러스 게놈의 아넬로바이러스 ORF1, ORF2, 및/또는 ORF2/3 유전자 중 하나 이상(예를 들어, 1, 2, 또는 3개)은 기능적 단백질을 인코딩하지 않는(예를 들어, 아넬로바이러스 ORF1, ORF2, 및/또는 ORF2/3 유전자의 하나 이상(예를 들어, 1, 2, 또는 3개)은 비활성 돌연변

이, 예를 들어, 조기성숙 정지 코돈 돌연변이, 프레임시프트 돌연변이, 또는 개시 코돈을 변경하거나 결실시키는 돌연변이를 포함하는), 핵산 작제물.

[0362] 126. 구현예 117 내지 125 중 어느 하나에 있어서, 제2 아넬로바이러스 게놈은 외인성 이펙터를 인코딩하는 서열을 추가로 포함하고;

[0363] 선택적으로 외인성 이펙터를 포함하는 서열은 길이가 약 50 내지 100, 100 내지 200, 200 내지 300, 300 내지 400, 400 내지 500, 500 내지 600, 600 내지 700, 700 내지 800, 800 내지 900, 900 내지 950, 950 내지 1000, 1000 내지 1100, 1100 내지 1200, 1200 내지 1400, 1400 내지 1600, 1600 내지 1800, 1800 내지 2000, 2000 내지 2200, 또는 2200 내지 2400개 뉴클레오티드(예를 들어, 길이가 약 986개 뉴클레오티드)인, 핵산 작제물.

[0364] 127. 핵산(예를 들어, DNA) 작제물로서, 5'에서 3' 순서로

[0365] a) 선택적으로 야생형 아넬로바이러스 게놈 서열에 대해 돌연변이를 포함하는, 제1 아넬로바이러스 게놈,

[0366] b) 선택적으로, 스페이서 서열, 및

[0367] c) 제1 아넬로바이러스 게놈과 나란히 배치되고, 외인성 이펙터를 인코딩하는 서열을 포함하고, 선택적으로 야생형 아넬로바이러스 게놈 서열에 대해 돌연변이를 포함하는, 제2 아넬로바이러스 게놈

[0368] 을 포함하고,

[0369] 선택적으로 외인성 이펙터를 포함하는 서열은 길이가 약 50 내지 100, 100 내지 200, 200 내지 300, 300 내지 400, 400 내지 500, 500 내지 600, 600 내지 700, 700 내지 800, 800 내지 900, 900 내지 950, 950 내지 1000, 1000 내지 1100, 1100 내지 1200, 1200 내지 1400, 1400 내지 1600, 1600 내지 1800, 1800 내지 2000, 2000 내지 2200, 또는 2200 내지 2400개 뉴클레오티드(예를 들어, 길이가 약 986개 뉴클레오티드)인 핵산 작제물.

[0370] 128. 구현예 127에 있어서, 제2 아넬로바이러스 게놈은

[0371] (i) 아넬로바이러스 ORF2 유전자를 포함하지 않고/않거나,

[0372] (ii) 아넬로바이러스 ORF2/3 유전자의 제1 엑손을 포함하지 않고/않거나,

[0373] (iii) 절단된 아넬로바이러스 ORF1 유전자를 포함하는,

[0374] 핵산 작제물.

[0375] 129. 구현예 127에 있어서, 제2 아넬로바이러스 게놈은

[0376] (i) 절단된 아넬로바이러스 ORF2 유전자를 포함하고/하거나,

[0377] (ii) 아넬로바이러스 ORF2/3 유전자의 절단된 제1 엑손을 포함하고/하거나,

[0378] (iii) 절단된 아넬로바이러스 ORF1 유전자를 포함하는,

[0379] 핵산 작제물.

[0380] 130. 구현예 128 또는 129에 있어서, 절단은 아넬로바이러스 ORF1 유전자의 5' 말단에서 발생하는(예를 들어, 아넬로바이러스 ORF1 유전자의 약 5 내지 10, 10 내지 20, 20 내지 30, 30 내지 40, 40 내지 50, 50 내지 60, 60 내지 70, 70 내지 80, 80 내지 90, 90 내지 100, 100 내지 150, 150 내지 200, 200 내지 300, 300 내지 400, 400 내지 500, 500 내지 600, 600 내지 700, 700 내지 800, 800 내지 900, 900 내지 1000, 1000 내지 1200, 1200 내지 1400, 1400 내지 1600, 1600 내지 1800, 또는 1800 내지 2000개가 절단되는), 핵산 작제물.

[0381] 131. 구현예 128 또는 129에 있어서, 절단은 아넬로바이러스 ORF1 유전자 내에서 발생하는(예를 들어, 아넬로바이러스 ORF1 유전자의 약 5 내지 10, 10 내지 20, 20 내지 30, 30 내지 40, 40 내지 50, 50 내지 60, 60 내지 70, 70 내지 80, 80 내지 90, 90 내지 100, 100 내지 150, 150 내지 200, 200 내지 300, 300 내지 400, 400 내지 500, 500 내지 600, 600 내지 700, 700 내지 800, 800 내지 900, 900 내지 1000, 1000 내지 1200, 1200 내지 1400, 1400 내지 1600, 1600 내지 1800, 또는 1800 내지 2000개가 절단되는), 핵산 작제물.

[0382] 132. 구현예 127에 있어서, 제2 아넬로바이러스 게놈은

[0383] (i) 아넬로바이러스 ORF2/3 유전자의 제2 엑손을 포함하지 않거나 아넬로바이러스 ORF2/3 유전자의 절단된 제2 엑손을 포함하고/하거나,

[0384] (ii) GC-풍부 영역을 포함하지 않거나 절단된 GC-풍부 영역을 포함하고/하거나,

[0385] (iii) 절단된 아넬로바이러스 ORF1 유전자를 포함하는,

[0386] 핵산 작제물.

[0387] 133. 구현예 132에 있어서, 절단은 아넬로바이러스 ORF1 유전자 내에서 발생하는(예를 들어, 아넬로바이러스 ORF1 유전자의 약 5 내지 10, 10 내지 20, 20 내지 30, 30 내지 40, 40 내지 50, 50 내지 60, 60 내지 70, 70 내지 80, 80 내지 90, 90 내지 100, 100 내지 150, 150 내지 200, 200 내지 300, 300 내지 400, 400 내지 500, 500 내지 600, 600 내지 700, 700 내지 800, 800 내지 900, 900 내지 1000, 1000 내지 1200, 1200 내지 1400, 1400 내지 1600, 1600 내지 1800, 또는 1800 내지 2000개가 절단되는), 핵산 작제물.

[0388] 134. 구현예 132에 있어서, 절단은 아넬로바이러스 ORF1 유전자의 3' 말단에서 발생하는(예를 들어, 아넬로바이러스 ORF1 유전자의 약 5 내지 10, 10 내지 20, 20 내지 30, 30 내지 40, 40 내지 50, 50 내지 60, 60 내지 70, 70 내지 80, 80 내지 90, 90 내지 100, 100 내지 150, 150 내지 200, 200 내지 300, 300 내지 400, 400 내지 500, 500 내지 600, 600 내지 700, 700 내지 800, 800 내지 900, 900 내지 1000, 1000 내지 1200, 1200 내지 1400, 1400 내지 1600, 1600 내지 1800, 또는 1800 내지 2000개가 절단되는), 핵산 작제물.

[0389] 135. 구현예 127 내지 134 중 어느 하나에 있어서, 제2 아넬로바이러스 게놈의 아넬로바이러스 ORF1, ORF2, 및/또는 ORF2/3 유전자 중 하나 이상(예를 들어, 1, 2, 또는 3개)은 기능적 단백질을 인코딩하지 않는(예를 들어, 아넬로바이러스 ORF1, ORF2, 및/또는 ORF2/3 유전자의 하나 이상(예를 들어, 1, 2, 또는 3개)은 비활성 돌연변이, 예를 들어, 조기성숙 정지 코돈 돌연변이, 프레임시프트 돌연변이, 또는 개시 코돈을 변경하거나 결실시키는 돌연변이를 포함하는), 핵산 작제물.

[0390] 136. 구현예 127 내지 135 중 어느 하나에 있어서, 제1 아넬로바이러스 게놈은 외인성 이펙터를 인코딩하는 서열을 추가로 포함하고;

[0391] 선택적으로 외인성 이펙터를 포함하는 서열은 길이가 약 50 내지 100, 100 내지 200, 200 내지 300, 300 내지 400, 400 내지 500, 500 내지 600, 600 내지 700, 700 내지 800, 800 내지 900, 900 내지 950, 950 내지 1000, 1000 내지 1100, 1100 내지 1200, 1200 내지 1400, 1400 내지 1600, 1600 내지 1800, 1800 내지 2000, 2000 내지 2200, 또는 2200 내지 2400개 뉴클레오티드(예를 들어, 길이가 약 986개 뉴클레오티드)인, 핵산 작제물.

[0392] 137. 구현예 117 내지 136 중 어느 하나에 있어서, 예를 들어, 외인성 이펙터를 인코딩하는 서열에, 프로모터(예를 들어, SV40 프로모터), 코작(Kozak) 서열, 및/또는 폴리-A 서열(예를 들어, SV40 폴리-A 서열) 중 하나 이상(예를 들어, 1, 2, 또는 3개)을 추가로 포함하는, 핵산 작제물.

[0393] 138. 구현예 117 내지 137 중 어느 하나에 있어서, 박테리아 복제 원점을 추가로 포함하는, 핵산 작제물.

[0394] 139. 구현예 117 내지 138 중 어느 하나에 있어서, 선택 가능한 마커(예를 들어, 내성 유전자, 예를 들어, 항생제 내성 유전자, 예를 들어, 스펙티노마이신 내성 유전자)를 추가로 포함하는, 핵산 작제물.

[0395] 140. 구현예 114 내지 139 중 어느 하나에 있어서, 제1 아넬로바이러스 게놈은 야생형 아넬로바이러스의 게놈 서열(예를 들어, 본 명세서에 기재된 바와 같음)에 대하여, 또는 적어도 50, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000, 2100, 2200, 2300, 2400, 2500, 2600, 2700, 2800, 2900, 또는 3000개 뉴클레오티드의 길이를 갖는 이의 연속 부분(예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어, 야생형 아넬로바이러스 게놈의 요소)에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 핵산 서열을 포함하는, 핵산 작제물.

[0396] 141. 구현예 114 내지 139 중 어느 하나에 있어서, 제1 아넬로바이러스 게놈은 야생형 아넬로바이러스의 게놈 서열(예를 들어, 본 명세서에 기재된 바와 같음)에 대하여, 또는 약 50 내지 100, 100 내지 200, 200 내지 300, 300 내지 400, 400 내지 500, 500 내지 600, 600 내지 700, 700 내지 800, 800 내지 900, 900 내지 1000, 1000 내지 1100, 1100 내지 1200, 1200 내지 1300, 1300 내지 1400, 1400 내지 1500, 1500 내지 1600, 1600 내지 1700, 1700 내지 1800, 1800 내지 1900, 1900 내지 2000, 2000 내지 2100, 2100 내지 2200, 2200 내지 2300, 2300 내지 2400, 2400 내지 2500, 2500 내지 2600, 2600 내지 2700, 2700 내지 2800, 2800 내지 2900, 또는 2900 내지 3000개 뉴클레오티드의 길이를 갖는 이의 연속 부분(예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어, 야생형 아넬로바이러스 게놈의 요소)에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 핵산 서열을 포함하는, 핵산 작제물.

[0397] 142. 구현예 114 내지 140 중 어느 하나에 있어서, 제2 아넬로바이러스 게놈은 야생형 아넬로바이러스의 게놈 서열(예를 들어, 본 명세서에 기재된 바와 같음)에 대하여, 또는 적어도 50, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000, 2100, 2200, 2300, 2400, 2500, 2600, 2700, 2800, 2900, 또는 3000개 뉴클레오티드의 길이를 갖는 이의 연속 부분(예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어, 야생형 아넬로바이러스 게놈의 요소)에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 핵산 서열을 포함하는, 핵산 작제물.

[0398] 143. 구현예 114 내지 141 중 어느 하나에 있어서, 제2 아넬로바이러스 게놈은 야생형 아넬로바이러스의 게놈 서열(예를 들어, 본 명세서에 기재된 바와 같음)에 대하여, 또는 약 50 내지 100, 100 내지 200, 200 내지 300, 300 내지 400, 400 내지 500, 500 내지 600, 600 내지 700, 700 내지 800, 800 내지 900, 900 내지 1000, 1000 내지 1100, 1100 내지 1200, 1200 내지 1300, 1300 내지 1400, 1400 내지 1500, 1500 내지 1600, 1600 내지 1700, 1700 내지 1800, 1800 내지 1900, 1900 내지 2000, 2000 내지 2100, 2100 내지 2200, 2200 내지 2300, 2300 내지 2400, 2400 내지 2500, 2500 내지 2600, 2600 내지 2700, 2700 내지 2800, 2800 내지 2900, 또는 2900 내지 3000개 뉴클레오티드의 길이를 갖는 이의 연속 부분(예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어, 야생형 아넬로바이러스 게놈의 요소)에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 핵산 서열을 포함하는, 핵산 작제물.

[0399] 144. 구현예 114 내지 143 중 어느 하나에 있어서,

[0400] 제1 아넬로바이러스 게놈은 외인성 이펙터를 인코딩하고,

[0401] 제2 아넬로바이러스 게놈은 아넬로바이러스 게놈(예를 들어, Ring2)의 뉴클레오티드 1 내지 2812, 또는 이에 대하여 적어도 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 서열을 포함하는 절단된 게놈이고;

[0402] 제2 아넬로바이러스 게놈은 제1 아넬로바이러스 게놈의 3'에 있는, 핵산 작제물.

[0403] 145. 구현예 114 내지 143 중 어느 하나에 있어서,

[0404] 제1 아넬로바이러스 게놈은 외인성 이펙터를 인코딩하고,

[0405] 제2 아넬로바이러스 게놈은 아넬로바이러스 게놈(예를 들어, Ring2)의 뉴클레오티드 1 내지 2583, 또는 이에 대하여 적어도 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 서열을 포함하는 절단된 게놈이고;

[0406] 제2 아넬로바이러스 게놈은 제1 아넬로바이러스 게놈의 3'에 있는, 핵산 작제물.

[0407] 146. 구현예 114 내지 143 중 어느 하나에 있어서,

[0408] 제1 아넬로바이러스 게놈은 외인성 이펙터를 인코딩하고,

[0409] 제2 아넬로바이러스 게놈은 아넬로바이러스 게놈(예를 들어, Ring2)의 뉴클레오티드 1 내지 2264, 또는 이에 대하여 적어도 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 서열을 포함하는 절단된 게놈이고;

[0410] 제2 아넬로바이러스 게놈은 제1 아넬로바이러스 게놈의 3'에 있는, 핵산 작제물.

[0411] 147. 구현예 114 내지 143 중 어느 하나에 있어서,

[0412] 제1 아넬로바이러스 게놈은 외인성 이펙터를 인코딩하고,

[0413] 제2 아넬로바이러스 게놈은 아넬로바이러스 게놈(예를 들어, Ring2)의 뉴클레오티드 1 내지 723, 또는 이에 대하여 적어도 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 서열을 포함하는 절단된 게놈이고;

[0414] 제2 아넬로바이러스 게놈은 제1 아넬로바이러스 게놈의 3'에 있는, 핵산 작제물.

[0415] 148. 구현예 114 내지 143 중 어느 하나에 있어서,

[0416] 제1 아넬로바이러스 게놈은 외인성 이펙터를 인코딩하고,

[0417] 제2 아넬로바이러스 게놈은 아넬로바이러스 게놈(예를 들어, Ring2)의 5' 최말단의 723 내지 2264, 2264 내지 2583, 또는 2583 내지 2812개 뉴클레오티드, 또는 이에 대하여 적어도 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 서열을 포함하는 절단된 게놈이고;

[0418] 제2 아넬로바이러스 게놈은 제1 아넬로바이러스 게놈의 3'에 있는, 핵산 작제물.

[0419] 149. 구현예 114 내지 143 중 어느 하나에 있어서,

[0420] 제1 아넬로바이러스 게놈은 외인성 이펙터를 인코딩하고,

[0421] 제2 아넬로바이러스 게놈은 아넬로바이러스 게놈(예를 들어, Ring2)의 5' 최말단의 700 내지 800, 800 내지 1000, 1000 내지 1500, 1500 내지 2000, 2000 내지 2500, 또는 2500 내지 2900개 뉴클레오티드, 또는 이에 대하여 적어도 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 서열을 포함하는 절단된 게놈이고;

[0422] 제2 아넬로바이러스 게놈은 제1 아넬로바이러스 게놈의 3'에 있는, 핵산 작제물.

[0423] 150. 구현예 114 내지 143 중 어느 하나에 있어서,

[0424] 제1 아넬로바이러스 게놈은 외인성 이펙터를 인코딩하고,

[0425] 제2 아넬로바이러스 게놈은 아넬로바이러스 게놈(예를 들어, Ring2)의 3' 최말단의 2712개 뉴클레오티드(예를 들어, 뉴클레오티드 256 내지 2979), 또는 이에 대하여 적어도 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 서열을 포함하는 절단된 게놈이고;

[0426] 제2 아넬로바이러스 게놈은 제1 아넬로바이러스 게놈의 5'에 있는, 핵산 작제물.

[0427] 151. 구현예 114 내지 143 중 어느 하나에 있어서,

[0428] 제1 아넬로바이러스 게놈은 외인성 이펙터를 인코딩하고,

[0429] 제2 아넬로바이러스 게놈은 아넬로바이러스 게놈(예를 들어, Ring2)의 3' 최말단의 2556개 뉴클레오티드(예를 들어, 뉴클레오티드 423 내지 2979), 또는 이에 대하여 적어도 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 서열을 포함하는 절단된 게놈이고;

[0430] 제2 아넬로바이러스 게놈은 제1 아넬로바이러스 게놈의 5'에 있는, 핵산 작제물.

[0431] 152. 구현예 114 내지 143 중 어느 하나에 있어서,

[0432] 제1 아넬로바이러스 게놈은 외인성 이펙터를 인코딩하고,

[0433] 제2 아넬로바이러스 게놈은 아넬로바이러스 게놈(예를 들어, Ring2)의 3' 최말단의 2256개 뉴클레오티드(예를 들어, 뉴클레오티드 723 내지 2979), 또는 이에 대하여 적어도 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 서열을 포함하는 절단된 게놈이고;

[0434] 제2 아넬로바이러스 게놈은 제1 아넬로바이러스 게놈의 5'에 있는, 핵산 작제물.

[0435] 153. 구현예 114 내지 143 중 어느 하나에 있어서,

[0436] 제1 아넬로바이러스 게놈은 외인성 이펙터를 인코딩하고,

[0437] 제2 아넬로바이러스 게놈은 아넬로바이러스 게놈(예를 들어, Ring2)의 3' 최말단의 706개 뉴클레오티드(예를 들어, 뉴클레오티드 2273 내지 2979), 또는 이에 대하여 적어도 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 서열을 포함하는 절단된 게놈이고;

[0438] 제2 아넬로바이러스 게놈은 제1 아넬로바이러스 게놈의 5'에 있는, 핵산 작제물.

[0439] 153. 구현예 114 내지 143 중 어느 하나에 있어서,

[0440] 제1 아넬로바이러스 게놈은 외인성 이펙터를 인코딩하고,

[0441] 제2 아넬로바이러스 게놈은 아넬로바이러스 게놈(예를 들어, Ring2)의 3' 최말단의 706 내지 2256, 2256 내지 2556, 또는 2556 내지 2712개 뉴클레오티드, 또는 이에 대하여 적어도 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 서열을 포함하는 절단된 게놈이고;

[0442] 제2 아넬로바이러스 게놈은 제1 아넬로바이러스 게놈의 5'에 있는, 핵산 작제물.

[0443] 154. 구현예 114 내지 143 중 어느 하나에 있어서,

[0444] 제1 아넬로바이러스 게놈은 외인성 이펙터를 인코딩하고,

[0445] 제2 아넬로바이러스 게놈은 아넬로바이러스 게놈(예를 들어, Ring2)의 3' 최말단의 700 내지 800, 800 내지 1000, 1000 내지 1500, 1500 내지 2000, 2000 내지 2500, 또는 2500 내지 2800개 뉴클레오티드, 또는 이에 대

하여 적어도 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 서열을 포함하는 절단된 게놈이고;

[0446] 제2 아넬로바이러스 게놈은 제1 아넬로바이러스 게놈의 5'에 있는, 핵산 작제물.

[0447] 155. 제114항 내지 제154항 중 어느 한 항에 있어서, 제2 아넬로바이러스 게놈은 ORF1, ORF2, ORF2/2, ORF2/3, ORF1/1, 또는 ORF1/2 중 하나 이상을 인코딩하는, 핵산 작제물.

[0448] 156. 단백질성 외부에 봉입된 유전 요소를 포함하는 아넬로벡터의 제조 방법으로서,

[0449] a) 구현예 114 내지 155 중 어느 하나의 핵산 작제물 및 아넬로바이러스 유전 요소(예를 들어, 여기서 아넬로바이러스 유전 요소는 핵산 작제물로부터 증폭된 것임)의 하나 이상의 카피를 포함하는 세포(예를 들어, 포유동물 숙주 세포, 예를 들어, MOLT-4 세포)를 제공하는 단계;

[0450] b) 아넬로바이러스 유전 요소가 세포의 단백질성 외부(예를 들어, 아넬로바이러스 ORF1 분자를 포함함)에 봉입될 수 있도록 하는 조건 하에서 세포를 인큐베이션하는 단계

[0451] 를 포함하여, 이에 의해 아넬로벡터를 제조하는 방법.

[0452] 157. 구현예 156에 있어서, 단백질성 외부는 시스 또는 트랜스로 제공되는, 방법.

[0453] 158. 구현예 156 또는 157에 있어서, 단백질성 외부에 봉입된 아넬로바이러스 유전 요소는 감염성 입자, 예를 들어, 바이러스 입자를 형성하는, 방법.

[0454] 159. 구현예 156 내지 158 중 어느 하나에 있어서, 단계 a) 이전에, 세포를 핵산 작제물로 형질감염시키는 단계를 추가로 포함하는, 방법.

[0455] 160. 구현예 156 내지 159 중 어느 하나에 있어서, 단계 b) 후에, 세포를 용해시키는 단계를 추가로 포함하는, 방법.

[0456] 161. 구현예 160에 있어서, 세포 용해물을 벤조나제(예를 들어, 실온에서, 예를 들어, 약 60, 70, 80, 90, 100, 110, 또는 120분 동안, 예를 들어, 100 U/ml의 벤조나제)로 처리하는 단계를 추가로 포함하고, 선택적으로 세포 용해물을 정화시키는 단계 및/또는 등밀도 원심분리 단계를 추가로 포함하는, 방법.

[0457] 162. 구현예 160 또는 161에 있어서, 용해물을 분획화하는 단계 및 아넬로벡터를 함유하는 분획을 수집하는 단계를 추가로 포함하는, 방법.

[0458] 163. 구현예 156 내지 162에 있어서, 아넬로벡터는 세포(예를 들어, 포유동물 세포, 예를 들어, 인간 세포)를 감염시킬 수 있는, 방법

[0459] 164. 외인성 이펙터를 세포로 운반하는 방법으로서, 구현예 99 내지 106 또는 156 내지 163 중 어느 하나의 방법에 의해 제조된 아넬로벡터를 세포 내로 도입하는 단계 및 외인성 이펙터의 발현에 적합한 조건 하에서 세포를 인큐베이션하는 단계를 포함하는 방법.

[0460] 본 발명의 다른 특징, 목적 및 이점이 설명 및 도면, 및 청구범위로부터 명백해질 것이다.

[0461] 달리 정의되지 않는 한, 본 명세서에 사용되는 모든 기술 및 과학 용어는 본 발명이 속하는 분야의 당업자에 의해 일반적으로 이해되는 바와 동일한 의미를 갖는다. 본 명세서에 언급된 모든 간행물, 특허 출원, 특허 및 다른 참고문헌은 전문이 참조로 포함된다. 추가적으로, 물질, 방법 및 실시예는 단지 예시적인 것일 뿐이며, 제한하려는 것이 아니다.

도면의 간단한 설명

[0462] 도 1a 내지 1c는 탠덤 아넬로바이러스 플라스미드가 아넬로바이러스 생성을 증가시킬 수 있음을 나타내는 일련의 다이어그램이다. (도 1a) 예시적인 탠덤 아넬로바이러스 플라스미드에 대한 플라스미드 맵. (도 1b) 탠덤 아넬로바이러스 플라스미드로 MOLT-4 세포를 형질감염시켜 야생형 크기의 아넬로바이러스 게놈을 회수하였다. (도 1c) 탠덤 아넬로바이러스 플라스미드로부터 MOLT-4 세포에서 생성된 아넬로바이러스 게놈은 캡슐화된 바이러스 입자에 대한 예상 밀도로 이동한다. GCR = GC-풍부 영역. 박테리아 SM = 박테리아 선택 마커. 박테리아 ori = 박테리아 복제 원점. ORF = 오픈-리딩 프레임. Prom. = 프로모터. 5CD = 5' 비번역 영역의 보존된 도메인.

도 2a 내지 2e는 Ring2 게놈에 기반한 예시적인 탠덤 작제물을 나타내는 일련의 다이어그램이다. (도 2a) 유전 요소의 제1 카피와 제1 카피에 대해 3'에 위치한 유전 요소의 전체 또는 부분적 제2 카피를 포함하는 탠덤 작제

물. 각각의 연속적인 작제물은 제2 카피의 3' 말단의 더 큰 절단을 포함한다. 작제물은, 예를 들어 표시된 바와 같이 5CD(5' UTR 보존된 도메인)를 포함하는, 하류 복제-촉진 서열(dRFS)을 포함할 수 있다. (도 2b) 유전 요소의 제1 카피와 제1 카피에 대해 5'에 위치한 유전 요소의 전체 또는 부분적 제2 카피를 포함하는 탠덤 작제물. 각각의 연속적인 작제물은 제2 카피의 5' 말단의 더 큰 절단을 포함한다. (도 2c) 유전 요소의 부분적인 제1 카피(예를 들어, uRFS를 포함함) 및 제1 카피에 대해 5'에 위치한 유전 요소의 부분적인 제2 카피(예를 들어, dRFS를 포함함)를 포함하는 탠덤 작제물. 각각의 연속적인 작제물은 제1 카피의 5' 말단의 더 큰 절단과 제2 카피의 3' 말단의 더 큰 비율을 포함한다. (도 2d) 도 2a 및 도 2b에 나타낸 작제물로 형질감염된 MOLT-4 세포로부터 수확된 총 DNA에 대한 서던 블롯으로서, 야생형 길이 아넬로바이러스 게놈의 회수를 입증한다. (도 2e) CsCl 밀도 구배로부터의 아넬로바이러스 게놈의 DNase-보호 qPCR로서, 도 2a 및 도 2b에 나타낸 작제물로 MOLT-4 세포에서 생성된 아넬로바이러스 게놈의 봉입을 입증한다.

도 2f는 Ring1 백본의 제1 카피에서, eGFP-ORF1 융합 단백질을 인코딩하는 서열을 인코딩하는 탠덤 Ring1 작제물을 포함하여, 다양한 Ring1 작제물(나타낸 바와 같음)로 형질감염된 Jurkat 세포로부터 전장의 Ring1 ORF1 mRNA에 대한 긴 RNA 판독물을 나타내는 일련의 다이어그램이다.

도 2g는 Ring2 탠덤 작제물이 뉴클레오펙션에 의해 도입된 MOLT-4 세포에서 ORF1 단백질 발현의 검출을 나타내는 일련의 다이어그램이다.

도 2h는 나란히 배열된 2개의 Ring2 게놈을 포함하는 예시적인 바큘로바이러스 작제물을 나타내는 다이어그램이다.

도 2i는 바큘로바이러스를 통한 탠덤 Ring2 게놈의 Sf9 세포로의 운반을 나타내는 일련의 다이어그램이다.

도 3은 TTMiniV("아넬로벡터 1")의 LY1 균주를 인코딩하는 카나마이신 벡터의 개략도를 도시한다.

도 4는 TTMiniV("아넬로벡터 2")의 LY2 균주를 인코딩하는 카나마이신 벡터의 개략도를 도시한다.

도 5는 293T 및 A549 세포에서 합성 아넬로벡터의 형질감염 효율을 도시한다.

도 6a 및 6b는 합성 아넬로벡터에 의한 293T 세포의 성공적인 감염을 예시하는 정량적 PCR 결과를 도시한다.

도 7a 및 7b는 합성 아넬로벡터에 의한 A549 세포의 성공적인 감염을 예시하는 정량적 PCR 결과를 도시한다.

도 8a 및 8b는 합성 아넬로벡터에 의한 Raji 세포의 성공적인 감염을 예시하는 정량적 PCR 결과를 도시한다.

도 9a 및 9b는 합성 아넬로벡터에 의한 Jurkat 세포의 성공적인 감염을 예시하는 정량적 PCR 결과를 도시한다.

도 10a 및 10b는 합성 아넬로벡터에 의한 Chang 세포의 성공적인 감염을 예시하는 정량적 PCR 결과를 도시한다.

도 11은 아넬로벡터(예를 들어, 본 명세서에 기재된 바와 같은 복제-적격 또는 복제-결핍 아넬로벡터)의 생성을 위한 예시적인 작업 흐름을 나타내는 개략도이다.

도 12는 지시된 플라스미드로 형질감염된 HEK293T 세포에서 miR-625 발현의 배수 변화를 나타내는 그래프이다.

도 13은 n-myc 상호작용 단백질(NMI)을 표적화하는 miRNA를 인코딩하는 아넬로벡터를 이용한 Raji B 세포의 감염을 나타내는 다이어그램이다. Raji B 세포(화살표) 또는 대조군 세포를 NMI miRNA-인코딩 아넬로벡터로 감염시킨 후 검출된 아넬로벡터의 게놈 등가물의 정량화가 나타나 있다.

도 14는 n-myc 상호작용 단백질(NMI)을 표적화하는 miRNA를 인코딩하는 아넬로벡터를 이용한 Raji B 세포의 감염을 나타내는 다이어그램이다. 웨스턴 블롯은 NMI에 대한 miRNA를 인코딩하는 아넬로벡터가 Raji B 세포에서 NMI 단백질 발현을 감소시키는 반면, miRNA가 결여된 아넬로벡터로 감염된 Raji B 세포는 대조군과 비슷한 NMI 단백질 발현을 나타내었음을 보여준다.

도 15는 내인성 miRNA-인코딩 서열을 포함하는 아넬로벡터 및 내인성 miRNA-인코딩 서열이 결실된 상응하는 아넬로벡터로 감염시킨 후 숙주 세포에서 생성된 아넬로벡터 입자의 정량화를 나타내는 일련의 그래프이다.

도 16a 및 16b는 나노-루시퍼라제를 발현하는 아넬로벡터를 생성하는 데 사용되는 작제물(도 16a) 및 세포를 형질감염시키는 데 사용되는 일련의 아넬로벡터/플라스미드 조합(도 16b)을 나타내는 일련의 다이어그램이다.

도 17a 내지 17c는 아넬로벡터가 주입된 마우스에서 나노-루시퍼라제 발현을 나타내는 일련의 다이어그램이다. (도 17a) 주사 후 제0일 내지 제9일에 마우스에서의 나노-루시퍼라제 발현. (도 17b) 표시된 바와 같이 다양한 아넬로벡터/플라스미드 작제물 조합을 주사한 마우스에서의 나노-루시퍼라제 발현. (도 17c) 주사 후 마우스에

서 검출된 나노-루시퍼라제 발광의 정량화. 그룹 A는 TTMV-LY2 벡터 ± 나노-루시퍼라제를 받았다. 그룹 B는 나노 루시퍼라제 단백질과 TTMV-LY2 ORF를 받았다.

도 18a는 TTMV-LY2 플라스미드 pVL46-063 및 pVL46-240의 환화를 나타내는 겔 전기영동 이미지이다.

도 18b는 크기 배제 크로마토그래피(SEC)에 의해 결정된 선형 및 환형 TTMV-LY2 작제물에 대한 카피 수를 나타내는 크로마토그램이다.

도 18c는 아넬로바이러스 ORF1 분자의 도메인 및 상이한 아넬로바이러스로부터의 초가변 도메인으로 대체될 초가변 영역을 나타내는 개략도이다.

도 18d는 비-아넬로바이러스 공급원으로부터의 관심 단백질 또는 펩티드(POI)로 대체될 초가변 영역 및 ORF1의 도메인을 나타내는 개략도이다.

도 19는 인간 에리트로포이에틴(hEpo)을 인코딩하는 서열을 함유하도록 엔지니어링된 tth8 또는 LY2에 기반한 아넬로벡터가 포유동물 세포로 기능성 이식유전자를 운반할 수 있음을 나타내는 그래프이다.

도 20a 및 20b는 마우스에 투여된 엔지니어링된 아넬로벡터가 정맥 주사 7일 후에 검출 가능했음을 나타내는 일련의 그래프이다.

도 21은 hGH를 인코딩하는 엔지니어링된 아넬로벡터의 정맥 투여 7일 후 전혈의 세포 분획에서 hGH mRNA가 검출되었음을 나타내는 그래프이다.

도 22는 HEK293T 세포에서 예상 밀도로 TTV-tth8 게놈 카피를 생성하는 플라스미드에서, TTV-tth8 게놈과 비교하여 시험관 내 환화(IVC) TTV-tth8 게놈(IVC TTV-tth8)의 능력을 나타내는 그래프이다.

도 23은 Jurkat 세포에서 예상 밀도로 LY2 게놈 카피를 생성하는 플라스미드(WT LY2 플라스미드)에서 시험관 내 환화(IVC) LY2 게놈(WT LY2 IVC) 및 야생형 LY2 게놈의 능력을 나타내는 일련의 그래프이다.

도 24a 및 24b는 각각 야생형 Ring2 게놈의 2개 카피를 포함하는 예시적인 탠덤 작제물을 나타내는 일련의 다이어그램이다. 제1 작제물에서, Ring2 게놈의 카피는 둘 다 야생형 서열이다. 나머지 작제물에서, Ring2 게놈의 하나의 카피는 야생형이고, Ring2 게놈의 다른 카피는 삽입된 나노-루시퍼라제 카세트를 포함한다. 나노-루시퍼라제 카세트를 갖는 Ring2 게놈의 카피는 추가로 돌연변이되어 Ring2 게놈의 야생형 카피만이 ORF 유전자 발현이 가능하도록 개시 코돈을 돌연변이시킴으로써 내부 ORF의 유전자 발현을 녹아웃시킨다.

도 25는 등밀도 초원심분리에 의해 탠덤 작제물로 형질감염된 MOLT-4 세포로부터 Ring2 입자의 구제를 입증하는 대표적인 데이터를 나타내는 그래프이다. 각각의 분획에 대한 밀도는 검은색으로 도시되고 각각의 분획에 대한 DNAse 보호 RING2 역가는 회색으로 나타나 있다.

도 26a 및 26b는 (도 26a) qPCR, 웨스턴 블롯, 및 쿠마씨 염색; 및 (도 26b) 투과 전자 현미경(TEM) 분석을 사용한 정제된 Ring2 입자의 특성화를 나타내는 일련의 다이어그램이다.

본 발명의 구현예의 하기의 상세한 설명은 첨부된 도면과 함께 읽으면 더 잘 이해될 것이다. 본 발명의 예시 목적을 위하여, 본 명세서에 예시된 구현예가 도면에 나타나 있다. 그러나, 본 발명이 도면에 나타낸 구현예의 정확한 배열 및 수단으로 제한되지 않음을 이해해야 한다.

발명을 실시하기 위한 구체적인 내용

[0463] 정의

[0464] 본 발명은 특정 구현예에 대해 그리고 특정 도면을 참조하여 설명될 것이지만 본 발명은 이에 한정되지 않고, 청구범위에 의해서만 한정된다. 이하에 기재된 용어는 일반적으로 다르게 지시되지 않는 한 그의 일반적인 의미로 이해될 것이다.

[0465] 발명을 실시하기 위한 구체적인 내용 및 청구범위에서 "포함하다"라는 용어가 사용되는 경우, 이는 다른 요소를 배제하지 않는다. 본 발명의 목적을 위하여, "로 이루어지다"라는 용어는 "구성된다"라는 용어의 바람직한 구현예인 것으로 간주된다. 이하에서 그룹이 적어도 특정 개수의 구현예를 포함하는 것으로 정의되는 경우, 이는 또한 바람직하게는 이들 구현예만으로 이루어지는 그룹을 개시하는 것으로 이해될 것이다.

[0466] 단수 명사를 언급할 때, 부정 관사 또는 정관사, 예를 들어, 하나의("a", "an") 또는 상기("the")가 사용되는 경우, 이는 그 밖의 어떤 것이 특별히 언급되지 않는 한, 그 명사의 복수형을 포함한다.

[0467] 용어 "치료, 조절 등을 위한 화합물, 조성물, 생성물 등"은 표기된 치료, 조절 등의 목적에 적합한 화합물, 조성물, 생성물 등 그 자체를 지칭함을 이해할 것이다. 용어 "치료, 조절 등을 위한 화합물, 조성물, 생성물 등"은 추가로 일 구현예로서, 이러한 화합물, 조성물, 생성물 등이 치료, 조절 등에 사용하기 위한 것임을 개시한다.

[0468] 용어 "...에 사용하기 위한 화합물, 조성물, 생성물 등", "...를 위한 의약, 약제학적 조성물, 수의학적 조성물, 진단적 조성물 등의 제조에서의 화합물, 조성물, 생성물 등의 용도", 또는 "의약으로서 사용하기 위한 화합물, 조성물, 생성물 등..."은 이러한 화합물, 조성물, 생성물 등이 인간 또는 동물 신체에서 실시될 수 있는 치료적 방법에 사용될 것임을 나타낸다. 그들은 치료 방법 등에 관한 구현예 및 청구범위의 동등한 개시로서 간주된다. 따라서, 구현예 또는 청구범위가 "질환을 앓는 것으로 의심되는 인간 또는 동물을 치료하는 데 사용하기 위한 화합물"을 지칭한다면, 이는 또한 "질환을 앓는 것으로 의심되는 인간 또는 동물을 치료하기 위한 의약의 제조에서의 화합물의 용도" 또는 "질환을 앓는 것으로 의심되는 인간 또는 동물로의 화합물의 투여에 의한 치료 방법"의 개시인 것으로 간주된다. 용어 "치료, 조절 등을 위한 화합물, 조성물, 생성물 등"은 표기된 치료, 조절 등의 목적에 적합한 화합물, 조성물, 생성물 등 그 자체를 지칭함을 이해할 것이다.

[0469] 기간, 값, 수 등의 하기의 예가 괄호 안에 제공된다면, 이는 괄호 안에 언급된 예가 구현예를 구성할 수 있다는 표시로서 이해해야 한다. 예를 들어, "구현예에서, 핵산 분자가 표 1의 아넬로바이러스 ORF1-인코딩 뉴클레오티드 서열(예를 들어, 표 1의 핵산 서열의 뉴클레오티드 571 내지 2613)에 대하여 적어도 약 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100% 서열 동일성을 갖는 핵산 서열을 포함한다"고 언급된다면, 일부 구현예는 표 1의 핵산 서열의 뉴클레오티드 571 내지 2613에 대하여 적어도 약 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100% 서열 동일성을 갖는 핵산 서열을 포함하는 핵산 분자에 관한 것이다.

[0470] 본 명세서에 사용되는 바와 같이, 용어 "증폭"은 핵산 분자 또는 이의 일부(예를 들어, 유전 요소 또는 유전 요소 영역)의 하나 이상의 추가의 카피를 생성하기 위한 핵산 분자 또는 이의 일부의 복제를 지칭한다. 일부 구현예에서, 증폭은 핵산 서열의 부분 복제를 초래한다. 일부 구현예에서, 증폭은 회전환 복제(rolling circle replication)를 통해 발생한다.

[0471] 본 명세서에 사용되는 바와 같이, 용어 "아넬로벡터"는 단백질성 외부 내에 봉입된 유전 요소, 예를 들어, 환형 DNA를 포함하는 비히클을 지칭하며, 예를 들어 유전 요소는 단백질성 외부에 의해 DNAse I을 이용한 분해로부터 실질적으로 보호된다. 본 명세서에 사용되는 바와 같이, "합성 아넬로벡터"는 일반적으로, 천연 발생이 아닌, 예를 들어, 야생형 바이러스(예를 들어, 본 명세서에 기재된 바와 같은 야생형 아넬로바이러스)에 비하여 상이한 서열을 갖는 아넬로벡터를 지칭한다. 일부 구현예에서, 합성 아넬로벡터는 엔지니어링되거나 또는 재조합이며, 예를 들어, 야생형 바이러스 게놈(예를 들어, 본 명세서에 기재된 바와 같은 야생형 아넬로바이러스 게놈)에 비하여 차이 또는 변형을 포함하는 유전 요소를 포함한다. 일부 구현예에서, 단백질성 외부 내의 봉입은 단백질성 외부에 의한 100% 커버리지, 및 100% 미만, 예를 들어, 95%, 90%, 85%, 80%, 70%, 60%, 50% 또는 그 미만의 커버리지를 포함한다. 유전 요소가 예를 들어, 숙주 세포로의 유입 이전에 단백질성 외부에 보유되거나 DNAse I을 이용한 분해로부터 보호되는 한, 예를 들어, (예를 들어, 단백질성 외부를 물, 이온, 펩티드 또는 소분자에 대해 투과 가능하게 만드는) 갭 또는 불연속부가 단백질성 외부에 존재할 수 있다. 일부 구현예에서, 아넬로벡터는 정제되며, 예를 들어, 아넬로벡터는 이의 원래의 공급원으로부터 분리되고/분리되거나 다른 성분이 실질적으로 없다(50% 초과, 60% 초과, 70% 초과, 80% 초과, 90% 초과). 일부 구현예에서, 아넬로벡터는 유전 요소를 (예를 들어, 감염을 통해) 표적 세포로 도입할 수 있다. 일부 구현예에서, 아넬로벡터는 감염성 합성 아넬로바이러스 바이러스 입자이다.

[0472] 본 명세서에 사용되는 바와 같이, 용어 "항체 분자"는 단백질, 예를 들어, 적어도 하나의 면역글로불린 가변 도메인 서열을 포함하는 면역글로불린 쇄 또는 이의 단편을 지칭한다. 용어 "항체 분자"는 전장 항체 및 항체 단편(예를 들어, scFv)을 포함한다. 일부 구현예에서, 항체 분자는 다중특이적 항체 분자이며, 예를 들어, 항체 분자는 복수의 면역글로불린 가변 도메인 서열을 포함하며, 복수의 것들 중 제1 면역글로불린 가변 도메인 서열은 제1 에피토프에 대한 결합 특이성을 가지며, 복수의 것들 중 제2 면역글로불린 가변 도메인 서열은 제2 에피토프에 대한 결합 특이성을 갖는다. 구현예에서, 다중특이적 항체 분자는 이중특이적 항체 분자이다. 이중특이적 항체 분자는 일반적으로 제1 에피토프에 대한 결합 특이성을 갖는 제1 면역글로불린 가변 도메인 서열 및 제2 에피토프에 대한 결합 특이성을 갖는 제2 면역글로불린 가변 도메인 서열에 의해 특성화된다.

[0473] 본 명세서에 사용되는 바와 같이, "하류 복제-촉진 서열"(dRFS)은, 유전 요소 서열의 하류에 위치할 때(예를 들어, 유전 요소는 dRFS에 대해 5'임), dRFS가 없는 다른 유사한 유전 요소 서열과 비교하여 유전 요소 서열의 복

제를 증가시키는 유전 요소(예를 들어, 본 명세서에 기재된 바와 같음)의 서열의 단편을 지칭한다. 일반적으로, 생성된 복제된 가닥은 아넬로벡터(예를 들어, 본 명세서에 기재된 바와 같음)를 형성하기 위해 단백질성 외부에 봉입될 수 있는 기능적 유전 요소이다. 일부 구현예에서, dRFS는 Rep 단백질(예를 들어, 아넬로바이러스 Rep 단백질)에 대한 변위 부위를 포함한다. 일부 구현예에서, dRFS는 아넬로바이러스 3' UTR 서열 또는 이의 단편, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 서열 동일성을 갖는 서열을 포함한다. 일부 구현예에서, dRFS는 5' UTR(예를 들어, 헤어핀 루프를 포함함)을 포함한다. 일부 구현예에서, dRFS는 복제 원점을 포함한다.

[0474] 본 명세서에 사용되는 바와 같이, "상류 복제-촉진 서열"(uRFS)은, 유전 요소 서열의 상류에 위치할 때(예를 들어, 유전 요소는 uRFS에 대해 3'임), uRFS가 없는 다른 유사한 유전 요소 서열과 비교하여 유전 요소 서열의 복제를 증가시키는 유전 요소(예를 들어, 본 명세서에 기재된 바와 같음)의 서열의 단편을 지칭한다. 일반적으로, 생성된 복제된 가닥은 아넬로벡터(예를 들어, 본 명세서에 기재된 바와 같음)를 형성하기 위해 단백질성 외부에 봉입될 수 있는 기능적 유전 요소이다. 일부 구현예에서, uRFS는 Rep 단백질(예를 들어, 아넬로바이러스 Rep 단백질)에 대한 결합 및/또는 인식 부위를 포함한다. 일부 구현예에서, uRFS는 아넬로바이러스 5' UTR 서열 또는 이의 단편, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 서열 동일성을 갖는 서열을 포함한다. 일부 구현예에서, uRFS는 5' UTR(예를 들어, 헤어핀 루프를 포함함)을 포함한다. 일부 구현예에서, uRFS는 복제 원점을 포함한다.

[0475] 본 명세서에 사용되는 바와 같이, "인코딩"하는 핵산은 아미노산 서열 또는 기능적 폴리뉴클레오티드(예를 들어, 비-코딩 RNA, 예를 들어, siRNA 또는 miRNA)를 인코딩하는 핵산 서열을 지칭한다.

[0476] 본 명세서에 사용되는 바와 같은 "외인성" 작용제(예를 들어, 이펙터, 핵산(예를 들어, RNA), 유전자, 페이로드, 단백질)는 상응하는 야생형 바이러스, 예를 들어, 본 명세서에 기재된 바와 같은 아넬로바이러스에 의해 포함되지 않거나, 그에 의해 인코딩되지 않는 작용제를 지칭한다. 일부 구현예에서, 외인성 작용제, 예컨대, 천연 발생 단백질 또는 핵산에 비하여 (예를 들어, 삽입, 결실 또는 치환에 의해) 변경된 서열을 갖는 단백질 또는 핵산은 자연적으로 존재하지 않는다. 일부 구현예에서, 외인성 작용제는 숙주 세포에서 자연적으로 존재하지 않는다. 일부 구현예에서, 외인성 작용제는 숙주 세포에서 자연적으로 존재하지만, 바이러스에 대하여 외인성이다. 일부 구현예에서, 외인성 작용제는 숙주 세포에서 자연적으로 존재하지만, 요망되는 수준으로 또는 요망되는 시간에 존재하지 않는다.

[0477] 또 다른 작용제 또는 요소(예를 들어, 이펙터, 핵산 서열, 아미노산 서열)에 관하여 본 명세서에 사용되는 바와 같은, "이종" 작용제 또는 요소(예를 들어, 이펙터, 핵산 서열, 아미노산 서열)는 천연적으로, 예를 들어, 야생형 바이러스, 예를 들어, 아넬로바이러스에서 함께 관찰되지 않는 작용제 또는 요소를 지칭한다. 일부 구현예에서, 이종 핵산 서열은 천연 발생 핵산 서열(예를 들어, 아넬로바이러스에서 천연적으로 발생하는 서열)과 동일한 핵산에 존재할 수 있다. 일부 구현예에서, 이종 작용제 또는 요소는 아넬로벡터의 다른(예를 들어, 나머지) 요소가 기반으로 하는 아넬로바이러스에 비하여 외인성이다.

[0478] 본 명세서에 사용되는 바와 같이, 용어 "유전 요소"는, 예를 들어, 본 명세서에 기재된 바와 같은 아넬로벡터를 형성하기 위해, 단백질성 외부이거나 단백질성 외부 내에 둘러싸일 수 있는(예를 들어, 단백질성 외부에 의해 DNAse I로부터 보호될 수 있는) 핵산 분자를 지칭한다. 유전 요소가 네이키드 DNA로서 생성되고, 선택적으로 추가로 단백질성 외부 내로 조립될 수 있는 것이 이해된다. 또한, 아넬로벡터는 이의 유전 요소를 세포 내로 삽입하여, 유전 요소가 세포 내에 존재하게 하고, 단백질성 외부가 반드시 세포에 유입되는 것은 아니라는 점이 이해된다.

[0479] 본 명세서에 사용되는 바와 같이, "유전 요소 작제물"은 적어도 하나(예를 들어, 2개)의 유전 요소 서열(들), 또는 이의 단편을 포함하는 핵산 작제물(예를 들어, 플라스미드, 박미드, 코스미드, 또는 미니서클)을 지칭한다. 일부 구현예에서, 본 명세서에 기재된 바와 같은 탠덤 작제물은 (예를 들어, 본 명세서에 기재된 바와 같이) 나란히 배열된, 2개 이상의 유전 요소 서열, 또는 이의 단편을 포함하는 유전 요소 작제물이다. 일부 구현예에서, 유전 요소 작제물은 적어도 하나의 전장 유전 요소 서열을 포함한다. 일부 구현예에서, 유전 요소는 전장 유전 요소 서열 및 부분적인 유전 요소 서열을 포함한다. 일부 구현예에서, 유전 요소는 2개 이상의 부분적인 유전 요소 서열(예를 들어, PCT/US19/65995의 도 2C에 나타낸 바와 같이, 예를 들어, 5'에서 3' 순서로, 3'-절단된 유전 요소 서열과 나란히 배열된 5'-절단된 유전 요소 서열)을 포함한다.

[0480] 본 명세서에 사용되는 바와 같이, 용어 "유전 요소 영역"은 유전 요소의 서열을 포함하는 작제물의 영역을 지칭한다. 일부 구현예에서, 유전 요소 영역은 야생형 아넬로바이러스 서열 또는 이의 단편에 대해 충분한 동일성을

갖는 서열(예를 들어, 야생형 아넬로바이러스 서열 또는 이의 단편에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100% 서열 동일성을 갖는 서열)을 포함하여 단백질성 외부로 둘러싸임으로써 아넬로벡터를 형성한다. 구현예에서, 유전 요소 영역은, 예를 들어, 본 명세서에 기재된 바와 같은 단백질 결합 서열(예를 들어, 본 명세서에 기재된 바와 같은 5' UTR, 3' UTR 및/또는 GC-풍부 영역, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 서열)을 포함한다. 일부 구현예에서, 유전 요소 영역은 회전환 복제를 겪을 수 있다. 일부 구현예에서, 유전 요소 영역은 uRFS를 포함한다. 일부 구현예에서, 유전 요소 영역은 dRFS를 포함한다. 일부 구현예에서, 유전 요소는 Rep 단백질 결합 부위를 포함한다. 일부 구현예에서, 유전 요소는 Rep 단백질 변위 부위를 포함한다. 일부 구현예에서, 유전 요소 영역을 포함하는 작제물은 단백질성 외부에 봉입되지 않지만, 작제물로부터 생성된 유전 요소는 단백질성 외부에 봉입될 수 있다. 일부 구현예에서, 유전 요소 영역을 포함하는 작제물은 제2 uRFS 또는 제2 dRFS를 추가로 포함한다. 일부 구현예에서, 유전 요소 영역을 포함하는 작제물은 벡터 백본을 추가로 포함한다.

[0481] 본 명세서에 사용되는 바와 같이, 용어 "돌연변이체"는 게놈(예를 들어, 아넬로바이러스 게놈) 또는 이의 단편과 관련하여 사용될 때 상응하는 야생형 아넬로바이러스 서열에 대해 적어도 하나의 변화를 갖는 서열을 지칭한다. 일부 구현예에서, 돌연변이체 게놈 또는 이의 단편은 상응하는 야생형 아넬로바이러스 서열에 대해 적어도 하나의 단일 뉴클레오티드 다형성, 추가, 결실, 또는 프레임시프트를 포함한다. 일부 구현예에서, 돌연변이체 게놈 또는 이의 단편은 상응하는 야생형 아넬로바이러스 서열에 대해 적어도 하나의 아넬로바이러스 ORF(예를 들어, ORF1, ORF2, ORF2/2, ORF2/3, ORF1/1, 및/또는 ORF1/2 중 하나 이상)의 결실을 포함한다. 일부 구현예에서, 돌연변이체 게놈 또는 이의 단편은 상응하는 야생형 아넬로바이러스 서열에 대해 모든 아넬로바이러스 ORF(예를 들어, ORF1, ORF2, ORF2/2, ORF2/3, ORF1/1, 및 ORF1/2 전부)의 결실을 포함한다. 일부 구현예에서, 돌연변이체 게놈 또는 이의 단편은 상응하는 야생형 아넬로바이러스 서열에 대해 적어도 하나의 아넬로바이러스 비코딩 영역(예를 들어, 5' UTR, 3' UTR, 및/또는 GC-풍부 영역 중 하나 이상)의 결실을 포함한다. 일부 구현예에서, 돌연변이체 게놈 또는 이의 단편은 외인성 이펙터를 포함하거나 인코딩한다.

[0482] 본 명세서에 사용되는 바와 같이, 용어 "ORF1 분자"는 아넬로바이러스 ORF1 단백질(예를 들어, 본 명세서에 기재된 바와 같은 아넬로바이러스 ORF1 단백질)의 활성 및/또는 구조적 특징을 갖는 폴리펩티드 또는 이의 기능적 단편을 지칭한다. ORF1 분자는 일부 예에서, 하기 중 하나 이상(예를 들어, 이 중 1, 2, 3 또는 4개)을 포함할 수 있다: 적어도 60%의 염기성 잔기(예를 들어, 적어도 60%의 아르기닌 잔기)를 포함하는 제1 영역, 적어도 약 6개의 베타 가닥(예를 들어, 적어도 4, 5, 6, 7, 8, 9, 10, 11 또는 12개의 베타 가닥)을 포함하는 제2 영역, 아넬로바이러스 N22 도메인(예를 들어, 본 명세서에 기재된 바와 같음, 예를 들어, 본 명세서에 기재된 바와 같은 아넬로바이러스 ORF1 단백질로부터의 N22 도메인)의 구조 또는 활성을 포함하는 제3 영역 및/또는 아넬로바이러스 C-말단 도메인(CTD)(예를 들어, 본 명세서에 기재된 바와 같음, 예를 들어, 본 명세서에 기재된 바와 같은 아넬로바이러스 ORF1 단백질로부터의 CTD)의 구조 또는 활성을 포함하는 제4 영역. 일부 예에서, ORF1 분자는 N-말단에서 C-말단 순서로, 제1, 제2, 제3 및 제4 영역을 포함한다. 일부 예에서, 아넬로벡터는 N-말단에서 C-말단 순서로, 제1, 제2, 제3 및 제4 영역을 포함하는 ORF1 분자를 포함한다. ORF1 분자는 일부 예에서, 아넬로바이러스 ORF1 핵산에 의해 인코딩된 폴리펩티드를 포함할 수 있다. ORF1 분자는 일부 예에서, 이종 서열, 예를 들어, 초가변 영역(HVR), 예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어, 아넬로바이러스 ORF1 단백질로부터의 HVR을 추가로 포함할 수 있다. 본 명세서에 사용되는 바와 같이, "아넬로바이러스 ORF1 단백질"은 아넬로바이러스 게놈(예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어, 야생형 아넬로바이러스 게놈)에 의해 인코딩된 ORF1 단백질을 지칭한다.

[0483] 본 명세서에 사용되는 바와 같이, 용어 "ORF2 분자"는 아넬로바이러스 ORF2 단백질(예를 들어, 본 명세서에 기재된 바와 같은 아넬로바이러스 ORF2 단백질)의 활성 및/또는 구조적 특징을 갖는 폴리펩티드 또는 이의 기능적 단편을 지칭한다. 본 명세서에 사용되는 바와 같이, "아넬로바이러스 ORF2 단백질"은 아넬로바이러스 게놈(예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어, 야생형 아넬로바이러스 게놈)에 의해 인코딩된 ORF2 단백질을 지칭한다.

[0484] 본 명세서에 사용되는 바와 같이, 용어 "단백질성 외부"는 주로(예를 들어, 50% 초과, 60% 초과, 70% 초과, 80% 초과, 90% 초과) 단백질인 외부 구성성분을 지칭한다.

[0485] 본 명세서에 사용되는 바와 같이, 용어 "조절 핵산"은 발현 생성물을 인코딩하는 DNA 서열의 발현, 예를 들어, 전사 및/또는 번역을 변경시키는 핵산 서열을 지칭한다. 구현예에서, 발현 생성물은 RNA 또는 단백질을 포함한다.

[0486] 본 명세서에 사용되는 바와 같이, 용어 "조절 서열"은 표적 유전자 생성물의 전사를 변경시키는 핵산 서열을 지칭한다. 일부 구현예에서, 조절 서열은 프로모터 또는 인핸서이다.

[0487] 본 명세서에 사용되는 바와 같이, 용어 "Rep" 또는 "복제 단백질"은 바이러스 게놈 복제를 촉진하는 단백질, 예를 들어, 바이러스 단백질을 지칭한다. 일부 구현예에서, 복제 단백질은 아넬로바이러스 Rep 단백질이다.

[0488] 본 명세서에 사용되는 바와 같이, 용어 "Rep 결합 부위"는 Rep 단백질(예를 들어, 아넬로바이러스 Rep 단백질)에 의해 인식되고 결합되는 핵산 분자 내의 핵산 서열을 지칭한다. 일부 구현예에서, Rep 결합 부위는 5' UTR(예를 들어, 헤어핀 루프를 포함함)을 포함한다. 일부 구현예에서, Rep 결합 부위는 복제 원점(ORI)을 포함한다.

[0489] 본 명세서에 사용되는 바와 같이, 용어 "Rep 변위 부위"는 핵산 분자와 회합된(예를 들어, 이에 결합된) Rep 단백질(예를 들어, 아넬로바이러스 Rep 단백질)을 야기하여 Rep 변위 부위에 도달하면 핵산 분자를 방출할 수 있는 핵산 분자 내의 핵산 서열을 지칭한다. 일부 구현예에서, Rep 변위 부위는 5' UTR(예를 들어, 헤어핀 루프를 포함함)을 포함한다. 일부 구현예에서, Rep 변위 부위는 복제 원점(ORI)을 포함한다.

[0490] 본 명세서에 사용되는 바와 같이, "실질적으로 비-병원성인" 유기체, 입자 또는 성분은 예를 들어, 숙주 유기체, 예를 들어, 포유동물, 예를 들어, 인간에서 허용 불가능한 질환 또는 병원성 병태를 야기하거나 유도하지 않는 유기체, 입자(예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어, 바이러스 또는 아넬로벡터) 또는 이의 성분을 지칭한다. 일부 구현예에서, 대상체로의 아넬로벡터의 투여는 치료 기준의 일부로서 허용 가능한 부차 반응 또는 부작용을 초래할 수 있다.

[0491] 본 명세서에 사용되는 바와 같이, 용어 "비-병원성"은 숙주 유기체, 예를 들어, 포유동물, 예를 들어, 인간에서 허용 불가능한 질환 또는 병원성 병태를 야기하거나 유도하지 않는 유기체 또는 이의 성분을 지칭한다.

[0492] 본 명세서에 사용되는 바와 같이, "실질적으로 비-통합성인" 유전 요소는 숙주 세포(예를 들어, 진핵 세포) 또는 유기체(예를 들어, 포유동물, 예를 들어, 인간) 내로 유입되는 유전 요소 중 약 0.01%, 0.05%, 0.1%, 0.5% 또는 1% 미만이 게놈 내로 통합되는 유전 요소, 예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어, 바이러스 또는 아넬로벡터 내의 유전 요소를 지칭한다. 일부 구현예에서, 유전 요소는 예를 들어, 숙주 세포의 게놈 내로 검출 가능하게 통합되지 않는다. 일부 구현예에서, 게놈 내로의 유전 요소의 통합은 본 명세서에 기재된 바와 같은 기법, 예를 들어, 핵산 서열결정, PCR 검출 및/또는 핵산 혼성화를 사용하여 검출될 수 있다. 일부 구현예에서, 통합 빈도는, 예를 들어, 문헌[Wang et al., 2004, *Gene Therapy* 11: 711-721](본 명세서에 전문이 참조로 포함됨)에 기재된 바와 같이, 자유 벡터에서 분리된 게놈 DNA의 정량적 겔 정제 분석에 의해 결정된다.

[0493] 본 명세서에 사용되는 바와 같이, "실질적으로 비-면역원성인" 유기체, 입자 또는 성분은 예를 들어, 숙주 조직 또는 유기체(예를 들어, 포유동물, 예를 들어, 인간)에서 요망되지 않는 또는 비표적화된 면역 반응을 야기하거나 유도하지 않는 유기체, 입자(예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어, 바이러스 또는 아넬로벡터) 또는 이의 성분을 지칭한다. 구현예에서, 실질적으로 비-면역원성인 유기체, 입자 또는 성분은 임상적으로 유의한 면역 반응을 생성하지 않는다. 구현예에서, 실질적으로 비-면역원성인 아넬로벡터는 아넬로바이러스 또는 아넬로벡터 유전 요소의 아미노산 서열을 포함하거나, 핵산 서열에 의해 인코딩되는 단백질에 대하여 임상적으로 유의한 면역 반응을 생성하지 않는다. 구현예에서, 면역 반응(예를 들어, 요망되지 않는 또는 비표적화된 면역 반응)은 예를 들어, 문헌[Tsuda et al., 1999; *J. Virol. Methods* 77: 199-206](본 명세서에 참조로 포함됨)에 기재된 항-TTV 항체 검출 방법 및/또는 문헌[Kakkola et al., 2008; *Virology* 382: 182-189](본 명세서에 참조로 포함됨)에 기재된 항-TTV IgG 수준의 결정 방법에 따라 대상체에서 항체(예를 들어, 중화 항체) 존재 또는 수준(예를 들어, 항-아넬로벡터 항체의 존재 또는 수준, 예를 들어, 본 명세서에 기재된 바와 같은 아넬로벡터에 대한 항체의 존재 또는 수준)을 검정함으로써 검출된다. 아넬로바이러스 또는 그에 기초한 아넬로벡터에 대한 항체(예를 들어, 중화 항체)는 또한 항-바이러스 항체를 검출하기 위한 당업계의 방법, 예를 들어, 문헌[Calcedo et al., 2013; *Front. Immunol.* 4(341): 1-7](본 명세서에 참조로 포함됨)에 기재된 바와 같은 예를 들어, 항-AAV 항체의 검출 방법에 의해 검출될 수 있다.

[0494] 본 명세서에 사용되는 바와 같은 "하위서열"은 각각 더 큰 핵산 서열 또는 아미노산 서열에 포함되는 핵산 서열 또는 아미노산 서열을 지칭한다. 일부 예에서, 하위서열은 더 큰 서열의 도메인 또는 기능적 단편을 포함할 수 있다. 일부 예에서, 하위서열은 더 큰 서열로부터 단리되는 경우, 더 큰 서열의 나머지와 함께 존재할 때 하위서열에 의해 형성되는 2차 및/또는 3차 구조와 유사한 2차 및/또는 3차 구조를 형성할 수 있는 더 큰 서열의 단

편을 포함할 수 있다. 일부 예에서, 하위서열은 또 다른 서열(예를 들어, 외인성 서열 또는 더 큰 서열의 나머지에 대하여 이종인 서열을 포함하는 하위서열, 예를 들어, 상이한 아넬로바이러스 유래의 상응하는 하위서열)에 의해 대체될 수 있다.

[0495] 본 발명은 일반적으로 아넬로벡터, 예를 들어, 합성 아넬로벡터 및 이의 용도에 관한 것이다. 본 개시내용은 아넬로벡터, 아넬로벡터를 포함하는 조성물, 및 아넬로벡터의 제조 또는 이용 방법을 제공한다. 아넬로벡터는 일반적으로 예를 들어, 치료제를 진핵 세포로 운반하기 위한 운반 비히클로서 유용하다. 일반적으로, 아넬로벡터는 단백질성 외부 내에 봉입된 (예를 들어, 이펙터, 예를 들어, 외인성 이펙터 또는 내인성 이펙터를 인코딩하는) 핵산 서열을 포함하는 유전 요소를 포함할 것이다. 아넬로벡터는 (예를 들어, 본 명세서에 기재된 바와 같은) 아넬로바이러스 서열에 비한 서열(예를 들어, 본 명세서에 기재된 바와 같은 영역 또는 도메인)의 하나 이상의 결실을 포함할 수 있다. 아넬로벡터는 예를 들어, 세포를 포함하는 대상체에서 질환 또는 장애를 치료하기 위하여, 유전 요소 또는 그 내에 인코딩된 이펙터(예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어, 폴리펩티드 또는 핵산 이펙터)를 진핵 세포 내로 운반하기 위한 실질적으로 비-면역원성인 비히클로서 사용될 수 있다.

[0496] **목차**

[0497] I. 아넬로벡터를 제조하기 위한 조성물 및 방법

[0498] A. 아넬로벡터의 구성성분 및 조립

[0499] i. 아넬로벡터의 조립을 위한 ORF1 분자

[0500] ii. 아넬로벡터의 조립을 위한 ORF2 분자

[0501] B. 유전 요소 작제물

[0502] i. 플라스미드

[0503] ii. 환형 핵산 작제물

[0504] iii. 시험관 내 환화

[0505] iv. 시스/트랜스 작제물

[0506] v. 발현 카세트

[0507] vi. 유전 요소 작제물의 설계 및 생성

[0508] C. 이펙터

[0509] D. 숙주 세포

[0510] i. 숙주 세포 내로 유전 요소의 도입

[0511] ii. 시스 또는 트랜스로 아넬로바이러스 단백질(들)을 제공하는 방법

[0512] iii. 헬퍼

[0513] iv. 예시적인 세포 유형

[0514] E. 배양 조건

[0515] F. 수확

[0516] G. 농축 및 정제

[0517] II. 아넬로벡터

[0518] A. 아넬로바이러스

[0519] B. ORF1 분자

[0520] C. ORF2 분자

[0521] D. 유전 요소

[0522] E. 단백질 결합 서열

[0523] F. 5' UTR 영역

[0524] G. GC-풍부 영역

[0525] H. 이펙터

[0526] I. 조절 서열

[0527] J. 복제 단백질

[0528] K. 기타 서열

[0529] L. 단백질성 외부

[0530] III. 핵산 작제물

[0531] IV. 조성물

[0532] V. 숙주 세포

[0533] VI. 사용 방법

[0534] VII. 투여/운반

[0535] **I. 아넬로벡터를 제조하기 위한 조성물 및 방법**

[0536] 본 개시내용은, 일부 양태에서, 예를 들어, 본 명세서에 기재된 바와 같이, 아넬로벡터를 생성하는 데 사용될 수 있는 핵산 탠덤 작제물을 제공한다. 탠덤 작제물은 일반적으로, 핵산 작제물의 나머지에 연결되지 않고/않거나 환형, 단일-가닥 DNA 분자로 전환되지 않는 경우, 단백질성 외부 내에 봉입되어 아넬로벡터를 생성할 수 있는 제1 유전 요소 영역을 포함한다. 탠덤 작제물은 제2 유전 요소 영역, 또는 이의 일부를 추가로 포함할 수 있다. 일부 구현예에서, 본 명세서에 기재된 탠덤 작제물은, 예를 들어, 회전환 증폭에 의해, 단백질성 외부(예를 들어, ORF1 핵산에 의해 인코딩되는 폴리펩티드를 포함함)에 봉입하기에 적합한 유전 요소를 생성하는 데 사용될 수 있다. 일부 구현예에서, 단백질성 외부에 봉입하기에 적합한 유전 요소는 제1 유전 요소 영역의 회전환 증폭을 통해 생성된다.

[0537] 일부 구현예에서, 탠덤 작제물은 제1 카피의 유전 요소 서열(예를 들어, 유전 요소 영역) 및 제2 카피의 유전 요소 서열(예를 들어, uRFS 또는 dRFS를 포함함)의 적어도 일부를 포함하는 핵산 작제물이다. 일부 구현예에서, 제2 카피는 유전 요소의 전체 서열을 포함한다. 일부 구현예에서, 제2 카피는 유전 요소의 부분 서열(예를 들어, 유전 요소 서열의 5' 말단 또는 3' 말단으로부터, 예를 들어, 유전 요소 서열의 적어도 1%, 2%, 3%, 4%, 5%, 10%, 15%, 20%, 25%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 또는 95%)을 포함한다. 일부 구현예에서, 제1 카피의 유전 요소 서열 및 제2 카피의 유전 요소 서열은 동일한 유전 요소 서열(예를 들어, 동일한 아넬로바이러스 서열)이거나, 이로부터 유래된다. 일부 구현예에서, 제1 카피의 유전 요소 서열 및 제2 카피의 유전 요소 서열은 상이한 유전 요소 서열(예를 들어, 상이한 아넬로바이러스 유래의 서열)이거나, 이로부터 유래된다. 일부 구현예에서, 제1 카피의 유전 요소 서열 및 제2 카피의 유전 요소 서열은 핵산 작제물에서 서로 인접하여 위치한다. 다른 구현예에서, 제1 카피의 유전 요소 서열 및 제2 카피의 유전 요소 서열은, 예를 들어, 스페이서 영역에 의해 분리될 수 있다. 일부 구현예에서, 제2 카피의 유전 요소 서열 또는 이의 일부(예를 들어, uRFS를 포함함)는 제1 카피의 유전 요소 서열에 대해 5'에 위치한다. 일부 구현예에서, 제2 카피의 유전 요소 서열 또는 이의 일부(예를 들어, dRFS를 포함함)는 제1 카피의 유전 요소 서열에 대해 3'에 위치한다.

[0538] 이론에 결부시키고자 하지 않고, 회전환 증폭은 (예를 들어, 제2 유전 요소 영역에서, 예를 들어, uRFS 또는 dRFS에서) 제1 유전 요소 영역에 대해 5'에(또는 제1 유전 요소 영역의 5' 영역 내에) 위치된 Rep 결합 부위(예를 들어, 본 명세서에 기재된 바와 같이, 예를 들어, 5' UTR을 포함함, 예를 들어, 헤어핀 루프 및/또는 복제 원점을 포함함)에 대한 Rep 단백질 결합을 통해 발생할 수 있다. 그 다음 Rep 단백질은 제1 유전 요소 영역을 통해 진행되어, 유전 요소의 합성을 초래할 수 있다. 일부 구현예에서, 제2 유전 요소 영역 또는 이으 일부는 제1 유전 요소 영역에 대해 3'에 위치한다. 이론에 결부시키고자 하지 않고, Rep 단백질은 3'에 위치된 제2 유전 요소 영역에 도달 시, 예를 들어, 제2 유전 요소 영역의 Rep 결합 부위(예를 들어, 5' UTR, 헤어핀 루프, 및/또는 제2 유전 요소 서열의 복제 원점)에 도달 시 탠덤 작제물로부터 탈착되어, 합성된 유전 요소를 방출할 수 있음이 고려된다. 그 다음 방출된 유전 요소는 환화된 다음, 단백질성 외부 내에 봉입되어 아넬로벡터를 형성할

수 있다.

[0539] ***아넬로벡터의 구성성분 및 조립***

[0540] 본 명세서의 조성물 및 방법은 아넬로벡터를 생성하는 데 사용될 수 있다. 본 명세서에 기재된 바와 같이, 아넬로벡터는 일반적으로 단백질성 외부(예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어, 아넬로바이러스 ORF1 핵산에 의해 인코딩되는 폴리펩티드를 포함함) 내에 봉입된 유전 요소(예를 들어, 본 명세서에 기재된 바와 같은 5' UTR 영역을 포함하는, 예를 들어, 단일-가닥 환형 DNA 분자)를 포함한다. ORF1 핵산, 예를 들어 본 명세서에 기재된 바와 같음). 일부 구현예에서, 유전 요소는 아넬로바이러스 ORF(예를 들어, ORF1, ORF2, ORF2/2, ORF2/3, ORF1/1, 또는 ORF1/2 중 하나 이상)를 인코딩하는 하나 이상의 서열을 포함한다. 본 명세서에 사용된 바와 같이, 아넬로바이러스 ORF 또는 ORF 분자(예를 들어, 아넬로바이러스 ORF1, ORF2, ORF2/2, ORF2/3, ORF1/1, 또는 ORF1/2)는, 예를 들어 PCT/US2018/037379 또는 PCT/US19/65995(이들 각각은 본 명세서에 전문이 참조로 포함됨)에 기재된 바와 같이, 상응하는 아넬로바이러스 ORF 서열에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 아미노산 서열을 포함하는 폴리펩티드를 포함한다. 구현예에서, 유전 요소는 아넬로바이러스 ORF1, 또는 이의 스플라이스 변이체 또는 기능적 단편(예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어, 젤리-롤 영역)을 인코딩하는 서열을 포함한다. 일부 구현예에서, 단백질성 외부는 아넬로바이러스 ORF1 핵산(예를 들어, 아넬로바이러스 ORF1 분자 또는 이의 스플라이스 변이체 또는 기능적 단편)에 의해 인코딩되는 폴리펩티드를 포함한다.

[0541] 일부 구현예에서, 아넬로벡터는 단백질성 외부(예를 들어, 본 명세서에 기재된 바와 같음) 내에 유전 요소(예를 들어, 본 명세서에 기재된 바와 같음)를 봉입함으로써 조립된다. 일부 구현예에서, 유전 요소는 숙주 세포(예를 들어, 본 명세서에 기재된 바와 같음)에서 단백질성 외부 내에 봉입된다. 일부 구현예에서, 숙주 세포는 단백질성 외부에 포함된 하나 이상의 폴리펩티드(예를 들어, 아넬로바이러스 ORF1 핵산에 의해 인코딩되는 폴리펩티드, 예를 들어 ORF1 분자)를 발현한다. 예를 들어, 일부 구현예에서, 숙주 세포는 아넬로바이러스 ORF1 분자, 예를 들어, 아넬로바이러스 ORF1 폴리펩티드의 스플라이스 변이체 또는 기능적 단편(예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어, 야생형 아넬로바이러스 ORF1 단백질 또는 야생형 아넬로바이러스 ORF1 핵산에 의해 인코딩되는 폴리펩티드)을 인코딩하는 핵산 서열을 포함한다. 구현예에서, 아넬로바이러스 ORF1 분자를 인코딩하는 핵산 서열은 숙주 세포에 포함된 핵산 작제물(예를 들어, 플라스미드, 바이러스 벡터, 바이러스, 미니서클, 박미드, 또는 인공 염색체)에 포함된다. 구현예에서, 아넬로바이러스 ORF1 분자를 인코딩하는 핵산 서열은 숙주 세포의 게놈에 통합된다.

[0542] 일부 구현예에서, 숙주 세포는 유전 요소 및/또는 유전 요소의 서열을 포함하는 핵산 작제물을 포함한다. 일부 구현예에서, 핵산 작제물은 플라스미드, 바이러스 핵산, 미니서클, 박미드, 또는 인공 염색체로부터 선택된다. 일부 구현예에서, 유전 요소는 핵산 작제물로부터 절단되고, 선택적으로, 이중-가닥 형태에서 단일-가닥 형태로(예를 들어, 변성에 의해) 전환된다. 일부 구현예에서, 유전 요소는 핵산 작제물의 주형 서열에 기반한 중합효소에 의해 생성된다. 일부 구현예에서, 중합효소는 유전 요소 서열의 단일-가닥 카피를 생성하며, 이는 본 명세서에 기재된 바와 같은 유전 요소를 형성하기 위해 선택적으로 환화될 수 있다. 다른 구현예에서, 핵산 작제물은 시험관 내에서 유전 요소의 핵산 서열의 환화에 의해 생성된 이중-가닥 미니서클이다. 구현예에서, 시험관 내-환화(IVC) 미니서클은 숙주 세포 내로 도입되고, 여기서 본 명세서에 기재된 바와 같이 단백질성 외부에 봉입하기에 적합한 단일-가닥 유전 요소로 전환된다.

[0543] ***예를 들어, 아넬로벡터의 조립을 위한 ORF1 분자***

[0544] 아넬로벡터는, 예를 들어, 단백질성 외부 내에 유전 요소를 봉입함으로써 제조될 수 있다. 아넬로벡터의 단백질성 외부는 일반적으로 아넬로바이러스 ORF1 핵산(예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어, 아넬로바이러스 ORF1 분자 또는 이의 스플라이스 변이체 또는 기능적 단편)에 의해 인코딩되는 폴리펩티드를 포함한다. ORF1 분자는 일부 구현예에서 하기 중 하나 이상을 포함할 수 있다: 아르기닌-풍부 영역, 예를 들어, 적어도 60%의 염기성 잔기(예를 들어, 적어도 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 또는 100%의 염기성 잔기; 예를 들어, 60% 내지 90%, 60% 내지 80%, 70% 내지 90% 또는 70 내지 80%의 염기성 잔기)를 갖는 영역을 포함하는 제1 영역, 및 젤리-롤 도메인, 예를 들어, 적어도 6개의 베타 가닥(예를 들어, 4, 5, 6, 7, 8, 9, 10, 11 또는 12개의 베타 가닥)을 포함하는 제2 영역. 구현예에서, 단백질성 외부는 아넬로바이러스 ORF1 아르기닌-풍부 영역, 젤리-롤 영역, N22 도메인, 초가변 영역, 및/또는 C-말단 도메인 중 하나 이상(예를 들어, 1, 2, 3, 4개, 또는 5개 전부)을 포함한다. 일부 구현예에서, 단백질성 외부는 아넬로바이러스 ORF1 젤리-롤 영역(예를 들어, 본 명세서에 기재된 바와 같음)을 포함한다. 일부 구현예에서, 단백질성 외부는 아넬로바이러스 ORF1 아

르기닌-풍부 영역(예를 들어, 본 명세서에 기재된 바와 같음)을 포함한다. 일부 구현예에서, 단백질성 외부는 아넬로바이러스 ORF1 N22 도메인(예를 들어, 본 명세서에 기재된 바와 같음)을 포함한다. 일부 구현예에서, 단백질성 외부는 아넬로바이러스 초가변 영역(예를 들어, 본 명세서에 기재된 바와 같음)을 포함한다. 일부 구현예에서, 단백질성 외부는 아넬로바이러스 ORF1 C-말단 도메인(예를 들어, 본 명세서에 기재된 바와 같음)을 포함한다.

[0545] 일부 구현예에서, 아넬로벡터는 ORF1 분자 및/또는 ORF1 분자를 인코딩하는 핵산을 포함한다. 일반적으로, ORF1 분자는 아넬로바이러스 ORF1 단백질(예를 들어, 본 명세서에 기재된 바와 같은 아넬로바이러스 ORF1 단백질)의 구조적 특징 및/또는 활성을 갖는 폴리펩티드 또는 이의 기능적 단편을 포함한다. 일부 구현예에서, ORF1 분자는 아넬로바이러스 ORF1 단백질(예를 들어, 본 명세서에 기재된 바와 같은 아넬로바이러스 ORF1 단백질)에 비해 절단을 포함한다. 일부 구현예에서, ORF1 분자는 아넬로바이러스 ORF1 단백질의 적어도 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 또는 700개 아미노산만큼 절단된다. 일부 구현예에서, ORF1 분자는, 예를 들어 본 명세서에 기재된 바와 같은, 알파토르크바이러스, 베타토르크바이러스, 또는 감마토르크바이러스 ORF1 단백질에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 서열 동일성을 갖는 아미노산 서열을 포함한다. ORF1 분자는 일반적으로 핵산 분자, 예컨대, DNA(예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어, 유전 요소)에 결합할 수 있다. 일부 구현예에서, ORF1 분자는 세포의 핵에 국소화된다. 특정 구현예에서, ORF1 분자는 세포의 핵소체에 국소화된다.

[0546] 이론에 결부시키고자 하지 않고, ORF1 분자는 다른 ORF1 분자에 결합하여, 예를 들어, (예를 들어, 본 명세서에 기재된 바와 같은) 단백질성 외부를 형성할 수 있다. 이러한 ORF1 분자는 캡시드를 형성하는 능력을 갖는 것으로서 설명될 수 있다. 일부 구현예에서, 단백질성 외부는 핵산 분자(예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어, 탠덤 작제물을 사용하여 생성된, 예를 들어, 본 명세서에 기재된 바와 같은 유전 요소)를 봉입할 수 있다. 일부 구현예에서, 복수의 ORF1 분자는 다량체를 형성하여, 예를 들어, 단백질성 외부를 생성할 수 있다. 일부 구현예에서, 다량체는 동종다량체일 수 있다. 다른 구현예에서, 다량체는 이종다량체일 수 있다.

[0547] ***예를 들어, 아넬로벡터의 조립을 위한 ORF2 분자***

[0548] 본 명세서에 기재된 조성물 또는 방법을 사용하여 아넬로벡터를 생성하는 것은 아넬로바이러스 ORF2 분자(예를 들어, 본 명세서에 기재된 바와 같음), 또는 이의 스플라이스 변이체 또는 기능적 단편의 발현을 수반할 수 있다. 일부 구현예에서, 아넬로벡터는 ORF2 분자, 또는 이의 스플라이스 변이체 또는 기능적 단편, 및/또는 ORF2 분자를 인코딩하는 핵산, 또는 이의 스플라이스 변이체 또는 기능적 단편을 포함한다. 일부 구현예에서, 아넬로벡터는 ORF2 분자, 또는 이의 스플라이스 변이체 또는 기능적 단편, 및/또는 ORF2 분자를 인코딩하는 핵산, 또는 이의 스플라이스 변이체 또는 기능적 단편을 포함하지 않는다. 일부 구현예에서, 아넬로벡터를 생성하는 것은 ORF2 분자, 또는 이의 스플라이스 변이체 또는 기능적 단편의 발현을 포함하지만, ORF2 분자는 아넬로벡터에 포함되지 않는다.

[0549] ***예를 들어, 아넬로벡터의 조립을 위한 유전 요소 작제물***

[0550] 본 명세서에 기재된 바와 같은 아넬로벡터의 유전 요소는 유전 요소 영역 및 선택적으로 벡터 백본과 같은 다른 서열을 포함하는 유전 요소 작제물로부터 생성될 수 있다. 일반적으로, 유전 요소 작제물은 아넬로바이러스 5' UTR(예를 들어, 본 명세서에 기재된 바와 같음)을 포함한다. 유전 요소 작제물은 유전 요소가 단백질성 외부 내에 봉입될 수 있는 숙주 세포로 유전 요소의 서열을 운반하기에 적합한 임의의 핵산 작제물일 수 있다. 일부 구현예에서, 유전 요소 작제물은 프로모터를 포함한다. 일부 구현예에서, 유전 요소 작제물은 선형 핵산 분자이다. 일부 구현예에서, 유전 요소 작제물은 환형 핵산 분자(예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어, 플라스미드, 박미드, 또는 미니서클)이다. 일부 구현예에서, 유전 요소 작제물은 (예를 들어, 유전 요소 작제물을 포함하는 곤충 세포가 유전 요소 작제물의 유전 요소 서열, 또는 이의 단편을 포함하는 바큘로바이러스를 생성할 수 있도록) 바큘로바이러스 서열을 포함한다. 유전 요소 작제물은 일부 구현예에서 이중-가닥일 수 있다. 다른 구현예에서, 유전 요소는 단일-가닥이다. 일부 구현예에서, 유전 요소 작제물은 DNA를 포함한다. 일부 구현예에서, 유전 요소 작제물은 RNA를 포함한다. 일부 구현예에서, 유전 요소 작제물은 하나 이상의 변형된 뉴클레오티드를 포함한다.

[0551] 일부 구현예에서, 유전 요소 작제물은 유전 요소 서열의 하나의 카피를 포함한다. 일부 구현예에서, 유전 요소는 유전 요소 서열의 복수 카피(예를 들어, 유전 요소 서열의 2개 카피)를 포함한다. 일부 구현예에서, 유전은 유전 요소 서열의 하나의 전장 카피 및 적어도 하나의 부분적 유전 요소 서열을 포함한다. 일부 구현예에서, 유전 요소 서열의 2개 카피(예를 들어, 전장 및/또는 부분적인 유전 요소 서열)는 유전 요소 작제물(예를 들어,

본 명세서에 기재된 바와 같음) 내에 나란히 위치한다.

[0552] 일부 양태에서, 본 개시내용은 (예를 들어, 세포 배양 시스템에서) 본 명세서에 기재된 바와 같은 아넬로벡터의 복제 및 증식을 위한 방법을 제공하며, 이는 하기의 단계 중 하나 이상을 포함할 수 있다: (a) 유전 요소(예를 들어, 선형화됨)를 아넬로벡터 감염에 감수성인 세포주 내로 도입하는(예를 들어, 형질감염시키는) 단계; (b) 세포를 수확하고, 선택적으로 유전 요소의 존재를 나타내는 세포를 단리하는 단계; (c) 단계 (b)에서 수득되는 세포를 실험 조건 및 유전자 발현에 따라, (예를 들어, 적어도 3일, 예컨대 적어도 1주 이상 동안) 배양하는 단계; 및 (d) 예를 들어, 본 명세서에 기재된 바와 같이, 단계 (c)의 세포를 수확하는 단계.

[0553] *플라스미드*

[0554] 일부 구현예에서, 유전 요소 작제물은 플라스미드이다. 플라스미드는 일반적으로 본 명세서에 기재된 바와 같은 유전 요소의 서열뿐만 아니라 숙주 세포에서의 복제에 적합한 복제 원점(예를 들어, 박테리아 세포에서의 복제를 위한 박테리아 복제 원점) 및 선택 가능한 마커(예를 들어, 항생제 내성 유전자)를 포함한다. 일부 구현예에서, 유전 요소의 서열은 플라스미드로부터 절제될 수 있다. 일부 구현예에서, 플라스미드는 박테리아 세포에서 복제할 수 있다. 일부 구현예에서, 플라스미드는 포유동물 세포(예를 들어, 인간 세포)에서 복제할 수 있다. 일부 구현예에서, 플라스미드는 길이가 적어도 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000 또는 5000 bp이다. 일부 구현예에서, 플라스미드는 길이가 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 또는 10,000 bp 미만이다. 일부 구현예에서, 플라스미드는 300 내지 400, 400 내지 500, 500 내지 600, 600 내지 700, 700 내지 800, 800 내지 900, 900 내지 1000, 1000 내지 1500, 1500 내지 2000, 2000 내지 2500, 2500 내지 3000, 3000 내지 4000, 또는 4000 내지 5000 bp의 길이를 갖는다. 일부 구현예에서, 유전 요소는, 예를 들어, (예를 들어, 시험관 내 환화에 의해) 플라스미드로부터 절제되어, 예를 들어, 본 명세서에 기재된 바와 같이, 미니서클을 형성할 수 있다. 구현예에서, 유전 요소의 절제는 유전 요소 서열을 플라스미드 백본으로부터 분리한다(예를 들어, 유전 요소를 박테리아 백본으로부터 분리함).

[0555] *작은 환형 핵산 작제물*

[0556] 일부 구현예에서, 유전 요소 작제물은, 예를 들어 백본이 결여된(예를 들어, 박테리아 복제 원점 및/또는 선택 가능한 마커가 결여된) 환형 핵산 작제물이다. 구현예에서, 유전 요소는 이중-가닥 환형 핵산 작제물이다. 구현예에서, 이중-가닥 환형 핵산 작제물은, 예를 들어, 본 명세서에 기재된 바와 같이, 시험관 내 환화(IVC)에 의해 생성된다. 구현예에서, 이중-가닥 환형 핵산 작제물은 숙주 세포 내로 도입될 수 있으며, 여기서, 예를 들어 본 명세서에 기재된 바와 같이, 단일-가닥 환형 유전 요소를 생성하기 위한 주형으로 전환되거나 주형으로서 사용될 수 있다. 일부 구현예에서, 환형 핵산 작제물은 플라스미드 백본 또는 이의 기능적 단편을 포함하지 않는다. 일부 구현예에서, 환형 핵산 작제물은 길이가 적어도 2000, 2100, 2200, 2300, 2400, 2500, 2600, 2700, 2800, 2900, 3000, 3100, 3200, 3300, 3400, 3500, 3600, 3700, 3800, 3900, 4000, 4100, 4200, 4300, 4400, 또는 4500 bp이다. 일부 구현예에서, 환형 핵산 작제물은 길이가 2900, 3000, 3100, 3200, 3300, 3400, 3500, 3600, 3700, 3800, 3900, 4000, 4100, 4200, 4300, 4400, 4500, 4600, 4700, 4800, 4900, 5000, 5500, 또는 6000 bp 미만이다. 일부 구현예에서, 환형 핵산 작제물은 길이가 2000 내지 2100, 2100 내지 2200, 2200 내지 2300, 2300 내지 2400, 2400 내지 2500, 2500 내지 2600, 2600 내지 2700, 2700 내지 2800, 2800 내지 2900, 2900 내지 3000, 3000 내지 3100, 3100 내지 3200, 3200 내지 3300, 3300 내지 3400, 3400 내지 3500, 3500 내지 3600, 3600 내지 3700, 3700 내지 3800, 3800 내지 3900, 3900 내지 4000, 4000 내지 4100, 4100 내지 4200, 4200 내지 4300, 4300 내지 4400, 또는 4400 내지 4500 bp이다. 일부 구현예에서, 환형 핵산 작제물은 미니서클이다.

[0557] *시험관 내 환화*

[0558] 일부 예에서, 단백질성 외부로 패키징되는 유전 요소는 단일 가닥 환형 DNA이다. 유전 요소는 일부 예에서 단일 가닥 환형 DNA 이외의 형태를 갖는 유전 요소 작제물을 통해 숙주 세포로 도입될 수 있다. 예를 들어, 유전 요소 작제물은 이중-가닥 환형 DNA일 수 있다. 그 다음 이중-가닥 환형 DNA는 숙주 세포(예를 들어, 문헌 [Wawrzyniak et al. 2017, *Front. Microbiol.* 8: 2353](열거된 효소와 관련하여 본 명세서에 참조로 포함됨)에 기재된 바와 같은, 예를 들어, 회전환 복제에 적합한 효소, 예를 들어, 아넬로바이러스 Rep 단백질, 예를 들어, Rep68/78, Rep60, RepA, RepB, Pre, MobM, TraX, TrwC, Mob02281, Mob02282, NikB, ORF50240, NikK, TecH, OrfJ, 또는 TraI를 포함하는 숙주 세포)에서 단일-가닥 환형 DNA로 전환될 수 있다. 일부 구현예에서, 이중-가닥 환형 DNA는, 예를 들어, 실시예 20에 기재된 바와 같이, 시험관 내 환화(IVC)에 의해 생성된다.

[0559] 일반적으로, 시험관 내 환화된 DNA 작제물은 패키징될 유전 요소의 서열을 포함하는 플라스미드를 유전 요소 서열이 선형 DNA 분자로서 절제되도록 분해함으로써 생성될 수 있다. 그 다음, 생성된 선형 DNA는, 예를 들어, DNA 리가제를 사용하여 결찰되어, 이중-가닥 환형 DNA를 형성할 수 있다. 일부 예에서, 시험관 내 환화에 의해 생성되는 이중-가닥 환형 DNA는, 예를 들어, 본 명세서에 기재된 바와 같이, 회전환 복제를 겪을 수 있다. 이론에 결부시키고자 하지 않고, 시험관 내 환화가 추가의 변형 없이 회전환 복제를 겪을 수 있는 이중-가닥 DNA 작제물을 생성함으로써, 예를 들어, 본 명세서에 기재된 바와 같이, 아넬로벡터 내로 패키징되기에 적합한 크기의 단일-가닥 환형 DNA를 생성할 수 있는 것으로 고려된다. 일부 구현예에서, 이중-가닥 DNA 작제물은 플라스미드(예를 들어, 박테리아 플라스미드)보다 더 작다. 일부 구현예에서, 이중-가닥 DNA 작제물은 플라스미드(예를 들어, 박테리아 플라스미드)로부터 절제된 다음, 예를 들어, 시험관 내 환화에 의해 환화된다.

[0560] *시스/트랜스 작제물*

[0561] 일부 구현예에서, 본 명세서에 기재된 바와 같은 유전 요소 작제물은 하나 이상의 아넬로바이러스 ORF, 예를 들어, 단백질성 외부 구성성분(예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어, 아넬로바이러스 ORF1 핵산에 의해 인코딩되는 폴리펩티드)을 인코딩하는 하나 이상의 서열을 포함한다. 예를 들어, 유전 요소 작제물은 아넬로바이러스 ORF1 분자를 인코딩하는 핵산 서열을 포함할 수 있다. 이러한 유전 요소 작제물은 유전 요소 및 아넬로바이러스 ORF(들)를 시스로 숙주 세포로 도입하는 데 적합할 수 있다. 다른 구현예에서, 본 명세서에 기재된 바와 같은 유전 요소 작제물은 하나 이상의 아넬로바이러스 ORF, 예를 들어, 단백질성 외부 구성성분(예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어, 아넬로바이러스 ORF1 핵산에 의해 인코딩된 폴리펩티드)을 인코딩하는 서열을 포함하지 않는다. 예를 들어, 유전 요소 작제물은 아넬로바이러스 ORF1 분자를 인코딩하는 핵산 서열을 포함하지 않을 수 있다. 이러한 유전 요소 작제물은 (예를 들어, 하나 이상의 아넬로바이러스 ORF를 인코딩하는 제2 핵산 작제물의 도입을 통해, 또는 숙주 세포의 게놈에 통합된 아넬로바이러스 ORF 카세트를 통해) 트랜스로 제공되는 하나 이상의 아넬로바이러스 ORF와 함께, 유전 요소를 숙주 세포로 도입하는 데 적합할 수 있다.

[0562] 일부 구현예에서, 유전 요소 작제물은 아넬로바이러스 ORF1 분자, 또는 이의 스플라이스 변이체 또는 기능적 단편(예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어, 젤리-롤 영역)을 인코딩하는 서열을 포함한다. 구현예에서, 유전 요소의 서열을 포함하지 않는 유전 요소의 부분은 (예를 들어, 프로모터 및 아넬로바이러스 ORF1 분자, 또는 이의 스플라이스 변이체 또는 기능적 단편을 인코딩하는 서열을 포함하는 카세트에) 아넬로바이러스 ORF1 분자, 또는 이의 스플라이스 변이체 또는 기능적 단편을 인코딩하는 서열을 포함한다. 추가 구현예에서, 유전 요소 서열을 포함하는 작제물의 일부는 아넬로바이러스 ORF1 분자, 또는 이의 스플라이스 변이체 또는 기능적 단편(예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어, 젤리-롤 영역)을 인코딩하는 서열을 포함한다. 구현예에서, 단백질성 외부(예를 들어, 본 명세서에 기재된 바와 같음)에 이러한 유전 요소의 봉입은 복제-구성성분 아넬로벡터(예를 들어, 세포 감염 시, 예를 들어, 본 명세서에 기재된 바와 같은 하나 이상의 아넬로바이러스 ORF를 인코딩하는, 추가 핵산 작제물을 세포에 도입하지 않고 세포가 아넬로벡터의 추가 카피를 생성할 수 있게 하는 아넬로벡터)를 생성한다.

[0563] 다른 구현예에서, 유전 요소는 아넬로바이러스 ORF1 분자, 또는 이의 스플라이스 변이체 또는 기능적 단편(예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어, 젤리-롤 영역)을 인코딩하는 서열을 포함하지 않는다. 구현예에서, 단백질성 외부(예를 들어, 본 명세서에 기재된 바와 같음)에 이러한 유전 요소의 봉입은 복제-부적격 아넬로벡터(예를 들어, 세포 감염 시, 예를 들어, 본 명세서에 기재된 바와 같은 하나 이상의 아넬로바이러스 ORF를 인코딩하는 하나 이상의 추가 작제물의 부재 시, 감염된 세포가 추가 아넬로벡터를 생성할 수 없게 하는 아넬로벡터)를 생성한다.

[0564] *발현 카세트*

[0565] 일부 구현예에서, 유전 요소 작제물은 폴리펩티드 또는 비코딩 RNA(예를 들어, miRNA 또는 siRNA)의 발현을 위한 하나 이상의 카세트를 포함한다. 일부 구현예에서, 유전 요소 작제물은 이펙터(예를 들어, 외인성 또는 내인성 이펙터), 예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어, 폴리펩티드 또는 비코딩 RNA의 발현을 위한 카세트를 포함한다. 일부 구현예에서, 유전 요소 작제물은 아넬로바이러스 단백질(예를 들어, 아넬로바이러스 ORF1, ORF2, ORF2/2, ORF2/3, ORF1/1, 또는 ORF1/2, 또는 이의 기능적 단편)의 발현을 위한 카세트를 포함한다. 발현 카세트는 일부 구현예에서 유전 요소 서열 내에 위치할 수 있다. 구현예에서, 이펙터에 대한 발현 카세트는 유전 요소 서열 내에 위치한다. 구현예에서, 아넬로바이러스 단백질에 대한 발현 카세트는 유전 요소 서열 내에 위치한다. 다른 구현예에서, 발현 카세트는 유전 요소의 서열 외부에 있는 유전 요소 작제물 내의

위치(예를 들어, 백본 내)에 위치한다. 구현예에서, 아넬로바이러스 단백질에 대한 발현 카세트는 유전 요소의 서열 외부에 있는 유전 요소 작제물 내의 위치(예를 들어, 백본)에 위치한다.

[0566] 폴리펩티드 발현 카세트는 일반적으로 프로모터 및 폴리펩티드, 예를 들어 이펙터(예를 들어, 본 명세서에 기재된 바와 같은 외인성 또는 내인성 이펙터) 또는 아넬로바이러스 단백질을 인코딩하는 코딩 서열(예를 들어, 아넬로바이러스 ORF1, ORF2, ORF2/2, ORF2/3, ORF1/1, 또는 ORF1/2, 또는 이의 기능적 단편을 인코딩하는 서열)을 포함한다. (예를 들어, 폴리펩티드의 발현을 구동하기 위해) 폴리펩티드 발현 카세트에 포함될 수 있는 예시적인 프로모터는 제한 없이, 예를 들어, 본 명세서에 기재된 바와 같은, 구성적 프로모터(예를 들어, CMV, RSV, PGK, EF1a, 또는 SV40), 세포 또는 조직-특이적 프로모터(예를 들어, 골격 α-액틴 프로모터, 미오신 경쇄 2A 프로모터, 디스트로핀 프로모터, 근육 크레아틴 키나제 프로모터, 간 알부민 프로모터, B형 간염 바이러스 코어 프로모터, 오스테오칼신 프로모터, 골 시알로단백질 프로모터, CD2 프로모터, 면역글로불린 중쇄 프로모터, T 세포 수용체 a 쇄 프로모터, 뉴런-특이적 에놀라제(NSE) 프로모터, 또는 신경필라멘트 경쇄 프로모터), 및 유도성 프로모터(예를 들어, 아연-유도성 양 메탈로티오닌(MT) 프로모터; 덱사메타손(Dex)-유도성 마우스 유방 종양 바이러스(MMTV) 프로모터; T7 중합효소 프로모터 시스템, 테트라사이클린-억제성 시스템, 테트라사이클린-유도성 시스템, RU486-유도성 시스템, 라파마이신-유도성 시스템)를 포함한다. 일부 구현예에서, 발현 카세트는, 예를 들어, 본 명세서에 기재된 바와 같은, 인핸서를 추가로 포함한다.

[0567] *유전 요소 작제물의 설계 및 생성*

[0568] 유전 요소 작제물을 합성하기 위해 다양한 방법이 이용 가능하다. 예를 들어, 유전 요소 작제물 서열은 합성하기에 더 쉬운 더 작은 중첩하는 조각(예를 들어, 약 100 bp 내지 약 10 kb 범위의 세그먼트 또는 개별 ORF)으로 나뉠 수 있다. 이들 DNA 세그먼트는 중첩하는 단일-가닥 올리고뉴클레오티드의 세트로부터 합성된다. 그 다음 생성되는 중첩 신톤(synthon)은 더 큰 조각의 DNA, 예를 들어, 유전 요소 작제물로 조립된다. 세그먼트 또는 ORF는 유전 요소 작제물, 예를 들어, 5' 및 3' 말단에서 시험관 내 재조합 또는 특유한 제한 부위 내로 조립되어, 결찰을 가능하게 할 수 있다.

[0569] 유전 요소 작제물은 작제물 서열을 올리고-길이 단편으로 분석하는 설계 알고리즘으로 합성하여, 서열 공간의 복잡성을 고려하는 합성을 위한 적합한 설계 조건을 생성할 수 있다. 그 다음 올리고는 반도체-기반의 고밀도 칩 상에서 화학적으로 합성되며, 칩마다 200,000개 초과의 개별 올리고가 합성된다. 올리고를 조립 기법, 예컨대, BioFab®으로 조립하여, 더 작은 올리고로부터 더 긴 DNA 세그먼트를 구축한다. 이는 병렬 방식으로 행해지며, 따라서 수백 내지 수천개의 합성 DNA 세그먼트가 한번에 구축된다.

[0570] 각각의 유전 요소 작제물 또는 유전 요소 작제물의 세그먼트는 검증된 서열일 수 있다. 일부 구현예에서, RNA 또는 DNA의 고-처리량 서열결정은 AnyDot.chips(Genovoxx, 독일)를 사용하여 일어날 수 있으며, 이는 생물학적 과정(예를 들어, miRNA 발현 또는 대립형질 가변성(SNP 검출))의 모니터링을 가능하게 한다. 다른 고-처리량 서열결정 시스템은 문헌[Venter, J., et al. Science 16 Feb. 2001]; 문헌[Adams, M. et al, Science 24 Mar. 2000]; 및 문헌[M. J, Levene, et al. Science 299:682-686, January 2003]; 및 미국 공개 출원 20030044781 및 2006/0078937에 개시된 것을 포함한다. 전체적으로 이러한 시스템은 핵산 분자에서 측정되는 중합 반응을 통해 염기를 일시적으로 첨가함으로써 복수의 염기를 갖는 표적 핵산 분자를 서열결정하는 것을 수반하며, 즉, 서열결정될 주형 핵산 분자 상의 핵산 중합 효소의 활성이 실시간으로 추적된다. 일부 구현예에서, 샷건(shotgun) 서열결정이 수행된다.

[0571] 유전 요소 작제물은 복제 또는 패키징을 위한 인자가 유전 요소에 비하여 시스 또는 트랜스로 공급될 수 있도록 설계될 수 있다. 예를 들어, 시스로 공급되는 경우, 유전 요소는, 예를 들어, 본 명세서에 기재된 바와 같은, 아넬로바이러스 ORF1, ORF1/1, ORF1/2, ORF2, ORF2/2, ORF2/3, 또는 ORF2t/3을 인코딩하는 하나 이상의 유전자를 포함할 수 있다. 일부 구현예에서, 복제 및/또는 패키징 신호는, 예를 들어, 증폭 및/또는 캡슐화를 유도하기 위하여 유전 요소 내로 혼입될 수 있다. 일부 구현예에서, 이펙터는 게놈 내의 특정 부위 내로 삽입된다. 일부 구현예에서, 하나 이상의 바이러스 ORF는 이펙터로 대체된다.

[0572] 또 다른 예에서, 복제 또는 패키징 인자가 트랜스로 공급되는 경우, 유전 요소는, 예를 들어, 본 명세서에 기재된 바와 같은, 아넬로바이러스 ORF1, ORF1/1, ORF1/2, ORF2, ORF2/2, ORF2/3, 또는 ORF2t/3 중 하나 이상을 인코딩하는 유전자가 결여될 수 있으며; 이 단백질 또는 단백질들은, 예를 들어, 또 다른 핵산, 예를 들어, 헬퍼 핵산에 의해 공급될 수 있다. 일부 구현예에서, 최소 시스 신호(예를 들어, 5' UTR 및/또는 GC-풍부 영역)가 유전 요소 내에 존재한다. 일부 구현예에서, 유전 요소는 복제 또는 패키징 인자(예를 들어, 복제효소 및/또는 캡시드 단백질)를 인코딩하지 않는다. 이러한 인자는 일부 구현예에서, 하나 이상의 헬퍼 핵산(예를 들어, 숙주

세포 게놈 내로 통합되는 헬퍼 바이러스 핵산, 헬퍼 플라스미드 또는 헬퍼 핵산)에 의해 공급될 수 있다. 일부 구현예에서, 헬퍼 핵산은 증폭 및/또는 패키징을 유도하기에 충분한 단백질 및/또는 RNA를 발현하지만, 이들 자체의 패키징 신호가 결여될 수 있다. 일부 구현예에서, 유전 요소 및 헬퍼 핵산은 (예를 들어, 동시에 또는 개별적으로) 숙주 세포 내로 도입되어, 유전 요소의 증폭 및/또는 패키징을 초래하지만, 헬퍼 핵산의 증폭 및/또는 패키징을 초래하지 않는다.

[0573] 일부 구현예에서, 유전 요소 작제물은 컴퓨터-지원 설계 도구를 사용하여 설계될 수 있다.

[0574] 작제물을 제조하는 일반적인 방법은, 예를 들어, 문헌[Khudyakov & Fields, *Artificial DNA: Methods and Applications*, CRC Press (2002)]; 문헌[Zhao, *Synthetic Biology: Tools and Applications*, (First Edition), Academic Press (2013)]; 및 문헌[Egli & Herdewijn, *Chemistry and Biology of Artificial Nucleic Acids*, (First Edition), Wiley-VCH (2012)]에 기재되어 있다.

[0575] ***이펙터***

[0576] 본 명세서에 기재된 조성물 및 방법은, 예를 들어, 본 명세서에 기재된 바와 같은, 이펙터(예를 들어, 외인성 이펙터 또는 내인성 이펙터)를 인코딩하는 서열을 포함하는 아넬로벡터의 유전 요소를 생성하는 데 사용될 수 있다. 이펙터는 일부 예에서, 내인성 이펙터 또는 외인성 이펙터일 수 있다. 일부 구현예에서, 이펙터는 치료용 이펙터이다. 일부 구현예에서, 이펙터는 폴리펩티드(예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어, 치료용 폴리펩티드 또는 펩티드)를 포함한다. 일부 구현예에서, 이펙터는 비코딩 RNA(예를 들어, miRNA, siRNA, shRNA, mRNA, lncRNA, RNA, DNA, 안티센스 RNA, 또는 gRNA)를 포함한다. 일부 구현예에서, 이펙터는, 예를 들어, 본 명세서에 기재된 바와 같은, 조절 핵산을 포함한다.

[0577] 일부 구현예에서, 이펙터-인코딩 서열은, TATA 박스 상류의 5' 비코딩 영역, 5' UTR, 폴리-A 신호의 하류, 또는 GC-풍부 영역의 상류의 3' 비코딩 영역에서, 예를 들어, 비코딩 영역, 예를 들어, 오픈-리딩 프레임의 3' 및 유전 요소의 GC-풍부 영역의 5'에 배치된 비코딩 영역에서, 유전 요소에 삽입될 수 있다. 일부 구현예에서, 이펙터-인코딩 서열은, 예를 들어, 코딩 서열(예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어, 아넬로바이러스 ORF1, ORF1/1, ORF1/2, ORF2, ORF2/2, ORF2/3, 및/또는 ORF2t/3을 인코딩하는 서열)에서, 유전 요소에 삽입될 수 있다. 일부 구현예에서, 이펙터-인코딩 서열은 오픈-리딩 프레임의 전부 또는 일부를 대체한다. 일부 구현예에서, 유전 요소는 이펙터-인코딩 서열에 작동 가능하게 연결된 조절 서열(예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어, 프로모터 또는 인핸서)을 포함한다.

[0578] 일부 구현예에서, 이펙터를 포함하는 유전 요소는, 예를 들어, 유전 요소 작제물 상에 배치된 유전 요소 서열의 회전환 증폭에 의해, 본 명세서에 기재된 바와 같은 유전 요소 작제물(예를 들어, 탠덤 작제물)로부터 생성된다. 일부 구현예에서, 탠덤 작제물은 이펙터-인코딩 서열의 정확히 하나의 카피를 포함한다. 일부 구현예에서, 탠덤 작제물은 이펙터-인코딩 서열의 2개 이상(예를 들어, 2, 3, 4, 5, 6, 7, 8, 9, 10개 이상)의 카피를 포함한다. 일부 구현예에서, 탠덤 작제물은 이펙터-인코딩 서열의 하나의 전장 카피 및 이펙터-인코딩 서열의 적어도 하나(예를 들어, 1, 2, 3, 4, 5, 6, 7, 8, 9개 이상)의 부분 카피(예를 들어, 이펙터-인코딩 서열의 5' 절단 또는 3' 절단을 포함하는 부분 카피)를 포함한다.

[0579] ***숙주 세포***

[0580] 본 명세서에 기재된 아넬로벡터는, 예를 들어, 숙주 세포에서 생성될 수 있다. 일반적으로, 아넬로벡터 유전 요소 및 아넬로벡터 단백질성 외부(예를 들어, 아넬로바이러스 ORF1 핵산 또는 아넬로바이러스 ORF1 분자에 의해 인코딩되는 폴리펩티드)의 구성성분을 포함하는 숙주 세포가 제공된다. 그 다음 숙주 세포는 단백질성 외부 내에 유전 요소를 봉입하기에 적합한 조건(예를 들어, 본 명세서에 기재된 바와 같은 배양 조건) 하에서 인큐베이션된다. 일부 구현예에서, 숙주 세포는 아넬로벡터를 숙주 세포로부터, 예를 들어 주변 상청액으로 방출하기에 적합한 조건 하에서 추가로 인큐베이션된다. 일부 구현예에서, 숙주 세포는 세포 용해물로부터 아넬로벡터의 수확을 위해 용해된다. 일부 구현예에서, 아넬로벡터는 높은 세포 밀도로 성장한 숙주 세포주에 도입될 수 있다.

[0581] *숙주 세포 내로의 유전 요소의 도입*

[0582] 유전 요소, 또는 핵산 작제물은 유전 요소의 서열을 포함하고, 숙주 세포 내로 도입될 수 있다. 일부 구현예에서, 유전 요소 자체가 숙주 세포 내로 도입된다. 일부 구현예에서, 유전 요소(예를 들어, 본 명세서에 기재된 바와 같음)의 서열을 포함하는 유전 요소 작제물이 숙주 세포 내로 도입된다. 유전 요소 또는 유전 요소 작제물은, 예를 들어 당업계에 알려진 방법을 사용하여 숙주 세포 내로 도입될 수 있다. 예를 들어, 유전 요소 또는 유전 요소 작제물은 형질감염(예를 들어, 안정한 형질감염 또는 일시적 형질감염)에 의해 숙주 세포 내로 도입

될 수 있다. 구현예에서, 유전 요소 또는 유전 요소 작제물은 lipofectamine 형질감염에 의해 숙주 세포 내로 도입된다. 구현예에서, 유전 요소 또는 유전 요소 작제물은 칼슘 포스페이트 형질감염에 의해 숙주 세포 내로 도입된다. 일부 구현예에서, 유전 요소 또는 유전 요소 작제물은 전기천공에 의해 숙주 세포 내로 도입된다. 일부 구현예에서, 유전 요소 또는 유전 요소 작제물은 유전자 총을 사용하여 숙주 세포 내로 도입된다. 일부 구현예에서, 유전 요소 또는 유전 요소 작제물은 뉴클레오펙션에 의해 숙주 세포 내로 도입된다. 일부 구현예에서, 유전 요소 또는 유전 요소 작제물은 PEI 형질감염에 의해 숙주 세포 내로 도입된다. 일부 구현예에서, 유전 요소는 숙주 세포를 유전 요소를 포함하는 아넬로벡터와 접촉시킴으로써 숙주 세포 내로 도입된다.

[0583] 구현예에서, 유전 요소 작제물은 일단 숙주 세포 내로 도입되면 복제할 수 있다. 구현예에서, 유전 요소는 일단 숙주 세포 내로 도입되면 유전 요소 작제물로부터 생성될 수 있다. 일부 구현예에서, 유전 요소는, 예를 들어, 주형으로서 유전 요소 작제물을 사용하여, 중합효소에 의해 숙주 세포에서 생성된다.

[0584] 일부 구현예에서, 유전 요소 또는 유전 요소를 포함하는 벡터는 아넬로벡터의 발현을 달성하기 위해 바이러스 중합효소 단백질을 발현하는 세포주 내로 도입(예를 들어, 형질감염)된다. 이를 위해, 아넬로벡터 중합효소 단백질을 발현하는 세포주가 적절한 숙주 세포로 이용될 수 있다. 숙주 세포는 다른 바이러스 기능 또는 추가 기능을 제공하도록 유사하게 엔지니어링될 수 있다.

[0585] 본 명세서에 개시된 아넬로벡터를 제조하기 위해, 유전 요소 작제물이 복제 및 생성에 필요한 아넬로벡터 단백질 및 기능을 제공하는 세포를 형질감염시키는 데 사용될 수 있다. 대안적으로, 세포는 본 명세서에 개시된 유전 요소 또는 유전 요소를 포함하는 벡터에 의한 형질감염 전, 도중 또는 후에 아넬로벡터 단백질 및 기능을 제공하는 제2 작제물(예를 들어, 바이러스)로 형질감염될 수 있다. 일부 구현예에서, 제2 작제물은 불완전한 바이러스 입자의 생성을 보완하는 데 유용할 수 있다. 제2 작제물(예를 들어, 바이러스)은 숙주 범위 제한 또는 온도 민감성과 같은 조건부 성장 결함을 가질 수 있으며, 이는 예를 들어 형질감염 바이러스의 후속 선택을 가능하게 한다. 일부 구현예에서, 제2 작제물은 아넬로벡터의 발현을 달성하기 위해 숙주 세포에 의해 이용되는 하나 이상의 복제 단백질을 제공할 수 있다. 일부 구현예에서, 숙주 세포는 하나 이상의 복제 단백질과 같은 바이러스 단백질을 인코딩하는 벡터로 형질감염될 수 있다. 일부 구현예에서, 제2 작제물은 항바이러스 민감성을 포함한다.

[0586] 본 명세서에 개시된 유전 요소 또는 유전 요소를 포함하는 벡터는 일부 예에서 당업계에 알려진 기법을 사용하여 아넬로벡터로 복제 및 생성될 수 있다. 예를 들어, 다양한 바이러스 배양 방법이, 예를 들어, 미국 특허 4,650,764; 미국 특허 5,166,057; 미국 특허 5,854,037; 유럽 특허 공개 EP 0702085A1; 미국 특허 출원 일련번호 09/152,845; 국제 특허 공개 PCT WO97/12032; WO96/34625; 유럽 특허 공개 EP-A780475; WO 99/02657; WO 98/53078; WO 98/02530; WO 99/15672; WO 98/13501; WO 97/06270; 및 EPO 780 47SA1에 기재되어 있으며, 이들 각각은 본 명세서에 전문이 참조로 포함된다.

[0587] *시스 또는 트랜스로 아넬로바이러스 단백질(들)을 제공하는 방법*

[0588] 일부 구현예(예를 들어, 본 명세서에 기재된 시스 구현예)에서, 유전 요소 작제물은 아넬로바이러스 ORF(예를 들어, 아넬로바이러스 ORF1, ORF2, ORF2/2, ORF2/3, ORF1/1, 또는 ORF1/2, 또는 이의 기능적 단편)에 대한 코딩 서열을 포함하는 하나 이상의 발현 카세트를 추가로 포함한다. 구현예에서, 유전 요소 작제물은 아넬로바이러스 ORF1, 또는 이의 스플라이스 변이체 또는 기능적 단편에 대한 코딩 서열을 포함하는 발현 카세트를 포함한다. 이펙터뿐만 아니라 하나 이상의 아넬로바이러스 ORF에 대한 발현 카세트를 포함하는 이러한 유전 요소 작제물은, 숙주 세포 내로 도입될 수 있다. 이러한 유전 요소 작제물을 포함하는 숙주 세포는, 일부 예에서, 추가의 핵산 작제물 또는 발현 카세트의 숙주 세포 게놈 내로의 통합을 요구하지 않으면서, 단백질성 외부에 대한, 그리고 단백질성 외부 내의 유전 요소의 봉입을 위한 유전 요소 및 구성성분을 생성할 수 있다. 다시 말하면, 이러한 유전 요소 작제물은, 예를 들어, 본 명세서에 기재된 바와 같이, 숙주 세포에서 시스 아넬로벡터 생성 방법에 사용될 수 있다.

[0589] 일부 구현예(예를 들어, 본 명세서에 기재된 트랜스 구현예)에서, 유전 요소는 하나 이상의 아넬로바이러스 ORF(예를 들어, 아넬로바이러스 ORF1, ORF2, ORF2/2, ORF2/3, ORF1/1, 또는 ORF1/2, 또는 이의 기능적 단편)에 코딩 서열을 포함하는 발현 카세트를 포함하지 않는다. 구현예에서, 유전 요소 작제물은 아넬로바이러스 ORF1, 또는 이의 스플라이스 변이체 또는 기능적 단편에 대한 코딩 서열을 포함하는 발현 카세트를 포함하지 않는다. 이펙터에 대한 발현 카세트를 포함하지만 하나 이상의 아넬로바이러스 ORF(예를 들어, 아넬로바이러스 ORF1 또는 이의 스플라이스 변이체 또는 기능적 단편)에 대한 발현 카세트가 결여된 이러한 유전 요소 작제물은 숙주 세포 내로 도입될 수 있다. 이러한 유전 요소 작제물을 포함하는 숙주 세포는, 일부 예에서, 아넬로벡터의 하나

이상의 구성성분(예를 들어, 단백질성 외부 단백질)의 생성을 위해 추가의 핵산 작제물 또는 발현 카세트의 숙주 세포 게놈으로의 통합을 필요로 할 수 있다. 일부 구현예에서, 이러한 유전 요소 작제물을 포함하는 숙주 세포는 아넬로바이러스 ORF1 분자를 인코딩하는 추가의 핵산 작제물의 부재시 단백질성 외부 내에 유전 요소를 봉입할 수 없다. 다시 말하면, 이러한 유전 요소 작제물은, 예를 들어, 본 명세서에 기재된 바와 같이, 숙주 세포에서 트랜스 아넬로벡터 생성 방법에 사용될 수 있다.

[0590] *헬퍼*

[0591] 일부 구현예에서, 헬퍼 작제물은 숙주 세포(예를 들어, 본 명세서에 기재된 바와 같은 유전 요소 작제물 또는 유전 요소를 포함하는 숙주 세포) 내로 도입된다. 일부 구현예에서, 헬퍼 작제물은 유전 요소 작제물의 도입 이전에 숙주 세포 내로 도입된다. 일부 구현예에서, 헬퍼 작제물은 유전 요소 작제물의 도입과 동시에 숙주 세포 내로 도입된다. 일부 구현예에서, 헬퍼 작제물은 유전 요소 작제물의 도입 후에 숙주 세포 내로 도입된다.

[0592] *예시적인 세포 유형*

[0593] 아넬로벡터의 생성에 적합한 예시적인 숙주 세포는 제한 없이, 포유동물 세포 및 곤충 세포를 포함한다. 일부 구현예에서, 숙주 세포는 인간 세포 또는 세포주이다. 일부 구현예에서, 세포는 면역 세포 또는 세포주, 예를 들어, T 세포 또는 세포주, 암 세포주, 간 세포 또는 세포주, 뉴런, 신경아교세포, 피부 세포, 상피 세포, 중간엽 세포, 혈액 세포, 내피 세포, 눈 세포, 위장 세포, 전구 세포, 전구체 세포, 줄기 세포, 폐 세포, 심장 세포, 또는 근육 세포이다. 일부 구현예에서, 숙주 세포는 동물 세포(예를 들어, 마우스 세포, 래트 세포, 토끼 세포, 햄스터 세포, 또는 곤충 세포)이다.

[0594] 일부 구현예에서, 숙주 세포는 림프 세포이다. 일부 구현예에서, 숙주 세포는 T 세포 또는 무한증식 세포이다. 구현예에서, 숙주 세포는 Jurkat 세포이다. 구현예에서, 숙주 세포는 MOLT 세포(예를 들어, MOLT-4 또는 MOLT-3 세포)이다. 구현예에서, 숙주 세포는 MOLT-4 세포이다. 구현예에서, 숙주 세포는 MOLT-3 세포이다. 일부 구현예에서, 숙주 세포는 급성 림프구성 백혈병(ALL) 세포, 예를 들어 MOLT 세포, 예를 들어 MOLT-4 또는 MOLT-3 세포이다. 일부 구현예에서, 숙주 세포는 B 세포 또는 무한증식 B 세포이다. 일부 구현예에서, 숙주 세포는 유전 요소 작제물, 예를 들어 탠덤 작제물(예를 들어, 본 명세서에 기재된 바와 같음)을 포함한다.

[0595] 일부 구현예에서, 숙주 세포는 MOLT 세포(예를 들어, MOLT-4 또는 MOLT-3 세포)이다.

[0596] 일부 구현예에서, 숙주 세포는 급성 림프구성 백혈병(ALL) 세포, 예를 들어 MOLT 세포, 예를 들어 MOLT-4 또는 MOLT-3 세포이다.

[0597] 일 양태에서, 본 개시내용은 단백질성 외부에 봉입되는 유전 요소를 포함하는 아넬로벡터를 제조하는 방법을 제공하며, 방법은 아넬로벡터 유전 요소를 포함하는 MOLT-4 세포를 제공하는 단계, 및 아넬로벡터 유전 요소가 MOLT-4 세포의 단백질성 외부에 봉입될 수 있도록 하는 조건 하에서 MOLT-4 세포를 인큐베이션하는 단계를 포함한다. 일부 구현예에서, MOLT-4 세포는 단백질성 외부의 일부 또는 전부를 형성하는 하나 이상의 아넬로바이러스 단백질(예를 들어, 아넬로바이러스 ORF1 분자)을 추가로 포함한다. 일부 구현예에서, 아넬로벡터 유전 요소는 MOLT-4 세포에서, 예를 들어 유전 요소 작제물(예를 들어, 본 명세서에 기재된 바와 같음)로부터 생성된다. 일부 구현예에서, 유전 요소 작제물은 탠덤 작제물(예를 들어, 본 명세서에 기재된 바와 같음)이다. 일부 구현예에서, 유전 요소 작제물은 탠덤 작제물(예를 들어, 본 명세서에 기재된 바와 같음)이 아니다. 일부 구현예에서, 방법은 MOLT-4 세포 내로 아넬로벡터 유전 요소 작제물을 도입하는 단계를 추가로 포함한다.

[0598] 일 양태에서, 본 개시내용은 단백질성 외부에 봉입되는 유전 요소를 포함하는 아넬로벡터를 제조하는 방법을 제공하며, 방법은 아넬로벡터 유전 요소를 포함하는 MOLT-3 세포를 제공하는 단계, 및 아넬로벡터 유전 요소가 MOLT-3 세포의 단백질성 외부에 봉입될 수 있도록 하는 조건 하에서 MOLT-3 세포를 인큐베이션하는 단계를 포함한다. 일부 구현예에서, MOLT-3 세포는 단백질성 외부의 일부 또는 전부를 형성하는 하나 이상의 아넬로바이러스 단백질(예를 들어, 아넬로바이러스 ORF1 분자)을 추가로 포함한다. 일부 구현예에서, 아넬로벡터 유전 요소는 MOLT-3 세포에서, 예를 들어 유전 요소 작제물(예를 들어, 본 명세서에 기재된 바와 같음)로부터 생성된다. 일부 구현예에서, 유전 요소 작제물은 탠덤 작제물(예를 들어, 본 명세서에 기재된 바와 같음)이다. 일부 구현예에서, 유전 요소 작제물은 탠덤 작제물(예를 들어, 본 명세서에 기재된 바와 같음)이 아니다. 일부 구현예에서, 방법은 MOLT-3 세포 내로 아넬로벡터 유전 요소 작제물을 도입하는 단계를 추가로 포함한다.

[0599] 일부 구현예에서, 숙주 세포는 HEK293T 세포, HEK293F 세포, A549 세포, Jurkat 세포, Raji 세포, Chang 세포, HeLa 세포 Phoenix 세포, MRC-5 세포, NCI-H292 세포 또는 Wi38 세포이다. 일부 구현예에서, 숙주 세포는 비-인간 영장류 세포(예를 들어, Vero 세포, CV-1 세포, 또는 LLCMK2 세포)이다. 일부 구현예에서, 숙주 세포는 뮤

린 세포(예를 들어, McCoy 세포)이다. 일부 구현예에서, 숙주 세포는 햄스터 세포(예를 들어, CHO 세포 또는 BHK 21 세포)이다. 일부 구현예에서, 숙주 세포는 MARC-145, MDBK, RK-13, 또는 EEL 세포이다. 일부 구현예에서, 숙주 세포는 상피 세포(예를 들어, 상피 계통의 세포주)이다.

[0600] 일부 구현예에서, 아넬로벡터는 연속 동물 세포주(예를 들어, 연속 증식될 수 있는 무한증식 세포주)에서 배양된다. 본 발명의 일 구현예에 따르면, 세포주는 돼지 세포주를 포함할 수 있다. 본 발명의 맥락에서 고려되는 세포주는 돼지 신장 상피 세포주 PK-15 및 SK, 단골수성 세포주 3D4/31 및 고환 세포주 ST와 같은 무한증식 돼지 세포주를 포함하지만 이에 제한되지 않는다.

[0601] **배양 조건**

[0602] 단백질성 외부의 유전 요소 및 구성성분을 포함하는 숙주 세포는 단백질성 외부 내에 유전 요소를 봉입하기에 적합한 조건 하에서 인큐베이션되어, 아넬로벡터를 생성할 수 있다. 적합한 배양 조건은, 예를 들어, 실시예 9, 10, 12 내지 16, 또는 20 중 임의의 것에 기재된 것을 포함한다. 일부 구현예에서, 숙주 세포는 액체 배지(예를 들어, Grace's Supplemented(TNM-FH), IPL-41, TC-100, Schneider's Drosophila, SF-900 II SFM, 또는 EXPRESS-FIVE™ SFM)에서 인큐베이션된다. 일부 구현예에서, 숙주 세포는 부착 배양에서 인큐베이션된다. 일부 구현예에서, 숙주 세포는 현탁 배양에서 인큐베이션된다. 일부 구현예에서, 숙주 세포는 튜브, 병, 마이크로캐리어 또는 플라스크에서 인큐베이션된다. 일부 구현예에서, 숙주 세포는 접시 또는 웰(예를 들어, 플레이트 상의 웰)에서 인큐베이션된다. 일부 구현예에서, 숙주 세포는 숙주 세포의 증식에 적합한 조건 하에서 인큐베이션된다. 일부 구현예에서, 숙주 세포는 숙주 세포가 내부에서 생성된 아넬로벡터를 주변 상청액으로 방출하기에 적합한 조건 하에서 인큐베이션된다.

[0603] 본 발명에 따른 아넬로벡터-함유 세포 배양물의 생성은 상이한 규모(예를 들어, 플라스크, 롤러 보틀(roller bottle) 또는 생물반응기)에서 수행될 수 있다. 감염될 세포의 배양을 위해 사용되는 배지는 세포 생존에 필요한 표준 영양물질을 포함하지만, 또한 세포 유형에 따라 추가의 영양물질을 포함할 수 있다. 선택적으로, 배지는 단백질-부재 및/또는 혈청-부재일 수 있다. 세포 유형에 따라, 세포는 현탁액 중에서 또는 기질 상에서 배양될 수 있다. 일부 구현예에서, 숙주 세포의 성장을 위하여, 그리고 아넬로벡터의 생성을 위하여 상이한 배지가 사용된다.

[0604] ***수확***

[0605] 숙주 세포에 의해 생성된 아넬로벡터는, 예를 들어 당업계에 알려진 방법에 따라 수확될 수 있다. 예를 들어, 배양 중인 숙주 세포에 의해 주변 상청액으로 방출된 아넬로벡터는 상청액으로부터 수확될 수 있다(예를 들어, 실시예 9에 기재된 바와 같음). 일부 구현예에서, 상청액은 아넬로벡터를 얻기 위해 숙주 세포로부터 분리된다. 일부 구현예에서, 숙주 세포는 수확 전 또는 수확 동안 용해된다. 일부 구현예에서, 아넬로벡터는 숙주 세포 용해물로부터 수확된다(예를 들어, 실시예 15에 기재된 바와 같음). 일부 구현예에서, 아넬로벡터는 숙주 세포 용해물 및 상청액 둘 다로부터 수확된다. 일부 구현예에서, 아넬로벡터의 정제 및 단리는, 예를 들어, 문헌 [Rinaldi, et al., DNA Vaccines: Methods and Protocols (Methods in Molecular Biology), 3rd ed. 2014, Humana Press](본 명세서에 전문이 참조로 포함됨)에 기재된 바와 같이, 바이러스 생성에서 알려진 방법에 따라 수행된다. 일부 구현예에서, 아넬로벡터는 약제학적 부형제를 이용한 제형화 이전에, 생물물리학적 특성에 기반한 용질의 분리, 예를 들어, 이온 교환 크로마토그래피 또는 접선 유동 여과에 의해 수확되고/수확되거나 정제될 수 있다.

[0606] ***농축 및 정제***

[0607] 수확된 아넬로벡터는, 예를 들어 아넬로벡터 제제를 생성하기 위해, 정제 및/또는 농축될 수 있다. 일부 구현예에서, 수확된 아넬로벡터는, 예를 들어, 바이러스 입자를 정제하기 위해 당업계에 알려진 방법(예를 들어, 침강, 크로마토그래피 및/또는 한외여과에 의한 정제)

[0608] 벡터는 예를 들어 아넬로벡터 제제를 생산하기 위해 정제 및/또는 농축될 수 있다. 일부 구현예에서, 수확된 아넬로벡터는 예를 들어 바이러스 입자를 정제하기 위해 당업계에 알려진 방법(예를 들어 침강, 크로마토그래피 및/또는 한외여과에 의한 정제)을 사용하여, 수확 용액에 존재하는 다른 구성요소 또는 오염물로부터 단리된다. 일부 구현예에서, 정제 단계는 제제로부터 혈청, 숙주 세포 DNA, 숙주 세포 단백질, 유전 요소가 결여된 입자, 및/또는 페놀 레드 중 하나 이상을 제거하는 것을 포함한다. 일부 구현예에서, 수확된 아넬로벡터는, 예를 들어, 바이러스 입자를 농축시키기 위해 당업계에 알려진 방법을 사용하여, 수확 용액에 존재하는 다른 구성요소 또는 오염물에 비해 농축된다.

[0609] 일부 구현예에서, 생성된 제제 또는 제제를 포함하는 약제학적 조성물은 허용 가능한 기간 및 온도에 걸쳐 안정할 것이고/이거나, 요망되는 투여 경로 및/또는 투여 경로가 필요할 임의의 디바이스, 예를 들어, 니들 또는 주사기와 양립 가능할 것이다.

[0610] **II. 아넬로벡터**

[0611] 일부 양태에서, 본 명세서에 기재된 발명은 아넬로벡터를 사용 및 제조하기 위한 조성물 및 방법, 아넬로벡터 제제, 및 치료용 조성물을 포함한다. 일부 구현예에서, 아넬로벡터는 본 명세서에 기재된 바와 같은 탠덤 작제물을 사용하여 제조된다. 특정 구현예에서, 아넬로벡터의 유전 요소. 일부 구현예에서, 아넬로벡터는 아넬로바이러스(예를 들어, 본 명세서에 기재된 바와 같은 아넬로바이러스) 또는 이의 단편 또는 부분, 또는 다른 실질적으로 비-병원성인 바이러스, 예를 들어, 상리공생 바이러스(symbiotic virus), 편리공생 바이러스(commensal virus), 천연 바이러스에 기반한 서열, 구조 및/또는 기능을 포함하는 하나 이상의 핵산 또는 폴리펩티드를 포함한다. 일부 구현예에서, 아넬로바이러스-기반 아넬로벡터는 해당 아넬로바이러스에 대하여 외인성인 적어도 하나의 요소, 예를 들어 외인성 이펙터 또는 아넬로벡터의 유전 요소 내에 배치된 외인성 이펙터를 인코딩하는 핵산 서열을 포함한다. 일부 구현예에서, 아넬로바이러스-기반 아넬로벡터는 해당 아넬로바이러스 유래의 또 다른 요소에 대하여 이종인 적어도 하나의 요소, 예를 들어, 프로모터 요소와 같은 또 다른 연결된 핵산 서열에 대하여 이종인 이펙터-인코딩 핵산 서열을 포함한다. 일부 구현예에서, 아넬로벡터는 유전 요소의 나머지 및/또는 단백질성 외부에 비하여 이종인 적어도 하나의 요소(예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어, 이펙터를 인코딩하는 외인성 요소)를 포함하는 유전 요소(예를 들어, 환형 DNA, 예를 들어, 단일 가닥 DNA)를 포함한다. 아넬로벡터는 페이로드를 위한 숙주, 예를 들어, 인간 내로의 운반 비히클(예를 들어, 실질적으로 비-병원성 운반 비히클)일 수 있다. 일부 구현예에서, 아넬로벡터는 진핵 세포, 예를 들어, 포유동물 세포, 예를 들어, 인간 세포에서 복제할 수 있다. 일부 구현예에서, 아넬로벡터는 포유동물(예를 들어, 인간) 세포에서 실질적으로 비-병원성 및/또는 실질적으로 비통합성이다. 일부 구현예에서, 아넬로벡터는 포유동물, 예를 들어, 인간에서 실질적으로 비-면역원성이다. 일부 구현예에서, 아넬로벡터는 복제-결핍이다. 일부 구현예에서, 아넬로벡터는 복제-적격이다.

[0612] 일부 구현예에서, 아넬로벡터는, 예를 들어, PCT 출원 PCT/US2018/037379(이는 본 명세서에 전문이 참조로 포함됨)에 기재된 바와 같은, 큐론 또는 이의 구성성분(예를 들어, 이펙터를 인코딩하는 서열, 및/또는 단백질성 외부를 포함하는, 예를 들어, 유전 요소)을 포함한다. 일부 구현예에서, 아넬로벡터는, 예를 들어, PCT 출원 PCT/US19/65995(이는 본 명세서에 전문이 참조로 포함됨)에 기재된 바와 같은, 아넬로벡터, 또는 이의 구성성분(예를 들어, 이펙터를 인코딩하는 서열, 및/또는 단백질성 외부를 포함하는, 예를 들어, 유전 요소)을 포함한다.

[0613] 일 양태에서, 본 발명은 (i) 프로모터 요소, 이펙터(예를 들어, 내인성 이펙터 또는 외인성 이펙터, 예를 들어, 페이로드)를 인코딩하는 서열 및 단백질 결합 서열(예를 들어, 외부 단백질 결합 서열, 예를 들어, 패키징 신호)을 포함하는 유전 요소로서, 유전 요소가 단일-가닥 DNA이며, 하기의 특성 중 하나 또는 둘 모두를 갖는 유전 요소: 환형이고/환형이거나 세포에 유입되는 유전 요소의 약 0.001%, 0.005%, 0.01%, 0.05%, 0.1%, 0.5%, 1%, 1.5% 또는 2% 미만의 빈도로 진핵 세포의 게놈 내로 통합됨; 및 (ii) 단백질성 외부를 포함하는 아넬로벡터를 포함하며; 유전 요소는 단백질성 외부 내에 봉입되며; 아넬로벡터는 유전 요소를 진핵 세포 내로 운반할 수 있다.

[0614] 본 명세서에 기재된 아넬로벡터의 일부 구현예에서, 유전 요소는 세포에 유입되는 유전 요소의 약 0.001%, 0.005%, 0.01%, 0.05%, 0.1%, 0.5%, 1%, 1.5% 또는 2% 미만의 빈도로 통합된다. 일부 구현예에서, 대상체에 투여되는 복수의 아넬로벡터로부터 약 0.01%, 0.05%, 0.1%, 0.5%, 1%, 2%, 3%, 4% 또는 5% 미만의 유전 요소는 대상체 내의 하나 이상의 숙주 세포의 게놈 내로 통합될 것이다. 일부 구현예에서, 예를 들어, 본 명세서에 기재된 바와 같은 아넬로벡터의 집단의 유전 요소는 유사한 AAV 바이러스의 집단의 것보다 더 낮은 빈도로, 예를 들어, 유사한 AAV 바이러스의 집단보다 약 50%, 60%, 70%, 75%, 80%, 85%, 90%, 95%, 100%, 또는 그 이상 더 낮은 빈도로 숙주 세포의 게놈 내로 통합된다.

[0615] 일 양태에서, 본 발명은 (i) 프로모터 요소 및 이펙터(예를 들어, 내인성 이펙터 또는 외인성 이펙터, 예를 들어, 페이로드)를 인코딩하는 서열, 및 단백질 결합 서열(예를 들어, 외부 단백질 결합 서열)을 포함하는 유전 요소로서, 유전 요소가 야생형 아넬로바이러스 서열(예를 들어, 야생형 토르크 테노 바이러스(TTV), 토르크 테노 미니 바이러스(TTMV) 또는 TTMDV 서열, 예를 들어, 본 명세서에 기재된 바와 같은 야생형 아넬로바이러스 서열)에 대하여 적어도 75%(예를 들어, 적어도 75, 76, 77, 78, 79, 80, 90, 91, 92, 93, 94, 95, 96, 97, 98,

99 또는 100%) 서열 동일성을 갖는 유전 요소; 및 (ii) 단백질성 외부를 포함하는 아넬로벡터를 포함하며; 유전 요소는 단백질성 외부 내에 봉입되며; 아넬로벡터는 유전 요소를 진핵 세포 내로 운반할 수 있다.

[0616] 일 양태에서, 본 발명은

[0617] a) (i) 외부 단백질(예를 들어, 비-병원성 외부 단백질)을 인코딩하는 서열, (ii) 유전 요소를 비-병원성 외부 단백질에 결합시키는 외부 단백질 결합 서열, 및 (iii) 이펙터(예를 들어, 내인성 또는 외인성 이펙터)를 인코딩하는 서열을 포함하는 유전 요소; 및

[0618] b) 유전 요소와 회합된, 예를 들어, 이를 둘러싸거나 봉입하는 단백질성 외부를 포함하는 아넬로벡터를 포함한다.

[0619] 일부 구현예에서, 아넬로벡터는 비-외피, 환형, 단일-가닥 DNA 바이러스로부터의(또는 이와 70% 초과, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 100%의 상동성을 갖는) 서열 또는 발현 생성물을 포함한다. 동물 환형 단일-가닥 DNA 바이러스는 일반적으로 진핵 비-식물 숙주를 감염시키고, 환형 게놈을 갖는 단일 가닥 DNA(ssDNA) 바이러스의 하위군을 지칭한다. 따라서, 동물 환형 ssDNA 바이러스는 원핵생물을 감염시키는 ssDNA 바이러스(즉, 마이크로바이러스과 및 이노바이러스과) 및 식물을 감염시키는 ssDNA 바이러스(즉, 제미니바이러스과 및 나노바이러스과)와 구별 가능하다. 그들은 또한 비-식물 진핵생물을 감염시키는 선형 ssDNA 바이러스(즉, 파보바이러스과)와 구별 가능하다.

[0620] 일부 구현예에서, 아넬로벡터는 숙주 세포 기능을 예를 들어, 일시적으로 또는 장기간 조절한다. 특정 구현예에서, 세포 기능은 안정하게 변경되며, 예컨대 조절은 적어도 약 1시간 내지 약 30일, 또는 적어도 약 2시간, 6시간, 12시간, 18시간, 24시간, 2일, 3일, 4일, 5일, 6일, 7일, 8일, 9일, 10일, 11일, 12일, 13일, 14일, 15일, 16일, 17일, 18일, 19일, 20일, 21일, 22일, 23일, 24일, 25일, 26일, 27일, 28일, 29일, 30일, 60일 이상 또는 그 사이의 임의의 시간 동안 지속된다. 특정 구현예에서, 세포 기능은 일시적으로 변경되며, 예를 들어, 예컨대 조절은 약 30분 이하 내지 약 7일 또는 약 1시간 이하, 2시간, 3시간, 4시간, 5시간, 6시간, 7시간, 8시간, 9시간, 10시간, 11시간, 12시간, 13시간, 14시간, 15시간, 16시간, 17시간, 18시간, 19시간, 20시간, 21시간, 22시간, 24시간, 36시간, 48시간, 60시간, 72시간, 4일, 5일, 6일, 7일 또는 그 사이의 임의의 시간 동안 지속된다.

[0621] 일부 구현예에서, 유전 요소는 프로모터 요소를 포함한다. 구현예에서, 프로모터 요소는 RNA 중합효소 II-의존성 프로모터, RNA 중합효소 III-의존성 프로모터, PGK 프로모터, CMV 프로모터, EF-1α 프로모터, SV40 프로모터, CAGG 프로모터 또는 UBC 프로모터, TTV 바이러스 프로모터, 조직 특이적, U6(polIII), 활성화 단백질(TetR-VP16, Gal4-VP16, dCas9-VP16 등)에 대한 상류 DNA 결합 부위가 있는 최소 CMV 프로모터로부터 선택된다. 구현예에서, 프로모터 요소는 TATA 박스를 포함한다. 구현예에서, 프로모터 요소는 예를 들어, 본 명세서에 기재된 바와 같은 야생형 아넬로바이러스에 대하여 내인성이다.

[0622] 일부 구현예에서, 유전 요소는 하기의 특징 중 하나 이상을 포함한다: 단일-가닥, 환형, 음성 가닥 및/또는 DNA. 구현예에서, 유전 요소는 에피솜을 포함한다. 일부 구현예에서, 이펙터를 배제한 유전 요소의 부분은 약 2.5 내지 5 kb(예를 들어, 약 2.8 내지 4kb, 약 2.8 내지 3.2kb, 약 3.6 내지 3.9kb 또는 약 2.8 내지 2.9kb), 약 5kb 미만(예를 들어, 약 2.9kb, 3.2 kb, 3.6kb, 3.9kb 또는 4kb 미만) 또는 적어도 100개의 뉴클레오티드(예를 들어, 적어도 1kb)의 조합된 크기를 갖는다.

[0623] 본 명세서에 기재된 바와 같은 아넬로벡터, 아넬로벡터를 포함하는 조성물, 이러한 아넬로벡터의 사용 방법 등은 일부 예에서 부분적으로, 상이한 이펙터, 예를 들어, (예를 들어, IFN 또는 miR-625에 대한) miRNA, shRNA 등 및 단백질 결합 서열, 예를 들어, 캡시드 단백질, 예컨대 Q99153에 결합하는 DNA 서열을 단백질성 외부, 예를 들어, 문헌[Arch Virol (2007) 152: 1961-1975]에 개시된 캡시드와 조합하여, 이어서 이펙터를 세포(예를 들어, 동물 세포, 예를 들어, 인간 세포 또는 비-인간 동물 세포, 예컨대 돼지 또는 마우스 세포)로 운반하는 데 사용될 수 있는 아넬로벡터를 생성하는 방법을 예시하는 실시예에 기초한다. 구현예에서, 이펙터는 인터페론과 같은 인자의 발현을 침묵화시킬 수 있다. 실시예에는 이펙터를 예를 들어, 아넬로바이러스로부터 유래된 서열 내로 삽입함으로써 아넬로벡터가 제조될 수 있는 방법이 추가로 기재된다. 이들 실시예에 기초하여 이하의 설명은 실시예에서 고려되는 특정 발견 및 조합의 다양한 변형을 고려한다. 예를 들어, 당업자는 특정 miRNA가 단지 이펙터의 예로서 사용되며, 다른 이펙터가 예를 들어, 다른 조절 핵산 또는 치료적 펩티드일 수 있음을 실시예로부터 이해할 것이다. 유사하게, 실시예에 사용되는 특정 캡시드는 하기 기재되는 실질적으로 비-병원성인 단백질에 의해 대체될 수 있다. 실시예에 기재된 특정 아넬로바이러스 서열은 또한 하기 기재되는 아넬로바이러

스 서열에 의해 대체될 수 있다. 이들 고려사항은 단백질 결합 서열, 조절 서열, 예컨대 프로모터 등에 유사하게 적용된다. 그와 독립적으로, 당업자는 특히 실시예와 밀접하게 관련이 있는 구현예를 고려할 것이다.

[0624] 일부 구현예에서, 아넬로벡터 또는 아넬로벡터에 포함되는 유전 요소는 세포(예를 들어, 인간 세포) 내로 도입된다. 일부 구현예에서, 예를 들어, 일단 아넬로벡터 또는 유전 요소가 세포 내로 도입되면, 예를 들어, 아넬로벡터의 유전 요소에 의해 인코딩되는 이펙터(예를 들어, RNA, 예를 들어, miRNA)는 세포(예를 들어, 인간 세포)에서 발현된다. 구현예에서, 세포 내로의 아넬로벡터 또는 거기에 포함된 유전 요소의 도입은 예를 들어, 세포에 의한 표적 분자의 발현 수준을 변경시킴으로써 세포에서 표적 분자(예를 들어, 표적 핵산, 예를 들어, RNA 또는 표적 폴리펩티드)의 수준을 조절한다(예를 들어, 증가시키거나 감소시킨다). 구현예에서, 아넬로벡터 또는 거기에 포함된 유전 요소의 도입은 세포에 의해 생성되는 인터페론의 수준을 감소시킨다. 구현예에서, 세포 내로의 아넬로벡터 또는 거기에 포함된 유전 요소의 도입은 세포의 기능을 조절한다(예를 들어, 증가시키거나 감소시킨다). 구현예에서, 세포 내로의 아넬로벡터 또는 거기에 포함된 유전 요소의 도입은 세포의 생존력을 조절한다(예를 들어, 증가시키거나 감소시킨다). 구현예에서, 세포 내로의 아넬로벡터 또는 거기에 포함된 유전 요소의 도입은 세포(예를 들어, 암 세포)의 생존력을 감소시킨다.

[0625] 일부 구현예에서, 본 명세서에 기재된 아넬로벡터(예를 들어, 합성 아넬로벡터)은 70% 미만의 항체 출현율(예를 들어, 약 60%, 50%, 40%, 30%, 20% 또는 10% 미만의 항체 출현율)을 유도한다. 구현예에서, 항체 출현율은 당업계에 알려진 방법에 따라 결정된다. 구현예에서, 항체 출현율은 예를 들어, 문헌[Tsuda et al., 1999; *J. Virol. Methods* 77: 199-206](본 명세서에 참조로 포함됨)에 기재된 항-TTV 항체 검출 방법 및/또는 문헌[Kakkola et al., 2008; *Virology* 382: 182-189](본 명세서에 참조로 포함됨)에 기재된 항-TTV IgG 혈청학적 출현율의 결정 방법에 따라 생물학적 샘플에서 아넬로바이러스(예를 들어, 본 명세서에 기재된 바와 같음) 또는 그에 기초한 아넬로벡터에 대한 항체를 검출함으로써 결정된다. 아넬로바이러스 또는 그에 기초한 아넬로벡터에 대한 항체는 또한 항-바이러스 항체를 검출하기 위한 당업계의 방법, 예를 들어, 문헌[Calcedo et al., 2013; *Front. Immunol.* 4(341): 1-7](본 명세서에 참조로 포함됨)에 기재된 바와 같은 예를 들어, 항-AAV 항체의 검출 방법에 의해 검출될 수 있다.

[0626] 일부 구현예에서, 복제 결핍, 복제 결함 또는 복제 비적격 유전 요소는 유전 요소의 복제에 필요한 필수 기구 또는 구성성분의 전부를 인코딩하지 않는다. 일부 구현예에서, 복제 결함 유전 요소는 복제 인자를 인코딩하지 않는다. 일부 구현예에서, 복제 결함 유전 요소는 하나 이상의 ORF(예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어, ORF1, ORF1/1, ORF1/2, ORF2, ORF2/2, ORF2/3, 및/또는 ORF2t/3)를 인코딩하지 않는다. 일부 구현예에서, 유전 요소에 의해 인코딩되지 않는 기구 또는 구성성분은 (예를 들어, 헬퍼, 예를 들어, 헬퍼 바이러스 또는 헬퍼 플라스미드를 사용하여, 또는 숙주 세포에 의해 포함되는, 예를 들어, 숙주 세포의 게놈 내로 통합되는 핵산에 인코딩되어) 트랜스로 제공되어, 예를 들어, 유전 요소가 트랜스로 제공되는 기구 또는 구성성분의 존재 하에 복제를 겪을 수 있게 할 수 있다.

[0627] 일부 구현예에서, 패키징 결핍, 패키징 결함 또는 패키징 비적격 유전 요소는 단백질성 외부(예를 들어, 단백질성 외부는 예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어, ORF1 핵산에 의해 인코딩된 폴리펩티드를 포함하는, 캡시드 또는 이의 부분을 포함함) 내에 패키징될 수 없다. 일부 구현예에서, 패키징 결핍 유전 요소는 (예를 들어, 본 명세서에 기재된 바와 같은) 야생형 아넬로바이러스에 비하여 10% 미만(예를 들어, 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, 1%, 0.5%, 0.1%, 0.01% 또는 0.001% 미만)의 효율로 단백질성 외부 내로 패키징된다. 일부 구현예에서, 패키징 결함 유전 요소는 심지어 (예를 들어, 본 명세서에 기재된 바와 같은) 야생형 아넬로바이러스의 유전 요소의 패키징을 허용할 인자(예를 들어, ORF1, ORF1/1, ORF1/2, ORF2, ORF2/2, ORF2/3 또는 ORF2t/3)의 존재 하에서도 단백질성 외부 내에 패키징될 수 없다. 일부 구현예에서, 패키징 결핍 유전 요소는 심지어 (예를 들어, 본 명세서에 기재된 바와 같은) 야생형 아넬로바이러스의 유전 요소의 패키징을 허용할 인자(예를 들어, ORF1, ORF1/1, ORF1/2, ORF2, ORF2/2, ORF2/3 또는 ORF2t/3)의 존재 하에서도, (예를 들어, 본 명세서에 기재된 바와 같은) 야생형 아넬로바이러스에 비하여 10% 미만(예를 들어, 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, 1%, 0.5%, 0.1%, 0.01% 또는 0.001% 미만)의 효율로 단백질성 외부 내로 패키징된다.

[0628] 일부 구현예에서, 패키징 적격 유전 요소는 단백질성 외부(예를 들어, 단백질성 외부는 예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어, ORF1 핵산에 의해 인코딩된 폴리펩티드를 포함하는, 캡시드 또는 이의 부분을 포함함) 내로 패키징될 수 있다. 일부 구현예에서, 패키징 적격 유전 요소는 (예를 들어, 본 명세서에 기재된 바와 같은) 야생형 아넬로바이러스에 비하여 적어도 20%(예를 들어, 적어도 20%, 30%, 40%, 50%, 60%, 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 100% 이상)의 효율로 단백질성 외부 내에 패키징된다. 일부 구현예에서, 패키징 적격 유전 요소는 (예를 들어, 본 명세서에 기재된 바와 같은) 야생형 아넬로바이러스의 유전 요소의 패

키징을 허용할 인자(예를 들어, ORF1, ORF1/1, ORF1/2, ORF2, ORF2/2, ORF2/3 또는 ORF2t/3)의 존재 하에서, 단백질성 외부 내로 패키징될 수 있다. 일부 구현예에서, 패키징 적격 유전 요소는 (예를 들어, 본 명세서에 기재된 바와 같은) 야생형 아넬로바이러스의 유전 요소의 패키징을 허용할 인자(예를 들어, ORF1, ORF1/1, ORF1/2, ORF2, ORF2/2, ORF2/3 또는 ORF2t/3)의 존재 하에서, (예를 들어, 본 명세서에 기재된 바와 같은) 야생형 아넬로바이러스에 비하여 적어도 20%(예를 들어, 적어도 20%, 30%, 40%, 50%, 60%, 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 100% 이상)의 효율로 단백질성 외부 내로 패키징된다.

[0629] *아넬로바이러스*

[0630] 일부 구현예에서, 예를 들어, 본 명세서에 기재된 바와 같은 아넬로벡터는 아넬로바이러스로부터 유래된 서열 또는 발현 생성물을 포함한다. 일부 구현예에서, 아넬로벡터는 아넬로바이러스에 비하여 외인성인 하나 이상의 서열 또는 발현 생성물을 포함한다. 일부 구현예에서, 아넬로벡터는 아넬로바이러스에 비하여 내인성인 하나 이상의 서열 또는 발현 생성물을 포함한다. 일부 구현예에서, 아넬로벡터는 아넬로벡터 내의 하나 이상의 다른 서열 또는 발현 생성물에 비하여 이종인 하나 이상의 서열 또는 발현 생성물을 포함한다. 아넬로바이러스는 일반적으로 음의 극성을 갖는 단일-가닥 환형 DNA 게놈을 갖는다. 아넬로바이러스는 일반적으로 임의의 인간 질환과 연관된 바 없다. 그러나, 아넬로바이러스 감염을 인간 질환과 연관시키려는 시도는 대조군 코호트 집단(들)에서의 무증상 아넬로바이러스 바이러스혈증의 높은 유병률, 아넬로바이러스 과 내의 현저한 게놈 다양성, 종래에 시험관 내에서 작용제를 전파할 수 없음, 아넬로바이러스 질환의 동물 모델(들)의 결여에 의해 혼동되었다(문헌[Yzebe et al., Panminerva Med. (2002) 44:167-177]; 문헌[Biagini, P., Vet. Microbiol. (2004) 98:95-101]).

[0631] 아넬로바이러스는 일반적으로 구비강 또는 배변-구강 감염, 모에서 유아로의 전염 및/또는 자궁 내의 전염에 의해 전염된다(문헌[Gerner et al., Ped. Infect. Dis. J. (2000) 19:1074-1077]). 감염된 사람은 일부 예에서, 장기의(수개월 내지 수년) 아넬로바이러스 바이러스혈증이라는 특징이 나타날 수 있다. 인간은 한가지 초과의 유전자군 또는 바이러스주로 동시-감염될 수 있다(문헌[Saback, et al., Scad. J. Infect. Dis. (2001) 33:121-125]). 이들 유전자군이 감염된 인간 내에서 재조합될 수 있다는 제안이 존재한다(문헌[Rey et al., Infect. (2003) 31:226-233]). 이중 가닥 아이소폼(복제성) 중간체는 몇몇의 조직, 예컨대 간, 말초 혈액 단핵 세포 및 골수에서 관찰된 바 있다(문헌[Kikuchi et al., J. Med. Virol. (2000) 61:165-170]; 문헌[Okamoto et al., Biochem. Biophys. Res. Commun. (2002) 270:657-662]; 문헌[Rodriguez-Inigo et al., Am. J. Pathol. (2000) 156:1227-1234]).

[0632] 일부 구현예에서, 유전 요소는 아미노산 서열 또는 이의 기능적 단편, 또는 본 명세서에 기재된 아미노산 서열, 예를 들어, 아넬로바이러스 아미노산 서열 중 어느 하나에 대하여 적어도 약 60%, 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100% 서열 동일성을 갖는 서열을 인코딩하는 뉴클레오티드 서열을 포함한다.

[0633] 일부 구현예에서, 본 명세서에 기재된 바와 같은 아넬로벡터는 예를 들어, 본 명세서에 기재된 바와 같은 아넬로바이러스 서열 또는 이의 단편에 대하여 적어도 약 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100% 서열 동일성을 갖는 서열을 포함하는 하나 이상의 핵산 분자(예를 들어, 본 명세서에 기재된 바와 같은 유전 요소)를 포함한다.

[0634] 일부 구현예에서, 본 명세서에 기재된 바와 같은 아넬로벡터는, 예를 들어, 본 명세서에 기재된 바와 같은, 아넬로바이러스의 TATA 박스, cap 부위, 개시인자 요소, 전사 시작 부위, 5' UTR 보존된 도메인, ORF1, ORF1/1, ORF1/2, ORF2, ORF2/2, ORF2/3, ORF2t/3, 3개의 오픈-리딩 프레임 영역, 폴리(A) 신호, GC-풍부 영역 중 하나 이상 또는 이의 임의의 조합에 대하여 적어도 약 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 서열을 포함하는 하나 이상의 핵산 분자(예를 들어, 본 명세서에 기재된 바와 같은 유전 요소)를 포함한다. 일부 구현예에서, 핵산 분자는 캡시드 단백질을 인코딩하는 서열, 예를 들어, 본 명세서에 기재된 아넬로바이러스 중 임의의 것의 ORF1, ORF1/1, ORF1/2, ORF2, ORF2/2, ORF2/3, ORF2t/3 서열을 포함한다. 구현예에서, 핵산 분자는 아넬로바이러스 ORF1 단백질(또는 이의 스플라이스 변이체 또는 기능적 단편) 또는 아넬로바이러스 ORF1 핵산에 의해 인코딩되는 폴리펩티드에 대하여 적어도 약 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 아미노산 서열을 포함하는 캡시드 단백질을 인코딩하는 서열을 포함한다.

[0635] 구현예에서, 핵산 분자는 표 A1의 아넬로바이러스 ORF1 핵산 서열에 대하여 적어도 약 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 핵산 분자는 표 A1의 아넬로바이러스 ORF1/1 뉴클레오티드 서열에 대하여 적어도 약 70%, 75%, 80%, 85%, 90%, 95%,

96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 핵산 분자는 표 A1의 아넬로바이러스 ORF1/2 뉴클레오티드 서열에 대하여 적어도 약 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 핵산 분자는 표 A1의 아넬로바이러스 ORF2 뉴클레오티드 서열에 대하여 적어도 약 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 핵산 분자는 표 A1의 아넬로바이러스 ORF2/2 뉴클레오티드 서열에 대하여 적어도 약 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 핵산 분자는 표 A1의 아넬로바이러스 ORF2/3 뉴클레오티드 서열에 대하여 적어도 약 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 핵산 분자는 표 A1의 아넬로바이러스 ORF2t/3 뉴클레오티드 서열에 대하여 적어도 약 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 핵산 분자는 표 A1의 아넬로바이러스 TATA 박스 뉴클레오티드 서열에 대하여 적어도 약 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 핵산 분자는 표 A1의 아넬로바이러스 개시인자 요소 뉴클레오티드 서열에 대하여 적어도 약 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 핵산 분자는 표 A1의 아넬로바이러스 전사 시작 부위 뉴클레오티드 서열에 대하여 적어도 약 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 핵산 분자는 표 A1의 아넬로바이러스 5' UTR 보존된 도메인 뉴클레오티드 서열에 대하여 적어도 약 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 핵산 분자는 표 A1의 아넬로바이러스 3개의 오픈-리딩 프레임 영역 뉴클레오티드 서열에 대하여 적어도 약 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 핵산 분자는 표 A1의 아넬로바이러스 폴리(A) 신호 뉴클레오티드 서열에 대하여 적어도 약 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 핵산 분자는 표 A1의 아넬로바이러스 GC-풍부 뉴클레오티드 서열에 대하여 적어도 약 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 핵산 서열을 포함한다.

[0636] 구현예에서, 핵산 분자는 표 B1의 아넬로바이러스 ORF1 핵산 서열에 대하여 적어도 약 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 핵산 분자는 표 B1의 아넬로바이러스 ORF1/1 뉴클레오티드 서열에 대하여 적어도 약 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 핵산 분자는 표 B1의 아넬로바이러스 ORF1/2 뉴클레오티드 서열에 대하여 적어도 약 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 핵산 분자는 표 B1의 아넬로바이러스 ORF2 뉴클레오티드 서열에 대하여 적어도 약 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 핵산 분자는 표 B1의 아넬로바이러스 ORF2/2 뉴클레오티드 서열에 대하여 적어도 약 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 핵산 분자는 표 B1의 아넬로바이러스 ORF2/3 뉴클레오티드 서열에 대하여 적어도 약 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 핵산 분자는 표 B1의 아넬로바이러스 TATA 박스 뉴클레오티드 서열에 대하여 적어도 약 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 핵산 분자는 표 B1의 아넬로바이러스 개시인자 요소 뉴클레오티드 서열에 대하여 적어도 약 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 핵산 분자는 표 B1의 아넬로바이러스 전사 시작 부위 뉴클레오티드 서열에 대하여 적어도 약 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 핵산 분자는 표 B1의 아넬로바이러스 5' UTR 보존된 도메인 뉴클레오티드 서열에 대하여 적어도 약 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 핵산 분자는 표 B1의 아넬로바이러스 3개의 오픈-리딩 프레임 영역 뉴클레오티드 서열에 대하여 적어도 약 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 핵산 분자는 표 B1의 아넬로바이러스 폴리(A) 신호 뉴클레오티드 서열에 대하여 적어도 약 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 핵산 분자는 표 B1의 아넬로바이러스 GC-풍부 뉴클레오티드 서열에 대하여 적어도 약 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 핵산 서열을 포함한다.

[0637] 구현예에서, 핵산 분자는 표 C1의 아넬로바이러스 ORF1 핵산 서열에 대하여 적어도 약 70%, 75%, 80%, 85%,

90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 핵산 분자는 표 C1의 아넬로바이러스 ORF1/1 뉴클레오티드 서열에 대하여 적어도 약 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 핵산 분자는 표 C1의 아넬로바이러스 ORF1/2 뉴클레오티드 서열에 대하여 적어도 약 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 핵산 분자는 표 C1의 아넬로바이러스 ORF2 뉴클레오티드 서열에 대하여 적어도 약 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 핵산 분자는 표 C1의 아넬로바이러스 ORF2/2 뉴클레오티드 서열에 대하여 적어도 약 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 핵산 분자는 표 C1의 아넬로바이러스 ORF2/3 뉴클레오티드 서열에 대하여 적어도 약 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 핵산 분자는 표 C1의 아넬로바이러스 TAIP 뉴클레오티드 서열에 대하여 적어도 약 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 핵산 분자는 표 C1의 아넬로바이러스 TATA 박스 뉴클레오티드 서열에 대하여 적어도 약 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 핵산 분자는 표 C1의 아넬로바이러스 개시인자 요소 뉴클레오티드 서열에 대하여 적어도 약 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 핵산 분자는 표 C1의 아넬로바이러스 전사 시작 부위 뉴클레오티드 서열에 대하여 적어도 약 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 핵산 분자는 표 C1의 아넬로바이러스 5' UTR 보존된 도메인 뉴클레오티드 서열에 대하여 적어도 약 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 핵산 분자는 표 C1의 아넬로바이러스 3개의 오픈-리딩 프레임 영역 뉴클레오티드 서열에 대하여 적어도 약 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 핵산 분자는 표 C1의 아넬로바이러스 폴리(A) 신호 뉴클레오티드 서열에 대하여 적어도 약 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 핵산 분자는 표 C1의 아넬로바이러스 GC-풍부 뉴클레오티드 서열에 대하여 적어도 약 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 핵산 서열을 포함한다.

[0638] 일부 구현예에서, 유전 요소는 아미노산 서열 또는 이의 기능적 단편을 인코딩하는 뉴클레오티드 서열 또는 본 명세서에 기재된 아미노산 서열, 예를 들어, 아넬로바이러스 아미노산 서열 중 임의의 하나에 대하여 적어도 약 60%, 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 서열을 포함한다.

[0639] 일부 구현예에서, 본 명세서에 기재된 바와 같은 아넬로벡터는, 예를 들어 본 명세서에 기재된 바와 같은, 아넬로바이러스 서열, 또는 이의 단편에 대하여 적어도 약 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 서열을 포함하는 하나 이상의 핵산 분자(예를 들어, 본 명세서에 기재된 바와 같은 유전 요소)를 포함한다. 구현예에서, 아넬로벡터는 표 A1 내지 M2 중 임의의 것에 나타낸 바와 같은 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 서열로부터 선택되는 핵산 서열을 포함한다. 구현예에서, 아넬로벡터는 표 A2 내지 M2 중 임의의 것에 나타낸 바와 같은 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 서열을 포함하는 폴리펩티드를 포함한다.

[0640] 일부 구현예에서, 본 명세서에 기재된 바와 같은 아넬로벡터는 본 명세서에 기재된 아넬로바이러스(예를 들어, 표 A 내지 M 중 어느 하나에 주석이 달리거나, 그에 열거된 서열에 의해 인코딩되는 바와 같은 아넬로바이러스 서열) 중 어느 하나의 TATA 박스, 캡 부위, 개시인자 요소, 전사 시작 부위, 5' UTR 보존된 도메인, ORF1, ORF1/1, ORF1/2, ORF2, ORF2/2, ORF2/3, ORF2t/3, 3개의 오픈-리딩 프레임 영역, 폴리(A) 신호, GC-풍부 영역 중 하나 이상 또는 이의 임의의 조합에 대하여 적어도 약 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100% 서열 동일성을 갖는 서열을 포함하는 하나 이상의 핵산 분자(예를 들어, 본 명세서에 기재된 바와 같은 유전 요소)를 포함한다. 일부 구현예에서, 핵산 분자는 캡시드 단백질을 인코딩하는 서열, 예를 들어, 본 명세서에 기재된 아넬로바이러스(예를 들어, 표 A 내지 M 중 어느 하나에 주석이 달리거나, 그에 열거된 서열에 의해 인코딩되는 바와 같은 아넬로바이러스 서열) 중 어느 하나의 ORF1, ORF1/1, ORF1/2, ORF2, ORF2/2, ORF2/3, ORF2t/3 서열을 포함한다. 구현예에서, 핵산 분자는 아넬로바이러스 ORF1 또는 ORF2 단백질(예를 들어, 표 A2 내지 M2 중 어느 하나에 나타낸 바와 같은 ORF1 또는 ORF2 아미노산 서열, 또는 표 A1 내지 M1 중 어느 하나에 나타낸 바와 같은 핵산 서열에 의해 인코딩되는 ORF1 또는 ORF2 아미노산 서열)에 대하여 적어도 약 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100% 서열 동일성을 갖는 아미노산 서열을 포함하는

캡시드 단백질을 인코딩하는 서열을 포함한다. 구현예에서, 핵산 분자는 아넬로바이러스 ORF1 단백질(예를 들어, 표 A2 내지 M2 중 어느 하나에 나타낸 바와 같은 ORF1 아미노산 서열 또는 표 A1 내지 M1 중 어느 하나에 나타낸 바와 같은 핵산 서열에 의해 인코딩되는 ORF1 아미노산 서열)에 대하여 적어도 약 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100% 서열 동일성을 갖는 아미노산 서열을 포함하는 캡시드 단백질을 인코딩하는 서열을 포함한다.

[0641] 일부 구현예에서, 본 명세서에 기재된 바와 같은 아넬로벡터는 키메라 아넬로벡터이다. 일부 구현예에서, 키메라 아넬로벡터는 아넬로바이러스 이외의 바이러스 유래의 하나 이상의 요소, 폴리펩티드, 또는 핵산을 추가로 포함한다.

[0642] 구현예에서, 키메라 아넬로벡터는 복수의 상이한 아넬로바이러스(예를 들어, 본 명세서에 기재된 바와 같음)로부터의 서열을 포함하는 복수의 폴리펩티드(예를 들어, 아넬로바이러스 ORF1, ORF1/1, ORF1/2, ORF2, ORF2/2, ORF2/3, 및/또는 ORF2t/3)를 포함한다. 예를 들어, 키메라 아넬로벡터는 하나의 아넬로바이러스 유래의 ORF1 분자(예를 들어, Ring1 ORF1 분자, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또한 99% 아미노산 서열 동일성을 갖는 ORF1 분자) 및 상이한 아넬로바이러스 유래의 ORF2 분자(예를 들어, Ring2 ORF2 분자, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 아미노산 서열 동일성을 갖는 ORF2 분자)를 포함할 수 있다. 또 다른 예에서, 키메라 아넬로벡터는 아넬로바이러스 유래의 제1 ORF1 분자(예를 들어, Ring1 ORF1 분자, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 아미노산 서열 동일성을 갖는 ORF1 분자) 및 상이한 아넬로바이러스 유래의 제2 ORF1 분자(예를 들어, Ring2 ORF1 분자, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 아미노산을 갖는 ORF1 분자)를 포함할 수 있다.

[0643] 일부 구현예에서, 아넬로벡터는, 예를 들어, 아넬로바이러스(예를 들어, 본 명세서에 기재된 바와 같음) 유래의 적어도 하나의 부분 및 상이한 바이러스 유래의 적어도 하나의 부분(예를 들어, 본 명세서에 기재된 바와 같음)을 포함하는, 키메라 폴리펩티드(예를 들어, 아넬로바이러스 ORF1, ORF1/1, ORF1/2, ORF2, ORF2/2, ORF2/3, 및/또는 ORF2t/3)를 포함한다.

[0644] 일부 구현예에서, 아넬로벡터는, 예를 들어 하나의 아넬로바이러스(예를 들어, 본 명세서에 기재된 바와 같음) 유래의 적어도 하나의 부분 및 상이한 아넬로바이러스(예를 들어, 본 명세서에 기재된 바와 같음) 유래의 적어도 하나의 부분을 포함하는, 키메라 폴리펩티드(예를 들어, 아넬로바이러스 ORF1, ORF1/1, ORF1/2, ORF2, ORF2/2, ORF2/3, 및/또는 ORF2t/3)를 포함한다. 구현예에서, 아넬로벡터는 하나의 아넬로바이러스(예를 들어, 본 명세서에 기재된 바와 같음) 유래의 ORF1 분자, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 아미노산 서열 동일성을 갖는 ORF1 분자의 적어도 하나의 부분, 및 상이한 아넬로바이러스(예를 들어, 본 명세서에 기재된 바와 같음) 유래의 ORF1 분자, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 아미노산 서열 동일성을 갖는 ORF1 분자의 적어도 하나의 부분을 포함하는 키메라 ORF1 분자를 포함한다. 구현예에서, 키메라 ORF1 분자는 하나의 아넬로바이러스 유래의 ORF1 젤리-롤 도메인, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 서열 동일성을 갖는 서열, 및 상이한 아넬로바이러스 유래의 ORF1 아미노산 서열(예를 들어, 본 명세서에 기재된 바와 같음), 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 서열 동일성을 갖는 서열을 포함한다. 구현예에서, 키메라 ORF1 분자는 하나의 아넬로바이러스 유래의 ORF1 아르기닌-풍부 영역, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 서열 동일성을 갖는 서열, 및 상이한 아넬로바이러스 유래의 ORF1 아미노산 서열(예를 들어, 본 명세서에 기재된 바와 같음), 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 서열 동일성을 갖는 서열을 포함한다. 구현예에서, 키메라 ORF1 분자는 하나의 아넬로바이러스 유래의 ORF1 초가변 도메인, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 서열 동일성을 갖는 서열, 및 상이한 아넬로바이러스 유래의 ORF1 아미노산 서열(예를 들어, 본 명세서에 기재된 바와 같음), 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 서열 동일성을 갖는 서열을 포함한다. 구현예에서, 키메라 ORF1 분자는 하나의 아넬로바이러스 유래의 ORF1 N22 도메인, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 서열 동일성을 갖는 서열, 및 상이한 아넬로바이러스 유래의 ORF1 아미노산 서열(예를 들어, 본 명세서에 기재된 바와 같음), 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 서열 동일성을 갖는 서열을 포함한다. 구현예에서, 키메라 ORF1 분자는 하나의 아넬로바이러스 유래의 ORF1 C-말단 도메인, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 서열 동일성을 갖는 서열, 및 상이한 아넬로바이러스 유래의 ORF1 아미노산 서열(예를 들어, 본 명세서에 기재된 바와 같음), 또는 이에 대하여 적어도 75%, 80%,

85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 서열 동일성을 갖는 서열을 포함한다.

[0645] 구현예에서, 아넬로벡터는 하나의 아넬로바이러스(예를 들어, 본 명세서에 기재된 바와 같음) 유래의 ORF1/1 분자, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 아미노산 서열 동일성을 갖는 ORF1/1 분자의 적어도 하나의 부분, 및 상이한 아넬로바이러스(예를 들어, 본 명세서에 기재된 바와 같음) 유래의 ORF1/1 분자, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 아미노산 서열 동일성을 갖는 ORF1/1 분자의 적어도 하나의 부분을 포함하는 키메라 ORF1/1 분자를 포함한다. 구현예에서, 아넬로벡터는 하나의 아넬로바이러스(예를 들어, 본 명세서에 기재된 바와 같음) 유래의 ORF1/2 분자, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 아미노산 서열 동일성을 갖는 ORF1/2 분자의 적어도 하나의 부분, 및 상이한 아넬로바이러스(예를 들어, 본 명세서에 기재된 바와 같음) 유래의 ORF1/2 분자, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 아미노산 서열 동일성을 갖는 ORF1/2 분자의 적어도 하나의 부분을 포함하는 키메라 ORF1/2 분자를 포함한다. 구현예에서, 아넬로벡터는 하나의 아넬로바이러스(예를 들어, 본 명세서에 기재된 바와 같음) 유래의 ORF2 분자, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 아미노산 서열 동일성을 갖는 ORF2 분자의 적어도 하나의 부분, 및 상이한 아넬로바이러스(예를 들어, 본 명세서에 기재된 바와 같음) 유래의 ORF2 분자, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 아미노산 서열 동일성을 갖는 ORF2 분자의 적어도 하나의 부분을 포함하는 키메라 ORF2 분자를 포함한다. 구현예에서, 아넬로벡터는 하나의 아넬로바이러스(예를 들어, 본 명세서에 기재된 바와 같음) 유래의 ORF2/2 분자, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 아미노산 서열 동일성을 갖는 ORF2/2 분자의 적어도 하나의 부분, 및 상이한 아넬로바이러스(예를 들어, 본 명세서에 기재된 바와 같음) 유래의 ORF2/2 분자, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 아미노산 서열 동일성을 갖는 ORF2/2 분자의 적어도 하나의 부분을 포함하는 키메라 ORF2/2 분자를 포함한다. 구현예에서, 아넬로벡터는 하나의 아넬로바이러스(예를 들어, 본 명세서에 기재된 바와 같음) 유래의 ORF2/3 분자, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 아미노산 서열 동일성을 갖는 ORF2/3 분자의 적어도 하나의 부분, 및 상이한 아넬로바이러스(예를 들어, 본 명세서에 기재된 바와 같음) 유래의 ORF2/3 분자, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 아미노산 서열 동일성을 갖는 ORF2/3 분자의 적어도 하나의 부분을 포함하는 키메라 ORF2/3 분자를 포함한다. 구현예에서, 아넬로벡터는 하나의 아넬로바이러스(예를 들어, 본 명세서에 기재된 바와 같음) 유래의 ORF2T/3 분자, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 아미노산 서열 동일성을 갖는 ORF2T/3 분자의 적어도 하나의 부분, 및 상이한 아넬로바이러스(예를 들어, 본 명세서에 기재된 바와 같음) 유래의 ORF2T/3 분자, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 아미노산 서열 동일성을 갖는 ORF2T/3 분자의 적어도 하나의 부분을 포함하는 키메라 ORF2T/3 분자를 포함한다.

[0646] 본 명세서에 포함된 서열 또는 하위서열이 (예를 들어, 본 명세서에 기재된 바와 같이, 예를 들어, 아넬로벡터의 유전 요소를 형성하기 위해 또는 탠덤 작제물에서 유전 요소 서열로서) 본 명세서에 기재된 조성물 및 방법에서 이용될 수 있는 추가의 예시적인 아넬로바이러스 게놈은, 예를 들어, PCT 출원 PCT/US2018/037379 및 PCT/US19/65995(본 명세서에 전문이 참조로 포함됨)에 기재되어 있다. 일부 구현예에서, 예시적인 아넬로바이러스 서열은 PCT/US19/65995(본 명세서에 참조로 포함됨)의 표 A1, A3, A5, A7, A9, A11, B1 내지 B5, 1, 3, 5, 7, 9, 11, 13, 15, 또는 17 중 임의의 것에 열거된 바와 같은 핵산 서열을 포함한다. 일부 구현예에서, 예시적인 아넬로바이러스 서열은 PCT/US19/65995(본 명세서에 참조로 포함됨)의 표 A2, A4, A6, A8, A10, A12, C1 내지 C5, 2, 4, 6, 8, 10, 12, 14, 16, 또는 18 중 임의의 것에 열거된 바와 같은 아미노산 서열을 포함한다. 일부 구현예에서, 예시적인 아넬로바이러스 서열은, 예를 들어, PCT/US19/65995(본 명세서에 참조로 포함됨)의 표 21, 23, 25, 27, 29, 31, 33, 35, D2, D4, D6, D8, D10, 또는 37A 내지 37C 중 임의의 것에 열거된 바와 같은, ORF1 분자 서열, 또는 이를 인코딩하는 핵산 서열을 포함한다.

[0647] **[표 A1] 예시적인 아넬로바이러스 핵산 서열(알파토르크바이러스, 클레이드 3)**

명칭	Ring1
속/클레이드	*알파토르크바이러스*, 클레이드 3
수탁번호	AJ620231.1

전체 서열: 3753 bp

```
1         10         20         30         40         50
|         |          |          |          |          |
TGCTACGTCACTAACCCACGTGTCCTCTACAGGCCAATCGCAGTCTATGT
CGTGCACTTCCTGGGCATGGTCTACATAATTATATAAATGCTTGCACTTC
CGAATGGCTGAGTTTTTGCTGCCCGTCCGCGGAGAGGAGCCACGGCAGGG
GATCCGAACGTCCTGAGGGCGGGTGCCGGAGGTGAGTTTACACACCGAAG
TCAAGGGGCAATTCGGGCTCAGGACTGGCCGGGCTTTGGGCAAGGCTCTT
AAAAATGCACTTTTCTCGAATAAGCAGAAAGAAAAGGAAAGTGCTACTGC
TTTGCGTGCCAGCAGCTAAGAAAAAACCAACTGCTATGAGCTTCTGGAAA
CCTCCGGTACACAATGTCACGGGGATCCAACGCATGTGGTATGAGTCCTT
TCACCGTGGCCACGCTTCTTTTTGTGGTTGTGGGAATCCTATACTTCACA
TTACTGCACTTGCTGAAACATATGGCCATCCAACAGGCCCGAGACCTTCT
GGGCCACCGGGAGTAGACCCCAACCCCCACATCCGTAGAGCCAGGCCTGC
CCCGGCCGCTCCGGAGCCCTCACAGGTTGATTCGAGACCAGCCCTGACAT
GGCATGGGGATGGTGGAAGCGACGGAGGCGCTGGTGGTTCCGGAAGCGGT
GGACCCGTGGCAGACTTCGCAGACGATGGCCTCGATCAGCTCGTCGCCGC
CCTAGACGACGAAGAGTAAGGAGGCGCAGACGGTGGAGGAGGGGGAGACG
AAAAACAAGGACTTACAGACGCAGGAGACGCTTTAGACGCAGGGGACGAA
AAGCAAAACTTATAATAAAACTGTGGCAACCTGCAGTAATTAAAAGATGC
AGAATAAAGGGATACATACCACTGATTATAAGTGGGAACGGTACCTTTGC
CACAAACTTTACCAGTCACATAAATGACAGAATAATGAAAGGCCCCTTCG
GGGGAGGACACAGCACTATGAGGTTCAGCCTCTACATTTTGTTTGAGGAG
CACCTCAGACACATGAACTTCTGGACCAGAAGCAACGATAACCTAGAGCT
AACCAGATACTTGGGGGCTTCAGTAAAAATATACAGGCACCCAGACCAAG
ACTTTATAGTAATATACAACAGAAGAACCCCTCTAGGAGGCAACATCTAC
ACAGCACCCTCTCTACACCCAGGCAATGCCATTTTAGCAAAACACAAAAT
ATTAGTACCAAGTTTACAGACAAGACCAAAGGGTAGAAAAGCAATTAGAC
TAAGAATAGCACCCCCCACACTCTTTACAGACAAGTGGTACTTTCAAAAG
GACATAGCCGACCTCACCCTTTTCAACATCATGGCAGTTGAGGCTGACTT
GCGGTTTCCGTTCTGCTCACCACAAACTGACAACACTTGCATCAGCTTCC
AGGTCCTTAGTTCCGTTTACAACAACTACCTCAGTATTAATACCTTTAAT
AATGACAACTCAGACTCAAAGTTAAAAGAATTTTTAAATAAAGCATTTCC
```

[0648]

AACAACAGGCACAAAAGGAACAAGTTTAAATGCACTAAATACATTTAGAA
CAGAAGGATGCATAAGTCACCCACAACTAAAAAAACCAAACCCACAAATA
AACAAACCATTAGAGTCACAATACTTTGCACCTTTAGATGCCCTCTGGGG
AGACCCCATATACTATAATGATCTAAATGAAAACAAAAGTTTGAACGATA
TCATTGAGAAAATACTAATAAAAAACATGATTACATACCATGCAAAACTA
AGAGAATTTCCAAATTCATACCAAGGAAACAAGGCCTTTTGCCACCTAAC
AGGCATATACAGCCCACCATACCTAAACCAAGGCAGAATATCTCCAGAAA
TATTTGGACTGTACACAGAAATAATTTACAACCCTTACACAGACAAAGGA
ACTGGAAACAAAGTATGGATGGACCCACTAACTAAAGAGAACAACATATA
TAAAGAAGGACAGAGCAAATGCCTACTGACTGACATGCCCCTATGGACTT
TACTTTTTGGATATACAGACTGGTGTAAAAAGGACACTAATAACTGGGAC
TTACCACTAAACTACAGACTAGTACTAATATGCCCTTATACCTTTCCAAA
ATTGTACAATGAAAAAGTAAAAGACTATGGGTACATCCCGTACTCCTACA
AATTCGGAGCGGGTCAGATGCCAGACGGCAGCAACTACATACCCTTTCAG
TTTAGAGCAAAGTGGTACCCCACAGTACTACACCAGCAACAGGTAATGGA
GGACATAAGCAGGAGCGGGCCCTTTGCACCTAAGGTAGAAAAACCAAGCA
CTCAGCTGGTAATGAAGTACTGTTTTAACTTTAACTGGGGCGGTAACCCT
ATCATTGAACAGATTGTTAAAGACCCCAGCTTCCAGCCCACCTATGAAAT
ACCCGGTACCGGTAACATCCCTAGAAGAATACAAGTCATCGACCCGCGGG
TCCTGGGACCGCACTACTCGTTCCGGTCATGGGACATGCGCAGACACACA
TTTAGCAGAGCAAGTATTAAGAGAGTGTCAGAACAACAAGAAACTTCTGA
CCTTGTATTCTCAGGCCCAAAAAAGCCTCGGGTCGACATCCCAAAACAAG
AAACCCAAGAAGAAAGCTCACATTCACTCCAAAGAGAATCGAGACCGTGG
GAGACCGAGGAAGAAAGCGAGACAGAAGCCCTCTCGCAAGAGAGCCAAGA
GGTCCCCTTCCAACAGCAGTTGCAGCAGCAGTACCAAGAGCAGCTCAAGC
TCAGACAGGGAATCAAAGTCCTCTTCGAGCAGCTCATAAGGACCCAACAA
GGGGTCCATGTAAACCCATGCCTACGGTAGGTCCCAGGCAGTGGCTGTTT
CCAGAGAGAAAGCCAGCCCCAGCTCCTAGCAGTGGAGACTGGGCCATGGA
GTTTCTCGCAGCAAAAATATTTGATAGGCCAGTTAGAAGCAACCTTAAAG
ATACCCCTTACTACCCATATGTTAAAAACCAATACAATGTCTACTTTGAC
CTTAAATTTGAATAAACAGCAGCTTCAAACTTGCAAGGCCGTGGGAGTTT
CACTGGTCGGTGTCTACCTCTAAAGGTCACTAAGCACTCCGAGCGTAAGC
GAGGAGTGCGACCCTCCCCCCTGGAACAACTTCTTCGGAGTCCGGCGCTA
CGCCTTCGGCTGCGCCGGACACCTCAGACCCCCCCTCCACCCGAAACGCT
TGCGCGTTTCGGACCTTCGGCGTCGGGGGGGTCGGGAGCTTTATTAAACG
GACTCCGAAGTGCTCTTGGACACTGAGGGGGTGAACAGCAACGAAAGTGA
GTGGGGCCAGACTTCGCCATAAGGCCTTTATCTTCTTGCCATTTGTCAGT

[0649] GTCCGGGGTCGCCATAGGCTTCGGGCTCGTTTTTAGGCCTTCCGGACTAC

AAAAATCGCCATTTTGGTGACGTCACGGCCGCCATCTTAAGTAGTTGAGG
CGGACGGTGGCGTGAGTTCAAAGGTCACCATCAGCCACACCTACTCAAAA
TGGTGGACAATTTCTTCCGGGTCAAAGGTTACAGCCGCCATGTTAAAACA
CGTGACGTATGACGTCACGGCCGCCATTTTGTGACACAAGATGGCCGACT
TCCTTCCTCTTTTTCAAAAAAAAGCGGAAGTGCCGCCGCGGCGGCGGGGG
GCGGCGCGCTGCGCGCGCCGCCCAGTAGGGGGAGCCATGCGCCCCCCCCC
GCGCATGCGCGGGGCCCCCCCCCGCGGGGGGCTCCGCCCCCCGGCCCCCC
CCG (서열번호 16)

주석:

추정의 도메인	*염기 범위*
TATA 박스	83 – 88
Cap 부위	104 – 111
전사 시작 부위	111
5' UTR 보존된 도메인	170 – 240
ORF2	336 – 719
ORF2/2	336 – 715 ; 2363 – 2789
ORF2/3	336 – 715 ; 2565 – 3015
ORF2t/3	336 – 388 ; 2565 – 3015
ORF1	599 – 2830
ORF1/1	599 – 715 ; 2363 – 2830
ORF1/2	599 – 715 ; 2565 – 2789
3개의 오픈-리딩 프레임 영역	2551 – 2786
폴리(A) 신호	3011 – 3016
GC-풍부 영역	3632 – 3753

[0650]

[0651] [표 A2] 예시적인 아넬로바이러스 아미노산 서열(알파토르크바이러스, 클레이드 3)

Ring1 (*알파토르크바이러스* 클레이드 3)	
ORF2	MSFWKPPVHNVTGIQRMWYESFHRGHASFCGCGNPILHITALAETYGHPTGPRPSG PPGVDPNPHIRRARPAPAAPEPSQVDSRPALTWHGDGGSDGGAGGSGSGGPVADFA DDGLDQLVAALDDEE (서열번호 17)
ORF2/2	MSFWKPPVHNVTGIQRMWYESFHRGHASFCGCGNPILHITALAETYGHPTGPRPSG PPGVDPNPHIRRARPAPAAPEPSQVDSRPALTWHGDGGSDGGAGGSGSGGPVADFA DDGLDQLVAALDDEELLKTPASSPPMKYPVPVTSLEEYKSSTRGSWDRTTRSGHGT CADTHLAEQVLRECQNNKKLLTLYSQAQKSLGSTSQNKKPKKKAHIHSKENRDRG RPRKKARQKPSRKRAKRSPSNSSCSSSTKSSSSSDRESKSSSSSS (서열번호 18)
ORF2/3	MSFWKPPVHNVTGIQRMWYESFHRGHASFCGCGNPILHITALAETYGHPTGPRPSG PPGVDPNPHIRRARPAPAAPEPSQVDSRPALTWHGDGGSDGGAGGSGSGGPVADFA DDGLDQLVAALDDEEPKKASGRHPKTRNPRRKLTFTPKRIETVGDRGRKRDRSPLA REPRGPLPTAVAAAVPRAAQAQTGNQSPLRAAHKDPTRGPCKPMPTVGPRQWLFP ERKPAPAPSSGDWAMEFLAAKIFDRPVRSNLKDTPYYPYVKNQYNVYFDLKFE (서열번호 19)
ORF2t/3	MSFWKPPVHNVTGIQRMWPKKASGRHPKTRNPRRKLTFTPKRIETVGDRGRKRDR SPLAREPRGPLPTAVAAAVPRAAQAQTGNQSPLRAAHKDPTRGPCKPMPTVGPRQ WLFPERKPAPAPSSGDWAMEFLAAKIFDRPVRSNLKDTPYYPYVKNQYNVYFDLK FE (서열번호 20)
ORF1	MAWGWWKRRRRWWFRKRWTRGRLRRRWPRSARRRPRRRRVRRRRRWRRGRRK TRTYRRRRRFRRRGRKAKLIIKLWQPAVIKRCRIKGYIPLIISGNGTFATNFTSHINDR IMKGPFGGGHSTMRFSLYILFEEHLRHMNFWTRSNDNLELTRYLGASVKIYRHPDQ DFIVIYNRRTPLGGNIYTAPSLHPGNAILAKHKILVPSLQTRPKGRKAIRLRIAPPTLFT DKWYFQKDIADLTLFNIMAVEADLRFPFCSPQTDNTCISFQVLSSVYNNYLSINTFN NDNSDSKLKEFLNKAFPTTGTKGTSLNALNTFRTEGCISHPQLKKPNPQINKPLESQ YFAPLDALWGDPIYYNDLNENKSLNDIIEKILIKNMITYHAKLREFPNSYQGNKAFC HLTGIYSPPYLNQGRISPEIFGLYTEIIYNPYTDKGTGNKVWMDPLTKENNIYKEGQS KCLLTDMPLWTLLFGYTDWCKKDTNNWDLPLNYRLVLICPYTFPKLYNEKVKDY GYIPYSYKFGAGQMPDGSNYIPFQFRAKWYPTVLHQQQVMEDISRSGPFAPKVEKP STQLVMKYCFNFNWGGNPIIEQIVKDPSFQPTYEIPGTGNIPRRIQVIDPRVLGPHYSF RSWDMRRHTFSRASIKRVSEQQETSDLVFSGPKKPRVDIPKQETQEESSHSLQRESR PWETEEESETEALSQESQEVPFQQQLQQQYQEQLKLRQGIKVLFEQLIRTQQGVHV NPCLR (서열번호 21)
ORF1/1	MAWGWWKRRRRWWFRKRWTRGRLRRRWPRSARRRPRRRRIVKDPSFQPTYEIPG TGNIPRRIQVIDPRVLGPHYSFRSWDMRRHTFSRASIKRVSEQQETSDLVFSGPKKPR VDIPKQETQEESSHSLQRESRPWETEEESETEALSQESQEVPFQQQLQQQYQEQLKL RQGIKVLFEQLIRTQQGVHVNPCLR (서열번호 22)
ORF1/2	MAWGWWKRRRRWWFRKRWTRGRLRRRWPRSARRRPRRRRAQKSLGSTSQNKK PKKKAHIHSKENRDRGRPRKKARQKPSRKRAKRSPSNSSCSSSTKSSSSSDRESKSSS SSS (서열번호 23)

[0652]

[0653] [표 B1] 예시적인 아넬로바이러스 핵산 서열(베타토르크바이러스)

명칭	Ring2
속/클레이드	*베타토르크바이러스*
수탁번호	JX134045.1

전체 서열: 2797 bp

```
1         10         20         30         40         50
|         |          |          |          |          |
TAATAAATATTCAACAGGAAAACCACCTAATTTAAATTGCCGACCACAAA
CCGTCACTTAGTTCCCCTTTTTGCAACAACTTCTGCTTTTTTCCAACTGC
CGGAAAACCACATAATTTGCATGGCTAACCACAAACTGATATGCTAATTA
ACTTCCACAAAACAACTTCCCCTTTTAAAACCACACCTACAAATTAATTA
TTAAACACAGTCACATCCTGGGAGGTACTACCACACTATAATACCAAGTG
CACTTCCGAATGGCTGAGTTTATGCCGCTAGACGGAGAACGCATCAGTTA
CTGACTGCGGACTGAACTTGGGCGGGTGCCGAAGGTGAGTGAAACCACCG
AAGTCAAGGGGCAATTCGGGCTAGTTCAGTCTAGCGGAACGGGCAAGAAA
CTTAAAATTATTTTATTTTTCAGATGAGCGACTGCTTTAAACCAACATGC
TACAACAACAAAACAAAGCAAACTCACTGGATTAATAACCTGCATTTAAC
CCACGACCTGATCTGCTTCTGCCCAACACCAACTAGACACTTATTACTAG
CTTTAGCAGAACAACAAGAAACAATTGAAGTGTCTAAACAAGAAAAAGAA
AAAATAACAAGATGCCTTATTACTACAGAAGAAGACGGTACAACTACAGA
CGTCCTAGATGGTATGGACGAGGTTGGATTAGACGCCCTTTTCGCAGAAG
ATTTCGAAGAAAAAGAAGGGTAAGACCTACTTATACTACTATTCCTCTAA
AGCAATGGCAACCGCCATATAAAAGAACATGCTATATAAAAGGACAAGAC
TGTTTAATATACTATAGCAACTTAAGACTGGGAATGAATAGTACAATGTA
TGAAAAAAGTATTGTACCTGTACATTGGCCGGGAGGGGGTTCTTTTTCTG
TAAGCATGTTAACTTTAGATGCCTTGTATGATATACATAAACTTTGTAGA
AACTGGTGGACATCCACAAACCAAGACTTACCACTAGTAAGATATAAAGG
ATGCAAAATAACATTTTATCAAAGCACATTTACAGACTACATAGTAAGAA
TACATACAGAACTACCAGCTAACAGTAACAAACTAACATACCCAAACACA
CATCCACTAATGATGATGATGTCTAAGTACAAACACATTATACCTAGTAG
ACAAACAAGAAGAAAAAAGAAACCATACACAAAAATATTTGTAAAACCAC
CTCCGCAATTTGAAAACAAATGGTACTTTGCTACAGACCTCTACAAAATT
CCATTACTACAAATACACTGCACAGCATGCAACTTACAAAACCCATTTGT
AAAACCAGACAAATTATCAAACAATGTTACATTATGGTCACTAAACACCA
TAAGCATACAAAATAGAAACATGTCAGTGGATCAAGGACAATCATGGCCA
TTTAAAATACTAGGAACACAAAGCTTTTATTTTTACTTTTACACCGGAGC
```

[0654]

AAACCTACCAGGTGACACAACACAAATACCAGTAGCAGACCTATTACCAC
TAACAAACCCAAGAATAAACAGACCAGGACAATCACTAAATGAGGCAAAA
ATTACAGACCATATTACTTTCACAGAATACAAAAACAAATTTACAAATTA
TTGGGGTAACCCATTTAATAAACACATTCAAGAACACCTAGATATGATAC
TATACTCACTAAAAAGTCCAGAAGCAATAAAAAACGAATGGACAACAGAA
AACATGAAATGGAACCAATTAAACAATGCAGGAACAATGGCATTAACACC
ATTTAACGAGCCAATATTCACACAAATACAATATAACCCAGATAGAGACA
CAGGAGAAGACACTCAATTATACCTACTCTCTAACGCTACAGGAACAGGA
TGGGACCCACCAGGAATTCCAGAATTAATACTAGAAGGATTTCCACTATG
GTTAATATATTGGGGATTTGCAGACTTTCAAAAAAACCTAAAAAAAGTAA
CAAACATAGACACAAATTACATGTTAGTAGCAAAAACAAAATTTACACAA
AAACCTGGCACATTCTACTTAGTAATACTAAATGACACCTTTGTAGAAGG
CAATAGCCCATATGAAAAACAACCTTTACCTGAAGACAACATTAAATGGT
ACCCACAAGTACAATACCAATTAGAAGCACAAAACAAACTACTACAAACT
GGGCCATTTACACCAAACATACAAGGACAACTATCAGACAATATATCAAT
GTTTTATAAATTTTACTTTAAATGGGGAGGAAGCCCACCAAAAGCAATTA
ATGTTGAAAATCCTGCCCACCAGATTCAATATCCCATACCCCGTAACGAG
CATGAAACAACTTCGTTACAGAGTCCAGGGGAAGCCCCAGAATCCATCTT
ATACTCCTTCGACTATAGACACGGGAACTACACAACAACAGCTTTGTCAC
GAATTAGCCAAGACTGGGCACTTAAAGACACTGTTTCTAAAATTACAGAG
CCAGATCGACAGCAACTGCTCAAACAAGCCCTCGAATGCCTGCAAATCTC
GGAAGAAACGCAGGAGAAAAAAGAAAAAGAAGTACAGCAGCTCATCAGCA
ACCTCAGACAGCAGCAGCAGCTGTACAGAGAGCGAATAATATCATTATTA
AAGGACCAATAACTTTTAACTGTGTAAAAAAGGTGAAATTGTTTGATGAT
AAACCAAAAAAACCGTAGATTTACACCTGAGGAATTTGAAACTGAGTTACA
AATAGCAAAATGGTTAAAGAGACCCCCAAGATCCTTTGTAAATGATCCTC
CCTTTTACCCATGGTTACCACCTGAACCTGTTGTAAACTTTAAGCTTAAT
TTTACTGAATAAAGGCCAGCATTAATTCACTTAAGGAGTCTGTTTATTTA
AGTTAAACCTTAATAAACGGTCACCGCCTCCCTAATACGCAGGCGCAGAA
AGGGGGCTCCGCCCCCTTTAACCCCCAGGGGGCTCCGCCCCCTGAAACCC
CCAAGGGGGCTACGCCCCCTTACACCCCC (서열번호 54)

주석:

추정의 도메인	*염기 범위*
TATA 박스	237– 243
Cap 부위	260 – 267
전사 시작 부위	267

[0655]

5' UTR 보존된 도메인	323 – 393
ORF2	424 – 723
ORF2/2	424 – 719 ; 2274 – 2589
ORF2/3	424 – 719 ; 2449 – 2812
ORF1	612 – 2612
ORF1/1	612 – 719 ; 2274 – 2612
ORF1/2	612 – 719 ; 2449 – 2589
3개의 오픈-리딩 프레임 영역	2441 – 2586
폴리(A) 신호	2808 – 2813
GC-풍부 영역	2868 – 2929

[0656]

[0657] [표 B2] 예시적인 아넬로바이러스 아미노산 서열(베타토르크바이러스)

Ring2 (*베타토르크바이러스*)	
ORF2	MSDCFKPTCYNNKTKQTHWINNLHLTHDLICFCPTPTRHLLLALAEQQETIEVSKQE KEKITRCLITTEEDGTTTDVLDGMDEVGLDALFAEDFEEKEG (서열번호 55)
ORF2/2	MSDCFKPTCYNNKTKQTHWINNLHLTHDLICFCPTPTRHLLLALAEQQETIEVSKQE KEKITRCLITTEEDGTTTDVLDGMDEVGLDALFAEDFEEKEGFNIPYPVTSMKQLRY RVQGKPQNPSYTPSTIDTGTTQQQLCHELAKTGHLKTLFLKLQSQIDSNCSNKPSNA CKSRKKRRRKKKKKYSSSSATSDSSSSCTESE (서열번호 56)
ORF2/3	MSDCFKPTCYNNKTKQTHWINNLHLTHDLICFCPTPTRHLLLALAEQQETIEVSKQE KEKITRCLITTEEDGTTTDVLDGMDEVGLDALFAEDFEEKEGARSTATAQTSPRMP ANLGRNAGEKRKRSTAAHQQPQTAAAAVQRANNIIIKGPITFNCVKKVKLFDDKPK NRRFTPEEFETELQIAKWLKRPPRSFVNDPPFYPWLPPEPVVNFKLNFTE (서열번호 57)
ORF1	MPYYYRRRRYNYRRPRWYGRGWIRRPFRRRFRRKRRVRPTYTTIPLKQWQPPYKR TCYIKGQDCLIYYSNLRLGMNSTMYEKSIVPVHWPGGGSFSVSMLTLDALYDIHKL CRNWWTSTNQDLPLVRYKGCKITFYQSTFTDYIVRIHTELPANSNKLTYPNTHPLM MMMSKYKHIIPSRQTRRKKKPYTKIFVKPPPQFENKWYFATDLYKIPLLQIHCTACN LQNPFVKPDKLSNNVTLWSLNTISIQNRNMSVDQGQSWPFKILGTQSFYFYFYTGA NLPGDTTQIPVADLLPLTNPRINRPGQSLNEAKITDHITFTEYKNKFTNYWGNPFNK HIQEHLDMILYSLKSPEAIKNEWTTENMKWNQLNNAGTMALTPFNEPIFTQIQYNP DRDTGEDTQLYLLSNATGTGWDPPGIPELILEGFPLWLIYWGFADFQKNLKKVTNID TNYMLVAKTKFTQKPGTFYLVILNDTFVEGNSPYEKQPLPEDNIKWYPQVQYQLEA QNKLLQTGPFTPNIQGQLSDNISMFYKFYFKWGGSPPKAINVENPAHQIQYPIPRNE HETTSLQSPGEAPESILYSFDYRHGNYTTTALSRISQDWALKDTVSKITEPDRQQLLK QALECLQISEETQEKKEKEVQQLISNLRQQQQLYRERIISLLKDQ (서열번호 58)
ORF1/1	MPYYYRRRRYNYRRPRWYGRGWIRRPFRRRFRRKRRIQYPIPRNEHETTSLQSPGE APESILYSFDYRHGNYTTTALSRISQDWALKDTVSKITEPDRQQLLKQALECLQISEE TQEKKEKEVQQLISNLRQQQQLYRERIISLLKDQ (서열번호 59)
ORF1/2	MPYYYRRRRYNYRRPRWYGRGWIRRPFRRRFRRKRRSQIDSNCSNKPSNACKSRK KRRRKKKKKYSSSSATSDSSSSCTESE (서열번호 60)

[0658]

[0659] [표 C1] 예시적인 아넬로바이러스 핵산 서열(감마토르크바이러스)

명칭	Ring4
속/클레이드	*감마토르크바이러스*
수탁번호	

전체 서열: 3176 bp

```
1         10         20         30         40         50
|         |          |          |          |          |
TAAAATGGCGGGAGCCAATCATTTTATACTTTCACTTTCCAATTAAAAAT
GGCCACGTCACAAACAAGGGGTGGAGCCATTTAAACTATATAACTAAGTG
GGGTGGCGAATGGCTGAGTTTACCCCGCTAGACGGTGCAGGGACCGGATC
GAGCGCAGCGAGGAGGTCCCCGGCTGCCCATGGGCGGGAGCCGAGGTGAG
TGAAACCACCGAGGTCTAGGGGCAATTCGGGCTAGGGCAGTCTAGCGGAA
CGGGCAAGAAACTTAAAACAATATTTGTTTTACAGATGGTTAGTATATCC
TCAAGTGATTTTTTTAAGAAAACGAAATTTAATGAGGAGACGCAGAACCA
AGTATGGATGTCTCAAATTGCTGACTCTCATGATAATATCTGCAGTTGCT
GGCATCCATTTGCTCACCTTCTTGCTTCCATATTTCCTCCTGGCCACAAA
GATCGTGATCTTACTATTAACCAAATTCTTCTAAGAGATTATAAAGAAAA
ATGCCATTCTGGTGGAGAAGAAGGAGAAAATTCTGGACCAACAACAGGTT
TAATTACACCAAAAGAAGAAGATATAGAAAAAGATGGCCCAGAAGGCGCC
GCAGAAGAAGACCATACAGACGCCCTGTTCGCCGCCGCCGTAGAAAACTT
CGAAAGGTAAAGAGAAAAAAAAAAATCTTTAATTGTTAGACAATGGCAACC
AGACAGTATAAGAACTTGTAAAATTATAGGACAGTCAGCTATAGTTGTTG
GGGCTGAAGGAAAGCAAATGTACTGTTATACTGTCAATAAGTTAATTAAT
GTGCCCCCAAAAACACCATATGGGGGAGGCTTTGGAGTAGACCAATACAC
ACTGAAATACTTATATGAAGAATACAGATTTGCACAAAACATTTGGACAC
AATCTAATGTACTGAAAGACTTATGCAGATACATAAATGTTAAGCTAATA
TTCTACAGAGACAACAAAACAGACTTTGTCCTTTCCTATGACAGAAACCC
ACCTTTTCAACTAACAAAATTTACATACCCAGGAGCACACCCACAACAAA
TCATGCTTCAAAAACACCACAAATTCATACTATCACAAATGACAAAGCCT
AATGGAAGACTAACAAAAAAACTCAAAATTAAACCTCCTAAACAAATGCT
TTCTAAATGGTTCTTTTCAAAACAATTCTGTAAATACCCTTTACTATCTC
TTAAAGCTTCTGCACTAGACCTTAGGCACTCTTACCTAGGCTGCTGTAAT
GAAAATCCACAGGTATTTTTTTATTATTTAAACCATGGATACTACACAAT
AACAAACTGGGGAGCACAATCCTCAACAGCATACAGACCTAACTCCAAGG
TGACAGACACAACATACTACAGATACAAAAATGACAGAAAAAATATTAAC
```

[0660]

ATTAAAAGCCATGAATACGAAAAAAGTATATCATATGAAAACGGTTATTT
TCAATCTAGTTTCTTACAAACACAGTGCATATATACCAGTGAGCGTGGTG
AAGCCTGTATAGCAGAAAAACCACTAGGAATAGCTATTTACAATCCAGTA
AAAGACAATGGAGATGGTAATATGATATACCTTGTAAGCACTCTAGCAAA
CACTTGGGACCAGCCTCCAAAAGACAGTGCTATTTTAATACAAGGAGTAC
CCATATGGCTAGGCTTATTTGGATATTTAGACTACTGTAGACAAATTAAA
GCTGACAAAACATGGCTAGACAGTCATGTACTAGTAATTCAAAGTCCTGC
TATTTTTACTTACCCAAATCCAGGAGCAGGCAAATGGTATTGTCCACTAT
CACAAAGTTTTATAAATGGCAATGGTCCGTTTAATCAACCACCTACACTG
CTACAAAAAGCAAAGTGGTTTCCACAAATACAATACCAACAAGAAATTAT
TAATAGCTTTGTAGAATCAGGACCATTTGTTCCCAAATATGCAAATCAAA
CTGAAAGCAACTGGGAACTAAAATATAAATATGTTTTTACATTTAAGTGG
GGTGGACCACAATTCCATGAACCAGAAATTGCTGACCCTAGCAAACAAGA
GCAGTATGATGTCCCCGATACTTTCTACCAAACAATACAAATTGAAGATC
CAGAAGGACAAGACCCCAGATCTCTCATCCATGATTGGGACTACAGACGA
GGCTTTATTAAAGAAAGATCTCTTAAAAGAATGTCAACTTACTTCTCAAC
TCATACAGATCAGCAAGCAACTTCAGAGGAAGACATTCCCAAAAAGAAAA
AGAGAATTGGACCCCAACTCACAGTCCCACAACAAAAAGAAGAGGAGACA
CTGTCATGTCTCCTCTCTCTCTGCAAAAAAGATACCTTCCAAGAAACAGA
GACACAAGAAGACCTCCAGCAGCTCATCAAGCAGCAGCAGGAGCAGCAGC
TCCTCCTCAAGAGAAACATCCTCCAGCTCATCCACAAACTAAAAGAGAAT
CAACAAATGCTTCAGCTTCACACAGGCATGTTACCTTAACCAGATTTAAA
CCTGGATTTGAAGAGCAAACAGAGAGAGAATTAGCAATTATATTTCATAG
GCCCCCTAGAACCTACAAAGAGGACCTTCCATTCTATCCCTGGCTACCAC
CTGCACCCCTTGTACAATTTAACCTTAACTTCAAAGGCTAGGCCAACAAT
GTACACTTAGTAAAGCATGTTTATTAAAGCACAACCCCCAAAATAAATGT
AAAAATAAAAAAAAAAAAAAAAAAAATAAAAAATTGCAAAAATTCGGCGCT
CGCGCGCATGTGCGCCTCTGGCGCAAATCACGCAACGCTCGCGCGCCCGC
GTATGTCTCTTTACCACGCACCTAGATTGGGGTGCGCGCGCTAGCGCGCG
CACCCCAATGCGCCCCGCCCTCGTTCCGACCCGCTTGCGCGGGTCGGACC
ACTTCGGGCTCGGGGGGGCGCGCCTGCGGCGCTTTTTTACTAAACAGACT
CCGAGCCGCCATTTGGCCCCCTAAGCTCCGCCCCCCTCATGAATATTCAT
AAAGGAAACCACATAATTAGAATTGCCGACCACAAACTGCCATATGCTAA
TTAGTTCCCCTTTTACAAAGTAAAAGGGGAAGTGAACATAGCCCCACACC
CGCAGGGGCAAGGCCCCGCACCCCTACGTCACTAACCACGCCCCCGCCGC
CATCTTGGGTGCGGCAGGGCGGGGGC (서열번호 886)

주석:

추정의 도메인	*염기 범위*
TATA 박스	87– 93
Cap 부위	110 – 117
전사 시작 부위	117
5' UTR 보존된 도메인	185 – 254
ORF2	286 – 660
ORF2/2	286 – 656 ; 1998 – 2442
ORF2/3	286 – 656 ; 2209 – 2641
TAIP	385 - 484
ORF1	501 – 2489
ORF1/1	501 – 656 ; 1998 – 2489
ORF1/2	501 – 656 ; 2209 – 2442
3개의 오픈-리딩 프레임 영역	2209 – 2439
폴리(A) 신호	2672 – 2678
GC-풍부 영역	3076 – 3176

[0661]

[0662]

[0663] [표 C2] 예시적인 아넬로바이러스 아미노산 서열(감마토르크바이러스)

Ring4 (*감마토르크바이러스*)	
ORF2	MVSISSSDFFKKTKFNEETQNQVWMSQIADSHDNICSCWHPFAHLLASIFPP GHKDRDLTINQILLR DYKEKCHSGGEEGENSGPTTGLITPKEEDIEKDGPEGAAEEDHTDALFAAAV ENFER (서열번호 887)
ORF2/2	MVSISSSDFFKKTKFNEETQNQVWMSQIADSHDNICSCWHPFAHLLASIFPP GHKDRDLTINQILLRDYKEKCHSGGEEGENSGPTTGLITPKEEDIEKDGPEGA AEEDHTDALFAAAVENFESGVDHNSMNQKLLTLANKSSMMSPILSTKQYKL KIQKDKTPDLSSMIGTTDEALLKKDLLKECQLTSQLIQISKQLQRKTFPKRKR ELDPNSQSHNKKKRRHCHVSSLSAKKIPSKKQRHKKTSSSSSSSSRSSSSSSR ETSSSSSTN (서열번호 888)
ORF2/3	MVSISSSDFFKKTKFNEETQNQVWMSQIADSHDNICSCWHPFAHLLASIFPP GHKDRDLTINQILLRDYKEKCHSGGEEGENSGPTTGLITPKEEDIEKDGPEGA AEEDHTDALFAAAVENFERSASNFRGRHSQKEKENWTPTHSPTTKRRGDTV MSPLSLQKRYLPRNRDTRRPPAAHQAAAGAAAPPQEKHPPAHPQTKRESTN ASASHRHVTLTRFKPGFEEQTERELAIIFHRPPRTYKEDLPFYPWLPPAPLVQ FNLNFKG (서열번호 889)
TAIP	MRRRRTKYGCLKLLTLMIISAVAGIHLLTFLLPYFLLATKIVILLLTKFF (서열 번호 890)
ORF1	MPFWWRRRRKFWTNNRFNYTKRRRYRKRWPRRRRRRRPYRRPVRRRRRK LRKVKRKKKSLIVRQWQPDSIRTCKIIGQSAIVVGAEGKQMYCYTVNKLINV PPKTPYGGGFGVDQYTLKYLYEEYRFAQNIWTQSNVLKDLCRYINVKLIFY RDNKTDFVLSYDRNPPFQLTKFTYPGAHPQQIMLQKHHKFILSQMTKPNGR LTKKLKIKPPKQMLSKWFFSKQFCKYPLLSLKASALDLRHSYLGCCNENPQ VFFYYLNHGYYTITNWGAQSSTAYRPNSKVTDTTYYRYKNDRKNINIKSHE YEKSISYENGYFQSSFLQTQCIYTSERGEACIAEKPLGIAIYNPVKDNGDGNM IYLVSTLANTWDQPPKDSAILIQGVPIWLGLFGYLDYCRQIKADKTWLDSHV LVIQSPAIFTYPNPGAGKWYCPLSQSFINGNGPFNQPPTLLQKAKWFPQIQYQ QEIINSFVESGPFVPKYANQTESNWELKYKYVFTFKWGGPQFHEPEIADPSK QEQYDVPDTFYQTIQIEDPEGQDPRSLIHDWDYRRGFIKERSLKRMSTYFST HTDQQATSEEDIPKKKKRIGPQLTVPQQKEEETLSCLLSLCKKDTFQETETQE DLQQLIKQQQEQQLLLKRNILQLIHKLKENQQMLQLHTGMLP (서열번호 891)
ORF1/1	MPFWWRRRRKFWTNNRFNYTKRRRYRKRWPRRRRRRRPYRRPVRRRRRK LRKWGGPQFHEPEIADPSKQEQYDVPDTFYQTIQIEDPEGQDPRSLIHDWDY RRGFIKERSLKRMSTYFSTHTDQQATSEEDIPKKKKRIGPQLTVPQQKEEETL SCLLSLCKKDTFQETETQEDLQQLIKQQQEQQLLLKRNILQLIHKLKENQQM LQLHTGMLP (서열번호 892)
ORF1/2	MPFWWRRRRKFWTNNRFNYTKRRRYRKRWPRRRRRRRPYRRPVRRRRRK LRKISKQLQRKTFPKRKR ELDPNSQSHNKKKRRHCHVSSLSAKKIPSKKQRHKKTSSSSSSSSRSSSSSSR ETSSSSSTN (서열번호 893)

[0664]

[0665] 일부 구현예에서, 아넬로벡터는 본 명세서에 전문이 참조로 포함된 PCT 출원 PCT/US2018/037379에 열거된 서열을 포함하는 핵산을 포함한다. 일부 구현예에서, 아넬로벡터는 본 명세서에 전문이 참조로 포함된 PCT 출원 PCT/US2018/037379에 열거된 서열을 포함하는 폴리펩티드를 포함한다. 일부 구현예에서, 아넬로벡터는 본 명세서에 전문이 참조로 포함된 PCT 출원 PCT/US19/65995에 열거된 서열을 포함하는 핵산을 포함한다. 일부 구현예에서, 아넬로벡터는 본 명세서에 전문이 참조로 포함된 PCT 출원 PCT/US19/65995에 열거된 서열을 포함하는 폴리펩티드를 포함한다.

[0666] *ORF1 분자*

[0667] 일부 구현예에서, 아넬로벡터는 ORF1 분자 및/또는 ORF1 분자를 인코딩하는 핵산을 포함한다. 일반적으로, ORF1 분자는 아넬로바이러스 ORF1 단백질(예를 들어, 본 명세서에 기재된 바와 같은 아넬로바이러스 ORF1 단백질)의 구조적 특징 및/또는 활성을 갖는 폴리펩티드를 포함한다. 일부 구현예에서, ORF1 분자는 아넬로바이러스 ORF1

단백질(예를 들어, 본 명세서에 기재된 바와 같은 아넬로바이러스 ORF1 단백질)에 비해 절단을 포함한다. ORF1 분자는, 예를 들어, 단백질성 외부(예를 들어, 본 명세서에 기재된 바와 같음), 예를 들어 캡시드를 형성하기 위해, 다른 ORF1 분자에 결합할 수 있다. 일부 구현예에서, 단백질성 외부는 핵산 분자(예를 들어, 본 명세서에 기재된 바와 같은 유전 요소)를 봉입할 수 있다. 일부 구현예에서, 복수의 ORF1 분자는, 예를 들어, 단백질성 외부를 형성하기 위해, 다량체를 형성할 수 있다. 일부 구현예에서, 다량체는 동종다량체일 수 있다. 다른 구현예에서, 다량체는 이종다량체일 수 있다.

[0668] ORF1 분자는 일부 구현예에서, 하기 중 하나 이상을 포함할 수 있다: 아르기닌-풍부 영역, 예를 들어, 적어도 60%의 염기성 잔기(예를 들어, 적어도 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% 또는 100%의 염기성 잔기; 예를 들어, 60% 내지 90%, 60% 내지 80%, 70% 내지 90% 또는 70 내지 80%의 염기성 잔기)를 갖는 영역을 포함하는 제1 영역, 및 젤리-롤 도메인, 예를 들어, 적어도 6개의 베타 가닥(예를 들어, 4, 5, 6, 7, 8, 9, 10, 11 또는 12개의 베타 가닥)을 포함하는 제2 영역.

[0669] 아르기닌-풍부 영역

[0670] 아르기닌-풍부 영역은 본 명세서에 기재된 아르기닌-풍부 영역 서열에 대하여 적어도 70%(예를 들어, 적어도 약 70, 80, 90, 95, 96, 97, 98, 99 또는 100%) 서열 동일성을 갖거나, 적어도 60%, 70% 또는 80%의 염기성 잔기(예를 들어, 아르기닌, 라이신 또는 이의 조합)를 포함하는 적어도 약 40개 아미노산의 서열을 갖는다.

[0671] 젤리 롤 도메인

[0672] 젤리-롤 도메인 또는 영역은 하기의 특징 중 하나 이상(예를 들어, 1, 2 또는 3개)을 포함하는 폴리펩티드(예를 들어, 더 큰 폴리펩티드에 포함된 도메인 또는 영역)를 포함한다(예를 들어, 이로 이루어진다):

[0673] (i) 젤리-롤 도메인의 적어도 30%(예를 들어, 적어도 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 90% 이상)의 아미노산이 하나 이상의 β-시트의 부분임;

[0674] (ii) 젤리-롤 도메인의 2차 구조는 적어도 4개(예를 들어, 적어도 4, 5, 6, 7, 8, 9, 10, 11 또는 12개)의 β-가닥을 포함함; 및/또는

[0675] (iii) 젤리-롤 도메인의 3차 구조는 적어도 2개(예를 들어, 적어도 2, 3 또는 4개)의 β-시트를 포함함; 및/또는

[0676] (iv) 젤리-롤 도메인은 적어도 2:1, 3:1, 4:1, 5:1, 6:1, 7:1, 8:1, 9:1 또는 10:1의 β-시트 대 α-나선의 비를 포함함.

[0677] 특정 구현예에서, 젤리-롤 도메인은 2개의 β-시트를 포함한다.

[0678] 특정 구현예에서, 하나 이상(예를 들어, 1, 2, 3, 4, 5, 6, 7, 8, 9 또는 10개)의 β-시트는 약 8개(예를 들어, 4, 5, 6, 7, 8, 9, 10, 11 또는 12개)의 β-가닥을 포함한다. 특정 구현예에서, 하나 이상(예를 들어, 1, 2, 3, 4, 5, 6, 7, 8, 9 또는 10개)의 β-시트는 8개의 β-가닥을 포함한다. 특정 구현예에서, 하나 이상(예를 들어, 1, 2, 3, 4, 5, 6, 7, 8, 9 또는 10개)의 β-시트는 7개의 β-가닥을 포함한다. 특정 구현예에서, 하나 이상(예를 들어, 1, 2, 3, 4, 5, 6, 7, 8, 9 또는 10개)의 β-시트는 6개의 β-가닥을 포함한다. 특정 구현예에서, 하나 이상(예를 들어, 1, 2, 3, 4, 5, 6, 7, 8, 9 또는 10개)의 β-시트는 5개의 β-가닥을 포함한다. 특정 구현예에서, 하나 이상(예를 들어, 1, 2, 3, 4, 5, 6, 7, 8, 9 또는 10개)의 β-시트는 4개의 β-가닥을 포함한다.

[0679] 일부 구현예에서, 젤리-롤 도메인은 제2 β-시트에 대한 역평행 배향의 제1 β-시트를 포함한다. 특정 구현예에서, 제1 β-시트는 약 4개(예를 들어, 3, 4, 5 또는 6개)의 β-가닥을 포함한다. 특정 구현예에서, 제2 β-시트는 약 4개(예를 들어, 3, 4, 5 또는 6개)의 β-가닥을 포함한다. 구현예에서, 제1 및 제2 β-시트는 총 약 8개(예를 들어, 6, 7, 8, 9, 10, 11 또는 12개)의 β-가닥을 포함한다.

[0680] 특정 구현예에서, 젤리-롤 도메인은 캡시드 단백질(예를 들어, 본 명세서에 기재된 바와 같은 ORF1 분자)의 구성성분이다. 특정 구현예에서, 젤리-롤 도메인은 자가-조립 활성을 갖는다. 일부 구현예에서, 젤리-롤 도메인을 포함하는 폴리펩티드는 젤리-롤 도메인을 포함하는 또 다른 카피의 폴리펩티드에 결합한다. 일부 구현예에서, 제1 폴리펩티드의 젤리-롤 도메인은 제2 카피의 폴리펩티드의 젤리-롤 도메인에 결합한다.

[0681] *N22 도메인*

[0682] ORF1 분자는 또한, 아넬로바이러스 N22 도메인(예를 들어, 본 명세서에 기재된 바와 같음, 예를 들어, 본 명세서에 기재된 바와 같은 아넬로바이러스 ORF1 단백질로부터의 N22 도메인)의 구조 또는 활성을 포함하는 제3 영역, 및/또는 아넬로바이러스 C-말단 도메인(CTD)(예를 들어, 본 명세서에 기재된 바와 같음, 예를 들어, 본 명세서에 기재된 바와 같은 아넬로바이러스 ORF1 단백질로부터의 CTD)의 구조 또는 활성을 포함하는 제4 영역을 포함할 수 있다. 일부 구현예에서, ORF1 분자는 N-말단에서 C-말단 순서로, 제1, 제2, 제3 및 제4 영역을 포함한다.

[0683] *초가변 영역(HVR)*

[0684] ORF1 분자는 일부 구현예에서, 초가변 영역(HVR), 예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어, 아넬로바이러스 ORF1 단백질로부터의 HVR을 추가로 포함할 수 있다. 일부 구현예에서, HVR은 제2 영역과 제3 영역 사이에 위치한다. 일부 구현예에서, HVR은 적어도 약 55개(예를 들어, 적어도 약 45, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60 또는 65개)의 아미노산(예를 들어, 약 45 내지 160, 50 내지 160, 55 내지 160, 60 내지 160, 45 내지 150, 50 내지 150, 55 내지 150, 60 내지 150, 45 내지 140, 50 내지 140, 55 내지 140 또는 60 내지 140개의 아미노산)을 포함한다.

[0685] *예시적인 ORF1 서열*

[0686] 예시적인 아넬로바이러스 ORF1 아미노산 서열, 및 예시적인 ORF1 도메인의 서열이 하기 표에 제공된다. 일부 구현예에서, 본 명세서에 기재된 폴리펩티드(예를 들어, ORF1 분자)는, 예를 들어, 표 N 내지 Z 중 임의의 것에 기재된 바와 같은, 하나 이상의 아넬로바이러스 ORF1 하위서열에 대하여 적어도 약 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, 본 명세서에 기재된 아넬로벡터는, 예를 들어, 표 N 내지 Z 중 임의의 것에 기재된 바와 같은, 하나 이상의 아넬로바이러스 ORF1 하위서열에 대하여 적어도 약 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 아미노산 서열을 포함하는 ORF1 분자를 포함한다. 일부 구현예에서, 본 명세서에 기재된 아넬로벡터는, 예를 들어, 표 N 내지 Z 중 임의의 것에 기재된 바와 같은, 하나 이상의 아넬로바이러스 ORF1 하위서열에 대하여 적어도 약 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 아미노산 서열을 포함하는 ORF1 분자를 인코딩하는 핵산 분자(예를 들어, 유전 요소)를 포함한다.

[0687] 일부 구현예에서, 하나 이상의 아넬로바이러스 ORF1 하위서열은 아르기닌(Arg)-풍부 도메인, 젤리-롤 도메인, 초가변 영역(HVR), N22 도메인, 또는 C-말단 도메인(CTD)(예를 들어, 표 N 내지 Z 중 임의의 것에 열거된 바와 같음), 또는 이에 대하여 적어도 약 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100% 서열 동일성을 갖는 서열 중 하나 이상을 포함한다. 일부 구현예에서, ORF1 분자는 상이한 아넬로바이러스 유래의 복수의 하위서열(예를 들어, 표 N 내지 Z에 열거된 알파토르크바이러스 클레이드 1 내지 7 하위서열로부터 선택되는 ORF1 하위서열의 임의의 조합)을 포함한다. 구현예에서, ORF1 분자는 하나의 아넬로바이러스 유래의 Arg-풍부 도메인, 젤리-롤 도메인, N22 도메인, 및 CTD 중 하나 이상, 및 또 다른 아넬로바이러스 유래의 HVR을 포함한다. 구현예에서, ORF1 분자는 하나의 아넬로바이러스 유래의 젤리-롤 도메인, HVR, N22 도메인, 및 CTD 중 하나 이상, 및 또 다른 아넬로바이러스 유래의 Arg-풍부 도메인을 포함한다. 구현예에서, ORF1 분자는 하나의 아넬로바이러스 유래의 Arg-풍부 도메인, HVR, N22 도메인, 및 CTD 중 하나 이상, 및 또 다른 아넬로바이러스 유래의 젤리-롤 도메인을 포함한다. 구현예에서, ORF1 분자는 하나의 아넬로바이러스 유래의 Arg-풍부 도메인, 젤리-롤 도메인, HVR, 및 CTD 중 하나 이상, 및 또 다른 아넬로바이러스 유래의 N22 도메인을 포함한다. 구현예에서, ORF1 분자는 하나의 아넬로바이러스 유래의 Arg-풍부 도메인, 젤리-롤 도메인, HVR, 및 N22 도메인 중 하나 이상, 및 또 다른 아넬로바이러스 유래의 CTD를 포함한다.

[0688] ORF1 분자, 또는 이의 스플라이스 변이체 또는 기능적 단편이 (예를 들어, 유전 요소를 봉입함으로써, 예를 들어 아넬로벡터의 단백질성 외부를 형성하기 위해) 본 명세서에 기재된 조성물 및 방법에 이용될 수 있는 추가의 예시적인 아넬로바이러스는, 예를 들어 PCT 출원 PCT/US2018/037379 및 PCT/US19/65995(본 명세서에 전문이 참조로 포함됨)에 기재되어 있다.

[0689] **[표 N] 예시적인 아넬로바이러스 ORF1 아미노산 하위서열(알파토르크바이러스, 클레이드 3)**

명칭	Ring1
속/클레이드	*알파토르크바이러스*, 클레이드 3
수탁번호	AJ620231.1
단백질 수탁 번호	CAF05750.1

전체 서열: 743 AA

```
1         10         20         30         40         50
|         |          |          |          |          |
MAWGWWKRRRRWWFRKRWTRGRLRRRWPRSARRRPRRRRVRRRRRWRRGR
RKTRTYRRRRRFRRRGRKAKLIIKLWQPAVIKRCRIKGYIPLIISGNGTF
ATNFTSHINDRIMKGPFGGGHSTMRFSLYILFEEHLRHMNFWTRSNDNLE
LTRYLGASVKIYRHPDQDFIVIYNRRTPLGGNIYTAPSLHPGNAILAKHK
ILVPSLQTRPKGRKAIRLRIAPPTLFTDKWYFQKDIADLTLFNIMAVEAD
LRFPFCSPQTDNTCISFQVLSSVYNNYLSINTFNNDNSDSKLKEFLNKAF
PTTGTKGTSLNALNTFRTEGCISHPQLKKPNPQINKPLESQYFAPLDALW
GDPIYYNDLNENKSLNDIIEKILIKNMITYHAKLREFPNSYQGNKAFCHL
TGIYSPPYLNQGRISPEIFGLYTEIIYNPYTDKGTGNKVWMDPLTKENNI
YKEGQSKCLLTDMPLWTLLFGYTDWCKKDTNNWDLPLNYRLVLICPYTFP
KLYNEKVKDYGYIPYSYKFGAGQMPDGSNYIPFQFRAKWYPTVLHQQQVM
EDISRSGPFAPKVEKPSTQLVMKYCFNFNWGGNPIIEQIVKDPSFQPTYE
IPGTGNIPRRIQVIDPRVLGPHYSFRSWDMRRHTFSRASIKRVSEQQETS
DLVFSGPKKPRVDIPKQETQEESSHSLQRESRPWETEEESETEALSQESQ
EVPFQQQLQQQYQEQLKLRQGIKVLFEQLIRTQQGVHVNPCLR
( 서열번호    185)
```

주석:

추정의 도메인	*AA 범위*
Arg-풍부 영역	1 – 68
젤리-롤 도메인	69 - 280
초가변 영역	281 - 413
N22	414 – 579
C-말단 도메인	580 - 743

[0690]

[0691] [표 0] 예시적인 아넬로바이러스 ORF1 아미노산 하위서열(알파토르크바이러스, 클레이드 3)

Ring1 ORF1 (*알파토르크바이러스* 클레이드 3)	
Arg-풍부 영역	MAWGWWKRRRRWWFRKRWTRGRLRRRWPRSARRRPRRRRVRRRR RWRRGRRKTRTYRRRRRFRRRGRK (서열번호 186)
젤리-롤 도메인	AKLIIKLWQPAVIKRCRIKGYIPLIISGNGTFATNFTSHINDRIMKGPFGG GHSTMRFSLYILFEEHLRHMNFWTRSNDNLELTRYLGASVKIYRHPDQ DFIVIYNRRTPLGGNIYTAPSLHPGNAILAKHKILVPSLQTRPKGRKAIRL RIAPPTLFTDKWYFQKDIADLTLFNIMAVEADLRFPFCSPQTDNTCISFQ VLSSVYNNYLSI (서열번호 187)
초가변 도메인	NTFNNDNSDSKLKEFLNKAFPTTGTKGTSLNALNTFRTEGCISHPQLKK PNPQINKPLESQYFAPLDALWGDPIYYNDLNENKSLNDIIEKILIKNMIT YHAKLREFPNSYQGNKAFCHLTGIYSPPYLNQGR (서열번호 188)
N22	ISPEIFGLYTEIIYNPYTDKGTGNKVWMDPLTKENNIYKEGQSKCLLTD MPLWTLLFGYTDWCKKDTNNWDLPLNYRLVLICPYTFPKLYNEKVKD YGYIPYSYKFGAGQMPDGSNYIPFQFRAKWYPTVLHQQQVMEDISRSG PFAPKVEKPSTQLVMKYCFNFN (서열번호 189)
C-말단 도메인	WGGNPIIEQIVKDPSFQPTYEIPGTGNIPRRIQVIDPRVLGPHYSFRSWD MRRHTFSRASIKRVSEQQETSDLVFSGPKKPRVDIPKQETQEESSHSLQR ESRPWETEEESETEALSQESQEVPFQQQLQQQYQEQLKLRQGIKVLFEQ LIRTQQGVHVNPCLR (서열번호 190)

[0692]

[0693] **[표 P] 예시적인 아넬로바이러스 ORF1 아미노산 하위서열(베타토르크바이러스)**

명칭	Ring2
속/클레이드	*베타토르크바이러스*
수탁번호	JX134045.1
단백질 수탁 번호	AGG91484.1

전체 서열: 666 AA

```
1         10         20         30         40         50
|         |          |          |          |          |
MPYYYRRRRYNYRRPRWYGRGWIRRPFRRRFRRKRRVRPTYTTIPLKQWQ
PPYKRTCYIKGQDCLIYYSNLRLGMNSTMYEKSIVPVHWPGGGSFSVSML
TLDALYDIHKLCRNWWTSTNQDLPLVRYKGCKITFYQSTFTDYIVRIHTE
LPANSNKLTYPNTHPLMMMMSKYKHIIPSRQTRRKKKPYTKIFVKPPPQF
ENKWYFATDLYKIPLLQIHCTACNLQNPFVKPDKLSNNVTLWSLNTISIQ
NRNMSVDQGQSWPFKILGTQSFYFYFYTGANLPGDTTQIPVADLLPLTNP
RINRPGQSLNEAKITDHITFTEYKNKFTNYWGNPFNKHIQEHLDMILYSL
KSPEAIKNEWTTENMKWNQLNNAGTMALTPFNEPIFTQIQYNPDRDTGED
TQLYLLSNATGTGWDPPGIPELILEGFPLWLIYWGFADFQKNLKKVTNID
TNYMLVAKTKFTQKPGTFYLVILNDTFVEGNSPYEKQPLPEDNIKWYPQV
QYQLEAQNKLLQTGPFTPNIQGQLSDNISMFYKFYFKWGGSPPKAINVEN
PAHQIQYPIPRNEHETTSLQSPGEAPESILYSFDYRHGNYTTTALSRISQ
DWALKDTVSKITEPDRQQLLKQALECLQISEETQEKKEKEVQQLISNLRQ
QQQLYRERIISLLKDQ ( 서열번호    215)
```

주석:

추정의 도메인	*AA 범위*
Arg-풍부 영역	1 – 38
젤리-롤 도메인	39 - 246
초가변 영역	247 - 374
N22	375 – 537
C-말단 도메인	538 – 666

[0694]

[0695] [표 Q] 예시적인 아넬로바이러스 ORF1 아미노산 하위서열(베타토르크바이러스)

Ring2 ORF1 (*베타토르크바이러스*)	
Arg-풍부 영역	MPYYYRRRRYNYRRPRWYGRGWIRRPFRRRFRRKRRVR (서열번호 216)
젤리-롤 도메인	PTYTTIPLKQWQPPYKRTCYIKGQDCLIYYSNLRLGMNSTMYEKSIVPV HWPGGGSFSVSMLTLDALYDIHKLCRNWWTSTNQDLPLVRYKGCKIT FYQSTFTDYIVRIHTELPANSNKLTYPNTHPLMMMMSKYKHIIPSRQTR RKKKPYTKIFVKPPPQFENKWYFATDLYKIPLLQIHCTACNLQNPFVKP DKLSNNVTLWSLNT (서열번호 217)
초가변 도메인	ISIQNRNMSVDQGQSWPFKILGTQSFYFYFYTGANLPGDTTQIPVADLL PLTNPRINRPGQSLNEAKITDHITFTEYKNKFTNYWGNPFNKHIQEHLD MILYSLKSPEAIKNEWTTENMKWNQLNNAG (서열번호 218)
N22	TMALTPFNEPIFTQIQYNPDRDTGEDTQLYLLSNATGTGWDPPGIPELIL EGFPLWLIYWGFADFQKNLKKVTNIDTNYMLVAKTKFTQKPGTFYLVI LNDTFVEGNSPYEKQPLPEDNIKWYPQVQYQLEAQNKLLQTGPFTPNI QGQLSDNISMFYKFYFK (서열번호 219)
C-말단 도메인	WGGSPPKAINVENPAHQIQYPIPRNEHETTSLQSPGEAPESILYSFDYRH GNYTTTALSRISQDWALKDTVSKITEPDRQQLLKQALECLQISEETQEK KEKEVQQLISNLRQQQQLYRERIISLLKDQ (서열번호 220)

[0696]

[0697] [표 R] 예시적인 아넬로바이러스 ORF1 아미노산 하위서열(감마토르크바이러스)

명칭	Ring4
속/클레이드	*감마토르크바이러스*
수탁번호	
단백질 수탁 번호	

전체 서열: 662 AA

```
1         10         20         30         40         50
|         |          |          |          |          |
MPFWWRRRRKFWTNNRFNYTKRRRYRKRWPRRRRRRRPYRRPVRRRRRKL
RKVKRKKKSLIVRQWQPDSIRTCKIIGQSAIVVGAEGKQMYCYTVNKLIN
VPPKTPYGGGFGVDQYTLKYLYEEYRFAQNIWTQSNVLKDLCRYINVKLI
FYRDNKTDFVLSYDRNPPFQLTKFTYPGAHPQQIMLQKHHKFILSQMTKP
NGRLTKKLKIKPPKQMLSKWFFSKQFCKYPLLSLKASALDLRHSYLGCCN
ENPQVFFYYLNHGYYTITNWGAQSSTAYRPNSKVTDTTYYRYKNDRKNIN
IKSHEYEKSISYENGYFQSSFLQTQCIYTSERGEACIAEKPLGIAIYNPV
KDNGDGNMIYLVSTLANTWDQPPKDSAILIQGVPIWLGLFGYLDYCRQIK
ADKTWLDSHVLVIQSPAIFTYPNPGAGKWYCPLSQSFINGNGPFNQPPTL
LQKAKWFPQIQYQQEIINSFVESGPFVPKYANQTESNWELKYKYVFTFKW
GGPQFHEPEIADPSKQEQYDVPDTFYQTIQIEDPEGQDPRSLIHDWDYRR
GFIKERSLKRMSTYFSTHTDQQATSEEDIPKKKKRIGPQLTVPQQKEEET
LSCLLSLCKKDTFQETETQEDLQQLIKQQQEQQLLLKRNILQLIHKLKEN
          QQMLQLHTGMLP ( 서열번호     925)
```

주석:

추정의 도메인	*AA 범위*
Arg-풍부 영역	1 – 58
젤리-롤 도메인	59 - 260
초가변 영역	261 - 339
N22	340 – 499
C-말단 도메인	500 – 662

[0698]

[0699] [표 S] 예시적인 아넬로바이러스 ORF1 아미노산 하위서열(감마토르크바이러스)

Ring4 (*감마토르크바이러스*)	
Arg-풍부 영역	MPFWWRRRRKFWTNNRFNYTKRRRYRKRWPRRRRRRRPYRRPVRRR RRKLRKVKRKKK (서열번호 926)
젤리-롤 도메인	SLIVRQWQPDSIRTCKIIGQSAIVVGAEGKQMYCYTVNKLINVPPKTPY GGGFGVDQYTLKYLYEEYRFAQNIWTQSNVLKDLCRYINVKLIFYRDN KTDFVLSYDRNPPFQLTKFTYPGAHPQQIMLQKHHKFILSQMTKPNGR LTKKLKIKPPKQMLSKWFFSKQFCKYPLLSLKASALDLRHSYLGCCNE NPQVFFYYL (서열번호 927)
초가변 도메인	NHGYYTITNWGAQSSTAYRPNSKVTDTTYYRYKNDRKNINIKSHEYEK SISYENGYFQSSFLQTQCIYTSERGEACIAE (서열번호 928)
N22	KPLGIAIYNPVKDNGDGNMIYLVSTLANTWDQPPKDSAILIQGVPIWLG LFGYLDYCRQIKADKTWLDSHVLVIQSPAIFTYPNPGAGKWYCPLSQS FINGNGPFNQPPTLLQKAKWFPQIQYQQEIINSFVESGPFVPKYANQTES NWELKYKYVFTFK (서열번호 929)
C-말단 도메인	WGGPQFHEPEIADPSKQEQYDVPDTFYQTIQIEDPEGQDPRSLIHDWDY RRGFIKERSLKRMSTYFSTHTDQQATSEEDIPKKKKRIGPQLTVPQQKE EETLSCLLSLCKKDTFQETETQEDLQQLIKQQQEQQLLLKRNILQLIHKL KENQQMLQLHTGMLP (서열번호 930)

[0700]

[0701] 일부 구현예에서, 제1 영역은 핵산 분자(예를 들어, DNA)에 결합할 수 있다. 일부 구현예에서, 염기성 잔기는 아르기닌, 히스티딘 또는 라이신 또는 이의 조합으로부터 선택된다. 일부 구현예에서, 제1 영역은 적어도 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% 또는 100%의 아르기닌 잔기(예를 들어, 60% 내지 90%, 60% 내지 80%, 70% 내지 90% 또는 70 내지 80%의 아르기닌 잔기)를 포함한다. 일부 구현예에서, 제1 영역은 약 30 내지 120개의 아미노산(예를 들어, 약 40 내지 120, 40 내지 100, 40 내지 90, 40 내지 80, 40 내지 70, 50 내지 100, 50 내지 90, 50 내지 80, 50 내지 70, 60 내지 100, 60 내지 90 또는 60 내지 80개의 아미노산)을 포함한다. 일부 구현예에서, 제1 영역은 바이러스 ORF1 아르기닌-풍부 영역(예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어, 아넬로바이러스 ORF1 단백질로부터의 아르기닌-풍부 영역)의 구조 또는 활성을 포함한다. 일부 구현예에서, 제1 영역은 핵 국소화 신호를 포함한다.

[0702] 일부 구현예에서, 제2 영역은 젤리-롤 도메인, 예를 들어, 바이러스 ORF1 젤리-롤 도메인(예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어, 아넬로바이러스 ORF1 단백질로부터의 젤리-롤 도메인)의 구조 또는 활성을 포함한다. 일부 구현예에서, 제2 영역은 또 다른 ORF1 분자의 제2 영역에 결합하여, 예를 들어, 단백질성 외부(예를 들어, 캡시드) 또는 이의 부분을 형성할 수 있다.

[0703] 일부 구현예에서, 제4 영역은 단백질성 외부(예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어, ORF1 분자의 다량체를 포함하는 단백질성 외부)의 표면 상에 노출된다.

[0704] 일부 구현예에서, 제1 영역, 제2 영역, 제3 영역, 제4 영역 및/또는 HVR은 각각 4개 미만(예를 들어, 0, 1, 2 또는 3개)의 베타 시트를 포함한다.

[0705] 일부 구현예에서, 제1 영역, 제2 영역, 제3 영역, 제4 영역, 및/또는 HVR 중 하나 이상은 이종 아미노산 서열(예를 들어, 이종 ORF1 분자로부터의 상응하는 영역)에 의해 대체될 수 있다. 일부 구현예에서, 이종 아미노산 서열은 예를 들어, 본 명세서에 기재된 바와 같은, 요망되는 작용성을 갖는다.

[0706] 일부 구현예에서, ORF1 분자는 (예를 들어, PCT/US19/65995의 도 34에 나타낸 바와 같이) 복수의 보존된 모티프(예를 들어, 약 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100개, 또는 그 이상의 아미노산을 포함하는 모티프)를 포함한다. 일부 구현예에서, 보존된 모티프는 하나 이상의 야생형 아넬로바이러스 클레이드(예를 들어, 알파토르크바이러스, 클레이드 1; 알파토르크바이

러스, 클레이드 2; 알파토르크바이러스, 클레이드 3; 알파토르크바이러스, 클레이드 4; 알파토르크바이러스, 클레이드 5; 알파토르크바이러스, 클레이드 6; 알파토르크바이러스, 클레이드 7; 베타토르크바이러스; 및/또는 감마토르크바이러스)의 ORF1 단백질에 대하여 60, 70, 80, 85, 90, 95 또는 100% 서열 동일성을 보일 수 있다. 구현예에서, 보존된 모티프는 각각 1 내지 1000개(예를 들어, 5 내지 10, 5 내지 15, 5 내지 20, 10 내지 15, 10 내지 20, 15 내지 20, 5 내지 50, 5 내지 100, 10 내지 50, 10 내지 100, 10 내지 1000, 50 내지 100, 50 내지 1000 또는 100 내지 1000개)의 아미노산의 길이를 갖는다. 특정 구현예에서, 보존된 모티프는 약 2 내지 4%(예를 들어, 약 1 내지 8%, 1 내지 6%, 1 내지 5%, 1 내지 4%, 2 내지 8%, 2 내지 6%, 2 내지 5% 또는 2 내지 4%)의 ORF1 분자의 서열로 이루어지며, 각각은 야생형 아넬로바이러스 클레이드의 ORF1 단백질 내의 상응하는 모티프에 대하여 100% 서열 동일성을 보인다. 특정 구현예에서, 보존된 모티프는 약 5 내지 10%(예를 들어, 약 1 내지 20%, 1 내지 10%, 5 내지 20% 또는 5 내지 10%)의 ORF1 분자의 서열로 이루어지며, 각각은 야생형 아넬로바이러스 클레이드의 ORF1 단백질 내의 상응하는 모티프에 대하여 80% 서열 동일성을 보인다. 특정 구현예에서, 보존된 모티프는 약 10 내지 50%(예를 들어, 약 10 내지 20%, 10 내지 30%, 10 내지 40%, 10 내지 50%, 20 내지 40%, 20 내지 50% 또는 30 내지 50%)의 ORF1 분자의 서열로 이루어지며, 각각은 야생형 아넬로바이러스 클레이드의 ORF1 단백질 내의 상응하는 모티프에 대하여 60% 서열 동일성을 보인다. 일부 구현예에서, 보존된 모티프는 표 19에 열거된 바와 같은 하나 이상의 아미노산 서열을 포함한다.

[0707] 일부 구현예에서, ORF1 분자는 예를 들어, 본 명세서에 기재된 바와 같은, 야생형 ORF1 단백질에 비하여 적어도 하나의 차이(예를 들어, 돌연변이, 화학적 변형 또는 후성적 변화)를 포함한다.

[0708] *N22 도메인 내의 보존된 ORF1 모티프*

[0709] 일부 구현예에서, 본 명세서에 기재된 폴리펩티드(예를 들어, ORF1 분자)는 아미노산 서열 $YNPX^2DXGX^2N$(서열번호 829)을 포함하며, X^n은 임의의 n개의 아미노산의 인접 서열이다. 예를 들어, X^2는 임의의 2개의 아미노산의 인접 서열을 나타낸다. 일부 구현예에서, $YNPX^2DXGX^2N$(서열번호 829)은 예를 들어, 본 명세서에 기재된 바와 같은, ORF1 분자의 N22 도메인 내에 포함된다. 일부 구현예에서, 본 명세서에 기재된 유전 요소는 아미노산 서열 $YNPX^2DXGX^2N$(서열번호 829)을 인코딩하는 핵산 서열(예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어, ORF1 분자를 인코딩하는 핵산 서열)을 포함하며, X^n은 임의의 n개의 아미노산의 인접 서열이다.

[0710] 일부 구현예에서, 폴리펩티드(예를 들어, ORF1 분자)는 예를 들어, N22 도메인 내의, 예를 들어, $YNPX^2DXGX^2N$(서열번호 829) 모티프의 부분을 포함하고/포함하거나 이에 측접하는 보존된 2차 구조를 포함한다. 일부 구현예에서, 보존된 2차 구조는 제1 베타 가닥 및/또는 제2 베타 가닥을 포함한다. 일부 구현예에서, 제1 베타 가닥은 약 5 내지 6개(예를 들어, 3, 4, 5, 6, 7 또는 8개)의 아미노산 길이이다. 일부 구현예에서, 제1 베타 가닥은 $YNPX^2DXGX^2N$(서열번호 829) 모티프의 N-종말단에 티로신(Y) 잔기를 포함한다. 일부 구현예에서, $YNPX^2DXGX^2N$(서열번호 829) 모티프는 랜덤 코일(예를 들어, 약 8 내지 9개 아미노산의 랜덤 코일)을 포함한다. 일부 구현예에서, 제2 베타 가닥은 약 7 내지 8개(예를 들어, 5, 6, 7, 8, 9 또는 10개) 아미노산 길이이다. 일부 구현예에서, 제2 베타 가닥은 $YNPX^2DXGX^2N$(서열번호 829) 모티프의 C-종말단에 아스파라긴(N) 잔기를 포함한다.

[0711] 예시적인 $YNPX^2DXGX^2N$(서열번호 829) 모티프-측접 2차 구조는 PCT/US19/65995(본 명세서에 전문이 참조로 포함됨)의 실시예 47 및 도 48에 기재되어 있다. 일부 구현예에서, ORF1 분자는 PCT/US19/65995의 도 48에 나타낸 2차 구조 요소(예를 들어, 베타 가닥)의 하나 이상(예를 들어, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10개 또는 전부)을 포함하는 영역을 포함한다. 일부 구현예에서, ORF1 분자는 (예를 들어, 본 명세서에 기재된 바와 같은) $YNPX^2DXGX^2N$(서열번호 829) 모티프에 측접하는, PCT/US19/65995의 도 48에 나타낸 2차 구조 요소(예를 들어, 베타 가닥) 중 하나 이상(예를 들어, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10개 또는 전부)을 포함하는 영역을 포함한다.

[0712] *ORF1 젤리-롤 도메인 내의 보존된 2차 구조 모티프*

[0713] 일부 구현예에서, 본 명세서에 기재된 폴리펩티드(예를 들어, ORF1 분자)는 (예를 들어, 본 명세서에 기재된 바와 같은) 아넬로바이러스 ORF1 단백질에 의해 포함되는 하나 이상의 2차 구조 요소를 포함한다. 일부 구현예에서, ORF1 분자는 (예를 들어, 본 명세서에 기재된 바와 같은) 아넬로바이러스 ORF1 단백질의 젤리-롤 도메인에 의해 포함되는 하나 이상의 2차 구조 요소를 포함한다. 일반적으로, ORF1 젤리-롤 도메인은 N-말단에서 C-말단 방향의 순서로, 제1 베타 가닥, 제2 베타 가닥, 제1 알파 나선, 제3 베타 가닥, 제4 베타 가닥, 제5 베타 가닥,

제2 알파 나선, 제6 베타 가닥, 제7 베타 가닥, 제8 베타 가닥 및 제9 베타 가닥을 포함하는 2차 구조를 포함한다. 일부 구현예에서, ORF1 분자는 N-말단에서 C-말단 방향의 순서로, 제1 베타 가닥, 제2 베타 가닥, 제1 알파 나선, 제3 베타 가닥, 제4 베타 가닥, 제5 베타 가닥, 제2 알파 나선, 제6 베타 가닥, 제7 베타 가닥, 제8 베타 가닥 및/또는 제9 베타 가닥을 포함하는 2차 구조를 포함한다.

[0714] 일부 구현예에서, 보존된 2차 구조 요소(즉, 베타 가닥 및/또는 알파 나선)의 쌍은 예를 들어, 랜덤 코일 서열, 베타 가닥 또는 알파 나선 또는 이의 조합을 포함하는 개재 아미노산 서열에 의해 분리된다. 보존된 2차 구조 요소 사이의 개재 아미노산 서열은 예를 들어, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30개, 또는 그 이상의 아미노산을 포함할 수 있다. 일부 구현예에서, ORF1 분자는 하나 이상의 추가의 베타 가닥 및/또는 알파 나선을 (예를 들어, 젤리-롤 도메인 내에) 추가로 포함할 수 있다. 일부 구현예에서, 연속 베타 가닥 또는 연속 알파 나선이 조합될 수 있다. 일부 구현예에서, 제1 베타 가닥 및 제2 베타 가닥은 더 큰 베타 가닥 내에 포함된다. 일부 구현예에서, 제3 베타 가닥 및 제4 베타 가닥은 더 큰 베타 가닥 내에 포함된다. 일부 구현예에서, 제4 베타 가닥 및 제5 베타 가닥은 더 큰 베타 가닥 내에 포함된다. 일부 구현예에서, 제6 베타 가닥 및 제7 베타 가닥은 더 큰 베타 가닥 내에 포함된다. 일부 구현예에서, 제7 베타 가닥 및 제8 베타 가닥은 더 큰 베타 가닥 내에 포함된다. 일부 구현예에서, 제8 베타 가닥 및 제9 베타 가닥은 더 큰 베타 가닥 내에 포함된다.

[0715] 일부 구현예에서, 제1 베타 가닥은 약 5 내지 7개(예를 들어, 3, 4, 5, 6, 7, 8, 9 또는 10개)의 아미노산 길이이다. 일부 구현예에서, 제2 베타 가닥은 약 15 내지 16개(예를 들어, 13, 14, 15, 16, 17, 18 또는 19개)의 아미노산 길이이다. 일부 구현예에서, 제1 알파 나선은 약 15 내지 17개(예를 들어, 13, 14, 15, 16, 17, 18, 19 또는 20개)의 아미노산 길이이다. 일부 구현예에서, 제3 베타 가닥은 약 3 내지 4개(예를 들어, 1, 2, 3, 4, 5 또는 6개)의 아미노산 길이이다. 일부 구현예에서, 제4 베타 가닥은 약 10 내지 11개(예를 들어, 8, 9, 10, 11, 12 또는 13개)의 아미노산 길이이다. 일부 구현예에서, 제5 베타 가닥은 약 6 내지 7개(예를 들어, 4, 5, 6, 7, 8, 9 또는 10개)의 아미노산 길이이다. 일부 구현예에서, 제2 알파 나선은 약 8 내지 14개(예를 들어, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16 또는 17개)의 아미노산 길이이다. 일부 구현예에서, 제2 알파 나선은 (예를 들어, 랜덤 코일 서열에 의해 분리된) 2개의 더 작은 알파 나선으로 파단될 수 있다. 일부 구현예에서, 2개의 더 작은 알파 나선의 각각은 약 4 내지 6개(예를 들어, 2, 3, 4, 5, 6, 7 또는 8개)의 아미노산 길이이다. 일부 구현예에서, 제6 베타 가닥은 약 4 내지 5개(예를 들어, 2, 3, 4, 5, 6 또는 7개)의 아미노산 길이이다. 일부 구현예에서, 제7 베타 가닥은 약 5 내지 6개(예를 들어, 3, 4, 5, 6, 7, 8 또는 9개)의 아미노산 길이이다. 일부 구현예에서, 제8 베타 가닥은 약 7 내지 9개(예를 들어, 5, 6, 7, 8, 9, 10, 11, 12 또는 13개)의 아미노산 길이이다. 일부 구현예에서, 제9 베타 가닥은 약 5 내지 7개(예를 들어, 3, 4, 5, 6, 7, 8, 9 또는 10개)의 아미노산 길이이다.

[0716] 예시적인 젤리-롤 도메인 2차 구조는 PCT/US19/65995의 실시예 47 및 도 47에 기재되어 있다. 일부 구현예에서, ORF1 분자는 PCT/US19/65995의 도 47에 나타낸 젤리-롤 도메인 2차 구조 중 임의의 것의 2차 구조 요소(예를 들어, 베타 가닥 및/또는 알파 나선)의 하나 이상(예를 들어, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10개 또는 전부)을 포함하는 영역을 포함한다.

[0717] *컨센서스 ORF1 도메인 서열*

[0718] 일부 구현예에서, 예를 들어, 본 명세서에 기재된 바와 같은, ORF1 분자는 젤리-롤 도메인, N22 도메인, 및/또는 C-말단 도메인(CTD) 중 하나 이상을 포함한다. 일부 구현예에서, 젤리-롤 도메인은 본 명세서에 기재된 바와 같은(예를 들어, 표 37A 내지 37C 중 어느 하나에 열거된 바와 같은) 젤리-롤 도메인 컨센서스 서열을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, N22 도메인은 본 명세서에 기재된 바와 같은(예를 들어, 표 37A 내지 37C 중 어느 하나에 열거된 바와 같은) N22 도메인 컨센서스 서열을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, CTD 도메인은 본 명세서에 기재된 바와 같은(예를 들어, 표 37A 내지 37C 중 어느 하나에 열거된 바와 같은) CTD 도메인 컨센서스 서열을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, 표 37A 내지 37C 중 어느 하나에 형식 "(X_{a-b})"로 열거된 아미노산은 인접 아미노산의 연속물을 포함하며, 여기서, 연속물은 적어도 a개 및 최대 b개의 아미노산을 포함한다. 특정 구현예에서, 연속물에서 모든 아미노산은 동일하다. 다른 구현예에서, 연속물은 적어도 2개(예를 들어, 적어도 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 또는 21개)의 상이한 아미노산을 포함한다.

[0719] [표 37A] 알파토르크바이러스 ORF1 도메인 컨센서스 서열

도메인	서열	서열번호
젤리-롤	LVLTQWQPNTVRRCYIRGYLPLIICGEN(X_{0-3})TTSRNYATHS DDTIQKGPFGGGMSTTTFSLRVLYDEYQRFMNRWTYSNED LDLARYLGCKFTFYRHPDXDFIVQYNTNPPFKDTKLTAPSIH P(X_{1-5})GMLMLSKRKILIPSLKTRPKGKHYVKVRIGPPKLFED KWYTQSDLCDVPLVXLYATAADLQHPFGSPQTDNPCVTFQ VLGSXYNKHLSISP; 여기서 X = 임의의 아미노산임.	227
N22	SNFEFPGAYTDITYNPLTDKGVGNMVWIQYLTKPDTIXDKT QS(X_{0-3})KCLIEDLPLWAALYGYVDFCEKETGDSAIIXNXGRV LIRCPYTKPPLYDKT(X_{0-4})NKGFVPYSTNFGNGKMPGGSGY VPIYWRARWYPTLFHQKEVLEDIVQSGPFAYKDEKPSTQLV MKYCFNFN; 여기서 X = 임의의 아미노산임.	228
CTD	WGGNPISQQVVRNPCKDSG(X_{0-3})SGXGRQPRSVQVVDPKY MGPEYTFHSWDWRRGLFGEKAIKRMSEQPTDDEIFTGGXPK RPRRDPPTXQXPEE(X_{1-4})QKESSSFR(X_{2-14})PWESSSQEXESES QEEEE(X_{0-30})EQTVQQQLRQQLREQRRLRVQLQLLFQQLLKT (X_{0-4})QAGLHINPLLLSQA(X_{0-40})*; 여기서 X = 임의의 아미노산임.	229

[0720]

[0721] [표 37B] 베타토르크바이러스 ORF1 도메인 컨센서스 서열

도메인	서열	서열번호
젤리-롤	LKQWQPSTIRKCKIKGYLPLFQCGKGRISNNYTQYKESIVPH HEPGGGGWSIQQFTLGALYEEHLKLRNWWTKSNDGLPLVR YLGCTIKLYRSEDTDYIVTYQRCYPMTATKLTYLSTQPSRM LMNKHKIIVPSKXT(X_{1-4})NKKKKPYKKIFIKPPSQMQNKWYF QQDIANTPLLQLTXTACSLDRMYLSSDSISNNITFTSLNTNFF QNPNFQ; 여기서 X = 임의의 아미노산임.	230
N22	(X_{4-10})TPLYFECRYNPFKDKGTGNKVYLVSNN(X_{1-8})TGWDPP TDPDLIIEGFPLWLLLWGWLDWQKKLGKIQNIDTDYILVIQS XYYIPP(X_{1-3})KLPYYVPLDXD(X_{0-2})FLHGRSPY(X_{3-16})PSDKQH WHPKVRFQXETINNIALTGPGTPKLPNQKSIQAHMKYKFYF K; 여기서 X = 임의의 아미노산임.	231
CTD	WGGCPAPMETITDPCKQPKYPIPNNLLQTTSLQXPTTPIETYL YKFDERRGLLTKKAAKRIKKDXTTETTLFTDTGXXTSTTLPT XXQTETTQEEXTSEEE(X_{0-5})ETLLQQLQQLRRKQKQLRXRIL QLLQLLXLL(X_{0-26})*; 여기서 X = 임의의 아미노산임.	232

[0722]

[0723] [표 37C] 감마토르크바이러스 ORF1 도메인 컨센서스 서열

도메인	서열	서열번호
젤리-롤	TIPLKQWQPESIRKCKIKGYGTLVLGAEGRQFYCYTNEKDE YTPPKAPGGGGFGVELFSLEYLYEQWKARNNIWTKSNXYK DLCRYTGCKITFYRHPTTDFIVXYSRQPPFEIDKXTYMXXHP QXLLLRKHKKIILSKATNPKGKLKKKIKIKPPKQMLNKWFF QKQFAXYGLVQLQAAACBLRYPRLGCCNENRLITLYYLN; 여기서 X = 임의의 아미노산임.	233
N22	LPIVVARYNPAXDTGKGNKXWLXSTLNGSXWAPPTTDKDL IIEGLPLWLALYGYWSYJKKVKKDKGILQSHMFVVKSPAIQP LXTATTQXTFYPXIDNSFIQGKXPYDEPJTXNQKKLWYPTLE HQQETINAIVESGPYVPKLDNQKNSTWELXYXYTFYFK; 여기서 X = 임의의 아미노산임.	234
CTD	WGGPQIPDQPVEDPKXQGTYPVPDTXQQTIQIXNPLKQKPE TMFHDWDYRRGIITSTALKRMQENLETDSSFXSDSEETP(X_{0-2})KKKKRLTXELPXPQEETEEIQSCLLSLCEESTCQEE(X_{1-6})ENL QQLIHQQQQQQQQLKHNILKLLSDLKZKQRLLQLQTGILE(X_{1-10})*; 여기서 X = 임의의 아미노산임.	235

[0724]

[0725] 일부 구현예에서, 젤리-롤 도메인은 표 21, 23, 25, 27, 29, 31, 33, 35, D2, D4, D6, D8, D10, 또는 37A 내지 37C 중 임의의 것에 열거된 바와 같은 젤리-롤 도메인 아미노산 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 8%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, N22 도메인은 표 21, 23, 25, 27, 29, 31, 33, 35, D2, D4, D6, D8, D10, 또는 37A 내지 37C 중 임의의 것에 열거된 바와 같은 N22 도메인 아미노산 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 8%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100% 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, CTD 도메인은 표 21, 23, 25, 27, 29, 31, 33, 35, D2, D4, D6, D8, D10, 또는 37A 내지 37C 중 임의의 것에 열거된 바와 같은 CTD 도메인 아미노산 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 8%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 아미노산 서열을 포함한다.

[0726] *ORF1 단백질 서열의 확인*

[0727] 일부 구현예에서, 아넬로바이러스 ORF1 단백질 서열, 또는 ORF1 단백질을 인코딩하는 핵산 서열은 아넬로바이러스의 게놈(예를 들어, 핵산 서열결정 기법, 예를 들어, 심층 서열결정 기법에 의해 확인된, 예를 들어 추정의 아넬로바이러스 게놈)으로부터 확인될 수 있다. 일부 구현예에서, ORF1 단백질 서열은 하기 선택 기준 중 하나 이상(예를 들어, 1, 2개, 또는 3개 전부)에 의해 확인된다:

[0728] *(i) 길이 선택:* 단백질 서열(예를 들어, 하기 (ii) 또는 (iii)에 기재된 기준을 통과하는 추정 아넬로바이러스 ORF1 서열)은 추정 아넬로바이러스 ORF1 단백질을 확인하기 위해 약 600개 초과의 아미노산 잔기에 대해 크기-선택될 수 있다. 일부 구현예에서, 아넬로바이러스 ORF1 단백질 서열은 길이가 적어도 약 600, 650, 700, 750, 800, 850, 900, 950, 또는 1000개의 아미노산 잔기이다. 일부 구현예에서, 알파토르크바이러스 ORF1 단백질 서열은 길이가 적어도 약 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 900, 또는 1000개 아미노산 잔기이다. 일부 구현예에서, 베타토르크바이러스 ORF1 단백질 서열은 길이가 적어도 약 650, 660, 670, 680, 690, 700, 750, 800, 900, 또는 1000개 아미노산 잔기이다. 일부 구현예에서, 감마토르크바이러스 ORF1 단백질 서열은 길이가 적어도 약 650, 660, 670, 680, 690, 700, 750, 800, 900, 또는 1000개의 아미노산 잔기이다. 일부 구현예에서, 아넬로바이러스 ORF1 단백질을 인코딩하는 핵산 서열은 길이가 적어도 약 1800, 1900, 2000,

2100, 2200, 2300, 2400, 또는 2500개 뉴클레오티드이다. 일부 구현예에서, 알파토르크바이러스 ORF1 단백질 서열을 인코딩하는 핵산 서열은 길이가 적어도 약 2100, 2150, 2200, 2250, 2300, 2400, 또는 2500개 뉴클레오티드이다. 일부 구현예에서, 베타토르크바이러스 ORF1 단백질 서열을 인코딩하는 핵산 서열은 길이가 적어도 약 1900, 1950, 2000, 2500, 2100, 2150, 2200, 2250, 2300, 2400, 또는 2500개 또는 1000개 뉴클레오티드이다. 일부 구현예에서, 감마토르크바이러스 ORF1 단백질 서열을 인코딩하는 핵산 서열은 길이가 적어도 약 1900, 1950, 2000, 2500, 2100, 2150, 2200, 2250, 2300, 2400, 또는 2500 또는 1000개 뉴클레오티드이다.

[0729] *(ii) ORF1 모티프의 존재:* 단백질 서열(예를 들어, 상기 (i) 또는 하기 (iii)에 기재된 기준을 통과하는 추정 아넬로바이러스 ORF1 서열)을 필터링하여 상기 기재된 N22 도메인에 보존된 ORF1 모티프를 함유하는 것들을 확인할 수 있다. 일부 구현예에서, 추정 아넬로바이러스 ORF1 서열은 서열 YNPXXDXGXXN을 포함한다. 일부 구현예에서, 추정 아넬로바이러스 ORF1 서열은 서열 Y[NCS]P*XX*D*X*[GASKR]*XX*[NTSVAK]을 포함한다.

[0730] *(iii) 아르기닌-풍부 영역의 존재:* 단백질 서열(예를 들어, 상기 (i) 및/또는 (ii)에 기재된 기준을 통과하는 추정 아넬로바이러스 ORF1 서열)은 아르기닌-풍부 영역(예를 들어, 본 명세서에 기재된 바와 같음)을 포함하는 것에 대해 필터링될 수 있다. 일부 구현예에서, 추정 아넬로바이러스 ORF1 서열은 적어도 30%(예를 들어, 적어도 약 20%, 25%, 30%, 35%, 40%, 45%, 또는 50%) 아르기닌 잔기를 포함하는 적어도 약 30, 35, 40, 45, 50, 55, 60, 65, 또는 70개 아미노산의 인접 서열을 포함한다. 일부 구현예에서, 추정 아넬로바이러스 ORF1 서열은 적어도 30%(예를 들어, 적어도 약 20%, 25%, 30%, 35%, 40%, 45%, 또는 50%) 아르기닌 잔기를 포함하는 약 35 내지 40, 40 내지 45, 45 내지 50, 50 내지 55, 55 내지 60, 60 내지 65, 또는 65 내지 70개 아미노산의 인접 서열을 포함한다. 일부 구현예에서, 아르기닌-풍부 영역은 추정 아넬로바이러스 ORF1 단백질의 시작 코돈의 적어도 약 30, 40, 50, 60, 70 또는 80개 아미노산 하류에 위치한다. 일부 구현예에서, 아르기닌-풍부 영역은 추정 아넬로바이러스 ORF1 단백질의 시작 코돈의 적어도 약 50개 아미노산 하류에 위치한다.

[0731] *ORF2 분자*

[0732] 일부 구현예에서, 아넬로벡터는 ORF2 분자 및/또는 ORF2 분자를 인코딩하는 핵산을 포함한다. 일반적으로, ORF2 분자는 아넬로바이러스 ORF2 단백질(예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어, 표 A2, A4, A6, A8, A10, A12, C1 내지 C5, 2, 4, 6, 8, 10, 12, 14, 16 또는 18 중 어느 하나에 열거된 바와 같은 아넬로바이러스 ORF2 단백질)의 구조적 특징 및/또는 활성을 갖는 폴리펩티드 또는 이의 기능적 단편을 포함한다. 일부 구현예에서, ORF2 분자는 표 A2, A4, A6, A8, A10, A12, C1 내지 C5, 2, 4, 6, 8, 10, 12, 14, 16 또는 18 중 어느 하나에 나타낸 바와 같은 아넬로바이러스 ORF2 단백질 서열에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100% 서열 동일성을 갖는 아미노산 서열을 포함한다.

[0733] 일부 구현예에서, ORF2 분자는 알파토르크바이러스, 베타토르크바이러스 또는 감마토르크바이러스 ORF2 단백질에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, ORF2 분자(예를 들어, 알파토르크바이러스 ORF2 단백질에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 갖는 ORF2 분자)는 250개 이하의 아미노산(예를 들어, 약 150 내지 200개의 아미노산)의 길이를 갖는다. 일부 구현예에서, ORF2 분자(예를 들어, 베타토르크바이러스 ORF2 단백질에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 갖는 ORF2 분자)는 약 50 내지 150개의 아미노산의 길이를 갖는다. 일부 구현예에서, ORF2 분자(예를 들어, 감마토르크바이러스 ORF2 단백질에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% 또는 99% 서열 동일성을 갖는 ORF2 분자)는 약 100 내지 200개의 아미노산(예를 들어, 약 100 내지 150개의 아미노산)의 길이를 갖는다. 일부 구현예에서, ORF2 분자는 나선-회전-나선 모티프(예를 들어, 회전 영역에 측접하는 2개의 알파 나선을 포함하는 나선-회전-나선 모티프)를 포함한다. 일부 구현예에서, ORF2 분자는 TTV 분리주 TA278 또는 TTV 분리주 SANBAN의 ORF2 단백질의 아미노산 서열을 포함하지 않는다. 일부 구현예에서, ORF2 분자는 단백질 포스파타제 활성을 갖는다. 일부 구현예에서, ORF2 분자는, 예를 들어, 본 명세서에 기재된 바와 같은(예를 들어, 표 A2, A4, A6, A8, A10, A12, C1 내지 C5, 2, 4, 6, 8, 10, 12, 14, 16 또는 18 중 어느 하나에 나타낸 바와 같은) 야생형 ORF2 단백질에 비하여 적어도 하나의 차이(예를 들어, 돌연변이, 화학적 변형 또는 후성적 변화)를 포함한다.

[0734] *보존된 ORF2 모티프*

[0735] 일부 구현예에서, 본 명세서에 기재된 폴리펩티드(예를 들어, ORF2 분자)는 아미노산 서열 [W/F]X^7HX^3CX^1CX^5H(서열번호 949)를 포함하며, 여기서, X^n은 임의의 N개의 아미노산의 인접 서열이다. 구현예에서, X^7은 임의의 7개의 아미노산의 인접 서열을 나타낸다. 구현예에서, X^3은 임의의 3개의 아미노산의 인접 서열을 나타낸다. 구현예에

서, X^1은 임의의 단일의 아미노산을 나타낸다. 구현예에서, X^5는 임의의 5개의 아미노산의 인접 서열을 나타낸다. 일부 구현예에서, [W/F]는 트립토판 또는 페닐알라닌 중 어느 하나일 수 있다. 일부 구현예에서, $[W/F]X^7HX^3CX^1CX^5H$(서열번호 949)는 예를 들어, 본 명세서에 기재된 바와 같은 ORF2 분자의 N22 도메인 내에 포함된다. 일부 구현예에서, 본 명세서에 기재된 유전 요소는 아미노산 서열 $[W/F]X^7HX^3CX^1CX^5H$(서열번호 949)를 인코딩하는 핵산 서열(예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어, ORF2 분자를 인코딩하는 핵산 서열)을 포함하며, 여기서, X^n은 임의의 n개의 아미노산의 인접 서열이다.

[0736] *유전 요소*

[0737] 일부 구현예에서, 아넬로벡터는 유전 요소를 포함한다. 일부 구현예에서, 유전 요소는 하기의 특징 중 하나 이상을 갖는다: 숙주 세포의 게놈과 실질적으로 비-통합성임, 에피솜 핵산임, 단일 가닥 DNA임, 환형임, 약 1 내지 10 kb임, 세포의 핵 내에 존재함, 내인성 단백질에 의해 결합될 수 있음, 이펙터, 예컨대 숙주 또는 표적 세포의 유전자, 활성 또는 기능을 표적화하는 폴리펩티드 또는 핵산(예를 들어, RNA, iRNA, 마이크로RNA)를 생성함. 일 구현예에서, 유전 요소는 실질적으로 비-통합성 DNA이다. 일부 구현예에서, 유전 요소는 패키징 신호, 예를 들어, 캡시드 단백질에 결합하는 서열을 포함한다. 일부 구현예에서, 패키징 또는 캡시드-결합 서열의 외측에서, 유전 요소는 야생형 아넬로바이러스 핵산 서열에 대하여 70%, 60%, 50%, 40%, 30%, 20%, 10%, 5% 미만의 서열 동일성을 가지며, 예를 들어, 본 명세서에 기재된 바와 같은, 아넬로바이러스 핵산 서열에 대하여 70%, 60%, 50%, 40%, 30%, 20%, 10%, 5% 미만의 서열 동일성을 갖는다. 일부 구현예에서, 패키징 또는 캡시드-결합 서열의 외측에서, 유전 요소는 아넬로바이러스 핵산 서열과 적어도 70%, 75%, 80%, 8%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100% 동일한 500, 450, 400, 350, 300, 250, 200, 150 또는 100개 미만의 인접 뉴클레오티드를 갖는다. 특정 구현예에서, 유전 요소는 프로모터 서열, 치료적 이펙터 및 캡시드 결합 단백질을 인코딩하는 서열을 포함하는 환형, 단일 가닥 DNA이다.

[0738] 일부 구현예에서, 유전 요소는 20 kb 미만(예를 들어, 약 19 kb, 18 kb, 17 kb, 16 kb, 15 kb, 14 kb, 13 kb, 12 kb, 11 kb, 10 kb, 9 kb, 8 kb, 7 kb, 6 kb, 5 kb, 4 kb, 3 kb, 2 kb, 1 kb 또는 그 미만)의 길이를 갖는다. 일부 구현예에서, 유전 요소는 독립적으로 또는 부가적으로 1000b 초과(예를 들어, 적어도 약 1.1 kb, 1.2 kb, 1.3 kb, 1.4 kb, 1.5 kb, 1.6 kb, 1.7 kb, 1.8 kb, 1.9 kb, 2 kb, 2.1 kb, 2.2 kb, 2.3 kb, 2.4 kb, 2.5 kb, 2.6 kb, 2.7 kb, 2.8 kb, 2.9 kb, 3 kb, 3.1 kb, 3.2 kb, 3.3 kb, 3.4 kb, 3.5 kb, 3.6 kb, 3.7 kb, 3.8 kb, 3.9 kb, 4 kb, 4.1 kb, 4.2 kb, 4.3 kb, 4.4 kb, 4.5 kb, 4.6 kb, 4.7 kb, 4.8 kb, 4.9 kb, 5 kb 또는 그 초과)의 길이를 갖는다. 일부 구현예에서, 유전 요소는 약 2.5 내지 4.6, 2.8 내지 4.0, 3.0 내지 3.8 또는 3.2 내지 3.7 kb의 길이를 갖는다. 일부 구현예에서, 유전 요소는 약 1.5 내지 2.0, 1.5 내지 2.5, 1.5 내지 3.0, 1.5 내지 3.5, 1.5 내지 3.8, 1.5 내지 3.9, 1.5 내지 4.0, 1.5 내지 4.5 또는 1.5 내지 5.0 kb의 길이를 갖는다. 일부 구현예에서, 유전 요소는 약 2.0 내지 2.5, 2.0 내지 3.0, 2.0 내지 3.5, 2.0 내지 3.8, 2.0 내지 3.9, 2.0 내지 4.0, 2.0 내지 4.5 또는 2.0 내지 5.0 kb의 길이를 갖는다. 일부 구현예에서, 유전 요소는 약 2.5 내지 3.0, 2.5 내지 3.5, 2.5 내지 3.8, 2.5 내지 3.9, 2.5 내지 4.0, 2.5 내지 4.5 또는 2.5 내지 5.0 kb의 길이를 갖는다. 일부 구현예에서, 유전 요소는 약 3.0 내지 5.0, 3.5 내지 5.0, 4.0 내지 5.0 또는 4.5 내지 5.0 kb의 길이를 갖는다. 일부 구현예에서, 유전 요소는 약 1.5 내지 2.0, 2.0 내지 2.5, 2.5 내지 3.0, 3.0 내지 3.5, 3.1 내지 3.6, 3.2 내지 3.7, 3.3 내지 3.8, 3.4 내지 3.9, 3.5 내지 4.0, 4.0 내지 4.5 또는 4.5 내지 5.0 kb의 길이를 갖는다. 일부 구현예에서, 유전 요소는 약 3.6 내지 3.9 kb의 길이를 갖는다. 일부 구현예에서, 유전 요소는 약 2.8 내지 2.9 kb의 길이를 갖는다. 일부 구현예에서, 유전 요소는 약 2.0 내지 3.2 kb의 길이를 갖는다.

[0739] 일부 구현예에서, 유전 요소는 본 명세서에 기재된 특징, 예를 들어, 실질적으로 비-병원성인 단백질을 인코딩하는 서열, 단백질 결합 서열, 조절 핵산을 인코딩하는 하나 이상의 서열, 하나 이상의 조절 서열, 복제 단백질을 인코딩하는 하나 이상의 서열, 및 기타 서열 중 하나 이상을 포함한다.

[0740] 구현예에서, 유전 요소를 (예를 들어, 시험관 내 환화에 의해 생성되는) 이중-가닥 환형 DNA로부터 생성하였다. 일부 구현예에서, 유전 요소를 이중-가닥 환형 DNA로부터 회전환 복제에 의해 생성하였다. 구현예에서, 회전환 복제는 세포(예를 들어, 숙주 세포, 예를 들어, 포유동물 세포, 예를 들어, 인간 세포, 예를 들어, HEK293T 세포, A549 세포 또는 Jurkat 세포)에서 발생한다. 구현예에서, 유전 요소는 세포에서 회전환 복제에 의해 지수적으로 복제될 수 있다. 구현예에서, 유전 요소는 세포에서 회전환 복제에 의해 선형적으로 복제될 수 있다. 구현예에서, 이중-가닥 환형 DNA 또는 유전 요소는 세포에서 회전환 복제에 의해 원래의 양의 적어도 2, 4, 8, 16,

32, 64, 128, 256, 518, 1024배 이상을 제공할 수 있다. 구현예에서, 이중-가닥 환형 DNA를 예를 들어, 본 명세서에 기재된 바와 같이, 세포 내로 도입하였다.

[0741] 일부 구현예에서, 이중-가닥 환형 DNA 및/또는 유전 요소는 하나 이상의 박테리아 플라스미드 요소(예를 들어, 박테리아 복제 원점 또는 선택 가능한 마커, 예를 들어, 박테리아 내성 유전자)를 포함하지 않는다. 일부 구현예에서, 이중-가닥 환형 DNA 및/또는 유전 요소는 박테리아 플라스미드 백본을 포함하지 않는다.

[0742] 일 구현예에서, 본 발명은 (i) 실질적으로 비-병원성인 외부 단백질, (ii) 유전 요소를 실질적으로 비-병원성인 외부 단백질에 결합시키는 외부 단백질 결합 서열 및 (iii) 조절 핵산을 인코딩하는 핵산 서열(예를 들어, DNA 서열)을 포함하는 유전 요소를 포함한다. 이러한 구현예에서, 유전 요소는 고유 바이러스 서열(예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어, 고유 아넬로바이러스 서열)에 대한 뉴클레오티드 서열 중 어느 하나에 대하여 적어도 약 60%, 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98% 및 99% 뉴클레오티드 서열 동일성을 갖는 하나 이상의 서열을 포함할 수 있다.

[0743] 단백질 결합 서열

[0744] 많은 바이러스에 의해 사용되는 전략은 바이러스 캡시드 단백질이 이의 게놈 내의 특정 단백질 결합 서열을 인식하는 것이다. 예를 들어, 비세그먼트화된 게놈을 갖는 바이러스, 예컨대 효모의 L-A 바이러스에서, 게놈의 5' 말단에 이차 구조(스템-루프) 및 특정 서열이 존재하며, 이는 둘 모두 바이러스 캡시드 단백질에 결합하는 데 사용된다. 그러나, 세그먼트화된 게놈을 갖는 바이러스, 예컨대 레오바이러스과, 오르쏘믹소바이러스과(*Orthomyxoviridae*)(인플루엔자), 부니아바이러스(*Bunyaviruse*) 및 아레나바이러스(*Arenaviruse*)는 게놈 세그먼트의 각각을 패키징할 필요가 있다. 일부 바이러스는 세그먼트의 상보성 영역을 사용하여, 바이러스가 각각의 게놈 분자 중 하나를 포함시키는 것을 보조한다. 다른 바이러스는 상이한 세그먼트의 각각에 대하여 특이적인 결합 부위를 갖는다. 예를 들어, 문헌[Curr Opin Struct Biol. 2010 Feb; 20(1): 114-120]; 및 문헌[Journal of Virology (2003), 77(24), 13036-13041]을 참조한다.

[0745] 일부 구현예에서, 유전 요소는 실질적으로 비-병원성인 단백질에 결합하는 단백질 결합 서열을 인코딩한다. 일부 구현예에서, 단백질 결합 서열은 단백질성 외부로의 유전 요소의 패키징을 용이하게 한다. 일부 구현예에서, 단백질 결합 서열은 실질적으로 비-병원성인 단백질의 아르기닌-풍부 영역에 특이적으로 결합한다. 일부 구현예에서, 유전 요소는 PCT/US19/65995의 실시예 8에 기재된 바와 같은 단백질 결합 서열을 포함한다.

[0746] 일부 구현예에서, 유전 요소는, 예를 들어, 본 명세서에 기재된 바와 같은, 아넬로바이러스 서열의 5' UTR 보존된 도메인 또는 GC-풍부 도메인에 대하여 적어도 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 단백질 결합 서열을 포함한다.

[0747] 구현예에서, 단백질 결합 서열은, 예를 들어, 본 명세서에 기재된 바와 같은, 아넬로바이러스 5' UTR 보존된 도메인 뉴클레오티드 서열에 대하여 적어도 약 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는다.

[0748] *5' UTR 영역*

[0749] 일부 구현예에서, 본 명세서에 기재된 바와 같은 핵산 분자(예를 들어, 유전 요소, 유전 요소 작제물, 또는 유전 요소 영역)는 5' UTR 서열, 예를 들어, 본 명세서에 기재된 바와 같은 5' UTR 보존된 도메인 서열(예를 들어, 표 A1, B1, 또는 C1의 임의의 것), 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 서열 동일성을 갖는 서열을 포함한다.

[0750] 일부 구현예에서, 5' UTR 서열은 핵산 서열 AGGTGAGTGAAACCACCGAAGTCAAGGGGCAATTCGGGCTAGGGX_1CAGTCT, 또는 이에 대하여 적어도 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 서열 동일성을 갖는 핵산 서열을 포함한다. 일부 구현예에서, 5' UTR 서열은 핵산 서열 AGGTGAGTGAAACCACCGAAGTCAAGGGGCAATTCGGGCTAGGGX_1CAGTCT, 또는 이에 대하여 1, 2, 3, 4, 5, 6, 7, 8, 9, 또는 10개 이하의 뉴클레오티드 차이(예를 들어, 치환, 결실, 또는 추가)를 갖는 핵산 서열을 포함한다. 구현예에서, X_1은 A이다. 구현예에서, X_1은 존재하지 않는다.

[0751] 일부 구현예에서, 5' UTR 서열은 알파토르크바이러스(예를 들어, Ring1)의 5' UTR의 핵산 서열, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 서열 동일성을 갖는 서열을 포함한다. 구현예에서, 5' UTR 서열은 표 A1에 열거된 5' UTR 보존된 도메인의 핵산 서열, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 서열 동일성을 갖는 서열을 포함한다. 일부 구현예에서, 핵산 분자는

표 A1에 열거된 5' UTR 보존된 도메인에 대하여 적어도 95% 서열 동일성을 갖는 핵산 서열을 포함한다. 일부 구현예에서, 핵산 분자는 표 A1에 열거된 5' UTR 보존된 도메인에 대하여 적어도 95.775% 서열 동일성을 갖는 핵산 서열을 포함한다. 일부 구현예에서, 핵산 분자는 표 A1에 열거된 5' UTR 보존된 도메인에 대하여 적어도 97% 서열 동일성을 갖는 핵산 서열을 포함한다. 일부 구현예에서, 핵산 분자는 표 A1에 열거된 5' UTR 보존된 도메인에 대하여 적어도 97.183% 서열 동일성을 갖는 핵산 서열을 포함한다. 일부 구현예에서, 5' UTR 서열은 핵산 서열 AGGTGAGTTTACACACCGCAGTCAAGGGGCAATTCGGGCTCGGGACTGGC, 또는 이에 대하여 적어도 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 서열 동일성을 갖는 핵산 서열을 포함한다. 일부 구현예에서, 5' UTR 서열은 핵산 서열 AGGTGAGTTTACACACCGCAGTCAAGGGGCAATTCGGGCTCGGGACTGGC, 또는 이에 대하여 1, 2, 3, 4, 5, 6, 7, 8, 9, 또는 10개 이하의 뉴클레오티드 차이(예를 들어, 치환, 결실, 또는 추가)를 갖는 핵산 서열을 포함한다.

[0752] 일부 구현예에서, 5' UTR 서열은 베타토르크바이러스(예를 들어, Ring2)의 5' UTR의 핵산 서열, 또는 이에 대하여 적어도 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99% 서열 동일성을 갖는 서열을 포함한다. 구현예에서, 5' UTR 서열은 표 B1에 열거된 5' UTR 보존된 도메인의 핵산 서열, 또는 이에 대하여 적어도 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99% 서열 동일성을 갖는 서열을 포함한다. 일부 구현예에서, 핵산 분자는 표 B1에 열거된 5' UTR 보존된 도메인에 대하여 적어도 85% 서열 동일성을 갖는 핵산 서열을 포함한다. 일부 구현예에서, 핵산 분자는 표 B1에 열거된 5' UTR 보존된 도메인에 대하여 적어도 87% 서열 동일성을 갖는 핵산 서열을 포함한다. 일부 구현예에서, 핵산 분자는 표 B1에 열거된 5' UTR 보존된 도메인에 대하여 적어도 87.324% 서열 동일성을 갖는 핵산 서열을 포함한다. 일부 구현예에서, 핵산 분자는 표 B1에 열거된 5' UTR 보존된 도메인에 대하여 적어도 88% 서열 동일성을 갖는 핵산 서열을 포함한다. 일부 구현예에서, 핵산 분자는 표 B1에 열거된 5' UTR 보존된 도메인에 대하여 적어도 88.732% 서열 동일성을 갖는 핵산 서열을 포함한다. 일부 구현예에서, 핵산 분자는 표 B1에 열거된 5' UTR 보존된 도메인에 대하여 적어도 91% 서열 동일성을 갖는 핵산 서열을 포함한다. 일부 구현예에서, 핵산 분자는 표 B1에 열거된 5' UTR 보존된 도메인에 대하여 적어도 91.549% 서열 동일성을 갖는 핵산 서열을 포함한다. 일부 구현예에서, 핵산 분자는 표 B1에 열거된 5' UTR 보존된 도메인에 대하여 적어도 92% 서열 동일성을 갖는 핵산 서열을 포함한다. 일부 구현예에서, 핵산 분자는 표 B1에 열거된 5' UTR 보존된 도메인에 대하여 적어도 92.958% 서열 동일성을 갖는 핵산 서열을 포함한다. 일부 구현예에서, 핵산 분자는 표 B1에 열거된 5' UTR 보존된 도메인에 대하여 적어도 94% 서열 동일성을 갖는 핵산 서열을 포함한다. 일부 구현예에서, 핵산 분자는 표 B1에 열거된 5' UTR 보존된 도메인에 대하여 적어도 94.366% 서열 동일성을 갖는 핵산 서열을 포함한다. 일부 구현예에서, 핵산 분자는 표 B1에 열거된 5' UTR 보존된 도메인에 대하여 적어도 95% 서열 동일성을 갖는 핵산 서열을 포함한다. 일부 구현예에서, 핵산 분자는 표 B1에 열거된 5' UTR 보존된 도메인에 대하여 적어도 95.775% 서열 동일성을 갖는 핵산 서열을 포함한다. 일부 구현예에서, 핵산 분자는 표 B1에 열거된 5' UTR 보존된 도메인에 대하여 적어도 97% 서열 동일성을 갖는 핵산 서열을 포함한다. 일부 구현예에서, 핵산 분자는 표 B1에 열거된 5' UTR 보존된 도메인에 대하여 적어도 97.183% 서열 동일성을 갖는 핵산 서열을 포함한다. 일부 구현예에서, 5' UTR 서열은 핵산 서열 AGGTGAGTGAAACCACCGAAGTCAAGGGGCAATTCGGGCTAGATCAGTCT, 또는 이에 대하여 적어도 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 서열 동일성을 갖는 핵산 서열을 포함한다. 일부 구현예에서, 5' UTR 서열은 핵산 서열 AGGTGAGTGAAACCACCGAAGTCAAGGGGCAATTCGGGCTAGATCAGTCT, 또는 이에 대하여 1, 2, 3, 4, 5, 6, 7, 8, 9, 또는 10개 이하의 뉴클레오티드 차이(예를 들어, 치환, 결실, 또는 추가)를 갖는 핵산 서열을 포함한다.

[0753] 일부 구현예에서, 5' UTR 서열은 감마토르크바이러스(예를 들어, Ring4)의 5' UTR의 핵산 서열, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 서열 동일성을 갖는 서열을 포함한다. 구현예에서, 5' UTR 서열은 표 C1에 열거된 5' UTR 보존된 도메인의 핵산 서열, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 서열 동일성을 갖는 서열을 포함한다. 일부 구현예에서, 핵산 분자는 표 C1에 열거된 5' UTR 보존된 도메인에 대하여 적어도 97% 서열 동일성을 갖는 핵산 서열을 포함한다. 일부 구현예에서, 핵산 분자는 표 C1에 열거된 5' UTR 보존된 도메인에 대하여 적어도 97.183% 서열 동일성을 갖는 핵산 서열을 포함한다. 일부 구현예에서, 5' UTR 서열은 핵산 서열 AGGTGAGTGAAACCACCGAGGTCTAGGGGCAATTCGGGCTAGGGCAGTCT, 또는 이에 대하여 적어도 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 서열 동일성을 갖는 핵산 서열을 포함한다. 일부 구현예에서, 5' UTR 서열은 핵산 서열 AGGTGAGTGAAACCACCGAGGTCTAGGGGCAATTCGGGCTAGGGCAGTCT, 또는 이에 대하여 1, 2, 3, 4, 5, 6, 7, 8, 9, 또는 10개 이하의 뉴클레오티드 차이(예를 들어, 치환, 결실, 또는 추가)를 갖는 핵산 서열을 포함한다.

[0754] 일부 구현예에서, 유전 요소(예를 들어, 유전 요소의 단백질-결합 서열)는 아넬로바이러스 5' UTR 서열, 예를 들어 표 38에 나타낸 핵산 서열에 대하여 적어도 약 75%(예를 들어, 적어도 75%, 80%, 85%, 90%, 95%, 96%,

97%, 98%, 99% 또는 100%) 동일성을 갖는 핵산 서열을 포함한다. 일부 구현예에서, 유전 요소(예를 들어, 유전 요소의 단백질-결합 서열)는 표 38에 나타낸 컨센서스 5' UTR 서열의 핵산 서열을 포함하며, X_1, X_2, X_3, X_4 및 X_5는 각각 독립적으로, 임의의 뉴클레오티드이며, 예를 들어, X_1 = G 또는 T이며, X_2 = C 또는 A이며, X_3 = G 또는 A이며, X_4 = T 또는 C이며, X_5 = A, C 또는 T이다. 구현예에서, 유전 요소(예를 들어, 유전 요소의 단백질-결합 서열)는 표 38에 나타낸 컨센서스 5' UTR 서열에 대하여 적어도 약 75%(예를 들어, 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%) 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 유전 요소(예를 들어, 유전 요소의 단백질-결합 서열)는 표 38에 나타낸 예시적인 TTV 5' UTR 서열에 대하여 적어도 약 75%(예를 들어, 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%) 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 유전 요소(예를 들어, 유전 요소의 단백질-결합 서열)는 표 38에 나타낸 TTV-CT30F 5' UTR 서열에 대하여 적어도 약 75%(예를 들어, 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%) 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 유전 요소(예를 들어, 유전 요소의 단백질-결합 서열)는 표 38에 나타낸 TTV-HD23a 5' UTR 서열에 대하여 적어도 약 75%(예를 들어, 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%) 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 유전 요소(예를 들어, 유전 요소의 단백질-결합 서열)는 표 38에 나타낸 TTV-JA20 5' UTR 서열에 대하여 적어도 약 75%(예를 들어, 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%) 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 유전 요소(예를 들어, 유전 요소의 단백질-결합 서열)는 표 38에 나타낸 TTV-TJN02 5' UTR 서열에 대하여 적어도 약 75%(예를 들어, 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%) 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 유전 요소(예를 들어, 유전 요소의 단백질-결합 서열)는 표 38에 나타낸 TTV-tth8 5' UTR 서열에 대하여 적어도 약 75%(예를 들어, 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%) 동일성을 갖는 핵산 서열을 포함한다.

[0755] 구현예에서, 유전 요소(예를 들어, 유전 요소의 단백질-결합 서열)는 표 38에 나타낸 알파토르크바이러스 컨센서스 5' UTR 서열에 대하여 적어도 약 75%(예를 들어, 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%) 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 유전 요소(예를 들어, 유전 요소의 단백질-결합 서열)는 표 38에 나타낸 알파토르크바이러스 클레이드 1 5' UTR 서열에 대하여 적어도 약 75%(예를 들어, 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%) 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 유전 요소(예를 들어, 유전 요소의 단백질-결합 서열)는 표 38에 나타낸 알파토르크바이러스 클레이드 2 5' UTR 서열에 대하여 적어도 약 75%(예를 들어, 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%) 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 유전 요소(예를 들어, 유전 요소의 단백질-결합 서열)는 표 38에 나타낸 알파토르크바이러스 클레이드 3 5' UTR 서열에 대하여 적어도 약 75%(예를 들어, 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%) 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 유전 요소(예를 들어, 유전 요소의 단백질-결합 서열)는 표 38에 나타낸 알파토르크바이러스 클레이드 4 5' UTR 서열에 대하여 적어도 약 75%(예를 들어, 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%) 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 유전 요소(예를 들어, 유전 요소의 단백질-결합 서열)는 표 38에 나타낸 알파토르크바이러스 클레이드 5 5' UTR 서열에 대하여 적어도 약 75%(예를 들어, 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%) 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 유전 요소(예를 들어, 유전 요소의 단백질-결합 서열)는 표 38에 나타낸 알파토르크바이러스 클레이드 6 5' UTR 서열에 대하여 적어도 약 75%(예를 들어, 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%) 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 유전 요소(예를 들어, 유전 요소의 단백질-결합 서열)는 표 38에 나타낸 알파토르크바이러스 클레이드 7 5' UTR 서열에 대하여 적어도 약 75%(예를 들어, 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%) 동일성을 갖는 핵산 서열을 포함한다.

[0756] **[표 38] 아넬로바이러스 유래의 예시적인 5' UTR 서열**

공급원	서열	서열번호
컨센서스	CGGGTGCCGX$_1$AGGTGAGTTTACACACCGX$_2$AGT CAAGGGGCAATTCGGGCTCX$_3$GGACTGGCCGGG CX$_4$X$_5$TGGG X$_1$ = G 또는 T X$_2$ = C 또는 A X$_3$ = G 또는 A X$_4$ = T 또는 C X$_5$ = A, C, 또는 T	105
예시적인 TTV 서열	CGGGTGCCGGAGGTGAGTTTACACACCGCAGTC AAGGGGCAATTCGGGCTCGGGACTGGCCGGGCT WTGGG	106
TTV-CT30F	CGGGTGCCGTAGGTGAGTTTACACACCGCAGTC AAGGGGCAATTCGGGCTCGGGACTGGCCGGGCT ATGGG	107
TTV-HD23a	CGGGTGCCGGAGGTGAGTTTACACACCGCAGTC AAGGGGCAATTCGGGCTCGGGACTGGCCGGGCC CTGGG	108
TTV-JA20	CGGGTGCCGGAGGTGAGTTTACACACCGCAGTC AAGGGGCAATTCGGGCTCGGGACTGGCCGGGCT TTGGG	109
TTV-TJN02	CGGGTGCCGGAGGTGAGTTTACACACCGCAGTC AAGGGGCAATTCGGGCTCGGGACTGGCCGGGCT ATGGG	110
TTV-tth8	CGGGTGCCGGAGGTGAGTTTACACACCGAAGTC AAGGGGCAATTCGGGCTCAGGACTGGCCGGGCT TTGGG	111
알파토르크바이러스 컨센서스 5' UTR	CGGGTGCCGGAGGTGAGTTTACACACCGCAGTC AAGGGGCAATTCGGGCTCGGGACTGGCCGGGC X$_1$X$_2$TGGG; 여기서 X$_1$은 T 또는 C를 포함하고, X$_2$는 A, C, 또는 T를 포함함.	112
알파토르크바이러스 클레이드 1 5' UTR (예: TTV-CT30F)	CGGGTGCCGTAGGTGAGTTTACACACCGCAGTC AAGGGGCAATTCGGGCTCGGGACTGGCCGGGCT ATGGG	113
알파토르크바이러스 클레이드 2 5' UTR (예: TTV-P13-1)	CGGGTGCCGGAGGTGAGTTTACACACCGCAGTC AAGGGGCAATTCGGGCTCGGGACTGGCCGGGCC CGGG	114
알파토르크바이러스 클레이드 3 5' UTR (예: TTV-tth8)	CGGGTGCCGGAGGTGAGTTTACACACCGAAGTC AAGGGGCAATTCGGGCTCAGGACTGGCCGGGCT TTGGG	115
알파토르크바이러스 클레이드 4 5' UTR (예: TTV-HD20a)	CGGGTGCCGGAGGTGAGTTTACACACCGCAGTC AAGGGGCAATTCGGGCTCGGGAGGCCGGGCCAT GGG	116
알파토르크바이러스 클레이드 5 5' UTR (예: TTV-16)	CGGGTGCCGGAGGTGAGTTTACACACCGCAGTC AAGGGGCAATTCGGGCTCGGGACTGGCCGGGCC CCGGG	117
알파토르크바이러스 클레이드 6 5' UTR (예: TTV-TJN02)	CGGGTGCCGGAGGTGAGTTTACACACCGCAGTC AAGGGGCAATTCGGGCTCGGGACTGGCCGGGCT ATGGG	118
알파토르크바이러스 클레이드 7 5' UTR (예: TTV-HD16d)	CGGGTGCCGAAGGTGAGTTTACACACCGCAGTC AAGGGGCAATTCGGGCTCGGGACTGGCCGGGCT ATGGG	119

[0757]

[0758] 일부 구현예에서, 아넬로바이러스 5' UTR 서열은 아넬로바이러스의 게놈(예를 들어, 핵산 서열결정 기법, 예를 들어, 심층 서열결정 기법에 의해 확인된, 예를 들어, 추정의 아넬로바이러스 게놈) 내에서 확인될 수 있다. 일부 구현예에서, 아넬로바이러스 5' UTR 서열은 하기 단계 중 하나 또는 둘 다에 의해 확인된다:

[0759] *(i) 환화 접합점의 확인:* 일부 구현예에서, 5' UTR은 전장의 환화된 아넬로바이러스 게놈의 환화 접합점 근처에 위치할 것이다. 예를 들어, 서열의 중첩 영역을 확인함으로써 환화 접합점을 확인할 수 있다. 일부 구현예에서, 서열의 중첩 영역은 환화된 전장 아넬로바이러스 게놈 서열을 생성하기 위해 서열로부터 트리밍될 수 있다. 일부 구현예에서, 게놈 서열은 소프트웨어를 사용하여 이러한 방식으로 환화된다. 이론에 결부시키고자 하지 않고, 게놈을 컴퓨터로 환화하면 서열의 시작 위치가 비-생물학적으로 배향될 수 있다. 서열 내의 표지점을 사용하여 서열을 올바른 방향으로 재배향할 수 있다. 예를 들어, 표지점 서열은 본 명세서에 기재된 바와 같은 아넬로바이러스 게놈 내의 하나 이상의 요소(예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어 아넬로바이러

스의 TATA 박스, cap 부위, 개시인자 요소, 전사 시작 부위, 5' UTR 보존된 도메인, ORF1, ORF1/1, ORF1/2, ORF2, ORF2/2, ORF2/3, ORF2t/3, 3개의 오픈-리딩 프레임 영역, 폴리(A) 신호, 또는 GC-풍부 영역 중 하나 이상)에 실질적인 상동성을 갖는 서열을 포함할 수 있다.

[0760] *(ii) 5' UTR 서열의 확인:* 일단 추정 아넬로바이러스 게놈 서열이 얻어지면, 서열(또는 이의 부분, 예를 들어, 약 40 내지 50, 50 내지 60, 60 내지 70, 70 내지 80, 80 내지 90, 또는 90 내지 100개 뉴클레오티드의 길이를 가짐)을 하나 이상의 아넬로바이러스 5' UTR 서열(예를 들어, 본 명세서에 기재된 바와 같음)과 비교하여 이에 대한 실질적인 상동성을 갖는 서열을 확인할 수 있다. 일부 구현예에서, 추정 아넬로바이러스 5' UTR 영역은 본 명세서에 기재된 바와 같은 아넬로바이러스 5' UTR에 대하여 적어도 50%, 60%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는다.

[0761] *GC-풍부 영역*

[0762] 일부 구현예에서, 유전 요소(예를 들어, 유전 요소의 단백질-결합 서열)는 표 39에 나타낸 핵산 서열에 대하여 적어도 약 75%(예를 들어, 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%) 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 유전 요소(예를 들어, 유전 요소의 단백질-결합 서열)는 표 39에 나타낸 GC-풍부 서열에 대하여 적어도 약 75%(예를 들어, 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%) 동일성을 갖는 핵산 서열을 포함한다.

[0763] 구현예에서, 유전 요소(예를 들어, 유전 요소의 단백질-결합 서열)는 표 39에 나타낸 바와 같은 36-뉴클레오티드 GC-풍부 서열(예를 들어, 36-뉴클레오티드 컨센서스 GC-풍부 영역 서열 1, 36-뉴클레오티드 컨센서스 GC-풍부 영역 서열 2, TTV 클레이드 1 36-뉴클레오티드 영역, TTV 클레이드 3 36-뉴클레오티드 영역, TTV 클레이드 3 분리주 GH1 36-뉴클레오티드 영역, TTV 클레이드 3 sle1932 36-뉴클레오티드 영역, TTV 클레이드 4 ctdc002 36-뉴클레오티드 영역, TTV 클레이드 5 36-뉴클레오티드 영역, TTV 클레이드 6 36-뉴클레오티드 영역 또는 TTV 클레이드 7 36-뉴클레오티드 영역)에 대하여 적어도 약 75%(예를 들어, 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%) 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 유전 요소(예를 들어, 유전 요소의 단백질-결합 서열)는 표 39에 나타낸 바와 같은 36-뉴클레오티드 GC-풍부 서열(예를 들어, 36-뉴클레오티드 컨센서스 GC-풍부 영역 서열 1, 36-뉴클레오티드 컨센서스 GC-풍부 영역 서열 2, TTV 클레이드 1 36-뉴클레오티드 영역, TTV 클레이드 3 36-뉴클레오티드 영역, TTV 클레이드 3 분리주 GH1 36-뉴클레오티드 영역, TTV 클레이드 3 sle1932 36-뉴클레오티드 영역, TTV 클레이드 4 ctdc002 36-뉴클레오티드 영역, TTV 클레이드 5 36-뉴클레오티드 영역, TTV 클레이드 6 36-뉴클레오티드 영역 또는 TTV 클레이드 7 36-뉴클레오티드 영역)의 적어도 10, 15, 20, 25, 30, 31, 32, 33, 34, 35 또는 36개의 인접 뉴클레오티드를 포함하는 핵산 서열을 포함한다.

[0764] 구현예에서, 유전 요소(예를 들어, 유전 요소의 단백질-결합 서열)는 예를 들어, 표 39에 열거된 바와 같은, 예를 들어, TTV-CT30F, TTV-P13-1, TTV-tth8, TTV-HD20a, TTV-16, TTV-TJN02 또는 TTV-HD16d로부터 선택되는 알파토르크바이러스 GC-풍부 영역 서열에 대하여 적어도 약 75%(예를 들어, 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%) 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 유전 요소(예를 들어, 유전 요소의 단백질-결합 서열)는 예를 들어, 표 39에 열거된 바와 같은, 예를 들어, TTV-CT30F, TTV-P13-1, TTV-tth8, TTV-HD20a, TTV-16, TTV-TJN02 또는 TTV-HD16d로부터 선택되는 알파토르크바이러스 GC-풍부 영역 서열의 적어도 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 104, 105, 108, 110, 111, 115, 120, 122, 130, 140, 145, 150, 155 또는 156개의 인접 뉴클레오티드를 포함하는 핵산 서열을 포함한다.

[0765] 구현예에서, 36-뉴클레오티드 GC-풍부 서열은 하기로부터 선택된다:

[0766] (i) CGCGCTGCGCGCGCCGCCCAGTAGGGGGAGCCATGC(서열번호 160),

[0767] (ii) GCGCTX$_1$CGCGCGCGCGCCGGGGGGCTGCGCCCCCCC(서열번호 164), 여기서, X$_1$은 T, G 또는 A로부터 선택됨;

[0768] (iii) GCGCTTCGCGCGCCGCCCACTAGGGGGCGTTGCGCG(서열번호 165);

[0769] (iv) GCGCTGCGCGCGCCGCCCAGTAGGGGGCGCAATGCG(서열번호 166);

[0770] (v) GCGCTGCGCGCGCGGCCCCCGGGGGAGGCATTGCCT(서열번호 167);

[0771] (vi) GCGCTGCGCGCGCGCGCCGGGGGGGCGCCAGCGCCC(서열번호 168);

[0772] (vii) GCGCTTCGCGCGCGCGCCGGGGGGCTCCGCCCCCCC(서열번호 169);

[0773] (viii) GCGCTTCGCGCGCGCGCCGGGGGGCTGCGCCCCCCC(서열번호 170);

[0774] (ix) GCGCTACGCGCGCGCGCCGGGGGGCTGCGCCCCCCC(서열번호 171); 또는

[0775] (x) GCGCTACGCGCGCGCGCCGGGGGGCTCTGCCCCCCC(서열번호 172).

[0776] 구현예에서, 유전 요소(예를 들어, 유전 요소의 단백질-결합 서열)는 핵산 서열 CGCGCTGCGCGCGCCGCCCAGTAGGGGGAGCCATGC(서열번호 160)를 포함한다.

[0777] 구현예에서, 유전 요소(예를 들어, 유전 요소의 단백질-결합 서열)는 표 39에 나타낸 컨센서스 GC-풍부 서열의 핵산 서열을 포함하며, X_1, X_4, X_5, X_6, X_7, X_{12}, X_{13}, X_{14}, X_{15}, X_{20}, X_{21}, X_{22}, X_{26}, X_{29}, X_{30} 및 X_{33}은 각각 독립적으로 임의의 뉴클레오티드이며, X_2, X_3, X_8, X_9, X_{10}, X_{11}, X_{16}, X_{17}, X_{18}, X_{19}, X_{23}, X_{24}, X_{25}, X_{27}, X_{28}, X_{31}, X_{32} 및 X_{34}는 각각 독립적으로 부재이거나, 임의의 뉴클레오티드이다. 일부 구현예에서, X_1 내지 X_{34} 중 하나 이상(예를 들어, 이들 전부)는 각각 독립적으로 표 39에 특정된 뉴클레오티드(또는 부재)이다. 구현예에서, 유전 요소(예를 들어, 유전 요소의 단백질-결합 서열)는 표 39에 나타낸 예시적인 TTV GC-풍부 서열(예를 들어, 전체 서열, 단편 1, 단편 2, 단편 3 또는 이의 임의의 조합, 예를 들어, 순서대로 단편 1 내지 3)에 대하여 적어도 약 75%(예를 들어, 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%) 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 유전 요소(예를 들어, 유전 요소의 단백질-결합 서열)는 표 39에 나타낸 TTV-CT30F GC-풍부 서열(예를 들어, 전체 서열, 단편 1, 단편 2, 단편 3, 단편 4, 단편 5, 단편 6, 단편 7, 단편 8 또는 이의 임의의 조합, 예를 들어, 순서대로 단편 1 내지 7)에 대하여 적어도 약 75%(예를 들어, 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%) 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 유전 요소(예를 들어, 유전 요소의 단백질-결합 서열)는 표 39에 나타낸 TTV-HD23a GC-풍부 서열(예를 들어, 전체 서열, 단편 1, 단편 2, 단편 3, 단편 4, 단편 5, 단편 6 또는 이의 임의의 조합, 예를 들어, 순서대로 단편 1 내지 6)에 대하여 적어도 약 75%(예를 들어, 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%) 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 유전 요소(예를 들어, 유전 요소의 단백질-결합 서열)는 표 39에 나타낸 TTV-JA20 GC-풍부 서열(예를 들어, 전체 서열, 단편 1, 단편 2 또는 이의 임의의 조합, 예를 들어, 순서대로 단편 1 및 2)에 대하여 적어도 약 75%(예를 들어, 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%) 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 유전 요소(예를 들어, 유전 요소의 단백질-결합 서열)는 표 39에 나타낸 TTV-TJN02 GC-풍부 서열(예를 들어, 전체 서열, 단편 1, 단편 2, 단편 3, 단편 4, 단편 5, 단편 6, 단편 7, 단편 8 또는 이의 임의의 조합, 예를 들어, 순서대로 단편 1 내지 8)에 대하여 적어도 약 75%(예를 들어, 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%) 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 유전 요소(예를 들어, 유전 요소의 단백질-결합 서열)는 표 39에 나타낸 TTV-tth8 GC-풍부 서열(예를 들어, 전체 서열, 단편 1, 단편 2, 단편 3, 단편 4, 단편 5, 단편 6, 단편 7, 단편 8, 단편 9 또는 이의 임의의 조합, 예를 들어, 순서대로 단편 1 내지 6)에 대하여 적어도 약 75%(예를 들어, 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%) 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 유전 요소(예를 들어, 유전 요소의 단백질-결합 서열)는 표 39에 나타낸 단편 7에 대하여 적어도 약 75%(예를 들어, 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%) 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 유전 요소(예를 들어, 유전 요소의 단백질-결합 서열)는 표 39에 나타낸 단편 8에 대하여 적어도 약 75%(예를 들어, 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%) 동일성을 갖는 핵산 서열을 포함한다. 구현예에서, 유전 요소(예를 들어, 유전 요소의 단백질-결합 서열)는 표 39에 나타낸 단편 9에 대하여 적어도 약 75%(예를 들어, 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%) 동일성을 갖는 핵산 서열을 포함한다.

[0778] [표 39] 아넬로바이러스 유래의 예시적인 GC-풍부 서열

공급원	서열	서열번호
컨센서스	CGGCGGX$_1$GGX$_2$GX$_3$X$_4$X$_5$CGCGCTX$_6$CGCGC GCX$_7$X$_8$X$_9$X$_{10}$CX$_{11}$X$_{12}$X$_{13}$X$_{14}$GGGGX$_{15}$X$_{16}$X$_{17}$X$_{18}$ X$_{19}$X$_{20}$X$_{21}$GCX$_{22}$X$_{23}$X$_{24}$X$_{25}$CCCCCCCX$_{26}$CGCGC ATX$_{27}$X$_{28}$GCX$_{29}$CGGGX$_{30}$CCCCCCCCCX$_{31}$X$_{32}$X $_{33}$GGGGGGCTCCGX$_{34}$CCCCCCGGCCCCCC X$_1$ = G 또는 C X$_2$ = G, C, 또는 부재 X$_3$ = C 또는 부재 X$_4$ = G 또는 C X$_5$ = G 또는 C X$_6$ = T, G, 또는 A X$_7$ = G 또는 C X$_8$ = G 또는 부재 X$_9$ = C 또는 부재 X$_{10}$ = C 또는 부재 X$_{11}$ = G, A, 또는 부재 X$_{12}$ = G 또는 C X$_{13}$ = C 또는 T	120

[0779]

		X_{14} = G 또는 A X_{15} = G 또는 A X_{16} = A, G, T, 또는 부재 X_{17} = G, C, 또는 부재 X_{18} = G, C, 또는 부재 X_{19} = C, A, 또는 부재 X_{20} = C 또는 A X_{21} = T 또는 A X_{22} = G 또는 C X_{23} = G, T, 또는 부재 X_{24} = C 또는 부재 X_{25} = G, C, 또는 부재 X_{26} = G 또는 C X_{27} = G 또는 부재 X_{28} = C 또는 부재 X_{29} = G 또는 A X_{30} = G 또는 T X_{31} = C, T, 또는 부재 X_{32} = G, C, A, 또는 부재 X_{33} = G 또는 C X_{34} = C 또는 부재	
예시적인 TTV 서열	전체 서열	GCCGCCGCGGCGGCGGSGGNGNSGCGCGCT DCGCGCGCSNNNCRCCRGGGGGNNNNCWG CSNCNCCCCCCCCCGCGCATGCGCGGGKCC CCCCCCCNNCGGGGGGCTCCGCCCCCCGGC CCCCCCCCGTGCTAAACCCACCGCGCATGC GCGACCACGCCCCCGCCGCC	121
	단편 1	GCCGCCGCGGCGGCGGSGGNGNSGCGCGCT DCGCGCGCSNNNCRCCRGGGGGNNNNCWG CSNCNCCCCCCCCCGCGCAT	122
	단편 2	GCGCGGGKCCCCCCCCCNNCGGGGGGCTC CG	123

[0780]

	단편 3	CCCCCCGGCCCCCCCCCGTGCTAAACCCAC CGCGCATGCGCGACCACGCCCCCGCCGCC	124
TTV-CT30F	전체 서열	GCGGCGG-GGGGGCG-GCCGCG-TTCGCGCGCCGCCCACCAGGGGGTG--CTGCG-CGCCCCCCCCCGCGCAT GCGCGGGGCCCCCCCCC--GGGGGGGCTCCGCCCCCCCGGCCCCCCCCC GTGCTAAACCCACCGCGCATGCGCGACCAC GCCCCCGCCGCC	125
	단편 1	GCGGCGG	126
	단편 2	GGGGGCG	127
	단편 3	GCCGCG	128
	단편 4	TTCGCGCGCCGCCCACCAGGGGGTG	129
	단편 5	CTGCG	130
	단편 6	CGCCCCCCCCCGCGCAT	131
	단편 7	GCGCGGGGCCCCCCCCC	132
	단편 8	GGGGGGGCTCCGCCCCCCCGGCCCCCCCCC GTGCTAAACCCACCGCGCATGCGCGACCAC GCCCCCGCCGCC	133
TTV-HD23a	전체 서열	CGGCGGCGGCGGCG-CGCGCGCTGCGCGCGCG---CGCCGGGGGGGCGCCAGCG-CCCCCCCCCCCCGCGCAT GCACGGGTCCCCCCCCCCCACGGGGGGCTCC G CCCCCCGGCCCCCCCCC	134
	단편 1	CGGCGGCGGCGGCG	135
	단편 2	CGCGCGCTGCGCGCGCG	136
	단편 3	CGCCGGGGGGGCGCCAGCG	137
	단편 4	CCCCCCCCCCCGCGCAT	138
	단편 5	GCACGGGTCCCCCCCCCCACGGGGGGCTCC G	139
	단편 6	CCCCCCGGCCCCCCCCC	140

[0781]

TTV-JA20	전체 서열	CCGTCGGCGGGGGGGCCGCGCGCTGCGCG CGCGGCCC- CCGGGGGAGGCACAGCCTCCCCCCCCCGCG CGCATGCGCGCGGGTCCCCCCCCCTCCGGG GGGCTCCGCCCCCCGGCCCCCCCC	141
	단편 1	CCGTCGGCGGGGGGGCCGCGCGCTGCGCG CGCGGCCC	142
	단편 2	CCGGGGGAGGCACAGCCTCCCCCCCCCGCG CGCATGCGCGCGGGTCCCCCCCCCTCCGGG GGGCTCCGCCCCCCGGCCCCCCCC	143
TTV-TJN02	전체 서열	CGGCGGCGGCG-CGCGCGCTACGCGCGCG-- -CGCCGGGGGG----CTGCCGC- CCCCCCCCCGCGCAT GCGCGGGGCCCCCCCCC- GCGGGGGGCTCCG CCCCCCGGCCCCCC	144
	단편 1	CGGCGGCGGCG	145
	단편 2	CGCGCGCTACGCGCGCG	146
	단편 3	CGCCGGGGGG	147
	단편 4	CTGCCGC	148
	단편 5	CCCCCCCCCGCGCAT	149
	단편 6	GCGCGGGGCCCCCCCCC	150
	단편 7	GCGGGGGGCTCCG	151
	단편 8	CCCCCCGGCCCCCC	152
TTV-tth8	전체 서열	GCCGCCGCGGCGGCGGGGG- GCGGCGCGCTGCGCGCGCCGCCCAGTAGG GGGAGCCATGCG---CCCCCCCCCGCGCAT GCGCGGGGCCCCCCCCC- GCGGGGGGCTCCG CCCCCCGGCCCCCCCCG	153
	단편 1	GCCGCCGCGGCGGCGGGGG	154
	단편 2	GCGGCGCGCTGCGCGCGCCGCCCAGTAGG GGGAGCCATGCG	155
	단편 3	CCCCCCCCCGCGCAT	156

[0782]

	단편 4	GCGCGGGGCCCCCCCCC	157
	단편 5	GCGGGGGGCTCCG	158
	단편 6	CCCCCCGGCCCCCCCCG	159
	단편 7	CGCGCTGCGCGCGCCGCCCAGTAGGGGGA GCCATGC	160
	단편 8	CCGCCATCTTAAGTAGTTGAGGCGGACGGT GGCGTGAGTTCAAAGGTCACCATCAGCCAC ACCTACTCAAAATGGTGG	161
	단편 9	CTTAAGTAGTTGAGGCGGACGGTGGCGTGA GTTCAAAGGTCACCATCAGCCACACCTACT CAAAATGGTGGACAATTTCTTCCGGGTCAA AGGTTACAGCCGCCATGTTAAAACACGTGA CGTATGACGTCACGGCCGCCATTTTGTGAC ACAAGATGGCCGACTTCCTTCC	162
추가의 GC-풍부 서열	36- 뉴클레오티드 컨센서스 GC-풍부 영역 서열 1	CGCGCTGCGCGCGCCGCCCAGTAGGGGGA GCCATGC	163
	36-뉴클레오티드 영역 컨센서스 서열 2	GCGCTX$_1$CGCGCGCGCCGGGGGGCTGCG CCCCCCC, 여기서, X$_1$은 T, G, 또는 A로부터 선택됨	164
	TTV 클레이드 1 36-뉴클레오티드 영역	GCGCTTCGCGCGCCGCCCACTAGGGGGCGT TGCGCG	165
	TTV 클레이드 3 36-뉴클레오티드 영역	GCGCTGCGCGCGCCGCCCAGTAGGGGGCG CAATGCG	166
	TTV 클레이드 3 단리 GH1 36-뉴클레오티드 영역	GCGCTGCGCGCGCGGCCCCCGGGGGAGGC ATTGCCT	167

[0783]

	TTV 클레이드 3 sle1932 36-뉴클레오티드 영역	GCGCTGCGCGCGCGCCGGGGGGGCGCC AGCGCCC	168
	TTV 클레이드 4 ctdc002 36-뉴클레오티드 영역	GCGCTTCGCGCGCGCCGGGGGGCTCCGC CCCCCC	169
	TTV 클레이드 5 36-뉴클레오티드 영역	GCGCTTCGCGCGCGCCGGGGGGCTGCGC CCCCCC	170
	TTV 클레이드 6 36-뉴클레오티드 영역	GCGCTACGCGCGCGCCGGGGGGCTGCG CCCCCCC	171
	TTV 클레이드 7 36-뉴클레오티드 영역	GCGCTACGCGCGCGCCGGGGGGCTCTGC CCCCCC	172
추가의 알파토르크바이러스 GC-풍부 영역 서열	TTV-CT30F	GCGGCGGGGGGGCGGCCGCGTTCGCGCGC CGCCCACCAGGGGGTGCTGCGCGCCCCCCC CCGCGCATGCGCGGGGCCCCCCCCCGGGG GGGCTCCGCCCCCCCGGCCCCCCCCCGTGC TAAACCCACCGCGCATGCGCGACCACGCCC CCGCCGCC	801
	TTV-P13-1	CCGAGCGTTAGCGAGGAGTGCGACCCTACC CCCTGGGCCCACTTCTTCGGAGCCGCGCGC TACGCCTTCGGCTGCGCGCGGCACCTCAGA CCCCCGCTCGTGCTGACACGCTTGCGCGTG TCAGACCACTTCGGGCTCGCGGGGGTCGGG	802
	TTV-tth8	GCCGCCGCGGCGGCGGGGGGCGGCGCGCT GCGCGCGCCGCCCAGTAGGGGGAGCCATG CGCCCCCCCCCGCGCATGCGCGGGGCCCCC CCCCGCGGGGGGCTCCGCCCCCCGGCCCCC CCCG	803

[0784]

	TTV-HD20a	CGGCCCAGCGGCGGCGCGCGCTTCGCGC GCGCGCCGGGGGGCTCCGCCCCCCCCCCGCG CATGCGCGGGGCCCCCCCCCCGCGGGGGGCT CCGCCCCCCGGTCCCCCCCCG	804
	TTV-16	CGGCCGTGCGGCGGCGCGCGCTTCGCGC GCGCGCCGGGGGCTGCCGCCCCCCCCCCGCG CATGCGCGCGGGGCCCCCCCCCCGCGGGGG GCTCCGCCCCCCGGCCCCCCCCCCG	805
	TTV-TJN02	CGGCGGCGGCGCGCGCGCTACGCGCGCGC GCCGGGGGGCTGCCGCCCCCCCCCCGCGCA TGCGCGGGGCCCCCCCCCCGCGGGGGGCTCC GCCCCCCGGCCCCCC	806
	TTV-HD16d	GGCGGCGGCGCGCGCGCTACGCGCGCGCG CCGGGGAGCTCTGCCCCCCCCCCGCGCATGC GCGCGGGTCCCCCCCCCCGCGGGGGGCTCCG CCCCCCGGTCCCCCCCCCG	807

[0785]

[0786] *이펙터*

[0787] 일부 구현예에서, 유전 요소는 이펙터, 예를 들어 기능적 이펙터, 예를 들어, 내인성 이펙터 또는 외인성 이펙터, 예를 들어, 치료적 폴리펩티드 또는 핵산, 예를 들어, 세포독성 또는 세포용해 RNA 또는 단백질을 인코딩하는 하나 이상의 서열을 포함할 수 있다. 일부 구현예에서, 기능적 핵산은 비-코딩 RNA이다. 일부 구현예에서, 기능적 핵산은 코딩 RNA이다. 이펙터는 생물학적 활성을 조절할 수 있으며, 예를 들어, 효소 활성, 유전자 발현, 세포 신호전달 및 세포 또는 기관 기능을 증가시키거나 감소시킬 수 있다. 이펙터 활성은 또한 조절 단백질에 결합하여, 조절제의 활성, 예컨대 전사 또는 번역을 조절하는 것을 포함할 수 있다. 이펙터 활성은 또한, 활성화제 또는 저해제 기능을 포함할 수 있다. 예를 들어, 이펙터는 효소에서 증가된 기질 친화성을 촉발시킴으로써 효소 활성을 유도할 수 있으며, 예를 들어, 프룩토스 2,6-비스포스페이트는 포스포프룩토키나제 1을 활성화시키고, 인슐린에 반응하여 당분해 속도를 증가시킨다. 또 다른 예에서, 이펙터는 수용체로의 기질 결합을 저해하고, 이의 활성화를 저해할 수 있으며, 예를 들어, 날트렉손(naltrexone) 및 날록손(naloxone)은 오피오이드 수용체를 활성화시키지 않고 결합하며, 수용체가 오피오이드에 결합하지 못하게 한다. 이펙터 활성은 또한, 단백질 안정성/분해 및/또는 전사물 안정성/분해를 조절하는 것을 포함할 수 있다. 예를 들어, 단백질은 분해를 위해 폴리펩티드 보조-인자인 유비퀴틴에 의해 단백질 상에 표적화되어, 분해에 대해 표시된다. 또 다른 예에서, 이펙터는 효소의 활성 부위를 차단함으로써 효소 활성을 저해하며, 예를 들어, 메토트렉세이트는 천연 기질보다 1000배 더 단단하게 다이하이드로폴레이트 환원효소에 결합하고, 뉴클레오티드 염기 합성을 저해하는, 효소 다이하이드로폴레이트 환원효소에 대한 보조효소인 테트라하이드로폴레이트의 구조적 유사체이다.

[0788] 일부 구현예에서, 이펙터를 인코딩하는 서열은 유전 요소의 부분이며, 예를 들어, 그것은 본 명세서에 기재된 바와 같이 삽입 부위에서 삽입될 수 있다. 구현예에서, 이펙터를 인코딩하는 서열은 비코딩 영역, 예를 들어, 오픈-리딩 프레임의 3'에, 그리고 유전 요소의 GC-풍부 영역의 5'에 배치된 비코딩 영역에서, TATA 박스의 상류의 5' 비코딩 영역 내에서, 5' UTR 내에서, 폴리-A 신호의 하류 또는 GC-풍부 영역의 상류의 3' 비코딩 영역에서 유전 요소 내로 삽입된다. 구현예에서, 이펙터를 인코딩하는 서열은 예를 들어, 본 명세서에 기재된 바와 같이 TTV-tth8 플라스미드의 뉴클레오티드 약 3588에서, 또는 예를 들어, 본 명세서에 기재된 바와 같이 TTMV-LY2 플라스미드의 뉴클레오티드 약 2843에서 유전 요소 내로 삽입된다. 구현예에서, 이펙터를 인코딩하는 서열은 예를 들어, 본 명세서에 기재된 바와 같이 TTV-tth8 플라스미드의 뉴클레오티드 336 내지 3015에 또는 그 내에, 또는 예를 들어, 본 명세서에 기재된 바와 같이 TTV-LY2 플라스미드의 뉴클레오티드 242 내지 2812에 또는 그 내에 유전 요소 내로 삽입된다. 일부 구현예에서, 이펙터를 인코딩하는 서열은 오픈-리딩 프레임(예를 들어, 본 명세서에 기재된 바와 같은 ORF, 예를 들어, ORF1, ORF1/1, ORF1/2, ORF2, ORF2/2, ORF2/3 및/또는 ORF2t/3)의 부분 또는 전부를 대체한다.

[0789] 일부 구현예에서, 이펙터를 인코딩하는 서열은 100 내지 2000개, 100 내지 1000개, 100 내지 500개, 100 내지 200개, 200 내지 2000개, 200 내지 1000개, 200 내지 500개, 500 내지 1000, 500 내지 2000 또는 1000 내지

2000개의 뉴클레오티드를 포함한다. 일부 구현예에서, 이펙터는 예를 들어, 본 명세서에 기재된 바와 같이 핵산 또는 단백질 페이로드이다.

[0790] *조절 핵산*

[0791] 일부 구현예에서, 이펙터는 조절 핵산이다. 조절 핵산은 내인성 유전자 및/또는 외인성 유전자의 발현을 변경시킨다. 일 구현예에서, 조절 핵산은 숙주 유전자를 표적화한다. 조절 핵산은 내인성 유전자(예를 들어, 본 명세서의 다른 곳에 기재된 바와 같은 miRNA, siRNA, mRNA, lncRNA, RNA, DNA, 안티센스 RNA, gRNA)에 혼성화되는 핵산, 외인성 핵산, 예컨대 바이러스 DNA 또는 RNA에 혼성화하는 핵산, RNA에 혼성화하는 핵산, 유전자 전사를 간섭하는 핵산, RNA 번역을 간섭하는 핵산, RNA를 안정화시키거나, 예컨대 분해에 대한 표적화를 통하여 RNA를 불안정화시키는 핵산, 및 DNA 또는 RNA 결합 인자를 조절하는 핵산을 포함할 수 있지만, 이에 제한되지 않는다. 구현예에서, 조절 핵산은 miRNA를 인코딩한다. 일부 구현예에서, 조절 핵산은 야생형 아넬로바이러스에 대해 내인성이다. 일부 구현예에서, 조절 핵산은 야생형 아넬로바이러스에 대해 외인성이다.

[0792] 일부 구현예에서, 조절 핵산은 통상적으로 5 내지 500개의 염기쌍(특정 RNA 구조에 따라, 예를 들어, miRNA 5 내지 30 bp, lncRNA 200 내지 500 bp)을 함유하는 RNA 또는 RNA-유사 구조를 포함하며, 세포 내의 발현된 표적 유전자의 코딩 서열 또는 세포 내의 발현된 표적 유전자를 인코딩하는 서열과 동일하거나(또는 상보성이거나) 거의 동일한(또는 실질적으로 상보성인) 핵염기 서열을 가질 수 있다.

[0793] 일부 구현예에서, 조절 핵산은 핵산 서열, 예를 들어, 가이드 RNA(gRNA)를 포함한다. 일부 구현예에서, DNA 표적화 모이어티는 가이드 RNA 또는 가이드 RNA를 인코딩하는 핵산을 포함한다. gRNA 짧은 합성 RNA는 불완전한 이펙터 모이어티로의 결합에 필요한 "스캐폴드" 서열 및 게놈 표적에 대한 사용자-정의된 약 20개 뉴클레오티드 표적화 서열로 구성될 수 있다. 실시에서, 가이드 RNA 서열은 일반적으로 17 내지 24개의 뉴클레오티드(예를 들어, 19, 20 또는 21개의 뉴클레오티드)의 길이를 갖도록 설계되고 표적화된 핵산 서열에 대하여 상보성이다. 커스텀 gRNA 생성자 및 알고리즘은 효율적인 가이드 RNA의 설계에 사용하기 위하여 상업적으로 이용 가능하다. 유전자 편집은 또한, 천연 발생 crRNA-tracrRNA 복합체를 모방하고 tracrRNA(뉴클레아제에 결합) 및 적어도 하나의 crRNA(뉴클레아제를 편집에 대해 표적화된 서열로 가이드함) 둘 모두를 함유하는 엔지니어링된(합성) 단일 RNA 분자인, 키메라 "단일 가이드 RNA"("sgRNA")를 사용하여 달성되었다. 화학적으로 변형된 sgRNA는 또한, 게놈 편집에서 효율적인 것으로 입증되었으며; 예를 들어, 문헌[Hendel et al. (2015) Nature Biotechnol., 985 - 991]을 참조한다.

[0794] 조절 핵산은 특정 DNA 서열(예를 들어, 프로모터, 유전자의 인핸서, 침묵자 또는 억제자에 인접하거나 그 내의 서열)을 인식하는 gRNA를 포함한다.

[0795] 특정 조절 핵산은 RNA 간섭(RNAi)의 생물학적 과정을 통하여 유전자 발현을 저해할 수 있다. RNAi 분자는 통상적으로 15 내지 50개의 염기쌍(예컨대 약 18 내지 25개의 염기쌍)을 함유하고, 세포 내의 발현된 표적 유전자의 코딩 서열과 동일하거나(상보성이거나) 또는 거의 동일한(실질적으로 상보성인) 핵염기 서열을 갖는 RNA 또는 RNA-유사 구조를 포함한다. RNAi 분자는 짧은 간섭 RNA(siRNA), 이중-가닥 RNA(dsRNA), 마이크로 RNA(miRNA), 짧은 헤어핀 RNA(shRNA), 메로듀플렉스(meroduplex) 및 다이서(dicer) 기질(미국 특허 8,084,599, 8,349,809 및 8,513,207)을 포함하나 이들에 제한되지 않는다.

[0796] 긴 비-코딩 RNA(lncRNA)는 100개 초과의 뉴클레오티드의 비-단백질 코딩 전사물로서 정의된다. 이러한 다소 임의의 제한은 lncRNA를 소형 조절 RNA, 예컨대 마이크로RNA(miRNA), 짧은 간섭 RNA(siRNA) 및 다른 짧은 RNA로부터 구별한다. 일반적으로, 대다수(약 78%)의 lncRNA는 조직-특이적인 것으로 특성화된다. 근처의 단백질-코딩 유전자에 대하여 반대 방향으로 전사되는(포유동물 게놈 내에 총 lncRNA의 유의미한 비율 약 20%를 포함하는) 분지 lncRNA는 아마도 근처의 유전자의 전사를 조절할 수 있다.

[0797] 유전 요소는 내인성 유전자 또는 유전자 생성물(예를 들어, mRNA)의 전부 또는 이의 단편에 실질적으로 상보성이거나, 완전히 상보성인 서열을 갖는 조절 핵산을 인코딩할 수 있다. 조절 핵산은 인트론과 엑손 사이의 경계에서의 서열에 상보성이어서, 특정 유전자의 새로 생성된 핵 RNA 전사물이 전사를 위해 mRNA로 성숙되는 것을 방지할 수 있다. 특정 유전자에 상보성인 조절 핵산은 유전자에 대한 mRNA와 혼성화하고, 이의 번역을 방지할 수 있다. 안티센스 조절 핵산은 DNA, RNA 또는 이의 유도체 또는 혼성물일 수 있다.

[0798] 관심 전사물에 혼성화하는 조절 핵산의 길이는 5 내지 30개의 뉴클레오티드, 약 10 내지 30개의 뉴클레오티드 또는 약 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30개 이상의 뉴클레오티드일 수 있다. 표적화된 전사물에 대한 조절 핵산의 동일성의 정도는 적어도 75%, 적어도 80%, 적어도 85%,

적어도 90% 또는 적어도 95%여야 한다.

[0799] 유전 요소는 조절 핵산, 예를 들어, 표적 유전자의 약 5 내지 약 25개의 인접 뉴클레오티드와 동일한 마이크로RNA(miRNA) 분자를 인코딩할 수 있다. 일부 구현예에서, miRNA 서열은 mRNA를 표적화하며, 다이뉴클레오티드 AA로 시작하며, 약 30 내지 70%(약 30 내지 60%, 약 40 내지 60% 또는 약 45% 내지 55%)의 GC-함량을 포함하며, 예를 들어, 표준 BLAST 검색에 의해 결정시, 그것이 도입될 포유동물의 게놈 내의 표적 이외의 임의의 뉴클레오티드 서열에 대하여 높은 동일성 백분율을 갖지 않는다.

[0800] siRNA 및 shRNA는 내인성 마이크로RNA(miRNA) 유전자의 가공 경로 내의 중간체와 유사하다(문헌[Bartel, Cell 116:281-297, 2004]). 일부 구현예에서, siRNA는 miRNA로서 기능할 수 있으며, 그 역도 그러하다(문헌[Zeng et al., Mol Cell 9:1327-1333, 2002]; 문헌[Doench et al., Genes Dev 17:438-442, 2003]). siRNA와 같이 마이크로RNA는 RISC를 사용하여, 표적 유전자를 하향조절하지만, siRNA와 달리, 대부분의 동물 miRNA는 mRNA를 절단하지 않는다. 대신에, miRNA는 번역 억제 또는 폴리A 제거 및 mRNA 분해를 통해 단백질 출력물을 감소시킨다(문헌[Wu et al., Proc Natl Acad Sci USA 103:4034-4039, 2006]). 알려진 miRNA 결합 부위는 mRNA 3' UTR 내에 존재하며; miRNA는 miRNA의 5' 말단으로부터 뉴클레오티드 2 내지 8에 대하여 거의 완벽하게 상보성을 갖는 부위를 표적화하는 것으로 보인다(문헌[Rajewsky, Nat Genet 38 Suppl:S8-13, 2006]; 문헌[Lim et al., Nature 433:769-773, 2005]). 이러한 영역은 시드 영역으로 공지되어 있다. siRNA 및 miRNA가 상호 교환 가능하기 때문에, 외인성 siRNA는 siRNA에 대하여 시드 상보성을 갖는 mRNA를 하향조절한다(문헌[Birmingham et al., Nat Methods 3:199-204, 2006]). 3' UTR 내의 다수의 표적 부위는 더 강한 하향조절을 제공한다(문헌[Doench et al., Genes Dev 17:438-442, 2003]).

[0801] 알려진 miRNA 서열의 목록은 특히 Wellcome Trust Sanger Institute, Penn Center for Bioinformatics, Memorial Sloan Kettering Cancer Center, 및 European Molecule Biology Laboratory와 같은 연구 기관에 의해 유지되는 데이터베이스에서 확인될 수 있다. 또한, 알려진 효과적인 siRNA 서열 및 동족(동족) 결합 부위는 관련 문헌에서 잘 나타나 있다. RNAi 분자는 당업계에 알려진 기술에 의해 쉽게 설계되고 생성된다. 또한, 효과적이고 특이적인 서열 모티프를 찾는 기회를 증가시키는 컴퓨터 기반 도구가 존재한다(문헌[Lagana et al., Methods Mol. Bio., 2015, 1269:393-412]).

[0802] 조절 핵산은 유전자에 의해 인코딩되는 RNA의 발현을 조절할 수 있다. 다수의 유전자가 서로 어느 정도의 서열 상동성을 공유할 수 있기 때문에, 일부 구현예에서, 조절 핵산은 충분한 서열 상동성을 갖는 유전자의 부류를 표적화하도록 설계될 수 있다. 일부 구현예에서, 조절 핵산은 상이한 유전자 표적 간에 공유되는 또는 특정 유전자 표적에 특유한 서열에 상보성을 갖는 서열을 함유할 수 있다. 일부 구현예에서, 조절 핵산은 몇몇의 유전자 간에 상동성을 갖는 RNA 서열의 보존된 영역을 표적화하여, 그에 의해, 유전자 과 내의 몇몇의 유전자(예를 들어, 상이한 유전자 아이소폼, 슬라이스 변이체, 돌연변이 유전자 등)를 표적화하도록 설계될 수 있다. 일부 구현예에서, 조절 핵산은 단일 유전자의 특정 RNA 서열에 특유한 서열을 표적화하도록 설계될 수 있다.

[0803] 일부 구현예에서, 유전 요소는 하나 이상의 유전자의 발현을 조절하는 조절 핵산을 인코딩하는 하나 이상의 서열을 포함할 수 있다.

[0804] 일 구현예에서, 본 명세서의 다른 곳에 기재된 gRNA는 유전자 편집을 위한 CRISPR 시스템의 부분으로서 사용된다. 유전자 편집의 목적을 위하여, 아넬로벡터는 요망되는 표적 DNA 서열에 상응하는 하나의 또는 다수의 가이드 RNA 서열을 포함하도록 설계될 수 있으며; 예를 들어, 문헌[Cong et al. (2013) Science, 339:819-823]; 문헌[Ran et al. (2013) Nature Protocols, 8:2281 - 2308]을 참조한다. 적어도 약 16 또는 17개 뉴클레오티드의 gRNA 서열은 일반적으로 Cas9-매개된 DNA 절단이 발생하게 하며; Cpf1에 있어서, 적어도 약 16개 뉴클레오티드의 gRNA 서열이 검출 가능한 DNA 절단을 달성하는 데 필요하다.

[0805] *치료용 이펙터(예를 들어, 펩티드 또는 폴리펩티드)*

[0806] 일부 구현예에서, 유전 요소는 치료용 발현 서열, 예를 들어 치료용 펩티드 또는 폴리펩티드, 예를 들어 세포내 펩티드 또는 세포내 폴리펩티드, 분비 폴리펩티드, 또는 단백질 대체 치료제를 인코딩하는 서열을 포함한다. 일부 구현예에서, 유전 요소는 단백질, 예를 들어, 치료용 단백질을 인코딩하는 서열을 포함한다. 치료용 단백질의 일부 예는 호르몬, 사이토카인, 효소, 항체(예를 들어, 적어도 중쇄 또는 경쇄를 인코딩하는 하나 또는 복수의 폴리펩티드), 전사 인자, 수용체(예를 들어, 막 수용체), 리간드, 막 수송체, 분비 단백질, 펩티드, 담체 단백질, 구조 단백질, 뉴클레아제, 또는 이의 구성성분을 포함할 수 있지만 이에 제한되지 않는다.

[0807] 일부 구현예에서, 유전 요소는 펩티드, 예를 들어, 치료용 펩티드를 인코딩하는 서열을 포함한다. 펩티드는 선

형 또는 분지형일 수 있다. 펩티드는 약 5 내지 약 500개의 아미노산, 약 15 내지 약 400개의 아미노산, 약 20 내지 약 325개의 아미노산, 약 25 내지 약 250개의 아미노산, 약 50 내지 약 200개의 아미노산, 또는 이들 사이의 임의의 범위의 길이를 갖는다.

[0808] 일부 구현예에서, 치료용 발현 서열에 의해 인코딩되는 폴리펩티드는 상기 중 임의의 것의 기능적 변이체 또는 이의 단편, 예를 들어, UniProt ID를 참조하여 본 명세서의 표에 개시된 단백질 서열에 대하여 적어도 80%, 85%, 90%, 95%, 967%, 98%, 99% 동일성을 갖는 단백질일 수 있다.

[0809] 일부 구현예에서, 치료용 발현 서열은 상기 중 임의의 것에 결합하는 항체 또는 항체 단편, 예를 들어, UniProt ID를 참조하여 본 명세서의 표에 개시된 단백질 서열에 대하여 적어도 80%, 85%, 90%, 95%, 967%, 98%, 99% 동일성을 갖는 단백질에 대한 항체를 인코딩할 수 있다. 본 명세서에서 용어 "항체"는 가장 넓은 의미로 사용되며, 원하는 항원-결합 활성을 나타내는 한 모노클로날 항체, 폴리클로날 항체, 다중특이적 항체(예를 들어, 이중 특이적 항체), 및 항체 단편을 포함하지만 이에 제한되지 않는 다양한 항체 구조를 포함한다. "항체 단편"은 적어도 하나의 중쇄 또는 경쇄를 포함하고 항원에 결합하는 분자를 지칭한다. 항체 단편의 예는 Fv, Fab, Fab', Fab'-SH, $F(ab')_2$; 디아바디; 선형 항체; 단쇄 항체 분자(예를 들어, scFv); 및 항체 단편으로부터 형성된 다중특이적 항체를 포함하지만 이에 제한되지 않는다.

[0810] **예시적인 세포내 폴리펩티드 이펙터**

[0811] 일부 구현예에서, 이펙터는 세포질 폴리펩티드 또는 세포질 펩티드를 포함한다. 일부 구현예에서, 이펙터는 DPP-4 억제제, GLP-1 신호전달의 활성화제, 또는 호중구 엘라스타제의 억제제인 세포질 펩티드를 포함한다. 일부 구현예에서, 이펙터는 성장 인자 또는 이의 수용체(예를 들어, FGF 수용체, 예를 들어, FGFR3)의 수준 또는 활성을 증가시킨다. 일부 구현예에서, 이펙터는 n-myc 상호작용 단백질 활성의 억제제(예를 들어, n-myc 상호작용 단백질 억제제); EGFR 활성의 억제제(예를 들어, EGFR 억제제); IDH1 및/또는 IDH2 활성의 억제제(예를 들어, IDH1 억제제 및/또는 IDH2 억제제); LRP5 및/또는 DKK2 활성의 억제제(예를 들어, LRP5 및/또는 DKK2 억제제); KRAS 활성의 억제제; HTT 활동의 활성화제; 또는 DPP-4 활성의 억제제(예를 들어, DPP-4 억제제)를 포함한다.

[0812] 일부 구현예에서, 이펙터는 조절 세포내 폴리펩티드를 포함한다. 일부 구현예에서, 조절 세포내 폴리펩티드는 표적 세포에 대해 내인성인 하나 이상의 분자(예를 들어, 단백질 또는 핵산)에 결합한다. 일부 구현예에서, 조절 세포내 폴리펩티드는 표적 세포에 대해 내인성인 하나 이상의 분자(예를 들어, 단백질 또는 핵산)의 수준 또는 활성을 증가시킨다. 일부 구현예에서, 조절 세포내 폴리펩티드는 표적 세포에 대해 내인성인 하나 이상의 분자(예를 들어, 단백질 또는 핵산)의 수준 또는 활성을 감소시킨다.

[0813] **예시적인 분비 폴리펩티드 이펙터**

[0814] 예시적인 분비 치료제는 본 명세서에, 예를 들어 하기 표에 기재되어 있다.

[0815] [표 50] 예시적인 사이토카인 및 사이토카인 수용체

사이토카인	사이토카인 수용체(들)	Entrez Gene ID	UniProt ID
IL-1α, IL-1β, 또는 이의 이종이량체	IL-1 타입 1 수용체 , IL-1 타입 2 수용체	3552, 3553	P01583, P01584
IL-1Ra	IL-1 타입 1 수용체 , IL-1 타입 2 수용체	3454, 3455	P17181, P48551
IL-2	IL-2R	3558	P60568
IL-3	IL-3 수용체 α + β c (CD131)	3562	P08700
IL-4	IL-4R 타입 I, IL-4R 타입 II	3565	P05112
IL-5	IL-5R	3567	P05113
IL-6	IL-6R (sIL-6R) gp130	3569	P05231
IL-7	IL-7R 및 sIL-7R	3574	P13232
IL-8	CXCR1 및 CXCR2	3576	P10145
IL-9	IL-9R	3578	P15248
IL-10	IL-10R1/IL-10R2 복합체	3586	P22301
IL-11	IL-11Rα 1 gp130	3589	P20809
IL-12 (예: p35, p40, 또는 이의 이종이량체)	IL-12Rβ1 및 IL-12Rβ2	3593, 3592	P29459, P29460
IL-13	IL-13R1α1 및 IL-13R1α2	3596	P35225
IL-14	IL-14R	30685	P40222
IL-15	IL-15R	3600	P40933
IL-16	CD4	3603	Q14005
IL-17A	IL-17RA	3605	Q16552
IL-17B	IL-17RB	27190	Q9UHF5
IL-17C	IL-17RA 내지 IL-17RE	27189	Q9P0M4
e	SEF	53342	Q8TAD2
IL-17F	IL-17RA, IL-17RC	112744	Q96PD4
IL-18	IL-18 수용체	3606	Q14116
IL-19	IL-20R1/IL-20R2	29949	Q9UHD0

[0816]

IL-20	L-20R1/IL-20R2 및 IL-22R1/ IL-20R2	50604	Q9NYY1
IL-21	IL-21R	59067	Q9HBE4
IL-22	IL-22R	50616	Q9GZX6
IL-23 (예: p19, p40, 또는 이의 이종이량체)	IL-23R	51561	Q9NPF7
IL-24	IL-20R1/IL-20R2 및 IL-22R1/ IL-20R2	11009	Q13007
IL-25	IL-17RA 및 IL-17RB	64806	Q9H293
IL-26	IL-10R2 사슬 및 IL-20R1 사슬	55801	Q9NPH9
IL-27 (예: p28, EBI3, 또는 이의 이종이량체)	WSX-1 및 gp130	246778	Q8NEV9
IL-28A, IL-28B, 및 IL29	IL-28R1/IL-10R2	282617, 282618	Q8IZI9, Q8IU54
IL-30	IL6R/gp130	246778	Q8NEV9
IL-31	IL-31RA/OSMRβ	386653	Q6EBC2
IL-32		9235	P24001
IL-33	ST2	90865	O95760
IL-34	콜로니-자극 인자 1 수용체	146433	Q6ZMJ4
IL-35 (예: p35, EBI3, 또는 이의 이종이량체)	IL-12Rβ2/gp130; IL-12Rβ2/IL-12Rβ2; gp130/gp130	10148	Q14213
IL-36	IL-36Ra	27179	Q9UHA7
IL-37	IL-18Rα 및 IL-18BP	27178	Q9NZH6
IL-38	IL-1R1, IL-36R	84639	Q8WWZ1
IFN-α	IFNAR	3454	P17181
IFN-β	IFNAR	3454	P17181
IFN-γ	IFNGR1/IFNGR2	3459	P15260
TGF-β	TβR-I 및 TβR-II	7046, 7048	P36897, P37173
TNF-α	TNFR1, TNFR2	7132, 7133	P19438, P20333

[0817]

[0818] 일부 구현예에서, 본 명세서에 기재된 이펙터는 표 50의 사이토카인, 또는 이의 기능적 변이체, 예를 들어 이의 상동체(예를 들어, 오솔로그 또는 파라로그) 또는 단편을 포함한다. 일부 구현예에서, 본 명세서에 기재된 이펙터는 UniProt ID를 참조하여 표 50에 열거된 아미노산 서열에 대하여 적어도 80%, 85%, 90%, 95%, 967%, 98%, 99% 서열 동일성을 갖는 단백질을 포함한다. 일부 구현예에서, 기능적 변이체는 동일한 조건 하에서 동일한 수용체에 대하여 상응하는 야생형 사이토카인의 Kd보다 10%, 20%, 30%, 40%, 또는 50% 이하로 더 높거나 더 낮은 Kd로 상응하는 사이토카인 수용체에 결합한다. 일부 구현예에서, 이펙터는 제1 영역(예를 들어, 표 50의 사이토카인 폴리펩티드 또는 이의 기능적 변이체 또는 단편) 및 제2의 이종 영역을 포함하는 융합 단백질을 포함한다. 일부 구현예에서, 제1 영역은 표 50의 제1 사이토카인 폴리펩티드이다. 일부 구현예에서, 제2 영역은 표 50의 제2 사이토카인 폴리펩티드이며, 여기서 제1 및 제2 사이토카인 폴리펩티드는 야생형 세포에서 서로 사이토카인 이종이량체를 형성한다. 일부 구현예에서, 표 50의 폴리펩티드 또는 이의 기능적 변이체는 신호 서열, 예를 들어, 이펙터에 대해 내인성인 신호 서열, 또는 이종 신호 서열을 포함한다. 일부 구현예에서, 표 50의 사이토카인, 또는 이의 기능적 변이체를 인코딩하는 아넬로벡터는 본 명세서에 기재된 질환 또는 장애의 치료에 사용된다.

[0819] 일부 구현예에서, 본 명세서에 기재된 이펙터는 표 50의 사이토카인에 결합하는 항체 분자(예를 들어, scFv)를 포함한다. 일부 구현예에서, 본 명세서에 기재된 이펙터는 표 50의 사이토카인 수용체에 결합하는 항체 분자(예를 들어, scFv)를 포함한다. 일부 구현예에서, 항체 분자는 신호 서열을 포함한다.

[0820] 예시적인 사이토카인 및 사이토카인 수용체는, 예를 들어, 문헌[Akdis et al., "Interleukins (from IL-1 to IL-38), interferons, transforming growth factor β, and TNF-α: Receptors, functions, and roles in diseases" October 2016 Volume 138, Issue 4, Pages 984-1010]에 기재되어 있으며, 이는 본 명세서에 표 I을 포함하여 전문이 참조로 포함된다.

[0821] **[표 51] 예시적인 폴리펩티드 호르몬 및 수용체**

호르몬	수용체	Entrez Gene ID	UniProt ID
나트륨이뇨 펩티드, 예를 들어, 심방 나트륨이뇨 펩티드(ANP)	NPRA, NPRB, NPRC	4878	P01160
뇌 나트륨이뇨 펩티드(BNP)	NPRA, NPRB	4879	P16860
C 형 나트륨이뇨 펩티드(CNP)	NPRB	4880	P23582
성장 호르몬(GH)	GHR	2690	P10912
인간 성장 호르몬(hGH)	hGH 수용체(인간 GHR)	2690	P10912
프롤락틴(PRL)	PRLR	5617	P01236
갑상선-자극 호르몬(TSH)	TSH 수용체	7253	P16473
부신피질자극호르몬(ACTH)	ACTH 수용체	5443	P01189
난포-자극 호르몬(FSH)	FSHR	2492	P23945
황체 형성 호르몬(LH)	LHR	3973	P22888
항이뇨 호르몬(ADH)	바소프레신 수용체, 예를 들어, V2; AVPR1A; AVPR1B; AVPR3; AVPR2	554	P30518
옥시토신	OXTR	5020	P01178
칼시토닌	칼시노닌 수용체(CT)	796	P01258
부갑상선 호르몬(PTH)	PTH1R 및 PTH2R	5741	P01270
인슐린	인슐린 수용체(IR)	3630	P01308
글루카곤	글루카곤 수용체	2641	P01275

[0822]

[0823] 일부 구현예에서, 본 명세서에 기재된 이펙터는 표 51의 호르몬, 또는 이의 기능적 변이체, 예를 들어, 이의 상동체(예를 들어, 오솔로그 또는 파라로그) 또는 단편을 포함한다. 일부 구현예에서, 본 명세서에 기재된 이펙터는 UniProt ID를 참조하여 표 51에 열거된 아미노산 서열에 대하여 적어도 80%, 85%, 90%, 95%, 967%, 98%,

99% 서열 동일성을 갖는 단백질을 포함한다. 일부 구현예에서, 기능적 변이체는 동일한 조건 하에서 동일한 수용체에 대하여 상응하는 야생형 호르몬의 Kd보다 10%, 20%, 30%, 40%, 또는 50% 이하로 더 높은 Kd로 상응하는 수용체에 결합한다. 일부 구현예에서, 표 51의 폴리펩티드 또는 이의 기능적 변이체는 신호 서열, 예를 들어, 이펙터에 대해 내인성인 신호 서열, 또는 이종 신호 서열을 포함한다. 일부 구현예에서, 표 51의 호르몬, 또는 이의 기능적 변이체를 인코딩하는 아넬로벡터는 본 명세서에 기재된 질환 또는 장애의 치료에 사용된다.

[0824] 일부 구현예에서, 본 명세서에 기재된 이펙터는 표 51의 호르몬에 결합하는 항체 분자(예를 들어, scFv)를 포함한다. 일부 구현예에서, 본 명세서에 기재된 이펙터는 표 51의 호르몬 수용체에 결합하는 항체 분자(예를 들어, scFv)를 포함한다. 일부 구현예에서, 항체 분자는 신호 서열을 포함한다.

[0825] **[표 52] 예시적인 성장 인자**

성장 인자		**Entrez Gene ID**	**UniProt ID**
PDGF 패밀리			
PDGF(예를 들어, PDGF-1, PDGF-2, 또는 이의 이종이량체)	PDGF 수용체, 예를 들어, PDGFRα, PDGFRβ	5156	P16234
CSF-1	CSF1R	1435	P09603
SCF	CD117	3815	P10721
VEGF 패밀리			
VEGF(예를 들어, 아이소폼 VEGF 121, VEGF 165, VEGF 189, 및 VEGF 206)	VEGFR-1, VEGFR-2	2321	P17948
VEGF-B	VEGFR-1	2321	P17949
VEGF-C	VEGFR-2 및 VEGFR -3	2324	P35916
PlGF	VEGFR-1	5281	Q07326
EGF 패밀리			
EGF	EGFR	1950	P01133
TGF-α	EGFR	7039	P01135
앰피레귤린	EGFR	374	P15514
HB-EGF	EGFR	1839	Q99075
베타셀룰린	EGFR, ErbB-4	685	P35070
에피레귤린	EGFR, ErbB-4	2069	O14944
헤레귤린	EGFR, ErbB-4	3084	Q02297
FGF 패밀리			

[0826]

FGF-1, FGF-2, FGF-3, FGF-4, FGF-5, FGF-6, FGF-7, FGF-8, FGF-9	FGFR1, FGFR2, FGFR3, 및 FGFR4	2246, 2247, 2248, 2249, 2250, 2251, 2252, 2253, 2254	P05230, P09038, P11487, P08620, P12034, P10767, P21781, P55075, P31371
인슐린 패밀리			
인슐린	IR	3630	P01308
IGF-I	IGF-I 수용체, IGF-II 수용체	3479	P05019
IGF-II	IGF-II 수용체	3481	P01344
HGF 패밀리			
HGF	MET 수용체	3082	P14210
MSP	RON	4485	P26927
뉴로트로핀 패밀리			
NGF	LNGFR, trkA	4803	P01138
BDNF	trkB	627	P23560
NT-3	trkA, trkB, trkC	4908	P20783
NT-4	trkA, trkB	4909	P34130
NT-5	trkA, trkB	4909	P34130
안지오포이에틴 패밀리			
ANGPT1	HPK-6/TEK	284	Q15389
ANGPT2	HPK-6/TEK	285	O15123
ANGPT3	HPK-6/TEK	9068	O95841
ANGPT4	HPK-6/TEK	51378	Q9Y264

[0827]

[0828] 일부 구현예에서, 본 명세서에 기재된 이펙터는 표 52의 성장 인자, 또는 이의 기능적 변이체, 예를 들어, 이의 상동체(예를 들어, 오솔로그 또는 파라로그) 또는 단편을 포함한다. 일부 구현예에서, 본 명세서에 기재된 이펙터는 UniProt ID를 참조하여 표 52에 열거된 아미노산 서열에 대하여 적어도 80%, 85%, 90%, 95%, 967%, 98%, 99% 서열 동일성을 갖는 단백질을 포함한다. 일부 구현예에서, 기능적 변이체는 동일한 조건 하에서 동일한 수용체에 대하여 상응하는 야생형 성장 인자의 Kd보다 10%, 20%, 30%, 40%, 또는 50% 이하로 더 높은 Kd로 상응하는 수용체에 결합한다. 일부 구현예에서, 표 52의 폴리펩티드 또는 이의 기능적 변이체는 신호 서열, 예를 들어, 이펙터에 대해 내인성인 신호 서열, 또는 이종 신호 서열을 포함한다. 일부 구현예에서, 표 52의 성장 인자, 또는 이의 기능적 변이체를 인코딩하는 아넬로벡터는 본 명세서에 기재된 질환 또는 장애의 치료에 사용된다.

[0829] 일부 구현예에서, 본 명세서에 기재된 이펙터는 표 52의 성장 인자에 결합하는 항체 분자(예를 들어, scFv)를 포함한다. 일부 구현예에서, 본 명세서에 기재된 이펙터는 표 52의 성장 인자 수용체에 결합하는 항체 분자(예를 들어, scFv)를 포함한다. 일부 구현예에서, 항체 분자는 신호 서열을 포함한다.

[0830] 예시적인 성장 인자 및 성장 인자 수용체는, 예를 들어, 문헌[Bafico et al., "Classification of Growth Factors and Their Receptors" Holland-Frei Cancer Medicine. 6th edition]에 기재되어 있으며, 이는 본 명세서에 전문이 참조로 포함된다.

[0831] [표 53] 응고-연관 인자

이펙터	적응증	Entrez Gene ID	UniProt ID
인자 I (피브리노겐)	무피브리노겐혈증	2243, 2266, 2244	P02671, P02679, P02675
인자 II	인자 II 결핍증	2147	P00734
인자 IX	B 형 혈우병	2158	P00740
인자 V	오우렌병	2153	P12259
인자 VIII	A 형 혈우병	2157	P00451
인자 X	스튜어트-프라워 인자 결핍증	2159	P00742
인자 XI	C 형 혈우병	2160	P03951
인자 XIII	피브린 안정화 인자 결핍증	2162, 2165	P00488, P05160
vWF	폰 빌레브란트병	7450	P04275

[0832]

[0833] 일부 구현예에서, 본 명세서에 기재된 이펙터는 표 53의 폴리펩티드, 또는 이의 기능적 변이체, 예를 들어, 이의 상동체(예를 들어, 오솔로그 또는 파라로그) 또는 단편을 포함한다. 일부 구현예에서, 본 명세서에 기재된 이펙터는 UniProt ID를 참조하여 표 53에 열거된 아미노산 서열에 대하여 적어도 80%, 85%, 90%, 95%, 967%, 98%, 99% 서열 동일성을 갖는 단백질을 포함한다. 일부 구현예에서, 기능적 변이체는, 예를 들어, 야생형 단백질보다 10%, 20%, 30%, 40%, 또는 50% 이상 더 낮은 속도로, 상응하는 야생형 단백질과 동일한 반응을 촉매한다. 일부 구현예에서, 표 53의 폴리펩티드 또는 이의 기능적 변이체는 신호 서열, 예를 들어, 이펙터에 대해 내인성인 신호 서열, 또는 이종 신호 서열을 포함한다. 일부 구현예에서, 표 53의 폴리펩티드, 또는 이의 기능적 변이체를 인코딩하는 아넬로벡터는 표 53의 질환 또는 장애의 치료에 사용된다.

[0834] **예시적인 단백질 대체 치료제**

[0835] 예시적인 단백질 대체 치료제는 본 명세서에, 예를 들어 하기 표에 기재되어 있다.

[0836] [표 54] 예시적인 효소 이펙터 및 상응하는 적응증

이펙터	결핍증	Entrez Gene ID	UniProt ID
3-메틸크로토노일-CoA 카복실라제	3- 메틸크로토노일-CoA 카복실라제 결핍증	56922, 64087	Q96RQ3, Q9HCC0
아세틸-CoA-글루코사미니드 N-아세틸트랜스퍼라제	점액다당류증 MPS III(산필리포 증후군) 유형 III-C	138050	Q68CP4
ADAMTS13	혈전성 혈소판 감소성 자반증	11093	Q76LX8
아데닌 포스포리보실트랜스퍼라제	아데닌 포스포리보실트랜스퍼라제 결핍증	353	P07741
아데노신 데아미나제	아데노신 데아미나제 결핍증	100	P00813
ADP-리보스 단백질 하이드롤라제	글루타밀 리보스-5-포스페이트 축적병	26119, 54936	Q5SW96, Q9NX46
알파 글루코시다제	글리코겐 축적병 유형 2 (폼페병)	2548	P10253
아르기나제	가족성 아르기닌혈증	383, 384	P05089, P78540
아릴설파타제 A	이염백색질장애	410	P15289
카텝신 K	피크노디소토시스	1513	P43235
세라미다제	파아버병(지방육아종증)	125981, 340485, 55331	Q8TDN7, Q5QJU3, Q9NUN7
시스타티오닌 B 신타제	호모시스틴뇨증	875	P35520

[0837]

[0838]

돌리콜-P-만노스 신타제	N-글리코실화 CDG Ie 의 선천성 장애	8813, 54344	O60762, Q9P2X0
돌리코-P-Glc:Man9GlcNAc2-PP-돌리콜 글루코실트랜스퍼라제	N-글리코실화 CDG Ic 의 선천성 장애	84920	Q5BKT4
돌리코-P-Man:Man5GlcNAc2-PP-돌리콜 만노실트랜스퍼라제	N-글리코실화 CDG Id 의 선천성 장애	10195	Q92685
돌리킬-P-글루코스:Glc-1-Man-9-GlcNAc-2-PP-돌리킬-α-3-글루코실트랜스퍼라제	N-글리코실화 CDG Ih 의 선천성 장애	79053	Q9BVK2
돌리킬-P-만노스:Man-7-GlcNAc-2-PP-돌리킬-α-6-만노실트랜스퍼라제	N-글리코실화 CDG Ig 의 선천성 장애	79087	Q9BV10
인자 II	인자 II 결핍증	2147	P00734
인자 IX	B 형 혈우병	2158	P00740
인자 V	오우렌병	2153	P12259
인자 VIII	A 형 혈우병	2157	P00451
인자 X	스튜어트-프라워 인자 결핍증	2159	P00742
인자 XI	C 형 혈우병	2160	P03951
인자 XIII	피브린 안정화 인자 결핍증	2162, 2165	P00488, P05160
갈락토사민-6-설페이트 설파타제	점액다당류증 MPS IV(모르키오 증후군) 유형 IV-A	2588	P34059
갈락토실세라마이드 β-갈락토시다제	크라베병	2581	P54803

[0839]

강글리오시드 β-갈락토시다제	전신 GM1 강글리오시드증	2720	P16278
강글리오시드 β-갈락토시다제	GM2 강글리오시드증	2720	P16278
강글리오시드 β-갈락토시다제	스핑고지질증 유형 I	2720	P16278
강글리오시드 β-갈락토시다제	스핑고지질증 유형 II (소아형)	2720	P16278
강글리오시드 β-갈락토시다제	스핑고지질증 유형 III (성인형)	2720	P16278
글루코시다제 I	N-글리코실화 CDG IIb 의 선천성 장애	2548	P10253
글루코실세라마이드 β-글루코시다제	고셰병	2629	P04062
헤파란-S-설페이트 설파미다제	점액다당류증 MPS III(산필리포 증후군) 유형 III-A	6448	P51688
호모겐티세이트 옥시다제	알캅톤뇨증	3081	Q93099
히알루로니다제	점액다당류증 MPS IX(히알루로니다제 결핍증)	3373, 8692, 8372, 23553	Q12794, Q12891, O43820, Q2M3T9
이두로네이트 설페이트 설파타제	점액다당류증 MPS II (헌터 증후군)	3423	P22304
레시틴-콜레스테롤 아실트랜스퍼라제 (LCAT)	완전 LCAT 결핍증, 물고기눈병, 죽상동맥경화증, 고콜레스테롤혈증	3931	606967
라이신 옥시다제	글루타르산혈증 유형 I	4015	P28300
리소좀산 리파제	콜레스테릴 에스테르 축적병(CESD)	3988	P38571
리소좀산 리파제	리소좀산 리파제 결핍증	3988	P38571
리소좀산 리파제	볼만병	3988	P38571

[0840]

리소좀 펩스타틴-비감수성 펩티다제	세로이드 리포푸신증 후기 영아형(CLN2, 얀스키-빌쇼스키병)	1200	O14773
만노스(Man) 포스페이트 (P) 아이소머라제	N-글리코실화 CDG Ib 의 선천성 장애	4351	P34949
만노실-α-1,6-당단백질-β-1,2-N-아세틸글루코사미닐 트랜스퍼라제	N-글리코실화 CDG IIa 의 선천성 장애	4247	Q10469
메탈로프로테이나제-2	윈체스터 증후군	4313	P08253
메틸말로닐-CoA 뮤타제	메틸말론산혈증(비타민 b12 비-반응성)	4594	P22033
N-아세틸 갈락토사민 α-4-설페이트 설파타제 (아릴설파타제 B)	점액다당류증 MPS VI (마로토-라미 증후군)	411	P15848
N-아세틸-D-글루코사미니다제	점액다당류증 MPS III(산필리포 증후군) 유형 III-B	4669	P54802
N-아세틸-갈락토사미니다제	쉰들러병 유형 I(영아 중증형)	4668	P17050
N-아세틸-갈락토사미니다제	쉰들러병 유형 II(칸자키병, 성인-발병형)	4668	P17050
N-아세틸-갈락토사미니다제	쉰들러병 유형 III (중간형)	4668	P17050
N-아세틸-글루코사미닌-6-설페이트 설파타제	점액다당류증 MPS III(산필리포 증후군) 유형 III-D	2799	P15586

[0841]

N-아세틸글루코사미닐-1-포스포트랜스퍼라제	점액지질증 ML III(거짓-헐러 다발디스트로피)	79158	Q3T906
N-아세틸글루코사미닐-1-포스포트랜스퍼라제 촉매 서브유닛	점액지질증 ML II (I-세포병)	79158	Q3T906
N-아세틸글루코사미닐-1-포스포트랜스퍼라제, 기질-인식 서브유닛	점액지질증 ML III (거짓-헐러 다발디스트로피) 유형 III-C	84572	Q9UJJ9
N-아스파르틸글루코사미니다제	아스파르틸글루코사민뇨	175	P20933
뉴라미니다제 1 (시알리다제)	시알산증	4758	Q99519
팔미토일-단백질 티오에스테라제-1	세로이드 리포푸신증 성인형(CLN4, 쿠프스병)	5538	P50897
팔미토일-단백질 티오에스테라제-1	세로이드 리포푸신증 영아형(CLN1, 산타부리-할티아병)	5538	P50897
페닐알라닌 하이드록실라제	페닐케톤뇨증	5053	P00439
포스포만노뮤타제-2	N-글리코실화 CDG Ia 의 선천성 장애(오직 신경계 및 신경계-다기관 형태)	5373	O15305
포르포빌리노겐 데아미나제	급성 간헐성 포르피린증	3145	P08397
퓨린 뉴클레오시드 포르포릴라제	퓨린 뉴클레오시드 포스포릴라제 결핍증	4860	P00491
피리미딘 5' 뉴클레오티다제	용혈성 빈혈 및/또는 피리미딘 5' 뉴클레오티다제 결핍증	51251	Q9H0P0
스핑고미엘리나제	니만-픽병 유형 A	6609	P17405
스핑고미엘리나제	니만-픽병 유형 B	6609	P17405

스테롤 27-하이드록실라제	뇌건황색종증(콜레스타놀 지질증)	1593	Q02318
티미딘 포스포릴라제	미토콘드리아 신경위장 뇌근병증(MNGIE)	1890	P19971
트리헥소실세라마이드 α-갈락토시다제	파브리병	2717	P06280
티로시나제, 예를 들어, OCA1	백색증, 예를 들어, 눈백색증	7299	P14679
UDP-GlcNAc:돌리킬-P NAcGlc 포스포트랜스퍼라제	N-글리코실화 CDG Ij 의 선천성 장애	1798	Q9H3H5
UDP-N-아세틸글루코사민-2-에피머라제/N-아세틸만노사민 키나제, 시알린	시알뇨증(Sialuria) 프랑스형	10020	Q9Y223
우리카제	레슈-니한 증후군, 통풍	391051	단백질 없음
우리딘 디포스페이트 글루쿠로닐-트랜스퍼라제(예를 들어, UGT1A1)	크리글러-나자르 증후군	54658	P22309
α-1,2-만노실트랜스퍼라제	N-글리코실화 CDG Il 의 선천성 장애(608776)	79796	Q9H6U8
α-1,2-만노실트랜스퍼라제	N-글리코실화의 선천성 장애, 유형 I(pre-Golgi 글리코실화 결함)	79796	Q9H6U8
α-1,3-만노실트랜스퍼라제	N-글리코실화 CDG Ii 의 선천성 장애	440138	Q2TAA5
α-D-만노시다제	α-만노시도시스, 유형 I(중증) 또는 II(경증)	10195	Q92685
α-L-푸코시다제	푸코시드축적증	4123	Q9NTJ4

[0842]

α-l-이두로니다제	점액다당류증 MPS IH/S(헐러-샤이에 증후군)	2517	P04066
α-l-이두로니다제	점액다당류증 MPS I-H(헐러 증후군)	3425	P35475
α-l-이두로니다제	점액다당류증 MPS I-S(샤이에 증후군)	3425	P35475
β-1,4-갈락토실트랜스퍼라제	N-글리코실화 CDG IId 의 선천성 장애	3425	P35475
β-1,4-만노실트랜스퍼라제	N-글리코실화 CDG Ik 의 선천성 장애	2683	P15291
β-D-만노시다제	β-만노시도시스	56052	Q9BT22
β-갈락토시다제	점액다당류증 MPS IV(모르키오 증후군) 유형 IV-B	4126	O00462
β-글루쿠로니다제	점액다당류증 MPS VII(슬라이 증후군)	2720	P16278
β-헥소미니다제 A	테이-삭스병	2990	P08236
β-헥소미니다제 B	샌드호프병	3073	P06865

[0843]

[0844] 일부 구현예에서, 본 명세서에 기재된 이펙터는 표 54의 효소, 또는 이의 기능적 변이체, 예를 들어, 이의 상동체(예를 들어, 오솔로그 또는 파라로그) 또는 단편을 포함한다. 일부 구현예에서, 본 명세서에 기재된 이펙터는 UniProt ID를 참조하여 표 54에 열거된 아미노산 서열에 대하여 적어도 80%, 85%, 90%, 95%, 967%, 98%, 99% 서열 동일성을 갖는 단백질을 포함한다. 일부 구현예에서, 기능적 변이체는, 예를 들어, 야생형 단백질보다 10%, 20%, 30%, 40%, 또는 50% 이상 더 낮은 속도로, 상응하는 야생형 단백질과 동일한 반응을 촉매한다. 일부 구현예에서, 표 54의 효소, 또는 이의 기능적 변이체를 인코딩하는 아넬로벡터는 표 54의 질환 또는 장애의 치료에 사용된다. 일부 구현예에서, 아넬로벡터는 우리딘 디포스페이트 디포스페이트 글루쿠로닐-트랜스퍼라제 또는 이의 기능적 변이체를 표적 세포, 예를 들어, 간 세포로 운반하는 데 사용된다. 일부 구현예에서, 아넬로벡터는 OCA1 또는 이의 기능적 변이체를 표적 세포, 예를 들어, 망막 세포로 운반하는 데 사용된다.

[0845] [표 55] 예시적인 비-효소 이펙터 및 상응하는 적응증

이펙터	적응증	Entrez Gene ID	UniProt ID
생존 운동 뉴런 단백질(SMN)	척수성 근위축증	6606	Q16637
디스트로핀 또는 마이크로-디스트로핀	근이영양증(예를 들어, 뒤시엔느 근이영양증 또는 베커 근이영양증)	1756	P11532
보체 단백질, 예를 들어, 보체 인자 C1	보체 인자 I 결핍증	3426	P05156
보체 인자 H	비정형 용혈성 요독증후군	3075	P08603
시스틴증(리소좀 시스틴 수송체)	시스틴증	1497	O60931
부고환 분비 단백질 1(HE1; NPC2 단백질)	니만-픽병 유형 C2	10577	P61916
GDP-푸코스 수송체-1	N-글리코실화 CDG IIc 의 선천성 장애(람밤=-하샤론(Rambam-Hasharon) 증후군)	55343	Q96A29
GM2 활성화제 단백질	GM2 활성화제 단백질 결핍증(테이-삭스병 AB 변이형, GM2A)	2760	Q17900
리소좀 막횡단 CLN3 단백질	세로이드 리포푸신증 소아형(CLN3, 바텐병, 포그트-스필마이어병)	1207	Q13286
리소좀 막횡단 CLN5 단백질	세로이드 리포푸신증 변이형 후기 영아형, 핀란드형(CLN5)	1203	O75503

[0846]

Na 인산염 공동수송체, 시알린	영아 시알산 축적 장애	26503	Q9NRA2
Na 인산염 공동수송체, 시알린	시알뇨증 핀란드형(살라병)	26503	Q9NRA2
NPC1 단백질	니만-픽병 유형 C1/유형 D	4864	O15118
올리고머 골지 복합체-7	N-글리코실화 CDG IIe 의 선천성 장애	91949	P83436
프로사포신	프로사포신 결핍증	5660	P07602
보호 단백질/카텝신 A(PPCA)	갈락토시알리도증(골드버그 증후군, 조합된 뉴라미니다제 및 β-갈락토시다제 결핍증)	5476	P10619
만노스-P-돌리콜 이용에 수반되는 단백질	N-글리코실화 CDG If 의 선천성 장애	9526	O75352
사포신 B	사포신 B 결핍증(술파티드 활성화제 결핍증)	5660	P07602
사포신 C	사포신 C 결핍증(고셰 활성화제 결핍증)	5660	P07602
술파타제-변형 인자-1	뮤코술파티도시스(Mucosulfatidosis)(다발성 술파타제 결핍증)	285362	Q8NBK3
막횡단 CLN6 단백질	세로이드 리포푸신증 변이형 후기 영아형(CLN6)	54982	Q9NWW5
막횡단 CLN8 단백질	지적 장애를 갖는 세로이드 리포푸신증 진행성 간질	2055	Q9UBY8
vWF	폰 빌레브란트병	7450	P04275
인자 I(피브리노겐)	무피브리노겐혈증	2243, 2244, 2266	P02671, P02675, P02679
에리트로포이에틴(hEPO)			

[0847]

[0848] 일부 구현예에서, 본 명세서에 기재된 이펙터는 에리트로포이에틴(EPO), 예를 들어, 인간 에리트로포이에틴(hEPO), 또는 이의 기능적 변이체를 포함한다. 일부 구현예에서, 에리트로포이에틴, 또는 이의 기능적 변이체를 인코딩하는 아넬로벡터는 적혈구형성을 자극하는 데 사용된다. 일부 구현예에서, 에리트로포이에틴, 또는 이의 기능적 변이체를 인코딩하는 아넬로벡터는 질환 또는 장애, 예를 들어, 빈혈의 치료에 사용된다. 일부 구현예에서, 아넬로벡터는 EPO 또는 이의 기능적 변이체를 표적 세포, 예를 들어, 적혈구로 운반하는 데 사용된다.

[0849] 일부 구현예에서, 본 명세서에 기재된 이펙터는 표 55의 폴리펩티드, 또는 이의 기능적 변이체, 예를 들어, 이의 상동체(예를 들어, 오솔로그 또는 파라로그) 또는 단편을 포함한다. 일부 구현예에서, 본 명세서에 기재된 이펙터는 UniProt ID를 참조하여 표 55에 열거된 아미노산 서열에 대하여 적어도 80%, 85%, 90%, 95%, 967%, 98%, 99% 서열 동일성을 갖는 단백질을 포함한다. 일부 구현예에서, 표 55의 폴리펩티드, 또는 이의 기능적 변이체를 인코딩하는 아넬로벡터는 표 55의 질환 또는 장애의 치료에 사용된다. 일부 구현예에서, 아넬로벡터는 SMN 또는 이의 기능적 변이체를 표적 세포, 예를 들어, 척수 및/또는 운동 뉴런의 세포로 운반하는 데

사용된다. 일부 구현예에서, 아넬로벡터는 마이크로-디스트로핀을 표적 세포, 예를 들어, 근세포로 운반하는 데 사용된다.

[0850] 예시적인 마이크로-디스트로핀은 문헌[Duan, "Systemic AAV Micro-dystrophin Gene Therapy for Duchenne Muscular Dystrophy." Mol Ther. 2018 Oct 3;26(10):2337-2356. doi: 10.1016/j.ymthe.2018.07.011. Epub 2018 Jul 17]에 기재되어 있다.

[0851] 일부 구현예에서, 본 명세서에 기재된 이펙터는 응고 인자, 예를 들어, 본 명세서에서 표 54 또는 표 55에 열거된 응고 인자를 포함한다. 일부 구현예에서, 본 명세서에 기재된 이펙터는 돌연변이되는 경우 리소좀 축적 장애를 야기하는 단백질, 예를 들어, 본 명세서에서 표 54 또는 표 55에 열거된 단백질을 포함한다. 일부 구현예에서, 본 명세서에 기재된 이펙터는 수송체 단백질, 예를 들어, 본 명세서에서 표 55에 열거된 수송체 단백질을 포함한다.

[0852] 일부 구현예에서, 야생형 단백질의 기능적 변이체는 야생형 단백질의 하나 이상의 활성을 갖는 단백질을 포함하며, 예를 들어, 기능적 변이체는 예를 들어, 야생형 단백질보다 10%, 20%, 30%, 40% 또는 50% 이상 더 낮은 속도로, 상응하는 야생형 단백질과 동일한 반응을 촉매한다. 일부 구현예에서, 기능적 변이체는 야생형 단백질에 의해 결합되는 동일한 결합 파트너에, 예를 들어, 동일한 조건 하의 동일한 결합 파트너에 대한 상응하는 야생형 단백질의 Kd보다 10%, 20%, 30%, 40% 또는 50% 이하로 더 큰 Kd로 결합한다. 일부 구현예에서, 기능적 변이체는 야생형 폴리펩티드의 서열과 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% 또는 99% 동일한 폴리펩티드 서열을 갖는다. 일부 구현예에서, 기능적 변이체는 상응하는 야생형 단백질의 상동체(예를 들어, 오솔로그(오솔로그) 또는 파라로그(paralog))를 포함한다. 일부 구현예에서, 기능적 변이체는 융합 단백질이다. 일부 구현예에서, 융합물은 상응하는 야생형 단백질에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% 또는 99% 동일성을 갖는 제1 영역 및 제2 이종 영역을 포함한다. 일부 구현예에서, 기능적 변이체는 상응하는 야생형 단백질의 단편을 포함하거나, 이로 이루어진다.

[0853] **재생, 복구, 섬유증 인자**

[0854] 본 명세서에 기재된 치료용 폴리펩티드는 또한, 예를 들어, 표 56에 개시된 바와 같은, 성장 인자, 또는 이의 기능적 변이체, 예를 들어, UniProt ID를 참조하여 표 56에 개시된 단백질 서열에 대하여 적어도 80%, 85%, 90%, 95%, 967%, 98%, 99% 동일성을 갖는 단백질을 포함한다. 또한 재생 및 복구를 촉진하는 이러한 성장 인자, 또는 miRNA에 대한 항체 또는 이의 단편이 포함된다.

[0855] [표 56] 예시적인 재생, 복구, 및 섬유증 인자

표적	유전자 수탁 번호	단백질 수탁 번호
VEGF-A	NG_008732	NP_001165094
NRG-1	NG_012005	NP_001153471
FGF2	NG_029067	NP_001348594
FGF1	유전자 ID: 2246	NP_001341882
miR-199-3p	MIMAT0000232	
miR-590-3p	MIMAT0004801	
mi-17-92	MI0000071	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2732113/figure/F1/
miR-222	MI0000299	
miR-302-367	MIR302A And MIR367	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4400607/

[0856]

[0857] **형질전환 인자**

[0858] 본 명세서에 기재된 치료용 폴리펩티드는 또한, 예를 들어, 섬유아세포를 분화 세포로 형질전환시키는 단백질 인자, 예를 들어, 표 57에 개시된 인자 또는 이의 기능적 변이체, 예를 들어, UniProt ID를 참조하여 표 57에 개시된 단백질 서열에 대하여 적어도 80%, 85%, 90%, 95%, 967%, 98%, 99% 동일성을 갖는 단백질을 포함한다.

[0859] [표 57] 예시적인 형질전환 인자

표적	적응증	유전자 수탁 번호	단백질 수탁 번호
MESP1	섬유아세포 형질전환에 의한 장기 복구	유전자ID: 55897	EAX02066
ETS2	섬유아세포 형질전환에 의한 장기 복구	유전자 ID: 2114	NP_005230
HAND2	섬유아세포 형질전환에 의한 장기 복구	유전자 ID: 9464	NP_068808
MYOCARDIN	섬유아세포 형질전환에 의한 장기 복구	유전자 ID: 93649	NP_001139784
ESRRA	섬유아세포 형질전환에 의한 장기 복구	유전자 ID: 2101	AAH92470
miR-1	섬유아세포 형질전환에 의한 장기 복구	MI0000651	n/a
miR-133	섬유아세포 형질전환에 의한 장기 복구	MI0000450	n/a
TGFb	섬유아세포 형질전환에 의한 장기 복구	유전자 ID: 7040	NP_000651.3
WNT	섬유아세포 형질전환에 의한 장기 복구	유전자 ID: 7471	NP_005421
JAK	섬유아세포 형질전환에 의한 장기 복구	유전자 ID: 3716	NP_001308784
NOTCH	섬유아세포 형질전환에 의한 장기 복구	유전자 ID: 4851	XP_011517019

[0860]

[0861] **세포 재생을 자극하는 단백질**

[0862] 본 명세서에 기재된 치료용 폴리펩티드는 또한 세포 재생을 자극하는 단백질, 예를 들어 표 58에 개시된 단백질 또는 이의 기능적 변이체, 예를 들어, UniProt ID를 참조하여 표 58에 개시된 단백질 서열에 대하여 적어도 80%, 85%, 90%, 95%, 967%, 98%, 99% 동일성을 갖는 단백질을 포함한다.

[0863] [표 58] 세포 재생을 자극하는 예시적인 단백질

표적	유전자 수탁 번호	단백질 수탁 번호
MST1	NG_016454	NP_066278
STK30	유전자 ID: 26448	NP_036103
MST2	유전자 ID: 6788	NP_006272
SAV1	유전자 ID: 60485	NP_068590
LATS1	유전자 ID: 9113	NP_004681
LATS2	유전자 ID: 26524	NP_055387
YAP1	NG_029530	NP_001123617
CDKN2b	NG_023297	NP_004927
CDKN2a	NG_007485	NP_478102

[0864]

[0865] STING 조절제 이펙터

[0866] 일부 구현예에서, 본 명세서에 기재된 분비 이펙터는 STING/cGAS 신호전달을 조절한다. 일부 구현예에서, STING 조절제는 폴리펩티드, 예를 들어, 바이러스 폴리펩티드 또는 이의 기능적 변이체이다. 예를 들어, 이펙터는 문헌[Maringer et al. "Message in a bottle: lessons learned from antagonism of STING signalling during RNA virus infection" Cytokine & Growth Factor Reviews Volume 25, Issue 6, December 2014, Pages 669-679]에 기재된 STING 조절제(예를 들어, 억제제)를 포함할 수 있으며, 이는 본 명세서에 전문이 참조로 포함된다. 추가의 STING 조절제(예를 들어, 활성화제)는, 예를 들어, 문헌[Wang et al. "STING activator c-di-GMP enhances the anti-tumor effects of peptide vaccines in melanoma-bearing mice." Cancer Immunol Immunother. 2015 Aug;64(8):1057-66. doi: 10.1007/s00262-015-1713-5. Epub 2015 May 19]; 문헌[Bose "cGAS/STING Pathway in Cancer: Jekyll and Hyde Story of Cancer Immune Response" Int J Mol Sci. 2017 Nov; 18(11): 2456]; 및 문헌[Fu et al. "STING agonist formulated cancer vaccines can cure established tumors resistant to PD-1 blockade" Sci Transl Med. 2015 Apr 15; 7(283): 283ra52]에 기재되어 있으며, 이들 각각은 본 명세서에 전문이 참조로 포함된다.

[0867] 펩티드의 일부 예는 형광 태그 또는 마커, 항원, 펩티드 치료제, 천연적인 생물활성 펩티드로부터의 합성 또는 유사체 펩티드, 효능제 또는 길항제 펩티드, 항-미생물 펩티드, 표적화 또는 세포독성 펩티드, 분해 또는 자가-파괴 펩티드, 및 분해 또는 자가-파괴 펩티드들을 포함하지만, 이에 제한되지 않는다. 본 명세서에 기재된 본 발명에 유용한 펩티드는 또한, 항원-결합 펩티드, 예를 들어, 항원 결합 항체 또는 항체-유사 단편, 예컨대 단일 쇄 항체, 나노바디(예를 들어, 문헌[Steeland et al. 2016. Nanobodies as therapeutics: big opportunities for small antibodies. Drug Discov Today: 21(7):1076-113] 참조)를 포함한다. 이러한 항원 결합 펩티드는 시토졸 항원, 핵 항원 또는 소기관-내 항원에 결합할 수 있다.

[0868] 일부 구현예에서, 유전 요소는 작은 펩티드, 펩티드모방체(예를 들어, 펩토이드(peptoid)), 아미노산, 및 아미노산 유사체를 인코딩하는 서열을 포함한다. 이러한 치료제는 일반적으로 몰당 약 5,000 그램 미만의 분자량, 몰당 약 2,000 그램 미만의 분자량, 몰당 약 1,000 그램 미만의 분자량, 몰당 약 500 그램 미만의 분자량, 및 이러한 화합물의 염, 에스테르 및 다른 약제학적으로 허용 가능한 형태를 갖는다. 이러한 치료제는 신경전달물질, 호르몬, 약물, 독소, 바이러스 또는 미생물 입자, 합성 분자, 및 이의 효능제 또는 길항제를 포함할 수 있지만, 이에 제한되지 않는다.

[0869] 일부 구현예에서, 본 명세서에 기재된 조성물 또는 아넬로벡터는 특정 위치, 조직, 또는 세포를 표적화할 수 있는 리간드에 연결된 폴리펩티드를 포함한다.

[0870] *유전자 편집 구성성분*

[0871] 아넬로벡터의 유전 요소는 유전자 편집 시스템의 구성성분을 인코딩하는 하나 이상의 유전자를 포함할 수 있다. 예시적인 유전자 편집 시스템은 클러스터링된 규칙적으로 산재된 짧은 회문 반복부(CRISPR) 시스템, 아연 핑거 뉴클레아제(ZFN), 및 전사 활성화제-유사 이펙터-기반 뉴클레아제(TALEN)를 포함한다. ZFN, TALEN, 및 CRISPR-기반 방법은, 예를 들어, 문헌[Gaj et al. Trends Biotechnol. 31.7(2013):397-405]에 기재되어 있으며; 유전자 편집의 CRISPR 방법은, 예를 들어, 문헌[Guan et al., Application of CRISPR-Cas system in gene therapy: Pre-clinical progress in animal model. DNA Repair 2016 Oct;46:1-8. doi: 10.1016/j.dnarep.2016.07.004]; 문헌[Zheng et al., Precise gene deletion and replacement using the CRISPR/Cas9 system in human cells. BioTechniques, Vol. 57, No. 3, September 2014, pp. 115-124]에 기재되어 있다.

[0872] CRISPR 시스템은 원래 박테리아 및 조류에서 발견되는 적응 방어 시스템이다. CRISPR 시스템은 CRISPR-회합 또는 "Cas" 엔도뉴클레아제(예를 들어, Cas9 또는 Cpf1)로 명명되는 RNA-가이드된 뉴클레아제를 사용하여, 외래 DNA를 절단한다. 통상적인 CRISPR/Cas 시스템에서, 엔도뉴클레아제는 단일- 또는 이중-가닥 DNA 서열을 표적화하는 서열-특이적 비-코딩 "가이드 RNA"에 의해 표적 뉴클레오티드 서열(예를 들어, 서열-편집될 게놈 내의 부위)에 지향된다. CRISPR 시스템의 3가지 부류(I 내지 III)가 확인된 바 있다. 부류 II CRISPR 시스템은 (다수의 Cas 단백질이 아닌) 단일의 Cas 엔도뉴클레아제를 사용한다. 하나의 부류 II CRISPR 시스템은 II형 Cas 엔도뉴클레아제, 예컨대 Cas9, CRISPR RNA("crRNA"), 및 트랜스-활성화 crRNA("tracrRNA")를 포함한다. crRNA는 "가이드 RNA", 통상적으로 표적 DNA 서열에 상응하는 약 20-뉴클레오티드 RNA 서열을 함유한다. crRNA는 또한, tracrRNA에 결합하여, RNase III에 의해 절단되는 부분 이중-가닥 구조를 형성하여, crRNA/tracrRNA 혼성물을 초래하는 영역을 함유한다. 그 다음, crRNA/tracrRNA 혼성물은 Cas9 엔도뉴클레아제가 표적 DNA 서열을 인식하고 절단하게 한다. 표적 DNA 서열은 일반적으로 주어진 Cas 엔도뉴클레아제에 특이적인 "프로토스페이서 인접 모티프"("PAM")에 인접해야 하지만; PAM 서열은 주어진 게놈의 전체에 걸쳐 나타난다.

[0873] 일부 구현예에서, 아넬로벡터는 CRISPR 엔도뉴클레아제에 대한 유전자를 포함한다. 예를 들어, 다양한 원핵생물 종으로부터 확인되는 일부 CRISPR 엔도뉴클레아제는 독특한 PAM 서열 요건을 가지며; PAM 서열의 예는 5'-NGG(스트렙토코커스 피오게네스(Streptococcus pyogenes)), 5'-NNAGAA(스트렙토코커스 써모필러스(Streptococcus thermophilus) CRISPR1), 5'-NGGNG(스트렙토코커스 써모필러스 CRISPR3), 및 5'-NNNGATT(네이세리아 메닌지디티스(Neisseria meningiditis))를 포함한다. 일부 엔도뉴클레아제, 예를 들어, Cas9 엔도뉴클레아제는 G-풍부 PAM 부위, 예를 들어, 5'-NGG와 회합되며, PAM 부위(이로부터 5')로부터 3개 뉴클레오티드 상류의 위치에서 표적 DNA의 블런트-말단 절단을 수행한다. 또 다른 부류 II CRISPR 시스템은 Cas9보다 더 작은 V형 엔도뉴클레아제 Cpf1을 포함하며; 예는 AsCpf1(아시드아미노코커스(Acidaminococcus) 종 유래) 및 LbCpf1(라크노스피라세아에(Lachnospiraceae) 종 유래)을 포함한다. Cpf1 엔도뉴클레아제는 T-풍부 PAM 부위, 예를 들어, 5'-TTN과 회합된다. 또한, Cpf1은 5'-CTA PAM 모티프를 인식할 수 있다. Cpf1은 4- 또는 5-뉴클레오티드 5' 오버행을 갖는 오프셋 또는 스태거형 이중-가닥 파단부를 도입함으로써, 예를 들어, 코딩 가닥 상의 PAM 부위(이로부터 3')로부터 18개 뉴클레오티드 하류, 및 상보성 가닥 상의 PAM 부위로부터 23개 뉴클레오티드 하류에 위치한 5-뉴클레오티드 오프셋 또는 스태거형 절단으로 표적 DNA를 절단함으로써 표적 DNA를 절단하며; 이러한 오프셋 절단으로부터 초래되는 5-뉴클레오티드 오버행은 상동성 재조합에 의한 DNA 삽입에 의해 블런트-말단 절단된 DNA에서의 삽입에 의한 것보다 더 정밀한 게놈 편집을 가능하게 한다. 예를 들어, 문헌[Zetsche et al. (2015) Cell, 163:759 - 771]을 참조한다.

[0874] 다양한 CRISPR 회합(Cas) 유전자가 아넬로벡터 내에 포함될 수 있다. 유전자의 구체적인 예는 Cas1, Cas2, Cas3, Cas4, Cas5, Cas6, Cas7, Cas8, Cas9, Cas10, Cpf1, C2C1, 또는 C2C3을 포함하는 부류 II 시스템으로부터의 Cas 단백질을 인코딩하는 것이다. 일부 구현예에서, 아넬로벡터는 Cas 단백질, 예를 들어, 다양한 원핵생물 종 중 임의의 것의 유래일 수 있는 Cas9 단백질을 인코딩하는 유전자를 포함한다. 일부 구현예에서, 아넬로벡터는 특정 Cas 단백질, 예를 들어, 특정 프로토스페이서-인접 모티프(PAM) 서열을 인식하도록 선택되는 특정 Cas9 단백질을 인코딩하는 유전자를 포함한다. 일부 구현예에서, 아넬로벡터는 세포, 접합체, 배아 또는 동물 내로 도입되어, 예를 들어, 동일한, 유사한 또는 상이한 PAM 모티프를 포함하는 부위의 인식 및 변형을 가능하게 할 수 있는 2가지 이상의 상이한 Cas 단백질 또는 2가지 이상의 Cas 단백질을 인코딩하는 핵산을 포함한다. 일부 구현예에서, 아넬로벡터는 비활성화된 뉴클레아제를 갖는 변형된 Cas 단백질, 예를 들어, 뉴클레아제-결핍

Cas9를 인코딩하는 유전자를 포함한다.

[0875] 야생형 Cas9 단백질이 gRNA에 의해 표적화된 특정 DNA 서열에서 이중-가닥 파단부(DSB)를 생성하는 한편, 변형된 기능을 갖는 수많은 CRISPR 엔도뉴클레아제가 알려져 있으며, 예를 들어, "닉카제" 버전의 Cas엔도뉴클레아제(예를 들어, Cas9)는 오직 단일-가닥 파단만을 생성하며; 촉매적으로 비활성인 Cas 엔도뉴클레아제, 예를 들어, Cas9("dCas9")는 표적 DNA를 절단하지 않는다. dCas9를 인코딩하는 유전자를 이펙터 도메인을 인코딩하는 유전자와 융합시켜, 표적 유전자의 발현을 억제(CRISPRi)하거나 활성화(CRISPRa)시킬 수 있다. 예를 들어, 유전자는 전사 침묵화제(예를 들어, KRAB 도메인) 또는 전사 활성화제(예를 들어, dCas9-VP64 융합물)를 갖는 Cas9 융합물을 인코딩할 수 있다. 2개의 gRNA에 상동성인 표적 서열에 DSB를 생성하기 위하여, FokI 뉴클레아제에 융합된 촉매적 비활성 Cas9(dCas9)("dCas9-FokI")를 인코딩하는 유전자가 포함될 수 있다. 예를 들어, 애드진 레포지터리(Addgene repository)(Addgene, 75 Sidney St., Suite 550A, Cambridge, MA 02139; addgene.org/crispr/)에 개시되고, 이로부터 공개적으로 이용 가능한 수많은 CRISPR/Cas9 플라스미드를 참조한다. 각각 개별 가이드 RNA에 의해 유도되는, 2개의 별개 이중-가닥 파단을 도입하는 "이중 닉카제" Cas9는 더 정확한 게놈 편집을 달성하는 것으로 문헌[Ran et al. (2013) Cell, 154:1380 - 1389]에 기재되어 있다.

[0876] 진핵생물의 유전자를 편집하기 위한 CRISPR 기술은 미국 특허 출원 공개 2016/0138008A1 및 US2015/0344912A1, 및 미국 특허 8,697,359, 8,771,945, 8,945,839, 8,999,641, 8,993,233, 8,895,308, 8,865,406, 8,889,418, 8,871,445, 8,889,356, 8,932,814, 8,795,965, 및 8,906,616에 개시되어 있다. Cpf1 엔도뉴클레아제 및 상응하는 가이드 RNA 및 PAM 부위는 미국 특허 출원 공개 2016/0208243 A1에 개시되어 있다.

[0877] 일부 구현예에서, 아넬로벡터는 본 명세서에 기재된 폴리펩티드, 예를 들어, 표적화된 뉴클레아제, 예를 들어, Cas9, 예를 들어, 야생형 Cas9, 닉카제 Cas9(예를 들어, Cas9 D10A), 데드(dead) Cas9(dCas9), eSpCas9, Cpf1, C2C1 또는 C2C3, 및 gRNA를 인코딩하는 유전자를 포함한다. 뉴클레아제 및 gRNA(들)를 인코딩하는 유전자의 선택은 표적화된 돌연변이가 뉴클레오티드의 결실인지, 치환인지 또는 부가인지, 예를 들어, 표적화된 서열에 대한 뉴클레오티드의 결실인지, 치환인지 또는 부가인지에 의해 결정된다. (하나 이상의) 이펙터 도메인(예를 들어, VP64)의 전부 또는 일부(예를 들어, 이의 생물학적 활성 부분)와 테더링된 촉매적 비활성 엔도뉴클레아제, 예를 들어, 데드 Cas9(dCas9, 예를 들어, D10A; H840A)를 인코딩하는 유전자는 하나 이상의 표적 핵산 서열의 활성 및/또는 발현을 조절할 수 있는 키메라 단백질을 생성한다.

[0878] 일부 구현예에서, 아넬로벡터는 dCas9와 하나 이상의 이펙터 도메인의 전부 또는 일부(예를 들어, 전장 야생형 이펙터 도메인, 또는 이의 단편 또는 변이체, 예를 들어, 이의 생물학적 활성 부분)의 융합물을 인코딩하는 유전자를 포함하여 본 명세서에 기재된 방법에 유용한 키메라 단백질을 생성한다. 따라서, 일부 구현예에서, 아넬로벡터는 dCas9-메틸라제 융합물을 인코딩하는 유전자를 포함한다. 다른 일부 구현예에서, 아넬로벡터는 내인성 유전자를 표적화하기 위해 부위-특이적 gRNA와 함께 dCas9-효소 융합물을 인코딩하는 유전자를 포함한다.

[0879] 다른 양태에서, 아넬로벡터는 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20개, 또는 그 이상의 dCas9와 융합된 이펙터 도메인(전체 또는 생물학적 활성 부분)을 인코딩하는 유전자를 포함한다.

[0880] *조절 서열*

[0881] 일부 구현예에서, 유전 요소는 이펙터를 인코딩하는 서열에 작동 가능하게 연결된 조절 서열, 예를 들어, 프로모터 또는 인핸서를 포함한다.

[0882] 일부 구현예에서, 프로모터는 발현 생성물을 인코딩하는 DNA 서열에 인접하여 위치한 DNA 서열을 포함한다. 프로모터는 인접 DNA 서열에 작동 가능하게 연결될 수 있다. 프로모터는 통상적으로 프로모터가 존재하지 않는 경우 발현되는 생성물의 양에 비하여 DNA 서열로부터 발현되는 생성물의 양을 증가시킨다. 하나의 유기체로부터의 프로모터를 사용하여 또 다른 유기체로부터 기원한 DNA 서열로부터의 생성물 발현을 향상시킬 수 있다. 예를 들어, 척추동물 프로모터는 척추동물에서 해파리 GFP의 발현을 위해 사용될 수 있다. 추가로, 하나의 프로모터 요소는 탠덤으로 부착된 다수의 DNA 서열에 대하여 발현되는 생성물의 양을 증가시킬 수 있다. 그러므로, 하나의 프로모터 요소는 하나 이상의 생성물의 발현을 향상시킬 수 있다. 다수의 프로모터 요소는 당업자에게 널리 알려져 있다.

[0883] 일 구현예에서, 높은 수준의 구성적 발현이 요망된다. 이러한 프로모터의 예는 제한 없이, 레트로바이러스 라우스 육종 바이러스(RSV) 긴 말단 반복부(LTR) 프로모터/인핸서, 사이토메갈로바이러스(CMV) 즉시 조기 프로모터/인핸서(예를 들어, 문헌[Boshart et al, Cell, 41:521-530 (1985)] 참조), SV40 프로모터, 다이하이드로폴레이트 환원효소 프로모터, 세포질 β-액틴 프로모터 및 포스포글리세롤 키나제(PGK) 프로모터를 포함한다.

[0884] 또 다른 구현예에서, 유도성 프로모터가 요망될 수 있다. 유도성 프로모터는 제한 없이, 아연-유도성 양 메탈로티오닌(MT) 프로모터; 덱사메타손(Dex)-유도성 마우스 유방 종양 바이러스(MMTV) 프로모터; T7 중합효소 프로모터 시스템(WO 98/10088호); 테트라사이클린-억제성 시스템(문헌[Gossen et al, Proc. Natl. Acad. Sci. USA, 89:5547-5551 (1992)]); 테트라사이클린-유도성 시스템(문헌[Gossen et al., Science, 268:1766-1769 (1995)]; 또한, 문헌[Harvey et al., Curr. Opin. Chem. Biol., 2:512-518 (1998)] 참조); RU486-유도성 시스템(문헌[Wang et al., Nat. Biotech., 15:239-243 (1997)] 및 문헌[Wang et al., Gene Ther., 4:432-441 (1997)]); 및 라파마이신-유도성 시스템(문헌[Magari et al., J. Clin. Invest., 100:2865-2872 (1997)]; 문헌[Rivera et al., Nat. Medicine. 2:1028-1032 (1996)])을 포함하여, 예를 들어 시스로 또는 트랜스로 제공되는, 외인적으로 공급되는 화합물에 의해 조절되는 것들이다. 이러한 맥락에서 유용할 수 있는 다른 유형의 유도성 프로모터는 특정 생리학적 상태, 예를 들어, 온도, 급성 단계에 의해 또는 오직 복제하는 세포에서만 조절되는 것들이다.

[0885] 일부 구현예에서, 관심 유전자 또는 핵산 서열에 대한 고유 프로모터가 사용된다. 유전자 또는 핵산 서열의 발현이 고유 발현을 모방하는 것이 요망되는 경우에 고유 프로모터가 사용될 수 있다. 유전자 또는 다른 핵산 서열의 발현이 일시적으로 또는 발달적으로, 또는 조직-특이적 방식으로 또는 특정 전사 자극에 반응하여 조절되어야 하는 경우에, 고유 프로모터가 사용될 수 있다. 추가의 구현예에서, 다른 고유 발현 제어 요소, 예컨대, 인핸서 요소, 폴리아데닐화 부위 또는 코작(Kozak) 컨센서스 서열도 또한 고유 발현을 모방하기 위해 사용될 수 있다.

[0886] 일부 구현예에서, 유전 요소는 조직-특이적 프로모터에 작동 가능하게 연결된 유전자를 포함한다. 예를 들어, 골격근에서의 발현이 요망되는 경우, 근육에서 활성인 프로모터가 사용될 수 있다. 이들은 골격 α-액틴, 미오신 경쇄 2A, 디스트로핀, 근육 크레아틴 키나제를 인코딩하는 유전자로부터의 프로모터 및 자연-발생 프로모터보다 더 높은 활성을 갖는 합성 근육 프로모터를 포함한다. 문헌[Li et al., Nat. Biotech., 17:241-245 (1999)]을 참조한다. 특히 간 알부민, 문헌[Miyatake et al. J. Virol., 71:5124-32 (1997)]; B형 간염 바이러스 코어 프로모터, 문헌[Sandig et al., Gene Ther. 3:1002-9 (1996)]; 알파-태아단백질(AFP), 문헌[Arbuthnot et al., Hum. Gene Ther., 7:1503-14 (1996)], 골(오스테오칼신(osteocalcin), 문헌[Stein et al., Mol. Biol. Rep., 24:185-96 (1997)]; 골 시알로단백질(sialoprotein), 문헌[Chen et al., J. Bone Miner. Res. 11:654-64 (1996)]), 림프구(CD2, 문헌[Hansal et al., J. Immunol., 161:1063-8 (1998)]; 면역글로불린 중쇄; T 세포 수용체 a 쇄), 뉴런(뉴런-특이적 에놀라제(NSE) 프로모터, 문헌[Andersen et al. Cell. Mol. Neurobiol., 13:503-15 (1993)]; 신경필라멘트 경쇄 유전자, 문헌[Piccioli et al., Proc. Natl. Acad. Sci. USA, 88:5611-5 (1991)]; 뉴런-특이적 vgf 유전자, 문헌[Piccioli et al., Neuron, 15:373-84 (1995)]을 위한 조직-특이적인 프로모터의 예가 알려져 있다.

[0887] 유전 요소는 인핸서, 예를 들어, 유전자를 인코딩하는 DNA 서열에 인접하여 위치한 DNA 서열을 포함할 수 있다. 인핸서 요소는 통상적으로 프로모터 요소의 상류에 위치하거나, 코딩 DNA 서열(예를 들어, 생성물 또는 생성물들로 전사 또는 번역되는 DNA 서열)의 하류에 또는 그 내에 위치할 수 있다. 그러므로, 인핸서 요소는 생성물을 인코딩하는 DNA 서열의 100개 염기쌍, 200개 염기쌍 또는 300개 이상의 염기쌍 상류 또는 하류에 위치할 수 있다. 인핸서 요소는 DNA 서열로부터 발현되는 재조합 생성물의 양을 프로모터 요소에 의해 제공되는 증가된 발현을 초과하여 증가시킬 수 있다. 다중의 인핸서 요소는 당업자에게 용이하게 이용 가능하다.

[0888] 일부 구현예에서, 유전 요소는 본 명세서에 기재된 발현 생성물을 인코딩하는 서열을 측접하는 하나 이상의 역위 말단 반복부(ITR)를 포함한다. 일부 구현예에서, 유전 요소는 본 명세서에 기재된 발현 생성물을 인코딩하는 서열을 측접하는 하나 이상의 긴 말단 반복부(LTR)를 포함한다. 사용될 수 있는 프로모터 서열의 예는 시미안 바이러스 40(SV40) 조기 프로모터, 마우스 유방 종양 바이러스(MMTV), 인간 면역결핍 바이러스(HIV) 긴 말단 반복부(LTR) 프로모터, MoMuLV 프로모터, 조류 백혈병 바이러스 프로모터, 엡스테인-바 바이러스 즉시 조기 프로모터 및 라우스 육종 바이러스 프로모터를 포함하나 이들에 제한되지 않는다.

[0889] *복제 단백질*

[0890] 일부 구현예에서, 아넬로벡터, 예를 들어, 합성 아넬로벡터의 유전 요소는 하나 이상의 복제 단백질을 인코딩하는 서열을 포함할 수 있다. 일부 구현예에서, 아넬로벡터는 회전환 복제 방법에 의해 복제할 수 있으며, 예를 들어, 선도 가닥 및 지연 가닥의 합성은 분리된다. 이러한 구현예에서, 아넬로벡터는 3가지의 추가의 요소를 포함한다: i) 개시 단백질을 인코딩하는 유전자, ii) 이중 가닥 원점, 및 iii) 단일 가닥 원점. 복제 단백질을 포함하는 회전환 복제(RCR) 단백질 복합체는 선도 가닥에 결합하며, 복제 원점을 불안정화시킨다. RCR 복합체는

게놈을 절단하여, 유리 3'OH 말단을 생성한다. 세포 DNA 중합효소는 유리 3'OH 말단으로부터 바이러스 DNA 복제를 개시한다. 게놈이 복제된 후에, RCR 복합체는 루프를 공유적으로 폐쇄시킨다. 이는 양성 환형 단일-가닥 모 DNA 분자 및 음성 모체 가닥 및 새로 합성된 양성 가닥으로 구성된 환형 이중-가닥 DNA 분자의 방출을 야기한다. 단일-가닥 DNA 분자는 캡시드화되거나 제2차의 복제에 수반될 수 있다. 예를 들어, 문헌[Virology Journal 2009, 6:60 doi:10.1186/1743-422X-6-60]을 참조한다.

[0891] 유전 요소는 중합효소, 예를 들어, RNA 중합효소 또는 DNA 중합효소를 인코딩하는 서열을 포함할 수 있다.

[0892] *기타 서열*

[0893] 일부 구현예에서, 유전 요소는 생성물(예를 들어, 리보자임, 단백질을 인코딩하는 치료적 mRNA, 외인성 유전자)을 인코딩하는 핵산을 추가로 포함한다.

[0894] 일부 구현예에서, 유전 요소는 아넬로벡터의 종 및/또는 조직 및/또는 세포 향성(예를 들어, 캡시드 단백질 서열), 감염성(예를 들어, 캡시드 단백질 서열), 면역억제/활성화(예를 들어, 조절 핵산), 바이러스 게놈 결합 및/또는 패키징, 면역 회피(비-면역원성 및/또는 관용), 약동학, 엔도시토시스 및/또는 세포 부착, 핵 유입, 세포내 조절 및 국소화, 엑소시토시스 조절, 증식 및 숙주 또는 숙주 세포에서의 아넬로벡터의 핵산 보호에 영향을 미치는 하나 이상의 서열을 포함한다.

[0895] 일부 구현예에서, 유전 요소는 DNA, RNA 또는 인공 핵산을 포함하는 다른 서열을 포함할 수 있다. 다른 서열은 게놈 DNA, cDNA, 또는 tRNA, mRNA, rRNA, miRNA, gRNA, siRNA 또는 다른 RNAi 분자를 인코딩하는 서열을 포함할 수 있지만, 이들에 제한되지 않는다. 일 구현예에서, 유전 요소는 조절 핵산과 동일한 유전자 발현 생성물의 상이한 유전자좌를 표적화하기 위하여 siRNA를 인코딩하는 서열을 포함한다. 일 구현예에서, 유전 요소는 조절 핵산과 상이한 유전자 발현 생성물을 표적화하기 위하여 siRNA를 인코딩하는 서열을 포함한다.

[0896] 일부 구현예에서, 유전 요소는 하기의 서열 중 하나 이상을 추가로 포함한다: 하나 이상의 miRNA를 인코딩하는 서열, 하나 이상의 복제 단백질을 인코딩하는 서열, 외인성 유전자를 인코딩하는 서열, 치료제를 인코딩하는 서열, 조절 서열(예를 들어, 프로모터, 인핸서), 내인성 유전자를 표적화하는 하나 이상의 조절 서열(siRNA, lncRNA, shRNA)을 인코딩하는 서열, 및 치료적 mRNA 또는 단백질을 인코딩하는 서열.

[0897] 다른 서열은 약 2 내지 약 5000개의 뉴클레오티드, 약 10 내지 약 100개의 뉴클레오티드, 약 50 내지 약 150개의 뉴클레오티드, 약 100 내지 약 200개의 뉴클레오티드, 약 150 내지 약 250개의 뉴클레오티드, 약 200 내지 약 300개의 뉴클레오티드, 약 250 내지 약 350개의 뉴클레오티드, 약 300 내지 약 500개의 뉴클레오티드, 약 10 내지 약 1000개의 뉴클레오티드, 약 50 내지 약 1000개의 뉴클레오티드, 약 100 내지 약 1000개의 뉴클레오티드, 약 1000 내지 약 2000개의 뉴클레오티드, 약 2000 내지 약 3000개의 뉴클레오티드, 약 3000 내지 약 4000개의 뉴클레오티드, 약 4000 내지 약 5000개의 뉴클레오티드 또는 그 사이의 임의의 범위의 길이를 가질 수 있다.

[0898] *인코딩된 유전자*

[0899] 예를 들어, 유전 요소는 신호전달 생화학 경로와 연관된 유전자, 예를 들어, 신호전달 생화학 경로-연관 유전자 또는 폴리뉴클레오티드를 포함할 수 있다. 예에는 질환 연관 유전자 또는 폴리뉴클레오티드가 포함된다. "질환-연관" 유전자 또는 폴리뉴클레오티드는 비 질환 대조군의 조직 또는 세포와 비교하여, 질환-이환된 조직으로부터 유래된 세포에서 전사 또는 번역 생성물을 비정상적인 수준 또는 비정상적인 형태로 제공하는 임의의 유전자 또는 폴리뉴클레오티드를 지칭한다. 그것은 비정상적으로 높은 수준으로 발현되는 유전자일 수 있으며; 비정상적으로 낮은 수준으로 발현되는 유전자일 수 있으며, 여기서, 변경된 발현은 질환의 발생 및/또는 진행과 상호연관된다. 질환-연관 유전자는 또한, 질환의 병인의 직접적인 원인이 되거나, 질환의 병인을 담당하는 유전자(들)와 연관 불균형인 돌연변이(들) 또는 유전적 변이를 지니는 유전자를 지칭한다.

[0900] 질환-연관 유전자 및 폴리뉴클레오티드의 예는 맥쿠식-나탄스(McKusick-Nathans) 유전 의학 기관, 존스 홉킨스 대학(미국 메릴랜드주 볼티모어) 및 국립 생명과학 정보 센터, 국립 의학 도서관(미국 메릴랜드주 베데스다)로부터 이용 가능하다. 질환-연관 유전자 및 폴리뉴클레오티드의 예는 본 명세서에 전문이 참조로 포함되는 미국 특허 8,697,359의 표 A 및 B에 열거되어 있다. 질환 특이적 정보는 맥쿠식-나탄스 유전 의학 기관, 존스 홉킨스 대학(미국 메릴랜드주 볼티모어) 및 국립 생명과학 정보 센터, 국립 의학 도서관(미국 메릴랜드주 베데스다)로부터 이용 가능하다. 신호전달 생화학 경로-연관 유전자 및 폴리뉴클레오티드의 예는 본 명세서에 전문이 참조로 포함되는 미국 특허 8,697,359의 표 A 내지 C에 열거되어 있다.

[0901] 또한, 유전 요소는 본 명세서의 다른 곳에 기재된 바와 같이 표적화 모이어티를 인코딩할 수 있다. 이는 예를 들어, 당, 당지질 또는 단백질, 예컨대 항체를 인코딩하는 폴리뉴클레오티드를 삽입함으로써 달성될 수 있다. 당업자는 표적화 모이어티의 추가의 생성 방법을 안다.

[0902] *바이러스 서열*

[0903] 일부 구현예에서, 유전 요소는 적어도 하나의 바이러스 서열을 포함한다. 일부 구현예에서, 서열은 단일 가닥 DNA 바이러스, 예를 들어, 아넬로바이러스, 비드나바이러스, 써코바이러스, 제미니바이러스, 게노모바이러스, 이노바이러스, 마이크로바이러스, 나노바이러스, 파보바이러스 및 스피라바이러스로부터의 하나 이상의 서열에 대하여 상동성 또는 동일성을 갖는다. 일부 구현예에서, 서열은 이중 가닥 DNA 바이러스, 예를 들어, 아데노바이러스, 암풀라바이러스, 아스코바이러스, 아스파바이러스, 배큘로바이러스, 푸셀로바이러스, 글로불로바이러스, 구타바이러스, 하이트로사바이러스, 헤르페스바이러스, 이리도바이러스, 리포트릭스바이러스, 니마바이러스 및 폭스바이러스로부터의 하나 이상의 서열에 대하여 상동성 또는 동일성을 갖는다. 일부 구현예에서, 서열은 RNA 바이러스, 예를 들어, 알파바이러스, 푸로바이러스, 간염 바이러스, 호르데이바이러스, 토바모바이러스, 토브라바이러스, 트리코르나바이러스, 루비바이러스, 비르나바이러스, 시스토바이러스, 파르티티바이러스 및 레오바이러스로부터의 하나 이상의 서열에 대하여 상동성 또는 동일성을 갖는다.

[0904] 일부 구현예에서, 유전 요소는 비-병원성 바이러스, 예를 들어, 상리공생 바이러스, 예를 들어, 편리공생 바이러스, 예를 들어, 고유 바이러스, 예를 들어, 아넬로바이러스로부터의 하나 이상의 서열을 포함할 수 있다. 명명법의 최근의 변화는 인간 세포를 감염시킬 수 있는 3가지 아넬로바이러스를 알파토르크바이러스(TT), 베타토르크바이러스(TTM) 및 감마토르크바이러스(TTMD) 속의 아넬로바이러스과의 바이러스로 분류하였다. 현재까지, 아넬로바이러스는 임의의 인간 질환과 연관된 바 없다. 일부 구현예에서, 유전 요소는 토르크 테노 바이러스(TT), 환형 음성-센스 게놈을 갖는 비-외피, 단일-가닥 DNA 바이러스에 대하여 상동성 또는 동일성을 갖는 서열을 포함할 수 있다. 일부 구현예에서, 유전 요소는 SEN 바이러스, 센티넬(Sentinel) 바이러스, TTV-유사 미니 바이러스 및 TT 바이러스에 대하여 상동성 또는 동일성을 갖는 서열을 포함할 수 있다. TT 바이러스 유전형 6, TT 바이러스 군, TTV-유사 바이러스 DXL1 및 TTV-유사 바이러스 DXL2를 포함하는 상이한 유형의 TT 바이러스가 기재된 바 있다. 일부 구현예에서, 유전 요소는 더 작은 바이러스, 토르크 테노-유사 미니 바이러스(TTM) 또는 토르크 테노-유사 미디 바이러스(TTMD)로 지칭되는 TTV와 TTMV 사이의 게놈 크기를 갖는 제3 바이러스에 대하여 상동성 또는 동일성을 갖는 서열을 포함할 수 있다. 일부 구현예에서, 유전 요소는 본 명세서에, 예를 들어, 표 19에 기재된 뉴클레오티드 서열 중 어느 하나에 대하여 적어도 약 60%, 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98% 및 99% 뉴클레오티드 서열 동일성을 갖는 비-병원성 바이러스로부터의 하나 이상의 서열 또는 서열의 단편을 포함할 수 있다.

[0905] 일부 구현예에서, 유전 요소는 본 명세서에, 예를 들어, 표 41에 기재된 뉴클레오티드 서열 중 어느 하나에 대하여 적어도 약 60%, 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98% 및 99% 뉴클레오티드 서열 동일성을 갖는 실질적으로 비-병원성인 바이러스로부터의 하나 이상의 서열 또는 서열의 단편을 포함할 수 있다.

[0906] **[표 41] 아넬로바이러스 및 이들의 서열의 예.** 수탁번호 및 관련 서열 정보는 2018년 12월 11일에 참조한 바와 같이 www.ncbi.nlm.nih.gov/genbank/에서 수득될 수 있다.

수탁번호	설명
AB017613.1	토르크 테노 바이러스 16 DNA, 완전한 게놈, 단리주: TUS01
AB026345.1	ORF1 및 ORF2 에 대한 TT 바이러스 유전자, 완전한 코딩 서열, 단리주:TRM1
AB026346.1	ORF1 및 ORF2 에 대한 TT 바이러스 유전자, 완전한 코딩 서열, 단리주:TK16
AB026347.1	ORF1 및 ORF2 에 대한 TT 바이러스 유전자, 완전한 코딩 서열, 단리주:TP1-3
AB028669.1	ORF1 및 ORF2 에 대한 TT 바이러스 유전자, 완전한 게놈, 단리주:TJN02
AB030487.1	PORF2a, pORF2b, pORF1 에 대한 TT 바이러스 유전자, 완전한 코딩 서열, 클론:JaCHCTC19
AB030488.1	PORF2a, pORF2b, pORF1 에 대한 TT 바이러스 유전자, 완전한 코딩 서열, 클론:JaBD89
AB030489.1	PORF2a, pORF2b, pORF1 에 대한 TT 바이러스 유전자, 완전한 코딩 서열, 클론:JaBD98
AB038340.1	ORF2s, ORF1, ORF3 에 대한 TT 바이러스 유전자, 완전한 코딩 서열
AB038622.1	ORF2, ORF1, ORF3 에 대한 TT 바이러스 유전자, 완전한 코딩 서열, 단리주:TTVyon-LC011
AB038623.1	ORF2, ORF1, ORF3 에 대한 TT 바이러스 유전자, 완전한 코딩 서열, 단리주:TTVyon-KC186
AB038624.1	ORF2, ORF1, ORF3 에 대한 TT 바이러스 유전자, 완전한 코딩 서열, 단리주:TTVyon-KC197
AB041821.1	VP1 에 대한 TT 바이러스 mRNA, 완전한 코딩 서열
AB050448.1	ORF1, ORF2, ORF3, ORF4 에 대한 토르크 테노 바이러스 유전자, 완전한 코딩 서열, 단리주: TYM9
AB060592.1	ORF1, ORF2, ORF3, ORF4 에 대한 토르크 테노 바이러스 유전자, 클론: SAa-39
AB060593.1	ORF1, ORF2, ORF3, ORF4 에 대한 토르크 테노 바이러스 유전자, 완전한 코딩 서열, 클론: SAa-38
AB060595.1	ORF1, ORF2, ORF3, ORF4 에 대한 TT 바이러스 유전자, 완전한 코딩 서열, 클론:SAj-30
AB060596.1	ORF1, ORF2, ORF3, ORF4 에 대한 TT 바이러스 유전자, 완전한 코딩 서열, 클론:SAf-09
AB064596.1	토르크 테노 바이러스 DNA, 완전한 게놈, 단리주: CT25F
AB064597.1	토르크 테노 바이러스 DNA, 완전한 게놈, 단리주: CT30F
AB064599.1	토르크 테노 바이러스 DNA, 완전한 게놈, 단리주: JT03F
AB064600.1	토르크 테노 바이러스 DNA, 완전한 게놈, 단리주: JT05F
AB064601.1	토르크 테노 바이러스 DNA, 완전한 게놈, 단리주: JT14F
AB064602.1	토르크 테노 바이러스 DNA, 완전한 게놈, 단리주: JT19F
AB064603.1	토르크 테노 바이러스 DNA, 완전한 게놈, 단리주: JT41F
AB064604.1	토르크 테노 바이러스 DNA, 완전한 게놈, 단리주: CT39F
AB064606.1	토르크 테노 바이러스 DNA, 완전한 게놈, 단리주: JT33F

[0907]

[0908]

AB290918.1	토르크 테로 미디 바이러스 1 DNA, 완전한 게놈, 단리주: MD1-073
AF079173.1	TT 바이러스 주 TTVCHN1, 완전한 게놈
AF116842.1	TT 바이러스 주 BDH1, 완전한 게놈
AF122914.3	TT 바이러스 단리주 JA20, 완전한 게놈
AF122917.1	TT 바이러스 단리주 JA4, 완전한 게놈
AF122919.1	TT 바이러스 단리주 JA10 미공지 유전자
AF129887.1	TT 바이러스 TTVCHN2, 완전한 게놈
AF247137.1	TT 바이러스 단리주 TUPB, 완전한 게놈
AF254410.1	TT 바이러스 ORF2 단백질 및 ORF1 단백질 유전자, 완전한 코딩 서열
AF298585.1	TT 바이러스 폴리시(Polish) 단리주 P/1C1, 완전한 게놈
AF315076.1	TTV-유사 바이러스 DXL1 미공지 유전자
AF315077.1	TTV-유사 바이러스 DXL2 미공지 유전자
AF345521.1	TT 바이러스 단리주 TCHN-G1 Orf2 및 Orf1 유전자, 완전한 코딩 서열
AF345522.1	TT 바이러스 단리주 TCHN-E Orf2 및 Orf1 유전자, 완전한 코딩 서열
AF345525.1	TT 바이러스 단리주 TCHN-D2 Orf2 및 Orf1 유전자, 완전한 코딩 서열
AF345527.1	TT 바이러스 단리주 TCHN-C2 Orf2 및 Orf1 유전자, 완전한 코딩 서열
AF345528.1	TT 바이러스 단리주 TCHN-F Orf2 및 Orf1 유전자, 완전한 코딩 서열
AF345529.1	TT 바이러스 단리주 TCHN-G2 Orf2 및 Orf1 유전자, 완전한 코딩 서열
AF371370.1	TT 바이러스 ORF1, ORF3, 및 ORF2 유전자, 완전한 코딩 서열
AJ620212.1	토르크 테노 바이러스, 단리주 tth6, 완전한 게놈
AJ620213.1	토르크 테노 바이러스, 단리주 tth10, 완전한 게놈
AJ620214.1	토르크 테노 바이러스, 단리주 tth11g2, 완전한 게놈
AJ620215.1	토르크 테노 바이러스, 단리주 tth18, 완전한 게놈
AJ620216.1	토르크 테노 바이러스, 단리주 tth20, 완전한 게놈
AJ620217.1	토르크 테노 바이러스, 단리주 tth21, 완전한 게놈
AJ620218.1	토르크 테노 바이러스, 단리주 tth3, 완전한 게놈
AJ620219.1	토르크 테노 바이러스, 단리주 tth9, 완전한 게놈
AJ620220.1	토르크 테노 바이러스, 단리주 tth16, 완전한 게놈
AJ620221.1	토르크 테노 바이러스, 단리주 tth17, 완전한 게놈
AJ620222.1	토르크 테노 바이러스, 단리주 tth25, 완전한 게놈
AJ620223.1	토르크 테노 바이러스, 단리주 tth26, 완전한 게놈
AJ620224.1	토르크 테노 바이러스, 단리주 tth27, 완전한 게놈
AJ620225.1	토르크 테노 바이러스, 단리주 tth31, 완전한 게놈

AJ620226.1	토르크 테노 바이러스, 단리주 tth4, 완전한 게놈
AJ620227.1	토르크 테노 바이러스, 단리주 tth5, 완전한 게놈
AJ620228.1	토르크 테노 바이러스, 단리주 tth14, 완전한 게놈
AJ620229.1	토르크 테노 바이러스, 단리주 tth29, 완전한 게놈
AJ620230.1	토르크 테노 바이러스, 단리주 tth7, 완전한 게놈
AJ620231.1	토르크 테노 바이러스, 단리주 tth8, 완전한 게놈
AJ620232.1	토르크 테노 바이러스, 단리주 tth13, 완전한 게놈
AJ620233.1	토르크 테노 바이러스, 단리주 tth19, 완전한 게놈
AJ620234.1	토르크 테노 바이러스, 단리주 tth22g4, 완전한 게놈
AJ620235.1	토르크 테노 바이러스, 단리주 tth23, 완전한 게놈
AM711976.1	TT 바이러스 sle1957 완전한 게놈
AM712003.1	TT 바이러스 sle1931 완전한 게놈
AM712004.1	TT 바이러스 sle1932 완전한 게놈
AM712030.1	TT 바이러스 sle2057 완전한 게놈
AM712031.1	TT 바이러스 sle2058 완전한 게놈
AM712032.1	TT 바이러스 sle2072 완전한 게놈
AM712033.1	TT 바이러스 sle2061 완전한 게놈
AM712034.1	TT 바이러스 sle2065 완전한 게놈
AY026465.1	TT 바이러스 단리주 L01ORF2및 ORF1유전자, 완전한 코딩 서열
AY026466.1	TT 바이러스 단리주 L02ORF2및 ORF1유전자, 완전한 코딩 서열
DQ003341.1	토르크 테노 바이러스 클론 P2-9-02 ORF2 (ORF2), ORF1A (ORF1A), 및 ORF1B (ORF1B) 유전자, 완전한 코딩 서열
DQ003342.1	토르크 테노 바이러스 클론 P2-9-07 ORF2 (ORF2), ORF1A (ORF1A), 및 ORF1B (ORF1B) 유전자, 완전한 코딩 서열
DQ003343.1	토르크 테노 바이러스 클론 P2-9-08 ORF2 (ORF2), ORF1A (ORF1A), 및 ORF1B (ORF1B) 유전자, 완전한 코딩 서열
DQ003344.1	토르크 테노 바이러스 클론 P2-9-16 ORF2 (ORF2), ORF1A (ORF1A), 및 ORF1B (ORF1B) 유전자, 완전한 코딩 서열
DQ186994.1	토르크 테노 바이러스 클론 P601 ORF2 (ORF2) 및 ORF1 (ORF1) 유전자, 완전한 코딩 서열
DQ186995.1	토르크 테노 바이러스 클론 P605 ORF2 (ORF2) 및 ORF1 (ORF1) 유전자, 완전한 코딩 서열

[0909]

DQ186996.1	토르크 테노 바이러스 클론 BM1A-02 ORF2 (ORF2) 및 ORF1 (ORF1) 유전자, 완전한 코딩 서열
DQ186997.1	토르크 테노 바이러스 클론 BM1A-09 ORF2 (ORF2) 및 ORF1 (ORF1) 유전자, 완전한 코딩 서열
DQ186998.1	토르크 테노 바이러스 클론 BM1A-13 ORF2 (ORF2) 및 ORF1 (ORF1) 유전자, 완전한 코딩 서열
DQ186999.1	토르크 테노 바이러스 클론 BM1B-05 ORF2 (ORF2) 및 ORF1 (ORF1) 유전자, 완전한 코딩 서열
DQ187000.1	토르크 테노 바이러스 클론 BM1B-07 ORF2 (ORF2) 및 ORF1 (ORF1) 유전자, 완전한 코딩 서열
DQ187001.1	토르크 테노 바이러스 클론 BM1B-11 ORF2 (ORF2) 및 ORF1 (ORF1) 유전자, 완전한 코딩 서열
DQ187002.1	토르크 테노 바이러스 클론 BM1B-14 ORF2 (ORF2) 및 ORF1 (ORF1) 유전자, 완전한 코딩 서열
DQ187003.1	토르크 테노 바이러스 클론 BM1B-08 ORF2 (ORF2) 유전자, 완전한 코딩 서열; 및 비기능적 ORF1 (ORF1) 유전자, 완전한 서열
DQ187004.1	토르크 테노 바이러스 클론 BM1C-16 ORF2 (ORF2) 및 ORF1 (ORF1) 유전자, 완전한 코딩 서열
DQ187005.1	토르크 테노 바이러스 클론 BM1C-10 ORF2 (ORF2) 및 ORF1 (ORF1) 유전자, 완전한 코딩 서열
DQ187007.1	토르크 테노 바이러스 클론 BM2C-25 ORF2 (ORF2) 유전자, 완전한 코딩 서열; 및 비기능적 ORF1 (ORF1) 유전자, 완전한 서열
DQ361268.1	토르크 테노 바이러스 단리주 ViPi04 ORF1 유전자, 완전한 코딩 서열
EF538879.1	토르크 테노 바이러스 단리주 CSC5ORF2및 ORF1유전자, 완전한 코딩 서열
EU305675.1	토르크 테노 바이러스 단리주 LTT7 ORF1 유전자, 완전한 코딩 서열
EU305676.1	토르크 테노 바이러스 단리주 LTT10 ORF1 유전자, 완전한 코딩 서열
EU889253.1	토르크 테노 바이러스 단리주 ViPi08 비기능적 ORF1 유전자, 완전한 서열
FJ392105.1	토르크 테노 바이러스 단리주 TW53A25 ORF2 유전자, 부분 코딩 서열; 및 ORF1 유전자, 완전한 코딩 서열
FJ392107.1	토르크 테노 바이러스 단리주 TW53A27 ORF2 유전자, 부분 코딩 서열; 및 ORF1 유전자, 완전한 코딩 서열
FJ392108.1	토르크 테노 바이러스 단리주 TW53A29 ORF2 유전자, 부분 코딩 서열; 및 ORF1 유전자, 완전한 코딩 서열

[0910]

FJ392111.1	토르크 테노 바이러스 단리주 TW53A35 ORF2 유전자, 부분 코딩 서열; 및 ORF1 유전자, 완전한 코딩 서열
FJ392112.1	토르크 테노 바이러스 단리주 TW53A39 ORF2 유전자, 부분 코딩 서열; 및 ORF1 유전자, 완전한 코딩 서열
FJ392113.1	토르크 테노 바이러스 단리주 TW53A26 ORF2 유전자, 완전한 코딩 서열; 및 비기능적 ORF1 유전자, 완전한 서열
FJ392114.1	토르크 테노 바이러스 단리주 TW53A30 ORF2및 ORF1유전자, 완전한 코딩 서열
FJ392115.1	토르크 테노 바이러스 단리주 TW53A31 ORF2및 ORF1유전자, 완전한 코딩 서열
FJ392117.1	토르크 테노 바이러스 단리주 TW53A37 ORF1 유전자, 완전한 코딩 서열
FJ426280.1	토르크 테노 바이러스 주 SIA109, 완전한 게놈
FR751500.1	토르크 테노 바이러스 완전한 게놈, 단리주 TTV-HD23a (rheu215)
GU797360.1	토르크 테노 바이러스 클론 8-17, 완전한 게놈
HC742700.1	특허 WO2010044889 로부터의 서열 7
HC742710.1	특허 WO2010044889 로부터의 서열 17
JX134044.1	TTV-유사 미니 바이러스 단리주 TTMV_LY1, 완전한 게놈
JX134045.1	TTV-유사 미니 바이러스 단리주 TTMV_LY2, 완전한 게놈
KU243129.1	TTV-유사 미니 바이러스 단리주 TTMV-204, 완전한 게놈
KY856742.1	TTV-유사 미니 바이러스 단리주 zhenjiang, 완전한 게놈
LC381845.1	토르크 테노 바이러스 Human/Japan/KS025/2016 DNA, 완전한 게놈
MH648892.1	아넬로비리대 종 단리주 ctdc048, 완전한 게놈
MH648893.1	아넬로비리대 종 단리주 ctdh007, 완전한 게놈
MH648897.1	아넬로비리대 종 단리주 ctcb038, 완전한 게놈
MH648900.1	아넬로비리대 종 단리주 ctfc019, 완전한 게놈
MH648901.1	아넬로비리대 종 단리주 ctbb022, 완전한 게놈
MH648907.1	아넬로비리대 종 단리주 ctcf040, 완전한 게놈
MH648911.1	아넬로비리대 종 단리주 cthi018, 완전한 게놈
MH648912.1	아넬로비리대 종 단리주 ctea38, 완전한 게놈
MH648913.1	아넬로비리대 종 단리주 ctbg006, 완전한 게놈
MH648916.1	아넬로비리대 종 단리주 ctbg020, 완전한 게놈
MH648925.1	아넬로비리대 종 단리주 ctci019, 완전한 게놈
MH648932.1	아넬로비리대 종 단리주 ctid031, 완전한 게놈
MH648946.1	아넬로비리대 종 단리주 ctdb017, 완전한 게놈
MH648957.1	아넬로비리대 종 단리주 ctch017, 완전한 게놈

[0911]

MH648958.1	아넬로비리대 종 단리주 ctbh011, 완전한 게놈
MH648959.1	아넬로비리대 종 단리주 ctbc020, 완전한 게놈
MH648962.1	아넬로비리대 종 단리주 ctif015, 완전한 게놈
MH648966.1	아넬로비리대 종 단리주 ctei055, 완전한 게놈
MH648969.1	아넬로비리대 종 단리주 ctjg000, 완전한 게놈
MH648976.1	아넬로비리대 종 단리주 ctcj064, 완전한 게놈
MH648977.1	아넬로비리대 종 단리주 ctbj022, 완전한 게놈
MH648982.1	아넬로비리대 종 단리주 ctbf014, 완전한 게놈
MH648983.1	아넬로비리대 종 단리주 ctbd027, 완전한 게놈
MH648985.1	아넬로비리대 종 단리주 ctch016, 완전한 게놈
MH648986.1	아넬로비리대 종 단리주 ctbd020, 완전한 게놈
MH648989.1	아넬로비리대 종 단리주 ctga035, 완전한 게놈
MH648990.1	아넬로비리대 종 단리주 cthf001, 완전한 게놈
MH648995.1	아넬로비리대 종 단리주 ctbd067, 완전한 게놈
MH648997.1	아넬로비리대 종 단리주 ctce026, 완전한 게놈
MH648999.1	아넬로비리대 종 단리주 ctfb058, 완전한 게놈
MH649002.1	아넬로비리대 종 단리주 ctjj046, 완전한 게놈
MH649006.1	아넬로비리대 종 단리주 ctcf030, 완전한 게놈
MH649008.1	아넬로비리대 종 단리주 ctbg025, 완전한 게놈
MH649011.1	아넬로비리대 종 단리주 ctbh052, 완전한 게놈
MH649014.1	아넬로비리대 종 단리주 ctba003, 완전한 게놈
MH649017.1	아넬로비리대 종 단리주 ctbb016, 완전한 게놈
MH649022.1	아넬로비리대 종 단리주 ctch023, 완전한 게놈
MH649023.1	아넬로비리대 종 단리주 ctbd051, 완전한 게놈
MH649028.1	아넬로비리대 종 단리주 ctbf9, 완전한 게놈
MH649038.1	아넬로비리대 종 단리주 ctbi030, 완전한 게놈
MH649039.1	아넬로비리대 종 단리주 ctca057, 완전한 게놈
MH649040.1	아넬로비리대 종 단리주 ctch033, 완전한 게놈
MH649042.1	아넬로비리대 종 단리주 ctjd005, 완전한 게놈
MH649045.1	아넬로비리대 종 단리주 ctdc021, 완전한 게놈
MH649051.1	아넬로비리대 종 단리주 ctdg044, 완전한 게놈
MH649056.1	아넬로비리대 종 단리주 ctcc062, 완전한 게놈
MH649061.1	아넬로비리대 종 단리주 ctid009, 완전한 게놈

[0912]

MH649062.1	아넬로비리대 종 단리주 ctdc018, 완전한 게놈
MH649063.1	아넬로비리대 종 단리주 ctbf012, 완전한 게놈
MH649068.1	아넬로비리대 종 단리주 ctcc066, 완전한 게놈
MH649070.1	아넬로비리대 종 단리주 ctda011, 완전한 게놈
MH649077.1	아넬로비리대 종 단리주 ctbh034, 완전한 게놈
MH649083.1	아넬로비리대 종 단리주 ctdg028, 완전한 게놈
MH649084.1	아넬로비리대 종 단리주 ctii061, 완전한 게놈
MH649085.1	아넬로비리대 종 단리주 cteh021, 완전한 게놈
MH649092.1	아넬로비리대 종 단리주 ctbg012, 완전한 게놈
MH649101.1	아넬로비리대 종 단리주 ctif053, 완전한 게놈
MH649104.1	아넬로비리대 종 단리주 ctei657, 완전한 게놈
MH649106.1	아넬로비리대 종 단리주 ctca015, 완전한 게놈
MH649114.1	아넬로비리대 종 단리주 ctbf050, 완전한 게놈
MH649122.1	아넬로비리대 종 단리주 ctdc002, 완전한 게놈
MH649125.1	아넬로비리대 종 단리주 ctbb15, 완전한 게놈
MH649127.1	아넬로비리대 종 단리주 ctba013, 완전한 게놈
MH649137.1	아넬로비리대 종 단리주 ctbb000, 완전한 게놈
MH649141.1	아넬로비리대 종 단리주 ctbc019, 완전한 게놈
MH649142.1	아넬로비리대 종 단리주 ctid026, 완전한 게놈
MH649144.1	아넬로비리대 종 단리주 ctfj004, 완전한 게놈
MH649152.1	아넬로비리대 종 단리주 ctcj13, 완전한 게놈
MH649156.1	아넬로비리대 종 단리주 ctci006, 완전한 게놈
MH649157.1	아넬로비리대 종 단리주 ctbd025, 완전한 게놈
MH649158.1	아넬로비리대 종 단리주 ctbf005, 완전한 게놈
MH649161.1	아넬로비리대 종 단리주 ctcf045, 완전한 게놈
MH649165.1	아넬로비리대 종 단리주 ctcc29, 완전한 게놈
MH649169.1	아넬로비리대 종 단리주 ctib021, 완전한 게놈
MH649172.1	아넬로비리대 종 단리주 ctbh857, 완전한 게놈
MH649174.1	아넬로비리대 종 단리주 ctbj049, 완전한 게놈
MH649178.1	아넬로비리대 종 단리주 ctfc006, 완전한 게놈
MH649179.1	아넬로비리대 종 단리주 ctbe000, 완전한 게놈
MH649183.1	아넬로비리대 종 단리주 ctbb031, 완전한 게놈
MH649186.1	아넬로비리대 종 단리주 ctcb33, 완전한 게놈

[0913]

[0914]

MH649189.1	아넬로비리대 종 단리주 ctcc12, 완전한 게놈
MH649196.1	아넬로비리대 종 단리주 ctci060, 완전한 게놈
MH649199.1	아넬로비리대 종 단리주 ctbb017, 완전한 게놈
MH649203.1	아넬로비리대 종 단리주 cthc018, 완전한 게놈
MH649204.1	아넬로비리대 종 단리주 ctbj003, 완전한 게놈
MH649206.1	아넬로비리대 종 단리주 ctbg010, 완전한 게놈
MH649208.1	아넬로비리대 종 단리주 ctid008, 완전한 게놈
MH649209.1	아넬로비리대 종 단리주 ctbg056, 완전한 게놈
MH649210.1	아넬로비리대 종 단리주 ctda001, 완전한 게놈
MH649212.1	아넬로비리대 종 단리주 ctcf004, 완전한 게놈
MH649217.1	아넬로비리대 종 단리주 ctbe029, 완전한 게놈
MH649223.1	아넬로비리대 종 단리주 ctci016, 완전한 게놈
MH649224.1	아넬로비리대 종 단리주 ctce11, 완전한 게놈
MH649228.1	아넬로비리대 종 단리주 ctcf013, 완전한 게놈
MH649229.1	아넬로비리대 종 단리주 ctcb036, 완전한 게놈
MH649241.1	아넬로비리대 종 단리주 ctda027, 완전한 게놈
MH649242.1	아넬로비리대 종 단리주 ctbf003, 완전한 게놈
MH649254.1	아넬로비리대 종 단리주 ctjb007, 완전한 게놈
MH649255.1	아넬로비리대 종 단리주 ctbb023, 완전한 게놈
MH649256.1	아넬로비리대 종 단리주 ctca002, 완전한 게놈
MH649258.1	아넬로비리대 종 단리주 ctcg010, 완전한 게놈
MH649263.1	아넬로비리대 종 단리주 ctgh3, 완전한 게놈
MK012439.1	아넬로비리대 종 단리주 cthe000, 완전한 게놈
MK012440.1	아넬로비리대 종 단리주 ctjd008, 완전한 게놈
MK012448.1	아넬로비리대 종 단리주 ctch012, 완전한 게놈
MK012457.1	아넬로비리대 종 단리주 ctda009, 완전한 게놈
MK012458.1	아넬로비리대 종 단리주 ctcd015, 완전한 게놈
MK012485.1	아넬로비리대 종 단리주 ctfd011, 완전한 게놈
MK012489.1	아넬로비리대 종 단리주 ctba003, 완전한 게놈
MK012492.1	아넬로비리대 종 단리주 ctbb005, 완전한 게놈
MK012493.1	아넬로비리대 종 단리주 ctcj014, 완전한 게놈
MK012500.1	아넬로비리대 종 단리주 ctcb001, 완전한 게놈
MK012504.1	아넬로비리대 종 단리주 ctcj010, 완전한 게놈

MK012516.1	아넬로비리대 종 단리주 ctcf003, 완전한 게놈
NC_038336.1	토르크 테노 바이러스 5 단리주 TCHN-C1 Orf2 및 Orf1 유전자, 완전한 코딩 서열
NC_038338.1	토르크 테노 바이러스 11 단리주 TCHN-D1 Orf2 및 Orf1 유전자, 완전한 코딩 서열
NC_038339.1	토르크 테노 바이러스 13 단리주 TCHN-A Orf2 및 Orf1 유전자, 완전한 코딩 서열
NC_038340.1	토르크 테노 바이러스 20 ORF4, ORF3, ORF2, ORF1 유전자, 완전한 코딩 서열, 클론: SAa-10
NC_038341.1	토르크 테노 바이러스 21 단리주 TCHN-B Orf2 및 Orf1 유전자, 완전한 코딩 서열
NC_038342.1	토르크 테노 바이러스 23 ORF2, ORF1 유전자, 완전한 코딩 서열, 단리주: s-TTV CH65-2
NC_038343.1	토르크 테노 바이러스 24 ORF4, ORF3, ORF2, ORF1 유전자, 완전한 코딩 서열, 클론: SAa-01
NC_038344.1	토르크 테노 바이러스 29 ORF2, ORF1, ORF3 유전자, 완전한 코딩 서열, 단리주: TTVyon-KC009
NC_038345.1	토르크 테노 미니 바이러스 10 단리주 LIL-y1 ORF2, ORF1, ORF3, 및 ORF4 유전자, 완전한 코딩 서열
NC_038346.1	토르크 테노 미니 바이러스 11 단리주 LIL-y2 ORF2, ORF1, 및 ORF3 유전자, 완전한 코딩 서열
NC_038347.1	토르크 테노 미니 바이러스 12 단리주 LIL-y3 ORF2, ORF1, ORF3, 및 ORF4 유전자, 완전한 코딩 서열
NC_038350.1	토르크 테로 미디 바이러스 3 단리주 2PoSMA Orf2 및 Orf1 유전자, 완전한 코딩 서열
NC_038351.1	토르크 테로 미디 바이러스 4 단리주 6PoSMA ORF2, ORF1, 및 ORF3 유전자, 완전한 코딩 서열
NC_038352.1	토르크 테로 미디 바이러스 5 DNA, 완전한 게놈, 단리주: MDJHem2
NC_038353.1	토르크 테로 미디 바이러스 6 DNA, 완전한 게놈, 단리주: MDJHem3-1
NC_038354.1	토르크 테로 미디 바이러스 7 DNA, 완전한 게놈, 단리주: MDJHem3-2
NC_038355.1	토르크 테로 미디 바이러스 8 DNA, 완전한 게놈, 단리주: MDJN1
NC_038356.1	토르크 테로 미디 바이러스 9 DNA, 완전한 게놈, 단리주: MDJN2
NC_038357.1	토르크 테로 미디 바이러스 10 DNA, 완전한 게놈, 단리주: MDJN14
NC_038358.1	토르크 테로 미디 바이러스 11 DNA, 완전한 게놈, 단리주: MDJN47
NC_038359.1	토르크 테로 미디 바이러스 12 DNA, 완전한 게놈, 단리주: MDJN51
NC_038360.1	토르크 테로 미디 바이러스 13 DNA, 완전한 게놈, 단리주: MDJN69
NC_038361.1	토르크 테로 미디 바이러스 14 DNA, 완전한 게놈, 단리주: MDJN97
NC_038362.1	토르크 테로 미디 바이러스 15 DNA, 완전한 게놈, 단리주: Pt-TTMDV210

[0915]

[0916] 일부 구현예에서, 유전 요소는 하나 이상의 비-아넬로바이러스, 예를 들어, 아데노바이러스, 헤르페스바이러스, 폭스 바이러스, 백시니아 바이러스, SV40, 파필로마 바이러스, RNA 바이러스, 예컨대 레트로바이러스, 예를 들어, 렌티 바이러스, 단일-가닥 RNA 바이러스, 예를 들어, 간염 바이러스, 또는 이중-가닥 RNA 바이러스, 예를 들어, 로타바이러스로부터의 하나 이상의 서열에 대하여 상동성 또는 동일성을 갖는 하나 이상의 서열을 포함한다. 일부 구현예에서, 재조합 레트로바이러스가 결함이 있기 때문에, 감염성 입자를 생성하도록 지원이 제공될 수 있다. 이러한 지원은 예를 들어, LTR 내의 조절 서열의 제어 하에 레트로바이러스의 구조적 유전자의 전부를 인코딩하는 플라스미드를 함유하는 헬퍼 세포주를 사용함으로써 제공될 수 있다. 본 명세서에 기재된 아넬로벡터를 복제하는 데 적합한 세포주는 당업계에 알려진 세포주, 예를 들어, 본 명세서에 기재된 바와 같이 변형될 수 있는 A549 세포를 포함한다. 상기 유전 요소는 선택 가능한 마커를 인코딩하는 유전자를 추가로 함유하여, 요망되는 유전 요소가 확인되게 할 수 있다.

[0917] 일부 구현예에서, 유전 요소는 서열이 제1 뉴클레오티드 서열에 의해 인코딩되는 폴리펩티드와 적어도 약 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% 또는 99% 동일하게 유지되거나, 또는 다르게 본 발명을 실시하는 데 유용하다면, 인코딩된 폴리펩티드에 아미노산 차이를 초래하는 비-침묵 돌연변이, 예를 들어, 염기 치환, 결실 또는 부가를 포함한다. 이와 관련하여, 일반적으로 전체 단백질 기능을 비활성화시키지 않는 것으로 인식되는 특정 보존적 아미노산 치환이 이루어질 수 있다: 예컨대 양으로 하전된 아미노산(및 이의 역)에 관하여, 라이신, 아르기닌 및 히스티딘; 음으로 하전된 아미노산(및 이의 역)에 관하여, 아스파르트산 및 글루탐산; 및 특정 군의 중성으로 하전된 아미노산(및 모든 경우에, 또한 이의 역)에 관하여, (1) 알라닌 및 세린, (2) 아스파라긴, 글루타민 및 히스티딘, (3) 시스테인 및 세린, (4) 글리신 및 프롤린, (5) 이소류신, 류신 및 발린, (6) 메티오닌, 류신 및 이소류신, (7) 페닐알라닌, 메티오닌, 류신 및 티로신, (8) 세린 및 트레오닌, (9) 트립토판 및 티로신, (10) 및 예를 들어, 티로신, 트립토판 및 페닐알라닌. 아미노산은 물리적 특성 및 이차 및 삼차 단백질 구조에 대한 기여에 따라 분류될 수 있다. 보존적 치환은 하나의 아미노산의 유사한 특성을 갖는 또 다른 아미노산으로의 치환으로서 당업계에 인식된다.

[0918] 동일한 또는 특정 백분율의 동일한 뉴클레오티드 또는 아미노산 잔기를 갖는 2개 이상의 핵산 또는 폴리펩티드

서열의 동일성(예를 들어, 비교 윈도우 또는 표기된 영역에 걸쳐 최대 상응성을 위해 비교되고 정렬되는 경우, 특정 영역에 걸쳐 약 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 이상의 동일성)은 하기 기재된 디폴트 파라미터와 함께 BLAST 또는 BLAST 2.0 서열 비교 알고리즘을 사용하여 또는 수동의 정렬 및 육안의 검사(예를 들어, NCBI 웹 사이트 www.ncbi.nlm.nih.gov/BLAST/ 등 참조)에 의해 측정될 수 있다. 동일성은 또한 시험 서열의 상보물을 지칭할 수 있거나 이에 적용될 수 있다. 동일성은 또한 결실 및/또는 부가를 갖는 서열 및 치환을 갖는 서열을 포함한다. 본 명세서에 기재된 바와 같이, 알고리즘은 갭 등을 설명한다. 동일성은 적어도 약 10개 아미노산 또는 뉴클레오티드 길이, 약 15개 아미노산 또는 뉴클레오티드 길이, 약 20개 아미노산 또는 뉴클레오티드 길이, 약 25개 아미노산 또는 뉴클레오티드 길이, 약 30개 아미노산 또는 뉴클레오티드 길이, 약 35개 아미노산 또는 뉴클레오티드 길이, 약 40개 아미노산 또는 뉴클레오티드 길이, 약 45개 아미노산 또는 뉴클레오티드 길이, 약 50개 아미노산 또는 뉴클레오티드 길이 이상인 영역에 걸쳐 존재할 수 있다. 유전 암호가 축퇴성이기 때문에, 상동성 뉴클레오티드 서열은 임의의 수의 침묵 염기 변경, 즉, 그럼에도 불구하고 동일한 아미노산을 인코딩하는 뉴클레오티드 치환을 포함할 수 있다.

[0919] *단백질성 외부*

[0920] 일부 구현예에서, 아넬로벡터, 예를 들어, 합성 아넬로벡터는 유전 요소를 봉입하는 단백질성 외부를 포함한다. 단백질성 외부는 포유동물에서 원치 않는 면역 반응을 유도하지 못하는 실질적으로 비-병원성인 외부 단백질을 포함할 수 있다. 아넬로벡터의 단백질성 외부는 통상적으로 단백질성 외부를 구성하는 정20면체 형성으로 자가-조립될 수 있는 실질적으로 비-병원성인 단백질을 포함한다.

[0921] 일부 구현예에서, 단백질성 외부 단백질은 아넬로벡터의 유전 요소의 서열에 의해 인코딩된다(예를 들어, 유전 요소와 시스로 존재함). 다른 구현예에서, 단백질성 외부 단백질은 아넬로벡터의 유전 요소와 별개인 핵산에 의해 인코딩된다(예를 들어, 유전 요소와 트랜스로 존재함).

[0922] 일부 구현예에서, 단백질, 예를 들어, 실질적으로 비-병원성인 단백질 및/또는 단백질성 외부 단백질은 하나 이상의 글리코실화된 아미노산, 예를 들어, 2, 3, 4, 5, 6, 7, 8, 9, 10개 이상을 포함한다.

[0923] 일부 구현예에서, 단백질, 예를 들어, 실질적으로 비-병원성인 단백질 및/또는 단백질성 외부 단백질은 적어도 하나의 친수성 DNA-결합 영역, 아르기닌-풍부 영역, 트레오닌-풍부 영역, 글루타민-풍부 영역, N-말단 폴리아르기닌 서열, 가변 영역, C-말단 폴리글루타민/글루탐산염 서열, 및 하나 이상의 이황화 가교를 포함한다.

[0924] 일부 구현예에서, 단백질은 캡시드 단백질이며, 예를 들어, 본 명세서에 기재된 캡시드 단백질을 인코딩하는 뉴클레오티드 서열, 예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어, 아넬로바이러스 ORF1 분자 및/또는 캡시드 단백질 서열 중 어느 하나에 의해 인코딩되는 단백질에 대하여 적어도 약 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 서열을 갖는다. 일부 구현예에서, 단백질 또는 캡시드 단백질의 기능적 단편은, 예를 들어, 본 명세서에 기재된 바와 같은, 아넬로바이러스 ORF1 핵산에 대하여 적어도 약 60%, 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 뉴클레오티드 서열에 의해 인코딩된다.

[0925] 일부 구현예에서, 아넬로벡터는 캡시드 단백질 또는 캡시드 단백질의 기능적 단편을 인코딩하는 뉴클레오티드 서열, 또는 본 명세서에 기재된 바와 같은 아넬로바이러스 ORF1 분자에 대하여 적어도 약 60%, 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 갖는 서열을 포함한다.

[0926] 일부 구현예에서, 더 낮은 서열 동일성을 갖는 아미노산의 범위는 본 명세서에 기재된 특성 중 하나 이상 및 세포/조직/종 특이성(예를 들어, 향성)의 차이를 제공할 수 있다.

[0927] 일부 구현예에서, 아넬로벡터는 단백질성 외부 내에 지질이 결여된다. 일부 구현예에서, 아넬로벡터에는 지질 이중층, 예를 들어, 바이러스 외피가 결여된다. 일부 구현예에서, 아넬로벡터의 내부는 단백질성 외부에 의해 완전히 덮인다(예를 들어, 100% 커버리지). 일부 구현예에서, 아넬로벡터의 내부는 단백질성 외부에 의해 100% 미만, 예를 들어, 95%, 90%, 85%, 80%, 70%, 60%, 50% 이하의 커버리지로 덮인다. 일부 구현예에서, 단백질성 외부는 유전 요소가 아넬로벡터에 보유되는 한, 예를 들어, 물, 이온, 펩티드 또는 소분자에 대한 투과 가능성을 허용하는 갭 또는 불연속부를 포함한다.

[0928] 일부 구현예에서, 단백질성 외부는 숙주 세포 내로의 유전 요소의 유입을 매개하기 위하여 숙주 세포를 특이적으로 인식하고/인식하거나 이에 결합하는 하나 이상의 단백질 또는 폴리펩티드, 예를 들어, 상보성 단백질 또는 폴리펩티드를 포함한다.

[0929] 일부 구현예에서, 단백질성 외부는 하기 중 하나 이상을 포함한다: 예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어, ORF1 분자의 아르기닌-풍부 영역, 젤리-롤 영역, N22 도메인, 초가변 영역, 및/또는 C-말단 도메인. 일부 구현예에서, 단백질성 외부는 하기 중 하나 이상을 포함한다: 하나 이상의 글리코실화된 단백질, 친수성 DNA-결합 영역, 아르기닌-풍부 영역, 트레오닌-풍부 영역, 글루타민-풍부 영역, N-말단 폴리아르기닌 서열, 가변 영역, C-말단 폴리글루타민/글루탐산염 서열, 및 하나 이상의 이황화 가교. 예를 들어, 단백질성 외부는, 예를 들어, 본 명세서에 기재된 바와 같은, 아넬로바이러스 ORF1 핵산에 의해 인코딩된 단백질을 포함한다.

[0930] 일부 구현예에서, 단백질성 외부는 하기의 특성 중 하나 이상을 포함한다: 정20면체 대칭, 하나 이상의 숙주 세포 분자와 상호작용하는 분자를 인식하고/하거나 이에 결합하여, 숙주 세포로의 유입을 매개함, 지질 분자가 결여됨, 탄수화물이 결여됨, pH 및 온도 안정성임, 세제 저항성임, 및 숙주에서 실질적으로 비-면역원성이거나 비-병원성임.

[0931] **III. 핵산 작제물**

[0932] 본 명세서에 기재된 유전 요소는 핵산 작제물(예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어, 탠덤 작제물)에 포함될 수 있다.

[0933] 일 양태에서, 본 발명은 (i) 비-병원성 외부 단백질(예를 들어, 아넬로바이러스 ORF1 분자 또는 이의 스플라이스 변이체 또는 기능적 단편)을 인코딩하는 서열, (ii) 유전 요소를 비-병원성 외부 단백질에 결합시키는 외부 단백질 결합 서열, 및 (iii) 이펙터를 인코딩하는 서열을 포함하는 유전 요소를 포함하는 핵산 유전 요소 작제물(예를 들어, 탠덤 작제물)을 포함한다. 일부 구현예에서, 유전 요소 작제물은 제2 카피의 유전 요소, 또는 이의 단편(예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어, uRFS 또는 dRFS를 포함함)을 추가로 포함한다.

[0934] 유전 요소 또는 유전 요소 내의 임의의 서열은 임의의 적합한 방법을 사용하여 수득될 수 있다. 다양한 재조합 방법, 예를 들어, 바이러스 서열을 갖는 세포로부터 라이브러리의 스크리닝, 이를 포함하는 것으로 알려져 있는 핵산 작제물로부터 서열의 유도, 또는 표준 기법을 사용한 이를 함유하는 세포 및 조직으로부터의 직접적인 단리가 당업계에 알려져 있다. 대안적으로 또는 조합하여, 유전 요소의 일부 또는 전부는 클로닝보다는 합성에 의해 생성될 수 있다.

[0935] 일부 구현예에서, 핵산 작제물은 조절 요소, 표적 유전자에 상동성인 핵산 서열, 및 생존 가능한 세포 내에서 및/또는 세포내 분자가 표적 세포 내에 존재하는 경우 리포터 분자의 발현을 야기하기 위한 다양한 리포터 작제물을 포함한다.

[0936] 잠재적으로 형질감염된 세포를 확인하고, 조절 서열의 기능을 평가하기 위한 리포터 유전자가 사용된다. 일반적으로, 리포터 유전자는 수여자 유기체 또는 조직에 존재하거나 그에 의해 발현되지 않고, 발현이 다소 용이하게 검출 가능한 특성, 예를 들어, 효소 활성에 의해 나타나는 폴리펩티드를 인코딩하는 유전자이다. 리포터 유전자의 발현은 DNA가 수여자 세포 내로 도입된 후 적합한 시간에 검정된다. 적합한 리포터 유전자는 루시퍼라제, 베타-갈락토시다제, 클로람페니콜 아세틸 트랜스퍼라제, 분비 알칼리성 포스파타제 또는 녹색 형광 단백질 유전자를 인코딩하는 유전자를 포함할 수 있다(예를 들어, 문헌[Ui-Tei et al., 2000 FEBS Letters 479: 79-82]). 적합한 발현 시스템은 널리 알려져 있으며, 알려진 기법을 사용하여 제조되거나 상업적으로 수득될 수 있다. 일반적으로, 가장 높은 리포터 유전자의 발현 수준을 보이는 최소 5' 측접 영역을 갖는 작제물은 프로모터로서 확인된다. 이러한 프로모터 영역은 리포터 유전자에 연결되고, 프로모터-유도 전사를 조절하는 능력에 대하여 작용제를 평가하기 위해 사용될 수 있다.

[0937] 일부 구현예에서, 핵산 작제물은 실질적으로 비-병원성이고/이거나 숙주 세포 내에 실질적으로 비-통합되거나, 숙주에서 실질적으로 비-면역원성이다.

[0938] 일부 구현예에서, 핵산 작제물은 표현형, 바이러스 수준, 유전자 발현, 다른 바이러스와의 경쟁, 병태 등 중 하나 이상을 적어도 약 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50% 이상 조절하기에 충분한 양으로 존재한다.

[0939] **IV. 조성물**

[0940] 본 명세서에 기재된 아넬로벡터는 또한 예를 들어, 본 명세서에 기재된 바와 같이, 약제학적 부형제와 함께 약제학적 조성물에 포함될 수 있다. 일부 구현예에서, 약제학적 조성물은 적어도 10^5, 10^6, 10^7, 10^8, 10^9, 10^{10}, 10^{11}, 10^{12}, 10^{13}, 10^{14} 또는 10^{15}개의 아넬로벡터를 포함한다. 일부 구현예에서, 약제학적 조성물은 약 10^5 내지

10^{15}, 10^{5} 내지 10^{10} 또는 10^{10} 내지 10^{15}개의 아넬로벡터를 포함한다. 일부 구현예에서, 약제학적 조성물은 약 10^{8}개(예를 들어, 약 10^{5}, 10^{6}, 10^{7}, 10^{8}, 10^{9} 또는 10^{10}개) 게놈 당량/㎖의 아넬로벡터를 포함한다. 일부 구현예에서, 약제학적 조성물은 예를 들어, PCT/US19/65995의 실시예 18의 방법에 따라 결정시, 10^{5} 내지 10^{10}, 10^{6} 내지 10^{10}, 10^{7} 내지 10^{10}, 10^{8} 내지 10^{10}, 10^{9} 내지 10^{10}, 10^{5} 내지 10^{6}, 10^{5} 내지 10^{7}, 10^{5} 내지 10^{8}, 10^{5} 내지 10^{9}, 10^{5} 내지 10^{11}, 10^{5} 내지 10^{12}, 10^{5} 내지 10^{13}, 10^{5} 내지 10^{14}, 10^{5} 내지 10^{15} 또는 10^{10} 내지 10^{15}개 게놈 당량/㎖의 아넬로벡터를 포함한다. 일부 구현예에서, 약제학적 조성물은 세포당 아넬로벡터에 포함되는 적어도 1, 2, 5 또는 10, 100, 500, 1000, 2000, 5000, 8,000, 1×10^{4}, 1×10^{5}, 1×10^{6}, 1×10^{7}개 이상의 카피의 유전 요소를 진핵 세포의 집단에 운반하기에 충분한 아넬로벡터를 포함한다. 일부 구현예에서, 약제학적 조성물은 세포당 아넬로벡터에 포함되는 적어도 약 1×10^{4}, 1×10^{5}, 1×10^{6}, $1 \times$ 또는 10^{7} 또는 약 1×10^{4} 내지 1×10^{5}, 1×10^{4} 내지 1×10^{6}, 1×10^{4} 내지 1×10^{7}, 1×10^{5} 내지 1×10^{6}, 1×10^{5} 내지 1×10^{7} 또는 1×10^{6} 내지 1×10^{7}개 카피의 유전 요소를 진핵 세포의 집단에 운반하기에 충분한 아넬로벡터를 포함한다.

[0941] 일부 구현예에서, 약제학적 조성물은 하기의 특징 중 하나 이상을 갖는다: 약제학적 조성물이 약제학적 또는 우수 제조 기준(GMP) 표준을 충족하거나; 약제학적 조성물을 우수 제조 기준(GMP)에 따라 제조하거나; 약제학적 조성물이 사전결정된 기준 값 미만의 병원체 수준을 갖거나, 예를 들어, 실질적으로 병원체가 없거나; 약제학적 조성물이 사전결정된 기준 값 미만의 오염물질 수준을 갖거나, 예를 들어, 실질적으로 오염물질이 없거나; 또는 약제학적 조성물이 낮은 면역원성을 갖거나, 예를 들어, 본 명세서에 기재된 바와 같이 실질적으로 비-면역원성임.

[0942] 일부 구현예에서, 약제학적 조성물은 임계값 미만의 양의 하나 이상의 오염물질을 포함한다. 약제학적 조성물에서 바람직하게는 배제되거나 최소화되는 예시적인 오염물질은 제한 없이, 숙주 세포 핵산(예를 들어, 숙주 세포 DNA 및/또는 숙주 세포 RNA), 동물-유래 성분(예를 들어, 혈청 알부민 또는 트립신), 복제-적격 바이러스, 비-감염성 입자, 유리 바이러스 캡시드 단백질, 외래 작용제 및 응집물을 포함한다. 구현예에서, 오염물질은 숙주 세포 DNA이다. 구현예에서, 조성물은 용량당 약 10 ng 미만의 숙주 세포 DNA를 포함한다. 구현예에서, 조성물 중 숙주 세포 DNA의 수준은 여과 및/또는 숙주 세포 DNA의 효소적 분해에 의해 감소된다. 구현예에서, 약제학적 조성물은 중량 기준 10% 미만(예를 들어, 약 10%, 5%, 4%, 3%, 2%, 1%, 0.5% 또는 0.1% 미만)의 오염물질로 이루어진다.

[0943] 일 양태에서, 본 명세서에 기재된 본 발명은 하기를 포함하는 약제학적 조성물을 포함한다:

[0944] a) (i) 비-병원성 외부 단백질을 인코딩하는 서열, (ii) 유전 요소를 비-병원성 외부 단백질에 결합하는 외부 단백질 결합 서열, 및 (iii) 조절 핵산을 인코딩하는 서열을 포함하는 유전 요소; 및 유전 요소와 회합된, 예를 들어, 이를 봉입하거나 둘러싸는 단백질성 외부를 포함하는 아넬로벡터; 및

[0945] b) 약제학적 부형제.

[0946] *소낭*

[0947] 일부 구현예에서, 조성물은 운반체 성분, 예를 들어, 마이크로입자, 리포좀, 소낭 또는 엑소좀을 추가로 포함한다. 일부 구현예에서, 리포좀은 내부 수성 구획을 둘러싸는 단일- 또는 다중라멜라 지질 이중층 및 상대적으로 불투과성인 외부 친지성 인지질 이중층으로 구성된 구형 소낭 구조를 포함한다. 리포좀은 음이온성, 중성 또는 양이온성일 수 있다. 리포좀은 일반적으로 생체적합성, 비독성이며, 친수성 및 친지성 약물 분자 둘 모두를 운반하고, 혈장 효소에 의한 분해로부터 그들의 카고(cargo)를 보호하고, 생물학적 막을 가로질러 그들의 로드를 수송할 수 있다(예를 들어, 검토를 위하여 문헌[Spuch and Navarro, Journal of Drug Delivery, vol. 2011, Article ID 469679, 12 pages, 2011. doi:10.1155/2011/469679] 참조).

[0948] 소낭은 몇몇 상이한 유형의 지질로부터 제조될 수 있지만; 약물 담체로서 리포좀을 생성하는 데 인지질이 가장 흔히 사용된다. 소낭은 제한 없이, DOTMA, DOTAP, DOTIM, DDAB를 단독으로 또는 콜레스테롤과 함께 포함하여, DOTMA 및 콜레스테롤, DOTAP 및 콜레스테롤, DOTIM 및 콜레스테롤, 및 DDAB 및 콜레스테롤을 제공할 수 있다. 다중라멜라 소낭 지질의 제조 방법은 당업계에 알려져 있다(예를 들어, 다중라멜라 소낭 지질 제제에 관한 교시가 본 명세서에 참조로 포함되는 미국 특허 6,693,086 참조). 지질 필름이 수용액과 혼합되는 경우 소낭 형성은 자발적일 수 있지만, 이는 또한 균질화기, 초음파처리기, 또는 압출 장치를 사용함으로써 진탕 형태로 힘을 가

함으로써 더 신속히 처리될 수 있다(예를 들어, 검토를 위해 문헌[Spuch and Navarro, Journal of Drug Delivery, vol. 2011, Article ID 469679, 12 pages, 2011. doi:10.1155/2011/469679] 참조). 압출된 지질은 압출된 지질 제제에 관한 교시가 본 명세서에 참조로 포함되는 문헌[Templeton et al., Nature Biotech, 15:647-652, 1997]에 기재된 바와 같이 감소하는 크기의 필터를 통해 압출시킴으로써 제조될 수 있다.

[0949] 본 명세서에 기재된 바와 같이, 첨가제를 소낭에 첨가하여, 그들의 구조 및/또는 특성을 변형시킬 수 있다. 예를 들어, 콜레스테롤 또는 스핑고미엘린을 혼합물에 첨가하여, 구조의 안정화 및 내부 카고의 누출의 방지를 도울 수 있다. 추가로, 소낭은 수소화된 난(egg) 포스파티딜콜린 또는 난 포스파티딜콜린, 콜레스테롤 및 디세틸 포스페이트로부터 제조될 수 있다(예를 들어, 검토를 위하여 문헌[Spuch and Navarro, Journal of Drug Delivery, vol. 2011, Article ID 469679, 12 pages, 2011. doi:10.1155/2011/469679] 참조). 또한, 소낭은 합성 동안 또는 그 후에 수여자 세포 상의 반응성 기에 상보성인 반응성 기를 포함하도록 표면 변형될 수 있다. 이러한 반응성 기는 제한 없이, 말레이미드 기를 포함한다. 일 예로서, 소낭은 말레이미드 컨쥬게이트된 인지질, 예컨대 제한 없이 DSPE-MaL-PEG2000을 포함하도록 합성될 수 있다.

[0950] 소낭 형성은 주로 천연 인지질 및 지질, 예컨대 1,2-다이스테아로릴-sn-글리세로-3-포스파티딜 콜린(DSPC), 스핑고미엘린, 난 포스파티딜콜린 및 모노시알로강글리오시드로 구성될 수 있다. 인지질만으로 제조된 제형은 혈장에서 덜 안정하다. 그러나, 콜레스테롤을 사용한 지질 막의 조작은 캡슐화된 카고의 신속한 방출을 감소시키거나, 1,2-다이올레오일-sn-글리세로-3-포스포에탄올아민(DOPE)은 안정성을 증가시킨다(예를 들어, 검토를 위하여 문헌[Spuch and Navarro, Journal of Drug Delivery, vol. 2011, Article ID 469679, 12 pages, 2011. doi:10.1155/2011/469679] 참조).

[0951] 구현예에서, 지질은 지질 마이크로입자를 형성하기 위해 사용될 수 있다. 지질은 DLin-KC2-DMA4, C12-200 및 공지질 디스테로일포스파티딜 콜린을 포함하지만 이들로 제한되지 않으며, 콜레스테롤 및 PEG-DMG는 자발적 소낭 형성 절차를 사용하여 제형화될 수 있다(예를 들어, 문헌[Novobrantseva, Molecular Therapy-Nucleic Acids (2012) 1, e4; doi:10.1038/mtna.2011.3] 참조). 성분 몰비는 약 50/10/38.5/1.5(DLin-KC2-DMA 또는 C12-200/디스테로일포스파티딜 콜린/콜레스테롤/PEG-DMG)일 수 있다. 테크미라(Tekmira)는 모두가 본 발명에 사용될 수 있고/있거나 조정될 수 있는 지질 마이크로입자 및 지질 마이크로입자 제형의 다양한 양태에 관한, 미국 및 해외에서의 대략 95개의 특허 패밀리의 포트폴리오를 갖는다(예를 들어, 미국 특허 7,982,027; 7,799,565; 8,058,069; 8,283,333; 7,901,708; 7,745,651; 7,803,397; 8,101,741; 8,188,263; 7,915,399; 8,236,943 및 7,838,658 및 유럽 특허 1766035; 1519714; 1781593 및 1664316 참조).

[0952] 일부 구현예에서, 마이크로입자는 무작위 방식으로 배열된 하나 이상의 고형화된 폴리머(들)를 포함한다. 마이크로입자는 생분해성일 수 있다. 생분해성 마이크로입자는 예를 들어, 제한 없이, 용매 증발, 핫 멜트 마이크로캡슐화, 용매 제거 및 분무 건조를 포함하는 당업계에 알려져 있는 방법을 사용하여 합성될 수 있다. 마이크로입자를 합성하기 위한 예시적인 방법은 문헌[Bershteyn et al., Soft Matter 4:1787-1787, 2008] 및 US 2008/0014144 A1에 기재되어 있으며, 마이크로입자 합성에 관한 이의 구체적인 교시는 본 명세서에 참조로 포함된다.

[0953] 생분해성 마이크로입자를 형성하기 위해 사용될 수 있는 예시적인 합성 폴리머는 제한 없이, 지방족 폴리에스테르, 폴리 (락트산)(PLA), 폴리 (글리콜산)(PGA), 락트산 및 글리콜산의 코폴리머(PLGA), 폴리카프로락톤(PCL), 다가무수물(polyanhydride), 폴리(오르토)에스테르, 폴리우레탄, 폴리(부티르산), 폴리(발레르산) 및 폴리(락티드-코-카프로락톤) 및 천연 폴리머, 예컨대 알부민, 알긴산염 및 덱스트란 및 셀룰로스를 포함하는 기타 다당류, 콜라겐, 예컨대, 예를 들어, 알킬, 알킬렌과 같은 화학 기의 치환, 부가, 하이드록실화, 산화 및 당업자에 의해 일상적으로 이루어지는 기타 변형을 포함하는 이의 화학적 유도체, 알부민 및 기타 친수성 단백질, 젠인 및 기타 프롤라민 및 소수성 단백질, 이의 코폴리머 및 혼합물을 포함한다. 일반적으로, 이들 물질은 효소적 가수분해 또는 물로의 노출에 의해, 표면 또는 벌크 침식에 의해 분해된다.

[0954] 마이크로입자의 직경은 0.1 내지 1000 마이크로미터(㎛) 범위이다. 일부 구현예에서, 그들의 직경은 크기가 1 내지 750 ㎛, 또는 50 내지 500 ㎛, 또는 100 내지 250 ㎛의 범위이다. 일부 구현예에서, 그들의 직경은 크기가 50 내지 1000 ㎛, 50 내지 750 ㎛, 50 내지 500 ㎛, 또는 50 내지 250 ㎛의 범위이다. 일부 구현예에서, 그들의 직경은 크기가 .05 내지 1000 ㎛, 10 내지 1000 ㎛, 100 내지 1000 ㎛, 또는 500 내지 1000 ㎛의 범위이다. 일부 구현예에서, 그들의 직경은 약 0.5 ㎛, 약 10 ㎛, 약 50 ㎛, 약 100 ㎛, 약 200 ㎛, 약 300 ㎛, 약 350 ㎛, 약 400 ㎛, 약 450 ㎛, 약 500 ㎛, 약 550 ㎛, 약 600 ㎛, 약 650 ㎛, 약 700 ㎛, 약 750 ㎛, 약 800 ㎛, 약 850 ㎛, 약 900 ㎛, 약 950 ㎛ 또는 약 1000 ㎛이다. 마이크로입자 직경의 맥락에서 사용되는 바와 같이, 용

어 "약"은 언급된 절대값의 +/-5%를 의미한다.

[0955] 일부 구현예에서, 리간드는 입자의 표면 상에 존재하고, 부착될 리간드 상에 존재하는 작용성 화학기(카복실산, 알데히드, 아민, 술프하이드릴 및 하이드록실)를 통해 마이크로입자의 표면에 컨쥬게이트된다. 작용성은 예를 들어, 마이크로입자의 에멀젼 제조 동안 작용성 화학 기를 갖는 안정화제의 혼입에 의해 마이크로입자 내로 도입될 수 있다.

[0956] 작용성 기를 마이크로입자에 도입하는 또 다른 예는 입자 제조 후 동안, 입자 및 리간드를 동종- 또는 이종 이작용성 가교결합제와 직접 가교결합시킴으로써 이루어진다. 이러한 절차는 적합한 화학물질 및 소정의 부류의 가교결합제(하기에 더욱 상세히 논의된 바와 같은 CDI, EDAC, 글루타르알데히드 등) 또는 제조 후에 입자 표면의 화학적 변형을 통해 리간드를 입자 표면에 커플링시키는 임의의 다른 가교결합제를 사용할 수 있다. 이는 또한 양친매성 분자, 예컨대 지방산, 지질 또는 기능적 안정화제가 입자 표면에 수동으로 흡수되고 부착되어, 그에 의해 리간드로의 테더링을 위한 작용성 말단 기를 도입할 수 있는 과정을 포함한다.

[0957] 일부 구현예에서, 마이크로입자는 그들의 외부 표면 상에 하나 이상의 표적화 기를 포함하도록 합성하여, 특정 세포 또는 조직 유형(예를 들어, 심근세포)을 표적화할 수 있다. 이들 표적화 기는 제한 없이, 수용체, 리간드, 항체 등을 포함한다. 이들 표적화 기는 세포의 표면 상의 그들의 파트너와 결합한다. 일부 구현예에서, 마이크로입자는 세포 표면을 포함하는 지질 이중층 내로 통합될 것이며, 미토콘드리아는 세포로 운반된다.

[0958] 마이크로입자는 또한, 그들의 최외각 표면 상에 지질 이중층을 포함할 수 있다. 이러한 이중층은 동일하거나 상이한 유형의 하나 이상의 지질로 구성될 수 있다. 예는 인지질, 예컨대 포스포콜린 및 포스포이노시톨을 포함하지만, 이에 제한되지 않는다. 구체적인 예는 제한 없이, DMPC, DOPC, DSPC 및 다양한 다른 지질, 예컨대 리포좀에 대하여 본 명세서에 기재된 것들을 포함한다.

[0959] 일부 구현예에서, 담체는 예를 들어, 본 명세서에 기재된 바와 같은 나노입자를 포함한다.

[0960] 일부 구현예에서, 본 명세서에 기재된 소낭 또는 마이크로입자는 진단제로 작용화된다. 진단제의 예는 양전자 방사 단층촬영(PET), 컴퓨터 보조 단층촬영(CAT), 단일 광자 방출 컴퓨터 단층촬영, x-선, 형광투시법 및 자기 공명 영상화(MRI)에 사용되는 구매할 수 있는 영상화제; 및 조영제를 포함하지만, 이에 제한되지 않는다. MRI에서 조영제로서 사용하기에 적합한 물질의 예는 가돌리늄 킬레이트, 및 철, 마그네슘, 망간, 구리 및 크롬을 포함한다.

[0961] *담체*

[0962] 본 명세서에 기재된 조성물(예를 들어, 약제학적 조성물)은 담체를 포함하고/하거나, 이와 함께 제형화되고/되거나, 이 내에서 운반될 수 있다. 일 양태에서, 본 발명은 본 명세서에 기재된 조성물(예를 들어, 본 명세서에 기재된 아넬로벡터, 아넬로바이러스, 또는 유전 요소)을 포함하는(예를 들어, 캡슐화하는) 담체(예를 들어, 소낭, 리포좀, 지질 나노입자, 엑소좀, 적혈구, 엑소좀(예를 들어, 포유동물 또는 식물 엑소좀), 푸소좀)를 포함하는 조성물, 예를 들어, 약제학적 조성물을 포함한다.

[0963] 일부 구현예에서, 본 명세서에 기재된 조성물 및 시스템은 리포좀 또는 다른 유사한 소낭 내에 제형화될 수 있다. 일반적으로, 리포좀은 내부 수성 구획을 둘러싸는 단일- 또는 다중라멜라 지질 이중층 및 상대적으로 불투과성인 외부 친지성 인지질 이중층으로 구성된 구형 소낭 구조이다. 리포좀은 음이온성, 중성 또는 양이온성일 수 있다. 리포좀은 일반적으로 하기의 특징 중 하나 이상(예를 들어, 전부)을 갖는다: 생체적합성, 비독성, 친수성 및 친지성 약물 분자 둘 모두를 운반할 수 있음, 혈장 효소에 의한 분해로부터 그들의 카고(cargo)를 보호할 수 있음 및 생물학적 막 및 혈액 뇌 장벽(BBB)을 가로질러 그들의 로드를 수송할 수 있음(예를 들어, 문헌[Spuch and Navarro, Journal of Drug Delivery, vol. 2011, Article ID 469679, 12 pages, 2011. doi:10.1155/2011/469679]; 및 문헌[Zylberberg & Matosevic. 2016. Drug Delivery, 23:9, 3319-3329, doi: 10.1080/10717544.2016.1177136] 참조).

[0964] 소낭은 몇몇 상이한 유형의 지질로부터 제조될 수 있지만; 약물 담체로서 리포좀을 생성하는 데 인지질이 가장 흔히 사용된다. 다중라멜라 소낭 지질의 제조 방법은 공지되어 있다(예를 들어, 미국 특허 6,693,086을 참조하며, 다중라멜라 소낭 지질 제제에 관한 이의 교시는 본 명세서에 참조로 포함됨). 소낭 형성은 지질막을 수성 용액과 혼합하는 경우에 자발적일 수 있지만, 그것은 균질화기, 초음파기 또는 압출 장치를 사용함으로써 진탕 형태의 힘을 인가함으로써 촉진될 수도 있다(예를 들어, 검토를 위하여 문헌[Spuch and Navarro, Journal of Drug Delivery, vol. 2011, Article ID 469679, 12 pages, 2011. doi:10.1155/2011/469679] 참조). 압출된 지질은 예를 들어, 문헌[Templeton et al., Nature Biotech, 15:647-652, 1997]에 기재된 바와 같이, 감소하는

크기의 필터를 통한 압출에 의해 제조될 수 있다.

[0965] 지질 나노입자(LNP)는 본 명세서에 기재된 약제학적 조성물을 위한 생체적합성 및 생분해성 운반 시스템을 제공하는 담체의 또 다른 예이다. 예를 들어, 문헌[Gordillo-Galeano et al. European Journal of Pharmaceutics and Biopharmaceutics. Volume 133, December 2018, Pages 285-308]을 참조한다. 나노구조화된 지질 담체(NLC)는 변형된 고체 지질 나노입자(SLN)의 특징을 보유하고, 약물 안정성 및 로딩 능력을 개선시키고, 약물 누출을 방지하는 변형된 고체 지질 나노입자(SLN)이다. 폴리머 나노입자(PNP)는 약물 운반의 중요한 구성성분이다. 이들 나노입자는 특정 표적으로의 약물 운반을 효율적으로 유도하고, 약물 안정성 및 제어된 약물 방출을 개선시킬 수 있다. 리포좀과 폴리머를 조합한 새로운 유형의 담체인, 지질-중합체 나노입자(PLN)도 또한 사용될 수 있다. 이들 나노입자는 PNP와 리포좀의 보완적인 이점을 보유한다. PLN는 코어(core)-쉘(shell) 구조로 구성된다; 폴리머 코어는 안정한 구조를 제공하며, 인지질 쉘(shell)은 우수한 생체적합성을 제공한다. 이와 같이, 2가지 구성성분은 약물 캡슐화 효율을 증가시키고, 표면 변형을 용이하게 하고, 수용성 약물의 누출을 방지한다. 검토를 위하여, 예를 들어, 문헌[Li et al. 2017, Nanomaterials 7, 122; doi:10.3390/nano7060122]을 참조한다.

[0966] 엑소좀도 또한, 본 명세서에 기재된 조성물 및 시스템을 위한 약물 운반 비히클로서 사용될 수 있다. 검토를 위하여, 문헌[Ha et al. July 2016. Acta Pharmaceutica Sinica B. Volume 6, Issue 4, Pages 287-296; doi.org/10.1016/j.apsb.2016.02.001]을 참조한다.

[0967] 생체 외 분화된 적혈구도 또한, 본 명세서에 기재된 조성물을 위한 담체로서 사용될 수 있다. 예를 들어, WO2015073587; WO2017123646; WO2017123644; WO2018102740; WO2016183482; WO2015153102; WO2018151829; WO2018009838; 문헌[Shi et al. 2014. Proc Natl Acad Sci USA. 111(28): 10131-10136]; 미국 특허 9,644,180; 문헌[Huang et al. 2017. Nature Communications 8: 423]; 문헌[Shi et al. 2014. Proc Natl Acad Sci USA. 111(28): 10131-10136]을 참조한다.

[0968] 예를 들어, WO2018208728에 기재된 바와 같은 푸소좀 조성물도 또한 본 명세서에 기재된 조성물을 운반하기 위한 담체로서 사용될 수 있다.

[0969] *막 투과성 폴리펩티드*

[0970] 일부 구현예에서, 조성물은 세포 내로 또는 막, 예를 들어, 세포 또는 핵 막을 가로질러 성분을 운반하도록 막 투과성 폴리펩티드(MPP)를 추가로 포함한다. 막을 가로질러 물질의 수송을 용이하게 할 수 있는 막 투과성 폴리펩티드는 세포-투과 펩티드(CPP)(예를 들어, 미국 특허 제8,603,966호 참조), 식물 세포내 운반을 위한 융합 펩티드(예를 들어, 문헌[Ng et al., PLoS One, 2016, 11:e0154081] 참조), 단백질 변환 도메인, 트로얀(Trojan) 펩티드, 및 막 전위 신호(MTS)(예를 들어, 문헌[Tung et al., Advanced Drug Delivery Reviews 55:281-294 (2003)] 참조)를 포함하지만, 이에 제한되지 않는다. 일부 MPP는 양으로 하전된 측쇄를 갖는 아미노산, 예컨대 아르기닌이 풍부하다.

[0971] 막 투과성 폴리펩티드는 전신 투여 시에 생체 내에서 다중의 조직의 세포 내에서, 성분의 막 투과를 유도하고, 거대분자 전위를 가능하게 하는 능력을 갖는다. 막 투과성 폴리펩티드는 또한, 적절한 조건 하에서 세포와 접촉되는 경우, 외부 환경으로부터 세포질, 세포소기관, 예컨대 미토콘드리아 또는 세포의 핵을 포함하는 세포내 환경에, 수동 확산으로 달성될 것보다 유의미하게 더 큰 양으로 통과하는 펩티드를 지칭할 수 있다.

[0972] 막을 가로질러 수송되는 성분은 막 투과성 폴리펩티드와 가역적으로 또는 비가역적으로 연결될 수 있다. 링커는 화학적 결합, 예를 들어, 하나 이상의 공유 결합 또는 비-공유 결합일 수 있다. 일부 구현예에서, 링커는 펩티드 링커이다. 이러한 링커는 2 내지 30개 아미노산 이상일 수 있다. 링커는 유연성, 강성 또는 절단 가능한 링커를 포함한다.

[0973] *조합*

[0974] 일 양태에서, 본 명세서에 기재된 아넬로벡터 또는 아넬로벡터를 포함하는 조성물은 또한, 하나 이상의 이종성 모이어티를 포함할 수 있다. 일 양태에서, 본 명세서에 기재된 아넬로벡터 또는 아넬로벡터를 포함하는 조성물은 또한, 융합물 내에 하나 이상의 이종성 모이어티를 포함할 수 있다. 일부 구현예에서, 이종성 모이어티는 유전 요소와 연결될 수 있다. 일부 구현예에서, 이종성 모이어티는 아넬로벡터의 일부로서 단백질성 외부에 봉입될 수 있다. 일부 구현예에서, 이종성 모이어티는 아넬로벡터와 함께 투여될 수 있다.

[0975] 일 양태에서, 본 발명은 본 명세서에 기재된 아넬로벡터 및 이종성 모이어티 중 어느 하나를 포함하는 세포 또

는 조직을 포함한다.

[0976] 또 다른 양태에서, 본 발명은 본 명세서에 기재된 아넬로벡터 및 이종성 모이어티를 포함하는 약제학적 조성물을 포함한다.

[0977] 일부 구현예에서, 이종성 모이어티는 바이러스(예를 들어, 이펙터(예를 들어, 약물, 소분자)), 표적화제(예를 들어, DNA 표적화제, 항체, 수용체 리간드), 태그(예를 들어, 형광단, 감광성 작용제, 예컨대 킬러레드(KillerRed)) 또는 본 명세서에 기재된 편집 또는 표적화 모이어티일 수 있다. 일부 구현예에서, 본 명세서에 기재된 막 전위 폴리펩티드는 하나 이상의 이종성 모이어티에 연결된다. 일 구현예에서, 이종성 모이어티는 소분자(예를 들어, 펩티드 모방체 또는 2000 달톤 미만의 분자량을 갖는 작은 유기 분자), 펩티드 또는 폴리펩티드(예를 들어, 항체 또는 이의 항원-결합 단편), 나노입자, 압타머 또는 약리학적 작용제(pharmacoagent)이다.

[0978] 바이러스

[0979] 일부 구현예에서, 아넬로벡터 또는 조성물(예를 들어, 본 명세서에 기재된 바와 같음)은, 예를 들어, 이종성 모이어티로서, 아넬로바이러스 이외의 바이러스, 예를 들어, 단일 가닥 DNA 바이러스, 예를 들어, 비드나바이러스, 써코바이러스, 제미니바이러스, 게노모바이러스, 이노바이러스, 마이크로바이러스, 나노바이러스, 파보바이러스 및 스피라바이러스 유래의 하나 이상의 구성성분 또는 요소(예를 들어, 핵산 또는 폴리펩티드)를 추가로 포함할 수 있다. 일부 구현예에서, 조성물은 이중 가닥 DNA 바이러스, 예를 들어, 아데노바이러스, 암폴라바이러스, 아스코바이러스, 아스파바이러스, 배큘로바이러스, 푸셀로바이러스, 글로불로바이러스, 구타바이러스, 하이트로사바이러스, 헤르페스바이러스, 이리도바이러스, 리포트릭스바이러스, 니마바이러스 및 폭스바이러스를 추가로 포함할 수 있다. 일부 구현예에서, 조성물은 RNA 바이러스, 예를 들어, 알파바이러스, 푸로바이러스, 간염 바이러스, 호르데이바이러스, 토바모바이러스, 토브라바이러스, 트리코르나바이러스, 루비바이러스, 비르나바이러스, 시스토바이러스, 파르티티바이러스 및 레오바이러스를 추가로 포함할 수 있다. 일부 구현예에서, 아넬로벡터는 이종성 모이어티로서 바이러스와 함께 투여된다.

[0980] 일부 구현예에서, 이종성 모이어티는 비-병원성, 예를 들어, 상리공생, 편리공생, 고유 바이러스를 포함할 수 있다. 일부 구현예에서, 비-병원성 바이러스는 하나 이상의 아넬로바이러스, 예를 들어, 알파토르크바이러스(TT), 베타토르크바이러스(TTM) 및 감마토르크바이러스(TTMD)이다. 일부 구현예에서, 아넬로바이러스는 토르크 테노 바이러스(TT), SEN 바이러스, 센티넬 바이러스, TTV-유사 미니 바이러스, TT 바이러스, TT 바이러스 유전형 6, TT 바이러스 군, TTV-유사 바이러스 DXL1, TTV-유사 바이러스 DXL2, 토르크 테노-유사 미니 바이러스(TTM) 또는 토르크 테노-유사 미디 바이러스(TTMD)를 포함할 수 있다. 일부 구현예에서, 비-병원성 바이러스는 본 명세서에 기재된 뉴클레오티드 서열 중 어느 하나에 대하여 적어도 약 60%, 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98% 및 99% 뉴클레오티드 서열 동일성을 갖는 하나 이상의 서열을 포함한다.

[0981] 일부 구현예에서, 이종성 모이어티는 대상체에서 결여된 것으로 확인된 하나 이상의 바이러스를 포함할 수 있다. 예를 들어, 디스비로시스를 갖는 것으로 확인된 대상체에는 대상체에서 불균형인 아넬로벡터 및 하나 이상의 바이러스 성분 또는 바이러스를 포함하는, 또는 기준 값, 예를 들어, 건강한 대상체와 상이한 비를 갖는 조성물이 투여될 수 있다.

[0982] 일부 구현예에서, 이종성 모이어티는 하나 이상의 비-아넬로바이러스, 예를 들어, 아데노바이러스, 헤르페스 바이러스, 폭스 바이러스, 백시니아 바이러스, SV40, 파필로마 바이러스, RNA 바이러스, 예컨대 레트로바이러스, 예를 들어, 렌티 바이러스, 단일-가닥 RNA 바이러스, 예를 들어, 간염 바이러스 또는 이중-가닥 RNA 바이러스, 예를 들어, 로타바이러스를 포함할 수 있다. 일부 구현예에서, 아넬로벡터 또는 바이러스는 결함이 있거나, 감염성 입자를 생성하기 위하여 지원을 필요로 한다. 이러한 지원은 예를 들어, LTR 내의 조절 서열의 제어 하에 복제 결함 아넬로벡터 또는 바이러스의 구조적 유전자 중 하나 이상(예를 들어, 이의 전부)를 인코딩하는, 핵산, 예를 들어, 게놈 내로 통합되는 플라스미드 또는 DNA를 함유하는 헬퍼 세포주를 사용함으로써 제공될 수 있다. 본 명세서에 기재된 아넬로벡터를 복제하는 데 적합한 세포주는 해당 분야에 알려진 세포주, 예를 들어, 본 명세서에 기재된 바와 같이 변형될 수 있는 A549 세포를 포함한다.

[0983] 표적화 모이어티

[0984] 일부 구현예에서, 본 명세서에 기재된 조성물 또는 아넬로벡터는 표적화 모이어티, 예를 들어, 표적 세포 상에 존재하는 관심 분자에 특이적으로 결합하는 표적화 모이어티를 추가로 포함할 수 있다. 표적화 모이어티는 관심 분자 또는 세포의 특정 기능을 조절하거나, 특정 분자(예를 들어, 효소, 단백질 또는 핵산), 예를 들어, 경로 내의 관심 분자의 하류의 특정 분자를 조절하거나, 표적에 특이적으로 결합하여, 아넬로벡터 또는 유전 요소를

국소화시킬 수 있다. 예를 들어, 표적화 모이어티는 특정 관심 분자와 상호작용하여 이의 기능을 증가시키거나, 감소시키거나, 또는 다르게 조절하는 치료제를 포함할 수 있다.

[0985] 태그화 또는 모니터링 모이어티

[0986] 일부 구현예에서, 본 명세서에 기재된 조성물 또는 아넬로벡터는 태그를 추가로 포함하여, 본 명세서에 기재된 아넬로벡터 또는 유전 요소를 표지하거나 모니터링할 수 있다. 태그화 또는 모니터링 모이어티는 화학적 작용제 또는 효소적 절단, 예컨대 단백질 분해 또는 인테인 스플라이싱에 의해 제거할 수 있다. 친화성 태그는 친화성 기법을 사용하여 태그화된 폴리펩티드를 정제하는 데 유용할 수 있다. 일부 예에는 키틴 결합 단백질(CBP), 말토스 결합 단백질(MBP), 글루타티온-S-트랜스퍼라제(GST) 및 폴리(His) 태그가 포함된다. 가용화 태그는 샤페론-결함 종, 예컨대 에스케리키아 콜라이(*E. coli*)에서 발현되는 재조합 단백질이 단백질에서 적절하게 폴딩되는 것을 지원하는 것을 돕고, 그들이 침전하지 못하게 하는 데 유용할 수 있다. 일부 예는 티오레독신(TRX) 및 폴리(NANP)를 포함한다. 태그화 또는 모니터링 모이어티는 감광성 태그, 예를 들어, 형광을 포함할 수 있다. 형광 태그는 시각화에 유용하다. GFP 및 이의 변이체는 형광 태그로서 흔히 사용되는 몇몇의 예이다. 단백질 태그는 특정 효소적 변형(예컨대, 비오틴 리가제에 의한 비오티닐화) 또는 화학적 변형(예컨대 형광 영상화를 위한 FlAsH-EDT2와의 반응)이 일어나게 할 수 있다. 종종 태그화 또는 모니터링 모이어티를 조합하여, 단백질을 다수의 다른 성분과 연결시킨다. 태그화 또는 모니터링 모이어티는 또한 특정 단백질분해 또는 효소적 절단에 의해(예를 들어, TEV 프로테아제, 트롬빈, 인자 Xa 또는 엔테로펩티다제에 의해) 제거될 수 있다.

[0987] 나노입자

[0988] 일부 구현예에서, 본 명세서에 기재된 조성물 또는 아넬로벡터는 나노입자를 추가로 포함할 수 있다. 나노입자는 약 1 내지 약 1000 나노미터, 약 1 내지 약 500 나노미터 크기, 약 1 내지 약 100 nm, 약 50 nm 내지 약 300 nm, 약 75 nm 내지 약 200 nm, 약 100 nm 내지 약 200 nm, 및 그 사이의 임의의 범위의 크기를 갖는 무기 물질을 포함한다. 나노입자는 일반적으로, 나노규모 크기의 복합 구조를 갖는다. 일부 구현예에서, 나노입자는 통상적으로 구형이지만, 나노입자 조성에 따라 상이한 형태가 가능하다. 나노입자의 외부 환경과 접촉하는 나노입자의 부분은 일반적으로 나노입자의 표면으로서 확인된다. 본 명세서에 기재된 나노입자에서, 크기 제한은 2차원으로 제한될 수 있으며, 이에 따라, 나노입자는 약 1 내지 약 1000 nm의 직경을 갖는 복합 구조를 포함하게 되며, 여기서 특정 직경은 나노입자 조성물 및 실험 설계에 따라 의도된 나노입자의 용도에 좌우된다. 예를 들어, 치료적 응용에 사용되는 나노입자는 통상적으로 약 200 nm 이하의 크기를 갖는다.

[0989] 나노입자의 바람직한 추가의 특성, 예컨대 표면 전하 및 입체적 안정화는 또한, 특정 관심 응용을 고려하여 달라질 수 있다. 임상적 응용, 예컨대 암 치료에 바람직할 수 있는 예시적인 특성은 이의 전문이 각각 본 명세서에 참조로 포함되는 문헌[Davis et al, Nature 2008 vol. 7, pages 771-782]; 문헌[Duncan, Nature 2006 vol. 6, pages 688-701]; 및 문헌[Allen, Nature 2002 vol. 2 pages 750-763]에 기재되어 있다. 추가의 특성은 본 개시내용의 판독 시에 당업자에 의해 확인 가능하다. 나노입자 치수 및 특성은 해당 분야에 알려진 기법에 의해 검출될 수 있다. 입자 치수를 검출하기 위한 예시적인 기법은 동적 광 산란(DLS) 및 다양한 현미경관찰, 예컨대 투과 전자 현미경관찰(TEM) 및 원자력 현미경관찰(AFM)을 포함하지만, 이에 제한되지 않는다. 입자 형태를 검출하기 위한 예시적인 기법은 TEM 및 AFM을 포함하지만 이들에 제한되지 않는다. 나노입자의 표면 전하를 검출하기 위한 예시적인 기법은 제타 전위 방법을 포함하지만, 이에 제한되지 않는다. 다른 화학적 특성을 검출하는 데 적합한 추가의 기법은 ^{1}H, ^{11}B 및 ^{13}C 및 ^{19}F NMR, UV/Vis 및 적외선/라만 분광학 및 형광 분광학(나노입자가 형광 표지와 조합하여 사용되는 경우) 및 당업자에 의해 확인 가능한 추가의 기법을 포함한다.

[0990] 소분자

[0991] 일부 구현예에서, 본 명세서에 기재된 조성물 또는 아넬로벡터는 소분자를 추가로 포함할 수 있다. 소분자 모이어티는 작은 펩티드, 펩티드 모방체(예를 들어, 펩토이드), 아미노산, 아미노산 유사체, 합성 폴리뉴클레오티드, 폴리뉴클레오티드 유사체, 뉴클레오티드, 뉴클레오티드 유사체, 일반적으로 몰당 약 5,000 그램 미만의 분자량을 갖는 유기 및 무기 화합물(이종유기 및 유기금속 화합물 포함), 예를 들어, 몰당 약 2,000 그램 미만의 분자량을 갖는 유기 또는 무기 화합물, 예를 들어, 몰당 약 1,000 그램 미만의 분자량을 갖는 유기 또는 무기 화합물, 예를 들어, 몰당 약 500 그램 미만의 분자량을 갖는 유기 또는 무기 화합물, 및 이러한 화합물의 염, 에스테르 및 다른 약제학적으로 허용 가능한 형태를 포함하지만, 이에 제한되지 않는다. 소분자는 신경전달물질, 호르몬, 약물, 독소, 바이러스 또는 미생물 입자, 합성 분자, 및 효능제 또는 길항제를 포함할 수 있지만, 이들에 제한되지 않는다.

[0992] 적합한 소분자의 예는 모두 본 명세서에 참조로 포함되는 문헌["The Pharmacological Basis of Therapeutics," Goodman and Gilman, McGraw-Hill, New York, N.Y., (1996), Ninth edition]의 다음과 같은 섹션에 기재된 것들을 포함한다: 시냅스 및 신경이펙터 연결 부위에서 작용하는 약물; 중추신경계에서 작용하는 약물; 오타코이드: 염증의 약물 요법; 물, 염 및 이온; 신장 기능 및 전해질 대사에 영향을 미치는 약물; 심혈관 약물; 위장 기능에 영향을 미치는 약물; 자궁 운동에 영향을 미치는 약물; 기생충 감염의 화학요법; 미생물 질환의 화학요법; 신생물 질환의 화학요법; 면역억제를 위해 사용되는 약물; 혈액-형성 기관에 작용하는 약물; 호르몬 및 호르몬 길항제; 비타민, 피부과학 및 독성학. 소분자의 일부 예는 프리온 약물, 예컨대 타크롤리무스, 유비퀴틴 리가제 또는 HECT 리가제 저해제, 예컨대 헤클린(heclin), 히스톤 변형 약물, 예컨대 부티르산나트륨, 효소적 저해제, 예컨대 5-아자-시티딘, 안트라사이클린, 예컨대 독소루비신, 베타-락탐, 예컨대 페니실린, 항-박테리아제, 화학요법제, 항-바이러스제, 다른 유기체로부터의 조절제, 예컨대 VP64, 및 불충분한 생체 이용 가능성을 갖는 약물, 예컨대 결함이 있는 약동학을 갖는 화학치료제를 포함하지만, 이에 제한되지 않는다.

[0993] 일부 구현예에서, 소분자는 후성적 변형제, 예를 들어 문헌[de Groote et al. Nuc. Acids Res. (2012):1-18]에 기재된 것들이다. 예시적인 소분자 후성적 변형제는 예를 들어, 본 명세서에 참조로 포함되는 문헌[Lu et al. J. Biomolecular Screening 17.5(2012):555-71], 예를 들어, 표 1 또는 2에 기재되어 있다. 일부 구현예에서, 후성적 변형제는 보리노스타트(vorinostat) 또는 로미뎁신(romidepsin)을 포함한다. 일부 구현예에서, 후성적 변형제는 부류 I, II, III 및/또는 IV 히스톤 데아세틸라제(HDAC)의 저해제를 포함한다. 일부 구현예에서, 후성적 변형제는 SirTI의 활성화제를 포함한다. 일부 구현예에서, 후성적 변형제는 가르시놀(Garcinol), Lys-CoA, C646, (+)-JQI, I-BET, BICI, MS120, DZNep, UNC0321, EPZ004777, AZ505, AMI-I, 피라졸 아미드 7b, 벤조[d]이미다졸 17b, 아실화된 답손 유도체(예를 들어, PRMTI), 메틸스타트(methylstat), 4,4'-디카복시-2,2'-비피리딘, SID 85736331, 하이드록사메이트 유사체 8, 타닐사이프로미(tanylcypromie), 비스구아니딘 및 비구아니드 폴리아민 유사체, UNC669, 비다자(Vidaza), 데시타빈(decitabine), 소듐 페닐 부티레이트(SDB), 리포산(LA), 케르세틴(quercetin), 발프로산, 하이드랄라진(hydralazine), 박트림(bactrim), 녹차 추출물(예를 들어, 에피갈로카테킨 갈레이트(EGCG)), 커큐민, 술포르판(sulforphane) 및/또는 알리신(allicin)/디알릴 디술피드를 포함한다. 일부 구현예에서, 후성적 변형제는 DNA 메틸화를 저해하며, 예를 들어, DNA 메틸트랜스퍼라제의 저해제이다(예를 들어, 5-아자시티딘이고/이거나 데시타빈이다). 일부 구현예에서, 후성적 변형제는 히스톤 변형, 예를 들어, 히스톤 아세틸화, 히스톤 메틸화, 히스톤 수모화(sumoylation) 및/또는 히스톤 인산화를 변형시킨다. 일부 구현예에서, 후성적 변형제는 히스톤 데아세틸라제의 저해제이다(예를 들어, 보리노스타트 및/또는 트리코스타틴 A이다).

[0994] 일부 구현예에서, 소분자는 약제학적 활성 작용제이다. 일 구현예에서, 소분자는 대사 활성 또는 성분의 저해제이다. 유용한 부류의 약제학적 활성 작용제는 항생제, 항-염증 약물, 혈관신생 또는 혈관작용제, 성장 인자 및 화학치료제(항-신생물)제(예를 들어, 종양 억제제)를 포함하지만, 이에 제한되지 않는다. 본 명세서에 기재된 범주 및 예로부터의 또는 문헌[(Orme-Johnson 2007, Methods Cell Biol. 2007;80:813-26)]으로부터의 분자 중 하나 또는 조합이 사용될 수 있다. 일 구현예에서, 본 발명은 항생제, 항-염증 약물, 혈관신생 또는 혈관작용제, 성장 인자 또는 화학치료제를 포함하는 조성물을 포함한다.

[0995] 펩티드 또는 단백질

[0996] 일부 구현예에서, 본 명세서에 기재된 조성물 또는 아넬로벡터는 펩티드 또는 단백질을 추가로 포함할 수 있다. 펩티드 모이어티는 펩티드 리간드 또는 항체 단편(예를 들어, 수용체, 예컨대 세포외 수용체에 결합하는 항체 단편), 신경펩티드, 호르몬 펩티드, 펩티드 약물, 독성 펩티드, 바이러스 또는 미생물 펩티드, 합성 펩티드 및 효능제 또는 길항제 펩티드를 포함할 수 있지만, 이들에 제한되지 않는다.

[0997] 펩티드 모이어티는 선형 또는 분지형일 수 있다. 펩티드는 약 5 내지 약 200개의 아미노산, 약 15 내지 약 150개의 아미노산, 약 20 내지 약 125개의 아미노산, 약 25 내지 약 100개의 아미노산 또는 그 사이의 임의의 범위의 길이를 갖는다.

[0998] 펩티드의 일부 예는 형광 태그 또는 마커, 항원, 항체, 항체 단편, 예컨대 단일 도메인 항체, 리간드 및 수용체, 예컨대 글루카곤-유사 펩티드-1(GLP-1), GLP-2 수용체 2, 콜레시스토키닌(cholecystokinin) B(CCKB) 및 소마토스타틴 수용체, 펩티드 치료제, 예컨대 특정 세포 표면 수용체, 예컨대 G 단백질-커플링된 수용체(GPCR) 또는 이온 채널에 결합되는 것들, 천연적으로 생물활성인 펩티드로부터의 합성 또는 유사체 펩티드, 항-미생물 펩티드, 포어-형성 펩티드, 종양 표적화 또는 세포독성 펩티드, 및 분해 또는 자가-파괴 펩티드, 예컨대 아폽토시스-유도 펩티드 신호 또는 감광제 펩티드를 포함하지만 이들에 제한되지 않는다.

[0999] 본 명세서에 기재된 본 발명에서 유용한 펩티드는 또한 작은 항원-결합 펩티드, 예를 들어, 항원 결합 항체 또는 항체-유사 단편, 예컨대 단일 쇄 항체, 나노바디를 포함한다(예를 들어, 문헌[Steeland et al. 2016. Nanobodies as therapeutics: big opportunities for small antibodies. Drug Discov Today: 21(7):1076-113] 참조). 이러한 작은 항원 결합 펩티드는 사이토졸 항원, 핵 항원, 세포소기관-내 항원에 결합할 수 있다.

[1000] 일부 구현예에서, 본 명세서에 기재된 조성물 또는 아넬로벡터는 특정 위치, 조직 또는 세포를 표적화할 수 있는 리간드에 연결된 폴리펩티드를 포함한다.

[1001] 올리고뉴클레오티드 압타머

[1002] 일부 구현예에서, 본 명세서에 기재된 조성물 또는 아넬로벡터는 올리고뉴클레오티드 압타머를 추가로 포함할 수 있다. 압타머 모이어티는 올리고뉴클레오티드 또는 펩티드 압타머이다. 올리고뉴클레오티드 압타머는 단백질 및 펩티드를 포함하는 사전-선택된 표적에 높은 친화성 및 특이성으로 결합할 수 있는 단일-가닥 DNA 또는 RNA(ssDNA 또는 ssRNA) 분자이다.

[1003] 올리고뉴클레오티드 압타머는 다양한 분자 표적, 예컨대 소분자, 단백질, 핵산 및 심지어 세포, 조직 및 유기체에 결합하도록 반복된 회차의 시험관 내 선택 또는 동등하게, SELEX(지수적 농축에 의한 리간드의 계통 진화)를 통해 엔지니어링될 수 있는 핵산 종이다. 압타머는 차별적인 분자 인식을 제공하며, 화학적 합성에 의해 생성될 수 있다. 또한, 압타머는 바람직한 저장 특성을 갖고, 치료적 응용에서 면역원성을 거의 유도하지 않거나, 전혀 유도하지 않을 수 있다.

[1004] DNA 및 RNA 압타머는 둘 모두 다양한 표적에 대하여 강력한 결합 친화성을 보여줄 수 있다. 예를 들어, DNA 및 RNA 압타머는 t 라이소자임, 트롬빈, 인간 면역결핍 바이러스 트랜스-작용 반응성 요소(HIV TAR),(en.wikipedia.org/wiki/Aptamer - cite_note-10 참조), 헤민(hemin), 인터페론 γ, 혈관 내피 성장 인자(VEGF), 전립선 특이적 항원(PSA), 도파민 및 비-정규 종양유전자, 열 충격 인자 1(HSF1)을 위하여 선택된다.

[1005] 펩티드 압타머

[1006] 일부 구현예에서, 본 명세서에 기재된 조성물 또는 아넬로벡터는 펩티드 압타머를 추가로 포함할 수 있다. 펩티드 압타머는 낮은 분자량, 12 내지 14 kDa을 갖는 펩티드를 포함하는 하나의(또는 그 이상의) 짧은 가변 펩티드 도메인을 갖는다. 펩티드 압타머는 세포에 특이적으로 결합하고, 세포 내측의 단백질-단백질 상호작용을 간섭하도록 설계될 수 있다.

[1007] 펩티드 압타머는 특이적 표적 분자에 결합하도록 선택되거나 엔지니어링된 인공 단백질이다. 이들 단백질은 가변 서열의 하나 이상의 펩티드 루프를 포함한다. 이들은 통상적으로 조합 라이브러리로부터 단리되며, 종종 유도된 돌연변이 또는 가변 영역 돌연변이유발 및 선택의 회차에 의해 후속적으로 개선된다. 생체 내에서, 펩티드 압타머는 세포 단백질 표적에 결합할 수 있고, 그들의 표적화된 분자와 다른 단백질의 정상적인 단백질 상호작용의 간섭을 포함하는 생물학적 효과를 발휘할 수 있다. 특히, 전사 인자 결합 도메인에 부착된 가변 펩티드 압타머 루프는 전사 인자 활성화 도메인에 부착된 표적 단백질에 대해 스크리닝한다. 이러한 선택 전략을 통해 펩티드 압타머의 표적으로의 펩티드 압타머의 생체 내 결합은 하류 효모 마커 유전자의 발현으로서 검출된다. 이러한 실험은 표현형을 유발하기 위해 압타머에 의해 결합되는 특정 단백질 및 압타머가 파괴하는 단백질 상호작용을 확인한다. 또한, 적절한 기능적 모이어티로 유도체화된 펩티드 압타머는 그들의 표적 단백질의 특정 번역-후 변형을 유발할 수 있거나, 표적의 하위세포 국소화를 변경시킬 수 있다.

[1008] 펩티드 압타머는 또한 시험관 내에서 표적을 인식할 수 있다. 그들은 바이오센서에서 항체 대신에 이용되며, 비활성 및 활성 단백질 형태 둘 모두를 함유하는 집단으로부터 단백질의 활성 아이소폼을 검출하기 위해 이용된다. 펩티드 압타머 "헤드"가 독특한 서열의 이중-가닥 DNA "테일(tail)"에 공유적으로 연결된 태드폴(tadpole)로서 알려져 있는 유도체는 (예를 들어, 정량적 실시간 중합효소 연쇄 반응을 사용하여) 그들의 DNA 테일의 PCR에 의해 혼합물 중 드문 표적 분자의 정량화를 가능하게 한다.

[1009] 펩티드 압타머 선택은 상이한 시스템을 사용하여 이루어질 수 있지만, 현재 효모 2-하이브리드 시스템이 가장 많이 사용된다. 펩티드 압타머는 또한 파지 디스플레이 또는 다른 표면 디스플레이 기술, 예컨대 mRNA 디스플레이, 리보좀 디스플레이, 박테리아 디스플레이 및 효모 디스플레이에 의해 구축된 조합 펩티드 라이브러리로부터 선택될 수 있다. 이들 실험 절차는 또한 바이오패닝(biopanning)으로서 알려져 있다. 바이오패닝으로부터 수득되는 펩티드 중에, 미모토프(mimotope)는 펩티드 압타머의 한 종류로서 간주될 수 있다. 조합 펩티드 라이브러리로부터 패닝된 모든 펩티드는 명칭 MimoDB를 갖는 특수 데이터에 저장되어 있다.

[1010] **V. 숙주 세포**

[1011] 본 발명은 또한 본 명세서에 기재된 아넬로벡터를 포함하는 숙주 또는 숙주 세포에 관한 것이다. 일부 구현예에서, 숙주 또는 숙주 세포는 식물, 곤충, 박테리아, 진균, 척추동물, 포유동물(예를 들어, 인간), 또는 기타 유기체 또는 세포이다. 특정 구현예에서, 본 명세서에서 확인된 바와 같이, 제공된 아넬로벡터는 다양한 상이한 숙주 세포를 감염시킨다. 표적 숙주 세포는 중배엽, 내배엽, 또는 외배엽 기원의 세포를 포함한다. 표적 숙주 세포는, 예를 들어, 상피 세포, 근육 세포, 백혈구 세포(예를 들어, 림프구), 신장 조직 세포, 폐 조직 세포를 포함한다.

[1012] 일부 구현예에서, 아넬로벡터는 숙주에서 실질적으로 비-면역원성이다. 아넬로벡터 또는 유전 요소는 숙주의 면역계에 의해 요망되지 않는 상당한 반응을 생성하지 않는다. 일부 면역 반응은 체액성 면역 반응(예를 들어, 항원-특이적 항체의 생성) 및 세포-매개의 면역 반응(예를 들어, 림프구 증식)을 포함하지만, 이에 제한되지 않는다.

[1013] 일부 구현예에서, 숙주 또는 숙주 세포는 아넬로벡터와 접촉된다(예를 들어, 이로 감염됨). 일부 구현예에서, 숙주는 포유동물, 예컨대, 인간이다. 숙주 내의 아넬로벡터의 양은 투여 후 임의의 시간에 측정될 수 있다. 특정 구현예에서, 배양물 중 아넬로벡터 성장의 시간 경과가 결정된다.

[1014] 일부 구현예에서, 아넬로벡터, 예를 들어, 본 명세서에 기재된 바와 같은 아넬로벡터는 유전 가능하다. 일부 구현예에서, 아넬로벡터는 유체 및/또는 세포 중에서, 모체에서 아이에게 연속적으로 전달된다. 일부 구현예에서, 원래의 숙주 세포로부터의 딸 세포는 아넬로벡터를 포함한다. 일부 구현예에서, 모체는 아넬로벡터를 적어도 25%, 50%, 60%, 70%, 80%, 85%, 90%, 95%, 또는 99%의 효율로 아이에게, 또는 적어도 25%, 50%, 60%, 70%, 80%, 85%, 90%, 95%, 또는 99%의 전달 효율로 숙주 세포로부터 딸 세포로 전달한다. 일부 구현예에서, 숙주 세포 내의 아넬로벡터는 감수분열 동안 25%, 50%, 60%, 70%, 80%, 85%, 90%, 95%, 또는 99%의 전달 효율을 갖는다. 일부 구현예에서, 숙주 세포 내의 아넬로벡터는 유사분열 동안 적어도 25%, 50%, 60%, 70%, 80%, 85%, 90%, 95%, 또는 99%의 전달 효율을 갖는다. 일부 구현예에서, 세포 내의 아넬로벡터는 약 10% 내지 20%, 20% 내지 30%, 30% 내지 40%, 40% 내지 50%, 50% 내지 60%, 60% 내지 70%, 70% 내지 75%, 75% 내지 80%, 80% 내지 85%, 85% 내지 90%, 90% 내지 95%, 95% 내지 99% 또는 그 사이의 임의의 백분율의 전달 효율을 갖는다.

[1015] 일부 구현예에서, 아넬로벡터, 예를 들어, 아넬로벡터는 숙주 세포 내에서 복제한다. 일 구현예에서, 아넬로벡터는 포유동물 세포, 예를 들어, 인간 세포에서 복제할 수 있다. 다른 구현예에서, 아넬로벡터는 복제 결핍 또는 복제 비적격이다.

[1016] 일부 구현예에서, 아넬로벡터가 숙주 세포에서 복제되지만, 아넬로벡터는 숙주의 게놈 내로, 예를 들어, 숙주의 염색체와 통합되지 않는다. 일부 구현예에서, 아넬로벡터와 숙주의 염색체와의 재조합 빈도는 무시할 수 있는 정도이다. 일부 구현예에서, 아넬로벡터는, 예를 들어, 숙주의 염색체와, 예를 들어, 약 1.0 cM/Mb, 0.9 cM/Mb, 0.8 cM/Mb, 0.7 cM/Mb, 0.6 cM/Mb, 0.5 cM/Mb, 0.4 cM/Mb, 0.3 cM/Mb, 0.2 cM/Mb, 0.1 cM/Mb 미만의 재조합 빈도를 갖는다.

[1017] **VI. 사용 방법**

[1018] 본 명세서에 기재된 아넬로벡터 및 아넬로벡터를 포함하는 조성물은 예를 들어, 질환, 장애 또는 병태의 치료를 필요로 하는 대상체(예를 들어, 포유동물 대상체, 예를 들어, 인간 대상체)에서의 질환, 장애 또는 병태의 치료 방법에 사용될 수 있다. 본 명세서에 기재된 약제학적 조성물의 투여는 예를 들어, 비경구(정맥내, 종양내, 복강내, 근육내, 강내 및 피하 포함) 투여로 이루어질 수 있다. 아넬로벡터는 단독으로 또는 약제학적 조성물로서 제형화되어 투여될 수 있다.

[1019] 아넬로벡터는 단위-용량 조성물, 예컨대 단위 용량 비경구 조성물의 형태로 투여될 수 있다. 이러한 조성물은 일반적으로 혼합에 의해 제조되며, 비경구 투여를 위하여 적합하게 조정될 수 있다. 이러한 조성물은 예를 들어, 주사 및 주입 용액 또는 현탁액 또는 좌제 또는 에어로졸의 형태로 존재할 수 있다.

[1020] 일부 구현예에서, 예를 들어, 본 명세서에 기재된 바와 같은 아넬로벡터 또는 이를 포함하는 조성물의 투여는 아넬로벡터에 의해 포함되는 유전 요소가 예를 들어, 대상체 내의 표적 세포로 운반되게 할 수 있다.

[1021] 예를 들어, 이펙터(예를 들어, 내인성 또는 외인성 이펙터)를 포함하는 본 명세서에 기재된 아넬로벡터 또는 이의 조성물을 사용하여 이펙터를 세포, 조직 또는 대상체에 운반할 수 있다. 일부 구현예에서, 아넬로벡터 또는 이의 조성물을 사용하여, 이펙터를 골수, 혈액, 심장, GI 또는 피부에 운반한다. 본 명세서에 기재된 아넬로벡

터 조성물의 투여에 의한 이펙터의 운반은 세포, 조직 또는 대상체에서 비코딩 RNA 또는 폴리펩티드의 발현 수준을 조절할 수 있다(예를 들어, 증가 또는 감소시킬 수 있다). 이러한 방식으로의 발현 수준의 조절은 이펙터가 운반되는 세포에서 기능적 활성의 변경을 초래할 수 있다. 일부 구현예에서, 조절된 기능적 활성은 성질이 효소적, 구조적 또는 조절적일 수 있다.

[1022] 일부 구현예에서, 아넬로벡터 또는 이의 카피는 세포로의 운반 후 24시간(예를 들어, 1일, 2일, 3일, 4일, 5일, 6일, 1주, 2주, 3주, 4주, 30일 또는 1개월)에 세포에서 검출 가능하다. 구현예에서, 아넬로벡터 또는 이의 조성물은 표적 세포 상에서 효과를 매개하며, 효과는 적어도 1, 2, 3, 4, 5, 6 또는 7일, 2, 3 또는 4주 또는 1, 2, 3, 6 또는 12개월 동안 지속된다. (예를 들어, 아넬로벡터 또는 이의 조성물이 외인성 단백질을 인코딩하는 유전 요소를 포함하는) 일부 구현예에서, 효과는 1, 2, 3, 4, 5, 6 또는 7일, 2, 3 또는 4주 또는 1, 2, 3, 6 또는 12개월 미만 동안 지속된다.

[1023] 본 명세서에 기재된 아넬로벡터 또는 아넬로벡터를 포함하는 조성물로 치료될 수 있는 질환, 장애 및 병태의 예는 제한 없이, 면역 장애, 인터페론병증(예를 들어, I형 인터페론병증), 감염성 질환, 염증성 장애, 자가면역 병태, 암(예를 들어, 고형 종양, 예를 들어, 폐암, 비-소세포 폐암, 예를 들어, mIR-625, 예를 들어, 카스파제-3에 반응성인 유전자를 발현하는 종양) 및 위장 장애를 포함한다. 일부 구현예에서, 아넬로벡터는 아넬로벡터가 접촉되는 세포 내에서 활성 또는 기능을 조절한다(예를 들어, 증가 또는 감소시킨다). 일부 구현예에서, 아넬로벡터는 아넬로벡터가 접촉되는 세포에서 분자(예를 들어, 핵산 또는 단백질)의 수준 또는 활성을 조절한다(예를 들어, 증가 또는 감소시킨다). 일부 구현예에서, 아넬로벡터는 아넬로벡터가 접촉되는 세포, 예를 들어, 암 세포의 생존력을 예를 들어, 적어도 약 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 99% 이상 감소시킨다. 일부 구현예에서, 아넬로벡터는 아넬로벡터가 접촉되는 세포, 예를 들어, 암 세포의 생존력을 예를 들어, 적어도 약 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 99% 이상 감소시키는 이펙터, 예를 들어, miRNA, 예를 들어, miR-625를 포함한다. 일부 구현예에서, 아넬로벡터는 예를 들어, 카스파제-3 활성을 증가시킴으로써 아넬로벡터가 접촉되는 세포, 예를 들어, 암 세포의 아폽토시스를 예를 들어, 적어도 약 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 99% 이상 증가시킨다. 일부 구현예에서, 아넬로벡터는 예를 들어, 카스파제-3 활성을 증가시킴으로써, 아넬로벡터가 접촉되는 세포, 예를 들어, 암 세포의 아폽토시스를 예를 들어, 적어도 약 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 99% 이상 증가시키는 이펙터, 예를 들어, miRNA, 예를 들어, miR-625를 포함한다.

[1024] **VII. 투여/운반**

[1025] 조성물(예를 들어, 본 명세서에 기재된 바와 같은 아넬로벡터를 포함하는 약제학적 조성물)은 약제학적으로 허용 가능한 부형제를 포함하도록 제형화될 수 있다. 약제학적 조성물은 선택적으로 하나 이상의 추가의 활성 물질, 예를 들어, 치료적 및/또는 예방적 활성 물질을 포함할 수 있다. 본 발명의 약제학적 조성물은 멸균 및/또는 발열원-부재일 수 있다. 약제의 제형화 및/또는 제조에서의 일반적인 고려사항은 예를 들어, 문헌 [Remington: The Science and Practice of Pharmacy 21st ed., Lippincott Williams & Wilkins, 2005](본 명세서에 참조로 포함됨)에서 찾을 수 있다.

[1026] 본 명세서에 제공되는 약제학적 조성물의 설명이 주로 인간으로의 투여에 적합한 약제학적 조성물에 관한 것이지만, 이러한 조성물이 일반적으로 임의의 다른 동물로의, 예를 들어, 비-인간 동물, 예를 들어, 비-인간 포유동물로의 투여에 적합하다는 것을 당업자라면 이해할 것이다. 조성물이 다양한 동물로의 투여에 적합하게 만들기 위해 인간으로의 투여에 적합한 약제학적 조성물을 변형하는 것이 충분히 이해되며, 숙련된 수의학 약리학자는 존재한다면, 단지 일상적인 실험을 사용하여 이러한 변형을 설계하고/설계하거나 수행할 수 있다. 약제학적 조성물의 투여가 고려되는 대상체는 인간 및/또는 기타 영장류; 포유동물, 예컨대 상업적으로 관련된 포유동물, 예컨대 소, 돼지, 말, 양, 고양이, 개, 마우스 및/또는 랫트; 및/또는 조류, 예컨대 상업적으로 관련된 조류, 예컨대 가금류, 닭, 오리, 거위 및/또는 칠면조를 포함하지만, 이에 제한되지 않는다.

[1027] 본 명세서에 기재된 약제학적 조성물의 제형은 약리학 분야에 알려져 있거나, 이후에 개발되는 임의의 방법에 의해 제조될 수 있다. 일반적으로, 이러한 제조 방법은 활성 성분이 부형제 및/또는 하나 이상의 다른 보조 성분과 회합되게 한 다음, 필요하면 그리고/또는 바람직하면, 생성물을 나누고/나누거나, 성형하고/하거나 패키징하는 단계를 포함한다.

[1028] 일 양태에서, 본 발명은 대상체로 아넬로벡터를 운반하는 방법을 특징으로 한다. 방법은 본 명세서에 기재된 바와 같은 아넬로벡터를 포함하는 약제학적 조성물을 대상체로 투여하는 단계를 포함한다. 일부 구현예에서, 투여된 아넬로벡터는 대상체에서 복제한다(예를 들어, 대상체의 바이러스체의 일부가 됨).

[1029] 약제학적 조성물은 야생형 또는 고유 바이러스 요소 및/또는 변형된 바이러스 요소를 포함할 수 있다. 아넬로벡터는 하나 이상의 아넬로바이러스 서열(예를 들어, 핵산 서열 또는 이의 아미노산 서열을 인코딩하는 핵산 서열) 또는 이에 대하여 적어도 약 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% 및 99% 뉴클레오티드 서열 동일성을 갖는 서열을 포함할 수 있다. 아넬로벡터는 하나 이상의 아넬로바이러스 서열(예를 들어, 아넬로바이러스 ORF1 핵산 서열)에 대하여 적어도 약 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% 및 99% 서열 동일성을 갖는 핵산 서열을 포함하는 핵산 분자를 포함할 수 있다. 아넬로벡터는 아넬로바이러스 아미노산 서열(예를 들어, 아넬로바이러스 ORF1 분자의 아미노산 서열)에 대하여 적어도 약 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% 및 99% 서열 동일성을 갖는 아미노산 서열을 인코딩하는 핵산 분자를 포함할 수 있다. 아넬로벡터는 아넬로바이러스 아미노산 서열(예를 들어, 아넬로바이러스 ORF1 분자의 아미노산 서열)에 대하여 적어도 약 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% 및 99% 서열 동일성을 갖는 아미노산 서열을 포함하는 폴리펩티드를 포함할 수 있다.

[1030] 일부 구현예에서, 아넬로벡터는 내인성 유전자 및 단백질 발현을 기준, 예를 들어, 건강한 대조군과 비교하여, 예를 들어, 적어도 약 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 또는 그 이상 증가(자극)시키기에 충분하다. 특정 구현예에서, 아넬로벡터는 내인성 유전자 및 단백질 발현을 기준, 예를 들어, 건강한 대조군과 비교하여, 예를 들어, 적어도 약 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 또는 그 이상 감소(저해)시키기에 충분하다.

[1031] 일부 구현예에서, 아넬로벡터는 숙주 또는 숙주 세포에서 하나 이상의 바이러스 특성, 예를 들어, 향성, 감염성, 면역억제/활성화를 기준, 예를 들어, 건강한 대조군에 비하여 예를 들어, 적어도 약 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 또는 그 이상 저해/향상시킨다.

[1032] 일부 구현예에서, 바이러스 유전 정보에서 표현되지 않는 하나 이상의 바이러스 주를 추가로 포함하는 약제학적 조성물이 대상체에게 투여된다.

[1033] 일부 구현예에서, 본 명세서에 기재된 아넬로벡터를 포함하는 약제학적 조성물은 바이러스 감염을 조절하기에 충분한 용량 및 시간으로 투여된다. 바이러스 감염의 일부 비제한적인 예는 아데노-연관 바이러스, 아이치(Aichi) 바이러스, 호주 박쥐 리사 바이러스(lyssavirus), BK 폴리오마바이러스(polyomavirus), 반나 바이러스(Banna virus), 바르마 포레스트(Barmah forest) 바이러스, 부니아무에라(Bunyamwera) 바이러스, 부니아바이러스 라 크로세(Bunyavirus La Crosse), 부니아바이러스 스노우슈 하레(Bunyavirus snowshoe hare), 세르코피테신 헤르페스바이러스(Cercopithecine herpesvirus), 찬디푸라 바이러스(Chandipura virus), 치쿤구니아 바이러스(Chikungunya virus), 코사바이러스(Cosavirus) A, 우두 바이러스, 콕사키바이러스(Coxsackievirus), 크리민-콩고 출혈성 열 바이러스(Crimean-Congo hemorrhagic fever virus), 뎅기(Dengue) 바이러스, 도리(Dhori) 바이러스, 더그비(Dugbe) 바이러스, 두벤하게(Duvenhage) 바이러스, 동부 말 뇌염 바이러스, 에볼라바이러스, 에코바이러스(Echovirus), 뇌심근염(Encephalomyocarditis) 바이러스, 엡스테인-바 바이러스, 유럽 박쥐 리사바이러스, GB 바이러스 C/G형 간염 바이러스, 한탄(Hantaan) 바이러스, 헨드라(Hendra) 바이러스, A형 간염 바이러스, B형 간염 바이러스, C형 간염 바이러스, E형 간염 바이러스, 간염 델타 바이러스, 마두 바이러스, 인간 아데노바이러스, 인간 아스트로바이러스, 인간 코로나바이러스, 인간 사이토메갈로바이러스, 인간 엔테로바이러스 68, 인간 엔테로바이러스 70, 인간 헤르페스바이러스 1, 인간 헤르페스바이러스 2, 인간 헤르페스바이러스 6, 인간 헤르페스바이러스 7, 인간 헤르페스바이러스 8, 인간 면역결핍 바이러스, 인간 파필로마바이러스 1, 인간 파필로마바이러스 2, 인간 파필로마바이러스 16, 인간 파필로마바이러스 18, 인간 파라인플루엔자, 인간 파보바이러스 B19, 인간 호흡기 융합 바이러스, 인간 리노바이러스, 인간 SARS 코로나바이러스, 인간 스푸마레트로바이러스(spumaretrovirus), 인간 T-림포트로픽 바이러스, 인간 토로바이러스(torovirus), 인플루엔자 A 바이러스, 인플루엔자 B 바이러스, 인플루엔자 C 바이러스, 이스파한(Isfahan) 바이러스, JC 폴리오마바이러스, 일본 뇌염 바이러스, 주닌 아레나바이러스(Junin arenavirus), KI 폴리오마바이러스, 쿤진(Kunjin) 바이러스, 라고스(Lagos) 박쥐 바이러스, 레이크 빅토리아 마르부르그바이러스(Lake Victoria marburgvirus), 란가트(Langat) 바이러스, 라사(Lassa) 바이러스, 로즈데일(Lordsdale) 바이러스, 라우핑 일(Louping ill) 바이러스, 림프구 맥락수막염 바이러스, 마추포(Machupo) 바이러스, 마이아로(Mayaro) 바이러스, MERS 코로나바이러스, 홍역 바이러스, 멘고(Mengo) 뇌심근염 바이러스, 메르켈 세포 폴리오마바이러스(Merkel cell polyomavirus), 모콜라(Mokola) 바이러스, 전염성 연속종 바이러스, 원두 바이러스, 볼거리 바이러스, 무레이 밸리(Murray valley) 뇌염 바이러스, 뉴욕 바이러스, 니파(Nipah) 바이러스, 노르워크(Norwalk) 바이러스, 오뇽-뇽(O'nyong-nyong) 바이러스, 올프(Orf) 바이러스, 오로퓨스(Oropouche) 바이러스, 피친데(Pichinde) 바이러스, 폴리오바이러스, 푼타 토로 플레보바이러스(Punta toro phlebovirus), 푸말라(Puumala) 바이러스, 광견병 바이러스, 리프트계곡

열 바이러스, 로사바이러스(Rosavirus) A, 로스 강(Ross river) 바이러스, 로타바이러스 A, 로타바이러스 B, 로타바이러스 C, 풍진 바이러스, 사기야마(Sagiyama) 바이러스, 살리바이러스(Salivirus) A, 모래파리 열 시실리안(Sandfly fever sicilian) 바이러스, 삿포로(Sapporo) 바이러스, 셈리키 포레스트(Semliki forest) 바이러스, 서울(Seoul) 바이러스, 유인원 포미(Simian foamy) 바이러스, 유인원 바이러스 5, 신드비스(Sindbis) 바이러스, 사우샘프턴(Southampton) 바이러스, 세인트 루이스 뇌염(St. louis encephalitis) 바이러스, 틱-본 포와산(Tick-borne powassan) 바이러스, 토르크 테노 바이러스, 토스카나(Toscana) 바이러스, 우우쿠니에미(Uukuniemi) 바이러스, 백시니아 바이러스, 바리셀라-조스터(Varicella-zoster) 바이러스, 바리올라(Variola) 바이러스, 베네주엘란(Venezuelan) 말 뇌염 바이러스, 수포성 구내염 바이러스, 서부 말 뇌염 바이러스, WU 폴리오마바이러스, 웨스트 나일(West Nile) 바이러스, 야바(Yaba) 원숭이 종양 바이러스, 야바-유사 질병 바이러스, 황열 바이러스 및 지카 바이러스를 포함한다. 특정 구현예에서, 아넬로벡터는 대상체에 이미 존재하는 바이러스를 예를 들어, 기준에 비하여 적어도 약 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 또는 그 이상 능가하고/능가하거나 대체하기에 충분하다. 특정 구현예에서, 아넬로벡터는 만성 또는 급성 바이러스 감염과 경쟁하기에 충분하다. 특정 구현예에서, 아넬로벡터는 바이러스 감염으로부터 보호하기 위하여 예방적으로 투여될 수 있다(예를 들어, 바이러스전(provirotic)). 일부 구현예에서, 아넬로벡터는 (예를 들어, 표현형, 바이러스 수준, 유전자 발현, 다른 바이러스와의 경쟁, 질환 상태 등을 적어도 약 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 또는 그 이상) 조절하기에 충분한 양으로 존재한다. 일부 구현예에서, "치료", "치료하는" 및 이의 동족어는 (예를 들어, 아넬로벡터, 예를 들어, 본 명세서에 기재된 바와 같이 제조된 아넬로벡터를 투여함으로써) 예를 들어, 질환, 병리학적 병태, 또는 장애를 개선하거나, 호전시키거나, 안정화시키거나, 예방하거나 치유하기 위한 의도를 갖는 대상체의 의학적 관리를 포함한다. 일부 구현예에서, 치료는 적극 치료(질환, 병리학적 병태, 또는 장애를 개선하고자 하는 치료), 원인 치료(연관 질환, 병리학적 병태, 또는 장애의 원인을 대상으로 하는 치료), 완화 치료(증상의 경감을 위해 설계된 치료), 예방적 치료(연관 질환, 병리학적 병태, 또는 장애의 예방, 최소화, 또는 이의 발생을 부분적으로 또는 완전히 억제하고자 하는 치료), 및/또는 지지 치료(또 다른 치료법을 보충하기 위해 이용되는 치료)를 포함한다.

[1034] 본 명세서에 언급된 모든 참고문헌 및 간행물은 본 명세서에 참조로 포함된다.

[1035] 하기의 실시예는 본 발명의 일부 구현예를 추가로 예시하기 위해 제공되지만, 본 발명의 범주를 제한하고자 하지 않으며; 당업자에게 알려져 있는 다른 절차, 방법, 또는 기법이 대안적으로 사용될 수 있다는 것이 실시예의 예시적인 성질에 의해 이해될 것이다.

[1036] **실시예**

[1037] **목차**

[1038] 실시예 1: 아넬로바이러스 게놈의 탠덤 카피

[1039] 실시예 2: 텐담 아넬로벡터 작제물로부터 아넬로벡터의 효율적인 복제

[1040] 실시예 3: 예시적인 탠덤 아넬로벡터 작제물 설계

[1041] 실시예 4: 포유동물 세포에서 탠덤 아넬로바이러스 작제물로부터의 유전자 전사

[1042] 실시예 5: 포유동물 세포에서 탠덤 아넬로바이러스 작제물로부터 생성된 ORF1 및 ORF2 단백질

[1043] 실시예 6: 탠덤 아넬로벡터의 감염성 평가

[1044] 실시예 7: 바큘로바이러스를 통한 Sf9 곤충 세포로의 탠덤 아넬로바이러스 게놈의 운반

[1045] 실시예 8: 합성 아네로벡터의 제조

[1046] 실시예 9: 아넬로벡터의 조립 및 감염

[1047] 실시예 10: 아넬로벡터의 선택성

[1048] 실시예 11: 복제-결핍 아넬로벡터 및 헬퍼 바이러스

[1049] 실시예 12: 복제-적격 아넬로벡터의 제조 과정

[1050] 실시예 13: 복제-결핍 아넬로벡터의 제조 과정

[1051] 실시예 14: 현탁 세포를 사용한 아넬로벡터의 생성

[1052] 실시예 15: 마우스에서 외인성 단백질을 발현하기 위한 아넬로벡터의 이용

[1053] 실시예 16: 외인성 마이크로RNA 서열을 발현하는 아넬로벡터의 기능적 효과

[1054] 실시예 17: 외인성 비-코딩 RNA 발현을 위한 아넬로벡터의 제조 및 생성

[1055] 실시예 18: 아넬로벡터로부터 내인성 miRNA의 발현 및 내인성 miRNA의 결실

[1056] 실시예 19: 생체 내에서의 외인성 단백질의 아넬로벡터 운반

[1057] 실시예 20: 시험관 내 환화된 아넬로바이러스 게놈

[1058] 실시예 21: 상이한 토르크 테노 바이러스 주 유래의 초가변 도메인을 갖는 키메라 ORF1을 함유하는 아넬로벡터의 생성

[1059] 실시예 22: 초가변 도메인 대신에 비-TTV 단백질/펩티드를 함유하는 키메라 ORF1의 생성

[1060] 실시예 23: tth8 및 LY2에 기반한 아넬로벡터는 각각 EPO 유전자를 폐암 세포 내로 성공적으로 형질도입하였다

[1061] 실시예 24: 치료용 이식유전자를 갖는 아넬로벡터는 정맥내(i.v.) 투여 후 생체 내에서 검출될 수 있다

[1062] 실시예 25: 시험관 내에서 아넬로벡터를 생성하기 위한 투입 물질로서 시험관 내 환화된 게놈

[1063] **실시예 1: 아넬로바이러스 게놈의 탠덤 카피**

[1064] 본 실시예는 상류 게놈의 GC-풍부 영역이 하류 게놈의 5' 영역 근처에 있도록 나란히 배열된, 단일 아넬로바이러스 게놈의 2개 카피를 보유하는 플라스미드-기반 발현 벡터를 기재한다(도 1a).

[1065] 일부 구현예에서, 아넬로바이러스는 회전환을 통해 복제할 수 있으며, 여기에서 복제효소(Rep) 단백질은 아넬로바이러스 Rep 결합 부위(예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어, 5' UTR을 포함함, 예를 들어, 헤어핀 루프 및/또는 복제 원점을 포함함)에서 게놈에 결합하고 환 주위에서 DNA 합성을 개시한다. 플라스미드 백본에 함유된 아넬로바이러스 게놈의 경우, 이는 통상적으로, 천연 바이러스 게놈보다 더 긴 전체 플라스미드 길이의 복제, 또는 최소한의 백본을 가진 게놈을 포함하는 더 작은 환을 생성하는 플라스미드의 재조합을 수반한다. 따라서, 플라스미드를 벗어난 바이러스 복제는 비효율적일 수 있다. 바이러스 게놈 복제 효율을 개선하기 위해, 플라스미드를 TTMV-LY2의 탠덤 카피로 엔지니어링하였다. 이론에 결부시키고자 하지 않고, 이들 플라스미드는 Rep 단백질이 결합하는 위치에 관계없이 상류 아넬로바이러스 Rep 결합 부위로부터 하류 아넬로바이러스 Rep 변위 부위(예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어, 5' UTR을 포함함, 예를 들어, 헤어핀 루프 및/또는 복제 원점을 포함함)로 바이러스 게놈의 복제를 구동시킬 수 있도록, 아넬로바이러스 게놈의 환형 순열을 제시했을 수 있다.

[1066] 탠덤 TTMV-LY2를 Golden-gate 조립을 통해 조립하였고, 동시에 2개의 게놈 카피를 백본에 통합하고, 게놈 사이에 추가 뉴클레오티드를 남기지 않았다. 탠덤 TTMV-LY2 플라스미드는 첫 번째 5' NCR에서 시작하여 첫 번째 GC-풍부 영역까지 이어지고 바로 두 번째 5' NCR부터 두 번째 GC-풍부 영역까지 아넬로바이러스 게놈의 2개의 동일한 카피를 포함하였다(도 1a). 플라스미드는 또한 박테리아 기원 및 선택 가능한 마커를 갖는 박테리아 백본을 포함하였다.

[1067] TTMV-LY2의 탠덤 카피를 보유하는 플라스미드를 뉴클레오펙션을 통해 MOLT-4 세포로 형질감염시켰다. TTMV-LY2 게놈의 단일 카피를 갖는 플라스미드를 대조군으로서 유사하게 형질감염시켰다. 세포를 4일 동안 인큐베이션한 다음, 세포 펠렛을 수집하였다. 각각의 세포 펠렛의 일부를 서던 블롯팅에 사용하였다. 총 DNA를 Qiagen DNeasy 혈액 및 조직 키트를 사용하여 세포에서 단리하였다. TTMV-LY2 게놈 및 플라스미드에 대해 서로 다른 효과로 게놈 DNA를 분해하는 제한 엔도뉴클레아제를 사용하여, 각각 총 DNA 샘플의 10 μg에 대해 4가지의 대체 분해를 수행하였다: 한 가지 분해는 게놈 내에서 절단하지 않거나 플라스미드를 절단하지 않고; 두 번째 분해는 박테리아 백본 내의 단일에서 절단하지만, 아넬로바이러스 게놈을 절단하지 않고; 세 번째 분해는 TTMV-LY2 게놈 내의 단일 유전자좌를 절단하지만, 박테리아 백본 내에서 절단하지 않고; 마지막 분해는 TTMV-LY2 게놈 내에서 절단하지만 박테리아 백본 내에서 절단하지 않고, 또한 박테리아에서 생성된 유입 플라스미드 DNA만을 분해하고 포유동물 세포에서 복제된 DNA 내에서 절단하지 않을 메틸화-민감성 DpnI 효소를 포함하였다. 분해를 3시간 동안 0.5 V/cm으로 1×TAE에서 7 mm 두께의 1% 아가로스 겔에서 실행하였다. 그 다음 겔을 처리하여 DNA를 탈퓨린화하고 변성시켰다. 그 다음 DNA를 밤새 모세관 이동을 통해 양전하 나일론 막으로 옮겼다. DNA를 자외선을 통해 막에 가교시켰다. 그 다음 비오틴-dUTP를 프로브에 통합하여 TTMV-LY2 게놈에 대해 무작위-헥사머 생성 단편으

로 블롯을 프로브하였다. 프로브를 스트렙타비딘-접합 IRDye-800을 사용하여 검출하고, LiCor Odyssey 이미저에서 영상화하였다.

[1068] 서던 블롯팅은 탠덤 TTMV-LY2 플라스미드가 야생형 크기의 환형 이중-가닥 아넬로바이러스 게놈을 복제할 수 있음을 입증하였다(도 1b). TTMV-LY2 게놈의 단일 카피를 보유하는 플라스미드의 경우, 4 내지 10 kb 사이의 미절단 수퍼코일 DNA가 관찰되었으며(레인 1), 이는 플라스미드 백본(레인 2) 또는 TTMV- LY2 게놈(레인 3) 내에서 절단될 경우 5.1 kb로 선형화되었다. 환형 또는 선형의 회수된 야생형 길이 TTMV-LY2 게놈과 일치하는 밴드는 TTMV-LY2 게놈의 단일 카피를 가진 플라스미드로부터 관찰되지 않았다. 선형화된 플라스미드의 DpnI-내성 카피 분해에 의해 관찰된 바와 같이(레인 4), 단일 카피를 갖는 전체 플라스미드는 MOLT-4 세포에서 복제하였다. 그러나, 단일-카피 TTMV-LY2 플라스미드로부터 어떠한 야생형 길이 게놈도 회수되지 않았다.

[1069] TTMV-LY2 게놈의 탠덤 카피를 보유하는 플라스미드의 경우, 4 kb와 10 kb 사이의 슈퍼코일 플라스미드가 관찰되었으며(레인 5), 이는 플라스미드 백본에서 절단될 경우 8.8 kb로 선형화되었다(레인 6). 중요하게도, 야생형 TTMV-LY2 게놈(레인 5 및 6)의 회복과 일치하는 이중 가닥 DNA TTMV-LY2 게놈의 단일 카피와 일치하는 대략 1.8 kb 밴드가 미절단 및 백본 절단 레인에서 관찰되었다. 이는 TTMV-LY2 게놈 내에서 절단하는 효소로 분해되었을 때 1.8 kb 밴드가 선형화된 TTMV-LY2 게놈 DNA와 일치하는 3.0 kb 밴드로 대체되었다(레인 7). 이 선형화된 TTMV-LY2 게놈 밴드는 DpnI 내성이었으며, 이는 탠덤 DNA의 재조합(레인 8)을 통해 생성되는 것이 아니라 포유동물 세포 내에서 복제되었음을 나타낸다. 함께 이러한 데이터는 야생형 길이의 TTMV-LY2 게놈이 MOLT-4 세포의 탠덤 TTMV-LY2 플라스미드로부터 회수되었음을 입증하였다.

[1070] 탠덤 TTMV-LY2 플라스미드로 형질감염된 추가 세포 펠렛을 0.5% 트리톤의 존재 하에 동결/해동에 의해 용해시킨 다음, 패키징되지 않은 DNA로부터 바이러스 입자를 분리하기 위해 선형 CsCl 구배에서 실행하였다. 선형 구배로부터 분획을 취하고, TTMV-LY2 게놈 서열에 대해 Taqman 프로브를 사용하여 qPCR을 수행하였다. TTMV-LY2 게놈의 피크는 아넬로바이러스 크기의 입자가 발견될 것으로 예상되는 1.30 내지 1.35 g/cm^{3}의 CsCl 밀도에서 관찰되었다(도 1c). 이는 MOLT-4 세포에서 생성된 TTMV-LY2 게놈이 성공적으로 바이러스 입자로 패키징되었음을 나타내었다. 전반적으로, 이 데이터는 엔지니어링 탠덤 아넬로바이러스 게놈이 바이러스 게놈 복제를 증가시킬 수 있고 아넬로바이러스 생성을 증가시키기 위한 전략으로 사용될 수 있음을 입증하였다.

[1071] **실시예 2: 텐담 아넬로벡터 작제물로부터 아넬로벡터의 효율적인 복제**

[1072] 본 실시예에서, HEK293 또는 MOLT-4 세포와 같은 포유동물 숙주 세포에서의 증폭에 대해 탠덤 아넬로벡터를 분석한다. 탠덤 아넬로벡터 작제물은 아넬로바이러스 게놈의 2개의 전장 카피(예를 들어, 본 명세서에 기재된 바와 같은, Ring1, Ring2, 또는 Ring4)를 포함하도록 구축한다. 게놈의 각각의 카피는 5'에서 3' 순서로, 고도로 보존된 도메인을 포함하는 5' 비-코딩 영역, 천연 아넬로바이러스 오픈-리딩 프레임을 대체하는 카고 서열을 포함하는 영역, 및 GC-풍부 영역을 포함하는 3' UTR을 포함하였다. 제1 게놈 카피의 3' 말단과 제2 게놈 카피의 5' 말단을 뉴클레오티드 개입 없이 서로 직접 부착한다.

[1073] 간략하게, 작제물을 PEI 형질감염 시약 또는 뉴클레오펙션에 의해 HEK293 또는 MOLT-4 세포 내로 도입한다. 아넬로바이러스 ORF1을 포함한 트랜스 복제 및 패키징 요소를 별도의 플라스미드에서 트랜스로 제공한다. 형질감염된 세포를 37℃에서 4일 동안 인큐베이션한다. 아넬로바이러스 게놈의 복제를 qPCR 및 서던 블롯으로 측정한다. 음성 대조군의 경우, 트랜스 요소 없이 아넬로벡터 및 탠덤 아넬로벡터의 단일 카피를 보유하는 플라스미드를 포함하였다.

[1074] **실시예 3: 예시적인 탠덤 아넬로벡터 작제물 설계**

[1075] 하기 기재된 실시예에서, 탠덤 아넬로바이러스에 대한 다수의 예시적인 작제물 설계를 MOLT-4 숙주 세포에서 회전환 증폭을 겪는 능력에 대해 테스트하였다. 이론에 결부시키고자 하지 않고, 아넬로바이러스 회전환 증폭은 복제효소-결합 부위(예를 들어, 헤어핀 루프 및/또는 복제 원점을 포함하는, 예를 들어, 5' UTR)에서 시작되고 끝나는 것으로 고려된다. 환화된 단일 아넬로바이러스 게놈에서, 동일한 복제효소-결합 부위는 시작 부위와 중지 부위 둘 다로 작용할 수 있다. 탠덤 아넬로바이러스는뿐만 아니라, 본 실시예에 기재된 대안 설계는, 게놈이 환화된 단일-카피 게놈처럼 효과적으로 작동하도록, 복제될 게놈의 양쪽 말단에 이러한 복제효소-결합 부위를 배치한다.

[1076] *3' 말단에 부분적인 아넬로바이러스 게놈을 갖는 작제물*

[1077] 본 실시예에서, 아넬로바이러스 게놈의 전장 카피가 부분적인 아넬로바이러스 게놈에 비해 5'에 위치하는 예시

적인 탠덤 아넬로벡터를 설계하였다. 도 2a에 나타낸 바와 같이, 제1 대안 작제물(pRTx-843)은 5'에서 3' 순서로, 전장 카피의 아넬로바이러스 게놈(Ring2), 이어서 5' NCR로 이루어진 부분적인 아넬로바이러스 게놈, 바이러스 오픈-리딩 프레임의 전체 세트를 포함하는 영역, 및 GC-풍부 영역이 결여된 3' NCR을 포함하였다. 도 2a에 나타낸 바와 같이, 제2 대안 작제물(pRTx-844)은 5'에서 3' 순서로, 전장 카피의 아넬로바이러스 게놈(Ring2), 이어서 5' NCR로 이루어진 부분적인 아넬로바이러스 게놈, 및 Ring2의 뉴클레오티드 1에서 2812의 바이러스 오픈-리딩 프레임의 전체 세트를 포함하는 영역을 포함하였다. 도 2a에 나타낸 바와 같이, 제3 대안 작제물(pRTx-845)은 5'에서 3' 순서로, 전장 카피의 아넬로바이러스 게놈(Ring2), 이어서 5' NCR로 이루어진 부분적인 아넬로바이러스 게놈, 및 Ring2의 뉴클레오티드 1에서 2583의 바이러스 오픈-리딩 프레임의 일부만을 포함하는 영역을 포함하였다. 도 2a에 나타낸 바와 같이, 제4 대안 작제물(pRTx-846)은 5'에서 3' 순서로, 전장 카피의 아넬로바이러스 게놈(Ring2), 이어서 5' NCR로 이루어진 부분적인 아넬로바이러스 게놈, 및 Ring2의 뉴클레오티드 1에서 2264의 바이러스 오픈-리딩 프레임의 일부만을 포함하는 영역을 포함하였다. 도 2a에 나타낸 바와 같이, 제5 대안 작제물(pRTx-847)은 5'에서 3' 순서로, 전장 카피의 아넬로바이러스 게놈(Ring2), 이어서 5' NCR로 이루어진 부분적인 아넬로바이러스 게놈, 및 Ring2의 뉴클레오티드 1에서 723의 바이러스 오픈-리딩 프레임의 일부만을 포함하는 영역을 포함하였다. 도 2a에 나타낸 바와 같이, 제6 대안 작제물(pRTx-848)은 5'에서 3' 순서로, 전장 카피의 아넬로바이러스 게놈(Ring2), 이어서 5' NCR로 이루어진 부분적인 아넬로바이러스 게놈, 및 Ring2의 뉴클레오티드 1에서 423을 포함하였다. 도 2a에 나타낸 바와 같이, 제7 대안 작제물(pRTx-849)은 5'에서 3' 순서로, 전장 카피의 아넬로바이러스 게놈(Ring2), 이어서 5' NCR의 일부로 이루어진 부분적인 아넬로바이러스 게놈, 및 Ring2의 뉴클레오티드 1에서 267을 포함하였다.

[1078] 간략하게, 각각의 탠덤 작제물을 뉴클레오펙션에 의해 MOLT-4 세포에 도입하였다. 회전환 증폭을 위한 복제효소 단백질을 완전한 바이러스 게놈에 의해 시스로 제공하였다. ORF1 단백질을 완전한 바이러스 게놈에 의해 시스로 제공하였다.

[1079] 2개의 전체 게놈(pVL46-257)을 포함하는 전장 탠덤 Ring2 작제물을 바이러스 복제 및 패키징을 위한 양성 대조군으로 사용하였다. 음성 대조군의 경우, Ring2 게놈의 단일 카피를 보유하는 플라스미드(pVL46-240)가 사용된다. 형질감염된 세포를 37°에서 4일 동안 인큐베이션한 다음, 서던 블롯 및 qPCR 분석을 위해 세포를 수확하였다. 서던 블롯의 경우, Qiagen DNeasy 혈액 및 조직 키트를 사용하여 세포에서 전체 DNA를 단리하고, 플라스미드 백본에서 한 번 절단하는 효소와 DpnI로 총 DNA 10 μg을 분해하여 박테리아에서 생성된 임의의 투입 DNA를 분해하였다. 분해를 3시간 동안 0.5 V/cm으로 1×TAE에서 7 mm 두께의 1% 아가로스 겔에서 실행하였다. 그 다음 겔을 처리하여 DNA를 탈퓨린화하고 변성시켰다. 그 다음 DNA를 밤새 모세관 이동을 통해 양전하 나일론 막으로 옮겼다. DNA를 자외선을 통해 막에 가교시켰다. 그 다음 비오틴-dUTP를 프로브에 통합하여 TTMV-LY2 게놈에 대해 무작위-헥사머 생성 단편으로 블롯을 프로브하였다. 프로브를 스트렙타비딘-접합 IRDye-800을 사용하여 검출하고, LiCor Odyssey 이미저에서 영상화하였다. 플라스미드 pRTx-845의 샘플을 서던 블롯으로 테스트하지 않았음에 유의한다. 복제된 환형 이중-가닥 DNA Ring2 게놈의 회수는 pRTx-843 및 844의 경우 관찰되었지만, pRTx-846 내지 849에 대해서는 관찰되지 않았다(도 2d). pRTx-843, 844 및 848의 경우 플라스미드 DNA의 복제가 관찰되었으며, 이는 단일-카피 게놈 플라스미드에서 관찰되는 것과 유사하였다.

[1080] 추가 세포 펠렛을 동결/해동 및 0.5% 트리톤을 사용하여 용해하였다. 용해물을 염화세슘 단계 구배에 통과시키고 아넬로바이러스-함유 분획을 수집하였다. 아넬로바이러스 게놈의 복제를 DNase-보호 qPCR에 의해 측정하였다. pRTx-843 내지 846은 전체 탠덤 pRTx-257에서 관찰된 바와 같이 세포당 유사한 수준의 Ring2 바이러스 게놈을 생성하였으며, 이는 캡슐화된 바이러스의 성공적인 생성을 나타낸다(도 2e). pRTx-847은 또한 전체 탠덤에서 관찰된 것보다 적지만 보호된 게놈을 생성한 반면, pRTx-848 및 849는 qPCR에 의해 테스트되지 않았다.

[1081] *5' 말단에 부분적인 아넬로바이러스 게놈을 갖는 작제물*

[1082] 본 실시예에서, 아넬로바이러스 게놈의 전장 카피가 부분적인 아넬로바이러스 게놈에 비해 3'에 위치하는 예시적인 탠덤 아넬로벡터를 설계하였다. 도 2b에 나타낸 바와 같이, 하기와 같은 부분적인 Ring 게놈에 이어 전장 Ring2 게놈을 갖는 일련의 작제물을 테스트하였다: pRTx-836, 고도로 보존된 5' NCR 도메인, 아넬로바이러스 오픈-리딩 프레임의 전체 세트, 및 GC-풍부 영역(Ring2 뉴클레오티드 267 내지 2979)을 포함하는 3' NCR로 이루어진 부분적인 아넬로바이러스 게놈을 포함함; pRTx-837, 아넬로바이러스 오픈-리딩 프레임의 전체 세트 및 GC-풍부 영역(Ring2 뉴클레오티드 423 내지 2979)을 포함하는 3' NCR로 이루어진 부분적인 아넬로바이러스 게놈을 포함함; pRTx-838, 아넬로바이러스 오픈-리딩 프레임의 일부 및 GC-풍부 영역을 포함하는 3' NCR로 이루어진 부분적인 아넬로바이러스 게놈(Ring2 뉴클레오티드 723 내지 2979); pRTx-839, 아넬로바이러스 오픈-리딩 프레임의

일부 및 GC-리치 영역(Ring2 뉴클레오티드 723 내지 2979)을 포함하는 3' NCR로 이루어진 부분적인 아넬로바이러스 게놈을 포함함; pRTx-840, 아넬로바이러스 오픈-리딩 프레임의 일부 및 GC-풍부 영역(Ring2 뉴클레오티드 2812 내지 2979)을 포함하는 3' NCR로 이루어진 부분적인 아넬로바이러스 게놈을 포함함; pRTx-841, GC-풍부 영역(Ring2 뉴클레오티드 2812 내지 2979)을 포함하는 3' NCR로 이루어진 부분 아넬로바이러스 게놈을 포함함; 및 pRTx-842, GC-풍부 영역(Ring2 뉴클레오티드 2867 내지 2979)으로 이루어진 부분적인 아넬로바이러스 게놈을 포함함.

[1083] 간락하게, 각각의 탠덤 작제물을 뉴클레오펙션에 의해 MOLT-4 세포 내로 도입하였다. 회전환 증폭을 위한 복제 효소 단백질을 완전한 바이러스 게놈에 의해 시스로 제공하였다. ORF1 단백질을 완전한 바이러스 게놈에 의해 시스로 제공하였다. 2개의 전체 게놈(pVL46-257)을 포함하는 전장 탠덤 Ring2 작제물을 바이러스 복제 및 패키징을 위한 양성 대조군으로 사용하였다. 음성 대조군의 경우, Ring2 게놈의 단일 카피를 보유하는 플라스미드(pVL46-240)가 사용된다. 형질감염된 세포를 37°에서 4일 동안 인큐베이션한 다음, 서던 블롯 및 qPCR 분석을 위해 세포를 수확하였다. 서던 블롯의 경우, Qiagen DNeasy 혈액 및 조직 키트를 사용하여 세포에서 전체 DNA를 단리하고, 플라스미드 백본에서 한 번 절단하는 효소와 DpnI로 총 DNA 10 μg을 분해하여 박테리아에서 생성된 임의의 투입 DNA를 분해하였다. 분해를 3시간 동안 0.5 V/cm으로 1×TAE에서 7 mm 두께의 1% 아가로스 겔에서 실행하였다. 그 다음 겔을 처리하여 DNA를 탈퓨린화하고 변성시켰다. 그 다음 DNA를 밤새 모세관 이동을 통해 양전하 나일론 막으로 옮겼다. DNA를 자외선을 통해 막에 가교시켰다. 그 다음 비오틴-dUTP를 프로브에 통합하여 TTMV-LY2 게놈에 대해 무작위-헥사머 생성 단편으로 블롯을 프로브하였다. 프로브를 스트렙타비딘-접합 IRDye-800을 사용하여 검출하고, LiCor Odyssey 이미저에서 영상화하였다. 복제된 환형 이중-가닥 DNA Ring2 게놈의 회수는 pRTx-836 내지 839의 경우 관찰되었지만, pRTx-840 내지 842에 대해서는 관찰되지 않았다(도 2d).

[1084] 추가 세포 펠렛을 동결/해동 및 0.5% 트리톤을 사용하여 용해하였다. 용해물을 염화세슘 단계 구배에 통과시키고 아넬로바이러스-함유 분획을 수집하였다. 아넬로바이러스 게놈의 복제를 DNase-보호 qPCR에 의해 측정하였다. pRTx-836 내지 840은 전체 탠덤 pRTx-257에서 관찰된 바와 같이 세포당 유사한 수준의 Ring2 바이러스 게놈을 생성하였으며, 이는 캡슐화된 바이러스의 성공적인 생성을 나타낸다(도 2e). pRTx-841 및 842에 대해서는 보호된 바이러스 게놈이 거의 내지 전혀 관찰되지 않았다.

[1085] *2개의 부분적인 아넬로바이러스 게놈을 갖는 작제물*

[1086] 본 실시예에서, 탠덤 구조의 구조를 충분히 모방하여 효율적인 회전환 증폭을 허용하도록 배열된, 아넬로바이러스 게놈의 2개의 부분적인 카피를 포함하는 예시적인 탠덤 아넬로벡터를 설계하였다. 6개의 이러한 순열이 도 2c에 나타나 있다: 5'에서 3'으로, 전체 Ring2 오픈-리딩 프레임 및 GC-풍부 영역(Ring2 뉴클레오티드 267 내지 2979)을 포함하는 3' NCR을 갖는 5' NCR 보존된 영역에서 시작하는 부분적인 Ring2 게놈, 이어서 5' NCR 및 고도로 보존된 영역(Ring 뉴클레오티드 1 내지 423)을 갖는 부분적인 Ring2 게놈을 포함하는 순열 1; 5'에서 3'으로, 전체 Ring2 오픈-리딩 프레임 및 GC-풍부 영역(Ring2 뉴클레오티드 423 내지 2979)을 포함하는 3' NCR로 시작하는 부분적인 Ring2 게놈, 이어서 고도로 보존된 영역을 갖는 5' NCR 및 오픈-리딩 프레임의 일부(Ring2 뉴클레오티드 1 내지 723)를 갖는 부분적인 Ring2 게놈을 포함하는 순열 2; 5'에서 3'으로, Ring2 오픈-리딩 프레임의 일부 및 GC-풍부 영역(Ring2 뉴클레오티드 723 내지 2979)을 포함하는 3' NCR로 시작하는 부분적인 Ring2 게놈, 이어서 5' NCR 및 아넬로바이러스 오픈-리딩 프레임의 일부(Ring2 뉴클레오티드 1 내지 2273)를 갖는 부분적인 Ring2 게놈을 포함하는 순열 3; 5'에서 3'으로, 부분적인 Ring2 오픈-리딩 프레임 및 GC-풍부 영역(Ring2 뉴클레오티드 2273 내지 2979)을 포함하는 3' NCR로 시작하는 부분적인 Ring2 게놈, 이어서 5' NCR 및 아넬로바이러스 오픈-리딩 프레임의 일부(Ring2 뉴클레오티드 1 내지 2452)를 갖는 부분적인 Ring2 게놈을 포함하는 순열 4; 5'에서 3'으로, 부분적인 Ring2 오픈-리딩 프레임 및 GC-풍부 영역(Ring2 뉴클레오티드 2452 내지 2979)을 포함하는 3' NCR로 시작하는 부분적인 Ring2 게놈, 이어서 5' NCR 및 전체 Ring2 오픈-리딩 프레임(Ring2 뉴클레오티드 1 내지 2812)를 갖는 부분적인 Ring2 게놈을 포함하는 순열 5; 및 5'에서 3'으로, GC-풍부 영역(Ring2 뉴클레오티드 2812 내지 2979)을 포함하는 3' NCR에서 시작하는 부분적인 Ring 2 게놈, 이어서 5' NCR 및 전체 Ring2 오픈-리딩 프레임 및 GC-풍부 영역(Ring2 뉴클레오티드 1 내지 2867)이 없는 3' NCR을 갖는 부분적인 Ring2 게놈을 포함하는 순열 6.

[1087] 간락하게, 각각의 탠덤 작제물을 뉴클레오펙션에 의해 MOLT-4 세포 내로 도입한다. Rep 인자 및 Ring2 ORF1을 포함하여, 회전환 증폭 및 바이러스 패키징을 위한 단백질을 다른 플라스미드로부터 트랜스로 제공한다. 형질감염된 세포를 37°에서 4일 동안 인큐베이션한다. 아넬로바이러스 게놈의 복제를 qPCR 및 서던 블롯에 의해 측정한다. 2개의 전체 게놈을 갖는 전장 탠덤 Ring2 작제물(pVL46-257) 을 바이러스 복제 및 패키징을 위한 양성 대조군으로 사용한다. 음성 대조군의 경우, Ring2 게놈의 단일 카피를 보유하는 플라스미드(pVL46-240)가 사용된

다.

[1088] **실시예 4: 포유동물 세포에서 탠덤 아넬로바이러스 작제물로부터 유전자의 전사**

[1089] 본 실시예에서, 백본으로서 Ring1을 기반으로 하여 일련의 아넬로벡터 작제물을 생성하였다(도 2f에 나타낸 바와 같음). 작제물은 eGFP-ORF1 융합 단백질(최적화된 코돈)을 암호화하는 Ring1 서열 및 탠덤 Ring1 서열을 포함하는 탠덤 작제물을 포함하였다. 그 다음 작제물을 Jurkat 세포에 형질감염시켰다. 그 다음 아넬로바이러스(Ring1) ORF1의 전사를 긴 RNA 판독물의 서열을 결정함으로써 평가하였다.

[1090] 도 2f에 나타낸 바와 같이, 대안 작제물로 형질감염된 Jurkat 세포와 비교하여 Ring1-기반 탠덤 GFP 작제물로 형질감염된 Jurkat 세포에서 더 많은 양의 전장 Ring1 ORF1 전사체를 검출하였다.

[1091] **실시예 5: 포유동물 세포에서 탠덤 아넬로바이러스 작제물로부터 생성된 ORF1 및 ORF2 단백질**

[1092] 본 실시예에서, 백본으로서 아넬로바이러스 Ring2를 기반으로 하여 일련의 아넬로벡터 작제물을 생성하였다(도 2g에 나타낸 바와 같음). 작제물은 제 Ring2 서열 및 제2 Ring2 서열을 나란히 포함하는 탠덤 작제물을 포함하였다. 작제물을 MOLT4 세포(인간 T 림프아구 세포주)에 형질감염시키고 그 다음 Ring2 ORF1 단백질을 웨스턴 블롯에 의해 검출하였다. 간략하게, 탠덤 Ring2 게놈(Rep)을 함유하는 플라스미드 또는 Ring2 게놈의 149 bp를 함유하는 음성 대조군 플라스미드 중 하나의 25 ug으로 1E07 MOLT4 세포를 뉴클레오펙션시켰다. 각각의 뉴클레오펙션된 샘플을 25 ml 성장 배지(RPMI + 10% FBS + 0.01% Polyaxmer + 1 mM 피루브산나트륨)에 접종하였다. 뉴클레오펙션 후 제1일부터 제3일까지 매일 각각의 샘플에서 배양액 1 ml를 펠렛화하였다. 세포를 50 ul 용해 완충액(0.5% 트리톤, 300 mM NaCl, 50 mM 트리스 pH 8.0)에 재현탁한 후 2회 동결 해동하여 펠렛화된 세포를 용해하였다. 그 다음 용해물을 10,000×g에서 30분 동안 회전시켜 정화하였다. 20 ul의 정화된 용해물을 웨스턴 블롯 분석에 사용하여 Ring2 ORF1에 대해 생성된 2개의 토끼 폴리클로날 항체의 칵테일을 사용하여 Ring2 ORF1 단백질을 검출하였다.

[1093] 도 2g에 나타낸 바와 같이, 뉴클레오펙션 후 제2일 및 제3일에 Ring2-기반 탠덤 GFP 작제물로 뉴클레오펙션된 MOLT-4 세포에서 Ring2 ORF1 단백질을 검출하였다.

[1094] **실시예 6: 탠덤 아넬로벡터의 감염성 평가**

[1095] 본 실시예에서는, 탠덤 아넬로벡터를 외인성 유전자를 인코딩하는 유전 요소를 캡슐화하는 단백질성 외부로 생성한다. 실시예 1 내지 4 중 임의의 것에 기재된 바와 같이, 탠덤 아넬로벡터를 생성한다. 간단히 말해서, 숙주 세포를 탠덤 아넬로벡터 DNA로 형질감염시키고 탠덤 아넬로벡터 유전 요소의 복제 및 단백질성 외부 내의 캡슐화에 적합한 조건 하에서 인큐베이션한다. 그 다음 캡슐화된 아넬로벡터를, 예를 들어, 본 명세서에 기재된 바와 같이, 배양물로부터 단리한다. 그 다음 아넬로벡터를 세포의 감염에 적합한 조건 하에서 세포(예를 들어, MOLT-4 또는 Jurkat 세포)와 접촉시킨다.

[1096] 예를 들어, 감염된 세포에서 아넬로바이러스 핵산을 검출하기 위해 정량적 실시간 PCR(qPCR)을 사용하여 감염성을 평가할 수 있다. 예를 들어, 감염된 세포를 DNA를 위해 수확할 수 있으며, 그 다음 아넬로바이러스-특이적 서열에 특이적인 프라이머를 사용하여 qPCR을 수행한다. 예를 들어, GAPDH의 게놈 DNA 서열에 특이적인 프라이머에 대한 qPCR이 정규화에 사용될 수 있다. qPCR은 검출된 아넬로바이러스 DNA의 게놈 등가물에 따라 감염성을 정량화하는 데 사용될 수 있다.

[1097] 대안적으로, 외인성 유전자의 발현 또는 외인성 유전자의 하류 활성을 검출함으로써 감염성을 평가할 수 있다. 예를 들어, GFP 또는 나노-루시퍼라제와 같은 외인성 형광 마커는, 예를 들어 형광을 검출하거나 마커를 인식하는 항체를 사용하는 면역분석에 의해 검출될 수 있다.

[1098] **실시예 7: 바큘로바이러스를 통한 Sf9 곤충 세포로의 탠덤 아넬로바이러스 게놈의 운반**

[1099] 본 실시예에서, Ring2 게놈의 탠덤 카피를 보유하는 바큘로바이러스를 제조하고 Sf9 세포로 운반하였다. 탠덤 Ring2 게놈을 상기 기재한 바와 같이 조립하였다. 유형 IIS 제한 부위를 추가하는 PCR을 통해 전장 Ring2 게놈을 증폭하고 이를 박테리아 복제 원점 및 선택 가능한 마커를 갖는 플라스미드 백본에 골든 게이트 조립을 통해 삽입하였다. 생성된 플라스미드는 5' 비-코딩 영역에서 GC-풍부 영역까지의 제1 게놈과, 이어서 5' 비-코딩 영역에서 GC-영역까지의 제2 게놈으로 배열된 2개의 완전한 Ring2 게놈을 개재 뉴클레오티드 없이 서로 다음에 포함하였다. 플라스미드 백본에서 AsiSI 및 PacI 제한 효소 부위가 게놈 쌍에 측접하였다.

[1100] 탠덤 Rong2 게놈을 바큘로바이러스에 삽입하기 위해, 먼저 변형된 pFastBac을 조립하였다. 변형된 pFastBac에서

곤충-세포 프로모터를 제거하고, 프로모터 및 표준 다중 클로닝 부위를 AsiSI 및 PacI 부위를 함유하는 맞춤형 다중 클로닝 부위로 대체하였다. 탠덤 Ring2 게놈 작제물을 AsiSI 및 PacI에 의한 분해를 통해 pFastBac 플라스미드에 클로닝한 다음, 결찰을 수행하였다. 최종 pFastBac-TandemRing2 플라스미드는 Tn7L 재조합 서열, 탠덤 Ring2 게놈, 젠타마이신 내성 유전자, 및 Tn7R 재조합 서열을 포함한 다음, 박테리아 복제 원점 및 암피실린-내성 마커를 갖는 플라스미드 백본을 포함하였다(도 2h). 탠덤 Ring2 게놈의 포함을 서열결정 및 PCR 생성물 분석에 의해 확인하였다. pFastBac을 사용하여 탠덤 Ring2 게놈을 보유하는 박미드를 생성한 다음, 상기 기재된 바와 같이 바큘로바이러스를 생성하였다.

[1101] 탠덤 Ring2 게놈을 보유하는 바큘로바이러스를 사용하여 MOI 1에서 Sf9 세포를 감염시켰다. 또한, 샘플을 Ring2 ORF1-발현 바큘로바이러스 단독으로 감염되거나 Ring2 탠덤 게놈 바큘로바이러스 및 Ring2 ORF1-발현 바큘로바이러스로 동시-감염된 Sf9 세포에 포함하였다. 3일 후, Sf9 세포를 원심분리에 의해 펠렛화하였다. Qiagen DNeasy 혈액 및 조직 키트를 사용하여 세포에서 전체 DNA를 수확하였다. 탠덤 Ring2 게놈에 바로 측접하는 바큘로바이러스 내에서 절단하는 Esp3I 제한 효소로 총 DNA 10 μg을 분해하였다(도 2i 참조). 분해된 DNA를 아가로스 겔에서 실행하였다. 그 다음 DNA를 화학적으로 변성시키고 탈퓨린화한 다음, 모세관 이동에 의해 양전하 나일론 막으로 옮겼다. DNA를 자외선에 의해 막에 가교시킨 다음, Ring2 게놈에 대해 설계된 비오틴-함유 프로브와 혼성화하였다. 프로브를 스트렙타비딘-IRDye800으로 검출하고, LiCor Odyssey 이미저에서 영상화하였다.

[1102] 탠덤 Ring2 게놈 크기와 일치하는 밴드가 탠덤 Ring2 배큘로바이러스로 감염된 모든 샘플에서 관찰되었으며, 이는 Sf9 세포로 탠덤 Ring2 게놈의 성공적인 운반을 입증한다(도 2i). 또한, 바큘로바이러스로부터 단리된 Ring2 게놈의 단일 카피와 일치하는 밴드가 관찰되었으며, 이는 바큘로바이러스 생성 중에 일부 DNA 재조합이 발생하여 바큘로바이러스 집단의 일부에서 Ring2 게놈의 하나의 카피가 손실되었음을 나타낸다. 바큘로바이러스의 대략 50%는 탠덤 카피가 아닌 단일 카피 Ring2 게놈을 나타내었다. 환형 Ring2 게놈은 바큘로바이러스로부터 검출되지 않았다(MOLT-4 세포 내로 도입된 탠덤 Ring2 작제물과 대조적으로, 여기서 환형 단일-카피 dsDNA 게놈이 검출됨; 도 2i). 그러나, 이 재조합은 탠덤 게놈 카피의 SF9 세포로의 성공적인 운반을 억제하지 않았다.

[1103] **실시예 8: 합성 아넬로벡터의 제조**

[1104] 본 실시예는 합성 아넬로벡터의 시험관 내 생성을 입증한다.

[1105] EcoRV 제한 효소 부위 사이의 TTMiniV의 LY1과 LY2 균주 유래의 DNA 서열(Eur Respir J. 2013 Aug;42(2):470-9)을 카나마이신 벡터(Integrated DNA Technologies) 내로 클로닝하였다. TTMiniV의 LY1과 LY2 균주 유래의 DNA 서열에 기반한 생성된 유전 요소 작제물을 실시예 6 및 7에서 각각 아넬로벡터 1(아넬로 1) 및 아넬로벡터 2(아넬로 2)로 지칭한다. 클로닝된 작제물을 10-베타 적격 이.콜라이(New England Biolabs Inc.)로 형질전환한 다음, 제조업체의 프로토콜에 따라 플라스미드 정제(Qiagen)를 수행하였다.

[1106] DNA 작제물(도 3 및 도 4)를 37℃에서 6시간 동안 EcoRV 제한 분해(New England Biolabs, Inc.)를 이용하여 선형화하여, TTMiniV 게놈을 함유하고 박테리아 백본 요소(예컨대, 복제 원점 및 선택 가능한 마커)를 배제한 이중-가닥 선형 DNA 단편을 산출하였다. 이어서 아가로스 겔 전기영동, TTMiniV 게놈 단편(2.9 킬로염기쌍)에 대한 올바른 크기 DNA의 절제, 및 제조업체의 프로토콜에 따른 겔 추출 키트(Qiagen)를 사용한 절제된 아가로스 밴드로부터 DNA의 겔 정제를 수행하였다.

[1107] **실시예 9: 아넬로벡터의 조립 및 감염**

[1108] 본 실시예에서는 실시예 5에 기재된 바와 같은 합성 DNA 서열을 사용한 감염성 아넬로벡터의 성공적인 시험관 내 생성이 입증된다.

[1109] 이중-가닥 선형화된 겔-정제된 아넬로바이러스 게놈 DNA(실시예 5에서 수득됨)를 지질 형질감염 시약(Thermo Fisher Scientific)을 사용하여, 온전한 플라스미드에서 또는 선형화된 형태에서, HEK293T 세포(인간 배아 신장 세포주) 또는 A549 세포(인간 폐 암종 세포주) 중 어느 하나로 형질감염시켰다. T25 플라스크 내의 70% 컨플루언트 세포의 형질감염을 위하여 6 ㎍의 플라스미드 또는 1.5 ㎍의 선형화된 아넬로바이러스 게놈 DNA를 사용하였다. 아넬로벡터에 포함된 바이러스 서열이 결여된 빈 벡터 백본을 음성 대조군으로서 사용하였다. 형질감염 후 6시간에, 세포를 PBS로 2회 세척하고, 37℃ 및 5% 이산화탄소에서 신선한 성장 배지에서 성장하게 하였다. 인간 Eflalpha 프로모터에 이어서 YFP 유전자를 인코딩하는 DNA 서열을 IDT로부터 합성하였다. 이러한 DNA 서열을 클로닝 벡터(Thermo Fisher Scientific) 내로 블런트 말단 결찰시켰다. 생성된 벡터를 형질감염 효율을 평가하기 위한 대조군으로서 사용하였다. YFP를 형질감염 72시간 후에 세포 영상화 시스템(Thermo Fisher Scientific)을 사용하여 검출하였다. HEK293T 및 A549 세포의 형질감염 효율은 각각 85% 및 40%로 계산되었다

(도 5).

[1110] 아넬로벡터로 형질감염된 293T 및 A549 세포의 상청액을 형질감염 후 96시간에 수집하였다. 수집된 상청액을 2000 rpm에서 4℃에서 10분 동안 회전 침강시켜, 임의의 세포 데브리스를 제거하였다. 수집된 상청액의 각각을 사용하여 24 웰 플레이트 내의 웰에서 70% 컨플루언트인 새로운 293T 및 A549 세포를 각각 감염시켰다. 37℃ 및 5% 이산화탄소에서의 24시간의 인큐베이션 후에 상청액을 세척해낸 후, 2회의 PBS의 세척 및 신선한 성장 배지로 대체를 행하였다. 이들 세포를 37℃ 및 5% 이산화탄소에서 또 다시 48시간 동안 인큐베이션한 후에, 게놈 DNA 추출을 위하여 세포를 개별적으로 수집하였다. 각각의 샘플로부터의 게놈 DNA를 제조업체의 프로토콜에 따라 게놈 DNA 추출 키트(Thermo Fisher Scientific)를 사용하여 수집하였다.

[1111] 시험관 내 생성된 아넬로벡터에 의한 293T 및 A549 세포의 성공적인 감염을 확인하기 위하여, 본 명세서에 기재된 바와 같이 수집된 100 ng의 게놈 DNA를 사용하여, LY2 특이적 서열 또는 베타-토르크바이러스에 대하여 특이적인 프라이머를 사용한 정량적 중합효소 연쇄 반응(qPCR)을 수행하였다. 제조업체의 프로토콜에 따라 SYBR 그린 시약(Thermo Fisher Scientific)을 사용하여 qPCR을 수행하였다. GAPDH의 게놈 DNA 서열에 대하여 특이적인 프라이머에 대한 qPCR을 정규화를 위해 사용하였다. 사용되는 모든 프라이머에 대한 서열은 표 42에 열거되어 있다.

[1112] **[표 42]**

[1113]

	프라이머 서열 (5' > 3')	
표적	**정방향**	**역방향**
베타토르크바이러스	ATTCGAATGGCTGAGTTTATGC (서열번호 690)	CCTTGACTACGGTGGTTTCAC (서열번호 693)
LY2 TTMiniV 균주	CACGAATTAGCCAAGACTGGGCAC (서열번호 691)	TGCAGGCATTCGAGGGCTTGTT (서열번호 694)
GAPDH	GCTCCCACTCCTGATTTCTG (서열번호 692)	TTTAACCCCCTAGTCCCAGG (서열번호 695)

[1114] 도 6a, 6b, 7a, 및 7b에 도시된 qPCR 결과에 나타낸 바와 같이, 시험관 내에서 생성되고 본 실시예에 기재된 바와 같은 아넬로벡터는 감염성이었다.

[1115] **실시예 10: 아넬로벡터의 선택성**

[1116] 본 실시예에서는 시험관 내에서 생성된 합성 아넬로벡터가 다양한 조직 기원의 세포주를 감염시키는 능력이 입증된다.

[1117] 감염성 TTMiniV 아넬로벡터(실시예 5에 기재됨)을 갖는 상청액을 24 웰 플레이트의 웰 내에서 37℃ 및 5% 이산화탄소에서 70% 컨플루언트 293T, A549, Jurkat(급성 T 세포 백혈병 세포주), Raji(버킷 림프종 B 세포주) 및 Chang 세포주와 인큐베이션하였다. 세포를 감염 24시간 후에 PBS로 2회 세척한 후, 신선한 성장 배지로 교체하였다. 그 다음, 세포를 다시 또 다른 48시간 동안 37℃ 및 5% 이산화탄소에서 인큐베이션한 후, 게놈 DNA 추출을 위한 수집을 행하였다. 각각의 샘플로부터의 게놈 DNA를 제조업체의 프로토콜에 따라 게놈 DNA 추출 키트(Thermo Fisher Scientific)를 사용하여 수집하였다.

[1118] 이전의 실시예에서 생성된 아넬로벡터에 의한 이들 세포주의 성공적인 감염을 확인하기 위하여, 본 명세서에 기재된 바와 같이 수집된 100 ng의 게놈 DNA를 사용하여, LY2 특이적 서열 또는 베타-토르크바이러스에 대하여 특이적인 프라이머를 사용한 정량적 중합효소 연쇄 반응(qPCR)을 수행하였다. 제조업체의 프로토콜에 따라 SYBR 그린 시약(Thermo Fisher Scientific)을 사용하여 qPCR을 수행하였다. GAPDH의 게놈 DNA 서열에 대하여 특이적인 프라이머에 대한 qPCR을 정규화를 위해 사용하였다. 사용되는 모든 프라이머에 대한 서열은 표 42에 열거되어 있다.

[1119] 도 6a 내지 도 10b에 도시된 qPCR 결과에 나타나 있는 바와 같이, 시험관 내에서 생성되는 아넬로벡터는 감염성일 뿐만 아니라, 그들은 상피 세포, 폐 조직 세포, 간 세포, 암종 세포, 림프구, 림프모구, T 세포, B 세포 및 신장 세포의 예들을 비롯한 다양한 세포주를 감염시킬 수 있었다. 또한, 합성 아넬로벡터가 HepG2 세포(간 세포주)를 감염시켜, 대조군에 비하여 100배 초과의 증가를 초래할 수 있는 것이 관찰되었다.

[1120] **실시예 11: 복제-결핍 아넬로벡터**

[1121] 아넬로벡터의 복제 및 패키징을 위해, 일부 요소는 트랜스로 제공될 수 있다. 이는 DNA 복제 또는 패키징을 지시하거나 지원하는 단백질 또는 비-코딩 RNA를 포함한다. 일부 예에서, 트랜스 요소는, 바이러스, 플라스미드와 같은 아넬로벡터에 대해 대안적인 공급원, 또는 세포 게놈으로부터 제공될 수 있다.

[1122] 다른 요소는 통상적으로 시스로 제공된다. 이러한 요소는, 예를 들어, 복제 원점(예를 들어, 아넬로벡터 DNA의 증폭을 가능하게 하기 위함) 또는 패키징 신호(예를 들어, 단백질에 결합하여 게놈을 캡시드 내로 로딩하기 위함)의 역할을 하는 아넬로벡터 DNA의 서열 또는 구조일 수 있다. 일반적으로, 복제 결핍 바이러스 또는 아넬로벡터는 이러한 요소 중 하나 이상이 누락되어, 다른 요소가 트랜스로 제공되더라도 DNA가 감염성 비리온 또는 아넬로벡터로 패키징될 수 없다.

[1123] 복제 결핍 바이러스는, 예를 들어, 동일한 세포에서 아넬로벡터(예를 들어, 복제-결핍 또는 패키징-결핍 아넬로벡터)의 복제를 제어하는 데 유용할 수 있다. 일부 예에서, 복제-결핍 바이러스는 시스 복제 또는 패키징 요소가 결여되지만, 트랜스 요소, 예컨대 단백질 및 비-코딩 RNA를 발현할 것이다. 일반적으로, 치료용 아넬로벡터는 이들 트랜스 요소의 일부 또는 전부가 결여될 것이며, 이에 따라, 이의 자체가 복제할 수 없을 것이지만, 시스 요소를 보유할 것이다. 세포 내로 동시-형질감염/감염되는 경우, 복제-결핍 바이러스는 아넬로벡터의 증폭 및 패키징을 유도할 것이다. 이에 따라, 수집된, 패키징된 입자는 트랜스 요소를 제공하는 바이러스에 의한 오염 없이, 단지 치료용 아넬로벡터만으로 구성될 것이다.

[1124] 복제 결핍 아넬로벡터를 개발하기 위하여, 아넬로바이러스의 비-코딩 영역 내의 보존된 요소를 제거한다. 특히, 보존된 5' UTR 도메인 및 GC-풍부 도메인의 결실을 개별적으로, 그리고 함께 시험할 것이다. 두 요소는 모두 바이러스 복제 또는 패키징에 중요한 것으로 고려된다. 또한, 전체 비-코딩 영역에 걸쳐 일련의 결실을 수행하여, 이전에 공지되지 않은 관심 영역을 확인할 것이다.

[1125] 복제 요소의 성공적인 결실은 예를 들어, qPCR에 의해 측정시, 세포 내에서 아넬로벡터 DNA 증폭의 감소를 초래할 것이지만, 예를 들어, qPCR, 웨스턴 블롯, 형광 검정 또는 발광 검정 중 임의의 것 또는 전부를 포함할 수 있는 감염된 세포에서의 검정에 의해 모니터링시, 일부 감염성 아넬로벡터 생성을 뒷받침할 것이다. 패키징 요소의 성공적인 결실은 아넬로벡터 DNA 증폭을 파괴하지 않을 것이며, 그래서, qPCR에 의해 형질감염된 세포에서 아넬로벡터 DNA의 증가가 관찰될 것이다. 그러나, 아넬로벡터 게놈은 캡슐화되지 않을 것이며, 그래서 감염성 아넬로벡터 생성이 관찰되지 않을 것이다.

[1126] **실시예 12: 복제-적격 아넬로벡터의 제조 과정**

[1127] 본 실시예는 복제-적격 아넬로벡터의 회수 및 이의 생성의 대량화 방법을 기재한다. 아넬로벡터는 이의 게놈 내에 세포에서 복제하는 데 필요한, 모든 요구되는 유전 요소 및 ORF를 인코딩하는 경우에 복제 적격이 있다. 이들 아넬로벡터는 이의 복제에 결함이 없기 때문에, 트랜스로 제공되는 보완 활성을 필요로 하지 않는다. 그러나, 이들 아넬로벡터는 헬퍼 활성, 예컨대 전사의 인핸서(예를 들어, 부티르산나트륨) 또는 바이러스 전사 인자(예를 들어, 아데노바이러스 E1, E2 E4, VA; HSV Vp16 및 즉시 조기 단백질)를 필요로 할 것이다.

[1128] 이러한 실시예에서, 선형 또는 환형 형태로 합성 아넬로벡터의 완전한 서열을 인코딩하는 이중-가닥 DNA는 화학적 형질감염에 의해 T75 플라스크 내의 5E+05개의 부착 포유동물 세포 내로 또는 전기천공법에 의해 현탁액 중 5E+05개의 세포 내로 도입된다. 최적의 기간 후에(예를 들어, 형질감염 3 내지 7일 후에), 세포를 상청액 배지 내로 스크랩핑하여, 세포 및 상청액을 수집한다. 연성 세제, 예를 들어, 담즙염을 0.5%의 최종 농도로 첨가하고, 37℃에서 30분 동안 인큐베이션한다. 칼슘 및 마그네슘 클로라이드를 각각 0.5 mM 및 2.5 mM의 최종 농도로 첨가한다. 엔도뉴클레아제(예를 들어, DNAse I, 벤조나제)를 첨가하고, 25 내지 37℃에서 0.5 내지 4시간 동안 인큐베이션한다. 아넬로벡터 현탁액을 1000 × g에서 4℃에서 10분 동안 원심분리한다. 투명해진 상청액을 새로운 튜브로 옮기고, 동결보호 완충액(안정화 완충액으로도 공지되어 있음)으로 1:1 희석하고, 요망되는 경우 -80℃에 보관한다. 이는 아넬로벡터의 계대 0(P0)을 생성한다. 세제의 농도를 배양된 세포에서 사용될 안전한 한계 미만이 되게 하기 위하여, 이러한 접종물을 아넬로벡터 역가에 따라 무혈청 배지(SFM)에 적어도 100배 이상 희석한다.

[1129] T225 플라스크 내의 신선한 단층의 포유동물 세포에 배양물 표면을 덮기에 충분한 최소 부피를 오버레이하고, 부드럽게 로킹(rocking)시키면서 37℃ 및 5% 이산화탄소에서 90분 동안 인큐베이션한다. 이러한 단계를 위해 사용되는 포유동물 세포는 P0 회수를 위해 사용된 것과 동일한 유형의 세포이거나, 그렇지 않을 수 있다. 이러한 인큐베이션 후에, 접종물을 40 ㎖의 무혈청, 동물 기원-부재의 배양 배지로 대체하였다. 세포를 37℃ 및 5% 이

산화탄소에서 3 내지 7일 동안 인큐베이션한다. 이전에 사용된 동일한 연성 세제 10X 용액 4 ㎖을 첨가하여, 0.5%의 최종 세제 농도를 달성한 다음, 혼합물을 부드럽게 교반하면서 37℃에서 30분 동안 인큐베이션한다. 엔도뉴클레아제를 첨가하고, 25 내지 37℃에서 0.5 내지 4시간 동안 인큐베이션한다. 그 다음, 배지를 수집하고, 1000 × g에서 4℃에서 10분 동안 원심분리한다. 투명해진 상청액을 40 ㎖의 안정화 완충제와 혼합하고, -80℃에서 보관한다. 이는 시드 스톡을 또는 아넬로벡터의 계대 1(P1)을 생성한다.

[1130] 스톡의 역가에 따라, 그를 SFM 중에 100배 이상 희석하고, 필요한 크기의 다층 플라스크 상에서 성장한 세포에 첨가한다. 더 작은 규모에서 감염 다중도(MOI) 및 인큐베이션 시간을 최적화시켜, 최대 아넬로벡터 생성을 보장한다. 그 다음, 수집 후에, 아넬로벡터를 필요한 대로 정제하고 농축시킬 수 있다. 예를 들어, 본 실시예에 기재된 바와 같은 작업흐름을 보여주는 개략도는 도 11에 제공되어 있다.

[1131] **실시예 13: 복제-결핍 아넬로벡터의 제조 과정**

[1132] 본 실시예에는 복제-결핍 아넬로벡터의 회수 및 이의 생성의 대량화 방법이 기재되어 있다.

[1133] 아넬로벡터는 복제에 수반되는 하나 이상의 ORF(예를 들어, ORF1, ORF1/1, ORF1/2, ORF2, ORF2/2, ORF2/3 및/또는 ORF2t/3)의 결실에 의해 복제-결핍이 될 수 있다. 복제-결핍 아넬로벡터는 보완하는 세포주에서 성장할 수 있다. 이러한 세포주는 아넬로벡터 성장을 촉진하지만, 아넬로벡터의 게놈에서 소실되거나 비작용성인 구성성분을 구성적으로 발현한다.

[1134] 일 예에서, 아넬로벡터 증량에 수반되는 임의의 ORF(들)의 서열(들)을, 선택 마커를 인코딩하는 안정한 세포주의 생성에 적합한 렌티바이러스 발현 시스템 내로 클로닝하고, 렌티바이러스 벡터를 본 명세서에 기재된 바와 같이 생성한다. 아넬로벡터 증량을 지원할 수 있는 포유동물 세포주를 이러한 렌티바이러스 벡터로 감염시키고, 클로닝된 ORF가 안정적으로 통합된 세포 집단을 선택하는 선택 마커(예를 들어, 푸로마이신 또는 임의의 다른 항생제)에 의해 선택압으로 처리한다. 이러한 세포주가 특성화되고, 엔지니어링된 아넬로벡터에서 결함을 보완하고, 이에 따라 이러한 아넬로벡터의 성장 및 증식을 지원하는 것으로 증명되면, 그것을 증량시키고, 저온 보관에 저장한다. 이들 세포의 증량 및 유지 동안, 선택 항생제를 배양 배지에 첨가하여 선택압을 유지한다. 아넬로벡터가 이들 세포 내로 도입되면, 선택 항생제를 보류할 수 있다.

[1135] 일단 이러한 세포주가 확립되면, 복제-결핍 아넬로벡터의 성장 및 생성을 예를 들어, 실시예 15에 기재된 바와 같이 수행한다.

[1136] **실시예 14: 현탁 세포를 사용한 아넬로벡터의 생성**

[1137] 본 실시예에는 현탁액 중의 세포에서의 아넬로벡터의 생성이 기재되어 있다.

[1138] 본 실시예에서, 현탁액 조건에서 성장하는 것으로 조정된 A549 또는 293T 생성자 세포주를 웨이브(WAVE) 생물반응기 백에서, 37℃ 및 5% 이산화탄소에서 동물 성분-부재 및 항생제-부재 현탁액 배지(Thermo Fisher Scientific) 중에 성장시킨다. 1 × 10^6개의 생존 가능한 세포/㎖로 시딩한 이들 세포를 현재의 우수 제조 기준(cGMP) 하에 리포펙타민 2000(Thermo Fisher Scientific)을 사용하여, (예를 들어, 실시예 16에 기재된 바와 같이, 예를 들어, 복제-결핍 아넬로벡터의 경우에) 아넬로벡터를 패키징하는 데 적합하거나 필요한 임의의 보완 플라스미드와 함께 아넬로벡터 서열을 포함하는 플라스미드로 형질감염시킨다. 보완 플라스미드는 일부 예에서 아넬로벡터 게놈(예를 들어, 본 명세서에 기재된 바와 같이, 예를 들어, 바이러스 게놈, 예를 들어, 아넬로바이러스 게놈에 기초한 아넬로벡터 게놈)으로부터 결실되지만, 아넬로벡터의 복제 및 패키징에 유용하거나 필요한 바이러스 단백질을 인코딩할 수 있다. 형질감염된 세포를 웨이브 생물반응기 백에서 성장시키고, 상청액을 하기의 시점에 수집한다: 형질감염 48, 72 및 96시간 후. 상청액을 원심분리를 사용하여 각 샘플에 대한 세포 펠렛으로부터 분리한다. 이어서, 패키징된 아넬로벡터 입자를 이온 교환 크로마토그래피를 사용하여 수집된 상청액 및 용해된 세포 펠렛으로부터 정제한다.

[1139] 아넬로벡터의 정제된 제제에서 게놈 당량을 예를 들어, 실시예 18에 기재된 바와 같이, 바이러스 게놈 추출 키트(Qiagen)에 이어서, 아넬로벡터 DNA 서열에 대하여 표적화된 프라이머 및 프로브를 사용한 qPCR을 사용하여 아넬로벡터 게놈을 수집하기 위하여, 작은 정제된 제제의 분취액을 사용함으로써 결정할 수 있다.

[1140] 정제된 제제의 연속 희석물을 제조하여, 새로운 A549 세포를 감염시킴으로써 정제된 제제 내의 아넬로벡터의 감염성을 정량화할 수 있다. 이들 세포를 형질감염 72시간 후에 수집한 후에, 아넬로벡터 DNA 서열에 특이적인 프라이머 및 프로브를 사용하여 게놈 DNA에서 qPCR 검정을 행하였다.

[1141] **실시예 15: 마우스에서 외인성 단백질을 발현하기 위한 아넬로벡터의 이용**

[1142] 본 실시예에는 토르크 테노 미니 바이러스(TTMV) 게놈이 마우스에서 반딧불이 루시퍼라제 단백질을 발현하도록 엔지니어링된 아넬로벡터의 이용이 기재된다.

[1143] 반딧불이-루시퍼라제 유전자를 인코딩하는 엔지니어링된 TTMV의 DNA 서열을 인코딩하는 플라스미드를 화학적 형질감염에 의해 A549 세포(인간 폐 암종 세포주) 내로 도입한다. 18 ㎍의 플라스미드 DNA를 10 cm 조직 배양 플레이트에서 70% 컨플루언트 세포의 형질감염을 위해 사용한다. TTMV 서열이 결여된 빈 벡터 백본을 음성 대조군으로서 사용한다. 형질감염 5시간 후에, 세포를 PBS로 2회 세척하고, 37℃ 및 5% 이산화탄소에서 신선한 성장 배지에서 성장하게 한다.

[1144] 형질감염된 A549 세포를 그들의 상청액과 함께, 형질감염 96시간 후에 수집한다. 수집된 물질을 37℃에서 1시간 동안 0.5% 데옥시콜레이트(부피 중 중량)로 처리한 후에, 엔도뉴클레아제로 처리한다. 아넬로벡터 입자를 이온 교환 크로마토그래피를 사용하여 이러한 용해물로부터 정제한다. 아넬로벡터 농도를 결정하기 위하여, 아넬로벡터 스톡의 샘플을 바이러스 DNA 정제 키트를 통해 시행하고, ㎖당 게놈 당량을 아넬로벡터 DNA 서열에 대하여 표적화된 프라이머 및 프로브를 사용한 qPCR에 의해 측정한다.

[1145] 1× 인산염-완충 염수 중 소정의 용량-범위의 게놈 당량의 아넬로벡터를 8 내지 10주령 마우스에서 다양한 주사 경로(예를 들어, 정맥내, 복강내, 피하, 근육내)를 통해 수행한다. 배측 및 등측 생물발광 영상화를 주사 후 3, 7, 10 및 15일에 각 동물에서 수행한다. 제조업체의 프로토콜에 따라, 루시퍼라제 기질(퍼킨-엘머)을 표기된 시점에 각 동물에 복강내로 첨가한 후 생체내 영상화에 의해 영상화를 수행한다.

[1146] **실시예 16: 외인성 마이크로RNA 서열을 발현하는 아넬로벡터의 기능적 효과**

[1147] 본 실시예는 고유 프로모터를 사용한 아넬로벡터 게놈으로부터의 외인성 miRNA(miR-625)의 성공적인 발현을 보여준다.

[1148] 500 ng의 하기의 플라스미드 DNA를 24 웰 플레이트 내의 60% 컨플루언트 HEK293T 세포의 웰 내로 형질감염시켰다:

[1149] i) 빈 플라스미드 백본

[1150] ii) 내인성 miRNA가 녹아웃된(KO) TTV-tth8 게놈을 함유하는 플라스미드

[1151] iii) 내인성 miRNA가 비-표적화 스크램블 miRNA로 대체된 TTV-tth8

[1152] iv) 내인성 miRNA 서열이 miR-625를 인코딩하는 miRNA로 대체된 TTV-tth8

[1153] 형질감염 72시간 후에, 전체 miRNA를 Qiagen miRNeasy 키트를 사용하여 형질감염된 세포로부터 수집한 후, miRNA 스크립트(Script) RT II 키트를 사용하여 역전사를 행하였다. 정량적 PCR을 miRNA-625 또는 RNU6 소형 RNA를 특이적으로 검출할 프라이머를 사용하여 역 전사된 DNA 상에서 수행하였다. RNU6 소형 RNA를 하우스키핑(housekeeping) 유전자로서 사용하였으며, 데이터는 빈 벡터에 비한 배수 변화로서 도 12에 플롯팅되어 있다. 도 12에 나타낸 바와 같이, miR-625 아넬로벡터는 miR-625 발현의 대략 100-배 증가를 초래한 한편, 빈 벡터, miR-녹아웃(KO) 및 스크램블드 miR에 대하여 신호가 검출되지 않았다.

[1154] **실시예 17: 외인성 비-코딩 RNA 발현을 위한 아넬로벡터의 제조 및 생성**

[1155] 본 실시예에는 외인성 소형 비-코딩 RNA를 발현하기 위한 아넬로벡터의 합성 및 생성이 기재된다.

[1156] TTV의 tth8 주(Jelcic et al, *Journal of Virology*, 2004)로부터 DNA 서열을 합성하고, 박테리아 복제 원점 및 박테리아 항생제 내성 유전자를 함유하는 벡터 내로 클로닝한다. 이러한 벡터에서, TTV miRNA 헤어핀을 인코딩하는 DNA 서열을 외인성 소형 비-코딩 RNA, 예컨대 miRNA 또는 shRNA를 인코딩하는 DNA 서열로 대체한다. 그 다음, 엔지니어링된 작제물을 전기-적격 박테리아로 형질전환시킨 후, 제조업체의 프로토콜에 따라 플라스미드 정제 키트를 사용하여 플라스미드 단리를 행한다.

[1157] 외인성 소형 비-코딩 RNA를 인코딩하는 아넬로벡터 DNA를 진핵 생성자 세포주 내로 형질감염시켜, 아넬로벡터 입자를 생성한다. 아넬로벡터 입자를 함유하는 형질감염된 세포의 상청액을 형질감염 후 상이한 시점에 수확한다. 여과된 상청액으로부터의 또는 정제 이후의 아넬로벡터 입자를, 예를 들어, 본 명세서에 기재된 바와 같이, 하류 적용에 사용한다.

[1158] **실시예 18: 아넬로벡터로부터 내인성 miRNA의 발현 및 내인성 miRNA의 결실**

[1159] 일 예에서, TTV-tth8 게놈을 실시예 27에 기재된 바와 같은 GC-풍부 영역 내의 결실을 사용하여 변형시킨 변형된 TTV-tth8 게놈을 포함하는 아넬로벡터를 사용하여 배양 중 Raji B 세포를 감염시켰다. 이들 아넬로벡터는 n-myc 상호작용 단백질(NMI)을 인코딩하는 mRNA를 표적화하는 miRNA인 TTV-tth8 아넬로바이러스의 내인성 페이로드를 인코딩하는 서열을 포함하였으며, 아넬로바이러스 게놈을 포함하는 플라스미드를 숙주 세포 내로 도입함으로써 생성하였다. NMI는 JAK/STAT 경로의 하류에서 작동하여, 인터페론-자극된 유전자, 증식 및 성장 유전자, 및 염증 반응의 매개자를 포함하는 다양한 세포내 신호의 전사를 조절한다. 도 13에 나타낸 바와 같이, 바이러스 게놈이 표적 Raji B 세포에서 검출되었다. 대조군 세포에 비하여 표적 Raji B 세포에서 NMI의 성공적인 녹다운이 또한 관찰되었다(도 14). NMI에 대한 miRNA를 포함하는 아넬로벡터는 대조군 세포에 비하여 NMI 단백질 수준의 75% 초과의 감소를 유도하였다. 본 실시예는 고유 아넬로바이러스 miRNA를 갖는 아넬로벡터가 숙주 세포에서 표적 분자를 녹다운시킬 수 있는 것을 보여준다.

[1160] 또 다른 예에서, 아넬로바이러스-기반의 아넬로벡터의 내인성 miRNA를 결실시켰다. 그 다음, 생성된 아넬로벡터(ΔmiR)을 숙주 세포와 함께 인큐베이션하였다. 그 다음, ΔmiR 아넬로벡터 유전 요소의 게놈 당량을 내인성 miRNA를 보유하는 상응하는 아넬로벡터의 것과 비교하였다. 도 15에 나타낸 바와 같이, 내인성 miRNA가 결실된 아넬로벡터 게놈은 내인성 miRNA가 여전히 존재하는 아넬로벡터 게놈에 대하여 관찰되는 것과 유사한 수준으로 세포에서 검출되었다. 이러한 실시예는 아넬로바이러스-기반의 아넬로벡터의 내인성 miRNA가 완전히 돌연변이되거나 결실될 수 있으며, 아넬로벡터 게놈이 표적 세포에서 여전히 검출될 수 있는 것을 보여준다.

[1161] **실시예 19: 생체 내에서의 외인성 단백질의 아넬로벡터 운반**

[1162] 본 실시예는 투여 후의 아넬로벡터의 생체 내 이펙터 기능(예를 들어, 단백질의 발현)을 보여준다.

[1163] 나노-루시퍼라제(nLuc)를 인코딩하는 이식유전자를 포함하는 아넬로벡터(도 16a 및 도 16b)을 제조하였다. 약술하여, TTMV-LY2 비-코딩 영역 및 nLuc 발현 카세트를 갖는 이중-가닥 DNA 플라스미드를 트랜스 복제 및 패키징 인자로서 작용하도록 전장 TTMV-LY2 게놈을 인코딩하는 이중-가닥 DNA 플라스미드와 함께, HEK293T 세포 내로 형질감염시켰다. 형질감염 후에, 세포를 인큐베이션하여, 아넬로벡터 생성을 가능하게 하였으며, 아넬로벡터 물질을 수집하고, 뉴클레아제 처리, 한외여과/정용여과 및 멸균 여과를 통해 농축시켰다. 추가의 HEK293T 세포를 "비-바이러스" 음성 대조군으로서 작용하도록 nLuc 발현 카세트 및 TTMV-LY2 ORF 형질감염 카세트를 갖지만, 복제 및 패키징에 필수적인 비-코딩 도메인이 결여된 비-복제성 DNA 플라스미드로 형질감염시켰다. 비-바이러스 샘플을 아넬로벡터 물질과 동일한 프로토콜에 따라 제조하였다.

[1164] 아넬로벡터 제제를 3마리 건강한 마우스의 코호트에 근육내로 투여하고, 9일의 경과에 걸쳐 IVIS 루미나(Lumina) 영상화(브루커(Bruker))에 의해 모니터링하였다(도 17a). 비-바이러스 대조군으로서, 비-복제성 제제를 추가 3마리의 마우스에 투여하였다(도 17b). 25 ㎕의 아넬로벡터 또는 비-바이러스 제제의 주사를 제0일에, 왼쪽 뒷다리에 투여하고, 제4일에 오른쪽 뒷다리에 재-투여하였다(도 17a 및 도 17b에서 화살표 참조). 제9일의 IVIS 영상화 후에, 더 많은 nLuc 발광 신호의 발생이 비-바이러스 제제(도 17b)보다 아넬로벡터 제제(도 17a)를 주사한 마우스에서 관찰되었으며, 이는 생체 내 아넬로벡터 형질도입 후의 트랜스 유전자 발현과 일치한다.

[1165] **실시예 20: 시험관 내 환화된 아넬로바이러스 게놈**

[1166] 본 실시예에는 최소 비-바이러스 DNA와 함께 환형, 이중 가닥 아넬로바이러스 게놈 DNA를 포함하는 작제물이 기재된다. 이들 환형 바이러스 게놈은 야생형 아넬로바이러스 복제 동안 관찰되는 이중-가닥 DNA 중간체와 더욱 근접하게 일치한다. 이러한 환형, 이중 가닥 아넬로바이러스 게놈 DNA가 최소 비-바이러스 DNA와 함께 세포 내로 도입되는 경우, 회전환 복제를 겪어, 예를 들어, 본 명세서에 기재된 바와 같은 유전 요소를 생성할 수 있다.

[1167] 일 예에서, TTV-tth8 변이체 및 TTMV-LY2를 보유하는 플라스미드를 게놈 DNA에 측접하는 제한 엔도뉴클레아제 인식 부위를 사용하여 분해하였다. 이어서, 생성되는 선형화된 게놈을 결찰시켜, 환형 DNA를 형성하였다. 이들 결찰 반응을 다양한 DNA 농도를 사용하여 행하여, 분자내 결찰을 최적화시켰다. 결찰된 환을 포유동물 세포 내에 바로 형질감염시키거나, 플라스미드 백본을 절단하기 위한 제한 엔도뉴클레아제 및 선형 DNA를 분해하기 위한 엑소뉴클레아제를 사용한 분해에 의해 추가로 처리하여, 비-환형 게놈 DNA를 제거하였다. TTV-tth8에 있어서, XmaI 엔도뉴클레아제를 사용하여, DNA를 선형화시켰으며; 결찰된 환은 GC-풍부 영역과 5' 비-코딩 영역 사이에 53 bp의 비-바이러스 DNA를 함유하였다. TTMV-LY2에 있어서, IIS형 제한 효소 Esp3I을 사용하여, 비-바이러스 DNA를 갖지 않는 바이러스 게놈 DNA 환을 제공하였다. 이 프로토콜을 이전에 공개된 TTV-tth8의 환화로

부터 조정하였다(문헌[Kincaid et al., 2013, PLoS Pathogens 9(12): e1003818]). 아넬로바이러스 생성의 개선을 입증하기 위하여, 환화된 TTV-tth8 및 TTMV-LY2를 HEK293T 세포 내로 형질감염시켰다. 7일의 인큐베이션 후에, 세포를 용해시키고, qPCR을 수행하여, 환화된 및 플라스미드-기반의 아넬로바이러스 게놈 간에 아넬로바이러스 게놈의 수준을 비교하였다. 증가된 아넬로바이러스 게놈의 수준은 바이러스 DNA의 환화가 아넬로바이러스 생성을 증가시키기 위한 유용한 전략인 것을 보여준다.

[1168] 또 다른 예에서, TTMV-LY2 플라스미드(pVL46-240) 및 TTMV-LY2-nLuc를 각각 Esp3I 또는 EcoRV-HF를 사용하여 선형화시켰다. 분해된 플라스미드를 1% 아가로스 겔 상에 정제한 후에, 전기용리 또는 Qiagen 컬럼 정제 및 T4 DNA 리가제를 사용한 결찰을 행하였다. 환화된 DNA를 형질감염 전에 100 kDa UF/DF 멤브레인 상에서 농축시켰다. 환화를 도 18a에 나타낸 바와 같이, 겔 전기영동에 의해 확인하였다. 리포펙타민 2000을 사용한 리포펙션 1일 전에, T-225 플라스크에, 3 x 10^4개 세포/㎠의 HEK293T를 시딩하였다. 9 마이크로그램의 환화된 TTMV-LY2 DNA 및 50 ㎍의 환화된 TTMV-LY2-nLuc를 플라스크 시딩 1일 후에 동시-형질감염시켰다. 비교로서, 추가의 T-225 플라스크에, 50 ㎍의 선형화된 TTMV-LY2 및 50 ㎍의 선형화된 TTMV-LY2-nLuc를 동시-형질감염시켰다.

[1169] 아넬로벡터 생성을 트리톤 X-100 수집 완충제 중에서의 세포 수집 이전에 8일 동안 진행하였다. 일반적으로, 아넬로벡터를 예를 들어, 숙주 세포의 용해, 용해물의 정화, 여과 및 크로마토그래피에 의해 농축시킬 수 있다. 본 실시예에서, 수집된 세포를 염화나트륨 조정 및 1.2 ㎛/ 0.45 ㎛ 정상 유동 여과 전에 뉴클레아제 처리하였다. 정화된 수집액을 농축시키고, 750 kDa MWCO mPES 중공사막 상에서 PBS로 완충제 교환하였다. TFF 투석유물을 0.45 ㎛ 필터를 사용하여 여과한 후에, PBS 중에 사전-평형화된 세파크릴(Sephacryl) S-500 HR SEC 컬럼 상에 로딩하였다. 아넬로벡터를 30 cm/hr로 SEC 컬럼에서 처리하였다. 개별 분획을 수집하고, 도 18b에 나타낸 바와 같이, 바이러스 게놈 카피수 및 이식유전자 카피수에 대하여 qPCR에 의해 검정하였다. 바이러스 게놈 및 이식유전자 카피는 SEC 크로마토그램의 공극 부피인 분획 7에서 시작하여 관찰되었다. 잔류 플라스미드 피크는 분획 15에서 관찰되었다. TTMV-LY2 게놈 및 TTMV-LY2-nLuc 이식유전자에 대한 카피수는 분획 7 내지 분획 10에서 nLuc 이식이유전자를 함유하는 팩킹된 아넬로벡터를 나타내는 환화된 투입 DNA를 사용하여 생성된 아넬로벡터에 대하여 우수하게 일치하였다. SEC 분획을 풀링하고, 100 kDa MWCO PVDF 멤브레인을 사용하여 농축시킨 다음, 생체 내 투여 이전에 0.2 ㎛ 여과하였다.

[1170] 투입 아넬로벡터 DNA의 환화는 선형화된 아넬로벡터 DNA와 비교하는 경우, 정제 과정을 통한 뉴클레아제 보호된 게놈의 회수 백분율의 3배 증가를 초래하였으며, 이는 표 46에 나타낸 바와 같은 환화된 투입 아넬로벡터 DNA를 사용하여 개선된 제조 효율을 나타낸다.

[1171] **[표 46] 정제 과정 수율**

단계	선형화된 TTMV-LY2		환화된 TTMV-LY2	
	총 바이러스 게놈 카피	**총 nLuc 이식유전자 게놈 카피**	**총 바이러스 게놈 카피**	**총 nLuc 이식유전자 게놈 카피**
사전-뉴클레아제 수확	2.78E+12	2.17E+12	1.04E+11	4.39E+11
정화된 수확	9.96E+09	5.48E+09	6.55E+08	9.81E+08
TFF	1.01E+10	7.66E+09	2.58E+08	3.56E+08
SEC	3.18E+07	8.73E+06	9.16E+06	7.75E+06
UF/DF	8.82E+06	3.25E+06	1.78E+06	2.73E+06
멸균 여과	5.60E+06	2.64E+06	8.66E+05	1.63E+06
정제 과정 수율(%)	0.0002%	0.0001%	0.0006%	0.0004%

[1172]

[1173] **실시예 21: 상이한 토르크 테노 바이러스 주 유래의 초가변 도메인을 갖는 키메라 ORF1을 함유하는 아넬로벡터의 생성**

[1174] 본 실시예에는 하나의 TTV 주의 ORF1 아르기닌-풍부 영역, 젤리-롤 도메인, N22 및 C-말단 도메인 및 상이한 TTV 주의 ORF1 단백질로부터의 초가변 도메인을 함유하는 키메라 아넬로벡터를 생성하기 위한 ORF1의 초가변 영역의 도메인 스와핑이 기재된다.

[1175] 베타토르크바이러스의 전장 게놈 LY2 주를 포유동물 세포에서의 발현을 위하여 발현 벡터 내에 클로닝하였다. 이 게놈을 돌연변이시켜, LY2의 초가변 도메인을 제거하고, 그것을 먼 관계의 베타토르크바이러스의 초가변 도메인으로 대체하였다(도 18c). 이어서, 스와핑된 초가변 도메인을 갖는 LY2 게놈을 함유하는 플라스미드(pTTMV-LY2-HVRa-z)를 이전에 공개된 방법(문헌[Kincaid et al., PLoS Pathogens 2013])을 사용하여 선형화시키고 환화시킨다. HEK293T 세포를 환화된 게놈으로 형질감염시키고, 5 내지 7일 동안 인큐베이션하여, 아넬로벡터 생성을 가능하게 한다. 인큐베이션 기간 후에, 아넬로벡터를 구배 초원심분리에 의해 형질감염된 세포의 상청액 및 세포 펠렛으로부터 정제한다.

[1176] 키메라 아넬로벡터가 여전히 감염성인지를 결정하기 위하여, 단리된 바이러스 입자를 비감염된 세포에 첨가한다. 세포를 5 내지 7일 동안 인큐베이션하여, 바이러스 복제를 가능하게 한다. 인큐베이션 후에, 감염을 확립하기 위한 키메라 아넬로벡터의 능력은 면역형광, 웨스턴 블롯 및 qPCR에 의해 모니터링될 것이다. 키메라 바이러스의 구조적 온전성을 음성 염색 및 초저온-전자 현미경법에 의해 평가한다. 키메라 아넬로벡터를 추가로 생체 내에서 세포를 감염시키는 능력에 대하여 시험할 수 있다. 초가변 도메인 스와핑을 통하여 기능적 키메라 아넬로벡터를 생성하는 능력의 확립은 향성을 변경시키고, 면역 검출을 잠재적으로 회피하기 위한 바이러스의 엔지니어링을 가능하게 할 수 있었다.

[1177] **실시예 22: 초가변 도메인 대신에 비-TTV 단백질/펩티드를 함유하는 키메라 ORF1의 생성**

[1178] 본 실시예에는 하나의 TTV 주의 아르기닌-풍부 영역, 젤리-롤 도메인, N22 및 C-말단 도메인, 및 초가변 도메인 대신에 비-TTV 단백질/펩티드를 함유하는 키메라 ORF1 단백질을 생성하기 위하여, ORF1의 초가변 영역을 다른

관심 단백질 또는 펩티드로 대체하는 것이 기재된다.

[1179] 실시예 B에 나타낸 바와 같이, LY2의 초가변 도메인을 게놈으로부터 결실시키고, 관심 단백질 또는 펩티드를 이 영역 내로 삽입할 수 있다(도 18d). 이 영역 내로 도입될 수 있는 서열의 유형의 예는 친화성 태그, 항체의 단일 쇄 가변 영역(scFv) 및 항원 펩티드를 포함하지만, 이에 제한되지 않는다. 플라스미드(pTTMV-LY2-ΔHVR-POI) 내의 돌연변이된 게놈을 실시예 B에 기재된 바와 같이 선형화시키고 환화시킨다. 환화된 게놈을 HEK293T 세포 내로 형질감염시키고, 5 내지 7일 동안 인큐베이션한다. 인큐베이션 후에, POI를 함유하는 키메라 아넬로벡터를 적절한 경우 초원심분리 및/또는 친화성 크로마토그래피를 통해 상청액 및 세포 펠렛으로부터 정제한다.

[1180] POI를 함유하는 기능적 키메라 아넬로벡터를 생성하기 위한 능력을 다양한 기법을 사용하여 평가한다. 먼저, 정제된 바이러스를 비감염된 세포에 첨가하여, 키메라 아넬로벡터가 페이로드를 복제하고/복제하거나 페이로드를 나이브 세포에 운반할 수 있는지를 결정한다. 추가로, 키메라 아넬로벡터의 구조적 온전성을 전자 현미경법을 사용하여 평가한다. 시험관 내에서 기능성인 키메라 아넬로벡터에 있어서, 생체 내에서의 페이로드의 복제/운반 능력도 또한 평가한다.

[1181] **실시예 23: tth8 및 LY2에 기반한 아넬로벡터는 각각 EPO 유전자를 폐암 세포 내로 성공적으로 형질도입하였다**

[1182] 본 실시예에서, 비-소세포 폐암 세포주(EKVX)를 에리트로포이에틴 유전자(EPO)를 보유하는 2가지 상이한 아넬로벡터를 사용하여 형질도입시켰다. 아넬로벡터를 본 명세서에 기재된 바와 같이, 시험관 내 환화에 의해 생성하였으며, 이는 LY2 또는 tth8 백본 중 어느 하나에 기반한 2가지 유형의 아넬로벡터를 포함하였다. LY2-EPO 및 tth8-EPO 아넬로벡터 각각은, 예를 들어, 실시예 39에 기재된 바와 같이, 각각 LY2 또는 tth8 게놈의 EPO-인코딩 카세트 및 비-코딩 영역(5' UTR, GC-풍부 영역)을 포함하지만, 아넬로바이러스 ORF를 포함하지 않는 유전 요소를 포함하였다. 세포에 정제된 아넬로벡터 또는 양성 대조군(고 용량 또는 아넬로벡터와 동일한 용량의 AAV2-EPO)을 접종하고, 7일 동안 인큐베이션하였다. 아넬로바이러스 ORF를 개별 시험관 내 환화된 DNA에 트랜스로 제공하였다. 배양 상청액을 접종 후 제3일, 제5.5일 및 제7일에 샘플링하고, 상용의 ELISA 키트를 사용하여 분석하여, EPO를 검출하였다. LY2-EPO 및 tth8-EPO 아넬로벡터 둘 다는 세포를 성공적으로 형질도입시켰으며, 이는 미처리(음성) 대조군 세포와 비교하여 유의하게 더 높은 EPO 역가를 나타내었다(모든 시점에 $P < 0.013$)(도 19).

[1183] **실시예 24: 치료용 이식유전자를 갖는 아넬로벡터는 정맥내(i.v.) 투여 후 생체 내에서 검출될 수 있다**

[1184] 본 실시예에서, 인간 성장 호르몬(hGH)을 인코딩하는 아넬로벡터는 정맥내(i.v.) 투여 후에 생체 내에서 검출되었다. LY2 백본에 기반한, 및 외인성 hGH(LY2-hGH)를 인코딩하는 복제-결핍 아넬로벡터를 본 명세서에 기재된 바와 같은 시험관 내 환화에 의해 생성하였다. LY2-hGH 아넬로벡터의 유전 요소는, 예를 들어, 실시예 39에 기재된 바와 같이, LY2 비-코딩 영역(5' UTR, GC-풍부 영역) 및 hGH-인코딩 카세트를 포함하였지만, 아넬로바이러스 ORF를 포함하지 않았다. LY2-hGH 아넬로벡터를 마우스에 정맥내 투여하였다. 아넬로바이러스 ORF를 개별 시험관 내 환화된 DNA에서 트랜스로 제공하였다. 간략하게, 아넬로벡터(LY2-hGH) 또는 PBS를 제0일에 정맥내 주사하였다(n=4 마우스/그룹). 아넬로벡터를 마우스당 4.66E+07개의 아넬로벡터 게놈으로 독립적인 동물 그룹에 투여하였다.

[1185] 제1의 예에서, 아넬로벡터 바이러스 게놈 DNA 카피를 검출하였다. 제7일에, 혈액 및 혈장을 수집하고, qPCR에 의해 hGH DNA 앰플리콘에 대하여 분석하였다. LY2-hGH 아넬로벡터는 생체 내에서 감염 7일 후에 전혈의 세포 분획에 존재하였다(도 20a). 추가로, 혈장 중 아넬로벡터의 부재는 이들 아넬로벡터가 생체 내에서 복제될 수 없음을 입증하였다(도 20b).

[1186] 제2의 예에서, hGH mRNA 전사물을 생체 내 형질도입 후에 검출하였다. 제7일에, 혈액을 수집하고, qRT-PCR에 의해 hGH mRNA 전사물 앰플리콘에 대하여 분석하였다. GAPDH를 대조군 하우스키핑 유전자로서 사용하였다. hGH mRNA 전사물을 전혈의 세포 분획에서 측정하였다. 아넬로벡터-인코딩된 이식유전자로부터의 mRNA를 생체 내에서 검출하였다(도 21).

[1187] **실시예 25: 시험관 내에서 아넬로벡터를 생성하기 위한 투입 물질로서 시험관 내 환화된 게놈**

[1188] 본 실시예는 본 명세서에 기재된 바와 같은 아넬로벡터 유전 요소에 대한 공급원 물질로서, 시험관 내 환화된(IVC) 이중 가닥 아넬로바이러스 DNA가 예상되는 밀도의 패키징된 아넬로벡터 게놈을 제공하는 데 있어 플라스미드 내의 아넬로바이러스 게놈 DNA보다 더욱 강력한 것을 보여준다.

[1189] T75 플라스크 내의 1.2E+07개의 HEK293T 세포(인간 배아 신장 세포주)를 11.25 μg의 (i) 시험관 내 환화된 이중

가닥 TTV-tth8 게놈(IVC TTV-tth8), (ii) 플라스미드 백본 내의 TTV-tth8 게놈 또는 (iii) TTV-tth8의 ORF1 서열만을 함유하는 플라스미드(비-복제성 TTV-tth8) 중 어느 하나로 형질감염시켰다. 세포를 형질감염 7일 후에 수집하고, 0.1% 트리톤으로 용해시키고, mℓ당 100 유닛의 벤조나제로 처리하였다. 용해물을 염화세슘 밀도 분석을 위해 사용하였으며; 염화세슘 선형 기울기의 각각의 분획에 대하여 밀도를 측정하였으며, TTV-tth8 카피 정량화를 수행하였다. 도 22에 나타낸 바와 같이, IVC TTV-tth8은 예상되는 1.33의 밀도에서, TTV-tth8 플라스미드에 비하여 현저하게 더 많은 바이러스 게놈 카피를 제공하였다.

[1190] 1E+07개의 Jurkat 세포(인간 T 림프구 세포주)를 플라스미드 내의 시험관 내의 환화된 LY2 게놈(LY2 IVC) 또는 LY2 게놈 중 어느 하나를 사용하여 뉴클레오펙션시켰다. 세포를 형질감염 4일 후에 수집하고, 0.5% 트리톤 및 300 mM 염화나트륨을 함유하는 완충제를 사용하여 용해시킨 후, 2회의 즉각적인 동결-해동을 행하였다. 용해물을 100 유닛/mℓ의 벤조나제를 사용하여 처리한 후, 염화세슘 밀도 분석을 행하였다. 밀도 측정 및 LY2 게놈 정량화를 염화세슘 선형 기울기의 각각의 분획 상에서 수행하였다. 도 23에 나타낸 바와 같이, Jurkat 세포 내의 시험관 내 환화된 LY2 게놈의 형질감염은 도 23에서 검출 가능한 피크를 보이지 않았던 LY2 게놈을 함유하는 플라스미드의 형질감염에 비하여, 예상되는 밀도에서 날카로운 피크를 야기하였다.

[1191] **실시예 26: Ring2 탠덤 아넬로벡터 작제물의 설계 및 구축**

[1192] 본 실시예는 RING2 아넬로벡터를 구조하는 데 적합한 예시적인 탠덤 작제물의 설계를 기재한다. 간단히 말해서, 야생형 RING2 게놈의 2개 카피를 머리부터 꼬리까지 나란히 포함하는 플라스미드 작제물을 설계하고 구축한다. 플라스미드는 또한 대장균 세포에서 증식하기 위한 박테리아 복제 원점 및 스펙티노마이신 내성 유전자를 인코딩한다. 이들 플라스미드의 16개 유도체를 또한 설계 및 구축하는데, 여기서 RING2 게놈의 986개 염기쌍을 탠덤 플라스미드의 제1 카피 또는 제2 카피에서 300개 염기쌍마다 nLuc 카세트로 대체한다(개략도는 도 24a 내지 24b에 나타낸 바와 같음). 986-염기쌍 nLuc 카세트는 SV40 프로모터, 코작 서열, 나노 루시퍼라제 유전자에 대한 코딩 서열 및 SV40 폴리 A 서열로 이루어진다. "nLuc 카세트"를 함유하는 게놈 카피에서 ORF의 발현은 개시 코돈을 돌연변이시킴으로써 녹아웃되는 반면, 탠덤의 다른 카피는 여전히 모든 ORF를 발현한다. 클로닝 후, 이러한 작제물은 반복 요소를 함유하는 플라스미드의 단리를 위해 엔지니어링된 이. 콜라이 세포에서 증식된다(New England Biolabs). endofree 플라스미드 추출 키트(Qiagen)를 사용하여 플라스미드를 추출한다.

[1193] **실시예 27: 탠덤 작제물을 사용한 RING2 아넬로벡터의 회수**

[1194] 본 실시예는 MOLT4 세포에서 탠덤 작제물을 사용한 재조합 RING2 아넬로벡터의 회수를 기재한다.

[1195] 실시예 26에 기재된 각각의 탠덤 플라스미드 100 ug을 1E7 MOLT4 세포에 개별적으로 형질감염시켰다. 형질감염된 MOLT4 세포를 완전한 성장 배지(10% FBS, 100 mM 피루브산나트륨 및 0.01% 플루로닉을 함유하는 RPMI)에 ml당 4E5로 접종하고 5% 이산화탄소가 있는 37도 인큐베이터에서 125 RPM으로 진탕 배양한다. 형질감염 4일 후, 세포를 펠렛으로 수확하고 성장 배지를 버린다. 세포 펠렛을 0.5% 트리톤-X100, 50 mM 트리스 pH 8.0 및 300 mM 염화나트륨을 함유하는 용해 완충액에 재현탁한다. 세포를 2회 동결 해동에 의해 용해한 후, 100 U/ml의 벤조나제로 실온에서 90분 동안 처리한다. 그 다음 벤조나제 처리 용해물을 4도에서 30분 동안 10,000×g에서 회전시켜 정화하였다. 정화된 용해물에 대해 등밀도 원심분리를 수행하였다. 등밀도 원심분리 후, 튜브 상단에서 1 ml의 분획을 수집한다. 수집된 각각의 분획을 밀도에 대해 분석하고 qPCR 분석으로 DNAse 보호 nLuc 카피를 정량화한다. RING2 아넬로벡터의 패키징은 RING2 입자에 대한 예상 밀도의 역가를 기반으로 결정되며 RING2 아넬로벡터를 함유하는 분획을 사용하여 시험관 내 형질도입을 테스트한다. 추가 돌연변이체를 제조함으로써 아넬로벡터를 구제할 수 있는 탠덤 작제물을 더 최적화한다.

[1196] **실시예 28: 탠덤 작제물을 사용한 MOLT4 세포에서 RING2 입자의 구조**

[1197] 본 실시예는 등밀도 원심분리에 의해 결정된 바와 같이, 탠덤 작제물을 사용한 야생형 RING2 입자의 성공적인 구조를 기재한다.

[1198] 머리부터 꼬리까지 나란히 RING2 게놈을 함유하는 2 mg의 플라스미드로 2E9 MOLT4 세포를 형질감염시켰다. 형질감염된 세포를 5% 이산화탄소가 있는 37도 인큐베이터에서 100 RPM으로 진탕하는 완전 성장 배지(10% FBS, 100 mM 피루브산나트륨 및 0.01% 플루로닉을 함유하는 RPMI)에서 ml당 4E5개 세포로 접종하였다. 형질감염된 세포를 형질감염 4일 후 펠렛으로 수확하였다. 세포를 재현탁 완충액(50 mM 트리스 pH 8.0, 100 mM 염화나트륨)에 재현탁시키고 미세유동화기를 사용하여 10,000 PSI에서 용해를 수행하였다. 용해된 세포를 실온에서 90분 동안 100 U/ml의 벤조나제로 처리한 후, 실온에서 45분 동안 0.5% 트리톤-X100으로 처리하였다. 세제 처리 용해물을 30분 동안 10,000×g에서 회전시켜 정화시키고 분획화하였다. 정화된 용해물에 등밀도 초원심분리를

수행하였다. 회전 후, 튜브 상단에서 1 ml의 분획을 수집하였다. 각각의 분획을 밀도에 대해 분석하고 qPCR에 의해 RING2 바이러스 게놈의 DNAse 내성 정량화를 수행하였다. 등밀도 초원심분리에 대한 대표적인 데이터는 도 25에 나타나 있다.

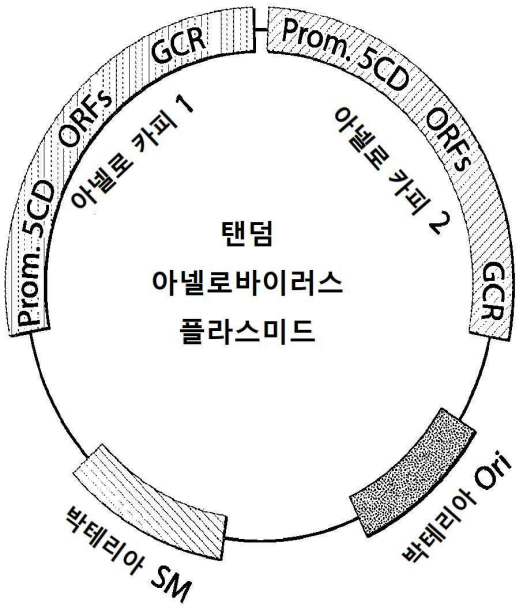
[1199] **실시예 29: MOLT4 세포에서 생성된 RING2 입자의 정제**

[1200] 본 실시예는 실시예 28에 기재된 탠덤 플라스미드 작제물을 사용하여 MOLT4 세포에서 구조된 RING2 입자의 정제를 기재한다.

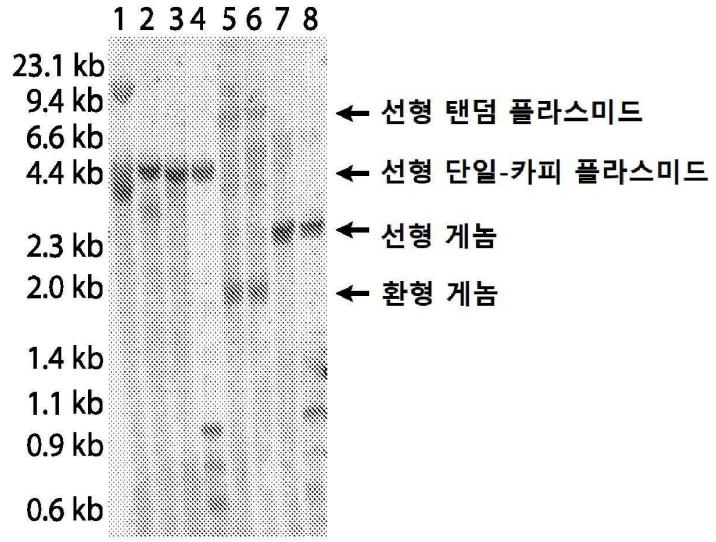
[1201] RING2 입자를 함유하는 실시예 28에서 생성된 분획을 함께 모았다. 추가 크로마토그래피 정제를 위해, Cellufine Max DexS Hbp 흡착제와 Mustang Q 음이온 교환 필터를 선택하였는데, 그 이유는 이들이 양호한 정제와 회수를 제공하기 때문이다. 50 mL(1.6 × 25 cm) Cellufine Max DexS Hbp 컬럼을 다음과 같이 실행하였다: 50 mM 트리스 pH 7.5 + 150 mM NaCl + 0.01% Tween 80에서 컬럼을 평형화하였다. 풀을 평형 완충액으로 1:1로 희석하고(pH 및 전도도를 조정하기 위함) 컬럼에 로딩하였다. 로딩 후, 컬럼을 평형 완충액으로 A280 기준선으로 세척하고 결합된 단백질을 2.5 M NaCl로 용출한다. RING2 입자는 통과액 분획에 함유되었다. 1/10 부피의 1 M 트리스 pH 9.0을 사용하여 이 분획의 pH를 조정하고 앞서 50 mM 트리스 pH 9.0 + 0.01% Tween 80에서 평형화한 Mustang Q XT 필터(3 mL 흡착제 등가물)를 통해 로딩하였다. 로딩 후, 컬럼을 A280 기준선으로 세척하고 생형물을 1 M NaCl로 용출하였다. 이 용출 분획을 Microsep-10 원심 농축기를 사용하여 10 내지 100× 농축시키고, SDS-PAGE, 웨스턴 블롯, qPCR, 및 EM에 의해 최종 풀을 분석하였다(도 26a 및 26b).

도면

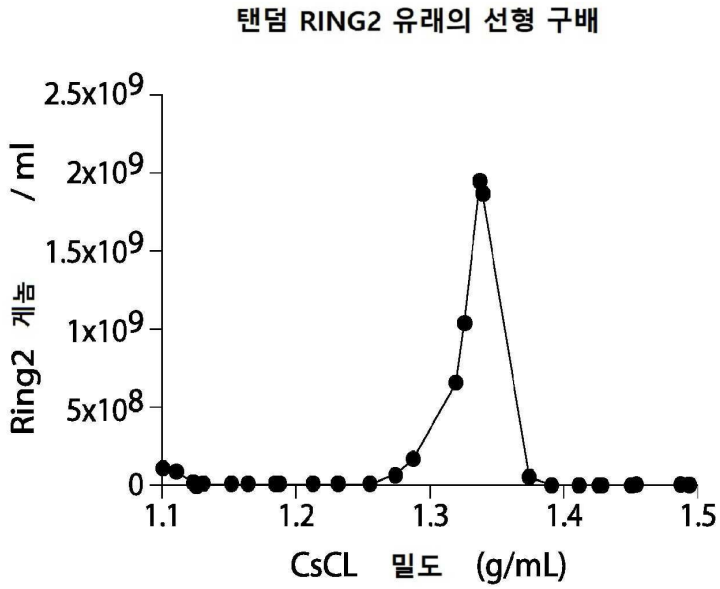
도면1a



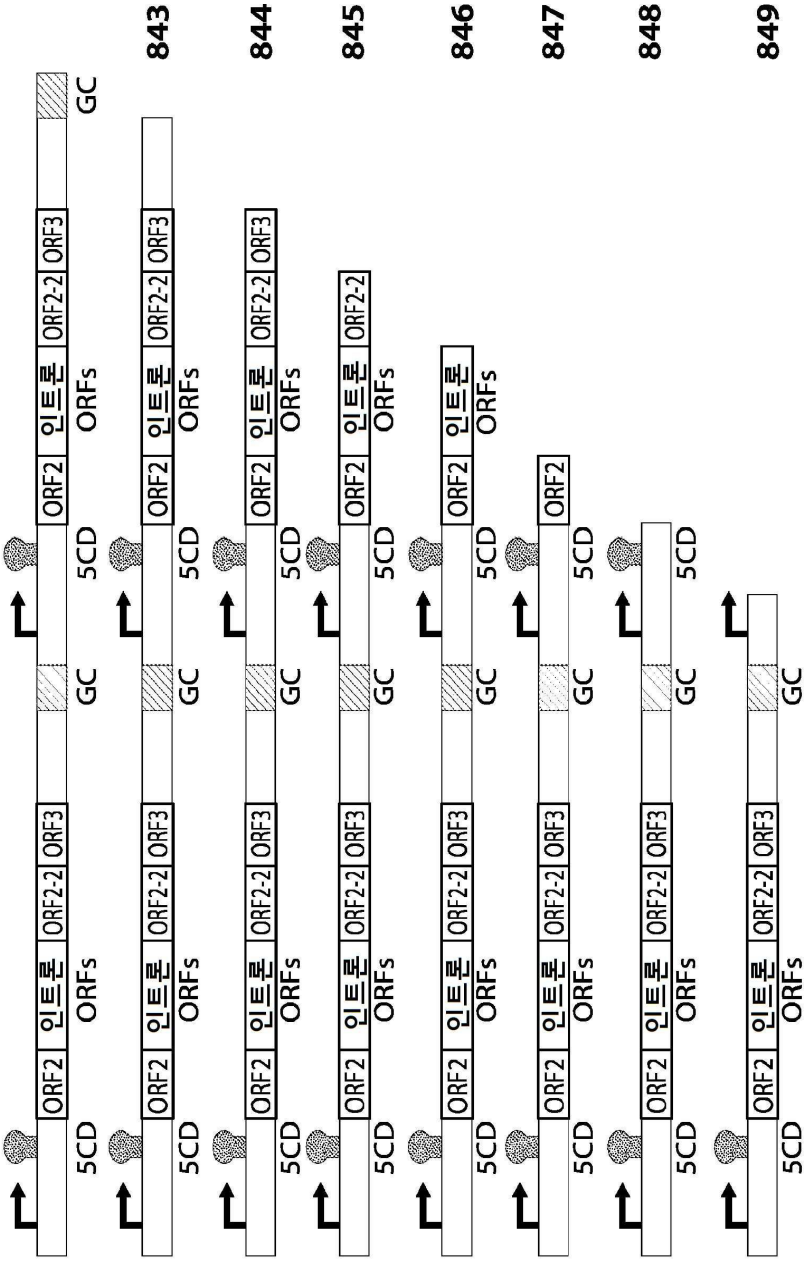
도면1b


1 2 3 4 5 6 7 8
23.1 kb
9.4 kb
6.6 kb
4.4 kb
2.3 kb
2.0 kb
1.4 kb
1.1 kb
0.9 kb
0.6 kb
선형 탠덤 플라스미드
선형 단일-카피 플라스미드
선형 게놈
환형 게놈

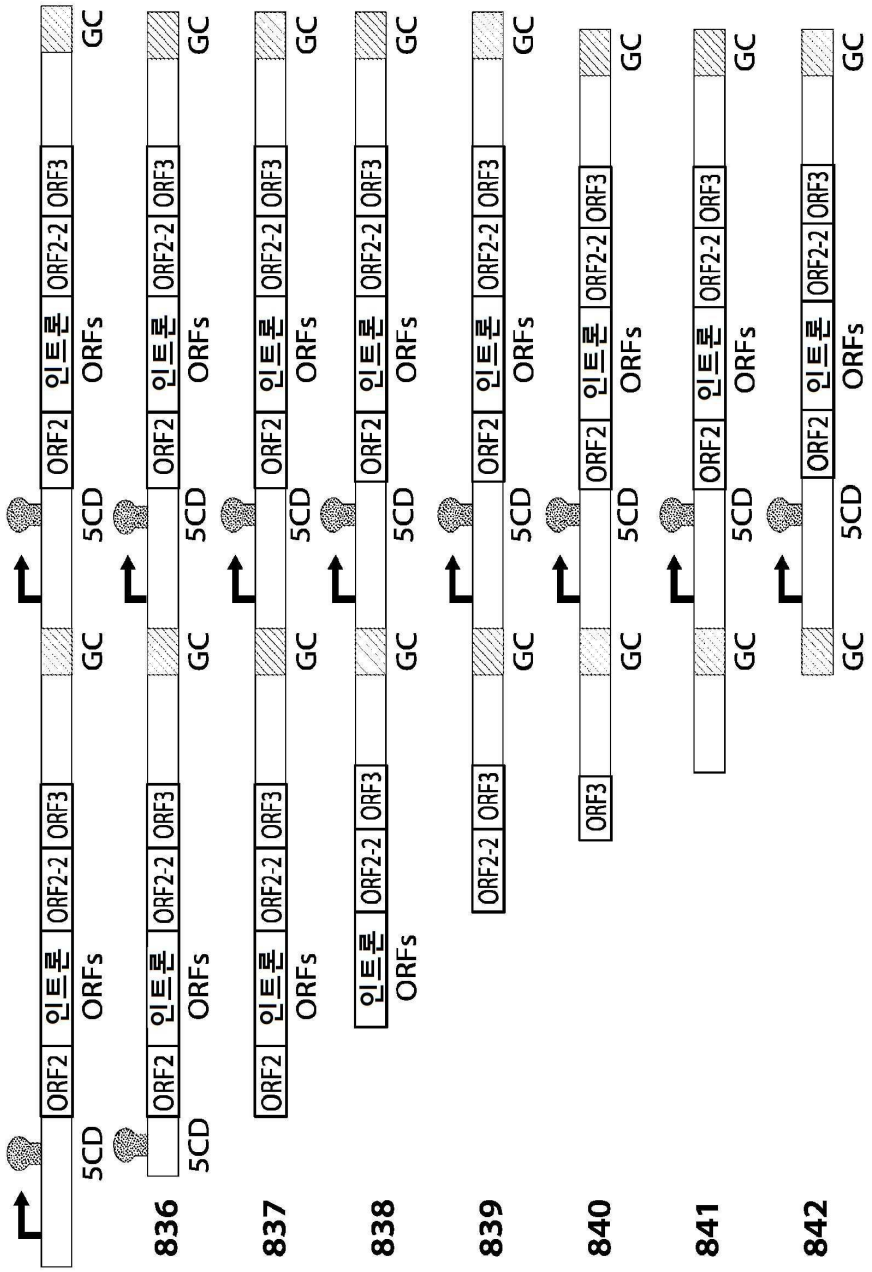
도면1c


탠덤 RING2 유래의 선형 구배
Ring2 게놈 / ml
2.5x10^9
2x10^9
1.5x10^9
1x10^9
5x10^8
0
1.1
1.2
1.3
1.4
1.5
CsCL 밀도 (g/mL)

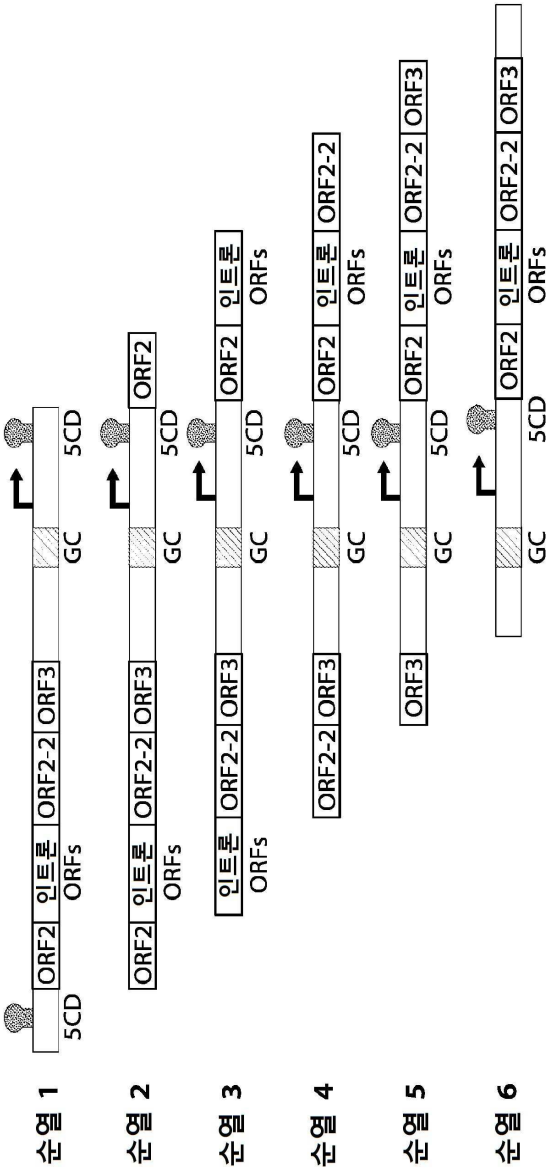
도면2a



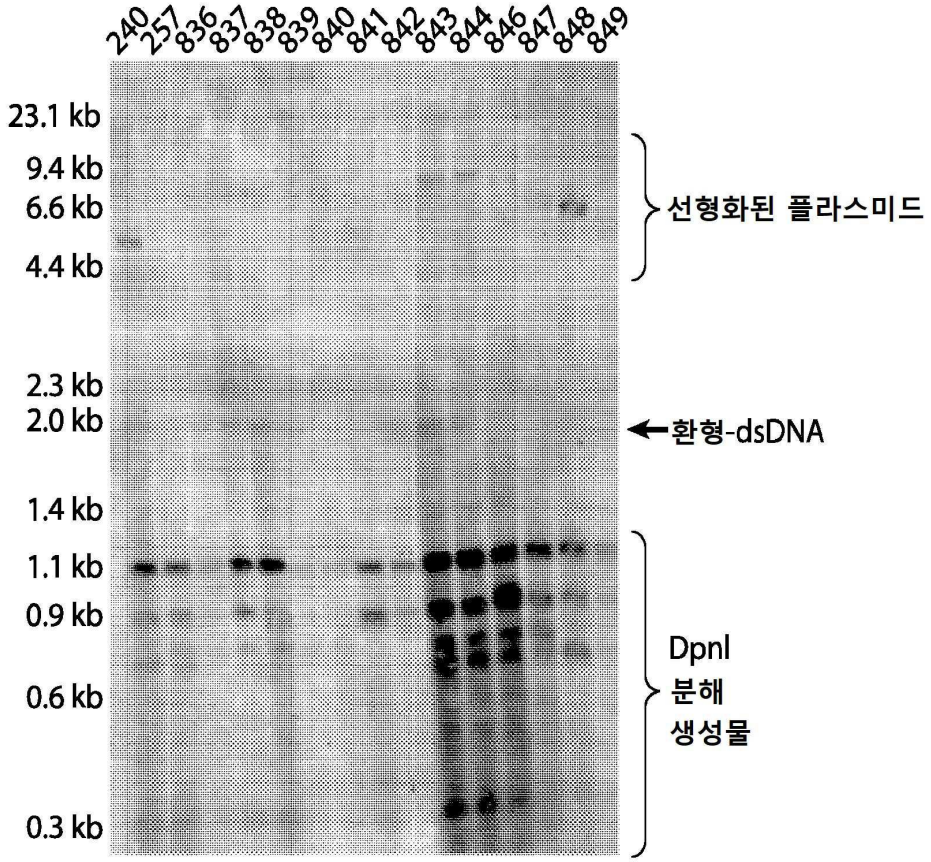
도면2b



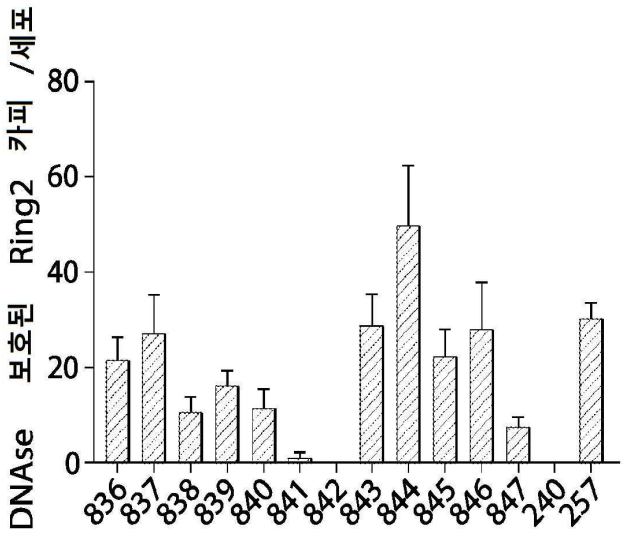
도면2c


순열 1
5CD
ORF2
인트론
ORF2-2
ORF3
ORFs
GC
5CD
순열 2
ORF2
인트론
ORF2-2
ORF3
ORFs
GC
5CD
ORF2
순열 3
인트론
ORF2-2
ORF3
ORFs
GC
5CD
ORF2
인트론
순열 4
ORF2-2
ORF3
ORFs
GC
5CD
ORF2
인트론
ORF2-2
순열 5
ORF3
ORFs
GC
5CD
ORF2
인트론
ORF2-2
ORF3
순열 6
ORFs
GC
5CD

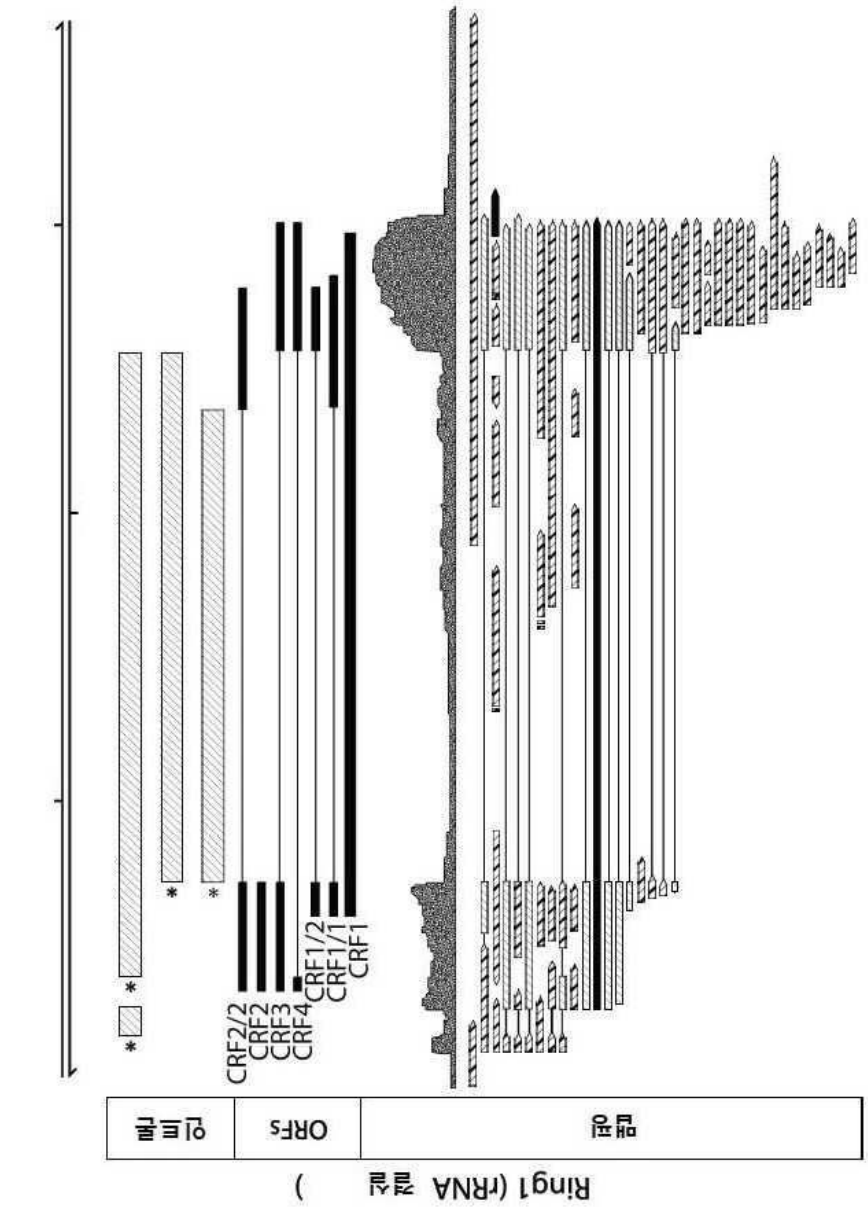
도면2d


240
257
836
837
838
839
840
841
842
843
844
846
847
848
849
23.1 kb
9.4 kb
6.6 kb
4.4 kb
2.3 kb
2.0 kb
1.4 kb
1.1 kb
0.9 kb
0.6 kb
0.3 kb
선형화된 플라스미드
환형-dsDNA
DpnI
분해
생성물

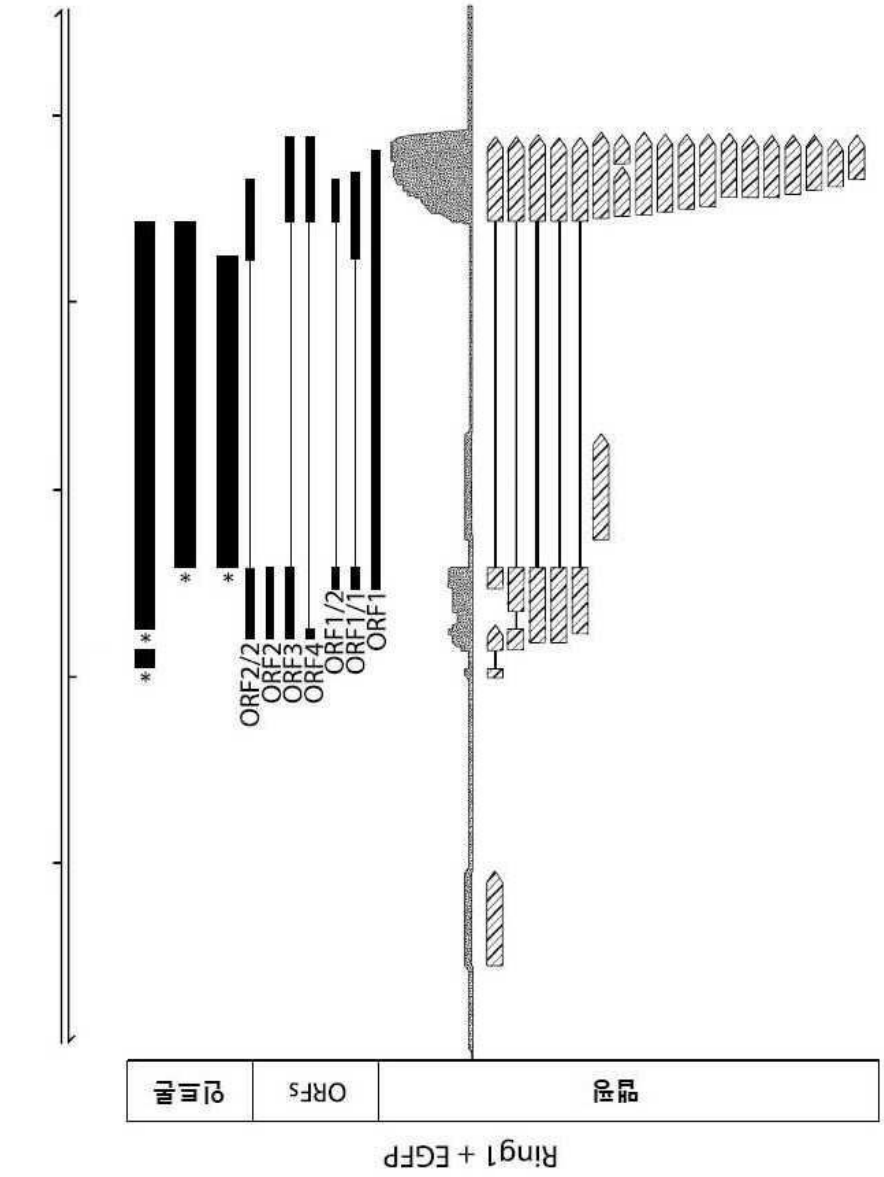
도면2e


DNAse 보호된 Ring2 카피 /세포
80
60
40
20
0
836
837
838
839
840
841
842
843
844
845
846
847
240
257

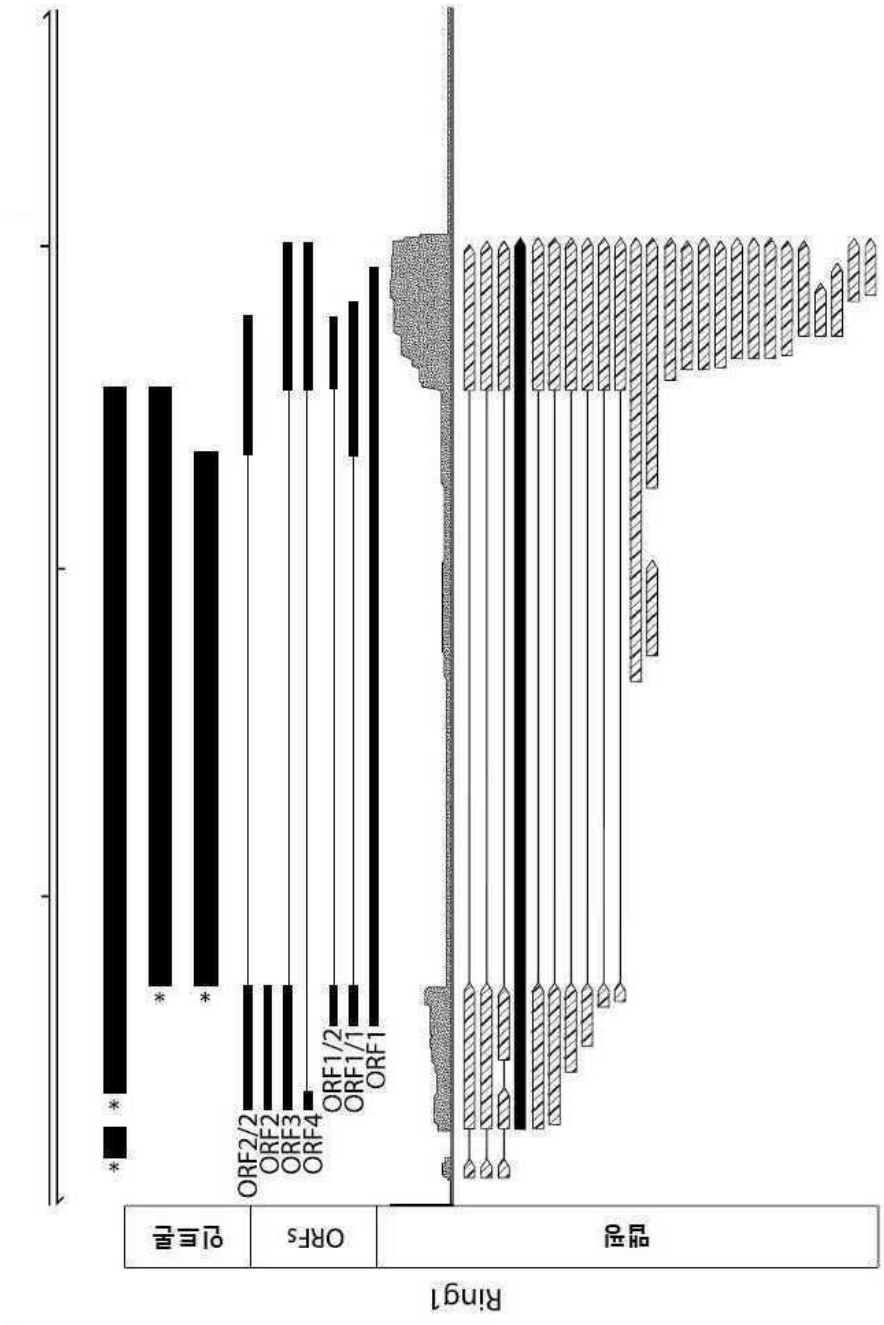
도면2fa


Ring1 (rRNA 결실)
인트론
ORFs
맵핑
CRF2/2
CRF2
CRF3
CRF4
CRF1/2
CRF1/1
CRF1

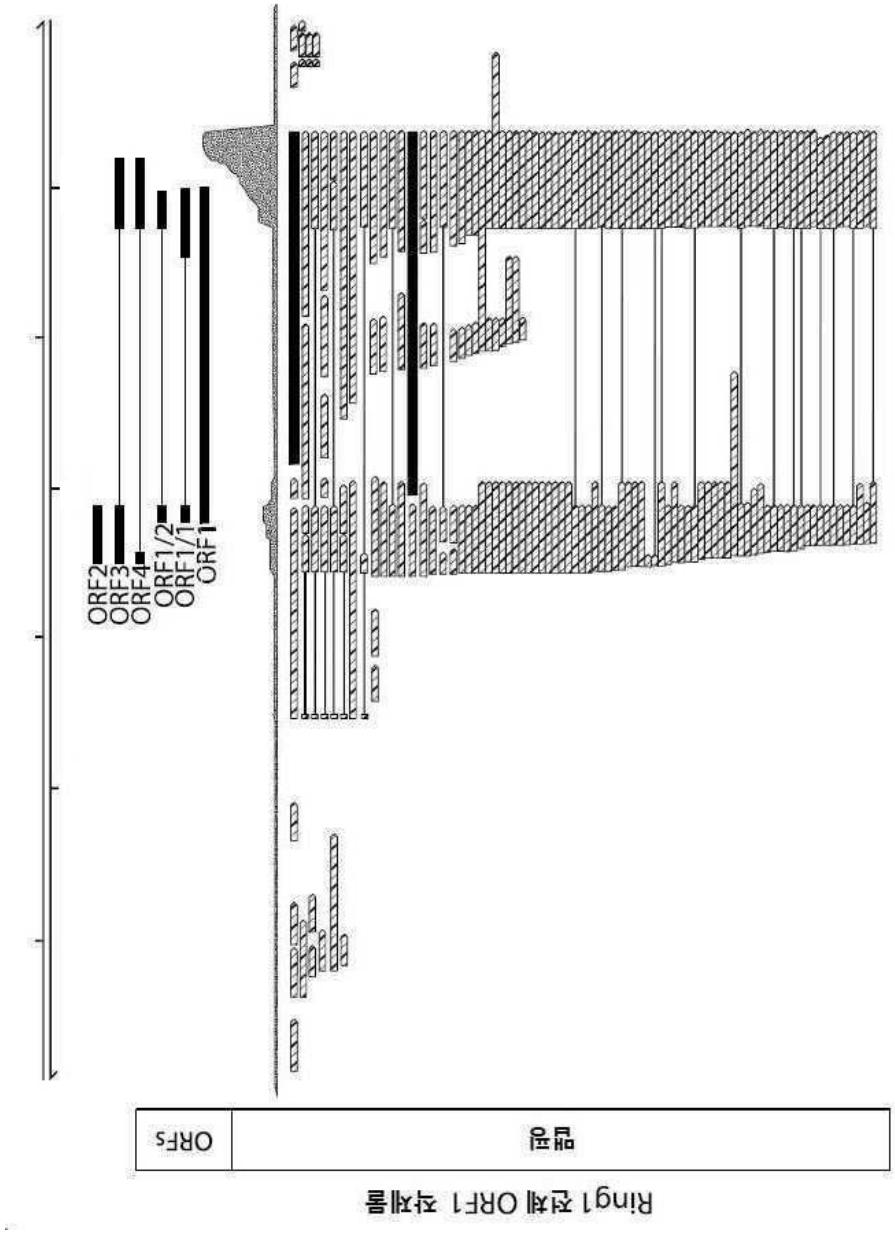
도면2fb



도면2fc

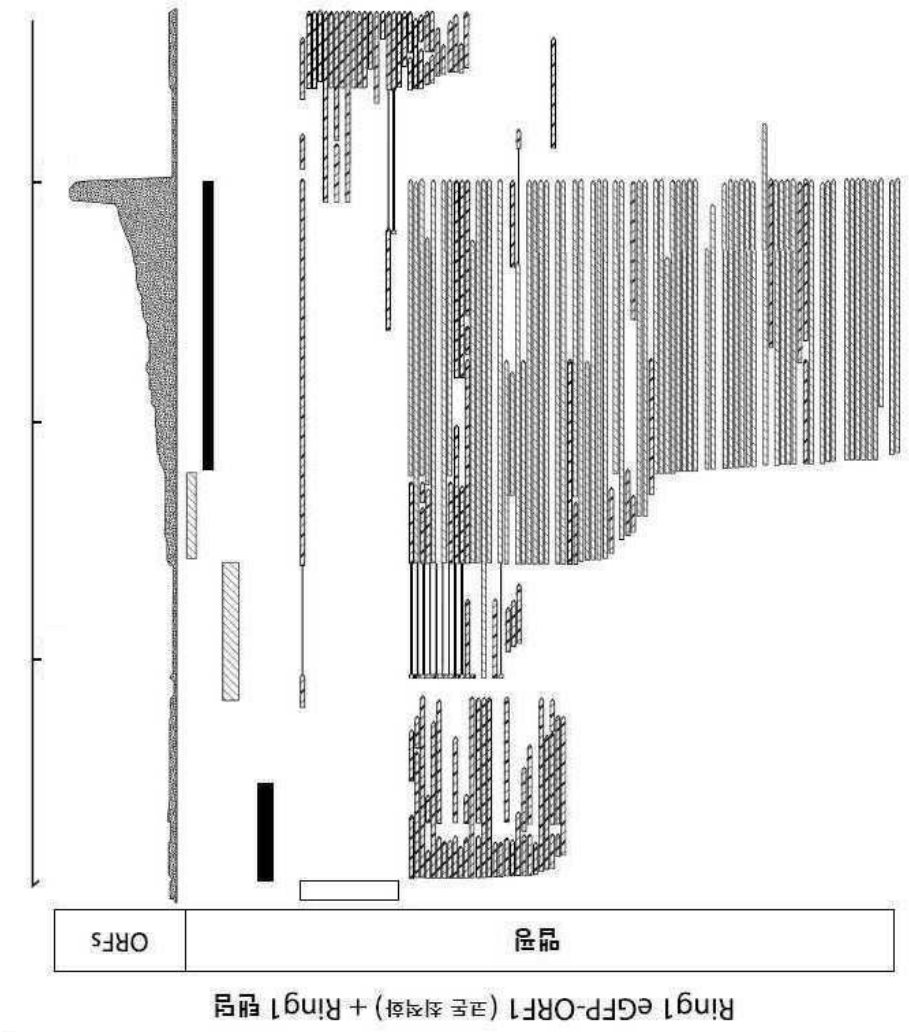


도면2fd



Ring1 전체 ORF1 작제물

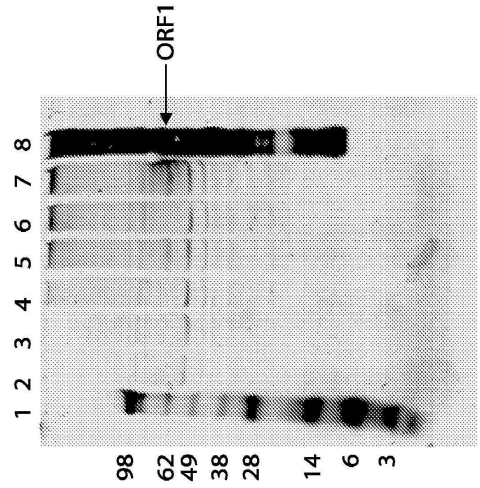
도면2fe



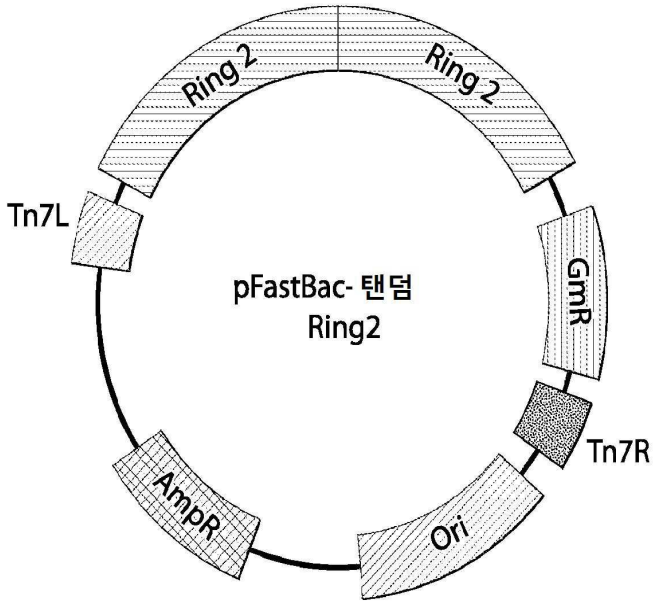
도면2g

웨스턴 블롯 ORF1 발현

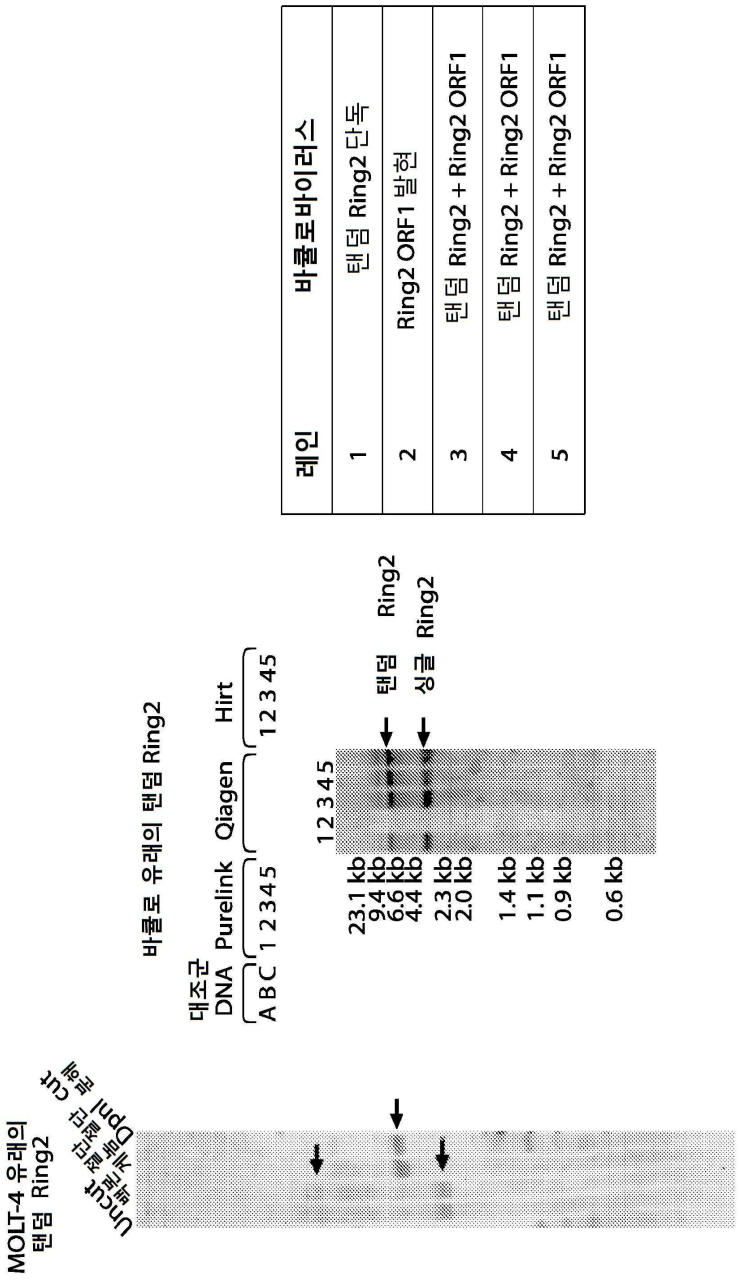
레인 번호	샘플
1	래더
2	제1일 비-복제
3	제1일 복제
4	제2일 비-복제
5	제2일 복제
6	제3일 비-복제
7	제3일 복제
8	Ring2 ORF1 바큘로 양성 대조군



도면2h

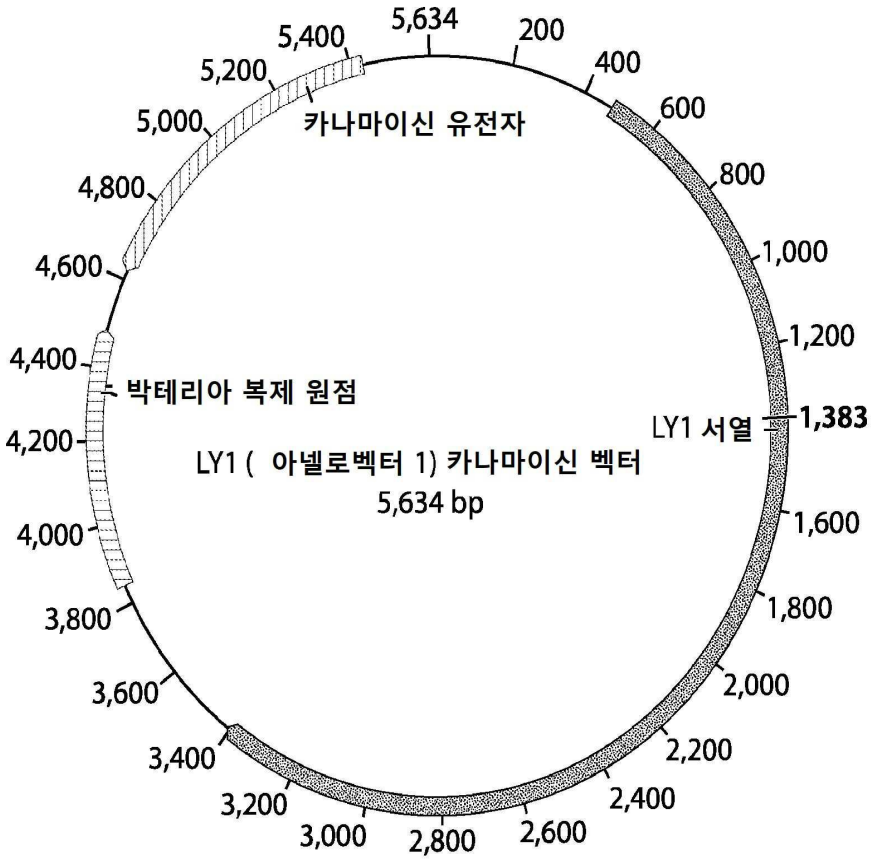


도면2i



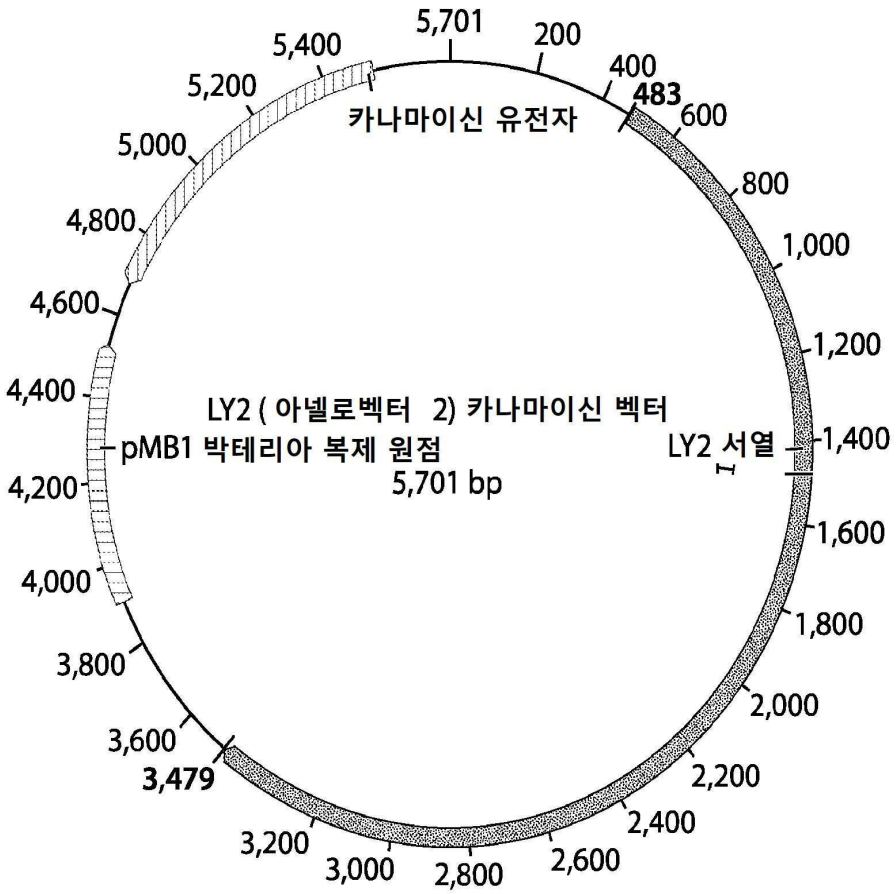
레인	바큘로바이러스
1	탠덤 Ring2 단독
2	Ring2 ORF1 발현
3	탠덤 Ring2 + Ring2 ORF1
4	탠덤 Ring2 + Ring2 ORF1
5	탠덤 Ring2 + Ring2 ORF1

도면3



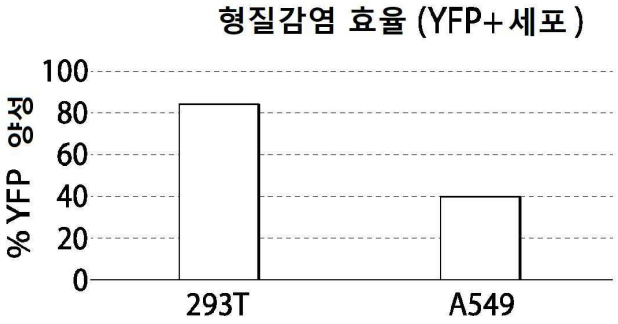
TTMiniV의 LY1 균주를 인코딩하는 카나마이신 벡터의 개략도
(아넬로좀 1)

도면4

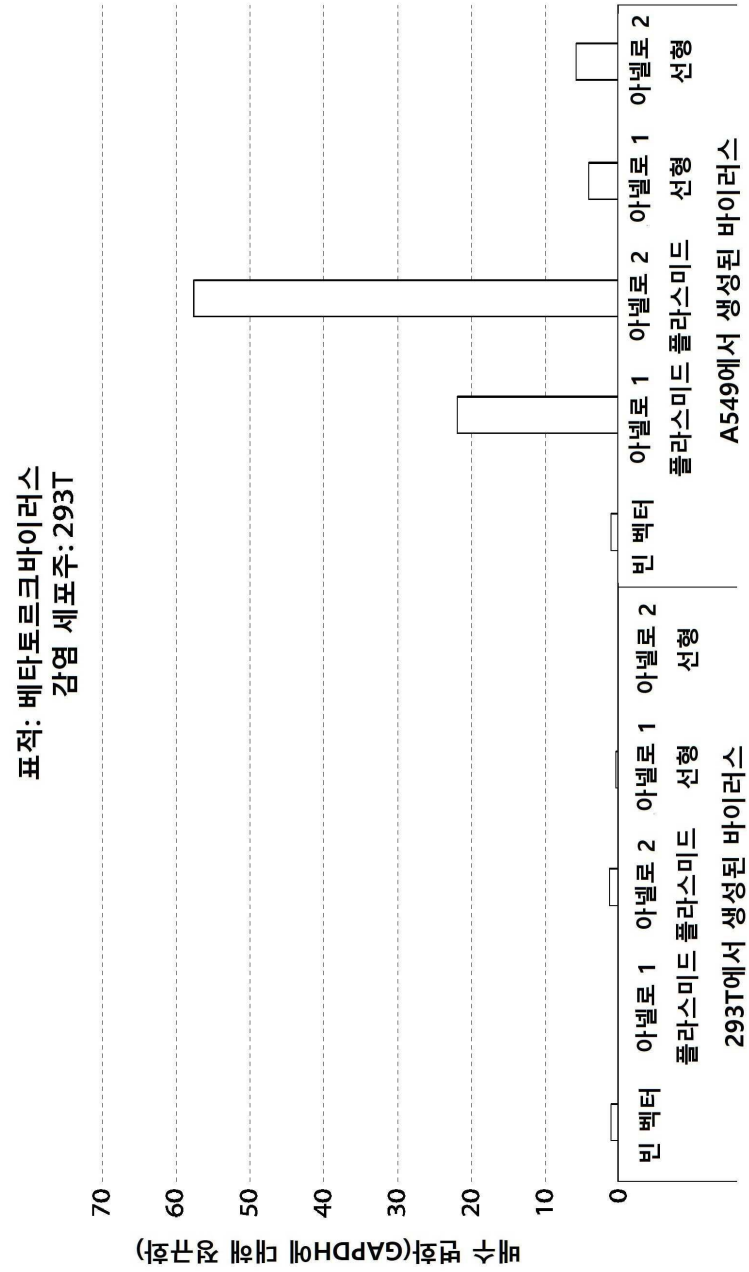


TTMiniV의 LY2 균주를 인코딩하는 카나마이신 벡터의 개략도 (아넬로좀 2)

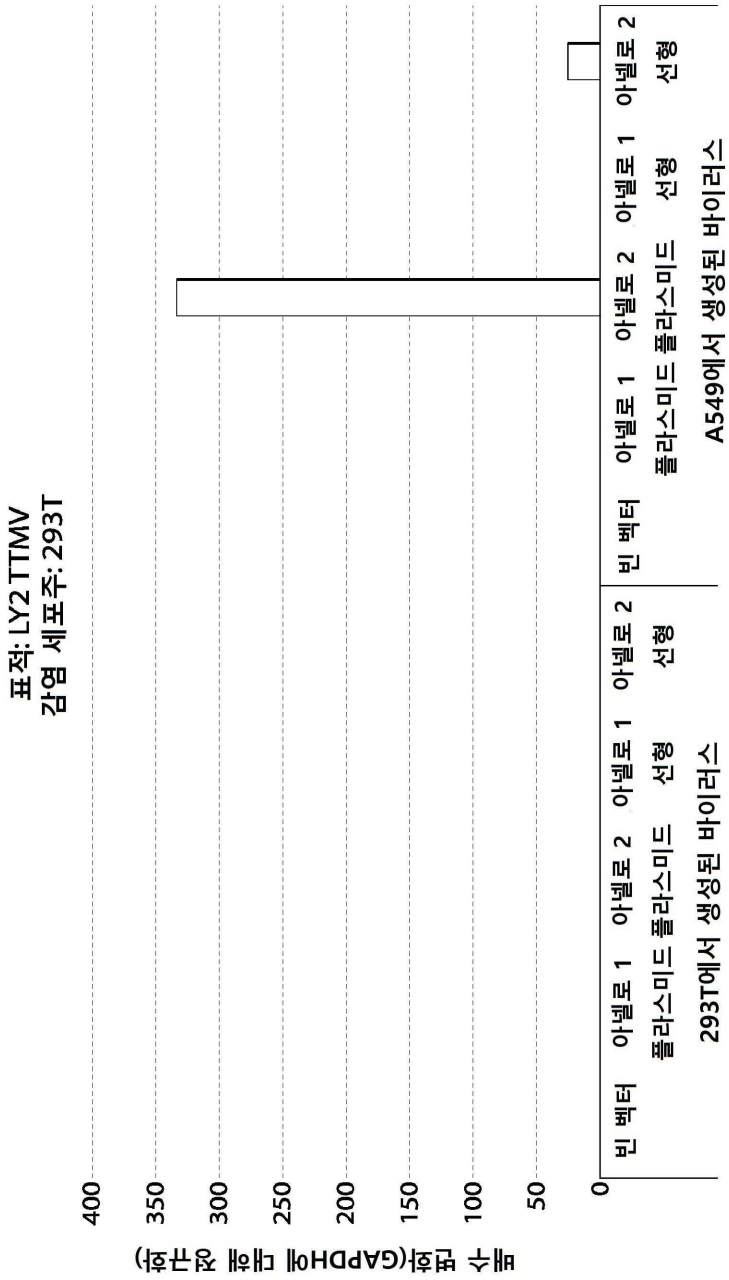
도면5



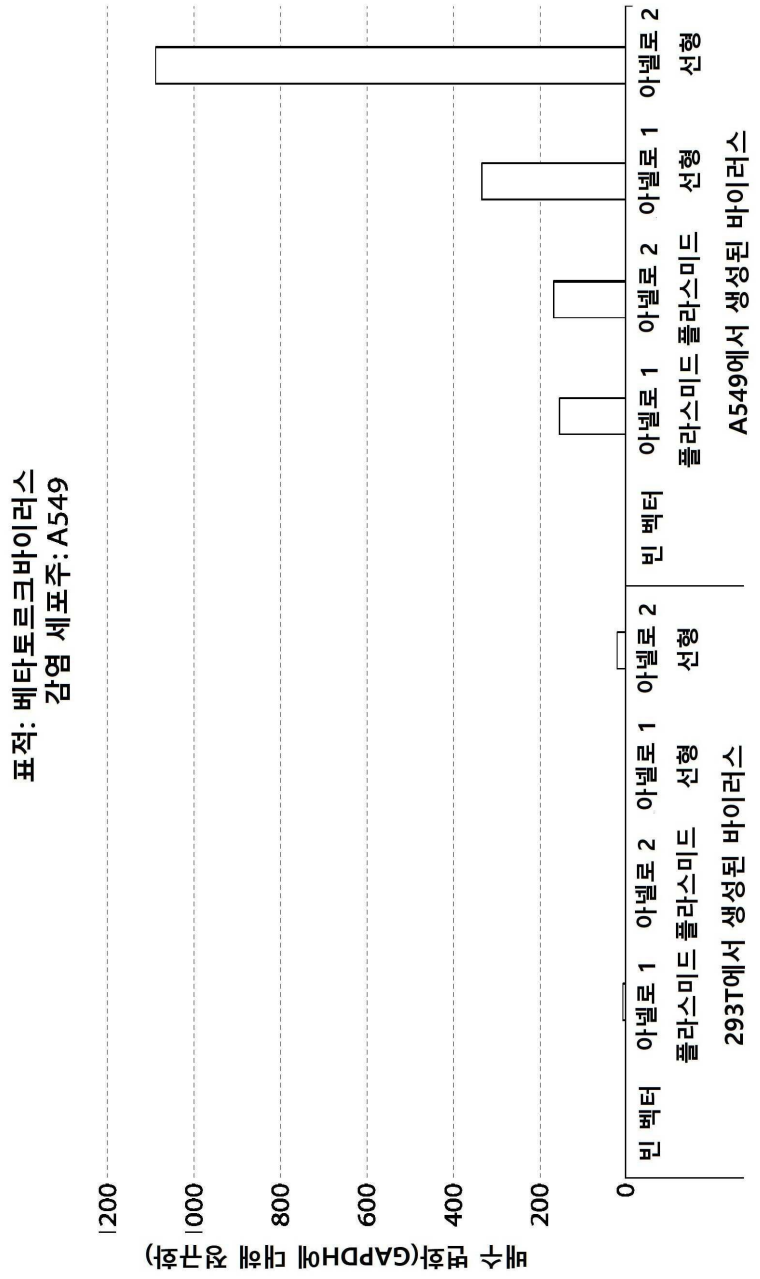
도면6a


표적: 베타토르크바이러스
감염 세포주: 293T
복수 변화(GAPDH에 대해 정규화)
70
60
50
40
30
20
10
0
빈 벡터
아넬로 1
아넬로 2
아넬로 1 플라스미드
아넬로 2 플라스미드
선형
293T에서 생성된 바이러스
빈 벡터
아넬로 1
아넬로 2
아넬로 1 플라스미드
아넬로 2 플라스미드
선형
A549에서 생성된 바이러스

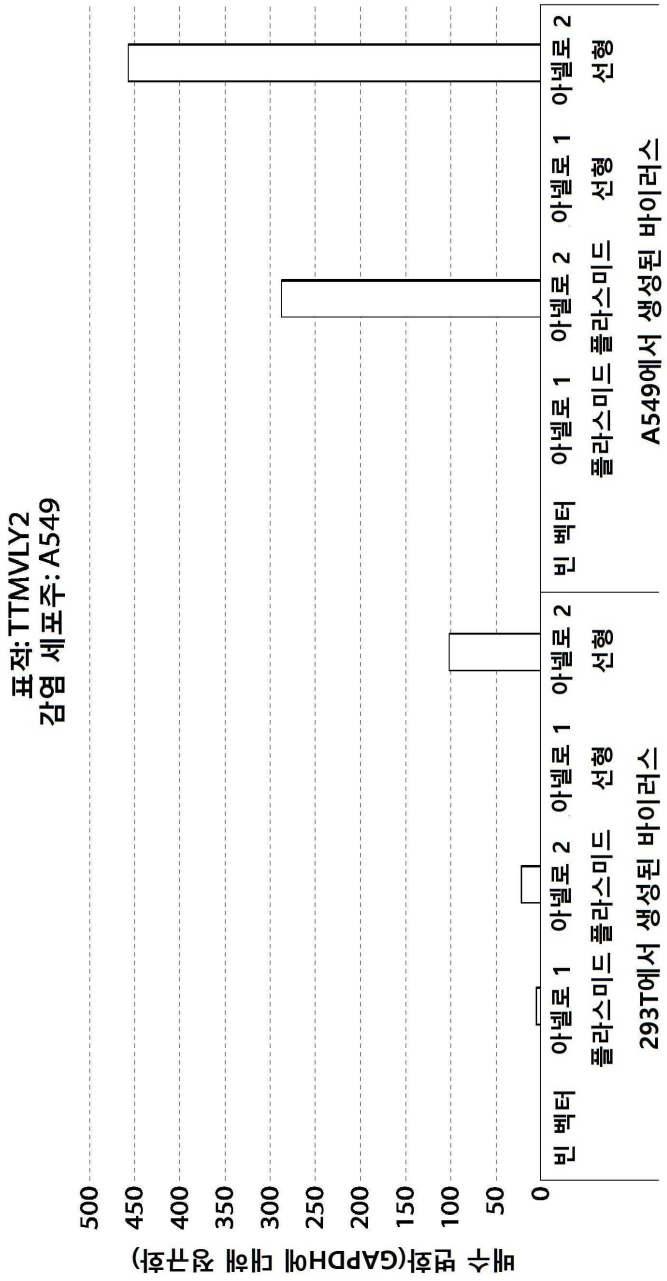
도면6b


표적: LY2 TTMV
감염 세포주: 293T
뷸수 변화(GAPDH에 대해 정규화)
400
350
300
250
200
150
100
50
0
빈 벡터
아넬로 1
아넬로 2
아넬로 1 선형
아넬로 2 선형
293T에서 생성된 바이러스
빈 벡터
아넬로 1
아넬로 2
아넬로 1 선형
아넬로 2 선형
A549에서 생성된 바이러스

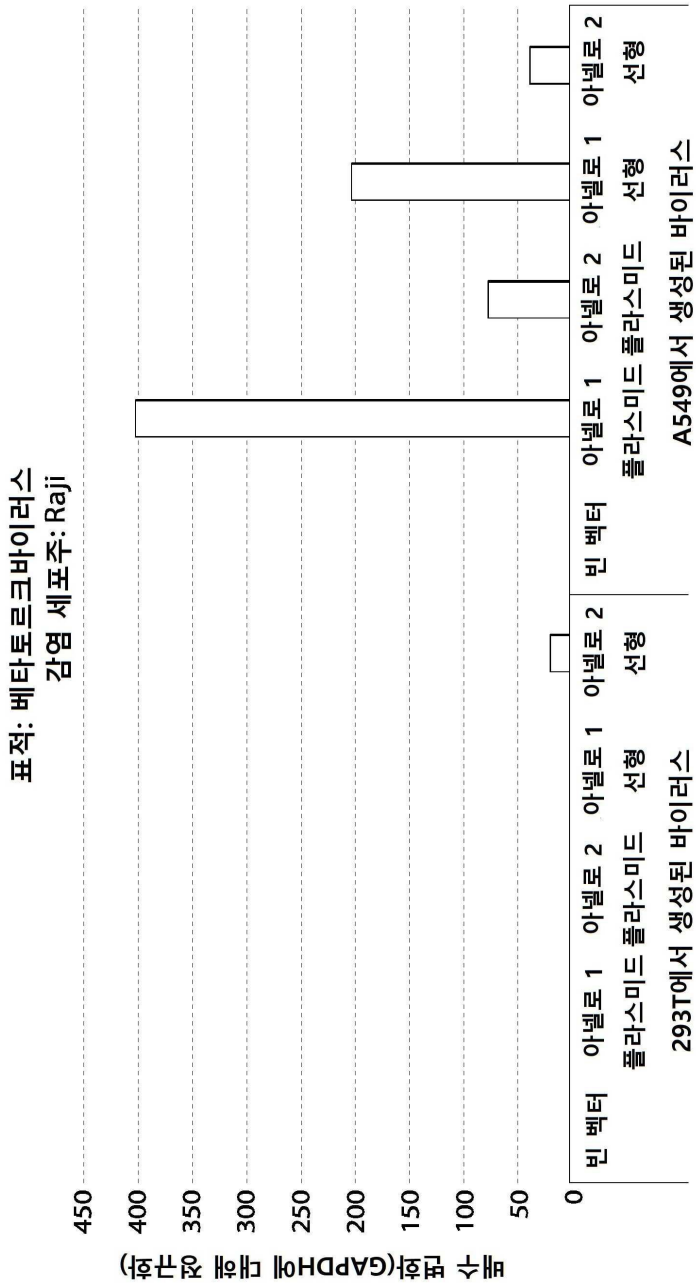
도면7a


표적: 베타토르크바이러스
감염 세포주: A549
1200
1000
800
600
400
200
0
빈 벡터
아넬로 1
아넬로 2
아넬로 1 플라스미드
아넬로 2 플라스미드
293T에서 생성된 바이러스
빈 벡터
아넬로 1
아넬로 2
아넬로 1 플라스미드
아넬로 2 플라스미드
A549에서 생성된 바이러스
선형
선형
선형
선형
복수 변화(GAPDH에 대해 정규화)

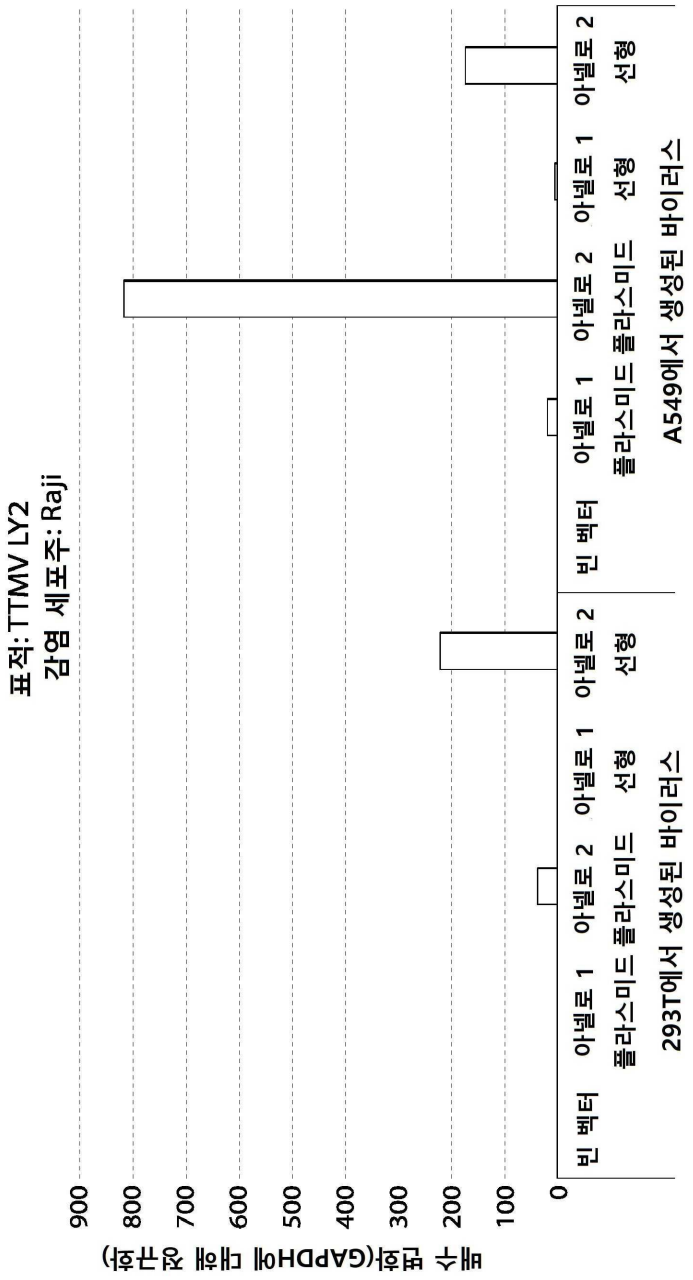
도면7b


표적: TTMVLY2
감염 세포주: A549
뷰수 변화(GAPDH에 대해 정규화)
500
450
400
350
300
250
200
150
100
50
0
빈 벡터
아넬로 1
아넬로 2 플라스미드
아넬로 1 선형
아넬로 2 선형
293T에서 생성된 바이러스
빈 벡터
아넬로 1
아넬로 2 플라스미드
아넬로 2
아넬로 1 선형
아넬로 2 선형
A549에서 생성된 바이러스

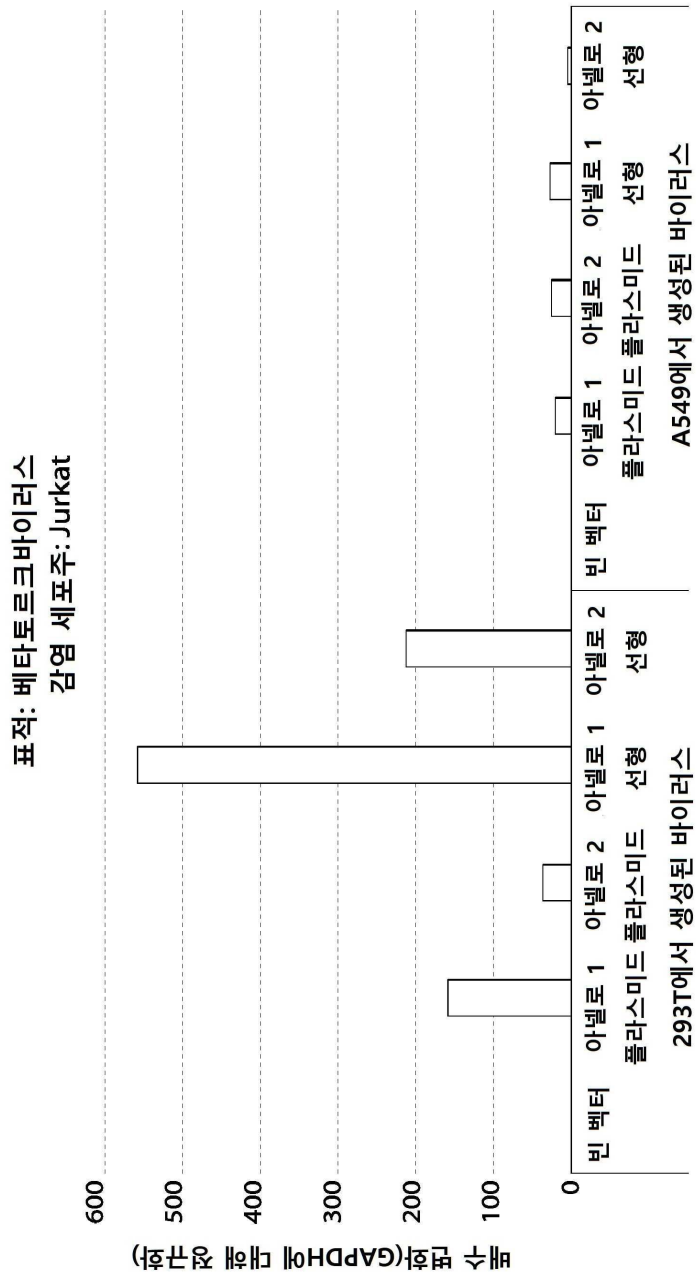
도면8a


표적: 베타토르크바이러스
감염 세포주: Raji
뉴클레오티드 변화(GAPDH에 대해 정규화)
450
400
350
300
250
200
150
100
50
0
빈 벡터
아넬로 1 폴라스미드
아넬로 2 폴라스미드
아넬로 1 선형
아넬로 2 선형
293T에서 생성된 바이러스
빈 벡터
아넬로 1 폴라스미드
아넬로 2 폴라스미드
아넬로 1 선형
아넬로 2 선형
A549에서 생성된 바이러스

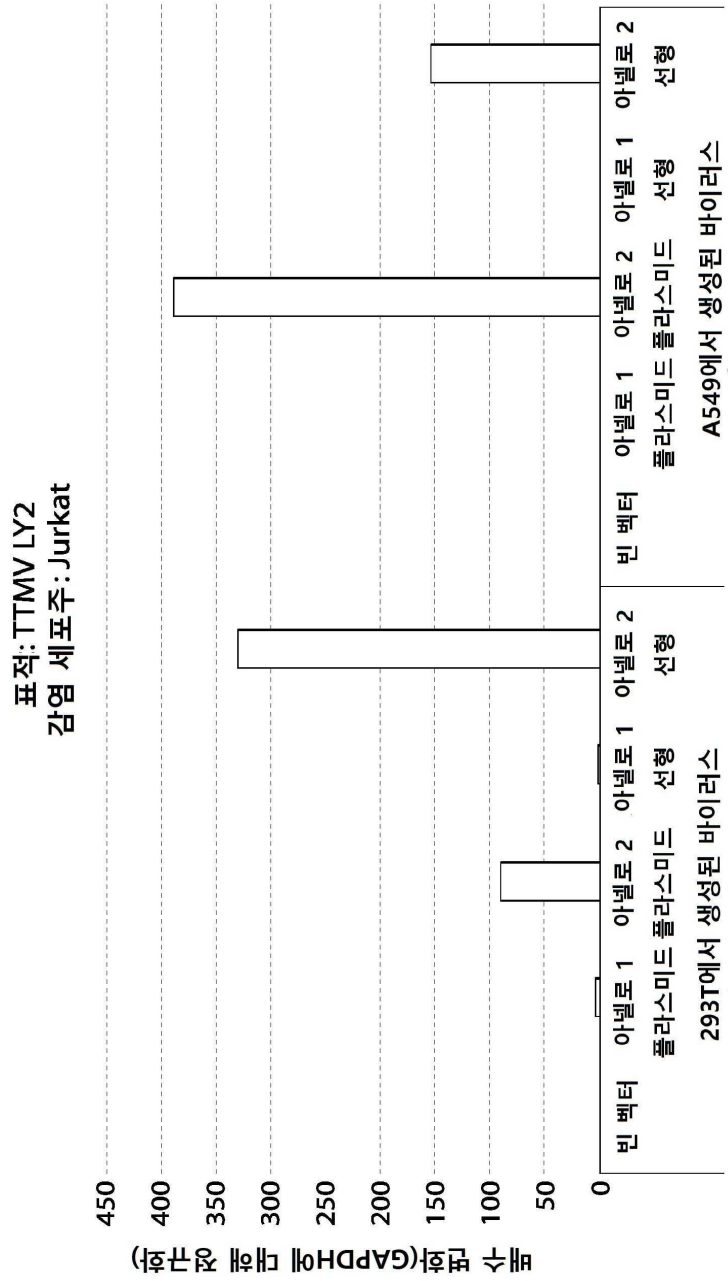
도면8b


표적: TTMV LY2
감염 세포주: Raji
복수 변화(GAPDH에 대해 정규화)
900
800
700
600
500
400
300
200
100
0
빈 벡터
아넬로 1
아넬로 2 플라스미드
아넬로 1 선형
아넬로 2 선형
293T에서 생성된 바이러스
빈 벡터
아넬로 1
아넬로 2 플라스미드
아넬로 1 선형
아넬로 2 선형
A549에서 생성된 바이러스

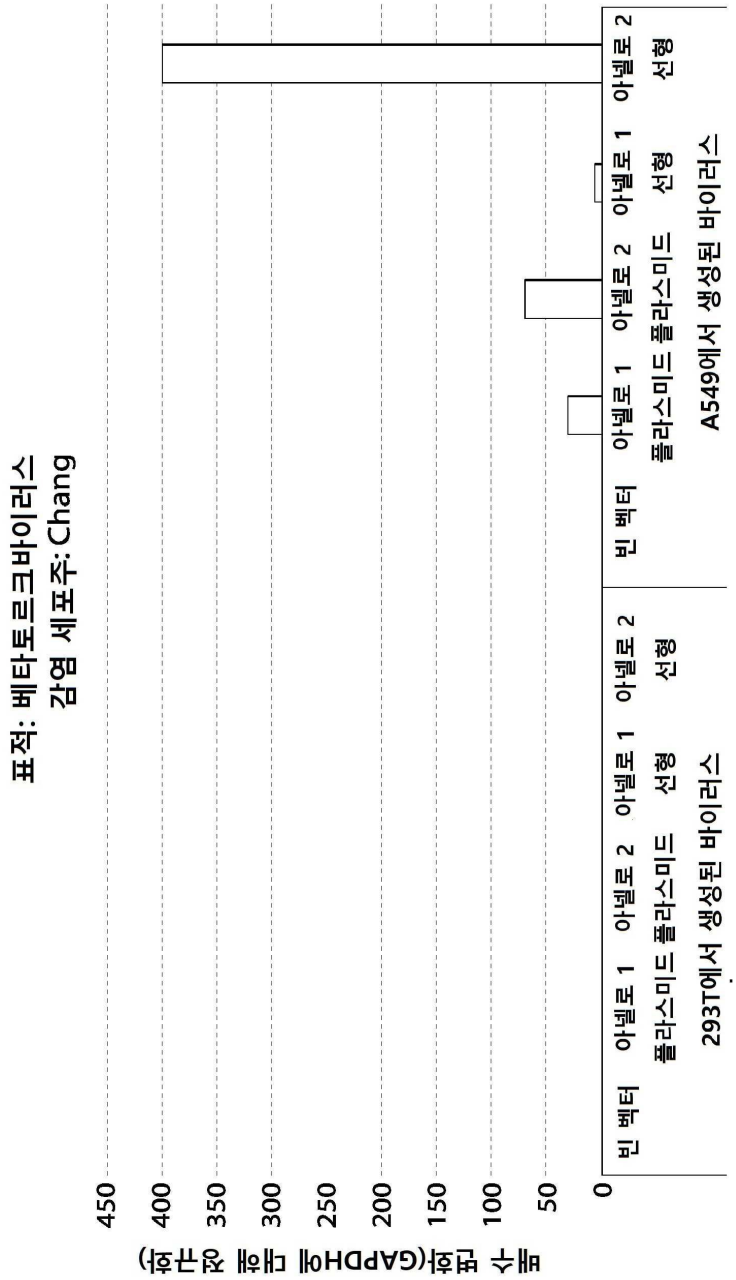
도면9a


표적: 베타토르크바이러스
감염 세포주: Jurkat
뉴클레오캡시드 단백질 발현량 (GAPDH에 대한 정규화)
600
500
400
300
200
100
0
빈 벡터
아넬로 1 플라스미드
아넬로 2 플라스미드
아넬로 1 선형
아넬로 2 선형
293T에서 생성된 바이러스
빈 벡터
아넬로 1 플라스미드
아넬로 2 플라스미드
아넬로 1 선형
아넬로 2 선형
A549에서 생성된 바이러스

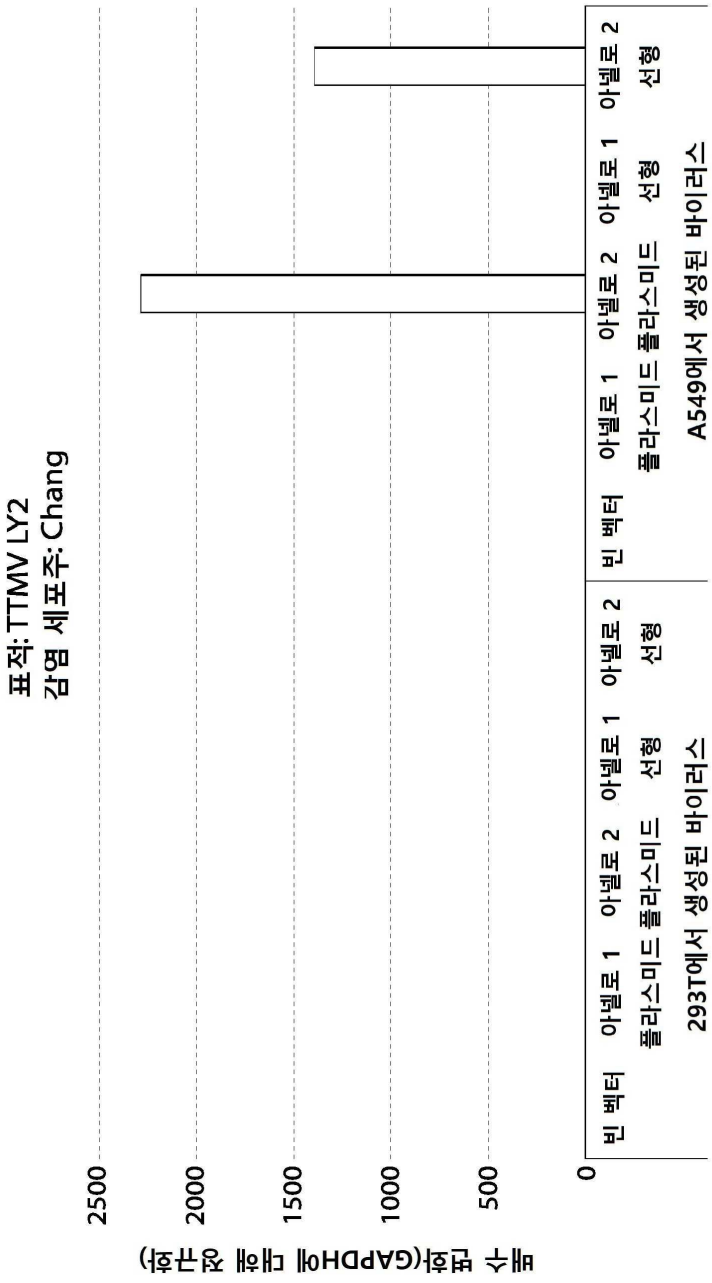
도면9b


표적: TTMV LY2
감염 세포주: Jurkat
복수 변화(GAPDH에 대해 정규화)
450
400
350
300
250
200
150
100
50
0
빈 벡터
아넬로 1
아넬로 2 플라스미드
아넬로 1 선형
아넬로 2 선형
293T에서 생성된 바이러스
빈 벡터
아넬로 1
아넬로 2 플라스미드
아넬로 2
아넬로 1 선형
아넬로 2 선형
A549에서 생성된 바이러스

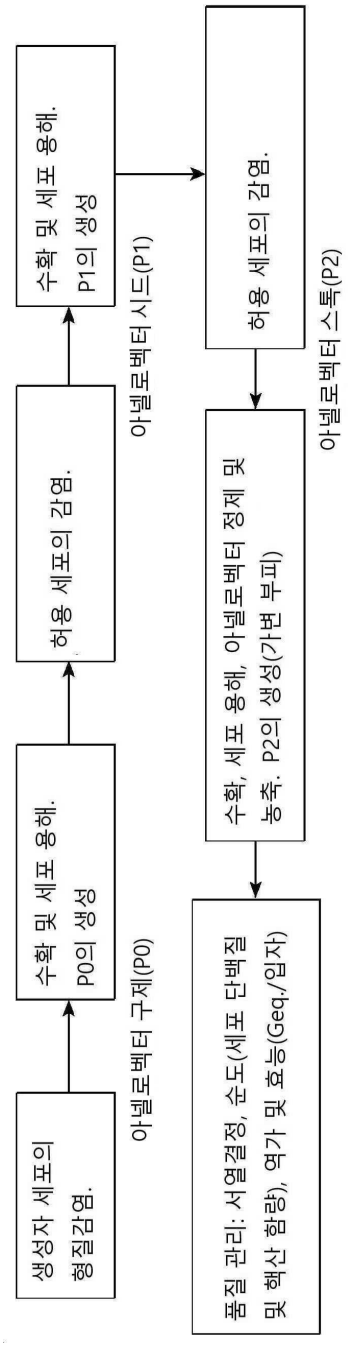
도면10a


표적: 베타토르크바이러스
감염 세포주: Chang
벡수 변화(GAPDH에 대해 정규화)
450
400
350
300
250
200
150
100
50
0
빈 벡터
아넬로 1 폴라스미드
아넬로 2 폴라스미드
아넬로 1 선형
아넬로 2 선형
293T에서 생성된 바이러스
빈 벡터
아넬로 1 폴라스미드
아넬로 2 폴라스미드
아넬로 1 선형
아넬로 2 선형
A549에서 생성된 바이러스

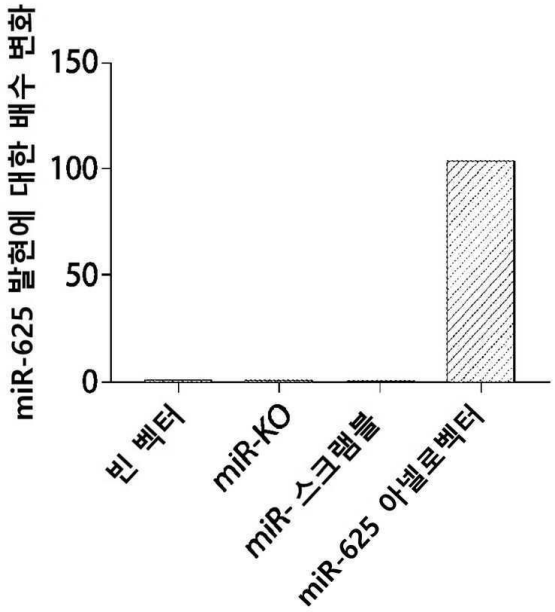
도면10b


표적: TTMV LY2
감염 세포주: Chang
벡수 변화(GAPDH에 대해 정규화)
2500
2000
1500
1000
500
0
빈 벡터
아넬로 1 플라스미드
아넬로 2 플라스미드
아넬로 1 선형
아넬로 2 선형
293T에서 생성된 바이러스
빈 벡터
아넬로 1 플라스미드
아넬로 2 플라스미드
아넬로 1 선형
아넬로 2 선형
A549에서 생성된 바이러스

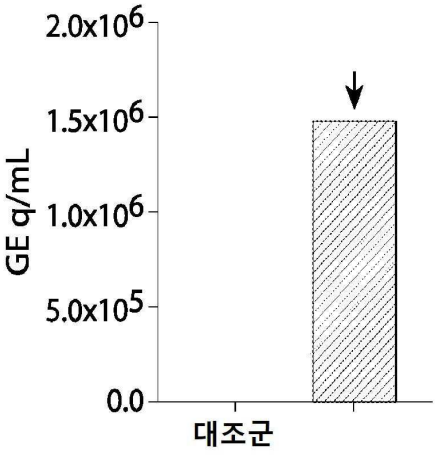
도면11


생성자 세포의
형질감염.
수확 및 세포 용해.
P0의 생성
허용 세포의 감염.
수확 및 세포 용해.
P1의 생성
아넬로벡터 구제(P0)
아넬로벡터 시드(P1)
품질 관리: 서열결정, 순도(세포 단백질
및 핵산 함량), 역가 및 효능(Geq./입자)
수확, 세포 용해, 아넬로벡터 정제 및
농축. P2의 생성(가변 부피)
허용 세포의 감염.
아넬로벡터 스톡(P2)

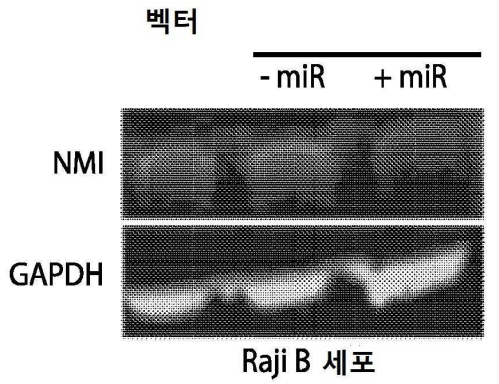
도면12


miR-625 발현에 대한 배수 변화
150
100
50
0
빈 벡터
miR-KO
miR-스크램블
miR-625 아넬로벡터

도면13

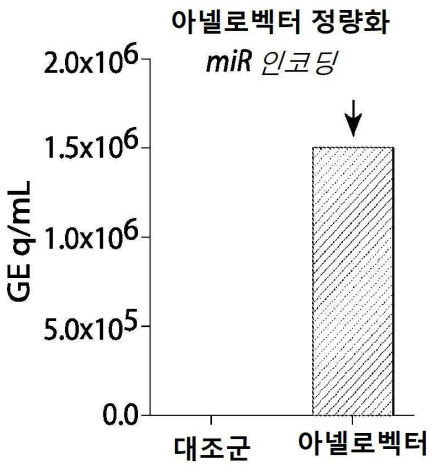

GE q/mL
2.0x10^6
1.5x10^6
1.0x10^6
5.0x10^5
0.0
대조군

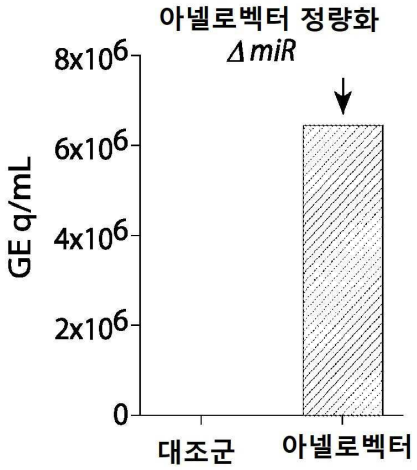
도면14



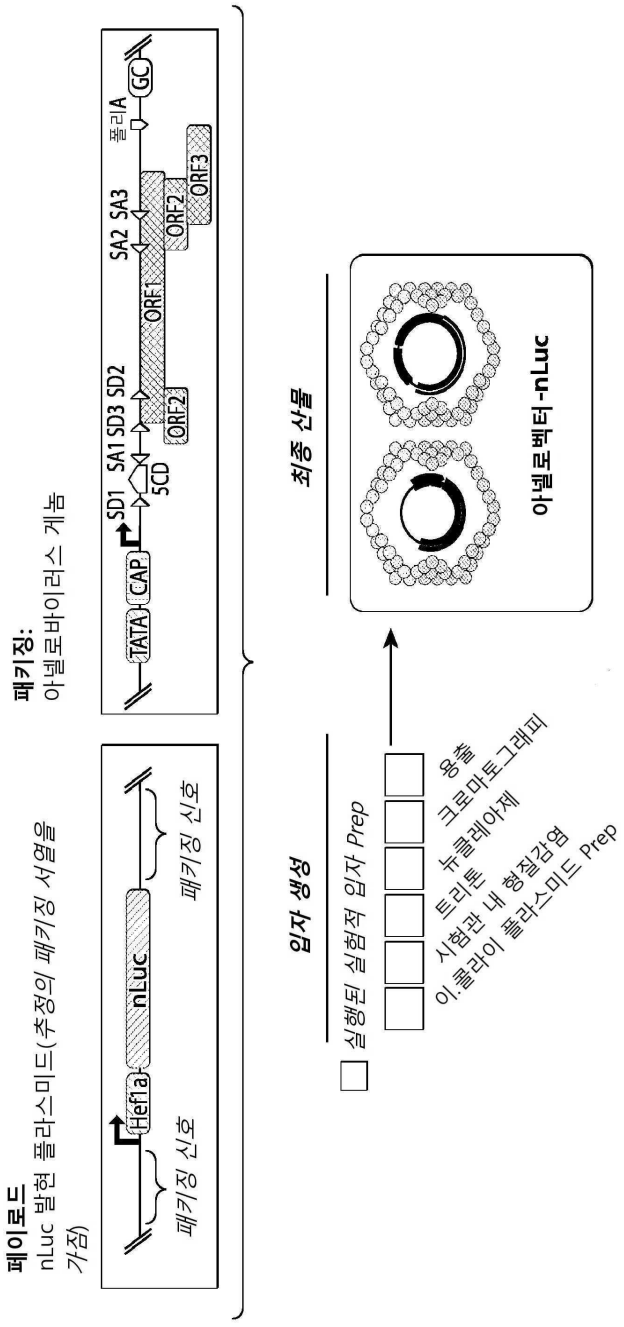
단백질 발현의 75% 초과 감소

도면15

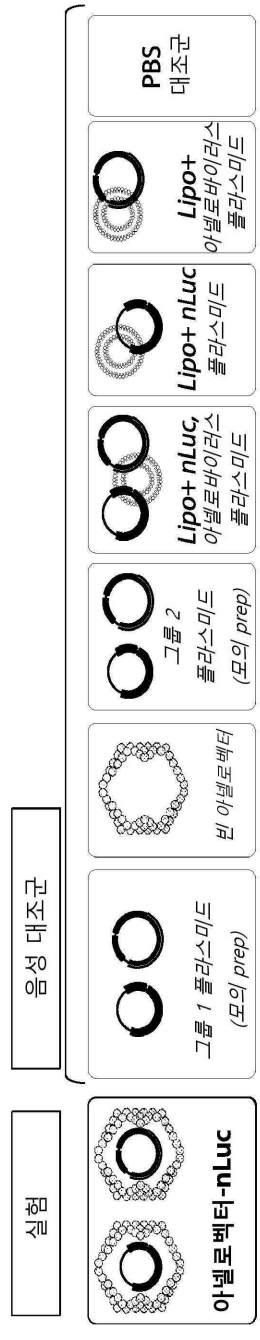




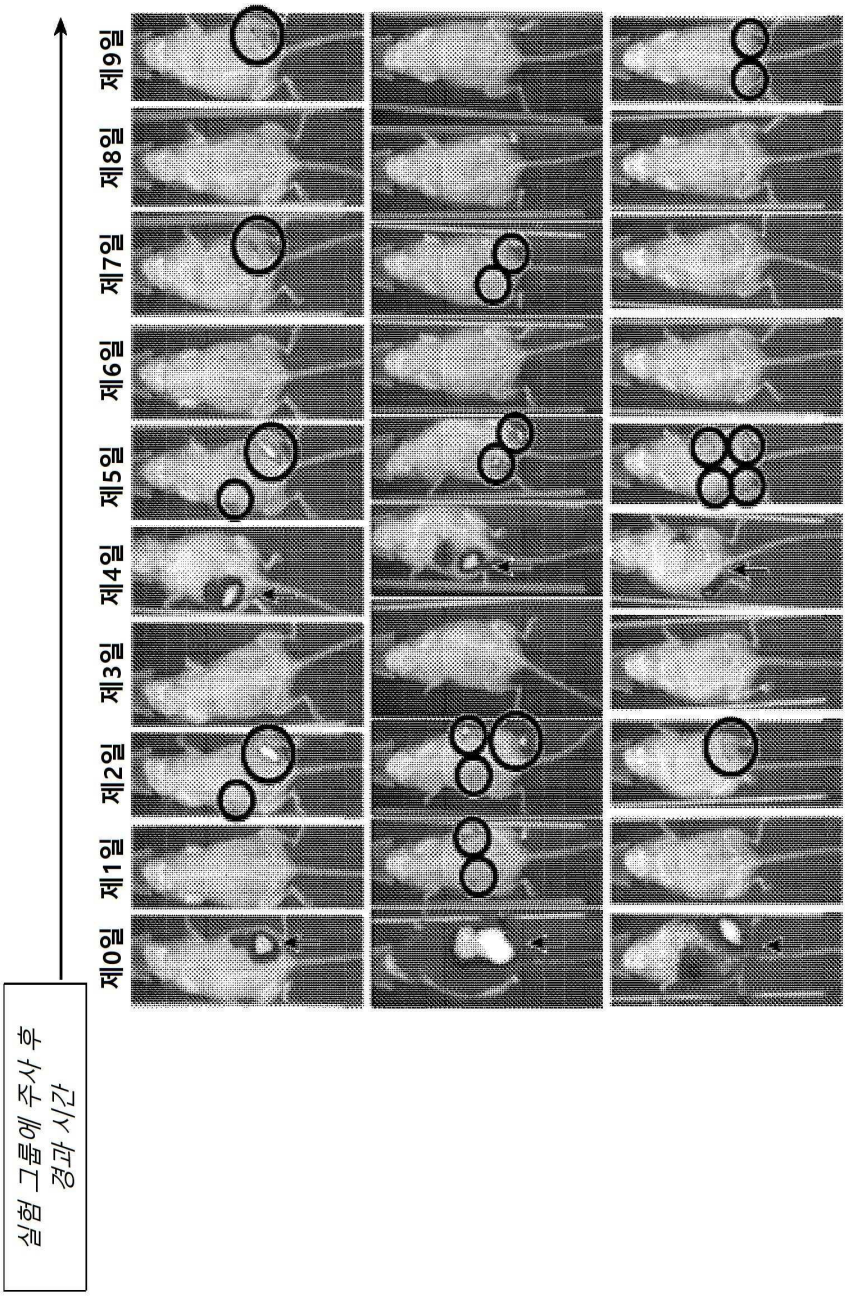
도면16a


페이로드
nLuc 발현 플라스미드(추정의 패키징 서열을 가짐)
패키징 신호
Hef1a
nLuc
패키징 신호
패키징:
아넬로바이러스 게놈
TATA
CAP
SD1
SA1
SD3
SD2
5CD
ORF2
ORF1
ORF2
SA2
SA3
ORF3
폴리A
GC
입자 생성
실행된 실험적 입자 Prep
옹출
크로마토그래피
뉴클레아제
트리톤
시험관 내 형질감염
이.콜라이 플라스미드 Prep
최종 산물
아넬로벡터-nLuc

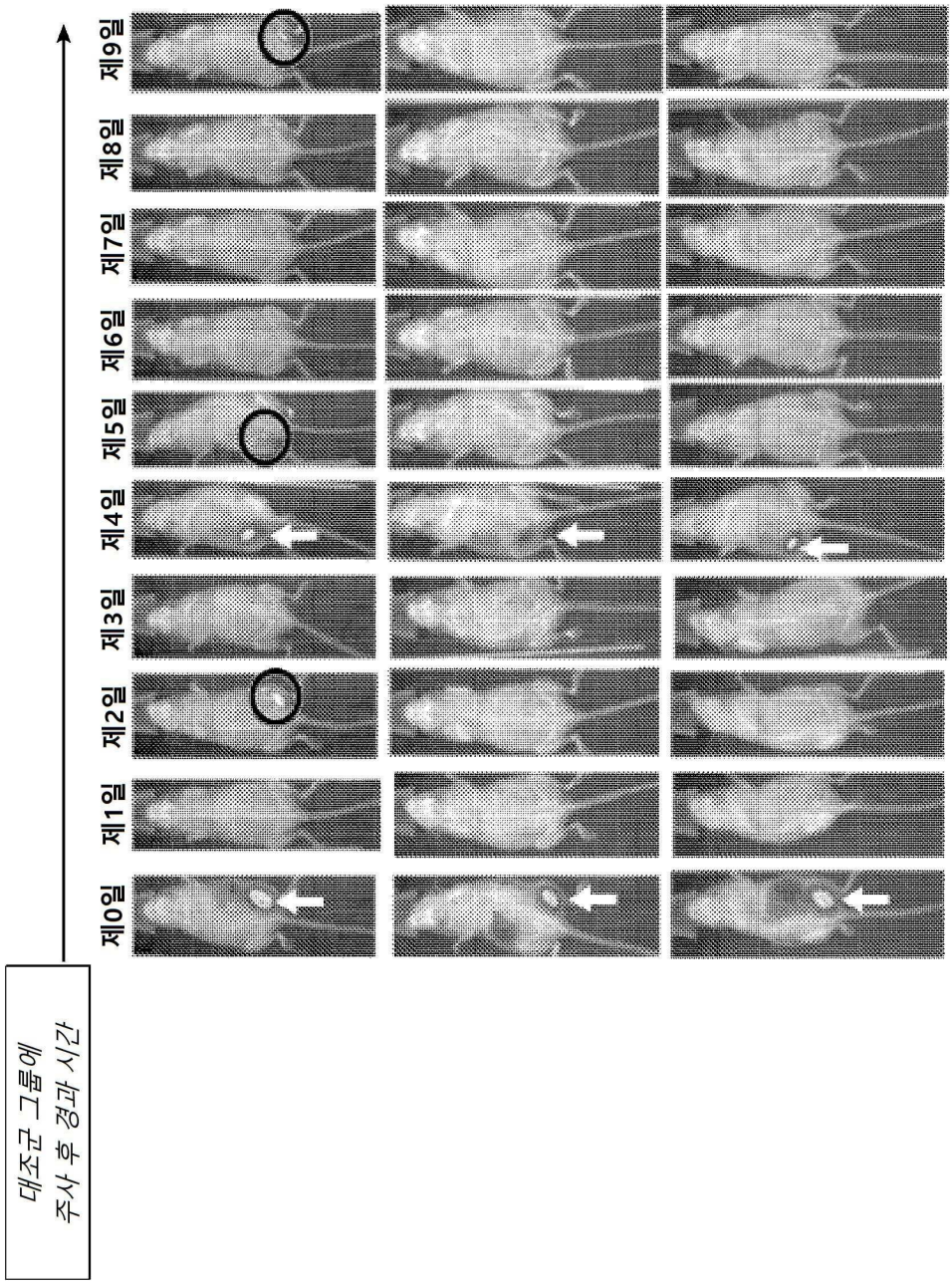
도면16b


실험
아넬로 벡터-nLuc
음성 대조군
그룹 1 플라스미드 (모의 prep)
비 아넬로 벡터
그룹 2 플라스미드 (모의 prep)
Lipo+ nLuc, 아넬로바이러스 플라스미드
Lipo+ nLuc 플라스미드
Lipo+ 아넬로바이러스 플라스미드
PBS 대조군

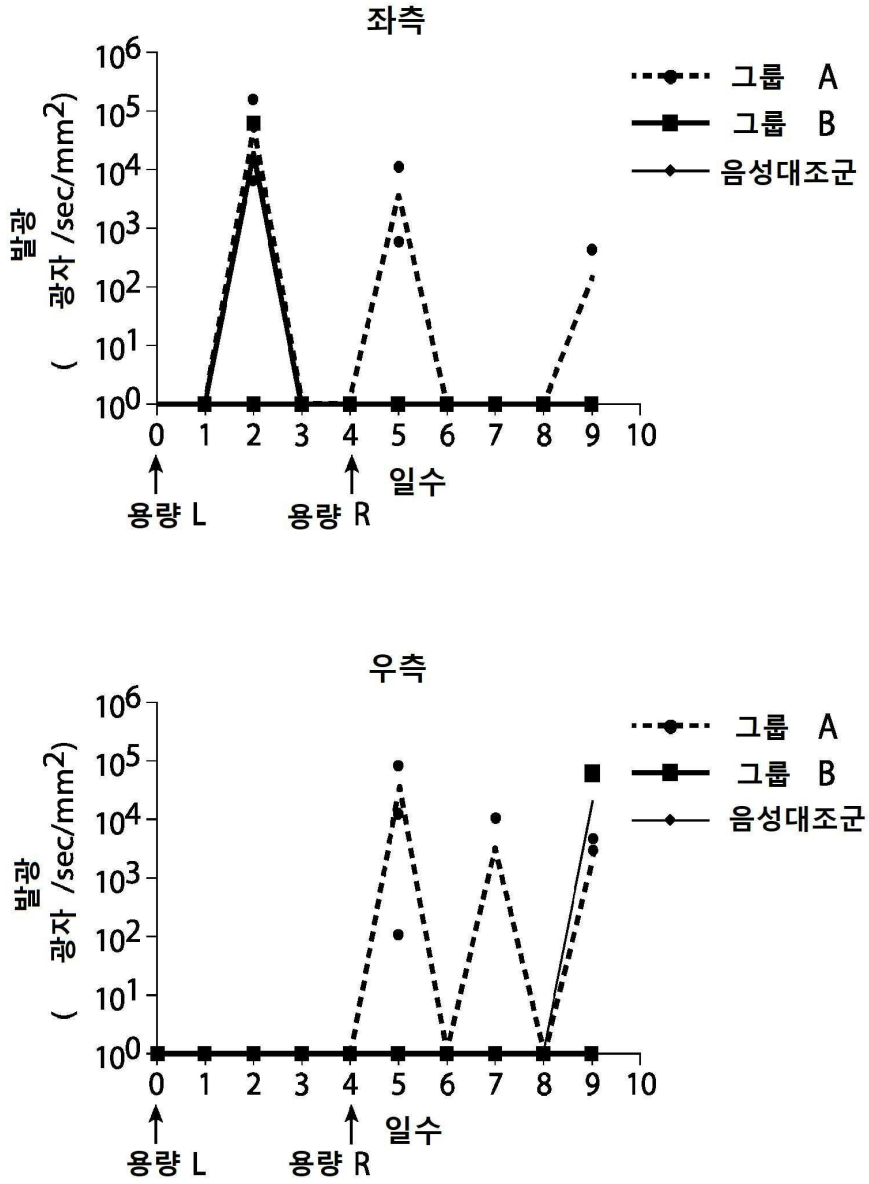
도면17a



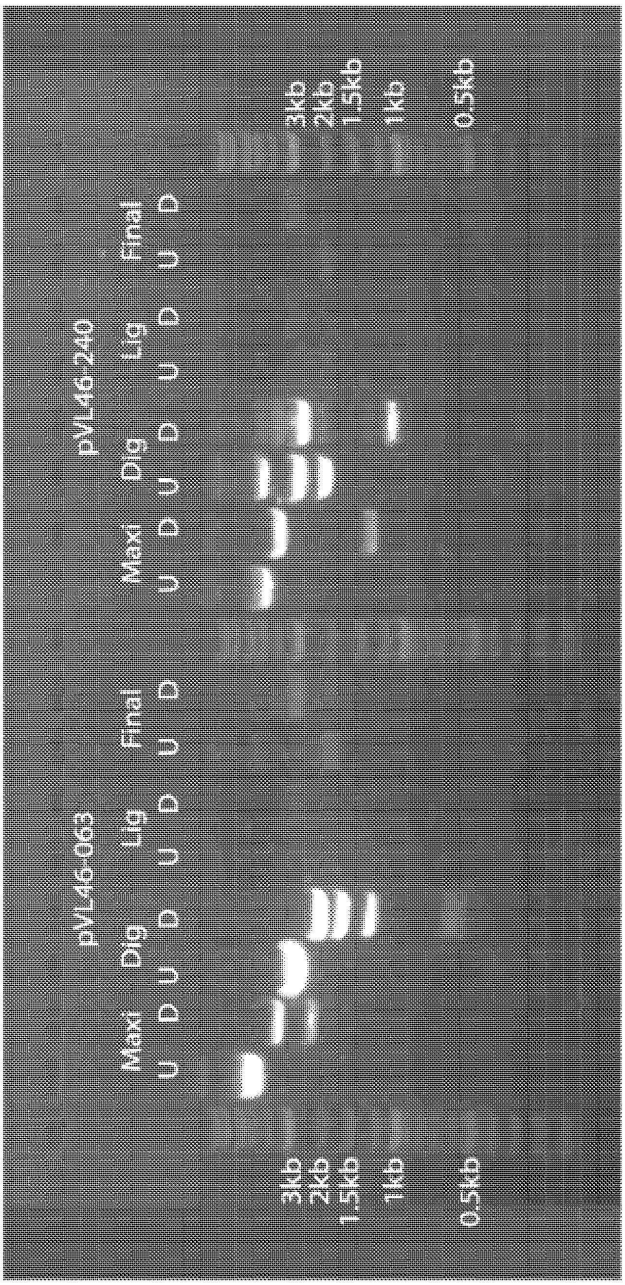
도면17b



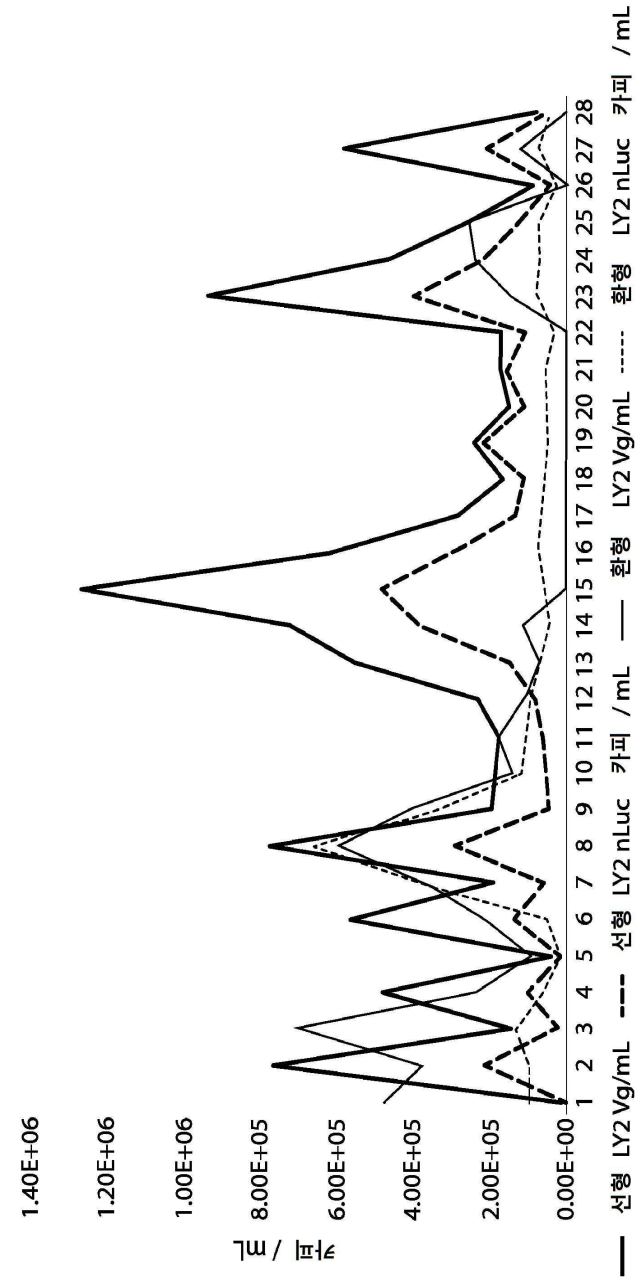
도면17c


좌측
그룹 A
그룹 B
음성대조군
발광 (광자 /sec/mm2)
10^6
10^5
10^4
10^3
10^2
10^1
10^0
0 1 2 3 4 5 6 7 8 9 10
일수
용량 L
용량 R
우측
그룹 A
그룹 B
음성대조군
발광 (광자 /sec/mm2)
10^6
10^5
10^4
10^3
10^2
10^1
10^0
0 1 2 3 4 5 6 7 8 9 10
일수
용량 L
용량 R

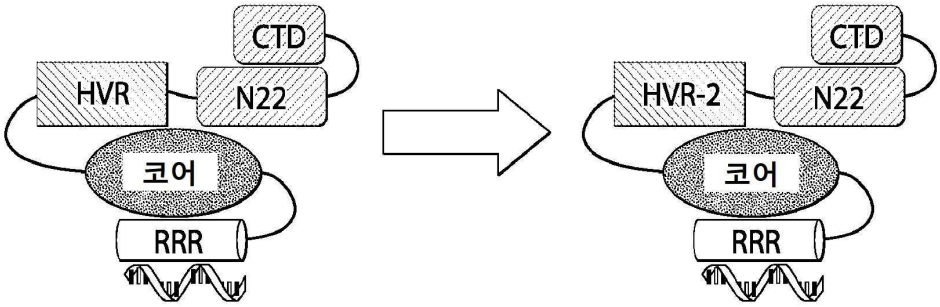
도면18a


pVL46-063
Maxi
U D
Dig
U D
Lig
U D
Final
U D
pVL46-240
Maxi
U D
Dig
U D
Lig
U D
Final
U D
3kb
2kb
1.5kb
1kb
0.5kb
3kb
2kb
1.5kb
1kb
0.5kb

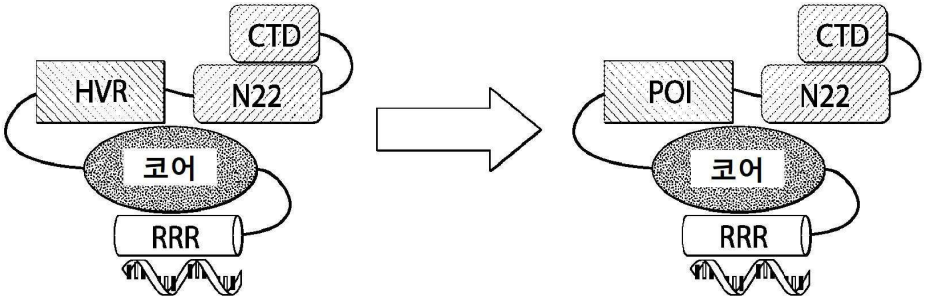
도면18b


선형 LY2 Vg/mL
선형 LY2 nLuc 카피 / mL
환형 LY2 Vg/mL
환형 LY2 nLuc 카피 / mL
카피 / mL
1.40E+06
1.20E+06
1.00E+06
8.00E+05
6.00E+05
4.00E+05
2.00E+05
0.00E+00
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28

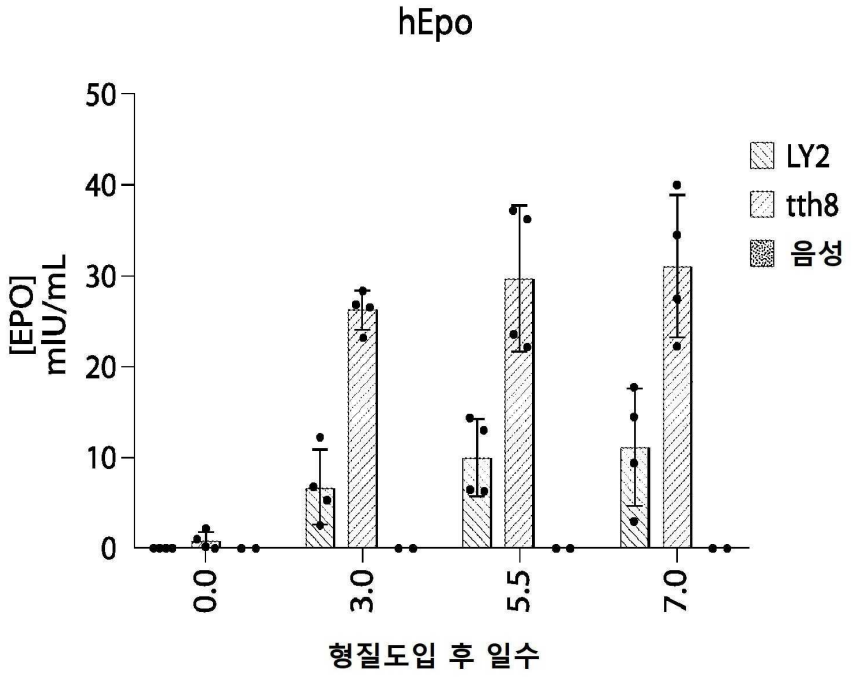
도면18c


CTD
HVR
N22
코어
RRR
CTD
HVR-2
N22
코어
RRR

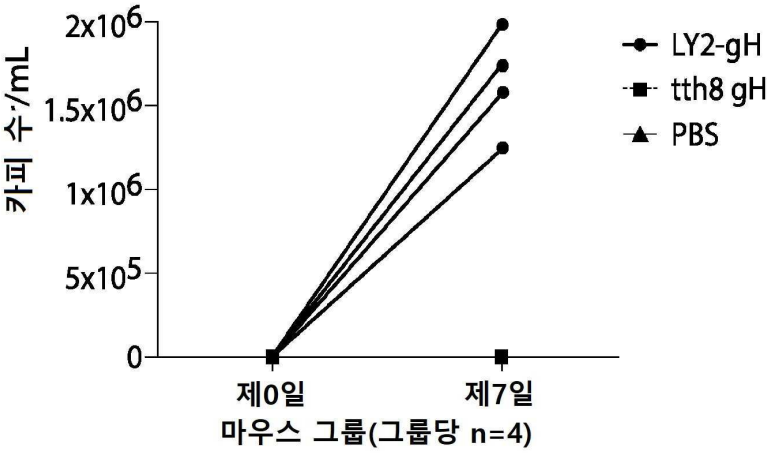
도면18d



도면19

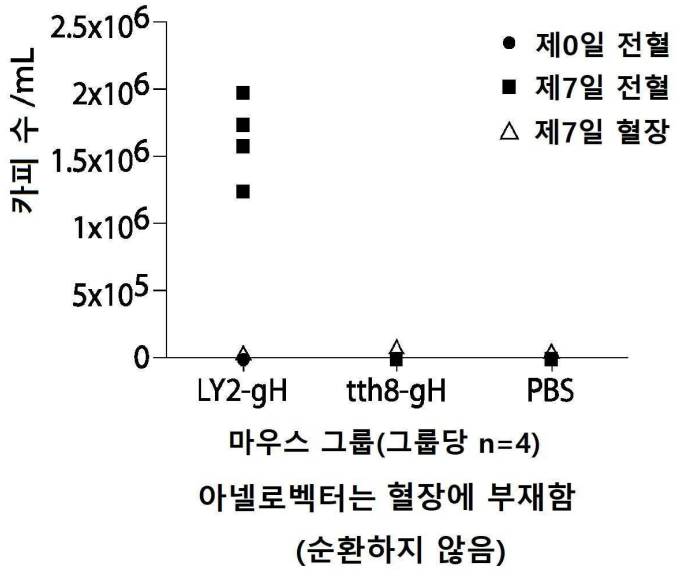


도면20a

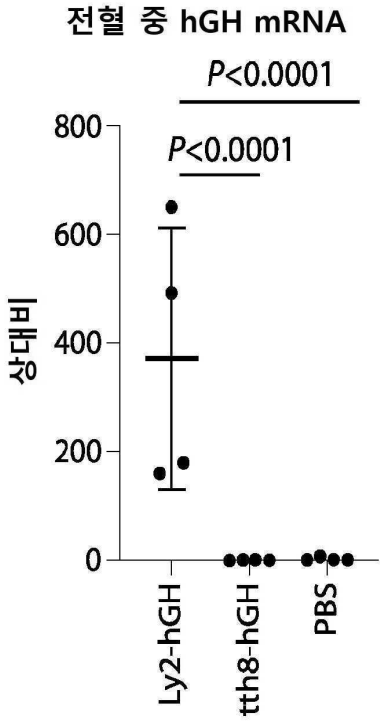


전혈 세포 분획에서 검출된 아넬로벡터

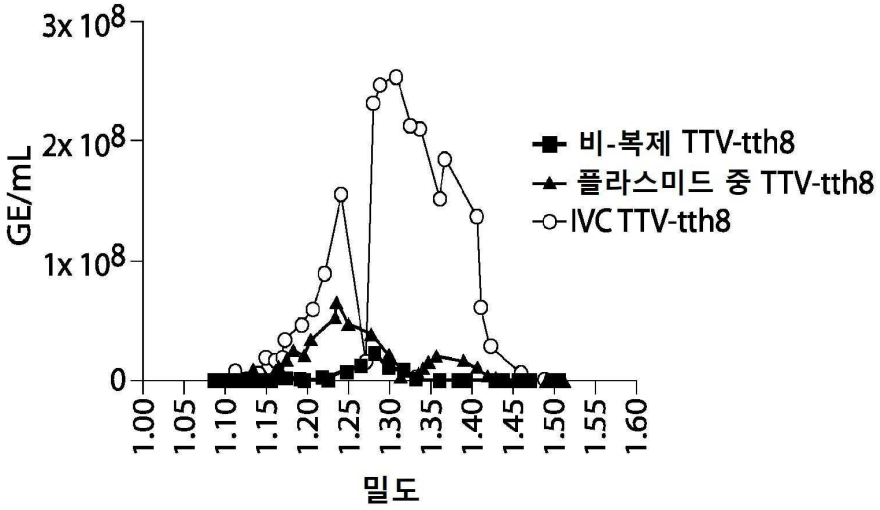
도면20b


2.5x10^6
2x10^6
1.5x10^6
1x10^6
5x10^5
0
카피 수 /mL
● 제0일 전혈
■ 제7일 전혈
△ 제7일 혈장
LY2-gH
tth8-gH
PBS
마우스 그룹(그룹당 n=4)
아넬로벡터는 혈장에 부재함
(순환하지 않음)

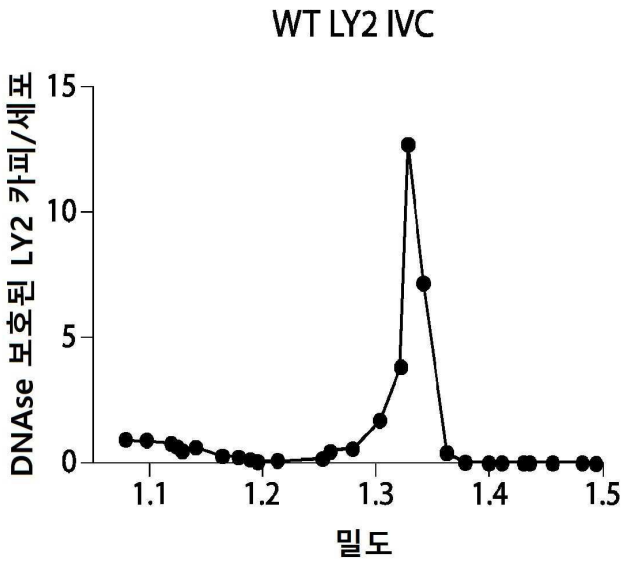
도면21

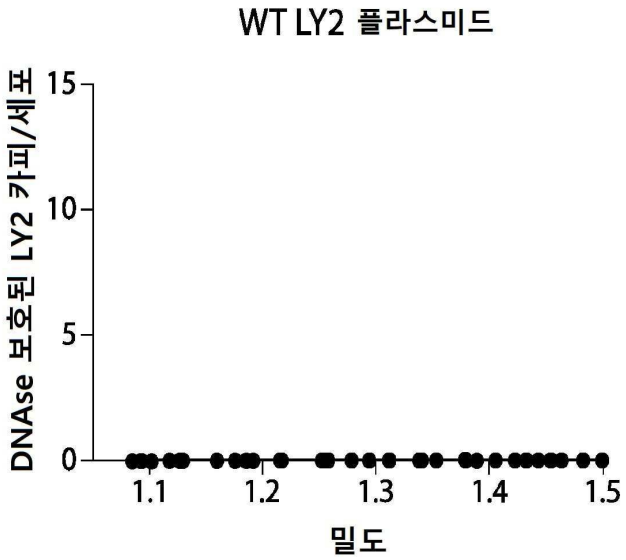

전혈 중 hGH mRNA
P<0.0001
P<0.0001
800
600
400
200
0
상대비
Ly2-hGH
tth8-hGH
PBS

도면22

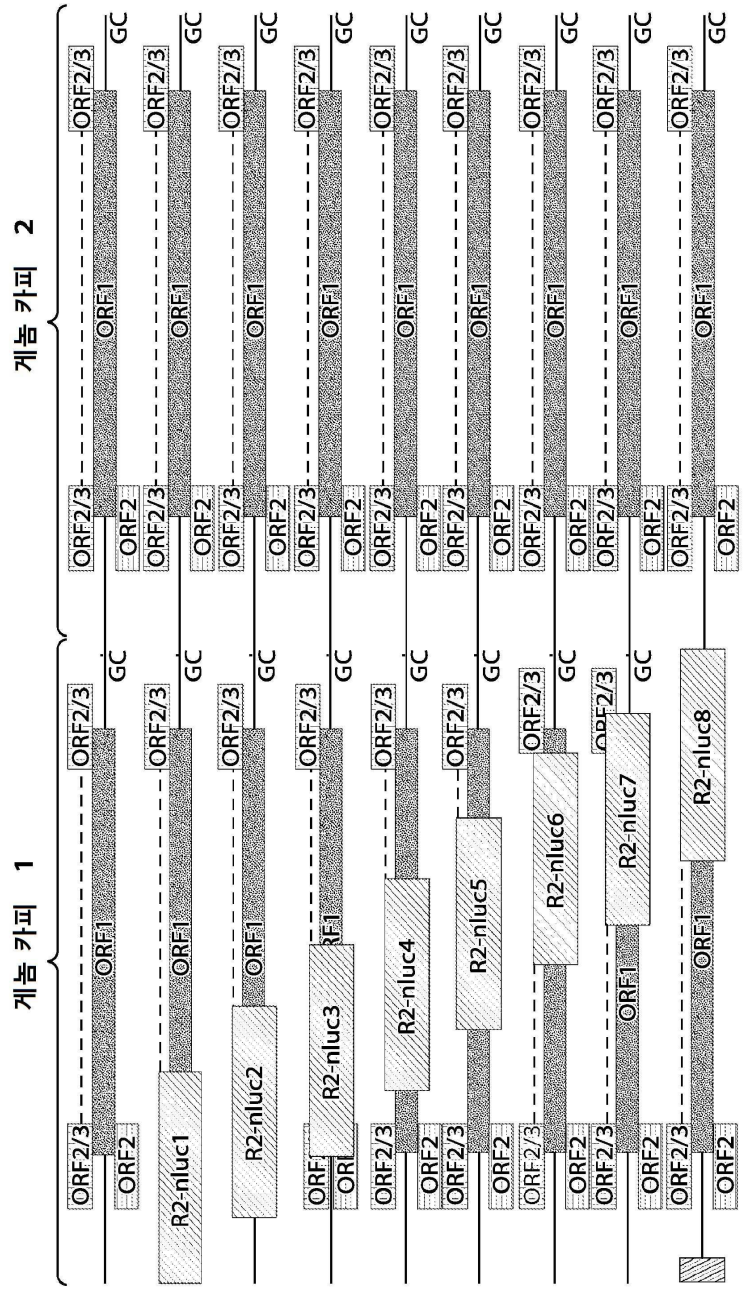

3x 10^8
2x 10^8
1x 10^8
0
GE/mL
1.00
1.05
1.10
1.15
1.20
1.25
1.30
1.35
1.40
1.45
1.50
1.55
1.60
밀도
비-복제 TTV-tth8
플라스미드 중 TTV-tth8
IVC TTV-tth8

도면23

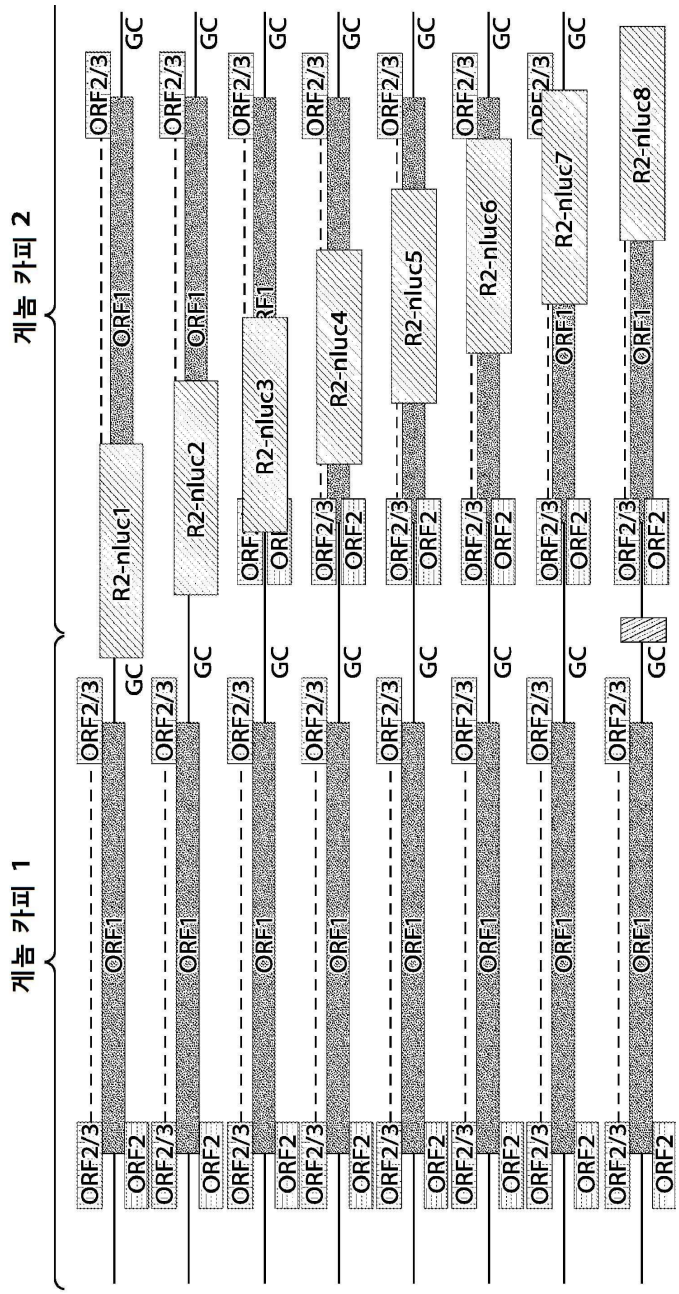

WT LY2 IVC
DNAse 보호된 LY2 카피/세포
15
10
5
0
1.1
1.2
1.3
1.4
1.5
밀도


WT LY2 플라스미드
DNAse 보호된 LY2 카피/세포
15
10
5
0
1.1
1.2
1.3
1.4
1.5
밀도

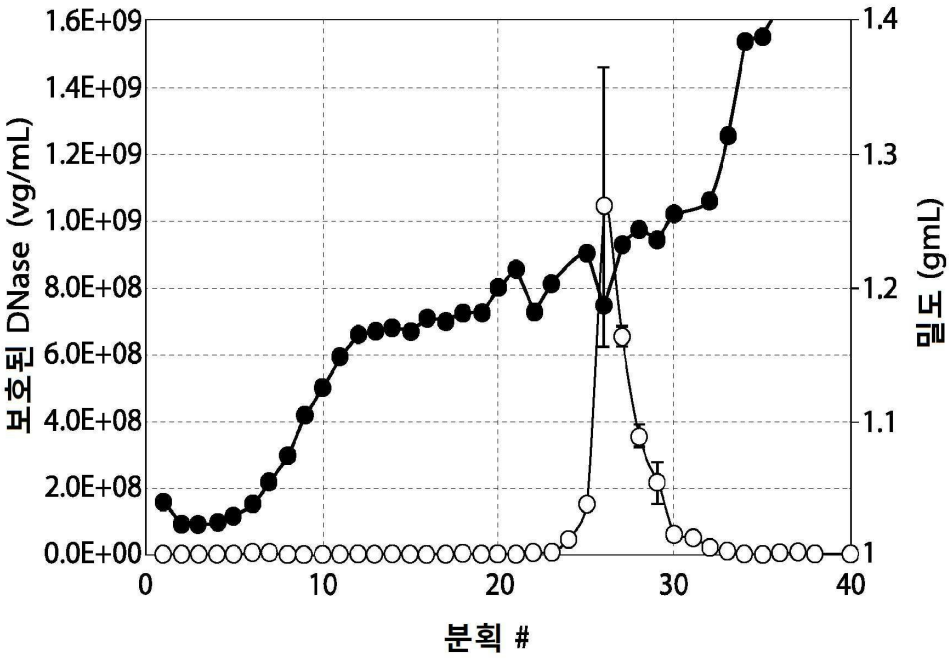
도면24a


게놈 카피 1
게놈 카피 2
ORF2/3
ORF1
ORF2
GC
R2-nluc1
R2-nluc2
R2-nluc3
R2-nluc4
R2-nluc5
R2-nluc6
R2-nluc7
R2-nluc8

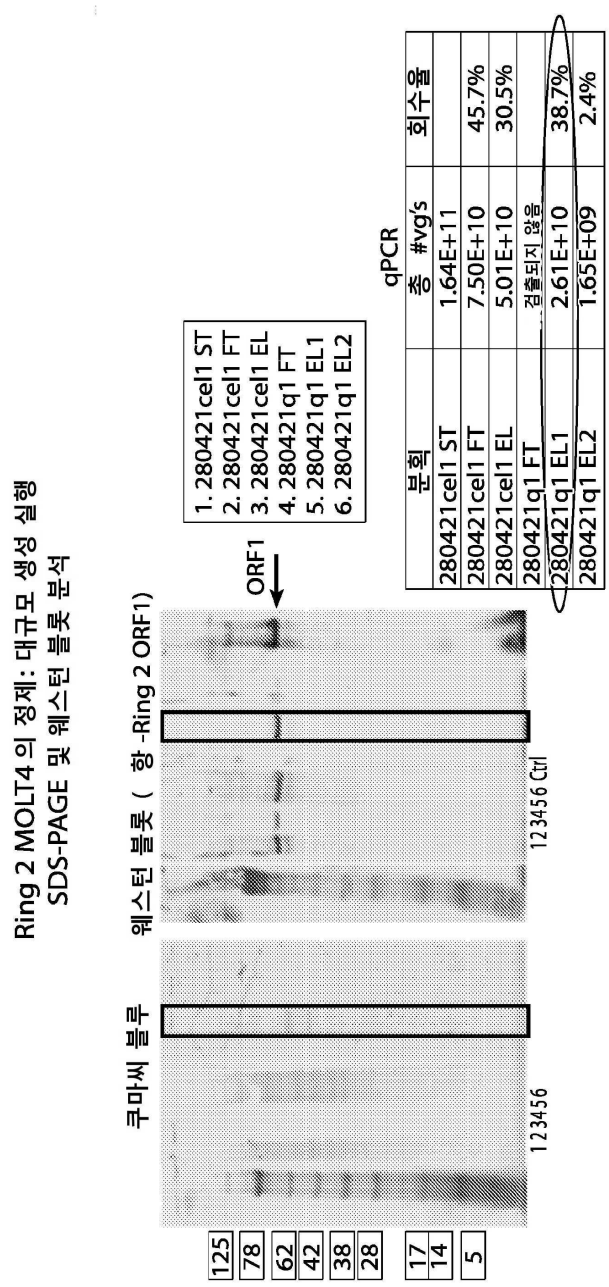
도면24b


게놈 카피 1
게놈 카피 2
ORF2/3
ORF1
ORF2
GC
R2-nluc1
R2-nluc2
R2-nluc3
R2-nluc4
R2-nluc5
R2-nluc6
R2-nluc7
R2-nluc8

도면25


1.6E+09
1.4E+09
1.2E+09
1.0E+09
8.0E+08
6.0E+08
4.0E+08
2.0E+08
0.0E+00
보호된 DNase (vg/mL)
1.4
1.3
1.2
1.1
1
밀도 (gmL)
0
10
20
30
40
분획 #

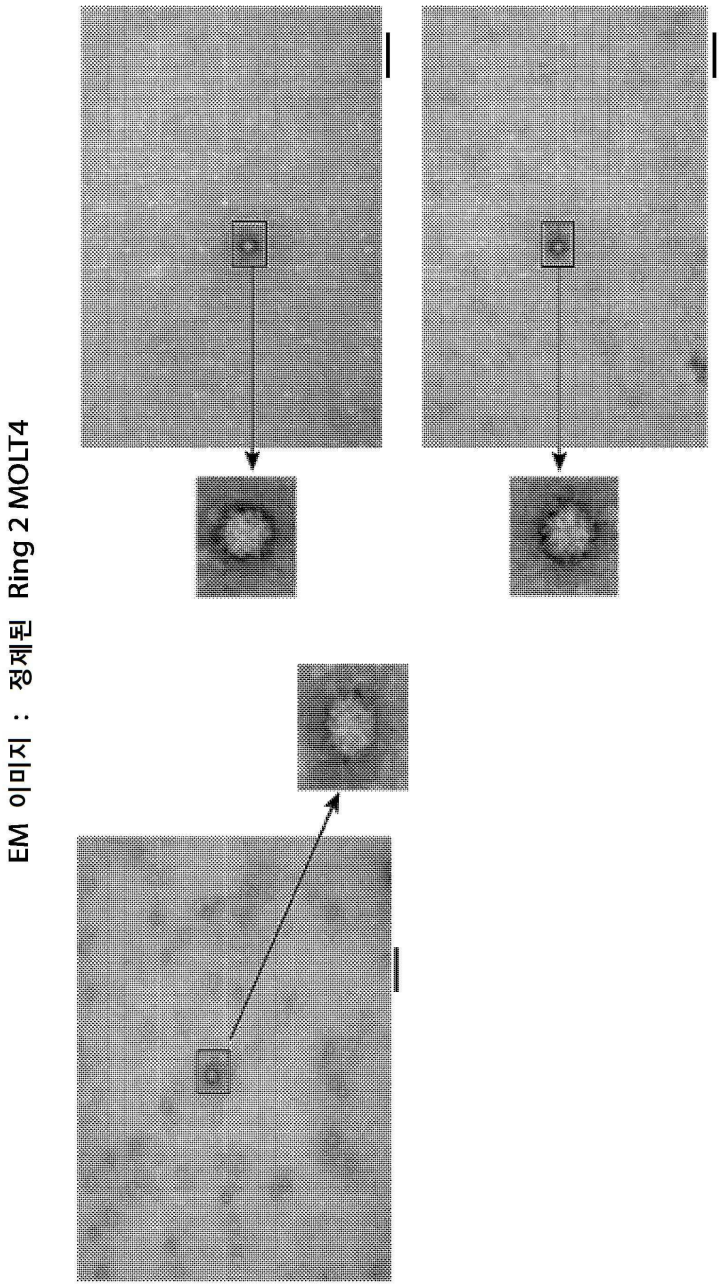
도면26a



qPCR

분획	총 #vg's	회수율
280421cel1 ST	1.64E+11	
280421cel1 FT	7.50E+10	45.7%
280421cel1 EL	5.01E+10	30.5%
280421q1 FT	검출되지 않음	
280421q1 EL1	2.61E+10	38.7%
280421q1 EL2	1.65E+09	2.4%

도면26b



서 열 목 록

SEQUENCE LISTING

<110> FLAGSHIP PIONEERING INNOVATIONS V, INC.

<120> TANDEM ANELLOVIRUS CONSTRUCTS

<130> V2057-7011WO

<140> PCT/US2021/037091

<141> 2021-06-11

<150> 63/146,963

<151> 2021-02-08

<150> 63/038,483

<151> 2020-06-12

<160> 953

<170> PatentIn version 3.5

<210> 1

<400> 1

000

<210> 2

<400> 2

000

<210> 3

<400> 3

000

<210> 4

<400> 4

000

<210> 5

<400> 5

000

<210> 6

<400> 6

000

<210> 7

<400> 7

000

<210> 8

<400> 8

000

<210> 9

<400> 9

000

<210> 10

<400> 10

000

<210> 11

<400> 11

000

<210> 12

<400> 12

000

<210> 13

<400> 13

000

<210> 14

<400> 14

000

<210> 15

<400> 15

000

<210> 16

<211> 3753

<212> DNA

<213> Alphatorquevirus sp.

<400> 16

tgctacgtca ctaacccacg tgtcctctac aggccaatcg cagtctatgt cgtgcacttc 60

ctgggcatgg tctacataat tatataaatg cttgcacttc cgaatggctg agtttttgct 120

gcccgtccgc ggagaggagc cacggcaggg gatccgaacg tcctgagggc gggtgccgga 180

ggtgagttta cacaccgaag tcaaggggca attcgggctc aggactggcc gggctttggg 240

caaggctctt aaaaatgcac ttttctcgaa taagcagaaa gaaaaggaaa gtgctactgc 300

tttgcgtgcc agcagctaag aaaaaaccaa ctgctatgag cttctggaaa cctccggtac 360

acaatgtcac ggggatccaa cgcatgtggt atgagtcctt tcaccgtggc cacgcttctt 420

tttgtggttg tgggaatcct atacttcaca ttactgcact tgctgaaaca tatggccatc 480

caacaggccc gagaccttct gggccaccgg gagtagaccc caacccccac atccgtagag 540

ccaggcctgc cccggccgct ccggagccct cacaggttga ttcgagacca gccctgacat 600

ggcatgggga tggtggaagc gacggaggcg ctggtggttc cggaagcggt ggacccgtgg 660

cagacttcgc agacgatggc ctcgatcagc tcgtcgccgc cctagacgac gaagagtaag 720

gaggcgcaga cggtggagga gggggagacg aaaaacaagg acttacagac gcaggagacg 780

ctttagacgc aggggacgaa aagcaaaact tataataaaa ctgtggcaac ctgcagtaat 840

taaaagatgc agaataaagg gatacatacc actgattata agtgggaacg gtacctttgc 900

cacaaacttt accagtcaca taaatgacag aataatgaaa ggccccttcg ggggaggaca 960

cagcactatg aggttcagcc tctacatttt gtttgaggag cacctcagac acatgaactt 1020

ctggaccaga agcaacgata acctagagct aaccagatac ttgggggctt cagtaaaaat 1080

atacaggcac ccagaccaag actttatagt aatatacaac agaagaaccc ctctaggagg 1140

caacatctac acagcaccct ctctacaccc aggcaatgcc attttagcaa aacacaaaat 1200

attagtacca agtttacaga caagaccaaa gggtagaaaa gcaattagac taagaatagc 1260

acccccccaca ctctttacag acaagtggta ctttcaaaag gacatagccg acctcaccct 1320

tttcaacatc atggcagttg aggctgactt gcggtttccg ttctgctcac cacaaactga 1380

caacacttgc atcagcttcc aggtccttag ttccgtttac aacaactacc tcagtattaa 1440

taccttttaat aatgacaact cagactcaaa gttaaaagaa tttttaaata aagcatttcc 1500

aacaacaggc acaaaaggaa caagtttaaa tgcactaaat acatttagaa cagaaggatg 1560

cataagtcac ccacaactaa aaaaaccaaa cccacaaata aacaaaccat tagagtcaca 1620

atactttgca cctttagatg ccctctgggg agaccccata tactataatg atctaaatga 1680

aaacaaaagt ttgaacgata tcattgagaa aatactaata aaaaacatga ttacatacca 1740

tgcaaaacta agagaatttc caaattcata ccaaggaaac aaggcctttt gccacctaac 1800

aggcatatac agcccaccat acctaaacca aggcagaata tctccagaaa tatttggact 1860

gtacacagaa ataatttaca acccttacac agacaaagga actggaaaca aagtatggat 1920

ggacccacta actaaagaga acaacatata taaagaagga cagagcaaat gcctactgac 1980

tgacatgccc ctatggactt tacttttgg atatacagac tggtgtaaaa aggacactaa 2040

taactgggac ttaccactaa actacagact agtactaata tgcccttata cctttccaaa 2100

attgtacaat gaaaaagtaa aagactatgg gtacatcccg tactcctaca aattcggagc 2160

gggtcagatg ccagacggca gcaactacat acccttttcag tttagagcaa agtggtaccc 2220

cacagtacta caccagcaac aggtaatgga ggacataagc aggagcgggc cctttgcacc 2280

taaggtagaa aaaccaagca ctcagctggt aatgaagtac tgttttaact ttaactgggg 2340

cggtaaccct atcattgaac agattgttaa agaccccagc ttccagccca cctatgaaat 2400

acccggtacc ggtaacatcc ctagaagaat acaagtcatc gacccgcggg tcctgggacc 2460

gcactactcg ttccggtcat gggacatgcg cagacacaca tttagcagag caagtattaa 2520

gagagtgtca gaacaacaag aaacttctga ccttgtattc tcaggcccaa aaaagcctcg 2580

ggtcgacatc ccaaaacaag aaacccaaga agaaagctca cattcactcc aaagagaatc 2640

gagaccgtgg gagaccgagg aagaaagcga gacagaagcc ctctcgcaag agagccaaga 2700

ggtccccttc caacagcagt tgcagcagca gtaccaagag cagctcaagc tcagacaggg 2760

aatcaaagtc ctcttcgagc agctcataag gacccaacaa ggggtccatg taaacccatg 2820

cctacggtag gtcccaggca gtggctgttt ccagagagaa agccagcccc agctcctagc 2880

agtggagact gggccatgga gtttctcgca gcaaaaatat ttgataggcc agttagaagc 2940

aaccttaaag ataccccta ctacccatat gttaaaaacc aatacaatgt ctactttgac 3000

cttaaatttg aataaacagc agcttcaaac ttgcaaggcc gtgggagttt cactggtcgg 3060

tgtctacctc taaaggtcac taagcactcc gagcgtaagc gaggagtgcg accctccccc 3120

ctggaacaac ttcttcggag tccggcgcta cgccttcggc tgcgccggac acctcagacc 3180

ccccctccac ccgaaacgct tgcgcgtttc ggaccttcgg cgtcgggggg gtcgggagct 3240

ttattaaacg gactccgaag tgctcttgga cactgagggg gtgaacagca acgaaagtga 3300

gtggggccag acttcgccat aaggccttta tcttcttgcc atttgtcagt gtccggggtc 3360

gccataggct tcgggctcgt ttttaggcct tccggactac aaaaatcgcc attttggtga 3420

cgtcacggcc gccatcttaa gtagttgagg cggacggtgg cgtgagttca aaggtcacca 3480

tcagccacac ctactcaaaa tggtggacaa tttcttccgg gtcaaaggtt acagccgcca 3540

tgttaaaaca cgtgacgtat gacgtcacgg ccgccatttt gtgacacaag atggccgact 3600

tccttcctct ttttcaaaaa aaagcggaag tgccgccgcg gcggcggggg gcggcgcgct 3660

gcgcgcgccg cccagtaggg ggagccatgc gccccccccc gcgcatgcgc ggggcccccc 3720

cccgcggggg gctccgcccc ccggcccccc ccg 3753

<210> 17

<211> 127

<212> PRT

<213> Alphatorquevirus sp.

<400> 17

Met Ser Phe Trp Lys Pro Pro Val His Asn Val Thr Gly Ile Gln Arg
1 5 10 15

Met Trp Tyr Glu Ser Phe His Arg Gly His Ala Ser Phe Cys Gly Cys
20 25 30

Gly Asn Pro Ile Leu His Ile Thr Ala Leu Ala Glu Thr Tyr Gly His
35 40 45

Pro Thr Gly Pro Arg Pro Ser Gly Pro Pro Gly Val Asp Pro Asn Pro

50 55 60

His Ile Arg Arg Ala Arg Pro Ala Pro Ala Ala Pro Glu Pro Ser Gln

65 70 75 80

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

85 90 95

Gly Gly Ala Gly Gly Ser Gly Ser Gly Gly Pro Val Ala Asp Phe Ala

100 105 110

Asp Asp Gly Leu Asp Gln Leu Val Ala Ala Leu Asp Asp Glu Glu

115 120 125

<210> 18

<211> 268

<212> PRT

<213> Alphatorquevirus sp.

<400> 18

Met Ser Phe Trp Lys Pro Pro Val His Asn Val Thr Gly Ile Gln Arg

1 5 10 15

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

20 25 30

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

35 40 45

Pro Thr Gly Pro Arg Pro Ser Gly Pro Pro Gly Val Asp Pro Asn Pro

50 55 60

His Ile Arg Arg Ala Arg Pro Ala Pro Ala Ala Pro Glu Pro Ser Gln

65 70 75 80

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

85 90 95

Gly Gly Ala Gly Gly Ser Gly Ser Gly Gly Pro Val Ala Asp Phe Ala

100 105 110

Asp Asp Gly Leu Asp Gln Leu Val Ala Ala Leu Asp Asp Glu Glu Leu

115 120 125

Leu Lys Thr Pro Ala Ser Ser Pro Pro Met Lys Tyr Pro Val Pro Val

130 135 140

Thr Ser Leu Glu Glu Tyr Lys Ser Ser Thr Arg Gly Ser Trp Asp Arg

145 150 155 160

Thr Thr Arg Ser Gly His Gly Thr Cys Ala Asp Thr His Leu Ala Glu

165 170 175

Gln Val Leu Arg Glu Cys Gln Asn Asn Lys Lys Leu Leu Thr Leu Tyr

180 185 190

Ser Gln Ala Gln Lys Ser Leu Gly Ser Thr Ser Gln Asn Lys Lys Pro

195 200 205

Lys Lys Lys Ala His Ile His Ser Lys Glu Asn Arg Asp Arg Gly Arg

210 215 220

Pro Arg Lys Lys Ala Arg Gln Lys Pro Ser Arg Lys Arg Ala Lys Arg

225 230 235 240

Ser Pro Ser Asn Ser Ser Cys Ser Ser Ser Thr Lys Ser Ser Ser Ser

245 250 255

Ser Asp Arg Glu Ser Lys Ser Ser Ser Ser Ser Ser

260 265

<210> 19

<211> 276

<212> PRT

<213> Alphatorquevirus sp.

<400> 19

Met Ser Phe Trp Lys Pro Pro Val His Asn Val Thr Gly Ile Gln Arg

1 5 10 15

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

20 25 30

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

35 40 45

Pro Thr Gly Pro Arg Pro Ser Gly Pro Pro Gly Val Asp Pro Asn Pro

50 55 60

His Ile Arg Arg Ala Arg Pro Ala Pro Ala Ala Pro Glu Pro Ser Gln

65 70 75 80

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

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

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

260 265 270
Leu Lys Phe Glu
275

<210> 20
<211> 167
<212> PRT
<213> Alphatorquevirus sp.
<400> 20
Met Ser Phe Trp Lys Pro Pro Val His Asn Val Thr Gly Ile Gln Arg

1 5 10 15

Met Trp Pro Lys Lys Ala Ser Gly Arg His Pro Lys Thr Arg Asn Pro

20 25 30

Arg Arg Lys Leu Thr Phe Thr Pro Lys Arg Ile Glu Thr Val Gly Asp

35 40 45

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

50 55 60

Pro Leu Pro Thr Ala Val Ala Ala Ala Val Pro Arg Ala Ala Gln Ala

65 70 75 80

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

85 90 95

Arg Gly Pro Cys Lys Pro Met Pro Thr Val Gly Pro Arg Gln Trp Leu

100 105 110

Phe Pro Glu Arg Lys Pro Ala Pro Ala Pro Ser Ser Gly Asp Trp Ala

115 120 125

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

130 135 140

Leu Lys Asp Thr Pro Tyr Tyr Pro Tyr Val Lys Asn Gln Tyr Asn Val

145 150 155 160

Tyr Phe Asp Leu Lys Phe Glu

165

<210> 21

<211> 743

<212> PRT

<213> Alphatorquevirus sp.

<400> 21

Met Ala Trp Gly Trp Trp Lys Arg Arg Arg Arg Trp Trp Phe Arg Lys

1 5 10 15

Arg Trp Thr Arg Gly Arg Leu Arg Arg Arg Trp Pro Arg Ser Ala Arg

20 25 30

Arg Arg Pro Arg Arg Arg Arg Val Arg Arg Arg Arg Arg Trp Arg Arg

35 40 45

Gly Arg Arg Lys Thr Arg Thr Tyr Arg Arg Arg Arg Arg Phe Arg Arg
50 55 60

Arg Gly Arg Lys Ala Lys Leu Ile Ile Lys Leu Trp Gln Pro Ala Val
65 70 75 80
Ile Lys Arg Cys Arg Ile Lys Gly Tyr Ile Pro Leu Ile Ile Ser Gly
85 90 95
Asn Gly Thr Phe Ala Thr Asn Phe Thr Ser His Ile Asn Asp Arg Ile
100 105 110
Met Lys Gly Pro Phe Gly Gly Gly His Ser Thr Met Arg Phe Ser Leu
115 120 125

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

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

Ser Pro Gln Thr Asp Asn Thr Cys Ile Ser Phe Gln Val Leu Ser Ser
260 265 270
Val Tyr Asn Asn Tyr Leu Ser Ile Asn Thr Phe Asn Asn Asp Asn Ser
275 280 285
Asp Ser Lys Leu Lys Glu Phe Leu Asn Lys Ala Phe Pro Thr Thr Gly

290 295 300

Thr Lys Gly Thr Ser Leu Asn Ala Leu Asn Thr Phe Arg Thr Glu Gly
305 310 315 320

Cys Ile Ser His Pro Gln Leu Lys Lys Pro Asn Pro Gln Ile Asn Lys
325 330 335

Pro Leu Glu Ser Gln Tyr Phe Ala Pro Leu Asp Ala Leu Trp Gly Asp
340 345 350

Pro Ile Tyr Tyr Asn Asp Leu Asn Glu Asn Lys Ser Leu Asn Asp Ile
355 360 365

Ile Glu Lys Ile Leu Ile Lys Asn Met Ile Thr Tyr His Ala Lys Leu
370 375 380

Arg Glu Phe Pro Asn Ser Tyr Gln Gly Asn Lys Ala Phe Cys His Leu
385 390 395 400

Thr Gly Ile Tyr Ser Pro Pro Tyr Leu Asn Gln Gly Arg Ile Ser Pro
405 410 415

Glu Ile Phe Gly Leu Tyr Thr Glu Ile Ile Tyr Asn Pro Tyr Thr Asp
420 425 430

Lys Gly Thr Gly Asn Lys Val Trp Met Asp Pro Leu Thr Lys Glu Asn
435 440 445

Asn Ile Tyr Lys Glu Gly Gln Ser Lys Cys Leu Leu Thr Asp Met Pro
450 455 460

Leu Trp Thr Leu Leu Phe Gly Tyr Thr Asp Trp Cys Lys Lys Asp Thr
465 470 475 480

Asn Asn Trp Asp Leu Pro Leu Asn Tyr Arg Leu Val Leu Ile Cys Pro
485 490 495

Tyr Thr Phe Pro Lys Leu Tyr Asn Glu Lys Val Lys Asp Tyr Gly Tyr
500 505 510

Ile Pro Tyr Ser Tyr Lys Phe Gly Ala Gly Gln Met Pro Asp Gly Ser
515 520 525

Asn Tyr Ile Pro Phe Gln Phe Arg Ala Lys Trp Tyr Pro Thr Val Leu
530 535 540

His Gln Gln Gln Val Met Glu Asp Ile Ser Arg Ser Gly Pro Phe Ala
545 550 555 560
Pro Lys Val Glu Lys Pro Ser Thr Gln Leu Val Met Lys Tyr Cys Phe
565 570 575

Asn Phe Asn Trp Gly Gly Asn Pro Ile Ile Glu Gln Ile Val Lys Asp
580 585 590
Pro Ser Phe Gln Pro Thr Tyr Glu Ile Pro Gly Thr Gly Asn Ile Pro
595 600 605
Arg Arg Ile Gln Val Ile Asp Pro Arg Val Leu Gly Pro His Tyr Ser
610 615 620
Phe Arg Ser Trp Asp Met Arg Arg His Thr Phe Ser Arg Ala Ser Ile
625 630 635 640

Lys Arg Val Ser Glu Gln Gln Glu Thr Ser Asp Leu Val Phe Ser Gly
645 650 655
Pro Lys Lys Pro Arg Val Asp Ile Pro Lys Gln Glu Thr Gln Glu Glu
660 665 670
Ser Ser His Ser Leu Gln Arg Glu Ser Arg Pro Trp Glu Thr Glu Glu
675 680 685
Glu Ser Glu Thr Glu Ala Leu Ser Gln Glu Ser Gln Glu Val Pro Phe
690 695 700

Gln Gln Gln Leu Gln Gln Gln Tyr Gln Glu Gln Leu Lys Leu Arg Gln
705 710 715 720
Gly Ile Lys Val Leu Phe Glu Gln Leu Ile Arg Thr Gln Gln Gly Val
725 730 735
His Val Asn Pro Cys Leu Arg
740

<210> 22
<211> 194
<212> PRT
<213> Alphatorquevirus sp.
<400> 22
Met Ala Trp Gly Trp Trp Lys Arg Arg Arg Arg Trp Trp Phe Arg Lys

1 5 10 15

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

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

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

<210> 23
<211> 113
<212> PRT
<213> Alphatorquevirus sp.
<400> 23
Met Ala Trp Gly Trp Trp Lys Arg Arg Arg Arg Trp Trp Phe Arg Lys

1 5 10 15

Arg Trp Thr Arg Gly Arg Leu Arg Arg Arg Trp Pro Arg Ser Ala Arg

20 25 30

Arg Arg Pro Arg Arg Arg Arg Ala Gln Lys Ser Leu Gly Ser Thr Ser

35 40 45

Gln Asn Lys Lys Pro Lys Lys Lys Ala His Ile His Ser Lys Glu Asn

50 55 60

Arg Asp Arg Gly Arg Pro Arg Lys Lys Ala Arg Gln Lys Pro Ser Arg

65 70 75 80

Lys Arg Ala Lys Arg Ser Pro Ser Asn Ser Ser Cys Ser Ser Ser Thr

85 90 95

Lys Ser Ser Ser Ser Ser Asp Arg Glu Ser Lys Ser Ser Ser Ser Ser

100 105 110

Ser

<210> 24

<400> 24

000

<210> 25

<400> 25

000

<210> 26

<400> 26

000

<210> 27

<400> 27

000

<210> 28

<400> 28

000

<210> 29

<400> 29

000

<210> 30

<400> 30

000

<210> 31

<400> 31

000

<210> 32

<400> 32

000

<210> 33

<400> 33

000

<210> 34

<400> 34

000

<210> 35

<400> 35

000

<210> 36

<400> 36

000

<210> 37

<400> 37

000

<210> 38

<400> 38

000

<210> 39

<400> 39

000

<210> 40

<400> 40

000

<210> 41

<400> 41

000

<210> 42

<400> 42

000

<210> 43

<400> 43

000

<210> 44

<400> 44

000

<210> 45

<400> 45

000

<210> 46

<400> 46

000

<210> 47

<400> 47

000

<210> 48

<400> 48

000

<210> 49

<400> 49

000

<210> 50

<400> 50

000

<210> 51

<400> 51

000

<210> 52

<400> 52

000

<210> 53

<400> 53

000

<210> 54

<211> 2979

<212> DNA

<213> Betatorquevirus sp.

<400> 54

taataaatat tcaacaggaa aaccacctaa tttaaattgc cgaccacaaa ccgtcactta 60

gttccccttt ttgcaacaac ttctgctttt ttccaactgc cggaaaacca cataatttgc 120

atggctaacc acaaactgat atgctaatta acttccacaa aacaacttcc ccttttaaaa 180

ccacacctac aaattaatta ttaaacacag tcacatcctg ggaggtacta ccacactata 240

ataccaagtg cacttccgaa tggctgagtt tatgccgcta gacggagaac gcatcagtta 300

ctgactgcgg actgaacttg ggcgggtgcc gaaggtgagt gaaaccaccg aagtcaaggg 360

gcaattcggg ctagttcagt ctagcggaac gggcaagaaa cttaaaatta ttttatttt 420

cagatgagcg actgctttaa accaacatgc tacaacaaca aaacaaagca aactcactgg 480

attaataacc tgcatttaac ccacgacctg atctgcttct gcccaacacc aactagacac 540

ttattactag ctttagcaga acaacaagaa acaattgaag tgtctaaaca agaaaaagaa 600

aaaataacaa gatgccttat tactacagaa gaagacggta caactacaga cgtcctagat 660

ggtatggacg aggttggatt agacgccctt ttcgcagaag atttcgaaga aaaagaaggg 720

taagacctac ttatactact attcctctaa agcaatggca accgccatat aaaagaacat 780

gctatataaa aggacaagac tgtttaatat actatagcaa cttaagactg ggaatgaata 840

gtacaatgta tgaaaaaagt attgtacctg tacattggcc gggaggggggt tctttttctg 900

taagcatgtt aactttagat gccttgtatg atatacataa actttgtaga aactggtgga 960

catccacaaa ccaagactta ccactagtaa gatataaagg atgcaaaata acattttatc 1020

aaagcacatt tacagactac atagtaagaa tacatacaga actaccagct aacagtaaca 1080

aactaacata cccaaacaca catccactaa tgatgatgat gtctaagtac aaacacatta 1140

tacctagtag acaaacaaga agaaaaaaga aaccatacac aaaaatattt gtaaaaccac 1200

ctccgcaatt tgaaaacaaa tggtactttg ctacagacct ctacaaaatt ccattactac 1260

aaatacactg cacagcatgc aacttacaaa acccatttgt aaaaccagac aaattatcaa 1320

acaatgttac attatggtca ctaaacacca taagcataca aaatagaaac atgtcagtgg 1380

atcaaggaca atcatggcca tttaaaatac taggaacaca aagcttttat ttttactttt 1440

acaccggagc aaacctacca ggtgacacaa cacaaatacc agtagcagac ctattaccac 1500

taacaaaccc aagaataaac agaccaggac aatcactaaa tgaggcaaaa attacagacc 1560
atattacttt cacagaatac aaaaacaaat ttacaaatta ttggggtaac ccatttaata 1620
aacacattca agaacaccta gatatgatac tatactcact aaaaagtcca gaagcaataa 1680
aaaacgaatg gacaacagaa aacatgaaat ggaaccaatt aaacaatgca ggaacaatgg 1740
cattaacacc atttaacgag ccaatattca cacaaataca atataaccca gatagagaca 1800
caggagaaga cactcaatta tacctactct ctaacgctac aggaacagga tgggacccac 1860
caggaattcc agaattaata ctagaaggat ttccactatg gttaatatat tggggatttg 1920

cagactttca aaaaaaccta aaaaaagtaa caaacataga cacaaattac atgttagtag 1980
caaaaacaaa atttacacaa aaacctggca cattctactt agtaatacta aatgacacct 2040
ttgtagaagg caatagccca tatgaaaaac aacctttacc tgaagacaac attaaatggt 2100
acccacaagt acaataccaa ttagaagcac aaaacaaact actacaaact gggccattta 2160
caccaaacat acaaggacaa ctatcagaca atatatcaat gttttataaa ttttacttta 2220
aatggggagg aagcccacca aaagcaatta atgttgaaaa tcctgcccac cagattcaat 2280
atcccatacc ccgtaacgag catgaaacaa cttcgttaca gagtccaggg gaagccccag 2340

aatccatctt atactccttc gactatagac acgggaacta cacaacaaca gctttgtcac 2400
gaattagcca agactgggca cttaaagaca ctgtttctaa aattacagag ccagatcgac 2460
agcaactgct caaacaagcc ctcgaatgcc tgcaaatctc ggaagaaacg caggagaaaa 2520
aagaaaaaga agtacagcag ctcatcagca acctcagaca gcagcagcag ctgtacagag 2580
agcgaataat atcattatta aaggaccaat aacttttaac tgtgtaaaaa aggtgaaatt 2640
gtttgatgat aaaccaaaaa accgtagatt tacacctgag gaatttgaaa ctgagttaca 2700
aatagcaaaa tggttaaaga gacccccaag atcctttgta aatgatcctc ccttttaccc 2760

atggttacca cctgaacctg ttgtaaactt taagcttaat tttactgaat aaaggccagc 2820
attaattcac ttaaggagtc tgtttattta agttaaacct taataaacgg tcaccgcctc 2880
cctaatacgc aggcgcagaa agggggctcc gccccttta acccccaggg ggctccgccc 2940
cctgaaaccc ccaaggggggc tacgcccccc tacaccccc 2979

<210> 55

<211> 99

<212> PRT

<213> Betatorquevirus sp.

<400> 55

Met Ser Asp Cys Phe Lys Pro Thr Cys Tyr Asn Asn Lys Thr Lys Gln
1 5 10 15

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

Met Asp Glu Val Gly Leu Asp Ala Leu Phe Ala Glu Asp Phe Glu Glu
85 90 95
Lys Glu Gly

<210> 56
<211> 203
<212> PRT
<213> Betatorquevirus sp.
<400> 56
Met Ser Asp Cys Phe Lys Pro Thr Cys Tyr Asn Asn Lys Thr Lys Gln
1 5 10 15
Thr His Trp Ile Asn Asn Leu His Leu Thr His Asp Leu Ile Cys Phe
20 25 30
Cys Pro Thr Pro Thr Arg His Leu Leu Leu Ala Leu Ala Glu Gln Gln

35 40 45
Glu Thr Ile Glu Val Ser Lys Gln Glu Lys Glu Lys Ile Thr Arg Cys
50 55 60
Leu Ile Thr Thr Glu Glu Asp Gly Thr Thr Thr Asp Val Leu Asp Gly
65 70 75 80
Met Asp Glu Val Gly Leu Asp Ala Leu Phe Ala Glu Asp Phe Glu Glu
85 90 95

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

100 105 110

Arg Tyr Arg Val Gln Gly Lys Pro Gln Asn Pro Ser Tyr Thr Pro Ser

115 120 125

Thr Ile Asp Thr Gly Thr Thr Gln Gln Gln Leu Cys His Glu Leu Ala

130 135 140

Lys Thr Gly His Leu Lys Thr Leu Phe Leu Lys Leu Gln Ser Gln Ile

145 150 155 160

Asp Ser Asn Cys Ser Asn Lys Pro Ser Asn Ala Cys Lys Ser Arg Lys

165 170 175

Lys Arg Arg Arg Lys Lys Lys Lys Lys Tyr Ser Ser Ser Ser Ala Thr

180 185 190

Ser Asp Ser Ser Ser Ser Cys Thr Glu Ser Glu

195 200

<210> 57

<211> 219

<212> PRT

<213> Betatorquevirus sp.

<400> 57

Met Ser Asp Cys Phe Lys Pro Thr Cys Tyr Asn Asn Lys Thr Lys Gln

1 5 10 15

Thr His Trp Ile Asn Asn Leu His Leu Thr His Asp Leu Ile Cys Phe

20 25 30

Cys Pro Thr Pro Thr Arg His Leu Leu Leu Ala Leu Ala Glu Gln Gln

35 40 45

Glu Thr Ile Glu Val Ser Lys Gln Glu Lys Glu Lys Ile Thr Arg Cys

50 55 60

Leu Ile Thr Thr Glu Glu Asp Gly Thr Thr Thr Asp Val Leu Asp Gly

65 70 75 80

Met Asp Glu Val Gly Leu Asp Ala Leu Phe Ala Glu Asp Phe Glu Glu

85 90 95

Lys Glu Gly Ala Arg Ser Thr Ala Thr Ala Gln Thr Ser Pro Arg Met
100 105 110
Pro Ala Asn Leu Gly Arg Asn Ala Gly Glu Lys Arg Lys Arg Ser Thr
115 120 125
Ala Ala His Gln Gln Pro Gln Thr Ala Ala Ala Ala Val Gln Arg Ala
130 135 140
Asn Asn Ile Ile Ile Lys Gly Pro Ile Thr Phe Asn Cys Val Lys Lys

145 150 155 160
Val Lys Leu Phe Asp Asp Lys Pro Lys Asn Arg Arg Phe Thr Pro Glu
165 170 175
Glu Phe Glu Thr Glu Leu Gln Ile Ala Lys Trp Leu Lys Arg Pro Pro
180 185 190
Arg Ser Phe Val Asn Asp Pro Pro Phe Tyr Pro Trp Leu Pro Pro Glu
195 200 205
Pro Val Val Asn Phe Lys Leu Asn Phe Thr Glu
210 215

<210> 58
<211> 666
<212> PRT
<213> Betatorquevirus sp.
<400> 58
Met Pro Tyr Tyr Tyr Arg Arg Arg Arg Tyr Asn Tyr Arg Arg Pro Arg
1 5 10 15
Trp Tyr Gly Arg Gly Trp Ile Arg Arg Pro Phe Arg Arg Arg Phe Arg
20 25 30
Arg Lys Arg Arg Val Arg Pro Thr Tyr Thr Thr Ile Pro Leu Lys Gln
35 40 45
Trp Gln Pro Pro Tyr Lys Arg Thr Cys Tyr Ile Lys Gly Gln Asp Cys

50 55 60
Leu Ile Tyr Tyr Ser Asn Leu Arg Leu Gly Met Asn Ser Thr Met Tyr
65 70 75 80
Glu Lys Ser Ile Val Pro Val His Trp Pro Gly Gly Gly Ser Phe Ser

85 90 95

Val Ser Met Leu Thr Leu Asp Ala Leu Tyr Asp Ile His Lys Leu Cys

100 105 110

Arg Asn Trp Trp Thr Ser Thr Asn Gln Asp Leu Pro Leu Val Arg Tyr

115 120 125

Lys Gly Cys Lys Ile Thr Phe Tyr Gln Ser Thr Phe Thr Asp Tyr Ile

130 135 140

Val Arg Ile His Thr Glu Leu Pro Ala Asn Ser Asn Lys Leu Thr Tyr

145 150 155 160

Pro Asn Thr His Pro Leu Met Met Met Met Ser Lys Tyr Lys His Ile

165 170 175

Ile Pro Ser Arg Gln Thr Arg Arg Lys Lys Lys Pro Tyr Thr Lys Ile

180 185 190

Phe Val Lys Pro Pro Pro Gln Phe Glu Asn Lys Trp Tyr Phe Ala Thr

195 200 205

Asp Leu Tyr Lys Ile Pro Leu Leu Gln Ile His Cys Thr Ala Cys Asn

210 215 220

Leu Gln Asn Pro Phe Val Lys Pro Asp Lys Leu Ser Asn Asn Val Thr

225 230 235 240

Leu Trp Ser Leu Asn Thr Ile Ser Ile Gln Asn Arg Asn Met Ser Val

245 250 255

Asp Gln Gly Gln Ser Trp Pro Phe Lys Ile Leu Gly Thr Gln Ser Phe

260 265 270

Tyr Phe Tyr Phe Tyr Thr Gly Ala Asn Leu Pro Gly Asp Thr Thr Gln

275 280 285

Ile Pro Val Ala Asp Leu Leu Pro Leu Thr Asn Pro Arg Ile Asn Arg

290 295 300

Pro Gly Gln Ser Leu Asn Glu Ala Lys Ile Thr Asp His Ile Thr Phe

305 310 315 320

Thr Glu Tyr Lys Asn Lys Phe Thr Asn Tyr Trp Gly Asn Pro Phe Asn

325 330 335

Lys His Ile Gln Glu His Leu Asp Met Ile Leu Tyr Ser Leu Lys Ser
340 345 350
Pro Glu Ala Ile Lys Asn Glu Trp Thr Thr Glu Asn Met Lys Trp Asn
355 360 365
Gln Leu Asn Asn Ala Gly Thr Met Ala Leu Thr Pro Phe Asn Glu Pro

370 375 380
Ile Phe Thr Gln Ile Gln Tyr Asn Pro Asp Arg Asp Thr Gly Glu Asp
385 390 395 400
Thr Gln Leu Tyr Leu Leu Ser Asn Ala Thr Gly Thr Gly Trp Asp Pro
405 410 415
Pro Gly Ile Pro Glu Leu Ile Leu Glu Gly Phe Pro Leu Trp Leu Ile
420 425 430
Tyr Trp Gly Phe Ala Asp Phe Gln Lys Asn Leu Lys Lys Val Thr Asn

435 440 445
Ile Asp Thr Asn Tyr Met Leu Val Ala Lys Thr Lys Phe Thr Gln Lys
450 455 460
Pro Gly Thr Phe Tyr Leu Val Ile Leu Asn Asp Thr Phe Val Glu Gly
465 470 475 480
Asn Ser Pro Tyr Glu Lys Gln Pro Leu Pro Glu Asp Asn Ile Lys Trp
485 490 495
Tyr Pro Gln Val Gln Tyr Gln Leu Glu Ala Gln Asn Lys Leu Leu Gln

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

565 570 575
Glu Ser Ile Leu Tyr Ser Phe Asp Tyr Arg His Gly Asn Tyr Thr Thr

580 585 590

Thr Ala Leu Ser Arg Ile Ser Gln Asp Trp Ala Leu Lys Asp Thr Val

595 600 605

Ser Lys Ile Thr Glu Pro Asp Arg Gln Gln Leu Leu Lys Gln Ala Leu

610 615 620

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

625 630 635 640

Val Gln Gln Leu Ile Ser Asn Leu Arg Gln Gln Gln Gln Leu Tyr Arg

645 650 655

Glu Arg Ile Ile Ser Leu Leu Lys Asp Gln

660 665

<210> 59

<211> 148

<212> PRT

<213> Betatorquevirus sp.

<400> 59

Met Pro Tyr Tyr Tyr Arg Arg Arg Arg Tyr Asn Tyr Arg Arg Pro Arg

1 5 10 15

Trp Tyr Gly Arg Gly Trp Ile Arg Arg Pro Phe Arg Arg Arg Phe Arg

20 25 30

Arg Lys Arg Arg Ile Gln Tyr Pro Ile Pro Arg Asn Glu His Glu Thr

35 40 45

Thr Ser Leu Gln Ser Pro Gly Glu Ala Pro Glu Ser Ile Leu Tyr Ser

50 55 60

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

65 70 75 80

Ser Gln Asp Trp Ala Leu Lys Asp Thr Val Ser Lys Ile Thr Glu Pro

85 90 95

Asp Arg Gln Gln Leu Leu Lys Gln Ala Leu Glu Cys Leu Gln Ile Ser

100 105 110

Glu Glu Thr Gln Glu Lys Lys Glu Lys Glu Val Gln Gln Leu Ile Ser

115 120 125

Asn Leu Arg Gln Gln Gln Gln Leu Tyr Arg Glu Arg Ile Ile Ser Leu

130 135 140

Leu Lys Asp Gln

145

<210> 60

<211> 82

<212> PRT

<213> Betatorquevirus sp.

<400> 60

Met Pro Tyr Tyr Tyr Arg Arg Arg Arg Tyr Asn Tyr Arg Arg Pro Arg

1 5 10 15

Trp Tyr Gly Arg Gly Trp Ile Arg Arg Pro Phe Arg Arg Arg Phe Arg

20 25 30

Arg Lys Arg Arg Ser Gln Ile Asp Ser Asn Cys Ser Asn Lys Pro Ser

35 40 45

Asn Ala Cys Lys Ser Arg Lys Lys Arg Arg Arg Lys Lys Lys Lys Lys

50 55 60

Tyr Ser Ser Ser Ser Ala Thr Ser Asp Ser Ser Ser Ser Cys Thr Glu

65 70 75 80

Ser Glu

<210> 61

<400> 61

000

<210> 62

<400> 62

000

<210> 63

<400> 63

000

<210> 64

<400> 64

000

<210> 65

<400> 65

000

<210> 66

<400> 66

000

<210> 67

<400> 67

000

<210> 68

<400> 68

000

<210> 69

<400> 69

000

<210> 70

<400> 70

000

<210> 71

<

400> 71

000

<210> 72

<400> 72

000

<210> 73

<400> 73

000

<210> 74

<400> 74

000

<210> 75

<400> 75

000

<210> 76

<400> 76

000

<210> 77

<400> 77

000

<210> 78

<400> 78

000

<210> 79

<400> 79

000

<210> 80

<400> 80

000

<210> 81

<400> 81

000

<210> 82

<400> 82

000

<210> 83

<400> 83

000

<210> 84

<400> 84

000

<210> 85

<400> 85

000

<210> 86

<400> 86

000

<210> 87

<400> 87

000

<210> 88

<400> 88

000

<210> 89

<400> 89

000

<210> 90

<400> 90

000

<210> 91

<400> 91

000

<210> 92

<400> 92

000

<210> 93

<400> 93

000

<210> 94

<400> 94

000

<210> 95

<400> 95

000

<210> 96

<400> 96

000

<210> 97

<400> 97

000

<210> 98

<400> 98

000

<210> 99

<400> 99

000

<210> 100

<400> 100

000

<210> 101

<400> 101

000

<210> 102

<400> 102

000

<210> 103

<400> 103

000

<210> 104

<400> 104

000

<210> 105

<211> 71

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 105

cgggtgccgk aggtgagttt acacaccgma gtcaaggggc aattcgggct crggactggc 60

cgggcyhtgg g 71

<210> 106

<211> 71

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 106

cgggtgccgg aggtgagttt acacaccgca gtcaaggggc aattcgggct cgggactggc 60

cgggctwtgg g 71

<210> 107

<211> 71

<212> DNA

<213> Artificial Sequence

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

<400> 107

cgggtgccgt aggtgagttt acacaccgca gtcaaggggc aattcgggct cgggactggc 60

cgggctatgg g 71

<210> 108

<211> 71

<212> DNA

<213> Artificial Sequence

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

<400> 108

cgggtgccgg aggtgagttt acacaccgca gtcaaggggc aattcgggct cgggactggc 60

cgggccctgg g 71

<210> 109

<211> 71

<212> DNA

<213> Artificial Sequence

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

<400> 109

cgggtgccgg aggtgagttt acacaccgca gtcaaggggc aattcgggct cgggactggc 60

cgggctttgg g 71

<210> 110

<211> 71

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 110

cgggtgccgg aggtgagttt acacaccgca gtcaaggggc aattcgggct cgggactggc 60

cgggctatgg g 71

<210> 111

<211> 71

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 111

cgggtgccgg aggtgagttt acacaccgaa gtcaaggggc aattcgggct caggactggc 60

cgggctttgg g 71

<210> 112

<211> 71

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 112

cgggtgccgg aggtgagttt acacaccgca gtcaaggggc aattcgggct cgggactggc 60

cgggcyhtgg g 71

<210> 113

<211> 71

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 113

cgggtgccgt aggtgagttt acacaccgca gtcaaggggc aattcgggct cgggactggc 60

cgggctatgg g 71

<210> 114

<211> 70

<212> DNA

<213> Artificial Sequence

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

<400> 114

cgggtgccgg aggtgagttt acacaccgca gtcaaggggc aattcgggct cgggactggc 60

cgggcccggg 70

<210> 115

<211> 71

<212> DNA

<213> Artificial Sequence

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

<400> 115

cgggtgccgg aggtgagttt acacaccgaa gtcaaggggc aattcgggct caggactggc 60

cgggctttgg g 71

<210> 116

<211> 69

<212> DNA

<213> Artificial Sequence

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

<400> 116

cgggtgccgg aggtgagttt acacaccgca gtcaaggggc aattcgggct cgggaggccg 60

ggccatggg 69

<210> 117

<211> 71

<212> DNA

<213> Artificial Sequence

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

<400> 117

cgggtgccgg aggtgagttt acacaccgca gtcaaggggc aattcgggct cgggactggc 60

cgggccccgg g 71

<210> 118

<211> 71

<212> DNA

<213> Artificial Sequence

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

<400> 118

cgggtgccgg aggtgagttt acacaccgca gtcaaggggc aattcgggct cgggactggc 60

cgggctatgg g 71

<210> 119

<211> 71

<212> DNA

<213> Artificial Sequence

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

<400> 119

cgggtgccga aggtgagttt acacaccgca gtcaaggggc aattcgggct cgggactggc 60

cgggctatgg g 71

<210> 120

<211> 117

<212> DNA

<213> Artificial Sequence

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

<220><221> misc_feature

<222> (10)..(10)

<223> May or may not be present

<220><221> misc_feature

<222> (12)..(12)

<223> May or may not be present

<220><221> misc_feature

<222> (30)..(32)

<223> May or may not be present

<220><221> misc_feature

<222> (34)..(34)

<223> May or may not be present

<220><221> misc_feature

<222> (43)..(46)

<223> May or may not be present

<220><221> misc_feature

<222> (52)..(54)

<223> May or may not be present

<220><221> misc_feature

<222> (70)..(71)

<223> May or may not be present

<220><221> misc_feature

<222> (89)..(90)

<223> May or may not be present

<220><221> misc_feature

<222> (103)..(103)

<223> May or may not be present

<400> 120

cggcggsggs gcsscgcgct dcgcgcgcsg cccrsyrggg grdssmmwgc skcsccccccc 60

cscgcgcatg cgcrcgggkc ccccccccyv sgggggggctc cgccccccccg gcccccc 117

<210> 121

<211> 169

<212> DNA

<213> Artificial Sequence

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

polynucleotide

<220><221> modified_base

<222> (20)..(20)

<223> a, c, t, g, unknown or other

<220><221> modified_base

<222> (22)..(22)

<223> a, c, t, g, unknown or other

<220><221> modified_base

<222> (40)..(42)

<223> a, c, t, g, unknown or other

<220><221> modified_base

<222> (53)..(56)

<223> a, c, t, g, unknown or other

<220><221> modified_base

<222> (62)..(62)

<223> a, c, t, g, unknown or other

<220><221> modified_base

<222> (64)..(64)

<223> a, c, t, g, unknown or other

<220><221> modified_base

<222> (97)..(98)

<223> a, c, t, g, unknown or other

<400> 121

gccgccgcgg cggcggsggn gnsgcgcgct dcgcgcgcsn nncrccrggg ggnnnncwgc 60

sncnccccccc cccgcgcatg cgcgggkccc ccccccnncg gggggctccg ccccccggcc 120

cccccccgtg ctaaacccac cgcgcatgcg cgaccacgcc cccgccgcc 169

<210> 122

<211> 79

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<220><221> modified_base

<222> (20)..(20)

<223> a, c, t, g, unknown or other

<220><221> modified_base

<222> (22)..(22)

<223> a, c, t, g, unknown or other

<220><221> modified_base

<222> (40)..(42)

<

223> a, c, t, g, unknown or other

<220><221> modified_base

<222> (53)..(56)

<223> a, c, t, g, unknown or other

<220><221> modified_base

<222> (62)..(62)

<223> a, c, t, g, unknown or other

<220><221> modified_base

<222> (64)..(64)

<223> a, c, t, g, unknown or other

<400> 122

gccgccgcgg cggcggsggn gnsgcgcgct dcgcgcgcsn nncrccrggg ggnnnncwgc 60

sncncccccc cccgcgcat 79

<210> 123

<211> 31

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<220><221> modified_base

<222> (18)..(19)

<223> a, c, t, g, unknown or other

<400> 123

gcgcgggkcc ccccccnnc gggggctcc g 31

<210> 124

<211> 59

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 124

ccccccggcc ccccccccgtg ctaaacccac cgcgcatgcg cgaccacgcc cccgccgcc 59

<210> 125

<211> 156

<212> DNA

<213> Artificial Sequence

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

<400> 125

gcggcggggg ggcggccgcg ttcgcgcgcc gcccaccagg gggtgctgcg cgccccccccc 60

cgcgcatgcg cggggccccc ccccgggggg gctccgcccc cccggccccc ccccgtgcta 120

aacccaccgc gcatgcgcga ccacgccccc gccgcc 156

<210> 126

<211> 7

<212> DNA

<213> Artificial Sequence

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

<400

> 126

gcggcgg 7

<210> 127

<211> 7

<212> DNA

<213> Artificial Sequence

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

<400> 127

ggggggcg 7

<210> 128

<211> 6

<212> DNA

<213> Artificial Sequence

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

<400> 128

gccgcg 6

<210> 129

<211> 25

<212> DNA

<213> Artificial Sequence

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

<400> 129

ttcgcgcgcc gcccaccagg gggtg 25

<210> 130

<211> 5

<212> DNA

<213> Artificial Sequence

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

<400> 130

ctgcg 5

<210> 131

<211> 17

<212> DNA

<213> Artificial Sequence

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

<400> 131

cgccccccc cgcgcat 17

<210> 132

<211> 17

<212> DNA

<213> Artificial Sequence

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

<400> 132

gcgcggggcc ccccccc 17

<210> 133

<211> 72

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 133

gggggggctc cgcccccccg gccccccccc gtgctaaacc caccgcgcat gcgcgaccac 60

gcccccgccg cc 72

<210> 134

<211> 115

<212> DNA

<213> Artificial Sequence

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

polynucleotide

<400> 134

cggcggcggc ggcgcgcgcg ctgcgcgcgc gcgccggggg ggcgccagcg ccccccccccc 60

cgcgcatgca cgggtccccc cccccacggg gggctccgcc ccccggcccc ccccc 115

<210> 135

<211> 14

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 135

cggcggcggc ggcg 14

<210> 136

<211> 17

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 136

cgcgcgctgc gcgcgcg 17

<210> 137

<211> 19

<212> DNA

<213> Artificial Sequence

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

<400> 137

cgccgggggg gcgccagcg 19

<210> 138

<211> 17

<212> DNA

<213> Artificial Sequence

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

<400> 138

ccccccccc cgcgcat 17

<210> 139

<211> 31

<212> DNA

<213> Artificial Sequence

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

<400> 139

gcacgggtcc ccccccccac gggggctcc g 31

<210> 140

<211> 17

<212> DNA

<213> Artificial Sequence

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

<400> 140

ccccccggcc cccccc 17

<210> 141

<211> 121

<212> DNA

<213> Artificial Sequence

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

<400> 141

ccgtcggcgg gggggccgcg cgctgcgcgc gcggcccccg ggggaggcac agcctccccc 60

ccccgcgcgc atgcgcgcgg gtcccccccc ctccgggggg ctccgccccc cggccccccc 120

c 121

<210> 142

<211> 37

<212> DNA

<213> Artificial Sequence

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

<400> 142

ccgtcggcgg gggggccgcg cgctgcgcgc gcggccc 37

<210> 143

<211> 84

<212> DNA

<213> Artificial Sequence

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

<400> 143

ccgggggagg cacagcctcc ccccccccgcg cgcatgcgcg cgggtccccc cccctccggg 60

gggctccgcc ccccggcccc cccc 84

<210> 144

<211> 104

<212> DNA

<213> Artificial Sequence

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

polynucleotide

<400> 144

cggcggcggc gcgcgcgcta cgcgcgcgcg ccgggggggct gccgcccccc ccccgcgcat 60

gcgcggggcc cccccccgcg gggggctccg cccccggcc cccc 104

<210> 145

<211> 11

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 145

cggcggcggc g 11

<210> 146

<211> 17

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 146

cgcgcgctac gcgcgcg 17

<210> 147

<211> 10

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 147

cgccgggggg 10

<210> 148

<211> 7

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 148

ctgccgc 7

<210> 149

<211> 15

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 149

ccccccccg cgcat 15

<210> 150

<211> 17

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 150

gcgcggggcc ccccccc 17

<210> 151

<211> 13

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 151

gcgggggct ccg 13

<210> 152

<211> 14

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 152

cccccggcc cccc 14

<210> 153

<211> 122

<212> DNA

<213> Artificial Sequence

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

<400> 153

gccgccgcgg cggcgggggg cggcgcgctg cgcgcgccgc ccagtagggg gagccatgcg 60

ccccccccg cgcatgcgcg gggccccccc ccgcgggggg ctccgccccc cggccccccc 120

cg 122

<210> 154

<211> 19

<212> DNA

<213> Artificial Sequence

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

<400> 154

gccgccgcgg cggcggggg 19

<210> 155

<211> 41

<212> DNA

<213> Artificial Sequence

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

<400> 155

gcggcgcgct gcgcgcgccg cccagtaggg ggagccatgc g 41

<210> 156

<211> 15

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 156

ccccccccg cgcat 15

<210> 157

<211> 17

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 157

gcgcggggcc ccccccc 17

<210> 158

<211> 13

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 158

gcgggggct ccg 13

<210> 159

<211> 17

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 159

cccccggcc ccccccg 17

<210> 160

<211> 36

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 160

cgcgctgcgc gcgccgccca gtaggggggag ccatgc 36

<210> 161

<211> 78

<212> DNA

<213> Artificial Sequence

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

<400> 161

ccgccatctt aagtagttga ggcggacggt ggcgtgagtt caaaggtcac catcagccac 60

acctactcaa aatggtgg 78

<210> 162

<211> 172

<212> DNA

<213> Artificial Sequence

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

<400> 162

cttaagtagt tgaggcggac ggtggcgtga gttcaaaggt caccatcagc cacacctact 60

caaaatggtg gacaatttct tccgggtcaa aggttacagc cgccatgtta aaacacgtga 120

cgtatgacgt cacggccgcc attttgtgac acaagatggc cgacttcctt cc 172

<210> 163

<211> 36

<212> DNA

<213> Artificial Sequence

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

<400> 163

cgcgctgcgc gcgccgccca gtaggggggag ccatgc 36

<210> 164

<211> 36

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 164

gcgctdcgcg cgcgcgccgg ggggctgcgc cccccc 36

<210> 165

<211> 36

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 165

gcgcttcgcg cgccgcccac taggggggcgt tgcgcg 36

<210> 166

<211> 36

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 166

gcgctgcgcg cgccgcccag taggggggcgc aatgcg 36

<210> 167

<211> 36

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 167

gcgctgcgcg cgcggccccc gggggaggca ttgcct 36

<210> 168

<211> 36

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 168

gcgctgcgcg cgcgcgccgg gggggcgcca gcgccc 36

<210> 169

<211> 36

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 169

gcgcttcgcg cgcgcgccgg ggggctccgc cccccc 36

<210> 170

<211> 36

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 170

gcgcttcgcg cgcgcgccgg ggggctgcgc cccccc 36

<210> 171

<211> 36

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 171

gcgctacgcg cgcgcgccgg ggggctgcgc cccccc 36

<210> 172

<211> 36

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 172

gcgctacgcg cgcgcgccgg ggggctctgc cccccc 36

<210> 173

<400> 173

000

<210> 174

<400> 174

000

<210> 175

<400> 175

000

<210> 176

<400> 176

000

<210> 177

<400> 177

000

<210> 178

<400

> 178

000

<210> 179

<400> 179

000

<210> 180

<400> 180

000

<210> 181

<400> 181

000

<210> 182

<400> 182

000

<210> 183

<400> 183

000

<210> 184

<400> 184

000

<210> 185

<211> 743

<212> PRT

<213> Alphatorquevirus sp.

<400> 185

Met Ala Trp Gly Trp Trp Lys Arg Arg Arg Arg Trp Trp Phe Arg Lys
1 5 10 15

Arg Trp Thr Arg Gly Arg Leu Arg Arg Arg Trp Pro Arg Ser Ala Arg
20 25 30

Arg Arg Pro Arg Arg Arg Arg Val Arg Arg Arg Arg Arg Trp Arg Arg
35 40 45

Gly Arg Arg Lys Thr Arg Thr Tyr Arg Arg Arg Arg Arg Phe Arg Arg
50 55 60

Arg Gly Arg Lys Ala Lys Leu Ile Ile Lys Leu Trp Gln Pro Ala Val
65 70 75 80

Ile Lys Arg Cys Arg Ile Lys Gly Tyr Ile Pro Leu Ile Ile Ser Gly
85 90 95

Asn Gly Thr Phe Ala Thr Asn Phe Thr Ser His Ile Asn Asp Arg Ile
100 105 110

Met Lys Gly Pro Phe Gly Gly Gly His Ser Thr Met Arg Phe Ser Leu
115 120 125

Tyr Ile Leu Phe Glu Glu His Leu Arg His Met Asn Phe Trp Thr Arg
130 135 140

Ser Asn Asp Asn Leu Glu Leu Thr Arg Tyr Leu Gly Ala Ser Val Lys
145 150 155 160

Ile Tyr Arg His Pro Asp Gln Asp Phe Ile Val Ile Tyr Asn Arg Arg
165 170 175

Thr Pro Leu Gly Gly Asn Ile Tyr Thr Ala Pro Ser Leu His Pro Gly
180 185 190

Asn Ala Ile Leu Ala Lys His Lys Ile Leu Val Pro Ser Leu Gln Thr
195 200 205

Arg Pro Lys Gly Arg Lys Ala Ile Arg Leu Arg Ile Ala Pro Pro Thr
210 215 220

Leu Phe Thr Asp Lys Trp Tyr Phe Gln Lys Asp Ile Ala Asp Leu Thr
225 230 235 240

Leu Phe Asn Ile Met Ala Val Glu Ala Asp Leu Arg Phe Pro Phe Cys
245 250 255

Ser Pro Gln Thr Asp Asn Thr Cys Ile Ser Phe Gln Val Leu Ser Ser
260 265 270

Val Tyr Asn Asn Tyr Leu Ser Ile Asn Thr Phe Asn Asn Asp Asn Ser
275 280 285

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

Thr Lys Gly Thr Ser Leu Asn Ala Leu Asn Thr Phe Arg Thr Glu Gly
305 310 315 320

Cys Ile Ser His Pro Gln Leu Lys Lys Pro Asn Pro Gln Ile Asn Lys
325 330 335

Pro Leu Glu Ser Gln Tyr Phe Ala Pro Leu Asp Ala Leu Trp Gly Asp
340 345 350

Pro Ile Tyr Tyr Asn Asp Leu Asn Glu Asn Lys Ser Leu Asn Asp Ile
355 360 365

Ile Glu Lys Ile Leu Ile Lys Asn Met Ile Thr Tyr His Ala Lys Leu
370 375 380

Arg Glu Phe Pro Asn Ser Tyr Gln Gly Asn Lys Ala Phe Cys His Leu
385 390 395 400

Thr Gly Ile Tyr Ser Pro Pro Tyr Leu Asn Gln Gly Arg Ile Ser Pro
405 410 415

Glu Ile Phe Gly Leu Tyr Thr Glu Ile Ile Tyr Asn Pro Tyr Thr Asp
420 425 430

Lys Gly Thr Gly Asn Lys Val Trp Met Asp Pro Leu Thr Lys Glu Asn

435 440 445

Asn Ile Tyr Lys Glu Gly Gln Ser Lys Cys Leu Leu Thr Asp Met Pro

450 455 460

Leu Trp Thr Leu Leu Phe Gly Tyr Thr Asp Trp Cys Lys Lys Asp Thr

465 470 475 480

Asn Asn Trp Asp Leu Pro Leu Asn Tyr Arg Leu Val Leu Ile Cys Pro

485 490 495

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

500 505 510

Ile Pro Tyr Ser Tyr Lys Phe Gly Ala Gly Gln Met Pro Asp Gly Ser

515 520 525

Asn Tyr Ile Pro Phe Gln Phe Arg Ala Lys Trp Tyr Pro Thr Val Leu

530 535 540

His Gln Gln Gln Val Met Glu Asp Ile Ser Arg Ser Gly Pro Phe Ala

545 550 555 560

Pro Lys Val Glu Lys Pro Ser Thr Gln Leu Val Met Lys Tyr Cys Phe

565 570 575

Asn Phe Asn Trp Gly Gly Asn Pro Ile Ile Glu Gln Ile Val Lys Asp

580 585 590

Pro Ser Phe Gln Pro Thr Tyr Glu Ile Pro Gly Thr Gly Asn Ile Pro

595 600 605

Arg Arg Ile Gln Val Ile Asp Pro Arg Val Leu Gly Pro His Tyr Ser

610 615 620

Phe Arg Ser Trp Asp Met Arg Arg His Thr Phe Ser Arg Ala Ser Ile

625 630 635 640

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

645 650 655

Pro Lys Lys Pro Arg Val Asp Ile Pro Lys Gln Glu Thr Gln Glu Glu

660 665 670

Ser Ser His Ser Leu Gln Arg Glu Ser Arg Pro Trp Glu Thr Glu Glu

675 680 685

Glu Ser Glu Thr Glu Ala Leu Ser Gln Glu Ser Gln Glu Val Pro Phe
690 695 700

Gln Gln Gln Leu Gln Gln Gln Tyr Gln Glu Gln Leu Lys Leu Arg Gln
705 710 715 720

Gly Ile Lys Val Leu Phe Glu Gln Leu Ile Arg Thr Gln Gln Gly Val
725 730 735

His Val Asn Pro Cys Leu Arg
740

<210> 186

<211> 68

<212> PRT

<213> Alphatorquevirus sp.

<400> 186

Met Ala Trp Gly Trp Trp Lys Arg Arg Arg Arg Trp Trp Phe Arg Lys
1 5 10 15

Arg Trp Thr Arg Gly Arg Leu Arg Arg Arg Trp Pro Arg Ser Ala Arg
20 25 30

Arg Arg Pro Arg Arg Arg Arg Arg Val Arg Arg Arg Arg Arg Trp Arg Arg
35 40 45

Gly Arg Arg Lys Thr Arg Thr Tyr Arg Arg Arg Arg Arg Phe Arg Arg
50 55 60

Arg Gly Arg Lys
65

<210> 187

<211> 212

<212> PRT

<213> Alphatorquevirus sp.

<400> 187

Ala Lys Leu Ile Ile Lys Leu Trp Gln Pro Ala Val Ile Lys Arg Cys
1 5 10 15

Arg Ile Lys Gly Tyr Ile Pro Leu Ile Ile Ser Gly Asn Gly Thr Phe
20 25 30

Ala Thr Asn Phe Thr Ser His Ile Asn Asp Arg Ile Met Lys Gly Pro
35 40 45
Phe Gly Gly Gly His Ser Thr Met Arg Phe Ser Leu Tyr Ile Leu Phe
50 55 60
Glu Glu His Leu Arg His Met Asn Phe Trp Thr Arg Ser Asn Asp Asn
65 70 75 80
Leu Glu Leu Thr Arg Tyr Leu Gly Ala Ser Val Lys Ile Tyr Arg His
85 90 95

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

Lys Trp Tyr Phe Gln Lys Asp Ile Ala Asp Leu Thr Leu Phe Asn Ile
165 170 175
Met Ala Val Glu Ala Asp Leu Arg Phe Pro Phe Cys Ser Pro Gln Thr
180 185 190
Asp Asn Thr Cys Ile Ser Phe Gln Val Leu Ser Ser Val Tyr Asn Asn
195 200 205
Tyr Leu Ser Ile
210
<210> 188
<211> 133
<212> PRT
<213> Alphatorquevirus sp.
<400> 188

Asn Thr Phe Asn Asn Asp Asn Ser Asp Ser Lys Leu Lys Glu Phe Leu
1 5 10 15
Asn Lys Ala Phe Pro Thr Thr Gly Thr Lys Gly Thr Ser Leu Asn Ala

20 25 30

Leu Asn Thr Phe Arg Thr Glu Gly Cys Ile Ser His Pro Gln Leu Lys

35 40 45

Lys Pro Asn Pro Gln Ile Asn Lys Pro Leu Glu Ser Gln Tyr Phe Ala

50 55 60

Pro Leu Asp Ala Leu Trp Gly Asp Pro Ile Tyr Tyr Asn Asp Leu Asn

65 70 75 80

Glu Asn Lys Ser Leu Asn Asp Ile Ile Glu Lys Ile Leu Ile Lys Asn

85 90 95

Met Ile Thr Tyr His Ala Lys Leu Arg Glu Phe Pro Asn Ser Tyr Gln

100 105 110

Gly Asn Lys Ala Phe Cys His Leu Thr Gly Ile Tyr Ser Pro Pro Tyr

115 120 125

Leu Asn Gln Gly Arg

130

<210> 189

<211> 166

<212> PRT

<213> Alphatorquevirus sp.

<400> 189

Ile Ser Pro Glu Ile Phe Gly Leu Tyr Thr Glu Ile Ile Tyr Asn Pro

1 5 10 15

Tyr Thr Asp Lys Gly Thr Gly Asn Lys Val Trp Met Asp Pro Leu Thr

20 25 30

Lys Glu Asn Asn Ile Tyr Lys Glu Gly Gln Ser Lys Cys Leu Leu Thr

35 40 45

Asp Met Pro Leu Trp Thr Leu Leu Phe Gly Tyr Thr Asp Trp Cys Lys

50 55 60

Lys Asp Thr Asn Asn Trp Asp Leu Pro Leu Asn Tyr Arg Leu Val Leu

65 70 75 80

Ile Cys Pro Tyr Thr Phe Pro Lys Leu Tyr Asn Glu Lys Val Lys Asp

85 90 95

Tyr Gly Tyr Ile Pro Tyr Ser Tyr Lys Phe Gly Ala Gly Gln Met Pro
100 105 110

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

<210> 190
<211> 164
<212> PRT
<213> Alphatorquevirus sp.

<400> 190
Trp Gly Gly Asn Pro Ile Ile Glu Gln Ile Val Lys Asp Pro Ser Phe
1 5 10 15
Gln Pro Thr Tyr Glu Ile Pro Gly Thr Gly Asn Ile Pro Arg Arg Ile
20 25 30
Gln Val Ile Asp Pro Arg Val Leu Gly Pro His Tyr Ser Phe Arg Ser
35 40 45
Trp Asp Met Arg Arg His Thr Phe Ser Arg Ala Ser Ile Lys Arg Val
50 55 60

Ser Glu Gln Gln Glu Thr Ser Asp Leu Val Phe Ser Gly Pro Lys Lys
65 70 75 80
Pro Arg Val Asp Ile Pro Lys Gln Glu Thr Gln Glu Glu Ser Ser His
85 90 95
Ser Leu Gln Arg Glu Ser Arg Pro Trp Glu Thr Glu Glu Glu Ser Glu
100 105 110
Thr Glu Ala Leu Ser Gln Glu Ser Gln Glu Val Pro Phe Gln Gln Gln
115 120 125

Leu Gln Gln Gln Tyr Gln Glu Gln Leu Lys Leu Arg Gln Gly Ile Lys

130 135 140

Val Leu Phe Glu Gln Leu Ile Arg Thr Gln Gln Gly Val His Val Asn

145 150 155 160

Pro Cys Leu Arg

<210> 191

<400> 191

000

<210> 192

<400> 192

000

<210> 193

<400> 193

000

<210> 194

<400> 194

000

<210> 195

<400> 195

000

<210> 196

<400> 196

000

<210> 197

<400> 197

000

<

210> 198

<400> 198

000

<210> 199

<400> 199

000

<210> 200

<400> 200

000

<210> 201

<400> 201

000

<210> 202

<400> 202

000

<210> 203

<400> 203

000

<210> 204

<400> 204

000

<210> 205

<400> 205

000

<210> 206

<400> 206

000

<210> 207

<400> 207

000

<210> 208

<400> 208

000

<210> 209

<400> 209

000

<210> 210

<400> 210

000

<210> 211

<400> 211

000

<210> 212

<400> 212

000

<210> 213

<400> 213

000

<210> 214

<400> 214

000

<210

> 215

<211> 666

<212> PRT

<213> Betatorquevirus sp.

<400> 215

Met Pro Tyr Tyr Tyr Arg Arg Arg Arg Tyr Asn Tyr Arg Arg Pro Arg
1 5 10 15
Trp Tyr Gly Arg Gly Trp Ile Arg Arg Pro Phe Arg Arg Arg Phe Arg
20 25 30
Arg Lys Arg Arg Val Arg Pro Thr Tyr Thr Thr Ile Pro Leu Lys Gln
35 40 45
Trp Gln Pro Pro Tyr Lys Arg Thr Cys Tyr Ile Lys Gly Gln Asp Cys

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

115 120 125
Lys Gly Cys Lys Ile Thr Phe Tyr Gln Ser Thr Phe Thr Asp Tyr Ile
130 135 140
Val Arg Ile His Thr Glu Leu Pro Ala Asn Ser Asn Lys Leu Thr Tyr

145 150 155 160

Pro Asn Thr His Pro Leu Met Met Met Met Ser Lys Tyr Lys His Ile

165 170 175

Ile Pro Ser Arg Gln Thr Arg Arg Lys Lys Lys Pro Tyr Thr Lys Ile

180 185 190

Phe Val Lys Pro Pro Pro Gln Phe Glu Asn Lys Trp Tyr Phe Ala Thr

195 200 205

Asp Leu Tyr Lys Ile Pro Leu Leu Gln Ile His Cys Thr Ala Cys Asn

210 215 220

Leu Gln Asn Pro Phe Val Lys Pro Asp Lys Leu Ser Asn Asn Val Thr

225 230 235 240

Leu Trp Ser Leu Asn Thr Ile Ser Ile Gln Asn Arg Asn Met Ser Val

245 250 255

Asp Gln Gly Gln Ser Trp Pro Phe Lys Ile Leu Gly Thr Gln Ser Phe

260 265 270

Tyr Phe Tyr Phe Tyr Thr Gly Ala Asn Leu Pro Gly Asp Thr Thr Gln

275 280 285

Ile Pro Val Ala Asp Leu Leu Pro Leu Thr Asn Pro Arg Ile Asn Arg

290 295 300

Pro Gly Gln Ser Leu Asn Glu Ala Lys Ile Thr Asp His Ile Thr Phe

305 310 315 320

Thr Glu Tyr Lys Asn Lys Phe Thr Asn Tyr Trp Gly Asn Pro Phe Asn

325 330 335

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

340 345 350

Pro Glu Ala Ile Lys Asn Glu Trp Thr Thr Glu Asn Met Lys Trp Asn

355 360 365

Gln Leu Asn Asn Ala Gly Thr Met Ala Leu Thr Pro Phe Asn Glu Pro

370 375 380

Ile Phe Thr Gln Ile Gln Tyr Asn Pro Asp Arg Asp Thr Gly Glu Asp

385 390 395 400

Thr Gln Leu Tyr Leu Leu Ser Asn Ala Thr Gly Thr Gly Trp Asp Pro
405 410 415
Pro Gly Ile Pro Glu Leu Ile Leu Glu Gly Phe Pro Leu Trp Leu Ile
420 425 430
Tyr Trp Gly Phe Ala Asp Phe Gln Lys Asn Leu Lys Lys Val Thr Asn

435 440 445
Ile Asp Thr Asn Tyr Met Leu Val Ala Lys Thr Lys Phe Thr Gln Lys
450 455 460
Pro Gly Thr Phe Tyr Leu Val Ile Leu Asn Asp Thr Phe Val Glu Gly
465 470 475 480
Asn Ser Pro Tyr Glu Lys Gln Pro Leu Pro Glu Asp Asn Ile Lys Trp
485 490 495
Tyr Pro Gln Val Gln Tyr Gln Leu Glu Ala Gln Asn Lys Leu Leu Gln

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

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

625 630 635 640
Val Gln Gln Leu Ile Ser Asn Leu Arg Gln Gln Gln Gln Leu Tyr Arg

645 650 655

Glu Arg Ile Ile Ser Leu Leu Lys Asp Gln

660 665

<210> 216

<211> 38

<212> PRT

<213> Betatorquevirus sp.

<400> 216

Met Pro Tyr Tyr Tyr Arg Arg Arg Arg Tyr Asn Tyr Arg Arg Pro Arg

1 5 10 15

Trp Tyr Gly Arg Gly Trp Ile Arg Arg Pro Phe Arg Arg Arg Phe Arg

20 25 30

Arg Lys Arg Arg Val Arg

35

<210> 217

<211> 208

<212> PRT

<213> Betatorquevirus sp.

<400> 217

Pro Thr Tyr Thr Thr Ile Pro Leu Lys Gln Trp Gln Pro Pro Tyr Lys

1 5 10 15

Arg Thr Cys Tyr Ile Lys Gly Gln Asp Cys Leu Ile Tyr Tyr Ser Asn

20 25 30

Leu Arg Leu Gly Met Asn Ser Thr Met Tyr Glu Lys Ser Ile Val Pro

35 40 45

Val His Trp Pro Gly Gly Gly Ser Phe Ser Val Ser Met Leu Thr Leu

50 55 60

Asp Ala Leu Tyr Asp Ile His Lys Leu Cys Arg Asn Trp Trp Thr Ser

65 70 75 80

Thr Asn Gln Asp Leu Pro Leu Val Arg Tyr Lys Gly Cys Lys Ile Thr

85 90 95

Phe Tyr Gln Ser Thr Phe Thr Asp Tyr Ile Val Arg Ile His Thr Glu

100 105 110

Leu Pro Ala Asn Ser Asn Lys Leu Thr Tyr Pro Asn Thr His Pro Leu

115 120 125

Met Met Met Met Ser Lys Tyr Lys His Ile Ile Pro Ser Arg Gln Thr

130 135 140

Arg Arg Lys Lys Lys Pro Tyr Thr Lys Ile Phe Val Lys Pro Pro Pro

145 150 155 160

Gln Phe Glu Asn Lys Trp Tyr Phe Ala Thr Asp Leu Tyr Lys Ile Pro

165 170 175

Leu Leu Gln Ile His Cys Thr Ala Cys Asn Leu Gln Asn Pro Phe Val

180 185 190

Lys Pro Asp Lys Leu Ser Asn Asn Val Thr Leu Trp Ser Leu Asn Thr

195 200 205

<210> 218

<211> 128

<212> PRT

<213> Betatorquevirus sp.

<400> 218

Ile Ser Ile Gln Asn Arg Asn Met Ser Val Asp Gln Gly Gln Ser Trp

1 5 10 15

Pro Phe Lys Ile Leu Gly Thr Gln Ser Phe Tyr Phe Tyr Phe Tyr Thr

20 25 30

Gly Ala Asn Leu Pro Gly Asp Thr Thr Gln Ile Pro Val Ala Asp Leu

35 40 45

Leu Pro Leu Thr Asn Pro Arg Ile Asn Arg Pro Gly Gln Ser Leu Asn

50 55 60

Glu Ala Lys Ile Thr Asp His Ile Thr Phe Thr Glu Tyr Lys Asn Lys

65 70 75 80

Phe Thr Asn Tyr Trp Gly Asn Pro Phe Asn Lys His Ile Gln Glu His

85 90 95

Leu Asp Met Ile Leu Tyr Ser Leu Lys Ser Pro Glu Ala Ile Lys Asn

100 105 110

Glu Trp Thr Thr Glu Asn Met Lys Trp Asn Gln Leu Asn Asn Ala Gly
115 120 125

<210> 219

<211> 163

<212> PRT

<213> Betatorquevirus sp.

<400> 219

Thr Met Ala Leu Thr Pro Phe Asn Glu Pro Ile Phe Thr Gln Ile Gln

1 5 10 15

Tyr Asn Pro Asp Arg Asp Thr Gly Glu Asp Thr Gln Leu Tyr Leu Leu
20 25 30

Ser Asn Ala Thr Gly Thr Gly Trp Asp Pro Pro Gly Ile Pro Glu Leu
35 40 45

Ile Leu Glu Gly Phe Pro Leu Trp Leu Ile Tyr Trp Gly Phe Ala Asp
50 55 60

Phe Gln Lys Asn Leu Lys Lys Val Thr Asn Ile Asp Thr Asn Tyr Met

65 70 75 80

Leu Val Ala Lys Thr Lys Phe Thr Gln Lys Pro Gly Thr Phe Tyr Leu
85 90 95

Val Ile Leu Asn Asp Thr Phe Val Glu Gly Asn Ser Pro Tyr Glu Lys
100 105 110

Gln Pro Leu Pro Glu Asp Asn Ile Lys Trp Tyr Pro Gln Val Gln Tyr
115 120 125

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

130 135 140

Asn Ile Gln Gly Gln Leu Ser Asp Asn Ile Ser Met Phe Tyr Lys Phe
145 150 155 160

Tyr Phe Lys

<210> 220

<211> 129

<212> PRT

<213> Betatorquevirus sp.

<400> 220

Trp Gly Gly Ser Pro Pro Lys Ala Ile Asn Val Glu Asn Pro Ala His
1 5 10 15

Gln Ile Gln Tyr Pro Ile Pro Arg Asn Glu His Glu Thr Thr Ser Leu
20 25 30

Gln Ser Pro Gly Glu Ala Pro Glu Ser Ile Leu Tyr Ser Phe Asp Tyr
35 40 45

Arg His Gly Asn Tyr Thr Thr Thr Ala Leu Ser Arg Ile Ser Gln Asp
50 55 60

Trp Ala Leu Lys Asp Thr Val Ser Lys Ile Thr Glu Pro Asp Arg Gln
65 70 75 80

Gln Leu Leu Lys Gln Ala Leu Glu Cys Leu Gln Ile Ser Glu Glu Thr
85 90 95

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

Gln Gln Gln Gln Leu Tyr Arg Glu Arg Ile Ile Ser Leu Leu Lys Asp
115 120 125

Gln

<210> 221

<400> 221

000

<210> 222

<400> 222

000

<210> 223

<400> 223

000

<210> 224

<400> 224

000

<210> 225

<400> 225

000

<210> 226

<400> 226

000

<210> 227

<211> 220

<212> PRT

<213> Artificial Sequence

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

<220><221> MOD_RES

<222> (29)..(31)

<223> Any amino acid

<220><221> SITE

<222> (29)..(31)

<223> This region may encompass 0-3 residues

<220><221> MOD_RES

<222> (100)..(100)

<223> Any amino acid

<220><221> MOD_RES

<222> (125)..(129)

<223> Any amino acid

<220><221> SITE

<222> (125)..(129)

<223> This region may encompass 1-5 residues

<220><221> MOD_RES

<222> (181)..(181)

<223> Any amino acid

<220><

221> MOD_RES

<222> (211)..(211)

<223> Any amino acid

<400> 227

Leu Val Leu Thr Gln Trp Gln Pro Asn Thr Val Arg Arg Cys Tyr Ile
1 5 10 15
Arg Gly Tyr Leu Pro Leu Ile Ile Cys Gly Glu Asn Xaa Xaa Xaa Thr
20 25 30
Thr Ser Arg Asn Tyr Ala Thr His Ser Asp Asp Thr Ile Gln Lys Gly
35 40 45
Pro Phe Gly Gly Gly Met Ser Thr Thr Thr Phe Ser Leu Arg Val Leu

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

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

180 185 190
Phe Gly Ser Pro Gln Thr Asp Asn Pro Cys Val Thr Phe Gln Val Leu
195 200 205
Gly Ser Xaa Tyr Asn Lys His Leu Ser Ile Ser Pro
210 215 220
<210> 228
<211> 172
<212> PRT
<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
polypeptide
<220><221> MOD_RES
<222> (38)..(38)
<223> Any amino acid
<220><221
> MOD_RES
<222> (44)..(46)
<223> Any amino acid
<220><221> SITE
<222> (44)..(46)
<223> This region may encompass 0-3 residues
<220><221> MOD_RES
<222> (77)..(77)
<223> Any amino acid
<220><221> MOD_RES
<222> (79)..(79)
<223> Any amino acid
<220><221> MOD_RES
<222> (98)..(101)
<223> Any amino acid
<220><221> SITE
<222> (98)..(101)
<223> This region may encompass 0-4 residues
<400> 228

Ser Asn Phe Glu Phe Pro Gly Ala Tyr Thr Asp Ile Thr Tyr Asn Pro
1 5 10 15

Leu Thr Asp Lys Gly Val Gly Asn Met Val Trp Ile Gln Tyr Leu Thr
20 25 30
Lys Pro Asp Thr Ile Xaa Asp Lys Thr Gln Ser Xaa Xaa Xaa Lys Cys
35 40 45
Leu Ile Glu Asp Leu Pro Leu Trp Ala Ala Leu Tyr Gly Tyr Val Asp
50 55 60
Phe Cys Glu Lys Glu Thr Gly Asp Ser Ala Ile Ile Xaa Asn Xaa Gly

65 70 75 80

Arg Val Leu Ile Arg Cys Pro Tyr Thr Lys Pro Pro Leu Tyr Asp Lys
85 90 95

Thr Xaa Xaa Xaa Xaa Asn Lys Gly Phe Val Pro Tyr Ser Thr Asn Phe
100 105 110

Gly Asn Gly Lys Met Pro Gly Gly Ser Gly Tyr Val Pro Ile Tyr Trp
115 120 125

Arg Ala Arg Trp Tyr Pro Thr Leu Phe His Gln Lys Glu Val Leu Glu
130 135 140

Asp Ile Val Gln Ser Gly Pro Phe Ala Tyr Lys Asp Glu Lys Pro Ser
145 150 155 160

Thr Gln Leu Val Met Lys Tyr Cys Phe Asn Phe Asn
165 170

<210> 229

<211> 258

<212> PRT

<213> Artificial Sequence

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

<220><221> MOD_RES

<222> (20)..(22)

<223> Any amino acid

<220><221> SITE

<222> (20)..(22)

<223> This region may encompass 0-3 residues

<220><221> MOD_RES

<222> (25)..(25)

<223> Any amino acid

<220><221> MOD_RES

<222> (78)..(78)

<223> Any amino acid

<220><221> MOD_RES

<222> (89)..(89)

<223> Any amino acid

<220><221> MOD_RES

<222> (91)..(91)

<223> Any amino acid

<220><221> MOD_RES

<222> (95)..(98)

<223> Any amino acid

<220><221> SITE

<222> (95)..(98)

<223> This region may encompass 1-4 residues

<220><221> MOD_RES

<222> (107)..(120)

<223> Any amino acid

<220><221> SITE

<222> (107)..(120)

<223> This region may encompass 2-14 residues

<220><221> MOD_RES

<222> (129)..(129)

<223> Any amino acid

<220><221> MOD_RES

<222> (139)..(168)

<223> Any amino acid

<220><221> SITE

<222> (139)..(168)

<223> This region may encompass 0-30 residues

<220><221> MOD_RES

<222> (201)..(204)

<223> Any amino acid

<220><221> SITE

<222> (201)..(204)

<223> This region may encompass 0-4 residues

<220><221> MOD_RES

<222> (219)..(258)

<223> Any amino acid

<220><221> SITE

<222> (219)..(258)

<223> This region may encompass 0-40 residues

<400> 229

Trp Gly Gly Asn Pro Ile Ser Gln Gln Val Val Arg Asn Pro Cys Lys
1 5 10 15

Asp Ser Gly Xaa Xaa Xaa Ser Gly Xaa Gly Arg Gln Pro Arg Ser Val
20 25 30

Gln Val Val Asp Pro Lys Tyr Met Gly Pro Glu Tyr Thr Phe His Ser
35 40 45

Trp Asp Trp Arg Arg Gly Leu Phe Gly Glu Lys Ala Ile Lys Arg Met
50 55 60

Ser Glu Gln Pro Thr Asp Asp Glu Ile Phe Thr Gly Gly Xaa Pro Lys
65 70 75 80

Arg Pro Arg Arg Asp Pro Pro Thr Xaa Gln Xaa Pro Glu Glu Xaa Xaa
85 90 95

Xaa Xaa Gln Lys Glu Ser Ser Ser Phe Arg Xaa Xaa Xaa Xaa Xaa Xaa
100 105 110

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Pro Trp Glu Ser Ser Ser Gln Glu
115 120 125

Xaa Glu Ser Glu Ser Gln Glu Glu Glu Glu Xaa Xaa Xaa Xaa Xaa Xaa
130 135 140

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
145 150 155 160

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Glu Gln Thr Val Gln Gln Gln Leu
165 170 175

Arg Gln Gln Leu Arg Glu Gln Arg Arg Leu Arg Val Gln Leu Gln Leu
180 185 190

Leu Phe Gln Gln Leu Leu Lys Thr Xaa Xaa Xaa Xaa Gln Ala Gly Leu
195 200 205

His Ile Asn Pro Leu Leu Leu Ser Gln Ala Xaa Xaa Xaa Xaa Xaa Xaa
210 215 220
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
225 230 235 240
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
245 250 255

Xaa Xaa

<210> 230
<211> 214
<212> PRT
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic polypeptide
<220><221> MOD_RES
<222> (136)..(136)
<223> Any amino acid
<220><221> MOD_RES
<222> (138)..(141)
<223> Any amino acid
<220><221> SITE
<222> (138)..(141)
<223> This region may encompass 1-4 residues
<220><221> MOD_RES
<222> (179)..(179)
<223> Any amino acid
<400> 230
Leu Lys Gln Trp Gln Pro Ser Thr Ile Arg Lys Cys Lys Ile Lys Gly

1 5 10 15
Tyr Leu Pro Leu Phe Gln Cys Gly Lys Gly Arg Ile Ser Asn Asn Tyr
20 25 30
Thr Gln Tyr Lys Glu Ser Ile Val Pro His His Glu Pro Gly Gly Gly
35 40 45

Gly Trp Ser Ile Gln Gln Phe Thr Leu Gly Ala Leu Tyr Glu Glu His
50 55 60
Leu Lys Leu Arg Asn Trp Trp Thr Lys Ser Asn Asp Gly Leu Pro Leu

65 70 75 80
Val Arg Tyr Leu Gly Cys Thr Ile Lys Leu Tyr Arg Ser Glu Asp Thr
85 90 95
Asp Tyr Ile Val Thr Tyr Gln Arg Cys Tyr Pro Met Thr Ala Thr Lys
100 105 110
Leu Thr Tyr Leu Ser Thr Gln Pro Ser Arg Met Leu Met Asn Lys His
115 120 125
Lys Ile Ile Val Pro Ser Lys Xaa Thr Xaa Xaa Xaa Xaa Asn Lys Lys

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

195 200 205
Gln Asn Pro Asn Phe Gln
210
<210> 231
<211> 187
<212> PRT
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
polypeptide
<220><221> MOD_RES
<222> (1)..(10)
<223> Any amino acid
<220><221> SITE

<222> (1)..(10)

<223> This region may encompass 4-10 residues

<220><221> MOD_RES

<222> (38)..(45)

<223> Any amino acid

<220><221> SITE

<222> (38)..(45)

<223

> This region may encompass 1-8 residues

<220><221> MOD_RES

<222> (94)..(94)

<223> Any amino acid

<220><221> MOD_RES

<222> (100)..(102)

<223> Any amino acid

<220><221> SITE

<222> (100)..(102)

<223> This region may encompass 1-3 residues

<220><221> MOD_RES

<222> (112)..(112)

<223> Any amino acid

<220><221> MOD_RES

<222> (114)..(115)

<223> Any amino acid

<220><221> SITE

<222> (114)..(115)

<223> This region may encompass 0-2 residues

<220><221> MOD_RES

<222> (124)..(139)

<223> Any amino acid

<220><221> SITE

<222> (124)..(139)

<223> This region may encompass 3-16 residues

<220><221> MOD_RES

<222> (154)..(154)

<223> Any amino acid

<400> 231

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Thr Pro Leu Tyr Phe Glu
1 5 10 15

Cys Arg Tyr Asn Pro Phe Lys Asp Lys Gly Thr Gly Asn Lys Val Tyr
20 25 30

Leu Val Ser Asn Asn Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Thr Gly Trp
35 40 45

Asp Pro Pro Thr Asp Pro Asp Leu Ile Ile Glu Gly Phe Pro Leu Trp
50 55 60

Leu Leu Leu Trp Gly Trp Leu Asp Trp Gln Lys Lys Leu Gly Lys Ile
65 70 75 80

Gln Asn Ile Asp Thr Asp Tyr Ile Leu Val Ile Gln Ser Xaa Tyr Tyr
85 90 95

Ile Pro Pro Xaa Xaa Xaa Lys Leu Pro Tyr Tyr Val Pro Leu Asp Xaa
100 105 110

Asp Xaa Xaa Phe Leu His Gly Arg Ser Pro Tyr Xaa Xaa Xaa Xaa Xaa
115 120 125

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Pro Ser Asp Lys Gln
130 135 140

His Trp His Pro Lys Val Arg Phe Gln Xaa Glu Thr Ile Asn Asn Ile
145 150 155 160

Ala Leu Thr Gly Pro Gly Thr Pro Lys Leu Pro Asn Gln Lys Ser Ile
165 170 175

Gln Ala His Met Lys Tyr Lys Phe Tyr Phe Lys
180 185

<210> 232

<211> 163

<212> PRT

<213> Artificial Sequence

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

polypeptide

<220><221> MOD_RES

<222> (34)..(34)

<223> Any amino acid

<220><221> MOD_RES

<222> (65)..(65)

<223> Any amino acid

<220><221> MOD_RES

<222> (77)..(78)

<223> Any amino acid

<220><221> MOD_RES

<222> (86)..(87)

<223> Any amino acid

<220><221> MOD_RES

<222> (96)..(96)

<223> Any amino acid

<220><221> MOD_RES

<222> (102)..(106)

<223> Any amino acid

<220><221> SITE

<222> (102)..(106)

<223> This region may encompass 0-5 residues

<220><221> MOD_RES

<222> (125)..(125)

<223> Any amino acid

<220><221> MOD_RES

<222> (135)..(135)

<223> Any amino acid

<220><221> MOD_RES

<222> (138)..(163)

<223> Any amino acid

<220><221> SITE

<222> (138)..(163)

<223> This region may encompass 0-26 residues

<400> 232

Trp Gly Gly Cys Pro Ala Pro Met Glu Thr Ile Thr Asp Pro Cys Lys
1 5 10 15

Gln Pro Lys Tyr Pro Ile Pro Asn Asn Leu Leu Gln Thr Thr Ser Leu
20 25 30

Gln Xaa Pro Thr Thr Pro Ile Glu Thr Tyr Leu Tyr Lys Phe Asp Glu
35 40 45

Arg Arg Gly Leu Leu Thr Lys Lys Ala Ala Lys Arg Ile Lys Lys Asp
50 55 60

Xaa Thr Thr Glu Thr Thr Leu Phe Thr Asp Thr Gly Xaa Xaa Thr Ser
65 70 75 80

Thr Thr Leu Pro Thr Xaa Xaa Gln Thr Glu Thr Thr Gln Glu Glu Xaa
85 90 95

Thr Ser Glu Glu Glu Xaa Xaa Xaa Xaa Xaa Glu Thr Leu Leu Gln Gln
100 105 110

Leu Gln Gln Leu Arg Arg Lys Gln Lys Gln Leu Arg Xaa Arg Ile Leu
115 120 125

Gln Leu Leu Gln Leu Leu Xaa Leu Leu Xaa Xaa Xaa Xaa Xaa Xaa Xaa
130 135 140

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
145 150 155 160

Xaa Xaa Xaa

<210> 233

<211> 203

<212> PRT

<213> Artificial Sequence

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

<220><221> MOD_RES

<222> (79)..(79)

<223> Any amino acid

<220><221> MOD_RES

<222> (104)..(104)

<223> Any amino acid

<220><221> MOD_RES

<222> (116)..(116)

<223> Any amino acid

<220><221> MOD_RES

<222> (120)..(121)

<223> Any amino acid

<220><221> MOD_RES

<222> (125)..(125)

<223> Any amino acid

<220><221> MOD_RES

<222> (170)..(170)

<223> Any amino acid

<400> 233

Thr Ile Pro Leu Lys Gln Trp Gln Pro Glu Ser Ile Arg Lys Cys Lys
1 5 10 15

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

20 25 30

Tyr Cys Tyr Thr Asn Glu Lys Asp Glu Tyr Thr Pro Pro Lys Ala Pro
35 40 45

Gly Gly Gly Gly Phe Gly Val Glu Leu Phe Ser Leu Glu Tyr Leu Tyr
50 55 60

Glu Gln Trp Lys Ala Arg Asn Asn Ile Trp Thr Lys Ser Asn Xaa Tyr
65 70 75 80

Lys Asp Leu Cys Arg Tyr Thr Gly Cys Lys Ile Thr Phe Tyr Arg His

85 90 95

Pro Thr Thr Asp Phe Ile Val Xaa Tyr Ser Arg Gln Pro Pro Phe Glu
100 105 110

Ile Asp Lys Xaa Thr Tyr Met Xaa Xaa His Pro Gln Xaa Leu Leu Leu

115 120 125

Arg Lys His Lys Lys Ile Ile Leu Ser Lys Ala Thr Asn Pro Lys Gly

130 135 140

Lys Leu Lys Lys Lys Ile Lys Ile Lys Pro Pro Lys Gln Met Leu Asn

145 150 155 160

Lys Trp Phe Phe Gln Lys Gln Phe Ala Xaa Tyr Gly Leu Val Gln Leu

165 170 175

Gln Ala Ala Ala Cys Asx Leu Arg Tyr Pro Arg Leu Gly Cys Cys Asn

180 185 190

Glu Asn Arg Leu Ile Thr Leu Tyr Tyr Leu Asn

195 200

<210> 234

<211> 162

<212> PRT

<213> Artificial Sequence

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

polypeptide

<220><221> MOD_RES

<222> (12)..(12)

<223> Any amino acid

<220><221> MOD_RES

<222> (20)..(20)

<223> Any amino acid

<220><221> MOD_RES

<222> (23)..(23)

<223> Any amino acid

<220><221> MOD_RES

<222> (30)..(30)

<223> Any amino acid

<220><221> MOD_RES

<222> (58)..(58)

<223> I or L

<220><221> MOD_RES

<222> (84)..(84)

<223> Any amino acid

<220><221> MOD_RES

<222> (90)..(90)

<223> Any amino acid

<220><221> MOD_RES

<222> (95)..(95)

<223> Any amino acid

<220><221> MOD_RES

<222> (105)..(105)

<223> Any amino acid

<220><221> MOD_RES

<222> (111)..(111)

<223> I or L

<220><221> MOD_RES

<222> (113)..(113)

<223> Any amino acid

<220><221> MOD_RES

<222> (154)..(154)

<223> Any amino acid

<220><221> MOD_RES

<222> (156)..(156)

<223> Any amino acid

<400> 234

Leu Pro Ile Val Val Ala Arg Tyr Asn Pro Ala Xaa Asp Thr Gly Lys
1 5 10 15

Gly Asn Lys Xaa Trp Leu Xaa Ser Thr Leu Asn Gly Ser Xaa Trp Ala
20 25 30

Pro Pro Thr Thr Asp Lys Asp Leu Ile Ile Glu Gly Leu Pro Leu Trp
35 40 45

Leu Ala Leu Tyr Gly Tyr Trp Ser Tyr Xaa Lys Lys Val Lys Lys Asp
50 55 60

Lys Gly Ile Leu Gln Ser His Met Phe Val Val Lys Ser Pro Ala Ile

65 70 75 80

Gln Pro Leu Xaa Thr Ala Thr Thr Gln Xaa Thr Phe Tyr Pro Xaa Ile

85 90 95

Asp Asn Ser Phe Ile Gln Gly Lys Xaa Pro Tyr Asp Glu Pro Xaa Thr

100 105 110

Xaa Asn Gln Lys Lys Leu Trp Tyr Pro Thr Leu Glu His Gln Gln Glu

115 120 125

Thr Ile Asn Ala Ile Val Glu Ser Gly Pro Tyr Val Pro Lys Leu Asp

130 135 140

Asn Gln Lys Asn Ser Thr Trp Glu Leu Xaa Tyr Xaa Tyr Thr Phe Tyr

145 150 155 160

Phe Lys

<210> 235

<211> 177

<212> PRT

<213> Artificial Sequence

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

<220><221> MOD_RES

<222> (16)..(16)

<223> Any amino acid

<220><221> MOD_RES

<222> (26)..(26)

<223> Any amino acid

<220><221> MOD_RES

<222> (33)..(33)

<223> Any amino acid

<220><221> MOD_RES

<222> (73)..(73)

<223> Any amino acid

<220><221> MOD_RES

<222> (81)..(82)

<223> Any amino acid

<220><221> SITE

<222> (81)..(82)

<223> This region may encompass 0-2 residues

<220><221> MOD_RES

<222> (90)..(90)

<223> Any amino acid

<220><221> MOD_RES

<222> (94)..(94)

<223> Any amino acid

<220><221> MOD_RES

<222> (119)..(124)

<223> Any amino acid

<220><221> SITE

<222> (119)..(124)

<223> This region may encompass 1-6 residues

<220><221> MOD_RES

<222> (168)..(177)

<223> Any amino acid

<220><221> SITE

<222> (168)..(177)

<223> This region may encompass 1-10 residues

<400> 235

Trp Gly Gly Pro Gln Ile Pro Asp Gln Pro Val Glu Asp Pro Lys Xaa
1 5 10 15

Gln Gly Thr Tyr Pro Val Pro Asp Thr Xaa Gln Gln Thr Ile Gln Ile
20 25 30

Xaa Asn Pro Leu Lys Gln Lys Pro Glu Thr Met Phe His Asp Trp Asp
35 40 45

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

Asn Leu Glu Thr Asp Ser Ser Phe Xaa Ser Asp Ser Glu Glu Thr Pro
65 70 75 80
Xaa Xaa Lys Lys Lys Lys Arg Leu Thr Xaa Glu Leu Pro Xaa Pro Gln
85 90 95
Glu Glu Thr Glu Glu Ile Gln Ser Cys Leu Leu Ser Leu Cys Glu Glu
100 105 110
Ser Thr Cys Gln Glu Glu Xaa Xaa Xaa Xaa Xaa Xaa Glu Asn Leu Gln
115 120 125

Gln Leu Ile His Gln Gln Gln Gln Gln Gln Gln Gln Leu Lys His Asn
130 135 140
Ile Leu Lys Leu Leu Ser Asp Leu Lys Glx Lys Gln Arg Leu Leu Gln
145 150 155 160
Leu Gln Thr Gly Ile Leu Glu Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
165 170 175
Xaa

<210> 236
<400> 236
000
<210> 237
<400> 237
000
<210> 238
<400> 238
000
<210> 239

<400> 239
000
<210> 240
<400> 240
000
<210> 241
<400> 241

000

<210> 242

<400> 242

000

<210> 243

<400> 243

000

<210> 244

<400> 244

000

<210> 245

<400> 245

000

<210> 246

<400> 246

000

<210> 247

<400> 247

000

<210> 248

<400> 248

000

<210> 249

<400> 249

000

<210> 250

<400> 250

000

<210> 251

<400> 251

000

<210> 252

<400> 252

000

<210> 253

<400> 253

000

<210> 254

<400> 254

000

<210> 255

<400> 255

000

<210> 256

<400

> 256

000

<210> 257

<400> 257

000

<210> 258

<400> 258

000

<210> 259

<400> 259

000

<210> 260

<400> 260

000

<210> 261

<400> 261

000

<210> 262

<400> 262

000

<210> 263

<400> 263

000

<210> 264

<400> 264

000

<210> 265

<400> 265

000

<210> 266

<400> 266

000

<210> 267

<400> 267

000

<210> 268

<400> 268

000

<210> 269

<400> 269

000

<210> 270

<400> 270

000

<210> 271

<400> 271

000

<210> 272

<400> 272

000

<210> 273

<400> 273

000

<210> 274

<400> 274

000

<210> 275

<400> 275

000

<210> 276

<400> 276

000

<210> 277

<400> 277

000

<210> 278

<400> 278

000

<210> 279

<400> 279

000

<210> 280

<400> 280

000

<210> 281

<400> 281

000

<210> 282

<400> 282

000

<210> 283

<400> 283

000

<210> 284

<400> 284

000

<210> 285

<400> 285

000

<210> 286

<400> 286

000

<210> 287

<400> 287

000

<210> 288

<400> 288

000

<210> 289

<400> 289

000

<210> 290

<400> 290

000

<210> 291

<400> 291

000

<210> 292

<400> 292

000

<210> 293

<400> 293

000

<210> 294

<400> 294

000

<210> 295

<400> 295

000

<210> 296

<400> 296

000

<210> 297

<400> 297

000

<210> 298

<400> 298

000

<210> 299

<400> 299

000

<210> 300

<400> 300

000

<210> 301

<400> 301

000

<210> 302

<400> 302

000

<210> 303

<400> 303

000

<210> 304

<400> 304

000

<210> 305

<400> 305

000

<210> 306

<400> 306

000

<210> 307

<400> 307

000

<210

> 308

<400> 308

000

<210> 309

<400> 309

000

<210> 310

<400> 310

000

<210> 311

<400> 311

000

<210> 312

<400> 312

000

<210> 313

<400> 313

000

<210> 314

<400> 314

000

<210> 315

<400> 315

000

<210> 316

<400> 316

000

<210> 317

<400> 317

000

<210> 318

<400> 318

000

<210> 319

<400> 319

000

<210> 320

<400> 320

000

<210> 321

<400> 321

000

<210> 322

<400> 322

000

<210> 323

<400> 323

000

<210> 324

<400> 324

000

<210> 325

<400> 325

000

<210> 326

<400> 326

000

<210> 327

<400> 327

000

<210> 328

<400> 328

000

<210> 329

<400> 329

000

<210> 330

<400> 330

000

<210> 331

<400> 331

000

<210> 332

<400> 332

000

<210> 333

<400> 333

000

<210> 334

<400> 334

000

<210> 335

<400> 335

000

<210> 336

<400> 336

000

<210> 337

<400> 337

000

<210> 338

<400> 338

000

<210> 339

<400> 339

000

<210> 340

<400> 340

000

<210> 341

<400> 341

000

<210> 342

<400

> 342

000

<210> 343

<400> 343

000

<210> 344

<400> 344

000

<210> 345

<400> 345

000

<210> 346

<400> 346

000

<210> 347

<400> 347

000

<210> 348

<400> 348

000

<210> 349

<400> 349

000

<210> 350

<400> 350

000

<210> 351

<400> 351

000

<210> 352

<400> 352

000

<210> 353

<400> 353

000

<210> 354

<400> 354

000

<210> 355

<400> 355

000

<210> 356

<400> 356

000

<210> 357

<400> 357

000

<210> 358

<400> 358

000

<210> 359

<400> 359

000

<210> 360

<400> 360

000

<210> 361

<400> 361

000

<210> 362

<400> 362

000

<210> 363

<400> 363

000

<210> 364

<400> 364

000

<210> 365

<400> 365

000

<210> 366

<400> 366

000

<210> 367

<400> 367

000

<210> 368

<400> 368

000

<210> 369

<400> 369

000

<210> 370

<400> 370

000

<210> 371

<400> 371

000

<210> 372

<400> 372

000

<210> 373

<400> 373

000

<210> 374

<400> 374

000

<210> 375

<400> 375

000

<210> 376

<400> 376

000

<210> 377

<400> 377

000

<210> 378

<400> 378

000

<210> 379

<400> 379

000

<210> 380

<400> 380

000

<210> 381

<400> 381

000

<210> 382

<400> 382

000

<210> 383

<400> 383

000

<210> 384

<400> 384

000

<210> 385

<400> 385

000

<210> 386

<400> 386

000

<210> 387

<400> 387

000

<210> 388

<400> 388

000

<210> 389

<400> 389

000

<210> 390

<400> 390

000

<210> 391

<400> 391

000

<210> 392

<400> 392

000

<210> 393

<400> 393

000

<210

> 394

<400> 394

000

<210> 395

<400> 395

000

<210> 396

<400> 396

000

<210> 397

<400> 397

000

<210> 398

<400> 398

000

<210> 399

<400> 399

000

<210> 400

<400> 400

000

<210> 401

<400> 401

000

<210> 402

<400> 402

000

<210> 403

<400> 403

000

<210> 404

<400> 404

000

<210> 405

<400> 405

000

<210> 406

<400> 406

000

<210> 407

<400> 407

000

<210> 408

<400> 408

000

<210> 409

<400> 409

000

<210> 410

<400> 410

000

<210> 411

<400> 411

000

<210> 412

<400> 412

000

<210> 413

<400> 413

000

<210> 414

<400> 414

000

<210> 415

<400> 415

000

<210> 416

<400> 416

000

<210> 417

<400> 417

000

<210> 418

<400> 418

000

<210> 419

<400> 419

000

<210> 420

<400> 420

000

<210> 421

<400> 421

000

<210> 422

<400> 422

000

<210> 423

<400> 423

000

<210> 424

<400> 424

000

<210> 425

<400> 425

000

<210> 426

<400> 426

000

<210> 427

<400> 427

000

<210> 428

<400

> 428

000

<210> 429

<400> 429

000

<210> 430

<400> 430

000

<210> 431

<400> 431

000

<210> 432

<400> 432

000

<210> 433

<400> 433

000

<210> 434

<400> 434

000

<210> 435

<400> 435

000

<210> 436

<400> 436

000

<210> 437

<400> 437

000

<210> 438

<400> 438

000

<210> 439

<400> 439

000

<210> 440

<400> 440

000

<210> 441

<400> 441

000

<210> 442

<400> 442

000

<210> 443

<400> 443

000

<210> 444

<400> 444

000

<210> 445

<400> 445

000

<210> 446

<400> 446

000

<210> 447

<400> 447

000

<210> 448

<400> 448

000

<210> 449

<400> 449

000

<210> 450

<400> 450

000

<210> 451

<400> 451

000

<210> 452

<400> 452

000

<210> 453

<400> 453

000

<210> 454

<400> 454

000

<210> 455

<400> 455

000

<210> 456

<400> 456

000

<210> 457

<400> 457

000

<210> 458

<400> 458

000

<210> 459

<400> 459

000

<210> 460

<400> 460

000

<210> 461

<400> 461

000

<210> 462

<400> 462

000

<210> 463

<400> 463

000

<210> 464

<400> 464

000

<210> 465

<400> 465

000

<210> 466

<400> 466

000

<210> 467

<400> 467

000

<210> 468

<400> 468

000

<210> 469

<400> 469

000

<210> 470

<400> 470

000

<210> 471

<400> 471

000

<210> 472

<400> 472

000

<210> 473

<400> 473

000

<210> 474

<400> 474

000

<210> 475

<400> 475

000

<210> 476

<400> 476

000

<210> 477

<400> 477

000

<210> 478

<400> 478

000

<210> 479

<400> 479

000

<210

> 480

<400> 480

000

<210> 481

<400> 481

000

<210> 482

<400> 482

000

<210> 483

<400> 483

000

<210> 484

<400> 484

000

<210> 485

<400> 485

000

<210> 486

<400> 486

000

<210> 487

<400> 487

000

<210> 488

<400> 488

000

<210> 489

<400> 489

000

<210> 490

<400> 490

000

<210> 491

<400> 491

000

<210> 492

<400> 492

000

<210> 493

<400> 493

000

<210> 494

<400> 494

000

<210> 495

<400> 495

000

<210> 496

<400> 496

000

<210> 497

<400> 497

000

<210> 498

<400> 498

000

<210> 499

<400> 499

000

<210> 500

<400> 500

000

<210> 501

<400> 501

000

<210> 502

<400> 502

000

<210> 503

<400> 503

000

<210> 504

<400> 504

000

<210> 505

<400> 505

000

<210> 506

<400> 506

000

<210> 507

<400> 507

000

<210> 508

<400> 508

000

<210> 509

<400> 509

000

<210> 510

<400> 510

000

<210> 511

<400> 511

000

<210> 512

<400> 512

000

<210> 513

<400> 513

000

<210> 514

<400

> 514

000

<210> 515

<400> 515

000

<210> 516

<400> 516

000

<210> 517

<400> 517

000

<210> 518

<400> 518

000

<210> 519

<400> 519

000

<210> 520

<400> 520

000

<210> 521

<400> 521

000

<210> 522

<400> 522

000

<210> 523

<400> 523

000

<210> 524

<400> 524

000

<210> 525

<400> 525

000

<210> 526

<400> 526

000

<210> 527

<400> 527

000

<210> 528

<400> 528

000

<210> 529

<400> 529

000

<210> 530

<400> 530

000

<210> 531

<400> 531

000

<210> 532

<400> 532

000

<210> 533

<400> 533

000

<210> 534

<400> 534

000

<210> 535

<400> 535

000

<210> 536

<400> 536

000

<210> 537

<400> 537

000

<210> 538

<400> 538

000

<210> 539

<400> 539

000

<210> 540

<400> 540

000

<210> 541

<400> 541

000

<210> 542

<400> 542

000

<210> 543

<400> 543

000

<210> 544

<400> 544

000

<210> 545

<400> 545

000

<210> 546

<400> 546

000

<210> 547

<400> 547

000

<210> 548

<400> 548

000

<210> 549

<400> 549

000

<210> 550

<400> 550

000

<210> 551

<400> 551

000

<210> 552

<400> 552

000

<210> 553

<400> 553

000

<210> 554

<400> 554

000

<210> 555

<400> 555

000

<210> 556

<400> 556

000

<210> 557

<400> 557

000

<210> 558

<400> 558

000

<210> 559

<400> 559

000

<210> 560

<400> 560

000

<210> 561

<400> 561

000

<210> 562

<400> 562

000

<210> 563

<400> 563

000

<210> 564

<400> 564

000

<210> 565

<400> 565

000

<210

> 566

<400> 566

000

<210> 567

<400> 567

000

<210> 568

<400> 568

000

<210> 569

<400> 569

000

<210> 570

<400> 570

000

<210> 571

<400> 571

000

<210> 572

<400> 572

000

<210> 573

<400> 573

000

<210> 574

<400> 574

000

<210> 575

<400> 575

000

<210> 576

<400> 576

000

<210> 577

<400> 577

000

<210> 578

<400> 578

000

<210> 579

<400> 579

000

<210> 580

<400> 580

000

<210> 581

<400> 581

000

<210> 582

<400> 582

000

<210> 583

<400> 583

000

<210> 584

<400> 584

000

<210> 585

<400> 585

000

<210> 586

<400> 586

000

<210> 587

<400> 587

000

<210> 588

<400> 588

000

<210> 589

<400> 589

000

<210> 590

<400> 590

000

<210> 591

<400> 591

000

<210> 592

<400> 592

000

<210> 593

<400> 593

000

<210> 594

<400> 594

000

<210> 595

<400> 595

000

<210> 596

<400> 596

000

<210> 597

<400> 597

000

<210> 598

<400> 598

000

<210> 599

<400> 599

000

<210> 600

<400

> 600

000

<210> 601

<400> 601

000

<210> 602

<400> 602

000

<210> 603

<400> 603

000

<210> 604

<400> 604

000

<210> 605

<400> 605

000

<210> 606

<400> 606

000

<210> 607

<400> 607

000

<210> 608

<400> 608

000

<210> 609

<400> 609

000

<210> 610

<400> 610

000

<210> 611

<400> 611

000

<210> 612

<400> 612

000

<210> 613

<400> 613

000

<210> 614

<400> 614

000

<210> 615

<400> 615

000

<210> 616

<400> 616

000

<210> 617

<400> 617

000

<210> 618

<400> 618

000

<210> 619

<400> 619

000

<210> 620

<400> 620

000

<210> 621

<400> 621

000

<210> 622

<400> 622

000

<210> 623

<400> 623

000

<210> 624

<400> 624

000

<210> 625

<400> 625

000

<210> 626

<400> 626

000

<210> 627

<400> 627

000

<210> 628

<400> 628

000

<210> 629

<400> 629

000

<210> 630

<400> 630

000

<210> 631

<400> 631

000

<210> 632

<400> 632

000

<210> 633

<400> 633

000

<210> 634

<400> 634

000

<210> 635

<400> 635

000

<210> 636

<400> 636

000

<210> 637

<400> 637

000

<210> 638

<400> 638

000

<210> 639

<400> 639

000

<210> 640

<400> 640

000

<210> 641

<400> 641

000

<210> 642

<400> 642

000

<210> 643

<400> 643

000

<210> 644

<400> 644

000

<210> 645

<400> 645

000

<210> 646

<400> 646

000

<210> 647

<400> 647

000

<210> 648

<400> 648

000

<210> 649

<400> 649

000

<210> 650

<400> 650

000

<210> 651

<400> 651

000

<210

> 652

<400> 652

000

<210> 653

<400> 653

000

<210> 654

<400> 654

000

<210> 655

<400> 655

000

<210> 656

<400> 656

000

<210> 657

<400> 657

000

<210> 658

<400> 658

000

<210> 659

<400> 659

000

<210> 660

<400> 660

000

<210> 661

<400> 661

000

<210> 662

<400> 662

000

<210> 663

<400> 663

000

<210> 664

<400> 664

000

<210> 665

<400> 665

000

<210> 666

<400> 666

000

<210> 667

<400> 667

000

<210> 668

<400> 668

000

<210> 669

<400> 669

000

<210> 670

<400> 670

000

<210> 671

<400> 671

000

<210> 672

<400> 672

000

<210> 673

<400> 673

000

<210> 674

<400> 674

000

<210> 675

<400> 675

000

<210> 676

<400> 676

000

<210> 677

<400> 677

000

<210> 678

<400> 678

000

<210> 679

<400> 679

000

<210> 680

<400> 680

000

<210> 681

<400> 681

000

<210> 682

<400> 682

000

<210> 683

<400> 683

000

<210> 684

<400> 684

000

<210> 685

<400> 685

000

<210> 686

<400

> 686

000

<210> 687

<400> 687

000

<210> 688

<400> 688

000

<210> 689

<400> 689

000

<210> 690

<211> 22

<212> DNA

<213> Artificial Sequence

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

primer

<400> 690

attcgaatgg ctgagtttat gc 22

<210> 691

<211> 24

<212> DNA

<213> Artificial Sequence

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

primer

<400> 691

cacgaattag ccaagactgg gcac 24

<210> 692

<211> 20

<212> DNA

<213> Artificial Sequence

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

primer

<400> 692

gctcccactc ctgatttctg 20

<210> 693

<211> 21

<212> DNA

<213> Artificial Sequence

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

primer

<400> 693

ccttgactac ggtggtttca c 21

<210> 694

<211> 22

<212> DNA

<213> Artificial Sequence

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

primer

<400> 694

tgcaggcatt cgagggcttg tt 22

<210> 695

<211> 20

<212> DNA

<213> Artificial Sequence

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

primer

<400> 695

tttaaccccc tagtcccagg 20

<210> 696

<400> 696

000

<210> 697

<400> 697

000

<210> 698

<400> 698

000

<210> 699

<400> 699

000

<210

> 700

<400> 700

000

<210> 701

<400> 701

000

<210> 702

<400> 702

000

<210> 703

<400> 703

000

<210> 704

<400> 704

000

<210> 705

<400> 705

000

<210> 706

<400> 706

000

<210> 707

<400> 707

000

<210> 708

<400> 708

000

<210> 709

<400> 709

000

<210> 710

<400> 710

000

<210> 711

<400> 711

000

<210> 712

<400> 712

000

<210> 713

<400> 713

000

<210> 714

<400> 714

000

<210> 715

<400> 715

000

<210> 716

<400> 716

000

<210> 717

<400> 717

000

<210> 718

<400> 718

000

<210> 719

<400> 719

000

<210> 720

<400> 720

000

<210> 721

<400> 721

000

<210> 722

<400> 722

000

<210> 723

<400> 723

000

<210> 724

<400> 724

000

<210> 725

<400> 725

000

<210> 726

<400> 726

000

<210> 727

<400> 727

000

<210> 728

<400> 728

000

<210> 729

<400> 729

000

<210> 730

<400> 730

000

<210> 731

<400> 731

000

<210> 732

<400> 732

000

<210> 733

<400> 733

000

<210> 734

<400

> 734

000

<210> 735

<400> 735

000

<210> 736

<400> 736

000

<210> 737

<400> 737

000

<210> 738

<400> 738

000

<210> 739

<400> 739

000

<210> 740

<400> 740

000

<210> 741

<400> 741

000

<210> 742

<400> 742

000

<210> 743

<400> 743

000

<210> 744

<400> 744

000

<210> 745

<400> 745

000

<210> 746

<400> 746

000

<210> 747

<400> 747

000

<210> 748

<400> 748

000

<210> 749

<400> 749

000

<210> 750

<400> 750

000

<210> 751

<400> 751

000

<210> 752

<400> 752

000

<210> 753

<400> 753

000

<210> 754

<400> 754

000

<210> 755

<400> 755

000

<210> 756

<400> 756

000

<210> 757

<400> 757

000

<210> 758

<400> 758

000

<210> 759

<400> 759

000

<210> 760

<400> 760

000

<210> 761

<400> 761

000

<210> 762

<400> 762

000

<210> 763

<400> 763

000

<210> 764

<400> 764

000

<210> 765

<400> 765

000

<210> 766

<400> 766

000

<210> 767

<400> 767

000

<210> 768

<400> 768

000

<210> 769

<400> 769

000

<210> 770

<400> 770

000

<210> 771

<400> 771

000

<210> 772

<400> 772

000

<210> 773

<400> 773

000

<210> 774

<400> 774

000

<210> 775

<400> 775

000

<210> 776

<400> 776

000

<210> 777

<400> 777

000

<210> 778

<400> 778

000

<210> 779

<400> 779

000

<210> 780

<400> 780

000

<210> 781

<400> 781

000

<210> 782

<400> 782

000

<210> 783

<400> 783

000

<210> 784

<400> 784

000

<210> 785

<400> 785

000

<210

> 786

<400> 786

000

<210> 787

<400> 787

000

<210> 788

<400> 788

000

<210> 789

<400> 789

000

<210> 790

<400> 790

000

<210> 791

<400> 791

000

<210> 792

<400> 792

000

<210> 793

<400> 793

000

<210> 794

<400> 794

000

<210> 795

<400> 795

000

<210> 796

<400> 796

000

<210> 797

<400> 797

000

<210> 798

<400> 798

000

<210> 799

<400> 799

000

<210> 800

<400> 800

000

<210> 801

<211> 156

<212> DNA

<213> Artificial Sequence

<220

><223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 801

gcggcggggg ggcggccgcg ttcgcgcgcc gcccaccagg gggtgctgcg cgccccccgc 60

cgcgcatgcg cggggccccc ccccgggggg gctccgcccc cccggccccc ccccgtgcta 120

aacccaccgc gcatgcgcga ccacgccccc gccgcc 156

<210> 802

<211> 150

<212> DNA

<213> Artificial Sequence

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

<400> 802

ccgagcgtta gcgaggagtg cgaccctacc ccctgggccc acttcttcgg agccgcgcgc 60

tacgccttcg gctgcgcgcg gcacctcaga cccccgctcg tgctgacacg cttgcgcgtg 120

tcagaccact tcgggctcgc gggggtcggg 150

<210> 803

<211> 122

<212> DNA

<213> Artificial Sequence

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

<400> 803

gccgccgcgg cggcgggggg cggcgcgctg cgcgcgccgc ccagtagggg gagccatgcg 60

ccccccccg cgcatgcgcg gggcccccc ccgcgggggg ctccgccccc cggcccccc 120

cg 122

<210> 804

<211> 111

<212> DNA

<213> Artificial Sequence

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

<400> 804

cggcccagcg gcggcgcgcg cgcttcgcgc gcgcgccggg gggctccgcc ccccccgcg 60

catgcgcggg gccccccccc gcgggggct ccgcccccg gtcccccccc g 111

<210> 805

<211> 115

<212> DNA

<213> Artificial Sequence

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

<400> 805

cggccgtgcg gcggcgcgcg cgcttcgcgc gcgcgccggg ggctgccgcc ccccccgcg 60

catgcgcgcg gggcccccc ccgcgggggg ctccgccccc cggcccccc ccccg 115

<210> 806

<211> 104

<212> DNA

<213> Artificial Sequence

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

polynucleotide

<400> 806

cggcggcggc gcgcgcgcta cgcgcgcgcg ccgggggggct gccgcccccc ccccgcgcat 60

gcgcggggcc cccccccgcg gggggctccg cccccggcc cccc 104

<210> 807

<211> 108

<212> DNA

<213> Artificial Sequence

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

polynucleotide

<400> 807

ggcggcggcg cgcgcgctac gcgcgcgcgc cggggagctc tgcccccccc cgcgcatgcg 60

cgcgggtccc cccccccgcgg ggggctccgc cccccggtcc cccccccg 108

<210> 808

<400> 808

000

<210> 809

<400> 809

000

<210> 810

<400> 810

000

<210> 811

<400> 811

000

<210> 812

<400> 812

000

<210> 813

<400> 813

000

<210> 814

<400> 814

000

<210> 815

<400> 815

000

<210> 816

<400> 816

000

<210> 817

<400> 817

000

<210> 818

<400> 818

000

<210> 819

<400>

819

000

<210> 820

<400> 820

000

<210> 821

<400> 821

000

<210> 822

<400> 822

000

<210> 823

<400> 823

000

<210> 824

<400> 824

000

<210> 825

<400> 825

000

<210> 826

<400> 826

000

<210> 827

<400> 827

000

<210> 828

<400> 828

000

<210> 829

<211> 11

<212> PRT

<213> Artificial Sequence

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

peptide

<220><221> MOD_RES

<222> (4)..(5)

<223> Any amino acid

<220><221> MOD_RES

<222> (7)..(7)

<223> Any amino acid

<220><221> MOD_RES

<222> (9)..(10)

<223> Any amino acid

<400> 829

Tyr Asn Pro Xaa Xaa Asp Xaa Gly Xaa Xaa Asn

1 5 10

<210> 830

<400> 830

000

<210> 831

<400> 831

000

<210> 832

<400> 832

000

<210> 833

<400> 833

000

<210> 834

<400> 834

000

<210> 835

<400> 835

000

<210> 836

<400> 836

000

<210> 837

<400> 837

000

<210> 838

<400> 838

000

<210> 839

<400> 839

000

<210> 840

<400> 840

000

<210> 841

<400> 841

000

<210> 842

<400> 842

000

<210> 843

<400> 843

000

<210> 844

<400> 844

000

<210> 845

<400> 845

000

<210> 846

<400> 846

000

<210> 847

<400> 847

000

<210> 848

<400> 848

000

<210> 849

<400> 849

000

<210> 850

<400> 850

000

<210> 851

<400> 851

000

<210> 852

<400> 852

000

<210> 853

<400> 853

000

<210> 854

<400> 854

000

<210> 855

<400> 855

000

<210> 856

<400> 856

000

<210> 857

<400> 857

000

<210> 858

<400> 858

000

<210> 859

<400> 859

000

<210> 860

<400> 860

000

<210> 861

<400> 861

000

<210> 862

<400> 862

000

<210> 863

<400> 863

000

<210> 864

<400> 864

000

<210> 865

<400> 865

000

<210> 866

<400> 866

000

<210> 867

<400> 867

000

<210> 868

<400> 868

000

<210> 869

<400> 869

000

<210> 870

<400> 870

000

<210> 871

<400> 871

000

<210> 872

<400> 872

000

<210> 873

<400> 873

000

<210> 874

<400> 874

000

<210

> 875

<400> 875

000

<210> 876

<400> 876

000

<210> 877

<400> 877

000

<210> 878

<400> 878

000

<210> 879

<400> 879

000

<210> 880

<400> 880

000

<210> 881

<400> 881

000

<210> 882

<400> 882

000

<210> 883

<400> 883

000

<210> 884

<400> 884

000

<210> 885

<400> 885

000

<210> 886

<211> 3176

<212> DNA

<213> Gammatorquevirus sp.

<400> 886

taaaatggcg ggagccaatc attttatact ttcactttcc aattaaaaat ggccacgtca 60

caaacaaggg gtggagccat ttaaactata taactaagtg gggtggcgaa tggctgagtt 120

taccccgcta gacggtgcag ggaccggatc gagcgcagcg aggaggtccc cggctgccca 180

tgggcgggag ccgaggtgag tgaaaccacc gaggtctagg ggcaattcgg gctagggcag 240

tctagcggaa cgggcaagaa acttaaaaca atatttgttt tacagatggt tagtatatcc 300

tcaagtgatt tttttaagaa aacgaaattt aatgaggaga cgcagaacca agtatggatg 360

tctcaaattg ctgactctca tgataatatc tgcagttgct ggcatccatt tgctcacctt 420

cttgcttcca tatttcctcc tggccacaaa gatcgtgatc ttactattaa ccaaattctt 480

ctaagagatt ataaagaaaa atgccattct ggtggagaag aaggagaaaa ttctggacca 540

acaacaggtt taattacacc aaaagaagaa gatatagaaa aagatggccc agaaggcgcc 600

gcagaagaag accatacaga cgccctgttc gccgccgccg tagaaaactt cgaaaggtaa 660

agagaaaaaa aaaatcttta attgttagac aatggcaacc agacagtata agaacttgta 720

aaattatagg acagtcagct atagttgttg gggctgaagg aaagcaaatg tactgttata 780

ctgtcaataa gttaattaat gtgcccccaa aaacaccata tgggggaggc tttggagtag 840

accaatacac actgaaatac ttatatgaag aatacagatt tgcacaaaac atttggacac 900

aatctaatgt actgaaagac ttatgcagat acataaatgt taagctaata ttctacagag 960

acaacaaaac agactttgtc ctttcctatg acagaaaccc acctttcaa ctaacaaaat 1020

ttacataccc aggagcacac ccacaacaaa tcatgcttca aaaacaccac aaattcatac 1080

tatcacaaat gacaaagcct aatggaagac taacaaaaaa actcaaaatt aaacctccta 1140

aacaaatgct ttctaaatgg ttcttttcaa aacaattctg taaataccct ttactatctc 1200

ttaaagcttc tgcactagac cttaggcact cttacctagg ctgctgtaat gaaaatccac 1260

aggtattttt ttattattta aaccatggat actacacaat aacaaactgg ggagcacaat 1320

cctcaacagc atacagacct aactccaagg tgacagacac aacatactac agatacaaaa 1380

atgacagaaa aaatattaac attaaaagcc atgaatacga aaaaagtata tcatatgaaa 1440

acggttattt tcaatctagt ttcttacaaa cacagtgcat atataccagt gagcgtggtg 1500

aagcctgtat agcagaaaaa ccactaggaa tagctattta caatccagta aaagacaatg 1560

gagatggtaa tatgatatac cttgtaagca ctctagcaaa cacttgggac cagcctccaa 1620

aagacagtgc tattttaata caaggagtac ccatatggct aggcttattt ggatatttag 1680

actactgtag acaaattaaa gctgacaaaa catggctaga cagtcatgta ctagtaattc 1740

aaagtcctgc tatttttact tacccaaatc caggagcagg caaatggtat tgtccactat 1800

cacaaagttt tataaatggc aatggtccgt ttaatcaacc acctacactg ctacaaaaag 1860

caaagtggtt tccacaaata caataccaac aagaaattat taatagcttt gtagaatcag 1920

gaccatttgt tcccaaatat gcaaatcaaa ctgaaagcaa ctgggaacta aaatataaat 1980

atgtttttac atttaagtgg ggtggaccac aattccatga accagaaatt gctgacccta 2040

gcaaacaaga gcagtatgat gtccccgata ctttctacca aacaatacaa attgaagatc 2100

cagaaggaca agaccccaga tctctcatcc atgattggga ctacagacga ggctttatta 2160

aagaaagatc tcttaaaaga atgtcaactt acttctcaac tcatacagat cagcaagcaa 2220

cttcagagga agacattccc aaaaagaaaa agagaattgg accccaactc acagtcccac 2280
aacaaaaaga agaggagaca ctgtcatgtc tcctctctct ctgcaaaaaa gataccttcc 2340
aagaaacaga gacacaagaa gacctccagc agctcatcaa gcagcagcag gagcagcagc 2400
tcctcctcaa gagaaacatc ctccagctca tccacaaact aaaagagaat caacaaatgc 2460
ttcagcttca cacaggcatg ttaccttaac cagatttaaa cctggatttg aagagcaaac 2520
agagagagaa ttagcaatta tatttcatag gccccctaga acctacaaag aggaccttcc 2580
attctatccc tggctaccac ctgcacccct tgtacaattt aaccttaact tcaaaggcta 2640

ggccaacaat gtacacttag taaagcatgt ttattaaagc acaaccccca aaataaatgt 2700
aaaaataaaa aaaaaaaaaa aaaaataaaa aattgcaaaa attcggcgct cgcgcgcatg 2760
tgcgcctctg gcgcaaatca cgcaacgctc gcgcgcccgc gtatgtctct ttaccacgca 2820
cctagattgg ggtgcgcgcg ctagcgcgcg caccccaatg cgccccgccc tcgttccgac 2880
ccgcttgcgc gggtcggacc acttcgggct cgggggggcg cgcctgcggc gctttttac 2940
taaacagact ccgagccgcc atttggcccc ctaagctccg ccccccctcat gaatattcat 3000
aaaggaaacc acataattag aattgccgac cacaaactgc catatgctaa ttagttcccc 3060

ttttacaaag taaaagggga agtgaacata gccccacacc cgcaggggca aggccccgca 3120
cccctacgtc actaaccacg cccccgccgc catcttgggt gcggcagggc gggggc 3176
<210> 887
<211> 124
<212> PRT
<213> Gammatorquevirus sp.
<400> 887
Met Val Ser Ile Ser Ser Ser Asp Phe Phe Lys Lys Thr Lys Phe Asn
1 5 10 15
Glu Glu Thr Gln Asn Gln Val Trp Met Ser Gln Ile Ala Asp Ser His
20 25 30
Asp Asn Ile Cys Ser Cys Trp His Pro Phe Ala His Leu Leu Ala Ser

35 40 45
Ile Phe Pro Pro Gly His Lys Asp Arg Asp Leu Thr Ile Asn Gln Ile
50 55 60
Leu Leu Arg Asp Tyr Lys Glu Lys Cys His Ser Gly Gly Glu Glu Gly

65 70 75 80

Glu Asn Ser Gly Pro Thr Thr Gly Leu Ile Thr Pro Lys Glu Glu Asp

85 90 95

Ile Glu Lys Asp Gly Pro Glu Gly Ala Ala Glu Glu Asp His Thr Asp

100 105 110

Ala Leu Phe Ala Ala Ala Val Glu Asn Phe Glu Arg

115 120

<210> 888

<211> 271

<212> PRT

<213> Gammatorquevirus sp.

<400> 888

Met Val Ser Ile Ser Ser Ser Asp Phe Phe Lys Lys Thr Lys Phe Asn

1 5 10 15

Glu Glu Thr Gln Asn Gln Val Trp Met Ser Gln Ile Ala Asp Ser His

20 25 30

Asp Asn Ile Cys Ser Cys Trp His Pro Phe Ala His Leu Leu Ala Ser

35 40 45

Ile Phe Pro Pro Gly His Lys Asp Arg Asp Leu Thr Ile Asn Gln Ile

50 55 60

Leu Leu Arg Asp Tyr Lys Glu Lys Cys His Ser Gly Gly Glu Glu Gly

65 70 75 80

Glu Asn Ser Gly Pro Thr Thr Gly Leu Ile Thr Pro Lys Glu Glu Asp

85 90 95

Ile Glu Lys Asp Gly Pro Glu Gly Ala Ala Glu Glu Asp His Thr Asp

100 105 110

Ala Leu Phe Ala Ala Ala Val Glu Asn Phe Glu Ser Gly Val Asp His

115 120 125

Asn Ser Met Asn Gln Lys Leu Leu Thr Leu Ala Asn Lys Ser Ser Met

130 135 140

Met Ser Pro Ile Leu Ser Thr Lys Gln Tyr Lys Leu Lys Ile Gln Lys

145 150 155 160

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

165 170 175

Leu Leu Lys Lys Asp Leu Leu Lys Glu Cys Gln Leu Thr Ser Gln Leu

180 185 190

Ile Gln Ile Ser Lys Gln Leu Gln Arg Lys Thr Phe Pro Lys Arg Lys

195 200 205

Arg Glu Leu Asp Pro Asn Ser Gln Ser His Asn Lys Lys Lys Arg Arg

210 215 220

His Cys His Val Ser Ser Leu Ser Ala Lys Lys Ile Pro Ser Lys Lys

225 230 235 240

Gln Arg His Lys Lys Thr Ser Ser Ser Ser Ser Ser Ser Ser Arg Ser

245 250 255

Ser Ser Ser Ser Ser Arg Glu Thr Ser Ser Ser Ser Ser Thr Asn

260 265 270

<210> 889

<211> 267

<212> PRT

<213> Gammatorquevirus sp.

<400> 889

Met Val Ser Ile Ser Ser Ser Asp Phe Phe Lys Lys Thr Lys Phe Asn

1 5 10 15

Glu Glu Thr Gln Asn Gln Val Trp Met Ser Gln Ile Ala Asp Ser His

20 25 30

Asp Asn Ile Cys Ser Cys Trp His Pro Phe Ala His Leu Leu Ala Ser

35 40 45

Ile Phe Pro Pro Gly His Lys Asp Arg Asp Leu Thr Ile Asn Gln Ile

50 55 60

Leu Leu Arg Asp Tyr Lys Glu Lys Cys His Ser Gly Gly Glu Glu Gly

65 70 75 80

Glu Asn Ser Gly Pro Thr Thr Gly Leu Ile Thr Pro Lys Glu Glu Asp

85 90 95

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

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

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

<210> 890

<211> 50

<212> PRT

<213> Gammatorquevirus sp.

<400> 890

Met Arg Arg Arg Arg Thr Lys Tyr Gly Cys Leu Lys Leu Leu Thr Leu
1 5 10 15
Met Ile Ile Ser Ala Val Ala Gly Ile His Leu Leu Thr Phe Leu Leu
20 25 30
Pro Tyr Phe Leu Leu Ala Thr Lys Ile Val Ile Leu Leu Leu Thr Lys

35 40 45

Phe Phe

50

<210> 891

<211> 662

<212> PRT

<213

> Gammatorquevirus sp.

<400> 891

Met Pro Phe Trp Trp Arg Arg Arg Arg Lys Phe Trp Thr Asn Asn Arg

1 5 10 15

Phe Asn Tyr Thr Lys Arg Arg Arg Tyr Arg Lys Arg Trp Pro Arg Arg

20 25 30

Arg Arg Arg Arg Arg Pro Tyr Arg Arg Pro Val Arg Arg Arg Arg Arg

35 40 45

Lys Leu Arg Lys Val Lys Arg Lys Lys Lys Ser Leu Ile Val Arg Gln

50 55 60

Trp Gln Pro Asp Ser Ile Arg Thr Cys Lys Ile Ile Gly Gln Ser Ala

65 70 75 80

Ile Val Val Gly Ala Glu Gly Lys Gln Met Tyr Cys Tyr Thr Val Asn

85 90 95

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

100 105 110

Val Asp Gln Tyr Thr Leu Lys Tyr Leu Tyr Glu Glu Tyr Arg Phe Ala

115 120 125

Gln Asn Ile Trp Thr Gln Ser Asn Val Leu Lys Asp Leu Cys Arg Tyr

130 135 140

Ile Asn Val Lys Leu Ile Phe Tyr Arg Asp Asn Lys Thr Asp Phe Val

145 150 155 160

Leu Ser Tyr Asp Arg Asn Pro Pro Phe Gln Leu Thr Lys Phe Thr Tyr

165 170 175

Pro Gly Ala His Pro Gln Gln Ile Met Leu Gln Lys His His Lys Phe

180 185 190

Ile Leu Ser Gln Met Thr Lys Pro Asn Gly Arg Leu Thr Lys Lys Leu
195 200 205

Lys Ile Lys Pro Pro Lys Gln Met Leu Ser Lys Trp Phe Phe Ser Lys
210 215 220

Gln Phe Cys Lys Tyr Pro Leu Leu Ser Leu Lys Ala Ser Ala Leu Asp
225 230 235 240

Leu Arg His Ser Tyr Leu Gly Cys Cys Asn Glu Asn Pro Gln Val Phe
245 250 255

Phe Tyr Tyr Leu Asn His Gly Tyr Tyr Thr Ile Thr Asn Trp Gly Ala
260 265 270

Gln Ser Ser Thr Ala Tyr Arg Pro Asn Ser Lys Val Thr Asp Thr Thr
275 280 285

Tyr Tyr Arg Tyr Lys Asn Asp Arg Lys Asn Ile Asn Ile Lys Ser His
290 295 300

Glu Tyr Glu Lys Ser Ile Ser Tyr Glu Asn Gly Tyr Phe Gln Ser Ser
305 310 315 320

Phe Leu Gln Thr Gln Cys Ile Tyr Thr Ser Glu Arg Gly Glu Ala Cys
325 330 335

Ile Ala Glu Lys Pro Leu Gly Ile Ala Ile Tyr Asn Pro Val Lys Asp
340 345 350

Asn Gly Asp Gly Asn Met Ile Tyr Leu Val Ser Thr Leu Ala Asn Thr
355 360 365

Trp Asp Gln Pro Pro Lys Asp Ser Ala Ile Leu Ile Gln Gly Val Pro
370 375 380

Ile Trp Leu Gly Leu Phe Gly Tyr Leu Asp Tyr Cys Arg Gln Ile Lys
385 390 395 400

Ala Asp Lys Thr Trp Leu Asp Ser His Val Leu Val Ile Gln Ser Pro
405 410 415

Ala Ile Phe Thr Tyr Pro Asn Pro Gly Ala Gly Lys Trp Tyr Cys Pro
420 425 430

Leu Ser Gln Ser Phe Ile Asn Gly Asn Gly Pro Phe Asn Gln Pro Pro
435 440 445

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

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

Glu Asp Ile Pro Lys Lys Lys Lys Arg Ile Gly Pro Gln Leu Thr Val
580 585 590
Pro Gln Gln Lys Glu Glu Glu Thr Leu Ser Cys Leu Leu Ser Leu Cys
595 600 605
Lys Lys Asp Thr Phe Gln Glu Thr Glu Thr Gln Glu Asp Leu Gln Gln
610 615 620
Leu Ile Lys Gln Gln Gln Glu Gln Gln Leu Leu Leu Lys Arg Asn Ile
625 630 635 640

Leu Gln Leu Ile His Lys Leu Lys Glu Asn Gln Gln Met Leu Gln Leu
645 650 655
His Thr Gly Met Leu Pro
660

<210> 892

<211> 215

<212> PRT

<213> Gammatorquevirus sp.

<400> 892

Met Pro Phe Trp Trp Arg Arg Arg Arg Lys Phe Trp Thr Asn Asn Arg
1 5 10 15

Phe Asn Tyr Thr Lys Arg Arg Arg Tyr Arg Lys Arg Trp Pro Arg Arg
20 25 30

Arg Arg Arg Arg Arg Pro Tyr Arg Arg Pro Val Arg Arg Arg Arg Arg
35 40 45

Lys Leu Arg Lys Trp Gly Gly Pro Gln Phe His Glu Pro Glu Ile Ala
50 55 60

Asp Pro Ser Lys Gln Glu Gln Tyr Asp Val Pro Asp Thr Phe Tyr Gln
65 70 75 80

Thr Ile Gln Ile Glu Asp Pro Glu Gly Gln Asp Pro Arg Ser Leu Ile
85 90 95

His Asp Trp Asp Tyr Arg Arg Gly Phe Ile Lys Glu Arg Ser Leu Lys
100 105 110

Arg Met Ser Thr Tyr Phe Ser Thr His Thr Asp Gln Gln Ala Thr Ser
115 120 125

Glu Glu Asp Ile Pro Lys Lys Lys Lys Arg Ile Gly Pro Gln Leu Thr
130 135 140

Val Pro Gln Gln Lys Glu Glu Glu Thr Leu Ser Cys Leu Leu Ser Leu
145 150 155 160

Cys Lys Lys Asp Thr Phe Gln Glu Thr Glu Thr Gln Glu Asp Leu Gln
165 170 175

Gln Leu Ile Lys Gln Gln Gln Glu Gln Gln Leu Leu Leu Lys Arg Asn
180 185 190

Ile Leu Gln Leu Ile His Lys Leu Lys Glu Asn Gln Gln Met Leu Gln
195 200 205

Leu His Thr Gly Met Leu Pro
210 215

<210> 893

<211> 129

<212> PRT

<213> Gammatorquevirus sp.

<400> 893

Met Pro Phe Trp Trp Arg Arg Arg Arg Lys Phe Trp Thr Asn Asn Arg
1 5 10 15

Phe Asn Tyr Thr Lys Arg Arg Arg Tyr Arg Lys Arg Trp Pro Arg Arg
20 25 30

Arg Arg Arg Arg Arg Pro Tyr Arg Arg Pro Val Arg Arg Arg Arg Arg
35 40 45

Lys Leu Arg Lys Ile Ser Lys Gln Leu Gln Arg Lys Thr Phe Pro Lys
50 55 60

Arg Lys Arg Glu Leu Asp Pro Asn Ser Gln Ser His Asn Lys Lys Lys
65 70 75 80

Arg Arg His Cys His Val Ser Ser Leu Ser Ala Lys Lys Ile Pro Ser
85 90 95

Lys Lys Gln Arg His Lys Lys Thr Ser Ser Ser Ser Ser Ser Ser Ser
100 105 110

Arg Ser Ser Ser Ser Ser Ser Arg Glu Thr Ser Ser Ser Ser Ser Thr
115 120 125

Asn

<210> 894

<400> 894

000

<210> 895

<400> 895

000

<210> 896

<400> 896

000

<210> 897

<400> 897

000

<210> 898

<400> 898

000

<210> 899

<400> 899

000

<210> 900

<400> 900

000

<210> 901

<400> 901

000

<210> 902

<400> 902

000

<210> 903

<400> 903

000

<210> 904

<400> 904

000

<210> 905

<400> 905

000

<210> 906

<400> 906

000

<210> 907

<400> 907

000

<210> 908

<400> 908

000

<210> 909

<400> 909

000

<210> 910

<400> 910

000

<210> 911

<400> 911

000

<210> 912

<400> 912

000

<210> 913

<400> 913

000

<210> 914

<400> 914

000

<210> 915

<400> 915

000

<210> 916

<400> 916

000

<210> 917

<400> 917

000

<210> 918

<400> 918

000

<210> 919

<400> 919

000

<210> 920

<400> 920

000

<210> 921

<400> 921

000

<210> 922

<400> 922

000

<210> 923

<400> 923

000

<210> 924

<400> 924

000

<210> 925

<211> 662

<212> PRT

<213> Gammatorquevirus sp.

<400> 925

Met Pro Phe Trp Trp Arg Arg Arg Arg Lys Phe Trp Thr Asn Asn Arg

1 5 10 15

Phe Asn Tyr Thr Lys Arg Arg Arg Tyr Arg Lys Arg Trp Pro Arg Arg

20 25 30

Arg Arg Arg Arg Arg Pro Tyr Arg Arg Pro Val Arg Arg Arg Arg Arg

35 40 45

Lys Leu Arg Lys Val Lys Arg Lys Lys Lys Ser Leu Ile Val Arg Gln

50 55 60

Trp Gln Pro Asp Ser Ile Arg Thr Cys Lys Ile Ile Gly Gln Ser Ala

65 70 75 80

Ile Val Val Gly Ala Glu Gly Lys Gln Met Tyr Cys Tyr Thr Val Asn

85 90 95

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

100 105 110

Val Asp Gln Tyr Thr Leu Lys Tyr Leu Tyr Glu Glu Tyr Arg Phe Ala
115 120 125
Gln Asn Ile Trp Thr Gln Ser Asn Val Leu Lys Asp Leu Cys Arg Tyr

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

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

260 265 270
Gln Ser Ser Thr Ala Tyr Arg Pro Asn Ser Lys Val Thr Asp Thr Thr
275 280 285
Tyr Tyr Arg Tyr Lys Asn Asp Arg Lys Asn Ile Asn Ile Lys Ser His
290 295 300
Glu Tyr Glu Lys Ser Ile Ser Tyr Glu Asn Gly Tyr Phe Gln Ser Ser
305 310 315 320
Phe Leu Gln Thr Gln Cys Ile Tyr Thr Ser Glu Arg Gly Glu Ala Cys

325 330 335
Ile Ala Glu Lys Pro Leu Gly Ile Ala Ile Tyr Asn Pro Val Lys Asp
340 345 350
Asn Gly Asp Gly Asn Met Ile Tyr Leu Val Ser Thr Leu Ala Asn Thr

355 360 365

Trp Asp Gln Pro Pro Lys Asp Ser Ala Ile Leu Ile Gln Gly Val Pro

370 375 380

Ile Trp Leu Gly Leu Phe Gly Tyr Leu Asp Tyr Cys Arg Gln Ile Lys

385 390 395 400

Ala Asp Lys Thr Trp Leu Asp Ser His Val Leu Val Ile Gln Ser Pro

405 410 415

Ala Ile Phe Thr Tyr Pro Asn Pro Gly Ala Gly Lys Trp Tyr Cys Pro

420 425 430

Leu Ser Gln Ser Phe Ile Asn Gly Asn Gly Pro Phe Asn Gln Pro Pro

435 440 445

Thr Leu Leu Gln Lys Ala Lys Trp Phe Pro Gln Ile Gln Tyr Gln Gln

450 455 460

Glu Ile Ile Asn Ser Phe Val Glu Ser Gly Pro Phe Val Pro Lys Tyr

465 470 475 480

Ala Asn Gln Thr Glu Ser Asn Trp Glu Leu Lys Tyr Lys Tyr Val Phe

485 490 495

Thr Phe Lys Trp Gly Gly Pro Gln Phe His Glu Pro Glu Ile Ala Asp

500 505 510

Pro Ser Lys Gln Glu Gln Tyr Asp Val Pro Asp Thr Phe Tyr Gln Thr

515 520 525

Ile Gln Ile Glu Asp Pro Glu Gly Gln Asp Pro Arg Ser Leu Ile His

530 535 540

Asp Trp Asp Tyr Arg Arg Gly Phe Ile Lys Glu Arg Ser Leu Lys Arg

545 550 555 560

Met Ser Thr Tyr Phe Ser Thr His Thr Asp Gln Gln Ala Thr Ser Glu

565 570 575

Glu Asp Ile Pro Lys Lys Lys Lys Arg Ile Gly Pro Gln Leu Thr Val

580 585 590

Pro Gln Gln Lys Glu Glu Glu Thr Leu Ser Cys Leu Leu Ser Leu Cys

595 600 605

Lys Lys Asp Thr Phe Gln Glu Thr Glu Thr Gln Glu Asp Leu Gln Gln

610 615 620

Leu Ile Lys Gln Gln Gln Glu Gln Gln Leu Leu Leu Lys Arg Asn Ile

625 630 635 640

Leu Gln Leu Ile His Lys Leu Lys Glu Asn Gln Gln Met Leu Gln Leu

645 650 655

His Thr Gly Met Leu Pro

660

<210> 926

<211> 58

<212> PRT

<213> Gammatorquevirus sp.

<400> 926

Met Pro Phe Trp Trp Arg Arg Arg Arg Lys Phe Trp Thr Asn Asn Arg

1 5 10 15

Phe Asn Tyr Thr Lys Arg Arg Arg Tyr Arg Lys Arg Trp Pro Arg Arg

20 25 30

Arg Arg Arg Arg Arg Pro Tyr Arg Arg Pro Val Arg Arg Arg Arg Arg

35 40 45

Lys Leu Arg Lys Val Lys Arg Lys Lys Lys

50 55

<210> 927

<211> 202

<212> PRT

<213> Gammatorquevirus sp.

<400> 927

Ser Leu Ile Val Arg Gln Trp Gln Pro Asp Ser Ile Arg Thr Cys Lys

1 5 10 15

Ile Ile Gly Gln Ser Ala Ile Val Val Gly Ala Glu Gly Lys Gln Met

20 25 30

Tyr Cys Tyr Thr Val Asn Lys Leu Ile Asn Val Pro Pro Lys Thr Pro

35 40 45

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

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

165 170 175
Lys Ala Ser Ala Leu Asp Leu Arg His Ser Tyr Leu Gly Cys Cys Asn
180 185 190
Glu Asn Pro Gln Val Phe Phe Tyr Tyr Leu
195 200

<210> 928
<211> 79
<212> PRT
<213> Gammatorquevirus sp.
<400> 928
Asn His Gly Tyr Tyr Thr Ile Thr Asn Trp Gly Ala Gln Ser Ser Thr
1 5 10 15
Ala Tyr Arg Pro Asn Ser Lys Val Thr Asp Thr Thr Tyr Tyr Arg Tyr

20 25 30
Lys Asn Asp Arg Lys Asn Ile Asn Ile Lys Ser His Glu Tyr Glu Lys
35 40 45
Ser Ile Ser Tyr Glu Asn Gly Tyr Phe Gln Ser Ser Phe Leu Gln Thr

50 55 60

Gln Cys Ile Tyr Thr Ser Glu Arg Gly Glu Ala Cys Ile Ala Glu

65 70 75

<210> 929

<211> 160

<212> PRT

<213> Gammatorquevirus sp.

<400> 929

Lys Pro Leu Gly Ile Ala Ile Tyr Asn Pro Val Lys Asp Asn Gly Asp

1 5 10 15

Gly Asn Met Ile Tyr Leu Val Ser Thr Leu Ala Asn Thr Trp Asp Gln

20 25 30

Pro Pro Lys Asp Ser Ala Ile Leu Ile Gln Gly Val Pro Ile Trp Leu

35 40 45

Gly Leu Phe Gly Tyr Leu Asp Tyr Cys Arg Gln Ile Lys Ala Asp Lys

50 55 60

Thr Trp Leu Asp Ser His Val Leu Val Ile Gln Ser Pro Ala Ile Phe

65 70 75 80

Thr Tyr Pro Asn Pro Gly Ala Gly Lys Trp Tyr Cys Pro Leu Ser Gln

85 90 95

Ser Phe Ile Asn Gly Asn Gly Pro Phe Asn Gln Pro Pro Thr Leu Leu

100 105 110

Gln Lys Ala Lys Trp Phe Pro Gln Ile Gln Tyr Gln Gln Glu Ile Ile

115 120 125

Asn Ser Phe Val Glu Ser Gly Pro Phe Val Pro Lys Tyr Ala Asn Gln

130 135 140

Thr Glu Ser Asn Trp Glu Leu Lys Tyr Lys Tyr Val Phe Thr Phe Lys

145 150 155 160

<210> 930

<211> 163

<212> PRT

<213> Gammatorquevirus sp.

<400> 930

Trp Gly Gly Pro Gln Phe His Glu Pro Glu Ile Ala Asp Pro Ser Lys
1 5 10 15
Gln Glu Gln Tyr Asp Val Pro Asp Thr Phe Tyr Gln Thr Ile Gln Ile

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

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

145 150 155 160
Met Leu Pro

<210> 931

<400> 931

000

<210> 932

<400> 932

000

<210> 933

<400> 933

000

<210> 934

<400> 934

000

<210> 935

<400> 935

000

<210> 936

<400> 936

000

<210> 937

<400> 937

000

<210> 938

<400> 938

000

<210> 939

<400> 939

000

<210> 940

<400> 940

000

<210> 941

<400> 941

000

<210> 942

<400> 942

000

<210> 943

<400> 943

000

<210> 944

<400> 944

000

<210> 945

<400> 945

000

<210> 946

<400> 946

000

<210> 947

<400> 947

000

<210> 948

<400> 948

000

<210> 949

<211> 21

<212> PRT

<213> Artificial Sequence

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

peptide

<220><221> MOD_RES

<222> (1)..(1)

<223> W or F

<220><221> MOD_RES

<222> (2)..(8)

<223> Any amino acid

<220><221> MOD_RES

<222> (10)..(12)

<223> Any amino acid

<220><221> MOD_RES

<222> (14)..(14)

<223> Any amino acid

<220><221> MOD_RES

<222> (16)..(20)

<223> Any amino acid

<400> 949

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa His Xaa Xaa Xaa Cys Xaa Cys Xaa

1 5 10 15

Xaa Xaa Xaa Xaa His

20

<210> 950

<211> 51

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<220><221> modified_base

<222> (45)..(45)

<223> a, c, t, g, unknown or other

<220>

<223> See specification as filed for detailed description of

substitutions and preferred embodiments

<400> 950

aggtgagtga aaccaccgaa gtcaaggggc aattcgggct agggncagtc t 51

<210> 951

<211> 50

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 951

aggtgagttt acacaccgca gtcaaggggc aattcgggct cgggactggc 50

<210> 952

<211> 50

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 952

aggtgagtga aaccaccgaa gtcaaggggc aattcgggct agatcagtct 50

<210> 953

<211> 50

<212> DNA

<213> Artificial Sequence

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

<400> 953

aggtgagtga aaccaccgag gtctaggggc aattcgggct agggcagtct 50