



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2017년03월27일

(11) 등록번호 10-1720063

(24) 등록일자 2017년03월21일

(51) 국제특허분류(Int. Cl.)

C12N 15/67 (2006.01) A61K 48/00 (2006.01)

C07K 14/435 (2006.01) C12N 15/11 (2006.01)

C12N 15/65 (2006.01)

(21) 출원번호 10-2013-7005831

(22) 출원일자(국제) 2011년08월12일

심사청구일자 2015년09월25일

(85) 번역문제출일자 2013년03월06일

(65) 공개번호 10-2013-0108282

(43) 공개일자 2013년10월02일

(86) 국제출원번호 PCT/EP2011/004077

(87) 국제공개번호 WO 2012/019780

국제공개일자 2012년02월16일

(30) 우선권주장

PCT/EP2010/004998 2010년08월13일 세계지적재

산권기구(WIPO)(WO)

(56) 선행기술조사문헌

US20070172949 A9\*

\*는 심사관에 의하여 인용된 문헌

(73) 특허권자

큐어백 아게

독일, 72076 뒤빙엔, 폴 에르리츠 스트라쎄 15

(72) 발명자

테스, 안드레아스

독일, 72127 쿠스테르딘겐, 하인든부르크스트라쎄 9

스크라케, 토마스

독일, 79194 군델핀겐, 가르텐웨그 1

프롭스트, 조챈

독일, 72649 볼프쉬루겐, 인 텐 키르쵸프란데른 19

(74) 대리인

특허법인필엔온지

전체 청구항 수 : 총 18 항

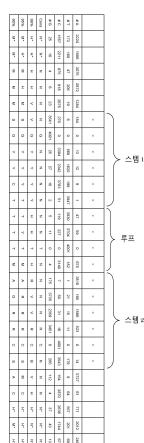
심사관 : 김현태

(54) 발명의 명칭 암호화된 단백질의 발현을 증가시키기 위한 히스톤 스템-루프 및 폴리(A) 서열 또는 폴리아데닐레이션 신호를 포함하거나 코딩하는 핵산

## (57) 요 약

본 출원은 암호화된 단백질의 발현 증가를 위한, 히스톤 스템-루프 및 폴리(A) 서열 또는 폴리아데닐레이션 신호를 포함하는, 코딩 핵산 서열, 특히 메신저 RNA(mRNA) 및 이들의 용도에 관한 것이다. 또한 예를 들어 종양 및 암 질환, 심혈관 질환, 전염성 질환, 자가면역 질환 또는 유전자적 질환의 치료 또는 유전자 치료법에 사용을 위한, 약학적 조성물, 특히 백신의 제조를 위한 그것의 용도를 개시한다. 본 발명은 또한, 시험관 내 전사 방법, 히스톤 스템-루프 및 폴리(A) 서열 또는 폴리아데닐레이션 신호를 포함하거나 코딩하는 핵산을 이용하는 단백질의 발현을 증가시키기 위한 시험관 내 방법 및 생체 밖 및 생체 내 방법에 대한 것이다.

대 표 도 - 도1



## 명세서

### 청구범위

#### 청구항 1

다음을 포함하거나 코딩하는 핵산을 포함하는, 암 또는 종양 질환, 전염성 질환, 자가면역 질환, 알러지 또는 알레르기성 질환, 일원발생적 질환, 심혈관 질환, 뉴런의 질환, 호흡계 질환, 소화계 질환, 피부 질환, 근골격 질환, 결합 조직 질환, 종양, 면역 결핍, 내분비, 영양 및 대사 질환, 안 질환, 또는 귀 질환의 치료 또는 예방 용 약학적 조성물: \_

a) 웨타이드 또는 단백질을 암호화하는 코딩 영역;

여기서, 상기 코딩 영역이 히스톤 단백질, EGFP 및 루시페라아제로부터 선택된 리포터 단백질, 그리고 알파-글로빈, 갈락토키나아제 및 잔틴 구아닌 포스포리보실 전이효소(GPT)로부터 선택된 마커 또는 선별 단백질을 코딩 하지 않음,

b) 적어도 하나의 히스톤 스템-루프; 및

c) 폴리(A) 서열 또는 폴리아데닐레이션 신호.

#### 청구항 2

제1항에 있어서, 상기 전염성 질환은 바이러스성, 박테리아성 또는 원생동물성 전염성 질환인 약학적 조성물.

#### 청구항 3

제1항에 있어서, 상기 핵산은 자가-스플라이싱 리보자임을 포함하는 리보자임을 암호화하는 서열, 바이러스성 핵산 서열, 생쥐 히스톤 H2A614 유전자로부터 유래된 히스톤-스템 루프 가공 서열을 포함하는 히스톤 스템-루프 가공 신호, 네오 유전자, 불활성화된 프로모터 서열 및 불활성화된 인핸서 서열로 이루어진 그룹의 구성성분들 중 적어도 하나를 포함하지 않는 약학적 조성물.

#### 청구항 4

제1항에 있어서, 상기 핵산은

자가-스플라이싱 리보자임을 포함하는 리보자임, 그리고

네오 유전자, 불활성화된 프로모터 서열, 불활성화된 인핸서 서열, 및 생쥐 히스톤 H2A614 유전자로부터 유래된 히스톤-스템 루프 가공 서열을 포함하는 히스톤 스템-루프 가공 신호로 이루어진 그룹의 하나를 포함하지 않는 약학적 조성물.

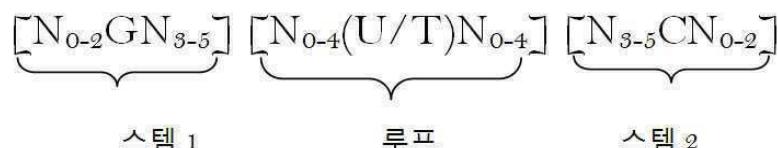
#### 청구항 5

제1항에 있어서, 상기 핵산은 mRNA를 포함하는 RNA인 약학적 조성물.

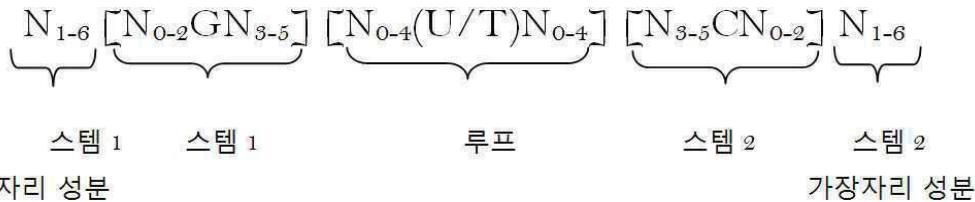
#### 청구항 6

제1항에 있어서, 상기 적어도 하나의 히스톤 스템-루프는 다음 식(I) 또는 (II)로부터 선택되는 약학적 조성물:

식 (I) (스템 가장자리 성분이 없는 스템-루프 서열):



식 (II) (스텝 가장자리 성분이 있는 스템-루프 서열):



상기:

스텝1 또는 스템2 가장자리 성분  $N_{1-6}$  은 1 내지 6N의 연속적인 서열이며, 상기 각각의 N은 A, U, T, G 및 C로부터 선택된 뉴클레오티드로부터 서로 독립적으로 선택되며;

스텝1  $[N_{0-2}GN_{3-5}]$  은 성분 스템2와 역으로 상보적으로 또는 부분적으로 역으로 상보적이며, 그리고 5 내지 7 뉴클레오티드 사이의 연속적인 서열이고;

상기  $N_{0-2}$  는 0 내지 2N이며, 상기 각각의 N은 A, U, T, G 및 C로부터 선택된 뉴클레오티드로부터 서로 독립적으로 선택되며;

상기  $N_{3-5}$  는 3 내지 5N이며, 상기 각각의 N은 A, U, T, G 및 C으로부터 선택된 뉴클레오티드로부터 서로 독립적으로 선택되며, 그리고

상기 G는 구아노신, 또는 스템2에서 그것의 상보적인 뉴클레오티드 시티딘이 구아노신으로 치환되면, 시티딘 또는 이의 유사체로 치환되고;

루프 서열  $[N_{0-4}(U/T)N_{0-4}]$  은 성분 스템1과 스템2 사이에 위치하며 3 내지 5 뉴클레오티드의 연속적인 서열이고;

상기 각각의  $N_{0-4}$  은 0 내지 4N의 다른 연속적인 서열로부터 독립적으로 선택되며, 상기 각각의 N은 A, U, T, G 및 C로부터 선택된 뉴클레오티드로부터 서로 독립적으로 선택되며; 그리고

상기 U/T는 우리딘, 또는 티미딘을 나타내며;

스텝2  $[N_{3-5}CN_{0-2}]$  은 성분 스템1과 역으로 상보적인 또는 부분적으로 역으로 상보적이며, 그리고 5 내지 7 뉴클레오티드의 연속적인 서열이고;

상기  $N_{3-5}$  은 3 내지 5N의 연속적인 서열이며, 상기 각각의 N은 A, U, T, G 및 C로부터 선택된 뉴클레오티드로부터 서로 독립적으로 선택되며;

상기  $N_{0-2}$  는 0 내지 2N의 연속적인 서열이며, 상기 각각의 N은 A, U, T, G 또는 C로부터 선택된 뉴클레오티드로부터 서로 독립적으로 선택되고; 그리고

상기 C는 시티딘이며, 또는 스템1에서 그것의 상보적인 뉴클레오티드 구아노신이 시티딘으로 치환되면, 구아노신으로 치환되며;

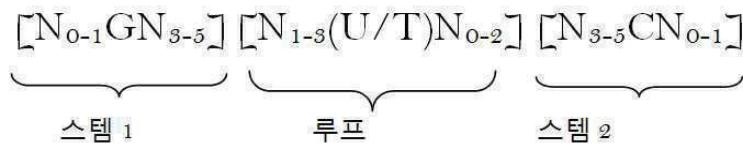
상기

스텝1 및 스템2는 역의 상보 서열을 형성하면서, 서로 서로와 염기 짹짓기(base paring)를 할 수 있고, 상기 염기 짹짓기는 스템1 및 스템2 사이에서 일어날 수 있으며, 또는

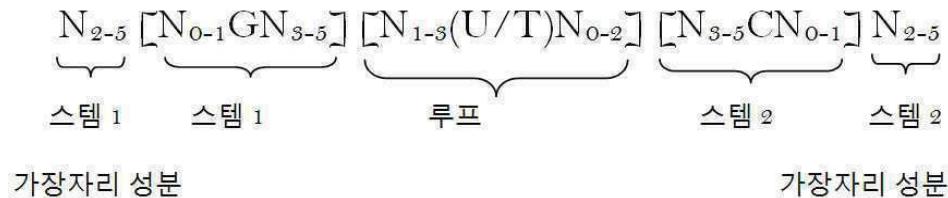
부분적으로 역의 상보 서열을 형성하면서, 서로 서로와 염기 짹짓기를 할 수 있고, 불완전한 염기 짹짓기는 스템1 및 스템2 사이에서 일어날 수 있다.

## 청구항 7

제6항에 있어서, 상기 적어도 하나의 히스톤 스템-루프는 다음의 식 (Ia) 또는 (IIa)로부터 선택되는 약학적 조성물:



식 (Ia) (스텝 가장자리 성분을 갖지 않는 스텝-루프 서열)



### 식 (IIa) (스텝

## 가장자리 성분

적 (IIa) (스템 가장자리 정준을 갖는 스템-루프 서열)

## 정구항 8

제1항에 있어서, 상기 풀리(A) 서열은 25 대지 400 아데노신 뉴클레오티드 서열을 포함하는 약학적 조성을.

청구항 9

제1항에 있어서, 상기 풀리(A) 서열은 60 내지 250 아데노신 뉴클레오티드 서열을 포함하는 약학적 조성물.

청구항 10

제1항에 있어서, 상기 폴리아데닐레이션 신호는 공통서열 NNUANA을 포함하는 약학적 조성물.

청구항 11

제1항에 있어서, 상기 폴리아데닐레이션 신호는 공통서열 AAUAAA 또는 AUUAAA을 포함하는 약학적 조성물.

## 청구항 12

제1항에 있어서, 상기 핵산은 치료적으로 활성인 단백질 또는 펩타이드, 애쥬반트 단백질, 항원, 종양 항원, 병원성 항원, 동물 항원, 바이러스성 항원, 원생동물성 항원, 박테리아성 항원, 알레르기성 항원, 자가면역 항원, 알레르겐, 항체, 면역자극성 단백질 또는 펩타이드 또는 항원-특이적 T-세포 수용체를 암호화하는 약학적 조성물.

청구항 13

제1항에 있어서, 상기 해산은 모노시스트론성, 바이시스트론성 또는 멀티시스트론성 해산이 약학적 조성물.

청구항 14

제1항에 있어서, 상기 핵산은 암호화된 펩타이드 또는 단백질의 발현 수준을 증가시키기 위한 것인 약학적 조성물

첨구학 15

제1학기 입학선서서에서 산기 약학전 조선물을 약학전으로 혼용가능한 단체를 포함하는 약학전 조선물

첨국한 16

한의원은 컴퓨터화하거나 한의원은 포함하는 다음은 포함하는 컴퓨터화하거나 컴퓨터화하는 행사를 포함하는 백신:

8) 페터인도 또는 단백질은 와우 히치는 코더 영역:

여기서, 상기 코딩 영역이 허스톤 단백질, EGFP 및 루시페라아제로부터 선택된 리포터 단백질, 그리고 알파-글로빈, 갈락토키나아제 및 잔틴 구아닌 포스포리보실 전이효소(GPT)로부터 선택된 마커 또는 선별 단백질을 코딩 하지 않은

- b) 적어도 하나의 히스톤 스템-루프; 및
- c) 폴리(A) 서열 또는 폴리아데닐레이션 신호.

### 청구항 17

제16항에 있어서, 상기 항원은 종양 항원, 알레르기성 항원, 자가-면역 자가-항원, 병원성 항원, 바이러스성 항원, 박테리아성 항원, 곰팡이성 항원, 원생동물성 항원, 동물 항원 또는 알레르기성 항원으로부터 선택되는 백신.

### 청구항 18

다음 단계를 포함하는 단백질 발현의 증가를 위한 시험관 내 또는 생체 외 방법:

- 1) 다음을 포함하거나 코딩하는 핵산을 준비하는 단계,

- a) 펩타이드 또는 단백질을 암호화하는 코딩 영역;

여기서, 상기 코딩 영역이 히스톤 단백질, EGFP 및 루시페라아제로부터 선택된 리포터 단백질, 그리고 알파-글로빈, 갈락토키나아제 및 잔틴 구아닌 포스포리보실 전이효소(GPT)로부터 선택된 마커 또는 선별 단백질을 코딩하지 않음,

- b) 적어도 하나의 히스톤 스템-루프; 및

- c) 폴리(A) 서열 또는 폴리아데닐레이션 신호

- 2) 무세포 발현계, 세포, 조직 또는 유기체에 핵산을 적용하거나 투여하는 단계.

### 발명의 설명

#### 기술 분야

[0001]

본 발명은 암호화된 단백질의 발현을 증가시키기 위한 코딩 핵산 서열, 특히 히스톤 스템-루프 및 폴리(A) 서열 또는 폴리아데닐레이션 신호를 포함하거나 코딩하는 메신저(messenger) RNA (mRNA) 및 이들의 용도에 관한 것이다. 또한 본 발명은, 예를 들어 종양 및 암 질환, 심혈관 질환, 감염성 질환, 자가면역(autoimmune) 질환 또는 유전자(genetic) 질환 또는 유전자 치료법에서 사용을 위한 것과 같은 약학적 조성물, 특히 백신의 제조를 위한 그것의 용도를 개시한다. 본 발명은 또한 시험관 내 전사(transcription) 방법, 시험관 내 히스톤 스템-루프 및 폴리(A) 서열 또는 폴리아데닐레이션 신호를 포함하거나 코딩하는 핵산을 이용한 단백질의 발현을 증가시키기 위한 방법 및 생체 내 방법을 개시한다.

#### 배경 기술

[0002]

심혈관 질환 및 감염성 질환을 제외하고, 종양 및 암 질환의 존재는 현대 사회에서 사망을 유발하는 가장 흔한 질환 중 하나이고 대부분의 경우에, 치료 및 후속 회복 조치 기간 동안에 상당히 많은 비용이 요구된다. 종양 및 암 질환의 치료는 예를 들어, 발생한 종양의 타입, 나이, 치료가 필요한 환자의 암 세포의 분포 등과 같은 것에 상당히 의존한다. 암 치료는 요즘 일반적으로 방사선 치료 또는 화학적 요법뿐만 아니라 수술적 방법의 사용을 통해서 수행된다. 그러나, 일반적으로 그러한 전형적인 치료법들은 면역 시스템에 예기치 못한 스트레스를 주고 오직 제한적인 범위에서 몇몇의 경우에만 사용될 수 있다. 게다가, 이러한 전형적인 치료법들은 면역 시스템이 재생하기 위하여 각각의 치료들 사이에 오랜 간격을 요구한다.

[0003]

그러므로, 그러한 "전형적인 치료"에 의해 면역 시스템이 영향을 받는 것을 피하거나 적어도 감소시키기 위하여 전형적인 치료뿐만 아니라 추가적인 전략들이 최근에 연구되고 있다. 한가지 그러한 추가적인 치료는 특히 유전자 치료적 접근법 또는 유전자 예방접종(genetic vaccination)을 포함하며, 이들은 이미 치료에 매우 효과적이거나 그러한 전형적인 치료법들을 도와 치료에 효과를 나타낼 수 있다는 것이 밝혀졌다.

[0004]

유전자 치료법 및 유전자 예방접종은 질병의 치료 및 예방 분야에서 이미 증명되고 일반적으로 데일리 의료행위, 특히 상기 언급된 질병들의 치료에서 큰 효과를 나타내는 분자 의학의 방법이다. 유전자 치료는 또한

다른 의학 분야, 즉 유전자 질병 분야에서 이용되며, 다시 말해 정의된 유전자 결함에 의해 야기되고 멘델의 유전 법칙에 따라 유전되는 (유전성) 질병 분야에서 이용된다. 그러한 유전자 질환을 대표하는 것으로 가장 잘 알려진 것은 그 중에서도 점액성점착증(mucoviscidosis(cystic fibrosis)) 및 겹상 적혈구 빈혈(sickle cell anaemia)이다. 두 가지 방법들, 유전자 치료 및 유전자 백신은 환자의 세포 또는 조직 내로 핵산의 도입 및 세포 또는 조직 안으로 도입되는 핵산에 의해 암호화된 정보의 후속 공정들, 즉 목적으로 하는 폴리펩타이드의 (단백질) 발현에 기초한다.

[0005] 비록 RNA는 최근 연구에서 잘 알려져 있지만, 유전자 치료 접근법에서, 일반적으로 DNA가 사용된다. 중요하게, 모든 이러한 유전자 치료 접근법에서 mRNA는 DNA, 바이러스성(viral) RNA 또는 mRNA가 사용되는지에 관계없이, 암호화된 단백질의 서열 정보를 위한 메신저로써 기능한다.

[0006] 일반적으로 RNA는 불안정한 분자로 생각된다: RNA 가수분해 효소(RNases)는 도처에 존재하고 있고 불활성화하기 어렵기로 악명이 높다. 게다가, RNA는 또한 DNA보다 화학적으로 더 변형되기 쉽다. 따라서, 진핵생물의 세포에서 mRNA의 "디폴트 상태(Default state)"는 상대적인 안정성에 의해 규정되고 특징적인 신호들이 각각의 mRNA의 분해를 촉진하는데 필요하다는 것은 아마 놀라운 것이다. 이러한 발견을 위한 주요 원인은 세포내에서 mRNA의 분해는 거의 두드러지게 핵산말단가수분해효소(exonucleases)에 의해 촉진되는 것처럼 보인다. 그러나, 진핵생물의 mRNAs 끝은 특정 말단 구조 및 그들의 관련된 단백질들에 의해 이러한 효소들로부터 보호된다: 5' 끝에 m7GpppN CAP 그리고 일반적으로 3' 끝에 폴리(A) 서열. 이를 두개의 말단 변형의 제거는 따라서 mRNA 분해 속도를 제한하는 것으로 생각된다. 비록 요소를 안정화하는 것을 알파-글로빈 mRNA의 3' UTR에서 특징지어지나, 일반적으로 진핵생물의 RNA 서열의 턴오버(turnover)에 영향을 미치는 RNA 서열은 탈아데닐레이션(deadenylation)을 촉진함으로써 항상 분해의 프로모터로 작용한다(reviewed in Meyer, S., C. Temme, *et al.* (2004), Crit Rev Biochem Mol Biol 39(4): 197-216.).

[0007] 상기 언급한 것처럼, 진핵생물의 mRNAs의 5' 끝은 일반적으로 메틸화된 CAP 구조, 예를 들어 m7GpppN을 운반하기 위하여 전사후에(posttranscriptionally) 변형된다. RNA 스플라이싱(splicing), 안정화 및 수송에서 역할을 제외하고, CAP 구조는 분명히 전사 개시 동안 mRNA의 5' 끝에 40S 리보솜 단위(ribosomal subunit)의 보충을 증가시킨다. 후자의 기능은 진핵생물의 개시 인자 복합체 eIF4F에 의한 CAP 구조의 인지를 요구한다. 추가적으로 폴리(A) 서열은 mRNA에 증가된 40S 리보솜 단위의 보충을 통하여 번역을 자극하고, 폴리(A) 결합 단백질(PABP)의 개입을 요하는 결과를 나타낸다. 결국, PABP는 최근에 CAP-결합 eIF4F 복합체의 부분인 eIF4G와 자연법칙에 따라 상호작용하는 것이 증명되었다. 따라서, 씌워지고(capped), 폴리아데닐화된 mRNAs에서 번역 개시의 단한 루프 모델이 주장되었다(Michel, Y. M., D. Poncet, *et al.* (2000), J Biol Chem 275(41): 32268-76.).

[0008] 그러한 폴리(A) 서열을 갖는 거의 모든 진핵생물의 mRNAs 끝은 도처에 존재하는 절단/폴리아데닐레이션 장치에 의하여 그들의 3' 끝에 추가된다. 3' 끝에 폴리(A) 서열의 존재는 진핵생물의 mRNAs의 가장 식별되는 특징 중 하나이다. 절단후에, 복제-의존적 히스톤 전사를 제외하고, 대부분의 프리-mRNAs는 폴리아데닐화된 꼬리(tail)를 얻는다. 이와 관련하여, 3' 가공은 핵에서 세포질로 mRNA의 수송을 향상시키는 핵의 동시전사(co-transcriptional) 과정이며 mRNAs의 안정성 및 번역에 영향을 미친다. 이러한 3' 끝의 형성은 절단/폴리아데닐화 장치에 의한 두가지 단계로 일어나고 mRNA 전구체(precursors) (pre-mRNAs)에서 두 가지 서열 요소의 존재에 의존한다; 매우 잘 보존된 핵산뉴클레오티드 AAUAAA(폴리아데닐레이션 신호) 및 다운스트림(downstream) G/U-풍부 서열. 첫 번째 단계에서, 프리-mRNAs는 이러한 두 가지 요소들 사이에서 절단된다. 첫 번째 단계와 긴밀하게 연결되어 있는 두 번째 단계에서 새롭게 형성된 3' 끝은 mRNA 수송(export), 안정화 및 번역을 포함하는 mRNA 대사의 뒤따라 일어나는 모든 측면에 영향을 미치는 200-250 아데닐산(adenylates)으로 이루어진 폴리(A) 서열의 부가에 의해서 확장된다(Dominski, Z. and W. F. Marzluff (2007), Gene 396(2): 373-90.).

[0009] 유일하게 알려진 이러한 규칙에 대한 예외는 폴리(A) 서열 대신에 말단에 히스톤 스템-루프를 가지는 복제-의존적 히스톤 mRNA이다. 전형적인 히스톤 스템-루프 서열은 Lopez *et al.*에 기술되어 있다. (Dalia Lopez, M., &

Samuelsson, T. (2008), RNA (New York, N.Y.), 14(1), 1-10. doi:10.1261/rna.782308.).

[0010] 히스톤 프리-mRNAs에서 스템-루프들은 일반적으로 히스톤 다운스트림 요소(histone downstream element (HDE))로 알려진 퓨린(purine)-풍부 서열에 의해서 뒤따른다. 이러한 프리-mRNAs는 HDE를 갖는 U7 snRNP의 염기 짹짓기(base pairing)를 통하여 U7 snRNP에 의해 촉매화되는, 스템-루프의 대략 5 뉴클레오티드 다운스트림에서 단일 내부가수분해효소 절단에 의해 핵에서 가공된다.

[0011] 염색질(chromatin)내로 새롭게 합성된 DNA를 포장하는데 필요하기 때문에, 히스톤 합성은 세포 주기와 협력하여 조절된다. S 단계 동안에 증가된 히스톤 단백질의 합성은 히스톤 mRNA 수준의 전사후 조절 뿐만 아니라 히스톤 유전자들의 전사 활성에 의해서도 이루어진다. 히스톤 스템-루프는 히스톤 발현 조절의 모든 전사후 단계에 필수적인 것으로 보여진다. 그것은 효과적인 가공, 세포질 안으로 mRNA의 수송, 폴리리보솜(polyribosomes)에 로딩 및 mRNA의 안정성 조절에 필수적이다.

[0012] 상기와 관련하여, 32 kDa 단백질은 핵 및 세포질 모두에서 히스톤 메시지(messages)의 3'- 끝에 있는 히스톤 스템-루프와 관련된 것으로 밝혀졌다. 이러한 스템-루프 결합 단백질(SLBP)의 발현 수준은 세포주기를 조절하고 히스톤 mRNA 수준이 증가될 때 S- 단계 동안에 가장 높게 나타난다. SLBP는 U7 snRNP에 의한 히스톤 프리-mRNA의 효과적인 3'-끝 가공에 필수적이다. 가공의 완료된 후에, SLBP는 성숙한 히스톤 mRNAs의 끝에 있는 스템-루프와 연관되어 존재하고, 세포질안에서 히스톤 단백질로 번역을 촉진한다. (Dominski, Z. and W. F. Marzluff (2007), Gene 396(2): 373-90). 흥미롭게도, SLBP의 RNA 결합 도메인은 후생동물(metazoa) 및 원생동물(protozoa)에서 보존되고 (Daila Lopez, M., & Samuelsson, T. (2008), RNA (New York, N.Y.), 14(1), 1-10. doi:10.1261/rna.782308) 히스톤 스템-루프 서열에 그것의 결합은 스템-루프 구조에 의존하는 것으로 보여질 것이며 최소한의 결합 부위는 스템-루프의 적어도 3개의 뉴클레오티드를 포함하는 5' 및 2개의 뉴클레오티드를 포함하는 3'을 포함할 것이다(Pandey, N. B., et al. (1994), *Molecular and Cellular Biology*, 14(3), 1709-1720 and Williams, A. S., & Marzluff, W. F., (1995), *nucleic acids Research*, 23(4), 654-662.).

[0013] 비록 히스톤 유전자는 일반적으로 히스톤 스템-루프에서 mRNA 종결을 가져다주는 "복제-의존적(replication-dependent)"으로 또는 폴리(A)-꼬리(tail)을 갖는 mRNA로 되는 "치환형(replacement-type)"으로 분류되나, 이들의 3'에 히스톤 스템-루프 및 폴리(A) 또는 올리고(A)를 모두 포함하는 자연 발생적인 mRNA는 매우 드문 경우에서 발견되었다. Sanchez et al.는 리포터 단백질로써 루시페라제를 이용한 제노포스(Xenopus) 난형성(oogenesis) 동안 히스톤 mRNA의 히스톤 스템-루프의 3'에 부가되는 자연발생적인 올리고(A) 꼬리의 효과를 시험하였고 올리고(A) 꼬리는 난형성 동안 히스톤 mRNA를 침묵하게 하는 번역 억제 메카니즘의 활성 부분이며 그 것의 제거는 히스톤 mRNAs의 번역을 활성화하는 메카니즘의 부분이라는 것을 발견하였다(Sanchez, R. and W. F. Marzluff (2004), Mol Cell Biol 24(6): 2513-25).

[0014] 게다가, 프리-mRNA 가공 및 mRNA 안정성의 정도에 복제 의존적 히스톤의 조절에 대한 필요는 마커 단백질 알파글로빈을 코딩하는 인공적인 복합체(constructs)를 이용하여 조사되었으며, 글로빈 유전자는 인트론-부족 히스톤 유전자와 대조적으로 인트론을 포함한다는 사실을 이용하였다. 이러한 목적으로 알파 글로빈 코딩 서열에서 만들어지는 복합체들 다음에 히스톤 스템-루프 신호(히스톤 스템-루프 다음에 히스톤 다운스트림 요소(downstream element)가 뒤따르는) 및 폴리아데닐레이션 신호가 뒤따른다(Whitelaw, E., et al. (1986). 핵산 Research, 14(17), 7059-7070.; Pandey, N. B., & Marzluff, W. F. (1987). Molecular and Cellular Biology, 7(12), 4557-4559.; Pandey, N. B., et al. (1990). nucleic acids Research, 18(11), 3161-3170).

[0015] 다른 관점에서 Lucher et al. 등은 재조합 히스톤 H4 유전자의 세포-주기 의존성 조절에 대해 조사하였다. 복합체는 히스톤 스템-루프 신호 및 폴리아데닐레이션 신호가 뒤따르는 H4 코딩 서열에서 만들어지고, 두 가공 신호는 갈락토키나아제 코딩 서열에 의해서 우연히 분리되었다. (Lucher, B. et al., (1985). Proc. Natl. Acad.

Sci. USA, 82(13), 4389-4393).

[0016] 추가적으로, Stauber *et al.*은 최소한의 서열이 히스톤 H4 mRNA 단계에서 세포-주기 조절을 위해 필요하다는 것을 밝혀내었다. 이러한 조사를 위해서 복합체들이 사용되었고, 이들은 폴리아데닐레이션 신호에 의해 뒤이어 나타나는 히스톤 스템-루프 신호에 앞서 선별 마커 잔틴(Xanthine):구아닌 포스포리보실 전이효소(GPT)를 코딩하는 서열을 포함한다(Stauber, C. *et al.*, (1986). EMBO J, 5(12), 3297-3303).

[0017] 히스톤 프리-mRNA 가공을 연구한 Wagner *et al.*은 EGFP는 히스톤 프리-mRNA 가공에 방해를 받는 경우에만 오직 발현되기에 히스톤 스템-루프 신호와 폴리아데닐레이션 신호 사이에 EGFP를 자리하게 하는 리포터 복합체를 사용하여 히스톤 프리-mRNA의 절단에 필요한 패터를 찾아내었다(Wagner, E. J. *et al.*, (2007). Mol Cell 28(4), 692-9).

[0018] 주목할 것은, 폴리아데닐화된 mRNA의 번역은 보통 5' CAP 근처로 근접하게 하는 3' 폴리(A) 서열을 필요로 한다는 것이다. 이것은 폴리(A) 결합 단백질과 진핵생물의 개시 패터 eIF4G 사이의 단백질-단백질 상호작용을 통하여 매개된다. 복제-의존적 히스톤 mRNAs에 대해서 유사한 메카니즘은 밝혀지지 않았다. 이와 관련하여, Gallie *et al.*은 히스톤 스템-루프는 번역 효율을 증가시키고, 그리고 충분한 번역이 이루어지도록 하기 위해 5'-CAP에 상호의존적인 경우에 폴리(A) 서열과 기능적으로 유사하다는 것을 보여준다. 그들은 히스톤 스템-루프는 형질주입한 중국 햄스터 난소세포주(Chinese hamster ovary cells)에서 리포터 mRNA의 번역을 증가시키는데 충분하고 필요하지만 가장 바람직하게 작용하기 위해서는 3'- 말단에 위치해야만 한다는 것을 보여주었다. 그러므로, 다른 mRNA의 폴리(A) 꼬리와 유사하게, 이들 히스톤 mRNA의 3' 말단은 생체 내에서 번역을 위해 필수적으로 나타나고 기능적으로 폴리(A) 꼬리와 유사하다(Gallie, D. R., Lewis, N. J., & Marzluff, W. F. (1996), nucleic acids Research, 24(10), 1954-1962).

[0019] 추가적으로, SLBP는 세포질의 히스톤 mRNA에 결합되고 그것의 번역을 위해 필요하다는 것을 보여줄 수 있었다. 비록 SLBP는 직접적으로 eIF4G와 반응하지 않지만, 히스톤 mRNA의 번역에 필요한 도메인은 최근에 알려진 단백질 SLIP1와 작용한다. 다른 단계에서, SLIP1은 eIF4G 와 상호작용하고 히스톤 mRNA를 원형화하도록 하고 폴리아데닐화된 mRNA의 번역과 유사한 메카니즘에 의해서 히스톤 mRNA의 충분한 번역을 돋도록 한다.

[0020] 상기 언급한 것처럼, 유전자 치료법은 보통 세포 안으로 코딩 정보를 전달하기 위하여 DNA를 사용하며, 그 다음, mRNA의 자연적으로 발생하는 구성요소들, 특히 암호화된 치료적 단백질의 발현을 가능하도록 하는 5'-CAP 구조와 3' 폴리(A) 서열을 갖는, mRNA로 전사된다.

[0021] 그러나, 많은 경우에서, 이러한 핵산에 의해 환자들의 세포 또는 조직 안으로 그러한 핵산의 도입 및 목적으로 하는 코딩된 폴리펩타이드의 뒤이은 발현에 근거한 발현 시스템은 DNA가 사용되었는지 RNA가 사용되었는지와 관계없이 충분한 치료가 가능하도록 하는, 목적으로 하는 또는 필요로 하는 발현 수준을 나타내지 못한다.

[0022] 선행 기술에서는, 다른 시도들이 암호화된 단백질의 발현 정도를 증가시키기 위하여 지금까지 시도들이 있으며, 특히 시험관 내 및/또는 생체 내 모두에서, 향상된 발현 시스템의 사용에 의해서 있어왔다. 이란적을 종래 기술에서 언급된 발현 증가를 위한 방법들은 특이적인 프로모터들과 대응하는 조절 요소들을 포함하는 발현 벡터 또는 카세트(cassettes)의 이용에 근거하는 것이 일반적이었다. 이러한 발현 벡터 도는 카세트들은 일반적으로 특정 세포 시스템에 제한되기 때문에, 이들 발현 시스템은 다른 세포 시스템에서 이용되기 위하여 조작이 필요하다. 그러한 조작된 발현 벡터 도는 카세트들은 그 다음 세포안으로 대개는 형질주입되고 일반적으로 특정 세포주에 의존하여 처리된다. 그러므로, 선호도는 우선적으로 세포내에서 유전되는 시스템에 의한 타겟 세포내 암호화된 단백질을 발현할 수 있는 그러한 핵산 분자들에게 주어지며, 특정 세포 타입에 특이적인 프로모터 및 조절 요소들에는 의존하지 않는다. 이와 관련하여, mRNA 안정화 요소 및 mRNA의 번역 효율을 증가시키는 요소는

구별될 수 있다.

[0023] 그들의 코딩 서열에 적합한 mRNA 및 일반적으로 그러한 목적에 적합한 mRNA 는 WO 02/098443 (CureVac GmbH) 출원에 개시되어 있다. 예를 들어, WO 02/098443는 일반적인 형태에서 안정화되고, 그들의 코딩 영역에서 전자에 최적화되는 mRNA를 개시한다. mRNA는 또한 서열 변형을 결정하는 방법을 개시한다. WO 02/098443는 추가적으로 서열의 구아닌/시토신(G/C) 함량을 증가시키기 위해 mRNA 서열내의 아데닌 및 우라실 뉴클레오티드 치환을 가능성을 개시한다. WO 02/098443에 따르면, G/C 함량 증가를 위한 치환 및 적응(adaptations)은 암 또는 전염성 질환의 치료에서 유전자 백신 뿐만 아니라 유전자 치료 어플리케이션(applications)에도 이용될 수 있다. 이와 관련하여, WO 02/098443는 일반적으로 치료가 필요한 환자에서, 예를 들어 전혀 또는 적절하지 않게 또는 결함을 가지고, 변역되는 적어도 하나의 생물학적 활성 웨პ타이드 또는 폴리웨პ타이드를 코딩하는 변형된 mRNA에서 그러한 변형을 위한 염기 서열과 같은 서열들을 언급한다. 그 대신, WO 02/098443는 그러한 변형을 위한 염기 서열로써, 종양 항원 또는 바이러스성 항원과 같은 항원들을 코딩하는 mRNA를 제안한다.

[0024] 암호화된 단백질의 발현을 증가시키기 위한 다른 접근법에서 WO 2007/036366 출원은 긴 폴리(A) 서열(특히 120bp보다 더 긴)의 긍정적인 효과 및 mRNA 안정화 및 변역 활성에서 베타 글로빈 유전자의 적어도 두 개의 3' 비변역 영역의 조합을 개시한다.

[0025] 그러나, 비록 이러한 최근의 종래 기술들이 이미 유전자 치료 접근법 및 추가적으로 향상된 mRNA 안정화 및 변역 활성을 위한 효과적인 수단들을 꽤 제공하려고 있지만, RNA-기반 어플리케이션과 DNA 백신의 일반적으로 낮은 안정성에 대한 문제와 DNA 기반 유전자 치료법에 대한 문제가 여전히 남아있다. 따라서, 여전히 생체 내에서 암호화된 단백질의 더 나은 제공을 가능하게 하는, 예를 들어, 다른 향상된 mRNA 안정성 및/또는 변역 활성을 통하여, 바람직하게는 유전자 치료를 위해, 앞서 언급된 일반적인 처치를 위한 보충적 치료법으로써 또는 유전자 치료 접근법 및 유전자 예방접종을 위한 향상된 수단을 제공하기 위한 필요가 존재한다.

## 발명의 내용

### 해결하려는 과제

[0026] 그러므로, 본 발명의 기초를 이루는 목적은 암호화된 단백질의 발현, 바람직하게는 또 다른 mRNA의 안정화 및/ 또는 치료(예를 들어, 유전자 치료법 및 유전자 예방접종)분야에서 이용을 위해 종래 알려진 그러한 핵산들에 대하여 mRNA의 전사의 효과적인 전사의 증가를 통하여, 암호화된 단백질의 전사를 증가시키기 위한 추가적인 및/또는 대체적인 방법들을 제공하는 것이다.

### 과제의 해결 수단

[0027] 이러한 목적은 특허청구범위의 기술적 사상에 의하여 해결된다. 특히, 본 발명의 기초를 이루는 목적은

[0028] a) 코딩 영역, 바람직하게 웨პ타이드 또는 단백질을 암호화하는;

[0029] b) 적어도 하나의 히스톤 스템-루프. 및

[0030] c) 선택적으로 폴리(A) 서열 또는 폴리아데닐레이션 신호,

[0031] 바람직하게 암호화된 단백질의 발현 수준을 증가시키기 위하여, 상기 암호화된 단백질은 바람직하게 히스톤 단백질이 아니고, 리포터 단백질(예를 들어, 루시페라아제, GFP, EGFP,  $\beta$ -갈اكت오시데이즈(Galactosidase), 특히 EGFP)이 아니며, 그리고 마커 또는 선별 단백질(예를 들어, 알파-글로빈, 갈락토키나아제(Galactokinase) 및 잔틴(Xanthine): 구아닌 포스포리보실 전이효소(guanine phosphoribosyl transferase (GPT)))이 아닌 단백질,

[0032] 을 포함하거나 코딩하는 독창적인 핵산 서열에 의한 첫 번째 실시예에 따라 해결된다

[0033] 이와 관련하여, 특히 바람직하게는 본 발명의 첫 번째 실시예에 따른 독창적인 핵산은 적어도 부분적으로 DNA

또는 RNA 합성에 의해 생산되며 분리된(isolated) 핵산이다.

[0034] 본 발명은, 폴리(A) 서열 또는 폴리아데닐레이션 신호 및 적어도 하나의 히스톤 스템-루프의 조합은, 비록 두 개가 자연에서는 선택적인(alternative) 메커니즘을 나타내지만, 이러한 조합은 각각의 요소들에서 관찰되는 수준이상으로 다양한 단백질 발현 증가를 나타냄으로써 상승 작용에 의해 행동한다는, 본 발명자들의 놀라운 발견에 기초한다. 폴리(A) 및 적어도 하나의 히스톤 스템-루프의 조합의 시너지 효과는 폴리(A) 및 히스톤 스템-루프의 배열과 무관하게 보여지고 폴리(A) 서열의 길이와 무관하게 보여진다.

[0035] 그러므로 특히 바람직한 것은 독창적인 핵산 분자는 a) 코딩영역, 바람직하게 웨타이드 또는 단백질을 암호화하는; b) 적어도 하나의 히스톤 스템-루프, 및 c) 폴리(A) 서열 또는 폴리아데닐레이션 서열; 바람직하게는 암호화된 단백질의 발현 수준 증가를 위하여, 상기 암호화된 단백질은 바람직하게 히스톤 단백질이 아니고, 리포터 단백질(예를 들어, 루시페라아제, GFP, EGFP,  $\beta$ -갈اكتоз이드(Galactosidase), 특히 EGFP)이 아니며, 그리고 마커 또는 선별 단백질(예를 들어, 알파-글로빈, 갈락토키나아제(Galactokinase) 및 잔틴(Xanthine): 구아닌 포스포리보실 전이효소(guanine phosphoribosyl transferase (GPT)))가 아닌 단백질을 포함하거나 코딩한다.

[0036] 본 발명의 첫 번째 실시예의 다른 측면에 따르면, 상기 독창적인 핵산은 히스톤 다운스트림 요소(HDE)를 포함하지 않는다.

[0037] 이와 관련하여, 특히 바람직한 것은 상기 독창적인 핵산은 5'에서 3' 방향으로

[0038] a) 코딩 영역, 바람직하게 웨타이드 또는 단백질을 암호화하는;

[0039] b) 적어도 하나의 히스톤 스템-루프. 선택적으로 히스톤 스템-루프에 3' 히스톤 다운스트림 요소 없이

[0040] c) 폴리(A) 서열 또는 폴리아데닐레이션 신호를 포함하거나 코딩한다.

[0041] 용어 "히스톤 다운스트림 요소(HDE)"는 성숙한 히스톤 mRNA으로 히스톤 프리-mRNA의 가공에 수반되는 U7 snRNA를 위한 결합 부위를 대표하는 자연적으로 발생한 스템-루프의 3' 뉴클레오티드의 약 15 내지 20의 퓨린-풍부 폴리뉴클레오티드 구역을 일컫는다. 예를 들어, 성계에서 HDE는 CAAGAAAGA이다 (Dominksi, Z. and W. F. Marzluff (2007), Gene 396(2): 373-90).

[0042] 게다가, 바람직하게는 본 발명의 첫 번째 실시예에 따른 독창적인 핵산은 인트론(intron)을 포함하지 않는다.

[0043] 다른 특히 더 바람직한 실시예에서, 본 발명의 첫 번째 실시예에 따른 독창적인 핵산 서열은 5'에서 3'으로

[0044] a) 코딩 영역, 바람직하게 웨타이드 또는 단백질을 암호화하는;

[0045] b) 폴리(A) 서열; 및

[0046] c) 적어도 하나의 히스톤 스템-루프를 포함하거나 코딩한다.

[0047] 본 발명의 첫 번째 실시예에 따른 독창적인 핵산 서열은 예를 들어 어느 (단일-가닥 또는 이중-가닥의)DNA, 바람직하게, 이들에 제한없이, 예를 들어, 유전체(genomic) DNA, 단일가닥 DNA 분자, 이중가닥 DNA 분자로 부터, 선택된 어느 적절한 핵산을 포함하며, 예를, 들어, 어느 PNA(웨타이드 핵산)으로부터 선택될 수 있으며 또는 예를 들어, 어느(단일가닥 또는 이중가닥) RNA, 바람직하게 메신저 RNA (mRNA); 등으로 부터 선택될 수 있다. 상기 독창적인 핵산 분자는 또한 바이러스성 RNA(vRNA)를 또한 포함할 수 있다. 그러나, 상기 독창적인 핵산 서열은 바이러스성 RNA 가 아닐 수 있거나 바이러스성 RNA 를 포함하지 않을 수 있다. 더 구체적으로, 상기 독창적인 핵산 서열은 바이러스성 서열 요소, 예를 들어 바이러스성 인핸서(enhancers) 또는 바이러스성 프로모터

(promotors)(예를 들어, 불활성화된 바이러스성 프로모터 또는 서열 요소가 아닌, 더 구체적으로 교체 전략(replacement strategies)에 의해 불화성되지 않은) 또는 다른 바이러스성 서열 요소, 또는 바이러스성 또는 레트로바이러스성(retroviral) 핵산 서열을 포함하지 않을 수 있다. 더 구체적으로, 상기 독창적인 핵산 서열은 레트로바이러스성 또는 바이러스성 벡터 또는 변형된 레트로바이러스성 또는 바이러스성 벡터가 아닐 수 있다.

[0048] 어느 경우에는. 상기 독창적인 핵산 서열은 변형될 수 있거나 그렇지 않은 또는 활성화될 수 있거나 그렇지 않은 인핸서 및/또는 프로모터 서열을 포함하거나 또는 포함하지 않을 수 있다. 상기 인핸서 및/또는 프로모터는 식물에서 발현가능하거나 발현가능하지 않을 수 있고, 그리고/또는 진핵 생물에서 발현가능하거나 발현가능하지 않을 수 있고 그리고/또는 원핵 생물에서 발현가능하거나 발현가능하지 않을 수 있다. 상기 독창적인 핵산 분자는 (셀프-스플라이싱(self-splicing))리보자임(ribozyme)을 암호화하는 서열을 포함하거나 포함하지 않을 수 있다.

[0049] 바람직하게, 상기 독창적인 핵산 분자는 RNA 이다.

[0050] 특히 본 발명의 첫 번째 실시예의 측면에서, 상기 독창적인 핵산은 시험관 내 전사에 적절한 핵산에 포함된, 특히, 적절한 시험관 내 전사 벡터에 포함된 핵산 서열이다(예를 들어, T3, T7 또는 Sp6 프로모터와 같은 시험관 내 전사를 위한 특정한 프로모터를 포함하는 플라스미드 또는 선형 핵산 서열).

[0051] 본 발명의 첫 번째 실시예의 특히 더 바람직한 측면에서, 상기 독창적인 핵산은 발현 시스템(예를 들어, 발현 벡터 또는 플라스미드에서), 특히 원핵생물(예를 들어, E.coli와 같은 박테리아) 또는 진행 생물(예를 들어, CHO 세포와 같은 포유류 세포, 효모 세포 또는 곤충 세포 또는 식물 또는 동물과 같은 전체 유기체)의 발현 시스템에서 전사 및/또는 번역에 적절한 핵산에 포함된다.

[0052] 상기 용어 "발현 시스템"은 웨타이드, 단백질 또는 RNA, 특히 mRNA의 생산에 적절한 시스템(세포 배양 또는 전체 유기체)을 의미한다.

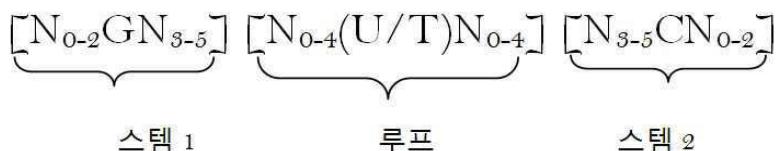
[0053] 본 발명의 첫 번째 실시예에 따른 독창적인 핵산 서열은 적어도 하나의 히스톤 스템-루프를 포함하거나 코딩한다. 본 발명과 관련하여, 그러한 히스톤 스템-루프는 일반적으로 히스톤 유전자로부터 유래되고, 두 개의 이웃한 전체적으로 또는 부분적으로 역의 상보서열들(reverse complementary sequences)의 분자내 염기 짹짓기를 포함하며, 이로 인해 스템-루프를 형성한다. 스템-루프는 단일가닥 DNA 또는 더 흔히는, RNA에서 발생할 수 있다. 상기 구조는 헤어핀(hairpin) 또는 헤어핀 루프로 알려져 있고 보통은 보존적 서열 내에서 스템 및 (말단의) 루프로 이루어져 있으며, 상기 스템은 스템-루프 구조의 루프를 만드는 얼마간의 스페이서(spacer) 같은 짧은 서열에 의해 분리된 두 개의 이웃한 전체적인 또는 부분적인 역의 상보서열에 의해 형성된다. 상기 두 개의 이웃한 전체적인 또는 부분적인 역의 상보서열은 예를 들어, 스템 루프 요소 스템1(stem1) 및 스템2(stem2)와 같이 정의될 수 있다. 상기 스템 루프는 이러한 두 개의 이웃한 전체적인 또는 부분적인 역의 상보서열, 예를 들어 스템 루프 요소들 스템1 및 스템2가 서로서로 염기-짝을 형성할 때 만들어지며, 보존적 서열에서 스템 루프 요소 스템1 및 스템2 사이에 위치한 짧은 서열에 의해 만들어진 그것의 말단 끝에서 짹짓지 않은 루프를 포함하는 이중 가닥 핵산 서열을 이끌어낸다. 그러므로 짹짓지 않은 루프는 일반적으로 이러한 스템 루프 요소들 중 어느 한쪽과 염기 짹짓기를 할 수 없는 핵산의 영역을 대표한다. 상기 결과적인 막대사탕 형태의 구조는 많은 RNA 이차 구조의 핵심적인 빌딩 블록(building block)이다. 스템-루프 구조의 형성은 따라서 결과적인 스템 및 루프 영역의 안정성에 의존하며, 가장 우선시되는 필수 조건은 일반적으로 짹짓은 이중 가닥을 형성하기 위하여 그 자체에 접힐 수 있는 서열이 존재하는 것이다. 짹짓은 스템 루프의 안정성은 길이, 그것을 포함하는 미스매치 또는 불룩하게 튀어 나온 형태의 수(적은 수의 미스매치는 일반적으로 허용할만하고, 특히 길이가 긴 이중가닥에서), 및 짹짓은 영역의 염기 구성에 의해서 결정된다. 본 발명과 관련하여, 바람직한 루프 길이는 3-10 염기, 더 바람직하게는 3 내지 8, 3 내지 7, 3 내지 6 또는 더욱 바람직하게는 4 내지 5, 그리고 가장 바람직하

계는 4 염기이다.

[0054] 본 발명에 따르면, 청구항 1의 구성요소 (b)에 따른 히스톤 스템 루프 서열은 생쥐(mouse)로부터 유래되지 않을 수 있다. 더 구체적으로, 히스톤 스템 루프 서열은 생쥐 히스톤 유전자 H2A614로부터 유래되지 않을 것이다. 또한, 본 발명의 핵산은 생쥐 히스톤 스템 루프 서열도 생쥐 히스톤 유전자 H2A614도 포함하지 않을 수 있다. 게다가, 상기 독창적인 핵산 서열은 스템-루프 가공 신호, 더 구체적으로, 생쥐 히스톤 가공 신호를 포함하지 않을 수 있고, 그리고 가장 구체적으로, 생쥐 스템 루프 가공 신호 H2kA614을 포함하지 않을 수 있다. 또한, 만약 상기 독창적인 핵산 문자가 적어도 하나의 포유동물의 히스톤 유전자를 포함할 수 있다면. 그러나, 적어도 하나의 포유동물의 히스톤 유전자는 WO 01/12824의 Seq. ID No. 7이 아닐 수 있다.

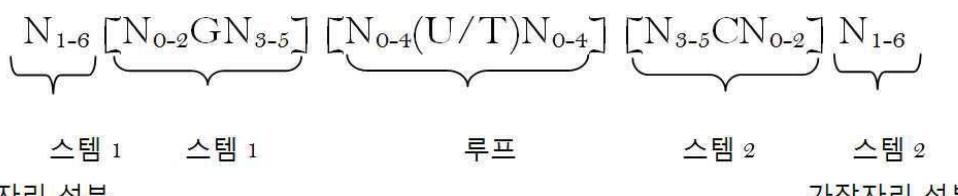
[0055] 첫 번째 독창적인 실시예의 한가지 바람직한 측면에 따르면, 상기 독창적인 핵산 문자 서열은 적어도 하나의 히스톤 스템-루프 서열을 포함하거나 코딩하며, 바람직하게는 다음의 식(I) 또는 (II) 중 적어도 하나에 따른다:

[0056] 식 (I) (스템 가장자리 성분이 없는 스템-루프 서열):



[0057]

[0058] 식 (II) (스템 가장자리 성분이 있는 스템-루프 서열):



[0059]

[0060] 상기:

[0061] 스템1 또는 스템2 가장자리 성분  $N_{1-6}$  은 1 내지 6의 연속적인 서열, 바람직

하게는 2 내지 6의, 더 바람직하게는

2내지 5의, 더욱 더 바람직하게는 3

내지 5의, 가장 바람직하게는 4 내지

5 또는 5N이며, 상기 각각의 N은 A,

U, T, G 및 C 또는 이들의 뉴클레오티드로부터 서로 독립적으로 선택

드 유사

되며;

[0062] 스템1  $[N_{0-2}GN_{3-5}]$  은 성분 스템2와 역으로 상보적으로

또는 부분적으로 역으로 상보적이며,

그리고 5 내지 7 뉴클레오티드 사이의

연속적인 서열이고;

[0063]

[0064]

상기  $N_{0-2}$  는 0 내지 2, 바람직하게는

0 내지 1, 더 바람직하게는 1 N의 연속적인 서열이며, 상기 각각의 N은 A, U, T, G 및 C로부터 선택된 뉴클레오티드 또는 이들의 뉴클레오티드 유사체로부터 서로 독립적으로 선택되며;

[0065]

[0066]

상기 N<sub>3-5</sub> 는 3 내지 5, 바람직하게는 4 내지 5, 더 바람직하게는 4N의 연속적인 서열이며, 상기 각각의 N은 A, U, T, G 및 C으로부터 선택된 뉴클레오티드 또는 이들의 뉴클레오티드 유사체로부터 서로 독립적으로 선택되며, 그리고

[0067]

상기 G는 구아노신 또는 이의 유사체, 그리고 만약 스템2에서 그것의 상보적인 뉴클레오티드 시티딘이 구아노신으로 치환되면, 선택적으로 시티딘 또는 이의 유사체로 치환될 수 있고;

[0068]

루프 서열 [N<sub>0-4</sub>(U/T)N<sub>0-4</sub>] 은 성분 스템1과 스템2 사이에 위치하며 3 내지 5 뉴클레오티드, 더 바람직하게는 4 뉴클레오티드의 연속적인 서열이고;

[0069]

상기 각각의 N<sub>0-4</sub> 은 0 내지 4의, 바람직하게 1 내지 3의, 더 바람직하게 1 내지 2N의 다른 연속적인 서열로부터 독립적으로 선택되며, 상기 각각의 N은 A, U, T, G 및 C로부터 선택된 뉴클레오티드 또는 이들의 뉴클레오티드 유사체로부터 서로 독립적으로 선택되며, 그리고

[0070]

상기 U/T 는 우리딘, 또는 선택적으로 티미딘을 나타내며;

[0071]

스템2 [N<sub>3-5</sub>CN<sub>0-2</sub>] 은 성분 스템1과 역으로 상보적인 또 는 부분적으로 역으로 상보적이며, 그리고 5 내지 7 뉴클레오티드의 연속적인 서열이고;

[0072]

[0073]

상기 N<sub>3-5</sub> 은 3 내지 5의, 바람직하게는 4 내지 5의, 더 바람직하게는 4N의 연속적인 서열이며, 상기 각각의 N은

A, U, T, G 및 C로부터 선택된 뉴클레오티드 또는 이들의 뉴클레오티드 유사체로부터 서로 독립적으로 선택되며;

[0074] 상기  $N_{0-2}$  는 0 내지 2, 바람직하게는 0 내지 1, 더 바람직하게는 1N의 연속적인 서열이며, 상기 각각의 N은 A, U, T, G 또는 C로부터 선택된 뉴클레오티드 유사체로부터 서로 독립적으로 선택되고; 그리고

[0075] 상기 C는 시티딘 또는 이의 유사체이며, 그리고 만약 스템1에서 그것의 상보적인 뉴클레오티드 구아노신이 시티딘으로 치환되면, 선택적으로 구아노신 또는 이의 유사체로 치환될 수 있으며;

[0076] 상기

[0077] 스템1 및 스템2는 역의 상보서열을 형성하면서 서로 염기 짹짓기를 할 수 있으며, 상기 염기 짹짓기는 스템1 및 스템2 사이에서 발생할 수 있으며, 예를 들어, 뉴클레오티드 A 와 U/T 또는 G 와 C의 Watson-Crick 염기 짹짓기에 의해 또는 예를 들어 wobble 염기 짹짓기, 역(reverse) Watson-Crick 염기 짹짓기, Hoogsteen 염기 짹짓기, 역 Hoogsteen 염기 짹짓기와 같은 비-Watson-Crick 염기 짹짓기에 의해 발생할 수 있거나 또는 부분적으로 역의 상보서열을 형성하면서 서로 염기 짹짓기를 할 수 있으며, 하나의 스템에서 하나 또는 더 많은 염기들은 다른 스템의 역의 상보 서열에서 상보적인 염기를 갖지 않음에 근거하여, 상기 불완전한 염기 짹짓기는 스템1 및 스템2 사이에서 발생할 수 있다.

[0078] 이와 관련하여, wobble 염기 짹짓기는 일반적으로 두 개의 뉴클레오티드 사이 비-Watson-Crick 염기 짹짓기이다. 본 문맥에서 사용될 수 있는 네 개의 주요 wobble 염기 짹짓기는 구아노신-우리딘, 이노신-우리딘, 이노신-아데노신, 이노신-시티딘(G-U/T, I-U/T, I-A and I-C) 및 아데노신-시티딘(A-C)이다.

[0079] 따라서, 본 발명과 관련하여 wobble 염기는 앞서 기술한 바와 같이 다른 염기와 wobble 염기쌍을 형성할 수 있는 염기이다. 그러므로 비-Watson-Crick 염기 짹짓기, 예를 들어 wobble 염기 짹짓기는 본 발명에 따른 히스톤 스템-루프 구조의 스템에서 발생할 수 있다.

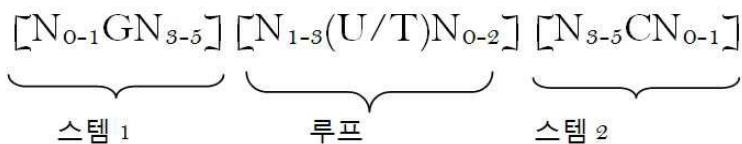
[0080] 이와 관련하여 부분적으로 역의 상보서열은 스템1 및 스템2의 염기쌍에 의해 형성된 스템-루프 서열의 스템-구조에서 최대로 2, 바람직하게는 오직 하나의 미스매치(mismatch)를 포함한다. 다른 말로, 스템1 및 스템2는 바람직하게 스템1 및 스템2의 전체 서열을 통틀어 서로서로간에 (완전한) 염기 짹짓기를 할 수 있으며(가능한 100% 정확한 Watson-Crick 또는 비-Watson-Crick 염기쌍), 이로 인해 역의 상보서열을 형성하며, 상기 각각의 염기는 상보적인 결합 파트너로써 그것의 정확한 Watson-Crick 또는 비-Watson-Crick 염기 짹(pendant)을 가진다. 아니면 그 대신, 스템1 및 스템2는 바람직하게 스템1 및 스템2의 전체 서열을 통틀어 서로 부분적인 염기 짹짓기를 할 수 있으며, 상기 적어도 가능한 100%의 대략 70%, 75%, 80%, 85%, 90%, 또는 95% 정확한 Watson-Crick 또는 비-Watson-Crick 염기 짹짓기는 정확한 Watson-Crick 또는 비-Watson-Crick가 차지하고 기껏해야 남아 있는 염기들의 대략 30%, 25%, 20%, 15%, 10%, 또는 5% 만이 쌍을 형성하지 않는다.

[0081] 첫 번째 독창적인 실시예의 바람직한 측면에 따르면, 본원에 정의된 독창적인 핵산 서열의 (스텝 가장자리 성분을 갖는) 적어도 하나의 히스톤 스템-루프 서열은 대략 15 내지 대략 45 뉴클레오티드 길이를 포함하며, 바람직하게는 대략 15 내지 40 뉴클레오티드 길이, 바람직하게는 대략 15 내지 대략 35 뉴클레오티드의 길이, 바람직하게는 대략 15 내지 대략 30 뉴클레오티드 길이 및 더 바람직하게는 대략 20 내지 대략 30의 길이 그리고 가장 바람직하게는 24 내지 28 뉴클레오티드의 길이를 포함한다.

[0082] 첫 번째 독창적인 실시예의 바람직한 측면에 따르면, 본원에 정의된 독창적인 핵산 서열의 (스텝 가장자리 성분을 갖지 않는) 적어도 하나의 히스톤 스템-루프 서열은 대략 10 내지 대략 30 뉴클레오티드 길이를 포함하며, 바람직하게는 대략 10 내지 20 뉴클레오티드 길이, 바람직하게는 대략 12 내지 대략 20 뉴클레오티드의 길이, 바람직하게는 대략 14 내지 대략 20 뉴클레오티드 길이 및 더 바람직하게는 대략 16 내지 대략 17의 길이 그리고 가장 바람직하게는 16 뉴클레오티드의 길이를 포함한다.

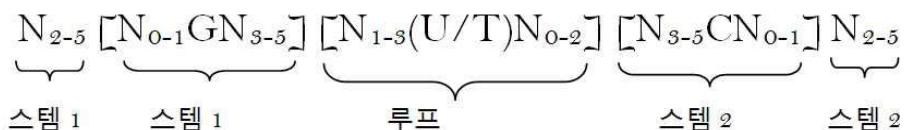
[0083] 첫 번째 독창적인 실시예의 다른 바람직한 측면에 따르면, 본 발명의 첫 번째 실시예에 따른 독창적인 핵산 서열은 다음의 구체적인 식(Ia) 또는 (IIa):의 적어도 하나에 따른 적어도 하나의 히스톤 스템-루프 서열을 포함하거나 코딩할 수 있다;

[0084] 식 (Ia) (스텝 가장자리 성분을 갖지 않는 스템-루프 서열):



[0085]

식 (IIa) (스텝 가장자리 성분을 갖는 스템-루프 서열):



[0087]

가장자리 성분

가장자리 성분

[0088]

상기 :

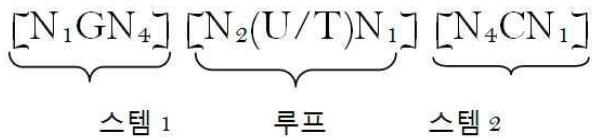
[0089]

N, C, G, T 및 U 는 상기 정의된 바와 같다.

[0090]

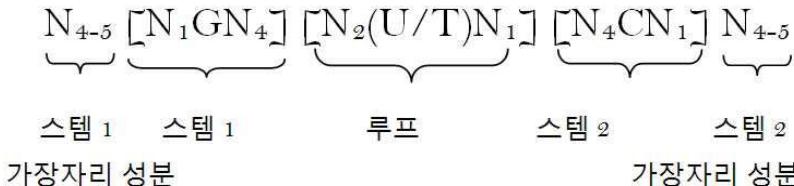
첫 번째 실시예의 더 특허나 바람직한 측면에 따르면, 상기 독창적인 핵산 서열은 다음의 구체적인 식 (Ib) 또는 (IIb):의 적어도 하나에 따른 적어도 하나의 히스톤 스템-루프 서열을 포함하거나 코딩할 수 있다;

[0091] 식 (Ib) (스텝 가장자리 성분들을 갖지 않는 스텝-루프 서열):



[0092]

[0093] 식 (IIb) (스텝 가장자리 성분을 갖는 스텝-루프 서열):



[0094]

[0095] 상기:

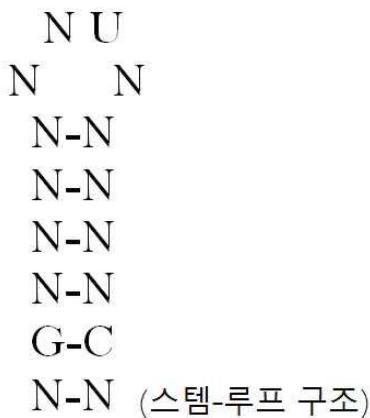
[0096]

N, C, G, T 및 U은 상기 정의된 바와 같다.

[0097]

첫 번째 독창적인 실시예의 더더욱 바람직한 측면에 따르면, 본 발명의 첫 번째 실시예에 따른 독창적인 핵산 서열은 다음의 구체적인 식 (Ic) 내지 (Ih) 또는 (IIc) 내지 (IIh)의 적어도 하나에 따른 적어도 하나의 히스톤 스텝-루프 서열을 포함하거나 코딩할 수 있고, 아니면, 상기 서열은 그것의 스텝-루프 구조에서 보여지며 실시 예 1에 따라 만들어지는 것처럼 히스톤 스텝-루프 서열을 대표하는 선형 서열로써 나타난다:

[0098] 식 (Ic): (스텝 가장자리 성분들을 갖지 않는 후생동물 및 원생동물의 히스톤 스텝-루프 연속적 서열)



NGNNNNNNUNNNNNCN

(선형 서열) (SEQ ID NO: 1)

[0099]

[0100] 식 (IIc): (스텝 가장자리 성분들을 갖는 후생동물 및 원생동물의 히스톤 스템-루프 연속적 서열):

N U  
N N  
N-N  
N-N  
N-N  
N-N  
G-C

N\*N\*NNNN-NNNN\*N\*N\* (스텝-루프 구조)

N\*N\*NNNNGNNNNNNUNNNNNCNNNN\*N\*N\*

(선형 서열) (SEQ ID NO: 2)

[0101]

[0102] 식 (Id): (스텝 가장자리 성분이 없는)

N U  
N N  
N-N  
N-N  
N-N  
N-N  
C-G  
N-N (스텝-루프 구조)

NCNNNNNNUNNNNNNGN

(선형 서열) (SEQ ID NO: 3)

[0103]

[0104] 식 (IIId): (스템 가장자리 성분들을 갖는)

N U  
 N N  
 N-N  
 N-N  
 N-N  
 N-N  
 C-G

N\*N\*NNNN-NNNN\*N\*N\* (스템-루프 구조)

N\*N\*NNNCNNNNNNUNNNNNNGNNNN\*N\*N\*

(선형 서열) (SEQ ID NO: 4)

[0105]

[0106] 식 (Ie): (스템 가장자리 성분을 갖지 않는 원생동물의 히스톤 스템-루프 연속적 서열)

N U  
 N N  
 N-N  
 N-N  
 N-N  
 N-N  
 G-C  
 D-H (스템-루프 구조)

DGNNNNNNUNNNNNCH

(선형 서열) (SEQ ID NO: 5)

[0107]

[0108]

[0109] 식 (IIe): (스템 가장자리 성분을 갖는 원생동물의 히스톤 스템-루프 연속적 서열)

N U  
 N N  
 N-N  
 N-N  
 N-N  
 N-N  
 G-C

N\*N\*NNND-HNNN\*N\*N\* (스템-루프 구조)

N\*N\*NNNDGNNNNNNUNNNNNCHNNN\*N\*N\*

(선형 서열) (SEQ ID NO: 6)

[0110] 식 (If): (스템 가장자리 성분을 갖지 않는 후생동물의 히스톤 스템-루프 연속적 서열)

N U  
 N N  
 Y-V  
 Y-N  
 B-D  
 N-N  
 G-C  
 N-N (스템-루프 구조)

NGNBYYNNUNVNDNCN

(선형 서열) (SEQ ID NO: 7)

[0111]

[0112] 식 (IIf): (스텝 가장자리 성분을 갖는 후생동물의 히스톤 스템-루프 연속적 서열)

N U  
N N  
Y-V  
Y-N  
B-D  
N-N  
G-C

N\*N\*NNNN-NNNN\*N\*N\* (스텝-루프 구조)

N\*N\*NNNNGNBYYNNUNVNDNCNNNN\*N\*N\*

(선형 서열) (SEQ ID NO: 8)

[0113]

[0114] 식 (Ig): (스텝 가장자리 성분을 갖지 않는 척추동물의 히스톤 스템-루프 연속적 서열)

N U  
D H  
Y-A  
Y-B  
Y-R  
H-D  
G-C  
N-N (스텝-루프 구조)

NGHYYDNUHABRDCN

(선형 서열) (SEQ ID NO: 9)

[0115]

[0116]

식 (IIg): (스템 가장자리 성분을 갖는 척추동물의 히스톤 스템-루프 연속적 서열)

N U  
D H  
Y-A  
Y-B  
Y-R  
H-D  
G-C

N\*N\*HNNN-NNNN\*N\*H\* (스템-루프 구조)

N\*N\*HNNNGHYYDNUHABRDCNNNN\*N\*H\*

(선형 서열) (SEQ ID NO: 10)

[0117]

식 (Ih): (스템 가장자리 성분을 갖지 않는 사람의 히스톤 스템-루프 연속적 서열(Homo sapiens))

Y U  
D H  
U-A  
C-S  
Y-R  
H-R  
G-C  
D-C (스템-루프 구조)

DGHYCUDYUHASRRCC

(선형 서열) (SEQ ID NO: 11)

[0119]

[0120] 식 (IIh): (스템 가장자리 성분을 갖는 사람의 히스톤 스템-루프 연속적 서열(Homo sapiens))

Y U  
D H  
U-A  
C-S  
Y-R  
H-R  
G-C

N\*H\*AAHD-CVHB\*N\*H\* (스템 루프 구조)

N\*H\*AAHDGHYCUDYUHASRRCCVHB\*N\*H\*

(선형 서열) (SEQ ID NO: 12)

[0121]

[0122] 상기 각각의 식 (Ic) 내지 (Ih) 또는 (IIc) 내지 (IIh)에서:

[0123] N, C, G, A, T 및 U 은 상기 정의한 바와 같고;

[0124] 각각의 U는 T에 의하여 치환될 수 있으며;

[0125] 만약, 유사한 스템에서 그것의 상보적인 뉴클레오티드가 동시에 그것의 상보적인 뉴클레오티드에 의해 치환된다면, 스템 성분 1 및 2에서 각각의 (매우) 연속적인 G 또는 C는 그것의 상보적인 뉴클레오티드 염기 C 또는 G에 의해 치환될 수 있으며; 그리고/또는

[0126] G, A, T, U, C, R, Y, M, K, S, W, H, B, V, D, 및 N 은 다음의 표에서 정의된 뉴클레오티드 염기들이다:

### 표 1

약어	뉴클레오티드 염기	참고사항
G	G	Guanine
A	A	Adenine
T	T	Thymine
U	U	Uracile
C	C	Cytosine
R	G or A	Purine
Y	T/U or C	Pyrimidine
M	A or C	Amino
K	G or T/U	Keto
S	G or C	Strong (3H bonds)
W	A or T/U	Weak (2H bonds)
H	A or C or T/U	Not G
B	G or T/U or C	Not A
V	G or C or A	Not T/U
D	G or A or T/U	Not C
N	G or C or T/U or A	Any base
*	Present or not	Base may be present or not

[0128] 이와 관련하여, 특히 바람직한 것은 본 발명의 식 (I) 또는 (Ia) 내지 (Ih) 또는 (II) 또는 (IIa) 내지 (IIh) 중 적어도 하나에 따른 히스톤 스템-루프 서열은 자연적으로 발생한 히스톤 스템 루프 서열로부터 선택되며, 더 특히 바람직하게는 원생동물의 또는 후생동물의 히스톤 스템-루프 서열로부터, 그리고 더 특히 바람직하게는 척

추동물로부터 그리고 가장 바람직하게는 포유류 히스톤 스템-루프 서열, 특히 사람의 히스톤 스템-루프 서열로부터 선택된다.

- [0129] 첫 번째 실시예의 특히 바람직한 측면에 따르면, 본 발명의 구체적인 식 (I) 또는 (Ia) 내지 (Ih) 또는 (II) 또는 (IIa) 내지 (IIh) 중 적어도 하나에 따른 히스톤 스템-루프 서열은 도 1 내지 5에서 보여지는 바와 같은, 각각의 뉴클레오티드 위치에서 가장 빈번하게 나타나는 뉴클레오티드, 또는 가장 빈번하거나 두번째로 가장 빈번하게 나타나는, 후생동물 및 원생동물(도 1), 원생동물(도 2), 후생동물(도 3), 척추동물(도 4) 및 사람(도 5)에서 자연적으로 발생하는 히스톤 스템-루프 서열들의 뉴클레오티드를 포함하는 히스톤 스템-루프 서열이다. 이와 관련하여 특히 바람직한 것은 모든 뉴클레오티드들의 적어도 80%, 바람직하게는 적어도 85%, 또는 가장 바람직하게는 적어도 90%는 자연적으로 발생하는 히스톤 스템-루프 서열들의 가장 빈번하게 존재하는 뉴클레오티드와 일치하는 것이다.
- [0130] 첫 번째 실시예의 다른 측면에서, 본 발명의 구체적인 식 (I) 또는 (Ia) 내지 (Ih)의 적어도 하나에 따른 히스톤 스템-루프 서열은 실시예 1에 따라 만들어지는 것과 같은 히스톤 스템-루프 서열을 대표하는 다음의 히스톤 스템-루프 서열들(스템-가장자리 성분들을 갖지 않는)로부터 선택된다:
- [0131] VGYYYYHHTHRVVRCB (식 (Ic)에 따른 SEQ ID NO: 13)
- [0132] SGYYTTTYTMARRRCS (식 (Ic)에 따른 SEQ ID NO: 14)
- [0133] SGYYCTTTMAGRRCs (식 (Ic)에 따른 SEQ ID NO: 15)
- [0134] DGNNNBNNTHVNNNCH (식 (Ie)에 따른 SEQ ID NO: 16)
- [0135] RGNNNYHBTHRDNNCY (식 (Ie)에 따른 SEQ ID NO: 17)
- [0136] RGNDBYHYTHRHDHNCY (식 (Ie)에 따른 SEQ ID NO: 18)
- [0137] VGYYTYHHTHRVRRCB (식 (If)에 따른 SEQ ID NO: 19)
- [0138] SGYYCTTYTMAGRRCs (식 (If)에 따른 SEQ ID NO: 20)
- [0139] SGYYCTTTMAGRRCs (식 (If)에 따른 SEQ ID NO: 21)
- [0140] GGYYCTTYTHAGRRC (식 (Ig)에 따른 SEQ ID NO: 22)
- [0141] GGCYCTTYTMAGRGC (식 (Ig)에 따른 SEQ ID NO: 23)
- [0142] GGCTCTTTMAGRGC (식 (Ig)에 따른 SEQ ID NO: 24)
- [0143] DGHYCTDYTHASRRCC (식 (Ih)에 따른 SEQ ID NO: 25)
- [0144] GGCYCTTTTHAGRGC (식 (Ih)에 따른 SEQ ID NO: 26)
- [0145] GGCYCTTTMAGRGC (식 (Ih)에 따른 SEQ ID NO: 27)
- [0146] 게다가 이와 관련하여 구체적인 식 (II) 또는 (IIa) 내지 (IIh) 중 하나에 따른 실시예 1에 따라 만들어지는 것과 같은 (스템 가장자리 성분을 갖는) 다음의 히스톤 스템-루프 서열들은 특히 바람직하다:
- [0147] H\*H\*HHVGYYYYHHTHRVVRCBVHH\*N\*N\* (식 (IIc)에 따른 SEQ ID NO: 28)

[0148] M\*H\*MHMSGYYYTTYTMARRCSMCH\*H\*H\* (식 (IIc)에 따른 SEQ ID NO: 29)

[0149] M\*M\*MMMSGYYCTTTMAGRCSACH\*M\*H\* (식 (IIc)에 따른 SEQ ID NO: 30)

[0150] N\*N\*NNNDGNNNBNNTHVNNNCHNH\*N\*N\* (식 (IIe)에 따른 SEQ ID NO: 31)

[0151] N\*N\*HHNRGNNNYHBTHRDNNCYDHH\*N\*N\* (식 (IIe)에 따른 SEQ ID NO: 32)

[0152] N\*H\*HHVRGNDBYHYTHRDIHNCYRHH\*H\*H\* (식 (IIe)에 따른 SEQ ID NO: 33)

[0153] H\*H\*MHMVGYYYTYHTHRVRRCBVMH\*H\*N\* (식 (IIf)에 따른 SEQ ID NO: 34)

[0154] M\*M\*MMMSGYYCTTYTMAGRCSMCH\*H\*H\* (식 (IIf)에 따른 SEQ ID NO: 35)

[0155] M\*M\*MMMSGYYCTTTMAGRCSACH\*M\*H\* (식 (IIf)에 따른 SEQ ID NO: 36)

[0156] H\*H\*MAMGGYYCTTYTHAGRCCVHN\*N\*M\* (식 (IIg)에 따른 SEQ ID NO: 37)

[0157] H\*H\*AAMGGCYCTTYTMAGRGCCVCH\*H\*M\* (식 (IIg)에 따른 SEQ ID NO: 38)

[0158] M\*M\*AAMGGCTTTMAGRGCCMCY\*M\*M\* (식 (IIg)에 따른 SEQ ID NO: 39)

[0159] N\*H\*AAHDGHYCTDYTHASRRCCVHB\*N\*H\* (식 (IIh)에 따른 SEQ ID NO: 40)

[0160] H\*H\*AAMGGCYCTTTAGRGCCVMY\*N\*M\* (식 (IIh)에 따른 SEQ ID NO: 41)

[0161] H\*M\*AAAGGCYCTTTMAGRGCCRMY\*H\*M\* (식 (IIh)에 따른 SEQ ID NO: 42)

[0162] 첫 번째 독창적인 실시예의 더 바람직한 측면에 따르면, 상기 독창적인 핵산 서열은 구체적인 식 (I) 또는 (Ia) 내지 (Ih) 또는 (II) 또는 (IIa) 내지 (IIh) 또는 자연적으로 발생하는 히스톤 스템-루프 서열 중 적어도 하나에 따른 히스톤 스템-루프 서열들에서 100% 연속적 뉴클레오티드가 아닌 것과 적어도 약 80%, 바람직하게는 적어도 약 85%, 더 바람직하게는 적어도 약 90%, 또는 더욱 더 바람직하게는 적어도 약 95%, 서열 상동성을 보이는 적어도 하나의 히스톤 스템-루프 서열을 포함하거나 코딩한다.

[0163] 본 발명의 첫 번째 실시예에 따른 독창적인 핵산 서열은 선택적으로 폴리(A) 서열을 포함하거나 코딩한다. 상기 폴리 (A) 서열이 존재할 때, 그러한 폴리(A) 서열은 약 25 내지 약 400 아데노신 뉴클레오티드 서열, 바람직하게 약 50 내지 약 400 아데노신 뉴클레오티드, 더 바람직하게는 약 50 내지 약 300 아데노신 뉴클레오티드 서열, 더욱 더 바람직하게는 약 50 내지 약 250 아데노신 뉴클레오티드 서열, 가장 바람직하게는 약 60 내지 약 250 아데노신 뉴클레오티드 서열을 포함한다. 이와 관련하여 용어 "약"은 부착되는 값(들)의 ±10%의 편차를 의미한다.

[0164] 바람직하게, 본 발명에 따른 핵산은 다음으로 이루어진 그룹의 구성성분들 중 하나 또는 둘 또는 적어도 하나 또는 모두를 포함하는 것이 아니라 하나 또는 모두를 포함한다: 리보자임(바람직하게 자가-스플라이싱 리보자임), 바이러스성 핵산 서열, 히스톤 스템-루프 가공 신호, 특히 생쥐 히스톤 H2A614 유전자로부터 유래된 히스톤 스템 루프 가공 서열, 네오 유전자(Neo), 불활성화된 프로모터 서열 및 불활성화된 인핸서 서열. 더 바람직하게, 본 발명에 따른 핵산은 리보자임, 바람직하게는 자가-스플라이싱 리보자임 및 다음으로 이루어진 그룹 중 하나를 포함하지 않는다: 네오(Neo) 유전자, 불활성화된 프로모터 서열, 불활성화된 인핸서 서열, 히스톤 스템-루프 가공 신호, 특히 생쥐 히스톤 H2A614 유전자로부터 유래된 히스톤 스템 루프 가공 서열. 따라서, 바람직한 모드에서, 상기 핵산은 리보자임, 바람직하게 자가-스플라이싱 리보자임 뿐만 아니라 네오 유전자를 포함하지 않으며 또는 그 대신에 리보자임, 바람직하게 자가-스플라이싱 리보자임 뿐만 아니라 어떠한 저항성 유전

자(예를 들어, 일반적으로 선별에 이용되는(selection))도 포함하지 않는다. 다른 바람직한 모드에서, 본 발명의 핵산은 리보자임, 바람직하게는 자가-스플라이싱 리보자임 뿐만 아니라, 히스톤 스템-루프 가공 신호, 특히 생쥐 히스톤 H2A614 유전자로부터 유래된 히스톤-스템 루프 가공 신호도 포함하지 않는다.

[0165] 택일적으로, 본 발명의 첫 번째 실시예에 따르면, 상기 독창적인 핵산 서열은 선택적으로 특정한 단백질 팩터(예를 들어, 절단 및 폴리아데닐레이션 특이성 팩터(cleavage and polyadenylation specificity factor (CPSF)), 절단 자극 팩터(cleavage stimulation factor (CstF)), 절단 팩터 I 및 II (CF I 및 CF II), 폴리(A) 중합효소 (PAP))에 의하여 mRNA에 폴리아데닐레이션을 나타내는 신호로써 본원에 정의된 폴리아데닐레이션 신호를 포함한다. 이와 관련하여 공통 폴리아데닐레이션 신호는 바람직하게는 NNUANA 공통 서열을 포함한다. 더 바람직한 측면은 상기 폴리아데닐레이션 신호는 다음의 서열들 중 하나를 포함한다: AAUAAA 또는 AUUAAA.

[0166] 본 발명의 첫 번째 실시예에 따른 독창적인 핵산 서열은 게다가, 이들에 제한되지 않고, 애쥬반트 단백질을 포함하는 치료적으로 활성인 단백질 또는 웨타이드로부터, 항원, 예를 들어, 종양 항원, 병원성 항원(예를 들어, 동물 항원으로부터, 바이러스성 항원으로부터, 원생동물의 항원으로부터, 박테리아성 항원으로부터 선택된), 알레르기성 항원, 자가면역 항원, 또는 다른 항원들로부터, 알레르겐으로부터, 항체로부터, 면역자극성 단백질 또는 웨타이드로부터, 항원-특이적 T-세포 수용체들로부터, 또는 특정(치료적) 이용에 적절한 다른 단백질 또는 웨타이드로부터 선택될 수 있는 단백질 또는 웨타이드를 암호화하며, 상기 독창적인 핵산은 세포(예를 들어, 발현 숙주 세포 또는 체세포(somatic cell), 조직 또는 유기체안으로 수송될 수 있고 상기 단백질은 이러한 세포, 조직 또는 유기체내에서 동시에 발현될 수 있다.

[0167] 본 발명의 첫 번째 실시예에 따른 독창적인 핵산의 코딩 영역은 모노-(mono-), 다이-(di-) 또는 삼지어 멀티 시스트론의 핵산, 즉 하나, 둘 또는 더 많은 단백질 또는 웨타이드의 코딩 서열을 가지는 핵산으로 존재할 수 있다. 다이 또는 멀티 시스트론의 핵산에서 코딩 서열은, 예를 들어 본원에 정의된 바와 같이 또는 몇개의 단백질 또는 웨타이드를 포함하는 최종 폴리웨타이드의 절단을 유도하는 신호 웨타이드와 같은, 적어도 하나의 내부리 보솜유입점(internal ribosome entry site (IRES))에 의해 분리될 수 있다.

[0168] 본 발명의 첫 번째 실시예의 특히 바람직한 측면에서 암호화된 웨타이드 또는 단백질들은 사람, 바이러스의, 박테리아의, 원생동물의 단백질 또는 웨타이드로부터 선택된다.

[0169] 본 발명과 관련하여, 독창적인 핵산 분자에 의해 암호화된 치료적으로 활성인 단백질은 이들에 제한되지 않고, 바람직하게는 본원에 정의된 바와 같이, 치료, 예방 또는 지방 치료에 효과를 갖는 단백질들로부터 선택될 수 있고, 또는 각 개체들이 필요로 하는 단백질들이다. 그러한 단백질들은 종래 기술로부터 통상의 기술자들에게 알려진 자연적으로 또는 인공적으로 계획하여 나타나는 재조합 또는 분리된 단백질들로부터 선택될 수 있다. 이들에 제한되지 않고 치료적으로 활성인 단백질들은 세포내에 신호 전달을 자극하거나 저지할 수 있는 단백질들, 예를 들어, 사이토카인, 림포카인, 모노카인, 성장 인자, 수용체들, 신호전달 분자, 전사 인자, 등과 같은 단백질; 항응혈제(anticoagulants); 안티트롬빈(antithrombins); 항알레르기성 단백질(antiallergic proteins); 세포사멸 인자(apoptotic factor) 또는 세포사멸과 관련된 단백질, 치료적 활성 효소 및 어떤 후천적 질병 또는 어떤 유전적 질병과 연관된 단백질을 포함할 수 있다.

[0170] 바람직하게, 독창적인 핵산 분자에 의해 암호화될 수 있는 치료적으로 활성인 단백질은 또한 애쥬반트 단백질일 수 있다. 이와 관련하여, 애쥬반트 단백질은 바람직하게, 본원에 정의된 선천성 면역 반응을 유도할 수 있는 어떠한 단백질로써 이해될 수 있다. 바람직하게, 그러한 선천성 면역 반응은 예를 들어, 인간 TLR1 내지 TLR10 또는 쥐의 톨 유사 수용체 TLR1 내지 TLR13로부터 선택된 톨 유사 수용체(Toll-like receptor (TLR))를 포함하는 톨 유사 수용체 패밀리로부터 선택된 수용체와 같은, 패턴 인식 수용체(Pattern recognition receptors)의 활성을 포함한다. 더 바람직하게, 애쥬반트 단백질은 인간 애쥬반트 단백질들로부터 또는 병원성 애쥬반트 단백질들로부터 선택되거나 또는 이들에 제한되지 않고, 박테리아성 단백질들, 원생동물성 단백질들, 바이러스성 단

백질들 또는 곰팡이성 단백질들, 동물성 단백질들로 이루어진 그룹으로부터 선택되며, 특히 박테리아성 애쥬반트 단백질들로부터 선택된다. 게다가, 애쥬반트 효과에 관계되는 인간 단백질들(예를 들어, 패턴 인식 수용체들의 리간드, 패턴 인식 수용체, 신호 전달 과정의 단백질들, 전사 인자 또는 사이토카인들)을 암호화하는 핵산들은 마찬가지로 사용될 수 있다.

- [0171] 상기 독창적인 핵산 분자는 그렇지 않으면, 항원을 암호화할 수 있다. 본 발명에 따르면, 용어 "항원"은 면역 시스템에 의해 인식되는 물질을 가르키고, 그리고 예를 들어, 후천적 면역 반응의 한 부분으로써 항체 또는 항원-특이적 T-세포의 형성에 의한 항원-특이적 면역 반응을 유발할 수 있는 물질을 나타낸다.
- [0172] 이와 관련하여 항원 에피토프, 단백질의 절편 또는 웹타이드는 특히 B 세포, 항체 또는 T 세포 각각에 의해 인식될 수 있는 B 세포 및 T 세포 에피토프를 의미한다.
- [0173] 본 발명과 관련하여, 독창적인 핵산 분자에 의해 암호화될 수 있는 항원들은 일반적으로 상기 정의된 어떤 항원, 항원의 에피토프, 항원의 절편 또는 항원의 웹타이드를 포함하며, 더 바람직하게는 예를 들어 종양 항원, 알레르기 항원 또는 알레르겐, 자가 면역 자가-항원, 병원성 항원 등과 같은 단백질 및 웹타이드 항원을 포함한다.
- [0174] 특히 독창적인 핵산 분자에 의해 암호화되는 항원들은 세포 바깥에서 만들어지는 항원들, 더 일반적으로 숙주 유기체(예를 들어, 인간) 그 자체(즉 비-자가 항원)로부터 유래된 것이 아니라 오히려 숙주 유기체 외부의 숙주 세포들로부터 유래된 항원들, 예를 들어 바이러스성 항원, 박테리아성 항원, 곰팡이 항원, 원생동물성 항원, 동물성 항원, 알러지성 항원 등으로부터 유래된 항원들일 수 있다. 알레르기 항원들(알러지 항원 또는 알레르겐)은 일반적으로 인간의 알러지를 유발하는 항원이며 인간 또는 다른 출처로부터 유래될 수 있다. 추가적으로, 독창적인 핵산 분자에 의해 암호화된 항원들은 게다가 세포, 조직 또는 몸 전체 내부에서 만들어질 수 있다. 그러한 항원들은 숙주 유기체(예를 들어 인간) 그 자체로부터 유래된 항원들, 예를 들어 종양 항원들, 자가-면역 자가-항원(auto-immune self-antigens)과 같은 자기-항원(self-antigens or auto-antigens)으로부터 유래된 항원을 포함하며, 그러나 또한 본원에 정의된 (비-자기) 항원은 이는 원래 숙주 유기체 바깥의 숙주 세포들로부터 유래되었으나, 예를 들어 (단백질 분해효소) 분해, 대사 등에 의하여 몸 전체, 조직 또는 세포 내에서 조각나거나 분해된다.
- [0175] 독창적인 핵산 분자에 의해 암호화될 수 있는 항원들 중 한가지 종류는 종양 항원을 포함한다. "종양 항원"은 바람직하게 (종양) 세포의 표면에 위치한다. 종양 항원은 정상 세포와 비교하여 종양 세포에서 과발현되는 단백질들로부터 또한 선택될 수 있다. 게다가, 종양 항원은 또한 분해된 그들 자신 (또는 최초에 그들 자신이 아닌)이 아니라 가정한 종양과 관련되는 세포에서 발현되는 항원을 포함한다. 종양 공급 혈관(tumour-supplying vessel) 또는 이들의 (재)형성과 관련된 항원들, 특히 신혈관생성(neovascularization), 예를 들어 VEGF, bFGF 등과 같은 성장 인자와 관련된 항원들은 또한 이에 포함된다. 게다가 종양과 관련된 항원들은 세포들 또는 조직, 일반적으로 종양 포매(embedding)로부터 얻은 항원을 포함한다. 게다가 몇몇 물질들(보통 단백질 또는 웹타이드)은 (알게 모르게) 암 질환을 겪는 환자들에서 발현되고 그들은 상기 환자들의 체액내에서 증가된 농도를 나타낸다. 이러한 물질들은 또한 "종양 항원"으로 언급되나, 면역 반응 유도 물질이라는 엄격한 의미에서 항원은 아니다. 종양 항원의 종류는 종양-특이적 항원(tumour-specific antigens(TSAs))과 종양-관련 항원(tumour-associated-antigens(TAAs))으로 다시 나눌 수 있다. TSAs는 오직 종양 세포에 의해서 제시될 수 있고 결코 정상의 "건강한" 세포에 의해서는 제시될 수 없다. 그들은 전형적으로 종양 특이적 돌연변이에 기인한 것이다. 더 흔한 TAAs는 보통 종양 세포 및 건강한 세포 모두에 의해 제시된다. 이러한 항원들은 인식되며 항원-제시 세포(antigen-presenting cell)는 세포독성 T 세포에 의해서 파괴될 수 있다. 추가적으로, 종양 항원들은 또한 예를 들어, 돌연변이된 수용체와 같은 형태로 종양의 표면에 나타날 수 있다. 이러한 경우에, 그들은 항체들에 의해서 인식될 수 있다.
- [0176] 다른 대안에 따르면, 독창적인 핵산 분자에 의해 암호화될 수 있는 또 다른 항원의 종류는 알레르기성 항원을

포함한다. 그러한 알레르기성 항원들은 다른 출처로부터, 예를 들어, 동물, 식물, 곰팡이, 박테리아 등으로부터 유래된 항원들로부터 선택될 수 있다. 이와 관련하여 알레르겐은 예를 들어, 풀, 꽃가루, 곰팡이, 약물 또는 다양한 환경 자극(triggers) 등을 포함한다. 알레르기성 항원들은 일반적으로 핵산 및 그들의 절편, 단백질 또는 펩타이드 및 그들의 절편, 탄수화물(carbohydrates), 다당류(polysaccharides), 당분(sugars), 지질(lipids), 인지질(phospholipids) 등과 같은 다른 종류의 화합물들에 속한다. 본 발명과 관련하여 특히 관심이 있는 것은 본원에 정의된 독창적인 핵산 분자, 즉 단백질 또는 펩타이드 항원 및 그들의 절편 또는 에피토프, 또는 핵산 및 그들의 절편들, 특히 그러한 단백질 또는 펩타이드 항원 및 그들의 절편 또는 에피토프를 암호화하는 핵산 및 그들의 절편에 의해 암호화될 수 있는 항원들이다.

[0177] 다른 대안에 따르면, 독창적인 핵산 분자는 항체 또는 항체 절편을 암호화할 수 있다. 본 발명에 따르면, 그러한 항체들은 예를 들어, 기술분야에서 알려진, 재조합하여 생산된 또는 자연적으로 발생한 항체, 특히 치료, 진단 또는 학문적 목적에 적합한 항체들 또는 특정 암 질환에 관련된 것으로 밝혀진 항체들과 같은 어떤 항체로부터 선택될 수 있다. 여기서, 용어 "항체"는 그것의 가장 넓은 의미로 사용되며 구체적으로 단일클론성(monoclonal) 및 다클론성(polyclonal) 항체들(작용제, 길항제, 그리고 차단 또는 중화하는 항체들) 및 다중에피토프 특이성(polyepitopic specificity)을 갖는 항체 종류들을 특히 포괄한다. 본 발명에 따르면, 용어 "항체"는 키메라 항체(chimeric antibodies), 인간 항체, 인간화된 항체(humanized antibodies), 이중 특이성 항체(bispecific antibodies), 내부체(intrabodies), 즉 세포에서 발현되고 선택적으로 특정 세포 구획에 위치한 항체들, 및 언급한 항체들의 절편 및 변이체 뿐만 아니라, 자연적으로 발생하는 항체들, 숙주 유기체에서 면역화에 의해 생성된 항체들, 자연적으로 발생한 항체 또는 숙주 유기체에서 면역화에 의해 생성되고 해당 기술분야에서 잘 알려진 생체 분자적 방법에 의해 재조합적으로 생산된 항체들로부터 분리되거나 확인된 항체들을 포함한다. 일반적으로 항체는 가변부와 불변부를 갖는 경쇄 및 중쇄로 이루어진다. 경쇄는 N-말단 가변부  $V_L$  및 C-말단 불변부  $C_L$ 로 이루어진다. 대조적으로, 예를 들어 IgG 항체의 중쇄는 N-말단 가변부  $V_H$ , 및 세 개의 불변부  $C_H1$ ,  $C_H2$  및  $C_H3$ 로 이루어진다.

[0178] 본 발명과 관련하여, 독창적인 핵산 분자에 의해 암호화되어지는 항체는 바람직하게 완전한 길이의(full-length) 항체, 즉, 상기 기술된 바와 같이 완전한 중쇄 및 완전한 경쇄를 포함하는 항체를 포함한다. 그러나 항체 절편, 변이체 또는 부가체와 같은 항체의 유도체들은 또한 독창적인 핵산 분자에 의해 암호화될 수 있다. 항체 절편들은 바람직하게 앞서 언급된 (완전한-길이) 항체들의  $Fab$ ,  $Fab'$ ,  $F(ab')_2$ ,  $Fc$ ,  $Facb$ ,  $pFc'$ ,  $Fd$  및  $Fv$  절편들로부터 선택된다. 일반적으로, 항체 절편들은 해당 기술분야에서 잘 알려져 있다. 예를 들어,  $Fab$  ("절편(fragment)", 항원 결합(antigen binding)") 절편은 각각의 경쇄 및 중쇄의 하나의 불변부 및 하나의 가변부로 이루어진다. 상기 두 개의 가변부는 특이적 항원에 위치한 에피토프와 결합한다. 상기 두 개의 사슬은 이황화 결합을 통하여 연결된다. 예를 들어,  $scFv$  ("단일 사슬 변이 절편(Single chain variable fragment)") 절편은 일반적으로 경쇄 및 중쇄의 가변부로 이루어진다. 상기 영역은 인공적인(artificial) 결합에 의해 연결되며, 일반적으로 15-25 글리신, 프롤린 및/또는 세린 잔기로 구성된 펩타이드같은 폴리펩타이드 결합에 의해 연결된다.

[0179] 이와 관련하여, 바람직한 것은 항체 또는 항체 절편의 다른 사슬들은 다중시스트론성 핵산 분자에 의해서 암호화된다. 선택적으로, 항체 또는 항체 절편의 다른 변종들(strains)은 몇몇의 단일시스트론성(monocistronic) 핵산(들)(서열)에 의해서 암호화된다.

[0180] 본 발명의 첫 번째 실시예에 따르면, 독창적인 핵산 서열은 바람직하게는 펩타이드 또는 단백질을 암호화하는 코딩 영역을 포함한다. 바람직하게, 암호화된 단백질은 히스톤 단백질이 아니다. 본 발명과 관련하여, 그러한 히스톤 단백질은 일반적으로 DNA 를 뉴클레오좀이라 불리는 구조 단위 안으로 포장하고 배열하는 진핵 세포 핵에서 발견되는 강력한 알칼리성의 단백질이다. 히스톤 단백질은 염색질의 중요 단백질 구성성분이며, DNA 가 움직이지 못하도록 그 주위를 감는 역할을 하고 유전자 조절 역할을 담당한다. 히스톤이 없는, 염색체에서 풀어진 상태의 DNA는 매우 길다(인간 DNA에서 길이 대비 넓이 비율은 천만 이상 대 1의 비율이다). 예를 들어, 각각의 인간 세포는 약 1.8 미터의 DNA를 가지고 있지만 히스톤에 묶여 있으며, 유사분열(mitosis) 동안 복제되고 응축

될 때 약 120 마이크로미터의 크로모좀이 되는 약 90 밀리미터의 크로마틴을 가진다. 더 바람직하게, 본 발명과 관련하여 그러한 히스톤 단백질은 일반적으로 다음의 주요 히스톤 종류 : H1/H5, H2A, H2B, H3, 및 H4"로부터 선택된, 바람직하게는 포유류 히스톤으로부터 선택된, 더 바람직하게는 인간 히스톤 또는 히스톤 단백질로부터 선택된 꽤나 보존적인 단백질로 정의된다. 그러한 히스톤 또는 히스톤 단백질들은 일반적으로 H2A, H2B, H3 및 H4을 포함하는 코어(core) 히스톤 및 히스톤 H1 및 H5을 포함하는 링커(linker) 히스톤으로 정의되는 두 개의 슈퍼-클래스(super-classes)로 분류된다.

[0181] 이와 관련하여, 링커 히스톤은, 바람직하게는 계류중인 발명의 보호 범위로 부터 제외된, 바람직하게는 포유류의 링커 히스톤, 더 바람직하게는 사람 링커 히스톤은 일반적으로 특히 HIST1H1A, HIST1H1B, HIST1H1C, HIST1H1D, HIST1H1E, HIST1H1T을 포함하는 H1H1 및 H1F0, H1FNT, H1FOO, H1FX을 특히 포함하는 H1F을 포함하는 H1으로부터 선택되며; 그리고

[0182] 게다가, 코어 히스톤은 바람직하게는 계류중인 발명의 보호 범위로부터 제외된, 바람직하게는 포유류의 코어 히스톤, 더 바람직하게는 인간 코어 히스톤은 일반적으로 특히 H2AFB1, H2AFB2, H2AFB3, H2AFJ, H2AFV, H2AFX, H2AFY, H2AFY2, H2AFZ을 포함하는 H2A 및 특히 HIST1H2AA, HIST1H2AB, HIST1H2AC, HIST1H2AD, HIST1H2AE, HIST1H2AG, HIST1H2AI, HIST1H2AJ, HIST1H2AK, HIST1H2AL, HIST1H2AM을 포함하는 H2A1 및 특히 HIST2H2AA3, HIST2H2AC을 포함하는 H2A2을 포함하는 H2A; 특히 H2BFM, H2BF0, H2BFS, H2BFWT을 포함하는 H2BF, 특히 HIST1H2BA, HIST1H2BB, HIST1H2BC, HIST1H2BD, HIST1H2BE, HIST1H2BF, HIST1H2BG, HIST1H2BH, HIST1H2BI, HIST1H2BJ, HIST1H2BK, HIST1H2BL, HIST1H2BM, HIST1H2BN, HIST1H2B0을 포함하는 H2B1 및 특히 HIST2H2BE을 포함하는 H2B2을 포함하는 H2B; 특히 HIST1H3A, HIST1H3B, HIST1H3C, HIST1H3D, HIST1H3E, HIST1H3F, HIST1H3G, HIST1H3H, HIST1H3I, HIST1H3J을 포함하는 H3A1 및 특히 HIST2H3C을 포함하는 H3A2 및 특히 HIST3H3을 포함하는 H3A3을 포함하는 H3; 특히 HIST1H4A, HIST1H4B, HIST1H4C, HIST1H4D, HIST1H4E, HIST1H4F, HIST1H4G, HIST1H4H, HIST1H4I, HIST1H4J, HIST1H4K, HIST1H4L을 포함하는 H41 및 특히 HIST4H4을 포함하는 H44을 포함하는 H4; 및 H5로부터 선택된다.

[0183] 본 발명의 첫 번째 실시예에 따르면, 상기 독창적인 핵산 서열은 코딩 영역, 바람직하게는 펩타이드 또는 단백질을 암호화하는 영역을 포함한다. 바람직하게, 암호화된 단백질은 리포터 단백질(예를 들어, 루시페라아제, 녹색 형광 단백질(Green Fluorescent Protein (GFP)), 향상된 녹색 형광 단백질(Enhanced Green Fluorescent Protein(EGFP)),  $\beta$ -갈락토시다아제)이 아니고 마커 또는 선별 단백질(예를 들어, 알파-글로빈, 갈락토키나아제 및 잔틴:구아닌 포스포리보실 전이효소)가 아니다. 바람직하게, 본 발명의 상기 핵산 서열은 (박테리아성) 네오유전자 서열(네오마이신 저항성 유전자)를 포함하지 않는다.

[0184] 상기 정의된 바와 같이 상기 독창적인 핵산은 a) 바람직하게는 펩타이드 또는 단백질을 암호화하는 코딩 영역; b) 적어도 하나의 히스톤 스템-루프 및 c) 선택적으로 폴리(a) 서열 또는 폴리아데닐레이션 신호를 포함하거나 코딩함; 바람직하게는 암호화된 단백질 발현 수준을 증가시키기 위하여, 상기 정의된 바와 같이 상기 암호화된 단백질은 바람직하게는 히스톤 단백질이 아니며, 리포터 단백질이 아니고 그리고/또는 마커 또는 선별 단백질이 아니다. 상기 독창적인 핵산의 구성요소 b) 내지 c)는 어떠한 배열로든 상기 독창적인 핵산내에서 존재할 수 있으며, 즉, 구성요소 a), b) 및 c)는 상기 독창적인 핵산 서열에서 5'에서 3' 방향으로 a), b) 및 c) 또는 a), c) 및 b)로 배열될 수 있으며, 상기 본원에 기술된 다른 구성요소들, 5'-CAP 구조, 폴리(C)서열, 안정화 서열, IRES 서열 등과 같은 것들이 또한 포함될 수 있다. 독창적인 핵산의 구성요소 a) 내지 c) 각각, 특히 다이 또는 멀티시스템트론 구조에서 a) 및/또는 독창적인 핵산의 구성요소 b) 및 c), 더 바람직하게는 b)는 또한 상기 독창적인 핵산에서 적어도 한 번, 바람직하게는 두 번 또는 더 많이 반복될 수 있다. 예로써, 상기 독창적인 핵산은 그것의 서열 구성요소 a), b) 및 선택적으로 c)를 다음과 같은 배열로 보여줄 수 있다:

[0185] 5'-코딩 영역- 히스톤 스템-루프 -3': 또는

[0186] 5'-코딩 영역- 코딩 영역 - 히스톤 스템-루프 -3': 또는

[0187] 5'-코딩 영역- IRES - 코딩 영역 - 히스톤 스템-루프 -3': 또는

- [0188] 5'-코딩 영역- 히스톤 스템-루프 -폴리(A) 서열-3': 또는
- [0189] 5'-코딩 영역- 히스톤 스템-루프 -폴리아데닐레이션 신호-3': 또는
- [0190] 5'-코딩 영역- 코딩 영역 - 히스톤 스템-루프 - 폴리아데닐레이션 신호-3': 또는
- [0191] 5'-코딩 영역- 히스톤 스템-루프 -히스톤 스템-루프 -3': 또는
- [0192] 5'-코딩 영역- 히스톤 스템-루프 -히스톤 스템-루프 -폴리(A) 서열-3': 또는
- [0193] 5'-코딩 영역- 히스톤 스템-루프 -히스톤 스템-루프 -폴리아데닐레이션 신호-3': 또는
- [0194] 5'-코딩 영역- 히스톤 스템-루프 -폴리(A) 서열-히스톤 스템-루프-3': 또는
- [0195] 5'-코딩 영역- 폴리(A) 서열 -히스톤 스템-루프-3': 또는
- [0196] 5'-코딩 영역- 폴리(A) 서열 -히스톤 스템-루프 -히스톤 스템-루프-3': 등.
- [0197] 이와 관련하여 특히 바람직하게는 암호화된 단백질 빌현 수준을 증가시키며 바람직하게는 상기 독창적인 핵산 분자는 a) 바람직하게 켙타이드 또는 단백질을 암호화하는 코딩 영역; b) 적어도 하나의 히스톤 스템-루프, 및 c) 폴리(A) 서열 또는 폴리아데닐레이션 서열을 포함하거나 코딩하며, 상기 암호화된 단백질은 바람직하게 히스톤 단백질이 아니고, 리포터 단백질(예를 들어, 루시페라아제, GFP, EGFP,  $\beta$ -갈락토시다아제, 특히 EGFP)이 아니며 그리고/또는 마커 또는 선별 단백질(예를 들어, 알파-글로빈, 갈락토키나아제 및 잔틴:구아닌 포스포리보실 전이효소(GPT))이 아니다.
- [0198] 첫 번째 실시예의 더 바람직한 측면에서 본원에 정의된 상기 독창적인 핵산 분자는 또한 변형된 핵산의 형태로 나타날 수 있다.
- [0199] 첫 번째 실시예의 한가지 측면에 따르면, 본원에 정의된 독창적인 핵산 분자는 "안정화된 핵산"으로 제공될 수 있으며, 바람직하게는 안정화된 RNA, 더 바람직하게는 생체 내 분해에(예를 들어, 핵산말단가수분해효소에 의해 또는 핵산내부가수분해효소에 의해) 저항성을 갖는데 필수적인 RNA로 제공될 수 있다.
- [0200] 이와 관련하여, 본원에 정의된 독창적인 핵산 분자는 유사체(analogues)/변형들(modifications), 예를 들어 뼈대(backbone) 변형, 당 변형 또는 염기 변형들을 포함할 수 있다. 본 발명과 관련된 뼈대 변형은 본원에 기술된 독창적인 핵산 분자에 포함된 뉴클레오티드들의 뼈의 인삼염이 화학적으로 변형된 변형이다. 본 발명과 관련된 당 변형은 본원에 정의된 독창적인 핵산 분자의 뉴클레오티드의 당의 화학적 변형이다. 게다가, 본 발명과 관련된 염기 변형은 독창적인 핵산 분자의 핵산 분자의 뉴클레오티드의 염기 모이어티(moiety)의 화학적 변형이다. 이와 관련하여 뉴클레오티드 유사체 또는 변형은 바람직하게는 전사 및/또는 번역에 이용될 수 있는 뉴클레오티드 유사체들로부터 선택된다.
- [0201] 본원에 정의된 본 발명의 첫 번째 실시예의 특히 바람직한 측면에서 뉴클레오티드 유사체/변형들은 추가적으로 암호화된 단백질의 발현을 증가시키는 염기 변형들로부터 선택되고 바람직하게는 2-아미노(amino)-6-클로로퓨린리보시드(chloropurineriboside)-5'-트리포스페이트(triphosphate), 2-아미노아데노신(aminoadenosine)-5'-트리포스페이트, 2-티오시티딘(thiocytidine)-5'-트리포스페이트, 2-티오우리딘(thiouridine)-5'-트리포스페이트, 4-티오우리딘-5'-트리포스페이트, 5-아미노알릴시티딘(aminoallylcytidine)-5'-트리포스페이트, 5-아미노알릴우리딘(aminoallyluridine)-5'-트리포스페이트, 5-브로모시티딘(bromocytidine)-5'-트리포스페이트, 5-브로모우리딘(bromouridine)-5'-트리포스페이트, 5-아이오도시티딘(iodocytidine)-5'-트리포스페이트, 5-아이오도우리딘(iodouridine)-5'-트리포스페이트, 5-메틸시티딘(methylcytidine)-5'-트리포스페이트, 5-메틸우리딘(methyluridine)-5'-트리포스페이트, 6-아자시티딘(azacytidine)-5'-트리포스페이트, 6-아자우리딘(azauridine)-5'-트리포스페이트, 6-클로로퓨린리보시드(chloropurineriboside)-5'-트리포스페이트, 7-데아자아데노신(deazaadenosine)-5'-트리포스페이트, 7-데아자구아노신(deazaguanosine)-5'-트리포스페이트, 8-아자

아데노신(azaadenosine)-5'-트리포스페이트, 8-아지도아데노신(azidoadenosine)-5'-트리포스페이트, 벤즈이미다졸(benzimidazole)-리보시드(riboside)-5'-트리포스페이트, N1-메틸아데노신(methyladenosine)-5'-트리포스페이트, N1-메틸구아노신(methylguanosine)-5'-트리포스페이트, N6-메틸아데노신-5'-트리포스페이트, 06-메틸구아노신-5'-트리포스페이트, 슈도우리딘(pseudouridine)-5'-트리포스페이트, 또는 퓨로마이신(puromycin)-5'-트리포스페이트, 크산토신(xanthosine)-5'-트리포스페이트로 부터 선택된다. 특히 선호되는 것은 5-메틸시티딘-5'-트리포스페이트, 7-데아자구아노신-5'-트리포스페이트, 5-브로모시티딘-5'-트리포스페이트, 및 슈도우리딘-5'-트리포스페이트로 이루어진 염기-변형된 뉴클레오티드 그룹으로부터 선택된 염기 변형 뉴클레오티디이다.

[0202]

다른 측면에 따르면, 본원에 정의된 독창적인 핵산 분자는 리피드(lipid) 변형을 포함할 수 있다. 그러한 리피드-변형된 핵산은 일반적으로 본원에 정의된 핵산을 포함한다. 본원에 정의된 독창적인 핵산 분자의 리피드-변형된 핵산 분자는 일반적으로 그러한 핵산 분자와 공유적으로 결합된 적어도 하나의 링커 및 각각의 링커와 공유적으로 결합된 적어도 하나의 링커를 더 포함한다. 아니면 그 대신, 리피드-변형된 핵산 분자는 본원에 정의된 적어도 하나의 핵산 분자 및 (링커없이) 핵산 분자와 공유적으로 결합된 적어도 하나의 (두 기능을 갖는) 리피드를 포함한다. 세 번째 대안에 따르면, 리피드-변형된 핵산 분자는 본원에 정의된 핵산 분자, 상기 핵산 분자와 공유적으로 결합된 적어도 하나의 링커, 및 각각의 링커와 공유적으로 결합된 적어도 하나의 리피드 그리고 또한 (링커 없이) 상기 핵산 분자와 공유적으로 결합된 적어도 하나의 (두 기능을 갖는) 리피드를 포함한다. 이와 관련하여 특히 바람직한 것은 리피드 변형은 선형의 독창적인 핵산 분자의 말단에 나타나는 것이다.

[0203]

본 발명의 첫 번째 실시예의 다른 바람직한 측면에 따르면, 본원에 정의된 독창적인 핵산 분자, 특히 (m)RNA라면, 소위 "5'-CAP" 구조의 추가에 의해서 RNA가수분해효소(RNases)에 의한 분해에 대해서 안정화될 수 있다.

[0204]

본 발명의 첫 번째 실시예의 다른 바람직한 측면에 따르면, 본원에 정의된 독창적인 핵산 분자는 적어도 10 시티딘(cytidines), 바람직하게는 적어도 20 시티딘, 더 바람직하게는 적어도 30 시티딘 서열(소위 "폴리(C) 서열"로 불리는)에 의해서 변형될 수 있다. 특히, 독창적인 핵산 분자는 일반적으로 약 10 내지 200 시티딘 뉴클레오티드, 바람직하게 약 10 내지 100 시티딘 뉴클레오티드, 더 바람직하게 약 10 내지 70 시티딘 뉴클레오티드 또는 더 바람직하게는 약 20 내지 50 또는 20 내지 30 시티딘 뉴클레오티드의 폴리(C) 서열을 포함하거나 코딩할 수 있다. 이러한 폴리(C) 서열은 바람직하게 본 발명의 첫 번째 실시예에 따른 독창적인 핵산에 포함된 코딩 영역의 3'에 위치한다.

[0205]

본 발명의 첫 번째 실시예의 다른 바람직한 측면에 따르면, 본원에 정의된 독창적인 핵산 분자는 바람직하게 적어도 하나의 5' 및/또는 3' 안정화 서열을 가진다. 5' 및/또는 3' 비 번역 영역에 있는 이러한 안정화 서열들은 사이토졸(cytosol) 내에서 핵산의 반감기를 증가시키는 효과를 가진다. 이러한 안정화 서열은 바이러스, 박테리아 및 진핵생물들에서 나타나는 자연발생적인 서열들과 100% 서열 상동성을 가질 수 있으나 또한 부분적으로 또는 완전히 합성될 수 있다. 예를 들어, 호모 사피엔스(*Homo sapiens*) 또는 제노푸스 래비스(*Xenopus laevis*)로부터 유래한, (알파-)글로빈 유전자의 비번역 서열(UTR)은 안정화된 핵산을 위하여 본 발명에서 사용될 수 있는 안정화 서열의 예로써 언급될 수 있다. 안정화 서열의 다른 예는 (알파-)글로빈, 타입(I)-콜라겐, 15-리폭시제나아제(lipoxygenase) 또는 티로신 수산화효소(tyrosine hydroxylase) (cf. Holcik *et al.*, Proc. Natl. Acad. Sci. USA 1997, 94: 2410 to 2414)를 코딩하는, 매우 안정한 RNAs의 3'-UTRs에 포함된 일반식 (C/U)CCAN<sub>x</sub>CCC(U/A)Py<sub>x</sub>UC(C/U)CC (SEQ ID NO: 55)을 가진다. 그러한 안정화 서열은 물론 개별적으로 또는 서로 조합하여 이용될 수 있고 또한 기술분야에서 통상의 기술자에게 알려진 다른 안정화 서열들과 조합하여 이용될 수 있다. 이와 관련하여 특히 바람직하게 알파 글로빈 유전자의 3' UTR 서열은 본 발명의 첫 번째 실시예에 따른 독창적인 핵산에 포함된 코딩 서열의 3'에 위치한다.

[0206]

염기들의 치환, 부가 또는 삭제는 잘 알려진 위치 선택적 돌연변이(site directed mutagenesis)에 의한 또는 올리고뉴클레오티드 결합(oligonucleotide ligation) 전략(see e.g. Maniatis *et al.*, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, 3rd ed., Cold Spring Harbor, NY, 2001)로, 핵산

분자의 준비를 위한 DNA 매트릭스(matrix)를 이용하여 본원에 정의된 독창적인 핵산 분자로 수행된다. 본원에 정의된 독창적인 핵산 분자의 준비를 위하여, 그러한 과정에서, 특히 상기 핵산이 mRNA의 형태라면, 대응하는 DNA 분자는 시험관 내에서 전사될 수 있다. 이러한 DNA 매트릭스는, 준비될 수 있는 핵산 분자, 예를 들어 mRNA를 위한 목적으로 하는 뉴클레오티드 서열에 의해 뒤따르는, 시험관 내 전사를 위한, 적절한 프로모터, 예를 들어 T7 또는 SP6 프로모터를 포함하며 시험관내 전사를 위한 종결 신호를 포함한다. 적어도 하나의 관심있는 RNA 매트릭스를 형성하는 DNA 분자는 박테리아에서 복제될 수 있는 플라스미드의 부분으로써 fermentative 증식 및 뒤따르는 분리에 의해서 준비될 수 있다. 본 발명에 적합한 것으로 언급될 수 있는 플라스미드는 예를 들어 플라스미드 pT7Ts (GenBank accession number U26404; Lai *et al.*, Development 1995, 121: 2349 to 2360), 예를 들어, pGEM<sup>®</sup> 1 (GenBank accession number X65300; from Promega)와 같은 pGEM<sup>®</sup> 시리즈 및 pSP64 (GenBank accession number X65327)이며; 반면 또한, Mezei and Storts, Purification of PCR Products, in: Griffin and Griffin (ed.), PCR Technology: Current Innovation, CRC Press, Boca Raton, FL, 2001.

[0207] 본원에 정의된 본 발명에 따라 이용된 핵산 분자는 시험관내 전사 반응과 같은 시험관 내 방법들 또는 박테리아에서 DNA 플라스미드의 생체 내 증식과 같은 생체 내 반응들뿐만 아니라, 고상 합성(solid phase synthesis)과 같은 합성 방법들을 포함한, 기술 분야에서 이미 잘 알려진 어떠한 방법을 이용하여 준비될 수 있다.

[0208] 상기 변형들 중 어떤 것이든 본원에 정의된 독창적인 핵산 분자에 이용될 수 있고 본 발명과 관련하여 이용되는 어떤 핵산에든지 또한 이용될 수 있으며, 만약 적절하거나 필요할 경우에는, 이러한 변형들의 조합이 각각의 핵산에서 서로서로 방해하지 않는다면, 어떠한 조합이든 가능할 수 있다. 기술 분야에서 통상의 기술자들은 따라서 그들의 선택에 따라 수행할 수 있을 것이다.

[0209] 이러한 핵산 분자에 의해 암호화된 단백질 또는 웨타이드 뿐만 아니라 본원에 정의된 독창적인 핵산 분자는 그러한 서열들의 절편 또는 변이체를 포함할 수 있다. 그러한 절편 또는 변이체들은 적어도 하나의 핵산 분자에 의해 암호화되어진다면, 핵산 수준 또는 아미노산 수준 모두에서, 상기 언급된 핵산 중 하나와 또는, 단백질 또는 웨타이드 서열 중 하나와 전체 야생형 서열의 적어도 5%, 10%, 20%, 30%, 40%, 50%, 60%, 바람직하게는 적어도 70%, 더 바람직하게는 적어도 80%, 마찬가지로 더 바람직하게 적어도 85%, 더더욱 바람직하게 90% 및 가장 바람직하게는 적어도 95% 또는 97%, 98% 또는 99%, 서열 상동성을 갖는 서열을 포함할 수 있다.

[0210] 본 발명의 첫 번째 실시예의 더 바람직한 측면에서 독창적인 핵산 서열은 상기 독창적인 핵산 서열의 형질주입(transfection) 효율을 증가시키는 비히클, 형질주입 또는 착화제(complexation agent)와 관련되어 있다. 특히 형질주입 효율을 증가시키는데 적절한 본 발명과 관련된 바람직한 제제는 프로타민(protamine), 뉴클레오린(nucleoline), 스퍼민(spermine) 또는 스퍼미딘(spermidine)을 포함하는, 양이온 또는 다중양이온 화합물, 또는 폴리-L-리신 (PLL), 기본적인 폴리웨타이드, 올리고 또는 폴리-아르기닌, HIV-결합 웨타이드를 포함하는 세포 투과 웨타이드(cell penetrating peptides (CPPs)), HIV-1 Tat (HIV), Tat-유도 웨타이드, 페네트라틴(Penetratin), VP22 유도 또는 유사 웨타이드, HSV, VP22 (헤르페스(Herpes simplex)), MAP, KALA 또는 단백질 변환 도메인(PTDs), PpT620, 프롤린-풍부 웨타이드, 아르기닌-풍부 웨타이드, 리신-풍부 웨타이드, MPG 웨타이드, Pep-1, L-올리고머, 칼시토닌 웨타이드(Calcitonin peptide(s)), 안테나페디아-유도 웨타이드(Antennapedia-derived peptides) (특히, *Drosophila antennapedia*로부터 유래된), pAntp, pIsl, FGF, 락토페린(Lactoferrin) 트랜스포르탄(Transportan), Buforin-2, Bac715-24, SynB, SynB(1), pVEC, hCT-유도 웨타이드, SAP 또는 히스톤(histones)과 같은 다른 양이온 웨타이드 또는 단백질이다. 추가적으로, 바람직한 양이온 또는 다중양이온 단백질 또는 웨타이드는 다음의 전체 식을 갖는 다음의 단백질 또는 웨타이드로부터 선택될 수 있다: (Arg)<sub>1</sub>;(Lys)<sub>m</sub>;(His)<sub>n</sub>;(Orn)<sub>o</sub>;(Xaa)<sub>x</sub>에서, 만약 Arg, Lys, His 및 Orn의 전체 함량이 올리고웨타이드의 전체 아미노산의 적어도 50%를 나타낸다면, 상기  $1 + m + n + o + x = 8-15$ , 이고 1, m, n 또는 o 은 각각 독립적으로 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 또는 15로부터 선택된 어떤 숫자일 수 있으며; 그리고 Xaa는 Arg, Lys, His 또는 Orn를 제외한 천연의(=자연적으로 발생한) 또는 비-천연의 아미노산으로부터 선택된 어떤 아미노산일 수 있고; 그리고 만약 Xaa의 전체 함량이 올리고웨타이드의 전체 아미노산의 50%를 초과하지 않는다면, x는 0, 1, 2, 3 또는 4로부터 선택된 어떤 숫자일 수 있다. 특히 본 발명과 관련하여 바랄모직

한 양이온 웹타이드는 예를 들어, Arg<sub>7</sub>, Arg<sub>8</sub>, Arg<sub>9</sub>, H<sub>3</sub>R<sub>9</sub>, R<sub>9</sub>H<sub>3</sub>, H<sub>3</sub>R<sub>9</sub>H<sub>3</sub>, YSSR<sub>9</sub>SSY, (RKH)<sub>4</sub>, Y(RKH)<sub>2</sub>R, 등이다. 형질주입 제제로 사용될 수 있는 더 바람직한 양이온 또는 다중양이온 화합물은 예를 들어 키토산과 같은 다당류, 폴리브레(polybrene), 폴리에틸렌이민(polyethyleneimine (PEI))과 같은 양이온성 고분자, DOTMA: [1-(2,3-sioleyoxy)propyl]-N,N,N-트리메틸암모늄 클로라이드(trimethylammonium chloride), DMRIE, 다이(di)-C14-아미딘(amidine), DOTIM, SAINT, DC-Chol, BGTC, CTAP, DOPC, DODAP, DOPE: 디올레일 포스파티딜에탄올-아민(Dioleyl phosphatidylethanol-amine), DOSPA, DODAB, DOIC, DMEPC, DOGS: 다이옥타데실아미도글리실스페민(Dioctadecylamidoglycylspermin), DIMRI: 디미리스토-옥시프로필 다이메틸 하이드록시에틸 암모늄 브로마이드(Dimyristo-oxypropyl dimethyl hydroxyethyl ammonium bromide), DOTAP: 다이오레오일옥시-3-(트리메틸암모니오)프로판(dioleoyloxy-3-(trimethylammonio)propane), DC-6-14: 0,0-다이테트라데카노일-N-( $\alpha$ -트리메틸암모니오아세틸)다이에탄올아민 클로라이드(0,0-ditetradecanoyl-N-( $\alpha$ -trimethylammonioacetyl)diethanolamine chloride), CLIP1: rac-[(2,3-다이옥타데실옥시프로필)(2-하이드록시에틸)]-다이메틸암모니움 클로라이드(rac-[(2,3-dioctadecyloxypropyl)(2-hydroxyethyl)]-dimethylammonium chloride), CLIP6: rac-[2(2,3-다이헥사데실옥시프로필-옥시메틸옥시)에틸]트리메틸암모니움(rac-[2(2,3-dihexadecyloxypropyl-oxymethoxy)ethyl]trimethylammonium), CLIP9: rac-[2(2,3-다이헥사데실옥시프로필-옥시숙시닐옥시)에틸]-트리메틸암모니움(rac-[2(2,3-dihexadecyloxypropyl-oxysuccinyl)ethyl]-trimethylammonium), 올리고펙타민과 같은 양이온성 리피드 또는  $\beta$ -아미노산-고분자 또는 리버스 폴리아미드(reversed polyamides) 등과 같은 변형된 폴리아미노산, PVP (폴리(N-에틸-4-비닐파리디늄 브로마이드) 등과 같은 변형된 폴리에틸렌, pDMAEMA (폴리(다이메틸아미노에틸 메틸아크릴레이트(methylaminoethyl methylacrylate))) 등과 같은 변형된 아크릴레이트, pAMAM (폴리(아미도아민)) 등과 같은 변형된 아미도아민, 다이아민 말단 변형된 1,4 부탄 다이아크릴레이트-코-5-1-펜탄올(1,4-butanediol diacrylate-co-5-amino-1-pentanol) 고분자 등과 같은 변형된 폴리베타아미노에스터(polybetaaminoester (PBAE)), 폴리프로필아민 덴드리머(polypropylamine dendrimers) 또는 덴드리머 기반 pAMAM 등과 같은 덴드리머(dendrimers), PEI: 폴리(에틸렌이민), 폴리(프로필렌이민) 등과 같은 폴리이민(들), 폴리알릴아민(polyallylamine), 사이클로덱스트린(cyclodextrin) 기반 중합체, 덱스트란(dextran) 기반 중합체, 키토산 등과 같은 당류 백본(backbone) 기반 중합체, PMOXA-PDMS 코폴리머 등과 같은 실란(silan) 백본 기반 중합체, 하나 또는 더 많은 양이온 블럭(예를 들어, 상기 언급된 양이온 중합체로부터 선택된) 및 하나 또는 더 많은 친수성 또는 소수성 블럭(폴리에틸렌글리콜(polyethyleneglycole)과 같은)의 조합으로 이루어진 블럭고분자(blockpolymers); 등과 같은 양이온 또는 다중양이온 고분자를 포함할 수 있다.

[0211] 바람직한 것은 본 발명의 핵산 서열은 착화되지 않은(naked) 형태 또는, 예를 들어 화학적 구조가 무엇이든 이들의 다중양이온 화합물들, 바람직하게 다중양이온 (폴리)웹타이드 또는 합성된 다중양이온 화합물들에 의해서 착화된 형태로 제공된다. 바람직하게, 상기 핵산 서열은 패키징 세포(packaging cell)와 함께 제공되지 않는다.

[0212] 다른 실시예에 따르면, 본 발명은 또한 예를 들어, a) 본원에 정의된 독창적인 핵산을 준비하는 단계, b) 발현 시스템, 예를 들어, 무세포 발현계(cell-free expression system), 세포(예를 들어, 숙주 세포 또는 체세포 발현), 조직 또는 유기체와 같은 발현 시스템에 본원에 정의된 단백질 또는 웹타이드를 암호화하는 독창적인 핵산을 이용하거나 투여하는 단계를 포함하는 암호화된 단백질/웹타이드의 발현 수준을 증가시키는 방법을 제공한다. 상기 방법은 실험실 용으로, 연구를 위해, 진단, 웹타이드 또는 단백질의 상업적 생산 및/또는 치료 목적을 위해 적용될 수 있다. 이와 관련하여, 일반적으로 본원에 정의된 독창적인 핵산을 준비한 후에, 바람직하게는 착화되지 않은 형태로 또는 본원에 정의된 약학적 조성물 또는 백신으로써, 바람직하게는 형질주입을 통하여 또는 본원에 정의된 투여 방식들 중 어떤 것이든지 사용하여 일반적으로 무세포 발현계, 세포(예를 들어 숙주 세포 또는 체세포 발현과 같은), 조직 또는 유기체에 적용하거나 또는 투여된다. 상기 방법은 시험관 내, 생체 내 또는 생체 밖에서 수행될 수 있다. 게다가 상기 방법은 특정 질병, 바람직하게는 본원에 정의된 질병들의 치료와 관련하여 수행될 수 있다.

[0213] 이와 관련하여, 시험관 내는 유기체 바깥의 배지에서 세포내로 독창적인 핵산의 형질주입(transfection) 또는 형질전환(transduction)으로 본원에서 정의되며; 생체 내는 전체 유기체 또는 개체에 독창적인 핵산의 적용에 의해서 세포 안으로 독창적인 핵산의 형질주입 또는 형질전환으로 본원에서 정의되며 그리고 생체 밖은 유기체 또는 개체의 바깥에서 세포 안으로 독창적인 핵산의 형질주입 또는 형질전환 및 뒤이어 유기체 또는 개체에 형

질주입된 세포의 적용으로 본원에 정의된다.

- [0214] 마찬가지로, 다른 실시예에 따르면, 본 발명은 무세포 발현계, 세포(예를 들어 숙주 세포 또는 체세포 발현과 같은), 조직 또는 유기체에 예를 들어, 본원에 정의된 단백질 또는 웨타이드를 암호화하는 독창적인 핵산을 적용하거나 또는 투여함으로써 암호화된 단백질/웨타이드의 발현 수준을 증가시키기 위하여, 본원에 정의된 독창적인 핵산의 용도를 또한 제공하며, 바람직하게는 진단 또는 치료 목적을 위한 용도를 제공한다. 상기 용도는 실험실 용으로, 연구를 위해, 진단, 웨타이드 또는 단백질의 상업적 생산 및/또는 치료 목적을 위해 적용될 수 있다. 이와 관련하여, 본원에 정의된 독창적인 핵산을 준비한 이후, 일반적으로 무세포 발현계, 세포(예를 들어 숙주 세포 또는 체세포 발현과 같은), 조직 또는 유기체에 적용 또는 투여되며, 바람직하게는 착화되지 않은 형태로 또는 본원에 정의된 투여 방식들 중 어떤 것이든지 사용하여 일반적으로 무세포 발현계, 세포(예를 들어 숙주 세포 또는 체세포 발현과 같은), 조직 또는 유기체에 적용하거나 또는 투여된다. 상기 용도는 시험관 내, 생체 내 또는 생체 밖에서 수행될 수 있다. 게다가 상기 방법은 특정 질병, 바람직하게는 본원에 정의된 질병들의 치료와 관련하여 수행될 수 있다.
- [0215] 다른 실시예에서 본 발명은 또한 본 발명의 첫 번째 실시예에 따른 독창적인 핵산 또는 발현 벡터 또는プラス 미드를 포함하는 독창적인 발현 시스템과 관련된다. 이와 관련하여 발현 시스템은 무세포 발현계(예를 들어 시험관 내 전사/번역 시스템), 세포의 발현 시스템(예를 들어, CHO 세포와 같은 포유류 세포, 곤충 세포, 효모 세포, *E.coli*와 같은 박테리아성 세포) 또는 웨타이드 또는 단백질의 발현을 위해 이용되는 유기체(예를 들어, 식물 또는 소(cows)와 같은 동물들).
- [0216] 추가적으로, 다른 실시예에 따르면, 본 발명은 또한 암호화된 단백질/웨타이드의 발현 수준 증가를 위한 약학적 조성물의 제조를 위한, 예를 들어 본원에 정의된 질병의 치료를 위한, 예를 들어 세포(예를 들어 숙주 세포 또는 체세포 발현), 조직 또는 유기체에 본원에 정의된 독창적인 핵산을 적용 또는 투여하기 위해, 바람직하게는 착화되지 않은 형태 또는 착화된 형태로 또는 본원에 정의된 약학적 조성물 또는 백신으로써, 더바람직하게는 본원에 정의된 투여 방식 중 어떤 것이든지 이용한 본원에 정의된 독창적인 핵산의 용도와 관련된다.
- [0217] 따라서, 특히 바람직한 실시예에서, 본 발명은 또한 본원에 정의된 독창적인 핵산 및 선택적으로 약학적으로 허용가능한 담체 및/또는 비히클을 포함하는 약학적 조성물을 제공한다.
- [0218] 첫 번째 요소로써, 독창적인 약학적 조성물은 본원에 정의된 독창적인 핵산을 포함한다.
- [0219] 두 번째 요소로써, 독창적인 약학적 조성물은 적어도 하나의 추가적인 약학적으로 활성인 성분을 포함할 수 있다. 이와 관련하여 약학적으로 활성인 성분은 본원에 언급된 특정 질환, 바람직하게는 암 질환, 자가 면역 질환, 알러지 또는 전염성 질환, 심혈관 질환, 순환계 질환, 소화계 질환, 피부 질환, 근골격 질환, 결합 조직(connective tissue) 질환, 종양(neoplasms), 면역 결핍, 내분비, 영양 및 대사 질환, 신경증후 질환, 안 질환, 귀 질환 및 유전성 질환을 치료, 개선 또는 예방하는 치료적 효과를 가지는 화합물이다. 그러한 화합물들은 이들에 제한되지 않고, 바람직하게 본원에 정의된 바와 같이, 웨타이드 또는 단백질, 바람직하게 본원에 정의된 바와 같이 핵산, (치료적으로 활성인) 저분자량 유기 또는 무기 화합물(분자량 5000 미만, 바람직하게는 1000 미만), 당류, 바람직하게 본원에 정의된 바와 같이 항원 또는 항체, 해당 기술분야에 이미 알려진 치료적 제제, 항원성 세포, 항원성 세포 조각들, 세포 파편들; 세포벽 성분들(예를 들어, 다당류), 변형된, 절단된 또는 불-활성화된 (예를 들어 화학적으로 또는 방사선에 의해) 병원체(바이러스, 박테리아 등), 바람직하게 본원에 정의된 바와 같은 애쥬반트 등을 포함한다.
- [0220] 게다가, 상기 독창적인 약학적 조성물은 약학적으로 허용가능한 담체 및/또는 비히클을 포함할 수 있다. 본 발

명과 관련하여, 약학적으로 허용가능한 담체는 일반적으로 독창적인 약학적 조성물의 액체 또는 비-액체 베이스를 포함한다. 만약 상기 독창적인 약학적 조성물이 액체 형태로 제공된다면, 담체는 일반적으로 발열물질(pyrogen)이 없는 증류수; 예를 들어 등장성 식염수(isotonic saline) 또는 인산염, 시트르산염 등과 같은 완충용액일 것이다. 주입 버퍼는 특정 레퍼런스 배지를 참고하여 고장액, 등장액 또는 저장액일 수 있으며, 즉 상기 배지는 특정 레퍼런스 배지를 참고할 때 더 높거나, 동일하거나, 더 낮은 염 농도를 가질 수 있으며, 상기 바람직한 앞서 언급된 염의 농도는 삼투압 또는 다른 농도 영향 때문에 세포들에 손상을 가하지 않을 정도로 사용될 수 있다. 레퍼런스 배지는 예를 들어, 혈액, 럼프액, 세포질 액 또는 다른 체액과 같은 "생체 내" 방법에서 나타나는 액체들 또는 예를 들어 일반적인 완충액 또는 액체들과 같은 "시험관 내" 방법에서 레퍼런스 배지로 이용될 수 있는 액체들이다. 그러한 일반적인 완충액 또는 액체들은 통상의 기술자에게 알려져 있다. 링거액(Ringer-Lactate solution)은 액체 베이스로써 특히 바람직하다.

[0221] 그러나, 하나 또는 많은 호환 가능한 고체 또는 액체 충전물(fillers) 또는 희석제(diluents) 캡슐화 화합물(encapsulating compounds)은 치료가 필요한 환자에 투여되기 적합한 독창적인 약학적 조성물을 위해 또한 이용될 수 있다. 상기 용어 "호환 가능한"은 여기서 독창적인 약학적 조성물의 성분들이 일반적인 이용상태하에서 독창적인 약학적 조성물의 약학적 효과를 상당히 감소시킬 수 있는 어떤 상호 작용이 일어나지 않는 방식으로 본원에 정의된 독창적인 핵산과 혼합될 수 있다는 것을 의미한다.

[0222] 특정 측면에 따르면, 상기 독창적인 약학적 조성물은 애쥬반트를 포함할 수 있다. 이와 관련하여, 애쥬반트는 선천성 면역 시스템의 면역 반응, 즉, 비-특이적 면역 반응의 개시 또는 증가에 적절한 어떤 화합물로써 이해될 수 있다. 다른 말로, 투여되었을 때, 상기 독창적인 약학적 조성물은 선택적으로 거기에 포함된 애쥬반트로 인하여 선천성 면역 반응을 바람직하게 유도한다. 바람직하게, 그러한 애쥬반트는 통상의 기술자에게 잘 알려지고 본 사안에 적절한, 즉 포유류에서 선천성 면역 반응의 유도를 돋는, 예를 들어 상기 정의된 애쥬반트 단백질 또는 다음에 정의된 애쥬반트와 같은, 애쥬반트로부터 선택될 수 있다.

[0223] 특히 저장(depot) 및 수송에 적절한 애쥬반트로 바람직한 것은 비히클, 형질주입 또는 착화제로써 독창적인 핵산 서열로 상기 정의된 양이온 또는 다중양이온 화합물들이다.

[0224] 독창적인 약학적 조성물은 원활 경우, 그것의 면역원성(immunogenicity) 또는 면역자극적 능력의 증가를 위하여 하나 이상의 보조제(auxilliary substances)를 추가적으로 포함할 수 있다. 본원에 정의된 독창적인 핵산과 추가적으로 독창적인 약학적 조성물에 포함될 수 있는 보조제의 시너지 활성은 그로인해 얻을 수 있다. 보조제들의 다양한 종류에 따라, 다양한 메카니즘은 이러한 측면에서 고려될 수 있다. 예를 들어, 수지상 세포(dendritic cells (DCs))의 성숙분열을 가능하게 하는 화합물들, 예를 들어 지질다당류, TNF-알파 또는 CD40 리간드는 적절한 보조제의 첫 번째 부류를 이룬다. 일반적으로, 목적으로 하는 방식으로 면역 반응을 증가시키거나 향상시키도록 하는 "위험한 신호(danger signal)"(LPS, GP96, 등) 또는 GM-CFS와 같은, 사이토카인 방식으로 면역 시스템에 영향을 미치는 어떤 제제를 보조제로써 이용하는 것이 가능하다. 특히 바람직한 보조제는 IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-12, IL-13, IL-14, IL-15, IL-16, IL-17, IL-18, IL-19, IL-20, IL-21, IL-22, IL-23, IL-24, IL-25, IL-26, IL-27, IL-28, IL-29, IL-30, IL-31, IL-32, IL-33, IFN-알파, IFN-베타, IFN-감마, GM-CSF, G-CSF, M-CSF, LT-베타 또는 TNF-알파, hGH와 같은 성장인자 같은, 선천성 면역 반응을 더욱 증가시키는 모노카인, 림포카인, 인터루킨 또는 케모카인과 같은 사이토카인들이다.

[0225] 상기 독창적인 약학적 조성물에 포함될 수 있는 추가적인 첨가제들은 예를 들어 Tween<sup>®</sup>와 같은 유화제; 예를 들어 소듐 라우릴 설페이트(sodium lauryl sulfate)와 같은 습윤제; 착색제; 감미제(aste-imparting agents), 약학적 담체; 정제-형성제(tablet-forming agents); 안정제(stabilizers); 항산화제; 방부제이다.

[0226] 상기 독창적인 약학적 조성물은 또한 인간 톨-유사 수용체 Toll-like receptors TLR1, TLR2, TLR3, TLR4,

TLR5, TLR6, TLR7, TLR8, TLR9, TLR10에 (리간드로써) 그것의 결합 친화력 때문에 또는 주의 틀-유사 수용체 TLR1, TLR2, TLR3, TLR4, TLR5, TLR6, TLR7, TLR8, TLR9, TLR10, TLR11, TLR12 또는 TLR13에 (리간드로써) 그것의 결합 친화력 때문에 면역자극성이 될는 것으로 알려진 다른 화합물을 추가적으로 더 포함할 수 있다.

[0227] 상기 독창적인 약학적 조성물은 구강으로(orally), 비경구투여(parenterally), 흡입제 스프레이(inhalation spray)로, 국소적으로(topically), 직장으로(rectally), 코로(nasally), 구강으로(buccally), 질로(vaginally) 또는 이식 장치(implanted reservoir)를 통하여 투여될 수 있다. 본원에 사용된 용어 비경구적은 피하지방의(subcutaneous), 정맥내의(intravenous), 근육내의(intramuscular), 관절내의(intra-articular), 활액내의(intra-synovial), 흉골내의(intrasternal), 포막내의(intrathecal), 간장내의(intrahepatic), 상처내(intralesional), 두개내의(intracranial), 경피의(transdermal), 내피의(intradermal), 폐내의(intrapulmonary), 복막내의(intraperitoneal), 심장내의(intracardial), 동맥내의(intraarterial), 그리고 혀 아래(sublingual) 주사 또는 투입기술(infusion techniques)을 포함한다.

[0228] 바람직하게 상기 독창적인 약학적 조성물은 비경구적으로, 더 바람직하게는 피하지방의, 정맥내의, 근육내의, 관절내의, 활액내의, 흉골내의, 포막내의, 간장내의, 상처내의, 두개내의, 경피의, 내피의, 폐내의, 심장내의, 동맥내의, 그리고 혀 아래 주사 또는 투입 기술을 통하여 투여될 수 있다. 특히 바람직하게는, 내피의 그리고 근육내의 주사이다. 독창적인 약학적 조성물의 살균한 주사가능한 형태는 수성의(aqueous) 또는 유성의(oleaginous) 혼탁액일 수 있다. 이러한 혼탁액은 적절한 분산제 또는 습윤제 및 서스펜션화제를 이용하여 기술분야에서 잘 알려진 기술에 따라 만들어질 수 있다.

[0229] 본원에 정의된 독창적인 약학적 조성물은 캡슐, 정제, 수성 서스펜션(aqueous suspension) 또는 용액을 포함하나, 이들에 제한되지 않고, 경구적으로 허용가능한 투약 형태로 경구적으로 투여될 수 있다.

[0230] 상기 독창적인 약학적 조성물은 또한 국소적으로, 특히 치료의 목표가, 예를 들어 피부의 또는 다른 접근가능한 상피 조직의 질병을 포함한 국소 처리법(topical application)에 의해 쉽게 접근 가능한 부위 또는 조직을 포함할 때, 투여될 수 있다. 적절한 국부적 제제는 이러한 부위 또는 조직을 위해 쉽게 준비된다. 국소 처리법을 위하여, 상기 독창적인 약학적 조성물은 하나 이상의 담체에 띠 있거나 용해되어 있는 본원에 정의된 독창적인 핵산을 포함하는, 적절한 연고(ointment)의 형태로 제형화될 수 있다.

[0231] 독창적인 약학적 조성물은 일반적으로 독창적인 약학적 조성물의 성분들의, 특히 본원에 정의된 독창적인 핵산의 "안전하고 효과적인 양"을 포함한다. 본원에 사용된 바와 같이 "안전하고 효과적인 양"은 본원에 정의된 독창적인 핵산의 양이 본원에 정의된 질병 또는 장애의 긍정적인 변화를 유도하는데 충분히 효과적이라는 의미이다. 그러나 동시에 "안전하고 효과적인 양"은 심각한 부작용을 피하고 이점과 위험 사이에서 적절한 관계를 유지하기에 충분할 만큼 적다는 것이다. 일반적으로 이러한 한계를 결정하는 것은 합리적인 의학적 판단의 범위에 해당한다.

[0232] 독창적인 약학적 조성물은 인간을 위해 이용될 수 있고 또한, 수의용(veterinary medical) 목적으로 이용될 수 있으며, 바람직하게는 일반적인 약학적 조성물로써 또는 백신으로써, 인간의 의료용으로 이용될 수 있다.

[0233] 특히 바람직한 다른 실시예에 따르면, 독창적인 약학적 조성물(또는 본원에 정의된 독창적인 핵산)은 백신으로써 제공되거나 이용될 수 있다. 일반적으로, 그러한 백신은 약학적 조성물로 상기와 같이 정의된다. 추가적으로, 그러한 백신은 일반적으로, 상기 정의되어진 항원을 바람직하게는 암호화하는, 본원에 정의된 독창적인 핵산을 포함한다. 아니면 그 대신, 그러한 백신은 본원에 정의된 독창적인 핵산 및 추가적인 항원을 포함할 수 있고, 바람직하게는 단백질 또는 웹타이드로써 또는 본원에 정의된 바와 같이 항원을 암호화하는 핵산으로써 또는 본원에 정의된 어떤 항원의 형태로써 또는 이들의 모든 가능한 조합을 포함할 수 있다.

- [0234] 상기 독창적인 백신은 독창적인 약학적 조성물을 위하여 본원에 정의된 약학적으로 허용가능한 담체, 애쥬반트 및/또는 비히클을 또한 포함할 수 있다. 독창적인 백신과 특히 관련하여, 약학적으로 허용가능한 담체의 선택은 독창적인 백신이 투여되는 방식에 의해서 원칙적으로 결정된다. 예를 들어, 상기 독창적인 백신은 전신으로 (systemically) 또는 국부적으로(locally) 투여될 수 있다. 상기 독창적인 백신은 예를 들어, 온몸에 또는 국소적으로 투여될 수 있다. 일반적으로 전신에 영향을 주는 투여를 위한 경로는 예를 들어, 피하지방의, 정맥내의, 근육내의, 동맥의 그리고 복막내의 주입을 포함한, 경피적인, 경구적인, 비경구적인 경로 및 비강내의 투여 경로를 포함한다. 일반적으로 국소(local) 투여를 위한 경로는 국소적 투여 경로 뿐만 아니라 내피의, 경피의, 피하의 또는 근육내의 주입 또는 상처내, 두개내의, 폐내의, 심장내의, 그리고 혀 아래의 주입을 포함한다. 더 바람직하게 백신은 내피의, 피하지방의 또는 근육내 경로에 의해서 투여될 수 있다. 그러므로 독창적인 백신은 바람직하게 액체 형태 (또는 때때로 고체 형태로) 제제화된다.
- [0235] 상기 독창적인 백신은, 원활 경우, 그것의 면역원성 또는 면역자극적 능력을 증가시키기 위하여 추가적으로 하나 이상의 보조제를 포함할 수 있다. 특히 바람직한 것은 약학적 조성물을 위하여 보조제로써 애쥬반트 또는 정의되어진 첨가제들이다.
- [0236] 게다가 본 발명은 본원에 정의된 독창적인 핵산, 독창적인 약학적 조성물, 독창적인 백신, 상기 독창적인 백신을 포함하는 약학적 조성물과 백신들의 몇 가지 어플리케이션(applications) 및 용도 또는 이들을 포함하는 키트를 제공한다.
- [0237] 한 가지 특정 실시예에 따르면, 본 발명은 약물로써, 바람직하게는 유전자 치료 분야에서 면역자극 제제, 애쥬반트 또는 백신으로써 본원에 정의된 독창적인 핵산의 첫 번째 의학적 용도에 관한 것이다.
- [0238] 다른 실시예에 따르면, 본 발명은 본원에 정의된 질병들의 치료를 위한, 본원에 정의된 핵산의 두 번째 의학적 용도, 바람직하게는 본원에 정의된 독창적인 핵산의 용도에 관한 것이며, 본원에 정의된 다양한 질병들의 예방, 치료 및/또는 개선을 위한 약물의 제조를 위해 이들을 포함하는 약학적 조성물 또는 백신의 용도 또는 이들을 포함하는 키트의 용도에 관한 것이다. 바람직하게, 약학적 조성물 또는 백신은 이러한 목적을 위해 이들이 필요한 환자에게 이용되거나 투여될 수 있다.
- [0239] 바람직하게, 본원에 언급된 질병들은 암 또는 종양 질환, 전염성 질환, 바람직하게 (바이러스성, 박테리아성 또는 원생동물성) 전염성 질환, 자가면역 질환, 알러지 또는 알러지성 질환, 일원발생적 질환(monogenetic diseases), 즉, (유전적) 질환 또는 일반적으로 유전자(genetic) 질환, 유전되는 배경을 갖고 정의되는 유전자 결합에 의해 일반적으로 야기되는 질병 및 멘델의 유전 법칙에 의해 유전되는 질병, 심혈관 질환, 뉴런의 (neuronal) 질환, 호흡계 질환, 소화계 질환, 피부 질환, 근골격 질환, 결합 조직(connective tissue) 질환, 종양(neoplasms), 면역 결핍, 내분비, 영양 및 대사 질환, 안 질환, 귀 질환 및 본 발명에 의해 영향을 받을 수 있는 어떠한 질병으로부터 선택된다.
- [0240] 바람직하게는 상기 언급된 암 또는 종양 질환들은 바람직하게, 예를 들어, 대장암종(colon carcinomas), 흑색소 세포종(melanomas), 신장 암종(renal carcinomas), 림프종(lymphomas), 급성 골수성 백혈병(acute myeloid leukaemia (AML)), 급성 림프구성 백혈병(acute lymphoid leukaemia (ALL)), 만성 골수성 백혈병(chronic myeloid leukaemia (CML)), 만성 림프구성 백혈병(chronic lymphocytic leukaemia (CLL)), 위장관 종양(gastrointestinal tumors), 폐암(pulmonary carcinomas), 신경교종(gliomas), 갑상선 종양(thyroid tumours), 유선종양(mammary carcinomas), 전립선 종양(prostate tumours), 간종양(hepatomas), 예를 들어, 유두종(papilloma) 바이러스-유도 종양(자궁경부암 종양과 같은) 다양한 바이러스-유도 종양들, 선암(adenocarcinomas), 헤르페스 바이러스-유도 종양(버킷 림프종(Burkitt's lymphoma), EBV-유도 B-세포 림프

종)), B형 간염-유도 종양(hepatocell carcinoma), HTLV-1- 및 HTLV-2-유발 림프종, 청 신경종/신경초종(acoustic neuromas/neurinomas), 자궁경부암(cervical cancer), 폐암, 인두암(pharyngeal cancer), 항문암(anal carcinoma), 교도세포종(glioblastoma), 직장암종(rectal carcinoma), 성상세포종(astrocytoma), 뇌종양(brain tumors), 위암(stomach cancer), 망막아세포종(retinoblastoma), 기저세포암(basalioma), 뇌전이암(brain metastases), 수아세포종(medulloblastomas), 질암(vaginal cancer), 췌장암(pancreatic cancer), 정소암(testicular cancer), 흑색소 세포종(melanomas), 갑상선 종양(thyroidal carcinomas), 방광암(Bladder cancer), 호지킨 증후군(Hodgkin's syndrome), 슈네베르거 질병(Schneeberger disease), 기관지 암종(bronchial carcinoma), 뇌하수체 종양(hypophysis tumor), 균상식육종(Mycosis fungoides), 식도암(oesophageal cancer), 유방암(breast cancer), 유암종(carcinoids), 신경섬유초종(neurinomas), 스파날리오마(spinaliomas), 베킷 림프종(Burkitt's lymphoma), 후두암(laryngeal cancer), 신장암(renal cancer), 흉선종(thymoma), 자궁체부암(corpus carcinoma), 뼈암(bone cancer), 비-호지킨 림프종(non-Hodgkin's lymphomas), 요도암(urethral cancer), CUP 증후군(CUP syndrome), 머리/목 종양(head/neck tumors), 펩지교종(oligodendrogloma), 음문암(vulval cancer), 장암(intestinal cancer), 결장암(colon carcinoma), 식도암종(oesophageal carcinomas), 사마귀 관련(wart involvement), 소장의 종양(tumors of the small intestine), 두개인두종(craniopharyngeomas), 난소암(ovarian carcinomas), 연조직 종양 및 육종(Soft tissue tumours/sarcomas), 난소암(ovarian cancer), 간암(liver cancer), 췌장암종(pancreatic carcinoma), 자궁경부암종(cervical carcinoma), 자궁내막 암종(endometrial carcinoma), 간 전이 암(liver metastases), 음경암(penile cancer), 설암(tongue cancer), 담낭암(gall bladder cancer), 백혈병(leukaemia), 형질세포종(plasmacytoma), 자궁암(uterine cancer), 눈꺼풀 종양(lid tumor), 전립선암(prostate cancer) 등을 포함한다.

[0241]

추가적으로, 상기와 관련하여, 전염성 질병은 바람직하게, 독감(influenza), 말라리아(malaria), SARS, 황열병(yellow fever), AIDS, 라임 보렐리아증(Lyme borreliosis), 리슈마니아증(Leishmaniasis), 탄저병(anthrax), 수막염(meningitis), AIDS, 첨규사마귀(Condyloma acuminata), 야원 사마귀(hollow warts), 땅기열(Dengue fever), 삼일열(three-day fever), 에볼라 바이러스(Ebola virus), 오한(cold), 초여름 수막뇌염(early summer meningoencephalitis (FSME)), 유행성감기(flu), 대상포진(shingles), 간염(hepatitis), 헤르페스 바이러스 I형(herpes simplex type I), 헤르페스 바이러스 II형(herpes simplex type II), 대상포진(Herpes zoster), 인플루엔자(influenza), 일본뇌염(Japanese encephalitis), 라사열(Lassa fever), 마르부르크병 바이러스(Marburg virus), 홍역(measles), 구제역(foot-and-mouth disease), 단핵종(mononucleosis), 유행성 이하선염(mumps), 노워크 바이러스 감염(Norwalk virus infection), 파이퍼 샘열(Pfeiffer's glandular fever), 천연두(smallpox), 소아마비(polio (childhood lameness)), 가성-크룹병(pseudo-croup), 제5병(fifth disease), 공수병(rabies), 사마귀(warts), 서부 나일강 열(West Nile fever), 수두(chickenpox), 거대 세포성 바이러스(cytomegalic virus (CMV))와 같은 바이러스성 전염병, 유산(miscarriage (전립선 염증(prostate inflammation))), 탄저병(anthrax), 맹장염(appendicitis), 보렐리아(borreliosis), 보툴리누스 중독(botulism), 캄필로박ter(Camphylobacter), 클라미디아 트라코마티스(Chlamydia trachomatis (요도(urethra)의 염증, 결막염(conjunctivitis)), 콜레라(cholera), 디프테리아(diphtheria), 도나바노시스(donavanosis), 후두개염(epiglottitis), 발진티푸스(typhus fever), 가스 과저병(gas gangrene), 임질(gonorrhoea), 야토병(rabbit fever), 헬리코박ter 파일로리(Heliobacter pylori), 백일해(whooping cough), 기후성가래톳(climatic bubo), 골수염(osteomyelitis), 재향 군인병(Legionnaire's disease), 나병(leprosy), 리스테리아증(listeriosis), 폐렴(pneumonia), 뇌수막염(meningitis), 박테리아 수막염(bacterial meningitis), 탄저병(anthrax), 중이염(otitis media), 마이코플라즈마 호미니스(Mycoplasma hominis), 신생아 폐혈증(neonatal sepsis(흉모양막염( Chorioamnionitis))), 과저성 구내염(noma), 파라티푸스(paratyphus), 역병(plague), 라이터증후군(Reiter's syndrome), 로키산 흥반열(Rocky Mountain spotted fever), 살모넬라 파라티푸스(Salmonella paratyphus), 살모넬라 티푸스(Salmonella typhus), 성홍열(scarlet fever), 매독(syphilis), 파상풍(tetanus), 임질(tripper), 쓰쓰가무시병(tsutsugamushi disease), 결핵(tuberculosis), 티푸스(typhus), 질염(vaginitis (colpitis)), 연성 하감(soft chancre)과 같은 박테리아성 전염병, 그리고 아메바증(amoebiasis), 빌하르쯔주혈흡충증(bilharziosis), 샤가스병(Chagas disease), 효모균 반점(yeast fungus spots), 옴(scabies), 말라리아(malaria), 회선 사상충증(onchocercosis (river blindness)) 또는 곰팡이성 질병, 톡소플라스마증(toxoplasmosis), 질 트리코모나스증(trichomoniasis), 트리파노소마증(trypanosomiasis (sleeping sickness)), 내장의 레이쉬마니아증(visceral Leishmaniosis), 기저귀 발진(nappy/diaper

dermatitis), 주혈 흡충병(schistosomiasis), 어독(fish poisoning (Ciguatera)), 칸디다증(candidosis), 피부 리슈만판모충증(cutaneous Leishmaniosis), 람블편모충증(giardiasis), 또는 수면병(sleeping sickness)과 같은 기생충, 원생동물 또는 곰팡이에 의해 야기되는 전염성 질병으로부터 또는 포낭충(Echinococcus), 웬머리 촌충(fish tapeworm), 개 촌충(canine tapeworm), 이(lice), 소 촌충(bovine tapeworm), 돼지 촌충(porcine tapeworm), 작은 촌충(miniature tapeworm)에 의해 야기되는 전염성 질병으로부터 선택된다.

[0242]

상기와 관련하여, 게다가 알려지는 일반적으로 이들 항원 또는 알레르겐에 대한 국소적 또는 전신적 염증 반응의 결과이고 이들 알레르겐에 대항하여 신체 내에서 면역원을 이끌어낸다. 이와 관련한 알레르겐은 예를 들어, 풀, 꽃가루, 곰팡이, 약물 또는 다양한 환경 자극(triggers) 등을 포함한다. 가설에 얹매이지 않고, 몇 가지 다른 질병 메카니즘들이 알려지의 개발에 수반될 수 있는 것으로 지지된다. P. Gell and R. Coombs에 의한 분류 체계에 따르면, 용어 "알러지(allergy)"는 전통적인 IgE 메카니즘에 의해 야기되는 type I 과민성 알레르기로 제한된다. type I 과민성 알레르기는 IgE에 의한 비만 세포와 호염기 세포(basophils)의 과도한 활성이 특징이며, 계속적으로 콧물이 흐를만큼 온화한(benign) 증상을 나타내고, 생명에 위협적인 과민성 쇼크 그리고 죽음으로 이끌 수 있는 전신성 염증성 반응으로 결과를 나타낸다. 잘 알려진 알러지 타입들은 알러지성 천식(코의 점막을 부풀어 오르게 하는), 알러지 결막염(결막이 붉어지고 가렵게되는), 알레르기 비염("고초열"), 알레르기 반응, 혈관부종(angiodema), 아토피 피부염(atopic dermatitis (습진(eczema))), 두드러기(urticaria (hives)), 호산구 증가(eosinophilia), 호흡기성(respiratory), 콧从中에 쏘여서 발생하는 알러지, 피부 알러지(습진, 발진(두드러기) 그리고 (접촉성)피부염과 같은 다양한 발진을 포함하고 유도하는), 음식 알러지, 약물 알러지 등을 포함하며, 이들에 한정하지는 않는다. 본 발명과 관련하여, 단백질, 상기 정의된 발명의 핵산과 조합된 단백질 알레르겐을 암호화하는 핵산으로써 또는 독창적인 핵산으로써 알레르겐(예를 들어, 고양이 알레르겐, 면지 알레르겐, 진드기 항원, 식물 항원(예를 들어 자작나무 항원) 등으로부터 유래한) 알레르겐을 암호화하거나 포함하는, 예를 들어 독창적인 핵산, 약학적 조성물 또는 백신이 제공된다. 본 발명의 약학적 조성물은 더 강력한 TH1 반응으로 (과도한(exceeding)) 면역 반응을 이동할 수 있으며, 이로 인해 원하지 않는 IgE 반응을 억제하거나 저지할 수 있다.

[0243]

추가적으로, 자가면역 질병들은 대략적으로 각 질병들의 주요 임상병리적인 특징에 의존하는 전신성과 조직 특이적 또는 국소적 자가면역 장애로 나눌 수 있다. 자가면역 질환들은 전신성홍반성낭창(systemic lupus erythematosus(SLE)), 쇼그伦 증후군(Sjogren's syndrome), 피부 경화증(Scleroderma), 류마티스 관절염(Rheumatoid Arthritis) 및 다발성 근육염(polymyositis)을 포함한 전신성 증후군(systemic syndromes) 또는 내분비의(endocrinologic (DM Type 1, 갑상선염(Hashimoto's thyroiditis), 애디슨병(Addison's disease etc.)), 피부 외과의(dermatologic(심상성천포창(pemphigus vulgaris))), 혈액적(haematologic (자가면역성 용혈성 빈혈(autoimmune haemolytic anaemia))), 뉴런성 (neural (다발성 경화증(multiple sclerosis)))일 수 있는 국소적 증후군이거나 실질적으로 신체조직의 덩어리를 제한하는데(circumscribed) 관여할 수 있는 국소적 증후군으로 나눌 수 있다. 치료되는 자가면역 질병들은 예를 들어, 다발성 경화증(multiple sclerosis (MS)), 류마티스 관절염(rheumatoid arthritis), 당뇨병(diabetes), 제1형 당뇨병(type I diabetes (Diabetes mellitus)), 전신성홍반성낭창(SLE), 만성적 다발성 관절염(chronic polyarthritis), 바제도병(Basedow's disease), 만성 간염의 자가면역 형태(autoimmune forms of chronic hepatitis), 궤양성 대장염(colitis ulcerosa), 제1형 알레르기 질환(type I allergy diseases), 제2형 알레르기 질환(type II allergy diseases), 제3형 알레르기 질환(type III allergy diseases), 제4형 알레르기 질환(type IV allergy diseases), 섬유근육통(fibromyalgia), 탈모(hair loss), 베히테레브 병(Bechterew's disease), 크론병(Crohn's disease), 중증 근무력증(Myasthenia gravis), 아토피(neurodermitis), 다발성 근육통(Polymyalgia rheumatica), 진행성 전신 경화증(progressive systemic sclerosis (PSS)), 건선(psoriasis), 라이터증후군(Reiter's syndrome), 류마티스 관절염(rheumatic arthritis), 건선(psoriasis), 맥관염(vasculitis) 등 또는 제2형 당뇨병과 같은 제1형 자가면역 질환 또는 제2형 자가면역 질환 또는 제3형 자가면역 질환 또는 제4형 자가면역 질환으로 이루어진 그룹에서 선택될 수 있다. 왜 면역 시스템이 자가항원에 대항한 면역 반응을 유도하는지에 관한 정확한 방식은 지금까지 밝혀지지 않았지만, 몇 가지 병인학(etiology)에 관한 발견들이 존재한다. 그에 따라, 아마 자가반응은 T-세포 바이패스(Bypass)때문일 것이다. 전자가 많은 항체들을 생산하기 전에 정상적인 면역 시스템은 T-세포에 의한 B-세포의 활성을 요구한다. T-세포의 이러한 요구는 비-특이적 방식으로 T-세포 수용체의 어떤 서브유닛에

직접적으로 결합하는 B-cells 또는 심지어 T-cells의 다클론성 활성화를 개시하는 능력이 있는 슈퍼-항원을 생산하는 유기체들에 의한 감염과 같은 드문 경우에 견너될 수 있다. 다른 설명은 분자 모방(molecular mimicry)으로부터 자가면역 질병들을 밝힌다. 외인적인(exogenous) 항원은 특정한 숙주 항원과 구조적인 유사성을 공유할 수 있다; 따라서, 이러한 항원에 대항하여 생산된 어떤 항체(자가-항원들과 흡사한)는 또한 이론상으로는, 숙주 항원에 결합할 수 있고 면역 반응을 증폭시킬 수 있다. 분자 모방의 가장 두드러진 형태는 그룹 A 베타-용혈성 연쇄상구균에서 관찰되며, 이는 인간 심근과 공통되는 항원을 가지며 류마티스 열의 심장병 정후의 원인이 되는 것이다. 그러므로 본 발명은 일반적으로 면역 반응이 둔감해지도록 하는 자가항원(자가항원 단백질을 암호화하는 단백질, mRNA 또는 DNA 로씨)과 같은 것을 암호화하거나 포함하는 본원에 정의된 핵산 또는 약학적 조성물 또는 백신을 제공한다.

[0244] 추가적인 실시예에 따르면, 본 발명은 유전자 치료의 수단으로써 본원에 정의된 질병들의 치료를 위한, 본원에 정의된 독창적인 핵산의 두 번째 의학적 용도에 관한 것이다.

[0245] 더 바람직한 실시예에서, 상기 독창적인 핵산은 특히 본원에 정의된 목적을 위하여 약학적 조성물 또는 백신의 준비를 위해 이용될 수 있다.

[0246] 상기 독창적인 약학적 조성물 또는 백신은 게다가 본원에 정의된 질병 또는 질환들의 치료를 위해 이용될 수 있다.

[0247] 마지막 실시예에 따르면, 본 발명은 또한 키트, 특히 키트의 부분을 제공한다. 그러한 키트, 특히 부품들의 키트(kits of parts)는 일반적으로 단독으로 또는 본원에 정의된 다른 구성성분들을 본원에 정의된 적어도 하나의 독창적인 핵산, 독창적인 핵산을 포함하는 독창적인 약학적 조성물 또는 백신과 조합하여 구성성분으로써 포함한다. 본원에 정의된 적어도 하나의 독창적인 핵산은, 선택적으로 본원에 정의된 다른 성분들과 조합하여, 독창적인 약학적 조성물 및/또는 독창적인 백신의 키트의 한 곳 또는 다른 부분에서 나타날 수 있다. 예를 들어, 예를 들어 적어도 하나의 키트의 부분은 본원에 정의된 적어도 하나의 독창적인 핵산, 및 본원에 정의된 적어도 하나의 다른 성분으로써 적어도 하나의 다른 키트의 다른 부분을 포함할 수 있으며, 예를 들어 적어도 하나의 다른 키트의 부분은 적어도 하나의 약학적 조성물 또는 백신 또는 이들의 부분을 포함할 수 있으며, 예를 들어 적어도 하나의 키트의 부분은 본원에 정의된 독창적인 핵산, 본원에 정의된 적어도 하나의 다른 성분으로써 키트의 적어도 하나의 다른 부분, 독창적인 약학적 조성물 또는 백신의 적어도 하나의 성분으로써 키트의 적어도 하나의 부분 또는 전체로써 독창적인 약학적 조성물 또는 백신 및 적어도 하나의 키트의 다른 부분, 예를 들어 적어도 하나의 항원, 적어도 하나의 약학적 담체 또는 비히클 등을 포함할 수 있다. 키트 또는 부품들의 키트는 게다가 독창적인 핵산의 투여 및 주사에 대한 정보가 있는 기술 설명서, 독창적인 약학적 조성물 또는 독창적인 백신을 포함하며 또는 예를 들어 만약 키트가 부품들의 키트로 준비된다면, 어떤 구성성분들 또는 부품들의 투여 및 주사에 대한 정보가 있는 기술 설명서를 포함할 수 있다.

[0248] 본 발명에서, 언급되지 않았다면, 실시예들 및 다른 예들의 몇 가지 특징들은 적절한 곳에서, 서로 조합될 수 있다. 게다가, 상기 용어 "포함하는(comprising)"은 만약 구체적으로 언급되지 않았다면, "이루는(consisting of)"의 의미로 이해되어서는 안된다. 그러나, 본 발명과 관련하여, 용어 "comprising"은 적절한 곳에서, 용어 "이루는(consisting of)"으로 대체될 수 있다.

### 도면의 간단한 설명

[0249] 다음의 도면들은 본 발명을 실례로써 더욱 명확하게 하려고 하는 것이다. 이들은 본 발명의 기술적 사상을 한정 하려 하지 않는다.

도 1: 후생동물 및 원생동물의 스템 루프 서열로부터 만들어진 히스톤 스템-루프 공통서열(consensus sequence)을 나타낸다(as reported by Davila Lopez, M., & Samuelsson, T. (2008), RNA (New York, N.Y.),

14(1), 1-10. doi:10.1261/rna.782308). 후생동물 및 원생동물로부터 유래한 4001 히스톤 스템-루프 서열은 정렬되고, 나타나는 뉴클레오티드의 양은 스템-루프 서열 내의 모든 위치를 나타낸다. 모든 뉴클레오티드를 대표하는 형성된 공통 서열은 분석된 서열내에 존재하며, 이것은 단일-문자 (single-letter) 뉴클레오티드 코드를 이용하여 주어진다. 게다가 공통 서열에서, 서열은 분석된 서열 내에서 존재하는 뉴클레오티드의 적어도 99%, 95% 및 90%를 대표하는 것으로 보여진다.

**도 2:** 원생동물의 스템 루프 서열로부터 만들어진 히스톤 스템-루프 공통 서열을 나타낸다(as reported by Davila Lopez, M., & Samuelsson, T. (2008), RNA (New York, N.Y.), 14(1), 1-10. doi:10.1261/rna.782308). 원생동물로부터 유래한 131 히스톤 스템-루프 서열은 정렬되고, 나타나는 뉴클레오티드의 양은 스템-루프 서열 내의 모든 위치를 나타낸다. 모든 뉴클레오티드를 대표하는 형성된 공통 서열은 분석된 서열내에 존재하며, 이것은 단일-문자 (single-letter) 뉴클레오티드 코드를 이용하여 주어진다. 게다가 공통 서열에서, 서열은 분석된 서열내에서 존재하는 뉴클레오티드의 적어도 99%, 95% 및 90%를 대표하는 것으로 보여진다.

**도 3:** 후생동물의 스템 루프 서열로부터 만들어진 히스톤 스템-루프 서열을 나타낸다(as reported by Davila Lopez, M., & Samuelsson, T. (2008), RNA (New York, N.Y.), 14(1), 1-10. doi:10.1261/rna.782308). 후생동물로부터 유래한 3870 히스톤 스템-루프 서열은 정렬되고, 나타나는 뉴클레오티드의 양은 스템-루프 서열 내의 모든 위치를 나타낸다. 모든 뉴클레오티드를 대표하는 형성된 공통 서열은 분석된 서열내에 존재하며, 이것은 단일-문자 (single-letter) 뉴클레오티드 코드를 이용하여 주어진다. 게다가 공통 서열에서, 서열은 분석된 서열내에서 존재하는 뉴클레오티드의 적어도 99%, 95% 및 90%를 대표하는 것으로 보여진다.

**도 4:** 척추동물의 스템 루프 서열로부터 만들어진 히스톤 스템-루프 서열을 나타낸다 (as reported by Davila Lopez, M., & Samuelsson, T. (2008), RNA (New York, N.Y.), 14(1), 1-10. doi:10.1261/rna.782308). 척추동물로부터 유래한 1333 히스톤 스템-루프 서열은 정렬되고, 나타나는 뉴클레오티드의 양은 스템-루프 서열 내의 모든 위치를 나타낸다. 모든 뉴클레오티드를 대표하는 형성된 공통 서열은 분석된 서열내에 존재하며, 이것은 단일-문자 (single-letter) 뉴클레오티드 코드를 이용하여 주어진다. 게다가 공통 서열에서, 서열은 분석된 서열내에서 존재하는 뉴클레오티드의 적어도 99%, 95% 및 90% 를 대표하는 것으로 보여진다.

**도 5:** 사람의(*Homo sapiens*) 스템 루프 서열로부터 만들어진 히스톤 스템-루프 서열을 나타낸다 (as reported by Davila Lopez, M., & Samuelsson, T. (2008), RNA (New York, N.Y.), 14(1), 1-10. doi:10.1261/rna.782308). 사람으로부터 유래한 84 히스톤 스템-루프 서열은 정렬되고, 나타나는 뉴클레오티드의 양은 스템-루프 서열 내의 모든 위치를 나타낸다. 모든 뉴클레오티드를 대표하는 형성된 공통 서열은 분석된 서열내에 존재하며, 이것은 단일-문자 (single-letter) 뉴클레오티드 코드를 이용하여 주어진다. 게다가 공통 서열에서, 서열은 분석된 서열내에서 존재하는 뉴클레오티드의 적어도 99%, 95% 및 90% 를 대표하는 것으로 보여진다.

**도 6 내지 17:** 시험관 내 전사를 통한 mRNAs를 나타낸다.

시험관 내 전사에 의해 얻어진 mRNAs의 서열과 명칭이 주어졌다. 다음의 약어들이 사용되었다:

ppLuc (GC): *Photinus pyralis* 루시페라아제를 코딩하는 GC-풍부 mRNA 서열

ag: 알파 글로빈 유전자의 3'- 비번역 영역 (UTR)

A64: 64 아데닐산(adenylate)을 갖는 폴리(A)-서열

A120: 120 아데닐산(adenylate)을 갖는 폴리(A)-서열

히스톤(histone)SL: 히스톤 스템-루프

aCPSL: aCP-2KL 단백질의 특정 결합 라이브러리에서 선택된 스템 루프

폴리오(Polio)CL: 폴리오 바이러스 유전체 RNA로부터 유래한 5' 클로버 잎(5' clover leaf)

G30: 30 구아닐산(guanylates)을 갖는 폴리(G) 서열

U30: 30 우리딜산(uridylates)을 갖는 폴리(U) 서열

SL: 비특이적/인공적 스템-루프

N32: 비특이적 32 뉴클레오티드 서열

서열 내에서, 다음의 요소들은 강조하여 나타내었다: ppLuc(GC) ORF (대문자), ag (볼드체), 히스톤SL (밑줄), 테스트한 다른 독특한 서열(이탤릭체).

도 6: ppLuc(GC) - ag (SEQ ID NO: 43) 서열의 mRNA를 나타낸다.

알파-글로빈 3'-UTR (ag)이 즉시 뒤이어 나타나는 제한효소 절단 부위에서 오리지날 벡터의 선형화 (linearization)에 의해서, mRNA는 부족한 폴리(A) 서열이 얻어진다.

도 7: ppLuc(GC) - ag - A64 (SEQ ID NO: 44) 서열의 mRNA를 나타낸다.

A64 폴리(A)-서열이 즉시 뒤이어 나타나는 제한효소 절단 부위에서 오리지날 벡터의 선형화에 의해서, mRNA는 A64 폴리(A) 서열을 갖는 말단이 얻어진다.

도 8: ppLuc(GC) - ag - 히스톤SL (SEQ ID NO: 45) 서열의 mRNA를 나타낸다.

A64 폴리(A) 서열은 히스톤SL에 의해서 치환된다. 예로 사용된 히스톤 스템-루프 서열은 Cakmakci *et al.* (2008). Molecular and Cellular Biology, 28(3), 1182-1194.로부터 얻었다.

도 9: ppLuc(GC) - ag - A64 - 히스톤SL (SEQ ID NO: 46) 서열의 mRNA를 나타낸다.

히스톤SL은 A64 폴리(A)의 3'에 부가되었다.

도 10: ppLuc(GC) - ag - A120 (SEQ ID NO: 47) 서열의 mRNA를 나타낸다.

A64 폴리(A) 서열은 A120 폴리(A) 서열에 의해서 치환되었다.

도 11: ppLuc(GC) - ag - A64 - ag (SEQ ID NO: 48) 서열의 mRNA를 나타낸다.

두 번째 알파-글로빈 3'-UTR은 A64 폴리(A)의 3'에 부가되었다.

도 12: ppLuc(GC) - ag - A64 - aCPSL (SEQ ID NO: 49) 서열의 mRNA를 나타낸다.

스템 루프는 A64 폴리(A)의 3'에 부가되었다. 스템 루프는 aCP-2KL 단백질의 특정 결합을 위한 라이브러리로부터 선택된다 (Thisted *et al.*, (2001), The Journal of Biological Chemistry, 276(20), 17484-17496). aCP-2KL은 aCP-2의 동형 단백질(isoform)이고, 가장 강력하게 발현되는 aCP 단백질(알파-글로빈 mRNA 폴리(C) 결합 단백질) (Makeyev *et al.*, (2000), Genomics, 67(3), 301-316), 알파-글로빈 3'-UTR (Chkheidze *et al.*, (1999), Molecular and Cellular Biology, 19(7), 4572-4581)에 결합하는 RNA 결합 단백질의 그룹이다.

도 13: ppLuc(GC) - ag - A64 - CL (SEQ ID NO: 50) 서열의 mRNA를 나타낸다.

폴리오 바이러스 유전체 RNA로부터 유래한 5' 클로버 잎은 A64 폴리(A)의 3'에 부가되었다.

도 14: ppLuc(GC) - ag - A64 - G30 (SEQ ID NO: 51) 서열의 mRNA를 나타낸다.

30 구아닐산 구간은 A64 폴리(A)의 3'에 부가되었다.

도 15: ppLuc(GC) - ag - A64 - U30 (SEQ ID NO: 52) 서열의 mRNA를 나타낸다.

30 우리딜산 구간은 A64 폴리(A)의 3'에 부가되었다.

도 16: ppLuc(GC) - ag - A64 - SL (SEQ ID NO: 53) 서열의 mRNA를 나타낸다.

스템 루프는 A64 폴리(A)의 3'에 부가되었다. 스템의 상위 부분과 루프는 (Babendure *et al.*, (2006), RNA (New York, N.Y. 12(5), 851-861))로부터 얻었다. 스템 루프는 17 염기쌍 길이, CG-풍부 스템 및 6 염기 길이 루프로 이루어진다.

도 17: ppLuc(GC) - ag - A64 - N32 (SEQ ID NO: 54) 서열의 mRNA를 나타낸다.

다른 제한효소 절단 부위에서 오리지날 벡터의 선형화에 의해서, mRNA는 폴리(A) 다음에 추가적으로 32 뉴클레오티드를 가진다.

도 18: 폴리(A)와 히스톤SL의 조합은 시너지 효과로 mRNA로부터 단백질 발현을 증가시킴을 보여준다.

mRNA로부터 루시페라아제 발현에서 폴리(A) 서열, 히스톤SL 및 폴리(A)와 히스톤SL의 조합을 실험하였다. 따라

서 다른 mRNAs들이 HeLa cells 내로 전기천공(electroporated)되었다. 루시페라아제 수준은 형질주입후 6, 24, 그리고 48 시간에 측정하였다. 폴리(A) 서열 뿐만 아니라 히스톤SL도 갖기 않은 mRNA로부터 적은 루시페라아제가 발현되었다. 폴리(A) 서열 또는 히스톤SL이 루시페라아제 발현을 증가시킨다. 그러나, 폴리(A) 및 히스톤SL의 조합은 각각의 요소들로 측정한 수준 이상으로 더 강력하게 루시페라아제 수준을 증가시켰으며, 따라서 시너지효과를 나타낸다. 그러나 폴리(A) 및 히스톤SL의 조합은 더 강력하게 루시페라아제 발현 수준, 각각의 구성요소들을 가지고 관찰한 다양한 수준을 증가시키고, 따라서 시너지효과를 가지고 작용한다. 데이터는 3번 형질주입하여 RLU ± SD (relative light units ± standard deviation)로 나타내었다. 구체적인 RLU는 실시예 11.2에 정리하였다.

**도 19:** 폴리(A)와 히스톤SL의 조합은 그들의 배열과 무관하게 mRNA로부터 단백질 발현을 증가시킨다.

폴리(A) 서열, 히스톤SL, 폴리(A) 서열 및 히스톤SL의 조합과 그들의 배열에 대한 mRNA로부터 루시페라아제 발현에 대한 영향을 시험하였다. 각각의 mRNAs는 HeLa cells 내로 리포크션하였다. 루시페라아제 수준은 형질주입을 시작한 후 6, 24, 및 48 시간에 측정하였다. A64 폴리(A) 서열 또는 히스톤SL은 유사한 루시페라아제 수준을 보여주었다. A64에서 A120로 또는 A300으로 폴리(A) 서열의 길이 증가는 루시페라아제 발현을 적절하게 증가시킨다. 대조적으로, 폴리(A) 및 히스톤SL의 조합은 폴리(A) 서열을 연장하는 것보다 훨씬 더 많이 루시페라아제 수준을 증가시킨다. 폴리(A) 및 히스톤SL의 조합은 그것이 각각의 요소들을 가지고 측정된 수준 이상으로 루시페라아제 발현을 증가시키는 것처럼 시너지효과를 나타낸다. 폴리(A) 및 히스톤SL 조합의 시너지 효과는 폴리(A) 및 히스톤SL의 조합의 배열과는 무관하고 A64-히스톤SL 또는 히스톤SL-A250 mRNA를 갖는 폴리(A) 서열과는 무관한 것으로 보인다. 데이터는 3번 형질주입하여 RLU ± SD로 나타내었다. 구체적인 RLU는 실시예 11.3에 정리하였다.

**도 20:** 폴리(A) 및 히스톤SL의 조합에 의한 단백질 발현의 증가는 특이적임을 보여준다.

mRNA로부터 루시페라아제 발현에 대한 폴리(A) 및 히스톤SL의 조합 또는 폴리(A) 및 다른 서열들의 조합의 영향을 시험하였다. 따라서 다른 mRNAs가 HeLa 세포 안으로 전기천공되었다. 루시페라아제 수준은 형질주입 후 6, 24 및 48 시간 경과 후 측정하였다. 폴리(A) 서열 또는 히스톤SL은 유사하게 루시페라아제 증가를 나타냈었다. 폴리(A) 및 히스톤SL의 조합은 각각의 요소들을 가지고 측정한 수준 이상으로 루시페라아제 수준을 강력하게 증가시키고, 따라서 시너지 효과를 가진다. 대조적으로 다른 서열들 중에서 어떤 것과 폴리(A)를 조합한 것은 단지 폴리(A) 서열만 포함하는 mRNA와 비교해서 루시페라아제 수준에 어떤 영향도 미치지 않는다. 따라서, 폴리(A) 및 히스톤SL의 조합은 특이적으로(specifically) 그리고 시너지적으로(synergistically) 작용한다. 데이터는 3번 형질주입하여 RLU ± SD로 나타내었다. 구체적인 RLU는 실시예 11.4에 정리하였다.

**도 21:** 폴리(A) 및 히스톤SL의 조합은 생체 내에서 동반 상승효과를 가지고 mRNA로부터 단백질 발현을 증가시키는 것을 보여준다.

생체 내에서 mRNA로부터 루시페라아제 발현에 대한 폴리(A) 서열, 히스톤SL, 그리고 폴리(A)와 히스톤SL의 조합의 영향을 시험하였다. 그러므로 다른 mRNA들이 생쥐 피부 안으로 주입되었다. 생쥐들은 주입 후 16 시간 후 죽였고, 주입 부위에서 루시페라아제 발현을 측정하였다. 폴리(A) 서열 또는 히스톤SL을 갖는 mRNA로부터 루시페라아제 발현을 측정하였다. 그러나, 폴리(A) 및 히스톤SL의 조합은 각각의 요소들을 가지고 측정한 수준 이상으로 루시페라아제 수준을 강력하게 증가시키고, 따라서 시너지 효과를 가진다. 데이터는 3번 형질주입하여 RLU ± SD (relative light units ± standard deviation)로 나타내었다. 구체적인 RLU는 실시예 11.5에 정리하였다.

**도 22:** 폴리(A) 및 히스톤SL의 조합은 mRNA로부터 NY-ESO-1 단백질 발현을 증가시킨다.

mRNA로부터 NY-ESO-1 발현에 대한 폴리(A) 서열 및 폴리(A)와 히스톤SL의 조합의 영향을 시험하였다. 따라서 다른 mRNAs가 HeLa 세포 안으로 전기천공되었다. NY-ESO-1 수준은 유세포분석기(flow cytometry)에 의해서 형질주입 후 24시간 경과 후에 측정하였다. NY-ESO-1은 폴리(A) 서열만 갖는 mRNA로부터 발현되었다. 그러나, 폴리(A)와 히스톤SL의 조합은 오로지 폴리(A) 서열만을 가지고 측정한 수준 이상으로, NY-ESO-1 수준을 강력하게 증가시켰다. 데이터는 형광강도(fluorescence intensity)를 측정하여 그래프로 나타내었다. Median fluorescence intensities (MFI)는 실시예 11.6에 정리하였다.

**도 23:** 폴리(A) 및 히스톤SL의 조합은 mRNA를 이용한 예방접종에 의해 유발되는 항체의 수준을 증가시킴을 보여준다.

mRNA를 이용한 예방접종에 의해 유발되는 항 NY-ESO-1의 유도에 대한 폴리(A) 서열 및 폴리(A)와 히스톤SL의 조

합의 영향을 시험하였다. 따라서 C57BL/6 쥐는 프로타민으로 착화된 다른 mRNA를 이용해 괴부 안으로 예방접종하였다. 예방접종된 그리고 대조군 쥐에서 NY-ESO-1 특이적 항체의 수준은 순차적인 혈청의 희석액을 이용하여 ELISA 분석을 하였다. 항 NY-ESO-1 IgG2a[b]은 오로지 폴리(A) 서열만 갖는 mRNA에 의해서 유도된다. 그러나, 폴리(A)와 히스톤SL의 조합은 오로지 폴리(A) 서열만을 가지고 측정한 수준 이상으로, 항 NY-ESO-1 IgG2a[b] 수준을 강력하게 증가시킨다. 데이터는 측정 임계 역가(endpoint titers)를 이용해 그래프로 나타내었다. 측정 임계 역가는 실시예 11.7에 정리하였다.

### 발명을 실시하기 위한 구체적인 내용

[0250] 다음의 실시예들은 본 발명을 더욱 상세히 설명하고자 하는 것이다. 이들은 이들에 의해 본 발명의 기술적 사상을 제한하지 않는다.

[0251] 1. 히스톤-스템-루프 공통서열의 발생

[0252] 실험하기 앞서, 히스톤 스템-루프 공통서열은 후생동물 및 원생동물의 히스톤 스템-루프 서열에 근거하여 결정되었다. 서열들은 유전체 서열 및 발현된 서열 태그를 연구함으로써 자연의 수 많은 히스톤 스템-루프 서열들을 발견한, Lopez *et al.* (Davila Lopez, M., & Samuelsson, T. (2008), RNA (New York, N.Y.), 14(1), 1-10. doi:10.1261/rna.782308)에 의해 제공된 서플리먼트(supplement)로부터 얻었다. 첫 번째, 후생동물 및 원생동물로부터 얻은 모든 서열들(4001 서열), 또는 원생동물로부터 얻은(131 서열) 또는 아니면, 후생동물로부터 얻은(3870 서열), 또는 척추동물로부터 얻은(1333 서열) 또는 사람으로부터 얻은(84 서열) 모든 서열들은 그룹화되고 정렬되었다. 그 다음, 나타나는 뉴클레오티드의 양은 모든 위치를 위해 결정되었다. 얻어진 표에 근거하여, 5개의 다른 서열 그룹으로 공통서열은 분석된 서열에 나타나는 모든 뉴클레오티드를 나타낸다. 게다가, 더욱 제한적인 공통서열이 또한 얻어지며, 더욱더 보존된 뉴클레오티드를 강조한다.

[0253] 2. DNA-템플레이트(templates)의 준비

[0254] 시험관 내 전사를 위한 벡터는 *Photinus pyralis* 루시페라아제 (ppLuc(GC))를 코딩하는 GC-풍부 서열에 이어 T7 프로모터, 알파-글로빈 (ag)의 3' 비번역 영역(UTR)의 중심 부분, 그리고 폴리(A) 서열을 포함하여 만들어졌다. 폴리(A) 서열은 A64 폴리(A) 서열에서 mRNA 말단을 얻기 위하여 시험관 내 전사 전에 벡터의 선형화를 위해 사용되는 제한효소 절단 부위에 뒤이어 즉시 나타난다. 따라서 시험관 내 전사에 의한 이러한 벡터로부터 얻어지는 mRNA는 "ppLuc(GC) - ag - A64"로 표기된다.

[0255] 시험관 내 전사 이전에 다른 제한효소 절단 부위에서 이러한 벡터의 선형화는 A64의 3'에 추가적인 뉴클레오티드에 의해 확장된 mRNA를 얻거나 또는 A64가 없는 mRNA를 얻었다. 게다가, 올리지날 벡터는 다른 서열들을 포함하도록 변형되었다. 요약하자면, 다음의 mRNAs는 시험관 내 전사에 의해 이러한 벡터들로부터 얻어지는 것들이다.(mRNA 서열은 도 6 내지 17에 주어진다):

[0256]	ppLuc(GC) - ag	(SEQ ID NO: 43)
[0257]	ppLuc(GC) - ag - A64	(SEQ ID NO: 44)
[0258]	ppLuc(GC) - ag - histoneSL	(SEQ ID NO: 45)
[0259]	ppLuc(GC) - ag - A64 - histoneSL	(SEQ ID NO: 46)
[0260]	ppLuc(GC) - ag - A120	(SEQ ID NO: 47)
[0261]	ppLuc(GC) - ag - A64 - ag	(SEQ ID NO: 48)
[0262]	ppLuc(GC) - ag - A64 - aCPSL	(SEQ ID NO: 49)
[0263]	ppLuc(GC) - ag - A64 - PolioCL	(SEQ ID NO: 50)
[0264]	ppLuc(GC) - ag - A64 - G30	(SEQ ID NO: 51)

[0265]	ppLuc(GC) - ag - A64 - U30	(SEQ ID NO: 52)
[0266]	ppLuc(GC) - ag - A64 - SL	(SEQ ID NO: 53)
[0267]	ppLuc(GC) - ag - A64 - N32	(SEQ ID NO: 54)

### [0268] 3. 시험관 내 전사

[0269] 실시예 2에 따른 DNA-템플레이트는 T7-중합효소를 이용하여 시험관 내에서 선형화되고 전사된다. DNA-템플레이트는 그다음 DNase-처리에 의해서 절단된다. 모든 mRNA-전사체(transcript)는 전사 반응에 N7-메틸-구아노신-5'-트리포스페이트-5'-구아노신을 과량 추가하여 얻어진 5'-CAP 구조를 포함하였다. 따라서 획득한 mRNA는 정제되었고 중류수내에서 재부유(resuspende)시켰다.

### [0270] 4. mRNA의 효소 아데닐화(Enzymatic adenylation)

[0271] 두 개의 mRNAs가 효소처리에 의해서 아데닐화된다:

[0272] ppLuc(GC) - ag - A64 그리고 ppLuc(GC) - ag - 히스톤SL.

[0273] 이들을 종결시키기 위해, RNA는 제조자의 지시에 따라서 *E. coli* 폴리(A)-중합효소 및 ATP (Poly(A) Polymerase Tailing Kit, Epicentre, Madison, USA)로 인큐베이터하였다. 연장된 폴리(A) 서열을 갖는 mRNAs는 정제되고 중류수에서 재부유되었다. 폴리(A) 서열의 길이는 아가로스 겔 전기영동에 의해서 결정되었다. 처음 mRNAs는 대략 250 아데닐산에 의해 연장되었고, 얻어진 mRNAs는 다음과 같이 표기된다.

[0274] 각각, ppLuc(GC) - ag - A300 그리고 ppLuc(GC) - ag - 히스톤SL - A250.

### [0275] 5. mRNA 전기천공법에 의한 루시페라아제 발현

[0276] HeLa 세포는 트립신화되고 opti-MEM에서 세척되었다. 200  $\mu$ l opti-MEM에서  $1 \times 10^5$  세포를 각각은 0.5  $\mu$ g ppLuc-암호화 mRNA로 전기천공되었다. 대조적으로, ppLuc을 암호화하지 않는 mRNA는 각각 전기천공되었다. 전기천공된 세포들은 1 ml RPMI 1640 배지에서 24-웰 플레이트 내에 접종되었다. 형질주입 6,24, 또는 48 시간 후, 배지를 뺄아들이고, 세포는 200  $\mu$ l 용해 버퍼(25 mM Tris, pH 7.5 (HC1), 2 mM EDTA, 10% glycerol, 1% Triton X-100, 2 mM DTT, 1 mM PMSF)에서 용해되었다. 용해물은 ppLuc 활성이 측정될 때까지 -20°C에서 저장되었다.

### [0277] 6. mRNA 리포펙션에 의한 루시페라아제 발현

[0278] HeLa 세포는 웨일당  $2 \times 10^4$  세포의 농도로 96 웰 플레이트에 접종되었다. 다음날, 세포는 opti-MEM에서 세척되었고, 150  $\mu$ l opti-MEM에서 0.25  $\mu$ g 리포펙틴(Lipofectin)-착화된 ppLuc-암호화 mRNA로 형질주입하였다. 대조적으로, ppLuc를 코딩하지 않는 mRNA는 각각 리포펙션하였다. 몇몇 웰에서, opti-MEM은 뺄아들여지고, 세포들은 형질주입 시작 6시간 후에 200  $\mu$ l 용해 버퍼에서 용해되었다. 그때 남아있는 웰에서, opti-MEM은 RPMI 1640 배지로 교환되었다. 이를 웰에서, 배지는 뺄아들여지고, 세포들은 형질주입 시작 24 또는 48 시간 후에 200  $\mu$ l 용해 버퍼에서 용해되었다. 용해물은 ppLuc 활성이 측정될 때까지 -20°C에서 저장되었다.

### [0279] 7. 루시페라아제 측정

[0280] ppLuc 활성은 50  $\mu$ l 용해물 및 200  $\mu$ l 루시페린 버퍼 (25 mM Glycylglycin, pH 7.8 (NaOH), 15 mM MgSO<sub>4</sub>, 2 mM ATP, 75  $\mu$ M luciferin)를 이용하여 측정 시간 5초에서 BioTek SynergyHT 플레이트 리더에서 상대적 광단위(relative light unit (RLU))로 측정된다. 구체적인 RLU는 전체 RLU에서 대조군 RNA의 RLU를 뺌으로써 계산된다.

## [0281] 8. 내피 mRNA 주입에 의한 투시페라아제 발현 (생체 내 투시페라아제 발현)

[0282] 생쥐는 럼풀(Rompun) 및 케타빗(Ketavet)의 혼합물로 마취시켰다. 각각의 ppLuc-암호화 mRNA는 피부 내로 주입되었다(주입시 50  $\mu\text{l}$  내에 mRNA 0.5  $\mu\text{g}$ ). 대조적으로, ppLuc를 코딩하지 않는 mRNA는 각각 주입되었다. 주입 16시간 경과 후, 생쥐로부터 조직을 채취하였다. 조직 샘플은 액체 질소로 급속 냉동하였고 800 $\mu\text{l}$  용해 버퍼(25 mM Tris, pH 7.5 (HCl), 2 mM EDTA, 10% glycerol, 1% Triton X-100, 2 mM DTT, 1 mM PMSF)에서 조직 용해제 (Qiagen)로 용해하였다. 그 후에 샘플은 10분 동안 4°C에서 13500 rpm으로 원심분리하였다. 용해물들은 ppLuc 활성이 측정될 때까지 -80°C에서 저장되었다(참고 7. 투시페라아제 측정).

## [0283] 9. mRNA 전기천공법에 의한 NY-ESO-1 발현

[0284] HeLa 세포는 트립신화되고 opti-MEM에서 세척되었다. 200  $\mu\text{l}$  opti-MEM내의  $2 \times 10^5$  세포들은 10  $\mu\text{g}$  NY-ESO-1-암호화 mRNA로 전기천공되었다. 세번의 전기천공법을 통한 세포들은 혼합되고 2 ml RPMI 1640 배지에서 6-웰 플레이트 내에 접종되었다. 형질주입 24시간 후, 세포들은 채취하고 96-웰 V-형 밑바닥 플레이트로 옮겼다(2 wells per mRNA). 세포들은 인산완충식염수(phosphate buffered saline(PBS))로 세척하였고, Cytofix/Cytoperm (Becton Dickinson (BD)) 웰당 200  $\mu\text{l}$ 로 침투시켰다. 15분 후에, 세포들은 PermWash (BD)로 세척되었다. 그 다음, 세포들은 생쥐 anti-NY-ESO-1 IgG1 또는 아이소타입 대조군(isotype control) (20  $\mu\text{g}/\text{ml}$ )으로 실온에서 1시간 동안 배양하였다. 세포들은 PermWash로 다시 두번 씩 세척하였다. 다음으로, 세포들은 1:500 비율로 회석된, goat-anti-mouse IgG으로 커플링된 Alexa-647으로 4°C에서 1시간 동안 배양되었다. 마지막으로, 세포들은 PermWash로 두 번 세척하였다. 세포들은 200 $\mu\text{l}$  버퍼(PBS, 2% FCS, 2 mM EDTA, 0.01% sodium azide)에서 재부유하였다. NY-ESO-1 발현은 형광강도(MFI)로써 유세포분석기에 의해 정량화되었다.

## [0285] 10. mRNA로 예방접종에 의한 anti NY-ESO-1 항체의 유도

[0286] C57BL/6 mice의 피부 내로 프로타민으로 착화된 NY-ESO-1-암호화 mRNA를 예방접종하였다(14일 동안 5번). 대조군 생쥐는 버퍼로 처리하였다. 예방접종한 쥐와 대조군 쥐에서 NY-ESO-1-특이적 항체들의 수준은 마지막 접종후 8일 뒤에 ELISA에 의해 분석되었고: 96 웰 ELISA 플레이트(Nunc)는 4°C에서 16시간 동안 웰당 100  $\mu\text{l}$ 로 10  $\mu\text{g}/\text{ml}$  재조합 NY-ESO-1 단백질을 덮었다. 플레이트는 세척 버퍼(PBS, 0.05% Tween-20)로 두번 세척하였다. 비특이적 결합을 막기 위하여, 플레이트는 그 다음 블라킹 버퍼(blocking buffer(PBS, 0.05% Tween-20, 1% BSA))로 2시간 동안 37°C에서 배양하였다. 블라킹 이후에, 웰당 100  $\mu\text{l}$ 의 순차적으로 회석된 쥐 혈청이 추가되었고 실온에서 4시간 동안 배양되었다. 플레이트는 그 다음 세척 버퍼로 세 번 세척되었다. 다음으로, 웰당 100 $\mu\text{l}$ 의 블라킹 버퍼에서 1:600으로 회석된 비오틴 결합된 rat anti-mouse IgG2a[b] 검출 항체(BD Biosciences)는 실온에서 1시간 동안 결합되도록 하였다. 플레이트는 다시 세척 버퍼로 세번 세척되었고, 웰당 100 $\mu\text{l}$ 의 겨자무과산화효소-결합된 스트렙트아비딘(horseradish peroxidase-coupled streptavidin)으로 실온에서 30분 동안 배양하였다. 세척 버퍼로 4번 세척한 후에, 웰당 100  $\mu\text{l}$ 의 3,3',5,5'-테트라메틸벤지딘(tetramethylbenzidine) (Thermo Scientific)이 추가되었다. 색깔이 변하고 웰당 100  $\mu\text{l}$ 의 20% 황산이 추가되었다. 흡광도는 405nm에서 측정하였다.

## [0287] 11. 결과

## [0288] 11.1 히스톤 스템-루프 서열:

[0289] 히스톤 스템-루프 서열을 규정하기 위하여, 후생동물 및 원생동물로부터(4001 sequences), 또는 원생동물로부터(131 서열) 또는 그 대신 후생동물로부터(3870 서열), 또는 척추동물로부터(1333 서열) 또는 인간으로부터(84 서열) 유래한 서열들을 그룹화하고 정렬하였다. 그 다음, 존재하는 뉴클레오티드의 양은 모든 위치를 위해 측정되었다. 이와 같이 획득한 표에 근거하여, 서열들의 5가지 다른 그룹의 공통서열들이 만들어졌고, 이들은 분석된 서열들에 나타나는 모든 뉴클레오티드를 대표한다. 조합된 후생동물 및 원생동물의 공통 서열 내에서, 염기 쌍을 형성하는, 3 뉴클레오티드들, 루프 내의 T/U, 그리고 스템 내의 G 와 C은 보존되었다. 구조적으로, 6 염기-쌍 스템 및 4 뉴클레오티드 루프가 만들어지는 것이 일반적이다. 그러나, 변형된 구조들이 일반적이다: 84 사람 히스톤 스템-루프의, 둘은 4 4 염기-쌍 및 하나의 미스매치를 포함하는 오직 5 뉴클레오티드의 스템을 포함한다. 다른 사람 히스톤 스템-루프는 오직 5 염기-쌍의 스템을 포함한다. 4개의 더 많은 사람 히스톤 스템-

루프는 6 뉴클레오티드 길이 스템을 포함하지만, 각각, 세계의 다른 위치에서, 하나의 미스매치를 포함한다. 게다가, 네 개의 사람 히스톤 스템-루프는 각각, 두 개의 다른 위치들에서 하나의 위를 염기-쌍을 포함한다. 루프에 관해서는, 5 뉴클레오티드의 루프가 *D. discoideum*에서 발견되어진 것처럼, 4 뉴클레오티드의 길이가 염격하게 요구되는 것처럼 보이지 않는다.

- [0290] 공통 서열과 더불어 모든 뉴클레오티드를 대표하는 것이 분석된 서열에 나타나며, 더 제한적인 공통서열이 또한 얻어지며, 더욱 더 보존된 뉴클레오티드를 강조한다. 요약하면, 다음 서열들이 얻어진다:

[0291] (Cons): 전체 뉴클레오티드 존재를 대표한다

[0292] (99%): 모든 뉴클레오티드 존재의 적어도 99%를 대표한다

[0293] (95%): 모든 뉴클레오티드 존재의 적어도 95%를 대표한다

[0294] (90%): 모든 뉴클레오티드 존재의 적어도 90%를 대표한다

[0295] 히스톤 스템-루프 서열의 분석 결과는 다음 표 1 내지 5에 요약하였다(도 1 내지 5 참조):

豆 2

- [0296] 후생동물 및 원생동물의 히스톤 스템-루프 공통 서열:(4001 후생동물 및 원생동물의 히스톤 스템-루프 서열의 배열을 기초로 함) (도 1 참조)

					<	<	<	<	<				>	>	>	>	>									
#	222	158	307	287	128	184	0	13	12	9	1	47	59	0	675	381	195	159	523	0	14	372	61	771	201	
A	4	6	5	2	4										8	6	6	6			7	7	2	4	9	9
#	172	188	47	205	19	6	0	569	162	199	394	383	370	400	182	1	21	15	11	0	179	8	64	557	201	
T									0	7	0	0	1	1										6	9	0
#	155	221	875	918	267	270	0	339	234	378	51	119	227	0	314	7	50	31	16	400	354	154	387	263	174	
C	7	1			5			4	2	3					0		1	3	1	0	0	0	0	6	7	4
#	25	16	4	6	23	354	400	1	25	27	10	2	5	11	0	4	175	373	235	345	0	265	112	4	37	43
G						1	1									5	9	1	1	0				1	3	8
C	N*	N*	N	N	N	N	G	N	N	N	N	N	N	T	N	N	N	N	N	C	N	N	N	N*	N*	
9	H*	H*	H	H	V	V	G	Y	Y	Y	Y	H	H	T	H	R	V	V	R	C	B	V	H	H*	N*	N*
9	M*	H*	M	H	M	S	G	Y	Y	Y	T	T	Y	T	M	A	R	R	R	C	S	M	C	H*	H*	H*
9	M*	M*	M	M	M	S	G	Y	Y	C	T	T	T	T	M	A	G	R	R	C	S	A	C	H*	M*	H*

丑 3

- [0297] 원생동물의 히스톤 스템-루프 공통 서열:(131 원생동물의 히스톤 스템-루프 서열의 배열을 기초로 함) (도 2  
참조)

						<	<	<	<	<	<				>	>	>	>	>	>						
# A	52	32	71	82	76	13	0	12	12	9	1	46	3	0	75	82	53	79	20	0	4	94	17	35	74	56
# T	20	32	37	21	8	3	0	21	85	58	86	70	65	131	28	1	17	13	10	0	15	7	31	32	20	28
# C	45	59	20	25	38	0	0	86	8	54	42	13	58	0	27	2	6	31	10	131	112	5	82	58	30	40
# G	14	8	3	3	9	115	131	12	26	10	2	2	5	0	1	46	55	8	91	0	0	25	1	6	7	7
C ons	N*	N*	N	N	N	D	G	N	N	N	N	N	N	T	N	N	N	N	C	H	N	N	N*	N*	N*	
99%	N*	N*	N	N	N	D	G	N	N	N	B	N	N	T	H	V	N	N	C	H	N	H	N*	N*	N*	
95%	N*	N*	H	H	N	R	G	N	N	N	Y	H	B	T	H	R	D	N	N	C	Y	D	H	H*	N*	N*
90%	N*	H*	H	H	V	R	G	N	D	B	Y	H	Y	T	H	R	D	H	N	C	Y	R	H	H*	H*	H*

## 표 4

[0298]

후생동물의 히스톤 스템-루프 공통 서열: (3870 원생동물의 히스톤 스템-루프 서열의 배열을 기초로 함(1333 척 추동물 서열을 포함)) (도 3 참조)

						<	<	<	<	<	<				>	>	>	>	>	>							
# A	21	15	30	27	12	17	0	1	0	0	0	1	56	0	60	37	14	15	50	0	10	36	44	73	19	24	
	72	54	04	90	08	1								0	36	2	17	3	3	0	33	36	6	38	43		
# T	15	15	10	18	11	3	0	54	15	14	38	37	36	38	15	0	4	2	1	0	16	1	33	52	18	66	
	2	6		4																	4	5	1	2			
# C	15	21	85	89	26	27	0	33	23	37	9	10	16	0	31	5	44	0	6	38	34	14	37	25	17	63	
	12	52	5	3	37	0									13				70	31	9	88	78	14	14		
# G	11	8	1	3	14	34	38	13	1	0	0	3	6	0	3	12	36	23	33	0	26	87	3	31	36	13	1
																9	80	51	60		5						
C ons	N*	N*	N	N	N	N	G	N	B	Y	Y	N	N	T	N	V	N	D	N	C	N	N	N*	N*	N*		
99%	H*	H*	M	H	M	V	G	Y	Y	Y	T	Y	H	T	H	R	V	R	R	C	B	V	M	H*	H*	N*	
95%	M*	M*	M	M	M	S	G	Y	Y	C	T	T	Y	T	M	A	G	R	R	C	S	M	C	H*	H*	H*	
90%	M*	M*	M	M	M	S	G	Y	Y	C	T	T	T	T	M	A	G	R	R	C	S	A	C	H*	M*	H*	

## 표 5

						<	<	<	<	<	<	*	*	*	*	*	*	*	*	*	*	*	*	*	*	
# A	661	146	1315	1323	920	8	0	1	0	0	0	1	4	0	441	1333	0	1199	21	0	1	1126	26	81	880	960
# T	63	121	2	2	6	2	0	39	1217	2	1331	1329	1207	1333	30	0	1	0	1	0	2	1	22	91	91	12
# C	601	1062	16	6	403	1	0	1293	116	1331	2	0	121	0	862	0	2	0	0	1333	1328	128	1284	1143	834	361
# G	8	4	0	2	4	1322	1333	0	0	0	0	3	1	0	0	0	1330	134	1311	0	2	78	1	18	28	0
Cons	N*	N*	H	N	N	N	G	H	Y	Y	D	N	T	H	A	B	R	D	C	N	N	N*	N*	H*		
99%	H*	H*	M	A	M	G	G	Y	Y	C	T	T	Y	T	H	A	G	R	R	C	C	V	H	N*	N*	M*
95%	H*	H*	A	A	M	G	G	C	Y	C	T	T	Y	T	M	A	G	R	G	C	C	V	C	H*	H*	M*
90%	M*	M*	A	A	M	G	G	C	T	C	T	T	T	T	M	A	G	R	G	C	C	M	C	Y*	M*	M*

[0299]

[0300] 척추동물의 히스톤 스템-루프 공통 서열: (1333 척추동물의 히스톤 스템-루프 서열의 배열을 기초로 함) (도 4 참조)

豆 6

					<	<	<	<	<	•	•	•	•	>	>	>	>	>			
# A	10	17	84	84	76	1	0	1	0	0	0	1	0	0	12	84	0	65	3	0	0
# T	8	6	0	0	2	2	0	1	67	0	84	80	81	84	5	0	0	0	0	0	4
# C	62	61	0	0	6	0	0	82	17	84	0	0	3	0	67	0	1	0	0	84	5
# G	4	0	0	0	0	81	84	0	0	0	0	5	0	0	0	0	83	19	81	0	0
Cons	N*	H*	A	A	H	D	G	H	Y	C	T	D	Y	T	H	A	S	R	R	C	C
99%	N*	H*	A	A	H	D	G	H	Y	C	T	D	Y	T	H	A	S	R	R	C	C
95%	H*	H*	A	A	M	G	G	C	Y	C	T	T	T	T	H	A	G	R	G	C	C
90%	H*	M*	A	A	A	G	G	C	Y	C	T	T	T	T	M	A	G	R	G	C	C

[0301]

[0302] 호모 사피엔스(*Homo sapiens*) 히스톤 스템-루프 consensus sequence: (84 사람 히스톤 스템-루프 서열의 배열에 기초함) (도 5 참조)

五 7

[0303]

abbreviation	Nucleotide bases	remark
G	G	Guanine
A	A	Adenine
T	T	Thymine
U	U	Uracile
C	C	Cytosine
R	G or A	Purine
Y	T/U or C	Pyrimidine
M	A or C	Amino
K	G or T/U	Keto
S	G or C	Strong (3H bonds)
W	A or T/U	Weak (2H bonds)
H	A or C or T/U	Not G
B	G or T/U or C	Not A
V	G or C or A	Not T/U
D	G or A or T/U	Not C
N	G or C or T/U or A	Any base
*	present or not	Base may be present or not

[0304] 11.2 폴리(A)와 히스톤SL의 조합은 시너지효과를 가지고 mRNA에서 단백질 발현을 증가시킨다.

[0305] 폴리(A) 및 히스톤SL의 조합이 mRNA에서 단백질 발현에 미치는 효과를 조사하기 위하여, 알파-글로빈 3'-UTR의 3'DP 서로 다른 서열들을 갖는 mRNA들이 합성되었다: 3'-UTR의 3' 그냥 종결되고, 따라서 폴리(A) 서열과 히스톤SL 모두가 부족한 mRNA, 또는 A64 폴리(A) 서열 또는 대신 히스톤SL이 포함된 mRNA, 또는 3'-UTR의 3'에 A64 폴리(A)와 히스톤SL이 모두 포함된 mRNA. 루시페라아제-암호화 mRNA 또는 대조군 mRNA들은 HeLa 세포 안으로 전기천공되었다. 루시페라아제 수준은 형질주입 6, 24, 및 48 시간 후에 측정하였다(다음의 표 8 및 도 18 참조).

표 8

mRNA	RLU at 6 hours	RLU at 24 hours	RLU at 48 hours
ppLuc(GC)-ag-A64-histoneSL	466553	375169	70735
ppLuc(GC)-ag-histoneSL	50947	3022	84
ppLuc(GC)-ag-A64	10471	19529	4364
ppLuc(GC)-ag	997	217	42

[0306] [0307] 폴리(A) 서열도 히스톤SL도 갖지 않는 mRNA에서 적은 루시페라아제가 발현되었다. 폴리(A) 서열 또는 히스톤SL은 모두 유사한 정도로 루시페라아제 수준을 증가시켰다. 이들 mRNA 모두는 폴리(A) 및 히스톤SL이 모두 부족한 mRNA에서 보다 훨씬 더 많이 루시페라아제 수준을 증가시켰다. 그러나, 폴리(A)와 히스톤SL의 조합은 각각의 요소들을 가지고 측정한 수준 이상 더욱 더 많이 루시페라아제 수준을 증가시켰다. 같은 mRNA에서 폴리(A)와 히스톤SL의 조합 때문에 루시페라아제 수준이 증가되는 정도는 그들이 동반상승효과를 가지고 작용한다는 것을 증명한다.

[0308] [0309] 폴리(A)와 히스톤SL 사이에서의 시너지는 poly(A)-히스톤SL mRNA (+/+)로부터 나온 신호를 히스톤SL mRNA (-/+)와 poly(A) mRNA (+/-)로부터 나온 신호들을 더한 전체 신호들로 나누어 정량화하였다(다음의 표 9 참조)

표 9

	A64	히스톤SL	RLU at 6 hours	RLU at 24 hours	RLU at 48 hours
	+	+	466553	375169	70735
		+	50947	3022	84
	+		10471	19529	4364
Synergy			7.6	16.6	15.9

[0310] 따라서 계산된 팩터는 폴리(A)와 히스톤SL이 조합된 mRNA로부터의 루시페라아제 수준은 폴리(A)와 히스톤SL의 영향이 순수하게 추가된다면 기대되는 것보다 더 많은 루시페라아제 발현 수준을 보이는 것을 나타낸다. 폴리(A)와 히스톤SL이 조합된 mRNA로부터의 루시페라아제 수준은 그들의 효과가 순수하게 추가되는 것 보다 16.6배 더 높았다. 이러한 결과는 폴리(A)와 히스톤SL의 조합은 단백질 발현에서 분명히 동반상승 효과를 가지고 증가시키는 것임을 확실히 보여준다.

[0311] 11.3 폴리(A)와 히스톤SL의 조합은 그들의 배열과는 무관하게 mRNA로부터 단백질 발현을 증가시킨다.

[0312] 폴리(A)와 히스톤SL의 조합의 효과는 폴리(A) 서열의 길이 및 폴리(A)와 히스톤SL의 배열에 의존할 것으로 생각되었다. 따라서, 증가된 폴리(A) 서열 길이를 갖는 mRNAs와 반대로 배열된 폴리(A) 서열과 히스톤SL을 갖는 mRNAs를 합성하였다: 두 mRNAs는 3'UTR의 3'에 A120 또는 A300 폴리(A) 서열을 포함하였다. 다른 하나의 mRNAs는 3'UTR의 3'에 처음 A250 폴리(A) 서열 다음에 히스톤SL을 포함하였다. 루시페라아제-암호화 mRNAs 또는 대조군 mRNAs는 HeLa 세포 안으로 리포택션되었다. 루시페라아제 수준은 형질주입을 시작한 6, 24 및 48 시간이 지난 후에 측정하였다(다음의 표 10 및 도 19 참조).

표 10

mRNA	RLU at 6 hours	RLU at 24 hours	RLU at 48 hours
ppLuc(GC)-ag-histoneSL-A250	98472	734222	146479
ppLuc(GC)-ag-A64-histoneSL	123674	317343	89579
ppLuc(GC)-ag-histoneSL	7291	4565	916

ppLuc(GC)-ag-A300	4357	38560	11829
ppLuc(GC)-ag-A120	4371	45929	10142
ppLuc(GC)-ag-A64	1928	26781	537

[0314]

A64 폴리(A) 서열 또는 히스톤SL은 유사하게 루시페라아제 수준을 증가시켰다. 이전의 실험에서처럼 A64와 히스톤SL의 조합은 각각의 요소들로 측정한 수준 이상으로 강력하게 루시페라아제 수준을 증가시켰다. 같은 mRNA에서 폴리(A)와 히스톤SL의 조합으로 인한 루시페라아제 수준의 증가 정도는 그들이 동반상승 효과를 가지고 작용한다는 것을 증명한다. A64와 히스톤SL 사이에서 시너지는 A64-히스톤SL, A64, 및 히스톤SL mRNA의 루시페라아제 수준에 근거하여 앞서 실시한 것처럼 정량화하였다(다음의 표 11 참조). 폴리(A)와 히스톤SL의 영향이 순수하게 추가된다면, 그것보다 A64와 히스톤SL이 조합된 mRNA에서 루시페라아제 수준은 61.7배 더 높게 나타났다.

표 11

	A64	히스톤SL	RLU at 6 hours	RLU at 24 hours	RLU at 48 hours
	+	+	123674	317343	89579
		+	7291	4565	916
	+		1928	26781	537
Synergy			13.4	10.1	61.7

[0315]

대조적으로, A64에서 A120 또는 A300까지 폴리(A) 서열의 길이 증가는 루시페라아제 수준을 적절히 증가시켰다(표 9 및 도 19 참조). 가장 긴 폴리(A) 서열, A300을 갖는 mRNA는 히스톤SL과 결합된 유사한 길이의 폴리(A) 서열을 갖는 mRNA, 히스톤SL-A250와 또한 비교하였다. 긴 폴리(A) 서열을 갖는 것 이외에, 히스톤SL과 폴리(A)의 배열은 A64-히스톤SL mRNA와 관련된 이러한 mRNA에서 뒤집어진다. A250 및 히스톤SL의 조합은 히스톤SL 또는 A300으로 측정한 수준 이상으로 루시페라아제 수준을 강력하게 증가시켰다. 다시, A250 와 히스톤SL 사이의 시너지는 히스톤SL-A250 mRNA로부터의 RLU와 A300 mRNA 더하기 히스톤SL mRNA로부터의 RLU를 비교하여 이전에 한 것처럼 정량화하였다(다음의 표 12 참조). 폴리(A)와 히스톤SL의 영향이 순수하게 추가된다면, 그것보다 A250와 히스톤SL이 조합된 mRNA에서 루시페라아제 수준은 17.0배 더 높게 나타났다.

표 12

	히스톤SL	A250/A300	RLU at 6 hours	RLU at 24 hours	RLU at 48 hours
	+	+	98472	734222	146479
	+		7291	4565	916
		+	4357	38560	11829
Synergy			8.5	17.0	11.5

[0318]

요약하면, mRNA에서 단백질 발현에 대한 폴리(A)와 히스톤SL의 조합의 높은 동반상승 효과는 폴리(A)의 다른 길이와 폴리(A)와 히스톤SL의 배열과는 무관함을 증명한다.

[0319]

11.4 폴리(A)와 히스톤SL의 조합에 의한 단백질 발현 증가는 특이적이다.

[0320]

mRNA에서 단백질 발현에 대한 폴리(A)와 히스톤SL의 조합의 효과가 특이적인지를 조사하기 위하여, 폴리(A)가 결합된 다른 서열들을 갖는 mRNA가 합성되었다: 이러한 mRNA는 A64의 3'에 각각 7개의 별개의 서열 중 하나를 포함하였다. 루시페라아제-암호화 mRNA 또는 대조군 mRNA는 HeLa 세포 안으로 전기천공되었다. 루시페라아제 수준은 형질주입 후에 6, 24 및 48 시간 후에 측정하였다(다음의 표 13 및 도 20 참조).

표 13

mRNA	RLU at 6 hours	RLU at 24 hours	RLU at 48 hours
ppLuc(GC)-ag-A64-N32	33501	38979	2641
ppLuc(GC)-ag-A64-SL	28176	20364	874
ppLuc(GC)-ag-A64-U30	41632	54676	3408
ppLuc(GC)-ag-A64-G30	46763	49210	3382
ppLuc(GC)-ag-A64-PolioCL	46428	26090	1655
ppLuc(GC)-ag-A64-aCPSL	34176	53090	3338
ppLuc(GC)-ag-A64-ag	18534	18194	989
ppLuc(GC)-ag-A64-히스톤SL	282677	437543	69292
ppLuc(GC)-ag-히스톤SL	27597	3171	0
ppLuc(GC)-ag-A64	14339	48414	9357

[0321] 폴리(A) 서열 또는 히스톤SL는 유사하게 루시페라아제 수준을 증가시켰다. 다시, 폴리(A)와 히스톤SL의 조합은 각각의 요소들로 측정한 수준 이상으로 강력하게 루시페라아제 수준을 증가시켰고, 따라서 동반상승 효과를 가지고 작용하였다. 대조적으로, 다른 서열들 중 어떤 것과 조합된 폴리(A)는 오로지 폴리(A) 서열을 포함하는 mRNA과 비교하여 루시페라아제 수준에 영향이 없었다. 따라서, 폴리(A)와 히스톤SL의 조합은 시너지 효과를 가지고 mRNA에서 단백질 발현을 증가시키고 이러한 효과는 특이적이다.

[0323] 11.5 폴리(A)와 히스톤SL의 조합은 생체 내에서 동반상승 효과를 가지고 mRNA에서 단백질 발현을 증가시킨다.

[0324] mRNA에서 단백질 발현에 대한 생체 내에서 폴리(A)와 히스톤SL의 조합의 효과를 조사하기 위하여, 알파-글로빈 3' UTR의 3'에 다른 서열을 갖는 루시페라아제-암호화 mRNAs 또는 대조군 mRNA가 생쥐의 피부 안으로 투여되었다: mRNAs는 A64 폴리(A) 서열 또는 히스톤SL 또는 A64 폴리(A)와 히스톤SL을 3' UTR의 3'에 포함하였다. 루시페라아제 수준은 투여 16 시간 후에 측정하였다(표 14 및 도 21 참조)

표 14

mRNA	RLU at 16 hours
ppLuc(GC)-ag-A64-히스톤SL	38081
ppLuc(GC)-ag-히스톤SL	137
ppLuc(GC)-ag-A64	4607

[0326] 루시페라아제는 히스톤SL 또는 폴리(A) 서열을 갖는 mRNA로부터 발현되었다. 그러나, 폴리(A)와 히스톤SL의 조합은 각각의 요소들로 측정한 수준 이상으로 강력하게 루시페라아제 수준을 증가시켰다. 동일한 mRNA에서 폴리(A)와 히스톤SL의 조합으로 인한 루시페라아제 수준의 증가 정도는 그들이 시너지 효과를 가지고 작용했다는 것을 증명한다.

[0327] 폴리(A)와 히스톤SL 사이에서의 시너지는 poly(A)-히스톤SL mRNA (+/+)로부터 나온 신호를 히스톤SL mRNA (-/+)와 poly(A) mRNA (+/-)로부터 나온 신호들을 더한 전체 신호들로 나누어 정량화하였다(다음의 표 15 참조)

표 15

[0328]

	A64	히스톤SL	RLU at 16 hours
	+	+	38081
		+	137
	+		4607
Synergy			8.0

[0329]

따라서 계산된 팩터는 폴리(A)와 히스톤SL이 조합된 mRNA로부터의 루시페라아제 수준은 폴리(A)와 히스톤SL의 영향이 순수하게 추가된다면 기대되는 것보다 더 많은 루시페라아제 발현 수준을 보이는 것을 나타낸다. 폴리(A)와 히스톤SL이 조합된 mRNA로부터의 루시페라아제 수준은 그들의 효과가 순수하게 추가되는 것보다 8배 더 높았다. 이러한 결과는 폴리(A)와 히스톤SL의 조합은 생체 내 단백질 발현에서 분명히 동반상승 효과를 가지고 증가시키는 것임을 확실히 보여준다.

[0330]

11.6 폴리(A)와 히스톤SL의 조합은 mRNA에서 NY-ESO-1 단백질 발현을 증가시킨다.

[0331]

mRNA에서 단백질 발현에 대한 폴리(A)와 히스톤SL의 조합의 효과를 조사하기 위하여, 알파-글로빈 3' UTR의 3'에 다른 서열을 갖는 NY-ESO-1-암호화 mRNAs를 합성하였다: mRNAs는 A64 폴리(A) 서열을 포함하거나 또는 3'UTR의 3'에 A64 폴리(A)와 히스톤SL을 모두 포함하였다. NY-ESO-1-암호화 mRNAs는 HeLa 세포 안으로 전기전공되었다. NY-ESO-1 수준은 유세포분석기에 의해 형질주입한 후 24시간 이후에 측정하였다(표 16 및 도 22 참조).

표 16

[0332]

mRNA	MFI at 24 hours	
	anti-NY-ESO-1	isotype control
NY-ESO-1(GC)-ag-A64-히스톤SL	15600	1831
NY-ESO-1(GC)-ag-A64	1294	849

[0333]

NY-ESO-1는 오로지 폴리(A) 서열만을 갖는 mRNA로부터 발현되었다. 그러나, 폴리(A)와 히스톤SL의 조합은 오로지 폴리(A) 서열만으로 측정한 수준 이상으로 NY-ESO-1 수준을 강하게 증가시켰다.

[0334]

11.7 폴리(A)와 히스톤SL의 조합은 mRNA로 예방접종함으로써 유발된 항체의 수준을 증가시켰다.

[0335]

mRNA로 예방접종함으로써 유발된 항체의 유도에 대한 폴리(A)와 히스톤SL의 조합의 효과를 조사하기 위하여, C57BL/6 쥐는 알파-글로빈 3'UTR의 3'에 다른 서열들을 갖는, 프로타민-착화된 NY-ESO-1-암호화 mRNAs로 피부내 예방접종하였다. mRNAs는 3'UTR의 3'에 A64 폴리(A) 서열을 포함하거나 또는 A64 폴리(A) 와 히스톤SL을 모두 포함하였다. 예방접종된 쥐와 대조군 쥐에서 NY-ESO-1-특이적 항체의 수준은 단계적으로 회색된 혈청을 이용하여 ELISA로 분석하였다(표 17 및 도 23 참조)

표 17

[0336]

mRNA	mean IgG2a[b] endpoint titer
NY-ESO-1(GC)-ag-A64-histoneSL	763
NY-ESO-1(GC)-ag-A64	20

[0337]

항 NY-ESO-1 IgG2a[b]는 오로지 폴리(A) 서열을 갖는 mRNA에 의해서 유도되었다. 그러나, 폴리(A)와 히스톤SL의 조합은 오로지 폴리(A) 서열만으로 측정한 수준 이상으로 항 NY-ESO-1 IgG2a[b] 수준을 강하게 증가시켰다.

## 도면

## 도면1

		1										2																
		<					<					<					<					>						
				<					<					<					<					>				
#A	2224	1536	3075	2872	1284	184	0	13	12	9	1	47	59	0	675	3818	195	1595	523	0	14	3727	61	771	2012	2499		
#T	172	168	47	205	19	6	0	569	1620	199	3947	3830	3704	4001	182	1	21	15	11	0	179	8	64	557	201	690		
#C	1537	2211	875	918	2075	270	0	3394	2342	3753	51	119	227	0	3749	7	50	31	16	4001	3563	154	380	2036	1744	674		
#G	25	16	4	6	23	3541	4001	25	27	10	2	5	11	0	4	175	3735	2359	3451	0	265	112	4	37	43	138		
Cos	N*	N	N	N	N	G	N	N	N	N	N	N	N	N	N	N	N	N	C	N	N	N*	N*	N*	N*			
99%	H*	H*	H	H	V	V	G	Y	Y	Y	H	H	T	H	R	V	R	C	B	V	H	H*	N*	N*	N*			
95%	M*	H*	M	H	M	S	G	Y	Y	T	T	Y	T	M	A	R	R	C	S	M	C	H*	H*	H*	H*			
90%	M*	M*	M	M	S	G	Y	Y	C	T	T	T	T	M	A	G	R	R	C	S	A	C	H*	M*	H*			

## 도면2

### 도면3

## 도면4

## 도면5

				<	<	<	<	<	*	*	*	*	>	>	>	>	>		
#A	10	17	84	84	76	1	0	1	0	0	0	1	0	0	12	84	0	65	3
#T	8	6	0	0	2	2	0	1	67	0	84	80	81	84	5	0	0	0	0
#C	62	61	0	0	6	0	0	82	17	84	0	0	3	0	67	0	1	0	84
#G	4	0	0	0	0	81	84	0	0	0	0	0	3	0	0	0	83	19	81
Cors	N*	H*	A	A	H	D	G	H	Y	C	T	D	Y	T	H	A	S	R	R
99%	N*	H*	A	A	H	D	G	H	Y	C	T	D	Y	T	H	A	S	R	R
95%	H*	H*	A	A	M	G	G	C	Y	C	T	T	T	T	H	A	G	R	G
90%	H*	M*	A	A	G	G	C	Y	C	T	T	T	T	T	M	A	G	R	G

## 도면6

## ppLuc(GC) – ag

gggagaaagcuugaggAUGGAGGACGCCAAGAACAUCAAGAAGGGCCGGGCCUUCUA  
 CCCGCUGGAGGACGGGACCGCCGGCAGCAGCUCCACAAGGCCAUGAAGCGGUACGCCU  
 GGUGGCCGGCACGAUCGCCUUCACCGACGCCACAUCGAGGUCGACAUCACCUACGCCA  
 GUACUUCGAGAUGAGCGUGGCCUGGCCAGGCCAUGAAGCGGUACGCCUGAACACCAA  
 CCACCGGAUCGUGGUGUGCUCGGAGAACAGCCUGCGAGUUCUCAUGCCGGUGCUGGCC  
 CCUCUUCAUCGGCGUGGCCUGCCGGCAACGACAUCUACAACGAGCGGGAGCUGCU  
 GAACAGCAUGGGGAUCAGCCAGCGACCGUGGUGUUCGUGAGCAAGAACGGCCUGCAGAA  
 GAUCCUGAACGUGCAGAAGAACUGCCCAUCAUCCAGAACAUCAUCAUCAUGGACAGCAA  
 GACCGACUACCAGGGCUUCCAGUGAUGUACACGUUCGUGACCAGCCACCUCCGCC  
 CUUCAACGAGUACGACUUCGUUCGGAGAGCUUCGACCGGGACAAGACCAUCGCC  
 CAUGAACAGCAGCGCAGCACCGGCCUGCGAAGGGGUGGCCUGCCGACCGGACCG  
 CUGCGUGCGCUUCUCGCACGCCGGGACCCAUCUUCGGCAACCAGAACUACCCGG  
 CGCCAUCUGAGCGUGGUGCCGUUCCACCGACGGCUUCGGCAUGUUCAGCACCC  
 CCUCAUCUGGGCUUCCGGGUGGUCCUGAUGUACCGGUUCGAGGAGGACGUUC  
 GAGCCUGCAGGACUACAAGAACCCAGAGCGCUGCUCGUGCCGACCC  
 CGUUCAGCAGCACCCUGAUCGACAAGUACGACCUGUCGAACC  
 CUGCAGGAGAACGCGGACAC  
 GGGCGCCCGCUGAGCAAGGAGGUGGGCAGGCCAAGCGGUU  
 CAUCCGCCAGGGCUACGCCUGACCGAGACCACGAGCG  
 GGACGACAAGCCGGCGCCGUGGGCAAGGUGGUCCGUUC  
 CGUUCAGGCAAGGUGGUGGCCAGGCCAAGCG  
 CGCUGGCUGCACAGCGCGACAUCGCCUACUGGGACGAGG  
 GAGCAUCCUGCUCCAGCACCCAA  
 CGACGCCGGCGAGCUGCCGGCGGGGUGGGUG  
 GAAGGAGAUCGUCGACUACGUGGCCAGCAGG  
 CGUGGUGUUCGUGGAGCAGGUCCGAAGGGC  
 CCGCGAGAUCUGAUCAGGCCAAGAACGG  
 agacugacua**GCCCGAUGGGCCUCCCAACGGGCCUCCUCCCCUCCUUGCACCG**  
 Gagauua  
 auagauc-3'

## 도면7

## ppLuc(GC) – ag – A64

gggagaaagcuugaggAUGGAGGACGCCAAGAACAUCAAGAAGGGCCGGCCUUCUA  
 CCCGCUGGAGGACGGGACCGCCGGCAGCAGCUCCACAAGGCCAUGAAGCGGUACGCCU  
 GGUGGCCGGCACGAUCGCCUUCACCGACGCCACAUCGAGGUCGACAUCACCUACCGGA  
 GUACUUCGAGAUGAGCGUGGCCUGGCCAGGCCAUGAAGCGGUACGCCUGAACACCAA  
 CCACCGGAUCGUGGUGUGCUCGGAGAACAGCCUGCGAGUUCUCAUGCCGGUGCUGGGC  
 CCUCUUCAUCGGCGUGGCCGUCGCCGGCGAACGACAUCUACAACGAGCGGGAGCUGCU  
 GAACAGCAUGGGGAUCAGCCAGCGACCGUGGUGUUCGUGAGCAAGAACGGCCUGCAGAA  
 GAUCCUGAACGUGCAGAAGAACUGGCCAUCAUCCAGAACAUCAUCAUGGACAGCAA  
 GACCGACUACCAGGGCUUCCAGUCGAUGUACACGUUCGUGACCAGCCACCUCCGCCGG  
 CUUCAACGAGUACGACUUCGUUCGGAGAGCUUCGACCGGGACAAGACCAUCGCCUGAU  
 CAUGAACAGCAGCGGCAGCACCGGCCUGCCGAAGGGGUGGCCUGCCGACCGGACCG  
 CUGCGUGCGCUUCUCGACGCCGGGACCCAUCUUCGCAACCAGAACUACCCGGACAC  
 CGCCAUCUGAGCGUGGUGCCGUUCCACACGGCUUCGCAUGUUCAGCACCGACCG  
 CCUCAUCUGGGCUUCCGGGUGGUCCUGAUGUACCGGUUCGAGGAGGAGCUUCCUGCG  
 GAGCCUGCAGGACUACAAGAACUCCAGAGCGCUGCUCGUGCCGACCCUGUUCAGCUU  
 CGCCAAGAGCACCCUGAUCGACAAGUACGACCUGUCGAACCUGCAGGAGAACGCCAGCG  
 GGGCGCCCGCUGAGCAAGGAGGUGGGCAGGCCAAGCGGUUCCACCUCCGGAGGG  
 CAUCCGCCAGGGCUACGCCUGACCGAGACCACGAGCGGAUCCUGAACCCCCGAGGG  
 GGACGACAAGCCGGCGCCGUGGGCAAGGUGGUCCGUUUCUGAGGCCAAGGUGGUGGA  
 CCUGGACACCGCAAGACCCUGGGCUGAACCAGCGGGCGAGCUGUGCGUGCGGGGCC  
 GAUGAUCAUGAGCGCUACGUGAACAAACCCGGAGGCCACCAACGCCUCAUCGACAAGGA  
 CGGCUGGCUGCACAGCGCGACAUCGCCUACUGGGACGAGGACGAGCACUUCUCAUCGU  
 CGACCGGCUGAAGUCGCUAAGUACAAGGCCUACCGAGGUGGCCGGCGAGCUGGA  
 GAGCAUCCUGCUCCAGCACCCAAACAUUCGACGCCGGCUGGCCGGCUGCCGGACGA  
 CGACGCCGGCGAGCUGCCGGCGGGUGGGUGGUGUGCUGGAGCACGGCAAGACCAUGACGGA  
 GAAGGAGAACGUGACUACGUGGCCAGCCAGGUGACCCGCCAAGAACUGCGGGCG  
 CGUGGUGUUCGUGGAGCAGGUCCGAAGGCCUGACCGGGAGCUCGGACGCCGGAGAU  
 CCGCGAGAACCUCAAGGCCAAGAACGGCGGCCAAGAACUACGCCUGUA<sub>Agacuaguua</sub>  
<sub>agacugacua</sub>**GCCCGAUGGGCCUCCCAACGGGCCUCCUCCCCUCCUUGCACCGagauua**  
<sub>au</sub>aa  
 AAAAAA-3'

## 도면8

## ppLuc(GC) – ag – histoneSL

gggagaaagcuugaggAUGGAGGACGCCAAGAACAUCAAGAACGGCCGGCGCCUUCUA  
 CCCGCUGGAGGACGGGACCGCCGGCGAGCAGCUCCACAAGGCCAUGAACCGGUACGCCU  
 GGUGCCGGGCACGAUCGCCUUCACCGACGCCACAUCAUCGAGGUCGACAUCACCUACGCCA  
 GUACUUCGAGAUGAGCUGGCCUGGCCAGGCCAUGAACGGGUACGGCCUGAACACCAA  
 CCACCGGAUCGUGGUGUGCUCGGAGAACGCCUCCAGUUCUCAUGCCGGUGCUGGGCGC  
 CCUCUUCAGGGCUGGCCUGCCUGGCCGGCGAACGACAUCUACACGGAGCUGCU  
 GAACAGCAUGGGGAUCAGCCAGCGACCGUGGUUCGUGAGCAAGAACGGCCUGCAGAA  
 GAUCCUGAACGUGCAGAACAGCUGCCAUCAUCCAGAACAUCAUCAUGGACAGCAA  
 GACCGACUACCAGGGCUUCCAGUCGAUGUACACGUUCGUGACCAGCCACCUCCGCCGG  
 CUUCAAGAGUACGACUUCGUCCCGAGAGCUUCGACCGGGACAAGACCAUCGCCUGAU  
 CAUGAACAGCAGGGCAGCACGGCCUGCCGAAGGGGGUGGCCUGCCGACCGGACCG  
 CUGCGUGCGCUUCUCGCACGCCGGGACCCAUCUUCGGCAACCAGAACUACCCGGACAC  
 CGCCAUCUGAGCGUGGUUCGACCGGGCAUGUACAGACCCUGCCUGGGCAGCG  
 CCUCAUCUGCGCUUCCGGGUUGGUCCUGAUGUACCGGUUCGAGGAGGAGCUGUUCUGCG  
 GAGCCUGCAGGACUACAAGAACUCCAGAGCGCUGCUCGUGCCGACCCUGUUCAGCUU  
 CGCCAAGAGCACCUGAUCGACAAGUACGACCUGUGCCAGGAGAACGCG  
 GGGCGCCCGCUGAGCAAGGAGGUGGGCGAGGCCAAGCGGUUCCACCUCCCGAGGG  
 CAUCCGCCAGGGUACGCCUGACCGAGACCACGAGCGCAUCCUGAACACCCCGAGGG  
 GGACGACAAGCCGGCGCCGUGGGCAAGGUGGUCCGUUCUUCGAGGCCAAGGUGGUGGA  
 CCUGGACCCCGAAGACCCUGGGCUGAACCAGCAGCGGGCAGCUGUGCGUGCGGGGCC  
 GAUGAUCAUGAGCGGUACGUGAACACCCGGAGGCCAACGCCUACGACAAGGA  
 CGGCUGGCUGCACAGCGGCAGAACUCCGCUACUGGGACGAGGACGAGCACUU  
 CGACCGCUGAAGUCGCUAGUACAAAGGCCUACCGAGGUGGCCGGCGAGCUGGA  
 GAGCAUCCUGCUCCAGCACCCAAACAUUCGACGCCGGCUGGCCGGCUGCCGGACGA  
 CGACGCCGGCGAGCUGCCGGCGUGGUGGUCCGUGGAGCACGGCAAGACCAUGACGGA  
 GAAGGAGAUCGUCGACUACGUGGCCAGCCAGGUGACCACGCCAAGAACUGCGGGCG  
 CGUGGUGUUCGUGGACGAGGUCCCGAAGGCCUGACCGGGAAAGCUCGACGCCGGAAAGAU  
 CCGCGAGAUCGUCAAGGCCAAGAACGGCGCAAGAACGCCUGUAAgacuaguua  
 agacugacua**GCCC****GA****UGGGCCUCCCAACGGCCUCCUCCUCCUJGCACCGagauua**  
 auagaucu**CAAAGGCUCUUUCAGAGCCACCA****-3'**

## 도면9

## ppLuc(GC) – ag – A64 – histoneSL

gggagaaagcuugaggAUGGAGGACGCCAAGAACAUCAAGAACGGCCGGCGCCUUCUA  
 CCCGCUGGAGGACGGGACCGCCGGCGAGCAGCUCCACAAGGCCAUGAAGCGGUACGCCU  
 GGUGGCCGGCACGAUCGCCUUCACCGACGCCACAUCGAGGUCGACAUCAUCACCACCGA  
 GUACUUCGAGAUGAGCUGGCCUGGCCGAGGCCAUGAAGCGGUACGCCUGAACACCAA  
 CCACCGGAUCGUGGUGUGCUCCGGAGAACAGCCUGCAGUUCUCAUGCCGGUGCUGGCC  
 CCUCUUCAUCGGCGUGGCCUGCCGGCGAACGACAUCUACAAAGCAGCGGGAGCUGCU  
 GAACAGCAUGGGGAUCAGCCGACCGUGGUUCUGUGAGCAAGAAGGCCUGCAGAA  
 GAUCCUGAACGUGCAGAAGAACGCCAUCAUCCAGAAGAUCAUCAUCAUGGACAGCAA  
 GACCGACUACCAGGGCUUCCAGUCGAUGUACACGUUCUGACCAGCACCUCCGCC  
 CUUCAAGCAGUACGACUUCGUCCCGAGAGCUUCGACCGGGACAAGACCAUCGCCUGAU  
 CAUGAACAGCAGGGCAGCACCGCCUGCCGAAGGGGUGGCCUGCCGACCGGACCG  
 CUGCGUGCGCUUCUCGCACGCCGGGACCCAUCUUCGCAACCAGAUCAUCCCGACAC  
 CGCCAUCUGAGCGUGGUCCACCGCCUUCGCAUGUUCACGACCCUGGCC  
 CCUCAUCUGCGCUUCCGGGUCCUGAUGUACCGGUUCGAGGAGGAGCUGUUCUGCG  
 GAGCCUGCAGGACUACAAGAUCCAGAGCGCGCUGCUCGUGCCGACCGUUCAGCUUU  
 CGCCAAGAGCACCCUGAUCGACAAGUACGACCUGUCGAACCUGCACGAGAACGCC  
 GGGCGCCCGCUGAGCAAGGAGGGCGAGGCCUGGGCAAGCGGUUCCACCUCCCG  
 CAUCCGCCAGGGCUACGCCUGACCGAGAACACGAGCGCAUCCUGAUCAACCCCGAGGG  
 GGACGACAAGCCGGCGCCUGGGCAAGGUGGUCCGUUCUUCGAGGCCAAGGUGGUGGA  
 CCUGGACACCGGAAGACCCUGGGCGUGAACCAGCGGGCGAGCUGUGCGUGCGGGGCC  
 GAUGAUCAUGAGCGCUACGUGAACAAACCCGGAGGCCACCAACGCCUCAUCGACAAGGA  
 CGGCUGGCUGCACAGCGGCACAUCCGCUACUGGGACGAGGACGAGCACUUCUUCAUCGU  
 CGACCGCGUGAAGUCGCUGAUCAAGUACAAGGGCUACCAGGUGGCCGGCGAGCUGGA  
 GAGCAUCCUGCUCCAGCACCCAAACAUUUCGACGCCGGCUGGCCGGCUGCCGGACGA  
 CGACGCCGGCGAGCUGCCGGCGUGGGUGUGCUGGAGCACGGCAAGACCAUGACGGA  
 GAAGGAGAUCGUCGACUACGUGGCCAGCCAGGUGACCACCGCCAAGAACUGCGGGCG  
 CGUGGUGUUCGUGGAGCAGGGCUGACGGGGCAGCCGGGAAGCUCGACGCCGGAAAGAU  
 CCGCGAGAUCCUGAUCAAGGCCAAGAACGGGGCAAGAACGCGUGUA~~Agacuaguua~~  
 agacugacua**GCCCGAUGGGCCUCCCAACGGGCCUCCUCCCCUCCUUGCACCGagauua**  
 auAA  
 AAAAAAAugcauCAAAGGCUCUUUCAGAGCCACCA-3'

## 도면10

ppLuc(GC) – ag – A120

## 도면11

## ppLuc(GC) – ag – A64 – ag

gggagaaaagcugaggAUGGAGGACGCCAAGAACAUCAAGAAGGGCCGGGCCUUCUA  
 CCCGUGGAGGACGGGACCGCCGGCGAGCAGCUCCACAAGGCAUGAAGCGGUACGCCU  
 GGUGCCGGCACGAUCGCCUUCACCGACGCCACAUCGGAGGUCGACAUCACCUACGCC  
 GUACUUCGAGAUGAGCUGGCCUGGCCAGGCCAUGAAGCGGUACGCCUGAACACCAA  
 CCACCGGAUCGUGGUGUGUCGGAGAACAGCCUGCAGUUCUCAUGCCGGUGCUGGCC  
 CCUCUUCAUCGGCGUGGCCUGGCCCGCGAACGACAUCUACAACGAGCGGGAGCUGCU  
 GAACAGCAUGGGGAUCAGCAGCCGACCGUGGUUCGGAGCAAGAAGGCCUGCAGAA  
 GAUCCUGAACGUGCAGAAGAACGUGGCCAUCAUCCAGAAAGAUCAUCAUCAUGGACAGCAA  
 GACCGACUACCAGGGCUCCAGUCGAUGUACACGUUCGGACAGCACCUCGCCGG  
 CUUCAACGAGUACGACUUCGUCCGGAGAGCUUCGACCGGACAAGACCAUCGCCUGAU  
 CAUGAACAGCAGCGCAGCACCGGCCUGCGAAGGGGGUGGCCUGCCGACCGACCG  
 CUGCGUGCGCUCUCGACGCCGGGACCCAUCUUCGGAACAGAUCAUCCGGACAC  
 CGCCAUCUGAGCUGGCCGUUCCACCGACGGCUUCGGCAUGUUCACGACCCUGGGCUA  
 CCUCAUCUGCGCUUCCGGUGGUCCUGAUGUACCGGUUCGGAGGAGCUGUUCUGCG  
 GACCCUGCAGGACUACAAGAUCCAGACGGCGCUGCUCGGACCCUGUUCAGCUU  
 CGCCAAGAGCACCCUGAUCGACAAGUACGACCUUCGACGAGAACGACGCCAGCG  
 GGGCGCCCGCUGAGCAAGGAGGUGGGCGAGGCCAGCGGUUCCACCUCCGG  
 CAUCCGCCAGGGCUACGCCUGACCGAGACCAACGAGCGCAUCCUGAUCACCCCCGAGGG  
 GGACGACAAGCCGGCGCCUGGGCAAGGUGGUCCGUUUCUGAGGCCAAGGUGGUGGA  
 CCUGGACACCGCAAGACCCUGGGCGUGAACCCAGCGGGCGAGCUGUGCGUGCGGG  
 GAUGAUCAUGAGCGGCUACGUGAACACCCGGAGGCCACCAACGCCUCAUCGACAAGGA  
 CGGCUGGCUGCACAGCGGACAUCCUACUGGGACGGAGCACGGCACUUCUCAUCGU  
 CGACCGGCGUGAACGUGCGUGAACUAGAACAGGGCUACCCAGGGUGGCCGGAGCUGGA  
 GAGCAUCCUGCUCCAGCACCCAAACAUUCGACGCCGGCUGGCCGGCUGCCGGACGA  
 CGACGCCGGCGAGCUGCCGGCGGGUGGUGCGUGGAGCACGGCAAGACCAUGACGGA  
 GAAGGAGAACGUGACUACGUGGCCAGCCAGGUGACCACCGCCAAGAACGUGCGGGCG  
 CGUGGUGUUCGUGGAGCAGGGUCCCGAAGGGCUGACCGGGAAAGCUCGACGCCGGAAAGAU  
 CCGCGAGAACCCUGAUCAAGGCCAAGAACGGCGCAAGAACGCGUGUAAagacuaguua  
 agacugacuaGCCCGAUGGGCCUCCCAACGGGCCUCCUCCCCUCCUUGCACCGagauua  
 auAAA  
 AAAAAGugcauCCUGCCGAUGGGCCUCCCAACGGGCCUCCUCCCCUCCUUGCACCG3'

## 도면12

## ppLuc(GC) – ag – A64 – aCPSL

gggagaaaagcuuaggaggAUGGAGGACGCCAAGAACAUCAAGAAGGGCCGGCGCCUUCUA  
 CCCGCUGGAGGACGGGACCGCCGGCAGCAGCUCCACAAGCCAUGAAGCGGUACGCCU  
 GGUGCCGGGCACGAUCGCCUACCGACGCCACAUCGAGGUCGACAUCACCUACGCCA  
 GUACUUCGAGAUGAGCUGGCCUGGCCAGGCCAUGAAGCGGUACGCCUGAACACCAA  
 CCACCGGAUCGGGUGUGCUCGGAGAACAGCCUGCAGUUCUCAUGCCGGUGCUGGGC  
 CCUCUUCAUCGGCGUGGCCUGGCCGGGAACGACAUCUACAACGAGCAGGGACUGCU  
 GAACAGCAUGGGGAUCGCCAGCGACCGUGGUGUUCGUGAGCAAGAACGGCCUGCAGAA  
 GAUCUGAACGUGCAGAAGAACGUGCCCAUCAUCCAGAAGAACAUCAUCAUCAUGGACAGCAA  
 GACCGACUACCAGGGCUUCCAGUCGAUGUACACGUUCGUGACCAGCCACCUCCGCCGG  
 CUUCAACGAGUACGACUUCGUUCCGGAGAGCUUCGACCGGACAAGACCAUCGCCUGAU  
 CAUGAACAGCAGCGCAGCACCGGCCUGCCGAAGGGGGUGGCCUGCCGACCGGACCG  
 CUGCGUGCGCUUCUCGCACGCCGGGACCCAUCUUCGCAACCAGAUCAUCCGGACAC  
 CGCCAUCUGAGCGUGGUGCCGUUCCACCGCCUUCGCAUGUACAGCAGGGACUGUCCUG  
 CCUCAUCUGCGCUUCCGGGUUGGUCCUGAUGUACCGGUUCGAGGAGGAGCUGUUCUGCG  
 GAGCCUGCAGGACUACAAGAACUCCAGAGCCGCUGCUCGUGCCGACCCUGUUCAGCUU  
 CGCCAAGAGCACCCUGAUCGACAAGUACGACCUGUCGAACCUGCAGGAGAACGCCAGCG  
 GGGCGCCCGCUGAGCAAGGAGGUGGGCGAGGCCUGGCAAGCGGUUCCACCUCCGCC  
 CAUCCGCCAGGGCUACGCCUGACCGAGACCACGAGCGGAUCCUGAACACCCCGAGGG  
 GGACGACAAGCCGGCGCCUGGGCAAGGUGGUCCGUUCUUCGAGGCCAAGGUGGUGGA  
 CCUGGACACCGGAAGACCCUGGGCUGAACACCAGGAGGCCACCAACGCCUACUGACAAGGA  
 CGGCUGGCUGCACAGCGGCACAUCCUGGACGAGGACGAGCACUUUUCAUCGU  
 CGACCGGCUGAAGUCGCUGAUCAAGUACAAGGGUACCCAGGUGGCCGGCGAGCUGGA  
 GAGCAUCCUGCUCCAGCACCCAAACAUUCGACGCCGGCUGGCCGGCGAGCUGGA  
 CGACGCCGGCAGCUGCCGGCGUGGUGGUCCUGGAGCACGGCAAGACCAUGACCGA  
 GAAGGAGAACGUGACUACGUGGCCAGCAGGUGACCACGCCAAGAACUGCGGGCG  
 CGUGGUGUUCGUGGACGAGGUCCCGAAGGCCUGACGGGAAGCUCGACGCCGGAAAGAU  
 CCGCGAGAACUCCUGAUCAAGGCCAAGAACGGGGCAAGAACUGCCGUGUA~~Agacuaguua~~  
 agacugacua~~GCCCGAUGGCCUCCCAACGGCCCUCUCCCCUCCUUGCACCG~~~~agauua~~  
 auAAA  
 AAAAAAugcauCAAUUCCUACACGUGAGGCCUGUGAUUCCUAUCCCCUUCAUUCCU  
 AUACAUUAGCACAGGCCAUUGCAUGUAGGAAUU-3'

## 도면13

## ppLuc(GC) – ag – A64 – PolioCL

gggagaaagcuuaggAUGGAGGACGCCAAGAACAUCAAGAAGGGCCGGCCCCUUCUA  
 CCCGCUGGAGGACGGGACCGCCGGCAGCAGCUCCACAAGCCAUGAAGCGGUACGCCU  
 GGUGCCGGGCACGAUCGCCUUCACCGACGCCACAUCGAGGUCGACAUCACCUACGCCA  
 GUACUUCGAGAUGAGCUGGCCUGGCCAGGCCAUGAAGCGGUACGGCCUGAACACCAA  
 CCACCGGAUCGUUGGUGUGCUCCGGAGAACAGCCUGCAGUUCUCAUGCCGGUGCUGGCC  
 CCUCUUCAUCGGCGUGGCCUGGCCGGCGAACGACAUCUACAACGAGCAGGGAGCUGCU  
 GAACAGCAUGGGGAUCGCCAGCGACCGUGGUGUUCGUGAGCAAGAAGGCCUGCAGAA  
 GAUCCUGAACGUGCAGAAGAACGUGGCCAUCAUCCAGAAGAACAUCAUCAUGGACAGCAA  
 GACCGACUACCGGGCUUCCAGUCGAUGUACACGUUCGUGACCAGCCACCUCCGCCGG  
 CUUCAACGAGUACGACUUCGUCCGGAGAGCUUCGACCGGGACAAGACCAUCGCCUGAU  
 CAUGAACAGCAGCGCAGCACGCCUGCGAAGGGGGUGGCCUGCCGACCGGACCGC  
 CUGCGUGCGCUUCUCGCACGCCGGGACCCAUCUUCGCAACCAGAUCAUCCGGACAC  
 CGCCAUCUGAGCGUGGUGCCGUUCCACACGGCUUCCGAUGUUCACGACCCUGGGC  
 CCUCAUCUGCGGUUCCGGUGGUCCUGAUGUACCGGUUCGAGGAGGAGCUGUUCUGCG  
 GAGCCUGCAGGACUACAAGAACUCCAGAGCGCUGCUGCGACCCUGUUCAGCUU  
 CGCCAAGAGCACCCUGAUCGACAAGUACGACCUGUCGAACCUGCAGGAGAACGCCAGCG  
 GGGCGCCCGCUGAGCAAGGAGGUGGGCGAGGCCUGGGCAAGCGGUUCCACCCGG  
 CAUCCGCCAGGGCUACGCCUGACCAGGACACGAGCGCAUCCUGAACACCCGGAGGG  
 GGACGACAAGCCGGCGCCUGGGCAAGGUGGUCCGUUCUUCGAGGCCAAGGUGGUGGA  
 CCUGGACACCGCAAGACCCUGGGCGUGAACCAGCGGGCGAGCUGCGUGCGGGGCC  
 GAUGAUCAUGAGCGCUACGUGAACAAACCCGGAGGCCAACGCCUACGACAAGGA  
 CGGCUGGCUGCACAGCGCGACAUCGCCUACUGGGACGAGGACGAGCACUUCUCAUCGU  
 CGACCGCGCUGAAGUCGCUAGUAUCAAGGGCUACCAGGUGGCCGGCAGCUGGA  
 GAGCAUCCUGCUCCAGCACCCAAACAUUCGACGCCGGCUGGCCGGCUGCCGGACGA  
 CGACGCCGGCAGCUGCCGGCGUGGUGGUCCAGGCCAAGAACGUGCGGGCG  
 GAAGGAGAACGUGACUACGUGGCCAGCCAGGUGACCAAGGCCAAGAACGUGCGGGCG  
 CGUGGUGUUCGUGGAGCAGGUCCCGAAGGGCUGACGGGAAGCUGCGACGCCGGAAAGAU  
 CCGCGAGAUCCUGAUCAAGGCCAAGAACGGCGCAAGAACUGCCGUGUAA  
 agacugacua**GCCCGAUGGGCCUCCCAACGGCCCUCCUCCCCUCCUUGCACCG**agacua  
 auAA  
 AAAAAAAAugcauCAAUUCUAAAACAGCUCUGGGGUUGUACCCACCCAGAGGCCACGUGG  
 CGGUAGUACUCCGGUAUUGCGGUACCCUUGUACGCCUGUUUAGAAUU-3'

### 도면14

ppLuc(GC) – ag – A64 – G30

## 도면15

ppLuc(GC) – ag – A64 – U30

## 도면16

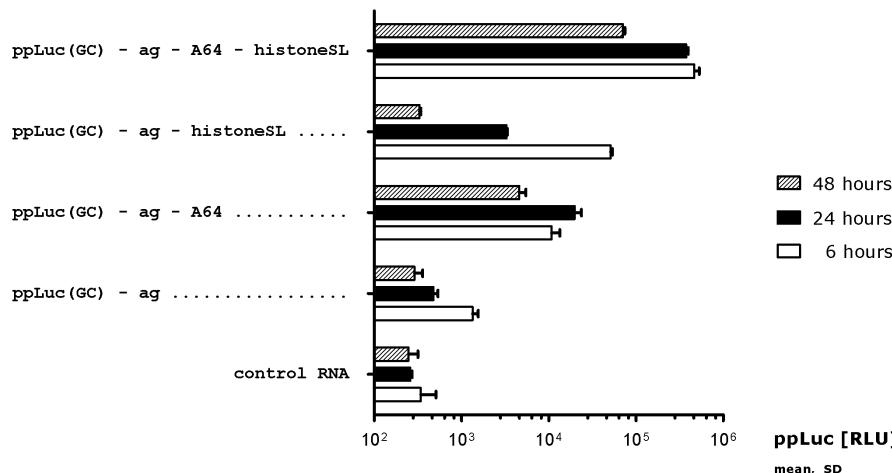
## ppLuc(GC) – ag – A64 – SL

gggagaaaagcuugaggAUGGAGGACGCCAAGAACAUCAAGAAGGGCCGGCCUUCUA  
 CCCGUGGAGGACGGGACCGCCGGCAGCAGCUCCACAAGCCAUGAAGCGGUACGCCU  
 GGUGGCCGGCACGAUCGCCUACCCGACGCCACAUCGAGGUCGACAUCACCUACCGGA  
 GUACUUCGAGAUGAGCGUGGCCUGGCCAGGCCAUGAAGCGGUACGCCUGAACACCAA  
 CCACCGGAUCGUGGUGUGCUCGGAGAACAGCCUGCAGUUCUCAUGCCGGUGCUGGGC  
 CCUCUUCAUCGGCGUGGCCUGGCCGGCAACGACAUCUACAACGAGCGGGAGCUGCU  
 GAACAGCAUGGGGAUCAGCCAGCGACCGUGGUGUUCGUGAGCAAGAACGCCUGCAGAA  
 GAUCCUGAACGUGCAGAAGAACCCAUCAUCCAGAAGAACAUCAUCAUGGACAGCAA  
 GACCGACUACCAGGGCUUCCAGUCGAUGUACACGUUCGUGACCAGCCACCUCCGCC  
 CUUCAACGAGUACGACUUCGUUCGGAGAGCUUCGACCGGGACAAGACCAUCGCC  
 CAUGAACAGCAGCGCAGCACCGGCCUGCGAAGGGGUGGCCUGCCGACCGGACCG  
 CUGCGUGCGCUUCUCGCACGCCGGGACCCAUCUUCGGCAACCAGAACUACCCGG  
 CGCCAUCUGAGCGUGGUGCCGUUCCACCGCCUUCGGCAUGUUCAGCACCCUGGG  
 CCUCAUCUGCGCUUCCGGGUGGUCCUGAUGUACCGGUUCGAGGAGGACGUUC  
 GAGCCUGCAGGACUACAAGAACUCCAGAGCGCGCUGCUCGUGCCGACCC  
 CGUUCAGCACCGACUAGAACUACGACCUGCAGCAGGAGAACUAGCCAGCG  
 GGGCGCCCGCUGAGCAAGGAGGUGGGCAGGCCUGGCCAACGGUUCC  
 CAUCCGCCAGGGCUACGCCUGACCGAGACCACGAGCGGAUCCUGAAC  
 GGACGACAAGCCGGCGCCUGGGCAAGGUGGUCCGUUUCUGAGGCC  
 GGAGGAGAUCGUGACUACGUGGCCAGCAGGUGACCC  
 CGACGCCGGCGAGCUGCCGGCGGGGUGGUGGUGCUGGAGCACGG  
 GAAGGAGAUCGUGACUACGUGGCCAGCAGGUGACCC  
 CGUGGUGUUCGUGGAGCAGGUCCGAAGGGCUGACGGGAAGC  
 CCGCGAGAUCUGAUCAAGGCCAAGAACGGCG  
 agacugacua**GCCCGAUGGGCCUCCCAACGGCCCUCCUCCCCUCCUUGCACCGagauua**  
 auaa  
 AAAAAGugcauUAUGGCGGCCGUGGUCCACCGAUAUCACCGUGGUGGACGCC-3'

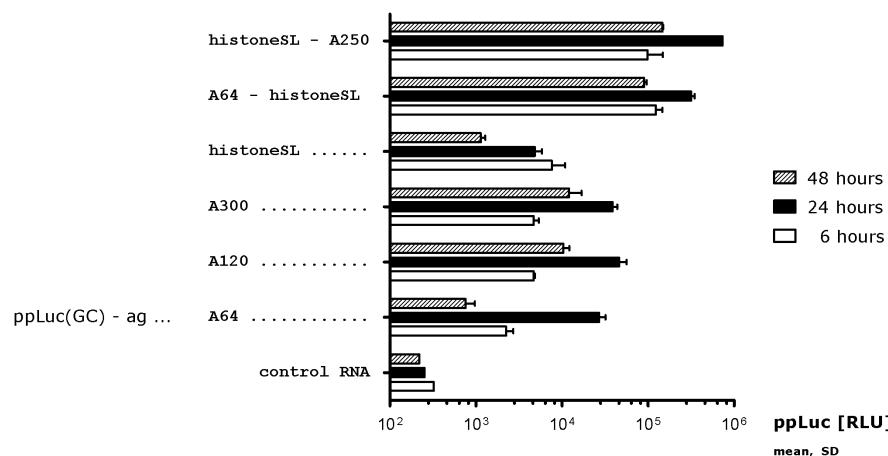
## 도면17

ppLuc(GC) – ag – A64 – N32

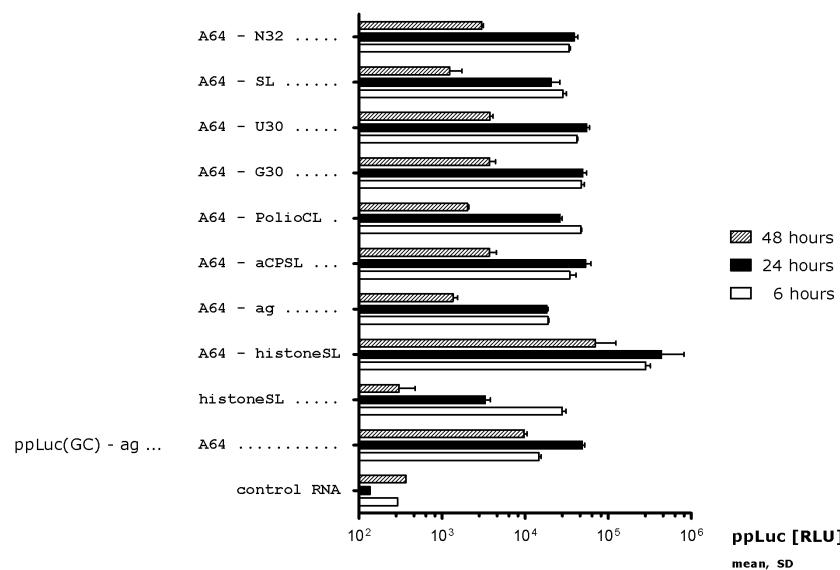
## 도면18



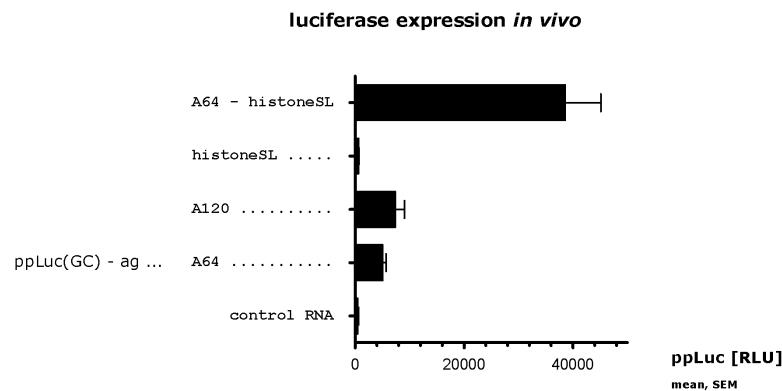
도면19



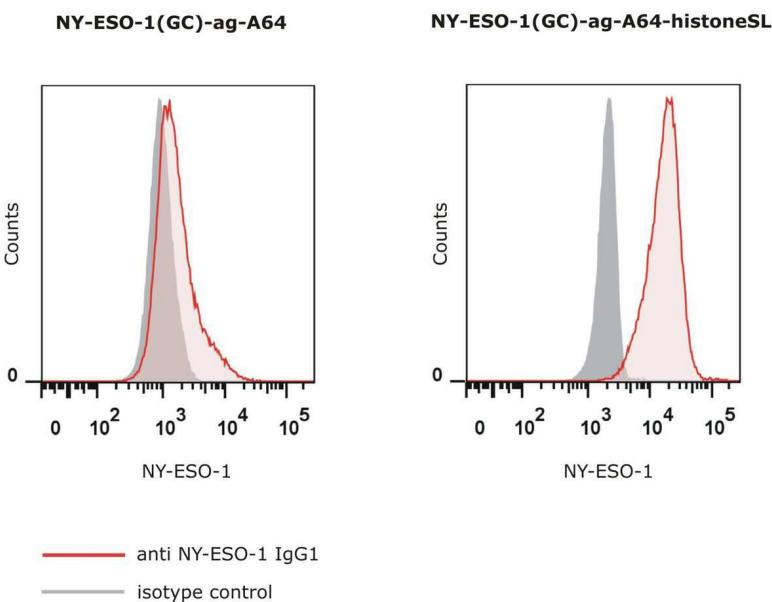
도면20



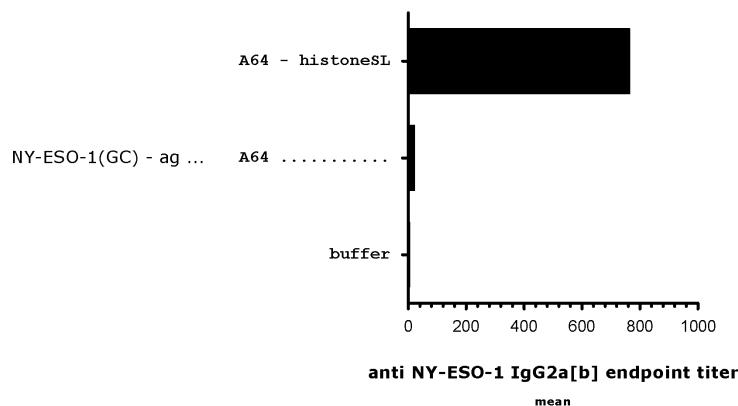
도면21



## 도면22



## 도면23



## 서 열 목록

## SEQUENCE LISTING

&lt;110&gt; CureVac GmbH

<120> Nucleic acid comprising a histone stem loop and optionally a polyadenylation sequence for increasing expression of an encoded protein

&lt;130&gt; CU01P088W01

&lt;160&gt; 95

&lt;170&gt; PatentIn version 3.5

&lt;210&gt; 1

<211> 16

<212> DNA

<213> Artificial Sequence

<220><223> stem loop structure according to formula (Ic)

<220><221> variation

<222> (1)..(1)

<223> /replace="guanine" /replace="adenine" /replace="thymine"

/replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (1)..(1)

<223> n is a, c, g, t or u

<220><221> variation

<222> (3)..(8)

<223> /replace="guanine" /replace="adenine" /replace="thymine"

/replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (3)..(8)

<223> n is a, c, g, t or u

<220><221> variation

<222> (10)..(14)

<223> /replace="guanine" /replace="adenine" /replace="thymine"

/replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (10)..(14)

<223> n is a, c, g, t or u

<220><221> variation

<222> (16)..(16)

<223> /replace="guanine" /replace="adenine" /replace="thymine"

/replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (16)..(16)

<223> n is a, c, g, t or u

<400> 1

ngnnnnnnun nnnncn

16

<210> 2

<211> 26

<212> DNA

<213> Artificial Sequence

<220><223> stem loop structure according to formula (IIC)

<220><221> variation

<222> (1)..(6)

<223> /replace="guanine" /replace="adenine" /replace="thymine"  
 /replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (1)..(6)

<223> n is a, c, g, t or u

<220><221> variation

<222> (8)..(13)

<223> /replace="guanine" /replace="adenine" /replace="thymine"  
 /replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (8)..(13)

<223> n is a, c, g, t or u

<220><221> variation

<222> (15)..(19)

<223> /replace="guanine" /replace="adenine" /replace="thymine"  
 /replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (15)..(19)

<223> n is a, c, g, t or u

<220><221> variation

<222> (21)..(26)

<223> /replace="guanine" /replace="adenine" /replace="thymine"  
 /replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (21)..(26)

<223> n is a, c, g, t or u

<400> 2

nnnnnnngnnn nnnunnnnnnc nnnnnn

26

<210> 3

<211> 16

<212> DNA

<213> Artificial Sequence

<220><223> stem loop structure according to formula (Id)

<220><221> variation

<222> (1)..(1)

<223> /replace="guanine" /replace="adenine" /replace="thymine"  
 /replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (1)..(1)

<223> n is a, c, g, t or u

<220><221> variation

<222> (3)..(8)

<223> /replace="guanine" /replace="adenine" /replace="thymine"  
 /replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (3)..(8)

<223> n is a, c, g, t or u

<220><221> variation

<222> (10)..(14)

<223> /replace="guanine" /replace="adenine" /replace="thymine"  
 /replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (10)..(14)

<223> n is a, c, g, t or u

<220><221> variation

<222> (16)..(16)

<223> /replace="guanine" /replace="adenine" /replace="thymine"  
 /replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (16)..(16)

<223> n is a, c, g, t or u

<400> 3

ncnnnnnnun nnnngn

16

<210> 4

<211> 26

<212> DNA

<213> Artificial Sequence

<220><223> stem loop structure according to formula (IId)

<220><221> variation

<222> (1)..(6)

<223> /replace="guanine" /replace="adenine" /replace="thymine"

/replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (1)..(6)

<223> n is a, c, g, t or u

<220>

><221> variation

<222> (8)..(13)

<223> /replace="guanine" /replace="adenine" /replace="thymine"

/replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (8)..(13)

<223> n is a, c, g, t or u

<220><221> variation

<222> (15)..(19)

<223> /replace="guanine" /replace="adenine" /replace="thymine"

/replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (15)..(19)

<223> n is a, c, g, t or u

<220><221> variation

<222> (21)..(26)

<223> /replace="guanine" /replace="adenine" /replace="thymine"  
 /replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (21)..(26)

<223> n is a, c, g, t or u

<400> 4

nnnnnncnnn nnnunnnnng nnnnnn

26

<210> 5

<211> 16

<212> DNA

<213> Artificial Sequence

<220><223> stem loop structure according to formula (Ie)

<220><221> variation

<222> (3)..(8)

<223> /replace="guanine" /replace="adenine" /replace="thymine"

/replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (3)..(8)

<223> n is a, c, g, t or u

<220><221> variation

<222> (10)..(14)

<223> /replace="guanine" /replace="adenine" /replace="thymine"

/replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (10)..(14)

<223> n is a, c, g, t or u

<400> 5

dgnnnnnnun nnnnch

16

<210> 6

<211> 26

<212> DNA

<213

> Artificial Sequence

<220><223> stem loop structure according to formula (IIe)

<220><221> variation

<222> (1)..(5)

<223> /replace="guanine" /replace="adenine" /replace="thymine"  
           /replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (1)..(5)

<223> n is a, c, g, t or u

<220><221> variation

<222> (8)..(13)

<223> /replace="guanine" /replace="adenine" /replace="thymine"  
           /replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (8)..(13)

<223> n is a, c, g, t or u

<220><221> variation

<222> (15)..(19)

<223> /replace="guanine" /replace="adenine" /replace="thymine"  
           /replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (15)..(19)

<223> n is a, c, g, t or u

<220><221> variation

<222> (22)..(26)

<223> /replace="guanine" /replace="adenine" /replace="thymine"  
           /replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (22)..(26)

<223> n is a, c, g, t or u

<400> 6

nnnnndgnnn nnnunnnnnnc hnnnnn

<210> 7

<211> 16

<212> DNA

<213> Artificial Sequence

<220><223> stem loop structure according to formula (If)

<220><221> variation

<222> (1)..(1)

<223> /replace="guanine" /replace="adenine" /replace="thymine"  
           /replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (1)..(1)

<223> n is a, c, g, t or u

<220><221> variation

<

222> (3)..(3)

<223> /replace="guanine" /replace="adenine" /replace="thymine"  
           /replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (3)..(3)

<223> n is a, c, g, t or u

<220><221> variation

<222> (7)..(8)

<223> /replace="guanine" /replace="adenine" /replace="thymine"  
           /replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (7)..(8)

<223> n is a, c, g, t or u

<220><221> variation

<222> (10)..(10)

<223>

> /replace="guanine" /replace="adenine" /replace="thymine"  
           /replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (10)..(10)

<223> n is a, c, g, t or u

<220><221> variation

<222> (12)..(12)

```

<223> /replace="guanine" /replace="adenine" /replace="thymine"
      /replace="uracile" /replace="cytosine" /replace=""
<220><221> misc_feature
<222> (12)..(12)
<223> n is a, c, g, t or u
<220><221> variation
<222> (14)..(14)
<223> /replace="guanine" /replace="adenine" /replace="thymine"
      /replace="uracile" /replace="cytosine" /replace=""
<220><221> misc_feature
<222> (14)..(14)
<223> n is a, c, g, t or u
<220><221> variation
<222> (16)..(16)
<223> /replace="guanine" /replace="adenine" /replace="thymine"
      /replace="uracile" /replace="cytosine" /replace=""
<220><221> misc_feature
<222> (16)..(16)
<223> n is a, c, g, t or u
<400> 7
ngnbyynnun rndncn 16
<210> 8
<211> 26
<212> DNA
<213
> Artificial Sequence
<220><223> stem loop structure according to formula (IIf)
<220><221> variation
<222> (1)..(6)
<223> /replace="guanine" /replace="adenine" /replace="thymine"
      /replace="uracile" /replace="cytosine" /replace=""
<220><221> misc_feature
<222> (1)..(6)
<223> n is a, c, g, t or u

```

<220><221> variation

<222> (8)..(8)

<223> /replace="guanine" /replace="adenine" /replace="thymine"  
 /replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (8)..(8)

<223> n is a, c, g, t or u

<220><221> variation

<222> (12)..(13)

<223> /replace="guanine" /replace="adenine" /replace="thymine"  
 /replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (12)..(13)

<223> n is a, c, g, t or u

<220><221> variation

<222> (15)..(15)

<223> /replace="guanine" /replace="adenine" /replace="thymine"  
 /replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (15)..(15)

<223> n is a, c, g, t or u

<220><221> variation

<222> (17)..(17)

<223> /replace="guanine" /replace="adenine" /replace="thymine"  
 /replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (17)..(17)

<223> n is a, c, g, t or u

<220><221> variation

<222> (19)..(19)

<223> /replace="guanine" /replace="adenine" /replace="thymine"  
 /replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (19)..(19)

<223> n is a, c, g, t or u

<220><221> variation

<222> (21)..(26)

<223> /replace="guanine" /replace="adenine" /replace="thymine"  
 /replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (21)..(26)

<223> n is a, c, g, t or u

<400> 8

nnnnnngnby ynnunrndnc nnnnnn

26

<210> 9

<211> 16

<212> DNA

<213> Artificial Sequence

<220><223> stem loop structure according to formula (Ig)

<220><221> variation

<222> (1)..(1)

<223> /replace="guanine" /replace="adenine" /replace="thymine"

/replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (1)..(1)

<223> n is a, c, g, t or u

<220><221> variation

<222> (8)..(8)

<223> /replace="guanine" /replace="adenine" /replace="thymine"

/replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (8)..(8)

<223> n is a, c, g, t or u

<220><221> variation

<222> (16)..(16)

<223> /replace="guanine" /replace="adenine" /replace="thymine"

/replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (16)..(16)

<223> n is a, c, g, t or u

<400> 9

nghyyydnth abrdcn

16

<210> 10

<211> 26

<212> DNA

<213> Artificial Sequence

<220><223> stem loop structure according to formula (IIg)

<220><221> variation

<222> (1)..(2)

<223> /replace="guanine" /replace="adenine" /replace="thymine"  
 /replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (1)..(2)

<223> n is a, c, g, t or u

<220><221> variation

<222> (4)..(6)

<223> /replace="guanine" /replace="adenine" /replace="thymine"  
 /replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (4)..(6)

<223> n is a, c, g, t or u

<220><221> variation

<222> (13)..(13)

<223> /replace="guanine" /replace="adenine" /replace="thymine"  
 /replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (13)..(13)

<223> n is a, c, g, t or u

<220><221> variation

<222> (21)..(25)

<223> /replace="guanine" /replace="adenine" /replace="thymine"  
 /replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (21)..(25)

<223> n is a, c, g, t or u

<400> 10

nhnnnnghyy ydnthabrdc nnnnnh

26

<210> 11

<211> 16

<212> DNA

<213> Artificial Sequence

<220><223> stem loop structure according to formula (Ih)

<400> 11

dghyctdyuh asrrcc

16

<210> 12

<211> 26

<212> DNA

<213> Artificial Sequence

<220><223> stem loop structure according to formula (IIh)

<220><221> variation

<222> (1)..(1)

<223> /replace="guanine" /replace="adenine" /replace="thymine"  
 /replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (1)..(1)

<223> n is a, c, g, t or u

<220><221> variation

<222> (25)..(25)

<223> /replace="guanine" /replace="adenine" /replace="thymine"  
 /replace="uracile" /replace="cytosine" /replace=""

<220><221> misc_feature	
<222> (25)..(25)	
<223> n is a, c, g, t or u	
<400> 12	
nhaahdghyc tdyuhasrrc cvhbnh	26
<210> 13	
<211> 16	
<212> DNA	
<213> Artificial Sequence	
<220><223> histon stem loop sequence according to formula (Ic)	
<220><221> variation	
<222> (4)..(4)	
<223> /replace="guanine" /replace="adenine" /replace="thymine"	
/replace="uracile" /replace="cytosine" /replace=""	
<220><221> misc_feature	
<222> (4)..(4)	
<223> n is a, c, g, t or u	
<220><221> variation	
<222> (13)..(13)	
<223> /replace="guanine" /replace="adenine" /replace="thymine"	
/replace="uracile" /replace="cytosine" /replace=""	
<220><221> misc_feature	
<222> (13)..(13)	
<223> n is a, c, g, t or u	
<400> 13	
ggcnctttc agngcc	16
<210> 14	
<211> 16	
<212> DNA	
<213> Artificial Sequence	
<220><223> histon stem loop sequence according to formula (Ie)	
<220><221> variation	
<222> (5)..(5)	

<223> n = g or c or t/u or a (any base or not present)

<220><221> misc\_feature

<222> (5)..(5)

<223> n is a, c, g, t or u

<220><221> variation

<222> (7)..(8)

<223> n = g or c or t/u or a (any base or not present)

<220><221> misc\_feature

<222> (7)..(8)

<223> n is a, c, g, t or u

<220><221> variation

<222> (10)..(10)

<223> n = g or c or t/u or a (any base or not present)

<220><221> misc\_feature

<222> (10)..(10)

<223> n is a, c, g, t or u

<220><221>

> variation

<222> (12)..(13)

<223> n = g or c or t/u or a (any base or not present)

<220><221> misc\_feature

<222> (12)..(13)

<223> n is a, c, g, t or u

<400> 14

ggctnnntnn anngcc

16

<210> 15

<211> 16

<212> DNA

<213> Artificial Sequence

<220><223> histon stem loop sequence according to formula (If)

<400> 15

ggcbctttc agdgcc

16

<210> 16

<211> 16

<212> DNA

<213> Artificial Sequence

<220><223> histon stem loop sequence according to formula (Ig)

<400> 16

ggctttth agagcc

16

<210> 17

<211> 16

<212> DNA

<213> Artificial Sequence

<220><223> histon stem loop sequence according to formula (Ih)

<400> 17

ggcyctttth agrgcc

16

<210> 18

<211> 26

<212> DNA

<213> Artificial Sequence

<220><223> histon stem loop sequence according to formula (IIc)

<220><221> variation

<222> (1)..(2)

<223> /replace="guanine" /replace="adenine" /replace="thymine"  
 /replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (1)..(2)

<223> n is a, c, g, t or u

<220><221> variation

<222> (4)..(5)

<223> /replace="guanine" /replace="adenine" /replace="thymine"  
 /replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (4)..(5)

<223> n is a, c, g, t or u

<220><221> variation

<222> (9)..(9)

<223> /replace="guanine" /replace="adenine" /replace="thymine"  
   /replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (9)..(9)

<223> n is a, c, g, t or u

<220><221> variation

<222> (18)..(18)

<223> /replace="guanine" /replace="adenine" /replace="thymine"  
   /replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (18)..(18)

<223> n is a, c, g, t or u

<220><221> variation

<222> (24)..(26)

<223> /replace="guanine" /replace="adenine" /replace="thymine"  
   /replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (24)..(26)

<223> n is a, c, g, t or u

<400> 18

nnanngcnc	tttcagngc	cacnnn	26
-----------	-----------	--------	----

<210> 19

<211> 26

<212> DNA

<213> Artificial Sequence

<220><223> histon stem loop sequence according to formula (IIe)

<220><221> variation

<222> (1)..(3)

<223> /replace="guanine" /replace="adenine" /replace="thymine"

  /replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (1)..(3)

<223> n is a, c, g, t or u

<220><221> variation

<222> (5)..(5)

<223> /replace="guanine" /replace="adenine" /replace="thymine"  
           /replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (5)..(5)

<223> n is a, c, g, t or u

<220><221> variation

<222> (10)..(10)

<223> /replace="guanine" /replace="adenine" /replace="thymine"  
           /replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (10)..(10)

<223> n is a, c, g, t or u

<220><221> variation

<222> (12)..(13)

<223> /replace="guanine" /replace="adenine" /replace="thymine"  
           /replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (12)..(13)

<223> n is a, c, g, t or u

<220><221> variation

<222> (15)..(15)

<223> /replace="guanine" /replace="adenine" /replace="thymine"  
           /replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (15)..(15)

<223> n is a, c, g, t or u

<220><221> variation

<222> (17)..(18)

<223> /replace="guanine" /replace="adenine" /replace="thymine"  
           /replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (17)..(18)

<223> n is a, c, g, t or u

<220><221> variation

<222> (24)..(26)

<223> /replace="guanine" /replace="adenine" /replace="thymine"  
 /replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (24)..(26)

<223> n is a, c, g, t or u

<400> 19

nnnanggctn tnnntnanngc cacnnn 26

<210> 20

<211> 26

<212> DNA

<213> Artificial Sequence

<220><223> histon stem loop sequence according to formula (IIf)

<220><221> variation

<222> (1)..(2)

<223> /replace="guanine" /replace="adenine" /replace="thymine"  
 /replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (1)..(2)

<223> n is a, c, g, t or u

<220><221> variation

<222> (4)..(5)

<223> /replace="guanine" /replace="adenine" /replace="thymine"  
 /replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (4)..(5)

<223> n is a, c, g, t or u

<220><221> variation

<222> (24)..(26)

<223> /replace="guanine" /replace="adenine" /replace="thymine"  
     /replace="uracile" /replace="cytosine" /replace=""  
 <220><221> misc\_feature  
 <222> (24)..(26)  
 <223> n is a, c, g, t or u  
 <400> 20  
 nnanngcbc tttcagdgc cacnnn

26

<210> 21  
 <211> 26  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> histon stem loop sequence according to formula (IIg)  
 <220><221> variation  
 <222> (1)..(1)  
 <223> /replace="guanine" /replace="adenine" /replace="thymine"  
     /replace="uracile" /replace="cytosine" /replace=""  
 <220><221> misc\_feature  
 <222> (1)..(1)  
 <223> n is a, c, g, t or u  
 <220><221> variation  
 <222> (5)..(5)  
 <223> /replace="guanine" /replace="adenine" /replace="thymine"  
     /replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature  
 <222> (5)..(5)  
 <223> n is a, c, g, t or u  
 <220><221> variation  
 <222> (25)..(25)  
 <223> /replace="guanine" /replace="adenine" /replace="thymine"  
     /replace="uracile" /replace="cytosine" /replace=""  
 <220><221> misc\_feature  
 <222> (25)..(25)  
 <223> n is a, c, g, t or u

<400> 21	
ncaanggctc tttthagagc caccnh	26
<210> 22	
<211> 26	
<212> DNA	
<213> Artificial Sequence	
<220><223> histon stem loop sequence according to formula (IIh)	
<220><221> variation	
<222> (1)..(1)	
<223> /replace="guanine" /replace="adenine" /replace="thymine"	
/replace="uracile" /replace="cytosine" /replace=""	
<220><221> misc_feature	
<222> (1)..(1)	
<223> n is a, c, g, t or u	
<220><221> variation	
<222> (25)..(25)	
<223> /replace="guanine" /replace="adenine" /replace="thymine"	
/replace="uracile" /replace="cytosine" /replace=""	
<220><221> misc_feature	
<222> (25)..(25)	
<223> n is a, c, g, t or u	
<400> 22	
nhaahggcyc tttthagrcg cvgbnh	26
<210> 23	
<211> 16	
<212> DNA	
<213> Artificial Sequence	
<220><223> histon stem loop sequence according to formula (Ic)	
<400> 23	
vgyyyyhhth rvvrcb	16
<210> 24	
<211> 16	
<212> DNA	

<213> Artificial Sequence

<220><223> histon stem loop sequence according to formula (Ic)

<400> 24

rgyyytttm agrrcs

16

<210> 25

<211> 16

<212> DNA

<213> Artificial Sequence

<220>

<223> histon stem loop sequence according to formula (Ic)

<400> 25

rgyyytttm agrrcs

16

<210> 26

<211> 16

<212> DNA

<213> Artificial Sequence

<220><223> histon stem loop sequence according to formula (Ic)

<400> 26

rgyyyyyytm rrrrcs

16

<210> 27

<211> 16

<212> DNA

<213> Artificial Sequence

<220><223> histon stem loop sequence according to formula (Ic)

<400> 27

ggcyctttc agrgcc

16

<210> 28

<211> 16

<212> DNA

<213> Artificial Sequence

<220><223> histon stem loop sequence according to formula (Ic)

<400> 28

ggccctttc agggcc

16

<210> 29

<211> 16

<212> DNA

<213> Artificial Sequence

<220><223> histon stem loop sequence according to formula (Ie)

<220><221> variation

<222> (3)..(5)

<223> /replace="guanine" /replace="adenine" /replace="thymine"

/replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (3)..(5)

<223> n is a, c, g, t or u

<220><221> variation

<222> (7)..(8)

<223> /replace="guanine" /replace="adenine" /replace="thymine"

/replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (7)..(8)

<223> n is a, c, g, t or u

<220><221> variation

<222> (12)..(14)

<223> /replace="guanine" /replace="adenine" /replace="thymine"

/replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (12)..(14)

<223> n is a, c, g, t or u

<400> 29

dgnnnbnnth vnnnch

16

<210> 30

<211> 16

<212> DNA

<213> Artificial Sequence

<220><223> histon stem loop sequence according to formula (Ie)

<220><221> variation

<222> (3)..(5)

<223> /replace="guanine" /replace="adenine" /replace="thymine"  
 /replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (3)..(5)

<223> n is a, c, g, t or u

<220><221> variation

<222> (13)..(14)

<223> /replace="guanine" /replace="adenine" /replace="thymine"  
 /replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (13)..(14)

<223> n is a, c, g, t or u

<400> 30

rgnnnyhbth rdnncy

16

<210> 31

<211> 16

<212> DNA

<213> Artificial Sequence

<220><223> histon stem loop sequence according to formula (Ie)

<220><221> variation

<222> (3)..(3)

<223> /replace="guanine" /replace="adenine" /replace="thymine"  
 /replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (3)..(3)

<223> n is a, c, g, t or u

<220><221> variation

<222> (14)..(14)

<223> /replace="guanine" /replace="adenine" /replace="thymine"  
 /replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (14)..(14)

<223> n is a, c, g, t or u

<400> 31

rgndbyhyth rdhncy

16

<210> 32

<211> 16

<212> DNA

<213> Artificial Sequence

<220><223> histon stem loop sequence according to formula (Ie)

<400> 32

rgykywytw rrmrcy

16

<210> 33

<211> 16

<212> DNA

<213> Artificial Sequence

<220><223> histon stem loop sequence according to formula (Ie)

<400> 33

ggctytwytw armgcc

16

<210> 34

<211> 16

<212> DNA

<213> Artificial Sequence

<220>

<223> histon stem loop sequence according to formula (Ie)

<400> 34

ggctttta agagcc

16

<210> 35

<211> 16

<212> DNA

<213> Artificial Sequence

<220><223> histon stem loop sequence according to formula (If)

<400> 35

vgyyytyhth ryrrcb

16

<210> 36

<211> 16

<212> DNA

<213> Artificial Sequence

<220><223> histon stem loop sequence according to formula (If)

<400> 36

sgyycttytm agrrcs

16

<210> 37

<211> 16

<212> DNA

<213> Artificial Sequence

<220><223> histon stem loop sequence according to formula (If)

<400> 37

sgyyctttm agrrcs

16

<210> 38

<211> 16

<212> DNA

<213> Artificial Sequence

<220><223> histon stem loop sequence according to formula (If)

<400> 38

sgyyyyytm rrrrcs

16

<210> 39

<211> 16

<212> DNA

<213> Artificial Sequence

<220><223> histon stem loop sequence according to formula (If)

<400> 39

ggcyctttc agrgcc

16

<210> 40

<211> 16

<212> DNA

<213> Artificial Sequence

<220><223> histon stem loop sequence according to formula (If)

<400> 40

ggccctttc agggcc

16

<210> 41  
 <211> 16  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> histon stem loop sequence according to formula (Ig)  
 <400> 41  
 ggyycttyth agrrcc 16  
 <210> 42  
 <211> 16  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> histon stem loop sequence according to formula (Ig)  
 <400> 42  
 ggcycyttm agrgcc 16  
 <210> 43  
 <211> 16  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> histon stem loop sequence according to formula (Ig)  
 <400> 43  
 ggcttttm agrgcc 16  
 <210> 44  
 <211> 16  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> histon stem loop sequence according to formula (Ig)  
 <400> 44  
 rgyyykytm asrrcb 16  
 <210> 45  
 <211> 16  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> histon stem loop sequence according to formula (Ig)

<400> 45	
ggctctttm agagcc	16
<210> 46	
<211> 16	
<212> DNA	
<213> Artificial Sequence	
<220><223> histon stem loop sequence according to formula (Ig)	
<400> 46	
ggctctttc agagcc	16
<210> 47	
<211> 16	
<212> DNA	
<213> Artificial Sequence	
<220><223> histon stem loop sequence according to formula (Ih)	
<400> 47	
kgcyctryth agrrcc	16
<210> 48	
<211> 16	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> histon stem loop sequence according to formula (Ih)	
<400> 48	
ggcycttth agrgcc	16
<210> 49	
<211> 16	
<212> DNA	
<213> Artificial Sequence	
<220><223> histon stem loop sequence according to formula (Ih)	
<400> 49	
ggcycttth agrgcc	16
<210> 50	
<211> 16	
<212> DNA	

<213> Artificial Sequence

<220><223> histon stem loop sequence according to formula (Ih)

<400> 50

kghyctkytm asrrcc

16

<210> 51

<211> 16

<212> DNA

<213> Artificial Sequence

<220><223> histon stem loop sequence according to formula (Ih)

<400> 51

ggcyctttm agrgcc

16

<210> 52

<211> 16

<212> DNA

<213> Artificial Sequence

<220><223> histon stem loop sequence according to formula (Ih)

<400> 52

ggctctttc agagcc

16

<210> 53

<211> 26

<212> DNA

<213> Artificial Sequence

<220><223> histon stem loop sequence according to formula (IIc)

<220><221> variation

<222> (26)..(26)

<223> /replace="guanine" /replace="adenine" /replace="thymine"

/replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (26)..(26)

<223> n is a, c, g, t or u

<400> 53

hhhhhvgyyy yhhthrvvrc bvhhhn

26

<210> 54

&lt;211&gt; 26

&lt;212&gt; DNA

&lt;213&gt;

Artificial Sequence

&lt;220&gt;&lt;223&gt; histon stem loop sequence according to formula (IIc)

&lt;400&gt; 54

hhmmmrgrgyyy ttttmagrrc sachhh 26

&lt;210&gt; 55

&lt;211&gt; 26

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;&lt;223&gt; histon stem loop sequence according to formula (IIc)

&lt;400&gt; 55

hmmhmrgrgyyy ttttmagrrc sachhh 26

&lt;210&gt; 56

&lt;211&gt; 26

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;&lt;223&gt; histon stem loop sequence according to formula (IIc)

&lt;400&gt; 56

mmmmmmrgrgyyy yytymrrrc smymmmw 26

&lt;210&gt; 57

&lt;211&gt; 26

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;&lt;223&gt; histon stem loop sequence according to formula (IIc)

&lt;400&gt; 57

mmammmggcyc tttcagrgc cacmmw 26

&lt;210&gt; 58

&lt;211&gt; 26

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;&lt;223&gt; histon stem loop sequence according to formula (IIc)

&lt;400&gt; 58

acaacggccc tttcaggc cacaa

26

&lt;210&gt; 59

&lt;211&gt; 26

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;&lt;223&gt; histon stem loop sequence according to formula (IIe)

&lt;220&gt;&lt;221&gt; variation

&lt;222&gt; (1)..(5)

<223> /replace="guanine" /replace="adenine" /replace="thymine"  
/replace="uracile" /replace="cytosine" /replace=""

&lt;220&gt;&lt;221&gt; misc\_feature

&lt;222&gt; (1)..(5)

&lt;223&gt; n is a, c, g, t or u

&lt;220&gt;&lt;221&gt; variation

&lt;222&gt; (8)..(10)

<223> /replace="guanine" /replace="adenine" /replace="thymine"  
/replace="uracile" /replace="cytosine" /replace=""

&lt;220&gt;&lt;221&gt; misc\_feature

&lt;222&gt; (8)..(10)

&lt;223&gt; n is a, c, g, t or u

&lt;220&gt;&lt;221&gt; variation

&lt;222&gt; (12)..(13)

<223> /replace="guanine" /replace="adenine" /replace="thymine"  
/replace="uracile" /replace="cytosine" /replace=""

&lt;220&gt;&lt;221&gt; misc\_feature

&lt;222&gt; (12)..(13)

&lt;223&gt; n is a, c, g, t or u

&lt;220&gt;&lt;221&gt; variation

&lt;222&gt; (17)..(19)

<223> /replace="guanine" /replace="adenine" /replace="thymine"  
/replace="uracile" /replace="cytosine" /replace=""

&lt;220&gt;&lt;221&gt; misc\_feature

<222> (17)..(19)

<223> n is a, c, g, t or u

<220><221> variation

<222> (22)..(22)

<223> /replace="guanine" /replace="adenine" /replace="thymine"

/replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (22)..(22)

<223> n is a, c, g, t or u

<220><221> variation

<222> (24)..(26)

<223> /replace="guanine" /replace="adenine" /replace="thymine"

/replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (24)..(26)

<223> n is a, c, g, t or u

<400> 59

nnnnndgnnn bnnthvnnnc hhnnn

26

<210> 60

<211> 26

<212> DNA

<213> Artificial Sequence

<220><223> histon stem loop sequence according to formula (IIe)

<220><221> variation

<222> (1)..(2)

<223> /replace="guanine" /replace="adenine" /replace="thymine"

/replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (1)..(2)

<223> n is a, c, g, t or u

<220><221> variation

<222> (8)..(10)

<223> /replace="guanine" /replace="adenine" /replace="thymine"

```

/replace="uracile" /replace="cytosine" /replace=""

<220><221> misc_feature
<222> (8)..(10)
<223> n is a, c, g, t or u

<220><221> variation
<222> (18)..(19)
<223> /replace="guanine" /replace="adenine" /replace="thymine"
      /replace="uracile" /replace="cytosine" /replace=""

<220><221> misc_feature
<222> (18)..(19)
<223> n is a, c, g, t or u

<400> 60

```

nnhhhrgnnn yhbthrdnnc ydhhhh 26

```

<210> 61
<211> 26
<212> DNA
<213> Artificial Sequence
<220><223> histon stem loop sequence according to formula (IIe)
<220><221> variation
<222> (1)..(1)
<223> /replace="guanine" /replace="adenine" /replace="thymine"
      /replace="uracile" /replace="cytosine" /replace=""

<220><221> misc_feature
<222> (1)..(1)
<223> n is a, c, g, t or u

<220><221> variation
<222> (8)..(8)
<223> /replace="guanine" /replace="adenine" /replace="thymine"
      /replace="uracile" /replace="cytosine" /replace=""

<220><221> misc_feature
<222> (8)..(8)
<223> n is a, c, g, t or u

<220><221> variation

```

<222> (19)..(19)	
<223> /replace="guanine" /replace="adenine" /replace="thymine"	
/replace="uracile" /replace="cytosine" /replace=""	
<220><221> misc_feature	
<222> (19)..(19)	
<223> n is a, c, g, t or u	
<400> 61	
nhhhhrgnb yhythrdhnc yrhhh	26
<210> 62	
<211> 26	
<212> DNA	
<213> Artificial Sequence	
<220><223> histon stem loop sequence according to formula (IIe)	
<400> 62	
mhwmrgrgyky ywytwrrmrc yrymmmm	26
<210> 63	
<211> 26	
<212> DNA	
<213> Artificial Sequence	
<220><223> histon stem loop sequence according to formula (IIe)	
<400> 63	
mhwamggcty twytwrmgc cacmmm	26
<210> 64	
<211> 26	
<212> DNA	
<213> Artificial Sequence	
<220><223> histon stem loop sequence according to formula (IIe)	
<400> 64	
acaaaggctt tttaagagc caccaa	26
<210> 65	
<211> 26	
<212> DNA	
<213> Artificial Sequence	

<220><223> histon stem loop sequence according to formula (IIIf)

<220><221> variation

<222> (26)..(26)

<223> /replace="guanine" /replace="adenine" /replace="thymine"  
 /replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature

<222> (26)..(26)

<223> n is a, c, g, t or u

<400> 65

hhhhmvgyyy tyhthryrc bvmhhn

26

<210> 66

<211> 26

<212> DNA

<213>

Artificial Sequence

<220><223> histon stem loop sequence according to formula (IIIf)

<400> 66

mmmmmsgyyyc ttytmagrcc smchhh

26

<210> 67

<211> 26

<212> DNA

<213> Artificial Sequence

<220><223> histon stem loop sequence according to formula (IIIf)

<400> 67

mmmmmsgyyyc tttmagrrc sachmh

26

<210> 68

<211> 26

<212> DNA

<213> Artificial Sequence

<220><223> histon stem loop sequence according to formula (IIIf)

<400> 68

mmmmmsgyyy yytmtrrrc smmmmw

26

<210> 69

<211> 26

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;&lt;223&gt; histon stem loop sequence according to formula (IIf)

&lt;400&gt; 69

mmmmmmggcyc tttcagrgc cacmmw

26

&lt;210&gt; 70

&lt;211&gt; 26

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;&lt;223&gt; histon stem loop sequence according to formula (IIf)

&lt;400&gt; 70

acaacggccc tttcagggc cacaa

26

&lt;210&gt; 71

&lt;211&gt; 26

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;&lt;223&gt; histon stem loop sequence according to formula (IIg)

&lt;220&gt;&lt;221&gt; variation

&lt;222&gt; (24)..(25)

<223> /replace="guanine" /replace="adenine" /replace="thymine"  
/replace="uracile" /replace="cytosine" /replace=""

&lt;220&gt;&lt;221&gt; misc\_feature

&lt;222&gt; (24)..(25)

&lt;223&gt; n is a, c, g, t or u

&lt;400&gt; 71

hhvamggcyc ttythagrrc cvhnnm

26

&lt;210&gt; 72

&lt;211&gt; 26

&lt;212&gt; DNA

&lt;213&gt;

Artificial Sequence

&lt;220&gt;&lt;223&gt; histon stem loop sequence according to formula (IIg)

&lt;400&gt; 72

hhaamggcyc ttytmagrgc cvchhm

26

<210> 73  
 <211> 26  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> histon stem loop sequence according to formula (IIg)  
 <400> 73  
 mhaamggctc tttmagrc cmcymmm 26  
 <210> 74  
 <211> 26  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> histon stem loop sequence according to formula (IIg)  
 <400> 74  
 mmmmmrgyyy ykytmasrrc bmmymmm 26  
 <210> 75  
 <211> 26  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> histon stem loop sequence according to formula (IIg)  
 <400> 75  
 mcaamggctc tttmagagc caccmmm 26  
 <210> 76  
 <211> 26  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> histon stem loop sequence according to formula (IIg)  
 <400> 76  
 acaaaggctc tttcagagc caccca 26  
 <210> 77  
 <211> 26  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> histon stem loop sequence according to formula (IIh)

<220><221> variation  
 <222> (1)..(1)  
 <223> /replace="guanine" /replace="adenine" /replace="thymine"  
       /replace="uracile" /replace="cytosine" /replace=""  
 <220><221> misc\_feature  
 <222> (1)..(1)  
 <223> n is a, c, g, t or u  
 <220><221> variation  
 <222> (25)..(25)  
 <223> /replace="guanine" /replace="adenine" /replace="thymine"  
       /replace="uracile" /replace="cytosine" /replace=""

<220><221> misc\_feature  
 <222> (25)..(25)  
 <223> n is a, c, g, t or u  
 <400> 77  
 nhaahkgcyc trythagrcc cvhbnh 26  
 <210> 78  
 <211> 26  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> histon stem loop sequence according to formula (IIh)  
 <220><221> variation  
 <222> (25)..(25)  
 <223> /replace="guanine" /replace="adenine" /replace="thymine"  
       /replace="uracile" /replace="cytosine" /replace=""  
 <220><221> misc\_feature  
 <222> (25)..(25)

<223> n is a, c, g, t or u  
 <400> 78  
 hhaamggcyc ttthagrgc cvmynm 26  
 <210> 79  
 <211> 26  
 <212> DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;&lt;223&gt; histon stem loop sequence according to formula (IIh)

&lt;400&gt; 79

hmaaaggcyc tttthagrc crmyhm 26

&lt;210&gt; 80

&lt;211&gt; 26

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;&lt;223&gt; histon stem loop sequence according to formula (IIh)

&lt;400&gt; 80

mmaamkghyc tkytmasrrc crmyym 26

&lt;210&gt; 81

&lt;211&gt; 26

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;&lt;223&gt; histon stem loop sequence according to formula (IIh)

&lt;400&gt; 81

mmaamggcyc tttmagrc crmyym 26

&lt;210&gt; 82

&lt;211&gt; 26

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;&lt;223&gt; histon stem loop sequence according to formula (IIh)

&lt;400&gt; 82

ccaaaggctc tttcagagc caccca 26

&lt;210&gt; 83

&lt;211&gt; 1747

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;

&lt;220&gt;&lt;223&gt; ppLuc(GC) - ag

&lt;400&gt; 83

gggagaaagg uugaggaagg aggacgccaa gaacaucaag aaggccccgg cgcccuucua 60

ccgcuggag gacgggaccg cggcgagca gcuccacaag gccaugaagc gguacgccc 120

ggugccgggc acgaucgccu ucacccgacgc ccacaucgag gucgacauca ccuacgcgga	180
guacuucgag augagcgugc gccuggccga ggcuaugaaag cgguaacggcc ugaacaccaa	240
ccaccggaua guggugugcu cggagaacag ccugcaguuc uucaugccgg ugcugggcgc	300
ccucuucauc ggcguggccg ucgccccggc gaacgacaua uacaacgagc gggagcugcu	360
gaacagcaug gggauacagcc agccgaccgu gguguucgug agcaagaagg gccugcagaa	420
gauccugaac gugcagaaga agcugccau cauccagaag aucaucauca uggacagcaa	480
gaccgacuac cagggcuucc agucgaugua cacguucgug accagccacc ucccgccgg	540
cuucaacgag uacgacuucg ucccgagag cuucgaccgg gacaagacca ucgcccugau	600
caugaacagc agcggcagca cggccugcc gaagggggug gcccugccgc accggaccgc	660
cugcugcgc uucucgcacg cccgggaccc caucuucggc aaccagauca uccggacac	720
cgccauccug agcguggugc cguuccacca cggcuuucggc auguucacga cccugggcua	780
ccucaucugc ggcuuccggg ugguccugau guaccgguuc gaggaggagc uguuccugcg	840
gagccugcag gacuacaaga uccagagcgc gcugcucgug ccgaccugu ucagcuuu	900
cgcuaagagc acccugaucg acaaguacga ccugucgaac cugcagcaga ucgccaggg	960
gggcgccccg cugagcaagg aggugggcga gggcguggcc aagcgguucc accuccggg	1020
cauccgcccag ggcuacggcc ugaccgagac cacgagcgcg auccugauca ccccccgggg	1080
ggacgacaag cggggcgccg uggaagggu gguccgguuc uucgaggcca aggugggugga	1140
ccuggacacc ggcagaagaccc ugcccugaa ccagcgggac gagcugugcg ugcggggcc	1200
gaugaucaug agcgccuacg ugaacaaccc ggaggccacc aacgcccua ucgacaagga	1260
cggcuggcug cacagcgccg acaucgccua cugggacgag gacgagcacu ucuucaucgu	1320
cgaccggcug aagucgcuga ucaaguacaa gggcuaccag guggcgccgg ccgagcugga	1380
gagcauccug cuccagcacc ccaacauuu cgacgcccgc guggccggc ugccggacga	1440
cgacgcccggc gagcugccgg cgcgguggu ggugcuggag cacggcaaga ccaugacgga	1500
gaaggagauc gucgacuacg ugcccagcca ggugaccacc gccaagaagc ugcggggccg	1560
cgugguguuc guggacgagg ucccgaaggg ccugaccggg aagcucgacg cccggaagau	1620
ccgcgagauc cugaucaagg ccaagaaggg cggcaagauc gccguguaag acuaguua	1680
agacugacua gcccgauggg ccucccaacg ggcccuccuc cccuccuugc accgagauua	1740
auagauac	1747
<210> 84	
<211> 1806	
<212> DNA	

&lt;213&gt; Artificial Sequence

&lt;220&gt;&lt;223&gt; ppLuc(GC) - ag - A64

&lt;400&gt; 84

gggagaaagg uugaggaugg aggacgcca gaacaucaag aaggcccgg cgccuucua	60
cccgcuggag gacgggaccg cggcgagca gcuccacaag gcagaagg gguacgccc	120
ggugccggc acgaucgcc uacccgacgc ccacaucgag gugacauca ccuacgcca	180
guacuucgag augagcgugc gccuggccga ggcagaagg cgguacggcc ugaacaccaa	240
ccaccggau guggugugcu cggagaacag ccugcaguuc uucaugccgg ugcugggcgc	300
ccucuucauc ggcguggccg ugcggccgc gaacgacau uacaacgagc gggagcugcu	360
gaacagcaug gggaucaagcc agccgaccgu gguguucgug agcaagaagg gccugcagaa	420
gauccugaac gugcagaaga agcugccau cauccagaag aucaucauca uggacagcaa	480
gaccgacuac cagggcuucc agucgaugua cacguucgug accagccacc uccggccgg	540
cuucaacgag uacgacuucg uccggagag cuucgaccgg gacaagacca ugcggugau	600
caugaacagc agcggcagca cggccugcc gaagggggug gcccugccgc accggaccgc	660
cugcugcgc uucucgcacg cccgggaccc caucuucggc aaccagauca uccggacac	720
cgc当地ug agcguggugc cguuccacca cggcuucggc auguucacga cccuggcua	780
ccucaucugc gcuuuccggg ugguccugau guaccgguuc gaggaggagc uguuccugcg	840
gagccugcag gacuacaaga uccagagcgc gcugcucgug ccgaccugu ucagcuiuu	900
cgc当地agc acccugauag acaaguacga ccugucgaac cugcagcaga ugcggcaggg	960
ggggcccccg cugagcaagg agggggcga ggccguggcc aagcgguiuc accuccggg	1020
cauccgcccag gcuacggcc ugaccgagac cacgagcgcg auccugauca ccccccggg	1080
ggacgacaag cggggcgccg uggaagggu gguccguuc uucgaggcca agguggugga	1140
ccuggacacc gcaagaccc uggaaggugaa ccagcggggc gagcugugcg ugcggggcc	1200
gaugaucaug agcggcuacg ugaacaaccc ggaggccacc aacgcccua ucgacaagga	1260
cggcuggcug cacagcggcgc acaucgccua cugggacgag gacgacacu ucuucaucgu	1320
cgaccggcug aagucgcuga ucaaguacaa gggcuaccag guggccggc cggagcugga	1380
gagcauccug cuccagcacc ccaacaucuu cgacgcccgc guggccggc ugccggacga	1440
cgacgcccgc gagcugccgg cgcgguggu ggugcuggag cacggcaaga ccaugacgga	1500
gaaggagauc gucgaucacg ugcccagcca ggugaccacc gccaagaagc ugcggggcgg	1560
cgguguguuc guggacgagg ucccgaaagg ccugaccggg aagcucgacg cccggaagau	1620
ccgcgagauc cugaucaagg ccaagaagg cggcaagauc gccguguaag acuaguua	1680

agacugacua	gcccgaauggg	ccucccaacg	ggcccuccuc	ccuccuugc	accgagauua	1740
auaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	1800
aaaaaaaa						1806
<210>	85					
<211>	1772					
<212>	DNA					
<213>	Artificial Sequence					

<220><223> ppLuc(GC) - ag - histoneSL

<400>	85					
gggagaaagc	uugaggaugg	aggacgcca	gaacaucaag	aaggcccgg	cgccuucua	60
ccgcuggag	gacgggaccg	ccggcgagca	gcuccacaag	gccaugaagc	gguacgccc	120
ggugccggc	acgaucgccu	ucaccgacgc	ccacaucgag	gucgacauca	ccuacgcgga	180
guacuucgag	augagcgugc	gccuggccga	ggccaugaag	cgguacggcc	ugaacaccaa	240
ccaccggau	guggugugcu	cggagaacag	ccugcaguuc	uucaugccgg	ugcugggcgc	300
ccucuucau	ggcguggccg	ucgccccggc	gaacgacauc	uacaacgagc	gggagcugcu	360

gaacagcaug	gggaucagcc	agccgaccgu	gguguucgug	agcaagaagg	gccugcagaa	420
gauccugaac	gugcagaaga	agcugccau	cauccagaag	aucaucauca	uggacagcaa	480
gaccgacuac	cagggciucc	agucgaugua	cacguucgug	accagccacc	ucccgccggg	540
ciucaacgag	uacgacuucg	ucccgagag	cuucgaccgg	gacaagacca	ucgcccugau	600
caugaacagc	agcggcagca	ccggccugcc	gaagggggug	gcccugccgc	accggaccgc	660
cugcugcgc	uucucgcacg	cccgggaccc	caucuucggc	aaccagauca	ucccggacac	720
cgccauccug	agcguggugc	cguuccacca	cggciucggc	auguucacga	cccugggcua	780

ccucaucugc	ggciuccggg	ugguccugau	guaccgguuc	gaggaggagc	uguuccugcg	840
gagccugcag	gacuacaaga	uccagagcgc	gcugcucgug	ccgaccugu	ucagciucuu	900
cgc当地	accugauacg	acaaguacga	ccugucgaac	cugcagcaga	ucgcccagcgg	960
gggc当地	cugagcaagg	aggugggcga	ggccguggcc	aagcgguucc	accuccggg	1020
cauccggccag	ggcuacggcc	ugaccgagac	cacgagcgcg	aucugauca	cccccgaggg	1080
ggacgacaag	ccggccgccc	ugggcaagg	ggcccguuc	uucgaggcca	agguggugga	1140
ccuggacacc	ggcaagaccc	ugggcugua	ccagcggggc	gagcugugc	ugcggggcc	1200

gaugaucaug	agcgccuacg	ugaacaaccc	ggaggccacc	aacgcccua	ucgacaagga	1260
cggcuggcug	cacagcggcgc	acaucgccua	cugggacgag	gacgagcacu	uciucaucgu	1320

cgaccggcug aagucgcuga ucaaguacaa gggcuaccag guggcgccgg ccgagcugga	1380
gagcauccug cuccagcacc ccaacaucuu cgacgcggc guggccggc ugccggacga	1440
cgacgcggc gagcugccgg ccgcgguggu ggugcuggag cacggcaaga ccaugacgga	1500
gaaggagauc gucgcacuacg uggccagcca ggugaccacc gccaagaagc ugcggggcgg	1560
cgugguguuc guggacgagg ucccgaaggg ccugaccggg aagcucgacg cccggaagau	1620
ccgcgagauc cugaucaagg ccaagaaggg cggcaagauc gccguguaag acuaguuaaua	1680
agacugacua gcccgauuggg ccucccaacg ggcccuccuc cccuccuugc accgagauua	1740
auagaucuca aaggcucuuu ucagagccac ca	1772
<210> 86	
<211> 1835	
<212> DNA	
<213> Artificial Sequence	
<220><223> ppLuc(GC) - ag - A64 - histoneSL	
<400> 86	
gggagaaagc uugaggaugg aggacgccaa gaacaucaag aaggcccgg cggccuucua	60
cccgcuggag gacgggaccg ccggcgagca gcuccacaag gccaugaagc gguacgccc	120
ggugccggc acgaucgccc ucacccgacgc ccacaucgag gucgcacuca ccuacgccc	180
guacuucgag augagcgugc gccuggccga ggcuaugaaag cgguaacggcc ugaacaccaa	240
ccaccggaua guggugugcu cggagaacag ccugcaguuc uucaugccgg ugcuggggcgc	300
ccucuucau ggcguggccg ucgcggccgc gaacgacauc uacaacgagc gggagcugcu	360
gaacagcaug gggauacgaccc agccgaccgu gguguucgug agcaagaagg gccugcagaa	420
gauccugaac gugcagaaga agcugcccau cauccagaag aucaucauca uggacagcaa	480
gaccgacuac cagggcuiucc agucgaugua cacguucgug accagccacc ucccggccgg	540
cuucaacgag uacgacuucg ucccggagag cuucgaccgg gacaagacca ucgcccugau	600
caugaacagc agcggcagca cggccugcc gaagggggug gcccugccgc accggaccgc	660
cugcgugcgc uucucgcacg cccgggaccc caucuucggc aaccagauca ucccggacac	720
cgccauccug agcguggugc cguuccacca cggcuucggc auguucacga cccuggggcua	780
ccucaucugc ggcuuccggg ugguccugau guaccgguuc gaggaggagc uguuccugcg	840
gagccugcag gacuacaaga uccagagcgc gcugcucgug ccgaccuugc ucagcuuuu	900
cggcaagagc acccugacuacg acaaguacga ccugucgaac cugcagcaga ucgccagcgg	960
gggcgccccg cugagcaagg aggugggcga ggccguggcc aagcgguucc accucccggg	1020

cauccggccag ggcuacggcc ugaccgagac cacgagcgcg auccugauca ccccccgggg	1080
ggacgacaag ccggggcgccg ugggcaaggu ggccccguuc uucgaggcca agguggguga	1140
ccuggacacc ggcagaagaccc ugggcgugaa ccagcggggc gagcugugcg ugcggggcc	1200
gaugaucaug agcggcuacg ugaacaaccc ggaggccacc aacgcccua ucgacaagga	1260
cggcuggcug cacagcggcg acaucgccua cugggacgag gacgagcacu ucuucaucgu	1320
cgaccggcug aagucgcuga ucaaguacaa gggcuaccag guggccggg ccgagcugga	1380
gagcauccug cuccagcacc ccaacaucuu cgacgcccggc guggccgggc ugccggacga	1440

cgacgcccggc gagcugccgg ccgcgguggu ggugcuggag cacggcaaga ccaugacgga	1500
gaaggagauc gucaciucg uggccagcca ggugaccacc gccaagaagc ugcggggcg	1560
cgugguguuc guggacgagg ucccgaaggg ccugaccggg aagcucgacg cccggaagau	1620
ccgcgagauc cugaucaagg ccaagaaggg cggcaagauc gccguguaag acuaguua	1680
agacugacua gcccgauggg ccucccaacg ggcccuccuc cccuccuugc accgagauua	1740
auaaaaaaaaaaaaaaa aaaaaaaaaaaaaaaa aaaaaaaaaaaaaaaa aaaaaaaaaaaaaaaa	1800
aaaaaaaaaugca ucaaaggcuc uuuucagagc cacca	1835

&lt;210&gt; 87

&lt;211&gt; 1869

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;&lt;223&gt; ppLuc(GC) - ag - A120

&lt;400&gt; 87

gggagaaaagc uugaggaugg aggacgccaa gaacaucaag aaggcccgg cgccuucua	60
cccgcuggag gacgggaccg ccggcgagca gcuccacaag gccaugaagc gguacgccc	120
ggugccgggc acgaucgccc ucacccgacgc ccacaucgag gucgacauca ccuacgccc	180
guacuucgag augagcgugc gccuggccga ggcuaugaaag cgguacggcc ugaacaccaa	240
ccacccggauc guggugugcu cggagaacag ccugcaguuc uucaugccgg ugcuggggc	300
ccucuucauc ggccuggccg ucgcggccgc gaacgacaua uacaacgagc gggagcugcu	360

gaacagcaug gggauacagcc agccgaccgu gguguucgug agcaagaagg gcccugcagaa	420
gauccugaac gugcagaaga agcugcccau cauccagaag aucaucauca uggacagcaa	480
gaccgacuac cagggcuucc agucgaugua cacguucgug accagccacc ucccgccgg	540
cuucaacgag uacgacuucg ucccgagag cuucgaccgg gacaagacca ucgcggugau	600
caugaaacgc agcggcagca ccggccugcc gaagggggug gcccugccgc accggaccgc	660
cugcugcgc uucucgcacg cccgggaccc caucuucggc aaccagauca ucccgacac	720

cgccauccug agcguggugc cguuccacca cggcuucggc auguucacga cccugggcua	780
ccucaucugc ggcuuccggg ugguccugau guaccgguuc gaggaggagc uguuccugcg	840
gagccugcag gacuacaaga uccagagcgc gcugcucug cgcaccugu ucagcuucuu	900
cgcgaagagc acccugaucg acaaguacga ccugugcaac cugcagcaga ugcgcagcgg	960
gggcgcggcug cugagcaagg aggugggcga ggccguggcc aagcgguuc accucccgaa	1020
cauccggcag ggcuacggcc ugaccgagac cacgagcgcg auccugauca ccccgaggg	1080
ggacgacaag ccgggcgcgg ugugcaaggu ggucggcguuc uucgaggcca aggugggug	1140
ccuggacacc ggcagaagaccc ugugcugugaa ccagcgggc gagcugugcg ugcggggcc	1200
gaugaucaug agcggcuacg ugaacaaccc ggaggccacc aacgcccua ucgacaagga	1260
cgccuggcug cacagcgccg acaucgcccua cugggacgag gacgagcacu ucuucaucgu	1320
cgaccggcug aagucgcuga ucaaguacaa gggcuaccag guggcgcgg ccgagcugga	1380
gagcauccug cuccagcacc ccaacaucuu cgacgcggc guggccggc ugccggacga	1440
cgacgcggc gacgugccgg ccgcgguggu ggugcuggag cacggcaaga ccaugacgga	1500
gaaggagauc gucgacuacg ugugccagca ggugaccacc gccaagaagc ugcggggcg	1560
cgugguguuc guggacgagg ucccgaaagg ccugaccggg aagcucgacg cccggaagau	1620
ccgcgagauc cugaucaagg ccaagaaggg cggcaagauc gcccuguaag acuaguuaaua	1680
agacugacua gcccgauggg ccucccaacg ggcccuccuc cccuccuugc accgagauua	1740
auagaucuaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa	1800
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa	1860
aaaaaaaaaa	1869
<210> 88	
<211> 1858	
<212> DNA	
<213> Artificial Sequence	
<220><223> ppLuc(GC) - ag - A64 - ag	
<400> 88	
gggagaaagc uugaggaugg aggacgccaa gaacaucaag aaggcccgg cgcccuucua	60
ccgcuggag gacgggacgg cggcgagca gcuccacaag gccaugaagc gguacgccc	120
ggugccggc acgaucgccc ucacccgacgc ccacaucgag gucgacauca ccuacgcgg	180
guacuucgag augagcgugc gccuggccga ggcuaugaaag cgguaacggcc ugaacaccaa	240
ccaccggauc guggugugcu cggagaacag ccugcaguuc uucaugccgg ugcugggcgc	300

cccuucauc ggcguggccg ucggccggc gaacgacauc uacaacgagc gggagcugcu	360
gaacagcaug gggauacgac agccgaccgu gguguucgug agcaagaagg gccugcagaa	420
gauccugaac gugcagaaga agcugccau cauccagaag aucaucaua uggacagcaa	480
gaccgacuac cagggcuucc agucgaugua cacguucgug accagccacc ucccgccggg	540
cuucaacgag uacgacuucg ucccgagag cuucgaccgg gacaagacca ucggccugau	600
caugaacagc agcggcagca ccggccugcc gaagggggug gcccugccgc accggaccgc	660
cugcgugcgc uucucgcacg cccgggaccc caucuucggc aaccagaua ucccgacac	720
cgc当地ugugc acguggugc cguuccacca cggcuucggc auguucacga cccuggcua	780
ccucaucugc ggc当地ccggg ugguccugau guaccgguuc gaggaggagc uguuccugcg	840
gagccugcag gacuacaaga uccagagcgc gcugcucgug ccgaccuug ucaucuu	900
cgc当地agac acccugauacg acaaguacga ccugucgaac cugcagcaga ucggccagcgg	960
gggc当地cccg cugagcaagg aggugggcga ggccguggcc aagcgguucc accuccccggg	1020
cauccgcccag ggc当地acggcc ugaccgagac cacgagcgcg auccugaua ccccccgggg	1080
ggacgacaag cc当地ggcgc当地 ugga当地agggu ggucgguuc uucgaggcca aggugggugga	1140
ccuggacacc ggcaagaccc ugggc当地ugaa cc当地ggggc gagcugugc ugccc当地ggcc	1200
gaugaucaug agcggcuacg ugaacaaccc ggaggccacc aacgccc当地ca ucgacaagga	1260
cggcuggcug cacagcggc当地 acaucggccua cugggacgag gacgagcacu ucuucaucgu	1320
cgaccggcug aagucgcuga ucaaguacaa gggcuaccag guggccggg cc当地gugcugga	1380
gagcauccug cuccagcacc ccaacaucuu cgacgccc当地 guggccggg ugccc当地gacga	1440
cgacgccc当地 gagcugccgg cc当地ggguggu ggugcuggag cacggcaaga ccaugacgga	1500
gaaggagaauc gugc当地acg ugga当地agggu ggugaccacc gccaagaagc ugccc当地gggg	1560
cgugguguuc guggacgagg uccc当地aggg cc当地ggggcc aagcucgacg cccggaaagau	1620
ccg当地gagaauc cugaucaagg ccaagaagg ggcaagac gccguguaag acuaguua	1680
agacugacua gccc当地aggg cc当地ggggcc aacgagccuugc accgagauua	1740
auaaaaaaaaaaaa aaaaaaaaaaaaa aaaaaaaaaaaaa aaaaaaaaaaaaa aaaaaaaaaaaaa	1800
aaaaaaaaugca uccugcccgaa ugccc当地ucc aacggggccu cc当地ggggccu uugcaccg	1858
<210> 89	
<211> 1894	
<212> DNA	
<213> Artificial Sequence	
<220><223> ppLuc(GC) - ag - A64 - aCPSL	

&lt;400&gt; 89

gggagaaagc uugaggaugg aggacgcca gaacaucaag aagggccgg cgccciucua	60
cccgcuggag gacgggaccg ccggcgagca gcuccacaag gcuaugaagc gguacgccc	120
ggugccggc acgaucgcu ucaccgacgc ccacaucgag gucgacauca ccuacgcca	180
guacuucgag augagcgugc gccuggccga ggccaugaag cgguacggcc ugaacaccaa	240
ccaccggauc guggugugcu cggagaacag ccugcaguuc uucaugccgg ugcugggcgc	300
ccucuucauc ggcguggccg ucgccccggc gaacgacauc uacaacgagc gggagcugcu	360
gaacagcaug gggaucagcc agccgaccgu gguguucgug agcaagaagg gccugcagaa	420
gauccugaac gugcagaaga agcugccau cauccagaag aucaucauca uggacagcaa	480
gaccgacuac cagggciucc agucgaugua cacguucgug accagccacc uccgccccgg	540
cuucaacgag uacgacuucg uccggagag cuucgaccgg gacaagacca ugcuccugau	600
caugaacagc agcggcagca ccggccugcc gaagggggug gcccugccgc accggaccgc	660
cugcgugcgc uucucgcacg cccgggaccc caucuucggc aaccagauca uccggacac	720
cgc当地ugcugc agcguggugc cguuccacca cggciucggc auguucacga cccugggcua	780
ccucaucugc ggc当地ccggg ugguccugau guaccgguuc gaggaggagc uguuccugcg	840
gagccugcag gacuacaaga uccagagcgc gcugcucgug ccgaccugu ucagciucuu	900
cgc当地agc acccugaucg acaaguacga ccugucgaac cugcagcaga ugc当地cgg	960
gggcccccg cugagcaagg aggugggcga ggccguggcc aagcgguiucc accuccccgg	1020
cauccgccag ggc当地acggcc ugaccgagac cacgagcgcg auccugauca ccccccggg	1080
ggacgacaag ccggccgccc uggaaggu gguccgguuc uucgaggcca agguggugga	1140
ccuggacacc ggc当地agaccc uggaugaa ccagcggggc gagcugugcg ugc当地ggcc	1200
gaugaucaug agcggcuacg ugaacaaccc ggaggccacc aacgcccua uc当地acaagga	1260
cggcuggcug cacagcggcgc acaucgcca cugggacgag gacgagcacu uciucuacgu	1320
cgaccggcug aagucgcuga ucaaguacaa gggciaccag guggccggc cc当地agcugga	1380
gagcauccug cuccagcacc ccaacaucuu cgacgcccgc guggccggc ugccggacga	1440
cgacgcccgc gagcugccgg cc当地gguggu ggugcuggag cacggcaaga ccaugacgga	1500
gaaggagauc guc当地acuacg ugcccagcca ggugaccacc gccaagaagc ugc当地ggccgg	1560
cguggugiuuc guggacgagg uccgaaagg ccugaccggg aagcucgacg cccggaagau	1620
ccglocalgagauc cugaucaagg ccaagaaggg cggcaagauc gccguguaag acuaguuaaua	1680
agacugacua gcccgaauggg ccuuccaacg ggccuccuc cccuucuugc accgagauua	1740
auaaaaaaaaaaaaaaa aaaaaaaaaaaaaaaa aaaaaaaaaaaaaaaa aaaaaaaaaaaaaaaa	1800

aaaaaaugca ucaaauuccua cacgugaggc gcugugauuc ccuaucccc uucaauuccu	1860
auacauuagc acagcgccau ugcauguagg aauu	1894
<210> 90	
<211> 1909	
<212> DNA	
<213> Artificial Sequence	
<220><223> ppLuc(GC) - ag - A64 - PolioCL	
<400> 90	
gggagaaagc uugaggaugg aggacgccaa gaacaucaag aaggcccgg cgccciucua	60
cccgcuggag gacgggaccg cggcgagca gcuccacaag gccaugaagc gguacgccc	120
ggugccgggc acgaucgccc uacccgacgc ccacaucgag gucgacauca ccuacgccc	180
guacuucgag augagcugc gccuggccga ggcuaugaag cgguacggcc ugaacaccaa	240
ccaccggauc guggugugcu cggagaacag ccugcaguuc uucaugccgg ugcugggcgc	300
ccucuucauc ggcguggccg ucgcggccgc gaacgacauc uacaacgagc gggagcugcu	360
gaacagcaug gggaucaagcc agccgaccgu gguguucgug agcaagaagg gccugcagaa	420
gaucucugaac gugcagaaga agcugccau cauccagaag aucaucauca uggacagcaa	480
gaccgacuac cagggciucc agucgaugua cacguucgug accagccacc ucccgccgg	540
cuucaacgag uacgacuucg ucgggagag cuucgaccgg gacaagacca ucgcugcugau	600
caugaacagc agcggcagca ccggccugcc gaagggggug gcccugccgc accggaccgc	660
cugcugcgc uucucgcacg cccgggaccc caucuucggc aaccagauca ucgggacac	720
cgccauccug agcguggugc cguuccacca cggcuucggc auguucacga cccugggcua	780
ccucaucugc gcuuuccggg ugguccugau guaccgguuc gaggaggagc uguuccugcg	840
gagccugcag gacuacaaga uccagagcgc gcugcucgug cggaccgu ucagcuiuu	900
cgcggcggc acccugaucg acaaguacga ccugcugcaac cugcagcaga ucggccagcg	960
ggggcccccg cugagcaagg aggugggcga ggccguggcc aagcgguiuc accuccgg	1020
cauccggcag ggcuacggcc ugaccgagac cacgagcgcg auccugauca ccccgaggg	1080
ggacgacaag ccggccggc ugggcaagg gucccguuc uucgaggcca aggugggugga	1140
ccuggacacc ggcagaagaccc ugcccugaa ccagcggggc gagcugugcg ugcggggcc	1200
gaugaucaug agcggcuacg ugaacaaccc ggaggccacc aacgcccua ucgacaagga	1260
cgccuggcug cacagcggcgc acaucgccua cugggacgag gacgagcacu ucuucaucgu	1320
cgaccggcug aagucgcuga ucaaguacaa ggcuaccag guggcggcgg ccgagcugga	1380

gagcauccug cuccagcacc ccaacaucuu cgacgcccggc guggccgggc ugccggacga 1440  
 cgacgcccggc gagcugccgg ccgcgguggu ggugcuggag cacggcaaga ccaugacgga 1500  
 gaaggagauc gucgacuacg uggccagcca ggugaccacc gccaagaagc ugcggggcgg 1560  
 cgugguguuc guggacgagg ucccgaaggg ccugaccggg aagcucgacg cccggaagau 1620

ccgcgagauc cugaucaagg ccaagaaggg cggcaagauc gccguguaag acuaguuaaua 1680  
 agacugacua gcccgauggg ccucccaacg ggcuccuc cccuccuugc accgagauua 1740  
 auaaaaaaaaaaaaaaa aaaaaaaaaaaaaaaa aaaaaaaaaaaaaaaa aaaaaaaaaaaaaaaa 1800  
 aaaaaaaaaugca ucaauucuaa aacagcucug ggguuguacc caccggagag gcccacgugg 1860  
 cgguaguac uccgguaauug cgguacccuu guacgcuu uuuagaauu 1909

&lt;210&gt; 91

&lt;211&gt; 1841

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;&lt;223&gt; ppLuc(GC) - ag - A64 - G30

&lt;400&gt; 91

gggagaaagc uugaggaugg aggacgccaa gaacaucaag aaggggccgg cgccuucua 60

cccgcuggag gacgggaccg cggcgagca gcuccacaag gccaugaagc gguacgccc 120  
 ggugccgggc acgaucgccc ucacccgacgc ccacaucgag gucgacauca ccuacgcgg 180  
 guacuucgag augagcgcugc gccuggccga ggcuaugaaag cgguacggcc ugaacaccaa 240  
 ccacccggauc guggugugcu cggagaacag ccugcaguuc uucaugccgg ugcugggcgc 300  
 ccucuucauc ggcguggccg ucgcggccgc gaacgacaua uacaacgagc gggagcugcu 360  
 gaacagcaug gggauacagcc agccgaccgu ggugiuucgug agcaagaagg gccugcagaa 420  
 gauccugaac gugcagaaga agcugcccau cauccagaag aucaucaua uggacagcaa 480

gaccgacuac cagggcuucc agucgaugua cacguucgug accagccacc uccgcggg 540  
 cuucaacgag uacgacuucg uccggagag cuucgaccgg gacaagacca ucgcggcugau 600  
 caugaacagc agcggcagca cggccugcc gaagggggug gcccugccgc accggaccgc 660  
 cugcugcgc uucucgcacg cccgggaccc caucuucggc aaccagauca uccggacac 720  
 cgccauccug agcguggugc cguuccacca cggcuucggc auguucacga cccugggcua 780  
 ccucaucugc ggciuccggg ugguccugau guaccgguiuc gaggaggagc uguuccugcg 840  
 gagccugcag gacuacaaga uccagagcgc gcugcucgug cccgaccgu ucagciucuu 900

cgccaagagc acccugaucg acaaguacga ccugucgaac cugcagcaga ucggccagcgg 960

gggcgcggc cugagcaagg aggugggcga ggccguggcc aagcgguucc accuccggg 1020  
 cauccggcag ggcuacggcc ugaccgagac cacgagcgcg auccugaua ccccgagg 1080  
 ggacgacaag ccggcgccg uggcaagggu ggcccguuc uucgaggcca agguggg 1140  
 ccuggacacc ggcagaaccc uggcuguaa ccagcgggc gagcugugcg ugcggggcc 1200  
 gaugaucaug agcggcuacg ugaacaaccc ggaggccacc aacgcccua ucgacaagga 1260  
 cggcuggcug cacagcgccg acaucgccua cugggacgag gagcagcacu ucuucaugu 1320

cgaccggcug aagucgcuga ucaaguacaa gggcuaccag guggcgcgg ccgagcugga 1380  
 gagcauccug cuccagcacc ccaacaucuu cgacgcccgc guggccggc ugccggacga 1440  
 cgacgcccgc gagcugccgg ccgcgguggu ggugcuggag cacggcaaga ccaugacgga 1500  
 gaaggagauc gucgacuacg uggccagcca ggugaccacc gccaagaagc ugcggggcgg 1560  
 cgugguguuc guggacgagg ucccgaaggg ccugaccggg aagcucgacg cccggaagau 1620  
 ccgcgagauc cugaucaagg ccaagaaggg cggcaagauc gccguguaag acuaguua 1680  
 agacugacua gcccgauaggg ccucccaacg ggcccuccuc cccuccuugc accgagauua 1740

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1800  
 aaaaaaaaaugca ugaaaaaaaaa gggggggggg gggggggggg gggggggggg g 1841

&lt;210&gt; 92

&lt;211&gt; 1841

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;&lt;223&gt; ppLuc(GC) - ag - A64 - U30

&lt;400&gt; 92

gggagaaagc uugaggaugg aggacgccaa gaacaucaag aaggcccgg cgcccuucua 60  
 cccgcuggag gacgggaccg ccggcgagca gcuccacaag gccaugaagc gguacgccc 120  
 ggugccggc acgaucgccu ucacgcacgc ccacaucgag gucgacauca ccuacgcgga 180  
 guacuucgag augagcgugc gccuggccga ggccaugaag cguacggcc ugaacaccaa 240

ccaccggauc guggugugcu cggagaacag ccugcaguuc uucaugccgg ugcuggcgc 300  
 ccucuucauc ggcguggccg ugcggccgc gaacgacauc uacaacgagc gggagcugcu 360  
 gaacagcaug gggaucagcc agccgaccgu gguguucgug agcaagaagg gccugcagaa 420  
 gauccugaac gugcagaaga agcugccau cauccagaag aucaucauca uggacagcaa 480  
 gaccgacuac cagggcuucc agucgaugua cacguucgug accagccacc uccgcggg 540  
 cuucaacgag uacgacuucg uccggagag cuucgaccgg gacaagacca ugcggcugau 600

caugaacagc agcggcagca cggccugcc gaagggggug gcccugccgc accggaccgc 660

cugcgugcgc uucucgcacg cccgggaccc caucuucggc aaccagauca uccggacac 720

cgc当地ug agcguggugc cguuccacca cggcuucggc auguucacga cccugggcua 780

ccucaucugc ggc当地ccggg ugguccugau guaccgguuc gaggaggagc uguuccugcg 840

gagccugcag gacuacaaga uccagagcgc gc当地cugug cccgacccgu ucagcuucuu 900

cgc当地agc acccugaucg acaaguacga ccugucgaac cugcagcaga uc当地cagcgg 960

gggc当地ccggc cugagcaagg aggugggcga ggccguggcc aagcgguucc accuccccgg 1020

cauccgcccag ggc当地acggcc ugaccgagac cacgagcgcg auccugauca ccccccggg 1080

ggacgacaag cggccggccg ugga当地caagg guccccc当地 uucgaggcca agguggugga 1140

ccuggacacc gcaagaccc ugga当地cugaa ccagccggc gagcugugc ugc当地ggggcc 1200

gaugaucaug agcggcuacg ugaacaaccc ggaggccacc aacgccc当地 ucgacaagga 1260

cggc当地ggcug cacagccggc acaucgc当地 cugggacgag gacgagc当地 uciucaucgu 1320

cgaccggcug aagucgc当地 ucaaguacaa ggccuaccag guggccggg cc当地agcugga 1380

gagcaucugc cuccagcacc ccaaca当地u cgacgccc当地 guggccggg ugccggacga 1440

cgacgccc当地 gagecugccgg cc当地ggguggu ggugcuggag cacggcaaga ccaugacgga 1500

gaaggagauc guc当地acuacg ugga当地cagcca ggugaccacc gcaagaagc ugc当地ggggccg 1560

cgugguguuc guggacgagg uccc当地aagg ccugaccgg aagcucgc当地 cccggaagau 1620

ccg当地gagauc cugaucaagg ccaagaagg ggcaagauc gccguguaag acuaguua 1680

agacugacua gccc当地auggg cc当地cccaacg ggccc当地cuc cccuucuugc accgagauua 1740

auaaaaaaaaaaaa aaaaaaaaaaaaa aaaaaaaaaaaaa aaaaaaaaaaaaa aaaaaaaaaaaaa 1800

aaaaaaaaugca uuuuuuuuuuuuu uuuuuuuuuuuuu uuuuuuuuuuuuu u 1841

<210> 93

<211> 1857

<212> DNA

<213> Artificial Sequence

<220><223> ppLuc(GC) - ag - A64 - SL

<400> 93

gggagaaagc uugaggaugg aggacgccaa gaacaucaag aaggccgg cggccuucua 60

cccgcuggag gacgggaccg cggc当地gagca gc当地ccacaag gccaugaaagc gguacgccc当地 120

ggugccggc acgaucgc当地 ucacccgacgc ccacauccgag guc当地acauc ccaucgc当地 180

guacuucgag augagc当地gugc gccuggccga ggc当地augaaag cgguaacggcc ugaacaccaa 240

ccaccggau cuggugugcu cggagaacag ccugcaguuc uucaugccgg ugcugggcgc	300
ccucuucauc ggcguggccg ucgccccggc gaacgacauc uacaacgagc gggagcugcu	360
gaacagcaug gggauacagcc agccgaccgu gguguucgug agcaagaagg gccugcagaa	420
gauccugaac gugcagaaga agcugccau cauccagaag aucaucauca uggacagcaa	480
gaccgacuac cagggcuucc agucgaugua cacguucgug accagccacc uccgcggg	540
cuucaacgag uacgacuucg ucccgagag cuucgaccgg gacaagacca ugcggugau	600
caugaacagc agcggcagca cccggcugcc gaaggggug gcccugccgc accggaccgc	660
cugcgugcgc uucucgcacg cccgggaccc caucuucggc aaccagauca uccggacac	720
cgccauccug agcguggugc cguuccacca cggcuucggc auguucacga cccuggcua	780
ccucaucugc ggcuuccggg ugguccugau guaccgguiuc gaggaggagc uguuccugcg	840
gagccugcag gacuacaaga uccagagcgc gcugcucgug ccgaccuugu ucagcuuuu	900
cgcgaagagc acccugaucg acaaguacga ccugucgaac cugcacgaga ugcgcagcgg	960
gggcgcggc cugagcaagg aggugggcga ggccguggcc aagcgguucc accuccggg	1020
cauccgcag ggcuacggcc ugaccgagac cacgagcgcg auccugauca ccccgaggg	1080
ggacgacaag cccggcgcgg uggaagggcgg uggcaagggu ggcccguuc uucgaggcca agguguggga	1140
ccuggacacc ggcagaagaccc ugcccggugaa ccagcggggc gagcugugcg ugcggggcc	1200
gaugaucaug agcggcuacg ugaacaaccc ggaggccacc aacgcccua ucgacaagga	1260
cggcuggcug cacagcggcg acaucgccua cugggacgag gacgagcacu ucuucaucgu	1320
cgaccggcug aagugcuga ucaaguacaa gggcuaccag guggcggcgg ccgagcugga	1380
gagcauccug cuccagcacc ccaacaucuu cgacgcccggc guggccggc ugccggacga	1440
cgacgcccggc gagcugccgg ccgcccgggu ggugcuggag cacggcaaga ccaugacgga	1500
gaaggagauc gucgacuacg ugcccggccca ggugaccacc gccaagaagc ugcggggcgg	1560
cgugguguuuc guggacgagg ucccggaaagg ccugaccggg aagcucgacg cccggaagau	1620
ccgcgagauc cugaucaagg ccaagaaggg cggcaagaua ggcguguaag acuaguuaaua	1680
agacugacua gcccgauggg ccuuccaacg gcccuccuc cccuccuugc accgagauua	1740
auaaaaaaaaaaaa aaaaaaaaaaaaa aaaaaaaaaaaaa aaaaaaaaaaaaa aaaaaaaaaaaaa	1800
aaaaaaaaugca uuauuggcggc cguguccacc acggauauca ccguggugga cgcggcc	1857
<210> 94	
<211> 1838	
<212> DNA	
<213> Artificial Sequence	

&lt;220&gt;&lt;223&gt; ppLuc(GC) - ag - A64 - N32

&lt;400&gt; 94

gggagaaaagc uugaggaugg aggacgccaa gaacaucaag aaggcccgg cgccuucua	60
cccgcuggag gacgggaccg ccggcgagca gcuccacaag gccaugaagc gguacgccc	120

ggugccgggc acgaucgccc ucacccgacgc ccacauccgag gucgacauca ccuacgccc	180
guacuucgag augagcgugc gccuggccga ggcuaugaaag cgguacggcc ugaacaccaa	240
ccaccggau cuggugugcu cggagaacag ccugcaguuc uucaugccgg ugcugggcgc	300
ccucuucauc ggcguggccg ucgcggccgc gaacgacauc uacaacgagc gggagcugcu	360
gaacagcaug gggauacagcc agccgaccgu gguguucgug agcaagaagg gccugcagaa	420
gauccugaac gugcagaaga agcugcccau cauccagaag aucaucauca uggacagcaa	480
gaccgacuac cagggcuucc agucgaugua cacguucgug accagccacc ucccggccgg	540

cuucaacgag uacgacuucg ucccggagag cuucgaccgg gacaagacca ucggccugau	600
caugaacagc agcggcagca ccggccugcc gaagggggug gcccugccgc accggaccgc	660
cugcugcgc uucugcagc cccgggaccc caucuucggc aaccagauca uccggacac	720
cgcacuccug agcguggugc cguuccacca cggcuucggc auguucacga cccuggcua	780
ccucaucugc ggcuuccggg ugguccugau guaccgguuc gaggaggagc uguuccugcg	840
gagccugcag gacuacaaga uccagagcgc gcugcucgug ccgaccugu ucagcuuuu	900
cgcacccaggc acccugauacg acaaguacga ccugucgaac cugcagcaga ucggccagg	960

ggcgccccg cugagcaagg aggugggcga ggccguggcc aagcgguucc accuccggg	1020
cauccggccag ggcuacggcc ugaccgagac cacgagcgcg auccugauca ccccccgggg	1080
ggacgacaag ccggggcgccg uggaagggu ggcccggguuuc uucgaggcca aggugggugga	1140
ccuggacacc ggcacccggcc ugggcugaa ccagcggggc gagcugugcg ugcggggcc	1200
gaugaucaug agcgccuacg ugaacaaccc ggaggccacc aacgcccua ucgacaagga	1260
cggcuggcug cacagcggcg acaucgccua cugggacgag gacgagcacu ucuucaucgu	1320
cgcacccggc aagucgcuga ucaaguacaa gggcuaccag guggccggccgg ccgagcugga	1380

gagcauccug cuccagcacc ccaacaucuu cgacgcccggc guggccgggc ugcccggacga	1440
cgacgcccggc gagcugccgg ccggcguggu ggugcuggag cacggcaaga ccaugacgga	1500
gaaggagauc gucgacuacg ugcccggccaa ggugaccacc gccaagaagc ugcggggcg	1560
cgugguguuc guggacgagg ucccggggccu ccugaccggg aagcucgacg cccggaagau	1620
ccgcgagauc cugaucagg ccaagaagg ccggcaagauc gccguguaag acuaguuaaua	1680
agacugacua gcccgauuggg ccucccaacg ggcccuccuc cccuccuugc accgagauua	1740

aaaaaaaaaaaa aaaaaaaaaaa aaaaaaaaaaa aaaaaaaaaaa aaaaaaaaaaa aaaaaaaaaaa 1800

aaaaaaaaugca uccccucua gacaauugga auuccaua 1838

<210> 95

<211> 15

<212> RNA

<213> Artificial

<220><223> Description of sequence: generic stabilizing sequence of the

formula (C/U)CCANxCCC(U/A)PyxUC(C/U)CC

<220><221> variation

<222> (1)..(1)

<223> /replace="cytosine" /replace="uracile"

<220><221> misc\_feature

<222> (1)..(1)

<223> nucleic acid = cytosine or uracil

<220><221> misc\_feature

<222> (5)..(5)

<223> Nx = a, g, c or u or any other nucleic acid

<220><221> variation

<222> (5)..(5)

<223> /replace="cytosine" /replace="uracile" /replace="guanosine"

/replace="adenosine", or any other nucleic acid

<220><221> repeat\_unit

<222> (5)..(5)

<223> x = any number

<220><221> misc\_feature

<222> (9)..(9)

<223> nucleic acid = uracil or adenosine

<220><221> variation

<222> (9)..(9)

<223> /replace="uracile" /replace="adenosine"

<220><221> misc\_feature

<222> (10)..(10)

<223> Py = pyrimidine

<220><221> repeat\_unit

<222> (10)..(10)

<223> x = any number

<220><221> variation

<222> (10)..(10)

<223> /replace="pyrimidine"

<220><221> misc\_feature

<222> (13)..(13)

<223> nucleic acid = cytosine or uracil

<220><221> variation

<222> (13)..(13)

<223> /replace="cytosine" /replace="uracile"

<400> 95

ncccanccnn ucncc

15