



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2021년07월02일

(11) 등록번호 10-2271291

(24) 등록일자 2021년06월24일

(51) 국제특허분류(Int. Cl.)
C12N 9/22 (2006.01) *C12N 15/63* (2006.01)
C12N 15/85 (2006.01) *C12N 9/10* (2006.01)
(52) CPC특허분류
C12N 9/22 (2013.01)
C12N 15/63 (2013.01)
(21) 출원번호 10-2021-7002427(분할)
(22) 출원일자(국제) 2014년03월14일
심사청구일자 2021년01월25일
(85) 번역문제출일자 2021년01월25일
(65) 공개번호 10-2021-0013302
(43) 공개일자 2021년02월03일
(62) 원출원 특허 10-2015-7029170
원출원일자(국제) 2014년03월14일
심사청구일자 2019년03월12일
(86) 국제출원번호 PCT/US2014/027335
(87) 국제공개번호 WO 2014/152432
국제공개일자 2014년09월25일
(30) 우선권주장
61/799,647 2013년03월15일 미국(US)
(뒷면에 계속)
(56) 선행기술조사문헌
Cell, Vol.152, pp.1173-1183 (2013.02.28)
Nature Methods. Vol.10, pp.243-245
(2013.02.10)
US20070020627 A1
US20110236894 A1

(73) 특허권자
더 제너럴 하스피탈 코퍼레이션
미국, 메사추세츠 02114, 보스턴 프룻트 스트리트 55
(72) 발명자
중, 제이. 키쓰
미국 01890 매사추세츠주 윈체스터 매그놀리아 웨이 1
미더, 모르간
미국 02446 매사추세츠주 브루클린 비컨 스트리트 1569 아파트먼트 32
(74) 대리인
양영준, 김영

전체 청구항 수 : 총 12 항

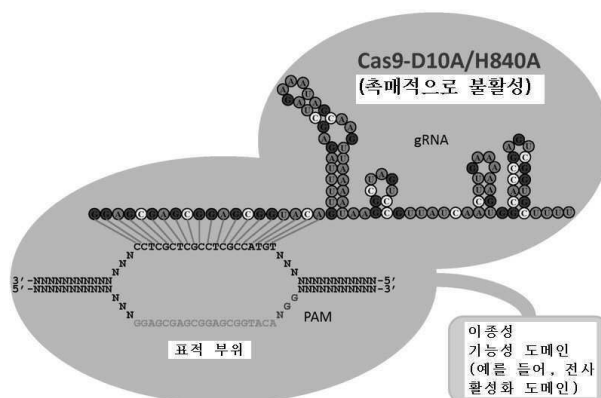
심사관 : 한지혜

(54) 발명의 명칭 특정 게놈 좌위에 대한 유전적 및 후성적 조절 단백질의 RNA-안내 표적화

(57) 요약

특정 게놈 좌위에 전사 활성제와 같은 이중성 기능성 도메인의 RNA-안내 표적화를 위한 방법 및 작제물.

대표도 - 도1c



(52) CPC특허분류

C12N 15/85 (2013.01)
C12N 9/1007 (2013.01)
C12Y 114/11 (2013.01)
C12Y 201/01 (2013.01)
C12Y 301/00 (2013.01)
C12Y 301/21004 (2013.01)
C07K 2319/00 (2013.01)
C12N 2710/00033 (2013.01)
C12N 2770/00033 (2013.01)

(30) 우선권주장

61/838,148	2013년06월21일	미국(US)
61/838,178	2013년06월21일	미국(US)
61/921,007	2013년12월26일	미국(US)

명세서

청구범위

청구항 1

포유동물 세포를 이중성 기능성 도메인에 연결된 촉매적으로 불활성인 스트렙토코커스 피오게네스(*S. pyogenes*) CRISPR 관련 9 (Cas9) 단백질을 포함하는 융합 단백질을 암호화하는 핵산에 작동가능하게 연결된 프로모터를 포함하는 발현 벡터, 및 표적 유전자의 프로모터로 지향된 하나 이상의 안내(guide) RNA를 암호화하는 핵산 서열에 작동가능하게 연결된 프로모터를 포함하는 하나 이상의 발현 벡터와 시험관내(in vitro) 또는 생체외(ex vivo)에서 접촉시키는 것을 포함하고,

여기서 촉매적으로 불활성인 Cas9 단백질은 D10 및 H840에 돌연변이를 포함하고, 이중성 기능성 도메인은 전사 활성화 도메인이며,

접촉된 포유동물 세포는 Cas9 융합 단백질 및 하나 이상의 안내 RNA를 발현하여, 발현된 Cas9 융합 단백질이 발현된 하나 이상의 안내 RNA에 의해 표적 유전자의 프로모터로 지향되며, 전사 활성화 도메인은 증가된 발현을 매개하는 것인, 포유동물 세포에서 표적 유전자의 발현을 증가시키는 방법.

청구항 2

포유동물 세포에서 이중성 기능성 도메인에 연결된 촉매적으로 불활성인 스트렙토코커스 피오게네스(*S. pyogenes*) CRISPR 관련 9 (Cas9) 단백질을 포함하는 융합 단백질을 발현시키고, 표적 유전자로 지향된 하나 이상의 안내 RNA를 발현시키는 것을 포함하며,

여기서 촉매적으로 불활성인 Cas9 단백질은 D10 및 H840에 돌연변이를 포함하고, 이중성 기능성 도메인은 전사 활성화 도메인이며,

발현된 Cas9 융합 단백질은 하나 이상의 안내 RNA에 의해 표적 유전자로 지향되며, 전사 활성화 도메인은 증가된 발현을 매개하는 것인, 시험관내 또는 생체외에서 포유동물 세포 중 표적 유전자의 발현을 증가시키는 방법.

청구항 3

제2항에 있어서, 전사 활성화 도메인이 DNA의 메틸화 상태를 변형시키는 효소인 방법.

청구항 4

제2항에 있어서, 전사 활성화 도메인이 히스톤을 변형시키는 효소인 방법.

청구항 5

제1항에 있어서, 전사 활성화 도메인이 VP64 및 NF- κ B p65로 이루어진 군으로부터 선택되는 전사 활성화 도메인인 방법.

청구항 6

제2항에 있어서, 전사 활성화 도메인이 VP64 및 NF- κ B p65로 이루어진 군으로부터 선택되는 전사 활성화 도메인인 방법.

청구항 7

제1항 내지 제6항 중 어느 한 항에 있어서, 촉매적으로 불활성인 Cas9 단백질의 돌연변이가 (i) D10A 및 (ii) H840A인 방법.

청구항 8

제1항 내지 제6항 중 어느 한 항에 있어서, 이중성 기능성 도메인이 촉매적으로 불활성인 Cas9 단백질의 N 말단 또는 C 말단에 선택적 개재 링커(intervening linker)를 개재하여 연결되되, 상기 링커는 상기 융합 단백질의

활성을 방해하지 않는 것인 방법.

청구항 9

제1항 내지 제6항 중 어느 한 항에 있어서, 융합 단백질이 N-말단 또는 C-말단 상에 또는 촉매적으로 불활성인 Cas9 단백질과 이중성 기능성 도메인 사이에, 선택적으로 하나 이상의 개재 링커를 개재해서, 핵 국소화 서열 및 하나 이상의 에피토프 태그 중 하나 또는 둘 다를 더 포함하는 것인 방법.

청구항 10

제9항에 있어서, 상기 에피토프 태그는 c-myc, 6His 및 FLAG로 이루어진 군으로부터 선택되는 것인 방법.

청구항 11

제1항 내지 제6항 중 어느 한 항에 있어서, 포유동물 세포에서 융합 단백질을 발현시키는 것이, 포유동물 세포를 융합 단백질을 암호화하는 핵산에 작동가능하게 연결된 프로모터를 포함하는 발현 벡터와 접촉시키는 것을 포함하는 방법.

청구항 12

제1항 내지 제6항 중 어느 한 항에 있어서, 포유동물 세포에서 표적 유전자로 지향된 하나 이상의 안내 RNA를 발현시키는 것이, 포유동물 세포를 표적 유전자로 지향된 하나 이상의 안내 RNA를 암호화하는 핵산 서열에 작동가능하게 연결된 프로모터를 포함하는 하나 이상의 발현 벡터와 접촉시키는 것을 포함하는 방법.

발명의 설명

기술 분야

[0001] 우선권의 주장

[0002] 본 출원은 2013년 3월 15일자로 출원된 미국 특허 출원 제61/799,647호; 2013년 6월 21일자로 출원된 제61/838,178호; 2013년 6월 21일자로 출원된 제61/838,148호; 및 2013년 12월 26일자로 출원된 제61/921,007호의 유익을 주장한다. 이들 기초출원의 전문은 본 명세서에 참고로 포함된다.

[0003] 연방정부 지원된 연구 또는 개발

[0004] 본 발명은 미국국립보건원에 의해 부여된 등록번호 DP1 GM105378 및 미국 국방부의 미국 방위 고등 연구 계획국(Defense Advanced Research Projects Agency: DARPA)에 의해 부여된 W911NF-11-2-0056 하에서 정부 지원으로 이루어졌다. 정부는 본 발명에 특정 권리를 가진다.

[0005] 본 발명의 기술분야

[0006] 본 발명은 특정 계놈 좌위에 대한 유전적 및 후성적 조절 단백질, 예를 들어, 전사 활성제, 히스톤 변형 효소, DNA 메틸화 변형제의 RNA-안내 표적화를 위한 방법 및 작제물에 관한 것이다.

배경 기술

[0007] CRISPR/Cas 시스템으로 지칭되는 군집된, 주기적으로 간격을 둔 짧은 회문구조 반복부(regularly interspaced, short palindromic repeat: CRISPR) 및 CRISPR-관련(Cas) 유전자는 바이러스 및 다른 외래 핵산에 대한 방어를 매개하기 위해 다양한 박테리아 및 고세균에 의해 사용된다. 이들 시스템은 서열-특이적 방식으로 외래 핵산을 검출하고 침묵시키기 위해 소 RNA를 사용한다.

[0008] 3가지 유형의 CRISPR/Cas 시스템이 기재되었다(Makarova et al., Nat. Rev. Microbiol. 9, 467 (2011); Makarova et al., Biol. Direct 1, 7 (2006); Makarova et al., Biol. Direct 6, 38 (2011)). 최근의 연구는 II형 CRISPR/Cas 시스템이 DNA 표적 부위 및 Cas9 뉴클레아제에 대해 상보성을 지니는 단일 "안내 RNA(guide RNA)"를 사용함으로써 특이적 서열에 대해 시험관내에서 표적화된 이중-가닥 DNA 파손을 지시하도록 공학적으로 조작될 수 있다는 것을 나타내었다(Jinek et al., Science 2012; 337:816-821). 이 표적가능한 Cas9-기반 시스템은 또한 내인성 유전자 내로 표적화된 변용(alteration)을 유도하기 위해 배양된 인간 세포에서(Mali et al., Science. 2013 Feb 15;339(6121):823-6; Cong et al., Science. 2013 Feb 15;339(6121):819-23) 그리고 제브

라피쉬 생체내에서(Hwang and Fu et al., Nat Biotechnol. 2013 Mar;31(3):227-9) 작용한다.

발명의 내용

해결하려는 과제

- [0009] 적어도 부분적으로, 본 발명은 돌연변이에 의해 불활성화된 뉴클레아제 활성을 갖는 Cas9 뉴클레아제(또한 "dCas9"로서 알려짐)에 융합된 이종성 기능성 도메인(heterologous functional domain)(예를 들어, 전사 활성화 도메인)을 포함하는 융합 단백질의 개발에 기반한다. 공개된 연구가 특정 계능 좌위에 대해 촉매적으로 활성 및 불활성의 Cas9 뉴클레아제 단백질을 표적화하도록 안내 RNA를 사용하였지만, 추가적인 효과기 도메인을 동원하기 위해 이 시스템의 용도에 적합화된 작업은 아직 없었다. 이 작업은 또한 표적 유전자의 발현 수준에서의 (감소보다는) 증가를 초래하는 RNA-안내 과정의 첫 번째 입증을 제공한다.
- [0010] 추가로, 본 개시내용은 다중복합 gRNA가 단일 프로모터에 대한 다수의 dCas9-VP64 융합을 초래함으로써 전사의 상승적 활성화를 야기하기 위해 사용될 수 있다는 첫 번째 입증을 제공한다.

과제의 해결 수단

- [0011] 따라서, 제1 양태에서, 본 발명은 유전자 발현, 히스톤 또는 DNA를 변형시키는 이종성 기능성 도메인(HFD)에 연결된 촉매적으로 불활성인 CRISPR 관련 9(dCas9) 단백질을 포함하는 융합 단백질, 예를 들어, 전사 활성화 도메인, 전사 리프레서(예를 들어, 사일런서(silencer), 예컨대 이질염색질 단백질 1(HP1), 예를 들어, HP1 α 또는 HP1 β , 또는 전사 억제 도메인, 예를 들어, 크루펠-관련 박스(Krueppel-associated box: KRAB) 도메인, ERF 리프레서 도메인(ERD) 또는 mSin3A 상호작용 도메인(SID)), DNA의 메틸화 상태를 변형시키는 효소(예를 들어, DNA 메틸트랜스퍼라제(DNMT) 또는 텐-일레븐 전위(Ten-Eleven Translocation: TET) 단백질, 예를 들어, Tet 메틸사이토신 2산소화효소 1로도 알려진 TET1) 또는 히스톤 서브유닛을 변형시키는 효소(예를 들어, 히스톤 아세틸트랜스퍼라제(HAT), 히스톤 탈아세틸라제(HDAC) 또는 히스톤 탈메틸효소)를 제공한다. 일부 실시형태에서, 이종성 기능성 도메인은 전사 활성화 도메인, 예를 들어, VP64 또는 NF- κ B p65로부터의 전사 활성화 도메인; DNA 탈메틸화를 촉매하는 효소, 예를 들어, TET; 또는 히스톤 변형(예를 들어, LSD1, 히스톤 메틸트랜스퍼라제, HDAC 또는 HAT) 또는 예를 들어, 이질염색질 단백질 1(HP1), 예를 들어, HP1 α 또는 HP1 β 로부터의 전사 침묵 도메인; 또는 생물학적 테더(biological tether), 예를 들어, CRISPR/Cas 아형 Ypest 단백질 4(Csy4), MS2 또는 람다 N 단백질이다.
- [0012] 일부 실시형태에서, 촉매적으로 불활성인 Cas9 단백질은 스트렙토코커스 피오게네스(*S. pyogenes*)로부터 유래된다.
- [0013] 일부 실시형태에서, 촉매적으로 불활성인 Cas9 단백질은 D10, E762, H983 또는 D986에서; 및 H840 또는 N863에서, 예를 들어, D10 및 H840에서의 돌연변이, 예를 들어, D10A 또는 D10N 및 H840A 또는 H840N 또는 H840Y의 돌연변이를 포함한다.
- [0014] 일부 실시형태에서, 이종성 기능성 도메인은 촉매적으로 불활성인 Cas9 단백질의 N 말단 또는 C 말단에 선택적 개재 링커(intervening linker)를 개재해서 연결되되, 링커는 융합 단백질의 활성을 방해하지 않는다.
- [0015] 일부 실시형태에서, 융합 단백질은, N-말단, C-말단 상에 또는 촉매적으로 불활성인 CRISPR 관련 9(Cas9) 단백질과 이종성 기능성 도메인 사이에 선택적으로 하나 이상의 개재 링커를 개재해서 핵 국소화 서열과 하나 이상의 에피토프 태그 중 하나 또는 둘 다, 예를 들어, c-myc, 6His 또는 FLAG 태그를 포함한다.
- [0016] 추가 양태에서, 본 발명은 본 명세서에 기재된 융합 단백질을 암호화하는 핵산뿐만 아니라 핵산을 포함하는 발현 벡터 및 융합 단백질을 발현시키는 숙주 세포를 제공한다.
- [0017] 추가적인 양태에서, 본 발명은 세포 내에서 표적 유전자의 발현을 증가시키기 위한 방법을 제공한다. 상기 방법은, 예를 들어 융합 단백질을 암호화하는 서열을 포함하는 발현 벡터와 세포를 접촉시킴으로써 세포 내에서 본 명세서에 기재된 바와 같은 Cas9-HFD 융합 단백질을 발현시키는 단계, 및 또한, 예를 들어 하나 이상의 안내 RNA를 암호화하는 핵산 서열을 포함하는 하나 이상의 발현 벡터와 세포를 접촉시킴으로써 표적 유전자에 대해 상보성을 지니는 하나 이상의 안내 RNA를 세포 내에서 발현시키는 단계를 포함한다.
- [0018] 달리 정의되지 않는 한, 본 명세서에서 사용되는 모든 기술적 및 과학적 용어는 본 발명이 속하는 기술분야의 당업자에 의해 통상적으로 이해되는 것과 동일한 의미를 가진다. 방법 및 물질은 본 발명에서 사용을 위해 본

명세서에 기재되며; 당업계에 공지된 다른, 적합한 방법 및 물질이 또한 사용될 수 있다. 물질, 방법 및 예는 단지 예시적이며, 제한하는 것으로 의도되지 않는다. 모든 간행물, 특허출원, 특허, 서열, 데이터베이스 엔트리 및 본 명세서에 언급된 다른 참고문헌은 그들의 전문이 참고로 포함된다. 상충되는 경우에, 정의를 포함하는 본 명세서로 조절할 것이다.

[0019] 본 발명의 다른 특징 및 이점은 다음의 상세한 설명 및 도면으로부터 그리고 청구범위로부터 명확하게 될 것이다.

발명의 효과

[0020] 본 발명은 특정 게놈 좌위에 전사 활성제와 같은 이중성 기능성 도메인의 RNA-안내 표적화를 위한 방법 및 작제물을 제공한다.

도면의 간단한 설명

[0021] 본 특허 또는 출원은 컬러로 실행된 적어도 하나의 도면을 포함한다. 컬러 도면(들)을 지니는 본 특허 또는 특허 출원 공개의 사본은 요청 시 그리고 필요한 비용을 지불하였을 때 사무국에 의해 제공될 것이다.

도 1a는 특이적 DNA 서열에 Cas9 뉴클레아제를 동원하고, 이에 의해 표적화된 변용을 도입하는 단일 안내 RNA(sgRNA)를 도시하는 개략적 도면. 나타낸 안내 RNA의 서열은 GGAGCGAGCGGAGCGGUACAGUUUAGAGCUAGAAAAGCAAGUUAAGGCUAGUCCG(서열번호 9)이다.

도 1b는 특이적 DNA 서열에 Cas9 뉴클레아제를 동원하고, 이에 의해 표적화된 변용을 도입하기 위해 사용되는 더 긴 형태의 sgRNA를 도시하는 개략적 도면. 나타낸 안내 RNA의 서열은 GGAGCGAGCGGAGCGGUACAGUUUAGAGCUAGAAAAGCAAGUUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU(서열번호 10)이다.

도 1c는 전사 활성화 도메인에 융합되고 sgRNA에 의해 특이적 DNA 서열에 동원된 단백질 촉매적으로 불활성인 뉴클레아제 부분을 제공하기 위해 D10A 및 H840A 돌연변이를 함유하는 Cas9 단백질을 도시하는 개략적 도면. 나타낸 안내 RNA의 서열은 GGAGCGAGCGGAGCGGUACAGUUUAGAGCUAGAAAAGCAAGUUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU(서열번호 10)이다.

도 1d는 키메라 sgRNA에 의한 특정 게놈 표적 서열에 대한 dCas9-VP64 융합 단백질의 동원을 개략적으로 도시한 도면.

도 1e는 내인성 인간 VEGFA 유전자 프로모터에 대해 표적화된 16개의 sgRNA의 위치 및 배향을 도시한 다이어그램. 작은 수평 화살표는 5'에서 3'으로 향하는 화살표를 지니는 게놈 DNA 서열에 대해 상보성인 gRNA의 처음 20개 nt를 나타낸다. 회색 막대는 전사 시작 부위(직각 화살표)에 대해 넘버링한 인간 293 세포에서 이전에 정한 DNaseI 초민감 부위를 표시한다(Liu et al., J Biol Chem. 2001 Apr 6;276(14):11323-34).

도 2a는 각각 dCas9-VP64와 함께 발현되거나(회색 막대) 또는 이것 없이 발현된(검정색 막대), 다양한 sgRNA에 의한 293 세포 내 VEGFA 단백질 발현의 활성화를 나타내는 막대 그래프를 도시한 도면. 방법에 기재하는 바와 같이 표적을 벗어난 sgRNA 대조군에 대해 VEGFA의 배수적 활성화를 계산하였다. 각각의 실험을 3회 중복하여 수행하였고, 오차 막대는 평균의 표준 오차를 나타낸다. 별표는 양측, 단측 t-검정에 의해 결정된 바와 같은 표적을 벗어난 대조군 이상으로 유의하게 상승된 샘플을 나타낸다($p < 0.05$).

도 2b는 다중복합 sgRNA 발현이 dCas9-VP64 단백질에 의한 VEGFA 단백질 발현의 상승적 활성화를 유도한다는 것을 나타내는 막대 그래프를 도시한 도면. sgRNA의 표시된 조합이 dCas9-VP64와 함께 공동발현된 293 세포 내 VEGFA 단백질의 배수적 활성화를 나타낸다. 모든 이들 실험에서 형질감염을 위해 사용한 각각의 개개 sgRNA 발현 플라스미드의 양이 동일하다는 것을 주의한다. 배수적 활성화 값을 2A에서 도시한 바와 같이 계산하고, 회색 막대로 나타낸다. 개개의 sgRNA에 의해 유도되는 평균 배수적 활성화 값의 계산된 합계를 검정색 막대로서 각각의 조합에 대해 나타낸다. 별표는 분산분석(ANOVA)($p < 0.05$)에 의해 결정되는 예상 합계보다 상당히 더 큰 것으로 발견된 모든 조합을 표시한다.

도 3a는 내인성 인간 *NTF3* 유전자 프로모터에 표적화된 6개의 sgRNA의 위치 및 배향을 도시하는 다이어그램. 수평 화살표는 5'에서 3'으로 향하는 화살표를 지니는 게놈 DNA 서열에 대해 상보성인 sgRNA의 처음 20개의 nt를 나타낸다. 회색선은 제1 전사 엑손을 나타내는 더 굵은 막대 부분을 지니는 UCSC 게놈 브라우저 상에서의

ENCODE DNaseI 초민감 트랙으로부터 동정된 잠재적 개방 염색질 영역을 표시한다. 전사 시작 부위(+1, 직각 화살표)에 대한 넘버링을 나타낸다.

도 3b는 293 세포 내 sgRNA-안내된 dCas9-VP64에 의한 *NTF3* 유전자 발현의 활성화를 나타내는 막대 그래프를 도시한 도면. 정량적 RT-PCR에 의해 검출되고 *GAPDH* 대조군(델타Ct x 10⁴)에 대해 정규화된 *NTF3* mRNA의 상대적 발현을 표시량의 dCas9-VP64 및 *NTF3*-표적화된 sgRNA 발현 플라스미드와 함께 공동형질감염시킨 293 세포에 대해 나타낸다. 모든 실험을 평균 표준 오차를 나타내는 오차 막대를 이용하여 3회 중복으로 수행하였다. 별표는 양측, 단측 T-검정(P<0.05)에 의해 결정된 표적을 벗어난 gRNA 대조군보다 상당히 더 큰 샘플을 나타낸다.

도 3c는 다중복합 gRNA 발현이 dCas9-VP64 단백질에 의한 *NTF3* mRNA 발현의 상승적 활성화를 유도한다는 막대 그래프를 도시한 도면. 정량적 RT-PCR에 의해 검출되고, *GAPDH* 대조군(델타Ct x 104)에 의해 정규화된 *NTF3* mRNA의 상대적 발현을 dCas9-VP64 및 *NTF3*-표적화된 gRNA 발현 플라스미드의 표시된 조합과 함께 공동 형질감염된 293 세포에 대해 나타낸다. 모든 이들 실험에서 형질감염에 대해 사용한 각각의 개개 gRNA 발현 플라스미드의 양이 동일하다는 것을 주의한다. 모든 실험을 평균 표준 오차를 나타내는 오차 막대를 이용하여 3회 중복으로 수행하였다. 개개의 gRNA에 의해 유도된 평균 배수적 활성화 값의 계산된 총계를 각각의 조합에 대해 나타낸다.

도 4는 sgRNA 발현 벡터의 예시적 서열을 도시한 도면.

도 5는 CMV-T7-Cas9 D10A/H840A-3XFLAG-VP64 발현 벡터의 예시적 서열을 도시한 도면.

도 6은 CMV-T7-Cas9 보고 D10A/H840A-3XFLAG-VP64 발현 벡터의 예시적 서열을 도시한 도면.

도 7은 Cas9-HFD, 즉, Cas9-활성제의 예시적 서열을 도시한 도면. 선택적 3xFLAG 서열을 밑줄 표시하며; 핵 국소화 신호 PKKKRKVS(서열번호 11)는 소문자 이고; 두 링커는 볼드체이며; VP64 전사 활성제 서열인 DALDDFDLMDLGSALDDFDLMDLGSALDDFDLMDLGSALDDFDLMDL(서열번호 12)를 박스표시한다.

도 8a 내지 도 8b는 (8a) dCas9-NLS-3XFLAG-HP1알파 및 (8b) dCas9-NLS-3XFLAG-HP1베타의 예시적 서열을 도시한 도면. 박스 = 핵 국소화 신호; 밑줄 = 3중 플래그 태그; 이중 밑줄 = HP1알파 힌지 및 크로모솜도 도메인(Chromo shadow domain).

도 9는 dCas9-TET1의 예시적 서열을 도시한 도면.

도 10은 다양한 dCas9-VP64 융합 작제물로 얻은 결과를 나타내는 막대 그래프를 도시한 도면. 시험한 것 중에서, 최적화된 dCas9-VP64 구조는 N-말단 NLS(NFN) 및 dCas9와 VP64 사이에 놓이는 추가적인 NLS(N) 또는 FLAG 태그/NLS(NF)를 포함하였다. 인간 HEK293 세포 내 VEGFA 유전자의 발현을 RNA-안내 dCas9-VP64 융합에 의해 매개한 전사 활성화에 의해 활성화시켰다. dCas9-VP64의 변이체를 암호화하는 발현 플라스미드를 VEGFA 시작 코돈 상류의 영역 내 부위를 표적화한 3개의 gRNA를 발현시킨 플라스미드와 함께 공동형질감염시켰다(이 실험에서, gRNA를 단일 gRNA로부터 발현시키고 나서, Csy4 엔도리보뉴클레아제에 의해 처리하였다. VEGFA 단백질 발현을 ELISA에 의해 측정하고, 두 복제물의 평균을 평균의 표준 오차를 나타내는 오차 막대로 나타낸다.

도 11A 내지 도 11B는 Cas9 기능을 촉매적으로 불활성화시키기 위해 대안의 치환 돌연변이를 보유하는 dCas9-VP64 활성제의 활성을 나타내는 막대 그래프를 도시한 도면. (11A) 잔기 D10 및 H840에 대해 다양한 Cas9 불활성화 치환을 보유하는 dCas9-VP64 단백질을 발현시키는 플라스미드를 각각 단일 gRNA 또는 VEGFA 상류 영역을 표적화하는 3개의 별도로 표적화된 gRNA(각각 청색 및 적색 막대) 중 하나와 함께 HEK293 세포 내로 공동형질감염시켰다. (11B) 이들 dCas9-VP64 변이체를 발현시키는 플라스미드를 단일의 VEGFA-표적화된 gRNA를 안정하게 발현시키는 HEK293 세포주 내로 형질감염시켰다. 2개의 복제의 평균 및 나타낸 평균 표준 오차(오차 막대)를 이용하여 VEGFA 단백질 수준을 결정하였다.

발명을 실시하기 위한 구체적인 내용

[0022] 세포 및 살아있는 유기체 내 특정 게놈 위치에 이들 기능성 도메인의 RNA-가이드 표적화를 가능하게 할 목적을 위해 Cas9 단백질의 촉매적으로 불활성화된 형태에 융합된 이중성 기능성 도메인(예를 들어, 전사 활성화 도메인)의 융합 단백질이 본 명세서에 기재된다.

[0023] CRISPR/Cas 시스템은 침입 플라스미드 및 바이러스에 대해 보호하기 위한 방어 메커니즘으로서 박테리아 내에서 발달시켰다. 외래 핵산으로부터 유래된 짧은 프로토스페이서를 CRISPR 좌위 내로 혼입시키고, 후속적으로 전사시키고 나서, 짧은 CRISPR RNA(crRNA) 내로 처리하였다. 이들 crRNA를 제2의 tracrRNA와 복합체화시키고, 이어

서, 침입 핵산에 대한 그들의 서열 상보성을 사용하여 Cas9-매개된 절단, 및 그 결과로 일어나는 외래 핵산의 파괴를 안내한다. 2012년에, 다우드나(Doudna)와 동료들은 tracrRNA와 crRNA의 융합을 구성한 단일 안내 RNA(sgRNA)가 시험관내 특이적 DNA 서열에 대한 Cas9 뉴클레아제의 동원을 매개할 수 있다는 것을 입증하였다 (도 1c; 문헌[Jinek et al., Science 2012]).

[0024] 더 최근에, sgRNA의 더 긴 형태는 인간 세포와 제브라피쉬에서 표적화된 변용을 도입하기 위해 사용되었다(도 1b; 문헌[Mali et al. Science 2013, Hwang and Fu et al., Nat Biotechnol. 2013 Mar;31(3):227-9]). Qi 등은 Cas9의 촉매적으로 불활성인 돌연변이체 형태(dCas9로서 지칭됨)의 gRNA-매개된 동원이 이콜라이(*E. coli*)에 서의 특이적 내인성 유전자의 억제뿐만 아니라 인간 세포에서 EGFP 수용체 유전자의 억제를 야기할 수 있다는 것을 입증하였다(Qi et al., Cell 152, 1173-1183 (2013)). 이 연구는 유전자 발현의 조절을 위한 RNA-안내 Cas9 기법에 적합하게 될 가능성을 입증하였지만, 이중성 기능성 도메인(예를 들어, -전사 활성화 도메인)이 프로그래밍 가능한 sgRNA 또는 이중 gRNA(dgRNA - 즉- 맞춤형(customized) crRNA 및 tracrRNA)에 의해 특정 게놈 부위에 동원될 능력을 방해하는 일 없이 dCas9에 융합될 수 있었는지의 여부는 시험 또는 입증하지 않았다.

[0025] 본 명세서에 기재된 바와 같이, Cas9-매개된 뉴클레아제 활성을 안내하는 것에 추가로, 게놈 내 특이적 부위에 대해 Cas9에 융합된 표적 이중성 기능성 도메인 (Cas9-HFD)에 대해 CRISPR-유래 RNA를 사용할 수 있다(도 1c). 예를 들어, 본 명세서에 기재된 바와 같이, 특이적 유전자의 프로모터에 대한 표적 Cas9-HFD, 예를 들어, Cas9-전사 활성화제(본 명세서에서 이후에 Cas9-활성제로서 지칭함)에 대해 단일 안내 RNA(sgRNA)를 사용함으로써 표적 유전자의 발현을 증가시킬 수 있다. 따라서, Cas9-HFD는 안내 RNA의 서열 상보성에 의해 정해진 표적 특이성으로 게놈 내 부위에 국소화될 수 있다. 표적 서열은 또한 PAM 서열(RNA에 의해 구체화된 서열에 인접한 Cas9 단백질에 의해 구체화된 2 내지 5개의 뉴클레오타이드 서열)을 포함한다.

[0026] Cas9-HFD는 이중성 기능성 도메인(예를 들어, 전사 활성화 도메인, 예를 들어, VP64 또는 NF- κ B p65로부터)를 촉매적으로 불활성인 Cas9 단백질의 N-말단 또는 C-말단에 융합시킴으로써 생성된다.

[0027] **Cas9**

[0028] 다수의 박테리아는 Cas9 단백질 변이체를 발현시킨다. 스트렙토코커스 피오게네스(*Streptococcus pyogenes*)로부터의 Cas9는 현재 가장 통상적으로 사용되며; 일부 다른 Cas9 단백질은 스트렙토코커스 피오게네스 Cas9와 고수준의 서열 동일성을 가지며, 동일한 안내 RNA를 사용한다. 다른 것들은 더 다양하며, 상이한 gRNA를 사용하고, 마찬가지로 상이한 PAM 서열을 인식한다(RNA에 의해 구체화되는 서열에 인접한 단백질에 의해 구체화된 2 내지 5개의 뉴클레오타이드 서열). 실린스키(Chylinski) 등은 박테리아의 거대한 군으로부터 Cas9 단백질을 분류하였고(RNA Biology 10:5, 1-12; 2013), 매우 다수의 Cas9 단백질은 본 명세서에 참고로 포함되는 본 명세서의 보충적 도 1에서 그리고 보충적 표 1에서 열거된다. 추가적인 Cas9 단백질은 문헌[Esvelt et al., Nat Methods. 2013 Nov; 10(11):1116-21 및 Fonfara et al., "Phylogeny of Cas9 determines functional exchangeability of dual-RNA and Cas9 among orthologous type II CRISPR-Cas systems." Nucleic Acids Res. 2013 Nov 22. [Epub ahead of print] doi:10.1093/nar/gkt1074]에 기재되어 있다.

[0029] 다양한 종의 Cas9 분자가 본 명세서에 기재된 방법 및 조성물에서 사용될 수 있다. 스트렙토코커스 피오게네스 및 스트렙토코커스 써모필루스(*S. thermophilus*) Cas9 분자는 본 명세서의 다수의 개시내용이지만, 본 명세서에 열거된 종의 Cas9 단백질의, 또는 이로부터 유래되거나 또는 이를 기반으로 한 Cas9 분자가 마찬가지로 사용될 수 있다. 다시 말해서, 본 명세서의 다수의 설명이 스트렙토코커스 피오게네스 및 스트렙토코커스 써모필루스 Cas9 분자를 사용하지만, 다른 종으로부터의 Cas9 분자는 그들을 대신할 수 있다. 이러한 종은 문헌[Chylinski et al., 2013]의 보충적 도 1에 기반하여 만들어진 다음의 표에 제시되는 것을 포함한다.

대안의 Cas9 단백질	
제네뱅크 등록번호	박테리아
303229466	베일로넬라 아티피카(<i>Veillonella atypica</i>) ACS-134-V-Col7a
34762592	푸소박테리움 누클리아툼 아종 빈세티(<i>Fusobacterium nucleatum subsp. vincentii</i>)
374307738	필리팩터 알코시스(<i>Filifactor alocis</i>) ATCC 35886
320528778	솔로박테리움 무레이(<i>Solobacterium moorei</i>) F0204
291520705	코프로코커스 카투스(<i>Coprococcus catus</i>) GD-7
42525843	트레포네마 덴티콜라(<i>Treponema denticola</i>) ATCC 35405
304438954	펩토니필루스 두에르데니(<i>Peptoniphilus duerdenii</i>) ATCC BAA-1640
224543312	카테니박테리움 미츠오카이(<i>Catenibacterium mitsuokai</i>) DSM 15897
24379809	스트렙토코커스 뮤탄스(<i>Streptococcus mutans</i>) UA159
15675041	스트렙토코커스 껌오제네스 SR370
16801805	리스테리아 이노쿠아(<i>Listeria innocua</i>) Clip11262
116628213	스트렙토코커스 서모필루스(<i>Streptococcus thermophilus</i>) LMD-9
323463801	스타필로코커스 슈딘테르메디우스(<i>Staphylococcus pseudintermedius</i>) ED99
352684361	아시다미노코커스 인테스티니(<i>Acidaminococcus intestinalis</i>) RvC-MR95
302336020	올세넬라 울리(<i>Olsenella uli</i>) DSM 7084
366983953	오에노코커스 키타하라(<i>Oenococcus kitaharae</i>) DSM 17330
310286728	비피도박테리움 비피둠(<i>Bifidobacterium bifidum</i>) S17
258509199	락토바실러스 람노저스(<i>Lactobacillus rhamnosus</i>) GG
300361537	락토바실러스 가세리(<i>Lactobacillus gasseri</i>) JV-V03

169823755	피네골디아 마그나(<i>Finigoldia magna</i>) ATCC 29328
47458868	마이코플라즈마 모빌레(<i>Mycoplasma mobile</i>) 163K
284931710	마이코플라즈마 갈리셉티쿰(<i>Mycoplasma gallisepticum</i>) 군주 F
363542550	마이코플라즈마 오비뉴모니아(<i>Mycoplasma ovipneumoniae</i>) SC01
384393286	마이코플라즈마 카니스(<i>Mycoplasma canis</i>) PG 14
71894592	마이코플라즈마 시노비아(<i>Mycoplasma synoviae</i>) 53
238924075	유박테리움 렉탈리스(<i>Eubacterium rectale</i>) ATCC 33656
116627542	스트렙토코커스 서모필루스 LMD-9
315149830	엔테로코커스 패칼리스(<i>Enterococcus faecalis</i>) TX0012
315659848	스타필로코커스 루그두넨시스(<i>Staphylococcus lugdunensis</i>) M23590
160915782	유박테리움 돌리쿰(<i>Eubacterium dolichum</i>) DSM 3991
336393381	락토바실러스 코리니포르미스 아종 토르쿠엔스(<i>Lactobacillus coryniformis</i> subsp. <i>torquens</i>)
310780384	일리오박터 폴리트로푸스(<i>Ilyobacter polytropus</i>) DSM 2926
325677756	루미노코커스 알부스(<i>Ruminococcus albus</i>) 8
187736489	아커만시아 뮤시니필라(<i>Akkermansia muciniphila</i>) ATCC BAA-835
117929158	아시도테르무스 셀룰로리티쿠스(<i>Acidothermus cellulolyticus</i>) 11B
189440764	비피도박테리움 롱검(<i>Bifidobacterium longum</i>) DJ010A
283456135	비피도박테리움 덴티움 Bd1
38232678	코리네박테리움 디프테리아(<i>Corynebacterium diphtheriae</i>) NCTC 13129

187250660	엘루시미크로비움 미누툼 (<i>Elusimicrobium minutum</i>) <i>Peil01</i>
319957206	니트라티프락터 살수기니스 (<i>Nitratifractor</i> <i>salsuginis</i>) <i>DSM 16511</i>
325972003	스파에로케타 글로부스 균주 버디 (<i>Sphaerochaeta</i> <i>globus</i> str. <i>Buddy</i>)
261414553	피프로박터 속시노게네스 아종 속시노게네스 (<i>Fibrobacter succinogenes</i> subsp. <i>succinogenes</i>)
60683389	박테로이데스 프라길리스 (<i>Bacteroides fragilis</i>) <i>MTC 8343</i>
256819408	캡노사이토파가 오크라세아 (<i>Capnocytophaga</i> <i>ochracea</i>) <i>DSM 7271</i>
90425961	로도슈도모나스 팔루스트리스 (<i>Rhodospseudomonas</i> <i>palustris</i>) <i>BisB18</i>
373501184	프레보텔라 미칸스 (<i>Prevotella micans</i>) <i>R0438</i>
294674019	프레보텔라 루미니콜라 (<i>Prevotella ruminicola</i>) <i>23</i>
365959402	플라보박테리움 칼럼나트 <i>ATCC 49512</i>
312879015	아미노모나스 파우시보란스 (<i>Aminomonas</i> <i>paucivorans</i>) <i>DSM 12260</i>
83591793	로도스피릴룸 루브룸 (<i>Rhodospirillum rubrum</i>) <i>ATCC</i> <i>11170</i>
294086111	칸디다투스 푸니세이스피릴룸 (<i>Candidatus</i> <i>Puniceispirillum marinum</i>) <i>IMCC1322</i>
121608211	벨미네프로박터 에이세니아 (<i>Verminephrobacter</i> <i>eiseniae</i>) <i>EF01-2</i>
344171927	랄스토니아 시지기 (<i>Ralstonia syzygii</i>) <i>R24</i>
159042956	디노로세오박터 시바 (<i>Dinoroseobacter shibae</i>) <i>DFL</i> <i>12</i>
288957741	아조스피릴룸 (<i>Azospirillum</i>) sp- <i>B510</i>
92109262	니트로박터 함버겐시스 (<i>Nitrobacter hamburgensis</i>) <i>X14</i>

[0032]

148255343	브레디아조비움 (<i>Bradyrhizobium</i>) sp- <i>BTA11</i>
34557790	울리넬라 속시노게네스 (<i>Wolinella succinogenes</i>) <i>DSM 1740</i>
218563121	캠필로박터 제주니 아종 제주니 (<i>Campylobacter</i> <i>jejuni</i> subsp. <i>jejuni</i>)
291276265	헬리코박터 무스텔라 (<i>Helicobacter mustelae</i>) <i>12198</i>
229113166	바실러스 세레우스 (<i>Bacillus cereus</i>) <i>Rock1-15</i>
222109285	아시도보락스 에브레우스 (<i>Acidovorax ebreus</i>) <i>TPSY</i>
189485225	미배양 흰개미군 1 (<i>uncultured Termite group 1</i>)
182624245	클로스트리디움 페르프린겐스 디 균주 (<i>Clostridium</i> <i>perfringens</i> D str.)
220930482	클로스트리디움 셀룰로리틱움 (<i>Clostridium</i> <i>cellulolyticum</i>) <i>H10</i>
154250555	파비바쿨룸 라바멘티보란스 <i>DS-1</i>
257413184	로세부리아 인테스티날리스 <i>LI-82</i>
218767588	나이세리아 메닝기티디스 (<i>Neisseria meningitidis</i>) <i>Z2491</i>
15602992	파스테우렐라 톨토시다 아종 톨토시다 (<i>Pasteurella</i> <i>multocida</i> subsp. <i>multocida</i>)
319941583	수테렐라 와즈워텐시스 (<i>Sutterella wadsworthensis</i>) <i>3.1</i>
254447899	감마 프로테오박테리움 (<i>gamma proteobacterium</i>) <i>HTCC5015</i>
54296138	레지오넬라 뉴모필라 균주 파리 (<i>Legionella</i> <i>pneumophila</i> str. <i>Paris</i>)
331001027	파라수테렐라 엑스크레멘티호미니 (<i>Parasutterella</i> <i>excrementihominis</i>) <i>YIT 11859</i>
34557932	울리넬라 속시노게네스 <i>DSM 1740</i>
118497352	프란시셀라 노비시다 (<i>Francisella novicida</i>) <i>U112</i>

[0033]

[0034]

본 명세서에 기재된 작제물 및 방법은 임의의 해당 Cas9 단백질, 및 그들의 대응하는 안내 RNA 또는 양립가능한 다른 안내 RNA를 포함할 수 있다. 스트렙토코커스 서모필루스 LMD-9 CRISPR1 시스템으로부터의 Cas9는 쿡(Cong) 등(Science 339, 819 (2013))에서 인간 세포 내에서 작용하는 것으로 나타났다. 추가적으로 지네크(Jinek) 등은 시험관 내에서, 비록 약간 감소된 효율을 지니기는 하지만, 스트렙토코커스 써모필루스 및 리스테리아 이노쿠아로부터의 Cas9 오솔로그(그러나 상이한 안내 RNA를 사용할 가능성이 있는 나이세리아 메닝기티디스 또는 캠필로박터 제주니로부터의 오솔로그는 아님)가 이중 스트렙토코커스 피오게네스 gRNA에 의해 안내되어 표적 플라스미드 DNA를 절단할 수 있다는 것을 나타내었다.

[0035]

일부 실시형태에서, 본 발명의 시스템은 촉매적으로 불활성인 단백질의 뉴클레아제 부분을 제공하기 위해 D10, E762, H983 또는 D986 및 H840 또는 N863, 예를 들어, D10A/D10N 및 H840A/H840N/H840Y에서의 돌연변이를 함유하여 박테리아에서 암호화된 바와 같은 또는 포유류 세포에서 발현을 위해 코돈 최적화된 스트렙토코커스 피오게네스로부터의 Cas9 단백질을 이용하며; 이들 위치에서의 치환은 알라닌일 수 있고(그들은 문헌[Nishimasu et al., Cell 156, 935-949 (2014)]에서와 같음) 또는 그들은 다른 잔기, 예를 들어, 글루타민, 아스파라긴, 타이로신, 세린 또는 아스파르트산염일 수 있었다, 예를 들어, E762Q, H983N, H983Y, D986N, N863D, N863S 또는 N863H(도 1c). 본 명세서에서 기재되는 방법 및 조성물에서 사용될 수 있는 촉매적으로 불활성인 스트렙토코커스 피오게네스 Cas9의 서열은 다음과 같으며; D10A 및 H840A의 예시적인 돌연변이는 볼드체이며 밑줄 표시되어 있다.

```

10      20      30      40      50      60
MDKKYSIGLA IGTNSVGWAV ITDEYKVPESK KFKVLGNTDR HSIKKNLIGA LLFDSGETAE

70      80      90      100     110     120
ATRLKRTARR RYTRRNRIAC YLQEIFSNEM AKVDDSFHR LEESFLVEED KKHERHPIFG

130     140     150     160     170     180
NIVDEVAYHE KYPTIYHLRK KLVDSTDKAD LRLIYLALAH MIKFRGHFLI EGDLPDNDSD

190     200     210     220     230     240
VDKLFIQLVQ TYNQLFENP INASGVDAKA ILSARLSKSR RLENLIAQLP GEKKNGLFGN

250     260     270     280     290     300
LIALSLGLTP NFKSNFDLAE DAKLQLSKDT YDDDLNLLA QIGDQYADLF LAAKNLSDAI

310     320     330     340     350     360
LLSDILRVNT BITKAPLSAS MIKRYDEHHQ DLTLKALVR QQLPEKYKEI FFDQSKNGYA

370     380     390     400     410     420
GYIDGGASQE EFYKFIKPIE EKMDGTEELL VKLNREDLLR KQRTFDNGSI PHQIHLGELH

430     440     450     460     470     480
AILRRQEDFY PFLKDNREKI EKILTFRIPY YVGPLARGNS RFAWMTRKSE ETITPWNFEE

490     500     510     520     530     540
VVDKGASAGS FIERMTNEDK NLPNEKVLPK HSLLYEYFTV YNELTKVKYV TEGMRKPAFL

550     560     570     580     590     600
SGEQKKAIVD LLFKTNRKVT VKQLKEDYFK KIECFDSVEI SGVEDRFNAS LGTYHDLKI

610     620     630     640     650     660
IKDKDFLDNE ENEDILEDIV LTLTLFEDRE MIEERLKTYA HLFDDKVMKQ LKRRRYTGWG

670     680     690     700     710     720
RLSRKLINGI RDKQSGKTIL DFLKSDGFAN RNFMQLIHDD SLTFKEDIQK AQVSGQGDLSL

730     740     750     760     770     780
HEHIANLAGS PAIKKGILQT VKVVDLVKV MGRHKPENIV IEMARENQTT QKGQKNSRER

790     800     810     820     830     840
MKRIEIGIKE LGSQILKEHP VENTQLQNEK LYLYYLQNGR DMYVDQELDI NRLSDYDVDAA

850     860     870     880     890     900
IVPQSFLKDD SIDNKVLTRS DKNRGKSDNV PSEEVVKKMK NYWRQLLNAK LITQRKFDNL

910     920     930     940     950     960
TKAERGGLSE LDKAGFIKRQ LVETRQITKH VAQILDSRMN TKYDENDKLI REVKVITLKS

970     980     990     1000    1010    1020
KLVSDFRKDF QFYKVVREINN YHHAHDAYLN AVVGTAIIKK YPKLESEFVY GDYKQYDVRK

1030    1040    1050    1060    1070    1080
MIAKSEQEIG KATAKYFFYS NIMNFFKTEI TLANGIIRK PLIETNGETG EIVWDKGRDF

1090    1100    1110    1120    1130    1140
ATVRKVLSPM QVNIKKTEV QTGGFSKESI LPKRNSDKLI ARKKDWDPKK YGGFDSPTVA

1150    1160    1170    1180    1190    1200
YSVLVVAKEV KGSKKLKV KELLGITIME RSSFEKNPID FLEAKGYKEV KKDIIKLPK

1210    1220    1230    1240    1250    1260
YSLFELENGR KRMLASAGEL QKGNEALPS KYVNFYLLAS HYEKLKGSPE DNEQKQLFVE

1270    1280    1290    1300    1310    1320
QHKHYLDEII EQISEFSKRV ILADANLDKV LSAYNKHRDK PIREQAENII HLFITLNLGA

1330    1340    1350    1360
PAAFKYFDTT IDRKYTSTK EVLDATLIHQ SITGLYETRI DLSQLGGD (서열번호 13)

```

[0036]

```

1150    1160    1170    1180    1190    1200
YSVLVVAKEV KGSKKLKV KELLGITIME RSSFEKNPID FLEAKGYKEV KKDIIKLPK

1210    1220    1230    1240    1250    1260
YSLFELENGR KRMLASAGEL QKGNEALPS KYVNFYLLAS HYEKLKGSPE DNEQKQLFVE

1270    1280    1290    1300    1310    1320
QHKHYLDEII EQISEFSKRV ILADANLDKV LSAYNKHRDK PIREQAENII HLFITLNLGA

1330    1340    1350    1360
PAAFKYFDTT IDRKYTSTK EVLDATLIHQ SITGLYETRI DLSQLGGD (서열번호 13)

```

[0037]

- [0038] 일부 실시형태에서, 본 명세서에서 사용되는 Cas9 뉴클레아제는 스트렙토코커스 피오게네스 Cas9의 서열에 대해 적어도 약 50%, 즉, 서열번호 13에 대해 적어도 50% 동일하다. 일부 실시형태에서, 뉴클레오타이드 서열은 서열번호 13에 대해 약 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 99% 또는 100% 동일하다.
- [0039] 일부 실시형태에서, 본 명세서에서 사용된 촉매적으로 불활성인 Cas9는 촉매적으로 불활성인 스트렙토코커스 피오게네스 Cas9의 서열에 대해 적어도 약 50% 동일하고, 즉, 서열번호 13에 대해 적어도 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 99% 또는 100% 동일하되, D10 및 H840에서의 돌연변이, 예를 들어, D10A/D10N 및 H840A/H840N/H840Y가 유지된다.
- [0040] 일부 실시형태에서, 일부 실시형태에서, 서열번호 13으로부터의 임의의 차이는 문헌[Chylinski et al., RNA Biology 10:5, 1-12; 2013(예를 들어, 이의 보충적 도 1 및 보충적 표 1에서); Esvelt et al., Nat Methods. 2013 Nov;10(11):1116-21 및 Fonfara et al., Nucl. Acids Res. (2014) 42 (4): 2577-2590. [Epub ahead of print 2013 Nov 22] doi:10.1093/nar/gkt1074]에 제시된 서열의 서열 정렬에 의해 동정되는 비보존적 영역이며, D10 및 H840에서의 돌연변이, 예를 들어, D10A/D10N 및 H840A/H840N/H840Y가 유지된다.
- [0041] 두 서열의 동일성 백분율을 결정하기 위해, 서열을 최적의 비교 목적을 위해 정렬한다(갭은 최적의 정렬에 필요한 제1 아미노산 및 제2 아미노산 또는 핵산 서열 중 하나 또는 둘 다에 도입되며, 비-상동성 서열은 비교 목적을 위해 무시될 수 있다). 비교 목적을 위해 정렬된 기준 서열의 길이는 적어도 50%(일부 실시형태에서, 기준 서열 길이의 약 50%, 55%, 60%, 65%, 70%, 75%, 85%, 90%, 95% 또는 100%가 정렬됨)이다. 이어서, 대응하는 위치에서 뉴클레오타이드 또는 잔기가 비교된다. 제1 서열 내 위치가 제2 서열 내 대응하는 위치와 동일한 뉴클레오타이드 또는 잔기에 의해 점유된다면, 분자는 해당 위치에서 동일하다. 두 서열 간의 동일성 백분율은 서열에 의해 공유되는 동일한 위치의 함수이며, 두 서열의 최적의 정렬을 위해 도입될 필요가 있는 갭의 수 및 각각의 갭의 길이를 고려한다.
- [0042] 서열의 비교 및 두 서열 간의 백분율 동일성의 결정은 수학적 알고리즘을 이용하여 수행될 수 있다. 본 출원의 목적을 위해, 두 아미노산 서열 간의 백분율 동일성은 갭 페널티 12, 갭 연장 페널티 4 및 틀림 갭 페널티 5를 이용하는 블로섬(Blossum) 62 스코어링 매트릭스를 이용하여, GCG 소프트웨어 패키지 내 GAP 프로그램에 포함된 니들만(Needleman) 및 분슈(Wunsch)((1970) J. Mol. Biol. 48:444-453) 알고리즘을 이용하여 결정된다.
- [0043] **이종성 기능성 도메인**
- [0044] 전사 활성화 도메인은 Cas9의 N 또는 C 말단에 융합될 수 있다. 추가로, 본 설명은 전사 활성화 도메인, 다른 이종성 기능성 도메인(예를 들어, 전사 리프레서(예를 들어, KRAB, ERD, SID 및 기타, 예를 들어, *ets2* 리프레서 인자(ERF) 리프레서 도메인(ERD)의 아미노산 473 내지 530, KOX1의 KRAB 도메인의 아미노산 1 내지 97 또는 Mad mSIN3 상호작용 도메인(SID)의 아미노산 1 내지 36을 예시하며; 문헌[Beerli et al., PNAS USA 95:14628-14633 (1998)] 참조) 또는 사일런서, 예컨대 이질염색질 단백질 1(HP1, 또한 swi6으로서 알려짐), 예를 들어, HP1 α 또는 HP1 β ; 고정된 RNA 결합 서열에 융합된 긴 비-암호 RNA(lncRNA)를 동원할 수 있는 단백질 또는 펩타이드, 예컨대 MS2 코딩 단백질에 의해 결합되는 것, 엔도리보뉴클레아제 Csy4 또는 람다 N 단백질; DNA의 메틸화 상태를 변형시키는 효소(예를 들어, DNA 메틸트랜스퍼라제(DNMT) 또는 TET 단백질); 또는 당업계에 공지된 바와 같은 히스톤 서브유닛을 변형시키는 효소(예를 들어, 히스톤 아세틸트랜스퍼라제(HAT), 히스톤 탈아세틸라제(HDAC), 히스톤 메틸트랜스퍼라제(예를 들어, 라이신 또는 알기닌 잔기의 메틸화를 위함) 또는 히스톤 탈메틸 효소(예를 들어, 라이신 또는 알기닌 잔기의 탈메틸화를 위함))가 또한 사용될 수 있다. 이러한 도메인, 예를 들어 DNA에서 메틸화된 사이토신의 수산화를 촉매하는 도메인에 대한 다수의 서열은 당업계에 공지되어 있다. 예시적인 단백질은 텐-일레븐-전위(TET)1-3 패밀리, DNA에서 5-메틸사이토신(5-mC)을 5-하이드록시메틸사이토신(5-hmC)으로 전환시키는 효소를 포함한다.

[0045] 인간 TET1-3에 대한 서열은 당업계에 공지되어 있으며, 다음의 표에 나타낸다:

유전자	젠뱅크 등록번호	
	아미노산	핵산
TET1	NP_085128.2	NM_030625.2
TET2*	NP_001120680.1(변이체 1)	NM_001127208.2
	NP_060098.3(변이체 2)	NM_017628.4
TET3	NP_659430.1	NM_144993.1

* 변이체 (1)은 더 긴 전사체를 나타내며, 더 긴 동형(isoform)(a)을 암호화한다. 변이체 (2)는 5' UTR에서 그리고 3' UTR에서 상이하며 변이체 1과 비교되는 서열을 암호화한다. 얻어진 동형(b)은 더 짧으며, 동형 a와 비교하여 별도의 C-말단을 가진다.

[0046]

[0047] 일부 실시형태에서, 촉매적 도메인의 모든 또는 일부의 전장 서열이 포함될 수 있으며, 예를 들어, 촉매적 모듈은 시스테인-풍부 연장부 및 7개의 고도로 보존된 엑손에 의해 암호화된 20GFeD0 도메인, 예를 들어, 아미노산 1580 내지 2052를 포함하는 Tet1 촉매적 도메인, 아미노산 1290 내지 1905를 포함하는 Tet2 및 아미노산 966 내지 1678을 포함하는 Tet3을 포함한다. 예를 들어, 모두 3개의 Tet 단백질 및 전장 서열(예를 들어, 서열 2c 참조)에 대해 이의 보충적 물질(ftp 사이트 ftp.ncbi.nih.gov/pub/aravind/DONS/supplementary_material_DONS.html에서 입수가 가능)에서 중요한 촉매적 잔기를 예시하는 정렬에 대해 문헌[Iyer et al., Cell 주기. 2009 Jun 1;8(11):1698-710. Epub 2009 Jun 27]의 도 1을 참조하며; 일부 실시형태에서, 서열은 Tet1의 아미노산 1418 내지 2136 또는 Tet2/3에서의 대응하는 영역을 포함한다.

[0048] 다른 촉매적 모듈은 문헌[Iyer et al., 2009]에서 동정된 단백질로부터 유래될 수 있다.

[0049] 일부 실시형태에서, 이중성 기능성 도메인은 생물학적 테더이며, 모두 또는 일부의 MS2 외피 단백질, 엔도리보뉴클레아제 Csy4 또는 람다 N 단백질(예를 들어, 이들로부터의 DNA 결합 도메인)을 포함한다. 이들 단백질은 dCas9 gRNA 표적화 서열에 의해 구획화된 자리에 특이적 스템-루프(stem-loop) 구조를 함유하는 RNA 분자를 동원하기 위해 사용될 수 있다. 예를 들어, MS2 외피 단백질, 엔도리보뉴클레아제 Csy4 또는 람다 N에 융합된 dCas9는 Csy4, MS2 또는 람다 N 결합 서열에 연결되는 긴 비-암호 RNA(lncRNA), 예컨대 XIST 또는 HOTAIR을 동원하기 위해 사용될 수 있으며; 예를 들어, 문헌[Keryer-Bibens et al., Biol. Cell 100:125-138 (2008)]을 참조한다. 대안적으로, Csy4, MS2 또는 람다 N 단백질 결합 서열은, 예를 들어, 문헌[Keryer-Bibens et al., 상기 참조]에 기재된 바와 같은 다른 단백질에 연결될 수 있으며, 단백질은 본 명세서에 기재된 방법 및 조성물을 이용하여 dCas9 결합 부위로 표적화될 수 있다. 일부 실시형태에서, Csy4는 촉매적으로 불활성이다.

[0050] 일부 실시형태에서, 융합 단백질은 dCas9와 이중성 기능성 도메인 사이에 링커를 포함한다. 이들 융합 단백질에서(또는 연쇄된 구조에서의 융합 단백질 사이에서) 사용될 수 있는 링커는 융합 단백질의 기능을 방해하지 않는 임의의 서열을 포함할 수 있다. 바람직한 실시형태에서, 링커는 짧으며, 예를 들어, 2 내지 20개의 아미노산이고, 전형적으로 가요성이다(즉, 글라이신, 알라닌 및 세린과 같이 고자유도를 지니는 아미노산을 포함함). 일부 실시형태에서, 링커는 GGGG(서열번호 14) 또는 GGGGS(서열번호 15)로 이루어진 하나 이상의 단위, 예를 들어, GGGG(서열번호 14) 또는 GGGGS(서열번호 15) 단위의 2, 3, 4개 이상의 반복부를 포함한다. 다른 링커 서열이 또한 사용될 수 있다.

[0051] 사용방법

[0052] 기재된 Cas9-HFD 시스템은 내인성 유전자의 발현을 변형시키기 위한 유용하고 다재다능한 도구이다. 이를 달성하기 위한 현재의 방법은 표적화될 부위 각각에 대해 신규한 공학적으로 조작된 DNA-결합 단백질(예컨대 공학적으로 조작된 아연 핑거 또는 전사 활성제-유사 효소가 DNA 결합 도메인)의 생성을 필요로 한다. 이 방법은 각각의 표적 부위에 결합하도록 특이적으로 공학적으로 조작된 거대 단백질의 발현을 필요로 하기 때문에, 그들은 다중복합화를 위한 그들의 능력을 제한한다. 그러나, Cas9-HFD는 다중 짧은 gRNA의 발현에 의해 게놈 내 다수 부위에 표적화될 수 있는 단일 Cas9-HFD 단백질만의 발현을 필요로 한다. 이 시스템은 따라서 다수의 유전자의 발현을 동시에 유도하기 위해 또는 단일 유전자, 프로모터 또는 인핸서에 대해 다수의 Cas9-HFD를 동원하기 위해 용이하게 사용될 수 있었다. 이 능력은, 예를 들어 기본적 생물학적 연구에 대해 넓은 효용을 가질 것이며, 이는 유전자 기능을 연구하기 위해 그리고 단일 경로에서 그리고 연구자들이 다중 입력 신호에 반응하는 세포 내 회로를 생성할 수 있는 합성 생물학에서 다수 유전자의 발현을 조작하기 위해 사용될 수 있다. 이 기술이 실행되고 다중복합화에 적용될 수 있는 상대적 용이함은 그것을 다수의 넓은 범위의 적용을 지니는 광범위하게 유

용한 기술로 만들 것이다.

[0053] 본 명세서에 기재된 방법은 본 명세서에 기재된 Cas9-HFD를 암호화하는 핵산 및 선택 유전자로 향하는 하나 이상의 안내 RNA를 암호화하는 핵산과 세포를 접촉시킴으로써 해당 유전자의 발현을 조절하는 단계를 포함한다.

[0054] 안내 RNA(gRNA)

[0055] 안내 RNA는 일반적으로 말해서 2가지 상이한 시스템을 시작한다: 별도의 crRNA 및 tracrRNA를 사용하고 Cas9에 의한 절단을 안내하도록 함께 작용하는 시스템 1, 및 단일 시스템에서 2개의 별도의 안내 RNA를 합하는 키메라 crRNA-tracrRNA 혼성체를 사용하는 시스템 2(단일 안내 RNA 또는 sgRNA로서 지칭됨, 또한 문헌[Jinek et al., Science 2012; 337:816-821] 참조). tracrRNA는 가변적으로 절단될 수 있으며, 길이 범위는 별개의 시스템(시스템 1)과 키메라 gRNA 시스템(시스템 2) 둘 다에서 작용하는 것으로 나타났다. 예를 들어, 일부 실시형태에서, tracrRNA는 적어도 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35 또는 40개의 nt만큼 그의 3' 말단으로부터 절단될 수 있다. 일부 실시형태에서, tracrRNA 분자는 적어도 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35 또는 40개의 nt만큼 그의 5' 말단으로부터 절단될 수 있다. 대안적으로, tracrRNA 분자는 5'과 3' 말단 둘 다로부터, 예를 들어, 5' 말단 상에서 적어도 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15 또는 20개의 nt만큼 그리고 3' 말단 상에서 적어도 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35 또는 40개의 nt만큼 절단될 수 있다. 예를 들어, 문헌[Jinek et al., Science 2012; 337:816-821; Mali et al., Science. 2013 Feb 15;339(6121):823-6; Cong et al., Science. 2013 Feb 15;339(6121):819-23; 및 Hwang and Fu et al., Nat Biotechnol. 2013 Mar;31(3):227-9; Jinek et al., Elife 2, e00471 (2013))] 참조. 시스템 2에 대해, 일반적으로 더 긴 길이의 키메라 gRNA가 더 큰 표적 상의 활성을 나타내었지만, 다양한 길이의 gRNA의 상대적 특이성은 현재 정해지지 않은 채로 남아있으며, 따라서 더 짧은 gRNA를 사용하는 특정 예에서 바람직할 수 있다. 일부 실시형태에서, gRNA는 전사 시작 부위 상류의 약 100 내지 800bp 내에 있는 영역에 상보적이며, 예를 들어, 전사 시작 부위 상류의 약 500bp 내에 있고, 전사 시작 부위를 포함하거나, 또는 전사 시작 부위 하류의 약 100 내지 800bp 내에, 예를 들어, 약 500bp 내를 포함한다. 일부 실시형태에서, 하나 이상의 gRNA를 암호화하는 벡터(예를 들어, 플라스미드), 예를 들어 동일 영역의 표적 유전자 내에서 상이한 부위로 향하는 2, 3, 4, 5개 이상의 gRNA를 암호화하는 플라스미드가 사용된다.

[0056] Cas9 뉴클레아제는 게놈 DNA 표적 부위의 상보성 가닥에 상보성인 5' 말단에서 17 내지 20개의 nt를 함유하는 안내 RNA, 예를 들어, 단일 gRNA 또는 tracrRNA/crRNA를 이용하여 서열 NGG의 추가적인 근위의 프로토스페이서 인접 모티프(PAM)를 보유하는 특이적 17 내지 20개의 nt 게놈 표적에 안내될 수 있다. 따라서, 본 발명의 방법은 프로토스페이서 인접 모티프(PAM), 예를 들어, NGG, NAG 또는 NNGG의 5' 바로 옆의 표적 서열에 대한 상보성 가닥의, 예를 들어, 25 내지 17, 선택적으로는 20개 이하의 뉴클레오타이드(nt), 예를 들어, 20, 19, 18 또는 17 nt, 바람직하게는 17 또는 18개의 표적 서열에 대해 상보성인 5' 말단에서 서열을 지니는 정상적으로 트랜스-암호화된 tracrRNA, 예를 들어, 문헌[Mali et al., Science 2013 Feb 15; 339(6121):823-6]에 기재된 바와 같은 단일 Cas9 안내 RNA에 융합된 crRNA를 포함하는 단일 안내 RNA의 사용을 포함할 수 있다. 일부 실시형태에서, 단일 Cas9 안내 RNA는 하기 서열로 이루어진다:

[0057] $(X_{17-20})GUUUUAGAGCUAGAAUAGCAAGUAAAAUAAGGCUAGUCCG(X_N)$ (서열번호 1);

[0058] $(X_{17-20})GUUUUAGAGCUAUGCUGAAAAGCAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUC(X_N)$ (서열번호 2);

[0059] $(X_{17-20})GUUUUAGAGCUAUGCUGUUUUGGAAACAAAACAGCAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUC(X_N)$ (서열번호 3);

[0060] $(X_{17-20})GUUUUAGAGCUAGAAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC(X_N)$ (서열번호 4),

[0061] $(X_{17-20})GUUUUAGAGCUAGAAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC$ (서열번호 5);

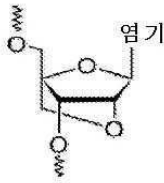
[0062] $(X_{17-20})GUUUUAGAGCUAUGCUGGAAACAGCAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC$ (서열번호 6); 또는

[0063] $(X_{17-20})GUUUUAGAGCUAUGCUGGAAACAGCAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC$ (서열번호 7);

[0064] 여기서, X_{17-20} 은 표적 서열의 17 내지 20개의 연속적 뉴클레오타이드에 상보성인 뉴클레오타이드 서열이다. 단일 안내 RNA를 암호화하는 DNA는 문헌에서 이미 기재되었다(Jinek et al., Science. 337(6096):816-21 (2012) 및

Jinek et al., Elife. 2:e00471 (2013)).

- [0065] 안내 RNA는 Cas9에 대한 리보핵산의 결합을 방해하지 않는 임의의 서열일 수 있는 X_N 을 포함할 수 있으며, 여기서 (RNA에서의) N은 0 내지 200, 예를 들어, 0 내지 100, 0 내지 50, 또는 0 내지 20일 수 있다.
- [0066] 일부 실시형태에서, 안내 RNA는 3' 말단 상에서 하나 이상의 아데닌(A) 또는 유라실(U) 뉴클레오타이드를 포함한다. 일부 실시형태에서, RNA는 RNA PolIII 전사를 종결시키는 종결 신호로서 사용되는 하나 이상의 T의 선택적 존재의 결과로서 분자의 3' 말단에서 하나 이상의 U, 예를 들어, 1 내지 8개 또는 그 이상의 U(예를 들어, U, UU, UUU, UUUU, UUUUU, UUUUUU, UUUUUUU, UUUUUUUU)를 포함한다.
- [0067] 본 명세서에 기재된 일부 예는 단일 gRNA를 이용하지만, 상기 방법은 또한 이중 gRNA(예를 들어, 자연적으로 생기는 시스템 내에서 발견되는 crRNA 및 tracrRNA)와 함께 사용될 수 있다. 이 경우에, 단일 tracrRNA는 본 시스템, 예를 들어, 하기를 이용하여 발현된 다수의 상이한 crRNA 및 tracrRNA 서열과 함께 사용되었다:
- [0068] (X_{17-20})GUUUUAGAGCUA (서열번호 102);
- [0069] (X_{17-20}) GUUUUAGAGCUAUGCUGUUUUG (서열번호 103); 또는
- [0070] (X_{17-20})GUUUUAGAGCUAUGCU (서열번호 104). 이 경우에, crRNA는 본 명세서에 기재된 방법 및 분자에서 안내 RNA로서 사용되며, tracrRNA는 동일 또는 상이한 DNA 분자로부터 발현될 수 있다. 일부 실시형태에서, 상기 방법은 서열 GGAACCAUUCAAAAACAGCAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC(서열번호 8) 또는 이의 활성 부분(활성 부분은 Cas9 또는 dCas9와 복합체를 형성하는 능력을 보유하는 것임)을 포함하거나 또는 이루어진 tracrRNA와 세포를 접촉시키는 단계를 포함한다. 일부 실시형태에서, tracrRNA 분자는 적어도 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35 또는 40개의 nt에 의해 그의 3' 말단으로부터 절단될 수 있다. 다른 실시형태에서, tracrRNA 분자는 적어도 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35 또는 40개의 nt에 의해 그의 5' 말단으로부터 절단될 수 있다. 대안적으로, tracrRNA 분자는 5'과 3' 말단 둘 다로부터, 예를 들어, 5'말단 상에서 적어도 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15 또는 20개의 nt만큼 그리고 3' 말단 상에서 적어도 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35 또는 40개의 nt만큼 절단될 수 있다. 서열번호 8에 추가적인 예시적인 tracrRNA 서열은 다음을 포함한다: UAGCAAGUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC(서열번호 105) 또는 이의 활성 부분; 또는
- [0071] AGCAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC(서열번호 106) 또는 이의 활성 부분.
- [0072] 일부 실시형태에서, (X_{17-20})GUUUUAGAGCUAUGCUGUUUUG(서열번호 102)가 crRNA로서 사용될 때, 다음의 tracrRNA가 사용된다: GGAACCAUUCAAAAACAGCAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC (서열번호 8) 또는 이의 활성 부분.
- [0073] 일부 실시형태에서, when (X_{17-20})GUUUUAGAGCUA(서열번호 102)가 crRNA로서 사용될 때, 다음의 tracrRNA가 사용된다: UAGCAAGUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC (서열번호 105) 또는 이의 활성 부분.
- [0074] 일부 실시형태에서, (X_{17-20}) GUUUUAGAGCUAUGCU(서열번호 104)가 crRNA로서 사용될 때, 다음의 tracrRNA가 사용된다: AGCAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC (서열번호 106) 또는 이의 활성 부분.
- [0075] 일부 실시형태에서, gRNA는 표적을 벗어난 효과를 최소화하기 위해 게놈의 나머지에서 임의의 서열과 상이한 적어도 3개 이상의 미스매치인 부위에 표적화된다.
- [0076] 변형된 RNA 올리고뉴클레오타이드, 예컨대 잠긴 핵산(LNA)은 더 바람직한(안정한) 입체형태에서 변형된 올리고뉴클레오타이드를 잠금으로써 RNA-DNA 혼성화의 특이성을 증가시키는 것으로 입증되었다. 예를 들어, 2'-O-메틸 RNA는 올리고뉴클레오타이드 내로 포함될 때 전반적인 열 안정성 및 선택성을 개선시킬 수 있는 2' 산소와 4' 탄소 사이의 추가적인 공유 결합이 있는 변형된 염기이다(**화학적 I**).



화학식 I - 잠긴 핵산

[0077]

[0078]

따라서, 일부 실시형태에서, 본 명세서에 개시된 tru-gRNA는 하나 이상의 변형된 RNA 올리고뉴클레오타이드를 포함할 수 있다. 예를 들어, 본 명세서에 기재된 절단된 안내 RNA 분자는 표적 서열에 상보적인 안내 RNA 영역 중 하나, 일부 또는 모두를 가질 수 있으며, 이는 변형되고, 예를 들어 (2'-O-4'-C 메틸렌 브릿지), 5'-메틸사이티딘, 2'-O-메틸-슈도유리딘이거나, 또는 리보솜 인산염 백분이 폴리아마이드쇄(펩타이드 핵산), 예를 들어, 합성 리보핵산에 의해 대체되었다.

[0079]

다른 실시형태에서, tru-gRNA 서열의 뉴클레오타이드 중 하나, 일부 또는 모두는 변형될 수 있으며, 예를 들어, 잠긴(2'-O-4'-C 메틸렌 브릿지), 5'-메틸사이티딘, 2'-O-메틸-슈도유리딘이거나, 또는 여기서 리보스 인산염 백분은 폴리아마이드쇄(펩타이드 핵산), 예를 들어 합성 리보핵산에 의해 대체되었다.

[0080]

일부 실시형태에서, 단일 안내 RNA 및/또는 crRNA 및/또는 tracrRNA는 3' 말단 상에서 하나 이상의 아데닌(A) 또는 유라실(U) 뉴클레오타이드를 포함할 수 있다.

[0081]

기존의 Cas9-기반 RGN는 표적화를 관심 대상의 게놈 부위에 안내하기 위한 gRNA-DNA 이형이중가닥(heteroduplex) 형성을 사용한다. 그러나, RNA-DNA 이형이중가닥은 그들의 DNA-DNA 상대보다 더 혼잡한 범위의 구조를 형성할 수 있다. 효과에서, DNA-DNA 이중가닥은 미스매치에 더 민감한데, 이는 DNA-안내된 뉴클레아제가 표적을 벗어난 서열에 용이하게 결합하지 않아서 그들을 RNA-안내 뉴클레아제보다 상대적으로 더 특이적으로 만들 수 있다는 것을 시사한다. 따라서, 본 명세서에 기재된 방법에서 유용한 안내 RNA는 혼성체일 수 있고, 즉, 하나 이상의 데옥시리보뉴클레오타이드, 예를 들어, 짧은 DNA 올리고뉴클레오타이드는 모든 또는 일부의 gRNA, 예를 들어 모든 또는 일부의 gRNA의 상보성 영역을 대체한다. 이 DNA-기반 분자는 단일 gRNA 시스템에서 모든 또는 일부의 gRNA를 대체할 수 있거나 혹은 대안적으로 이중 crRNA/tracrRNA 시스템에서 모든 또는 일부의 crRNA 및/또는 tracrRNA를 대체할 수 있다. DNA를 상보성 영역 내로 혼입시키는 이러한 시스템은 RNA-DNA 이중가닥에 비해 미스매치에 대해 DNA-DNA 이중가닥의 일반적 불용인에 기인하여 의도된 게놈 DNA 서열을 더 신뢰할 수 있게 표적화하여야 한다. 이러한 이중가닥의 제조방법은 당업계에 공지되어 있으며, 예를 들어, 문헌[Barker et al., BMC Genomics. 2005 Apr 22;6:57; 및 Sugimoto et al., Biochemistry. 2000 Sep 19;39(37):11270-81] 참조.

[0082]

추가로, 별도의 crRNA 및 tracrRNA를 사용하는 시스템에서, 하나 또는 둘 다는 합성일 수 있고, 하나 이상의 변형된(예를 들어, 잠긴) 뉴클레오타이드 또는 데옥시리보뉴클레오타이드를 포함한다.

[0083]

세포 내용에서, 이들 합성 gRNA와 Cas9의 복합체는 CRISPR/Cas9 뉴클레아제 시스템의 게놈 전장 특이성을 개선시키기 위해 사용될 수 있었다.

[0084]

본 명세서에 기재된 방법은 세포 내에서 발현시키는 단계 또는 Cas9 gRNA + 본 명세서에 기재된 융합 단백질과 세포를 접촉시키는 단계를 포함할 수 있다.

[0085]

발현 시스템

[0086]

본 명세서에 기재된 융합 단백질 및 안내 RNA를 사용하기 위해, 그들을 암호화하는 핵산으로부터 그들을 발현시키는 것이 바람직할 수 있다. 이는 다양한 방법으로 수행될 수 있다. 예를 들어, 안내 RNA 또는 융합 단백질을 암호화하는 핵산은 복제 및/또는 발현을 위해 원핵 또는 진핵세포로 형질전환하기 위한 중간체 벡터 내로 클로닝될 수 있다. 중간체 벡터는 융합 단백질의 생성을 위해 융합 단백질을 암호화하는 핵산의 저장 또는 조작을 위해, 전형적으로 원핵생물 벡터, 예를 들어, 플라스미드 또는 서틀 벡터 또는 곤충 벡터이다. 안내 RNA 또는 융합 단백질을 암호화하는 핵산은 또한 식물 세포, 동물 세포, 바람직하게는 포유류 세포 또는 인간 세포, 진균 세포, 박테리아 세포 또는 원생동물 세포에 투여를 위해 발현 벡터 내로 클로닝될 수 있다.

[0087]

발현을 얻기 위해, 안내 RNA 또는 융합 단백질을 암호화하는 서열은 전형적으로 전사를 지시하기 위해 프로모터를 함유하는 발현 벡터 내로 서브클로닝된다. 적합한 박테리아 및 진핵생물 프로모터는 당업계에 잘 공지되어

있으며, 예를 들어, 문헌[Sambrook et al., Molecular Cloning, A Laboratory Manual (3d ed. 2001); Kriegler, Gene Transfer and Expression: A Laboratory Manual (1990); 및 Current Protocols in Molecular Biology (Ausubel et al., eds., 2010)]에 기재되어 있다. 유전자 조작된 단백질을 발현시키기 위한 박테리아 발현 시스템은, 예를 들어, 이콜라이(*E. coli*), 바실러스 종(*Bacillus* sp.) 및 살모넬라(*Salmonella*)(Palva et al., 1983, Gene 22:229-235)에서 이용가능하다. 이러한 발현 시스템에 대한 키트는 상업적으로 입수가능하다. 포유류 세포, 효모 및 곤충 세포에 대한 진핵생물 발현 시스템은 당업계에 잘 공지되어 있고, 또한 상업적으로 입수가능하다.

[0088] 핵산의 발현을 지시하기 위해 사용되는 프로모터는 특정 적용에 의존한다. 예를 들어, 강한 구성적 프로모터는 전형적으로 융합 단백질의 발현 및 정제를 위해 사용된다. 대조적으로, 융합 단백질이 유전자 조절을 위해 생체 내에서 투여될 때, 융합 단백질의 특정 용도에 따라서 구성적 또는 유도성 프로모터 중 하나가 사용될 수 있다. 추가로, 융합 단백질의 투여를 위한 바람직한 프로모터는 약한 프로모터, 예컨대 유사한 활성을 갖는 HSV TK 또는 프로모터일 수 있다. 프로모터는 또한 전사촉진(transactivation)에 반응성인 구성요소, 예를 들어, 저산소증 반응 구성요소, Gal4 반응 구성요소, lac 리프레서 반응 구성요소 및 소분자 제어 시스템, 예컨대 테트라사이클린-조절 시스템 및 RU-486 시스템을 포함할 수 있다(예를 들어, 문헌[Gossen & Bujard, 1992, Proc. Natl. Acad. Sci. USA, 89:5547; Oligino et al., 1998, Gene Ther., 5:491-496; Wang et al., 1997, Gene Ther., 4:432-441; Neering et al., 1996, Blood, 88:1147-55; 및 Rendahl et al., 1998, Nat. Biotechnol., 16:757-761] 참조).

[0089] 프로모터에 추가로, 발현 벡터는 전형적으로 원핵생물 또는 진핵생물인 숙주 세포 내 핵산의 발현에 필요한 모든 추가적인 구성요소를 함유하는 전사 단위 또는 발현 카세트를 함유한다. 따라서 전형적인 발현 카세트는, 예를 들어, 융합 단백질을 암호화하는 핵산 서열에 작동가능하게 연결된 프로모터, 및 예를 들어, 전사체의 효율적인 폴리아데닐화, 전사 종결, 리보솜 결합 부위 또는 번역 종결을 위해 필요한 임의의 신호를 함유한다. 카세트의 추가적인 구성요소는, 예를 들어, 인핸서 및 이중성 스플라이싱된 인트론 신호를 포함할 수 있다.

[0090] 세포 내로 유전자 정보를 수송하기 위해 사용되는 특정 발현 벡터는 융합 단백질의 의도된 용도, 예를 들어 식물, 동물, 박테리아, 진균, 원생동물 등에서의 발현과 관련하여 선택된다. 표준 박테리아 발현 벡터는 플라스미드, 예컨대 pBR322 기반 플라스미드, pSKF, pET23D 및 상업적으로 입수가능한 태그-융합 발현 시스템, 예컨대 GST 및 LacZ를 포함한다. 바람직한 태그-융합 단백질은 말토스 결합 단백질(MBP)이다. 이러한 태그-융합 단백질은 공학적으로 조작된 TALE 반복부 단백질의 정제를 위해 사용될 수 있다. 에피토프 태그, 예를 들어, c-myc 또는 FLAG는 발현을 모니터링하며, 세포 및 하위세포 국소화를 모니터링하기 위해 편리한 단리 방법을 제공하기 위하여 제조합 단백질에 첨가될 수 있다.

[0091] 진핵생물 바이러스로부터의 조절 구성요소를 함유하는 발현 벡터는 진핵생물 발현 벡터, 예를 들어, SV40 벡터, 유두종 바이러스 벡터 및 엡스타인-바르 바이러스로부터 유래된 벡터에서 종종 사용된다. 다른 예시적인 진핵생물 벡터는 SV40 초기 프로모터, SV40 후기 프로모터, 메탈로티오네인 프로모터, 무린 유방 종양 바이러스 프로모터, 라우스 육종 바이러스 프로모터, 폴리헤드린 프로모터 또는 진핵세포에서의 발현에 효과적인 것으로 나타난 다른 프로모터의 지시 하에서 단백질을 발현시키는 pMSG, pAV009/A+, pMT010/A+, pMAMneo-5, 바콜로바이러스 pDSVE 및 임의의 다른 벡터를 포함한다.

[0092] 안내 RNA를 발현시키기 위한 벡터는 안내 RNA의 발현을 구동시키기 위한 RNA Pol III 프로모터, 예를 들어, H1, U6 또는 7SK 프로모터를 포함할 수 있다. 이들 인간 프로모터는 플라스미드 형질감염 후에 포유류 세포에서 gRNA를 발현시킨다. 대안적으로, T7 프로모터는, 예를 들어, 시험관내 전사를 위해 사용될 수 있으며, RNA는 시험관 내에서 전사 및 정제될 수 있다. 짧은 RNA, 예를 들어, siRNA, shRNA 또는 다른 소 RNA의 발현에 적합한 벡터가 사용될 수 있다.

[0093] 일부 발현 시스템은 안정하게 형질감염된 세포주의 선택을 위한 마커, 예컨대 티미딘, 키나제, 하이그로마이신 B, 포스포트랜스퍼라제 및 다이하이드로엽산생성효소 환원효소를 가진다. 고수율 발현 시스템, 예컨대 폴리헤드린 프로모터 또는 다른 강한 바콜로바이러스 프로모터의 지시 하에서 융합 단백질 암호화 서열을 지니는 곤충 세포에서의 바콜로바이러스 벡터를 이용하는 것이 또한 적합하다.

[0094] 발현 벡터 내에 전형적으로 포함된 구성요소는 또한 이콜라이에서 기능하는 레플리콘, 제조합 플라스미드를 보유하는 박테리아를 선택을 허용하는 유전자 암호화 항생제 내성 및 제조합 서열을 삽입하게 하는 플라스미드의 비필수 영역 내 독특한 제한 부위를 포함한다.

- [0095] 표준 형질감염 방법은 다량의 단백질을 발현되고, 이어서, 표준 기법을 이용하여 정제되는 박테리아, 포유류, 효모 또는 곤충 세포주를 생성하기 위해 사용된다(예를 들어, 문헌[Colley et al., 1989, J. Biol. Chem., 264:17619-22; Guide to Protein Purification, in Methods in Enzymology, vol. 182 (Deutscher, ed., 1990)] 참조). 진핵생물 및 원핵생물 세포의 형질전환은 표준 기법에 따라 수행된다(예를 들어, 문헌[Morrison, 1977, J. Bacteriol. 132:349-351; Clark-Curtiss & Curtiss, Methods in Enzymology 101:347-362 (Wu et al., eds, 1983)] 참조).
- [0096] 외래의 뉴클레오타이드 서열을 숙주 세포에 도입하기 위한 임의의 공지된 절차가 이용될 수 있다. 이들은 인간 칼슘 형질감염, 폴리브렌(polybrene), 프로토플라스트 융합, 전기천공법, 뉴클레오펙션, 리포솜, 미세주입법, 네이키드 DNA(naked DNA), 플라스미드 벡터, 바이러스 벡터, 에피솜 및 게놈 DNA, cDNA, 합성 DNA 또는 다른 외래 유전자 물질을 숙주 세포 내로 도입하기 위한 통합적이며 임의의 다른 잘 공지된 방법의 사용을 포함한다(예를 들어, 문헌[Sambrook et al.], 상기 참조). 사용되는 특정 유전자 조작 절차는 적어도 하나의 유전자를 선택 단백질을 발현시킬 수 있는 숙주 세포 내로 성공적으로 도입할 유일한 필요가 있다.
- [0097] 일부 실시형태에서, 융합 단백질은 핵에 전위될 단백질을 제공하는 핵 국소화 도메인을 포함한다. 몇몇 핵 국소화 서열(NLS)은 공지되어 있으며, 임의의 적합한 NLS가 사용될 수 있다. 예를 들어, 다수의 NLS는 2부(bipartite) 염기성 반복부로서 지칭되는 복수의 염기성 아미노산을 가진다(문헌[Garcia-Bustos et al, 1991, Biochim. Biophys. Acta, 1071:83-101]에서 검토). 2부 염기성 반복부를 함유하는 NLS는 키메라 단백질의 임의의 부분에 위치될 수 있으며, 키메라 단백질이 핵 내부에 국소화되도록 초래한다. 바람직한 실시형태에서, 핵 국소화 도메인은 최종 융합 단백질 내로 혼입되는데, 본 명세서에 기재된 융합 단백질의 궁극적 기능은 전형적으로 핵 내에서 국소화될 단백질을 필요로 할 것이다. 그러나, DBD 도메인 그 자체 또는 최종 키메라 단백질 내에서 다른 기능성 도메인이 본래의 핵 전위 기능을 갖는 경우에 별도의 핵 국소화 도메인을 첨가할 필요가 없을 수도 있다.
- [0098] 본 발명은 벡터, 및 벡터를 포함하는 세포를 포함한다.
- [0099] 실시예
- [0100] 본 발명은 청구범위에 기재되는 본 발명의 범주를 제한하지 않는 다음의 실시예에서 추가로 기재된다.
- [0101] 실시예 1. 공학적 조작 CRISPR/Cas 활성화 시스템:
- [0102] RNA-안내 전사 활성화제는 촉매적으로 불활성화된 dCas9 단백질의 카복시-말단에 대해 강한 합성 VP64 활성화 도메인을 융합시킴으로써 생성할 수 있다는 가설을 세웠다(Beerli et al., Proc Natl Acad Sci USA 95, 14628-14633 (1998))(도 1d).
- [0103] 인간 세포 내에서 안내 RNA(gRNA)를 발현시키기 위해, U6 프로모터에 의해 구동된 전장 키메라 gRNA(문헌[Jinek et al. (Science 2012)]에 의해 본질적으로 기재된 crRNA 및 tracrRNA의 융합)를 발현시키는 벡터를 공학적으로 조작하였다. gRNA 발현 플라스미드의 구성을 다음과 같이 수행하였다. 가변의 20개 nt gRNA 표적 서열을 암호화하는 DNA 올리고뉴클레오타이드 쌍을 함께 어닐링하여 4bp 돌출부를 지니는 짧은 이중가닥 DNA를 생성하기 위해 함께 어닐링하였다(표 1).

표 1

gRNA 발현 플라스미드를 구성하기 위해 사용한 VEGFA 및 NTF3 유전자 표적 부위 및 관련된 올리고뉴클레오타이드		
gRNA	표적 부위(PAM 포함)	서열번호
V1	GTGTGCAGACGGCAGTCACTAGG	16.
V2	GAGCAGCGTCTTCGAGAGTGAGG	17.
V3	GGTGAGTGAGTGTGTGCGTGTGG	18.
V4	GTTGGAGCGGGGAGAAGGCCAGG	19.
V5	GGGTGGGGGAGTTTGCTCCTGG	20.
V6	GGCTTTGGAAGGGGTGGGGGG	21.
V7	GGGCGGGGTCCCGCGGGGCGG	22.
V8	GCTCGGAGGTCGTGGCGCTGGGG	23.
V9	GACTCACCGGCCAGGGCGCTCGG	24.

V10	GGCGCAGCGGTTAGGTGGACCGG	25.
V11	GGCGCATGGCTCCGCCCCGCCGG	26.
V12	GCCACGACCTCCGAGCTACCCGG	27.
V13	GCGGCGTGAGCCCTCCCCCTTGG	28.
V14	GGAGGCGGGGTGGAGGGGGTCGG	29.
V15	GGGCTACGCCGCGCTCCGGCGG	30.
V16	GACCCCTCCACCCCGCCTCCGG	31.
N1	GAGCGCGGAGCCATCTGGCCGGG	32.
N2	GCGCGCGCGGAAGGGTTAAGG	33.
N3	GCGGCGGCGCGGGCCGGCGGG	34.
N4	GCCGCGCGCCCTCCCCGCCGG	35.
N5	GCGTTATAACCAGCCAACCCGG	36.
N6	GTGCGGAGCTGTTTCGGAAGGG	37.
gRNA	상부 올리고	서열번호
V1	ACACCGTGTGCAGACGGCAGTCACTG	38.
V2	ACACCGAGCAGCGTCTTCGAGAGTGG	39.
V3	ACACCGGTGAGTGAGTGTGTGCGTGG	40.
V4	ACACCGTTGGAGCGGGGAGAAGGCCG	41.
V5	ACACCGGTGGGGGGAGTTTGCTCCG	42.
V6	ACACCGGCTTTGAAAGGGGTGGGG	43.
V7	ACACGGGGCGGGGTCCCGCGGGGG	44.
V8	ACACCGCTCGGAGGTCGTGGCGCTGG	45.
V9	ACACCGACTCACCGCCAGGGCGCTG	46.
V10	ACACCGGCGCAGCGGTTAGGTGGACG	47.
V11	ACACCGGCGCATGGCTCCGCCCCGCG	48.
V12	ACACCGCCACGACCTCCGAGCTACCG	49.
V13	ACACCGCGCGTGAGCCCTCCCCCTG	50.
V14	ACACCGGAGCGGGGTGGAGGGGGTG	51.
V15	ACACGGGCTCACGCCGCGTCCGGG	52.
V16	ACACGACCCCTCCACCCCGCCTCG	53.
N1	ACACGAGCGCGGAGCCATCTGGCCG	54.
N2	ACACCGCGCGCGCGGAAGGGTTAG	55.
N3	ACACCGCGCGCGCGCGGGCCGGCG	56.
N4	ACACCGCGCGCCGCTCCCCGCG	57.
N5	ACACCGCGTTATAACCAGCCAACCG	58.
N6	ACACGTGCGCGGAGCTGTTTCGGAAG	59.
gRNA	하부 올리고	서열번호
V1	AAAACAGTGACTGCCGTCTGCACACG	60.
V2	AAAACCACTCTCGAAGACGCTGCTCG	61.
V3	AAAACACGCACACTACTCACCG	62.
V4	AAAACGGCCTTCTCCCCGCTCCAACG	63.
V5	AAAACGGAGCAAATCCCCCACCCG	64.
V6	AAAACCCACCCCTTTCCAAAGCCG	65.
V7	AAAACCCCGCGGGACCCCGCCCCG	66.
V8	AAAACAGCGCCACGACCTCCGAGCG	67.
V9	AAAACAGCGCCTGGCCGGTGAGTCG	68.
V10	AAAACGTCCACCTAACCGCTGCGCCG	69.
V11	AAAACGCGGGCGGAGCCATGCGCCG	70.
V12	AAAACGGTAGCTCGGAGGTCGTGGCG	71.
V13	AAAACAGGGGAGGGCTACGCCGCG	72.
V14	AAAACACCCCTCCACCCCGCCTCCG	73.

V15	AAAACCCGGAGCGCGCGTGAGCCCG	74.
V16	AAAACGAGGCGGGGTGGAGGGGTCG	75.
N1	AAAACGCCAGATGGCTCCGCGCTCG	76.
N2	AAAACCTAACCCCTTCCGCGCGCGCG	77.
N3	AAAACGCGGCGCCGCGCGCGCGCG	78.
N4	AAAACGCGGGGAGGCGCGCGCGCG	79.
N5	AAAACGGTTGGCTGGTTATAACCGCG	80.
N6	AAAACCTCCGAACAGCTCCGCGCACG	81.

- [0105] 이들 단편을 BsmBI-분해된 플라스미드 pMLM3636에 결합시켜 인간 U6 프로모터에 의해 발현된 키메라 ~102개의 nt 단일체 안내 RNA를 암호화하는 DNA를 획득하였다(Mali et al., Science. 2013 Feb 15;339(6121):823-6; Hwang et al., Nat Biotechnol. 2013 Mar;31(3):227-9). pMLM3636 플라스미드 및 그의 전체 DNA 서열은 애드진(Addgene)으로부터 입수가능하다. 도 4 참조.Cas9-활성제를 유전자 조작하기 위해, D10A, H840A 촉매적 돌연변이(문헌[Jinek et al., 2012; 및 Qi et al., 2013]에 이전에서 기재됨)를 야생형 또는 코돈-최적화된 Cas9 서열 중 하나에 도입하였다(도 5). 이들 돌연변이는 Cas9 촉매적 불활성을 제공하며, 따라서 더 이상 이중-가닥 파손을 유도하지 못할 것이다. 하나의 작제물에서, 삼중 플래그 태그, 핵 국소화 신호 및 VP64 활성화 도메인을 불활성 Cas9의 C-말단에 융합시켰다(도 6). 이 융합 단백질의 발현을 CMV 프로모터에 의해 구동시켰다.
- [0106] dCas-VP64 발현 플라스미드의 구성을 다음에 따라 수행하였다. 불활성화 D10A/H840A 돌연변이를 보유하는 Cas9 뉴클레아제(dCas9)를 암호화하는 DNA를 Cas9 암호 서열의 카복시-말단의 단부에서 시작 코돈 및 핵 국소화 신호에 대해 T7 프로모터 부위 5'을 첨가한 프라이머를 사용하여 플라스미드 pMJ841(애드진 플라스미드 #39318)으로부터 PCR에 의해 증폭시켰고, 앞서 기재한 바와 같은 CMV 프로모터를 함유하는 플라스미드 내로 클로닝시켜서(Hwang et al., Nat Biotechnol 31, 227-229 (2013)) 플라스미드 pMLM3629를 획득하였다. 삼중 FLAG 에피토프를 암호화하는 올리고뉴클레오타이드를 어닐링시키고 나서, 플라스미드 pMLM3629 내 XhoI 및 PstI 부위 내로 클로닝시켜 C-말단의 플래그 FLAG 태그를 지니는 dCas9를 발현시키는 플라스미드 pMLM3647을 생성하였다. Gly₄Ser 링커 다음에 합성 VP64 활성화 도메인을 암호화하는 DNA 서열을 플라스미드 pMLM3647 내 FLAG-태그된 dCas9의 하류에 도입하여 플라스미드 pSL690을 획득하였다. D10A/H840A 돌연변이를 또한 인간 세포 내에서 발현을 위해 코돈 최적화된 FLAG-태그된 Cas9 서열을 암호화하는 플라스미드 pJDS247 내로 퀵체인지(QuikChange) 부위-지정 돌연변이유발(애질런트(Agilent))에 의해 도입하여 플라스미드 pMLM3668을 획득하였다. 이어서, Gly₄Ser 링커 및 VP64 활성화 도메인을 암호화하는 DNA 서열을 pMLM3668 내로 클로닝시켜 pMLM3705로 칭해지는 코돈-최적화된 dCas9-VP64 발현 벡터를 획득하였다.
- [0107] 세포 배양물, 형질감염 및 ELISA 분석을 다음과 같이 수행하였다. Flp-In T-Rex 293 세포를 10% FBS, 1% 펜스트랩(penstrep) 및 1% 글루타맥스(Glutamax)(인비트로젠(Invitrogen))로 보충한 어드벤스트 DMEM(Advanced DMEM)에서 유지하였다. 세포를 제조업자의 지시에 따라 리포펙타민(Lipofectamine) LTX(인비트로젠)에 의해 형질감염시켰다. 간단하게, 160,000개의 293 세포를 24-웰 플레이트 내에 파종하고 나서, 다음날 250ng gRNA 플라스미드, 250ng Cas9-VP64 플라스미드, 30ng pmaxGFP 플라스미드(론자(Lonza)), 0.5 μ l 플러스 시약(Plus Reagent) 및 1.65 μ l 리포펙타민 LTX와 함께 형질감염시켰다. 형질감염시킨 293 세포로부터의 조직 배양 배지를 형질감염 후 40시간에 채취하고 나서, 알앤디 시스템즈(R&D System's)의 인간 VEGF-A ELISA 키트 "인간 VEGF 면역분석"을 이용하여 분석한 VEGF-A 단백질을 분비시켰다.
- [0108] 16개의 sgRNA를 293 세포 내 인간 VEGFA 유전자의 전사 시작 부위의 상류, 하류 또는 전사 시작 부위에 위치한 3개의 DNase I 초-민감 부위(HSS) 내의 표적 서열에 대해 구성하였다(도 1e).
- [0109] 신규한 dCas9-VP64 융합 단백질을 동원하기 위해 16개의 VEGFA-표적화된 gRNA의 능력을 시험하기 전에, 각각의 이들 gRNA를, 인간 293 세포 내에서 그의 의도된 표적 부위로 Cas9 뉴클레아제를 보내는 그의 능력에 대해 우선 평가하였다. 이 목적을 위해, gRNA 및 Cas9 발현 벡터를 1:3 비로 형질감염시켰는데, 이전의 최적화 실험이 플라스미드의 이런 비를 사용하여 U2OS 세포 내 고수준의 Cas9-유도 DNA 절단을 입증하였기 때문이다.
- [0110] VEGFA-표적화된 gRNA를 암호화하는 125ng의 플라스미드 및 375 ng의 활성 Cas9 뉴클레아제를 암호화하는 플라스미드(pMLM3639)로 형질감염시킨 세포를 제외하고, dCas9-VP16 VEGFA 실험에 대해 상기 기재한 바와 같이 293 세포의 형질감염을 수행하였다. 형질감염 후 40시간에, 제조업자의 지침에 따라 QIAamp DNA 블러드 미니 키트(DNA

Blood Mini kit)(퀴아젠)를 이용하여 게놈 DNA를 단리시켰다. 3% DMSO 및 다음의 터치다운 PCR 주기를 이용하는 퓨전 핫 스타트 II(Phusion Hot Start II) 고충실도 DNA 증합효소(NEB)를 이용하여 *VEGFA* 프로모터 내 3개의 상이한 표적화된 영역의 PCR 증폭을 수행하였다: 10주기의 98℃, 10초; 72 내지 62℃, -1℃/주기, 15초; 72℃, 30초, 다음에 25주기의 98℃, 10초; 62℃, 15초; 72℃, 30초. 프라이머 oFYF434(5'-TCCAGATGGCACATTGTCAG-3'(서열번호 82)) 및 oFYF435(5'-AGGGAGCAGGAAAGTGAGGT-3'(서열번호 83))를 이용하여 -500 영역을 증폭시켰다. 프라이머 oFYF438(5'-GCACGTAACCTCACTTTCCT-3'(서열번호 84)) 및 oFYF439(5'-CTTGCTACCTCTTTCCTCTTCT-3'(서열번호 85))을 이용하여 전사 시작 부위 주위의 영역을 증폭시켰다. 프라이머 oFYF444(5'-AGAGAAGTCGAGGAAGAGAGAG-3'(서열번호 86)) 및 oFYF445(5'-CAGCAGAAAGTTCATGGTTTCG-3'(서열번호 87))를 이용하여 +500 영역을 증폭시켰다. 앰퓨어 XP(Ampure XP) 비드(에이전코트)를 이용하여 PCR 산물을 정제하고 나서, T7 엔도뉴클레아제 I 분석을 수행하였고, 앞서 기재한 바와 같이 퀴악셀(QIAXCEL) 모세관 전기이동 시스템 상에서 분석하였다(Reyon et al., Nat Biotech 30, 460-465 (2012)).

[0111] 모두 16개의 gRNA는 앞서 기재한 T7E1 게노타이핑 분석을 이용하여 평가한 바와 같이 각각의 표적 부위에서 Cas9 뉴클레아제-유도 삽입결실 돌연변이의 효율적인 도입을 매개할 수 있었다(표 2). 따라서, 모두 16개의 gRNA를 Cas9 뉴클레아제와 복합체화되고, 인간 세포에서 특이적 표적 게놈에 대한 그의 활성을 지시할 수 있다.

표 2

VEGFA- 표적화된 gRNA 및 Cas9 뉴클레아제에 의해 유도된 삽입결실 돌연변이의 빈도

gRNA	평균 삽입결실 돌연변이 빈도 (%) ± SEM
V1	18.05 ± 0.47
V2	41.48 ± 0.62
V3	33.22 ± 1.05
V4	16.97 ± 0.06
V5	7.46 ± 0.50
V6	16.99 ± 0.51
V7	1.42 ± 0.11
V8	34.07 ± 0.90
V9	24.53 ± 1.40
V10	35.65 ± 1.35
V11	4.45 ± 0.22
V12	23.95 ± 0.41
V13	9.45 ± 0.74
V14	12.17 ± 0.36
V15	14.28 ± 0.54
V16	18.82 ± 1.48

[0112]

[0113] dCas9-VP64 단백질이 이들 동일한 gRNA에 의해 인간 세포 내 특정 게놈 부위에 표적화될 수 있었는지의 여부를 시험하기 위해, *VEGFA* 단백질의 효소-결합 면역 블롯 분석을 다음과 같이 수행하였다. *VEGFA*-표적화된 sgRNA 및 dCas9-VP64를 암호화하는 플라스미드로 형질감염시킨 Flp-In T-Rex HEK293 세포의 배양 배지를 형질감염 후 40 시간에 채취하였고, *VEGFA* 단백질 발현을 앞서 기재한 바와 같이 측정하였다(Maeder et al., Nat Methods 10, 243-245 (2013)). *VEGFA* 발현의 배수적 활성화를 sgRNA와 dCas9-VP64를 둘 다 발현시킨 세포로부터의 배지 내 *VEGFA* 단백질의 농도를 표적을 벗어난 sgRNA(*EGFP* 리포터 유전자 내 서열에 대해 표적화됨)와 dCas9-VP64가 발현된 세포로부터의 배지 내에서의 *VEGFA* 단백질의 농도에 의해 나눔으로써 *VEGFA* 발현의 배수적 활성화를 계산하였다.

[0114] 시험한 16개의 gRNA 중 15개는 인간 293 세포 내에서 dCas9-VP64와 공동발현시켰을 때 *VEGFA* 단백질 발현에서의 상당한 증가를 유도하였다(도 2a). 관찰한 *VEGFA* 유도의 규모는 2 내지 18.7의 배수적 활성화의 범위(평균 5배 활성화)에 있었다. 대조군 실험은 각각 "표적을 벗어난" gRNA와 함께 16개 gRNA 단독, dCas9-VP64 단독 및 dCas9-VP64의 발현이 상승된 *VEGFA* 발현을 유도하지 못한 모든 *EGFP* 리포터 유전자 서열에 결합하도록 설계되었

다는 것을 나타내었는데(도 2a), 특이적 gRNA와 dCas9-VP64 단백질의 공동 발현은 둘 다 프로모터 활성화에 필요하다라는 것을 입증한다. 따라서, dCas9-VP64는 안정하게 발현되며, 인간 세포 내 특정 게놈 좌위의 전사를 활성화시키기 위해 gRNA에 의해 지시될 수 있다. VEGFA에서의 가장 큰 증가를 18.7배만큼 단백질 발현을 유도한 gRNA3으로 형질감염시킨 세포 내에서 관찰하였다. 흥미롭게도, 3종의 최상의 gRNA 및 3배 이상만큼 발현을 유도할 수 있는 9종의 gRNA 중 6종은 -500 영역(전사 시작 부위 상류의 ~500bp)을 유도할 수 있다.

[0115] 일 양태에서 본 명세서에 기재된 시스템은 통상적인 dCas9-VP64 활성화제 융합을 동원하기 위해 가변적 gRNA를 사용하기 때문에, 단일 세포 내에서 다수의 안내 RNA의 발현은 내인성 유전자 표적의 다중복합 또는 조합적 활성화를 가능하게 할 수 있었다. 이 가능성을 시험하기 위해, 293 세포를 VEGFA 프로모터로부터의 발현을 각각 개별적으로 유도한 4종의 gRNA(V1, V2, V3 및 V4)에 대한 발현 플라스미드와 함께 dCas9-VP64 발현 플라스미드로 형질감염시켰다. dCas9-VP64와 함께 모두 4종의 gRNA의 공동발현은 VEGFA 단백질 발현의 상승적 활성화를 유도하였다(즉, 각각의 개개 활성화제의 예상된 상가적 효과보다 더 큰 배수적 활성화)(도 2b). 추가로, 이들 4종의 활성화제 중 3종의 다양한 조합은 또한 VEGFA 프로모터를 상승적으로 활성화시켰다(도 2b). 전사의 상승적 활성화는 단일 프로모터에 다수의 활성화제 도메인의 동원으로부터 초래되는 것으로 믿어지기 때문에, 다수의 gRNA/dCas9-VP64 복합체는 이들 실험에서 VEGFA 프로모터에 동시에 결합될 가능성이 있다.

[0116] 이들 실험은 Cas9-HFD, 예를 들어, Cas9-활성제 단백질(VP64 전사 활성화 도메인을 보유)과 인간 HEK293 세포 내 인간 VEGF-A 프로모터 내 부위에 대해 20개 nt의 서열 상보성을 지니는 sgRNA의 공동 발현이 VEGF-A 발현의 상향조절을 초래할 수 있다는 것을 입증한다. VEGF-A 단백질에서의 증가를 ELISA 분석에 의해 측정하였고, 개개 gRNA가 Cas9-활성제 융합 단백질과 함께 작용하여 ~18배까지만큼 VEGF-A 단백질 수준을 증가시킬 수 있다는 것을 발견하였다(도 2a). 추가적으로, Cas9-활성제 융합 단백질과 함께 동일한 프로모터 내의 다양한 부위를 표적화하는 다수의 gRNA를 도입함으로써 전사 상승효과를 통해 심지어 활성화에서의 더 큰 증가를 달성할 수 있었다(도 2b).

[0117] 실시예 2. 내인성 인간 *NTF3* 유전자를 표적화하는 공학적 조작 CRISPR/Cas 활성화제 시스템

[0118] 본 발견의 일반성을 연장하기 위해, 본 발명자들은 RNA-안내 활성화제 플랫폼이 인간 *NTF3* 유전자의 발현을 유도하는데 사용될 수 있었는지의 여부를 시험하였다. 이를 행하기 위해, 인간 *NTF3* 프로모터 내에서 예측된 DNase I 초민감 부위(hypersensitive site: HSS)에 대해 6종의 sgRNA를 설계하였고, 각각의 이들 gRNA를 발현시키는 플라스미드를 인간 세포 발현을 위해 코돈 최적화된 dCas9-VP64 단백질을 암호화하는 플라스미드와 함께 공동형질감염시켰다(도 3a).

[0119] 시험한 모두 6종의 gRNA는 정량적 RT-PCR에 의해 검출되는 바와 같이 *NTF3* 전사체 수준에서의 상당한 증가를 유도하였다(도 3b). 이들 6종의 RNA-안내 활성화제에 대한 배수적 활성화 값을 정확하게 계산할 수 없었지만(전사체의 기저 수준은 본질적으로 검출가능하지 않았기 때문), 활성화된 *NTF3* mRNA 발현의 평균 수준은 4배 범위에 걸쳐 변화하였다. 형질감염시킨 gRNA 및 dCas9-VP64 발현 플라스미드의 양을 감소시키는 것은 *NTF3* 유전자의 더 적은 활성화를 초래하였는데(도 3b), 이는 분명한 용량-의존적 효과를 입증한다.

[0120] 추가로, 293 세포를 dCas9-VP64 및 *NTF3*-표적화된 gRNA 발현 플라스미드 단독으로, 단일 및 이중 조합물로 공동형질감염시켰다. *NTF3* mRNA의 상대적 발현을 정량적 RT-PCR에 의해 검출하고 나서, GAPDH 대조군(델타Ct x 10⁴)에 대해 정규화시켰다. 모든 이들 실험에서, 형질감염에 대해 사용한 각각의 개개 gRNA 발현 플라스미드의 양은 동일하였다. 도 3b는 이 다중복합 gRNA 발현이 dCas9-VP64 단백질에 의한 *NTF3* mRNA 발현의 상승적 활성화를 유도하였다는 것을 나타낸다.

[0121] 실시예 3. CRISPR/Cas-MS2, -Csy4 및 -람다 N 융합 시스템의 공학적 조작 - 생물학적 테더의 생성

[0122] MS2 외피 단백질, Csy4 뉴클레아제(바람직하게는 촉매적으로 불활성인 Csy4, 예를 들어, 문헌[Haurwitz et al. 329(5997):1355-8 (2010)]에 기재된 H29A 돌연변이체) 또는 람다 N가 불활성화된 dCas9의 N- 또는 C-말단에 융합된 융합 단백질을 생성하였다. MS2 및 람다 N은 특이적 RNA 서열에 결합하는 박테리오파지 단백질이며, 따라서 dCas9 단백질을 특이적 MS2 또는 람다 N RNA 결합 서열로 태그된 이중성 RNA 서열에 테더링하기 위한 연결자로서 사용될 수 있다. dCas9-MS2 융합 또는 dCas9-람다 N 융합을 5' 또는 3' 말단 상에서 MS2 또는 람다 N 스템 루프 인식 서열에 융합되는 키메라의 긴 비-암호 RNA(lncRNA)와 함께 공동발현시켰다. 키메라 Xist 또는 키메라 RepA lncRNA는 dCas9 융합에 의해 특이적으로 동원될 것이며, 표적화된 침묵을 유도하기 위한 이 전략의 능력을 표적 유전자 발현을 측정함으로써 분석할 것이다. 외피 단백질 및 키메라 RNA에 대해 다양한 변형을 시험함으로써 시스템을 최적화할 것이다. MS2 외피 단백질에 대한 N55K 및 델타FG 돌연변이는 단백질 응집을 예방하고 스

템-루프 RNA에 대한 친화도를 증가시키는 것으로 이전에 입증되었다. 추가적으로, 본 발명자들은 MS2 외피 단백질에 대한 친화도를 증가시키는 것으로 보고된 고친화도 C-루프 RNA 돌연변이체를 시험할 것이다. MS2 및 람다 N 단백질에 대한 예시적 서열을 이하에 제공한다; MS2는 이량체로서 작용하며, 따라서 MS2 단백질은 융합된 단일체 이량체 서열을 포함할 수 있다.

[0123] 1. dCas9의 N-말단 또는 C-말단에 대해 단일 MS2 외피 단백질(wt, N55K 또는 델타FG)의 융합을 위한 예시적 서열.

MS2 외피 단백질 아미노산 서열:

MASNFTQFVLVDNGGTGDVTVPSPNFANGVAEWISSNSRSQAYKVTCSVRQSSAQ
NRKYTIKVEVPKVATQTVGGVELPVAAWRSYLNMEITPIFATNSDCELVKAMQG
LLKDGNIPIPSAIAANSIGY (서열번호 88)

MS2 N55K:

MASNFTQFVLVDNGGTGDVTVPSPNFANGVAEWISSNSRSQAYKVTCSVRQSSAQ
KRKYTIKVEVPKVATQTVGGVELPVAAWRSYLNMEITPIFATNSDCELVKAMQG
LLKDGNIPIPSAIAANSIGY (서열번호 89)

MS2 델타FG:

MASNFTQFVLVDNGGTGDVTVPSPNFANGIAEWISSNSRSQAYKVTCSVRQSSAQ
NRKYTIKVEVPKGAWRSYLNMEITPIFATNSDCELVKAMQGLLKDGNIPIPSAIAA
NSIGY (서열번호 90)

[0124]

[0125] 2. dCas9의 N-말단 또는 C-말단에 대해 융합된 이량체 MS2 외피 단백질(wt, N55K 또는 델타FG)의 융합에 대한 예시적 서열.

이량체 MS2 외피 단백질:

MASNFTQFVLVDNGGTGDVTVPSPNFANGVAEWISSNSRSQAYKVTCSVRQSSAQ
NRKYTIKVEVPKVATQTVGGVELPVAAWRSYLNMEITPIFATNSDCELVKAMQG
LLKDGNIPIPSAIAANSGLYGAMASNFTQFVLVDNGGTGDVTVPSPNFANGVAEWI
SSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKVATQTVGGVELPVAAWRSYLN
MEITPIFATNSDCELVKAMQGLLKDGNIPIPSAIAANSLIN (서열번호 91)

[0126]

MASNFTQFVLVDNGGTGDVTVPSPNFANGVAEWISSNSRSQAYKVTCSVRQSSAQ
KRKYTIKVEVPKVATQTVGGVELPVAAWRSYLNMEITPIFATNSDCELVKAMQG
LLKDGNIPIPSAIAANSGLYGAMASNFTQFVLVDNGGTGDVTVPSPNFANGVAEWI
SSNSRSQAYKVTCSVRQSSAQKRKYTIKVEVPKVATQTVGGVELPVAAWRSYLN
MEITPIFATNSDCELVKAMQGLLKDGNIPIPSAIAANSLIN (SEQ ID NO:92)

이량체 MS2델타FG:

MASNFTQFVLVDNGGTGDVTVPSPNFANGVAEWISSNSRSQAYKVTCSVRQSSAQ
KRKYTIKVEVPKGAWRSYLNMEITPIFATNSDCELVKAMQGLLKDGNIPIPSAIAA
NSGLYGAMASNFTQFVLVDNGGTGDVTVPSPNFANGVAEWISSNSRSQAYKVTCS
VRQSSAQKRKYTIKVEVPKGAWRSYLNMEITPIFATNSDCELVKAMQGLLKDGNI
PIPSAIAANSLIN (서열번호 93)

[0127]

[0128] 3. dCas9의 N-말단 또는 C-말단에 대해 람다 N의 융합을 위한 예시적 서열.

람다 N 아미노 서열:

MDAQTRRRERRAEKQAQWKAAN (서열번호 94) 또는
MDAQTRRRERRAEKQAQWKAANPLLGVSAKPVNRPILSLNRKPKSRVESALNPI
DLTVLAHEYHKQIESNLQRIERKNQRTWYSKPGERGITCSGRQKIKGKSIPLI
(서열번호 95)

[0129]

[0130] 4. dCas9의 N-말단 또는 C-말단에 대해 Csy4의 융합을 위한 예시적 서열

[0131] Cys4에 대한 예시적 서열을 문헌[Haurwitz et al. 329(5997):1355-8 (2010)]에서, 예를 들어, 불활성화 형태로 제공한다.

[0132] 작제물을 또한 5' 또는 3' 말단 중 하나에서 람다 N 또는 MS2에 대해 동족 스템-루프 인식 서열과 융합된 조절 RNA, 예를 들어 긴 비-암호 RNA(lncRNA), 예컨대 HOTAIR, HOTTIP, XIST 또는 XIST RepA를 발현시키는 세포에서 발현시켰다. MS2에 대한 야생형 및 고친화도 서열은 각각 AAACAUGAGGAUACCCAUGUCG(서열번호 96) 및 AAACAUGAGGAUACCCAUGUCG(서열번호 97)(문헌[Keryer-Bibens et al.], 상기 참조, 도 2); 람다 N이 결합하는

nutL 및 nutR BoxB 서열은 각각 GCCCUGAAGAAGGGC(서열번호 98) 및 GCCCUGAAAAAGGGC(서열번호 99)이다. Csy4가 결합하는 서열은 GTTCACTGCCGTATAGGCAG(절단된 20개의 nt)(서열번호 100) 또는 GUUCACUGCCGUAGGCAGCUAAGAAA(서열번호 101)이다.

[0133] MS2-결합 서열 태그된 lncRNA를 발현시키는 세포 내에서 표적 부위에 대한 dCas9/MS2의 결합은 dCas9 결합 부위에 대해 해당 lncRNA를 동원하며; lncRNA가 리프레서, 예를 들어 XIST인 경우, dCas9 결합 부위 근처의 유전자는 억제된다. 유사하게, 람다N-결합 서열 태그된 lncRNA를 발현시키는 세포 내에서 표적 부위에 대한 dCas9/람다N의 결합은 dCas9 결합 부위에 대해 lncRNA를 동원한다.

[0134] **실시예 4. CRISPR/Cas-HP1 융합 시스템-서열-특이적 침묵의 공학적 조작**

[0135] 본 명세서에 기재된 dCas9 융합 단백질은 표적 침묵 도메인, 예를 들어, 이질염색질 단백질 1(HP1, 또한 swi6으로서 알려짐), 예를 들어, HP1 α 또는 HP1 β 에 사용될 수 있다. 크로모도메인이 제거된 HP1 α 또는 HP1 β 의 절단된 형태를 이질염색질 형태 및 유전자 침묵을 유도하기 위해 특이적 좌위에 표적화시킬 수 있다. dCas9에 융합된 절단된 HP1의 예시적인 서열을 도 8a 내지 도 8b에 나타낸다. HP1 서열을 상기 기재한 바와 같은 불활성화된 dCas9의 N- 또는 C-말단에 융합시킬 수 있다.

[0136] **실시예 5. CRISPR/Cas-TET 융합 시스템-서열-특이적 탈메틸화의 공학적 조작**

[0137] 본 명세서에 기재된 dCas9 융합 단백질을 또한 DNA의 메틸화 상태를 변형시키는 효소(예를 들어, DNA 메틸트랜스퍼라제(DNMT) 또는 TET 단백질)를 표적화하기 위해 사용할 수 있다. TET1의 절단 형태를 DNA 탈메틸화를 촉매하기 위해 특이적 좌위에 표적화할 수 있다. dCas9에 융합된 절단된 TET1의 예시적 서열을 도 9에 나타낸다. TET1 서열을 상기 기재한 바와 같이 불활성화된 dCas9의 N- 또는 C-말단에 융합시킬 수 있다.

[0138] **실시예 6. 최적화된 CRISPR/Cas-VP64 융합의 공학적 조작**

[0139] VP64 활성화 도메인을 보유하는 dCas9-기반 전사 활성제의 활성을 핵 국소화 신호(들)(NLS) 및 3xFLAG-태그의 수 및 위치를 달리함으로써 최적화시켰다(도 10). N-말단의 NLS와, dCas9와 VP64 서열 사이에 놓인 NLS를 둘 다 함유하는 dCas9-VP64 융합은 활성제의 향상된 핵 국소화로부터 초래될 가능성이 있는 고수준의 표적 유전자 활성화를 지속적으로 유도한다(도 10). 더 나아가, 3xFLAG 태그가 dCas9 C-말단과 VP64의 N-말단에 위치될 때, 훨씬 더 큰 수준의 활성화를 관찰하였다. 3xFLAG 태그는 인공적 링커로서 작용하여 dCas9와 VP64 사이의 필요한 간격을 제공하고, VP64 도메인의 더 양호한 폴딩(근처의 dCas9를 제약할 때 가능하지 않을 수도 있음) 또는 RNA 중합효소 II를 동원하는 전사 매개체 복합체에 의한 VP64의 더 양호한 인식을 제공할 수 있다. 대안적으로, 음으로 하전된 3xFLAG 태그는 또한 우연한 전사 활성화 도메인으로서 작용하여 VP64 도메인의 효과를 향상시킬 수 있다.

[0140] **실시예 7. 최적화된 촉매적으로 불활성인 Cas9 단백질(dCas9)**

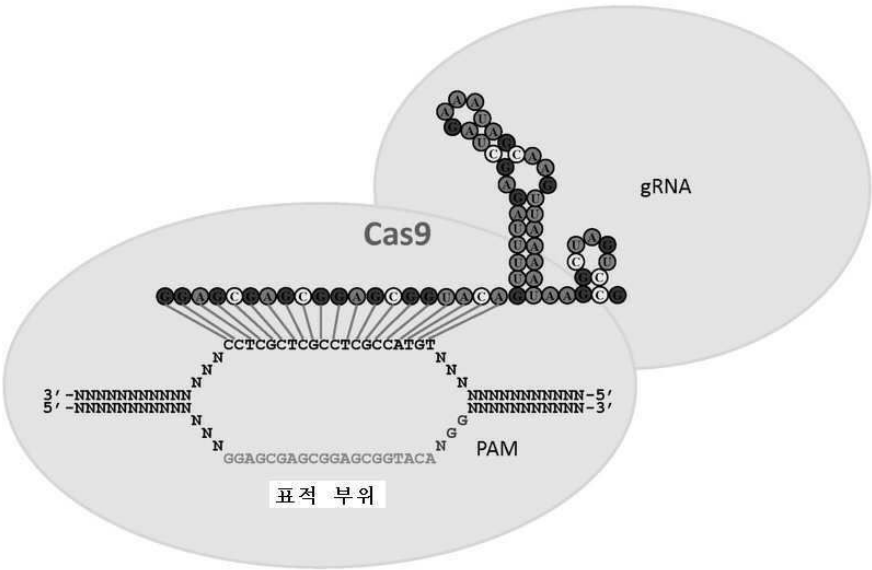
[0141] dCas9-VP64 활성제의 활성의 추가적인 최적화를 dCas9 도메인 내 Cas9의 뉴클레아제 활성을 없애는 불활성화 돌연변이의 특성을 변화시킴으로써 수행하였다(도 11A 내지 도 11B). 지금까지 공개된 연구에서, 촉매적 잔기 D10 및 H840을 알라닌으로 돌연변이시켜(D10A 및 H840A) DNA의 가수분해를 매개하는 활성 부위 네트워크를 붕괴시켰다. 이들 위치에서의 알라닌 치환은 dCas9의 탈안정화 및 따라서 차선의 활성을 초래할 수 있는 것으로 가설을 세웠다. 따라서, D10 또는 H840에서 더 구조적으로 보존적인 치환(예를 들어, 아스파라긴 또는 타이로신 잔기에 대해: D10N, H840N 및 840Y)을 시험하여 그들이 이들 상이한 돌연변이를 보유하는 dCas9-VP64 융합에 의해 더 큰 유전자 활성화를 야기할 수 있는지의 여부를 알게 되었다. 이들 변이체 치환을 보유하는 dCas9-VP64 변이체를 내인성 인간 VEGFA 유전자의 상류 영역을 표적화하는 3개의 gRNA와 함께 HEK293 세포 내로 공동형질감염시켰을 때, 이들 변이체 중 하나를 제외한 모두에 대해 큰 VEGFA 단백질 발현을 관찰하였다(도 11A). 그러나, 이 효과는 dCas9-VP64 변이체가 이들 gRNA 중 단지 하나와 함께 공동형질감염되었을 때(도 11A) 또는 단일 VEGFA-표적화된 gRNA를 발현시키는 HEK293 유도체 세포주 내로 형질감염되었을 때(도 11B) 유의하지 않았다.

[0142] **다른 실시형태**

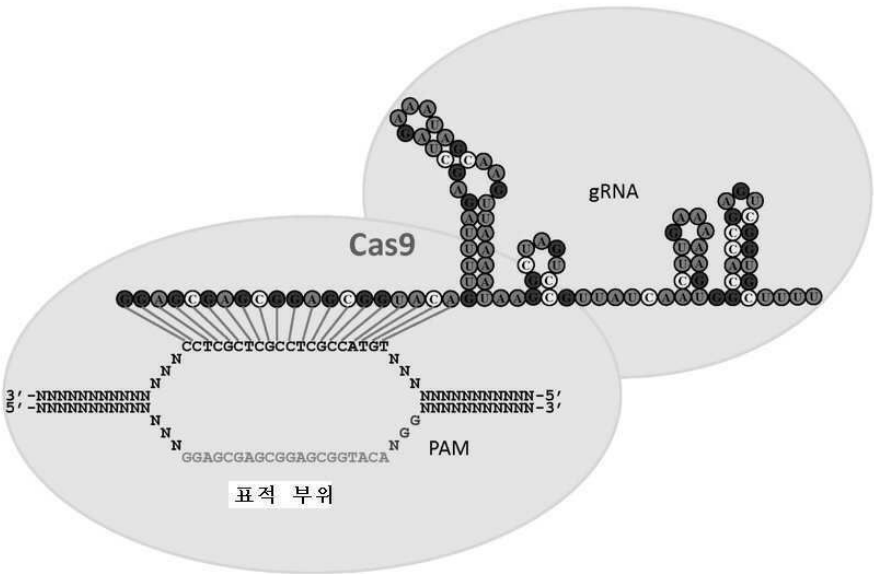
[0143] 본 발명이 이의 상세한 설명과 함께 기재되었지만, 앞서 언급한 설명은 예시적인 것이며 첨부하는 청구범위의 범주에 의해서 정해지는 본 발명의 범주를 제한하지 않는 것으로 의도된다는 것이 이해되어야 한다. 다른 양태, 이점 및 변형이 다음의 청구범위의 범주 내에 있다.

도면

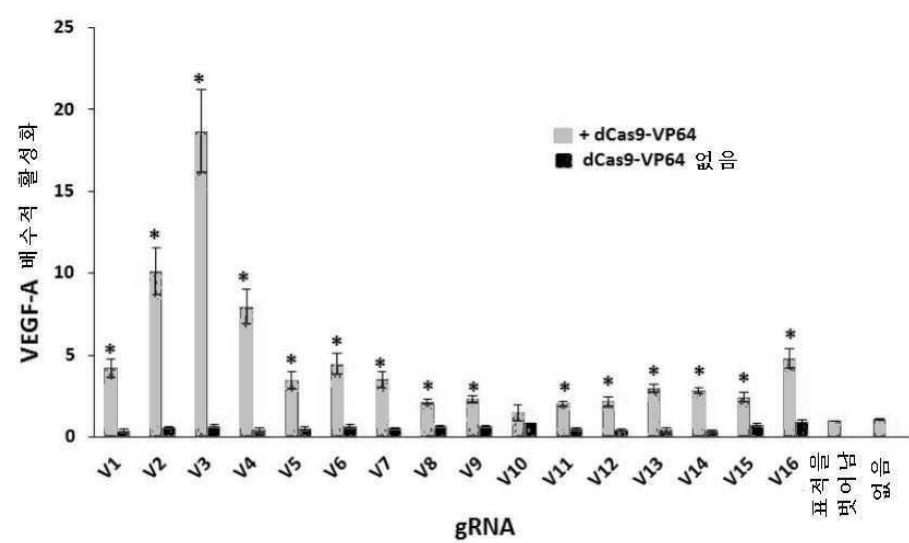
도면1a



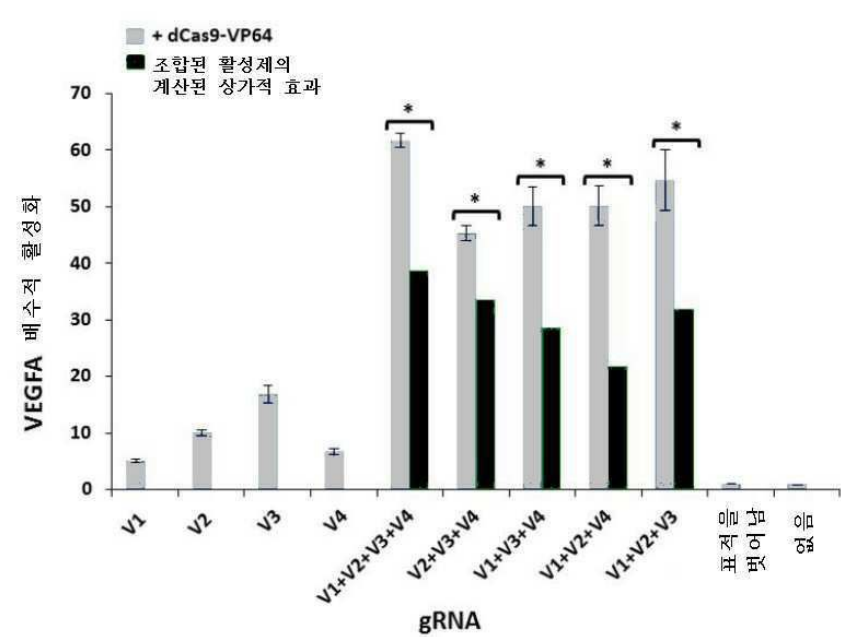
도면1b



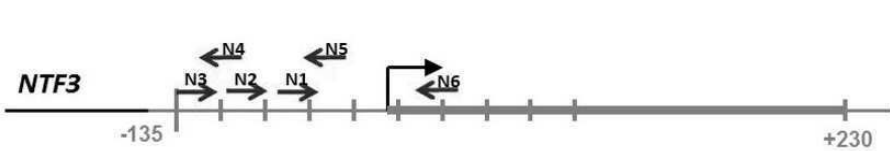
도면2a



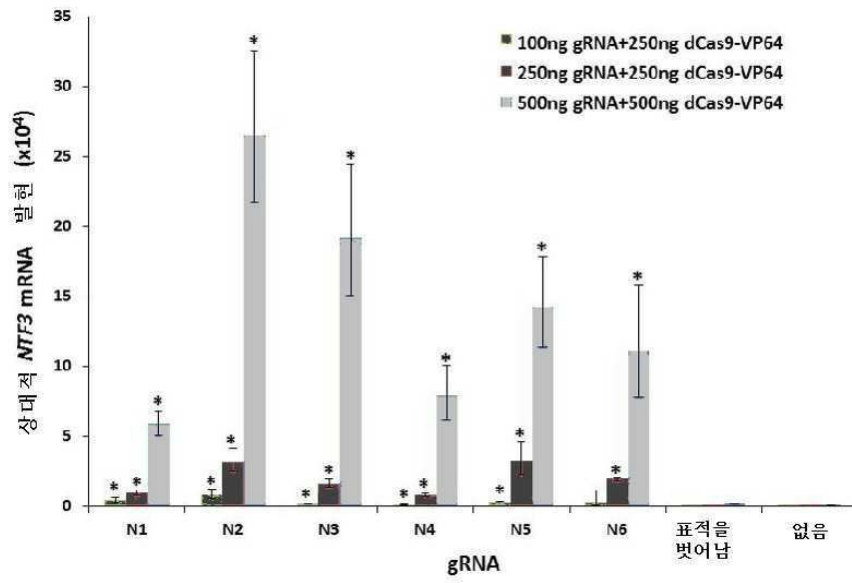
도면2b



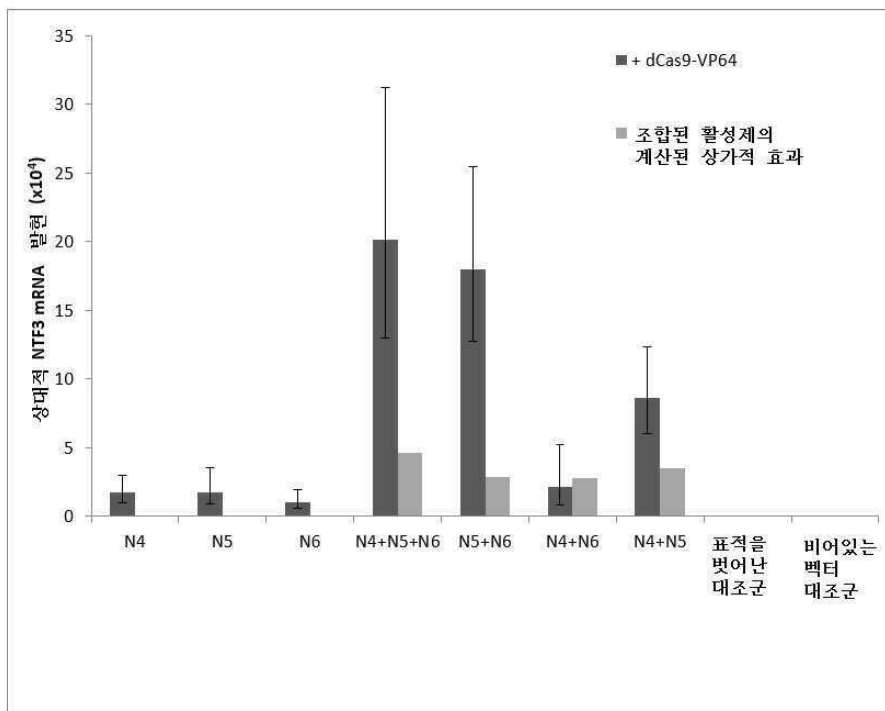
도면3a



도면3b



도면3c



도면4

안내 RNA 발현 벡터 서열
GACGTCGCTAGTGTACAAAAAGCAGGCTTTAAAGGAACCAATTTCAGTCGACTGGATCCGGTACCAA GGTCGGGCAGGAAGAGGGCCTATTTCCCATGATTCCTTCATATTTGCATATACGATACAAGGCTGTTA GAGAGATAATTAGAATTAATTTGACTGTAAACACAAAGATATTAGTACAAAATACGTGACGTAGAAAG TAATAATTTCTTGGGTAGTTTGCAGTTTAAAAATTATGTTTTAAATGGACTATCATATGCTTACCGT AACTTGAAAGTATTTGATTTCTTGCTTTATATATCTTGTGAAAGGACGAAACACCCNNNNNNNNNN NNNNNNNNNGTTTTAGAGCTAGAAATAGCAAGTTAAAAATAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGCTTTTTTAAAGCTTGGCCGCTCGAGGTACCTCTCTACATATGACATGTGA GCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTGTCTGGCGTTTTTCCATAGGCTCCG CCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAA GATACCAGGCGTTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGA TACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAG TTCCGTTGATAGTTCGTTCCGCTCAAGCTGGGCTGTGTGCAGAACCCCGTTTACGCCCCGACCGCTGCG CCTTATCCGTTAACTATCGTCTTGATCCAAACCCGTTAAGACACGACTTATCGCCACTGGCAGCAGCC ACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAA CTACGGGTACACTAGAGAAGCAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGA AAAA GAGTTGGTAGCTCTTGATCCGGCAAAACAAACCCGCTGGTAGCGGTGGTTTTTTTGTTCAGAGCAG CAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCA GTGAACGAAAACTCAGCTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCC TTTTAAATTA AAAATGAAGTTTAAATCAATCTAAAGTATATATGAGTAACTTGGTCTGACAGTTAC CAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGCTATTTTCGTTTATCCATAGTTGCCTGACT CCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGC GAGACCCACGCTCACCAGCTCCAGATTTATCAGCAATAAACCCAGCCAGCCGGAAGGGCCGAGCGCAGA AGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAG TTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTACGCTCGTCGT TTGGTATGGCTTCATTACGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGC AAAAAAGCGGTTAGCTCCTTCGGTCTCCGATCGTTGTGAGAAAGTAAAGTTGGCCGAGTGTATCACT CATGGTTATGGCAGCACTGCATAATCTCTTACTGTCTATGCCATCCGTAAGATGCTTTTCTGTGACTG GTGAGTACTCAACCAAGTCATTCTGAGAATAGTGATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCA ATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGG GCGAAAACCTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCAACCAACT GATCTTCAGCATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCA AAAAAGGGAATAAGGGCGACACGGAATGTTGAATACTCATACTCTTCTTTTCAATATTATTGAAG CATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAG GGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCT (서열번호 107)

도면5a

CMV-T7-Cas9 D10A/H840A-3XFLAG-VP64:
ATATGCCAAGTACGCCCTATTGACGTCAATGACGGTAAATGGCCCGCTGGCATTATGCCAGTACATGAC CTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGC AGTACATCAATGGGCGTGGATAGCGGTTTGA CTACGGGGATTTC AAGTCTCCACCCATTGACGTCAATGG GAGTTTGTGGTGGCACC AAAATCAACGGGACTTCCAAAATGTCGTAACAACTCCGCCCATGACGCAAAATG GCGGGTAGGCGGTGTACGGTGGGAGGTCTATATAAGCAGAGCTGGTTTAGTGAACCGTCAGATCCGCTAGAG ATCCGCGGCCGCTAATACGACTCACTATAGGGAGAGCCGCCACCATGGATAAGAAATACTCAATAGGCTTAGc TATCGGCACAAATAGCGTCGGATGGGCGGTGATCACTGATGAATATAAGGTTCCGTCTAAAAAGTTCAAGGTT CTGGGAAATACAGACCGCCACAGTATCAAAAAAATCTTATAGGGGCTCTTTATTTGACAGTGGAGAGACAG CGGAAGCGACTCGTCTAAACGGACAGCTCGTAGAAGGTATACACGTCGGAAGAATCGTATTGTTATCTACA GGAGATTTTTCAAATGAGATGGCGAAAGTAGATGATGTTCTTTTCATCGACTTGAAGAGTCTTTTTGGTGG AAGAAGACAAGAAGCATGAACGTCATCTATTTTTGGAAATATAGTAGATGAAGTTGCTTATCATGAGAAAAA TCCAACATCTATCATCTGCGAAAAAATGGTAGATTCTACTGATAAAGCGGATTTGCGCTTAATCTATTTGG CCTTAGCGCATATGATTAAGTTTCGTGGTCATTTTTGATTGAGGGAGATTAAATCCTGATAATAGTGATGTG GACAACTATTTATCCAGTTGGTACAAACCTACAATCAATTATTTGAAGAAAACCTATTAACGCAAGTGGAGT AGATGCTAAAGCGATTCTTTCTGCACGATTGAGTAAATCAAGACGATTAGAAAATCTCATTGCTCAGCTCCCG GTGAGAAGAAAAATGGCTATTTGGGAATCTCATTGCTTTGTCATTGGGTTGACCCCTAATTTAAATCAAAT TTTGATTTGGCAGAAGATGCTAAATTACAGCTTTCAAAAGATACCTACGATGATGATTAGATAATTTATTGGC GCAAAATGGAGATCAATATGCTGATTTGTTTTGGCAGCTAAGAAATTTATCAGATGCTATTTACTTTTCAGATAT CCTAAGAGTAAATCTGAAATAACTAAGGCTCCCTATCAGCTTCAATGATTAAACGCTACGATGAACATCATC AAGACTTGACTCTTTAAAAAGCTTTAGTTTCGACAACAACCTCCAGAAAAGTATAAAGAAATCTTTTTGATCAAT CAAAAAACGGATATGCAGGTTATATTGATGGGGAGCTAGCCAAGAAGATTTTATAAATTTATCAAACCAAT TTTAGAAAAATGGATGGTACTGAGGAATTATTGGTGAAGTAACTCGTGAAGATTGCTGCGCAAGCAACG GACCTTTGACAACGGCTCTATTTCCCATCAAATTCATTGGGTGAGCTGCATGCTATTTTGAGAAGACAAGAA GACTTTTATCCATTTTAAAAAGACAATCGTGAGAAGATTGAAAAAATCTTGACTTTTCGAATTCCTTATTATGTT GGTCCATTGGCGCGTGGCAATAGTCGTTTTGTCATGGATGACTCGGAAGTCTGAAGAAACAATTACCCCATGGGA ATTTTGAAGAAGTTGTCGATAAAGGTGCTTCACTCAATCATTTATTGAACGCATGACAACTTTGATAAAAAAT CTTCCAAATGAAAAAGTACTACCAAAACATAGTTTGCTTTATGAGTATTTTACGGTTTATAACGAATTGACAAA GGTCAAATATGTTACTGAAGGAATGCGAAAACAGCATTCTTTTCAGGTGAACAGAAGAAAGCCATTGTTGAT TTACTCTTCAAAACAATCGAAAAGTAACCGTTAAGCAATTAAAAGAAGATTATTTCAAAAAATAGAATGTTT TGATAGTGTGAAATTTCAAGAGTTGAAGATAGATTTAATGCTTCATTAGGTACCTACCATGATTGCTAAAAA TTATTAAGATAAAGATTTTTGGATAATGAAGAAAATGAAGATATCTTAGAGGATATTGTTTTAACATTGACC TTATTTGAAGATAGGGAGATGATTGAGGAAAGACTTAAACATATGCTCACCTCTTTGATGATAAGGTGATGA AACAGCTTAAACGTCGCCGTTATACTGGTTGGGGACGTTTGTCTCGAAAATTGATTAAATGGTATTAGGGATAA GCAATCTGGCAAAACAATTATGATTTTTGAAATCAGATGGTTTTGCCAATCGCAATTTTATGCAGCTGATCC ATGATGATAGTTTGACATTTAAAGAAGACATTCAAAAAGCACAAAGTGTCTGGACAAGGCGATAGTTTACATGA ACATATTGCAAAATTTAGCTGGTAGCCCTGCTATTAATAAAGGTATTTTACAGACTGTAAAAAGTTGTTGATGAAT TGGTCAAAGTAATGGGGCGGCATAAGCCAGAAAAATATCGTTATTGAAATGGCACGTGAAAAATCAGACAACCTC AAAAGGGCCAGAAAAATTCGCGAGAGCGTATGAAACGAATCGAAGAAGGTATCAAAGAATTAGGAAGTCAG ATTCTTAAAGAGCATCTGTTGAAAATACTCAATTGCAAAATGAAAAGCTCTATCTATTATCTCCAAAATGG AAGAGACATGTATGTGGACCAAGAATTAGATATTAATCGTTTAAAGTGATTATGATGTCGATgcATTGTTCCAC AAAGTTTCTTAAAGACGATTCAATAGACAATAAGGTCTTAACGCGTTCTGATAAAAAATCGTGGTAAATCGGA TAACGTTCCAAGTGAAGAAAGTAGTCAAAAAGATGAAAACTATTGGAGACAACCTCTAAACGCCAAGTTAATC ACTCAACGTAAGTTTGATAATTTAACGAAAGCTGAACGTGGAGGTTTGAAGTGAACCTGATAAAGCTGGTTTTA TCAAACGCCAATTGGTTGAAACTGCCAAATCACTAAGCATGTGGCACAATTTTGGATAGTCGCATGAATAC TAAATACGATGAAAATGATAAATCTATTCGAGAGGTTAAAGTGATTACCTTAAATCTAAATTAGTTTCTGACT TCCGAAAAGATTTCGAATCTATAAAGTACGTGAGATTAACAATTACCATCATGCCCATGATGCGTATCTAAAT GCCGTCGTTGGAACTGCTTTGATTAAGAAATATCCAAAACCTGAATCGGAGTTTGTCTATGGTGATTATAAAGT

도면5b

CMV-T7-Cas9 D10A/H840A-3XFLAG-VP64:
<p>TTATGATGTTCTGTAAGTCTGAGCAAGAAATAGGCAAAGCAACCGCAAAATATTTCTTTACT CTAATATCATGAACCTCTCAAAACAGAAATTACACTTGCAAATGGAGAGATTGCGAAACGCCCTCTAATCGAA ACTAATGGGGAACCTGGAGAAATTGTCTGGGATAAAGGGCGAGATTTGCCACAGTGCAGCAAGTATTGTCC ATGCCCAAGTCAATATTGTCAAGAAAACAGAAAGTACAGACAGGCGGATTCTCAAGGAGTCAATTTTACCAA AAAGAAATTCGGACAGCTTATTGCTCGTAAAAAAGACTGGGATCAAAAAAATATGGTGGTTTTGATAGTCC AACGGTAGCTTATTAGTCTAGTGGTGTCTAAGGTGAAAAAGGAAATCGAAGAAGTTAAATCCGTAA AGAGTTACTAGGGATCACAATTATGGAAGAAAGTTCTTTGAAAAAATCCGATTGACTTTTGAAGCTAAA GGATATAAGGAAGTAAAAAAGACTTAATCATTAACTACCTAAATATAGTCTTTTGAAGTTAGAAAACGGTC GTAACGGGATGCTGCTAGTGCAGGAGAAATTACAAAAAGGAAATGAGCTGGCTCTGCCAAGCAAAATATGTGA ATTTTTATATTTAGCTAGTCATTATGAAAAGTTGAAGGGTAGTCCAGAAGATAACGAACAAAAACAATTGTTT GTGGAGCAGCATAAGCATTATTTAGATGAGATTATTGAGCAATCAGTGAATTTCTAAGCGTGTATTTTATGC AGATGCCAATTTAGATAAAGTTCTTAGTGATATAACAACATAGAGACAAACCAATACGTGAACAAGCAGAA AATATTATTCATTTATTACGTTGACGAATCTGGAGCTCCGCTGCTTTAAATATTTGATACAACAATTGAT CGTAAACGATATACGCTACAAAAGAAAGTTTAGATGCCACTCTATCCATCAATCCATCACTGGTCTTTATGA AACACGCATTGATTGAGTCAGTAGGAGGTGACGGTCTCCCAAGAAGAAGAGGAAAGTCTCGAGCGACTA CAAAGACCATGACGGTGATTATAAAGATCATGACATCGATTACAAGGATGACGATGACAAGGctgaggaggcgg tggaagcGGGCGCGCCGACGCGCTGGACGATTTCGATCTCGACATGCTGGGTTCTGATGCCCTCGATGACTTTG ACCTGGATATGTTGGGAAGCGACGCTTGGATGACTTTGATCTGGACATGCTCGGCTCCGATGCTCTGGACGA TTTCGATCTCGATATGTTATAccggTCATCATCACCATCACCATTGAGTTTAAACCCGCTGATCAGCCTCGACTG TGCCCTCTAGTTGCCAGCCTCTGTTGTTGCCCTCCCCCGTGCCTTCTTACCTGGAAGGTGCCACTCCCA CTGCTCTTCTTAATAAAGAGGAAATGTCATGCTGCTGAGTAGTGTCTATTCTATTTCTGGGGGGTGG GGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCT ATGGCTCTGAGGGCGAAAGAACAGCTGGGGCTCGATACCGTCGACCTCTAGCTAGAGCTTGGCGTAATCA TGGTCATAGCTGTTTCTGTGTAATTTGTTATCCGCTCACAATTCACACAAACATACGAGCCGGAAGCATAAA GTGTAAAGCCTAGGGTGCTTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGTTTCCAGT AGCTCCCTCGTGCCTCTCTGTTCCGACCTGCCGTTACCGGATACCTGTCGCTTCTTCCCTTCGGGAAGC GTGGCGCTTCTCAATGCTCAGCTGTAGGTATCTCAGTTCCGGTGTAGGTCTGCTCCAAGCTGGGCTGTGT GCACGAACCCCCGTTACGCCCCACCGTGTGCGCTTATCCGGTAACATCTGCTTGTAGTCCAACCCGGTAAGA CACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACA GAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGC CAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACCCGCTGGTAGCGGTGGTTTTTT TGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTGATCTTTTCTACGGGGTCT GACGCTCAGTGGAAACGAAAACCTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTACCTAGA TCCTTTTAAATTAATAAAGTATTAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAAT GCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCTGTCATCCATAGTTGCTGACTCCCCGTCTGT AGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGTGCAATGATACCGCAGAGCCACGCTCACC GGCTCCAGATTTATCAGCAATAAACAGCCAGCCGGAAGGGCCGAGCGCAGAAAGTGGTCTGCAACTTTATC CGCTCCATCCAGTCTATTAATTTGTTGCCGGAAGCTAGAGTAAGTAGTTCCGCAATTAAGTTTGCACAAC GTTGTTGCCATTGCTACAGGCATCGTGGTGTACGCTCTGCTGTTGGTATGGCTTCACTCAGCTCCGGTCCCA ACGATCAAGGCGAGTTACATGATCCCCATGTTGTGCAAAAAAGCGTTAGCTCTTCCGCTCCTCGATCGTT CTCAGAAAGTAAGTTGGCCGAGCAAAATTAAGCTACAACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCT GCTTAGGGTTAGGCGTTTTGCGCTGCTTCCGATGTACGGGCCAGATACGCGTTGACATTGATTATTGACT AGTTATTAATAGTAATCAATTACGGGGTCATTAGTTTCATAGCCATATATGAGTCCGCGTTACATAACTTAC GGTAATAGGGCCGCTGGCTGACCGCCCAACGACCCCGCCATTGACGTCAATAATGACGTATGTTCCCAT GTAAAGCAATAGGGACTTTCCATTGACGTCAATGGGTGACTATTTACGGTAACTGCCACTTGGCAGTAC ATCAAGTGATCC (서열번호 108)</p>

도면5c

CMV-T7-Cas9 D10A/H840A-3XFLAG-VP64:
<p>GCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCTTTTCAATATTATTGAA GCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAACAAATAGGGGT TCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTGACGGATCGGGAGATCGATCTCCCGATCCCCAGGG TCGACTCTCAGTACAATCTGCTCTGATGCCGATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGT CGCTAGTAGTGCGCGAGCAAAATTTAAGCTACAACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCT GCTTAGGGTTAGGCGTTTTGCGCTGCTTCCGATGTACGGGCCAGATACGCGTTGACATTGATTATTGACT AGTTATTAATAGTAATCAATTACGGGGTCATTAGTTTCATAGCCATATATGAGTCCGCGTTACATAACTTAC GGTAATAGGGCCGCTGGCTGACCGCCCAACGACCCCGCCATTGACGTCAATAATGACGTATGTTCCCAT GTAAAGCAATAGGGACTTTCCATTGACGTCAATGGGTGACTATTTACGGTAACTGCCACTTGGCAGTAC ATCAAGTGATCC (서열번호 108)</p>

도면6a

MV-T7-Cas9 보고	D10A/H840A-3XFLAG-VP64
ATATGCCAAGTACGCCCTATTGACGTCATGACGGTAAATGGCCGCTGGCATTATGCCAGTACATGAC CTTATGGGACTTTCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGC AGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCATTGACGTCAATGG GAGTTTGTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACCTCCGCCCATTTGACGCAAATG GGCGGTAGGCGGTGACGGTGGGAGGTCTATATAAGCAGAGCTGGTTTAGTGAACCGTCAGATCCGCTAGAG ATCCGCGGCGCTAATACGACTCACTATAGGGAGAGCCGCCACCATGGATAAAAAGTATTCTATTGGTTAGC CATCGGCACTAATTCGGTGGATGGGCTGTCTATAACCGATGAATACAAAGTACCTTCAAAGAAATTTAAGGTG TTGGGGAACACAGACCGTCATTCGATTAAGAAGTCTTATCGGTGCCCTCTATTCGATAGTGGCGAAACGG CAGAGGCGACTCGCTGAAACGAACCGCTCGGAGAAGGTATACACGTCGCAAGAACCGAATATGTTACTTAC AAGAAATTTTAGCAATGAGATGGCCAAAGTTGACGATTCTTTCTTACCCTTTGGAAGAGTCCTTCCTTGTG GAAGAGGACAAAGAACATGAACGGCACCCCATCTTTGAAACATAGTAGATGAGGTGGCATATCATGAAAAG TACCAACGATTATCACCTCAGAAAAAGCTAGTTGACTCAACTGATAAAGCGGACCTGAGGTTAATCTACTT GGCTCTTGGCCATATGATAAAGTTCCGTGGGCACTTTCTCATTGAGGGTGATCTAAATCCGGACAACCTCGGAT GTCGACAACTGTTTATCCAGTTAGTACAAACCTATAATCAGTTGTTGAAGAGAACCCTATAAATGCAAGTG GCGTGGATGCGAAGGCTATTCTAGCGCCGCTCTTAAATCCGACGGCTAGAAAACTGATCGCACAATT ACCCGGAGAGAAGAAAAATGGGTTGTTGCGTAACCTTATAGCGCTCTCACTAGGCTGACACCAAAATTTAAG TCGAACTTCGACTTAGCTGAAGATGCCAAATGCGAGCTTAGTAAGGACACGTACGATGACGATCTCGACAATC TACTGGCACAATTTGGAGATCAGTATGCGGACTTATTTTGGCTGCCAAAACCTTAGCGATGCAATCTCCTA TCTGACATACTGAGAGTTAATACTGAGATTACCAAGGCGCCGTTATCCGCTTCAATGATCAAAAGGTACGATG AACATCACCAAGACTTGACACTTCTCAAGGCCCTAGTCCGTCAGCAACTGCCTGAGAAATATAAGGAAATATT CTTTGACGATCGAAAAACGGGTACGCAAGTTATATTGACGGCGGAGCGAGTCAAGAGGAATTTCTACAAGTT TATCAAAACCATATTAGAGAAGATGGATGGGACGGAAGAGTTGCTTGTAAGTCAATCGCGAAGATCTACT GCGAAAGCAGCGGACTTTGCAACACGGTAGCATTCCACATCAAATCCACTTAGGCGAATTGCATGCTATACTT AGAAGGCAGGAGGATTTTATCCGTTCTCAAGACAATCGTGAAAAGATTGAGAAAACTTAACCTTTTCGCA TACCTTACTATGTGGGACCCCTGGCCGAGGGAAGTCTCGGTTGCGATGGATGACAAGAAAGTCCGAAGAAA CGATTATCCATGGAATTTTGAAGGAAGTTGCGATAAAGGTGCGTCAGCTCAATCGTTTATCAGAGGATGAC CACTTTGACAAGAATTTACCGAACGAAAAAGTATTGCCTAAGCACAGTTTACTTTACGAGTATTTACAGTGT ACAATGAACCTACGAAAAGTTAAGTATGTCACTGAGGGCATGCGTAAACCCGCTTTCTAAGCGGAGAACAGA AGAAAGCAATAGTAGTCTGTTATTCAAGACCAACCGCAAAGTGACAGTTAAGCAATTGAAAGAGGACTACTT TAAGAAAATTGAATGCTTCGATTCTGTCGAGATCTCCGGGTAGAAGATCGATTAAATGCGTCACTTGGTACG TATCATGACCTCTTAAAGATAATTAAGATAAGGACTTCTGGATAACGAAGAGAATGAAGATATCTTAGAAG ATATAGTGTGACTCTTACCCTCTTGAAGATCGGGAATGATTGAGGAAAGACTAAAAACATACGCTCACCT GTTCGACGATAAGGTTATGAACAGTTAAAGAGGCGTCGCTATACGGGCTGGGACGATTGTCGCGGAAACT TATCAACGGGATAAGAGACAAGCAAAGTGGTAAACTATTCTCGATTTCTAAGAGCGACGGCTTCGCCAAT AGGAACCTTATGCAGCTGATCCATGATGACTCTTAACTTCAAAGAGGATATACAAAAGGCACAGGTTTCCG GACAAGGGGACTCATTGACGAACATATTGCGAATCTTGCTGGTTTCGCCAGCCATCAAAAAGGGCATACTCCA GACAGTCAAAGTAGTGATGAGCTAGTTAAGGTCATGGGACGTCACAAACCGGAAAAACATTGTAATCGAGAT GGCACGCGAAAATCAAACGACTCAGAAGGGGCAAAAAACAGTCGAGAGCGGATGAAGAGAATAGAAGAG GGTATTAAAGAACTGGGAGCCAGATCTTAAAGGAGCATCTGTGGAATAACCCAAATGCGAACGAGAAA CTTTACCTCTATTACCTACAAAATGGAAGGGACATGTATGTTGATCAGGAACTGGACATAAACCGTTTATCTGA TTACGACGTCGATgCATTGTACCCCAATCCTTTTGAAGGACGATTCAATCGACAATAAAGTGCTTACACGCTC GGATAAGAACCAGGGAAAAAGTGACAATGTTCCAAGCGAGGAAGTCGTAAGAAAAATGAAGAACTATTGGC GGCAGCTCCTAAATGCGAAACTGATAACGCAAGAAAAGTTGATAACTTAACTAAAGCTGAGAGGGGTGGCT TGTCTGAACCTGACAAGGCCGATTATTAAACGTGAGCTCGTGGAACCCGCCAAATCACAAAGCATGTTGC ACAGATACTAGATTCGGAATGAATACGAAAATACGACGAGAACGATAAGCTGATTCGGGAAGTCAAAGTAAT CACTTTAAAGTCAAATTTGGTGTGCGACTTCAGAAAGGATTTTCAATCTATAAAGTTAGGGAGATAAATACT ACCACCATGCGCAGCAGCTTATCTTAATGCCGTCGTAGGGACCGCACTTAAAGAAATACCGAAGCTAGA AAGTGAGTTTGTGATGGTGATTACAAAGTTTATGACGTCCGTAAGATGATCGCGAAAAGCGAACAGGAGAT	

도면6b

MV-T7-Cas9 보고 D10A/H840A-3XFLAG-VP64
AGGCAAGGCTACAGCCAAATACCTCTTTTATTCTAACATTATGAATTTCTTTAAGACGGAAATCACTCTGGCAA ACGGAGAGATACGCAACACGACCTTTAATTGAAACCAATGGGGAGACAGGTGAAATCGTATGGGATAAGGGC CGGGACTTCGCGACGGTGAGAAAAGTTTTGTCCATGCCCAAGTCAACATAGTAAAGAAAAGTGAAGTGCAG ACCGGAGGGTTTTCAAAGGAATCGATTCTTCCAAAAAGGAATAGTGATAAGCTCATCGCTCGTAAAAAGGACT GGGACCCGAAAAAGTACGGTGGCTTCGATAGCCCTACAGTTGCCTATTCTGTCTAGTAGTGGCAAAAGTTGA GAAGGGAAAAATCCAAGAACTGAAGTCAGTCAAAAGAATTATTGGGGATAACGATTATGGAGCGCTCGTCTTT TGAAAAAGAACCCCATCGACTTCTTGAGGCGAAAGGTTACAAGGAAGTAAAAAGGATCTCATAATTAAACTA CCAAAGTATAGTCTGTTTGAAGTAGAAAATGGCCGAAAACGGATGTTGGCTAGCGCCGGAGAGCTTCAAAAG GGGAACGAACTCGCACTACCGTCTAAATACGTGAATTTCTGTATTTAGCGTCCCATTACGAGAAGTTGAAAG GTTACCTGAAGATAACGAACAGAAAGCACTTTTTGTTGAGCAGCACAAACATTATCTCGACGAAATCATAGA GCAAAATTCGGAATTCAGTAAGAGAGTCATCCTAGCTGATGCCAATCTGGACAAAGTATTAAGCGCATACAAC AAGCACAGGGATAAACCCATACGTGAGCAGGCGGAAAAATATTATCCATTTGTTTACTCTTACCAACCTCGGCG CTCCAGCCGATTCAAAGTATTTTGACACAACGATAGATCGCAACGATACACTTCTACCAAGGAGGTGCTAGA CGCGACACTGATTACCAATCCATCACGGGATTATATGAAACTCGGATAGATTGTACAGCTTGGGGGTGAC GGATCCCCAAGAAGAAGAGGAAAGTCTCGAGCGACTACAAAGACCATGACGGTGATTATAAAGATCATGAC ATCGATTACAAGGATGACGATGACAAAGGctcaggaggcggtggaagcGGGCGCGCCGACGCGCTGGACGATTTCG ATCTCGACATGCTGGGTTCTGATGCCCTCGATGACTTTGACCTGGATATGTTGGGAAGCGACGCTTGGATGA CTTTGATCTGGACATGCTCGGCTCCGATGCTCTGGACGATTTCTGATCTCGATATGTTATAAccggtCATCATACC ATCACCATTTAGTTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTGCCCT CCCCGTCGCTTCTTACCTTGGAGGTTGCCACTCCACTGTCTTCTTAATAAAATGAGGAAATTGCATCG CATTTGCTGAGTAGGTTGCTATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGA AGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACAGCTGGGGCT CGATACCGTCGACTCTAGCTAGAGCTTGGCGTAATCATGGTCATAGCTTTCTGTGTGAAATTGTTATCCG CTCACAATTCACACAACATACGAGCCGGAAGCATAAAGTGTAAGCCTAGGGTGCTTAATGAGTGAGCTAA CTCACAATTAATTGCGTTGCGCTCACTGCCGCTTTCCAGTCGGGAAACCTGCTGCGCAGCTGCATTAATGAAT CGGCCAACGCGCGGGGAGAGCGGTTTGCCTATTGGGCGCTCTTCCGCTTCTCGCTCACTGACTCGCTGCGC TCGGTCTGTTCCGCTGCGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGG GATAACGCAAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCTTGCT GGCGTTTTTCCATAGGCTCCGCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGAAAC CCGACAGGACTATAAAGATACGAGCGTTTCCCCCTGGAAGCTCCCTCGTGCCTCTCCTGTTCCGACCTGCC GCTTACCGGATCTGTCCGCTTCTCCCTCGGGAGCGTGGCGCTTCTCAATGCTCAGCTGTAGGTATC TCAGTTGGGTAGGTGCTGCTCAAGCTGGGCTGTGTGCACGAACCCCCGTTACGCCGACCGCTGCGC CTTATCCGGTAATATCGTCTTGAAGTCCAACCCGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGT AACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTAACTACGGCTAC ACTAGAAGGACAGTATTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTT GATCCGGCAAAACAACACCGCTGTGTAGCGGTGTTTTTTTTGTTGCAAGCAGCAGATTACGCGCAGAAAAA AAGGATCTCAAGAAGATCCTTTGATCTTTCTACGGGGTCTGACGCTCAGTGGAAACGAAACTCACGTTAAGG GATTTTGGTCATGAGATTATCAAAAAGGATCTTACCTAGATCCTTTTAAATTAATAATGAAGTTTAAATCAA TCTAAAGTATATAGAGTAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATC TGCTATTTCCTCATCCATAGTTGCTGACTCCCCGTGCTGATAGATACTACGATACGGGAGGGCTTACCATC TGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTACCGGCTCCAGATTTATCAGCAATAAACAGCCA GCCGGAAGGGCCGAGCGCAGAAGTGGTCTGCACTTTATCCGCTCCATCCAGTCTATTAATTGTTGCCGGG AAGCTAGAGTAAGTAGTTCCGCAAGTAAAGTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTC ACGCTCGTCTGTTGGTATGGCTTCATTGAGCTCCGGTTCCTCAACGATCAAGGCGAGTTACATGATCCCCATGT TGTCGCAAAAAGCGGTTAGCTCCTCGGTCTCCGATCGTTGTGAGAAGTAAGTTGGCGCAGTGTTATCACT CATGGTTATGGCAGCACTGCATAATTCTTACTGTGATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGT ACTCAACCAAGTCATTCTGAGAATAGTGATGCGGCGACCGAGTTGCTCTTGGCCGGCTCAATACGGGATAA TACCGCGCCACATAGCAGAATTAAAAAGTGCTCATATTGGAAAACGTTCTCGGGGCGAAAACTCTCAAGG ATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGACCCCACTGATCTTCAGCATCTTTACTTTT ACCAGCGTTCTGGGTGAGCAAAACGGAAGGCAAAATGCCGCAAAAAGGGAATAAGGGCGACACGGAA

도면6c

MV-T7-Cas9 보고 D10A/H840A-3XFLAG-VP64
ATGTTGAATACTCATACTCTTCTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATA CATATTTGAATGATTTAGAAAAATAACAAATAGGGGTTCCGCGCACATTTCCCGAAAAAGTGCCACCTGAC GTCGACGGATCGGGAGATCGATCTCCCGATCCCTAGGGTCTGACTCTCAGTACAACTGCTCTGATGCCGAT AGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTTGAGGTCGCTGAGTAGTGCGCGAGCAAAATTAAGCTAC AACAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGAT GTACGGGCCAGATATACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGT TCATAGCCCATATATGGAGTTCCGCGTTACATAACTACGGTAAATGGCCCGCTGCTGACCGCCCAACGAC CCCCGCCCATTTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATG GGTGGACTATTTACGGTAAACTGCCACTTGGCAGTACATCAAGTGTATC (서열번호 109)

도면7

Cas9- 활성제 단백질
MDKKYSIGLAIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEAT RLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNIVD EVAYHEKYPTIYHLRKKLV DSTDKADRLRIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFI QLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLLENLIAQLPGEKKNGLFGNLIALSLGL TPNFKSNFDLAEDAKLQLSKDITYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNT EITKAPLSASMIKRYDEHHQDLTLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEF YKFIKPILEKMDGTEELLVKNLREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLK DNREKIEKILTRIPYVVGPLARGNSRFAMTRKSEETITPWNFEEVVDKGASAQSFIERMT NFDKNLPNEKVLPHKSLLEYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRK VTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLKIIKDKDFLDNEENEDILEDIV LTTLTFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDF LKSDGFANRNFQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVV DELVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEIEGKELGSQILKEHPVENTQL QNEKLYLYYLQNGRDMYVDQELDINRLSDYDVAIVPQSFLKDDSIDNKVLRSDKNRGKSD NVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTAKERGGSELKAGFIKRQLVETRQITKH VAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVRINNYHHAHDAYLNAV VGTALIKKYPKLESEFVYGDYKVDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLAN GEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNIKKTEVQTGGFSKESILPKRNS DKLIARKKDWDPKKYGGFDSPTVAYSVLVAKVEKGKSKKLKSVKELLGITIMERSSEKPNP IDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNEALALPSKYVNFYLAS HYEKLKGSPEDEQKQLFVEQHKHYLDEIEQISEFSKRVLADANLDKVL SAYNKHRDKPI REQAENI IHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDL SQ LGGDGS pkkkkrkvs DYKDHGDYKDHIDYKDDDDKAAGGGGS GRADALDDFDLMDLGSDA LDDFDLMDLGS DALDDFDLMDLGS DALDDFDLMDL (서열번호 110)

도면8a

dCas9-NLS-3XFLAG-HP1알파
MDKKYSIGLAIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEAT RLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNIVD EVAYHEKYPTIYHLRKKLV DSTDKADRLRIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFI QLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLLENLIAQLPGEKKNGLFGNLIALSLGL TPNFKSNFDLAEDAKLQLSKDITYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNT EITKAPLSASMIKRYDEHHQDLTLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEF YKFIKPILEKMDGTEELLVKNLREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLK DNREKIEKILTRIPYVVGPLARGNSRFAMTRKSEETITPWNFEEVVDKGASAQSFIERMT NFDKNLPNEKVLPHKSLLEYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRK VTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLKIIKDKDFLDNEENEDILEDIV LTTLTFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDF LKSDGFANRNFQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVV DELVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEIEGKELGSQILKEHPVENTQL QNEKLYLYYLQNGRDMYVDQELDINRLSDYDVAIVPQSFLKDDSIDNKVLRSDKNRGKSD NVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTAKERGGSELKAGFIKRQLVETRQITKH VAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVRINNYHHAHDAYLNAV VGTALIKKYPKLESEFVYGDYKVDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLAN GEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNIKKTEVQTGGFSKESILPKRNS DKLIARKKDWDPKKYGGFDSPTVAYSVLVAKVEKGKSKKLKSVKELLGITIMERSSEKPNP IDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNEALALPSKYVNFYLAS HYEKLKGSPEDEQKQLFVEQHKHYLDEIEQISEFSKRVLADANLDKVL SAYNKHRDKPI REQAENI IHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDL SQ LGGDGS PKKKRKV SSDYKDHGDYKDHIDYKDDDDKAAGGGGS MKEGENNKPREKSES KS NFNSADDIKSKKKREQSNDIARGFERGLEPEKII GATDSCGDL MFLMKWKDTDEADLVL AKEANVKCPQIVIAFYEERLTWHAYPEDAENKEKETAKS (서열번호 111)

도면 8b

dCas9-NLS-3XFLAG-HP1 베타
MDKKYSISGLAIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEAT RLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNIVD EVAYHEKYPTIYHLRKKLVDSTDKADLRILIYALAHMIKFRGHFLIEGDLNPDNSDVKLFI QLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLLENLIAQLPGEKKNGLFGNLIALSLGL TPNFKSNFDLAEDAKLQLSKDITYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNT EITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEF YKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLK DNREKIEKILTFRIPIYYVGPLARGNSRFAMTRKSEETITPWNFEVVVDKGASAQSFIERMT NFDKNLPNEKVLPHKSLLEYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRK VTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLKIIKDKDFLDNEENEDILEDIV LTTLTFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDF LKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDLSLHEHIANLAGSPAICKGILQTVKVV DELVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGKELGSQILKEHPVENTQL QNEKLYLYYLQNGRDMYVDQELDINRLSDYDVAIVPQSFLKDDSIDNKVLTRSDKNRGKSD NVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGSELKAGFIKRQLVETRQITKH VAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVRINNYHHAHDAYLNAV VGTALIKKYPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLAN GEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNIKKTEVQTGGFSKESILPKRNS DKLIARKKDWDPKKYGGFDSPTVAISVLVAKVEKGKSKKLKSVKELLGITIMERSSSFENP IDFLEAKGYEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLAS HYEKLKGSPEDEQKQLFVEQHKHYLDEIEQISEFSKRVLADANLDKVL SAYNKHDKPI REQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDTLHQSITGLYETRIDLSDQ LGGDGS[PKKKRKV]SSDYKDHGDYKDHDIDYKDDDDKAAGGGGSTAHETDKSEGKKRKADSD <u>SEDKGEESKPKKKKEESEKPRGFARGLEPERIIGATDSSGELMFLMKWKNSEADLVPKEA</u> <u>NVKCPQVVISFYEERLTWHSYPSEDDDKDDKN. (서열번호 112)</u>

박스 = 핵 국소화 신호

밑줄 = 삼중 플래그 태그

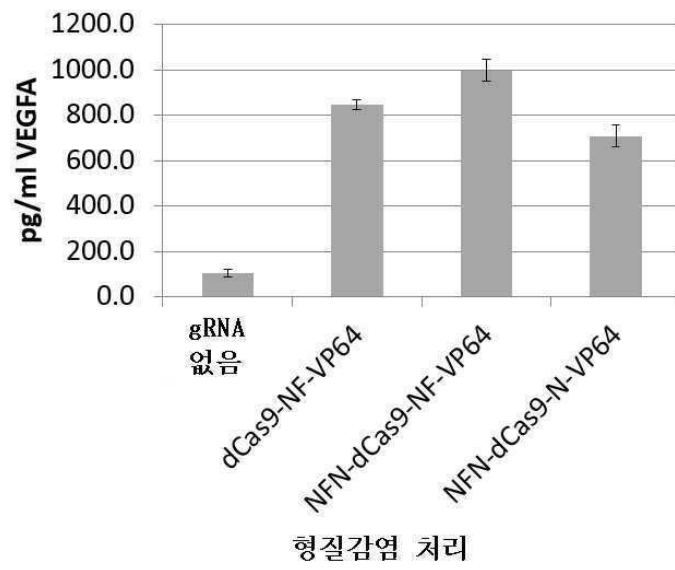
이중 밑줄 = HP1알파 힌지 및 크로모솜도 도메인

도면9

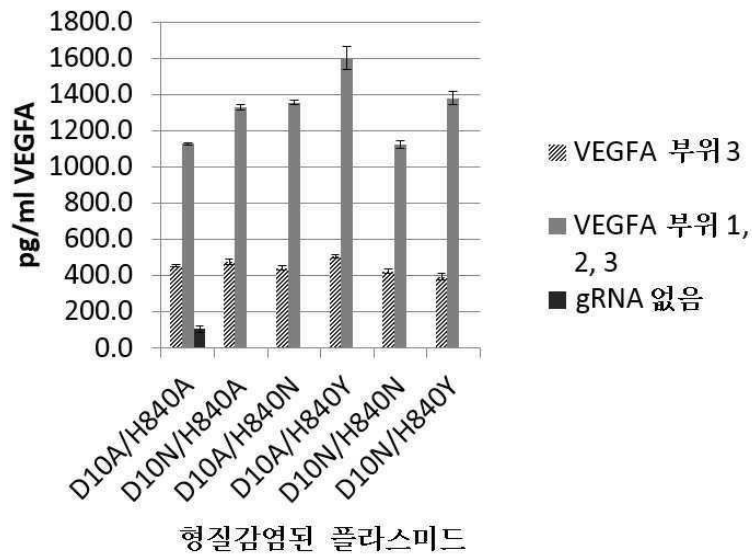
dCas9-3XFLAG-TET1CD
MDKKYSIGLAIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEAT RLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFFHRLSEESFLVEEDKKHERHPIFGNIVD EVAYHEKYPTIYHLRKKLV DSTDKADLR LIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFI QLVQTYNQLFEEENPINASGVDAKAILSARLSKSRRLLENLIAQLPGEKKNGLFGNLIALSLGL TPNFKSNFDLAEDAKLQLSKD TYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNT EITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEF YKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLK DNREKIEKILTFRIPIYYVGPLARGNSRFAMTRKSEETITPWNFEVVDKGASAQSFIERMT NFDKNLPNEKVLPHKSLLEYEFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRK VTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI IKDKDFLDNEENEDILEDIV LTTLTFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDQSGKTILDF LKSDGDFANRNFQQLIHDDSLTFKEDIQKAQVSGQGDLSLHEHIANLAGSPAIKKGILQTVKVV DELVKVMGRHKPENIVIEARENQTTQKGQKNSRERMKRIEIEGKELGSQILKEHPVENTQL QNEKLYLYLLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSD NVPSEEVVKMKNYWRQLLNAKLITQRKFDNLTKAERGGSELDKAGFIKRQLVETRQITKH VAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVVREINNYHHAHDAYLNAV VGTALIKKYPKLESEFYGDYKVDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLAN GEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNS DKLIARKKDWDPKKYGGFDSPTVAYSVLVAKVEKGKSKKLKSVKELLGITIMERSSSFENP IDFLEAKGYKEVKKDLIKLPKYSLEFELENGRKRMLASAGELQKGNELALPSKYVNFLYLAS HYEKLKGSPEDEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKHDKPI REQAENI IHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDL SQ LGGDGS <u>PKKKRKV</u> SSDYKDHDGDYKDHDIDYKDDDDKAAGGGGSLPTCSCLD RVIQKDKGPY <u>YTHLGAGPSVAAREIMENRYGQKGNAIRIEIVVYTGKEGKSSHGCPIAKWVLRSSDEEKV</u> <u>LCLVRQRTGHHCTAVMVVIMVWDGIPLPMADRLYTELTENLKS YNGHPTDRRCTLNENRT</u> <u>CTCQGIDPETCGASFSGCSWSMYFNGCKFGRSPSPRRFRIDPSSPLHEKNLEDNLQSLATR</u> <u>LAPIYQYAPVAYQNQVEYENVARECRLGSKEGRPFSGVTACLDCAHPHRDIHNMNNGSTV</u> <u>VCTLTREDNRSLGVI PQDEQLHVLPYKLSDTDEFGSKEGMEAKIKSGAIEVLAPRRKKRTC</u> <u>FTQPVPRSGKKRAAMTEVLAHKIRAVEKKPIPRIKRKNNSTTTNNSKPSSLPTLGSNTETV</u> <u>QPEVKSETEPHFILKSSDNTKTYSLMPSAPHPVKEASPGFSWSPKTASATPAPLKN DATASC</u> <u>GFERSSTPHCTMPSGRLSGANAAAADGPGISOLGEVAPLPTLSAPVMEPLINSEPTGVTE</u> <u>PLTPHQPNHQPSFLTSPQDLASSPMEDEQHS EADEPPSDEPLSDDPLSPAEEKLPHIDEYW</u> <u>SDSEHIFLDANIGGVAIAPAHGSVLIECARRELHATTPVEHPNRNHPTRL SLVFYQHKNLNK</u> <u>POHGFE LNKIKFEAKEAKNKKMKASEQKDQAA NEGPESSEVNELNQIPSHKALT LTHDNVV</u> <u>TVSPYALTHVAGPYNHVV</u>
(서열번호 113)

박스 = 핵 국소화 신호
밀줄 = 삼중 플래그 태그
이중 밀줄 = TET1CD

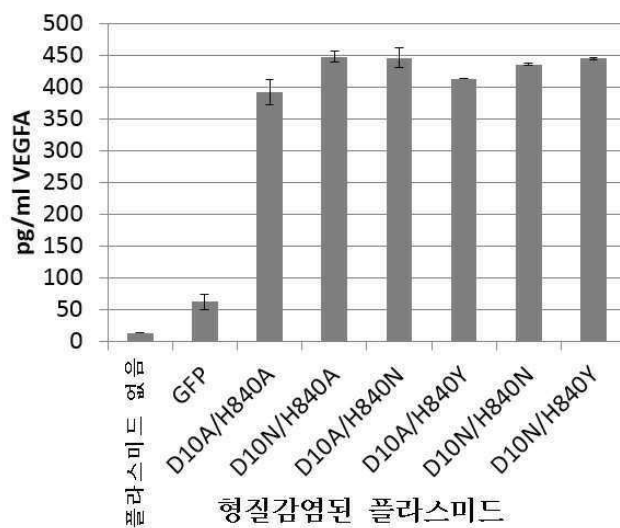
도면10



도면11



A



B

서열 목록

SEQUENCE LISTING

<110> THE GENERAL HOSPITAL CORPORATION

<120> RNA-GUIDED TARGETING OF GENETIC AND EPIGENOMIC REGULATORY
PROTEINS TO SPECIFIC GENOMIC LOCI

<130> 00786-0882W01

<140> PCT/US2014/027335

<141> 2014-03-14

<150> 61/921,007

<151> 2013-12-26
 <150> 61/838,178
 <151> 2013-06-21
 <150> 61/838,148
 <151> 2013-06-21
 <150> 61/799,647
 <151> 2013-03-15
 <160> 113
 <170> PatentIn version 3.5
 <210> 1
 <211> 262
 <212> RNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic

 guide polynucleotide
 <220><221> modified_base
 <222> (1)..(20)
 <223> a, c, u, g, unknown or other and this region may encompass
 17-20 nucleotides, wherein some positions may be absent
 <220><221> modified_base
 <222> (63)..(262)
 <223> a, c, u, g, unknown or other and this region may encompass
 0-200 nucleotides, wherein some positions may be absent
 <400> 1
 nnnnnnnnnn nnnnnnnnnn guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60
 cgnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 120

 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 180
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 240
 nnnnnnnnnn nnnnnnnnnn nn 262
 <210> 2
 <211> 275
 <212> RNA
 <213> Artificial Sequence

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

guide polynucleotide

<220><221> modified_base

<222> (1)..(20)

<223> a, c, u, g, unknown or other and this region may encompass
17-20 nucleotides, wherein some positions may be absent

<220><221> modified_base

<222> (76)..(275)

<223> a, c, u, g, unknown or other and this region may encompass
0-200 nucleotides, wherein some positions may be absent

<400> 2

nnnnnnnnnn nnnnnnnnnn guuuuagagc uaugcugaaa agcauagcaa guuaaaauaa	60
ggcuaguccg uuauccccnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	120
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	180
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	240
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnn	275

<210> 3

<211> 287

<212> RNA

<213> Artificial Sequence

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

guide polynucleotide

<220><221> modified_base

<222> (1)..(20)

<223> a, c, u, g, unknown or other and this region may encompass
17-20 nucleotides, wherein some positions may be absent

<220><221> modified_base

<222> (88)..(287)

<223> a, c, u, g, unknown or other and this region may encompass
0-200 nucleotides, wherein some positions may be absent

<400> 3

nnnnnnnnnn nnnnnnnnnn guuuuagagc uaugcuguuu uggaacacaaa acagcauagc	60
--	----

aaguuaaaau aaggcuaguc cguuauacnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 120
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 180
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 240
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnn 287

<210> 4

<211> 296

<212> RNA

<213> Artificial Sequence

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

guide polynucleotide

<220><221> modified_base

<222> (1)..(20)

<223> a, c, u, g, unknown or other and this region may encompass

17-20 nucleotides, wherein some positions may be absent

<220><221> modified_base

<222> (97)..(296)

<223> a, c, u, g, unknown or other and this region may encompass

0-200 nucleotides, wherein some positions may be absent

<400> 4

nnnnnnnnnn nnnnnnnnnn guuuuagagc uagaaaauagc aaguuaaaau aaggcuaguc 60
 cguuaucaac uugaaaaagu ggcaccgagu cggugcnnnn nnnnnnnnnn nnnnnnnnnn 120

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 180

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 240

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnn 296

<210> 5

<211> 96

<212> RNA

<213> Artificial Sequence

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

guide oligonucleotide

<220><221> modified_base

<222> (1)..(20)

<223> a, c, u, g, unknown or other and this region may encompass

17-20 nucleotides, wherein some positions may be absent

<400> 5

nnnnnnnnnn nnnnnnnnnn guuuuagagc uagaaaauagc aaguuuuuuu aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugc 96

<210> 6

<211> 106

<212> RNA

<213> Artificial Sequence

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

guide polynucleotide

<220><221> modified_base

<222> (1)..(20)

<223> a, c, u, g, unknown or other and this region may encompass

17-20 nucleotides, wherein some positions may be absent

<400> 6

nnnnnnnnnn nnnnnnnnnn guuuuagagc uagcuggaa acagcauagc aaguuuuuuu 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 7

<211> 106

<212> RNA

<213> Artificial Sequence

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

guide polynucleotide

<220><221> modified_base

<222> (1)..(20)

<223> a, c, u, g, unknown or other and this region may encompass

17-20 nucleotides, wherein some positions may be absent

<400> 7

nnnnnnnnnn nnnnnnnnnn guuuuagagc uagcuggaa acagcauagc aaguuuuuuu 60

aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106

<210> 8

<211> 79

<212> RNA

<213> Artificial Sequence

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

guide oligonucleotide

<400> 8

ggaaccuuc aaaacagcau agcaaguuaa aaauaggcua guccguuau aacuugaaaa 60

aguggcaccg agucggugc 79

<210> 9

<211> 62

<212> RNA

<213> Artificial Sequence

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

guide oligonucleotide

<400> 9

ggagcgagcg gagcgguaca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cg 62

<210> 10

<211> 100

<212> RNA

<213> Artificial Sequence

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

guide polynucleotide

<400> 10

ggagcgagcg gagcgguaca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 11

<211> 8

<212> PRT

<213> Artificial Sequence

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

nuclear localization signal peptide

<400> 11

Pro Lys Lys Lys Arg Lys Val Ser

1 5

<210> 12

<211> 50

<212> PRT

<213> Artificial Sequence

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

VP64 domain polypeptide

<400> 12

Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu

1 5 10 15

Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe

 20 25 30

Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp

 35 40 45

Met Leu

 50

<210> 13

<211> 1368

<212> PRT

<213> Streptococcus pyogenes

<400> 13

Met Asp Lys Lys Tyr Ser Ile Gly Leu Ala Ile Gly Thr Asn Ser Val

1 5 10 15

Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe

 20 25 30

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

 35 40 45

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

 50 55 60

Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys

65 70 75 80

Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser

 85 90 95

Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys

100 105 110

His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr

115 120 125

His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp

130 135 140

Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His

145 150 155 160

Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro

165 170 175

Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr

180 185 190

Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala

195 200 205

Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn

210 215 220

Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn

225 230 235 240

Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe

245 250 255

Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp

260 265 270

Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp

275 280 285

Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp

290 295 300

Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser

305 310 315 320

Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys

325 330 335

Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe

340 345 350
Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser

355 360 365
Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp

370 375 380
Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg

385 390 395 400
Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu

405 410 415
Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe

420 425 430
Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile

435 440 445
Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp

450 455 460
Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu

465 470 475 480
Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr

485 490 495
Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser

500 505 510
Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys

515 520 525
Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln

530 535 540
Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr

545 550 555 560
Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp

565 570 575
Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly

580 585 590

Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp
 595 600 605
 Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr
 610 615 620
 Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala
 625 630 635 640
 His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr
 645 650 655
 Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp
 660 665 670
 Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe
 675 680 685
 Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe
 690 695 700
 Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu
 705 710 715 720
 His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly
 725 730 735
 Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly
 740 745 750
 Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln
 755 760 765
 Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile
 770 775 780
 Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro
 785 790 795 800
 Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu
 805 810 815
 Gln Asn Gly Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg
 820 825 830
 Leu Ser Asp Tyr Asp Val Asp Ala Ile Val Pro Gln Ser Phe Leu Lys

835	840	845	
Asp Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg			
850	855	860	
Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys			
865	870	875	880
Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys			
885	890	895	
Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp			
900	905	910	
Lys Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr			
915	920	925	
Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp			
930	935	940	
Glu Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser			
945	950	955	960
Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg			
965	970	975	
Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val			
980	985	990	
Val Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe			
995	1000	1005	
Val Tyr Gly Asp Tyr Lys Val Tyr Asp Val Arg Lys Met Ile Ala			
1010	1015	1020	
Lys Ser Glu Gln Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe			
1025	1030	1035	
Tyr Ser Asn Ile Met Asn Phe Phe Lys Thr Glu Ile Thr Leu Ala			
1040	1045	1050	
Asn Gly Glu Ile Arg Lys Arg Pro Leu Ile Glu Thr Asn Gly Glu			
1055	1060	1065	
Thr Gly Glu Ile Val Trp Asp Lys Gly Arg Asp Phe Ala Thr Val			
1070	1075	1080	

Arg Lys	Val Leu Ser Met Pro	Gln Val Asn Ile Val	Lys Lys Thr
1085	1090	1095	
Glu Val	Gln Thr Gly Gly Phe	Ser Lys Glu Ser Ile	Leu Pro Lys
1100	1105	1110	
Arg Asn	Ser Asp Lys Leu Ile	Ala Arg Lys Lys Asp	Trp Asp Pro
1115	1120	1125	
Lys Lys	Tyr Gly Gly Phe Asp	Ser Pro Thr Val Ala	Tyr Ser Val
1130	1135	1140	
Leu Val	Val Ala Lys Val Glu	Lys Gly Lys Ser Lys	Lys Leu Lys
1145	1150	1155	
Ser Val	Lys Glu Leu Leu Gly	Ile Thr Ile Met Glu	Arg Ser Ser
1160	1165	1170	
Phe Glu	Lys Asn Pro Ile Asp	Phe Leu Glu Ala Lys	Gly Tyr Lys
1175	1180	1185	
Glu Val	Lys Lys Asp Leu Ile	Ile Lys Leu Pro Lys	Tyr Ser Leu
1190	1195	1200	
Phe Glu	Leu Glu Asn Gly Arg	Lys Arg Met Leu Ala	Ser Ala Gly
1205	1210	1215	
Glu Leu	Gln Lys Gly Asn Glu	Leu Ala Leu Pro Ser	Lys Tyr Val
1220	1225	1230	
Asn Phe	Leu Tyr Leu Ala Ser	His Tyr Glu Lys Leu	Lys Gly Ser
1235	1240	1245	
Pro Glu	Asp Asn Glu Gln Lys	Gln Leu Phe Val Glu	Gln His Lys
1250	1255	1260	
His Tyr	Leu Asp Glu Ile Ile	Glu Gln Ile Ser Glu	Phe Ser Lys
1265	1270	1275	
Arg Val	Ile Leu Ala Asp Ala	Asn Leu Asp Lys Val	Leu Ser Ala
1280	1285	1290	
Tyr Asn	Lys His Arg Asp Lys	Pro Ile Arg Glu Gln	Ala Glu Asn
1295	1300	1305	
Ile Ile	His Leu Phe Thr Leu	Thr Asn Leu Gly Ala	Pro Ala Ala

1310 1315 1320
Phe Lys Tyr Phe Asp Thr Thr Ile Asp Arg Lys Arg Tyr Thr Ser
1325 1330 1335
Thr Lys Glu Val Leu Asp Ala Thr Leu Ile His Gln Ser Ile Thr
1340 1345 1350
Gly Leu Tyr Glu Thr Arg Ile Asp Leu Ser Gln Leu Gly Gly Asp

1355 1360 1365
<210> 14
<211> 4
<212> PRT
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
linker peptide

<400> 14
Gly Gly Gly Ser
1

<210> 15
<211> 5
<212> PRT
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
linker peptide

<400> 15
Gly Gly Gly Gly Ser
1 5

<210> 16
<211> 23
<212> DNA
<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
target binding site oligonucleotide

<400> 16
gtgtgcagac ggcagtcact agg

<210> 17

<211> 23

<212> DNA

<213> Artificial Sequence

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

target binding site oligonucleotide

<400> 17

gagcagcgtc ttcgagagtg agg 23

<210> 18

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223>

> Description of Artificial Sequence: Synthetic

target binding site oligonucleotide

<400> 18

ggtgagttag tgtgtgcgtg tgg 23

<210> 19

<211> 23

<212> DNA

<213> Artificial Sequence

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

target binding site oligonucleotide

<400> 19

gttgagcgg ggagaaggcc agg 23

<210> 20

<211> 23

<212> DNA

<213> Artificial Sequence

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

target binding site oligonucleotide

<400> 20

gggtgggggg agtttgctcc tgg 23

<210> 21

<211> 23

<212> DNA

<213> Artificial Sequence

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

target binding site oligonucleotide

<400> 21

ggcttttgaa agggggtggg ggg 23

<210> 22

<211> 23

<212> DNA

<213> Artificial Sequence

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

target binding site oligonucleotide

<400> 22

ggggcgggggt cccggcgggg cgg 23

<210> 23

<211> 23

<212> DNA

<213> Artificial Sequence

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

target binding site oligonucleotide

<400> 23

gctcggaggt cgtggcgctg ggg 23

<210> 24

<211> 23

<212> DNA

<213> Artificial Sequence

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

target binding site oligonucleotide

<400> 24

gactcaccgg ccaggcgct cgg 23

<210> 25

<211> 23

<212> DNA

<213> Artificial Sequence

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

target binding site oligonucleotide

<400> 25

ggcgcagcgg ttaggtggac cgg 23

<210> 26

<211> 23

<212> DNA

<213> Artificial Sequence

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

target binding site oligonucleotide

<400> 26

ggcgcacatggc tccgccccgc cgg 23

<210> 27

<211> 23

<212> DNA

<213> Artificial Sequence

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

target binding site oligonucleotide

<400> 27

gccacgacct ccgagctacc cgg 23

<210> 28

<211> 23

<212> DNA

<213> Artificial Sequence

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

target binding site oligonucleotide

<400> 28

gcggcgtgag ccctccccct tgg 23

<210> 29

<211> 23

<212> DNA

<213> Artificial Sequence

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

target binding site oligonucleotide

<400> 29

ggaggcgggg tggaggggt cgg 23

<210> 30

<211> 23

<212> DNA

<213> Artificial Sequence

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

target binding site oligonucleotide

<400> 30

gggctcacgc cgcgtccgg cgg 23

<210> 31

<211> 23

<212> DNA

<213> Artificial Sequence

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

target binding site oligonucleotide

<400> 31

gacccctcc acccgctc cgg 23

<210> 32

<211> 23

<212> DNA

<213> Artificial Sequence

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

target binding site oligonucleotide

<400> 32

gagcgaggag ccatctggc ggg 23

<210> 33

<211> 23

<212> DNA

<213> Artificial Sequence

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

target binding site oligonucleotide

<400> 33

gcgcggcgcg gaaggggtta agg

23

<210> 34

<211> 23

<212> DNA

<213> Artificial Sequence

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

target binding site oligonucleotide

<400> 34

gcggcgcggc gcgggccggc ggg

23

<210> 35

<211> 23

<212> DNA

<213> Artificial Sequence

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

target binding site oligonucleotide

<400> 35

gccgcgccgc cctccccgc cgg

23

<210> 36

<211> 23

<212> DNA

<213> Artificial Sequence

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

target binding site oligonucleotide

<400> 36

gcggttataa ccagccaacc cgg

23

<210> 37

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223>

> Description of Artificial Sequence: Synthetic

target binding site oligonucleotide

<400> 37

gtgcgcggag ctgttcggaa ggg 23

<210> 38

<211> 26

<212> DNA

<213> Artificial Sequence

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

target binding site oligonucleotide

<400> 38

acaccgtgtg cagacggcag tcactg 26

<210> 39

<211> 26

<212> DNA

<213> Artificial Sequence

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

target binding site oligonucleotide

<400> 39

acaccgagca gcgtcttcga gagtgg 26

<210> 40

<211> 26

<212> DNA

<213> Artificial Sequence

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

target binding site oligonucleotide

<400> 40

acaccgtga gtgagtgtgt gcgtgg 26

<210> 41

<211> 26

<212> DNA

<213> Artificial Sequence

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

target binding site oligonucleotide

<400> 41
acaccgttgg agcggggaga aggccg 26
<210> 42
<211> 26
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
target binding site oligonucleotide
<400> 42
acaccgggtg gggggagttt gctccg 26
<210> 43
<211> 26
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
target binding site oligonucleotide
<400> 43

acaccggctt tggaaagggg gtgggg 26
<210> 44
<211> 26
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
target binding site oligonucleotide
<400> 44
acaccggggc ggggtcccg cggggg 26
<210> 45
<211> 26
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
target binding site oligonucleotide
<400> 45

acaccgctcg gaggtcgtgg cgctgg 26

<210> 46

<211> 26

<212> DNA

<213> Artificial Sequence

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

target binding site oligonucleotide

<400> 46

acaccgactc accggccagg gcgctg 26

<210> 47

<211> 26

<212> DNA

<213> Artificial Sequence

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

target binding site oligonucleotide

<400> 47

acaccggcgc agcggttagg tggacg 26

<210> 48

<211> 26

<212> DNA

<213> Artificial Sequence

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

target binding site oligonucleotide

<400> 48

acaccggcgc atggctccgc cccgcg 26

<210> 49

<211> 26

<212> DNA

<213> Artificial Sequence

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

target binding site oligonucleotide

<400> 49

acaccgccac gacctccgag ctaccg 26

<210> 50

<211> 26

<212> DNA

<213> Artificial Sequence

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

target binding site oligonucleotide

<400> 50

acaccgcggc gtgagccctc cccctg 26

<210> 51

<211> 26

<212> DNA

<213> Artificial Sequence

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

target binding site oligonucleotide

<400> 51

acaccggagg cggggtggag ggggtg 26

<210> 52

<211> 26

<212> DNA

<213> Artificial Sequence

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

target binding site oligonucleotide

<400> 52

acaccgggct cagccgcgc tccggg 26

<210> 53

<211> 26

<212> DNA

<213> Artificial Sequence

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

target binding site oligonucleotide

<400> 53

acaccgaccc cctccacccc gcctcg 26

<210> 54

<211> 26

<212> DNA

<213> Artificial Sequence

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

target binding site oligonucleotide

<400> 54

acaccgagcg cggagccatc tggccg

26

<210> 55

<211> 26

<212> DNA

<213> Artificial Sequence

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

target binding site oligonucleotide

<400> 55

acaccgcgcg gcgcggaagg ggtag

26

<210> 56

<211> 26

<212> DNA

<213> Artificial Sequence

<220><223>

> Description of Artificial Sequence: Synthetic

target binding site oligonucleotide

<400> 56

acaccgcggc gcggcgcggg ccggcg

26

<210> 57

<211> 26

<212> DNA

<213> Artificial Sequence

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

target binding site oligonucleotide

<400> 57

acaccgccgc gccgcctcc cccgcg

26

<210> 58

<211> 26

<212> DNA

<213> Artificial Sequence

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

target binding site oligonucleotide

<400> 58

acaccgcggt tataaccagc caaccg

26

<210> 59

<211> 26

<212> DNA

<213> Artificial Sequence

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

target binding site oligonucleotide

<400> 59

acaccgtgcg cggagctgtt cggaag

26

<210> 60

<211> 26

<212> DNA

<213> Artificial Sequence

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

target binding site oligonucleotide

<400> 60

aaaacagtga ctgccgtctg cacacg

26

<210> 61

<211> 26

<212> DNA

<213> Artificial Sequence

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

target binding site oligonucleotide

<400> 61

aaaaccactc tcgaagacgc tgctcg

26

<210> 62

<211> 26

<212> DNA

<213> Artificial Sequence

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

target binding site oligonucleotide

<400> 62

aaaaccacgc acacactcac tcaccg

26

<210> 63

<211> 26

<212> DNA

<213> Artificial Sequence

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

target binding site oligonucleotide

<400> 63

aaaacggcct tctccccgct ccaacg

26

<210> 64

<211> 26

<212> DNA

<213> Artificial Sequence

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

target binding site oligonucleotide

<400> 64

aaaacggagc aaactcccc caccg

26

<210> 65

<211> 26

<212> DNA

<213> Artificial Sequence

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

target binding site oligonucleotide

<400> 65

aaaacccac ccctttcca aagccg

26

<210> 66

<211> 26

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
target binding site oligonucleotide
<400> 66
aaaacccccg ccgggacccc gccccg 26
<210> 67

<211> 26
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
target binding site oligonucleotide
<400> 67
aaaaccagcg ccacgacctc cgagcg 26
<210> 68
<211> 26
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
target binding site oligonucleotide
<400> 68
aaaacagcgc cctggccggt gagtcg 26
<210> 69
<211> 26

<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
target binding site oligonucleotide
<400> 69
aaaacgtcca cctaaccgct gcgccg 26
<210> 70
<211> 26
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic

target binding site oligonucleotide

<400> 70

aaaacgcggg gcggagccat gcgccg 26

<210> 71

<211> 26

<212> DNA

<213> Artificial Sequence

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

target binding site oligonucleotide

<400> 71

aaaacggtag ctcgagggtc gtggcg 26

<210> 72

<211> 26

<212> DNA

<213> Artificial Sequence

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

target binding site oligonucleotide

<400> 72

aaaacagggg gagggctcac gccgcg 26

<210> 73

<211> 26

<212> DNA

<213> Artificial Sequence

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

target binding site oligonucleotide

<400> 73

aaaacacccc ctccaccccg cctccg 26

<210> 74

<211> 26

<212> DNA

<213> Artificial Sequence

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

target binding site oligonucleotide

<400> 74
aaaacccgga gcgcggcgtg agcccg 26
<210> 75
<211> 26
<212> DNA
<213> Artificial Sequence
<220><223>
> Description of Artificial Sequence: Synthetic
target binding site oligonucleotide
<400> 75
aaaacgaggc ggggtggagg gggtcg 26
<210> 76
<211> 26
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
target binding site oligonucleotide
<400> 76
aaaacggcca gatggctccg cgctcg 26
<210> 77
<211> 26
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
target binding site oligonucleotide
<400> 77
aaaactaacc ccttccgcgc cgcgcg 26
<210> 78
<211> 26
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
target binding site oligonucleotide
<400> 78

aaaacgccgg cccgcgccgc gccgcg 26

<210> 79

<211> 26

<212> DNA

<213> Artificial Sequence

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

target binding site oligonucleotide

<400> 79

aaaacgcggg ggaggcgcc gcggcg 26

<210> 80

<211> 26

<212> DNA

<213> Artificial Sequence

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

target binding site oligonucleotide

<400> 80

aaaacggttg gctggttata accgcg 26

<210> 81

<211> 26

<212> DNA

<213> Artificial Sequence

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

target binding site oligonucleotide

<400> 81

aaaacttccg aacagctccg cgcacg 26

<210> 82

<211> 20

<212> DNA

<213> Artificial Sequence

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

primer

<400> 82

tccagatggc acattgtcag 20

<210> 83
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic
 primer
 <400> 83
 agggagcagg aaagtgaggt 20

<210> 84
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic
 primer
 <400> 84
 gcacgtaacc tcactttcct 20

<210> 85
 <211> 23
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic
 primer
 <400> 85
 cttgctacct ctttcctctt tct 23

<210> 86
 <211> 22
 <212> DNA
 <213> Artificial Sequence
 <220>
 ><223> Description of Artificial Sequence: Synthetic
 primer
 <400> 86
 agagaagtcg aggaagagag ag 22

<210> 87

<211> 22

<212> DNA

<213> Artificial Sequence

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

primer

<400> 87

cagcagaaag ttcatggttt cg

22

<210> 88

<211> 130

<212> PRT

<213> Enterobacteria phage lambda

<400> 88

Met Ala Ser Asn Phe Thr Gln Phe Val Leu Val Asp Asn Gly Gly Thr

1 5 10 15
Gly Asp Val Thr Val Ala Pro Ser Asn Phe Ala Asn Gly Val Ala Glu

20 25 30
Trp Ile Ser Ser Asn Ser Arg Ser Gln Ala Tyr Lys Val Thr Cys Ser

35 40 45
Val Arg Gln Ser Ser Ala Gln Asn Arg Lys Tyr Thr Ile Lys Val Glu

50 55 60
Val Pro Lys Val Ala Thr Gln Thr Val Gly Gly Val Glu Leu Pro Val

65 70 75 80
Ala Ala Trp Arg Ser Tyr Leu Asn Met Glu Leu Thr Ile Pro Ile Phe

85 90 95
Ala Thr Asn Ser Asp Cys Glu Leu Ile Val Lys Ala Met Gln Gly Leu

100 105 110
Leu Lys Asp Gly Asn Pro Ile Pro Ser Ala Ile Ala Ala Asn Ser Gly

115 120 125
Ile Tyr

130

<210> 89

<211> 130

<212>

> PRT

<213> Artificial Sequence

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

lambda bacteriophage MS2 N55K mutant polypeptide

<400> 89

Met Ala Ser Asn Phe Thr Gln Phe Val Leu Val Asp Asn Gly Gly Thr

1 5 10 15

Gly Asp Val Thr Val Ala Pro Ser Asn Phe Ala Asn Gly Val Ala Glu

20 25 30

Trp Ile Ser Ser Asn Ser Arg Ser Gln Ala Tyr Lys Val Thr Cys Ser

35 40 45

Val Arg Gln Ser Ser Ala Gln Lys Arg Lys Tyr Thr Ile Lys Val Glu

50 55 60

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

65 70 75 80

Ala Ala Trp Arg Ser Tyr Leu Asn Met Glu Leu Thr Ile Pro Ile Phe

85 90 95

Ala Thr Asn Ser Asp Cys Glu Leu Ile Val Lys Ala Met Gln Gly Leu

100 105 110

Leu Lys Asp Gly Asn Pro Ile Pro Ser Ala Ile Ala Ala Asn Ser Gly

115 120 125

Ile Tyr

130

<210> 90

<211> 117

<212> PRT

<213> Artificial Sequence

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

lambda bacteriophage MS2 deltaFG mutant polypeptide

<400> 90

Met Ala Ser Asn Phe Thr Gln Phe Val Leu Val Asp Asn Gly Gly Thr

1 5 10 15

Gly Asp Val Thr Val Ala Pro Ser Asn Phe Ala Asn Gly Ile Ala Glu

20 25 30

Trp Ile Ser Ser Asn Ser Arg Ser Gln Ala Tyr Lys Val Thr Cys Ser

35 40 45

Val Arg Gln Ser Ser Ala Gln Asn Arg Lys Tyr Thr Ile Lys Val Glu

50 55 60

Val Pro Lys Gly Ala Trp Arg Ser Tyr Leu Asn Met Glu Leu Thr Ile

65 70 75 80

Pro Ile Phe Ala Thr Asn Ser Asp Cys Glu Leu Ile Val Lys Ala Met

85 90 95

Gln Gly Leu Leu Lys Asp Gly Asn Pro Ile Pro Ser Ala Ile Ala Ala

100 105 110

Asn Ser Gly Ile Tyr

115

<210> 91

<211> 262

<212> PRT

<213> Artificial Sequence

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

dimeric MS2 coat polypeptide

<400> 91

Met Ala Ser Asn Phe Thr Gln Phe Val Leu Val Asp Asn Gly Gly Thr

1 5 10 15

Gly Asp Val Thr Val Ala Pro Ser Asn Phe Ala Asn Gly Val Ala Glu

20 25 30

Trp Ile Ser Ser Asn Ser Arg Ser Gln Ala Tyr Lys Val Thr Cys Ser

35 40 45

Val Arg Gln Ser Ser Ala Gln Asn Arg Lys Tyr Thr Ile Lys Val Glu

50 55 60

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

65 70 75 80

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

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

Glu Leu Pro Val Ala Ala Trp Arg Ser Tyr Leu Asn Met Glu Leu Thr
210 215 220
Ile Pro Ile Phe Ala Thr Asn Ser Asp Cys Glu Leu Ile Val Lys Ala
225 230 235 240
Met Gln Gly Leu Leu Lys Asp Gly Asn Pro Ile Pro Ser Ala Ile Ala
245 250 255
Ala Asn Ser Leu Ile Asn
260

<210> 92

<211> 262

<212> PRT

<213> Artificial Sequence

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

dimeric MS2 N55K mutant coat polypeptide

<400> 92

Met Ala Ser Asn Phe Thr Gln Phe Val Leu Val Asp Asn Gly Gly Thr

1	5	10	15
Gly Asp Val Thr	Val Ala Pro Ser Asn Phe	Ala Asn Gly Val	Ala Glu
20	25	30	
Trp Ile Ser Ser	Asn Ser Arg Ser Gln Ala Tyr Lys	Val Thr Cys Ser	
35	40	45	
Val Arg Gln Ser	Ser Ala Gln Lys Arg Lys Tyr Thr	Ile Lys Val Glu	
50	55	60	
Val Pro Lys Val	Ala Thr Gln Thr Val Gly Gly Val	Glu Leu Pro Val	
65	70	75	80
Ala Ala Trp Arg	Ser Tyr Leu Asn Met Glu Leu Thr	Ile Pro Ile Phe	
85	90	95	
Ala Thr Asn Ser	Asp Cys Glu Leu Ile Val Lys Ala Met	Gln Gly Leu	
100	105	110	
Leu Lys Asp Gly	Asn Pro Ile Pro Ser Ala Ile Ala Ala	Asn Ser Gly	
115	120	125	
Leu Tyr Gly Ala	Met Ala Ser Asn Phe Thr Gln Phe Val	Leu Val Asp	
130	135	140	
Asn Gly Gly Thr	Gly Asp Val Thr Val Ala Pro Ser	Asn Phe Ala Asn	
145	150	155	160
Gly Val Ala Glu	Trp Ile Ser Ser Asn Ser Arg Ser	Gln Ala Tyr Lys	
165	170	175	
Val Thr Cys Ser	Val Arg Gln Ser Ser Ala Gln Lys Arg	Lys Tyr Thr	
180	185	190	
Ile Lys Val Glu	Val Pro Lys Val Ala Thr Gln Thr	Val Gly Gly Val	
195	200	205	
Glu Leu Pro Val	Ala Ala Trp Arg Ser Tyr Leu Asn Met	Glu Leu Thr	
210	215	220	
Ile Pro Ile Phe	Ala Thr Asn Ser Asp Cys Glu Leu Ile	Val Lys Ala	
225	230	235	240
Met Gln Gly Leu	Leu Lys Asp Gly Asn Pro Ile Pro Ser	Ala Ile Ala	
245	250	255	

Ala Asn Ser Leu Ile Asn

260

<210> 93

<211> 236

<212> PRT

<213> Artificial Sequence

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

dimeric MS2 deltaFG mutant coat polypeptide

<400> 93

Met Ala Ser Asn Phe Thr Gln Phe Val Leu Val Asp Asn Gly Gly Thr

1 5 10 15

Gly Asp Val Thr Val Ala Pro Ser Asn Phe Ala Asn Gly Val Ala Glu

20 25 30

Trp Ile Ser Ser Asn Ser Arg Ser Gln Ala Tyr Lys Val Thr Cys Ser

35 40 45

Val Arg Gln Ser Ser Ala Gln Lys Arg Lys Tyr Thr Ile Lys Val Glu

50 55 60

Val Pro Lys Gly Ala Trp Arg Ser Tyr Leu Asn Met Glu Leu Thr Ile

65 70 75 80

Pro Ile Phe Ala Thr Asn Ser Asp Cys Glu Leu Ile Val Lys Ala Met

85 90 95

Gln Gly Leu Leu Lys Asp Gly Asn Pro Ile Pro Ser Ala Ile Ala Ala

100 105 110

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

115 120 125

Leu Val Asp Asn Gly Gly Thr Gly Asp Val Thr Val Ala Pro Ser Asn

130 135 140

Phe Ala Asn Gly Val Ala Glu Trp Ile Ser Ser Asn Ser Arg Ser Gln

145 150 155 160

Ala Tyr Lys Val Thr Cys Ser Val Arg Gln Ser Ser Ala Gln Lys Arg

165 170 175

Lys Tyr Thr Ile Lys Val Glu Val Pro Lys Gly Ala Trp Arg Ser Tyr

180 185 190
 Leu Asn Met Glu Leu Thr Ile Pro Ile Phe Ala Thr Asn Ser Asp Cys
 195 200 205
 Glu Leu Ile Val Lys Ala Met Gln Gly Leu Leu Lys Asp Gly Asn Pro

210 215 220
 Ile Pro Ser Ala Ile Ala Ala Asn Ser Leu Ile Asn
 225 230 235

<210> 94

<211> 22

<212> PRT

<213> Enterobacteria phage lambda

<400> 94

Met Asp Ala Gln Thr Arg Arg Arg Glu Arg Arg Ala Glu Lys Gln Ala

1 5 10 15

Gln Trp Lys Ala Ala Asn

20

<210> 95

<211> 107

<212> PRT

<213> Enterobacteria phage lambda

<400> 95

Met Asp Ala Gln Thr Arg Arg Arg Glu Arg Arg Ala Glu Lys Gln Ala

1 5 10 15

Gln Trp Lys Ala Ala Asn Pro Leu Leu Val Gly Val Ser Ala Lys Pro

20 25 30

Val Asn Arg Pro Ile Leu Ser Leu Asn Arg Lys Pro Lys Ser Arg Val

35 40 45

Glu Ser Ala Leu Asn Pro Ile Asp Leu Thr Val Leu Ala Glu Tyr His

50 55 60

Lys Gln Ile Glu Ser Asn Leu Gln Arg Ile Glu Arg Lys Asn Gln Arg

65 70 75 80

Thr Trp Tyr Ser Lys Pro Gly Glu Arg Gly Ile Thr Cys Ser Gly Arg

	85	90	95
Gln Lys Ile Lys Gly Lys Ser Ile Pro Leu Ile			
	100	105	
<210> 96			
<211> 23			
<212> RNA			
<213> Enterobacteria phage lambda			
<400> 96			
aaacaugagg auuacccaug ucg			23
<210> 97			
<211> 23			
<212> RNA			
<213> Artificial Sequence			
<220><223> Description of Artificial Sequence: Synthetic			
high affinity MS2 binding oligonucleotide			
<400> 97			
aaacaugagg aucacccaug ucg			23
<210> 98			
<211> 15			
<212> RNA			
<213> Enterobacteria phage lambda			
<400> 98			
gcccuagaaga agggc			15
<210> 99			
<211> 15			
<212> RNA			
<213> Enterobacteria phage lambda			
<400> 99			
gcccuagaaaa agggc			15
<210> 100			
<211> 20			
<212> DNA			
<213> Artificial Sequence			

<220><223> Description of Artificial Sequence: Synthetic
truncated Csy4 binding site oligonucleotide

<220><223> Description of combined DNA/RNA molecule: Synthetic
truncated Csy4 binding site oligonucleotide

<400> 100
gttcactgcc gtataggcag 20

<210> 101
<211> 28
<212> RNA
<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
Csy4 binding site oligonucleotide

<400> 101
guucacugcc guauaggcag cuaagaaa 28

<210> 102
<211> 32
<212> RNA
<213> Artificial Sequence

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

<220><221> modified_base
<222> (1)..(20)
<223> a, c, u, g, unknown or other and this region may encompass
17-20 nucleotides, wherein some positions may be absent

<400> 102
nnnnnnnnnn nnnnnnnnnn guuuuagagc ua 32

<210> 103

<211> 42
<212> RNA
<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
crRNA oligonucleotide
<220><221> modified_base

<222> (1)..(20)

<223> a, c, u, g, unknown or other and this region may encompass
17-20 nucleotides, wherein some positions may be absent

<400> 103

nnnnnnnnnn nnnnnnnnnn guuuuagagc uaucuguuu ug 42

<210> 104

<211> 36

<212> RNA

<213> Artificial Sequence

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

crRNA oligonucleotide

<220><221> modified_base

<222> (1)..(20)

<223> a, c, u, g, unknown or other and this region may encompass
17-20 nucleotides, wherein some positions may be absent

<400> 104

nnnnnnnnnn nnnnnnnnnn guuuuagagc uaugcu 36

<210> 105

<211> 60

<212> RNA

<213> Artificial Sequence

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

tracrRNA oligonucleotide

<400> 105

uagcaaguua aaauaagguu aguccguuau caacuugaaa aaguggcacc gagucggugc 60

<210> 106

<211> 64

<212> RNA

<213> Artificial Sequence

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

tracrRNA oligonucleotide

<400> 106

agcauagcaa guaaaaauaa ggcuaguccg uuaucaacuu gaaaaagugg caccgagucg 60

gugc	64
<210> 107	
<211> 2279	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic	
RNA expression vector polynucleotide	
<220><221> modified_base	
<222> (331)..(350)	
<223> a, c, t, g, unknown or other	
<400> 107	
gacgtcgcta gcgtacaaa aaagcaggct ttaaaggaac caattcagtc gactggatcc	60
ggtagcaagg tcgggcagga agagggccta tttcccatga ttccttcata ttgcatata	120
cgatacaagg ctgttagaga gataattaga attaatgtga ctgtaaacac aaagatatta	180
gtacaaaata cgtgacgtag aaagtaataa tttcttgggt agtttgcagt tttaaaatta	240
tgttttaaaa tggactatca tatgcttacc gtaacttgaa agtatcttga tttcttggct	300
ttatatatct tgttgaaagg acgaaacacc nnnnnnnnnn nnnnnnnnnn gtttttagagc	360
tagaaatagc aagttaaaat aaggctagtc cggtatcaac ttgaaaaagt ggcaccgagt	420
cgggtgctttt ttttagcttg ggccgctcga ggtacctctc tacatatgac atgtgagcaa	480
aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt tcccataggc	540
tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg cgaaaccga	600
caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc tctcctgttc	660
cgacctgcc gcttaccgga taccgtccg cctttctccc ttcgggaagc gtggcgcttt	720
ctcatagctc acgctgtagg tatctcagtt cgggttaggt cgctcgtcc aagctgggct	780
gtgtgcacga accccccgtt cagcccgacc gctgcgcctt atccggtaac tatcgtcttg	840
agtccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt aacaggatta	900
gcagagcgag gtatgtaggc ggtgtctacag agttcttgaa gtggtggcct aactacggct	960
acactagaag aacagtatct ggtatctcgc ctctgctgaa gccagttacc ttcggaaaaa	1020
gagttggtag ctcttgatcc ggcaaaaaa ccaccgctgg tagcgggtgt tttttgttt	1080
gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatcctttg atcttttcta	1140
cggggtctga cgctcagtg aacgaaaact cacgttaagg gattttggtc atgagattat	1200

caaaaaggat cttcacctag atccttttaa attaaaaatg aagttttaaa tcaatctaaa 1260
 gtatatatga gtaaacttgg tctgacagtt accaatgctt aatcagttag gcacctatct 1320
 cagcgatctg tctatcttctg tcatccatag ttgctgact ccccgctctg tagataacta 1380
 cgatacggga gggcttaccg tctggcccca gtgctgcaat gataccgcga gaccacagct 1440
 caccggtccc agatttatca gcaataaacc agccagccgg aaggcccgag cgcagaagtg 1500
 gtcttgaac tttatccgcc tccatccagt ctattaattg ttgccgggaa gctagagtaa 1560
 gtagttcgcc agttaatagt ttgcgcaacg ttgttgccat tgctacaggc atcgtggtgt 1620

cagctctgtc gtttggatg gcttcattca gctccggttc ccaacgatca aggcgagtta 1680
 catgatcccc catgttctgc aaaaaagcgg ttagctcctt cggctcctcg atcgttctca 1740
 gaagtaagtt ggccgcagtg ttatcactca tggttatggc agcactgcat aattctctta 1800
 ctgtcatgcc atccgtaaga tgcttttctg tgactggtga gtactcaacc aagtcattct 1860
 gagaatagtg tatcgccgca ccgagttgct cttgcccggc gtcaatacgg gataataccg 1920
 cgccacatag cagaacttta aaagtgtca tcattggaaa acgttcttcg gggcgaaaac 1980
 tctcaaggat cttaccgctg ttgagatcca gtctgatgta acccactcgt gcaccaact 2040

gatcttcagc atcttttact ttcaccagcg tttctgggtg agcaaaaaca ggaaggcaaa 2100
 atgccgcaaa aaagggaata agggcgacac ggaaatgttg aatactcata ctcttccttt 2160
 ttcaatatta ttgaagcatt tatcagggtt attgtctcat gagcggatag atatttgaat 2220
 gtatttagaa aaataaaca ataggggttc cgcgcacatt tccccgaaa gtgccacct 2279

<210> 108

<211> 7786

<212> DNA

<213> Artificial Sequence

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

CMV-T7-Cas9 D10A/H840A-3xFlag-VP64 polynucleotide

<400> 108

atatgccaag tacgccccct attgacgtca atgacggtaa atggcccgcc tggcattatg 60

cccagtacat gaccttatgg gactttctca cttggcagta catctacgta ttagtcacg 120
 ctattacat ggtagatcgg ttttggcagt acatcaatgg gcgtggatag cggtttgact 180
 cacggggatt tccaagtcct caccctattg acgtcaatgg gagtttggtt tggcaccaaa 240
 atcaacggga ctttccaaaa tgtcgtaaca actccgcccc attgacgcaa atgggcggta 300
 ggctgtacg gtgggaggtc tatataagca gagctggttt agtgaaccgt cagatccgct 360
 agagatccgc ggccgctaata acgactcact ataggagag cgcaccat ggataagaaa 420

tactcaatag gcttagctat cggcacaaat agcgtcggat gggcggatgat cactgatgaa	480
tataaggttc cgtctaaaaa gttcaagggt ctgggaaata cagaccgcca cagtatcaaa	540
aaaaatctta taggggctct tttatttgac agtggagaga cagcggaaagc gactcgtctc	600
aaacggacag ctctagaag gtatacacgt cggaagaatc gtatttgta tctacaggag	660
atTTTTTcaa atgagatggc gaaagtagat gatagtttct ttcacgact tgaagagtct	720
TTTTTggTgg aagaagacaa gaagcatgaa cgtcatccta TTTTTggaaa tatagtagat	780
gaagttgctt atcatgagaa atatccaact atctatcatc tgcgaaaaaa attggtagat	840
tctactgata aagcggattt gcgcttaatc tatttggcct tagcgcatat gattaagttt	900
cgTggTcatt ttttgattga gggagattta aatcctgata atagtgatgt ggacaaacta	960
tttatccagt tggTacaac ctacaatcaa ttatttgaag aaaaccctat taacgcaagt	1020
ggagtagatg ctaaagcgat tctttctgca cgattgagta aatcaagacg attagaaaaat	1080
ctcattgctc agtccccgg tgagaagaaa aatggcttat ttgggaatct cattgctttg	1140
tcatTgggtt tgaccctaa ttttaaatca aattttgatt tggcagaaga tgctaaatta	1200
cagctttcaa aagatactta cgatgatgat ttagataatt tattggcgca aattggagat	1260
caatatgctg atttgTTTTT ggcagctaag aatttatcag atgctatTTT actttcagat	1320
atcctaagag taaatactga aataactaag gctccctat cagcttcaat gattaaacgc	1380
tacgatgaac atcatcaaga cttgactctt ttaaaagctt tagttcgaca acaacttcca	1440
gaaaagtata aagaaatctt ttttgatcaa tcaaaaaacg gatatgcagg ttatatTgat	1500
gggggagcta gccaagaaga attttataaa tttatcaaac caattttaga aaaaatggat	1560
ggtactgagg aattattggt gaaactaaat cgtgaagatt tgctgcgcaa gcaacggacc	1620
tttgacaacg gctctattcc ccatcaaat cacttgggtg agtgcgatgc tattttgaga	1680
agacaagaag acttttatcc atttttaaaa gacaatcgtg agaagattga aaaaatcttg	1740
acttttcgaa ttccttatta tgTtggtcca ttggcgcgtg gcaatagtcg ttttgcatgg	1800
atgactcgga agtctgaaga aacaattacc ccatggaatt ttgaagaagt tgTcgataaa	1860
ggtgcttcag ctcaatcatt tattgaacgc atgacaaact ttgataaaaa tcttccaaat	1920
gaaaaagtac taccaaaaca tagtttgctt tatgagtatt ttacggttta taacgaattg	1980
acaaaggTca aatatgttac tgaaggaatg cgaaaaccag catttcttTc aggtgaacag	2040
aagaaagcca ttgttgattt actcttcaaa acaaatcgaa aagtaaccgt taagcaatta	2100
aaagaagatt atttcaaaaa aatagaatgt ttgatagtg ttgaaatttc aggagtTgaa	2160

gatagattta atgcttcatt aggtacctac catgatttgc taaaaattat taaagataaa	2220
gatttttttg ataatagaaga aaatgaagat atcttagagg atattgtttt aacattgacc	2280
ttatttgaag ataggagat gattgaggaa agacttaaaa catatgctca cctctttgat	2340
gataaggatga tgaacacagct taaacgtcgc cgttatactg gttggggacg tttgtctcga	2400
aaattgatta atggtattag ggataagcaa tctggcaaaa caatattaga ttttttgaaa	2460
tcagatggtt ttgccaatcg caattttatg cagctgatcc atgatgatag ttigacattt	2520
aaagaagaca ttcaaaaagc acaagtgtct ggacaaggcg atagtttaca tgaacatatt	2580
gcaaatttag ctggtagccc tgctattaaa aaaggtatit tacagactgt aaaagttgtt	2640
gatgaattgg tcaaagtaat ggggcggcat aagccagaaa atatcgttat tgaaatggca	2700
cgtgaaaatc agacaactca aaagggccag aaaaattcgc gagagcgtat gaaacgaatc	2760
gaagaaggta tcaagaattt aggaagtcag attcttaaag agcatcctgt tgaaaatact	2820
caattgcaaa atgaaaagct ctatctctat tatctccaaa atggaagaga catgtatgtg	2880
gaccaagaat tagatattaa tcgtttaagt gattatgatg tcgatgcat tgttccacaa	2940
agtttcctta aagacgattc aatagacaat aaggtcttaa cgcgttctga taaaaatcgt	3000
ggtaaatcgg ataacgttcc aagtgaagaa gtagtcaaaa agatgaaaa ctattggaga	3060
caacttctaa acgccaagtt aatcactcaa cgtaagtittg ataatttaac gaaagctgaa	3120
cgtggagggtt tgagtgaact tgataaagct ggttttatca aacgccaatt ggttgaaact	3180
cgccaaatca ctaagcatgt ggcacaaatt ttggatagtc gcatgaatac taaatacgat	3240
gaaaatgata aacttattcg agaggttaaa gtgattacct taaaatctaa attagtttct	3300
gacttccgaa aagatttcca attctataaa gtacgtgaga ttaacaatta ccatcatgcc	3360
catgatgcgt atctaaatgc cgtcgttgga actgctttga ttaagaaata tccaaaactt	3420
gaatcggagt ttgtctatgg tgattataaa gtttatgatg ttcgtaaaat gattgctaag	3480
tctgagcaag aaataggcaa agcaaccgca aaatatttct tttactctaa tatcatgaac	3540
ttcttcaaaa cagaaattac acttgcaaat ggagagattc gcaaacgccc tctaactgaa	3600
actaatgggg aaactggaga aattgtctgg gataaagggc gagattttgc cacagtgcgc	3660
aaagtattgt ccatgcccc agtcaatat gtcaagaaaa cagaagtaca gacaggcgga	3720
ttctccaagg agtcaatttt accaaaaaga aattcggaca agcttattgc tcgtaaaaaa	3780
gactgggac caaaaaata tgggtggttt gatagtccaa cggtagctta ttcagtccta	3840
giggttgcta aggttgaaaa agggaaatcg aagaagttaa aatccgttaa agagttacta	3900
gggatcacia ttatggaaag aagttccttt gaaaaaaatc cgattgactt tttagaagct	3960
aaaggatata aggaagttaa aaaagactta atcattaaac tacctaaata tagtcttttt	4020

gagttagaaa acggtcgtaa acggatgctg gctagtgccg gagaattaca aaaaggaaat 4080
 gagctggctc tgccaagcaa atatgtgaat tttttatatt tagctagtca ttatgaaaag 4140
 ttgaagggtg gtccagaaga taacgaacaa aaacaattgt ttgtggagca gcataagcat 4200
 tatttagatg agattattga gcaaatcagt gaattttcta agcgtgttat tttagcagat 4260

 gccaatntag ataaagtctt tagtgcatat aacaaacata gagacaaacc aatacgtgaa 4320
 caagcagaaa atattattca tttatttacg ttgacgaatc ttggagctcc cgctgctttt 4380
 aaatattttg atacaacaat tgatcgtaaa cgatatacgt ctacaaaaga agtttttagat 4440
 gccactctta tccatcaatc catcactggt ctttatgaaa cacgcattga tttgagtcag 4500
 ctaggagggtg acggtttctc caagaagaag aggaaagtct cgagcgacta caaagacat 4560
 gacggtgatt ataaagatca tgacatcgat tacaaggatg acgatgacaa ggctgcagga 4620
 ggcggtggaa gcgggcgcgc cgacgcgctg gacgatttcg atctcgacat gctgggttct 4680

 gatgccctcg atgactttga cctggatatg ttgggaagcg acgcattgga tgactttgat 4740
 ctggacatgc tcggctccga tgctctggac gatttcgatc tcgatatgtt ataaccggtc 4800
 atcatcacca tcaccattga gtttaaaccg gctgatcagc ctgactgtg ctttctagtt 4860
 gccagccatc tgttgtttgc cctcccccg tgccttcctt gaccttgaa ggtgccactc 4920
 ccactgtcct ttctaataa aatgaggaaa ttgcatcgca ttgtctgagt aggtgtcatt 4980
 ctattctggg gggtggggtg gggcaggaca gcaaggggga ggattgggaa gacaatagca 5040
 ggcatgctgg ggatgcggtg ggctctatgg cttctgaggc ggaaagaacc agctggggct 5100

 cgataccgtc gacctctagc tagagcttgg cgtaatcatg gtcatactg tttcctgtgt 5160
 gaaattgtta tccgctcaca attccacaca acatacgagc cggaagcata aagtgtaaag 5220
 cctagggtgc ctaatgagtg agctaactca cattaattgc gttgcgctca ctgcccgctt 5280
 tccagtcggg aaacctgtcg tgccagctgc attaatgaat cggccaacgc gcggggagag 5340
 gcggtttgcg tattgggcgc tcttcgctt cctcgctcac tgactcgctg cgctcggtcg 5400
 ttcggctcgc gcgagcggtg tcagctcact caaaggcggt aatacgttga tccacagaat 5460
 caggggataa cgcaggaaaag aacatgtgag caaaaggcca gcaaaaggcc aggaaccgta 5520

 aaaaggccgc gttgctggcg tttttccata ggctccgcc ccctgacgag catcacaaaa 5580
 atcgacgtc aagtcagagg tggcgaaacc cgacaggact ataaagatac caggcgtttc 5640
 cccctggaag ctccctcgtg cgctctccig ttccgacct gccgcttacc ggatacctgt 5700
 ccgcctttct ccttcggga agcgtggcgc tttctcaatg ctcacgctgt aggtatctca 5760
 gttcggtgta ggtcgttcgc tccaagctgg gctgtgtgca cgaaccccc gttcagcccc 5820
 accgtcgcgc cttatccggt aactatcgtc ttgagtccaa cccgtaaga cagcattat 5880

cgccactggc agcagccact ggtaacagga ttagcagagc gaggtatgta ggccgtgcta 5940

cagagtctct gaagtgggtg cctaactacg gctacactag aaggacagta ttggtatct 6000

gcgtctgct gaagccagtt accttcggaa aaagagtgg tagctcttga tccggcaaac 6060

aaaccaccgc tggtagcggg ggtttttttg ttgcaagca gcagattacg cgcagaaaaa 6120

aaggatctca agaagatcct ttgatctttt ctacggggtc tgacgctcag tggacgaaa 6180

actcacgtta agggattttg gtcatgagat tatcaaaaag gatcttcacc tagatccttt 6240

taaattaaaa atgaagtttt aaatcaatct aaagtatata tgaglaaact tggcttgaca 6300

gttaccaatg cttaatcagt gaggcaccta tctcagcgat ctgtctatct cgttcacca 6360

tagttgcctg actccccgtc gtgtagataa ctacgatacg ggagggtta ccatctggcc 6420

ccagtgtgc aatgataccg cgagaccac gctaccggc tccagattta tcagcaataa 6480

accagccagc cggaagggcc gagcgagaa gtggtcctgc aactttatcc gcctccatcc 6540

agtctattaa ttgttgcgg gaagctagag taagtagttc gccagttaat agtttgcgca 6600

acgttgttgc cattgtaca ggcatcgtgg tgcacgctc gtcgtttgt atggcttcat 6660

tcagctccgg ttcccaacga tcaaggcgag ttacatgac cccatgttg tgcaaaaaag 6720

cggttagctc cticggctct ccatcgttg tcagaagtaa gttggccgca gtgttatcac 6780

tcatgttat ggcagcactg cataattctc ttactgtcat gccatccgta agatgctttt 6840

ctgtgactgg tgagtactca accaagtcac tctgagaata gtgtatgcgg cgaccgagtt 6900

gctcttggcc ggcgtcaata cgggataata ccgcgccaca tagcagaact ttaaaagtgc 6960

tcatcattgg aaaacgttct tccgggagaa aactctcaag gatcttacg ctgttgagat 7020

ccagttcgat gtaaccact cgtgcacca actgatcttc agcatcttt actttacca 7080

gcgtttctgg gtgagcaaaa acaggaaggc aaaatgccg aaaaaaggga ataaggcgca 7140

cacggaaatg ttgaatactc atactcttc tttttcaata ttattgaagc atttatcagg 7200

gttattgtct catgagcgga tacatatctg aatgtattta gaaaaataa caaatagggg 7260

ttccgcgcac atttccccga aaagtccac ctgacgtcga cggatcgga gatcgatctc 7320

ccgatccct agggtcgact ctcagtacaa tctgctctga tgccgcatag ttaagccagt 7380

atctgtccc tgcctgtgtg ttggaggctg ctgagtagtg cgcgagcaaa atttaagcta 7440

caacaaggca aggtttgacc gacaattgca tgaagaatct gcttagggtt aggcgttttg 7500

cgctgcttcg cgatgtacgg gccagatata cgcgttgaca ttgattattg actagtatt 7560

aatagtaac aattacgggg tcattagttc atagccata tatggagttc cgcgttacat 7620

aacttacggt aaatggcccg cctggctgac cgcccaacga ccccgccca ttgacgtcaa 7680
 taatgacgta tgttcccata gtaacgcaa tagggacttt ccattgacgt caatgggtgg 7740
 actatttacg gtaaacgcc cacttggcag tacatcaagt gtatcc 7786

<210> 109

<211> 7785

<212> DNA

<213> Artificial Sequence

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

MV-T7-Cas9 recoded D10A/H840A-3xFLAG-VP64 polynucleotide

<400> 109

atatgccaag tacgccccct attgacgtca atgacggtaa atggcccgcc tggcattatg 60
 cccagtagat gaccttatgg gactttccta cttggcagta catctacgta ttagtcatcg 120

ctattacat ggtgatgcgg ttttggcagt acatcaatgg gcgtggatag cggtttgact 180
 cacggggatt tccaagtctc caccacattg acgtcaatgg gagtttggtt tggcaccaaa 240
 atcaacggga ctttccaaaa tgtctgaaca actccgcccc attgacgcaa atgggcggta 300
 ggcggtgtag gtgggaggtc tatataagca gagctgggtt agtgaaccgt cagatccgct 360
 agagatccgc ggccgcta at acgactcact ataggagag cggccacat ggataaaaag 420
 tattctattg gtttagccat cggcactaat tccgttggat gggctgtcat aaccgatgaa 480
 taaaaagtag cttcaaagaa atttaagggtg ttggggaaca cagaccgtca ttcgattaaa 540

aagaatctta tcggtgcctt cctattcgat agtggcgaaa cggcagaggc gactcgcctg 600
 aaacgaaccg ctcgagaag gtatacacgt cgcaagaacc gaatatgtta cttacaagaa 660
 attttttagca atgagatggc caaagttgac gattctttct ttcaccgttt ggaagagtcc 720
 ttccttgtcg aagaggacaa gaaacatgaa cggcacccca tctttggaaa catagtagat 780
 gaggtggcat atcatgaaaa gtacccaacg atttatcacc tcagaaaaaa gctagttgac 840
 tcaactgata aagcggacct gaggttaatc tacttggctc ttgcccatat gataaagttc 900
 cgtgggcact ttctcattga gggatgatc aatccggaca actcggatgt cgacaaactg 960

ttcatccagt tagtacaac ctataatcag ttgtttgaag agaaccctat aaatgcaagt 1020
 ggcggtgatg cgaaggctat tcttagcgcc cgcctctcta aatcccgcg gctagaaaaac 1080
 ctgatcgac aattaccgg agagaagaaa aatgggttgt tcggtaacct tatagcgctc 1140
 tcactaggcc tgacaccaa ttttaagtcg aacttcgact tagctgaaga tgccaaattg 1200
 cagcttagta aggacacgta cgatgacgat ctcgacaatc tactggcaca aattggagat 1260
 cagtatgcgg acttatTTTT ggctgcaaaa aaccttagcg atgcaatcct cctatctgac 1320

atactgagag ttaatactga gattaccaag gcgccgttat ccgcttcaat gatcaaaagg 1380

tacgatgaac atcaccaaga cttgacactt ctcaaggccc tagtccgtca gcaactgcct 1440

gagaaatata aggaaatatt ctttgatcag tcgaaaaacg ggtacgcagg ttatattgac 1500

ggcggagcga gtcaagagga attctacaag ttatcaaac ccatattaga gaagatggat 1560

gggacggaag agttgcttgt aaaactcaat cgcgaagatc tactgcgaaa gcagcggact 1620

ttcgacaacg gtagcattcc acatcaaadc cacttaggcg aattgcatgc tatacttaga 1680

aggcaggagg atttttatcc gttcctcaaa gacaatcgtg aaaagattga gaaaatccta 1740

acctttcgca taccttacta tgtgggaccc ctggcccgag ggaactctcg gttcgcatgg 1800

atgacaagaa agtccgaaga aacgattact ccatggaatt ttgaggaagt tgtcgataaa 1860

ggtgcgtcag ctcaatcgtt catcgagagg atgaccaact ttgacaagaa ttaccgaac 1920

gaaaaagtat tgctaagca cagtttactt tacgagtatt tcacagtgtc caatgaactc 1980

acgaaagtta agtatgtcac tgagggcacg cgtaaacccg cttttctaag cggagaacag 2040

aagaaagcaa tagtagatct gttattcaag accaaccgca aagtgcacgt taagcaattg 2100

aaagaggact actttaagaa aattgaatgc ttcgattctg tcgagatctc cggggtagaa 2160

gatcgattta atgcgtcact tggtagctat catgacctcc taaagataat taaagataag 2220

gacttcctgg ataacgaaga gaatgaagat atcttagaag atatagtgtt gactcttacc 2280

ctctttgaag atcgggaaat gattgaggaa agactaaaaa catacgtca cctgttcgac 2340

gataaggta tgaacagtt aaagaggcgt cgctatacgg gctggggacg attgtcgcg 2400

aaacttatca acgggataag agacaagcaa agtggtaaaa ctattctcga ttttctaag 2460

agcgacggct tcgccaatag gaactttaig cagctgatcc atgatgactc tttaaccttc 2520

aaagaggata tacaaaaggc acaggtttcc ggacaagggg actcattgca cgaacatatt 2580

gcgaatcttg ctggttcgcc agccatcaaa aagggcatac tccagacagt caaagtagtg 2640

gatgagctag ttaaggtcat gggacgtcac aaaccgaaa acattgtaat cgagatggca 2700

cgcaaaaac aaacgactca gaaggggcaa aaaaacagtc gagagcggat gaagagaata 2760

gaagagggtta ttaaagaact gggcagccag atcttaaagg agcatcctgt ggaaaatacc 2820

caattgcaga acgagaaact ttacctat tacctacaaa atggaaggga catgtatgtt 2880

gatcaggaac tggacataaa ccgtttatct gattacgacg tcgatgcat tgtaccccaa 2940

tcctttttga aggacgattc aatcgacaat aaagtgccta cacgtcggga taagaaccga 3000

gggaaaagtg acaatgttcc aagcgaggaa gtcgtaaaga aaatgaagaa ctattggcgg 3060

cagctcctaa atgcgaaact gataacgcaa agaaagtctg ataacttaac taaagctgag 3120
 aggggtggct tgtctgaact tgacaaggcc ggatttatta aacgtcagct cgtggaaacc 3180
 cgccaaatca caaagcatgt tgcacagata ctagattccc gaatgaatac gaaatacgac 3240
 gagaacgata agctgattcg ggaagtcaaa gtaatcactt taaagtcaaa attgggtgctg 3300
 gacttcagaa aggattttca attctataaa gttagggaga taaataacta ccacatgcg 3360
 cacgacgctt atcttaatgc cgtcgtaggg accgcactca ttaagaaata cccgaagcta 3420
 gaaagtgagt ttgtgtatgg tgattacaaa gtttatgacg tccgtaagat gatcgcgaaa 3480

 agcgaacagg agataggcaa ggctacagcc aaatacttct tttattctaa cattatgaat 3540
 ttctttaaga cggaatcac tctggcaaac ggagagatac gcaaacgacc ttaattgaa 3600
 accaatgggg agacaggatga aatcgtatgg gataagggcc gggacttcgc gacggtgaga 3660
 aaagttttgt ccatgcccc agtcaacata gtaagaaaaa ctgagggtgca gaccggaggg 3720
 ttttcaaagg aatcgattct tccaaaaagg aatagtata agctcatcgc tcgtaaaaag 3780
 gactgggacc cgaaaaagta cggtaggcttc gatagcccta cagtgccta ttctgtccta 3840
 gtagtggcaa aagttgagaa gggaaaatcc aagaaactga agtcagtcaa agaattattg 3900

 gggataacga ttatggagcg ctctctttt gaaaagaacc ccatcgactt ccttgaggcg 3960
 aaaggttaca aggaagtaaa aaaggatctc ataattaaac taccaaagta tagtctgttt 4020
 gagttagaaa atggccgaaa acggatgttg gctagcgccg gagagcttca aaagggaac 4080
 gaactcgac tacctctaa atacgtgaat ttctgtatt tagcgtcca ttacgagaag 4140
 ttgaaaggtt cacctgaaga taacgaacag aagcaacttt ttgttgagca gcacaaacat 4200
 tatctcgacg aaatcataga gcaaatctcg gaattcagta agagagtcac cctagctgat 4260
 gccaatctgg acaaagtatt aagcgcatc aacaagcaca gggataaacc catagtgag 4320

 caggcggaaa atattatcca ttgttttact cttaccaacc tcggcgctcc agccgcattc 4380
 aagtattttg acacaacgat agatcgcaaa cgatacactt ctaccaagga ggtgctagac 4440
 gcgacactga ttaccaatc catcacggga ttatatgaaa ctcgataga ttgtcacag 4500
 ctgggggtg acggatcccc caagaagaag aggaaagtct cgagcgacta caaagacat 4560
 gacggtgatt ataaagatca tgacatgat tacaaggatg acgatgaca ggctgcagga 4620
 ggcggtgaa gggggcgcg cgacgcgtg gacgatttcg atctcgacat gctgggttct 4680
 gatgccctcg atgactttga cctggatatg ttgggaagcg acgattgga tgactttgat 4740

 ctggacatgc tcggctccga tgctctggac gatttcgac tcgatatgtt ataaccggtc 4800
 atcatcacca tcaccattga gtttaaaccc gctgatcagc ctgactgtg cttctagtt 4860
 gccagccatc tgttgtttgc cctcccccg tgccttcctt gacctggaa ggtgccactc 4920

ccactgtcct ttctaataa aatgaggaaa ttgcatcgca ttgtctgagt aggtgtcatt	4980
ctattctggg ggggtgggtg gggcaggaca gcaaggggga ggattgggaa gacaatagca	5040
ggcatgctgg ggatgcggtg ggctctatgg cttctgaggc ggaaagaacc agctggggct	5100
cgataccgtc gacctctagc tagagcttgg cgtaatcatg gtcatactgt tttcctgtgt	5160
gaaattgtta tccgtcaca attccacaca acatacgagc cggaagcata aagtgtaaag	5220
cctaggggtgc ctaatgagtg agctaactca cattaattgc gttgcgtca ctgcccgctt	5280
tccagtcggg aaacctgtcg tgccagctgc attaatgaat cgccaacgc gcggggagag	5340
gcggtttgcg tattgggcgc ttttccgctt cctcgctcac tgactcgtg cgctcggtcg	5400
ttcggtgctg gcgagcggtg tcagctcact caaaggcggg aatacggtta tccacagaat	5460
caggggataa cgcaggaaag aacatgtgag caaaaggcca gcaaaaggcc aggaaccgta	5520
aaaaggccgc gttgctggcg tttttccata ggctccgccc cctgacgag catcacaaaa	5580
atcgacgtc aagtcagagg tggcgaaacc cgacaggact ataaagatac caggcgtttc	5640
cccttgaag ctccctcgtg cgctctctg ttcgacctt gccgttacc ggatacctgt	5700
ccgcttttct ccttcggga agcgtggcgc tttctcaatg ctacgctgt aggtatctca	5760
gttcggtgta ggtcgttcgc tccaagctgg gctgtgtgca cgaaccccc gttcagcccc	5820
accgtgcgc cttatccgtt aactatctc ttgagtccaa cccgtaaga cagacttat	5880
cgccactggc agcagccact ggtaacagga ttagcagagc gaggtatgta ggcggtgcta	5940
cagagttctt gaagtgggtg cctaactacg gctacactag aaggacagta tttggtatct	6000
gcgtctgct gaagccagtt accttcggaa aaagagtgg tagctcttga tccggcaaac	6060
aaaccaccgc tggtagcgtt ggtttttttg ttgcaagca gcagattacg cgcagaaaaa	6120
aaggatctca agaagatcct ttgatctttt ctacggggtc tgacgctcag tggaacgaaa	6180
actcagttta agggattttg gtcatgagat tatcaaaaag gatcttcacc tagatccttt	6240
taaattaaaa atgaagtttt aaatcaatct aaagtatata tgagttaaact tggcttgaca	6300
gttaccaatg cttaatcagt gaggcaccta tctcagcgat ctgtctattt cgttcatcca	6360
tagttgcctg actccccgtc gtgtagataa ctacgatacg ggagggtta ccatctggcc	6420
ccagtgtgc aatgataccg cgagaccac gtcaccggc tccagattta tcagcaataa	6480
accagccagc cggaagggcc gagcgagaa gtggtcctgc aactttatcc gcctccatcc	6540
agtcatttaa ttgttgccgg gaagctagag taagtagttc gccagttaat agtttgcgca	6600
acgttgttgc cattgtaca ggcatcgtgg tgtcacgtc gtcgtttggt atggcttcat	6660
tcagctccgg ttcccaacga tcaaggcgag ttacatgac cccatgttg tgcaaaaaag	6720
cggttagctc cticggctct ccgatcgttg tcagaagtaa gttggccgca gtgttatcac	6780

tcatggttat ggcagactg cataattctc ttactgtcat gccatccgta agatgctttt 6840

ctgtgactgg tgagtactca accaagtcac tctgagaata gtgtatgcgg cgaccgagtt 6900

gctcttgccc ggcgtcaata cgggataata ccgcgccaca tagcagaact ttaaaagtgc 6960

tcatcattgg aaaacgttct tcggggcgaa aactctcaag gatcttaccg ctgttgagat 7020

ccagttcgat gtaaccact cgtgcacca actgatcttc agcatctttt actttcacca 7080

gcgtttctgg gtgagcaaaa acaggaaggc aaaatgccgc aaaaaaggga ataagggcga 7140

cacggaaatg ttgaatactc atactcttcc tttttcaata ttattgaagc atttatcagg 7200

gttattgtct catgagcgga tacatatttg aatgtattta gaaaaataaa caaatagggg 7260

ttccgcgcac atttccccga aaagtgccac ctgacgtcga cggatcggga gatcgatctc 7320

ccgatccctt agggctgact ctcatgacaa tctgctctga tgccgcatag ttaagccagt 7380

atctgtctcc tgcctgtgtg ttggaggctg ctgagtagtg cgcgagcaaa atttaagcta 7440

caacaaggca aggcctgacc gacaattgca tgaagaatct gcttaggggtt aggcgttttg 7500

cgctgcttcg cgatgtacgg gccagatata cgcgttgaca ttgattattg actagttatt 7560

aatagtaate aattacgggg tcattagttc atagcccata tatggagtgc cgcgttacat 7620

aacttacggt aaatggcccg cctggctgac cgcccaacga ccccgccca ttgacgtcaa 7680

taatgacgta tgttcccata gtaacccaa tagggacttt ccattgacgt caatgggtgg 7740

actatttacg gtaaaactgcc cacttggcag tacatcaagt gtatc 7785

<210> 110

<211> 1461

<212> PRT

<213> Artificial Sequence

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

Cas9-activator polypeptide

<400> 110

Met Asp Lys Lys Tyr Ser Ile Gly Leu Ala Ile Gly Thr Asn Ser Val

1 5 10 15

Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe

20 25 30

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

35 40 45

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

50 55 60
 Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys
 65 70 75 80
 Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser

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

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

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

 275 280 285
 Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp
 290 295 300

Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser
305 310 315 320
Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys
325 330 335
Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe
340 345 350
Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser
355 360 365
Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp
370 375 380
Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg
385 390 395 400
Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu
405 410 415
Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe
420 425 430
Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile
435 440 445
Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp
450 455 460
Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu
465 470 475 480
Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr
485 490 495
Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser
500 505 510
Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys
515 520 525
Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln
530 535 540
Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr

545 550 555 560
 Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp
 565 570 575
 Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly
 580 585 590
 Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp

 595 600 605
 Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr
 610 615 620
 Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala
 625 630 635 640
 His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr
 645 650 655
 Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp

 660 665 670
 Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe
 675 680 685
 Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe
 690 695 700
 Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu
 705 710 715 720
 His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly

 725 730 735
 Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly
 740 745 750
 Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln
 755 760 765
 Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile
 770 775 780
 Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro

 785 790 795 800

Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu
805 810 815
Gln Asn Gly Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg
820 825 830
Leu Ser Asp Tyr Asp Val Asp Ala Ile Val Pro Gln Ser Phe Leu Lys
835 840 845
Asp Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg
850 855 860
Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys
865 870 875 880
Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys
885 890 895
Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp
900 905 910
Lys Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr
915 920 925
Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp
930 935 940
Glu Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser
945 950 955 960
Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg
965 970 975
Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val
980 985 990
Val Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe
995 1000 1005
Val Tyr Gly Asp Tyr Lys Val Tyr Asp Val Arg Lys Met Ile Ala
1010 1015 1020
Lys Ser Glu Gln Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe
1025 1030 1035
Tyr Ser Asn Ile Met Asn Phe Phe Lys Thr Glu Ile Thr Leu Ala

1040	1045	1050
Asn Gly Glu Ile Arg Lys Arg	Pro Leu Ile Glu Thr	Asn Gly Glu
1055	1060	1065
Thr Gly Glu Ile Val Trp Asp	Lys Gly Arg Asp Phe	Ala Thr Val
1070	1075	1080
Arg Lys Val Leu Ser Met Pro	Gln Val Asn Ile Val	Lys Lys Thr
1085	1090	1095
Glu Val Gln Thr Gly Gly Phe	Ser Lys Glu Ser Ile	Leu Pro Lys

1100	1105	1110
Arg Asn Ser Asp Lys Leu Ile	Ala Arg Lys Lys Asp	Trp Asp Pro
1115	1120	1125
Lys Lys Tyr Gly Gly Phe Asp	Ser Pro Thr Val Ala	Tyr Ser Val
1130	1135	1140
Leu Val Val Ala Lys Val Glu	Lys Gly Lys Ser Lys	Lys Leu Lys
1145	1150	1155
Ser Val Lys Glu Leu Leu Gly	Ile Thr Ile Met Glu	Arg Ser Ser

1160	1165	1170
Phe Glu Lys Asn Pro Ile Asp	Phe Leu Glu Ala Lys	Gly Tyr Lys
1175	1180	1185
Glu Val Lys Lys Asp Leu Ile	Ile Lys Leu Pro Lys	Tyr Ser Leu
1190	1195	1200
Phe Glu Leu Glu Asn Gly Arg	Lys Arg Met Leu Ala	Ser Ala Gly
1205	1210	1215
Glu Leu Gln Lys Gly Asn Glu	Leu Ala Leu Pro Ser	Lys Tyr Val

1220	1225	1230
Asn Phe Leu Tyr Leu Ala Ser	His Tyr Glu Lys Leu	Lys Gly Ser
1235	1240	1245
Pro Glu Asp Asn Glu Gln Lys	Gln Leu Phe Val Glu	Gln His Lys
1250	1255	1260
His Tyr Leu Asp Glu Ile Ile	Glu Gln Ile Ser Glu	Phe Ser Lys
1265	1270	1275

Arg Val Ile Leu Ala Asp Ala Asn Leu Asp Lys Val Leu Ser Ala

1280 1285 1290

Tyr Asn Lys His Arg Asp Lys Pro Ile Arg Glu Gln Ala Glu Asn

1295 1300 1305

Ile Ile His Leu Phe Thr Leu Thr Asn Leu Gly Ala Pro Ala Ala

1310 1315 1320

Phe Lys Tyr Phe Asp Thr Thr Ile Asp Arg Lys Arg Tyr Thr Ser

1325 1330 1335

Thr Lys Glu Val Leu Asp Ala Thr Leu Ile His Gln Ser Ile Thr

1340 1345 1350

Gly Leu Tyr Glu Thr Arg Ile Asp Leu Ser Gln Leu Gly Gly Asp

1355 1360 1365

Gly Ser Pro Lys Lys Lys Arg Lys Val Ser Ser Asp Tyr Lys Asp

1370 1375 1380

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

1385 1390 1395

Asp Asp Lys Ala Ala Gly Gly Gly Gly Ser Gly Arg Ala Asp Ala

1400 1405 1410

Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp

1415 1420 1425

Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe

1430 1435 1440

Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu

1445 1450 1455

Asp Met Leu

1460

<210> 111

<211> 1527

<212> PRT

<213> Artificial Sequence

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

dCas9-NLS-3xFLAG-HP1alpha polypeptide

<400> 111

Met Asp Lys Lys Tyr Ser Ile Gly Leu Ala Ile Gly Thr Asn Ser Val

1 5 10 15

Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe

20 25 30

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

35 40 45

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

50 55 60

Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys

65 70 75 80

Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser

85 90 95

Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys

100 105 110

His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr

115 120 125

His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp

130 135 140

Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His

145 150 155 160

Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro

165 170 175

Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr

180 185 190

Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala

195 200 205

Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn

210 215 220

Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn

225 230 235 240
 Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe
 245 250 255
 Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp
 260 265 270
 Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp
 275 280 285
 Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp
 290 295 300

 Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser
 305 310 315 320
 Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys
 325 330 335
 Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe
 340 345 350
 Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser
 355 360 365

 Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp
 370 375 380
 Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg
 385 390 395 400
 Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu
 405 410 415
 Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe
 420 425 430

 Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile
 435 440 445
 Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp
 450 455 460
 Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu
 465 470 475 480

Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr
485 490 495

Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser
500 505 510

Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys
515 520 525

Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln
530 535 540

Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr
545 550 555 560

Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp
565 570 575

Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly
580 585 590

Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp
595 600 605

Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr
610 615 620

Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala
625 630 635 640

His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr
645 650 655

Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp
660 665 670

Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe
675 680 685

Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe
690 695 700

Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu
705 710 715 720

His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly

	725	730	735
Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly			
	740	745	750
Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln			
	755	760	765
Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile			
	770	775	780
Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro			
785	790	795	800
Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu			
	805	810	815
Gln Asn Gly Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg			
	820	825	830
Leu Ser Asp Tyr Asp Val Asp Ala Ile Val Pro Gln Ser Phe Leu Lys			
	835	840	845
Asp Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg			
	850	855	860
Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys			
865	870	875	880
Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys			
	885	890	895
Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp			
	900	905	910
Lys Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr			
	915	920	925
Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp			
	930	935	940
Glu Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser			
945	950	955	960
Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg			
	965	970	975

Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val			
980	985	990	
Val Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe			
995	1000	1005	
Val Tyr Gly Asp Tyr Lys Val Tyr Asp Val Arg Lys Met Ile Ala			
1010	1015	1020	
Lys Ser Glu Gln Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe			
1025	1030	1035	
Tyr Ser Asn Ile Met Asn Phe Phe Lys Thr Glu Ile Thr Leu Ala			
1040	1045	1050	
Asn Gly Glu Ile Arg Lys Arg Pro Leu Ile Glu Thr Asn Gly Glu			
1055	1060	1065	
Thr Gly Glu Ile Val Trp Asp Lys Gly Arg Asp Phe Ala Thr Val			
1070	1075	1080	
Arg Lys Val Leu Ser Met Pro Gln Val Asn Ile Val Lys Lys Thr			
1085	1090	1095	
Glu Val Gln Thr Gly Gly Phe Ser Lys Glu Ser Ile Leu Pro Lys			
1100	1105	1110	
Arg Asn Ser Asp Lys Leu Ile Ala Arg Lys Lys Asp Trp Asp Pro			
1115	1120	1125	
Lys Lys Tyr Gly Gly Phe Asp Ser Pro Thr Val Ala Tyr Ser Val			
1130	1135	1140	
Leu Val Val Ala Lys Val Glu Lys Gly Lys Ser Lys Lys Leu Lys			
1145	1150	1155	
Ser Val Lys Glu Leu Leu Gly Ile Thr Ile Met Glu Arg Ser Ser			
1160	1165	1170	
Phe Glu Lys Asn Pro Ile Asp Phe Leu Glu Ala Lys Gly Tyr Lys			
1175	1180	1185	
Glu Val Lys Lys Asp Leu Ile Ile Lys Leu Pro Lys Tyr Ser Leu			
1190	1195	1200	
Phe Glu Leu Glu Asn Gly Arg Lys Arg Met Leu Ala Ser Ala Gly			

1205	1210	1215
Glu Leu Gln Lys Gly Asn Glu	Leu Ala Leu Pro Ser	Lys Tyr Val
1220	1225	1230
Asn Phe Leu Tyr Leu Ala Ser	His Tyr Glu Lys Leu	Lys Gly Ser
1235	1240	1245
Pro Glu Asp Asn Glu Gln Lys	Gln Leu Phe Val Glu	Gln His Lys
1250	1255	1260
His Tyr Leu Asp Glu Ile Ile	Glu Gln Ile Ser Glu	Phe Ser Lys
1265	1270	1275
Arg Val Ile Leu Ala Asp Ala	Asn Leu Asp Lys Val	Leu Ser Ala
1280	1285	1290
Tyr Asn Lys His Arg Asp Lys	Pro Ile Arg Glu Gln	Ala Glu Asn
1295	1300	1305
Ile Ile His Leu Phe Thr Leu	Thr Asn Leu Gly Ala	Pro Ala Ala
1310	1315	1320
Phe Lys Tyr Phe Asp Thr Thr	Ile Asp Arg Lys Arg	Tyr Thr Ser
1325	1330	1335
Thr Lys Glu Val Leu Asp Ala	Thr Leu Ile His Gln	Ser Ile Thr
1340	1345	1350
Gly Leu Tyr Glu Thr Arg Ile	Asp Leu Ser Gln Leu	Gly Gly Asp
1355	1360	1365
Gly Ser Pro Lys Lys Lys Arg	Lys Val Ser Ser Asp	Tyr Lys Asp
1370	1375	1380
His Asp Gly Asp Tyr Lys Asp	His Asp Ile Asp Tyr	Lys Asp Asp
1385	1390	1395
Asp Asp Lys Ala Ala Gly Gly	Gly Gly Ser Met Lys	Glu Gly Glu
1400	1405	1410
Asn Asn Lys Pro Arg Glu Lys	Ser Glu Ser Asn Lys	Arg Lys Ser
1415	1420	1425
Asn Phe Ser Asn Ser Ala Asp	Asp Ile Lys Ser Lys	Lys Lys Arg
1430	1435	1440

Glu Gln Ser Asn Asp Ile Ala Arg Gly Phe Glu Arg Gly Leu Glu
 1445 1450 1455
 Pro Glu Lys Ile Ile Gly Ala Thr Asp Ser Cys Gly Asp Leu Met
 1460 1465 1470
 Phe Leu Met Lys Trp Lys Asp Thr Asp Glu Ala Asp Leu Val Leu
 1475 1480 1485

Ala Lys Glu Ala Asn Val Lys Cys Pro Gln Ile Val Ile Ala Phe
 1490 1495 1500
 Tyr Glu Glu Arg Leu Thr Trp His Ala Tyr Pro Glu Asp Ala Glu
 1505 1510 1515
 Asn Lys Glu Lys Glu Thr Ala Lys Ser
 1520 1525

<210> 112

<211> 1521

<212> PRT

<213> Artificial Sequence

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

dCas9-NLS-3xFLAG-HP1beta polypeptide

<

400> 112

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

Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys
 65 70 75 80
 Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser
 85 90 95
 Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys

100	105	110	
His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr			
115	120	125	
His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp			
130	135	140	
Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His			
145	150	155	160
Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro			
165	170	175	
Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr			
180	185	190	
Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala			
195	200	205	
Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn			
210	215	220	
Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn			
225	230	235	240
Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe			
245	250	255	
Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp			
260	265	270	
Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp			
275	280	285	
Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp			
290	295	300	
Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser			
305	310	315	320
Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys			
325	330	335	
Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe			
340	345	350	

Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser
 355 360 365
 Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp
 370 375 380

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

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

 Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys
 515 520 525
 Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln
 530 535 540
 Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr
 545 550 555 560
 Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp
 565 570 575

 Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly
 580 585 590
 Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp

595	600	605	
Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr			
610	615	620	
Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala			
625	630	635	640
His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr			
645	650	655	
Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp			
660	665	670	
Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe			
675	680	685	
Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe			
690	695	700	
Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu			
705	710	715	720
His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly			
725	730	735	
Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly			
740	745	750	
Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln			
755	760	765	
Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile			
770	775	780	
Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro			
785	790	795	800
Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu			
805	810	815	
Gln Asn Gly Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg			
820	825	830	
Leu Ser Asp Tyr Asp Val Asp Ala Ile Val Pro Gln Ser Phe Leu Lys			
835	840	845	

Asp Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg

850 855 860

Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys

865 870 875 880

Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys

885 890 895

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

900 905 910

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

915 920 925

Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp

930 935 940

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

945 950 955 960

Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg

965 970 975

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

980 985 990

Val Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe

995 1000 1005

Val Tyr Gly Asp Tyr Lys Val Tyr Asp Val Arg Lys Met Ile Ala

1010 1015 1020

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

1025 1030 1035

Tyr Ser Asn Ile Met Asn Phe Phe Lys Thr Glu Ile Thr Leu Ala

1040 1045 1050

Asn Gly Glu Ile Arg Lys Arg Pro Leu Ile Glu Thr Asn Gly Glu

1055 1060 1065

Thr Gly Glu Ile Val Trp Asp Lys Gly Arg Asp Phe Ala Thr Val

1070 1075 1080

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

1085	1090	1095
Glu Val Gln Thr Gly Gly Phe	Ser Lys Glu Ser Ile	Leu Pro Lys
1100	1105	1110
Arg Asn Ser Asp Lys Leu Ile	Ala Arg Lys Lys Asp	Trp Asp Pro
1115	1120	1125
Lys Lys Tyr Gly Gly Phe Asp	Ser Pro Thr Val Ala	Tyr Ser Val
1130	1135	1140
Leu Val Val Ala Lys Val Glu	Lys Gly Lys Ser Lys	Lys Leu Lys
1145	1150	1155
Ser Val Lys Glu Leu Leu Gly	Ile Thr Ile Met Glu	Arg Ser Ser
1160	1165	1170
Phe Glu Lys Asn Pro Ile Asp	Phe Leu Glu Ala Lys	Gly Tyr Lys
1175	1180	1185
Glu Val Lys Lys Asp Leu Ile	Ile Lys Leu Pro Lys	Tyr Ser Leu
1190	1195	1200
Phe Glu Leu Glu Asn Gly Arg	Lys Arg Met Leu Ala	Ser Ala Gly
1205	1210	1215
Glu Leu Gln Lys Gly Asn Glu	Leu Ala Leu Pro Ser	Lys Tyr Val
1220	1225	1230
Asn Phe Leu Tyr Leu Ala Ser	His Tyr Glu Lys Leu	Lys Gly Ser
1235	1240	1245
Pro Glu Asp Asn Glu Gln Lys	Gln Leu Phe Val Glu	Gln His Lys
1250	1255	1260
His Tyr Leu Asp Glu Ile Ile	Glu Gln Ile Ser Glu	Phe Ser Lys
1265	1270	1275
Arg Val Ile Leu Ala Asp Ala	Asn Leu Asp Lys Val	Leu Ser Ala
1280	1285	1290
Tyr Asn Lys His Arg Asp Lys	Pro Ile Arg Glu Gln	Ala Glu Asn
1295	1300	1305
Ile Ile His Leu Phe Thr Leu	Thr Asn Leu Gly Ala	Pro Ala Ala
1310	1315	1320

Phe Lys	Tyr Phe Asp Thr Thr	Ile Asp Arg Lys Arg	Tyr Thr Ser
1325	1330	1335	
Thr Lys	Glu Val Leu Asp Ala	Thr Leu Ile His Gln	Ser Ile Thr
1340	1345	1350	
Gly Leu	Tyr Glu Thr Arg Ile	Asp Leu Ser Gln Leu	Gly Gly Asp
1355	1360	1365	
Gly Ser	Pro Lys Lys Lys Arg	Lys Val Ser Ser Asp	Tyr Lys Asp
1370	1375	1380	

His Asp	Gly Asp Tyr Lys Asp	His Asp Ile Asp Tyr	Lys Asp Asp
1385	1390	1395	
Asp Asp	Lys Ala Ala Gly Gly	Gly Gly Ser Thr Ala	His Glu Thr
1400	1405	1410	
Asp Lys	Ser Glu Gly Gly Lys	Arg Lys Ala Asp Ser	Asp Ser Glu
1415	1420	1425	
Asp Lys	Gly Glu Glu Ser Lys	Pro Lys Lys Lys Lys	Glu Glu Ser
1430	1435	1440	

Glu Lys	Pro Arg Gly Phe Ala	Arg Gly Leu Glu Pro	Glu Arg Ile
1445	1450	1455	
Ile Gly	Ala Thr Asp Ser Ser	Gly Glu Leu Met Phe	Leu Met Lys
1460	1465	1470	
Trp Lys	Asn Ser Asp Glu Ala	Asp Leu Val Pro Ala	Lys Glu Ala
1475	1480	1485	
Asn Val	Lys Cys Pro Gln Val	Val Ile Ser Phe Tyr	Glu Glu Arg
1490	1495	1500	

Leu Thr	Trp His Ser Tyr Pro	Ser Glu Asp Asp Asp	Lys Lys Asp
1505	1510	1515	
Asp Lys	Asn		
1520			

<210> 113

<211> 2126

<212> PRT

<213> Artificial Sequence

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

dCas9-3xFLAG-TET1CD polypeptide

<400> 113

Met Asp Lys Lys Tyr Ser Ile Gly Leu Ala Ile Gly Thr Asn Ser Val

1 5 10 15

Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe

20 25 30

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

35 40 45

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

50 55 60

Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys

65 70 75 80

Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser

85 90 95

Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys

100 105 110

His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr

115 120 125

His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp

130 135 140

Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His

145 150 155 160

Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro

165 170 175

Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr

180 185 190

Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala

195 200 205

Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn

210 215 220

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

 275 280 285
 Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp
 290 295 300
 Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser
 305 310 315 320
 Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys
 325 330 335
 Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe

 340 345 350
 Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser
 355 360 365
 Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp
 370 375 380
 Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg
 385 390 395 400
 Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu

 405 410 415
 Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe
 420 425 430
 Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile
 435 440 445
 Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp
 450 455 460
 Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu

465 470 475 480
 Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr
 485 490 495
 Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser
 500 505 510
 Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys
 515 520 525
 Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln

 530 535 540
 Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr
 545 550 555 560
 Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp
 565 570 575
 Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly
 580 585 590
 Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp

 595 600 605
 Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr
 610 615 620
 Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala
 625 630 635 640
 His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr
 645 650 655
 Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp

 660 665 670
 Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe
 675 680 685
 Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe
 690 695 700
 Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu
 705 710 715 720

His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly

725 730 735

Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly

740 745 750

Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln

755 760 765

Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile

770 775 780

Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro

785 790 795 800

Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu

805 810 815

Gln Asn Gly Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg

820 825 830

Leu Ser Asp Tyr Asp Val Asp Ala Ile Val Pro Gln Ser Phe Leu Lys

835 840 845

Asp Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg

850 855 860

Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys

865 870 875 880

Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys

885 890 895

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

900 905 910

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

915 920 925

Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp

930 935 940

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

945 950 955 960

Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg

965	970	975
Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val		
980	985	990
Val Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe		
995	1000	1005
Val Tyr Gly Asp Tyr Lys Val Tyr Asp Val Arg Lys Met Ile Ala		
1010	1015	1020
Lys Ser Glu Gln Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe		
1025	1030	1035
Tyr Ser Asn Ile Met Asn Phe Phe Lys Thr Glu Ile Thr Leu Ala		
1040	1045	1050
Asn Gly Glu Ile Arg Lys Arg Pro Leu Ile Glu Thr Asn Gly Glu		
1055	1060	1065
Thr Gly Glu Ile Val Trp Asp Lys Gly Arg Asp Phe Ala Thr Val		
1070	1075	1080
Arg Lys Val Leu Ser Met Pro Gln Val Asn Ile Val Lys Lys Thr		
1085	1090	1095
Glu Val Gln Thr Gly Gly Phe Ser Lys Glu Ser Ile Leu Pro Lys		
1100	1105	1110
Arg Asn Ser Asp Lys Leu Ile Ala Arg Lys Lys Asp Trp Asp Pro		
1115	1120	1125
Lys Lys Tyr Gly Gly Phe Asp Ser Pro Thr Val Ala Tyr Ser Val		
1130	1135	1140
Leu Val Val Ala Lys Val Glu Lys Gly Lys Ser Lys Lys Leu Lys		
1145	1150	1155
Ser Val Lys Glu Leu Leu Gly Ile Thr Ile Met Glu Arg Ser Ser		
1160	1165	1170
Phe Glu Lys Asn Pro Ile Asp Phe Leu Glu Ala Lys Gly Tyr Lys		
1175	1180	1185
Glu Val Lys Lys Asp Leu Ile Ile Lys Leu Pro Lys Tyr Ser Leu		
1190	1195	1200

Phe Glu	Leu Glu Asn Gly Arg	Lys Arg Met Leu Ala	Ser Ala Gly
1205	1210	1215	
Glu Leu	Gln Lys Gly Asn Glu	Leu Ala Leu Pro Ser	Lys Tyr Val
1220	1225	1230	
Asn Phe	Leu Tyr Leu Ala Ser	His Tyr Glu Lys Leu	Lys Gly Ser
1235	1240	1245	
Pro Glu	Asp Asn Glu Gln Lys	Gln Leu Phe Val Glu	Gln His Lys
1250	1255	1260	
His Tyr	Leu Asp Glu Ile Ile	Glu Gln Ile Ser Glu	Phe Ser Lys
1265	1270	1275	
Arg Val	Ile Leu Ala Asp Ala	Asn Leu Asp Lys Val	Leu Ser Ala
1280	1285	1290	
Tyr Asn	Lys His Arg Asp Lys	Pro Ile Arg Glu Gln	Ala Glu Asn
1295	1300	1305	
Ile Ile	His Leu Phe Thr Leu	Thr Asn Leu Gly Ala	Pro Ala Ala
1310	1315	1320	
Phe Lys	Tyr Phe Asp Thr Thr	Ile Asp Arg Lys Arg	Tyr Thr Ser
1325	1330	1335	
Thr Lys	Glu Val Leu Asp Ala	Thr Leu Ile His Gln	Ser Ile Thr
1340	1345	1350	
Gly Leu	Tyr Glu Thr Arg Ile	Asp Leu Ser Gln Leu	Gly Gly Asp
1355	1360	1365	
Gly Ser	Pro Lys Lys Lys Arg	Lys Val Ser Ser Asp	Tyr Lys Asp
1370	1375	1380	
His Asp	Gly Asp Tyr Lys Asp	His Asp Ile Asp Tyr	Lys Asp Asp
1385	1390	1395	
Asp Asp	Lys Ala Ala Gly Gly	Gly Gly Ser Leu Pro	Thr Cys Ser
1400	1405	1410	
Cys Leu	Asp Arg Val Ile Gln	Lys Asp Lys Gly Pro	Tyr Tyr Thr
1415	1420	1425	
His Leu	Gly Ala Gly Pro Ser	Val Ala Ala Val Arg	Glu Ile Met

1430	1435	1440
Glu Asn Arg Tyr Gly Gln Lys	Gly Asn Ala Ile Arg	Ile Glu Ile
1445	1450	1455
Val Val Tyr Thr Gly Lys Glu	Gly Lys Ser Ser His	Gly Cys Pro
1460	1465	1470
Ile Ala Lys Trp Val Leu Arg	Arg Ser Ser Asp Glu	Glu Lys Val
1475	1480	1485
Leu Cys Leu Val Arg Gln Arg	Thr Gly His His Cys	Pro Thr Ala
1490	1495	1500
Val Met Val Val Leu Ile Met	Val Trp Asp Gly Ile	Pro Leu Pro
1505	1510	1515
Met Ala Asp Arg Leu Tyr Thr	Glu Leu Thr Glu Asn	Leu Lys Ser
1520	1525	1530
Tyr Asn Gly His Pro Thr Asp	Arg Arg Cys Thr Leu	Asn Glu Asn
1535	1540	1545
Arg Thr Cys Thr Cys Gln Gly	Ile Asp Pro Glu Thr	Cys Gly Ala
1550	1555	1560
Ser Phe Ser Phe Gly Cys Ser	Trp Ser Met Tyr Phe	Asn Gly Cys
1565	1570	1575
Lys Phe Gly Arg Ser Pro Ser	Pro Arg Arg Phe Arg	Ile Asp Pro
1580	1585	1590
Ser Ser Pro Leu His Glu Lys	Asn Leu Glu Asp Asn	Leu Gln Ser
1595	1600	1605
Leu Ala Thr Arg Leu Ala Pro	Ile Tyr Lys Gln Tyr	Ala Pro Val
1610	1615	1620
Ala Tyr Gln Asn Gln Val Glu	Tyr Glu Asn Val Ala	Arg Glu Cys
1625	1630	1635
Arg Leu Gly Ser Lys Glu Gly	Arg Pro Phe Ser Gly	Val Thr Ala
1640	1645	1650
Cys Leu Asp Phe Cys Ala His	Pro His Arg Asp Ile	His Asn Met
1655	1660	1665

Asn Asn Gly Ser Thr Val Val Cys Thr Leu Thr Arg Glu Asp Asn		
1670	1675	1680
Arg Ser Leu Gly Val Ile Pro Gln Asp Glu Gln Leu His Val Leu		
1685	1690	1695
Pro Leu Tyr Lys Leu Ser Asp Thr Asp Glu Phe Gly Ser Lys Glu		
1700	1705	1710
Gly Met Glu Ala Lys Ile Lys Ser Gly Ala Ile Glu Val Leu Ala		
1715	1720	1725
Pro Arg Arg Lys Lys Arg Thr Cys Phe Thr Gln Pro Val Pro Arg		
1730	1735	1740
Ser Gly Lys Lys Arg Ala Ala Met Met Thr Glu Val Leu Ala His		
1745	1750	1755
Lys Ile Arg Ala Val Glu Lys Lys Pro Ile Pro Arg Ile Lys Arg		
1760	1765	1770
Lys Asn Asn Ser Thr Thr Thr Asn Asn Ser Lys Pro Ser Ser Leu		
1775	1780	1785
Pro Thr Leu Gly Ser Asn Thr Glu Thr Val Gln Pro Glu Val Lys		
1790	1795	1800
Ser Glu Thr Glu Pro His Phe Ile Leu Lys Ser Ser Asp Asn Thr		
1805	1810	1815
Lys Thr Tyr Ser Leu Met Pro Ser Ala Pro His Pro Val Lys Glu		
1820	1825	1830
Ala Ser Pro Gly Phe Ser Trp Ser Pro Lys Thr Ala Ser Ala Thr		
1835	1840	1845
Pro Ala Pro Leu Lys Asn Asp Ala Thr Ala Ser Cys Gly Phe Ser		
1850	1855	1860
Glu Arg Ser Ser Thr Pro His Cys Thr Met Pro Ser Gly Arg Leu		
1865	1870	1875
Ser Gly Ala Asn Ala Ala Ala Ala Asp Gly Pro Gly Ile Ser Gln		
1880	1885	1890
Leu Gly Glu Val Ala Pro Leu Pro Thr Leu Ser Ala Pro Val Met		

1895	1900	1905
Glu Pro Leu Ile Asn Ser	Glu Pro Ser Thr Gly Val	Thr Glu Pro
1910	1915	1920
Leu Thr Pro His Gln Pro Asn	His Gln Pro Ser Phe	Leu Thr Ser
1925	1930	1935
Pro Gln Asp Leu Ala Ser Ser	Pro Met Glu Glu Asp	Glu Gln His
1940	1945	1950
Ser Glu Ala Asp Glu Pro Pro	Ser Asp Glu Pro Leu	Ser Asp Asp
1955	1960	1965
Pro Leu Ser Pro Ala Glu Glu	Lys Leu Pro His Ile	Asp Glu Tyr
1970	1975	1980
Trp Ser Asp Ser Glu His Ile	Phe Leu Asp Ala Asn	Ile Gly Gly
1985	1990	1995
Val Ala Ile Ala Pro Ala His	Gly Ser Val Leu Ile	Glu Cys Ala
2000	2005	2010
Arg Arg Glu Leu His Ala Thr	Thr Pro Val Glu His	Pro Asn Arg
2015	2020	2025
Asn His Pro Thr Arg Leu Ser	Leu Val Phe Tyr Gln	His Lys Asn
2030	2035	2040
Leu Asn Lys Pro Gln His Gly	Phe Glu Leu Asn Lys	Ile Lys Phe
2045	2050	2055
Glu Ala Lys Glu Ala Lys Asn	Lys Lys Met Lys Ala	Ser Glu Gln
2060	2065	2070
Lys Asp Gln Ala Ala Asn Glu	Gly Pro Glu Gln Ser	Ser Glu Val
2075	2080	2085
Asn Glu Leu Asn Gln Ile Pro	Ser His Lys Ala Leu	Thr Leu Thr
2090	2095	2100
His Asp Asn Val Val Thr Val	Ser Pro Tyr Ala Leu	Thr His Val
2105	2110	2115
Ala Gly Pro Tyr Asn His Trp	Val	
2120	2125	