



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

(11) 공개번호 10-2014-0137455  
(43) 공개일자 2014년12월02일

- |  |  |
|--|--|
| <p>(51) 국제특허분류(Int. Cl.)<br/> <i>C12N 15/11</i> (2006.01) <i>C12N 15/67</i> (2006.01)<br/> <i>C12N 15/85</i> (2006.01) <i>A61K 31/713</i> (2006.01)<br/> <i>A61K 35/12</i> (2006.01)</p> <p>(21) 출원번호 10-2014-7030132</p> <p>(22) 출원일자(국제) 2013년03월27일<br/>                 심사청구일자 없음</p> <p>(85) 번역문제출일자 2014년10월27일</p> <p>(86) 국제출원번호 PCT/EP2013/000936</p> <p>(87) 국제공개번호 WO 2013/143698<br/>                 국제공개일자 2013년10월03일</p> <p>(30) 우선권주장<br/>                 PCT/EP2012/001337 2012년03월27일<br/>                 유럽특허청(EPO)(EP)<br/>                 PCT/EP2012/002446 2012년06월08일<br/>                 유럽특허청(EPO)(EP)</p> | <p>(71) 출원인<br/>                 큐어백 게임바하<br/>                 독일, 72076 튀빙엔, 폴 에르리츠 스트라세 15</p> <p>(72) 발명자<br/>                 테스, 안드레아스<br/>                 독일, 72127 쿠스테르던젠, 하인든부르크스트라세 9</p> <p>칼렌, 칼-요셉<br/>                 독일 50226 퀴니히스도르프 마리엔호퍼르 웨그 24</p> <p>(74) 대리인<br/>                 특허법인필엔은지</p> |
|--|--|

전체 청구항 수 : 총 58 항

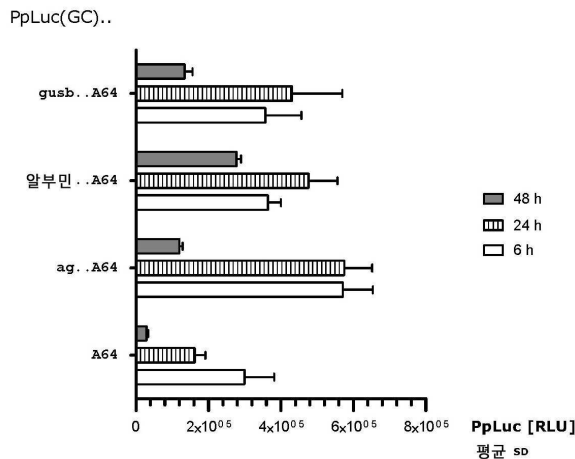
(54) 발명의 명칭 **인공 핵산 분자**

**(57) 요약**

본 발명은 개방 해독 틀 하나 이상과, 알부민 유전자의 3'-UTR 또는 알부민 유전자의 3'-UTR의 변이체로부터 유래하는 핵산을 포함하는 3'-UTR 요소 하나 이상을 포함하는 인공 핵산 분자에 관한 것이다. 본 발명은 또한 유전자 치료법 및/또는 유전자 백신화에 있어서 이와 같은 인공 핵산 분자의 용도에 관한 것이기도 하다. 뿐만 아니라, 본 발명은 이와 같은 3'-UTR 요소를 포함하는 핵산 서열로부터의 단백질 발현을 안정화 및/또는 연장하기 위한, 알부민 유전자의 3'-UTR 또는 알부민 유전자의 3'-UTR의 변이체로부터 유래하는 핵산 서열을 포함하는 3'-UTR 요소의 용도에 관한 것이기도 하다.

**대표도 - 도1**

상이한 3'-UTR을 가지는 mRNA들로부터의 PpLuc(HeLa)



## 특허청구의 범위

### 청구항 1

- a. 개방 해독 틀(ORF) 하나 이상; 및
- b. 알부민 유전자의 3'-UTR 또는 알부민 유전자의 3'-UTR의 변이체로부터 유래하는 핵산 서열을 포함하는 3'-비번역 영역 요소(3'-UTR 요소) 하나 이상

을 포함하는 인공 핵산 분자로서,

상기 개방 해독 틀은, 특히 3'-UTR 요소가 래트 알부민 유전자의 3'-UTR로부터 유래하는 경우, 베타-글로빈을 암호화하지 않는 것이 바람직하고,

상기 개방 해독 틀은, 특히 3'-UTR 요소가 사람 알부민 유전자의 3'-UTR로부터 유래하는 경우, 사람 인자 IX를 암호화하지 않는 것이 바람직한 인공 핵산 분자.

### 청구항 2

제1항에 있어서, 상기 3'-UTR 요소 하나 이상은 상기 인공 핵산 분자로부터의 단백질 생산을 안정화/연장하는 것인 인공 핵산 분자.

### 청구항 3

제1항 또는 제2항에 있어서, 상기 3'-UTR 요소 하나 이상은, 척추동물 알부민 유전자의 3'-UTR 또는 이의 변이체, 바람직하게는 포유동물 알부민 유전자, 예를 들어 마우스 알부민 유전자의 3'-UTR 또는 이의 변이체, 더욱 바람직하게는 영장류 알부민 유전자, 특히 사람 알부민 유전자 또는 올리브 개코 원숭이 유래 알부민 유전자의 3'-UTR 또는 이것들의 변이체, 더더욱 바람직하게는 유전자 은행 수탁 번호 NM\_000477.5에 의한 사람 알부민 유전자의 3'-UTR 또는 이의 변이체로부터 유래하는 핵산 서열을 포함하는 것인 인공 핵산 분자.

### 청구항 4

제1항 내지 제3항 중 어느 하나의 항에 있어서, 상기 3'-UTR 요소 하나 이상은 서열 번호 1, 서열 번호 2, 서열 번호 32, 서열 번호 33, 서열 번호 34 또는 서열 번호 35에 따른 핵산 서열에 대한 동일성이 약 40% 이상, 바람직하게는 약 50% 이상, 바람직하게는 약 60% 이상, 바람직하게는 약 70% 이상, 더욱 바람직하게는 약 80% 이상, 더욱 바람직하게는 약 90% 이상, 더더욱 바람직하게는 약 95% 이상, 더더욱 바람직하게는 약 99% 이상인 핵산 서열을 포함하거나 이 핵산 서열로 이루어져 있거나, 또는 상기 3'-UTR 요소 하나 이상은 서열 번호 1, 서열 번호 2, 서열 번호 32, 서열 번호 33, 서열 번호 34 또는 서열 번호 35에 따른 핵산 서열에 대한 동일성이 약 40% 이상, 바람직하게는 약 50% 이상, 바람직하게는 약 60% 이상, 바람직하게는 약 70% 이상, 더욱 바람직하게는 약 80% 이상, 더욱 바람직하게는 약 90% 이상, 더더욱 바람직하게는 약 95% 이상, 더더욱 바람직하게는 약 99% 이상인 핵산 서열의 단편을 포함하거나 이 단편으로 이루어져 있는 것인 인공 핵산 분자.

### 청구항 5

제4항에 있어서, 상기 단편의 길이는 약 50 뉴클레오티드 이상, 바람직하게는 약 75 뉴클레오티드 이상, 더욱 바람직하게는 약 100 뉴클레오티드 이상, 더더욱 바람직하게는 약 125 뉴클레오티드 이상, 가장 바람직하게는 약 150 뉴클레오티드 이상인 것인 인공 핵산 분자.

### 청구항 6

제1항 내지 제5항 중 어느 하나의 항에 있어서, 상기 3'-UTR 요소 하나 이상의 길이는 약 50 뉴클레오티드 이상, 바람직하게는 약 75 뉴클레오티드 이상, 더욱 바람직하게는 약 100 뉴클레오티드 이상, 더더욱 바람직하게는 약 125 뉴클레오티드 이상, 가장 바람직하게는 약 150 뉴클레오티드 이상인 것인 인공 핵산 분자.

### 청구항 7

제1항 내지 제6항 중 어느 하나의 항에 있어서, 상기 인공 핵산 분자는

c. 폴리(A) 서열 및/또는 폴리아데닐화 신호  
를 추가로 포함하는 것인 인공 핵산 분자.

**청구항 8**

제7항에 있어서, 상기 폴리아데닐화 신호는 3'-UTR 요소의 내부에 위치하는 것인 인공 핵산 분자.

**청구항 9**

제7항 또는 제8항에 있어서, 상기 폴리아데닐화 신호는 공통 서열 NN(U/T)ANA[여기서, N은 A 또는 U임], 바람직하게는 AA(U/T)AAA 또는 A(U/T)(U/T)AAA를 포함하는 것인 인공 핵산 분자.

**청구항 10**

제7항 내지 제9항 중 어느 하나의 항에 있어서, 상기 폴리아데닐화 신호, 바람직하게는 공통 서열 NNUANA는 3'-UTR의 3'-말단 쪽 약 50 뉴클레오티드 미만만큼 상류에 위치하는 것인 인공 핵산 분자.

**청구항 11**

제7항 내지 제10항 중 어느 하나의 항에 있어서, 상기 폴리(A) 서열의 길이는 약 20~약 300개 아데닌 뉴클레오티드, 바람직하게는 약 40~약 200개 아데닌 뉴클레오티드, 더욱 바람직하게는 약 50~약 100개 아데닌 뉴클레오티드, 더더욱 바람직하게는 약 60~약 70개 아데닌 뉴클레오티드인 것인 인공 핵산 분자.

**청구항 12**

제1항 내지 제11항 중 어느 하나의 항에 있어서, 상기 개방 해독 틀은 알부민, 바람직하게는 사람 알부민을 암호화하지 않는 것인 인공 핵산 분자.

**청구항 13**

제1항 내지 제12항 중 어느 하나의 항에 있어서, 상기 인공 핵산 분자는 5'-캡 구조, 폴리(C) 서열, 히스톤 스템-루프 및/또는 IRES 모티프를 추가로 포함하는 것인 인공 핵산 분자.

**청구항 14**

제1항 내지 제13항 중 어느 하나의 항에 있어서, 상기 핵산은 부가의 5'-요소, 바람직하게는 5'-UTR, 프로모터, 또는 5'-UTR과 프로모터 포함 서열을 포함하는 것인 인공 핵산 분자.

**청구항 15**

제1항 내지 제14항 중 어느 하나의 항에 있어서, 상기 인공 핵산 분자, 바람직하게 개방 해독 틀은 G/C에 적어도 부분적으로 변화가 가하여지며, 개방 해독 틀의 G/C 함량은, 바람직하게 야생형 개방 해독 틀의 G/C 함량과 비교되게 증가한 것인 인공 핵산 분자.

**청구항 16**

제1항 내지 제15항 중 어느 하나의 항에 있어서, 상기 개방 해독 틀은 코돈 최적화된 영역을 포함하고, 바람직하게 상기 개방 해독 틀은 코돈 최적화되는 것인 인공 핵산 분자.

**청구항 17**

제1항 내지 제16항 중 어느 하나의 항에 있어서, 상기 인공 핵산 분자는 RNA 분자, 바람직하게는 mRNA 분자인 것인 인공 핵산 분자.

**청구항 18**

a. 개방 해독 틀 및/또는 클로닝 위치; 및

b. 알부민 유전자의 3'-UTR로부터 유래하거나 또는 알부민 유전자의 3'-UTR의 변이체로부터 유래하는 핵산 서열을 포함하는 3'-비번역 영역 요소(3'-UTR 요소) 하나 이상

을 포함하는 벡터로서,

상기 개방 해독 틀은, 특히 3'-UTR 요소가 래트 알부민 유전자의 3'-UTR로부터 유래하는 경우, 베타-글로빈을 암호화하지 않는 것이 바람직하고,

상기 개방 해독 틀은, 특히 3'-UTR 요소가 사람 알부민 유전자의 3'-UTR로부터 유래하는 경우, 사람 인자 IX를 암호화하지 않는 것이 바람직한 벡터.

**청구항 19**

제18항에 있어서, 상기 3'-UTR 요소 하나 이상은, 척추동물 알부민 유전자의 3'-UTR 또는 이의 변이체, 바람직하게는 포유동물 알부민 유전자, 예를 들어 마우스 알부민 유전자의 3'-UTR 또는 이의 변이체, 더욱 바람직하게는 영장류 알부민 유전자, 특히 사람 알부민 유전자 또는 올리브 개코 원숭이 유래 알부민 유전자의 3'-UTR 또는 이의 변이체, 더더욱 바람직하게는 유전자 은행 수탁 번호 NM\_000477.5에 의한 사람 알부민 유전자의 3'-UTR 또는 이의 변이체로부터 유래하는 핵산 서열을 포함하거나 이 핵산 서열로 이루어진 것인 벡터.

**청구항 20**

제18항 또는 제19항에 있어서, 상기 3'-UTR 요소 하나 이상은 서열 번호 1, 서열 번호 2, 서열 번호 32, 서열 번호 33, 서열 번호 34 또는 서열 번호 35에 따른 핵산 서열에 대한 동일성이 약 40% 이상, 바람직하게는 약 50% 이상, 바람직하게는 약 60% 이상, 바람직하게는 약 70% 이상, 더욱 바람직하게는 약 80% 이상, 더욱 바람직하게는 약 90% 이상, 더더욱 바람직하게는 약 95% 이상, 더더욱 바람직하게는 약 99% 이상인 핵산 서열을 포함하거나 이 핵산 서열로 이루어져 있거나, 또는 상기 3'-UTR 요소 하나 이상은 서열 번호 1, 서열 번호 2, 서열 번호 32, 서열 번호 33, 서열 번호 34 또는 서열 번호 35에 따른 핵산 서열에 대한 동일성이 약 40% 이상, 바람직하게는 약 50% 이상, 바람직하게는 약 60% 이상, 바람직하게는 약 70% 이상, 더욱 바람직하게는 약 80% 이상, 더욱 바람직하게는 약 90% 이상, 더더욱 바람직하게는 약 95% 이상, 더더욱 바람직하게는 약 99% 이상인 핵산 서열의 단편을 포함하거나 이 단편으로 이루어져 있는 것인 벡터.

**청구항 21**

제20항에 있어서, 상기 단편의 길이는 약 50 뉴클레오티드 이상, 바람직하게는 약 75 뉴클레오티드 이상, 더욱 바람직하게는 약 100 뉴클레오티드 이상, 더더욱 바람직하게는 약 125 뉴클레오티드 이상, 가장 바람직하게는 약 150 뉴클레오티드 이상인 것인 벡터.

**청구항 22**

제18항 내지 제21항 중 어느 하나의 항에 있어서, 상기 3'-UTR 요소 하나 이상의 길이는 약 50 뉴클레오티드 이상, 바람직하게는 약 75 뉴클레오티드 이상, 더욱 바람직하게는 약 100 뉴클레오티드 이상, 더더욱 바람직하게는 약 125 뉴클레오티드 이상, 가장 바람직하게는 약 150 뉴클레오티드 이상인 것인 벡터.

**청구항 23**

제18항 내지 제22항 중 어느 하나의 항에 있어서, 상기 벡터는

c. 폴리(A) 서열 및/또는 폴리아데닐화 신호

를 추가로 포함하는 것인 벡터.

**청구항 24**

제23항에 있어서, 상기 폴리아데닐화 신호는 3'-UTR 요소의 내부에 위치하는 것인 벡터.

**청구항 25**

제23항 또는 제24항에 있어서, 상기 폴리아데닐화 신호는 공통 서열 NN(U/T)ANA[여기서, N은 A 또는 U임], 바람직하게는 AA(U/T)AAA 또는 A(U/T)(U/T)AAA를 포함하는 것인 벡터.

**청구항 26**

제23항 내지 제25항 중 어느 하나의 항에 있어서, 상기 폴리아데닐화 신호, 바람직하게는 공통 서열 NNUANA는

3'-UTR 요소의 3'-말단 쪽 약 50 뉴클레오티드 미만만큼 상류에 위치하는 것인 벡터.

**청구항 27**

제23항 내지 제26항 중 어느 하나의 항에 있어서, 상기 폴리(A) 서열의 길이는 약 20~약 300개 아데닌 뉴클레오티드, 바람직하게는 약 40~약 200개 아데닌 뉴클레오티드, 더욱 바람직하게는 약 50~약 100개 아데닌 뉴클레오티드, 더더욱 바람직하게는 약 60~약 70개 아데닌 뉴클레오티드인 것인 벡터.

**청구항 28**

제18항 내지 제27항 중 어느 하나의 항에 있어서, 상기 개방 해독 틀은 알부민, 바람직하게는 사람 알부민을 암호화하지 않는 것인 벡터.

**청구항 29**

제18항 내지 제28항 중 어느 하나의 항에 있어서, 상기 개방 해독 틀은 사람 인자 IX 또는 리포터 단백질, 바람직하게는 GFP 단백질, 루시페라제 단백질 또는 글로빈 단백질, 또는 이것들의 변이체들을 암호화하지 않는 것인 벡터.

**청구항 30**

제18항 내지 제29항 중 어느 하나의 항에 있어서, 상기 벡터는 폴리(C) 서열, 히스톤 스템-루프 및/또는 IRES-모티프를 추가로 포함하는 것인 벡터.

**청구항 31**

제18항 내지 제30항 중 어느 하나의 항에 있어서, 상기 벡터는 5' 요소, 바람직하게는 5'-UTR, 프로모터, 또는 5'-UTR과 프로모터를 포함하는 서열을 추가로 포함하는 것인 벡터.

**청구항 32**

제18항 내지 제31항 중 어느 하나의 항에 있어서, 상기 벡터는 적어도 부분적으로 G/C에 변화가 가하여지고, 바람직하게 상기 개방 해독 틀은 적어도 부분적으로 G/C에 변화가 가하여지며, 바람직하게 상기 개방 해독 틀의 G/C 함량은, 야생형 개방 해독 틀의 G/C 함량과 비교되게 증가한 것인 벡터.

**청구항 33**

제18항 내지 제32항 중 어느 하나의 항에 있어서, 상기 개방 해독 틀은 코돈 최적화된 영역을 포함하고, 바람직하게 상기 개방 해독 틀은 코돈 최적화되는 것인 벡터.

**청구항 34**

제18항 내지 제33항 중 어느 하나의 항에 있어서, 상기 벡터는 DNA 벡터인 것인 벡터.

**청구항 35**

제18항 내지 제34항 중 어느 하나의 항에 있어서, 상기 벡터는 플라스미드 벡터 또는 바이러스 벡터, 바람직하게는 플라스미드 벡터인 것인 벡터.

**청구항 36**

제18항 내지 제35항 중 어느 하나의 항에 있어서, 상기 벡터는 제1항 내지 제17항 중 어느 하나의 항에 의한 인공 핵산 분자를 포함하는 것인 벡터.

**청구항 37**

제18항 내지 제36항 중 어느 하나의 항에 있어서, 상기 벡터는 원형 분자인 것인 벡터.

**청구항 38**

제37항에 있어서, 암호화 사슬의 폴리(A) 서열 또는 3'-UTR 요소 뒤에는 원형 벡터 분자의 선형화에 사용될 제

한 위치가 5'→3' 방향으로 따르는 것인 벡터.

**청구항 39**

제1항 내지 제17항 중 어느 하나의 항에 의한 인공 핵산 분자 또는 제18항 내지 제38항 중 어느 하나의 항에 의한 벡터를 포함하는 세포.

**청구항 40**

제39항에 있어서, 상기 세포는 포유동물 세포인 것인 세포.

**청구항 41**

제39항 또는 제40항에 있어서, 상기 세포는 포유동물 개체의 세포, 바람직하게는 포유동물 개체, 바람직하게는 사람 개체의 분리된 세포인 것인 세포.

**청구항 42**

제1항 내지 제17항 중 어느 하나의 항에 의한 인공 핵산 분자, 제18항 내지 제38항 중 어느 하나의 항에 의한 벡터, 또는 제39항 내지 제41항 중 어느 하나의 항에 의한 세포를 포함하는 약학 조성물.

**청구항 43**

제42항에 있어서, 상기 약학 조성물은 하나 이상의 약학적으로 허용 가능한 희석제 및/또는 부형제 및/또는 하나 이상의 보조제를 추가로 포함하는 것인 약학 조성물.

**청구항 44**

의약품으로서 사용되기 위한, 제1항 내지 제17항 중 어느 하나의 항에 의한 인공 핵산 분자, 제18항 내지 제38항 중 어느 하나의 항에 의한 벡터, 제39항 내지 제41항 중 어느 하나의 항에 의한 세포, 또는 제42항 또는 제43항에 의한 약학 조성물.

**청구항 45**

백신으로서 사용되거나 유전자 치료법에서 사용되기 위한, 제1항 내지 제17항 중 어느 하나의 항에 의한 인공 핵산 분자, 제18항 내지 제38항 중 어느 하나의 항에 의한 벡터, 제39항 내지 제41항 중 어느 하나의 항에 의한 세포, 또는 제42항 또는 제43항에 의한 약학 조성물.

**청구항 46**

제1항 내지 제17항 중 어느 하나의 항에 의한 인공 핵산 분자, 제18항 내지 제38항 중 어느 하나의 항에 의한 벡터, 제39항 내지 제41항 중 어느 하나의 항에 의한 세포, 또는 제42항 또는 제43항에 의한 약학 조성물을, 이것들의 투여를 필요로 하는 개체에 투여하는 단계를 포함하는, 질환을 치료 또는 예방하는 방법.

**청구항 47**

제1항 내지 제17항 중 어느 하나의 항에 의한 인공 핵산 분자 또는 제18항 내지 제38항 중 어느 하나의 항에 의한 벡터를 세포에 형질 감염시키는 단계를 포함하는, 질환을 치료 또는 예방하는 방법.

**청구항 48**

제47항에 있어서, 상기 세포의 형질 감염은 시험관 내/생체 외에서 수행되고, 상기 형질 감염된 세포는 이것의 투여를 필요로 하는 개체, 바람직하게는 사람인 환자에 투여되는 것인 방법.

**청구항 49**

제48항에 있어서, 상기 시험관 내 형질 감염될 세포는 개체, 바람직하게는 사람인 환자의 분리된 세포인 것인 방법.

**청구항 50**

제46항 내지 제49항 중 어느 하나의 항에 있어서, 상기 방법은 백신화 방법 또는 유전자 치료법인 것인 방법.

**청구항 51**

핵산 분자, 바람직하게는 mRNA 분자 또는 벡터로부터의 단백질 생산을 안정화 및/또는 연장하는 방법으로서, 상기 방법은 핵산 분자, 바람직하게는 mRNA 분자 또는 벡터와 3'-UTR 요소를 결합하는 단계를 포함하고, 상기 3'-UTR 요소는 알부민 유전자의 3'-UTR 또는 알부민 유전자의 3'-UTR의 변이체로부터 유래하는 핵산 서열을 포함하거나 이 핵산 서열로 이루어진 방법.

**청구항 52**

핵산 분자, 바람직하게는 mRNA 분자 또는 벡터로부터 단백질 생산을 안정화 및/또는 연장하기 위한 3'-UTR의 용도로서, 상기 3'-UTR 요소는 알부민 유전자의 3'-UTR 또는 알부민 유전자의 3'-UTR의 변이체로부터 유래하는 핵산 서열을 포함하거나 이 핵산 서열로 이루어진 용도.

**청구항 53**

제51항에 의한 방법 또는 제52항에 의한 용도로서, 상기 3'-UTR 요소는 척추동물 알부민 유전자의 3'-UTR 또는 이의 변이체, 바람직하게는 포유동물 알부민 유전자, 예를 들어 마우스 알부민 유전자의 3'-UTR 또는 이의 변이체, 더욱 바람직하게는 영장류 알부민 유전자, 특히 사람 알부민 유전자 또는 올리브 개코 원숭이 유래 알부민 유전자의 3'-UTR 또는 이의 변이체, 더더욱 바람직하게는 유전자 은행 수탁 번호 NM\_000477.5에 의한 사람 알부민 유전자의 3'-UTR 또는 이의 변이체로부터 유래하는 핵산 서열을 포함하거나 이 핵산 서열로 이루어진 것인 방법 또는 용도.

**청구항 54**

제51항 내지 제53항 중 어느 하나의 항에 있어서, 상기 3'-UTR 요소는 서열 번호 1, 서열 번호 2, 서열 번호 32, 서열 번호 33, 서열 번호 34 또는 서열 번호 35에 따른 핵산 서열에 대한 동일성이 약 40% 이상, 바람직하게는 약 50% 이상, 바람직하게는 약 60% 이상, 바람직하게는 약 70% 이상, 더욱 바람직하게는 약 80% 이상, 더욱 바람직하게는 약 90% 이상, 더더욱 바람직하게는 약 95% 이상, 더더욱 바람직하게는 약 99% 이상인 핵산 서열을 포함하거나 이 핵산 서열로 이루어져 있거나, 또는 상기 3'-UTR 요소는 서열 번호 1, 서열 번호 2, 서열 번호 32, 서열 번호 33, 서열 번호 34 또는 서열 번호 35에 따른 핵산 서열에 대한 동일성이 약 40% 이상, 바람직하게는 약 50% 이상, 바람직하게는 약 60% 이상, 바람직하게는 약 70% 이상, 더욱 바람직하게는 약 80% 이상, 더욱 바람직하게는 약 90% 이상, 더더욱 바람직하게는 약 95% 이상, 더더욱 바람직하게는 약 99% 이상인 핵산 서열의 단편을 포함하거나 이 단편으로 이루어져 있는 것인 방법 또는 용도.

**청구항 55**

제54항에 있어서, 상기 단편의 길이는 약 50 뉴클레오티드 이상, 바람직하게는 약 75 뉴클레오티드 이상, 더욱 바람직하게는 약 100 뉴클레오티드 이상, 더더욱 바람직하게는 약 125 뉴클레오티드 이상, 가장 바람직하게는 약 150 뉴클레오티드 이상인 것인 방법 또는 용도.

**청구항 56**

제51항 내지 제55항 중 어느 하나의 항에 있어서, 상기 3'-UTR 요소의 길이는 약 50 뉴클레오티드 이상, 바람직하게는 약 75 뉴클레오티드 이상, 더욱 바람직하게는 약 100 뉴클레오티드 이상, 더더욱 바람직하게는 약 125 뉴클레오티드 이상, 가장 바람직하게는 약 150 뉴클레오티드 이상인 것인 방법 또는 용도.

**청구항 57**

제1항 내지 제17항 중 어느 하나의 항에 의한 인공 핵산 분자, 제18항 내지 제38항 중 어느 하나의 항에 의한 벡터, 제39항 내지 제41항 중 어느 하나의 항에 의한 세포 및/또는 제42항 또는 제43항에 의한 약학 조성물을 포함하는 키트 또는 이것을 포함하는 부분들로 이루어진 키트.

**청구항 58**

제57항에 있어서, 상기 키트는 사용 지침, 형질 감염용 세포, 보조제, 약학 조성물 투여 수단, 인공 핵산 분자, DNA 벡터 또는 약학 조성물을 용해 또는 희석함에 있어 약학적으로 허용 가능한 캐리어 및/또는 약학적으로 허

용 가능한 용액을 추가로 포함하는 것인 키트.

**명세서**

**기술분야**

[0001] 본 발명은 개방 해독 틀, 즉 3'-UTR 요소와, 임의로는 폴리(A) 서열 및/또는 폴리아데닐화 신호를 포함하는 인공 핵산 분자에 관한 것이다. 본 발명은 또한, 바람직하게 유전자 치료법 및/또는 유전자 백신화 분야에서 사용될, 3'-UTR 요소를 포함하는 벡터, 상기 인공 핵산 분자를 포함하는 약학 조성물, 그리고 인공 핵산 분자, 이 인공 핵산 분자를 포함하는 벡터 및/또는 약학 조성물을 포함하는 키트에 관한 것이기도 하다.

**배경기술**

[0002] 유전자 치료법과 유전자 백신화는 가장 전도 유망하고도 급속도로 발전하고 있는 현대 의학의 방법들에 속한다. 이 방법들은 다양한 질병을 치료하기 위해 매우 특이적이며 구체적인 사양을 제공할 수 있다. 특히 선천성 유전 병뿐만 아니라 자가 면역 질병, 암 또는 종양 관련 질병, 그리고 염증성 질병이 이러한 치료 접근법의 대상이 될 수 있다. 또한 이와 같은 접근법들에 의해 상기와 같은 질병들의 (조기) 발병을 예방하는 것이 구현되기도 한다.

[0003] 유전자 치료법의 저변에 있는 주요 개념상 근거는, 특정 질병의 병리적 병태와 연관된, 손상 유전자 발현을 적당히 조정하는 것이다. 병리적으로 변형된 유전자 발현은 필수 유전자 생산물, 예를 들어 신호 전달 인자, 예를 들어 호르몬, 항존 인자(housekeeping factor), 대사 효소 또는 구조 단백질 등의 결실 또는 과 생산을 초래할 수 있다. 변형된 유전자 발현은 전사 및/또는 번역의 잘못된 조절(misregulation)에 기인할 수 있을 뿐만 아니라 특정 단백질을 암호화하는 ORF 내 돌연 변이에 기인할 수도 있다. 병리적 돌연 변이들은, 예를 들어 염색체 이상, 또는 더욱 특이적인 돌연 변이, 예를 들어 점 돌연 변이 또는 해독 틀 이동 돌연 변이에 의하여 발생할 수 있는데, 이 돌연 변이들은 모두 유전자 생산물의 기능 상실과, 잠재적으로는 이 유전자 생산물 기능의 완전 상실을 초래한다. 그러나 전사 또는 번역의 잘못된 조절은, 만일 돌연 변이들이 세포의 전사 또는 번역 기구에 관여하는 단백질을 암호화하는 유전자에 영향을 미치면 발생할 수도 있다. 이와 같은 돌연 변이들은 (보통 말하는) 기능을 가지는 유전자들의 병리적 상향 조절 또는 하향 조절을 유도할 수 있다. 이와 같은 조절 기능을 발휘하는 유전자 생산물을 암호화하는 유전자들은, 예를 들어 전사 인자, 신호 수용체 또는 메신저 단백질 등일 수 있다. 그러나 이와 같은 조절 단백질 암호화 유전자들의 기능 상실은, 임의의 경우 손상된 유전자 생산물의 추가 하류로서 작용을 하는 기타 인자들을 인공적으로 도입함으로써 회복될 수 있다. 이와 같은 유전자 결함들은 또한 손상된 유전자 자체의 치환을 통한 유전자 치료법에 의하여 만회될 수도 있다.

[0004] 유전자 백신화는 선택된 항원, 예를 들어 박테리아 표면, 바이러스 입자 또는 종양 항원 등의 특징을 이루는 구성 요소들에 대하여 원하는 면역 반응을 일으킬 수 있다. 일반적으로 백신화는 현대 의학의 중심축을 이루는 성과물들 중 하나이다. 그러나 현재는 단지 소수의 질병들에 대한 유효 백신들만이 시판되고 있는 실정이다. 또한 백신화를 통해 예방될 수 없는 감염들이 매년 수 백만명의 사람들을 여전히 침노하고 있다.

[0005] 보통 백신은 "제1 세대", "제2 세대" 또는 "제3 세대" 백신으로 세분될 수 있다. "제1 세대 백신"은 통상 전 유기체 백신(whole organism vaccine)이다. 상기 제1 세대 백신은 살아있으면서 약독화된 병원체 또는 사멸된 병원체, 예를 들어 바이러스 또는 박테리아 등을 기반으로 한다. 살아있으면서 약독화된 백신의 주요 단점은 병원체가 생명을 위협하는 변이체로 역전될 위험이 있다는 점이다. 그러므로 이와 같이 약독화된 병원체는 본질적으로 예측 불가능한 위험을 여전히 보유하고 있을 수 있다. 사멸된 병원체는 특이적 면역 반응을 유발하는데 요망될 만큼 효과적일 수 없다. 이와 같은 위험들을 최소화하기 위해 "제2 세대" 백신들이 개발되었다. 상기 제2 세대 백신은 통상 병원체로부터 유래하는 한정된 항원들 또는 재조합 단백질 구성 성분들로 이루어진 서브유닛 백신이다.

[0006] 유전자 백신, 즉 유전자 백신화용 백신은 보통 "제3 세대" 백신인 것으로 이해된다. 상기 제3 세대 백신은 통상 생체 내에서 병원체나 종양 항원의 특징을 나타내는 펩티드 또는 단백질 (항원) 단편들을 발현시킬 수 있는 유전자 조작된 핵산 분자로 구성된다. 유전자 백신은 환자에게 투여되었을 때 수용성 세포에 의해 흡수됨으로써 발현된다. 투여된 핵산이 발현되면 암호화된 단백질이 생산된다. 이러한 단백질이 환자의 면역계에 의해 외래 물질로서 인지되는 경우, 면역 반응은 촉발된다.

- [0007] 진술된 사항들로부터 알 수 있는 바와 같이, 유전자 치료법과 유전자 백신화 방법은 둘 다 본질적으로 핵산 분자를 환자에 투여한 다음, 암호화된 유전 정보를 전사 및/또는 번역시키는 것을 바탕으로 한다. 대안적으로 유전자 백신화 또는 유전자 치료법은 또한 치료될 환자로부터 특이적 체세포를 분리한 다음, 이 세포를 시험관 내에서 형질 감염시키고, 이와 같이 처리된 세포들을 환자에게 다시 투여하는 단계들을 포함할 수도 있다.
- [0008] RNA 뿐만 아니라 DNA는 유전자 치료법 또는 유전자 백신화 과정 중 투여되는 핵산 분자로서 사용될 수 있다. DNA는 비교적 안정적이고 다루기 용이한 것으로 알려져 있다. 그러나 DNA를 사용함에 있어서는, 투여된 DNA 단편이 환자의 유전체에 원치 않게 삽입되어, 손상된 유전자의 기능이 잠재적으로 상실될 위험이 있다. 추가의 위험으로서는, 항 DNA 항체들이 원치 않게 생산된다는 점이 있다. 또 다른 단점으로는, DNA가 투여되어 전사/번역이 진행될 때 달성될 수 있는 암호화 펩티드 또는 단백질의 발현 수준이 제한된다는 점이 있다. 다른 어떤 이 유들보다도, 투여된 DNA의 발현 수준은, DNA 전사를 조절하는 특이적 전사 인자들의 존재에 좌우될 것이다. 이러한 인자들이 존재하지 않으면 DNA 전사를 통해서만 만족할만한 양의 RNA가 생산되지 않을 것이다. 결과적으로, 번역된 펩티드 또는 단백질의 수준은 제한된다.
- [0009] 유전자 치료법이나 유전자 백신화에 DNA 대신 RNA가 사용되면, 이 DNA가 원치 않게 유전자에 통합되어 항 DNA 항체가 생산될 위험이 최소화되거나 없어질 수 있다. 그러나 RNA는 오히려 도처에 존재하는 RNA 분해효소(ubiquitous RNAase)에 의해 용이하게 분해될 수 있는 불안정 분자 중인 것으로 간주한다.
- [0010] 생체 내에서의 RNA 분해는 RNA 반감기(RNA half-life time) 조절에 기여한다. 이러한 효과는 진핵 생물 유전자 발현의 조절을 훌륭하게 변조(tuning)하는 것으로 간주 및 입증되었다[Friedel *et al.*, Conserved principles of mammalian transcriptional regulation revealed by RNA half-life, *Nucleic Acid Research*, 2009, 1-12]. 즉 자연 발생 mRNA 각각은 자체의 개별 반감기를 가지는데, 이 반감기는 mRNA가 유도된 유전자에 의존한다. 반감기로 말미암아 이와 같은 유전자의 발현 수준이 조절되는 것이다. 불안정한 RNA는 개별 시점들에서 유전자를 일시적으로 발현시킴에 있어 중요한 역할을 한다. 그러나 반감기가 긴 RNA(long-lived RNA)는 변별적 단백질들의 축적 또는 유전자의 연속 발현과 연관될 수 있다. 생체 내 mRNA의 반감기는 또한, 예를 들어 인슐린 유사 성장 인자 I, 액틴 및 알부민 mRNA의 경우에서 보인 바와 같이, 환경적 요인들, 예를 들어 호르몬 처리에 의존적일 수도 있다[Johnson *et al.*, Newly synthesized RNA: Simultaneous measurement in intact cells of transcription rates and RNA stability of insulin-like growth factor I, actin, and albumin in growth hormone-stimulated hepatocytes, *Proc. Natl. Acad. Sci.*, Vol. 88, pp. 5287-5291, 1991].
- [0011] 유전자 치료법과 유전자 백신화에는 보통 안정한 RNA가 요망된다. 이는, 한편으로는 RNA 서열에 의해 암호화된 생산물이 생체 내에 축적될 것이라는 사실 때문이다. 다른 한편으로, RNA는, 그것이 축적되는 과정 중에 적당한 투여형으로 제조될 때와 투여될 때, 자체의 구조적 및 기능적 일체성을 유지하여야 한다. 그러므로 RNA 분자가 조기에 분해 또는 붕괴하는 것이 방지되도록, 유전자 치료법 또는 유전자 백신화에 안정한 RNA 분자들을 제공하는데 상당한 주의가 기울여졌다.
- [0012] 핵산 분자의 G/C 함량은 핵산 분자의 안정성에 영향을 미칠 수 있음이 보고된 바 있다. 그러므로 구아닌(G) 및/또는 시토신(C) 잔기들의 양이 증가한 핵산은, 아데닌(A) 및 티민(T) 또는 우라실(U) 뉴클레오티드들을 다량으로 함유하는 핵산보다 기능상으로 더욱 안정할 수 있다. 이와 같은 내용에 있어서, WO 제02/098443호는 번역된 영역 내 서열 변형을 통하여 안정화된 mRNA를 포함하는 약학 조성물을 제공한다. 이와 같은 서열 변형은 유전 암호의 축퇴성(degeneracy)을 이용한다. 그러므로 덜 유리한(RNA 안정성의 관점에서 덜 유리한) 뉴클레오티드 조합을 포함하는 코돈들은, 암호화된 아미노산 서열을 변형시키지 않고 대안적 코돈들에 의해 치환될 수 있다. 이와 같은 RNA 안정화 방법은, 원하는 아미노산 서열의 공간이 남는 것이 허용되지 않는 단일 RNA 분자 각각의 특이적 뉴클레오티드 서열이 제공됨으로 인하여 제한된다. 또한 이러한 접근법은 RNA의 암호화 영역에 제한된다.
- [0013] mRNA 안정화에 대한 대안적 사양으로서, 자연 발생 진핵 생물 mRNA 분자들이 특징적인 안정화 요소들을 포함한다는 사실이 파악되었다. 예를 들어 자연 발생 진핵 생물 mRNA 분자들은 자체의 5' 말단 및/또는 3'-말단에 소위 비번역 영역(UTR)(각각 5'-UTR 및/또는 3'-UTR)을 포함할 수 있을 뿐만 아니라, 기타 구조적 특징들, 예를 들어 5'-캡 구조 또는 3'-폴리(A) 미부를 포함할 수도 있다. 상기 5'-UTR 및 3'-UTR은 둘 다 통상 게놈 DNA로부터 전사되므로, 미성숙 mRNA의 요소이다. 성숙 mRNA의 특징을 이루는 구조적 특징들, 예를 들어 5'-캡 및 3'-폴리(A) 미부(폴리(A) 미부 또는 폴리(A) 서열이라고도 칭하여짐)는 보통 mRNA 가공 중 전사된(미성숙) mRNA에 부가된다.
- [0014] 3'-폴리(A) 미부는 통상 전사된 mRNA의 3'-미부에 부가된, 아데닌 뉴클레오티드들의 단조로운 서열 구획이다.

상기 3'-폴리(A) 미부는 약 400개 이하의 아데닌 뉴클레오티드를 포함할 수 있다. 이와 같은 3'-폴리(A) 미부의 길이는 개별 mRNA의 안정성에 잠재적으로 중요한 요소라는 사실이 파악되었다.

[0015] 또한 α-글로빈 mRNA의 3'-UTR은 α-글로빈 mRNA의 안정성에 기여하는 것으로 널리 알려진 중요한 인자일 수 있다[Rodgers *et al.*, Regulated α-globin mRNA decay is a cytoplasmic event proceeding through 3'-to-5' exosome-dependent decapping, RNA, 8, pp. 1526-1537, 2002]. α-글로빈 mRNA의 3'UTR은, 그 존재 자체가 시험관 내 mRNA 안정성과 상관 있는 특이적 리보핵산단백질-복합체, 즉 α-복합체의 형성에 관여하는 것이 분명하다[Wang *et al.*, An mRNA stability complex functions with poly(A)-binding protein to stabilize mRNA *in vitro*, Molecular and Cellular biology, Vol 19, No. 7, July 1999, p. 4552-4560].

[0016] mRNA 안정성에 영향을 미치는 인자들과는 상관없이, 투여된 핵산 분자의 표적 세포 또는 조직에 의한 효과적인 번역은, 유전자 치료비용 또는 유전자 백신화용 핵산 분자를 사용하는 임의의 접근법에 중요하다. mRNA 안정성 조절과 아울러, 다량의 mRNA의 번역은 구조적 특징들, 예를 들어 UTR, 5'-캡 및 3'-폴리(A) 미부에 의해 조절된다. 이와 같은 내용에 있어서, 폴리(A) 미부의 길이는 또한 번역 효율에 중요한 역할을 할 수 있는 것으로 보고된 바 있다. 그러나 3'-요소가 안정화되면 번역을 저해하는 효과가 나타날 수도 있다.

### 발명의 내용

#### 해결하려는 과제

[0017] 유전자 치료법 및/또는 유전자 백신화를 수행하는데 적당할 수 있는 핵산 분자를 제공하는 것이 본 발명의 목적이다. 특히 번역 효율에 있어서 유의적인 기능 상실을 나타내지 않고 조기 분해 또는 붕괴에 대해 안정화된 mRNA 종을 제공하는 것이 본 발명의 목적이다. 이와 같이 유전자 치료법 및/또는 유전자 백신화에 사용함에 있어 순응할 수 있는, 우수한 mRNA 종을 암호화하는 핵산 분자를 제공하는 것이 본 발명의 다른 목적이다. 유전자 치료법 및/또는 유전자 백신화에 사용할 약학 조성물을 제공하는 것이 본 발명의 추가의 목적이다. 요약하면, 비용 효율적이면서 간단한 접근법에 의해 선행 기술의 전술된 단점들을 해결하는, 개선된 핵산 종을 제공하는 것이 본 발명의 목적인 것이다.

#### 과제의 해결 수단

[0018] 유전자 치료법 및/또는 유전자 백신화를 수행하는데 적당할 수 있는 핵산 분자를 제공하는 것이 본 발명의 목적이다. 특히 번역 효율에 있어서 유의적인 기능 상실을 나타내지 않고 조기 분해 또는 붕괴에 대해 안정화된 mRNA 종을 제공하는 것이 본 발명의 목적이다. 이와 같이 유전자 치료법 및/또는 유전자 백신화에 사용함에 있어 순응할 수 있는, 우수한 mRNA 종을 암호화하는 핵산 분자를 제공하는 것이 본 발명의 다른 목적이다. 유전자 치료법 및/또는 유전자 백신화에 사용할 약학 조성물을 제공하는 것이 본 발명의 추가의 목적이다. 요약하면, 비용 효율적이면서 간단한 접근법에 의해 선행 기술의 전술된 단점들을 해결하는, 개선된 핵산 종을 제공하는 것이 본 발명의 목적인 것이다.

[0019] 본 발명의 바탕을 이루는 목적은 청구된 특허 대상들에 의해 해결된다.

[0020] 명확성과 이해의 용이함을 위하여 다음과 같은 정의들이 제공된다. 이러한 정의들에 대해서 언급된 임의의 기술적 특징은 본 발명의 각각의 구체에 및 모든 구체에 대해 적용될 수 있다. 부가의 정의 및 설명은 이러한 구체에의 내용에 구체적으로 제공될 수 있다.

[0021] 적응성 면역 반응: 적응성 면역 반응은 통상 면역계의 항원 특이적 반응인 것으로 이해된다. 항원 특이성은 특이 병원체 또는 병원체 감염 세포에 대해 맞추어진 반응들을 일으킬 수 있다. 이와 같이 맞추어진 반응들을 확립하는 능력은 보통 체내에서 "기억 세포"에 의해 유지된다. 만일 병원체가 1회 이상 체내 감염되면, 특이적 기억 세포들은 상기 병원체를 신속하게 제거하는데 사용된다. 이와 같은 내용에 있어서, 적응성 면역 반응의 제1 단계는, 항원 제시 세포에 의해 항원 특이적 면역 반응을 유도할 수 있는 무경험(naive) 항원 특이 T 세포들 또는 상이한 면역 세포들의 활성화다. 이와 같은 활성화는 무경험 T 세포가 일정하게 통과하는 림프양 조직 및 장기 내에서 일어난다. 항원 제시 세포로서 사용될 수 있는 세포의 3가지 유형으로서는, 수지상 세포, 대식 세포 및 B 세포가 있다. 이러한 세포들 각각은 면역 반응을 유도함에 있어서 변별적 기능을 가진다. 수지상 세포는 식세포 작용 및 대음세포작용에 의해 항원을 흡수할 수 있으며, 예를 들어 외래 항원과의 접촉을 통해 자극되어 국소 림프양 조직으로 이동한 다음, 여기에서 성숙한 수지상 세포로 분화될 수 있다. 대식 세포는 입자형 항원, 예를 들어 박테리아를 섭취한 후, 감염성 체제나 기타 적당한 자극에 의해 MHC 분자들을 발현하도록 유도된다. B 세포 특유의 능력, 즉 가용성 단백질 항체 자체의 수용기를 통해 이 단백질 항원을 결합 및 내부화하는 능력

은 또한 T 세포를 유도하는데 중요하게 작용할 수도 있다. MHC 분자는 통상 T 세포에 항원을 제시하는데 관여한다. 여기서 MHC 분자가 항원을 제시하게 되면, T 세포가 활성화되고, 그 결과 T 세포가 증식되어 무장된(armed) 효과기 T 세포로 분화된다. 효과기 T 세포의 가장 중요한 기능은, CD8+ 세포 독성 T 세포에 의해 감염된 세포를 사멸하는 것과, Th1 세포에 의해 대식 세포를 활성화함으로써 상기 대식 세포와 함께 세포 매개 면역을 나타내는 것, 그리고 Th2 및 Th1 세포 둘 다에 의해 B 세포가 활성화됨으로써 상이한 항체 군들이 생산되도록 하고, 이를 통해 체액 면역 반응을 일으키는 것이다. T 세포는, 항원을 직접 인지하여 항원과 결합하지 않는 대신에, 병원체 유래 단백질 항원의 짧은 펩티드 단편(예를 들어 소위 에피토프)(기타 세포의 표면 위에 있는 MHC 분자들에 결합하는 부분)을 인지하는, 자체의 T 세포 수용기에 의해 항원을 인식한다.

[0022] 적응성 면역계: 적응성 면역계는 본질적으로 병원체의 생장을 막거나 예방하는데 기여한다. 적응성 면역계는 통상 척추 동물 면역계에 (면역성을 나타내기 위해) 특이 병원체를 인지 및 기억하는 능력과, 병원체와 마주칠 때마다 더욱 강력하게 병원체를 공격할 태세를 갖추는 능력을 제공함으로써 적응성 면역 반응을 조절한다. 상기 적응성 면역계는, 체세포 과 돌연 변이(가속화된 체세포 돌연 변이 과정), V(D)J 재조합(항원 수용기 유전자 분절들의 비가역적 유전자 재조합)으로 인해 순응성이 크다. 이러한 기작은 소수의 유전자들이, 각각의 개별 림프구 상에서만 특유하게 발현되는 다수의 상이한 항원 수용기들을 생산하도록 만든다. 유전자 재배열은 각각의 세포의 DNA 내에 비가역적인 변화를 초래하므로, 이러한 세포, 예를 들어 지속 기간이 긴 특이적 면역성에 대해 중요한 역할을 담당하는 기억 B 세포 및 기억 T 세포의 자손(후손)들은 모두 동일한 수용기 특이성을 암호화하는 유전자들을 물려줄 것이다.

[0023] 보조제/보조제 구성 성분: 보조제 또는 보조제 구성 성분은 최 광의에서 통상 기타 제제, 예를 들어 약물 또는 백신의 효과를 개선(예를 들어 증강)할 수 있는 약학적 및/또는 면역학적 제제이다. 이는 최 광의에서 해석될 것이며, 광범위한 물질들을 말한다. 통상적으로 이와 같은 물질들은 항원의 면역원성을 증가시킬 수 있다. 예를 들어 보조제는 선천성 면역계에 의해 인지될 수 있으며, 예를 들어 선천성 면역 반응을 유도할 수도 있다. "보조제"는 통상 적응성 면역 반응은 유도하지 않는다. 이러한 관점에서 "보조제"는 항원으로서 간주하지 않는다. 보조제의 작용 방식은 적응성 면역 반응을 일으키는 항원에 의해 촉발되는 효과들과는 구분된다.

[0024] 항원: 본 발명의 내용에 있어서, "항원"이란 통상 면역계, 바람직하게는 적응성 면역계에 의해 인지될 수 있으며, 예를 들어 적응성 면역 반응의 일환으로 항체 및/또는 항원 특이 T 세포를 생산함으로써 항원 특이 면역 반응을 촉발할 수 있는 물질을 말한다. 통상적으로 항원은 MHC에 의해 T 세포에 제시될 수 있는 펩티드 또는 단백질일 수 있거나 또는 이와 같은 펩티드 또는 단백질을 포함할 수 있다.

[0025] 인공 핵산 분자: 인공 핵산 분자는 통상 자연 발생하지 않는 핵산 분자, 예를 들어 DNA 또는 RNA로서 이해될 수 있다. 다시 말해서, 인공 핵산 분자는 비 자연 핵산 분자로서 이해될 수 있다. 이러한 핵산 분자는 자체의 개별 서열(자연적으로 발생하지 않는 서열) 및/또는 기타 자연 발생하지 않는 뉴클레오티드의 변형, 예를 들어 구조적 변형으로 인해 비 자연 생산되는 것일 수 있다. 인공 핵산 분자는 DNA 분자, RNA 분자, 또는 DNA 부분들과 RNA 부분들을 포함하는 하이브리드 분자일 수 있다. 통상적으로 인공 핵산 분자는 뉴클레오티드의 원하는 인공 서열(이중 서열)에 상응하도록 유전자 조작 방법에 의해 디자인 및/또는 생산될 수 있다. 이와 같은 내용 중, 인공 서열은 보통 자연 발생할 수 없는 서열, 즉 야생형 서열과 하나 이상의 뉴클레오티드만큼이 상이한 서열이다. 용어 "야생형 서열"은 자연에서 발생하는 서열인 것으로 이해될 수 있다. 또한 용어 "인공 핵산 분자"는 "하나의 단일 분자"를 의미하는 것으로 한정되지 않고, 통상 동일 분자들의 총체를 포함하는 것으로 이해된다. 그러므로 이는 특정 분취량 만큼에 포함된 동일한 분자들 다수 개와 관련될 수 있다.

[0026] 2 시스트론형 RNA, 다중 시스트론형 RNA: 2 시스트론형 또는 다중 시스트론형 RNA는 통상 개방 해독 틀(ORF)을 2개 가질 수 있거나(2 시스트론형) 그 이상 가질 수 있는(다중 시스트론형) RNA, 바람직하게는 mRNA이다. 이와 같은 내용에 있어서 개방 해독 틀은 펩티드 또는 단백질로 번역되는 코돈들의 서열이다.

[0027] 캐리어/중합체 캐리어: 본 발명의 내용 중 캐리어는 통상 다른 화합물(화물; cargo)의 운반 및/또는 복합체화를 촉진하는 화합물일 수 있다. 중합체 캐리어는 통상 중합체로 이루어진 캐리어이다. 캐리어는 공유 또는 비 공유 상호 작용에 의해 자체의 화물과 결합할 수 있다. 캐리어는 핵산, 예를 들어 RNA 또는 DNA를 표적 세포에 운반할 수 있다. 캐리어는 (몇몇 구체예를 위해서) 양이온 구성 성분일 수 있다.

[0028] 양이온 구성 성분: 용어 "양이온 구성 성분"이란, 통상적으로 pH값이 통상적으로 1-9일 때, 바람직하게는 pH값이 9 또는 그 이하일 때(예를 들어 5-9일 때), 또는 8 또는 그 이하일 때(예를 들어 5-8일 때), 또는 7 또는 그 이하일 때(예를 들어 5-7일 때), 가장 바람직하게는 생리적 pH 범위(예를 들어 7.3-7.4)일 때 양으로 하전되는(양이온인) 하전 분자를 말한다. 그러므로 양이온 구성 성분은 임의의 양으로 하전된 화합물 또는 중합체, 바람

직하게는 생리적 조건, 특히 생체 내 생리적 조건 하에서 양으로 하전되는 양이온 펩티드 또는 단백질일 수 있다. "양이온 펩티드 또는 단백질"은, 예를 들어 Arg, His, Lys 또는 Orn으로부터 선택되는, 하나 이상의 양 하전 아미노산 또는 하나보다 많은 양 하전 아미노산을 포함할 수 있다. 그러므로 "다가 양이온" 구성 성분도 소정의 조건 하에서 하나보다 많은 양전하를 나타내는 구성 성분의 범위 내에 포함된다.

[0029] **5'-캡:** 5'-캡은, 일반적으로 성숙한 mRNA의 5'-말단을 "캡핑(capping)"하는 실체, 통상적으로 변형된 뉴클레오티드 실체이다. 5'-캡은 통상 변형된 뉴클레오티드, 특히 구아닌 뉴클레오티드의 유도체에 의해 형성될 수 있다. 바람직하게 5'-캡은 5'-5'-트리포스페이트 결합을 통해 5'-말단에 결합한다. 5'-캡은 메틸화되어, 예를 들어 m7GpppN[식 중, N은 5'-캡을 운반하는 핵산의 5' 말단 뉴클레오티드, 통상적으로 RNA의 5'-말단임]이 될 수 있다. 5'-캡 구조에 관한 추가의 예들로서는 글리세릴, 역전 데옥시 무 염기 잔기(부), 4',5'-메틸렌 뉴클레오티드, 1-(베타-D-에리트로푸라노실) 뉴클레오티드, 4'-티오 뉴클레오티드, 탄소환 뉴클레오티드, 1,5-무수 핵시톨 뉴클레오티드, L-뉴클레오티드, 알파-뉴클레오티드, 변형 염기 뉴클레오티드, 트레오-펜토푸라노실 뉴클레오티드, 비환형 3',4'-세코 뉴클레오티드, 비환형 3,4-디하이드록시부틸 뉴클레오티드, 비환형 3,5-디하이드록시펜틸 뉴클레오티드, 3'-3'-역전 뉴클레오티드 부, 3'-3'-역전 무 염기 부, 3'-2'-역전 뉴클레오티드 부, 3'-2'-역전 무 염기 부, 1,4-부탄디올포스페이트, 3'-포스포라미데이트, 핵실포스페이트, 아미노핵실포스페이트, 3'-포스페이트, 3'-포스포티오에이트, 포스포로디티오에이트, 또는 가교형 또는 비 가교형 메틸포스포네이트 부를 포함한다.

[0030] **세포성 면역/세포성 면역 반응:** 세포성 면역는 통상 대식 세포, 자연 살해(NK) 세포, 항원 특이적 세포 독성 T 림프구의 활성화와, 항원에 반응하여 다양한 시토킨들을 방출하는 것과 관련된다. 더욱 일반적인 관점에서, 세포성 면역는 항체에 바탕을 두지 않고, 면역계 세포의 활성화에 바탕을 둔다. 통상적으로 세포성 면역 반응은, 예를 들어 세포, 예를 들어 수지상 세포나 기타 세포와 같은 특이적 면역 세포 내에서 세포 자살을 유도할 수 있는 항원 특이적 세포 독성 T 림프구를 활성화하는 것과, 외래 항원의 에피토프를 자체의 표면에 전시하도록 만드는 것을 특징으로 할 수 있다. 이러한 세포들은 바이러스 감염될 수 있거나, 세포 내 박테리아로 감염될 수 있거나, 또는 종양 항원을 전시하는 암 세포일 수 있다. 추가의 특징들로서는, 대식 세포 및 자연 살해 세포를 활성화하는 것과, 이 세포들이 병원체를 파괴할 수 있는 능력, 그리고 이 세포를 자극하여 다양한 시토킨(적응성 면역 반응 및 선천성 면역 반응에 관여하는 기타 세포들의 기능에 영향을 미침)을 분비하게 하는 것을 들 수 있다.

[0031] **DNA:** DNA는 데옥시리보핵산의 일반적인 축약어이다. 상기 DNA는 핵산 분자, 즉 뉴클레오티드들로 이루어진 중합체이다. 이와 같은 뉴클레오티드들은 보통 당 부(데옥시리보스), 염기 부 및 포스페이트 부로 이루어져 있고, 특징적인 주쇄 구조에 의해 중합되는, 데옥시-아데노신-모노포스페이트, 데옥시-티미딘-모노포스페이트, 데옥시-구아노신-모노포스페이트 및 데옥시-시티딘-모노포스페이트 단량체이다. 상기 주쇄 구조는 통상 뉴클레오티드의 당 부, 즉 제1 단량체의 데옥시리보스와, 인접하는 제2 단량체의 포스페이트 부 사이의 포스포디에스테르 결합에 의해 형성된다. 단량체의 구체적인 순서, 즉 당/포스페이트-주쇄에 결합된 염기들의 순서는 DNA 서열이라 칭하여진다. DNA는 단일 사슬형 또는 이중 사슬형일 수 있다. 이중 사슬 형태에 있어서, 제1 사슬의 뉴클레오티드들은 통상, 예를 들어 A/T-염기쌍 형성 및 G/C-염기쌍 형성에 의해 제2 사슬의 뉴클레오티드들과 혼성화된다.

[0032] **에피토프:** 에피토프("항원 결정기"라고도 칭하여짐)는 T 세포 에피토프 및 B 세포 에피토프로 구분될 수 있다. 본 발명의 내용에 있어서 단백질의 일부 또는 T 세포 에피토프는, 바람직하게 길이가 약 6-약 20개 아미노산, 또는 그 이상의 아미노산인 단편들, 예를 들어 제I군 MHC 분자에 의해 가공 및 제시되는 단편들, 바람직하게 길이가 약 8-약 10개 아미노산, 예를 들어 8, 9 또는 10개 아미노산(또는 심지어 11 또는 12개 아미노산)인 단편들, 또는 제II군 MHC 분자에 의해 가공 및 제시되는 단편들, 바람직하게 길이가 약 13개 이상 아미노산, 예를 들어 13, 14, 15, 16, 17, 18, 19, 20개 또는 그 이상의 아미노산인 단편들을 포함할 수 있는데, 이 경우, 이와 같은 단편들은 아미노산 서열의 임의의 부분으로부터 선택될 수 있다. 이러한 단편들은 통상적으로 T 세포에 의해 펩티드 단편과 MHC 분자로 이루어진 복합체의 형태로서 인지되는데, 즉 이러한 단편들은 통상적으로 그 자체의 천연 형태로서 인지되지 않는다. B 세포 에피토프는 통상 항체에 의해 자체의 천연 형태로서 인지될 수 있는, 본원에 정의된 (천연) 단백질 또는 펩티드 항원의 외부 표면에 위치하는 단편으로서, 바람직하게는 5-15개의 아미노산, 더욱 바람직하게는 5-12개의 아미노산, 더더욱 바람직하게는 6-9개의 아미노산을 가지는 단편이다.

[0033] 뿐만 아니라 단백질 또는 펩티드의 이러한 에피토프는, 이 단백질 또는 펩티드의 변이체들로서 본원에 예시된 것들 중 임의의 것들로부터 선택될 수도 있다. 이와 같은 내용에 있어서, 항원 결정기는, 본원에 정의된 바와 같은 단백질 또는 펩티드의 분절들로 이루어진 형태학적 또는 불연속적 에피토프로서, 본원에 정의된 바와 같은

단백질 또는 펩티드의 아미노산 서열에는 불연속적으로 존재하되, 다만 모여있을 때에는 3차원 구조를 형성하는 에피토프이거나, 또는 단일 폴리펩티드 사슬로 이루어진 연속적 또는 선형 에피토프일 수 있다.

[0034] 서열 단편: 서열 단편은 통상 전장 서열(예를 들어 핵산 분자 또는 아미노산 서열)의 전장 서열 중 길이가 짧은 부분일 수 있다. 그러므로 단편은 통상 전장 서열 내 상응하는 구획과 동일한 서열로 이루어져 있는 것이다. 본 발명의 내용에 있어서 서열의 바람직한 단편은, 이 단편이 유래한 분자 내 실체의 연속 구획에 상응하는 실체(예를 들어 뉴클레오티드 또는 아미노산)의 연속 구획으로 이루어져 있으며, 이와 같은 단편은 이 단편이 유래한 전체 분자(즉 전장 분자)의 20% 이상, 바람직하게는 30% 이상, 더욱 바람직하게는 40% 이상, 더욱 바람직하게는 50% 이상, 더더욱 바람직하게는 60% 이상, 더더욱 바람직하게는 70% 이상, 그리고 가장 바람직하게는 80% 이상에 해당한다.

[0035] G/C 변형 핵산: GC 변형 핵산은 통상, 바람직하게 야생형 서열에 비하여 구아노신 및/또는 시토신 뉴클레오티드의 수가 증가한, 변형 야생형 서열을 기반으로 하는 핵산, 바람직하게는 본원에 정의된 인공 핵산 분자일 수 있다. 이와 같이 구아노신 및/또는 시토신 뉴클레오티드 수의 증가는, 아데노신 또는 티미딘 뉴클레오티드를 포함하는 코돈을, 구아노신 또는 시토신 뉴클레오티드를 포함하는 코돈으로 치환함으로써 이루어질 수 있다. 만일 DNA 또는 RNA의 암호화 영역 중 G/C 함량이 증가하게 되면, 유전 암호의 축퇴성이 적용된다. 그러므로 코돈 치환은 암호화된 아미노산 잔기들을 변형하지 않되, 핵산 분자의 G/C 함량을 배타적으로 증가시키는 것이 바람직하다.

[0036] 유전자 치료법: 유전자 치료법은 통상 환자의 몸 또는 환자의 몸에서 분리된 구성 요소, 예를 들어 분리된 조직/세포를 펩티드나 단백질을 암호화하는 핵산으로 치료하는 것을 의미하는 것으로 이해될 수 있다. 유전자 치료법은 통상 a) 핵산, 바람직하게는 본원에 정의된 바와 같은 인공 핵산 분자를 (투여 경로가 무엇이든 간에) 환자에게 직접 투여하거나, 또는 환자로부터 분리된 세포/조직에 투여하여(시험관 내 투여) 환자의 세포를 생체 내/생체 외 또는 시험관 내 형질 감염시키는 단계; b) 도입된 핵산 분자를 전사 및/또는 번역하는 단계; 그리고 임의로는 c) 만일 핵산이 환자에게 직접 투여되지 않는다면, 분리되어 형질 감염된 세포를 환자에 다시 투여하는 단계와 같은 단계들 중 하나 이상의 단계를 포함할 수 있다.

[0037] 유전자 백신화: 유전자 백신화는 통상 항원이나 면역원 또는 이것들의 단편을 암호화하는 핵산 분자의 투여에 의한 백신화로서 이해될 수 있다. 핵산 분자는 개체의 체내에 투여될 수 있거나, 또는 개체로부터 분리된 세포에 투여될 수 있다. 항원 또는 면역원으로, 개체의 몸을 구성하는 임의의 세포를 형질 감염시키거나 개체로부터 분리된 세포를 형질 감염시켰을 때, 상기 항원 또는 면역원은 이와 같이 형질 감염된 세포에 의해 발현될 수 있으며, 그 결과 적응성 면역 반응, 즉 항원 특이적 면역 반응을 유도하는 면역계에 제시될 수 있다. 그러므로 유전자 백신화는 통상 a) 본원에 정의된 바와 같은 핵산, 바람직하게는 인공 핵산 분자를 개체, 바람직하게는 환자에 투여하거나, 또는 개체, 바람직하게는 환자로부터 분리된 세포에 투여하여, 일반적으로 개체의 세포를 생체 내 또는 시험관 내 형질 감염시키는 단계; b) 도입된 핵산 분자를 전사 및/또는 번역하는 단계; 그리고 임의로는 c) 만일 핵산이 환자에게 직접 투여되지 않는다면, 분리되어 형질 감염된 세포를 개체, 바람직하게는 환자에 다시 투여하는 단계와 같은 단계들 중 하나 이상의 단계를 포함한다.

[0038] 이중 서열: 만일 2개의 서열이 동일한 유전자로부터 유래할 수 없는 것이면(다시 말해서 비록 이중 서열이 동일한 유기체로부터 유래할 수 있는 것일지라도, 이 서열들이 동일한 핵산 분자, 예를 들어 동일한 mRNA 내에 자연적으로(자연에서) 발생하지 않는 것이면), "이중"인 것으로 이해된다.

[0039] 체액성 면역/체액성 면역 반응: 체액성 면역이란 통상적으로 항체의 생산, 임의로는 항체의 생산을 수반하는 부수적 과정을 말한다. 체액성 면역 반응은 통상, 예를 들어 Th2 활성화 및 시토킨 생산, 배 중심 형성 및 동형 전환(isotype switching), 친화성 성숙 및 기억 세포 생산을 특징으로 할 수 있다. 체액성 면역은 또한 통상적으로 병원체 및 독소 중화, 고전적 보체 활성화, 그리고 식세포 작용의 흡소닌에 의한 촉진 및 병원체 제거를 포함하는 항체의 효과기 기능을 말하는 것일 수도 있다.

[0040] 면역원: 본 발명의 내용 중, 면역원은 통상 면역 반응을 자극할 수 있는 화합물로서 이해될 수 있다. 면역원은 펩티드, 폴리펩티드 또는 단백질인 것이 바람직하다. 특히 바람직한 구체예에서, 본 발명의 관점에서의 면역원은 제공된 핵산 분자, 바람직하게는 본원에 정의된 바와 같은 인공 핵산 분자의 번역 생산물이다. 통상적으로 면역원은 최소한 적응성 면역 반응을 유도한다.

[0041] 면역 자극성 조성물: 본 발명의 내용에 있어서, 면역 자극성 조성물은 통상 면역 반응을 유도할 수 있는 성분 하나 이상을 포함하는 조성물, 또는 면역 반응을 유도할 수 있는 성분이 유래할 수 있는 조성물로서 이해될 수

있다. 이러한 면역 반응은 선천성 면역 반응일 수 있거나, 또는 적응성 면역 반응 및 선천성 면역 반응의 조합일 수 있는 것이 바람직하다. 바람직하게 본 발명의 내용에 있어서, 면역 자극성 조성물은 인공 핵산 분자, 더욱 바람직하게는 RNA, 예를 들어 mRNA 분자 하나 이상을 포함한다. 면역 자극성 성분, 예를 들어 mRNA는 적당한 캐리어와 복합체를 형성할 수 있다. 그러므로 면역 자극성 조성물은 mRNA/캐리어-복합체를 포함할 수 있다. 뿐만 아니라, 면역 자극성 조성물은 면역 자극성 성분, 예를 들어 mRNA에 적당한 비이클 및/또는 보조제를 포함할 수 있다.

[0042] **면역 반응:** 면역 반응은 통상적으로 특정 항원에 대한 적응성 면역계의 특이적 반응(소위 특이적 또는 적응성 면역 반응이라 칭함) 또는 선천성 면역계의 비특이적 반응(소위 비특이적 또는 선천성 면역 반응이라 칭함), 또는 이것들의 조합일 수 있다.

[0043] **면역계:** 면역계는 유기체를 감염으로부터 보호할 수 있다. 만일 병원체가 유기체의 물리적 장벽을 통과하는데에 성공하여 이 유기체 내에 들어왔다면, 선천성 면역계는 즉각적이되 비특이적인 반응을 제공한다. 만일 병원체가 이 선천성 면역 반응을 피하였다면, 척추 동물은 제2의 보호층, 즉 적응성 면역계를 가동시키게 된다. 여기서, 면역계는 감염이 진행되는 동안 자체의 반응을 순응시키고, 그 결과, 면역계의 병원체 인지 능력이 개선된다. 이와 같이 개선된 반응은 병원체가 제거된 이후에도 면역 기억의 형태로 유지되며, 이후 동일한 병원체와 마주 칠 때마다 적응성 면역계는 더욱 신속하고 강력하게 병원체를 공격하는 기작을 확립할 수 있다. 이에 따라서 면역계는 선천성 및 적응성 면역계를 포함한다. 통상적으로 이와 같은 두 가지 경로들 각각은, 소위 체액성 구성 성분과 세포성 구성 성분을 포함한다.

[0044] **면역 자극성 RNA:** 본 발명의 내용에 있어서 면역 자극성 RNA(isRNA)는 통상적으로 선천성 면역 반응을 유도할 수 있는 RNA일 수 있다. 상기 면역 자극성 RNA는 보통 개방 해독 틀을 가지지 않으므로 펩티드 항원 또는 면역 원을 제공하지 않지만, 예를 들어 특정 종류의 톨-유사 수용기(TLR) 또는 기타 적당한 수용기와 결합함으로써 말미암아 면역 반응을 유도한다. 그러나 물론 개방 해독 틀을 가지고 펩티드/단백질을 암호화하는 mRNA도 선천성 면역 반응을 유도할 수 있으므로, 이와 같은 mRNA도 면역 자극성 RNA일 수 있다.

[0045] **선천성 면역계:** 선천성 면역계(비 특이적(또는 비특이적) 면역계라고도 알려짐)는 통상적으로 숙주를, 다른 유기체로 인한 감염으로부터 비특이적 방식으로 방어하는 기구와 세포를 포함한다. 이는, 선천성 면역계의 세포들은 유전적 방식으로 병원체를 인지하여 이에 반응할 수 있지만, 적응성 면역계와는 달리, 오랜 기간 동안 지속되거나 보호 작용을 하는 면역성을 숙주에 부여하지 않는다는 것을 의미한다. 선천성 면역계는, 예를 들어 톨-유사 수용기(TLR)의 리간드 또는 기타 보조적 물질, 예를 들어 리포다당체, TNF-알파, CD40 리간드 또는 시토킨, 모노킨, 림포카인, 인터루킨 또는 케모카인, IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, 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, TNF-알파, 성장 인자 및 hGH, 사람 톨-유사 수용기 TLR1, TLR2, TLR3, TLR4, TLR5, TLR6, TLR7, TLR8, TLR9, TLR10의 리간드, 척추 동물 톨-유사 수용기 TLR1, TLR2, TLR3, TLR4, TLR5, TLR6, TLR7, TLR8, TLR9, TLR10, TLR11, TLR12 또는 TLR13의 리간드, NOD 유사 수용기의 리간드, RIG-I 유사 수용기의 리간드, 면역 자극성 핵산, 면역 자극성 RNA(isRNA), CpG-DNA, 항 박테리아 제제 또는 항 바이러스 제제에 의해 활성화될 수 있다. 본 발명에 의한 약학 조성물은 이러한 물질을 하나 이상 포함할 수 있다. 통상적으로 선천성 면역계의 반응은, 화학적 인자들, 예를 들어 특화된 화학적 매개체(시토킨이라 칭하여짐)의 생산을 통한 면역 세포의 감염 위치로의 보충; 보체 캐스케이드의 활성화; 특화된 백혈구에 의한 장기, 조직, 혈액 및 림프 내 외래 물질의 동정 및 제거; 적응성 면역계의 활성화; 및/또는 감염성 제제에 대한 물리적 및 화학적 장벽으로서의 작용을 포함한다.

[0046] **클로닝 위치:** 클로닝 위치는 통상적으로 핵산 서열, 예를 들어 개방 해독 틀을 포함하는 핵산 서열의 삽입에 적당한 핵산 분자의 본질인 것으로 이해된다. 삽입은, 당 업자에게 알려진 임의의 분자 생물학적 방법, 예를 들어 제한(restriction) 및 결찰(ligation)에 의해 수행될 수 있다. 클로닝 위치는 통상 제한 효소 인지 위치(제한 위치)를 하나 이상 포함한다. 이와 같은 하나 이상의 제한 위치들은 이 위치들에서 DNA를 절단하는 제한 효소에 의해 인지될 수 있다. 하나 이상의 제한 위치를 포함하는 클로닝 위치는 또한 다중 클로닝 위치(MCS) 또는 폴리링커(polylinker)라고도 칭하여질 수 있다.

[0047] **핵산 분자:** 핵산 분자는 핵산 구성 성분들을 포함하는 분자, 바람직하게 핵산 구성 성분들로 이루어진 분자이다. 용어 핵산 분자는 바람직하게 DNA 분자 또는 RNA 분자를 말한다. 상기 핵산 분자는 바람직하게 용어 "폴리뉴클레오티드"와 같은 의미로 사용된다. 바람직하게 핵산 분자는 당/포스페이트-주쇄의 포스포디에스테르

결합에 의해 서로 공유 결합하고 있는 뉴클레오티드 단량체들을 포함하거나 이러한 단량체들로 이루어진 중합체이다. 용어 "핵산 분자"는 또한 변형된 핵산 분자, 예를 들어 염기 변형, 당 변형 또는 주쇄 변형된 DNA 또는 RNA 분자를 포함하기도 한다.

- [0048] **개방 해독 틀:** 본 발명의 내용에 있어서 개방 해독 틀(ORF)은 통상 펩티드나 단백질로 번역될 수 있는 몇몇 뉴클레오티드 삼중체의 서열일 수 있다. 개방 해독 틀은 바람직하게 서열의 5'-말단에 개시 코돈, 즉 일반적으로 아미노산 메티오닌을 암호화하는 연속 3개 뉴클레오티드 조를 포함하고, 그 다음에는 보통 뉴클레오티드들 3개가 1조를 이루는 것들(조들) 다수 개로 이루어진, 특정 길이의 영역이 뒤따른다. ORF는 종결 코돈(예를 들어 TAA, TAG, TGA)에 의해 종결되는 것이 바람직하다. 통상적으로 이는 개방 해독 틀의 종결 코돈일 뿐이다. 그러므로 본 발명의 내용에 있어서 개방 해독 틀은, 바람직하게 뉴클레오티드들 3개가 1조를 이룰 수 있는 것들(조들) 다수 개로 이루어져 있으며, 개시 코돈(예를 들어 ATG)으로 개시되어, 바람직하게는 종결 코돈(예를 들어 TAA, TGA 또는 TAG)으로 종결되는 뉴클레오티드 서열이다. 개방 해독 틀은 분리될 수 있거나 길이가 더욱 긴 핵산 서열, 예를 들어 벡터나 mRNA에 통합될 수 있다. 개방 해독 틀은 또한 "단백질 암호화 영역"이라고도 칭하여질 수 있다.
- [0049] **펩티드:** 펩티드 또는 폴리펩티드는 통상 펩티드 결합에 의해 결합하고 있는 아미노산 단량체들의 중합체이다. 펩티드 또는 폴리펩티드는 통상 50개 미만의 단량체 단위들을 포함한다. 그럼에도 불구하고, 용어 "펩티드"는 50개 이상의 단량체 단위를 가지는 분자를 제외하는 것은 아니다. 길이가 긴 펩티드도 폴리펩티드라고 불리기도 하는데, 이는 통상 50~600개의 단량체 단위들을 가진다.
- [0050] **약학적 유효량:** 본 발명의 내용에 있어서 약학적 유효량은 통상, 예를 들어 병리적 조건에서 약학적 효과, 예를 들어 면역 반응, 발현된 펩티드 또는 단백질의 병리적 수준 변경, 또는 결실된 유전자 생산물의 치환을 유도하기에 충분한 양으로 이해된다.
- [0051] **단백질:** 단백질은 통상 펩티드 또는 폴리펩티드를 하나 이상 포함한다. 단백질은 통상 3차원 형태로 접히는데, 이와 같은 형태는 단백질이 자체의 생물학적 기능을 발휘하는데 필요할 수 있다.
- [0052] **폴리(A) 서열:** 폴리(A) 미부 또는 3'-폴리(A) 미부라고도 불리는 폴리(A) 서열은 통상 아데닌 뉴클레오티드, 예를 들어 약 400개 이하의 아데닌 뉴클레오티드, 예를 들어 약 20~약 400개, 바람직하게 약 50~약 400개, 더욱 바람직하게 약 50~약 300개, 더더욱 바람직하게 약 50~약 250개, 가장 바람직하게는 약 60~약 250개 아데닌 뉴클레오티드로 이루어진 서열로서 이해된다. 폴리(A) 서열은 통상 mRNA의 3'-말단에 위치한다. 본 발명의 내용에 있어서, 폴리(A) 서열은 mRNA 또는 기타 임의의 핵산 분자, 예를 들어 벡터, 예를 들어 RNA, 바람직하게는 mRNA 생산에 주형으로서 사용되는 벡터 내에, 예를 들어 벡터의 전사를 통하여 위치하게 될 수 있다.
- [0053] **폴리아데닐화:** 폴리아데닐화는 통상 폴리(A) 서열을 핵산 분자, 예를 들어 RNA 분자, 예를 들어 미성숙 mRNA에 부가하는 것으로 이해된다. 폴리아데닐화는 소위 **폴리아데닐화 신호**에 의해 유도될 수 있다. 이 신호는 바람직하게 폴리아데닐화될 핵산 분자, 예를 들어 RNA 분자의 3'-말단에 있는 뉴클레오티드들의 구획 내에 위치한다. 폴리아데닐화 신호는 통상 아데닌 및 우라실/티민 뉴클레오티드로 이루어진 6량체, 바람직하게는 6량체 서열인 AAUAAA를 포함한다. 기타 서열, 바람직하게 6량체 서열도 고려될 수 있다. 폴리아데닐화는 통상 전구체 mRNA(미성숙 mRNA라고도 칭하여짐)의 가공 중에 발생한다. 통상적으로 (전구체 mRNA에서 성숙 mRNA로의) RNA 성숙은 폴리아데닐화 단계를 포함한다.
- [0054] **제한 위치:** 제한 효소 인지 위치라고도 칭하여지는 제한 위치는 제한 효소에 의해 인지되는 뉴클레오티드 서열이다. 제한 위치는 통상 길이가 짧은 서열, 바람직하게는 앞뒤 역 순상 동서열의(palindromic) 뉴클레오티드 서열, 예를 들어 4~8개의 뉴클레오티드를 포함하는 서열이다. 제한 위치는 제한 효소에 의해 특이적으로 인지되는 것이 바람직하다. 제한 효소는, 통상 제한 위치에서 이 제한 위치를 포함하는 뉴클레오티드 서열을 절단한다. 이중 사슬 뉴클레오티드 서열, 예를 들어 이중 사슬 DNA 서열에 있어서, 제한 효소는 통상 뉴클레오티드 서열의 양 사슬들을 절단한다.
- [0055] **RNA, mRNA:** RNA는 리보핵산의 일반적인 축약어이다. RNA는 핵산 분자, 즉 뉴클레오티드들로 이루어진 중합체이다, 이와 같은 뉴클레오티드는 일반적으로 소위 주쇄를 따라서 서로 연결되어 있는 아데노신-모노포스페이트, 우리딘-모노포스페이트, 구아노신-모노포스페이트 및 시티딘-모노포스페이트 단량체이다. 주쇄는 제1 단량체의 당, 즉 리보스와, 이 제1 단량체에 인접하는 제2 단량체의 포스페이트 부 사이의 포스포디에스테르 결합에 의해 형성된다. 단량체들의 특이적인 연결 관계(succession)는 RNA 서열이라고 칭하여진다. 보통 RNA는, 예를 들어 세포 내에서 DNA 서열의 전사에 의해 생산될 수 있다. 진핵 생물 세포에 있어서, 전사는 통상 핵 또는 미토콘드

리아 내에서 수행된다. DNA의 생체 내 전사는 소위 메신저 RNA(일반적으로 mRNA로 약칭됨)로 가공될, 소위 미성숙 RNA를 생산한다. 예를 들어 진핵 생물 유기체 내 미성숙 RNA의 가공은, 다수의 상이한 전사 후 변형들, 예를 들어 스플라이싱(splicing), 5'-캡핑(5'-capping), 폴리아데닐화, 그리고 핵 또는 미토콘드리아로부터의 배출 등을 포함한다. 이러한 가공 과정들은 통틀어서 RNA의 성숙이라고도 칭하여진다. 성숙한 메신저 RNA는 보통 특정 펩티드 또는 단백질의 아미노산 서열로 번역될 수 있는 뉴클레오티드 서열을 제공한다. 통상적으로, 성숙한 mRNA는 5'-캡, 5'-UTR, 개방 해독 틀, 3'-UTR 및 폴리(A) 서열을 포함한다. 메신저 RNA와는 별도로, 전사 및/또는 번역의 조절에 관여할 수 있는 RNA의 비 암호화 유형 몇 가지가 존재한다.

[0056] **핵산 분자의 서열:** 핵산 분자의 서열은 통상 자체의 뉴클레오티드의 구체적인 고도 개별적인 순서, 즉 연결 관계인 것으로 이해된다. 단백질 또는 펩티드의 서열은 통상 자체의 아미노산의 순서, 즉 연결 관계인 것으로 이해된다.

[0057] **서열 동일성:** 만일 2개 이상의 서열들이 동일한 길이와 동일한 뉴클레오티드 또는 아미노산 순서를 나타낸다면 그 서열들은 동일한 것이다. 동일성%는 통상 2개의 서열들이 동일한 정도를 말해주는 것인데, 다시 말해서, 동일성%는 통상 뉴클레오티드의 서열 위치 중 기준 서열의 동일한 뉴클레오티드들과 일치하는 뉴클레오티드들의 백분율을 말하는 것이다. 동일성 정도가 측정됨에 있어서, 비교될 서열들은 일단 길이(즉 비교될 서열들 중 가장 긴 서열의 길이)가 동일한 것으로 간주한다. 이는, 8개의 뉴클레오티드로 이루어진 제1 서열은 제1 서열을 포함하여 10개의 뉴클레오티드로 이루어진 제2 서열과 80% 동일하다는 것을 의미한다. 다시 말해서, 본 발명의 내용에 있어서, 서열의 동일성은 바람직하게 길이가 동일한 2개 이상의 서열들 중 동일한 위치를 차지하는 서열 내 뉴클레오티드들의 백분율과 관련이 있다. 갭(gap)은, 보통 배열 중 이 갭 자체의 실제 위치와는 상관없이 동일하지 않은 위치인 것으로 간주한다.

[0058] **안정화된 핵산 분자:** 안정화된 핵산 분자는 변형되어, 예를 들어 환경적 요인들 또는 효소에 의한 분해, 예를 들어 엑소뉴클레아제 또는 엔도뉴클레아제 분해에 의한 분해 또는 붕괴에 대해, 변형이 일어나지 않은 핵산 분자보다 더욱 안정하게 된 핵산 분자, 바람직하게는 DNA 또는 RNA 분자이다. 본 발명의 내용에 있어서 안정화된 핵산 분자는 바람직하게 세포, 예를 들어 원핵 생물 세포 또는 진핵 생물 세포, 바람직하게는 포유동물 세포, 예를 들어 사람 세포 내에서 안정화된다. 안정화 효과는 또한 세포 외부, 예를 들어 안정화된 핵산 분자를 포함하는 약학 조성물을 제조하는 방법에 있어서, 예를 들어 완충 용액 중에서 발휘될 수도 있다.

[0059] **형질 감염:** 용어 "형질 감염"은 핵산 분자, 예를 들어 DNA 또는 RNA(예를 들어 mRNA) 분자가 세포, 바람직하게는 진핵 생물 세포에 도입되는 것을 말한다. 본 발명의 내용에 있어서, 용어 "형질 감염"은 핵산 분자를 세포, 바람직하게는 진핵 생물 세포, 예를 들어 포유 동물 세포에 도입하는 방법으로서 당 업자에게 알려진 임의의 방법을 포함한다. 이와 같은 방법들로서는, 예를 들어 전기 천공법, 리포펙션(lipofection), 예를 들어 양이온 지질 및/또는 리포좀 기반 리포펙션, 칼슘 포스페이트 침전법, 나노입자를 기반으로 하는 형질 감염, 바이러스를 기반으로 하는 형질 감염, 또는 양이온 중합체, 예를 들어 DEAE-덱스트란 또는 폴리에틸렌이민 등을 기반으로 하는 형질 감염을 포함한다. 도입은 비 바이러스성인 것이 바람직하다.

[0060] **백신:** 백신은 통상 하나 이상의 항원, 바람직하게는 면역원을 제공하는 예방 물질 또는 치료 물질인 것으로 이해된다. 항원 또는 면역원은 백신화에 적당한 임의의 물질로부터 유래할 수 있다. 예를 들어 항원 또는 면역원은 병원체, 예를 들어 박테리아 또는 바이러스 입자 등으로부터 유래할 수 있거나, 또는 종양이나 암 조직으로부터 유래할 수도 있다. 항원 또는 면역원은 체내 적응성 면역계를 자극하여 적응성 면역 반응을 제공한다.

[0061] **벡터:** 용어 "벡터"란, 핵산 분자, 바람직하게는 인공 핵산 분자를 말한다. 본 발명의 내용에 있어서 벡터는 원하는 핵산 서열, 예를 들어 개방 해독 틀을 포함하는 핵산 서열을 통합(incorporating) 또는 담지(harboring)하는데 적당하다. 이와 같은 벡터로서는 저장 벡터(storage vector), 발현 벡터(expression vector), 클로닝 벡터(cloning vector), 운반 벡터(transfer vector) 등이 있을 수 있다. 저장 벡터는 핵산 분자, 예를 들어 mRNA 분자를 편리하게 저장할 수 있는 벡터이다. 그러므로 상기 저장 벡터는, 예를 들어 원하는 mRNA 서열 또는 이것의 일부에 상응하는 서열, 예를 들어 mRNA의 개방 해독 틀 및 3'-UTR에 상응하는 서열을 포함할 수 있다. 발현 벡터는 발현 생산물, 예를 들어 RNA, 예를 들어 mRNA 또는 펩티드, 폴리펩티드 또는 단백질을 생산하는데 사용될 수 있다. 예를 들어 발현 벡터는 벡터의 서열 구획을 전사하는데 필요한 서열, 예를 들어 프로모터 서열, 예를 들어 RNA 프로모터 서열을 포함할 수 있다. 클로닝 벡터는 통상 핵산 서열들을 벡터에 통합하는데 사용될 수 있는 클로닝 위치를 포함하는 벡터이다. 클로닝 벡터는, 예를 들어 플라스미드 벡터 또는 박테리오파지 벡터일 수 있다. 운반 벡터는 핵산 분자를 세포나 유기체에 운반하는데 적당한 벡터, 예를 들어 바이러스 벡터일 수 있다. 본 발명의 내용에 있어서 벡터는, 예를 들어 RNA 벡터 또는 DNA 벡터일 수 있다. 바람직하게 벡터는 DNA 분

자이다. 바람직하게 본 출원의 관점에서 벡터는 클로닝 위치, 선별 마커, 예를 들어 항생제 내성 인자, 그리고 벡터의 증식에 적당한 서열, 예를 들어 복제 기원을 포함한다. 바람직하게 본 출원의 내용에 있어서 벡터는 플라스미드 벡터이다.

[0062] **비이클:** 비이클은 통상 화합물, 예를 들어 약학적으로 활성인 화합물을 저장, 운반 및/또는 투여하는데 적당한 물질인 것으로 이해된다. 예를 들어 비이클은 생리적으로 허용 가능한 액체로서, 약학적으로 활성인 화합물을 저장, 운반 및/또는 투여하는데 적당한 액체일 수 있다.

[0063] **3'-비번역 영역(3'-UTR):** 3'-UTR은 통상 mRNA의 단백질 암호화 영역(개방 해독 틀(ORF) 또는 암호화 서열(CDS))과 폴리(A) 서열 사이에 위치하는 mRNA의 일부이다. mRNA의 3'-UTR은 아미노산 서열로 번역되지 않는다. 3'-UTR 서열은 보통 유전자 발현 과정에서 각각의 mRNA로 전사되는 유전자에 의해 암호화된다. 게놈 서열은 우선 임의의 인트론들을 포함하는 미성숙 mRNA로 전사된다. 이후, 미성숙 mRNA는 성숙 과정을 통해 성숙한 mRNA로 추가 가공된다. 이와 같은 성숙 과정은, 5'-캡핑 단계, 미성숙 mRNA를 스플라이싱하여 임의의 인트론들을 잘라내는 단계, 그리고 3'-말단의 변형 단계, 예를 들어 미성숙 mRNA의 3'-말단의 폴리아데닐화 단계와, 임의의 엔도뉴클레아제/엑소뉴클레아제 절단 단계 등을 포함한다. 본 발명의 내용에 있어서, 3'-UTR은, 단백질 암호화 영역의 종결 코돈에 대해 3'-말단 쪽, 바람직하게는 단백질 암호화 영역의 종결 코돈에 대해 바로 3'-말단 쪽에 위치하며, 폴리(A) 서열의 5'-말단 쪽, 바람직하게는 폴리(A) 서열에 대해 바로 5'-말단 쪽으로 뻗어있는, 성숙한 mRNA의 서열에 상응한다. 용어 "-에 상응하다"란, 3'-UTR 서열이 RNA 서열, 예를 들어 3'-UTR 서열을 한정하는데 사용되는 mRNA 서열 내 RNA 서열, 또는 이러한 RNA 서열에 상응하는 DNA 서열일 수 있음을 의미한다. 본 발명의 내용에 있어서, 용어 "유전자의 3'-UTR", 예를 들어 "알부민 유전자의 3'-UTR"은, 이 유전자로부터 유래하는 성숙한 mRNA(즉 유전자의 전사 및 미성숙 mRNA의 성숙을 통해 생산된 mRNA)의 3'-UTR에 상응하는 서열이다. 용어 "유전자의 3'-UTR"은, 3'-UTR의 DNA 서열 및 RNA 서열을 포함한다.

[0064] 제1 양태에 있어서, 본 발명은 다음과 같은 요소들을 포함하는 인공 핵산 분자에 관한 것이다:

[0065] a. 개방 해독 틀(ORF) 하나 이상; 및

[0066] b. 알부민 유전자의 3'-UTR 또는 알부민 유전자의 3'-UTR의 변이체로부터 유래하는 핵산 서열을 포함하거나 이 핵산 서열로 이루어진 3'-비번역 영역 요소(3'-UTR 요소) 하나 이상.

[0067] 용어 "3'-UTR 요소"란, 3'-UTR 또는 3'-UTR 변이체로부터 유래하는 핵산 서열을 포함하거나 이 핵산 서열로 이루어진 핵산 서열을 말한다. "3'-UTR 요소"란 바람직하게 인공 핵산 서열, 예를 들어 인공 mRNA의 3'-UTR을 나타내거나 인공 핵산 분자의 3'-UTR을 암호화하는 핵산 서열을 말한다. 그러므로 본 발명의 관점에 있어서, 바람직하게 3'-UTR 요소는 mRNA의 3'-UTR, 바람직하게는 인공 mRNA의 3'-UTR일 수 있거나, 아니면 mRNA의 3'-UTR에 대한 전사 주형일 수 있다. 그러므로 3'-UTR 요소는 바람직하게 mRNA의 3'-UTR, 바람직하게는 인공 mRNA, 예를 들어 유전자 조작된 벡터 구성물의 전사에 의해 생산된 mRNA의 3'-UTR에 상응하는 핵산 서열이다. 바람직하게 본 발명의 관점에 있어서 3'-UTR 요소는 3'-UTR로서 작용을 하거나, 또는 3'-UTR의 작용을 수행하는 뉴클레오티드 서열을 암호화한다.

[0068] 바람직하게 하나 이상의 개방 해독 틀과 하나 이상의 3'-UTR 요소는 이종이다. 이와 같은 내용에 있어서 용어 "이종"이란, 개방 해독 틀과 3'-UTR 요소가 이 조합으로는 자연적으로(자연에서) 발생하지 않는다는 것을 의미한다. 바람직하게 3'-UTR 요소는 개방 해독 틀과는 상이한 유전자로부터 유래한다. 예를 들어 ORF는, 예를 들어 상이한 단백질 또는 동일한 단백질을 암호화하되 상이한 종의 것인, 3'-UTR 요소와 상이한 유전자로부터 유래할 수 있다. 바람직하게 개방 해독 틀은 사람 알부민을 암호화하지 않으며, 바람직하게 개방 해독 틀은 알부민을 암호화하지 않는다. 개방 해독 틀은, 예를 들어 글로빈 단백질(특히 베타-글로빈), 루시페라제 단백질, GFP 단백질 또는 이것들의 변이체, 예를 들어 글로빈 단백질, 루시페라제 단백질 또는 GFP 단백질에 대한 서열 동일성이 70% 이상인 변이체로 이루어진 균으로부터 선택되는 리포터 단백질을 암호화하지 않는 것이 바람직하다. 특히 특정 구체예에서, 개방 해독 틀은, 특히 3'-UTR 요소가 래트 알부민 3'-UTR 또는 이의 변이체로부터 유래하는 경우, 베타-글로빈을 암호화하지 않는 것이 바람직하거나, 또는 더욱 구체적으로 토끼 베타-글로빈 또는 이것들의 변이체를 암호화하지 않는 것이 바람직하다. 더욱이 특정 구체예에서, 본 발명의 인공 핵산 분자는 신호 서열(이는 또한 국소화 신호 또는 표적화 신호라고도 명명됨)을 암호화하지 않는다(다시 말해서 상기 신호 서열을 암호화하는 분절을 포함하지 않는다). 상기와 같은 신호 서열은 통상 암호화된 아미노산 서열의 5'-말단에 제공된다. 특히 본 발명의 인공 핵산은 "신호 아미노산 서열", 구체적으로 암호화된 단백질을 소포체 막에 결합하고 있는 폴리솜(polysome)에 보내고/보내거나, 본 발명의 핵산에 의해 암호화된 관심 단백질을 소포체 막을 통과시켜 이동하게 하는 신호 서열을 (인공적으로나 자연적으로) 포함하는 단백질을 암호화하지 않는다. 특히

본 발명의 핵산 분자의 암호화된 단백질은 알부민 신호 서열을 포함하지 않으며, 더욱 구체적으로는 래트의 알부민 신호 서열을 포함하지 않는다. 특히 암호화 영역이 글로빈, 더욱 구체적으로는 베타-글로빈, 더더욱 구체적으로는 토끼의 베타-글로빈을 암호화하면, 통상적으로 본 발명의 인공 핵산 분자에 의해 암호화된 단백질은 또한 우유 단백질이나 성장 호르몬 신호 서열을 포함하지 않는다. 다른 구체예에서, 만일 암호화 영역이 글로빈 서열, 더욱 구체적으로는 베타-글로빈 또는 이의 변이체를 암호화하면, 본 발명의 인공 핵산 서열은 글로빈 5'-UTR 서열, 구체적으로 베타-글로빈 유전자 유래 5'-UTR 서열, 더욱 구체적으로는 토끼의 글로빈 5'-UTR 서열을 포함하지 않는다. 만일 인공 핵산 분자가 5'-UTR 서열을 포함하면, 이 5'-UTR 서열은, 알부민 유전자 유래 5'-UTR 서열이 아닌 것, 특히 래트 알부민 유전자 유래 5'-UTR 서열이 아닌 것으로 선택될 수 있다. 본 발명의 인공 핵산 분자가 3'-UTR을 하나 이상 포함하는 경우, 다른 3'-UTR(들)은 글로빈 3'-UTR 및 c-myc 3'-UTR로 이루어진 군으로부터 선택되지 않는 것이 바람직하다. 다른 구체예에서, 특히 암호화 영역이 글로빈, 더욱 구체적으로는 베타-글로빈을 암호화하면, 본 발명의 핵산 분자의 3'-UTR은 래트 알부민 3'-UTR에 상응하지 않는다.

[0069] 더욱이 개방 해독 틀은 사람 인자 IX 또는 이의 변이체, 예를 들어 사람 인자 IX에 대한 서열 동일성이 70% 이상인 변이체를 암호화하지 않는 것이 특히 바람직하다. 특히 본 발명의 핵산 분자의 암호화 영역이 전술된 바와 같이 사람 인자 IX 또는 이의 변이체를 암호화하면, 본 발명의 핵산 분자는 알부민 프로모터, 구체적으로 점 돌연 변이가 일어난 알부민 프로모터, 더욱 구체적으로 G52A 점 돌연 변이가 일어난 알부민 프로모터를 포함하지 않는다.

[0070] 다른 구체예에서, 특히 암호화 영역이 내성 유전자, 구체적으로 네오마이신 내성 유전자를 암호화하면, 본 발명의 인공 핵산 분자는 트랜스포손 요소(transposon element), 예를 들어 트랜스포손 플라스미드와 상응하지 않거나, 또는 트랜스포손(구체적으로 Tn5 트랜스포손 또는 TN5 모자이크 요소)를 포함하지 않는다. 본 발명의 핵산 분자는 기능상 트랜스포사제(transposase)와 상호 작용할 수 없으며, 특히 이러한 경우 Tn5 트랜스포사제와 상호 작용할 수 없다. 기능면에서 말하면, (본 발명의 인공 핵산은 트랜스포손을 포함하지 않으므로) 본 발명의 핵산 분자는 통상 핵산과, 트랜스포손이 어떠한 것이든 간에 이 트랜스포손에 특이적인 트랜스포사제 사이에 복합체를 형성하지 않을 것이다. 특히 본 발명의 핵산 분자가 기능상 트랜스포사제와 상호 작용을 한다면, 본 발명의 핵산 분자의 암호화 영역(ORF)은 siRNA를 암호화하지 않는다.

[0071] 이와 같은 내용에 있어서, 특정 구체예에서, 특히 개방 해독 틀이 사람 인자 IX 또는 이의 변이체를 암호화하는 경우, 개방 해독 틀은 인트론을 포함하지 않는 것이 특히 바람직하다.

[0072] 뿐만 아니라 이와 같은 내용에 있어서, 특정 구체예에서, 특히 3'-UTR 요소가 사람 알부민 3'-UTR 또는 이의 변이체로부터 유래하는 경우, 개방 해독 틀은 사람 인자 IX 또는 이의 변이체를 암호화하지 않는 것이 바람직하다.

[0073] 본 발명의 인공 핵산 분자는 (본 발명의 특정 구체예로서) 발현 카세트가 아니다. 그러므로 본 발명의 핵산 분자는, 예를 들어 3'-프로모터 또는 프로모터의 3'-말단을 포함하지 않는다. 이러한 구체예의 관점에서, 본 발명의 핵산 분자는 신호 서열, 바람직하게는 3'-프로모터를 포함하지 않으므로, "분비 카세트"인 것도 아니다. 특히 본 발명의 핵산 분자는 ORF 및 3'-UTR 영역을 포함하고, 임의로는 5'-UTR 영역도 포함하는 단일 핵산 분자 하나로 이루어져 있다. 그러므로 본 발명의 핵산 분자는, 독립된 유전 요소 하나 이상으로 이루어진 분비 카세트에 상응하지 않으며, 특히, 예를 들어 키트의 일부로서 독립적으로 제공되는, 암호화 영역의 상류에 있는 영역을 나타내는 제1 유전 요소(ORF)와, ORF의 하류에 있는 영역을 나타내는 독립된 제2 유전 요소에 상응하지 않는다.

[0074] 바람직하게 3'-UTR 요소 하나 이상은 ORF에 기능상 결합한다. 이는, 바람직하게 3'-UTR 요소는 ORF와 결합할 때, 특정 기능, 예를 들어 ORF 발현에 대한 안정화 기능 또는 인공 핵산 분자에 대한 안정화 기능을 발휘할 수 있다는 것을 의미한다. 바람직하게 ORF 및 3'-UTR 요소는 5'→3' 방향으로 결합한다. 그러므로 본 발명의 인공 핵산 분자는 구조 5'-ORF-(임의의) 링커-3'-UTR 요소-3'[여기서, 링커는 존재할 수 있거나 존재하지 않을 수 있음]을 포함하는 것이 바람직하다. 예를 들어 링커는, 예를 들어 제한 효소 인지 위치(제한 위치) 하나 이상을 포함하거나 이 위치 하나 이상으로 이루어진 뉴클레오티드 하나 이상, 예를 들어 1~50개 또는 1~20개 뉴클레오티드로 이루어진 구획일 수 있다.

[0075] 바람직하게 하나 이상의 3'-UTR 요소는, 척추 동물 알부민 유전자의 3'-UTR 또는 이의 변이체, 바람직하게는 포유동물 알부민 유전자의 3'-UTR, 예를 들어 마우스 알부민 유전자, 올리브 개코 원숭이의 알부민 유전자 또는 사람 알부민 유전자의 3'-UTR, 또는 이것들의 변이체로부터 유래하는 핵산 서열을 포함하거나 이 핵산 서열로 이루어져 있다. 더욱 바람직하게 하나 이상의 3'-UTR 요소는 영장류 알부민 유전자, 특히 사람 알부민 유전자

또는 올리브 개코 원숭이의 알부민 유전자의 3'-UTR, 또는 이것들의 변이체, 더욱 바람직하게는 유전자 은행 수탁 번호 NM\_000477.5에 의한 사람 알부민 유전자의 3'-UTR 또는 이의 변이체로부터 유래하는 핵산 서열을 포함하거나 이 핵산 서열로 이루어져 있다. 바람직한 구체예에서, 3'-UTR 요소는 손톱 개구리속의 알부민 유전자의 3'-UTR로부터 유래하지 않는다. 바람직하게 3'-UTR 요소는 손톱 개구리속 알부민 유전자 유래 3'-UTR의 폴리(A) 제한 요소 B(PLEB)를 포함하지 않는다. 바람직하게 3'-UTR 요소는 손톱 개구리속 알부민 유전자 유래 3'-UTR의 PLEB로 이루어져 있지 않다.

[0076] 용어 "[...] 알부민 유전자의 3'-UTR로부터 유래하는 핵산 서열"이란, 바람직하게 알부민 유전자의 3'-UTR 서열, 이의 단편 또는 일부를 기반으로 하는 핵산 서열을 말한다. 상기 용어는 알부민 유전자의 전체 3'-UTR 서열, 즉 알부민 유전자의 전장 3'-UTR 서열에 상응하는 서열과, 알부민 유전자의 3'-UTR 서열의 단편에 상응하는 서열을 포함한다. 바람직하게 알부민 유전자의 3'-UTR 단편은 알부민 유전자의 전장 3'-UTR 내 뉴클레오티드 연속 구획에 상응하는 뉴클레오티드 연속 구획으로서, 알부민 유전자의 전장 3'-UTR의 20% 이상, 바람직하게는 30% 이상, 더욱 바람직하게는 40% 이상, 더욱 바람직하게는 50% 이상, 더더욱 바람직하게는 60% 이상, 더더욱 바람직하게는 70% 이상, 더더욱 바람직하게는 80% 이상, 그리고 가장 바람직하게는 90% 이상에 해당하는 뉴클레오티드 연속 구획으로 이루어져 있다. 이러한 단편은 본 발명의 관점에서 본원에 기술된 바와 같이 기능성 단편인 것이 바람직하다. 용어 "알부민 유전자의 3'-UTR"이란, 바람직하게 자연 발생 알부민 유전자의 3'-UTR을 말한다.

[0077] 알부민 유전자의 3'-UTR에 관한 내용에 있어서 용어 "알부민 유전자의 3'-UTR의 변이체" 및 "이의 변이체"란, 자연 발생 알부민 유전자의 3'-UTR의 변이체, 바람직하게는 척추 동물 알부민 유전자의 3'-UTR의 변이체, 더욱 바람직하게는 포유동물 알부민 유전자의 3'-UTR, 예를 들어 마우스 알부민 유전자의 3'-UTR의 변이체, 더더욱 바람직하게는 영장류 알부민 유전자, 특히 전술된 바와 같은 사람 알부민 유전자 또는 올리브 개코 원숭이의 알부민 유전자의 3'-UTR의 변이체를 말한다. 이와 같은 변이체는 알부민 유전자의 변형된 3'-UTR일 수 있다. 예를 들어 변이체 3'-UTR은, 이 변이체가 유래하는 자연 발생 3'-UTR과 비교되게, 뉴클레오티드 결실, 삽입, 부가 및/또는 치환을 하나 이상 가질 수 있다. 바람직하게 알부민 유전자의 3'-UTR의 변이체는, 이 변이체가 유래하는 자연 발생 3'-UTR에 대하여 40% 이상, 바람직하게는 50% 이상, 더욱 바람직하게는 60% 이상, 더욱 바람직하게는 70% 이상, 더더욱 바람직하게는 80% 이상, 더더욱 바람직하게는 90% 이상, 가장 바람직하게는 95% 이상 동일하다. 바람직하게 변이체는 본원에 기술된 바와 같은 기능성 변이체이다.

[0078] 용어 "알부민 유전자의 3'-UTR의 변이체로부터 유래하는 핵산 서열"이란, 전술된 바와 같이 바람직하게는 알부민 유전자의 3'-UTR 서열의 변이체, 이의 단편 또는 일부를 기반으로 하는 핵산 서열을 말한다. 상기 용어는, 알부민 유전자의 3'-UTR의 변이체의 전체 서열(즉 알부민 유전자의 전장 변이체 3'-UTR 서열)에 상응하는 서열과, 알부민 유전자의 변이체 3'-UTR 서열의 단편에 상응하는 서열을 포함한다. 바람직하게 알부민 유전자의 3'-UTR의 변이체의 단편은, 알부민 유전자의 3'-UTR의 전장 변이체 내 뉴클레오티드 연속 구획에 상응하는 뉴클레오티드 연속 구획으로서, 알부민 유전자의 3'-UTR의 전장 변이체의 20% 이상, 바람직하게는 30% 이상, 더욱 바람직하게는 40% 이상, 더욱 바람직하게는 50% 이상, 더더욱 바람직하게는 60% 이상, 더더욱 바람직하게는 70% 이상, 더더욱 바람직하게는 80% 이상, 그리고 가장 바람직하게는 90% 이상에 해당하는 뉴클레오티드 연속 구획으로 이루어져 있다. 본 발명의 관점에서 이러한 변이체 단편은 본원에 기술된 바와 같은 변이체의 기능성 단편인 것이 바람직하다.

[0079] 본 발명의 내용에 있어서 용어 "기능성 변이체", "기능성 단편" 및 "변이체의 기능성 단편"("기능성 변이체 단편"이라고도 칭하여짐)이란, 알부민 유전자의 3'-UTR의 단편, 3'-UTR의 변이체, 또는 3'-UTR의 변이체의 단편으로서, 상기 변이체, 단편 또는 변이체의 단편이 유래하는 알부민 유전자의 자연 발생 3'-UTR의 기능을 하나 이상, 바람직하게는 하나보다 많이 수행하는 변이체, 단편 또는 변이체의 단편을 의미한다. 이러한 기능은, 바람직하게 포유동물 세포, 예를 들어 사람 세포 내에서, 예를 들어 mRNA를 안정화하는 것, 및/또는 mRNA로부터의 단백질 생산을 안정화하고/하거나 연장하는 것, 및/또는 mRNA로부터의 단백질 발현량 또는 단백질 총 생산량을 증가시키는 것일 수 있다. 본 발명의 내용에 있어서 변이체, 단편 및 변이체의 단편은, 바람직하게 포유동물 세포, 예를 들어 사람 세포 내에서 mRNA를, 기준 3'-UTR을 포함하거나 3'-UTR이 결실된 mRNA와 비교되게 안정화하는 기능, 및/또는 바람직하게 포유동물 세포, 예를 들어 사람 세포 내에서 mRNA로부터의 단백질 생산을, 기준 3'-UTR을 포함하거나 3'-UTR이 결실된 mRNA로부터의 단백질 생산과 비교되게 안정화 및/또는 연장하는 기능, 및/또는 바람직하게 포유동물 세포, 예를 들어 사람 세포 내에서 mRNA로부터의 단백질 생산량을, 기준 3'-UTR을 포함하거나 3'-UTR이 결실된 mRNA로부터의 단백질 생산량과 비교되게 증가시키는 기능을 수행하는 것이 특히 바람직하다. 기준 3'-UTR은, 예를 들어 ORF와 함께 자연 발생하는 3'-UTR일 수 있다. 더욱이 알부민 유전자의 3'-

UTR의 기능성 변이체, 기능성 단편 또는 기능성 변이체 단편은, 이와 같은 변이체, 단편 또는 변이체 단편이 유래하는 야생형 3'-UTR과 비교되게, 상기 3'-UTR의 변이체, 단편 또는 변이체 단편을 포함하는 mRNA의 번역 효율을 실질적으로 감소시키는 효과를 나타내지 않는다. 본 발명의 내용에 있어서 알부민 유전자의 3'-UTR의 "기능성 단편", "기능성 변이체" 또는 "변이체의 기능성 단편"의 특히 바람직한 기능은, 전술된 바와 같은 기능성 단편, 기능성 변이체 또는 변이체의 기능성 단편을 운반하는 mRNA를 발현시킴으로써 단백질 생산을 안정화 및/또는 연장하는 것이다.

[0080] 바람직하게 상기 기능성 변이체, 상기 기능성 단편 또는 상기 기능성 변이체 단편에 의해 발휘되는 기능들 하나 이상의 효율, 예를 들어 mRNA 및/또는 단백질 생산 안정화 효율 및/또는 단백질 생산량 증가 효율은, 상기 변이체, 단편 또는 변이체 단편이 유래하는 알부민 유전자의 자연 발생 3'-UTR에 의해 발휘되는 mRNA 및/또는 단백질 생산 안정화 효율 및/또는 단백질 생산량 증가 효율의 40% 이상, 더욱 바람직하게는 50% 이상, 더욱 바람직하게는 60% 이상, 더더욱 바람직하게는 70% 이상, 더더욱 바람직하게는 80% 이상, 가장 바람직하게는 90% 이상에 해당한다.

[0081] 본 발명의 내용에 있어서, 알부민 유전자의 3'-UTR의 단편 또는 알부민 유전자의 3'-UTR의 변이체의 단편의 길이는 바람직하게 약 50 뉴클레오티드 이상, 바람직하게는 약 75 뉴클레오티드 이상, 더욱 바람직하게는 약 100 뉴클레오티드 이상, 더더욱 바람직하게는 약 125 뉴클레오티드 이상, 가장 바람직하게는 약 150 뉴클레오티드 이상이다. 바람직하게 알부민 유전자의 3'-UTR의 단편 또는 알부민 유전자의 3'-UTR의 변이체의 단편은 전술된 바와 같이 기능성 단편이다.

[0082] 바람직하게 본 발명에 의한 인공 핵산 분자의 3'-UTR 요소 하나 이상은, 알부민 유전자의 "기능성 단편", "기능성 변이체" 또는 "변이체의 기능성 단편"을 포함하거나 이것들로 이루어져 있다.

[0083] 바람직하게 본 발명에 의한 인공 핵산 분자의 3'-UTR 요소 하나 이상은, 3'-UTR이 결실되었거나 기준 3'-UTR, 예를 들어 ORF와 함께 자연 발생하는 3'-UTR을 포함하는 각각의 mRNA(기준 mRNA)와 비교되게, 본 발명에 의한 인공 핵산 분자의 안정성, 예를 들어 본 발명에 의한 mRNA의 안정성을 증가시킨다. 바람직하게 본 발명에 의한 인공 핵산 분자의 3'-UTR 요소 하나 이상은 본 발명에 의한 인공 핵산 분자, 예를 들어 본 발명에 의한 mRNA로부터의 단백질 생산의 안정성을, 3'-UTR이 결실되었거나 기준 3'-UTR, 예를 들어 ORF와 함께 자연 발생하는 3'-UTR을 포함하는 각각의 mRNA로부터의 단백질 생산의 안정성과 비교되게 증가시킨다. 바람직하게 본 발명에 의한 인공 핵산 분자의 3'-UTR 요소 하나 이상은, 본 발명에 의한 인공 핵산 분자, 예를 들어 본 발명에 의한 mRNA로부터의 단백질 생산을, 3'-UTR이 결실되었거나 기준 3'-UTR, 예를 들어 ORF와 함께 자연 발생하는 3'-UTR을 포함하는 각각의 mRNA로부터의 단백질 생산과 비교되게 연장한다. 바람직하게 본 발명에 의한 인공 핵산 분자의 3'-UTR 요소 하나 이상은, 본 발명에 의한 인공 핵산 분자, 예를 들어 본 발명에 의한 mRNA로부터의 단백질 발현량 및/또는 단백질 총 생산량을, 3'-UTR이 결실되었거나 기준 3'-UTR, 예를 들어 ORF와 함께 자연 발생하는 3'-UTR을 포함하는 각각의 mRNA로부터의 단백질 발현량 및/또는 단백질 총 생산량과 비교되게 증가시킨다. 바람직하게 본 발명에 의한 인공 핵산 분자의 3'-UTR 요소 하나 이상은 mRNA의 번역 효율에, 3'-UTR이 결실되었거나 기준 3'-UTR, 예를 들어 ORF와 함께 자연 발생하는 3'-UTR을 포함하는 각각의 mRNA의 번역 효율과 비교되게 약 영향을 미치지 않는다. 이와 같은 내용에 있어서 용어 "각각의 mRNA"란, 기준 mRNA가 상이한 3'-UTR들과는 별개로 본 발명의 3'-UTR 요소를 포함하는 mRNA와 거의 비슷하거나, 바람직하게는 동일하다는 것을 의미한다.

[0084] 용어, 인공 핵산 분자, 예를 들어 인공 mRNA로부터 "단백질 생산을 안정화 및/또는 연장함"이라 함은, 바람직하게는 포유동물 발현계, 예를 들어 HeLa 세포 또는 HDF 세포 내에서, 인공 핵산 분자, 예를 들어 인공 mRNA로부터의 단백질 생산이, 예를 들어 기준 3'-UTR을 포함하거나 3'-UTR이 결실된 기준 핵산 분자, 예를 들어 기준 mRNA로부터의 단백질 생산과 비교되게 안정화 및/또는 연장됨을 의미한다. 그러므로 인공 핵산 분자, 예를 들어 인공 mRNA로부터 생산되는 단백질은, 기준 핵산 분자로부터 생산되는 단백질이 관찰될 수 있는 기간보다 더욱 오랜 기간 동안 관찰될 수 있다. 다시 말해서, 인공 핵산 분자, 예를 들어 인공 mRNA로부터 생산되는 단백질의 경시적으로 측정되는 양은, 기준 핵산 분자, 예를 들어 기준 mRNA로부터 생산되는 단백질의 경시적으로 측정되는 양과 비교되게, 더욱 늦은 시점에서 역치 값을 밀돌게 된다. 이와 같은 역치 값은, 예를 들어 발현의 초기 시점, 예를 들어 발현 개시 후(예를 들어 핵산 분자의 형질 감염 후) 1, 2, 3, 4, 5 또는 6시간 경과시에 측정되는 단백질의 양일 수 있다(도 17).

[0085] 포유 동물 발현계, 예를 들어 포유동물 세포, 예를 들어 HeLa 세포 또는 HDF 세포 내에서, 예를 들어 인공 핵산 분자, 예를 들어 인공 mRNA로부터의 단백질 생산은, 기준 핵산 분자, 예를 들어 기준 mRNA로부터의 단백질 생산과 비교되게, 발현의 초기 시점, 예를 들어 발현 개시 후, 예를 들어 핵산 분자의 형질 감염 후 1, 2, 3, 4, 5

또는 6시간 경과시에 관찰되는 양 이상의 양으로, 약 5시간 이상, 바람직하게는 약 10시간 이상, 더욱 바람직하게는 약 24시간 이상 연장된다. 그러므로, 본 발명에 의한 인공 핵산 분자는, 3'-UTR이 결실되었거나 기준 3'-UTR을 포함하는 기준 핵산 분자와 비교되게 단백질 생산을, 발현의 초기 시점, 예를 들어 발현 개시 후, 예를 들어 형질 감염 후 1, 2, 3, 4, 5 또는 6시간 경과시에 관찰되는 양 이상의 양으로, 약 5시간 이상, 바람직하게는 약 10시간 이상, 더욱 바람직하게는 약 24시간 이상 연장하는 것이 바람직하다.

[0086] 바람직한 구체예에서, 본 발명에 의한 인공 핵산 분자로부터의 단백질 생산은, 3'-UTR이 결실되었거나 기준 3'-UTR을 포함하는 기준 핵산 분자로부터의 단백질 생산과 비교되게, 1.5배 이상, 바람직하게는 2배 이상, 더욱 바람직하게는 2.5배 이상 연장된다.

[0087] 이와 같은 단백질 생산 연장 효과는, (i) 바람직하게 포유동물 발현계, 예를 들어 HeLa 세포 또는 HDF 세포 내에서, 예를 들어 리포터 단백질, 예를 들어 루시퍼라제를 암호화하는 ORF의 발현에 의해 경시적으로 생산되는 단백질 양을 측정하고, (ii) 단백질 양이, 예를 들어 발현 개시 후 1, 2, 3, 4, 5 또는 6시간 경과시, 예를 들어 인공 핵산 분자의 형질 감염 후 1, 2, 3, 4, 5 또는 6시간 경과시에 관찰되는 단백질의 양을 밀도는 시점을 측정하며, (iii) 단백질 양이, 발현 개시 후 1, 2, 3, 4, 5 또는 6시간 경과시에 관찰되는 단백질의 양을 밀도는 시점을, 3'-UTR이 결실되었거나 기준 3'-UTR을 포함하는 핵산 분자에 대해 측정된 상기 시점과 비교함으로써 확인될 수 있다(도 17).

[0088] 바람직하게 이와 같은 단백질 생산에 대한 안정화 및/또는 연장 효과는, 예를 들어 48시간 또는 72시간에 걸친 기간 이내에 본 발명에 의한 인공 핵산 분자로부터 생산되는 단백질의 총량이, 3'-UTR이 결실되었거나 기준 3'-UTR, 예를 들어 인공 핵산 분자의 ORF와 함께 자연 발생하는 3'-UTR을 포함하는 기준 핵산 분자로부터 생산되는 단백질 양의 이상이 될 때 달성된다. 그러므로 본 발명은 포유동물 발현계, 예를 들어 포유동물 세포, 예를 들어 상기 특정된 바와 같이 HeLa 세포 또는 HDF 세포 내에서 단백질 생산을 연장 및/또는 안정화할 수 있는 인공 핵산 분자를 제공하는데, 이 경우, 예를 들어 48시간 또는 72시간에 걸친 기간 이내에 상기 인공 핵산 분자로부터 생산되는 단백질의 총량은, 예를 들어 상기 기간 이내에 3'-UTR이 결실되었거나 기준 3'-UTR, 예를 들어 인공 핵산 분자의 ORF와 함께 자연 발생하는 3'-UTR을 포함하는 기준 핵산 분자로부터 생산되는 단백질의 총량 이상이다.

[0089] 그러므로 "안정화된 단백질 발현"이란, 바람직하게 소정의 기간, 예를 들어 24시간, 더욱 바람직하게는 48시간, 더더욱 바람직하게 72시간에 걸친 본 발명에 의한 인공 핵산 분자로부터의 단백질 생산이, 기준 핵산 분자, 예를 들어 기준 3'-UTR을 포함하거나 3'-UTR이 결실된 mRNA로부터의 단백질 생산과 비교되었을 때 더욱 균일함을 의미한다. 따라서, 예를 들어 포유동물 계 내 본 발명에 의한 3'-UTR 요소를 포함하는 인공 핵산 분자, 예를 들어 본 발명에 의한 mRNA로부터의 단백질 생산 수준은, 전술된 바와 같은 기준 핵산 분자, 예를 들어 기준 mRNA에 대해서 관찰되는 단백질 생산 수준 정도로 떨어지지 않는 것이 바람직하다. 예를 들어 발현 개시 후 6시간 경과시, 예를 들어 본 발명에 의한 인공 핵산 분자의 세포, 예를 들어 포유동물 세포로의 형질 감염 후 6시간 경과시에 관찰되는 (ORF에 의해 암호화된) 단백질의 양은, 발현 개시 후 48시간 경과시, 예를 들어 형질 감염 후 48시간 경과시 관찰되는 단백질의 양과 거의 동일할 수 있다. 그러므로 발현 개시 후 48시간 경과시, 예를 들어 형질 감염 후 48시간 경과시 관찰되는, ORF에 의해 암호화되는 단백질, 예를 들어 리포터 단백질, 예를 들어 루시퍼라제의 양 대 발현 개시 후 6시간 경과시, 예를 들어 형질 감염 후 6시간 경과시 관찰되는 단백질의 양의 비율은, 바람직하게 약 0.4 이상, 더욱 바람직하게 약 0.5 이상, 더욱 바람직하게 약 0.6 이상, 더더욱 바람직하게 약 0.7 이상이다. 바람직하게 본 발명에 의한 핵산 분자에 있어서 상기 비율은 약 0.4~약 4, 바람직하게 약 0.65~약 3, 더욱 바람직하게 약 0.7~약 2이다. 각각의 기준 핵산 분자, 예를 들어 기준 3'-UTR을 포함하거나 3'-UTR이 결실된 mRNA에 있어서, 상기 비율은 약 0.05~약 0.3일 수 있다.

[0090] 그러므로 본 발명은 전술된 바와 같이 ORF와 3'-UTR 요소를 포함하는 인공 핵산 분자를 제공하는데, 여기서, 바람직하게 포유동물 발현계, 예를 들어 포유동물 세포, 예를 들어 HeLa 세포 내 발현 개시 후 48시간 경과시에 관찰되는 (리포터) 단백질의 양, 예를 들어 루시퍼라제의 양 대 발현 개시 후 6시간 경과시에 관찰되는 (리포터) 단백질의 양의 비율은 바람직하게 약 0.4 이상, 더욱 바람직하게 약 0.5 이상, 더욱 바람직하게 약 0.6 이상, 더더욱 바람직하게 약 0.7 이상, 예를 들어 약 0.4~약 4, 바람직하게 약 0.65~약 3, 더욱 바람직하게 약 0.7~2이고, 이때, 예를 들어 48시간의 기간 이내에 상기 인공 핵산 분자로부터 생산되는 단백질의 총량은, 예를 들어 상기 기간 이내에 3'-UTR이 결실되었거나 기준 3'-UTR, 예를 들어 인공 핵산 분자의 ORF와 함께 자연 발생하는 기준 3'-UTR을 포함하는 기준 핵산 분자로부터 생산되는 단백질의 총량 이상인 것이 바람직하다. 바람직한 구체예에서, 본 발명은 전술된 바와 같은 ORF 및 3'-UTR 요소를 포함하는 인공 핵산 분자를 제공하는데, 여기서, 포유동물 발현계, 예를 들어 포유동물 세포, 예를 들어 HeLa 세포 내 발현 개시 후 72시간 경과시 관찰

되는 상기 (리포터) 단백질의 양, 예를 들어 루시퍼라제의 양 대 발현 개시 후 6시간 경과시 관찰되는 상기 (리포터) 단백질의 양의 비율은, 바람직하게 약 0.4 이상, 더욱 바람직하게 약 0.5 이상, 더욱 바람직하게 약 0.6 이상, 더더욱 바람직하게 약 0.7 이상, 예를 들어 약 0.4~1.5, 바람직하게 약 0.65~약 1.15, 더욱 바람직하게 약 0.7~약1.0인 것이 바람직하고, 여기서, 예를 들어 72시간의 기간 이내에 상기 인공 핵산 분자로부터 생산되는 단백질의 총량은, 예를 들어 상기와 같은 기간 이내에 3'-UTR이 결실되었거나 기준 3'-UTR, 예를 들어 인공 핵산 분자의 ORF와 함께 자연 발생하는 3'-UTR을 포함하는 기준 핵산 분자로부터 생산되는 단백질의 총량 이상인 것이 바람직하다.

[0091] 본 발명의 내용에 있어서 "증가된 단백질 발현량"이란, 발현 개시 후 어느 특정 시점에서의 단백질 발현량이, 기준 분자의 경우와 비교되게 증가하였음을 의미한다. 그러므로 본 발명에 의한 인공 핵산 분자의 발현 개시 후, 예를 들어 형질 감염 후, 예를 들어 본 발명에 의한 mRNA의 형질 감염 후 임의의 시점, 예를 들어 형질 감염 후 48시간 또는 72시간 경과시에 관찰되는 단백질 수준은, 기준 핵산 분자, 예를 들어 기준 3'-UTR을 포함하거나 3'-UTR이 결실된 기준 mRNA의 발현 개시 후, 예를 들어 형질 감염 후 상기와 동일한 시점에 관찰되는 단백질 수준보다 더 높은 것이 바람직하다.

[0092] 인공 핵산 분자로부터의 "증가한 단백질 총 생산량"이란, 일정 기간에 걸친 단백질 생산량이 증가하였음을 말하는데, 여기서, 상기 단백질은, 바람직하게 포유동물 발현계, 예를 들어 포유동물 세포, 예를 들어 HeLa 세포 또는 HDF 세포 내에서 인공 핵산 분자로부터 생산된다. 그러므로 "단백질 총 생산량"이란, 경시적 단백질 생산량을 나타내는 곡선 아래 면적(AUC)을 말하는 것이 바람직하다.

[0093] 상기 인공 핵산 분자의 안정성 증가, 상기 단백질 생산량 증가, 상기 단백질 생산 연장 및/또는 상기 단백질 발현량의 증가 및/또는 단백질 총 생산량은, 3'-UTR이 결실된 각각의 기준 핵산 분자, 예를 들어 3'-UTR이 결실된 mRNA, 또는 기준 3'-UTR, 예를 들어 전술된 바와 같이 ORF와 함께 자연 발생하는 3'-UTR을 포함하는 기준 핵산 분자의 경우와 비교됨으로써 분석되는 것이 바람직하다.

[0094] 알부민 유전자의 3'-UTR의 변이체, 단편 및/또는 변이체 단편의 mRNA 및/또는 단백질 생산 안정화 효과 및 효율 및/또는 단백질 생산량 증가 효과 및 효율과, 본 발명에 의한 인공 핵산 분자의 3'-UTR 요소 하나 이상의 mRNA 및/또는 단백질 생산 안정화 효과 및 효율, 및/또는 단백질 생산량 증가 효과 및 효율은, 이와 같은 효과 및 효율의 측정용으로서 적당하다고 당 업자에게 알려진 임의의 방법에 의해 분석될 수 있다. 예를 들어 리포터 단백질, 예를 들어 루시퍼라제를 암호화하는 암호화 서열을 포함하고, 3'-UTR, 즉 자연 발생 알부민 유전자로부터 유래하는 3'-UTR, 기준 유전자로부터 유래하는 3'-UTR(즉 기준 3'-UTR, 예를 들어 ORF와 함께 자연 발생하는 3'-UTR), 알부민 유전자의 3'-UTR의 변이체, 자연 발생 알부민 유전자의 단편, 또는 알부민 유전자의 3'-UTR의 변이체의 단편은 3'-UTR로서 포함되지 않는 인공 mRNA 분자가 생산될 수 있다. 이와 같은 mRNA는, 예를 들어 T7 프로모터와 각각의 mRNA 서열을 암호화하는 서열을 포함하는 각각의 벡터, 예를 들어 플라스미드 벡터를 시험관 내 전사함으로써 생산될 수 있다. 생산된 mRNA 분자는 mRNA를 형질 감염하기 적당한 임의의 형질 감염 방법을 통해 세포에 형질 감염될 수 있는데, 예를 들어 생산된 mRNA 분자는 포유동물 세포, 예를 들어 HeLa 세포에 전기 천공에 의해 도입될 수 있고, 샘플은 형질 감염 후 임의의 시점, 예를 들어 형질 감염 후 6시간, 24시간, 48시간 및 72시간 경과시에 분석될 수 있다. 상기 샘플은 당 업자에게 널리 알려진 방법들을 통하여 mRNA의 양 및/또는 단백질의 양에 대해 분석될 수 있다. 예를 들어 샘플 채취 시점에 세포 내에 존재하는 리포터 mRNA의 양은 정량적 PCR 방법에 의해 측정될 수 있다. 각각의 mRNA에 의해 암호화되는 리포터 단백질의 양은 사용된 리포터 단백질에 따라서, 예를 들어 ELISA 분석법 또는 리포터 분석법, 예를 들어 루시퍼라제 분석법을 통해 측정될 수 있다. 단백질 발현의 안정화 효과 및/또는 단백질 발현의 연장 효과는, 예를 들어 형질 감염 후 48시간 경과시에 관찰되는 단백질 수준과 형질 감염 후 6시간 경과시에 관찰되는 단백질 수준의 비율을 측정함으로써 분석될 수 있다. 상기 비율이 1에 가까울수록 상기 시기 이내에서의 단백질 발현은 더욱 안정한 것이다. 이와 같은 측정은 물론 72시간 이상 경과된 후에도 수행될 수 있으며, 형질 감염 후 72시간 경과시에 관찰되는 단백질 수준과 형질 감염 후 6시간 경과시에 관찰되는 단백질 수준의 비율은 단백질 발현의 안정성을 분석하기 위해 측정될 수 있다.

[0095] 바람직하게 본 발명에 의한 인공 핵산 분자의 3'-UTR 요소 하나 이상은, 알부민 유전자의 3'-UTR의 핵산 서열, 예를 들어 이하에 보인 서열 번호 1, 서열 번호 2, 서열 번호 32, 서열 번호 33, 서열 번호 34 또는 서열 번호 35에 의한 핵산 서열에 대한 동일성이 약 40% 이상, 바람직하게는 약 50% 이상, 바람직하게는 약 60% 이상, 바람직하게는 약 70% 이상, 더욱 바람직하게는 약 80% 이상, 더욱 바람직하게는 약 90% 이상, 더더욱 바람직하게는 약 95% 이상, 더더욱 바람직하게는 약 99% 이상, 가장 바람직하게는 100%인 핵산 서열을 포함하거나 이 핵산 서열로 이루어져 있는데, 이 경우, (예를 들어 동일성이 40% 이상인) 서열의 변이체는 전술된 기능성 변이체인

것이 바람직하다:

CAUCACAUUU AAAAGCAUCU CAGCCUACCA UGAGAAUAAG AGAAAGAAA UGAAGAUCAA  
 AAGCUUAUUC AUCUGUUUUU CUUUUUCGUU GGUGUAAAGC CAACACCCUG UCUAAAAAC  
 AUAAAAUUCU UUAUCAUUU UGCCUCUUU CUCUGUGCUU CAAUUAAUAA AAAAUGGAAA  
 GAAUCU (서열 번호 2).

AAACATCACA ATTAAGAACA TCTCAGCCTA CCATGAGAAC AAGAGAAATA AAATGAAGAT  
CAAAAGCTTA TTCATCTGTT TTTCTTTTTT ATTGGTATAA AGCCAACACC CTGTCTAAAA  
AACTATAAAT TTCTTTAATC ATTTTGCCTC TTTTCTCTGT GCTTCAATTA ATAAAAAATG  
GAAAGAATCT AGATCTAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA  
 AAAA (서열 번호 32)

AAACAUCACA AUUAAGAACA UCUCAGCCUA CCAUGAGAAC AAGAGAAUA AAAUGAAGAU  
CAAAAGCUUA UUCAUCUGUU UUUCUUUUUC AUUGGUUUAU AGCCAACACC CUGUCUAAAA  
AACUAUAAAU UUCUUUAAUC AUUUUGCCUC UUUUCUCUGU GCUUCAAUUA AUAAAAAUG  
GAAAGAAUCU AGAUCUAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA  
 AAAA (서열 번호 33)

[0096]

ACACATCACA ACCACAACCT TCTCAGGCTA CCCTGAGAAA AAAAGACATG AAGACTCAGG  
ACTCATCTTT TCTGTTGGTG TAAAATCAAC ACCCTAAGGA ACACAAATTT CTTTAAACAT  
TTGACTTCTT GTCTCTGTGC TGCAATTAAT AAAAAATGGA AAGAATCTAC AGATCTAAAA  
 AAAA (서열 번호 34)

ACACAUCACA ACCACAACCU UCUCAGGCUA CCCUGAGAAA AAAAGACAUG AAGACUCAGG  
ACUCAUCUUU UCUGUUGGUG UAAAAUCAAC ACCCUAAGGA ACACAAAUUU CUUUAACAU  
UUGACUUCUU GUCUCUGUGC UGCAAUUAAU AAAAAAUGGA AAGAAUCUAC AGAUCUAAAA  
 AAAA (서열 번호 35)

[0097]

[0098]

본 발명에 의한 인공 핵산 분자의 3'-UTR 요소 하나 이상은 또한, 알부민 유전자의 3'-UTR의 핵산 서열, 예를 들어 서열 번호 1, 서열 번호 2, 서열 번호 32, 서열 번호 33, 서열 번호 34 또는 서열 번호 35에 따른 핵산 서열에 대한 동일성이 약 40% 이상, 바람직하게는 약 50% 이상, 바람직하게는 약 60% 이상, 바람직하게는 약 70% 이상, 더욱 바람직하게는 약 80% 이상, 더욱 바람직하게는 약 90% 이상, 더더욱 바람직하게는 약 95% 이상, 더더욱 바람직하게는 약 99% 이상, 가장 바람직하게는 100%인 핵산 서열의 단편을 포함하거나 이의 단편으로 이루어져 있는데, 이 경우, 상기 단편은 전술된 기능성 단편 또는 기능성 변이체 단편인 것이 바람직하다. 이와 같은 단편의 길이는 약 50 뉴클레오티드 이상, 바람직하게는 약 75 뉴클레오티드 이상, 더욱 바람직하게는 약 100 뉴클레오티드 이상, 더더욱 바람직하게는 약 125 뉴클레오티드 이상, 가장 바람직하게는 약 150 뉴클레오티드 이상인 것이 바람직하다.

[0099]

예를 들어 이러한 단편은 서열 번호 18~30에 따른 핵산 서열, 예를 들어 다음과 같은 서열들 또는 이에 상응하는 RNA 서열, 또는 이와 같은 핵산 서열과 40% 이상, 바람직하게는 약 50% 이상, 바람직하게는 약 60% 이상, 바람직하게는 약 70% 이상, 더욱 바람직하게는 약 80% 이상, 더욱 바람직하게는 약 90% 이상, 더더욱 바람직하게는 약 95% 이상, 더더욱 바람직하게는 약 99% 이상 동일한 핵산 서열 또는 이에 상응하는 RNA 서열을 나타낼 수 있다:

AAAAGCATCT CAGCCTACCA TGAGAATAAG AGAAAGAAAA TGAAGATCAA AAGCTTATTC  
 ATCTGTTTTT CTTTTTCGTT GGTGTAAAGC CAACACCCTG TCTAAAAAAC ATAAATTTCT  
 TTAATCATT TGCCTCTTTT CTCTGTGCTT CAATT (서열 번호 18)

[0100]

CATCACATTT AAAAGCATCT CAGCCTACCA TGAGAATAAG AGAAAGAAAA TGAAGATCAA  
 AAGCTTATTC ATCTGTTTTT CTTTTTCGTT GGTGTAAAGC CAACACCCTG (서열 번호  
 19)

AAAAGCATCT CAGCCTACCA TGAGAATAAG AGAAAGAAAA TGAAGATCAA AAGCTTATTC  
 ATCTGTTTTT CTTTTTCGTT GGTGTAAAGC CAACACCCTG TCTAAAAAAC (서열 번호  
 20)

CAGCCTACCA TGAGAATAAG AGAAAGAAAA TGAAGATCAA AAGCTTATTC ATCTGTTTTT  
 CTTTTTCGTT GGTGTAAAGC CAACACCCTG TCTAAAAAAC ATAAATTTCT (서열 번호  
 21)

TGAGAATAAG AGAAAGAAAA TGAAGATCAA AAGCTTATTC ATCTGTTTTT CTTTTTCGTT  
 GGTGTAAAGC CAACACCCTG TCTAAAAAAC ATAAATTTCT TTAATCATT (서열 번호  
 22)

AGAAAGAAAA TGAAGATCAA AAGCTTATTC ATCTGTTTTT CTTTTTCGTT GGTGTAAAGC  
 CAACACCCTG TCTAAAAAAC ATAAATTTCT TTAATCATT TGCCTCTTTT (서열 번호  
 23)

TGAAGATCAA AAGCTTATTC ATCTGTTTTT CTTTTTCGTT GGTGTAAAGC CAACACCCTG  
 TCTAAAAAAC ATAAATTTCT TTAATCATT TGCCTCTTTT CTCTGTGCTT (서열 번호  
 24)

AAGCTTATTC ATCTGTTTTT CTTTTTCGTT GGTGTAAAGC CAACACCCTG TCTAAAAAAC  
 ATAAATTTCT TTAATCATT TGCCTCTTTT CTCTGTGCTT CAATTAATAA (서열 번호  
 25)

ATCTGTTTTT CTTTTTCGTT GGTGTAAAGC CAACACCCTG TCTAAAAAAC ATAAATTTCT  
 TTAATCATT TGCCTCTTTT CTCTGTGCTT CAATTAATAA AAAATGGAAA (서열 번호  
 26)

[0101]

CAGCCTACCA TGAGAATAAG AGAAAGAAAA TGAAGATCAA AAGCTTATTC ATCTGTTTTT  
 CTTTTTCGTT GGTGTAAAGC CAACACCCTG TCTAAAAAAC ATAAATTTCT TTAATCATT  
 TGCCTCTTTT CTCTGTGCTT CAATTAATAA A (서열 번호 27)

TGAAGATCAA AAGCTTATTC ATCTGTTTTT CTTTTTCGTT GGTGTAAAGC CAACACCCTG  
 TCTAAAAAAC ATAAATTTCT TTAATCATT TGCCTCTTTT CTCTGTGCTT CAATTAATAA  
 A (서열 번호 28)

CTTTTTTCGTT GGTGTAAAGC CAACACCCTG TCTAAAAAAC ATAAATTTCT TTAATCATT  
 TGCCTCTTTT CTCTGTGCTT CAATTAATAA A (서열 번호 29)

AAGCTTATTC ATCTGTTTTT CTTTTTCGTT GGTGTAAAGC CAACACCCTG TCTAAAAAAC  
 (서열 번호 30)

[0102]

[0103]

그러므로 본 발명에 의한 인공 핵산 분자의 3'-UTR 요소 하나 이상은, 전술된 바와 같은 핵산 단편을 포함하거나 이 단편으로 이루어질 수 있다. 분명한 점은, 서열 번호 18~30에 의한 단편 내에 포함된 티미딘 뉴클레오티드들은 우리딘 뉴클레오티드들로 치환될 수 있다는 점이다.

[0104]

바람직하게 상기 변이체, 단편 또는 변이체 단편은, 바람직하게 서열 번호 1, 서열 번호 2, 서열 번호 32, 서열 번호 33, 서열 번호 34 또는 서열 번호 35에 의한 핵산 서열에 의해 발휘되는 단백질 생산량 증가 효율 및/또는 단백질 생산 안정화 효율의 40% 이상, 더욱 바람직하게는 50% 이상, 더욱 바람직하게는 60% 이상, 더더욱 바람직하게는 70% 이상, 더더욱 바람직하게는 80% 이상, 가장 바람직하게는 90% 이상에 해당하는 효율로, 전술된 바와 같이 서열 번호 1, 서열 번호 2, 서열 번호 32, 서열 번호 33, 서열 번호 34 또는 서열 번호 35에 의한 핵산 서열의 기능, 예를 들어 본 발명에 의하여 인공 핵산 분자를 안정화하는 기능, 본 발명에 의한 인공 핵산 분자로부터의 단백질 발현을 안정화 및/또는 연장하는 기능, 및/또는 단백질 생산량을 증가시키는 기능을 하나 이상 발휘하는 기능성 변이체, 기능성 단편 또는 기능성 변이체 단편이다.

[0105]

바람직하게 본 발명에 의한 인공 핵산 분자의 3'-UTR 요소 하나 이상의 길이는 약 50 뉴클레오티드 이상, 바람직하게는 약 75 뉴클레오티드 이상, 더욱 바람직하게는 약 100 뉴클레오티드 이상, 더더욱 바람직하게는 약 125 뉴클레오티드 이상, 가장 바람직하게는 약 150 뉴클레오티드 이상이다. 예를 들어 3'-UTR 요소의 길이는 약 50~약 300 뉴클레오티드, 바람직하게는 약 100~약 250 뉴클레오티드, 더욱 바람직하게는 약 150~약 200 뉴클레오티드일 수 있다.

[0106]

뿐만 아니라 본 발명에 의한 인공 핵산 분자는 전술된 바와 같은 3'-UTR 요소를 하나 이상 포함할 수 있다. 예를 들어 본 발명에 의한 인공 핵산 분자는 1개, 2개, 3개, 4개 또는 그 이상의 3'-UTR 요소를 포함할 수 있는데, 이 경우, 각각의 3'-UTR 요소는 동일할 수 있으며, 아니면 상기 3'-UTR 요소는 상이할 수 있다. 예를 들어 본 발명에 의한 인공 핵산 분자는 전술된 바와 같은 본질적으로 동일한 3'-UTR 요소 2개, 예를 들어 알부민 유전자의 3'-UTR 또는 알부민 유전자의 3'-UTR의 변이체로부터 유래하는 핵산 서열, 예를 들어 서열 번호 1, 2, 32, 33, 34 또는 35에 의한 핵산 서열, 이의 기능성 변이체, 이의 기능성 단편 또는 이의 기능성 변이체 단편을 포함하거나 이것들로 이루어져 있는 3'-UTR 요소 2개를 포함할 수 있다.

[0107]

놀랍게도 본 발명의 발명자들은, 전술된 바와 같은 3'-UTR을 포함하는 인공 핵산 분자가 단백질 생산을 연장 및/또는 안정화할 수 있는 mRNA 분자를 나타낼 수 있거나 제공할 수 있음을 발견하게 되었다. 그러므로 본원에 기술된 바와 같은 3'-UTR은 mRNA 분자로부터의 단백질 발현의 안정성을 개선할 수 있으며/개선할 수 있거나 번역 효율을 증가시킬 수 있다.

[0108]

본 발명에 의한 인공 핵산 분자는 RNA, 예를 들어 mRNA, DNA, 예를 들어 DNA 벡터일 수 있거나, 또는 변형된 RNA 또는 DNA 분자일 수 있다. 본 발명에 의한 인공 핵산 분자는 센스 사슬과 안티센스 사슬을 가지는 이중 사슬 분자, 예를 들어 센스 사슬과 안티센스 사슬을 가지는 DNA 분자로서 제공될 수 있다.

[0109]

본 발명에 의한 인공 핵산 분자는 임의로 5'-UTR 및/또는 5'-캡을 추가로 포함할 수 있다. 임의의 5'-캡 및/또

는 5'-UTR은 본 발명에 의한 인공 핵산 분자 내 ORF에 대해 5' 말단 쪽에 위치하는 것이 바람직하다.

[0110] 바람직하게 본 발명에 의한 인공 핵산 분자는 폴리(A) 서열 및/또는 폴리아데닐화 신호를 추가로 포함한다. 바람직하게 임의의 폴리(A) 서열은 3'-UTR 요소 하나 이상에 대해 3'-말단 쪽에 위치하며, 바람직하게 3'-UTR 요소의 3'-말단에 연결되어 있다. 연결 관계는, 예를 들어 하나 이상의 제한 위치를 포함하거나 이 제한 위치로 이루어져 있는, 2, 4, 6, 8, 10, 20개 등의 뉴클레오티드로 이루어진 구획, 예를 들어 1-50개, 바람직하게는 1-20개의 뉴클레오티드로 이루어진 링커를 통한 직접적 또는 간접적 연결 관계일 수 있다.

[0111] 하나의 구체예에서, 임의의 폴리아데닐화 신호는 3'-UTR 요소 내에 위치한다. 바람직하게 폴리아데닐화 신호는 공통 서열인 NN(U/T)ANA[여기서, N은 A 또는 U임], 바람직하게 AA(U/T)AAA 또는 A(U/T)(U/T)AAA를 포함한다. 이와 같은 공통 서열은 대부분의 동물 및 박테리아 세포 계, 예를 들어 폴리아데닐화 인자, 예를 들어 CstF, PAP, PAB2, CFI 및/또는 CFII와 공동으로 작용하는 절단/폴리아데닐화 특이성 인자(CPSF)에 의해 인지될 수 있다. 바람직하게 폴리아데닐화 신호, 바람직하게 공통 서열 NNUANA는 3'-UTR 요소의 3'-말단 쪽 약 50 뉴클레오티드 미만, 더욱 바람직하게는 약 30개 염기 미만, 가장 바람직하게는 약 25개 염기 미만, 예를 들어 21개 염기 만큼 상류에 위치한다.

[0112] 3'-UTR 요소 내에 폴리아데닐화 신호를 포함하는 본 발명에 의한 인공 핵산 분자의 전사, 예를 들어 인공 DNA 분자의 전사는, 이 핵산 분자 자체의 3'-UTR 요소 내에 폴리아데닐화 신호를 포함하는 미성숙 RNA를 생산해낼 것이다. 예를 들어 서열 번호 1, 즉

```
CATCACATTT AAAAGCATCT CAGCCTACCA TGAGAATAAG AGAAAGAAAA TGAAGATCAA
AAGCTTATTC ATCTGTTTTT CTTTTTCGTT GGTGTAAAGC CAACACCCTG TCTAAAAAAC
ATAAATTTCT TTAATCATTT TGCCTCTTTT CTCTGTGCTT CAATTAATAA AAAATGGAAA
GAATCT (서열 번호 1)
```

[0113] 에 의한 3'-UTR 요소를 포함하는 DNA 분자의 전사는 서열

```
CAUCACAUUU AAAAGCAUCU CAGCCUACCA UGAGAAUAAG AGAAAGAAAA UGAAGAUCAA
AAGCUUAUUC AUCUGUUUUU CUUUUUCGUU GGUGUAAAGC CAACACCUG UCUAAAAAAC
AUAAAUUUCU UUAUCAUUU UGCCUCUUUU CUCUGUGCUU CAUUUAUAA AAAAUGGAAA
GAAUCU (서열 번호 2)
```

[0115] 에 의한 3'-UTR 요소를 가지는 RNA를 생산해낼 것이다.

[0117] 적당한 전사 계를 이용하면 미성숙 RNA에 폴리(A) 서열이 부착될 것이다. 예를 들어 본 발명의 인공 핵산 분자는 전술된 바와 같은 3'-UTR 요소와 폴리아데닐화 신호를 포함하는 DNA 분자일 수 있으며, 이 DNA 분자가 전사됨에 따라서 RNA는 폴리아데닐화될 수 있는 것이다. 그러므로 생산된 RNA는 본 발명의 3'-UTR 요소와 폴리(A) 서열을 이 순서대로 함께 포함할 수 있다.

[0118] 잠재적인 전사 계는 시험관 내 전사 계 또는 세포 전사 계 등이다. 그러므로 본 발명에 의한 인공 핵산 분자의 전사, 예를 들어 개방 해독 틀, 3'-UTR 요소 및 폴리아데닐화 신호를 포함하는 인공 핵산 분자의 전사는, 개방 해독 틀, 3'-UTR 요소 및 폴리(A) 서열을 포함하는 mRNA 분자를 생산해낼 수 있다.

[0119] 그러므로 본 발명은 또한 개방 해독 틀, 전술된 3'-UTR 요소 및 폴리(A) 서열을 포함하는 mRNA 분자인 인공 핵산 분자를 제공하기도 한다.

[0120] 하나의 구체예에서, 본 발명은 전술된 바와 같이 개방 해독 틀과, 서열 번호 1에 의한 서열 또는 서열 번호 1의 서열에 대한 동일성이 약 40% 이상 또는 그보다 높은 서열 또는 이것들의 단편을 포함하는 인공 DNA 분자인 인공 핵산 분자를 제공한다. 뿐만 아니라 본 발명은 전술된 바와 같이 개방 해독 틀과, 서열 번호 2에 의한 서열 또는 서열 번호 2의 서열에 대한 동일성이 약 40% 이상 또는 그보다 높은 서열 또는 이것들의 단편을 포함하는 인공 RNA 분자인 인공 핵산 분자를 제공한다.

[0121] 하나의 추가 구체예에서 본 발명은 전술된 바와 같이 개방 해독 틀과, 서열 번호 32에 의한 서열 또는 서열 번호 32의 서열에 대한 동일성이 약 40% 이상 또는 그보다 높은 서열 또는 이것들의 단편을 포함하는 인공 DNA 분자인 인공 핵산 분자를 제공한다. 뿐만 아니라 본 발명은 전술된 바와 같이 개방 해독 틀과, 서열 번호 33에 의

한 서열 또는 서열 번호 33의 서열에 대한 동일성이 약 40% 이상 또는 그보다 높은 서열 또는 이것들의 단편을 포함하는 인공 RNA 분자인 인공 핵산 분자를 제공한다.

[0122] 더욱이 본 발명은 전술된 바와 같이 개방 해독 통과, 서열 번호 34에 의한 서열 또는 서열 번호 34의 서열에 대한 동일성이 약 40% 이상 또는 그보다 높은 서열 또는 이것들의 단편을 포함하는 인공 DNA 분자인 인공 핵산 분자를 제공한다. 뿐만 아니라 본 발명은 전술된 바와 같이 개방 해독 통과, 서열 번호 35에 의한 서열 또는 서열 번호 35의 서열에 대한 동일성이 약 40% 이상 또는 그보다 높은 서열 또는 이것들의 단편을 포함하는 인공 RNA 분자인 인공 핵산 분자를 제공한다.

[0123] 그러므로 본 발명은 RNA 분자, 바람직하게는 mRNA 분자에 대한 주형일 수 있는 인공 핵산 분자로서, 번역 효율의 관점에서 안정화 및 최적화된 인공 핵산 분자를 제공한다. 다시 말해서, 본 발명의 인공 핵산 분자는 mRNA의 생산에 사용될 수 있는 DNA 또는 RNA일 수 있다. 생산될 수 있는 mRNA는 이후 번역되며, 그 결과 개방 해독 통과에 의해 암호화되는, 원하는 펩티드 또는 단백질이 생산된다. 만일 인공 핵산 분자가 DNA이면, 상기 인공 핵산 분자는, 예를 들어 mRNA의 시험관 내 또는 생체 내 연속적이고도 반복적인 생산에 사용될 이중 사슬 저장형으로서 사용될 수 있다.

[0124] 하나의 구체예에서, 본 발명에 의한 인공 핵산 분자는 폴리(A) 서열을 추가로 포함한다. 폴리(A) 서열의 길이는 다양할 수 있다. 예를 들어 폴리(A) 서열의 길이는 약 20개 아데닌 뉴클레오티드~약 300개 이하의 아데닌 뉴클레오티드, 바람직하게는 약 40개~약 200개 아데닌 뉴클레오티드, 더욱 바람직하게 약 50개~약 100개 아데닌 뉴클레오티드, 예를 들어 약 60, 70, 80, 90 또는 100개 아데닌 뉴클레오티드일 수 있다.

[0125] 예를 들어 본 발명에 의한 인공 핵산 분자는 다음과 같은 DNA 서열들에 상응하는 핵산 서열을 포함할 수 있다:

CATCACATTT AAAAGCATCT CAGCCTACCA TGAGAATAAG AGAAAGAAA TGAAGATCAA  
 AAGCTTATTC ATCTGTTTTT CTTTTTCGTT GGTGTAAAGC CAACACCCTG TCTAAAAAAC  
 ATAAATTTCT TTAATCATTT TGCCTCTTTT CTCTGTGCTT CAATTAATAA AAAATGGAAA  
 GAATCTAGAT CTAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA  
 AAAAAAAAAA AAAAAA (서열 번호 3),

AAACATCACA ATTAAGAACA TCTCAGCCTA CCATGAGAAC AAGAGAAATA AAATGAAGAT  
CAAAAGCTTA TTCATCTGTT TTTCTTTTTT ATTGGTATAA AGCCAACACC CTGTCTAAAA  
AACTATAAAT TTCTTTAATC ATTTTGCCTC TTTTCTCTGT GCTTCAATTA ATAAAAAATG  
GAAAGAATCT AGATCTAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA  
 AAAAAAAAAA AAAAAAAAAA (서열 번호 36)

또는

ACACATCACA ACCACAACCT TCTCAGGCTA CCCTGAGAAA AAAAGACATG AAGACTCAGG  
ACTCATCTTT TCTGTTGGTG TAAAATCAAC ACCCTAAGGA ACACAAATTT CTTTAAACAT  
TTGACTTCTT GTCTCTGTGC TGCAATTAAT AAAAAATGGA AAGAATCTAC AGATCTAAAA  
 AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA  
 (서열 번호 38)

[0126]

[0127] 이와 같은 서열들이 전사되면, 다음과 같은 서열들을 포함하는 인공 핵산 분자들이 생산될 수 있다:

CAUCACAUUU AAAAGCAUCU CAGCCUACCA UGAGAAUAG AGAAAGAAA UGAAGAUCAA  
 AAGCUUAUUC AUCUGUUUUU CUUUUUCGUU GGUGUAAAAG CAACACCCUG UCUAAAAAAC  
 AUAAAUUUCU UAAAUCAUUU UGCCUCUUUU CUCUGUGCUU CAAUUAUAA AAAAUGGAAA  
 GAAUCUAGAU CUAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA  
 AAAAAAAAAA AAAAAA (서열 번호 4),

AAACAUCACA AUUAAGAACA UCUCAGCCUA CCAUGAGAAC AAGAGAAUA AAAUGAAGAU  
CAAAAGCUUA UUCAUCUGUU UUUCUUUUUC AUUGGUUAA AGCCAACACC CUGUCUAAAA  
AACUAUAAA UUCUTTAATC ATTTTGCCTC TTTTCTCTGT GCTTCAATTA ATAAAAATG

[0128]

GAAAGAATCT AGATCTAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA  
 AAAAAAAAAA AAAAAAAAAA (서열 번호 37)

또는

ACACAUCACA ACCACAACCU UCUCAGGCUA CCCUGAGAAA AAAAGACAUG AAGACUCAGG  
ACUCAUCUUU UCGUUGGUG UAAAAUCAAC ACCCUAAGGA ACACAAUUU CUUUAACAU  
UUGACUUCUU GUCUCUGUGC UGCAAUUAAU AAAAAUGGA AAGAAUCUAC AGAUCUAAAA  
 AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA  
 (서열 번호 39)

[0129]

[0130] 이와 같은 인공 RNA 분자, 즉 서열 번호 4, 37 또는 39에 의한 서열을 포함하는 인공 핵산 분자는 또한 DNA 전구체(DNA progenitor)로부터의 전사 과정을 반드시 거치지 않고서도 통상의 화학적 합성 방법에 의해 시험관 내에서 생산될 수도 있다.

[0131] 특히 바람직한 구체예에서, 본 발명에 의한 인공 핵산 분자는 5'→3' 방향 개방 해독 틀, 전술된 3'-UTR 및 폴리(A) 서열을 포함하는 RNA 분자, 바람직하게 mRNA 분자이다.

[0132] 바람직한 구체예에서, 개방 해독 틀은 알부민, 특히 사람 알부민, 마우스 알부민 또는 올리브 개코 원숭이 유래 알부민을 암호화하지 않는데, 다만 3'-UTR 요소는 사람 알부민, 마우스 알부민 또는 올리브 개코 원숭이 유래 알부민의 3'-UTR 각각과 동일하다. 몇몇 추가의 구체예에서, 개방 해독 틀이 유전자 은행 수탁 번호 NM\_000477.5에 의한 사람 알부민을 암호화하지 않되, 다만 3'-UTR 요소는 사람 알부민의 3'-UTR과 동일한 것이 바람직할 수 있다. 몇몇 추가의 구체예에서, 개방 해독 틀이 알부민 또는 이의 변이체를 암호화하지 않되, 다만 3'-UTR 요소는 서열 번호 1, 서열 번호 2, 서열 번호 32, 서열 번호 33, 서열 번호 34 또는 서열 번호 35와 동일한 것이 바람직할 수 있다. 뿐만 아니라 몇몇 구체예에서, 개방 해독 틀은 사람 인자 IX 또는, 예를 들어 글로빈 단백질(특히 베타-글로빈), 루시페라제 단백질, GFP 단백질 또는 이것들의 변이체, 예를 들어 글로빈 단백질, 루시페라제 단백질 또는 GFP 단백질에 대한 서열 동일성이 70% 이상인 변이체로 이루어진 균으로부터 선택되는 리포터 단백질을 암호화하지 않는 것이 바람직하다. 또한 특정 구체예에서, 특히 개방 해독 틀이 사람 인자 IX를 암호화하는 경우, 개방 해독 틀은 인트론을 포함하지 않는 것이 바람직하다.

[0133] 하나의 구체예에서, 본 발명은 개방 해독 틀, 바람직하게는 알부민 이외의 펩티드 또는 단백질을 암호화하는 개방 해독 틀; 서열 번호 1, 32 또는 34의 서열에 대한 서열 동일성이 약 60% 이상, 바람직하게는 약 70% 이상, 더욱 바람직하게는 약 80% 이상, 더욱 바람직하게는 약 90% 이상, 더더욱 바람직하게는 약 95% 이상, 더더욱 바람직하게는 99% 이상, 더더욱 바람직하게는 100%인 서열을 포함하거나 이 서열로 이루어져 있는 3'-UTR 요소; 그리고 폴리아데닐화 신호 및/또는 폴리(A) 서열을 포함하는 인공 핵산 분자를 제공한다. 뿐만 아니라 본 발명은 개방 해독 틀, 바람직하게는 알부민 이외의 임의의 펩티드 또는 단백질을 암호화하는 개방 해독 틀; 서열 번호 3, 36 또는 38의 서열에 대한 서열 동일성이 약 60% 이상, 바람직하게는 약 70% 이상, 더욱 바람직하게는 약 80% 이상, 더욱 바람직하게는 약 90% 이상, 더더욱 바람직하게는 약 95% 이상, 더더욱 바람직하게는 99% 이상, 더더욱 바람직하게는 100%인 서열을 포함하거나 이 서열로 이루어져 있는 3'-UTR 요소를 포함하는 인공 DNA 분

자를 제공한다.

[0134] 또한 본 발명은 개방 해독 틀, 바람직하게는 알부민 이외의 펩티드 또는 단백질을 암호화하는 개방 해독 틀; 서열 번호 2, 33 또는 35의 서열에 대한 서열 동일성이 약 60% 이상, 바람직하게는 약 70% 이상, 더욱 바람직하게는 약 80% 이상, 더욱 바람직하게는 약 90% 이상, 더더욱 바람직하게는 약 95% 이상, 더더욱 바람직하게는 99% 이상, 더더욱 바람직하게는 100%인 서열을 포함하거나 이 서열로 이루어져 있는 3'-UTR 요소; 그리고 폴리아데닐화 신호 및/또는 폴리(A) 서열을 포함하는 인공 RNA 분자, 바람직하게는 인공 mRNA 분자 또는 인공 바이러스 RNA 분자를 제공한다. 뿐만 아니라 본 발명은 개방 해독 틀, 바람직하게는 알부민 이외의 펩티드 또는 단백질을 암호화하는 개방 해독 틀; 서열 번호 4, 37 또는 39의 서열에 대한 서열 동일성이 약 60% 이상, 바람직하게는 약 70% 이상, 더욱 바람직하게는 약 80% 이상, 더욱 바람직하게는 약 90% 이상, 더더욱 바람직하게는 약 95% 이상, 더더욱 바람직하게는 99% 이상, 더더욱 바람직하게는 100%인 서열을 포함하거나 이 서열로 이루어져 있는 3'-UTR 요소를 포함하는 인공 RNA 분자, 바람직하게는 인공 mRNA 분자 또는 인공 바이러스 RNA 분자를 제공한다.

[0135] 본 발명은 암호화된 펩티드 또는 단백질의 발현 안정성이 증강되었고 발현이 연장된 것을 특징으로 할 수 있는 인공 핵산 분자, 바람직하게는 인공 mRNA를 제공한다. 어떠한 이론에도 국한되지 않을 때, 단백질 발현의 증강된 안정성과 이로 인한 단백질 발현의 연장은, 본 발명에 의한 인공 핵산 분자, 예를 들어 인공 mRNA 분자의 분해를 감소시키는 것에서 기인할 수 있다. 그러므로 본 발명의 3'-UTR 요소는 인공 핵산이 분해 및 붕괴되는 것을 막을 수 있다.

[0136] 몇몇 구체예에서, 3'-UTR 요소는 히스톤 스템-루프를 이루어져 있지 않은 것이 바람직하며, 상기 3'-UTR 요소는 히스톤 스템-루프를 포함하지 않는 것이 바람직하다. 하나의 구체예에서, 본 발명에 의한 인공 핵산 분자는 히스톤 스템-루프를 포함하지 않는다. 그러나 몇몇 구체예에서, 본 발명의 인공 핵산 분자는 알부민 유전자의 3'-UTR로부터 유래하는 핵산 서열 이외에 히스톤 스템-루프를 포함할 수 있다. 이와 같이 본 발명에 의한 인공 핵산 분자는, 예를 들어 ORF와, 바람직하게는 폴리아데닐화 신호를 포함하는 본 발명의 3'-UTR 요소, 임의의 히스톤 스템-루프 및 임의의 폴리(A) 서열을 5'→3' 방향으로 포함할 수 있다. 또한 본 발명의 인공 핵산 분자는 ORF와, 예를 들어 폴리아데닐화 신호를 포함하는 본 발명의 3'-UTR 요소, 폴리(A) 서열 및 임의의 히스톤 스템-루프를 5'→3' 방향으로 포함할 수도 있다.

[0137] 본 발명의 내용에 있어서, 이와 같은 히스톤 스템-루프는 통상 히스톤 유전자로부터 유래하고, 2개의 이웃하는 전체 또는 부분 역전 상보성 서열들의 분자 내 염기쌍을 포함함으로써 스템-루프를 형성할 수 있다. 스템-루프는 단일 사슬 DNA, 더욱 일반적으로는 RNA에서 형성될 수 있다. 상기 구조는 또한 헤어핀 또는 헤어핀 루프라고도 알려져 있으며, 보통 연속 서열 내 (말단) 루프 부와 스템 부로 이루어져 있는데, 상기 스템은 2개의 이웃하는 전체 또는 부분 역전 상보성 서열들(스템-루프 구조 중 루프를 이루는 짧은 서열, 즉 일종의 스페이서(spacer)에 의해 떨어져 있음)에 의해 형성된다. 상기 2개의 이웃하는 전체 또는 부분 역전 상보성 서열들은, 예를 들어 스템-루프 요소 스템 1 및 스템 2로서 정의될 수 있다. 스템-루프는, 상기 2개의 이웃하는 전체 또는 부분 역전 상보성 서열들, 예를 들어 스템-루프 요소 스템 1 및 스템 2가 서로 염기 쌍을 이루어, 쌍을 형성하지 않은 루프(연속 서열상 스템-루프 요소 스템 1 및 스템 2 사이에 위치하는 짧은 서열에 의해 형성됨)를 자체의 말단에 포함하는 이중 사슬 핵산 서열이 형성될 때 만들어진다. 이로써 상기 쌍을 형성하지 않은 루프는 통상 상기된 바와 같은 스템-루프 요소들 중 어느 하나와 염기 쌍을 형성할 수 없는 핵산의 한 영역을 이룬다. 형성된 롤리팝 형상의 구조는 다수의 RNA 2차 구조를 이루는 핵심 구성 블록(building block)이다. 그러므로 스템-루프 구조의 형성은 생성된 스템 및 루프 영역들의 안정성에 좌우되는데, 이 경우, 제1 전체 조건은 통상 서열 자체가 폴딩백(folding back)되어, 쌍을 이루는 이중 사슬을 이룰 수 있는 서열이 존재하는 것이다. 쌍을 이루는 스템-루프 요소들의 안정성은 길이, 이중 사슬이 포함하고 있는 미스매치부(mismatch) 또는 돌출부(bulge)의 수(특히 길이가 긴 이중 사슬에서 미스매치부가 소수 개 존재하는 것은 통상 관용됨), 그리고 쌍을 이루는 영역의 염기 조성에 의해 결정된다. 본 발명의 내용에 있어서, 최적 루프 길이는 3~10개 염기, 더욱 바람직하게는 3~8개, 3~7개, 3~6개 염기, 또는 더욱 바람직하게 4~5개 염기, 그리고 가장 바람직하게는 4개 염기이다.

[0138] 히스톤 스템-루프 서열에 관한 일례로서는 서열 번호 31에 의한 서열(CAAAGGCTCTTTTCAGAGCCACCA) 또는 이에 상응하는 RNA 서열이 있다.

[0139] 그러므로 몇몇 구체예에서, 본 발명에 의한 인공 핵산 분자는 (a) 개방 해독 틀 하나 이상; (b) 전술된 바와 같은 3'-UTR 요소 하나 이상, 그리고 (c) 예를 들어 서열 번호 31에 의한 서열에 대한 서열 동일성이 약 75% 이상, 바람직하게는 약 80% 이상, 바람직하게는 약 85% 이상, 더욱 바람직하게는 약 90% 이상, 더더욱 바람직하

게는 약 95% 이상인 서열 또는 이에 상응하는 RNA 서열을 포함할 수 있거나 이와 같은 서열로 이루어져 있을 수 있는 히스톤-스텝 루프 하나 이상을 포함하는데, 이 경우, 서열 번호 31에 의한 서열에 대한 서열 동일성이 약 75% 이상, 바람직하게는 약 80% 이상, 바람직하게는 약 85% 이상, 더욱 바람직하게는 약 90% 이상, 더더욱 바람직하게는 약 95% 이상인 서열 또는 이에 상응하는 RNA 서열의 6, 13 및 20번 위치들은 보존되는 것(즉 서열 번호 31에 의한 서열의 6, 13 및 20번 위치의 뉴클레오티드들과 동일한 것)이 바람직하다.

[0140] 몇몇 구체예에서, 본 발명의 인공 핵산 분자는 추가의 요소들, 예를 들어 5'-캡, 폴리(C) 서열 및/또는 IRES-모티프를 포함한다. 5'-캡은 전사 후에 RNA의 5' 말단에 부가될 수 있다. 뿐만 아니라 본 발명의 인공 핵산 분자는 (특히 이 핵산이 mRNA의 형태이거나 mRNA를 암호화한다면) 10개 이상의 시티딘, 바람직하게는 20개 이상의 시티딘, 더욱 바람직하게는 30개 이상의 시티딘으로 이루어진 서열(소위 "폴리(C) 서열")에 의해 변형될 수 있다. 특히 본 발명의 핵산 분자는 (특히 이 핵산이 (m)RNA의 형태이거나 mRNA를 암호화한다면) 통상 약 10~200개의 시티딘 뉴클레오티드, 바람직하게는 약 10~100개의 시티딘 뉴클레오티드, 더욱 바람직하게는 약 10~70개의 시티딘 뉴클레오티드 또는 더더욱 바람직하게는 약 20~50개 또는 심지어 20~30개의 시티딘 뉴클레오티드로 이루어진 폴리(C) 서열을 포함할 수 있다. 그러므로 본 발명에 의한 인공 핵산 분자는 바람직하게 ORF, 전술된 바와 같은 3'-UTR 요소 하나 이상, 폴리(A) 서열 또는 폴리(아데닐화 신호, 그리고 폴리(C) 서열을 5'→3' 방향으로 포함하는 것이 바람직하다.

[0141] 예를 들어 만일 본 발명의 인공 핵산 분자가 2개 이상의 켈티드 또는 단백질을 암호화한다면, 내부 리보솜 도입 위치(Internal Ribosome Entry Site; IRES) 서열 또는 IRES-모티프는 몇몇 개방 해독 틀을 격리할 수 있다. 만일 mRNA가 2 시스트론형 RNA이거나 다중 시스트론형 RNA라면 IRES-서열은 특히 유리할 수 있다.

[0142] 뿐만 아니라 본 발명의 인공 핵산 분자는 부가의 5'-요소, 바람직하게는 5'-UTR, 프로모터, 또는 5'-UTR과 프로모터를 포함하는 서열을 포함할 수 있다. 프로모터는 본 발명에 의한 인공 핵산 분자, 예를 들어 본 발명에 의한 인공 DNA 분자의 전사를 가동 및/또는 조절할 수 있다. 뿐만 아니라 5'-UTR은 본 발명의 3'-UTR 요소와 상호작용을 할 수 있으므로, 본 발명의 3'-UTR 요소의 안정화 효과를 뒷받침할 수 있다. 이와 같은 요소는 안정성과 번역 효율을 추가로 뒷받침할 수 있다. 그러므로 몇몇 구체예에서, 본 발명은 5'→3' 방향으로 다음과 같은 구조들 중 하나 이상을 포함하는 인공 핵산 분자, 바람직하게는 mRNA 분자를 제공한다:

[0143] 5'-캡—5'-UTR—ORF—3'-UTR 요소—히스톤 스텝-루프—폴리(A) 서열

[0144] 5'-캡—5'-UTR—ORF—3'-UTR 요소—폴리(A) 서열—히스톤 스텝-루프

[0145] 5'-캡—5'-UTR—ORF—IRES—ORF—3'-UTR 요소—히스톤 스텝-루프—폴리(A) 서열

[0146] 5'-캡—5'-UTR—ORF—IRES—ORF—3'-UTR 요소—히스톤 스텝-루프—폴리(A) 서열—폴리(C) 서열

[0147] 5'-캡—5'-UTR—ORF—IRES—ORF—3'-UTR 요소—폴리(A) 서열—히스톤 스텝-루프

[0148] 5'-캡—5'-UTR—ORF—IRES—ORF—3'-UTR 요소—폴리(A) 서열—폴리(C) 서열—히스톤 스텝-루프

[0149] 5'-캡—5'-UTR—ORF—3'-UTR 요소—폴리(A) 서열—폴리(C) 서열

[0150] 5'-캡—5'-UTR—ORF—3'-UTR 요소—폴리(A) 서열—폴리(C) 서열—히스톤 스텝-루프

[0151] 바람직하게 본 발명에 의한 인공 핵산 분자, 바람직하게 개방 해독 틀은 적어도 부분적으로 G/C에 변화가 있다. 그러므로 본 발명의 인공 핵산 분자는 분자의 G(구아닌)/C(시티딘) 함량을 바꾸어줌으로써 열역학적으로 안정화될 수 있다. 본 발명에 의한 인공 핵산 분자의 개방 해독 틀의 G/C 함량은, 바람직하게는 유전 암호의 축퇴성이 이용됨으로써, 상응하는 야생형 서열의 개방 해독 틀의 G/C 함량과 비교되게 증가할 수 있다. 그러므로 핵산 분자의 암호화된 아미노산 서열은 G/C에 변화가 가하여졌을 때, 해당 핵산 분자의 야생형 서열의 암호화된 아미노산 서열과 비교되게 변형되지 않는 것이 바람직하다. 따라서 암호화 서열 또는 전체 핵산 분자, 예를 들어 mRNA의 코돈들은 야생형 암호화 서열의 경우와 비교되게 다양할 수 있으므로, 상기 코돈들에 있어서는 번역된 아미노산 서열은 유지되면서 G/C 뉴클레오티드의 양만 증가하게 된다. 몇몇 코돈들은 하나의 아미노산 및 이와 동일한 아미노산을 암호화한다는 사실(소위 유전 암호의 축퇴성)과 관련하여, 안정성에 가장 유리한 코돈들이 결정될 수 있다(소위 대안적 코돈 선호도).

[0152] 본원에 정의된 바와 같은 본 발명의 핵산 분자의 암호화 영역에 의해 암호화될 아미노산에 따라서, 상기 핵산 분자의 야생형 암호화 영역과 비교되게, 핵산 서열, 예를 들어 개방 해독 틀이 변형될 여러 가지 가능성이 존재한다. 배타적으로 G 또는 C 뉴클레오티드만을 포함하는 코돈에 의해 암호화되는 아미노산의 경우, 코돈 변형은

필수적이지 않다. 그러므로 Pro(CCC 또는 CCG), Arg(CGC 또는 CGG), Ala(GCC 또는 GCG) 및 Gly(GGC 또는 GGG)와 같은 코돈들은 변형을 필요로 하지 않는데, 그 이유는 A 또는 U/T가 존재하기 때문이다.

- [0153] 이와는 대조적으로 A 및/또는 U/T 뉴클레오티드를 포함하는 코돈들은, 동일한 아미노산을 암호화하되 A 및/또는 U/T는 포함하지 않는 다른 코돈들의 치환을 통해 변형될 수 있다. 예를 들어
- [0154] Pro에 대한 코돈은 CC(U/T) 또는 CCA에서 CCC 또는 CCG로 변형될 수 있고;
- [0155] Arg에 대한 코돈은 CG(U/T) 또는 CGA 또는 AGA 또는 AGG에서 CGC 또는 CGG로 변형될 수 있으며;
- [0156] Ala에 대한 코돈은 GC(U/T) 또는 GCA에서 GCC 또는 GCG로 변형될 수 있고;
- [0157] Gly에 대한 코돈은 GG(U/T) 또는 GGA에서 GGC 또는 GGG로 변형될 수 있다.
- [0158] 다른 경우에 있어서, 비록 A 또는 (U/T) 뉴클레오티드들이 코돈들로부터 제외될 수 없다고는 하지만, A 및/또는 (U/T) 뉴클레오티드들의 함량이 낮은 코돈들을 이용함으로써 A 및 (U/T) 함량을 낮추는 것이 가능하다. 이에 관한 예들로서는 다음과 같은 경우들이 있다:
- [0159] Phe에 대한 코돈은 (U/T)(U/T)(U/T)에서 (U/T)(U/T)C로 변형될 수 있고;
- [0160] Leu에 대한 코돈은 (U/T)(U/T)A, (U/T)(U/T)G, C(U/T)(U/T) 또는 C(U/T)A에서 C(U/T)C 또는 C(U/T)G로 변형될 수 있으며;
- [0161] Ser에 대한 코돈은 (U/T)C(U/T) 또는 (U/T)CA 또는 AG(U/T)로부터 (U/T)CC, (U/T)CG 또는 AGC로 변형될 수 있고;
- [0162] Tyr에 대한 코돈은 (U/T)A(U/T)에서 (U/T)AC로 변형될 수 있으며;
- [0163] Cys에 대한 코돈은 (U/T)G(U/T)에서 (U/T)GC로 변형될 수 있고;
- [0164] His에 대한 코돈은 CA(U/T)에서 CAC로 변형될 수 있으며;
- [0165] Gln에 대한 코돈은 CAA에서 CAG로 변형될 수 있고;
- [0166] Ile에 대한 코돈은 A(U/T)(U/T) 또는 A(U/T)A에서 A(U/T)C로 변형될 수 있으며;
- [0167] Thr에 대한 코돈은 AC(U/T) 또는 ACA에서 ACC 또는 ACG로 변형될 수 있고;
- [0168] Asn에 대한 코돈은 AA(U/T)에서 AAC로 변형될 수 있으며;
- [0169] Lys에 대한 코돈은 AAA에서 AAG로 변형될 수 있으며;
- [0170] Val에 대한 코돈은 G(U/T)(U/T) 또는 G(U/T)A에서 G(U/T)C 또는 G(U/T)G로 변형될 수 있고;
- [0171] Asp에 대한 코돈은 GA(U/T)에서 GAC로 변형될 수 있으며;
- [0172] Glu에 대한 코돈은 GAA에서 GAG로 변형될 수 있고;
- [0173] 종결 코돈 (U/T)AA는 (U/T)AG 또는 (U/T)GA로 변형될 수 있다.
- [0174] 다른 한편, 코돈 Met(A(U/T)G) 및 Trp((U/T)GG)의 경우, 암호화된 아미노산 서열이 변형되지 않고 서열 변형이 일어날 가능성은 없다.
- [0175] 상기 나열된 치환들은, 본원에 정의된 본 발명의 핵산 서열의 개방 해독들의 G/C 함량, 상기 핵산 서열의 구체적인 야생형 개방 해독들(즉 원 서열)과 비교되게 증가하도록, 개별적으로 또는 가능한 모든 조합에 따라서 이용될 수 있다. 그러므로, 예를 들어 야생형 서열에 발생하는 Thr에 대한 모든 코돈들은 ACC(또는 ACG)로 변형될 수 있다.
- [0176] 바람직하게 본원에 정의된 본 발명의 인공 핵산 분자의 개방 해독들의 G/C 함량은, 암호화된 아미노산 서열이 변형되지 않으면서(즉 유전 암호의 축퇴성이 이용되면서) 야생형 암호화 영역의 G/C 함량과 비교되게 7% 이상, 더욱 바람직하게는 15% 이상, 특히 바람직하게는 20% 이상 증가한다. 특정 구체예에 의하면, 본 발명의 인공 핵산 분자, 이의 단편, 변이체 또는 유도체의 개방 해독들 내 치환 가능한 코돈들 중 5% 이상, 10% 이상, 20% 이상, 30% 이상, 40% 이상, 50% 이상, 60% 이상, 더욱 바람직하게는 70% 이상, 더더욱 바람직하게는 80% 이상, 그리고 가장 바람직하게는 90% 이상, 95% 이상 또는 심지어 100%가 치환되면, 상기 개방 해독들의 G/C 함량은 증

가하게 된다.

- [0177] 이와 같은 내용에 있어서, 본원에 정의된 본 발명의 핵산 서열의 개방 해독 틀의 G/C 함량은, 암호화된 아미노산 서열은 변형되지 않으면서 야생형 개방 해독 틀의 경우와 비교되게 최대치(즉 치환 가능한 코돈들 100%)로 증가하는 것이 특히 바람직하다.
- [0178] 더욱이 개방 해독 틀은 최소한 부분적으로 코돈 최적화된 것이 바람직하다. 코돈 최적화는, 번역 효율은 세포 내에 운반 RNA(tRNA)가 발생할 때 상이한 발생 빈도에 의해 결정될 수 있다는 사실에 바탕을 둔다. 그러므로 소위 "희귀 코돈(rare codon)"이 본원에 정의된 본 발명의 인공 핵산 분자의 암호화 영역 내에 다수 존재한다면, 상응하는 변형 핵산 서열의 번역은, 비교적 "자주 발생하는(frequent)" tRNA를 암호화하는 코돈이 존재하는 경우보다 그 효율이 떨어지게 된다.
- [0179] 그러므로 본 발명의 핵산 서열의 개방 해독 틀은, 세포 내에서 비교적 희귀한 tRNA를 암호화하는 야생형 서열의 코돈 하나 이상이, 세포 내에서 비교적 자주 발생하고 상기 비교적 희귀한 tRNA가 운반하는 아미노산과 동일한 아미노산을 운반하는 tRNA를 암호화하는 코돈으로 치환되도록, 상응하는 야생형 암호화 영역과 비교되게 변형되는 것이 바람직하다. 이와 같은 변형을 통해서 본원에 정의된 본 발명의 인공 핵산 분자의 개방 해독 틀은, 자주 발생하는 tRNA가 운반할 수 있는 코돈들이 희귀한 tRNA에 상응하는 코돈들을 대체할 수 있도록 변형된다. 다시 말해서 본 발명에 의하면, 이와 같은 변형을 통해, 희귀한 tRNA를 암호화하는 야생형 개방 해독 틀 코돈 전부는, 세포 내에서 더욱 자주 발생하고 상기 희귀한 tRNA가 운반하는 아미노산과 동일한 아미노산을 운반하는 tRNA를 암호화하는 코돈으로 치환될 수 있는 것이다. 어느 tRNA가 세포 내에서 비교적 자주 발생하는지와, 이와는 대조적으로 어느 tRNA가 비교적 희귀하게 발생하는지에 관하여는 당 업자에게 알려져 있다[Akashi, Curr. Opin. Genet. Dev. 2001, 11(6): 660-666]. 그러므로 개방 해독 틀은, 바람직하게 본 발명에 의한 핵산 분자가 발현될 계, 더욱 바람직하게는 본 발명에 의한 핵산 분자가 번역될 계에 대해서 코돈 최적화되는 것이 바람직하다. 바람직하게 개방 해독 틀의 코돈 선호도는 포유동물 코돈 선호도, 더욱 바람직하게는 사람 코돈 선호도에 따라서 코돈 최적화된다. 개방 해독 틀은 코돈 최적화되고 G/C 함량에 변화가 일어나는 것이 바람직하다.
- [0180] 분해 저항성, 예를 들어 엑소뉴클레아제 또는 엔도뉴클레아제에 의한 생체 내 분해 저항성을 더욱 개선하기 위하여/위하거나, 본 발명에 의한 인공 핵산 분자로부터의 단백질 발현의 안정성을 더욱 개선하기 위해서, 본 발명의 인공 핵산 분자는 변형, 예를 들어 주쇄 변형, 당 변형 및/또는 염기 변형, 예를 들어 지질 변형 등을 추가로 포함할 수 있다. 본 발명에 의한 인공 핵산 분자의 전사 및/또는 번역은 상기와 같은 변형들에 의해 거의 저해되지 않는 것이 바람직하다.
- [0181] 본 발명에서 사용될 수 있는 뉴클레오티드 유사체/변형체는, 예를 들어 2-아미노-6-클로로퓨린리보시드-5'-트리포스페이트, 2-아미노아데노신-5'-트리포스페이트, 2-티오시티딘-5'-트리포스페이트, 2-티오우리딘-5'-트리포스페이트, 4-티오우리딘-5'-트리포스페이트, 5-아미노알릴시티딘-5'-트리포스페이트, 5-아미노알릴우리딘-5'-트리포스페이트, 5-브로모시티딘-5'-트리포스페이트, 5-브로모우리딘-5'-트리포스페이트, 5-요오도시티딘-5'-트리포스페이트, 5-요오도우리딘-5'-트리포스페이트, 5-메틸시티딘-5'-트리포스페이트, 5-메틸우리딘-5'-트리포스페이트, 6-아자시티딘-5'-트리포스페이트, 6-아자우리딘-5'-트리포스페이트, 6-클로로퓨린리보시드-5'-트리포스페이트, 7-테아자아데노신-5'-트리포스페이트, 7-테아자구아노신-5'-트리포스페이트, 8-아자아데노신-5'-트리포스페이트, 8-아지도아데노신-5'-트리포스페이트, 벤지미다졸-리보시드-5'-트리포스페이트, N1-메틸아데노신-5'-트리포스페이트, N1-메틸구아노신-5'-트리포스페이트, N6-메틸아데노신-5'-트리포스페이트, O6-메틸구아노신-5'-트리포스페이트, 슈도우리딘-5'-트리포스페이트 또는 퓨로마이신-5'-트리포스페이트, 잔토신-5'-트리포스페이트로부터 선택될 수 있다. 5-메틸시티딘-5'-트리포스페이트, 7-테아자구아노신-5'-트리포스페이트, 5-브로모시티딘-5'-트리포스페이트 및 슈도우리딘-5'-트리포스페이트로 이루어진 염기 변형 뉴클레오티드들의 군으로부터 선택되는 염기 변형에 대한 뉴클레오티드가 특히 바람직하다.
- [0182] 또한 지질 변형 인공 핵산 분자는 통상 본 발명의 인공 핵산 분자와 공유 결합하고 있는 링커 하나 이상, 그리고 이 링커와 공유 결합하고 있는 지질 하나 이상을 포함할 수 있다. 대안적으로 지질 변형 인공 핵산 분자는 본원에 정의된 인공 핵산 분자 하나 이상과, 링커를 매개로 하지 않고 상기 인공 핵산 분자와 공유 결합하고 있는 2 기능성 지질 하나 이상을 포함할 수 있다. 제3 대안에 의하면, 지질 변형 인공 핵산 분자는 본원에 정의된 인공 핵산 분자, 이 인공 핵산 분자와 공유 결합하고 있는 링커 하나 이상, 이 링커와 공유 결합하고 있는 지질 하나 이상, 그리고 추가로, 바람직하게는 링커를 매개로 하지 않고 상기 인공 핵산 분자와 공유 결합하고 있는 2 기능성 지질 하나 이상을 포함할 수 있다.
- [0183] 추가의 양태에서, 본 발명은 다음과 같은 요소들을 포함하는 벡터를 제공한다:

- [0184] a. 개방 해독 틀(ORF) 및/또는, 예를 들어 개방 해독 틀 또는 개방 해독 틀을 포함하는 서열을 삽입하기 위한 클로닝 위치; 및
- [0185] b. 알부민 유전자의 3'-UTR로부터 유래하거나 또는 알부민 유전자의 3'-UTR의 변이체로부터 유래하는 핵산 서열을 포함하는 3'-비번역 영역 요소(3'-UTR 요소) 하나 이상.
- [0186] 3'-UTR 요소 하나 이상과 ORF는 본 발명에 의한 인공 핵산 분자에 대해 전술된 바와 같다. 클로닝 위치는 개방 해독 틀 또는 개방 해독 틀을 포함하는 서열을 도입하는데 적당한 임의의 서열, 예를 들어 하나 이상의 제한 위치일 수 있다. 그러므로 클로닝 위치를 포함하는 벡터는 개방 해독 틀을 벡터에 삽입하기 적당한 것, 바람직하게는 개방 해독 틀(5')을 3'-UTR 요소에 삽입하기 적당한 것이 바람직하다. 클로닝 위치 또는 ORF는 3'-UTR 요소에 대해 5' 말단 쪽, 바람직하게는 3'-UTR 요소의 5'-말단에 인접한 곳에 위치하는 것이 바람직하다. 예를 들어 클로닝 위치 또는 ORF는 3'-UTR 요소의 5'-말단에 직접 결합할 수 있거나, 아니면 뉴클레오티드 구획, 예를 들어 본 발명에 의한 인공 핵산 분자에 대해 전술된 바와 같은 2, 4, 6, 8, 10 또는 20개 등의 뉴클레오티드들로 이루어진 구획을 통하여 3'-UTR 요소의 5'-말단에 결합할 수 있다.
- [0187] 바람직하게 본 발명에 의한 벡터는, 예를 들어 임의로 개방 해독 틀 또는 이 개방 해독 틀을 포함하는 서열을 벡터에 삽입한 다음, 이 벡터를 전사시킴으로써 본 발명에 의한 인공 핵산 분자, 바람직하게는 본 발명에 의한 인공 mRNA를 생산하는데 적당하다. 그러므로 벡터는 전사에 필요한 요소들, 예를 들어 프로모터, 예를 들어 RNA 중합 효소 프로모터를 포함하는 것이 바람직하다. 벡터는 진핵생물, 원핵생물, 바이러스 또는 파지 전사 계, 예를 들어 진핵생물 세포, 원핵생물 세포, 또는 진핵생물, 원핵생물, 바이러스 또는 파지 시험관 내 전사 계가 사용되는 전사한 것이 바람직하다. 따라서, 예를 들어 벡터는 중합 효소, 예를 들어 RNA 중합 효소, 예를 들어 진핵생물, 원핵생물, 바이러스 또는 파지 RNA 중합 효소에 의해 인지되는 프로모터 서열을 포함할 수 있다. 바람직한 구체예에서, 벡터는 파지 RNA 중합 효소 프로모터, 예를 들어 SP6 또는 T7, 바람직하게는 T7 프로모터를 포함한다. 바람직하게 벡터는 파지 기반 시험관 내 전사 계, 예를 들어 T7 RNA 중합 효소 기반 시험관 내 전사 계가 사용되는 시험관 내 전사에 적당하다.
- [0188] 벡터는 본 발명에 의한 인공 핵산 분자에 대해 전술된 바와 같이 폴리(A) 서열 및/또는 폴리아데닐화 신호를 추가로 포함할 수 있다.
- [0189] 벡터는 RNA 벡터 또는 DNA 벡터일 수 있다. 벡터는 DNA 벡터가 바람직하다. 벡터는 당 업자에게 알려진 임의의 벡터, 예를 들어 바이러스 벡터 또는 플라스미드 벡터일 수 있다. 바람직하게 벡터는 플라스미드 벡터, 바람직하게는 DNA 플라스미드 벡터이다.
- [0190] 바람직한 구체예에서, 본 발명에 의한 벡터는 본 발명에 의한 인공 핵산 분자를 포함한다.
- [0191] 바람직하게 본 발명에 의한 DNA 벡터는 서열 번호 1, 서열 번호 3, 서열 번호 32, 서열 번호 34, 서열 번호 36 또는 서열 번호 38에 의한 서열이나, 또는 이 서열 번호 1, 서열 번호 3, 서열 번호 32, 서열 번호 34, 서열 번호 36 또는 서열 번호 38에 의한 핵산 서열에 대한 서열 동일성이 약 40% 이상, 바람직하게는 약 50% 이상, 바람직하게는 약 60% 이상, 바람직하게는 약 70% 이상, 더욱 바람직하게는 약 80% 이상, 더욱 바람직하게는 약 90% 이상, 더더욱 바람직하게는 약 95% 이상, 더더욱 바람직하게는 약 99% 이상인 서열이나, 또는 전술된 바와 같은 이것들의 단편, 바람직하게는 이것들의 기능성 단편을 포함한다.
- [0192] 바람직하게 본 발명에 의한 RNA 벡터는 서열 번호 2, 서열 번호 4, 서열 번호 33, 서열 번호 35, 서열 번호 37 또는 서열 번호 39에 의한 서열이나, 또는 이 서열 번호 2, 서열 번호 4, 서열 번호 33, 서열 번호 35, 서열 번호 37 또는 서열 번호 39에 의한 핵산 서열에 대한 서열 동일성이 약 40% 이상, 바람직하게는 약 50% 이상, 바람직하게는 약 60% 이상, 바람직하게는 약 70% 이상, 더욱 바람직하게는 약 80% 이상, 더욱 바람직하게는 약 90% 이상, 더더욱 바람직하게는 약 95% 이상, 더더욱 바람직하게는 약 99% 이상인 서열이나, 또는 이것들의 단편, 바람직하게는 이것들의 기능성 단편을 포함한다.
- [0193] 바람직하게 벡터는 원형 분자이다. 바람직하게 벡터는 이중 사슬 분자, 예를 들어 이중 사슬 DNA 분자이다. 이와 같이 원형이면서, 바람직하게는 이중 사슬인 DNA 분자는 본 발명의 인공 핵산 분자에 대한 저장 형태로서 편리하게 사용될 수 있다. 뿐만 아니라 벡터는 세포, 예를 들어 배양된 세포의 형질 감염에 사용될 수 있다. 또한 벡터는 본 발명에 의한 인공 RNA 분자를 생산하기 위한 시험관 내 전사에 사용될 수도 있다.
- [0194] 바람직하게 벡터, 바람직하게 원형 벡터는, 예를 들어 제한 효소 분해를 통해 선형화된다. 바람직한 구체예에서, 벡터는, 3'-UTR 요소에 대해 3'-말단 쪽에 바로 붙어있거나, 또는 (만일 존재한다면) 폴리(A) 서열

또는 폴리아데닐화 신호에 대해 3'-말단 쪽에 위치하거나, 또는 (만일 존재한다면) 폴리(C) 서열에 대해 3'-말단 쪽에 위치하거나, 또는 (만일 존재한다면) 히스톤 스템-루프에 대해 3'-말단 쪽에 위치하는 절단 위치, 예를 들어 제한 위치, 바람직하게는 특유의(unique) 절단 위치를 포함한다. 그러므로, 벡터를 선형화하여 제조된 생산물의 3'-말단은, 3'-UTR 요소의 3'-말단으로 종결되거나, 또는 (만일 존재한다면) 폴리(A) 서열 또는 폴리아데닐화 신호의 3'-말단으로 종결되거나, 또는 (만일 존재한다면) 폴리(C) 서열의 3'-말단으로 종결된다. 본 발명에 의한 벡터가 본 발명에 의한 인공 핵산 분자를 포함하는 구체예에서, 제한 위치, 바람직하게는 특유의 제한 위치는 인공 핵산 분자의 3'-말단에 대해 3'-말단 쪽에 바로 붙어 있다.

[0195] 추가의 양태에서, 본 발명은, 본 발명에 의한 인공 핵산 분자 또는 본 발명에 의한 벡터를 포함하는 세포에 관한 것이다. 세포는 임의의 세포, 예를 들어 박테리아 세포, 곤충 세포, 식물 세포, 척추 동물 세포, 예를 들어 포유동물 세포일 수 있다. 이와 같은 세포는, 예를 들어 박테리아 세포 내에서, 예를 들어 본 발명의 벡터를 복제하는데 사용될 수 있다. 더욱이 세포는 본 발명에 의한 인공 핵산 분자 또는 벡터를 전사하고/전사하거나 본 발명에 의한 인공 핵산 분자 또는 벡터의 개방 해독 틀을 번역하는데 사용될 수 있다. 예를 들어 세포는 재조합 단백질 생산에 사용될 수 있다.

[0196] 본 발명에 의한 세포는, 예를 들어 표준 핵산 운반 방법, 예를 들어 표준 형질 감염, 형질 도입 또는 형질 전환 방법을 통해 생산될 수 있다. 예를 들어 본 발명에 의한 인공 핵산 분자 또는 벡터는 전기 천공법, 예를 들어 양이온 지질 및/또는 리포솜을 기반으로 하는 리포펙션, 칼슘 인산염 침전법, 나노입자 기반 형질 감염, 바이러스 기반 형질 감염 또는 양이온 중합체, 예를 들어 DEAE-덱스트란 또는 폴리에틸렌이민 기반 방법을 통해 세포에 운반될 수 있다.

[0197] 세포는 포유동물 세포, 예를 들어 사람 개체의 세포, 가축의 세포, 실험실 동물의 세포, 예를 들어 마우스나 래트의 세포인 것이 바람직하다. 세포는 사람 세포인 것이 바람직하다. 세포는 확립된 세포주의 세포, 예를 들어 CHO, BHK, 293T, COS-7, HELA, HEK 등일 수 있거나, 또는 세포는 1차 세포, 예를 들어 HDF 세포 등, 바람직하게는 유기체로부터 분리된 세포일 수 있다. 바람직한 구체예에서, 세포는 포유동물 개체, 바람직하게는 사람 개체의 분리된 세포이다. 예를 들어 세포는 바람직하게 포유동물 개체, 바람직하게는 사람 개체의 면역 세포, 예를 들어 수지상 세포, 암 또는 종양 세포, 또는 임의의 체세포 등일 수 있다.

[0198] 추가의 양태에서, 본 발명은, 본 발명에 의한 인공 핵산 분자, 본 발명에 의한 벡터 또는 본 발명에 의한 세포를 포함하는 약학 조성물을 제공한다. 본 발명에 의한 약학 조성물은, 예를 들어 백신, 예를 들어 유전자 백신 화용 백신으로서 사용될 수 있다. 그러므로 ORF는, 예를 들어 백신화를 위해 환자에게 투여될 항원을 암호화할 수 있다. 그러므로 바람직한 구체예에서, 본 발명에 의한 약학 조성물은 백신이다. 뿐만 아니라, 본 발명에 의한 약학 조성물은, 예를 들어 유전자 치료법에 사용될 수 있다.

[0199] 바람직하게 본 발명의 약학 조성물은 하나 이상의 약학적으로 허용 가능한 부형제, 비이클, 충전제 및/또는 희석제를 추가로 포함한다. 본 발명의 내용에 있어서, 약학적으로 허용 가능한 비이클은 통상 본 발명의 약학 조성물에 대한 액체 또는 비 액체 베이스를 포함한다. 하나의 구체예에서, 본 발명의 약학 조성물은 액체의 형태로 제공된다. 이와 같은 내용에 있어서, 비이클은 물, 예를 들어 무발연원 물, 등장성 염수 또는 완충한 (수성) 용액, 예를 들어 인산염 및 시트르산염 등으로 완충한 용액을 기반으로 한다. 완충액은 특정 기준 매질과 관련하여 고장성, 등장성 또는 저장성일 수 있는데, 다시 말해서, 완충액은 특정 기준 매질과 관련하여 염의 함량이 더욱 높을 수 있거나, 동일할 수 있거나, 낮을 수 있는데, 여기서, 전술된 염의 농도, 즉 삼투 작용이나 기타 농도 효과로 인해 포유동물 세포를 손상시키지 않는 농도가 적용될 수 있는 것이 바람직하다. 기준 매질은, 예를 들어 "생체 내" 방법에서 사용되는 액체, 예를 들어 혈액, 림프액, 세포 기질액 또는 기타 체액이거나, 또는, 예를 들어 "시험관 내" 방법에서 기준 매질로서 사용될 수 있는 액체, 예를 들어 통상의 완충액이나 액체이다. 이와 같은 통상의 완충액 또는 액체는 당 업자에게 알려져 있다. 링거-젯산 염 용액이 액체 베이스로서 특히 바람직하다.

[0200] 환자에 투여되기 적당한 혼화성 고체 또는 액체 충전제나 희석제 또는 피포화 화합물(encapsulating compound) 하나 이상도 또한 본 발명의 약학 조성물에 사용될 수 있다. 본원에 사용된 용어 "혼화성"이란, 바람직하게 본 발명의 약학 조성물의 구성 성분들이, 통상의 사용 조건 하에서 본 발명의 약학 조성물의 약학 효능을 실질적으로 떨어뜨리게 될 상호 작용이 일어나지 않는 양상으로, 본원에 정의된 본 발명의 핵산, 벡터 또는 세포와 혼합될 수 있는 경우를 의미한다.

[0201] 임의로, 본 발명에 의한 약학 조성물은 추가의 약학 활성 성분들 하나 이상을 추가로 포함할 수도 있다. 이러한 내용에 있어서 약학 활성 성분은 특정 징후나 질병을 치유, 완화 또는 예방하는 치료 효과를 발휘하는 화합물이

다. 이러한 화합물로서는 펩티드 또는 단백질, 핵산, (치료 활성) 저 분자량 유기 또는 무기 화합물(분자량 5000 미만, 바람직하게는 1000 미만), 당, 항원 또는 항체, 이미 선행 기술로서 공지되어 있는 치료제, 항원성 세포(antigenic cell), 항원성 세포 단편, 세포 분획, 세포 벽 구성 성분(예를 들어 다당체), (예를 들어 화학적으로나 조사에 의해) 변형, 약독화 또는 불활성화된 병원체(바이러스, 박테리아 등)를 포함하나, 여기에는 어떠한 제한이 내포되어 있는 것은 아니다.

[0202] 더욱이 본 발명의 약학 조성물은 인공 핵산 분자 또는 벡터용 캐리어를 포함할 수 있다. 이러한 캐리어는 생리적으로 허용 가능한 액체 중 용해, 약학 활성 인공 핵산 분자 또는 벡터의 이동과 세포 내 흡수를 매개하는데 적당할 수 있다. 그러므로 이러한 캐리어는 본 발명에 의한 인공 핵산 분자 또는 벡터의 저장 및 전달에 적당할 수 있는 성분일 수 있다. 이러한 성분들은, 예를 들어 형질 감염 제제 또는 복합체화 제제로서 사용될 수 있는 양이온 또는 다가 양이온 캐리어 또는 화합물일 수 있다.

[0203] 이와 같은 내용에 있어서 특히 바람직한 형질 감염 제제 또는 복합체화 제제는, 양이온 또는 다가 양이온 화합물, 예를 들어 프로타민, 뉴클레올린, 스퍼민 또는 스퍼미딘이나, 또는 기타 양이온 펩티드 또는 단백질, 예를 들어 폴리-L-리신(PLL), 폴리-아르기닌, 염기성 폴리펩티드, 세포 투과성 펩티드(CPP), 예를 들어 HIV-결합 펩티드, HIV-1 Tat(HIV), Tat-유래 펩티드, 페넨트라틴, VP22 유래 펩티드 또는 VP22 유사체 펩티드, HSV VP22 (단순 포진 바이러스), MAP, KALA 또는 단백질 형질 도입 도메인(PTD), PpT620, 프롤린-풍부 펩티드, 아르기닌-풍부 펩티드, 리신-풍부 펩티드, MPG-펩티드(들), Pep-1, L-올리고머, 칼시토닌 펩티드(들), 안테나페디아 유래 펩티드(특히 드로스필라 안테나페디아(*Drosophila antennapedia*) 유래 펩티드), pAntp, pIsl, FGF, 락토펜린, 트랜스포탄, 버포린-2, Bac715-24, SynB, SynB(1), pVEC, hCT-유래 펩티드, SAP 또는 히스톤이다.

[0204] 뿐만 아니라, 이러한 양이온 또는 다가 양이온 화합물 또는 캐리어는, 바람직하게 -SH부를 하나 이상 포함하거나 이것을 포함하도록 추가로 변형된 양이온 또는 다가 양이온 펩티드 또는 단백질일 수 있다. 바람직하게 양이온 또는 다가 양이온 캐리어는 다음과 같은 화학식(I)을 가지는 양이온 펩티드들로부터 선택된다:

[0205] [화학식 I]

[0206]  $\{(Arg)_1;(Lys)_m;(His)_n;(Orn)_o;(Xaa)_x\};$

[0207] 상기 식 중,

[0208]  $l + m + n + o + x = 3\sim 100$ 으로서,  $l, m, n$  또는  $o$ 는 서로 독립적으로 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21~30, 31~40, 41~50, 51~60, 61~70, 71~80, 81~90 및 91~100으로부터 선택되는 임의의 수이되, 다만 Arg(아르기닌), Lys(리신), His(히스티딘) 및 Orn(오르니틴)의 총 함량은 올리고펩티드를 구성하는 모든 아미노산의 10% 이상에 해당하고; Xaa는 Arg, Lys, His 또는 Orn을 제외한 천연(자연 발생) 또는 비 천연 아미노산으로부터 선택되는 임의의 아미노산이며;  $x$ 는 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21~30, 31~40, 41~50, 51~60, 61~70, 71~80, 81~90으로부터 선택되는 임의의 수이되, 다만 Xaa의 총 함량은 올리고펩티드를 구성하는 모든 아미노산의 90%를 넘지 않는다. 아미노산 Arg, Lys, His, Orn 및 Xaa 중 임의의 것은 펩티드의 임의의 위치에 배치될 수 있다. 이와 같은 내용에서, 7~30개 범위의 아미노산으로 이루어진 양이온 펩티드 또는 단백질이 특히 바람직하다.

[0209] 뿐만 아니라 -SH부를 하나 이상 포함하거나 이 -SH부를 하나 이상 포함하도록 추가로 변형된 양이온 또는 다가 양이온 펩티드 또는 단백질이 상기 보인 화학식  $\{(Arg)_1;(Lys)_m;(His)_n;(Orn)_o;(Xaa)_x\}$ (화학식 I)에 의해서 정의될 때, 이 양이온 또는 다가 양이온 펩티드 또는 단백질은, 다음과 같은 하위 화학식 Ia의 펩티드 또는 단백질로부터 선택될 수 있으나, 이에 한정되는 것은 아니다:

[0210] [화학식 Ia]

[0211]  $\{(Arg)_1;(Lys)_m;(His)_n;(Orn)_o;(Xaa')_x(Cys)_y\}$

[0212] 상기 식 중,

[0213]  $(Arg)_1;(Lys)_m;(His)_n;(Orn)_o$ ; 및  $x$ 는 본원에 정의된 바와 같고, Xaa'는 Arg, Lys, His, Orn 또는 Cys을 제외한 천연(자연 발생) 또는 비 천연 아미노산으로부터 선택되는 임의의 아미노산이며,  $y$ 는 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21~30, 31~40, 41~50, 51~60, 61~70, 71~80 및 81~90로부터 선택되는 임의의 수이되, 다만 Arg(아르기닌), Lys(리신), His(히스티딘) 및 Orn(오르니틴)의 총 함량은 올리고펩티드를 구성하는 모든 아미노산의 10% 이상에 해당한다. 뿐만 아니라 양이온 또는 다가 양이온 펩티드

는 하기 하위 화학식 Ib의 펩티드로부터 선택될 수 있다:

- [0214] [화학식 Ib]
- [0215]  $Cys_1 \{(Arg)_1;(Lys)_m;(His)_n;(Orn)_o;(Xaa)_x\} Cys_2$
- [0216] 실험식  $\{(Arg)_1;(Lys)_m;(His)_n;(Orn)_o;(Xaa)_x\}$ (화학식 III)는 본원에 정의된 바와 같고, (반 실험) 화학식 III에 의한 아미노산 서열의 코어를 형성하며,  $Cys_1$  및  $Cys_2$ 는  $(Arg)_1;(Lys)_m;(His)_n;(Orn)_o;(Xaa)_x$ 가 있는 쪽으로 되도록 가까이 존재하거나 이의 양 말단을 이루는 시스템이다.
- [0217] 형질 감염 제제 또는 복합체화 제제로서 사용될 수 있는 추가의 바람직한 양이온 또는 다가 양이온 화합물로서는, 양이온 다당체, 예를 들어 키토산, 폴리브렌, 양이온 중합체, 예를 들어 폴리에틸렌이민(PEI), 양이온 지질, 예를 들어 DOTMA: [1-(2,3-시올레일옥시)프로필]-N,N,N-트리메틸암모늄 염화물, DMRIE, 디-C14-아미딘, DOTIM, SAINT, DC-Chol, BGTC, CTAP, DOPC, DODAP, DOPE: 디올레일 포스파티딜에탄올아민, DOSPA, DODAB, DOIC, DMEPC, DOGS: 디옥타데실아미도글리실스퍼민, DIMRI: 디미리스토옥시프로필 디메틸 하이드록시에틸 암모늄 브롬화물, DOTAP: 디올레오일옥시-3-(트리메틸암모니오)프로판, DC-6-14: 0,0-디테트라데카노일-N-( $\alpha$ -트리메틸암모니오아세틸)디에탄올아민 염화물, CLIP1: rac-[(2,3-디옥타데실옥시프로필)(2-하이드록시에틸)]-디메틸 암모늄 염화물, CLIP6: rac-[2(2,3-디헥사데실옥시프로필옥시메틸옥시)에틸]트리메틸암모늄, CLIP9: rac-[2(2,3-디헥사데실옥시프로필옥시숙시닐옥시)에틸]-트리메틸암모늄, 올리고펩타민이나, 양이온 또는 다가 양이온 중합체, 예를 들어 변형된 폴리아미노산, 예를 들어  $\beta$ -아미노산-중합체 또는 역전된 폴리아미드 등, 변형된 폴리에틸렌, 예를 들어 PVP (폴리(N-에틸-4-비닐피리디늄 브롬화물)) 등, 변형된 아크릴산염, 예를 들어 pDMAEMA (폴리(디메틸아미노에틸 메타크릴산염)) 등, 변형된 아미도아민, 예를 들어 pAMAM (폴리(아미도아민)) 등, 변형된 폴리베타아미노에스테르 (PBAE), 예를 들어 디아민 말단 변형된 1,4-부탄디올 디아크릴산염-코-5-아미노-1-펜탄올 중합체 등, 덴드리머, 예를 들어 폴리프로필아민 덴드리머 또는 pAMAM 기반 덴드리머 등, 폴리이민(들), 예를 들어 PEI: 폴리(에틸렌이민), 폴리(프로필렌이민) 등, 폴리알릴아민, 당 주쇄 기반 중합체, 예를 들어 사이클로덱스트린 기반 중합체, 텍스트란 기반 중합체, 키토산 등, 실란 주쇄 기반 중합체, 예를 들어 PMOXA-PDMS 공중합체 등, 하나 이상의 양이온 블록(예를 들어 전술된 바와 같은 양이온 중합체로부터 선택되는 양이온 블록)과 하나 이상의 친수성 또는 소수성 블록(예를 들어 폴리에틸렌글리콜)의 조합으로 이루어진 블록 중합체 등을 포함할 수 있다.
- [0218] 이러한 내용에 있어서, 본 발명의 인공 핵산 분자 또는 본 발명의 벡터는 적어도 부분적으로 양이온 또는 다가 양이온 화합물, 바람직하게는 양이온 단백질 또는 펩티드와 복합체를 형성하는 것이 특히 바람직하다. "부분적으로"란, 본 발명의 인공 핵산 분자 또는 본 발명의 벡터의 일부만이 양이온 또는 다가 양이온 화합물과 복합체를 형성하고, 본 발명의 인공 핵산 분자 또는 본 발명의 벡터의 나머지는 복합체를 형성하지 않은 형태("자유형")를 하고 있는 경우를 의미한다. 바람직하게 복합체를 형성한 핵산 대 자유형 핵산의 비율은 약 5:1(w/w)~약 1:10(w/w), 더욱 바람직하게 약 4:1(w/w)~약 1:8(w/w), 더더욱 바람직하게 약 3:1(w/w)~약 1:5(w/w) 또는 1:3(w/w)으로부터 선택되고, 가장 바람직하게 복합체를 형성한 핵산 대 자유형 핵산의 비율은 약 1:1(w/w)의 비율로부터 선택된다.
- [0219] 본 발명에 의한 약학 조성물은 임의로는 보조제, 예를 들어 선천성 면역계를 자극하거나 인공 핵산 분자 또는 벡터의 세포 내 흡수를 촉진하는 보조제 하나 이상을 추가로 포함할 수도 있다. 이와 같은 내용에 있어서, 보조제는 선천성 면역계의 면역 반응, 즉 비특이적 면역 반응을 개시 또는 증강하는데 적당한 임의의 화합물인 것으로 이해될 수 있다. 다시 말해서, 본 발명의 약학 조성물이 투여될 때 상기 조성물은, 임의로 이 조성물 중에 포함된 상기 보조제로 말미암아 선천성 면역 반응을 유도하는 것이 바람직하다. 바람직하게 이와 같은 보조제는 포유동물에서 선천성 면역 반응의 유도를 뒷받침하는 보조제일 수 있다. 이러한 보조제는, 예를 들어 면역 자극성 핵산, 즉 톨-유사 수용기 등에 결합할 수 있는 핵산, 바람직하게는 면역 자극성 RNA 등일 수 있다.
- [0220] 이러한 보조제, 바람직하게는 이러한 면역 자극성 핵산은 본 발명의 약학 조성물, 바람직하게는 백신의 인공 핵산 분자에 의해 암호화되는 펩티드 또는 단백질, 즉 항원에 대한 특이적(즉 적응성) 면역 반응을 뒷받침할 수 있는 선천성(즉 비특이적) 면역 반응을 유도할 수 있다.
- [0221] 본 발명의 약학 조성물은 또한 추가로 임의의 추가 화합물(화합물 자체의 사람 톨-유사 수용기들 TLR1, TLR2, TLR3, TLR4, TLR5, TLR6, TLR7, TLR8, TLR9, TLR10에 대한 (리간드로서의) 결합 친화성, 또는 쥐과 동물 톨-유사 수용기들 TLR1, TLR2, TLR3, TLR4, TLR5, TLR6, TLR7, TLR8, TLR9, TLR10, TLR11, TLR12 또는 TLR13에 대한 (리간드로서의) 결합 친화성으로 말미암아 면역 자극성을 가진다고 알려져 있는 화합물)을 포함할 수도

있다.

- [0222] 본 발명의 약학 조성물 중에 포함될 수 있는 추가의 첨가제로서는, 예를 들어 유화제, 예를 들어 트윈(Tween)®; 습윤제, 예를 들어 소듐 라우릴 황산염; 착색제; 풍미 부여제, 약학 캐리어; 정제 형성 제제; 안정화제; 항산화제; 보존제 등이 있다.
- [0223] 본 발명에 의한 약학 조성물은 바람직하게 약학 조성물의 성분들, 특히 본원에 정의된 바와 같은 본 발명의 핵산 서열, 벡터 및/또는 세포를 "안전량 및 유효량"만큼 포함한다. 본원에 사용된 "안전량 및 유효량"이란, 본원에 정의된 질병 또는 질환의 긍정적 개선을 유의적으로 유도하는데 충분한 양을 의미한다. 그러나, 이와 동시에 "안전량 및 유효량"은 종종 부작용이 발생하는 것을 막고, 이점과 위험성 간 분별 가능한 관계를 허용한다. 이러한 한계들은 통상적으로 현명한 의료상 판단 범위 내에서 결정된다.
- [0224] 추가의 양태에서, 본 발명은 유전자 치료법 또는 (유전자 백신화에서) 의약품, 예를 들어 백신으로서 사용될, 본 발명에 의한 인공 핵산 분자, 본 발명에 의한 벡터, 본 발명에 의한 세포 또는 본 발명에 의한 약학 조성물을 제공한다.
- [0225] 본 발명에 의한 인공 핵산 분자, 본 발명에 의한 벡터, 본 발명에 의한 세포 또는 본 발명에 의한 약학 조성물은 펩티드, 폴리펩티드 또는 단백질의 치료 작용이나 치료 효과를 이용하는 임의의 의학 분야에서, 또는 특정 펩티드 또는 단백질의 보충이 필요한 경우에 특히 적당하다. 그러므로 본 발명은, 펩티드, 폴리펩티드 또는 단백질의 치료 작용 또는 치료 효과를 바탕으로 한 치료에 순응성이거나, 또는 특정 펩티드, 폴리펩티드 또는 단백질의 보충을 바탕으로 한 치료에 순응성인 질병 또는 질환의 치료 또는 예방에 사용되는, 본 발명에 의한 인공 핵산 분자, 본 발명에 의한 벡터, 본 발명에 의한 세포 또는 본 발명에 의한 약학 조성물을 제공한다. 예를 들어 본 발명에 의한 인공 핵산 분자, 본 발명에 의한 벡터, 본 발명에 의한 세포 또는 본 발명에 의한 약학 조성물은 유전병, 자가 면역 질환, 암 또는 종양 관련 질병, 감염성 질병 또는 만성 질병 등의, 예를 들어 유전자 백신화 또는 유전자 치료법에 의한 치료 또는 예방에 사용될 수 있다.
- [0226] 특히 치료 대상인 개체에서 치료 펩티드, 폴리펩티드 또는 단백질이 안정적으로 존재하고 그 존재 기간이 연장됨으로 인해 이익을 얻는, 이와 같은 치료적 처치들은 본 발명의 내용에 있어서 의료 수단으로서 사용되기에 특히 적당한다, 그 이유는 본 발명의 3'-UTR 요소가 본 발명의 핵산 분자의 ORF 발현을 안정화 및 연장해 주기 때문이다. 그러므로 본 발명에 의한 인공 핵산 분자, 본 발명에 의한 벡터, 본 발명에 의한 세포 또는 본 발명에 의한 약학 조성물에 특히 적당한 의료적 사용 방법은 백신화이다. 그러므로 본 발명은, 개체, 바람직하게는 포유동물 개체, 더욱 바람직하게는 사람 개체의 백신화를 위한, 본 발명에 의한 인공 핵산 분자, 본 발명에 의한 벡터, 본 발명에 의한 세포 또는 본 발명에 의한 약학 조성물을 제공한다. 바람직한 백신화 치료법으로서는 감염성 질병, 예를 들어 박테리아, 원생 동물 또는 바이러스 감염에 대한 백신화와 항종양 백신화가 있다. 이러한 백신화 치료법은 예방용 또는 치료용일 수 있다.
- [0227] 치료 또는 예방될 질병에 따라서 ORF가 선택될 수 있다. 예를 들어 개방 해독 틀은 단백질의 기능이 완전히 상실되었거나 적어도 부분적으로 상실된 환자, 예를 들어 유전병을 앓고 있는 환자에게 공급되어야 하는 단백질을 암호화할 수 있다. 뿐만 아니라 개방 해독 틀은 개체의 질병이나 병태에 유익한 영향을 미치는 펩티드 또는 단백질을 암호화하는 ORF로부터 선택될 수 있다. 더욱이 개방 해독 틀은 자연적 펩티드 또는 단백질의 병리적 과생산을 하향 조절하거나 병리적으로 단백질 또는 펩티드를 발현하는 세포를 제거하는 펩티드 또는 단백질을 암호화할 수 있다. 이와 같이 기능의 결실, 상실 또는 과 생산은, 예를 들어 종양 및 신생 물질 생성, 자가 면역 질환, 알레르기, 감염 또는 만성 질병 등에서 일어날 수 있는 현상이다. 뿐만 아니라 개방 해독 틀은 항원 또는 면역원, 예를 들어 병원체의 에피토프 또는 종양 항원을 암호화할 수도 있다. 그러므로 바람직한 구체예에서, 본 발명에 의한 인공 핵산 분자 또는 벡터는 항원 또는 면역원, 예를 들어 병원체의 에피토프 또는 종양 연관 항원, 전술된 3'-UTR 요소, 그리고 임의의 추가 구성 성분들, 예를 들어 폴리(A) 서열 등을 포함하거나 이것으로 이루어진 아미노산 서열을 암호화하는 ORF를 포함한다.
- [0228] 의료적 사용 방법, 특히 백신화에 있어서, 본 발명에 의한 인공 핵산 분자는 RNA, 바람직하게는 mRNA인 것이 바람직한다, 그 이유는 DNA는 항DNA 면역 반응을 유도할 위험을 보유하고 있으면서 게놈 DNA 내에 삽입되는 경향이 있기 때문이다. 그러나 몇몇 구체예에 의하면, 예를 들어 유전자 치료 처치법에서, 예를 들어 바이러스 전달 비이클, 예를 들어 아데노바이러스 전달 비이클이 본 발명에 의한 인공 핵산 분자 또는 벡터의 전달에 사용되면, 상기 인공 핵산 분자 또는 벡터는 DNA 분자인 것이 요망될 수 있다.
- [0229] 본 발명에 의한 인공 핵산 분자, 본 발명에 의한 벡터, 본 발명에 의한 세포 또는 본 발명에 의한 약학 조성물

은 경구, 비 경구, 흡입 스프레이, 국소, 직장, 비강, 협측, 질 내 투여될 수 있거나 또는 이식된 저장체 (reservoir)를 통해 투여될 수 있다. 본원에 사용된 용어 "비 경구 투여"에는 피하, 정맥 내, 근육 내, 관절 내, 활액 내, 흉골 내, 초 내, 간 내, 병변 내, 두개골 내, 경 피, 피 내, 폐 내, 복막 내, 심장 내, 동맥 내 및 설 하 주사 또는 주입 기법을 포함한다.

- [0230] 바람직하게 본 발명에 의한 인공 핵산 분자, 본 발명에 의한 벡터, 본 발명에 의한 세포 또는 본 발명에 의한 약학 조성물은 비 경구 투여, 예를 들어 비 경구 주사, 더욱 바람직하게는 피하, 정맥 내, 근육 내, 관절 내, 활액 내, 흉골 내, 초 내, 간 내, 병변 내, 두개골 내, 경 피, 피 내, 폐 내, 복막 내, 심장 내, 동맥 내 및 설 하 주사 또는 주입 기법을 통하여 투여된다. 피 내 및 근육 내 주사가 특히 바람직하다. 본 발명의 약학 조성물의 멸균 주사 제형은 수성 현탁액 또는 유질 현탁액일 수 있다. 이러한 현탁액은 당 업계에 알려진 기법에 따라서 적당한 분산제 또는 습윤제 및 현탁제를 사용하여 제형될 수 있다.
- [0231] 본 발명에 의한 인공 핵산 분자, 본 발명에 의한 벡터, 본 발명에 의한 세포 또는 본 발명에 의한 약학 조성물은 또한 경구적으로 허용 가능한 투여형, 예를 들어 캡슐, 정제, 수성 현탁액 또는 용액(이에 한정되는 것은 아님)으로 경구 투여될 수도 있다.
- [0232] 본 발명에 의한 인공 핵산 분자, 본 발명에 의한 벡터, 본 발명에 의한 세포 또는 본 발명에 의한 약학 조성물은 또한, 특히 치료 대상이 국소 적용에 의해 용이하게 접근 가능한 부위이거나 장기를 포함할 때(예를 들어 치료 대상이 피부병 또는 기타 접근 가능한 임의의 상피 조직의 질병일 때) 국소 투여될 수도 있다. 적당한 국소 투여용 제형들은 이와 같은 부위들 또는 장기들 각각에 사용될 목적으로 용이하게 제조된다. 국소 적용을 위해, 본 발명에 의한 인공 핵산 분자, 본 발명에 의한 벡터, 본 발명에 의한 세포 또는 본 발명에 의한 약학 조성물은 하나 이상의 캐리어 중에 현탁 또는 용해된 적당한 연고로서 제형될 수 있다.
- [0233] 하나의 구체예에서, 의약품으로서의 사용 방법은, 포유동물 세포를 형질 감염시키는 단계, 바람직하게는 포유동물 세포를 시험관 내 형질 감염시키는 단계, 더욱 바람직하게는 의약품으로 치료될 개체의 분리된 세포들을 시험관 내 형질 감염시키는 단계를 포함한다. 만일 상기 사용 방법이 분리된 세포를 시험관 내 형질 감염시키는 단계를 포함하면, 의약품으로서의 사용 방법은 또한 형질 감염된 세포들을 환자에 (재)투여하는 단계를 추가로 포함할 수 있다. 본 발명의 인공 핵산 분자 또는 벡터의 의약품으로서의 사용 방법은 연이어 형질 감염된 분리 세포들을 선택하는 단계를 추가로 포함할 수 있다. 그러므로, 본 발명의 인공 핵산 분자 또는 벡터의 의약품으로서의 사용 방법은 만일 벡터가 선별 마커를 추가로 포함하면 유리할 수 있다. 또한 의약품으로서의 사용 방법은, 분리된 세포를 시험관 내 형질 감염시키는 단계와, 이러한 세포들로부터 발현 생산물, 즉 암호화된 펩티드 또는 단백질을 정제하는 단계를 포함할 수 있다. 이처럼 정제된 펩티드 또는 단백질은 추후 이의 투여를 필요로 하는 개체에 투여될 수 있다.
- [0234] 본 발명은 또한, 본 발명에 의한 인공 핵산 분자, 본 발명에 의한 벡터, 본 발명에 의한 세포 또는 본 발명에 의한 약학 조성물을 이것들의 투여를 필요로 하는 개체에 투여하는 단계를 포함하는, 전술된 바와 같은 질병 또는 질환을 치료 또는 예방하는 방법도 제공한다.
- [0235] 뿐만 아니라 본 발명은, 본 발명에 의한 인공 핵산 분자 또는 본 발명에 의한 벡터로 세포를 형질 감염시키는 단계를 포함하는, 질병 또는 질환을 치료 또는 예방하는 방법을 제공한다. 상기 형질 감염은 시험관 내 또는 생체 내에서 수행될 수 있다. 바람직한 구체예에서, 세포의 형질 감염은 시험관 내에서 수행되며, 형질 감염된 세포는 치료 또는 예방을 필요로 하는 개체, 바람직하게는 사람인 환자에 투여된다. 바람직하게 시험관 내 형질 감염될 세포는 개체, 바람직하게는 사람인 환자의 분리된 세포이다. 그러므로 본 발명은, 개체, 바람직하게는 사람인 환자로부터 세포를 분리하는 단계, 이 분리된 세포를 본 발명에 의한 인공 핵산 분자 또는 본 발명에 의한 벡터로 형질 감염시키는 단계, 그리고 형질 감염된 세포를 개체, 바람직하게는 사람인 환자에 투여하는 단계를 포함하는 치료 방법을 제공한다.
- [0236] 본 발명에 의한 질환의 치료 또는 예방 방법은 전술된 바와 같은 백신화 방법 및/또는 유전자 치료법인 것이 바람직하다.
- [0237] 전술된 바와 같이, 본 발명의 3'-UTR 요소는 mRNA 분자를 안정화할 수 있고/있거나 mRNA 분자로부터의 단백질 생산을 안정화 및/또는 연장할 수 있다. 그러므로 추가의 양태에서, 본 발명은, RNA 분자, 바람직하게는 mRNA 분자 또는 이 RNA 분자를 암호화하는 벡터를, 알부민 유전자의 3'-UTR 또는 알부민 유전자의 3'-UTR의 변이체가 유래하는 핵산 서열을 포함하거나 이 핵산 서열로 이루어진 3'-UTR 요소, 바람직하게는 전술된 바와 같은 3'-UTR 요소와 결합시키는 단계를 포함하는, RNA 분자, 바람직하게는 mRNA 분자를 안정화하는 방법에 관한 것이다.

- [0238] 뿐만 아니라, 본 발명은 본 발명의 인공 핵산 분자 또는 벡터, 바람직하게는 mRNA 분자로부터의 단백질 생산량을 증가시키고/증가시키거나, 본 발명의 인공 핵산 분자 또는 벡터, 바람직하게는 mRNA 분자로부터의 단백질 생산을 안정화 및/또는 연장하는 방법에 관한 것인데, 여기서, 이 방법은 본 발명의 인공 핵산 분자 또는 벡터, 바람직하게는 mRNA 분자를, 알부민 유전자의 3'-UTR로부터 유래하거나 알부민 유전자의 3'-UTR의 변이체로부터 유래하는 핵산 서열을 포함하거나 이 핵산 서열로 이루어진 3'-UTR 요소, 바람직하게는 전술된 바와 같은 3'-UTR 요소를 결합시키는 단계를 포함한다.
- [0239] 본 발명의 내용 중 용어 "본 발명의 인공 핵산 분자 또는 벡터와 3'-UTR 요소를 결합시키는 것"이란, 본 발명의 인공 핵산 분자 또는 벡터와 3'-UTR 요소를 기능상 결합시키는 것 또는 기능상 합하는 것을 의미한다. 이는, 인공 핵산 분자 또는 벡터와 3'-UTR 요소, 바람직하게는 전술된 3'-UTR 요소가 결합 또는 커플링되어, 3'-UTR 요소의 기능, 예를 들어 RNA 및/또는 단백질 생산 안정화 기능이 발휘되는 것을 의미한다. 통상적으로 이는, 3'-UTR 요소가 인공 핵산 분자 또는 벡터, 바람직하게는 mRNA 분자 내에 개방 해독 틀에 대해 3'-말단 쪽에 통합되는 것, 바람직하게는 개방 해독 틀에 대해 3'-말단 쪽에 바로 붙어 통합되는 것, 바람직하게는 개방 해독 틀과 폴리(A) 서열 또는 폴리아데닐화 신호 사이에 통합되는 것을 의미한다. 바람직하게 3'-UTR 요소는 인공 핵산 분자 또는 벡터, 바람직하게는 mRNA에 3'-UTR로서 통합되는데, 다시 말해서 상기 3'-UTR 요소는 인공 핵산 분자 또는 벡터의 3'-UTR, 바람직하게는 mRNA의 3'-UTR로서, 개방 해독 틀의 3'-말단 쪽으로부터 폴리(A) 서열 또는 폴리아데닐화 신호의 5'-말단 쪽(임의로는 짧은 링커, 예를 들어 하나 이상의 제한 위치를 포함하거나 이것으로 이루어진 서열을 통해 상기 개방 해독 틀의 3'-말단과 연결되어 있기도 함)까지 뻗어있다. 그러므로 상기 용어 "본 발명의 인공 핵산 분자 또는 벡터와 3'-UTR 요소를 결합시키는 것"이란, 본 발명의 인공 핵산 분자 또는 벡터, 바람직하게는 mRNA 분자 내에 존재하는 개방 해독 틀과 3'-UTR 요소를 기능상 결합시키는 것을 의미한다. 상기 3'-UTR 및 ORF는 본 발명에 의한 인공 핵산 분자에 대해 전술된 바와 같은데, 예를 들어 상기 ORF 및 3'-UTR은 전술된 바와 같이 이중인 것(예를 들어 상이한 유전자들로부터 유래하는 것)이 바람직하다.
- [0240] 추가의 양태에서, 본 발명은, RNA 분자, 바람직하게는 mRNA 분자의 안정성을 증가시키기 위한, 3'-UTR 요소, 바람직하게는 전술된 바와 같은 3'-UTR 요소의 용도를 제공하는데, 여기서, 상기 3'-UTR 요소는 알부민 유전자의 3'-UTR 또는 알부민 유전자의 3'-UTR의 변이체로부터 유래하는 핵산 서열을 포함하거나 이 핵산 서열로 이루어져 있다.
- [0241] 뿐만 아니라, 본 발명은, 본 발명의 인공 핵산 분자 또는 벡터, 바람직하게는 mRNA 분자로부터의 단백질 생산량을 증가시키고/증가시키거나, 본 발명의 인공 핵산 분자 또는 벡터 분자, 바람직하게는 mRNA 분자로부터의 단백질 생산을 안정화 및/또는 연장하기 위한, 3'-UTR 요소, 바람직하게는 전술된 3'-UTR 요소의 용도를 제공하는데, 여기서, 상기 3'-UTR 요소는 전술된 바와 같이 알부민 유전자의 3'-UTR 또는 알부민 유전자의 3'-UTR의 변이체로부터 유래하는 핵산 분자를 포함하거나 이 핵산 분자로 이루어져 있다.
- [0242] 본 발명에 의한 용도는 바람직하게 본 발명의 인공 핵산 분자, 벡터 또는 RNA를 전술된 3'-UTR 요소와 결합하는 것을 포함한다.
- [0243] 본 발명의 약학 조성물 중 화합물들 및 성분들은 또한 서로 별도로 제조 및 거래될 수도 있다. 그러므로 본 발명은, 본 발명에 의한 인공 핵산 분자, 본 발명에 의한 벡터, 본 발명에 의한 세포 및/또는 본 발명에 의한 약학 조성물을 포함하는 키트 또는 이것들을 포함하는 부분들로 이루어진 키트에 관한 것이기도 하다. 상기 키트 또는 상기된 바와 같은 부분들로 이루어진 키트는 사용 지침, 형질 감염용 세포, 보조제, 약학 조성물 투여 수단, 본 발명의 인공 핵산 분자, 벡터, 세포 또는 약학 조성물을 용해 또는 희석함에 있어 약학적으로 허용 가능한 캐리어 및/또는 약학적으로 허용 가능한 용액을 추가로 포함할 수 있다.

### 도면의 간단한 설명

[0244] 이하 도면들, 서열들 및 실시예들은 본 발명을 추가로 설명하기 위한 것이다. 이와 같은 것들은 본 발명에 관한 특허 대상을 제한하고자 하는 것이 아니다.

도 1은, 사람  $\alpha$ -글로빈 3'-UTR, 사람 알부민 3'-UTR 및 사람  $\beta$ -글루쿠로니다제 3'-UTR의, 인공 mRNA로부터의 루시퍼라제 발현에 대한 영향력을 보여주는 것이다. 이 도면에 있어서, 사람 알부민 3'-UTR을 포함하는 mRNA는 본 발명에 의한 mRNA이다. 상기 mRNA는 포토누스 피랄리스(*Photinus pyralis*)의 루시퍼라제를 암호화하는 개방 해독 틀, 서열 번호 2에 의한 3'-UTR 요소(5'→3' 방향), 그리고 길이 64개 아데닌인 폴리(A) 서열을, 이 순서대로 포함한다. 인공 mRNA(서열 번호 2의 서열에 상응하는 사람 알부민 3'-UTR 포함)로부터의 단백질 발현이 확연히 연장되었음이 관찰될 수 있다.

사람  $\alpha$ -글로빈 3'-UTR, 사람 알부민 3'-UTR 및 사람  $\beta$ -글루쿠로니다제 3'-UTR의, mRNA로부터의 루시페라제 발현에 대한 영향력이, 3'-UTR이 결실된 mRNA로부터의 루시페라제 발현의 경우와 비교되면서 관찰되었다. 즉 상이한 mRNA들이 전기 천공에 의해 HeLa 세포에 도입되었다. 형질 감염 후 6, 24 및 48시간 경과시 루시페라제 수준이 측정되었다. 3'-UTR이 결실된 mRNA로부터 합성되는 루시페라제 수준은 형질 감염 후 6-48시간 사이에 감소하는데, 형질 감염 후 48시간 경과시에는 6시간 경과시 신호의 10%로 유지된다.  $\alpha$ -글로빈 3'-UTR은 mRNA로부터의 루시페라제 발현을 중간 정도로 안정화한다. 그러나 놀랍게도 본 발명의 사람 알부민 3'-UTR은 mRNA로부터의 루시페라제 발현을 확연히 더욱 연장한다. 이와는 대조적으로, 안정한  $\beta$ -글루쿠로니다제 mRNA의 3'-UTR은 루시페라제 발현을, 알부민 3'-UTR이 존재할 때 관찰되는 정도로 연장하지는 않는데, 이는 곧 알부민 3'-UTR이 mRNA로부터의 단백질 발현을 연장함에 있어서 특히 효율적임을 입증하는 것이다. 3회 실시한 형질 감염에 대한 데이터를 소재로 그래프가 작성된다[평균 RLU  $\pm$  SD(상대적 광 단위(Relative Light Unit)  $\pm$  표준 편차)]. RLU는 실시예 5.1에 요약되어 있다.

도 2는, 사람 알부민 3'-UTR의 mRNA로부터의 루시페라제 발현에 대한 영향력을, 사람  $\alpha$ -글로빈 3'-UTR 또는 몇몇 상이한 안정형 mRNA들 각각의 3'-UTR들을 포함하는 mRNA로부터의 루시페라제 발현의 경우와 비교하면서 보여주는 것이다. 즉 상이한 mRNA들이 전기 천공에 의해 HeLa 세포에 도입되었다. 형질 감염 후 6, 24 및 48시간 경과시 루시페라제 수준이 측정되었다. 3'-UTR이 결실된 mRNA로부터 합성되는 루시페라제 수준은 형질 감염 후 6-48시간 사이에 감소하는데, 형질 감염 후 48시간 경과시에는 6시간 경과시 신호의 14%로 유지된다.  $\alpha$ -글로빈 3'-UTR은 mRNA로부터의 루시페라제 발현을 중간 정도로 안정화한다. 그러나 놀랍게도 본 발명의 사람 알부민 3'-UTR은 mRNA로부터의 루시페라제 발현을 확연히 더욱 연장한다. 알부민 3'-UTR과 비교되게, 몇몇 상이한 안정형 mRNA의 3'-UTR들은 mRNA로부터의 루시페라제 발현에 훨씬 덜 유리한 양상으로 영향을 미친다: atp5o 및 atp5l 3'-UTR은 루시페라제 발현을 알부민 3'-UTR보다 훨씬 덜 안정화한다. 더욱이, atp5o 및 atp5l 3'-UTR은 알부민 3'-UTR보다 루시페라제 수준을 상당히 감소시킨다. 안정한 ndufa1 mRNA의 3'-UTR은 루시페라제 발현을 확연히 안정화한다. 그러나 상기 ndufa1 3'-UTR도 루시페라제 수준을 상당히 감소시킨다. 알부민 3'-UTR은, 총 단백질 발현량은 유지하면서 단백질 발현을 연장함에 있어서 특유의 요소이다. 3회 실시한 형질 감염에 대한 데이터를 소재로 그래프가 작성된다[평균 RLU  $\pm$  SD(상대적 광 단위  $\pm$  표준 편차)]. RLU는 실시예 5.2에 요약되어 있다.

도 3은, 사람 알부민 3'-UTR로부터 HindIII 및/또는 XbaI 제한 위치 및/또는 T7 종결 신호가 제거되는 점 돌연변이의, 사람 알부민 3'-UTR 포함 mRNA로부터의 루시페라제 발현에 대한 영향력을 보여주는 것이다. 즉 상이한 mRNA들이 전기 천공에 의해 HeLa 세포에 도입되었다. 형질 감염 후 6, 24, 48 및 72시간 경과시 루시페라제 수준이 측정되었다. PpLuc 신호가, 공동 형질 감염된 RrLuc의 신호에 의하여 형질 감염 효율에 대해 보정되었다.  $\alpha$ -글로빈 3'-UTR은 mRNA로부터의 루시페라제 발현을 단지 중간 정도로만 안정화한다. 이와는 대조적으로, 알부민 3'-UTR의 모든 변이체들은 mRNA로부터의 루시페라제 발현을 확연히 연장한다. 3회 실시한 형질 감염에 대한 데이터를 소재로 그래프가 작성된다[평균 RLU  $\pm$  SD(상대적 광 단위  $\pm$  표준 편차)]. RLU는 실시예 5.4에 요약되어 있다.

도 4는, PpLuc(GC)—A64의 3'-UTR 결실 mRNA 서열을 보여주는 것이다, 도 5는, PpLuc(GC)—알부민—A64의 mRNA 서열을 보여주는 것이다. 사람 알부민의 3'-UTR은 ORF와 폴리(A) 사이에 삽입되었다. 서열은 문헌[Dugaiczky et al. 1982; Proc Natl Acad Sci U S A. Jan;79(1):71-5]에 따라 입수되었다.

도 6은, PpLuc(GC)—알부민2—A64의 mRNA 서열을 보여주는 것이다. 단일 점 돌연 변이에 의해 T7 종결 신호가 제거된 사람 알부민의 3'-UTR이 ORF와 폴리(A) 사이에 삽입되었다.

도 7은, PpLuc(GC)—알부민3—A64의 mRNA 서열을 보여주는 것이다. 단일 점 돌연 변이에 의해 T7 종결 신호가 제거된 사람 알부민의 3'-UTR이 ORF와 폴리(A) 사이에 삽입되었다.

도 8은, PpLuc(GC)—알부민4—A64의 mRNA 서열을 보여주는 것이다. 단일 점 돌연 변이에 의해 T7 종결 신호가 제거된 사람 알부민의 3'-UTR이 ORF와 폴리(A) 사이에 삽입되었다.

도 9는, PpLuc(GC)—알부민5—A64의 mRNA 서열을 보여주는 것이다. 2회의 연속 점 돌연 변이에 의해 T7 종결 신호가 제거된 사람 알부민의 3'-UTR이 ORF와 폴리(A) 사이에 삽입되었다.

도 10은, PpLuc(GC)—알부민6—A64의 mRNA 서열을 보여주는 것이다. 2회의 단일 점 돌연 변이에 의해 HindIII 및 XbaI 제한 위치가 제거된 사람 알부민의 3'-UTR이 ORF와 폴리(A) 사이에 삽입되었다.

도 11은, PpLuc(GC)—알부민7—A64의 mRNA 서열을 보여주는 것이다. 3회의 단일 점 돌연 변이에 의해 T7 종결

신호뿐만 아니라 HindIII 및 XbaI 제한 위치도 제거된 사람 알부민의 3'-UTR이 ORF와 폴리(A) 사이에 삽입되었다.

도 12는, PpLuc(GC)—ag—A64의 mRNA 서열을 보여주는 것이다. 사람  $\alpha$ -글로빈의 3'-UTR의 중앙  $\alpha$ -복합체-결합부가 ORF와 폴리(A) 사이에 삽입되었다.

도 13은, PpLuc(GC)—gusb—A64의 mRNA 서열을 보여주는 것이다. 사람  $\beta$ -글루쿠로니다제의 3'-UTR이 ORF와 폴리(A) 사이에 삽입되었다.

도 14는, PpLuc(GC)—atp5o—A64의 mRNA 서열을 보여주는 것이다. 사람 ATP 신타제 서브유닛 0의 3'-UTR이 ORF와 폴리(A) 사이에 삽입되었다.

도 15는, PpLuc(GC)—ndufa1—A64의 mRNA 서열을 보여주는 것이다. 사람 NADH 탈수소효소[유비퀴논] 1 $\alpha$  종속 복합체 서브유닛 1의 3'-UTR이 ORF와 폴리(A) 사이에 삽입되었다.

도 16은, PpLuc(GC)—atp5l—A64의 mRNA 서열을 보여주는 것이다. 사람 ATP 신타제 서브유닛 g의 3'-UTR이 ORF와 폴리(A) 사이에 삽입되었다.

도 17은, 본 발명에 의한 3'-UTR 요소의 단백질 생산 안정화 효과 및/또는 단백질 생산 연장 효과를 예를 들어 도시하는 것이다. 곡선은, 예를 들어 포유동물 세포 내에서 핵산 분자로부터 생산되는 단백질의 양(경시적으로 측정됨)을 나타내는 것이다. 실선은 본 발명에 의한 인공 핵산 분자, 예를 들어 인공 mRNA로부터의 단백질 생산량을 나타내는 것이고, 점선은 기준 핵산 분자(예를 들어 3'-UTR이 결실되었거나 기준 3'-UTR, 예를 들어 리포터 단백질을 암호화하는 ORF와 함께 자연 발생하는 3'-UTR을 포함하는 기준 핵산 분자)로부터의 단백질 생산량을 나타내는 것이다. 수평의 굵은 실선은 역치 값을 나타내는 것이다. 이는, 예를 들어 발현 개시 후, 예를 들어 핵산 분자의 형질 감염 후 1, 2, 3, 4, 5 또는 6시간 경과시 측정된 단백질의 양일 수 있다. 기준 핵산 분자로부터 생산되는 단백질의 양은, 발현 개시 후, 예를 들어 형질 감염 후 약 32시간 경과시의 역치 값을 밑도는 반면에, 본 발명에 의한 인공 핵산 분자로부터 생산되는 단백질의 양은, 발현 개시 후, 예를 들어 형질 감염 후 약 68시간 경과시의 역치 값을 밑돈다. 단백질의 총 생산량은 곡선 아래의 면적(AUC)과 같다. 바람직하게 본 발명에 의한 인공 핵산 분자로부터 생산된 단백질의 총량은, 3'-UTR이 결실된 기준 핵산 분자로부터 생산된 단백질의 총량보다 많다.

도 18은, 영장류로부터 유래하는 상이한 알부민 3'-UTR들의, 인공 mRNA로부터의 루시페라제 발현에 대한 영향력을, 3'-UTR이 결실된 mRNA로부터의 루시페라제 발현의 경우와 비교하면서 보여주는 것이다. 본 도면에서, 사람 알부민 3'-UTR을 포함하는 mRNA(알부민 8)와, 올리브 개코 원숭이 유래 알부민의 3'-UTR을 포함하는 mRNA(알부민 8)은 본 발명에 의한 mRNA이다. 상기 mRNA들은 포티누스 피랄리스의 루시페라제를 암호화하는 개방 해독 틀, 서열 번호 2 또는 서열 번호 33에 의한 3'-UTR 요소(5'→3' 방향), 그리고 길이 64개 아데닌인 폴리(A) 서열을, 이 순서대로 포함한다. 인공 mRNA(서열 번호 2 또는 서열 번호 33의 서열에 상응하는 알부민 3'-UTR 포함)로부터의 단백질 발현이 확연히 연장되었음이 관찰될 수 있다.

사람 알부민 3'-UTR 및 올리브 개코 원숭이 유래 알부민의 3'-UTR의, mRNA로부터의 루시페라제 발현에 대한 영향력이, 3'-UTR이 결실된 mRNA로부터의 루시페라제 발현의 경우와 비교되면서 관찰되었다. 루시페라제 발현을 관찰하기 위해 상이한 mRNA들이 전기 천공으로 사람 진피 섬유 아세포(HDF)에 도입되었다. 형질 감염 후 6, 24, 48 및 72시간 경과시 루시페라제 수준이 측정되었다. 3'-UTR이 결실된 mRNA로부터 합성되는 루시페라제 수준은 형질 감염 후 6~72시간 사이에 감소하였는데, 형질 감염 후 72시간 경과시에는 6시간 경과시 신호의 5%로 유지된다. 또한, 본 발명의 사람 알부민 3'-UTR은 mRNA로부터의 루시페라제 발현을 확연히 연장한다. 올리브 개코 원숭이 유래 알부민 3'-UTR(알부민 8)은 mRNA로부터의 루시페라제 발현을, 사람 알부민 3'-UTR 서열의 경우와 동일한 정도로 연장한다. 그러므로 영장류 유래 알부민 3'-UTR은 mRNA로부터의 단백질 발현을 연장하는데 특히 적당하다. 3회 실시한 형질 감염에 대한 데이터를 소재로 그래프가 작성된다[평균 RLU  $\pm$  SD(상대적 광 단위  $\pm$  표준 편차)]. RLU는 실시예 5.4 및 표 7에 요약되어 있다.

도 19는, 영장류로부터 유래하는 상이한 알부민 3'-UTR들의, mRNA로부터의 루시페라제 발현에 대한 영향력을, 3'-UTR이 결실된 mRNA로부터의 루시페라제 발현의 경우와 비교하면서 보여주는 것이다. 즉 상이한 mRNA들이 전기 천공에 의해 사람 진피 섬유 아세포(HDF)에 도입되었다. 형질 감염 후 6, 24, 48 및 72시간 경과시 루시페라제 수준이 측정되었다. 3'-UTR이 결실된 mRNA로부터 합성되는 루시페라제 수준은 형질 감염 후 6~48시간 사이에, 형질 감염 후 6시간 경과시 신호의 82%로 감소한다. 또한, 사람 알부민 3'-UTR은 mRNA로부터의 루시페라제 발현을 확연히 연장하는데, 형질 감염 후 48시간 경과시의 신호는 형질 감염 후 6시간 경과시의 신호보다

켰다. 올리브 개코 원숭이 유래 알부민 3'-UTR(알부민 8)은 mRNA로부터의 루시퍼라제 발현을, 사람 알부민 3'-UTR 서열과 동일한 정도로 연장한다. 그러므로 영장류 유래 알부민 3'-UTR은 mRNA로부터의 단백질 발현을 연장하는데 특히 적당하다. 3회 실시한 형질 감염에 대한 데이터를 소재로 그래프가 작성된다[평균 RLU ± SD(상대적 광 단위 ± 표준 편차)]. RLU는 실시예 5.4 및 표 9에 요약되어 있다.

도 20은, 영장류로부터 유래하는 상이한 알부민 3'-UTR들의, 마우스 세포 내 mRNA로부터의 루시퍼라제 발현에 대한 영향력을, 3'-UTR이 결실된 mRNA로부터의 루시퍼라제 발현의 경우와 비교하면서 보여주는 것이다. 즉 상이한 mRNA들이 리포펙션에 의해 L-929 세포, 즉 쥐과 동물 섬유 아세포 세포주에 도입되었다. 형질 감염 후 6, 24, 48 및 72시간 경과시 루시퍼라제 수준이 측정되었다. 3'-UTR이 결실된 mRNA로부터 합성되는 루시퍼라제 수준은 형질 감염 후 6~48시간 사이에 감소하였는데, 형질 감염 후 48시간 경과시에는 형질 감염 후 6시간 경과시 신호의 23%로 유지된다. 또한, 쥐과 동물 세포 내에서조차도 사람 알부민 3'-UTR은 mRNA로부터의 루시퍼라제 발현을 확연히 연장한다. 올리브 개코 원숭이 유래 알부민 3'-UTR(알부민 8)은 mRNA로부터의 루시퍼라제 발현을 유사하게 연장한다. 그러므로 영장류 유래 알부민 3'-UTR은 포유동물 세포 류 내에서 mRNA로부터의 단백질 발현을 연장하는데 특히 적당하다. 3회 실시한 형질 감염에 대한 데이터를 소재로 그래프가 작성된다[평균 RLU ± SD(상대적 광 단위 ± 표준 편차)]. RLU는 실시예 5.5 및 표 11에 요약되어 있다.

도 21은, 포유동물로부터 유래하는 상이한 알부민 3'-UTR들의, mRNA로부터의 루시퍼라제 발현에 대한 영향력을, 3'-UTR이 결실된 mRNA로부터의 루시퍼라제 발현의 경우와 비교하면서 보여주는 것이다. 즉 상이한 mRNA들이 리포펙션에 의해 HeLa 세포에 도입되었다. 형질 감염 후 6, 24, 48 및 72시간 경과시 루시퍼라제 수준이 측정되었다. 3'-UTR이 결실된 mRNA로부터 합성되는 루시퍼라제 수준은 형질 감염 후 6~48시간 사이에 감소하였는데, 형질 감염 후 48시간 경과시에는 형질 감염 후 6시간 경과시의 신호의 49%로 유지된다. 또한 사람 알부민 3'-UTR 및 올리브 개코 원숭이 유래 알부민 3'-UTR(알부민 8)은 mRNA로부터의 루시퍼라제 발현을 확연히 연장하는데, 형질 감염 후 48시간 경과시의 신호는 형질 감염 후 6시간 경과시의 신호보다 컸다. 중요한 점은, 마우스 유래 알부민 3'-UTR(알부민 9) 역시 mRNA로부터의 루시퍼라제 발현을 유사하게 연장한다는 점이다. 그러므로 포유동물 유래 알부민 3'-UTR은 mRNA로부터의 단백질 발현을 연장하는데 특히 적당하다. 3회 실시한 형질 감염에 대한 데이터를 소재로 그래프가 작성된다[평균 RLU ± SD(상대적 광 단위 ± 표준 편차)]. RLU는 실시예 5.6 및 표 13에 요약되어 있다.

도 22는, PpLuc(GC)—알부민 8—A64의 mRNA 서열을 보여주는 것이다. 올리브 개코 원숭이 알부민의 3'-UTR은 ORF와 폴리(A) 사이에 삽입되었다. 이 서열은 NCBI 기준 서열 XM\_003898783.1로부터 입수되었다.

도 23은, PpLuc(GC)—알부민 9—A64의 mRNA 서열을 보여주는 것이다. 마우스 알부민의 3'-UTR이 ORF와 폴리(A) 사이에 삽입되었다. 이 서열은 NCBI 기준 서열 XM\_009654.3으로부터 입수되었다.

**발명을 실시하기 위한 구체적인 내용**

[0245]

**실시예**

[0246]

**1. DNA 주형의 제조**

[0247]

T7 프로모터, 포터누스 피랄리스 루시퍼라제를 암호화하는 GC-풍부 서열(PpLuc(GC)) 및 A64 폴리(A) 서열을 이 순서대로 포함하는 시험관 내 전사용 벡터를 제조하였다. 폴리(A) 서열을, 벡터를 선형화하는데 사용되는 제한 위치 바로 앞에 붙인 다음 시험관 내 전사를 수행하여, A64 폴리(A) 서열로 종결되는 mRNA를 제조하였다. 시험관 내 전사를 통해 상기 벡터로부터 생산된 mRNA를 "PpLuc(GC)—A64"라 명명하였다.

[0248]

상기 벡터가 개방 해독 틀의 3' 말단 비번역 서열(3'-UTR)을 포함하도록 변형시켰다. 다음과 같은 mRNA 암호화 서열들을 포함하는 벡터들을 제조하였다(도 4~16 및 도 22 및 23):

[0249]

서열 번호 5(도 4): PpLuc(GC)—A64

[0250]

서열 번호 6(도 5): PpLuc(GC)—알부민—A64

[0251]

서열 번호 7(도 6): PpLuc(GC)—알부민 2—A64

[0252]

서열 번호 8(도 7): PpLuc(GC)—알부민 3—A64

[0253]

서열 번호 9(도 8): PpLuc(GC)—알부민 4—A64

[0254]

서열 번호 10(도 9): PpLuc(GC)—알부민 5—A64

[0255] 서열 번호 11(도 10): PpLuc(GC)—알부민 6—A64

[0256] 서열 번호 12(도 11): PpLuc(GC)—알부민 7—A64

[0257] 서열 번호 13(도 12): PpLuc(GC)—ag—A64

[0258] 서열 번호 14(도 13): PpLuc(GC)—gusb—A64

[0259] 서열 번호 15(도 14): PpLuc(GC)—atp5o—A64

[0260] 서열 번호 16(도 15): PpLuc(GC)—ndufa1—A64

[0261] 서열 번호 17(도 16): PpLuc(GC)—apt51—A64

[0262] 서열 번호 40(도 22): PpLuc(GC)—알부민 8—A64

[0263] 서열 번호 41(도 23): PpLuc(GC)—알부민 9—A64

[0264] 실시예들에서 사용된 mRNA들은 상기 벡터들의 시험관 내 전사를 통해 생산된 것들이다.

[0265] **2. 시험관 내 전사**

[0266] 실시예 1에 의한 DNA 주형을 선형화한 후 T7 중합 효소를 사용하여 시험관 내 전사하였다. 이후 상기 DNA 주형을 DNA 분해효소 처리로 절단하였다. mRNA 전사체는, 과량의 N7-메틸-구아노신-5'-트리포스페이트-5'-구아노신을 전사 반응에 첨가함으로써 생산된 5'-캡을 포함하였다. 이와 같이 제조된 mRNA를 정제하고 나서, 물에 다시 현탁하였다.

[0267] **3. 세포의 형질 감염**

[0268] **3.1 mRNA 전기 천공**

[0269] 세포를 트립신으로 처리한 다음, 옵티-MEM(opti-MEM)에서 세정하였다. 옵티-MEM 200 $\mu$ l 중  $5 \times 10^4$ 개 세포 또는  $1 \times 10^4$ 개 세포 각각을 전기 천공시켜 PpLuc 암호화 mRNA 0.3 $\mu$ g 또는 1 $\mu$ g을 도입하였다. 대조군으로서 PpLuc를 암호화하지 않는 mRNA를 별도로 전기 천공에 의해 도입하였다. 몇몇 실험에서, 레닐라 레니포르미스(*Renilla reniformis*) 루시페라제(RrLuc)를 암호화하는 mRNA(RrLuc mRNA 0.1 $\mu$ g)를 PpLuc mRNA와 함께 전기 천공으로 도입하여, 형질 감염 효율을 제어하였다. 전기 천공된 세포를 배지 1ml 중 24웰 평판에 접종하였다. 형질 감염 후 6, 24 또는 48시간 경과시(몇몇 실험에서는 72시간 경과시), 배지를 흡인해낸 다음, 세포를 용해 완충액[25 mM Tris, pH 7.5(HCl), 2mM EDTA, 10% 글리세롤, 1% Triton X-100, 2mM DTT, 1mM PMSF, 대안적으로는 수동적 용해 완충액(Passive Lysis Buffer)(프로메가(Promega)사)] 200 $\mu$ l 중에서 용해하였다. 루시페라제 활성을 측정할 때까지 용해물을 -20 $^{\circ}$ C에 보관하여 두었다.

[0270] **3.2 mRNA 리포펙션**

[0271] 형질 감염하기 3일 전에 세포를 96웰 평판에 접종하였다(웰당 2500 또는 5000개 세포). 리포펙션을 수행하기 직전에 세포를 옵티-MEM 중에서 세정하였다. 웰당 PpLuc 암호화 mRNA(리포펙타민 2000으로 복합체화된 mRNA) 25ng을 사용하여 세포를 리포펙션하였다. 레닐라 레니포르미스 루시페라제(RrLuc)를 암호화하는 mRNA를 PpLuc mRNA와 함께 공동 형질 감염하여, 형질 감염 효율을 제어하였다(웰당 RrLuc mRNA 2.5ng). 형질 감염 후 6, 24, 48 또는 72시간 경과시, 배지를 흡인해낸 다음, 세포를 용해 완충액[프로메가사의 수동적 용해 완충액] 100 $\mu$ l 중에서 용해하였다. 루시페라제 활성을 측정할 때까지 용해물을 -80 $^{\circ}$ C에 보관하여 두었다.

[0272] **4. 루시페라제 측정**

[0273] 바이오텍 시너지HT(BioTek SynergyHT) 평판 판독기에서 상대적 광 단위(RLU)로서 루시페라제 활성을 측정하였다. 측정 시간 5초 경과시 용해물 50 $\mu$ l와 루시페린 완충액[75  $\mu$ M 루시페린, 25mM 글리실글리신, pH 7.8(NaOH), 15mM MgSO<sub>4</sub>, 2mM ATP] 200 $\mu$ l를 사용하여 PpLuc 활성을 측정하였다. 측정 시간 5초 경과시 용해물 50  $\mu$ l와 코엘렌테라진(coelenterazin) 완충액[40  $\mu$ M 코엘렌테라진, 2.2mM EDTA, 220mM KH<sub>2</sub>PO<sub>4</sub>/K<sub>2</sub>HPO<sub>4</sub> pH 5.0, 1.1M NaCl, 1.3mM Na<sub>3</sub>, 0.44 g/l BSA] 200 $\mu$ l를 사용하여 RrLuc 활성을 측정하였다.

[0274] 대안적으로, 하이텍스 카멜레온(Hidex Chameleon) 평판 판독기에서 루시페라제 활성을 상대적 광단위(RLU)로서 측정하였다. 측정 시간 2초 경과시 용해물 20 $\mu$ l와 루시페린 완충액[비틀 주스(Beetle-Juice), PJK GmbH] 100 $\mu$ l

를 사용하여 PpLuc 활성을 측정하였다. 측정 시간 2초 경과시 용해물 20 $\mu$ l와 코엘렌테라진 완충액[레닐라-쥬스 (Renilla-Juice), PJK GmbH] 100 $\mu$ l를 사용하여 RrLuc 활성을 측정하였다.

[0275] 5. 결과

[0276] 5.1 알부민 3'-UTR은, 널리 알려진  $\alpha$ -글로빈 3'-UTR보다 mRNA로부터의 단백질 발현을 확연히 연장함

[0277] 3' 비번역 영역들의, mRNA로부터의 단백질 발현에 대한 영향력을 연구하기 위해, 상이한 3'-UTR들을 가지는 mRNA들을 실시예 1~2에 따라서 합성하였다: mRNA는 사람  $\alpha$ -글로빈의 3'-UTR의 중앙  $\alpha$ -복합체 결합부(서열 번호 13에 의한 PpLuc(GC)—ag—A64)를 포함시켰는데, 그 이유는  $\alpha$ -글로빈 3'-UTR은 암호화 영역 서열과는 독립되게 mRNA를 안정화한다고 보고된바 있기 때문이었다[Rodgers, N.D., Wang, Z. & Kiledjian, M., 2002. Regulated alpha-globin mRNA decay is a cytoplasmic event proceeding through 3'-to-5'exosome-dependent decapping. *RNA*, 8(12), S.1526 -1537]. 대안적으로, mRNA는 사람 알부민의 3'-UTR(서열 번호 6에 의한 PpLuc(GC)—알부민—A64)을 포함시키기도 하였다. 사람 알부민 mRNA는 안정한 것으로 보고된 바 있다[Johnson, T.R. et al., 1991. Newly synthesized RNA: simultaneous measurement in intact cells of transcription rates and RNA stability of insulin-like growth factor I, actin, and albumin in growth hormone-stimulated hepatocytes. *Proceedings of the National Academy of Sciences*, 88(12), S.5287-5291]. 마지막으로 사람  $\beta$ -글루쿠로니다제의 3'-UTR을 포함하는 mRNA(서열 번호 14에 의한 PpLuc(GC)—gusb—A64)를 사용하였다. 이와 같은 사람의  $\beta$ -글루쿠로니다제 mRNA도 또한 안정한 것으로 보고된 바 있다[Watson, G. & Paigen, K., 1987. Genetic variations in kinetic constants that describe beta-glucuronidase mRNA induction in androgen-treated mice. *Molecular and Cellular Biology*, 7(3), S.1085 -1090]. 비교를 위해서, 3'-UTR이 결실된 mRNA(서열 번호 5에 의한 PpLuc(GC)—A64)도 사용하였다. 루시페라제 암호화 mRNA를 전기 천공으로 HeLa 세포에 도입하였다. 형질 감염 후 6, 24 및 48시간 경과시에 루시페라제 수준을 측정하였다(이하 표 1 및 도 1 참조).

표 1

[0278]

mRNA	6시간 경과시 RLU	24시간 경과시 RLU	48시간 경과시 RLU
PpLuc(GC)-A64	299993	162445	29168
PpLuc(GC)-ag-A64	571131	574896	120029
PpLuc(GC)-알부민-A64	364580	476567	277317
PpLuc(GC)-gusb-A64	357513	431134	134463

[0279]

3'-UTR이 결실된 mRNA로부터 합성되는 루시페라제 수준은 형질 감염 후 6~48시간 사이에 감소하였는데, 48시간 경과시에는 형질 감염 후 6시간 경과시 신호의 10%로 유지되었다.  $\alpha$ -글로빈 3'-UTR은 mRNA로부터의 루시페라제 발현을 단지 중간 정도로만 안정화하였다. 그러나 놀랍게도, 사람 알부민 3'-UTR은 mRNA로부터의 루시페라제 발현을 확연히 더 연장하였다. 이와는 대조적으로,  $\beta$ -글루쿠로니다제의 3'-UTR은 알부민 3'-UTR에 대해 관찰되는 정도로 루시페라제 발현을 연장하지 않았다.

[0280]

형질 감염 후 48시간 및 6시간 경과시 루시페라제 수준들의 비율을 산정하였다(그 수치가 클수록 단백질 발현이 안정화되었음을 말해주는 것임). 임의의 3'-UTR이 단백질 발현의 시간 경과를 얼마나 많이 안정화하였는지 말해주는 데이터를 이하 표 2에 요약하였다.

표 2

[0281]

mRNA	48시간 경과시 RLU / 6시간 경과시 RLU
PpLuc(GC)-A64	0,10
PpLuc(GC)-ag-A64	0,21
PpLuc(GC)-알부민-A64	0,76
PpLuc(GC)-gusb-A64	0,38

[0282]

알부민 3'-UTR은, 널리 알려진  $\alpha$ -글로빈 3'-UTR보다 훨씬 많이, 그리고 안정한  $\beta$ -글루쿠로니다제 mRNA의 3'-UTR보다 많이 단백질 발현을 안정화하였다. 이 결과는, 알부민 3'-UTR이 mRNA로부터의 단백질 발현을 연장하는

데 특히 효율적임을 입증해준다.

[0283] **5.2 알부민 3'-UTR은 총 단백질 발현량을 유지하면서 단백질 발현을 연장함에 있어 특유의 역할을 함**

[0284] (실시에 5.1에서 관찰된 α-글로빈 3'-UTR의 매우 제한된 효과는 이러한 논쟁을 입증해주는 않지만) 알부민 3'-UTR에 의한 단백질 발현이 연장됨은 안정한 mRNA 3'-UTR의 일반 효과를 말해줄 수 있는 것이었다. 그러므로 안정한 mRNA의 3'-UTR을 포함하는 mRNA를 합성하였다[Friedel, C.C. et al., 2009. Conserved principles of mammalian transcriptional regulation revealed by RNA half-life. Nucleic Acids Research, 37(17), S.e115.]. atp5o 또는 ndufa1 또는 atp5l은 각각 사람 ATP 신타제 서브유닛 0, 사람 NADH 탈수소효소 [유비퀴논] 1α 하위 복합체 서브유닛 1, 또는 사람 ATP 신타제 서브유닛 g이다. 루시퍼라제 암호화 mRNA를 전기 천공에 의해 HeLa 세포에 도입하였다. 형질 도입 후 6, 24 및 48시간 경과시 루시퍼라제 수준을 측정하였다(이하 표 3 및 도 2 참조).

**표 3**

[0285]

mRNA	6시간 경과시 RLU	24시간 경과시 RLU	48시간 경과시 RLU
PpLuc(GC)-A64	529804	725602	72348
PpLuc(GC)-ag-A64	1065036	1790023	263484
PpLuc(GC)-알부민-A64	548821	1261832	523000
PpLuc(GC)-atp5o-A64	239418	402629	79566
PpLuc(GC)-ndufa1-A64	116139	277149	133723
PpLuc(GC)-atp5l-A64	58610	56553	9728

[0286] 3'-UTR이 결실된 mRNA로부터 합성되는 루시퍼라제 수준은 형질 감염 후 6~48시간 사이에는 감소하였으며, 형질 감염 후 48시간 경과시에는 6시간 경과시 신호의 14%로 유지되었다. α-글로빈 3'-UTR은 mRNA로부터의 루시퍼라제 발현을 중간 정도로 안정화하였다. 그러나 놀랍게도, 사람 알부민 3'-UTR은 mRNA로부터의 루시퍼라제 발현을 확연히 더 연장하였다. 알부민 3'-UTR이 그러했던 것과는 대조적으로, 몇몇 상이한 안정형 mRNA의 3'-UTR은 mRNA로부터의 루시퍼라제 발현에 훨씬 덜 유리한 양상으로 영향을 미쳤다: atp5o 및 atp5l 3'-UTR들은 알부민 3'-UTR보다 루시퍼라제 발현을 훨씬 덜 안정화하였다. 또한 atp5o 및 atp5l 3'-UTR이 존재할 때의 루시퍼라제 수준은 알부민 3'-UTR이 존재할 때의 루시퍼라제 수준보다 상당히 낮았다. 안정한 ndufa1 mRNA의 3'-UTR은 루시퍼라제 발현을 확연히 안정화하였다. 그러나 ndufa1 3'-UTR도 루시퍼라제 수준을 상당히 감소시켰다.

[0287] 형질 감염 후 48시간 및 6시간 경과시 루시퍼라제 수준들의 비율을 산정하였다(그 수치가 클수록 단백질 발현이 안정화되었음을 말해주는 것임). 임의의 3'-UTR이 단백질 발현의 시간 경과를 얼마나 많이 안정화하였는지 말해주는 데이터를 이하 표 4에 요약하였다.

**표 4**

[0288]

mRNA	48시간 경과시 RLU / 6시간 경과시 RLU
PpLuc(GC)-A64	0,14
PpLuc(GC)-ag-A64	0,25
PpLuc(GC)-알부민-A64	0,95
PpLuc(GC)-atp5o-A64	0,33
PpLuc(GC)-ndufa1-A64	1,15
PpLuc(GC)-atp5l-A64	0,17

[0289] 알부민 3'-UTR은 단백질의 총 발현량을 유지하면서 단백질 발현을 연장함에 있어서 특유의 역할을 하였다. 알부민 3'-UTR은, 널리 알려진 α-글로빈 3'-UTR과 몇몇 상이한 안정형 mRNA의 3'-UTR과 비교되게, 가장 늦은 시점에서의 단백질 발현량을 상당히 증가시켰다. 이 결과는, 알부민 3'-UTR이 mRNA로부터의 단백질 발현을 연장하는데 특히 적당함을 입증해준다.

[0290] **5.3 알부민 3'-UTR의 상이한 변이체들은 mRNA로부터의 단백질 발현을 연장함**

[0291] 사람 알부민 3'-UTR은 HindIII 제한 위치, XbaI 제한 위치 및 T7 종결 신호를 포함하였다. 그러므로 HindIII 및/또는 XbaI 제한 위치 및/또는 T7 종결 신호가 점 돌연 변이(들)에 의해 제거되었으며, 사람 알부민 3'-UTR의

변이체를 포함하는 mRNA를 합성하였다[서열 번호 17~12에 의한 PpLuc(GC)-알부민 2-7]. 루시페라제 암호화 mRNA를 전기 천공에 의해 HeLa 세포에 도입하였다. 형질 전환 후 6, 24, 48 및 72 시간 경과시 루시페라제 수준을 측정하였다. PpLuc 신호를 공동 형질 감염된 RrLuc의 신호에 의해 형질 감염 효율에 대해 보정하였다(이하 표 5 및 도 3 참조).

표 5

[0292]

mRNA	6시간 경과시 RLU	24시간 경과시 RLU	48시간 경과시 RLU	72시간 경과시 RLU
PpLuc(GC)-A64	382909	576557	122118	20962
PpLuc(GC)-ag-A64	281262	536346	118000	20356
PpLuc(GC)-알부민-A64	199494	499804	282475	134271
PpLuc(GC)-알부민2-A64	258516	655711	351888	186869
PpLuc(GC)-알부민3-A64	219365	547307	292511	124330
PpLuc(GC)-알부민4-A64	236873	576151	298229	139260
PpLuc(GC)-알부민5-A64	223815	576899	289954	131145
PpLuc(GC)-알부민6-A64	180412	455039	240086	99802
PpLuc(GC)-알부민7-A64	174371	417171	216048	68887

[0293]

형질 감염 후 72시간 및 6시간 경과시 루시페라제 수준들의 비율을 산정하였다(그 수치가 클수록 단백질 발현이 안정화되었음을 말해주는 것임). 임의의 3'-UTR이 단백질 발현의 시간 경과를 얼마나 많이 안정화하였는지 말해주는 데이터를 이하 표 6에 요약하였다.

표 6

[0294]

mRNA	72시간 경과시 RLU / 6시간 경과시 RLU
PpLuc(GC)-A64	0,05
PpLuc(GC)-ag-A64	0,07
PpLuc(GC)-알부민-A64	0,67
PpLuc(GC)-알부민2-A64	0,72
PpLuc(GC)-알부민3-A64	0,57
PpLuc(GC)-알부민4-A64	0,59
PpLuc(GC)-알부민5-A64	0,59
PpLuc(GC)-알부민6-A64	0,55
PpLuc(GC)-알부민7-A64	0,40

[0295]

3'-UTR이 결실된 mRNA로부터 합성되는 루시페라제 수준은 형질 감염 후 6~72시간 사이에 감소하였는데, 72시간 경과시에는 형질 감염 후 6시간 경과시 신호의 5%로 유지되었다. α-글로빈 3'-UTR은 mRNA로부터의 루시페라제 발현을 단지 중간 정도로만 안정화하였다. 이와는 대조적으로, 알부민 3'-UTR의 모든 변이체는 mRNA로부터의 루시페라제 발현을 확연히 연장하였다.

[0296]

**5.4 영양류의 알부민 3'-UTR은 mRNA로부터의 단백질 발현을 연장함**

[0297]

알부민 3'-UTR에 의한 단백질 발현의 연장은 종 특이적일 것이다. 상이한 영양류들로부터 유래하는 알부민 3'-UTR들을 비교하였을 때, 올리브 개코 원숭이 알부민 3'-UTR은 사람 알부민 3'-UTR과의 상동성이 가장 작았다(통상의 침팬지와와의 동일성: 99%, 피그미 침팬지와와의 동일성: 99%, 수마트라 오랑우탄과의 동일성: 99%, 올리브 개코 원숭이와의 동일성: 96%). 그러므로 올리브 개코 원숭이 알부민 유전자의 3'-UTR을 포함하는 mRNA를 합성하였다[서열 번호 40에 의한 PpLuc(GC)-알부민8-A64]. 루시페라제 암호화 mRNA들을 전기 천공에 의해 사람 진피 섬유 아세포(HDF)에 도입하였다. 형질 감염 후 6, 24, 48 및 72시간 경과시 루시페라제 수준을 측정하였다. PpLuc 신호를 공동 형질 감염된 RrLuc의 신호에 의해 형질 감염 효율에 대해 보정하였다(이하 표 7 및 도 18 참조).

표 7

[0298]

mRNA	6시간 경과시 RLU	24시간 경과시 RLU	48시간 경과시 RLU	72시간 경과시 RLU
PpLuc(GC)-A64	130469	72629	26267	6637
PpLuc(GC)-알부민-A64	70661	74152	38209	17648
PpLuc(GC)-알부민8-A64	71463	123361	51361	18373

[0299]

형질 감염 후 72시간 및 6시간 경과시 루시페라제 수준들의 비율을 산정하였다(그 수치가 클수록 단백질 발현이 안정화되었음을 말해주는 것임). 임의의 3'-UTR이 경시적으로 단백질 발현의 시간 경과를 얼마나 많이 안정화하였는지 말해주는 데이터를 이하 표 8에 요약하였다.

표 8

[0300]

mRNA	72시간 경과시 RLU / 6시간 경과시 RLU
PpLuc(GC)-A64	0.05
PpLuc(GC)-알부민-A64	0.25
PpLuc(GC)-알부민8-A64	0.26

[0301]

3'-UTR이 결실된 mRNA로부터 합성되는 루시페라제 수준은 형질 감염 후 6~72시간 사이에 감소하였는데, 72시간 경과시에는 형질 감염 후 6시간 경과시 신호의 5%로 유지되었다. 사람 알부민 3'-UTR은 mRNA로부터의 루시페라제 발현을 확연히 연장하였다. 올리브 개코 원숭이 유래 알부민 3'-UTR은 사람 서열의 경우와 동일하게 mRNA로부터의 루시페라제 발현을 연장하였다. 이 결과는, 영장류 유래 알부민 3'-UTR이 mRNA로부터의 단백질 발현을 연장하는데 특히 적당함을 입증해 주는 것이다.

[0302]

만일 mRNA가 전기 천공에 의하기보다는 리포펙션에 의해 도입되면, 사람 유래 알부민 3'-UTR 및 올리브 개코 원숭이 유래 알부민 3'-UTR 둘 다에 의해서도 단백질 발현이 연장되는 것이 관찰되었다. 루시페라제 암호화 mRNA를 HDF에 리포펙션으로 도입하였다. 형질 감염 후 6, 24, 48 및 72시간 경과시 루시페라제 수준을 측정하였다. PpLuc 신호를 공동 형질 감염된 RrLuc의 신호에 의해 형질 감염 효율에 대해 보정하였다(이하 표 9 및 도 19 참조).

표 9

[0303]

mRNA	6시간 경과시 RLU	24시간 경과시 RLU	48시간 경과시 RLU	72시간 경과시 RLU
PpLuc(GC)-A64	3285	7946	2725	1266
PpLuc(GC)-알부민-A64	3743	9525	5466	3381
PpLuc(GC)-알부민8-A64	3259	7367	4044	1892

[0304]

형질 감염 후 48시간 및 6시간 경과시 루시페라제 수준들을 산정하였다(그 수치가 클수록 단백질 발현이 안정화되었음을 말해주는 것임). 임의의 3'-UTR이 단백질 발현의 시간 경과를 얼마나 많이 안정화하였는지 말해주는 데이터를 이하 표 10에 요약하였다.

표 10

[0305]

mRNA	48시간 경과시 RLU / 6시간 경과시 RLU
PpLuc(GC)-A64	0.82
PpLuc(GC)-알부민-A64	1.46
PpLuc(GC)-알부민8-A64	1.24

[0306]

전기 천공보다는 리포펙션을 수행하였을 때, 사람 알부민 3'-UTR과 올리브 개코 원숭이 유래 알부민 3'-UTR 둘 다는 mRNA로부터의 루시페라제 발현을 확연히 연장하였다. 이 결과는, 영장류 유래 알부민 3'-UTR이 mRNA로부터

의 단백질 발현을 연장하는데 특히 적당함을 확인시켜 주는 것이다.

[0307] **5.5 영장류의 알부민 3'-UTR은 마우스 세포 내에서 mRNA로부터의 단백질 발현을 연장함**

[0308] 알부민 3'-UTR에 의한 단백질 발현의 연장은 종 특이적일 것이다. 그러므로 사람 알부민 3'-UTR과 올리브 개코 원숭이 유래 알부민 3'-UTR이 마우스 세포 내에서 mRNA로부터의 루시페라제 발현을 연장하는지 여부에 대해 테스트하였다. 루시페라제 암호화 mRNA를 L-929 세포, 즉 쥐과 동물 섬유 아세포 세포주에 리포펙션으로 도입하였다. 형질 감염 후 6, 24, 48 및 72시간 경과시 루시페라제 수준을 측정하였다. PpLuc 신호를 공동 형질 감염된 RrLuc의 신호에 의해 형질 감염 효율에 대해 보정하였다(이하 표 11 및 도 20 참조).

**표 11**

[0309]

mRNA	6시간 경과시 RLU	24시간 경과시 RLU	48시간 경과시 RLU	72시간 경과시 RLU
PpLuc(GC)-A64	177805	128658	40414	12593
PpLuc(GC)-알부민-A64	160478	244279	101177	22605
PpLuc(GC)-알부민8-A64	151076	178839	68786	16969

[0310] 형질 감염 후 48시간 및 6시간 경과시 루시페라제 수준들의 비율을 산정하였다(그 수치가 클수록 단백질 발현이 안정화되었음을 말해주는 것임). 임의의 3'-UTR이 단백질 발현의 시간 경과를 얼마나 많이 안정화하였는지 말해주는 데이터를 이하 표 12에 요약하였다.

**표 12**

[0311]

mRNA	48시간 경과시 RLU / 6시간 경과시 RLU
PpLuc(GC)-A64	0.23
PpLuc(GC)-알부민-A64	0.63
PpLuc(GC)-알부민8-A64	0.46

[0312] 3'-UTR이 결실된 mRNA로부터 합성되는 루시페라제 수준은 형질 감염 후 6~48시간 사이에 감소하였는데, 48시간 경과시에는 형질 감염 후 6시간 경과시 신호의 23%로 유지되었다. 사람 알부민 3'-UTR은 마우스 세포주 내 mRNA로부터의 루시페라제 발현을 확연히 연장하였다. 올리브 개코 원숭이 유래 알부민 3'-UTR도 역시 마우스 세포주 내 mRNA로부터의 루시페라제 발현을 연장하였다. 이 결과는, 영장류 유래 알부민 3'-UTR이 포유동물 세포 류에서 mRNA로부터의 단백질 발현을 연장하는데 특히 적당함을 입증해 주는 것이다.

[0313] **5.6 포유동물의 알부민 3'-UTR은 mRNA로부터의 단백질 발현을 연장함**

[0314] 알부민 3'-UTR에 의한 단백질 발현의 연장은 종 특이적일 것이다. 상이한 포유동물들로부터 유래하는 알부민 3'-UTR들을 비교하였을 때, 마우스 알부민 3'-UTR은 사람 알부민 3'-UTR과의 상동성이 가장 작았다(말과의 동일성: 86%, 애완용 개와의 동일성: 84%, 소와의 동일성: 74%, 래트와의 동일성: 73%, 마우스와의 동일성: 72%). 그러므로 마우스 알부민 유전자의 3'-UTR을 포함하는 mRNA를 합성하였다[서열 번호 41에 의한 PpLuc(GC)-알부민 9-A64]. 루시페라제 암호화 mRNA들을 리포펙션에 의해 HeLa 세포에 도입하였다. 형질 감염 후 6, 24, 48 및 72 시간 경과시 루시페라제 수준을 측정하였다. PpLuc 신호를 공동 형질 감염된 RrLuc의 신호에 의해 형질 감염 효율에 대해 보정하였다(이하 표 13 및 도 21 참조).

**표 13**

[0315]

mRNA	6시간 경과시 RLU	24시간 경과시 RLU	48시간 경과시 RLU	72시간 경과시 RLU
PpLuc(GC)-A64	46533	200168	22702	3001
PpLuc(GC)-알부민-A64	42931	224315	109190	25314
PpLuc(GC)-알부민8-A64	39783	205950	82918	17447
PpLuc(GC)-알부민9-A64	42500	210365	60893	8380

[0316] 형질 감염 후 48시간 및 6시간 경과시 루시페라제 수준들의 비율을 산정하였다(그 수치가 클수록 단백질 발현이

안정화되었음을 말해주는 것임). 임의의 3'-UTR이 단백질 발현의 시간 경과를 얼마나 많이 안정화하였는지 말해주는 데이터를 이하 표 14에 요약하였다.

표 14

[0317]

mRNA	48시간 경과시 RLU / 6시간 경과시 RLU
PpLuc(GC)-A64	0.49
PpLuc(GC)-알부민-A64	2.54
PpLuc(GC)-알부민8-A64	2.08
PpLuc(GC)-알부민9-A64	1.43

[0318]

3'-UTR이 결실된 mRNA로부터 합성되는 루시퍼라제 수준은 형질 감염 후 6~48시간 사이에 감소하였는데, 48시간 경과시에는 형질 감염 후 6시간 경과시 신호의 49%로 유지되었다. 사람 알부민 3'-UTR과 올리브 개코 원숭이 유래 알부민 3'-UTR은 mRNA로부터의 루시퍼라제 발현을 확연히 연장하였다. 중요한 점은, 마우스 유래 알부민 3'-UTR은 사람 HeLa 세포주 내에서 mRNA로부터의 루시퍼라제 발현을 유사하게 연장하였다는 점이다. 이 결과는, 포유동물 유래 알부민 3'-UTR이 mRNA로부터의 단백질 발현을 연장하는데 특히 적합함을 입증해 주는 것이다.

[0319]

서열:

서열 번호 1:

CATCACATTT AAAAGCATCT CAGCCTACCA TGAGAATAAG AGAAAGAAAA TGAAGATCAA  
AAGCTTATTC ATCTGTTTTT CTTTTTCGTT GGTGTAAAGC CAACACCCCTG TCTAAAAAAC  
ATAAATTTCT TTAATCATTT TGCCTCTTTT CTCTGTGCTT CAATTAATAA AAAATGGAAA  
GAATCT

서열 번호 2:

CAUCACAUUU AAAAGCAUCU CAGCCUACCA UGAGAAUAAG AGAAAGAAAA UGAAGAUCAA  
AAGCUUAUUC AUCUGUUUUU CUUUUUCGUU GGUGUAAAGC CAACACCCUG UCUAAAAAAC  
AUAAAUUUCU UUAUCAUUU UGCCUCUUU CUCUGUGCUU CAUUUAAUAA AAAAUGGAAA  
GAAUCU

서열 번호 3:

CATCACATTT AAAAGCATCT CAGCCTACCA TGAGAATAAG AGAAAGAAAA TGAAGATCAA  
AAGCTTATTC ATCTGTTTTT CTTTTTCGTT GGTGTAAAGC CAACACCCCTG TCTAAAAAAC  
ATAAATTTCT TTAATCATTT TGCCTCTTTT CTCTGTGCTT CAATTAATAA AAAATGGAAA  
GAATCTAGAT CTAATAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA  
AAAAAAAAA AAAAAA

서열 번호 4:

CAUCACAUUU AAAAGCAUCU CAGCCUACCA UGAGAAUAAG AGAAAGAAAA UGAAGAUCAA  
AAGCUUAUUC AUCUGUUUUU CUUUUUCGUU GGUGUAAAGC CAACACCCUG UCUAAAAAAC  
AUAAAUUUCU UUAUCAUUU UGCCUCUUU CUCUGUGCUU CAUUUAAUAA AAAAUGGAAA  
GAAUCUAGAU CUAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA  
AAAAAAAAA AAAAAA

서열 번호 5:

GGGAGAAAGCTTGAGGATGAGGACGCCAAGAACATCAAGAAGGGCCCGGCCCTTCTA

[0320]

CCCGCTGGAGGACGGGACCGCCGGCAGCAGCTCCACAAGGCCATGAAGCGGTACGCCCT  
 GGTGCCGGGCACGATCGCCTTCACCGACGCCACATCGAGGTCGACATCACCTACGCGGA  
 GTACTTCGAGATGAGCGTGCCTGGCCGAGGCCATGAAGCGGTACGGCCTGAACACCAA  
 CCACCGGATCGTGGTGTGCTCGGAGAACAGCCTGCAGTTCTTCATGCCGGTGTGGGCGC  
 CCTCTTCATCGCGGTGGCCGTGCCCCGGCGAACGACATCTACAACGAGCGGGAGCTGCT  
 GAACAGCATGGGGATCAGCCAGCCGACCGTGGTGTTCGTGAGCAAGAAGGGCCTGCAGAA  
 GATCCTGAACGTGCAGAAGAAGCTGCCATCATCCAGAAGATCATCATGACAGCAA  
 GACCGACTACCAGGGCTTCCAGTCGATGTACACGTTTCGTGACCAGCCACCTCCCGCCGGG  
 CTTCAACGAGTACGACTTCGTCCCGGAGAGCTTCGACCGGGACAAGACCATCGCCCTGAT  
 CATGAACAGCAGCGGCAGCACCAGCCCTGCCGAAGGGGTGGCCCTGCCGCACCGGACCGC  
 CTGCGTGCCTTCTCGCAGCCCGGGACCCCATCTTCGGCAACCAGATCATCCCGGACAC  
 CGCCATCCTGAGCGTGGTGCCTTCCACCAGGCTTCGGCATGTTACGACCTGGGCTA  
 CCTCATTCGCGGCTTCCGGGTGGTCTGATGTACCGGTTTCGAGGAGGAGCTGTTCTGCG  
 GAGCCTGCAGGACTACAAGATCCAGAGCGGCTGCTCGTGCCGACCTGTTACGCTTCTT  
 CGCCAAGAGCACCTGATCGACAAGTACGACCTGTGAACTGCACGAGATCGCCAGCGG  
 GGGCGCCCCGCTGAGCAAGGAGGTGGGCGAGGCCGTGGCCAAGCGGTTCCACCTCCCGGG  
 CATCCGCCAGGGCTACGGCCTGACCAGACCACGAGCGGATCCTGATCACCCCGAGGG  
 GGACGACAAGCCGGGGCGCCGTGGGCAAGTGGTCCCGTTCTTCGAGGCCAAGGTGGTGA  
 CCTGGACACCGCAAGACCTGGGCGTGAACCAGCGGGCGAGCTGTGCGTCCGGGGCC  
 GATGATCATGAGCGGTACGTGAACAACCCGGAGGCCACCAACGCCCTCATCGACAAGGA  
 CGGCTGGCTGCACAGCGGCACATCGCCTACTGGGACGAGGACGAGCACTTCTTCATCGT  
 CGACCGGCTGAAGTTCGTGATCAAGTACAAGGGCTACCAGTGGCGCCGGCCGAGCTGGA  
 GAGCATCCTGCTCCAGCACCCCAACATCTTCGACCGCGGCTGGCCGGGCTGCCGGACGA  
 CGACCGCGCGAGCTGCCGGCCGGTGGTGGTGTGGAGCACGGCAAGACCATGACGGA  
 GAAGGAGATCGTGCAGTACGTGGCCAGCCAGGTGACCACCGCCAAGAAGCTGCGGGGCGG  
 CGTGGTGTTCGTGGACGAGTCCCGAAGGGCTGACCGGGAAGCTCGACGCCCGGAAGAT  
 CCGCGAGATCCTGATCAAGGCCAAGAAGGGCGGCAAGATCGCCGTGAAGACTAGTAGAT  
 CTAA  
 AAAAAA

서열 번호 6:

GGGAGAAAGCTTGGAGTGGAGGACGCCAAGAACATCAAGAAGGGCCCGCGCCCTTCTA  
 CCCGCTGGAGGACGGGACCGCCGGCAGCAGCTCCACAAGGCCATGAAGCGGTACGCCCT  
 GGTGCCGGGCACGATCGCCTTCACCGACGCCACATCGAGGTCGACATCACCTACGCGGA  
 GTACTTCGAGATGAGCGTGCCTGGCCGAGGCCATGAAGCGGTACGGCCTGAACACCAA  
 CCACCGGATCGTGGTGTGCTCGGAGAACAGCCTGCAGTTCTTCATGCCGGTGTGGGCGC  
 CCTCTTCATCGCGGTGGCCGTGCCCCGGCGAACGACATCTACAACGAGCGGGAGCTGCT  
 GAACAGCATGGGGATCAGCCAGCCGACCGTGGTGTTCGTGAGCAAGAAGGGCCTGCAGAA  
 GATCCTGAACGTGCAGAAGAAGCTGCCATCATCCAGAAGATCATCATGACAGCAA  
 GACCGACTACCAGGGCTTCCAGTCGATGTACACGTTTCGTGACCAGCCACCTCCCGCCGGG  
 CTTCAACGAGTACGACTTCGTCCCGGAGAGCTTCGACCGGGACAAGACCATCGCCCTGAT  
 CATGAACAGCAGCGGCAGCACCAGCCCTGCCGAAGGGGTGGCCCTGCCGCACCGGACCGC  
 CTGCGTGCCTTCTCGCAGCCCGGGACCCCATCTTCGGCAACCAGATCATCCCGGACAC  
 CGCCATCCTGAGCGTGGTGCCTTCCACCAGGCTTCGGCATGTTACGACCTGGGCTA  
 CCTCATCTGCGGCTTCCGGGTGGTCTGATGTACCGGTTTCGAGGAGGAGCTGTTCTGCG  
 GAGCCTGCAGGACTACAAGATCCAGAGCGGCTGCTCGTGCCGACCTGTTACGCTTCTT  
 CGCCAAGAGCACCTGATCGACAAGTACGACCTGTGAACTGCACGAGATCGCCAGCGG  
 GGGCGCCCCGCTGAGCAAGGAGGTGGGCGAGGCCGTGGCCAAGCGGTTCCACCTCCCGGG  
 CATCCGCCAGGGCTACGGCCTGACCAGACCACGAGCGGATCCTGATCACCCCGAGGG  
 GGACGACAAGCCGGGGCGCCGTGGGCAAGTGGTCCCGTTCTTCGAGGCCAAGGTGGTGA  
 CCTGGACACCGCAAGACCCTGGGCGTGAACCAGCGGGCGAGCTGTGCGTCCGGGGCC  
 GATGATCATGAGCGGTACGTGACAACCCGGAGGCCACCAACGCCCTCATCGACAAGGA

[0321]

CGGCTGGGTGCACAGCGGGACATCGCCTACTGGGACGAGGACGAGCACTTCTTCATCGT  
CGACCGGTGAAGTCGCTGATCAAGTACAGGGCTACCAGGTGGCGCCGGCCGAGCTGGA  
GAGCATCCTGCTCCAGCACCCCAACATCTTCGACGCCGGCGTGGCCGGGCTGCCGGACGA  
CGACGCCGGCGAGCTGCCGGCCGGTGGTGGTGGTGGAGCACGGCAAGACCATGACGGGA  
GAAGGAGATCGTCGACTACGTGGCCAGCCAGGTGACCACCCCAAGAAGCTGCGGGGGCGG  
CGTGGTGTTCGTGGACGAGGTCCCGAAGGGCCTGACCGGGAAGCTCGACGCCGGGAAGAT  
CCGCGAGATCCTGATCAAGGCCAAGAAGGGCGCAAGATCGCCGTGTAAGACTAGTGCAT  
CACATTTAAAAGCATCTCAGCCTACCATGAGAATAAGAGAAAAGAAATGAAGATCAAAG  
CTTATTCATCTGTTTTCTTTCTTTCTGTTGGTGTAAAGCCAACACCCTGTCTAAAAACATA  
AATTTCTTAATCATTTTGCCTCTTTCTCTGTGCTTCAATTAATAAAAAATGGAAGAA  
TCTAGATCTAA  
AAAAAAAAAAAAA

서열 번호 7:

GGGAGAAAGCTTGAGGATGGAGGACGCCAAGAACATCAAGAAGGGCCCGGCGCCCTTCTA  
CCCCTGGAGGACGGGACCGCCGGCGAGCAGCTCCACAAGGCCATGAAGCGGTACGCCCT  
GGTCCGGGCACGATCGCCTTCACCGACGCCACATCGAGGTGACATCACCTACGCGGA  
GTACTTCGAGATGAGCGTGCCCTGGCCGAGGCCATGAAGCGGTACGCCCTGAACACCAA  
CCACCGGATCGTGGTGTGCTCGGAGAACAGCCTGCAGTTCTTCATGCCGGTGTGGGCGC  
CCTTTCATCGCGTGGCCGTCGCCCGGGCAACGACATCTACAACGAGCGGGAGCTGCT  
GAAGCAGATGGGATCAGCCAGCCAGCCGTGGTGTTCGTGAGCAAGAGGGCCTGCAGAA  
GATCCTGAACGTGCAGAAGAAGCTGCCATCATCCAGAAGATCATCATGACGACGAA  
GACCGACTACCAGGGCTTCCAGTCGATGTACAGTTCTGTGACCGCCACCTCCCGCGGG  
CTTCAACGAGTACGACTTCGTCCCGGAGAGCTTCGACCGGGACAAGACCATCGCCCTGAT  
CATGAACAGCAGCGGCAGCACCGCCCTGCCAAGGGGGTGGCCCTGCCGCACGGGACCGC  
CTGCGTGGCCTTCTCGACGCCCGGGACCCATCTTCGGCAACCAGATCATCCCGGACAC  
CGCCATCCTGAGCGTGGTGGCCTCCACCACGGCTTCGGCATGTTACGACCCCTGGGCTA  
CCTCATCTGCGGCTCCGGGTGGTCTGATGTACCGGTTTCGAGGAGGAGCTGTTCTCTGCG  
GAGCCTGCAGGACTACAAGATCCAGAGCGCGCTGCTGTCGCGACCCCTGTTTCAGCTTCTT  
CGCCAAGAGCACCTGATCGACAAGTACGACCTGTCAACCTGCACGAGATCGCCAGCGG  
GGCGCCCCGCTGAGCAAGGAGGTGGGCGAGCCGTGGCCAAGCGGTTCCACCTCCCGGG  
CATCCGCCAGGGCTACGGCTGACCGAGACCACGAGCGCGATCCTGATCACCCCGAGGG  
GGACGACAGCCGGGCGCCGTGGGCAAGGTGGTCCCCTTTCGAGGCCAAGGTGGTGGGA  
CCTGGACACCGGCAAGACCTGGGCGTGAACAGCGGGCGAGCTGTGCGTGGGGGGCC  
GATGATCATGAGCGGCTACGTGAACAACCCGGAGGCCACCAACGCCCTCATCGACAAGGA  
CGGCTGGTGCACAGCGGCACATCGCCTACTGGGACGAGGACGAGCACTTCTTCATCGT  
CGACCGGTGAAGTCGCTGATCAAGTACAAGGGCTACCAGGTGGCGCCGGCCGAGCTGGA  
GAGCATCCTGCTCCAGCACCCCAACATCTTCGACGCCGGCGTGGCCGGGCTGCCGGACGA  
CGACGCCGGCGAGCTGCCGGCCGGTGGTGGTGGTGGAGCACGGCAAGACCATGACGGGA  
GAAGGAGATCGTCGACTACGTGGCCAGCCAGGTGACCACCCCAAGAAGCTGCGGGGGCGG  
CGTGGTGTTCGTGGACGAGGTCCCGAAGGGCCTGACCGGGAAGCTCGACGCCCGGAAGAT  
CCGCGAGATCCTGATCAAGGCCAAGAAGGGCGGCAAGATCGCCGTGTAAGACTAGTGCAT  
CACATTTAAAAGCATCTCAGCCTACCATGAGAATAAGAGAAAAGAAATGAAGATCAAAG  
CTTATTCGCTGTTTTCTTTCTTTCTGTTGGTGTAAAGCCAACACCCTGTCTAAAAACATA  
AATTTCTTAATCATTTTGCCTCTTTCTCTGTGCTTCAATTAATAAAAAATGGAAGAA  
TCTAGATCTAA  
AAAAAAAAAAAAA

서열 번호 8:

GGGAGAAAGCTTGAGGATGGAGGACGCCAAGAACATCAAGAAGGGCCCGGCGCCCTTCTA  
CCCCTGGAGGACGGGACCGCCGGCGAGCAGCTCCACAAGGCCATGAAGCGGTACGCCCT  
GGTCCGGGCACGATCGCCTTCACCGACGCCACATCGAGGTGACATCACCTACGCGGA

[0322]

GTACTTCGAGATGAGCGTGCCTGGCCGAGGCCATGAAGCGGTACGGCCTGAACACCAA  
CCACCGGATCGTGGTGTGCTCGGAGAACAGCCTGCAGTTCTTCATGCCGGTGTGGGCGC  
CCTTTCATCGGCGTGGCCGTGCCCCGGCGAACGACATCTACAACGAGCGGGAGCTGCT  
GAACAGCATGGGGATCAGCCAGCCGACCGTGGTGTTCGTGAGCAAGAAGGGCCTGCAGAA  
GATCCTGAACGTGCAGAAGAAGCTGCCATCATCCAGAAGATCATCATATGGACAGCAA  
GACCGACTACCAGGGCTTCCAGTTCGATGTACACGTTTCGTGACCAGCCACCTCCCGCCGGG  
CTTCAACGAGTACGACTTCGTCCCGGAGAGCTTCGACCGGGACAAGACCATCGCCCTGAT  
CATGAACAGCAGCGGCAGCACCGGCCCTGCCGAAGGGGGTGGCCCTGCCGACCGGACCGC  
CTGCGTGCCTTCTCGCACGCCCGGGACCCCATCTTCGGCAACCAGATCATCCCGGACAC  
CGCCATCCTGAGCGTGGTGCCTTCCACCAGGCTTCGGCATGTTACGACCCCTGGGCTA  
CCTCATCTGCGGCTCCCGGTGGTCTGATGTACCAGGTTTCGAGGAGGAGCTGTTCTGCG  
GAGCCTGCAGGACTACAAGATCCAGAGCGCGTGTTCGTGCCGACCTGTTACGCTTCTT  
CGCCAAGAGCACCTGATCGACAAGTACGACCTGTTCGAACCTGCACGAGATCGCCAGCGG  
GGCGCCCCGCTGAGCAAGGAGGTGGGCGAGGCCGTGGCCAAGCGGTTCCACCTCCCGGG  
CATCCGCCAGGGCTACGGCCTGACCGAGACCACGAGCGCATCCTGATCACCCCGAGGG  
GGACACAAGCCGGCGCCGTGGGCAAGGTGGTCCCGTTCCTCGAGGCCAAGGTGGTGA  
CCTGGACACCGGCAAGACCTGGGCGTGAACCAGCGGGCGAGCTGTGCGTGCAGGGGGC  
GATGATCATGAGCGGCTACGTGAACAACCCGAGGCCACCAACGCCCTCATCGACAAGGA  
GGCTGGCTGCACAGCGCGCATCGCCTACTGGGACGAGGACGAGCACTTCTTCATCGT  
CGACCGGCTGAAGTTCGTGATCAAGTACAAGGGCTACCAGGTGGCGCCGGCCGAGCTGGA  
GAGCATCCTGCTCCAGCACCCCAACATCTTCGACGCGCGGTGGCCGGCTGCCGACGA  
CGACCGCGGAGCTGCCGCGCGGTGGTGGTGTGAGCAGCGCAAGACCATGACGGGA  
GAAGGAGATCGTGCAGTACGTGGCCAGCCAGGTGACCACCGCCAAGAAGCTGCCGGCGG  
CGTGGTGTTCGTGGACGAGGTCCCGAAGGGCTGACCGGGAAGCTCGACCGCCGGAAGAT  
CCCGGAGATCCTGATCAAGGCCAAGAAGGGCGGCAAGATCGCCGTGTAAGACTAGTGAT  
CACATTTAAAGATCTCAGCCTACCATGAGAATAAGAGAAAGAAATGAAGTCAAAAG  
CTTATTCATCAGTTTTTCTTTTCGTGGTGTAAAGCCAACACCTGTCTAAAAACATA  
AATTTCTTTAATCATTTTGCCTTTTCTCTGTGCTTCAATTAATAAAAAATGGAAAGAA  
TCTAGATCTAA  
AAAAAAAAAAAAA

서열 번호 9:

GGGAGAAAGCTTGAGGATGGAGGACGCCAAGAACATCAAGAAGGGCCCGGCGCCTTCTA  
CCCCTGGAGGACGGGACCGCCGGGAGCAGCTCCACAAGCCATGAAGCGGTACGCCCT  
GGTGGCGGACGATCGCCTTACCGACGCCACATCGAGGTTCGACATCACTACGGGA  
GTACTTCGAGATGAGCGTGCCTGGCCGAGGCCATGAAGCGGTACGGCCTGAACACCAA  
CCACCGGATCGTGGTGTGCTCGGAGAACAGCCTGCAGTTCTTCATGCCGGTGTGGGCGC  
CCTTTCATCGGCGTGGCCGTGCCCCGGCGAACGACATCTACAACGAGCGGGAGCTGCT  
GAACAGCATGGGGATCAGCCAGCCGACCGTGGTGTTCGTGAGCAAGAAGGGCCTGCAGAA  
GATCCTGAACGTGCAGAAGAAGCTGCCATCATCCAGAAGATCATCATATGGACAGCAA  
GACCGACTACCAGGGCTTCCAGTTCGATGTACACGTTTCGTGACCAGCCACCTCCCGCCGG  
CTTCAACGAGTACGACTTCGTCCCGGAGAGCTTCGACCGGGACAAGACCATCGCCTGAT  
CATGAACAGCAGCGGCAGCACCGGCCCTGCCGAAGGGGGTGGCCCTGCCGACCGGACCGC  
CTGCGTGCCTTCTCGCACGCCCGGGACCCCATCTTCGGCAACCAGATCATCCCGGACAC  
CGCCATCCTGAGCGTGGTGCCTTCCACCAGGCTTCGGCATGTTACGACCCCTGGGCTA  
CCTCATCTGCGGCTTCCGGGTGGTCTGATGTACCAGGTTTCGAGGAGGAGCTGTTCTGCG  
GAGCCTGCAGGACTACAAGATCCAGAGCGCGTGTTCGTGCCGACCTGTTACGCTTCTT  
CGCCAAGAGCACCTGATCGACAAGTACGACCTGTTCGAACCTGCACGAGATCGCCAGCGG  
GGGCGCCCCGCTGAGCAAGGAGGTGGGCGAGGCCGTGGCCAAGCGGTTCCACCTCCCGGG  
CATCCGCCAGGGCTACGGCCTGACCGAGACCACGAGCGCATCCTGATCACCCCGAGGG  
GGACGACAAGCCGGCGCCGTGGGCAAGGTGGTCCCGTTCCTCGAGGCCAAGGTGGTGA  
CCTGGACACCGGCAAGACCTGGGCGTGAACCAGCGGGCGAGCTGTGCGTGCAGGGGGC

[0323]



GGTGCCGGGCACGATCGCCTTACCAGCCACATCGAGGTCGACATCACCTACGCGGA  
GTACTTCGAGATGAGCGTGCCTGGCCGAGGCCATGAAGCGGTACGGCCTGAACACCAA  
CCACCGGATCGTGGTGTGCTCGGAGAACAGCCTGCAGTCTTTCATGCCGGTGTGGGCGC  
CCTTTCATCGCGTGGCGTGCCTCCGCGGGAACGACATCTACAACGAGCGGGAGCTGCT  
GAACAGCATGGGATCAGCCAGCCGACCGTGGTGTTCGTGAGCAAGAAGGGCCTGCAGAA  
GATCCTGAACGTGCAGAAGAAGCTGCCATCATCCAGAAGATCATCATCATGGACAGCAA  
GACCGACTACCAGGGTTCAGTTCGATGTACACGTTTCGTGACCAGCCACCTCCCGCGGG  
TTCAACGAGTACGACTTCGTCCCGGAGAGCTTCGACCGGGACAAGACCATCGCCCTGAT  
CATGAACAGCAGCGGCAGCACCGGCCCTGCCGAAGGGGGTGGCCCTGCCGCACCGGACCGC  
CTGCGTGCCTTTCGACCGCCCGGACCCCATCTTCGGCAACCAGATCATCCCGGACAC  
CGCCATCCTGAGCGTGGTCCGTTCCACCACGGCTTCGGCATGTTACGACCTGGGCTA  
CCTCATCTGCGGCTCCGGGTGGTCTGATGTACCGGTTTCGAGGAGGAGCTGTTCTGCG  
GAGCCTGCAGACTACAAGATCCAGAGCGCGTGTGCTGCGGACCTGTTAGCTTCTT  
CGCCAGAGCACCCTGATCGACAAGTACGACCTGTGCAACCTGCACGAGATCGCCAGCGG  
GGGCGCCCGCTGAGCAAGGAGGTGGGCGAGGCCGTGGCCAAGCGGTTCCACCTCCCGGG  
CATCCGCGAGGGTACGGCCTGACCAGACACGAGCGCGATCCTGATCACCCCGAGGG  
GGACGACAAGCCGGGCGCCGTGGGCAAGGTGGTCCCGTTCCTCGAGGCCAAGGTGGTGA  
CCTGGACACCGCAAGACCTGGCGTGAACCAGCGGGGCGAGCTGTGCGTGGGGGGGCG  
GATGATCATGAGCGGCTACGTGAACAACCCGGAGGCCACCAACGCCCTCATCGACAAGGA  
CGGCTGGCTGCACAGCGGCGACATCGCCTACTGGGACGAGGACGAGCACTTCTTCATCGT  
CGACCGGCTGAAGTCGCTGATCAAGTACAAGGGCTACCAGGTGGCGCCGGCGAGCTGGA  
GAGCATCTGCTCCAGCACCCCAACATCTTCGACCGCGCGTGGCCGGGCTGCCGAGCA  
CGACCGCGGAGCTGCCGCGCGTGGTGGTGTGAGCAGCGCAAGACCATGACGGGA  
GAAGGAGATCGTGCATACGTGGCCAGCGAGTGCACCCGCAAGAAGCTGCCGGGGCGG  
CGTGGTGTAAATCATTTGGCTCCCGAAGGGCTGACCGGGAAGCTCGACGCGCGGAAGAT  
CCGCGAGATCCTGATCAAGGCCAAGAAGGGCGGCAAGATCGCCGTGAAGACTAGTGCAT  
CACATTTAAAGCATCTCAGCCTACCATGAGAATAAGAGAAAAGAAAATGAAGATCAATAG  
CTTATTCATCTGTTTTCTTTTCGTTGGTGTAAAGCCAACACCTGTCTAAAAACATA  
AATTTCTTTAATCATTTGGCTCTTTCTCTGTGCTTCAATTAATAAAAAATGGAAGAA  
CCTAGATCTAAAAA  
AA  
AAAAAAAAAAAAA

서열 번호 12:

GGGAGAAAGCTTGAGGATGGAGGACGCCAAGAACATCAAGAAGGGCCCGCGCCCTTCTA  
CCCCTGGAGGACGGGACCGCCGGCGAGCAGCTCCACAAGGCCATGAAGCGGTACGCCCT  
GGTGCCGGGCACGATCGCCTTACCAGCCACATCGAGGTCGACATCACCTACGCGGA  
GTACTTCGAGATGAGCGTGCCTGGCCGAGGCCATGAAGCGGTACGGCCTGAACACCAA  
CCACCGGATCGTGGTGTGCTCGGAGAACAGCCTGCAGTCTTTCATGCCGGTGTGGGCGC  
CCTTTCATCGCGTGGCGTGCCTCCGCGGGAACGACATCTACAACGAGCGGGAGCTGCT  
GAACAGCATGGGATCAGCCAGCCGACCGTGGTGTTCGTGAGCAAGAAGGGCCTGCAGAA  
GATCCTGAACGTGCAGAAGAAGCTGCCATCATCCAGAAGATCATCATCATGGACAGCAA  
GACCGACTACCAGGGTTCAGTTCGATGTACACGTTTCGTGACCAGCCACCTCCCGCGGG  
CTTCAACGAGTACGACTTCGTCCCGGAGAGCTTCGACCGGGACAAGACCATCGCCCTGAT  
CATGAACAGCAGCGGCAGCACCGGCCCTGCCGAAGGGGGTGGCCCTGCCGCACCGGACCGC  
CTGCGTGCCTTTCGACCGCCCGGACCCCATCTTCGGCAACCAGATCATCCCGGACAC  
CGCCATCCTGAGCGTGGTCCGTTCCACCACGGCTTCGGCATGTTACGACCTGGGCTA  
CCTCATCTGCGGCTCCGGGTGGTCTGATGTACCGGTTTCGAGGAGGAGCTGTTCTGCG  
GAGCCTGCAGACTACAAGATLCAGAGCGCGTGTGCTGCGGACCTGTTAGCTTCTT  
CGCCAGAGCACCCTGATCGACAAGTACGACCTGTGCAACCTGCACGAGATCGCCAGCGG  
GGGCGCCCGCTGAGCAAGGAGGTGGGCGAGGCCGTGGCCAAGCGGTTCCACCTCCCGGG  
CATCCGCGAGGGTACGGCCTGACCAGACACGAGCGCGATCCTGATCACCCCGAGGG  
GGACGACAAGCCGGGCGCCGTGGGCAAGGTGGTCCCGTTCCTTCGAGGCCAAGGTGGTGA

[0325]

CCTGGACACCGGCAAGACCCCTGGGCGTGAACCAGCGGGGCGAGCTGTGCGTGCGGGGGCC  
GATGATCATGAGCGGCTACGTGAACAACCCGAGGCCACCAACGCCCTCATCGACAAGGA  
CGGCTGGCTGCACAGCGCGACATCGCCTACTGGGACGAGGACGAGCATTCTTCATCGT  
CGACCGGCTGAAGTCGCTGATCAAGTACAAGGGCTACCAGGTGGCGCCGGCCGAGCTGGA  
GAGCATCCTGCTCCAGCACCCCAACATCTTCGACGCCGGCGTGGCCGGGTGCCGGACGA  
CGACGCCGGCGAGCTGCCGGCCGGGTGGTGGTGTGGAGCACGGCAAGACCATGACGGA  
GAAGGAGATCGTCGACTACGTGGCCAGCCAGGTGACCAACGCCAAGAAGCTGCCGGGCGG  
CGTGGTGTTCGTGGACGAGGTCCCGAAGGGCCTGACCGGGAAGCTCGACGCCCGGAAGAT  
CCCGGAGATCCTGATCAAGGCCAAGAAGGGCGGCAAGATCGCCGTGTAAGACTAGTGCAT  
CACATTTAAAGCATCTCAGCCTACCATGAGAATAAGAGAAAATAAGATCAATAG  
CTTATTCATCTCTTTTCTTTTCTGTTGGTGTAAAGCCAACACCCTGTCTAAAAACATA  
AATTTCTTTAATCATTTTGCCTCTTTCTCTGTGCTTCAATTAATAAAAAATGGAAAGAA  
CCTAGATCTAAAAAATAA  
AAAAAAAAAAAAA

서열 번호 13:

GGGAGAAAGCTTGAGGATGGAGGACGCCAAGAACATCAAGAAGGGCCCGCGCCTTCTA  
CCCGCTGGAGGACGGGACCGCCGGCGAGCAGCTCCACAAGGCCATGAAGCGGTACGCCCT  
GGTGCCGGGACGATCGCCTTCACCGACGCCACATCGAGGTGCATCACCTACGCGGA  
GTACTTCGAGATGAGCGTGCCTGGCCGAGGCCATGAAGCGGTACGGCCTGAACACCAA  
CCACCGGATCGTGGTGTGCTCGGAGAACAGCCTGCAGTCTTCATGCCGGTGTGGGCGC  
CCTCTTCATCGGCGTGGCCGTGCCCCGGCGAACGACATCTACAACGAGCGGGAGCTGCT  
GAACAGCATGGGATCAGCCAGCCGACCGTGGTGTTCGTGAGCAAGAAGGGCCTGCAGAA  
GATCCTGAACGTGCAGAAGAAGCTGCCATCATCCAGAAGATCATCATGAGACAGCAA  
GACCGACTACCAGGGCTTCCAGTTCGATACACGTTTCGTGACCGACCCCTCCCGCGGG  
CTTCAACGAGTACGACTTCGTCCCGGAGAGCTTCGACCGGGACAAGACCATCGCCCTGAT  
CATGAACAGCAGCGGCGAGCAGCCGCGCTGCCAAGGGGGTGGCCCTGCCGACCGGACCGC  
CTGCGTGCCTTCTCGCACGCCCGGGACCCATCTTCGGCAACCAGATCATCCGGACAC  
CGCCATCCTGAGCGTGGTCCCGTCCACCACGGCTTCGGCATGTTACGACCCCTGGGCTA  
CCTCATCTGCGGCTTCCGGTGGTCTGATGTACCGGTTTCGAGGAGGAGCTGTCTCTGCG  
GAGCCTGCAGACTACAAGATCCAGAGCGCGTGTCTGCGCACCCCTGTTAGCTTCTT  
CGCAAGAGCACCTGATCGACAAGTACGACTGTGCAACCTGCACGAGATCGCCAGCGG  
GGCGCCCCGCTGAGCAAGGAGGTGGGCGAGCCGTGGCCAAGCGGTTCCACCTCCCGGG  
CATCCGCCAGGGCTACGGCCTGACCGAGACCAGAGCGGATCCTGATCACCCCGAGGG  
GGACGACAAGCCGGGCGCCTGGGCAAGGTGGTCCCGTCTTCGAGGCCAAGTGGTGA  
CCTGGACACCGGCAAGACCCCTGGGCGTGAACCAGCGGGGCGAGCTGTGCGTGCGGGGCC  
GATGATCATGAGCGGCTACGTGAACAACCCGAGGCCACCAACGCCCTCATCGACAAGGA  
CGGCTGGCTGCACAGCGGCGACATCGCCTACTGGGACGAGGACGACACTTCTTCATCGT  
CGACCGGCTGAAGTCGCTGATCAAGTACAAGGGCTACCAGGTGGCGCCGGCCGAGCTGGA  
GAGCATCCTGCTCCAGCACCCCAACATCTTCGACGCCGGCGTGGCCGGGTGCCGGACGA  
CGACGCCGGCGAGCTGCCGGCCGGTGGTGGTGTGGAGCACGGCAAGACCATGACGGA  
GAAGGAGATCGTCGACTACGTGGCCAGCCAGGTGACCAACGCCAAGAAGCTGCCGGGCGG  
CGTGGTGTTCGTGGACGAGGTCCCGAAGGGCCTGACCGGGAAGCTCGACGCCCGGAAGAT  
CCCGGAGATCCTGATCAAGGCCAAGAAGGGCGGCAAGATCGCCGTGTAAGACTAGTTATA  
AGACTGACTAGCCGATGGGCTCCCAACGGGCCCTCCTCCCTCCTTGCACCGAGATTA  
ATAGATCTAAAAAATAA  
AAAAAAAAAAAAA

서열 번호 14:

GGGAGAAAGCTTGAGGATGGAGGACGCCAAGAACATCAAGAAGGGCCCGCGCCTTCTA  
CCCGCTGGAGGACGGGACCGCCGGCGAGCAGCTCCACAAGGCCATGAAGCGGTACGCCCT  
GGTGCCGGGACGATCGCCTTCACCGACGCCACATCGAGGTGCATCACCTACGCGGA

[0326]

GTACTTCGAGATGAGCGTGGCCCTGGCCGAGGCCATGAAGCGGTACGGCCTGAACACCAA  
 CCACCGGATCGTGGTGTGCTCGGAGAACAGCCTGCAGTTCTTCATGCCGGTGTGGGGCG  
 CCTTTCATCGGCGTGGCCGTGCCCGCGGAACGACATCTACAACGAGCGGGAGCTGCT  
 GAACAGCATGGGGATCAGCCAGCCGACCGTGGTGTTCGTGAGCAAGAAGGGCCTGCAGAA  
 GATCCTGAACGTGCAGAAGAAGCTGCCATCATCCAGAAGATCATCATGAGCAGCAA  
 GACCGACTACCAGGGCTCCAGTTCAGTGTACACGTTTCGTGACCAGCCACCTCCCGCCGGG  
 CTTCAACGAGTACGACTTCGTCCCGGAGAGCTTCGACCCGGGACAAGACCATCGCCCTGAT  
 CATGAACAGCAGCGGCAGCACCAGCCGCTGCCGAAGGGGGTGGCCCTGCCGACCCGGACCGC  
 CTGCGTGGCCTTCTCGCACGCCCGGGACCCCATCTTCGGCAACCAGATCATCCCGGACAC  
 CGCCATCCTGAGCGTGGTGGCGTTCACCACCGGCTTCGGCATGTTACGACCCCTGGGGTA  
 CCTCATCTGCGGCTTCCGGTGGTCTGATGTACCGGTTTCGAGGAGGAGCTGTTCTCTGCG  
 GAGCCTGCAGGACTACAAGATCCAGAGCGCGTGTCTGTCGCCGACCCCTGTTAGCTTCTT  
 CGCCAAGAGCACCTGATCGACAAGTACGACCTGTCGAACCTGCACGAGATCGCCAGCGG  
 GGGCGCCCGCTGAGCAAGGAGGTGGGGCAGGCGGTGCCAAGCGGTTCCACCTCCCGGG  
 CATCCGCCAGGGCTACGGCTGACCGAGACCACGAGCGCATCCTGATCACCCCGAGGG  
 GGACGACAAGCCGGGCGCCGTGGGCAAGGTGGTCCCGTTCCTCGAGGCCAAGGTGGTGGGA  
 CCTGGACACCGGCAAGACCTGGGGCTGAACAGCGGGGCGAGCTGTGCGTGGCGGGGCC  
 GATGATCATGAGCGGTACGTTGAAACAACCCGAGGCCACCAACCCCTCATCGACAAGGA  
 CGGCTGGCTGCACAGCGGCACATCGCCTACTGGGACGAGGACGACACTTCTTCATCGT  
 CGACCGGTGAAGTTCGTGATCAAGTACAGGGCTACCAAGTGGCCCGGCCGAGCTGGA  
 GAGCATCCTGCTCCAGCACCCCAACATCTTCGACCGCGGCTGGCCGGGTGCCGGACGA  
 CGACGCCCGGCGAGCTGCCGGCCGCGTGGTGGTGTGGAGCACGGCAAGACCTGACCGGA  
 GAAGGAGATCGTGCATACGTGGCCAGCAGGTGACCAACCGCAAGAAGTGGCGGGCGG  
 CGTGGTGTTCGTGGACGAGTCCCGAAGGGCCTGACCGGGAAGCTCGACGCCCGGAAGAT  
 CCGCGAGATCCTGATCAAGGCCAAGAAGGGCGCAAGATCGCCGTGAAGACTAGTGCAA  
 GACTGATACCACTGCGTGTCCCTTCCCTCCCGAGTCAGGGCGACTTCCACAGCAGCAGA  
 ACAAGTCCCTCCTGGACTGTTACGGCAGACCAGAACGTTTCTGGCTGGGTTTGTGGT  
 CATCTATTCTAGCAGGGAACACTAAAGGTGAAATAAAAGATTTTCTATTATGGAATAA  
 AGAGTTGGCATGAAAGTGGCTACTGAGATCTAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

서열 번호 15:

GGGAGAAAGCTTGAGGATGGAGGACGCCAAGAACATCAAGAAGGGCCCGGCCCTTCTA  
 CCCGCTGGAGGACGGGACCGCGGCGAGCAGCTCCACAAGGCCATGAAGCGGTACGCCCT  
 GGTGCCGGGCACGATCGCCTTACCAGACGCCACATCGAGGTGCACATCACCTACGCGGA  
 GTACTTCGAGATGAGCGTGGCCCTGGCCGAGGCCATGAAGCGGTACGGCCTGAACACCAA  
 CCACCGGATCGTGGTGTGCTCGGAGAACAGCCTGCAGTTCTTCATGCCGGTGTGGGGCG  
 CCTTTCATCGGCGTGGCCGTGCCCGCGGAACGACATCTACAACGAGCGGGAGCTGCT  
 GAACAGCATGGGGATCAGCCAGCCGACCGTGGTGTTCGTGAGCAAGAAGGGCCTGCAGAA  
 GATCCTGAACGTGCAGAAGAAGCTGCCATCATCCAGAAGATCATCATGAGCAGCAA  
 GACCGACTACCAGGGCTTCCAGTGCATGTACACGTTTCGTGACCAGCCACCTCCCGCCGGG  
 CTTCAACGAGTACGACTTCGTCCCGGAGAGCTTCGACCCGGGACAAGACCATCGCCCTGAT  
 CATGAACAGCAGCGGCAGCACCAGCCGCTGCCGAAGGGGGTGGCCCTGCCGACCCGACCGC  
 CTGCGTGGCCTTCTCGCACGCCCGGGACCCCATCTTCGGCAACCAGATCATCCCGGACAC  
 CGCCATCCTGAGCGTGGTGGCGTTCACCACCGGCTTCGGCATGTTACGACCCCTGGGGTA  
 CCTCATCTGCGGCTTCCGGTGGTCTGATGTACCGGTTTCGAGGAGGAGCTGTTCTCTGCG  
 GAGCCTGCAGGACTACAAGATCCAGAGCGCGTGTCTGTCGCCGACCCCTGTTAGCTTCTT  
 CGCCAAGAGCACCTGATCGACAAGTACGACCTGTCGAACCTGCACGAGATCGCCAGCGG  
 GGGCGCCCGCTGAGCAAGGAGGTGGGGCAGGCGGTGCCAAGCGGTTCCACCTCCCGGG  
 CATCCGCCAGGGCTACGGCTGACCGAGACCACGAGCGCATCCTGATCACCCCGAGGG  
 GGACGACAAGCCGGGCGCGTGGGCAAGGTGGTCCCGTTCCTCGAGGCCAAGGTGGTGGGA  
 CCTGGACACCGGCAAGACCTGGGGCTGAACAGCGGGGCGAGCTGTGCGTGGCGGGCC

[0327]

GATGATCATGAGCGGCTACGTGAACAACCCGGAGGCCACCAACGCCCTCATCGACAAGGA  
CGGCTGGCTGCACAGCGGGACATCGCCTACTGGGACGAGGACGAGCACTTCTTCATCGT  
CGACCGGCTGAAGTTCGCTGATCAAGTACAAGGGCTACCAGGTGGCGCCGGCCGAGCTGGA  
GAGCATCCTGCTCCAGCACCCCAACATCTTCGACGCCGGCGTGGCCGGGCTGCCGGACGA  
CGACGCCGGCGAGCTGCCGGCCGGTGGTGGTGGTGGAGCACGGCAAGACCATGACGGA  
GAAGGAGATCGTTCGACTACGTGGCCAGCCAGGTGACCACGCCAAGAAGCTGCCGGGCGG  
CGTGGTGTTCGTTGGAGAGGTCCCGAAGGGCCTGACCGGGAAGCTCGACGCCCGAAGAT  
CCGCGAGATCCTGATCAAGGCCAAGAAGGGCGGCAAGATCGCCGTGTAAGACTAGTAAGT  
GTTGGTTTTCTGCATCAGTGAATACTTAAACTGGAGCAACAATAAAAAGCTTCCAG  
AACAGATCAGATCTAAAAA  
AA  
AAAAAAAAAAAAAAAAAAAA

서열 번호 16:

GGGAGAAAGCTTGAGGATGGAGGACGCCAAGAACATCAAGAAGGGCCCGGCCCTTCTA  
CCCCTGGAGGACGGGACCGCCGCGAGCAGCTCCACAAGGCCATGAAGCGGTACGCCCT  
GGTCCGGGACGATCGCCTTACCCGACGCCACATCGAGGTGACATCACCTACGCGGA  
GTACTTCGAGATGAGCGTGGCCCTGGCCGAGGCCATGAAGCGGTACGGCCTGAACACCAA  
CCACCGGATCGTGGTGTGCTCGGAGAACAGCCTGCAGTTCTTCATGCCGGTGTGGGCGC  
CCTTTCATCGCGTGGCCGTCGCCCGGCGAACGACATCTACAACGAGCGGGAGCTGCT  
GAACAGCATGGGGATCAGCCAGCCGACCGTGGTGTTCGTTGAGCAAGAAGGGCCTGCAGAA  
GATCTGAACGTGCAGAAGAGCTGCCATCATCCAGAAGATCATCATATGGACAGCAA  
GACCGACTACCAGGGCTTCCAGTCGATGTACAGTTCGTGACCAGCCACCTCCGCGGGG  
CTTCAACGAGTACGACTTCGTCGCCGAGAGCTTCGACCGGGACAAGACCATCGCCCTGAT  
CATGAACAGCAGCGGCGACCCGGCTGCCGAAGGGGTGGCCCTGCCGACCCGACCGC  
CTGCGTGGCTTCTCGCACCGCCCGGGACCCCATCTTCGGCAACCAGATCATCCCGGACAC  
CGCCATCCTGAGCGTGGTGGCTTCCACCAGGCTTCGGCATGTTACGACCCCTGGGCTA  
CCTCATTCGCGCTTCCGGTGGTTCCTGATGTACCGGTTTCGAGGAGGAGCTGTCTCTGCG  
GAGCCTGCAGGACTACAAGATCCAGAGCGCGCTGCTCGTGGCAGCCCTGTTACGTTCTT  
CGCCAAGAGCACCTGATCGACAAGTACGACCTGTCGAACCTGCACGAGATCGCCAGCGG  
GGGCGCCCGCTGAGCAAGGAGGTGGGCGAGGCGTGGCCAAGCGGTCCACCTCCCGGG  
CATCCGCGAGGCTACGGCTGACCGAGACCACGAGCGGATCCTGATCACCCCGGAGGG  
GGACGACAAGCCGGGCGCGTGGCAAGGTGGTCCCGTTCCTCGAGGCCAAGGTGGTGGAA  
CCTGGACACCGGCAAGACCCTGGGCGTGAACCAGCGGGGCGAGCTGTGCGTGGGGGGCC  
GATGATCATGAGCGGCTACGTGAACAACCCGGAGGCCACCAACGCCCTCATCGACAAGGA  
CGGCTGGCTGCACAGCGGGACATCGCCTACTGGGACGAGGACGAGCACTTCTTCATCGT  
CGACCGGCTGAAGTTCGCTGATCAAGTACAAGGGCTACCAGGTGGCGCCGGCCGAGCTGGA  
GAGCATCCTGCTCCAGCACCCCAACATCTTCGACGCCGGCGTGGCCGGGCTGCCGAGCA  
CGACCGCGGAGCTGCCGGCCGGTGGTGGTGGTGGAGCACGGCAAGACCATGACGGA  
GAAGGAGATCGTTCGACTACGTGGCCAGCCAGGTGACCACGCCAAGAAGCTGCCGGGCGG  
CGTGGTGTTCGTTGGAGAGGTCCCGAAGGGCCTGACCGGGAAGCTCGACGCCCGGAAGAT  
CCGCGAGATCCTGATCAAGGCCAAGAAGGGCGGCAAGATCGCCGTGTAAGACTAGTGGAA  
GCATTTTCTGATTGATGAAAAAATAACTCAGTTATGGCCATACCCCTGCTAGAAGG  
TTACAGTGTATTATGTAGCATGCAATGTGTTATGTAGTGCTTAATAAAAAATAAATGAAA  
AAAAATGCAGATCTAAAAA  
AA  
AAAAAAAAAAAAAAAAAAAA

서열 번호 17:

GGGAGAAAGCTTGAGGATGGAGGACGCCAAGAACATCAAGAAGGGCCCGGCCCTTCTA  
CCCCTGGAGGACGGGACCGCCGCGAGCAGCTCCACAAGGCCATGAAGCGGTACGCCCT  
GGTCCGGGACGATCGCCTTACCCGACGCCACATCGAGGTGACATCACCTACGCGGA  
GTACTTCGAGATGAGCGTGGCCCTGGCCGAGGCCATGAAGCGGTACGGCCTGAACACCAA  
CCACCGGATCGTGGTGTGCTCGGAGAACAGCCTGCAGTTCTTCATGCCGGTGTGGGCGC

[0328]

CCTTTCATCGGCGTGGCCGTCGCCCCGGCGAACGACATCTACAACGAGCGGGAGCTGCT
GAACAGCATGGGGATCAGCCAGCCGACCGTGGTGTTCGTGAGCAAGAAGGGCCTGCAGAA
GATCCTGAACGTGCAGAAAGCTGCCCATCATCCAGAAGATCATCATGGACAGCAA
GACCGACTACCAGGGCTTCCAGTCGATGTACACGTTTCGTGACCAGCCACCTCCCAGGGG
CTTCAACGAGTACGACTTCGTCCCGGAGAGCTTCGACCGGGACAGACCATCGCCCTGAT
CATGAACAGCAGCGGCAGCACCGGCCCTGCCAAGGGGGTGGCCCTGCCGACCCGGACCGC
CTGCGTGCCTTCGACAGCCCGGGACCCCATCTCGGCAACCAGATCATCCCGGACAC
CGCCATCCTGAGCGTGGTCCGTTCCACCAGCGCTTCGGCATGTTACAGACCCTGGGCTA
CCTCATCTGCGGCTTCGCGGTGGTCTGATGTACCGGTTTCGAGGAGGAGCTGTTCCTGCG
GAGCCTGCAGACTACAAGATCCAGAGCGCGTCTGCTGCGGACCCCTGTTACGTTCTT
CGCCAGAGCACCTGATCGACAAGTACGACCTGTCGAACCTGCACAGATCGCCAGCGG
GGGGCCCGCTGAGCAAGGAGTGGGCGAGGCGTGGCCAAAGCGTTCCACCTCCCAGG
CATCCGCGAGGGCTACGGCTGACCGAGACCAGAGCGCATCCTGATCACCCCGAGGG
GGACGACAAGCCCGGCCCTGGGCAAGGTGGTCCCGTTCGAGGGCAAGGTGGTGGGA
CCTGGACACCGGCAAGACCCTGGGCGTGAACCAGCGGGGGAGCTGTGCGTGGGGGGCC
GATGATCATGAGCGGCTACGTGAACAACCCGGAGGCCACCAACGCCCTCATCGACAAGGA
CGGCTGGCTGCACAGCGGCACATCGCCTACTGGGACGAGGACGAGCACTTCTTCATCGT
CGACCGGCTGAAGTCGCTGATCAAGTACAAGGGTACAGGTGGCGCCGGCCGAGCTGGA
GAGCATCTGCTCCAGCACCCCAATCTTCGACGCGCGCTGGCCGGGCTGCCGGACGA
CGACCGCGGGGAGCTGCCGGCCGGTGGTGGTCTGGAGCAGCGCAAGACCATGACGGA
GAAGGATCGTCGACTACGTGGCCAGCCAGGTGACCACGCCAAGAAGCTGCCGGGGCG
CGTGGTTCGTGGACAGGTCCCGAAGGGCCTGACCGGGAAGCTCGACGCCCGGAAGAT
CCGCGAGATCCTGATCAAGGCCAAGAAGGGCGGCAAGATCGCCGTGTAAGACTAGTAGAC
CAATCTTAAACATCTGATTATATTTGATTTATTTAGTGTGTTGGACCATGTGTGA
TCAGACTGCTATCTGAATAAAATAAGATTTGTCAAACCTCAGTGTTTTCCATCAGACA
CTCCATGAAAGGTCACAAATTTCTCTTGATATTAAGCTGGGTGTCTTTAAACAACCCATA
ATACAGCTCTGTTTAGCCCGCAATGGAAAGGATATATGTGGCAATATTAACCTGGTACA
TGAATATATGGGATAACATTTAATTTGAAGGTTTGGAAATATATATTTAAGCTTTAT
TTCCAGAACAGTGAAGGTTAGGTCTTGGGAAAACATAAAGTTCGCAAGTAGAAGAAATA
GTAGTACCATATGCCAAGTATAGAGATGAATCATGTCAGTAGTTAGAATAACATTTCA
ACTGTTTCTTTGTTAAATCAGAAAGCCCTATTGACAACATCTATGTCTGTAAAAA
TGTTAGAGTACTTGTATCTTGAATATAGCCTCCCAAGAGAGAACAGGGTGGTATTCTA
AGTATGTTTCTTTGTAACATCTTTAGCAGTAGGACAGGCCATACATGTGAAATCTGATT
TTTATGTGTGTTATTCGTTTGTCTGGTTTTACTACCTTTGCAAAAACAAAAACCCAAA
GATATTTAAACAAGGTTATAATTTAGCATCTTCCTGGATCTAAATAGTATATTATATCC
TGAATAAATGAAATGATTGCTATAGATCTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

서열 번호 18:
AAAAGCATCT CAGCCTACCA TGAGAATAAG AGAAAGAAAA TGAAGATCAA AAGCTTATTC
ATCTGTTTTT CTTTTTCGTT GGTGTAAAGC CAACACCCCTG TCTAAAAAAC ATAAATTTCT
TTAATCATT TGCCTTTTT CTCTGTGCTT CAATT

서열 번호 19:
CATCACATTT AAAAGCATCT CAGCCTACCA TGAGAATAAG AGAAAGAAAA TGAAGATCAA
AAGCTTATTC ATCTGTTTTT CTTTTTCGTT GGTGTAAAGC CAACACCCCTG

서열 번호 20:
AAAAGCATCT CAGCCTACCA TGAGAATAAG AGAAAGAAAA TGAAGATCAA AAGCTTATTC
ATCTGTTTTT CTTTTTCGTT GGTGTAAAGC CAACACCCCTG TCTAAAAAAC

서열 번호 21:

[0329]

CAGCCTACCA TGAGAATAAG AGAAAGAAAA TGAAGATCAA AAGCTTATTC ATCTGTTTTT  
 CTTTTTCGTT GGTGTAAAGC CAACACCCTG TCTAAAAAAC ATAAATTCTT

서열 번호 22:

TGAGAATAAG AGAAAGAAAA TGAAGATCAA AAGCTTATTC ATCTGTTTTT CTTTTTCGTT  
 GGTGTAAAGC CAACACCCTG TCTAAAAAAC ATAAATTCTT TTAATCATT

서열 번호 23:

AGAAAGAAAA TGAAGATCAA AAGCTTATTC ATCTGTTTTT CTTTTTCGTT GGTGTAAAGC  
 CAACACCCTG TCTAAAAAAC ATAAATTCTT TTAATCATT TGCCTCTTTT

서열 번호 24:

TGAAGATCAA AAGCTTATTC ATCTGTTTTT CTTTTTCGTT GGTGTAAAGC CAACACCCTG  
 TCTAAAAAAC ATAAATTCTT TTAATCATT TGCCTCTTTT CTCTGTGCTT

서열 번호 25:

AAGCTTATTC ATCTGTTTTT CTTTTTCGTT GGTGTAAAGC CAACACCCTG TCTAAAAAAC  
 ATAAATTCTT TTAATCATT TGCCTCTTTT CTCTGTGCTT CAATTAATAA

서열 번호 26:

ATCTGTTTTT CTTTTTCGTT GGTGTAAAGC CAACACCCTG TCTAAAAAAC ATAAATTCTT  
 TTAATCATT TGCCTCTTTT CTCTGTGCTT CAATTAATAA AAAATGAAA

서열 번호 27:

CAGCCTACCA TGAGAATAAG AGAAAGAAAA TGAAGATCAA AAGCTTATTC ATCTGTTTTT  
 CTTTTTCGTT GGTGTAAAGC CAACACCCTG TCTAAAAAAC ATAAATTCTT TTAATCATT  
 TGCCTCTTTT CTCTGTGCTT CAATTAATAA A

서열 번호 28:

TGAAGATCAA AAGCTTATTC ATCTGTTTTT CTTTTTCGTT GGTGTAAAGC CAACACCCTG  
 TCTAAAAAAC ATAAATTCTT TTAATCATT TGCCTCTTTT CTCTGTGCTT CAATTAATAA  
 A

서열 번호 29:

CTTTTTTCGTT GGTGTAAAGC CAACACCCTG TCTAAAAAAC ATAAATTCTT TTAATCATT  
 TGCCTCTTTT CTCTGTGCTT CAATTAATAA A

서열 번호 30:

AAGCTTATTC ATCTGTTTTT CTTTTTCGTT GGTGTAAAGC CAACACCCTG TCTAAAAAAC

서열 번호 31:

CAAAGGCTCTTTTCAGAGCCACCA

서열 번호 32:

AAACATCACA ATTAAGAACA TCTCAGCCTA CCATGAGAAC AAGAGAAATA AAATGAAGAT  
CAAAGCCTTA TTCATCTGTT TTTCTTTTTC ATTGGTATAA AGCCAACACC CTGTCTAAAA  
AACTATAAAT TTCTTTAATC ATTTTGCCTC TTTTCTCTGT GCTTCAATTA ATAAAAAATG  
GAAAGAATCT AGATCTAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA  
 AAAA

[0330]

서열 번호 33:

AAACAUCACA AUUAAGAACA UCUCAGCCUA CCAUGAGAAC AAGAGAAUA AAAUGAAGAU  
CAAAGCUUA UUCAUCUGUU UUUCUUUUUC AUUGGUUUA AGCCAACACC CUGUCUAAAA  
AACUAUAAAU UUCUUUAUC AUUUUGCCUC UUUUCUCUGU GCUUCAUUA AUAAAAAUG  
GAAAGAAUCU AGAUCUAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA  
 AAAAA

서열 번호 34:

ACACATCACA ACCACAACCT TCTCAGGCTA CCCTGAGAAA AAAAGACATG AAGACTCAGG  
ACTCATCTTT TCTGTTGGTG TAAATCAAC ACCCTAAGGA ACACAAATTT CTTTAAACAT  
TTGACTTCTT GTCTCTGTGC TGCAATTAAT AAAAAATGGA AAGAATCTAC AGATCTAAAA  
 AAAAA

서열 번호 35:

ACACAUCACA ACCACAACCU UCUCAGGCUA CCCUGAGAAA AAAAGACAUG AAGACUCAGG  
ACUCAUCUUU UCUGUUGGUG UAAAAUCAAC ACCCUAAGGA ACACAAUUUU CUUUAACAU  
UUGACUUCUU GUCUCUGUGC UGCAUUUAU AAAAAAUGGA AAGAAUCUAC AGAUCUAAAA  
 AAAAA

서열 번호 36: AAACATCACA ATTAAGAACA TCTCAGCCTA CCATGAGAAC AAGAGAAATA

AAATGAAGAT CAAAAGCTTA TTCATCTGTT TTTCTTTTC ATTGGTATAA AGCCAACACC  
CTGTCTAAAA AACTATAAAT TTCTTTAATC ATTTTGCTC TTTTCTCTGT GCTTCAATTA  
ATAAAAAATG GAAAGAATCT AGATCTAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA  
 AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA

서열 번호 37:

AAACAUCACA AUUAAGAACA UCUCAGCCUA CCAUGAGAAC AAGAGAAUA AAAUGAAGAU  
CAAAGCUUA UUCAUCUGUU UUUCUUUUUC AUUGGUUUA AGCCAACACC CUGUCUAAAA  
AACUAUAAAU UUCUTTAATC ATTTTGCTC TTTTCTCTGT GCTTCAATTA ATAAAAATG  
GAAAGAATCT AGATCTAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA  
 AAAAAAAAAA AAAAAAAAAA

서열 번호 38:

ACACATCACA ACCACAACCT TCTCAGGCTA CCCTGAGAAA AAAAGACATG AAGACTCAGG  
ACTCATCTTT TCTGTTGGTG TAAATCAAC ACCCTAAGGA ACACAAATTT CTTTAAACAT

[0331]

TTGACTTCTT GTCTCTGTGC TGCAATTAAT AAAAAATGGA AAGAATCTAC AGATCTAAAA  
AAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA

서열 번호 39:

ACACAUCACA ACCACAACCU UCUCAGGCUA CCCUGAGAAA AAAAGACAUG AAGACUCAGG  
ACUCAUCUUU UCUGUUGGUG UAAAAUCAAC ACCCUAAGGA ACACAAAUUU CUUUAACAU  
UUGACUUCUU GUCUCUGUGC UGCAAUUAAU AAAAAUUGGA AAGAAUCUAC AGAUCUAAAA  
AAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA

서열 번호 40:

GGGAGAAAGC TTGAGGATGG AGGACGCCAA GAACATCAAG AAGGCCCGG CGCCCTTCTA  
CCCCTGGAG GACGGGACCG CCGGCGAGCA GCTCCACAAG GCCATGAAG GGTACGCCCT  
GGTGCCGGGG ACGATCGCCT TCACCGACGC CCACATCGAG GTCGACATCA CCTACGGCGA  
GTACTTCGAG ATGAGCGTGC GCCTGGCCGA GGCCATGAAG CGGTACGGCC TGAACACCAA  
CCACCGGATC GTGGTGTGCT CGGAGAACAG CCTGCAGTTC TTCATGCCGG TGCTGGGGCC  
CCTTTCATC GCGTGGCCG TCGCCCCGC GAACGACATC TACAACGAGC GGGAGCTGCT  
GAACAGCATG GGGATCAGCC AGCCGACCGT GGTGTTCTGT AGCAAGAAGG GCCTGCAGAA  
GATCCTGAAC GTGCAGAAGA AGCTGCCCAT CATCCAGAAG ATCATCATCA TGGACAGCAA  
GACCGACTAC CAGGCTTCC AGTCGATGTA CACGTTCTGT ACCAGCCACC TCCCGCCGGG  
CTTCAACGAG TACGACTTCG TCCCGGAGAG CTTGACCCGG GACAAGACCA TCGCCCTGAT  
CATGAACAGC AGCGGCAGCA CCGGCTGCC GAAGGGGGTG GCCCTGCCGC ACCGGACCCG  
CTGCGTGGCC TTCTCGCAG CCCGGGACCC CATCTCGGC AACCAGATCA TCCCGGACAC  
CGCCATCCTG AGCGTGGTGC CGTTCACCA CGGCTTCGGC ATGTTACGA CCTGGGCTA  
CCTCATCTGC GGCTCCGGG TGGTCTGAT GTACCGGTTG GAGGAGGAGC TGTTCCTGCC  
GAGCCTGCAG GACTACAAGA TCCAGAGGC GCTGCTCGTG CCGACCCTGT TCAGCTTCTT  
CGCCAAGAGC ACCCTGATCG ACAAGTACGA CCTGTGGAAC CTGCACGAGA TCGCCAGCGG  
GGCGCCCGG CTGAGCAAGG AGGTGGGCGA GGCCGTGGCC AAGCGGTTCC ACCTCCCGG  
CATCCGCAG GGCTACGGCC TGACCGAGAC CACGAGCGG ATCCTGATCA CCCCAGAGG  
GGACGACAAG CCGGGCGCC TGGGCAAGGT GGTCCCGTTC TTCGAGGCCA AGGTGGTGG  
CCTGGACACC GGCAAGACC TGGGCGTGAA CCAGCGGGC GAGCTGTGCG TCGGGGGCC  
GATGATCATG AGCGGCTACG TGAACAACCC GGAGGCCACC AACGCCCTCA TCGACAAGGA  
CGGCTGGCTG CACAGCGCG ACATCGCCTA CTGGGACGAG GACGAGCACT TCTTCATCGT  
CGACCGGCTG AAGTCGCTGA TCAAGTACAA GGGCTACCAG GTGGCGCCGG CCGAGCTGGA  
GAGCATCCTG CTCCAGCACC CCAACATCTT CGACGCCGGC GTGGCCGGG TGCCGGACGA  
CGACGCCGGC GAGCTGCCGG CCGCGGTGGT GGTGCTGGAG CACGGCAAGA CCATGACGGA  
GAAGGAGATC GTCGACTACG TGGCCAGCCA GGTGACCACC GCCAAGAAGC TGCGGGGGCG

[0332]

CGTGGTGTTC GTGGACGAGG TCCCGAAGGG CCTGACCGGG AAGCTCGACG CCCGGAAGAT  
CCGCGAGATC CTGATCAAGG CCAAGAAGGG CGGCAAGATC GCCGTGTAAG ACTAGTAAAC  
ATCACAATTA AGAACATCTC AGCCTACCAT GAGAACAAGA GAAATAAAAT GAAGATCAAA  
AGCTTATTCA TCTGTTTTTC TTTTTCATTG GTATAAAGCC AACACCCTGT CTAAAAACT  
ATAAATTCT TTAATCATTT TGCCCTCTTT CTCTGTGCTT CAATTAATAA AAAATGGAAA  
GAATCTAGAT CTAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA  
AAAAAAAAA AAAAAA

서열 번호 41:

GGGAGAAAGC TTGAGGATGG AGGACGCCAA GAACATCAAG AAGGGCCCGG CGCCCTTCTA  
 CCCGCTGGAG GACGGGACCG CCGGCGAGCA GCTCCACAAG GCCATGAAGC GTTACGCCCT  
 GGTGCCGGGG ACGATCGCCT TCACCGACGC CCACATCGAG GTCGACATCA CCTACGCGGA  
 GTACTTCGAG ATGAGCGTGC GCCTGGCCGA GGCCATGAAG CGGTACGGCC TGAACACCAA  
 CCACCGGATC GTGGTGTGCT CGGAGAACAG CCTGCAGTTC TTCATGCCGG TGCTGGGCGC  
 CCTTTCATC GGGGTGGCCG TCGCCCGGC GAACGACATC TACAACGAGC GGGAGCTGCT  
 GAACAGCATG GGGATCAGCC AGCCGACCGT GGTGTTCTGT AGCAAGAAGG GCCTGCAGAA  
 GATCCTGAAC GTGCAGAAGA AGCTGCCCAT CATCCAGAAG ATCATCATCA TGGACAGCAA  
 GACCGACTAC CAGGGCTTCC AGTCGATGTA CACGTTCTGT ACCAGCCACC TCCC GCCGGG  
 CTTCAACGAG TAGGACTTCG TCCCGGAGAG CTTGACCGG GACAAGACCA TCGCCCTGAT  
 CATGAACAGC AGCGGCAGCA CCGGCCTGCC GAAGGGGGTG GCCCTGCCGC ACCGGACCGC  
 CTGCGTGGCG TTCTCGCAGC CCCGGGACCC CATCTTCGGC AACAGATCA TCCCGGACAC  
 CGCCATCCTG AGCGTGGTGC CGTTCACCA CGGCTTCGGC ATGTTACGA CCCTGGGCTA  
 CCTCATCTGC GGCTCCGGG TGGTCTGAT GTACCGGTTT GAGGAGGAGC TGTTCTGCG  
 GAGCCTGCAG GACTACAAGA TCCAGAGCGC GCTGCTCGTG CCGACCCTGT TCAGCTTCTT  
 CGCCAAGAGC ACCCTGATCG ACAAGTACGA CCTGTCGAAC CTGCACGAGA TCGCCAGCGG  
 GGGCGCCCCG CTGAGCAAGG AGGTGGGCGA GGCCGTGGCC AAGCGGTTCC ACCTCCCGGG  
 CATCCGCCAG GGTAACGGCC TGACCGAGAC CACGAGCGCG ATCCTGATCA CCCCCGAGG  
 GGACGACAAG CCGGGCGCCG TGGGCAAGGT GGTCCCGTTC TTCGAGGCCA AGGTGGTGGA  
 CCTGGACACC GGCAAGACC TGGGCGTGAA CCAGCGGGG GAGCTGTGCG TCGGGGGGCC  
 GATGATCATG AGCGGTACG TGAACAACCC GGAGGCCACC AACGCCCTCA TCGACAAGGA  
 CGGCTGGCTG CACAGCGCGC ACATCGCCTA CTGGGACGAG GACGAGCACT TCTTCATCGT  
 CGACCGGCTG AAGTCGCTGA TCAAGTACAA GGGCTACCAG GTGGCGCCGG CCGAGCTGGA  
 GAGCATCCTG CTCCAGCACC CCAACATCTT CGACGCCGGC GTGGCCGGGC TGCCGGACGA  
 CGACGCCGGC GAGCTGCCGG CCGCGGTGGT GGTGCTGGAG CACGGCAAGA CCATGACGGA  
 GAAGGAGATC GTCGACTACG TGGCCAGCCA GGTGACCACC GCCAAGAAGC TGCGGGGCGG

[0333]

CGTGGTGTTC GTGGACGAGG TCCCGAAGGG CCTGACCGGG AAGCTCGACG CCCGGAAGAT  
 CCGCGAGATC CTGATCAAGG CCAAGAAGGG CGGCAAGATC GCCGTGTAAG ACTAGTACAC  
 ATCACAACCA CAACCTTCTC AGGCTACCCT GAGAAAAAAAA GACATGAAGA CTCAGGACTC  
 ATCTTTTCTG TTGGTGTAFAA ATCAACACCC TAAGGAACAC AAATTTCTTT AACATTTGA  
 CTTCTTGCTCT CTGTGCTGCA ATTAATAAAA AATGGAAAGA ATCTACAGAT CTAAAAAAA  
 AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAA

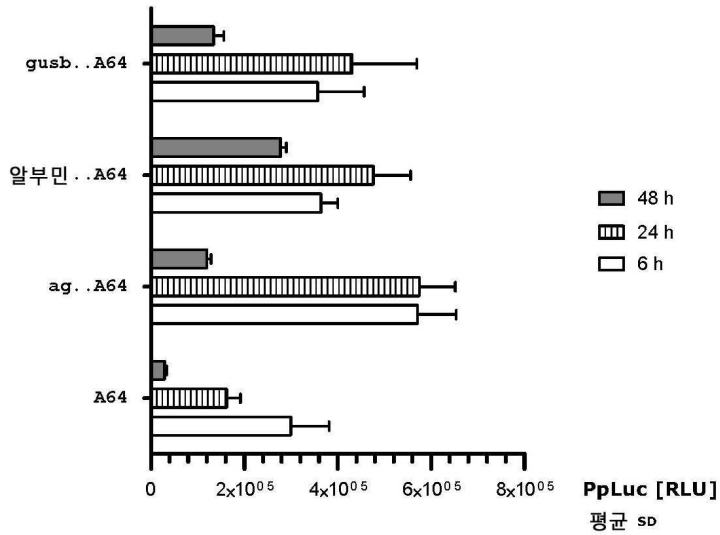
[0334]

도면

도면1

상이한 3'-UTR을 가지는 mRNA들로부터의 PpLuc(HeLa)

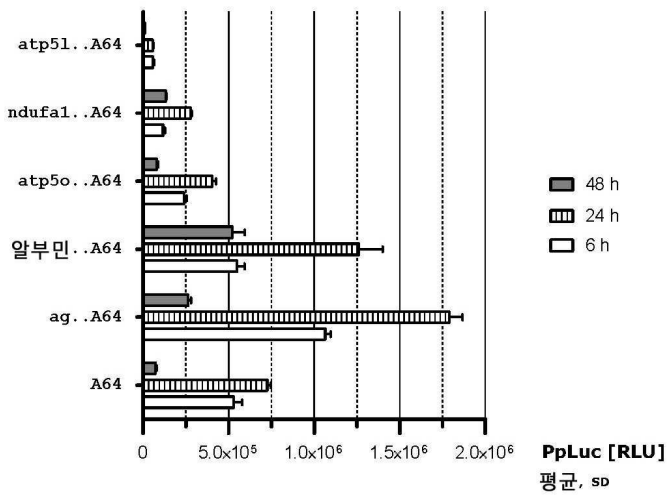
PpLuc(GC)..



도면2

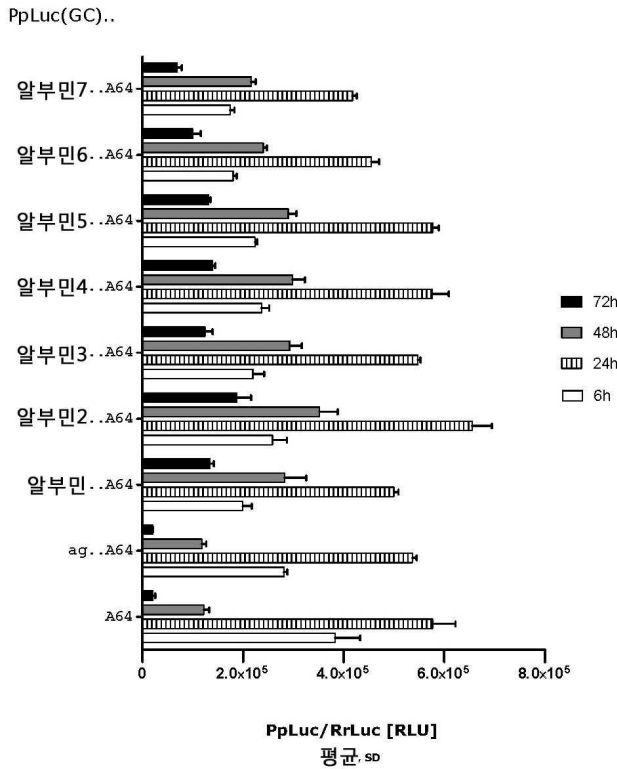
상이한 3'-UTR을 가지는 mRNA들로부터의 PpLuc(HeLa)

PpLuc(GC) ..



도면3

상이한 알부민 3'-UTR 변이체들을 가지는 mRNA로부터의 PpLuc(HeLa)



도면4

3'-UTR 결실 PpLuc(GC) - A64

5'-GGGAGAAAGCTTGAGGATGGAGGACGCCAAGAACATCAAGAAGGGCCCGGCGCCCTTCTA  
 CCCGCTGGAGGACGGGACCGCCGGCGAGCAGCTCCACAAGGCCATGAAGCGGTACGCCCT  
 GGTGCCGGGACGATCGCCTTACCGAGCCACATCGAGGTCGACATCACCTACGCGGA  
 GTACTTCGAGATGAGCCTGCGCTGGCCGAGGCCATGAAGCGGTACGGCCTGAACACCAA  
 CCACCGGATCGTGGTGTGCTCGGAGAACAGCCTGCAGTTCCTCATGCCGGTGTGGGCGC  
 CCTCTTCATCGGCGTGGCCGTGCCCCGGGGAACGACATCTACAACGAGCGGGAGCTGCT  
 GAACAGCATGGGGATCAGCCAGCCGACCCGTGGTGTTCGTGAGCAAGAAGGGCCTGCAGAA  
 GATCCTGAACGTGCAGAAGAAGCTGCCCATCATCCAGAAGATCATCATCATGGACAGCAA  
 GACCGACTACCAGGGCTTCAGTCGATGTACACGTTGCTGACCAGCCACCTCCCGCCGGG  
 CTTCAACGAGTACGACTTCGTCCCGGAGAGCTTCGACCCGGGACAAGACCATCGCCCTGAT  
 CATGAACAGCAGCGGCAGCACCCGGCCTGCCGAAGGGGGTGGCCCTGCCGCACCGGACCGC  
 CTGCGTGCCTTCTCGCACGCCCGGGACCCCATCTTCGGCAACCAGATCATCCCGGACAC  
 CGCCATCCTGAGCGTGGTCCGTTCCACCAGGCTTCGGCATGTTACGACCCCTGGGCTA  
 CCTCATCTGCGGCTTCGGGTGGTCTGATGTACCGGTTGAGGAGGAGCTGTTCTCGG  
 GAGCCTGCAGGACTACAAGATCCAGAGCGCGCTGCTCGTGCCGACCCCTGTTAGCTTCTT  
 CGCAAGAGCACCTGATCGACAAGTACGACCTGTGCAACCTGCACGAGATCGCCAGCGG  
 GGGCGCCCGCTGAGCAAGGAGGTGGGCGAGGCCGTGGCCAAAGCGGTTCCACCTCCCGGG  
 CATCCGCCAGGGCTACGGCCTGACCGAGACCAGAGCGGATCCTGATCACCCCGAGGG  
 GGACGACAAGCCGGGCGCGTGGCAAGGTGGTCCCGTTCTTCGAGGCCAAGGTGGTGG  
 CCTGGACACCGGCAAGACCCTGGCGTGAACCCAGCGGGCGAGCTGTGCGTCCGGGGCC  
 GATGATCATGAGCGGCTACGTGAACAACCCGGAGGCCACCAACGCCCTCATCGACAAGGA  
 CGCTGGCTGCACAGCGGCAGCATCGCTACTGGGACGAGGACGAGCACTTCTTCATCGT  
 CGACCGGCTGAAGTCCGTGATCAAGTACAAGGGCTACCAGGTGGCGCCGGCCGAGCTGGA  
 GAGCATCTGCTCCAGCACCCCAACATCTTCGACGCGCGGCGTGGCGGGGTGCCGGACGA  
 CGACGCCGGGAGCTGCCGGCCGCGGTGGTGGTGGTGGAGCACGGCAAGACCATGACGGA  
 GAAGGAGATCGTACGACTACGTGGCCAGCCAGGTGACCAACCGCAAGAAGCTGCGGGGCGG  
 CGTGGTGTTCGTGGACGAGGTCCCAGGGGCTGACCGGGAAGCTCGACGCCCGGAAGAT  
 CCGCGAGATCCTGATCAAGGCCAAGAAGGGCGGCAAGATCGCCGTGTAAGACTAGTAGAT  
 CTAAA  
 AAAAAA-3'

도면5

PpLuc(GC) - 알부민 - A64

5'-GGGAGAAAGCTTGAGGATGGAGGACGCCAAGAACATCAAGAAGGGCCCGGCGCCCTTCTA  
 CCCGCTGGAGGACGGGACCGCCGCGAGCAGCTCCACAAGGCCATGAAGCGGTACGCCCT  
 GGTGCCGGGCACGATCGCCTTCAACCGACGCCACATCGAGGTGACATCACCTACGCCGA  
 GTACTTCGAGATGAGCGTGCCTGGCCGAGGCCATGAAGCGGTACGCCCTGAACACCAA  
 CCACCGGATCGTGGTGTCTCGGAGAACAGCCTGCAGTTCCTCATGCCGGTGTGGGCGC  
 CCTCTTCATCGGCGTGGCCGTCGCCCGGCGAACGACATCTACAACGAGCGGAGCTGCT  
 GAACAGCATGGGGATCAGCCAGCCGACCGTGGTGTTCGTGAGCAAGAAGGGCCTGCAGAA  
 GATCCTGAACGTTTCGACCGCCCGGACCCATCTCCAGAAGATCATCATCATGACAGCAA  
 GACCAGTACCAGGGCTTCAGTCGATGTACAGTTCGTGACCAGCCACCTCCCGCCGGG  
 CTTCAACGAGTACGACTTCGTCCCGGAGAGCTTCGACCGGGACAAGACCATCGCCCTGAT  
 CATGAACAGCAGCGGCAGCACCGGCCCTGCCAAGGGGGTGGCCCTGCCGACCCGGACCGC  
 CTGCGTGCCTTTCGACCGCCCGGACCCCATCTTCGGCAACCAGATCATCCCGGACAC  
 CGCCATCCTGAGCGTGGTCCGCTCCACCACGGCTTCGGCATGTTACGACCCTGGGCTA  
 CCTCATCTGCGGCTTCGGGTGGTCTGATGTACCGGTCGAGGAGGAGCTGTTCTTCTGCG  
 GAGCCTGCAGGACTACAAGATCCAGAGCGCGTGTCTGTCGCCACCTGTTTCTGCTTCTT  
 CGCCAAGAGCACCTGATCGACAAGTACGACCTGTGCAACCTGCACGAGATCGCCAGCGG  
 GGGCGCCCGCTGAGCAAGGAGGTGGGCGAGGCCGTGGCCAAGCGGTTCCACCTCCCGGG  
 CATCCCGCAGGGCTACGGCTGACCGAGACCAGGCGCGATCCTGATCACCCCGAGGG  
 GGACGACAAGCCGGGCGCCGTGGGCAAGGTGGTCCCGTTCTTCGAGGCCAAGGTGGTGG  
 CCTGGACACCGGCAAGACCTGGGCGTGAACCAGCGGGCGAGCTGTGCGTGCGGGGGCC  
 GATGATCATGAGCGGCTACGTGAACAACCCGGAGGCCACCAACGCCCTCATCGACAAGGA  
 CGGCTGGCTGCACAGCGCGACATCGCCTACTGGGACGAGGACGAGCACTTCTTCATCGT  
 CGACCGGCTGAAGTCGCTGATCAAGTACAAGGGCTACCAGGTGGCGCGGCGAGCTGGA  
 GAGCATCCTGCTCCAGCACCCCAACATCTTCGACGCGCGGCGTGGCCGGGCTGCCGGA  
 CGACGCGCGGAGCTGCCGCGCGGTTGGTGGTGTGGAGCACGGCAAGACCATGACGGA  
 GAAGGAGATCGTCACTACGTGGCCAGCCAGGTGACCACCGCAAGAAGCTGCCGGGCGG  
 CGTGGTGTTCGTGGACGAGTCCCAGGGCCCTGACCGGGAAGCTCGACGCCCGGAAGAT  
 CCGCGAGATCCTGATCAAGGCCAAGAAGGGCGCAAGATCGCCGTGTAAGACTAGTGCA  
 CACATTTAAAGCATCTCAGCCTACCATGAGAATAAGAGAAAGAAATGAAGATCAAAAG  
 CTATTCATCTGTTTTCTTTTTCTGTTGGTGTAAAGCCAACCCCTGTCTAAAAACATA  
 AATTTCTTAAATCATTTTGCCTCTTTTCTCTGTGCTTCAATTAATAAAAAATGGAAAGAA  
 TCTAGATCTAAAAA  
 AA  
 AAAAAAAAAAAAAA-3'

도면6

PpLuc(GC) - 알부민2 - A64

5'-GGGAGAAAGCTTGAGGATGGAGGACGCCAAGAACATCAAGAAGGGCCCGGCGCCCTTCTA  
 CCCGCTGGAGGACGGGACCGCCGGCGAGCAGCTCCACAAGGCCATGAAGCGGTACGCCCT  
 GGTGCCGGGCACGATCGCCTTACCCGACGCCCACATCGAGGTCGACATCACCTACGCGGA  
 GTACTTCGAGATGAGCGTGCCTGGCCGAGGCCATGAAGCGGTACGGCTGAACACCAA  
 CCACCGGATCGTGGTGTCTCGGAGAACAGCCTGCAGTTCCTCATGCCGGTGTGGGCGC  
 CCTTTCATCGGCGTGGCCGTGCCCCGGCGAACGACATCTAACGAGCGGGAGCTGCT  
 GAACAGCATGGGGATCAGCCAGCCGACCGTGGTGTTCGTGAGCAAGAAGGGCCTGCAGAA  
 GATCCTGAACGTGCAGAAGAAGTGCCTCATCCAGAAGATCATCATGACAGCAA  
 GACCAGTACCAGGGCTTCAGTCGATGTACACGTTTCGTGACCAGCCACTCCCGCGGG  
 CTTCAACAGTACGACTTCGTCCCGGAGAGCTTCGACCGGGACAAGACCATCGCCTGAT  
 CATGAACAGCAGCGGCAGCACCGGCTGCCGAAGGGGTGGCCCTGCCGACCCGGACCGC  
 CTGCGTGCCTTTCGACAGCCCGGGACCCCATCTTCGCAACCAGATCATCCCGACAC  
 CGCCATCCTGAGCGTGGTCCGTTCCACCAGGCTTCGGCATGTTACGACCCTGGGCTA  
 CCTCATCTGCGGCTTCGCGGTGGTCTGATGTACCGGTTGAGGAGGAGCTGTTCTGCG  
 GAGCCTGCAGGACTACAAGATCCAGAGCGCGTGTCTCGTCCGACCCCTGTTAGCTTCTT  
 CGCCAAGAGCACCCCTGATCGACAAGTACGACCTGTCGAACCTGCACGAGATCGCCAGCGG  
 GGGCGCCCCGTGAGCAAGGAGGTGGGCGAGGCCGTGGCCAAGCGGTTCCACCTCCCGGG  
 CATCCGCCAGGGCTACGGCCTGACCGAGACCACGAGCGCGATCCTGATCACCCCGAGGG  
 GGACGACAAGCCGGCGCCGTGGGCAAGGTGGTCCCGTTCCTCGAGGCCAAGGTGGTGGGA  
 CCTGGACACCGGC AAGACCTGGGCGTGAACCAAGCGGGCGAGCTGTGCGTGGGGGGCC  
 GATGATCATGAGCGGTACGTGAACAACCCGGAGGCCACCAACGCCCTCATCGACAAGGA  
 CGCTGGCTGCACAGCGCGACATCGCCTACTGGGACGAGGACGAGCACTTCTTCATCGT  
 CGACCGGTGAAGTCGCTGATCAAGTACAAGGGTACCAGGTGGCGCCGGCCGAGCTGGA  
 GAGCATCTGCTCCAGCACCCCAACATCTTCGACGCGCGGTGGCCGGGCTGCCGGACGA  
 CGACGCCGGCGAGCTGCCGGCCGGTGGTGGTGTGGAGCACGGCAAGACCATGACGGGA  
 GAAGGAGATCGTCGACTACGTGGCCAGCCAGGTGACCACCGCCAAGAAGCTGCCGGGCGG  
 CGTGGTGTTCGTGGACGAGGTCCCGAAGGGCCTGACCGGGAAGTTCGACGCGCCGGAAGAT  
 CCGCGAGATCCTGATCAAGCC AAGAAGGGCGGCAAGATCGCCGTGTAAGACTAGTGCA  
 CACATTTAAAAGCATCTCAGCCTACCATGAGAA TAAGAGAAA GAAAATGAAGATCAAAG  
 CTTATTCGCTGTTTTTCTTTTTCGTTGGTGTAAAGCCAACACCTGTCTAAAAAACATA  
 AATTTCTTTAATCATTTTGCCTCTTTTCTCTGTGCTTCAATTAATAAAAAATGGAAAGAA  
 TCTAGATCTAAA  
 AAAAAAAAAAAAAA-3'

도면7

PpLuc(GC) - 알부민3 - A64

5'-GGGAGAAAGCTTGAGGATGGAGGACGCCAAGAACATCAAGAAGGGCCCGGCGCCCTTCTA  
 CCCGCTGGAGGACGGGACCGCCGGCGAGCAGCTCCACAAGGCCATGAAGCGGTACGCCCT  
 GGTGCCGGGCACGATCGCCTTCAACCGACGCCACATCGAGGTCGACATCACCCTACCGGGA  
 GTACTTCGAGATGAGCGTGCCTGGCCGAGGCCATGAAGCGGTACGGCCTGAACACCAA  
 CCACCGGATCGTGGTGTGCTCGGAGAACAGCCTGCAGTTCTTCATGCCGGTGTGGGCGC  
 CCTTTCATCGCGTGGCCGTGCCCCGGCGAACGACATCTACAACGAGCGGGAGCTGCT  
 GAACAGCATGGGATCAGCCAGCCGACCGTGGTGTTCGTGAGCAAGAAGGGCTGCAGAA  
 GATCCTGAACGTGCAGAAGAAGCTGCCCATCATCCAGAAGATCATCATCATGGACAGCAA  
 GACCACTACAGGGCTTCCAGTCGATGTACACGTTTCGTGACCAGCCACCTCCCGCGGG  
 CTTC AACGAGTACGACTTCGTCCCGGAGAGCTTCGACCGGGACAAGACCATCGCCCTGAT  
 CATGAACAAGCGCGCAGCACCGGCTGCGGAAGGGGGTGGCCCTGCCGACCGGACCGC  
 CTGCGTGCCTTTCGCACGCCCGGACCCCATCTTCGGCAACCAGATCATCCCGGACAC  
 CGCCATCCTGAGCGTGGTCCGTTCCACCACGGCTTCGGCATGTTACGACCCCTGGGCTA  
 CCTCATCTGCGGCTTCGGGTGGTCTGATGTACCGGTTTCGAGGAGGAGCTGTTCTGCG  
 GAGCCTGCAGGACTACAAGATCCAGAGCGCGCTGCTCGTGCCGACCTGTTTCAGCTTCTT  
 CGCCAAGAGCACCTGATCGACAAGTACGACCTGTGCAACCTGCACGAGATCGCCAGCGG  
 GGGCGCCCGCTGAGCAAGGAGGTGGGCGAGGCGTGGCCAAGCGGTTCCACCTCCCGGG  
 CATCCGCCAGGGCTACGGCCTGACCGAGACCACGAGCGGATCCTGATCACCCCGAGGG  
 GGACGACAAGCCGGCGCCTTGGGCAAGGTGGTCCGTTCTTCGAGGCAAGGTGGTGGGA  
 CCTGGACACCGGCAAGACCCTGGGCGTGAACACGCGGGGCGAGCTGTGCGTGGGGGGCC  
 GATGATCATGAGCGGTACGTGAACAACCCGGAGGCCACC AACGCCTCATCGACAAGGA  
 CGGCTGGCTGCACAGCGGCACATCGCCTACTGGGACGAGGACGAGCATTCTTCATCGT  
 CGACCGGCTGAAGTCGCTGATCAAGTACAAGGGTACCAGGTGGCGCCGGCGGAGCTGGA  
 GAGCATCCTGCTCCAGCACCCCAACATCTTCGACGCGGGCGTGGCCGGGCTGCCGGACGA  
 CGACGCCGGGAGCTGCCGCGCGGTGGTGGTGTGGAGCACGGCAAGACCATGACGGA  
 GAAGGAGATCGTGCCTACGTGGCCAGCCAGGTGACCACCGCCAAGAAGCTGCGGGGCGG  
 CGTGGTGTTCGTGGACGAGTCCC GAAGGGCTGACCGGGAAGCTCGACGCCCGGAAGAT  
 CCGCGAGATCCTGATCAAGGCCAAGAAGGGCGCAAGATCGCCGTGTAAGACTAGTGAT  
 CACATTTAAAGCATCTCAGCCTACCATGAGAATAAGAGAAAGAAAATGAAGATCAAAG  
 CTTATTCATCAGTTTTTCTTTTTCGTGGTGTAAAGCCAACACCTGTCTAAAAACATA  
 AATTTCTTTAATCATTTTGCTCTTTCTCTGTGCTTCAATTAATAAAAAATGGAAGAA  
 TCTAGATCTAAAAAATAA  
 AAAAAAAAAAAAAA-3'

도면8

PpLuc(GC) - 알부민4 - A64

5'-GGGAGAAAGCTTGAGGATGGAGGACGCCAAGAACATCAAGAAGGGCCCGGCGCCCTTCTA  
 CCGCTGGAGGACGGGACCGCCGGCGAGCAGCTCCACAAGGCCATGAAGCGGTACGCCCT  
 GGTGCCGGGCAGGATCGCTTTCACCGACGCCACATCGAGGTCGACATCAGCTACGCGGA  
 GTACTTCGAGATGAGCGTGCCTGGCCGAGGCCATGAAGCGGTACGCCCTGAACACCAA  
 CCACCGGATCGTGGTGTGCTCGGAGAACAGCCTGCAGTTCCTCATGCCGGTGTGGGCGC  
 CCTCTTCATCGGCGTGGCCGTGCCCCGGCGAACGACATCTACAACGAGCGGGAGCTGCT  
 GAACAGCATGGGATCAGCCAGCCGACCGTGGTGTTCGTGAGCAAGAAGGGCTGCAGAA  
 GATCCTGAACGTGCAGAAGAAGCTGCCATCATCCAGAAGATCATCATCATGGACAGCAA  
 GACCGACTACCAGGGCTTCAGTCGATGTACACGTTTCGTGACAGCCACCTCCCGCGGG  
 CTTCAACGAGTACGACTTCGTCCCGGAGAGCTTCGACCGGGACAAGACCATCGCCCTGAT  
 CATGAACAGCAGCGGCAGCACCGCCCTGCCGAAGGGGGTGGCCCTGCCGCAACCGACCGC  
 CTGCGTGGCTTCTCGCACGCCCGGGACCCCATCTTCGGCAACCAGATCATCCCGACAC  
 CGCCATCCTGAGCGTGGTGCCTTCCACCACGGCTTCGGCATGTTACAGACCTCGGGCTA  
 CCTCATCTGCGGCTTCCGGTGGTCTGATGTACCGGTTGAGGAGGAGCTGTTCTGCG  
 GAGCCTGCAGGACTACAAGATCCAGAGCGCGCTGCTCGTCCGACCCCTGTTGAGCTTCTT  
 CGCCAAGAGCACCCCTGATCGACAAGTACGACCTGTGAACTGCACGAGATCGCCAGCGG  
 GGGCGCCCGCTGAGCAAGGAGTGGGCGAGGCGGTGGCCAAGCGGTTCCACCTCCGGG  
 CATCCGCGAGGGCTACGGCCTGACCGAGACCACGAGCGGATCCTGATCACCCCGAGGG  
 GGACGACAAGCCGGCGCGTGGCAAGGTGGTCCCGTTCGAGGCCAAGGTGGTGGGA  
 CCTGGACACCGCAAGACCCCTGGGCGTGAACCAGCGGGCGAGCTGTGCGTGGGGGGCC  
 GATGATCATGAGCGGCTACGTGAACAACCCGGAGGCCACCAACGCCCTCATCGACAAGGA  
 CGGCTGGCTGCACAGCGGCGACATCGCCTACTGGGACGAGGACGAGCACTTCTTCATCGT  
 CGACCGGCTGAAGTCGTGATCAAGTACAAGGGTACCAGGTGGCGCCGGCCGAGCTGGA  
 GAGCATCCTGCTCCAGCACCCCAACATCTTCGACGCCCGGCTGGCCGGGCTGCCGACGA  
 CGACGCCGGCGAGCTGCCGGCCGGTGGTGGTGGTGGAGCACGGCAAGACCATGACGGA  
 GAAGGAGATCGTCGACTACGTGGCCAGCCAGGTGACCACCGCAAGAAGCTGCCGGGCGG  
 CGTGGTGTTCGTGGACGAGGTCCCGAAGGGCCTGACCGGAAGCTCGACGCCCGGAAGAT  
 CCGCGAGATCCTGATCAAGGCCAAGAAGGGCGGCAAGATCGCCGTGTAAGACTAGTGCA  
 CACATTTAAAAGCATCTCAGCCTACCATGAGAA TAAGAGAAAGAAAATGAAGATCAAAG  
 CTTATTCAICTCTTTTTCTTTTCGTTGGTGTAAAGCCAACCCCTGTCTAAAAAACATA  
 AATTTCTTTAATCATTTTGCCTCTTTTCTCTGTGCTTCAATTAATAAAAAATGGAAAGAA  
 TCTAGATCTAAAAAATAA  
 AAAAAAAAAAAAAA-3'

도면9

PpLuc(GC) - 알부민5 - A64

5'-GGGAGAAAGCTTGAGGATGGAGGACGCCAAGAACATCAAGAAGGGCCCGGCGCCCTTCTA  
 CCCGCTGGAGGACGGGACCGCCGGCGAGCAGCTCCACAAGGCCATGAAGCGGTACGCCCT  
 GGTGCCGGGCACGATCGCCTTCACCGACGCCACATCGAGGTCGACATCACCTACGCGGA  
 GTACTTCGAGATGAGCGTGGCCCTGGCCGAGGCCATGAAGCGGTACGGCCTGAACACCAA  
 CCACCGGATCGTGGTGTGCTCGGAGAACAGCCTGCAGTTCTTCATGCCGGTCTGGGCGC  
 CCTCTTCATCGGCGTGCCGTGCCTCCCGCGGAACGACATCTACAACGAGCGGGAGCTGCT  
 GAACAGCATGGGATCAGCCAGCCGACCGTGGTGTTCGTGAGCAAGAAGGGCTGCAGAA  
 GATCCTGAACGTGCAGAAGAAGCTGCCCATCATCCAGAAGATCATCATCATGGACAGCAA  
 GACCGACTACCAGGGCTTCCAGTCGATGTACACGTTCTGTGACCAGCCACCTCCCGCCGGG  
 CTTCAACGAGTACGACTTCGTCCCGGAGAGCTTCGACCGGGACAAGACCATCGCCCTGAT  
 CATGAACAGCAGCGCAGCACCGCCCTGCGCAAGGGGGTGGCCCTGCCCGACCGGACCGC  
 CTGCGTGGCGTCTTCGCACGCCCGGACCCCATCTTCGGCAACCAGATCATCCCGGACAC  
 CGCCATCCTGAGCGTGGTGGCGTTCCACCACGGCTTCGGCATGTTACAGACCTGGGCTA  
 CCTCATCTGCGGCTTCGGGTGGTCTGATGTACCGGTTTCGAGGAGGAGCTGTTCTGCG  
 GAGCCTGCAGGACTACAAGATCCAGAGCGCGCTGCTCGTGCCGACCCGTTCAGCTTCTT  
 CGCCAAGAGCACCCCTGATCGACAAGTACGACCTGTGCAACCTGCACGAGATCGCCAGCGG  
 GGGCGCCCGCTGAGCAAGGAGGTGGGCGAGGCCGTGGCCAAGCGGTTCCACCTCCCGGG  
 CATCCGCCAGGGCTACGGCCTGACCAGACCACGAGCGGATCCTGATCACCCCGAGGG  
 GGACGACAAGCCGGGCGCCGTGGCAAGGTGGTCCCGTTCTTCGAGGCCAAGGTGGTGA  
 CCTGGACACCGGCAAGACCCTGGGCGTGAACCAGCGGGCGAGCTGTGCGTGGGGGGCC  
 GATGATCATGAGCGGCTACGTGAACAACCCCGAGGCCACCAACGCCCTCATCGACAAGGA  
 CGGCTGGCTGCACAGCGGCGACATCGCCTACTGGGACGAGGACGAGCACTTCTTCATCGT  
 CGACCGGCTGAAGTCGCTGATCAAGTACAAGGGCTACCAGGTGGCGCCGGCCGAGCTGGA  
 GAGCATCTGCTCCAGCACCCCAACATCTTCGACGCCGGCGTGGCCGGGCTGCCGGACGA  
 CGACGCCGGCGAGCTGCCGGCCGGTGGTGGTGGTGGAGCACGGCAAGACCATGACGGA  
 GAAGGAGATCGTCGACTACGTGGCCAGCCAGGTGACCACCGCCAAGAAGCTGCGGGGCGG  
 CGTGGTGTTCGTGGACGAGGTCCCGAAGGGCCTGACCGGGAAGCTCGACGCCCGGAAAGAT  
 CCGCGAGATCCTGATCAAGGCCAAGAAGGGCGGCAAGATCGCCGTGTAAGACTAGTGCAT  
 CACATTTAAAAGCATCTCAGCCTACCATGAGAATAAGAGAAAAGAAAATGAAGATCAAAAG  
 CTTATTTCATCTGTTGGTCTTTTTCGTTGGTGTAAAGCCAACACCCTGTCTAAAAACATA  
 AATTTCTTTAATCATTTTGCTCTTTTCTCTGTGCTTCAATTAATAAAAAATGGAAAGAA  
 TCTAGATCTAAAAAATAA  
 AAAAAAAAAAAAAA-3'

도면10

PpLuc(GC) - 알부민6 - A64

5'-GGGAGAAAAGCTTGAGGATGGAGGACGCCAAGAACATCAAGAAGGGCCCGGCGCCCTTCTA  
 CCGCTGGAGGACGGGACCGCCGGCGAGCAGCTCCACAAGGCCATGAAGCGGTACGCCCT  
 GTGCGCGGCACGATCGCCTTACCGACGCCACATCGAGGTCGACATCAGCTACGCGGA  
 GTACTTCGAGATGAGCGTGGCCCTGGCCGAGGCCATGAAGCGGTACGGCCTGAACACCAA  
 CCACCGGATCGTGGTGTGCTCGGAGAACAGCCTGCAGTTCTTATGCCGGTGTGGGCGC  
 CCTTTCATCGGCGTGGCCGTGCGCCCGGCGAACGACATCTACAAAGAGCGGGAGCTGCT  
 GAACAGCATGGGATCAGCCAGCCGACCGTGGTGTTCGTGAGCAAGAAGGGCCTGCAGAA  
 GATCCTGAACGTGCAGAAGAAGCTGCCATCATCCAGAAGATCATCATCATGGACAGCAA  
 GACCGACTACCAGGGCTTCCAGTCGATGTACACGTTCCGTGACCAGCCACCTCCCGCCGGG  
 CTTCACAGGATACGACTTCGTCCCGGAGAGCTTCGACCGGGACAAGACCATCGCCCTGAT  
 CATGAACAGCAGGGCAGCACCGCCCTGCCGAAGGGGGTGGCCCTGCCGACCCGGACCGC  
 CTGGGTGCGCTTCTGACGCGCCGGGACCCCATCTTCGGCAACCAGATCATCCCGGACAC  
 CGCCATCCTGAGCGGTGCTGCTCCACCAGCGCTTCGGCATGTTACGACCCCTGGGCTA  
 CCTCATCTGCGGCTTCGGGTGGTCTGATGTACCGGTCGAGGAGGAGCTGTTCTCTGCG  
 GAGCCTGCAGGACTACAAGATCCAGAGCGCGCTGCTCGTCCGACCCCTGTTAGCTTCTT  
 CGCCAAGAGCACCTGATCGACAAGTACGACCTGTGAACTGCACGAGATCGCCAGCGG  
 GGGCGCCCGCTGAGCAAGGAGGTGGGCGAGGCGTGGCCAAGCGGTTCCACCTCCCGGG  
 CATCCGCCAGGGCTACGGCCTGACCGAGACCACGAGCGGATCCTGATCACCCCGAGGG  
 GGACGACAAGCCGGGCGCCGTGGGCAAGGTGGTCCCGTTCCTCGAGGCCAAGGTGGTGA  
 CCTGGACACCGGCAAGACCCTGGGCGTGAACCAGCGGGGCGAGCTGTGCGTGGCGGGGCC  
 GATGATCATGAGCGGCTACGTGAACAACCCGGAGGCCACCAACGCCCTCATCGACAAGGA  
 CGGCTGGCTGCACAGCGGACATCGCCTACTGGGACGAGGACGAGCACTTCTTCATCGT  
 CGACCGGCTGAAGTCGCTGATCAAGTACAAGGGTACCAGGTGGCGCCGGCCGAGCTGGA  
 GAGCATCCTGCTCCAGCACCCCAACATCTTCGACGCGCGGCTGGCCGGGCTGCCGGACGA  
 CGACCGCGCGAGCTGCGCGCCGCGGTGGTGGTGTGGAGCACGGCAAGACCATGACGGA  
 GAAGGAGATCGTGGACTACGTGGCCAGCCAGGTGACCACCGCCAAGAAGCTGGCGGGCGG  
 CGTGGTGTTCGTGGACGAGGTCCGAAGGGCCTGACCGGGAAGCTCGAGCCCGGAAGAT  
 CCGCGAGATCCTGATCAAGGCCAAGAAGGGCGGCAAGATCGCCGTGTAAGACTAGTGCA  
 CACATTTAAAAGCATCTCAGCCTACCATGAGAAATAAGAGAAAGAAAATGAAGATCAATAG  
 CTTATTCATCTGTTTTCTTTTTCTGTTGGTGTAAAGCCAACACCCCTGTCTAAAAAACATA  
 AATTTCTTTAATCATTTTGCTCTTTTTCTGTGCTTCAATTAATAAAAAATGGAAGAA  
 CCTAGATCTAAAAAATAA  
 AAAAAAAAAAAAAA-3'

도면11

PpLuc(GC) - 알부민7 - A64

5'-GGGAGAAAGCTTGAGGATGGAGGACGCCAAGAACATCAAGAAGGGCCCGGCGCCCTTCTA  
 CCCGCTGGAGGACGGGACCCCGGCGAGCAGCTCCACAAGGCCATGAAGCGGTACGCCCT  
 GGTGCCGGGCACGATCGCCCTTACCGAGCGCCACATCGAGGTGACATCACCTACGCGGA  
 GTACTTCGAGATGAGCGTGGCCCTGGCCGAGGCCATGAAGCGGTACGGCCTGAACACCAA  
 CCACCGGATCGTGGTGTGCTCGGAGAACAGCCTGCAGTTCTTCATGCCGGTGTGGGCGC  
 CCTCTTCATCGGCGTGGCCCTCGCCCGGCGAACGACATCTACAAGAGGGGAGCTGCT  
 GAACAGCATGGGATCAGCCAGCCGACCGTGGTGTTCGTGAGCAAGAAGGGCCTGCAGAA  
 GATCCTGAACGTGAGCAAGAAGCTGCCCATCATCCAGAAGATCATCATCATGGACAGCAA  
 GACCGACTACCAGGGCTTCAGTCGATGTACACGTTTCGTGACCAGCCACCTCCCGCCGGG  
 CTTCACAGGATACGACTTCGTTCCCGGAGAGCTTCGACCGGGACAAGACCATCGCCCTGAT  
 CATGAACAGCAGCGCAGCACCGGCTGCCGAAGGGGGTGGCCCTGCCGCACCGGACCGC  
 CTGCGTGGCTTCTCGCAGCCCGGACCCCATCTTCGGCAACCAGATCATCCCGGACAC  
 CGCCATCTGAGCGGTGCGCTCCACCACGGCTTCGGCATGTTACAGACCTGGGCTA  
 CCTCATCTGGGGTTCGGGTGGTCTGATGTACCGGTCGAGGAGGAGCTGTTCTGCG  
 GAGCTGCAGGACTACAAGATCCAGAGCGCGCTGCTGTCGCGACCTGTTTCAGCTTCTT  
 CGCCAAGAGCACCCTGATCGACAAGTACGACCTGTCGAACCTGCACGAGATCGCCAGCGG  
 GGGCGCCCGCTGAGCAAGGAGGTGGCGAGGCCGTGGCCAAAGCGGTTCCACCTCCCGGG  
 CATCCGCCAGGGCTAGGGCTGACCGAGACACGAGCGCGATCCTGATCACCCCGAGGG  
 GGACGACAAGCCGGCGCCGTGGGCAAGGTGGTCCCGTTCTTCGAGGCCAAGGTGGTGA  
 CACTGGACACCGCAAGACCTGGGCGTGAACCAGCGGGCGAGCTGTGCGTGGCGGGCC  
 GATGATCATGAGCGGTACGTGAACAACCCGAGGCCACCAACGCCCTCATCGACAAGGA  
 CGGCTGGCTGCACAGCGCGACATCGCCTACTGGGACGAGGACGAGCACTTCTTCATCGT  
 CGACCGCTGAAGTTCGCTGATCAAGTACAAGGGTACCAGGTGGCGCCGGCCGAGCTGGA  
 GAGCATCTGCTCCAGCACCCCAACATCTTCGACCGCGGCTGGCCGGCTGCCGGACGA  
 CGACCCCGGAGCTGCCCGCGCGTGGTGGTGTGGAGCACGGCAAGACCATGACGGA  
 GAAGGAGATCGTCGACTACGTGGCCAGCCAGGTGACCACCGCCAAGAAGCTGCGGGGGCG  
 CGTGGTGTTCGTGGACGAGTCCCAGGGCCCTGACCGGGAAGCTCGACCGCCGGAAGAT  
 CCGGAGATCCTGATCAAGCCCAAGAAGGGCGCAAGATCGCCGTGTAAGACTAGTGCAT  
 CACATTTAAAGCATCTCAGCCTACCATGAGAAATAAGAGAAAGAAAATGAAGATCAATAG  
 CTTATTCATCTCTTTTCTTTTTCGTTGGTAAAGCCAACACCTGTCTAAAAAACATA  
 AATTTCTTTAATCATTTTGGCTCTTTTCTCTGTGCTTCAATTAATAAAAATGGAAAGAA  
 CCTAGATCTAAAAAATAA  
 AAAAAAAAAAAAAA-3'

도면12

PpLuc(GC) - ag - A64

5'-GGGAGAAAGCTTGAGGATGGAGGACGCCAAGAACATCAAGAAGGGCCCGGCGCCCTTCTA  
 CCCGCTGGAGGACGGGACCCCGGCGAGCAGCTCCACAAGGCCATGAAGCGGTACGCCCT  
 GGTGCCGGGCACGATCGCCCTTACCGAGCGCCACATCGAGGTGACATCACCTACGCGGA  
 GTACTTCGAGATGAGCGTGGCCCTGGCCGAGGCCATGAAGCGGTACGGCCTGAACACCAA  
 CCACCGGATCGTGGTGTGCTCGGAGAACAGCCTGCAGTTCTTCATGCCGGTGTGGGCGC  
 CCTCTTCATCGGCGTGGCCCTCGCCCGGCGAACGACATCTACAAGAGCGGGAGCTGCT  
 GAACAGCATGGGGATCAGCCAGCCGACCGTGGTGTTCGTGAGCAAGAAGGGCCTGCAGAA  
 GATCTGAACGTGCAGAAGAAGCTGCCCATCATCCAGAAGATCATCATCATGGACAGCAA  
 GACCGACTACCAGGGCTTCAGTCGATGTACACGTTTCGTGACCAGCCACCTCCCGCCGGG  
 CTTCACAGGATACGACTTCGTTCCCGGAGAGCTTCGACCGGGACAAGACCATCGCCCTGAT  
 CATGAACAGCAGCGGCAGCACCGGCTGCCGAAGGGGGTGGCCCTGCCGCACCGGACCGC  
 CTGCGTGGCTTCTCGCACGCCCGGGACCCCATCTTCGGCAACCAGATCATCCCGGACAC  
 CGCCATCCTGAGCGTGTGCGGTTCCACCACGGCTTCGGCATGTTACAGACCTGGGCTA  
 CCTCATCTGCGGCTTCGGGTGGTCTGATGTACCGGTCGAGGAGGAGCTGTTCTGCG  
 GAGCCTGCAGGACTACAAGATCCAGAGCGCGTCTCTGCGCCACCTGTTTCAGCTTCTT  
 CGCCAAGAGCACCTGATCGACAAGTACGACCTGTCGAACCTGCACGAGATCGCCAGCGG  
 GGGCGCCCGCTGAGCAAGGAGGTGGGCGAGGCCGTGGCCAAGCGGTTCCACCTCCCGGG  
 CATCCCGCAGGGCTACGGCTGACCGAGACCACGAGCGCGATCCTGATCACCCCGAGGG  
 GGACGACAAGCCGGCGCCCTGGGCAAGGTGGTCCCCTTCTTCGAGGCCAAGGTGGTGA  
 CCTGGACACCGCAAGACCCTGGGCGTGAACCAGCGGGCGAGCTGTGCGTGGCGGGGCC  
 GATGATCATGAGCGGTACGTGAACAACCCGAGGCCACCAACGCCCTCATCGACAAGGA  
 CGGCTGGCTGCACAGCGCGACATCGCCTACTGGGACGAGGACGAGCACTTCTTCATCGT  
 CGACCGCTGAAGTTCGCTGATCAAGTACAAGGGTACCAGGTGGCGCCGGCCGAGCTGGA  
 GAGCATCTGCTCCAGCACCCCAACATCTTCGACCGCGGCTGGCCGGCTGCCGGACGA  
 CGACCGCGGCGAGCTGCCGGCCCGGTGGTGGTGTGGAGCACGGCAAGACCATGACGGA  
 GAAGGAGATCGTCGACTACGTGGCCAGCCAGGTGACCACCGCCAAGAAGCTGCGGGGGCG  
 CGTGGTGTTCGTGGACGAGTCCCAGGGCCCTGACCGGGAAGCTCGACCGCCGGAAGAT  
 CCGGAGATCCTGATCAAGGCCAAGAAGGGCGCAAGATCGCCGTGTAAGACTAGTTATA  
 AGACTGACTAGCCGATGGGCTCCCAACGGGCCCTCCTCCCTCCTTGACCCGAGATTA  
 ATAGATCTAAAAAATAA  
 AAAAAAAAAAAAAA-3'

도면13

**PpLuc(GC) - gusb - A64**

5' -GGGAGAAAGCTTGAGGATGGAGGACGCCAAGAACATCAAGAAGGGCCCGGGCCCTTCTA  
 CCCGCTGGAGGACGGGACCGCCGGCGAGCAGCTCCACAAGGCCATGAAGCGGTACGCCCT  
 GGTGCCGGGCACGATCGCCTTACCGACGCCACATCGAGGTCGACATCACCTACGCGGA  
 GTACTTCGAGATGAGCGTGGCCCTGGCCGAGGCCATGAAGCGGTACGGCTGAACACCAA  
 CCACCGGATCGTGGTGTGCTCGGAGAACAGCCTGCAGTTCTTCATGCCGGTGTGGGCGC  
 CCTCTTCATCGGCGTGGCCGTGCCCCGGCGAACGACATCTACAACGAGCGGGAGCTGCT  
 GAACAGCATGGGGATCAGCCAGCCGACCGTGGTGTTCGTGAGCAAGAAGGGCTGCAGAA  
 GATCTGAACGTGCAGAAGAAGCTGCCATCATCCAGAAGATCATCATCATGGACAGCAA  
 GACCGACTACCAGGGCTTCCAGTCGATGTACACGTTTCGTGACCAGCCACCTCCGCGCGGG  
 CTTCAACGAGTACGACTTCGTCCCGGAGAGCTTCGACCGGGACAAGACCATCGCCCTGAT  
 CATGAACAGCAGCGGCAGCACCGGCTGCGGAAGGGGTGGCCCTGCCGACCGGACCGC  
 CTGCGTGCCTTCTCGCACGCCCGGGACCCCATCTTCGGCAACCAGATCATCCGGACAC  
 CGCCATCTGAGCGTGGTCCGTTCCACCAGCGCTTCGGCATGTTACGACCTTGGGCTA  
 CCTCATCTGCGGCTTCCGGTGGTCTGATGTACCGGTTGAGGAGGAGCTGTTCTTGGC  
 GAGCCTGCAGGACTACAAGATCCAGAGCGCGCTGCTCGTCCGACCTGTTGAGCTTCTT  
 CGCCAAGAGCACCTGATCGACAAGTACGACCTGTGCAACCTGCACGAGATCGCCAGCGG  
 GGGCGCCCGCTGAGCAAGGAGGTGGGCGAGGCCGTGGCCAAAGCGGTTCCACCTCCCGGG  
 CATCCGCGAGGGCTACGGCCTGACCGAGACCACGAGCGCGATCCTGATCACCCCGAGGG  
 GGACGACAAGCCGGGCGCCCTGGGCAAGGTGGTCCCGTTCTTCGAGGCCAAGGTGGTGGGA  
 CCTGGACACCGCAAGACCCTGGGCTGAACCAGCGGGGCGAGCTGTGCGTGGGGGGCC  
 GATGATCATGAGCGGCTACGTGAACAACCCGAGGCCACCAACGCCCTCATCGACAAGGA  
 CGGCTGGTGCACAGCGGCGACATCGCCTACTGGGACGAGGACGAGCACTTCTTCATCGT  
 CGACCCGCTGAAGTCCGTGATCAAGTACAAGGGCTACCAGGTGGCCCGGCGAGCTGGA  
 GAGCATCTGCTCCAGCACCCCAACATCTTCGACCGCGGCTGGCCGGCTGCCGGACGA  
 CGACCGCGGCGAGCTGCCGGCCGCGGTGGTGGTGTGGAGCACGGCAAGACCATGACCGGA  
 GAAGGAGATCTGCGACTACGTGGCCAGCCAGGTGACCACCGCAAGAAGCTGCGGGCGG  
 CGTGGTGTTCGTGGACGAGGTCCCAGGGGCTGACCGGGAAGCTCGACGCCGGAAGAT  
 CCGCGAGATCTGATCAAGGCCAAGAAGGGCGGCAAGATCGCCGTGTAAGACTAGTGCAA  
 GACTGATACCACCTGCGTGTCCCTTCCCTCCCGAGTCAGGGCGACTTCCACAGCAGCAGA  
 ACAAGTCCCTCCTGGACTGTTACGCGCAGACCAGAACGTTTCTGCGCTGGGTTTGTGGT  
 CATCTATCTAGCAGGGAACACTAAAGGTGGAATAAAAAGATTTTCTATTATGGAAATAA  
 AGAGTTGGCATGAAAGTGGCTACTGAGATCTAAAAAATAAAAAAAAAAAAAAAAAAAAAA  
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA-3'

도면14

**PpLuc(GC) - atp50 - A64**

5' -GGGAGAAAGCTTGAGGATGGAGGACGCCAAGAACATCAAGAAGGGCCCGGGCCCTTCTA  
 CCCGCTGGAGGACGGGACCGCCGGCGAGCAGCTCCACAAGGCCATGAAGCGGTACGCCCT  
 GGTGCCGGGCACGATCGCCTTACCGACGCCACATCGAGGTCGACATCACCTACGCGGA  
 GTACTTCGAGATGAGCGTGGCCCTGGCCGAGGCCATGAAGCGGTACGGCTGAACACCAA  
 CCACCGGATCGTGGTGTGCTCGGAGAACAGCCTGCAGTTCTTCATGCCGGTGTGGGCGC  
 CCTCTTCATCGGCGTGGCCGTGCCCCGGCGAACGACATCTACAACGAGCGGGAGCTGCT  
 GAACAGCATGGGATCAGCCAGCCGACCGTGGTGTTCGTGAGCAAGAAGGGCTGCAGAA  
 GATCCTGAACGTGCAGAAGAAGCTGCCATCATCCAGAAGATCATCATCATGGACAGCAA  
 GACCGACTACCAGGGCTTCCAGTCGATGTACACGTTTCGTGACCAGCCACCTCCGCGCGG  
 CTTCAACGAGTACGACTTCGTCCCGGAGAGCTTCGACCGGGACAAGACCATCGCCCTGAT  
 CATGAACAGCAGCGGCAGCACCGCCCTGCCCCAAGGGGTGGCCCTGCCGCAACCGGACCGC  
 CTGCGTGCCTTCTCGCACGCCCGGGACCCCATCTTCGGCAACCAGATCATCCGGACAC  
 CGCCATCTGAGCGTGGTCCGTTCCACCAGCGCTTCGGCATGTTACGACCTGGGCTA  
 CCTCATCTGCGGCTTCCGGTGGTCTGATGTACCGGTTGAGGAGGAGCTGTTCTTGGC  
 GAGCCTGCAGGACTACAAGATCCAGAGCGCGCTGCTCGTCCGACCCCTGTTGAGCTTCTT  
 CGCCAAGAGCACCTGATCGACAAGTACGACCTGTGCAACCTGCACGAGATCGCCAGCGG  
 GGGCGCCCGCTGAGCAAGGAGGTGGGCGAGGCCGTGGCCAAAGCGGTTCCACCTCCCGGG  
 CATCCGCGAGGGCTACGGCCTGACCGAGACCACGAGCGCGATCCTGATCACCCCGAGGG  
 GGACGACAAGCCGGGCGCCGTGGGCAAGGTGGTCCCGTTCTTCGAGGCCAAGGTGGTGGGA  
 CCTGGACACCGCAAGACCCTGGGCTGAACCAGCGGGGCGAGCTGTGCGTGGGGGGCC  
 GATGATCATGAGCGGCTACGTGAACAACCCGAGGCCACCAACGCCCTCATCGACAAGGA  
 CGGCTGGTGCACAGCGGCGACATCGCCTACTGGGACGAGGACGAGCACTTCTTCATCGT  
 CGACCCGCTGAAGTCCGTGATCAAGTACAAGGGCTACCAGTGGCGCGGCGAGCTGGA  
 GAGCATCTGCTCCAGCACCCCAACATCTTCGACCGCGGCTGGCCGGCTGCCGGACGA  
 CGACCGCGGCGAGCTGCCGGCCGCGGTGGTGGTGTGGAGCACGGCAAGACCATGACCGGA  
 GAAGGAGATCTGCGACTACGTGGCCAGCCAGGTGACCACCGCAAGAAGCTGCGGGGCGG  
 GTGGTGTTCGTGGACGAGGTCCCAGGGGCTGACCGGGAAGCTGCACGCCGGAAGAT  
 CCGCGAGATCTGATCAAGGCCAAGAAGGGCGGCAAGATCGCCGTGTAAGACTAGTAAAGT  
 GTTGGTTTCTGCGCATCAGTAAAAATCTTAAACTTGGAGCAACAATAAAAAGCTTCCAG  
 AACAGATCAGATCTAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
 AAAAAAAAAAAAAAAAAAAAAA-3'

도면15

PpLuc(GC) - ndufa1 - A64

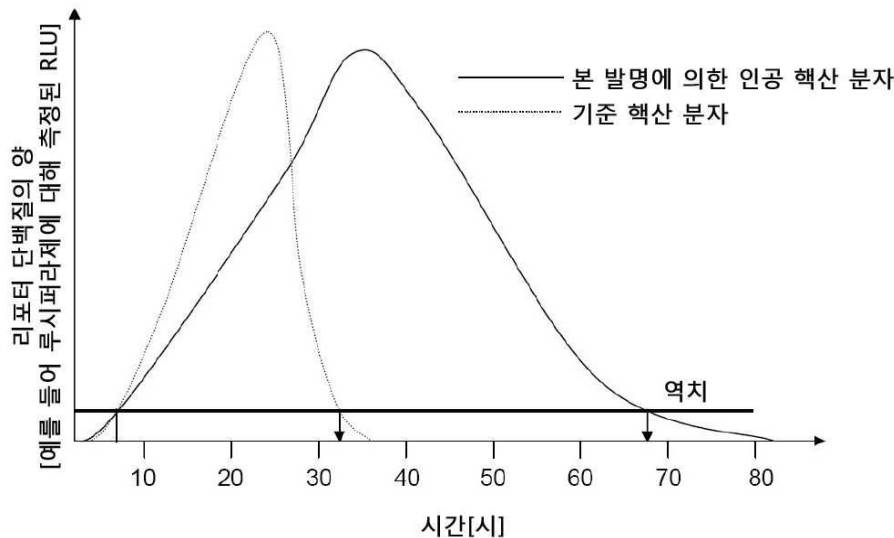
5'-GGGAGAAAGCTTGAGGATGGAGGACGCCAAGAACATCAAGAAGGGCCCGGCGCCCTTCTA  
 CCGCTGGAGGACGGGACCGCCGGCGAGCAGCTCCACAAGGCCATGAAGCGGTACGCCCT  
 GGTGCCGGGCACGATCGCCTTACCCGACGCCACATCGAGGTGACATCACCTACGCGGA  
 GTACTTCGAGATGAGCGTGGCCCTGGCCGAGGCCATGAAGCGGTACGGCCTGAACACCAA  
 CCACCGGATCGTGGTGTGCTCGGAGAAACAGCCTGCAGTTCTTCATGCCGGTGTGGGCGC  
 CCTTTCATCGGCGTGGCCGTCCGCCCCGGCGAACGACATCTACAACGAGCGGGAGCTGCT  
 GAACAGCATGGGGATCAGCCAGCCGACCGTGGTGTTCGTGAGCAAGAAGGGCCTGCAGAA  
 GATCCTGAACGTGCAGAAGAAGCTGCCATCATCCAGAAGATCATCATATGGACAGCAA  
 GACCGACTACCAGGGCTTCCAGTCGATGTACACGTTCTGTGACCAGCCACTCCCGCCGGG  
 CTTCAACGAGTACGACTTCGTCCCGGAGAGCTTCGACCGGGACAAGACCATCGCCCTGAT  
 CATGAACAGCAGCGGCAGCACCGCCCTGCCGAAGGGGGTGGCCCTGCCGACCGGACCGC  
 CTGCGTGGCCTTCTCGCACGCCCGGGACCCCATCTTCGGCAACCAGATCATCCCGGACAC  
 CGCCATCTGAGCGTGGTCCGCTTCCACCAGGCTTCGGCATGTTACGACCCCTGGGCTA  
 CCTCATCTGCGGCTTCCGGGTGGTCTGATGTACCGGTTGAGGAGGAGCTGTTCTGCG  
 GAGCCTGCAGGACTACAAGATCCAGAGCGCGTCTCGTCCGACCCCTGTTAGCTTCTT  
 CGCCAAGAGCACCTGATCGACAAGTACGACCTGTCGAACCTGCACGAGATCGCCAGCGG  
 GGGCGCCCGCTGAGCAAGGAGGTGGGCGAGGCCGTGGCCAAGCGGTTCCACCTCCCGG  
 CATCCCGCAGGGCTACGGCCTGACCGAGACCACGAGCGCGATCTGATCACCCCGAGGG  
 GGACGACAAGCCGGGCGCCGTGGGCAAGGTGGTCCCGTCTTCGAGGCCAAGGTGGTGA  
 CCTGGACACCGGCAAGACCCTGGGCGTGAACCAGCGGGGCGAGCTGTGCGTGGCGGGGCC  
 GATGATCATGAGCGGTACGTGAACAACCCGGAGGCCACCAACGCCCTCATCGACAAGGA  
 CGGCTGGCTGCACAGCGGACATCGCCTACTGGGACGAGGACGAGCACTTCTTCATCGT  
 CGACCGGCTGAAGTCGCTGATCAAGTACAAGGGCTACCAGGTGGCGCCCGCGAGCTGGA  
 GAGCATCTGCTCCAGCACCCCAACATCTTCGACGCGCGGCTGGCCTGGCTGCCGGACGA  
 CGACGCCCGCGAGCTGCGGCGCGGTTGGTGGTGTGGAGCACGGCAAGACCATGACGGA  
 GAAGGAGATCGTGCCTACGTGGCCAGCCAGGTGACCACCGCCAAGAAGCTGCGGGGCGG  
 CGTGGTGTTCGTGGACGAGGTCCCGAAGGGCCTGACCGGGAAGCTCGACGCCCGGAAGAT  
 CCGCGAGATCCTGATCAAGGCCAAGAAGGGCGGCAAGATCGCCGTGTAAGACTAGTGGAA  
 GCATTTTCCTGATTGATGAAAAAATAACTCAGTTATGGCCATCTACCCTGCTAGAAGG  
 TTACAGTGTATTATGTAGCATGCAATGTGTATGTAGTGCTTAATAAAAAATAAAATGAAA  
 AAAAAATGCAGATCTAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
 AAAAAAAAAAAAAAAAAA-3'

도면16

PpLuc(GC) - atp51 - A64

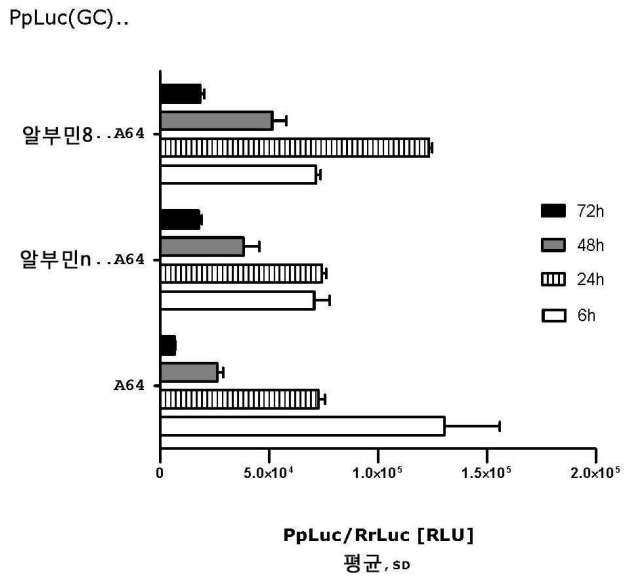
5'-GGGAGAAAGCTTGAGGATGGAGGACGCCAAGAACATCAAGAAGGGCCCGCGCCCTTCTA  
 CCCGCTGGAGGACGGGACCGCCGGCGAGCAGCTCCACAAGGCCATGAAGCGGTACGCCCT  
 GGTGCCGGGCACGATCGCCTTACCAGACGCCACATCGAGGTGCACATCACCTACGCGGA  
 GTACTTCGAGATGAGCGTGGCCCTGGCCGAGGCCATGAAGCGGTACGGCTGAACACCAA  
 CCACCGGATCGTGGTGTGCTCGGAGAAACAGCCTGCAGTTCCTCATGCCGGTGTGGGCGC  
 CCTCTTCATCGCGTGGCCGTGCCCGCGGAACGACATCTACAACGAGCGGGAGCTGCT  
 GAACAGCATGGGATCAGCCAGCCGACCCGTGGTGTTCGTGAGCAAGAAGGGCCTCAGAA  
 GATCCTGAACGTGCAGAAGAAGCTGCCCATCATCCAGAAGATCATCATGGACAGCAA  
 GACCGACTACCAGGGCTTCAGTCGATGTACACGTTTCGTGACCAGCCACTCCCGCCGGG  
 CTCAACGAGTACGACTTCGTCCCGGAGAGCTTCGACCGGGACAAGACCATCGCCCTGAT  
 CATGAACAGCGAGCGCAGCCAGCCGGCTGCCGAAGGGGGTGGCCCTGCCGACCCGACCGC  
 CTGGCTGGCGTTCGCGACGCCCGGACCCCATCTTCGCAACAGATCATCCGGACAC  
 CGCCATCCTGAGCGTGGTGGCGTCCACCAGGGCTTCGGCATGTTACGACCCCTGGGCTA  
 CCTCATCTGCGGCTTCGGGTGGTCTGATGTACCGGTTTCGAGGAGGAGCTGTTCTGCG  
 GAGCCTGCAGGACTACAAGATCCAGAGCGCGCTGCTCGTCCGACCCTGTTAGCTTCTT  
 CGCCAAGAGCACCTGATCGACAAGTACGACCTGTGAACTGCACGAGATCGCCAGCGG  
 GGGCGCCCGCTGAGCAAGGAGGTGGGCGAGGCCGTGGCCAAAGCGGTTCCACCTCCCGGG  
 CATCCCGCAGGGCTACGGCTGACCGAGACCACGAGCGGATCCTGATCACCCCGAGGG  
 GGACGACAAGCCGGCGCCGTGGGCAAGTGGTCCCGTTCGAGGCCAAGTGGTGGGA  
 CCTGGACACCGGCAAGACCTGGGCGTGAACCAGCGGGCGAGCTGTGCGTGGCGGGCC  
 GATGATCATGAGCGGTACGTGAACAACCCGAGGCCACCAACGCCCTCATCGACAAGGA  
 CGGCTGGCTGCACCGCGGACATCGCCTACTGGACGAGGACGAGCACTTCTTCATCGT  
 CGACCGGTGAAGTGGCTGATCAAGTACAAGGGTACCAGGTGGCGCCGGCCGAGCTGGA  
 GAGCATCCTGCTCCAGCACCCCAACATCTTCGACGCCGGCGTGGCCGGGCTGCCGGACGA  
 CGACCGCGCGAGCTGCCGGCCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT  
 GAAGGAGATCGTGGACTACGTGGCCAGCCAGGTGACCACCGCCAAGAAGCTGCCGGGCGG  
 CGTGGTGTTCGAGGACGAGTCCCGAAGGGCCTGACCGGGAAGCTCGACGCCCGGAAGAT  
 CCGCGAGATCTGATCAAGCCAAAGAAGGGCGGCAAGATCGCCGTGTAAGACTAGTAGAC  
 CAATCTTAAACATCTGATTATATTGATTTATTATTGAGTGTGTTGGACCATGTGTGA  
 TCAGACTGCTATCTGAATAAAATAAGATTTGTCAAACCTCAGTGTTCCTCCATCAGACA  
 CTCCATGAAGGTCACAATTTCTCTTGATATTAAGCTGGGTTGTCTTTAAACAACCTTAA  
 ATACACGCTCTGTTTAGCCCGCAATTGGAAAGGATATATGTGGCAATATAACCTGGTACA  
 TGAATATATGGGGATAACATTTTAATTTGAAGTTTGGAAATATATATTTAAGCTTTAT  
 TTCCAGAACAGTGGGTTAGGCTTTGGGAAAACTATAACTTGCCAAAGTAGAAGAAATA  
 GTAGTACCATATGCCAAAGTGATAGAGATGAATCATGTCAGTAGTLAGAATAACATTTCA  
 ACTGTTTCTTTGCTAAAAACACAGAAAGACCCATGACAACTCTATGCTCTGAAAAA  
 TGTTAGAGTACTTGTATCTTGAATATAGCCTCCCAAGAGAGAACAGGGTGGTATTCTA  
 AGTATGTTTCTTTGTAACATCTTTAGCAGTAGGACAGAGCCATACATGTGAAATCTGATT  
 TTTATGTGTGTTATTCGTTTGTCTGGTTTTACTACCTTTGCAAAAACAAAATACCCCAA  
 GATATTTAAACAAGGTATATAATTTAGCATCTTCCTGGATCTAAATAGTATATATATCC  
 TGAATAAAATGAAATGATTGCTATAGATCTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
 AA-3'

도면17



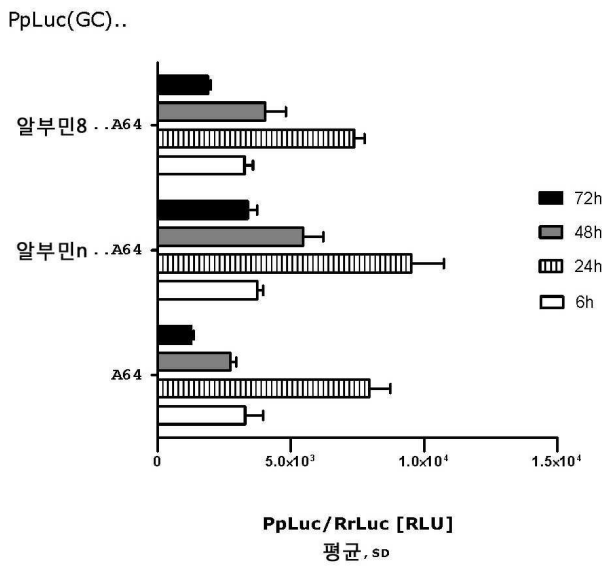
도면18

상이한 종으로부터 유래하는 알부민 3'-UTR을 포함하는 mRNA로부터의 PpLuc (HDF)



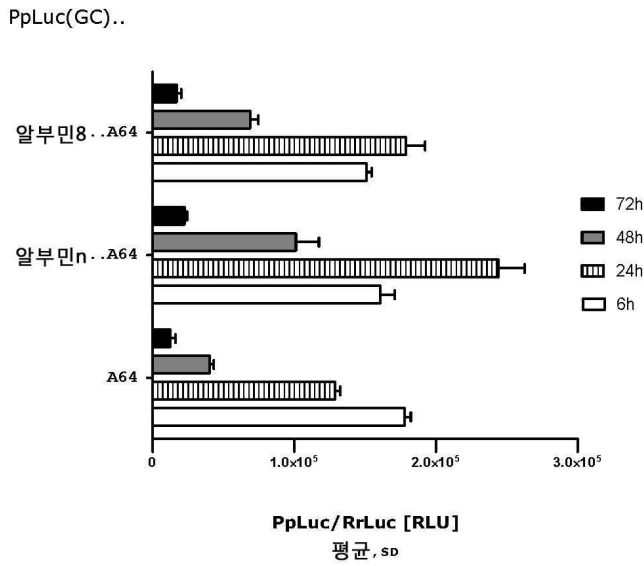
도면19

상이한 종으로부터 유래하는 알부민 3'-UTR을 포함하는 mRNA로부터의 PpLuc (HDF)



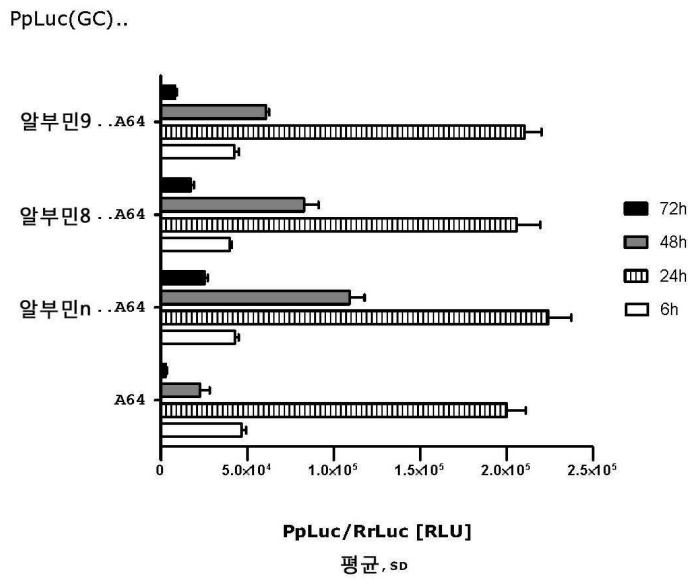
도면20

상이한 종으로부터 유래하는 알부민 3'-UTR을 포함하는 mRNA로부터의 PpLuc (L-929)



도면21

상이한 종으로부터 유래하는 알부민 3'-UTR을 포함하는 mRNA로부터의 PpLuc (HeLa)



도면22

PpLuc(GC) - 알부민8 - A64

5'-GGGAGAAAGCTTGAGGATGGAGGACGCCAAGAACATCAAGAAGGGCCCGGGCCCTTCTA  
 CCCGCTGGAGGACGGGACCGCGGCGAGCAGCTCCACAAGGCCATGAAGCGGTACGCCCT  
 GGTGCCGGGACAGATCGCCTTACCCGACGCCACATCGAGGTGACATCACCTACGCCGA  
 GTACTTCGAGATGAGCGTGGCCCTGGCCGAGGCCATGAAGCGGTACGGCCTGAACACCAA  
 CCACCGGATCGTGGTGTGCTCGGAGAACAGCCTGCAGTTCCTCATGCCGGTGTGGGCGC  
 CCTCTTCATCGGCGTGGCCGTGCCCCGGCGAACGACATCTACAACGAGCGGGAGCTGCT  
 GAACAGCATGGGGATCAGCCAGCCGACCGTGGTGTTCGTGAGCAAGAAGGGCCTGCAGAA  
 GATCCTGAACGTGCAGAAGAAGCTGCCCATCATCCAGAAGATCATCATCATGGACAGCAA  
 GACCGACTACCAGGCTTCCAGTCGATGTACACGTTTCGTGACCAGCCACTCCCGCCGGG  
 CTTCAACGAGTACGACTTCGTCCCGGAGAGCTTCGACCCGGGACAAGACCATCGCCCTGAT  
 CATGAACAGCAGCGGCAGCACCGGCTGCCGAAGGGGGTGGCCCTGCCGACCCGGACCGC  
 CTGCGTGGCCTTCTCGCACGCCCGGGACCCCATCTTCGGCAACCAGATCATCCGGACAC  
 CGCCATCCTGAGCGTGGTGCCTTCCACCAGGCTTCGGCATGTTACGACCCCTGGGCTA  
 CCTCATCTGCGGCTTCGGGTGGTCTGATGTACCGGTTTCGAGGAGGAGCTGTTCTGCG  
 GAGCCTGCAGGACTACAAGATCCAGAGCGCGTCTCGTGCCGACCCGTTTTCAGCTTCTT  
 CGCCAAGAGCACCTGATCGACAAGTACGACCTGTGCAACCTGCACGAGATCGCCAGCGG  
 GGGCGCCCGCTGAGCAAGGAGGTGGGCGAGGCCGTGGCCAAAGCGGTTCCACCTCCCGGG  
 CATCCGCGCAGGGCTACGGCTGACCCGAGACCACGAGCGCGATCCTGATCACCCCGAGGG  
 GGACGACAAGCGGCGCCCGTGGCAAGTGGTCCCCTTTCGAGGCCAAGGTGGTGGGA  
 CCTGGACACCGGCAAGACCTGGGCGTGAACAGCGGGCGAGCTGTGCGTGGGGGGCC  
 GATGATCATGAGCGGCTACGTGAACAACCCGGAGGCCACCAACGCCCTCATCGACAAGGA  
 CGGCTGGCTGCACAGCGGCGACATCGCCTACTGGGACGAGGACGAGCACTTCTTCATCGT  
 CGACCGGCTGAAGTTCGCTGATCAAGTACAAGGGCTACCAGGTGGCGCCGGCCGAGCTGGA  
 GAGCATCCTGCTCCAGCACCCCAACATCTTCGACGCGCGGTGGCCGGGCTGCCGGACGA  
 CGACGCCGGGAGCTGCGCGCCGCGGTGGTGGTGTGGAGCACGGCAAGACCATGACCGGA  
 GAAGGAGATCGTCGACTACGTGGCCAGCCAGGTGACCACCGCCAAGAAGCTGCGGGGCGG  
 CGTGGTGTTCGTGGACGAGGTCCCAGAGGCCCTGACCGGGAAGCTCGACGCCCGGAAGAT  
 CCGCGAGATCCTGATCAAGGCCAAGAAGGGCGGCAAGATCGCCGTGTAAGACTAGTAAAC  
 ATCACAATTAAGAACATCTCAGCCTACCATGAGAACAAGAGAATAAATAAGATCAA  
 AGCTTATTCATCTGTTTTCTTTTTCATTGGTATAAAGCCAACCCCTGTCTAAAAAACT  
 ATAAATTTCTTTAATCATTTTGCCTCTTTCTCTGTGCTTCAATTAATAAAAAATGGAAA  
 GAATCTAGATCTAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA  
 AAAAAAATAAAAAA-3'

도면23

PpLuc(GC) - 알부민9 - A64

5'-GGGAGAAAGCTTGAGGATGGAGGACGCCAAGAACATCAAGAAGGGCCCGGGCCCTTCTA  
 CCCGCTGGAGGACGGGACCGCGGCGAGCAGCTCCACAAGGCCATGAAGCGGTACGCCCT  
 GGTGCCGGGACAGATCGCCTTACCCGACGCCACATCGAGGTGACATCACCTACGCCGA  
 GTACTTCGAGATGAGCGTGGCCCTGGCCGAGGCCATGAAGCGGTACGGCCTGAACACCAA  
 CCACCGGATCGTGGTGTGCTCGGAGAACAGCCTGCAGTTCCTCATGCCGGTGTGGGCGC  
 CCTCTTCATCGGCGTGGCCGTGCCCCGGCGAACGACATCTACAACGAGCGGGAGCTGCT  
 GAACAGCATGGGGATCAGCCAGCCGACCGTGGTGTTCGTGAGCAAGAAGGGCCTGCAGAA  
 GATCCTGAACGTGCAGAAGAAGCTGCCCATCATCCAGAAGATCATCATCATGGACAGCAA  
 GACCGACTACCAGGCTTCCAGTTCGATGTACACGTTTCGTGACCAGCCACTCCCGCCGG  
 CTTCAACGAGTACGACTTCGTCCCGGAGAGCTTCGACCCGGGACAAGACCATCGCCCTGAT  
 CATGAACAGCAGCGGCAGCACCGGCTGCCGAAGGGGGTGGCCCTGCCGACCCGGACCGC  
 CTGCGTGGCCTTCTCGCACGCCCGGGACCCCATCTTCGGCAACCAGATCATCCGGACAC  
 CGCCATCCTGAGCGTGGTGCCTTCCACCAGGCTTCGGCATGTTACGACCCCTGGGCTA  
 CCTCATCTGCGGCTTCGGGTGGTCTGATGTACCGGTTTCGAGGAGGAGCTGTTCTGCG  
 GAGCCTGCAGGACTACAAGATCCAGAGCGCGTCTCGTGCCGACCCGTTTTCAGCTTCTT  
 CGCCAAGAGCACCTGATCGACAAGTACGACCTGTGCAACCTGCACGAGATCGCCAGCGG  
 GGGCGCCCGCTGAGCAAGGAGGTGGGCGAGGCCGTGGCCAAGCGGTTCCAATCCCGGG  
 CATCCGCGCAGGGCTACGGCCTGACCGAGACCACGAGCGCGATCCTGATCACCCCGAGGG  
 GGACGACAAGCCGGCGCCCGTGGCAAGTGGTCCCCTTTCGAGGCCAAGGTGGTGGGA  
 CCTGGACACCGGCAAGACCTGGGCGTGAACAGCGGGCGAGCTGTGCGTGGGGGGCC  
 GATGATCATGAGCGGCTACGTGAACAACCCGGAGGCCACCAACGCCCTCATCGACAAGGA  
 CGGCTGGCTGCACAGCGGCGACATCGCCTACTGGGACGAGGACGAGCACTTCTTCATCGT  
 CGACCGGCTGAAGTTCGCTGATCAAGTACAAGGGCTACCAGGTGGCGCCGGCCGAGCTGGA  
 GAGATCCTGCTCAGCACGCCCAACATCTTCGACGCGCGGTGGCCGGCTGCCGGACGA  
 CGACGCCGGGAGCTGCCGGCCGCGTGGTGGTGTGGAGCACGGCAAGACCATGACCGGA  
 GAAGGAGATCGTCGACTACGTGGCCAGCCAGGTGACCACCGCCAAGAAGCTGCGGGGCGG  
 CGTGGTGTTCGTGGACGAGGTCCCAGAGGCCCTGACCGGGAAGCTCGACGCCCGGAAGAT  
 CCGCGAGATCCTGATCAAGGCCAAGAAGGGCGGCAAGATCGCCGTGTAAGACTAGTACAC  
 ATCACAACCACAACCTTCTCAGGCTACCCTGAGAAAAAAGACATGAAGACTCAGGACTC  
 ATCTTTTCTGTGGTGTAAAATCAACCCCTAAGGAACACAATAATTTCTTTAATCATTTGA  
 CTCTTGTCTCTGTGCTGCAATTAATAAAAAATGGAAGAATCTACAGATCTAAAAAATA  
 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA  
 AAAAAAATAAAAAA-3'

서열 목록

SEQUENCE LISTING

<110> CureVac GmbH

<120> Artificial nucleic acid molecules

<130> CU01P126W01

<160> 41

<170> PatentIn version 3.5

<210> 1

<211> 186

<212> DNA

<213> Artificial Sequence

<220><223> fragment of 3'UTR of an albumin gene

<400> 1

```
catcacattt aaaagcatct cagcctacca tgagaataag agaaagaaaa tgaagatcaa      60
aagcttattc atctgttttt ctttttcggt ggtgtaaagc caacaccctg tctaaaaaac      120
ataaatttct ttaatcattt tgcctctttt cctctgtgctt caattaataa aaaatggaaa      180
```

```
gaatct      186
```

<210> 2

<211> 186

<212> RNA

<213> Artificial Sequence

<220><223> fragment of 3'UTR of an albumin gene

<400> 2

```
caucacauuu aaaagcaucu cagccuacca ugagaauaag agaaagaaaa ugaagaucaa      60
aagcuuuauuc aucuguuuuu cuuuuucguu gguguaaagc caacaccug ucuaaaaaaac      120
auaaaauuucu uaaaucauuu ugccucuuuu cucugugcuu cauuuaauaa aaaauggaaa      180
gaaucu      186
```

<210> 3

<211> 256

<212> DNA

<213> Artificial Sequence

<220><223> fragment of 3'UTR of albumin plus poly(A)

<400> 3  
 catcacatctt aaaagcatct cagcctacca tgagaataag agaaagaaaa tgaagatcaa 60  
 aagcttattc atctgttttt ctttttcggt ggtgtaaagc caacaccctg tctaaaaaac 120  
 ataaatttct ttaatcattt tgcctctttt ctctgtgctt caattaataa aaaatggaaa 180  
 gaatctagat ctaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 240  
 aaaaaaaaaa aaaaaa 256

<210> 4  
 <211> 256  
 <212> RNA  
 <213> Artificial Sequence  
 <220><223> fragment of 3'UTR of albumin plus poly(A)

<400> 4  
 caucacauuu aaaagcaucu cagccuacca ugagaauaag agaaagaaaa ugaagaucaa 60  
 aagcuuuuuc aucuguuuuu cuuuuucguu gguguaaagc caacaccug ucuaaaaaac 120  
 auaaaauucu uuaaucauuu ugccucuuuu cucugugcuu caauuaauaa aaaauggaaa 180  
 gaaucuagau cuaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 240  
 aaaaaaaaaa aaaaaa 256

<210> 5  
 <211> 1746  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> PpLuc(GC) A64

<400> 5  
 gggagaaagc ttgaggatgg aggacccaa gaacatcaag aagggcccg cgccttcta 60  
 cccgctggag gacgggaccg ccggcgagca gctccacaag gccatgaagc ggtacgcct 120  
 ggtgccgggc acgatcgct taccgacgc ccacatcgag gtcgacatca cctacgcgga 180  
 gtacttcgag atgagcgtgc gcctggccga ggccatgaag cggtagcgcc tgaacaccaa 240  
 ccaccggatc gtggtgtgct cggagaacag cctgcagttc tteatgccgg tgctgggcgc 300  
 cctcttcacg ggcgtggccg tcgccccggc gaacgacatc tacaacgagc gggagctgct 360  
 gaacagcatg gggatcagcc agccgaccgt ggtgttcgtg agcaagaagg gcctgcagaa 420  
 gatcctgaac gtgcagaaga agctgccccat catccagaag atcatcatca tggacagcaa 480  
 gaccgactac cagggttcc agtcgatgta cacgttcgtg accagccacc tcccgcggg 540

cttcaacgag tacgacttgc tcccggagag cttcgaccgg gacaagacca tcgccctgat 600  
 catgaacagc agcggcagca ccggcctgcc gaaggggggtg gccctgccgc accggaccgc 660  
 ctgctgtgcg ttctgcacg cccgggacce catcttcggc aaccagatca tcccggacac 720  
 cgccatcctg agcgtggtgc cgttccacca cggcttcggc atgttcacga ccctgggcta 780  
  
 cctcatctgc ggcttccggg tggctctgat gtaccggttc gaggaggagc tgttctctgcg 840  
 gagcctgcag gactacaaga tccagagcgc gctgctcgtg ccgaccctgt tcagcttctt 900  
 cgccaagagc accctgatcg acaagtacga cctgtcgaac ctgcacgaga tcgccagcgg 960  
 gggcgcccc ctgagcaagg aggtgggcga ggccgtggcc aagcggttcc acctcccggg 1020  
 catccgccag ggctaccgcc tgaccgagac cacgagcgcg atcctgatca cccccgaggg 1080  
 ggacgacaag ccgggcgcgg tgggcaaggt ggtcccgttc ttcgaggcca aggtggtgga 1140  
 cctggacacc ggcaagacc tgggcgtgaa ccagcggggc gagctgtgcg tgcggggggc 1200  
  
 gatgatcatg agcggctacg tgaacaacc ggaggccacc aacgcctca tcgacaagga 1260  
 cggctggctg cacagcggcg acatcgcta ctgggacgag gacgagcact tcttcatcgt 1320  
 cgaccggctg aagtctgta tcaagtacaa gggctaccag gtggcgcggc ccgagctgga 1380  
 gagcatcctg ctccagcacc ccaacatctt cgacgccggc gtggccgggc tgccggacga 1440  
 cgacgccggc gagctgccgg ccgcggtggt ggtgctggag cacggcaaga ccatgacgga 1500  
 gaaggagatc gtcgactacg tggccageca ggtgaccacc gccaagaagc tgcggggcgg 1560  
 cgtggtgttc gtggacgagg tcccgaaggg cctgaccggg aagctcgacg cccggaagat 1620  
  
 ccgcgagatc ctgatcaagg ccaagaaggg cggcaagatc gccgtgtaag actagtagat 1680  
 ctaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa 1740  
 aaaaaa 1746  
  
 <210> 6  
 <211> 1933  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> PpLuc(GC) albumin A64  
 <400> 6  
  
 gggagaaagc ttgaggatgg aggacgcaa gaacatcaag aaggccccg cgcccttcta 60  
 cccgctggag gacgggaccg ccggcgagca gctccacaag gccatgaagc ggtacgccct 120  
 ggtgccgggc acgatcgct taccgacgc ccacatcgag gtcgacatca cctacgcgga 180  
  
 gtacttcgag atgagcgtgc gcctggccga ggccatgaag cggtagcgcc tgaacaccaa 240

ccaccggatc gtgggtgtgct cggagaacag cctgcagttc ttcattgccg tgctgggcgc 300  
cctcttcate ggctggccg tcgccccggc gaacgacatc tacaacgagc gggagctgct 360  
gaacagcatg gggatcagcc agccgaccgt ggtgttcgtg agcaagaagg gcctgcagaa 420  
gatcctgaac gtgcagaaga agctgccccat catccagaag atcatcatca tggacagcaa 480  
gaccgactac cagggttcc agtcgatgta cacgttcgtg accagccacc tcccggcggg 540  
cttcaacgag tacgacttcc tcccggagag cttcgaccgg gacaagacca tcgacctgat 600  
  
catgaacagc agcggcagca ccggcctgcc gaaggggggtg gccctgccgc accggaccgc 660  
ctgctgtgce ttctgcacg cccgggaccc catcttcggc aaccagatca tcccggacac 720  
cgccatcctg agcgtgggtc cgttccacca cggcttcggc atgttcacga ccctgggcta 780  
cctcatctgc ggcttccggg tggctctgat gtaccggttc gaggaggagc tgttctctgc 840  
gagcctgcag gactacaaga tccagagcgc gctgctcgtg ccgaccctgt tcagcttctt 900  
cgccaagagc acctgatcg acaagtacga cctgtcgaac ctgcacgaga tcgccagcgg 960  
gggcgccccg ctgagcaagg aggtgggcga ggccgtggcc aagcggttcc acctcccggg 1020  
  
catccgccag ggctacggcc tgaccgagac caggagecgc atcctgatca cccccagggg 1080  
ggacgacaag ccgggcgccg tgggcaaggt ggtcccgttc ttcgaggcca aggtggtgga 1140  
cctggacacc ggcaagacc tgggcgtgaa ccagcggggc gagctgtgcg tgcgggggcc 1200  
gatgatcatg agcggctacg tgaacaacc ggaggccacc aacgccctca tcgacaagga 1260  
cggctggctg cacagcggcg acatgccta ctgggacgag gacgagcact tcttcatcgt 1320  
cgaccggctg aagtctgta tcaagtacaa gggctaccag gtggcggccg ccgagctgga 1380  
gagcatcctg ctccagcacc ccaacatctt cgacgccggc gtggccgggc tgccggacga 1440  
  
cgacgccggc gagctgccgg ccgcggtggt ggtgctggag cacggcaaga ccatgacgga 1500  
gaaggagatc gtcgactacg tggccagcca ggtgaccacc gccaagaagc tgcggggcgg 1560  
cgtggtgttc gtggacgagg tcccgaaggg cctgaccggg aagctcgacg cccggaagat 1620  
ccgcgagatc ctgatcaagg ccaagaaggg cgcaagatc gccgtgtaag actagtgcac 1680  
cacatttaa agcatctcag cctaccatga gaataagaga aagaaaatga agatcaaaag 1740  
cttattcate tgtttttctt tttcgttggf gtaaagccaa caccctgtct aaaaaacata 1800  
aatttcttta atcattttgc ctcttttctc tgtgcttcaa ttaataaaaa atggaaagaa 1860  
  
tctagatcta aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa 1920  
aaaaaaaaa aaa 1933

<210> 7

<211> 1933

<212> DNA

<213> Artificial Sequence

<220><223> PpLuc(GC) albumin2 A64

<400> 7

```

gggagaaagc ttgaggatgg aggacgcaa gaacatcaag aagggcccg cgccttcta      60
cccgtggag gacgggaccg ccggcgagca gctccacaag gccatgaag ggtacgcct      120
ggtgccgggc acgatcgct tcaccgacgc ccacatcgag gtcgacatca cctacgcgga      180
gtacttcgag atgagcgtgc gcctggccga ggccatgaag cggtagcgcc tgaacaccaa      240

ccaccggatc gtggtgtgct cggagaacag cctgcagttc ttcattgccg tgctgggcgc      300
cctcttcate ggcgtggccg tcgccccggc gaacgacatc tacaacgagc gggagctgct      360
gaacagcatg gggatcagcc agccgaccgt ggtgttcgtg agcaagaagg gcctgcagaa      420
gatcctgaac gtgcagaaga agctgccccat catccagaag atcatcatca tggacagcaa      480
gaccgactac cagggttcc agtcgatgta cacgttcgtg accagccacc tcccgcggg      540
cttcaacgag tacgacttcg tcccggagag cttcgaccgg gacaagacca tcgcctgat      600
catgaacagc agcggcagca ccggcctgcc gaagggggtg gccctgccgc accggaccgc      660

ctgcgtgcgc ttctgcacg cccgggacc catcttcggc aaccagatca tcccggacac      720
cgccatcctg agcgtgtgct cgttcacca cggcttcggc atgttcacga ccctgggcta      780
cctcatctgc ggcttccggg tggctctgat gtaccggttc gaggaggagc tgttcctgcg      840
gagcctgcag gactacaaga tccagagcgc gctgctcgtg ccgacctgt tcagcttctt      900
cgccaagagc acctgatcg acaagtacga cctgtcgaac ctgcacgaga tcgccagcgg      960
gggcgcccc ctgagcaagg aggtgggcga ggccgtggcc aagcggttcc acctcccggg      1020
catccgccag ggctacggcc tgaccgagac caggagecgc atcctgatca cccccaggg      1080

ggacgacaag ccgggcgccg tgggcaaggt ggtcccgttc ttcgaggcca aggtggtgga      1140
cctggacacc ggcaagacc tgggcgtgaa ccagcggggc gagctgtgcg tgcgggggcc      1200
gatgatcatg agcggctacg tgaacaacc ggaggccacc aacgccctca tcgacaagga      1260
cggctggctg cacagcggcg acatgccta ctgggacgag gacgagcact tcttcatcgt      1320
cgaccggctg aagtcgctga tcaagtataa gggctaccag gtggcggccg ccgagctgga      1380
gagcatcctg ctcagcacc ccaacatctt cgacgccggc gtggccgggc tgccggacga      1440
cgacccggc gagctgccgg ccgcggtggt ggtgctggag cacggcaaga ccatgacgga      1500

gaaggagatc gtcgactacg tggccagcca ggtgaccacc gccaagaagc tgcggggcgg      1560
cgtggtgttc gtggacgagg tcccgaaggg cctgaccggg aagctcgagc cccggaagat      1620

```

ccgcgagatc ctgatcaagg ccaagaaggg cggcaagatc gccgtgtaag actagtgcac 1680  
 cacatttaaa agcatctcag cctaccatga gaataagaga aagaaatga agatcaaaag 1740  
 cttattcgtc tgtttttctt tttcgttggg gtaaagccaa caccctgtct aaaaaacata 1800  
 aatttcttta atcattttgc ctcttttctc tgtgcttcaa ttaataaaaa atggaaagaa 1860  
 tctagatcta aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1920  
  
 aaaaaaaaaa aaa 1933  
 <210> 8  
 <211> 1933  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> PpLuc(GC) albumin3 A64  
 <400> 8  
 gggagaaagc ttgaggatgg aggacgcaa gaacatcaag aagggcccg gccccttcta 60  
 cccgctggag gacgggaccg ccggcgagca gctccacaag gccatgaagc ggtacgcct 120  
 ggtgccgggc acgatcgct taccgacgc ccacatcgag gtcgacatca cctacgcgga 180  
 gtacttcgag atgagcgtgc gcctggccga ggccatgaag cggtagcgcc tgaacaccaa 240  
 ccaccggatc gtggtgtgct cggagaacag cctgcagttc ttcattgccg tgctgggccc 300  
  
 cctcttcac ggcgtggccg tcgccccggc gaacgacatc tacaacgagc gggagctgct 360  
 gaacagcatg gggatcagcc agccgaccgt ggtgttcgtg agcaagaagg gcctgcagaa 420  
 gatcctgaac gtgcagaaga agctgccccat catccagaag atcatcatca tggacagcaa 480  
 gaccgactac cagggttcc agtcgatgta cacgttcgtg accagccacc tcccgcggg 540  
 cttaacgag tacgacttgc tcccggagag cttcgaccgg gacaagacca tcgcctgat 600  
 catgaacagc agcggcagca ccggcctgcc gaaggggggtg gccctgccgc accggaccgc 660  
 ctgcgtgcgc ttctgcacg cccgggacc catcttcggc aaccagatca tcccggacac 720  
  
 cgccatcctg agcgtggtgc cgttccacca cggcttcggc atgttcacga cctgggcta 780  
 cctcatctgc ggcttccggg tggctctgat gtaccggttc gaggaggagc tgttcctgcg 840  
 gagcctcgag gactacaaga tccagagcgc gctgctcgtg ccgacctgt tcagcttctt 900  
 cgccaagagc acctgatcg acaagtacga cctgtcgaac ctgcacgaga tcgccagcgg 960  
 gggcgccccg ctgagcaagg aggtgggcga ggccgtggcc aagcggttcc acctcccggg 1020  
 catccgccag ggctacggcc tgaccgagac caccgagcgc atcctgatca cccccaggg 1080  
 ggacgacaag ccgggcgccc tgggcaaggt ggtcccgttc ttcgaggcca aggtggtgga 1140



cgccatcctg agcgtgggtgc cgttccacca cggcttcggc atgttcacga ccctgggcta 780

cctcatctgc ggcttccggg tggctctgat gtaccggttc gaggaggagc tgttctcgcg 840

gagcctgcag gactacaaga tccagagcgc gctgctcgtg ccgaccctgt tcagcttctt 900

cgccaagagc accctgatcg acaagtacga cctgtcgaac ctgcacgaga tcgccagcgg 960

ggggccccc ctgagcaagg aggtgggcga ggccgtggcc aagcggttcc acctcccggg 1020

catccgccag ggctacggcc tgaccgagac cagcagcgcg atcctgatca cccccgaggg 1080

ggacgacaag ccgggcgcgg tgggcaaggt ggtcccgttc ttcgaggcca aggtggtgga 1140

cctggacacc ggcaagacc tgggcgtgaa ccagcggggc gagctgtgcg tgcgggggccc 1200

gatgatcatg agcggctacg tgaacaacc ggaggccacc aacgcctca tcgacaagga 1260

cggctggctg cacagcggcg acatcgcta ctgggacgag gacgagcact tcttcatcgt 1320

cgaccggctg aagtctgta tcaagtacaa gggctaccag gtggcgcggc ccgagctgga 1380

gagcatcctg ctccagcacc ccaacatctt cgacgccggc gtggccgggc tgccggacga 1440

cgacgccggc gagctgccgg ccgcggtggt ggtgctggag cacggcaaga ccatgacgga 1500

gaaggagatc gtcgactacg tggccagcca ggtgaccacc gccaagaagc tgcggggcgg 1560

cgtggtgttc gtggacgagg tcccgaaggg cctgaccggg aagctcgacg cccggaagat 1620

ccgcgagatc ctgatcaagg ccaagaaggg cggcaagatc gccgtgtaag actagtgcac 1680

cacatttaa agcatctcag cctaccatga gaataagaga aagaaatga agatcaaaag 1740

cttattcacc tcttttctt tttcgttggg gtaaagccaa caccctgtct aaaaaacata 1800

aatttcttta atcattttgc ctcttttctc tgtgcttcaa ttaataaaaa atggaaagaa 1860

tctagatcta aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa 1920

aaaaaaaaaaa aaa 1933

<210> 10

<211> 1933

<212> DNA

<213> Artificial Sequence

<220><223> PpLuc(GC) albumin5 A64

<400> 10

gggagaaagc ttgaggatgg aggacgcaa gaacatcaag aagggcccg cgcccttcta 60

cccgtggag gacgggaccg ccggcgagca gctccacaag gccatgaagc ggtacgcct 120

ggtgccgggc acgatgcct taccgacgc ccacatcgag gtcgacatca cctacgcgga 180

gtacttcgag atgagcgtgc gcctggccga ggccatgaag cggtagcgcc tgaacaccaa 240

ccaccggatc gtgggtgct cggagaacag cctgcagttc ttcattgccg tgctgggcgc 300  
 cctcttcate ggcgtagccg tcgccccggc gaacgacatc tacaacgagc gggagctgct 360  
 gaacagcatg gggatcagcc agccgaccgt ggtgttcgtg agcaagaagg gcctgcagaa 420

gatcctgaac gtgcagaaga agctgcccat catccagaag atcatcatca tggacagcaa 480  
 gaccgactac cagggttcc agtcgatgta cacgttcgtg accagccacc tcccgcggg 540  
 ctccaacgag tacgacttcg tcccggagag cttcgaccgg gacaagacca tcgacctgat 600  
 catgaacagc agcggcagca ccggcctgcc gaaggggggtg gccctgccgc accggaccgc 660  
 ctgctgctgc ttctgcacg cccgggaccc catcttcggc aaccagatca tcccggacac 720  
 cgccatcctg agcgtggcgc cgttcacca cggcttcggc atgttcacga cctgggcta 780  
 cctcatctgc ggcttccggg tggctctgat gtaccggttc gaggaggagc tgttcctgcg 840

gagcctgcag gactacaaga tccagagcgc gctgctcgtg cggacctgt tcagcttctt 900  
 cgccaagagc acctgatcg acaagtacga cctgtcgaac ctgcacgaga tcgccagcgg 960  
 gggcgccccg ctgagcaagg aggtgggcga ggccgtggcc aagcggttcc acctcccggg 1020  
 catccgccag ggctacggcc tgaccgagac cagcagcgcg atcctgatca cccccaggg 1080  
 ggacgacaag ccgggcgccg tgggcaaggt ggtcccgttc ttcgaggcca aggtggtgga 1140  
 cctggacacc ggcaagacc tgggcgtgaa ccagcggggc gagctgtgcg tgcgggggcc 1200  
 gatgatcatg agcggctacg tgaacaacc ggaggccacc aacgccctca tcgacaagga 1260

cggctggctg cacagcggcg acatgccta ctgggacgag gacgagcact tcttcatcgt 1320  
 cgaccggctg aagtctctga tcaagtataa gggctaccag gtggcggccg ccgagctgga 1380  
 gagcatcctg ctccagcacc ccaacatctt cgacgccggc gtggccgggc tgccggacga 1440  
 cgacgccgcg gagctgccg ccgcggtggt ggtgctggag cacggcaaga ccatgacgga 1500  
 gaaggagatc gtcgactacg tggccagcca ggtgaccacc gccaaagaagc tgcggggcgg 1560  
 cgtggtgttc gtggacgagg tcccgaagg cctgaccggg aagctcgacg cccggaagat 1620  
 ccgcgagatc ctgatcaagg ccaagaagg cgcaagatc gccgtgtaag actagtgcac 1680

cacatttaa agcatctcag cctaccatga gaataagaga aagaaatga agatcaaag 1740  
 cttattcate tgttggctt tttcgttgg gtaaagccaa caccctgtct aaaaaacata 1800  
 aatttttta atcattttc ctcttttctc tgtgcttcaa ttaataaaaa atggaaagaa 1860  
 tctagatcta aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa 1920  
 aaaaaaaaa aaa 1933

<210> 11

<211> 1933

<212> DNA

<213> Artificial Sequence

<220><223> PpLuc(GC) albumin6 A64

<400> 11

```

gggagaaagc ttgaggatgg aggacgcaa gaacatcaag aagggcccg cgccttcta      60

cccgtggag gacgggaccg ccggcgagca gtcacacaag gccatgaag ggtacgccct      120
ggtgccgggc acgatcgct tcaccgacgc ccacatcgag gtcgacatca cctacgcgga      180
gtacttcgag atgagcgtgc gcctggccga ggccatgaag cggtagccgc tgaacaccaa      240
ccaccggatc gtggtgtgct cggagaacag cctgcagttc ttcattgccg tgctgggcgc      300
cctcttcacg ggcgtggccg tcgccccggc gaacgacatc tacaacgagc gggagctgct      360
gaacagcatg gggatcagcc agccgaccgt ggtgttcgtg agcaagaagg gcctgcagaa      420
gatcctgaac gtgcagaaga agctgccccat catccagaag atcatcatca tggacagcaa      480

gaccgactac cagggttcc agtcgatgta cacgttcgtg accagccacc tcccgccggg      540
cttcaacgag tacgacttcg tcccggagag cttcgaccgg gacaagacca tcgacctgat      600
catgaacagc agcggcagca ccggcctgcc gaaggggggtg gccctgccgc accggaccgc      660
ctgcgtgcgc ttctgcacg cccgggacc ccatcttcggc aaccagatca tcccggacac      720
cgccatcctg agcgtggtgc cgttcacca cggcttcggc atgttcacga ccctgggcta      780
cctcatctgc ggcttccggg tggctctgat gtaccggttc gaggaggagc tgttcctgcg      840
gagcctgcag gactacaaga tccagagcgc gctgctcgtg ccgacctgt tcagcttctt      900

cgccaagagc acctgatcg acaagtacga cctgtcgaac ctgcacgaga tcgccagcgg      960
ggggccccc ctgagcaagg aggtgggcga ggccgtggcc aagcggttcc acctcccggg      1020
catccgccag ggtacggcc tgaccgagac cagcagcgcg atcctgatca cccccgaggg      1080
ggacgacaag ccgggcgccg tgggcaaggt ggtcccgttc ttcgaggcca aggtggtgga      1140
cctggacacc ggcaagacc tgggcgtgaa ccagcggggc gagctgtgcg tcggggggcc      1200
gatgatcatg agcggctacg tgaacaacc ggaggccacc aacgccctca tcgacaagga      1260
cggctggctg cacagcggcg acatgccta ctgggacgag gacgagcact tcttcatcgt      1320

cgaccggctg aagtgcctga tcaagtacaa gggctaccag gtggcgccgg ccgagctgga      1380
gagcatcctg ctccagcacc ccaacatctt cgacgccggc gtggccgggc tgccggacga      1440
cgacccggc gagctgccg ccgcggtgtt ggtgctggag cacggcaaga ccatgacgga      1500
gaaggagatc gtcgactacg tggccagcca ggtgaccacc gccaagaagc tcggggcgcg      1560
cgtggtgttc gtggacgagg tcccgaaggg cctgaccggg aagctcgagc cccggaagat      1620

```

ccgcgagatc ctgatcaagg ccaagaaggg cggcaagatc gccgtgtaag actagtgcac 1680  
cacatttaaa agcatctcag cctaccatga gaataagaga aagaaatga agatcaatag 1740

cttattcacc tgtttttctt tttcgttggc gtaaagecaa caccctgtct aaaaaacata 1800  
aatttcttta atcattttgc ctcttttctc tgtgcttcaa ttaataaaaa atggaagaa 1860  
cctagatcta aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1920  
aaaaaaaaaa aaa 1933

<210> 12  
<211> 1933  
<212> DNA  
<213> Artificial Sequence  
<220><223> PpLuc(GC) albumin7 A64  
<400> 12

gggagaaagc ttgaggatgg aggacccaa gaacatcaag aagggcccg cgccttcta 60  
cccgtggag gacgggaccg ccggcgagca gctccacaag gccatgaagc ggtacgcct 120

ggtgccgggc acgatcgct tcaccgacgc ccacatcgag gtcgacatca cctacgcgga 180  
gtacttcgag atgagcgtgc gcctggccga ggccatgaag cggtagcggc tgaacaccaa 240  
ccaccggate gtggtgtgct cggagaacag cctgcagttc tteatgccgg tgctgggcgc 300  
cctcttcacc ggcgtggccg tcgccccggc gaacgacatc tacaacgagc gggagctgct 360  
gaacagcatg gggatcagcc agccgaccgt ggtgttcgtg agcaagaagg gcctgcagaa 420  
gatcctgaac gtgcagaaga agctgccccat catccagaag atcatcatca tggacagcaa 480  
gaccgactac cagggttcc agtcgatgta cacgttcgtg accagccacc tcccgcggg 540

cttcaacgag tacgacttcg tcccggagag cttcgaccgg gacaagacca tcgacctgat 600  
catgaacagc agcggcagca ccggcctgcc gaaggggggtg gccctgccgc accggaccgc 660  
ctgctgctgc ttctcgcacg cccgggacc cacttcggc aaccagatca tcccggacac 720  
cgccatcctg agcgtggtgc cgttcacca cggcttcggc atgttcacga cctgggcta 780  
cctcatctgc ggcttccggg tggctctgat gtaccggttc gaggaggagc tgttctctgc 840  
gagcctcgag gactacaaga tccagagcgc gctgctctgt cggacctgt tcagcttctt 900  
cgccaagagc acctgatcg acaagtacga cctgtcgaac ctgcacgaga tcgccagcgg 960

gggcgccccg ctgagcaagg aggtgggcga ggccgtggcc aagcggttcc acctcccggg 1020  
catccgccag ggctacggcc tgaccgagac caccagcgcg atcctgatca cccccaggg 1080  
ggacgacaag ccgggcgccc tgggcaaggt ggtcccgttc ttcgaggcca aggtggtgga 1140

cctggacacc ggcaagacc tgggcgtgaa ccagcggggc gagctgtgcg tgcgggggcc 1200  
gatgatcatg agcggctacg tgaacaacc ggaggccacc aacgcctca tcgacaagga 1260  
cggctggctg cacagcggcg acatcgcta ctgggacgag gacgagcact tcttcatcgt 1320  
cgaccggctg aagtcgctga tcaagtacaa gggctaccag gtggcgccgg ccgagctgga 1380

gagcatcctg ctccagcacc ccaacatctt cgacgccggc gtggccgggc tgccggacga 1440  
cgacgccggc gagctgccgg ccgcggtggt ggtgctggag cacggcaaga ccatgacgga 1500  
gaaggagatc gtcgactacg tggccagcca ggtgaccacc gccaagaagc tgcggggcgg 1560  
cgtggtgttc gtggacgagg tcccgaaggg cctgaccggg aagctcgacg cccggaagat 1620  
ccgcgagatc ctgatcaagg ccaagaaggg cggcaagatc gccgtgtaag actagtgcac 1680  
cacatttaa agcatctcag cctaccatga gaataagaga aagaaaatga agatcaatag 1740  
cttattcacc tcttttctt tttcgttggg gtaaagccaa caccctgtct aaaaaacata 1800

aatttcttta atcattttgc ctcttttctc tgtgcttcaa ttaataaaaa atggaagaa 1860  
cctagatcta aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa 1920  
aaaaaaaaa aaa 1933

<210> 13  
<211> 1812  
<212> DNA  
<213> Artificial Sequence  
<220><223> PpLuc(GC) ag A64  
<400> 13

gggagaaagc ttgaggatgg aggacccaa gaacatcaag aagggcccg cgcccttcta 60  
cccgtggag gacgggaccg ccggcgagca gctccacaag gccatgaagc ggtacgccct 120  
ggtgccgggc acgatcgct taccgacgc ccacatcgag gtcgacatca cctacgcgga 180

gtacttcgag atgagcgtgc gctggccga ggccatgaag cggtacggcc tgaacaccaa 240  
ccaccggate gtggtgtgct cggagaacag cctgcagttc tteatgccgg tgctgggcgc 300  
cctcttcacc ggcgtggccg tcgccccggc gaacgacatc tacaacgagc gggagctgct 360  
gaacagcatg gggatcagcc agccgaccgt ggtgttcgtg agcaagaagg gctgcagaa 420  
gatcctgaac gtcagaaga agctgcccat catccagaag atcatcatca tggacagcaa 480  
gaccgactac cagggttcc agtcgatgta cacgttcgtg accagccacc tcccggggg 540  
cttcaacgag tacgacttcg tcccggagag cttcgaccgg gacaagacca tcgacctgat 600

catgaacagc agcggcagca ccggcctgcc gaagggggtg gccctgccgc accggaccgc 660

ctgctgctgc tctctgcacg cccgggaccc catcttcggc aaccagatca tcccggacac 720  
 cgccatcctg agcgtggtgc cgttccacca cggcttcggc atgttcacga ccctgggcta 780  
 cctcatctgc ggcttccggg tggctctgat gtaccggttc gaggaggagc tgttctctgc 840  
 gagcctgcag gactacaaga tccagagcgc gctgctctgt ccgaccctgt tcagcttctt 900  
 cgccaagagc accctgatcg acaagtacga cctgtcgaac ctgcacgaga tcgccagcgg 960  
 gggcgccccg ctgagcaagg aggtgggcca ggccgtggcc aagcggttcc acctcccggg 1020  
  
 catccgccag ggctacggcc tgaccgagac cagcagcgcg atcctgatca cccccgaggg 1080  
 ggacgacaag ccgggcgccc tgggcaaggt ggtcccgttc ttcgaggcca agtggtgga 1140  
 cctggacacc ggcaagacc tgggcgtgaa ccagcggggc gagctgtgcg tgcgggggcc 1200  
 gatgatcatg agcggctacg tgaacaacc ggaggccacc aacgccctca tcgacaagga 1260  
 cggctggctg cacagcggcg acatgccta ctgggacgag gacgagcact tcttcatcgt 1320  
 cgaccggctg aagtcgctga tcaagtacaa gggctaccag gtggcgccgg ccgagctgga 1380  
 gagcatcctg ctccagcacc ccaacatctt cgacgccggc gtggccgggc tgccggacga 1440  
  
 cgacgccgce gagctgccgg ccgcggtggt ggtgctggag cacggcaaga ccatgacgga 1500  
 gaaggagatc gtcgactacg tggccagcca ggtgaccacc gccaagaagc tgcggggcgg 1560  
 cgtggtgttc gtggacgagg tcccgaaggg cctgaccggg aagctcgacg cccggaagat 1620  
 ccgcgagatc ctgatcaagg ccaagaaggg cggcaagatc gccgtgtaag actagttata 1680  
 agactgacta gcccgatggg cctccaacg ggcctcctc ccctccttc accgagatta 1740  
 atagatctaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1800  
 aaaaaaaaaa aa 1812

<210> 14

<211> 1955

<212> DNA

<213> Artificial Sequence

<220><223> PpLuc(GC) gusb A64

<400> 14

gggagaaagc ttgaggatgg aggacccaa gaacatcaag aaggccccg cgccttcta 60  
 cccgctggag gacgggaccg ccggcgagca gctccacaag gccatgaagc ggtacgccct 120  
 ggtgccgggc acgatgcct tcaccgacgc ccacatcgag gtcgacatca cctacgcgga 180  
 gtacttcgag atgagcgtgc gcctggccga ggccatgaag cggtagccgc tgaacaccaa 240  
 ccaccggatc gtggtgtgct cggagaacag cctgcagttc ttcagccgg tgctggcgcc 300

cctcttcacg ggcgtggccg tcgccccggc gaacgacatc tacaacgagc gggagctgct 360

gaacagcatg gggatcagcc agccgaccgt ggtgttcgtg agcaagaagg gcctgcagaa 420

gatcctgaac gtgcagaaga agctgcccat catccagaag atcatcatca tggacagcaa 480

gaccgactac cagggcttcc agtcgatgta cacgttcgtg accagccacc tcccgcggg 540

cttcaacgag tacgacttcg tcccggagag cttcgaccgg gacaagacca tcgccctgat 600

catgaacagc agcggcagca ccggcctgcc gaaggggggtg gccctgccgc accggaccgc 660

ctgctgctgc ttctcgacg cccgggacc catcttcggc aaccagatca tcccggacac 720

cgccatcctg agcgtggtgc cgttcacca cggcttcggc atgttcacga ccctgggcta 780

cctcatctgc ggcttccggg tggctctgat gtaccggttc gaggaggagc tttctctgcg 840

gagcctgcag gactacaaga tccagagcgc gctgctcgtg ccgacctgt tcagcttctt 900

cgccaagagc acctgatcg acaagtacga cctgtcgaac ctgcacgaga tcgccagcgg 960

ggcgccccg ctgagcaagg aggtgggca ggccgtggcc aagcggttcc acctcccggg 1020

catccgccag ggctacggcc tgaccgagac cacgagcgcg atcctgatca cccccaggg 1080

ggacgacaag cgggcccgc tgggcaaggt ggtcccgttc ttcgaggcca aggtggtgga 1140

cctggacacc ggcaagacc tgggcgtgaa ccagcggggc gagctgtgcg tgcgggggcc 1200

gatgatcatg agcggctacg tgaacaacc ggaggccacc aacgcctca tcgacaagga 1260

cggctggctg cacagcggcg acatgcctta ctgggacgag gacgagcact tcttcatcgt 1320

cgaccggctg aagtcgtga tcaagtacaa gggtaccag gtggcggcg ccgagctgga 1380

gagcatcctg ctccagcacc ccaacatctt cgacgcggc gtggccgggc tgccggacga 1440

cgacccggc gagctgccgg ccgctgtgtt ggtgctggag cacggcaaga ccatgacgga 1500

gaaggagatc gtcgactac tggccagcca ggtgaccacc gccaagaagc tgcggggcgg 1560

cgtggtgttc gtggacgagg tcccgaagg cctgaccggg aagctcgac cccggaagat 1620

ccgcgagatc ctgatcaagg ccaagaagg cgcaagatc gccgtgtaag actagtcaa 1680

gactgatacc acctgcgtgt ccttctctcc ccgagtcagg gcgacttcca cagcagcaga 1740

acaagtgcct cctggactgt tcacggcaga ccagaacgtt tctggcctgg gttttgtgtt 1800

catctattct agcaggaac actaaagggt gaataaaaag attttctatt atggaataa 1860

agagttggca taaagtggc tactgagatc taaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1920

aaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaa 1955

<210> 15  
 <211> 1818  
 <212> DNA

<213> Artificial Sequence

<220><223> PpLuc(GC) atp5o A64

<400> 15

```

gggagaaagc ttgaggatgg aggacgcca gaacatcaag aagggcccg cgccttcta      60
cccgctggag gacgggaccg ccggcgagca gctccacaag gccatgaagc ggtacgcct      120
ggtgccgggc acgatcgctt taccgacgc ccacatcgag gtcgacatca cctacgcgga      180
gtacttcgag atgagcgtgc gcctggccga ggccatgaag cggtacggcc tgaacaccaa      240
ccaccggatc gtggtgtgct cggagaacag cctgcagttc ttcattgccg tgctgggcgc      300
cctcttcate ggcgtggccg tcgccccggc gaacgacatc tacaacgagc gggagctgct      360
gaacagcatg gggatcagcc agccgaccgt ggtgttcgtg agcaagaagg gcctgcagaa      420

gatcctgaac gtgcagaaga agctgcccat catccagaag atcatcatca tggacagcaa      480
gaccgactac cagggttcc agtcgatgta cacgttcgtg accagccacc tcccgcggg      540
cttcaacgag tacgacttcg tcccggagag cttcgaccgg gacaagacca tcgccctgat      600
catgaacage agcggcagca ccggcctgcc gaaggggggtg gccctgccgc accggaccgc      660
ctgcgtgcmc ttctgcacg cccgggacc catcttcggc aaccagatca tcccggacac      720
cgccatcctg agcgtggtgc cgttcacca cggcttcggc atgttcacga cctgggcta      780
cctcatctgc ggcttccggg tggctctgat gtaccggttc gaggaggagc tgttctcg      840

gagcctgcag gactacaaga tccagagcgc gctgctcgtg ccgacctgt tcagcttctt      900
cgccaagagc accctgatcg acaagtacga cctgtcgaac ctgcacgaga tcgccagcgg      960
gggcgccccg ctgagcaagg aggtgggcga ggccgtggcc aagcggttcc acctcccggg      1020
catccgccag ggctacggcc tgaccgagac cagcagcgcg atcctgatca cccccaggg      1080
ggacgacaag ccgggcgccg tgggcaaggt ggtcccgttc ttcgaggcca aggtggtgga      1140
cctggacacc ggcaagacc tgggcgtgaa ccagcggggc gagctgtgcg tgcgggggcc      1200
gatgatcatg agcggctacg tgaacaacc ggaggccacc aacgcctca tcgacaagga      1260

cggctggctg cacagcggcg acatcgctta ctgggacgag gacgagcact tcttcatcgt      1320
cgaccggctg aagtcgtcga tcaagtataa gggtaccag gtggcgcgg ccgagctgga      1380
gagcatcctg ctccagcacc ccaacatctt cgacgcggc gtggccgggc tgccggacga      1440
cgacccggc gagctgccgg ccgcggtggt ggtgctggag cacggcaaga ccatgacgga      1500
gaaggagatc gtcgactacg tggccagcca ggtgaccacc gccaagaagc tgcggggcgg      1560
cgtggtgttc gtggacgagg tcccgaagg cctgaccggg aagctcgagc cccggaagat      1620

```

ccgcgagatc ctgatcaagg ccaagaaggg cggcaagatc gccgtgtaag actagtaagt 1680  
  
 gttggttttc tgccatcagt gaaaattctt aaacttggag caacaataaa aagcttccag 1740  
 aacagatcag atctaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1800  
 aaaaaaaaaa aaaaaaaaaa 1818  
 <210> 16  
 <211> 1877  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> PpLuc(GC) ndufal A64  
 <400> 16  
  
 gggagaaagc ttgaggatgg aggacgcaa gaacatcaag aagggcccg cgccttcta 60  
 cccgctggag gacgggaccg ccggcgagca gctccacaag gccatgaagc ggtacgccct 120  
 ggtgccgggc acgatcgct tcaccgacgc ccacatcgag gtcgacatca cctacgcgga 180  
  
 gtacttcgag atgagcgtgc gcctggccga ggccatgaag cggtacggcc tgaacaccaa 240  
 ccaccggatc gtggtgtgct cggagaacag cctgcagttc ttcattgccg tgctgggcgc 300  
 cctcttcac ggcgtggccg tcgccccggc gaacgacatc tacaacgagc gggagctgct 360  
 gaacagcatg gggatcagcc agccgaccgt ggtgttcgtg agcaagaagg gcctgcagaa 420  
 gatcctgaac gtgcagaaga agctgccccat catccagaag atcatcatca tggacagcaa 480  
 gaccgactac cagggttcc agtcgatgta cacgttcgtg accagccacc tcccgcggg 540  
 cttcaacgag tacgacttcg tcccggagag cttcgaccgg gacaagacca tcgccctgat 600  
  
 catgaacagc agcggcagca ccggcctgcc gaaggggggtg gccctgccgc accggaccgc 660  
 ctgctgctgc ttctcgcaag cccgggacc cactctcggc aaccagatca tcccggacac 720  
 cgccatcctg agcgtggtgc cgttcacca cggcttcggc atgttcacga ccctgggcta 780  
 cctcatctgc ggcttccggg tggctctgat gtaccggttc gaggaggagc tgttctctgc 840  
 gagcctgcag gactacaaga tccagagcgc gctgctcgtg ccgaccctgt tcagcttctt 900  
 cgccaagagc accctgatcg acaagtacga cctgtcgaac ctgcacgaga tcgccagcgg 960  
 gggcgccccg ctgagcaagg aggtgggcga ggccgtggcc aagcggttcc acctcccggg 1020  
  
 catccgccag ggctacggcc tgaccgagac cagagcgcg atcctgatca cccccgaggg 1080  
 ggacgacaag ccgggcgccg tgggcaaggt ggtcccgttc ttcgaggcca aggtggtgga 1140  
 cctggacacc ggcaagacc tgggcgtgaa ccagcggggc gagetgtgcg tgcgggggccc 1200  
 gatgatcatg agcggctacg tgaacaacc ggaggccacc aacgcctca tcgacaagga 1260

cggctggctg cacagcggcg acatgcctia ctgggacgag gacgagcact tcttcatcgt 1320  
 cgaccggctg aagtcgctga tcaagtacaa gggctaccag gtggcgccgg ccgagctgga 1380  
 gagcatcctg ctccagcacc ccaacatctt cgacgccggc gtggccgggc tgccggacga 1440  
  
 cgacgccggc gagctgccgg ccgcggtggt ggtgctggag cacggcaaga ccatgacgga 1500  
 gaaggagatc gtcgactacg tggccagcca ggtgaccacc gccaagaagc tgcggggcgg 1560  
 cgtggtgttc gtggacgagg tcccgaaggg cctgaccggg aagctcgacg cccggaagat 1620  
 ccgcgagatc ctgatcaagg ccaagaaggg cggcaagatc gccgtgtaag actagtggaa 1680  
 gcattttcct gattgatgaa aaaaataact cagttagtgc catctacccc tgctagaagg 1740  
 ttacagtgta ttatgtagca tgcaatgtgt tatgtagtgc ttaataaaaa taaaatgaaa 1800  
 aaaaatgcaga tctaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1860  
  
 aaaaaaaaaa aaaaaaa 1877  
 <210> 17  
 <211> 2494  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> PpLuc(GC) atp5l A64  
 <400> 17  
 gggagaaagc ttgaggatgg aggacccaa gaacatcaag aagggcccg cgccttcta 60  
 cccgctggag gacgggaccg ccggcgagca gctccacaag gccatgaagc ggtacgcct 120  
 ggtgccgggc acgatcgct taccgacgc ccacatcgag gtcgacatca cctacgcgga 180  
 gtacttcgag atgagcgtgc gcctggccga ggccatgaag cggtagcgcc tgaacaccaa 240  
 ccaccggatc gtggtgtgct cggagaacag cctgcagttc ttcattgccg tgctgggcgc 300  
  
 cctcttcate ggcgtggccg tcgccccggc gaacgacatc tacaacgagc gggagctgct 360  
 gaacagcatg gggatcagcc agccgaccgt ggtgttcgtg agcaagaagg gcctgcagaa 420  
 gatcctgaac gtgcagaaga agctgcccat catccagaag atcatcatca tggacagcaa 480  
 gaccgactac cagggttcc agtcgatgta cacgttcgtg accagccacc tcccgcggg 540  
 cttaacgag tacgacttg tcccggagag cttcgaccgg gacaagacca tcgacctgat 600  
 catgaacagc agcggcagca ccggcctgcc gaagggggtg gccctgccgc accggaccgc 660  
 ctgctgctgc ttctgcacg cccgggacc catcttcggc aaccagatca tcccggacac 720  
  
 cgccatcctg agcgtggtgc cgttccacca cggcttcggc atgttcacga cctgggcta 780  
 cctcatctgc ggcttccggg tggctctgat gtaccggttc gaggaggagc tgttctctgc 840

gagcctgcag gactacaaga tccagagcgc gctgctcgtg ccgaccctgt tcagcttctt 900  
cgccaagagc accctgatcg acaagtacga cctgtcgaac ctgcacgaga tcgccagcgg 960  
gggcgccccg ctgagcaagg aggtgggcca ggccgtggcc aagcggttcc acctcccggg 1020  
catccgccag ggctacggcc tgaccgagac cagcagcgcg atcctgatca cccccgaggg 1080  
ggacgacaag ccgggcgccc tgggcaaggt ggtcccgttc ttcgaggcca aggtggtgga 1140

ctgggacacc ggcaagacc tgggcgtgaa ccagcggggc gagctgtgcg tgcgggggccc 1200  
gatgatcatg agcggctacg tgaacaacc ggaggccacc aacgcctca tcgacaagga 1260  
cggctggctg cacagcggcg acatcgcta ctgggacgag gacgagcact tcttcatcgt 1320  
cgaccggctg aagtcgctga tcaagtacaa gggctaccag gtggcgccgg ccgagctgga 1380  
gagcatcctg ctccagcacc ccaacatctt cgacgccggc gtggccgggc tgccggacga 1440  
cgacgccggc gagctgccgg ccgcggtggt ggtgctggag cacggcaaga ccatgacgga 1500  
gaaggagatc gtcgactacg tggccagcca ggtgaccacc gccaaagaagc tgcggggcgg 1560

cgtggtgttc gtggacgagg tcccgaaggg cctgaccggg aagctcgacg cccggaagat 1620  
ccgcgagatc ctgatcaagg ccaagaaggg cggcaagatc gccgtgtaag actagtagac 1680  
caatctttaa catctgatta tatttgattt attatttgag tgttgttgga ccatgtgtga 1740  
tcagactgct atctgaataa aataagattt gtcaaaaactc agtgttttct ccatcagaca 1800  
ctccatgaaa ggtcacaatt tctcttgata ttaagctggg ttgtctttaa acaaccctaa 1860  
atacagctct gtttagcccc caattggaaa ggatatatgt ggcaatatta acctggtaca 1920  
tgaatatatg gggataacat ttttaattga aggtttggaa tatatatatt taagctttat 1980

tccagaaca gtgagggtta ggtcttggga aaactataac ttgccaaagt agaagaata 2040  
gtagtaccat atgccaaagt gatagagatg aatcatgtca gtagttagaa taacatttca 2100  
actgttttct ttgctaaaaat cacagaaaaga ccctattgac aacatctatg tetgtaaaaa 2160  
tgtagagta ctigtcatct tgaatatagc ctcccaaga gagaacaggg tggatttcta 2220  
agtatgttcc tttgtaacat ctttagcagt aggacagagc catacatgtg aaatctgatt 2280  
tttatgtgtg ttatctgitt gtctggtttt actaccttg caaaaacaaa ataccceaaa 2340  
gatattttaa caaggttata atttagcatc ttcctggat ctaaatagta tattatatcc 2400

tgaaataaat gaaatgattg ctatagatct aaaaaaaaa aaaaaaaaa aaaaaaaaa 2460  
aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaa 2494

<210> 18  
<211> 155  
<212> DNA

<213> Artificial Sequence  
 <220><223> fragment of SEQ ID No. 1  
 <400> 18  
 aaaagcatct cagcctacca tgagaataag agaaagaaaa tgaagatcaa aagcttattc 60  
 atctgttttt ctttttcggt ggtgtaaagc caacaccctg tctaaaaaac ataaatttct 120  
 ttaatcattt tgcctctttt ctctgtgctt caatt 155  
 <210> 19

<211> 110  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> fragment of SEQ ID No. 1  
 <400> 19  
 catcacattt aaaagcatct cagcctacca tgagaataag agaaagaaaa tgaagatcaa 60  
 aagcttattc atctgttttt ctttttcggt ggtgtaaagc caacaccctg 110  
 <210> 20

<211> 110  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> fragment of SEQ ID No. 1  
 <400> 20  
 aaaagcatct cagcctacca tgagaataag agaaagaaaa tgaagatcaa aagcttattc 60  
 atctgttttt ctttttcggt ggtgtaaagc caacaccctg tctaaaaaac 110

<210> 21  
 <211> 110  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> fragment of SEQ ID No. 1  
 <400> 21  
 cagcctacca tgagaataag agaaagaaaa tgaagatcaa aagcttattc atctgttttt 60  
 ctttttcggt ggtgtaaagc caacaccctg tctaaaaaac ataaatttct 110

<210> 22  
 <211> 110  
 <212> DNA

<213> Artificial Sequence  
 <220><223> fragment of SEQ ID No. 1  
 <400> 22  
 tgagaataag agaaagaaaa tgaagatcaa aagcttattc atctgttttt ctttttcggt 60  
 ggtgtaaagc caacaccctg tctaaaaaac ataaatttct ttaatcattt 110

<210> 23  
 <211> 110  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> fragment of SEQ ID No. 1  
 <400> 23  
 agaaagaaaa tgaagatcaa aagcttattc atctgttttt ctttttcggt ggtgtaaagc 60  
 caacaccctg tctaaaaaac ataaatttct ttaatcattt tgccctctttt 110

<210> 24  
 <211> 110  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> fragment of SEQ ID No. 1  
 <400> 24  
 tgaagatcaa aagcttattc atctgttttt ctttttcggt ggtgtaaagc caacaccctg 60  
 tctaaaaaac ataaatttct ttaatcattt tgccctctttt cctctgtgctt 110

<210> 25  
 <211> 110  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> fragment of SEQ ID No. 1  
 <400> 25  
 aagcttattc atctgttttt ctttttcggt ggtgtaaagc caacaccctg tctaaaaaac 60  
 ataaatttct ttaatcattt tgccctctttt cctctgtgctt caattaataa 110

<210> 26  
 <211> 110  
 <212> DNA  
 <213> Artificial Sequence

<220><223> fragment of SEQ ID No. 1  
 <400> 26  
 atctgttttt ctttttcggtt ggtgtaaagc caacaccctg tctaaaaaac ataaatttct 60  
 ttaatcattt tgcccttttt ctctgtgctt caattaataa aaaatggaaa 110

<210> 27  
 <211> 151  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> fragment of SEQ ID No. 1  
 <400> 27  
 cagcctacca tgagaataag agaaagaaaa tgaagatcaa aagcttattc atctgttttt 60  
 ctttttcggtt ggtgtaaagc caacaccctg tctaaaaaac ataaatttct ttaatcattt 120  
 tgcccttttt ctctgtgctt caattaataa a 151

<210> 28  
 <211> 121  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> fragment of SEQ ID No. 1  
 <400> 28  
 tgaagatcaa aagcttattc atctgttttt ctttttcggtt ggtgtaaagc caacaccctg 60  
 tctaaaaaac ataaatttct ttaatcattt tgcccttttt ctctgtgctt caattaataa 120  
 a 121

<210> 29  
 <211> 91  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> fragment of SEQ ID No. 1  
 <400> 29  
 ctttttcggtt ggtgtaaagc caacaccctg tctaaaaaac ataaatttct ttaatcattt 60  
 tgcccttttt ctctgtgctt caattaataa a 91

<210> 30  
 <211> 60  
 <212> DNA

<213> Artificial Sequence  
 <220><223> fragment of SEQ ID No. 1

<400> 30  
 aagcttattc atctgttttt ctttttcggt ggtgtaaagc caacaccctg tctaaaaaac 60  
 <210> 31  
 <211> 24  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> example for a histone stem-loop sequence

<400> 31  
 caaaggctct tttcagagcc acca 24  
 <210> 32  
 <211> 244  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> fragment of 3'UTR of an albumin gene

<400> 32  
 aaacatcaca attaagaaca tctcagccta ccatgagaac aagagaaata aaatgaagat 60  
 caaaagctta ttcactgttt tttctttttc atfggtataa agccaacacc ctgtctaaaa 120  
 aactataaat ttctttaatc attttgccctc ttttctctgt gcttcaatta ataaaaaatg 180  
 gaaagaatct agatctaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 240  
 aaaa 244  
 <210> 33  
 <211> 244  
 <212> RNA  
 <213> Artificial Sequence  
 <220><223> fragment of 3'UTR of an albumin gene

<400> 33  
 aaacaucaca auuaagaaca ucucagccua ccaugagaac aagagaaaua aaaugaagau 60  
 caaaagcuua uucaucuguu uuucuuuuuc auugguauaa agccaacacc cugucuaaaa 120  
 aacuauaaa uucuuuaauc auuuugccuc uuucucugu gcuucaauua auaaaaaauug 180  
 gaaagaaucu agaucuaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 240

aaaa 244

<210> 34

<211> 184

<212> DNA

<213> Artificial Sequence

<220><223> fragment of 3'UTR of an albumin gene

<400> 34

acacatcaca accacaacct tctcaggcta ccctgagaaa aaaagacatg aagactcagg 60

actcatcttt tctgttggtg taaaatcaac acctaagga acacaattt cttaaacaat 120

ttgacttctt gtctctgtgc tgcaattaat aaaaaatgga aagaatctac agatctaaaa 180

aaaa 184

<210> 35

<211> 184

<212> RNA

<213> Artificial Sequence

<220><223> fragment of 3'UTR of an albumin gene

<400> 35

acacaucaca accacaaccu ucucaggcua ccugagaaa aaaagacaug aagacucagg 60

acucaucuuu ucuguuggug uaaaaucaac acccuaagga acacaaauuu cuuuaaacau 120

uugacuucuu gucucugugc ugcaauuaau aaaaaaugga aagaauacuac agaucuaaaa 180

aaaa 184

<210> 36

<211> 260

<212> DNA

<213> Artificial Sequence

<220><223> fragment of 3'UTR of albumin plus poly(A)

<400> 36

aaacatcaca attaagaaca tctcagccta ccatgagaac aagagaaata aatgaagat 60

caaaagctta ttcactgttt tttcttttc attggtataa agccaacacc ctgtctaaaa 120

aactataaat ttctttaatc attttgctc ttttctctgt gttcaatta ataaaaaatg 180

gaaagaatct agatctaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 240

aaaaaaaaaa aaaaaaaaaa 260

<210> 37  
 <211> 260  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> fragment of 3'utr of albumin plus poly(a)  
 <400> 37  
 aaacaucaca auuaagaaca ucucagccua ccaugagaac aagagaaaua aaaugaagau 60  
 caaaagcuua uucaucuguu uuucuuuuuc auugguauaa agccaacacc cugucuaaaa 120  
 aacuauaaau uucuttaatc attttgctc ttttctctgt gcttcaatta ataaaaaatg 180  
 gaaagaatct agatctaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 240  
  
 aaaaaaaaaa aaaaaaaaaa 260  
 <210> 38  
 <211> 240  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> fragment of 3'UTR of albumin plus poly(A)  
 <400> 38  
 acacatcaca accacaacct tctcaggcta ccctgagaaa aaaagacatg aagactcagg 60  
 actcatcttt tctgttggtg taaaatcaac accctaagga acacaaattt ctttaaacat 120  
 ttgacttctt gtctctgtgc tgcaattaat aaaaaatgga aagaatctac agatctaaaa 180  
 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 240  
  
 <210> 39  
 <211> 240  
 <212> RNA  
 <213> Artificial Sequence  
 <220><223> fragment of 3'UTR of albumin plus poly(A)  
 <400> 39  
 acacaucaca accacaaccu ucucaggcua ccugagaaa aaaagacaug aagacucagg 60  
 acucaucuuu ucuguuggug uaaaaucaac acccuaagga acacaaauuu cuuuaacau 120  
 uugacuucuu gucucugugc ugcaauuuau aaaaaaugga aagaauucac agaucuaaaa 180  
 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 240  
 <210> 40  
 <211> 1936

<212> DNA

<213> Artificial Sequence

<220

><223> PpLuc(GC) - albumin8 - A64

<400> 40

```

gggagaaagc ttgaggatgg aggacgcaa gaacatcaag aaggccccg cgccttcta      60
cccgctggag gacgggaccg ccggcgagca gctccacaag gccatgaag ggtacgccct      120
ggtgccgggc acgatcgctt tcaccgacgc ccacatcgag gtcgacatca cctacgcgga      180
gtacttcgag atgagcgtgc gcctggccga ggccatgaag cggtagccgc tgaacaccaa      240
ccaccggatc gtggtgtgct cggagaacag cctgcagttc ttcattgccg tgctgggcgc      300
cctcttcacg ggcgtggccg tcgccccggc gaacgacatc tacaacgagc gggagctgct      360
gaacagcatg gggatcagcc agccgaccgt ggtgttcgtg agcaagaagg gcctgcagaa      420

gatcctgaac gtcagaaga agctgccccat catccagaag atcatcatca tggacagcaa      480
gaccgactac cagggttcc agtcgatgta cacgttcgtg accagccacc tcccgcggg      540
cttcaacgag tacgacttgc tcccggagag cttcgaccgg gacaagacca tcgacctgat      600
catgaacagc agcggcagca ccggcctgcc gaaggggggtg gccctgccgc accggaccgc      660
ctgcgtgcgc ttctgcacg cccgggacc cactttcggc aaccagatca tcccggacac      720
cgccatcctg agcgtgtgct cgttccacca cggcttcggc atgttcacga ccctgggcta      780
cctcatctgc ggcttccggg tggctctgat gtaccggttc gaggaggagc tgttcctgcg      840

gagcctgcag gactacaaga tccagagcgc gctgctcgtg ccgaccctgt tcagcttctt      900
cgccaagagc acctgatcg acaagtacga cctgtcgaac ctgcacgaga tcgccagcgg      960
gggcgcccc ctgagcaagg aggtgggcga ggccgtggcc aagcggttcc acctcccggg      1020
catccgccag ggtacggcc tgaccgagac cagcagcgcg atcctgatca cccccgaggg      1080
ggacgacaag ccgggcgccg tgggcaaggt ggtcccgttc ttcgaggcca aggtggtgga      1140
cctggacacc ggcaagacc tgggcgtgaa ccagcggggc gagctgtgcg tgcgggggcc      1200
gatgatcatg agcggctacg tgaacaacc ggaggccacc aacgccctca tcgacaagga      1260

cggctggctg cacagcggcg acatgccta ctgggacgag gacgagcact tcttcatcgt      1320
cgaccggctg aagtctgta tcaagtacaa gggctaccag gtggcgccgg ccgagctgga      1380
gagcatcctg ctccagcacc ccaacatctt cgacgccggc gtggccgggc tgccggacga      1440
cgacccggc gagctgccgg ccgcggtgtt ggtgctggag cacggcaaga ccatgacgga      1500
gaaggagatc gtcgactacg tggccagcca ggtgaccacc gccaagaagc tgcggggcgg      1560
cgtggtgttc gtggacgagg tcccgaaggg cctgaccggg aagctcgacg cccggaagat      1620

```

ccgcgagatc ctgatcaagg ccaagaaggg cggcaagatc gccgtgtaag actagtaaac 1680

atcacaatta agaacatctc agcctacat gagaacaaga gaaataaaat gaagatcaaa 1740

agcttattca tctgtttttc tttttcattg gtataaagcc aacacctgt ctaaaaaact 1800

ataaatttct ttaatcattt tgcctctttt cctgtgtctt caattaataa aaaatggaaa 1860

gaatctagat ctaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1920

aaaaaaaaaa aaaaaa 1936

<210> 41

<211> 1916

<212> DNA

<213> Artificial Sequence

<220><223> PpLuc(GC) - albumin9 - A64

<400> 41

gggagaaagc ttgaggatgg aggacgcaa gaacatcaag aagggcccg cgccttcta 60

cccgtggag gacgggaccg ccggcgagca gtcacacaag gccatgaag ggtacgccct 120

ggtgccgggc acgatcgct tcaccgacgc ccacatcgag gtcgacatca cctacgcgga 180

gtacttcgag atgagcgtgc gcctggccga ggccatgaag cggtagcggc tgaacaccaa 240

ccaccggate gtggtgtgct cggagaacag cctgcagttc tteatgccgg tgctgggcgc 300

cctcttcate ggcgtggccg tcgccccggc gaacgacatc tacaacgagc gggagctgct 360

gaacagcatg gggatcagcc agccgaccgt ggtgttcgtg agcaagaagg gcctgcagaa 420

gatcctgaac gtgcagaaga agctgccccat catccagaag atcatcatca tggacagcaa 480

gaccgactac cagggttcc agtcgatgta cacgttcgtg accagccacc tcccgcggg 540

cttcaacgag tacgacttcg tcccggagag cttcgaccgg gacaagacca tcgacctgat 600

catgaacagc agcggcagca ccggcctgcc gaaggggggtg gccctgccgc accggaccgc 660

ctgcgtgcgc ttctcgcacg cccgggaccc catcttcggc aaccagatca tcccggacac 720

cgccatcctg agcgtggtgc cgttccacca cggcttcggc atgttcacga cctgggcta 780

cctcatctgc ggcttccggg tggctctgat gtaccggttc gaggaggagc tgttcctgcg 840

gagcctcgag gactacaaga tccagagcgc gctgctcgtg cggacctgt tcagcttctt 900

cgccaagagc accctgatcg acaagtacga cctgtcgaac ctgcacgaga tcgccagcgg 960

gggcgccccg ctgagcaagg aggtgggcga ggccgtggcc aagcggttcc acctcccggg 1020

catccgccag ggetacggcc tgaccgagac caccagcgcg atcctgatca cccccaggg 1080

ggacgacaag ccgggcgccc tgggcaaggt ggtcccgttc ttcgaggcca aggtggtgga 1140

cctggacacc ggcaagacc tgggcgtgaa ccagcggggc gagctgtgcg tgcgggggcc	1200
gatgatcatg agcggctacg tgaacaacc ggaggccacc aacgcctca tcgacaagga	1260
cggctggctg cacagcggcg acatcgcta ctgggacgag gacgagcact tcttcatcgt	1320
cgaccggctg aagtcgctga tcaagtaca gggctaccag gtggcggcgg ccgagctgga	1380
gagcatcctg ctccagcacc ccaacatctt cgacgccggc gtggccgggc tgccggacga	1440
cgacgccggc gagctgccgg ccgcggtggt ggtgctggag cacggcaaga ccatgacgga	1500
gaaggagatc gtcgactacg tggccagcca ggtgaccacc gccaagaagc tgcggggcgg	1560
cgtggtgttc gtggacgagg tcccgaagg cctgaccggg aagctcgacg cccggaagat	1620
ccgcgagatc ctgatcaagg ccaagaagg cgcaagatc gccgtgtaag actagtacac	1680
atcacaacca caaccttctc aggctacct gagaaaaaaaa gacatgaaga ctcaggactc	1740
atcttttctg ttggtgtaaa atcaacacc taaggaacac aaatttcttt aaacatttga	1800
cttcttgtct ctgtgctgca attaataaaa aatggaaga atctacagat ctaaaaaaaa	1860
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaa	1916