



**(19) 대한민국특허청(KR)**  
**(12) 공개특허공보(A)**

(11) 공개번호 10-2022-0093163  
(43) 공개일자 2022년07월05일

- (51) 국제특허분류(Int. Cl.)  
C07K 14/005 (2006.01) A61K 48/00 (2006.01)  
A61P 27/02 (2006.01) C12N 15/86 (2006.01)
- (52) CPC특허분류  
C07K 14/005 (2013.01)  
A61K 48/0075 (2013.01)
- (21) 출원번호 10-2022-7018274
- (22) 출원일자(국제) 2020년11월02일  
심사청구일자 없음
- (85) 번역문제출일자 2022년05월30일
- (86) 국제출원번호 PCT/EP2020/080704
- (87) 국제공개번호 WO 2021/084133  
국제공개일자 2021년05월06일
- (30) 우선권주장  
19206603.3 2019년10월31일  
유럽특허청(EPO)(EP)

- (71) 출원인  
유니버시태트 베른  
스위스 CH-3012 베른 호호슐슈트라쎄 6 퍼발통스 디레크티온  
메디치니쎄 호호슐레 하노버  
독일 30625 하노버 칼-노이베르크-슈트라쎄 1
- (72) 발명자  
클라인로젤, 존야  
스위스 3012 베른 호호슐슈트라쎄 6 유니버시태트 베른 내  
데이비드, 아난드  
인도 400 069 마하라시트라 뭄바이 안드헤리 (이스트) 군다발리 아자드 로드 니어 비엠편 오피스 비앙카 빌딩 룸 501  
뵤닝, 힐데가르트  
독일 30625 하노버 메디치니쎄 호호슐레 하노버 내
- (74) 대리인  
양영준, 이윤기

전체 청구항 수 : 총 17 항

(54) 발명의 명칭 **안구 유전자 전달을 위한 AAV 벡터 변이체**

**(57) 요약**

본 발명은 아데노-연관 바이러스 캡시드 폴리펩타이드 서열 그리고 눈 및 잠재적으로 다른 조직에 대한 치료용 트랜스유전자 전달에서의 이의 용도에 관한 것이다.

(52) CPC특허분류

**A61P 27/02** (2018.01)

**C12N 15/86** (2013.01)

*C12N 2750/14122* (2013.01)

*C12N 2750/14143* (2013.01)

*C12N 2810/40* (2013.01)

---

## 명세서

### 청구범위

#### 청구항 1

아데노-연관 바이러스(AAV) 혈청형 2 캡시드의 위치 453에 또는 위치 587 내지 592에, 구체적으로 위치 587 내지 592에, 보다 구체적으로 위치 587에 또는 또 다른 혈청형의 AAV에서 이와 동종인 위치에 펩타이드 삽입을 포함하는 AAV 캡시드 폴리펩타이드로서, 펩타이드 삽입이

SASEAST(Cap3; SEQ ID NO 10), DTRPHDQ(Cap5; SEQ ID NO 11), EHYNSTC(Cap7; SEQ ID NO 12), PNPNTL(Cap9; SEQ ID NO 13), TPPSITA(Cap11; SEQ ID NO 14), CGESSYL(Cap12; SEQ ID NO 15), PRTPHTA(Cap13; SEQ ID NO 16) 및 ELCDFGA(Cap14; SEQ ID NO 17)로부터 선택되며,

특히 삽입이 SASEAST(Cap3; SEQ ID NO 10), TPPSITA(Cap11; SEQ ID NO 14), PRTPHTA(Cap13; SEQ ID NO 16) 및 ELCDFGA(Cap14; SEQ ID NO 17)로부터 선택되는, AAV 캡시드 폴리펩타이드.

#### 청구항 2

SEQ ID NO 001과 적어도( $\geq$ ) 85% 동일성, 구체적으로  $\geq 90\%$ , 보다 구체적으로  $\geq 95\%$  동일성을 갖는 서열을 포함하거나 이로 본질적으로 구성되는, 아데노-연관 바이러스 캡시드 폴리펩타이드로서,

폴리펩타이드가 위치 587, 588, 589, 590, 591 또는 592에서, 구체적으로 위치 587, 588 또는 589에서, 보다 구체적으로 위치 587에서 삽입을 특징으로 하며,

삽입이 SEQ ID NO 10, 11, 12, 13, 14, 15, 16, 17 중 어느 하나로부터 선택되는 펩타이드 서열이거나 이를 포함하는, 아데노-연관 바이러스 캡시드 폴리펩타이드.

#### 청구항 3

제2항에 있어서, 폴리펩타이드가 SEQ ID NO 2 내지 SEQ ID NO 9로부터 선택되는 서열의 생물학적 활성의 적어도 90%를 특징으로 하는 아데노-연관 바이러스 캡시드 폴리펩타이드.

#### 청구항 4

제2항 또는 제3항에 있어서, 폴리펩타이드가 SEQ ID NO 2 내지 SEQ ID NO 9로부터 선택되는 서열이거나 이를 포함하는 아데노-연관 바이러스 캡시드 폴리펩타이드.

#### 청구항 5

제1항 내지 제4항 중 어느 한 항에 있어서, SEQ ID NO 2 내지 SEQ ID NO 9로부터 선택되는 아미노산 서열을 포함하거나, 이로 본질적으로 구성되는 아데노-연관 바이러스 캡시드 폴리펩타이드.

#### 청구항 6

제1항 또는 제2항에 있어서, AAV 캡시드 단백질이

a. 위치 252, 272, 444, 500, 700, 704 및 730에서, 구체적으로 252, 272, 444, 500, 704 및 730에서, 보다 구체적으로 위치 252, 272, 444, 500, 700 및 730 모두에서 하나 또는 몇몇의 티로신에서 페닐알라닌으로의 치환, 및/또는

b. 하나 또는 몇몇의 트레오닌에서 발린으로의 치환, 구체적으로 T491V를 특징으로 하는 AAV2 캡시드이거나, 캡시드가 AAV2 캡시드 이외의 캡시드이고 캡시드가 캡시드의 동종 위치에서 상기 a. 및 b.에 언급된 바와 같은 하나 또는 몇몇의 티로신에서 페닐알라닌으로의 치환 및/또는 하나 또는 몇몇의 트레오닌에서 발린으로의 치환을 특징으로 하는, 아데노-연관 바이러스 캡시드 폴리펩타이드.

#### 청구항 7

제1항 내지 제6항 중 어느 한 항의 AAV 캡시드 폴리펩타이드를 인코딩하는 핵산 서열.

**청구항 8**

제7항에 있어서, 서열이 자가-상보성 또는 단일 가닥 벡터 게놈 입체형태, 특히 자가-상보성 벡터 게놈인, 핵산 서열.

**청구항 9**

제7항 또는 제8항에 있어서, 트랜스유전자를 포함하는 핵산 서열.

**청구항 10**

제9항에 있어서, 트랜스유전자가 정확한 단백질, siRNA, shRNA 또는 CRISPR/Cas-gRNA 카세트의 서열을 인코딩하는, 핵산 서열.

**청구항 11**

제9항에 있어서, 트랜스유전자가 광-감수성 단백질을 인코딩하는, 핵산 서열.

**청구항 12**

제9항 내지 제11항 중 어느 한 항에 있어서, 트랜스유전자가 포유류 세포, 구체적으로, 망막 세포, 보다 구체적으로 인간 망막 세포에서 작동 가능한 프로모터 서열의 제어 하에 있는, 핵산 서열.

**청구항 13**

제12항에 있어서, 프로모터가 도처의 또는 세포-특이적 프로모터인, 핵산 서열.

**청구항 14**

제12항에 있어서, 프로모터가 CMV 최조기(immediate early) 프로모터 또는 hEfla 프로모터로부터 선택되는, 핵산 서열.

**청구항 15**

의약에서/약제로서 사용하기 위한, 제1항 내지 제6항 중 어느 한 항의 AAV 캡시드 폴리펩타이드 및 제7항 내지 제14항 중 어느 한 항의 핵산 서열로부터 선택되는 제제.

**청구항 16**

a. 망막 또는 망막 색소 상피 세포 및/또는  
 b. 광수용체, 두극 세포, 신경절 세포 또는 무축삭 세포  
 에 영향을 미치는 질병의 치료에서 사용하기 위한, 제1항 내지 제6항 중 어느 한 항의 AAV 캡시드 폴리펩타이드 및 제7항 내지 제14항 중 어느 한 항의 핵산 서열로부터 선택되는 제제.

**청구항 17**

제15항 또는 제16항에 명시된 바와 같은 임의의 용도를 위해, 제1항 내지 제6항 중 어느 한 항의 AAV 캡시드 폴리펩타이드 및 제7항 내지 제14항 중 어느 한 항의 핵산 서열로부터 선택되는 제제로서,  
 a. 유리체내 투여, 구체적으로 유리체내 주사, 또는  
 b. 망막하 투여  
 에 의해 투여되는, 제제.

**발명의 설명**

**기술 분야**

본 발명은 아데노-연관 바이러스 캡시드 폴리펩타이드 서열 그리고 광수용체 또는 망막 색소 상피 세포를 표적

[0001]

화하는 눈에 대한 치료용 트랜스유전자 전달에서의 이의 용도에 관한 것이다.

**배경 기술**

- [0002] 광-감지 망막 광수용체 세포(PR)의 영구 변성은 실명을 유발한다. 광수용체 사멸은 주로 PR 또는 망막 색소 상피(RPE) 세포에서, 예컨대 색소 망막염(RP)에서, 또는 광수용체의 건강에 영향을 미치는 다인자 질병에 의해, 예컨대 연령-관련 황반 변성(AMD) 등에서 국제화된 유전된 돌연변이에 의해 유도될 수 있다. 여러 형태의 광수용체 변성에 관여하는 유전적 돌연변이가 식별되어, 치료용 트랜스유전자의 투여에 의한 개입을 가능하게 한다. 핵산-기반 치료법을 표적화하는 기존 PR은 PR에 대한 손상이 최소인 경우 초기 질환 단계에만 효과적이다. 기능적 내부 망막 세포에 광 감수성을 부여하는 적합한 광유전자 단백질을 채택하는 유전자 치료법은 후기 질환 단계에서도 잠재적으로 시력을 복구할 수 있다. 이러한 접근법을 임상 환경에서 적용 가능하게 만들기 위해, 유전자 전달을 위해 더 우수한 망막 침투를 갖는 개선된 벡터가 요구된다.
- [0003] 아데노-연관 바이러스(AAV)에 기반한 재조합 벡터는 눈에서의 치료용 유전자 전달을 위한 후보이다. 최근, AAV 벡터는 전형적으로 망막하 적용된다. 그러나 망막하 적용은 망막 두께 및 시력의 감소를 야기하는 것으로 나타났다. 또한, 망막하 적용은 주사액 용적과 바로 인접한 세포에서만 유전자 전달 및 발현을 달성하는 반면, 유효 투여 방법은 망막의 전체 폭을 따라 세포에 도달해야 한다. 눈의 젤리-유사 충전물(유리체 방수) 내로의 유리체 내 주사는 전체 망막으로 벡터를 전달할 잠재력을 가질뿐만 아니라 또한 망막하 주사에 비해 훨씬 더 안전하고 기술적으로 훨씬 덜 까다롭다. 그러나, 신경 망막을 유리체 방수와 분리하는 내부 경계막(ILM)은 여러 AAV 혈청형의 천연 수용체에 풍부하며 이에 따라 AAV 벡터에 대한 엄청난 장벽을 제시한다. 유리체내 전달되는 경우 ILM 및 외부 망막 장벽을 통해 더 잘 침투하는 조작된 AAV 벡터가 안구 AAV 유전자 치료법의 효능 및 안전성을 증가시키기 위해 요구된다.
- [0004] 상기 언급된 최신 기술에 기반하여, 본 발명의 목적은 모든 망막 세포 유형, 특히 내부 망막 세포 및 광수용체의 표적화를 가능하게 하는 망막 내로의 유전자 전달을 위한 수단 및 방법을 제공하는 것이다. 이러한 목적은 본 명세서의 독립항의 요지-사안에 의해 획득된다.

**발명의 내용**

- [0005] 본 발명의 제1 양태는 각각 최고 및 두 번째 최고 캡시드 돌출부에 위치하는, 아데노-연관 바이러스(AAV) 혈청형 2 캡시드의 위치 453 또는 587/588에 또는 또 다른 혈청형의 AAV 캡시드에서 이와 동종인 위치(표 3: #1, #2)에 펩타이드 삽입을 포함하는 AAV 캡시드 폴리펩타이드에 관한 것이다. 펩타이드 삽입은 하기 서열로부터 선택된다:
- [0006] SASEAST(Cap3; SEQ ID NO 10), DTRPHDQ(Cap5; SEQ ID NO 11), EHYNSTC(Cap7; SEQ ID NO 12), PNPNTL(Cap9; SEQ ID NO 13), TPPSITA(Cap11; SEQ ID NO 14), CGESSYL(Cap12; SEQ ID NO 15), PRTPHTA(Cap13; SEQ ID NO 16) 및 ELCDGFA(Cap14; SEQ ID NO 17).
- [0007] 펩타이드에는 다른 아미노산의 짧은 신장물(1개, 2개, 3개, 4개, 5개 또는 심지어 6개 AA)이 측면에 있을 수 있다. 캡시드 서열 내의 나타난 위치에서 나타난 서열을 포매하기 위한 측면 아미노산의 구체예는 Ala, Leu, Gly, Ser, 및 Thr 으로부터 선택될 수 있다(그러나 이에 제한되지 않음).
- [0008] 이러한 펩타이드 삽입은 아데노-연관 바이러스의 감염 효능 및/또는 적어도 시험된 세포 유형 및 조직에 있어서 상기 나타난 위치 I-587에서 캡시드 표면 상에 펩타이드 삽입을 나타내는 아데노-연관 벡터의 형질도입 효능을 증가시킨다.
- [0009] 이는 그 세포 표면 노출로 인해 I-588 또는 I-453에서 잘 작용할 수 있다. 이론에 구애받지 않더라도, 본 발명자들은 적어도 일부 효과가 생체내 헤파란 설페이트 프로테오글리칸에 대한 더 낮은 결합에 기인하는 것으로 생각한다.
- [0010] 본 발명의 제2 양태는 제1 양태에 따른 AAV 캡시드 폴리펩타이드를 인코딩하는 핵산 서열에 관한 것이다. 특정 구현예에는 캡시드-조작 AAV 벡터의 생성을 위한 AAV cap 서열, 특히 AAV 혈청형 2 캡시드 서열에서 이러한 핵산 서열의 도입이 포함된다.
- [0011] 본 발명의 제3 양태는 망막 또는 RPE 세포에 영향을 미치는 질병의 치료를 위한 AAV 캡시드 폴리펩타이드, AAV 벡터, 및 핵산 서열로부터 선택되는 제제에 관한 것이다. 이러한 양태의 대안적 형태는 이를 필요로 하는 환자에 대한 본 발명에 따른 제제의 투여 단계를 포함하는, 광수용체 또는 RPE 세포에 영향을 미치는 질병의 치료

방법에 의해 구현된다.

[0012] 본 발명의 제제를 포함하는 투여 형태는 본 발명의 추가 양태이다.

**도면의 간단한 설명**

[0013] 도 1 선택된 캡시드 변이체에 대한 발현의 망막층 특이성.

도 2 유리체내 주사 후 외부 젓과 망막 전반의 범-망막 트랜스유전자 발현의 예.

도 3 마우스 눈 내로의 유리체내 주사 후 AAV2(M6) 및 AAV2(7m8)로의 신규한 캡시드에 의해 전달된 트랜스유전자 발현의 비교.

**발명을 실시하기 위한 구체적인 내용**

[0014] 용어 및 정의

[0015] 본 명세서의 맥락에서 약어 AAV는 아데노-연관 바이러스에 관한 것이다.

[0016] 본 명세서의 맥락에서 용어 AAV 벡터는 60개의 AAV 캡시드 단백질 및 캡시드화된 AAV 핵산으로 이루어진 바이러스 벡터에 관한 것이다. AAV 벡터는 AAV 비리온으로부터 유도되지만, AAV 벡터는 AAV 계통으로부터 rep 및 cap 유전자를 제거함으로써 헬퍼 바이러스의 존재 하에 복제-부적격으로 조작된다. 캡시드화된 AAV 핵산은 표적 세포 내로 전달될 트랜스유전자를 포함할 수 있다.

[0017] 달리 정의되지 않는 한, 본원에서 사용된 모든 기술 및 과학 용어는 (예를 들어, 세포 배양, 분자 유전학, 핵산 화학, 혼성화 기법 및 생화학에서) 당업자에 의해 통상적으로 이해되는 바와 동일한 의미를 갖는다. 표준 기법이 분자적, 유전적 및 생화학적 방법(일반적으로, 문헌[Sambrook et al., Molecular Cloning: A Laboratory Manual, 2d ed. (1989) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. 및 Ausubel et al., Short Protocols in Molecular Biology (1999) 4th Ed, John Wiley & Sons, Inc.] 참고) 및 화학적 방법에 대해 사용된다.

[0018] 본 명세서의 맥락에서 용어 AAV 캡시드는 본 발명자들에 의해 생성되고 아래에서 본원에 기재된 조작된 캡시드 (cap) 유전자에 의해 인코딩되는 폴리펩타이드에 관한 것이다. 본원에서 개시된 AAV 캡시드는 유전자 치료법을 위한 재조합 아데노-연관 바이러스 벡터를 조립하기 위해 사용될 수 있다.

[0019] 본 명세서의 맥락에서 AAV 캡시드에서의 아미노산 위치에 대한 언급은 SEQ ID NO. 1(표 2)의 서열을 갖는 아데노-연관 바이러스 2 캡시드 단백질 VP1의 캡시드 아미노산 서열에 관한 것이다(대응하는 핵산 서열의 GenBank 수탁 번호는 J01901.1임). 다른 동종 아데노-연관 바이러스 혈청형에 대한 대응하는 아미노산 위치를 표 3에 나타낸다.

[0020] 본 명세서의 맥락에서 용어 동종은 AAV의 상이한 혈청형으로부터 유래되는 핵산 또는 아미노산 서열에 관한 것이다. 상이한 아데노-연관 바이러스 혈청형의 대응하는 위치를 표 3에 나타낸다.

[0021] 본 명세서의 맥락에서 용어 트랜스유전자는 하나의 유기체에서 또 다른 유기체로 전달된 유전자 또는 유전 물질에 관한 것이다. 본 발명의 맥락에서, 이 용어는 또한 이것이 누락된 환자의 조직 내로 유전자 서열의 천연 또는 생리적으로 온전한 변이체의 전달을 나타낼 수 있다. 이는 또한 그 발현이 표적화된 조직에서 부재하거나 침묵화된 프로모터에 의해 유도되는, 천연 인코딩된 서열의 전달을 추가로 나타낼 수 있다.

[0022] 본 명세서의 맥락에서 용어 재조합은 하나의 또는 몇몇 클로닝, 제한 및/또는 결찰 단계의 산물이고 자연 발생 핵산과 상이한 핵산에 관한 것이다. 재조합 바이러스 입자는 재조합 핵산을 포함한다.

[0023] 본 명세서의 맥락에서 용어 유리체내 투여는 약학 제제, 예를 들어 바이러스 벡터의 투여 경로에 관한 것이며, 여기서 제제는 눈의 유리체 내로 전달된다. 유리체내 투여는 유리체강으로 불리는 눈의 후방 공간 내로 직접 투약을 배치하는 절차이며, 이는 유리체 방수 겔로 불리는 젤리-유사 유체로 충전된다.

[0024] 본 명세서의 맥락에서 용어 망막하 투여는 망막 색소 상피(RPE) 세포 및 광수용체 간 공간 내로의, 본 명세서의 맥락에서 약학 제제, 특히 바이러스 벡터의 투여 경로에 관한 것이다.

[0025] 본 명세서의 맥락에서 용어 폴리펩타이드는 선형 사슬을 형성하는 50개 이상의 아미노산으로 구성되는 분자에 관한 것이며 여기서 아미노산은 펩타이드 결합에 의해 연결된다. 폴리펩타이드의 아미노산 서열은 (생리적으로

확인된 바와 같은) 전체 단백질 또는 이의 단편의 아미노산 서열을 나타낼 수 있다. 용어 "폴리펩타이드" 및 "단백질"은 본원에서 상호 교환적으로 사용되며 단백질 및 이의 단편이 포함된다. 폴리펩타이드는 본원에서 아미노산 잔기 서열로 개시된다.

[0026] 아미노산 잔기 서열은 아미노 말단으로부터 카복실 말단으로 주어진다. 서열 위치에 대한 대문자는 L-아미노산을 한 글자 코드로 나타낸다(Stryer, Biochemistry, 3<sup>rd</sup> ed. p. 21). 아미노산 서열 위치에 대한 소문자는 대응하는 D- 또는 (2R)-아미노산을 나타낸다. 서열은 아미노 말단으로부터 카복시 말단의 방향으로 왼쪽에서 오른쪽으로 기재된다. 표준 명명에 따르면, 아미노산 잔기 서열은 하기에 나타낸 바와 같은 3글자 또는 1글자 코드로 명명된다: 알라닌(Ala, A), 아르기닌(Arg, R), 아스파라긴(Asn, N), 아스파르트산(Asp, D), 시스테인(Cys, C), 글루타민(Gln, Q), 글루탐산(Glu, E), 글리신(Gly, G), 히스티딘(His, H), 이소류신(Ile, I), 류신(Leu, L), 라이신(Lys, K), 메티오닌(Met, M), 페닐알라닌(Phe, F), 프롤린(Pro, P), 세린(Ser, S), 트레오닌(Thr, T), 트립토판(Trp, W), 티로신(Tyr, Y), 및 발린(Val, V).

[0027] 용어 *변이체*는 참조 폴리펩타이드와 상이하지만 본질적 특성을 보유하는 폴리펩타이드를 나타낸다. 폴리펩타이드의 전형적 변이체는 참조로 사용된 또 다른 폴리펩타이드와 그 일차 아미노산 서열이 상이하다. 일반적으로, 차이는 참조 폴리펩타이드 및 변이체의 서열이 전반적으로 밀접하게 유사하고, 여러 영역에서, 동일하도록 제한된다. 변이체 및 참조 폴리펩타이드는 하나 이상의 변형(예를 들어, 치환, 부가, 및/또는 결실)에 의해 아미노산 서열이 상이할 수 있다. 폴리펩타이드 변이체는 자연 발생, 예컨대 대립유전자 변이체일 수 있거나, 자연 발생하는 것으로 알려져 있지 않은 변이체일 수 있다.

[0028] 본 명세서의 맥락에서 용어 *생물학적 활성*은 본원에서 개시된 폴리펩타이드의 소정 변이체에 의해 나타나는 구체적으로 측정 가능한 양에 관한 것이다. 예를 들어, 망막 세포 내로의 유전자 전달을 촉진하는 바이러스 벡터 능력의 맥락에서, 캡시드 변이체의 생물학적 활성은 배양된 세포 내로의 또는 표준 검정에서 생체내 모델 유기체 내로의 트랜스유전자(예를 들어 반딧불이 루시퍼라제)의 발현을 측정함으로써 검정될 수 있다.

[0029] 본 명세서의 맥락에서, 용어 *서열 동일성 및 서열 동일성의 백분율*은 2개의 정렬된 서열을 비교함으로써 결정된 값을 나타낸다. 비교를 위한 서열의 정렬 방법은 당분야에서 잘 알려져 있다. 비교를 위한 서열의 정렬은, 문헌 [Smith and Waterman, Adv. Appl. Math. 2:482 (1981)]의 국소 상동성 알고리즘, 문헌 [Needleman and Wunsch, J. Mol. Biol. 48:443 (1970)]의 전체 정렬 알고리즘, 문헌 [Pearson and Lipman, Proc. Nat. Acad. Sci. 85:2444 (1988)]의 유사성 검색 방법, 또는 CLUSTAL, GAP, BESTFIT, BLAST, FASTA 및 TFASTA가 포함되지만 이에 제한되지 않는, 이들 알고리즘의 전산화된 구현에 의해 수행될 수 있다. BLAST 분석을 수행하기 위한 소프트웨어는, 예를 들어 미국 국립 생명공학 정보 센터(National Center for Biotechnology Information(<http://blast.ncbi.nlm.nih.gov/>))를 통해 공개적으로 이용 가능하다.

[0030] 아미노산 서열의 비교를 위한 하나의 예는 하기 디폴트 설정을 사용하는 BLASTP 알고리즘이다: 예상 역치: 10; 워드 크기: 3; 쿼리 범위에서의 최대 매치: 0; 매트릭스: BLOSUM62; 갭 코스트: 존재 11, 연장 1; 조성 조정: 조건별 조성 점수 매트릭스 조정. 핵산 서열의 비교를 위한 하나의 이러한 예는 하기 디폴트 설정을 사용하는 BLASTN 알고리즘이다: 예상 역치: 10; 워드 크기: 28; 쿼리 범위에서의 최대 매치: 0; 매치/미스매치 점수: 1.-2; 갭 코스트: 선형. 달리 언급되지 않는 한, 본원에서 제공된 서열 동일성 값은 각각 단백질 및 핵산 비교를 위한 상기 식별된 디폴트 파라미터를 사용하여 프로그램 BLAST 스위트를 사용해서 수득된 값을 나타낸다 (Altschul et al., J. Mol. Biol. 215:403-410 (1990)).

[0031] 본 명세서의 맥락에서, 용어 *아미노산 링커*는 단일쇄 폴리펩타이드를 생성하기 위해 2개의 폴리펩타이드를 연결하기 위해 사용되는 가변 길이의 올리고펩타이드를 나타낸다. 본원에서 명시된 본 발명의 실시를 위해 유용한 링커의 예시적 구현에는 1개 내지 6개 아미노산으로 구성되는 올리고펩타이드쇄이다. 아미노산 링커의 비제한적 예는 아래에서 예시된 바와 같은, AAA 또는 AA이다.

[0032] AAV는 인간 및 다른 영장류 종을 감염시키는 소형 비-병원성 바이러스이다.

[0033] AAV2 감염은 세포 표면 수용체 헤파란 설페이트 프로테오글리칸(HSPG)을 도킹함으로써 시작된다. 글리칸에 대한 그 저친화도 결합은 공동-수용체  $av\beta 5$  또는  $a5\beta 1$  인테그린에 대한 결합을 촉진하여 클래트린-코팅 오목부 (clathrin-coated pit)의 형성을 유도하는 캡시드의 가역적 구조 재배열을 유도한다. 클래트린-코팅 오목부는 세포내섭취를 통해 내재화되고 바이러스 입자는 핵으로 수송된다. pH는 엔도솜 구획의 산성화로 인해 감소하고, 이는 엔도솜 소수포 성숙의 특징이다. 산성화-유발 입체형태 변화가 캡시드에서 일어나며, 바이러스는 지방용해 포어 형성에 의해 후기 엔도솜으로부터 탈출한다.

- [0034] 야생형 AAV에서 게놈은 양성- 또는 음성-센스의 4.7킬로염기 길이 단일 가닥 DNA(ssDNA)로 구축된다. 게놈은 역방위 말단 반복(ITR)이 측면에 있는 3개의 개방 해독틀(ORF)을 포함한다. ITR은 AAV 게놈의 5' 및 3'-말단의 자가-상보성, CG-풍부, T-형 헤어핀이며, 재조합 벡터 게놈에 존재하는 유일하게 필요한 바이러스 성분이다. ITR에는 말단 분해 부위(TRS) 및 Rep 결합 요소(RBE)가 포함되며, 이는 바이러스 게놈의 복제 및 캡시드화를 촉진한다. ORF는 유전자 rep, cap, AAP를 인코딩한다. rep에 의해 인코딩된 4개의 다기능 비구조 Rep 단백질이 AAV 생활 주기를 위해 요구된다. Cap은 캡시드 단백질 VP1, VP2 및 VP3을 인코딩하며, 이는 함께 상호작용하여 정20면체 대칭의 캡시드, 및 어셈블리-활성화 단백질(AAP)을 형성하고, 이는 세포질로부터 세포 핵 내로의 새로 생성된 VP 단백질의 수송 및 안정화를 위해 요구된다. 3개의 VP 모두가 하나의 mRNA로부터 번역되고 상이하게 스플라이싱된다. 가장 큰 90 kDa VP1은 스플라이스되지 않은 전사체이며, 72 kDa VP2는 비-통상적 ACG 시작 코돈으로부터 번역되는 반면 가장 작은 60 kDa VP3은 AUG 코돈으로부터 번역된다. 3개의 VP 모두는 중첩되는 C-말단을 갖는다.
- [0035] VP3은 캡시드의 90%를 구성하며 VP1 및 VP2는 VP3와 공통 C-말단 아미노산 서열을 공유하지만 캡시드 내부에 매장된 채 유지되는 N-말단 연장을 갖는다. VP1 고유 N-말단 서열은 감염을 위해 요구되는 포스포리파제 A2(PLA2) 활성화, 및 핵 국제화 신호를 함유한다. 3개의 VP 단량체는 2-폴드, 3-폴드, 및 5-폴드 축 대칭으로 조립되어 60개 서브유닛의 AAV 캡시드를 생성한다. 본 발명의 제1 양태는 아데노-연관 바이러스(AAV) 혈청형 2 캡시드의 위치 453 또는 587에서의 피크 또는 스파이크성 돌출부에서 펩타이드 삽입을 포함하는 AAV 캡시드 폴리펩타이드에 관한 것이다. 펩타이드 삽입은 하기 서열로부터 선택되며:
- [0036] SASEAST(Cap3; SEQ ID NO 10), DTRPHDQ(Cap5; SEQ ID NO 11), EHYNSTC(Cap7; SEQ ID NO 12), PNPNTL(Cap9; SEQ ID NO 13), TPPSITA(Cap11; SEQ ID NO 14), CGESSYL(Cap12; SEQ ID NO 15), PRTPHTA(Cap13; SEQ ID NO 16) 및 ELCDGFA(Cap14; SEQ ID NO 17),
- [0037] 특히 삽입은 SASEAST(Cap3; SEQ ID NO 10), TPPSITA(Cap11; SEQ ID NO 14), PRTPHTA(Cap13; SEQ ID NO 16) 및 ELCDGFA(Cap14; SEQ ID NO 17)로부터 선택된다.
- [0038] 상기 제시된 바와 같은 펩타이드 삽입은 앞 단락에 주어진 삽입 서열의 N- 및/또는 C-말단에 삽입된 Ala, Leu, Gly, Ser, 및 Thr로부터 선택되지만 이에 제한되지 않는, 0 내지 6개 링커 아미노산이 측면에 있는(6개는 총 N 및 C 말단 측면 아미노산의 최대수임), 상기 서열 중 하나로 구성되는 7개 내지 13개에 포함될 수 있다.
- [0039] 본원에서 주어진 AAV2 서열에 대해 주어진 임의의 위치는 GenBank 엔트리 번호 J01901.1(아데노-연관 바이러스 2, 전체 게놈)에서 접근 가능한 참조 서열과 대비된다. 대응하는 아미노산 서열은 표 2에서 SEQ ID NO 1로 주어진 다.
- [0040] 스파이크성 돌출부(피크)는 캡시드의 가장 노출된 영역을 나타낸다. 최고 피크는 AAV2 캡시드 상에서 아미노산 위치 453에 및 두 번째 최고 피크는 위치 587에 위치한다. 이들 피크는 캡시드 어셈블리를 손상시키지 않는 펩타이드 삽입을 수용하며 비-허용성 세포를 표적화할 기회를 제공한다. 마찬가지로, 돌출부는 AAV 숙주 상호작용, 수용체 결합 및 면역원성의 중추적 부위를 나타낸다. 소정 구현예에서, 삽입이 위치 453에 삽입되는 경우 R585 및 588이 최적 효율을 달성하도록 돌연변이됨이 주지된다.
- [0041] 소정 피크 또는 스파이크성 돌출부 위치를 참조하면, 펩타이드 삽입이 도입되는 경우, 펩타이드 삽입은 이러한 위치의 (C-말단 방향으로) 바로 다음 또는 C-말단 방향으로 추가의 1개 내지 6개 아미노산 다음에 있을 수 있음이 이해되어야 한다. 예를 들어, 위치 587로 정의되는 펩타이드 삽입(삽입-전 위치 587 및 588 사이로 정의되는 삽입)은 위치 588 또는 589(C-말단 방향으로 추가의 1개 또는 2개 아미노산) 또는 590(C-말단 방향으로 추가의 3개 아미노산)에서 시작할 수 있다. 이들 삽입 위치 모두는 전반적 캡시드 구조를 손상시키지 않고, 오히려 형질도입 효율을 증가시킬 수 있는 펩타이드 삽입을 허용한다.
- [0042] 이러한 펩타이드 삽입은 유리체내 전달 후 바이러스의 망막 침투 및 형질도입 효율을 증가시킨다.
- [0043] 소정 구현예에서, 펩타이드 삽입은 SASEAST(Cap3; SEQ ID NO 10), DTRPHDQ(Cap5; SEQ ID NO 11), EHYNSTC(Cap7; SEQ ID NO 12), PNPNTL(Cap9; SEQ ID NO 13), TPPSITA(Cap11; SEQ ID NO 14), CGESSYL(Cap12; SEQ ID NO 15), PRTPHTA(Cap13; SEQ ID NO 16) 및 ELCDGFA(Cap14; SEQ ID NO 17)로부터 선택된다.
- [0044] 소정 구현예에서, 펩타이드 삽입은 SASEAST(Cap3; SEQ ID NO 10), TPPSITA(Cap11; SEQ ID NO 14), PRTPHTA(Cap13; SEQ ID NO 16) 및 ELCDGFA(Cap14; SEQ ID NO 17)로부터 선택된다.
- [0045] 펩타이드 삽입은 아미노 말단 및/또는 카복실 말단에서 그대로 또는 1개, 2개, 3개 또는 4개의 측면 스페이서

아미노산의 맥락에서 AAV 캡시드 폴리펩타이드에 존재할 수 있다. 적합한 스페이서 AA에는 알라닌, 류신, 글리신, 트레오닌 및 세린이 포함되지만 이에 제한되지 않는다.

- [0046] 본 발명자들의 예에서, 사용된 클로닝 전략으로 인해, AAV2 VP1의 N<sub>587</sub> 및 R<sub>588</sub> 간 삽입은 형태 R<sub>585</sub>G N A A A X<sub>1</sub> X<sub>2</sub> X<sub>3</sub> X<sub>4</sub> X<sub>5</sub> X<sub>6</sub> X<sub>7</sub>-A A R<sub>588</sub> Q A A의 12머이며, X<sub>1</sub> 내지 X<sub>7</sub>은 삽입된 헵타머 올리고를 나타내고 A는 측면 링커를 나타낸다. 이는 일레일 뿐이며, 본 발명은 정확히 이러한 측면 서열을 갖는 정확히 이러한 위치에서의 삽입으로 제한되지 않음이 이해되어야 한다.
- [0047] AAV2에서 587로 명명된 삽입 위치에 관한 구현예에서, 올리고는 항상 AAV2의 VP1 상 위치 587 및 588 간에 삽입된다. 이는 삽입으로, 이에 따라 VP1 AAV2 넘버링이 변하지 않고 이후 위치가 치환 등에 관해 "삽입 없이" 카운트된다. 삽입된 펩타이드는 본원에서 특정된 바와 같은 헵타머를 함유하지만 잠재적으로는 또한 기능하기 위해 더 짧거나 더 길 수 있다. 실시예에서 제시된 구현예에서, 본 발명자들은 형태 AAA-X1X2X3X4X5X7-AA의 12머를 삽입한 반면, 링커 아미노산은 Ala, Leu, Gly, Ser, Thr으로부터 자유롭게 선택될 수 있다. N- 및 C-말단 링커는 0 내지 5개 AA 길이일 수 있다. 소정 구현예에서, 삽입 펩타이드는 0 내지 6개 링커 아미노산을 포함하는, 5 내지 13개 아미노산 길이일 수 있다.
- [0048] 소정 구현예에서, 펩타이드 삽입(본 명세서의 다른 곳에서는 "올리고"로 나타냄)은 C-말단으로 1 내지 5AA 더 이동할 수 있다. AAV2에 있어서, 삽입 부위는 또한 아미노산 453에(예를 들어, 그 바로 C-말단에) 있을 수 있다.
- [0049] 본 발명의 이러한 제1 양태의 대안에 따르면, 상기 열거된 임의의 펩타이드 서열은 위치 587 대신 SEQ ID NO 001의 위치 453에 삽입된다. 또한, 삽입 위치는 변할 수 있고 펩타이드 서열은 상기 논의된 바와 같이 삽입이 측면에 있을 수 있고 5 내지 13개 AA로 구성될 수 있다. 위치 453에서의 펩타이드 삽입에 부가하여, R587A 및 R588A 돌연변이가 AAV2의 VP1 내로 도입될 수 있다(Boucas Jet al. J Gene Med. 2009 Dec;11(12):1103-13. doi: 10.1002/jgm.1392)
- [0050] 본 발명의 이러한 양태의 또 다른 대안에 따르면, AAV 혈청형 2에서 위치 587/588 또는 453과 동종인 위치가 상기 열거된 삽입된 펩타이드 서열 중 하나를 갖는 또 다른 혈청형의 AAV를 작제하기 위해 선택될 수 있다(표 3: #1, #2).
- [0051] 소정 구현예에서, AAV 캡시드 단백질은 티로신 잔기의 하나 또는 몇몇 티로신의 페닐알라닌으로의 치환을 특징으로 하며, 여기서 티로신 잔기는 위치 252, 272, 444, 500, 700, 704 및 730에서 야생형 캡시드 서열에서 일어난다.
- [0052] 본원에서 소정 캡시드 변이체는 티로신-변형 캡시드의 맥락에서 진화적 방법에 의해 선택되었다.
- [0053] 소정 구현예에서, AAV 캡시드 단백질은 위치 252, 272, 444, 500, 704 및 730에서 하나 또는 몇몇 티로신의 페닐알라닌으로의 치환을 특징으로 한다.
- [0054] 소정의 보다 구체적인 특정 구현예에서, AAV 캡시드 단백질은 위치 252, 272, 444, 500, 700 및 730 모두에서 몇몇 티로신의 페닐알라닌으로의 치환을 특징으로 한다.
- [0055] 본 발명의 이러한 양태의 또 다른 대안에 따르면, 티로신의 페닐알라닌으로의 치환을 갖는 또 다른 혈청형의 AAV를 작제하기 위해 AAV 혈청형 2에서 위치 252, 272, 444, 500, 700, 704 및 730 중 임의의 것과 동종인 티로신 위치가 페닐알라닌으로의 치환을 위해 선택될 수 있다(표 3: #3 내지 #9). 또한, 다른 티로신 위치가 페닐알라닌으로 치환될 수 있다.
- [0056] 소정 구현예에서, AAV 캡시드 단백질은 하나 또는 몇몇 트레오닌의 발린으로의 치환, 특히 T491V를 특징으로 한다.
- [0057] 본 발명의 이러한 양태의 또 다른 대안에 따르면, 트레오닌에서 발린으로의 치환을 갖는 또 다른 혈청형의 AAV를 작제하기 위해 AAV 혈청형 2에서 위치 491과 동종인 트레오닌 위치가 발린으로의 치환을 위해 선택될 수 있다(표 3: #10).
- [0058] 소정 구현예에서, 상기 언급된 티로신에서 페닐알라닌으로의 치환은 트레오닌에서 발린으로의 치환, 특히 T491V와 조합되어 있다.
- [0059] Y-F 및 T-V 돌연변이는 세포에 의해 내재화되면 바이러스 입자의 유비퀴틴화, 및 이에 따른 프로테아좀 분해를

감소시키고, 이에 따라 형질도입 효율을 증가시킨다.

- [0060] 소정 구현예에서, 아데노-연관 바이러스 캡시드 폴리펩타이드는 SEQ ID NO 2 내지 SEQ ID NO 9로부터 선택되는 아미노산 서열, 또는 이와 적어도 85% 동일성 및 SEQ ID NO 2 내지 SEQ ID NO 9로부터 선택되는 서열의 생물학적 활성의 적어도 90%를 갖는 서열을 포함한다.
- [0061] 생물학적 활성 검정:
- [0062] "(참조) 서열의 생물학적 활성의 적어도 소정 백분율을 갖는" 폴리펩타이드가 본원에서 언급되는 경우마다, 이러한 생물학적 활성은 하기와 같이 측정할 수 있다:
- [0063] HEK293 세포를 시험 또는 참조 폴리펩타이드를 인코딩하는 플라스미드로 공동-전달감염시키며, 플라스미드는 갈슘-포스페이트 침전 방법을 사용하여, 헬퍼 아데노바이러스 유전자 및 리포터로서 scCMV-mCitrine, scEfla-mCitrine, scCMV-EGFP, scEfla-EGFP로부터 선택되는 트랜스유전자 플라스미드를 보유한다. 트랜스유전자를 함유하는 벡터를 요오디사놀 구배(Axis-Shield, Oslo) 및 40% 요오디사놀 분획 상에서 밀도 정제에 의해 농축한 후, 예를 들어 아미콘 여과(Millipore)에 의해 완충액 교환한다. AAV 분획을 표준 벡터 대비 실시간 PCR에 의해 DN 아제-내성 벡터 게놈에 대해 적정한다. 3  $\mu$ l의 바이러스 벡터를 마취된(100 mg/kg 케타민 및 10 mg/kg 실라진) 야생형 C57BL/6J 마우스 내로 유리체내 주사한다. 주사를 위해, 바이러스를 E11 GC/ml로 역가-매칭하고 이의 3  $\mu$ l - 즉, 눈 당 3x10<sup>8</sup> vg를 주사한다. 마우스 망막을 면역조직화학 분석을 위해 주사 3주 후 제거한다. 마우스 눈 망막 외식편(형질도입 7일 후)을 실온(RT)에서 40분 동안 PBS 중 4%(wt/vol) 파라포름알데하이드로 고정하고, 등급화된 수크로스 용액(PBS 중 10%, 20%, 및 30% 수크로스) 중 4°C에서 연속 3일에 걸쳐 냉동 보호하고, O.C.T. 화합물(Sakura Finetek)과 함께 냉동몰드에서 포매하고, 액체 질소-냉각 2-메틸부탄 중에 냉동한다. 10  $\mu$ m 두께의 수직 절편을 크리오스타트 상에서 절단하고, SuperFrost 유리 슬라이드(Menzel) 상에 실장하고 적합한 일차 및 이차 항체로 면역조직화학적으로 표지하여 EGFP, YFP, mCitrine 또는 Turbo635 형광뿐만 아니라 특정 망막 세포 유형을 검출한다. 표적 세포 유형에서의 리포터 발현을, 전형적으로 형광에 의해 정량하고, 나머지 조직에서의 리포터의 양과 비교하여 조작된 바이러스 캡시드의 생물학적 활성을 평가한다. 조작된 바이러스 캡시드의 생물학적 활성에 대한 또 다른 척도는 형질도입된 표적 세포의 백분율 및 발현 조직의 면적이다. 의혹을 배제하기 위해, 생물학적 활성을 측정하는 경우 표적 세포 유형에서 리포터 발현을 정량하고 나머지 조직에서의 리포터 양과 비교하는 검정이 채택된다.
- [0064] 본 발명의 제2 양태는 제1 양태에 따른 AAV 캡시드 폴리펩타이드를 인코딩하는 핵산 서열에 관한 것이다.
- [0065] 특정 구현예에서, 본 발명에 따른 AAV 서열은 숙주 게놈 내로의 통합을 위해 요구되는 Rep 요소를 함유하지 않는다. 비-통합 바이러스가 인간에서의 적용을 위해 더 안전하다. 소정 구현예에서, 핵산 서열은 자가-상보적 AAV 벡터 플라스미드 개념에 따라 설계된다. 모든 경우에, 단일 가닥 DNA 서열을 사용한다.
- [0066] 소정 구현예에서, 핵산 서열은 조절 서열을 포함하거나 포함하지 않고 트랜스유전자를 포함한다.
- [0067] 소정 구현예에서, 트랜스유전자는 유전자 편집을 위해 정확한 단백질(즉 RPE65)의 서열, siRNA 또는 shRNA(우성 질환-유발 유전자의 야생형 또는 돌연변이체 RNA를 분해하기 위해 mRNA의 영역을 표적화하도록 설계됨), 또는 CRISPR/Cas-gRNA 카세트를 인코딩한다.
- [0068] 소정 구현예에서, 트랜스유전자는 광-감수성 단백질, 예컨대 무척추동물 또는 척추동물 옵션 또는 이의 변이체를 인코딩한다.
- [0069] 소정 구현예에서, 트랜스유전자는 채널로돕신-2 또는 이의 변이체를 인코딩한다.
- [0070] 소정 구현예에서, 트랜스유전자는 미생물 광-관문화 억제 이온 펌프, 예컨대 할로로돕신(eNpHR) 또는 원시로돕신(ArchT) 등을 인코딩한다.
- [0071] 소정 구현예에서, 트랜스유전자는 GPCR 옵션, 예컨대 인간 로돕신, 멜라놉신 또는 원뿔 옵션 등을 인코딩한다.
- [0072] 소정 구현예에서, 트랜스유전자는 광-스위치 가능한 리간드에 대한 수신체 단백질을 인코딩한다.
- [0073] 소정 구현예에서, 트랜스유전자는 포유류 세포에서 작동 가능한 프로모터 서열의 제어 하에 있다. 소정 구현예에서, 프로모터 서열은 인간 망막 세포에서 작동 가능하다.
- [0074] 소정 구현예에서, 프로모터는 도처의 프로모터이다. 소정 구현예에서, 프로모터는 세포-특이적 프로모터이다.
- [0075] 소정 구현예에서, 프로모터는 CMV 최조기(immediate early) 프로모터이다. 소정 구현예에서, 프로모터는 인간

Efla 프로모터이다. 소정 구현예에서, 프로모터는 광수용체 특이적 프로모터이다.

- [0076] 본 발명의 제3 양태는 의약에서/약제로서 사용하기 위한, 제1 양태에 따른 AAV 캡시드 폴리펩타이드, 제1 양태에 따른 캡시드 폴리펩타이드를 포함하는 AAV 벡터, 및 제2 양태에 따른 핵산 서열로부터 선택되는 제제를 나타낸다.
- [0077] 소정 구현예에서, AAV 벡터가 눈의 세포를 형질도입시키기 위해 사용된다.
- [0078] 눈은 다양한 해부학적 구조로 이루어진, 3개 층으로 구성된다. 섬유성 막은 최외층이며 각막 및 공막으로 이루어진다. 혈관성 막 또는 포도막은 중간층이며 맥락막, 모양체, 색소 상피 및 홍채로 구성된다. 망막은 최내층이며, 이는 맥락막의 혈관(후방으로)뿐만 아니라 망막 혈관(전방으로)으로부터 그 산소화를 얻는다.
- [0079] 각막 및 수정체 간 공간은 방수로 충전되어 있고, 수정체 뒤의 전체 후방강은 젤리-유사 성분인 유리체로 충전되어 있다. 유리체는 물, 콜라겐, 원섬유, 히알루론산 및 이온으로 이루어진다. 뉴런 또는 혈관에 의해 점유되지 않은 망막 내 공간은 모든 망막 세포층에 걸쳐 필러 아교 세포의 돌기(process)에 의해 충전되어 있다.
- [0080] 망막의 외부 경계막(OLM)은 필러 세포(MC) 및 광수용체 세포의 내부 구획 간 접합부로부터 형성되며 망막하 공간 간 대사 장벽으로 작용하여, 큰 분자의 통과를 제한한다. 망막의 내부 경계막(ILM)은 MC 종말단추 및 기저막 간 측면 연결에 의해 형성되고 유리체 방수 및 신경 망막 간 확산 장벽으로 작용한다.
- [0081] 망막은 황반, 시신경유두, 오목 및 말초 망막으로 이루어진다. 광수용체 세포는 망막에서 확인되는 신경상피 세포의 특화된 유형이다. 3개 유형의 광수용체 세포가 알려져 있다: 막대, 원뿔, 및 광감수성 망막 신경절 세포. 막대는 망막의 말초 영역에 분포되어 있다; 반면에 황반으로 불리는 중심 색소 영역은 원뿔 광수용체 세포가 농축되어 있다. 망막 색소 상피(RPE)는 영양을 제공하며 광수용체 세포의 건강을 유지한다. 광수용체 세포의 핵은 외부 핵층(ONL)을 구성하는 반면 두극, 무축삭 및 수평 세포의 핵은 내부 핵층(INL)에 위치한다.
- [0082] 소정 구현예에서, 제1 양태에 따른 AAV 캡시드 폴리펩타이드, 제1 양태에 따른 캡시드 폴리펩타이드를 포함하는 AAV 벡터, 및 제2 양태에 따른 핵산 서열로부터 선택되는 제제는 하기에 영향을 미치는 질병의 치료에서 사용하기 위한 것이다:
- [0083] a. 망막 또는 망막 색소 상피 세포 및/또는
- [0084] b. 망막의 광수용체, 두극 세포, 무축삭 세포 또는 신경절 세포.
- [0085] 소정 구현예에서, 제제는 녹내장, 색소 망막염, 황반 변성, 망막층간분리, 레베르 선천 흑암시, 당뇨병성 망막 병증, 완전 색맹, 또는 색맹, 흑색종-연관 망막병증, 선천성 부동 야맹증, 원뿔-막대 이영양증, 후기 단계 연령-관련 황반 변성, 황반병증, 초기 발병 중증 망막 이영양증, 완전 색맹, 안구 백색증, 눈피부 백색증, 슈트가르트병(Stargardt disease), 범맥락막위축, 척수소뇌성 실조증 유형 7(SCAT), 각막에 영향을 미치는 리소좀 저장 질환, 예컨대 점액다당류증(MPS) IV 및 MPS VII, 망막모세포종, 안구 흑색종, 고혈압 망막병증의 치료에서 사용하기 위한 것이다.
- [0086] 소정 구현예에서, 본 발명의 제제는 내이에 영향을 미치는 질환의 치료 또는 예방을 위해 채택될 수 있다.
- [0087] 소정 구현예에서, 제제는 하기에 의해 투여된다:
- [0088] a. 유리체내 투여에 의해, 특히 유리체내 주사에 의해, 또는
- [0089] b. 망막하 투여에 의해.
- [0090] 소정 구현예에서, 제제는 후방 구획, 전방 구획, 공막, 맥락막, 결막, 홍채, 수정체, 또는 각막으로 전달된다.
- [0091] 유사하게, 환자에 상기 기재에 따른 AAV 캡시드 및/또는 핵산 서열을 포함하는 바이러스 벡터를 투여하는 단계를 포함하는, 이를 필요로 하는 환자에서 색소 망막염을 포함하는 막대-원뿔 이영양증, 황반 변성을 포함하는 원뿔-막대 이영양증 및 선천성 부동 야맹증(CSBN1)으로부터 선택되는 질병을 치료하는 방법이 본 발명의 범위 내에 있다.
- [0092] 유사하게, 본 발명의 상기 양태 중 하나에 따른 AAV 캡시드 및/또는 핵산 서열을 포함하는, 원뿔-막대 이영양증, 막대-원뿔 이영양증 및 선천성 부동 야맹증으로부터 선택되는 질병의 예방 또는 치료를 위한 투여형이 제공된다.
- [0093] 본원에서 개시된 제제 및 방법은 또한, 본 발명에 의해 촉진되는 벡터가 임의의 기존 대안에 비해 더 효율적이



바이러스 게놈은 그 게놈이 인코딩하는 캡시드 단백질 변이체 내에 패키징되거나 캡시드화되도록 바이러스 라이브러리를 생성하였다(유전형-표현형 커플링). 또한, 모든 캡시드를 한 종류의 펩타이드만 60개 서브유닛 각각에 존재하는 방식으로 생성하였다. 이러한 방식으로, 선택을 통해 식별된 기능적 개선은 바이러스 캡시드 내에 함유된 이러한 개선된 기능을 인코딩하는 게놈 서열에 연관될 수 있다.

[0105] 실시예 2: 펩타이드 변이체의 선택

[0106] 망막 ON-두극 세포에서 빨간색 형광 마커 FP635를 발현하는, 본 실험실에서 생성된 C57BL/6<sub>Opto</sub>-mGluR6 마우스 (van Wyk et al., PLoS Biol, 2015. 13(5): p. e1002143) 내에서 이러한 라이브러리로 양성 선택을 거쳤다. ON-두극 세포가 망막 내에 깊이 위치하며 AAV 형질도입에 대해 허용성이 아닌 것으로 알려져서, 선택이 잘 침투하는 AAV 및 ON-두극 세포를 형질도입시키기 위해 유리한 특성을 갖는 변이체를 선호할 것이라는 논리에 기반하여 이러한 마우스 라인을 선택하였다. 간략하게, 4 내지 6주령의 5 내지 8마리의 트랜스제닉 마우스에 문헌(van Wyk et al., PLoS Biol, 2015. 13(5): p. e1002143; van Wyk et al., Front Neurosci, 2017. 11: p. 161)에 기재된 바와 같이 양안 내로 대략  $5 \times 10^{11}$  개 바이러스 게놈(vg)/ml 의 게놈 역가를 갖는 2.5  $\mu$ l의 요오딕산-정제 라이브러리를 유리체내 주사하였다.

[0107] 10일 후, 눈을 적출하고 가벼운 과과인 프로테아제 처리를 사용하여 망막을 해리한 후 ON-두극 세포 집단의 형광 활성화 세포 정렬(FACS)을 수행하였다. 이어서 성공적인 비리온을 DNA 추출로부터 PCR-증폭하고 추가 클로닝하고 후속 회의 주사를 위해 재패키징하였다.

[0108] 7머 올리고의 AAV2(M6) 게놈 내로의 각각의 클로닝 단계에, 4개의 병렬 걸찰을 수행하고 각각의 풀의 12개 클론을 플레이트접종하고 서열분석하여 각각의 생체내 선택 단계 후 라이브러리의 집합성을 추정하였다. 이러한 생체내 유도 진화 과정은 천연 진화 과정과 유사한, 양성 선택의 적용을 통해 ON-두극 세포 허용성 AAV 변이체를 생성하였다.

[0109] 생체내 선택된 라이브러리로부터, 81'738개 변이체의 캡시드 유전자를 차세대 시퀀싱에 의해 식별하였다. 상위 67개 순위로부터의 15개 펩타이드 서열을 높은 전반적 형질도입 효능(NGS에서의 축적), ON-두극 세포 표적화에 대한 선호 및 선택 회에서 초기에 나타나거나 이의 아미노산 서열의 관점에서 유망하게 보이는(즉, 음으로 하전된 잔기 또는 프롤린 잔기를 함유함) 모티프의 조합에 대해 표 1에 개략된 바와 같이 기능적 평가 및 특성 규명을 위해 선택하였다. 모든 라이브러리 변이체의 핵심 특징은 삽입을 통한 AAV2의 헤파란 설페이트 프로테오글리칸 결합 모티프의 양전하(R<sub>585</sub> 및 R<sub>588</sub>)의 분리이다. 그러나, 링커 서열의 2개 알라닌과 함께 펩타이드 서열은 헤파란 설페이트 프로테오글리칸에 결합하는 능력을 재구축할 수 있다(Uhrig... **Büning**, Gene Therapy). 7머 삽입 서열 중에는, 특정 위치에서, 예를 들어, 위치 1 또는 2 및/또는 위치 5 또는 6에서 하전된 아미노산 및 위치 7에서 극성 아미노산, 예컨대 Cys 또는 Ala의 중등도 선호가 존재하였다(표 1).

[0110] 실시예 3: 신규 캡시드 변이체의 평가

[0111] AAV2(M6)~7머~의 제조형 상태를 표 1에서와 같이 15개의 선택된 펩타이드 모두에 대해 클로닝하였다. 본 발명자들은 scCMV-EGFP 트랜스유전자(CMV: 사이토메갈로바이러스 프로모터)와 함께 패키징하였다. 2개의 변이체(Cap8 및 Cap10)를 제외하고, 모든 캡시드는  $10^{11}$  내지  $10^{12}$  vg/ml 간의 역가로 잘 패키징되었다. 성체 마우스 내로의 유리체내 주사 3주 후, 주로 광수용체에서 강력한 발현이 Cap3, Cap5, Cap7, Cap9, Cap 11, Cap 13 및 Cap 14에 대해 관찰되었으나(표 1, 도 1 내지 도 3), 또한 내부 망막 세포, 예컨대 두극 세포, 무축삭 세포 및 뿔러 세포 그리고 극히 소수의 망막 신경절 세포에서 일부 발현이 관찰되었다. 또한 진화된 AAV2 변이체(Dalkara et al., Sci Transl Med, 2013. 5: p. 189ra76.)이며 현재 유리체내 주사에 대한 최신 기술로 간주되는 AAV2(7m8)의 성능과의 직접 비교에서, 본 발명자들의 신규한 변이체는 망막을 통해 유의미하게 더 잘 침투하였고 유의미하게 더 많은 광수용체 및 유의미하게 더 적은 신경절 세포를 형질도입시켰다(도 1, 도 1 내지 도 3). 8개의 사전-선택된 변이체(Cap3, Cap5, Cap7, Cap9, Cap 11, Cap 13 및 Cap 14; 표 1)를 또한 schEfla-EGFP-mCitrine 트랜스유전자(hEfla: 인간 Efla 프로모터, "강력한" 천연 포유류의 도처의 프로모터)와 함께 패키징하고, 성체 마우스에 매칭된 역가  $1 \times 10^{11}$  vg/ml로 유리체내 주사함으로써 상기 결과를 확인하였다.

[0112] 동등하게, 새로운 변이체는 역가-매칭된 대조군 AAV2(M6) 및 AAV(7m8)에 비해 유의미하게 더 많은 수의 광수용체를 형질도입시켰다. 결과적으로, 선택된 변이체는 AAV2(M6) 및 AAV2(7m8)에 비해 더 우수한 내부 경계막 및 망막 침투 특성을 지니는 것으로 나타났고 외부 핵층(ONL)에서의 광수용체의 형질도입에서 유의미하게 더 효과

적이다(도 1).

[0113] scCMV-EGFP 및 schEf1a-EGFP가 패키징되고 3E+10 vg/ml로 주사된 Cap14는 동일한 저용량으로 주사된 AAV2(7m8)보다 유의미하게 더 우수하게(도 1, 도 2) 그리고 심지어 마우스에서 전형적으로 사용되는 1E+12 vg/ml의 최소 용량으로 AAV2(7m8)보다 더 우수하게 작용하였다. 이는 독성 및 면역 반응의 우려를 감소시키기 위해 환자에게서 더 낮은 용량을 사용하는 측면에서 유망하다.

[0114] [표 1]

[0115] 유도된 진화로부터 ON-두극 세포 단리된 변이체의 차세대 시퀀싱이 바이러스 라이브러리에서 고도의 집합성을 드러내었다. 7번째 삽입에는 본원에서 변이체 #3으로 예시되는 ~N<sub>587</sub>AAASEASTAA R<sub>588</sub>~(SEQ ID NO 29) 형태의 알려진 스페이서가 측면에 있었다.

[0116] ON-두극 세포 우세성 := 발생률[ON-두극 세포 / 비-ON-두극 세포]

클론	변이체 (Cap) #	ON - 두극 세포 분획		비-ON-두극 세포 분획		ON-두극 세포 우세성	SEQ ID NO
		F	O (%)	F	O (%)		
HHNDRAP#	15	6050	9.9	1121	5.4	1.8	18
DGNLHKS#	2	2837	4.7	1069	5.1	0.9	19
SASEAST*	3	2814	4.6	1511	7.2	0.6	10
HSDSSKP	4	1583	2.6	370	1.8	1.5	20
DTRPHDQ	5	1378	2.3	473	2.3	1.0	11
PNPNCTL	9	403	0.7	41	0.2	3.4	13
TLTGLCP	8	505	0.8	50	0.2	3.5	21
EHYNSTC	7	310	0.5	27	0.1	3.9	12
NHAPNHC	6	292	0.5	25	1	4.0	22
DRRPGIC	1	209	0.34	4	0.02	17.9	23
VDATCMQ	10	402	0.7	40	0.2	3.4	24

[0117]

TPPSITA*	11	342	0.6	35	0.2	3.3	14
CGESSYL	12	334	0.5	37	0.2	3.1	15
PRTPHTA*	13	227	0.4	27	0.1	2.9	16
ELCDGFA*	14	222	0.4	38	0.2	2.0	17

[0118] F = 빈도; O = 발생률

[0120] #제2 선택 회에서 드러남

- [0121] \*생체내 스크리닝으로부터 선택된 상위 후보
- [0122] 삽입 Nr. 8 및 Nr. 10을 가진 캡시드는 패키징하지 않음(회색 폰트)

[0123] **양으로 하전됨**

[0124] 음으로 하전됨

[0125] 15개의 선택된 클론 모두는 C57BL/6\_Opto-mGluR6 마우스(van Wyk et al., PLoS Biol, 2015. 13(5): p. e1002143)의 단리된 ON-두극 세포 분획으로부터 서열분석된 60'884개 펩타이드의 상위 67개 내의 순위였다.

[0126] [표 2]

[0127] 상이한 삽입을 가진 AAV2 캡시드 단백질 VP1의 아미노산 서열.

[0128] VP1과 중첩되는 VP3(회색 서열), 티로신의 페닐알라닌(Y-F)으로의 돌연변이를 진회색으로 밑줄쳐 강조하며, 아미노산 넘버링은 전체 VP1 서열을 나타낸다. 최고 피크는 **G<sub>453</sub>**에 및 두 번째 최고 피크는 **N587**에, 무작위 7머를 삽입하였고, 밑줄쳐서 회색 언더레이로 나타낸다. 삽입은 알라닌 링커를 포함하며, 이탤릭체로 나타내고, 상자로 표시한다.

[0129] SI: 펩타이드 삽입 서열의 SEQ ID NO; FS: 전체 캡시드 서열의 SEQ ID NO

SI	삽입	FS	서열
-	AAV2 (Y252, 272, 444, 500, 700, 730F)	1	<b>MAADGYLPDWLEDTLSEGI</b> RQWWKLKPGPPPKPAERHKDDSR GLVLPGYKYLGPFNGLDKGEPVNEADAAALEHDKAYDRQLDSG DNPYLYKNHADA <b>E</b> FQERLKEDTSFGGNLGRAVFQAKKRVLEPL GLVEEPVK <b>T</b> APGKKRPVEHSPVEPDSSSGTGKAGQQPARKRLN FGQ <b>T</b> GDADSVDPD <b>Q</b> PLGQPPAAP <b>S</b> GLGTNT <b>M</b> ATGSGAPMADNN EGADGVGNSSGN <b>W</b> HCDSTW <b>M</b> GDRVIT <b>T</b> STR <b>T</b> WALPT <b>F</b> NNHLYK QISSQSGASNDNH <b>E</b> FGY <b>S</b> TPWGYFDENRFHCHFS <b>P</b> RDWQRLIN NN <b>W</b> GF <b>R</b> PKRLNFKLENIQVKEVTQNDGTT <b>I</b> ANNLTSTVQVFT

[0130]

			DSEYQLPYVLGSAHQGCLPPFPADVEMVPOYGYLTLNNGSQAV GRSSFYCLEYFPSQMLRTGNNFTFSYTFEDVFFHSSYAHSQSL DRLMNPLIDQYLYFLSRNTNTPSGTTTQSRLOFSQAGASDIRDQ SRNWLPGPCYRQQRVSKTSADNNNSEFSWTGATKYHLNGRDSL VNP GPAMASHKDDEEKFFPQSGVLI FGKQGSEKTNVDIEKVM I TDEEEIRTTNPVATEQYGSVSTNLQRGNRQAATADVNTQGVL P GMVWQDRDVYLOGPIWAKIPH TDGHFHPSPLMGGFGLKHPP PQ ILIKNTPVPANPSTTFSAAKFASFITQYSTGQVSV EIEWELQK ENSKRWNPEIQFTSNYNKSVNVDFTVDTNGVYSEPRPIGTR E LTRNL
10	SASEAST	2	MAADGYLPDWLEDTLSEGI RQWWKLP GPPPPKPAERHKDDSR GLVLPGYKYLGPFNGLDKGEPVNEADAAA LEHDKAYDRQLDSG DNPYLKYNHADA EFQERLKEDTSFGGNLGRAVFQAKKRVL EPL GLVEEPVK TAPGKKRPVEHSPVEPDSSSGTGKAGQQPARKRLN FGQTDADSVDPDQPLGQPPAAPSGLGTNTMATGSGAPMADNN EGADGVGNSSGNW HCDSTWMDRVITTTSTRTWALPTENNHL YK QISSQSGASNDNH EFGYSTPWGYFDENRFHCHFSPRDQRLIN NNWGF RPKRLNFKLFNIQVKEVTQNDGTTTIANNLTSTVQVFT DSEYQLPYVLGSAHQGCLPPFPADVEMVPOYGYLTLNNGSQAV GRSSFYCLEYFPSQMLRTGNNFTFSYTFEDVFFHSSYAHSQSL DRLMNPLIDQYLYFLSRNTNTPSGTTTQSRLOFSQAGASDIRDQ SRNWLPGPCYRQQRVSKTSADNNNSEFSWTGATKYHLNGRDSL VNP GPAMASHKDDEEKFFPQSGVLI FGKQGSEKTNVDIEKVM I TDEEEIRTTNPVATEQYGSVSTNLQRGN AAASASEASTAARQAA TADVNTQGVLPGMVWQDRDVYLOGPIWAKIPH TDGHFHPSPLM GGFGLKHPP PQILIKNTPVPANPSTTFSAAKFASFITQYSTGQ VSV EIEWELQKENS KRWNPEIQFTSNYNKSVNVDFTVDTNGVY SEPRPIGTR E LTRNL
11	DTRPHDQ	3	MAADGYLPDWLEDTLSEGI RQWWKLP GPPPPKPAERHKDDSR GLVLPGYKYLGPFNGLDKGEPVNEADAAA LEHDKAYDRQLDSG DNPYLKYNHADA EFQERLKEDTSFGGNLGRAVFQAKKRVL EPL GLVEEPVK TAPGKKRPVEHSPVEPDSSSGTGKAGQQPARKRLN FGQTDADSVDPDQPLGQPPAAPSGLGTNTMATGSGAPMADNN EGADGVGNSSGNW HCDSTWMDRVITTTSTRTWALPTENNHL YK QISSQSGASNDNH EFGYSTPWGYFDENRFHCHFSPRDQRLIN

[0131]

			<p>NNWGRPKRLNFKLFENIQVKEVTQNDGTTTIANNLTSTVQVFT  DSEYQLPYVLGSAHQGCLPPFPADVEMVPOYGYLTLNNGSQAV  GRSSFYCLEYFPSQMLRTGNNFTFSYTFEDVFFHSSYAHSQSL  DRLMNPLIDQYLYFLSRNTNTPSGTTTQSRLOFSQAGASDIRDQ  SRNWLPGPCYRQQRVSKTSADNNNSEFSWTGATKYHLNGRDSL  VNP GPAMASHKDDEEKFFPQSGVLI FGKQGSEKTNVDIEKVM I  TDEEEIRTTNPVATEQYGSVSTNLQRGNAAADTRPHDQAAARQAA  TADVNTQGVLPGMVWQDRDVYLOGPIWAKIPHTDGHFHP SPLM  GGFGLKHPPPQILIKNTVPV PANPSTTF SAAKFA SFITQYSTGQ  VSVEIEWELQKENS KRWNPEIQFTSNYNKSVNVDFTVDTNGVY  SEPRPIGTRFLTRNL</p>
12	EHYNSTC	4	<p>MAADGYLPDWLEDTLSEGIRQWWKLP GPPPPKPAERHKDDSR  GLVLPGYKYLGPFNGLDKGEPVNEADAAA LEHDKAYDRQLDSG  DNPYLKYNHADA EFQERLKEDTSFGGNLGRAVFQAKKRVL EPL  GLVEEPVK TAPGKKRPVEHSPVEPDS SSGTGKAGQQPARKRLN  FGQTDADSVDPDQPLGQPPAAPSG LGTNTMATGSGAPMADNN  EGADGVGNSSGNW HCDSTWMGDRVIT TSTRTWALPTENN HLYK  QISSQSGASNDNHFFGYSTPWGYFDENRFHCHFS PRDWQRLIN  NNWGRPKRLNFKLFENIQVKEVTQNDGTTTIANNLTSTVQVFT  DSEYQLPYVLGSAHQGCLPPFPADVEMVPOYGYLTLNNGSQAV  GRSSFYCLEYFPSQMLRTGNNFTFSYTFEDVFFHSSYAHSQSL  DRLMNPLIDQYLYFLSRNTNTPSGTTTQSRLOFSQAGASDIRDQ  SRNWLPGPCYRQQRVSKTSADNNNSEFSWTGATKYHLNGRDSL  VNP GPAMASHKDDEEKFFPQSGVLI FGKQGSEKTNVDIEKVM I  TDEEEIRTTNPVATEQYGSVSTNLQRGNAAA EHYNSTCAAARQAA  TADVNTQGVLPGMVWQDRDVYLOGPIWAKIPHTDGHFHP SPLM  GGFGLKHPPPQILIKNTVPV PANPSTTF SAAKFA SFITQYSTGQ  VSVEIEWELQKENS KRWNPEIQFTSNYNKSVNVDFTVDTNGVY  SEPRPIGTRFLTRNL</p>
13	PNPNCTL	5	<p>MAADGYLPDWLEDTLSEGIRQWWKLP GPPPPKPAERHKDDSR  GLVLPGYKYLGPFNGLDKGEPVNEADAAA LEHDKAYDRQLDSG  DNPYLKYNHADA EFQERLKEDTSFGGNLGRAVFQAKKRVL EPL  GLVEEPVK TAPGKKRPVEHSPVEPDS SSGTGKAGQQPARKRLN  FGQTDADSVDPDQPLGQPPAAPSG LGTNTMATGSGAPMADNN  EGADGVGNSSGNW HCDSTWMGDRVIT TSTRTWALPTENN HLYK  QISSQSGASNDNHFFGYSTPWGYFDENRFHCHFS PRDWQRLIN</p>

[0132]

			<p>NNWGRPKRLNFKLFNIQVKEVTQNDGTTTIANNLTSTVQVFT  DSEYQLPYVLGSAHQGCLPPFPADVEMVPOYGYLTLNNGSQAV  GRSSFYCLEYFPSQMLRTGNNFTFSYTFEDVFFHSSYAHSQSL  DRLMNFLIDQYLYFLSRNTNTPSGTTTQSRLOFSQAGASDIRDQ  SRNWLPGPCYRQQRVSKTSADNNNSEFSWTGATKYHLNGRDSL  VNP GPAMASHKDDEEKFFPQSGVLI FGKQGSEKTNVDIEKVM I  TDEEEIRTTNPVATEQYGSVSTNLQRGNAAAPNPCTLAAARQAA  TADVNTQGVLPGMVWQDRDVYLOGPIWAKIPHTDGHFHP SPLM  GGFGLKHPPPQILIKNTVPV PANPSTTF SAAKFA SFITQYSTGQ  VSVEIEWELQKENS KRWNPEIQFTSNYNKSVNVDFTVDTNGVY  SEPRPIGTRFLTRNL</p>
14	TPPSITA	6	<p>MAADGYLPDWLEDTLSEGIRQWWKLPKPPPKPAERHKDDSR  GLVLPGYKYLGPFNGLDKGEPVNEADAAALEHDKAYDRQLDSG  DNPYLKYNHADA EFQERLKEDTSFGGNLGRAVFQAKKRVL EPL  GLVEEPVKTAPGKKRPVEHSPVEPDS SSGTGKAGQQPARKRLN  FGQTDADSVDPDPQLGQPPAAPSGLGTNTMATGSGAPMADNN  EGADGVGNSSGNWCDSTWMGDRVITSTRTWALPTENNHL YK  QISSQSGASNDNHFFGYSTPWGYFDENRFHCHFSPRDWQRLIN  NNWGRPKRLNFKLFNIQVKEVTQNDGTTTIANNLTSTVQVFT  DSEYQLPYVLGSAHQGCLPPFPADVEMVPOYGYLTLNNGSQAV  GRSSFYCLEYFPSQMLRTGNNFTFSYTFEDVFFHSSYAHSQSL  DRLMNFLIDQYLYFLSRNTNTPSGTTTQSRLOFSQAGASDIRDQ  SRNWLPGPCYRQQRVSKTSADNNNSEFSWTGATKYHLNGRDSL  VNP GPAMASHKDDEEKFFPQSGVLI FGKQGSEKTNVDIEKVM I  TDEEEIRTTNPVATEQYGSVSTNLQRGNAAATPPSITAAARQAA  TADVNTQGVLPGMVWQDRDVYLOGPIWAKIPHTDGHFHP SPLM  GGFGLKHPPPQILIKNTVPV PANPSTTF SAAKFA SFITQYSTGQ  VSVEIEWELQKENS KRWNPEIQFTSNYNKSVNVDFTVDTNGVY  SEPRPIGTRFLTRNL</p>
15	CGESSYL	7	<p>MAADGYLPDWLEDTLSEGIRQWWKLPKPPPKPAERHKDDSR  GLVLPGYKYLGPFNGLDKGEPVNEADAAALEHDKAYDRQLDSG  DNPYLKYNHADA EFQERLKEDTSFGGNLGRAVFQAKKRVL EPL  GLVEEPVKTAPGKKRPVEHSPVEPDS SSGTGKAGQQPARKRLN  FGQTDADSVDPDPQLGQPPAAPSGLGTNTMATGSGAPMADNN  EGADGVGNSSGNWCDSTWMGDRVITSTRTWALPTENNHL YK  QISSQSGASNDNHFFGYSTPWGYFDENRFHCHFSPRDWQRLIN</p>

[0133]

			<p>NNWGRPKRLNFKLFENIQVKEVTQNDGTTTIANNLTSTVQVFT  DSEYQLPYVLGSAHQGCLPPFPADVFMVPOYGYLTLNNGSQAV  GRSSFYCLEYFPSQMLRTGNNFTFSYTFEDVFPFHSSYAHSQSL  DRLMNPLIDQYLY<del>EL</del>SRNTNTPSG<del>CT</del>TTQSRLOFSQAGASDIRDQ  SRNWLPGPCYRQQRVSKTSADNNNSE<del>ES</del>SWTGATKYHLNGRDSL  VNP GPAMASHKDDEEKFFPQSGVLI<del>FG</del>KQSEKTNVDIEKVM  TDEEEIRTTNPVATEQYGSVSTNL<del>QR</del>GN<del>AAACGESSYLA</del>A<del>R</del>QAA  TADVNTQGVLPGMVWQDRD<del>VY</del>LOGPIWAKIPHTDGHFHPSPLM  GGFGLKHPPPQILIKNTVPANPSTTF<del>SA</del>AKFASFITQYSTGQ  VSVEIEWELQKENS<del>KRW</del>NPEIQ<del>ET</del>SNYNKSVNVDFTVDTNGVY  SEPRPIGTR<del>EL</del>TRNL</p>
16	PRTPHTA	8	<p>MAADGYLPDWLEDTLSEGIQWKKLPGPPPKPAERHKDDSR  GLVLPGYKYLGPFNGLDKGEPVNEADAAALEHDKAYDRQLDSG  DNPYLKYNHADAEFQERLKEDTSFGGNLGRAVFQAKKRVLEPL  GLVEEPVKTAPGKKRPVEHSPVEPDSSTGTGKAGQQPARKRLN  FGQTDADSVDPDQPLGQPPAAPSGLGTNTMATGSGAPMADNN  EGADGVGNSSGNWCDSTWMDRVITTTSTRTWALPT<del>EN</del>NHLYK  QISSQSGASNDNH<del>EF</del>GYSTPWGY<del>FD</del>ENRFHCHFSPRDWQRLIN  NNWGRPKRLNFKLFENIQVKEVTQNDGTTTIANNLTSTVQVFT  DSEYQLPYVLGSAHQGCLPPFPADVFMVPOYGYLTLNNGSQAV  GRSSFYCLEYFPSQMLRTGNNFTFSYTFEDVFPFHSSYAHSQSL  DRLMNPLIDQYLY<del>EL</del>SRNTNTPSG<del>CT</del>TTQSRLOFSQAGASDIRDQ  SRNWLPGPCYRQQRVSKTSADNNNSE<del>ES</del>SWTGATKYHLNGRDSL  VNP GPAMASHKDDEEKFFPQSGVLI<del>FG</del>KQSEKTNVDIEKVM  TDEEEIRTTNPVATEQYGSVSTNL<del>QR</del>GN<del>AAAPRTPHTAA</del>A<del>R</del>QAA  TADVNTQGVLPGMVWQDRD<del>VY</del>LOGPIWAKIPHTDGHFHPSPLM  GGFGLKHPPPQILIKNTVPANPSTTF<del>SA</del>AKFASFITQYSTGQ  VSVEIEWELQKENS<del>KRW</del>NPEIQ<del>ET</del>SNYNKSVNVDFTVDTNGVY  SEPRPIGTR<del>EL</del>TRNL</p>
17	ELCDGFA	9	<p>MAADGYLPDWLEDTLSEGIQWKKLPGPPPKPAERHKDDSR  GLVLPGYKYLGPFNGLDKGEPVNEADAAALEHDKAYDRQLDSG  DNPYLKYNHADAEFQERLKEDTSFGGNLGRAVFQAKKRVLEPL  GLVEEPVKTAPGKKRPVEHSPVEPDSSTGTGKAGQQPARKRLN  FGQTDADSVDPDQPLGQPPAAPSGLGTNTMATGSGAPMADNN  EGADGVGNSSGNWCDSTWMDRVITTTSTRTWALPT<del>EN</del>NHLYK  QISSQSGASNDNH<del>EF</del>GYSTPWGY<del>FD</del>ENRFHCHFSPRDWQRLIN</p>
			<p>NNWGRPKRLNFKLFENIQVKEVTQNDGTTTIANNLTSTVQVFT  DSEYQLPYVLGSAHQGCLPPFPADVFMVPOYGYLTLNNGSQAV  GRSSFYCLEYFPSQMLRTGNNFTFSYTFEDVFPFHSSYAHSQSL  DRLMNPLIDQYLY<del>EL</del>SRNTNTPSG<del>CT</del>TTQSRLOFSQAGASDIRDQ  SRNWLPGPCYRQQRVSKTSADNNNSE<del>ES</del>SWTGATKYHLNGRDSL  VNP GPAMASHKDDEEKFFPQSGVLI<del>FG</del>KQSEKTNVDIEKVM  TDEEEIRTTNPVATEQYGSVSTNL<del>QR</del>GN<del>AAAELCDGFAA</del>A<del>R</del>QAA  TADVNTQGVLPGMVWQDRD<del>VY</del>LOGPIWAKIPHTDGHFHPSPLM  GGFGLKHPPPQILIKNTVPANPSTTF<del>SA</del>AKFASFITQYSTGQ  VSVEIEWELQKENS<del>KRW</del>NPEIQ<del>ET</del>SNYNKSVNVDFTVDTNGVY  SEPRPIGTR<del>EL</del>TRNL</p>

[0134]

[0135]

[0136]

[0137]

[표 3]

상이한 AAV 혈청형에서 VP3의 캡시드 단백질 상에서, 각각 잠재적 올리고머 삽입 부위 및 Y-F 및 T-V 점 돌연변이 부위의 동중 위치.

위치	AAV1	AAV2	AAV5	AAV6	AAV7	AAV8	AAV9	AAV Rh10	AAV Anc80
#1	G454	G453	G446	G454	G456	G456	G455	G456	G454
#2	<u>P591</u>	R588	<u>S576</u>	<u>P591</u>	<u>T590</u>	<u>T591</u>	<u>A589</u>	<u>A591</u>	<u>T589</u>
#3	Y252	Y252	Y242	Y252	Y253	Y253	Y252	Y253	Y252
#4	Y273	Y272	Y263	Y273	Y274	Y275	Y274	Y275	Y273
#5	Y445	Y444	<u>R437</u>	Y445	Y446	Y447	Y446	Y447	Y445
#6	<u>F501</u>	Y500	<u>S487</u>	<u>F501</u>	<u>F503</u>	<u>F503</u>	<u>F501</u>	<u>F503</u>	<u>F501</u>
#7	Y701	Y700	Y689	Y701	Y702	Y703	Y701	Y703	Y701
#8	Y705	Y704	Y693	Y705	<u>F706</u>	Y707	Y705	Y707	Y705
#9	Y731	Y730	Y719	Y731	Y732	Y733	Y731	Y733	Y731
#10	T492	T491	<u>G478</u>	T492	T494	T494	T492	T494	T492

[0138]

[0139]

밑줄친 위치는 AAV2 대비 상이한 아미노산을 나타낸다.

[0140]

[표 4]

[0141]

AAV2(M6)로 나타내는, 아데노-연관 바이러스 2(Y252, 272, 444, 500, 700, 730F)의 캡시드 유전자 서열(VP1)

[0142]

SEQ ID NO	서열

27	<p>ATGGTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACACTCTCT          CTGAAGGAATAAGACAGTGGTGGAGCTCAAACCTGGCCACCACC          ACCAAAGCCCGCAGAGCGGCATAAGGACGACAGCAGGGGTCTTGTG          CTTCTGGGTACAAGTACCTCGGACCCTTCAACGGACTCGACAAGG          GAGAGCCGGTCAACGAGGCAGACGCCGCGGCCCTCGAGCACGACAA          AGCCTACGACCCGGCAGCTCGACAGCGGAGACAACCCGTACCTCAAG          TACAACCACGCCGACGCGGAGTTTCAGGAGCGCCTTAAAGAAGATA          CGTCTTTTGGGGCAACCTCGGACGAGCAGTCTTCCAGGCGAAAAA          GAGGGTTCTTGAACCTCTGGGCCGTTGAGGAACCTGTTAAGACG          GCTCCGGGAAAAAAGAGGCCGGTAGAGCACTCTCTGTGGAGCCAG          ACTCTCCTCGGGAACCGAAAGCGGGCCAGCAGCCTGCAAGAAA          AAGATTGAATTTTGGTCAGACTGGAGACGCACTCAGTACCTGAC          CCCCAGCCTCTCGGACAGCCACCAGCAGCCCCCTCTGGTCTGGGAA          CTAATACGATGGCTACAGGCAGTGGCGCACCAATGGCAGACAATAA          CGAGGGCGCCGACGGAGTGGGTAATTCCTCCGGAAATTGGCATTGC          GATTCCACATGGATGGGCGCAGAGTCATCACCACCAGCACC CGAA          CCTGGGCCCTGCCACCCTTCAACAACCACCTGTACAAACAAATTC          CAGCCAAATCAGGAGCTTCGAACGACAATCACTTCTTTGGCTACAGC          ACCCTTGGGGTATTTGACTTCAACAGATTCCTACTGCCACTTTT          CACCACGTGACTGGCAAAGACTCATCAACAACAAC TGGGGATTCCG          ACCCAAGAGACTCAACTTCAAGCTCTTTAACATTCAAGTCAAAGAG          GTCACGCAGAATGACGGTACGACGACGATTGCCAATAACCTTACCA          GCACGGTTCAGGTGTTTACTGACTCGGAGTACCAGCTCCCGTACGT          CCTCGGCTCGGCGCATCAAGGATGCCTCCCGCGTTCCAGCAGAC          GTCTTTCATGGTGCCACAGTATGGATACCTCACCTGAACAACGGGA          GTCAGGCAGTAGGACGCTCTTCAATTTACTGCCTGGAGTACTTTC          TTCTCAGATGCTGCGTACC GGAAACAAC TTTACCTT CAGCTACACT          TTTGAGGACGTTCC TTTCCACAGCAGCTACGCTCACAGCCAGAGTC          TGGACCGTCTCATGAATCCTCTCATCGACCAGTACCTGTATTTCTT          AAGCAGAACAACACTCCAAGTGAACCACCACGCAGTCAAGGCTT          CAGTTTCTCAGGCCGAGCGAGTGACATTCGGGACCAGTCTAGGA          ACTGGCTTCTGGACCCGT TACC GCCAGCAGCGAGTATCAAAGAC          ATCTGCGGATAACAACAACAGTGAATTC TCGTGACC GG TGCTACC          AAGTACCACCTCAATGGCAGAGACTCTCTGGTGAATCCGGGCCGG          CCATGGCAAGCCACAAGGACGATGAAGAAAAGTTTTTCTCTCAGAG          CGGGGTTCTCATCTTTGGGAAGCAAGGCTCAGAGAAAACAAATGTG          GACATTGAAAAGGTCATGATTACAGACGAAGAGGAAATCAGGACAA</p>
----	--

[0143]

	<p>CCAATCCCGTGGCTACGGAGCAGTATGGTTCGTATCTACCAACCT          CCAGAGAGGCAACAGACAAGCAGCTACCGCAGATGTCAACACACAA          GCGGTTCTTCCAGGCATGGTCTGGCAGGACAGAGATGTGTACCTTC          AGGGGCCCATCTGGGCAAAGATTCCACACACGGACGGACATTTTCA          CCCCTCTCCCTCATGGGTGGATTCCGACTTAAACACCCTCCTCCA          CAGATTCTCATCAAGAACACCCCGGTACCTGCGAATCCTTCGACCA          CCTTCAGTGGGCAAAGTTTGTCTCCTTCATCACAGTACTCCAC          GGGACAGGTGAGGTTGGAGATCGAGTGGGAGCTGCAGAAGGAAAAC          AGCAAACGCTGGAAATCCCGAAATTCAGTTCACTTCGAACTACAACA          AGTCTGTTAATGTGGACTTTACTGTGGACACTAATGGCGTGTATTC          AGAGCCTCGCCCCATGGTACCAGATTCCTGACTCGTAATCTGTAA</p>
--	--

[0144]

[0145]

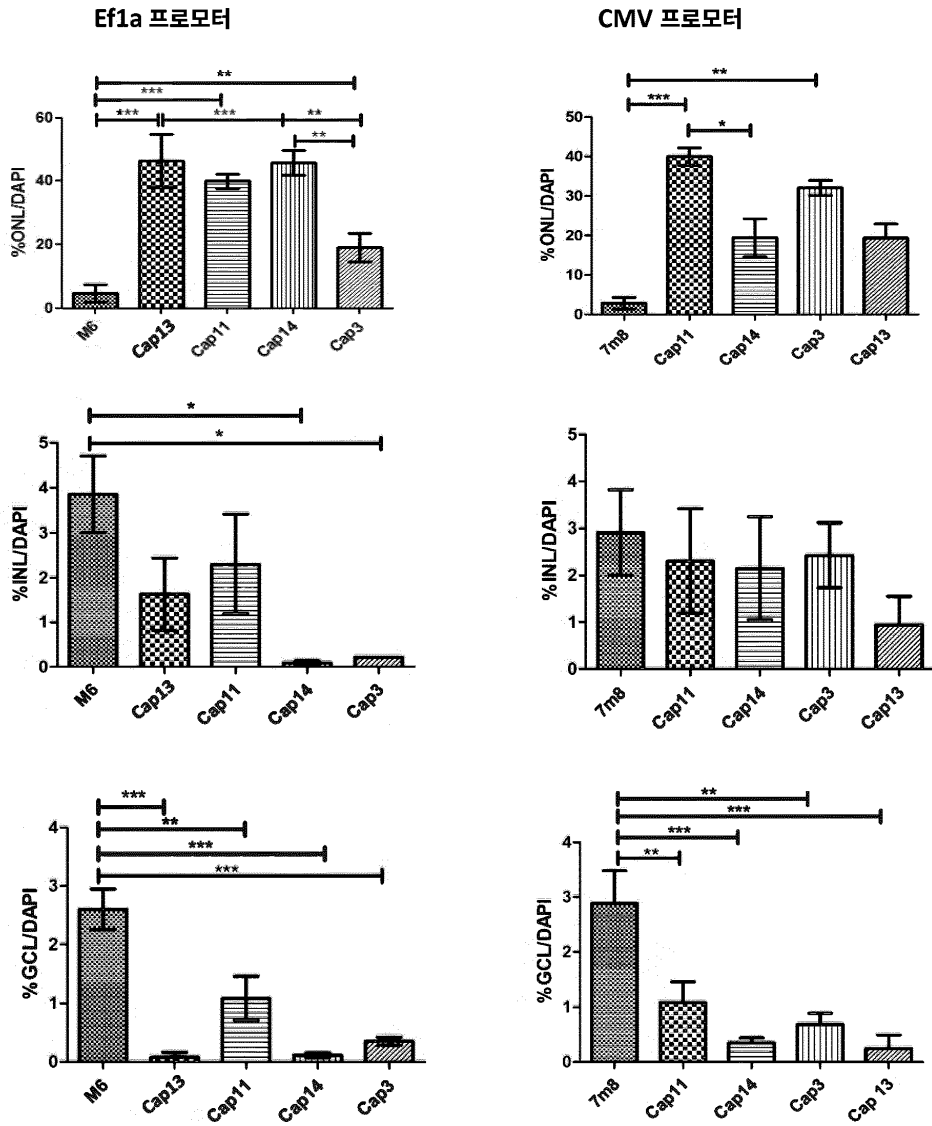
[0146]

AAV 2 전체 게놈 서열, GenBank 수탁 번호 J01901.1에 기반함

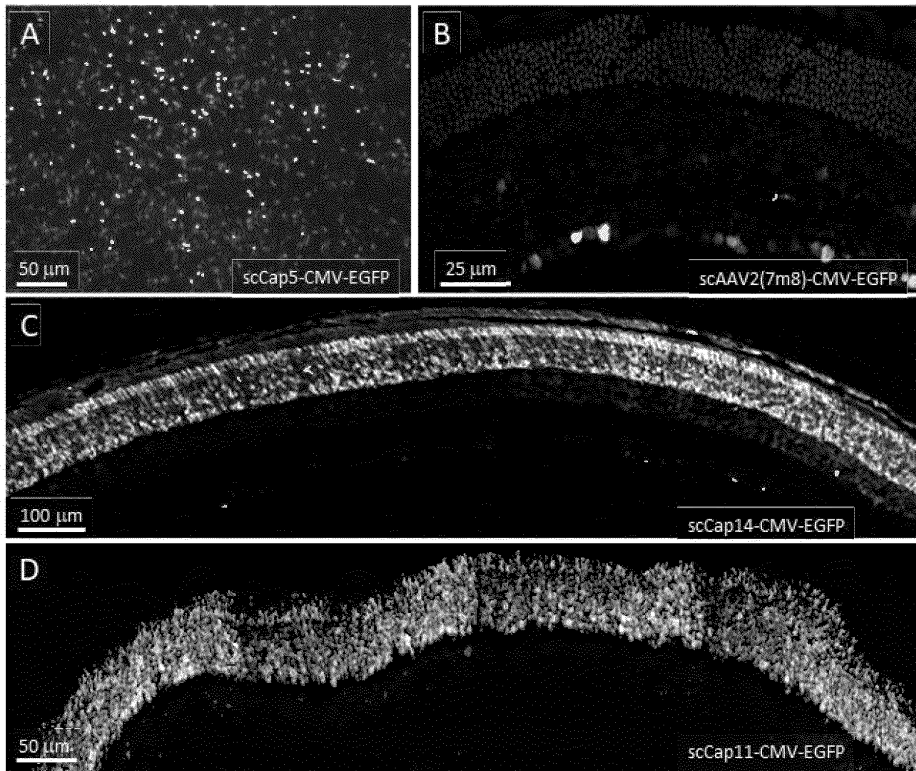
라이브러리의 Cap 유전자는 VP3(회색 서열)으로 구성되며, 티로신의 페닐알라닌(Y-F)으로의 돌연변이는 밑줄치고, 염기쌍 넘버링은 전체 VP1 서열을 나타낸다.

도면

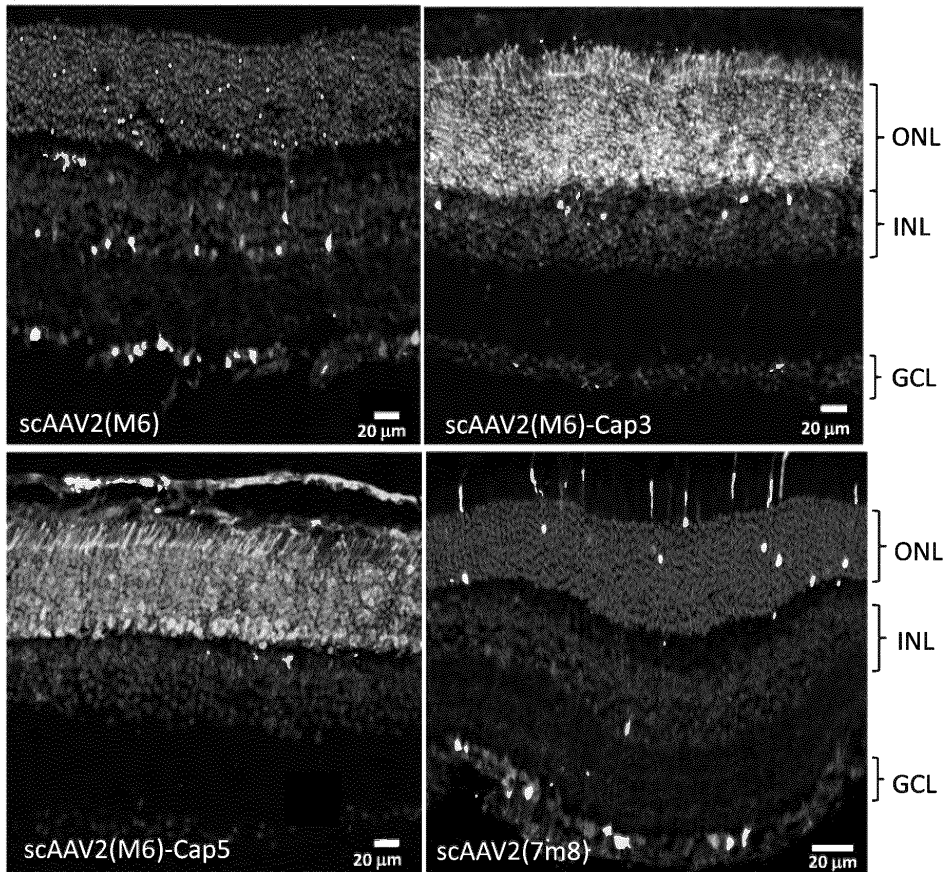
도면1



도면2



도면3



서열 목록

SEQUENCE LISTING

<110> Universit? Bern

<120> AAV Vector Variants for Ocular Gene Delivery

<130> uz374ep-a

<160> 28

<170> PatentIn version 3.5

<210> 1

<211> 735

<212> PRT

<213> Artificial Sequence

<220><223> AAV2 capsid protein VP1 with different mutations

<400> 1

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

1                    5                    10                    15

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

                  20                    25                    30

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

                  35                    40                    45

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

                  50                    55                    60

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

65                    70                    75                    80

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

                  85                    90                    95

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

                  100                    105                    110

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

                  115                    120                    125

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

                  130                    135                    140

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



Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe Glu  
 405 410 415

Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg  
 420 425 430

Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Phe Leu Ser Arg Thr  
 435 440 445

Asn Thr Pro Ser Gly Thr Thr Thr Gln Ser Arg Leu Gln Phe Ser Gln  
 450 455 460

Ala Gly Ala Ser Asp Ile Arg Asp Gln Ser Arg Asn Trp Leu Pro Gly  
 465 470 475 480

Pro Cys Tyr Arg Gln Gln Arg Val Ser Lys Thr Ser Ala Asp Asn Asn  
 485 490 495

Asn Ser Glu Phe Ser Trp Thr Gly Ala Thr Lys Tyr His Leu Asn Gly  
 500 505 510

Arg Asp Ser Leu Val Asn Pro Gly Pro Ala Met Ala Ser His Lys Asp  
 515 520 525

Asp Glu Glu Lys Phe Phe Pro Gln Ser Gly Val Leu Ile Phe Gly Lys  
 530 535 540

Gln Gly Ser Glu Lys Thr Asn Val Asp Ile Glu Lys Val Met Ile Thr  
 545 550 555 560

Asp Glu Glu Glu Ile Arg Thr Thr Asn Pro Val Ala Thr Glu Gln Tyr  
 565 570 575

Gly Ser Val Ser Thr Asn Leu Gln Arg Gly Asn Arg Gln Ala Ala Thr  
 580 585 590

Ala Asp Val Asn Thr Gln Gly Val Leu Pro Gly Met Val Trp Gln Asp  
 595 600 605

Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His Thr  
 610 615 620

Asp Gly His Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu Lys  
 625 630 635 640

His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala Asn

645 650 655  
 Pro Ser Thr Thr Phe Ser Ala Ala Lys Phe Ala Ser Phe Ile Thr Gln  
 660 665 670

Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln Lys  
 675 680 685

Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Phe Thr Ser Asn Tyr  
 690 695 700

Asn Lys Ser Val Asn Val Asp Phe Thr Val Asp Thr Asn Gly Val Tyr  
 705 710 715 720

Ser Glu Pro Arg Pro Ile Gly Thr Arg Phe Leu Thr Arg Asn Leu  
 725 730 735

<210> 2

<211> 747

<212> PRT

<213> Artificial Sequence

<220><223> AAV2 capsid protein VP1 with insertion

<400> 2

Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Thr Leu Ser  
 1 5 10 15

Glu Gly Ile Arg Gln Trp Trp Lys Leu Lys Pro Gly Pro Pro Pro Pro  
 20 25 30

Lys Pro Ala Glu Arg His Lys Asp Asp Ser Arg Gly Leu Val Leu Pro  
 35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro  
 50 55 60

Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp  
 65 70 75 80

Arg Gln Leu Asp Ser Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala  
 85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly  
 100 105 110

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



Thr Gln Gly Val Leu Pro Gly Met Val Trp Gln Asp Arg Asp Val Tyr  
 610 615 620

Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His Thr Asp Gly His Phe  
 625 630 635 640

His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu Lys His Pro Pro Pro  
 645 650 655

Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala Asn Pro Ser Thr Thr  
 660 665 670

Phe Ser Ala Ala Lys Phe Ala Ser Phe Ile Thr Gln Tyr Ser Thr Gly  
 675 680 685

Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys  
 690 695 700

Arg Trp Asn Pro Glu Ile Gln Phe Thr Ser Asn Tyr Asn Lys Ser Val  
 705 710 715 720

Asn Val Asp Phe Thr Val Asp Thr Asn Gly Val Tyr Ser Glu Pro Arg  
 725 730 735

Pro Ile Gly Thr Arg Phe Leu Thr Arg Asn Leu  
 740 745

<210> 3

<211> 747

<212> PRT

<213> Artificial Sequence

<220><223> AAV2 capsid protein VP1 with insertion

<400> 3

Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Thr Leu Ser  
 1 5 10 15

Glu Gly Ile Arg Gln Trp Trp Lys Leu Lys Pro Gly Pro Pro Pro Pro  
 20 25 30

Lys Pro Ala Glu Arg His Lys Asp Asp Ser Arg Gly Leu Val Leu Pro  
 35 40 45

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





545                      550                      555                      560  
 Asp Glu Glu Glu Ile Arg Thr Thr Asn Pro Val Ala Thr Glu Gln Tyr

                         565                      570                      575  
 Gly Ser Val Ser Thr Asn Leu Gln Arg Gly Asn Ala Ala Ala Asp Thr

                         580                      585                      590  
 Arg Pro His Asp Gln Ala Ala Arg Gln Ala Ala Thr Ala Asp Val Asn

                         595                      600                      605  
 Thr Gln Gly Val Leu Pro Gly Met Val Trp Gln Asp Arg Asp Val Tyr

                         610                      615                      620  
 Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His Thr Asp Gly His Phe

625                      630                      635                      640  
 His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu Lys His Pro Pro Pro

                         645                      650                      655  
 Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala Asn Pro Ser Thr Thr

                         660                      665                      670  
 Phe Ser Ala Ala Lys Phe Ala Ser Phe Ile Thr Gln Tyr Ser Thr Gly

                         675                      680                      685  
 Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys

                         690                      695                      700  
 Arg Trp Asn Pro Glu Ile Gln Phe Thr Ser Asn Tyr Asn Lys Ser Val

705                      710                      715                      720  
 Asn Val Asp Phe Thr Val Asp Thr Asn Gly Val Tyr Ser Glu Pro Arg

                         725                      730                      735  
 Pro Ile Gly Thr Arg Phe Leu Thr Arg Asn Leu

                         740                      745

<210> 4

<211> 747

<212> PRT

<213> Artificial Sequence

<220><223> AAV2 capsid protein VP1 with insertion

<400> 4

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

Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp  
 65                    70                    75                    80  
 Arg Gln Leu Asp Ser Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala  
                   85                    90                    95  
 Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly  
                   100                    105                    110  
 Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro  
                   115                    120                    125

Leu Gly Leu Val Glu Glu Pro Val Lys Thr Ala Pro Gly Lys Lys Arg  
                   130                    135                    140  
 Pro Val Glu His Ser Pro Val Glu Pro Asp Ser Ser Ser Gly Thr Gly  
 145                    150                    155                    160  
 Lys Ala Gly Gln Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln Thr  
                   165                    170                    175  
 Gly Asp Ala Asp Ser Val Pro Asp Pro Gln Pro Leu Gly Gln Pro Pro  
                   180                    185                    190

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

	245		250		255
Tyr Lys Gln Ile Ser Ser Gln Ser Gly Ala Ser Asn Asp Asn His Phe					
	260		265		270
Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His					
	275		280		285
Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp					
	290		295		300
Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile Gln Val					
305		310		315	320
Lys Glu Val Thr Gln Asn Asp Gly Thr Thr Thr Ile Ala Asn Asn Leu					
	325		330		335
Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr					
	340		345		350
Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp					
	355		360		365
Val Phe Met Val Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser					
	370		375		380
Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser					
385		390		395	400
Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe Glu					
	405		410		415
Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg					
	420		425		430
Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Phe Leu Ser Arg Thr					
	435		440		445
Asn Thr Pro Ser Gly Thr Thr Thr Gln Ser Arg Leu Gln Phe Ser Gln					
	450		455		460
Ala Gly Ala Ser Asp Ile Arg Asp Gln Ser Arg Asn Trp Leu Pro Gly					
465		470		475	480
Pro Cys Tyr Arg Gln Gln Arg Val Ser Lys Thr Ser Ala Asp Asn Asn					
	485		490		495

Asn Ser Glu Phe Ser Trp Thr Gly Ala Thr Lys Tyr His Leu Asn Gly  
 500 505 510

Arg Asp Ser Leu Val Asn Pro Gly Pro Ala Met Ala Ser His Lys Asp  
 515 520 525

Asp Glu Glu Lys Phe Phe Pro Gln Ser Gly Val Leu Ile Phe Gly Lys  
 530 535 540

Gln Gly Ser Glu Lys Thr Asn Val Asp Ile Glu Lys Val Met Ile Thr  
 545 550 555 560

Asp Glu Glu Glu Ile Arg Thr Thr Asn Pro Val Ala Thr Glu Gln Tyr  
 565 570 575

Gly Ser Val Ser Thr Asn Leu Gln Arg Gly Asn Ala Ala Ala Glu His  
 580 585 590

Tyr Asn Ser Thr Cys Ala Ala Arg Gln Ala Ala Thr Ala Asp Val Asn  
 595 600 605

Thr Gln Gly Val Leu Pro Gly Met Val Trp Gln Asp Arg Asp Val Tyr  
 610 615 620

Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His Thr Asp Gly His Phe  
 625 630 635 640

His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu Lys His Pro Pro Pro  
 645 650 655

Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala Asn Pro Ser Thr Thr  
 660 665 670

Phe Ser Ala Ala Lys Phe Ala Ser Phe Ile Thr Gln Tyr Ser Thr Gly  
 675 680 685

Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys  
 690 695 700

Arg Trp Asn Pro Glu Ile Gln Phe Thr Ser Asn Tyr Asn Lys Ser Val  
 705 710 715 720

Asn Val Asp Phe Thr Val Asp Thr Asn Gly Val Tyr Ser Glu Pro Arg  
 725 730 735

Pro Ile Gly Thr Arg Phe Leu Thr Arg Asn Leu

740 745

<210> 5

<211> 747

<212> PRT

<213> Artificial Sequence

<220><223> AAV2 capsid protein VP1 with insertion

<400> 5

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

1 5 10 15  
Glu Gly Ile Arg Gln Trp Trp Lys Leu Lys Pro Gly Pro Pro Pro Pro

20 25 30  
Lys Pro Ala Glu Arg His Lys Asp Asp Ser Arg Gly Leu Val Leu Pro

35 40 45  
Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro

50 55 60  
Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp

65 70 75 80  
Arg Gln Leu Asp Ser Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala

85 90 95  
Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly

100 105 110  
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro

115 120 125  
Leu Gly Leu Val Glu Glu Pro Val Lys Thr Ala Pro Gly Lys Lys Arg

130 135 140  
Pro Val Glu His Ser Pro Val Glu Pro Asp Ser Ser Ser Gly Thr Gly

145 150 155 160  
Lys Ala Gly Gln Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln Thr

165 170 175  
Gly Asp Ala Asp Ser Val Pro Asp Pro Gln Pro Leu Gly Gln Pro Pro

180 185 190

Ala Ala Pro Ser Gly Leu Gly Thr Asn Thr Met Ala Thr Gly Ser Gly  
 195 200 205  
 Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser  
 210 215 220  
 Ser Gly Asn Trp His Cys Asp Ser Thr Trp Met Gly Asp Arg Val Ile  
 225 230 235 240  
 Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Phe Asn Asn His Leu  
 245 250 255  
 Tyr Lys Gln Ile Ser Ser Gln Ser Gly Ala Ser Asn Asp Asn His Phe  
 260 265 270  
 Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His  
 275 280 285  
 Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp  
 290 295 300  
 Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile Gln Val  
 305 310 315 320  
 Lys Glu Val Thr Gln Asn Asp Gly Thr Thr Thr Ile Ala Asn Asn Leu  
 325 330 335  
 Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr  
 340 345 350  
 Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp  
 355 360 365  
 Val Phe Met Val Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser  
 370 375 380  
 Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser  
 385 390 395 400  
 Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe Glu  
 405 410 415  
 Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg  
 420 425 430  
 Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Phe Leu Ser Arg Thr



Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys  
 690 695 700

Arg Trp Asn Pro Glu Ile Gln Phe Thr Ser Asn Tyr Asn Lys Ser Val

705 710 715 720

Asn Val Asp Phe Thr Val Asp Thr Asn Gly Val Tyr Ser Glu Pro Arg  
 725 730 735

Pro Ile Gly Thr Arg Phe Leu Thr Arg Asn Leu  
 740 745

<210> 6  
 <211> 747  
 <212> PRT  
 <213> Artificial Sequence  
 <220><223> AAV2 capsid protein VP1 with insertion  
 <400> 6

Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Thr Leu Ser  
 1 5 10 15

Glu Gly Ile Arg Gln Trp Trp Lys Leu Lys Pro Gly Pro Pro Pro Pro  
 20 25 30

Lys Pro Ala Glu Arg His Lys Asp Asp Ser Arg Gly Leu Val Leu Pro  
 35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro  
 50 55 60

Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp  
 65 70 75 80

Arg Gln Leu Asp Ser Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala  
 85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly  
 100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro  
 115 120 125

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



Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser  
 385 390 395 400

Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe Glu  
 405 410 415

Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg  
 420 425 430

Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Phe Leu Ser Arg Thr  
 435 440 445

Asn Thr Pro Ser Gly Thr Thr Thr Gln Ser Arg Leu Gln Phe Ser Gln  
 450 455 460

Ala Gly Ala Ser Asp Ile Arg Asp Gln Ser Arg Asn Trp Leu Pro Gly  
 465 470 475 480

Pro Cys Tyr Arg Gln Gln Arg Val Ser Lys Thr Ser Ala Asp Asn Asn  
 485 490 495

Asn Ser Glu Phe Ser Trp Thr Gly Ala Thr Lys Tyr His Leu Asn Gly  
 500 505 510

Arg Asp Ser Leu Val Asn Pro Gly Pro Ala Met Ala Ser His Lys Asp  
 515 520 525

Asp Glu Glu Lys Phe Phe Pro Gln Ser Gly Val Leu Ile Phe Gly Lys  
 530 535 540

Gln Gly Ser Glu Lys Thr Asn Val Asp Ile Glu Lys Val Met Ile Thr  
 545 550 555 560

Asp Glu Glu Glu Ile Arg Thr Thr Asn Pro Val Ala Thr Glu Gln Tyr  
 565 570 575

Gly Ser Val Ser Thr Asn Leu Gln Arg Gly Asn Ala Ala Ala Thr Pro  
 580 585 590

Pro Ser Ile Thr Ala Ala Ala Arg Gln Ala Ala Thr Ala Asp Val Asn  
 595 600 605

Thr Gln Gly Val Leu Pro Gly Met Val Trp Gln Asp Arg Asp Val Tyr  
 610 615 620

Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His Thr Asp Gly His Phe

625                      630                      635                      640  
 His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu Lys His Pro Pro Pro  
                                  645                      650                      655  
  
 Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala Asn Pro Ser Thr Thr  
                                  660                      665                      670  
 Phe Ser Ala Ala Lys Phe Ala Ser Phe Ile Thr Gln Tyr Ser Thr Gly  
                                  675                      680                      685  
 Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys  
                                  690                      695                      700  
 Arg Trp Asn Pro Glu Ile Gln Phe Thr Ser Asn Tyr Asn Lys Ser Val  
 705                      710                      715                      720  
  
 Asn Val Asp Phe Thr Val Asp Thr Asn Gly Val Tyr Ser Glu Pro Arg  
                                  725                      730                      735  
 Pro Ile Gly Thr Arg Phe Leu Thr Arg Asn Leu  
                                  740                      745  
  
 <210> 7  
 <211> 747  
 <212> PRT  
 <213> Artificial Sequence  
 <220><223> AAV2 capsid protein VP1 with insertion  
 <400> 7  
 Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Thr Leu Ser  
 1                      5                      10                      15  
 Glu Gly Ile Arg Gln Trp Trp Lys Leu Lys Pro Gly Pro Pro Pro Pro  
  
                                  20                      25                      30  
 Lys Pro Ala Glu Arg His Lys Asp Asp Ser Arg Gly Leu Val Leu Pro  
                                  35                      40                      45  
 Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro  
                                  50                      55                      60  
 Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp  
 65                      70                      75                      80

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



Gly Ser Val Ser Thr Asn Leu Gln Arg Gly Asn Ala Ala Ala Cys Gly

580 585 590

Glu Ser Ser Tyr Leu Ala Ala Arg Gln Ala Ala Thr Ala Asp Val Asn

595 600 605

Thr Gln Gly Val Leu Pro Gly Met Val Trp Gln Asp Arg Asp Val Tyr

610 615 620

Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His Thr Asp Gly His Phe

625 630 635 640

His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu Lys His Pro Pro Pro

645 650 655

Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala Asn Pro Ser Thr Thr

660 665 670

Phe Ser Ala Ala Lys Phe Ala Ser Phe Ile Thr Gln Tyr Ser Thr Gly

675 680 685

Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys

690 695 700

Arg Trp Asn Pro Glu Ile Gln Phe Thr Ser Asn Tyr Asn Lys Ser Val

705 710 715 720

Asn Val Asp Phe Thr Val Asp Thr Asn Gly Val Tyr Ser Glu Pro Arg

725 730 735

Pro Ile Gly Thr Arg Phe Leu Thr Arg Asn Leu

740 745

<210> 8

<211> 747

<212> PRT

<213> Artificial Sequence

<220><223> AAV2 capsid protein VP1 with insertion

<400> 8

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

1 5 10 15

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



Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His  
 275 280 285  
  
 Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp  
 290 295 300  
 Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile Gln Val  
 305 310 315 320  
 Lys Glu Val Thr Gln Asn Asp Gly Thr Thr Thr Ile Ala Asn Asn Leu  
 325 330 335  
 Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr  
 340 345 350  
  
 Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp  
 355 360 365  
 Val Phe Met Val Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser  
 370 375 380  
 Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser  
 385 390 395 400  
 Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe Glu  
 405 410 415  
  
 Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg  
 420 425 430  
 Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Phe Leu Ser Arg Thr  
 435 440 445  
 Asn Thr Pro Ser Gly Thr Thr Thr Gln Ser Arg Leu Gln Phe Ser Gln  
 450 455 460  
 Ala Gly Ala Ser Asp Ile Arg Asp Gln Ser Arg Asn Trp Leu Pro Gly  
 465 470 475 480  
  
 Pro Cys Tyr Arg Gln Gln Arg Val Ser Lys Thr Ser Ala Asp Asn Asn  
 485 490 495  
 Asn Ser Glu Phe Ser Trp Thr Gly Ala Thr Lys Tyr His Leu Asn Gly  
 500 505 510  
 Arg Asp Ser Leu Val Asn Pro Gly Pro Ala Met Ala Ser His Lys Asp



<212> PRT

<213> Artificial Sequence

<220><223> AAV2 capsid protein VP1 with insertion

<400> 9

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

Ser Gly Asn Trp His Cys Asp Ser Thr Trp Met Gly Asp Arg Val Ile

225                            230                            235                            240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Phe Asn Asn His Leu

   245                            250                            255

Tyr Lys Gln Ile Ser Ser Gln Ser Gly Ala Ser Asn Asp Asn His Phe

   260                            265                            270

Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His

   275                            280                            285

Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp

   290                            295                            300

Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile Gln Val

305                            310                            315                            320

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

   325                            330                            335

Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr

   340                            345                            350

Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp

   355                            360                            365

Val Phe Met Val Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser

   370                            375                            380

Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser

385                            390                            395                            400

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

   405                            410                            415

Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg

   420                            425                            430

Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Phe Leu Ser Arg Thr

   435                            440                            445

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

   450                            455                            460

Ala Gly Ala Ser Asp Ile Arg Asp Gln Ser Arg Asn Trp Leu Pro Gly

465                      470                      475                      480  
 Pro Cys Tyr Arg Gln Gln Arg Val Ser Lys Thr Ser Ala Asp Asn Asn  
  
                                  485                      490                      495  
 Asn Ser Glu Phe Ser Trp Thr Gly Ala Thr Lys Tyr His Leu Asn Gly  
                                  500                      505                      510  
 Arg Asp Ser Leu Val Asn Pro Gly Pro Ala Met Ala Ser His Lys Asp  
                                  515                      520                      525  
 Asp Glu Glu Lys Phe Phe Pro Gln Ser Gly Val Leu Ile Phe Gly Lys  
                                  530                      535                      540  
 Gln Gly Ser Glu Lys Thr Asn Val Asp Ile Glu Lys Val Met Ile Thr  
  
 545                      550                      555                      560  
 Asp Glu Glu Glu Ile Arg Thr Thr Asn Pro Val Ala Thr Glu Gln Tyr  
                                  565                      570                      575  
 Gly Ser Val Ser Thr Asn Leu Gln Arg Gly Asn Ala Ala Ala Glu Leu  
                                  580                      585                      590  
 Cys Asp Gly Phe Ala Ala Ala Arg Gln Ala Ala Thr Ala Asp Val Asn  
                                  595                      600                      605  
 Thr Gln Gly Val Leu Pro Gly Met Val Trp Gln Asp Arg Asp Val Tyr  
  
                                  610                      615                      620  
 Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His Thr Asp Gly His Phe  
 625                      630                      635                      640  
 His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu Lys His Pro Pro Pro  
                                  645                      650                      655  
 Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala Asn Pro Ser Thr Thr  
                                  660                      665                      670  
 Phe Ser Ala Ala Lys Phe Ala Ser Phe Ile Thr Gln Tyr Ser Thr Gly  
  
                                  675                      680                      685  
 Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys  
                                  690                      695                      700  
 Arg Trp Asn Pro Glu Ile Gln Phe Thr Ser Asn Tyr Asn Lys Ser Val  
 705                      710                      715                      720

Asn Val Asp Phe Thr Val Asp Thr Asn Gly Val Tyr Ser Glu Pro Arg

725

730

735

Pro Ile Gly Thr Arg Phe Leu Thr Arg Asn Leu

740

745

<210> 10

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> heptamer insert

<400> 10

Ser Ala Ser Glu Ala Ser Thr

1

5

<210> 11

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> heptamer insert

<400> 11

Asp Thr Arg Pro His Asp Gln

1

5

<210> 12

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> heptamer insert

<400> 12

Glu His Tyr Asn Ser Thr Cys

1

5

<210> 13

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> heptamer insert

<400> 13

Pro Asn Pro Asn Cys Thr Leu

1                   5

<210> 14

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> heptamer insert

<400> 14

Thr Pro Pro Ser Ile Thr Ala

1                   5

<210> 15

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> heptamer insert

<400> 15

Cys Gly Glu Ser Ser Tyr Leu

1                   5

<210> 16

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> heptamer insert

<400> 16

Pro Arg Thr Pro His Thr Ala

1                   5

<210> 17

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> heptamer insert

<400> 17

Glu Leu Cys Asp Gly Phe Ala

1                    5  
<210> 18  
<211> 7  
<212> PRT  
<213> Artificial Sequence  
<220><223> heptamer insert  
<400> 18

His His Asn Asp Arg Ala Pro

1                    5  
<210> 19  
<211> 7  
<212> PRT

<213> Artificial Sequence  
<220><223> heptamer insert  
<400> 19

Asp Gly Asn Leu His Lys Ser

1                    5  
<210> 20  
<211> 7  
<212> PRT

<213> Artificial Sequence  
<220><223> heptamer insert  
<400> 20

His Ser Asp Ser Ser Lys Pro

1                    5  
<210> 21  
<211> 7  
<212> PRT

<213> Artificial Sequence  
<220><223> heptamer insert  
<400> 21

Thr Leu Thr Gly Leu Cys Pro

1                    5  
<210> 22

<211> 7  
 <212> PRT  
 <213> Artificial Sequence

<220><223> heptamer insert  
 <400> 22  
 Asn His Ala Pro Asn His Cys  
 1                    5

<210> 23  
 <211> 7  
 <212> PRT  
 <213> Artificial Sequence  
 <220><223> heptamer insert  
 <400> 23  
 Asp Arg Arg Pro Gly Ile Cys  
 1                    5

<210> 24  
 <211> 7  
 <212> PRT  
 <213> Artificial Sequence  
 <220><223> heptamer insert  
 <400> 24  
 Val Asp Ala Thr Cys Met Gln  
 1                    5

<210> 25  
 <211> 54  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223>  
 > single-stranded randomized 7mer pool with NNB codons  
 <220><221> misc\_feature  
 <222> (14)..(15)  
 <223> n is a, c, g, or t  
 <220><221> misc\_feature  
 <222> (17)..(18)

<223> n is a, c, g, or t  
 <220><221> misc\_feature  
 <222> (20)..(21)  
 <223> n is a, c, g, or t  
 <220><221> misc\_feature  
 <222> (23)..(24)  
 <223> n is a, c, g, or t  
 <220><221> misc\_feature  
 <222> (26)..(27)  
 <223> n is a, c, g, or t  
 <220><221> misc\_feature  
 <222> (29)..(30)  
 <223> n is a, c, g, or t  
 <220>  
 <221> misc\_feature  
 <222> (32)..(33)  
 <223> n is a, c, g, or t  
 <400> 25  
 ttggcgcgcc gcvnvnnvn nvnnvnnvn vnnggcggcc gctttttcc ttga 54  
 <210> 26  
 <211> 15  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> antisense primer  
 <400> 26  
 ctcaaggaaa aaagc 15  
 <210> 27  
 <211> 2208  
 <212> DNA  
 <213> adeno-associated virus 2  
 <400> 27  
 atggctgccg atggttatct tccagattgg ctcgaggaca ctctctctga aggaataaga 60  
 cagtgggtgga agctcaaacc tggcccacca ccaccaaagc cgcgagagcg gcataaggac 120  
 gacagcaggg gtcttgtgct tctgggtac aagtacctcg gacccttcaa cggactcgac 180

aagggagagc cgg1caacga ggcagacgcc gcggccctcg agcacgacaa agcctacgac 240  
cggcagctcg acagcggaga caacccttac ctcaagtaca accacgccga cgcggagttt 300  
caggagcggc ttaaagaaga tacgtctttt gggggcaacc tcggacgagc agtcttccag 360  
gcgaaaaaga ggg1tcttga acctctgggc ctgg1tgagg aacctgttaa gacggctccg 420  
ggaaaaaaga ggccggtaga gcactctcct gtggagccag actcctctc gggaaccgga 480  
aaggcggggc agcagcctgc aagaaaaaga ttgaattttg gtcagactgg agacgcagac 540  
  
tcagtacctg acccccagcc tctcggacag ccaccagcag cccctctgg tctgggaact 600  
aatacgatgg ctacaggcag tggcgcacca atggcagaca ataacgaggg cgcgcagga 660  
gtgggtaatt cctccgaaa ttggcattgc gattccacat ggatgggca cagagtcac 720  
accaccagca cccgaacctg ggccctgcc accttcaaca accacctgta caaacaatt 780  
tcagccaat caggagcttc gaacgacaat cacttctttg gctacagcac cccttggggg 840  
tattttgact tcaacagatt ccactgccac ttttaccac gtgactggca aagatcatc 900  
aacaacaact ggggattccg acccaagaga ctcaacttca agctctttaa cattcaagtc 960  
  
aaagggtca cgcagaatga cggtacgacg acgattgcca ataaccttac cagcacggtt 1020  
caggtgttta ctgactcggg gtaccagctc ccgtacgtcc tcggctcggc gcatcaagga 1080  
tgctcccgc cgttcccagc agacgtcttc atgg1gccac agtatggata cctcacctg 1140  
aacaacggga gtcaggcagt aggacgctct tcattttact gcctggagta ctttcttct 1200  
cagatgctgc gtaccgaaa caactttacc ttcagctaca cttttgagga cgttccttc 1260  
cacagcagct acgtcacag ccagagtctg gaccgtctca tgaatcctct catcgaccag 1320  
tacctgtatt tcttaagcag aaaaaaact ccaagtggaa ccaccagca gtcaaggctt 1380  
  
cagttttctc aggccggagc gagtgcatt cgggaccagt ctaggaactg gcttctgga 1440  
ccctgttacc gccagcagcg agtatcaaag acatctcggg ataacaaca cagtgaattc 1500  
tcgtggaccg gtgctacaa gtaccacctc aatggcagag actctctggt gaatccgggc 1560  
cggccatgg caagccaca ggacgatgaa gaaaagt111 ttcctcagag cgggg1tctc 1620  
atctttggga agcaaggctc agagaaaaca aatgtggaca ttgaaaaggt catgattaca 1680  
gacgaagagg aaatcaggac aaccaatcc gtggctacgg agcagtatgg ttctgtatct 1740  
accaactcc agagaggcaa cagacaagca gctaccgag atgtcaacac acaaggcgtt 1800  
  
cttccaggca tgg1ctgga ggacagagat gtgtacctc aggggccc11 ctgggcaaag 1860  
attccacaca cggacggaca ttttacc111 tctccc111ca tggg1ggatt cggact111aa 1920  
cacc111ctc cacagattct catcaagaac acccgg111tac ctgcgaatcc ttcgaccacc 1980  
ttcag111cgg caaag111tgc ttcct111catc acacag111tact ccacgggaca ggtcagcgtg 2040

gagatcgagt gggagctgca gaaggaaaac agcaaacgct ggaatcccga aattcagttc 2100  
acttcgaact acaacaagtc tgттаatgtg gactttactg tggacactaa tggcgtgtat 2160  
tcagagcctc gcccattgg taccagattc ctgactcgta atctgtaa 2208

<210> 28

<211> 12

<212> PRT

<213> Artificial Sequence

<220><223> alanin spacer

<400> 28

Ala Ala Ala Ser Ala Ser Glu Ala Ser Thr Ala Ala

1                    5                    10