(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2021-0124238
(43) 공개일자 2021년10월14일

(51) 국제특허분류(Int. Cl.)
C12N 15/62 (2006.01) *C12N 15/63* (2006.01)
C12N 9/22 (2006.01)
(52) CPC특허분류
C12N 15/62 (2013.01)
C12N 15/63 (2013.01)
(21) 출원번호 10-2021-7024731
(22) 출원일자(국제) 2020년01월03일
심사청구일자 없음
(85) 번역문제출일자 2021년08월04일
(86) 국제출원번호 PCT/US2020/012169
(87) 국제공개번호 WO 2020/142676
국제공개일자 2020년07월09일
(30) 우선권주장
62/788,571 2019년01월04일 미국(US)
(뒷면에 계속)

(71) 출원인
더 유니버서티 오브 시카고
미국 일리노이 60637 시카고 싸우스 엘리스 애버뉴 5801
(72) 발명자
디킨슨, 브라이언 씨.
미국, 60615 일리노이, 시카고 스위트 2씨, 하퍼 애비뉴 사우스 5113, 유니버시티 오브 시카고, 폴스카이 센터 포 엔터프리너십 앤 이노베이션
라우치, 시몬
미국, 60615 일리노이, 시카고 스위트 2씨, 하퍼 애비뉴 사우스 5113, 유니버시티 오브 시카고, 폴스카이 센터 포 엔터프리너십 앤 이노베이션
(74) 대리인
특허법인이룸리온

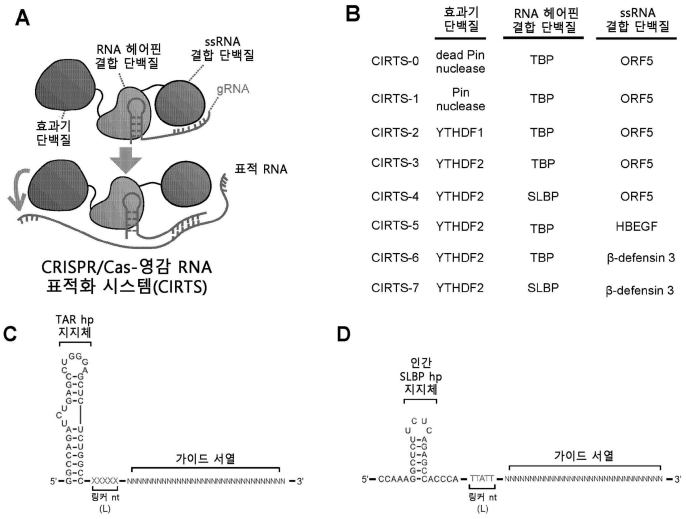
전체 청구항 수 : 총 94 항

(54) 발명의 명칭 **RNA 조절을 위한 시스템 및 방법**

(57) 요 약

RNA 조절을 위한 시스템 및 방법. 본 개시 내용의 측면은 RNA 조절 시스템과 관련되며, 상기 시스템은 i) RNA 헤어핀 결합 도메인; ii) RNA 표적화 영역 및 적어도 하나의 헤어핀 구조를 포함하는 RNA 표적화 분자(상기 RNA 표적화 분자의 헤어핀 구조는 i)에 특이적으로 결합함); 및 iii) RNA 조절 도메인; 각각 중 적어도 하나를 포함한다.

대 표 도



(52) CPC특허분류
C12N 9/22 (2013.01)
C07K 2319/09 (2013.01)
C07K 2319/35 (2013.01)
C07K 2319/85 (2013.01)
C12N 2310/351 (2013.01)
C12N 2310/531 (2013.01)

(30) 우선권주장
62/831,342 2019년04월09일 미국(US)
62/903,080 2019년09월20일 미국(US)
62/929,339 2019년11월01일 미국(US)

명 세 서

청구범위

청구항 1

RNA 조절 시스템으로서,

i) RNA 헤어핀 결합 도메인;

ii) RNA 표적화 영역 및 적어도 하나의 헤어핀 구조를 포함하는 RNA 표적화 분자 - 상기 RNA 표적화 분자의 상기 헤어핀 구조는 i)에 특이적으로 결합함 -; 및

iii) RNA 조절 도메인;

각각 중 적어도 하나를 포함하는, RNA 조절 시스템.

청구항 2

제1항에 있어서,

상기 RNA 헤어핀 결합 도메인 및 상기 RNA 조절 도메인은 작동 가능하게 연결되는 것을 특징으로 하는, RNA 조절 시스템.

청구항 3

제1항에 있어서,

i), ii), 및/또는 iii) 부분들은 인간의 것이거나 인간-유래인 것을 특징으로 하는, RNA 조절 시스템.

청구항 4

제2항 또는 제3항에 있어서,

i) 및 iii)은 펩타이드 결합을 통해 작동 가능하게 연결된 것을 특징으로 하는, RNA 조절 시스템.

청구항 5

제2항 또는 제3항에 있어서,

i) 및 iii)은 비공유 상호작용을 통해 작동 가능하게 연결된 것을 특징으로 하는, RNA 조절 시스템.

청구항 6

제1항 내지 제5항 중 어느 한 항에 있어서,

상기 RNA 조절 도메인은 제1 이량체화 도메인에 공유 결합되고 상기 RNA 헤어핀 결합 도메인은 제2 이량체화 도메인에 공유 결합되며, 상기 제1 및 제2 이량체화 도메인은 이량체화하여 비공유 또는 공유 연결을 형성할 수 있는 것을 특징으로 하는, RNA 조절 시스템.

청구항 7

제6항에 있어서,

상기 이량체화는 유도성(inducible)인 것을 특징으로 하는, RNA 조절 시스템.

청구항 8

제7항에 있어서,

상기 이량체화는 리간드-유도된 이량체화를 포함하는 것을 특징으로 하는, RNA 조절 시스템.

청구항 9

제8항에 있어서,

상기 제1 또는 제2 이량체화 도메인 중 하나는 PYR/PYR1-유사(PYL1)를 포함하고, 상기 제1 또는 제2 이량체화 도메인 중 다른 하나는 ABA 비민감성(insensitive) 1(ABI1)을 포함하고, 상기 리간드는 아브시스산(ABA) 또는 이의 유도체 또는 단편을 포함하는 것을 특징으로 하는, RNA 조절 시스템.

청구항 10

제8항에 있어서,

상기 제1 및/또는 제2 이량체화 도메인은 FKBP12를 포함하고 상기 리간드는 FK1012 또는 이의 유도체 또는 단편을 포함하는 것을 특징으로 하는, RNA 조절 시스템.

청구항 11

제8항에 있어서,

상기 제1 또는 제2 이량체화 도메인 중 하나는 FK506 결합 단백질(FKBP)을 포함하고, 상기 제1 또는 제2 도메인 중 다른 하나는 포유동물 Rap mTOR의 표적의 FKBP-Rap 결합 도메인(Frb)을 포함하고, 상기 리간드는 라파마이신(Rap) 또는 이의 유도체 또는 단편을 포함하는 것을 특징으로 하는, RNA 조절 시스템.

청구항 12

제1항 내지 제11항 중 어느 한 항에 있어서,

ii)는 적어도 두 개의 헤어핀 구조를 포함하는 것을 특징으로 하는, RNA 조절 시스템.

청구항 13

제1항 내지 제12항 중 어느 한 항에 있어서,

ii)는 하나 이상의 변형된 뉴클레오타이드를 포함하는 것을 특징으로 하는, RNA 조절 시스템.

청구항 14

제1항 내지 제13항 중 어느 한 항에 있어서,

상기 시스템은 안정화제 폴리펩타이드를 포함하되, 상기 안정화제 폴리펩타이드는 핵산에 비특이적으로 결합하는 양이온성 폴리펩타이드를 포함하는 것을 특징으로 하는, RNA 조절 시스템.

청구항 15

제14항에 있어서,

상기 안정화제 폴리펩타이드는 인간 유래인 것을 특징으로 하는, RNA 조절 시스템.

청구항 16

제1항 내지 제15항 중 어느 한 항에 있어서,

상기 시스템의 전체 크기는 150kDa 미만인 것을 특징으로 하는, RNA 조절 시스템.

청구항 17

제1항 내지 제16항 중 어느 한 항에 있어서,

i)은 U1A, SLBP, 또는 이의 변이체를 포함하는 것을 특징으로 하는, RNA 조절 시스템.

청구항 18

제1항 내지 제17항 중 어느 한 항에 있어서,

ii)는 서열번호: 1의 TAR 헤어핀 지지체(scaffold)를 포함하는 것을 특징으로 하는, RNA 조절 시스템.

청구항 19

제1항 내지 제18항 중 어느 한 항에 있어서,

ii)는 서열번호: 2의 SLBP 헤어핀 지지체를 포함하는 것을 특징으로 하는, RNA 조절 시스템.

청구항 20

제1항 내지 제19항 중 어느 한 항에 있어서,

ii)는 링커를 포함하는 것을 특징으로 하는, RNA 조절 시스템.

청구항 21

제20항에 있어서,

상기 링커는 적어도 5개 아미노산인 것을 특징으로 하는, RNA 조절 시스템.

청구항 22

제1항 내지 제21항 중 어느 한 항에 있어서,

상기 RNA 표적화 영역은 적어도 12개의 뉴클레오타이드를 포함하는 것을 특징으로 하는, RNA 조절 시스템.

청구항 23

제1항 내지 제22항 중 어느 한 항에 있어서,

iii)은 뉴클레아제, 메틸라아제, 데메틸라아제(demethylase), 번역 활성제, 번역 억제인자, 단일 가닥 RNA 절단 활성, 이중 가닥 RNA 절단 활성, 또는 RNA 결합 활성을 포함하는 것을 특징으로 하는, RNA 조절 시스템.

청구항 24

제1항 내지 제23항 중 어느 한 항에 있어서,

iii)은 Pin 뉴클레아제 도메인 또는 m6A 판독자 단백질(reader protein) 또는 이의 부분을 포함하는 것을 특징으로 하는, RNA 조절 시스템.

청구항 25

제24항에 있어서,

iii)은 YTHDF1, YTHDF2, 또는 ADAR을 포함하는 것을 특징으로 하는, RNA 조절 시스템.

청구항 26

제14항 내지 제25항 중 어느 한 항에 있어서,

상기 안정화제 단백질은 HBEGF, 베타-디펜신, 또는 이의 변이체 또는 부분을 포함하는 것을 특징으로 하는, RNA 조절 시스템.

청구항 27

제1항 내지 제26항 중 어느 한 항에 있어서,

ii)의 상기 RNA 표적화 영역은 원핵 또는 진핵 세포에서 표적 RNA에 혼성화(hybridize)되는 것을 특징으로 하는, RNA 조절 시스템.

청구항 28

제1항 내지 제27항 중 어느 한 항에 있어서,

i) 및/또는 iii)은 하나 이상의 핵 국소화 신호(NLS)를 포함하는 것을 특징으로 하는, RNA 조절 시스템.

청구항 29

제1항 내지 제28항 중 어느 한 항에 있어서,

상기 시스템은 각각의 i), ii), 및 iiii) 중에서 적어도 두 개를 포함하는 것을 특징으로 하는, RNA 조절 시스템.

청구항 30

제1항 내지 제29항 중 어느 한 항에 있어서,

상기 RNA 조절 도메인은 RNA를 절단하거나, RNA 번역을 촉진하거나, RNA 번역을 억제하거나, 또는 RNA의 염기서열을 변형하는 것을 특징으로 하는, RNA 조절 시스템.

청구항 31

하나 이상의 핵산 벡터를 포함하는 벡터 시스템으로서,

i) RNA 헤어핀 결합 도메인; ii) RNA 표적화 영역 및 적어도 하나의 헤어핀 구조를 포함하는 RNA 표적화 분자 - 상기 RNA 표적화 분자의 상기 헤어핀 구조는 i)에 특이적으로 결합함 -, 및 iii) RNA 조절 도메인;을 암호화하는 뉴클레오타이드를 포함하는, 벡터 시스템.

청구항 32

제31항에 있어서,

상기 RNA 헤어핀 결합 도메인 및 상기 RNA 조절 도메인은 작동 가능하게 연결되는 것을 특징으로 하는, 벡터 시스템.

청구항 33

제31항에 있어서,

i), ii), 및/또는 iii)을 암호화하는 뉴클레오타이드에 작동 가능하게 연결된 조절 요소를 더 포함하는, 벡터 시스템.

청구항 34

제31항 또는 제33항에 있어서,

상기 하나 이상의 핵산 벡터는 진핵 세포 내 발현에 최적화되는 것을 특징으로 하는, 벡터 시스템.

청구항 35

제31항 내지 제34항 중 어느 한 항에 있어서,

상기 발현은 구성적(constitutive) 또는 조건적(conditional)인 것을 특징으로 하는, 벡터 시스템.

청구항 36

제31항 내지 제35항 중 어느 한 항에 있어서,

i), ii), 및 iii)은 단일 벡터 상에 있는 것을 특징으로 하는, 벡터 시스템.

청구항 37

제31항 내지 제36항 중 어느 한 항에 있어서,

상기 벡터 중 하나 이상은 바이러스 벡터인 것을 특징으로 하는, 벡터 시스템.

청구항 38

제31항 내지 제37항 중 어느 한 항에 있어서,

상기 하나 이상의 벡터는 하나 이상의 레트로바이러스 벡터, 렌티바이러스 벡터, 아데노바이러스 벡터, 아데노연관 바이러스 벡터 또는 단순 헤르페스 바이러스 벡터를 포함하는 것을 특징으로 하는, 벡터 시스템.

청구항 39

제31항 내지 제36항 중 어느 한 항에 있어서,

상기 벡터 중 하나 이상은 비바이러스 벡터인 것을 특징으로 하는, 벡터 시스템.

청구항 40

RNA 표적화 분자에 작동 가능하게 연결된 RNA 조절 도메인을 포함하는 접합체로서, 상기 RNA표적화 분자는 RNA 표적화 영역 및 적어도 하나의 헤어핀 구조를 포함하는 것을 특징으로 하는, 접합체.

청구항 41

제40항에 있어서,

상기 RNA 조절 도메인은 인간 유래인 것을 특징으로 하는, 접합체.

청구항 42

제40항 또는 제41항에 있어서,

상기 RNA 조절 도메인 및 상기 RNA 표적화 분자는 펩타이드 결합을 통해 작동 가능하게 연결되는 것을 특징으로 하는, 접합체.

청구항 43

제40항 내지 제42항 중 어느 한 항에 있어서,

상기 폴리펩타이드는 하나 이상의 링커를 더 포함하는 것을 특징으로 하는, 접합체.

청구항 44

제40항 또는 제41항에 있어서,

상기 RNA 조절 도메인 및 상기 RNA 표적화 분자는 비공유 상호작용을 통해 작동 가능하게 연결되는 것을 특징으로 하는, 접합체.

청구항 45

제40항 내지 제44항 중 어느 한 항에 있어서,

상기 RNA 조절 도메인은 제1 이량체화 도메인에 공유 결합되고 상기 RNA 표적화 분자는 제2 이량체화 도메인에 공유 결합되며, 상기 제1 및 제2 이량체화 도메인은 이량체화하여 비공유 또는 공유 연결을 형성할 수 있는 것을 특징으로 하는, 접합체.

청구항 46

제45항에 있어서,

상기 이량체화는 유도성(inducible)인 것을 특징으로 하는, 접합체.

청구항 47

제46항에 있어서,

상기 이량체화는 리간드-유도된 이량체화를 포함하는 것을 특징으로 하는, 접합체.

청구항 48

제47항에 있어서,

상기 제1 또는 제2 이량체화 도메인 중 하나는 PYR/PYR1-유사(PYL1)를 포함하고, 상기 제1 또는 제2 이량체화 도메인 중 다른 하나는 ABA 비민감성(insensitive) 1(ABI1)을 포함하고, 상기 리간드는 아브시스산(ABA) 또는 이의 유도체 또는 단편을 포함하는 것을 특징으로 하는, 접합체.

청구항 49

제47항에 있어서,

상기 제1 및/또는 제2 이량체화 도메인은 FKBP12를 포함하고 상기 리간드는 FK1012 또는 이의 유도체 또는 단편을 포함하는 것을 특징으로 하는, 접합체.

청구항 50

제47항에 있어서,

상기 제1 또는 제2 이량체화 도메인 중 하나는 FK506 결합 단백질(FKBP)을 포함하고, 상기 제1 또는 제2 도메인 중 다른 하나는 포유동물 Rap mTOR의 표적의 FKBP-Rap 결합 도메인(Frb)을 포함하고, 상기 리간드는 라파마이신(Rap) 또는 이의 유도체 또는 단편을 포함하는 것을 특징으로 하는, 접합체.

청구항 51

제40항 내지 제50항 중 어느 한 항에 있어서,

상기 RNA 표적화 분자는 적어도 2개의 헤어핀 구조를 포함하는 것을 특징으로 하는, 접합체.

청구항 52

제40항 내지 제51항 중 어느 한 항에 있어서,

상기 RNA 표적화 분자는 하나 이상의 변형된 뉴클레오타이드를 포함하는 것을 특징으로 하는, 접합체.

청구항 53

제40항 내지 제52항 중 어느 한 항에 있어서,

상기 RNA 표적화 분자는 서열번호: 1의 TAR 헤어핀 지지체를 포함하는 것을 특징으로 하는, 접합체.

청구항 54

제40항 내지 제53항 중 어느 한 항에 있어서,

상기 RNA 표적화 분자는 서열번호: 2의 SLBP 헤어핀 지지체를 포함하는 것을 특징으로 하는, 접합체.

청구항 55

제40항 내지 제54항 중 어느 한 항에 있어서,

상기 RNA 표적화 분자는 링커를 포함하는 것을 특징으로 하는, 접합체.

청구항 56

제55항에 있어서,

상기 링커는 적어도 5개 아미노산을 포함하는 것을 특징으로 하는, 접합체.

청구항 57

제40항 내지 제56항 중 어느 한 항에 있어서,

상기 RNA 표적화 영역은 적어도 12개의 뉴클레오타이드를 포함하는 것을 특징으로 하는, 접합체.

청구항 58

제40항 내지 제57항 중 어느 한 항에 있어서,

상기 RNA 조절 도메인은 뉴클레아제, 메틸라아제, 데메틸라아제(demethylase), 번역 활성제, 번역 억제인자, 단일 가닥 RNA 절단 활성, 이중 가닥 RNA 절단 활성, 또는 RNA 결합 활성을 포함하는 것을 특징으로 하는, 접합체.

청구항 59

제40항 내지 제57항 중 어느 한 항에 있어서,

상기 RNA 조절 도메인은 Pin 뉴클레아제 도메인 또는 m6A 판독자 단백질(reader protein) 또는 이의 부분을 포함하는 것을 특징으로 하는, 접합체.

청구항 60

제40항 내지 제59항 중 어느 한 항에 있어서,

상기 RNA 조절 도메인은 YTHDF1, YTHDF2, 또는 ADAR을 포함하는 것을 특징으로 하는, 접합체.

청구항 61

제40항 내지 제60항 중 어느 한 항에 있어서,

상기 RNA 표적화 분자의 상기 RNA 표적화 영역은 원핵 또는 진핵 세포에서 표적 RNA에 혼성화(hybridize)되는 것을 특징으로 하는, 접합체.

청구항 62

제40항 내지 제61항 중 어느 한 항에 있어서,

상기 접합체는 하나 이상의 핵 국소화 신호(NLS)를 포함하는 것을 특징으로 하는, 접합체.

청구항 63

제40항 내지 제62항 중 어느 한 항에 있어서,

상기 RNA 조절 도메인은 RNA를 절단하거나, RNA 번역을 촉진하거나, RNA 번역을 억제하거나, 또는 RNA의 염기서열을 변형하는 것을 특징으로 하는, 접합체.

청구항 64

RNA 헤어핀 결합 도메인 및 RNA 조절 도메인을 포함하는, 융합 단백질.

청구항 65

RNA 조절 도메인 및 제1 이량체화 도메인을 포함하는, 융합 단백질.

청구항 66

RNA 헤어핀 결합 도메인 및 제2 이량체화 도메인을 포함하는, 융합 단백질.

청구항 67

제65항 또는 제66항에 있어서,

상기 제1 및/또는 제2 이량체화 도메인의 이량체화는 유도성(inducible)인 것을 특징으로 하는, 융합 단백질.

청구항 68

제67항에 있어서,

상기 이량체화는 리간드-유도된 이량체화를 포함하는 것을 특징으로 하는, 융합 단백질.

청구항 69

제68항에 있어서,

상기 제1 및 제2 이량체화 도메인은 PYL1 및 ABI1 중에서 선택되고, 상기 리간드는 ABA 또는 이의 유도체 또는 단편을 포함하는 것을 특징으로 하는, 융합 단백질.

청구항 70

제68항에 있어서,

상기 제1 및/또는 제2 이량체화 도메인은 FKBP12를 포함하고 상기 리간드는 FK1012 또는 이의 유도체 또는 단편을 포함하는 것을 특징으로 하는, 융합 단백질.

청구항 71

제68항에 있어서,

상기 제1 및 제2 이량체화 도메인은 FKBP 및 Frb 중에서 선택되고, 상기 리간드는 라파마이신 또는 이의 유도체 또는 단편을 포함하는 것을 특징으로 하는, 융합 단백질.

청구항 72

제64항 내지 제71항 중 어느 한 항에 있어서,

상기 RNA 헤어핀 결합 도메인 및/또는 상기 RNA 조절 도메인은 인간 유래인 것을 특징으로 하는, 융합 단백질.

청구항 73

제64항 내지 제72항 중 어느 한 항에 있어서,

상기 융합 단백질은 150kDa 미만인 것을 특징으로 하는, 융합 단백질.

청구항 74

제64항 내지 제73항 중 어느 한 항에 있어서,

상기 RNA 헤어핀 결합 도메인은 U1A, SLBP, 또는 이의 변이체 또는 단편을 포함하는 것을 특징으로 하는, 융합 단백질.

청구항 75

제64항 내지 제74항 중 어느 한 항에 있어서,

상기 RNA 조절 도메인은 뉴클레아제, 메틸라아제, 데메틸라아제(demethylase), 번역 활성제, 번역 억제인자, 단일 가닥 RNA 절단 활성, 이중 가닥 RNA 절단 활성, 또는 RNA 결합 활성을 포함하는 것을 특징으로 하는, 융합 단백질.

청구항 76

제75항에 있어서,

상기 RNA 조절 도메인은 Pin 뉴클레아제 도메인 또는 m6A 판독자 단백질 또는 이의 부분을 포함하는 것을 특징으로 하는, 융합 단백질.

청구항 77

제76항에 있어서,

상기 RNA 조절 도메인은 YTHDF1 또는 YTHDF2를 포함하는 것을 특징으로 하는, 융합 단백질.

청구항 78

제64항 내지 제77항 중 어느 한 항에 있어서,

하나 이상의 핵 국소화 신호(NLS)를 더 포함하는, 융합 단백질.

청구항 79

제64항 내지 제77항 중 어느 한 항에 있어서,

상기 RNA 조절 도메인은 RNA를 절단하거나, RNA 번역을 촉진하거나, RNA 번역을 억제하거나, 또는 RNA의 염기 서열을 변형하는 것을 특징으로 하는, 융합 단백질.

청구항 80

제64항 내지 제79항 중 어느 한 항의 융합 단백질을 암호화하는, 핵산.

청구항 81

제1항 내지 제39항 중 어느 한 항의 시스템, 제40항 내지 제63항 중 어느 한 항의 접합체, 또는 제64항 내지 제79항 중 어느 한 항의 융합 단백질을 포함하는, 전달 매개체.

청구항 82

제81항에 있어서,

상기 전달 매개체는 리포솜(들), 입자(들), 엑소좀(들), 미세소포(들), 유전자총 또는 하나 이상의 핵산 벡터(들)을 포함하는, 전달 매개체.

청구항 83

제1항 내지 제39항 중 어느 한 항의 시스템, 제40항 내지 제63항 중 어느 한 항의 접합체, 제64항 내지 제79항 중 어느 한 항의 융합 단백질, 또는 제81항 내지 제82항 중 어느 한 항의 전달 매개체를 포함하는, 조성물.

청구항 84

제1항 내지 제39항 중 어느 한 항의 시스템, 제40항 내지 제63항 중 어느 한 항의 접합체, 제64항 내지 제79항 중 어느 한 항의 융합 단백질, 제81항 내지 제82항 중 어느 한 항의 전달 매개체, 또는 제83항의 조성물을 포함하는, 세포.

청구항 85

적어도 하나의 표적 RNA를 조절하는 방법으로서,

상기 표적 RNA를 제1항 내지 제39항 중 어느 한 항의 시스템, 제40항 내지 제63항 중 어느 한 항의 접합체, 제64항 내지 제79항 중 어느 한 항의 융합 단백질, 제81항 내지 제82항 중 어느 한 항의 전달 매개체, 또는 제83항의 조성물과 접속시키는 단계를 포함하는, 방법.

청구항 86

제85항에 있어서,

상기 적어도 하나의 RNA를 조절하는 것은 상기 RNA를 절단하는 단계, 탈메틸화하는 단계, 메틸화하는 단계, 번역 활성화하는 단계, 번역 억제하는 단계, 분해 촉진하는 단계, 및/또는 그에 결합하는 단계를 포함하는 것을 특징으로 하는, 방법.

청구항 87

제85항 또는 제86항에 있어서,

상기 표적 RNA는 원핵 또는 진핵 세포 내에 있는 것을 특징으로 하는, 방법.

청구항 88

제87항에 있어서,

상기 표적 RNA는 인간 세포 내에 있는 것을 특징으로 하는, 방법.

청구항 89

제87항 또는 제88항에 있어서,

상기 표적 RNA는 시험관 내(vitro) 또는 생체 내(in vivo)에 있는 것을 특징으로 하는, 방법.

청구항 90

조절된 표적 RNA를 포함하는 세포 또는 이의 자손으로서,

상기 표적 RNA는 제85항 내지 제89항 중 어느 한 항에 따라서 조절된 것을 특징으로 하는, 세포 또는 이의 자손.

청구항 91

제90항에 따른 하나 이상의 세포를 포함하는, 다세포 생물.

청구항 92

제91항에 따른 하나 이상의 세포를 포함하는, 식물 또는 동물.

청구항 93

제1항 내지 제39항 중 어느 한 항의 시스템, 제40항 내지 제63항 중 어느 한 항의 접합체, 제64항 내지 제79항 중 어느 한 항의 융합 단백질, 제81항 내지 제82항 중 어느 한 항의 전달 매개체, 또는 제83항의 조성물을 포함하는, 키트.

청구항 94

대상체에서 표적 RNA를 조절하는 방법으로서,

상기 방법은 제40항 내지 제63항 중 어느 한 항의 접합체, 제64항 내지 제79항 중 어느 한 항의 융합 단백질, 제81항 내지 제82항 중 어느 한 항의 전달 매개체, 또는 제83항의 조성물을 투여하는 단계를 포함하는, 방법.

발명의 설명

기 술 분 야

[0001] 본 발명은 일반적으로 화학 및 의학 분야에 관한 것이다. 보다 구체적으로, 본 발명은 RNA를 조절하기 위한 시스템의 용도에 관한 것이다.

배 경 기 술

[0002] 프로그램 가능한 핵산-결합 단백질은 유전체 연구 및 편집 기술에 혁신을 일으켰고(Chandrasegaran and Carroll, 2016; Filipovska et al., 2011; Gootenberg et al., 2018; Hilton et al., 2015; Joung and Sander, 2012; Kearns et al., 2015; Strutt et al., 2018), 인간의 질병들을 치료하기 위한 새로운 치료 기회를 열어주고 있다(Liao et al., 2017; Monteys et al., 2017). 특히, 박테리아 면역 방어 메커니즘으로서 진화된 CRISPR/Cas9 시스템은 세포 DNA를 부위-특이적으로 연구하고 조작하는 능력을 완전히 바꾸어 놓았다(Cong et al., 2013; Jiang et al., 2013; O'Connell et al., 2014; Wiedenheft et al., 2012). 이전의 방법들(Desjarlais and Berg, 1993; Hockemeyer et al., 2011; Joung and Sander, 2012; Schierling et al., 2012)과 비교하여 CRISPR/Cas 시스템의 주요 이점은 이들이 실질적으로 임의의 관심 자리를 표적화하도록 용이하게 프로그램 가능하다는 것이다. CRISPR/Cas 시스템은 디스플레이된 가이드 RNA(gRNA)의 염기쌍 상호작용을 사용하여 표적 핵산 서열과 상호작용하는 리보핵산단백질 복합체이다. 염기쌍-가이드된 표적화의 단순한 성질은 가이드 가닥 상의 핵산 서열을 간단히 변화시킴으로써 규정된 핵산 서열과 상호작용하는 시스템을 프로그램할 가능성을 열어준다.

[0003] DNA를 직접적으로 표적화하는 것은 상당한 임상적 파급효과를 가질 것이지만, 많은 유전자에 대한 미묘한 변경을 수반하는 질병은 DNA 편집 기술을 사용하여 표적화하기 어려울 것이다(Fuxman Bass et al., 2015). 추가로, 영구적인 유전자 변경의 잠재적 부작용 또는 위험은 용인되지 않을 수 있다. 예를 들어, 향상된 상처 치유 반응을 활성화하기 위한 표적화를 원할 수 있는 유전자는 암 발생에 대한 위험이 있을 수 있는 표적일 가능성이 있으며, 이는 영구적인 DNA-기반 전략을 위험하게 만든다. RNA 수준에서 정보 흐름을 표적화하는 것은, 부작용이 나타나는 경우 치료를 중단시키는 능력, DNA 수준에서 변경되기에 너무 위험할 수 있는 유전자를 표적화하는 능력, 및 숙주 유전체에 대한 영구적인 변경 없이 유전자 발현을 조작하는 능력을 포함하는(그러나 이에 제한되지 않는), 치료적 개입을 위한 여러 기회를 제공한다. 유전체 수준에서 전사를 억제 또는 증강시키는 것은 유전자

발현을 제어하는 하나의 가능성을 제공하지만(Du et al., 2017; Fuxman Bass et al., 2015; Qi et al., 2013), 최근에 발견된 RNA 후성전사체 조절 메커니즘은 RNA 전사물의 편집, 분해, 수송 및 번역을 포함하는 표적에 대한 광범위한 RNA 조절 방법을 제공한다(Nishikura, 2010; Roundtree et al., 2017; Zhao et al., 2017). 이러한 후성전사체 조절층의 메커니즘 및 결과는 이제 밝혀지기 시작하고 있지만, RNA를 통한 정보 흐름이 엄격하게 조절되어, 기본적인 연구 발견 뿐만 아니라 치료 개발 둘 모두에 대한 많은 새로운 기회를 제공하는 것이 명백하다.

[0004] dCas9 DNA 표적화 시스템과 유사한 프로그램 가능한 RNA 표적화 도구는 후성전사체 조절 메커니즘을 연구하는데 그리고 치료 용도를 위해 매우 유망하다. RNA 표적화를 위한 현재의 도구는 큰 복합체의 전달을 포함하고 면역원성 문제를 제기한다. 기본적인 과학 관점에서 볼 때, 전달 매개체(또는, '전달체')의 큰 크기는 조사 중인 RNA에 잠재적인 교란을 일으켜, RNA 조절 메커니즘 연구를 복잡하게 만들 수 있다. 번역 관점에서, 큰 크기는 바이러스 패키징 또는 직접 단백질 전달에 대한 문제를 나타낸다. 또한 DNA 편집 요법은 일회성, 비가역적 치료로 구성될 가능성이 높지만, RNA 표적화 요법은 지속적으로 투여되어야 하므로 전달 문제가 특히 중요하다. 더욱이 최근에 사람들의 85%가 이미 CRISPR/Cas 단백질에 대한 순환 항체를 가지고 있다는 것이 발견되었으며(Kim et al., 2018; Wagner et al., 2018), 이는 면역원성 문제가 임상 적용에서 문제가 될 수 있음을 시사한다. 따라서, RNA를 표적화할 수 있고 면역 반응을 활성화지 않고 효율적으로 전달될 수 있는 개선된 시스템이 당업계에 필요하다.

발명의 내용

[0005] **관련 출원에 대한 상호 참조**

[0006] 본 출원은 2019년 1월 4일 출원된 미국 가출원 제62/788,571호, 2019년 4월 9일 출원된 미국 가출원 제62/831,342호, 2019년 9월 20일 출원된 미국 가출원 제62/903,080호, 및 2019년 11월 1일 출원된 미국 가출원 제62/929,339호의 우선권을 주장하고, 이의 내용들은 그 전체가 참조로 포함된다.

[0007] **정부 지원 진술**

[0008] 본 발명은 미국 국립보건원(The National Institutes of Health)에 의해 수여된 GM119840 및 HG008935 하의 정부 지원으로 이루어졌다. 정부는 본 발명에서 특정 권리를 갖는다.

[0009] 현재 RNA 표적화 시스템의 큰 크기와 미생물 유래 성질을 극복하기 위해, 본 발명자들은 프로그램 가능한 RNA 효과기 단백질(RNA effector protein)을 조작하기 위한 일반적인 방법인 CRISPR/Cas-영감 RNA 표적화 시스템(CIRTS)을 제시한다. CRISPR/Cas 기반 시스템과 유사하게, CIRTS는 왓슨-크릭-프랭클린(Watson-Crick-Franklin) 염기쌍 상호작용을 사용하여 전사체에서 단백질 카고(protein cargo) 부위를 선택적으로 전달하는 리보핵산단백질 복합체이다. 본 발명자들은 분해를 위한 뉴클레아제, 분해를 위한 탈아데닐화 조절 기구, 또는 강화된 단백질 생산을 위한 번역 활성화 기구를 포함하는 전사체에 일정 범위 또는 조절 단백질을 전달하는 CIRTS를 쉽게 조작할 수 있음을 보여준다. 그러나 CIRTS는 현재 가장 작은 CRISPR/Cas 시스템보다 최대 5배 작으며 전적으로 인간 부분으로부터 조작될 수 있다.

[0010] 본 개시 내용의 측면은 RNA 조절 시스템 또는 방법과 관련되며, 상기 시스템 또는 방법은 i) RNA 헤어핀 결합 도메인; ii) RNA 표적화 영역 및 적어도 하나의 헤어핀 구조를 포함하는 RNA 표적화 분자 -상기 RNA 표적화 분자의 헤어핀 구조는 i)에 특이적으로 결합함-; 및 iii) RNA 조절 도메인; 각각 중 적어도 하나를 포함한다. 일부 구현예에서, 다음이 포함된다: i) 및 ii), i) 및 iii), ii) 및 iii), 또는 i), ii), 및 iii). 본원에 개시된 일 구현예는 이 조합들 중 어느 하나를 함유한다.

[0011] 추가 측면은 하나 이상의 핵산 벡터를 포함하는 벡터 시스템에 관한 것으로, 상기 벡터 시스템은 i) RNA 헤어핀 결합 도메인; ii) RNA 표적화 영역 및 적어도 하나의 헤어핀 구조를 포함하는 RNA 표적화 분자 - 상기 RNA 표적화 분자의 상기 헤어핀 구조는 i)에 특이적으로 결합함 -, 및 iii) RNA 조절 도메인;을 암호화하는 뉴클레오타이드를 포함한다.

[0012] 추가 측면은 융합 단백질에 관한 것으로, 상기 융합 단백질은 RNA 헤어핀 결합 단백질 및 RNA 조절 도메인 및 상기 융합 단백질을 암호화하는 핵산을 포함한다.

[0013] 추가 측면은 RNA 표적화 분자에 작동 가능하게 연결된 RNA 조절 도메인을 포함하는 접합체에 관한 것으로, 상기 RNA표적화 분자는 RNA 표적화 영역 및 적어도 하나의 헤어핀 구조를 포함한다. 일부 구현예에서, 상기 RNA 조절 도메인 및 상기 RNA 표적화 분자는 펩타이드 결합을 통해 작동 가능하게 연결된다. 일부 구현예에서, 상기

폴리펩타이드는 하나 이상의 링커를 더 포함한다. 일부 구현예에서, 상기 RNA 조절 도메인 및 상기 RNA 표적화 분자는 비공유 상호작용을 통해 작동 가능하게 연결된다. 일부 구현예에서, 상기 RNA 조절 도메인은 제1 이량체화 도메인에 공유 결합되고 상기 RNA 표적화 분자는 제2 이량체화 도메인에 공유 결합되며 상기 제1 및 제2 이량체화 도메인은 이량체화하여 비공유 또는 공유 연결을 형성할 수 있다. 일부 구현예에서, 상기 접합체는 하나 이상의 핵 국소화 신호(NLS)를 포함한다.

[0014] 추가 측면은 본 개시 내용의 시스템을 포함하는 전달 매개체에 관한 것이다. 일부 구현예에서, 상기 전달 매개체는 리포솜(들), 입자(들), 엑소좀(들), 미세소포(들), 유전자총 또는 하나 이상의 핵산 벡터(들)을 포함한다.

[0015] 추가 측면은 본 개시 내용의 시스템, 전달 매개체, 또는 융합 단백질을 포함하는 조성물 또는 세포에 관한 것이다.

[0016] 추가 측면은 상기 표적 RNA를 본 개시 내용의 시스템, 조성물, 또는 융합 단백질과 접촉시키는 단계를 포함하는 적어도 하나의 RNA를 조절하는 방법에 관한 것이다. 일부 구현예에서, 적어도 하나의 RNA를 조절하는 것은 상기 RNA를 절단하는 단계, 탈메틸화하는 단계, 메틸화하는 단계, 번역 활성화하는 단계, 번역 억제하는 단계, 분해 촉진하는 단계, 또는 그에 결합하는 단계를 포함한다. 일부 구현예에서, 적어도 2개의 표적 RNA는 조절된다. 일부 구현예에서, 적어도 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 또는 25개(또는 여기에 속하는 임의의 추론가능한 범위)의 표적 RNA는 조절된다. 일부 구현예에서, 다수의 RNA는 동일한 RNA 조절 도메인에 의해 또는 동일한 활성을 가진 조절 도메인에 의해 조절된다. 일부 구현예에서, 상이한 RNA들은 예컨대 상기 RNA를 절단, 탈메틸화, 메틸화, 번역 활성화, 번역 억제, 분해 촉진, 또는 그에 결합함으로써, 상이한 활성으로 조절된다.

[0017] 일부 구현예에서, 상기 RNA 조절 도메인은 RNA에 결합하지 않는다. 일부 구현예에서, 상기 RNA 조절 도메인은 RNA 결합 활성을 갖지 않은 폴리펩타이드를 포함한다. 일부 구현예에서, 상기 RNA 조절 도메인은 변형된 RNA에 결합하지 않는다. 일부 구현예에서, 상기 RNA 조절 도메인은 m6A 변형된 RNA에 결합하지 않는다.

[0018] 추가 측면은 조절된 표적 RNA를 포함하는 세포 또는 이의 자손에 관한 것으로, 상기 표적 RNA는 본 개시 내용의 방법에 따라서 조절되었다. 추가 측면은 본 개시 내용의 하나 이상의 세포를 포함하는 다세포 생물에 관한 것이다. 추가 측면은 본 개시 내용의 하나 이상의 세포를 포함하는 식물 또는 동물에 관한 것이다. 추가 측면은 본 개시 내용의 시스템, 벡터, 전달 매개체, 또는 융합 단백질을 포함하는 키트에 관한 것이다.

[0019] 추가 측면은 대상체에서 표적 RNA를 조절하는 방법에 관한 것으로, 상기 방법은 본 개시 내용의 시스템 또는 조성물을 상기 대상체에 투여하는 단계를 포함한다.

[0020] 용어 "RNA 헤어핀"은 줄기-루프 분자내 염기쌍을 가진 RNA 분자를 지칭한다. 헤어핀은 동일한 가닥의 두 영역(일반적으로 반대 방향으로 읽을 때 뉴클레오타이드 서열에서 상보적)이 염기쌍을 형성하여 쌍을 이루지 않은 루프로 끝나는 이중 나선을 형성할 때 발생할 수 있다. 본 개시 내용은 RNA 표적화 영역 및 하나 이상의 헤어핀을 포함하는 조작된 RNA 표적화 분자에 관한 것이다. 따라서, 본 개시 내용의 조작된 RNA 분자는 비천연으로 발생하는 키메라 분자이다.

[0021] 용어 "RNA 표적화 영역"은 표적 RNA에 혼성화할 수 있는 RNA의 영역을 지칭한다. 표적 RNA는 질병 관련 RNA이거나 현재 시스템 및 방법에 따른 조절 표적인 RNA일 수 있다.

[0022] "RNA 조절 도메인"은 RNA에 대한 활성을 갖는 펩타이드 또는 폴리펩타이드를 지칭한다. 활성의 예로는 메틸화 활성, RNA-결합 활성, 뉴클레아제 활성, 및 번역 활성화 또는 억제 활성이 포함된다. RNA 조절 도메인을 포함하는 활성 및 단백질의 추가 예는 본 개시 내용 전반에 걸쳐 설명한다.

[0023] 일부 구현예에서, 상기 RNA 헤어핀 결합 도메인 및 상기 RNA 조절 도메인은 작동 가능하게 연결된다. "작동가능하게 연결된(operably linked)"이라는 용어는 공유 또는 비공유 상호작용을 통해 연결된 두 개의 단백질을 지칭한다. 예를 들어, 상기 두 개의 단백질은 펩타이드 결합을 통해 공유 연결될 수 있다. 일부 구현예에서, 상기 단백질은 비공유 연결된다. 본 개시 내용의 하나 이상의 단백질은 서로에 대해 강한 친화성을 갖는 한 쌍의 부속 단백질에 대한 연결을 통해 다른 단백질에 작동가능하게 연결될 수 있다. 상기 부속 단백질은 당업계에 공지되어 있다. 예를 들어, SunTag는 펩타이드에 대한 강한 친화성을 갖는 항체를 포함하는 상기와 같은 하나의 시스템이다. 본 개시 내용의 하나의 단백질, 폴리펩타이드 또는 도메인이 SunTag 펩타이드에 연결되고 본 개시 내용의 다른 단백질, 폴리펩타이드 또는 도메인이 항체에 연결되어 SunTag 펩타이드 및 항체의 상호작용을 통해 두 단백질, 폴리펩타이드 또는 도메인이 작동가능하게 연결되도록 할 수 있다. 추가 예에는 비오틴과 아

비딘/스트렙타비딘 및 스파이태그(spytag)와 스파이캐처(spycatcher)가 포함된다.

[0024] 일부 구현예에서, 시스템은 자극제의 존재시 연결되는 2개의 연결되지 않은 폴리펩타이드로서 RNA 조절 도메인 및 헤어핀 결합 도메인을 제공함으로써 유도될 수 있다(inducible). 유도는, 예를 들어 광 유도 또는 화학적 유도에 의한 것일 수 있다. 이러한 유도성(inducibility)은 원하는 순간에 RNA 조절의 활성화를 가능하게 한다. 일부 구현예에서, 상기 RNA 조절 도메인은 제1 이량체화 도메인에 공유 결합되고 상기 RNA 헤어핀 결합 도메인은 제2 이량체화 도메인에 공유 결합되며 상기 제1 및 제2 이량체화 도메인은 이량체화하여 비공유 또는 공유 연결을 형성할 수 있다. 일부 구현예에서, 이량체화는 유도성이다. 일부 측면에서, 이량체화는 리간드에 이량체화 도메인의 결합을 통해 유도된다. 유도성(inducible)이라는 말은 예를 들어 리간드, 화학물질, 온도 변화, 또는 빛과 같은 자극에 반응하여 형성되는 이량체화를 지칭한다.

[0025] 광 유도성(light inducibility)은 예를 들어 제1 및 제2 이량체화 도메인이 CRY2PHR 및 CIBN을 포함하는 융합 복합체를 설계함으로써 달성된다. 이 시스템은 살아있는 세포에서 단백질 상호작용의 광 유도에 특히 유용하며 본원에 참조로 포함된 문헌[Konermann S, et al. Nature. 2013;500:472-476]에 자세히 설명되어 있다.

[0026] 적합한 이량체화 도메인 및 상응하는 리간드는 당업계에 공지되어 있다. 예를 들어, 문헌[Liang, F.S., Ho, W.Q., 및 Crabtree, G.R. (2011)]을 참조한다. 본원에 참조로 포함된 문헌[Engineering the ABA plant stress pathway for regulation of induced proximity. Sci. Signal. 4, rs2]은 본 개시 내용의 구현예에서 유용한 적합한 이량체화/리간드 시스템을 설명한다. 일부 구현예에서, 상기 제1 또는 제2 이량체화 도메인 중 하나는 PYR/PYR1-유사(PYL1)를 포함하고, 상기 제1 또는 제2 이량체화 도메인 중 다른 하나는 ABA 비민감성(insensitive) 1(ABI1)을 포함하고, 상기 리간드는 아브시스산(ABA) 또는 이의 유도체 또는 단편을 포함한다. 이량체화 도메인은 전체 단백질의 단편 또는 부분일 수 있고 치환 또는 변형될 수 있다. 일부 구현예에서, 상기 제1 및/또는 제2 이량체화 도메인은 FKBP12를 포함하고 상기 리간드는 FK1012 또는 이의 유도체 또는 단편을 포함한다. 일부 구현예에서, 상기 제1 또는 제2 이량체화 도메인 중 하나는 FK506 결합 단백질(FKBP)을 포함하고, 상기 제1 또는 제2 도메인 중 다른 하나는 포유동물 Rap mTOR의 표적의 FKBP-Rap 결합 도메인(Frb)을 포함하고, 상기 리간드는 라파마이신(Rap) 또는 이의 유도체 또는 단편을 포함한다.

[0027] 유도체(derivative)는 각각 그들의 이량체화 도메인 또는 리간드에 대한 결합을 유지하거나 강화된 결합을 갖는 변형된 리간드 및 도메인을 지칭한다. 단편(fragment)은 리간드에 대한 결합을 유지하는 이량체화 도메인의 인접 부분을 지칭한다. 일부 구현예에서, 이량체화 도메인은 변형된 단편일 수 있다.

[0028] 일부 구현예에서, i), ii), 및/또는 iii)은 인간의 것이거나 인간 유래인 것이다. 일부 구현예에서, 시스템, 접합체, 및/또는 융합 단백질은 비면역원성이다. 인간 단백질, 폴리펩타이드, 도메인 또는 핵산은 인간 유전체에서 유래하지만 비인간 시스템에서 재조합으로 생산될 수 있는, 단백질, 폴리펩타이드, 도메인 또는 핵산을 지칭한다. "인간 유래(human-derived)"라는 말은 인간 유전체에서 유래하지만 비인간 시스템에서 재조합으로 생산될 수 있는, 단백질, 폴리펩타이드, 도메인 또는 핵산의 변이체 또는 단편인 단백질, 폴리펩타이드, 도메인, 또는 핵산을 지칭한다. 일부 구현예에서, 융합 단백질, 접합체, 시스템, 또는 i), ii), 및/또는 iii)과 같은 이의 부분은 인간에서 발현되거나 인간에게 투여될 때 비면역원성 및/또는 비독성이다.

[0029] 일부 구현예에서, 본 개시 내용의 핵산 또는 폴리펩타이드는 합성되고, 비천연이고/이거나 자연에서 자연적으로 발생하지 않는다.

[0030] 일부 구현예에서, 상기 시스템은 안정화제 폴리펩타이드를 포함하되, 상기 안정화제 폴리펩타이드는 핵산에 비특이적으로 결합하는 양이온성 폴리펩타이드를 포함한다. 일부 구현예에서, 안정화제 폴리펩타이드는 인간 유래이다. 일부 구현예에서, 안정화제 폴리펩타이드는 RNA 조절 도메인 및/또는 RNA 헤어핀 결합 도메인에 작동 가능하게 연결된다. 일부 구현예에서, 안정화제 폴리펩타이드는 ORF5 또는 이의 단편을 포함한다. 일부 구현예에서, 안정화제 폴리펩타이드는 서열번호: 5, 이의 변이체, 또는 서열번호: 5에 대해 적어도 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 75, 76, 77, 78, 79, 80, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 또는 99%(또는 여기에 속하는 임의의 추론가능한 범위)의 동일성 또는 상동성을 갖는 폴리펩타이드를 포함한다. 일부 구현예에서, 안정화제 폴리펩타이드는 HEBGF 또는 이의 단편을 포함한다. 일부 구현예에서, 안정화제 폴리펩타이드는 서열번호: 19, 이의 변이체, 또는 서열번호: 19에 대해 적어도 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 75, 76, 77, 78, 79, 80, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 또는 99%(또는 여기에 속하는 임의의 추론가능한 범위)의 동일성 또는 상동성을 갖는 폴리펩타이드를 포함한다. 일부 구현예에서, 안정화제 폴리펩타이드는 β-디펜신 3 또는 이의 단편을 포함한다. 일부 구현예에서, 안정화제 폴리펩타이드는 서열번호: 20, 이의 변이체, 또는 서열번호: 20에 대해 적어도 60, 61, 62, 63, 64,

65, 66, 67, 68, 69, 70, 75, 76, 77, 78, 79, 80, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 또는 99%(또는 여기에 속하는 임의의 추론가능한 범위)의 동일성 또는 상동성을 갖는 폴리펩타이드를 포함한다.

[0031] 일부 구현예에서, 안정화제 폴리펩타이드, 접합체, 융합 단백질, 접합체, RNA 조절 도메인 및/또는 RNA 헤어핀 결합 도메인은 175, 170, 165, 160, 155, 150, 145, 140, 135, 130, 125, 120, 115, 110, 105, 100, 95, 90, 85, 80, 75, 70, 65, 60, 55, 50, 45, 40, 35, 30, 25, 20, 15, 10, 또는 5kDa 미만이거나 초과이거나 또는 최대 또는 적어도 175, 170, 165, 160, 155, 150, 145, 140, 135, 130, 125, 120, 115, 110, 105, 100, 95, 90, 85, 80, 75, 70, 65, 60, 55, 50, 45, 40, 35, 30, 25, 20, 15, 10, 또는 5kDa(또는 여기에 속하는 임의의 추론가능한 범위)이다. 일부 구현예에서, RNA 조절 도메인 및 헤어핀 결합 도메인을 포함하는 총 복합체는 175, 170, 165, 160, 155, 150, 145, 140, 135, 130, 125, 120, 115, 110, 105, 100, 95, 90, 85, 80, 75, 70, 65, 60, 55, 50, 45, 40, 35, 30, 25, 20, 15, 10, 또는 5kDa 미만이거나 초과이거나 또는 최대 또는 적어도 175, 170, 165, 160, 155, 150, 145, 140, 135, 130, 125, 120, 115, 110, 105, 100, 95, 90, 85, 80, 75, 70, 65, 60, 55, 50, 45, 40, 35, 30, 25, 20, 15, 10, 또는 5kDa(또는 여기에 속하는 임의의 추론가능한 범위)이다. 일부 구현예에서, 안정화제 폴리펩타이드, RNA 조절 도메인 및 헤어핀 결합 도메인을 포함하는 총 복합체는 175, 170, 165, 160, 155, 150, 145, 140, 135, 130, 125, 120, 115, 110, 105, 100, 95, 90, 85, 80, 75, 70, 65, 60, 55, 50, 45, 40, 35, 30, 25, 20, 15, 10, 또는 5kDa 미만이거나 초과이거나 또는 최대 또는 적어도 175, 170, 165, 160, 155, 150, 145, 140, 135, 130, 125, 120, 115, 110, 105, 100, 95, 90, 85, 80, 75, 70, 65, 60, 55, 50, 45, 40, 35, 30, 25, 20, 15, 10, 또는 5kDa(또는 여기에 속하는 임의의 추론가능한 범위)이다.

[0032] 일부 구현예에서, RNA 헤어핀 결합 도메인은 U1A(TBP6.7), SLBP, 또는 이들의 변이체로부터 유래한 RNA 헤어핀 결합 도메인을 포함한다. 일부 구현예에서, RNA 헤어핀 결합 도메인은 U1A(TBP6.7), SLBP, Ku70, 뉴클레오린, 또는 이들의 변이체로부터 유래한 RNA 헤어핀 결합 도메인을 포함한다. 일부 구현예에서, RNA 헤어핀 결합 도메인은 서열번호: 7 또는 18, 이의 변이체, 또는 서열번호: 7 또는 18에 대해 적어도 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 75, 76, 77, 78, 79, 80, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 또는 99%(또는 여기에 속하는 임의의 추론가능한 범위)의 동일성 또는 상동성을 갖는 폴리펩타이드를 포함한다.

[0033] 일부 구현예에서, RNA 표적화 분자는 TAR 헤어핀 지지체를 포함한다. 일부 구현예에서, RNA 표적화 분자는 서열번호: 1 또는 서열번호: 1에 대해 적어도 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 75, 76, 77, 78, 79, 80, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 또는 99%(또는 여기에 속하는 임의의 추론가능한 범위)의 동일성을 갖는 뉴클레오타이드의 TAR 헤어핀 지지체를 포함한다. 일부 구현예에서, RNA 표적화 분자는 SLBP 헤어핀 지지체를 포함한다. 일부 구현예에서, RNA 표적화 분자는 서열번호: 2 또는 서열번호: 2에 대해 적어도 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 75, 76, 77, 78, 79, 80, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 또는 99%(또는 여기에 속하는 임의의 추론가능한 범위)의 동일성을 갖는 뉴클레오타이드의 SLBP 헤어핀 지지체를 포함한다.

[0034] 일부 구현예에서, RNA 표적화 분자는 정확히 하나의 헤어핀을 포함한다. 일부 구현예에서, RNA 표적화 분자는 적어도 하나의 헤어핀을 포함한다. 일부 구현예에서, RNA 표적화 분자는 정확히 두 개의 헤어핀을 포함한다. 일부 구현예에서, RNA 표적화 분자는 적어도 두 개의 헤어핀을 포함한다. 일부 구현예에서, RNA 표적화 분자는 정확히 세 개의 헤어핀을 포함한다. 일부 구현예에서, RNA 표적화 분자는 적어도 세 개의 헤어핀을 포함한다. 일부 구현예에서, RNA 표적화 분자는 정확히 네 개의 헤어핀을 포함한다. 일부 구현예에서, RNA 표적화 분자는 적어도 네 개의 헤어핀을 포함한다. 일부 구현예에서, RNA 표적화 분자는 정확히 다섯 개의 헤어핀을 포함한다. 일부 구현예에서, RNA 표적화 분자는 적어도 다섯 개의 헤어핀을 포함한다. 일부 구현예에서, RNA 표적화 분자는 정확히 1 내지 4개의 헤어핀을 포함한다. 일부 구현예에서, RNA 표적화 분자는 정확히 1 내지 3개의 헤어핀을 포함한다. 일부 구현예에서, RNA 표적화 분자는 정확히 1 내지 2개의 헤어핀을 포함한다. 일부 구현예에서, RNA 표적화 분자는 2 내지 4개의 헤어핀을 포함한다. 일부 구현예에서, RNA 표적화 분자는 2 내지 3개의 헤어핀을 포함한다. 일부 구현예에서, RNA 표적화 분자는 적어도, 최대, 또는 정확히 1, 2, 3, 4, 5, 또는 6개(또는 여기에 속하는 임의의 추론가능한 범위)의 헤어핀을 포함한다. 일부 구현예에서, RNA 표적화 분자는 RNA 헤어핀 결합 단백질에 결합하지 않는 적어도 하나의 헤어핀 및 RNA 헤어핀 결합 단백질에 결합하는 적어도 하나의 헤어핀을 포함한다. 일부 구현예에서, RNA 표적화 분자는 하나 초과의 RNA 결합 단백질에 결합한다. 일부 구현예에서, RNA 표적화 분자는 2개, 3개 또는 4개의 헤어핀 구조를 포함하고 적어도 2개의 RNA 결합 단백질에 결합한다. 일부 구현예에서, RNA 조절 시스템은 적어도 2개의 조절 도메인을 포함하고, 각각의 조절 도메

인은 상이한 RNA 결합 분자에 결합한다.

[0035] 일부 구현예에서, RNA 표적화 분자는 하나 이상의 변형된 뉴클레오타이드를 포함한다. 일부 구현예에서, 변형된 뉴클레오타이드는 포스포로티오에이트, 고정된 뉴클레오타이드, 에틸렌 가교된 뉴클레오타이드, 펩타이드 핵산, 5' E-VP와 같은 변형을 포함하거나 모르폴리노로 변형된다. 일부 구현예에서, 상기 변형은 본원에 기재된 것을 포함한다.

[0036] 일부 구현예에서, RNA 헤어핀 결합 도메인은 U1A, 이의 변이체 또는 서열번호: 7에 대해 적어도 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 75, 76, 77, 78, 79, 80, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 또는 99%(또는 여기에 속하는 임의의 추론가능한 범위)의 동일성 또는 상동성을 갖는 폴리펩타이드의 RNA 헤어핀 결합 도메인을 포함하고, RNA 표적화 분자는 TAR 헤어핀 지지체 또는 서열번호: 1에 대해 적어도 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 75, 76, 77, 78, 79, 80, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 또는 100%의 동일성을 갖는 뉴클레오타이드를 포함한다.

[0037] 일부 구현예에서, RNA 헤어핀 결합 도메인은 SLBP, 이의 변이체 또는 서열번호: 18에 대해 적어도 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 75, 76, 77, 78, 79, 80, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 또는 99%(또는 여기에 속하는 임의의 추론가능한 범위)의 동일성 또는 상동성을 갖는 폴리펩타이드의 RNA 헤어핀 결합 도메인을 포함하고, RNA 표적화 분자는 SLBP 헤어핀 지지체 또는 서열번호: 2에 대해 적어도 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 75, 76, 77, 78, 79, 80, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 또는 100%의 동일성을 갖는 뉴클레오타이드를 포함한다.

[0038] 일부 구현예에서, RNA 헤어핀 결합 도메인은 ku70 또는 이의 변이체의 RNA 헤어핀 결합 도메인을 포함하고, RNA 표적화 분자는 헤어핀 지지체 또는 서열번호: 83에 대해 적어도 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 75, 76, 77, 78, 79, 80, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 또는 100%(또는 여기에 속하는 임의의 추론가능한 범위)의 동일성을 갖는 뉴클레오타이드를 포함한다.

[0039] 일부 구현예에서, RNA 헤어핀 결합 도메인은 뉴클레오린 또는 이의 변이체의 RNA 헤어핀 결합 도메인을 포함하고, RNA 표적화 분자는 헤어핀 지지체 또는 서열번호: 84 내지 86 중 하나에 대해 적어도 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 75, 76, 77, 78, 79, 80, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 또는 100%(또는 여기에 속하는 임의의 추론가능한 범위)의 동일성을 갖는 뉴클레오타이드를 포함한다.

[0040] 일부 구현예에서, RNA 헤어핀 결합 도메인, 안정화제 폴리펩타이드, 또는 RNA 헤어핀 결합 도메인은 링커를 포함한다. 일부 구현예에서, 링커는 서열번호: 6, 21, 22, 23, 또는 25를 포함하는 폴리펩타이드 또는 서열번호: 6, 21, 22, 23, 또는 25에 대해 적어도 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 75, 76, 77, 78, 79, 80, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 또는 99%(또는 여기에 속하는 임의의 추론가능한 범위)의 동일성 또는 상동성을 갖는 폴리펩타이드를 포함한다. 일부 구현예에서, 링커는 경직성 링커(rigid linker)이다. 일부 구현예에서, 링커는 가요성 링커(flexible linker)이다. 일부 구현예에서, 링커는 글리신 및 세린 잔기를 포함한다. 일부 구현예에서, 링커는 적어도 4개의 아미노산이다. 일부 구현예에서, 링커는 적어도 또는 최대 또는 정확히 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43 44, 45, 46, 47, 48, 49, 또는 50개의 아미노산(또는 여기에 속하는 임의의 추론가능한 범위)이다.

[0041] 일부 구현예에서, 안정화제 폴리펩타이드는 cJun, HBEGF, HRX, NDEK, NHGF, 베타-디펜신 3, 또는 scGFP로부터 유래한 RNA-결합 폴리펩타이드와 같은 폴리펩타이드를 포함한다. 일부 구현예에서, RNA 조절 도메인은 RNA 조절 도메인의 카르복시 말단에서 안정화제 폴리펩타이드에 작동가능하게 연결된다. 일부 구현예에서, RNA 조절 도메인은 RNA 조절 도메인의 아미노 말단에서 안정화제 폴리펩타이드에 작동가능하게 연결된다. 일부 구현예에서, RNA 조절 도메인은 RNA 조절 도메인의 카르복시 말단에서 RNA 헤어핀 결합 도메인 폴리펩타이드에 작동가능하게 연결된다. 일부 구현예에서, RNA 조절 도메인은 RNA 조절 도메인의 아미노 말단에서 RNA 헤어핀 결합 도메인 폴리펩타이드에 작동가능하게 연결된다. 일부 구현예에서, RNA 헤어핀 결합 도메인은 RNA 헤어핀 결합 도메인 폴리펩타이드의 카르복시 말단에서 안정화제 폴리펩타이드에 작동가능하게 연결된다. 일부 구현예에서, RNA 헤어핀 결합 도메인은 RNA 헤어핀 결합 도메인 폴리펩타이드의 아미노 말단에서 안정화제 폴리펩타이드에 작동가능하게 연결된다.

[0042] 일부 구현예에서, RNA 표적화 영역은 적어도 12개의 뉴클레오타이드를 포함한다. 일부 구현예에서, RNA 표적화 영역은 적어도, 최대 또는 정확히 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25,

26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 또는 79개의 아미노산(또는 여기에 속하는 임의의 추론가능한 범위)을 포함한다.

[0043] 일부 구현예에서, RNA 조절 도메인은 뉴클레아제, 메틸라아제, 데메틸라아제(demethylase), 번역 활성제, 번역 억제인자, 단일 가닥 RNA 절단 활성, 이중 가닥 RNA 절단 활성, 또는 RNA 결합 활성을 포함한다. 일부 구현예에서, RNA 조절 도메인은 본원에 설명된 활성을 포함한다.

[0044] 일부 구현예에서, RNA 조절 도메인은 Pin 뉴클레아제 도메인 또는 m6A 판독자 단백질(reader protein) 또는 이의 부분을 포함한다. 일부 구현예에서, RNA 조절 도메인은 SMG6, YTHDF1, 또는 YTHDF2로부터 유래한 도메인 또는 폴리펩타이드를 포함한다. 일부 구현예에서, RNA 조절 도메인은 ADAR 단백질로부터 유래한 도메인 또는 폴리펩타이드를 포함한다. 일부 구현예에서, RNA 조절 도메인은 인간 ADAR 단백질로부터 유래한 도메인 또는 폴리펩타이드를 포함한다. 일부 구현예에서, RNA 조절 도메인은 서열번호: 9, 11, 15, 16, 17, 또는 123 내지 125에 대해 적어도, 최대, 또는 정확히 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 75, 76, 77, 78, 79, 80, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 또는 100%(또는 여기에 속하는 임의의 추론가능한 범위)의 동일성 또는 상동성을 갖는 폴리펩타이드를 포함한다. 일부 구현예에서, RNA 조절 도메인은 나선 영역을 더 포함한다. 일부 구현예에서, 나선 영역은 서열번호: 24에 대해 적어도, 최대, 또는 정확히 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 75, 76, 77, 78, 79, 80, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 또는 100%(또는 여기에 속하는 임의의 추론가능한 범위)의 동일성 또는 상동성을 갖는 폴리펩타이드를 포함한다.

[0045] 일부 구현예에서, RNA 조절 도메인은 표적 RNA의 번역을 증가시킨다. 일부 구현예에서, RNA 조절 도메인은 표적 RNA의 분해를 증가시킨다. 일부 구현예에서, RNA 조절 도메인은 표적 RNA의 국소화를 변형한다. 일부 구현예에서, RNA 조절 도메인은 표적 RNA의 처리를 변형한다.

[0046] 일부 구현예에서, RNA 조절 도메인은 IFIT2, eIF4a, eIF4e, PABP, PAIP, SLBP, BOLL, ICP27, YTHDF1, YTHDF2, 또는 YTHDF3로부터 유래한 RNA 조절 활성을 가진 폴리펩타이드와 같은 폴리펩타이드를 포함한다. 일부 구현예에서, RNA 조절 도메인은 YTHDF2, TOB2, ZFP36, CNOT7, RNaseA, RNaseL, RNaseP, RNase4, RNase1, RNaseU2, 또는 HRSP12로부터 유래한 RNA 조절 활성을 가진 폴리펩타이드와 같은 폴리펩타이드를 포함한다. 일부 구현예에서, RNA 조절 도메인은 표적 RNA에 의해 암호화되는 폴리펩타이드의 발현을 증가시키고 여기서 RNA 조절 도메인은 IFIT2, eIF4a, eIF4e, PABP, PAIP, SLBP, BOLL, ICP27, YTHDF1, 또는 YTHDF3을 포함한다. 일부 구현예에서, RNA 조절 도메인은 YTHDF2, TOB2, ZFP36, CNOT7, RNaseA, RNaseL, RNaseP, RNase4, RNase1, RNaseU2, 또는 HRSP12로부터 유래한 RNA 조절 활성을 가진 폴리펩타이드와 같은 폴리펩타이드를 포함한다. 일부 구현예에서, RNA 조절 도메인은 표적 RNA에 의해 암호화되는 폴리펩타이드의 발현을 감소시키고 여기서 RNA 조절 도메인은 YTHDF2, TOB2, ZFP36, CNOT7, RNaseA, RNaseL, RNaseP, RNase4, RNase1, RNaseU2, 또는 HRSP12를 포함한다.

[0047] 일부 구현예에서, 하나 이상의 핵 배출 신호(NES)가 RNA 조절 도메인, RNA 헤어핀 결합 도메인 및/또는 안정화제 폴리펩타이드에 융합된다. 일부 구현예에서, NES는 RNA 조절 도메인, RNA 헤어핀 결합 도메인 및/또는 안정화제 폴리펩타이드의 카르복시 말단에 있다. 일부 구현예에서, NES는 RNA 조절 도메인, RNA 헤어핀 결합 도메인 및/또는 안정화제 폴리펩타이드의 아미노 말단에 있다. 일부 구현예에서, NES는 서열번호: 8에 대해 적어도, 최대, 또는 정확히 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 75, 76, 77, 78, 79, 80, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 또는 100%(또는 여기에 속하는 임의의 추론가능한 범위)의 동일성 또는 상동성을 갖는 폴리펩타이드를 포함한다.

[0048] 일부 구현예에서, 하나 이상의 핵 국소화 신호(NLS)가 RNA 조절 도메인, RNA 헤어핀 결합 도메인 및/또는 안정화제 폴리펩타이드에 융합된다. 일부 구현예에서, NLS는 RNA 조절 도메인, RNA 헤어핀 결합 도메인 및/또는 안정화제 폴리펩타이드의 카르복시 말단에 있다. 일부 구현예에서, NLS는 RNA 조절 도메인, RNA 헤어핀 결합 도메인 및/또는 안정화제 폴리펩타이드의 아미노 말단에 있다. 일부 구현예에서, NES는 서열번호: 13에 대해 적어도, 최대, 또는 정확히 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 75, 76, 77, 78, 79, 80, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 또는 100%(또는 여기에 속하는 임의의 추론가능한 범위)의 동일성 또는 상동성을 갖는 폴리펩타이드를 포함한다.

[0049] 일부 구현예에서, ii)의 RNA 표적화 영역은 원핵 또는 진핵 세포에서 표적 RNA에 혼성화된다. 일부 구현예에서, 표적 RNA는 인간 세포 내에 있다. 일부 구현예에서, 표적 RNA는 시험관 내 또는 생체 내에 있다.

[0050] 일부 구현예에서, 본 시스템은 각 i), ii), 및 iii) 중 적어도 두 개를 포함한다. 일부 구현예에서, i), ii), 및 iii)의 적어도 두 개는 동일한 세포 내에서 발현된다. 일부 구현예에서, 본 방법은 적어도 두 개의 표적 RNA를 조절하는 단계를 포함한다. 일부 구현예에서, 본 시스템은 적어도, 최대, 또는 정확히 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 또는 50개 또는 그 이상(또는 여기에 속하는 임의의 추론가능한 범위)의 i), ii), 및 iii)을 포함한다. 일부 구현예에서, 적어도, 최대, 또는 정확히 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 또는 50개 또는 그 이상(또는 여기에 속하는 임의의 추론가능한 범위)의 표적 RNA가 세포 내에서 조절된다.

[0051] 일부 구현예에서, RNA 조절 도메인은 RNA를 절단하거나, RNA 번역을 촉진하거나, RNA 번역을 억제하거나, 또는 RNA의 염기 서열을 변형한다.

[0052] 일부 구현예에서, 본 개시 내용의 벡터는 i), ii), 및/또는 iii)을 암호화하는 뉴클레오타이드에 작동 가능하게 연결된 조절 요소를 더 포함한다. 이전에 설명된 NLS 및 NES에 더하여 조절 요소는 또한 프로모터, 폴리아데닐화 신호, 인핸서 등을 포함한다. 다른 조절 요소는 당업계에 공지되어 있고 본원에 기술되어 있으며 본 개시 내용의 구현예에서 사용될 수 있다. 일부 구현예에서, 하나 이상의 핵산 벡터는 진핵 세포 내 발현을 위해 최적화된다. 일부 구현예에서, 세포 내 또는 벡터로부터 유래한 도메인, RNA, 또는 폴리펩타이드의 발현은 구성적(constitutive)이다. 일부 구현예에서, 세포 내 또는 벡터로부터 유래한 도메인, RNA, 또는 폴리펩타이드의 발현은 조건적(conditional)이다. 일부 구현예에서, i), ii), 및 iii)은 단일 벡터 상에 있다. 일부 구현예에서, i), ii), iii), 및 안정화제 폴리펩타이드는 단일 벡터 상에서 암호화된다. 일부 구현예에서, i), iii), 및 안정화제 폴리펩타이드는 단일 벡터 상에서 암호화된다. 일부 구현예에서, 벡터들 중 하나 이상은 바이러스 벡터이다. 일부 구현예에서, 하나 이상의 벡터는 하나 이상의 레트로바이러스 벡터, 렌티바이러스 벡터, 아데노바이러스 벡터, 아데노 연관 바이러스 벡터 또는 단순헤르페스바이러스 벡터를 포함한다. 일부 구현예에서, 벡터들 중 하나 이상은 비바이러스 벡터이다. 일부 구현예에서, 본 시스템 또는 조성물은 비바이러스성이며, 이는 바이러스 성분을 함유하고 있지 않다는 의미이다.

[0053] 일부 구현예에서, 다음 성분들: RNA 조절 도메인을 포함하는 폴리펩타이드, RNA 결합 도메인을 포함하는 폴리펩타이드, 안정화제를 포함하는 폴리펩타이드, RNA 조절 도메인을 인코딩하는 핵산, RNA 결합 도메인을 인코딩하는 핵산, 안정화제를 인코딩하는 핵산, RNA 표적화 영역 및 적어도 하나의 헤어핀 구조를 포함하는 RNA 표적화 분자를 암호화하는 핵산; 본 개시 내용의 접합체; 본 개시 내용의 벡터, 본 개시 내용의 융합 단백질, 재조합 숙주 세포, 발현 작제물, 조작된 바이러스 벡터, 또는 조작된 감쇠 바이러스 중 하나 이상을 포함하는 시스템 또는 키트가 있다. 특정 구현예에서, 본 개시 내용의 폴리펩타이드는 이종 프로모터의 제어 하에 있다. 일부 구현예에서 사용되는 임의의 단백질 또는 폴리펩타이드 기능이 그 단백질 또는 폴리펩타이드 기능을 암호화하는 코딩하는 핵산에 사용될 수 있다는 것이 구체적으로 고려된다. 또한, 임의의 및 모든 폴리펩타이드, 단백질, 핵산 분자가 세포 또는 바이러스(예: 파지)와 같은 다른 생명체 내에 함유될 수 있다.

[0054] 키트는 멸균된 비반응성 용기와 같은 적절한 용기 수단 내에 분리되어 있거나 함께 있는 하나 이상의 성분을 포함할 수 있다. 일부 구현예에서, 본 개시 내용의 폴리펩타이드를 암호화하는 하나 이상의 핵산 작제물을 함유하는 세포 또는 바이러스가 제공된다. "프로모터"라는 용어는 분자 생물학 분야의 사람들에게 일반적인 의미로 사용되며; 이는 일반적으로 중합효소가 전사를 개시하기 위해 결합할 수 있는 핵산 상의 부위를 지칭한다. 구체적인 구현예에서, 프로모터는 T7 RNA 중합효소에 의해 인식된다.

[0055] 본 개시 내용의 조성물, 벡터, 시스템, 방법, 및 단백질은 다양한 임상 및 연구-관련 응용에 유용하다. 본 개시 내용의 구현예는 암 또는 자가면역과 같은 질병 또는 병태의 치료에 유용할 수 있다. 일부 구현예에서, 본 방법 및 조성물은 질병 또는 병태의 급성 치료를 위한 것이다. 일부 구현예에서, 본 방법 및 조성물은 RNA의 일시적인 조절에 유용하다. 일부 구현예에서는 유전자 활성의 영구적 변형은 배제한다. 일부 구현예에서, 본 방법 및 조성물은 RNA의 급성 조절 및/또는 생체 내 시스템의 발현을 제어하는 능력으로 인해 더 안전하다.

[0056] 본 명세서에 사용될 때, 단수 형태("a" 또는 "an")는 하나 이상을 의미할 수 있다. 청구범위(들)에서 본 명세서에 사용된 바와 같이, "포함하는"이라는 단어와 함께 사용될 때, 단수 형태("a" 또는 "an") 표현은 하나 또는 하나보다 많은 것을 의미할 수 있다.

[0057] 본 명세서에 사용될 때, 용어 "또는" 및 "및/또는"은 서로 조합되거나 배제하는 다수의 구성요소(성분)를 설명하는 데 사용된다. 예를 들어, "x, y 및/또는 z"는 "x" 단독, "y" 단독, "z" 단독, "x, y 및 z", "(x 및 y) 또는 z", "x 또는 (y 및 z)", 또는 "x 또는 y 또는 z"를 지칭할 수 있다. x, y 또는 z는 구현예에서 구체적으로 제외될 수 있음이 구체적으로 고려된다.

[0058] 본 출원 전반에 걸쳐, 용어 "약"은 세포 생물학 분야에서 그 값을 결정하기 위해 사용되는 장치 또는 방법에 대한 오차의 표준 편차를 포함하는 값을 나타내기 위해 평범하고 일반적인 의미에 따라 사용된다.

[0059] "포함하는(including)", "함유하는(containing)" 또는 "에 의해 특성화되는(characterized by)"과 동의어인 "포함하는(comprising)"이라는 용어는 포괄적이거나 개방형이며 추가의 인용되지 않은 요소 또는 방법 단계를 배제하지 않는다. "구성되는 또는 이루어진(consisting of)"이라는 문구는 지정되지 않은 요소, 단계 또는 성분을 제외한다. "필수적으로 포함하여 구성되는(consisting essentially of)"이라는 문구는 기술된 주제의 범위를 지정된 물질 또는 단계 및 기본적이고 새로운 특성에 물질적으로 영향을 미치지 않는 것으로 제한한다. "포함하는(comprising)"이라는 용어의 맥락에서 설명된 구현예는 또한 "구성되는 또는 이루어진(consisting of)" 또는 "필수적으로 포함하여 구성되는(consisting essentially of)"이라는 용어의 맥락에서 구현될 수 있는 것으로 고려된다.

[0060] 본 발명의 일 구현예와 관련하여 논의된 임의의 제한이 본 발명의 임의의 다른 구현예에 적용될 수 있다는 것이 구체적으로 고려된다. 또한, 본 발명의 임의의 조성물은 본 발명의 임의의 방법에 사용될 수 있고, 본 발명의 임의의 방법은 본 발명의 임의의 조성물을 생산하거나 사용하기 위해 사용될 수 있다. 실시예들에 제시된 구현예의 측면들은 또한 다른 실시예의 다른 곳 또는 "발명의 내용", "발명을 실시하기 위한 구체적인 내용", "청구범위", 및 "도면의 간단한 설명"과 같은 출원의 다른 곳에서 논의 되는 구현예의 맥락에서 구현될 수 있는 구현예이다.

[0061] 본 발명의 다른 목적, 특징 및 장점은 다음 상세한 설명으로부터 명확해질 것이다. 그러나, 본 발명의 사상 및 범위 내의 다양한 변경 및 수정이 이 상세한 설명으로부터 당업자에게 명백해질 것이므로, 상세한 설명 및 특정 실시예는 본 발명의 바람직한 구현예를 나타내면서 단지 예시로서 제공되는 것임을 이해해야 한다.

도면의 간단한 설명

[0062] 다음 도면들은 본 명세서의 일부를 형성하며 본 발명의 특정 측면들을 추가로 예시하기 위해 포함된다. 본 발명은 본원에 제시된 구체적인 구현예들의 상세한 설명과 함께 도면의 하나 이상을 참조하여 더 잘 이해될 수 있다.

도 1a-d. CRISPR/Cas-영감 RNA 표적화 시스템(CIRTS)의 설계. (A) 설계 전략의 개략도. CIRTS는 ssRNA 결합 단백질, RNA 헤어핀 결합 단백질, 효과기 단백질 및 가이드 RNA로 구성된다. (B) 본 연구에 사용된 모듈형 CIRTS 작제물의 목록. (C) TBP6.7에 대한 가이드 RNA의 설계. HIV TAR 헤어핀을 뉴클레오타이드 링커(L) 및 가이드 서열에 융합시켰다. (D) SLBP의 RRM에 대한 가이드 RNA의 설계. 인간 히스톤 mRNA 헤어핀을 가요성 5-뉴클레오타이드 링커 및 가이드 서열에 융합시켰다.

도 2a-c. CIRTS-1 시험관 내 결합 및 RNA 절단 분석. (A) 표지된 RNA 기질(R1)에 대한 MBP-CIRTS-1:표적내 gRNA(R3) 복합체의 결합 친화도를 평가하는 전기영동 이동성 분석(EMSA). EDTA를 반응 완충제에 보충하여 임의의 절단을 피하였다. (B) 2차 결합 방정식에 기질에 결합된 TBP6.7:gRNA의 분획을 피팅하여 결합 친화도의 계산. (C) 2시간의 인큐베이션 후에 변성 겔 상에서 수행된 절단 분석. 표지된 RNA 기질(R2)은 gRNA-의존적 방식으로 절단된다.

도 3a-g. CIRTS 포유류 세포 리포터 분석. (A) 이중 루시퍼라아제 분석의 일반적인 개요. 반딧불이 루시퍼라아제 및 레닐라 루시퍼라아제 둘 다를 함유하는 리포터 작제물을 모든 분석에 사용하였다. 본 발명자들은 형질주입 대조군으로서 사용하기 위해 레닐라 루시퍼라아제를 일정하게 유지하면서, 반딧불이 루시퍼라아제 전사물에 CIRTS를 표적화하였다. 모든 후속 분석을 위해, HEK293T 세포를 리포터 벡터, CIRTS 벡터, 및 gRNA 벡터로 형질주입시켰다. (B) 촉매 불활성 CIRTS-0을 대조군으로 사용하였다. 48시간의 인큐베이션 후, 본 발명자들은 단백질 판독치의 감소가 없음을 관찰하였다. (C) CIRTS-1과 Cas13b 뉴클레아제의 비교. CIRTS-1 또는 Cas13으로 형질주입된 세포 및 상응하는 gRNA 표적화는 인큐베이션 후 감소된 단백질 수준을 나타낸다. (D) CIRTS-2로 형질주입된 HEK293T 세포는 48시간 후 단백질 수준의 증가를 나타내지만, (E) CIRTS-3으로 형질주입된 세포는 단백질 수준의 예상된 감소를 나타낸다. (F) 헤어핀-결합 단백질을 SLBP로 전환시키는 것은 48시간 후에도 여전히 단백질 수준 감소를 초래한다. (G) 반딧불이 루시퍼라아제에 대한 완전 인간화 CIRTS(CIRTS-5 및 CIRTS-6) 시스템 및 표적내 gRNA로 형질주입된 세포는 다시 감소된 단백질 수준을 초래한다.

도 4a-c. CIRTS를 이용한 내인성 전사물의 표적화. (A) qPCR을 사용하여 분석된 바와 같이 CIRTS-1로 세포의 형질주입 시 5개의 내인성 전사물의 뉴클레아제-매개 녹다운. CIRTS-1은 gRNA를 상응하는 표적내 가이드 서열로 동시-형질주입시킴으로써 관심 내인성 전사물을 표적화하는데 사용될 수 있다. (B) CIRTS-3을 사용한 내인성

전사물의 비뉴클레아제 매개 녹다운의 qPCR 분석. CIRTS-3으로 형질주입된 세포는 시험된 5개의 전사물 모두에 대해 RNA 수준에서 gRNA-의존적 감소를 나타낸다. (C) 웨스턴 블롯을 이용한 CIRTS-2 또는 CIRTS-3에 의한 형질주입 후 단백질 수준의 분석. CIRTS-2는 단백질 수준의 증가를 유도할 수 있는 반면, CIRTS-3은 대조군으로서 단백질 수준의 예상된 감소를 나타낸다.

도 5a-b. CIRTS를 이용한 다차원 표적화. (A) 다중화 표적화에 사용되는 벡터의 개략도. CIRTS-6은 PPIB를 표적화하는 이의 gRNA 작제물로 동시-형질주입된 반면, CIRTS-7은 SMARCA4를 표적화하는 상응하는 gRNA 작제물로 동시-형질주입되었다. (B) 다중화 표적화의 녹다운을 보여주는 히트 맵. CIRTS-6 및 CIRTS-7 둘 모두 세포 내에 존재하는 경우, 표적내 또는 표적외 대조군 gRNA의 동시-형질주입은 CIRTS가 내인성 전사물을 감소시키도록 유도할 수 있다. 두 CIRTS가 PPIB 또는 SMARCA4에 대한 표적내 gRNA가 존재하는 경우, 두 전사물은 동일한 샘플에서 녹다운될 수 있다.

도 6. CIRTS를 다른 DNA 및 RNA-표적화 CRISPR/Cas 시스템과 비교하는 것. 현재 사용되는 Cas9, Cas13, 및 융합 단백질 시스템의 개략적 크기 비교. Cas9 및 Cas13-기반 전달 시스템은 조작된 CIRTS 시스템보다 실질적으로 더 크다.

도 7a-c. 대조군 EMSA 및 절단 분석. (A) 임의의 gRNA의 부재 하에 MBP-CIRTS-1에 의존적인 결합 이동을 평가하는 EMSA. (B) 단지 표지만된 기질(R1) 및 표적내 gRNA(R3)의 존재 하에 결합 이동을 분석하는 EMSA. (C) 도 2c의 본문에 나타낸 완전한 절단 겔.

도 8a-d. CIRTS 링커 및 gRNA 최적화. (A) 헤어핀-결합 단백질과 효과기 단백질 사이의 상이한 링커를 사용하는 CIRTS 뉴클레아제 시스템을 이용한 루시퍼라아제 분석. (B) 헤어핀-결합 단백질과 효과기 단백질 사이의 상이한 링커를 사용하는 CIRTS-YTHDF2-매개 붕괴를 이용한 루시퍼라아제 분석. (C) 도 1c에 도시된 설계에 기초한 TBP6.7에 대한 상이한 조작된 gRNA. 20개 및 40개 뉴클레오타이드의 2개의 상이한 표적화 길이가 헤어핀과 가이드 서열 사이의 상이한 수의 연결 뉴클레오타이드(L)와 조합되어 사용되었다. 뉴클레아제-매개된 붕괴를 평가하기 위해 이중 루시퍼라아제 분석을 사용하였다. (D) 도 8c에서와 동일한 조작된 gRNA를 CIRTS-3과 함께 사용하여 후성전사체-유도된 RNA 붕괴를 유도하였다. n = 3개 생물학적 복제물

도 9a-h. 대조군 루시퍼라아제 분석 및 RT-qPCR. (A) '데드(dead)' Pin 뉴클레아제 도메인 CIRTS(CIRTS-0)를 이용한 RNA 수준의 RT-qPCR 분석. (B) (CIRTS-8) 없는 TBP6.7-Pin 뉴클레아제 도메인의 뉴클레아제-매개된 붕괴를 (CIRTS-9) 있는 추가의 ssRNA 결합 단백질 ORF5와 비교하는 루시퍼라아제 분석. (C) CIRTS 뉴클레아제는 RNA의 감소를 매개할 수 있고, 이에 따라 핵 및 세포질 둘 다에서 단백질 수준 감소를 매개할 수 있다(n=6). (D) 세포를 CIRTS-1 및 활성 Cas13b 뉴클레아제로 형질주입시켰을 때의 RNA 수준의 비교. CIRTS-1-Pin 매개 RNA 절단은 Cas13b 시스템에 비해 실질적으로 더 적은 RNA 분해를 나타냈다. (E-H) 발명자들이 이중 루시퍼라아제 분석에서 시험한, 모든 조작된 CIRTS 시스템은 또한 RNA 수준의 변화를 평가하기 위해 RT-qPCR 분석을 거쳤다. 번역 활성화를 유도하는 YTHDF1 효과기 도메인을 함유하는 CIRTS-2는 RNA 수준에서 유의미한 변화를 나타내지 않은 반면, 모든 YTHDF2-함유 CIRTS는 RNA 수준의 예상된 감소를 나타낸다(S3F 및 S3H: n = 2 또는 3). 달리 나타내지 않는 한, n = 3개 생물학적 복제물. 스튜던트 t-검정: *P < 0.05, **P <0.01, ***P <0.001.

도 10a-c. 면역침강, 대조군 qPCR 웨스턴 블롯, Y2 절단. (A) 세포를 CIRTS-0-3xFLAG 및 PPIB, B4GALTN1 또는 NT에 대한 gRNA로 형질주입시켰다. 가교 및 FLAG IP 후, 풀다운된(pulled down) RNA를 RT-qPCR을 사용하여 정량화하였다. 어느 한 전사물에 대한 표적내 gRNA를 함유하는 반응은 이들 전사물에 대해 3.5 내지 5배 농축을 나타내었고, 이는 가이드된 RNA 표적화를 나타낸다(n = 2 또는 3). (B) 세포가 CIRTS-2로 형질주입되었을 때 RNA 수준의 RT-qPCR 분석. 예상한 바와 같이, YTHDF1-함유 단백질을 사용한 경우 RNA 수준의 유의한 변화가 관찰되지 않았다. (C) YTHDF2의 상이한 절단들을 분석하여 어느 것이 더 효율적인지를 결정하였다. 본 발명자들은 루시퍼라아제 데이터(좌측)를 qPCR 데이터(우측)와 비교하였고, 루시퍼라아제 분석을 위해 Y2(100-200) 작제물을 사용하고 내인성 표적화를 위해 Y2(1-200) 작제물을 사용하여 도구의 최상의 가능한 정량화를 가능하게 하는 것으로 결론지었다. 달리 나타내지 않는 한, n = 3개 생물학적 복제물. 스튜던트 t-검정: *P < 0.05, **P <0.01.

도 11a-c. CIRTS를 이용한 내인성 표적화. (A) PPIB에 대한 CIRTS5-7만의 형질주입 후에 평가한 RNA 수준의 변화. (B) 도 11a와 유사하게, 세포를 CIRTS5-7로 형질주입한 경우에 SMARCA4 수준을 분석하였다. (C) gRNA는 CIRTS-3을 사용하여 SMARCA4를 따라 스크리닝하여 gRNA-의존성 RNA 붕괴를 유도한다. 본 발명자들은 전사물이 표적화되는 곳(n = 2 또는 3)에 따라 유도 붕괴량의 유의한 변화를 관찰하였다. 달리 나타내지 않는 한, n = 3개 생물학적 복제물. 스튜던트 t-검정: *P < 0.05, **P <0.01, ***P <0.001.

도 12a-d. CRISPR/Cas-영감 RNA 표적화 시스템(CIRTS)의 설계. (A) 설계 전략의 개략도. CIRTS는 ssRNA 결합 단백질, RNA 헤어핀 결합 단백질, 효과기 단백질 및 가이드 RNA로 구성된다. (B) 본 연구에 사용된 핵심 CIRTS의 목록. (C) TBP6.7에 대한 가이드 RNA의 설계. HIV TAR 헤어핀을 뉴클레오타이드 링커(L) 및 가이드 서열에 융합시켰다. 뉴클레오타이드 링커는 최적화 동안 (지원 정보에 기재된 바와 같이) 변경되었지만, 이후 모든 작업을 위해 L = UUAUU를 사용하였다. (D) SLBP의 RNA 인식 모티프(RRM)에 대한 가이드 RNA의 설계. 인간 히스톤 mRNA 헤어핀을 가요성 5-뉴클레오타이드 링커 및 가이드 서열에 융합시켰다.

도 13a-b. CIRTS-1 *시험관 내* 결합 및 RNA 절단 분석. (A) 표지된 RNA 기질에 대한 표적내 gRNA 및 비-표적화 RNA 복합체와 MBP-CIRTS-1의 결합 친화도를 평가하는 필터 결합 분석. 2차 결합 방정식에 데이터를 맞추면 표적내:단백질 복합체에 대해 22 ± 7nM의 겉보기 K_D 및 표적외:단백질 상호작용에 대해 약 500nM의 겉보기 K_D를 나타냈다. (B) 0.5mM $MnCl_2$의 존재 하에 10% 변성 우레아 PAGE 겔 상에서 절단 분석을 수행하였다. IR800-표지된 RNA 기질은 gRNA-의존적 방식으로 절단된다.

도 14a-g. CIRTS 포유류 세포 리포터 분석. (A) 이중 루시퍼라아제 분석의 일반적인 개요. 반딧불이 루시퍼라아제 및 레닐라 루시퍼라아제 둘 다를 함유하는 리포터 작제물을 모든 분석에 사용하였다. 본 발명자들은 레닐라 루시퍼라아제를 내부 대조군으로 사용하면서, 반딧불이 루시퍼라아제 전사물에 CIRTS를 표적화하였다. 모든 후속 분석을 위해, HEK293T 세포를 리포터 벡터, CIRTS 벡터, 및 gRNA 벡터로 형질주입시켰다. (B) 촉매 불활성 CIRTS-0을 대조군으로 사용하였다. 48시간의 인큐베이션 후, 본 발명자들은 단백질 판독치의 감소가 없음을 관찰하였다. 값은 n = 3개 생물학적 복제물인 상태에서 평균 ± SEM으로 나타냈다. (C) CIRTS-1과 Cas13b 뉴클레아제의 비교. CIRTS-1 또는 Cas13으로 형질주입된 세포 및 상응하는 gRNA 표적화 Fluc는 인큐베이션 후 감소된 단백질 수준을 나타낸다. 값은 n = 3개 생물학적 복제물인 상태에서 평균 ± SEM으로 나타냈다. 스튜던트 t-검정: *P < 0.05, **P < 0.01. (D) CIRTS-2로 형질주입된 HEK293T 세포는 48시간 후에 단백질 수준의 증가를 나타낸다. 값은 n = 3개 생물학적 복제물인 상태에서 평균 ± SEM으로 나타냈다. 스튜던트 t-검정: **P < 0.01. (E) CIRTS-3으로 형질주입된 HEK293T 세포는 단백질 수준의 예상된 감소를 나타낸다. 값은 n = 3개 생물학적 복제물인 상태에서 평균 ± SEM으로 나타냈다. 스튜던트 t-검정: *P < 0.05. (F) 헤어핀-결합 단백질을 SLBP로 전환하면 48시간 후에도 여전히 단백질 수준의 감소를 초래한다. 값은 n = 3개 생물학적 복제물인 상태에서 평균 ± SEM으로 나타냈다. 스튜던트 t-검정: *P < 0.05. (G) 반딧불이 루시퍼라아제에 대한 완전 인간화 CIRTS(CIRTS-5 및 CIRTS-6) 시스템 및 표적내 gRNA로 형질주입된 세포는 다시 감소된 단백질 수준을 초래한다. 값은 n = 3개 생물학적 복제물인 상태에서 평균 ± SEM으로 나타냈다. 스튜던트 t-검정: *P < 0.05.

도 15a-b. RNA 편집을 위한 CIRTS. (A) 사용된 RNA 편집 리포터 분석의 개략도. 단일 G-to-A 돌연변이를 반딧불이 루시퍼라아제의 코딩 서열에 도입하여 W417X(X = STOP) 코돈 스위치를 낳았고 측정가능한 반딧불이 루시퍼라아제 신호를 나타내지 않았다(도 22j). (B) 표적내 gRNA를 이용한 CIRTS-7(hADAR2 wt) 및 CIRTS-8(hADAR E488Q)의 전달은 측정가능한 반딧불이 루시퍼라아제 신호를 초래하는 유의한 RNA 편집을 나타낸다. 과활성 hADAR2 돌연변이체인 CIRTS-8의 백그라운드 및 편집 효율 둘 다는 야생형에 비해 더 높은 것으로 밝혀졌다. 값은 n = 3개 생물학적 복제물인 상태에서 평균 ± SEM으로 나타냈다. 스튜던트 t-검정: ***P < 0.001.

도 16a-c. CIRTS를 이용한 내인성 전사물의 표적화. (A) qPCR을 사용하여 분석된 바와 같이 CIRTS-1로 세포의 형질주입 시 5개의 내인성 전사물의 뉴클레아제-매개 녹다운. CIRTS-1은 gRNA를 상응하는 표적내 가이드 서열로 동시-형질주입시킴으로써 관심 내인성 전사물을 표적화하는데 사용될 수 있다. 값은 n = 3개 생물학적 복제물인 상태에서 평균 ± SEM으로 나타냈다. 스튜던트 t-검정: *P < 0.05, **P <0.01, ***P <0.001. (B) CIRTS-3을 사용한 내인성 전사물의 YTHDF2 매개 녹다운의 qPCR 분석. CIRTS-3으로 형질주입된 세포는 시험된 5개의 전사물 모두에 대해 RNA 수준에서 gRNA-의존적 감소를 나타낸다. 값은 n = 3개 생물학적 복제물인 상태에서 평균 ± SEM으로 나타냈다. 스튜던트 t-검정: *P < 0.05. (C) 웨스턴 블롯에 의한 CIRTS-2 또는 CIRTS-3을 이용한 형질주입 후 단백질 수준의 분석. CIRTS-2는 단백질 수준의 증가를 유도하는 반면, CIRTS-3은 모두 gRNA-의존적 방식으로 단백질 수준의 예상된 감소를 나타낸다.

도 17. 표적화 접근성은 CIRTS의 녹다운 효율을 결정한다. gRNA는 CIRTS-3을 사용하여 SMARCA4를 따라 스크리닝하여 gRNA-의존성 RNA 붕괴를 유도한다. 본 발명자들은 전사물이 표적화되는 곳에 따라 유도 붕괴량의 유의한 변화를 관찰하였다(n = 2 또는 3).

도 18a-d. CIRTS를 이용한 다차원 표적화. CIRTS-6 및 3개의 gRNA의 전달의 개략도. (B) CIRTS-6은 PPIB, SMARCA4, 및 NRAS에 대한 3개의 별개의 gRNA와 함께 전달될 수 있고 동시에 세 모든 전사물의 녹다운을 야기할 수 있다. n = 5개 생물학적 복제물. 스튜던트 t-검정: **P <0.05, ***P <0.001. (C) 상이한 효과기 단백질들과

의 동시 CIRTS 전달의 개략도. (D) 세포를 Fluc 및 PPIB에 대해 각각 CIRTS-9(YTHDF1) 및 CIRTS-10(YTHDF2)과 gRNA들로 형질주입시켰을 때 루시퍼라아제 단백질 수준 및 PPIB 전사물 수준의 변화. 직교 CIRTS 둘 모두는 그들의 개별 기능을 유지하고 세포에서 동시에 작용한다. n = 5개 생물학적 복제물. 스튜던트 t-검정: *P < 0.1, **P <0.05.

도 19a-c. CIRTS의 AAV 전달. (A) CIRTS-6(YTHDF2) 뿐만 아니라 시스템의 gRNA 성분 둘 다를 함유하는 AAV 전달을 위한 전달 플라스미드. 2개의 역 말단 반복(ITR) 사이의 전체 삽입 크기는 2.7kb였다. (B) AAV-패키징된 CIRTS-6 및 루시퍼라아제를 표적화하는 gRNA를 HEK293T 세포로 전달하여 이중 루시퍼라아제 리포터 분석에서 반딧불이 루시퍼라아제를 녹다운시켰다. (C) AAV-패키징된 CIRTS-6 및 SMARCA4를 표적화하는 gRNA를 HEK293T 세포로 전달하여 내인성 유전자를 녹다운시켰고, 이는 일시적 형질주입에 의해 달성된 것에 필적하는 효율을 나타냈다. 값은 n = 3개 생물학적 복제물인 상태에서 평균 ± SEM으로 나타냈다. 스튜던트 t-검정: *P < 0.05.

도 20. CIRTS를 다른 DNA 및 RNA-표적화 CRISPR/Cas 시스템과 비교하는 것. 일반적으로 사용되는 Cas9, Cas12, Cas13, 및 융합 단백질 시스템의 개략적 크기 비교.

도 21. 도 12b로부터 계속된 CIRTS 목록. 본 연구에 사용된 모든 나머지 CIRTS의 참조 목록.

도 22a-j. 대조군 루시퍼라아제 분석 및 RT-qPCR. (A) (CIRTS-11) 없는 TBP6.7-Pin 뉴클레아제 도메인의 뉴클레아제-매개된 붕괴를 (CIRTS-12) 있는 추가의 ssRNA 결합 단백질 ORF5와 비교하는 루시퍼라아제 분석. (B) CIRTS 뉴클레아제는 RNA의 감소를 매개할 수 있고, 이에 따라 핵 및 세포질 둘 다에서 단백질 수준 감소를 매개할 수 있다(n=6). (C) '데드(dead)' Pin 뉴클레아제 도메인 CIRTS(CIRTS-0)를 이용한 RNA 수준의 RT-qPCR 분석. (D) 세포를 CIRTS-1 및 활성 Cas13b 뉴클레아제로 형질주입시켰을 때의 RNA 수준의 비교. CIRTS-1-Pin 매개 RNA 절단은 Cas13b 시스템에 비해 실질적으로 더 적은 RNA 분해를 나타냈다. (E-H) 이중 루시퍼라아제 분석에서 시험한, 모든 조작된 CIRTS 시스템은 또한 RNA 수준의 변화를 평가하기 위해 RT-qPCR 분석을 거쳤다. 번역 활성화를 유도하는 YTHDF1 효과기 도메인을 함유하는 CIRTS-2는 RNA 수준에서 유의미한 변화를 나타내지 않은 반면, 모든 YTHDF2-함유 CIRTS는 RNA 수준의 예상된 감소를 나타낸다. (I) 헤어핀 결합 단백질로서 PP7 이량체를 함유하는 조작된 CIRTS-18. qPCR에 의해 측정된 바와 같이 CIRTS-18로 형질주입 후 PPIB의 녹다운. (J) 리포터 단독 및 리포터와 비-표적화 또는 표적화 gRNA를 갖는 CIRTS-7(hADAR wt)의 비교(S3F 및 S3H: n = 2 또는 3). 달리 나타내지 않는 한, n = 3개 생물학적 복제물. 스튜던트 t-검정: *P < 0.05, **P <0.01, ***P <0.001.

도 23a-d. CIRTS 링커 및 gRNA 최적화. (A) 헤어핀-결합 단백질과 효과기 단백질 사이의 상이한 링커를 사용하는 CIRTS 뉴클레아제 시스템을 이용한 루시퍼라아제 분석. 이전에 공개된 L8 = SGSETPGTSESATPES(서열 번호: 133)(문헌 [Guilinger et al., 2014]), 10nm 나선형 링커 = EEEEKKKQQEEEAERLRRIQEEMEKERKRREEDEKRRRKEEEERRMKLEMEAKRKQEEEERKKREDDEKRKKK(서열 번호: 134). (B) 헤어핀-결합 단백질과 효과기 단백질 사이의 상이한 링커를 사용하는 CIRTS-YTHDF2-매개 붕괴를 이용한 루시퍼라아제 분석. (C) 도 1c에 도시된 설계에 기초한 TBP6.7에 대한 상이한 조작된 gRNA. 20개 및 40개 뉴클레오타이드의 2개의 상이한 표적화 길이가 헤어핀과 가이드 서열 사이의 상이한 수의 연결 뉴클레오타이드(L)와 조합되어 사용되었다. 뉴클레아제-매개된 붕괴를 평가하기 위해 이중 루시퍼라아제 분석을 사용하였다. NT = 상이한 링커 뉴클레아제를 함유하는 비표적화, Fluc gRNA(도 1), L2 = UU, L3 = UUU, L5 = UUAUU. (D) 도 S3c에서와 동일한 조작된 gRNA를 CIRTS-3과 함께 사용하여 후성전사체-유도된 RNA 붕괴를 유도하였다. NT = 상이한 링커 뉴클레아제를 함유하는 비표적화, Fluc gRNA(도 1), L2 = UU, L3 = UUU, L5 = UUAUU. n = 3개 생물학적 복제물.

도 24a-d. 대조군 qPCR, 웨스턴 블롯, YTHDF2 절단. (A) CIRTS-1은 mRNA 이외의 RNA 종으로 전달될 수 있다. 원리의 증명으로서, 본 발명자들은 세포를 CIRTS-1(Pin 뉴클레아제) 및 lncRNA MALAT1에 대한 2개의 상이한 gRNA로 형질주입시키고, RNA 수준을 RT-qPCR에 의해 평가하였다. (B) 세포가 CIRTS-2로 형질주입되었을 때 RNA 수준의 RT-qPCR 분석. 예상한 바와 같이, YTHDF1-함유 단백질을 사용한 경우 RNA 수준의 유의한 변화가 관찰되지 않았다. (C) 세포를 CIRTS-2 또는 CIRTS-3으로 형질주입시키고 PPIB로 표적화했을 때 웨스턴 블롯에 의해 측정된 단백질 수준의 정량화(n=3). (D) YTHDF2의 상이한 절단들을 분석하여 어느 것이 더 효율적인지를 결정하였다. 본 발명자들은 루시퍼라아제 데이터(좌측)를 qPCR 데이터(우측)와 비교하였고, 루시퍼라아제 분석을 위해 Y2(100-200) 작제물을 사용하고 내인성 표적화를 위해 Y2(1-200) 작제물을 사용하여 도구의 최상의 가능한 정량화를 가능하게 하는 것으로 결론지었다. 달리 나타내지 않는 한, n = 3개 생물학적 복제물. 스튜던트 t-검정: *P < 0.05, **P <0.01.

도 25a-h. CIRTS의 표적화 특이성. (A) KRAS4b-루시퍼라아제 미스매치 리포터 분석의 개략도. 본 발명자들은 설

계된 20nt 길이 gRNA에 대해 증가하는 수의 미스매치를 갖는 4개의 KRAS4b 변이체를 선택하였고, 이를 이중 루시퍼라아제 리포터에 N-말단에 융합시켰다. (B) 도 S5a에 기재된 바와 같이 gRNA와 표적 RNA 사이에 상이한 수의 미스매치를 갖는 KRAS4b-Fluc의 CIRTS-매개된 녹다운. CIRTS는 그의 가이드 서열의 중간에서 미스매치에 가장 민감한 것으로 밝혀졌다. (C) 상기 기재된 바와 동일한 KRAS4b-Fluc 리포터 분석에서의 Cas13b-매개 녹다운. Cas13b는 더 높은 녹다운 효율을 나타내지만 또한 도입된 미스매치에 덜 민감하다. CIRTS와 유사하게, Cas13b는 가이드 표적 이중체 영역의 중심에서의 미스매치에 의해 가장 영향을 받는 녹다운이다. (D) 40nt gRNA 길이를 사용할 때 KRAS4b-루시퍼라아제 미스매치 리포터에 대한 CIRTS의 녹다운 효율. gRNA에서 더 긴 안내 서열은 녹다운 효율의 손실의 일부를 구조할 수 있다. (E-F) CIRTS Pin 뉴클레아제(E) 또는 CIRTS YTHDF2(F)가 세포에서 SMARCA4에 전개될 때 $\log_2$(백만 당 전사물(TPM) +1) 단위의 전사체의 평균 발현 수준 (n=3). (G) RNA 서열분석에 의해 결정된 SMARCA4의 녹다운 수준. (H) 세포를 CIRTS-0-3xFLAG 및 PPIB, B4GALNT1 또는 NT에 대한 gRNA로 형질주입시켰다. 가교 및 FLAG IP 후, 풀다운된(pulled down) RNA를 RT-qPCR을 사용하여 정량화하였다. 어느 한 전사물에 대한 표적내 gRNA를 함유하는 반응은 이들 전사물에 대해 3.5 내지 5배 농축을 나타내었고, 이는 가이드된 RNA 표적화를 나타낸다(n = 2 또는 3).

도 26a-c. CIRTS를 이용한 내인성 표적화.(A) PPIB에 대한 CIRTS5-7만의 형질주입 후에 RT-qPCR에 의해 평가한 RNA 수준의 변화. (B) 도 S6a와 유사하게, 세포를 CIRTS5-7로 형질주입한 경우에 SMARCA4 수준을 분석하였다. C) PPIB에 대한 활성 Cas13b 뉴클레아제 또는 조작된 dCas13b-YTHDF2(1-200) 작제물의 전달 후 PPIB의 녹다운 수준의 비교(n=2 또는 3). 달리 나타내지 않는 한, n = 3개 생물학적 복제물. 스튜던트 t-검정: *P < 0.01, **P <0.05, ***P <0.01, ****P < 0.001.

도 27a-c. CIRTS를 이용한 다중화 표적화. (A) 다중화 표적화에 사용되는 벡터의 개략도. 세포를 CIRTS-6에 대한 발현 벡터, 및 CIRTS-10에 대한 발현 벡터, PPIB를 표적화하는 CIRTS-6 gRNA 작제물 또는 비-표적화 대조군에 대한 발현 벡터, 및 SMARCA4를 표적화하는 CIRTS-9 gRNA 또는 비-표적화 대조군에 대한 발현 벡터로 형질주입시켰다. (B) (A)에 기재된 다중화 표적화의 녹다운을 보여주는 히트 맵. 두 CIRTS가 PPIB 또는 SMARCA4에 대한 표적내 gRNA가 존재하는 경우, 두 전사물은 동일한 샘플에서 녹다운될 수 있다. 값은 GAPDH에 대한 각각의 표적 전사물의 평균 발현 수준으로서 제시되며, n = 8개 생물학적 복제물이다. (C) 면역원성의 컴퓨터 예측. 본 발명자들은 먼저 IEDB 데이터베이스를 사용하여 MHC I 결합제인 9량체 펩타이드를 예측하고, IEDB 면역원성 예측인자를 사용하여 결합제의 상위 1백분위수를 면역원성 예측에 적용하였다.

도 28. ADAR을 이용한 gRNA 스크리닝. 다수의 헤어핀을 특징으로 하는 가이드 RNA 설계가 CIRTS의 효능을 증가시키는지에 대한 시험. 좌측 패널은 gRNA 설계가 기원 설계, 또는 가이드의 어느 한 단부 상에 하나의 TAR 헤어핀을 갖는, 또는 한 단부 상에 2개의 헤어핀을 갖는, 가이드를 특징으로 함을 보여준다. 추가의 헤어핀 가이드는 세포의 ADAR 활성 분석에서 CIRTS의 효능을 증가시킨다. 우측 패널은 제2 헤어핀이 TAR 헤어핀일 필요가 있는지, 또는 단지 "안정화 헤어핀"이 기능할 수 있는지 여부를 시험하며 - 이는 CIRTS 단백질과 직접 상호작용하지 않지만 분해를 지연시키는 헤어핀을 의미한다. 데이터에서 볼 수 있는 바와 같이, 제2 헤어핀은 1개의 헤어핀 gRNA 설계와 비교하여 효능을 증가시킨다.

도 29a-b. C-to-U 편집기: A의 데이터는 C-to-U 염기 편집기에 기초한 CIRTS가 또한 포유동물 세포 리포터 분석을 사용하여 기능적이라는 것을 입증한다. (B) ssRNA 결합 단백질. B의 데이터는 다른 ssRNA 결합 단백질이 본 개시 내용의 시스템 및 방법에서 RNA 결합 단백질로서 기능할 수 있음을 입증한다.

도 30a-c. (A): 번역을 활성화하고, 리포터 RNA 상의 다양한 위치에 gRNA들을 배열하고 각각의 번역 활성화를 측정할 수 있는, RNA 조절 도메인-함유 단백질. (B-C) RNA를 잠재적으로 분해하거나 불안정화하고, 리포터 RNA 상의 다양한 위치에 gRNA들을 배열하고 각각의 RNA 분해를 측정할 수 있는, RNA 조절 도메인-함유 단백질.

도 31a-b. (A) 본 개시 내용의 시스템의 요소의 상이한 배향을 입증하는 구현예들. (B) 상이한 배향으로(A에 나타낸 바와 같이) 그리고 2개의 RNA 표적:루시퍼라아제 리포터(좌측) 및 내인성 RNA(우측) 상에서 CNOT7을 사용한 데이터. 여러 상이한 배향의 단백질들이 여전히 기능하고, 이는 단백질들이 효과기의 필요에 따라 상이한 순서로 조작될 수 있음을 나타낸다.

도 32a-d. 유도성 RNA 표적화를 위한 CIRTS 바이오센서. (A) 아브시스산(ABA) CIRTS 바이오센서 설계의 개략도. CIRTS의 gRNA 매개 표적화 성분은 ABA 이중이량체화 도메인들(ABI) 중 하나에 융합되는 반면 CIRTS의 효과기 성분은 그의 결합 파트너(PYL)에 융합된다. 저분자 ABA의 첨가 시, 2개의 CIRTS 성분이 이량체화되고 효과기를 표적화된 전사물에 근접하게 한다. (B) 형질주입 48시간 후 적색-형광 단백질(RFP) 리포터 전사물의 ABA-유도성 RNA 분해는 Pin 뉴클레아제 도메인 또는 YTHDF2에 의해 매개될 수 있다. (C) 형질주입 48시간 후 ABA-유도된

CIRTS-YTHDF1에 의한 RFP의 번역 활성화. (D) 돌연변이-불활성화된 루시퍼라아제 리포터로 형질주입된 세포(A-to-I 편집의 경우 FlucW417X, 또는 C-to-U 편집의 경우 GlucC82R)로 ABA의 존재 하에 표적내 gRNA를 이용한 CIRTS-hADAR의 전달은 ABA-의존성 RNA 편집을 유도한다.

발명을 실시하기 위한 구체적인 내용

[0063] 후성전사체 조절은 센트럴 도그마를 통해 정보 흐름을 제어하고 RNA 수준에서 세포 상태를 조작하기 위한 독특한 기회를 제공한다. 그러나, 기초적인 기계론적 연구 및 잠재적 번역 적용 둘 모두 효과기 단백질을 사용하여 특이적 RNA를 표적화하는 효과적인 방법의 결여로 인해 방해를 받는다. 여기서, 본 발명자들은 프로그램 가능한 RNA 조절 시스템을 구축하기 위한 새로운 단백질 조작 전략인 CRISPR/Cas-영감 RNA 표적화 시스템(CIRTS)의 설계 및 검증을 제시한다. 본 발명자들은 CIRTS가 뉴클레아제, 분해 기구, 및 번역 활성제를 포함하는 다양한 효과기 단백질을 표적 전사물에 전달하기 위한 단순하고 일반화 가능한 접근 방식임을 보여준다. CIRTS는 자연 발생 CRISPR/Cas 프로그램 가능한 RNA 결합 시스템보다 작을 뿐만 아니라, 전적으로 인간 단백질 부분으로부터 구축될 수 있다. CIRTS의 작은 크기 및 인간-유래 성질은 기초적인 RNA 조절 연구를 위한 덜 교란적인 방법뿐만 아니라 후성전사체-조절 요법에 적용할 때 면역 문제를 피하기 위한 잠재적인 전략을 제공한다.

[0064] I.RNA 조절 도메인

[0065] 임의의 RNA 조절 도메인이 본 개시 내용의 방법 및 시스템에서 사용될 수 있는 것으로 고려된다. 예를 들어, 하기 활성들: 메틸화(methylation), 5’-3’ 구아닐화(5’-3’ guanylylation), 포스포리보실화(phosphoribosylation), 탈아미노화(deamination), 카르바모일화(carbamoylation), 이소펜테닐화(isopentenylation), 아그마티닐화(agmatinylation), 아세틸화(acetylation), 리실화(lysylation), O/S 교환, 갈락토실화(galactosylation), 글루타밀화(glutamylation), 만노실화(mannosylation), 수소화(hydrogenation), 유사우리딘(pseudouridine) 형성, 카르복시메틸아미노메틸화(carboxymethylaminomethylation), 아미노메틸화(aminomethylation), 탈카르복시메틸화(decarboxymethylation), 탈수소화(dehydrogenation), 카르복시메틸화(carboxymethylation), 히드록실화(hydroxylation), 메틸티올화(methylthiolation), 3-아미노-3-카르복시프로필화(3-amino-3-carboxypropylation), 탈메틸화(demethylation), 5’-5’ 구아닐화(5’-5’ guanylylation), 및 탈인산화(dephosphorylation) 중 하나 이상을 갖는 RNA 조절 도메인이 사용될 수 있다.

[0066] 예시적인 RNA 조절 도메인은 표 1의 하기 단백질들로부터의 도메인들 (또는 이의 기능적 단편들)을 포함한다:

[0067] 표 1

머리글자어(Acronym)	전체 이름	유기체	활성
BCDIN3D	BCDIN3 도메인 함유	호모 사피엔스	메틸화
Nop1	rRNA 2′-O-메틸전달효소 피브릴라린	사카로미세스 세레비시아	메틸화
Nsr	rRNA (아데노신-2′-O-)-메틸전달효소	스트렙토마이세스 악투오서스(Streptomyces actuosus)	메틸화
Trm13	tRNA-2′-O-메틸전달효소 TRM13	사카로미세스 세레비시아	메틸화
Tsr	rRNA (아데노신-2′-O-)-메틸전달효소	스트렙토마이세스 키아네우스	메틸화
KamB	rRNA 아데닌 N-1-메틸전달효소	스트렙토마이세스 sp. DSM 40477	메틸화
NpmA	rRNA 아데닌 N-1-메틸전달효소	대장균(Escherichia coli)	메틸화
Rrp8	리보솜 RNA-처리 단백질 8	사카로미세스 세레비시아	메틸화
TRMT10C	tRNA 메틸전달효소 10 C, 미토콘드리아성	호모 사피엔스	메틸화
Trm61	tRNA (아데닌(58)-N(1))-메틸전달효소 촉매 서브유닛 TRM61	사카로미세스 세레비시아	메틸화
TrmI	tRNA (아데닌(57)-N(1)/아데닌(58)-N(1))-메틸전달효소 TrmI	피로코쿠스 아비시(Pyrococcus abyssi)	메틸화
TrmI	tRNA (아데닌(58)-N(1))-메틸전달효소 TrmI	테르무스 테르모필루스(Thermus thermophilus)	메틸화
TrmI	tRNA (아데닌(58)-N(1))-메틸전달효소 TrmI	결핵균(Mycobacterium tuberculosis)	메틸화

[0068]

[0069]

TrmK	tRNA (아데닌(22)-N(1))-메틸전달효소	고초균(Bacillus subtilis)	메틸화
Trmt61A	tRNA (아데닌(58)-N(1))-메틸전달효소 촉매 서브유닛 TRMT61A	호모 사피엔스	메틸화
Trmt61B	tRNA (아데닌(58)-N(1))-메틸전달효소, 미토콘드리아성	호모 사피엔스	메틸화
METTL14		호모 사피엔스	메틸화
METTL3	N6-아데노신-메틸전달효소 70 kDa 서브유닛	호모 사피엔스	메틸화
RsmA	디메틸아데노신 전달효소	사카로미세스 세레비시아	메틸화
ErmAM	rRNA 아데닌 N-6-메틸전달효소	폐렴연쇄구균(Streptococcus pneumoniae)	메틸화
Cfr	리보솜 RNA 대형 서브유닛 메틸전달효소 Cfr	포도상구균(Staphylococcus sciuri)	메틸화
Nop1	rRNA 2′-O-메틸전달효소 피브릴라린	사카로미세스 세레비시아	메틸화
Dnmt2	tRNA (시토신(38)-C(5))-메틸전달효소	호모 사피엔스	메틸화
Trm4	tRNA (시토신(34)-C(5))-메틸전달효소	호모 사피엔스	메틸화
MRM1	rRNA 메틸전달효소 1, 미토콘드리아성	호모 사피엔스	메틸화
RNMTL1	RNA 메틸전달효소-유사 단백질 1	호모 사피엔스	메틸화
TRM5	tRNA (구아닌(37)-N1)-메틸전달효소	호모 사피엔스	메틸화
TRMT10A	tRNA 메틸전달효소 10 동족체 A	호모 사피엔스	메틸화
TRMT10B	tRNA 메틸전달효소 10 동족체 B	호모 사피엔스	메틸화
TRMT10C	tRNA 메틸전달효소 10 C, 미토콘드리아성	호모 사피엔스	메틸화

TRMB	tRNA (구아닌-N(7)-)-메틸전달효소	호모 사피엔스	메틸화
Hcm1	RNA 캡 구아닌-N7 메틸전달효소	호모 사피엔스	메틸화
MRM2	rRNA 메틸전달효소 2, 미토콘드리아성	호모 사피엔스	메틸화
ALKBH8	알킬화 DNA 복구 단백질 alkB 동족체 8	호모 사피엔스	메틸화
BCDIN3D	BCDIN3 도메인 함유	호모 사피엔스	메틸화
Tgs1	트리메틸구아노신합성효소 1	호모 사피엔스	메틸화
Trm1	tRNA (구아닌(26)-N(2))-메틸전달효소	호모 사피엔스	메틸화
Thg1	tRNA(His) 구아닐전달효소(guanylyltransferase)	사카로미세스 세레비시아	5′-3′ 구아닐화
Rit1	tRNA A64-2′-O-리보실인산 전달효소	사카로미세스 세레비시아	포스포리보실화
Tad1	tRNA-특이적 아데노신 탈아미노효소 1	사카로미세스 세레비시아	탈아미노화
Tad1	tRNA-특이적 아데노신 탈아미노효소 1	호모 사피엔스	탈아미노화
Tad2	tRNA-특이적 아데노신 탈아미노효소 서브유닛 TAD2	사카로미세스 세레비시아	탈아미노화
Tad3	tRNA-특이적 아데노신 탈아미노효소 서브유닛 TAD3	사카로미세스 세레비시아	탈아미노화
TadA	tRNA-특이적 아데노신 탈아미노효소	대장균(Escherichia coli)	탈아미노화
TadA	tRNA-특이적 아데노신 탈아미노효소, 엽록체성	애기장대(Arabidopsis thaliana)	탈아미노화
CDAT8	tRNA-특이적 시티딘 탈아미노효소	메타노피루스 칸들레리(Methanopyrus kandleri)	탈아미노화
MiaA	tRNA (아데노신(37)-N6)-디메틸알릴전달효소	대장균(Escherichia coli)	이소펜테닐화

[0070]

[0071]

Mod5	tRNA 디메틸알릴전달효소, 미토콘드리아성	사카로미세스 세레비시아	이소펜테닐화
Mod5	tRNA 디메틸알릴전달효소, 미토콘드리아성	호모 사피엔스	이소펜테닐화
Sua5	tRNA 트레오닐카르바모일아데노신 생합성 단백질 SUA5	사카로미세스 세레비시아	카르바모일화
TsaD	tRNA N6-트레오닐카르바모일아데노신(37) 합성 단백질	대장균(Escherichia coli)	카르바모일화
TiaS	tRNA(Ile2) 2-아그마티닐시티딘 합성효소 TiaS	아키오글로부스 풀기두스(Archaeoglobus fulgidus)	아그마티닐화
NAT10	N-아세틸전달효소 10	호모 사피엔스	아세틸화
TilS	tRNA 리시딘(34) 합성효소	대장균(Escherichia coli)	리실화
TtcA	tRNA 2-티오시티딘(32) 합성효소	대장균(Escherichia coli)	O/S 교환
TrmU	미토콘드리아 tRNA-특이적 2-티오유리딜레이즈(thiouridylase) 1	호모 사피엔스	O/S 교환
Man/Gal-Q-전달효소	tRNA-퀴에우인 글리코실전달효소	다양한 진핵생물	갈락토실화
GluQRS	tRNA 글루타밀-Q(34) 합성효소	대장균(Escherichia coli)	글루타밀화
Man/Gal-Q-전달효소	tRNA-퀴에우인 글리코실전달효소	다양한 진핵생물	만노실화
Dus1	tRNA-디히드로우리딘 합성효소 1	사카로미세스 세레비시아	수소화
DusA	tRNA 디히드로우리딘 합성효소 A	대장균(Escherichia coli)	수소화
Pus1	tRNA 유사우리딘 합성효소 1, 미토콘드리아성	호모 사피엔스	유사우리딘 형성

Pus1	tRNA 유사우리딘 합성효소 1	사카로미세스 세레비시아	유사우리딘 형성
RluA	tRNA 유사우리딘(32) 합성효소 / 리보솜 대형 서브유닛 유사우리딘 합성효소 A	대장균(Escherichia coli)	유사우리딘 형성
MnmE	tRNA 우리딘(34) 5-카르복시메틸아미노메틸 합성 GTP 가수분해효소	대장균(Escherichia coli)	카르복시메틸아미노메틸화
MnmE	tRNA 우리딘(34) 5-카르복시메틸아미노메틸 합성 GTP 가수분해효소	대장균(Escherichia coli)	아미노메틸화
MnmCD	tRNA (5-메틸아미노메틸-2-티오우리딜레이트)-메틸전달효소 / FAD-의존성 cmnm(5)s(2)U34 산화환원효소	대장균(Escherichia coli)	탈카르복시메틸화
FTO	알파-케토글루타르산-의존성 2 산소화효소 FTO	호모 사피엔스	탈수소화
CmoB	tRNA 5-메톡시우리딘(34) 합성효소	대장균(Escherichia coli)	카르복시메틸화
MiaE	tRNA 2-메틸티오-N6-이소펜테닐 아데노신(37) 수산화효소	쥐티푸스균(Salmonella typhimurium)	히드록실화
FTO	알파-케토글루타르산-의존성 2 산소화효소 FTO	호모 사피엔스	히드록실화
ALKBH8	알킬화 DNA 복구 단백질 alkB 동족체 8	호모 사피엔스	히드록실화
MiaB	tRNA (N6-이소펜테닐 아데노신(37)-C2)-메틸티오전달효소	대장균(Escherichia coli)	메틸티올화
TYW2	tRNA 와이부토신-합성 단백질 2	사카로미세스 세레비시아	3-아미노-3-카르복시프로필화
ALKBH5		호모 사피엔스	탈메틸화

[0072]

FTO	알파-케토글루타르산-의존성 2 산소화효소 FTO	호모 사피엔스	탈메틸화
HCAP1	RNA-캐핑 효소	호모 사피엔스	5′-5′ 구아닐화
HCAP1	RNA-캐핑 효소	호모 사피엔스	탈인산화

[0073]

[0074] 추가의 RNA 조절 도메인은 표 2의 하기 인간 단백질들로부터의 기능적 도메인들을 포함한다:

[0075] 표 2

유전자 기호	동의어	유전자 설명
RBM5	LUCA15; H37	RNA 결합 모티프 단백질 5 [출처:HGNC Symbol;Acc:9902]
RBM6	DEF-3; 3G2; NY-LU-12; g16; DEF3	RNA 결합 모티프 단백질 6 [출처:HGNC Symbol;Acc:9903]
		추정 RNA 엑소뉴클레아제 NEF-sp (EC 3.1.-.-) [출처:UniProtKB/Swiss-Prot;Acc:Q96IC2]
YBX2	MSY2; CSDA3	Y 박스 결합 단백질 2 [출처:HGNC Symbol;Acc:17948]
CSDE1	D1S155E; UNR	저온 충격 도메인 함유 E1, RNA-결합 [출처:HGNC Symbol;Acc:29905]
PTBP1	HNRPI; HNRNP-I; PTB2; PTB3; PTB-1; PTB4; pPTB	폴리피리미딘 관 결합 단백질 1 [출처:HGNC Symbol;Acc:9583]
ZC3H3	KIAA0150	징크 핑거 CCCH-형 함유 3 [출처:HGNC Symbol;Acc:28972]
MATR3	KIAA0723; MGC9105	마트린 3 [출처:HGNC Symbol;Acc:6912]
SAMD4A	KIAA1053; DKFZP434H0350; 스마우그; SMG; SMGA; hSmaug1	멸균 알파 모티프 도메인 함유 4A [출처:HGNC Symbol;Acc:23023]
YTHDC2	FLJ2194; FLJ10053; DKFZp564A186	YTH 도메인 함유 2[출처:HGNC Symbol;Acc:24721]
CUGBP2		CUG 3 중 반복, RNA 결합 단백질 2 [출처:HGNC Symbol;Acc:2550]
PUM2	PUMH2; KIAA0235	푸밀리오 동족체 2 (초파리) [출처:HGNC Symbol;Acc:14958]
RC3H2	FLJ20301; FLJ20713; RNF164	링 핑거 및 CCCH-형 징크 핑거 도메인들 2 [출처:HGNC Symbol;Acc:21461]
ZC3H11A	KIAA0663	징크 핑거 CCCH-형 함유 11A [출처:HGNC Symbol;Acc:29093]
PARP12	FLJ22693; PARP-12; ZC3H1	폴리 (ADP-리보스) 중합효소 패밀리, 멤버 12 [출처:HGNC Symbol;Acc:21919]
CSDA	dbpA; ZONAB; CSDA1	저온 충격 도메인 단백질 A 유사유전자 1 [출처:HGNC Symbol;Acc:2429]

[0076]

SFRS8	SWAP	스플라이싱 인자, 아르기닌/세린-풍부 8 (흰색-살구-의-억제인자 동족체, 초파리) [출처:HGNC Symbol;Acc:10790]
EIF4B		진핵생물 번역 개시 인자 4B [출처:HGNC Symbol;Acc:3285]
U2AF2	U2AF65	U2 소핵 RNA 보조 인자 2 [출처:HGNC Symbol;Acc:23156]
SFRS14	KIAA0365	스플라이싱 인자, 아르기닌/세린-풍부 14 [출처:HGNC Symbol;Acc:18641]
SPEN	KIAA0929; MINT; SHARP; RBM15C	spen 동족체, 전사 조절인자(초파리) [출처:HGNC Symbol;Acc:17575]
ZC3H15	LEREPO4	징크 핑거 CCCH-형 함유 15 [출처:HGNC Symbol;Acc:29528]
YBX1	YB-1; YB1; DBPB; NSEP-1; MDR-NF1; BP-8; CSDB; CSDA2	Y 박스 결합 단백질 1 [출처:HGNC Symbol;Acc:8014]
ELAVL1	HuR; Hua; MelG	ELAV(미발달 치사, 비정상 시력, 초파리)-유사 1 (Hu 항원 R) [출처:HGNC Symbol;Acc:3312]
THUMPD1	FLJ20274	THUMP 도메인 함유 1[출처:HGNC Symbol;Acc:23807]
DHX8	HRH1; PRP22; PRPF22	DEAH (Asp-Glu-Ala-His) 박스 폴리펩타이드 8 [출처:HGNC Symbol;Acc:2749]
SRBD1	FLJ10379	S1 RNA 결합 도메인 1 [출처:HGNC Symbol;Acc:25521]
PABPC1	PABP1; PABPL1	폴리(A) 결합 단백질, 세포질 1 [출처:HGNC Symbol;Acc:8554]
DAZAP1	MGC19907	DAZ 관련 단백질 1 [출처:HGNC Symbol;Acc:2683]
IGF2BP2	IMP-2	인슐린-유사 성장 인자 2 mRNA 결합 단백질 2 [출처:HGNC Symbol;Acc:28867]
ZNF638	NP220; MGC26130; Zfp638	징크 핑거 단백질 638 [출처:HGNC Symbol;Acc:17894]
SART3	KIAA0156; RP11-13G14	T 세포에 의해 인식되는 편평세포암종 항원 3 [출처:HGNC Symbol;Acc:16860]
MKRN2	RNF62; HSPC070	마코린 링 핑거 단백질 2 [출처:HGNC Symbol;Acc:7113]

[0077]

[0078]

RBM7		RNA 결합 모티프 단백질 7 [출처:HGNC Symbol;Acc:9904]
RBMS2	SCR3	RNA 결합 모티프, 단일 가닥 상호작용 단백질 2 [출처:HGNC Symbol;Acc:9909]
MBNL3	CHCR; FLJ11316; MBLX39; MBXL	머슬블라인드(muscleblind)-유사 3 (초파리) [출처:HGNC Symbol;Acc:20564]
SNRPA	U1A; U1-A; Mud1	소핵 리보핵산단백질 폴리펩타이드 A [출처:HGNC Symbol;Acc:11151]
SYNJ2	INPP5H	시냅토자닌(synaptojanin) 2 [출처:HGNC Symbol;Acc:11504]
A2BP1		Fox-1 동족체 A (아탁신-2-결합 단백질 1) (헥사리보뉴클레오타이드-결합 단백질 1) [출처:UniProtKB/Swiss-Prot;Acc:Q9NWB1]
		트리뉴클레오타이드 반복-함유 유전자 6C 단백질 [출처:UniProtKB/Swiss-Prot;Acc:Q9HCJ0]
DDX43	HAGE; DKFZp434H2114; CT13	DEAD (Asp-Glu-Ala-Asp) 박스 폴리펩타이드 43 [출처:HGNC Symbol;Acc:18677]
KIAA0020	XTP5; PEN; PUF6	KIAA0020 [출처:HGNC Symbol;Acc:29676]
CNOT4	CLONE243; NOT4H	CCR4-NOT 전사 복합체, 서브유닛 4 [출처:HGNC Symbol;Acc:7880]
YTHDC1	YT521; KIAA1966; YT521-B	YTH 도메인 함유 1[출처:HGNC Symbol;Acc:30626]
PPIE	CyP-33; MGC3736; MGC111222	펩티딜프롤릴 이성화효소 E (시클로필린 E) [출처:HGNC Symbol;Acc:9258]
CHERP	ERPROT213-21; DAN16	칼슘 항상성 소포체 단백질 [출처:HGNC Symbol;Acc:16930]
RBM22	FLJ10290; ZC3H16; fSAP47; Cwc2	RNA 결합 모티프 단백질 22 [출처:HGNC Symbol;Acc:25503]
KHSRP	KSRP; FBP2; FUBP2	KH-형 스플라이싱 조절 단백질 [출처:HGNC Symbol;Acc:6316]
FUS	TLS; FUS1; hnRNP-P2	육종에 융합됨 [출처:HGNC Symbol;Acc:4010]
RBM41	FLJ11016	RNA 결합 모티프 단백질 41 [출처:HGNC Symbol;Acc:25617]
PCBP4	MCG10; LIP4	폴리(rC) 결합 단백질 4 [출처:HGNC Symbol;Acc:8652]

PABPC4	iPABP; APP-1	폴리(A) 결합 단백질, 세포질 4 (유도 가능 형태) [출처:HGNC Symbol;Acc:8557]
TNRC6A	CAGH26; KIAA1460; GW182	트리뉴클레오타이드 반복 함유 6A [출처:HGNC Symbol;Acc:11969]
RBM27	KIAA1311; ARRS1; Psc1	RNA 결합 모티프 단백질 27 [출처:HGNC Symbol;Acc:29243]
HNRNPC	hnRNPC	이종 핵 리보핵산단백질 C (C1/C2) [출처:HGNC Symbol;Acc:5035]
DAZL	DAZH; SPGYLA; MGC26406; DAZL1	무정자증-유사에서 결실됨 [출처:HGNC Symbol;Acc:2685]
HNRNPH3	2H9	이종 핵 리보핵산단백질 H3 (2H9) [출처:HGNC Symbol;Acc:5043]
SETD1A	KIAA0339; Set1; KMT2F	SET 도메인 함유 1A [출처:HGNC Symbol;Acc:29010]
CIRBP	CIRP	저온 유도성 RNA 결합 단백질 [출처:HGNC Symbol;Acc:1982]
HNRNPM	HTGR1; HNRNPM4; HNRPM4	이종 핵 리보핵산단백질 M [출처:HGNC Symbol;Acc:5046]
TRMT2A	HTF9C	TRM2 tRNA 메틸전달효소 2 동족체 A (사카로미세스 세레비시아) [출처:HGNC Symbol;Acc:24974]
SF3A1	SF3a120; SAP114; PRPF21; Prp21	스플라이싱 인자 3a, 서브유닛 1, 120kDa [출처:HGNC Symbol;Acc:10765]
SNRPD3	SMD3; Sm-D3	소핵 리보핵산단백질 D3 폴리펩타이드 18kDa [출처:HGNC Symbol;Acc:11160]
POLDIP3	PDIP46; KIAA1649	중합효소 (DNA-유도된), 델타 상호작용 단백질 3 [출처:HGNC Symbol;Acc:23782]
ZMAT5		징크 핑거, 마트린형 5 [출처:HGNC Symbol;Acc:28046]
RBM9	HNRBP2; FOX-2	RNA 결합 모티프 단백질 9 [출처:HGNC Symbol;Acc:9906]
ZC3H7B	RoXaN; FLJ13787; DKFZp434K0920; KIAA1031	징크 핑거 CCCH-형 함유 7B [출처:HGNC Symbol;Acc:30869]
RBM23	FLJ10482	RNA 결합 모티프 단백질 23 [출처:HGNC Symbol;Acc:20155]

[0079]

SFRS5	SRP40; HRS	스플라이싱 인자, 아르기닌/세린-풍부 5 [출처:HGNC Symbol;Acc:10787]
ZC3H14	FLJ11806; UKp68; NY-REN-37	징크 핑거 CCCH-형 함유 14 [출처:HGNC Symbol;Acc:20509]
ACIN1	KIAA0670; fSAP152	세포자연사 염색질 축합(condensation) 유도인자 1 [출처:HGNC Symbol;Acc:17066]
PABPN1	PAB2	폴리(A) 결합 단백질, 핵 1 [출처:HGNC Symbol;Acc:8565]
PABPC1L	dJ1069P2.3; PABPC1L1; ePAB	폴리(A) 결합 단백질, 세포질 1-유사 [출처:HGNC Symbol;Acc:15797]
BRUNOL4		브루노-유사 4, RNA 결합 단백질 (초파리) [출처:HGNC Symbol;Acc:14015]
CSTF2		절단 자극 인자, 3′ pre-RNA, 서브유닛 2, 64kDa [출처:HGNC Symbol;Acc:2484]
ZC3H12B	MCPIP2	징크 핑거 CCCH-형 함유 12B [출처:HGNC Symbol;Acc:17407]
FMR1	FMRP; FRAXA; MGC87458	취약 X 정신 지체 1 [출처:HGNC Symbol;Acc:3775]
HTATSF1	TAT-SF1	HIV-1 Tat 특이적 인자 1 [출처:HGNC Symbol;Acc:5276]
RBM3	IS1-RNPL	RNA 결합 모티프 (RNP1, RRM) 단백질 3 [출처:HGNC Symbol;Acc:9900]
ESRP2	FLJ21918	상피 스플라이싱 조절 단백질 2[출처:HGNC Symbol;Acc:26152]
MTHFSD	FLJ12998	메테닐테트라하이드로폴레이트 합성효소 도메인 함유 [출처:HGNC Symbol;Acc:25778]
DNAJC17	FLJ10634	DnaJ (Hsp40) 동족체, 하위패밀리 C, 멤버 17 [출처:HGNC Symbol;Acc:25556]
MYEF2	MEF-2; FLJ11213; KIAA1341; HsT18564	미엘린 발현 인자 2 [출처:HGNC Symbol;Acc:17940]
ESRP1	FLJ20171	상피 스플라이싱 조절 단백질 1[출처:HGNC Symbol;Acc:25966]
HNRNPL		이종 핵 리보핵산단백질 L [출처:HGNC Symbol;Acc:5045]
SNRNP70	U1-70K; Snp1	소핵 리보핵산단백질 70kDa (U1) [출처:HGNC Symbol;Acc:11150]

[0080]

SFRS16	SWAP2; CLASP	스플라이싱 인자, 아르기닌/세린-풍부 16 [출처:HGNC Symbol;Acc:17731]
TRMT1	FLJ20244	TRM1 tRNA 메틸전달효소 1 동족체 (사카로미세스 세레비시아) [출처:HGNC Symbol;Acc:25980]
NOVA2	ANOVA	신경-종양 복부 항원 2 [출처:HGNC Symbol;Acc:7887]
SF4	F23858; DKFZp434E2216; RBP	스플라이싱 인자 4 [출처:HGNC Symbol;Acc:18643]
ZC3HAV1	ZAP; FLB6421; FLJ13288; MGC48898; ZC3HDC2; ZC3H2; PARP13	징크 핑거 CCCH-형, 항바이러스성 1 [출처:HGNC Symbol;Acc:23721]
EIF3B	PRT1; eIF3b	진핵생물 번역 개시 인자 3, 서브유닛 B [출처:HGNC Symbol;Acc:3280]
RBM28	FLJ10377	RNA 결합 모티프 단백질 28 [출처:HGNC Symbol;Acc:21863]
LSM5	YER146W	LSM5 동족체, U6 소핵 RNA 관련 (사카로미세스 세레비시아) [출처:HGNC Symbol;Acc:17162]
EIF4H	WSCR1; KIAA0038	진핵생물 번역 개시 인자 4H [출처:HGNC Symbol;Acc:12741]
ELAVL2	HuB; HEL-N1	ELAV (미발달 치사, 비정상 시력, 초파리)-유사 2 (Hu 항원 B) [출처:HGNC Symbol;Acc:3313]
FUBP3		Far 상류부위 요소(FUSE) 결합 단백질 3 [출처:HGNC Symbol;Acc:4005]
CPEB3	KIAA0940	세포질 폴리아데닐화 요소 결합 단백질 3 [출처:HGNC Symbol;Acc:21746]
LARP4B		La 리보핵산단백질 도메인 패밀리, 멤버 4B [출처:HGNC Symbol;Acc:28987]
SUPT6H	KIAA0162; SPT6H	Ty 6 의 억제인자 동족체 (사카로미세스 세레비시아) [출처:HGNC Symbol;Acc:11470]
PPARGC1A	PGC1; PGC1A	퍼옥시솜 증식체 활성화 수용체 감마, 공활성화인자 1 알파 [출처:HGNC Symbol;Acc:9237]
CPSF6	CFIM; HPBRII-4; HPBRII-7; CFIM68	마이크로 RNA 1279 [출처:HGNC Symbol;Acc:35357]

[0081]

[0082]

KRR1	RIP-1	KRR1, 소형 서브유닛(SSU) processome 성분, 동족체 (효모) [출처:HGNC Symbol;Acc:5176]
SFRS9	SRp30c	스플라이싱 인자, 아르기닌/세린-풍부 9 [출처:HGNC Symbol;Acc:10791]
SFRS3	SRp20	스플라이싱 인자, 아르기닌/세린-풍부 3 [출처:HGNC Symbol;Acc:10785]
RBM24	FLJ30829; dJ259A10.1	RNA 결합 모티프 단백질 24 [출처:HGNC Symbol;Acc:21539]
KHDRBS2	SLM1; SLM-1; MGC26664	KH 도메인 함유, RNA 결합, 신호 전달 관련 2 [출처:HGNC Symbol;Acc:18114]
QKI	QK3	퀘이킹(quaking) 동족체, KH 도메인 RNA 결합 (마우스) [출처:HGNC Symbol;Acc:21100]
CPEB4	KIAA1673	세포질 폴리아데닐화 요소 결합 단백질 4 [출처:HGNC Symbol;Acc:21747]
FXR1		취약 X 정신 지체 , 상염색체 동족체 1 [출처:HGNC Symbol;Acc:4023]
NCL	C23	뉴클레오린 [출처:HGNC Symbol;Acc:7667]
		Pre-mRNA 분지 부위 단백질 p14 (SF3B 14 kDa 서브유닛) [출처:UniProtKB/Swiss-Prot;Acc:Q9Y3B4]
HDLBP	HBP	고밀도 지질단백질 결합 단백질 [출처:HGNC Symbol;Acc:4857]
SFRS7	9G8; ZCRB2; HSSG1; AAG3; RBM37; ZCCHC20	스플라이싱 인자, 아르기닌/세린-풍부 7, 35kDa [출처:HGNC Symbol;Acc:10789]
PNO1		NOB1 의 억제인자 동족체 (사카로미세스 세레비시아) [출처:HGNC Symbol;Acc:32790]
TIA1		TIA1 세포 독성 과립 관련 RNA 결합 단백질 [출처:HGNC Symbol;Acc:11802]
SFRS4	SRP75	스플라이싱 인자, 아르기닌/세린-풍부 4 [출처:HGNC Symbol;Acc:10786]
SFPQ	PSF	스플라이싱 인자 프롤린/글루타민-풍부 (폴리피리미딘 관 결합 단백질 관련) [출처:HGNC Symbol;Acc:10774]
SFRS11	p54; NET2	스플라이싱 인자, 아르기닌/세린-풍부 11 [출처:HGNC Symbol;Acc:10782]

PRPF3	Prp3; HPRP3	PRP3 pre-mRNA 처리 인자 3 동족체 (사카로미세스 세레비시아) [출처:HGNC Symbol;Acc:17348]
PTBP2	brPTB; nPTB; PTB; PTBLP	폴리피리미딘 관 결합 단백질 2 [출처:HGNC Symbol;Acc:17662]
ROD1	PTBP3	ROD1 분화 조절인자 1 (스키조사카로미세스 폼베) [출처:HGNC Symbol;Acc:10253]
RBM18	MGC2734	RNA 결합 모티프 단백질 18 [출처:HGNC Symbol;Acc:28413]
C14orf156	DC50	SRA 줄기-루프-상호작용 RNA-결합 단백질, 미토콘드리아 전구체 [출처:UniProtKB/Swiss-Prot;Acc:Q9GZT3]
RBM25	S164; fSAP94; NET52; Snu71	RNA 결합 모티프 단백질 25 [출처:HGNC Symbol;Acc:23244]
ENOX1	FLJ10094; PIG38; CNOX; cCNOX	엑토-NOX 디설파이드-티올 교환체 1 [출처:HGNC Symbol;Acc:25474]
TARDBP	TDP-43; ALS10	TAR DNA 결합 단백질 [출처:HGNC Symbol;Acc:11571]
AKAP1	AKAP121; AKAP149; SAKAP84; S-AKAP84; AKAP84	A 키나아제(PRKA) 고정 단백질 1 [출처:HGNC Symbol;Acc:367]
PSPC1	PSP1; FLJ10955	파라스펙클 성분 1 [출처:HGNC Symbol;Acc:20320]
ZCCHC17	PS1D; HSPC251; pNO40	징크 핑거, CCHC 도메인 함유 17 [출처:HGNC Symbol;Acc:30246]
KHDRBS1	Sam68; p62; FLJ34027	KH 도메인 함유, RNA 결합, 신호 전달 관련 1 [출처:HGNC Symbol;Acc:18116]
ZC3H7A	HSPC055; FLJ20318	징크 핑거 CCCH-형 함유 7A [출처:HGNC Symbol;Acc:30959]
HNRNPA2B1		이종 핵 리보핵산단백질 A2/B1 [출처:HGNC Symbol;Acc:5033]
BICC1		bicaudal C 동족체 1 (초파리) [출처:HGNC Symbol;Acc:19351]
RBM19	DKFZp586F1023; KIAA0682	RNA 결합 모티프 단백질 19 [출처:HGNC Symbol;Acc:29098]
ZC3H13	DKFZp434D1812	징크 핑거 CCCH-형 함유 13 [출처:HGNC Symbol;Acc:20368]

[0083]

SFRS6	SRP55; B52	스플라이싱 인자, 아르기닌/세린-풍부 6 [출처:HGNC Symbol;Acc:10788]
RNF113A	RNF113; Cwc24	링 핑거 단백질 113A [출처:HGNC Symbol;Acc:12974]
SNRPD2	Sm-D2	소핵 리보핵산단백질 D2 폴리펩타이드 16.5kDa [출처:HGNC Symbol;Acc:11159]
SNRPB	COD; SmB/SmB'; Sm-B/B'; snRNP-B	소핵 리보핵산단백질 폴리펩타이드 B 및 B1 [출처:HGNC Symbol;Acc:11153]
SNRPB2	Msl1	소핵 리보핵산단백질 폴리펩타이드 B" [출처:HGNC Symbol;Acc:11155]
HNRNPR	hnRNP-R	이종 핵 리보핵산단백질 R [출처:HGNC Symbol;Acc:5047]
RALY	P542; HNRPCL2	RNA 결합 단백질, 자가 항원성 (hnRNP-관련 치사 황색 동족체 (마우스)) [출처:HGNC Symbol;Acc:15921]
RBM42	MGC10433	RNA 결합 모티프 단백질 42 [출처:HGNC Symbol;Acc:28117]
HNRNPH2	hnRNPH'; FTP3; HNRPH'	이종 핵 리보핵산단백질 H2 (H') [출처:HGNC Symbol;Acc:5042]
ZFP36	RNF162A; TIS11; G0S24; TTP; NUP475	징크 핑거 단백질 36, C3H 형, 동족체 (마우스) [출처:HGNC Symbol;Acc:12862]
NAA38	YJR022W	N(알파)-아세틸전달효소 38, NatC 보조 서브유닛 [출처:HGNC Symbol;Acc:20471]
SNRPN	SMN; SM-D; HCERN3; SNRNP-N; SNURF-SNRPN; RT-LI	소핵 리보핵산단백질 폴리펩타이드 N [출처:HGNC Symbol;Acc:11164]
FXR2		취약 X 정신 지체 , 상염색체 동족체 2 [출처:HGNC Symbol;Acc:4024]
SAFB2	KIAA0138	지지체 부착 인자 B2 [출처:HGNC Symbol;Acc:21605]
LSM7	YNL147W	LSM7 동족체, U6 소핵 RNA 관련 (사카로미세스 세레비시아) [출처:HGNC Symbol;Acc:20470]
LSM4	YER112W	LSM4 동족체, U6 소핵 RNA 관련 (사카로미세스 세레비시아) [출처:HGNC Symbol;Acc:17259]
ZC3H4	KIAA1064	징크 핑거 CCCH-형 함유 4 [출처:HGNC Symbol;Acc:17808]

[0084]

EIF3G	eIF3-델타; eIF3-p44; eIF3g	진핵생물 번역 개시 인자 3, 서브유닛 G [출처:HGNC Symbol;Acc:3274]
DKC1	XAP101; 디스카린(dyskerin); NAP57; NOLA4	소핵소체 RNA, H/ACA 박스 56 [출처:HGNC Symbol;Acc:32650]
PPIL4		펩티딜프롤릴 이성화효소 (시클로필린)-유사 4 [출처:HGNC Symbol;Acc:15702]
RBM39	CC1.3; HCC1; CAPER; fSAP59	RNA 결합 모티프 단백질 39 [출처:HGNC Symbol;Acc:15923]
ANKHD1	MASK; FLJ20288; FLJ11979; FLJ10042; FLJ14127; KIAA1085	안키린 반복 및 KH 도메인 함유 1 [출처:HGNC Symbol;Acc:24714]
KHDRBS3	T-STAR; Etle; etoile; SALP; SLM2; SLM-2	KH 도메인 함유, RNA 결합, 신호 전달 관련 3 [출처:HGNC Symbol;Acc:18117]
RBM8A	ZNRP; BOV-1A; BOV-1B; BOV-1C; RBM8B; Y14	RNA 결합 모티프 단백질 8A [출처:HGNC Symbol;Acc:9905]
LIN28		lin-28 동족체 (예쁜꼬마선충) [출처:HGNC Symbol;Acc:15986]
GRSF1		G-풍부 RNA 서열 결합 인자 1 [출처:HGNC Symbol;Acc:4610]
ANKRD17	GTAR; KIAA0697; FLJ22206; NY-BR-16	안키린 반복 도메인 17 [출처:HGNC Symbol;Acc:23575]
UNK	KIAA1753	언켐프트 동족체(unkempt homolog) (초파리) [출처:HGNC Symbol;Acc:29369]
NIP7	CGI-37; FLJ10296; HSPC031; KD93	핵 이동(nuclear import) 7 동족체 (사카로미세스 세레비시아) [출처:HGNC Symbol;Acc:24328]
TOE1		EGR1 의 표적, 멤버 1 (핵) [출처:HGNC Symbol;Acc:15954]
RBM38	HSRNASEB; SEB4D; seb4B; dJ800J21.2	RNA 결합 모티프 단백질 38 [출처:HGNC Symbol;Acc:15818]
SRRM1	SRM160; POP101; MGC39488	세린/아르기닌 반복 매트릭스(repetitive matrix) 1 [출처:HGNC Symbol;Acc:16638]
MKRN1	RNF61	마코린 링 핑거 단백질 1 [출처:HGNC Symbol;Acc:7112]

[0085]

EIF2S1	EIF-2 알파; EIF2A	진핵생물 번역 개시 인자 2, 서브유닛 1 알파, 35kDa [출처:HGNC Symbol;Acc:3265]
THUMPD3	DKFZP434F091	THUMP 도메인 함유 3[출처:HGNC Symbol;Acc:24493]
RBMX2	CGI-79	RNA 결합 모티프 단백질, X-연관 2 [출처:HGNC Symbol;Acc:24282]
PUM1	PUMH1; KIAA0099	푸밀리오 동족체 1 (초파리) [출처:HGNC Symbol;Acc:14957]
MSI1		무사시(musashi) 동족체 1 (초파리) [출처:HGNC Symbol;Acc:7330]
SYNCRIP	NSAP1; GRY-RBP; dJ3J17.2; HNRPQ1; hnRNP-Q	시나프토타그민 결합, 세포질 RNA 상호작용 단백질 [출처:HGNC Symbol;Acc:16918]
ZC3H10	FLJ14451	징크 핑거 CCCH-형 함유 10 [출처:HGNC Symbol;Acc:25893]
HNRNPA1	hnRNPA1; hnRNP-A1	이종 핵 리보핵산단백질 A1 [출처:HGNC Symbol;Acc:5031]
RC3H1	KIAA2025; 로킨(roquin); RP5-1198E17.5; RNF198	링 핑거 및 CCCH-형 징크 핑거 도메인들 1 [출처:HGNC Symbol;Acc:29434]
IGF2BP3	IMP-3; CT98	인슐린-유사 성장 인자 2 mRNA 결합 단백질 3 [출처:HGNC Symbol;Acc:28868]
NUPL2	NLP_1; CG1; hCG1; H_RG271G13.9	뉴클레오포린(nucleoporin) 유사 2 [출처:HGNC Symbol;Acc:17010]
SFRS1	ASF; SF2; SRp30a; SF2p33; MGC5228	스플라이싱 인자, 아르기닌/세린-풍부 1 [출처:HGNC Symbol;Acc:10780]
TRA2B	Htra2-베타	형질전환유전자 2 베타 동족체) (초파리) [출처:HGNC Symbol;Acc:10781]
CPEB2		세포질 폴리아데닐화 요소 결합 단백질 2 [출처:HGNC Symbol;Acc:21745]
ALKBH8	MGC10235	alkB, 알킬화 복구 동족체 8 (대장균) [출처:HGNC Symbol;Acc:25189]
SLTM	Met; FLJ13213	SAFB-유사, 전사 조절인자 [출처:HGNC Symbol;Acc:20709]
PNPT1	PNPase; OLD35; old-35	폴리리보뉴클레오타이드 뉴클레오티딜전달효소 1 [출처:HGNC Symbol;Acc:23166]

[0086]

THUMPD2	MGC2454	THUMP 도메인 함유 2[출처:HGNC Symbol;Acc:14890]
ASCC1	CGI-18; ASC1p50; Em:AC022392.3	activating signal cointegrator 1 complex subunit 1 [출처:HGNC Symbol;Acc:24268]
SSB	LARP3; La	쇼그렌 증후군 항원 B (자가항원 La) [출처:HGNC Symbol;Acc:11316]
HNRNPD		이종 핵 리보핵산단백질 D (AU-풍부 요소 RNA 결합 단백질 1, 37kDa) [출처:HGNC Symbol;Acc:5036]
LARP1B	FLJ10378; DKFZp434K245; DKFZp686E0316	La 리보핵산단백질 도메인 패밀리, 멤버 1B [출처:HGNC Symbol;Acc:24704]
G3BP2	KIAA0660	GTP 가수분해효소 활성 단백질 (SH3 도메인) 결합 단백질 2 [출처:HGNC Symbol;Acc:30291]
ZCRB1	MADP-1; MADP1; RBM36; ZCCHC19	징크 핑거 CCHC-형 및 RNA 결합 모티프 1 [출처:HGNC Symbol;Acc:29620]
SNRPF	Sm-F	소핵 리보핵산단백질 폴리펩타이드 F [출처:HGNC Symbol;Acc:11162]
		이종 핵 리보핵산단백질 A1-유사 단백질 2 (hnRNP 코어 단백질 A1-유사 단백질 2) [출처:UniProtKB/Swiss-Prot;Acc:Q32P51]
SETD1B	KIAA1076; Set1B; KMT2G	SET 도메인 함유 1B [출처:HGNC Symbol;Acc:29187]
RBM26	PRO1777; SE70-2; FLJ20957; ZC3H17; ARRS2	RNA 결합 모티프 단백질 26 [출처:HGNC Symbol;Acc:20327]
MBNL2	MBLL; MBLL39	머슬블라인드(muscleblind)-유사 2 (초파리) [출처:HGNC Symbol;Acc:16746]
RNF113B	RNF161	링 핑거 단백질 113B [출처:HGNC Symbol;Acc:17267]
NOVA1		신경-종양 복부 항원 1 [출처:HGNC Symbol;Acc:7886]
BRUNOL6		브루노-유사 6, RNA 결합 단백질 (초파리) [출처:HGNC Symbol;Acc:14059]
DUS3L	DUS3; FLJ13896	디히드로우리딘 합성효소 3-유사 (사카로미세스 세레비시아) [출처:HGNC Symbol;Acc:26920]
SF3B4	SAP49; SF3b49; Hsh49	스플라이싱 인자 3b, 서브유닛 4, 49kDa [출처:HGNC Symbol;Acc:10771]
LGTN		ligatin [출처:HGNC Symbol;Acc:6583]

[0087]

[0088]

HNRPLL		이종 핵 리보핵산단백질 L-유사 [출처:HGNC Symbol;Acc:25127]
SNRPG	Sm-G	소핵 리보핵산단백질 폴리펩타이드 G [출처:HGNC Symbol;Acc:11163]
ZC3H8	Fliz1	징크 핑거 CCCH-형 함유 8 [출처:HGNC Symbol;Acc:30941]
RBMS3		RNA 결합 모티프, 단일 가닥 상호작용 단백질 [출처:HGNC Symbol;Acc:13427]
G3BP1	HDH-VIII; G3BP	GTP 가수분해효소 활성 단백질 (SH3 도메인) 결합 단백질 1 [출처:HGNC Symbol;Acc:30292]
NONO	NRB54; NMT55; P54NRB; P54	비-POU 도메인 함유, 옥타머-결합 [출처:HGNC Symbol;Acc:7871]
RBMX	RNMX; hnRNP-G	RNA 결합 모티프 단백질, X-연관 [출처:HGNC Symbol;Acc:9910]
A1CF	ACF; ASP; ACF64; ACF65; APOBEC1CF	APOBEC1 상보성 인자 [출처:HGNC Symbol;Acc:24086]
PPRC1	PRC; KIAA0595; MGC74642	퍼옥시솜 증식체 활성화 수용체 감마, 공활성화인자-관련 1 [출처:HGNC Symbol;Acc:30025]
PDCD11	KIAA0185; ALG-4	예정세포사멸 11 [출처:HGNC Symbol;Acc:13408]
TUT1	FLJ22347; FLJ22267; FLJ21850; PAPD2; TUTase	말단 우리딜일 전달효소 1, U6 snRNA-특이적 [출처:HGNC Symbol;Acc:26184]
CUGBP1		CUG 3 중 반복, RNA 결합 단백질 1 [출처:HGNC Symbol;Acc:2549]
RPS3	FLJ26283; FLJ27450; MGC87870	리보솜 단백질 S3 [출처:HGNC Symbol;Acc:10420]
ZC3H12C	KIAA1726; MCPIP3	징크 핑거 CCCH-형 함유 12C [출처:HGNC Symbol;Acc:29362]
CPSF7	FLJ12529	절단 및 폴리아데닐화 특이적 인자 7, 59kDa [출처:HGNC Symbol;Acc:30098]
YTHDF1	FLJ20391	YTH 도메인 패밀리, 멤버 1 [출처:HGNC Symbol;Acc:15867]
PABPC3	PABP3; tPABP	폴리(A) 결합 단백질, 세포질 3 [출처:HGNC Symbol;Acc:8556]

TIAL1	TIAR	TIA1 세포 독성 과립 관련 RNA 결합 단백질-유사 1 [출처:HGNC Symbol;Acc:11804]
RBM46	MGC27016; CT68	RNA 결합 모티프 단백질 46 [출처:HGNC Symbol;Acc:28401]
UHMK1	KIS; Kist	U2AF 상동성 모티프(UHM) 키나아제 1 [출처:HGNC Symbol;Acc:19683]
BOLL	BOULE	bol, boul-유사 (초파리) [출처:HGNC Symbol;Acc:14273]
ZFP36L2	ERF2; RNF162C; TIS11D	징크 핑거 단백질 36, C3H 형-유사 2 [출처:HGNC Symbol;Acc:1108]
PAN3		PAN3 폴리(A) 특이적 리보뉴클레아제 서브유닛 동족체 (사카로미세스 세레비시아) [출처:HGNC Symbol;Acc:29991]
MBNL1	KIAA0428; EXP42; EXP40; EXP35; EXP	머슬블라인드(muscleblind)-유사 (초파리) [출처:HGNC Symbol;Acc:6923]
HNRPDL	JKTBP; laAUF1	이종 핵 리보핵산단백질 D-유사 [출처:HGNC Symbol;Acc:5037]
CARHSP1	CRHSP-24; CSDC1	칼슘 조절 열 안정 단백질 1, 24kDa [출처:HGNC Symbol;Acc:17150]
RBMS1	SCR2; MSSP-1; MSSP-2; MSSP-3; YC1; HCC-4; DKFZp564H0764	RNA 결합 모티프, 단일 가닥 상호작용 단백질 1 [출처:HGNC Symbol;Acc:9907]
SFRS12	DKFZp564B176; SRrp86; SRrp508	스플라이싱 인자, 아르기닌/세린-풍부 12 [출처:HGNC Symbol;Acc:17882]
MSI2		무사시(musashi) 동족체 2 (초파리) [출처:HGNC Symbol;Acc:18585]
SFRS13B	SRrp35; SFRS19	스플라이싱 인자, 아르기닌/세린-풍부 13B [출처:HGNC Symbol;Acc:21220]
C4orf23	FLJ35725	추정 tRNA (우라실-O(2)-)-메틸전달효소 (EC 2.1.1.n2) [출처:UniProtKB/Swiss-Prot;Acc:Q8IYL2]
MKI67IP	Nopp34; NIFK	MKI67 (FHA 도메인) 상호작용 핵소체 인단백질 [출처:HGNC Symbol;Acc:17838]
LARP1	LARP; KIAA0731; MGC19556	La 리보핵산단백질 도메인 패밀리, 멤버 1 [출처:HGNC Symbol;Acc:29531]

[0089]

[0090]

RBM45	DRB1; FLJ44612	RNA 결합 모티프 단백질 45 [출처:HGNC Symbol;Acc:24468]
PPARGC1B	PERC; PGC1B	퍼옥시솜 증식체 활성화 수용체 감마, 공활성화인자 1 베타 [출처:HGNC Symbol;Acc:30022]
LSM11	FLJ38273	LSM11, U7 소핵 RNA 관련 [출처:HGNC Symbol;Acc:30860]
SFRS15	KIAA1172; DKFZp434E098; SRA4; SCAF4	스플라이싱 인자, 아르기닌/세린-풍부 15 [출처:HGNC Symbol;Acc:19304]
RBPMS	HERMES	다중 스플라이싱이 있는 RNA 결합 단백질 [출처:HGNC Symbol;Acc:19097]
ZC3H18	NHN1	징크 핑거 CCCH-형 함유 18 [출처:HGNC Symbol;Acc:25091]
SYNJ1	INPP5G	시냅토자닌(synaptojanin) 1 [출처:HGNC Symbol;Acc:11503]
IGF2BP1	IMP-1; ZBP1	인슐린-유사 성장 인자 2 mRNA 결합 단백질 1 [출처:HGNC Symbol;Acc:28866]
TNRC4		트리뉴클레오타이드 반복 함유 4 [출처:HGNC Symbol;Acc:11967]
U2AF1	U2AF35; RNU2AF1; RN	U2 소핵 RNA 보조 인자 1 [출처:HGNC Symbol;Acc:12453]
SAFB	HET; SAFB1	지지체 부착 인자 B [출처:HGNC Symbol;Acc:10520]
CPSF4	NAR; CPSF30	절단 및 폴리아데닐화 특이적 인자 4, 30kDa [출처:HGNC Symbol;Acc:2327]
BRUNOL5		브루노-유사 5, RNA 결합 단백질 (초파리) [출처:HGNC Symbol;Acc:14058]
U2AF1L4	MGC33901; U2af26	U2 소핵 RNA 보조 인자 1-유사 4 [출처:HGNC Symbol;Acc:23020]
SFRS2	SC-35; SC35; PR264; SFRS2A	스플라이싱 인자, 아르기닌/세린-풍부 2 [출처:HGNC Symbol;Acc:10783]
LARP4	PP13296	La 리보핵산단백질 도메인 패밀리, 멤버 4 [출처:HGNC Symbol;Acc:24320]
RAVER1	KIAA1978	리보핵산단백질, PTB-결합 1 [출처:HGNC Symbol;Acc:30296]
ELAVL4	PNEM	ELAV (미발달 치사, 비정상 시력, 초파리)-유사 4 (Hu 항원 D) [출처:HGNC Symbol;Acc:3315]

RAVER2	KIAA1579; FLJ10770	리보핵산단백질, PTB-결합 2 [출처:HGNC Symbol;Acc:25577]
FUBP1	FBP	Far 상류부위 요소(FUSE) 결합 단백질 1 [출처:HGNC Symbol;Acc:4004]
RBM15	OTT; OTT1	RNA 결합 모티프 단백질 15 [출처:HGNC Symbol;Acc:14959]
DHX57	DDX57	DEAH (Asp-Glu-Ala-Asp/His) 박스 폴리펩타이드 57 [출처:HGNC Symbol;Acc:20086]
TDRD10	DKFZp434M202	튜더 도메인 함유 10 [출처:HGNC Symbol;Acc:25316]
TIPARP	DKFZP434J214; DKFZp686N0351; DDF1; PARP7; PARP-7; PARP-1; pART14; RM1	TCDD-유도성 폴리(ADP-리보스) 중합효소 [출처:HGNC Symbol;Acc:23696]
RBM47	FLJ20273; NET18	RNA 결합 모티프 단백질 47 [출처:HGNC Symbol;Acc:30358]
SR140		U2-관련 protein SR140 (140 kDa Ser/Arg-풍부 도메인 단백질) [출처:UniProtKB/Swiss-Prot;Acc:O15042]
TTC14	FLJ00166; KIAA1980	테트라트라이코펩타이드 반복 도메인 14 [출처:HGNC Symbol;Acc:24697]
LSM6	YDR378C	LSM6 동족체, U6 소핵 RNA 관련 (사카로미세스 세레비시아) [출처:HGNC Symbol;Acc:17017]
TRA2A	htra-2-alpha; tra2a	형질전환유전자 2 알파 동족체 (초파리) [출처:HGNC Symbol;Acc:16645]
HNRNPK	CSBP; TUNP	마이크로 RNA 7-1 [출처:HGNC Symbol;Acc:31638]
ENOX2	APK1; tNOX	엑토-NOX 디설파이드-티올 교환체 2 [출처:HGNC Symbol;Acc:2259]
LARP6	아케론; FLJ11196	La 리보핵산단백질 도메인 패밀리, 멤버 6 [출처:HGNC Symbol;Acc:24012]
KIAA0430	LKAP	KIAA0430 [출처:HGNC Symbol;Acc:29562]
RBPMS2		다중 스플라이싱이 있는 RNA 결합 단백질 2 [출처:HGNC Symbol;Acc:19098]
SNRPD1	HsT2456; Sm-D1	소핵 리보핵산단백질 D1 폴리펩타이드 16kDa [출처:HGNC Symbol;Acc:11158]
SAMD14	FLJ36890	멸균 알파 모티프 도메인 함유 14 [출처:HGNC Symbol;Acc:27312]

[0091]

		Fox-1 동족체 C [출처:UniProtKB/Swiss-Prot;Acc:A6NFN3]
POLR2G	hRPB19; hsRPB7; RPB7	중합효소 (RNA) II (DNA 유도된) 폴리펩타이드 G [출처:HGNC Symbol;Acc:9194]
SF1	ZFM1	스플라이싱 인자 1 [출처:HGNC Symbol;Acc:12950]
HNRNPH1	hnRNPH	이종 핵 리보핵산단백질 H1 (H) [출처:HGNC Symbol;Acc:5041]
ZRSR2	U2AF1-RS2; URP	징크 핑거 (CCCH 형), RNA-결합 모티프 및 세린/아르기닌 풍부 2 [Source:HGNC Symbol;Acc:23019]
PCBP1	HNRPE1; hnRNP-E1; HNRPX; hnRNP-X	폴리(rC) 결합 단백질 1 [출처:HGNC Symbol;Acc:8647]
RBMY1F	MGC33094	RNA 결합 모티프 단백질, Y-연관, 패밀리 1, 멤버 F [출처:HGNC Symbol;Acc:23974]
HNRNPF		이종 핵 리보핵산단백질 F [출처:HGNC Symbol;Acc:5039]
HNRNPA3		이종 핵 리보핵산단백질 A3 [출처:HGNC Symbol;Acc:24941]
RBMXL2	HNRNPG-T; HNRPGT	RNA 결합 모티프 단백질, X-연관-유사 2 [출처:HGNC Symbol;Acc:17886]
LSM3	YLR438C; SMX4; USS2	LSM3 동족체, U6 소핵 RNA 관련 (사카로미세스 세레비시아) [출처:HGNC Symbol;Acc:17874]
NCBP2L		핵 캡 결합 단백질 서브유닛 2-유사 [출처:HGNC Symbol;Acc:31795]
CSDC2	PIPPin	저온 충격 도메인 함유 C2, RNA 결합 [출처:HGNC Symbol;Acc:30359]
TAF15	hTAFII68; RBP56; Npl3	TAF15 RNA 중합효소 II, TATA 박스 결합 단백질 (TBP)-연관 인자, 68kDa [출처:HGNC Symbol;Acc:11547]
RBM4B	MGC10871; ZCCHC15; RBM4L; ZCRB3B	RNA 결합 모티프 단백질 4B [출처:HGNC Symbol;Acc:28842]
RBM4	LARK; RBM4A; ZCRB3A; ZCCHC21	RNA 결합 모티프 단백질 4 [출처:HGNC Symbol;Acc:9901]

[0092]

LARP7	HDCMA18P; PIP7S; DKFZP564K112	La 리보핵산단백질 도메인 패밀리, 멤버 7 [출처:HGNC Symbol;Acc:24912]
PABPC5	PABP5	폴리(A) 결합 단백질, 세포질 5 [출처:HGNC Symbol;Acc:13629]
LSM1	CASM; YJL124C	LSM1 동족체, U6 소핵 RNA 관련 (사카로미세스 세레비시아) [출처:HGNC Symbol;Acc:20472]
RBMXL3	FLJ40249	RNA 결합 모티프 단백질, X-연관-유사 3 [출처:HGNC Symbol;Acc:26859]
MEX3C	FLJ38871; RNF194	mex-3 동족체 C (예쁜꼬마선충) [출처:HGNC Symbol;Acc:28040]
RBM44	FLJ40411	RNA 결합 모티프 단백질 44 [출처:HGNC Symbol;Acc:24756]
CSTF2T	DKFZp434C1013; KIAA0689; CstF-64T	절단 자극 인자, 3′ pre-RNA, 서브유닛 2, 64kDa, 타우 변이체 [출처:HGNC Symbol;Acc:17086]
HNRNPA0	hnRNPA0	이종 핵 리보핵산단백질 A0 [출처:HGNC Symbol;Acc:5030]
ZC3H12D	dJ281H8.1; MCPIP4	징크 핑거 CCCH-형 함유 12D [출처:HGNC Symbol;Acc:21175]
SAMD4B	FLJ10211; MGC99832; SMGB; hSmaug2	멸균 알파 모티프 도메인 함유 4B [출처:HGNC Symbol;Acc:25492]
HNRNPCL1		이종 핵 리보핵산단백질 C-유사 1 [출처:HGNC Symbol;Acc:29295]
MKRN3	RNF63; ZFP127; MGC88288	마코린 링 핑거 단백질 3 [출처:HGNC Symbol;Acc:7114]
RBM15B	HUMAGCGB; OTT3	RNA 결합 모티프 단백질 15B [출처:HGNC Symbol;Acc:24303]
PUF60	FIR; SIAHBP1; RoBPI	폴리-U 결합 스플라이싱 인자 60KDa [출처:HGNC Symbol;Acc:17042]
TRNAU1AP	SECP43; FLJ20503	tRNA 셀레노시스테인 1 관련 단백질 1 [출처:HGNC Symbol;Acc:30813]
SFRS2B	SRP46	스플라이싱 인자, 아르기닌/세린-풍부 2B [출처:HGNC Symbol;Acc:16988]
MEX3D	Tino; KIAA2031; OK/SW-cl.4; RNF193	mex-3 동족체 D (예쁜꼬마선충) [출처:HGNC Symbol;Acc:16734]

[0093]

[0094]

LSM10	MGC15749	LSM10, U7 소핵 RNA 관련 [출처:HGNC Symbol;Acc:17562]
SNRPE	Sm-E	소핵 리보핵산단백질 폴리펩타이드 E [출처:HGNC Symbol;Acc:11161]
TDRKH	TDRD2	튜더 및 KH 도메인 함유 [출처:HGNC Symbol;Acc:11713]
RBM10	DXS8237E; KIAA0122; GPATC9; ZRANB5; GPATCH9	RNA 결합 모티프 단백질 10 [출처:HGNC Symbol;Acc:9896]
LENG9		백혈구 수용체 클러스터(LRC) 멤버 9 [출처:HGNC Symbol;Acc:16306]
EWSR1	EWS	유잉육종 절단점 영역 1 [출처:HGNC Symbol;Acc:3508]
LSMD1	MGC14151; PFAAP2	LSM 도메인 함유 1[출처:HGNC Symbol;Acc:28212]
DND1	MGC34750; RBMS4	막다른(dead end) 동족체 1 (제브라피시) [출처:HGNC Symbol;Acc:23799]
MEX3B	DKFZp434J0617; RNF195	mex-3 동족체 B (예쁜꼬마선충) [출처:HGNC Symbol;Acc:25297]
PCBP3		폴리(rC) 결합 단백질 3 [출처:HGNC Symbol;Acc:8651]
THOC4	ALY; BEF	THO 복합체 4 [출처:HGNC Symbol;Acc:19071]
RBM12B		RNA 결합 모티프 단백질 12B [출처:HGNC Symbol;Acc:32310]
SNRNP35	U1SNRNPBP	소핵 리보핵산단백질 35kDa (U11/U12) [출처:HGNC Symbol;Acc:30852]
PABPC1L2B		폴리(A) 결합 단백질, 세포질 1-유사 2B [출처:HGNC Symbol;Acc:31852]
RALYL	HNRPCL3	RALY RNA 결합 단백질-유사 [출처:HGNC Symbol;Acc:27036]
DDX53	CAGE; CT26	DEAD (Asp-Glu-Ala-Asp) 박스 폴리펩타이드 53 [출처:HGNC Symbol;Acc:20083]
RBM33	DKFZp686F102; MGC20460; DKFZp434D1319	RNA 결합 모티프 단백질 33 [출처:HGNC Symbol;Acc:27223]
RBM43	FLJ45645	RNA 결합 모티프 단백질 43 [출처:HGNC Symbol;Acc:24790]

[0095]

RBM11		RNA 결합 모티프 단백질 11 [출처:HGNC Symbol;Acc:9897]
ZFP36L1	RNF162B; Berg36; ERF1; TIS11B; cMG1	징크 핑거 단백질 36, C3H 형-유사 1 [출처:HGNC Symbol;Acc:1107]
YTHDF3	FLJ31657	YTH 도메인 패밀리, 멤버 3 [출처:HGNC Symbol;Acc:26465]
RNPC3	KIAA1839; FLJ20008; RBM40	RNA-결합 영역 (RNP1, RRM) 함유 3 [출처:HGNC Symbol;Acc:18666]
PABPC1L2A		폴리(A) 결합 단백질, 세포질 1-유사 2A [출처:HGNC Symbol;Acc:27989]
DAZ3		무정자증에서 결실됨 3 [출처:HGNC Symbol;Acc:15965]
RDM1	MGC33977	RAD52 모티프 1 [출처:HGNC Symbol;Acc:19950]
LIN28B	FLJ16517; CSDD2	lin-28 동족체 B (예쁜꼬마선충) [출처:HGNC Symbol;Acc:32207]
CPSF4L		절단 및 폴리아데닐화 특이적 인자 4-유사 [출처:HGNC Symbol;Acc:33632]
DAZ1	SPGY	무정자증에서 결실됨 1 [출처:HGNC Symbol;Acc:2682]
ZC3H6	FLJ41410; FLJ45877; KIAA2035	징크 핑거 CCCH-형 함유 6 [출처:HGNC Symbol;Acc:24762]
SFRS13A	TASR1; TASR2; SRp38; SRrp40; SFRS13; FUSIP1; FUSIP2	스플라이싱 인자, 아르기닌/세린-풍부 13A [출처:HGNC Symbol;Acc:16713]
RBM34	KIAA0117	RNA 결합 모티프 단백질 34 [출처:HGNC Symbol;Acc:28965]
RRP7A	CGI-96	리보솜 RNA 처리 7 동족체 A (사카로미세스 세레비시아) [출처:HGNC Symbol;Acc:24286]
ELAVL3	HUC; PLE21; DKFZp547J036; HUCL; MGC20653	ELAV (미발달 치사, 비정상 시력, 초파리)-유사 3 (Hu 항원 C) [출처:HGNC Symbol;Acc:3314]
C14orf21		푸밀리오 도메인-함유 단백질 C14orf21 [출처:UniProtKB/Swiss-Prot;Acc:Q86U38]
PCBP2	HNRPE2; hnRNP-E2	폴리(rC) 결합 단백질 2 [출처:HGNC Symbol;Acc:8648]

[0096]

ZGPAT	dJ583P15.3; MGC44880; FLJ14972; ZC3HDC9; ZC3H9; GPATC6; GPATCH6; ZIP	징크 핑거, G 패치 도메인을 가진 CCCH-형 [출처:HGNC Symbol;Acc:15948]
HNRNPAB	ABBP1; FLJ40338	이종 핵 리보핵산단백질 A/B [출처:HGNC Symbol;Acc:5034]
NOL8	FLJ20736; Nop132	핵소체 단백질 8 [출처:HGNC Symbol;Acc:23387]
HELZ	KIAA0054; HUMORF5; DHRC	징크 핑거를 가진 나선효소 [출처:HGNC Symbol;Acc:16878]
YTHDF2	HGRG8; NY-REN-2	YTH 도메인 패밀리, 멤버 2 [출처:HGNC Symbol;Acc:31675]
SFMBT2	KIAA1617	4 개 mbt 도메인을 가진 Scm-유사 2 [출처:HGNC Symbol;Acc:20256]
L3MBTL3	KIAA1798	l(3)mbt-유사 3 (초파리) [출처:HGNC Symbol;Acc:23035]
MEX3A		mex-3 동족체 A (예쁜꼬마선충) [출처:HGNC Symbol;Acc:33482]
RBM20		RNA 결합 모티프 단백질 20 [출처:HGNC Symbol;Acc:27424]
DPPA5	Esg1	발생 전분화능(developmental pluripotency) 관련 5 [출처:HGNC Symbol;Acc:19201]
MIR1236		마이크로 RNA 1236 [출처:HGNC Symbol;Acc:33925]
		U6 snRNA-관련 Sm-유사 단백질 LSm2 (snRNP 코어 Sm-유사 단백질 Sm-x5)(소핵 리보핵 단백질 D 동족체)(단백질 G7b) [출처:UniProtKB/Swiss-Prot;Acc:Q9Y333]
PPP1R10		세린/트레오닌-단백질 포스파타아제 1 조절 서브유닛 10 (포스파타아제 1 핵 표적화 서브유닛)(MHC 클래스 I 영역 프롤린-풍부 단백질 CAT53)(FB19 단백질)(114 kDa 의 PP1-결합 단백질)(p99) [출처:UniProtKB/Swiss-Prot;Acc:Q96QC0]
PRR3		프롤린-풍부 단백질 3 (MHC 클래스 I 영역 프롤린-풍부 단백질 CAT56) [출처:UniProtKB/Swiss-Prot;Acc:P79522]

[0097]

PABPN1L	ePABP2	폴리(A) 결합 단백질, 핵 1-유사 (세포질) [출처:HGNC Symbol;Acc:37237]
RNPS1		RNA 결합 단백질 S1, 세린-풍부 도메인 [출처:HGNC Symbol;Acc:10080]
DAZ2	pDP1678; MGC126442	무정자증에서 결실됨 2 [출처:HGNC Symbol;Acc:15964]
ZRSR1		U2 소핵 리보핵산단백질 보조 인자 35 kDa 서브유닛-관련 단백질 1 (U2(RNU2) 소핵 RNA 보조 인자 1-유사 1)(CCCH 형 징크 핑거, RNA-결합 모티프 및 세린/아르기닌 풍부 단백질 1) [출처:UniProtKB/Swiss-Prot;Acc:Q15695]
RBM16	KIAA1116	RNA 결합 모티프 단백질 16 [출처:HGNC Symbol;Acc:20959]
		추정 비특성화 단백질(Putative uncharacterized protein) ENSP00000395989 [출처:UniProtKB/TrEMBL;Acc:C9JFH9]
RBMXL1	KAT3	RNA 결합 모티프 단백질, X-연관-유사 1 [출처:HGNC Symbol;Acc:25073]
CPEB1	FLJ13203; CPEB	세포질 폴리아데닐화 요소 결합 단백질 1 [출처:HGNC Symbol;Acc:21744]
		추정 비특성화 단백질 ENSP00000388079 단편 [출처:UniProtKB/TrEMBL;Acc:C9JXI7]
		추정 비특성화 단백질(Putative uncharacterized protein) ENSP00000383298 [출처:UniProtKB/TrEMBL;Acc:C9JCD7]
RBMY1J		RNA 결합 모티프 단백질, Y-연관, 패밀리 1, 멤버 J [출처:HGNC Symbol;Acc:23917]
SNRPEL1	bA390F4.4	소핵 리보핵산단백질 폴리펩타이드 E-유사 1 [출처:HGNC Symbol;Acc:20733]
		추정 비특성화 단백질 ENSP00000394145 단편 [출처:UniProtKB/TrEMBL;Acc:C9JK19]
MCTS1	MCT-1	악성 T 세포 증폭 서열 1 [출처:HGNC Symbol;Acc:23357]
PABPC4L		폴리(A) 결합 단백질, 세포질 4-유사 [출처:HGNC Symbol;Acc:31955]

[0098]

RBMY1A1	YRRM1; YRRM2	RNA 결합 모티프 단백질, Y-연관, 패밀리 1, 멤버 A1 [출처:HGNC Symbol;Acc:9912]
MKRNP5		마코린 링 핑거 단백질 위유전자 5 [출처:HGNC Symbol;Acc:7115]
RBM14	SIP; SYTIP1; COAA; DKFZp779J0927	RNA 결합 모티프 단백질 14 [출처:HGNC Symbol;Acc:14219]
NSUN6	FLJ23743	NOP2/SUN 도메인 패밀리, 멤버 6 [출처:HGNC Symbol;Acc:23529]
UNKL	ZC3HDC5L; ZC3H5L; FLJ23360	언켐프트 동족체(unkempt homolog) (초파리) [출처:HGNC Symbol;Acc:14184]
RBMY1E		RNA 결합 모티프 단백질, Y-연관, 패밀리 1, 멤버 E [출처:HGNC Symbol;Acc:23916]
HNRNPA1L2	LOC144983	이종 핵 리보핵산단백질 A1-유사 2 [출처:HGNC Symbol;Acc:27067]
RBMY1B		RNA 결합 모티프 단백질, Y-연관, 패밀리 1, 멤버 B [출처:HGNC Symbol;Acc:23914]
RBMY1D		RNA 결합 모티프 단백질, Y-연관, 패밀리 1, 멤버 D [출처:HGNC Symbol;Acc:23915]
RBM12	HRIHFB2091; KIAA0765; SWAN	RNA 결합 모티프 단백질 12 [출처:HGNC Symbol;Acc:9898]
DAZ4		무정자증에서 결실됨 4 [출처:HGNC Symbol;Acc:15966]
DGKQ	DAGK; DAGK7	디아실글리세롤 키나아제, 세타 110kDa [출처:HGNC Symbol;Acc:2856]
RBM17	SPF45; MGC14439	RNA 결합 모티프 단백질 17 [출처:HGNC Symbol;Acc:16944]
MCM3AP	Map80; KIAA0572; GANP	소염색체 관리 복합체 성분 3 관련 단백질 [출처:HGNC Symbol;Acc:6946]
PARN	DAN	폴리(A)-특이적 리보뉴클레아제 (탈아데닐화 뉴클레아제) [출처:HGNC Symbol;Acc:8609]
TLR2	TIL4; CD282	톨-유사 수용체 2 [출처:HGNC Symbol;Acc:11848]
PARP10	FLJ14464	폴리 (ADP-리보스) 중합효소 패밀리, 멤버 10 [출처:HGNC Symbol;Acc:25895]

[0099]

MOV10L1	DJ402G11.8; DKFZp434B0717	Mov10l1, 몰로니 백혈병 바이러스 10-유사 1, 동족체 (마우스) [출처:HGNC Symbol;Acc:7201]
CCAR1	FLJ10590; CARP-1; CARP1	세포 분열 주기 및 세포자연사 조절인자 1 [출처:HGNC Symbol;Acc:24236]
LEMD3	MAN1	LEM 도메인 함유 3[출처:HGNC Symbol;Acc:28887]
UPF3B	RENT3B; UPF3X; HUPF3B	넌센스 전사물의 UPF3 조절인자 동족체 B (효모) [출처:HGNC Symbol;Acc:20439]
RCAN2	ZAKI-4	칼시뉴린의 조절인자 2 [출처:HGNC Symbol;Acc:3041]
SRRT	Asr2; 톱니모양(serrate); ARS2	톱니모양 RNA 효과기 분자 동족체 (애기장대) [출처:HGNC Symbol;Acc:24101]
ZC3H12A	FLJ23231; MCPIP1	징크 핑거 CCCH-형 함유 12A [출처:HGNC Symbol;Acc:26259]
ZC3HAV1L	MGC14289	징크 핑거 CCCH-형, 항바이러스성 1-유사 [출처:HGNC Symbol;Acc:22423]
PARP14	KIAA1268; pART8	폴리 (ADP-리보스) 중합효소 패밀리, 멤버 14 [출처:HGNC Symbol;Acc:29232]
TMEM63A		막관통 단백질 63A [출처:HGNC Symbol;Acc:29118]
SLBP	HBP	줄기-루프 결합 단백질 [출처:HGNC Symbol;Acc:10904]
GTF3A		일반 전사 인자 IIIA [출처:HGNC Symbol;Acc:4662]
NUFIP1		핵 취약 X 정신 지체 단백질 상호작용 단백질 1 [출처:HGNC Symbol;Acc:8057]
SRRM2		세린/아르기닌 반복 매트릭스(repetitive matrix) 2 [출처:HGNC Symbol;Acc:16639]
APTX		아프라탁신 [출처:HGNC Symbol;Acc:15984]
ZFR		징크 핑거 RNA 결합 단백질 [출처:HGNC Symbol;Acc:17277]
UPF1		넌센스 전사물의 UPF1 조절인자 동족체 (효모) [출처:HGNC Symbol;Acc:9962]
KIN		KIN, recA 단백질의 항원 결정인자 동족체 (마우스) [출처:HGNC Symbol;Acc:6327]
ZNF239		징크 핑거 단백질 239 [출처:HGNC Symbol;Acc:13031]

ZNF74		징크 핑거 단백질 74 [출처:HGNC Symbol;Acc:13144]
SNRPC		소핵 리보핵산단백질 폴리펩타이드 C [출처:HGNC Symbol;Acc:11157]
C1orf25		TRM1-유사 단백질 [출처:UniProtKB/Swiss-Prot;Acc:Q7Z2T5]
ZFR2		징크 핑거 RNA 결합 단백질 2 [출처:HGNC Symbol;Acc:29189]
IREB2	ACO3; IRP2; IRP2AD; FLJ23381; IREB2	철-반응 요소 결합 단백질 2 [출처:HGNC Symbol;Acc:6115]
ACO1	IRP1; ACONS; IREB1; IREBP; IREBP1; ACO1	아코니타제 1, 가용성(soluble) [출처:HGNC Symbol;Acc:117]
TROVE2	RO60; SSA2; RORNP;	TROVE 도메인 패밀리, 멤버 2 [출처:HGNC Symbol;Acc:11313]
TEP1	p240; TLP1; TP1; TROVE1; VAULT2	텔로머라아제-관련 단백질 1 [출처:HGNC Symbol;Acc:11726]
GAPDH	GAPD	글리세르알데하이드-3-인산 탈수소효소 [출처:HGNC Symbol;Acc:4141]
ZRANB2		징크 핑거, RAN-결합 도메인 함유 2 [출처:HGNC Symbol;Acc:13058]

[0100]

[0101] RNA 조절 도메인은 표 1 또는 표 2로부터 선택된 단백질 또는 표 1 또는 2의 단백질 목록으로부터 선택된 단백질로부터의 기능성 도메인일 수 있다. 일부 구현예에서, RNA 조절 도메인은 표 1 또는 2의 단백질 목록으로부터 선택된 단백질로부터의 단편을 포함한다. 일부 구현예에서, RNA 조절 도메인은 표 1 또는 2의 단백질 또는 표 1 또는 2로부터의 단백질의 단편에 대해 적어도, 최대, 또는 정확히 100, 99, 98, 97, 96, 95, 94, 93, 92, 91, 90, 89, 88, 87, 86, 85, 84, 83, 82, 81, 80, 79, 78, 77, 76, 75, 74, 73, 72, 71, 70, 69, 68, 67, 66, 65, 64, 63, 62, 61 또는 60%(또는 여기에 속하는 임의의 추론가능한 범위)의 상동성 또는 서열 동일성을 갖는 단백질을 포함한다.

[0102] 일부 구현예에서, RNA 조절 도메인은 표 1 또는 표 2로부터의 단백질의 적어도, 최대, 또는 정확히 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, 1400, 1450, 1500, 1550, 1600, 1650, 1700, 1750, 1800, 1850, 1900, 1950, 또는 2000개의 인접 아미노산(또는 여기에 속하는 임의의 추론가능한 범위)을 포함한다.

[0103] 일부 구현예에서, RNA 조절 도메인은 표 1 또는 2로부터의 단백질의 단편을 포함하고, 여기서 상기 단편은 하기 활성들: 메틸화, 5'-3' 구아닐화, 포스포리보실화, 탈아미노화, 카르바모일화, O/S 교환, 갈락토실화, 글루타밀화, 만노실화, 수소화, 유사우리딘 형성, 카르복시메틸아미노메틸화, 아미노메틸화, 탈카르복시메틸화, 탈수소화, 카르복시메틸화, 히드록실화, 메틸티올화, 3-아미노-3-카르복시프로필화, 탈메틸화, 5'-5' 구아닐화, 탈인산화, 뉴클레아제, 편집, RNA 수송, 번역 활성화, 번역 억제, 단일 가닥 RNA 절단 활성, 이중 가닥 RNA 절단 활성, 및 RNA 결합 활성 중 하나 이상을 갖는다.

[0104] 일부 구현예에서, RNA 조절 도메인은 RNA 헤어핀 결합 단백질의 카르복시-말단에 또는 그 근처에 있다. 일부 구현예에서, RNA 조절 도메인은 RNA 헤어핀 결합 단백질의 아미노-말단에 또는 그 근처에 있다. 일부 구현예에서, RNA 조절 도메인은 펩타이드 결합에 의해 RNA 헤어핀 결합 단백질에 융합된다. 일부 구현예에서, RNA 조절 도메인은 링커 모이어티에 의해 RNA 헤어핀 결합 단백질에 연결된다.

[0105] II.RNA 헤어핀 결합 도메인 및 헤어핀 구조

[0106] 다른 RNA 헤어핀 결합 도메인 및 이들이 결합하는 헤어핀 구조는 당업계에 공지되어 있고, 본 개시 내용의 시스템, 조성물, 융합 단백질, 키트, 벡터 및 방법에서 사용될 수 있다. 예를 들어, 구현예들은 하기 표(표 3)에 따른 RNA 헤어핀 결합 도메인 및 헤어핀 구조를 포함하고, 이 표는 RNA 헤어틴 결합 도메인, 및 이들이 특이적으로 결합하는 헤어핀 구조를 포함하는 단백질을 열거한다:

[0107] 표 3: RNA 헤어핀 결합 도메인

표 3

[0108]

RNA 헤어핀 결합 도메인	헤어핀 구조
MS2 외피 단백질(MCP) (박테리오파지)	AAACAUGAGGAUUACCCAUGU(서열번호: 107)AAACAUGAGGAUCACCCAUGU(서열번호: 108)
λ N (박테리오파지)	GCCCUGAAGAAGGGC(서열번호: 109);GCCCUGAAAAAGGGC(서열번호: 110)
PP7 (슈도모나스 에루지노사)	UAAGGAGUUUAUAUGGAAACCCUUA(서열번호: 111)
철 반응 단백질(IRP)	CAGWGH, 여기서 W는 A 또는 U이고 H는 A, C, 또는 U이고;NNNNNUGCNNNNNCAGUGNNNNNNCNNNNN, 여기서 n은 임의의 뉴클레오타이드이다
Qβ	AAACAUGAGGAUUACCCAUGU(서열번호: 107);AAACAUGAGGAUCACCCAUGU(서열번호: 108); AUGCAUGUCUAAGACAGCAU(서열번호: 114)
GA	AAACAUGAGGAUUACCCAUGU(서열번호: 107);AAACAUGAGGAUCACCCAUGU(서열번호: 108); AAAACAUAAGGAAAACCUAUGUU(서열번호: 115)
소 면역결핍 바이러스(BIV) 전사의 전이활성인자 (tat)	GGCUCGUGUAGCUCAUUAGCUCCGAGCC(서열번호: 112)
U1A (인간)	AUUGCAC;GGAAUCCAUUGCACUCCGGAUUUCACTAG(서열번호: 113); GGCCAGAUCUGAGCCUGGGAGCUCUCUGGCC(서열번호: 1)
SLBP (인간)	CCAAAGGCUCUUCUCAGAGCCACCCA(서열번호: 2)
KU70 (인간)	GGCGUCCCUCCCGAAGCUGCGCGCUCGGUCGAACAGGACGACC(서열번호: 83)
뉴클레오린 (인간)	UCCCGA(서열번호: 84);GGCCGAAAUCCCGAAUGAGGCC(서열번호: 85); GGAUGCCUCCCGAGUGCAUCC(서열 번호: 86).

[0109] 다수의 RNA 헤어핀 결합 도메인 및/또는 RNA 조절 도메인이 상이한 헤어핀 구조에 결합하여 동일한 세포에서 다수의 상이한 RNA를 표적화하는 RNA 헤어틴 결합 도메인을 사용함으로써 다중화 방식으로 사용될 수 있는 것으로 고려된다. 상이한 RNA들은 동일한 방식으로 또는 상이한 방식으로 조절될 수 있다. 예를 들어, 하나의 RNA는 번역 활성화를 이용하여 조절될 수 있는 반면, 제2 RNA는 동일한 세포에서 번역 억제를 이용하여 조절될 수 있다. 따라서, 본 개시 내용의 시스템은 하나의 세포, 조직 또는 유기체에서 적어도 2, 3, 4, 5, 6, 7, 8, 9, 10개 또는 그 이상의 RNA의 조절을 위해 다중화된 방식으로 사용될 수 있다.

[0110] III.핵산

[0111] 특정 구현예에서, 본원에 기재된 단백질, 폴리펩타이드, 조절 도메인, 또는 RNA 표적화 분자를 암호화하는 재조합 핵산이 존재한다.

[0112] 본 출원에서 사용되는 용어 "폴리뉴클레오타이드"는 재조합되었거나 또는 전체 유전체 핵산이 없는 단리된 핵산 분자를 지칭한다. 용어 "폴리뉴클레오타이드"에는 올리고뉴클레오타이드(핵산 100개 이하의 잔기 길이), 예를 들어 플라스미드, 코스미드(cosmid), 파지(phage), 바이러스 등을 포함한 재조합 벡터가 포함된다. 폴리뉴클레오타이드는, 특정 측면에서, 그의 자연 발생 유전자 또는 단백질 암호화 서열로부터 실질적으로 멀리 단리된 조절 서열을 포함한다. 폴리뉴클레오타이드는 단일 가닥(코딩 또는 안티센스) 또는 이중 가닥일 수 있고, RNA, DNA(유전체, cDNA 또는 합성), 이들의 유사체, 또는 이들의 조합일 수 있다. 추가의 코딩 또는 비-코딩 서열이 폴리뉴클레오타이드 내에 존재할 수 있지만, 반드시 그럴 필요는 없다.

[0113] 이와 관련하여, 용어 "유전자", "폴리뉴클레오타이드" 또는 "핵산"은 단백질, 폴리펩타이드 또는 펩타이드(적절한 전사, 번역 후 변형 또는 국소화에 필요한 임의의 서열 포함)를 암호화하는 핵산을 지칭하는데 사용된다. 관련 기술분야의 통상의 기술자에 의해 이해되는 바와 같이, 이 용어는 단백질, 폴리펩타이드, 도메인, 펩타이드, 융합 단백질, 및 돌연변이체를 발현하거나 또는 발현하도록 적응될 수 있는 유전체 서열, 발현 카세트, cDNA 서열, 및 보다 작은 조작된 핵산 단편을 포괄한다. 폴리펩타이드의 전부 또는 일부를 암호화하는 핵산은 이러한 폴리펩타이드의 전부 또는 일부를 암호화하는 인접 핵산 서열을 함유할 수 있다. 또한, 특정 폴리펩타이드는 약간 상이한 핵산 서열을 갖는 변이를 함유하는 핵산에 의해 암호화될 수 있지만, 그럼에도 불구하고 동일하거나 실질적으로 유사한 단백질을 암호화할 수 있는 것으로 고려된다(상기 참조).

[0114] 특정 구현예에서, 상호작용 성분이 상호작용할 때 중합효소 도메인으로부터의 중합효소 활성에 의존적인 유전자 전사를 구동하는 폴리펩타이드(예를 들어, 중합효소, RNA 중합효소, 폴리펩타이드인 하나 이상의 절단된 중합효소 도메인 또는 상호작용 성분)를 암호화하는 핵산 서열이 혼입된 단리된 핵산 단편 및 재조합 벡터가 존재한다. 용어 "재조합"은 폴리펩타이드 또는 특정 폴리펩타이드의 이름과 함께 사용될 수 있고, 이는 일반적으로 시험관 내에서 조작된 또는 이러한 분자의 복제 생성물인 핵산 분자로부터 생성된 폴리펩타이드를 지칭

한다.

[0115] 핵산 단편은, 코딩 서열 자체의 길이에 관계없이, 다른 핵산 서열, 예컨대 프로모터, 폴리아데닐화 신호, 추가의 제한 효소 부위, 다중 클로닝 부위, 다른 코딩 분절 등과 조합되어, 그의 전체 길이가 상당히 달라질 수 있다. 따라서, 거의 임의의 길이의 핵산 단편이 사용될 수 있고, 총 길이는 바람직하게는 의도된 재조합 핵산 계획에서의 제조 및 사용의 용이성에 의해 제한되는 것으로 고려된다. 일부 경우에, 핵산 서열은, 예를 들어 폴리펩타이드의 정제, 수송, 분비, 번역후 변형을 가능하게 하기 위해, 또는 표적화 또는 효능과 같은 치료 이익을 위해 추가의 이종 코딩 서열을 갖는 폴리펩타이드 서열을 암호화할 수 있다. 상기 논의된 바와 같이, 꼬리표 또는 다른 이종 폴리펩타이드가 변형된 폴리펩타이드-암호화 서열에 첨가될 수 있고, 여기서 "이종(heterologous)"은 변형된 폴리펩타이드와 동일하지 않은 폴리펩타이드를 지칭한다.

[0116] 특정 구현예에서, 본원에 개시된 서열에 대해 실질적인 동일성을 갖는 폴리뉴클레오타이드 변이체들이 존재하며; 이들은 본원에 기재된 방법 (예를 들어, 표준 매개변수들을 사용한 BLAST 분석)을 사용하여 본원에 제공된 폴리뉴클레오타이드 서열과 비교하여, 하기 수치 사이의 모든 값 및 범위를 포함하여, 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 또는 그 이상의 서열 동일성을 포함한다. 특정 측면에서, 단리된 폴리뉴클레오타이드는 서열의 전체 길이에 걸쳐 본원에 기재된 아미노산 서열과 적어도 90%, 바람직하게는 95% 이상의 동일성을 갖는 폴리펩타이드를 암호화하는 뉴클레오타이드 서열; 또는 상기 단리된 폴리뉴클레오타이드에 상보적인 뉴클레오타이드 서열을 포함할 것이다.

[0117] A.벡터

[0118] 폴리펩타이드는 핵산 분자에 의해 암호화될 수 있다. 핵산 분자는 핵산 벡터의 형태일 수 있다. 용어 "벡터"는 이종 핵산 서열이 복제되고 발현될 수 있는 세포 내로의 도입을 위해 삽입될 수 있는 운반체 핵산 분자를 지칭하는 데 사용된다. 핵산 서열은 "이종성"일 수 있고, 이는 벡터가 도입되는 세포 또는 도입되는 핵산에 대해 외래인 맥락에 있음을 의미하며, 이는 세포 또는 핵산 내의 서열에 상동성이지만 숙주 세포 또는 핵산의 내에서 통상적으로 발견되지 않는 위치에 있는 서열을 포함한다. 벡터는 DNA, RNA, 플라스미드, 코스미드, 바이러스(박테리오파지, 동물 바이러스, 및 식물 바이러스), 및 인공 염색체(예를 들어, YAC)를 포함한다. 당업자는 표준 재조합 기술을 통해 벡터를 작제하기 위한 준비를 잘 갖출 수 있을 것이다(예를 들어, 문헌[Sambrook et al., 2001; Ausubel et al., 1996], 둘 다 본원에 참조로 포함됨). 벡터는 중합효소, RNA 중합효소, 하나 이상의 절단된 중합효소 도메인 또는 하나 이상의 절단된 RNA 중합효소 도메인에 융합, 부착 또는 연결되는 상호작용 성분을 생성하기 위해 숙주 세포에서 사용될 수 있다.

[0119] 용어 "발현 벡터"는 전사될 수 있는 유전자 생성물의 적어도 일부에 대한 핵산 서열 코딩을 함유하는 벡터를 지칭한다. 일부 경우에, RNA 분자는 이어서 단백질, 폴리펩타이드 또는 펩타이드로 번역된다. 발현 벡터는 특정 숙주 유기체에서 작동가능하게 연결된 코딩 서열의 전사 및 가능하게는 번역에 필요한 핵산 서열을 지칭하는 다양한 "제어 서열(control sequence)"을 함유할 수 있다. 전사 및 번역을 제어하는 제어 서열에 더하여, 벡터 및 발현 벡터는 또한 다른 기능을 하고 본원에 기재된 핵산 서열을 함유할 수 있다.

[0120] B.세포

[0121] 본 개시 내용은 구체적으로 원핵 세포, 진핵 세포, 조직, 기관 또는 유기체, 보다 구체적으로 포유동물 세포, 조직, 기관 또는 유기체에서 관심 표적 RNA를 변형시키는 방법을 제공한다. 표적 RNA는 세포 내의 핵산 분자에 포함될 수 있다. 일부 구현예에서, 표적 RNA는 포유동물 세포 또는 식물 세포와 같은 진핵 세포에 있다. 포유동물 세포는 인간, 비인간 영장류, 소과, 돼지과, 설치류 또는 마우스 세포일 수 있다. 세포는 가금류, 어류 또는 새우와 같은 비포유류 진핵 세포일 수 있다. 식물 세포는 카사바, 옥수수, 수수, 밀 또는 벼와 같은 작물 식물의 것일 수 있다. 식물 세포는 또한 조류(algae), 나무 또는 채소의 것일 수 있다. 본 개시 내용의 방법, 시스템, 및 조성물에 의해 세포에서 유도된 RNA의 조절은 세포 및 세포의 자손이 항체, 녹말, 알코올 또는 기타 원하는 세포 산물과 같은 생물학적 생성물의 개선된 생산을 위해 변경되도록 될 수 있다. 세포에서 유도된 RNA의 조절은 세포 및 세포의 자손이 생산된 생물학적 생성물을 변화시키는 변경을 포함하도록 될 수 있다.

[0122] 포유동물 세포는 인간 또는 비인간 포유동물, 예를 들어 영장류, 소과, 양과, 돼지과, 개과, 설치류, 토끼과, 예컨대 원숭이, 암소, 양, 돼지, 개, 토끼, 쥐 또는 마우스 세포일 수 있다. 세포는 가금류 조류(예: 닭), 척추동물 어류(예: 연어) 또는 조개류(예: 굴, 조개, 랍스터, 새우) 세포와 같은 비포유류 진핵 세포일 수 있다. 세포는 또한 식물 세포일 수 있다. 식물 세포는 단자엽 또는 쌍자엽 식물 또는 카사바, 옥수수, 수수, 대두, 밀, 귀리 또는 벼와 같은 작물 또는 곡물 식물의 것일 수 있다. 식물 세포는 또한 조류(algae), 나무 또는 생산

식물, 과일 또는 채소(예를 들어, 감귤류 나무, 예를 들어 오렌지, 그레이프프루트 또는 레몬 나무와 같은 나무; 복숭아 또는 천도 복숭아 나무; 사과 또는 배 나무; 아몬드, 호두 또는 피스타치오 나무와 같은 견과 나무; 가지속 식물; 배추속 식물; 상추속 식물; 시금치속 식물; 고추속 식물; 목화, 담배, 아스파라거스, 당근, 양배추, 브로콜리, 콜리플라워, 토마토, 가지, 후추, 상추, 시금치, 딸기, 블루베리, 라즈베리, 블랙베리, 포도, 커피, 코코아 등)의 것일 수 있다.

[0123] 본원에 사용되는 용어 "세포", "세포주" 및 "세포 배양물"은 상호교환적으로 사용될 수 있다. 이 모든 용어에는 모든 후속 세대인 자손도 포함된다. 모든 자손은 의도적이거나 의도하지 않은 돌연변이로 인해 동일하지 않을 수 있음을 이해해야 한다. 이종 핵산 서열을 발현하는 맥락에서, "숙주 세포"는 원핵 또는 진핵 세포를 지칭하며, 벡터를 복제하거나 벡터에 의해 암호화되는 이종 유전자를 발현할 수 있는 임의의 형질전환 가능 유기체를 포함한다. 숙주 세포는 벡터나 바이러스의 수용자(recipient)로 사용될 수 있고 사용되어 왔다. 숙주 세포는 재조합 단백질-암호화 서열과 같은 외인성 핵산이 숙주 세포 내로 전달되거나 도입되는 과정을 지칭하는 "형질주입" 또는 "형질전환"될 수 있다. 형질전환된 세포는 1차 대상체 세포 및 그의 자손을 포함한다.

[0124] 일부 벡터는 원핵 및 진핵 세포 모두에서 복제 및/또는 발현되도록 하는 제어 서열을 사용할 수 있다. 당업자는 상기 기재된 모든 숙주 세포를 인큐베이션하여 이들을 유지하고 벡터의 복제를 허용하는 조건을 추가로 이해할 것이다. 또한, 벡터 및 이들의 동족 폴리펩타이드, 단백질, 또는 펩타이드에 의해 암호화된 핵산의 생성 뿐만 아니라 벡터의 대규모 생성을 가능하게 하는 기술 및 조건은 공지되어 있으며 당업자가 이해 가능하다.

[0125] **C.발현 시스템**

[0126] 상기 논의된 조성물의 적어도 일부 또는 전부를 포함하는 수많은 발현 시스템이 존재한다. 원핵생물 및/또는 진핵생물 기반 시스템은 핵산 서열, 또는 그의 동족 폴리펩타이드, 단백질 및 펩타이드를 생성하기 위한 구현예와 함께 사용하기 위해 채용될 수 있다. 예를 들어, 본 개시 내용의 벡터, 융합 단백질, RNA 헤어핀 결합 단백질, RNA 표적화 분자, RNA 조절 도메인, 및 부속 단백질은 유도성 또는 구성적 발현 시스템과 같은 발현 시스템을 이용할 수 있다. 상기와 같은 많은 시스템이 상업적으로 널리 이용 가능하다.

[0127] 곤충 세포/배큘로바이러스(baculovirus) 시스템은 미국 특허 제5,871,986호, 제4,879,236호에 기술된 것과 같은 이종 핵산 단편의 높은 수준의 단백질 발현을 생성할 수 있으며, 두 특허 모두 본원에 참고로 포함되며, 예를 들어 INVITROGEN®의 MAXBAC® 2.0 및 CLONTECH®의 BACPACK™ BACULOVIRUS EXPRESSION SYSTEM이라는 이름으로 구입할 수 있다.

[0128] 개시된 발현 시스템에 더하여, 발현 시스템의 다른 예는 합성 엑디손-유도성 수용체를 포함하는 STRATAGENE®의 COMPLETE CONTROL 유도성 포유동물 발현 시스템, 또는 대장균 발현 시스템인 동사의 pET 발현 시스템을 포함한다. 유도성 발현 시스템의 또 다른 예는 전장(full-length) CMV 프로모터를 사용하는 유도성 포유동물 발현 시스템인 T-REX™(테트라사이클린 조절 발현) 시스템을 보유하는 INVITROGEN®에서 구입할 수 있다. INVITROGEN®은 또한 메틸올 자화 효모 *피치아 메탄올리카(Pichia methanolica)*에서 재조합 단백질의 높은 수준의 생산을 위해 설계된 *피치아 메탄올리카* 발현 시스템이라고 하는 효모 발현 시스템을 제공한다. 당업자는 핵산 서열 또는 그의 동족 폴리펩타이드, 단백질 또는 펩타이드를 생성하기 위해 발현 작제물과 같은 벡터를 발현하는 방법을 알고 있을 것이다.

[0129] **D.핵산과 폴리펩타이드의 접합**

[0130] 본 개시 내용의 구현예는 핵산을 폴리펩타이드에 접합하는 것에 관한 것이다. 폴리펩타이드에 핵산을 접합하는 방법은 당업계에 공지되어 있고 하기에 기술된 것을 포함한다. 본 개시 내용의 구현예는 핵산-폴리펩타이드 분자 및 핵산이 본원에 기재된 방법에 의해 폴리펩타이드에 접합된 분자 자체의 제조 방법에 관한 것이다. 상기와 같은 예에는 클릭 화학(click chemistry)이 포함된다. "클릭 화학"이라고도 알려진 "클릭 반응(click reaction)"은 1,2,3-트리아졸을 생성하기 위한 아지드 및 알킨의 휘스겐(Huisgen) 1,3-쌍극자 고리화 첨가의 단계적 변이체를 설명하는 데 자주 사용되는 이름이다. 이 반응은 대기 조건에서 또는 약한 마이크로파 조사하에, 일반적으로 Cu(I) 촉매의 존재 하에, Cu(I) 염의 촉매량에 의해 매개될 때 1,4-2기치환된 트리아졸에 대한 배타적 위치 선택성을 갖고 수행된다[V. Rostovtsev, L. G. Green, V. V. Fokin, K. B. Sharpless, Angew. Chem. Int. Ed. 2002, 41, 2596; H. C. Kolb, M. Finn, K. B. Sharpless, Angew Chem., Int. Ed. 2001, 40, 2004].

[0131] 다른 접합 방법에서, 돌연변이체 형태의 인간 DNA 복구 단백질 O6-알킬구아닌-DNA 알킬전달효소는 O6-벤질구아닌(BG), 및 벤질기에 연결된 큰 모이어티를 운반하는 유도체와 빠르게 특이적으로 반응한다. 구아닌을 이탈기

(leaving group)로 사용하면 벤질 모이어티가 효소의 활성 부위에 있는 시스테인에 공유 결합된다. 효소는 또한 유사한 방식으로 O6-벤질시토신(BC)에 특이적이 되도록 돌연변이되었다. 이러한 효소 도메인들(약 20kDa)은 각각 SNAP 및 CLIP 태그로 상업적으로 이용 가능하다.

[0132] 추가 접합 방법은 Halo 태그를 사용한다. Halo 태그는 진핵생물에 직교하는 화학 반응, 즉 할로알칸 리간드의 탈할로겐화를 이용하며, 따라서 태그의 고도로 특이적인 공유결합 라벨링으로 이어지고, 따라서 살아있는 세포 및 고정된 세포 모두에서 단백질로 이어진다.

[0133] **E.핵산 변형**

[0134] 본원에 기재된 RNA 표적화 분자 및 기타 핵산과 같은 본 개시 내용의 올리고뉴클레오타이드는 핵산의 안정성을 증가시키는 변형을 가질 수 있다. 일부 구현예에서, RNA 표적화 분자는 올리고뉴클레오타이드 유사체이다. 용어 "올리고뉴클레오타이드 유사체"는 올리고뉴클레오타이드와 유사하게 기능하지만 비천연 발생 부분을 갖는 화합물을 지칭한다. 올리고뉴클레오타이드 유사체는 변경된 당(sugar) 모이어티, 변경된 염기 모이어티 또는 변경된 당간(inter-sugar) 연결을 가질 수 있다. 용어 "올리고머"는 올리고뉴클레오타이드, 올리고뉴클레오타이드 유사체 또는 올리고뉴클레오사이드(oligonucleoside)를 포함하는 것으로 의도된다. 따라서, "올리고머"에 대해 언급할 때 천연 인산디에스테르 결합 또는 4개의 원자 링커를 포함하는 다른 연결을 통해 연결된 일련의 뉴클레오사이드 또는 뉴클레오사이드 유사체를 참조한다. 연결이 일반적으로 한 뉴클레오사이드의 3' 탄소에서 두 번째 뉴클레오사이드의 5' 탄소로 연결되지만, 용어 "올리고머"는 2'-5' 연결과 같은 다른 연결도 포함할 수 있다.

[0135] 올리고뉴클레오타이드 유사체는 또한 다른 변형, 구체적으로 뉴클레아제 내성을 증가시키고, 결합 친화도를 향상하고, 및/또는 결합 특이성을 향상하는 변형을 포함할 수 있다. 예를 들어, 뉴클레오사이드 또는 뉴클레오타이드의 당 부분이 탄소고리 모이어티로 대체되면 이는 더 이상 당이 아니다. 더욱이, 당간 인산디에스테르 결합에 대한 상기 치환과 같은 다른 치환이 이루어질 때, 생성된 물질은 더 이상 진정한 핵산 종이 아니다. 이러한 모든 화합물은 유사체로 간주된다. 본 명세서 전반에 걸쳐, 핵산 종의 당 부분에 대한 언급은 진정한 당 또는 야생형 핵산의 당의 구조적 위치를 차지하는 종을 언급하는 것으로 이해되어야 한다. 더욱이, 당간 연결에 대한 언급은 야생형 핵산의 방식으로 당 또는 당 유사체 부분을 연결하는 역할을 하는 모이어티를 포함하는 것으로 간주되어야 한다.

[0136] 본 개시 내용은 변형된 올리고뉴클레오타이드, 즉, 올리고뉴클레오타이드 유사체 또는 올리고뉴클레오사이드, 및 상기 변형을 수행하는 방법에 관한 것이다. 이러한 변형된 올리고뉴클레오타이드 및 올리고뉴클레오타이드 유사체는 자연 발생 대응물에 비해 증가된 화학적 및/또는 효소적 안정성을 나타낼 수 있다. 세포외 및 세포내 뉴클레아제는 일반적으로 백본 변형된 화합물을 인식하지 못하므로 그에 결합하지 않는다. 양성화된 산 형태로 존재할 때 음전하를 띤 백본이 결여되면 세포 침투가 촉진될 수 있다.

[0137] 변형된 뉴클레오사이드 연결은 천연 발생 인산디에스테르-5'-메틸렌 연결을 4개의 원자 연결기로 대체하여 생성된 화합물에 뉴클레아제 내성 및 향상된 세포 흡수력을 부여하기 위한 것이다. 바람직한 연결 구조는 CH2 --RA --NR1 CH2, CH2 --NR1 --RA --CH2, RA --NR1 --CH2 --CH2, CH2 --CH2 --NR1 --RA, CH2 --CH2 --RA --NR1, 또는 NR1 --RA --CH2 --CH2이고, 여기서 RA는 O 또는 NR2이다.

[0138] 변형은 DNA 합성기 분야의 숙련가에게 일반적으로 알려진 방법을 사용하여 DNA 합성기와 함께 사용하거나 수동으로 조작할 수 있는 고체 지지체를 사용하여 달성할 수 있다. 일반적으로, 절차는 선택된 서열에서 서로 인접할 2개의 뉴클레오사이드의 당 모이어티를 기능화하는 것을 포함한다. 5'에서 3'으로의 의미에서, "상류부위(upstream)" 신톤은 말단 3' 부위에서 변형되고, "하류부위(downstream)" 신톤은 말단 5' 부위에서 변형된다.

[0139] 히드라진, 히드록실아르닌(hydroxylarnine) 및 기타 연결기에 의해 연결된 올리고뉴클레오사이드는 5'-히드록실에서 디메톡시트리틸기에 의해 보호될 수 있고 3'-히드록실에서 시아노에틸디이소프로필-아인산염 모이어티와 결합하기 위해 활성화될 수 있다. 이러한 화합물은 표준, 고체상, 자동화된 DNA 합성 기술에 의해 원하는 서열에 삽입할 수 있다. 가장 널리 사용되는 공정 중 하나는 포스포라미디트 기술이다. 균일한 백본 연결을 함유하는 올리고뉴클레오타이드는 Applied Biosystems Inc. 380B 및 394 및 Milligen/Biosearch 7500 및 8800s와 같은 표준 핵산 합성 기계 및 CPG-고체 지지체를 사용하여 합성할 수 있다. 초기 뉴클레오타이드(3'-말단의 1번)는 조절된 공극 유리와 같은 고체 지지체에 부착된다. 서열 특이적 순서로, 각각의 새로운 뉴클레오타이드는 수동 조작에 의해 또는 자동화된 합성기 시스템에 의해 부착된다.

[0140] 유리 아미노기는 예를 들어 아세트산에서 아세톤 및 나트륨 시아노보로 수소화물(sodium cyanoboro hydride)로 알킬화될 수 있다. 알킬화 단계는 거대분자에 다른 유용한 기능 분자를 도입하는 데 사용할 수 있다. 이러한 유용한 기능적 분자는 리포터 분자, RNA 절단기(cleaving group), 올리고뉴클레오타이드의 약동학적 특성을 개선하기 위한 기(group), 및 올리고뉴클레오타이드의 약력학적 특성을 개선하기 위한 기(group)를 포함하지만 이에 제한되지는 않는다. 이러한 분자는 백본 연결에서 질소 원자에 대한 부착을 통해 거대분자에 부착되거나 접합될 수 있다. 대안적으로, 이러한 분자는 하나 이상의 뉴클레오타이드의 당 모이어티의 히드록실기로부터 연장되는 펜던트기에 부착될 수 있다. 이러한 다른 유용한 작용기의 예는 WO1993007883(본원에 참고로 포함됨) 및 상기 언급된 다른 특허 출원에 제공된다.

[0141] 고체 지지체에는 조절된 공극 유리(CPG), 옥살릴 조절된 공극 유리[53], TentaGel Support－아미노폴리에틸렌글리콜 유도된 지지체[54] 또는 Poros－폴리스티렌/디비닐벤젠의 공중합체를 포함하여 폴리뉴클레오타이드 합성에 대해 당업계에 알려진 모든 것이 포함될 수 있다. 뉴클레오타이드 및 올리고뉴클레오타이드의 부착 및 절단은 표준 절차를 통해 수행할 수 있다[55]. 본원에 사용된 바와 같이, 고체 지지체라는 용어는 성장하는 올리고뉴클레오사이드를 CPG와 같은 정지상(stationary phase)에 결합하는 데 사용되는 임의의 링커(예를 들어, 장쇄 알킬 아민 및 숙시닐 잔기)를 더 포함한다.

[0142] **1.고정된 뉴클레오타이드(Locked Nucleotide)**

[0143] 일부 구현예에서, RNA 표적화 분자와 같은 본 개시 내용의 핵산은 고정된 핵산(locked nucleic acid)을 포함한다. 접근 불가능한 RNA라고도 하는 고정된 핵산(LNA 또는 Ln)은 변형된 RNA 뉴클레오타이드이다. LNA 뉴클레오타이드의 리보스 모이어티는 2’ 산소와 4’ 탄소를 연결하는 여분의 가교로 변형된다. 가교(bridge)는 A형 이중체(duplex)에서 종종 발견되는 3’-엔도(북쪽) 형태로 리보스를 “고정(lock)”한다. LNA 뉴클레오타이드는 원하는 경우 언제든지 올리고뉴클레오타이드의 DNA 또는 RNA 잔기와 혼합될 수 있으며 Watson-Crick 염기쌍 규칙에 따라 DNA 또는 RNA와 혼성화될 수 있다. 이러한 올리고머는 화학적으로 합성되며 상업적으로 이용 가능하다. 고정된 리보스 형태는 염기 중첩(base stacking) 및 백본 사전 구성을 향상시킨다. 이것은 올리고뉴클레오타이드의 혼성화 특성(용융 온도)을 상당히 증가시킨다.

[0144] **2.에틸렌 가교된 뉴클레오타이드**

[0145] 일부 구현예에서, RNA 표적화 분자와 같은 본 개시 내용의 핵산은 하나 이상의 에틸렌 가교된 뉴클레오타이드(ethylene bridged nucleotide)를 포함한다. 에틸렌-가교된 핵산(ENA 또는 En)은 2’-O, 4’C 에틸렌 연결이 있는 변형된 뉴클레오타이드이다. 고정된 뉴클레오타이드와 마찬가지로 이 뉴클레오타이드는 당질 퍼커링(sugar puckering)을 RNA의 N 형태로 제한한다.

[0146] **3.펩타이드 핵산**

[0147] 일부 구현예에서, RNA 표적화 분자와 같은 본 개시 내용의 핵산은 하나 이상의 펩타이드 핵산을 포함한다. 펩타이드 핵산(PNA 또는 Pn)은 DNA의 거동을 모방하고 상보성 핵산 가닥에 결합한다. 본원에서 사용되는 용어 “펩타이드”는 또한 펩타이드 핵산을 지칭할 수 있다. PNA는 DNA 또는 RNA와 유사한 인공적으로 합성된 중합체이다. DNA와 RNA는 각각 디옥시리보스와 리보스 당 백본을 가지고 있는 반면, PNA의 백본은 펩타이드 결합으로 연결된 반복되는 N-(2-아미노에틸)-글리신 단위로 구성된다. 다양한 퓨린 및 피리미딘 염기는 메틸렌 가교(-CH2-) 및 카르보닐기(-(C=O)-)에 의해 백본에 연결된다. PNA는 첫 번째(왼쪽) 위치에 N-말단이 있고 마지막(오른쪽) 위치에 C-말단이 있는 펩타이드처럼 묘사된다.

[0148] PNA의 백본에는 전하를 띤 인산기가 없기 때문에 PNA/DNA 가닥 사이의 결합은 정전기적 반발력의 결여로 인해 DNA/DNA 가닥 사이보다 더 강력하다. PNA는 뉴클레아제나 프로테아제에 의해 쉽게 인식되지 않아 효소에 의한 분해에 저항성이 있다. PNA는 또한 넓은 pH 범위에서 안정적이다. 일부 측면에서, 본원에 기재된 PNA는 다른 PNA에 비해 개선된 세포액 전달력을 갖는다.

[0149] **4.5’(E)-비닐-포스포네이트(VP) 변형**

[0150] 일부 구현예에서, RNA 표적화 분자와 같은 본 개시 내용의 핵산은 하나 이상의 5’(E)-비닐-포스포네이트(VP) 변형을 포함한다. 5’(E)-비닐-포스포네이트 변형(대사적으로 안정된 인산염 모방체)은 올리고뉴클레오타이드의 대사 안정성과 효능을 향상시키는 것으로 보고되었다.

[0151] **5.모르폴리노(Morpholino)**

[0152] 일부 구현예에서, RNA 표적화 분자와 같은 본 개시 내용의 핵산은 모르폴리노(morpholino)를 포함한다. 모르폴

리노는 천연 핵산 구조를 재설계한 생성물인 합성 분자이다. 일반적으로 길이가 25개 염기인 이들은 표준 핵산 염기쌍에 의해 RNA 또는 단일 가닥 DNA의 상보성 서열에 결합한다. 구조면에서 모르폴리노와 DNA의 차이점은 모르폴리노에는 표준 핵산 염기가 있지만 이러한 염기는 인산염 대신 포스포로디아미데이트 기를 통해 연결된 메틸렌모르폴린 고리에 결합되어 있다는 것이다. 그림은 거기에 묘사된 두 가닥의 구조를 비교하는데, 하나는 RNA이고 다른 하나는 모르폴리노이다. 음이온성 인산염을 전하를 띠지 않는 포스포로디아미데이트 기로 대체하면 일반적인 생리학적 pH 범위에서 이온화가 제거되므로 유기체나 세포의 모르폴리노는 전하를 띠지 않는 분자이다. 모르폴리노의 전체 백본은 이러한 변형된 서브유닛들로 만들어진다.

[0153] **IV.전달 매개체**

[0154] 본 개시 내용은 대략 균일한 분포를 제공하고 제어가능한 방출 속도를 갖는 핵산과 양립가능한 여러 전달 시스템을 고려한다. 핵산 전달 시스템을 만드는 데 유용한 다양한 상이한 매질이 아래에서 설명된다. 임의의 하나의 매질 또는 운반체가 본 발명을 제한하는 것으로 의도되지 않는다. 모든 매질 또는 운반체는 다른 매질 또는 운반체와 결합될 수 있는데; 예를 들어, 일 구현예에서 화합물에 부착된 중합체 미세입자 운반체는 겔 매질과 결합될 수 있다.

[0155] 본 개시 내용에 의해 고려되는 운반체 또는 매질은 젤라틴, 콜라겐, 셀룰로오스 에스테르, 덱스트란 황산, 펜토산 폴리설페이트, 키틴, 당류, 알부민, 피브린 실란트, 합성 폴리비닐 피롤리돈, 폴리에틸렌 옥사이드, 폴리프로필렌 옥사이드, 폴리에틸렌 옥사이드와 폴리프로필렌 옥사이드의 블록 중합체, 폴리에틸렌 글리콜, 아크릴레이트, 아크릴아미드, 메타크릴레이트(2-히드록시에틸 메타크릴레이트를 포함하나 이에 제한되지 않음), 폴리(오르토 에스테르), 시아노아크릴레이트, 젤라틴-레조르시놀-알데히드 유형 생체접착물질, 폴리아크릴산 및 공중합체 및 이들의 블록 공중합체를 포함하는 군(group)으로부터 선택된 물질을 포함한다.

[0156] **A.미세입자**

[0157] 본 개시 내용의 일부 구현예는 미세입자를 포함하는 전달 시스템을 고려한다. 바람직하게는, 미세입자는 리포솜, 나노입자, 마이크로스피어, 나노스피어, 마이크로캡슐, 및 나노캡슐을 포함한다. 바람직하게는, 본 발명에 의해 고려되는 일부 미세입자는 폴리(락타이드-코-글리콜라이드), 폴리-글리콜산 및 폴리-락트산을 포함하나 이에 제한되지 않는 지방족 폴리에스테르, 히알루론산, 변형 다당류, 키토산, 셀룰로스, 덱스트란, 폴리우레탄, 폴리아크릴산, 슈도-폴리(아미노산), 폴리하이드록시부티레이트 관련 공중합체, 폴리무수물, 폴리메틸메타크릴레이트, 폴리(에틸렌 옥사이드), 레시틴 및 인지질을 포함한다.

[0158] **B.리포솜**

[0159] 본 개시 내용의 일 구현예는 본원에 기재된 바와 같은 핵산 접합체, 폴리펩타이드 및 융합 단백질을 부착 및 방출할 수 있는 리포솜을 고려한다. 리포솜은 인지질과 같은 양쪽 친매성 분자로 만들어진 수성 코어를 둘러싼 미세한 구형 지질 이중층이다. 예를 들어, 리포솜은 인지질 미셀의 소수성 꼬리 사이에 핵산을 포획할 수 있다. 수용성 제제는 코어에 포획될 수 있고 지용성 제제는 쉘-유사 이중층에 용해될 수 있다. 리포솜은 계면활성제나 다른 유화제를 사용하지 않고 매질에서 수용성 및 수불용성 화학물질을 함께 사용할 수 있다는 점에서 특별한 특징이 있다. 리포솜은 수성 매질에서 인지질을 강제로 혼합하여 자발적으로 형성할 수 있다. 수용성 화합물은 인지질을 수화시킬 수 있는 수용액에 용해된다. 따라서, 리포솜이 형성되면 이러한 화합물은 수성 리포솜 중심 내에 포획된다. 인지질 막인 리포솜 벽은 오일과 같은 지용성 물질을 보유한다. 리포솜은 통합된 화합물의 제어 방출을 제공한다. 또한, 리포솜은 약동학적 반감기를 증가시키기 위해 폴리에틸렌 글리콜과 같은 수용성 중합체로 코팅될 수 있다. 본 발명의 일 구현예는 리포솜 생산을 정제하여 특별히 설계된 구조적 특징을 갖는 안정한 단일박막(단일층) 리포솜을 생성하기 위한 초고전단 기술을 고려한다. 리포솜의 이러한 독특한 특성으로 인해 일반적으로 섞이지 않는 화합물의 동시 저장과 제어 방출 기능이 가능하다.

[0160] 일부 구현예에서, 본 개시 내용은 양이온성 및 음이온성 리포솜 뿐만 아니라 중성 지질을 갖는 리포솜을 고려한다. 바람직하게는, 양이온성 리포솜은 물질과 지방산 리포솜 성분을 혼합하고 이들이 전하-회합하도록 함으로써 음으로 하전된 물질을 포함한다. 분명히, 양이온성 또는 음이온성 리포솜의 선택은 최종 리포솜 혼합물의 원하는 pH에 따라 달라진다. 양이온성 리포솜의 예에는 리포펙틴(lipofectin), 리포펙타민(lipofectamine) 및 리포펙타스(lipofectace)가 포함된다.

[0161] 본 개시 내용의 일 구현예는 본원에 기재된 적어도 하나의 분자의 제어 방출을 제공하는 리포솜을 포함하는 전달 시스템을 고려한다. 바람직하게는, 제어 방출이 가능한 리포솜은: i) 생분해성이며 독성이 없고; ii) 수용성 및 지용성 화합물을 모두 운반하고; iii) 다루기 힘든 화합물을 가용화하고; iv) 화합물 산화를 방지하고; v)

단백질 안정화를 촉진하고; vi) 수화를 제어하고; vii) 지방산 사슬 길이, 지방산 지질 조성물, 포화 및 불포화 지방산의 상대적인 양, 물리적 배열과 같은 그러나 이에 제한되지 않는 이중층 조성물의 변화에 의해 화합물 방출을 제어하고; viii) 용매 의존성을 갖고; ix) pH 의존성을 갖고 x) 온도 의존성을 갖는다.

[0162] 리포솜의 조성은 크게 두 가지로 분류된다. 기존의 리포솜은 일반적으로 당지질을 포함하거나 포함하지 않을 수 있는 합성 동일쇄 인지질을 포함할 수 있는 안정화된 천연 레시틴(PC)의 혼합물이다. 특수 리포솜은: i) 양극성 지방산; ii) 조직 표적화 요법을 위한 항체 부착 능력; iii) 지단백질 및 탄수화물과 같은(그러나 이에 한정되지 않는) 물질로 코팅됨; iv) 다중 캡슐화 및 v) 에멀션 호환성을 포함할 수 있다.

[0163] 리포솜은 초음파 처리 및 진동과 같은(그러나 이에 한정되지 않는) 방법으로 실험실에서 쉽게 만들 수 있다. 대안적으로, 화합물 전달 리포솜은 상업적으로 이용 가능하다. 예를 들어, Collaborative Laboratories, Inc.는 특정 전달 요구 사항을 위해 맞춤 설계된 리포솜을 제조하는 것으로 알려져 있다.

[0164] **C.마이크로스피어, 미세입자 및 마이크로캡슐**

[0165] 마이크로스피어 및 마이크로캡슐은 일반적으로 균일한 분포를 유지하고 안정하게 제어된 화합물 방출을 제공하고 생산 및 분배하기에 경제적인 능력으로 인해 유용하다. 바람직하게는, 관련된 전달 겔 또는 화합물이 함침된(compound-impregnated) 겔은 투명하거나, 대안적으로 상기 겔은 의료진이 쉽게 볼 수 있도록 착색되어 있다.

[0166] 마이크로스피어는 상업적으로 입수가능하다(Prolease™, Alkerme's: Cambridge, Mass.). 예를 들어, 적어도 하나의 치료제를 포함하는 동결 건조 매질을 적합한 용매에서 균질화하고 분무하여 20 내지 90㎛ 범위의 마이크로스피어를 제조한다. 이어서 정제, 캡슐화 및 저장 단계 동안 지속 방출 무결성을 유지하는 기술을 따른다. 문헌[Scott et al., Improving Protein Therapeutics With Sustained Release Formulations, Nature Biotechnology, Volume 16:153-157 (1998)]을 참조한다. 생분해성 중합체의 사용에 의한 마이크로스피어 조성물의 변형은 핵산 방출 속도를 제어하는 능력을 제공할 수 있다. 문헌[Miller et al., Degradation Rates of Oral Resorbable Implants {Polylactates and Polyglycolates: Rate Modification and Changes in PLA/PGA Copolymer Ratios, J. Biomed. Mater. Res., Vol. 11:711-719 (1977)]을 참조한다.

[0167] 대안적으로, 생분해성 중합체 금속염의 유기 용매 용액이 먼저 제조되는 수중 건조 방법을 사용하여 지속 또는 제어 방출 마이크로스피어 제조용 물질을 제조한다. 이어서, 상기 생분해성 중합체 금속염 용액에 용해 또는 분산된 핵산 매질을 첨가한다. 핵산 대 생분해성 중합체 금속염의 중량비는 예를 들어 약 1:100000 내지 약 1:1, 바람직하게는 약 1:20000 내지 약 1:500, 보다 바람직하게는 약 1:10000 내지 약 1:500일 수 있다. 다음으로, 생분해성 중합체 금속염과 핵산을 함유하는 유기용매 용액을 수상에 부어 유/수(oil/water) 에멀션을 제조한다. 그런 다음 오일상의 용매를 증발시켜 마이크로스피어를 제공한다. 마지막으로 이러한 마이크로스피어를 회수하고 세척하고 동결 건조한다. 그 후, 마이크로스피어를 감압 하에 가열하여 잔류 물 및 유기 용매를 제거할 수 있다.

[0168] 생분해성 중합체 금속염 및 핵산 혼합물과 호환되는 마이크로스피어를 생산하는 데 유용한 다른 방법은: i) 코아세르베이트제의 점진적인 첨가 동안 상 분리; ii) 입자 덩어리 응집을 방지하기 위해 항응집제를 첨가하는 수중 건조 방법 또는 상 분리 방법 및 iii) 분무 건조 방법에 의한 것이 있다.

[0169] 일 구현예에서, 본 발명은 대략 1일 내지 6개월의 기간 동안 핵산의 제어 방출을 전달할 수 있는 마이크로스피어 또는 마이크로캡슐을 포함하는 배지를 고려한다. 일 구현예에서, 마이크로스피어 또는 미세입자는 의사가 매질이 분배될 때 매질을 명확하게 볼 수 있도록 착색될 수 있다. 또 다른 구현예에서, 마이크로스피어 또는 마이크로캡슐은 투명할 수 있다. 또 다른 구현예에서, 마이크로스피어 또는 미세입자는 방사선 불투과성 형광 염료로 함침된다.

[0170] 제어 방출 마이크로캡슐은 원심 압출, 팬 코팅 및 에어 서스펜션과 같은 공지된 캡슐화 기술을 사용하여 생산할 수 있다. 이러한 마이크로스피어 및/또는 마이크로캡슐은 원하는 방출 속도를 달성하도록 조작될 수 있다. 예를 들어, Oliosphere™(Macromed)는 제어 방출 마이크로스피어 시스템이다. 이러한 특정 마이크로스피어는 5-500μm 범위의 균일한 크기로 제공되며 생체 적합성 및 생분해성 중합체로 구성된다. 마이크로스피어의 특이적인 중합체 조성물은 버스트 효과(burst effect)의 효과적인 관리를 포함하여 맞춤형으로 설계된 마이크로스피어가 가능하도록 핵산 방출 속도를 제어할 수 있다. ProMaxx™(Epic Therapeutics, Inc.)는 단백질-매트릭스(protein-matrix) 전달 시스템이다. 이 시스템은 본질적으로 수성이며 표준 제약 전달 모델에 적용할 수 있다. 특히, ProMaxx™는 소형 및 거대 분자 약물을 모두 전달하는 생체침식성 단백질(bioerodible protein) 마이크로스피어이며, 마이크로스피어 크기와 원하는 방출 특성 모두에 대해 맞춤화할 수 있다.

[0171] 일 구현예에서, 마이크로스피어 또는 미세입자는 내부 장간막의 pH보다 낮은 pH에서 안정한 pH 민감성 캡슐화 물질을 포함한다. 내부 장간막의 일반적인 범위는 pH 7.6에서 pH 7.2이다. 따라서 마이크로캡슐은 pH 7 미만으로 유지되어야 한다. 그러나 pH 가변성이 예상되는 경우 마이크로캡슐의 용해에 필요한 다양한 pH 기준에 따라 pH 민감성 물질을 선택할 수 있다. 따라서 캡슐화된 핵산은 용해가 필요한 pH 환경에 대해 선택되고 안정성을 유지하기 위해 미리 선택된 pH에 저장된다. 캡슐화제로서 유용한 pH 민감성 물질의 예는 Eudragit™ L-100 또는 S-100(Rohm GMBH), 히드록시프로필 메틸셀룰로스 프탈레이트, 히드록시프로필 메틸셀룰로스 아세테이트 숙시네이트, 폴리비닐 아세테이트 프탈레이트, 셀룰로스 아세테이트 프탈레이트, 및 셀룰로스 아세테이트 트리멜리테이트가 있다. 일 구현예에서, 지질은 마이크로캡슐의 내부 코팅을 포함한다. 이들 조성에서, 이러한 지질은 지방산 및 헥시티올 무수물의 부분 에스테르, 및 트리글리세리드와 같은 식용 지방일 수 있지만, 이에 제한되지는 않는다. 문헌[Lew C. W., Controlled-Release pH Sensitive Capsule And Adhesive System And Method. 미국 특허 번호 제5,364,634호]을 참조한다(본원에 참조로 포함됨).

[0172] 일 구현예에서, 본 발명은 젤라틴, 또는 젤라틴과 유사한 전하 밀도를 갖는 다른 중합체성 양이온(즉, 폴리-L-리신)을 포함하는 미세입자를 고려하고, 이는 1차 미세입자를 형성하기 위한 복합체로서 사용된다. 1차 미세입자는 다음 조성: i) 젤라틴(60 블룸, 돼지 피부의 A형), ii) 콘드로이틴 4-설페이트(0.005%-0.1%), iii) 글루타르알데히드(25%, 등급 1), 및 iv) 1-에틸-3-(3-디메틸아미노프로필)-카르보디이미드 염산염(EDC 염산염), 및 초순수 수크로스(ultra-pure sucrose)의 혼합물로 생성된다(Sigma Chemical Co., St. Louis, Mo.). 젤라틴의 출처는 중요하지 않으며, 소, 돼지, 인간 또는 기타 동물 출처에서 유래할 수 있다. 일반적으로 중합 양이온은 19,000-30,000 돌턴 사이이다. 그런 다음 황산 콘드로이틴을 황산나트륨 또는 코아세르베이션제로서 에탄올과 함께 복합체에 첨가한다.

[0173] 미세입자의 형성 후, 핵산은 미세입자의 표면에 직접 결합되거나 "가교" 또는 "스페이서"를 사용하여 간접적으로 부착된다. 젤라틴 리신 기의 아미노 기는 화합물의 직접적인 결합 부위를 제공하기 위해 쉽게 유도된다. 대안적으로, 아비딘-비오틴과 같은 스페이서(즉, 표적 리간드 상의 연결 분자 및 유도체화 모이어티)는 표적 리간드를 미세입자에 간접적으로 결합하는 데에도 유용하다. 미세입자의 안정성은 EDC 염산염에 의해 유도된 글루타르알데히드-스페이서 가교의 양에 의해 제어된다. 제어 방출 매질은 또한 글루타르알데히드-스페이서 가교의 최종 밀도에 의해 경험적으로 결정된다.

[0174] 일 구현예에서, 본 발명은 피브리노겐 또는 트롬빈을 포함하는 조성물을 핵산으로 분무 건조함으로써 형성된 미세입자를 고려한다. 바람직하게는, 이들 미세입자는 가용성이고 선택된 단백질(즉, 피브리노겐 또는 트롬빈)은 미세입자의 벽을 생성한다. 결과적으로, 핵산은 미세입자의 단백질 벽 내부와 그 사이에 통합된다. 문헌[Heath et al., Microparticles And Their Use In Wound Therapy. 미국 특허 번호 제6,113,948호]을 참조한다(본원에 참조로 포함됨). 살아있는 조직에 미세입자를 적용한 후, 피브리노겐과 트롬빈 사이의 후속 반응은 조직 밀봉재(tissue sealant)를 생성하여 통합된 화합물을 바로 주변 영역으로 방출한다.

[0175] 당업자는 마이크로스피어의 형태가 정확히 구형일 필요는 없음을 이해할 것이며; 수술 부위(즉, 개방 또는 폐쇄된) 안으로 또는 수술 부위 상에 분무되거나 퍼질 수 있는 매우 작은 입자로서만 가능하다. 일 구현예에서, 미세입자는 폴리락타이드, 폴리글리콜라이드 및 락타이드/글리콜라이드의 공중합체(PLGA), 히알루론산, 변형된 다당류 및 기타 잘 알려진 물질로 구성된 군에서 선택되는 생체적합성 및/또는 생분해성 물질로 구성된다.

[0176] **V.단백질 조성물**

[0177] CIRT 융합 단백질, 안정화제 폴리펩타이드, 링커, RNA 헤어핀 결합 도메인, NES, RNA 조절 도메인, 태그, NLS, RNA 표적화 분자, RNA 표적화 분자의 헤어핀 영역, 나선 영역, 또는 RNA 표적화 분자의 표적화 영역과 같은, 본 개시 내용의 폴리펩타이드 또는 폴리뉴클레오타이드는 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 또는 50개 이상의 변이체 아미노산 또는 핵산 치환을 포함할 수 있거나 또는 서열번호: 1 내지 132의 적어도, 또는 최대 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154,

155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 300, 400, 500, 550, 1000개 이상의 인접 아미노산 또는 핵산, 또는 여기에 속하는 임의의 추론가능한 범위와 적어도 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 또는 100% 유사하거나, 동일하거나 또는 상동성일 수 있다.

[0178] CIRT 융합 단백질, 안정화제 폴리펩타이드, 링커, RNA 헤어핀 결합 도메인, NES, RNA 조절 도메인, 태그, NLS, RNA 표적화 분자, RNA 표적화 분자의 헤어핀 영역, 나선 영역, 또는 RNA 표적화 분자의 표적화 영역과 같은, 본 개시 내용의 폴리펩타이드 또는 폴리뉴클레오타이드는 서열번호: 1 내지 132의 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 300, 400, 500, 550, 1000개 이상의 인접 아미노산 또는 핵산, 또는 여기에 속하는 임의의 추론가능한 범위를 포함할 수 있다.

[0179] 일부 구현예에서, 융합 단백질은 서열번호: 87 내지 106 또는 128 내지 132의 아미노산 1 내지 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289, 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 302, 303, 304, 305, 306, 307, 308, 309, 310, 311, 312, 313, 314, 315, 316, 317, 318, 319, 320, 321, 322, 323, 324, 325, 326, 327, 328, 329, 330, 331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 344, 345, 346, 347, 348, 349, 350, 351, 352, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375, 376, 377, 378, 379, 380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, 400, 401, 402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, 413, 414, 415, 416, 417, 418, 419, 420, 421, 422, 423, 424, 425, 426, 427, 428, 429, 430, 431, 432, 433, 434, 435, 436, 437, 438, 439, 440, 441, 442, 443, 444, 445, 446, 447, 448, 449, 450, 451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461, 462, 463, 464, 465, 466, 467, 468, 469, 470, 471, 472, 473, 474, 475, 476, 477, 478, 479, 480, 481, 482, 483, 484, 485, 486, 487, 488, 489, 490, 491, 492, 493, 494, 495, 496, 497, 498, 499, 500, 501, 502, 503, 504, 505, 506,

507, 508, 509, 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 또는 615개(또는 여기에 속하는 임의의 추론가능한 범위)를 포함할 수 있다.

[0180] 일부 구현예에서, 융합 단백질은 서열번호: 87 내지 106 또는 128 내지 132의 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289, 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 302, 303, 304, 305, 306, 307, 308, 309, 310, 311, 312, 313, 314, 315, 316, 317, 318, 319, 320, 321, 322, 323, 324, 325, 326, 327, 328, 329, 330, 331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 344, 345, 346, 347, 348, 349, 350, 351, 352, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375, 376, 377, 378, 379, 380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, 400, 401, 402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, 413, 414, 415, 416, 417, 418, 419, 420, 421, 422, 423, 424, 425, 426, 427, 428, 429, 430, 431, 432, 433, 434, 435, 436, 437, 438, 439, 440, 441, 442, 443, 444, 445, 446, 447, 448, 449, 450, 451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461, 462, 463, 464, 465, 466, 467, 468, 469, 470, 471, 472, 473, 474, 475, 476, 477, 478, 479, 480, 481, 482, 483, 484, 485, 486, 487, 488, 489, 490, 491, 492, 493, 494, 495, 496, 497, 498, 499, 500, 501, 502, 503, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 또는 615개(또는 여기에 속하는 임의의 추론가능한 범위) 인접 아미노산을 포함할 수 있다.

[0181] 일부 구현예에서, 융합 단백질은 서열번호: 87 내지 106 또는 128 내지 132 중 하나와 적어도, 최대, 또는 정확히 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 또는 100% 유사한, 동일한, 또는 상동성인, 서열번호: 87 내지 106 또는 128 내지 132의 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166,

167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289, 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 302, 303, 304, 305, 306, 307, 308, 309, 310, 311, 312, 313, 314, 315, 316, 317, 318, 319, 320, 321, 322, 323, 324, 325, 326, 327, 328, 329, 330, 331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 344, 345, 346, 347, 348, 349, 350, 351, 352, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375, 376, 377, 378, 379, 380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, 400, 401, 402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, 413, 414, 415, 416, 417, 418, 419, 420, 421, 422, 423, 424, 425, 426, 427, 428, 429, 430, 431, 432, 433, 434, 435, 436, 437, 438, 439, 440, 441, 442, 443, 444, 445, 446, 447, 448, 449, 450, 451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461, 462, 463, 464, 465, 466, 467, 468, 469, 470, 471, 472, 473, 474, 475, 476, 477, 478, 479, 480, 481, 482, 483, 484, 485, 486, 487, 488, 489, 490, 491, 492, 493, 494, 495, 496, 497, 498, 499, 500, 501, 502, 503, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 또는 615개(또는 여기에 속하는 임의의 추론가능한 범위) 인접 아미노산을 포함할 수 있다.

[0182] 다른 구현예에서, 안정화제 폴리펩타이드는 서열번호: 5, 19, 또는 20의 아미노산 1 내지 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 또는 90개(또는 여기에 속하는 임의의 추론가능한 범위)를 포함할 수 있다.

[0183] 일부 구현예에서, 안정화제 폴리펩타이드는 서열번호: 5, 19, 또는 20의 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 또는 90개(또는 여기에 속하는 임의의 추론가능한 범위) 인접 아미노산을 포함할 수 있다.

[0184] 일부 구현예에서, 안정화제 폴리펩타이드는 서열번호: 5, 19, 또는 20 중 하나와 적어도, 최대, 또는 정확히 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 또는 100% 유사한, 동일한, 또는 상동성인, 서열번호: 5, 19, 또는 20의 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 또는 90개(또는 여기에 속하는 임의의 추론가능한 범위) 인접 아미노산을 포함할 수 있다.

[0185] 일부 구현예에서, RNA 헤어핀 결합 도메인은 서열번호: 7 또는 18의 아미노산 1 내지 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 또는 102개(또는 여기에 속하는 임의의 추론가능한 범위)를 포함할 수 있다.

[0186] 일부 구현예에서, RNA 헤어핀 결합 도메인은 서열번호: 7 또는 18의 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12,

13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 또는 102개(또는 여기에 속하는 임의의 추론가능한 범위) 인접 아미노산을 포함할 수 있다.

[0187] 일부 구현예에서, RNA 헤어핀 결합 도메인은 서열번호: 7 또는 18 중 하나와 적어도, 최대, 또는 정확히 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 또는 100% 유사한, 동일한, 또는 상동성인, 서열번호: 7 또는 18의 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 또는 102개(또는 여기에 속하는 임의의 추론가능한 범위) 인접 아미노산을 포함할 수 있다.

[0188] 일부 구현예에서, RNA 조절 도메인은 서열번호: 9, 11, 15 내지 17, 또는 123 내지 125의 아미노산 1 내지 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289, 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 302, 303, 304, 305, 306, 307, 308, 309, 310, 311, 312, 313, 314, 315, 316, 317, 318, 319, 320, 321, 322, 323, 324, 325, 326, 327, 328, 329, 330, 331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 344, 345, 346, 347, 348, 349, 350, 351, 352, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 또는 364개(또는 여기에 속하는 임의의 추론가능한 범위)를 포함할 수 있다.

[0189] 일부 구현예에서, RNA 조절 도메인은 서열번호: 9, 11, 15 내지 17, 또는 123 내지 125의 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289, 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 302, 303, 304, 305, 306, 307, 308, 309, 310, 311, 312, 313, 314, 315, 316, 317, 318, 319, 320, 321, 322, 323, 324, 325, 326, 327, 328, 329, 330, 331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 344, 345, 346,

347, 348, 349, 350, 351, 352, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 또는 364개(또는 여기에 속하는 임의의 추론가능한 범위) 인접 아미노산을 포함할 수 있다.

[0190] 일부 측면에서, 서열번호: 1 내지 132 중 어느 하나의 위치 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289, 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 302, 303, 304, 305, 306, 307, 308, 309, 310, 311, 312, 313, 314, 315, 316, 317, 318, 319, 320, 321, 322, 323, 324, 325, 326, 327, 328, 329, 330, 331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 344, 345, 346, 347, 348, 349, 350, 351, 352, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375, 376, 377, 378, 379, 380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, 400, 401, 402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, 413, 414, 415, 416, 417, 418, 419, 420, 421, 422, 423, 424, 425, 426, 427, 428, 429, 430, 431, 432, 433, 434, 435, 436, 437, 438, 439, 440, 441, 442, 443, 444, 445, 446, 447, 448, 449, 450, 451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461, 462, 463, 464, 465, 466, 467, 468, 469, 470, 471, 472, 473, 474, 475, 476, 477, 478, 479, 480, 481, 482, 483, 484, 485, 486, 487, 488, 489, 490, 491, 492, 493, 494, 495, 496, 497, 498, 499, 500, 501, 502, 503, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 또는 615에서 시작하고 서열번호: 1 내지 132 중 어느 하나의 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289, 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 302, 303, 304, 305, 306, 307, 308, 309, 310, 311, 312, 313, 314, 315, 316, 317, 318, 319, 320, 321, 322, 323, 324, 325, 326, 327, 328, 329, 330, 331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 344, 345, 346, 347, 348, 349, 350, 351, 352, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 368,

369, 370, 371, 372, 373, 374, 375, 376, 377, 378, 379, 380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, 400, 401, 402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, 413, 414, 415, 416, 417, 418, 419, 420, 421, 422, 423, 424, 425, 426, 427, 428, 429, 430, 431, 432, 433, 434, 435, 436, 437, 438, 439, 440, 441, 442, 443, 444, 445, 446, 447, 448, 449, 450, 451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461, 462, 463, 464, 465, 466, 467, 468, 469, 470, 471, 472, 473, 474, 475, 476, 477, 478, 479, 480, 481, 482, 483, 484, 485, 486, 487, 488, 489, 490, 491, 492, 493, 494, 495, 496, 497, 498, 499, 500, 501, 502, 503, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 또는 615개 인접 뉴클레오타이드 또는 폴리펩타이드를 포함하는 핵산 분자 또는 폴리펩타이드가 있다.

[0191] 일부 구현예에서, RNA 조절 도메인은 서열번호: 9, 11, 15 내지 17, 또는 123 내지 125 중 하나와 적어도, 최대, 또는 정확히 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 또는 100% 유사한, 동일한, 또는 상동성인, 서열번호: 9, 11, 15 내지 17, 또는 123 내지 125의 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289, 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 302, 303, 304, 305, 306, 307, 308, 309, 310, 311, 312, 313, 314, 315, 316, 317, 318, 319, 320, 321, 322, 323, 324, 325, 326, 327, 328, 329, 330, 331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 344, 345, 346, 347, 348, 349, 350, 351, 352, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 또는 364개(또는 여기에 속하는 임의의 추론가능한 범위) 인접 아미노산을 포함할 수 있다.

[0192] 일부 구현예에서, RNA 표적화 분자의 헤어핀 구조, 예컨대 줄기, 루프 또는 줄기 및 루프 둘 모두는 서열번호: 1, 2, 26, 28, 또는 83 내지 86의 핵산 1 내지 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 또는 43개(또는 여기에 속하는 임의의 추론가능한 범위)를 포함할 수 있다.

[0193] 일부 구현예에서, RNA 표적화 분자의 헤어핀 구조, 예컨대 줄기, 루프 또는 줄기 및 루프 둘 모두는 서열번호: 1, 2, 26, 28, 또는 83 내지 86의 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 또는 43개(또는 여기에 속하는 임의의 추론가능한 범위) 인접 핵산을 포함할 수 있다.

[0194] 일부 구현예에서, RNA 표적화 분자의 헤어핀 구조, 예컨대 줄기, 루프 또는 줄기 및 루프 둘 모두는 서열번호: 1, 2, 26, 28, 또는 83 내지 86 중 하나와 적어도, 최대, 또는 정확히 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 또는 100% 유사한, 동일한, 또는 상동성인, 서열번호: 1, 2, 26, 28, 또는 83 내지 86의 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16,

17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 또는 43개(또는 여기에 속하는 임의의 추론가능한 범위) 인접 핵산을 포함할 수 있다.

[0195] 일부 구현예에서, RNA 표적화 영역 또는 RNA 표적화 분자는 서열번호: 30 내지 62의 핵산 1 내지 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 또는 40개(또는 여기에 속하는 임의의 추론가능한 범위)를 포함할 수 있다.

[0196] 일부 구현예에서, RNA 표적화 영역 또는 RNA 표적화 분자는 서열번호: 30 내지 62의 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 또는 40개(또는 여기에 속하는 임의의 추론가능한 범위) 인접 핵산을 포함할 수 있다.

[0197] 일부 구현예에서, RNA 표적화 영역 또는 RNA 표적화 분자는 서열번호: 30 내지 62 중 하나와 적어도, 최대, 또는 정확히 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 또는 100% 유사한, 동일한, 또는 상동성인, 서열번호: 30 내지 62의 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 또는 40개(또는 여기에 속하는 임의의 추론가능한 범위) 인접 핵산을 포함할 수 있다.

[0198] 일부 측면에서, 안정화제 폴리펩타이드는 내인성 단백질에 대한 결합을 감소시키거나 또는 제거하는 하나 이상의 치환을 가질 수 있다. 일부 측면에서, 안정화제 폴리펩타이드는 내인성 단백질에 대한 활성을 감소시키거나 또는 제거하는 하나 이상의 치환을 가질 수 있다.

[0199] 본 개시 내용의 폴리펩타이드 및 핵산, 예컨대 CIRT 융합 단백질, 안정화제 폴리펩타이드, 링커, RNA 헤어핀 결합 도메인, NES, RNA 조절 도메인, 태그, NLS, RNA 표적화 분자, RNA 표적화 분자의 헤어핀 영역, 나선 영역 , 또는 RNA 표적화 분자의 표적화 영역은 적어도, 최대, 또는 정확히 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289, 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 302, 303, 304, 305, 306, 307, 308, 309, 310, 311, 312, 313, 314, 315, 316, 317, 318, 319, 320, 321, 322, 323, 324, 325, 326, 327, 328, 329, 330, 331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 344, 345, 346, 347, 348, 349, 350, 351, 352, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375, 376, 377, 378, 379, 380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, 400, 401, 402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, 413, 414, 415, 416, 417, 418, 419, 420, 421, 422, 423, 424, 425, 426, 427, 428, 429, 430, 431, 432, 433, 434, 435, 436, 437, 438, 439, 440, 441, 442, 443, 444, 445, 446, 447, 448, 449, 450, 451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461, 462, 463, 464, 465, 466, 467, 468, 469, 470, 471, 472, 473, 474, 475, 476, 477, 478, 479, 480, 481, 482, 483, 484, 485, 486, 487, 488, 489, 490, 491, 492, 493, 494, 495, 496, 497, 498, 499, 500, 501, 502, 503, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610,

611, 612, 613, 614, 또는 615개 치환을 포함할 수 있다.

[0200] 상기 치환은 서열번호: 1 내지 132 중 하나의 아미노산 위치 또는 핵산 위치 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289, 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 302, 303, 304, 305, 306, 307, 308, 309, 310, 311, 312, 313, 314, 315, 316, 317, 318, 319, 320, 321, 322, 323, 324, 325, 326, 327, 328, 329, 330, 331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 344, 345, 346, 347, 348, 349, 350, 351, 352, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375, 376, 377, 378, 379, 380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, 400, 401, 402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, 413, 414, 415, 416, 417, 418, 419, 420, 421, 422, 423, 424, 425, 426, 427, 428, 429, 430, 431, 432, 433, 434, 435, 436, 437, 438, 439, 440, 441, 442, 443, 444, 445, 446, 447, 448, 449, 450, 451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461, 462, 463, 464, 465, 466, 467, 468, 469, 470, 471, 472, 473, 474, 475, 476, 477, 478, 479, 480, 481, 482, 483, 484, 485, 486, 487, 488, 489, 490, 491, 492, 493, 494, 495, 496, 497, 498, 499, 500, 501, 502, 503, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 또는 615에 있을 수 있다.

[0201] 치환 변이체는 일반적으로 단백질 내의 하나 이상의 부위에서 하나의 아미노산을 다른 것으로 교환하는 것을 포함하고, 다른 기능 또는 특성의 손실이 있거나 또는 없는 상태로, 폴리펩타이드의 하나 이상의 특성을 조절하도록 설계될 수 있다. 치환은 보존적일 수 있는데, 즉, 하나의 아미노산이 유사한 모양과 전하를 갖는 아미노산으로 대체된다. 보존적 치환은 당업계에 잘 알려져 있으며, 예를 들어, 알라닌에서 세린으로; 아르기닌에서 리신으로; 아스파라긴에서 글루타민 또는 히스티딘으로; 아스파테이트에서 글루타메이트로; 시스테인에서 세린으로; 글루타민에서 아스파라긴으로; 글루타메이트에서 아스파르테이트로; 글리신에서 프롤린으로; 히스티딘에서 아스파라긴 또는 글루타민으로; 이소류신에서 류신 또는 발린으로; 류신에서 발린 또는 이소류신으로; 리신에서 아르기닌으로; 메티오닌에서 류신 또는 이소류신으로; 페닐알라닌에서 티로신, 류신 또는 메티오닌으로; 세린에서 트레오닌으로; 트레오닌에서 세린으로; 트립토판에서 티로신으로; 티로신에서 트립토판 또는 페닐알라닌으로; 및 발린에서 이소류신 또는 류신으로의 교체를 포함한다. 대안적으로, 치환은 폴리펩타이드의 기능 또는 활성이 영향을 받도록 비보존적일 수 있다. 비보존적 교체는 일반적으로 극성 또는 전하를 띤 아미노산을 비극성 또는 전하를 띠지 않는 아미노산으로 또는 그 반대로와 같이 화학적으로 유사하지 않은 것으로 잔기를 치환하는 것을 포함한다.

[0202] 단백질은 재조합되거나 *시험관 내*에서 합성될 수 있다. 대안적으로, 비재조합 또는 재조합 단백질은 박테리아로부터 단리될 수 있다. 이러한 변이체를 함유하는 박테리아가 조성물 및 방법에서 구현될 수 있음이 또한 고려된다. 따라서 단백질을 단리할 필요가 없다.

[0203] 용어 "기능적으로 등가인 코돈"은 아르기닌 또는 세린에 대한 6개의 코돈과 같이 동일한 아미노산을 암호화하

는 코돈을 지칭하기 위해 본원에서 사용되며, 또한 생물학적으로 등가인 아미노산을 암호화하는 코돈을 지칭한다.

[0204] 또한 아미노산 및 핵산 서열은 각각 추가 N- 또는 C-말단 아미노산, 또는 5' 또는 3' 서열과 같은 추가 잔기를 포함할 수 있고, 상기 서열이 단백질 발현이 관련된 생물학적 단백질 활성의 유지를 포함하여 상기 제시된 기준을 충족하는 한, 여전히 본질적으로 본원에 개시된 서열 중 하나에 제시된 바와 같을 수 있음이 이해될 것이다. 말단 서열의 추가는 특히, 예를 들어 코딩 영역의 5' 또는 3' 부분 중 하나에 인접하는 다양한 비-코딩 서열을 포함할 수 있는 핵산 서열에 적용된다.

[0205] 다음은 등가이거나 개선된 2세대 분자를 생성하기 위해 단백질의 아미노산을 교체하는 것에 기반한 논의이다. 예를 들어, 특정 아미노산은 상호작용 결합 능력의 상당한 손실 없이 단백질 구조의 다른 아미노산으로 치환될 수 있다. 예를 들어, 효소 촉매 도메인 또는 상호작용 성분과 같은 구조는 이러한 기능을 유지하기 위해 치환된 아미노산을 가질 수 있다. 단백질의 생물학적 기능적 활성을 정의하는 것은 단백질의 상호작용 능력과 성질이기 때문에, 단백질 서열과 그 기본 DNA 코딩 서열에서 특정 아미노산 치환이 이루어질 수 있지만 그럼에도 불구하고 유사한 특성을 가진 단백질을 생성할 수 있다. 따라서 본 발명자들은 생물학적 유용성 또는 활성의 상당한 손실 없이 유전자의 DNA 서열에 다양한 변경이 이루어질 수 있음을 고려한다.

[0206] 다른 구현예에서, 폴리펩타이드의 기능의 변경은 하나 이상의 치환을 도입함으로써 의도된다. 예를 들어, 특정 아미노산은 상호작용 성분의 상호작용 결합 능력을 변경하기 위한 의도로 단백질 구조의 다른 아미노산으로 치환될 수 있다. 예를 들어, 단백질 상호작용 도메인, 핵산 상호작용 도메인, 및 촉매 부위와 같은 구조는 이러한 기능을 변경하기 위해 치환된 아미노산을 가질 수 있다. 단백질의 생물학적 기능적 활성을 정의하는 것은 단백질의 상호작용 능력과 성질이기 때문에, 단백질 서열과 그 기본 DNA 코딩 서열에서 특정 아미노산 치환이 이루어질 수 있지만 그럼에도 불구하고 상이한 특성을 가진 단백질을 생성할 수 있다. 따라서 본 발명자들은 생물학적 유용성 또는 활성의 상당한 변경이 있도록 유전자의 DNA 서열에 다양한 변경이 이루어질 수 있음을 고려한다.

[0207] 이러한 변경을 수행할 때 아미노산의 소수성 지표를 고려할 수 있다. 단백질에 상호작용적 생물학적 기능을 부여하는 데 있어서 소수성 아미노산 지표의 중요성은 일반적으로 당업계에서 이해된다(Kyte and Doolittle, 1982). 아미노산의 상대적인 소수성 특성은 생성된 단백질의 2차 구조에 기여하고, 이는 결국 단백질과 다른 분자, 예를 들어 효소, 기질(substrate), 수용체, DNA, 항체, 항원 등과의 상호작용을 정의한다는 것이 인정된다.

[0208] 유사 아미노산의 치환은 친수성에 기초하여 효과적으로 이루어질 수 있음이 또한 당업계에서 이해된다. 본원에 참조로 포함된 미국 특허 제4,554,101호는 인접 아미노산의 친수성에 의해 좌우되는 단백질의 최대 로컬 평균 친수성이 단백질의 생물학적 특성과 상관관계가 있다고 기술하고 있다. 아미노산은 유사한 친수성 값을 갖는 다른 아미노산으로 치환될 수 있고 여전히 생물학적으로 등가이고 면역학적으로 등가인 단백질을 생성할 수 있는 것으로 이해된다.

[0209] 위에서 약술한 바와 같이, 아미노산 치환은 일반적으로 아미노산 측쇄 치환기의 상대적 유사성, 예를 들어 소수성, 친수성, 전하, 크기 등을 기반으로 한다. 전술한 다양한 특성을 고려한 예시적인 치환은 잘 알려져 있으며 아르기닌 및 리신; 글루타메이트 및 아스파르테이트; 세린 및 트레오닌; 글루타민 및 아스파라긴; 및 발린, 류신 및 이소류신을 포함한다.

[0210] 특정 구현예에서, 본원에 기재된 단백질의 전부 또는 일부는 또한 통상적인 기술에 따라 용액 내 또는 고체 지지체 상에서 합성될 수 있다. 다양한 자동 합성기가 시판되고 있으며 공지된 프로토콜에 따라 사용할 수 있다. 예를 들어 각각 본원에 참조로 포함된 문헌[Stewart and Young, (1984); Tam *et al.*, (1983); Merrifield, (1986); 및 Barany and Merrifield (1979)]을 참조 한다. 대안적으로, 펩타이드 또는 폴리펩타이드를 암호화하는 뉴클레오타이드 서열이 발현 벡터에 삽입되고, 적절한 숙주 세포로 형질전환되거나 형질주입되고, 발현에 적합한 조건 하에 배양되는 재조합 DNA 기술이 사용될 수 있다.

[0211] 한 구현예는 단백질의 생산 및/또는 제시를 위한 미생물을 비롯한 세포로의 유전자 전달의 사용을 포함한다. 관심 단백질에 대한 유전자는 적절한 숙주 세포로 전달된 후 적절한 조건 하에서 세포 배양이 이루어질 수 있다. 사실상 모든 폴리펩타이드를 암호화하는 핵산이 사용될 수 있다. 재조합 발현 벡터의 생성, 및 이에 포함된 요소들이 본원에서 논의된다. 대안적으로, 생산될 단백질은 단백질 생산에 사용되는 세포에 의해 정상적으로 합성되는 내인성 단백질일 수 있다.

[0212] VI.서열

[0213] TAR 헤어핀 지지체: ggccagaucugagccugggagcucucuggcc (서열번호: 1)

[0214] 인간 SLBP 헤어핀 지지체: ccaaaggcucuucucagagccaccca (서열번호: 2)

표 4

[0215]

이름:	설명:	
MBP-CIRTS-1	MBP 신호 전달 펩타이드	MKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIFWAHDRFGGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVDLIKNKHMNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRIAATMENAQKGEIMPNIPQMSAFWYAVRTAVINAASGRQTVDEALKDAQTNS (서열번호: 3)
		ITSLYKKAGSETVRFQSHHHHHHSSGVDLGTENLYFQSNA (서열번호: 4)
	ORF5	MDDPSFLTGRSTYAKRRRARRMNVCKCGAILHNNKDCRSSTISGHKLDRLRFVKEGRVALEGETPVYRTWVKWVETEYHINILETSDDEE (서열번호: 5)
	링커	GGSGGSGGSGGSGGSGGS (서열번호: 6)
	TBP6.7	MAVPETRPNHTIYINNLNSKIKKDELKKSLYAIFSQFGQILDILVPRQRTPRGQAFVIFKEVSSATNALRSMQGFPFYDKPMRIQYAKTDKRIPAKMKGTFV (서열번호: 7)
		GS
	HIV NES	LQLPPLERLTL (서열번호: 8)
	링커	GGSGGSGGSGGSGGSGGS (서열번호: 6)
	데드(dead) Pin 도메인	MELEIRPLFLVPDTNGFIDHLASLARLLESRKYILVVPLIVINELDGLAKGQETDHRAGGYARVVQEKARKSIEFLEQRFESRDSCLRALTSRGNELESIAFRSEDITGQLGNNDDLILSCCLHYCKDKAKDFMPASKEEPIRLLREVVLLTDDRNLRVKALTRNVPVRDIPAFLTWAQVG (서열번호: 9)
		SSG
	6H 태그	HHHHHH (서열번호: 10)
CIRTS-0	ORF5	MDDPSFLTGRSTYAKRRRARRMNVCKCGAILHNNKDCRSSTISGHKLDRLRFVKEGRVALEGETPVYRTWVKWVETEYHINILETSDDEE (서열번호: 5)
	링커	GGSGGSGGSGGSGGSGGS (서열번호: 6)
	TBP6.7	MAVPETRPNHTIYINNLNSKIKKDELKKSLYAIFSQFGQILDILVPRQRTPRGQAFVIFKEVSSATNALRSMQGFPFYDKPMRIQYAKTDKRIPAKMKGTFV (서열번호: 7)
		GS
	HIV NES	LQLPPLERLTL (서열번호: 8)
	링커	GGSGGSGGSGGSGGSGGS (서열번호: 6)
	데드(dead) Pin 도메인	MELEIRPLFLVPDTNGFIDHLASLARLLESRKYILVVPLIVINELDGLAKGQETDHRAGGYARVVQEKARKSIEFLEQRFESRDSCLRALTSRGNELESIAFRSEDITGQLGNNADLILSCCLHYCKDKAKDFMPASKEEPIRLLREVVLLTDDRNLRVKALTRNVPVRDIPAFLTWAQVG (서열번호: 11)
		GS
	3xFLAG	DYKDHDGDYKDHDIDYKDDDDK (서열번호: 12)
CIRTS-1	ORF5	MDDPSFLTGRSTYAKRRRARRMNVCKCGAILHNNKDCRSSTISGHKLDRLRFVKEGRVALEGETPVYRTWVKWVETEYHINILETSDDEE (서열번호: 5)
	링커	GGSGGSGGSGGSGGSGGS (서열번호: 6)
	TBP6.7	MAVPETRPNHTIYINNLNSKIKKDELKKSLYAIFSQFGQILDILVPRQRTPRGQAFVIFKEVSSATNALRSMQGFPFYDKPMRIQYAKTDKRIPAKMKGTFV (서열번호: 7)
	링커	GGSGGSGGSGGSGGSGGS (서열번호: 6)
	Pin 도메인	MELEIRPLFLVPDTNGFIDHLASLARLLESRKYILVVPLIVINELDGLAKGQETDHRAGGYARVVQEKARKSIEFLEQRFESRDSCLRALTSRGNELESIAFRSEDITGQLGNNDDLILSCCLHYCKDKAKDFMPASKEEPIRLLREVVLLTDDRNLRVKALTRNVPVRDIPAFLTWAQVG (서열번호: 9)
	NLS	KRPAATKKAGQAKKKK (서열번호: 13)
		GS
	태그	DYKDDDDK (서열번호: 14)

CIRTS-2	ORF5	MDDPSFLTGRSTYAKRRRARRMNVCKCGAILHNNKDCRSSTISGHKLDRLRFVKEGRVALEGETPVYRTWVKW VETEYHINILETSDDEE (서열번호: 5)
	링커	GGSGGSGGSGGSGGSGGS (서열번호: 6)
	TBP6.7	MAVPETRPNHTIYINNLNSKIKKDELKKSLYAIFSQFGQILDILVPRQRTPRGQAFVIFKEVSSATNALRSMQ GFPFYDKPMRIQYAKTDKRIPAKMKGTFV (서열번호: 7)
		GS
	HIV NES	LQLPPLERLTL (서열번호: 8)
	링커	GGSGGSGGSGGSGGSGGS (서열번호: 6)
	N 말단 YTHDF1	MSATSVDTQRTKGQDNKVQNGSLHQKDTVHDNDFEPYLTGQSNQSNSYPSMSDPYLSSYYPPSIGFPYSLNEA PWSTAGDPPIPYLTTYGQLSNGDHHFMHDAVFGQPGGLGNNIYQHRFNFFPENPAFSAWGTSGSQGQQTQSSA YGSSYTYPPSSLGGTVVDGQPGFHSDTLSKAPGMNSLEQGMVGLKIGDVSSSAVKTVGSVVSSVALTGVLSGN GGTNVNMPVSKPTSWAAIASKPAKPQPKMKTKSGPVMGGGLPPPPIKHNMDIGTWDNKGPVPKAPVPQQAPSP QAAPQPQQVAQPLPAQPPALAQPQYQSPQQPPQTRWVAPRNRNAAFGQSGGAGSDSNSPGNVQPNSAPSVES (서열번호: 15)
		GS
	태그	DYKDDDDK (서열번호: 14)
CIRTS-3_t	ORF5	MDDPSFLTGRSTYAKRRRARRMNVCKCGAILHNNKDCRSSTISGHKLDRLRFVKEGRVALEGETPVYRTWVKW VETEYHINILETSDDEE (서열번호: 5)
	링커	GGSGGSGGSGGSGGSGGS (서열번호: 6)
	TBP6.7	MAVPETRPNHTIYINNLNSKIKKDELKKSLYAIFSQFGQILDILVPRQRTPRGQAFVIFKEVSSATNALRSMQ GFPFYDKPMRIQYAKTDKRIPAKMKGTFV (서열번호: 7)
		GS
	HIV NES	LQLPPLERLTL (서열번호: 8)
	링커	GGSGGSGGSGGSGGSGGS (서열번호: 6)
	YTHDF2 (101-200)	DAMFGQPGALGSTPFLGQHGFNFFPSGIDFSAWGNNSSQGQSTQSSGYSSNYAYAPSSLGGAMIDGQSAFANE TLNKAPGMNTIDQGMAA (서열번호: 16)
		GS
	태그	DYKDDDDK (서열번호: 14)
CIRTS-3	ORF5	MDDPSFLTGRSTYAKRRRARRMNVCKCGAILHNNKDCRSSTISGHKLDRLRFVKEGRVALEGETPVYRTWVKW VETEYHINILETSDDEE (서열번호: 5)
	링커	GGSGGSGGSGGSGGSGGS (서열번호: 6)
	TBP6.7	MAVPETRPNHTIYINNLNSKIKKDELKKSLYAIFSQFGQILDILVPRQRTPRGQAFVIFKEVSSATNALRSMQ GFPFYDKPMRIQYAKTDKRIPAKMKGTFV (서열번호: 7)
		GS
	HIV NES	LQLPPLERLTL (서열번호: 8)
	링커	GGSGGSGGSGGSGGSGGS (서열번호: 6)
	N-말단 YTHDF2 (1-200)	MSASSLLEQRPKGQGNKVQNGSVHQKDGLNDDDFEPYLSPQARPNNAYTAMSDSYLPSYYSPSIGFSYSLGEA AWSTGGDTAMPYLTSYGQLSNGEPHFLPDAMFGQPGALGSTPFLGQHGFNFFPSGIDFSAWGNNSSQGQSTQS SGYSSNYAYAPSSLGGAMIDGQSAFANETLNKAPGMNTIDQGMAA (서열번호: 17)
		GS
	태그	DYKDDDDK (서열번호: 14)
CIRTS-4_t	ORF5	MDDPSFLTGRSTYAKRRRARRMNVCKCGAILHNNKDCRSSTISGHKLDRLRFVKEGRVALEGETPVYRTWVKW VETEYHINILETSDDEE (서열번호: 5)
	링커	GGSGGSGGSGGSGGSGGS (서열번호: 6)
	SLBP 124-198	ADFETDESVLMRRQKQINYGKNTIAYDRYIKEVPRHLRQPGIHPKTPNKFKKYSRRSWDQQIKLWKVALHFWD (서열번호: 18)
	HIV NES	LQLPPLERLTL (서열번호: 8)
	링커	GGSGGSGGSGGSGGSGGS (서열번호: 6)
	YTHDF2 (101-200)	DAMFGQPGALGSTPFLGQHGFNFFPSGIDFSAWGNNSSQGQSTQSSGYSSNYAYAPSSLGGAMIDGQSAFANE TLNKAPGMNTIDQGMAA (서열번호: 16)
		GS
	태그	DYKDDDDK (서열번호: 14)

CIRTS-4	ORF5	MDDPSFLTGRSTYAKRRRARRMNVCKCGAILHNNKDCRSSTISGHKLDRLRFVKEGRVALEGETPVYRTWVKW VETEYHINILETSDDEE (서열번호: 5)
	링커	GGSGGSGGSGGSGGSGGS (서열번호: 6)
	SLBP 124-198	ADFETDESVLMRRQKQINYGKNTIAYDRYIKEVPRHLRQPGIHPKTPNKFKKYSRRSWDQQIKLWKVALHFWD (서열번호: 18)
	HIV NES	LQLPPLERLTL (서열번호: 8)
	링커	GGSGGSGGSGGSGGSGGS (서열번호: 6)
	YTHDF2 (1-200)	MSASSLLEQRPKGQGNKVQNGSVHQKDGLNDDDFEPYLSPQARPNNAYTAMSDSYLPSYYSPSIGFSYSLGEA AWSTGGDTAMPYLTSYGQLSNGEPHFLPDAMFGQPGALGSTPFLGQHGFNFFPSGIDFSAWGNNSSQGQSTQS SGYSSNYAYAPSSLGGAMIDGQSAFANETLNKAPGMNTIDQGMAA (서열번호: 17)
		GS
	태그	DYKDDDDK (서열번호: 14)
CIRTS-5_t	HBEGF	MRVTLSSKPQALATPNKEEHGKRKKKGKGLGKKRDPCLRKYKDFCIHGECKYVKELRAPSCICHPGYHGERCH GLS (서열번호: 19)
	링커	GGSGGSGGSGGSGGSGGS (서열번호: 6)
	TBP6.7	MAVPETRPNHTIYINNLNSKIKKDELKKSLYAIFSQFGQILDILVPRQRTPRGQAFVIFKEVSSATNALRSMQ GFPFYDKPMRIQYAKTDKRIPAKMKGTFV (서열번호: 7)
		GS
	HIV NES	LQLPPLERLTL (서열번호: 8)
	링커	GGSGGSGGSGGSGGSGGS (서열번호: 6)
	YTHDF2 (101-200)	DAMFGQPGALGSTPFLGQHGFNFFPSGIDFSAWGNNSSQGQSTQSSGYSSNYAYAPSSLGGAMIDGQSAFANE TLNKAPGMNTIDQGMAA (서열번호: 16)
		GS
	태그	DYKDDDDK (서열번호: 14)
CIRTS-5	HBEGF	MRVTLSSKPQALATPNKEEHGKRKKKGKGLGKKRDPCLRKYKDFCIHGECKYVKELRAPSCICHPGYHGERCH GLS (서열번호: 19)
	링커	GGSGGSGGSGGSGGSGGS (서열번호: 6)
	TBP6.7	MAVPETRPNHTIYINNLNSKIKKDELKKSLYAIFSQFGQILDILVPRQRTPRGQAFVIFKEVSSATNALRSMQ GFPFYDKPMRIQYAKTDKRIPAKMKGTFV (서열번호: 7)
		GS
	HIV NES	LQLPPLERLTL (서열번호: 8)
	링커	GGSGGSGGSGGSGGSGGS (서열번호: 6)
	YTHDF2 (1-200)	MSASSLLEQRPKGQGNKVQNGSVHQKDGLNDDDFEPYLSPQARPNNAYTAMSDSYLPSYYSPSIGFSYSLGEA AWSTGGDTAMPYLTSYGQLSNGEPHFLPDAMFGQPGALGSTPFLGQHGFNFFPSGIDFSAWGNNSSQGQSTQS SGYSSNYAYAPSSLGGAMIDGQSAFANETLNKAPGMNTIDQGMAA (서열번호: 17)
		GS
	태그	DYKDDDDK (서열번호: 14)
CIRTS-6_t	β-디펜신	MGIINTLQKYYCRVRGGRCAVLSCLPKEEQIGKCSTRGRKCCRRKK (서열번호: 20)
	링커	GGSGGSGGSGGSGGSGGS (서열번호: 6)
	TBP6.7	MAVPETRPNHTIYINNLNSKIKKDELKKSLYAIFSQFGQILDILVPRQRTPRGQAFVIFKEVSSATNALRSMQ GFPFYDKPMRIQYAKTDKRIPAKMKGTFV (서열번호: 7)
		GS
	HIV NES	LQLPPLERLTL (서열번호: 8)
	링커	GGSGGSGGSGGSGGSGGS (서열번호: 6)
	YTHDF2 (101-200)	DAMFGQPGALGSTPFLGQHGFNFFPSGIDFSAWGNNSSQGQSTQSSGYSSNYAYAPSSLGGAMIDGQSAFANE TLNKAPGMNTIDQGMAA (서열번호: 16)
		GS
	태그	DYKDDDDK (서열번호: 14)
CIRTS-6	β-디펜신	MGIINTLQKYYCRVRGGRCAVLSCLPKEEQIGKCSTRGRKCCRRKK (서열번호: 20)
	링커	GGSGGSGGSGGSGGSGGS (서열번호: 6)
	TBP6.7	MAVPETRPNHTIYINNLNSKIKKDELKKSLYAIFSQFGQILDILVPRQRTPRGQAFVIFKEVSSATNALRSMQ GFPFYDKPMRIQYAKTDKRIPAKMKGTFV (서열번호: 7)
		GS
	HIV NES	LQLPPLERLTL (서열번호: 8)
	링커	GGSGGSGGSGGSGGSGGS (서열번호: 6)
	YTHDF2 (1-200)	MSASSLLEQRPKGQGNKVQNGSVHQKDGLNDDDFEPYLSPQARPNNAYTAMSDSYLPSYYSPSIGFSYSLGEA AWSTGGDTAMPYLTSYGQLSNGEPHFLPDAMFGQPGALGSTPFLGQHGFNFFPSGIDFSAWGNNSSQGQSTQS SGYSSNYAYAPSSLGGAMIDGQSAFANETLNKAPGMNTIDQGMAA (서열번호: 17)
		GS
	태그	DYKDDDDK (서열번호: 14)

CIRTS-7	β-디펜신	MGIINTLQKYYCRVRGGRCAVLSCLPKEEQIGKCSTRGRKCCRRKK (서열번호: 20)
	링커	GGSGGSGGSGGSGGSGGS (서열번호: 6)
	SLBP 124-198	ADFETDESVLMRRQKQINYGKNTIAYDRYIKEVPRHLRQPGIHPKTPNKFKKYSRRSWDQQIKLWKVALHFWD (서열번호: 18)
	HIV NES	LQLPPLERLTL (서열번호: 8)
	링커	GGSGGSGGSGGSGGSGGS (서열번호: 6)
	YTHDF2 (1-200)	MSASSLLEQRPKGQGNKVQNGSVHQKDGLNDDDFEPYLSPQARPNNAYTAMSDSYLPSYYSPSIGFSYSLGEAAWSTGGDTAMPYLTSYGQLSNGEPHFLPDAMFGQPGALGSTPFLGQHGFNFFPSGIDFSAWGNNSSQGQSTQSSGYSSNYAYAPSSLGGAMIDGQSAFANETLNKAPGMNTIDQGMAA (서열번호: 17)
		GS
	태그	DYKDDDDK (서열번호: 14)
CIRTS-8	TBP6.7	MAVPETRPNHTIYINNLNSKIKKDELKKSLYAIFSQFGQILDILVPRQRTPRGQAFVIFKEVSSATNALRSMQGFPFYDKPMRIQYAKTDKRIPAKMKGTFV (서열번호: 7)
		GS
	HIV NES	LQLPPLERLTL (서열번호: 8)
	링커	GGSGGSGGSGGSGGSGGS (서열번호: 6)
	Pin 도메인	MELEIRPLFLVPDTNGFIDHLASLARLLESRKYILVVPLIVINELDGLAKGQETDHRAGGYARVVQEKARKSIEFLEQRFESRDSCLRALTSRGNELESIAFRSEDITGQLGNNDDLILSCCLHYCKDKAKDFMPASKEEPIRLLREVVLLTDDRNLRVKALTRNVPVRDIPAFLTWAQVG (서열번호: 9)
		GS
	태그	DYKDDDDK (서열번호: 14)
CIRTS-9	ORF5	MDDPSFLTGRSTYAKRRRARRMNVCKCGAILHNNKDCRSSTISGHKLDRLRFVKEGRVALEGETPVYRTWVKWVETEYHINILETSDDEE (서열번호: 5)
	링커	GGSGGSGGSGGSGGSGGS (서열번호: 6)
	TBP6.7	MAVPETRPNHTIYINNLNSKIKKDELKKSLYAIFSQFGQILDILVPRQRTPRGQAFVIFKEVSSATNALRSMQGFPFYDKPMRIQYAKTDKRIPAKMKGTFV (서열번호: 7)
		GS
	HIV NES	LQLPPLERLTL (서열번호: 8)
	링커	GGSGGSGGSGGSGGSGGS (서열번호: 6)
	Pin 도메인	MELEIRPLFLVPDTNGFIDHLASLARLLESRKYILVVPLIVINELDGLAKGQETDHRAGGYARVVQEKARKSIEFLEQRFESRDSCLRALTSRGNELESIAFRSEDITGQLGNNDDLILSCCLHYCKDKAKDFMPASKEEPIRLLREVVLLTDDRNLRVKALTRNVPVRDIPAFLTWAQVG (서열번호: 9)
		GS
	태그	DYKDDDDK (서열번호: 14)
CIRTS-10	ORF5	MDDPSFLTGRSTYAKRRRARRMNVCKCGAILHNNKDCRSSTISGHKLDRLRFVKEGRVALEGETPVYRTWVKWVETEYHINILETSDDEE (서열번호: 5)
	링커	GGSGGSGGSGGSGGSGGS (서열번호: 6)
	TBP6.7	MAVPETRPNHTIYINNLNSKIKKDELKKSLYAIFSQFGQILDILVPRQRTPRGQAFVIFKEVSSATNALRSMQGFPFYDKPMRIQYAKTDKRIPAKMKGTFV (서열번호: 7)
		GS
	HIV NES	LQLPPLERLTL (서열번호: 8)
	링커	GGSGGSGGS (서열번호: 21)
	Pin 도메인	MELEIRPLFLVPDTNGFIDHLASLARLLESRKYILVVPLIVINELDGLAKGQETDHRAGGYARVVQEKARKSIEFLEQRFESRDSCLRALTSRGNELESIAFRSEDITGQLGNNDDLILSCCLHYCKDKAKDFMPASKEEPIRLLREVVLLTDDRNLRVKALTRNVPVRDIPAFLTWAQVG (서열번호: 9)
		GS
	태그	DYKDDDDK (서열번호: 14)
CIRTS-11	ORF5	MDDPSFLTGRSTYAKRRRARRMNVCKCGAILHNNKDCRSSTISGHKLDRLRFVKEGRVALEGETPVYRTWVKWVETEYHINILETSDDEE (서열번호: 5)
	링커	GGSGGSGGSGGSGGSGGS (서열번호: 6)
	TBP6.7	MAVPETRPNHTIYINNLNSKIKKDELKKSLYAIFSQFGQILDILVPRQRTPRGQAFVIFKEVSSATNALRSMQGFPFYDKPMRIQYAKTDKRIPAKMKGTFV (서열번호: 7)
		GS
	HIV NES	LQLPPLERLTL (서열번호: 8)
	L8 링커	SGSETPGTSESATPES (서열번호: 22)
	Pin 도메인	MELEIRPLFLVPDTNGFIDHLASLARLLESRKYILVVPLIVINELDGLAKGQETDHRAGGYARVVQEKARKSIEFLEQRFESRDSCLRALTSRGNELESIAFRSEDITGQLGNNDDLILSCCLHYCKDKAKDFMPASKEEPIRLLREVVLLTDDRNLRVKALTRNVPVRDIPAFLTWAQVG (서열번호: 9)
		GS
	태그	DYKDDDDK (서열번호: 14)

CIRTS-12	ORF5	MDDPSFLTGRSTYAKRRRARRMNVCKCGAILHNNKDCRSSTISGHKLDRLRFVKEGRVALEGETPVYRTWVKW VETEYHINILETSDDEE (서열번호: 5)
	링커	GGSGGSGGSGGSGGSGGS (서열번호: 6)
	TBP6.7	MAVPETRPNHTIYINNLNSKIKKDELKKSLYAIFSQFGQILDILVPRQRTPRGQAFVIFKEVSSATNALRSMQ GFPFYDKPMRIQYAKTDKRIPAKMKGTFV (서열번호: 7)
		GS
	HIV NES	LQLPPLERLTL (서열번호: 8)
	링커	GGSG (서열번호: 23)
	ER/K 10nm	EEEEKKKQQEEEAERLRRIQEEMEKERKRREEDEQRRRKEEEERRMKLEMEAKRKQEEEERKKREDDEKRKKK (서열번호: 24)
	링커	GSGGS (서열번호: 25)
	Pin 도메인	MELEIRPLFLVPDTNGFIDHLASLARLLESRKYILVVPLIVINELDGLAKGQETDHRAGGYARVVQEKARKSI EFLEQRFESRDSCLRALTSRGNELESIAFRSEDITGQLGNNDDLILSCCLHYCKDKAKDFMPASKEEPIRLLR EVVLLTDDRNLRVKALTRNVPVRDIPAFLTWAQVG (서열번호: 9)
		GS
	태그	DYKDDDDK (서열번호: 14)
CIRTS-13	ORF5	MDDPSFLTGRSTYAKRRRARRMNVCKCGAILHNNKDCRSSTISGHKLDRLRFVKEGRVALEGETPVYRTWVKW VETEYHINILETSDDEE (서열번호: 5)
	링커	GGSGGSGGSGGSGGSGGS (서열번호: 6)
	TBP6.7	MAVPETRPNHTIYINNLNSKIKKDELKKSLYAIFSQFGQILDILVPRQRTPRGQAFVIFKEVSSATNALRSMQ GFPFYDKPMRIQYAKTDKRIPAKMKGTFV (서열번호: 7)
		GS
	HIV NES	LQLPPLERLTL (서열번호: 8)
	L8 링커	SGSETPGTSESATPES (서열번호: 22)
	Y2(101-200)	DAMFGQPGALGSTPFLGQHGFNFFPSGIDFSAWGNNSSQGQSTQSSGYSSNYAYAPSSLGGAMIDGQSAFANE TLNKAPGMNTIDQGMAA (서열번호: 16)
		GS
	태그	DYKDDDDK (서열번호: 14)
CIRTS-14	ORF5	MDDPSFLTGRSTYAKRRRARRMNVCKCGAILHNNKDCRSSTISGHKLDRLRFVKEGRVALEGETPVYRTWVKW VETEYHINILETSDDEE (서열번호: 5)
	링커	GGSGGSGGSGGSGGSGGS (서열번호: 6)
	TBP6.7	MAVPETRPNHTIYINNLNSKIKKDELKKSLYAIFSQFGQILDILVPRQRTPRGQAFVIFKEVSSATNALRSMQ GFPFYDKPMRIQYAKTDKRIPAKMKGTFV (서열번호: 7)
		GS
	HIV NES	LQLPPLERLTL (서열번호: 8)
	링커	GGSG (서열번호: 23)
	ER/K 10nm	EEEEKKKQQEEEAERLRRIQEEMEKERKRREEDEQRRRKEEEERRMKLEMEAKRKQEEEERKKREDDEKRKKK (서열번호: 24)
	링커	GSGGS (서열번호: 25)
	Y2(101-200)	DAMFGQPGALGSTPFLGQHGFNFFPSGIDFSAWGNNSSQGQSTQSSGYSSNYAYAPSSLGGAMIDGQSAFANE TLNKAPGMNTIDQGMAA (서열번호: 16)
		GS
	태그	DYKDDDDK (서열번호: 14)
CIRTS-7X	b-디펜신 3	MGIINTLQKYYCRVRGGRCAVLSCLPKEEQIGKCSTRGRKCCRRKK (서열번호: 20)
	GGS_6	GGSGGSGGSGGSGGSGGS (서열번호: 6)
	TBP6.7	MAVPETRPNHTIYINNLNSKIKKDELKKSLYAIFSQFGQILDILVPRQRTPRGQAFVIFKEVSSATNALRSMQ GFPFYDKPMRIQYAKTDKRIPAKMKGTFV (서열번호: 7)
	GGS_6	GGSGGSGGSGGSGGSGGS (서열번호: 6)
	hADAR(299-701)	LHLDQTPSRQPIPSEGLQLHLPQVLADAVSRLVLGKFGDLTDNFSSPHARRKVLAGVVMTTGTDVKDAKVISV STGTKCINGEYMSDRGLALNDCHAEIISRRSLLRFLYTQLELYLNNKDDQKRSIFQKSERGGFRLKENVQFHL YISTSPCGDARIFSPHEPILEEPADRHPNRKARGQLRTKIESGEGTIPVRSNASIQTWDGVLQGERLLTMSCS DKIARWNVVGIQGSLLSIFVEPIYFSSIILGSLYHGDHLSRAMYQRISNIEDLPPLYTLNKPLLSGISNAEAR QPGKAPNFSVNWTVGDSAIEVINATTGKDELGRASRLCKHALYCRWMRVHGKVPSHLLRSKITKPNVYHESKL AAKEYQAAKARLFTAFIKAGLGAWVEKPTEQDQFSLTP (서열번호: 123)
	NLS	KRPAATKKAGQAKKKK (서열번호: 13)
		GS
	플래그 태그	DYKDDDDK (서열번호: 14)

CIRTS-8X	b-디펜신 3	MGIINTLQKYYCRVRGGRCAVLSCLPKEEQIGKCSTRGRKCCRRKK (서열번호: 20)
	GGS_6	GGSGGSGGSGGSGGSGGS (서열번호: 6)
	TBP6.7	MAVPETRPNHTIYINNLNSKIKKDELKKSLYAIFSQFGQILDILVPRQRTPRGQAFVIFKEVSSATNALRSMQGFPFYDKPMRIQYAKTDKRIPAKMKGTFV (서열번호: 7)
	GGS_6	GGSGGSGGSGGSGGSGGS (서열번호: 6)
	hADAR (299-701) E488Q	LHLDQTPSRQPIPSEGLQLHLPQVLADAVSRLVLGKFGDLTDNFSSPHARRKVLAGVVMTTGTDVKDAKVISVSTGTKCINGEYMSDRGLALNDCHAEIISRRSLLRFLYTQLELYLNNKDDQKRSIFQKSERGGFRLKENVQFHLYISTSPCGDARIFSPHEPILEEPADRHPNRKARGQLRTKIESGQGTIPVRSNASIQTWDGVLQGERLLTMSCSDKIARWNVVGIQGSLLSIFVEPIYFSSIILGSLYHGDHLSRAMYQRISNIEDLPPLYTLNKPLLSGISNAEARQPGKAPNFSVNWTVGDSAIEVINATTGKDELGRASRLCKHALYCRWMRVHGKVPSHLLRSKITKPNVYHESKLAAKEYQAAKARLFTAFIKAGLGAWVEKPTEQDQFSLTP (서열번호: 124)
	NLS	KRPAATKKAGQAKKKK (서열번호: 13)
		GS
	플래그 태그	DYKDDDDK (서열번호: 14)
CIRTS-9X	b-디펜신 3	MGIINTLQKYYCRVRGGRCAVLSCLPKEEQIGKCSTRGRKCCRRKK (서열번호: 20)
	GGS_6	GGSGGSGGSGGSGGSGGS (서열번호: 6)
	TBP6.7	MAVPETRPNHTIYINNLNSKIKKDELKKSLYAIFSQFGQILDILVPRQRTPRGQAFVIFKEVSSATNALRSMQGFPFYDKPMRIQYAKTDKRIPAKMKGTFV (서열번호: 7)
	NES	GSLQLPPLERLTL (서열번호: 124)
	GGS_6	GGSGGSGGSGGSGGSGGS (서열번호: 6)
	NYTHDF1	MSATSVDTQRTKGQDNKVQNGSLHQKDTVHDNDFEPYLTGQSNQSNSYPSMSDPYLSSYYPPSIGFPYSLNEAPWSTAGDPPIPYLTTYGQLSNGDHHFMHDAVFGQPGGLGNNIYQHRFNFFPENPAFSAWGTSGSQGQQTQSSAYGSSYTYPPSSLGGTVVDGQPGFHSDTLSKAPGMNSLEQGMVGLKIGDVSSSAVKTVGSVVSSVALTGVLSGNGGTNVNMPVSKPTSWAAIASKPAKPQPKMKTKSGPVMGGGLPPPPIKHNMDIGTWDNKGPVPKAPVPQQAPSPQAAPQPQQVAQPLPAQPPALAQPQYQSPQQPPQTRWVAPRNRNAAFGQSGGAGSDSNSPGNVQPNSAPSVES (서열번호: 15)
		GS
	플래그 태그	DYKDDDDK (서열번호: 14)
CIRTS-10X	b-디펜신 3	MGIINTLQKYYCRVRGGRCAVLSCLPKEEQIGKCSTRGRKCCRRKK (서열번호: 20)
	GGS_6	GGSGGSGGSGGSGGSGGS (서열번호: 6)
	SLBP	ADFETDESVLMRRQKQINYGKNTIAYDRYIKEVPRHLRQPGIHPKTPNKFKKYSRRSWDQQIKLWKVALHFWD (서열번호: 18)
	NES	LQLPPLERLTL (서열번호: 8)
	GGS_6	GGSGGSGGSGGSGGSGGS (서열번호: 6)
	YTHDF2(101-200)	MSASSLLEQRPKGQGNKVQNGSVHQKDGLNDDDFEPYLSPQARPNNAYTAMSDSYLPSYYSPSIGFSYSLGEAAWSTGGDTAMPYLTSYGQLSNGEPHFLPDAMFGQPGALGSTPFLGQHGFNFFPSGIDFSAWGNNSSQGQSTQSSGYSSNYAYAPSSLGGAMIDGQSAFANETLNKAPGMNTIDQGMAA (서열번호: 17)
		GS
	플래그 태그	DYKDDDDK (서열번호: 14)
CIRTS-18X	ORF5	MDDPSFLTGRSTYAKRRRARRMNVCKCGAILHNNKDCRSSTISGHKLDRLRFVKEGRVALEGETPVYRTWVKWVETEYHINILETSDDEE (서열번호: 5)
	GGS_6	GGSGGSGGSGGSGGSGGS (서열번호: 6)
	PP7	MAKTIVLAVGEATRTLTEIQSTADRQIFEEKVGPLVGRLRLTASLRQNGAKTAYRVNLKLDQADVVDASTSVAGELPKVRYTQVWSHDVTIVANSTEASRKSLYDLTKSLVATSQVEDLVVNLVPLGRSLEGGSGGMAKTIVLAVGEATRTLTEIQSTADRQIFEEKVGPLVGRLRLTASLRQNGAKTAYRVNLKLDQADVVDASTSVAGELPKVRYTQVWSHDVTIVANSTEASRKSLYDLTKSLVATSQVEDLVVNLVPLGRSLE (서열번호: 125)
	NES	LQLPPLERLTL (서열번호: 8)
	GGS_6	GGSGGSGGSGGSGGSGGS (서열번호: 6)
	Y2(100-200)	DAMFGQPGALGSTPFLGQHGFNFFPSGIDFSAWGNNSSQGQSTQSSGYSSNYAYAPSSLGGAMIDGQSAFANETLNKAPGMNTIDQGMAA (서열번호: 16)
		GS
	플래그 태그	DYKDDDDK (서열번호: 14)

[0216] 본 연구에 사용된 gRNA 서열(모든 gRNA는 hU6 프로모터로부터 발현됨)

표 5

[0217]

이름:		설명:	DNA 서열
TBP-OT	26-51	TBP6.7 gRNA (TAR)	GGCCAGATCTGAGCCTGGGAGCTCTCTGGCC (서열번호: 26)
			TTATT
		표적외 1 파지	Tgacagcccacatggcattccacttatcactggcatcctt (서열번호: 27)
SLBP-OT	24-18	SLBP gRNA	CCAAAGGCTCTTCTCAGAGCCACCCA (서열번호: 28)
			TTATT
		1 표적	GTGATAAGTGGAATGCCATG (서열번호: 29)
PP7-NT		PP7 비-표적화 gRNA 1 파지	TAAGGAGTTTATATGGAAACCCTTA (서열번호: 121)
			TTATT
		1 파지	GTGATAAGTGGAATGCCATG (서열번호: 122)

[0218] gRNA 가이드 서열분석; 모두 TBP-OT와 동일한 hU6 프로모터에 의해 발현됨

표 6

[0219]

표적 이름:		가이드 서열:	도면:
플루시페라아제-40	18-20	caggtcgactctagactcgaggctagcgagctcgtttaaa (서열번호: 30)	3
PPIB	23-26	cttggtgctctccaccttccgcaccacctccatgccctct (서열번호: 31)	4
SMARCA4	23-25	ccgatgcggtgggctcggtcctgcgcttgcaggtcctggt (서열번호: 32)	4
NFKB1	22-67	gcctccaccagctctctgactgtacccccagagacctcat (서열번호: 33)	4
NRAS	25-30	aagcatcttcaacaccctgtctggtcttggctgaggtttc (서열번호: 34)	4
B4GALTN1	24-50	cttcgcaccgcagcgcagcgcggctcagctcccggctcgt (서열번호: 35)	4
플루시페라아제-20		CTAGACTCGAGGCTAGCGAG (서열번호: 36)	8
플루시페라아제-30		CGACTCTAGACTCGAGGCTAGCGAGCTCGT (서열번호: 37)	8
KRAS-2	22-45	cttgtggtagttggagctga (서열번호: 116)	25B
KRAS-3	22-53	agttggagctgatggcgtag (서열번호: 117)	25B
KRAS-4	22-54	gttggagctgatggcgtagg (서열번호: 118)	25B
KRAS-5	22-65	tggcgtaggcaagagtgcct (서열번호: 119)	25B
KRAS-6	22-66	ggcgtaggcaagagtgcctt (서열번호: 120)	25B
SMARCA4-1	27-67	ggccctggcccttcccctggagccatgctgggccctagcc (서열번호: 38)	11C
SMARCA4-2	27-71	gcccaaccccatttaaccagaaccagctgcaccagctcag (서열번호: 39)	11C
SMARCA4-3	27-74	cctcccaagccctggcctgaaggacccatggcgaatgctg (서열번호: 40)	11C
SMARCA4-4	27-75	cgtcccacccgccgcctcgcccgtgatgccaccgcagacc (서열번호: 41)	11C
SMARCA4-5	27-78	ggaggtggtggtgtgcatgcggagggacacagcgctggag (서열번호: 42)	11C
SMARCA4-6	27-80	Tcacaggcaaaatccagaagctgaccaaggcagtggccac (서열번호: 43)	11C
SMARCA4-7	27-81	catggctgaagatgaggaggggtaccgcaagctcatcgac (서열번호: 44)	11C
SMARCA4-8	28-02	ctctggacgagaccagccagatgagcgacctcccggtgaa (서열번호: 45)	11C
SMARCA4-9	28-04	cccaccctgcccgtggaggagaagaagaagattccagatc (서열번호: 46)	11C
SMARCA4-10	28-05	aatatggcgtgtcccaggcccttgcacgtggcctgcagtc (서열번호: 47)	11C
SMARCA4-11	28-07	ctcatcacgtacctcatggagcacaaacgcatcaatgggc (서열번호: 48)	11C
SMARCA4-12	28-08	tggtgaaggtgtcttacaagggatccccagcagcaagacg (서열번호: 49)	11C
SMARCA4-13	28-10	accccgccgcctgctgctgacgggcacaccgctgcagaac (서열번호: 50)	11C
SMARCA4-14	28-11	cagtggtttaacgcacccttttgccatgaccggggaaaagg (서열번호: 51)	11C
SMARCA4-15	28-12	acgactcaagaaggaagtcgaggcccagttgcccgaaaag (서열번호: 52)	11C
SMARCA4-16	28-13	gtgctgctgactgatggctccgagaaggacaagaagggca (서열번호: 53)	11C
SMARCA4-17	28-16	ggaaccacgaaggcggaggaccggggcatgctgctgaaaa (서열번호: 54)	11C
SMARCA4-18	28-17	tccagtcggcagacactgtgatcatttttgacagcgactg (서열번호: 55)	11C
SMARCA4-19	28-20	cgccagcgggcgtcaaccccgacttggaggagccacctct (서열번호: 56)	11C
SMARCA4-20	28-22	gcggaggtggagcggctgacctgtgaggaggaggaggaga (서열번호: 57)	11C
SMARCA4-21	28-26	tgatcaagtacaaggacagcagcagtggacgtcagctcag (서열번호: 58)	11C
SMARCA4-22	28-27	cttcaagaagataaaggagcgcattcgcaaccacaagtac (서열번호: 59)	11C
SMARCA4-23	28-30	agggtcccgagccaagccggtcgtgagtgacgatgacagt (서열번호: 60)	11C
SMARCA4-24	28-32	tatttatacagcagagaagctgtaggactgtttgtgactg (서열번호: 61)	11C
SMARCA4-25	28-33	ggggaacacacgatacctgtttttcttttccgttgctggc (서열번호: 62)	11C

[0220] *시험관 내* 특성화에 사용되는 RNA 올리고뉴클레오타이드:

표 7

[0221]

유전자 이름:	올리고 이름:	순방향:	도면:
R1	EMSA에 대한 ssRNA 표적	ATAGGCCAGTGAATTCGAGCTCGAATATGGATTACTTGGTAGAACAGCAATCTACGCCGGAAGCATAAAG (서열번호: 63)	2A
R2	절단 분석에 대한 ssRNA 표적	GGCCAGTGAATTCGAGCTCGGTACCCGGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAGCAATCTACTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAG (서열번호: 64)	2C
R3	EMSA에 대한 gRNA	GGCCAGATCTGAGCCTGGGAGCTCTCTGGCCCagccaccacccacagagccgccaccaga (서열번호: 65)	2A
R4	절단 분석에 대한 gRNA	GGCCAGATCTGAGCCTGGGAGCTCTCTGGCCCTAGATTGCTGTTCTACCAAGTAATCCAT (서열번호: 66)	2C

[0222] qPCR 프라이머:

표 8

[0223]

유전자 이름:	순방향:	서열 번호:	역방향:	서열 번호:
GAPDH	GTCTCCTCTGACTTCAACAGCG	67	ACCACCCTGTTGCTGTAGCCAA	68
PPIB	AACGCAGGCAAAGACACCAACG	69	TCTGTCTTGGTGCTCTCCACCT	70
플루시페라아제	AGGTTACAACCGCCAAGAAGC	71	ATGAGAATCTCGCGGATCTTG	72
PPIB	AACGCAGGCAAAGACACCAACG	73	TCTGTCTTGGTGCTCTCCACCT	74
NFKB1	gcagcactacttcttgaccacc	75	tctgctcctgagcattgacgtc	76
NRAS	gaaacctcagccaagaccagac	77	ggcaatcccatacaaccctgag	78
B4GALTN1	tgaggctgctttcactatccgc	79	gaggaaggtcttggtggcaatc	80
SMARCA4	caaagacaagcacatcctcgcc	81	gccacatagtgcgtgttgagca	82
MALAT1	gaagaaggaaggagcgctaa	126	cctgctaccttcatcaccaa	127

[0224] RNA 헤어핀 구조: GGCGUCCCUCCCGAAGCUGCGCGCUCGGUCGAACAGGACGACC(서열번호: 83)

[0225] 뉴클레오린 인식 서열: UCCCGA (서열번호: 84).

[0226] RNA 헤어핀 구조: GGCCGAAAUCCCGAAUGAGGCC (서열번호: 85) 및 GGAUGCCUCCCGAGUGCAUCC (서열번호: 86).

표 9

[0227]

이름:	설명:	폴리펩타이드 서열
CIRTS-0	ORF5-TBP6.7-GGS_6-데드 Pin 도메인	MDDPSFLTGRSTYAKRRRARRMNVCKCGAILHNNKDCRSSTISGHKLDRLRFVKEGRVALEGETPVYRTWVKWVETEYHINILETSDDEEGGSGGSGGSGGSGGSGGSMAVPETRPNHTIYINNLNSKIKKDELKKSLYAIFSQFGQILDILVPRQRTPRGQAFVIFKEVSSATNALRSMQGFPFYDKPMRIQYAKTDKRIPAKMKGTFVGSLQLPPLERLTLGGSGGSGGSGGSGGSGGSMELEIRPLFLVPDTNGFIDHLASLARLLESRKYILVVPLIVINELDGLAKGQETDHRAGGYARVVQEKARKSIEFLEQRFESRDSCLRALTSRGNELESIAFRSEDITGQLGNNADLILSCCLHYCKDKAKDFMPASKEEPIRLLREVVLLTDDRNLRVKALTRNVPVRDIPAFLTWAQVGGSDYKDHDGDYKDHDIDYKDDDDK (서열번호: 87)
CIRTS-1	ORF5-TBP6.7-GGS_6-Pin 도메인-NLS	MDDPSFLTGRSTYAKRRRARRMNVCKCGAILHNNKDCRSSTISGHKLDRLRFVKEGRVALEGETPVYRTWVKWVETEYHINILETSDDEEGGSGGSGGSGGSGGSGGSMAVPETRPNHTIYINNLNSKIKKDELKKSLYAIFSQFGQILDILVPRQRTPRGQAFVIFKEVSSATNALRSMQGFPFYDKPMRIQYAKTDKRIPAKMKGTFVGGSGGSGGSGGSGGSGGSMELEIRPLFLVPDTNGFIDHLASLARLLESRKYILVVPLIVINELDGLAKGQETDHRAGGYARVVQEKARKSIEFLEQRFESRDSCLRALTSRGNELESIAFRSEDITGQLGNNDDLILSCCLHYCKDKAKDFMPASKEEPIRLLREVVLLTDDRNLRVKALTRNVPVRDIPAFLTWAQVGKRPAATKKAGQAKKKKGSDYKDDDDDK (서열번호: 88)

CIRTS-2	ORF5-TBP6.7-GGS$_6$-NYTHDF1	MDDPSFLTGRSTYAKRRRARRMNVCKCGAILHNNKDCRSSTISGHKLDRLRFVKEGRVALEGETPVYRTWVK WVETEYHINILETSDDEEGGSGGSGGSGGSGGSGGSMAVPETRPNHTIYINNLNSKIKKDELKKSLYAIFSQ FGQILDILVPRQRTPRGQAFVIFKEVSSATNALRSMQGFPFYDKPMRIQYAKTDKRIPAKMKGTFVGSLQLP PLERLTLGGSGGSGGSGGSGGSGGSMSATSVDTQRTKGQDNKVQNGSLHQKDTVHDNDFEPYLTGQSNQSNS YPSMSDPYLSSYYPPSIGFPYSLNEAPWSTAGDPPIPYLTTYGQLSNGDHHFMHDAVFGQPGGLGNNIYQHR FNFFPENPAFSAWGTSGSQGQQTQSSAYGSSYTYPPSSLGGTVVDGQPGFHSDTLSKAPGMNSLEQGMVGLK IGDVSSSAVKTVGSVVSSVALTGVLSGNGGTNVNMPVSKPTSWAAIASKPAKPQPKMKTKSGPVMGGGLPPP PIKHNMDIGTWDNKGPVPKAPVPQQAPSPQAAPQPQQVAQPLPAQPPALAQPQYQSPQQPPQTRWVAPRNRN AAFGQSGGAGSDSNSPGNVQPNSAPSVESGSDYKDDDDK (서열번호: 89)
CIRTS-3$_t$	ORF5-TBP6.7-GGS$_6$-YTHDF2 (100-200)	MDDPSFLTGRSTYAKRRRARRMNVCKCGAILHNNKDCRSSTISGHKLDRLRFVKEGRVALEGETPVYRTWVK WVETEYHINILETSDDEEGGSGGSGGSGGSGGSGGSMAVPETRPNHTIYINNLNSKIKKDELKKSLYAIFSQ FGQILDILVPRQRTPRGQAFVIFKEVSSATNALRSMQGFPFYDKPMRIQYAKTDKRIPAKMKGTFVGSLQLP PLERLTLGGSGGSGGSGGSGGSGGSDAMFGQPGALGSTPFLGQHGFNFFPSGIDFSAWGNNSSQGQSTQSSG YSSNYAYAPSSLGGAMIDGQSAFANETLNKAPGMNTIDQGMAAGSDYKDDDDK (서열번호: 90)
CIRTS-3	ORF5-TBP6.7-GGS$_6$-YTHDF2 (1-200)	MDDPSFLTGRSTYAKRRRARRMNVCKCGAILHNNKDCRSSTISGHKLDRLRFVKEGRVALEGETPVYRTWVK WVETEYHINILETSDDEEGGSGGSGGSGGSGGSGGSMAVPETRPNHTIYINNLNSKIKKDELKKSLYAIFSQ FGQILDILVPRQRTPRGQAFVIFKEVSSATNALRSMQGFPFYDKPMRIQYAKTDKRIPAKMKGTFVGSLQLP PLERLTLGGSGGSGGSGGSGGSGGSMSASSLLEQRPKGQGNKVQNGSVHQKDGLNDDDFEPYLSPQARPNNA YTAMSDSYLPSYYSPSIGFSYSLGEAAWSTGGDTAMPYLTSYGQLSNGEPHFLPDAMFGQPGALGSTPFLGQ HGFNFFPSGIDFSAWGNNSSQGQSTQSSGYSSNYAYAPSSLGGAMIDGQSAFANETLNKAPGMNTIDQGMAA GSDYKDDDDK (서열번호: 91)
CIRTS-4$_t$	ORF5-SLBP-GGS$_6$-YTHDF2 (100-200)	MDDPSFLTGRSTYAKRRRARRMNVCKCGAILHNNKDCRSSTISGHKLDRLRFVKEGRVALEGETPVYRTWVK WVETEYHINILETSDDEEGGSGGSGGSGGSGGSGGSADFETDESVLMRRQKQINYGKNTIAYDRYIKEVPRH LRQPGIHPKTPNKFKKYSRRSWDQQIKLWKVALHFWDLQLPPLERLTLGGSGGSGGSGGSGGSGGSDAMFGQ PGALGSTPFLGQHGFNFFPSGIDFSAWGNNSSQGQSTQSSGYSSNYAYAPSSLGGAMIDGQSAFANETLNKA PGMNTIDQGMAAGSDYKDDDDK (서열번호: 92)
CIRTS-4	ORF5-SLBP-GGS$_6$-YTHDF2 (1-200)	MDDPSFLTGRSTYAKRRRARRMNVCKCGAILHNNKDCRSSTISGHKLDRLRFVKEGRVALEGETPVYRTWVK WVETEYHINILETSDDEEGGSGGSGGSGGSGGSGGSADFETDESVLMRRQKQINYGKNTIAYDRYIKEVPRH LRQPGIHPKTPNKFKKYSRRSWDQQIKLWKVALHFWDLQLPPLERLTLGGSGGSGGSGGSGGSGGSMSASSL LEQRPKGQGNKVQNGSVHQKDGLNDDDFEPYLSPQARPNNAYTAMSDSYLPSYYSPSIGFSYSLGEAAWSTG GDTAMPYLTSYGQLSNGEPHFLPDAMFGQPGALGSTPFLGQHGFNFFPSGIDFSAWGNNSSQGQSTQSSGYS SNYAYAPSSLGGAMIDGQSAFANETLNKAPGMNTIDQGMAAGSDYKDDDDK (서열번호: 93)
CIRTS-5$_t$	HBEGF-TBP6.7-GGS$_6$-YTHDF2 (100-200)	MRVTLSSKPQALATPNKEEHGKRKKKGKGLGKKRDPCLRKYKDFCIHGECKYVKELRAPSCICHPGYHGERC HGLSGGSGGSGGSGGSGGSGGSMAVPETRPNHTIYINNLNSKIKKDELKKSLYAIFSQFGQILDILVPRQRT PRGQAFVIFKEVSSATNALRSMQGFPFYDKPMRIQYAKTDKRIPAKMKGTFVGSLQLPPLERLTLGGSGGSG GSGGSGGSGGSDAMFGQPGALGSTPFLGQHGFNFFPSGIDFSAWGNNSSQGQSTQSSGYSSNYAYAPSSLGG AMIDGQSAFANETLNKAPGMNTIDQGMAAGSDYKDDDDK (서열번호: 94)
CIRTS-5	HBEGF-TBP6.7-GGS$_6$-YTHDF2 (1-200)	MRVTLSSKPQALATPNKEEHGKRKKKGKGLGKKRDPCLRKYKDFCIHGECKYVKELRAPSCICHPGYHGERC HGLSGGSGGSGGSGGSGGSGGSMAVPETRPNHTIYINNLNSKIKKDELKKSLYAIFSQFGQILDILVPRQRT PRGQAFVIFKEVSSATNALRSMQGFPFYDKPMRIQYAKTDKRIPAKMKGTFVGSLQLPPLERLTLGGSGGSG GSGGSGGSGGSMSASSLLEQRPKGQGNKVQNGSVHQKDGLNDDDFEPYLSPQARPNNAYTAMSDSYLPSYYS PSIGFSYSLGEAAWSTGGDTAMPYLTSYGQLSNGEPHFLPDAMFGQPGALGSTPFLGQHGFNFFPSGIDFSA WGNNSSQGQSTQSSGYSSNYAYAPSSLGGAMIDGQSAFANETLNKAPGMNTIDQGMAAGSDYKDDDDK (서열번호: 95)
CIRTS-6$_t$	b-디펜신-TBP6.7-GGS$_6$-YTHDF2 (100-200)	MGIINTLQKYYCRVRGGRCAVLSCLPKEEQIGKCSTRGRKCCRRKKGGSGGSGGSGGSGGSGGSMAVPETRP NHTIYINNLNSKIKKDELKKSLYAIFSQFGQILDILVPRQRTPRGQAFVIFKEVSSATNALRSMQGFPFYDK PMRIQYAKTDKRIPAKMKGTFVGSLQLPPLERLTLGGSGGSGGSGGSGGSGGSDAMFGQPGALGSTPFLGQH GFNFFPSGIDFSAWGNNSSQGQSTQSSGYSSNYAYAPSSLGGAMIDGQSAFANETLNKAPGMNTIDQGMAAG SDYKDDDDK (서열번호: 96)
CIRTS-6	b-디펜신-TBP6.7-GGS$_6$-YTHDF2 (1-200)	MGIINTLQKYYCRVRGGRCAVLSCLPKEEQIGKCSTRGRKCCRRKKGGSGGSGGSGGSGGSGGSMAVPETRP NHTIYINNLNSKIKKDELKKSLYAIFSQFGQILDILVPRQRTPRGQAFVIFKEVSSATNALRSMQGFPFYDK PMRIQYAKTDKRIPAKMKGTFVGSLQLPPLERLTLGGSGGSGGSGGSGGSGGSMSASSLLEQRPKGQGNKVQ NGSVHQKDGLNDDDFEPYLSPQARPNNAYTAMSDSYLPSYYSPSIGFSYSLGEAAWSTGGDTAMPYLTSYGQ LSNGEPHFLPDAMFGQPGALGSTPFLGQHGFNFFPSGIDFSAWGNNSSQGQSTQSSGYSSNYAYAPSSLGGA MIDGQSAFANETLNKAPGMNTIDQGMAAGSDYKDDDDK (서열번호: 97)
CIRTS-7	b-디펜신-SLBP-GGS$_6$-YTHDF2 (1-200)	MGIINTLQKYYCRVRGGRCAVLSCLPKEEQIGKCSTRGRKCCRRKKGGSGGSGGSGGSGGSGGSADFETDES VLMRRQKQINYGKNTIAYDRYIKEVPRHLRQPGIHPKTPNKFKKYSRRSWDQQIKLWKVALHFWDLQLPPLE RLTLGGSGGSGGSGGSGGSGGSMSASSLLEQRPKGQGNKVQNGSVHQKDGLNDDDFEPYLSPQARPNNAYTA MSDSYLPSYYSPSIGFSYSLGEAAWSTGGDTAMPYLTSYGQLSNGEPHFLPDAMFGQPGALGSTPFLGQHGF NFFPSGIDFSAWGNNSSQGQSTQSSGYSSNYAYAPSSLGGAMIDGQSAFANETLNKAPGMNTIDQGMAAGSD YKDDDDK (서열번호: 98)
CIRTS-8	TBP6.7-GGS$_6$-Pin 도메인-NES	MAVPETRPNHTIYINNLNSKIKKDELKKSLYAIFSQFGQILDILVPRQRTPRGQAFVIFKEVSSATNALRSM QGFPFYDKPMRIQYAKTDKRIPAKMKGTFVGSLQLPPLERLTLGGSGGSGGSGGSGGSGGSMELEIRPLFLV PDTNGFIDHLASLARLLESRKYILVVPLIVINELDGLAKGQETDHRAGGYARVVQEKARKSIEFLEQRFESR DSCLRALTSRGNELESIAFRSEDITGQLGNNDDLILSCCLHYCKDKAKDFMPASKEEPIRLLREVVLLTDDR NLRVKALTRNVPVRDIPAFLTWAQVGGSDYKDDDDK (서열번호: 99)

CIRTS-9	ORF5-TBP6.7-GGS$_6$-Pin 도메인-NES	MDDPSFLTGRSTYAKRRRARRMNVCKCGAILHNNKDCRSSTISGHKLDRLRFVKEGRVALEGETPVYRTWVK WVETEYHINILETSDDEEGGSGGSGGSGGSGGSGGSMAVPETRPNHTIYINNLNSKIKKDELKKSLYAIFSQ FGQILDILVPRQRTPRGQAFVIFKEVSSATNALRSMQGFPFYDKPMRIQYAKTDKRIPAKMKGTFVGSLQLP PLERLTLGGSGGSGGSGGSGGSGGSMELEIRPLFLVPDTNGFIDHLASLARLLESRKYILVVPLIVINELDG LAKGQETDHRAGGYARVVQEKARKSIEFLEQRFESRDSCLRALTSRGNELESIAFRSEDITGQLGNNDDLIL SCCLHYCKDKAKDFMPASKEEPIRLLREVVLLTDDRNLRVKALTRNVPVRDIPAFLTWAQVGGSDYKDDDDK (서열번호: 100)
CIRTS-10	ORF5-TBP6.7-GGS$_3$-Pin 도메인	MDDPSFLTGRSTYAKRRRARRMNVCKCGAILHNNKDCRSSTISGHKLDRLRFVKEGRVALEGETPVYRTWVK WVETEYHINILETSDDEEGGSGGSGGSGGSGGSGGSMAVPETRPNHTIYINNLNSKIKKDELKKSLYAIFSQ FGQILDILVPRQRTPRGQAFVIFKEVSSATNALRSMQGFPFYDKPMRIQYAKTDKRIPAKMKGTFVGSLQLP PLERLTLGGSGGSGGSMELEIRPLFLVPDTNGFIDHLASLARLLESRKYILVVPLIVINELDGLAKGQETDH RAGGYARVVQEKARKSIEFLEQRFESRDSCLRALTSRGNELESIAFRSEDITGQLGNNDDLILSCCLHYCKD KAKDFMPASKEEPIRLLREVVLLTDDRNLRVKALTRNVPVRDIPAFLTWAQVGGSDYKDDDDK (서열번호: 101)
CIRTS-11	ORF5-TBP6.7-L8-Pin 도메인	MDDPSFLTGRSTYAKRRRARRMNVCKCGAILHNNKDCRSSTISGHKLDRLRFVKEGRVALEGETPVYRTWVK WVETEYHINILETSDDEEGGSGGSGGSGGSGGSGGSMAVPETRPNHTIYINNLNSKIKKDELKKSLYAIFSQ FGQILDILVPRQRTPRGQAFVIFKEVSSATNALRSMQGFPFYDKPMRIQYAKTDKRIPAKMKGTFVGSLQLP PLERLTLSGSETPGTSESATPESMELEIRPLFLVPDTNGFIDHLASLARLLESRKYILVVPLIVINELDGLA KGQETDHRAGGYARVVQEKARKSIEFLEQRFESRDSCLRALTSRGNELESIAFRSEDITGQLGNNDDLILSC CLHYCKDKAKDFMPASKEEPIRLLREVVLLTDDRNLRVKALTRNVPVRDIPAFLTWAQVGGSDYKDDDDK (서열번호: 102)
CIRTS-12	ORF5-TBP6.7-나선-Pin 도메인-NES	MDDPSFLTGRSTYAKRRRARRMNVCKCGAILHNNKDCRSSTISGHKLDRLRFVKEGRVALEGETPVYRTWVK WVETEYHINILETSDDEEGGSGGSGGSGGSGGSGGSMAVPETRPNHTIYINNLNSKIKKDELKKSLYAIFSQ FGQILDILVPRQRTPRGQAFVIFKEVSSATNALRSMQGFPFYDKPMRIQYAKTDKRIPAKMKGTFVGSLQLP PLERLTLGGSGEEEEKKKQQEEEAERLRRIQEEMEKERKRREEDEQRRRKEEEERRMKLEMEAKRKQEEEER KKREDDEKRKKKGSGGSMELEIRPLFLVPDTNGFIDHLASLARLLESRKYILVVPLIVINELDGLAKGQETD HRAGGYARVVQEKARKSIEFLEQRFESRDSCLRALTSRGNELESIAFRSEDITGQLGNNDDLILSCCLHYCK DKAKDFMPASKEEPIRLLREVVLLTDDRNLRVKALTRNVPVRDIPAFLTWAQVGGSDYKDDDDK (서열번호: 103)
CIRTS-13	ORF5-TBP6.7-L8-Y2(100-200)	MDDPSFLTGRSTYAKRRRARRMNVCKCGAILHNNKDCRSSTISGHKLDRLRFVKEGRVALEGETPVYRTWVK WVETEYHINILETSDDEEGGSGGSGGSGGSGGSGGSMAVPETRPNHTIYINNLNSKIKKDELKKSLYAIFSQ FGQILDILVPRQRTPRGQAFVIFKEVSSATNALRSMQGFPFYDKPMRIQYAKTDKRIPAKMKGTFVGSLQLP PLERLTLSGSETPGTSESATPESDAMFGQPGALGSTPFLGQHGFNFFPSGIDFSAWGNNSSQGQSTQSSGYS SNYAYAPSSLGGAMIDGQSAFANETLNKAPGMNTIDQGMAAGSDYKDDDDK (서열번호: 104)
CIRTS-14	ORF5-TBP6.7-나선-Y2(100-200)	MDDPSFLTGRSTYAKRRRARRMNVCKCGAILHNNKDCRSSTISGHKLDRLRFVKEGRVALEGETPVYRTWVK WVETEYHINILETSDDEEGGSGGSGGSGGSGGSGGSMAVPETRPNHTIYINNLNSKIKKDELKKSLYAIFSQ FGQILDILVPRQRTPRGQAFVIFKEVSSATNALRSMQGFPFYDKPMRIQYAKTDKRIPAKMKGTFVGSLQLP PLERLTLGGSGEEEEKKKQQEEEAERLRRIQEEMEKERKRREEDEQRRRKEEEERRMKLEMEAKRKQEEEER KKREDDEKRKKKGSGGSDAMFGQPGALGSTPFLGQHGFNFFPSGIDFSAWGNNSSQGQSTQSSGYSSNYAYA PSSLGGAMIDGQSAFANETLNKAPGMNTIDQGMAAGSDYKDDDDK (서열번호: 105)
MBP-CIRTS-1	MBP-ORF5-TBP6.7-GGS$_6$-데드 Pin 도메인	MDDPSFLTGRSTYAKRRRARRMNVCKCGAILHNNKDCRSSTISGHKLDRLRFVKEGRVALEGETPVYRTWVK WVETEYHINILETSDDEEGGSGGSGGSGGSGGSGGSMAVPETRPNHTIYINNLNSKIKKDELKKSLYAIFSQ FGQILDILVPRQRTPRGQAFVIFKEVSSATNALRSMQGFPFYDKPMRIQYAKTDKRIPAKMKGTFVGSLQLP PLERLTLGGSGGSGGSGGSGGSGGSMELEIRPLFLVPDTNGFIDHLASLARLLESRKYILVVPLIVINELDG LAKGQETDHRAGGYARVVQEKARKSIEFLEQRFESRDSCLRALTSRGNELESIAFRSEDITGQLGNNDDLIL SCCLHYCKDKAKDFMPASKEEPIRLLREVVLLTDDRNLRVKALTRNVPVRDIPAFLTWAQVGSSGHHHHHH (서열번호: 106)
CIRTS-7X	b-디펜신 3-TBP6.7-GGS$_6$-hADAR(299-701)	MGIINTLQKYYCRVRGGRCAVLSCLPKEEQIGKCSTRGRKCCRRKKGGSGGSGGSGGSGGSGGSMAVPETRP NHTIYINNLNSKIKKDELKKSLYAIFSQFGQILDILVPRQRTPRGQAFVIFKEVSSATNALRSMQGFPFYDK PMRIQYAKTDKRIPAKMKGTFVGGSGGSGGSGGSGGSGGSLHLDQTPSRQPIPSEGLQLHLPQVLADAVSRL VLGKFGDLTDNFSSPHARRKVLAGVVMTTGTDVKDAKVISVSTGTKCINGEYMSDRGLALNDCHAEIISRRS LLRFLYTQLELYLNNKDDQKRSIFQKSERGGFRLKENVQFHLYISTSPCGDARIFSPHEPILEEPADRHPNR KARGQLRTKIESGEGTIPVRSNASIQTWDGVLQGERLLTMSCSDKIARWNVVGIQGSLLSIFVEPIYFSSII LGSLYHGDHLSRAMYQRISNIEDLPPLYTLNKPLLSGISNAEARQPGKAPNFSVNWTVGDSAIEVINATTGK DELGRASRLCKHALYCRWMRVHGKVPSHLLRSKITKPNVYHESKLAAKEYQAAKARLFTAFIKAGLGAWVEK PTEQDQFSLTPKRPAATKKAGQAKKKKGSDYKDDDDK (서열번호: 128)
CIRTS-8X	b-디펜신 3-TBP6.7-GGS$_6$-hADAR (299-701) E488Q	MGIINTLQKYYCRVRGGRCAVLSCLPKEEQIGKCSTRGRKCCRRKKGGSGGSGGSGGSGGSGGSMAVPETRP NHTIYINNLNSKIKKDELKKSLYAIFSQFGQILDILVPRQRTPRGQAFVIFKEVSSATNALRSMQGFPFYDK PMRIQYAKTDKRIPAKMKGTFVGGSGGSGGSGGSGGSGGSLHLDQTPSRQPIPSEGLQLHLPQVLADAVSRL VLGKFGDLTDNFSSPHARRKVLAGVVMTTGTDVKDAKVISVSTGTKCINGEYMSDRGLALNDCHAEIISRRS LLRFLYTQLELYLNNKDDQKRSIFQKSERGGFRLKENVQFHLYISTSPCGDARIFSPHEPILEEPADRHPNR KARGQLRTKIESGQGTIPVRSNASIQTWDGVLQGERLLTMSCSDKIARWNVVGIQGSLLSIFVEPIYFSSII LGSLYHGDHLSRAMYQRISNIEDLPPLYTLNKPLLSGISNAEARQPGKAPNFSVNWTVGDSAIEVINATTGK DELGRASRLCKHALYCRWMRVHGKVPSHLLRSKITKPNVYHESKLAAKEYQAAKARLFTAFIKAGLGAWVEK PTEQDQFSLTPKRPAATKKAGQAKKKKGSDYKDDDDK (서열번호: 129)

CIRTS-9X	b-디펜신 3-SLBP-GGS_6-NYTHDF1	MGIINTLQKYYCRVRGGRCAVLSCLPKEEQIGKCSTRGRKCCRRKKGGSGGSGGSGGSGGSGGSMAVPETRPNHTIYINNLNSKIKKDELKKSLYAIFSQFGQILDILVPRQRTPRGQAFVIFKEVSSATNALRSMQGFPFYDKPMRIQYAKTDKRIPAKMKGTFVGSLQLPPLERLTLGGSGGSGGSGGSGGSGGSMSATSVDTQRTKGQDNKVQNGSLHQKDTVHDNDFEPYLTGQSNQSNSYPSMSDPYLSSYYPPSIGFPYSLNEAPWSTAGDPPIPYLTTYGQLSNGDHHFMHDAVFGQPGGLGNNIYQHRFNFFPENPAFSAWGTSGSQGQQTQSSAYGSSYTYPPSSLGGTVVDGQPGFHSDTLSKAPGMNSLEQGMVGLKIGDVSSSAVKTVGSVVSSVALTGVLSGNGGTNVNMPVSKPTSWAAIASKPAKPQPKMKTKSGPVMGGGLPPPPIKHNMDIGTWDNKGPVPKAPVPQQAPSPQAAPQPQQVAQPLPAQPPALAQPQYQSPQQPPQTRWVAPRNRNAAFGQSGGAGSDSNSPGNVQPNSAPSVESGSDYKDDDDK (서열번호: 130)
CIRTS-10X	b-디펜신 3-SLBP-GGS_6-YTHDF2 (1-200)	MGIINTLQKYYCRVRGGRCAVLSCLPKEEQIGKCSTRGRKCCRRKKGGSGGSGGSGGSGGSGGSADFETDESVLMRRQKQINYGKNTIAYDRYIKEVPRHLRQPGIHPKTPNKFKKYSRRSWDQQIKLWKVALHFWDLQLPPLERLTLGGSGGSGGSGGSGGSGGSMSASSLLEQRPKGQGNKVQNGSVHQKDGLNDDDFEPYLSPQARPNNAYTAMSDSYLPSYYSPSIGFSYSLGEAAWSTGGDTAMPYLTSYGQLSNGEPHFLPDAMFGQPGALGSTPFLGQHGFNFFPSGIDFSAWGNNSSQGQSTQSSGYSSNYAYAPSSLGGAMIDGQSAFANETLNKAPGMNTIDQGMAAGSDYKDDDDK (서열번호: 131)
CIRTS-18X	ORF5-PP7-Y2(100-200)	MDDPSFLTGRSTYAKRRRARRMNVCKCGAILHNNKDCRSSTISGHKLDRLRFVKEGRVALEGETPVYRTWVKWVETEYHINILETSDDEEGGSGGSGGSGGSGGSGGSMAKTIVLAVGEATRTLTEIQSTADRQIFEEKVGPLVGRLRLTASLRQNGAKTAYRVNLKLDQADVVDASTSVAGELPKVRYTQVWSHDVTIVANSTEASRKSLYDLTKSLVATSQVEDLVVNLVPLGRSLEGGSGGMAKTIVLAVGEATRTLTEIQSTADRQIFEEKVGPLVGRLRLTASLRQNGAKTAYRVNLKLDQADVVDASTSVAGELPKVRYTQVWSHDVTIVANSTEASRKSLYDLTKSLVATSQVEDLVVNLVPLGRSLELQLPPLERLTLGGSGGSGGSGGSGGSGGSDAMFGQPGALGSTPFLGQHGFNFFPSGIDFSAWGNNSSQGQSTQSSGYSSNYAYAPSSLGGAMIDGQSAFANETLNKAPGMNTIDQGMAAGSDYKDDDDK (서열번호: 132)

[0228] VII.실시예들

[0229] 하기 실시예들은 본 개시 내용의 바람직한 구현예를 설명하기 위해 포함된다. 하기 실시예들에 개시된 기술은 본 개시 내용의 실시에서 잘 기능하는 발명자들에 의해 발견된 기술을 나타내고, 따라서 이의 실시를 위한 바람직한 모드를 구성하는 것으로 간주될 수 있다는 것이 당업자에 의해 이해되어야 한다. 그러나, 당업자는 본 개시 내용에 비추어, 개시되는 특정 구현예에서 많은 변경이 이루어질 수 있고 개시 내용의 정신 및 범위를 벗어나지 않고 여전히 유사하거나 유사한 결과를 얻을 수 있음을 인식해야 한다.

[0230] 실시예 1 - 인간 부분으로부터 구축된 프로그램 가능한 RNA-가이드 RNA 효과기 단백질

[0231] A.도입

[0232] 현재 RNA 표적화 시스템의 큰 크기와 미생물 유래 성질을 극복하기 위해, 본 발명자들은 프로그램 가능한 RNA 효과기 단백질을 조작하기 위한 일반적인 방법인 CRISPR/Cas-영감 RNA 표적화 시스템(CIRTS)을 제시한다. 본 발명자들은 CIRTS가 기능적 부분에 대한 인간 단백질체(proteome)를 채굴하여 프로그램 가능한 RNA 조절 단백질을 구축할 수 있음을 보여준다. CIRTS는 왓슨-크릭-프랭클린(Watson-Crick-Franklin) 염기쌍 상호작용을 사용하여 전사체에서 단백질 카고(protein cargo) 부위를 선택적으로 전달하는 리보핵산단백질 복합체이다. 본 발명자들은 분해를 위한 뉴클레아제, 분해를 위한 탈아데닐화 조절 기구, 또는 강화된 단백질 생산을 위한 번역 활성화 기구를 포함하는 전사체에 일정 범위 또는 조절 단백질을 전달하는 CIRTS를 쉽게 조작할 수 있음을 보여준다. 그러나 CIRTS는 현재 가장 작은 CRISPR/Cas 시스템보다 최대 5배 작으며 전적으로 인간 부분으로부터 조작될 수 있다.

[0233] 본 발명자들은 최소 프로그램 가능한 RNA-표적화 시스템은 다음 성분들: 즉, (1) 시스템의 코어 역할을 하고 조작된 gRNA에 디스플레이되는 특정 RNA 구조에 대한 선택적 고친화성 결합제인 RNA 헤어핀 결합 단백질, (2) 조작된 헤어핀 결합 단백질과 상호작용하는 구조 및 관심 표적 RNA에 상보성을 갖는 서열 모두를 특징으로 하는 gRNA, 및 (3) 근접 의존적 방식으로 표적화된 RNA에 작용하는 뉴클레아제 또는 후성전사체 조절인자와 같은 효과기 단백질,을 필요로 할 것이라고 추론했다. 일부 구현예에서, 디스플레이된 gRNA 서열에 비특이적으로 결합하는 하전된 단백질은 표적 결합 전에 가이드 RNA를 안정화 및 보호하기 위해 사용된다. (**도 1a**). CIRTS는 조작된 시스템에서 이러한 기능을 수행하는 여러 단백질 도메인을 결합한다.

[0234] 여기에서 본 발명자들은 CIRTS의 설계 및 검증을 제시한다. 먼저, 본 발명자들은 *시험관 내* 및 포유동물 세포 리포터 분석 최적화 및 검증 모두에 사용되는 프로그램 가능한 CIRTS 리보뉴클레아제를 조작했다. 그런 다음 본 발명자들은 gRNA, 헤어핀 결합 단백질, ssRNA 결합 단백질 및 CIRTS-1의 효과기 도메인을 포함하는 4가지 구성 부분 모두가 다른 부분으로 대체될 수 있으며 5개의 추가 CIRTS가 발생될 수 있음을 보임으로써 CIRTS의 다목적성을 보여준다(**도 1b**). 본 발명자들은 프로그램 가능한 m^6A 판독기 단백질로서 조작된 CIRTS를 사용하여 분해

기구 및 번역 활성화를 포함하는 내인성 후성전사체 조절 경로를 표적화할 수 있다. 또한, 본 발명자들은 CIRTS를 사용하여 근접-의존성 뉴클레아제-매개 붕괴, 비-뉴클레아제-매개 붕괴 또는 번역 활성화를 위해 내인성 전사물을 표적화할 수 있음을 보여준다. 마지막으로, 본 발명자들은 직교 CIRTS를 사용하여 여러 유전자를 동시에 표적화할 수 있음을 보여준다. 종합하면, 이 작업은 CIRTS 전략이 작고 인간 부분에서 결집된 RNA 효과기 단백질을 조작하기 위한 실행 가능한 새로운 접근 방식으로 검증한다.

[0235] B.결과

[0236] 1.CIRTS-1의 개발 및 시험관 내 검증

[0237] 본 발명자의 1세대 시스템인 CIRTS-1의 경우, 본 발명자들은 HIV TAR 헤어핀에 결합하도록 이전에 조작되었고 내인성 인간 RNA 헤어핀 표적을 갖지 않는, 인간 헤어핀-결합 단백질 U1A 단백질(TBP6.7)의 진화된 버전을 사용했다(Blakeley and McNaughton, 2014; Crawford et al., 2016)(**도 1b**). 본 발명자들은 TAR 헤어핀 구조, 뉴클레오타이드 링커 서열(L) 및 가이드 서열을 포함하는 gRNA를 설계했다(**도 1c**). 시스템을 개발하고 검증하기 위해, 본 발명자들은 먼저 TBP6.7을 이전에 비특이적 근접 의존적 RNA 엔도뉴클레아제로 사용된, 인간 SMG6의 Pin 뉴클레아제 도메인에 융합하여 프로그램 가능한 뉴클레아제를 조작했다(Batra et al., 2017; Choudhury et al., 2012). 이 가장 단순한 설계가 이미 세포 기반 분석에서 gRNA 매개 전사물 분해를 디스플레이했지만(**도 9b, 왼쪽**), 성능은 상당히 안 좋았고, 본 발명자들은 이를 디스플레이된 표적화 서열의 잠재적인 분해의 결과로 보았다. Cas13 시스템의 단백질 표면과 헤어핀 채널은 높게 하전되는 경향이 있어, 가이드 RNA 서열에 비특이적으로 결합하고 이를 안정화할 가능성이 있다(Liu et al., 2017). 이 RNA 보호 기능을 본 발명자의 시스템에 조작하기 위해, 본 발명자들은 단일 가닥 RNA에 대해 비특이적이고 낮은 친화성을 갖는 RNA 결합 단백질을 추가하는 것을 목표로 하였다. 그러나 인간 단백질 툴박스는 주해가 달린 작고 비특이적인 단일 가닥 RNA 결합 단백질을 쉽게 포함하지 않았다. 따라서, 본 발명자들은 작은 바이러스 ssRNA 결합 단백질인 ORF5를 사용하여 CIRTS-1을 개발했다(Zhou et al., 2006). 전체적으로, CIRTS-1은 상응하는 gRNA와 함께 ORF5-TBP6.7-Pin 뉴클레아제 도메인의 단백질 융합 복합체로 구성된다(도 1a).

[0238] 본 발명자들은 먼저 모델 RNA 표적 기질에 대한 *시험관 내* CIRTS-1의 프로그램 가능한 RNA 결합 및 RNA 뉴클레아제 활성을 특성화하였다. 상기 뉴클레아제를 비활성화하기 위해 EDTA의 존재 하에 실행되는 전기영동 이동성 변화 분석(EMSA)에서 정제된 MBP-CIRTS-1 단백질 및 gRNA를 사용하여, 본 발명자들은 MBP-CIRTS-1이 105nM의 겉보기 K_D로 gRNA-의존적 방식으로 표적 RNA에 결합함을 발견하였다(**도 2a 및 2b**). gRNA(**도 7a**) 또는 CIRTS-1 단백질(**도 7b**) 중 하나가 분석에서 제외되면, 측정 가능한 결합이 검출되지 않아, 시스템에서 설계된 바와 같이 gRNA-의존적 결합 모드가 확인된다. 또한, 절단 분석에서, 본 발명자들은 CIRTS-1이 gRNA- 및 Mn^{2+}-의존적 방식으로 RNA 기질을 분해한다는 것을 발견하였다(**도 2c**). 종합적으로, 이러한 *시험관 내* 결과는 CIRTS-1 이면의 설계 원리를 검증하고 발명자들이 살아있는 세포에서 사용하기 위해 시스템을 최적화하려고 시도하도록 동기를 부여했다.

[0239] 2.살아있는 세포에서 CIRTS-1의 최적화

[0240] 살아있는 포유동물 세포에서 CIRTS-1의 표적 뉴클레아제 활성을 시험하기 위해, 본 발명자들은 표적 RNA 상의 gRNA-의존성 전사 변화에 대해 보고하는 이중 루시퍼라아제 리포터 분석을 확립했다(**도 3a**). 이 시스템을 사용하여, 본 발명자들은 상이한 단백질 링커 유형(**도 8a**), gRNA 구조(**도 8c**) 및 CIRTS-1 세포 국소화(**도 9c**)를 분석함으로써 CIRTS-1의 배치를 최적화하였다. 최적화 후, 본 발명자들은 표적 리포터 RNA를 분해하는 최적화된 CIRTS-1 시스템의 능력을 현재의 최첨단 Cas13 시스템과 비교하였다. 본 발명자들은 CIRTS-1 및 Cas13b 모두에 대한 이중 루시퍼라아제 리포터 분석에서 반딧불이 루시퍼라아제 mRNA를 표적으로 하는 gRNA는 물론, 각각의 프로그램 가능한 뉴클레아제에 대한 대조군, 표적외 gRNA를 설계했다. 전사물에 결합하는 것만으로도 전사물 수준이 변경되는지 여부를 시험하기 위해, 본 발명자들은 CIRTS-1의 뉴클레아제 도메인에 이전에 보고된 '죽은(dead)' 돌연변이를 함유하는(Eberle et al., 2009), CIRTS-0을 음성 대조군 역할을 하도록 조작했다. 본 발명자들은 CIRTS-0이 표적 전사물에 유의한 영향을 미치지 않는다는 것을 발견하였고(**도 3b 및 도 9a**), 이는 CIRTS 리보핵산단백질 표적 RNA 복합체가 관심 RNA에 대해 최소한으로 교란적이라는 것을 나타낸다. 다음으로, 본 발명자들은 활성 뉴클레아제를 갖는 CIRTS-1이 표적의 분해를 매개할 수 있는지 여부를 시험하였다. 실제로, 본 발명자들은 루시퍼라아제 활성에 의해 모니터링되는 단백질 수준(**도 3c**) 및 RT-qPCR에 의해 모니터링되는 mRNA 수준(**도 9d**) 모두에서 측정된 표적 유전자의 gRNA-의존성 분해를 발견하였다. 고무적으로, 본 발명자들은 CIRTS-1이 Cas13b에 의한 표적화된 뉴클레아제 분해와 비교하여 루시퍼라아제 리포터 유전자를 표적화하는데 있어서 약간 덜 효율적이라는 것을 발견하였고, 이는 CIRTS 전략이 살아있는 세포에서 실행 가능한 방법임을 나타

낸다. CIRTS-1의 성능에 고무된 본 발명가들은 다음으로 다양한 시스템 성분을 다른 부분으로 교체하고 CIRTS 기능을 유지할 수 있는지 여부를 시험하여 설계의 다목적성을 평가하려고 했다.

[0241] 3.CIRTS의 모듈성

[0242] CIRTS 설계의 다목적성을 탐구하기 위해, 본 발명자들은 다음으로 CIRTS 단백질 전달 시스템의 각 성분에 대해 상이한 단백질 도메인을 시험했다. 먼저, 본 발명자들은 CIRTS가 이전에 dCas13b 시스템을 사용하여 전달했던 RNA 후성전사체 조절 "판독자(reader)" 단백질을 전달할 수 있는지 여부를 분석했다.(Rauch et al., 2018). 본 발명자들은 CIRTS-1의 Pin 뉴클레아제 효과기 단백질을 번역 기구를 동원하는 세포질 N^6-메틸아데노신(m^6A) 판독자 단백질인 YTHDF1의 N-말단 도메인으로 교체하여, CIRTS-2를 생성하였다. CIRTS-2가 CIRTS-1 실험과 동일한 표적 서열에 전달되는 경우, RNA 수준은 상대적으로 변하지 않지만(도 9e), RNA로부터 단백질 수준의 상당한 증가가 발생한다(도 3d). 그런 다음, 본 발명자들은 YTHDF1 단편을 RNA 탈아데닐화 기구를 동원하고 RNA 분해를 유도하는 m^6A 판독자 단백질인 YTHDF2의 단편으로 교체하여, CIRTS-3을 생성하였다. 리포터 mRNA로 CIRTS-3을 전달하면 RNA(도 9f) 및 단백질 수준(도 3e) 모두에 의해 측정된 바와 같이 표적 전사물의 분해를 유도한다. CIRTS-1 내지 -3은 다양한 효과기 단백질 카고(effector protein cargo)를 살아있는 세포의 표적 RNA에 전달하는 설계 전략의 다목적성을 보여준다.

[0243] 효과기 도메인의 모듈성을 입증한 후, 본 발명자들은 다른 인간 부분이 RNA 헤어핀 결합 도메인 및 비특이적 ssRNA 결합 단백질에도 사용될 수 있는지 평가에 착수했다. 본 발명자들은 CIRTS-3의 TBP6.7을 인간 히스톤 줄기 루프 결합 단백질(SLBP)의 RNA 헤어핀 결합 도메인으로 대체하여 CIRTS-4를 생성하였다. 또한, 본 발명자들은 히스톤 mRNA 줄기 루프 구조를 기반으로 gRNA를 설계하였다(도 1d). 리포터 분석(도 3f)에서 그리고 RNA 수준을 평가하기 위한 RT-qPCR(도 9g)에 의한 CIRTS-4 분석은 CIRTS-3과 유사한 성능이 나타났다. 마지막으로, 본 발명자들은 CIRTS 시스템의 완전히 인간화된 버전을 조작하려고 하였다. 앞서 언급한 바와 같이, 본 발명자들은 바이러스 비특이적 단일 가닥 RNA 결합 단백질 ORF5를 기반으로 하는 첫 번째 개념 증명 시스템(proof-of-concept system)을 설계했다. 주해가 달린 인간 단일 가닥, 비특이적 RNA 결합 단백질은 없지만, 본 발명자들은 높게 하전된 양이온성 인간 단백질이 CIRTS 시스템에서 ORF5의 역할을 수행할 수 있다고 추론했다(Cronican et al., 2011). 따라서, 본 발명자들은 CIRTS-3에서 ORF5 대신에 2개의 양이온성 인간 단백질인 HBEGF 및 β-디펜신 3을 조작하여 CIRTS-5 및 CIRTS-6을 각각 생성하였다. 다시, 루시퍼라아제 리포터 분석에 이들 프로그램 가능한 효과기들을 배치하면 표적 유전자의 gRNA-의존성 분해가 나타났다(도 3g). 본 발명자들은 또한 qPCR에 의한 RNA 수준의 변화를 분석하고 CIRTS-5 및 CIRTS-6 모두에 대한 감소를 발견하였다(도 9h). 종합하면, 리포터 분석에서 CIRTS-1부터 CIRTS-6까지의 성능은 헤어핀 결합 도메인과 해당 gRNA, 단일 가닥 RNA 결합 단백질 및 효과기 단백질을 포함한 모든 차원에서 CIRTS 설계의 모듈성을 보여준다. 다음으로, 본 발명자들은 CIRTS가 내인성 전사물을 표적화하는 데에도 사용될 수 있는지 여부를 시험하고자 했다.

[0244] 4.CIRTS로 내인성 mRNA 표적화

[0245] 내인성 전사물을 선택적으로 표적화하기 위한 첫 번째 단계로서, 본 발명자들은 RNA 면역침강 후 RT-qPCR에 의한 분석에 의해 CIRTS-0이 표적 내인성 전사물에 결합할 수 있음을 확인하였다. 본 발명자들은 이전에 Cas13 시스템, PPIB 및 SMARCA4에 의해 표적화된 두 개의 내인성 전사물을 표적화하는 gRNA들을 설계했다(Cox et al., 2017; Konermann et al., 2018). 본 발명자들은 3x FLAG-태그에 융합된 CIRTS-0과 함께 각 gRNA를 개별적으로 전달하였다. 이어서, 본 발명자들은 항-Flag 항체로 면역침강된 전체 RNA를 단리하고, 단백질에 결합된 각각의 표적 RNA의 상대적인 양을 정량화하였다. 실제로, 내인성 전사물 둘 모두는 gRNA-의존적 방식으로 2.5배 내지 5배 농축되었고(도 10a), 이는 CIRTS가 내인성 전사물에서 프로그램 가능한 RNA-가이드 RNA 결합 단백질로서 기능할 수 있음을 입증한다.

[0246] 다음으로 본 발명자들은 CIRTS-1 프로그램 가능한 뉴클레아제 및 CIRTS-3 프로그램 가능한 YTHDF2-매개 붕괴 시스템을 예시로 사용하여 CIRTS 시스템이 표적 내인성 전사물에 효과기 단백질을 전달할 수 있는지 여부를 평가하고자 했다. 본 발명자들은 이전에 Cas13 표적으로 검증된 5개의 RNA 전사물을 선택하고, 이것들이 프로그램 가능한 RNA-결합 시스템에 의한 RNA 표적화에 접근 가능하다고 추론했다. 그런 다음 본 발명자들은 이전에 Cas13 실험에 사용된 표적상의 동일한 결합 부위를 사용하여 각 표적에 대한 gRNA를 설계했으며, 이러한 부위는 CIRTS 표적화에도 접근 가능할 수 있을 것이라고 상정했다(Konermann et al., 2018). 본 발명자들은 RT-qPCR에 의해 각 표적의 RNA 수준에 대한 CIRT 시스템의 효과를 분석하였다. 세포가 특정 gRNA 발현 벡터와 함께 CIRTS-1 또는 CIRTS-3으로 형질주입되었을 때, 본 발명자들은 5가지 내인성 전사물: PPIB, NFKB1, NRAS, B4GALTN1, 및 SMARCA4의 각각에 대해 RT-qPCR에 의한 RNA 수준의 상당한 감소를 보았다(도 4a 및 4b). 상대적 녹다운 효율

은 각 유전자에 따라 달랐으며, 이는 다른 RNA 표적화 시스템에서도 관찰되며 각 유전자에 특이적인 접근성 또는 기타 조절 경로에 의해 잠재적으로 매개된다. 그럼에도 불구하고, 이러한 결과는 CIRTS가 표적에 대한 활성 뉴클레아제 활성을 통해 또는 내인성 후성전사체 조절 경로를 촉발함으로써 내인성 전사물을 표적화하고 붕괴를 매개할 수 있음을 확인해준다.

[0247] 다음으로, 본 발명자들은 CIRTS-2가 YTHDF1-매개된 후성전사체 경로를 통해 내인성 전사물의 단백질 생산을 촉발할 수 있는지 여부를 평가하고자 했다. 본 발명자들은 웨스턴 블로팅(Western blotting)에 의한 단백질 생산 분석을 위해 보고되고 신뢰할 수 있는 항체가 있는 풍부한 전사물(CypB, PPIB의 단백질 생성물)를 선택했다. 실제로, CIRTS-2로 형질주입된 세포 및 표적내 gRNA는 RNA 수준(**도 10b**)의 변화 없이 단백질 수준(**도 4c**)의 증가를 나타내었고, 이는 전사물에 대한 YTHDF1 효과를 확인시켜주었다. 대조군으로서, YTHDF2를 전달하는 CIRTS-3으로 수행된 동일한 실험은 단백질 수준 감소로 나타났고, 이는 mRNA 수준의 감소와 상관관계가 있다(**도 3b**). 종합하면, 이러한 실험은 CIRTS 플랫폼이 gRNA 의존적 방식으로 내인성 전사물에서 기능적이며, 표적 전사물을 능동적으로 분해하거나, 분해 기구가 표적 전사물에 작용하도록 촉발하거나 또는 표적 전사물로부터 번역을 활성화하고 단백질 생산을 증가시킬 수 있음을 보여준다. CIRTS의 다목적성은 발명가들이 한 전사물 이상을 표적화하도록 그들을 동시에 배포하도록 영감을 주었다.

[0248] **5.CIRTS를 사용한 다중 내인성 RNA의 다중 표적화**

[0249] 마지막으로, 본 발명자들은 상이한 헤어핀 결합 도메인으로 조작된 CIRTS가 상이한 전사물을 선택적으로 표적화하기 위해 살아있는 세포에서 직교하게 기능할 수 있는지 여부를 평가하는 것을 목표로 하였다. 원칙적으로, 각각 별도로 조작된 gRNA(도 1c 및 1d)를 사용하는, TBP 헤어핀 결합 도메인으로부터 구축된 CIRTS와 SLBP 헤어핀 결합 도메인으로부터 구축된 CIRTS는 서로 직교해야 한다. 다중 표적화 CIRTS의 이러한 직교성 시험을 위해, 본 발명자들은 각각 상이한 헤어핀 구조에 결합하고 서로 간에 최소 혼선을 가져야 하는 YTHDF2, CIRTS-6 및 CIRTS-7을 각각 전달하는 2개의 완전히 인간화된 CIRTS를 배치했다. 본 발명자들은 PPIB를 표적화하기 위해 CIRTS-6을 사용하고 SMARCA4를 표적화하기 위해 CIRTS-7을 사용하는 것을 목표로 하였다(**도 5a**). 본 발명자들은 대조군, 비표적화 gRNA 서열, 또는 각각의 내인성 유전자를 표적화하는 gRNA 서열을 표시하는 각 CIRTS를 위한 gRNA에 대한 발현 벡터와 함께 두 CIRTS 단백질에 대한 발현 벡터로 세포를 동시-형질주입시켰다. 이어서, 본 발명자들은 RT-qPCR에 의해 2개의 표적 유전자의 상대적인 RNA 수준의 변화를 평가하였다. 실제로, 각각의 CIRTS 시스템은 두 시스템 사이에 어떠한 혼선도 없이 표적 전사물을 표적내 gRNA 의존적 방식으로 분해하였다(**도 5b**). 결정적으로, 본 발명자들은 두 CIRTS 단백질의 존재 하에 하나의 전사물 또는 다른 전사물을 선택적으로 분해할 수 있었고 또한 동시에 두 전사물을 분해할 수 있었다.

[0250] CIRTS-6의 RNA 분해는 CIRTS-6만이 세포에 전달될 때(**도 11a**)와 유사한 수준을 나타냈지만, CIRTS-7을 사용한 표적화는 개별 형질주입과 비교하여 RNA 분해를 약간 감소시켰다(**도 11b**). 다중 표적화 실험을 위해, 본 발명자들은 이틀 연속으로 매일 CIRTS 및 gRNA 벡터로 세포를 순차적으로 형질주입시켰다. 본 발명자들은 이러한 감소된 활성이 CIRTS-7로 형질주입된 지 2일째에 형질주입 효율이 감소한 결과로 본다. 또한 TBP6.7 기반 CIRTS에는 내인성 포유동물 세포 기구에 직교하는 RNA 헤어핀이 있지만, SLBP 기반 작제물은 이론적으로 여전히 세포 히스톤 mRNA에 결합할 수 있으며, 이는 전체 활성 감소로 이어질 수 있다. 흥미롭게도, 두 CIRTS 모두 표적내 gRNA와 함께 전달되었을 때 이 분석에서 최상의 성능을 보였다. 이 시점에서, 본 발명자들은 2개의 직교 CIRTS 시스템이 각각 서로 직교할 수 있고 각각 gRNA 의존적 방식으로 내인성 전사물을 동시에 표적화할 수 있다고 결론내린다. 본 발명자들은 현재, 더 이상 내인성 결합 파트너가 없고 잠재적으로 세포 이벤트 교란 없이 작용할 수 있는 조작 직교 CIRTS에 대해 연구하고 있다.

[0251] **C.논의**

[0252] 요약하면, 여기에서 본 발명자들은 프로그램 가능한 RNA 효과기 단백질을 조작하기 위한 새로운 전략인 CIRTS를 제시했다. CIRTS는 작고, 완전히 인간화될 수 있으며, 살아있는 세포의 내인성 RNA를 표적화 할 수 있으며, 다차원 제어를 위해 직교적 및 상승적으로 함께 작동할 수 있다. 연구 도구로서 CIRTS는 크기가 더 작기 때문에 이전 방법에 이점을 제공해야 한다. 예를 들어, CIRTS-2 및 CIRTS-3은 각각 65 및 36 kDa인 반면, 비교 가능한 Cas13b 기반 프로그램 가능한 YTHDF1 및 YTHDF2 시스템은 각각 155 및 126 kDa이다(**도 6**). CIRTS-1은 현재까지 발견된 가장 작은 DNA 표적화 Cas 단백질인 Cas14a보다 훨씬 작다(Harrington et al., 2018). 전달 시스템의 더 작은 크기는 살아있는 세포에서 RNA 조절 단백질의 역할을 밝혀낼 기회를 제시하는 연구 중인 내인성 전사물에 덜 교란적이어야 한다.

[0253] 번역의 관점에서, CIRTS는 몇 가지 주요 이점과 기회를 제공해야 한다. CIRTS의 인간화된 특성은 면역 반응을

회피하는 경로를 제공하여 지속-전달 치료법의 가능성을 열어줄 것이다. 인간β-디펜신 3과 같은 부속 단백질을 사용하는 것과 관련하여, 이 단백질은 광범위하게 연구되었으며 그 기능에 대한 구조적 기반이 해명되었다(Dhople et al., 2006; Kluver et al., 2005). 이러한 지식을 통해 당업자는 하전된 성질을 유지하지만 내인성 기능을 제거하는 β-디펜신 3 펩타이드를 조작할 수 있을 것이다. 현재, CIRTS의 작은 크기는 예를 들어 분해를 위해 하나의 전사물을 그리고 번역 활성화를 위해 다른 전사물을 표적화하기 위해 바이러스 전달 시스템에서 여러 조절 단백질을 동시에 전달이 가능하게 할 것이다. CIRTS의 다중화 능력은 전달될 수 있는 효과기 단백질의 작은 크기 및 다양성과 결합하여 다중 차원에서 한 번에 다중 유전자를 표적화함으로써 세포 재프로그램의 가능성을 열어준다.

[0254] 더 넓은 관점에서, CIRTS 플랫폼은 인간 단백질 툴박스에 포함된 부분을 결합하여 새로운 특성을 가진 단백질을 조작할 수 있는 가능성을 보여준다. CIRTS 시스템은 RNA 조절을 연구하고 활용하기 위한 새로운 접근 방식을 제공하며 질병 치료를 위한 세포 조절에 개입할 수 있는 많은 미래의 기회를 열어줄 것이다.

[0255] **실시예 2 - 인간 부분으로부터 구축된 프로그램 가능한 RNA-가이드 RNA 효과기 단백질**

[0256] **A.결과**

[0257] **1.CIRTS의 설계**

[0258] DNA 표적화 Cas9 기반 시스템은 DNA를 풀고 표적 서열에 어닐링하기(anneal) 위해 복잡한 생물물리학적 메커니즘을 사용하지만(Rutkauskas et al., 2015; Sternberg et al., 2014; Szczelkun et al., 2014), Cas13의 기계론적 연구는 RNA 표적화는 gRNA의 중심 시드(seed) 영역에 의해 시작됨을 보여주었다(Liu et al., 2017). 또한, Cas13 시스템은 개별 전사물에 대한 서열 맥락 표적화 가능성에서 상당한 가변성을 나타낸다(Abudayyeh et al., 2017; Cox et al., 2017; Konermann et al., 2018). 동시에 이러한 발견은 gRNA와 표적화된 전사물 사이의 서열 상보성과, 주어진 부위의 접근성이 RNA 표적화의 핵심 요건임을 시사한다. 본 발명자들은 과발현된 리포터 작제물을 사용한 이전의 RNA 테더링 분석과 유사하게, 표적 RNA에 단백질 카고를 전달하기 위해 gRNA 서열을 표시하는 정의된 단백질-RNA 상호작용을 사용하는 Cas13-영감 시스템을 조작하려고 했다(Coller and Wickens, 2007). 실제로, 헤어핀 결합 단백질과 공유 RNA 융합이 RNA 편집 기구를 전사물에 전달하는 데 사용되었다(Montiel-Gonzalez et al., 2016; Sinnamon et al., 2017; Vogel et al., 2018).

[0259] 현재 Cas13의 특성화(Abudayyeh et al., 2017; Cox et al., 2017; Gootenberg et al., 2018; Konermann et al., 2018; Liu et al., 2017)를 기반으로, 본 발명자들은 최소한의 프로그램 가능한 RNA 표적화 시스템은 다음의 4가지 성분: (1) 시스템의 코어 역할을 하고 조작된 gRNA에 표시된 특이적 RNA 구조에 대한 선택적 고친화성 결합제인 RNA 헤어핀-결합 단백질; (2) 조작된 헤어핀-결합 단백질과 상호작용하는 구조 및 관심 표적 RNA에 상보성을 갖는 서열 모두를 특징으로 하는 gRNA; (3) 표적 결합 전에 가이드 RNA를 안정화 및 보호하기 위해, 표시된 gRNA 서열에 비특이적으로 결합할 수 있는 하전된 단백질; 및 (4) 근접 의존 방식으로 표적화된 RNA에 작용하는 리보뉴클레아제 또는 후성전사체 조절인자와 같은 효과기 단백질(도 12a)을 가지는 것으로 추론하였다. Cas13은 단일 단백질 도메인에 이러한 모든 기능적 성분을 담고 있지만(Liu et al., 2017; Tambe et al., 2018), 본 발명자들은 CRISPR/Cas-영감 RNA 표적화 시스템(CIRTS)이라고 명명한, 각각이 이들 기능 중 하나를 수행하는 다중 단백질 도메인을 결합하는 시스템을 조작하는 것을 구상했다. CIRTS는 모듈 조성이 다양하며 도 12a 및 도 21에 나열된 바와 같이 고유하게 번호가 지정된다.

[0260] **2.CIRTS-1의 개발 및 시험관 내 검증**

[0261] 1세대 시스템인 CIRTS-1의 경우, 본 발명자들은 HIV 전이-활성화 반응(TAR) 헤어핀에 결합하도록 이전에 조작되었고 내인성 인간 RNA 헤어핀 표적을 갖지 않는, 인간 헤어핀-결합 단백질 U1A 단백질(TBP6.7)의 진화된 버전을 사용했다(Blakeley and McNaughton, 2014; Crawford et al., 2016)(도 12b). 본 발명자들은 TAR 헤어핀, 뉴클레오타이드 링커 서열(L) 및 가이드 서열을 포함하는 gRNA를 설계했다(도 12c). 시스템을 개발하고 검증하기 위해, 본 발명자들은 먼저 TBP6.7을 이전에 비특이적 근접 의존적 RNA 엔도뉴클레아제로 사용된, 인간 넌센스-매개 mRNA 붕괴 인자 SMG6의 Pin 뉴클레아제 도메인에 융합하여 프로그램 가능한 리보뉴클레아제를 조작했다(Batra et al., 2017; Choudhury et al., 2012). 이 가장 단순한 설계가 이미 세포 기반 루시퍼라아제 분석에서 유망한 gRNA 매개 전사물 분해를 디스플레이했지만(도 22a, 왼쪽), 성능은 상당히 안 좋았고, 본 발명자들은 이를 디스플레이된 가이드 서열의 분해의 결과로 보았다. Cas13 시스템의 단백질 표면과 헤어핀 채널은 높게 하전되는 경향이 있어, 가이드 RNA 서열에 비특이적으로 결합하고 이를 안정화할 가능성이 있다(Liu et al., 2017). 이 RNA 보호 기능을 상기 시스템에 조작하기 위해, 본 발명자들은 비특이적인 낮은 친화성 단일 가닥

RNA 결합 단백질(ssRNA 결합 단백질)을 추가하는 것을 목표로 하였다. 그러나, 인간 단백질체는 본 발명자들이 알고 있는 최선에서 주해가 달린 작은 비특이적 단일 가닥 RNA 결합 단백질을 쉽게 포함하지 않았다. 따라서, 본 발명자들은 작은 바이러스 ssRNA 결합 단백질인 ORF5를 사용하여 CIRTS-1을 개발했다(Zhou et al., 2006). 전체적으로, CIRTS-1은 상응하는 gRNA와 함께 ORF5-TBP6.7-Pin 뉴클레아제 도메인의 단백질 융합 복합체로 구성된다.

[0262] 본 발명자들은 먼저 모델 RNA 표적 기질에 대한 시험관 내 CIRTS-1의 프로그램 가능한 RNA 결합 및 RNA 뉴클레아제 활성을 특성화하였다. Pin 뉴클레아제 도메인은 이전에 Mn2+의 존재 하에서 활성인 것으로 나타났으며 활성은 EDTA의 첨가에 의해 소멸될 수 있었다(Choudhury et al., 2012). CIRTS-1을 직접 과발현하면 세포 펠릿에서 불용성 단백질이 생성되며, 이는 본 발명자들이 N-말단 MBP 태그를 CIRTS-1에 융합(MBP-CIRTS-1)하여 해결했다. 필터 결합 분석에서 정제된 MBP-CIRTS-1 단백질 및 gRNA를 사용하여, 본 발명자들은 MBP-CIRTS-1이 22nM의 겉보기 결합 해리 상수(K_D)로 gRNA-의존적 방식으로 표적 RNA에 결합함을 발견하였다(도 13a). 결정적으로, 본 발명자들이 시스템에 비-표적화 gRNA를 제공한다면, 본 발명자들은 약 50배 더 약한 결합을 본다. 더욱이, 절단 분석에서, 본 발명자들은 MBP-CIRTS-1이 gRNA- 및 Mn^{2+}-의존적 방식으로 RNA 기질을 분해한다는 것을 발견하였으며, 이는 융합 맥락에서 Pin 리보뉴클레아제 도메인의 활성을 확인해준다(도 13b). 종합적으로, 이러한 시험관 내 결과는 CIRTS-1 이면의 설계 원리를 검증하고 발명자들이 살아있는 세포에서 사용하기 위해 시스템을 최적화하도록 동기를 부여했다.

[0263] **3.살아있는 세포에서 CIRTS-1의 최적화**

[0264] 살아있는 포유동물 세포에서 CIRTS-1의 표적 뉴클레아제 활성을 시험하기 위해, 본 발명자들은 표적 반딧불이 루시퍼라아제(Fluc) RNA 상의 gRNA-의존성 전사 변화에 대해 보고하는 이중 루시퍼라아제 리포터 분석을 확립했다(도 14a). 이 시스템을 사용하여, 본 발명자들은 상이한 단백질 링커 유형(도 23a), gRNA 구조 및 길이(도 23c), 및 CIRTS 뉴클레아제 세포 국소화(도 22b)를 분석함으로써 Pin 뉴클레아제 CIRTS의 배치를 최적화하였다. 본 발명자들은 헤어핀-결합 단백질과 40개 뉴클레오타이드 길이의 gRNA 사이의 긴 가요성 링커가 최상의 녹다운 효율을 초래한다는 것을 발견하였다. 최적화 후, 본 발명자들은 표적 리포터 RNA를 분해하는 최적화된 CIRTS-1(Pin 뉴클레아제) 시스템의 능력을 Cas13b 시스템과 비교하였다(Cox et al., 2017). 본 발명자들은 CIRTS-1 및 Cas13b 모두에 대한 이중 루시퍼라아제 리포터 분석에서 반딧불이 루시퍼라아제 mRNA를 표적으로 하는 gRNA는 물론, 각각의 프로그램 가능한 뉴클레아제에 대한 대조군, 비표적화 gRNA(람다 파지 서열을 표적화)를 설계했다. 전사물에 결합하면 단백질 수준이 변경되는지 여부를 시험하기 위해, 본 발명자들은 CIRTS-1의 Pin 뉴클레아제 도메인에 이전에 보고된 '죽은(dead)' 돌연변이를 함유하는(Eberle et al., 2009), CIRTS-0을 음성 대조군 역할을 하도록 조작했다. 본 발명자들은 CIRTS-0이 표적 전사물의 발현 수준에 유의한 영향을 미치지 않는다는 것을 발견하였고(도 14b 및 도 22c), 이는 표적 RNA에 대한 CIRTS 리보핵산단백질의 결합이 표적화된 전사물에 최소한으로 교란적이라는 것을 나타낸다. 다음으로, 본 발명자들은 활성 뉴클레아제를 함유하는 CIRTS-1이 표적의 분해를 매개할 수 있는지 여부를 시험하였다. 실제로, 본 발명자들은 루시퍼라아제 활성에 의해 모니터링되는 단백질 수준(도 14c) 및 RT-qPCR에 의해 모니터링되는 mRNA 수준(도 22d) 모두에서 측정된 표적 Fluc mRNA의 gRNA-의존성 분해를 발견하였다. 두 결과 모두 CIRTS 전략이 살아있는 세포에서 실행 가능한 방법임을 시사한다. CIRTS-1은 Cas13b에 비해 리포터 유전자를 표적화하는 데 덜 효율적이지만, 특히 Cas13b 시스템이 이 녹다운 기능을 수행하도록 진화했다는 점을 고려할 때 성능은 크게 다르지 않았다. CIRTS-1(Pin 뉴클레아제)의 성능에 고무된 본 발명자들은 다음으로 gRNA, 헤어핀 결합 도메인, 비특이적 RNA 결합 도메인 및 효과기 단백질을 포함하여 시스템의 각 성분이 다양한 기능을 갖는 CIRTS를 달성하기 위해 다른 부분으로 교체될 수 있는지 여부를 시험하여 설계의 다목적성을 평가하려고 했다.

[0265] **4.리포터 전사물에 대한 CIRTS의 모듈성**

[0266] CIRTS 설계의 다목적성을 탐구하기 위해, 본 발명자들은 다음으로 CIRT 단백질 전달 시스템의 각 성분에 대해 상이한 단백질 도메인을 시험했다. 먼저, 본 발명자들은 CIRTS가 발명자들이 이전에 dCas13b 시스템을 사용하여 전달했던 RNA 후성전사체 조절 "판독자(reader)" 단백질을 전달할 수 있는지 여부를 분석했다(Rauch et al., 2018). 연구를 위해 본 발명자들은 가장 널리 퍼진 mRNA 변형인 N6-메틸아데노신의 조절 단백질에 집중하기로 했다. 평균적으로 각 전사물은 3' UTR에서 검출된 m6A 존재비가 높은 3개의 변형 부위를 함유하며, m6A는 스플라이싱(Xiao et al., 2016), 번역(Meyer et al., 2015; Wang et al., 2015) 및 안정성(Du et al., 2016; Wang et al., 2013)에서 조절 역할을 하는 것으로 나타났다. 본 발명자들은 CIRTS-1의 Pin 뉴클레아제 효과기 단백질을 번역 기구(Wang et al., 2015)를 동원하는 세포질 m6A 판독자 단백질인 YT521-B 상동성 도메인 패밀리 단백

질 1(YTHDF1)의 N-말단 도메인으로 교체하여, CIRTS-2를 생성했다. 본 발명자들의 현재 설계는 m6A 독립적인 방식으로 판독자 단백질의 연구를 단순화하고 CIRTS를 검증할 때 세포 mRNA의 다양한 m6A 수준으로 인한 혼란을 피하기 위해 m6A를 인식하는 YTHDF1의 C-말단 YTH 도메인을 포함하지 않는다는 것에 유의한다. CIRTS-2(YTHDF1)가 CIRTS-1 실험과 동일한 표적 서열에 전달될 때, RNA 수준은 상대적으로 변하지 않지만(도 22e), YTHDF1의 알려진 번역-활성화 역할에 기초하여 예상되는 바와 같이(Wang et al., 2015), RNA로부터 단백질 수준의 상당한 증가가 일어난다(도 14d). 그런 다음, 본 발명자들은 YTHDF1 단편을 RNA 탈아데닐화 기구를 동원하고 RNA 분해를 유도하는(Du et al., 2016; Wang et al., 2013) 상이한 m6A 판독자 단백질인 YTHDF2의 단편으로 교체하여, CIRTS-3을 생성하였다. 리포터 mRNA로 CIRTS-3(YTHDF2)을 전달하면 RNA(도 22f) 및 단백질 수준(도 14e) 모두에 의해 측정된 바와 같이 표적 전사물의 분해를 유도한다. CIRTS-1 내지 -3은 다양한 효과기 단백질 카고(effector protein cargo)를 살아있는 세포의 표적 RNA에 전달하는 설계 전략의 다목적성을 보여준다.

[0267] 효과기 도메인의 모듈성을 입증한 후, 본 발명자들은 다른 인간 부분이 RNA 헤어핀 결합 도메인 및 비특이적 ssRNA 결합 단백질에도 사용될 수 있는지 평가에 착수했다. 본 발명자들은 CIRTS-3(YTHDF2)의 TBP6.7을 인간 히스톤 줄기 루프 결합 단백질(SLBP)의 RNA 헤어핀 결합 도메인으로 대체하여 CIRTS-4(YTHDF2)를 생성하였다. 동시에, 본 발명자들은 히스톤 mRNA 줄기 루프 구조를 기반으로 gRNA를 설계하였다(도 12d). 리포터 분석(도 14f)에서 그리고 RNA 수준을 평가하기 위한 RT-qPCR(도 22g)에 의한 CIRTS-4(YTHDF2) 분석은 CIRTS-3과 유사한 성능을 나타냈고, 이는 다른 헤어핀 결합 도메인이 CIRTS의 코어로 사용될 수 있음을 확인해준다.

[0268] 다음으로, 본 발명자들은 CIRTS 시스템의 완전히 인간화된 버전을 조작하려고 하였다. 앞서 언급한 바와 같이, 본 발명자들은 바이러스 비특이적 단일 가닥 RNA 결합 단백질 ORF5를 기반으로 하는 첫 개념 증명 시스템(proof-of-concept system)을 설계했다. 주해가 달린 인간 단일 가닥, 비특이적 RNA 결합 단백질은 없지만, 본 발명자들은 높게 하전된 양이온성 인간 단백질이 CIRTS 시스템에서 ORF5의 역할을 수행할 수 있다고 추론했다(Cronican et al., 2011). 따라서, 본 발명자들은 CIRTS-3에서 ORF5 대신에 2개의 양이온성 인간 단백질인 HBEGF 및 β-디펜신 3을 조작하여 CIRTS-5 및 CIRTS-6을 각각 생성하였다. 다시, 루시퍼라아제 리포터 분석에 이들 프로그램 가능한 효과기들을 배치하면 YTHDF2 후성전사체 조절에 의해 매개된 표적 유전자의 gRNA-의존성 분해(도 14g 및 S2H)가 나타났다.

[0269] 마지막으로, 본 발명자들은 추가 효과기 단백질과 함께 기능 범위에서 CIRTS의 다목적성을 확인하기 위해 CIRTS를 사용하여 인간 ADAR2(hADAR2)의 촉매 도메인을 RNA 전사물에 전달하였다. 본 발명자들은 반딧불이 루시퍼라아제 코딩 영역에 G-to-A 돌연변이를 포함하는 이중 루시퍼라아제 리포터를 설계하여 번역이 조기에 중단되고 측정 가능한 반딧불이 루시퍼라아제 활성이 나타나지 않았다(도 15a 및 22j). 그런 다음 본 발명자들은 CIRTS를 배치하여 hADAR2의 알려진 과활성 돌연변이체(Kuttan and Bass, 2012)인 wt ADAR2(CIRTS-7) 또는 hADAR 2 E488Q(CIRTS-8)를 돌연변이된 위치에 전달했고, 그 결과 두 경우 모두 루시퍼라아제 활성의 gRNA-의존성 구조(rescue)가 나타났다(도 15b). 과활성 hADAR2 돌연변이체는 루시퍼라아제 분석을 기반으로 표적내 gRNA가 없을 때 더 높은 편집 효율과 더 높은 백그라운드를 보였다. 그러나 과잉 활성 돌연변이체를 사용하면 완화된 서열 제약을 갖기 때문에 더 넓은 기질 범위를 표적화하는 데 이롭다(Kuttan and Bass, 2012). 종합하면, 리포터 분석에서 이러한 다양한 CIRTS의 성능은 헤어핀 결합 도메인과 해당 gRNA, 단일 가닥 RNA 결합 단백질 및 효과기 단백질을 포함한 CIRTS 설계의 모듈성을 보여준다.

[0270] **5.CIRTS로 내인성 mRNA 표적화**

[0271] 다음으로 본 발명자들은 CIRTS-1 프로그램 가능한 뉴클레아제 및 CIRTS-3 프로그램 가능한 YTHDF2-매개 붕괴 시스템을 예시로 사용하여 CIRTS 시스템이 표적 내인성 전사물에 효과기 단백질을 전달할 수 있는지 여부를 평가하고자 했다. 본 발명자들은 이전에 Cas13 표적으로 검증된 5개의 RNA 전사물을 선택하고, 이것들이 프로그램 가능한 RNA-결합 시스템에 의한 RNA 표적화에 접근 가능하다고 추론했다. 그런 다음 본 발명자들은 이전에 Cas13 실험에 사용된 표적상의 동일한 결합 부위를 사용하여 각 표적에 대한 gRNA를 설계했다(Abudayyeh et al., 2017; Konermann et al., 2018). 본 발명자들은 RT-qPCR에 의해 각 표적의 RNA 수준에 대한 CIRT 시스템의 효과를 분석하였다. 세포가 특정 gRNA 발현 벡터와 함께 CIRTS-1 또는 CIRTS-3으로 형질주입되었을 때, 본 발명자들은 5가지 내인성 mRNA 전사물: PPIB, NFKB1, NRAS, B4GALNT1, 및 SMARCA4의 각각에 대해 RT-qPCR에 의한 RNA 수준의 상당한 감소를 관찰했다(도 16a 및 16b). mRNA를 표적화하는 것 외에도, 본 발명자들은 또한 본 발명자들이 CIRTS-1(Pin 뉴클레아제)을 MALAT1에 표적화함으로써 lncRNA와 같은 다른 RNA 종을 표적화할 수 있음을 확인하였다(도 24a). 상대적 녹다운 효율은 각 유전자에 따라 달랐으며, 이는 다른 RNA 표적화 시스템에서도 관찰되며 각 유전자에 특이적인 접근성 또는 기타 조절 경로에 의해 잠재적으로 매개된다. 그럼에도 불구하고, 이러한 결과는 CIRTS가 표적에 대한 활성 뉴클레아제 활성을 통해 또는 내인성 후성전사체 조절 경로를 촉

발함으로써 내인성 전사물을 표적화하고 붕괴를 매개할 수 있음을 확인해준다.

[0272] 다음으로, 본 발명자들은 CIRTS-2가 YTHDF1-매개된 후성전사체 경로를 통해 내인성 전사물의 단백질 생산을 촉발할 수 있는지 여부의 평가에 착수했다. 본 발명자들은 웨스턴 블로팅에 의한 CypB(PPIB의 단백질 생성물) 단백질 생산의 분석을 위해 보고되고 신뢰할 수 있는 항체가 있는 풍부한 전사물 PPIB를 선택했다. 실제로, CIRTS-2로 형질주입된 세포 및 표적내 gRNA는 RNA 수준(도 24b)의 변화 없이 단백질 수준(도 16c 및 도 24c)의 증가를 나타내었고, 이는 전사물에 대한 이전에 보고된 YTHDF1 효과와 일치한다(Wang et al., 2015). 대조군으로서, YTHDF2를 전달하는 CIRTS-3으로 수행된 동일한 실험은 약간의 단백질 수준 감소로 나타났고, 이는 mRNA 수준의 감소와 상관관계가 있다(도 16b 및 24c).

[0273] 마지막으로, 전사물 위치 특이적 효과의 첫 번째 시험으로서, 본 발명자들은 SMARCA4 mRNA를 따라 gRNA를 타일링하고 CIRTS-3에 의한 YTHDF2 매개 붕괴를 시험했다. 본 발명자들은 gRNA가 표적화된 mRNA에 도달하는 위치에 따라 시스템의 극적으로 다른 성능을 발견했으며(도 17), 이는 CIRTS 결합 접근성 및 조절 단백질 서열 요건 모두의 결과일 가능성이 높다. 종합하면, 이러한 실험은 CIRTS 플랫폼이 gRNA 의존적 방식으로 내인성 전사물에서 기능적이며, 표적 전사물을 능동적으로 분해하거나, 분해 기구가 표적 전사물에 작용하도록 촉발하거나 또는 표적 전사물로부터 번역을 활성화하고 단백질 생산을 증가시킬 수 있음을 보여준다.

[0274] **6.CIRTS의 표적화 특이성**

[0275] 특정 CIRTS가 RNA 기질을 표적화하는 방법에 대한 통찰력을 얻기 위해, 본 발명자들은 미스매치, 전사체 전반의 표적외 및 내인성 기질 표적화에 대한 gRNA의 민감도를 다루는 일련의 실험을 설계했다. 미스매치 내성(mismatch tolerance)을 평가하기 위해, 본 발명자들은 gRNA와 표적 RNA 사이에 형성된 이중체(duplex)에 1, 2 또는 3개의 미스매치를 도입할 때 본 발명자들이 표적화 효과를 분석할 수 있도록 하는 루시퍼라아제 기반 미스매치 실험을 설계했다. 본 발명자들은 질병 관련 KRAS4b 전사물을 루시퍼라아제 리포터에 융합하기로 결정하고 조작된 시스템이 암 관련 G12D(표적 1), 야생형(표적 2), G12C(표적 3) 및 G12W(표적 4) KRAS4b 변이체들을 구별할 수 있는지 질문했다(도 25a). 본 발명자들은 CIRTS가 G12D 및 야생형 변이체의 비슷한 녹다운을 산출한다는 것을 발견했는데, 이는 하나의 미스매치가 표적화 특이성에 큰 변화를 일으키지 않는다는 것을 나타낸다(도 25b). 그러나, 본 발명자들이 각각 2개 및 3개의 미스매치를 포함하는 G12C 및 G12W 리포터에 시스템을 표적화했을 때 CIRTS 녹다운 효율이 감소했다.

[0276] 다음으로, 본 발명자들은 분석에서 녹다운 효율에 가장 큰 영향을 보인 gRNA와 표적 RNA 사이에 형성된 이중체의 중앙 영역에 있는 미스매치들에 초점을 맞추어, gRNA 길이의 증가가 CIRTS의 미스매치 내성에 영향을 미칠 수 있는지 여부를 평가하였다. 더 짧은 20nt gRNA에서 관찰된 바와 같이, 본 발명자들은 본 발명자들이 미스매치가 없거나 하나인 리포터를 표적화할 때 녹다운 효율에 차이가 없음을 확인했다. 그러나, 더 긴 40nt gRNA는 2개-미스매치 변이체가 표적화되었을 때 일부 효과를 구조할 수 있으며, 이는 gRNA 길이가 시스템의 특이성에 기여함을 나타낸다(도 25d). 기존 기술과의 비교로서, 본 발명자들은 Cas13b에 동일한 미스매치 분석을 실시했으며, 이는 Cas13b가 일반적으로 불일치에 덜 민감함을 나타냈다. 1개 및 2개의 미스매치가 있는 리포터에 Cas13b를 표적화하면 녹다운 효율의 변화를 거의 나타내지 않은 반면, 3개의 미스매치는 녹다운 효율의 상당한 감소를 초래하였다(도 25c). Cas13b와 CIRTS는 모두 gRNA와 표적 RNA 사이에 형성된 이중체의 중심에서 미스매치된 염기쌍에 가장 민감하며, 이는 Cas13b에 대한 이전 연구와 잘 일치하는 연구결과이다(Abudayyeh et al., 2017).

[0277] 전사체 전반의 표적외를 평가하기 위해, 본 발명자들은 시스템을 RNA 서열분석에 적용하였다. 본 발명자들은 내인성 전사물 SMARCA4를 표적화하는 CIRTS Pin 뉴클레아제 및 CIRTS YTHDF2의 효과를 분석하였다(도 25e 및 25f). 두 경우 모두, 본 발명자들은 통계적으로 유의미한 표적외를 찾지 못했다. 그러나, 본 발명자들은 표적 전사물만을 볼 때((pval <0.1) 표적화된 전사물의 녹다운 및 CIRTS-3(YTHDF2)에 의한 심지어 통계적으로 유의한 녹다운을 보지만, 전사체 전반의 방식으로 평가했을 때 녹다운 수준은 통계적으로 유의한 영역으로 떨어지지 않는다(qval < 0.1). 이 시점에서 본 발명자들은, 시스템의 과발현이 전사체에 큰 교란을 유발하지 않는다는 결론을 내리며, 본 발명자들의 미스매치 연구와 함께 취한 결과는 표적화 특이성 및 그에 따른 녹다운 효율이 향후 연구에서 추가 최적화에 의해 추가로 최적화될 수 있음을 나타낸다.

[0278] CIRTS가 관심 전사물에 결합하는지 확인하기 위해, 본 발명자들은 또한 RNA 면역침강 후 RT-qPCR을 수행하였다. 본 발명자들은 이전에 Cas13 시스템, PPIB 및 B4GALNT1에 의해 표적화된 두 개의 내인성 전사물을 표적화하는 gRNA들을 설계했다(Cox et al., 2017; Konermann et al., 2018). 본 발명자들은 3x FLAG-태그에 융합된 CIRTS-0(죽은 뉴클레아제 CIRTS)과 함께 각 gRNA를 개별적으로 전달하였다. 이어서, 본 발명자들은 용해물에 항-Flag

항체로 면역침강을 수행하고, 단백질에 결합된 각각의 표적 RNA의 상대적인 양을 정량화하였다. 실제로, 내인성 전사물 둘 모두는 gRNA-의존적 방식으로 2.5배 내지 5배 농축되었고(도 25h), 이는 CIRTS가 내인성 전사물에서 프로그램 가능한 RNA-가이드 RNA 결합 단백질로서 기능함을 확인해준다.

[0279] **7.CIRTS를 사용한 다중 내인성 RNA의 다중 표적화**

[0280] CIRTS의 표적화 특이성과 모듈성은 함께 발명자들이 다중 표적화 방식으로 CIRTS의 적용을 확장하도록 영감을 주었다. 한 번에 단일 효과기 단백질을 전달하고 단일 전사물을 표적화하는 대신, 본 발명자들은 CIRTS가 하나 초과의 전사물을 표적화할 수 있는지 또는 동일한 샘플에서 하나 초과의 효과기 단백질을 전달할 수 있는지 시험에 착수했다. 원칙적으로, 각각 별도로 조작된 gRNA(도 12c 및 12d)를 사용하는, TBP 헤어핀 결합 도메인으로부터 구축된 CIRTS와 SLBP 헤어핀 결합 도메인으로부터 구축된 CIRTS는 서로 직교해야 하며, 이는 동일한 또는 심지어 상이한 CIRTS로 다수 전사물의 선택적 표적화를 가능하게 한다.

[0281] 먼저, 본 발명자들은 단일 CIRTS를 사용하여 다수 전사물을 동시에 표적화할 수 있는지 여부를 시험했다. 본 발명자들은 PPIB, SMARCA4 및 NRAS를 표적화하는 3개의 gRNA와 함께 CIRTS-6으로 세포를 동시-형질주입시키고 RT-qPCR에 의해 RNA 수준의 변화를 평가하였다. 예상대로, 본 발명자들은 3개의 표적화된 전사물 모두에 대한 RNA 수준의 감소를 관찰하였다(도 18a 및 18b). 그러나, 본 발명자들은 동일한 세포에 여러 gRNA 또는 CIRTS를 배치할 때 효율이 약간 감소하는 것을 관찰하였다. 본 발명자들은 이러한 감소를 4개의 플라스미드를 이용한 세포의 동시 형질주입의 결과로 본다. 단일 CIRTS/gRNA 실험의 성능을 회복하려면 단백질 및 gRNA 수준 및 벡터의 추가 최적화가 필요할 수 있다.

[0282] 다음으로, 두 가지 상이한 유형의 효과기가 동시에 사용될 수 있는지 여부를 시험하기 위해, 본 발명자들은 반딧불이 루시퍼라아제를 표적화하기 위해 완전히 인간화된 버전의 YTHDF1 작제물인 CIRTS-9를 배치하고 SMARCA4를 표적화하기 위해 CIRTS-10(YTHDF2)을 배치했다(도 18c 및 18d). 본 발명자들은 두 단백질 모두 활성이고 루시퍼라아제 단백질의 예상된 증가 및 RNA 수준의 감소를 각각 유도한다는 것을 발견하였다. 더욱이, 다중-표적화 CIRTS의 직교성을 추가로 확증하기 위해, 본 발명자들은 2개의 CIRTS, 즉 CIRTS-6(TBP6.7) 및 CIRTS-10(SLBP)(도 26a 및 26b)을 배치했지만, YTHDF2를 2개의 상이한 내인성 표적 mRNA에 전달하기 위해 상이한 헤어핀-결합 모듈을 사용하였다(도 27a). 각각의 CIRT 시스템은 두 시스템 사이에 최소한의 혼선으로 표적 전사물을 표적내 gRNA 의존적 방식으로 분해하였다(도 27b).

[0283] 이 시점에서, 본 발명자들은 TBP6.7 및 SLBP-기반 CIRTS가 각각 gRNA-의존적 방식으로 내인성 전사물을 동시에 표적화할 수 있다고 결론내렸다. 본 발명자들은 현재 TBP6.7에서 수행된 것처럼 인간 단백질을 새로운 특이성을 향해 진화시킴으로써 내인성 결합 파트너가 없는 직교 CIRTS를 조작하는 연구를 하고 있다. 인간 유래는 아니지만, 본 발명자들은 PP7과 같은 다른 헤어핀-결합 시스템이 CIRTS를 생성하는 데에도 사용될 수 있다는 것을 발견했으며, 이는 다양한 선택적 및 직교 시스템을 생성하는 것이 가능함을 시사한다(도 22i). CIRTS는 여러 조절 단백질이 동시에 전달될 수 있도록 하여, 예를 들어 분해를 위한 하나의 전사물과 번역 활성화를 위한 다른 전사물을 표적화하며, 이는 여러 차원에서 한 번에 여러 유전자를 표적화함으로써 세포 재프로그램의 가능성을 열어준다(Bao et al., 2017; Gao et al., 2016).

[0284] **8.AAV에 의한 CIRTS의 바이러스 전달**

[0285] CIRTS의 인간 유래 특성을 제외하고 또 다른 핵심 이점은 CIRTS의 크기가 작아 보다 효율적인 바이러스 패키징 및 전달이 가능하다는 것이다. 아데노바이러스 연관 바이러스(AAV)는 광범위한 이용 가능 혈청형(Gao et al., 2005), 낮은 면역 반응 자극(Vasileva and Jessberger, 2005), 및 낮은 유전체 삽입 위험(Gao et al., 2005; Naso et al., 2017)으로 인해 상이한 세포 유형에 전이 유전자 및 유전자 요법을 전달하는 다목적 전달 매개체이다. 그러나 약 4.7kb의 제한된 패키징 용량으로 인해 많은 Cas13 단백질을 패키징하고 전달하는 것이 어려웠었다(Wu et al., 2010). AAV에 의해 CIRTS가 전달될 가능성을 보여주기 위해, 본 발명자들은 이중 CIRTS-6/gRNA 전달 플라스미드를 설계하고 그것을 AAV 전달 매개체에 패키징했다. CIRTS 단백질 및 gRNA를 포함하는 전체 삽입물은 단지 2.7kb이다(도 19a). 본 발명자들은 생성된 바이러스를 사용한 HEK293T 세포의 형질도입이 루시퍼라아제 리포터 및 내인성 표적 둘 모두에서 CIRTS-6의 녹다운 효율을 반복한다는 것을 발견하였고(도 19b 및 19c), 이는 바이러스-전달된 CIRTS가 여전히 기능함을 확인해주고 임상 배치를 향한 경로를 제공한다. 미래의 응용 분야에서는 이전에 일시적 형질주입에 의해 나타난 바와 같이, 상향 조절을 위한 하나의 전사물과 분해를 위한 하나의 전사물을 동시에 표적화하기 위해 하나 초과의 CIRT 시스템을 AAV 전달 매개체에 패킹하는 것을 생각할 수 있다.

[0286] B.논의

[0287] 요약하면, 여기에서 본 발명자들은 프로그램 가능한 RNA 효과기 단백질을 조작하기 위한 다목적 전략인 CIRTS를 제시했다. CIRTS는 작고, 완전히 인간화될 수 있으며, 살아있는 세포의 내인성 RNA를 표적화 할 수 있으며, 다차원 전사체 제어에 작동할 수 있다. 연구 도구로서 CIRTS는 크기가 더 작기 때문에 이전 방법에 이점을 제공해야 한다. 예를 들어, CIRTS-2 및 CIRTS-3은 각각 65 및 36 kDa인 반면, 비교 가능한 Cas13b 기반 프로그램 가능한 YTHDF1 및 YTHDF2 시스템은 각각 155 및 126 kDa이다(도 20). CIRTS-1은 현재까지 발견된 가장 작은 DNA 표적화 Cas 단백질인 Cas14a 및 가장 작은 Cas12g RNA 표적화 단백질보다 훨씬 작다(Harrington et al., 2018; Yan et al., 2019).

[0288] 번역의 관점에서, CIRTS는 몇 가지 주요 이점과 기회를 제공해야 한다. CIRTS의 인간화된 특성은 면역 반응을 회피하는 경로를 제공하여 지속-전달 치료법의 가능성을 열어줄 것이다. CIRTS에서 인간 단백질 간의 융합은 면역계가 반응할 수 있는 설계에 잠재적인 한계를 나타내지만(Glaesner et al., 2010), 이는 원칙적으로 조작하여 해결될 수 있는 문제이다. 본 발명자들이 본 발명자들의 조작된 작제물에서 가장 가능성이 높은 MHC I 결합 펩타이드의 면역원성을 컴퓨터로 예측했을 때, 본 발명자들은 완전히 인간화된 CIRTS가 면역 반응을 일으키는 더 낮은 경향을 나타내지만(도 27c), 설계의 한계가 나타나는 위치를 발견하기 위해 추가 실험적 시험이 필요하다는 것을 발견했다.

[0289] 현재 CIRT 시스템에는 몇 가지 알려진 도전 과제가 남아 있다. 첫째, 현재 형태의 대체 헤어핀 결합 단백질 SLBP에는 내인성 RNA 헤어핀 결합 파트너가 있어 줄기 루프 RNA 트래픽킹에 영향을 미칠 수도 있다. 본 발명자들의 융합 작제물의 내인성 효과를 최소화하기 위해, 본 발명자들은 시스템에 헤어핀 인식에 필요한 최소 RNA 인식 모티프(RRM)만을 포함하고 다른 단백질 또는 핵산과의 잠재적인 상호작용 영역을 제외하였다. 마찬가지로, 본 발명자들은 잠재적인 내인성 상호작용을 피하기 위해 요구되는 RNA 헤어핀을 가능한 한 작게 유지하려고 노력했다. 줄기 루프 헤어핀은 이미 매우 짧아서 더 이상 자를 수 없었지만 본 발명자들은 TAR 헤어핀에 대한 TBP6.7 결합에 필요한 최소한으로 요구되는 영역만을 사용하도록 선택하여, 원래 헤어핀 길이의 절반 미만인 gRNA를 생성하였다. 둘째, 현재 형태의 양이온성 펩타이드인 β-디펜신 3은 이론적으로 여전히 세포내 결합 파트너와 상호작용하고 원치 않는 생물학적 반응을 끌어낼 수 있다. 그러나, 인간 β-디펜신 3은 광범위하게 연구되었으며 그 기능에 대한 구조적 기반이 해명되었다(Dhople et al., 2006; Kluver et al., 2005). 본 발명자들은 인간 부분 기반, 직교 RNA 표적화 시스템을 조작하기 위해 CIRTS에 필요한 높게 하전된 특성을 유지하지만 내인성 기능을 폐기하는 β-디펜신 3 돌연변이체를 조작하기 위한 기초로서 이러한 지식을 활용할 수 있다.

[0290] 더 넓은 관점에서, CIRTS 플랫폼은 인간 단백질 툴박스에 포함된 부분을 결합하여 새로운 특성을 가진 단백질을 조작할 수 있는 가능성을 보여준다. 제시된 CIRTS는 최소한의 단백질 조작 및 최적화 노력을 통해 만들어졌지만 자연적으로 진화한 CRISPR/Cas 시스템과 거의 같은 기능을 한다. 특히, 본 발명자들이 내인성 뉴클레아제에 의한 Cas13b 기반 녹다운과 YTHDF2(도 26c)를 CIRTS-매개된 녹다운에 전달함으로써 Cas13b 기반 녹다운을 비교했을 때, 본 발명자들은 Cas13b 뉴클레아제가 CIRTS 뉴클레아제보다 실질적으로 더 나은 성능을 나타내지만, 본 발명자들은 YTHDF2에 의해 매개되는 녹다운에 차이가 없음을 발견했으며, 이는 현재 상태의 CIRTS-3(YTHDF2)이 후성전사체를 효과적으로 연구하는 데 이미 사용될 수 있음을 시사한다. 유도 진화를 사용하여 CIRTS를 최적화하는 추가 작업은 포유류 시스템에서 향상된 성능을 가진 변이체를 산출할 가능성이 높다(Hu et al., 2018). 표적 부위 설계 및 효과기 단백질 맥락적 요건을 이해하면 CIRTS 시스템 성능도 향상될 것이며, 이용되고 있는 후성전사체 경로에 대한 더 나은 이해는 본 발명자들이 더 나은 시스템을 설계할 수 있게 해줄 것이다. 또한 고유한 RNA 제어 특성을 보유할 가능성이 있는 인간 단백질체 내에 함유된 다양한 다른 조절 단백질이 있으며(Dominguez et al., 2018), 이는 CIRTS와 결합하여 기능적 특성화 및 잠재적 번역 적용 모두에 대해 각 단백질의 프로그램 가능한 버전을 생성할 수 있다. CIRTS 시스템은 RNA 조절을 연구하고 활용하기 위한 새로운 접근 방식을 제공하며 질병 치료를 위한 세포 조절에 개입할 수 있는 미래의 기회를 열어줄 것이다.

[0291] * * *

[0292] 본원에 개시되고 청구되는 모든 방법은 본 개시 내용에 비추어 과도한 실험 없이 만들어지고 실행될 수 있다. 본 발명의 조성물 및 방법이 바람직한 구현예의 관점에서 설명되었지만, 본 발명의 개념, 사상 및 범위를 벗어나지 않고 본원에 기술된 방법의 방법 및 단계 또는 단계의 순서에 변형이 적용될 수 있음이 당업자에게 명백할 것이다. 보다 구체적으로, 화학적 및 생리학적으로 모두 관련된 특정 제제가 동일하거나 유사한 결과가 달성되는 본원에 기재된 제제를 대체할 수 있음이 명백할 것이다. 당업자에게 명백한 그러한 모든 유사한 대체물 및 수정물은 첨부된 청구범위에 의해 정의된 본 발명의 사상, 범위 및 개념 내에 있는 것으로 간주된다.

[0293] **참고문헌**

[0294] 예시적인 절차 또는 본원에 기술된 것에 보충적인 기타 세부사항을 제공하는 범위에서, 다음 참고문헌은 본원에 참고로 구체적으로 포함된다.

[0295] Abudayyeh, O.O., Gootenberg, J.S., Essletzbichler, P., Han, S., Joung, J., Belanto, J.J., Verdine, V., Cox, D.B.T., Kellner, M.J., Regev, A., et al. (2017). RNA targeting with CRISPR-Cas13. Nature 550, 280.

[0296] Abudayyeh, O.O., Gootenberg, J.S., Konermann, S., Joung, J., Slaymaker, I.M., Cox, D.B.T., Shmakov, S., Makarova, K.S., Semenova, E., Minakhin, L., et al. (2016). C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. Science 353, aaf5573.

[0297] Adamala, K.P., Martin-Alarcon, D.A., and Boyden, E.S. (2016). Programmable RNA-binding protein composed of repeats of a single modular unit. Proc Natl Acad Sci U S A 113, E2579-2588.

[0298] Batra, R., Nelles, D.A., Pirie, E., Blue, S.M., Marina, R.J., Wang, H., Chaim, I.A., Thomas, J.D., Zhang, N., Nguyen, V., et al. (2017). Elimination of Toxic Microsatellite Repeat Expansion RNA by RNA-Targeting Cas9. Cell 170, 899-912 e810.

[0299] Blakeley, B.D., and McNaughton, B.R. (2014). Synthetic RNA recognition motifs that selectively recognize HIV-1 trans-activation response element hairpin RNA. ACS Chem Biol 9, 1320-1329.

[0300] Chandrasegaran, S., and Carroll, D. (2016). Origins of Programmable Nucleases for Genome Engineering. J Mol Biol 428, 963-989.

[0301] Chavez, A., Scheiman, J., Vora, S., Pruitt, B.W., Tuttle, M., P R Iyer, E., Lin, S., Kiani, S., Guzman, C.D., Wiegand, D.J., et al. (2015). Highly efficient Cas9-mediated transcriptional programming. Nat Methods 12, 326.

[0302] Choudhury, R., Tsai, Y.S., Dominguez, D., Wang, Y., and Wang, Z. (2012). Engineering RNA endonucleases with customized sequence specificities. Nat Commun 3, 1147.

[0303] Chu, V.T., Weber, T., Wefers, B., Wurst, W., Sander, S., Rajewsky, K., and Kuhn, R. (2015). Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. Nat Biotechnol 33, 543-548.

[0304] Coller, J., and Wickens, M. (2007). Tethered function assays: an adaptable approach to study RNA regulatory proteins. Methods Enzymol 429, 299-321.

[0305] Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., et al. (2013). Multiplex genome engineering using CRISPR/Cas systems. Science 339, 819-823.

[0306] Cox, D.B.T., Gootenberg, J.S., Abudayyeh, O.O., Franklin, B., Kellner, M.J., Joung, J., and Zhang, F. (2017). RNA editing with CRISPR-Cas13. Science 358, 1019-1027.

[0307] Crawford, D.W., Blakeley, B.D., Chen, P.H., Sherpa, C., Le Grice, S.F., Laird-Offringa, I.A., and McNaughton, B.R. (2016). An Evolved RNA Recognition Motif That Suppresses HIV-1 Tat/TAR-Dependent Transcription. ACS Chem Biol 11, 2206-2215.

[0308] Cronican, J.J., Beier, K.T., Davis, T.N., Tseng, J.C., Li, W., Thompson, D.B., Shih, A.F., May, E.M., Cepko, C.L., Kung, A.L., et al. (2011). A class of human proteins that deliver functional proteins into mammalian cells in vitro and in vivo. Chem Biol 18, 833-838.

[0309] Desjarlais, J.R., and Berg, J.M. (1993). Use of a zinc-finger consensus sequence framework and specificity rules to design specific DNA binding proteins. Proceedings of the National Academy of Sciences 90, 2256.

[0310] Dhople, V., Krukemeyer, A., and Ramamoorthy, A. (2006). The human beta-defensin-3, an antibacterial

peptide with multiple biological functions. Biochim Biophys Acta 1758, 1499-1512.

[0311] Dominguez, D., Freese, P., Alexis, M.S., Su, A., Hochman, M., Palden, T., Bazile, C., Lambert, N.J., Van Nostrand, E.L., Pratt, G.A., et al. (2018). Sequence, Structure, and Context Preferences of Human RNA Binding Proteins. Mol Cell 70, 854-867.e859.

[0312] Du, D., Roguev, A., Gordon, D.E., Chen, M., Chen, S.-H., Shales, M., Shen, J.P., Ideker, T., Mali, P., Qi, L.S., et al. (2017). Genetic interaction mapping in mammalian cells using CRISPR interference. Nat Methods 14, 577.

[0313] Eberle, A.B., Lykke-Andersen, S., Muhlemann, O., and Jensen, T.H. (2009). SMG6 promotes endonucleolytic cleavage of nonsense mRNA in human cells. Nat Struct Mol Biol 16, 49-55.

[0314] Filipovska, A., Razif, M.F., Nygard, K.K., and Rackham, O. (2011). A universal code for RNA recognition by PUF proteins. Nat Chem Biol 7, 425-427.

[0315] Fuxman Bass, J.I., Sahni, N., Shrestha, S., Garcia-Gonzalez, A., Mori, A., Bhat, N., Yi, S., Hill, D.E., Vidal, M., and Walhout, A.J.M. (2015). Human gene-centered transcription factor networks for enhancers and disease variants. Cell 161, 661-673.

[0316] Gaudelli, N.M., Komor, A.C., Rees, H.A., Packer, M.S., Badran, A.H., Bryson, D.I., and Liu, D.R. (2017). Programmable base editing of A*?*T to G*C in genomic DNA without DNA cleavage. Nature 551, 464.

[0317] Gilbert, L.A., Larson, M.H., Morsut, L., Liu, Z., Brar, G.A., Torres, S.E., Stern-Ginossar, N., Brandman, O., Whitehead, E.H., Doudna, J.A., et al. (2013). CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. Cell 154, 442-451.

[0318] Glaesner, W., Vick, A.M., Millican, R., Ellis, B., Tschang, S.H., Tian, Y., Bokvist, K., Brenner, M., Koester, A., Porksen, N., et al. (2010). Engineering and characterization of the long-acting glucagon-like peptide-1 analogue LY2189265, an Fc fusion protein. Diabetes Metab Res Rev 26, 287-296.

[0319] Gootenberg, J.S., Abudayyeh, O.O., Kellner, M.J., Joung, J., Collins, J.J., and Zhang, F. (2018). Multiplexed and portable nucleic acid detection platform with Cas13, Cas12a, and Csm6. Science 360, 439.

[0320] Harrington, L.B., Burstein, D., Chen, J.S., Paez-Espino, D., Ma, E., Witte, I.P., Cofsky, J.C., Kyrpides, N.C., Banfield, J.F., and Doudna, J.A. (2018). Programmed DNA destruction by miniature CRISPR-Cas14 enzymes. Science 362, 839.

[0321] Hilton, I.B., D'Ippolito, A.M., Vockley, C.M., Thakore, P.I., Crawford, G.E., Reddy, T.E., and Gersbach, C.A. (2015). Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. Nat Biotechnol 33, 510.

[0322] Hockemeyer, D., Wang, H., Kiani, S., Lai, C.S., Gao, Q., Cassady, J.P., Cost, G.J., Zhang, L., Santiago, Y., Miller, J.C., et al. (2011). Genetic engineering of human pluripotent cells using TALE nucleases. Nat Biotechnol 29, 731-734.

[0323] Hu, J.H., Miller, S.M., Geurts, M.H., Tang, W., Chen, L., Sun, N., Zeina, C.M., Gao, X., Rees, H.A., Lin, Z., et al. (2018). Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. Nature 556, 57.

[0324] Jiang, W., Bikard, D., Cox, D., Zhang, F., and Marraffini, L.A. (2013). RNA-guided editing of bacterial genomes using CRISPR-Cas systems. Nat Biotechnol 31, 233.

[0325] Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science 337, 816-821.

[0326] Joung, J.K., and Sander, J.D. (2012). TALENs: a widely applicable technology for targeted genome editing. Nature Reviews Molecular Cell Biology 14, 49.

[0327] Kearns, N.A., Pham, H., Tabak, B., Genga, R.M., Silverstein, N.J., Garber, M., and Maehr, R. (2015). Functional annotation of native enhancers with a Cas9-histone demethylase fusion. Nat Methods 12, 401.

[0328] Kim, S., Koo, T., Jee, H.G., Cho, H.Y., Lee, G., Lim, D.G., Shin, H.S., and Kim, J.S. (2018). CRISPR RNAs trigger innate immune responses in human cells. Genome Res.

[0329] Kluver, E., Schulz-Maronde, S., Scheid, S., Meyer, B., Forssmann, W.G., and Adermann, K. (2005). Structure-activity relation of human beta-defensin 3: influence of disulfide bonds and cysteine substitution on antimicrobial activity and cytotoxicity. Biochemistry 44, 9804-9816.

[0330] Knott, G.J., and Doudna, J.A. (2018). CRISPR-Cas guides the future of genetic engineering. Science 361, 866-869.

[0331] Koblan, L.W., Doman, J.L., Wilson, C., Levy, J.M., Tay, T., Newby, G.A., Maianti, J.P., Raguram, A., and Liu, D.R. (2018). Improving cytidine and adenine base editors by expression optimization and ancestral reconstruction. Nat Biotechnol 36, 843-846.

[0332] Komor, A.C., Kim, Y.B., Packer, M.S., Zuris, J.A., and Liu, D.R. (2016). Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. Nature 533, 420.

[0333] Konermann, S., Lotfy, P., Brideau, N.J., Oki, J., Shokhirev, M.N., and Hsu, P.D. (2018). Transcriptome Engineering with RNA-Targeting Type VI-D CRISPR Effectors. Cell 173, 665-676.e614.

[0334] Liao, H.-K., Hatanaka, F., Araoka, T., Reddy, P., Wu, M.-Z., Sui, Y., Yamauchi, T., Sakurai, M., O'Keefe, D.D., Núñez-Delicado, E., et al. (2017). In Vivo Target Gene Activation via CRISPR/Cas9-Mediated Trans-epigenetic Modulation. Cell 171, 1495-1507.e1415.

[0335] Lin, S., Staahl, B.T., Alla, R.K., and Doudna, J.A. (2014). Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. Elife 3, e04766.

[0336] Liu, L., Li, X., Ma, J., Li, Z., You, L., Wang, J., Wang, M., Zhang, X., and Wang, Y. (2017). The Molecular Architecture for RNA-Guided RNA Cleavage by Cas13a. Cell 170, 714-726 e710.

[0337] Maeder, M.L., Linder, S.J., Cascio, V.M., Fu, Y., Ho, Q.H., and Joung, J.K. (2013). CRISPR RNA-guided activation of endogenous human genes. Nat Methods 10, 977-979.

[0338] Mali, P., Aach, J., Stranges, P.B., Esvelt, K.M., Moosburner, M., Kosuri, S., Yang, L., and Church, G.M. (2013a). CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. Nat Biotechnol 31, 833-838.

[0339] Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., and Church, G.M. (2013b). RNA-guided human genome engineering via Cas9. Science 339, 823-826.

[0340] Monteys, A.M., Ebanks, S.A., Keiser, M.S., and Davidson, B.L. (2017). CRISPR/Cas9 Editing of the Mutant Huntingtin Allele In Vitro and In Vivo. Mol Ther 25, 12-23.

[0341] Montiel-Gonzalez, M.F., Vallecillo-Viejo, I.C., and Rosenthal, J.J. (2016). An efficient system for selectively altering genetic information within mRNAs. Nucleic Acids Res 44, e157.

[0342] Nelles, D.A., Fang, M.Y., O'Connell, M.R., Xu, J.L., Markmiller, S.J., Doudna, J.A., and Yeo, G.W. (2016). Programmable RNA Tracking in Live Cells with CRISPR/Cas9. Cell 165, 488-496.

[0343] Nishikura, K. (2010). Functions and regulation of RNA editing by ADAR deaminases. Annu Rev Biochem 79, 321-349.

[0344] O'Connell, M.R., Oakes, B.L., Sternberg, S.H., East-Seletsky, A., Kaplan, M., and Doudna, J.A. (2014). Programmable RNA recognition and cleavage by CRISPR/Cas9. Nature 516, 263-266.

[0345] O'Connell, M.R., Oakes, B.L., Sternberg, S.H., East-Seletsky, A., Kaplan, M., and Doudna, J.A. (2014). Programmable RNA recognition and cleavage by CRISPR/Cas9. Nature 516, 263.

[0346] Perez-Pinera, P., Kocak, D.D., Vockley, C.M., Adler, A.F., Kabadi, A.M., Polstein, L.R., Thakore, P.I., Glass, K.A., Ousterout, D.G., Leong, K.W., et al. (2013). RNA-guided gene activation by CRISPR-Cas9-based transcription factors. Nat Methods 10, 973-976.

[0347] Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P., and Lim, W.A. (2013). Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell 152, 1173-1183.

[0348] Rauch, S., He, C., and Dickinson, B.C. (2018). Targeted m(6)A Reader Proteins To Study Epitranscriptomic Regulation of Single RNAs. J Am Chem Soc 140, 11974-11981.

[0349] Roundtree, I.A., Evans, M.E., Pan, T., and He, C. (2017). Dynamic RNA Modifications in Gene Expression Regulation. Cell 169, 1187-1200.

[0350] Rutkauskas, M., Sinkunas, T., Songailiene, I., Tikhomirova, Maria S., Siksnys, V., and Seidel, R. (2015). Directional R-Loop Formation by the CRISPR-Cas Surveillance Complex Cascade Provides Efficient Off-Target Site Rejection. Cell Reports 10, 1534-1543.

[0351] Saviζ, N., and Schwank, G. (2016). Advances in therapeutic CRISPR/Cas9 genome editing. Translational Research 168, 15-21.

[0352] Schierling, B., Dannemann, N., Gabsalilow, L., Wende, W., Cathomen, T., and Pingoud, A. (2012). A novel zinc-finger nuclease platform with a sequence-specific cleavage module. Nucleic Acids Res 40, 2623-2638.

[0353] Sinnamon, J.R., Kim, S.Y., Corson, G.M., Song, Z., Nakai, H., Adelman, J.P., and Mandel, G. (2017). Site-directed RNA repair of endogenous Mecp2 RNA in neurons. Proc Natl Acad Sci U S A 114, E9395-E9402.

[0354] Sternberg, S.H., Redding, S., Jinek, M., Greene, E.C., and Doudna, J.A. (2014). DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. Nature 507, 62.

[0355] Strutt, S.C., Torrez, R.M., Kaya, E., Negrete, O.A., and Doudna, J.A. (2018). RNA-dependent RNA targeting by CRISPR-Cas9. eLife 7, e32724.

[0356] Sundaram, G.M., Ismail, H.M., Bashir, M., Muhuri, M., Vaz, C., Nama, S., Ow, G.S., Vladimirovna, I.A., Ramalingam, R., Burke, B., et al. (2017). EGF hijacks miR-198/FSTL1 wound-healing switch and steers a two-pronged pathway toward metastasis. J Exp Med 214, 2889-2900.

[0357] Szczelkun, M.D., Tikhomirova, M.S., Sinkunas, T., Gasiunas, G., Karvelis, T., Pschera, P., Siksnys, V., and Seidel, R. (2014). Direct observation of R-loop formation by single RNA-guided Cas9 and Cascade effector complexes. Proceedings of the National Academy of Sciences 111, 9798.

[0358] Tambe, A., East-Seletsky, A., Knott, G.J., Doudna, J.A., and O'Connell, M.R. (2018). RNA Binding and HEPN-Nuclease Activation Are Decoupled in CRISPR-Cas13a. Cell Rep 24, 1025-1036.

[0359] Vogel, P., Moschref, M., Li, Q., Merkle, T., Selvasaravanan, K.D., Li, J.B., and Stafforst, T. (2018). Efficient and precise editing of endogenous transcripts with SNAP-tagged ADARs. Nat Methods 15, 535-538.

[0360] Wagner, D.L., Amini, L., Wendering, D.J., Burkhardt, L.-M., Akyüz, L., Reinke, P., Volk, H.-D., and Schmueck-Henneresse, M. (2018). High prevalence of Streptococcus pyogenes Cas9-reactive T cells within the adult human population. Nat Med.

[0361] Wiedenheft, B., Sternberg, S.H., and Doudna, J.A. (2012). RNA-guided genetic silencing systems in bacteria and archaea. Nature 482, 331.

[0362] Yan, W.X., Chong, S., Zhang, H., Makarova, K.S., Koonin, E.V., Cheng, D.R., and Scott, D.A. (2018). Cas13d Is a Compact RNA-Targeting Type VI CRISPR Effector Positively Modulated by a WYL-Domain-

Containing Accessory Protein. Mol Cell 70, 327-339 e325.

[0363] Yi, L., and Li, J. (2016). CRISPR-Cas9 therapeutics in cancer: promising strategies and present challenges. Biochim Biophys Acta 1866, 197-207.

[0364] Zhao, B.S., Roundtree, I.A., and He, C. (2017). Post-transcriptional gene regulation by mRNA modifications. Nat Rev Mol Cell Biol 18, 31-42.

[0365] Zhou, Z., Dell'Orco, M., Saldarelli, P., Turturo, C., Minafra, A., and Martelli, G.P. (2006). Identification of an RNA-silencing suppressor in the genome of Grapevine virus A. J Gen Virol 87, 2387-2395.

[0366] Guilinger, J.P., Thompson, D.B., and Liu, D.R. (2014). Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. Nat Biotechnol 32, 577-582.

[0367] Abil, Z., Denard, C.A., and Zhao, H. (2014). Modular assembly of designer PUF proteins for specific post-transcriptional regulation of endogenous RNA. J Biol Eng 8, 7.

[0368] Andreatta, M., and Nielsen, M. (2016). Gapped sequence alignment using artificial neural networks: application to the MHC class I system. Bioinformatics 32, 511-517.

[0369] Bao, Z., Jain, S., Jaroenpuntaruk, V., and Zhao, H. (2017). Orthogonal Genetic Regulation in Human Cells Using Chemically Induced CRISPR/Cas9 Activators. ACS Synth Biol 6, 686-693.

[0370] Boyle, E.A., Andreasson, J.O.L., Chircus, L.M., Sternberg, S.H., Wu, M.J., Guegler, C.K., Doudna, J.A., and Greenleaf, W.J. (2017). High-throughput biochemical profiling reveals sequence determinants of dCas9 off-target binding and unbinding. Proc Natl Acad Sci U S A 114, 5461-5466.

[0371] Bray, N.L., Pimentel, H., Melsted, P., and Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. Nat Biotechnol 34, 525.

[0372] Calis, J.J., Maybeno, M., Greenbaum, J.A., Weiskopf, D., De Silva, A.D., Sette, A., Kesmir, C., and Peters, B. (2013). Properties of MHC class I presented peptides that enhance immunogenicity. PLoS Comput Biol 9, e1003266.

[0373] Charlesworth, C., Deshpande, P., Dever, D., Dejene, B., Gomez-Ospina, N., Mantri, S., Pavel-Dinu, M., Camarena, J., Weinberg, K., and Porteus, M. Identification of Pre-Existing Adaptive Immunity to Cas9 Proteins in Humans. bioRxiv.

[0374] Du, H., Zhao, Y., He, J., Zhang, Y., Xi, H., Liu, M., Ma, J., and Wu, L. (2016). YTHDF2 destabilizes m6A-containing RNA through direct recruitment of the CCR4-NOT deadenylase complex. Nat Commun 7, 12626.

[0375] Gao, G., Vandenberghe, L.H., and Wilson, J.M. (2005). New recombinant serotypes of AAV vectors. Curr Gene Ther 5, 285-297.

[0376] Gao, Y., Xiong, X., Wong, S., Charles, E.J., Lim, W.A., and Qi, L.S. (2016). Complex transcriptional modulation with orthogonal and inducible dCas9 regulators. Nat Methods 13, 1043-1049.

[0377] Jurtz, V., Paul, S., Andreatta, M., Marcatili, P., Peters, B., and Nielsen, M. (2017). NetMHCpan-4.0: Improved Peptide-MHC Class I Interaction Predictions Integrating Eluted Ligand and Peptide Binding Affinity Data. J Immunol 199, 3360-3368.

[0378] Kearns, N.A., Pham, H., Tabak, B., Genga, R.M., Silverstein, N.J., Garber, M., and Maehr, R. (2015). Functional annotation of native enhancers with a Cas9-histone demethylase fusion. Nat Methods 12, 401.

[0379] Kuttan, A., and Bass, B.L. (2012). Mechanistic insights into editing-site specificity of ADARs. Proc Natl Acad Sci U S A 109, E3295-3304.

[0380] Meyer, Kate D., Patil, Deepak P., Zhou, J., Zinoviev, A., Skabkin, Maxim A., Elemento, O., Pestova, Tatyana V., Qian, S.-B., and Jaffrey, Samie R. (2015). 5′ UTR m6A Promotes Cap-Independent

Translation. Cell 163, 999-1010.

[0381] Moutaftsi, M., Peters, B., Pasquetto, V., Tscharke, D.C., Sidney, J., Bui, H.H., Grey, H., and Sette, A. (2006). A consensus epitope prediction approach identifies the breadth of murine T(CD8+)-cell responses to vaccinia virus. Nat Biotechnol 24, 817-819.

[0382] Naso, M.F., Tomkowicz, B., Perry, W.L., 3rd, and Strohl, W.R. (2017). Adeno-Associated Virus (AAV) as a Vector for Gene Therapy. BioDrugs 31, 317-334.

[0383] Peters, B., and Sette, A. (2005). Generating quantitative models describing the sequence specificity of biological processes with the stabilized matrix method. BMC Bioinformatics 6, 132.

[0384] Pimentel, H., Bray, N.L., Puente, S., Melsted, P., and Pachter, L. (2017). Differential analysis of RNA-seq incorporating quantification uncertainty. Nat Methods 14, 687.

[0385] Sidney, J., Assarsson, E., Moore, C., Ngo, S., Pinilla, C., Sette, A., and Peters, B. (2008). Quantitative peptide binding motifs for 19 human and mouse MHC class I molecules derived using positional scanning combinatorial peptide libraries. Immunome Res 4, 2.

[0386] Simhadri, V.L., McGill, J., McMahon, S., Wang, J., Jiang, H., and Sauna, Z.E. (2018). Prevalence of Pre-existing Antibodies to CRISPR-Associated Nuclease Cas9 in the USA Population. Mol Ther Methods Clin Dev 10, 105-112.

[0387] Vasileva, A., and Jessberger, R. (2005). Precise hit: adeno-associated virus in gene targeting. Nat Rev Microbiol 3, 837-847.

[0388] Wang, X., Lu, Z., Gomez, A., Hon, G.C., Yue, Y., Han, D., Fu, Y., Parisien, M., Dai, Q., Jia, G., et al. (2013). N6-methyladenosine-dependent regulation of messenger RNA stability. Nature 505, 117.

[0389] Wang, X., Zhao, Boxuan S., Roundtree, Ian A., Lu, Z., Han, D., Ma, H., Weng, X., Chen, K., Shi, H., and He, C. (2015). N6-methyladenosine Modulates Messenger RNA Translation Efficiency. Cell 161, 1388-1399.

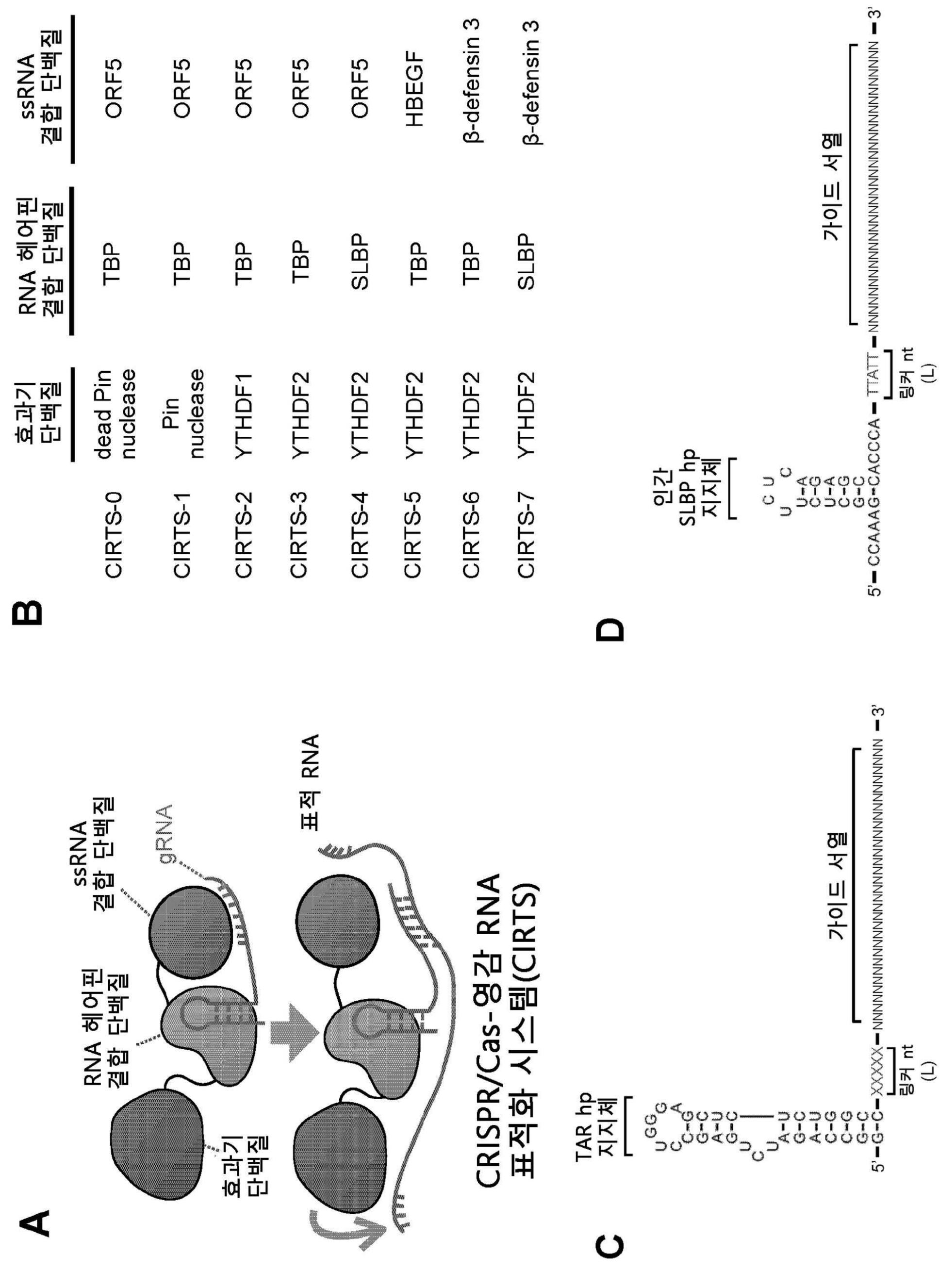
[0390] Wu, Z., Yang, H., and Colosi, P. (2010). Effect of genome size on AAV vector packaging. Mol Ther 18, 80-86.

[0391] Xiao, W., Adhikari, S., Dahal, U., Chen, Y.-S., Hao, Y.-J., Sun, B.-F., Sun, H.-Y., Li, A., Ping, X.-L., Lai, W.-Y., et al. (2016). Nuclear m6A Reader YTHDC1 Regulates mRNA Splicing. Mol Cell 61, 507-519.

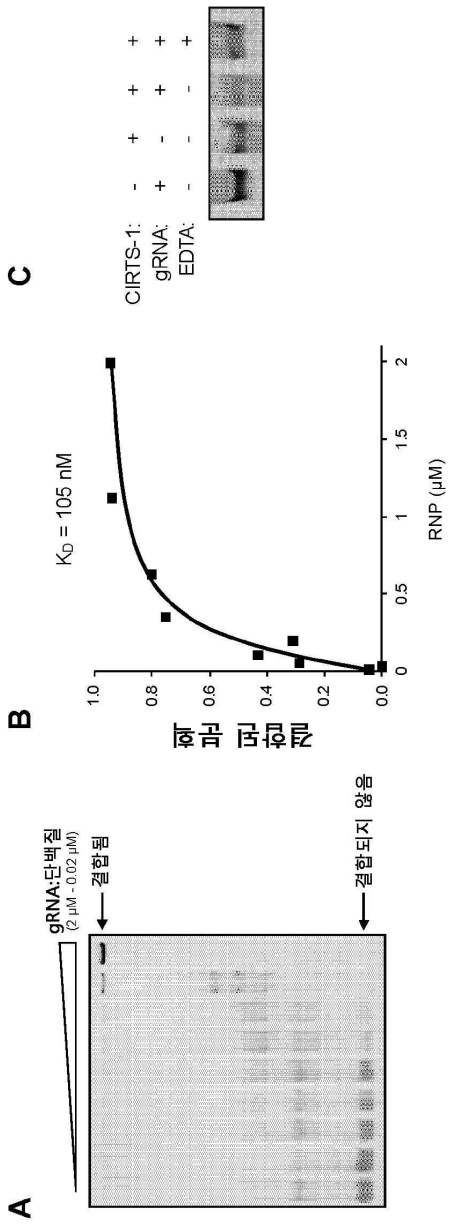
[0392] Yan, W.X., Hunnewell, P., Alfonse, L.E., Carte, J.M., Keston-Smith, E., Sothiselvam, S., Garrity, A.J., Chong, S., Makarova, K.S., Koonin, E.V., et al. (2019). Functionally diverse type V CRISPR-Cas systems. Science 363, 88-91.

도면

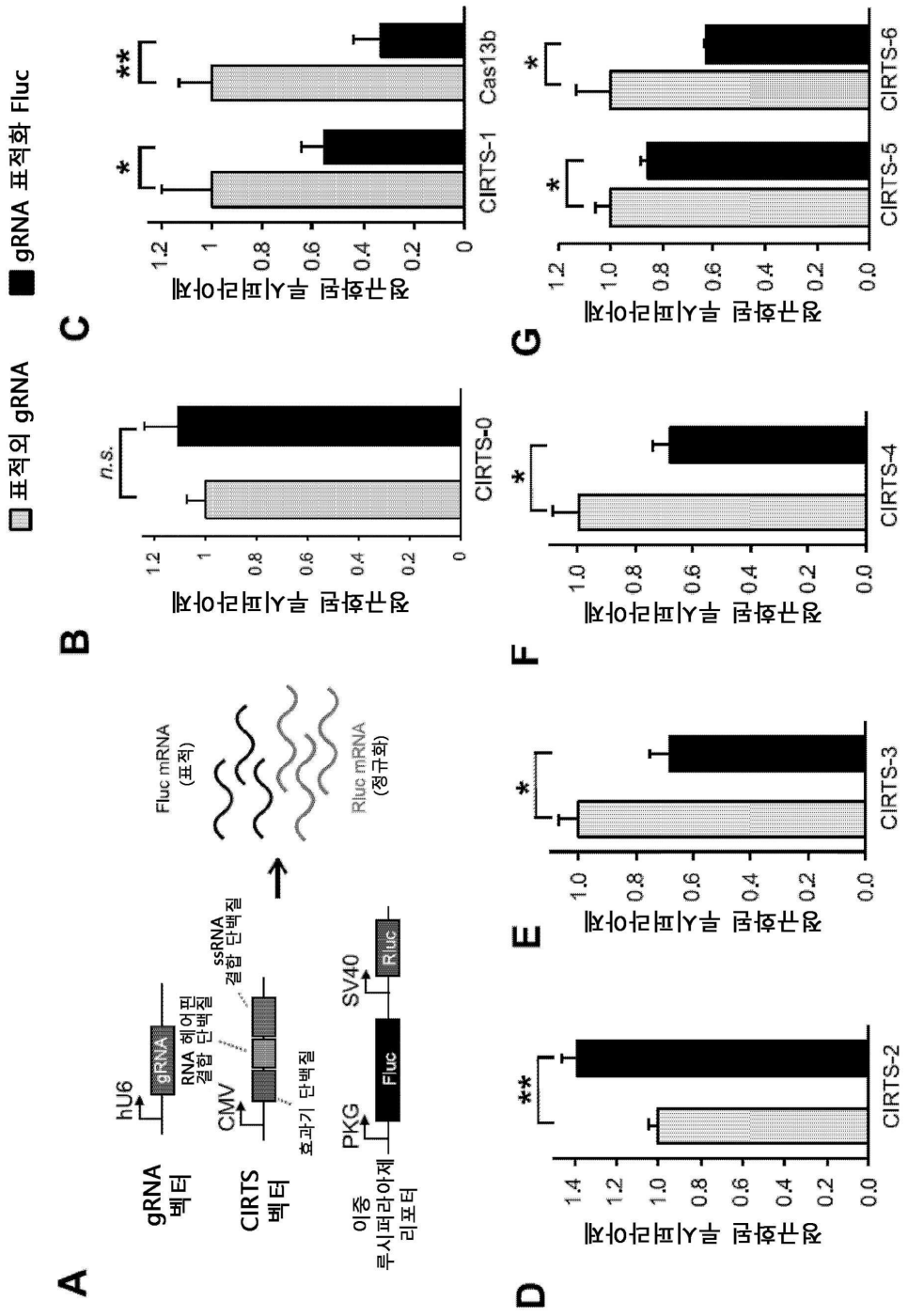
도면1


A
효과기 단백질
RNA 헤어핀 결합 단백질
ssRNA 결합 단백질
gRNA
표적 RNA
CRISPR/Cas-영감 RNA 표적화 시스템(CIRTS)
B
효과기 단백질
RNA 헤어핀 결합 단백질
ssRNA 결합 단백질
CIRTS-0 dead Pin nuclease TBP ORF5
CIRTS-1 Pin nuclease TBP ORF5
CIRTS-2 YTHDF1 TBP ORF5
CIRTS-3 YTHDF2 TBP ORF5
CIRTS-4 YTHDF2 SLBP ORF5
CIRTS-5 YTHDF2 TBP HBEGF
CIRTS-6 YTHDF2 TBP β-defensin 3
CIRTS-7 YTHDF2 SLBP β-defensin 3
C
TAR hp 지지체
5'–G–C–XXXXX–NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN–3'
링커 nt (L)
가이드 서열
D
인간 SLBP hp 지지체
5'–CCAAAG–CACCCA–TTATT–NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN–3'
링커 nt (L)
가이드 서열

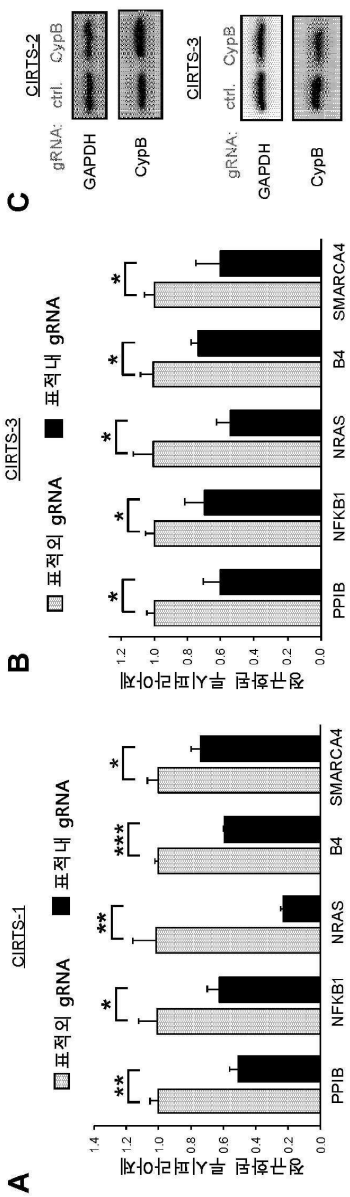
도면2


A
gRNA:단백질
(2 μM - 0.02 μM)
결합됨
결합되지 않음
B
결합된 분획
1.0
0.8
0.6
0.4
0.2
0.0
0
0.5
1
1.5
2
RNP (μM)
K_D = 105 nM
C
CIRTS-1:
gRNA:
EDTA:

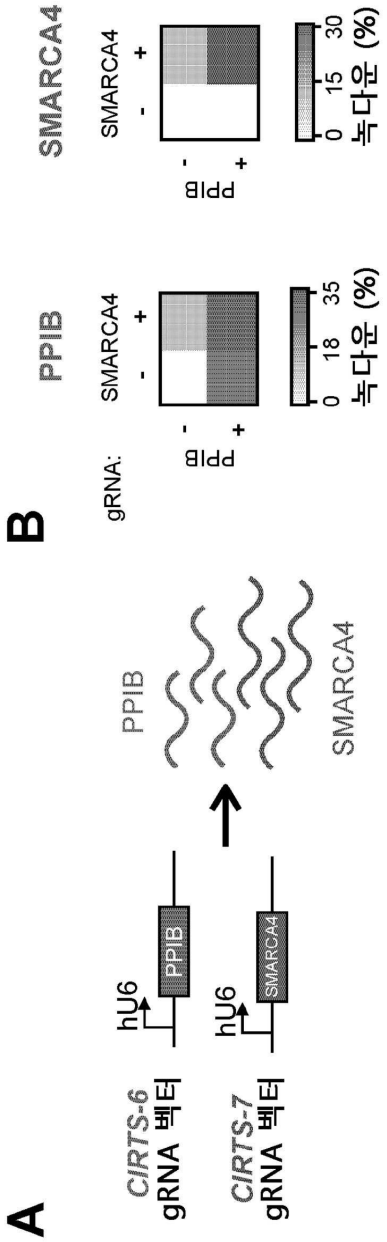
도면3


표적외 gRNA
gRNA 표적화 Fluc
A
gRNA 벡터
hU6
gRNA
CIRTS 벡터
CMV
RNA 헤어핀 결합 단백질
ssRNA 결합 단백질
효과기 단백질
이중 루시퍼라아제 리포터
PKG
Fluc
SV40
Rluc
Fluc mRNA (표적)
Rluc mRNA (정규화)
B
정규화된 루시퍼라아제
n.s.
CIRTS-0
C
CIRTS-1
Cas13b
D
CIRTS-2
E
CIRTS-3
F
CIRTS-4
G
CIRTS-5
CIRTS-6

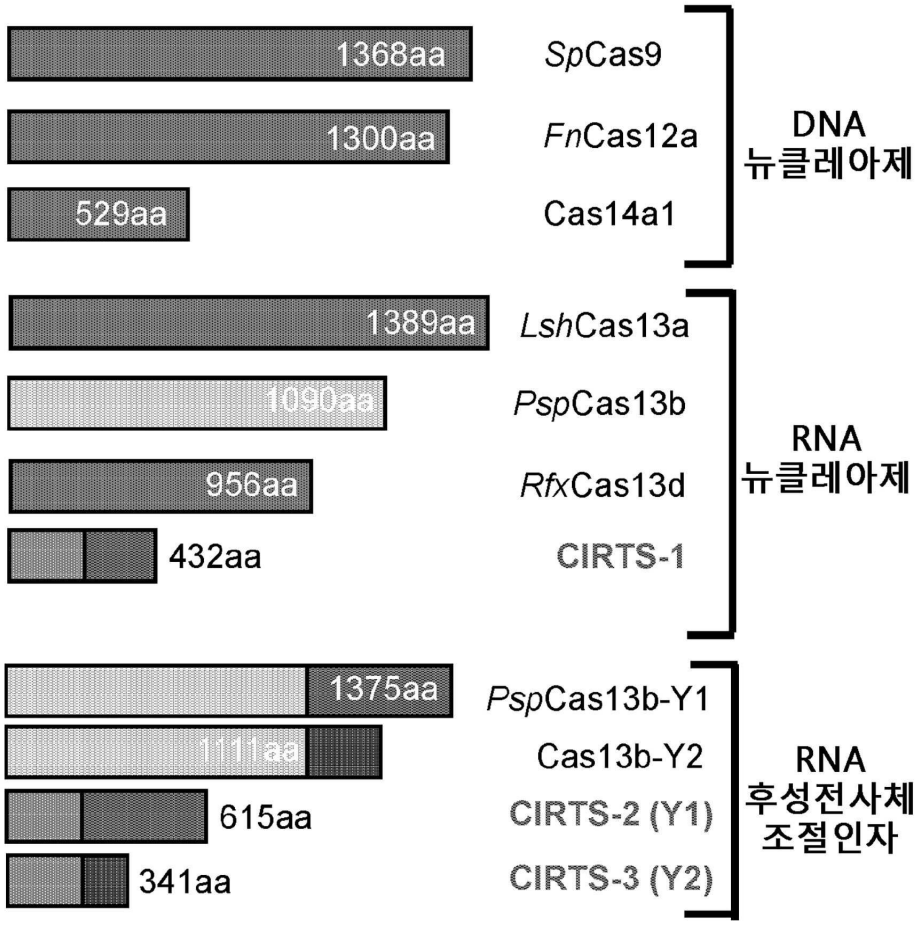
도면4


A
CIRTS-1
표적외 gRNA
표적내 gRNA
정규화된 루시퍼라아제
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
**
*
**
***
*
PPIB
NFKB1
NRAS
B4
SMARCA4
B
CIRTS-3
표적외 gRNA
표적내 gRNA
정규화된 루시퍼라아제
1.2
1.0
0.8
0.6
0.4
0.2
0.0
*
*
*
*
*
PPIB
NFKB1
NRAS
B4
SMARCA4
C
CIRTS-2
gRNA: ctrl. CypB
GAPDH
CypB
CIRTS-3
gRNA: ctrl. CypB
GAPDH
CypB

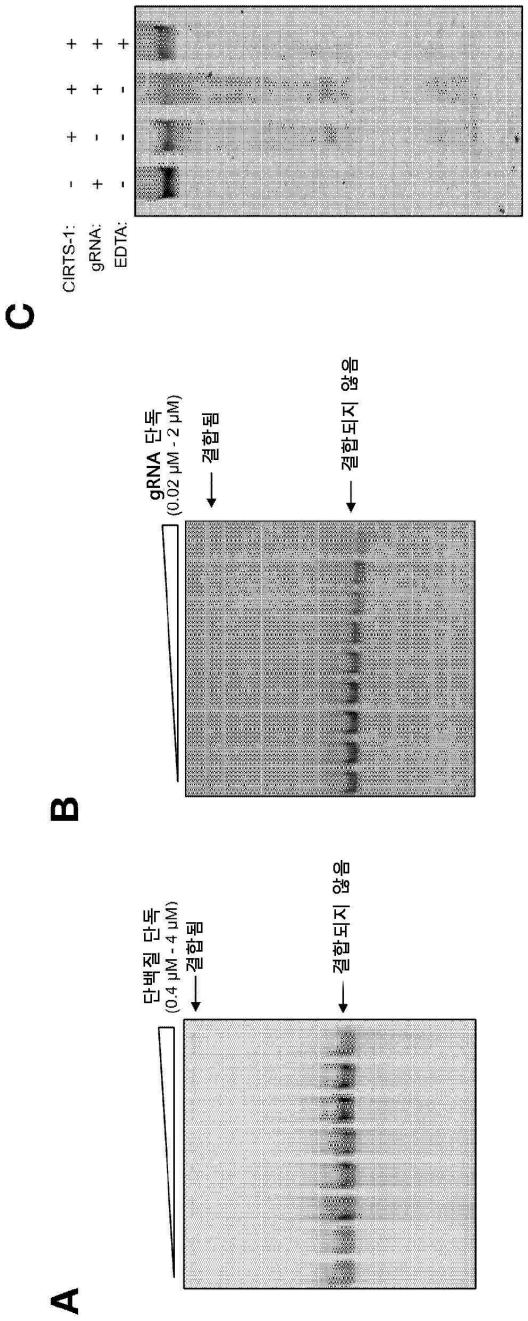
도면5


A
CIRTS-6
gRNA 벡터
hU6
PPIB
CIRTS-7
gRNA 벡터
hU6
SMARCA4
PPIB
SMARCA4
B
PPIB
gRNA:
SMARCA4
-
+
PPIB
-
+
0
18
35
녹다운 (%)
SMARCA4
SMARCA4
-
+
PPIB
-
+
0
15
30
녹다운 (%)

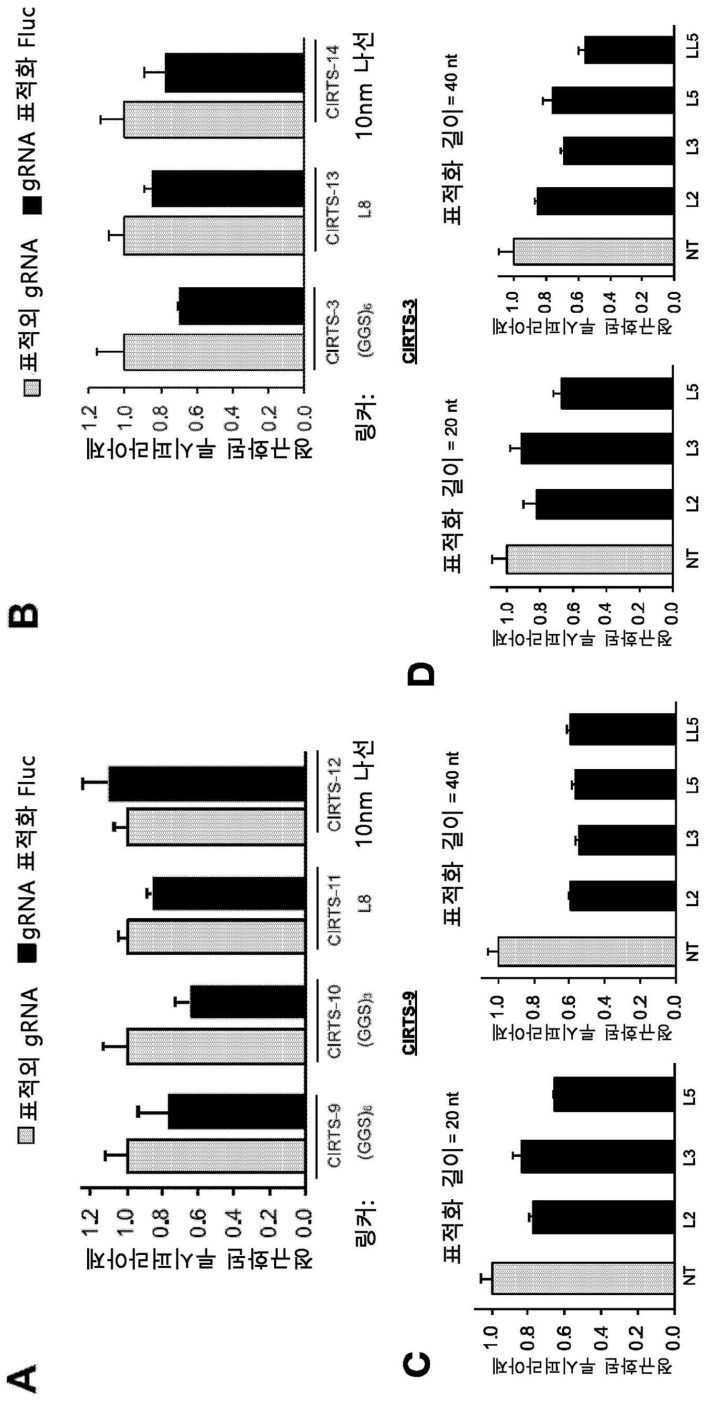
도면6


1368aa
SpCas9
1300aa
FnCas12a
529aa
Cas14a1
DNA
뉴클레아제
1389aa
LshCas13a
1090aa
PspCas13b
956aa
RfxCas13d
432aa
CIRTS-1
RNA
뉴클레아제
1375aa
PspCas13b-Y1
1111aa
Cas13b-Y2
615aa
CIRTS-2 (Y1)
341aa
CIRTS-3 (Y2)
RNA
후성전사체
조절인자

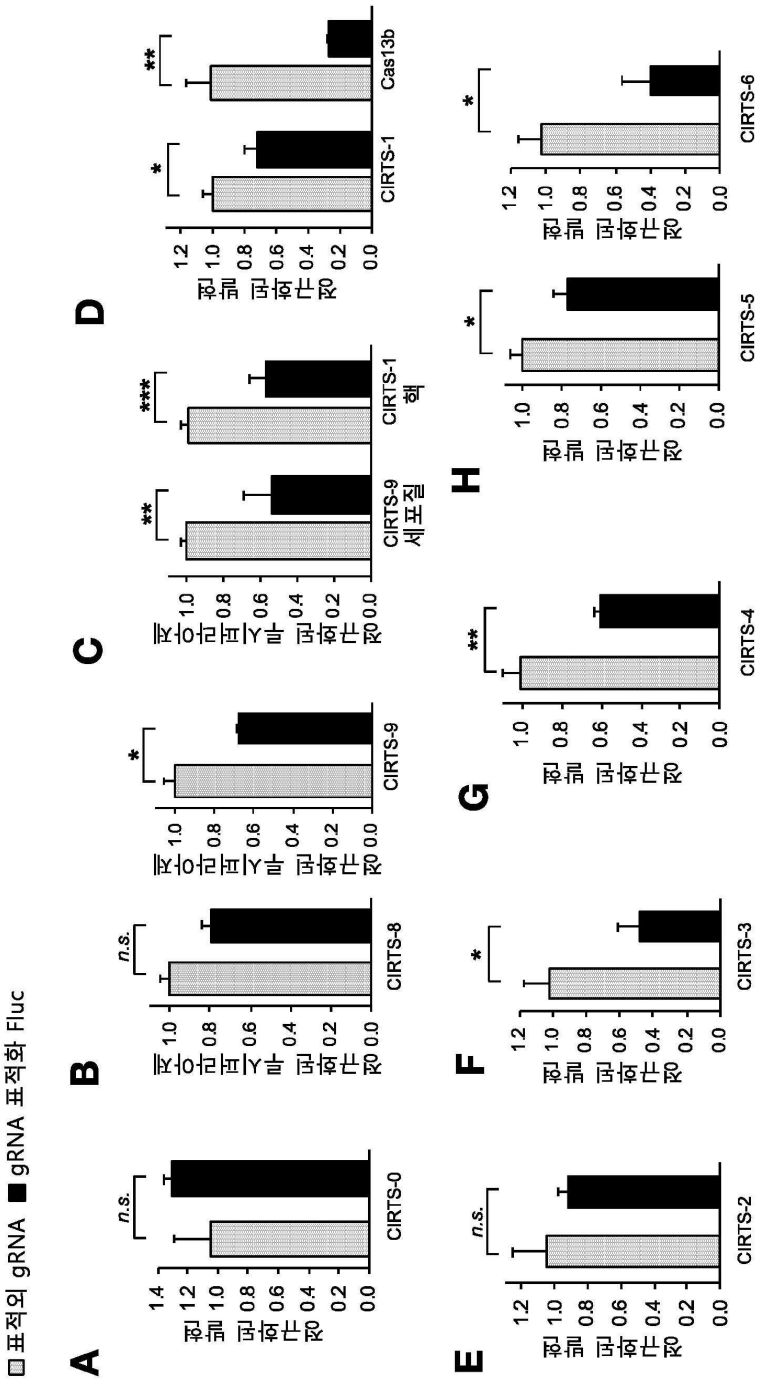
도면7


A
단백질 단독
(0.4 μM - 4 μM)
결합됨
결합되지 않음
B
gRNA 단독
(0.02 μM - 2 μM)
결합됨
결합되지 않음
C
CIRTS-1: - + + +
gRNA: + - + +
EDTA: - - - +

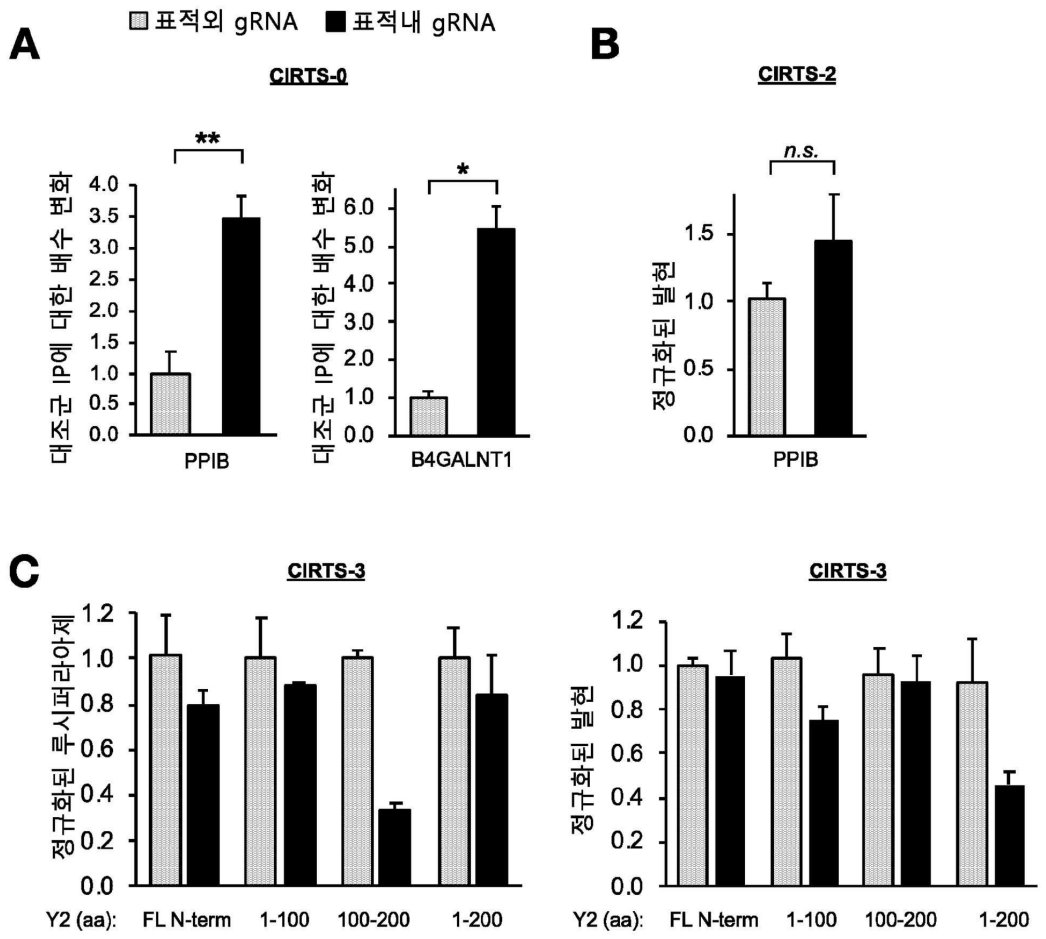
도면8


A
B
C
D
표적외 gRNA
gRNA 표적화 Fluc
CIRTS-9
CIRTS-10
CIRTS-11
CIRTS-12
CIRTS-3
CIRTS-13
CIRTS-14
(GGS)6
(GGS)3
L8
10nm 나선
링커:
표적화 길이 = 20 nt
표적화 길이 = 40 nt
NT
L2
L3
L5
LL5
CIRTS-9
CIRTS-3

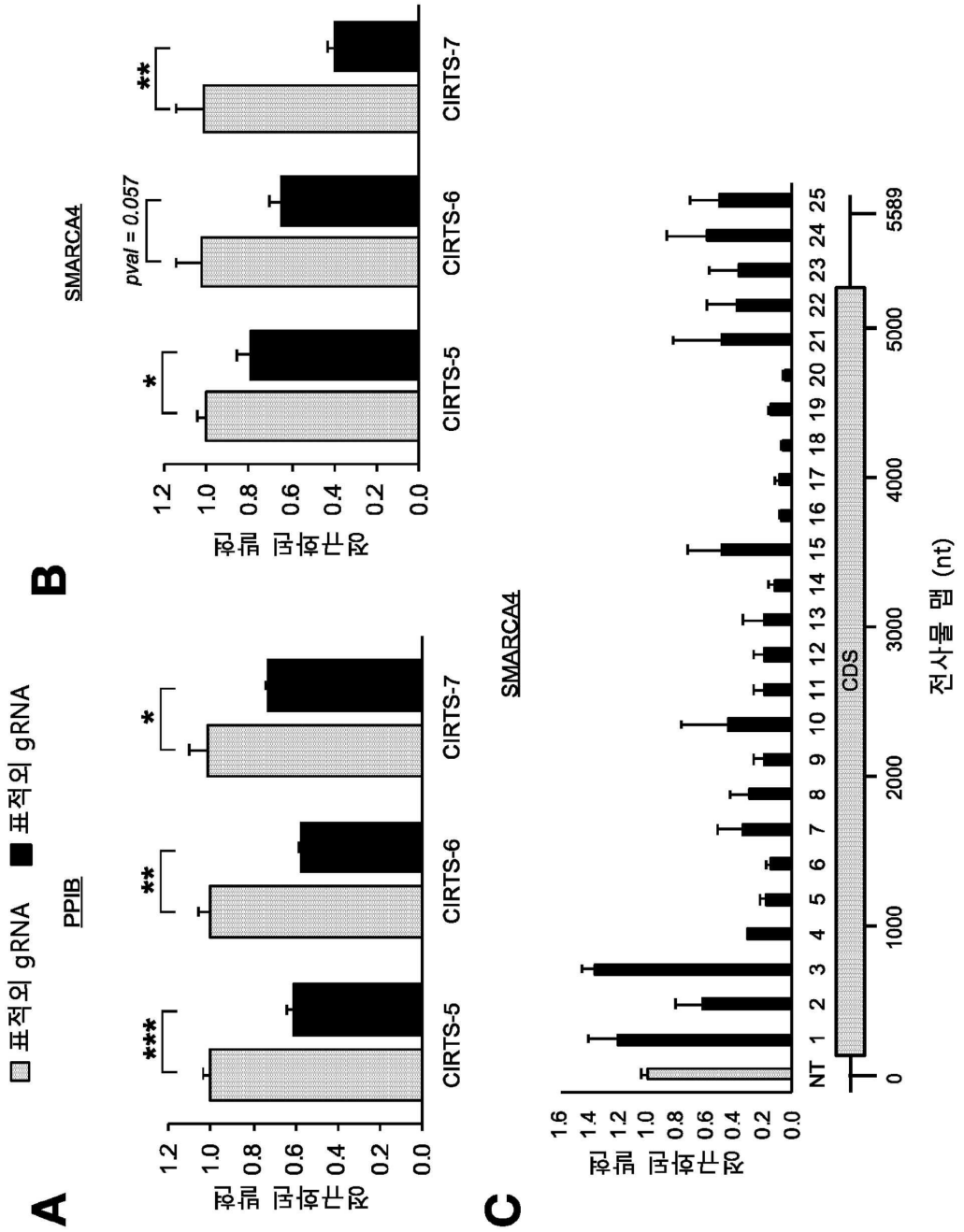
도면9


□ 표적외 gRNA ■ gRNA 표적화 Fluc
A
B
C
D
E
F
G
H
정규화된 발현
n.s.
*
**
***
CIRTS-0
CIRTS-8
CIRTS-9
CIRTS-9 세포질
CIRTS-1 핵
CIRTS-1
Cas13b
CIRTS-2
CIRTS-3
CIRTS-4
CIRTS-5
CIRTS-6

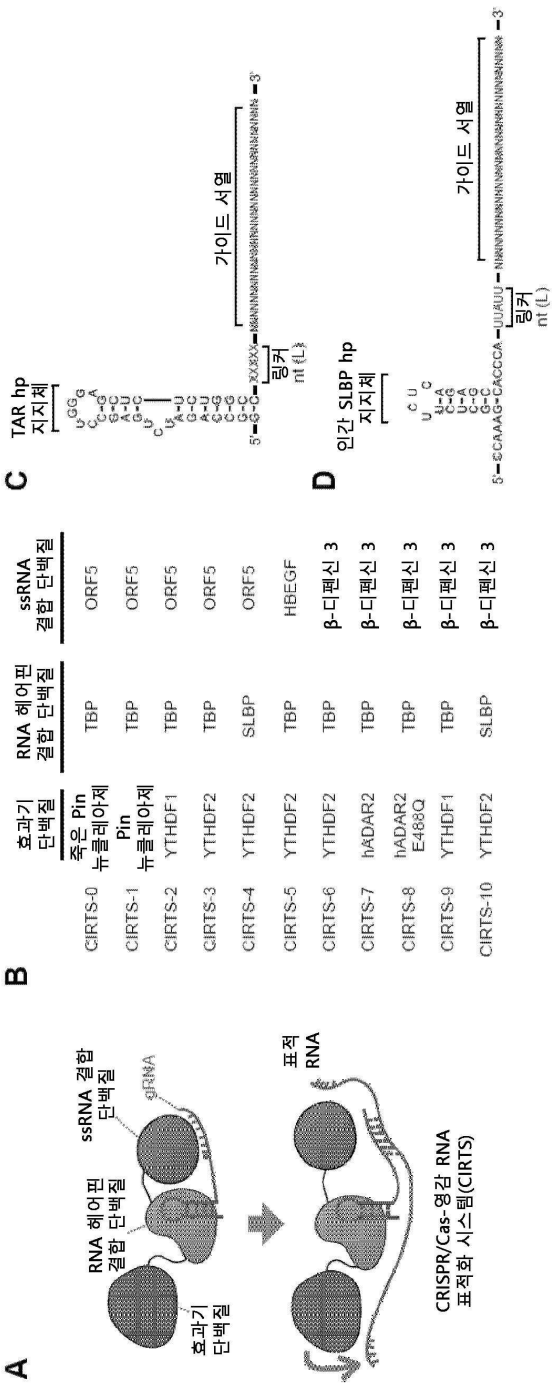
도면10


A
표적외 gRNA
표적내 gRNA
CIRTS-0
대조군 IP에 대한 배수 변화
PPIB
B4GALNT1
B
CIRTS-2
정규화된 발현
n.s.
PPIB
C
CIRTS-3
정규화된 루시퍼라아제
Y2 (aa):
FL N-term
1-100
100-200
1-200
CIRTS-3
정규화된 발현

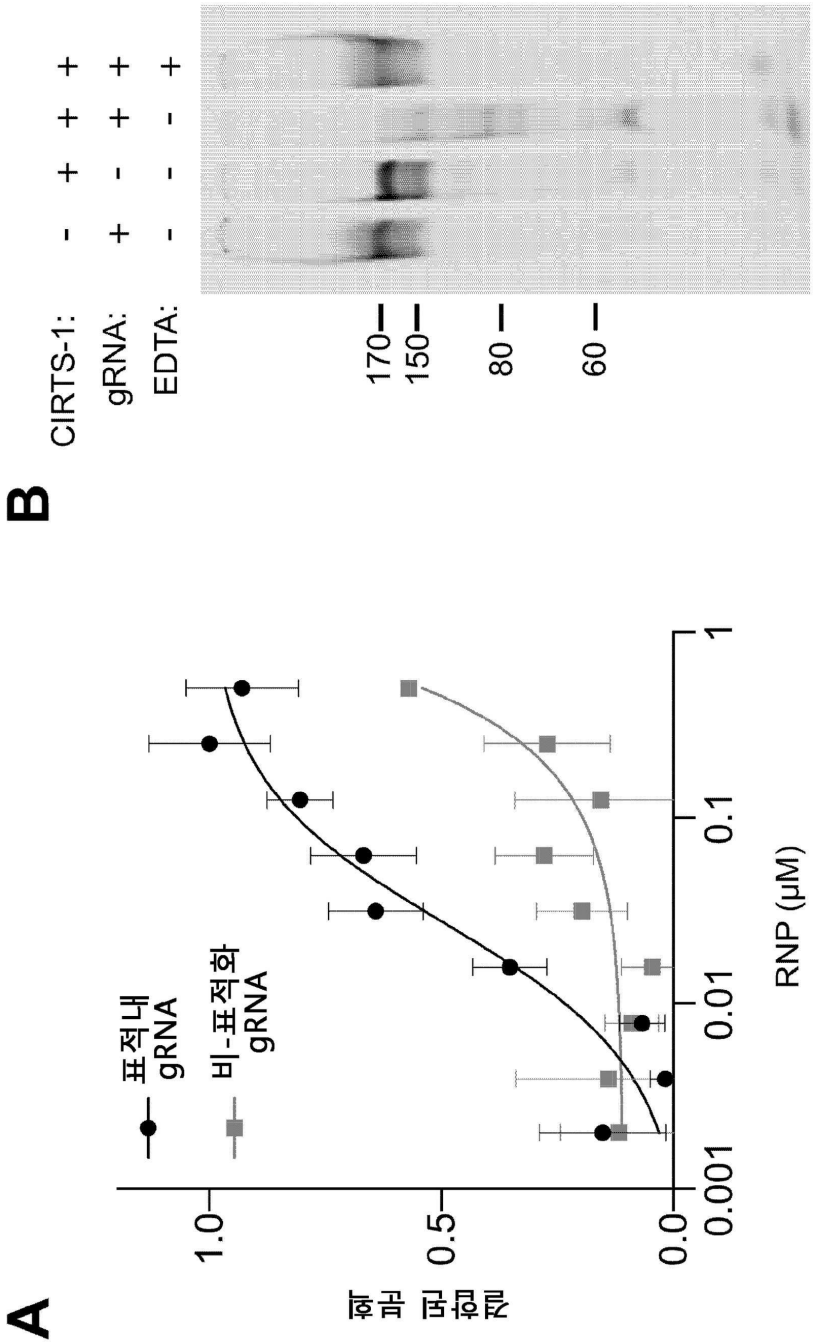
도면11


A
□ 표적외 gRNA ■ 표적외 gRNA
PPIB
정규화된 발현
1.2
1.0
0.8
0.6
0.4
0.2
0.0
***
**
*
CIRTS-5
CIRTS-6
CIRTS-7
B
SMARCA4
pval = 0.057
*
**
CIRTS-5
CIRTS-6
CIRTS-7
C
SMARCA4
정규화된 발현
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
NT 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25
CDS
0 1000 2000 3000 4000 5000 5589
전사물 맵 (nt)

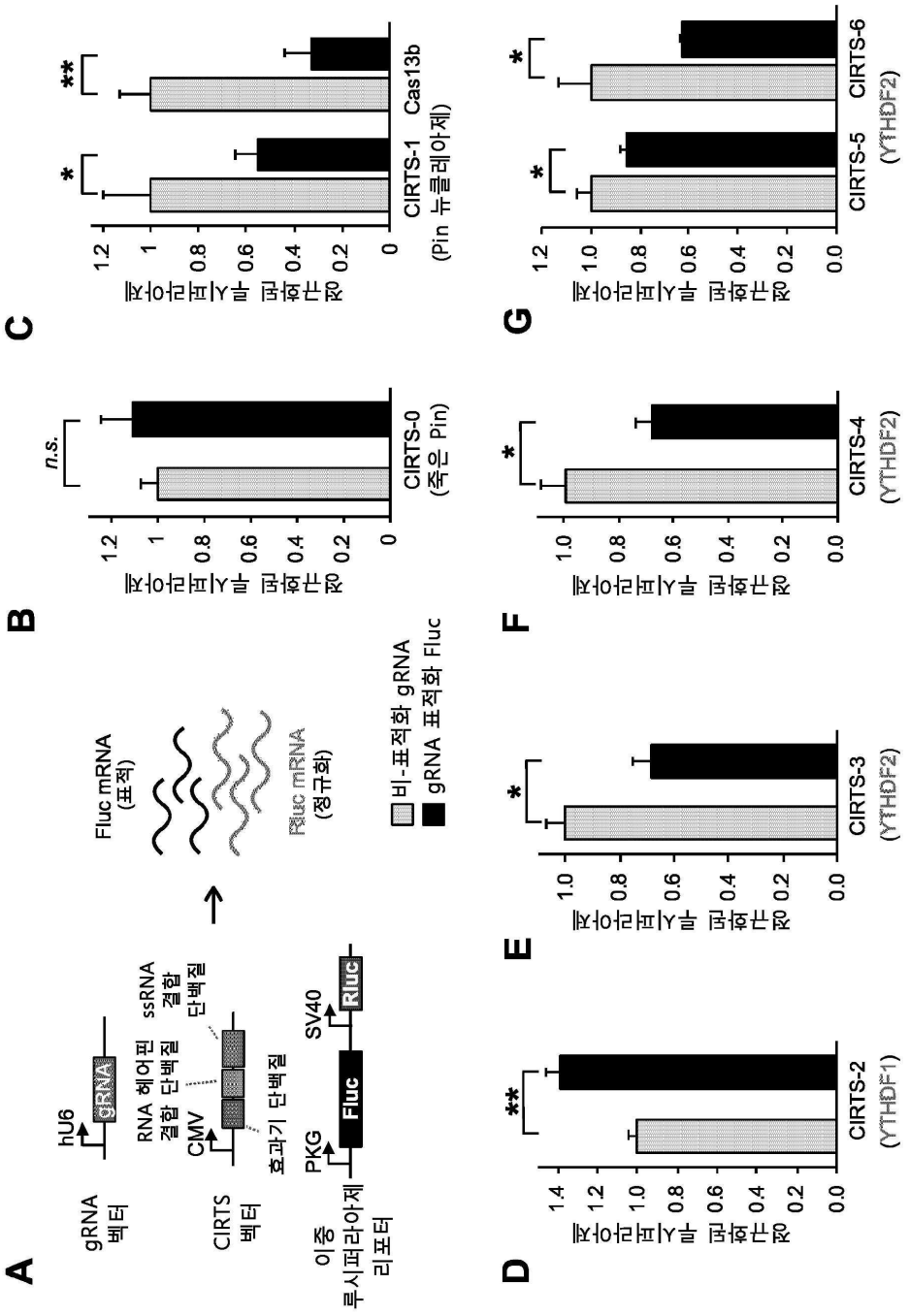
도면12


A
ssRNA 결합 단백질
RNA 헤어핀 결합 단백질
효과기 단백질
표적 RNA
CRISPR/Cas-영감 RNA 표적화 시스템(CIRTS)
B
효과기 단백질
RNA 헤어핀 결합 단백질
ssRNA 결합 단백질
CIRTS-0 죽은 Pin 뉴클레아제 TBP ORF5
CIRTS-1 Pin 뉴클레아제 TBP ORF5
CIRTS-2 YTHDF1 TBP ORF5
CIRTS-3 YTHDF2 TBP ORF5
CIRTS-4 YTHDF2 SLBP ORF5
CIRTS-5 YTHDF2 TBP HBEGF
CIRTS-6 YTHDF2 TBP β-디펜신 3
CIRTS-7 hADAR2 TBP β-디펜신 3
CIRTS-8 hADAR2 E488Q TBP β-디펜신 3
CIRTS-9 YTHDF1 TBP β-디펜신 3
CIRTS-10 YTHDF2 SLBP β-디펜신 3
C
TAR hp 지지체
가이드 서열
링커
nt (L)
D
인간 SLBP hp 지지체
가이드 서열
링커
nt (L)

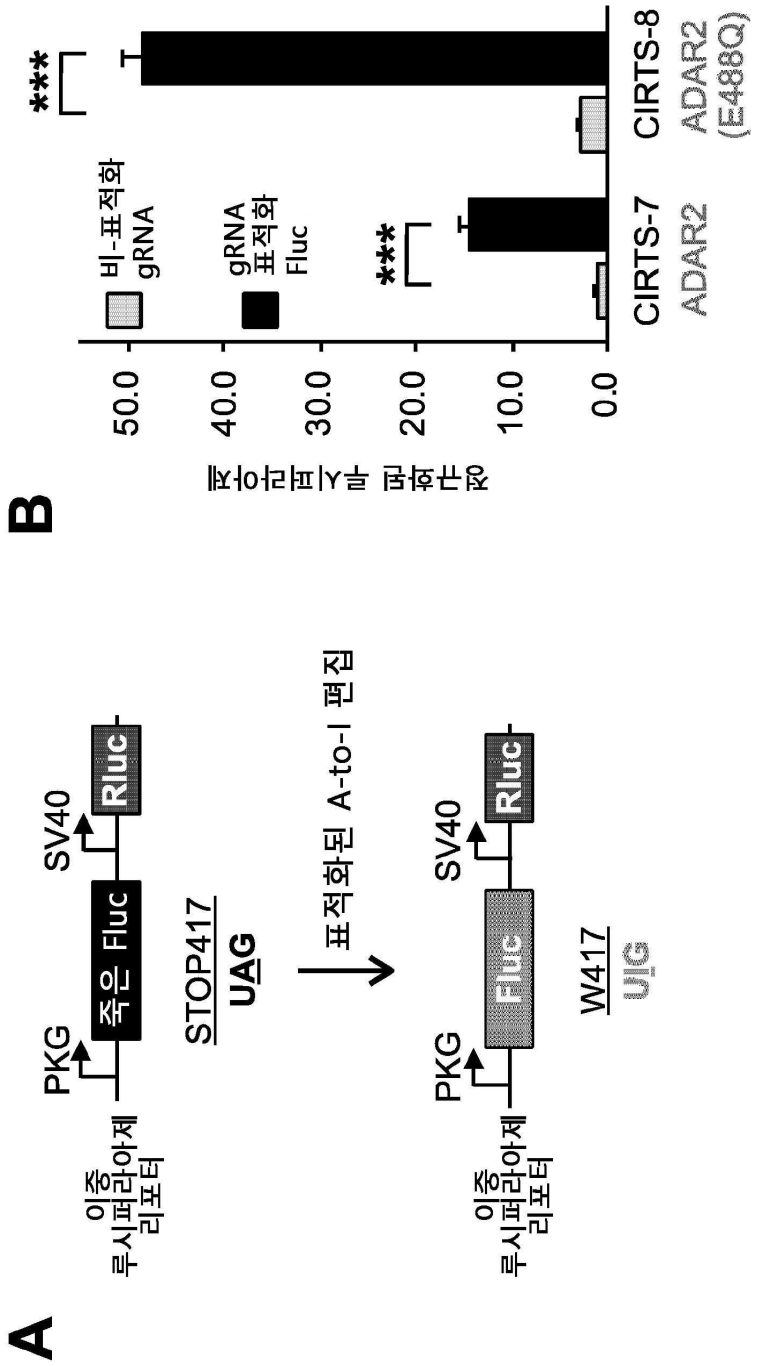
도면13


A
표적내 gRNA
비-표적화 gRNA
결합된 분획
1.0
0.5
0.0
0.001
0.01
0.1
1
RNP (μM)
B
CIRTS-1:
gRNA:
EDTA:
170
150
80
60

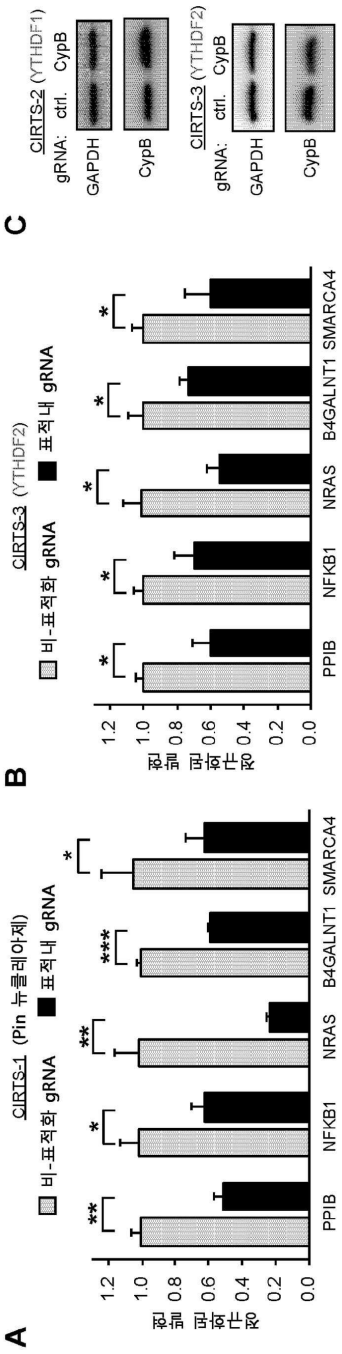
도면14


A
gRNA 벡터
hU6
gRNA
CIRTS 벡터
RNA 헤어핀 결합 단백질
ssRNA 결합 단백질
CMV
효과기 단백질
이중 루시퍼라아제 리포터
PKG
Fluc
SV40
Rluc
Fluc mRNA (표적)
Rluc mRNA (정규화)
비-표적화 gRNA
gRNA 표적화 Fluc
B
n.s.
정규화된 루시퍼라아제
CIRTS-0 (죽은 Pin)
C
*
**
CIRTS-1 (Pin 뉴클레아제)
Cas13b
D
CIRTS-2 (YTHDF1)
E
CIRTS-3 (YTHDF2)
F
CIRTS-4 (YTHDF2)
G
CIRTS-5 (YTHDF2)
CIRTS-6 (YTHDF2)

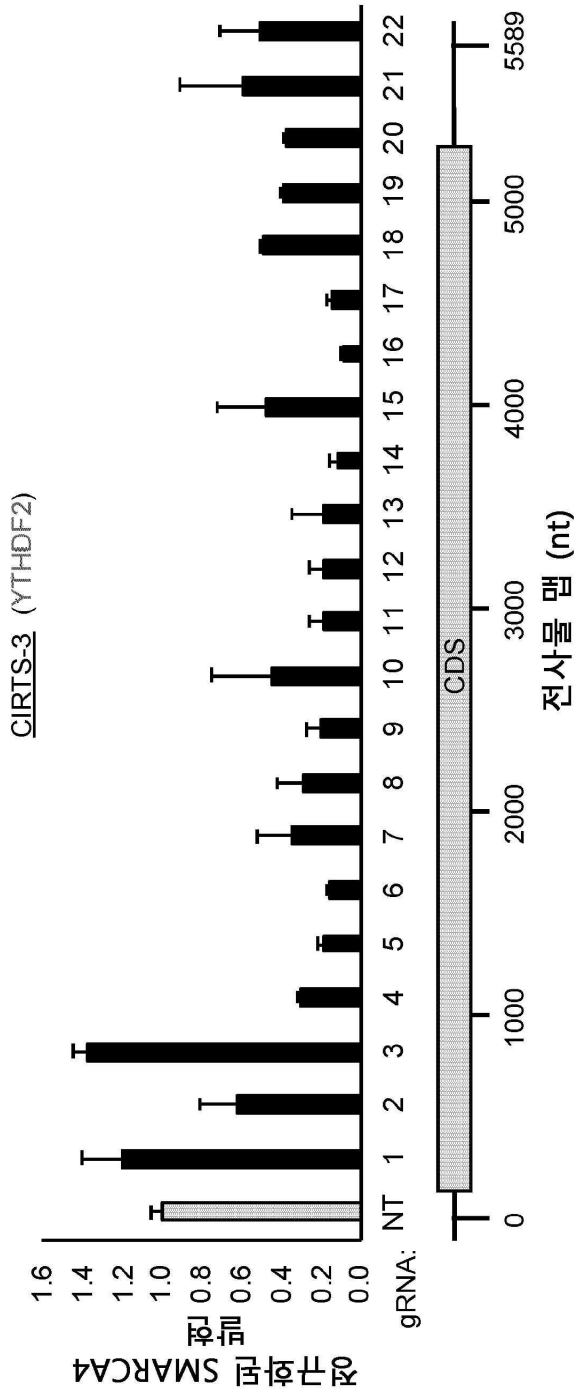
도면15


A
이중
루시퍼라아제
리포터
PKG
죽은 Fluc
SV40
Rluc
STOP417
UAG
표적화된 A-to-I 편집
이중
루시퍼라아제
리포터
PKG
Fluc
SV40
Rluc
W417
UIG
B
정규화된 루시퍼라아제
50.0
40.0
30.0
20.0
10.0
0.0
비-표적화
gRNA
gRNA
표적화
Fluc
***
***
CIRTS-7
ADAR2
CIRTS-8
ADAR2
(E488Q)

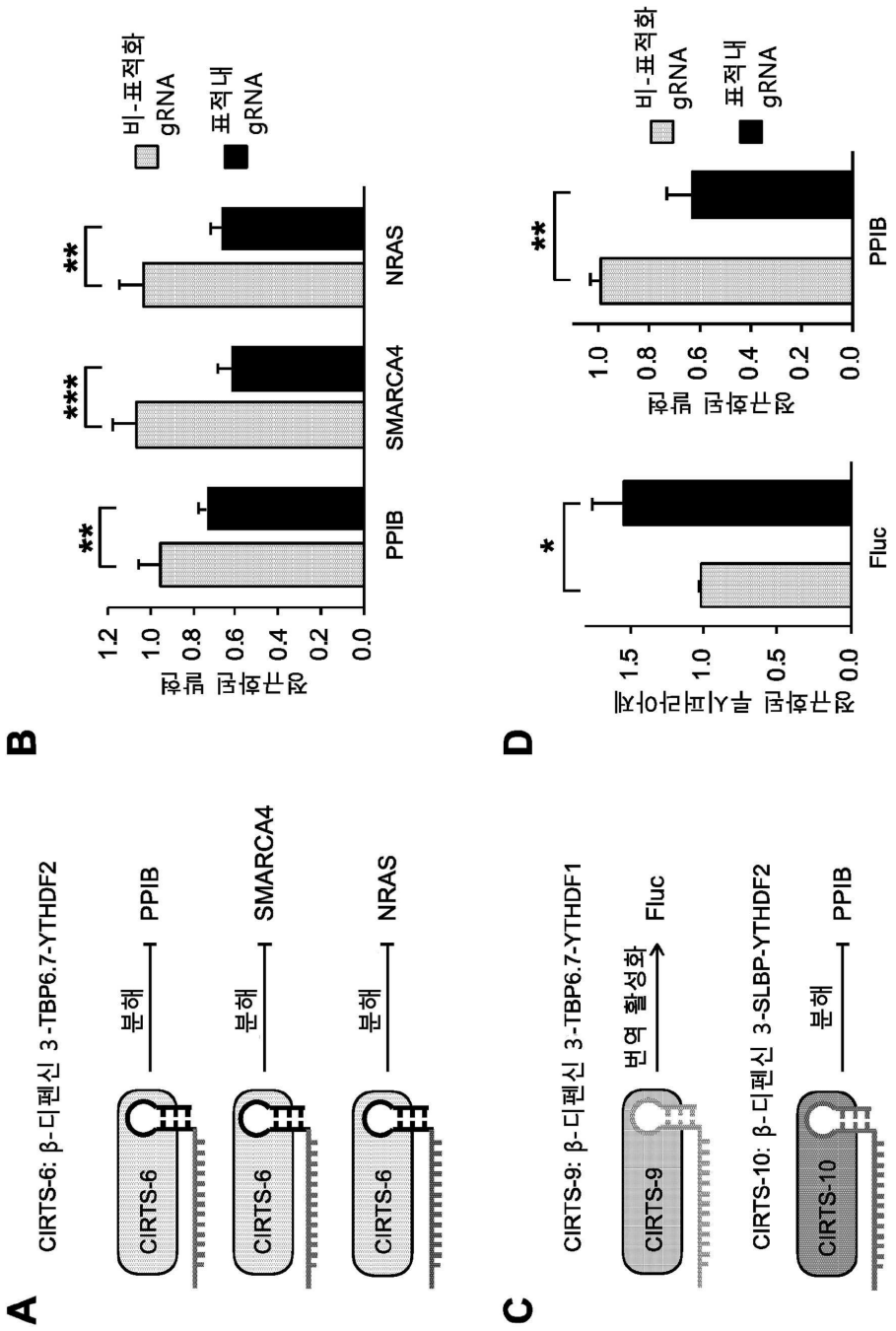
도면16


A
CIRTS-1 (Pin 뉴클레아제)
비-표적화 gRNA
표적내 gRNA
정규화된 발현
1.2
1.0
0.8
0.6
0.4
0.2
0.0
**
*
**
***
*
PPIB
NFKB1
NRAS
B4GALNT1
SMARCA4
B
CIRTS-3 (YTHDF2)
비-표적화 gRNA
표적내 gRNA
정규화된 발현
1.2
1.0
0.8
0.6
0.4
0.2
0.0
*
*
*
*
*
PPIB
NFKB1
NRAS
B4GALNT1
SMARCA4
C
CIRTS-2 (YTHDF1)
gRNA: ctrl. CypB
GAPDH
CypB
CIRTS-3 (YTHDF2)
gRNA: ctrl. CypB
GAPDH
CypB

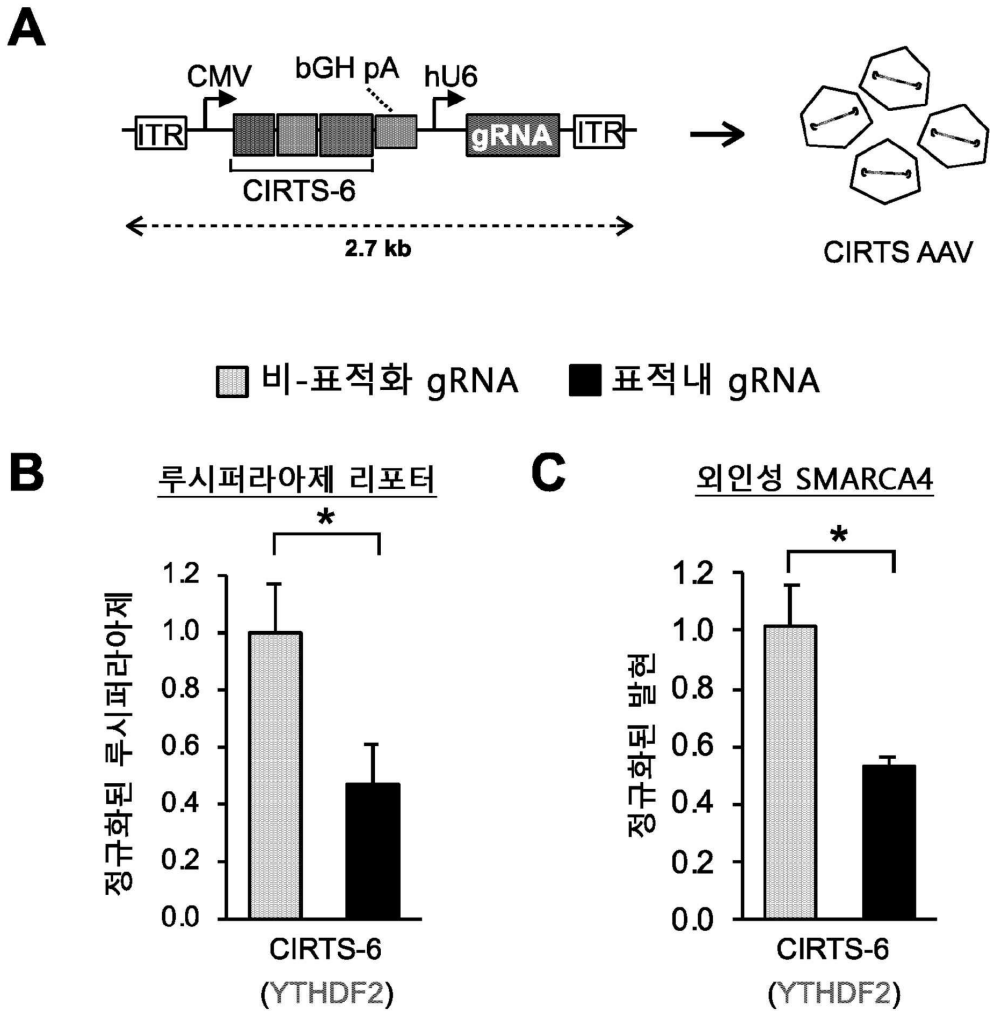
도면17


CIRTS-3 (YTHDF2)
정규화된 SMARCA4 발현
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
gRNA:
NT
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
CDS
0
1000
2000
3000
4000
5000
5589
전사물 맵 (nt)

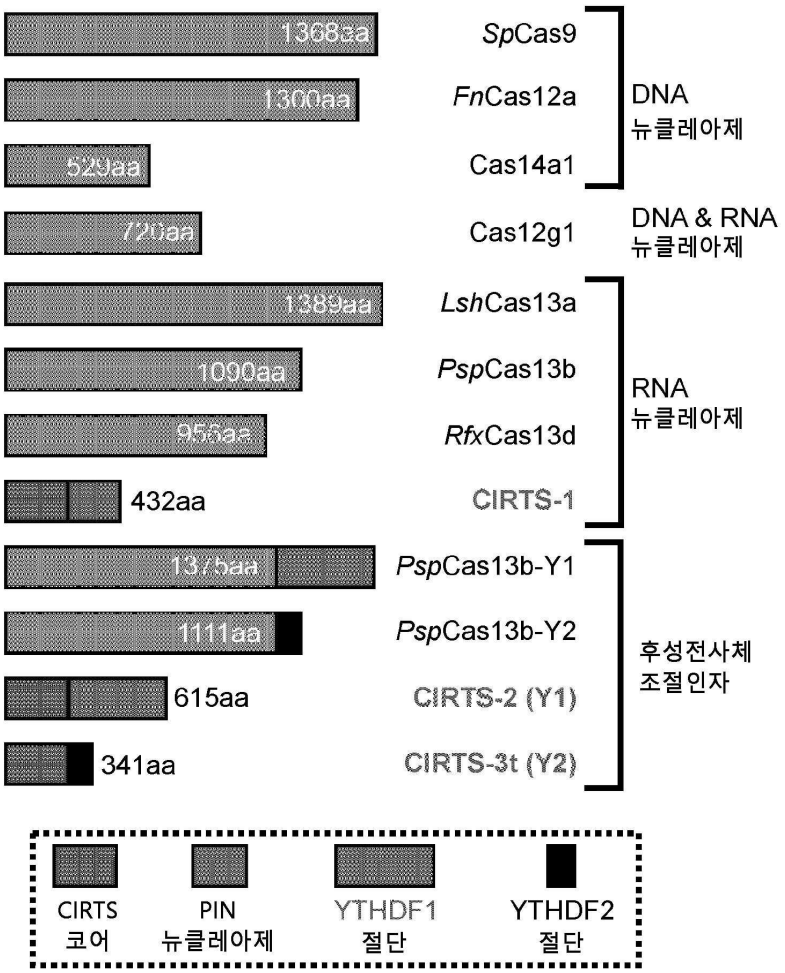
도면18


A
CIRTS-6: β-디펜신 3-TBP6.7-YTHDF2
CIRTS-6
분해
PPIB
CIRTS-6
분해
SMARCA4
CIRTS-6
분해
NRAS
B
비-표적화 gRNA
표적내 gRNA
정규화된 발현
1.2
1.0
0.8
0.6
0.4
0.2
0.0
**
***
**
PPIB
SMARCA4
NRAS
C
CIRTS-9: β-디펜신 3-TBP6.7-YTHDF1
번역 활성화
Fluc
CIRTS-9
CIRTS-10: β-디펜신 3-SLBP-YTHDF2
분해
PPIB
CIRTS-10
D
비-표적화 gRNA
표적내 gRNA
정규화된 루시퍼라아제
1.5
1.0
0.5
0.0
*
Fluc
정규화된 발현
1.0
0.8
0.6
0.4
0.2
0.0
**
PPIB

도면19


A
CMV
bGH pA
hU6
ITR
gRNA
ITR
CIRTS-6
2.7 kb
CIRTS AAV
비-표적화 gRNA
표적내 gRNA
B
루시퍼라아제 리포터
정규화된 루시퍼라아제
1.2
1.0
0.8
0.6
0.4
0.2
0.0
*
CIRTS-6
(YTHDF2)
C
외인성 SMARCA4
정규화된 발현
1.2
1.0
0.8
0.6
0.4
0.2
0.0
*
CIRTS-6
(YTHDF2)

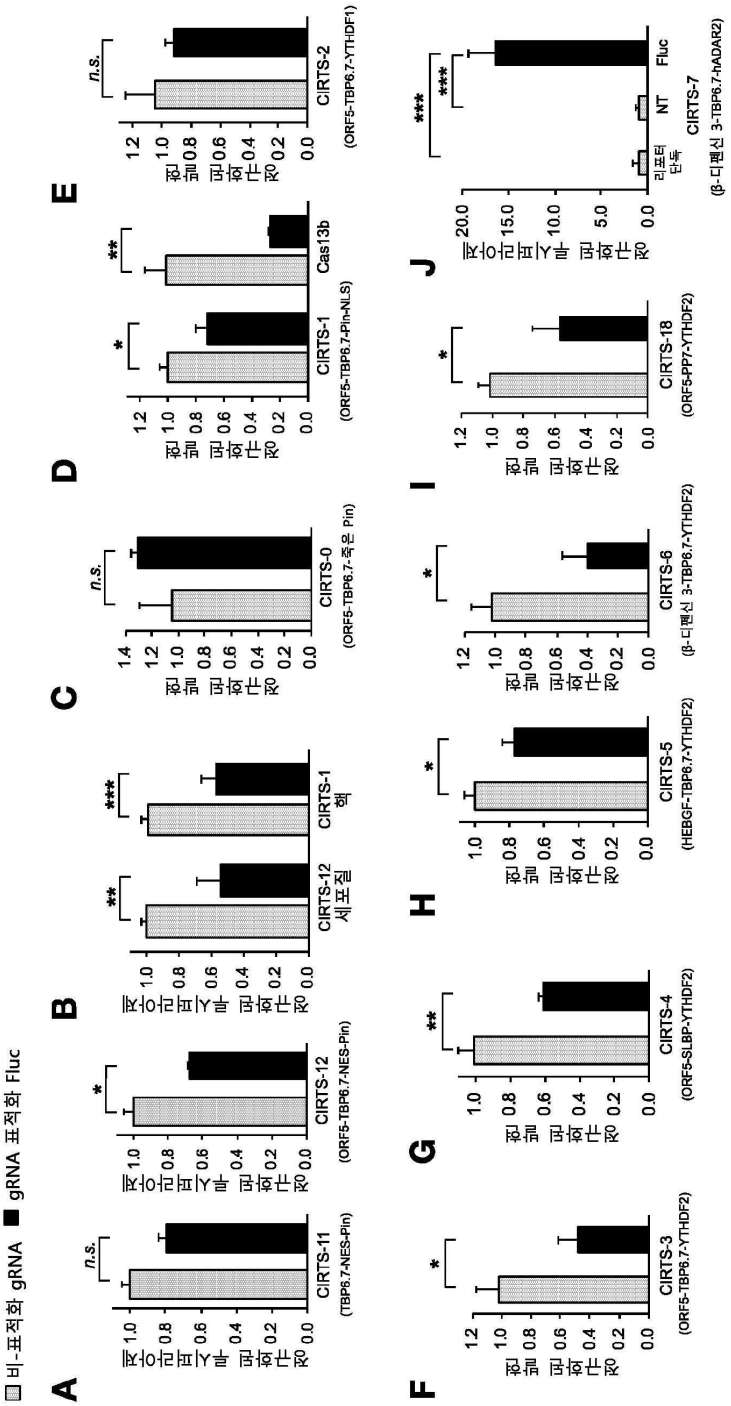
도면20



도면21

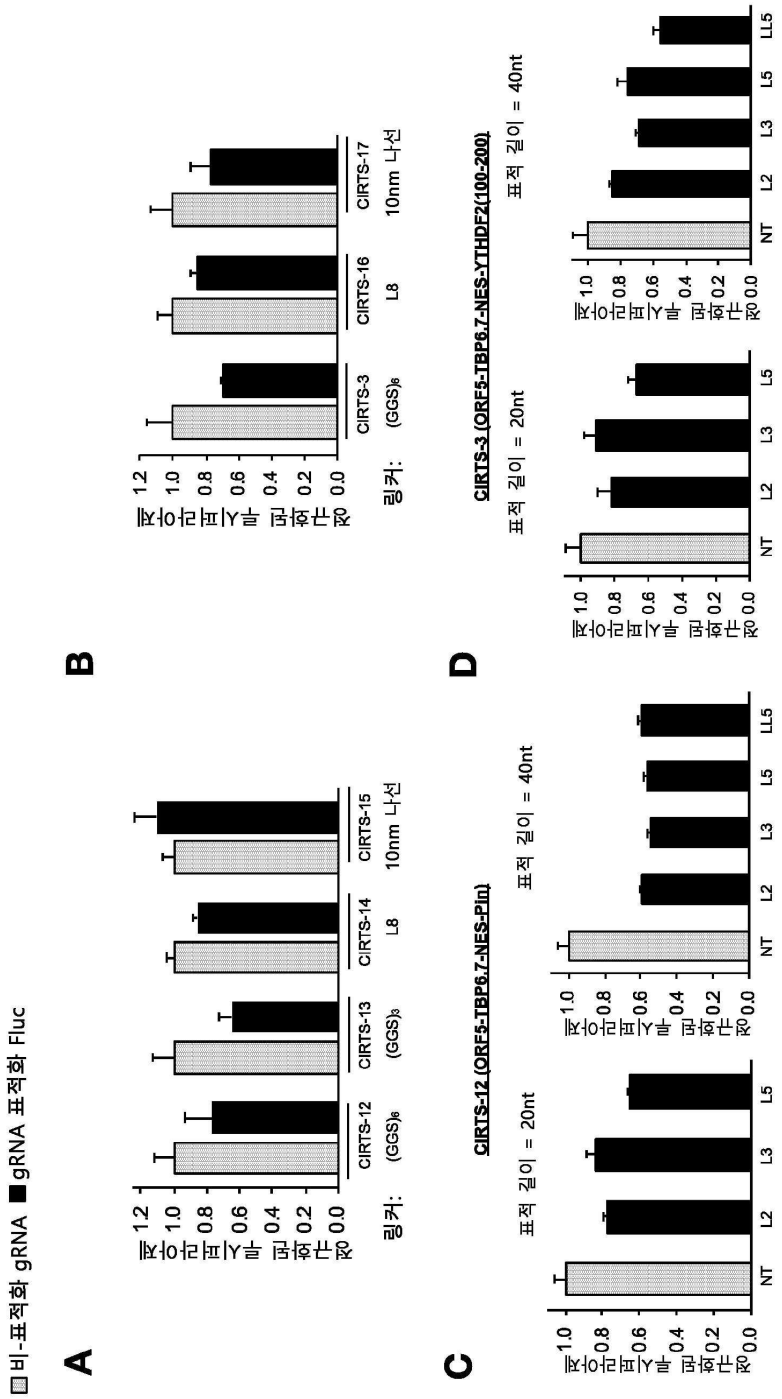
	효과기 단백질	링커	RNA 헤어핀 결합 단백질	ssRNA 결합 단백질
CIRTS-11	세포질 Pin 뉴클레아제	$(GGS)_6$	TBP	N/A
CIRTS-12	세포질 Pin 뉴클레아제	$(GGS)_6$	TBP	ORF5
CIRTS-13	Pin 뉴클레아제	$(GGS)_3$	TBP	ORF5
CIRTS-14	Pin 뉴클레아제	L8	TBP	ORF5
CIRTS-15	Pin 뉴클레아제	나선	TBP	ORF5
CIRTS-16	YTHDF2	L8	TBP	ORF5
CIRTS-17	YTHDF2	나선	TBP	ORF5
CIRTS-18	YTHDF2	$(GGS)_6$	PP7	ORF5

도면22

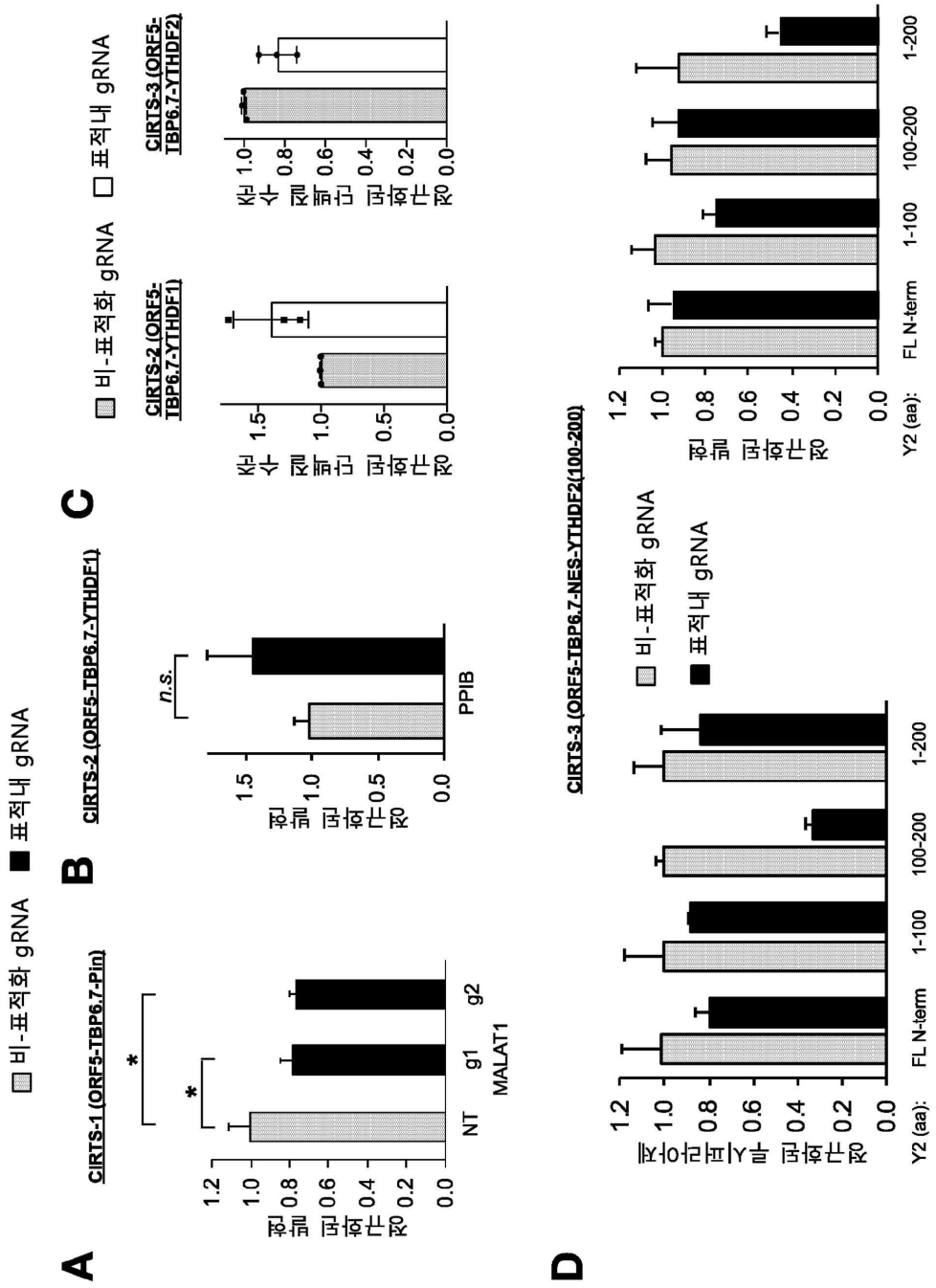


비-표적화 gRNA
gRNA 표적화 Fluc
A
B
C
D
E
F
G
H
I
J
CIRTS-11
(TBP6.7-NES-Pin)
CIRTS-12
(ORF5-TBP6.7-NES-Pin)
CIRTS-12
세포질
CIRTS-1
핵
CIRTS-0
(ORF5-TBP6.7-죽은 Pin)
CIRTS-1
(ORF5-TBP6.7-Pin-NLS)
Cas13b
CIRTS-2
(ORF5-TBP6.7-YTHDF1)
CIRTS-3
(ORF5-TBP6.7-YTHDF2)
CIRTS-4
(ORF5-SLBP-YTHDF2)
CIRTS-5
(HEBGF-TBP6.7-YTHDF2)
CIRTS-6
(β-디펜신 3-TBP6.7-YTHDF2)
CIRTS-18
(ORF5-PP7-YTHDF2)
리포터 단독
NT
Fluc
CIRTS-7
(β-디펜신 3-TBP6.7-hADAR2)

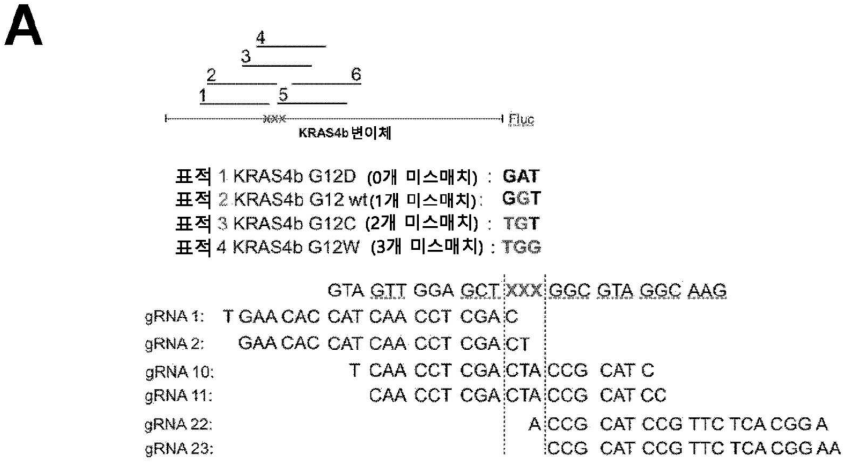
도면23


□비-표적화 gRNA ■gRNA 표적화 Fluc
A
정규화된 루시퍼라아제
1.2
1.0
0.8
0.6
0.4
0.2
0.0
CIRTS-12
CIRTS-13
CIRTS-14
CIRTS-15
링커: (GGS)6 (GGS)3 L8 10nm 나선
B
CIRTS-3
CIRTS-16
CIRTS-17
링커: (GGS)6 L8 10nm 나선
C
CIRTS-12 (ORF5-TBP6.7-NES-Pin)
표적 길이 = 20nt
표적 길이 = 40nt
NT
L2
L3
L5
LL5
D
CIRTS-3 (ORF5-TBP6.7-NES-YTHDF2(100-200))
표적 길이 = 20nt
표적 길이 = 40nt

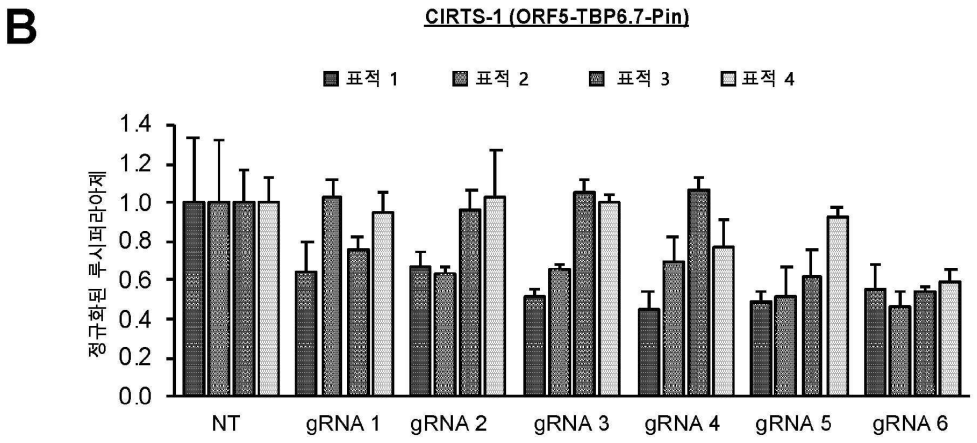
도면24


A
비-표적화 gRNA
표적내 gRNA
CIRTS-1 (ORF5-TBP6.7-Pin)
*
*
정규화된 발현
NT
g1
g2
MALAT1
B
CIRTS-2 (ORF5-TBP6.7-YTHDF1)
n.s.
PPIB
C
CIRTS-2 (ORF5-TBP6.7-YTHDF1)
CIRTS-3 (ORF5-TBP6.7-YTHDF2)
D
CIRTS-3 (ORF5-TBP6.7-NES-YTHDF2(100-200)
Y2 (aa):
FL N-term
1-100
100-200
1-200

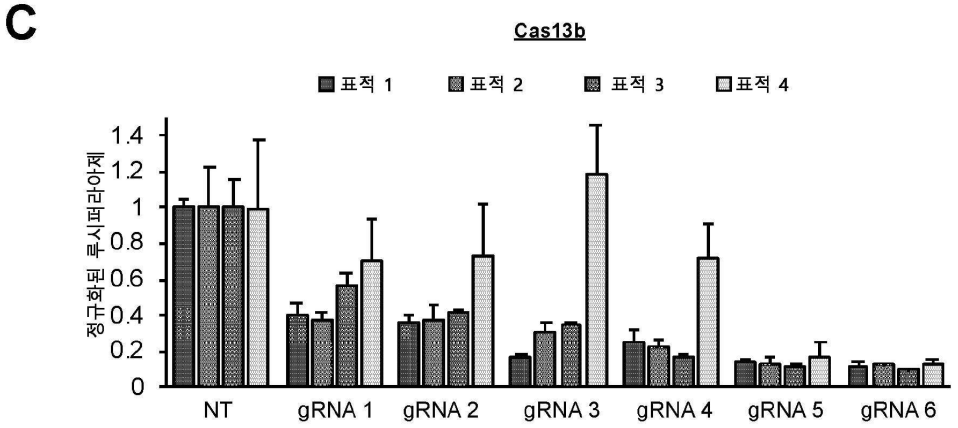
도면25a



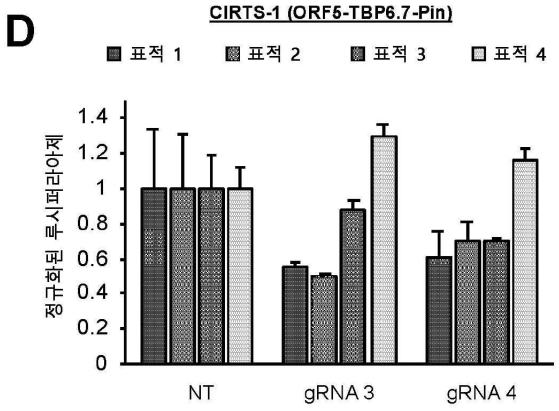
도면25b



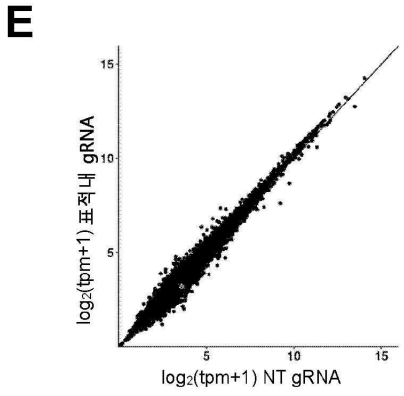
도면25c



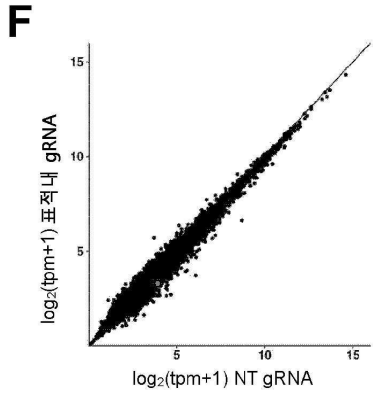
도면25d



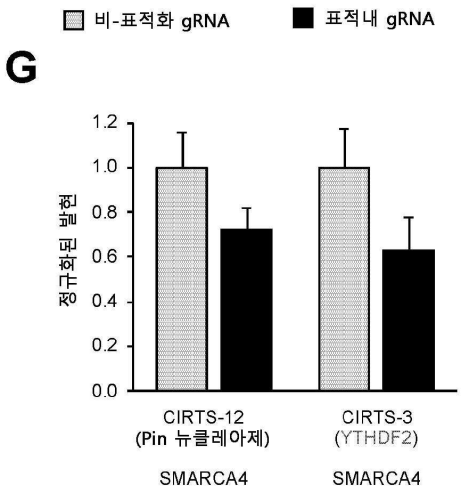
도면25e



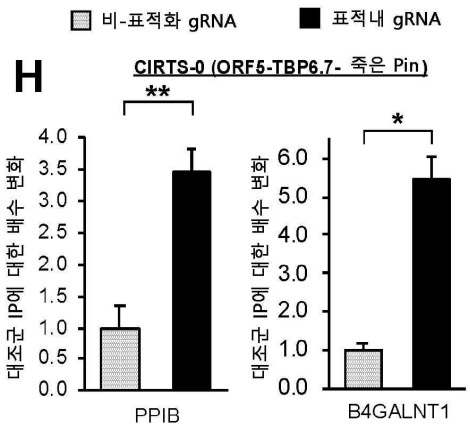
도면25f



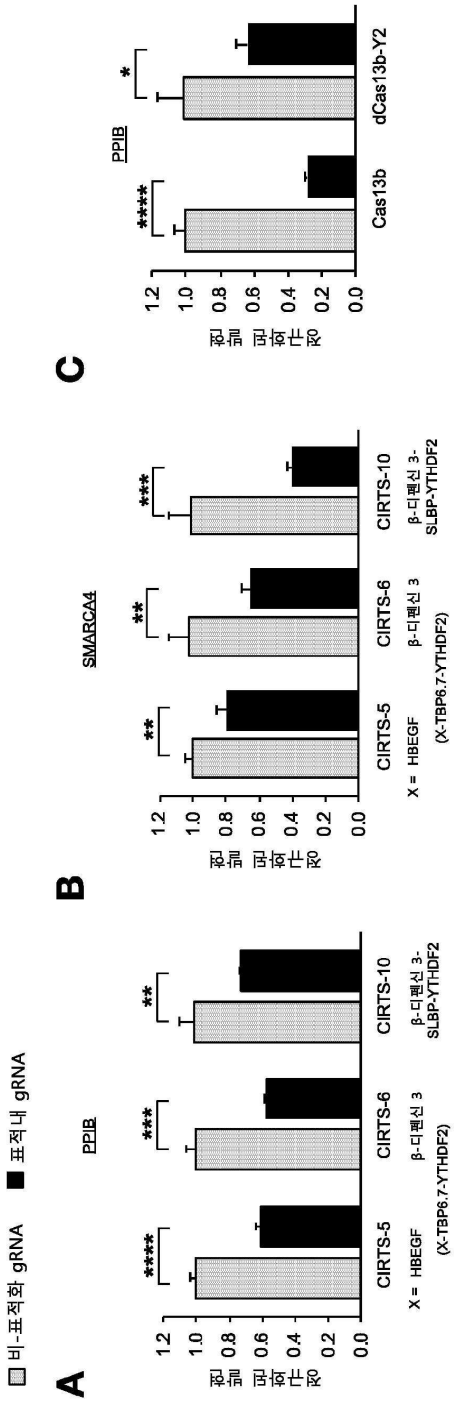
도면25g


비-표적화 gRNA
표적내 gRNA
G
정규화된 발현
1.2
1.0
0.8
0.6
0.4
0.2
0.0
CIRTS-12
(Pin 뉴클레아제)
SMARCA4
CIRTS-3
(YTHDF2)
SMARCA4

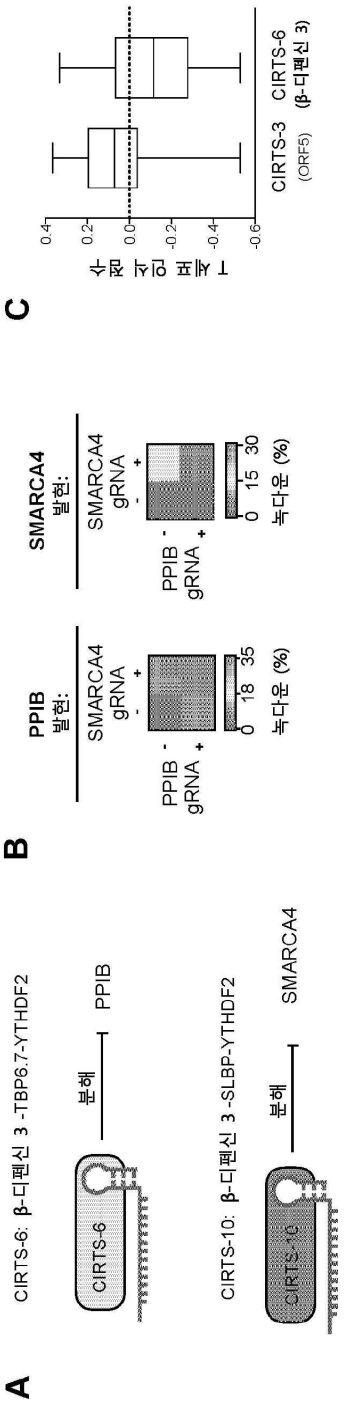
도면25h


비-표적화 gRNA
표적내 gRNA
H
CIRTS-0 (ORF5-TBP6.7- 죽은 Pin)
**
*
대조군 IP에 대한 배수 변화
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
PPIB
6.0
5.0
4.0
3.0
2.0
1.0
0.0
B4GALNT1

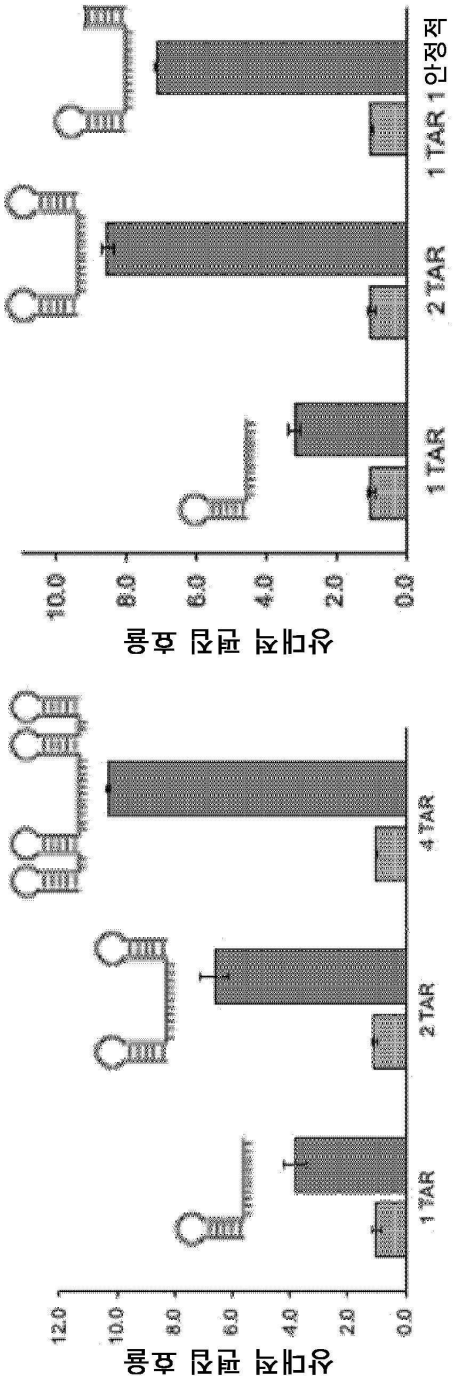
도면26


비-표적화 gRNA
표적내 gRNA
A
PPIB
정규화된 발현
1.2
1.0
0.8
0.6
0.4
0.2
0.0
****
***
**
CIRTS-5
CIRTS-6
CIRTS-10
X = HBEGF
β-디펜신 3
(X-TBP6.7-YTHDF2)
β-디펜신 3-
SLBP-YTHDF2
B
SMARCA4
정규화된 발현
1.2
1.0
0.8
0.6
0.4
0.2
0.0
**
**
***
CIRTS-5
CIRTS-6
CIRTS-10
X = HBEGF
β-디펜신 3
(X-TBP6.7-YTHDF2)
β-디펜신 3-
SLBP-YTHDF2
C
PPIB
정규화된 발현
1.2
1.0
0.8
0.6
0.4
0.2
0.0
****
*
Cas13b
dCas13b-Y2

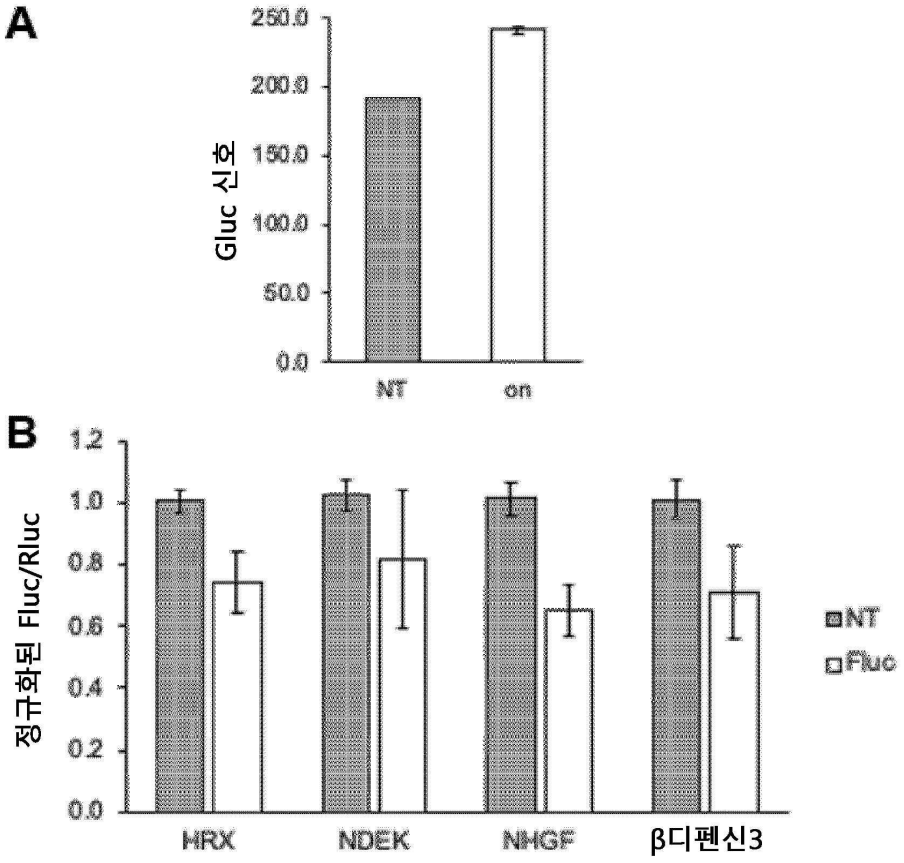
도면27


A
CIRTS-6: β-디펜신 3 -TBP6.7-YTHDF2
CIRTS-6
분해
PPIB
CIRTS-10: β-디펜신 3 -SLBP-YTHDF2
CIRTS-10
분해
SMARCA4
B
PPIB
발현:
SMARCA4 gRNA
-
+
PPIB gRNA
-
+
0 18 35
녹다운 (%)
SMARCA4
발현:
SMARCA4 gRNA
-
+
PPIB gRNA
-
+
0 15 30
녹다운 (%)
C
T 세포 인식 점수
0.4
0.2
0.0
-0.2
-0.4
-0.6
CIRTS-3 (ORF5)
CIRTS-6 (β-디펜신 3)

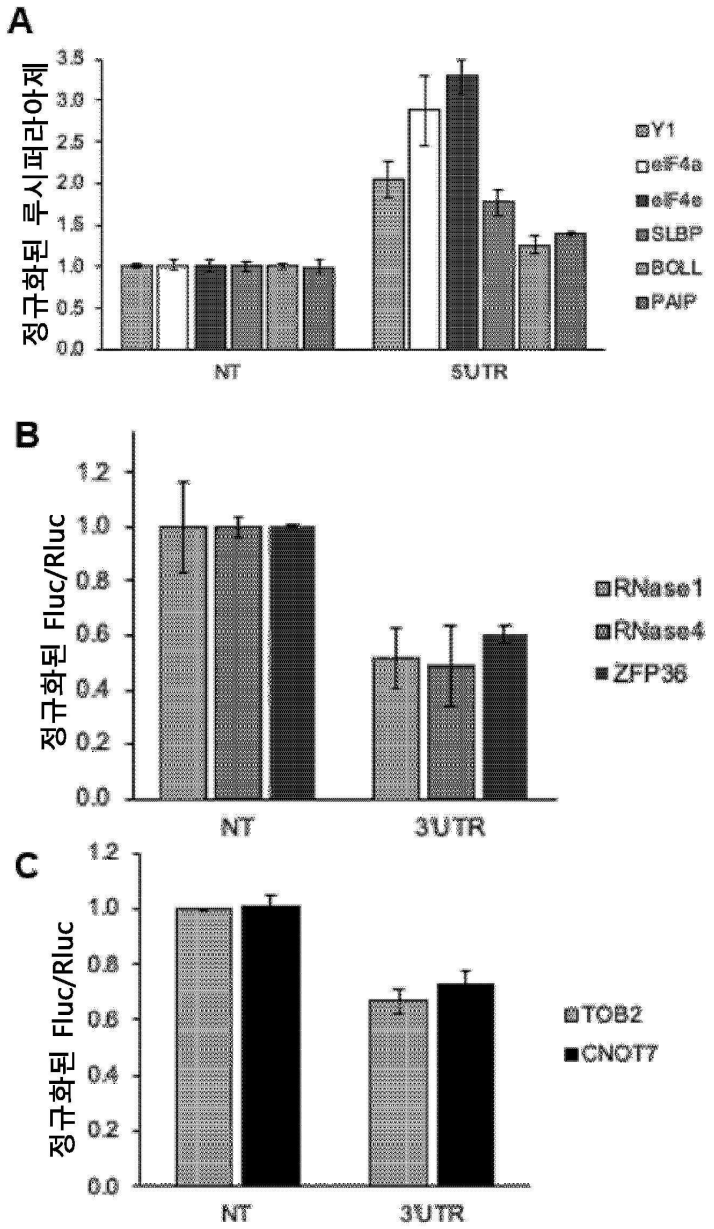
도면28


상대적 편집 효율
12.0
10.0
8.0
6.0
4.0
2.0
0.0
1 TAR
2 TAR
4 TAR
상대적 편집 효율
10.0
8.0
6.0
4.0
2.0
0.0
1 TAR
2 TAR
1 TAR 1 안정적

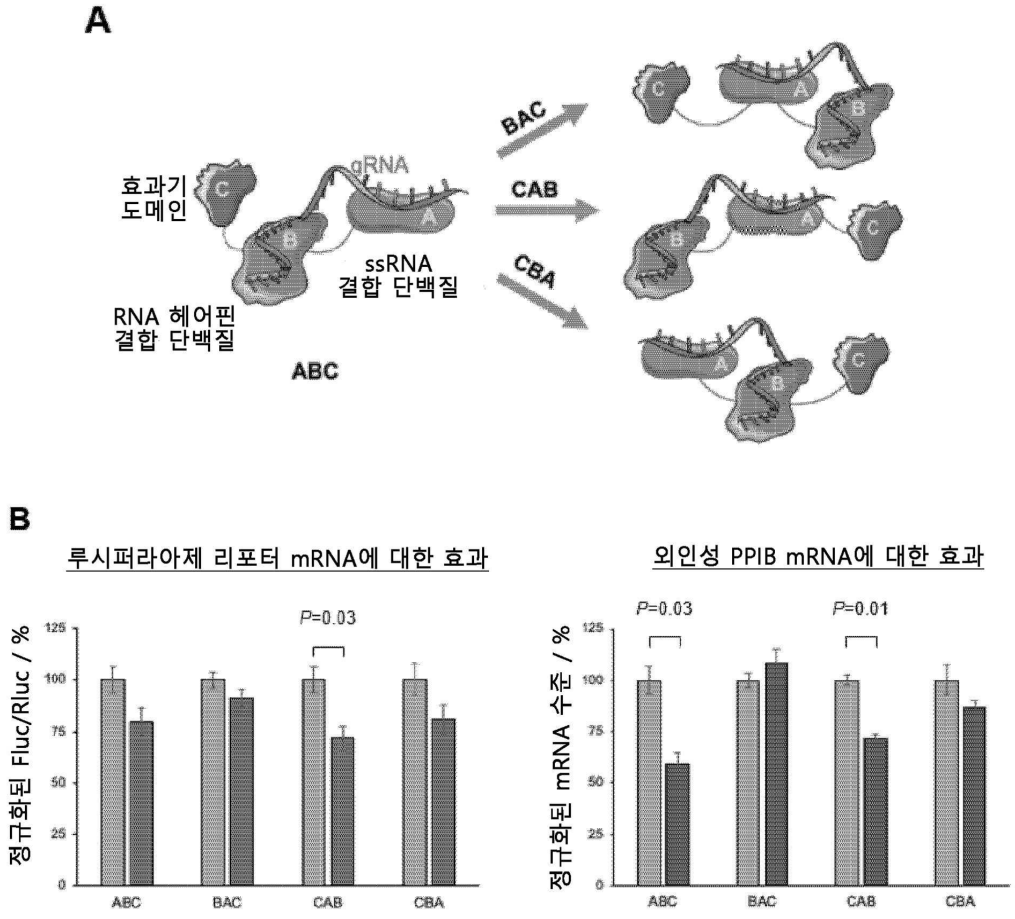
도면29


A
250.0
200.0
150.0
100.0
50.0
0.0
Gluc 신호
NT
on
B
1.2
1.0
0.8
0.6
0.4
0.2
0.0
정규화된 Fluc/Rluc
NT
Fluc
HRX
NDEK
NHGF
β디펜신3

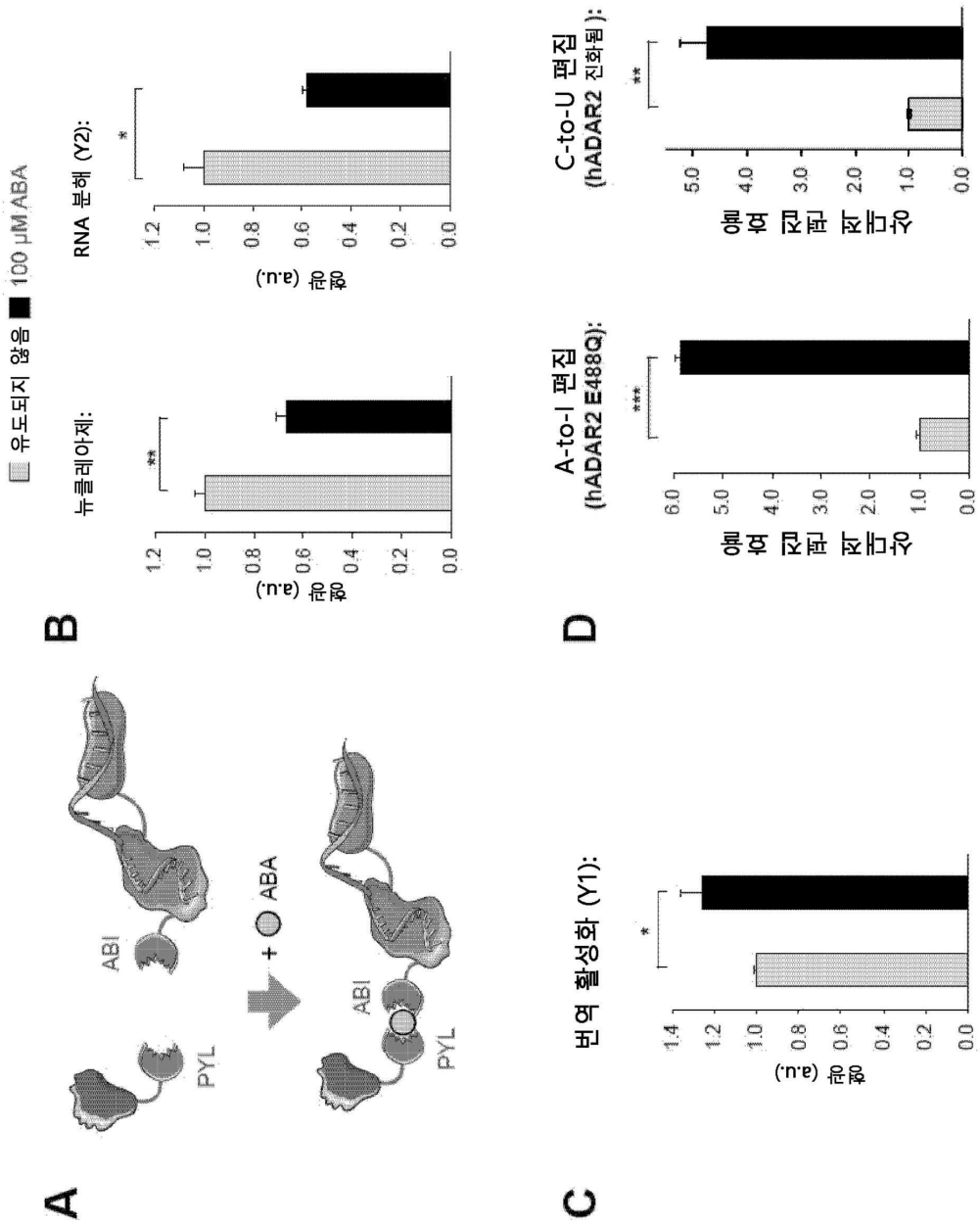
도면30


A
정규화된 루시퍼라아제
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
NT
5'UTR
Y1
eIF4a
eIF4e
SLBP
BCLL
PAIP
B
정규화된 Fluc/Rluc
1.2
1.0
0.8
0.6
0.4
0.2
0.0
NT
3'UTR
RNase1
RNase4
ZFP36
C
정규화된 Fluc/Rluc
1.2
1.0
0.8
0.6
0.4
0.2
0.0
NT
3'UTR
TOB2
CNOT7

도면31


A
효과기
도메인
gRNA
ssRNA
결합 단백질
RNA 헤어핀
결합 단백질
ABC
BAC
CAB
CBA
B
루시퍼라아제 리포터 mRNA에 대한 효과
정규화된 Fluc/Rluc / %
P=0.03
ABC
BAC
CAB
CBA
외인성 PPIB mRNA에 대한 효과
정규화된 mRNA 수준 / %
P=0.03
P=0.01

도면32



서 열 목 록

SEQUENCE LISTING

<110> THE UNIVERSITY OF CHICAGO

<120> SYSTEMS AND METHODS FOR MODULATING RNA

<130> ARCD-P0678WO-1001105787

<140> PCT/US2020/012169

<141> 2020-01-03

<150> 62/929,339

<151> 2019-11-01

<150> 62/903,080

<151> 2019-09-20

<150> 62/831,342

<151> 2019-04-09

<150> 62/788,571

<151> 2019-01-04

<160> 148

<170> PatentIn version 3.5

<210> 1

<211> 31

<212> RNA

<213> Homo sapiens

<400> 1

ggccagaucu gagccuggga gcucucuggc c 31

<210> 2

<211> 26

<212> RNA

<213> Homo sapiens

<400> 2

ccaaaggcuc uucucagagc caccca 26

<210> 3

<211> 369

<212> PRT

<213> Artificial Sequence

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

<400> 3

Met Lys Ile Glu Glu Gly Lys Leu Val Ile Trp Ile Asn Gly Asp Lys
1 5 10 15

Gly Tyr Asn Gly Leu Ala Glu Val Gly Lys Lys Phe Glu Lys Asp Thr
20 25 30

Gly Ile Lys Val Thr Val Glu His Pro Asp Lys Leu Glu Glu Lys Phe
35 40 45

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

Ala Val Arg Tyr Asn Gly Lys Leu Ile Ala Tyr Pro Ile Ala Val Glu
100 105 110
Ala Leu Ser Leu Ile Tyr Asn Lys Asp Leu Leu Pro Asn Pro Pro Lys
115 120 125
Thr Trp Glu Glu Ile Pro Ala Leu Asp Lys Glu Leu Lys Ala Lys Gly
130 135 140
Lys Ser Ala Leu Met Phe Asn Leu Gln Glu Pro Tyr Phe Thr Trp Pro
145 150 155 160

Leu Ile Ala Ala Asp Gly Gly Tyr Ala Phe Lys Tyr Glu Asn Gly Lys
165 170 175
Tyr Asp Ile Lys Asp Val Gly Val Asp Asn Ala Gly Ala Lys Ala Gly
180 185 190
Leu Thr Phe Leu Val Asp Leu Ile Lys Asn Lys His Met Asn Ala Asp
195 200 205
Thr Asp Tyr Ser Ile Ala Glu Ala Ala Phe Asn Lys Gly Glu Thr Ala
210 215 220

Met Thr Ile Asn Gly Pro Trp Ala Trp Ser Asn Ile Asp Thr Ser Lys
225 230 235 240
Val Asn Tyr Gly Val Thr Val Leu Pro Thr Phe Lys Gly Gln Pro Ser
245 250 255
Lys Pro Phe Val Gly Val Leu Ser Ala Gly Ile Asn Ala Ala Ser Pro
260 265 270
Asn Lys Glu Leu Ala Lys Glu Phe Leu Glu Asn Tyr Leu Leu Thr Asp
275 280 285

Glu Gly Leu Glu Ala Val Asn Lys Asp Lys Pro Leu Gly Ala Val Ala

290 295 300

Leu Lys Ser Tyr Glu Glu Glu Leu Ala Lys Asp Pro Arg Ile Ala Ala

305 310 315 320

Thr Met Glu Asn Ala Gln Lys Gly Glu Ile Met Pro Asn Ile Pro Gln

325 330 335

Met Ser Ala Phe Trp Tyr Ala Val Arg Thr Ala Val Ile Asn Ala Ala

340 345 350

Ser Gly Arg Gln Thr Val Asp Glu Ala Leu Lys Asp Ala Gln Thr Asn

355 360 365

Ser

<210> 4

<211> 40

<212> PRT

<213> Artificial Sequence

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

polypeptide

<400> 4

Ile Thr Ser Leu Tyr Lys Lys Ala Gly Ser Glu Thr Val Arg Phe Gln

1 5 10 15

Ser His His His His His His Ser Ser Gly Val Asp Leu Gly Thr Glu

20 25 30

Asn Leu Tyr Phe Gln Ser Asn Ala

35 40

<210> 5

<211> 90

<212> PRT

<213> Artificial Sequence

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

polypeptide

<400> 5

Met Asp Asp Pro Ser Phe Leu Thr Gly Arg Ser Thr Tyr Ala Lys Arg

1 5 10 15

Arg Arg Ala Arg Arg Met Asn Val Cys Lys Cys Gly Ala Ile Leu His

20 25 30

Asn Asn Lys Asp Cys Arg Ser Ser Thr Ile Ser Gly His Lys Leu Asp

35 40 45

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

50 55 60

Pro Val Tyr Arg Thr Trp Val Lys Trp Val Glu Thr Glu Tyr His Ile

65 70 75 80

Asn Ile Leu Glu Thr Ser Asp Asp Glu Glu

85 90

<210> 6

<211> 18

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 6

Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly

1 5 10 15

Gly Ser

<210> 7

<211> 102

<212> PRT

<213> Artificial Sequence

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

polypeptide

<400> 7

Met Ala Val Pro Glu Thr Arg Pro Asn His Thr Ile Tyr Ile Asn Asn

1 5 10 15

Leu Asn Ser Lys Ile Lys Lys Asp Glu Leu Lys Lys Ser Leu Tyr Ala

20 25 30

Ile Phe Ser Gln Phe Gly Gln Ile Leu Asp Ile Leu Val Pro Arg Gln

35 40 45

Arg Thr Pro Arg Gly Gln Ala Phe Val Ile Phe Lys Glu Val Ser Ser

50 55 60

Ala Thr Asn Ala Leu Arg Ser Met Gln Gly Phe Pro Phe Tyr Asp Lys

65 70 75 80

Pro Met Arg Ile Gln Tyr Ala Lys Thr Asp Lys Arg Ile Pro Ala Lys

85 90 95

Met Lys Gly Thr Phe Val

100

<210> 8

<211> 11

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 8

Leu Gln Leu Pro Pro Leu Glu Arg Leu Thr Leu

1 5 10

<210> 9

<211> 181

<212> PRT

<213> Artificial Sequence

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

polypeptide

<400> 9

Met Glu Leu Glu Ile Arg Pro Leu Phe Leu Val Pro Asp Thr Asn Gly

1 5 10 15

Phe Ile Asp His Leu Ala Ser Leu Ala Arg Leu Leu Glu Ser Arg Lys

20 25 30

Tyr Ile Leu Val Val Pro Leu Ile Val Ile Asn Glu Leu Asp Gly Leu

35 40 45

Ala Lys Gly Gln Glu Thr Asp His Arg Ala Gly Gly Tyr Ala Arg Val

50 55 60

Val Gln Glu Lys Ala Arg Lys Ser Ile Glu Phe Leu Glu Gln Arg Phe

65 70 75 80

Glu Ser Arg Asp Ser Cys Leu Arg Ala Leu Thr Ser Arg Gly Asn Glu

85 90 95

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

100 105 110

Asn Asn Asp Asp Leu Ile Leu Ser Cys Cys Leu His Tyr Cys Lys Asp

115 120 125

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

130 135 140

Leu Arg Glu Val Val Leu Leu Thr Asp Asp Arg Asn Leu Arg Val Lys

145 150 155 160

Ala Leu Thr Arg Asn Val Pro Val Arg Asp Ile Pro Ala Phe Leu Thr

165 170 175

Trp Ala Gln Val Gly

180

<210> 10

<211> 6

<212> PRT

<213> Artificial Sequence

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

6xHis tag

<400> 10

His His His His His His

1 5

<210> 11

<211> 181

<212> PRT

<213> Artificial Sequence

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

polypeptide

<400> 11

Met Glu Leu Glu Ile Arg Pro Leu Phe Leu Val Pro Asp Thr Asn Gly

1 5 10 15

Phe Ile Asp His Leu Ala Ser Leu Ala Arg Leu Leu Glu Ser Arg Lys

20 25 30

Tyr Ile Leu Val Val Pro Leu Ile Val Ile Asn Glu Leu Asp Gly Leu

35 40 45

Ala Lys Gly Gln Glu Thr Asp His Arg Ala Gly Gly Tyr Ala Arg Val

50 55 60

Val Gln Glu Lys Ala Arg Lys Ser Ile Glu Phe Leu Glu Gln Arg Phe

65 70 75 80

Glu Ser Arg Asp Ser Cys Leu Arg Ala Leu Thr Ser Arg Gly Asn Glu

85 90 95

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

100 105 110

Asn Asn Ala Asp Leu Ile Leu Ser Cys Cys Leu His Tyr Cys Lys Asp

115 120 125

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

130 135 140

Leu Arg Glu Val Val Leu Leu Thr Asp Asp Arg Asn Leu Arg Val Lys

145 150 155 160

Ala Leu Thr Arg Asn Val Pro Val Arg Asp Ile Pro Ala Phe Leu Thr

165 170 175

Trp Ala Gln Val Gly

180

<210> 12

<211> 22

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 12

Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp His Asp Ile Asp Tyr

1 5 10 15

Lys Asp Asp Asp Asp Lys

20

<210> 13

<211> 16

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

peptide

<400> 13

Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys Lys

1 5 10 15

<210> 14

<211> 8

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 14

Asp Tyr Lys Asp Asp Asp Asp Lys

1 5

<210> 15

<211> 364

<212> PRT

<213> Artificial Sequence

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

polypeptide

<400> 15

Met Ser Ala Thr Ser Val Asp Thr Gln Arg Thr Lys Gly Gln Asp Asn

1 5 10 15

Lys Val Gln Asn Gly Ser Leu His Gln Lys Asp Thr Val His Asp Asn

20 25 30

Asp Phe Glu Pro Tyr Leu Thr Gly Gln Ser Asn Gln Ser Asn Ser Tyr
35 40 45
Pro Ser Met Ser Asp Pro Tyr Leu Ser Ser Tyr Tyr Pro Pro Ser Ile
50 55 60

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

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

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

Gly Gly Gly Leu Pro Pro Pro Pro Ile Lys His Asn Met Asp Ile Gly
260 265 270
Thr Trp Asp Asn Lys Gly Pro Val Pro Lys Ala Pro Val Pro Gln Gln

275 280 285

Ala Pro Ser Pro Gln Ala Ala Pro Gln Pro Gln Gln Val Ala Gln Pro

290 295 300

Leu Pro Ala Gln Pro Pro Ala Leu Ala Gln Pro Gln Tyr Gln Ser Pro

305 310 315 320

Gln Gln Pro Pro Gln Thr Arg Trp Val Ala Pro Arg Asn Arg Asn Ala

325 330 335

Ala Phe Gly Gln Ser Gly Gly Ala Gly Ser Asp Ser Asn Ser Pro Gly

340 345 350

Asn Val Gln Pro Asn Ser Ala Pro Ser Val Glu Ser

355 360

<210> 16

<211> 90

<212> PRT

<213> Artificial Sequence

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

polypeptide

<400>

16

Asp Ala Met Phe Gly Gln Pro Gly Ala Leu Gly Ser Thr Pro Phe Leu

1 5 10 15

Gly Gln His Gly Phe Asn Phe Phe Pro Ser Gly Ile Asp Phe Ser Ala

20 25 30

Trp Gly Asn Asn Ser Ser Gln Gly Gln Ser Thr Gln Ser Ser Gly Tyr

35 40 45

Ser Ser Asn Tyr Ala Tyr Ala Pro Ser Ser Leu Gly Gly Ala Met Ile

50 55 60

Asp Gly Gln Ser Ala Phe Ala Asn Glu Thr Leu Asn Lys Ala Pro Gly

65 70 75 80

Met Asn Thr Ile Asp Gln Gly Met Ala Ala

85 90

<210> 17

<211> 191

<212> PRT

<213> Artificial Sequence

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

<400> 17

Met Ser Ala Ser Ser Leu Leu Glu Gln Arg Pro Lys Gly Gln Gly Asn
1 5 10 15

Lys Val Gln Asn Gly Ser Val His Gln Lys Asp Gly Leu Asn Asp Asp
20 25 30

Asp Phe Glu Pro Tyr Leu Ser Pro Gln Ala Arg Pro Asn Asn Ala Tyr
35 40 45

Thr Ala Met Ser Asp Ser Tyr Leu Pro Ser Tyr Tyr Ser Pro Ser Ile
50 55 60

Gly Phe Ser Tyr Ser Leu Gly Glu Ala Ala Trp Ser Thr Gly Gly Asp
65 70 75 80

Thr Ala Met Pro Tyr Leu Thr Ser Tyr Gly Gln Leu Ser Asn Gly Glu
85 90 95

Pro His Phe Leu Pro Asp Ala Met Phe Gly Gln Pro Gly Ala Leu Gly
100 105 110

Ser Thr Pro Phe Leu Gly Gln His Gly Phe Asn Phe Phe Pro Ser Gly
115 120 125

Ile Asp Phe Ser Ala Trp Gly Asn Asn Ser Ser Gln Gly Gln Ser Thr
130 135 140

Gln Ser Ser Gly Tyr Ser Ser Asn Tyr Ala Tyr Ala Pro Ser Ser Leu
145 150 155 160

Gly Gly Ala Met Ile Asp Gly Gln Ser Ala Phe Ala Asn Glu Thr Leu
165 170 175

Asn Lys Ala Pro Gly Met Asn Thr Ile Asp Gln Gly Met Ala Ala
180 185 190

<210> 18

<211> 73

<212> PRT

<213> Artificial Sequence

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

<400> 18

Ala Asp Phe Glu Thr Asp Glu Ser Val Leu Met Arg Arg Gln Lys Gln

1 5 10 15

Ile Asn Tyr Gly Lys Asn Thr Ile Ala Tyr Asp Arg Tyr Ile Lys Glu

20 25 30

Val Pro Arg His Leu Arg Gln Pro Gly Ile His Pro Lys Thr Pro Asn

35 40 45

Lys Phe Lys Lys Tyr Ser Arg Arg Ser Trp Asp Gln Gln Ile Lys Leu

50 55 60

Trp Lys Val Ala Leu His Phe Trp Asp

65 70

<210> 19

<211> 76

<212> PRT

<213> Artificial Sequence

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

<400> 19

Met Arg Val Thr Leu Ser Ser Lys Pro Gln Ala Leu Ala Thr Pro Asn

1 5 10 15

Lys Glu Glu His Gly Lys Arg Lys Lys Lys Gly Lys Gly Leu Gly Lys

20 25 30

Lys Arg Asp Pro Cys Leu Arg Lys Tyr Lys Asp Phe Cys Ile His Gly

35 40 45

Glu Cys Lys Tyr Val Lys Glu Leu Arg Ala Pro Ser Cys Ile Cys His

50 55 60

Pro Gly Tyr His Gly Glu Arg Cys His Gly Leu Ser

65 70 75

<210> 20

<211> 46

<212> PRT

<213> Artificial Sequence

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

polypeptide

<400> 20

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

1 5 10 15

Gly Arg Cys Ala Val Leu Ser Cys Leu Pro Lys Glu Glu Gln Ile Gly

20 25 30

Lys Cys Ser Thr Arg Gly Arg Lys Cys Cys Arg Arg Lys Lys

35 40 45

<210> 21

<211> 9

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 21

Gly Gly Ser Gly Gly Ser Gly Gly Ser

1 5

<210> 22

<211> 16

<212> PRT

<213>

Artificial Sequence

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

peptide

<400> 22

Ser Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser

1 5 10 15

<210> 23

<211> 4

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 23

Gly Gly Ser Gly

1

<210> 24

<211> 73

<212> PRT

<213> Artificial Sequence

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

polypeptide

<400> 24

Glu Glu Glu Glu Lys Lys Lys Gln Gln Glu Glu Glu Ala Glu Arg Leu

1 5 10 15

Arg Arg Ile Gln Glu Glu Met Glu Lys Glu Arg Lys Arg Arg Glu Glu

20 25 30

Asp Glu Gln Arg Arg Arg Lys Glu Glu Glu Glu Arg Arg Met Lys Leu

35 40 45

Glu Met Glu Ala Lys Arg Lys Gln Glu Glu Glu Glu Arg Lys Lys Arg

50 55 60

Glu Asp Asp Glu Lys Arg Lys Lys Lys

65 70

<210> 25

<211> 5

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 25

Gly Ser Gly Gly Ser

1 5

<210> 26

<211> 31

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 26

ggccagatct gagcctggga gctctctggc c 31

<210> 27

<211> 40

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 27

tgacagccca catggcattc cacttatcac tggcatcctt 40

<210> 28

<211> 26

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 28

ccaaaggctc ttctcagagc caccca 26

<210> 29

<211> 20

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 29

gtgataagtg gaatgccatg 20

<210> 30

<211> 40

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 30

caggtcgact ctagactcga ggctagcgag ctcgtttaaa 40

<210> 31

<211> 40

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 31

cttggtgctc tccaccttcc gcaccacctc catgccctct 40

<210> 32

<211> 40

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 32

ccgatgcggt gggctcggtc ctgcgcttgc aggtcctggt 40

<210> 33

<211> 40

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 33

gcctccacca gctctctgac tgtacccccca gagacctcat 40

<210> 34

<211> 40

<212> DNA

<213> Artificial Sequence

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

<400> 34

aagcatcttc aacaccctgt ctggtcttgg ctgaggtttc 40

<210> 35

<211> 40

<212> DNA

<213> Artificial Sequence

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

<400> 35

cttcgcaccg cagcgcagcg cggctcagct cccggctcgt 40

<210> 36

<211> 20

<212> DNA

<213> Artificial Sequence

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

<400> 36

ctagactcga ggctagcgag 20

<210> 37

<211> 30

<212> DNA

<213> Artificial Sequence

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

<400> 37

cgactctaga ctcgaggcta gcgagctcgt 30

<210> 38

<211> 40

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 38

ggccctggcc cttcccctgg agccatgctg ggccctagcc 40

<210> 39

<211> 40

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 39

gcccaacccc atttaaccag aaccagctgc accagctcag 40

<210> 40

<211> 40

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 40

cctcccaagc cctggcctga aggacccatg gcgaatgctg 40

<210> 41

<211> 40

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 41

cgtcccaccc gccgcctcgc ccgtgatgcc accgcagacc 40

<210> 42

<211> 40

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 42

ggaggtggtg gtgtgcatgc ggagggacac agcgctggag 40

<210> 43

<211> 40

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 43

tcacaggcaa aatccagaag ctgaccaagg cagtggccac 40

<210> 44

<211> 40

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 44

catggctgaa gatgaggagg ggtaccgcaa gctcatcgac 40

<210> 45

<211> 40

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 45

ctctggacga gaccagccag atgagcgacc tcccggtgaa 40

<210> 46

<211> 40

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 46

cccaccctgc ccgtggagga gaagaagaag attccagatc 40

<210> 47

<211> 40

<212> DNA

<213> Artificial Sequence

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

<400> 47

aatatggcgt gtcccaggcc cttgcacgtg gcctgcagtc 40

<210> 48

<211> 40

<212> DNA

<213> Artificial Sequence

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

<400> 48

ctcatcacgt acctcatgga gcacaaacgc atcaatgggc 40

<210> 49

<211> 40

<212> DNA

<213> Artificial Sequence

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

<400> 49

tggtgaaggt gtcttacaag ggatccccag cagcaagacg 40

<210> 50

<211> 40

<212> DNA

<213> Artificial Sequence

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

<400> 50

accccgccgc ctgctgctga cgggcacacc gctgcagaac 40

<210> 51

<211> 40

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 51

cagtggttta acgcaccctt tgccatgacc ggggaaaagg 40

<210> 52

<211> 40

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 52

acgactcaag aaggaagtcg aggcccagtt gcccgaaaag 40

<210> 53

<211> 40

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 53

gtgctgctga ctgatggctc cgagaaggac aagaagggca 40

<210> 54

<211> 40

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 54

ggaaccacga aggcggagga ccggggcatg ctgctgaaaa 40

<210> 55

<211> 40

<212> DNA

<213> Artificial Sequence

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

<400> 55

tccagtcggc agacactgtg atcatttttg acagcgactg 40

<210> 56

<211> 40

<212> DNA

<213> Artificial Sequence

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

<400> 56

cgccagcggg cgtcaacccc gacttggagg agccacctct 40

<210> 57

<211> 40

<212> DNA

<213> Artificial Sequence

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

<400> 57

gcggaggtgg agcggctgac ctgtgaggag gaggaggaga 40

<210> 58

<211> 40

<212> DNA

<213> Artificial Sequence

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

<400> 58

tgatcaagta caaggacagc agcagtggac gtcagctcag 40

<210> 59

<211> 40

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 59

cttcaagaag ataaaggagc gcattcgcaa ccacaagtac 40

<210> 60

<211> 40

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 60

agggtcccga gccaagccgg tcgtgagtga cgatgacagt 40

<210> 61

<211> 40

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 61

tatttataca gcagagaagc tgtaggactg tttgtgactg 40

<210> 62

<211> 40

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 62

ggggaacaca cgatacctgt ttttcttttc cgttgctggc 40

<210> 63

<211> 70

<212> DNA

<213> Artificial Sequence

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

<400> 63

ataggccagt gaattcgagc tcgaatatgg attacttggt agaacagcaa tctacgccgg 60

aagcataaag 70

<210> 64

<211> 173

<212> DNA

<213> Artificial Sequence

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

<400> 64

ggccagtgaa ttcgagctcg gtacccgggg atcctctaga aatatggatt acttggtaga 60

acagcaatct actcgacctg caggcatgca agcttggcgt aatcatggtc atagctgttt 120

cctgtgttta tccgctcaca attccacaca acatacgagc cggaagcata aag 173

<210> 65

<211> 60

<212> DNA

<213> Artificial Sequence

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

<400> 65

ggccagatct gagcctggga gctctctggc ccagccacca cccacagagc cgccaccaga 60

<210> 66

<211> 60

<212> DNA

<213> Artificial Sequence

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

<400> 66

ggccagatct gagcctggga gctctctggc cctagattgc tgttctacca agtaatccat 60

<210> 67

<211> 22

<212> DNA

<213> Artificial Sequence

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

primer

<400> 67

gtctcctctg acttcaacag cg 22

<210> 68

<211> 22

<212> DNA

<213> Artificial Sequence

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

primer

<400> 68

accaccctgt tgctgtagcc aa 22

<210> 69

<211> 22

<212> DNA

<213> Artificial Sequence

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

primer

<400> 69

aacgcaggca aagacaccaa cg 22

<210> 70

<211> 22

<212> DNA

<213> Artificial Sequence

<220

><223> Description of Artificial Sequence: Synthetic

primer

<400> 70

tctgtcttgg tgctctccac ct 22

<210> 71

<211> 21

<212> DNA

<213> Artificial Sequence

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

primer

<400> 71

aggttacaac cgccaagaag c 21

<210> 72

<211> 21

<212> DNA

<213> Artificial Sequence

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

primer

<400> 72

atgagaatct cgcggatctt g 21

<210> 73

<211> 22

<212> DNA

<213> Artificial Sequence

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

primer

<400> 73

aacgcaggca aagacaccaa cg 22

<210> 74

<211> 22

<212> DNA

<213> Artificial Sequence

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

primer

<400> 74

tctgtcttgg tgctctccac ct 22

<210> 75

<211> 22

<212> DNA

<213> Artificial Sequence

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

<400> 75

gcagcactac ttcttgacca cc 22

<210> 76

<211> 22

<212> DNA

<213> Artificial Sequence

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

<400> 76

tctgctcctg agcattgacg tc 22

<210> 77

<211> 22

<212> DNA

<213> Artificial Sequence

<220

><223> Description of Artificial Sequence: Synthetic primer

<400> 77

gaaacctcag ccaagaccag ac 22

<210> 78

<211> 22

<212> DNA

<213> Artificial Sequence

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

<400> 78

ggcaatccca tacaaccctg ag 22

<210> 79

<211> 22

<212> DNA

<213> Artificial Sequence

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

primer

<400> 79

tgaggctgct ttcactatcc gc 22

<210> 80

<211> 22

<212> DNA

<213> Artificial Sequence

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

primer

<400> 80

gaggaaggtc ttggtggcaa tc 22

<210> 81

<211> 22

<212> DNA

<213> Artificial Sequence

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

primer

<400> 81

caaagacaag cacatcctcg cc 22

<210> 82

<211> 22

<212> DNA

<213> Artificial Sequence

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

primer

<400> 82

gccacatagt gcgtgttgag ca 22

<210> 83

<211> 43

<212> RNA

<213> Homo sapiens

<400> 83

ggcgucccuc ccgaagcugc gcgcucgguc gaacaggacg acc 43

<210> 84

<211> 6

<212> RNA

<213> Homo sapiens

<400> 84

ucccga 6

<210

> 85

<211> 22

<212> RNA

<213> Homo sapiens

<400> 85

ggccgaaauc ccgaaugagg cc 22

<210> 86

<211> 21

<212> RNA

<213> Homo sapiens

<400> 86

ggaugccucc cgagugcauc c 21

<210> 87

<211> 446

<212> PRT

<213> Artificial Sequence

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

<400> 87

Met Asp Asp Pro Ser Phe Leu Thr Gly Arg Ser Thr Tyr Ala Lys Arg
1 5 10 15

Arg Arg Ala Arg Arg Met Asn Val Cys Lys Cys Gly Ala Ile Leu His
20 25 30

Asn Asn Lys Asp Cys Arg Ser Ser Thr Ile Ser Gly His Lys Leu Asp
35 40 45

Arg Leu Arg Phe Val Lys Glu Gly Arg Val Ala Leu Glu Gly Glu Thr
50 55 60

Pro Val Tyr Arg Thr Trp Val Lys Trp Val Glu Thr Glu Tyr His Ile
65 70 75 80

Asn Ile Leu Glu Thr Ser Asp Asp Glu Glu Gly Gly Ser Gly Gly Ser
85 90 95
Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Met Ala Val Pro
100 105 110
Glu Thr Arg Pro Asn His Thr Ile Tyr Ile Asn Asn Leu Asn Ser Lys
115 120 125
Ile Lys Lys Asp Glu Leu Lys Lys Ser Leu Tyr Ala Ile Phe Ser Gln
130 135 140

Phe Gly Gln Ile Leu Asp Ile Leu Val Pro Arg Gln Arg Thr Pro Arg
145 150 155 160
Gly Gln Ala Phe Val Ile Phe Lys Glu Val Ser Ser Ala Thr Asn Ala
165 170 175
Leu Arg Ser Met Gln Gly Phe Pro Phe Tyr Asp Lys Pro Met Arg Ile
180 185 190
Gln Tyr Ala Lys Thr Asp Lys Arg Ile Pro Ala Lys Met Lys Gly Thr
195 200 205

Phe Val Gly Ser Leu Gln Leu Pro Pro Leu Glu Arg Leu Thr Leu Gly
210 215 220
Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly
225 230 235 240
Ser Met Glu Leu Glu Ile Arg Pro Leu Phe Leu Val Pro Asp Thr Asn
245 250 255
Gly Phe Ile Asp His Leu Ala Ser Leu Ala Arg Leu Leu Glu Ser Arg
260 265 270

Lys Tyr Ile Leu Val Val Pro Leu Ile Val Ile Asn Glu Leu Asp Gly
275 280 285
Leu Ala Lys Gly Gln Glu Thr Asp His Arg Ala Gly Gly Tyr Ala Arg
290 295 300
Val Val Gln Glu Lys Ala Arg Lys Ser Ile Glu Phe Leu Glu Gln Arg

305 310 315 320

Phe Glu Ser Arg Asp Ser Cys Leu Arg Ala Leu Thr Ser Arg Gly Asn

325 330 335

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

340 345 350

Gly Asn Asn Ala Asp Leu Ile Leu Ser Cys Cys Leu His Tyr Cys Lys

355 360 365

Asp Lys Ala Lys Asp Phe Met Pro Ala Ser Lys Glu Glu Pro Ile Arg

370 375 380

Leu Leu Arg Glu Val Val Leu Leu Thr Asp Asp Arg Asn Leu Arg Val

385 390 395 400

Lys Ala Leu Thr Arg Asn Val Pro Val Arg Asp Ile Pro Ala Phe Leu

405 410 415

Thr Trp Ala Gln Val Gly Gly Ser Asp Tyr Lys Asp His Asp Gly Asp

420 425 430

Tyr Lys Asp His Asp Ile Asp Tyr Lys Asp Asp Asp Asp Lys

435 440 445

<210> 88

<211> 435

<212> PRT

<213> Artificial Sequence

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

polypeptide

<400> 88

Met Asp Asp Pro Ser Phe Leu Thr Gly Arg Ser Thr Tyr Ala Lys Arg

1 5 10 15

Arg Arg Ala Arg Arg Met Asn Val Cys Lys Cys Gly Ala Ile Leu His

20 25 30

Asn Asn Lys Asp Cys Arg Ser Ser Thr Ile Ser Gly His Lys Leu Asp

35 40 45

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

50 55 60

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

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

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

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

Glu Gln Arg Phe Glu Ser Arg Asp Ser Cys Leu Arg Ala Leu Thr Ser
305 310 315 320

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

Leu Arg Val Lys Ala Leu Thr Arg Asn Val Pro Val Arg Asp Ile Pro
385 390 395 400
Ala Phe Leu Thr Trp Ala Gln Val Gly Lys Arg Pro Ala Ala Thr Lys
405 410 415
Lys Ala Gly Gln Ala Lys Lys Lys Lys Gly Ser Asp Tyr Lys Asp Asp
420 425 430
Asp Asp Lys
435
<210> 89
<211> 615
<212> PRT
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic

polypeptide
<400> 89
Met Asp Asp Pro Ser Phe Leu Thr Gly Arg Ser Thr Tyr Ala Lys Arg
1 5 10 15
Arg Arg Ala Arg Arg Met Asn Val Cys Lys Cys Gly Ala Ile Leu His
20 25 30
Asn Asn Lys Asp Cys Arg Ser Ser Thr Ile Ser Gly His Lys Leu Asp
35 40 45
Arg Leu Arg Phe Val Lys Glu Gly Arg Val Ala Leu Glu Gly Glu Thr

50 55 60

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

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

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

Asn Lys Val Gln Asn Gly Ser Leu His Gln Lys Asp Thr Val His Asp
260 265 270
Asn Asp Phe Glu Pro Tyr Leu Thr Gly Gln Ser Asn Gln Ser Asn Ser
275 280 285
Tyr Pro Ser Met Ser Asp Pro Tyr Leu Ser Ser Tyr Tyr Pro Pro Ser
290 295 300

Ile Gly Phe Pro Tyr Ser Leu Asn Glu Ala Pro Trp Ser Thr Ala Gly
305 310 315 320

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

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

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

Gly Thr Trp Asp Asn Lys Gly Pro Val Pro Lys Ala Pro Val Pro Gln
515 520 525
Gln Ala Pro Ser Pro Gln Ala Ala Pro Gln Pro Gln Gln Val Ala Gln
530 535 540
Pro Leu Pro Ala Gln Pro Pro Ala Leu Ala Gln Pro Gln Tyr Gln Ser

545 550 555 560

Pro Gln Gln Pro Pro Gln Thr Arg Trp Val Ala Pro Arg Asn Arg Asn

565 570 575

Ala Ala Phe Gly Gln Ser Gly Gly Ala Gly Ser Asp Ser Asn Ser Pro

580 585 590

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

595 600 605

Tyr Lys Asp Asp Asp Asp Lys

610 615

<210> 90

<211> 341

<212> PRT

<213> Artificial Sequence

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

polypeptide

<400> 90

Met Asp Asp Pro Ser Phe Leu Thr Gly Arg Ser Thr Tyr Ala Lys Arg

1 5 10 15

Arg Arg Ala Arg Arg Met Asn Val Cys Lys Cys Gly Ala Ile Leu His

20 25 30

Asn Asn Lys Asp Cys Arg Ser Ser Thr Ile Ser Gly His Lys Leu Asp

35 40 45

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

50 55 60

Pro Val Tyr Arg Thr Trp Val Lys Trp Val Glu Thr Glu Tyr His Ile

65 70 75 80

Asn Ile Leu Glu Thr Ser Asp Asp Glu Glu Gly Gly Ser Gly Gly Ser

85 90 95

Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Met Ala Val Pro

100 105 110

Glu Thr Arg Pro Asn His Thr Ile Tyr Ile Asn Asn Leu Asn Ser Lys

115 120 125

Ile Lys Lys Asp Glu Leu Lys Lys Ser Leu Tyr Ala Ile Phe Ser Gln

130 135 140

Phe Gly Gln Ile Leu Asp Ile Leu Val Pro Arg Gln Arg Thr Pro Arg

145 150 155 160

Gly Gln Ala Phe Val Ile Phe Lys Glu Val Ser Ser Ala Thr Asn Ala

165 170 175

Leu Arg Ser Met Gln Gly Phe Pro Phe Tyr Asp Lys Pro Met Arg Ile

180 185 190

Gln Tyr Ala Lys Thr Asp Lys Arg Ile Pro Ala Lys Met Lys Gly Thr

195 200 205

Phe Val Gly Ser Leu Gln Leu Pro Pro Leu Glu Arg Leu Thr Leu Gly

210 215 220

Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly

225 230 235 240

Ser Asp Ala Met Phe Gly Gln Pro Gly Ala Leu Gly Ser Thr Pro Phe

245 250 255

Leu Gly Gln His Gly Phe Asn Phe Phe Pro Ser Gly Ile Asp Phe Ser

260 265 270

Ala Trp Gly Asn Asn Ser Ser Gln Gly Gln Ser Thr Gln Ser Ser Gly

275 280 285

Tyr Ser Ser Asn Tyr Ala Tyr Ala Pro Ser Ser Leu Gly Gly Ala Met

290 295 300

Ile Asp Gly Gln Ser Ala Phe Ala Asn Glu Thr Leu Asn Lys Ala Pro

305 310 315 320

Gly Met Asn Thr Ile Asp Gln Gly Met Ala Ala Gly Ser Asp Tyr Lys

325 330 335

Asp Asp Asp Asp Lys

340

<210> 91

<211> 442

<212> PRT

<213> Artificial Sequence

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

<400> 91

Met Asp Asp Pro Ser Phe Leu Thr Gly Arg Ser Thr Tyr Ala Lys Arg
1 5 10 15

Arg Arg Ala Arg Arg Met Asn Val Cys Lys Cys Gly Ala Ile Leu His
20 25 30

Asn Asn Lys Asp Cys Arg Ser Ser Thr Ile Ser Gly His Lys Leu Asp
35 40 45

Arg Leu Arg Phe Val Lys Glu Gly Arg Val Ala Leu Glu Gly Glu Thr
50 55 60

Pro Val Tyr Arg Thr Trp Val Lys Trp Val Glu Thr Glu Tyr His Ile
65 70 75 80

Asn Ile Leu Glu Thr Ser Asp Asp Glu Glu Gly Gly Ser Gly Gly Ser
85 90 95

Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Met Ala Val Pro
100 105 110

Glu Thr Arg Pro Asn His Thr Ile Tyr Ile Asn Asn Leu Asn Ser Lys
115 120 125

Ile Lys Lys Asp Glu Leu Lys Lys Ser Leu Tyr Ala Ile Phe Ser Gln
130 135 140

Phe Gly Gln Ile Leu Asp Ile Leu Val Pro Arg Gln Arg Thr Pro Arg
145 150 155 160

Gly Gln Ala Phe Val Ile Phe Lys Glu Val Ser Ser Ala Thr Asn Ala
165 170 175

Leu Arg Ser Met Gln Gly Phe Pro Phe Tyr Asp Lys Pro Met Arg Ile
180 185 190

Gln Tyr Ala Lys Thr Asp Lys Arg Ile Pro Ala Lys Met Lys Gly Thr
195 200 205

Phe Val Gly Ser Leu Gln Leu Pro Pro Leu Glu Arg Leu Thr Leu Gly

210 215 220

Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly
225 230 235 240
Ser Met Ser Ala Ser Ser Leu Leu Glu Gln Arg Pro Lys Gly Gln Gly
245 250 255
Asn Lys Val Gln Asn Gly Ser Val His Gln Lys Asp Gly Leu Asn Asp
260 265 270
Asp Asp Phe Glu Pro Tyr Leu Ser Pro Gln Ala Arg Pro Asn Asn Ala
275 280 285

Tyr Thr Ala Met Ser Asp Ser Tyr Leu Pro Ser Tyr Tyr Ser Pro Ser
290 295 300
Ile Gly Phe Ser Tyr Ser Leu Gly Glu Ala Ala Trp Ser Thr Gly Gly
305 310 315 320
Asp Thr Ala Met Pro Tyr Leu Thr Ser Tyr Gly Gln Leu Ser Asn Gly
325 330 335
Glu Pro His Phe Leu Pro Asp Ala Met Phe Gly Gln Pro Gly Ala Leu
340 345 350

Gly Ser Thr Pro Phe Leu Gly Gln His Gly Phe Asn Phe Phe Pro Ser
355 360 365
Gly Ile Asp Phe Ser Ala Trp Gly Asn Asn Ser Ser Gln Gly Gln Ser
370 375 380
Thr Gln Ser Ser Gly Tyr Ser Ser Asn Tyr Ala Tyr Ala Pro Ser Ser
385 390 395 400
Leu Gly Gly Ala Met Ile Asp Gly Gln Ser Ala Phe Ala Asn Glu Thr
405 410 415

Leu Asn Lys Ala Pro Gly Met Asn Thr Ile Asp Gln Gly Met Ala Ala
420 425 430
Gly Ser Asp Tyr Lys Asp Asp Asp Asp Lys
435 440
<210> 92
<211> 310

<212> PRT

<213> Artificial Sequence

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

<400> 92

Met Asp Asp Pro Ser Phe Leu Thr Gly Arg Ser Thr Tyr Ala Lys Arg
1 5 10 15

Arg Arg Ala Arg Arg Met Asn Val Cys Lys Cys Gly Ala Ile Leu His
20 25 30

Asn Asn Lys Asp Cys Arg Ser Ser Thr Ile Ser Gly His Lys Leu Asp
35 40 45

Arg Leu Arg Phe Val Lys Glu Gly Arg Val Ala Leu Glu Gly Glu Thr
50 55 60

Pro Val Tyr Arg Thr Trp Val Lys Trp Val Glu Thr Glu Tyr His Ile
65 70 75 80

Asn Ile Leu Glu Thr Ser Asp Asp Glu Glu Gly Gly Ser Gly Gly Ser
85 90 95

Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Ala Asp Phe Glu
100 105 110

Thr Asp Glu Ser Val Leu Met Arg Arg Gln Lys Gln Ile Asn Tyr Gly
115 120 125

Lys Asn Thr Ile Ala Tyr Asp Arg Tyr Ile Lys Glu Val Pro Arg His
130 135 140

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

Tyr Ser Arg Arg Ser Trp Asp Gln Gln Ile Lys Leu Trp Lys Val Ala
165 170 175

Leu His Phe Trp Asp Leu Gln Leu Pro Pro Leu Glu Arg Leu Thr Leu
180 185 190

Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly
195 200 205

Gly Ser Asp Ala Met Phe Gly Gln Pro Gly Ala Leu Gly Ser Thr Pro

210 215 220

Phe Leu Gly Gln His Gly Phe Asn Phe Phe Pro Ser Gly Ile Asp Phe

225 230 235 240

Ser Ala Trp Gly Asn Asn Ser Ser Gln Gly Gln Ser Thr Gln Ser Ser

245 250 255

Gly Tyr Ser Ser Asn Tyr Ala Tyr Ala Pro Ser Ser Leu Gly Gly Ala

260 265 270

Met Ile Asp Gly Gln Ser Ala Phe Ala Asn Glu Thr Leu Asn Lys Ala

275 280 285

Pro Gly Met Asn Thr Ile Asp Gln Gly Met Ala Ala Gly Ser Asp Tyr

290 295 300

Lys Asp Asp Asp Asp Lys

305 310

<210> 93

<211> 411

<212> PRT

<213> Artificial Sequence

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

polypeptide

<400> 93

Met Asp Asp Pro Ser Phe Leu Thr Gly Arg Ser Thr Tyr Ala Lys Arg

1 5 10 15

Arg Arg Ala Arg Arg Met Asn Val Cys Lys Cys Gly Ala Ile Leu His

20 25 30

Asn Asn Lys Asp Cys Arg Ser Ser Thr Ile Ser Gly His Lys Leu Asp

35 40 45

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

50 55 60

Pro Val Tyr Arg Thr Trp Val Lys Trp Val Glu Thr Glu Tyr His Ile

65 70 75 80

Asn Ile Leu Glu Thr Ser Asp Asp Glu Glu Gly Gly Ser Gly Gly Ser
85 90 95
Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Ala Asp Phe Glu
100 105 110
Thr Asp Glu Ser Val Leu Met Arg Arg Gln Lys Gln Ile Asn Tyr Gly
115 120 125
Lys Asn Thr Ile Ala Tyr Asp Arg Tyr Ile Lys Glu Val Pro Arg His
130 135 140

Leu Arg Gln Pro Gly Ile His Pro Lys Thr Pro Asn Lys Phe Lys Lys
145 150 155 160
Tyr Ser Arg Arg Ser Trp Asp Gln Gln Ile Lys Leu Trp Lys Val Ala
165 170 175
Leu His Phe Trp Asp Leu Gln Leu Pro Pro Leu Glu Arg Leu Thr Leu
180 185 190
Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly
195 200 205

Gly Ser Met Ser Ala Ser Ser Leu Leu Glu Gln Arg Pro Lys Gly Gln
210 215 220
Gly Asn Lys Val Gln Asn Gly Ser Val His Gln Lys Asp Gly Leu Asn
225 230 235 240
Asp Asp Asp Phe Glu Pro Tyr Leu Ser Pro Gln Ala Arg Pro Asn Asn
245 250 255
Ala Tyr Thr Ala Met Ser Asp Ser Tyr Leu Pro Ser Tyr Tyr Ser Pro
260 265 270

Ser Ile Gly Phe Ser Tyr Ser Leu Gly Glu Ala Ala Trp Ser Thr Gly
275 280 285
Gly Asp Thr Ala Met Pro Tyr Leu Thr Ser Tyr Gly Gln Leu Ser Asn
290 295 300
Gly Glu Pro His Phe Leu Pro Asp Ala Met Phe Gly Gln Pro Gly Ala
305 310 315 320
Leu Gly Ser Thr Pro Phe Leu Gly Gln His Gly Phe Asn Phe Phe Pro

325 330 335

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

Ala Gly Ser Asp Tyr Lys Asp Asp Asp Asp Lys
405 410

<210> 94
<211> 327
<212> PRT
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
polypeptide
<400> 94
Met Arg Val Thr Leu Ser Ser Lys Pro Gln Ala Leu Ala Thr Pro Asn
1 5 10 15
Lys Glu Glu His Gly Lys Arg Lys Lys Lys Gly Lys Gly Leu Gly Lys
20 25 30
Lys Arg Asp Pro Cys Leu Arg Lys Tyr Lys Asp Phe Cys Ile His Gly

35 40 45
Glu Cys Lys Tyr Val Lys Glu Leu Arg Ala Pro Ser Cys Ile Cys His
50 55 60
Pro Gly Tyr His Gly Glu Arg Cys His Gly Leu Ser Gly Gly Ser Gly
65 70 75 80
Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Met Ala
85 90 95
Val Pro Glu Thr Arg Pro Asn His Thr Ile Tyr Ile Asn Asn Leu Asn

100 105 110

Ser Lys Ile Lys Lys Asp Glu Leu Lys Lys Ser Leu Tyr Ala Ile Phe

115 120 125

Ser Gln Phe Gly Gln Ile Leu Asp Ile Leu Val Pro Arg Gln Arg Thr

130 135 140

Pro Arg Gly Gln Ala Phe Val Ile Phe Lys Glu Val Ser Ser Ala Thr

145 150 155 160

Asn Ala Leu Arg Ser Met Gln Gly Phe Pro Phe Tyr Asp Lys Pro Met

165 170 175

Arg Ile Gln Tyr Ala Lys Thr Asp Lys Arg Ile Pro Ala Lys Met Lys

180 185 190

Gly Thr Phe Val Gly Ser Leu Gln Leu Pro Pro Leu Glu Arg Leu Thr

195 200 205

Leu Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser

210 215 220

Gly Gly Ser Asp Ala Met Phe Gly Gln Pro Gly Ala Leu Gly Ser Thr

225 230 235 240

Pro Phe Leu Gly Gln His Gly Phe Asn Phe Phe Pro Ser Gly Ile Asp

245 250 255

Phe Ser Ala Trp Gly Asn Asn Ser Ser Gln Gly Gln Ser Thr Gln Ser

260 265 270

Ser Gly Tyr Ser Ser Asn Tyr Ala Tyr Ala Pro Ser Ser Leu Gly Gly

275 280 285

Ala Met Ile Asp Gly Gln Ser Ala Phe Ala Asn Glu Thr Leu Asn Lys

290 295 300

Ala Pro Gly Met Asn Thr Ile Asp Gln Gly Met Ala Ala Gly Ser Asp

305 310 315 320

Tyr Lys Asp Asp Asp Asp Lys

325

<210> 95

<211> 428

<212> PRT

<213> Artificial Sequence

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

<400> 95

Met Arg Val Thr Leu Ser Ser Lys Pro Gln Ala Leu Ala Thr Pro Asn
1 5 10 15

Lys Glu Glu His Gly Lys Arg Lys Lys Lys Gly Lys Gly Leu Gly Lys
20 25 30

Lys Arg Asp Pro Cys Leu Arg Lys Tyr Lys Asp Phe Cys Ile His Gly
35 40 45

Glu Cys Lys Tyr Val Lys Glu Leu Arg Ala Pro Ser Cys Ile Cys His
50 55 60

Pro Gly Tyr His Gly Glu Arg Cys His Gly Leu Ser Gly Gly Ser Gly
65 70 75 80

Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Met Ala
85 90 95

Val Pro Glu Thr Arg Pro Asn His Thr Ile Tyr Ile Asn Asn Leu Asn
100 105 110

Ser Lys Ile Lys Lys Asp Glu Leu Lys Lys Ser Leu Tyr Ala Ile Phe
115 120 125

Ser Gln Phe Gly Gln Ile Leu Asp Ile Leu Val Pro Arg Gln Arg Thr
130 135 140

Pro Arg Gly Gln Ala Phe Val Ile Phe Lys Glu Val Ser Ser Ala Thr
145 150 155 160

Asn Ala Leu Arg Ser Met Gln Gly Phe Pro Phe Tyr Asp Lys Pro Met
165 170 175

Arg Ile Gln Tyr Ala Lys Thr Asp Lys Arg Ile Pro Ala Lys Met Lys
180 185 190

Gly Thr Phe Val Gly Ser Leu Gln Leu Pro Pro Leu Glu Arg Leu Thr
195 200 205

Leu Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser

210 215 220

Gly Gly Ser Met Ser Ala Ser Ser Leu Leu Glu Gln Arg Pro Lys Gly

225 230 235 240

Gln Gly Asn Lys Val Gln Asn Gly Ser Val His Gln Lys Asp Gly Leu

245 250 255

Asn Asp Asp Asp Phe Glu Pro Tyr Leu Ser Pro Gln Ala Arg Pro Asn

260 265 270

Asn Ala Tyr Thr Ala Met Ser Asp Ser Tyr Leu Pro Ser Tyr Tyr Ser

275 280 285

Pro Ser Ile Gly Phe Ser Tyr Ser Leu Gly Glu Ala Ala Trp Ser Thr

290 295 300

Gly Gly Asp Thr Ala Met Pro Tyr Leu Thr Ser Tyr Gly Gln Leu Ser

305 310 315 320

Asn Gly Glu Pro His Phe Leu Pro Asp Ala Met Phe Gly Gln Pro Gly

325 330 335

Ala Leu Gly Ser Thr Pro Phe Leu Gly Gln His Gly Phe Asn Phe Phe

340 345 350

Pro Ser Gly Ile Asp Phe Ser Ala Trp Gly Asn Asn Ser Ser Gln Gly

355 360 365

Gln Ser Thr Gln Ser Ser Gly Tyr Ser Ser Asn Tyr Ala Tyr Ala Pro

370 375 380

Ser Ser Leu Gly Gly Ala Met Ile Asp Gly Gln Ser Ala Phe Ala Asn

385 390 395 400

Glu Thr Leu Asn Lys Ala Pro Gly Met Asn Thr Ile Asp Gln Gly Met

405 410 415

Ala Ala Gly Ser Asp Tyr Lys Asp Asp Asp Asp Lys

420 425

<210> 96

<211> 297

<212> PRT

<213> Artificial Sequence

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

polypeptide

<400> 96

Met Gly Ile Ile Asn Thr Leu Gln Lys Tyr Tyr Cys Arg Val Arg Gly
1 5 10 15

Gly Arg Cys Ala Val Leu Ser Cys Leu Pro Lys Glu Glu Gln Ile Gly
20 25 30

Lys Cys Ser Thr Arg Gly Arg Lys Cys Cys Arg Arg Lys Lys Gly Gly
35 40 45

Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser
50 55 60

Met Ala Val Pro Glu Thr Arg Pro Asn His Thr Ile Tyr Ile Asn Asn
65 70 75 80

Leu Asn Ser Lys Ile Lys Lys Asp Glu Leu Lys Lys Ser Leu Tyr Ala
85 90 95

Ile Phe Ser Gln Phe Gly Gln Ile Leu Asp Ile Leu Val Pro Arg Gln
100 105 110

Arg Thr Pro Arg Gly Gln Ala Phe Val Ile Phe Lys Glu Val Ser Ser
115 120 125

Ala Thr Asn Ala Leu Arg Ser Met Gln Gly Phe Pro Phe Tyr Asp Lys
130 135 140

Pro Met Arg Ile Gln Tyr Ala Lys Thr Asp Lys Arg Ile Pro Ala Lys
145 150 155 160

Met Lys Gly Thr Phe Val Gly Ser Leu Gln Leu Pro Pro Leu Glu Arg
165 170 175

Leu Thr Leu Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly
180 185 190

Gly Ser Gly Gly Ser Asp Ala Met Phe Gly Gln Pro Gly Ala Leu Gly
195 200 205

Ser Thr Pro Phe Leu Gly Gln His Gly Phe Asn Phe Phe Pro Ser Gly
210 215 220

Ile Asp Phe Ser Ala Trp Gly Asn Asn Ser Ser Gln Gly Gln Ser Thr

225 230 235 240

Gln Ser Ser Gly Tyr Ser Ser Asn Tyr Ala Tyr Ala Pro Ser Ser Leu

245 250 255

Gly Gly Ala Met Ile Asp Gly Gln Ser Ala Phe Ala Asn Glu Thr Leu

260 265 270

Asn Lys Ala Pro Gly Met Asn Thr Ile Asp Gln Gly Met Ala Ala Gly

275 280 285

Ser Asp Tyr Lys Asp Asp Asp Asp Lys

290 295

<210> 97

<211> 398

<212> PRT

<213> Artificial Sequence

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

polypeptide

<400> 97

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

1 5 10 15

Gly Arg Cys Ala Val Leu Ser Cys Leu Pro Lys Glu Glu Gln Ile Gly

20 25 30

Lys Cys Ser Thr Arg Gly Arg Lys Cys Cys Arg Arg Lys Lys Gly Gly

35 40 45

Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser

50 55 60

Met Ala Val Pro Glu Thr Arg Pro Asn His Thr Ile Tyr Ile Asn Asn

65 70 75 80

Leu Asn Ser Lys Ile Lys Lys Asp Glu Leu Lys Lys Ser Leu Tyr Ala

85 90 95

Ile Phe Ser Gln Phe Gly Gln Ile Leu Asp Ile Leu Val Pro Arg Gln

100 105 110

Arg Thr Pro Arg Gly Gln Ala Phe Val Ile Phe Lys Glu Val Ser Ser

115 120 125

Ala Thr Asn Ala Leu Arg Ser Met Gln Gly Phe Pro Phe Tyr Asp Lys

130 135 140

Pro Met Arg Ile Gln Tyr Ala Lys Thr Asp Lys Arg Ile Pro Ala Lys

145 150 155 160

Met Lys Gly Thr Phe Val Gly Ser Leu Gln Leu Pro Pro Leu Glu Arg

165 170 175

Leu Thr Leu Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly

180 185 190

Gly Ser Gly Gly Ser Met Ser Ala Ser Ser Leu Leu Glu Gln Arg Pro

195 200 205

Lys Gly Gln Gly Asn Lys Val Gln Asn Gly Ser Val His Gln Lys Asp

210 215 220

Gly Leu Asn Asp Asp Asp Phe Glu Pro Tyr Leu Ser Pro Gln Ala Arg

225 230 235 240

Pro Asn Asn Ala Tyr Thr Ala Met Ser Asp Ser Tyr Leu Pro Ser Tyr

245 250 255

Tyr Ser Pro Ser Ile Gly Phe Ser Tyr Ser Leu Gly Glu Ala Ala Trp

260 265 270

Ser Thr Gly Gly Asp Thr Ala Met Pro Tyr Leu Thr Ser Tyr Gly Gln

275 280 285

Leu Ser Asn Gly Glu Pro His Phe Leu Pro Asp Ala Met Phe Gly Gln

290 295 300

Pro Gly Ala Leu Gly Ser Thr Pro Phe Leu Gly Gln His Gly Phe Asn

305 310 315 320

Phe Phe Pro Ser Gly Ile Asp Phe Ser Ala Trp Gly Asn Asn Ser Ser

325 330 335

Gln Gly Gln Ser Thr Gln Ser Ser Gly Tyr Ser Ser Asn Tyr Ala Tyr

340 345 350

Ala Pro Ser Ser Leu Gly Gly Ala Met Ile Asp Gly Gln Ser Ala Phe

355 360 365

Ala Asn Glu Thr Leu Asn Lys Ala Pro Gly Met Asn Thr Ile Asp Gln

370 375 380

Gly Met Ala Ala Gly Ser Asp Tyr Lys Asp Asp Asp Asp Lys

385 390 395

<210> 98

<211> 367

<212> PRT

<213> Artificial Sequence

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

<400> 98

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

1 5 10 15

Gly Arg Cys Ala Val Leu Ser Cys Leu Pro Lys Glu Glu Gln Ile Gly

20 25 30

Lys Cys Ser Thr Arg Gly Arg Lys Cys Cys Arg Arg Lys Lys Gly Gly

35 40 45

Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser

50 55 60

Ala Asp Phe Glu Thr Asp Glu Ser Val Leu Met Arg Arg Gln Lys Gln

65 70 75 80

Ile Asn Tyr Gly Lys Asn Thr Ile Ala Tyr Asp Arg Tyr Ile Lys Glu

85 90 95

Val Pro Arg His Leu Arg Gln Pro Gly Ile His Pro Lys Thr Pro Asn

100 105 110

Lys Phe Lys Lys Tyr Ser Arg Arg Ser Trp Asp Gln Gln Ile Lys Leu

115 120 125

Trp Lys Val Ala Leu His Phe Trp Asp Leu Gln Leu Pro Pro Leu Glu

130 135 140

Arg Leu Thr Leu Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser

145 150 155 160

Gly Gly Ser Gly Gly Ser Met Ser Ala Ser Ser Leu Leu Glu Gln Arg

165 170 175

Pro Lys Gly Gln Gly Asn Lys Val Gln Asn Gly Ser Val His Gln Lys

180 185 190

Asp Gly Leu Asn Asp Asp Asp Phe Glu Pro Tyr Leu Ser Pro Gln Ala

195 200 205

Arg Pro Asn Asn Ala Tyr Thr Ala Met Ser Asp Ser Tyr Leu Pro Ser

210 215 220

Tyr Tyr Ser Pro Ser Ile Gly Phe Ser Tyr Ser Leu Gly Glu Ala Ala

225 230 235 240

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

245 250 255

Gln Leu Ser Asn Gly Glu Pro His Phe Leu Pro Asp Ala Met Phe Gly

260 265 270

Gln Pro Gly Ala Leu Gly Ser Thr Pro Phe Leu Gly Gln His Gly Phe

275 280 285

Asn Phe Phe Pro Ser Gly Ile Asp Phe Ser Ala Trp Gly Asn Asn Ser

290 295 300

Ser Gln Gly Gln Ser Thr Gln Ser Ser Gly Tyr Ser Ser Asn Tyr Ala

305 310 315 320

Tyr Ala Pro Ser Ser Leu Gly Gly Ala Met Ile Asp Gly Gln Ser Ala

325 330 335

Phe Ala Asn Glu Thr Leu Asn Lys Ala Pro Gly Met Asn Thr Ile Asp

340 345 350

Gln Gly Met Ala Ala Gly Ser Asp Tyr Lys Asp Asp Asp Asp Lys

355 360 365

<210> 99

<211> 324

<212> PRT

<213> Artificial Sequence

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

polypeptide

<400> 99

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

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

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

180 185 190
Gly Tyr Ala Arg Val Val Gln Glu Lys Ala Arg Lys Ser Ile Glu Phe
195 200 205
Leu Glu Gln Arg Phe Glu Ser Arg Asp Ser Cys Leu Arg Ala Leu Thr
210 215 220
Ser Arg Gly Asn Glu Leu Glu Ser Ile Ala Phe Arg Ser Glu Asp Ile
225 230 235 240
Thr Gly Gln Leu Gly Asn Asn Asp Asp Leu Ile Leu Ser Cys Cys Leu

245 250 255

His Tyr Cys Lys Asp Lys Ala Lys Asp Phe Met Pro Ala Ser Lys Glu

260 265 270

Glu Pro Ile Arg Leu Leu Arg Glu Val Val Leu Leu Thr Asp Asp Arg

275 280 285

Asn Leu Arg Val Lys Ala Leu Thr Arg Asn Val Pro Val Arg Asp Ile

290 295 300

Pro Ala Phe Leu Thr Trp Ala Gln Val Gly Gly Ser Asp Tyr Lys Asp

305 310 315 320

Asp Asp Asp Lys

<210> 100

<211> 432

<212> PRT

<213> Artificial Sequence

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

polypeptide

<400> 100

Met Asp Asp Pro Ser Phe Leu Thr Gly Arg Ser Thr Tyr Ala Lys Arg

1 5 10 15

Arg Arg Ala Arg Arg Met Asn Val Cys Lys Cys Gly Ala Ile Leu His

20 25 30

Asn Asn Lys Asp Cys Arg Ser Ser Thr Ile Ser Gly His Lys Leu Asp

35 40 45

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

50 55 60

Pro Val Tyr Arg Thr Trp Val Lys Trp Val Glu Thr Glu Tyr His Ile

65 70 75 80

Asn Ile Leu Glu Thr Ser Asp Asp Glu Glu Gly Gly Ser Gly Gly Ser

85 90 95

Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Met Ala Val Pro

100 105 110

Glu Thr Arg Pro Asn His Thr Ile Tyr Ile Asn Asn Leu Asn Ser Lys
115 120 125
Ile Lys Lys Asp Glu Leu Lys Lys Ser Leu Tyr Ala Ile Phe Ser Gln
130 135 140
Phe Gly Gln Ile Leu Asp Ile Leu Val Pro Arg Gln Arg Thr Pro Arg
145 150 155 160

Gly Gln Ala Phe Val Ile Phe Lys Glu Val Ser Ser Ala Thr Asn Ala
165 170 175
Leu Arg Ser Met Gln Gly Phe Pro Phe Tyr Asp Lys Pro Met Arg Ile
180 185 190
Gln Tyr Ala Lys Thr Asp Lys Arg Ile Pro Ala Lys Met Lys Gly Thr
195 200 205
Phe Val Gly Ser Leu Gln Leu Pro Pro Leu Glu Arg Leu Thr Leu Gly
210 215 220

Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly
225 230 235 240
Ser Met Glu Leu Glu Ile Arg Pro Leu Phe Leu Val Pro Asp Thr Asn
245 250 255
Gly Phe Ile Asp His Leu Ala Ser Leu Ala Arg Leu Leu Glu Ser Arg
260 265 270
Lys Tyr Ile Leu Val Val Pro Leu Ile Val Ile Asn Glu Leu Asp Gly
275 280 285

Leu Ala Lys Gly Gln Glu Thr Asp His Arg Ala Gly Gly Tyr Ala Arg
290 295 300
Val Val Gln Glu Lys Ala Arg Lys Ser Ile Glu Phe Leu Glu Gln Arg
305 310 315 320
Phe Glu Ser Arg Asp Ser Cys Leu Arg Ala Leu Thr Ser Arg Gly Asn
325 330 335
Glu Leu Glu Ser Ile Ala Phe Arg Ser Glu Asp Ile Thr Gly Gln Leu
340 345 350

Gly Asn Asn Asp Asp Leu Ile Leu Ser Cys Cys Leu His Tyr Cys Lys

355 360 365

Asp Lys Ala Lys Asp Phe Met Pro Ala Ser Lys Glu Glu Pro Ile Arg

370 375 380

Leu Leu Arg Glu Val Val Leu Leu Thr Asp Asp Arg Asn Leu Arg Val

385 390 395 400

Lys Ala Leu Thr Arg Asn Val Pro Val Arg Asp Ile Pro Ala Phe Leu

405 410 415

Thr Trp Ala Gln Val Gly Gly Ser Asp Tyr Lys Asp Asp Asp Asp Lys

420 425 430

<210> 101

<211> 423

<212> PRT

<213> Artificial Sequence

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

polypeptide

<400> 101

Met Asp Asp Pro Ser Phe Leu Thr Gly Arg Ser Thr Tyr Ala Lys Arg

1 5 10 15

Arg Arg Ala Arg Arg Met Asn Val Cys Lys Cys Gly Ala Ile Leu His

20 25 30

Asn Asn Lys Asp Cys Arg Ser Ser Thr Ile Ser Gly His Lys Leu Asp

35 40 45

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

50 55 60

Pro Val Tyr Arg Thr Trp Val Lys Trp Val Glu Thr Glu Tyr His Ile

65 70 75 80

Asn Ile Leu Glu Thr Ser Asp Asp Glu Glu Gly Gly Ser Gly Gly Ser

85 90 95

Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Met Ala Val Pro

100 105 110

Glu Thr Arg Pro Asn His Thr Ile Tyr Ile Asn Asn Leu Asn Ser Lys

115 120 125

Ile Lys Lys Asp Glu Leu Lys Lys Ser Leu Tyr Ala Ile Phe Ser Gln
130 135 140
Phe Gly Gln Ile Leu Asp Ile Leu Val Pro Arg Gln Arg Thr Pro Arg
145 150 155 160

Gly Gln Ala Phe Val Ile Phe Lys Glu Val Ser Ser Ala Thr Asn Ala
165 170 175
Leu Arg Ser Met Gln Gly Phe Pro Phe Tyr Asp Lys Pro Met Arg Ile
180 185 190
Gln Tyr Ala Lys Thr Asp Lys Arg Ile Pro Ala Lys Met Lys Gly Thr
195 200 205
Phe Val Gly Ser Leu Gln Leu Pro Pro Leu Glu Arg Leu Thr Leu Gly
210 215 220

Gly Ser Gly Gly Ser Gly Gly Ser Met Glu Leu Glu Ile Arg Pro Leu
225 230 235 240
Phe Leu Val Pro Asp Thr Asn Gly Phe Ile Asp His Leu Ala Ser Leu
245 250 255
Ala Arg Leu Leu Glu Ser Arg Lys Tyr Ile Leu Val Val Pro Leu Ile
260 265 270
Val Ile Asn Glu Leu Asp Gly Leu Ala Lys Gly Gln Glu Thr Asp His
275 280 285

Arg Ala Gly Gly Tyr Ala Arg Val Val Gln Glu Lys Ala Arg Lys Ser
290 295 300
Ile Glu Phe Leu Glu Gln Arg Phe Glu Ser Arg Asp Ser Cys Leu Arg
305 310 315 320
Ala Leu Thr Ser Arg Gly Asn Glu Leu Glu Ser Ile Ala Phe Arg Ser
325 330 335
Glu Asp Ile Thr Gly Gln Leu Gly Asn Asn Asp Asp Leu Ile Leu Ser
340 345 350

Cys Cys Leu His Tyr Cys Lys Asp Lys Ala Lys Asp Phe Met Pro Ala
355 360 365
Ser Lys Glu Glu Pro Ile Arg Leu Leu Arg Glu Val Val Leu Leu Thr

370 375 380

Asp Asp Arg Asn Leu Arg Val Lys Ala Leu Thr Arg Asn Val Pro Val

385 390 395 400

Arg Asp Ile Pro Ala Phe Leu Thr Trp Ala Gln Val Gly Gly Ser Asp

405 410 415

Tyr Lys Asp Asp Asp Asp Lys

420

<210> 102

<211> 430

<212> PRT

<213> Artificial Sequence

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

polypeptide

<400> 102

Met Asp Asp Pro Ser Phe Leu Thr Gly Arg Ser Thr Tyr Ala Lys Arg

1 5 10 15

Arg Arg Ala Arg Arg Met Asn Val Cys Lys Cys Gly Ala Ile Leu His

20 25 30

Asn Asn Lys Asp Cys Arg Ser Ser Thr Ile Ser Gly His Lys Leu Asp

35 40 45

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

50 55 60

Pro Val Tyr Arg Thr Trp Val Lys Trp Val Glu Thr Glu Tyr His Ile

65 70 75 80

Asn Ile Leu Glu Thr Ser Asp Asp Glu Glu Gly Gly Ser Gly Gly Ser

85 90 95

Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Met Ala Val Pro

100 105 110

Glu Thr Arg Pro Asn His Thr Ile Tyr Ile Asn Asn Leu Asn Ser Lys

115 120 125

Ile Lys Lys Asp Glu Leu Lys Lys Ser Leu Tyr Ala Ile Phe Ser Gln

130 135 140

Phe Gly Gln Ile Leu Asp Ile Leu Val Pro Arg Gln Arg Thr Pro Arg
145 150 155 160
Gly Gln Ala Phe Val Ile Phe Lys Glu Val Ser Ser Ala Thr Asn Ala

165 170 175
Leu Arg Ser Met Gln Gly Phe Pro Phe Tyr Asp Lys Pro Met Arg Ile
180 185 190
Gln Tyr Ala Lys Thr Asp Lys Arg Ile Pro Ala Lys Met Lys Gly Thr
195 200 205
Phe Val Gly Ser Leu Gln Leu Pro Pro Leu Glu Arg Leu Thr Leu Ser
210 215 220
Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Met

225 230 235 240
Glu Leu Glu Ile Arg Pro Leu Phe Leu Val Pro Asp Thr Asn Gly Phe
245 250 255
Ile Asp His Leu Ala Ser Leu Ala Arg Leu Leu Glu Ser Arg Lys Tyr
260 265 270
Ile Leu Val Val Pro Leu Ile Val Ile Asn Glu Leu Asp Gly Leu Ala
275 280 285
Lys Gly Gln Glu Thr Asp His Arg Ala Gly Gly Tyr Ala Arg Val Val

290 295 300
Gln Glu Lys Ala Arg Lys Ser Ile Glu Phe Leu Glu Gln Arg Phe Glu
305 310 315 320
Ser Arg Asp Ser Cys Leu Arg Ala Leu Thr Ser Arg Gly Asn Glu Leu
325 330 335
Glu Ser Ile Ala Phe Arg Ser Glu Asp Ile Thr Gly Gln Leu Gly Asn
340 345 350
Asn Asp Asp Leu Ile Leu Ser Cys Cys Leu His Tyr Cys Lys Asp Lys

355 360 365
Ala Lys Asp Phe Met Pro Ala Ser Lys Glu Glu Pro Ile Arg Leu Leu
370 375 380
Arg Glu Val Val Leu Leu Thr Asp Asp Arg Asn Leu Arg Val Lys Ala

385 390 395 400

Leu Thr Arg Asn Val Pro Val Arg Asp Ile Pro Ala Phe Leu Thr Trp

405 410 415

Ala Gln Val Gly Gly Ser Asp Tyr Lys Asp Asp Asp Asp Lys

420 425 430

<210> 103

<211> 496

<212> PRT

<213> Artificial Sequence

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

<400> 103

Met Asp Asp Pro Ser Phe Leu Thr Gly Arg Ser Thr Tyr Ala Lys Arg

1 5 10 15

Arg Arg Ala Arg Arg Met Asn Val Cys Lys Cys Gly Ala Ile Leu His

20 25 30

Asn Asn Lys Asp Cys Arg Ser Ser Thr Ile Ser Gly His Lys Leu Asp

35 40 45

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

50 55 60

Pro Val Tyr Arg Thr Trp Val Lys Trp Val Glu Thr Glu Tyr His Ile

65 70 75 80

Asn Ile Leu Glu Thr Ser Asp Asp Glu Glu Gly Gly Ser Gly Gly Ser

85 90 95

Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Met Ala Val Pro

100 105 110

Glu Thr Arg Pro Asn His Thr Ile Tyr Ile Asn Asn Leu Asn Ser Lys

115 120 125

Ile Lys Lys Asp Glu Leu Lys Lys Ser Leu Tyr Ala Ile Phe Ser Gln

130 135 140

Phe Gly Gln Ile Leu Asp Ile Leu Val Pro Arg Gln Arg Thr Pro Arg

145 150 155 160

Gly Gln Ala Phe Val Ile Phe Lys Glu Val Ser Ser Ala Thr Asn Ala

165 170 175

Leu Arg Ser Met Gln Gly Phe Pro Phe Tyr Asp Lys Pro Met Arg Ile

180 185 190

Gln Tyr Ala Lys Thr Asp Lys Arg Ile Pro Ala Lys Met Lys Gly Thr

195 200 205

Phe Val Gly Ser Leu Gln Leu Pro Pro Leu Glu Arg Leu Thr Leu Gly

210 215 220

Gly Ser Gly Glu Glu Glu Glu Lys Lys Lys Gln Gln Glu Glu Glu Ala

225 230 235 240

Glu Arg Leu Arg Arg Ile Gln Glu Glu Met Glu Lys Glu Arg Lys Arg

245 250 255

Arg Glu Glu Asp Glu Gln Arg Arg Arg Lys Glu Glu Glu Glu Arg Arg

260 265 270

Met Lys Leu Glu Met Glu Ala Lys Arg Lys Gln Glu Glu Glu Glu Arg

275 280 285

Lys Lys Arg Glu Asp Asp Glu Lys Arg Lys Lys Lys Gly Ser Gly Gly

290 295 300

Ser Met Glu Leu Glu Ile Arg Pro Leu Phe Leu Val Pro Asp Thr Asn

305 310 315 320

Gly Phe Ile Asp His Leu Ala Ser Leu Ala Arg Leu Leu Glu Ser Arg

325 330 335

Lys Tyr Ile Leu Val Val Pro Leu Ile Val Ile Asn Glu Leu Asp Gly

340 345 350

Leu Ala Lys Gly Gln Glu Thr Asp His Arg Ala Gly Gly Tyr Ala Arg

355 360 365

Val Val Gln Glu Lys Ala Arg Lys Ser Ile Glu Phe Leu Glu Gln Arg

370 375 380

Phe Glu Ser Arg Asp Ser Cys Leu Arg Ala Leu Thr Ser Arg Gly Asn

385 390 395 400

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

405 410 415

Gly Asn Asn Asp Asp Leu Ile Leu Ser Cys Cys Leu His Tyr Cys Lys

420 425 430

Asp Lys Ala Lys Asp Phe Met Pro Ala Ser Lys Glu Glu Pro Ile Arg

435 440 445

Leu Leu Arg Glu Val Val Leu Leu Thr Asp Asp Arg Asn Leu Arg Val

450 455 460

Lys Ala Leu Thr Arg Asn Val Pro Val Arg Asp Ile Pro Ala Phe Leu

465 470 475 480

Thr Trp Ala Gln Val Gly Gly Ser Asp Tyr Lys Asp Asp Asp Asp Lys

485 490 495

<210> 104

<211> 339

<212> PRT

<213> Artificial Sequence

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

<400> 104

Met Asp Asp Pro Ser Phe Leu Thr Gly Arg Ser Thr Tyr Ala Lys Arg

1 5 10 15

Arg Arg Ala Arg Arg Met Asn Val Cys Lys Cys Gly Ala Ile Leu His

20 25 30

Asn Asn Lys Asp Cys Arg Ser Ser Thr Ile Ser Gly His Lys Leu Asp

35 40 45

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

50 55 60

Pro Val Tyr Arg Thr Trp Val Lys Trp Val Glu Thr Glu Tyr His Ile

65 70 75 80

Asn Ile Leu Glu Thr Ser Asp Asp Glu Glu Gly Gly Ser Gly Gly Ser

85 90 95

Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Met Ala Val Pro

100 105 110

Glu Thr Arg Pro Asn His Thr Ile Tyr Ile Asn Asn Leu Asn Ser Lys

115 120 125

Ile Lys Lys Asp Glu Leu Lys Lys Ser Leu Tyr Ala Ile Phe Ser Gln

130 135 140

Phe Gly Gln Ile Leu Asp Ile Leu Val Pro Arg Gln Arg Thr Pro Arg

145 150 155 160

Gly Gln Ala Phe Val Ile Phe Lys Glu Val Ser Ser Ala Thr Asn Ala

165 170 175

Leu Arg Ser Met Gln Gly Phe Pro Phe Tyr Asp Lys Pro Met Arg Ile

180 185 190

Gln Tyr Ala Lys Thr Asp Lys Arg Ile Pro Ala Lys Met Lys Gly Thr

195 200 205

Phe Val Gly Ser Leu Gln Leu Pro Pro Leu Glu Arg Leu Thr Leu Ser

210 215 220

Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Asp

225 230 235 240

Ala Met Phe Gly Gln Pro Gly Ala Leu Gly Ser Thr Pro Phe Leu Gly

245 250 255

Gln His Gly Phe Asn Phe Phe Pro Ser Gly Ile Asp Phe Ser Ala Trp

260 265 270

Gly Asn Asn Ser Ser Gln Gly Gln Ser Thr Gln Ser Ser Gly Tyr Ser

275 280 285

Ser Asn Tyr Ala Tyr Ala Pro Ser Ser Leu Gly Gly Ala Met Ile Asp

290 295 300

Gly Gln Ser Ala Phe Ala Asn Glu Thr Leu Asn Lys Ala Pro Gly Met

305 310 315 320

Asn Thr Ile Asp Gln Gly Met Ala Ala Gly Ser Asp Tyr Lys Asp Asp

325 330 335

Asp Asp Lys

<210> 105

<211> 405

<212> PRT

<213> Artificial Sequence

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

<400> 105

Met Asp Asp Pro Ser Phe Leu Thr Gly Arg Ser Thr Tyr Ala Lys Arg
1 5 10 15

Arg Arg Ala Arg Arg Met Asn Val Cys Lys Cys Gly Ala Ile Leu His
20 25 30

Asn Asn Lys Asp Cys Arg Ser Ser Thr Ile Ser Gly His Lys Leu Asp
35 40 45

Arg Leu Arg Phe Val Lys Glu Gly Arg Val Ala Leu Glu Gly Glu Thr
50 55 60

Pro Val Tyr Arg Thr Trp Val Lys Trp Val Glu Thr Glu Tyr His Ile
65 70 75 80

Asn Ile Leu Glu Thr Ser Asp Asp Glu Glu Gly Gly Ser Gly Gly Ser
85 90 95

Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Met Ala Val Pro
100 105 110

Glu Thr Arg Pro Asn His Thr Ile Tyr Ile Asn Asn Leu Asn Ser Lys
115 120 125

Ile Lys Lys Asp Glu Leu Lys Lys Ser Leu Tyr Ala Ile Phe Ser Gln
130 135 140

Phe Gly Gln Ile Leu Asp Ile Leu Val Pro Arg Gln Arg Thr Pro Arg
145 150 155 160

Gly Gln Ala Phe Val Ile Phe Lys Glu Val Ser Ser Ala Thr Asn Ala
165 170 175

Leu Arg Ser Met Gln Gly Phe Pro Phe Tyr Asp Lys Pro Met Arg Ile
180 185 190

Gln Tyr Ala Lys Thr Asp Lys Arg Ile Pro Ala Lys Met Lys Gly Thr

195 200 205

Phe Val Gly Ser Leu Gln Leu Pro Pro Leu Glu Arg Leu Thr Leu Gly

210 215 220

Gly Ser Gly Glu Glu Glu Glu Lys Lys Lys Gln Gln Glu Glu Glu Ala

225 230 235 240

Glu Arg Leu Arg Arg Ile Gln Glu Glu Met Glu Lys Glu Arg Lys Arg

245 250 255

Arg Glu Glu Asp Glu Gln Arg Arg Arg Lys Glu Glu Glu Glu Arg Arg

260 265 270

Met Lys Leu Glu Met Glu Ala Lys Arg Lys Gln Glu Glu Glu Glu Arg

275 280 285

Lys Lys Arg Glu Asp Asp Glu Lys Arg Lys Lys Lys Gly Ser Gly Gly

290 295 300

Ser Asp Ala Met Phe Gly Gln Pro Gly Ala Leu Gly Ser Thr Pro Phe

305 310 315 320

Leu Gly Gln His Gly Phe Asn Phe Phe Pro Ser Gly Ile Asp Phe Ser

325 330 335

Ala Trp Gly Asn Asn Ser Ser Gln Gly Gln Ser Thr Gln Ser Ser Gly

340 345 350

Tyr Ser Ser Asn Tyr Ala Tyr Ala Pro Ser Ser Leu Gly Gly Ala Met

355 360 365

Ile Asp Gly Gln Ser Ala Phe Ala Asn Glu Thr Leu Asn Lys Ala Pro

370 375 380

Gly Met Asn Thr Ile Asp Gln Gly Met Ala Ala Gly Ser Asp Tyr Lys

385 390 395 400

Asp Asp Asp Asp Lys

405

<210> 106

<211> 431

<212> PRT

<213> Artificial Sequence

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

polypeptide

<400> 106

Met Asp Asp Pro Ser Phe Leu Thr Gly Arg Ser Thr Tyr Ala Lys Arg
1 5 10 15

Arg Arg Ala Arg Arg Met Asn Val Cys Lys Cys Gly Ala Ile Leu His
20 25 30

Asn Asn Lys Asp Cys Arg Ser Ser Thr Ile Ser Gly His Lys Leu Asp
35 40 45

Arg Leu Arg Phe Val Lys Glu Gly Arg Val Ala Leu Glu Gly Glu Thr
50 55 60

Pro Val Tyr Arg Thr Trp Val Lys Trp Val Glu Thr Glu Tyr His Ile
65 70 75 80

Asn Ile Leu Glu Thr Ser Asp Asp Glu Glu Gly Gly Ser Gly Gly Ser
85 90 95

Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Met Ala Val Pro
100 105 110

Glu Thr Arg Pro Asn His Thr Ile Tyr Ile Asn Asn Leu Asn Ser Lys
115 120 125

Ile Lys Lys Asp Glu Leu Lys Lys Ser Leu Tyr Ala Ile Phe Ser Gln
130 135 140

Phe Gly Gln Ile Leu Asp Ile Leu Val Pro Arg Gln Arg Thr Pro Arg
145 150 155 160

Gly Gln Ala Phe Val Ile Phe Lys Glu Val Ser Ser Ala Thr Asn Ala
165 170 175

Leu Arg Ser Met Gln Gly Phe Pro Phe Tyr Asp Lys Pro Met Arg Ile
180 185 190

Gln Tyr Ala Lys Thr Asp Lys Arg Ile Pro Ala Lys Met Lys Gly Thr
195 200 205

Phe Val Gly Ser Leu Gln Leu Pro Pro Leu Glu Arg Leu Thr Leu Gly
210 215 220

Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly

225 230 235 240

Ser Met Glu Leu Glu Ile Arg Pro Leu Phe Leu Val Pro Asp Thr Asn

245 250 255

Gly Phe Ile Asp His Leu Ala Ser Leu Ala Arg Leu Leu Glu Ser Arg

260 265 270

Lys Tyr Ile Leu Val Val Pro Leu Ile Val Ile Asn Glu Leu Asp Gly

275 280 285

Leu Ala Lys Gly Gln Glu Thr Asp His Arg Ala Gly Gly Tyr Ala Arg

290 295 300

Val Val Gln Glu Lys Ala Arg Lys Ser Ile Glu Phe Leu Glu Gln Arg

305 310 315 320

Phe Glu Ser Arg Asp Ser Cys Leu Arg Ala Leu Thr Ser Arg Gly Asn

325 330 335

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

340 345 350

Gly Asn Asn Asp Asp Leu Ile Leu Ser Cys Cys Leu His Tyr Cys Lys

355 360 365

Asp Lys Ala Lys Asp Phe Met Pro Ala Ser Lys Glu Glu Pro Ile Arg

370 375 380

Leu Leu Arg Glu Val Val Leu Leu Thr Asp Asp Arg Asn Leu Arg Val

385 390 395 400

Lys Ala Leu Thr Arg Asn Val Pro Val Arg Asp Ile Pro Ala Phe Leu

405 410 415

Thr Trp Ala Gln Val Gly Ser Ser Gly His His His His His His

420 425 430

<210> 107

<211> 21

<212> RNA

<213> Unknown

<220><223> Description of Unknown:

RNA hairpin binding domain

<400> 107

aaacaugagg auuacccaug u 21

<210> 108

<211> 21

<212> RNA

<213> Unknown

<220><223> Description of Unknown:

RNA hairpin binding domain

<400> 108

aaacaugagg aucacccaug u 21

<210> 109

<211> 15

<212> RNA

<213> Unknown

<220><223> Description of Unknown:

RNA hairpin binding domain

<400> 109

gcccugaaga agggc 15

<210> 110

<211> 15

<212> RNA

<213> Unknown

<220><223> Description of Unknown:

RNA hairpin binding domain

<400> 110

gcccugaaaa agggc 15

<210> 111

<211> 25

<212> RNA

<213> Pseudomonas aeroginosa

<400> 111

uaaggaguuu auauggaaac ccuua 25

<210> 112

<211> 28

<212> RNA

<213> Bovine immunodeficiency virus

<400> 112

ggcucgugua gcucauuagc uccgagcc 28

<210> 113

<211> 29

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<220><223> Description of Combined DNA/RNA Molecule: Synthetic

oligonucleotide

<400> 113

ggaauccauu gcacuccgga uuucactag 29

<210> 114

<211> 20

<212> RNA

<213> Unknown

<220><223> Description of Unknown:

RNA hairpin binding domain

<400> 114

augcaugucu aagacagcau 20

<210> 115

<211> 23

<212> RNA

<213> Unknown

<220><223> Description of Unknown:

RNA hairpin binding domain

<400> 115

aaaacauaag gaaaaccuau guu 23

<210> 116

<211> 20

<212> DNA

<213> Artificial Sequence

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

<400> 116

cttgtggtag ttggagctga 20

<210> 117

<211> 20

<212> DNA

<213> Artificial Sequence

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

<400> 117

agttggagct gatggcgtag 20

<210> 118

<211> 20

<212> DNA

<213> Artificial Sequence

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

<400> 118

gttggagctg atggcgtagg 20

<210> 119

<211> 20

<212> DNA

<213> Artificial Sequence

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

<400> 119

tggcgtaggc aagagtgcct 20

<210> 120

<211> 20

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 120

ggcgtaggca agagtgcctt 20

<210> 121

<211> 25

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 121

taaggagttt atatggaaac cctta 25

<210> 122

<211> 20

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 122

gtgataagtg gaatgccatg 20

<210> 123

<211> 403

<212> PRT

<213> Artificial Sequence

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

polypeptide

<400> 123

Leu His Leu Asp Gln Thr Pro Ser Arg Gln Pro Ile Pro Ser Glu Gly

1 5 10 15

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

20 25 30

Val Leu Gly Lys Phe Gly Asp Leu Thr Asp Asn Phe Ser Ser Pro His

35 40 45

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

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

Lys Ala Arg Gly Gln Leu Arg Thr Lys Ile Glu Ser Gly Glu Gly Thr
180 185 190
Ile Pro Val Arg Ser Asn Ala Ser Ile Gln Thr Trp Asp Gly Val Leu
195 200 205
Gln Gly Glu Arg Leu Leu Thr Met Ser Cys Ser Asp Lys Ile Ala Arg
210 215 220
Trp Asn Val Val Gly Ile Gln Gly Ser Leu Leu Ser Ile Phe Val Glu
225 230 235 240

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

Ala Glu Ala Arg Gln Pro Gly Lys Ala Pro Asn Phe Ser Val Asn Trp
290 295 300

Thr Val Gly Asp Ser Ala Ile Glu Val Ile Asn Ala Thr Thr Gly Lys
305 310 315 320
Asp Glu Leu Gly Arg Ala Ser Arg Leu Cys Lys His Ala Leu Tyr Cys
325 330 335
Arg Trp Met Arg Val His Gly Lys Val Pro Ser His Leu Leu Arg Ser
340 345 350
Lys Ile Thr Lys Pro Asn Val Tyr His Glu Ser Lys Leu Ala Ala Lys
355 360 365

Glu Tyr Gln Ala Ala Lys Ala Arg Leu Phe Thr Ala Phe Ile Lys Ala
370 375 380
Gly Leu Gly Ala Trp Val Glu Lys Pro Thr Glu Gln Asp Gln Phe Ser
385 390 395 400
Leu Thr Pro

<210> 124
<211> 13
<212> PRT
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
peptide
<400> 124
Gly Ser Leu Gln Leu Pro Pro Leu Glu Arg Leu Thr Leu
1 5 10

<210> 125
<211> 267
<212> PRT
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
polypeptide
<400> 125

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

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

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

Leu Thr Ala Ser Leu Arg Gln Asn Gly Ala Lys Thr Ala Tyr Arg Val
180 185 190
Asn Leu Lys Leu Asp Gln Ala Asp Val Val Asp Ala Ser Thr Ser Val
195 200 205
Ala Gly Glu Leu Pro Lys Val Arg Tyr Thr Gln Val Trp Ser His Asp
210 215 220
Val Thr Ile Val Ala Asn Ser Thr Glu Ala Ser Arg Lys Ser Leu Tyr
225 230 235 240

Asp Leu Thr Lys Ser Leu Val Ala Thr Ser Gln Val Glu Asp Leu Val

245 250 255

Val Asn Leu Val Pro Leu Gly Arg Ser Leu Glu

260 265

<210> 126

<211> 20

<212> DNA

<213> Artificial Sequence

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

primer

<400> 126

gaagaaggaa ggagcgctaa 20

<210> 127

<211> 20

<212> DNA

<213> Artificial Sequence

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

primer

<400> 127

cctgctacct tcatcaccaa 20

<210> 128

<211> 613

<212> PRT

<213> Artificial Sequence

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

polypeptide

<400> 128

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

1 5 10 15

Gly Arg Cys Ala Val Leu Ser Cys Leu Pro Lys Glu Glu Gln Ile Gly

20 25 30

Lys Cys Ser Thr Arg Gly Arg Lys Cys Cys Arg Arg Lys Lys Gly Gly

35 40 45

Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser

50 55 60

Met Ala Val Pro Glu Thr Arg Pro Asn His Thr Ile Tyr Ile Asn Asn

65 70 75 80

Leu Asn Ser Lys Ile Lys Lys Asp Glu Leu Lys Lys Ser Leu Tyr Ala

85 90 95

Ile Phe Ser Gln Phe Gly Gln Ile Leu Asp Ile Leu Val Pro Arg Gln

100 105 110

Arg Thr Pro Arg Gly Gln Ala Phe Val Ile Phe Lys Glu Val Ser Ser

115 120 125

Ala Thr Asn Ala Leu Arg Ser Met Gln Gly Phe Pro Phe Tyr Asp Lys

130 135 140

Pro Met Arg Ile Gln Tyr Ala Lys Thr Asp Lys Arg Ile Pro Ala Lys

145 150 155 160

Met Lys Gly Thr Phe Val Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly

165 170 175

Gly Ser Gly Gly Ser Gly Gly Ser Leu His Leu Asp Gln Thr Pro Ser

180 185 190

Arg Gln Pro Ile Pro Ser Glu Gly Leu Gln Leu His Leu Pro Gln Val

195 200 205

Leu Ala Asp Ala Val Ser Arg Leu Val Leu Gly Lys Phe Gly Asp Leu

210 215 220

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

225 230 235 240

Val Val Met Thr Thr Gly Thr Asp Val Lys Asp Ala Lys Val Ile Ser

245 250 255

Val Ser Thr Gly Thr Lys Cys Ile Asn Gly Glu Tyr Met Ser Asp Arg

260 265 270

Gly Leu Ala Leu Asn Asp Cys His Ala Glu Ile Ile Ser Arg Arg Ser

275 280 285

Leu Leu Arg Phe Leu Tyr Thr Gln Leu Glu Leu Tyr Leu Asn Asn Lys

290 295 300

Asp Asp Gln Lys Arg Ser Ile Phe Gln Lys Ser Glu Arg Gly Gly Phe
305 310 315 320
Arg Leu Lys Glu Asn Val Gln Phe His Leu Tyr Ile Ser Thr Ser Pro
325 330 335
Cys Gly Asp Ala Arg Ile Phe Ser Pro His Glu Pro Ile Leu Glu Glu

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

405 410 415
Ser Leu Leu Ser Ile Phe Val Glu Pro Ile Tyr Phe Ser Ser Ile Ile
420 425 430
Leu Gly Ser Leu Tyr His Gly Asp His Leu Ser Arg Ala Met Tyr Gln
435 440 445
Arg Ile Ser Asn Ile Glu Asp Leu Pro Pro Leu Tyr Thr Leu Asn Lys
450 455 460
Pro Leu Leu Ser Gly Ile Ser Asn Ala Glu Ala Arg Gln Pro Gly Lys

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

530 535 540
His Glu Ser Lys Leu Ala Ala Lys Glu Tyr Gln Ala Ala Lys Ala Arg

545 550 555 560

Leu Phe Thr Ala Phe Ile Lys Ala Gly Leu Gly Ala Trp Val Glu Lys

565 570 575

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

580 585 590

Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys Lys Gly Ser Asp Tyr Lys

595 600 605

Asp Asp Asp Asp Lys

610

<210> 129

<211> 613

<212> PRT

<213> Artificial Sequence

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

<400> 129

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

1 5 10 15

Gly Arg Cys Ala Val Leu Ser Cys Leu Pro Lys Glu Glu Gln Ile Gly

20 25 30

Lys Cys Ser Thr Arg Gly Arg Lys Cys Cys Arg Arg Lys Lys Gly Gly

35 40 45

Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser

50 55 60

Met Ala Val Pro Glu Thr Arg Pro Asn His Thr Ile Tyr Ile Asn Asn

65 70 75 80

Leu Asn Ser Lys Ile Lys Lys Asp Glu Leu Lys Lys Ser Leu Tyr Ala

85 90 95

Ile Phe Ser Gln Phe Gly Gln Ile Leu Asp Ile Leu Val Pro Arg Gln

100 105 110

Arg Thr Pro Arg Gly Gln Ala Phe Val Ile Phe Lys Glu Val Ser Ser

115 120 125

Ala Thr Asn Ala Leu Arg Ser Met Gln Gly Phe Pro Phe Tyr Asp Lys
130 135 140
Pro Met Arg Ile Gln Tyr Ala Lys Thr Asp Lys Arg Ile Pro Ala Lys
145 150 155 160

Met Lys Gly Thr Phe Val Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly
165 170 175
Gly Ser Gly Gly Ser Gly Gly Ser Leu His Leu Asp Gln Thr Pro Ser
180 185 190
Arg Gln Pro Ile Pro Ser Glu Gly Leu Gln Leu His Leu Pro Gln Val
195 200 205
Leu Ala Asp Ala Val Ser Arg Leu Val Leu Gly Lys Phe Gly Asp Leu
210 215 220

Thr Asp Asn Phe Ser Ser Pro His Ala Arg Arg Lys Val Leu Ala Gly
225 230 235 240
Val Val Met Thr Thr Gly Thr Asp Val Lys Asp Ala Lys Val Ile Ser
245 250 255
Val Ser Thr Gly Thr Lys Cys Ile Asn Gly Glu Tyr Met Ser Asp Arg
260 265 270
Gly Leu Ala Leu Asn Asp Cys His Ala Glu Ile Ile Ser Arg Arg Ser
275 280 285

Leu Leu Arg Phe Leu Tyr Thr Gln Leu Glu Leu Tyr Leu Asn Asn Lys
290 295 300
Asp Asp Gln Lys Arg Ser Ile Phe Gln Lys Ser Glu Arg Gly Gly Phe
305 310 315 320
Arg Leu Lys Glu Asn Val Gln Phe His Leu Tyr Ile Ser Thr Ser Pro
325 330 335
Cys Gly Asp Ala Arg Ile Phe Ser Pro His Glu Pro Ile Leu Glu Glu
340 345 350

Pro Ala Asp Arg His Pro Asn Arg Lys Ala Arg Gly Gln Leu Arg Thr
355 360 365
Lys Ile Glu Ser Gly Gln Gly Thr Ile Pro Val Arg Ser Asn Ala Ser

370 375 380

Ile Gln Thr Trp Asp Gly Val Leu Gln Gly Glu Arg Leu Leu Thr Met

385 390 395 400

Ser Cys Ser Asp Lys Ile Ala Arg Trp Asn Val Val Gly Ile Gln Gly

405 410 415

Ser Leu Leu Ser Ile Phe Val Glu Pro Ile Tyr Phe Ser Ser Ile Ile

420 425 430

Leu Gly Ser Leu Tyr His Gly Asp His Leu Ser Arg Ala Met Tyr Gln

435 440 445

Arg Ile Ser Asn Ile Glu Asp Leu Pro Pro Leu Tyr Thr Leu Asn Lys

450 455 460

Pro Leu Leu Ser Gly Ile Ser Asn Ala Glu Ala Arg Gln Pro Gly Lys

465 470 475 480

Ala Pro Asn Phe Ser Val Asn Trp Thr Val Gly Asp Ser Ala Ile Glu

485 490 495

Val Ile Asn Ala Thr Thr Gly Lys Asp Glu Leu Gly Arg Ala Ser Arg

500 505 510

Leu Cys Lys His Ala Leu Tyr Cys Arg Trp Met Arg Val His Gly Lys

515 520 525

Val Pro Ser His Leu Leu Arg Ser Lys Ile Thr Lys Pro Asn Val Tyr

530 535 540

His Glu Ser Lys Leu Ala Ala Lys Glu Tyr Gln Ala Ala Lys Ala Arg

545 550 555 560

Leu Phe Thr Ala Phe Ile Lys Ala Gly Leu Gly Ala Trp Val Glu Lys

565 570 575

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

580 585 590

Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys Lys Gly Ser Asp Tyr Lys

595 600 605

Asp Asp Asp Asp Lys

610

<210> 130

<211> 571

<212> PRT

<213> Artificial Sequence

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

<400> 130

Met Gly Ile Ile Asn Thr Leu Gln Lys Tyr Tyr Cys Arg Val Arg Gly
1 5 10 15

Gly Arg Cys Ala Val Leu Ser Cys Leu Pro Lys Glu Glu Gln Ile Gly
20 25 30

Lys Cys Ser Thr Arg Gly Arg Lys Cys Cys Arg Arg Lys Lys Gly Gly
35 40 45

Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser
50 55 60

Met Ala Val Pro Glu Thr Arg Pro Asn His Thr Ile Tyr Ile Asn Asn
65 70 75 80

Leu Asn Ser Lys Ile Lys Lys Asp Glu Leu Lys Lys Ser Leu Tyr Ala
85 90 95

Ile Phe Ser Gln Phe Gly Gln Ile Leu Asp Ile Leu Val Pro Arg Gln
100 105 110

Arg Thr Pro Arg Gly Gln Ala Phe Val Ile Phe Lys Glu Val Ser Ser
115 120 125

Ala Thr Asn Ala Leu Arg Ser Met Gln Gly Phe Pro Phe Tyr Asp Lys
130 135 140

Pro Met Arg Ile Gln Tyr Ala Lys Thr Asp Lys Arg Ile Pro Ala Lys
145 150 155 160

Met Lys Gly Thr Phe Val Gly Ser Leu Gln Leu Pro Pro Leu Glu Arg
165 170 175

Leu Thr Leu Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly
180 185 190

Gly Ser Gly Gly Ser Met Ser Ala Thr Ser Val Asp Thr Gln Arg Thr

195 200 205

Lys Gly Gln Asp Asn Lys Val Gln Asn Gly Ser Leu His Gln Lys Asp

210 215 220

Thr Val His Asp Asn Asp Phe Glu Pro Tyr Leu Thr Gly Gln Ser Asn

225 230 235 240

Gln Ser Asn Ser Tyr Pro Ser Met Ser Asp Pro Tyr Leu Ser Ser Tyr

245 250 255

Tyr Pro Pro Ser Ile Gly Phe Pro Tyr Ser Leu Asn Glu Ala Pro Trp

260 265 270

Ser Thr Ala Gly Asp Pro Pro Ile Pro Tyr Leu Thr Thr Tyr Gly Gln

275 280 285

Leu Ser Asn Gly Asp His His Phe Met His Asp Ala Val Phe Gly Gln

290 295 300

Pro Gly Gly Leu Gly Asn Asn Ile Tyr Gln His Arg Phe Asn Phe Phe

305 310 315 320

Pro Glu Asn Pro Ala Phe Ser Ala Trp Gly Thr Ser Gly Ser Gln Gly

325 330 335

Gln Gln Thr Gln Ser Ser Ala Tyr Gly Ser Ser Tyr Thr Tyr Pro Pro

340 345 350

Ser Ser Leu Gly Gly Thr Val Val Asp Gly Gln Pro Gly Phe His Ser

355 360 365

Asp Thr Leu Ser Lys Ala Pro Gly Met Asn Ser Leu Glu Gln Gly Met

370 375 380

Val Gly Leu Lys Ile Gly Asp Val Ser Ser Ser Ala Val Lys Thr Val

385 390 395 400

Gly Ser Val Val Ser Ser Val Ala Leu Thr Gly Val Leu Ser Gly Asn

405 410 415

Gly Gly Thr Asn Val Asn Met Pro Val Ser Lys Pro Thr Ser Trp Ala

420 425 430

Ala Ile Ala Ser Lys Pro Ala Lys Pro Gln Pro Lys Met Lys Thr Lys

435 440 445

Ser Gly Pro Val Met Gly Gly Gly Leu Pro Pro Pro Pro Ile Lys His
450 455 460
Asn Met Asp Ile Gly Thr Trp Asp Asn Lys Gly Pro Val Pro Lys Ala
465 470 475 480
Pro Val Pro Gln Gln Ala Pro Ser Pro Gln Ala Ala Pro Gln Pro Gln

485 490 495
Gln Val Ala Gln Pro Leu Pro Ala Gln Pro Pro Ala Leu Ala Gln Pro
500 505 510
Gln Tyr Gln Ser Pro Gln Gln Pro Pro Gln Thr Arg Trp Val Ala Pro
515 520 525
Arg Asn Arg Asn Ala Ala Phe Gly Gln Ser Gly Gly Ala Gly Ser Asp
530 535 540
Ser Asn Ser Pro Gly Asn Val Gln Pro Asn Ser Ala Pro Ser Val Glu

545 550 555 560
Ser Gly Ser Asp Tyr Lys Asp Asp Asp Asp Lys
565 570
<210> 131
<211> 367
<212> PRT
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
polypeptide
<400> 131
Met Gly Ile Ile Asn Thr Leu Gln Lys Tyr Tyr Cys Arg Val Arg Gly
1 5 10 15
Gly Arg Cys Ala Val Leu Ser Cys Leu Pro Lys Glu Glu Gln Ile Gly

20 25 30
Lys Cys Ser Thr Arg Gly Arg Lys Cys Cys Arg Arg Lys Lys Gly Gly
35 40 45
Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser
50 55 60
Ala Asp Phe Glu Thr Asp Glu Ser Val Leu Met Arg Arg Gln Lys Gln

65 70 75 80

Ile Asn Tyr Gly Lys Asn Thr Ile Ala Tyr Asp Arg Tyr Ile Lys Glu

85 90 95

Val Pro Arg His Leu Arg Gln Pro Gly Ile His Pro Lys Thr Pro Asn

100 105 110

Lys Phe Lys Lys Tyr Ser Arg Arg Ser Trp Asp Gln Gln Ile Lys Leu

115 120 125

Trp Lys Val Ala Leu His Phe Trp Asp Leu Gln Leu Pro Pro Leu Glu

130 135 140

Arg Leu Thr Leu Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser

145 150 155 160

Gly Gly Ser Gly Gly Ser Met Ser Ala Ser Ser Leu Leu Glu Gln Arg

165 170 175

Pro Lys Gly Gln Gly Asn Lys Val Gln Asn Gly Ser Val His Gln Lys

180 185 190

Asp Gly Leu Asn Asp Asp Asp Phe Glu Pro Tyr Leu Ser Pro Gln Ala

195 200 205

Arg Pro Asn Asn Ala Tyr Thr Ala Met Ser Asp Ser Tyr Leu Pro Ser

210 215 220

Tyr Tyr Ser Pro Ser Ile Gly Phe Ser Tyr Ser Leu Gly Glu Ala Ala

225 230 235 240

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

245 250 255

Gln Leu Ser Asn Gly Glu Pro His Phe Leu Pro Asp Ala Met Phe Gly

260 265 270

Gln Pro Gly Ala Leu Gly Ser Thr Pro Phe Leu Gly Gln His Gly Phe

275 280 285

Asn Phe Phe Pro Ser Gly Ile Asp Phe Ser Ala Trp Gly Asn Asn Ser

290 295 300

Ser Gln Gly Gln Ser Thr Gln Ser Ser Gly Tyr Ser Ser Asn Tyr Ala

305 310 315 320

Tyr Ala Pro Ser Ser Leu Gly Gly Ala Met Ile Asp Gly Gln Ser Ala
325 330 335
Phe Ala Asn Glu Thr Leu Asn Lys Ala Pro Gly Met Asn Thr Ile Asp

340 345 350
Gln Gly Met Ala Ala Gly Ser Asp Tyr Lys Asp Asp Asp Asp Lys
355 360 365
<210> 132
<211> 504
<212> PRT
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
polypeptide
<400> 132
Met Asp Asp Pro Ser Phe Leu Thr Gly Arg Ser Thr Tyr Ala Lys Arg
1 5 10 15
Arg Arg Ala Arg Arg Met Asn Val Cys Lys Cys Gly Ala Ile Leu His

20 25 30
Asn Asn Lys Asp Cys Arg Ser Ser Thr Ile Ser Gly His Lys Leu Asp
35 40 45
Arg Leu Arg Phe Val Lys Glu Gly Arg Val Ala Leu Glu Gly Glu Thr
50 55 60
Pro Val Tyr Arg Thr Trp Val Lys Trp Val Glu Thr Glu Tyr His Ile
65 70 75 80
Asn Ile Leu Glu Thr Ser Asp Asp Glu Glu Gly Gly Ser Gly Gly Ser

85 90 95
Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Met Ala Lys Thr
100 105 110
Ile Val Leu Ala Val Gly Glu Ala Thr Arg Thr Leu Thr Glu Ile Gln
115 120 125
Ser Thr Ala Asp Arg Gln Ile Phe Glu Glu Lys Val Gly Pro Leu Val
130 135 140

Gly Arg Leu Arg Leu Thr Ala Ser Leu Arg Gln Asn Gly Ala Lys Thr

145 150 155 160

Ala Tyr Arg Val Asn Leu Lys Leu Asp Gln Ala Asp Val Val Asp Ala

165 170 175

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

180 185 190

Trp Ser His Asp Val Thr Ile Val Ala Asn Ser Thr Glu Ala Ser Arg

195 200 205

Lys Ser Leu Tyr Asp Leu Thr Lys Ser Leu Val Ala Thr Ser Gln Val

210 215 220

Glu Asp Leu Val Val Asn Leu Val Pro Leu Gly Arg Ser Leu Glu Gly

225 230 235 240

Gly Ser Gly Gly Met Ala Lys Thr Ile Val Leu Ala Val Gly Glu Ala

245 250 255

Thr Arg Thr Leu Thr Glu Ile Gln Ser Thr Ala Asp Arg Gln Ile Phe

260 265 270

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

275 280 285

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

290 295 300

Asp Gln Ala Asp Val Val Asp Ala Ser Thr Ser Val Ala Gly Glu Leu

305 310 315 320

Pro Lys Val Arg Tyr Thr Gln Val Trp Ser His Asp Val Thr Ile Val

325 330 335

Ala Asn Ser Thr Glu Ala Ser Arg Lys Ser Leu Tyr Asp Leu Thr Lys

340 345 350

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

355 360 365

Pro Leu Gly Arg Ser Leu Glu Leu Gln Leu Pro Pro Leu Glu Arg Leu

370 375 380

Thr Leu Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly

385 390 395 400

Ser Gly Gly Ser Asp Ala Met Phe Gly Gln Pro Gly Ala Leu Gly Ser

405 410 415

Thr Pro Phe Leu Gly Gln His Gly Phe Asn Phe Phe Pro Ser Gly Ile

420 425 430

Asp Phe Ser Ala Trp Gly Asn Asn Ser Ser Gln Gly Gln Ser Thr Gln

435 440 445

Ser Ser Gly Tyr Ser Ser Asn Tyr Ala Tyr Ala Pro Ser Ser Leu Gly

450 455 460

Gly Ala Met Ile Asp Gly Gln Ser Ala Phe Ala Asn Glu Thr Leu Asn

465 470 475 480

Lys Ala Pro Gly Met Asn Thr Ile Asp Gln Gly Met Ala Ala Gly Ser

485 490 495

Asp Tyr Lys Asp Asp Asp Asp Lys

500

<210> 133

<211> 16

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 133

Ser Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser

1 5 10 15

<210> 134

<211> 73

<212> PRT

<213> Artificial Sequence

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

polypeptide

<400> 134

Glu Glu Glu Glu Lys Lys Lys Gln Gln Glu Glu Glu Ala Glu Arg Leu

1 5 10 15

Arg Arg Ile Gln Glu Glu Met Glu Lys Glu Arg Lys Arg Arg Glu Glu

20 25 30

Asp Glu Lys Arg Arg Arg Lys Glu Glu Glu Glu Arg Arg Met Lys Leu

35 40 45

Glu Met Glu Ala Lys Arg Lys Gln Glu Glu Glu Glu Arg Lys Lys Arg

50 55 60

Glu Asp Asp Glu Lys Arg Lys Lys Lys

65 70

<210> 135

<211> 403

<212> PRT

<213> Artificial Sequence

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

polypeptide

<400> 135

Leu His Leu Asp Gln Thr Pro Ser Arg Gln Pro Ile Pro Ser Glu Gly

1 5 10 15

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

20 25 30

Val Leu Gly Lys Phe Gly Asp Leu Thr Asp Asn Phe Ser Ser Pro His

35 40 45

Ala Arg Arg Lys Val Leu Ala Gly Val Val Met Thr Thr Gly Thr Asp

50 55 60

Val Lys Asp Ala Lys Val Ile Ser Val Ser Thr Gly Thr Lys Cys Ile

65 70 75 80

Asn Gly Glu Tyr Met Ser Asp Arg Gly Leu Ala Leu Asn Asp Cys His

85 90 95

Ala Glu Ile Ile Ser Arg Arg Ser Leu Leu Arg Phe Leu Tyr Thr Gln

100 105 110

Leu Glu Leu Tyr Leu Asn Asn Lys Asp Asp Gln Lys Arg Ser Ile Phe

115 120 125

Gln Lys Ser Glu Arg Gly Gly Phe Arg Leu Lys Glu Asn Val Gln Phe

130 135 140

His Leu Tyr Ile Ser Thr Ser Pro Cys Gly Asp Ala Arg Ile Phe Ser

145 150 155 160

Pro His Glu Pro Ile Leu Glu Glu Pro Ala Asp Arg His Pro Asn Arg

165 170 175

Lys Ala Arg Gly Gln Leu Arg Thr Lys Ile Glu Ser Gly Gln Gly Thr

180 185 190

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

195 200 205

Gln Gly Glu Arg Leu Leu Thr Met Ser Cys Ser Asp Lys Ile Ala Arg

210 215 220

Trp Asn Val Val Gly Ile Gln Gly Ser Leu Leu Ser Ile Phe Val Glu

225 230 235 240

Pro Ile Tyr Phe Ser Ser Ile Ile Leu Gly Ser Leu Tyr His Gly Asp

245 250 255

His Leu Ser Arg Ala Met Tyr Gln Arg Ile Ser Asn Ile Glu Asp Leu

260 265 270

Pro Pro Leu Tyr Thr Leu Asn Lys Pro Leu Leu Ser Gly Ile Ser Asn

275 280 285

Ala Glu Ala Arg Gln Pro Gly Lys Ala Pro Asn Phe Ser Val Asn Trp

290 295 300

Thr Val Gly Asp Ser Ala Ile Glu Val Ile Asn Ala Thr Thr Gly Lys

305 310 315 320

Asp Glu Leu Gly Arg Ala Ser Arg Leu Cys Lys His Ala Leu Tyr Cys

325 330 335

Arg Trp Met Arg Val His Gly Lys Val Pro Ser His Leu Leu Arg Ser

340 345 350

Lys Ile Thr Lys Pro Asn Val Tyr His Glu Ser Lys Leu Ala Ala Lys

355 360 365

Glu Tyr Gln Ala Ala Lys Ala Arg Leu Phe Thr Ala Phe Ile Lys Ala

370 375 380

Gly Leu Gly Ala Trp Val Glu Lys Pro Thr Glu Gln Asp Gln Phe Ser

385 390 395 400

Leu Thr Pro

<210> 136

<211> 30

<212> RNA

<213> Unknown

<220><223> Description of Unknown:

RNA hairpin binding domain

<220><221> modified_base

<222> (1)..(5)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (9)..(13)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (19)..(24)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (26)..(30)

<223> a, c, u, g, unknown or other

<400> 136

nnnnnugcnn nnncagugnn nnnncnnnnn 30

<210> 137

<211> 4

<212> PRT

<213> Homo sapiens

<400> 137

Asp Glu Ala His

1

<210> 138

<211> 4

<212> PRT

<213> Homo sapiens

<400> 138

Asp Glu Ala Asp

1

<210> 139

<211> 68

<212> RNA

<213> Artificial Sequence

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

oligonucleotide

<220><221> modified_base

<222> (32)..(68)

<223> a, c, u, g, unknown or other

<400> 139

ggccagaucu gagccuggga gcucucuggc cnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 60

nnnnnnnn 68

<210> 140

<211> 63

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<220><223> Description of Combined DNA/RNA Molecule: Synthetic

oligonucleotide

<220><221> modified_base

<222> (32)..(63)

<223> a, c, t, g, unknown or other

<400> 140

ccaaaggcuc uucucagagc cacccattat tnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 60

nnn 63

<210> 141

<211> 63

<212> RNA

<213> Artificial Sequence

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

oligonucleotide

<220><221> modified_base

<222> (32)..(63)

<223> a, c, u, g, unknown or other

<400> 141

ccaaaggcuc uucucagagc cacccauuau unnnnnnnnn nnnnnnnnnn nnnnnnnnnn 60

nnn 63

<210> 142

<211> 27

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<220><221> modified_base

<222> (13)..(15)

<223> a, c, t, g, unknown or other

<400> 142

gtagttggag ctnnnggcgt aggcaag 27

<210> 143

<211> 20

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 143

tgaacaccat caacctcgac 20

<210> 144

<211> 20

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 144

gaacaccatc aacctcgact 20

<210> 145

<211> 20

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 145

tcaacctcga ctaccgcatc 20

<210> 146

<211> 20

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 146

caacctcgac taccgcatcc 20

<210> 147

<211> 20

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 147

accgcatccg ttctcacgga 20

<210> 148

<211> 20

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 148

ccgcatccgt tctcacggaa 20