(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2025-0077608
(43) 공개일자 2025년05월30일

(51) 국제특허분류(Int. Cl.)
C12N 15/113 (2010.01) *C07K 14/47* (2006.01)
C12N 15/88 (2017.01) *C12N 9/10* (2006.01)
C12N 9/22 (2006.01)
(52) CPC특허분류
C12N 15/1131 (2013.01)
C07K 14/4703 (2013.01)
(21) 출원번호 10-2025-7008867
(22) 출원일자(국제) 2023년08월18일
심사청구일자 없음
(85) 번역문제출일자 2025년03월18일
(86) 국제출원번호 PCT/US2023/072525
(87) 국제공개번호 WO 2024/040254
국제공개일자 2024년02월22일
(30) 우선권주장
63/399,634 2022년08월19일 미국(US)
(뒷면에 계속)

(71) 출원인
튠 쎄라퓨틱스, 인코포레이티드
미국 워싱턴 98101 시애틀 보렌 애비뉴 1930 플로어 5
(72) 발명자
코스그로브, 브라이언
미국 노쓰 캐롤라이나 27701 더햄 스위트 320 모리스 스트리트 300 튠 쎄라퓨틱스 인코포레이티드 내
콩든, 켄드라
미국 노쓰 캐롤라이나 27701 더햄 스위트 320 모리스 스트리트 300 튠 쎄라퓨틱스 인코포레이티드 내
(뒷면에 계속)
(74) 대리인
박장원

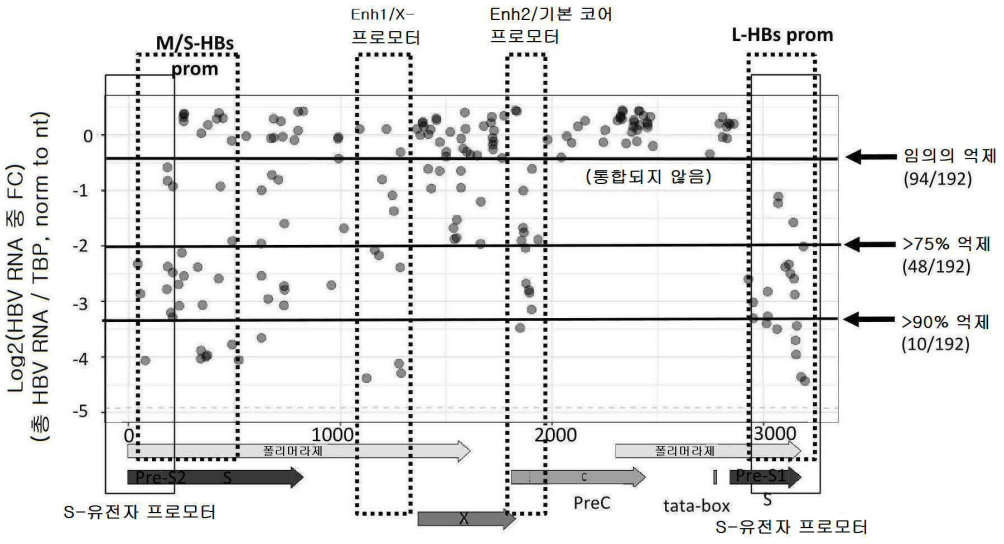
전체 청구항 수 : 총 350 항

(54) 발명의 명칭 **표적 유전자 억제를 통한 B형 간염 바이러스 조절용 조성물, 시스템 및 방법**

(57) 요 약

본 발명에 따라 세포 표현형을 촉진하여 HBV 감염 감소로 이어지는 B형 간염 바이러스(HBV) 유전자의 전사 억제를 위한 CRISPR-Cas/가이드 RNA(gRNA) 시스템과 같은 후성적 변형 DNA-표적화 시스템이 제공된다. 일부 구현예에서, 후성적 변형 DNA-표적화 시스템은 세포 내 B형 간염 바이러스 DNA 서열에서 적어도 하나의 유전자 또는 이의 조절 요소의 표적 부위에 결합하거나 표적화한다. 일부 측면에서, 제공된 시스템은 하나 이상의 B형 간염 바이러스 유전자 및/또는 그 조절 요소의 전사 억제와 관련된다. 일부 측면에서, 또한 본 발명에서는 제공된 조성물과 관련된 방법 및 용도, 예를 들어 B형 간염 감염과 관련하여 B형 간염 바이러스 복제 및 발현을 억제하는 용도가 제공된다.

대 표 도 - 도1



(52) CPC특허분류
C12N 15/88 (2013.01)
C12N 9/1007 (2013.01)
C12N 9/1007 (2013.01)
C12N 9/22 (2025.05)
C12N 9/22 (2025.05)
C12N 2310/20 (2025.05)
C12N 2310/315 (2013.01)
C12N 2310/321 (2013.01)
C12N 2310/351 (2013.01)

(72) 발명자
딘, 제이슨
미국 워싱턴 98101 시애틀 보렌 애비뉴 1930 플로어 5 튠 쎄라퓨틱스 인코포레이티드 내
고흐, 베로니카
미국 노쓰 캐롤라이나 27701 더햄 스위트 320 모리스 스트리트 300 튠 쎄라퓨틱스 인코포레이티드 내
블랙, 조슈아 비.
미국 노쓰 캐롤라이나 27701 더햄 스위트 320 모리스 스트리트 300 튠 쎄라퓨틱스 인코포레이티드 내
존스, 브리타
미국 노쓰 캐롤라이나 27701 더햄 스위트 320 모리스 스트리트 300 튠 쎄라퓨틱스 인코포레이티드 내

(30) 우선권주장
63/472,236 2023년06월09일 미국(US)
63/531,309 2023년08월07일 미국(US)

명 세 서

청구범위

청구항 1

하나 이상의 B형 간염 바이러스(HBV) 유전자의 전사를 억제하기 위한 적어도 하나의 DNA-표적화 모듈을 포함하는 후성적 변형 DNA-표적화 시스템으로서, 여기서 적어도 하나의 DNA-표적화 모듈 각각은:

(a) B형 간염 바이러스 DNA 서열의 표적 부위를 표적화하기 위한 DNA-결합 도메인; 및

(b) 적어도 하나의 전사 억제자 효과기 도메인

을 포함하는 융합 단백질을 포함하는 것인, 후성적 변형 DNA-표적화 시스템.

청구항 2

청구항 1에 있어서, 적어도 하나의 DNA-결합 도메인은 (a) Cas 단백질 또는 이의 변이체 및 (b) Cas 단백질과 결합하여 Cas 단백질을 표적 부위로 표적화할 수 있는 적어도 하나의 gRNA를 포함하는, 규칙적인 간격을 갖고 나타나는 짧은 회문 구조의 반복 서열(CRIPRE:Clustered Regularly Interspaced Short Palindromic Repeats) 관련 (Cas)-가이드(gRNA) 조합; 아연 집게 단백질(ZFP); 전사 활성화제 유사 효과기(TALE); 메가뉴클레아제; 호밍 엔도뉴클레아제; 또는 I-SceI 효소 또는 이의 변이체를 포함하며, 선택적으로 여기서 DNA-결합 도메인은 상기의 임의의 촉매적으로 비활성인 변이체를 포함하는 것인, 후성적 변형 DNA-표적화 시스템.

청구항 3

청구항 1 또는 2에 있어서, B형 간염 바이러스 DNA 서열은 HBV 유전자 또는 이의 조절 요소인, 후성적 변형 DNA-표적화 시스템.

청구항 4

청구항 1-3 중 어느 한 항에 있어서, 적어도 하나의 DNA-표적화 모듈은 하나 이상의 HBV 유전자 또는 이의 조절 요소의 복수의 표적 부위를 표적화하기 위한 복수의 DNA-표적화 모듈을 포함하는 것인, 후성적 변형 DNA-표적화 시스템.

청구항 5

청구항 4에 있어서, 복수의 DNA-표적화 모듈은 적어도 제1 DNA-표적화 모듈과 제2 DNA-표적화 모듈을 포함하고, 여기서:

(1) 제1 DNA-표적화 모듈은 제1 HBV 유전자의 전사를 억제하고, 여기서 제1 DNA-표적화 모듈은 (a) 제1 유전자 또는 이의 조절 DNA 요소의 표적 부위를 표적화하기 위한 DNA-결합 도메인; 및 (b) 적어도 하나의 전사 억제 도메인을 포함하는 제1 융합 단백질을 포함하고;

(2) 제2 DNA-표적화 모듈은 제2 HBV 유전자의 전사를 억제하고, 여기서 제2 DNA-표적화 모듈은 (a) 제2 유전자 또는 이의 조절 DNA 요소의 표적 부위를 표적화하기 위한 DNA-결합 도메인; 및 (b) 적어도 하나의 전사 억제 도메인을 포함하는 제2 융합 단백질을 포함하며, 선택적으로 여기서:

제1 DNA-표적화 모듈과 제2 DNA-표적화 모듈은 동일한 융합 단백질을 공유하여 제1 및 제2 융합 단백질은 동일하고, 여기서 융합 단백질의 DNA-결합 도메인은 제1 가이드 RNA(gRNA) 및 제2 gRNA와 연관될 수 있는 CRISPER 관련(Cas) 단백질 또는 그 변이체이며, 여기서

제1 DNA-표적화 모듈은 제1 HBV 유전자 또는 이의 조절 요소의 표적 부위를 표적으로 하는 제1 gRNA를 포함하고, 제2 DNA-표적화 모듈은 제2 HBV 유전자 또는 이의 조절 요소의 표적 부위를 표적으로 하는 제2 gRNA를 포함하는 것인, 후성적 변형 DNA-표적화 시스템.

청구항 6

하나 이상의 B형 간염 바이러스(HBV) 유전자의 전사를 억제하기 위한 후성적 변형 DNA-표적화 시스템으로서, 여기서 DNA-표적화 시스템은:

(a) CRISPR 관련 (Cas) 단백질 또는 이의 변이체 및 적어도 하나의 전사 억제자 효과기 도메인을 포함하는 융합 단백질; 및

(b) 적어도 제1 gRNA와 제2 gRNA를 포함하는 복수의 가이드 RNA(gRNA)로서, 여기서 제1 gRNA는 제1 HBV 유전자 또는 이의 조절 요소의 표적 부위를 표적으로 하고 제2 gRNA는 제2 HBV 유전자 또는 이의 조절 요소의 표적 부위를 표적으로 하는 것인 복수의 가이드 RNA(gRNA)

를 포함하고, 여기서 제1 및 제2 유전자 또는 이의 조절 요소는 B형 간염 바이러스 복제 및/또는 HBV 전사를 조절하는 것인, 후성적 변형 DNA-표적화 시스템.

청구항 7

청구항 6에 있어서, DNA-표적화 시스템은 B형 간염 바이러스 복제 및/또는 HBV 전사를 조절하는 제3 유전자 또는 이의 조절 요소의 표적 부위를 표적화하는 제3 gRNA를 더 포함하고,

선택적으로, 상기 시스템은 추가로 제4 유전자 또는 이의 조절 요소의 표적 부위를 표적으로 하는 제4 gRNA, 선택적으로 제5 유전자 또는 이의 조절 요소를 표적으로 하는 제5 gRNA, 및/또는 선택적으로 제6 유전자 또는 이의 조절 요소의 표적 부위를 표적으로 하는 제6 gRNA를 더 포함하고,

여기서 유전자 또는 이의 조절 요소는 B형 간염 바이러스 복제 및/또는 HBV 전사를 조절하는 것인, 후성적 변형 DNA-표적화 시스템.

청구항 8

청구항 7에 있어서, 제1, 제2, 제3, 제4, 제5 및/또는 제6 유전자 또는 이의 조절 요소는 상이한 것인, 후성적 변형 DNA-표적화 시스템.

청구항 9

하나 이상의 B형 간염 바이러스(HBV) 유전자의 전사를 억제하기 위한 후성적 변형 DNA-표적화 시스템으로서, 상기 DNA-표적화 시스템은:

(a) CRSISPR(Clustered Regularly Interspaced Short Palindromic Repeats) 관련(CAS) 단백질 또는 이의 변이체와 적어도 하나의 전사 억제자 효과기 도메인을 포함하는 융합 단백질; 및

(b) 복수의 유전자 또는 이의 조절 요소의 복수의 표적 부위를 표적으로 하는 복수의 가이드 RNA(gRNA)

를 포함하고, 여기서 상기 복수의 유전자 또는 이의 조절 요소는 B형 간염 바이러스의 복제 및/또는 HBV 전사를 조절하는 것인, 후성적 변형 DNA-표적화 시스템.

청구항 10

하나 초과의 B형 간염 바이러스(HBV) 유전자의 전사를 억제하기 위한 단일 DNA-표적화 모듈을 포함하는 후성적 변형 DNA-표적화 시스템으로서, 여기서 DNA-표적화 모듈은:

(a) CRSISPR(Clustered Regularly Interspaced Short Palindromic Repeats) 관련(CAS) 단백질 또는 이의 변이체와 적어도 하나의 전사 억제자 효과기 도메인을 포함하는 융합 단백질; 및

(b) 복수의 유전자 또는 이의 조절 요소의 복수의 표적 부위를 표적화하는 가이드 RNA(gRNA)

를 포함하고, 여기서 상기 복수의 유전자 또는 이의 조절 요소는 B형 간염 바이러스 복제 및/또는 HBV 전사를 조절하는 것인, 후성적 변형 DNA-표적화 시스템.

청구항 11

청구항 1-10 중 어느 한 항에 있어서, 전사 억제는 HBV 복제 감소 및/또는 HBV 단백질 수치 감소가 초래하는 것인, 후성적 변형 DNA-표적화 시스템.

청구항 12

청구항 1-11 중 어느 한 항에 있어서, DNA-표적화 시스템은 유전적 파괴 또는 DNA 파손을 도입하지 않는, 후성적 변형 DNA-표적화 시스템.

청구항 13

청구항 1-9 및 11-12 중 어느 한 항에 있어서, 적어도 하나의 DNA-결합 모듈은 HBV DNA 서열 내 복수의 표적 부위를 함께 표적화하는 복수의 DNA-결합 모듈을 포함하고, 선택적으로 각 DNA-결합 모듈은 HBV DNA 서열에서 다른 표적 부위를 표적하는, 후성적 변형 DNA-표적화 시스템.

청구항 14

청구항 4, 5 및 9-13 중 어느 한 항에 있어서, 복수의 표적 부위는 2, 3, 4, 5 또는 6개의 다른 표적 부위인, 후성적 변형 DNA-표적화 시스템.

청구항 15

청구항 4, 5 및 9-14 중 어느 한 항에 있어서, 복수의 표적 부위는 각각 다른 HBV 유전자 또는 이의 조절 요소에 존재하는, 후성적 변형 DNA-표적화 시스템.

청구항 16

청구항 1-14 중 어느 한 항에 있어서, 각 표적 부위는 동일한 HBV 유전자 또는 이의 조절 요소에 존재하는, 후성적 변형 DNA-표적화 시스템.

청구항 17

청구항 1-5 및 11-16 중 어느 한 항에 있어서, 시스템은 2 내지 10개의 DNA-표적화 모듈을 포함하는, 후성적 변형 DNA-표적화 시스템.

청구항 18

청구항 1-5 및 11-17 중 어느 한 항에 있어서, DNA-표적화 모듈 중 두 개 이상은 동일한 융합 단백질을 공유하거나 또는 DNA-표적화 모듈 중 두 개 이상은 서로 다른 융합 단백질을 포함하는, 후성적 변형 DNA-표적화 시스템.

청구항 19

청구항 1-5 및 11-18 중 어느 한 항에 있어서, 각 DNA-표적화 모듈의 DNA-결합 도메인은 CRISPER 관련(Cas) 단백질 또는 이의 변이체와 적어도 하나의 전사 억제자 효과기 도메인을 포함하는 융합 단백질을 포함하고, 각 DNA-표적화 모듈은 고유한 gRNA를 포함하는, 후성적 변형 DNA-표적화 시스템.

청구항 20

청구항 1-19 중 어느 한 항에 있어서, 표적 부위 또는 각각의 표적 부위는 공유적으로 폐쇄된 원형 DNA(cccDNA) 형태, 이완된 원형 DNA(rcDNA) 형태 및/또는 인간 게놈 DNA에 통합된 HBV 바이러스 DNA로 존재하는, 후성적 변형 DNA-표적화 시스템.

청구항 21

청구항 1-20 중 어느 한 항에 있어서, 표적 부위 또는 각각의 표적 부위는 HBV 복제 및/또는 HBV 전사를 제어하는데 관여하는 유전자 또는 이의 조절 요소에 존재하거나 그 근처에 존재하는, 후성적 변형 DNA-표적화 시스템.

청구항 22

청구항 21에 있어서, HBV 복제 및/또는 HBV 전사를 제어하는데 관여하는 유전자는 폴리머라제, 외피 단백질, 캡시드 단백질, 전사 인자 또는 전사 전사 활성화제를 인코딩하는, 후성적 변형 DNA-표적화 시스템.

청구항 23

청구항 21 또는 청구항 22에 있어서, HBV 복제 및/또는 HBV 전사를 제어하는데 관여하는 유전자는 폴리머라제 유전자, S-패밀리 유전자, X-유전자 또는 코어 패밀리 유전자인, 후성적 변형 DNA-표적화 시스템.

청구항 24

청구항 1-23 중 어느 한 항에 있어서, 적어도 하나의 표적 부위는 B형 간염 바이러스 단백질 X(HBx)를 인코딩하는 X 유전자의 유전자 또는 이의 조절 요소에 존재하는, 후성적 변형 DNA-표적화 시스템.

청구항 25

청구항 1-24 중 어느 한 항에 있어서, 표적 부위 또는 각각의 표적 부위는 HBV 복제 및/또는 HBV 전사를 제어하는 HBV 유전자의 조절 요소에 있거나 그 근처에 있는, 후성적 변형 DNA-표적화 시스템.

청구항 26

청구항 25에 있어서, 조절 요소는 프로모터 영역인, 후성적 변형 DNA-표적화 시스템.

청구항 27

청구항 26에 있어서, 프로모터 영역은 pre-S1 프로모터, pre-S2 프로모터, X 프로모터 또는 기저 코어 프로모터인, 후성적 변형 DNA-표적화 시스템.

청구항 28

청구항 25에 있어서, 조절 요소는 인핸서 영역인, 후성적 변형 DNA-표적화 시스템.

청구항 29

청구항 28에 있어서, 인핸서 영역은 Enh1 또는 Enh2 인핸서 영역인, 후성적 변형 DNA-표적화 시스템.

청구항 30

청구항 25에 있어서, 조절 요소는 전사 처리 제어 영역인, 후성적 변형 DNA-표적화 시스템.

청구항 31

청구항 1-24 중 어느 한 항에 있어서, 표적 부위 또는 각각의 표적 부위는 HBV 유전자의 코딩 영역에 있는, 후성적 변형 DNA-표적화 시스템.

청구항 32

청구항 1-31 중 어느 한 항에 있어서, 표적 부위 또는 각각의 표적 부위는 전사 개시 부위로부터 500개 염기쌍(bp) 이내, 1000bp 이내, 1500bp 이내에 위치하는, 후성적 변형 DNA-표적화 시스템.

청구항 33

청구항 1-32 중 어느 한 항에 있어서, 표적 부위 또는 각각의 표적 부위는 HBV 게놈의 0-3300 염기쌍(bp) 사이의 염기쌍에 위치한 표적 영역 내에 위치하며, 선택적으로 서열번호 650에 제시된 HBV 게놈을 참조로 위치에 상응하는 0-3182 bp 사이에 위치하는, 후성적 변형 DNA-표적화 시스템.

청구항 34

청구항 1-33 중 어느 한 항에 있어서, 표적 부위 또는 각각의 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 HBV 게놈의 43bp - 490bp, 1033bp - 1749bp, 1800bp - 1950bp 또는 2953bp - 3182bp 사이의 염기쌍에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내에 위치하는, 후성적 변형 DNA-표적화 시스템.

청구항 35

청구항 1-33 중 어느 한 항에 있어서, 표적 부위 또는 각각의 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 HBV 게놈의 1 bp - 42 bp, 491 bp - 1032 bp, 1750 bp - 1799 bp, 또는 1951 bp - 2952 bp 사이의 염기쌍에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내에 위치하는, 후성적 변형 DNA-표적화 시스템.

청구항 36

청구항 1-33 중 어느 한 항에 있어서, 표적 부위 또는 각각의 표적 부위는 HBV 게놈의 CpG 섬에 있는, 후성적 변형 DNA-표적화 시스템.

청구항 37

청구항 1-33 및 36 중 어느 한 항에 있어서, 표적 부위 또는 각각의 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 HBV 게놈의 67bp-392bp, 1033bp-1749bp 또는 2215bp-2490bp 사이의 염기쌍에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내에 위치하는, 후성적 변형 DNA-표적화 시스템.

청구항 38

청구항 1-33, 36 및 37 중 어느 한 항에 있어서, 표적 부위 또는 각각의 표적 부위는 서열번호 650의 뉴클레오타이드 위치를 참조로 B형 간염 바이러스 서열의 1033bp-1749bp 사이의 염기쌍에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내에 위치하는, 후성적 변형 DNA-표적화 시스템.

청구항 39

청구항 1-37 중 어느 한 항에 있어서, 표적 부위 또는 각각의 표적 부위는 B형 간염 X 단백질(HBx) 시작 코돈 상류의 300개 염기쌍 내에 위치한 표적 영역 내에 있는, 후성적 변형 DNA-표적화 시스템.

청구항 40

하나 이상의 B형 간염 바이러스(HBV) 유전자의 전사를 억제하기 위한 적어도 하나의 DNA-표적화 모듈을 포함하는 후성적 변형 DNA-표적화 시스템으로서, 여기서 적어도 하나의 DNA-표적화 모듈 각각은 다음을 포함하는 융합 단백질을 포함하는 것인, 후성적 변형 DNA-표적화 시스템:

(a) B형 간염 X 단백질(HBx) 시작 코돈 상류의 300개 염기쌍 내에 있는 표적 영역 내의 표적 부위를 표적화하기 위한 DNA-결합 도메인; 및

(b) 적어도 하나의 전사 억제자 효과기 도메인.

청구항 41

청구항 1-40 중 어느 한 항에 있어서, 표적 부위 또는 각각의 표적 부위는 HBx 기저 코어 프로모터 영역에 위치하는, 후성적 변형 DNA-표적화 시스템.

청구항 42

청구항 1-40 중 어느 한 항에 있어서, 표적 부위 또는 각각의 표적 부위는 HBx 프로모터/인핸서 영역 내에 위치하는, 후성적 변형 DNA-표적화 시스템.

청구항 43

청구항 1-42 중 어느 한 항에 있어서, 표적 부위 또는 각각의 표적 부위는 B형 간염 X 단백질(HBx) 시작 코돈 상류의 250개 염기쌍 내에 걸쳐 있는 표적 영역 내에 있는, 후성적 변형 DNA-표적화 시스템.

청구항 44

청구항 1-43 중 어느 한 항에 있어서, 표적 부위 또는 각각의 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 HBV 게놈의 1060-1480bp 사이의 염기쌍에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내에 있는, 후성적 변형 DNA-표적화 시스템.

청구항 45

청구항 1-44 중 4, 표적 부위 또는 각각의 표적 부위는 B형 간염 X 단백질(HBx) 시작 코돈 상류의 150개 염기쌍 내에 걸쳐 있는 표적 영역 내에 있는, 후성적 변형 DNA-표적화 시스템.

청구항 46

청구항 1-45 중 어느 한 항에 있어서, 표적 부위 또는 각각의 표적 부위는 B형 간염 X 단백질(HBx) 시작 코돈 상류의 120개 염기쌍 내에 걸쳐 있는 표적 영역 내에 있는, 후성적 변형 DNA-표적화 시스템.

청구항 47

청구항 1-46 중 어느 한 항에 있어서, 표적 부위 또는 각각의 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 HBV 게놈의 1250-1374bp 사이의 염기쌍에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내에 있는, 후성적 변형 DNA-표적화 시스템.

청구항 48

청구항 1-47 중 어느 한 항에 있어서, 표적 부위 또는 각각의 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 HBV 게놈의 1255-1302bp 사이의 염기쌍에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내에 있는, 후성적 변형 DNA-표적화 시스템.

청구항 49

청구항 1-48 중 어느 한 항에 있어서, 표적 부위 또는 각각의 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 HBV 게놈의 1260-1300bp 사이의 염기쌍에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내에 있는, 후성적 변형 DNA-표적화 시스템.

청구항 50

청구항 1-49 중 어느 한 항에 있어서, 표적 부위 또는 각각의 표적 부위는 모든 B형 간염 바이러스 게놈과 최소 70% 상동성을 갖는, 후성적 변형 DNA-표적화 시스템.

청구항 51

청구항 1-50 중 어느 한 항에 있어서, 표적 부위 또는 각각의 표적 부위는 최소 1000개의 B형 간염 바이러스 게놈과 최소 70% 상동성을 갖는, 후성적 변형 DNA-표적화 시스템.

청구항 52

청구항 1-51 중 어느 한 항에 있어서, 표적 부위 또는 각각의 표적 부위는 최소 1000개의 B형 간염 바이러스 게놈과 최소 70% 상동적이며 최대 2개의 불일치를 포함하는, 후성적 변형 DNA-표적화 시스템.

청구항 53

청구항 1-52 중 어느 한 항에 있어서, 표적 부위 또는 각각의 표적 부위는 서열번호 1-195 중 어느 하나에 제시된 서열, 적어도 14개 뉴클레오타이드(nt)의 이의 연속 부분 또는 상기 서열 중 어느 하나의 상보적 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

청구항 54

청구항 1-53 어느 한 항에 있어서, 표적 부위 또는 각각의 표적 부위는 서열번호 175, 138, 192, 152, 118, 125, 185, 63, 116, 124, 35, 82 중 어느 하나에 제시된 서열, 적어도 14개 뉴클레오타이드(nt)의 이의 연속 부분, 또는 상기 중 어느 것의 상보적 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

청구항 55

청구항 1-54 중 어느 한 항에 있어서, 표적 부위 또는 각각의 표적 부위는 서열번호 175, 138, 192, 152, 118, 125, 185, 63, 116, 124, 35, 82 중 어느 하나에 제시된 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

청구항 56

청구항 1-53 중 어느 한 항에 있어서, 표적 부위 또는 각각의 표적 부위는 서열번호 5, 6, 12, 18, 22, 26, 29, 38, 42, 43, 51, 56, 61, 63, 68, 72, 75, 79, 82, 84, 88, 89, 98, 99, 113, 116, 121, 124, 125, 118, 130, 133, 135, 138, 143, 150, 152, 155, 158, 164, 165, 175, 176, 182, 185, 189, 190, 192 중 어느 하나에 제시된 서열, 적어도 14개 뉴클레오타이드(nt)의 이의 연속 부분 또는 상기 중 어느 하나의 상보적 서열을 포함하

는, 후성적 변형 DNA-표적화 시스템.

청구항 57

청구항 1-53 및 56 중 어느 한 항에 있어서, 표적 부위 또는 각각의 표적 부위는 서열번호 5, 6, 12, 18, 22, 26, 29, 38, 42, 43, 51, 56, 61, 63, 68, 72, 75, 79, 82, 84, 88, 89, 98, 99, 113, 116, 121, 124, 125, 118, 130, 133, 135, 138, 143, 150, 152, 155, 158, 164, 165, 175, 176, 182, 185, 189, 190, 192 중 어느 하나에 제시된 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

청구항 58

청구항 1-53 및 58 중 어느 한 항에 있어서, 표적 부위 또는 각각의 표적 부위는 서열번호 12, 18, 20, 22, 26, 27, 46, 50, 63, 66, 73, 79, 185, 192 중 어느 하나에 제시된 서열, 적어도 14개 뉴클레오타이드의 이의 연속 부분 또는 상기 중 어느 하나의 상보적 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

청구항 59

청구항 1-53 및 58 중 어느 한 항에 있어서, 표적 부위 또는 각각의 표적 부위는 서열번호 12, 18, 20, 22, 26, 27, 46, 50, 63, 66, 73, 79, 185, 192 중 어느 하나에 제시된 것인, 후성적 변형 DNA-표적화 시스템.

청구항 60

청구항 1-53 및 56-59 중 어느 한 항에 있어서, 표적 부위 또는 각각의 표적 부위는 서열번호 22에 제시된 서열, 적어도 14개 뉴클레오타이드의 이의 연속 부분 또는 상기 중 어느 하나의 상보적 서열을 포함하고, 선택적으로 표적 부위는 서열번호 22에 제시된 것인, 후성적 변형 DNA-표적화 시스템.

청구항 61

청구항 1-53 및 56-59 중 어느 한 항에 있어서, 표적 부위 또는 각각의 표적 부위는 서열번호 63에 제시된 서열, 적어도 14개 뉴클레오타이드의 이의 연속 부분 또는 상기 중 어느 하나의 상보적 서열을 포함하고, 선택적으로 표적 부위는 서열번호 63에 제시된 것인, 후성적 변형 DNA-표적화 시스템.

청구항 62

청구항 2-53 중 어느 한 항에 있어서, gRNA 또는 각각의 gRNA는 서열번호 196-390 중 어느 하나에 제시된 서열을 포함하는 gRNA 스페이서 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

청구항 63

청구항 62에 있어서, gRNA 또는 각각의 gRNA는 서열번호 587에 제시된 서열을 추가로 포함하는, 후성적 변형 DNA-표적화 시스템.

청구항 64

청구항 2-53, 62 및 63 중 어느 한 항에 있어서, gRNA 또는 각각의 gRNA는 서열번호 391-585 중 어느 하나에 제시된 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

청구항 65

청구항 2-53 및 62-64 중 어느 한 항에 있어서, gRNA 또는 각각의 gRNA는 서열번호 391-585 중 어느 하나에 제시된 것인, 후성적 변형 DNA-표적화 시스템.

청구항 66

청구항 2-53 및 62-64 중 어느 한 항에 있어서, gRNA 또는 각각의 gRNA는 서열번호 370, 333, 387, 347, 313, 320, 380, 256, 258, 311, 319, 230, 272 중 어느 하나에 제시된 서열, 적어도 14개 뉴클레오타이드의 이의 연속 부분, 또는 상기 중 어느 하나의 상보적 서열을 포함하고, 선택적으로 gRNA 또는 각각의 gRNA는 서열번호 565, 528, 542, 508, 515, 575, 515, 453, 506, 514, 425 또는 472 중 어느 하나에 제시된 것인, 후성적 변형 DNA-표적화 시스템.

청구항 67

청구항 2-53 및 62-66 중 어느 한 항에 있어서, gRNA 또는 각각의 gRNA는 서열번호 370, 333, 387, 347, 313, 320, 380, 256, 258, 311, 319, 230, 272 중 어느 하나에 제시된 서열을 포함하고, 선택적으로 gRNA 또는 각각의 gRNA는 서열번호 565, 528, 542, 508, 515, 575, 515, 453, 506, 514, 425 또는 472 중 어느 하나에 제시된 것인, 후성적 변형 DNA-표적화 시스템.

청구항 68

청구항 2-53 및 62-65 중 어느 한 항에 있어서, gRNA 또는 각각의 gRNA는 서열번호 200, 201, 207, 217, 221, 224, 233, 237, 238, 246, 251, 256, 258, 263, 267, 274, 270, 277, 279, 283, 284, 293, 294, 308, 311, 313, 316, 319, 320, 325, 328, 330, 333, 338, 345, 347, 350, 353, 359, 360, 370, 371, 377, 380, 384, 385, 387 중 어느 하나에 제시된 서열. 적어도 14개 뉴클레오타이드의 이의 연속 부분 또는 상기 중 어느 하나의 상보적 서열을 포함하고, 선택적으로 여기서 gRNA 또는 각각의 gRNA는 서열번호 395, 402, 408, 412, 416, 419, 428, 432, 433, 441, 446, 451, 453, 458, 462, 465, 469, 472, 474, 478, 479, 488, 489, 503, 506, 508, 511, 514, 515, 520, 523, 525, 575, 528, 533, 540, 542, 545, 548, 554, 555, 565, 566, 572, 579, 580, 또는 582 중 어느 하나에 제시된 것인, 후성적 변형 DNA-표적화 시스템.

청구항 69

청구항 2-53, 62-65 및 68 중 어느 한 항에 있어서, gRNA 또는 각각의 gRNA는 서열번호 200, 201, 207, 217, 221, 224, 233, 237, 238, 246, 251, 256, 258, 263, 267, 274, 270, 277, 279, 283, 284, 293, 294, 308, 311, 313, 316, 319, 320, 325, 328, 330, 333, 338, 345, 347, 350, 353, 359, 360, 370, 371, 377, 380, 384, 385 또는 387 중 어느 하나에 제시된 서열을 포함하고, 선택적으로 gRNA 또는 각각의 gRNA는 서열번호 395, 402, 408, 412, 416, 419, 428, 432, 433, 441, 446, 451, 453, 458, 462, 465, 469, 472, 474, 478, 479, 488, 489, 503, 506, 508, 511, 514, 515, 520, 523, 525, 575, 528, 533, 540, 542, 545, 548, 554, 555, 565, 566, 572, 579, 580, 또는 582 중 어느 하나에 제시된 것인, 후성적 변형 DNA-표적화 시스템.

청구항 70

청구항 2-53 및 62-65 중 어느 한 항에 있어서, gRNA 또는 각각의 gRNA는 서열번호 207, 213, 215, 217, 221, 222, 241, 245, 258, 261, 268, 274, 380, 387 중 어느 하나에 제시된 서열, 적어도 14개 뉴클레오타이드의 이의 연속 부분, 또는 상기 중 어느 하나의 상보적 서열을 포함하고, 선택적으로 gRNA 또는 각각의 gRNA는 서열번호 402, 408, 410, 412, 416, 417, 436, 440, 453, 456, 463, 469, 575, 582 중 어느 하나에 제시된 것인, 후성적 변형 DNA-표적화 시스템.

청구항 71

청구항 2-53, 62-65 및 70 중 어느 한 항에 있어서, gRNA 또는 각각의 gRNA는 서열번호 207, 213, 215, 217, 221, 222, 241, 245, 258, 261, 268, 274, 380, 387 중 어느 하나에 제시된 서열을 포함하고, 선택적으로 gRNA 또는 각각의 gRNA는 서열번호 402, 408, 410, 412, 416, 417, 436, 440, 453, 456, 463, 469, 575, 582 중 어느 하나에 제시된 것인, 후성적 변형 DNA-표적화 시스템.

청구항 72

청구항 2-53, 62-65 및 68-71 중 어느 한 항에 있어서, gRNA 또는 각각의 gRNA는 서열번호 217에 제시된 서열, 적어도 14개 뉴클레오타이드의 이의 연속 부분 또는 상기 중 어느 하나의 상보적 서열을 포함하고, 선택적으로 gRNA 또는 각각의 gRNA는 서열번호 217에 제시된 서열을 포함하며, 선택적으로 gRNA 또는 각각의 gRNA는 서열번호 412 중 어느 하나에 제시된 것인, 후성적 변형 DNA-표적화 시스템.

청구항 73

청구항 2-53, 62-65 및 68-71 중 어느 한 항에 있어서, gRNA 또는 각각의 gRNA는 서열번호 258에 제시된 서열, 적어도 14개 뉴클레오타이드의 이의 연속 부분 또는 상기 중 어느 하나의 상보적 서열을 포함하고, 선택적으로 gRNA 또는 각각의 gRNA는 서열번호 258에 제시된 서열을 포함하며, 선택적으로 gRNA 또는 각각의 gRNA는 서열번호 453 중 어느 하나에 제시된 것인, 후성적 변형 DNA-표적화 시스템.

청구항 74

청구항 1-73 중 어느 한 항에 있어서, 표적 부위 또는 각각의 표적 부위는 모든 B형 간염 바이러스 게놈과 최소 90% 상동성인, 후성적 변형 DNA-표적화 시스템.

청구항 75

청구항 1-74 중 어느 한 항에 있어서, 표적 부위 또는 각각의 표적 부위는 최소 1000개의 B형 간염 바이러스 게놈과 최소 90% 상동성인, 후성적 변형 DNA-표적화 시스템.

청구항 76

청구항 1-75 중 어느 한 항에 있어서, 표적 부위 또는 각각의 표적 부위는 최소 1000개의 B형 간염 바이러스 게놈과 최소 90% 상동적이고 최대 2개의 불일치, 선택적으로 1개 또는 2개의 불일치를 포함하는, 후성적 변형 DNA-표적화 시스템.

청구항 77

청구항 1-53, 62-65 및 74-76 중 어느 한 항에 있어서, 표적 부위 또는 각각의 표적 부위는 서열번호 35-100 중 어느 하나에 제시된 서열, 적어도 14개 뉴클레오타이드(nt)의 이의 연속 부분 또는 상기 서열 중 어느 하나의 상보적 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

청구항 78

청구항 2-53, 62-65 및 74-77 중 어느 한 항에 있어서, gRNA 또는 각각의 gRNA는 서열번호 230-295 중 어느 하나에 제시된 서열을 포함하는 gRNA 스페이서 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

청구항 79

청구항 78에 있어서, gRNA 또는 각각의 gRNA는 추가로 서열번호 587에 제시된 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

청구항 80

청구항 2-53, 62-65 및 74-79 중 어느 한 항에 있어서, gRNA 또는 각각의 gRNA는 서열번호 230-295 중 어느 하나에 제시된 서열을 포함하고, 선택적으로 gRNA 또는 각각의 gRNA는 서열번호 425-490 중 어느 하나에 제시된 것인, 후성적 변형 DNA-표적화 시스템.

청구항 81

청구항 52-80 중 어느 한 항에 있어서, 최대 2개의 불일치는 gRNA 프로토스페이서의 5' 말단의 처음 12 nt에 위치하는, 후성적 변형 DNA-표적화 시스템.

청구항 82

청구항 1-51 및 53-73, 75 및 77-81 중 어느 한 항에 있어서, 표적 부위 또는 각각의 표적 부위는 최소 1000개의 B형 간염 바이러스 게놈과 최소 90% 상동적이며 불일치가 없는 것인, 후성적 변형 DNA-표적화 시스템.

청구항 83

청구항 1-53, 62-65, 74-76 및 82 중 어느 한 항에 있어서, 표적 부위는 서열번호 1-34 중 어느 하나에 제시된 서열, 적어도 14개 뉴클레오타이드(nt)의 이의 연속 부분 또는 상기 중 어느 하나의 상보적 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

청구항 84

청구항 2-53, 62-65, 74-76, 82 및 83 중 어느 한 항에 있어서, gRNA 또는 각각의 gRNA는 서열번호 196-229에 제시된 서열을 포함하는 gRNA 스페이서 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

청구항 85

청구항 62-84 중 어느 한 항에 있어서, gRNA 스페이서 서열은 길이가 14nt 내지 24nt 또는 16nt 내지 22nt인, 후성적 변형 DNA-표적화 시스템.

청구항 86

청구항 62-85 중 어느 한 항에 있어서, gRNA 스페이서 서열은 길이가 18 nt, 19 nt, 20 nt, 21 nt 또는 22 nt인, 후성적 변형 DNA-표적화 시스템.

청구항 87

청구항 62-86 중 어느 한 항에 있어서, gRNA 스페이서 서열은 안정성을 높이기 위해 변형된 뉴클레오타이드를 포함하는, 후성적 변형 DNA-표적화 시스템.

청구항 88

청구항 2-53, 62, 74-76 및 82-87 중 어느 한 항에 있어서, 상기 적어도 하나의 gRNA는 추가로 서열번호 587에 제시된 서열을 포함함.

청구항 89

청구항 2-53, 62, 74-76 및 82-88 중 어느 한 항에 있어서, gRNA 또는 각각의 gRNA는 서열번호 196-229 중 어느 하나에 제시된 서열을 포함하고, 선택적으로 gRNA 또는 각각의 gRNA는 서열번호 391-424 중 어느 하나에 제시된 것인, 후성적 변형 DNA-표적화 시스템.

청구항 90

청구항 2-53, 62, 74-76 및 82-89 중 어느 한 항에 있어서, gRNA 또는 각각의 gRNA는 서열번호 207, 213, 215, 217, 221, 222, 241, 245, 258, 261, 268, 274, 380, 387 중 어느 하나에 제시된 서열을 포함하고, 선택적으로 gRNA 또는 각각의 gRNA는 서열번호 402, 408, 410, 412, 416, 417, 436, 440, 453, 456, 463, 469, 575, 582 중 어느 하나에 제시된 것인, 후성적 변형 DNA-표적화 시스템.

청구항 91

청구항 2-53, 62, 74-76 및 82-90 중 어느 한 항에 있어서, gRNA는 서열번호 217에 제시된 서열을 포함하고, 선택적으로 gRNA는 서열번호 412에 제시된 것인, 후성적 변형 DNA-표적화 시스템.

청구항 92

청구항 2-91 중 후성적 변형 DNA-표적화 시스템, Cas 단백질 또는 이의 변이체는 Cas9 단백질 또는 이의 변이체인, 후성적 변형 DNA-표적화 시스템.

청구항 93

청구항 2-92 중 어느 한 항에 있어서, Cas 단백질 또는 이의 변이체는 Cas12 단백질 또는 이의 변이체인, 후성적 변형 DNA-표적화 시스템.

청구항 94

청구항 2-93 중 어느 한 항에 있어서, Cas 단백질 또는 이의 변이체는 변이체 Cas 단백질이고, 변이체 Cas 단백질은 뉴클레아제 활성이 없거나 비활성화된 Cas(dCas) 단백질인, 후성적 변형 DNA-표적화 시스템.

청구항 95

청구항 2-94 중 어느 한 항에 있어서. 변이체 Cas 단백질은 뉴클레아제 활성이 결여된 변이체 Cas9 단백질이거나 비활성화된 Cas9(dCas9) 단백질인, 후성적 변형 DNA-표적화 시스템

청구항 96

청구항 2-95 중 어느 한 항에 있어서, Cas9 단백질 또는 이의 변이체는 황색포도상구균(*Staphylococcus aureus*) Cas9 (SaCas9) 단백질 또는 이의 변이체인, 후성적 변형 DNA-표적화 시스템.

청구항 97

청구항 96에 있어서, 변이체 Cas9는 서열번호 596의 위치 번호를 참조로, D10A 및 N580A에서 선택된 적어도 하나의 아미노산 돌연변이를 포함하는 황색포도상구균 dCas9 단백질(dSaCas9)인, 후성적 변형 DNA-표적화 시스템.

청구항 98

청구항 96 또는 97에 이어서, 변이체 Cas9 단백질은 서열번호 597에 제시된 서열을 포함하거나, 서열번호 597과 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

청구항 99

청구항 2-95 중 어느 한 항에 있어서, Cas9 단백질 또는 이의 변이체는 화농성연쇄상구균(*Streptococcus pyogenes*) Cas9 (SpCas9) 단백질 또는 이의 변이체인, 후성적 변형 DNA-표적화 시스템

청구항 100

청구항 99에 있어서, 변이체 Cas9는 번호를 참조로 D10A 및 H840A 중에서 선택된 적어도 하나의 아미노산 돌연변이를 포함하는 화농성연쇄상구균(*Streptococcus pyogenes*) dCas9(dSpCas9) 단백질인, 후성적 변형 DNA-표적화 시스템.

청구항 101

청구항 99 또는 100에 있어서, 변이체 Cas9 단백질은 서열번호 599에 제시된 서열을 포함하거나, 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

청구항 102

청구항 1-5, 11-18, 20-33, 34-61 중 어느 한 항에 있어서, 상기 적어도 하나의 DNA-결합 도메인은 조작된 아연 집게 단백질(eZFP)을 포함하는, 후성적 변형 DNA-표적화 시스템.

청구항 103

청구항 1-5, 11-18, 20-33, 34-61 및 102 중 어느 한 항에 있어서, 적어도 하나의 DNA-결합 도메인은 eZFP인, 후성적 변형 DNA-표적화 시스템.

청구항 104

청구항 1-54, 11-18, 20-33, 34-61, 102 및 103 중 어느 한 항에 있어서, 표적 부위는 서열번호 1045, 1046, 1052 중 어느 하나에 제시된 뉴클레오타이드 서열, 최소 12nt의 이의 연속 부분, 또는 상기 중 어느 한 서열의 상보적 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

청구항 105

청구항 1-54, 11-18, 20-33, 34-61 및 102-104 중 어느 한 항에 있어서, 표적 부위는 서열번호 1045, 1046, 1052 중 어느 하나에 제시된 뉴클레오타이드 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

청구항 106

청구항 102-105 중 어느 한 항에 있어서, 아연 집게 단백질은 N-말단에서 C-말단 순서로 F1 내지 F6으로 표시되는 6개의 아연 집게를 포함하고, 여기서 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같은 것인 후성적 변형 DNA-표적화 시스템:

1) F1:SEADRSR(서열번호 720) F2:DRSNLTR(서열번호 721)F3:QSSDLSR(서열번호 722) F4:YHWYLKK(서열번호 723) F5:RSDSLSV(서열번호 724) F6:QNAANRKT(서열번호 725);

2) F1:RSDVLST(서열번호 726) F2:DNSSRTR(서열번호 727) F3:RPYTLRL(서열번호 728) F4:DSSHRTR(서열번호 729) F5:RSDHLSQ(서열번호 730) F6:DSSHRTR(서열번호 731);

3) F1:RSDHLSQ(서열번호 732) F2:QSADRTK(서열번호 733) F3:RSDHLSQ(서열번호 734) F4:RRSDLKR(서열번호 735) F5:RSDHLSR(서열번호 736) F6:QSSDLRR(서열번호 737);

4) F1:RSDNLSE(서열번호 738) F2:TSSNRKT(서열번호 739) F3:DRSHLTR(서열번호 740) F4:RSDALTQ(서열번호 741) F5:DRSALAR(서열번호 742) F6:RRFTLSK(서열번호 743);

5) F1:RSDHLSE(서열번호 744) F2:QYSGRYY(서열번호 745) F3:HGQTLNE(서열번호 746) F4:QSGNLAR(서열번호 747) F5:RSDSLLR(서열번호 748)F6:CREYRGK(서열번호 749);

6) F1:QSANRTT(서열번호 750) F2:RSANLTR(서열번호 751) F3:RSDVLSE(서열번호 752) F4:TSGHLSR(서열번호 753) F5:QSSDLSR(서열번호 754), F6:QWSTRKR(서열번호 755);

7) F1:QSGNLAR(서열번호 756) F2:ATCCLAH(서열번호 757) F3:RWQYLPT(서열번호 758) F4:DRSALAR(서열번호 759) F5:RSDNLSE(서열번호 760)F6:KRCNLRC(서열번호 761);

8) F1:NPANLTR(서열번호 762) F2:QNATRTK(서열번호 763) F3:QSGHLAR(서열번호 764) F4:NRHDRAK(서열번호 765) F5:RSDHLSE(서열번호 766), F6:QRRSRYK(서열번호 767);

9) F1:QSSDLSR(서열번호 768) F2:HRSTRNR(서열번호 769) F3:RSDVLSA(서열번호 770) F4:DSRTRKN(서열번호 771) F5:QSGSLTR(서열번호 772) F6:DQSGLAH(서열번호 773);

10) F1:QNPAQWR(서열번호 774) F2:RSADLSR(서열번호 775) F3:TSGSLSR(서열번호 776) F4:RSDHLSR(서열번호 777) F5:RSDSLLR(서열번호 778) F6:QSYDRFQ(서열번호 779);

11) F1:TSGSLSR(서열번호 780) F2:RSDHLSR(서열번호 781) F3:RSDSLLR(서열번호 782) F4:QSYDRFQ(서열번호 783) F5:RSDNLST(서열번호 784) F6:DNRDRIK(서열번호 785);

12) F1:DRSNLSR(서열번호 786) F2:LRQNLIM(서열번호 787) F3:ERGTLAR(서열번호 788) F4:RSDALTQ(서열번호 789) F5:RSDSLSQ(서열번호 790) F6:RKADRTR(서열번호 791);

13) F1:QYCCLTN(서열번호 792) F2:TSGNLTR(서열번호 793) F3:QSSDLSR(서열번호 794) F4:FRYYLKR(서열번호 795) F5:QSGDLTR(서열번호 796) F6:DKGNLTK(서열번호 797);

14) F1:TSGSLSR(서열번호 798) F2:RSDNLTT(서열번호 799) F3:QSGNLAR(서열번호 800) F4:DRTTLMR(서열번호 801) F5:QSGHLAR(서열번호 802) F6:QLTHLNS(서열번호 803);

15) F1:IKHDLHR(서열번호 804) F2:RSANLTR(서열번호 805) F3:RSDNLAR(서열번호 806) F4:QNVSRPR(서열번호 807) F5:RSDDLSK(서열번호 808) F6:DSSHRTR(서열번호 809);

16) F1:RSDNLAR(서열번호 810) F2:QNVSRPR(서열번호 811) F3:RSDDLSK(서열번호 812) F4:DSSHRTR(서열번호 813) F5:TSSNRKT(서열번호 814) F6:AQWTRAC(서열번호 815);

17) F1:RSDDLSK(서열번호 816) F2:DSSHRTR(서열번호 817) F3:TSSNRKT(서열번호 818) F4:AQWTRAC(서열번호 819) F5:RKQTRTT(서열번호 820) F6:HRSSLRR(서열번호 821);

18) F1:QSAHRKN(서열번호 822) F2:TSSNRKT(서열번호 823) F3:RSDNLSA(서열번호 824) F4:RNNDRKT(서열번호 825) F5:TSGSLSR(서열번호 826) F6:QAGHLAK(서열번호 827);

19) F1:RSDHLSQ(서열번호 828) F2:ASSTRTK(서열번호 829) F3:RSDDLTR(서열번호 830) F4:QKSNLSS(서열번호 831) F5:QSANRTT(서열번호 832) F6:QNATRTK(서열번호 833);

20) F1:RSDTLSE(서열번호 834) F2:RRWTLVG(서열번호 835) F3:DRSNLSR(서열번호 836) F4:QSGDLTR(서열번호 837) F5:QSSDLSR(서열번호 838) F6:YHWYLKK(서열번호 839);

21) F1:RSANLAR(서열번호 840) F2:RSDNLRE(서열번호 841) F3:RPYTLRL(서열번호 842) F4:HRSNLNK(서열번호 843) F5:QSGSLTR(서열번호 844) F6:TSANLSR(서열번호 845);

22) F1:RSDDLVR(서열번호 846) F2:TSGSLVR(서열번호 847) F3:RSDKLVR(서열번호 848) F4:RSDELVR(서열번호 849) F5:TSHSLTE(서열번호 850) F6:RADNLTE(서열번호 851);

23) F1:ERSHLRE(서열번호 852) F2:TSHSLTE(서열번호 853) F3:QAGHLAS(서열번호 854) F4:TSHSLTE(서열번호 855) F5:DPGHLVR(서열번호 856) F6:TSGNLVR(서열번호 857);

24) F1:RADNLTE(서열번호 858) F2:TSGSLVR(서열번호 859) F3:RKDNLKN(서열번호 860) F4:QSSSLVR(서열번호 861) F5:RSDKLVR(서열번호 862) F6:DSGNLRV(서열번호 863);

25) F1:QSSSLVR(서열번호 864) F2:QSGDLRR(서열번호 865) F3:RSDERKR(서열번호 866) F4:HRTTLTN(서열번호 867) F5:RSDHLTN(서열번호 868) F6:TSGELVR(서열번호 869);

26) F1:QSGDLRR(서열번호 870) F2:RSDERKR(서열번호 871) F3:HRTTLTN(서열번호 872) F4:RSDHLTN(서열번호 873) F5:TSGELVR(서열번호 874) F6:RSDDLVR(서열번호 875);

27) F1:QRAHLER(서열번호 876) F2:QLAHLRA(서열번호 877) F3:DPGHLVR(서열번호 878) F4:RRSACRR(서열번호 879) F5:RSDHLTT(서열번호 880) F6:QSSSLVR(서열번호 881); 및

28) F1:QSSNLVR(서열번호 882) F2:RSDDLVR(서열번호 883) F3:THLDLIR(서열번호 884) F4:TSGNLTE(서열번호 885) F5:RRSACRR(서열번호 886) F6:RNDTLTE(서열번호 887).

청구항 107

청구항 102-106 중 어느 한 항에 있어서, 아연 집게 단백질은 N-말단에서 C-말단 순서로 F1에서 F6으로 표시되는 6개의 아연 집게를 포함하고,

여기서 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같은 것인, 후성적 변형 DNA-표적화 시스템:

F1:QSAHRKN(서열번호 822) F2:TSSNRKT(서열번호 823) F3:RSDNLSA(서열번호 824) F4:RNNDRKT(서열번호 825) F5:TSGSLSR(서열번호 826) F6:QAGHLAK(서열번호 827).

청구항 108

청구항 102-106 중 어느 한 항에 있어서, 아연 집게 단백질은 N-말단에서 C-말단 순서로 F1에서 F6까지 표시된 6개의 아연 집게를 포함하고,

여기서 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같은 것인, 후성적 변형 DNA-표적화 시스템:

F1:RSDHLSQ(서열번호 828) F2:ASSTRTK(서열번호 829) F3:RSDDLTR(서열번호 830) F4:QKSNLSS(서열번호 831) F5:QSANRTT(서열번호 832) F6:QNATRTK(서열번호 833),

청구항 109

청구항 102-106 중 어느 한 항에 있어서, 아연 집게 단백질은 N-말단에서 C-말단의 순서로 F1에서 F6까지 표시된 6개의 아연 집게를 포함하고,

여기서 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같은 것인, 후성적 변형 DNA-표적화 시스템:

F1:QSSSLVR(서열번호 864) F2:QSGDLRR(서열번호 865) F3:RSDERKR(서열번호 866) F4:HRTTLTN(서열번호 867) F5:RSDHLTN(서열번호 868) F6:TSGELVR(서열번호 869).

청구항 110

후성적 변형 DNA-표적화 시스템으로서, 상기 후성적 변형 DNA-표적화 시스템으로서은 a) 하나 이상의 HBV 유전자 또는 이의 조절 요소의 표적 부위에 결합하는 eZFP 및 b) 하나 이상의 HBV 유전자의 전사를 억제하는 적어도 하나의 효과기 도메인을 포함하고, 여기서 아연 집게 단백질은 N-말단에서 C-말단 순서로 F1에서 F6까지 표시된 6개의 아연 집게를 포함하며, 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같은 것인, 후성적 변형 DNA-표적화 시스템:

1) F1:SEADRSR(서열번호 720) F2:DRSNLTR(서열번호 721)F3:QSSDLSR(서열번호 722) F4:YHWYLKK(서열번호 723) F5:RSDSLSV(서열번호 724) F6:QNAANRKT(서열번호 725);

2) F1:RSDVLST(서열번호 726) F2:DNSSRTR(서열번호 727) F3:RPYTLRL(서열번호 728) F4:DSSHRTR(서열번호 729) F5:RSDHLSQ(서열번호 730) F6:DSSHRTR(서열번호 731);

3) F1:RSDHLSQ(서열번호 732) F2:QSADRTK(서열번호 733) F3:RSDHLSQ(서열번호 734) F4:RRSDLKR(서열번호 735) F5:RSDHLSR(서열번호 736) F6:QSSDLRR(서열번호 737);

4) F1:RSDNLSE(서열번호 738) F2:TSSNRKT(서열번호 739) F3:DRSHLTR(서열번호 740) F4:RSDALTQ(서열번호 741) F5:DRSALAR(서열번호 742) F6:RRFTLSK(서열번호 743);

5) F1:RSDHLSE(서열번호 744) F2:QYSGRYY(서열번호 745) F3:HGQTLNE(서열번호 746) F4:QSGNLAR(서열번호 747) F5:RSDSLLR(서열번호 748)F6:CREYRGK(서열번호 749);

6) F1:QSANRTT(서열번호 750) F2:RSANLTR(서열번호 751) F3:RSDVLSE(서열번호 752) F4:TSGHLSR(서열번호 753) F5:QSSDLSR(서열번호 754), F6:QWSTRKR(서열번호 755);

7) F1:QSGNLAR(서열번호 756) F2:ATCCLAH(서열번호 757) F3:RWQYLPT(서열번호 758) F4:DRSALAR(서열번호 759) F5:RSDNLSE(서열번호 760)F6:KRCNLRC(서열번호 761);

8) F1:NPANLTR(서열번호 762) F2:QNATRTK(서열번호 763) F3:QSGHLAR(서열번호 764) F4:NRHDRAK(서열번호 765) F5:RSDHLSE(서열번호 766), F6:QRRSRYK(서열번호 767);

9) F1:QSSDLSR(서열번호 768) F2:HRSTRNR(서열번호 769) F3:RSDVLSA(서열번호 770) F4:DSRTRKN(서열번호 771) F5:QSGSLTR(서열번호 772) F6:DQSGLAH(서열번호 773);

10) F1:QNPAQWR(서열번호 774) F2:RSADLSR(서열번호 775) F3:TSGSLSR(서열번호 776) F4:RSDHLSR(서열번호 777) F5:RSDSLLR(서열번호 778) F6:QSYDRFQ(서열번호 779);

11) F1:TSGSLSR(서열번호 780) F2:RSDHLSR(서열번호 781) F3:RSDSLLR(서열번호 782) F4:QSYDRFQ(서열번호 783) F5:RSDNLST(서열번호 784) F6:DNRDRIK(서열번호 785);

12) F1:DRSNLSR(서열번호 786) F2:LRQNLIM(서열번호 787) F3:ERGTLAR(서열번호 788) F4:RSDALTQ(서열번호 789) F5:RSDSLSQ(서열번호 790) F6:RKADRTR(서열번호 791);

13) F1:QYCCLTN(서열번호 792) F2:TSGNLTR(서열번호 793) F3:QSSDLSR(서열번호 794) F4:FRYYLKR(서열번호 795) F5:QSGDLTR(서열번호 796) F6:DKGNLTK(서열번호 797);

14) F1:TSGSLSR(서열번호 798) F2:RSDNLTT(서열번호 799) F3:QSGNLAR(서열번호 800) F4:DRTTLMR(서열번호 801) F5:QSGHLAR(서열번호 802) F6:QLTHLNS(서열번호 803);

15) F1:IKHDLHR(서열번호 804) F2:RSANLTR(서열번호 805) F3:RSDNLAR(서열번호 806) F4:QNVSRPR(서열번호 807) F5:RSDDLSK(서열번호 808) F6:DSSHRTR(서열번호 809);

16) F1:RSDNLAR(서열번호 810) F2:QNVSRPR(서열번호 811) F3:RSDDLSK(서열번호 812) F4:DSSHRTR(서열번호 813) F5:TSSNRKT(서열번호 814) F6:AQWTRAC(서열번호 815);

17) F1:RSDDLSK(서열번호 816) F2:DSSHRTR(서열번호 817) F3:TSSNRKT(서열번호 818) F4:AQWTRAC(서열번호 819) F5:RKQTRTT(서열번호 820) F6:HRSSLRR(서열번호 821);

18) F1:QSAHRKN(서열번호 822) F2:TSSNRKT(서열번호 823) F3:RSDNLSA(서열번호 824) F4:RNNDRKT(서열번호 825) F5:TSGSLSR(서열번호 826) F6:QAGHLAK(서열번호 827);

19) F1:RSDHLSQ(서열번호 828) F2:ASSTRTK(서열번호 829) F3:RSDDLTR(서열번호 830) F4:QKSNLSS(서열번호 831) F5:QSANRTT(서열번호 832) F6:QNATRTK(서열번호 833);

20) F1:RSDTLSE(서열번호 834) F2:RRWTLVG(서열번호 835) F3:DRSNLSR(서열번호 836) F4:QSGDLTR(서열번호 837) F5:QSSDLSR(서열번호 838) F6:YHWYLKK(서열번호 839);

21) F1:RSANLAR(서열번호 840) F2:RSDNLRE(서열번호 841) F3:RPYTLRL(서열번호 842) F4:HRSNLNK(서열번호 843) F5:QSGSLTR(서열번호 844) F6:TSANLSR(서열번호 845);

22) F1:RSDDLVR(서열번호 846) F2:TSGSLVR(서열번호 847) F3:RSDKLVR(서열번호 848) F4:RSDELVR(서열번호 849) F5:TSHSLTE(서열번호 850) F6:RADNLTE(서열번호 851);

23) F1:ERSHLRE(서열번호 852) F2:TSHSLTE(서열번호 853) F3:QAGHLAS(서열번호 854) F4:TSHSLTE(서열번호 855) F5:DPGHLVR(서열번호 856) F6:TSGNLVR(서열번호 857);

24) F1:RADNLTE(서열번호 858) F2:TSGSLVR(서열번호 859) F3:RKDNLKN(서열번호 860) F4:QSSSLVR(서열번호 861) F5:RSDKLVR(서열번호 862) F6:DSGNLRV(서열번호 863);

25) F1:QSSSLVR(서열번호 864) F2:QSGDLRR(서열번호 865) F3:RSDERKR(서열번호 866) F4:HRTTLTN(서열번호 867) F5:RSDHLTN(서열번호 868) F6:TSGELVR(서열번호 869);

26) F1:QSGDLRR(서열번호 870) F2:RSDERKR(서열번호 871) F3:HRTTLTN(서열번호 872) F4:RSDHLTN(서열번호 873) F5:TSGELVR(서열번호 874) F6:RSDDLVR(서열번호 875);

27) F1:QRAHLER(서열번호 876) F2:QLAHLRA(서열번호 877) F3:DPGHLVR(서열번호 878) F4:RRSACRR(서열번호 879) F5:RSDHLTT(서열번호 880) F6:QSSSLVR(서열번호 881); 및

28) F1:QSSNLVR(서열번호 882) F2:RSDDLVR(서열번호 883) F3:THLDLIR(서열번호 884) F4:TSGNLTE(서열번호 885) F5:RRSACRR(서열번호 886) F6:RNDTLTE(서열번호 887).

청구항 111

청구항 102-110 중 어느 한 항에 있어서, 조작된 아연 집게 단백질은 서열번호 692-719 중 어느 하나에 제시된 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

청구항 112

청구항 102-111 중 어느 한 항에 있어서, 조작된 아연 집게 단백질은 서열번호 888-915 중 어느 하나에 제시된 서열 또는 이의 일부에 의해 인코딩되거나, 또는 상기 서열과 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩되는, 후성적 변형 DNA-표적화 시스템.

청구항 113

후성적 변형 DNA-표적화 시스템으로서, 상기 후성적 변형 DNA-표적화 시스템은 a) 하나 이상의 HBV 유전자 또는 이의 조절 요소의 표적 부위에 결합하는 조작된 아연 집게 단백질, 및 b) 하나 이상의 HBV 유전자의 전사를 억제하는 적어도 하나의 효과기 도메인을 포함하고, 여기서 아연 집게 단백질은 N-말단에서 C-말단 순서로 F1에서 F6까지 표시된 6개의 아연 집게를 포함하며,

여기서 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같은 것인 후성적 변형 DNA-표적화 시스템:

F1:QSAHRKN(서열번호 822) F2:TSSNRKT(서열번호 823) F3:RSDNLSA(서열번호 824) F4:RNNDRKT(서열번호 825) F5:TSGSLSR(서열번호 826) F6:QAGHLAK(서열번호 827).

청구항 114

청구항 102-113 중 어느 한 항에 있어서, eZFP는 서열번호 709에 제시된 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

청구항 115

청구항 102-114 중 어느 한 항에 있어서, 조작된 아연 집게 단백질은 서열번호 709 중 어느 하나에 제시된 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

청구항 116

청구항 102-115 중 어느 한 항에 있어서, 조작된 아연 집게 단백질은 서열번호 905에 제시된 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩되는, 후성적 변형 DNA-표적화 시스템.

청구항 117

청구항 102-116 중 어느 한 항에 있어서, 조작된 아연 집게 단백질은 서열번호 905 중 어느 하나에 제시된 서열에 의해 인코딩되는, 후성적 변형 DNA-표적화 시스템.

청구항 118

후성적 변형 DNA-표적화 시스템으로서, 상기 후성적 변형 DNA-표적화 시스템은 a) 하나 이상의 HBV 유전자 또는 이의 조절 요소의 표적 부위에 결합하는 조작된 아연 집게 단백질, 및 b) 하나 이상의 HBV 유전자의 전사를 억제하는 적어도 하나의 효과기 도메인을 포함하고,

여기서 아연 집게 단백질은 N-말단에서 C-말단의 순서로 F1에서 F6까지 표시된 6개의 아연 집게를 포함하며, 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같은 것인, 후성적 변형 DNA-표적화 시스템:

F1:RSDHLSQ(서열번호 828) F2:ASSTRTK(서열번호 829) F3:RSDDLTR(서열번호 830) F4:QKSNLSS(서열번호 831) F5:QSANRTT(서열번호 832) F6:QNATRTK(서열번호 833).

청구항 119

청구항 102-112 및 118 중 어느 한 항에 있어서, 조작된 아연 집게 단백질은 서열번호 710에 제시된 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

청구항 120

청구항 102-112, 118 또는 119 중 어느 한 항에 있어서, 조작된 아연 집게 단백질은 서열번호 710 중 어느 하나에 제시된 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

청구항 121

청구항 102-112 및 118-120 중 어느 한 항에 있어서, 조작된 아연 집게 단백질은 서열번호 906에 제시된 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩되는, 후성적 변형 DNA-표적화 시스템.

청구항 122

청구항 102-112 및 118-121 중 어느 한 항에 있어서, 조작된 아연 집게 단백질은 서열번호 906 중 어느 하나에 제시된 서열에 의해 인코딩되는, 후성적 변형 DNA-표적화 시스템.

청구항 123

후성적 변형 DNA-표적화 시스템으로서, 상기 후성적 변형 DNA-표적화 시스템은 a) 하나 이상의 HBV 유전자 또는 이의 조절 요소의 표적 부위에 결합하는 조작된 아연 집게 단백질, 및 b) 하나 이상의 HBV 유전자의 전사를 억제하는 적어도 하나의 효과기 도메인을 포함하고,

여기서 아연 집게 단백질은 N 말단에서 C 말단까지 순서대로 F1에서 F6까지 표시되는 6개의 아연 집게를 포함하며,

여기서 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같은 것인, 후성적 변형 DNA-표적화 시스템:

F1:QSSSLVR(서열번호 864) F2:QSGDLRR(서열번호 865) F3:RSDERKR(서열번호 866) F4:HRTTLTN(서열번호 867) F5:RSDHLTN(서열번호 868) F6:TSGELVR(서열번호 869).

청구항 124

청구항 102-112 및 123 중 어느 한 항에 있어서, 조작된 아연 집게 단백질은 서열번호 716에 제시된 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

청구항 125

청구항 102-112, 123 또는 124 중 어느 한 항에 있어서, 상기 조작된 아연 집게 단백질은 서열번호 716 중 어느 하나에 제시된 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

청구항 126

청구항 102-112 및 123-125 중 어느 한 항에 있어서, 조작된 아연 집게 단백질은 서열번호 912에 제시된 서열 또는 이의 일부에 의해 인코딩되거나, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는

99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩되는, 후성적 변형 DNA-표적화 시스템.

청구항 127

청구항 102-112 및 123-126 중 어느 한 항에 있어서, 조작된 아연 집게 단백질은 서열번호 912 중 어느 하나에 제시된 서열에 의해 인코딩되는, 후성적 변형 DNA-표적화 시스템.

청구항 128

청구항 1-127 중 어느 한 항에 있어서, 적어도 하나의 효과기 도메인은 전사 억제를 유도하는, 후성적 변형 DNA-표적화 시스템.

청구항 129

청구항 1-128 중 어느 한 항에 있어서, 적어도 하나의 효과기 도메인은 DNA 메틸트랜스퍼라제인, 후성적 변형 DNA-표적화 시스템.

청구항 130

청구항 1-129 중 어느 한 항에 있어서, 상기 적어도 하나의 효과기 도메인은 DNA 메틸트랜스퍼라제와 헤테로크로마틴 유도 인자를 동원할 수 있는 억제자 도메인을 포함하거나, 선택적으로 헤테로크로마틴 유도 인자는 히스톤 메틸트랜스퍼라제를 포함하는, 후성적 변형 DNA-표적화 시스템.

청구항 131

청구항 1-130 중 어느 한 항에 있어서, 적어도 하나의 효과기 도메인은 DNA 메틸트랜스퍼라제와 히스톤 메틸트랜스퍼라제를 포함하는, 후성적 변형 DNA-표적화 시스템.

청구항 132

청구항 1-131 중 어느 한 항에 있어서, 적어도 하나의 효과기 도메인은 KRAB 억제자 도메인, ERF 억제자 도메인, Mxi1 억제자 도메인, SID4X 억제자 도메인, Mad-SID 억제자 도메인, LSD1 억제자 도메인, 또는 DNMT3A, DNMT3A-3L, DNMT3A/L-KRAB 융합 억제자 도메인, DNMT3B 도메인 결합 단백질, EZH2 억제자 도메인, 또는 LSD1 억제자 도메인, 또는 상기 중 임의의 것의 변이체 중에서 선택되는, 후성적 변형 DNA-표적화 시스템.

청구항 133

청구항 1-132 중 어느 한 항에 있어서, 적어도 하나의 효과기 도메인은 서열번호 590, 또는 600-608, 651, 661, 664, 665, 666, 668 및 669 중 어느 하나로부터 선택된 서열 또는 이의 도메인, 이의 일부 또는 상기 중 어느 하나와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

청구항 134

청구항 1-133 중 어느 한 항에 있어서, 적어도 하나의 효과기 도메인은 KRAB 도메인 또는 그 변이체를 포함하는, 후성적 변형 DNA-표적화 시스템.

청구항 135

청구항 1-134 중 어느 한 항에 있어서, 적어도 하나의 효과기 도메인은 서열번호 590에 제시된 서열, 이의 일부 또는 상기 중 어느 것과 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

청구항 136

청구항 1-134 중 어느 한 항에 있어서, 적어도 하나의 효과기 도메인은 DNMT3A/L 도메인 또는 이의 변이체를 포함하는, 후성적 변형 DNA-표적화 시스템.

청구항 137

청구항 1-136 중 어느 한 항에 있어서;

적어도 하나의 효과기 도메인은 서열번호 604 및 607에 제시된 서열, 이의 일부 또는 앞서 언급한 서열 중 어느 하나와 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 이상의 서열 동일성을 갖는 아미노산 서열을 포함하거나; 또는

적어도 하나의 효과기 도메인은 서열번호 651에 제시된 서열, 이의 일부, 또는 서열번호 651과 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

청구항 138

청구항 1-137 중 어느 한 항에 있어서, 융합 단백질은 DNMT3A/3L-dCas9-KRAB 융합 단백질을 포함하는, 후성적 변형 DNA-표적화 시스템.

청구항 139

청구항 1-138 중 어느 한 항에 있어서, 융합 단백질은 서열번호 645에 제시된 서열, 이의 일부 또는 앞서 언급한 서열과 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 이상의 서열 동일성을 갖는 아미노산 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

청구항 140

청구항 1-139 중 어느 한 항에 있어서, 적어도 하나의 효과기 도메인은 DNA-결합 도메인 또는 이의 구성 요소의 N-말단, C-말단, 또는 N-말단과 C-말단 둘 다에 융합되는, 후성적 변형 DNA-표적화 시스템.

청구항 141

청구항 1-101, 또는 138-140 중 어느 한 항에 있어서, 융합 단백질은 서열번호 680에 제시된 서열, 이의 일부, 또는 앞서 언급한 서열과 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99% 이상의 서열 동일성을 갖는 아미노산 서열에 의해 인코딩되는, 후성적 변형 DNA-표적화 시스템.

청구항 142

청구항 1-101, 또는 138-140 중 어느 한 항에 있어서, 융합 단백질은 서열번호 680에 제시된 서열에 의해 인코딩되는, 후성적 변형 DNA-표적화 시스템.

청구항 143

청구항 1-5, 11-18, 20-33, 34-61 및 102-140 중 어느 한 항에 있어서, 융합 단백질은 서열번호 916-943 중 어느 하나에 제시된 서열, 이의 일부 또는 앞서 언급한 서열과 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 이상의 서열 동일성을 갖는 핵산 서열에 의해 인코딩되는, 후성적 변형 DNA-표적화 시스템.

청구항 144

청구항 1-5, 11-18, 20-33, 34-61, 102-140 및 143 중 어느 한 항에 있어서, 융합 단백질은 서열번호 944-971 중 어느 하나에 제시된 서열, 이의 일부 또는 앞서 언급한 서열과 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 이상의 서열 동일성을 갖는 아미노산 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

청구항 145

청구항 1-5, 11-18, 20-33, 34-61, 102-140 및 143-144 중 어느 한 항에 있어서, 융합 단백질은 서열번호 961, 962 또는 968 중 어느 하나에 제시된 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

청구항 146

청구항 1-5, 11-18, 20-33, 34-61, 102-140 및 143-145 중 어느 한 항에 있어서, 융합 단백질은 DNMT3A/3L-eZFP-KRAB 융합 단백질을 포함하는, 후성적 변형 DNA-표적화 시스템.

청구항 147

청구항 1-4, 11-18, 20-33, 34-61, 102-140 및 143-146 중 어느 한 항에 있어서, 융합 단백질은 서열번호 972-999 중 어느 하나에 제시된 서열, 이의 일부 또는 앞서 언급한 서열과 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 이상의 서열 동일성을 갖는 핵산 서열에 의해 인코딩되는, 후성적 변형 DNA-표적화 시스템.

청구항 148

청구항 1-5, 11-18, 20-33, 34-61, 102-140 및 143-147 중 어느 한 항에 있어서, 융합 단백질은 서열번호 933, 934 또는 940 중 어느 하나에 제시된 서열, 이의 일부 또는 앞서 언급한 서열과 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 이상의 서열 동일성을 갖는 핵산 서열에 의해 인코딩되는, 후성적 변형 DNA-표적화 시스템.

청구항 149

청구항 1-5, 11-18, 20-33, 34-61, 102-140 및 143-148 중 어느 한 항에 있어서, 융합 단백질은 서열번호 1000-1027 중 어느 하나에 제시된 서열, 이의 일부 또는 앞서 언급한 서열과 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 이상의 서열 동일성을 갖는 아미노산 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

청구항 150

청구항 1-4, 11-18, 20-33, 34-61, 102-140 및 143-149 중 어느 한 항에 있어서, 융합 단백질은 서열번호 1017, 1018 또는 1024 중 어느 하나에 제시된 서열, 이의 일부 또는 앞서 언급한 서열과 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 이상의 서열 동일성을 갖는 아미노산 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

청구항 151

청구항 1-150 중 어느 한 항에 있어서, 융합 단백질은 추가로 하나 이상의 핵 국소화 신호(NLS)를 포함하는, 후성적 변형 DNA-표적화 시스템.

청구항 152

청구항 1-151 중 어느 한 항에 있어서, 융합 단백질은 다음 중 두 개 이상을 연결하는 하나 이상의 링커를 더 포함하는, 후성적 변형 DNA-표적화 시스템: DNA-결합 도메인, 적어도 하나의 효과기 도메인 및 하나 이상의 핵 국소화 신호.

청구항 153

청구항 1-152 중 어느 한 항에 있어서, DNA-표적화 시스템은 모든 B형 간염 바이러스 게놈을 표적화하는, 후성적 변형 DNA-표적화 시스템.

청구항 154

청구항 1-153 중 어느 한 항에 있어서, DNA-표적화 시스템은 모든 B형 간염 바이러스 게놈의 최소 70%를 표적화하는, 후성적 변형 DNA-표적화 시스템.

청구항 155

청구항 1-154 중 어느 한 항에 있어서, DNA-표적화 시스템은 모든 B형 간염 바이러스 게놈의 최소 60%를 표적화하는, 후성적 변형 DNA-표적화 시스템.

청구항 156

청구항 1-155 중 어느 한 항에 있어서, DNA-표적화 시스템은 모든 B형 간염 바이러스 게놈의 최소 50%를 표적화하는, 후성적 변형 DNA-표적화 시스템.

청구항 157

청구항 1-156 중 어느 한 항에 있어서, DNA-표적화 시스템은 표적 부위 또는 그 근처에서 유전적 파괴 또는 DNA

파손을 도입할 수 없는, 후성적 변형 DNA-표적화 시스템

청구항 158

청구항 1-157 중 어느 한 항에 있어서, 하나 이상의 HBV 유전자의 전사 억제가 HBV DNA 서열의 RNA 수준 및/또는 단백질 수준의 감소를 초래하는, 후성적 변형 DNA-표적화 시스템.

청구항 159

청구항 1-158 중 어느 한 항에 있어서, 전사 억제는 총 B형 간염 바이러스 RNA 전사 수치의 감소를 포함하는, 후성적 변형 DNA-표적화 시스템.

청구항 160

청구항 1-159 중 어느 한 항에 있어서, 전사 억제는 B형 간염 pre-코어("preC"), pre-게놈("pgRNA"), preS1, preS2/S 및 HBx 수준의 감소를 포함하는, 후성적 변형 DNA-표적화 시스템.

청구항 161

청구항 1-160 중 어느 한 항에 있어서, 전사 억제는 HBx 수치의 감소를 포함하는, 후성적 변형 DNA-표적화 시스템.

청구항 162

청구항 1-161 중 어느 한 항에 있어서, 전사 억제는 B형 간염 표면 항원(HBsAg) 및/또는 B형 간염 바이러스 코어 관련 항원(HbcrAg) 단백질 수치의 감소를 포함하는, 후성적 변형 DNA-표적화 시스템.

청구항 163

청구항 1-162 중 어느 한 항에 있어서, 전사 억제는 HbsAg 전사체 및/또는 단백질 수치를 최소 90% 감소시키는 것을 포함하는, 후성적 변형 DNA-표적화 시스템.

청구항 164

청구항 1-163 중 어느 한 항에 있어서, 전사 억제는 cccDNA에서 HbcrAg 전사체 및/또는 단백질 수준이 최소 50% 감소하는 것을 포함하는, 후성적 변형 DNA-표적화 시스템.

청구항 165

B형 간염 바이러스 DNA 서열의 표적 부위에 결합하는 가이드 RNA(gRNA).

청구항 166

청구항 165에 있어서, B형 간염 바이러스 DNA 서열은 B형 간염(HBV) 유전자 또는 이의 조절 요소인, gRNA.

청구항 167

청구항 165 또는 청구항 166에 있어서, 표적 부위는 공유적으로 폐쇄된 원형, DNA(cccDNA) 형태, 이완된 원형 DNA(rcDNA) 형태로 존재하고/존재하거나 인간 게놈 DNA에 통합된, gRNA.

청구항 168

청구항 165-167 중 어느 한 항에 있어서, 표적 부위는 HBV 복제 및/또는 HBV 전사를 제어하는 데 관여하는 유전자 또는 이의 조절 요소에 있거나 그 근처에 있는, gRNA.

청구항 169

청구항 168에 있어서, HBV 복제 및/또는 HBV 전사를 제어하는데 관여하는 유전자는 폴리머라제, 외피 단백질, 캡시드 단백질, 전사 인자 또는 전사 전사 활성화제를 인코딩하는, gRNA.

청구항 170

청구항 168 또는 청구항 169에 있어서, HBV 복제 및/또는 HBV 전사를 제어하는 데 관여하는 유전자는 폴리머라제 유전자, S-패밀리 유전자, X-유전자 또는 코어-패밀리 유전자인, gRNA.

청구항 171

청구항 165-170 중 어느 한 항에 있어서, 표적 부위는 B형 간염 바이러스 단백질 X(HBx)를 인코딩하는 X 유전자의 유전자 또는 이의 조절 요소에 있는, gRNA.

청구항 172

청구항 165-171 중 어느 한 항에 있어서, 표적 부위는 HBV 복제 및/또는 HBV 전사를 제어하는 조절 요소에 있거나 그 근처에 있는, gRNA.

청구항 173

청구항 166-172 중 어느 한 항에 있어서, 조절 요소는 프로모터 영역인, gRNA.

청구항 174

청구항 173에 있어서, 프로모터 영역은 pre-S1 프로모터, pre-S2 프로모터, X 프로모터 또는 기저 코어 프로모터인, gRNA.

청구항 175

청구항 166-172 중 어느 한 항에 있어서, 조절 요소는 인핸서 영역인, gRNA.

청구항 176

청구항 175에 있어서, 인핸서 영역은 Enh1 또는 Enh2 인핸서 영역인, gRNA.

청구항 177

청구항 166-172 중 어느 한 항에 있어서, 조절 요소는 전사 처리 제어 영역인, gRNA

청구항 178

청구항 165-171 중 어느 한 항에 있어서, 표적 부위는 코딩 영역인, gRNA.

청구항 179

청구항 165-178 중 어느 한 항에 있어서, 표적 부위는 전사 개시 부위의 500bp 이내, 1000bp 이내, 1500bp 이내에 위치하는, gRNA.

청구항 180

청구항 165-179 중 어느 한 항에 있어서, 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 HBV 게놈의 0-3300 염기쌍(bp) 사이, 선택적으로 HBV 게놈의 0-3189 bp 사이에 위치한 염기쌍에 상응하는 서열을 갖는 표적 영역 내에 위치하는, gRNA.

청구항 181

청구항 165-180 중 어느 한 항에 있어서, 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 HBV 게놈의 43bp-490bp, 1033bp-1749bp, 1800bp-1950bp 또는 2953bp-3182bp 사이의 염기쌍에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내에 위치하는, gRNA.

청구항 182

청구항 165-181 중 어느 한 항에 있어서, 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 HBV 게놈의 1 bp-42 bp, 491 bp-1032 bp, 1750 bp-1799 bp 또는 1951 bp-2952 bp 사이의 염기쌍에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내에 위치하는, gRNA.

청구항 183

청구항 165-182 중 어느 한 항에 있어서, 표적 부위는 HBV 게놈의 CpG 섬에 있는, gRNA.

청구항 184

청구항 165-183 중 어느 한 항에 있어서, 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 HBV 게놈의 67bp-392bp, 1033bp-1749bp 또는 2215bp-2490bp 사이의 염기쌍에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내에 위치하는, gRNA.

청구항 185

청구항 165-184 중 어느 한 항에 있어서, 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 HBV 게놈의 1033bp-1749bp 사이의 염기쌍에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내에 위치하는, gRNA.

청구항 186

청구항 165-185 중 어느 한 항에 있어서, 표적 부위는 B형 간염 X 단백질(HBx) 시작 코돈 상류의 300개 염기쌍 내에 걸쳐 있는 표적 영역 내에 있는, gRNA.

청구항 187

B형 간염 X 단백질(HBx) 시작 코돈 상류의 300개 염기쌍 내에 있는 표적 영역 내의 표적 부위에 결합하는 gRNA(gRNA).

청구항 188

청구항 165-187 중 어느 한 항에 있어서, 표적 부위는 HBx 기저 코어 프로모터 영역에 위치하는, gRNA.

청구항 189

청구항 165-187 중 어느 한 항에 있어서, 표적 부위는 HBx 프로모터/인핸서 영역 내에 위치하는, gRNA.

청구항 190

청구항 165-189 중 어느 한 항에 있어서, 표적 부위는 B형 간염 X 단백질(HBx) 시작 코돈 상류의 250개 염기쌍 내에 걸쳐 있는 표적 영역 내에 있는, gRNA.

청구항 191

청구항 165-190 중 어느 한 항에 있어서, 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 HBV 게놈의 1060-1480bp 사이의 염기쌍에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내에 있는, gRNA.

청구항 192

청구항 165-191 중 어느 한 항에 있어서, 표적 부위는 B형 간염 X 단백질(HBx) 시작 코돈 상류의 150개 염기쌍 내에 걸쳐 있는 표적 영역 내에 있는, gRNA.

청구항 193

청구항 165-192 중 어느 한 항에 있어서, 표적 부위는 B형 간염 X 단백질(HBx) 시작 코돈 상류의 120개 염기쌍 내에 걸쳐 있는 표적 영역 내에 있는, gRNA.

청구항 194

청구항 165-193 중 어느 한 항에 있어서, 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 HBV 게놈의 1250-1374bp 사이의 염기쌍에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내에 있는, gRNA.

청구항 195

청구항 165-194 중 어느 한 항에 있어서, 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 HBV 게놈의

1255-1302bp 사이의 염기쌍에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내에 있는, gRNA.

청구항 196

청구항 165-195 중 어느 한 항에 있어서, 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 HBV 게놈의 1260-1300bp 사이의 염기쌍에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내에 있는, gRNA.

청구항 197

청구항 165-196 중 어느 한 항에 있어서, gRNA는 서열번호 200, 201, 207, 217, 221, 224, 233, 237, 238, 246, 251, 256, 258, 263, 267, 274, 270, 277, 279, 283, 284, 293, 294, 308, 311, 313, 316, 319, 320, 325, 328, 330, 333, 338, 345, 347, 350, 353, 359, 360, 369, 370, 371, 377, 380, 384, 385 또는 387 중 어느 하나에 제시된 서열, 적어도 14개 뉴클레오타이드의 이의 연속 부분 또는 상기 중 어느 하나의 상보적 서열을 포함하고, 선택적으로 여기서 gRNA 또는 각각의 gRNA는 서열번호 395, 402, 408, 412, 416, 419, 428, 432, 433, 441, 446, 451, 453, 458, 462, 465, 469, 472, 474, 478, 479, 488, 489, 503, 506, 508, 511, 514, 515, 520, 523, 525, 575, 528, 533, 540, 542, 545, 548, 554, 555, 565, 566, 572, 579, 580, 또는 582 중 어느 하나에 제시된 것인, gRNA.

청구항 198

청구항 165-196 중 어느 한 항에 있어서, gRNA는 서열번호 395, 402, 408, 412, 416, 419, 428, 432, 433, 441, 446, 451, 453, 458, 462, 465, 469, 472, 474, 478, 479, 488, 489, 503, 506, 508, 511, 514, 515, 520, 523, 525, 575, 528, 533, 540, 542, 545, 548, 554, 555, 565, 566, 572, 579, 580, 또는 582 중 어느 하나에 제시된, gRNA.

청구항 199

청구항 165-198 중 어느 한 항에 있어서, gRNA는 서열번호 207, 213, 215, 217, 221, 222, 241, 245, 258, 261, 268, 274, 380, 387 중 어느 하나에 제시된 서열, 적어도 14개 뉴클레오타이드의 이의 연속 부분, 또는 상기 중 어느 하나의 상보적 서열을 포함하고, 선택적으로 gRNA 또는 각각의 gRNA는 서열번호 402, 408, 410, 412, 416, 417, 436, 440, 453, 456, 463, 469, 575 또는 582 중 어느 하나에 명시된 것인, gRNA.

청구항 200

청구항 165-199 중 어느 한 항에 있어서, gRNA는 서열번호 402, 408, 410, 412, 416, 417, 436, 440, 453, 456, 463, 469, 575 또는 582 중 어느 하나에 제시된 서열을 포함하는, gRNA.

청구항 201

청구항 165-200 중 어느 한 항에 있어서, gRNA는 서열번호 217에 제시된 서열, 적어도 14개 뉴클레오타이드의 이의 연속 부분, 또는 상기 중 어느 하나의 상보적 서열을 포함하고, 선택적으로 해당 gRNA 또는 각각의 gRNA는 서열번호 412 중 어느 하나에 제시된 것인, gRNA.

청구항 202

청구항 165-201 중 어느 한 항에 있어서, gRNA는 서열번호 412에 제시된 서열을 포함하는, gRNA.

청구항 203

다음을 포함하는 CRISPR Cas-가이드 RNA(gRNA) 조합:

(a) 규칙적인 간격을 갖고 나타나는 짧은 회문 구조의 반복 서열 관련(Cas: Clustered Regularly Interspaced Short Palindromic Repeats associated) 단백질 또는 이의 변이체; 및

(b) Cas 단백질 또는 이의 변이체를 B형 간염 바이러스 DNA 서열의 표적 부위에 표적화하는 청구항 165-202 중 어느 한 청구항의 적어도 하나의 gRNA.

청구항 204

청구항 203에 있어서, Cas 단백질 또는 이의 변이체는 Cas9 단백질 또는 이의 변이체인, CRISPR Cas-gRNA 조합.

청구항 205

청구항 203 또는 청구항 204에 있어서, Cas 단백질 또는 이의 변이체는 변이체 Cas 단백질이고, 여기서 변이체 Cas 단백질은 뉴클레아제 활성이 없거나 비활성화된 Cas(dCas) 단백질인, CRISPR Cas-gRNA 조합.

청구항 206

청구항 203-205 중 어느 한 항에 있어서, 변이체 Cas 단백질은 뉴클레아제 활성이 결여된 변이체 Cas9 단백질이거나 비활성화된 Cas9(dCas9) 단백질인, CRISPR Cas-gRNA 조합.

청구항 207

청구항 203-206 중 어느 한 항에 있어서, Cas9 단백질 또는 이의 변이체는 황색포도상구균(*Staphylococcus aureus*) Cas9(SaCas9) 단백질 또는 이의 변이체인, CRISPR Cas-gRNA 조합.

청구항 208

청구항 207에 있어서, 변이체 Cas9는 서열번호 596의 위치 번호를 참조로 D10A 및 N580A에서 선택된 적어도 하나의 아미노산 돌연변이를 포함하는 황색포도상구균(*Staphylococcus aureus*) dCas9 단백질(dSaCas9)인, CRISPR Cas-gRNA 조합.

청구항 209

청구항 207 또는 청구항 208에 있어서, 변이체 Cas9 단백질은 서열번호 597에 제시된 서열을 포함하거나, 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는, CRISPR Cas-gRNA 조합.

청구항 210

청구항 203-206 중 어느 한 항에 있어서, Cas9 단백질 또는 이의 변이체는 화농성연쇄상구균(*Streptococcus pyogenes*) Cas9(SpCas9) 단백질 또는 이의 변이체인, CRISPR Cas-gRNA 조합.

청구항 211

청구항 210에 있어서, 변이체 Cas9는 서열번호 598의 위치 번호를 참조하여 D10A 및 H840A 중에서 선택된 적어도 하나의 아미노산 돌연변이를 포함하는 화농성연쇄상구균(*Streptococcus pyogenes*) dCas9(dSpCas9) 단백질인, CRISPR Cas-gRNA 조합.

청구항 212

청구항 210 또는 청구항 211에 있어서, 변이체 Cas9 단백질은 서열번호 599에 제시된 서열, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는, CRISPR Cas-gRNA 조합.

청구항 213

청구항 1-164 중 어느 한 청구항의 후성적 변형 DNA-표적화 시스템, 또는 청구항 1-164 중 어느 한 청구항의 DNA-표적화 시스템의 융합 단백질, 청구항 165-202 중 어느 한 청구항의 gRNA, 청구항 203-212 중 어느 한 청구항의 CRISPR Cas-gRNA 조합, 또는 상기의 일부 또는 구성 요소를 인코딩하는, 폴리뉴클레오타이드.

청구항 214

청구항 1-164 중 어느 한 청구항의 후성적 변형 DNA-표적화 시스템, 또는 청구항 1-164 중 어느 한 청구항의 DNA-표적화 시스템의 융합 단백질, 청구항 165-202 중 어느 한 청구항의 gRNA, 청구항 203-212 중 어느 한 청구항의 CRISPR Cas-gRNA 조합, 또는 상기의 일부 또는 구성 요소를 인코딩하는, 복수의 폴리뉴클레오타이드.

청구항 215

청구항 213의 폴리뉴클레오타이드를 포함하는 벡터.

청구항 216

청구항 214의 복수의 폴리뉴클레오타이드를 포함하는 벡터.

청구항 217

청구항 215 또는 216에 있어서, 벡터는 바이러스 벡터인, 벡터.

청구항 218

청구항 217에 있어서, 벡터는 아데노 연관 바이러스(AAV) 벡터인, 벡터.

청구항 219

청구항 218에 있어서, 벡터는 AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8 또는 AAV9 중에서 선택되는, 벡터.

청구항 220

청구항 219에 있어서, 벡터는 렌티바이러스 벡터인, 벡터.

청구항 221

구항 215 또는 청구항 216에 있어서, 벡터는 비바이러스 벡터인, 벡터.

청구항 222

청구항 221에 있어서, 비바이러스 벡터는 지질 나노입자, 리포좀, 엑소좀 또는 세포 침투 펩타이드 중에서 선택되는, 벡터.

청구항 223

청구항 215-222 중 어느 어느 한 항에 있어서, 벡터는 B형 간염 바이러스에 감염된 세포에 대해 향성을 나타내며, 선택적으로 감염된 세포는 간세포인, 벡터.

청구항 224

청구항 215-223 중 어느 한 항에 있어서, 벡터는 하나의 벡터, 또는 두 개 이상의 벡터를 포함하는, 벡터.

청구항 225

B형 간염 바이러스 서열의 표적 영역 내에서 후성적 변형을 촉진하는 방법으로서, 상기 방법은 B형 간염 바이러스 서열을 포함하는 HBV 감염 세포 내로 표적 영역 내의 표적 부위를 표적으로 하는 후성적 변형 DNA-표적화 시스템을 도입하는 단계를 포함하는, 방법.

청구항 226

B형 간염 바이러스 서열 내 표적 영역 내 CpG 메틸화를 증가시키는 방법으로서, 상기 방법은 B형 간염 바이러스 서열을 포함하는 HBV 감염 세포 내로 표적 영역 내 표적 부위를 표적으로 하는 후성적 변형 DNA-표적화 시스템을 도입하는 단계를 포함하는, 방법.

청구항 227

B형 간염 바이러스 서열의 표적 영역의 후성적 변형을 촉진하는 방법으로서, 상기 방법은 B형 간염 바이러스 서열을 포함하는 HBV 감염 세포 내로 청구항 1-164 중 어느 한 청구항의 후성적 변형 DNA-표적 시스템, 청구항 165-202 중 어느 한 청구항의 gRNA, 청구항 203-212 중 어느 한 청구항의 CRISPR Cas-gRNA 조합, 청구항 213의 폴리뉴클레오타이드, 청구항 214의 복수의 폴리뉴클레오타이드, 청구항 215-224 중 어느 한 청구항의 벡터, 또는 상기의 일부 또는 구성요소를 도입하는 것을 포함하는, 방법.

청구항 228

B형 간염 바이러스 서열의 표적 영역의 CpG 메틸화를 증가시키는 방법으로서, B형 간염 바이러스 서열을 포함하는 HBV 감염 세포에 청구항 1-164 중 어느 하나의 후성적 변형 DNA-표적화 시스템, 청구항 165-202 중 어느 하나의 gRNA, 청구항 203-212 중 어느 하나의 CRISPR Cas-gRNA 조합, 청구항 213의 폴리뉴클레오타이드, 청구항 214의 복수의 폴리뉴클레오타이드, 청구항 215-224 중 어느 하나의 벡터, 또는 상기의 일부 또는 구성요소를 도입하는 단계를 포함하는, 방법.

청구항 229

청구항 225-228 중 어느 한 항에 있어서, 표적 영역은 서열번호 650의 뉴클레오타이드 위치를 참조로 B형 간염 바이러스 서열의 1033bp-1749bp 사이의 염기쌍에 위치한 서열에 상응하는 서열 내의 연속된 뉴클레오타이드 서열을 포함하는, 방법.

청구항 230

B형 간염 바이러스 서열을 포함하는 HBV에 감염된 세포에서 하나 이상의 유전자 전사를 감소시키는 방법으로서, 상기 방법은 서열번호 650의 뉴클레오타이드 위치를 참조로 B형 간염 바이러스 서열 내에서 표적 CpG 메틸화를 유도하는 후성적 변형 DNA-표적화 시스템을 세포에 도입하는 것을 포함하는, 방법.

청구항 231

HBV에 감염된 세포에서 B형 간염 바이러스 감염을 감소시키는 방법으로서, B형 간염 바이러스 서열을 포함하는 세포 내로 서열번호 650의 뉴클레오타이드 위치를 참조로 B형 간염 바이러스 서열의 표적 영역 내에서 표적 CpG 메틸화를 유도하는 후성적 변형 DNA-표적화 시스템을 도입하는 단계를 포함하는, 방법.

청구항 232

청구항 225-231 중 어느 한 항에 있어서, 후성적 변형 DNA-표적화 시스템은 (a) B형 간염 바이러스 DNA 서열의 표적 부위를 표적화하기 위한 DNA-결합 도메인 및 (b) DNA 메틸트랜스퍼라제 효과기 도메인을 포함하는 적어도 하나의 효과기 도메인을 포함하는 융합 단백질을 포함하는 적어도 하나의 DNA-표적화 모듈을 포함하는, 방법.

청구항 233

청구항 226 및 228-232 중 어느 한 항에 있어서, CpG 메틸화 영역은 표적 영역의 500개 염기쌍 이내에 있는, 방법.

청구항 234

청구항 225-233 중 어느 한 항에 있어서, 도입은 대상의 생체 내 또는 생체 외에서 이루어지는, 방법.

청구항 235

청구항 225-234 중 어느 한 항에 있어서, 세포는 포유류 세포인, 방법.

청구항 236

청구항 225-235 중 어느 한 항에 있어서, 세포는 인간 세포인, 방법.

청구항 237

청구항 225-236 중 어느 한 항에 있어서, 세포는 통합된 HBV DNA를 포함하는, 방법.

청구항 238

청구항 225-237 중 어느 한 항에 있어서, 세포는 에피솜 HBV cccDNA 풀을 포함하는 간세포인, 방법.

청구항 239

청구항 238에 있어서, 간세포는 HBV 단백질을 발현하고, 여기서 HBV 단백질은 HBsAg, HBeAg 또는 HBcrAg이거나, 또는 상기의 조합인, 방법.

청구항 240

대상에서 B형 간염 바이러스 감염을 감소시키는 방법으로서, 상기 방법은 B형 간염 바이러스에 감염된 대상에게 B형 간염 바이러스 서열의 표적 영역 내에서 CpG 메틸화를 증가시키는 후성적 변형 DNA-표적화 시스템을 투여하는 단계를 포함하고, 여기서 후성적 변형 DNA-표적화 시스템은 (a) B형 간염 바이러스 DNA 서열의 표적 부위를 표적으로 삼기 위한 DNA-결합 도메인; 및 (b) DNA 메틸트랜스퍼라제 효과기 도메인을 포함하는 적어도 하나의 효과기 도메인을 포함하는 것인, 방법.

청구항 241

청구항 225-240 중 어느 한 항에 있어서, 표적 영역은 HBV 게놈의 CpG를 포함하는 영역인, 방법.

청구항 242

청구항 225-241 중 어느 한 항에 있어서, 표적 영역은 서열번호 650의 뉴클레오타이드 위치를 참조로 B형 간염 바이러스 서열에서 67bp-392bp, 1033bp-1749bp 또는 2215bp-2490bp 사이의 염기쌍에 위치한 서열에 상응하는 뉴클레오타이드의 연속 서열을 포함하는, 방법.

청구항 243

청구항 225-242 중 어느 한 항에 있어서, 표적 영역은 서열번호 650의 뉴클레오타이드 위치를 참조로 B형 간염 바이러스 서열의 1033bp-1749bp 사이의 염기쌍에 위치한 서열에 상응하는 뉴클레오타이드의 연속된 서열을 포함하는, 방법.

청구항 244

청구항 225-244 중 어느 한 항에 있어서, 표적 영역은 B형 간염 X 단백질(HBx) 시작 코돈 상류의 300개 염기쌍 이내에 위치하는, 방법.

청구항 245

청구항 225-244 중 어느 한 항에 있어서, 표적 영역은 HBx 기저 코어 프로모터 영역 내에 있는, 방법.

청구항 246

청구항 225-245 중 어느 한 항에 있어서, 표적 영역은 HBx 프로모터/인핸서 영역 내에 있는, 방법.

청구항 247

청구항 225-246 중 어느 한 항에 있어서, 표적 영역은 B형 간염 X 단백질(HBx) 시작 코돈 상류의 250개 염기쌍 이내에 있는, 방법.

청구항 248

청구항 225-247 중 어느 한 항에 있어서, 표적 영역은 서열번호 650의 뉴클레오타이드 위치를 참조로 B형 간염 바이러스 서열의 1060bp-1480bp 염기쌍 사이에 위치한 서열에 상응하는 뉴클레오타이드의 연속 서열을 포함하는, 방법.

청구항 249

청구항 225-248 중 어느 한 항에 있어서, 표적 영역은 B형 간염 X 단백질(HBx) 시작 코돈 상류의 150개 염기쌍 이내에 있는, 방법.

청구항 250

청구항 225-249 중 어느 한 항에 있어서, 표적 영역은 BX형 간염 단백질(HBx) 시작 코돈 상류의 120개 염기쌍 이내에 있는, 방법.

청구항 251

청구항 225-250 중 어느 한 항에 있어서, 표적 영역은 서열번호 650의 뉴클레오타이드 위치를 참조로 B형 간염

바이러스 서열의 염기쌍 1250bp-1374bp 사이에 위치한 서열에 상응하는 뉴클레오타이드의 연속된 서열을 포함하는, 방법.

청구항 252

청구항 225-251 중 어느 한 항에 있어서, 표적 영역은 서열번호 650의 뉴클레오타이드 위치를 참조로 B형 간염 바이러스 서열의 염기쌍 1260bp-1300bp 사이에 위치한 서열에 상응하는 뉴클레오타이드의 연속 서열을 포함하는, 방법.

청구항 253

청구항 225-252 중 어느 한 항에 있어서, DNA-결합 도메인은 (a) 규칙적인 간격을 갖고 나타나는 짧은 회문 구조의 반복 서열 관련(Cas) 단백질 또는 이의 변이체 및 (b) 적어도 하나의 gRNA를 포함하는 Cas-가이드 RNA(gRNA) 조합; 아연 집게 단백질(ZFP); 전사 활성화 인자 유사 효과기(TALE); 메가뉴클레아제; 호밍 엔도뉴클레아제; 또는 I-SceI 효소 또는 이의 변이체를 포함하고, 선택적으로 DNA-결합 도메인은 상기 중 임의의 것의 촉매적으로 비활성인 변이체를 포함하는, 방법.

청구항 254

청구항 225-254 중 어느 한 항에 있어서, CRISPR Cas-가이드 RNA(gRNA) 조합은:

(a) 규칙적인 간격을 갖고 나타나는 짧은 회문 구조의 반복 서열 관련(Cas) 단백질 또는 이의 변이체; 및

(b) B형 간염 바이러스 DNA 서열의 표적 부위에 Cas 단백질 또는 이의 변이체를 표적화하는 청구항 165-202 중 어느 한 청구항의 적어도 하나의 gRNA

를 포함하는, 방법.

청구항 255

청구항 225-254 중 어느 한 항에 있어서, 표적 부위 또는 각각의 표적 부위는 서열번호 5, 6, 12, 18, 22, 26, 29, 38, 42, 43, 51, 56, 61, 63, 68, 72, 75, 79, 82, 84, 88, 89, 98, 99, 113, 116, 121, 124, 125, 118, 130, 133, 135, 138, 143, 150, 152, 155, 158, 164, 165, 175, 176, 182, 185, 189, 190, 192 중 어느 하나에 제시된 서열, 적어도 14개 뉴클레오타이드(nt)의 이의 연속 부분 또는 상기 중 어느 하나의 상보적 서열을 포함하는, 방법.

청구항 256

청구항 225-255 중 어느 한 항에 있어서, 표적 부위 또는 각각의 표적 부위는 서열번호 12, 18, 20, 22, 26, 27, 46, 50, 63, 66, 73, 79, 185, 192 중 어느 하나에 제시된 서열, 적어도 14개 뉴클레오타이드(nt)의 이의 연속 부분, 또는 상기 중 어느 하나의 상보적 서열을 포함하는, 방법.

청구항 257

청구항 225-256 중 어느 한 항에 있어서, 표적 부위 또는 각각의 표적 부위는 서열번호 22에 제시된 서열, 적어도 14개 뉴클레오타이드(nt)의 이의 연속 부분, 또는 앞서 언급한 서열의 상보적 서열을 포함하는, 방법.

청구항 258

청구항 225-257 중 어느 한 항에 있어서, gRNA 또는 각각의 gRNA는 서열번호 200, 201, 207, 217, 221, 224, 233, 237, 238, 246, 251, 256, 258, 263, 267, 274, 270, 277, 279, 283, 284, 293, 294, 308, 311, 313, 316, 319, 320, 325, 328, 330, 333, 338, 345, 347, 350, 353, 359, 360, 370, 371, 377, 380, 384, 385, 387 중 어느 하나에 제시된 서열을, 적어도 14개 뉴클레오타이드의 이의 연속 부분 또는 상기 중 어느 하나의 상보적 서열을 포함하고, 선택적으로 gRNA 또는 각각의 gRNA는 서열번호 395, 402, 408, 412, 416, 419, 428, 432, 433, 441, 446, 451, 453, 458, 462, 465, 469, 472, 474, 478, 479, 488, 489, 503, 506, 508, 511, 514, 515, 520, 523, 525, 575, 528, 533, 540, 542, 545, 548, 554, 555, 565, 566, 572, 579, 580, 또는 582 중 어느 하나에 제시된 것인, 방법.

청구항 259

청구항 225-258 중 어느 한 항에 있어서, gRNA 또는 각각의 gRNA는 서열번호 207, 213, 215, 217, 221, 222, 241, 245, 258, 261, 268, 274, 380, 387 중 어느 하나에 제시된 서열, 적어도 14개 뉴클레오타이드의 이의 연속 부분, 또는 상기 중 어느 하나의 상보적 서열을 포함하고, 선택적으로 gRNA 또는 각각의 gRNA는 서열번호 402, 408, 410, 412, 416, 417, 436, 440, 453, 456, 463, 469, 575, 582 중 어느 하나에 제시된 것인, 방법.

청구항 260

청구항 225-259 중 어느 한 항에 있어서, gRNA 또는 각각의 gRNA는 서열번호 217 중 어느 하나에 제시된 서열, 적어도 14개 뉴클레오타이드의 이의 연속 부분, 또는 상기 중 어느 하나의 상보적 서열을 포함하고, 선택적으로 gRNA 또는 각각의 gRNA는 서열번호 412 중 어느 하나에 제시된 것인, 방법.

청구항 261

청구항 225-253 중 어느 한 항에 있어서, 적어도 하나의 DNA-결합 도메인은 조작된 아연 집게 단백질(eZFP)을 포함하는, 방법.

청구항 262

청구항 225-253 또는 261 중 어느 한 항에 있어서, 표적 부위는 서열번호 1045, 1046, 1052 중 어느 하나에 제시된 뉴클레오타이드 서열, 최소 12nt의 이의 연속 부분, 또는 앞서 언급한 것의 상보적 서열을 포함하는, 방법.

청구항 263

청구항 225-253, 또는 261-262 중 어느 한 항에 있어서, 표적 부위는 서열번호 1045, 1046, 1052 중 어느 하나에 제시된 뉴클레오타이드 서열을 포함하는, 방법.

청구항 264

청구항 225-263 중 어느 한 항에 있어서, 적어도 하나의 효과기 도메인은 DNA 메틸트랜스퍼라제인, 방법.

청구항 265

청구항 225-264 중 어느 한 항에 있어서, 적어도 하나의 효과기 도메인은 DNA 메틸트랜스퍼라제와 헤테로크로마틴 유도 인자를 동원할 수 있는 억제자 도메인을 포함하거나, 선택적으로 헤테로크로마틴 유도 인자는 히스톤 메틸트랜스퍼라제를 포함하는, 방법.

청구항 266

청구항 225-265 중 어느 한 항에 있어서, 적어도 하나의 효과기 도메인은 DNA 메틸트랜스퍼라제와 히스톤 메틸트랜스퍼라제를 포함하는, 방법.

청구항 267

청구항 225-266 중 어느 한 항에 있어서, 적어도 하나의 효과기 도메인은 DNMT3A/L 도메인 또는 그 변이체를 포함하는, 방법.

청구항 268

청구항 225-267 중 어느 한 항에 있어서, 적어도 하나의 효과기 도메인은 서열번호 651에 제시된 서열, 이의 일부, 또는 상기 중 임의의 것과 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는, 방법.

청구항 269

청구항 225-268 중 어느 한 항에 있어서, 적어도 하나의 효과기 도메인은 KRAB 도메인 또는 이의 변이체를 더 포함하는, 방법.

청구항 270

청구항 225-269 중 어느 한 항에 있어서, 적어도 하나의 효과기 도메인은 서열번호 590에 제시된 서열, 이의 일부, 또는 상기 중 임의의 것과 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 더 포함하는, 방법.

청구항 271

청구항 225-270 중 어느 한 항에 있어서, DNA-표적화 시스템은 DNMT3A/3L-dCas9-KRAB 도메인 또는 그 변이체를 포함하는, 방법.

청구항 272

청구항 225-271 중 어느 한 항에 있어서, DNA-표적화 시스템은 서열번호 645에 제시된 서열, 이의 일부 또는 상기 중 어느 것과 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는, 방법.

청구항 273

청구항 225-260 또는 264-272 중 어느 한 항에 있어서, DNA-표적화 시스템은 서열번호 680에 제시된 서열, 이의 일부 또는 상기 중 어느 것과 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 핵산 서열을 포함하는, 방법.

청구항 274

청구항 225-260 또는 264-273 중 어느 한 항에 있어서, DNA-표적화 시스템은 서열번호 680에 제시된 서열을 포함하는, 방법.

청구항 275

B형 간염 바이러스에 감염된 세포에서 하나 이상의 유전자의 전사를 억제하는 방법으로서, 상기 방법은 B형 간염 바이러스에 감염된 세포에 청구항 1-164 중 어느 한 청구항의 후성적 변형 DNA-표적화 시스템, 청구항 165-202 중 어느 하나의 gRNA, 청구항 203-212 중 어느 하나의 CRISPR Cas-gRNA 조합, 청구항 213의 폴리뉴클레오타이드, 청구항 214의 복수의 폴리뉴클레오타이드, 청구항 215-224 중 어느 하나의 벡터, 또는 상기의 일부 또는 구성요소를 도입하는 단계를 포함하는, 방법.

청구항 276

청구항 225-275 중 어느 한 항에 있어서, 하나 이상의 유전자는 DNA-표적화 시스템에 의해 후성적으로 변형되는. 방법.

청구항 277

청구항 225-275 중 어느 한 항에 있어서, 하나 이상의 유전자의 전사는 이 방법을 거치지 않은 비교 가능한 세포와 비교시 감소되는, 방법.

청구항 278

청구항 277에 있어서, 하나 이상의 유전자의 전사는 적어도 약 1.25배, 1.5배, 1.75배, 2.0배, 2.5배, 2.75배, 3.0배, 3.5배, 3.75배, 4.0배, 4.5배, 4.75배, 5.0배, 5.25배, 5.5배, 5.75배, 6배만큼 감소되는, 방법.

청구항 279

청구항 225-278 중 어느 한 항에 있어서, 하나 이상의 유전자의 전사의 억제는 HBV 복제 및/또는 HBV 전사의 감소를 초래하는, 방법.

청구항 280

청구항 225-279 중 어느 한 항에 있어서, HBV에 감염된 세포는 포유류 세포인, 방법.

청구항 281

청구항 225-280 중 어느 한 항에 있어서, HBV에 감염된 세포는 인간 세포인, 방법.

청구항 282

청구항 225-281 중 어느 한 항에 있어서, 세포는 통합된 HBV DNA를 포함하는, 방법.

청구항 283

청구항 225-282 중 어느 한 항에 있어서, 세포는 에피솜 HBV cccDNA 풀을 포함하는 간세포인, 방법.

청구항 284

청구항 225-283 중 어느 한 항에 있어서, 간세포는 HBV 단백질을 발현하고, 여기서 HBV 단백질은 HBsAg 및/또는 HBeAg인, 방법.

청구항 285

청구항 225-284 중 어느 한 항에 있어서, HBV에 감염된 세포는 대상에 존재하는, 방법.

청구항 286

청구항 225-285 중 어느 한 항에 있어서, 대상은 인간인, 방법.

청구항 287

청구항 225-286 중 어느 한 항에 있어서, 대상은 HBV 바이러스에 감염된, 방법.

청구항 288

청구항 225-288 중 어느 한 항에 있어서, 대상은 통합된 HBV DNA를 포함하는 간세포를 갖는, 방법.

청구항 289

청구항 225-288 중 어느 한 항에 있어서, 대상은 에피솜 HBV cccDNA 풀을 포함하는 간세포를 갖는, 방법.

청구항 290

청구항 225-289 중 어느 한 항에 있어서, 대상은 HBV 단백질을 발현하는 간세포를 갖고, 여기서 HBV 단백질은 HBsAg, HBeAg 또는 HBcrAg 및 이들의 조합인, 방법.

청구항 291

청구항 225-290 중 어느 한 항에 있어서, 대상은 HBV 바이러스 감염과 관련된 질병, 병태 또는 장애를 갖는, 방법.

청구항 292

청구항 291에 있어서, 질병, 병태 또는 장애는 간 질환 또는 암인, 방법.

청구항 293

청구항 291 또는 청구항 292에 있어서, 질병, 병태 또는 장애는 급성 간염, 만성 간염, 간부전 또는 간경변인, 방법.

청구항 294

청구항 291-293 중 어느 한 항에 있어서, 질병, 병태 또는 장애는 암이고, 선택적으로 암은 간세포암인, 방법.

청구항 295

청구항 215-224 중 어느 한 청구항의 벡터를 포함하는 약학 조성물.

청구항 296

청구항 295에 있어서, 벡터는 갈락토스의 아미노 당 유도체에 접합되고, 선택적으로 벡터는 N-아세틸갈락토사민(GalNAc) 부분에 접합된 것인, 약학 조성물.

청구항 297

청구항 1-164 중 어느 하나의 후성적 변형 DNA-표적화 시스템 또는 청구항 1-164 중 어느 하나의 융합 단백질, 청구항 165-202 중 어느 하나의 gRNA, 청구항 203-212 중 어느 하나의 CRISPR Cas-gRNA 조합, 청구항 213의 폴리뉴클레오타이드, 청구항 214의 복수의 폴리뉴클레오타이드, 청구항 215-224 중 어느 하나의 벡터, 또는 상기의 일부 또는 구성요소를 포함하는 약학 조성물.

청구항 298

청구항 295-297 중 어느 한 항에 있어서, 대상의 HBV 바이러스 감염을 치료하는 데 사용하기 위한, 약학 조성물.

청구항 299

청구항 295-298 중 어느 한 항에 있어서, 대상은 HBV 바이러스 감염과 관련된 질병, 병태 또는 장애를 갖는, 약학 조성물.

청구항 300

청구항 295-299 중 어느 한 항에 있어서, 대상에서 HBV 바이러스 감염과 관련된 질병, 장애 또는 상태를 치료하는 데 사용하기 위한, 약학 조성물.

청구항 301

청구항 295-300 중 어느 한 항에 있어서, 대상에서 HBV 바이러스 감염을 치료하기 위한 약물 제조에 사용하기 위한, 약학 조성물.

청구항 302

청구항 298-301 중 어느 한 항에 있어서, HBV 바이러스 감염은 질병, 장애 또는 병태와 연관된, 약학 조성물.

청구항 303

청구항 295-300 중 어느 한 항에 있어서, HBV 바이러스 감염과 관련된 대상에서 질병, 병태 또는 장애를 치료하기 위한 약물 제조에 사용하기 위한, 약학 조성물.

청구항 304

청구항 299, 300, 302 또는 303 중 어느 한 항에 있어서, 질병, 병태 또는 장애는 간 질환 또는 암인, 약학 조성물.

청구항 305

청구항 299, 300, 302 또는 303 중 어느 한 항에 있어서, 질병, 병태 또는 장애는 급성 간염, 만성 간염, 간부전 또는 간경변인, 약학 조성물.

청구항 306

청구항 299, 300, 302 303 또는 304 중 어느 한 항에 있어서, 질병, 병태 또는 장애는 암, 선택적으로 간세포암인, 약학 조성물.

청구항 307

청구항 295-306 중 어느 한 항에 있어서, 약학 조성물은 대상에게 생체 내로 투여되는, 약학 조성물.

청구항 308

청구항 295-307 중 어느 한 항에 있어서, 약학 조성물의 투여 후, 대상의 세포에서 하나 이상의 HBV 유전자의

전사가 억제되는, 약학 조성물.

청구항 309

청구항 308에 있어서, 하나 이상의 HBV 유전자는 HBV 복제 및/또는 HBV 전사를 제어하는 데 관여하는, 약학 조성물.

청구항 310

청구항 308 또는 309에 있어서, 하나 이상의 유전자는 폴리머라제 유전자, S-패밀리 유전자, X-유전자 또는 코어 패밀리 유전자인, 약학 조성물.

청구항 311

질병, 병태 또는 장애의 치료를 필요로 하는 대상에서 이를 치료하기 위한 방법으로서, 청구항 1-164 중 어느 하나의 후성적 변형 DNA-표적화 시스템, 청구항 165-202 중 어느 하나의 gRNA, 청구항 203-212 중 어느 하나의 CRISPR Cas-gRNA 조합, 청구항 213의 폴리뉴클레오타이드, 청구항 214의 복수의 폴리뉴클레오타이드, 청구항 215-224 중 어느 하나의 벡터, 청구항 297-310 중 어느 하나의 약학 조성물, 또는 상기의 일부 또는 구성요소를 대상에게 투여하는 단계를 포함하는, 방법.

청구항 312

대상에서 B형 간염 바이러스 감염을 감소시키는 방법으로서, B형 간염 바이러스 감염이 있는 대상에게 청구항 1-164 중 어느 하나의 후성적 변형 DNA-표적화 시스템, 청구항 165-202 중 어느 하나의 gRNA, 청구항 203-212 중 어느 하나의 CRISPR Cas-gRNA 조합, 청구항 213의 폴리뉴클레오타이드, 청구항 214의 복수의 폴리뉴클레오타이드, 청구항 215-224 중 어느 하나의 벡터, 청구항 297-310 중 어느 하나의 약학 조성물, 또는 상기의 일부 또는 구성요소를 투여하는 단계를 포함하는, 방법

청구항 313

하나 이상의 HBV 유전자 또는 이의 조절 요소의 표적 부위에 결합하는 조작된 아연 집게 단백질(eZFP)로서, 여기서 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 HBV 게놈의 1033bp-1749bp 사이의 염기쌍에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내에 있는, 조작된 아연 집게 단백질(eZFP).

청구항 314

청구항 313에 있어서, 표적 부위는 B형 간염 X 단백질(HBx) 시작 코돈 상류의 300개 염기쌍 걸쳐 있는 표적 영역 내에 있는, eZFP.

청구항 315

청구항 313 또는 청구항 314에 있어서, 표적 부위는 HBx 기저 코어 프로모터 영역에 위치하는, eZFP.

청구항 316

청구항 313 또는 청구항 314에 있어서, 표적 부위는 HBx 프로모터/인핸서 영역 내에 위치하는, eZFP.

청구항 317

청구항 313-316 중 어느 한 항에 있어서, 표적 부위는 B형 간염 X 단백질(HBx) 시작 코돈 상류의 250개 염기쌍 내에 걸쳐 있는 표적 영역 내에 있는, eZFP.

청구항 318

청구항 313-317 중 어느 한 항에 있어서, 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 HBV 게놈의 1060-1480bp 사이의 염기쌍에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내에 있는, eZFP.

청구항 319

청구항 313-318 중 어느 한 항에 있어서, 표적 부위는 형 간염 X 단백질(HBx) 시작 코돈 상류의 150개 염기쌍

내에 걸쳐 있는 표적 영역 내에 있는, eZFP.

청구항 320

청구항 313-319 중 어느 한 항에 있어서, 표적 부위는 HBx(간염 BX) 단백질 시작 코돈 상류의 120개 염기쌍 내에 걸쳐 있는 표적 영역 내에 있는, eZFP.

청구항 321

청구항 313-320 중 어느 한 항에 있어서, 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 HBV 게놈의 1250-1374bp 사이의 염기쌍에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내에 있는, eZFP.

청구항 322

청구항 313-321 중 어느 한 항에 있어서, 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 HBV 게놈의 1255-1302bp 사이의 염기쌍에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내에 있는, eZFP.

청구항 323

청구항 313-322 중 어느 한 항에 있어서, 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 HBV 게놈의 1260-1300bp 사이의 염기쌍에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내에 있는, eZFP.

청구항 324

청구항 323에 있어서, 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 HBV 게놈의 1255bp-1290bp 사이의 염기쌍에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내에 있는, eZFP.

청구항 325

청구항 313-324 중 어느 한 항에 있어서, 표적 부위는 서열번호 1028-1055 중 어느 하나에 제시된 뉴클레오타이드 서열, 최소 12nt의 이의 연속 부분, 또는 앞서 언급한 것의 상보적 서열을 포함하는, eZFP.

청구항 326

청구항 313-325 중 어느 한 항에 있어서, 표적 부위는 서열번호 1028-1055 중 하나에 명시된 뉴클레오타이드 서열을 포함하는, eZFP.

청구항 327

청구항 313-326 중 어느 한 항에 있어서, 표적 부위는 서열번호 1045, 1046 또는 1052 중 어느 하나에 명시된 뉴클레오타이드 서열, 최소 12nt의 이의 연속 부분, 또는 앞서 언급한 것의 상보적 서열을 포함하는, eZFP.

청구항 328

청구항 313-326 중 어느 한 항에 있어서, 표적 부위는 서열번호 1045, 1046 또는 1052 중 하나에 제시된 뉴클레오타이드 서열을 포함하는, eZFP.

청구항 329

청구항 313-328 중 어느 한 항에 있어서, 아연 집게 단백질은 N-말단에서 C-말단 순서로 F1 내지 F6으로 표시되는 6개의 아연 집게를 포함하고,

여기서 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같은 것인, eZFP:

1) F1:SEADRSR(서열번호 720) F2:DRSNLTR(서열번호 721)F3:QSSDLSR(서열번호 722) F4:YHWYLKK(서열번호 723) F5:RSDSLSV(서열번호 724) F6:QNAANRKT(서열번호 725);

2) F1:RSDVLST(서열번호 726) F2:DNSSRTR(서열번호 727) F3:RPYTLRL(서열번호 728) F4:DSSHRTR(서열번호 729) F5:RSDHLSQ(서열번호 730) F6:DSSHRTR(서열번호 731);

3) F1:RSDHLSQ(서열번호 732) F2:QSADRTK(서열번호 733) F3:RSDHLSQ(서열번호 734) F4:RRSDLKR(서열번호 735) F5:RSDHLSR(서열번호 736) F6:QSSDLRR(서열번호 737);

4) F1:RSDNLSE(서열번호 738) F2:TSSNRKT(서열번호 739) F3:DRSHLTR(서열번호 740) F4:RSDALTQ(서열번호 741) F5:DRSALAR(서열번호 742) F6:RRFTLSK(서열번호 743);

5) F1:RSDHLSE(서열번호 744) F2:QYSGRYY(서열번호 745) F3:HGQTLNE(서열번호 746) F4:QSGNLAR(서열번호 747) F5:RSDSLLR(서열번호 748)F6:CREYRGK(서열번호 749);

6) F1:QSANRTT(서열번호 750) F2:RSANLTR(서열번호 751) F3:RSDVLSE(서열번호 752) F4:TSGHLSR(서열번호 753) F5:QSSDLSR(서열번호 754), F6:QWSTRKR(서열번호 755);

7) F1:QSGNLAR(서열번호 756) F2:ATCCLAH(서열번호 757) F3:RWQYLPT(서열번호 758) F4:DRSALAR(서열번호 759) F5:RSDNLSE(서열번호 760)F6:KRCNLRC(서열번호 761);

8) F1:NPANLTR(서열번호 762) F2:QNATRTK(서열번호 763) F3:QSGHLAR(서열번호 764) F4:NRHDRAK(서열번호 765) F5:RSDHLSE(서열번호 766), F6:QRRSRYK(서열번호 767);

9) F1:QSSDLSR(서열번호 768) F2:HRSTRNR(서열번호 769) F3:RSDVLSA(서열번호 770) F4:DSRTRKN(서열번호 771) F5:QSGSLTR(서열번호 772) F6:DQSGLAH(서열번호 773);

10) F1:QNPAQWR(서열번호 774) F2:RSADLSR(서열번호 775) F3:TSGSLSR(서열번호 776) F4:RSDHLSR(서열번호 777) F5:RSDSLLR(서열번호 778) F6:QSYDRFQ(서열번호 779);

11) F1:TSGSLSR(서열번호 780) F2:RSDHLSR(서열번호 781) F3:RSDSLLR(서열번호 782) F4:QSYDRFQ(서열번호 783) F5:RSDNLST(서열번호 784) F6:DNRDRIK(서열번호 785);

12) F1:DRSNLSR(서열번호 786) F2:LRQNLIM(서열번호 787) F3:ERGTLAR(서열번호 788) F4:RSDALTQ(서열번호 789) F5:RSDSLSQ(서열번호 790) F6:RKADRTR(서열번호 791);

13) F1:QYCCLTN(서열번호 792) F2:TSGNLTR(서열번호 793) F3:QSSDLSR(서열번호 794) F4:FRYYLKR(서열번호 795) F5:QSGDLTR(서열번호 796) F6:DKGNLTK(서열번호 797);

14) F1:TSGSLSR(서열번호 798) F2:RSDNLTT(서열번호 799) F3:QSGNLAR(서열번호 800) F4:DRTTLMR(서열번호 801) F5:QSGHLAR(서열번호 802) F6:QLTHLNS(서열번호 803);

15) F1:IKHDLHR(서열번호 804) F2:RSANLTR(서열번호 805) F3:RSDNLAR(서열번호 806) F4:QNVSRPR(서열번호 807) F5:RSDDLSK(서열번호 808) F6:DSSHRTR(서열번호 809);

16) F1:RSDNLAR(서열번호 810) F2:QNVSRPR(서열번호 811) F3:RSDDLSK(서열번호 812) F4:DSSHRTR(서열번호 813) F5:TSSNRKT(서열번호 814) F6:AQWTRAC(서열번호 815);

17) F1:RSDDLSK(서열번호 816) F2:DSSHRTR(서열번호 817) F3:TSSNRKT(서열번호 818) F4:AQWTRAC(서열번호 819) F5:RKQTRTT(서열번호 820) F6:HRSSLRR(서열번호 821);

18) F1:QSAHRKN(서열번호 822) F2:TSSNRKT(서열번호 823) F3:RSDNLSA(서열번호 824) F4:RNNDRKT(서열번호 825) F5:TSGSLSR(서열번호 826) F6:QAGHLAK(서열번호 827);

19) F1:RSDHLSQ(서열번호 828) F2:ASSTRTK(서열번호 829) F3:RSDDLTR(서열번호 830) F4:QKSNLSS(서열번호 831) F5:QSANRTT(서열번호 832) F6:QNATRTK(서열번호 833);

20) F1:RSDTLSE(서열번호 834) F2:RRWTLVG(서열번호 835) F3:DRSNLSR(서열번호 836) F4:QSGDLTR(서열번호 837) F5:QSSDLSR(서열번호 838) F6:YHWYLKK(서열번호 839);

21) F1:RSANLAR(서열번호 840) F2:RSDNLRE(서열번호 841) F3:RPYTLRL(서열번호 842) F4:HRSNLNK(서열번호 843) F5:QSGSLTR(서열번호 844) F6:TSANLSR(서열번호 845);

22) F1:RSDDLVR(서열번호 846) F2:TSGSLVR(서열번호 847) F3:RSDKLVR(서열번호 848) F4:RSDELVR(서열번호 849) F5:TSHSLTE(서열번호 850) F6:RADNLTE(서열번호 851);

23) F1:ERSHLRE(서열번호 852) F2:TSHSLTE(서열번호 853) F3:QAGHLAS(서열번호 854) F4:TSHSLTE(서열번호 855) F5:DPGHLVR(서열번호 856) F6:TSGNLVR(서열번호 857);

24) F1:RADNLTE(서열번호 858) F2:TSGSLVR(서열번호 859) F3:RKDNLKN(서열번호 860) F4:QSSSLVR(서열번호 861) F5:RSDKLVR(서열번호 862) F6:DSGNLRV(서열번호 863);

25) F1:QSSSLVR(서열번호 864) F2:QSGDLRR(서열번호 865) F3:RSDERKR(서열번호 866) F4:HRTTLTN(서열번호 867) F5:RSDHLTN(서열번호 868) F6:TSGELVR(서열번호 869);

26) F1:QSGDLRR(서열번호 870) F2:RSDERKR(서열번호 871) F3:HRTTLTN(서열번호 872) F4:RSDHLTN(서열번호 873) F5:TSGELVR(서열번호 874) F6:RSDDLVR(서열번호 875);

27) F1:QRAHLER(서열번호 876) F2:QLAHLRA(서열번호 877) F3:DPGHLVR(서열번호 878) F4:RRSACRR(서열번호 879) F5:RSDHLTT(서열번호 880) F6:QSSSLVR(서열번호 881); 및

28) F1:QSSNLVR(서열번호 882) F2:RSDDLVR(서열번호 883) F3:THLDLIR(서열번호 884) F4:TSGNLTE(서열번호 885) F5:RRSACRR(서열번호 886) F6:RNDTLTE(서열번호 887).

청구항 330

청구항 313-329 중 어느 한 항에 있어서, 아연 집게 단백질은 N-말단에서 C-말단 순서로 F1 내지 F6으로 표시되는 6개의 아연 집게를 포함하고,

여기서 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같은 것인, eZFP:

F1:QSAHRKN(서열번호 822) F2:TSSNRKT(서열번호 823) F3:RSDNLSA(서열번호 824) F4:RNNDRKT(서열번호 825) F5:TSGSLSR(서열번호 826) F6:QAGHLAK(서열번호 827).

청구항 331

청구항 313-329 중 어느 한 항에 있어서, 아연 집게 단백질은 N-말단에서 C-말단 순서로 F1 내지 F6으로 표시되는 6개의 아연 집게를 포함하고,

여기서 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같은 것인, eZFP:

F1:RSDHLSQ(서열번호 828) F2:ASSTRTK(서열번호 829) F3:RSDDLTR(서열번호 830) F4:QKSNLSS(서열번호 831) F5:QSANRTT(서열번호 832) F6:QNATRTK(서열번호 833).

청구항 332

청구항 313-329 중 어느 한 항에 있어서, 아연 집게 단백질은 N-말단에서 C-말단 순서로 F1 내지 F6으로 표시되는 6개의 아연 집게를 포함하고,

여기서 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같은 것인, eZFP:

F1:QSSSLVR(서열번호 864) F2:QSGDLRR(서열번호 865) F3:RSDERKR(서열번호 866) F4:HRTTLTN(서열번호 867) F5:RSDHLTN(서열번호 868) F6:TSGELVR(서열번호 869).

청구항 333

하나 이상의 HBV 유전자 또는 이의 조절 요소의 표적 부위에 결합하는 eZFP로서, 상기 아연 집게 단백질은 N-말단에서 C-말단 순서로 F1에서 F6까지 표시되는 6개의 아연 집게를 포함하고,

여기서 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같은 것인, eZFP:

1) F1:SEADRSR(서열번호 720) F2:DRSNLTR(서열번호 721)F3:QSSDLSR(서열번호 722) F4:YHWYLKK(서열번호 723) F5:RSDSLSV(서열번호 724) F6:QNAANRKT(서열번호 725);

2) F1:RSDVLST(서열번호 726) F2:DNSSRTR(서열번호 727) F3:RPYTLRL(서열번호 728) F4:DSSHRTR(서열번호 729) F5:RSDHLSQ(서열번호 730) F6:DSSHRTR(서열번호 731);

3) F1:RSDHLSQ(서열번호 732) F2:QSADRTK(서열번호 733) F3:RSDHLSQ(서열번호 734) F4:RRSDLKR(서열번호 735) F5:RSDHLSR(서열번호 736) F6:QSSDLRR(서열번호 737);

4) F1:RSDNLSE(서열번호 738) F2:TSSNRKT(서열번호 739) F3:DRSHLTR(서열번호 740) F4:RSDALTQ(서열번호 741) F5:DRSALAR(서열번호 742) F6:RRFTLSK(서열번호 743);

5) F1:RSDHLSE(서열번호 744) F2:QYSGRYY(서열번호 745) F3:HGQTLNE(서열번호 746) F4:QSGNLAR(서열번호 747) F5:RSDSLLR(서열번호 748)F6:CREYRGK(서열번호 749);

6) F1:QSANRTT(서열번호 750) F2:RSANLTR(서열번호 751) F3:RSDVLSE(서열번호 752) F4:TSGHLSR(서열번호 753) F5:QSSDLSR(서열번호 754), F6:QWSTRKR(서열번호 755);

7) F1:QSGNLAR(서열번호 756) F2:ATCCLAH(서열번호 757) F3:RWQYLPT(서열번호 758) F4:DRSALAR(서열번호 759) F5:RSDNLSE(서열번호 760)F6:KRCNLRC(서열번호 761);

8) F1:NPANLTR(서열번호 762) F2:QNATRTK(서열번호 763) F3:QSGHLAR(서열번호 764) F4:NRHDRAK(서열번호 765) F5:RSDHLSE(서열번호 766), F6:QRRSRYK(서열번호 767);

9) F1:QSSDLSR(서열번호 768) F2:HRSTRNR(서열번호 769) F3:RSDVLSA(서열번호 770) F4:DSRTRKN(서열번호 771) F5:QSGSLTR(서열번호 772) F6:DQSGLAH(서열번호 773);

10) F1:QNPAQWR(서열번호 774) F2:RSADLSR(서열번호 775) F3:TSGSLSR(서열번호 776) F4:RSDHLSR(서열번호 777) F5:RSDSLLR(서열번호 778) F6:QSYDRFQ(서열번호 779);

11) F1:TSGSLSR(서열번호 780) F2:RSDHLSR(서열번호 781) F3:RSDSLLR(서열번호 782) F4:QSYDRFQ(서열번호 783) F5:RSDNLST(서열번호 784) F6:DNRDRIK(서열번호 785);

12) F1:DRSNLSR(서열번호 786) F2:LRQNLIM(서열번호 787) F3:ERGTLAR(서열번호 788) F4:RSDALTQ(서열번호 789) F5:RSDSLSQ(서열번호 790) F6:RKADRTR(서열번호 791);

13) F1:QYCCLTN(서열번호 792) F2:TSGNLTR(서열번호 793) F3:QSSDLSR(서열번호 794) F4:FRYYLKR(서열번호 795) F5:QSGDLTR(서열번호 796) F6:DKGNLTK(서열번호 797);

14) F1:TSGSLSR(서열번호 798) F2:RSDNLTT(서열번호 799) F3:QSGNLAR(서열번호 800) F4:DRTTLMR(서열번호 801) F5:QSGHLAR(서열번호 802) F6:QLTHLNS(서열번호 803);

15) F1:IKHDLHR(서열번호 804) F2:RSANLTR(서열번호 805) F3:RSDNLAR(서열번호 806) F4:QNVSRPR(서열번호 807) F5:RSDDLSK(서열번호 808) F6:DSSHRTR(서열번호 809);

16) F1:RSDNLAR(서열번호 810) F2:QNVSRPR(서열번호 811) F3:RSDDLSK(서열번호 812) F4:DSSHRTR(서열번호 813) F5:TSSNRKT(서열번호 814) F6:AQWTRAC(서열번호 815);

17) F1:RSDDLSK(서열번호 816) F2:DSSHRTR(서열번호 817) F3:TSSNRKT(서열번호 818) F4:AQWTRAC(서열번호 819) F5:RKQTRTT(서열번호 820) F6:HRSSLRR(서열번호 821);

18) F1:QSAHRKN(서열번호 822) F2:TSSNRKT(서열번호 823) F3:RSDNLSA(서열번호 824) F4:RNNDRKT(서열번호 825) F5:TSGSLSR(서열번호 826) F6:QAGHLAK(서열번호 827);

19) F1:RSDHLSQ(서열번호 828) F2:ASSTRTK(서열번호 829) F3:RSDDLTR(서열번호 830) F4:QKSNLSS(서열번호 831) F5:QSANRTT(서열번호 832) F6:QNATRTK(서열번호 833);

20) F1:RSDTLSE(서열번호 834) F2:RRWTLVG(서열번호 835) F3:DRSNLSR(서열번호 836) F4:QSGDLTR(서열번호 837) F5:QSSDLSR(서열번호 838) F6:YHWYLKK(서열번호 839);

21) F1:RSANLAR(서열번호 840) F2:RSDNLRE(서열번호 841) F3:RPYTLRL(서열번호 842) F4:HRSNLNK(서열번호 843) F5:QSGSLTR(서열번호 844) F6:TSANLSR(서열번호 845);

22) F1:RSDDLVR(서열번호 846) F2:TSGSLVR(서열번호 847) F3:RSDKLVR(서열번호 848) F4:RSDELVR(서열번호 849) F5:TSHSLTE(서열번호 850) F6:RADNLTE(서열번호 851);

23) F1:ERSHLRE(서열번호 852) F2:TSHSLTE(서열번호 853) F3:QAGHLAS(서열번호 854) F4:TSHSLTE(서열번호 855) F5:DPGHLVR(서열번호 856) F6:TSGNLVR(서열번호 857);

24) F1:RADNLTE(서열번호 858) F2:TSGSLVR(서열번호 859) F3:RKDNLKN(서열번호 860) F4:QSSSLVR(서열번호 861) F5:RSDKLVR(서열번호 862) F6:DSGNLRV(서열번호 863);

25) F1:QSSSLVR(서열번호 864) F2:QSGDLRR(서열번호 865) F3:RSDERKR(서열번호 866) F4:HRTTLTN(서열번호 867) F5:RSDHLTN(서열번호 868) F6:TSGELVR(서열번호 869);

26) F1:QSGDLRR(서열번호 870) F2:RSDERKR(서열번호 871) F3:HRTTLTN(서열번호 872) F4:RSDHLTN(서열번호 873) F5:TSGELVR(서열번호 874) F6:RSDDLVR(서열번호 875);

27) F1:QRAHLER(서열번호 876) F2:QLAHLRA(서열번호 877) F3:DPGHLVR(서열번호 878) F4:RRSACRR(서열번호 879) F5:RSDHLTT(서열번호 880) F6:QSSSLVR(서열번호 881); 및

28) F1:QSSNLVR(서열번호 882) F2:RSDDLVR(서열번호 883) F3:THLDLIR(서열번호 884) F4:TSGNLTE(서열번호 885) F5:RRSACRR(서열번호 886) F6:RNDTLTE(서열번호 887).

청구항 334

청구항 313-333 중 어느 한 항에 있어서, 조작된 아연 집게 단백질은 서열번호 692-719 중 어느 하나에 제시된 서열 또는 이의 일부를 포함하거나, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는, eZFP.

청구항 335

청구항 313-334 중 어느 한 항에 있어서, 조작된 아연 집게 단백질은 서열번호 888-915 중 어느 하나에 제시된 서열 또는 이의 일부에 의해 인코딩되거나, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩되는, eZFP.

청구항 336

하나 이상의 HBV 유전자 또는 이의 조절 요소의 표적 부위에 결합하는 조작된 아연 집게 단백질로서, 여기서 아연 집게 단백질은 N-말단에서 C-말단 순서로 F1에서 F6까지 표시된 6개의 아연 집게를 포함하고,

여기서 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같은 것인, eZFP:

F1:QSAHRKN(서열번호 822) F2:TSSNRKT(서열번호 823) F3:RSDNLSA(서열번호 824) F4:RNNDRKT(서열번호 825) F5:TSGSLSR(서열번호 826) F6:QAGHLAK(서열번호 827).

청구항 337

청구항 313-336 중 어느 한 항에 있어서, 조작된 아연 집게 단백질은 서열번호 709에 제시된 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는, eZFP.

청구항 338

청구항 313-337 중 어느 한 항에 있어서, 조작된 아연 집게 단백질은 서열번호 709 중 어느 하나에 제시된 서열을 포함하는, eZFP.

청구항 339

청구항 313-338 중 어느 한 항에 있어서, 조작된 아연 집게 단백질은 서열번호 905에 제시된 서열 또는 이의 일부에 의해 인코딩되거나, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩되는, eZFP.

청구항 340

청구항 313-339 중 어느 한 항에 있어서, 조작된 아연 집게 단백질은 서열번호 905 중 어느 하나에 제시된 서열에 의해 인코딩되는, eZFP.

청구항 341

하나 이상의 HBV 유전자 또는 이의 조절 요소의 표적 부위에 결합하는 조작된 아연 집게 단백질로서, 상기 아연 집게 단백질은 N-말단에서 C-말단 순서로 F1에서 F6까지 표시된 6개의 아연 집게를 포함하고,

여기서 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같은 것인, eZFP:

F1:RSDHLSQ(서열번호 828) F2:ASSTRTK(서열번호 829) F3:RSDDLTR(서열번호 830) F4:QKSNLSS(서열번호 831) F5:QSANRTT(서열번호 832) F6:QNATRTK(서열번호 833).

청구항 342

청구항 313-335 및 341 중 어느 한 항에 있어서, 조작된 아연 집게 단백질은 서열번호 710에 제시된 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는, eZFP.

청구항 343

청구항 313-335, 341 및 342 중 어느 한 항에 있어서, 조작된 아연 집게 단백질은 서열번호 710 중 어느 하나에 제시된 서열을 포함하는, eZFP.

청구항 344

청구항 313-335 및 341-343 중 어느 한 항에 있어서, 조작된 아연 집게 단백질은 서열번호 906에 제시된 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩되는, eZFP.

청구항 345

청구항 313-335 및 341-344 중 어느 한 항에 있어서, 조작된 아연 집게 단백질은 서열번호 906 중 어느 하나에 제시된 서열에 의해 인코딩되는, eZFP.

청구항 346

하나 이상의 HBV 유전자 또는 이의 조절 요소의 표적 부위에 결합하는 조작된 아연 집게 단백질로서, 여기서 아연 집게 단백질은 N-말단에서 C-말단 순서로 F1에서 F6까지 표시된 6개의 아연 집게를 포함하고,

여기서 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같은 것인, eZFP:

F1:QSSSLVR(서열번호 864) F2:QSGDLRR(서열번호 865) F3:RSDERKR(서열번호 866) F4:HRTTLTN(서열번호 867) F5:RSDHLTN(서열번호 868) F6:TSGELVR(서열번호 869).

청구항 347

청구항 313-335 및 346 중 어느 한 항에 있어서, 조작된 아연 집게 단백질은 서열번호 716에 제시된 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는, eZFP.

청구항 348

청구항 313-335, 346 및 347 중 어느 한 항에 있어서, 조작된 아연 집게 단백질은 서열번호 716 중 어느 하나에 제시된 서열을 포함하는, eZFP.

청구항 349

청구항 313-335 및 346-348 중 어느 한 항에 있어서, 조작된 아연 집게 단백질은 서열번호 912에 제시된 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩되는, eZFP.

청구항 350

청구항 313-335 및 346-349 중 어느 한 항에 있어서, 조작된 아연 집게 단백질은 서열번호 912 중 어느 하나에 제시된 서열에 의해 인코딩되는, eZFP.

발명의 설명

기 술 분 야

[0001] **관련 출원에 대한 상호 참조**

[0002] [0001] 본 출원은 2022년 8월 19일에 출원된 미국 가출원 제63/399,634호("표적 유전자 억제를 통한 B형 간염 바이러스의 조절을 위한 조성물, 시스템 및 방법"), 2023년 6월 9일에 출원된 미국 가출원 제63/472,236호("표

적 유전자 억제를 통한 B형 간염 바이러스의 조절을 위한 조성물, 시스템 및 방법"), 및 2023년 8월 7일에 출원된 미국 가출원 제63/531,309호("표적 유전자 억제를 통한 B형 간염 바이러스의 조절을 위한 조성물, 시스템 및 방법")의 우선권을 주장한다. 내용은 전체적으로 참고로 포함되어 있다.

[0003] **서열목록의 참조에 의한 통합**

[0004] [0002] 본 출원은 전자 형식의 서열목록과 함께 제출된다. 서열목록은 2023년 8월 18일에 생성된 224742002040SeqList.xml이라는 제목의 파일로 제공되며, 크기는 1,595,298바이트이다. 서열목록의 전자 형식의 정보는 그 전부가 참조 통합된다.

[0005] 분야

[0006] [0003] 본 개시는 일부 측면에서 B형 간염 바이러스(HBV) 유전자의 전사 억제를 위한 CRISPR-Cas/가이드 RNA(gRNA) 시스템과 같은 후성적(epigenetic) 변형 DNA-표적화 시스템에 관한 것으로, 이는 HBV 감소로 이어지는 세포 표현형을 촉진한다. 일부 구현예에서, 후성적 변형 DNA-표적화 시스템은 세포 내 B형 간염 바이러스 DNA 서열에서 적어도 하나의 유전자 또는 이의 조절 요소의 표적 부위에 결합하거나 이를 표적화한다. 일부 구현예에서, 시스템은 적어도 두 개의 유전자 또는 이의 조절 요소의 표적 부위에 결합하거나 이를 표적화하는 다중 시스템이다. 일부 측면에서, 본 개시의 시스템은 하나 이상의 B형 간염 바이러스 유전자의 전사 억제와 관련된다. 일부 측면에서, 본 개시는 예를 들어 B형 간염 감염에 대한 치료와 관련하여 B형 간염 바이러스 복제 및 발현을 억제하는 등, 제공된 조성물과 관련된 방법 및 용도에 관한 것이다.

배 경 기 술

[0007] **배경**

[0008] [0004] 미국에서만 100만 명, 전 세계적으로 2억 5천만 명으로 추산되는 대규모 환자 집단이 만성 B형 간염 감염을 다룬다. 그러나 뉴클레오사이드 유사체, PEG화 인터페론, 안티센스 올리고뉴클레오타이드 및 siRNA 접근법을 투여하는 것과 같은 바이러스 DNA 전사 억제를 포함한 현재의 표준 치료는 효능 및 안정성 측면에서 어려움에 직면해 있다. 따라서 이러한 어려움을 극복하기 위한 새롭고 개선된 방법이 필요한다. 본 개시는 이러한 필요성과 기타 필요성을 다룬다.

발명의 내용

[0009] 개요

[0010] [0005] 본 발명에서는 하나 이상의 B형 간염 바이러스(HBV) 유전자의 전사를 억제하기 위한 적어도 하나의 DNA-표적화 모듈을 포함하는 후성적 변형 DNA-표적화 시스템이 제공되며, 여기서 적어도 하나의 DNA-표적화 모듈 각각은 다음을 포함하는 융합 단백질을 포함한다: (a) B형 간염 바이러스 DNA 서열의 표적 부위를 표적화하기 위한 DNA-결합 도메인; 및 (b) 하나 이상의 전사 억제자 효과기 도메인. 일부 구현예에서, 적어도 하나의 DNA-결합 도메인은 (a) Cas 단백질 또는 이의 변이체 및 (b) 하나 이상의 gRNA를 포함하는 Clustered Regularly Interspaced Short Palindromic Repeats associated(CAS)-가이드 RNA(gRNA); 아연 집게 단백질(ZFP); 전사 활성화제 유사 효과기(transcription activator-like effector: TALE); 메가뉴클레아제; 호밍 엔도뉴클레아제; 또는 I-SceI 효소 또는 이의 변이체를 포함하며, 선택적으로 여기서 DNA-결합 도메인은 상기의 임의의 촉매적으로 비활성인 변이체를 포함한다. 일부 구현예에서, B형 간염 바이러스 DNA 서열은 HBV 유전자 또는 이의 조절 요소이다. 일부 구현예에서, 적어도 하나의 DNA-표적화 모듈은 하나 또는 복수의 유전자 또는 이의 조절 요소의 복수의 표적 부위를 표적화하기 위한 복수의 DNA-표적화 모듈을 포함한다. 일부 구현예에서, 복수의 DNA-표적화 모듈은 적어도 하나의 제1 DNA-표적화 모듈과 제2 DNA-표적화 모듈을 포함하고, 여기서:(1) 제1 DNA-표적화 모듈은 제1 HBV 유전자의 전사를 억제하고, 여기서 제1 DNA-표적화 모듈은 (a) 제1 유전자 또는 이의 조절 DNA 요소의 표적 부위를 표적화하기 위한 DNA-결합 도메인; 및 (b) 적어도 하나의 전사 억제자 도메인을 포함하는 제1 융합 단백질을 포함하고; (2) 제2 DNA-표적화 모듈은 제2 HBV 유전자의 전사를 억제하고, 여기서 제2 DNA-표적화 모듈은 (a) 제2 유전자 또는 이의 조절 DNA 요소의 표적 부위를 표적화하기 위한 DNA-결합 도메인; 및 (b) 적어도 하나의 전사 억제 도메인을 포함하는 제2 융합 단백질을 포함하되, 선택적으로 여기서: 제1 DNA-표적화 모듈 및 제2 DNA-표적화 모듈은 동일한 융합 단백질을 공유하여 제1 및 제2 융합 단백질이 동일하고, 융합 단백질의 DNA-결합 도메인은 Cas(Clustered Regularly Interspaced Short Palindromic Repeats) 관련 단백질 또는 이의 변이체이며; 제1 DNA-표적화 모듈은 제1 HBV 유전자 또는 이의 조절 요소의 표적 부위를 표적화하는 제1 가이드 RNA(gRNA)를 포함하고, 제2 DNA-표적화 모듈은 제2 HBV 유전자 또는 이의 조절 요소의 표적 부위를 표적

화하는 제2 gRNA를 포함한다.

[0011] [0006] 또한, 본 발명에서는 하나 이상의 B형 간염 바이러스(HBV) 유전자의 전사를 억제하기 위한 후성적 변형 DNA-표적화 시스템이 제공되며, 여기서 DNA-표적화 시스템은 (a) Cas(Clustered Regularly Interspaced Short Palindromic Repeats) 관련 단백질 또는 이의 변이체와 적어도 하나의 전사 억제자 효과기 도메인을 포함하는 융합 단백질; 및 (b) 적어도 제1 gRNA와 제2 gRNA를 포함하는 복수의 가이드 RNA(gRNA)을 포함하되, 여기서 제1 gRNA는 제1 HBV 유전자 또는 이의 조절 요소의 표적 부위를 표적으로 하고, 제2 gRNA는 제2 HBV 유전자 또는 이의 조절 요소의 표적 부위를 표적으로 하며, 여기서 제1 및 제2 유전자 또는 이의 조절 요소는 B형 간염 바이러스 복제 및/또는 HBV 전사를 조절한다. 일부 구현예에서, DNA-표적화 시스템은 또한 B형 간염 바이러스 복제 및/또는 HBV 전사를 조절하는 제3 유전자 또는 이의 조절 요소의 표적 부위를 표적으로 하는 제3 gRNA를 포함한다. 일부 구현예에서, 시스템은 또한 제4 유전자 또는 이의 조절 요소의 표적 부위를 표적으로 하는 제4 gRNA, 선택적으로 제5 유전자 또는 이의 조절 요소를 표적으로 하는 제5 gRNA, 및/또는 선택적으로 제6 유전자 또는 이의 조절 요소의 표적 부위를 표적으로 하는 제6 gRNA를 포함하며, 여기서 유전자 또는 이의 조절 요소는 B형 간염 바이러스 복제 및/또는 HBV 전사를 조절한다. 일부 구현예에서, 제1, 제2, 제3, 제4, 제5 및/또는 제6 유전자 또는 이의 조절 요소는 상이하다.

[0012] [0007] 또한 본 발명은 하나 이상의 B형 간염 바이러스(HBV) 유전자의 전사를 억제하기 위한 후성적 변형 DNA-표적화 시스템을 제공하는데, 여기서 DNA-표적화 시스템은:

[0013] (a) Cas(Clustered Regularly Interspaced Short Palindromic Repeats) 관련 단백질 또는 이의 변이체와 적어도 하나의 전사 억제자 효과기 도메인을 포함하는 융합 단백질; 및 (b) 복수의 유전자 또는 이의 조절 요소의 복수의 표적 부위를 표적화하는 복수의 가이드 RNA(gRNA)를 포함하고, 여기서 복수의 유전자 또는 이의 조절 요소는 B형 간염 바이러스 복제 및/또는 HBV 전사를 조절한다.

[0014] [0008] 또한, 본 발명에서는 하나 초과의 B형 간염 바이러스(HBV) 유전자의 전사를 억제하기 위한 단일 DNA-표적화 모듈을 포함하는 후성적 변형 DNA-표적화 시스템이 제공되며, 여기서 DNA-표적화 모듈은: (a) Cas(Clustered Regularly Interspaced Short Palindromic Repeats) 관련 단백질 또는 이의 변이체와 적어도 하나의 전사 억제자 효과기 도메인을 포함하는 융합 단백질; 및 (b) 복수의 유전자 또는 이의 조절 요소의 복수의 표적 부위를 표적화하는 가이드 RNA(gRNA)를 포함하고, 여기서 복수의 유전자 또는 이의 조절 요소는 B형 간염 바이러스 복제 및/또는 HBV 전사를 조절한다.

[0015] [0009] 본원의 임의의 구현예에서, 전사 억제는 HBV 복제 감소 및/또는 HBV 단백질 수치 감소를 초래한다.

[0016] [0010] 본원의 임의의 구현예에서, DNA-표적화 시스템은 유전적 파괴 또는 DNA 절단을 유발하지 않는다.

[0017] [0011] 본원의 임의의 구현예에서, 적어도 하나의 DNA-결합 모듈은 HBV DNA 서열 내의 복수의 표적 부위를 함께 표적으로 하는 복수의 DNA-결합 모듈을 포함하며, 선택적으로 각 DNA-결합 모듈은 HBV DNA 서열 내의 다른 표적 부위를 표적으로 한다.

[0018] [0012] 본원의 임의의 구현예에서, 복수의 표적 부위는 2, 3, 4, 5 또는 6개의 다른 표적 부위이다. 본원의 임의의 구현예에서, 복수의 표적 부위는 각각 다른 HBV 유전자 또는 이의 조절 요소에 있다.

[0019] [0013] 본원의 임의의 구현예에서, 각 표적 부위는 동일한 HBV 유전자 또는 이의 조절 요소에 존재한다.

[0020] [0014] 본원의 임의의 구현예에서, 시스템은 2 내지 10개의 DNA-표적화 모듈을 포함한다.

[0021] [0015] 본원의 임의의 구현예에서, DNA-표적화 모듈 중 둘 이상은 동일한 융합 단백질을 공유하거나, DNA-표적화 모듈 중 둘 이상은 서로 다른 융합 단백질을 포함한다.

[0022] [0016] 본원의 임의의 구현예에서, 각 DNA-표적화 모듈의 DNA-결합 도메인은 Cas 단백질 또는 이의 변이체와 적어도 하나의 전사 억제자 효과기 도메인을 포함하는 융합 단백질을 포함하고, 각 DNA-표적화 모듈은 고유한 gRNA를 포함한다.

[0023] [0017] 본원의 임의의 구현예에서, 표적 부위 또는 각각의 표적 부위는 공유적으로 폐쇄된 원형 DNA(cccDNA) 형태, 이완된 원형 DNA(rcDNA) 형태 및/또는 인간 게놈 DNA에 통합된 HBV 바이러스 DNA로 존재한다. 본원의 임의의 구현예에서, 표적 부위 또는 각각의 표적 부위는 HBV 복제 및/또는 HBV 전사를 제어하는 데 관여하는 유전자 또는 이의 조절 요소에 또는 그 근처에 존재한다.

[0024] [0018] 본원의 임의의 구현예에서, HBV 복제 및/또는 HBV 전사를 제어하는 데 관여하는 유전자는 폴리머라제,

외피 단백질, 캡시드 단백질, 전사 인자 또는 전사 전사활성화제를 인코딩한다. 본원의 임의의 구현예에서, HBV 복제 및/또는 HBV 전사를 제어하는 데 관여하는 유전자는 폴리머라제 유전자, S-패밀리 유전자, X-유전자 또는 코어 패밀리 유전자이다.

[0025] [0019] 본원의 임의의 구현예에서, 적어도 하나의 표적 부위는 B형 간염 바이러스 단백질 X(HBx)를 인코딩하는 X 유전자의 유전자 또는 이의 조절 요소에 존재한다.

[0026] [0020] 본원의 임의의 구현예에서,표적 부위 또는 각각의 표적 부위는 HBV 복제 및/또는 HBV 전사를 제어하는 데 관여하는 HBV 유전자의 조절 요소에 있거나 근처에 있다. 일부 구현예에서, 조절 요소는 프로모터 영역이다. 일부 구현예에서, 프로모터 영역은 pre-S1 프로모터, pre-S2 프로모터, X 프로모터 또는 기본 코어 프로모터이다. 일부 구현예에서, 조절 요소는 인핸서 영역이다. 일부 구현예에서, 인핸서 영역은 Enh1 또는 Enh2 인핸서 영역이다. 일부 구현예에서, 조절 요소는 전사 처리 제어 영역이다.

[0027] [0021] 본원의 임의의 구현예에서, 표적 부위 또는 각각의 표적 부위는 HBV 유전자의 코딩 영역에 있다. 본원의 임의의 구현예에서, 표적 부위 또는 각각의 표적 부위는 전사 개시 부위의 500개 염기쌍(bp) 이내, 1000bp 이내, 1500bp 이내에 위치한다. 본원의 임의의 구현예에서, 표적 부위 또는 각각의 표적 부위는 서열번호 650에 명시된 HBV 게놈을 기준으로 HBV 게놈의 0-3300 염기쌍(bp) 사이, 선택적으로 0-3182 bp 사이의 염기쌍에 위치한 표적 영역 내에 위치한다. 본원의 임의의 구현예에서, 표적 부위 또는 각각의 표적 부위는 HBV 게놈의 43 bp - 490 bp, 1033 bp - 1749 bp, 1800 bp - 1950 bp, 또는 2953 bp - 3182 bp 사이의 염기쌍에 위치한 표적 영역 내에 위치하며, 이는 서열번호 650에 명시된 HBV 게놈을 기준으로 하는 위치에 상응하는 것이다. 본원의 임의의 구현예에서, 표적 부위 또는 각각의 표적 부위는 HBV 게놈의 1 bp - 42 bp, 491 bp - 1032 bp, 1750 bp - 1799 bp, 또는 1951 bp - 2952 bp 사이의 염기쌍에 위치한 표적 영역 내에 위치하며, 이는 서열번호 650에 명시된 HBV 게놈을 기준으로 하는 위치에 상응하는 것이다. 본원의 임의의 구현예에서, 표적 부위 또는 각각의 표적 부위는 HBV 게놈의 CpG 섬에 있다. 본원의 임의의 구현예에서, 표적 부위 또는 각각의 표적 부위는 HBV 게놈의 67bp-392bp, 1033bp-1749bp 또는 2215bp-2490bp 사이의 염기쌍에 위치한 표적 영역 내에 위치하며, 이는 서열번호 650에 명시된 HBV 게놈을 기준으로 하는 위치에 상응하는 것이다. 본원의 임의의 구현예에서, 표적 부위 또는 각각의 표적 부위는 서열번호 650의 뉴클레오타이드 위치를 기준으로 B형 간염 바이러스 서열의 1033bp-1749bp 사이의 염기쌍에 위치한다. 본원의 임의의 구현예에서, 표적 부위 또는 각각의 표적 부위는 B형 간염 X 단백질(HBx) 시작 코돈 상류의 300개 염기쌍 내에 위치한 표적 영역 내에 위치한다.

[0028] [0022] 또한, 본 발명에서는 하나 이상의 B형 간염 바이러스(HBV) 유전자의 전사를 억제하기 위한 DNA-표적화 모듈을 포함하는 후성적 변형 DNA-표적화 시스템이 제공되며, 여기서 적어도 하나의 DNA-표적화 모듈 각각은 다음을 포함하는 융합 단백질을 포함한다: (a) B형 간염 X 단백질(HBx) 시작 코돈 상류의 300개 염기쌍 내에 있는 표적 영역 내의 표적 부위를 표적으로 삼기 위한 DNA-결합 도메인; 및 (b) 하나 이상의 전사 억제자 효과기 도메인.

[0029] [0023] 본원의 임의의 구현예에서, 표적 부위 또는 각각의 표적 부위는 HBx 기본 코어 프로모터 영역에 위치한다. 본원의 임의의 구현예에서, 표적 부위 또는 각각의 표적 부위는 HBx 프로모터/인핸서 영역 내에 위치한다.

[0030] [0024] 본원의 임의의 구현예에서, 표적 부위 또는 각각의 표적 부위는 B형 간염 X 단백질(HBx) 시작 코돈 상류에서 250개 염기쌍 내에 걸쳐 있는 표적 영역 내에 있다. 본원의 임의의 구현예에서, 표적 부위 또는 각각의 표적 부위는 서열번호 650에 명시된 HBV 게놈을 참조로 HBV 게놈의 1060-1480bp 사이의 염기쌍에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내에 있다. 본원의 임의의 구현예에서, 표적 부위 또는 각각의 표적 부위는 B형 간염 X 단백질(HBx) 시작 코돈 상류에서 150개 염기쌍 내에 걸쳐 있는 표적 영역 내에 있다. 본원의 임의의 구현예에서, 표적 부위 또는 각각의 표적 부위는 B형 간염 X 단백질(HBx) 시작 코돈 상류에서 120개 염기쌍 내에 걸쳐 있는 표적 영역 내에 있다. 본원의 임의의 구현예에서, 표적 부위 또는 각각의 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 HBV 게놈의 1250-1374bp에 걸친 서열에 상응하는 표적 영역 서열 내에 있다. 본원의 임의의 구현예에서, 표적 부위 또는 각각의 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 HBV 게놈의 1255-1302bp에 걸친 서열에 상응하는 표적 영역 서열 내에 있다. 본원의 임의의 구현예에서, 표적 부위 또는 각각의 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 HBV 게놈의 1260-1300bp에 걸친 서열에 상응하는 표적 영역 서열 내에 있다.

[0031] [0025] 본원의 임의의 구현예에서, 표적 부위 또는 각각의 표적 부위는 모든 B형 간염 바이러스 게놈과 적어도 70% 상동적이다. 본원의 임의의 구현예에서, 표적 부위 또는 각각의 표적 부위는 적어도 1000개의 B형 간염 바이러스 게놈과 적어도 70% 상동적이다. 본원의 임의의 구현예에서, 표적 부위 또는 각각의 표적 부위는 적어도

1000개의 B형 간염 바이러스 게놈과 적어도 70% 상동하고 최대 2개의 불일치(mismatches)를 포함한다.

[0032] [0026] 본원의 임의의 구현예에서, 표적 부위 또는 각각의 표적 부위는 서열번호 1-195 중 어느 하나에 제시된 서열, 적어도 14개 뉴클레오타이드(nt)의 그의 연속(contiguous) 부분, 또는 상기 중 어느 하나의 상보적 서열을 포함한다. 본원의 임의의 구현예에서, 표적 부위 또는 각각의 표적 부위는 서열번호 175, 138, 192, 152, 118, 125, 185, 63, 116, 124, 35, 82 중 어느 하나에 제시된 서열, 적어도 14개 뉴클레오타이드(nt)의 그의 연속 부분, 또는 상기 중 어느 하나의 상보적 서열을 포함한다. 본원의 임의의 구현예에서, 표적 부위 또는 각각의 표적 부위는 서열번호 175, 138, 192, 152, 118, 125, 185, 63, 116, 124, 35, 82 중 어느 하나에 제시된 서열을 포함한다. 본원의 임의의 구현예에서, 표적 부위 또는 각각의 표적 부위는 서열번호 5, 6, 12, 18, 22, 26, 29, 38, 42, 43, 51, 56, 61, 63, 68, 72, 75, 79, 82, 84, 88, 89, 98, 99, 113, 116, 121, 124, 125, 118, 130, 133, 135, 138, 143, 150, 152, 155, 158, 164, 165, 175, 176, 182, 185, 189, 190, 192 중 어느 하나에 제시된 서열, 적어도 14개 뉴클레오타이드(nt)의 그의 연속 부분, 또는 상기 중 어느 하나의 상보적 서열을 포함한다. 본 발명의 임의의 구현예에서, 표적 부위 또는 각각의 표적 부위는 서열번호 5, 6, 12, 18, 22, 26, 29, 38, 42, 43, 51, 56, 61, 63, 68, 72, 75, 79, 82, 84, 88, 89, 98, 99, 113, 116, 121, 124, 125, 118, 130, 133, 135, 138, 143, 150, 152, 155, 158, 164, 165, 175, 176, 182, 185, 189, 190, 192 중 어느 하나에 제시된 서열을 포함한다. 본원의 임의의 구현예에서, 표적 부위 또는 각각의 표적 부위는 서열번호 12, 18, 20, 22, 26, 27, 46, 50, 63, 66, 73, 79, 185, 192 중 어느 하나, 적어도 14개 뉴클레오타이드의 그의 연속 부분, 또는 상기 중 어느 하나의 상보적 서열에 제시되어 있다. 본원의 임의의 구현예에서, 표적 부위 또는 각각의 표적 부위는 서열번호 12, 18, 20, 22, 26, 27, 46, 50, 63, 66, 73, 79, 185, 192 중 어느 하나에 제시된다. 본원의 임의의 구현예에서, 표적 부위 또는 각각의 표적 부위는 서열번호 22에 제시된 서열, 적어도 14개 뉴클레오타이드의 이의 연속 부분 또는 상기 중 어느 것의 상보적 서열을 포함하며, 선택적으로 표적 부위는 서열번호 22에 제시된다. 본원의 임의의 구현예에서, 표적 부위 또는 각각의 표적 부위는 서열번호 63에 제시된 서열, 적어도 14개 뉴클레오타이드의 이의 연속 부분 또는 상기 중 어느 것의 상보적 서열을 포함하며, 선택적으로 표적 부위는 서열번호 63에 제시된다.

[0033] [0027] 본원의 임의의 구현예에서, gRNA 또는 각각의 gRNA는 서열번호 196-390 중 어느 하나에 제시된 서열을 포함하는 gRNA 스페이서 서열을 포함한다. 본원의 임의의 구현예에서, gRNA 또는 각각의 gRNA는 추가로 서열번호 587에 제시된 서열을 포함한다. 본원의 임의의 구현예에서, gRNA 또는 각각의 gRNA는 서열번호 196-390 중 어느 하나에 제시된 서열을 포함한다. 본원의 임의의 구현예에서, gRNA 또는 각각의 gRNA는 서열번호 391-585 중 어느 하나에 제시된다. 본원의 임의의 구현예에서, gRNA 또는 각각의 gRNA는 서열번호 370, 333, 387, 347, 313, 320, 380, 256, 258, 311, 319, 230, 272 중 어느 하나에 제시된 서열, 적어도 14개 뉴클레오타이드의 그의 연속 부분, 또는 상기 중 어느 하나의 상보적 서열을 포함하며, 선택적으로 gRNA 또는 각각의 gRNA는 서열번호 565, 528, 542, 508, 515, 575, 515, 453, 506, 514, 425 또는 472 중 어느 하나에 제시된다. 본원의 임의의 구현예에서, gRNA 또는 각각의 gRNA는 서열번호 370, 333, 387, 347, 313, 320, 380, 256, 258, 311, 319, 230, 272 중 어느 하나에 제시된 서열을 포함하며, 선택적으로 gRNA 또는 각각의 gRNA는 서열번호 565, 528, 542, 508, 515, 575, 515, 453, 506, 514, 425 또는 472 중 어느 하나에 제시된다. 본 발명의 임의의 구현예에서, gRNA 또는 각각의 gRNA는 서열번호 200, 201, 207, 217, 221, 224, 233, 237, 238, 246, 251, 256, 258, 263, 267, 274, 270, 277, 279, 283, 284, 293, 294, 308, 311, 313, 316, 319, 320, 325, 328, 330, 333, 338, 345, 347, 350, 353, 359, 360, 370, 371, 377, 380, 384, 385, 387 중 어느 하나에 제시된 서열, 적어도 14개 뉴클레오타이드의 그의 연속 부분 또는 상기 중 어느 하나의 상보적 서열을 포함하며, 선택적으로 여기서 gRNA 또는 각각의 gRNA는 서열번호 369, 395, 402, 408, 412, 416, 419, 428, 432, 433, 441, 446, 451, 453, 458, 462, 465, 469, 472, 474, 478, 479, 488, 489, 503, 506, 508, 511, 514, 515, 520, 523, 525, 575, 528, 533, 540, 542, 545, 548, 554, 555, 565, 566, 572, 579, 580 또는 582 중 어느 하나에 제시된다. 본원의 임의의 구현예에서, gRNA 또는 각각의 gRNA는 서열번호 200, 201, 207, 217, 221, 224, 233, 237, 238, 246, 251, 256, 258, 263, 267, 274, 270, 277, 279, 283, 284, 293, 294, 308, 311, 313, 316, 319, 320, 325, 328, 330, 333, 338, 345, 347, 350, 353, 359, 360, 370, 371, 377, 380, 384, 385 또는 387 중 어느 하나에 제시된 서열을 포함하며, 선택적으로 gRNA 또는 각각의 gRNA는 서열번호 369, 395, 402, 408, 412, 416, 419, 428, 432, 433, 441, 446, 451, 453, 458, 462, 465, 469, 472, 474, 478, 479, 488, 489, 503, 506, 508, 511, 514, 515, 520, 523, 525, 575, 528, 533, 540, 542, 545, 548, 554, 555, 565, 566, 572, 579, 580 또는 582 중 어느 하나에 제시된다. 본원의 모든 구현예에서, gRNA 또는 각각의 gRNA는 서열번호 207, 213, 215, 217, 221, 222, 241, 245, 258, 261, 268, 274, 380, 387 중 어느 하나에 제시된 서열, 적어도 14개 뉴클레오타이드의 그의 연속 부분 또는 상기 중 어느 하나의 상보적 서열을 포함하며, 선택적으로 gRNA 또는 각

각의 gRNA는 서열번호 402, 408, 410, 412, 416, 417, 436, 440, 453, 456, 463, 469, 575, 582 중 어느 하나에 제시된다. 본원의 임의의 구현예에서, gRNA 또는 각각의 gRNA는 서열번호 207, 213, 215, 217, 221, 222, 241, 245, 258, 261, 268, 274, 380, 387 중 어느 하나에 제시된 서열을 포함한다. 본원의 임의의 구현예에서, gRNA 또는 각각의 gRNA는 서열번호 402, 408, 410, 412, 416, 417, 436, 440, 453, 456, 463, 469, 575, 582 중 어느 하나에 제시된다. 본 발명의 임의의 구현예에서, gRNA 또는 각각의 gRNA는 서열번호 217에 제시된 서열, 적어도 14개 뉴클레오타이드의 그의 연속된 부분, 또는 상기 중 어느 하나의 상보적 서열을 포함한다. 본 발명의 임의의 구현예에서, gRNA 또는 각각 의 gRNA는 서열번호 217에 제시된 서열을 포함한다. 본원의 임의의 구현예에서, gRNA 또는 각각의 gRNA는 서열번호 412 중 어느 하나에 제시된다. 본원의 임의의 구현예에서, gRNA 또는 각각의 gRNA는 서열번호 258에 제시된 서열, 적어도 14개 뉴클레오타이드의 그의 연속된 부분, 또는 상기 중 어느 하나의 상보적 서열을 포함한다. 본원의 임의의 구현예에서, gRNA 또는 각각의 gRNA는 서열번호 258에 제시된 서열을 포함한다. 본원의 임의의 구현예에서, gRNA 또는 각각의 gRNA는 서열번호 453 중 어느 하나에 제시된다.

[0034] [0028] 본원의 임의의 구현예에서, 표적 부위 또는 표적 부위 각각은 모든 B형 간염 바이러스 게놈과 90% 이상 상동적이다. 본원의 임의의 구현예에서, 표적 부위 또는 표적 부위 각각은 최소 1000개의 B형 간염 바이러스 게놈과 90% 이상 상동적이다. 본원의 임의의 구현예에서, 표적 부위 또는 표적 부위 각각은 최소 1000개의 B형 간염 바이러스 게놈과 90% 이상 상동적이며 최대 2개의 불일치, 선택적으로 1개 또는 2개의 불일치를 포함한다.

[0035] [0029] 본원의 임의의 구현예에서, 표적 부위 또는 각각의 표적 부위는 서열번호 35-100 중 어느 하나에 제시된 서열, 적어도 14개 뉴클레오타이드(nt)의 그의 연속 부분, 또는 앞서 언급한 서열의 상보적 서열을 포함한다.

[0036] [0030] 본원의 임의의 구현예에서, gRNA 또는 각각의 gRNA는 서열번호 230-295 중 어느 하나에 제시된 서열을 포함하는 gRNA 스페이서 서열을 포함한다. 본원의 임의의 구현예에서, gRNA 또는 각각의 gRNA는 추가로 서열번호 587에 제시된 서열을 포함한다. 본원의 임의의 구현예에서, gRNA 또는 각각의 gRNA는 서열번호 230-295 중 어느 하나에 제시된 서열을 포함하고, 선택적으로 gRNA 또는 각각의 gRNA는 서열번호 425-490 중 어느 하나에 제시된다.

[0037] [0031] 본원의 임의의 구현예에서, 최대 두 개의 불일치는 프로토스페이서의 5' 말단의 처음 12개 nt에 위치한다.

[0038] [0032] 본원의 임의의 구현예에서, 표적 부위 또는 각각의 표적 부위는 적어도 1000개의 B형 간염 바이러스 유전체와 적어도 90% 상동성을 갖고, 불일치 사항은 전혀 없다.

[0039] [0033] 본원의 임의의 구현예에서, 표적 부위는 서열번호 1-34 중 어느 하나에 제시된 서열, 적어도 14개 뉴클레오타이드(nt)의 그의 연속 부분, 또는 상기 중 어느 하나의 상보적 서열을 포함한다. 본원의 임의의 구현예에서, gRNA 또는 각각의 gRNA는 서열번호 196-229에 제시된 서열을 포함하는 gRNA 스페이서 서열을 포함한다.

[0040] [0034] 본원의 임의의 구현예에서, gRNA 스페이서 서열은 길이가 14 nt와 24 nt 사이 또는 16 nt와 22 nt 사이이다. 본원의 임의의 구현예에서, gRNA 스페이서 서열은 길이가 18 nt, 19 nt, 20 nt, 21 nt 또는 22 nt이다.

[0041] [0035] 본원의 임의의 구현예에서, gRNA 스페이서 서열은 안정성을 높이기 위해 변형된 뉴클레오타이드를 포함한다.

[0042] [0036] 본원의 임의의 구현예에서, 적어도 하나의 gRNA는 추가로 서열번호 587에 제시된 서열을 포함한다. 본원의 임의의 구현예에서, gRNA 또는 각각의 gRNA는 서열번호 196-229 중 어느 하나에 제시된 서열을 포함하며, 선택적으로 gRNA 또는 각각의 gRNA는 서열번호 391-424 중 어느 하나에 제시된다. 본원의 임의의 구현예에서, gRNA 또는 각각의 gRNA는 서열번호 207, 213, 215, 217, 221, 222, 241, 245, 258, 261, 268, 274, 380, 387 중 어느 하나에 제시된 서열을 포함하며, 선택적으로 gRNA 또는 각각의 gRNA는 서열번호 402, 408, 410, 412, 416, 417, 436, 440, 453, 456, 463, 469, 575, 582 중 어느 하나에 제시된다. 본원의 임의의 구현예에서, gRNA는 서열번호 217에 제시된 서열을 포함하며, 선택적으로 gRNA는 서열번호 412에 제시된다.

[0043] [0037] 본원의 임의의 구현예에서, Cas 단백질 또는 이의 변이체는 Cas9 단백질 또는 이의 변이체이다. 본원의 임의의 구현예에서, Cas 단백질 또는 이의 변이체는 Cas12 단백질 또는 이의 변이체이다. 본원의 임의의 구현예에서, Cas 단백질 또는 이의 변이체는 변이체 Cas 단백질이며, 여기서 변이체 Cas 단백질은 뉴클레아제 활성이 없거나 비활성화된 Cas(dCas) 단백질이다. 본원의 임의의 구현예에서, 변이체 Cas 단백질은 뉴클레아제 활성이 없거나 비활성화된 Cas9(dCas9) 단백질인 변이체 Cas9 단백질이다. 본원의 임의의 구현예에서, Cas9 단백질 또는 이의 변이체는 *Staphylococcus aureus* Cas9(SaCas9) 단백질 또는 이의 변이체이다. 본원의 임의의 구현예에

서, 변이체 Cas9는 서열번호 596의 위치 넘버링을 참조로 D10A 및 N580A로부터 선택된 적어도 하나의 아미노산 돌연변이를 포함하는 *Staphylococcus aureus* dCas9 단백질(dSaCas9)이다. 본원의 임의의 구현예에서, 변이체 Cas9 단백질은 서열번호 597에 제시된 서열, 또는 서열번호 597과 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 본원의 임의의 구현예에서, Cas9 단백질 또는 이의 변이체는 화농성연쇄상구균(*Streptococcus pyogenes*) Cas9(SpCas9) 단백질 또는 이의 변이체이다. 본원의 임의의 구현예에서, 변이체 Cas9는 서열번호 598의 위치 넘버링을 참조로 D10A 및 H840A로부터 선택된 적어도 하나의 아미노산 돌연변이를 포함하는 화농성연쇄상구균(*Streptococcus pyogenes*) dCas9 (dSpCas9) 단백질이다. 본원의 임의의 구현예에서, 변이체 Cas9 단백질은 서열번호 599에 제시된 서열 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다.

[0044] [0038] 본원의 임의의 구현예에서, 적어도 하나의 DNA-결합 도메인은 조작된 아연 집게 단백질(eZFP)을 포함한다. 본원의 임의의 구현예에서, 적어도 하나의 DNA-결합 도메인은 eZFP이다. 본원의 임의의 구현예에서, 표적 부위는 서열번호 1045, 1046, 1052 중 어느 하나에 제시된 뉴클레오타이드 서열, 적어도 12nt의 그의 연속된 부분, 또는 상기 중 어느 하나의 상보적 서열을 포함한다. 본원의 임의의 구현예에서, 표적 부위는 서열번호 1045, 1046, 1052 중 어느 하나에 제시된 뉴클레오타이드 서열을 포함한다.

[0045] [0039] 본원의 임의의 구현예에서, 아연 집게 단백질은 N-말단에서 C-말단의 순서로 F1 내지 F6으로 표시되는 6개의 아연 집게를 포함하고, 여기서 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같다: 1) F1:SEADRSR(서열번호 720) F2:DRSNLTR(서열번호 721) F3:QSSDLSR(서열번호 722) F4:YHWYLKK(서열번호 723) F5:RSDSLSV(서열번호 724) F6:QNANRKT(서열번호 725); 2) F1:RSDVLST(서열번호 726) F2:DNSSRTR(서열번호 727) F3:RPYTLRL(서열번호 728) F4:DSSHRTR(서열번호 729) F5:RSDHLSQ(서열번호 730) F6:DSSHRTR(서열번호 731); 3) F1:RSDHLSQ (서열번호 732) F2:QSADRTK(서열번호 733) F3:RSDHLSQ(서열번호 734) F4:RRSDLKR(서열번호 735) F5:RSDHLSR(서열번호 736) F6:QSSDLRR(서열번호 737); 4) F1:RSDNLSE(서열번호 738) F2:TSSNRKT(서열번호 739) F3:DRSHLTR (서열번호 740) F4:RSDALTQ(서열번호 741) F5:DRSALAR(서열번호 742) F6:RRFTLSK(서열번호 743); 5) F1:RSDHLSE(서열번호 744) F2:QYSGRYY(서열번호 745) F3:HGQTLNE(서열번호 746) F4:QSGNLAR(서열번호 747) F5:RSDSLLR(서열번호 748) F6:CREYRGK(서열번호 749); 6) F1:QSANRTT(서열번호 750) F2:RSANLTR(서열번호 751) F3:RSDVLSE(서열번호 752) F4:TSGHLSR(서열번호 753) F5:QSSDLSR(서열번호 754), F6:QWSTRKR(서열번호 755); 7) F1:QSGNLAR(서열번호 756) F2:ATCCLAH(서열번호 757) F3:RWQYLPT(서열번호 758) F4:DRSALAR(서열번호 759) F5:RSDNLSE(서열번호 760) F6:KRCNLRC(서열번호 761); 8) F1:NPANLTR(서열번호 762) F2:QNATRTK(서열번호 763) F3:QSGHLAR(서열번호 764) F4:NRHDRAK(서열번호 765) F5:RSDHLSE(서열번호 766), F6:QRRSRYK(서열번호 767); 9) F1:QSSDLSR(서열번호 768) F2:HRSTRNR(서열번호 769) F3:RSDVLSA(서열번호 770) F4:DSRTRKN(서열번호 771) F5:QSGSLTR(서열번호 772) F6:DQSGLAH(서열번호 773); 10) F1:QNPAQWR(서열번호 774) F2:RSADLSR(서열번호 775) F3:TSGSLSR(서열번호 776) F4:RSDHLSR(서열번호 777) F5:RSDSLLR(서열번호 778) F6:QSYDRFQ(서열번호 779); 11) F1:TSGSLSR(서열번호 780) F2:RSDHLSR(서열번호 781) F3:RSDSLLR(서열번호 782) F4:QSYDRFQ(서열번호 783) F5:RSDNLST(서열번호 784) F6:DNRDRIK(서열번호 785); 12) F1:DRSNLSR(서열번호 786) F2:LRQNLIM(서열번호 787) F3:ERGTLAR(서열번호 788) F4:RSDALTQ(서열번호 789) F5:RSDSLSQ(서열번호 790) F6:RKADRTR(서열번호 791); 13) F1:QYCCLTN(서열번호 792) F2:TSGNLTR(서열번호 793) F3:QSSDLSR(서열번호 794) F4:FRYYLKR(서열번호 795) F5:QSGDLTR(서열번호 796) F6:DKGNLTK(서열번호 797); 14) F1:TSGSLSR(서열번호 798) F2:RSDNLTT(서열번호 799) F3:QSGNLAR(서열번호 800) F4:DRTTLMR(서열번호 801) F5:QSGHLAR(서열번호 802) F6:QLTHLNS(서열번호 803); 15) F1:IKHDLHR(서열번호 804) F2:RSANLTR(서열번호 805) F3:RSDNLAR(서열번호 806) F4:QNVSRPR(서열번호 807) F5:RSDDLSK(서열번호 808) F6:DSSHRTR(서열번호 809); 16) F1:RSDNLAR(서열번호 810) F2:QNVSRPR (서열번호 811) F3:RSDDLSK(서열번호 812) F4:DSSHRTR(서열번호 813) F5:TSSNRKT(서열번호 814) F6:AQWTRAC(서열번호 815); 17) F1:RSDDLSK(서열번호 816) F2:DSSHRTR(서열번호 817) F3:TSSNRKT(서열번호 818) F4:AQWTRAC (서열번호 819) F5:RKQTRTT(서열번호 820) F6:HRSSLRR(서열번호 821); 18) F1:QSAHRKN(서열번호 822) F2:TSSNRKT(서열번호 823) F3:RSDNLSA(서열번호 824) F4:RNNDRKT(서열번호 825) F5:TSGSLSR(서열번호 826) F6:QAGHLAK(서열번호 827); 19) F1:RSDHLSQ(서열번호 828) F2:ASSTRTK(서열번호 829) F3:RSDDLTR(서열번호 830) F4:QKSNLSS(서열번호 831) F5:QSANRTT(서열번호 832) F6:QNATRTK(서열번호 833); 20) F1:RSDTLSE(서열번호 834) F2:RRWTLVG(서열번호 835) F3:DRSNLSR(서열번호 836) F4:QSGDLTR(서열번호 837) F5:QSSDLSR(서열번호 838) F6:YHWYLKK(서열번호 839); 21) F1:RSANLAR(서열번호 840) F2:RSDNLRE(서열번호 841) F3:RPYTLRL(서열번호 842) F4:HRSNLNK(서열번호 843) F5:QSGSLTR(서열번호 844) F6:TSANLSR(서열번호 845); 22) F1:RSDDLVR(서열번호 846) F2:TSGSLVR(서열번호 847) F3:RSDKLVR(서열번호 848) F4:RSDELVR(서열번호 849) F5:TSHSLTE(서열번호 850) F6:RADNLTE(서열번호 851); 23) F1:ERSHLRE(서열번호 852) F2:TSHSLTE(서열번호 853) F3:QAGHLAS(서열

번호 854) F4:TSHSLTE(서열번호 855) F5:DPGHLVR(서열번호 856) F6:TSGNLVR(서열번호 857); 24) F1:RADNLTE(서열번호 858) F2:TSGSLVR(서열번호 859) F3:RKDNLKN(서열번호 860) F4:QSSSLVR(서열번호 861) F5:RSDKLVR(서열번호 862) F6:DSGNLRV(서열번호 863); 25) F1:QSSSLVR(서열번호 864) F2:QSGDLRR(서열번호 865) F3:RSDERKR(서열번호 866) F4:HRTTLTN(서열번호 867) F5:RSDHLTN(서열번호 868) F6:TSGELVR(서열번호 869); 26) F1:QSGDLRR(서열번호 870) F2:RSDERKR(서열번호 871) F3:HRTTLTN(서열번호 872) F4:RSDHLTN(서열번호 873) F5:TSGELVR(서열번호 874) F6:RSDDLVR(서열번호 875); 27) F1:QRAHLER(서열번호 876) F2:QLAHLRA(서열번호 877) F3:DPGHLVR(서열번호 878) F4:RRSACRR(서열번호 879) F5:RSDHLTT(서열번호 880) F6:QSSSLVR(서열번호 881); 및 28) F1:QSSNLVR(서열번호 882) F2:RSDDLVR(서열번호 883) F3:THLDLIR(서열번호 884) F4:TSGNLTE(서열번호 885) F5:RRSACRR(서열번호 886) F6:RNDTLTE(서열번호 887).

[0046] [0040] 본원의 임의의 구현예에서, 아연 집게 단백질은 N-말단에서 C-말단의 순서로 F1에서 F6까지 표시되는 6개의 아연 집게를 포함하고, 여기서 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같다: F1: QSAHRKN(서열번호 822) F2: TSSNRKT(서열번호 823) F3: RSDNLSA(서열번호 824) F4: RNNDRKT(서열번호 825) F5: TSGSLSR(서열번호 826) F6: QAGHLAK(서열번호 827).

[0047] [0041] 본원의 임의의 구현예에서, 아연 집게 단백질은 N-말단에서 C-말단의 순서로 F1에서 F6까지 표시되는 6개의 아연 집게를 포함하고, 여기서 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같다: F1: RSDHLSQ(서열번호 828) F2: ASSTRTK(서열번호 829) F3: RSDDLTR(서열번호 830) F4: QKSNLSS(서열번호 831) F5: QSANRTT(서열번호 832) F6: QNATRTK(서열번호 833).

[0048] [0042] 본원의 임의의 구현예에서, 아연 집게 단백질은 N-말단에서 C-말단의 순서로 F1에서 F6까지 표시되는 6개의 아연 집게를 포함하고, 여기서 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같다: F1:QSSSLVR(서열번호 864) F2:QSGDLRR(서열번호 865) F3:RSDERKR(서열번호 866) F4:HRTTLTN(서열번호 867) F5:RSDHLTN(서열번호 868) F6:TSGELVR(서열번호 869).

[0049] [0043] 또한, 본 발명에서는 다음을 포함하는 후성적 변형 DNA 표적 시스템이 제공된다: a) 하나 이상의 HBV 유전자 또는 이의 조절 요소의 표적 부위에 결합하는 eZFP 및 b) 하나 이상의 HBV 유전자의 전사를 억제하는 적어도 하나의 효과기 도메인, 여기서 아연 집게 단백질은 N-말단에서 C-말단 순서로 F1에서 F6까지 표시되는 6개의 아연 집게를 포함하고, 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같다: 1) F1:SEADRSR(서열번호 720) F2:DRSNLTR(서열번호 721) F3:QSSDLSR(서열번호 722) F4:YHWYLKK(서열번호 723) F5:RSDSLSV(서열번호 724) F6:QNANRKT(서열번호 725); 2) F1:RSDVLST(서열번호 726) F2:DNSSRTR(서열번호 727) F3:RPYTLRL(서열번호 728) F4:DSSHRTR(서열번호 729) F5:RSDHLSQ(서열번호 730) F6:DSSHRTR(서열번호 731); 3) F1:RSDHLSQ(서열번호 732) F2:QSADRTK(서열번호 733) F3:RSDHLSQ(서열번호 734) F4:RRSDLKR(서열번호 735) F5:RSDHLSR(서열번호 736) F6:QSSDLRR(서열번호 737); 4) F1:RSDNLSE(서열번호 738) F2:TSSNRKT(서열번호 739) F3:DRSHLTR(서열번호 740) F4:RSDALTQ(서열번호 741) F5:DRSALAR(서열번호 742) F6:RRFTLSK(서열번호 743); 5) F1:RSDHLSE(서열번호 744) F2:QYSGRYY(서열번호 745) F3:HGQTLNE(서열번호 746) F4:QSGNLAR(서열번호 747) F5:RSDSLLR(서열번호 748) F6:CREYRGK(서열번호 749); 6) F1:QSANRTT(서열번호 750) F2:RSANLTR(서열번호 751) F3:RSDVLSE(서열번호 752) F4:TSGHLSR(서열번호 753) F5:QSSDLSR(서열번호 754), F6:QWSTRKR(서열번호 755); 7) F1:QSGNLAR(서열번호 756) F2:ATCCLAH(서열번호 757) F3:RWQYLPT(서열번호 758) F4:DRSALAR(서열번호 759) F5:RSDNLSE(서열번호 760) F6:KRCNLRC(서열번호 761); 8) F1:NPANLTR(서열번호 762) F2:QNATRTK(서열번호 763) F3:QSGHLAR(서열번호 764) F4:NRHDRAK(서열번호 765) F5:RSDHLSE(서열번호 766), F6:QRRSRYK(서열번호 767); 9) F1:QSSDLSR(서열번호 768) F2:HRSTRNR(서열번호 769) F3:RSDVLSA(서열번호 770) F4:DSRTRKN(서열번호 771) F5:QSGSLTR(서열번호 772) F6:DQSGLAH(서열번호 773); 10) F1:QNPAQWR(서열번호 774) F2:RSADLSR(서열번호 775) F3:TSGSLSR(서열번호 776) F4:RSDHLSR(서열번호 777) F5:RSDSLLR(서열번호 778) F6:QSYDRFQ(서열번호 779); 11) F1:TSGSLSR(서열번호 780) F2:RSDHLSR(서열번호 781) F3:RSDSLLR(서열번호 782) F4:QSYDRFQ(서열번호 783) F5:RSDNLST(서열번호 784) F6:DNRDRIK(서열번호 785); 12) F1:DRSNLSR(서열번호 786) F2:LRQNLIM(서열번호 787) F3:ERGTLAR(서열번호 788) F4:RSDALTQ(서열번호 789) F5:RSDSLSQ(서열번호 790) F6:RKADRTR(서열번호 791); 13) F1:QYCCLTN(서열번호 792) F2:TSGNLTR(서열번호 793) F3:QSSDLSR(서열번호 794) F4:FRYYLKR(서열번호 795) F5:QSGDLTR(서열번호 796) F6:DKGNLTK(서열번호 797); 14) F1:TSGSLSR(서열번호 798) F2:RSDNLTT(서열번호 799) F3:QSGNLAR(서열번호 800) F4:DRTTLMR(서열번호 801) F5:QSGHLAR(서열번호 802) F6:QLTHLNS(서열번호 803); 15) F1:IKHDLHR(서열번호 804) F2:RSANLTR(서열번호 805) F3:RSDNLAR(서열번호 806) F4:QNVSRPR(서열번호 807) F5:RSDDLSK(서열번호 808) F6:DSSHRTR(서열번호 809); 16) F1:RSDNLAR(서열번호 810) F2:QNVSRPR(서열번호 811) F3:RSDDLSK(서열번호 812) F4:DSSHRTR(서열번호 813) F5:TSSNRKT(서열번호 814) F6:AQWTRAC(서열번호 815); 17)

F1:RSDDLSK(서열번호 816) F2:DSSHRTR(서열번호 817) F3:TSSNRKT(서열번호 818) F4:AQWTRAC(서열번호 819) F5:RKQTRTT(서열번호 820) F6:HRSSLRR(서열번호 821); 18) F1:QSAHRKN(서열번호 822) F2:TSSNRKT(서열번호 823) F3:RSDNLSA(서열번호 824) F4:RNNDRKT(서열번호 825) F5:TSGSLSR(서열번호 826) F6:QAGHLAK(서열번호 827); 19) F1:RSDHLSQ(서열번호 828) F2:ASSTRTK(서열번호 829) F3:RSDDLTR(서열번호 830) F4:QKSNLSS(서열번호 831) F5:QSANRTT(서열번호 832) F6:QNATRTK(서열번호 833); 20) F1:RSDTLSE(서열번호 834) F2:RRWTLVG(서열번호 835) F3:DRSNLSR(서열번호 836) F4:QSGDLTR(서열번호 837) F5:QSSDLSR(서열번호 838) F6:YHWYLKK(서열번호 839); 21) F1:RSANLAR(서열번호 840) F2:RSDNLRE(서열번호 841) F3:RPYTLRL(서열번호 842) F4:HRSNLNK(서열번호 843) F5:QSGSLTR(서열번호 844) F6:TSANLSR(서열번호 845); 22) F1:RSDDLVR(서열번호 846) F2:TSGSLVR(서열번호 847) F3:RSDKLVR(서열번호 848) F4:RSDELVR(서열번호 849) F5:TSHSLTE(서열번호 850) F6:RADNLTE(서열번호 851); 23) F1:ERSHLRE(서열번호 852) F2:TSHSLTE(서열번호 853) F3:QAGHLAS(서열번호 854) F4:TSHSLTE(서열번호 855) F5:DPGHLVR(서열번호 856) F6:TSGNLVR(서열번호 857); 24) F1:RADNLTE(서열번호 858) F2:TSGSLVR(서열번호 859) F3:RKDNLKN(서열번호 860) F4:QSSSLVR(서열번호 861) F5:RSDKLVR(서열번호 862) F6:DSGNLRV(서열번호 863); 25) F1:QSSSLVR(서열번호 864) F2:QSGDLRR(서열번호 865) F3:RSDERKR(서열번호 866) F4:HRTTLTN(서열번호 867) F5:RSDHLTN(서열번호 868) F6:TSGELVR(서열번호 869); 26) F1:QSGDLRR(서열번호 870) F2:RSDERKR(서열번호 871) F3:HRTTLTN(서열번호 872) F4:RSDHLTN(서열번호 873) F5:TSGELVR(서열번호 874) F6:RSDDLVR(서열번호 875); 27) F1:QRAHLER(서열번호 876) F2:QLAHLRA(서열번호 877) F3:DPGHLVR(서열번호 878) F4:RRSACRR(서열번호 879) F5:RSDHLTT(서열번호 880) F6:QSSSLVR(서열번호 881); 28) F1:QSSNLVR(서열번호 882) F2:RSDDLVR(서열번호 883) F3:THLDLIR(서열번호 884) F4:TSGNLTE(서열번호 885) F5:RRSACRR(서열번호 886) F6:RNDTLTE(서열번호 887).

[0050] [0044] 본원의 임의의 구현예에서, 조작된 아연 집게 단백질은 서열번호 692-719 중 어느 하나에 제시된 서열 또는 이의 일부, 또는 이에 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, 조작된 아연 집게 단백질은 서열번호 888-915 중 어느 하나에 제시된 서열 또는 이의 일부, 또는 이에 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩된다.

[0051] [0045] 또한, 본 발명에서는 다음을 포함하는 후성적 변형 DNA 표적 시스템이 제공된다: a) 하나 이상의 HBV 유전자 또는 이의 조절 요소의 표적 부위에 결합하는 조작된 아연 집게 단백질, 및 b) 하나 이상의 HBV 유전자의 전사를 억제하는 적어도 하나의 효과기 도메인, 여기서 아연 집게 단백질은 N-말단에서 C-말단의 순서로 F1에서 F6까지 표시되는 6개의 아연 집게를 포함하고, 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같다: F1:QSAHRKN(서열번호 822) F2:TSSNRKT(서열번호 823) F3:RSDNLSA(서열번호 824) F4:RNNDRKT(서열번호 825) F5:TSGSLSR(서열번호 826) F6:QAGHLAK(서열번호 827).

[0052] [0046] 본원의 임의의 구현예에서, eZFP는 서열번호 709에 제시된 서열 또는 이의 일부, 또는 그와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, 조작된 아연 집게 단백질은 서열번호 709 중 어느 하나에 제시된 서열을 포함한다.

[0053] [0047] 본원의 임의의 구현예에서, 조작된 아연 집게 단백질은 서열번호 905에 제시된 서열, 또는 그 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩된다. 본원의 임의의 구현예에서, 조작된 아연 집게 단백질은 서열번호 905 중 어느 하나에 제시된 서열에 의해 인코딩된다.

[0054] [0048] 또한, 본 발명은 다음을 포함하는 후성적 변형 DNA 표적 시스템을 제공한다: a) 하나 이상의 HBV 유전자 또는 이의 조절 요소의 표적 부위에 결합하는 조작된 아연 집게 단백질, 및 b) 하나 이상의 HBV 유전자의 전사를 억제하는 적어도 하나의 효과기 도메인, 여기서 아연 집게 단백질은 N-말단에서 C-말단의 순서로 F1에서 F6까지 표시되는 6개의 아연 집게를 포함하고, 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같다: F1:RSDHLSQ(서열번호 828) F2:ASSTRTK(서열번호 829) F3:RSDDLTR(서열번호 830) F4:QKSNLSS(서열번호 831) F5:QSANRTT(서열번호 832) F6:QNATRTK(서열번호 833). 일부 구현예에서, 조작된 아연 집게 단백질은 서열번호 710에 제시된 서열 또는 이의 일부, 또는 그와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, 조작된 아연 집게 단백질은 서열번호 710 중 어느 하나에 제시된 서열을 포함한다. 본원의 임의의 구현예에서, 조작된 아연 집게 단백질은 서열번호 906에 제시된 서열 또는 이의 일부, 또는 그와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩된다. 본원의 임의의 구현예에서, 설계된 아연 집

게 단백질은 서열번호 906 중 어느 하나에 제시된 서열에 의해 인코딩된다.

[0055] [0049] 또한, 본 발명에서는 다음을 포함하는 후성적 변형 DNA 표적 시스템이 제공된다: a) 하나 이상의 HBV 유전자 또는 이의 조절 요소의 표적 부위에 결합하는 조작된 아연 집게 단백질, 및 b) 하나 이상의 HBV 유전자의 전사를 억제하는 적어도 하나의 효과기 도메인, 여기서 아연 집게 단백질은 N-말단에서 C-말단의 순서로 F1에서 F6까지 표시되는 6개의 아연 집게를 포함하고, 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같다: F1: QSSSLVR(서열번호 864) F2: QSGDLRR(서열번호 865) F3: RSDERKR(서열번호 866) F4: HRTTLTN(서열번호 867) F5: RSDHLTN(서열번호 868) F6: TSGELVR(서열번호 869). 일부 구현예에서, 조작된 아연 집게 단백질은 서열번호 716에 제시된 서열 또는 이의 일부, 또는 그와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, 조작된 아연 집게 단백질은 서열번호 716 중 어느 하나에 제시된 서열을 포함한다. 일부 구현예에서, 조작된 아연 집게 단백질은 서열번호 912에 제시된 서열 또는 이의 일부, 또는 그와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩된다. 일부 구현예에서, 설계된 아연 집게 단백질은 서열번호 912 중 어느 하나에 제시된 서열에 의해 인코딩된다.

[0056] [0050] 본원의 임의의 구현예에서, 적어도 하나의 효과기 도메인은 전사 억제를 유도한다. 본원의 임의의 구현예에서, 적어도 하나의 효과기 도메인은 DNA 메틸트랜스퍼라제이다. 본원의 임의의 구현예에서, 적어도 하나의 효과기 도메인은 DNA 메틸트랜스퍼라제와 헤테로크로마틴 유도 인자를 동원할 수 있는 억제자 도메인을 포함하거나, 선택적으로 헤테로크로마틴 유도 인자는 히스톤 메틸트랜스퍼라제를 포함한다. 본원의 임의의 구현예에서, 적어도 하나의 효과기 도메인은 DNA 메틸트랜스퍼라제와 히스톤 메틸트랜스퍼라제를 포함한다. 본원의 임의의 구현예에서, 적어도 하나의 효과기 도메인은 KRAB 억제자 도메인, ERF 억제자 도메인, Mxi1 억제자 도메인, SID4X 억제자 도메인, Mad-SID 억제자 도메인, LSD1 억제자 도메인, 또는 DNMT3A, DNMT3A-3L, DNMT3A/L-KRAB 융합 억제자 도메인, DNMT3B 도메인 결합 단백질, EZH2 억제자 도메인, 또는 LSD1 억제자 도메인, 또는 상기 중 어느 것의 변이체로부터 선택된다. 본원의 임의의 구현예에서, 적어도 하나의 효과기 도메인은 서열번호 590, 또는 600-608, 651, 661, 664, 665, 666, 668 및 669 중 어느 하나로부터 선택된 서열 또는 이의 도메인, 이의 일부 또는 상기 중 어느 것과 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 본원의 임의의 구현예에서, 적어도 하나의 효과기 도메인은 KRAB 도메인 또는 그 변이체를 포함한다. 본원의 임의의 구현예에서, 적어도 하나의 효과기 도메인은 서열번호 590에 제시된 서열, 이의 일부 또는 상기 중 어느 것과 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 본원의 임의의 구현예에서, 적어도 하나의 효과기 도메인은 DNMT3A/L 도메인 또는 그 변이체를 포함한다. 본원의 임의의 구현예에서, 적어도 하나의 효과기 도메인은 서열번호 604 및 607에 제시된 서열, 이의 일부, 또는 앞서 언급된 것들 중 어느 하나와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함하며; 또는 적어도 하나의 효과기 도메인은 서열번호 651에 제시된 서열, 이의 일부, 또는 서열번호 651과 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다.

[0057] [0051] 본원의 임의의 구현예에서, 융합 단백질은 DNMT3A/3L-dSpCas9-KRAB 융합 단백질을 포함한다. 본원의 임의의 구현예에서, 융합 단백질은 서열번호 645에 제시된 서열, 이의 일부 또는 상기 중 어느 것과 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다.

[0058] [0052] 본원의 임의의 구현예에서, 적어도 하나의 효과기 도메인은 DNA-결합 도메인 또는 그 구성요소의 N-말단, C-말단, 또는 N-말단과 C-말단 둘 다에 융합된다.

[0059] [0053] 본원의 임의의 구현예에서, 융합 단백질은 서열번호 680에 제시된 서열, 이의 일부 또는 상기 중 어느 것과도 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 핵산 서열에 의해 인코딩된다. 본원의 임의의 구현예에서, 융합 단백질은 서열번호 680에 제시된 서열에 의해 인코딩된다.

[0060] [0054] 본원의 임의의 구현예에 있어서, 융합 단백질은 서열번호 916-943 중 어느 하나에 제시된 서열, 이의 일부 또는 상기 어느 것과 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 핵산 서열에 의해 인코딩된다. 본원의 임의의 구현예에 있어서, 융합 단백질은 서열번호 944-971 중 어느 하나에 제시된 서열, 이의 일부 또는 상기 어느 것과 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 본원의 임의의 구현예에서, 융합 단백질은 서열번호 961, 962 또는 968 중 어느 하나에 제시된 서열을 포함한다.

[0061] [0055] 본원의 임의의 구현예에서, 융합 단백질은 DNMT3A/3L-eZFP-KRAB 융합 단백질을 포함한다.

[0062] [0056] 본원의 임의의 구현예에 있어서, 융합 단백질은 서열번호 972-999 중 어느 하나에 제시된 서열, 이의 일부 또는 상기 어느 것과 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 핵산 서열에 의해 인코딩된다. 본원의 임의의 구현예에 있어서, 융합 단백질은 서열번호 933, 934 또는 940 중 어느 하나에 제시된 서열, 이의 일부 또는 상기 어느 것과 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 핵산 서열에 의해 인코딩된다. 본원의 임의의 구현예에 있어서, 융합 단백질은 서열번호 1000-1027 중 어느 하나에 제시된 서열, 이의 일부 또는 상기 어느 것과 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 본원의 임의의 구현예에 있어서, 융합 단백질은 서열번호 1017, 1018, 또는 1024 중 어느 하나에 제시된 서열, 이의 일부 또는 상기 어느 것과 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다.

[0063] [0057] 본원의 임의의 구현예에서, 융합 단백질은 하나 이상의 핵 국소화 신호(NLS)를 더 포함한다.

[0064] [0058] 본원의 임의의 구현예에서, 융합 단백질은 또한 다음 중 두 개 이상을 연결하는 하나 이상의 링커를 더 포함한다: DNA-결합 도메인, 최소한 하나의 효과기 도메인 및 하나 이상의 핵 국소화 신호.

[0065] [0059] 본원의 임의의 구현예에서, DNA 표적 시스템은 모든 B형 간염 바이러스 유전체를 표적으로 한다.

[0066] [0060] 본원의 임의의 구현예에서, DNA-표적화 시스템은 모든 B형 간염 바이러스 게놈의 최소 70%를 표적화한다. 본원의 임의의 구현예에서, DNA-표적화 시스템은 모든 B형 간염 바이러스 게놈의 최소 60%를 표적화한다. 본원의 임의의 구현예에서, DNA-표적화 시스템은 모든 B형 간염 바이러스 게놈의 최소 50%를 표적화한다.

[0067] [0061] 본원의 임의의 구현예에서, DNA-표적화 시스템은 표적 부위 또는 그 근처에서 유전적 파괴 또는 DNA 파손을 유도할 수 없다.

[0068] [0062] 본원의 임의의 구현예에서, 하나 이상의 HBV 유전자의 전사를 억제하면 HBV DNA 서열에서 RNA 수준 및/또는 단백질 수준이 감소한다. 본원의 임의의 구현예에서, 전사를 억제하는 것은 총 B형 간염 바이러스 RNA 전사 수준을 감소시키는 것을 포함한다. 본원의 임의의 구현예에서, 전사를 억제하는 것은 B형 간염 전코어("preC"), 전게놈("pgRNA"), preS1, preS2/S 및 HBx 수준을 감소시키는 것을 포함한다. 본원의 임의의 구현예에서, 전사를 억제하는 것은 HBx 수준을 감소시키는 것을 포함한다. 본원의 임의의 구현예에서, 전사를 억제하는 것은 B형 간염 표면 항원(HBsAg) 및/또는 B형 간염 바이러스 코어 관련 항원(HbcrAg) 단백질 수준을 감소시키는 것을 포함한다. 본원의 임의의 구현예에서, 전사를 억제하는 것은 HbsAg 전사체 및/또는 단백질 수준을 최소 90% 감소시키는 것을 포함한다. 본원의 임의의 구현예에서, 전사를 억제하는 것은 cccDNA에서 HbcrAg 전사체 및/또는 단백질 수준이 최소 50% 감소하는 것을 포함한다.

[0069] [0063] 또한, 본원에는 B형 간염 바이러스 DNA 서열의 표적 부위에 결합하는 가이드 RNA(gRNA)가 제공된다. 일부 구현예에서, B형 간염 바이러스 DNA 서열은 B형 간염(HBV) 유전자 또는 이의 조절 요소이다. 일부 구현예에서, 표적 부위는 공유적으로 폐쇄된 원형 DNA(cccDNA) 형태, 이완된 원형 DNA(rcDNA) 형태로 존재하거나 및/또는 인간 게놈 DNA에 통합된다. 본원의 임의의 구현예에서, 표적 부위는 HBV 복제 및/또는 HBV 전사를 제어하는 데 관여하는 유전자 또는 이의 조절 요소에 있거나 그 근처에 있다. 본원의 임의의 구현예에서, HBV 복제 및/또는 HBV 전사를 제어하는 데 관여하는 유전자는 폴리머라제, 외피 단백질, 캡시드 단백질, 전사 인자 또는 전사 전사 활성화제를 인코딩한다. 본원의 임의의 구현예에서, HBV 복제 및/또는 HBV 전사를 제어하는 데 관여하는 유전자는 폴리머라제 유전자, S-패밀리 유전자, X-유전자 또는 코어-패밀리 유전자이다. 본원의 임의의 구현예에서, 표적 부위는 B형 간염 바이러스 단백질 X(HBx)를 인코딩하는 X-유전자의 유전자 또는 이의 조절 요소에 있다. 본원의 임의의 구현예에서, 표적 부위는 HBV 복제 및/또는 HBV 전사를 제어하는 데 관여하는 조절 요소에 있거나 그 근처에 있다.

[0070] [0064] 본원의 임의의 구현예에서, 조절 요소는 프로모터 영역이다. 본원의 임의의 구현예에서, 프로모터 영역은 pre-S1 프로모터, pre-S2 프로모터, X 프로모터 또는 기본 코어 프로모터이다. 본원의 임의의 구현예에서, 조절 요소는 인핸서 영역이다. 본원의 임의의 구현예에서, 인핸서 영역은 Enh1 또는 Enh2 인핸서 영역이다. 본원의 임의의 구현예에서, 조절 요소는 전사 처리 제어 영역이다.

[0071] [0065] 본원의 임의의 구현예에서, 표적 부위는 코딩 영역이다. 본원의 임의의 구현예에서, 표적 부위는 전사 개시 부위의 500bp 이내, 1000bp 이내, 1500bp 이내에 위치한다. 본원의 임의의 구현예에서, 표적 부위는 서열번호 650에 명시된 HBV 게놈을 기준으로 HBV 게놈의 0-3300 염기쌍(bp) 사이, 선택적으로 0-3189 bp 사이의 염

기쌍에 위치한 표적 영역 내에 위치한다. 본원의 임의의 구현예에서, 표적 부위는 HBV 게놈의 43 bp - 490 bp, 1033 bp - 1749 bp, 1800 bp - 1950 bp, 또는 2953 bp - 3182 bp 사이의 염기쌍에 위치한 표적 영역 내에 위치하며, 이는 서열번호 650에 명시된 HBV 게놈을 기준으로 하는 위치에 상응하는 것이다. 본원의 임의의 구현예에서, 표적 부위는 서열번호 650에 명시된 HBV 게놈을 기준으로 1 bp - 42 bp, 491 bp - 1032bp, 1750 bp - 1799 bp, 또는 1951 bp -2952 bp 사이의 염기쌍에 위치한 표적 영역 내에 위치한다.

[0072] [0066] 본원의 임의의 구현예에서, 표적 부위 또는 각각의 표적 부위는 HBV 게놈의 CpG 섬에 존재한다. 본원의 임의의 구현예에서, 표적 부위는 서열번호 650에 제시된 HBV 게놈을 기준으로 하는 위치에 상응하는 HBV 게놈의 67bp-392bp, 1033bp-1749bp 또는 2215bp-2490bp 사이의 염기쌍에 위치한 표적 영역 내에 위치한다. 본원의 임의의 구현예에서, 표적 부위는 서열번호 650에 제시된 HBV 게놈을 기준으로 하는 위치에 상응하는 HBV 게놈의 1033bp-1749bp 사이의 염기쌍에 위치한 표적 영역 내에 위치한다. 본원의 임의의 구현예에서, 표적 부위 또는 각각의 표적 부위는 B형 간염 X 단백질(HBx) 시작 코돈 상류의 300개 염기쌍 내에 걸쳐 있는 표적 영역 내에 위치한다.

[0073] [0067] 또한, 여기에는 B형 간염 X 단백질(HBx) 시작 코돈 상류의 300개 염기쌍 내에 걸쳐 있는 표적 영역 내의 표적 부위에 결합하는 gRNA(gRNA)가 제공된다. 본원의 모든 구현예에서, 표적 부위는 HBx 기저 코어 프로모터 영역에 위치한다. 본원의 모든 구현예에서, 표적 부위는 HBx 프로모터/인핸서 영역 내에 위치한다. 본원의 모든 구현예에서, 표적 부위는 B형 간염 X 단백질(HBx) 시작 코돈 상류의 250개 염기쌍 내에 걸쳐 있는 표적 영역 내에 있다. 본원의 임의의 구현예에서, 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조하는 위치에 상응하는 HBV 게놈의 1060-1480bp에 걸친 표적 영역 내에 있다. 본원의 임의의 구현예에서, 표적 부위는 B형 간염 X 단백질(HBx) 시작 코돈 상류의 150개 염기쌍 내에 걸친 표적 영역 내에 있다. 본원의 임의의 구현예에서, 표적 부위는 B형 간염 X 단백질(HBx) 시작 코돈 상류의 120개 염기쌍 내에 걸친 표적 영역 내에 있다. 본원의 구현예 중 어느 하나에서, 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 HBV 게놈의 1250-1374bp에 걸친 서열에 상응하는 표적 영역 서열 내에 있다. 본원의 구현예 중 어느 하나에서, 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 HBV 게놈의 1255-1302bp에 걸친 서열에 상응하는 표적 영역 서열 내에 있다. 본원의 구현예 중 어느 하나에서, 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 HBV 게놈의 1260-1300bp에 걸친 서열에 상응하는 표적 영역 서열 내에 있다. 본원의 구현예 중 어느 하나에서, gRNA는 서열번호 200, 201, 207, 217, 221, 224, 233, 237, 238, 246, 251, 256, 258, 263, 267, 274, 270, 277, 279, 283, 284, 293, 294, 308, 311, 313, 316, 319, 320, 325, 328, 330, 333, 338, 345, 347, 350, 353, 359, 360, 369, 370, 371, 377, 380, 384, 385 또는 387, 적어도 14개 뉴클레오타이드의 이의 연속 부분 또는 상기 중 어느 하나의 상보적 서열 중 어느 하나에 제시된 서열을 포함하고, 선택적으로 gRNA 또는 각각의 gRNA는 서열번호 395, 402, 408, 412, 416, 419, 428, 432, 433, 441, 446, 451, 453, 458, 462, 465, 469, 472, 474, 478, 479, 488, 489, 503, 506, 508, 511, 514, 515, 520, 523, 525, 575, 528, 533, 540, 542, 545, 548, 554, 555, 565, 566, 572, 579, 580 또는 582 중 하나에 제시되어 있다. 본원의 임의의 구현예에서, gRNA는 서열번호 395, 402, 408, 412, 416, 419, 428, 432, 433, 441, 446, 451, 453, 458, 462, 465, 469, 472, 474, 478, 479, 488, 489, 503, 506, 508, 511, 514, 515, 520, 523, 525, 575, 528, 533, 540, 542, 545, 548, 554, 555, 565, 566, 572, 579, 580 또는 582 중 어느 하나에 제시된다. 본원의 모든 구현예에서, gRNA는 서열번호 207, 213, 215, 217, 221, 222, 241, 245, 258, 261, 268, 274, 380, 387 중 어느 하나에 제시된 서열, 적어도 14개 뉴클레오타이드의 이의 연속 부분, 또는 상기 중 어느 하나의 상보적 서열을 포함하며, 선택적으로 gRNA 또는 각각의 gRNA는 서열번호 402, 408, 410, 412, 416, 417, 436, 440, 453, 456, 463, 469, 575 또는 582 중 어느 하나에 제시된다. 본원의 모든 구현예에서, gRNA는 서열번호 402, 408, 410, 412, 416, 417, 436, 440, 453, 456, 463, 469, 575 또는 582 중 어느 하나에 제시된 서열을 포함한다. 본원의 모든 구현예에서, gRNA는 서열번호 217에 제시된 서열, 적어도 14개 뉴클레오타이드의 이의 연속 부분 또는 상기 중 어느 하나의 상보적 서열을 포함하며, 선택적으로 gRNA 또는 각각의 gRNA는 서열번호 412 중 하나에 제시된다. 본원의 모든 구현예에서, gRNA는 서열번호 412에 제시된 서열을 포함한다.

[0074] [0068] 또한, 본 발명에서는 다음을 포함하는 CRISPR Cas-가이드 RNA(gRNA) 조합이 제공된다: (a) 규칙적인 간격을 갖고 나타나는 짧은 회문 구조의 반복 서열 관련(Cas) 단백질 또는 이의 변이체; 및 (b) Cas 단백질 또는 이의 변이체를 B형 간염 바이러스 DNA 서열의 표적 부위에 표적 부위로 표적화하는 청구항 165-202 중 어느 한 항의 적어도 하나의 gRNA. 일부 구현예에서, Cas 단백질 또는 이의 변이체는 Cas9 단백질 또는 이의 변이체이다. 본 명세서의 임의의 구현예에서, Cas 단백질 또는 이의 변이체는 변이체 Cas 단백질이며, 여기서 변이체 Cas 단백질은 뉴클레아제 활성이 없거나 비활성화된 Cas(dCas) 단백질이다. 본 명세서의 임의의 구현예에서, 변이체 Cas 단백질은 뉴클레아제 활성이 없거나 비활성화된 Cas9(dCas9) 단백질인 변이체 Cas9 단백질이

다. 본 명세서의 임의의 구현예에서, Cas9 단백질 또는 이의 변이체는 황색포도상구균(*Staphylococcus aureus*) 이다. Cas9(SaCas9) 단백질 또는 이의 변이체. 일부 구현예에서, 변이체 Cas9는 서열번호 596의 위치 번호에 대한 참조로 D10A 및 N580A에서 선택된 적어도 하나의 아미노산 돌연변이를 포함하는 황색포도상구균 dCas9 단백질(dSaCas9)이다. 일부 구현예에서, 변이체 Cas9 단백질은 서열번호 597에 제시된 서열, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, Cas9 단백질 또는 이의 변이체는 화농성연쇄상구균(*Streptococcus pyogenes*) Cas9(SpCas9) 단백질 또는 이의 변이체이다. 일부 구현예에서, 변이체 Cas9는 서열번호 598의 위치 번호를 참조로 D10A 및 H840A에서 선택된 적어도 하나의 아미노산 돌연변이를 포함하는 화농성연쇄상구균(*Streptococcus pyogenes*) dCas9(dSpCas9) 단백질이다. 일부 구현예에서, 변이체 Cas9 단백질은 서열번호 599에 제시된 서열, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다.

[0075] [0069] 또한, 본원에 공개된 후성적 변형 DNA-표적화 시스템을 인코딩하는 폴리뉴클레오타이드 또는 본원에 공개된 DNA-표적화 시스템의 융합 단백질, 본원에 공개된 gRNA, 본원에 공개된 CRISPR Cas-gRNA 조합, 또는 앞서 언급한 것의 일부 또는 구성요소가 제공된다.

[0076] [0070] 또한, 본원에 공개된 후성적 변형 DNA-표적화 시스템 또는 본원에 공개된 DNA-표적화 시스템의 융합 단백질, 본원에 공개된 gRNA, 본원에 공개된 CRISPR Cas-gRNA 조합, 또는 앞서 언급한 것의 일부 또는 구성요소를 인코딩하는 복수의 폴리뉴클레오타이드가 제공된다.

[0077] [0071] 또한, 본원에 공개된 폴리뉴클레오타이드를 포함하는 벡터가 제공된다. 또한, 본원에 공개된 복수의 폴리뉴클레오타이드를 포함하는 벡터가 제공된다.

[0078] [0072] 본원의 임의의 구현예에서 벡터는 바이러스 벡터이다. 본원의 일부 구현예에서 벡터는 아데노-관련 바이러스(AAV) 벡터이다. 본원의 일부 구현예에서 벡터는 AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8 또는 AAV9 중에서 선택된다. 본원의 일부 구현예에서 벡터는 렌티바이러스 벡터이다. 본원의 일부 구현예에서 벡터는 비바이러스 벡터이다. 본원의 일부 구현예에서 비바이러스 벡터는 지질 나노입자, 리포좀, 엑소좀 또는 세포 침투 펩타이드 중에서 선택된다. 본원의 임의의 구현예에서 벡터는 B형 간염 바이러스에 감염된 세포에 대한 향성(tropism)을 나타낸다. 본원의 임의의 구현예에서 벡터는 하나의 벡터 또는 두 개 이상의 벡터를 포함한다.

[0079] [0073] 또한, 본 발명은 B형 간염 바이러스 서열의 표적 영역 내에서 후성적 변형을 촉진하는 방법을 제공하며, 이 방법은 B형 간염 바이러스 서열을 포함하는 HBV에 감염된 세포에 표적 영역 내의 표적 부위를 표적으로 하는 후성적 변형 DNA-표적화 시스템을 도입하는 단계를 포함한다.

[0080] [0074] 또한, 본 발명은 B형 간염 바이러스 서열 내의 표적 영역 내에서 CpG 메틸화를 증가시키는 방법을 제공하며, 이 방법은 B형 간염 바이러스 서열을 포함하는 HBV에 감염된 세포에 표적 영역 내의 표적 부위를 표적으로 하는 후성적 변형 DNA-표적화 시스템을 도입하는 단계를 포함한다.

[0081] [0075] 또한, 본 발명은 B형 간염 바이러스 서열의 표적 영역에 대한 후성적 변형을 촉진하는 방법을 제공하며, 이 방법은 B형 간염 바이러스 서열을 포함하는 HBV 감염 세포에 본 명세서에 개시된 후성적 변형 DNA-표적화 시스템, 본 명세서에 개시된 gRNA, 본 명세서에 개시된 CRISPR Cas-gRNA 조합, 본 명세서에 개시된 폴리뉴클레오타이드, 본 명세서에 개시된 복수의 폴리뉴클레오타이드, 본 명세서에 개시된 벡터, 또는 상기의 일부 또는 구성요소를 도입하는 단계를 포함한다.

[0082] [0076] 또한 본 발명은 B형 간염 바이러스 서열의 표적 영역의 CpG 메틸화를 증가시키는 방법을 제공하며, 이 방법은 B형 간염 바이러스 서열을 포함하는 HBV 감염 세포에 본 명세서에 개시된 후성적 변형 DNA-표적화 시스템, 본 명세서에 개시된 gRNA, 본 명세서에 개시된 CRISPR Cas-gRNA 조합, 본 명세서에 개시된 폴리뉴클레오타이드, 본 명세서에 개시된 폴리뉴클레오타이드의 복수, 본 명세서에 개시된 벡터, 또는 상기의 일부 또는 구성요소를 도입하는 단계를 포함한다.

[0083] [0077] 본원의 임의의 구현예에서, 표적 영역은 서열번호 650의 뉴클레오타이드 위치를 참조로 B형 간염 바이러스 서열의 1033bp-1749bp에 해당하는 서열 내의 연속된 뉴클레오타이드 서열을 포함한다.

[0084] [0078] 또한, 본 발명은 B형 간염 바이러스 서열을 포함하는 HBV에 감염된 세포에서 하나 이상의 유전자의 전사를 감소시키는 방법을 제공하며, 이 방법은 서열번호 650의 뉴클레오타이드 위치를 참조로 B형 간염 바이러스 서열 내에서 표적 CpG 메틸화를 유도하는 후성적 변형 DNA-표적화 시스템을 세포에 도입하는 단계를 포함한다.

[0085] [0079] 또한, 본 발명은 HBV에 감염된 세포에서 B형 간염 바이러스 감염을 감소시키는 방법을 제공하는데, 이는 B형 간염 바이러스 서열을 포함하는 세포에 서열번호 650의 뉴클레오타이드 위치를 참조로 B형 간염 바이러스 서열의 표적 영역 내에서 표적 CpG 메틸화를 유도하는 후성적 변형 DNA-표적화 시스템을 도입하는 단계를 포함한다.

[0086] [0080] 본원의 임의의 구현예에서, 후성적 변형 DNA-표적화 시스템은 (a) B형 간염 바이러스 DNA 서열의 표적 부위를 표적으로 하기 위한 DNA-결합 도메인; 및 (b) DNA 메틸트랜스퍼라제 효과기 도메인을 포함하는 적어도 하나의 효과기 도메인을 포함하는 융합 단백질을 포함하는 적어도 하나의 DNA-표적화 모듈을 포함한다.

[0087] [0081] 본원의 임의의 구현예에서, CpG 메틸화 영역은 표적 영역의 500개 염기쌍 내에 있다. 본원의 임의의 구현예에서, 도입은 대상의 생체 내 또는 생체 외에서 발생한다.

[0088] [0082] 본원의 임의의 구현예에서, 세포는 포유류 세포이다. 본원의 임의의 구현예에서, 세포는 인간 세포이다. 본원의 임의의 구현예에서, 세포는 통합된 HBV DNA를 포함한다. 본원의 임의의 구현예에서, 세포는 에피솜 HBV cccDNA 풀을 포함하는 간세포이다. 일부 구현예에서, 간세포는 HBV 단백질을 발현하는데, 여기서 HBV 단백질은 HBsAg, HBeAg 또는 HBcrAg이거나, 상기의 조합이다.

[0089] [0083] 또한, 본 발명은 B형 간염 바이러스 감염을 감소시키는 방법을 제공하는데, 이 방법은 B형 간염에 감염된 대상에게 B형 간염 바이러스 서열의 표적 영역 내에서 CpG 메틸화를 증가시키는 후성적 변형 DNA-표적화 시스템을 투여하는 단계를 포함하며, 여기서 후성적 변형 DNA-표적화 시스템은 (a) B형 간염 바이러스 DNA 서열의 표적 부위를 표적으로 하기 위한 DNA-결합 도메인; 및 (b) DNA 메틸트랜스퍼라제 효과기 도메인을 포함하는 적어도 하나의 효과기 도메인을 포함한다.

[0090] [0084] 본원의 임의의 구현예에서, 표적 영역은 HBV 게놈의 CpG를 포함하는 영역이다. 본원의 임의의 구현예에서, 표적 영역은 서열번호 650의 뉴클레오타이드 위치를 참조로 B형 간염 바이러스 서열에서 67bp - 392bp, 1033bp - 1749bp 또는 2215bp - 2490bp에 해당하는 서열 내의 뉴클레오타이드의 연속된 서열을 포함한다. 본원의 임의의 구현예에서, 표적 영역은 서열번호 650의 뉴클레오타이드 위치를 참조로 B형 간염 바이러스 서열에서 1033bp - 1749bp에 해당하는 서열 내의 뉴클레오타이드의 연속된 서열을 포함한다. 본원의 임의의 구현예에서, 표적 영역은 B형 간염 단백질(HBx) 시작 코돈 상류의 300개 염기쌍 내에 위치한다. 본원의 임의의 구현예에서, 표적 영역은 HBx 기저 코어 프로모터 영역 내에 있다. 본원의 임의의 구현예에서, 표적 영역은 HBx 프로모터/인핸서 영역 내에 있다. 본원의 임의의 구현예에서, 표적 영역은 B형 간염 X 단백질(HBx) 시작 코돈 상류 250개 염기쌍 내에 있다. 본원의 임의의 구현예에서, 표적 영역은 서열번호 650의 뉴클레오타이드 위치를 참조로 B형 간염 바이러스 서열에서 1060bp-1480bp에 해당하는 서열 내의 연속적인 뉴클레오타이드 서열을 포함한다. 본원의 임의의 구현예에서, 표적 영역은 B형 간염 X 단백질(HBx) 시작 코돈 상류 150개 염기쌍 내에 있다. 본원의 임의의 구현예에서, 표적 영역은 B형 간염 X 단백질(HBx) 시작 코돈 상류 120개 염기쌍 내에 있다. 본원의 임의의 구현예에 있어서, 표적 영역은 서열번호 650의 뉴클레오타이드 위치를 참조로 B형 간염 바이러스 서열의 1250bp-1374bp에 해당하는 서열 내의 연속적인 뉴클레오타이드 서열을 포함한다. 일부 구현예에 있어서, 표적 영역은 서열번호 1068에 제시된 서열을 갖는다. 본원의 임의의 구현예에서, 표적 영역은 서열번호 650의 뉴클레오타이드 위치를 참조로 B형 간염 바이러스 서열의 1260bp-1300bp에 해당하는 서열 내의 연속적인 뉴클레오타이드 서열을 포함한다. 일부 구현예에 있어서, 표적 영역은 서열번호 1070에 제시된 서열을 갖는다.

[0091] [0085] 본원의 임의의 구현예에서, DNA-결합 도메인은 (a) Cas 단백질 또는 이의 변이체; 및 (b) 적어도 하나의 gRNA를 포함하는 규칙적인 간격을 갖고 나타나는 짧은 회문 구조의 반복 서열 관련 (Cas)-가이드 RNA (gRNA) 조합; 아연 집게 단백질 (ZFP); 전사 활성화 인자 유사 효과기 (TALE); 메가뉴클레아제; 호밍 엔도뉴클레아제; 또는 I-SceI 효소 또는 이의 변이체를 포함하고, 선택적으로 DNA-결합 도메인은 상기 중 임의의 것의 촉매적으로 비활성인 변이체를 포함한다.

[0092] [0086] 본원의 임의의 구현예에서, 해당 방법은 다음을 포함하는 CRISPR Cas-가이드 RNA(gRNA) 조합을 포함한다: (a) 규칙적인 간격을 갖고 나타나는 짧은 회문 구조의 반복 서열 연관 (Cas) 단백질 또는 이의 변이체; 및 (b) B형 간염 바이러스 DNA 서열 내의 표적 부위에 Cas 단백질 또는 이의 변이체를 표적으로 하는 청구항 165-202 중 어느 한 항에 따른 적어도 하나의 gRNA.

[0093] [0087] 본 발명의 임의의 구현예에서, 표적 부위 또는 각각의 표적 부위는 서열번호 5, 6, 12, 18, 22, 26, 29, 38, 42, 43, 51, 56, 61, 63, 68, 72, 75, 79, 82, 84, 88, 89, 98, 99, 113, 116, 121, 124, 125, 118, 130, 133, 135, 138, 143, 150, 152, 155, 158, 164, 165, 175, 176, 182, 185, 189, 190, 중 어느 하나에 제시된

서열, 적어도 14개 뉴클레오타이드(nt)의 이의 연속 부분, 또는 상기 중 어느 하나의 상보적 서열을 포함한다. 본원의 구현예 중 어느 하나에서, 표적 부위 또는 각각의 표적 부위는 서열번호 12, 18, 20, 22, 26, 27, 46, 50, 63, 66, 73, 79, 185, 192t 중 어느 하나에 제시된 서열, 적어도 14개 뉴클레오타이드(nt)의 이의 연속 부분, 또는 상기 중 어느 하나의 상보적 서열을 포함한다. 본원의 구현예 중 어느 하나에서, 표적 부위 또는 각각의 표적 부위는 서열번호 22에 제시된 서열, 적어도 14개 뉴클레오타이드(nt)의 이의 연속 부분, 또는 상기 중 어느 하나의 상보적 서열을 포함한다. 본원의 임의의 구현예에서, gRNA 또는 각각의 gRNA는 서열번호 200, 201, 207, 217, 221, 224, 233, 237, 238, 246, 251, 256, 258, 263, 267, 274, 270, 277, 279, 283, 284, 293, 294, 308, 311, 313, 316, 319, 320, 325, 328, 330, 333, 338, 345, 347, 350, 353, 359, 360, 370, 371, 377, 380, 384, 385, 387 중 어느 하나에 제시된 서열을 포함하는 gRNA, 적어도 14개 뉴클레오타이드의 이의 연속 부분 또는 상기 중 어느 하나의 상보적 서열을 포함하며, 선택적으로 여기서 gRNA 또는 각각의 gRNA는 서열번호 395, 402, 408, 412, 416, 419, 428, 432, 433, 441, 446, 451, 453, 458, 462, 465, 469, 472, 474, 478, 479, 488, 489, 503, 506, 508, 511, 514, 515, 520, 523, 525 중 하나에 제시됨. 575, 528, 533, 540, 542, 545, 548, 554, 555, 565, 566, 572, 579, 580 또는 582 중 어느 하나에 제시된다. 본원의 모든 구현예에서, gRNA 또는 각각의 gRNA는 서열번호 207, 213, 215, 217, 221, 222, 241, 245, 258, 261, 268, 274, 380, 387 중 어느 하나에 제시된 서열, 적어도 14개 뉴클레오타이드의 이의 연속 부분 또는 상기 중 어느 하나의 상보적 서열을 포함하고, 선택적으로 gRNA 또는 각각의 gRNA는 서열번호 402, 408, 410, 412, 416, 417, 436, 440, 453, 456, 463, 469, 575, 582 중 어느 하나에 제시된다. 본원의 임의의 구현예에서, gRNA 또는 각각의 gRNA는 서열번호 217 중 어느 하나에 제시된 서열, 적어도 14개 뉴클레오타이드의 이의 연속 부분, 또는 상기 중 어느 하나의 상보적 서열을 포함하며, 선택적으로 gRNA 또는 각각의 gRNA는 서열번호 412 중 하나에 제시된다.

[0094] [0088] 본원의 임의의 구현예에서, 적어도 하나의 DNA-결합 도메인은 조작된 아연 집게 단백질(eZFP)을 포함한다. 본원의 임의의 구현예에서, 표적 부위는 서열번호 1045, 1046, 1052 중 어느 하나에 제시된 뉴클레오타이드 서열, 적어도 12nt의 이의 연속된 부분, 또는 상기 중 어느 하나의 상보적 서열을 포함한다. 본원의 임의의 구현예에서, 표적 부위는 서열번호 1045, 1046, 1052 중 어느 하나에 제시된 뉴클레오타이드 서열을 포함한다.

[0095] [0089] 본원의 임의의 구현예에서, 적어도 하나의 효과기 도메인은 DNA 메틸트랜스퍼라제이다. 본원의 임의의 구현예에서, 적어도 하나의 효과기 도메인은 DNA 메틸트랜스퍼라제와 헤테로크로마틴 유도 인자를 동원할 수 있는 억제자 도메인을 포함하거나, 선택적으로 헤테로크로마틴 유도 인자는 히스톤 메틸트랜스퍼라제를 포함한다. 본원의 임의의 구현예에서, 적어도 하나의 효과기 도메인은 DNA 메틸트랜스퍼라제와 히스톤 메틸트랜스퍼라제를 포함한다. 본원의 임의의 구현예에서, 적어도 하나의 효과기 도메인은 DNMT3A/L 도메인 또는 이의 변이체를 포함한다. 본원의 임의의 구현예에서, 적어도 하나의 효과기 도메인은 서열번호 604 및 607에 제시된 서열을 포함하는 효과기 도메인을 포함하거나, 그 일부 또는 상기 중 어느 것과 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 본원의 임의의 구현예에서, 적어도 하나의 효과기 도메인은 KRAB 도메인 또는 이의 변이체를 더 포함한다. 본원의 임의의 구현예에서, 적어도 하나의 효과기 도메인은 서열번호 590에 제시된 서열을 더 포함하거나, 그 일부 또는 상기 중 어느 것과 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 본원의 임의의 구현예에서, DNA-표적화 시스템은 DNMT3A/3L-d SpCas9-KRAB 도메인 또는 이의 변이체를 포함한다. 본원의 임의의 구현예에서, DNA-표적화 시스템은 서열번호 645에 제시된 서열, 이의 일부 또는 상기 중 어느 것과 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다.

[0096] [0090] 본원의 임의의 구현예에서, DNA-표적화 시스템은 서열번호 680에 제시된 서열, 이의 일부 또는 상기와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 핵산 서열을 포함한다. 본원의 임의의 구현예에서, DNA-표적화 시스템은 서열번호 680에 제시된 서열을 포함한다.

[0097] [0091] 또한, 본 발명에서는 B형 간염 바이러스에 감염된 세포에서 하나 이상의 유전자의 전사를 억제하는 방법이 제공되며, 이 방법은 B형 간염 바이러스에 감염된 세포에 본원에 공개된 후성적 변형 DNA-표적화 시스템, 본원에 공개된 gRNA, 본원에 공개된 CRISPR Cas-gRNA 조합, 본원에 공개된 폴리뉴클레오타이드, 본원에 공개된 복수의 폴리뉴클레오타이드, 본원에 공개된 벡터 또는 상기의 일부 또는 구성 요소를 도입하는 것을 포함한다. 일부 구현예에서, 하나 이상의 유전자는 DNA-표적화 시스템에 의해 후성적으로 변형된다. 일부 구현예에서, 하나 이상의 유전자의 전사는 이 방법을 거치지 않은 비교 가능한 세포와 비교하여 감소된다.

[0098] [0092] 본원의 임의의 구현예에서, 하나 이상의 유전자의 전사는 적어도 약 1.25배, 1.5배, 1.75배, 2.0배, 2.5

배, 2.75배, 3.0배, 3.5배, 3.75배, 4.0배, 4.5배, 4.75배, 5.0배, 5.25배, 5.5배, 5.75배, 6배 감소한다. 본원의 임의의 구현예에서, 하나 이상의 유전자의 전사를 억제하면 HBV 복제 및/또는 HBV 전사가 감소한다. 본원의 임의의 구현예에서, HBV에 감염된 세포는 포유류 세포이다.

[0099] [0093] 본원의 임의의 구현예에서, HBV에 감염된 세포는 인간 세포이다. 본원의 임의의 구현예에서, 세포는 통합된 HBV DNA를 포함한다. 본원의 임의의 구현예에서, 세포는 에피솜 HBV cccDNA 풀을 포함하는 간세포이다. 본원의 임의의 구현예에서, 간세포는 HBV 단백질을 발현하는데, 여기서 HBV 단백질은 HBsAg 및/또는 HBeAg이다. 본원의 임의의 구현예에서, HBV에 감염된 세포는 대상에게 존재한다.

[0100] [0094] 본원의 임의의 구현예에서, 대상은 인간이다. 본원의 임의의 구현예에서, 대상은 HBV 바이러스 감염을 갖는다. 본원의 임의의 구현예에서, 대상은 통합된 HBV DNA를 포함하는 간세포를 갖는다. 본원의 임의의 구현예에서, 대상은 에피솜 HBV cccDNA 풀을 포함하는 간세포를 갖는다. 본원의 임의의 구현예에서, 대상은 HBV 단백질을 발현하는 간세포를 갖는데, 여기서 HBV 단백질은 HBsAg, HBeAg 또는 HBcrAg 및 이들의 조합이다.

[0101] [0095] 본원의 임의의 구현예에서, 대상은 HBV 바이러스 감염과 관련된 질병, 병태 또는 장애를 가지고 있다. 본원의 임의의 구현예에서, 질병, 병태 또는 장애는 간 질환 또는 암이다. 본원의 임의의 구현예에서, 질병, 병태 또는 장애는 급성 간염, 만성 간염, 간부전 또는 간경변이다. 본원의 임의의 구현예에서, 질병, 병태 또는 장애는 암이고, 선택적으로 암은 간세포암이다.

[0102] [0096] 또한, 본 발명에서는 본원에 공개된 벡터를 포함하는 약학 조성물이 제공된다. 본 명세서에 있는 구현예 중 어느 하나에서, 벡터는 갈락토스의 아미노 당 유도체에 접합되고, 임의로 벡터는 N-아세틸갈락토사민(GalNAc) 부분에 접합된다.

[0103] [0097] 또한, 본 발명에 따라 본원에 공개된 후성적 변형 DNA-표적화 시스템 또는 본원에 공개된 융합 단백질, 본원에 공개된 gRNA, 본원에 공개된 CRISPR Cas-gRNA 조합, 본원에 공개된 폴리뉴클레오타이드, 본원에 공개된 복수의 폴리뉴클레오타이드, 본원에 공개된 벡터, 또는 상기의 일부 또는 구성요소를 포함하는 약학 조성물이 제공된다.

[0104] [0098] 또한, 본 발명에 따라 대상의 HBV 바이러스 감염을 치료하는 데 사용되는 약학 조성물이 제공된다. 본원의 모든 구현예에서, 대상은 HBV 바이러스 감염과 관련된 질병, 병태 또는 장애를 가지고 있다.

[0105] [0099] 또한, 본 발명은 HBV 바이러스 감염과 관련된 질병, 장애 또는 상태를 치료하는 데 사용하기 위한 약학 조성물을 제공한다.

[0106] [0100] 또한, 본 발명에 따라 대상의 HBV 바이러스 감염을 치료하기 위한 약물의 제조에 사용하기 위한 약학 조성물이 제공된다. 일부 구현예에서, HBV 바이러스 감염은 질병, 장애 또는 상태와 관련된다.

[0107] [0101] 또한, 본 발명은 HBV 바이러스 감염과 관련된 질병, 병태 또는 장애를 치료하기 위한 약물의 제조에 사용하기 위한 약학 조성물을 제공한다.

[0108] [0102] 본원에 개시된 임의의 구현예에서, 질병, 병태 또는 장애는 간 질환 또는 암이다. 본원에 개시된 임의의 구현예에서, 질병, 병태 또는 장애는 급성 간염, 만성 간염, 간부전 또는 간경변이다. 본원에 개시된 임의 구현예에서, 질병, 병태 또는 장애 암, 선택적으로 간세포암이다. 본원에 개시된 임의 구현예에서, 약학 조성물은 대상에게 생체 내로 투여된다.

[0109] [0103] 본원의 임의의 구현예에서, 약학 조성물을 투여한 후, 대상의 세포에서 하나 이상의 HBV 유전자의 전사가 억제된다. 본원의 임의의 구현예에서, 하나 이상의 HBV 유전자는 HBV 복제 및/또는 HBV 전사를 제어하는 데 관여한다. 본원의 임의의 구현예에서, 하나 이상의 유전자는 폴리머라제 유전자, S-패밀리 유전자, X-유전자 또는 코어 패밀리 유전자이다.

[0110] [0104] 또한, 본 발명은 질병, 병태 또는 장애를 필요로 하는 대상에게 이를 치료하는 방법을 제공하는데, 이는 본원에 공개된 후성적 변형 DNA-표적화 시스템, 본원에 공개된 gRNA, 본원에 공개된 CRISPR Cas-gRNA 조합, 본원에 공개된 폴리뉴클레오타이드, 본원에 공개된 복수의 폴리뉴클레오타이드, 본원에 공개된 벡터, 본원에 공개된 약학 조성물, 또는 상기의 일부 또는 성분을 대상에게 투여하는 단계를 포함한다.

[0111] [0105] 또한, 본 발명은 B형 간염 바이러스 감염을 감소시키는 방법을 제공하는데, 이 방법은 B형 간염 바이러스에 감염된 대상에게 본원에 공개된 후성적 변형 DNA-표적화 시스템, 본원에 공개된 gRNA, 본원에 공개된 CRISPR Cas-gRNA 조합, 본원에 공개된 폴리뉴클레오타이드, 본원에 공개된 복수의 폴리뉴클레오타이드, 본원에

공개된 벡터, 본원에 공개된 약학 조성물, 또는 상기의 일부 또는 성분을 투여하는 단계를 포함한다.

[0112] [0106] 또한, 본 발명에 따라 하나 이상의 HBV 유전자 또는 이의 조절 요소의 표적 부위에 결합하는 조작된 아연 집게 단백질(eZFP)이 제공되며, 여기서 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조하는 위치에 상응하는 HBV 게놈의 1033bp-1749bp에 걸친 표적 영역 내에 있다. 본 명세서의 모든 구현예에서, 표적 부위는 B형 간염 X 단백질(HBx) 시작 코돈 상류의 300개 염기쌍 내에 걸친 표적 영역 내에 있다. 본 명세서의 모든 구현예에서, 표적 부위는 HBx 기저 코어 프로모터 영역에 위치한다. 본 명세서의 모든 구현예에서, 표적 부위는 HBx 프로모터/인핸서 영역 내에 위치한다. 본 명세서의 모든 구현예에서, 표적 부위는 B형 간염 X 단백질(HBx) 시작 코돈 상류의 250개 염기쌍 내에 걸친 표적 영역 내에 있다. 본원의 임의의 구현예에서, 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조하는 위치에 상응하는 HBV 게놈의 1060-1480bp에 걸친 표적 영역 내에 있다. 본원의 임의의 구현예에서, 표적 부위는 B형 간염 X 단백질(HBx) 시작 코돈 상류의 150개 염기쌍 내에 걸친 표적 영역 내에 있다. 본원의 임의의 구현예에서, 표적 부위는 B형 간염 X 단백질(HBx) 시작 코돈 상류의 120개 염기쌍 내에 걸친 표적 영역 내에 있다. 본원의 구현예 중 어느 하나에서, 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 HBV 게놈의 1250-1374bp에 걸친 서열에 상응하는 표적 영역 서열 내에 있다. 일부 구현예에서, 표적 영역은 서열번호 1068에 제시된 서열을 갖는다. 본원의 구현예 중 어느 하나에서, 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 HBV 게놈의 1255-1302bp에 걸친 서열에 상응하는 표적 영역 서열 내에 있다. 일부 구현예에서, 표적 영역은 서열번호 1069에 제시된 서열을 갖는다. 본원의 구현예 중 어느 하나에서, 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 HBV 게놈의 1260-1300bp에 걸친 서열에 상응하는 표적 영역 서열 내에 있다. 일부 구현예에서, 표적 영역은 서열번호 1070에 제시되어 있다. 본원의 구현예 중 어느 하나에서, 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 HBV 게놈의 1255bp-1290bp에 걸친 서열에 상응하는 표적 영역 서열 내에 있다. 본원의 구현예 중 어느 하나에서, 표적 부위는 서열번호 1028-1055 중 어느 하나에 제시된 뉴클레오타이드 서열, 적어도 12nt의 이의 연속 부분, 또는 상기 중 어느 하나의 상보적 서열을 포함한다. 본원의 구현예 중 어느 하나에서, 표적 부위는 서열번호 1028-1055 중 어느 하나에 제시된 뉴클레오타이드 서열을 포함한다.

[0113] [0107] 본원의 임의의 구현예에서, 표적 부위는 서열번호 1045, 1046 또는 1052 중 어느 하나에 제시된 뉴클레오타이드 서열, 적어도 12nt의 이의 연속 부분, 또는 상기 중 어느 하나의 상보적 서열을 포함한다. 본원의 임의의 구현예에서, 표적 부위는 서열번호 1045, 1046 또는 1052 중 어느 하나에 제시된 뉴클레오타이드 서열을 포함한다.

[0114] [0108] 본원의 임의의 구현예에서, 아연 집게 단백질은 N-말단에서 C-말단의 순서로 F1 내지 F6으로 표시되는 6개의 아연 집게를 포함하고, 여기서 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같다: 1) F1:SEADRSR(서열번호 720) F2:DRSNLTR(서열번호 721) F3:QSSDLSR(서열번호 722) F4:YHWYLKK(서열번호 723) F5:RSDSLSV(서열번호 724) F6:QNANRKT(서열번호 725); 2) F1:RSDVLST(서열번호 726) F2:DNSSRTR(서열번호 727) F3:RPYTLRL(서열번호 728) F4:DSSHRTR(서열번호 729) F5:RSDHLSQ(서열번호 730) F6:DSSHRTR(서열번호 731); 3) F1:RSDHLSQ(서열번호 732) F2:QSADRTK(서열번호 733) F3:RSDHLSQ(서열번호 734) F4:RRSDLKR(서열번호 735) F5:RSDHLSR(서열번호 736) F6:QSSDLRR(서열번호 737); 4) F1:RSDNLSE(서열번호 738) F2:TSSNRKT(서열번호 739) F3:DRSHLTR(서열번호 740) F4:RSDALTQ(서열번호 741) F5:DRSALAR(서열번호 742) F6:RRFTLSK(서열번호 743); 5) F1:RSDHLSE(서열번호 744) F2:QYSGRYY(서열번호 745) F3:HGQTLNE(서열번호 746) F4:QSGNLAR(서열번호 747) F5:RSDSLLR(서열번호 748)F6:CREYRGK(서열번호 749); 6) F1:QSANRTT(서열번호 750) F2:RSANLTR(서열번호 751) F3:RSDVLSE(서열번호 752) F4:TSGHLSR(서열번호 753) F5:QSSDLSR(서열번호 754), F6:QWSTRKR(서열번호 755); 7) F1:QSGNLAR(서열번호 756) F2:ATCCLAH(서열번호 757) F3:RWQYLPT(서열번호 758) F4:DRSALAR(서열번호 759) F5:RSDNLSE(서열번호 760)F6:KRCNLRC(서열번호 761); 8) F1:NPANLTR(서열번호 762) F2:QNATRTK(서열번호 763) F3:QSGHLAR(서열번호 764) F4:NRHDRAK(서열번호 765) F5:RSDHLSE(서열번호 766), F6:QRRSRYK(서열번호 767); 9) F1:QSSDLSR(서열번호 768) F2:HRSTRNR(서열번호 769) F3:RSDVLSA(서열번호 770) F4:DSRTRKN(서열번호 771) F5:QSGSLTR(서열번호 772) F6:DQSGLAH(서열번호 773); 10) F1:QNPAQWR(서열번호 774) F2:RSADLSR(서열번호 775) F3:TSGSLSR(서열번호 776) F4:RSDHLSR(서열번호 777) F5:RSDSLLR(서열번호 778) F6:QSYDRFQ(서열번호 779); 11) F1:TSGSLSR(서열번호 780) F2:RSDHLSR(서열번호 781) F3:RSDSLLR(서열번호 782) F4:QSYDRFQ(서열번호 783) F5:RSDNLST(서열번호 784) F6:DNRDRIK(서열번호 785); 12) F1:DRSNLSR(서열번호 786) F2:LRQNLIM(서열번호 787) F3:ERGTLAR(서열번호 788) F4:RSDALTQ(서열번호 789) F5:RSDSLSQ(서열번호 790) F6:RKADRTR(서열번호 791); 13) F1:QYCCLTN(서열번호 792) F2:TSGNLTR(서열번호 793) F3:QSSDLSR(서열번호 794) F4:FRYYLKR(서열번호 795) F5:QSGDLTR(서열번호 796) F6:DKGNLTK(서열번호 797); 14) F1:TSGSLSR(서열번호 798) F2:RSDNLTT(서열번호 799) F3:QSGNLAR(서열번호 800) F4:DRTTLMR(서열번호 801) F5:QSGHLAR(서열번호 802) F6:QLTHLNS(서열

번호 803); 15) F1:IKHDLHR(서열번호 804) F2:RSANLTR(서열번호 805) F3:RSDNLAR(서열번호 806) F4:QNVSRPR(서열번호 807) F5:RSDDLSK(서열번호 808) F6:DSSHRTR(서열번호 809); 16) F1:RSDNLAR(서열번호 810) F2:QNVSRPR(서열번호 811) F3:RSDDLSK(서열번호 812) F4:DSSHRTR(서열번호 813) F5:TSSNRKT(서열번호 814) F6:AQWTRAC(서열번호 815); 17) F1:RSDDLSK(서열번호 816) F2:DSSHRTR(서열번호 817) F3:TSSNRKT(서열번호 818) F4:AQWTRAC(서열번호 819) F5:RKQTRTT(서열번호 820) F6:HRSSLRR(서열번호 821); 18) F1:QSAHRKN(서열번호 822) F2:TSSNRKT(서열번호 823) F3:RSDNLSA(서열번호 824) F4:RNNDRKT(서열번호 825) F5:TSGSLSR(서열번호 826) F6:QAGHLAK(서열번호 827); 19) F1:RSDHLSQ(서열번호 828) F2:ASSTRTK(서열번호 829) F3:RSDDLTR(서열번호 830) F4:QKSNLSS(서열번호 831) F5:QSANRTT(서열번호 832) F6:QNATRTK(서열번호 833); 20) F1:RSDTLSE(서열번호 834) F2:RRWTLVG(서열번호 835) F3:DRSNLSR(서열번호 836) F4:QSGDLTR(서열번호 837) F5:QSSDLSR(서열번호 838) F6:YHWYLKK(서열번호 839); 21) F1:RSANLAR(서열번호 840) F2:RSDNLRE(서열번호 841) F3:RPYTLRL(서열번호 842) F4:HRSNLNK(서열번호 843) F5:QSGSLTR(서열번호 844) F6:TSANLSR(서열번호 845); 22) F1:RSDDLVR(서열번호 846) F2:TSGSLVR(서열번호 847) F3:RSDKLVR(서열번호 848) F4:RSDELVR(서열번호 849) F5:TSHSLTE(서열번호 850) F6:RADNLTE(서열번호 851); 23) F1:ERSHLRE(서열번호 852) F2:TSHSLTE(서열번호 853) F3:QAGHLAS(서열번호 854) F4:TSHSLTE(서열번호 855) F5:DPGHLVR(서열번호 856) F6:TSGNLVR(서열번호 857); 24) F1:RADNLTE(서열번호 858) F2:TSGSLVR(서열번호 859) F3:RKDNLKN(서열번호 860) F4:QSSSLVR(서열번호 861) F5:RSDKLVR(서열번호 862) F6:DSGNLRV(서열번호 863);25) F1:QSSSLVR(서열번호 864) F2:QSGDLRR(서열번호 865) F3:RSDERKR(서열번호 866) F4:HRTTLTN(서열번호 867) F5:RSDHLTN(서열번호 868) F6:TSGELVR(서열번호 869); 26) F1:QSGDLRR(서열번호 870) F2:RSDERKR(서열번호 871) F3:HRTTLTN(서열번호 872) F4:RSDHLTN(서열번호 873) F5:TSGELVR(서열번호 874) F6:RSDDLVR(서열번호 875); 27) F1:QRAHLER(서열번호 876) F2:QLAHLRA(서열번호 877) F3:DPGHLVR(서열번호 878) F4:RRSACRR(서열번호 879) F5:RSDHLTT(서열번호 880) F6:QSSSLVR(서열번호 881); 28) F1:QSSNLVR(서열번호 882) F2:RSDDLVR(서열번호 883) F3:THLDLIR(서열번호 884) F4:TSGNLTE(서열번호 885) F5:RRSACRR(서열번호 886) F6:RNDTLTE(서열번호 887).

[0115] [0109] 본원의 임의의 구현예에서, 아연 집게 단백질은 N-말단에서 C-말단의 순서로 F1에서 F6까지 표시되는 6개의 아연 집게를 포함하고, 여기서 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같다: F1: QSAHRKN(서열번호 822) F2: TSSNRKT(서열번호 823) F3: RSDNLSA(서열번호 824) F4: RNNDRKT(서열번호 825) F5: TSGSLSR(서열번호 826) F6: QAGHLAK(서열번호 827).

[0116] [0110] 본원의 임의의 구현예에서, 아연 집게 단백질은 N-말단에서 C-말단의 순서로 F1에서 F6까지 표시되는 6개의 아연 집게를 포함하고, 여기서 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같다: F1: RSDHLSQ(서열번호 828) F2: ASSTRTK(서열번호 829) F3: RSDDLTR(서열번호 830) F4: QKSNLSS(서열번호 831) F5: QSANRTT(서열번호 832) F6: QNATRTK(서열번호 833).

[0117] [0111] 본원의 임의의 구현예에서, 아연 집게 단백질은 N-말단에서 C-말단의 순서로 F1에서 F6까지 표시되는 6개의 아연 집게를 포함하고, 여기서 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같다: F1:QSSSLVR(서열번호 864) F2:QSGDLRR(서열번호 865) F3:RSDERKR(서열번호 866) F4:HRTTLTN(서열번호 867) F5:RSDHLTN(서열번호 868) F6:TSGELVR(서열번호 869).

[0118] [0112] 또한, 본 발명은 하나 이상의 HBV 유전자 또는 이의 조절 요소의 표적 부위에 결합하는 eZFP를 제공하며, 여기서 아연 집게 단백질은 N-말단에서 C-말단의 순서로 F1 내지 F6으로 표시되는 6개의 아연 집게를 포함하고, 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같다: 1) F1:SEADRSR(서열번호 720) F2:DRSNLTR(서열번호 721) F3:QSSDLSR(서열번호 722) F4:YHWYLKK(서열번호 723) F5:RSDSLSV(서열번호 724) F6:QNANRKT(서열번호 725); 2) F1:RSDVLST(서열번호 726) F2:DNSSRTR(서열번호 727) F3:RPYTLRL(서열번호 728) F4:DSSHRTR(서열번호 729) F5:RSDHLSQ(서열번호 730) F6:DSSHRTR(서열번호 731); 3) F1:RSDHLSQ(서열번호 732) F2:QSADRTK(서열번호 733) F3:RSDHLSQ(서열번호 734) F4:RRSDLKR(서열번호 735) F5:RSDHLSR(서열번호 736) F6:QSSDLRR(서열번호 737); 4) F1:RSDNLSE(서열번호 738) F2:TSSNRKT(서열번호 739) F3:DRSHLTR(서열번호 740) F4:RSDALTQ(서열번호 741) F5:DRSALAR(서열번호 742) F6:RRFTLSK(서열번호 743); 5) F1:RSDHLSE(서열번호 744) F2:QYSGRYY(서열번호 745) F3:HGQTLNE(서열번호 746) F4:QSGNLAR(서열번호 747) F5:RSDSLLR(서열번호 748) F6:CREYRGK(서열번호 749); 6) F1:QSANRTT(서열번호 750) F2:RSANLTR(서열번호 751) F3:RSDVLSE(서열번호 752) F4:TSGHLSR(서열번호 753) F5:QSSDLSR(서열번호 754), F6:QWSTRKR(서열번호 755); 7) F1:QSGNLAR(서열번호 756) F2:ATCCLAH(서열번호 757) F3:RWQYLPT(서열번호 758) F4:DRSALAR(서열번호 759) F5:RSDNLSE(서열번호 760)F6:KRCNLRC(서열번호 761); 8) F1:NPANLTR(서열번호 762) F2:QNATRTK(서열번호 763) F3:QSGHLAR(서열번호 764) F4:NRHDRAK(서열번호 765) F5:RSDHLSE(서열번호 766), F6:QRRSRYK(서열번호 767); 9) F1:QSSDLSR(서열번

호 768) F2:HRSTRNR(서열번호 769) F3:RSDVLSA(서열번호 770) F4:DSRTRKN(서열번호 771) F5:QSGSLTR(서열번호 772) F6:DQSGLAH(서열번호 773); 10) F1:QNPAQWR(서열번호 774) F2:RSADLSR(서열번호 775) F3:TSGSLSR(서열번호 776) F4:RSDHLSR(서열번호 777) F5:RSDSLLR(서열번호 778) F6:QSYDRFQ(서열번호 779); 11) F1:TSGSLSR(서열번호 780) F2:RSDHLSR(서열번호 781) F3:RSDSLLR(서열번호 782) F4:QSYDRFQ(서열번호 783) F5:RSDNLST(서열번호 784) F6:DNRDRIK(서열번호 785); 12) F1:DRSNLSR(서열번호 786) F2:LRQNLIM(서열번호 787) F3:ERGTLAR(서열번호 788) F4:RSDALTQ(서열번호 789) F5:RSDSLSQ(서열번호 790) F6:RKADRTR(서열번호 791); 13) F1:QYCCLTN(서열번호 792) F2:TSGNLTR(서열번호 793) F3:QSSDLSR(서열번호 794) F4:FRYYLKR(서열번호 795) F5:QSGDLTR(서열번호 796) F6:DKGNLTK(서열번호 797); 14) F1:TSGSLSR(서열번호 798) F2:RSDNLTT(서열번호 799) F3:QSGNLAR(서열번호 800) F4:DRTTLMR(서열번호 801) F5:QSGHLAR(서열번호 802) F6:QLTHLNS(서열번호 803); 15) F1:IKHDLHR(서열번호 804) F2:RSANLTR(서열번호 805) F3:RSDNLAR(서열번호 806) F4:QNVSRPR(서열번호 807) F5:RSDDLSK(서열번호 808) F6:DSSHRTR(서열번호 809); 16) F1:RSDNLAR(서열번호 810) F2:QNVSRPR(서열번호 811) F3:RSDDLSK(서열번호 812) F4:DSSHRTR(서열번호 813) F5:TSSNRKT(서열번호 814) F6:AQWTRAC(서열번호 815); 17) F1:RSDDLSK(서열번호 816) F2:DSSHRTR(서열번호 817) F3:TSSNRKT(서열번호 818) F4:AQWTRAC(서열번호 819) F5:RKQTRTT(서열번호 820) F6:HRSSLRR(서열번호 821); 18) F1:QSAHRKN(서열번호 822) F2:TSSNRKT(서열번호 823) F3:RSDNLSA(서열번호 824) F4:RNNDRKT(서열번호 825) F5:TSGSLSR(서열번호 826) F6:QAGHLAK(서열번호 827); 19) F1:RSDHLSQ(서열번호 828) F2:ASSTRTK(서열번호 829) F3:RSDDLTR(서열번호 830) F4:QKSNLSS(서열번호 831) F5:QSANRTT(서열번호 832) F6:QNATRTK(서열번호 833); 20) F1:RSDTLSE(서열번호 834) F2:RRWTLVG(서열번호 835) F3:DRSNLSR(서열번호 836) F4:QSGDLTR(서열번호 837) F5:QSSDLSR(서열번호 838) F6:YHWYLKK(서열번호 839); 21) F1:RSANLAR(서열번호 840) F2:RSDNLRE(서열번호 841) F3:RPYTLRL(서열번호 842) F4:HRSNLNK(서열번호 843) F5:QSGSLTR(서열번호 844) F6:TSANLSR(서열번호 845); 22) F1:RSDDLVR(서열번호 846) F2:TSGSLVR(서열번호 847) F3:RSDKLVR(서열번호 848) F4:RSDELVR(서열번호 849) F5:TSHSLTE(서열번호 850) F6:RADNLTE(서열번호 851); 23) F1:ERSHLRE(서열번호 852) F2:TSHSLTE(서열번호 853) F3:QAGHLAS(서열번호 854) F4:TSHSLTE(서열번호 855) F5:DPGHLVR(서열번호 856) F6:TSGNLVR(서열번호 857); 24) F1:RADNLTE(서열번호 858) F2:TSGSLVR(서열번호 859) F3:RKDNLKN(서열번호 860) F4:QSSSLVR(서열번호 861) F5:RSDKLVR(서열번호 862) F6:DSGNLRV(서열번호 863); 25) F1:QSSSLVR(서열번호 864) F2:QSGDLRR(서열번호 865) F3:RSDERKR(서열번호 866) F4:HRTTLTN(서열번호 867) F5:RSDHLTN(서열번호 868) F6:TSGELVR(서열번호 869); 26) F1:QSGDLRR(서열번호 870) F2:RSDERKR(서열번호 871) F3:HRTTLTN(서열번호 872) F4:RSDHLTN(서열번호 873) F5:TSGELVR(서열번호 874) F6:RSDDLVR(서열번호 875); 27) F1:QRAHLER(서열번호 876) F2:QLAHLRA(서열번호 877) F3:DPGHLVR(서열번호 878) F4:RRSACRR(서열번호 879) F5:RSDHLTT(서열번호 880) F6:QSSSLVR(서열번호 881); 28) F1:QSSNLVR(서열번호 882) F2:RSDDLVR(서열번호 883) F3:THLDLIR(서열번호 884) F4:TSGNLTE(서열번호 885) F5:RRSACRR(서열번호 886) F6:RNDTLTE(서열번호 887).

[0119] [0113] 본원의 임의의 구현예에 있어서, 조작된 아연 집게 단백질은 서열번호 692-719 중 어느 하나에 제시된 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 본원의 임의의 구현예에 있어서, 조작된 아연 집게 단백질은 서열번호 888-915 중 어느 하나에 제시된 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩된다.

[0120] [0114] 또한, 본 발명은 하나 이상의 HBV 유전자 또는 이의 조절 요소의 표적 부위에 결합하는 조작된 아연 집게 단백질을 제공하며, 여기서 아연 집게 단백질은 N-말단에서 C-말단의 순서로 F1에서 F6까지 표시되는 6개의 아연 집게를 포함한다, F1:QSAHRKN(서열번호 822) F2:TSSNRKT(서열번호 823) F3:RSDNLSA(서열번호 824) F4:RNNDRKT(서열번호 825) F5:TSGSLSR(서열번호 826) F6:QAGHLAK(서열번호 827). 일부 구현예에서, 조작된 아연 집게 단백질은 서열번호 709에 제시된 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, 조작된 아연 집게 단백질은 서열번호 709 중 어느 하나에 제시된 서열을 포함한다. 일부 구현예에서, 조작된 아연 집게 단백질은 서열번호 905에 제시된 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩된다. 일부 구현예에서, 설계된 아연 집게 단백질은 서열번호 905 중 어느 하나에 제시된 서열에 의해 인코딩된다.

[0121] [0115] 또한, 본 발명은 하나 이상의 HBV 유전자 또는 이의 조절 요소의 표적 부위에 결합하는 조작된 아연 집게 단백질을 제공하며, 여기서 아연 집게 단백질은 N-말단에서 C-말단의 순서로 F1에서 F6까지 표시되는 6개의 아연 집게를 포함하고, 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같다: F1:RSDHLSQ(서열번호 828) F2:ASSTRTK(서열번호 829) F3:RSDDLTR(서열번호 830) F4:QKSNLSS(서열번호 831) F5:QSANRTT(서열번호 832)

F6:QNATRTK(서열번호 833). 일부 구현예에서, 조작된 아연 집게 단백질은 서열번호 710에 제시된 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, 조작된 아연 집게 단백질은 서열번호 710 중 어느 하나에 제시된 서열을 포함한다. 일부 구현예에서, 조작된 아연 집게 단백질은 서열번호 906에 제시된 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩된다. 일부 구현예에서, 설계된 아연 집게 단백질은 서열번호 906 중 어느 하나에 제시된 서열에 의해 인코딩된다.

[0122] [0116] 또한, 본 발명은 하나 이상의 HBV 유전자 또는 이의 조절 요소의 표적 부위에 결합하는 조작된 아연 집게 단백질을 제공하며, 여기서 아연 집게 단백질은 N-말단에서 C-말단의 순서로 F1에서 F6까지 표시되는 6개의 아연 집게를 포함하고, 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같다: F1:QSSSLVR(서열번호 864) F2:QSGDLRR(서열번호 865) F3:RSDERKR(서열번호 866) F4:HRTTLTN(서열번호 867) F5:RSDHLTN(서열번호 868) F6:TSGELVR(서열번호 869). 일부 구현예에서, 조작된 아연 집게 단백질은 서열번호 716에 제시된 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, 조작된 아연 집게 단백질은 서열번호 716 중 어느 하나에 제시된 서열을 포함한다. 일부 구현예에서, 조작된 아연 집게 단백질은 서열번호 912에 제시된 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩된다. 일부 구현예에서, 설계된 아연 집게 단백질은 서열번호 9 내지 12 중 어느 하나에 제시된 서열에 의해 인코딩된다.

도면의 간단한 설명

[0123] [0117] 도 1은 각 가이드 RNA와 dCas9-KRAB 효과기 융합 단백질에 의해 매개되는 총 HBV RNA의 배수 변화(fold change)를 도시한다. 배수 변화는 모든 HBV 유전자형에 걸쳐 각 gRNA의 중간 표적화 위치와 관련하여 도시된다.

[0118] 도 2는 HBV 게놈 유형(HBV0-HBV12) 전체에 걸친 가이드 RNA의 보존을 나타낸다.

[0119] 도 3A-3B는 형질도입 5일 후 스크린에서 나온 상위 gRNA 후보(HBVg_192(서열번호 582), HBVg_17 (서열번호 407), HBVg_63(서열번호 453))에 의해 매개되는 총 HBV RNA(도 3A) 및 HBsAg(도 3B)의 억제를 나타낸다.

[0120] 도 4는 각 가이드 RNA와 DNMT3A/L-dCas9-KRAB-효과기 융합 단백질에 의해 매개되는 총 HBV RNA의 배수 변화를 도시한다. 배수 변화는 HBV 게놈 전체에서 각 gRNA의 중간 표적화 위치와 관련하여 도시된다.

[0121] 도 5는 형질도입 41일 후 및 61일 후 스크린에서 나온 상위 gRNA 후보군(HBVg_142(서열번호 532), HBVg_138(서열번호 528), HBVg_185(서열번호 575), HBVg_152(서열번호 542))에 의한 총 HBV RNA 억제를 나타낸다.

[0122] 도 6A-6B는 gRNA HBVg_63(서열번호 453), HBVg_185(서열번호 575) 및 HBVg_56(서열번호 446)에 의해 매개되는 총 HBV RNA의 지속적이고 안정적인 억제를 도시한다.

[0123] 도 7은 대표적인 gRNA와 에피-에디터의 조합으로 재투여한 후 정규화된 총 HBV RNA 발현을 보여준다.

[0124] 도 8은 진성 감염 모델에서 각 가이드 RNA와 DNMT3A/L-dSpCas9-KRAB 효과기 융합 단백질을 매개로 한 3.5킬로베이스(kb) HBV RNA의 배수 변화를 도시한다. 배수 변화는 HBV 게놈 전체에서 각 gRNA의 중간 표적화 위치와 관련하여 도시된다.

[0125] 도 9는 두 개의 cccDNA 스크린 간의 상관관계 플롯을 사용하여 감염 간 일관성을 보여준다.

[0126] 도 10은 통합된 HBV, cccDNA 또는 통합된 HBV와 cccDNA 모두를 억제하는 gRNA를 보여준다.

[0127] 도 11은 HBV cccDNA를 표적으로 삼고 DNA를 통합하여 복수 유전자와 그 조절 요소에서 다중화된 HBV 억제를 달성한 단일 gRNA를 나열한다.

[0128] 도 12는 복수 유전자와 그 조절 요소에서 다중화된 HBV 억제를 달성한 단일 gRNA가 표적으로 삼은 서열을 나열한다.

[0129] 도 13은 각 가이드 RNA와 대표적인 에피-에디터에 의해 매개되는 총 HBV RNA의 배수 변화를 도시한다. 배수 변화는 HBV 게놈 전체에서 각 gRNA의 중간 표적화 위치와 관련하여 도시된다.

[0130] 도 14는 개별 gRNA에 의한 표적 HBV cccDNA 전사체의 억제를 보여준다.

[0131] 도 15A-15B는 1차 인간 간세포(PHH) 감염 모델에서 cccDNA의 억제를 나타낸다. 도 15A는 두 PHH 기증자의 HBV RNA에 대한 HBVg_22(서열번호 412)에 의한 억제를 보여준다. 도 15B는 2회의 HBV 투여로 감염된 PHH 세포에서 HBVg_22(서열번호 412)에 의한 억제를 보여준다.

[0132] 도 16A-16B는 HepG2.NTCP(도 16A) 및 PxB PHH(도 16B) 모델에서 dSpCas9-KRAB 단독(서열번호 595) 또는 DNMT3A/L-dSpCas9-KRAB("D3AL-K"; 서열번호 645) 융합과 조합된 HBVg_22(서열번호 412)에 의해 매개되는 억제를 비교한 도면이다.

[0133] 도 17은 Hep3B 세포의 HBV RNA 내 다양한 영역에 대한 다중 표적 전사 억제를 나타낸다.

[0134] 도 18은 PLC/PRF/5(Alexander) 세포 모델에서 개별 gRNA 또는 다중화된 gRNA에 의해 매개되는 억제를 보여준다.

[0135] 도 19는 PXB 1차 인간 간세포(PHH) 세포 모델에서 대표적인 dSpCas9 효과기와 두 개의 gRNA를 조합한 다중화 접근 방식을 보여준다.

[0136] 도 20A 및 도 20B는 cccDNA와 통합된 HBV DNA의 메틸 캡처 시퀀싱 분석을 보여준다.

[0137] 도 21은 다양한 gRNA와 함께 DNMT3A/L-dSpCas9-KRAB를 인코딩하는 mRNA를 전달한 후의 메틸화 패턴의 증가를 보여준다.

[0138] 도 22는 HBV에 감염된 PHH 기증자에 DNMT3A/L-dCas9-KRAB 및 HBVg_22(서열번호 412)를 인코딩하는 mRNA를 전달한 후의 지속적인 CpG 섬 2 메틸화 패턴을 보여준다.

[0139] 도 23A 및 도 23B는 유전자 발현에 대한 gRNA-의존적 및 에피-에디터-의존적 변화의 RNA 시퀀싱 분석을 보여준다. 도 23C는 지질 단독의 대조군과 비교하여 gRNA에 의해 매개된 차별적으로 발현된 유전자의 최소한의 변화를 보여준다. 도 23D는 어떠한 복용량이나 시점에서도 비표적화 gRNA와 HBVg_22(서열번호 412) 사이에 차별적으로 발현된 유전자가 없음을 보여준다.

[0140] 도 24는 인간 키메라 간 마우스에 대한 *생체 내* 연구의 개략적 표현을 보여준다.

[0141] 도 25는 투여 5일 후(D5)에 gRNA HBVg_22(서열번호 412)와 결합된 대표적인 DNMT3A/L-dCas9-KRAB 융합 단백질에 의해 매개되는 억제를 보여준다. 억제는 지질 나노입자 투여 2일 전(투여 전)과 투여 5일 후(D5)에 HBsAg 단백질 수치와 HBV DNA를 모니터링하여 측정하였다.

[0142] 도 26은 투여 후와 투여 전 지표의 배수 변화를 나타낸다.

[0143] 도 27A-27D는 HBVg_22(서열번호 412)를 단독으로 또는 HBVg_185(서열번호 575)와 다중화하여 재투여하거나, GalNAc와 접합된 LNP를 통해 전달하거나, dCas9-KRAB(도 27A-27C) 또는 DNMT3A/L-dCas9-KRAB(도 27D)를 인코딩하는 mRNA와 조합하여 전달한 후 FRG 마우스에서의 억제를 보여준다.

[0144] 도 28A-28C는 GalNAc 접합 PEG LNP를 전달한 후의 인간 키메라 FRG 마우스 연구를 보여준다. 도 28A는 33일차에 DNMT3A/3L-dSpCas9-KRAB 및 HBVg_22(서열번호 412)를 인코딩하는 mRNA를 포함하는 LNP를 단회 투여한 후의 안정적인 억제를 보여준다. 도 28B는 억제 정도에 대한 단일 마우스 추적을 보여준다. 도 28C는 비표적화 gRNA를 받은 마우스와 비교했을 때 HBVg_22(서열번호 412)를 전달한 마우스에서 pgRNA 신호가 현저히 감소한 조직 샘플을 보여준다.

[0145] 도 29A-29B는 KRAB 에피-에디터와 아연 집게 단백질(ZFP)을 포함하는 각 융합 단백질에 의해 매개되는 총 HBV RNA의 배수 변화를 나타낸다. ZFP-KRAB 융합 단백질은 두 벳치(도 29A 및 도 29B)에서 Hep3B 세포에서 HBV 억제에 대해 스크리닝되었다.

[0146] 도 30은 Hep3B 세포에서 ZFP-KRAB 융합 단백질과 dCas9-KRAB 융합 단백질을 gRNA와 함께 조합하여 유도된 억제 간의 예비 비교를 보여준다.

[0147] 도 31A-31B는 X-프로모터(HBx) 부위를 따라 각 ZFP의 표적화 위치를 보여준다. 도 31A 에서 강조된 영역은 가장 효과적인 gRNA(HBVg_22 및 HBVg_63)와 가장 효과적인 ZFP-KRAB 융합 단백질(eZFP_18, eZFP_19 및 eZFP_25)이 표적화하는 HBx 프로모터 내의 영역을 나타낸다. 도 31B는 gRNA HBVg_22 및 HBVg_63과 ZFP-KRAB 융합 단백질 eZFP_18, eZFP_19 및 eZFP_25가 표적화하는 HBx 영역의 부위를 도시한다.

[0148] 도 32A-32B는 HepG2.NTCP 세포에서 각 eZFP-KRAB 융합 단백질을 매개로 한 총 HBV RNA의 배수 변화를 나타낸다.

[0149] 도 33은 HBV 아형(subtype)에 걸쳐 융합 단백질의 보존과 함께 각 eZFP-KRAB 융합 단백질을 매개로 한 총 HBV RNA의 변화율을 보여준다.

[0150] 도 34는 형질도입 4일 후 및 15일 후 각 DNMT3A/L-eZFP-KRAB 융합 단백질을 매개로 한 총 HBV RNA의 변화를 나타낸다.

[0151] 도 35는 eZFP-KRAB 융합 단백질을 매개로 한 차등 발현 유전자의 변화가 지질 단독 대조군에 비해, 최소한으로 나타남을 보여준다.

[0152] 도 36A-36B는 GFP(도 36A) 또는 비표적화(NT) 대조군(도 36B)과 비교하여 유전자 발현에 대한 eZFP-KRAB 융합 단백질 의존적 변화의 RNA 시퀀싱 분석을 보여준다.

발명을 실시하기 위한 구체적인 내용

[0124] **상세한 설명**

[0125] [0153] B형 간염은 B형 간염 바이러스(HBV)로 인해 발생하는 생명을 위협할 수 있는 간 감염이다. HBV 감염은 만성 간 감염을 유발하고 간경변과 간암의 위험을 증가시키는 세계적인 공중 보건 문제이다. WHO는 2019년에 전 세계적으로 2억 9,600만 명, 미국에는 100만 명이 만성 B형 간염 감염을 앓고 있으며 매년 150만 건의 새로운 감염이 발생한다고 추정하였다. 2019년에 B형 간염은 대부분 간경변과 간세포암(원발성 간암)으로 인해 약 82만 명이 사망한 것으로 추정된다.

[0126] [0154] HBV는 작은 외피 간영양 DNA 바이러스 계열인 헤파데나비리대(Hepadnaviridae) 계열에 속한다(Wei L. 및 Ploss A. Nature communications 12(1591) 1-13 (2021)). HBV 바이러스 입자는 약 3.2kb의 컴팩트하고 부분적으로 이중 가닥인 이완된 원형 DNA(rcDNA) 게놈을 포함한다. 게놈에는 4개의 병변, 즉: 공유 결합된 HBV 폴리머라제와 마이너스 가닥의 5'-말단에 있는 10개 뉴클레오타이드(nt) DNA 플랩; 플러스 가닥에 있는 5'-캡 RNA 프라이머와 단일 가닥 DNA(ssDNA) 갭이 그것이다. HBV 게놈은 약 3.2킬로베이스 이중 가닥 DNA 분자이지만 특정 HBV 균주에 따라 더 길거나 짧을 수 있다(예컨대 크기가 최대 3300bp 이상). 예시적인 HBV 게놈은 B형 간염 바이러스 게놈(B형 간염 바이러스 아형 ayw, 완전 게놈, GenBank: U95551.1), 서열번호 650이다. 최소 10개의 유전형(A 내지 J)이 유전형 간 차이가 8%를 넘지 않는 것으로 확인되었다. HBV 유전형 A(A1 내지 A7), 유전형 B(B1 내지 B9), 유전형 C(C1 내지 16), 유전형 D(D1 내지 D8), 유전형 F(F1 내지 F4)로 분류되는 하위 유전형도 존재한다(Zhang et al. World J Gastroenterol., 2015, 22:126 내지 144). 하위 유전형 내에서 서열 동일성 차이는 약 4%에 불과하다. 제공된 시스템 및 방법은 특히 높은 서열 유사성을 감안할 때 복수의 HBV 게놈에 적용 가능한 것으로 이해된다. 본 명세서에서, 뉴클레오타이드 위치 번호에 대한 참조는 GenBank 접근 번호 U95551.1에 따라 기술된 HBV DNA 서열의 뉴클레오타이드(염기 쌍) 번호이며, 서열번호 650에 제시되어 있다. 통상의 기술자는 본 명세서에서 기술된 것과 같은 다른 HBV 게놈의 표적 부위 또는 염기 쌍 위치가 동일한 위치에 있지 않을 수 있지만, 그럼에도 불구하고 HBV 게놈 서열을 서열번호 650에 제시된 서열과 정렬하여 결정되는 것과 같이 상동 서열 또는 실질적으로 상동 서열(예컨대 1, 2 또는 3개의 불일치)일 수 있음을 이해할 것이다. 따라서 상응하는 위치 또는 위치들은 HBV 게놈 서열을 서열번호 650에 제시된 참조 서열과 정렬하여 쉽게 식별할 수 있다. B형 간염 바이러스는 원형 게놈을 포함하므로 본 명세서에서 선형 HBV DNA 서열의 뉴클레오타이드 위치 번호에 대한 참조는 선형 서열의 시작에 따라 몇 개의 뉴클레오타이드만큼 이동할 수 있다. 예를 들어, 서열번호 650 및 서열번호 1071에 제시된 서열은 동일한 서열이지만 선형 서열의 시작은 선형 서열의 시작의 차이로 인해 두 개의 뉴클레오타이드만큼 이동한다. HBV 게놈 서열의 다른 서열 사이에서 해당 서열 영역을 식별하는 것은 통상의 기술자의 수준 내에 있다.

[0127] [0155] HBV 라이프사이클에는 바이러스 유입, cccDNA 형성, 전사, 복제, 조립, 분비 및 통합과 같은 프로세스가 포함된다. 담즙산 수송체 NTCP11을 통해 바이러스가 간세포로 유입된 후, HBV rcDNA를 품은 바이러스 뉴클레오캡시드가 핵으로 운반된다. rcDNA가 방출되고 rcDNA의 네 가지 병변이 완전히 복구되어 초나선형 cccDNA 분자(미니염색체라고도 함)를 형성한다. 바이러스 복구 인자는 복구에 필요하지 않으며 cccDNA는 종종 TDP2, DNA 폴리머라제(POL)κ, POLα, DNA 리가제 1 및 3, 플랩 엔도뉴클레아제 1을 포함한 숙주 DNA 복구 머시너리에 의존한다. HBV는 cccDNA 전사 조절을 위해 숙주의 편재적이고 간이 풍부한 전사 인자를 활용한다. cccDNA는 만성 HBV 감염을 유발하는 주요 바이러스 저장소이며 모든 HBV 바이러스 전사체의 템플릿 역할을 한다. 숙주 내 HBV DNA의 또 다른 형태는 숙주 게놈에 안정적으로 통합된 HBV DNA이다(Zhao K. 등, Cell Press-The Innovation

1(2): 1-10(2020). 이중 가닥 선형 DNA(dslDNA)는 숙주 게놈으로 통합하기 위한 주요 기질이다. 바이러스 DNA와 세포 DNA 사이에 서열 상동성이 거의 없기 때문에 NHEJ DNA 복구 경로가 HBV DNA 통합을 위한 메커니즘으로 제안된다. HBV DNA 통합은 숙주 게놈 전체에서 이중 가닥이 끊어졌을 때 발생하며 통합된 HBV DNA에서 최대 200bp의 말단 삭제가 일반적이다. 환자 간에 특정 염색체 핫스팟이나 공통 반복 부위는 관찰되지 않았다. 종양 조직의 특정 게놈 부위에서 농축(enrichment)에 대한 증거가 있다(Sung W. 등, Nature Genetics 44(7):765-9(2012)). 자손 바이러스는 생성되지 않지만 통합된 HBV DNA는 바이러스 RNA 및 단백질을 생성할 수 있다. HBV DNA 통합은 정상 간 조직(30%)보다 간암 세포(84%)에서 더 자주 발생한다.

[0128] [0156] 현행 표준 치료에는 뉴클레오사이드 유사체(예컨대 라미부딘)와 PEG화 인터페론 요법이 포함된다. 뉴클레오사이드 유사체는 HBV 폴리머라제 활동을 억제하여 바이러스 복제를 감소시킨다. 그러나 장기간의 치료 기간, 바이러스 저항성 증가 및 돌연변이 균주의 출현으로 인해 뉴클레오사이드 요법의 효과가 감소하였다(Papatheodoridis GV et al., Am. J. Gastroenterol 97(7):1618-28 (2002). 단독으로 또는 뉴클레오사이드 유사체(예컨대 라미부딘)와 함께 사용되는 PEG화된 인터페론 요법은 바이러스 DNA 전사를 억제하는 것으로 테스트되었다. PEG화된 인터페론 요법은 선천적 면역 체계와 적응적 면역 체계에 다른 효과를 미치는 것으로 나타났으며, CD8 T 세포에 대한 효과가 현저히 감소하여 요법의 효능이 제한되었다(Micco L., et al., Journal of Hepatology 58(2): 225-233 (2013); Stelma F., et al., Journal f Infectious Disease 212(7):1042-51 (2015) marcellin P., et al., New England Journal of Medicine 351(12):1206-17 (2004)). 뉴클레오타이드 유사체도 PEG화된 치료법도 HBV 표면 항원(HBsAg)의 생성을 제거하거나 억제할 수 없으며, 이는 HBV 감염의 불량한 예후와 관련이 있다. HBsAg를 감소시켜 기능적 완치에 도달하는 데 중점을 둔 안티센스 올리고뉴클레오타이드(ASO) 및 siRNA 접근 방식을 포함한 다른 치료법(Billioud G. et al., Journal of Hepatology 64(4):781-9 (2015); Gane E. et al., Hepatology 74(4):1795-1808 (2021); Flisiak R. et al., Expert Opinion on Biology Therapy 18(6)609-617)은 HBsAg, HBeAg 및 HBV DNA 합성을 억제하는 데 유망한 것으로 나타났다. 그러나 이러한 치료법이 간 조직 재생에 미치는 기능적 이점은 불분명하다.

[0129] [0157] 현행의 항바이러스 요법은 세포질 HBV 게놈 복제를 억제하고 cccDNA 형태(이 형태는 HBV 복제 중간체 및 바이러스 지속성 저장소 역할을 함)를 직접 표적으로 삼지 않기 때문에 거의 치료할 수 없다(Yang G. et al., Theranostics 9(24):7345-58 (2019)). 핵산 분해 효소 또는 염기 편집자와 같은 게놈 엔지니어링 접근 방식은 감염을 기능적으로 치료하기 위해 cccDNA 풀의 제거 또는 돌연변이를 표적으로 삼는다. 그러나 이러한 핵산 분해 효소 기반 요법은 염색체 이상을 일으킬 가능성이 있으므로 선호되지 않으며, 이는 더 나은 HBV 치료법에 대한 필요성을 강조할 뿐이다.

[0130] [0158] 감염된 간세포에서 에피솜 cccDNA 풀의 지속성은 항-HBV 요법에 의한 완전한 제거에 있어 중요한 장애물로 남아 있다. cccDNA는 히스톤과 비히스톤에 의해 조립된 크로마틴 유사 cccDNA 미니크로모좀으로 핵에 축적된다. cccDNA는 비원래 상태로 인해 비정상적인 크로마틴 조절을 보인다. 예를 들어, cccDNA의 후성적 상태의 변화가 전사 활성을 결정하는 것으로 밝혀졌다(Yang G. et al., Theranostics 9(24):7345-58 (2019). 예를 들어, 숙주 핵소체 조립 장치(HAT1/CAF-1)는 cccDNA 조립에 기여하는 H4K5 및 H4K12 부위에서 히스톤 H4를 아세틸화한다. cccDNA의 히스톤에 있는 아세틸화 표시는 차례로 HBV 복제와 cccDNA 축적을 촉진한다. 이 전사 활성은 cccDNA에 활성화 후성적 표시가 있는지 없는지에 따라 크게 좌우된다; 억제적 히스톤 표시(예컨대 H3K27me3 및 H3K9me3)는 매우 미미한데, 이는 cccDNA에서 억제가 제한적임을 시사한다(Tropberger P. et al., PNAS, 112(42):E5715-E5724 (2015), Riviere L., et al., J Hepatol 15(00450):S0168-8278 (2015)).

[0131] [0159] 바람직한 임상 결과가 cccDNA 미니염색체 내의 주요 후성유전학적 특징과 관련이 있는 것으로 나타났다. 연구에 따르면 cccDNA에는 HBV의 거동과 관련된 메틸화 경향이 있는 CpG 섬이 포함되어 있는 것으로 밝혀졌다(Zhang Y., et al., PlosOne 9(10):e110442 (2014), Vivekanandan P, et al., Journal of infectious diseases, 199(9):1286-1291 (2009), Vivekanandan P, et al., Journal of Virology, 84(9):4321-4329 (2010), Vivekanandan P. et al., Journal of Viral hepatitis 15(2):103-107 (2008), Jain S., et al., Scientific Reports 5: 10478 (2015)). CpG 섬 II 및 III의 메틸화는 환자의 혈청 HBV DNA 및 HBsAg 역가의 낮은 수준과 상관 관계가 있는 것으로 나타났다(Zhang Y., et al., PlosOne 9(10):e110442 (2014)). HBV 유전자형, HBeAg 양성, 환자 연령 및 간 섬유화 단계는 cccDNA CpG 메틸화 상태와 상관 관계가 있는 것으로 밝혀졌다. 시험관 내 메틸화 연구 결과 CpG 섬 II 메틸화가 cccDNA 전사 및 그에 따른 바이러스 코어 DNA 복제를 현저히 감소시킬 수 있음이 추가로 확인되었는데(Zhang Y., et al., PlosOne 9(10):e110442 (2014)), 이는 cccDNA 조절을 위한 크로마틴의 중요성과 만성 HBV 감염 치료의 잠재적 표적임을 확립하는 것이다. 항바이러스제 및 IFN α와 같은 광범위한 후성적 변형제는 활성 히스톤 번역 후 변형을 감소시켜 cccDNA 전사를 전사적으로 하향조절

하는 것으로 알려져 있다(Tropberger P. et al., PNAS, 112(42):E5715-E5724 (2015), Belloni L, et al., Journal of Clinical investigation 122L529-537 (2012), Allweiss L., et al., Journal of Hepatology 60:500-507 (2014), Lucifora J., et al., Science 343: 1221-1228 (2014)).

[0132] [0160] 제공된 구현예는 cccDNA에 존재하는 유전자를 포함하여 하나 이상의 HBV 바이러스 유전자를 후성유전학적으로 침묵시키는 것이 HBV 감염을 치료하는 실행 가능한 치료적 접근 방식일 수 있다는 인식에 기초한다. 본 발명에 따라 cccDNA 형태, 이완된 원형 DNA(rcDNA) 형태 및 인간 게놈 DNA에 통합된 HBV의 정확한 후성유전학적 침묵을 통해 HBV로부터 감염을 개선하고 어떤 경우에는 잠재적으로 기능적 치료를 달성하는 접근 방식이 개시된다. 여기에서 설명하는 접근 방식은 높은 효능, 안전성 및 안정성을 보여준다. 일부 구현예에서, 접근 방식은 동일한 접근 방식으로 모든 형태의 HBV를 표적으로 삼고, 비돌연변이성 플랫폼을 활용하며, 하류 전사체가 아닌 전사 소스를 표적으로 삼는다. 메틸화는 세포 자손에 의해 유전될 수 있으므로, 에피-편집 접근 방식의 내구성은 HBV 감염 치료에 유망하다. 일부 구현예에서, 여기에서 설명하는 접근 방식은 바이러스 게놈의 여러 위치를 표적으로 삼아 HBV 변이체에 걸쳐 깊고 지속적인 반응을 보장한다. 일부 구현예에서, 후성적 접근법은 HBV 복제, HBV 전사 및 HBV DNA로부터 단백질 생성의 침묵을 초래한다. 제공된 구현예는 면역 재부팅이나 감염된 간세포 제거에 의존하지 않지만 직접적인 후성적 침묵(예컨대 HBV 억제)에 기반한다.

[0133] [0161] 본원에 제공된 구현예 중에는 하나 이상의 B형 간염 바이러스(HBV) 유전자 및/또는 그 조절 요소의 전사를 억제하기 위한 적어도 하나의 DNA-표적화 모듈을 포함하는 후성적-변형 DNA-표적화 시스템이 있으며; 여기서 적어도 하나의 DNA-표적화 모듈 각각은: (a) B형 간염 바이러스 DNA 서열의 표적 부위를 표적으로 하기 위한 DNA-결합 도메인; 및 (b) 적어도 하나의 전사 억제자 효과기 도메인을 포함하는 융합 단백질을 포함한다. 일부 구현예에서, 제공된 후성적-변형 DNA-표적화 시스템은 B형 간염 바이러스(HBV) 복제 및/또는 HBV 전사를 조절하는 복수의 다른 유전자 또는 이의 조절 요소의 다중화된 표적화 억제를 위한 것이다. 일부 구현예에서, 후성적-변형 DNA-표적화 시스템은 B형 간염 바이러스(HBV) 복제 및/또는 HBV 전사를 조절하는 복수의 유전자 또는 이의 조절 요소의 전사를 억제하기 위한 복수의 DNA-표적화 모듈을 포함한다. 일부 구현예에서, DNA-표적화 모듈은 ##(a) 규칙적인 간격을 갖고 나타나는 짧은 회문 구조의 반복 서열(CRISPR) 관련(Cas) 단백질 또는 그 변이체 및 적어도 하나의 전사 억제자 효과기 도메인을 포함하는 융합 단백질; 및 (b) 적어도 제1 gRNA 및 제2 gRNA를 포함하는 복수의 가이드 RNA(gRNA)를 포함한다. 일부 구현예에서, 제1 gRNA는 제1 유전자 또는 이의 조절 요소의 표적 부위를 표적으로 하고 제2 gRNA는 제2 유전자 또는 이의 조절 요소의 표적 부위를 표적으로 한다. 제1 및 제2 유전자 또는 이의 조절 요소는 B형 간염 바이러스 복제 및/또는 HBV 전사를 조절한다. 또한, DNA-표적화 시스템 또는 DNA-표적화 시스템의 융합 단백질을 인코딩하는 폴리뉴클레오타이드, 벡터 및 이를 포함하는 조성물이 본원에 제공된다.

[0134] [0162] 제공된 후성적 변형 DNA-표적화 시스템의 일부 구현예에서, DNA 결합 도메인은 뉴클레아제-비활성인 CRISPR 관련(Cas) 단백질 또는 이의 변이체, 예컨대 죽은 Cas(dCas, 예를 들어 dCas9)이고, DNA-표적화 시스템은 Cas와 복합체를 형성할 수 있는 적어도 하나의 gRNA를 더 포함한다. 일부 구현예에서, DNA-결합 도메인은 가이드 RNA[gRNA]와 복합체를 형성한 뉴클레아제-비활성 주기적으로 간격을 띠고 분포하는 짧은 회문구조 반복서열 관련(Cas) 단백질 또는 이의 변이체이다. 이러한 시스템에서, gRNA는 유전자의 표적 부위 또는 이의 조절 요소에 혼성화할 수 있는 스페이서 서열을 갖는다. 또한, 후성적 변형 DNA-표적화 시스템을 수반하거나 이와 관련된 Cas/gRNA 조합물, 폴리뉴클레오타이드, 조성물 및 방법을 포함한 관련 gRNA가 본원에 제공된다.

[0135] [0163] 제공된 후성적 변형 DNA-표적화 시스템의 일부 구현예에서, DNA-결합 도메인은 표적 부위에 대한 서열 특이적 결합을 위해 조작된 단백질 도메인이다. 예를 들어, 일부 구현예에서, DNA-결합 도메인은 본원에 기재된 바와 같이 아연 집게[ZFN]-기반 DNA-결합 도메인 또는 전사 활성인자 유사 효과기 DNA-결합 도메인이다.

[0136] [0164] 또한, 간 세포의 전사 또는 표현형을 조절하기 위해 후성적 변형 DNA-표적화 시스템을 사용하는 방법이 본원에 제공된다. 또한, HBV 복제 및/또는 단백질 수준을 억제하기 위해 후성적 변형 DNA-표적화 시스템을 사용하는 방법이 본원에 제공된다. 일부 구현예에서, 이 방법은 간염과 같은 HBV 감염을 치료하기 위한 요법에 사용될 수 있다.

[0137] [0165] 일부 구현예에서, 표적 부위는 공유적으로 폐쇄된 원형으로 존재한다. DNA(cccDNA), 이완된 원형 DNA(rcDNA) 및/또는 게놈 DNA에 통합된다. 일부 구현예에서, 표적 부위는 HBV 복제 및/또는 HBV 전사에 관여하는 유전자 또는 이의 조절 요소, 예를 들어 조절 요소 또는 코딩 영역 근처에 있다. 또한, 본 발명에서는 복수의 DNA-표적화 모듈과 멀티플렉싱되어 시스템이 이러한 유전자 또는 이의 조절 요소의 조합을 표적화할 수 있는 후성적 변형 DNA-표적화 시스템이 제공된다. 일부 구현예에서, DNA-표적화 시스템의 각 모듈은 다른 유전자의

전사를 억제한다. 또한, 본 발명에서는 HBV 복제 및/또는 전사를 감소시키기 위해 후성적 변형 DNA-표적화 시스템을 사용하는 방법이 제공된다. 일부 구현예에서, 이 방법은 간 질환(예컨대 간염), 암(예컨대 간세포암) 또는 HBV 감염(급성 또는 만성 간염)을 치료하는 데 사용될 수 있다.

[0138] [0166] 따라서 일부 구현예에서, DNA-표적화 시스템은 표적 방식으로 유전자의 전사를 조절, 예를 들어 감소 또는 억제할 수 있는 합성 전사 인자를 포함한다. 제공된 구현예에서, 제공된 후성적 변형 DNA-표적화 시스템은 유전자 및/또는 그 조절 요소 또는 복수의 유전자 및/또는 그 조절 요소의 전사를 감소시키고, 그에 따라 HBV 복제 및/또는 전사의 침묵을 촉진한다. 제공된 구현예는 감염된 환자에서 HBV를 치료하기 위한 다중 유전적 메커니즘을 표적화하는 데 사용될 수 있으며, 바이러스 저항성, 장기 치료와 관련된 비용 및 현재 조합 요법의 효능 저하를 피할 수 있다. 이 접근 방식은 바이러스 복제와 cccDNA 및 통합 HBV DNA 모두로부터의 전사를 감소시키고 현재 요법과 관련된 문제를 회피함으로써 HBV 감염 치료에 대한 실질적인 임상적 솔루션을 제공한다.

[0139] [0167] 본 출원에 언급된 특허 문서, 과학 기사 및 데이터베이스를 포함한 모든 간행물은 각각의 개별 간행물이 개별적으로 참조로 포함된 것과 동일한 정도로 모든 목적을 위해 그 전체가 참조로서 포함된다. 본원에 기재된 정의가 본원에 참조로 포함되는 특허, 출원, 공개 출원 및 기타 공개에 기재된 정의와 상반되거나, 다르게는 일치하지 않는 경우, 본원에 기재된 정의는 본원에 참조로 포함되는 정의보다 우선한다.

[0140] [0168] 본원에 사용된 섹션 제목은 구성 목적으로만 사용되고 기재된 주제를 제한하는 것으로 해석되어서는 아니된다.

[0141] **I. DNA-표적화 시스템**

[0142] [0169] 일부 구현예에서, (본원에서 표적 유전자로도 지칭되는) 적어도 하나의 유전자 또는 이의 DNA 조절 요소(예컨대 조절 요소)에서 표적 부위를 특이적으로 표적화할 수 있고, 적어도 하나의 유전자의 전사를 감소시킬 수 있는 DNA-표적화 시스템이 제공된다. 제공된 구현예에서, 표적화된 각각의 표적 유전자 또는 이의 조절 요소에 대해, DNA-표적화 시스템은 유전자 또는 이의 조절 요소 내의 표적 부위에 결합하는 DNA-결합 도메인을 포함한다. 일부 구현예에서, DNA-표적화 시스템은 유전자 또는 이의 조절 요소의 하나 이상의 DNA 염기를 후성적으로 변형시킬 수 있는 적어도 하나의 효과기 도메인을 추가로 포함하되, 후성적 변형은 유전자의 전사를 감소시킨다(예를 들어, DNA-표적화 시스템의 부재에 비해 유전자의 전사를 억제하거나 이의 전사를 감소시킨다). 그러므로 용어 DNA-표적화 시스템 및 후성적 변형 DNA-표적화 시스템은 본원에서 상호교환적으로 사용될 수 있다. 일부 구현예에서, DNA-표적화 시스템은 (a) 표적 부위로 표적화될 수 있는 적어도 하나의 DNA-결합 도메인 및 (b) 유전자의 전사를 감소시킬 수 있는 적어도 하나의 효과기 도메인을 포함한다. 예를 들어, 적어도 하나의 효과기 도메인은 전사 억제자 도메인이다.

[0143] [0170] 일부 구현예에서, DNA-표적화 시스템은 적어도 하나의 DNA-표적화 모듈을 포함하는데, 여기서 시스템의 각 DNA-표적화 모듈은 제공된 대로 표적 유전자 또는 이의 조절 요소에서 하나의 표적 부위를 독립적으로 표적화할 수 있는 DNA-표적화 시스템의 구성 요소이다. 일부 구현예에서, 각 DNA-표적화 모듈은 (a) HBV 복제 및/또는 HBV 전사를 조절하는 표적 유전자 또는 이의 조절 요소의 표적 부위를 표적으로 삼을 수 있는 DNA-결합 도메인 및 (b) 유전자의 전사를 감소시킬 수 있는 효과기 도메인을 포함한다.

[0144] [0171] 일부 구현예에서, DNA-표적화 시스템은 단일 유전자의 억제를 표적화하기 위한 단일 DNA-표적화 모듈을 포함한다. 일부 구현예에서, DNA-표적화 모듈은 (a) HBV 복제 및/또는 HBV 전사를 조절하는 표적 유전자 또는 이의 조절 요소의 표적 부위를 표적으로 삼을 수 있는 DNA-결합 도메인 및 (b) 유전자의 전사를 감소시킬 수 있는 효과기 도메인을 포함한다.

[0145] [0172] 일부 구현예에서, DNA-표적화 시스템은 하나 이상의 유전자 또는 이의 조절 요소의 억제를 표적화하기 위한 단일 DNA-표적화 모듈을 포함한다. 따라서 일부 구현예에서, 단일 DNA-표적화 모듈은 하나 이상의 유전자 또는 이의 조절 요소를 조절(예컨대 억제)하기 위해 표적화하는 다중화된 후성적 변형 DNA-표적화 시스템을 제공한다. 일부 구현예에서, DNA-표적화 모듈은 (a) HBV 복제 및/또는 HBV 전사를 조절하는 하나 이상의 표적 유전자 또는 이의 조절 요소의 표적 부위를 표적으로 삼을 수 있는 DNA-결합 도메인 및 (b) 유전자의 전사를 감소시킬 수 있는 효과기 도메인을 포함한다. 일부 구현예에서, DNA-표적화 시스템은 적어도 2개, 적어도 3개, 적어도 4개, 적어도 5개, 적어도 6개, 적어도 7개, 적어도 8개, 적어도 9개, 적어도 10개, 적어도 20개, 적어도 30개 이상의 유전자 또는 이의 조절 요소의 억제를 표적으로 삼기 위한 단일 DNA-표적화 모듈을 포함한다. 특정 구현예에서, DNA-표적화 모듈은 하나 이상의 유전자의 각 표적 부위에 교차 반응한다. 일부 구현예에서, 단일 DNA-표적화 모듈은 하나 이상의 유전자 또는 이의 조절 요소의 전사를 억제하는 다중화된 후성적 변형 DNA-표적

화 시스템을 제공한다. 일부 구현예에서, DNA-표적화 모듈은 적어도 2개, 적어도 3개, 적어도 4개, 적어도 5개, 적어도 6개, 적어도 7개, 적어도 8개, 적어도 9개, 적어도 10개, 적어도 20개, 적어도 30개 이상의 유전자 또는 이의 조절 요소의 전사를 억제한다.

[0146] [0173] 일부 구현예에서, DNA-표적화 시스템은 복수의 DNA-표적화 모듈을 포함하는 다중화된 DNA-표적화 시스템이며, 여기서 각 DNA-표적화 모듈은 하나 이상의 유전자 또는 이의 조절 요소의 다른 표적 부위를 표적화한다. 일부 구현예에서, 다른 표적 부위는 유전자 또는 이의 조절 요소의 동일한 영역에 있다. 일부 구현예에서, 상이한 표적 부위들이 프로모터와 같은 조절 요소에 존재한다. 일부 구현예에서는, 표적 부위들이 겹치므로, 임의의 두 개 이상의 DNA-표적화 모듈이 겹치는 표적 부위에 결합한다.

[0147] [0174] 일부 구현예에서, DNA-표적화 시스템은 복수의 DNA-표적화 모듈을 포함하는데, 여기서 각 DNA-표적화 모듈은 다른 유전자의 억제를 표적화하기 위한 것이다. 일부 구현예에서, DNA-표적화 시스템은 다중화된 DNA-표적화 시스템, 즉 두 개 이상의 유전자 또는 이의 조절 요소의 표적 부위를 표적화한다. DNA-표적화 시스템이라는 용어는 두 개 이상의 DNA-표적화 모듈을 포함하는 다중화된 후성적-변형 DNA-표적화 시스템을 포함할 수 있다. 일부 구현예에서, 다중화된 후성적-변형 DNA-표적화 시스템 내의 각 DNA-표적화 모듈은 다른 유전자 또는 이의 조절 요소의 표적 부위를 표적화하여 시스템의 다른 DNA-표적화 모듈과 다른 유전자를 각각 억제한다. 일부 구현예에서, 다중화된 후성적 변형 DNA-표적화 시스템 내의 각 DNA-표적화 모듈은 두 개 이상의 유전자 또는 이의 조절 요소의 표적화 부위를 표적화하여 두 개 이상의 유전자의 전사를 억제한다. 일부 구현예에서, 각 DNA-표적화 모듈은 적어도 2개, 적어도 3개, 적어도 4개, 적어도 5개, 적어도 6개, 적어도 7개, 적어도 8개, 적어도 9개, 적어도 10개, 적어도 20개, 적어도 30개 이상의 유전자의 전사를 억제한다.

[0148] [0175] 다중화된 후성적 변형 DNA-표적화 시스템은 적어도 2개, 적어도 3개, 적어도 4개, 적어도 5개, 적어도 6개, 적어도 7개, 적어도 8개, 적어도 9개, 적어도 10개, 적어도 20개, 적어도 30개의 DNA-표적화 모듈 또는 상기 중 임의의 값 사이를 포함한다. 일부 구현예에서, 다중화된 후성적 변형 DNA-표적화 시스템은 적어도 2개, 적어도 3개, 적어도 4개, 적어도 5개, 적어도 6개, 적어도 7개, 적어도 8개, 적어도 9개, 적어도 10개, 적어도 20개, 적어도 30개의 유전자의 전사를 억제한다.

[0149] [0176] 일부 구현예에서, DNA-표적화 시스템의 임의의 두 개의 DNA-표적화 모듈은 개별(즉, 비중첩) 성분을 포함한다. 일부 구현예에서, DNA-표적화 시스템의 DNA-표적화 모듈 각각은 개별(즉, 비중첩) 성분을 포함한다. 예를 들어, DNA-표적화 시스템은 제1 표적 부위를 표적화하는 DNA-결합 도메인(예를 들어, ZFN 또는 TALE-기반 DNA-결합 도메인)을 포함하는 제1 융합 단백질을 포함하는 제1 DNA-표적화 모듈, 및 제2 표적 부위를 표적화하는 제2 DNA-결합 도메인(예를 들어, ZFN 또는 TALE-기반 DNA-결합 도메인)을 포함하는 제2 융합 단백질을 포함하는 제2 DNA-표적화 모듈을 포함할 수 있다.

[0150] [0177] 일부 구현예에서, DNA-표적화 시스템의 임의의 두 개의 DNA-표적화 모듈은 공유(즉, 중첩) 성분을 포함할 수 있다. 일부 구현예에서, DNA-표적화 시스템의 DNA-표적화 모듈 각각은 공유(즉, 중첩) 성분을 포함한다. 예를 들어, DNA-표적화 시스템은 (a) Cas 단백질 및 전사 억제자 도메인을 포함하는 융합 단백질 및 (b) Cas 단백질과 복합체를 형성하고 제1 HBV 유전자 또는 이의 조절 요소의 제1 표적 부위를 표적화하는 제1 gRNA를 포함하는 제1 DNA-표적화 모듈, 및 (a) 제1 DNA-표적화 모듈의 융합 단백질, 및 (b) Cas 단백질과 복합체를 형성하고 제2 HBV 유전자 또는 이의 조절 요소의 제2 표적 부위를 표적화하는 제2 gRNA를 포함하는 제2 DNA-표적화 모듈을 포함할 수 있다. 주어진 Cas 단백질에 대해 두 개 이상의 상이한 gRNA를 제공하는 것은 동일한 Cas 단백질의 상이한 분자가 두 개 이상의 gRNA의 표적 부위로 표적화될 수 있게 한다는 것이 이해될 것이다. 반대로, 상이한 Cas 단백질 변이체(예를 들어, SpCas9 및 SaCas9)는 본원에 기재된 바와 같이 상이한 gRNA 스캐폴드 서열 및 PAM과 양립가능하다. 따라서, 다수의 비중첩 CRISPR/Cas-기반 DNA-표적화 모듈을 포함하는 단일 DNA-표적화 시스템을 조작하는 것이 가능하다.

[0151] [0178] 일부 측면에서, 본 발명에서는 HBV 복제 및/또는 HBV 전사를 조절하는 복수의 유전자의 전사를 억제하기 위한 복수의 DNA-표적화 모듈을 포함하는 후성적 변형 DNA-표적화 시스템이 제공된다. 일부 구현예에서, 복수의 DNA-표적화 모듈은 복수의 유전자 또는 이의 조절 요소의 제1 유전자의 전사를 억제하기 위한 제1 DNA-표적화 모듈과 복수의 유전자의 제2 유전자의 전사를 억제하기 위한 제2 DNA-표적화 모듈을 포함한다. 일부 구현예에서, 각 DNA-표적화 모듈은: (a) 복수의 유전자 중 하나의 표적 부위를 표적으로 하기 위한 DNA-결합 도메인, 및 (b) 적어도 하나의 전사 억제 도메인을 포함하는 융합 단백질을 포함한다. 일부 구현예에서, 표적 부위는 HBV 유전자 또는 이의 조절 요소에 있거나 그 근처에 있다. 일부 구현예에서, HBV 유전자 또는 이의 조절 요소는 HBV 복제 및/또는 HBV 전사를 제어하는 데 관여한다. 조절 요소는 프로모터 영역(예컨대 pre-S1 프로모

터, pre-S2 프로모터, X 프로모터 또는 기저 코어 프로모터), 인핸서 영역(예컨대 Enh1 또는 Enh2 인핸서 영역) 또는 기타 전사 처리 제어 영역(예컨대 5'-캐핑, 스플라이싱 및/또는 3' 폴리아데닐화에 관련된 영역)일 수 있다.

[0152] [0179] 일부 측면에서, 본 발명은 하나 이상의 B형 간염 바이러스(HBV) 유전자의 전사를 억제하기 위한 적어도 하나의 DNA-표적화 모듈을 포함하는 후성적 변형 DNA-표적화 시스템을 제공한다. 여기서 적어도 하나의 DNA-표적화 모듈 각각은: (a) HBV 유전자 또는 이의 조절 요소와 같은 B형 간염 바이러스 DNA 서열의 표적 부위를 표적으로 하기 위한 DNA-결합 도메인; 및 (b) 적어도 하나의 전사 억제자 효과기 도메인을 포함하는 융합 단백질을 포함한다.

[0153] [0180] 일부 측면에서, 본 발명에서는 하나 이상의 B형 간염 바이러스(HBV) 유전자의 전사를 억제하기 위한 적어도 하나의 DNA-표적화 모듈을 포함하는 후성적 변형 DNA-표적화 시스템이 제공되며, 여기서 적어도 하나의 DNA-표적화 모듈 각각은 다음을 포함한다: (a) CRISPR 관련(Cas) 단백질 또는 이의 변이체와 적어도 하나의 전사 억제자 효과기 도메인을 포함하는 융합 단백질; 및 (b) 복수의 유전자 또는 이의 조절 요소의 복수의 표적 부위를 표적으로 하는 복수의 가이드 RNA(gRNA), 여기서 복수의 유전자 또는 이의 조절 요소는 B형 간염 바이러스 복제 및/또는 HBV 전사를 조절한다. 제공된 구현예의 측면에서, 복수의 표적 부위는 2, 3, 4, 5 또는 6개의 다른 표적 부위이다. 제공된 구현예의 측면에서, 복수의 표적 부위는 각각 다른 HBV 유전자 또는 이의 조절 요소에 있다.

[0154] [0181] 제공된 구현예의 측면에서, 본원에서 제공된 DNA-표적화 시스템은 B형 간염 바이러스(HBV)에 감염된 세포에서 하나 이상의 HBV 유전자의 전사를 감소시키기 위해 유전자 또는 이의 조절 요소를 표적화하는데, 여기서 감소된 전사는 HBV에 감염된 세포의 하나 이상의 활동 또는 기능, 예를 들어 HBV RNA 및/또는 HBV 단백질의 발현을 조절한다. 일부 구현예에서, 유전자의 감소된 전사는 감염된 세포에서 유전자의 발현, 즉 감소된 유전자 발현의 감소를 초래한다. 일부 구현예에서, 유전자의 감소된 전사, 예를 들어 감소된 유전자 발현은 감염된 세포에서 단백질의 발현, 즉 감소된 단백질 발현의 감소를 초래한다.

[0155] [0182] 일부 측면에서, 세포는 간세포, 예를 들어 간세포, 간 성상세포(HSC), 쿠퍼세포 및 간 동모세혈관 내피세포와 같은 간세포이다. 예를 들어, 본원에서 제공되는 DNA-표적화 시스템은 표적 세포에서 HBV 유전자의 전사를 감소시키기 위해 유전자 또는 이의 조절 요소를 표적으로 하며, 여기서 감소된 전사는 HBV의 하나 이상의 활동 또는 기능, 예를 들어 HBV의 전사 및 단백질 발현을 조절한다. 일부 구현예에서, 유전자의 감소된 전사는 표적 세포에서 유전자 발현, 즉 감소된 유전자 발현의 감소를 초래한다. 일부 측면에서 세포는 간세포이다.

[0156] [0183] 일부 측면에서, 세포는 인간 대상으로부터 유래한다. 일부 측면에서, 세포는 대상 내의 세포(즉, 생체 내 세포)이다.

[0157] [0184] 일부 구현예에서, DNA-결합 도메인은 CRISPR 관련(Cas) 단백질, 아연 집게 단백질[ZFP], 전사 활성인자 유사 효과기[TALE], 메가뉴클레아제, 호밍 핵산내부분해효소, I-SceI 효소, 또는 이의 변이체를 포함하거나 이들에서의 것이다. 일부 구현예에서, DNA-결합 도메인은 전술한 것 중 임의의 촉매적 비활성(예를 들어, 뉴클레아제-비활성 또는 뉴클레아제-비활성화된) 변이체를 포함한다. 일부 구현예에서, DNA-결합 도메인은 촉매적으로 비활성화되어 뉴클레아제 활성에 대해 비활성이고 DNA를 절단할 수 없는 비활성화된 Cas9(dCas9) 단백질 또는 이의 변이체를 포함한다.

[0158] [0185] 일부 구현예에서, DNA-결합 도메인은 Cas 단백질 또는 그의 변이체, 예컨대 뉴클레아제-비활성 Cas 또는 dCas(예를 들어, dCas9)를 포함하거나 이것에서 유래하고, DNA-표적화 시스템은 하나 이상의 가이드 RNA[gRNA], 예컨대 gRNA의 조합물(예를 들어, 두 개의 gRNA 또는 세 개의 gRNA)을 포함한다. 일부 구현예에서, gRNA는 표적 부위를 표적화하고/하거나 표적 부위에 혼성화할 수 있는 스페이서 서열을 포함한다. 일부 구현예에서, gRNA는 Cas 단백질 또는 이의 변이체와 복합체를 형성할 수 있다. 일부 측면에서, gRNA는 Cas 단백질 또는 이의 변이체를 표적 부위로 가이드하거나 동원한다. 일부 구현예에서, 효과기 도메인은 전사 억제자 도메인을 포함하고/하거나 유전자의 전사를 감소시킬 수 있다. 일부 구현예에서, 효과기 도메인은 직접적으로 또는 간접적으로 유전자의 전사를 감소시킨다. 일부 구현예에서, 효과기 도메인은 전사 억제를 유도하거나 촉매화하거나 이로 이어진다. 일부 구현예에서, 효과기 도메인은 전사 억제를 유도한다. 일부 측면에서, 효과기 도메인은 KRAB 도메인, ERF 억제 인자 도메인, MXI1 도메인, SID4X 도메인, MAD-SID 도메인, DNMT 패밀리 단백질 도메인(예를 들어, DNMT3A 또는 DNMT3B), 하나 이상의 DNMT 패밀리 단백질 또는 이의 도메인의 융합체(예를 들어, DNMT3A/L, DNMT3A 및 DNMT3L 도메인의 융합체를 포함함), LSD1, EZH2, SunTag 도메인, EZH2 도메인, 전술한 것들 중 임의의 것의 부분 또는 완전 기능성 단편 또는 도메인, 또는 전술한 것들 중 임의의 것의 조합물에서 선택된다. 일

부 구현예에서, 효과기 도메인은 KRAB이다. 일부 구현예에서, 효과기 도메인은 DNMT3A/L이다.

[0159] [0186] 일부 구현예에서, DNA-표적화 시스템의 융합 단백질은 dCas9-KRAB 융합 단백질을 포함한다. 일부 구현예에서, DNA-표적화 시스템의 융합 단백질은 DNMT3A/L-dCas9-KRAB-융합 단백질을 포함한다. 일부 구현예에서, DNA-표적화 시스템의 융합 단백질은 KRAB-dCas9-DNMT3A/L-융합 단백질을 포함한다.

[0160] [0187] DNA 표적 지정 시스템의 대표적인 구성 요소와 기능은 아래 하위 섹션에서 제공된다.

[0161] **A. 표적 부위 및 표적 위치**

[0162] [0188] 본원의 임의의 구현예에서, 표적 부위는 B형 간염 바이러스(HBV) 게놈 내의 유전자 및/또는 그 조절 요소이다. 일부 구현예에서, 표적 부위는 공유적으로 폐쇄된 원형 DNA(cccDNA) 형태, 이완된 원형 DNA(rcDNA)에 존재하고/하거나 인간 게놈 DNA에 통합된다. 일부 구현예에서, 표적 부위는 B형 간염 바이러스 DNA 서열에 있다. 일부 구현예에서, B형 간염 바이러스 DNA 서열은 HBV 유전자 또는 이의 조절 요소이다. 일부 구현예에서, 표적 부위는 HBV 복제 및/또는 HBV 전사에 관여하는 유전자에 있거나 그 근처에 있다. 일부 구현예에서, 후성적 변형 DNA-표적화 시스템은 표적 부위를 표적으로 삼아 하나 이상의 HBV의 전사를 억제하기 위한 적어도 하나의 DNA-표적화 모듈을 포함한다. 일부 측면에서, HBV 유전자의 전사를 억제하는 것, 예를 들어 유전자 발현을 감소시키는 것은 HBV 복제(예컨대 감소된 HBV 복제) 및/또는 HBV 전사의 침묵을 초래한다.

[0163] [0189] 제공된 발명과 관련하여, HBV에 대해 양성(+)인 세포(예컨대 HBV에 감염된 세포)는 세포가 여기에 설명된 HBV 마커(예컨대 HBV RNA 전사체 및/또는 단백질) 중 하나를 발현한다는 것을 의미한다. 마찬가지로, 특정 마커에 대해 음성(-)인 세포는 검출할 수 없는 수준으로 마커를 발현하지 않는 세포라는 것을 의미한다. 항체 및 기타 결합 독립체는 주어진 세포 표면 마커를 식별하거나 검출하기 위한 마커 단백질의 발현 수준을 검출하는 데 사용될 수 있다. 적합한 항체는 다클론, 단클론, 단편(예컨대 Fab 단편), 단쇄 항체 및 기타 형태의 특이적 결합 분자를 포함할 수 있다. 위의 세포 표면 마커에 대한 항체 시약은 당업자에게 용이하게 알려져 있다. 표면 마커 또는 단백질의 발현 수준을 평가하기 위한 다수의 주지된 방법, 예컨대 친화성 기반 방법, 예를 들어 면역친화성 기반 방법에 의한 검출, 예를 들어 표면 마커의 맥락에서, 예컨대 유세포 분석에 의한 검출이 사용될 수 있다. 일부 구현예에서, 표지는 형광단이고, 세포(예를 들어, 간세포)의 세포 표면 마커의 검출 또는 확인을 위한 방법은 유세포 분석에 의한 것이다. 일부 구현예에서, 상이한 표지가 상이한 마커 각각에 대해 다색 유세포 분석에 의해 사용된다. 일부 구현예에서, 표면 발현은 유세포 분석에 의해, 예를 들어 마커에 특이적으로 결합하는 항체로 염색하고 마커에 대한 항체의 결합을 검출함으로써 결정될 수 있다.

[0164] [0190] 일부 구현예에서, 세포(예컨대 간세포)는 특정 마커가 세포 내 또는 세포 내에 검출 가능한 존재인 경우 특정 마커에 대해 양성(pos 또는 +)이며, 이는 세포 내 마커 또는 표면 마커(예컨대 HBeAg, HBsAg)일 수 있다. 일부 구현예에서, 흐름 세포 분석법에 의한 염색이 동일 조건에서 동형 매칭 대조군으로 동일한 절차를 수행하여 검출된 염색보다 상당히 높은 수준으로 검출되거나, 마커에 대해 양성인 것으로 알려진 세포와 상당히 유사하거나, 어떤 경우에는 그보다 더 높은 수준으로 검출되거나, 마커에 대해 음성인 것으로 알려진 세포보다 더 높은 수준으로 검출되는 경우 표면 발현이 양성이다. 일부 구현예에서, 본원에서 설명하는 DNA-표적화 시스템에 의해 접촉된 세포(예컨대 간세포)는 염색이 DNA-표적화 시스템에 의해 접촉되지 않은 유사한 세포보다 상당히 낮으면 특정 마커(예컨대 HBeAg)에 대한 발현이 감소한다.

[0165] [0191] 일부 구현예에서, 세포(예를 들어 간세포)는 세포 내 마커 또는 표면 마커일 수 있는 특정 마커의 세포상에 또는 세포 내에 검출가능한 존재가 없는 경우 특정 마커에 대해 음성(neg 또는 -)이다. 일부 구현예에서, 유세포 분석에 의한 염색이 달리 동일한 조건하에서 동종형-매칭된 대조군으로 동일한 절차를 수행하는 것으로 검출된 염색보다 실질적으로 높은 수준에서 및/또는 마커에 대해 양성인 것으로 공지된 세포와 실질적으로 낮은 수준에서 또는 마커에 대해 음성인 것으로 공지된 세포와 유사한 수준에서 검출 불가능한 경우 표면 발현은 음성이다.

[0166] [0192] 일부 구현예에서, 감염된 세포 및/또는 개인의 표현형이 기능적으로 특성화된다. 일부 측면에서, 표현형은 감염된 세포에서 HBV RNA 전사체의 존재로 특성화될 수 있다. 일부 측면에서, 표현형은 감염된 세포에서 HBV 단백질 중 하나 또는 조합의 존재로 특성화될 수 있다. 일부 측면에서, 표현형은 본원에서 기술된 마커 중 하나에 대한 항체의 존재로 특성화될 수 있다. 일부 측면에서, 항체는 항-HBc-IgM, 항-HBc 총 및 HBeAg에 대한 항체를 포함하지만 이에 국한되지 않는다. 일부 구현예에서, RNA 전사체, 단백질 및/또는 항체는 당해 기술 분야에 알려진 모든 적합한 기술에 의해 측정, 검출 및/또는 정량화된다. 예를 들어, RNA 전사체는 실시간 PCR 기술을 사용하여 측정, 검출 및/또는 정량화될 수 있다. HBV 단백질(예컨대 HBsAg, HBeAg 및/또는 HBcrAg)은 효소 결합

면역 흡착 검사(ELISA)를 사용하여 측정, 검출 및/또는 정량화할 수 있다.

[0167] [0193] 제공된 DNA-표적화 시스템, 여기에 포함된 다중화된 후성적 변형 DNA-표적화 시스템에 의한 조절을 위한 표적 유전자 및/또는 그 조절 요소에는 전사 및 발현이 세포(예컨대 HBV에 감염된 세포)에서 감소된 모든 것이 포함된다. 세포(예컨대 간세포)에서 유전자의 전사 또는 발현 수준을 특성화하기 위해 다양한 방법을 활용할 수 있다(예컨대 세포가 제공된 DNA-표적화 시스템과 접촉하거나 도입된 후). 일부 구현예에서, 유전자의 전사 활성 또는 발현을 분석하는 것은 RNA 분석에 의한 것일 수 있다. 일부 구현예에서, RNA 분석에는 RNA 정량화가 포함된다. 일부 구현예에서, RNA 정량화는 역전사 정량 PCR(RT-qPCR), 다중화된 qRT-PCR, 형광 현장 하이브리디제이션(FISH), RNA-시퀀싱(RNA-seq) 또는 이들의 조합에 의해 발생한다.

[0168] [0194] 일부 구현예에서, 유전자 또는 전사체는 세포(예컨대 HBV에 감염된 세포, 예를 들어 간세포)에서 유전자의 발현 또는 전사체의 존재가 제공된 DNA-표적화 시스템, 예를 들어 다중화된 후성적 DNA-표적화 시스템과 접촉 또는 도입된 후 감소된 것이다. 일부 측면에서, 복수의 유전자 또는 전사체는 다중화된 후성적 DNA-표적화 시스템, 예를 들어 하나 이상의 DNA-표적화 모듈에 의해 표적화된다. 이러한 시스템에서, 다중화된 DNA-표적화 시스템의 각 유전자 또는 전사체는 제공된 다중화된 후성적 DNA-표적화 시스템과 접촉 또는 도입된 후 세포에서 유전자의 발현 또는 전사체의 존재가 감소된 것이다. 일부 구현예에서, 세포(예컨대 HBV에 감염된 세포, 예를 들어 간세포)에서의 유전자 발현의 감소 또는 전사체 수준의 변화는 적어도 1.25배, 1.5배, 1.75배, 2.0배, 2.5배, 2.75배, 3.0배, 3.25배, 3.5배, 3.75배, 4.0배, 4.25배, 4.5배, 4.75배, 5.0배, 5.25배, 5.5배, 5.75배, 6.25배, 6.50배, 6.75배, 7.0배의 log2 배 변화이다. 7.25배, 7.50배, 7.75배, 8.0배, 8.25배, 8.5배, 8.75배, 9.0배 또는 앞서 언급한 것 중 어느 하나 사이의 임의의 값이며, 대조군 세포의 유전자 수준과 비교된다. 일부 구현예에서, 세포(예컨대 간세포와 같은 HBV 감염 세포)에서 유전자 발현의 감소 또는 전사체 수준의 변화는 약 10%, 약 20%, 약 30%, 약 40%, 약 50%, 약 60%, 약 70%, 약 80%, 약 90%, 약 100% 또는 앞서 언급한 것 중 어느 하나 사이의 임의의 값이다. 일부 구현예에서, 세포에서 유전자 발현의 감소 또는 전사체 수준의 변화는 대조군 세포의 유전자 수준과 비교되어 90%보다 크다. 일부 구현예에서, 가이드 RNA는 서열번호 565, 528, 542, 508, 515, 575, 515, 453, 506, 514, 425 또는 472에 제시된다. 일부 구현예에서, 세포에서 유전자 발현의 감소 또는 전사체 수준의 변화는 대조군 세포에서 유전자 수준에 비해 75% 이상이다. 일부 구현예에서, 가이드 RNA는 서열번호 565 (HBVg_175), 528 (HBVg_138), 582 (HBVg_192), 542 (HBVg_152), 508 (HBVg_118), 515 (HBVg_125), 575 (HBVg_185), 453 (HBVg_63), 506 (HBVg_116), 514 (HBVg_124), 395 (HBVg_5), 472 (HBVg_82), 451 (HBVg_61), 488 (HBVg_98), 540 (HBVg_150), 533 (HBVg_143), 572 (HBVg_182), 566 (HBVg_176), 489 (HBVg_99), 469 (HBVg_79), 408 (HBVg_18), 465 (HBVg_75), 402 (HBVg_12), 474 (HBVg_84), 525 (HBVg_135), 416 (HBVg_26), 396 (HBVg_6), 554 (HBVg_164), 419 (HBVg_29), 545 (HBVg_155), 446 (HBVg_56), 580 (HBVg_190), 555 (HBVg_165), 412 (HBVg_22), 428 (HBVg_38), 458 (HBVg_68), 548 (HBVg_158), 511 (HBVg_121), 432 (HBVg_42), 441 (HBVg_51), 433 (HBVg_43), 579 (HBVg_189), 479 (HBVg_89), 478 (HBVg_88), 520 (HBVg_130), 462 (HBVg_72), 523 (HBVg_133), 및 503 (HBVg_113)에 제시된다.

[0169] [0195] 제공된 구현예에서, cccDNA는 숙주 RNA 폴리머라제 하에서 5개의 HBV RNA(0.7kb, 2.1kb, 2.4kb, 더 길고 더 짧은 3.5kb RNA)를 전사한다. cccDNA의 전사는 4개의 프로모터 - 기저 코어, preS1, preS2 및 X 프로모터-와 2개의 인핸서(인핸서 I 및 II)에 의해 제어된다(도 1). 0.7kb RNA는 전사 조절자 역할을 하는 HBV X 단백질(HBx)로 번역될 수 있다. 2.1kb RNA는 HBV 소형 표면 단백질(S) 및 중간 표면 단백질(M)로 번역될 수 있다. 2.4kb RNA는 HBV 대형 표면 단백질(L)로 번역될 수 있다. L, M, S는 자체 조립되어 빈 아바이러스 입자(SVP)(구형 SVP 및 필라멘트형 SVP 포함)를 형성할 수 있으며, 이는 필라멘트형 SVP와 상당량의 L 단백질을 함유하는 바이러스 입자와 함께 분비된다. 구형 SVP는 구성적 분비 경로를 통해 분비된다. 필라멘트형 SVP는 다중포체(MVB: multivesicular bodies)를 통한 수송에 필요한 엔도솜 분류 복합체(ESCRT) 기구에 의해 분비된다. 더 긴 3.5kb RNA는 pre-코어 RNA(preC RNA)라고 하며, HBV e 항원(HBeAg)으로 더 잘 알려진 pre-코어 단백질로 번역될 수 있다. 더 짧은 3.5kb RNA는 pre-게놈 RNA(pgRNA)로, HBV 폴리머라제(Pol) 및 코어 단백질의 번역 템플릿과 Pol에 의한 캡시드 내(코어 단백질 중합에 의해 형성됨) 역전사를 위한 복제 템플릿의 두 가지 역할을 한다. 이러한 뉴클레오캡시드는 HBV 표면 단백질(L, M 및 S)에 의해 둘러싸여 성숙한 바이러스를 형성하고 ESCRT/MVB 경로를 통해 분비될 수 있다. 또는 이러한 뉴클레오캡시드는 핵으로 운반되어 cccDNA를 형성할 수도 있다. 일부 구현예에서, HBV 유전자의 전사 및/또는 번역을 억제하는 것, 예컨대 유전자 발현 감소는 다음 HBV 마커 중 어느 하나를 침묵시키는 결과를 가져온다: HBV HBV X 단백질(HBx), B형 간염 표면 항원(HBsAg), 예를 들어 작은 표면 단백질(S), 중간 표면 단백질(M) 또는 HBV 큰 표면 단백질(L), HBV e 항원(HBeAg). 일부 구현예에서, HBV 유전자의 전사 및/또는 번역을 억제하는 것(예컨대 유전자 발현 감소)은 HB 코어 관련 항원(HBcrAg)을 침묵시키는 결과를 가져온다. HBcrAg에는 B형 간염 코어 항원(HBcAg), HBeAg, 및 22kDA pre-코어 단백질(p22cr)을 포함한

3개의 pre-코어/코어 단백질 생성물이 포함된다. 일부 측면에서, cccDNA, HBV 총 DNA, 혈청 HBcrAg, HBsAg, HBeAg, B형 간염 코어 항체(항-HBc), HBV DNA, HBV RNA는 감소된 HBV 전사 및/또는 번역을 측정하기 위한 판독 값으로 정량화된다. 일부 구현예에서, 표적 부위는 HBV 단백질 중 하나를 인코딩하는 유전자에 있다. 일부 구현예에서, 표적 부위는 HBV 단백질 중 하나를 인코딩하는 유전자의 조절 요소(예컨대 프로모터 또는 인핸서)에 있다.

[0170] [0196] 일부 구현예에서, 후성적 변형 DNA-표적화 시스템의 표적 부위는 HBV 복제 및/또는 HBV 전사에 관여하는 유전자에 있다. 일부 측면에서, 후성적 변형 DNA-표적화 시스템의 표적 부위는 HBV 복제 및/또는 HBV 전사를 제어하는 유전자 또는 이의 조절 요소에 있거나 그 근처에 있다. 일부 구현예에서, HBV 복제 및/또는 HBV 전사에 관여하는 유전자는 폴리머라제 유전자, S-패밀리 유전자, X-유전자 및/또는 코어-패밀리 유전자이다. 일부 구현예에서, HBV 복제 및/또는 전사에 관여하는 유전자는 폴리머라제, 외피 단백질, 캡시드 단백질, 전사 인자 또는 전사 전사 활성화제를 인코딩한다. 일부 구현예에서, HBV 복제 및/또는 HBV 전사에 관여하는 그 조절 요소는 프로모터 영역, 인핸서 영역 및/또는 임의의 전사 처리 제어 영역이다. 일부 구현예에서, 프로모터 영역은 pre-S1 프로모터, pre-a S2 프로모터, X 프로모터 또는 기저 코어 프로모터이다. 일부 구현예에서, 인핸서 영역은 Enh1 인핸서 및/또는 Enh2 인핸서 영역이다. 일부 구현예에서, 전사 처리 제어 영역은 5'-말단 캡핑, 스플라이싱 및/또는 3'-말단 폴리아데닐화에 대한 신호를 인코딩하는 영역이다.

[0171] [0197] 일부 구현예에서, 표적 부위는 B형 간염 바이러스 게놈(B형 간염 바이러스 아형 ayw, 완전 게놈, GenBank: U95551.1), 서열번호 650을 참조로 HBV 게놈의 염기쌍(bp) 위치: 1 bp - 42 bp, 491 bp - 1032 bp, 1750 bp - 1799 bp, 또는 1951 bp -2952 bp 사이에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내의 서열이다. 일부 구현예에서, 표적 부위는 HBV 게놈의 표적 영역 내에 있으며, 여기서 표적 영역은 서열번호 1056에 제시된 서열을 갖거나 이의 상보적 서열이다. 일부 구현예에서, 표적 부위는 HBV 게놈의 표적 영역 내에 있으며, 여기서 표적 영역은 서열번호 1058에 제시된 서열을 갖거나 이의 상보적 서열이다. 일부 구현예에서, 표적 부위는 HBV 게놈의 표적 영역 내에 있으며, 여기서 표적 영역은 서열번호 1060에 제시된 서열을 갖거나 이의 상보적 서열이다. 일부 구현예에서, 표적 부위는 HBV 게놈의 표적 영역 내에 있으며, 여기서 표적 영역은 서열번호 1062에 제시된 서열을 갖거나 이의 상보적 서열이다.

[0172] [0198] 일부 구현예에서, 표적 부위는 B형 간염 바이러스 게놈(B형 간염 바이러스 아형 ayw, 완전 게놈, GenBank: U95551.1), 서열번호 650을 참조로 HBV 게놈의 염기쌍(bp) 위치: 43 bp - 490 bp, 1033 bp - 1749 bp, 1800 bp - 1950 bp, or 2953 bp - 3182bp 사이에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내의 서열이다. 일부 구현예에서, 표적 부위는 HBV 게놈의 표적 영역 내에 있으며, 여기서 표적 영역은 서열번호 1057에 제시된 서열을 갖거나 이의 상보적 서열이다. 일부 구현예에서, 표적 부위는 HBV 게놈의 표적 영역 내에 있으며, 여기서 표적 영역은 서열번호 1059에 제시된 서열을 갖거나 이의 상보적 서열이다. 일부 구현예에서, 표적 부위는 HBV 게놈의 표적 영역 내에 있으며, 여기서 표적 영역은 서열번호 1061에 제시된 서열을 갖거나 이의 상보적 서열이다. 일부 구현예에서, 표적 부위는 HBV 게놈의 표적 영역 내에 있으며, 여기서 표적 영역은 서열번호 1063에 제시된 서열을 갖거나 이의 상보적 서열이다.

[0173] [0199] 일부 구현예에서, 표적 부위는 B형 간염 바이러스 게놈(B형 간염 바이러스 아형 ayw, 완전 게놈, GenBank: U95551.1), 서열번호 650을 참조로 HBV 게놈의 염기쌍(bp) 위치: 67 bp - 392 bp (CpG 섬 1), 1033 bp -1749 bp (CpG Island 2), 또는 2215 bp - 2490 bp (CpG 섬 3) 사이에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내의 서열이다. 일부 구현예에서, 표적 부위는 HBV 게놈의 표적 영역 내에 있으며, 여기서 표적 영역은 서열번호 1064에 제시된 서열을 갖거나 이의 상보적 서열이다. 일부 구현예에서, 표적 부위는 HBV 게놈의 표적 영역 내에 있으며, 여기서 표적 영역은 서열번호 1059에 제시된 서열을 갖거나 이의 상보적 서열이다. 일부 구현예에서, 표적 부위는 HBV 게놈의 표적 영역 내에 있으며, 여기서 표적 영역은 서열번호 1066에 제시된 서열을 갖거나 이의 상보적 서열이다.

[0174] [0200] 일부 구현예에서, 표적 부위는 폴리머라제 유전자 또는 이의 조절 요소에 있다. 폴리머라제 유전자(P 유전자라고도 함)는 바이러스 RNA 게놈을 바이러스 세포질 캡시드의 dsDNA로 전환하는 다기능 효소(P 폴리머라제, HBVgp1, DNA 지향 DNA 폴리머라제라고도 함)를 인코딩한다. 폴리머라제는 DNA 또는 RNA 템플릿을 복사할 수 있는 DNA 폴리머라제 활동과 부분적으로 진행성인 3'-5'-엔도뉴클레아제 모드에서 RNA-DNA 헤테로듀플렉스의 RNA 가닥을 절단하는 리보뉴클레아제 H(RNase H) 활동을 나타낸다. 폴리머라제 유전자 ORF는 preS/S ORF와 완전히 중첩되고 코어 패밀리 및 X 유전자 ORF와 부분적으로 중첩친다. 일부 구현예에서, 표적 부위는 B형 간염 바이러스 게놈(B형 간염 바이러스 아형 ayw, 완전 게놈, GenBank: U95551.1), 서열번호 650을 참조로 HBV 게놈의 1 bp - 1621 bp, 1374 bp - 1838 bp 또는 2307 bp - 3182 bp 사이에 위치한 서열에 상응하는 서열을 갖는 표적

영역 내의 서열이다.

[0175] [0201] 일부 구현예에서, 표적 부위는 S-패밀리 유전자 또는 이의 조절 요소에 있다. 일부 구현예에서, 표적 부위는 S 유전자, pre-S1 프로모터 및/또는 pre-S2 프로모터 영역에 있다. S-패밀리 유전자는 세 가지 다른 구조적으로 관련된 엔벨로프 단백질을 인코딩하는데, 이는 대체 개시 코돈에서 합성되며 대형(L), 중형(M), 소형(S) B형 간염(HB) 단백질(각각 L-HB, M-HB 및 s-HB라고도 함)로 불린다. 이들 세 가지 단백질은 동일한 카르복시 말단을 공유하지만 아미노 말단 확장부는 다르다. 일부 구현예에서, 표적 부위는 B형 간염 바이러스 게놈(B형 간염 바이러스 아형 ayw, 완전 게놈, GenBank: U95551.1), 서열번호 650을 참조로 HBV 게놈의 1 bp - 837 bp, 1 bp - 155 bp 또는 2854 bp - 3182 bp 사이에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내의 서열이다.

[0176] [0202] 일부 구현예에서, 표적 부위는 X-유전자 또는 이의 조절 요소에 있다. X-유전자(HBx, HBVgp3, 펩타이드 X, pX로도 알려짐)는 숙주 인자와 직접 또는 간접적으로 상호 작용하여 전사 조절, 단백질 분해 경로, 세포 사멸, 신호 형질도입, 세포 주기 진행 및 유전적 안정성을 조절하는 다기능 단백질을 인코딩하는 유전자이다. X-유전자 단백질은 숙주 인자와 직접 또는 간접적으로 상호 작용하여 단백질 분해 경로, 세포 사멸, 전사, 신호 형질도입, 세포 주기 진행 및 유전적 안정성을 조절한다. 일부 구현예에서, 표적 부위는 B형 간염 바이러스 게놈(B형 간염 바이러스 아형 ayw, 완전 게놈, GenBank: U95551.1), 서열번호 650을 참조로 HBV 게놈의 1374bp-1838bp 사이에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내의 서열이다.

[0177] 일부 구현예에서, HBx 단백질을 인코딩하기 위한 시작 코돈(HBx 시작 코돈)은 서열번호 650에 제시된 HBV 게놈을 참조로 하는 위치에 상응하는 HBV 게놈의 잔여 염기 쌍 1376에 있다. 일부 구현예에서, HBx 단백질을 인코딩하기 위한 시작 코돈(HBx 시작 코돈)은 서열번호 1071에 제시된 HBV 게놈을 참조하여 위치에 상응하는 HBV 게놈의 잔여 염기 쌍 1374에 있다. 여기에서 제공된 후성적-변형 DNA 표적화 시스템을 이용하여 시작 코돈 상류의 이 영역의 X- 유전자의 표적 부위를 표적화하면 HBV 감염된 세포의 바이러스 복제 및 전사를 억압하는데 높은 활성을 나타내는 것으로 밝혀졌다. 일부 구현예에서, 표적 영역은 HBV 게놈의 CpG 섬에 있다. 일부 구현예에서, 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 1033-1749 bp 사이에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내의 서열이다. 일부 구현예에서, 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 1100-1350bp 사이에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내에 존재하는 것과 같이, HBx 프로모터/인핸서 #1 영역에 존재한다. 일부 구현예에서, 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 1600-1750bp 사이에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내에 존재하는 것과 같이, 기저 코어 프로모터 영역에 존재한다.

[0178] [0203] 일부 구현예에서, 표적 부위는 HBx 시작 코돈 상류에서 300개 염기쌍(bp) 이내, 250 bp 이내, 200 bp 이내, 150 bp 이내, 140 bp 이내, 130 bp 이내, 120 bp 이내, 110 bp 이내 또는 100 bp 이내에 걸쳐 있는 표적 영역 내에 있다. 일부 구현예에서, 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 1250-1374개 bp 사이에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내에 있다. 일부 구현예에서, 표적 부위는 표적 영역이 서열번호 1068에 제시된 서열을 갖거나 그 상보적 서열인 HBV 게놈의 표적 영역 내에 있다. 일부 구현예에서, 표적 부위는 서열번호 650에 제시된 HBV 게놈을 기준으로 1255-1302 bp 사이에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내의 서열이다. 일부 구현예에서, 표적 부위는 표적 영역이 서열번호 1069에 제시된 서열을 갖는 HBV 게놈의 표적 영역 내에 있거나, 이의 상보적 서열이다. 일부 구현예에서, 표적 부위는 서열번호 650에 제시된 HBV 게놈을 기준으로 1260-1300 bp 사이에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내의 서열이다. 일부 구현예에서, 표적 부위는 표적 영역이 서열번호 1070에 제시된 서열을 갖는 HBV 게놈의 표적 영역 내에 있거나, 이의 상보적 서열이다. 일부 구현예에서, 표적 부위 또는 각각의 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조하는 위치에 상응하는 HBV 게놈의 1060-1480 bp 사이의 염기쌍에 위치한 표적 영역 내에 있다. 일부 구현예에서, 표적 부위는 표적 영역이 서열번호 1067에 제시된 서열을 갖는 HBV 게놈의 표적 영역 내에 있다. 이러한 영역에서 표적 부위를 표적으로 삼기 위한 예시적인 DNA 결합 시스템이 본원에서 제공되며, 여기에는 CRISPR/Cas 시스템 및 ZFP를 포함한 다양한 DNA-결합 도메인을 갖는 시스템이 포함된다.

[0179] [0204] 일부 구현예에서, 표적 부위는 코어 패밀리 유전자 또는 이의 조절 요소에 있다. 일부 구현예에서, 그 조절 요소는 Enh2 프로모터이다. 일부 구현예에서, 그 조절 요소는 기저 코어 프로모터(BCP)이다. 바이러스 게놈의 코어 프로모터(CP) 영역은 바이러스의 복제 및 형태발생에서 중심 역할을 한다(Quarleri J, World Journal of Gastroenterology 20(2): 425-435(2014)). 코어 프로모터 영역은 pre-코어 mRNA와 pre-게놈 RNA(pgRNA)의 합성을 위한 전사 개시를 지시한다. CP 영역은 pre-코어 mRNA(preC, C 유전자, HBVgp4라고도 함) 및 pgRNA 전사를 개시하는 기저 코어 프로모터(BCP)로 구성되고, 프로모터 활성을 조절하는 양성 및 음성 조절 요소를 포함하는 상위 조절 영역(URR)으로 구성된다. 여러 전사 인자는 C/EBP, HNF1, HNF3/4, COUP-TF1과 같은 CP의 조절

서열 요소에 결합하여 pre-C mRNA와 pgRNA의 합성을 다르게 조절한다. CP 내의 AT가 풍부한 영역 또는 TATA-유사 박스의 존재도 pgRNA의 전사에 기인한다. pre-코어 mRNA는 외부 코어 항원(캡시드 단백질, pre-캡시드 단백질, HBeAg, precore 단백질, p25라고도 함)을 인코딩하며, 이는 자가 조립되어 바이러스 게놈을 포장하는 이십면체 캡시드를 형성한다. pgRNA는 폴리머라제, 뉴클레오캡시드 단백질 HBcAg 및 가용성 분비 HBeAg 단백질을 형성하도록 번역된다. pgRNA는 또한 자손 뉴클레오캡시드에 통합되고 공동 조립된 바이러스 폴리머라제에 의해 DNA로 역전사되어 새로운 HBV 비리온으로 변환된다. 이러한 성숙한 이완된 원형 DNA(rcDNA)를 포함하는 뉴클레오캡시드는 동일한 세포의 핵으로 게놈을 재전달하여 10-100개의 cccDNA 분자를 복사한 풀을 형성하거나 ER/골지체의 외피 단백질과 상호 작용하여 새로운 감염성 바이러스로 분비될 수 있다(Pollicino T., et al., Journal of hepatology, 61(2):P408-417 (2014)). 일부 구현예에서, 표적 부위는 코어 패밀리 유전자 또는 이의 조절 요소에 있다. 일부 측면에서, 코어 패밀리 유전자 또는 이의 조절 요소 내의 하나 이상의 부위를 표적화하는 것은 pgRNA 전사체의 억제를 포함한다. 일부 측면에서, pgRNA 전사체의 억제는 HBV 복제의 침묵을 포함한다. 일부 구현예에서, 표적 부위는 B형 간염 바이러스 게놈(B형 간염 바이러스 아형 ayw, 완전 게놈, GenBank: U95551.1), 서열번호 650을 참조로 HBV 게놈의 1590bp 내지 1815bp, 1636bp 내지 1744, 1751bp 내지 1769, 1814bp 내지 1900bp, 1816bp 내지 2455, 1800bp 내지 1950bp 사이에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내의 서열이다. 일부 구현예에서, 전사 감소는 B형 간염 바이러스 RNA 전사체 총량의 감소를 포함한다. 일부 구현예에서, 전사 감소는 B형 간염 전코어("preC") 및/또는 pre-게놈("pgRNA") RNA 수준의 감소를 포함한다.

[0180] [0205] 일부 측면에서, 표적 부위는 코딩 영역이다. 일부 구현예에서, HBV 복제 및/또는 HBV 전사에 관여하는 유전자는 HBV X 단백질(HBx), S 패밀리 단백질(HBsAg), 예를 들어 작은 표면 단백질(S-HBs), 중간 표면 단백질(M-HBs), 또는 HBV 큰 표면 단백질(L-HBs), pre-코어 단백질(HBeAg), HBV 코어 관련 항원(HBcrAg), 폴리머라제, 코어 및 pre-코어 단백질을 인코딩한다. 일부 측면에서, 표적 부위는 B형 간염 바이러스 게놈(B형 간염 바이러스 아형 ayw, 완전 게놈, GenBank: U95551.1), 서열번호 650을 참조로 HBV 게놈의 1 bp - 42 bp, 43 bp - 1090 bp, 1091 bp - 1849 bp, 또는 1850 bp - 2455 bp 또는 2455 bp - 3182 bp 사이에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내의 서열이다.

[0181] [0206] 일부 구현예에서, 전사 억제는 B형 간염 표면 항원(HBsAg) 및/또는 B형 간염 바이러스 코어 관련 항원(HBcrAg) 단백질 수치의 감소를 포함한다. 일부 구현예에서, 전사 억제는 HBsAg 전사체 및/또는 단백질 수치의 최소 90% 감소를 포함한다. 일부 구현예에서, 전사 억제 전사는 cccDNA에서 HBcrAg 전사체 및/또는 단백질 수준이 최소 50% 감소하는 것을 포함한다.

[0182] [0207] 일부 구현예에서, 전사 억제는 B형 간염 전핵(preC), pre-게놈(pgRNA), preS1, preS2/S 및 HBx 수준의 감소를 포함한다.

[0183] [0208] 일부 구현예에서, 다중 후성적 변형 DNA-표적화 시스템은 임의의 전술한 바와 같이 유전자 내의 표적 부위를 표적화하거나 이에 결합한다. 일부 구현예에서, 표적 부위는 세포(예를 들어, 간세포) 내 유전자의 조절 DNA 요소에 위치한다. 일부 구현예에서, 조절 DNA 요소는 유전자 조절 단백질이 결합하여 유전자의 전사에 영향을 미칠 수 있는 서열이다. 일부 구현예에서, 조절 DNA 요소는 유전자의 시스, 트랜스, 원위, 근위, 상류 또는 하류 조절 DNA 요소이다. 일부 구현예에서, 조절 DNA 요소는 유전자의 프로모터 또는 인핸서이다. 일부 구현예에서, 표적 부위는 유전자의 프로모터, 인핸서, 엑손, 인트론, 비번역 영역[UTR], 5' UTR 또는 3' UTR 내에 위치한다. 일부 구현예에서, 프로모터는 유전자의 전사를 시작하기 위해 RNA 폴리머라제가 결합하는 뉴클레오타이드 서열이다. 일부 구현예에서, 프로모터는 유전자의 전사 개시 부위의 100bp와 1000bp 사이에 일반적으로 위치한 뉴클레오타이드 서열이며, 예를 들어 유전자의 전사 개시 부위의 약 100bp, 약 500bp, 약 1000bp 내에 위치한다. 일부 구현예에서 표적 부위는 유전자의 발현을 제어할 수 있는 것으로 의심되는 알려지지 않았거나 알려진 기능의 서열 내에 위치한다.

[0184] [0209] 일부 구현예에서, 표적 부위는 전사 개시 지점의 약 50개 염기쌍(bp), 약 100bp, 약 150bp, 약 200bp, 약 250bp, 약 300bp, 약 350bp, 약 400bp, 약 450bp, 약 500bp, 약 600bp, 약 650bp, 약 700bp, 약 750bp, 약 800bp, 약 850bp, 약 900bp, 약 1000bp, 약 1050bp, 약 1100bp, 약 1200bp, 약 1250bp, 약 1300bp, 약 1350bp, 약 1400bp, 약 1450bp, 약 1500bp 내에 위치한다.

[0185] [0210] 일부 구현예에서, 표적 부위는 HBV 게놈의 1bp - 3300bp 사이의 염기쌍에 위치한 표적 영역 내에 위치한다. 일부 구현예에서, 표적 부위는 B형 간염 바이러스 게놈(B형 간염 바이러스 아형 ayw, 완전 게놈, GenBank: U95551.1), SEQ ID NO 650을 참조로 HBV 게놈의 43bp - 490bp, 1033bp - 1749bp, 1800bp - 1950bp 또는 2953bp - 3182bp 사이에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내의 서열이다. 일부 구현예에서, 표적

부위는 B형 간염 바이러스 게놈(B형 간염 바이러스 아형 ayw, 완전 게놈, GenBank: U95551.1), SEQ ID NO 650을 참조로 HBV 게놈의 1bp - 42bp, 491bp - 1032bp, 1750bp - 1799bp 또는 1951bp - 2952bp 또는 3198bp - 3182bp 사이에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내의 서열이다. 계통학적 분석과 서열 발산에 따르면, HBV는 완전한 뉴클레오타이드 서열에서 8% 이상의 군 간 발산을 기준으로 10개의 유전형(A 내지 J)으로 분류될 수 있다(Norder H, et.a., Complete genomes, phylogenetic relatedness, and structural proteins of six strains of the hepatitis B virus, four of which represent two new genotypes. Virology. 1994 Feb;198(2):489-503; Stuyver L, et.al., A new genotype of hepatitis B virus: complete genome and phylogenetic relatedness. J Gen Virol. 2000 Jan;81(Pt 1):67-74, Arauz-Ruiz P, et.al., Genotype H: a new Amerindian genotype of hepatitis B virus revealed in Central America. J Gen Virol. 2002 Aug;83(Pt 8):2059-2073)). HBV 유전자형이 임상 결과, pre-코어 및 코어 프로모터 영역의 돌연변이 패턴, HbeAg 혈청 전환율 및 인터페론 요법에 대한 반응에 영향을 미친다는 증거가 있다. 대부분 유전자형은 특정한 지리적 분포를 가지고 있다. 유전자형 A와 D는 서유럽과 북미에서 유행하고 유전자형 B와 C는 동아시아와 오세아니아에서 유행한다.

[0186] [0211] 일부 구현예에서, 표적 부위는 모든 B형 간염 바이러스 유전자형(예컨대 유전체)과 적어도 70% 상동성이다. 일부 구현예에서, 표적 부위는 적어도 500개, 적어도 1000개, 적어도 1500개, 적어도 2000개, 적어도 2500개, 적어도 3000개, 적어도 3500개, 적어도 4000개, 적어도 4500개, 적어도 5000개, 적어도 5500개, 적어도 6000개, 적어도 6500개, 적어도 7000개의 B형 간염 바이러스 유전체와 적어도 70% 상동성이다.

[0187] [0212] 일부 구현예에서, 표적 부위는 적어도 1000개의 B형 간염 바이러스 게놈과 적어도 70% 상동적이며 최대 2개의 불일치를 포함한다. 일부 구현예에서, 표적 부위는 서열번호 1-195 중 어느 하나에 제시된 서열, 서열번호 1-195 중 하나의 적어도 14개 뉴클레오타이드(nt)의 이의 연속 부분, 또는 상기 중 어느 하나의 상보적 서열을 포함한다. 일부 구현예에서, 표적 부위는 길이가 15, 16, 17, 18 또는 19개 뉴클레오타이드인 서열번호 1-195 중 어느 하나의 연속 부분, 또는 상기 중 어느 하나의 상보적 서열이다. 일부 구현예에서, 표적 부위는 상기 본원에서 설명한 표적 부위 서열의 전체 또는 연속 부분과 적어도 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 서열 동일성을 갖는 서열이다. 일부 구현예에서, 표적 부위는 서열번호 1-195 중 어느 하나에 제시된 서열이다.

[0188] [0213] 임의의 구현예 중 일부에서, 해당 표적 부위는 14 내지 22개의 뉴클레오타이드로 구성된 서열이다. 임의의 구현예 중 일부에서, 해당 표적 부위는 14 내지 19개의 뉴클레오타이드로 구성된 서열이다. 임의의 구현예 중 일부에서, 해당 표적 부위는 14개의 뉴클레오타이드로 구성된 서열이다. 임의의 구현예 중 일부에서, 해당 표적 부위는 15개의 뉴클레오타이드로 구성된 서열이다. 임의의 구현예 중 일부에서, 해당 표적 부위는 16개의 뉴클레오타이드로 구성된 서열이다. 임의의 구현예 중 일부에서, 해당 표적 부위는 17개의 뉴클레오타이드로 구성된 서열이다. 임의의 구현예 중 일부에서, 해당 표적 부위는 18개의 뉴클레오타이드로 구성된 서열이다. 임의의 구현예 중 일부에서, 해당 표적 부위는 19개의 뉴클레오타이드로 구성된 서열이다.

[0189] [0214] 본원에 제공된 임의의 구현예에서, 표적 부위는 참조된 서열(즉, 서열목록을 참조하여 서열번호로 제시된 특정 서열)에 상보적이다. 임의의 구현예 중 일부에서, 상보적 서열은 참조된 서열의 역 상보적이다.

[0190] [0215] 본원에 제공된 임의의 구현예에서, 표적 부위는 참조된 서열(즉, 서열 목록을 참조하여 서열번호에 의해 제시된 특정 서열)을 포함한다. 본원에 제공된 임의의 구현예에서, 표적 부위는 참조된 서열(즉, 서열 목록을 참조하여 서열번호에 의해 제시된 특정 서열)에 의해 제시된 서열이다.

[0191] [0216] 본원에 제공된 구현예 중 어느 하나에서, 표적 부위는 참조된 서열(즉, 서열 목록을 참조하여 서열번호로 제시된 특정 서열)의 적어도 14개 뉴클레오타이드(14nt)의 연속 부분이다. 일부 구현예에서, 연속 부분은 15개 뉴클레오타이드이다. 일부 구현예에서, 연속 부분은 16개 뉴클레오타이드이다. 일부 구현예에서, 연속 부분은 17개 뉴클레오타이드이다. 일부 구현예에서, 연속 부분은 18개 뉴클레오타이드이다. 일부 구현예에서, 연속 부분은 19개 뉴클레오타이드이다.

[0192] [0217] 일부 구현예에서, 표적 부위는 모든 B형 간염 바이러스 게놈과 적어도 90% 상동적이다. 일부 구현예에서, 표적 부위는 표적 부위는 최소 500개, 최소 1000개, 최소 1500개, 최소 2000개, 최소 2500개, 최소 3000개, 최소 3500개, 최소 4000개, 최소 4500개, 최소 5000개, 최소 5500개, 최소 6000개, 최소 6500개, 최소 7000개의 B형 간염 바이러스 게놈과 최소 90%의 상동성을 갖는다.

[0193] [0218] 일부 구현예에서, 표적 부위는 표적 부위는 최소 500개, 최소 1000개, 최소 1500개, 최소 2000개, 최소

2500개, 최소 3000개, 최소 3500개, 최소 4000개, 최소 4500개, 최소 5000개, 최소 5500개, 최소 6000개, 최소 6500개, 최소 7000개의 B형 간염 바이러스 게놈과 최소 90%의 상동성을 갖는다.

[0194] [0219] 일부 구현예에서, 표적 부위는 적어도 1000개의 B형 간염 바이러스 게놈과 적어도 90% 상동적이며, 하나 또는 두 개의 불일치를 포함한다. 일부 구현예에서, 표적 부위는 서열번호 35-100 중 어느 하나에 제시된 서열, 서열번호 35-100 중 어느 하나의 적어도 14nt의 연속된 부분, 또는 상기 중 어느 하나의 상보적 서열을 포함한다. 일부 구현예에서, 표적 부위는 길이가 15, 16, 17, 18 또는 19개의 뉴클레오타이드인 서열번호 35-100 중 어느 하나의 연속된 부분, 또는 상기 중 어느 하나의 상보적 서열이다. 일부 구현예에서, 표적 부위는 상기 본원에서 설명한 표적 부위 서열의 전체 또는 연속 부분과 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 이상의 서열 동일성을 갖는 서열이다. 일부 구현예에서, 표적 부위는 서열번호 35-100 중 어느 하나에 제시된 서열이다.

[0195] [0220] 일부 구현예에서, 불일치는 'nnnnnnnnnnnnnNNNNNNNN-NGG'에서 'n'으로 표현되는 것처럼 프로토스페이서 인접 모티프(PAM)의 5' 말단의 처음 12개 nt에 위치한다.

[0196] [0221] 일부 구현예에서, 표적 부위는 적어도 1000개의 B형 간염 바이러스 게놈과 적어도 90%의 상동성을 갖고, 불일치를 전혀 포함하지 않는다. 일부 구현예에서, 표적 부위는 서열번호 1-34 중 어느 하나에 제시된 서열, 서열번호 1-34 중 어느 하나의 적어도 14nt의 연속 부분, 또는 상기 중 어느 하나의 상보적 서열을 포함한다. 일부 구현예에서, 표적 부위는 길이가 15, 16, 17, 18 또는 19개의 뉴클레오타이드인 서열번호 1-34 중 어느 하나의 연속 부분, 또는 상기 중 어느 하나의 상보적 서열이다. 일부 구현예에서, 표적 부위는 상기 본원에서 설명한 표적 부위 서열의 전체 또는 연속 부분과 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 이상의 서열 동일성을 갖는 서열이다. 일부 구현예에서, 표적 부위는 서열번호 1-34 중 어느 하나에 제시된 서열이다.

[0197] [0222] 본원의 임의의 구현예에서, 표적 부위 또는 각각의 표적 부위는 서열번호 175, 138, 192, 152, 118, 125, 185, 63, 116, 124, 35, 82 중 어느 하나에 제시된 서열, 적어도 14개 뉴클레오타이드(nt)의 이의 연속 부분, 또는 상기 중 어느 하나의 상보적 서열을 포함한다. 일부 구현예에서, 표적 부위는 서열번호 175, 138, 192, 152, 118, 125, 185, 63, 116, 124, 35, 82 중 어느 하나의 연속 부분, 즉 길이가 15, 16, 17, 18 또는 19개 뉴클레오타이드이거나 상기 중 어느 하나의 상보적 서열이다. 일부 구현예에서, 표적 부위는 상기 본원에서 설명한 표적 부위 서열의 전체 또는 연속된 부분과 적어도 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100%의 서열 동일성을 갖는 서열이다. 일부 구현예에서, 표적 부위는 서열번호 175, 138, 192, 152, 118, 125, 185, 63, 116, 124, 35, 82 중 어느 하나에 제시된 서열이다. 일부 구현예에서, 세포에서 유전자 발현의 감소 또는 전사체 수준의 변화는 대조군 세포에서 유전자 수준에 비해 90% 이상이다.

[0198] [0223] 본원의 임의의 구현예에서, 표적 부위 또는 각각의 표적 부위는 서열번호 5, 10, 12, 18, 22, 26, 29, 38, 56, 61, 62, 63, 68, 72, 79, 80, 82, 84, 98, 99, 116, 118, 121, 124, 125, 135, 138, 143, 150, 152, 158, 164, 175, 176, 182, 185, 189, 190, 192 중 어느 하나에 제시된 서열, 적어도 14개 뉴클레오타이드(nt)의 이의 연속 부분, 또는 상기의 서열 중 임의의 서열의 상보적 서열을 포함한다. 일부 구현예에서, 표적 부위는 서열번호 5, 10, 12, 18, 22, 26, 29, 38, 56, 61, 62, 63, 68, 72, 75, 79, 80, 82, 84, 98, 99, 116, 118, 121, 124, 125, 135, 138, 143, 150, 152, 158, 164, 175, 176, 182, 185, 189, 190, 192 중 어느 하나의 연속적인 부분, 즉 길이가 15, 16, 17, 18 또는 19개의 뉴클레오타이드이거나, 상기 중 어느 하나의 상보적 서열이다. 일부 구현예에서, 표적 부위는 본원에서 상기 설명한 표적 부위 서열의 전체 또는 연속된 부분과 적어도 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100%의 서열 동일성을 갖는 서열이다. 일부 구현예에서, 표적 부위는 서열번호 5, 10, 12, 18, 22, 26, 29, 38, 56, 61, 62, 63, 68, 72, 75, 79, 80, 82, 84, 98, 99, 116, 118, 121, 124, 125, 135, 138, 143, 150, 152, 158, 164, 175, 176, 182, 185, 189, 190 또는 192 중 어느 하나에 제시된 서열이다. 일부 구현예에서, 세포에서 유전자 발현의 감소 또는 전사체 수준의 변화는 대조군 세포에서의 유전자 수준에 비해 80% 이상이다.

[0199] [0224] 본원의 임의의 구현예에서, 표적 부위 또는 각각의 표적 부위는 서열번호 5, 6, 12, 18, 22, 26, 29, 38, 42, 43, 51, 56, 61, 63, 68, 72, 75, 79, 82, 84, 88, 89, 98, 99, 113, 116, 121, 124, 125, 118, 130, 133, 135, 138, 143, 150, 152, 155, 158, 164, 165, 175, 176, 182, 185, 189, 190, 192 중 어느 하나에 제시된 서열, 적어도 14개 뉴클레오타이드(nt)로 이루어진 그 연속 부분 또는 앞서 언급한 것의 상보적 서열을 포함한다. 일부 구현예에서, 표적 부위는 서열번호 5, 6, 12, 18, 22, 26, 29, 38, 42, 43, 51, 56, 61, 63, 68,

72, 75, 79, 82, 84, 88, 89, 98, 99, 113, 116, 121, 124, 125, 118, 130, 133, 135, 138, 143, 150, 152, 155, 158, 164, 165, 175, 176, 182, 185, 189, 190, 192 중 어느 하나의 연속 부분, 즉 길이가 15, 16, 17, 18 또는 19개 뉴클레오타이드이거나, 상기 중 어느 것의 상보적 서열이다. 일부 구현예에서, 표적 부위는 본원에서 상기 설명한 표적 부위 서열의 전체 또는 연속된 부분과 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9% 또는 100% 이상의 서열 동일성을 갖는 서열이다. 일부 구현예에서, 표적 부위는 서열번호 5, 6, 12, 18, 22, 26, 29, 38, 42, 43, 51, 56, 61, 63, 68, 72, 75, 79, 82, 84, 88, 89, 98, 99, 113, 116, 121, 124, 125, 118, 130, 133, 135, 138, 143, 150, 152, 155, 158, 164, 165, 175, 176, 182, 185, 189, 190 또는 192 중 어느 하나에 제시된 서열이다. 일부 구현예에서, 세포의 유전자 발현 또는 전사체 수준의 변화는 대조군 세포의 유전자 수준과 비교하여 75% 이상인 경우를 말한다.

[0200] [0225] 본원의 임의의 구현예에서, 표적 부위 또는 각각의 표적 부위는 서열번호 22, 63, 75, 99, 116, 124, 138, 143, 150, 152, 175, 176, 192 중 어느 하나에 제시된 서열, 적어도 14개 뉴클레오타이드(nt)의 그의 연속 부분, 또는 상기 중 어느 것의 상보적 서열을 포함한다. 일부 구현예에서, 표적 부위는 서열번호 22, 63, 75, 99, 116, 124, 138, 143, 150, 152, 175, 176, 192 서열, 즉 길이가 15, 16, 17, 18 또는 19개 뉴클레오타이드인 서열, 또는 상기 중 어느 것의 상보적 서열을 포함한다. 일부 구현예에서, 표적 부위는 상기 본원에서 설명한 표적 부위 서열의 전체 또는 연속 부분과 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 이상의 서열 동일성을 갖는 서열이다. 일부 구현예에서, 표적 부위는 서열번호 22, 63, 75, 99, 116, 124, 138, 143, 150, 152, 175, 176, 또는 192 중 어느 하나에 제시된 서열이다.

[0201] [0226] 본원의 임의의 구현예에서, 표적 부위 또는 각각의 표적 부위는 서열번호 12, 18, 20, 22, 26, 27, 46, 50, 63, 66, 73, 79, 185, 192 중 어느 하나에 제시된 서열, 적어도 14개 뉴클레오타이드(nt)로 이루어진 그의 연속 부분, 또는 앞서 언급한 것의 상보적 서열을 포함한다. 일부 구현예에서, 표적 부위는 서열번호 12, 18, 20, 22, 26, 27, 46, 50, 63, 66, 73, 79, 185, 192 중 어느 하나의 연속 부분, 즉 길이가 15, 16, 17, 18 또는 19개 뉴클레오타이드이거나, 상기 중 어느 것의 상보적 서열이다. 일부 구현예에서, 표적 부위는 본원에서 상기 설명한 표적 부위 서열의 전체 또는 연속된 부분과 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9% 또는 100% 이상의 서열 동일성을 갖는 서열이다. 일부 구현예에서, 표적 부위는 서열번호 12, 18, 20, 22, 26, 27, 46, 50, 63, 66, 73, 79, 185 또는 192 중 어느 하나에 제시된 서열이다.

[0202] [0227] 일부 구현예에서, 표적 부위 또는 각각의 표적 부위는 서열번호 1028-1055 중 어느 하나에 제시된 뉴클레오타이드 서열, 적어도 12nt의 그의 연속 부분, 또는 상기 중 어느 하나의 상보적 서열을 포함한다. 일부 구현예에서, 표적 부위는 길이가 13, 14, 16, 16, 17 또는 18개 뉴클레오타이드인 서열번호 1028-1055 중 임의의 것의 연속 부분, 또는 상기 중 어느 하나의 상보적 서열이다. 일부 구현예에서, 표적 부위는 서열번호 1028-1055 중 어느 하나에 제시된 서열이다.

[0203] [0228] 일부 구현예에서, 표적 부위 또는 각각의 표적 부위는 서열번호 22에 제시된 서열, 적어도 14개 뉴클레오타이드(nt)로 이루어진 그의 연속 부분, 또는 상기 서열의 상보적 서열을 포함한다. 본원의 구현예 중 어느 하나에서, 표적 부위 또는 각각의 표적 부위는 길이가 14-19개 뉴클레오타이드(nt)인 서열번호 22에 제시된 서열의 연속 부분 또는 상기 서열의 상보적 서열을 포함한다. 일부 구현예에 있어서, 표적 부위는 서열번호 22에 제시된 서열이다. 일부 구현예에 있어서, 표적 부위는 본원에 제공된 DNA-표적화 시스템에 의해 표적화될 수 있다. 일부 구현예에 있어서, DNA-결합 도메인은 dSpCas9인 dCas9이며 표적 부위를 표적으로 삼기 위해 상보적 gRNA와 함께 사용된다. 일부 구현예에 있어서, gRNA는 서열번호 217에 제시된 스페이서 서열 또는 표적 부위에 상보적인 연속 부분을 갖는다. 일부 구현예에 있어서, gRNA는 또한 서열번호 587에 제시된 dSpCas9에 대한 스캐폴드 서열을 포함한다. 일부 구현예에서, DNA-표적화 시스템은 본원에서 설명하는 효과기 도메인을 갖는 dSpCas9 융합 단백질과 서열번호 22(예컨대 HBVg_22)에 명시된 gRNA를 포함한다.

[0204] [0229] 일부 구현예에서, 표적 부위 또는 각각의 표적 부위는 서열번호 63에 제시된 서열, 적어도 14개 뉴클레오타이드(nt)의 그의 연속 부분 또는 상기의 상보적 서열을 포함한다. 본원의 임의의 구현예에서, 표적 부위 또는 각각의 표적 부위는 길이가 14-20개 뉴클레오타이드(nt)인 서열번호 63에 제시된 서열의 연속 부분 또는 상기의 상보적 서열을 포함한다. 일부 구현예에서, 표적 부위는 서열번호 63에 제시된 서열이다. 일부 구현예에서, 표적 부위는 본원에 제공된 DNA-표적화 시스템에 의해 표적화될 수 있다. 일부 구현예에서, DNA-결합 도메인은 dSpCas9인 dCas9이며 표적 부위를 표적으로 삼기 위해 상보적 gRNA와 함께 사용된다. 일부 구현예에서, gRNA는 서열번호 217에 제시된 스페이서 서열 또는 표적 부위에 상보적인 그 연속 부분을 갖는다. 일부 구현예에서, gRNA는 또한 SEQ DI NO:587에 제시된 SpCas9에 대한 스캐폴드 서열을 포함한다. 일부 구현예에서, DNA-표적화 시스템은 본원에서 설명하는 효과기 도메인을 갖는 dSpCas9 융합 단백질 및 서열번호 63에 제시된

gRNA(예컨대 HBVg_63)를 포함한다.

[0205] [0230] 본원의 임의의 구현예에서, 표적 부위 또는 각각의 표적 부위는 서열번호 1045에 제시된 서열, 적어도 12개 뉴클레오타이드(nt)의 그의 연속 부분 또는 상기의 상보적 서열을 포함한다. 본원의 임의의 구현예에서, 표적 부위 또는 각각의 표적 부위는 길이가 12-18개 뉴클레오타이드(nt)인 서열번호 1045에 제시된 서열의 연속 부분 또는 상기의 상보적 서열을 포함한다. 일부 구현예에서, 표적 부위는 서열번호 1045에 제시된 서열이다. 일부 구현예에서, DNA-결합 도메인은 표적 부위를 표적화하기 위한 eZFP이다. 일부 구현예에서, ZFP는 서열번호 822, 823, 824, 825, 826 및 827에 제시된 인식 모티프를 포함한다. 일부 구현예에서, eZFP는 서열번호 709에 제시된 서열을 갖는다. 일부 구현예에서, eZFP는 eZFP_18로 지정된 eZFP이다.

[0206] [0231] 본원의 임의의 구현예에서, 표적 부위 또는 각각의 표적 부위는 서열번호 1046에 제시된 서열, 적어도 12개 뉴클레오타이드(nt)의 그의 연속 부분 또는 상기의 상보적 서열을 포함한다. 본원의 임의의 구현예에서, 표적 부위 또는 각각의 표적 부위는 길이가 12-18개 뉴클레오타이드(nt)인 서열번호 1046에 제시된 서열의 연속 부분 또는 상기의 상보적 서열을 포함한다. 일부 구현예에서, 표적 부위는 서열번호 1046에 제시된 서열이다. 일부 구현예에서, DNA-결합 도메인은 표적 부위를 표적화하기 위한 eZFP이다. 일부 구현예에서, ZFP는 서열번호 828, 829, 830, 831, 832 및 833에 제시된 인식 모티프를 포함한다. 일부 구현예에서, eZFP는 서열번호 710에 제시된 서열을 갖는다. 일부 구현예에서, eZFP는 eZFP_19로 지정된 eZFP이다.

[0207] [0232] 본원의 임의의 구현예에서, 표적 부위 또는 각각의 표적 부위는 서열번호 1052에 제시된 서열, 적어도 12개 뉴클레오타이드(nt)의 그의 연속 부분 또는 상기의 상보적 서열을 포함한다. 본원의 임의의 구현예에서, 표적 부위 또는 각각의 표적 부위는 길이가 12-18개 뉴클레오타이드(nt)인 서열번호 1052에 제시된 서열의 연속 부분 또는 상기의 상보적 서열을 포함한다. 일부 구현예에서, 표적 부위는 서열번호 1052에 제시된 서열이다. 일부 구현예에서, DNA-결합 도메인은 표적 부위를 표적화하기 위한 eZFP이다. 일부 구현예에서, ZFP는 서열번호 864, 865, 866, 867, 868 및 869에 제시된 인식 모티프를 포함한다. 일부 구현예에서, eZFP는 서열번호 716에 제시된 서열을 갖는다. 일부 구현예에서, eZFP는 eZFP_25로 지정된 eZFP이다.

[0208] [0233] 일부 구현예에서, 표적 부위는 공유적으로 폐쇄된 원형 DNA(cccDNA) 형태, 이완된 원형 DNA(rcDNA) 형태로 존재하고/존재하거나 인간 게놈 DNA에 통합된다. 일부 구현예에서, 표적 부위를 표적으로 삼으면 HBV 복제(예컨대 감소된 HBV 복제) 및/또는 HBV 전사가 침묵된다.

[0209] [0234] 일부 구현예에서, 본 발명에 따라 본 명세서에 설명된 적어도 두 개의 표적 유전자 또는 이의 조절 DNA 요소의 조합을 표적으로 하는 다중화된 후성적 변형 DNA-표적화 시스템이 제공된다. 일부 구현예에서, 다중화된 후성적 변형 DNA-표적화 시스템은 본 명세서에 설명된 두 개, 세 개, 네 개, 다섯 개, 여섯 개 또는 그 이상의 표적 유전자 또는 이의 조절 DNA 요소를 표적으로 한다.

[0210] [0235] 일부 구현예에서, 제공된 다중화된 후성적 변형 DNA-표적화 시스템에서 표적 부위는 각각 다른 HBV 유전자에 있다. 일부 구현예에서, 표적 부위는 각각 동일한 HBV 유전자에 있다.

[0211] [0236] 일부 구현예에서, 본 발명에 따라 본원에서 기술된 유전자 및/또는 그 조절 요소의 임의의 조합을 표적으로 하는 다중화된 후성적 변형 DNA-표적화 시스템이 제공된다.

[0212] [0237] 일부 구현예에서, 본원에 따라 제1 유전자 또는 이의 조절 요소와 제2 유전자 또는 이의 조절 요소를 표적으로 하는 다중화된 후성적 변형 DNA-표적화 시스템이 제공된다. 일부 구현예에서, 제1 유전자 또는 이의 조절 요소는 폴리머라제 유전자, S-패밀리 유전자, X-유전자, 코어 패밀리 유전자, pre-S1 프로모터, pre-S2 프로모터, X 프로모터, 기저 코어 프로모터, Enh1 인핸서, Enh2 인핸서, 전사체 처리 제어 영역 및 HBV 게놈 내의 모든 코딩 영역으로 구성된 목록에서 선택되고, 제2 유전자 또는 이의 조절 요소는 폴리머라제 유전자, S-패밀리 유전자, X-유전자, 코어 패밀리 유전자, pre-S1 프로모터, pre-S2 프로모터, X 프로모터, 기저 코어 프로모터, Enh1 인핸서, Enh2 인핸서, 전사체 처리 제어 영역 및 HBV 게놈 내의 모든 코딩 영역으로 구성된 목록에서 선택되며, 제1 유전자 또는 이의 조절 요소는 제2 유전자 또는 이의 조절 요소와 상이하다. 제1 및 제2 표적 부위는 전술한 어느 것이든 될 수 있다.

[0213] [0238] 일부 구현예에서, 본원에 따라 제1 조절 요소와 제2 조절 요소를 표적으로 하는 다중화된 후성적 변형 DNA-표적화 시스템이 제공된다. 일부 구현예에서, 제1 조절 요소와 제2 조절 요소는 표 1에 나열된 조합으로부터 선택된다.

표 1. 본원에 제공된 다중화된 후성적 변형 DNA 표적화 시스템이 표적으로 삼는 제 1 조절 요소와 제 2 조절 요소의 조합

제 1 조절 요소	제 2 조절 요소
L-HBs 프로모터	M-HBs 프로모터
L-HBs 프로모터	S-HBs 프로모터
L-HBs 프로모터	X-프로모터/Enh1 프로모터
L-HBs 프로모터	Basal 코어 프로모터/Enh2 인핸서
M-HBs 프로모터	S-HBs 프로모터
M-HBs 프로모터	X-프로모터/Enh1 프로모터
M-HBs 프로모터	Basal 코어 프로모터/Enh2 인핸서
S-HBs 프로모터	X-프로모터/Enh1 프로모터
S-HBs 프로모터	Basal 코어 프로모터/Enh2 인핸서
X-프로모터/Enh1 프로모터	Basal 코어 프로모터/Enh2 인핸서

[0214]

[0215] [0239] 일부 구현예에서, 본 발명은 제1 유전자 또는 이의 조절 요소, 제2 유전자 또는 이의 조절 요소, 그리고 제3 유전자 또는 이의 조절 요소를 표적으로 하는 다중화된 후성적 변형 DNA-표적화 시스템을 제공한다. 일부 구현예에서, 제1 유전자 또는 이의 조절 요소는 폴리머라제 유전자, S-패밀리 유전자, X-유전자, 코어 패밀리 유전자, pre-S1 프로모터, pre-S2 프로모터, X 프로모터, 기저 코어 프로모터, Enh1 인핸서, Enh2 인핸서, 전사체 처리 제어 영역 및 HBV 게놈 내의 임의의 코딩 영역으로 구성된 목록에서 선택되고, 제2 유전자 또는 이의 조절 요소는 폴리머라제 유전자, S-패밀리 유전자, X-유전자, 코어 패밀리 유전자, pre-S1 프로모터, pre-S2 프로모터, X 프로모터, 기저 코어 프로모터, Enh1 인핸서, Enh2 인핸서, 전사체 처리 제어 영역 및 HBV 게놈 내의 임의의 코딩 영역으로 구성된 목록에서 선택되며, 제3 유전자 또는 이의 조절 요소는 폴리머라제 유전자, S-패밀리 유전자, X-유전자, 코어 패밀리 유전자, pre-S1 프로모터, pre-S2 프로모터, X 프로모터, 기저 코어 프로모터, Enh1 인핸서, Enh2 인핸서, 전사 처리 제어 영역으로 구성된 목록에서 선택되고, 제1 유전자 또는 이의 조절 요소, 제2 유전자 또는 이의 조절 요소, 그리고 제3 유전자 또는 이의 조절 요소는 서로 다르다. 제1, 제2 및 제3 표적 부위는 전술한 임의의 것일 수 있다.

[0216] [0240] 일부 구현예에서, 본 발명에 따라 제1 조절 요소, 제2 조절 요소 및 제3 조절 요소를 표적으로 하는 다중화된 후성적 변형 DNA-표적화 시스템이 제공된다. 일부 구현예에서, 제1 조절 요소는 L-HBs 프로모터, M-HBs 프로모터, S-HBs 프로모터, X-프로모터, 기저 코어 프로모터, Enh1 인핸서 및 Enh2 인핸서로 구성된 목록에서 선택되고, 제2 조절 요소는 L-HBs 프로모터, M-HBs 프로모터, S-HBs 프로모터, X-프로모터, 기저 코어 프로모터, Enh1 인핸서 및 Enh2 인핸서로 구성된 목록에서 선택되며, 제3 조절 요소는 L-HBs 프로모터, M-HBs 프로모터, S-HBs 프로모터, X-프로모터, 기저 코어 프로모터, Enh1 인핸서 및 Enh2 인핸서로 구성된 목록에서 선택되고, 제1, 제2 및 제3 조절 요소는 서로 다르다. 일부 구현예에서, 제1 조절 요소는 Enh1 인핸서이고, 제2 조절 요소는 L-HBs 프로모터, M-HBs 프로모터, S-HBs 프로모터, X-프로모터, 기저 코어 프로모터, 및 Enh2 인핸서로 구성된 목록에서 선택되며, 제3 조절 요소는 L-HBs 프로모터, M-HBs 프로모터, S-HBs 프로모터, X-프로모터, 기저 코어 프로모터, 및 Enh2 인핸서로 구성된 목록에서 선택되고, 제2 조절 요소와 그 제3 조절 요소는 다르다. 일부 구현예에서, 제1 조절 요소는 Enh1 인핸서이고, 제2 조절 요소는 L-HBs이며, 제3 조절 요소는 M-HBs 프로모터, S-HBs 프로모터, X-프로모터, 기저 코어 프로모터, 및 Enh2 인핸서로 구성된 목록에서 선택된다.

[0217] [0241] 일부 구현예에서 제1 조절 요소, 제2 조절 요서, 및 제3 조절 요소는 표 2에 나열된 조합으로부터 선택된다.

표 2. 본원에 제공된 다중화된 후성적 변형 DNA 표적화 시스템이 표적으로 삼는 제 1 조절 요소, 제 2 조절 요소 및 제 3 조절 요소의 조합.

제 1 조절 요소	제2 조절 요소	제 3 조절 요소
L-HBs 프로모터	M-HBs 프로모터	S-HBs 프로모터
L-HBs 프로모터	M-HBs 프로모터	X-프로모터/Enh1 프로모터
L-HBs 프로모터	M-HBs 프로모터	Basal 코어 프로모터/Enh2 인핸서
L-HBs 프로모터	S-HBs 프로모터	X-프로모터/Enh1 프로모터
L-HBs 프로모터	S-HBs 프로모터	Basal 코어 프로모터/Enh2 인핸서
L-HBs 프로모터	X-프로모터/Enh1 프로모터	Basal 코어 프로모터/Enh2 인핸서
M-HBs 프로모터	S-HBs 프로모터	X-프로모터/Enh1 프로모터
M-HBs 프로모터	S-HBs 프로모터	Basal 코어 프로모터/Enh2 인핸서
M-HBs 프로모터	X-프로모터/Enh1 프로모터	Basal 코어 프로모터/Enh2 인핸서
S-HBs 프로모터	X-프로모터/Enh1 프로모터	Basal 코어 프로모터/Enh2 인핸서

[0218]

[0219] [0242] 일부 구현예에서, 본 발명은 제1 유전자 또는 이의 조절 요소, 제2 유전자 또는 이의 조절 요소, 제3 유전자 또는 이의 조절 요소, 그리고 제4 유전자 또는 이의 조절 요소를 표적으로 하는 다중화된 후성적 변형 DNA-표적화 시스템을 제공한다. 일부 구현예에서, 제1 유전자 또는 이의 조절 요소는 폴리머라제 유전자, S-패밀리 유전자, X-유전자, 코어 패밀리 유전자, pre-S1 프로모터, pre-S2 프로모터, X 프로모터, 기저 코어 프로모터, Enh1 인핸서, Enh2 인핸서, 전사체 처리 제어 영역 및 HBV 게놈 내의 임의의 코딩 영역으로 구성된 목록에서 선택되고, 제2 유전자 또는 이의 조절 요소는 폴리머라제 유전자, S-패밀리 유전자, X-유전자, 코어 패밀리 유전자, pre-S1 프로모터, pre-S2 프로모터, X 프로모터, 기저 코어 프로모터, Enh1 인핸서, Enh2 인핸서, 전사체 처리 제어 영역 및 HBV 게놈 내의 임의의 코딩 영역으로 구성된 목록에서 선택되며, 제3 유전자 또는 이의 조절 요소는 폴리머라제 유전자, S-패밀리 유전자, X-유전자, 코어 패밀리 유전자, pre-S1 프로모터, pre-S2 프로모터, X 프로모터, 기저 코어 프로모터, Enh1 인핸서, Enh2 인핸서, 전사체 처리 제어 영역으로 구성된 목록에서 선택되고, 제4 유전자 또는 이의 조절 요소는 폴리머라제 유전자, S-패밀리 유전자, X-유전자, 코어 패밀리 유전자, pre-S1 프로모터, pre-S2 프로모터, X 프로모터, 기저 코어 프로모터, Enh1 인핸서, Enh2 인핸서, 전사체 처리 제어 영역으로 구성된 목록에서 선택되고, 제1 유전자 또는 이의 조절 요소, 제2 유전자 또는 이의 조절 요소, 제3 유전자 또는 이의 조절 요소, 그리고 제4 유전자 또는 이의 조절 요소는 서로 다르다. 제1, 제2, 제3 및 제4 표적 부위는 전술한 임의의 것일 수 있다.

[0220] [0243] 일부 구현예에서, 본원에 따라 동일한 유전자 또는 이의 조절 요소를 표적으로 하는 다중화된 후성적 변형 DNA-표적화 시스템이 제공된다. 예를 들어, 두 개 이상의 다중화된 후성적 변형 DNA-표적화 시스템은 동일하거나 공통적인 유전자 또는 이의 조절 요소를 표적으로 한다. 일부 구현예에서, 유전자 또는 이의 조절 요소는 폴리머라제 유전자, S-패밀리 유전자, X-유전자, 코어 패밀리 유전자, pre-S1 프로모터, pre-S2 프로모터, X 프로모터, 기저 코어 프로모터, Enh1 인핸서, Enh2 인핸서, 전사체 처리 제어 영역 및 HBV 게놈 내의 임의의 코딩 영역으로 구성된 목록에서 선택되고, 제2 유전자 또는 이의 조절 요소는 폴리머라제 유전자, S-패밀리 유전자, X-유전자, 코어 패밀리 유전자, pre-S1 프로모터, pre-S2 프로모터, X 프로모터, 기저 코어 프로모터, Enh1 인핸서, Enh2 인핸서, 전사체 처리 제어 영역 및 HBV 게놈 내의 임의의 코딩 영역으로 구성된 목록에서 선택된다. 일부 구현예에서는, 적어도 2개, 적어도 3개, 적어도 4개, 적어도 5개, 적어도 6개, 적어도 7개, 적어도 8개, 적어도 9개, 적어도 10개, 적어도 20개, 적어도 30개의 다중화된 후성적 변형 DNA-표적화 시스템은 동일한 유전자 또는 이의 조절 요소를 표적으로 한다.

[0221] [0244] 일부 구현예에서, 다중화 편집을 위한 표적 부위(예컨대 본원에서 설명하는 다중화 후성적 변형 DNA-표적화 시스템)는 서열번호 12, 18, 20, 22, 26, 27, 46, 50, 63, 66, 73, 79, 185, 192 중 어느 하나에 제시된 서열이다. 일부 구현예에서, 다중화 편집을 위한 표적 부위(예컨대 본원에서 설명하는 다중화 후성적 변형 DNA-표적화 시스템)는 서열번호 22에 제시된 서열이다. 표적 부위는 상기 설명한 바와 같은 것일 수 있다.

[0222] **B. CRISPR 기반 DNA-표적화 시스템**

[0223] [0245] 본원에 따라 CRISPR/Cas 시스템, 즉 표적 유전자 또는 이의 조절 요소의 표적 부위에 결합할 수 있는 CRISPR/Cas 기반 DNA-표적화 시스템을 기반으로 하는 후성적 DNA-표적화 시스템이 제공된다. 일부 구현예에서, 제공된 후성적 DNA-표적화 시스템은 CRISPR/Cas 시스템, 즉 표적 유전자 또는 이의 조절 요소의 조합에서 부위를 표적할 수 있는 CRISPR/Cas 기반 DNA-표적화 시스템을 기반으로 하는 다중화된 후성적 DNA-표적화 시스템이

다.

[0224] [0246] 일부 구현예에서, CRISPR/Cas DNA-결합 도메인은 핵산 분해 효소 비활성이며, 예를 들어 dCas(예컨대 dCas9)를 포함하여 시스템이 표적 유전자 또는 이의 조절 요소의 표적 부위에 결합하면서 표적 부위에서 핵산 절단을 매개하지 않는다. 일부 구현예에서, DNA-표적화 시스템은 유전적 파괴 또는 DNA 파손을 도입하지 않는다. CRISPR/Cas 기반 DNA-표적화 시스템은 간세포와 같은 세포에서 표적 유전자의 발현을 조절하는 데 사용될 수 있다. 일부 구현예에서, 표적 유전자 또는 이의 조절 요소는 본원에서 기술한 바와 같은 임의의 것을 포함할 수 있으며, 여기에는 상기 섹션 I.A.에 기술된 임의의 것이 포함된다. 일부 구현예에서, 표적 유전자 또는 이의 조절 요소의 표적 부위는 섹션 I.A.에서 전술한 임의의 것을 포함하여, 본원에 기재된 임의의 것을 포함할 수 있다. 일부 구현예에서, CRISPR/Cas-기반 DNA-표적화 시스템은 임의의 공지된 Cas 효소 및 일반적으로 뉴클레아제 비활성 또는 dCas를 포함할 수 있다. 일부 구현예에서, CRISPR/Cas-기반 DNA-표적화 시스템은 뉴클레아제 비활성 Cas 단백질 또는 이의 변이체 및 유전자의 전사를 감소시키는 효과기 도메인(예를 들어, 전사 억제자)의 융합 단백질, 및 적어도 하나의 gRNA를 포함한다.

[0225] [0247] CRISPR 시스템(CRISPR/Cas 시스템 또는 CRISPR-Cas 시스템으로도 알려짐)은 침입하는 파지 및 플라스미드에 대해 후천성 면역의 형태를 제공하는 박테리아 및 고세균의 게놈에서 발견되는 보존된 미생물 뉴클레아제 시스템을 지칭한다. 주기적으로 간격을 띠고 분포하는 짧은 회문 구조 반복[Clustered Regularly Interspaced Short Palindromic Repeats, CRISPR]은 스페이서로 불리는 비반복 DNA 서열에 의해 분리되는 다중 반복 DNA 요소를 함유하는 유전자좌를 지칭한다. 스페이서는 과거 노출의 '기억' 역할을 하는 CRISPR 반복 사이의 게놈에 포함된 외래 DNA의 짧은 서열이다. 스페이서는 CRISPR 시스템으로 핵산 절단에 특이성을 부여하는 RNA 분자의 DNA-표적화 부분을 인코딩한다. CRISPR 유전자좌는 절단을 매개하기 위한 RNA-유도 뉴클레아제로서 작용할 수 있는 하나 이상의 CRISPR-관련(Cas) 유전자뿐만 아니라 CRISPR-매개 핵산 절단의 특이성을 프로그래밍할 수 있는 RNA 분자를 인코딩하는 비단백질 인코딩 DNA 요소를 함유하거나 이에 인접한다.

[0226] [0248] Cas 단백질 Cas9가 있는 II형 CRISPR/Cas 시스템에서 두 개의 RNA 분자와 Cas9 단백질은 Cas9 뉴클레아제 활성을 유도하기 위한 리보핵단백질(RNP) 복합체를 형성한다. CRISPR RNA(crRNA)는 표적 핵산 서열(표적 부위)에 상보적이고, 복합체의 서열 특이성을 인코딩하는 스페이서 서열을 함유한다. 트랜스 활성화 crRNA[trans-activating crRNA, tracrRNA]는 crRNA의 부분과 염기쌍을 이루고 Cas9 단백질과 복합체를 형성하여 Cas/RNA RNP 복합체를 형성한다.

[0227] [0249] Cas9를 갖는 시스템과 같은 천연 발생 CRISPR/Cas 시스템은 유전자 편집 및 유전자 발현의 조절 둘 모두를 위해 목적 세포에서 원하는 서열을 표적화하기 위해 Cas/RNA RNP의 효율적인 프로그래밍을 허용하도록 조작되었다. tracrRNA 및 crRNA는 예를 들어 WO 2013/176772 A1, WO 2014/093661 A2, WO 2014/093655 A2, 문헌[Jinek, M. et al. Science 337(6096):816-21 (2012), 또는 Cong, L. et al. Science 339(6121):819-23 (2013)]에 기재된 바와 같이, 보통 가이드 RNA[gRNA]로 지칭되는 단일 키메라 가이드 RNA 분자를 형성하도록 조작되었다. gRNA의 스페이서 서열은 Cas/gRNA RNP 복합체를 원하는 유전자좌, 예를 들어, 표적 유전자 및/또는 그의 조절 요소의 원하는 표적 부위로 표적화하기 위해 사용자에 의해 선택될 수 있다.

[0228] [0250] Cas 단백질은 또한 표적 부위에서 절단을 유도하지 않고 Cas/gRNA RNP를 표적화하도록 설계되었다. Cas 단백질에서의 돌연변이는 Cas 단백질의 뉴클레아제 활성을 감소시키거나 없애 Cas 단백질을 촉매적으로 비활성화시킬 수 있다. 감소되거나 없어진 뉴클레아제 활성을 갖는 Cas 단백질은 본원에서 상호교환적으로 지칭되는, 비활성화된 Cas(dCas), 또는 뉴클레아제-비활성 Cas(iCas) 단백질로 지칭된다. 황색포도상구균에서 유래한 예시적 비활성화된 Cas9(dCas9)는, 예를 들어 WO 2013/176772 A1, WO 2014/093661 A2, 문헌[Jinek, M. et al. Science 337(6096):816-21 (2012), 및 Qi, L. et al. Cell 152(5):1173-83 (2013)]에 기재된 바와 같이, RuvC 및 HNH 뉴클레아제 도메인(D10A 및 H840A)의 침묵화 돌연변이를 함유한다. Cas12 시스템(즉, Cpf1)에서 유래한 예시적인 dCas 변이체는 예를 들어 WO 2017/189308 A1 및 문헌[Zetsche, B. et al. Cell 163(3):759-71 (2015)]에 기재되어 있다. 핵산 절단을 매개하는 보존된 도메인, 예컨대 RuvC 및 HNH 내부뉴클레아제 도메인은 Cas 상동체(orthologues)에서 쉽게 확인가능하고, 예를 들어 문헌[Zetsche, B. et al. Cell 163(3):759-71 (2015)]에 기재된 바와 같이 돌연변이되어 비활성 변이체를 생성할 수 있다.

[0229] [0251] 전사 및/또는 후성적 조절자를 갖는 dCas-융합 단백질은 표적 세포에서 유전자 발현을 이소적으로 조절하기 위한 다용도 플랫폼으로 사용되어 왔다. 이는 Cas와 효과기 도메인, 예컨대 전사 활성인자 또는 전사 억제자의 융합을 포함한다. 예를 들어, dCas9를 전사 활성인자, 예컨대 VP64(단순 포진 바이러스의 16개 아미노산 트랜스활성화 도메인인 VP16의 4개의 탠덤 사본으로 구성된 폴리펩타이드)와 융합하면 유전자 발현이 강력하게

유도될 수 있다. 또한, dCas9를 전사 억제자, 예컨대 크루펠 연관 박스(Krueppel associated box, KRAB)와 융합하면 유전자 발현이 강력하게 억제될 수 있다. 전사 및 후성적 조절자를 갖는 다양한 dCas-융합 단백질은, 예를 들어 WO 2014/197748, WO 2016/130600, WO 2017/180915, WO 2021/226555, WO 2013/176772, WO 2014/152432, WO 2014/093661, WO 2021/247570, 문헌[Adli, M. Nat. Commun. 9, 1911 (2018), Perez-Pinera, P. et al. Nat. Methods 10, 973-976 (2013), Mali, P. et al. Nat. Biotechnol. 31, 833-838 (2013), Maeder, M. L. et al. Nat. Methods 10, 977-979 (2013), Gilbert, L. A. et al. Cell 154(2):442-451 (2013), 및 Nunez, J.K. et al. Cell 184(9):2503-2519 (2021)]에 기재된 바와 같이 유전자 발현의 조절을 위해 조작될 수 있다.

[0230] [0252] 일부 측면에서, 뉴클레아제 비활성 Cas 단백질 또는 이의 변이체를 포함하는 DNA-결합 도메인, 및 세포(예를 들어, 간세포) 내의 표적 유전자 또는 이의 조절 요소에 표적화되는 경우 전사를 감소시키거나 전사 억제를 유도하기 위한 효과기 도메인(즉, 전사 억제자)을 포함하는 융합 단백질을 포함하는 DNA-표적화 시스템이 제공된다. 이러한 구현예에서, DNA-표적화 시스템은 또한 표적 유전자 또는 이의 조절 요소의 표적 부위에 대한 DNA-표적화 시스템의 표적화를 위해 dCas 단백질 또는 이의 변이체와 조합되거나 복합체로서 제공되는 하나 이상의 gRNA를 포함한다. 일부 구현예에서, 융합 단백질은 가이드 RNA에 의해 표적 유전자 또는 이의 조절 요소의 특정 표적 부위 서열로 가이드되되, 효과기 도메인은 표적 유전자의 전사를 감소시키거나 억제하기 위해 표적화된 후성적 변형을 매개한다. 일부 구현예에서, gRNA의 조합은 융합 단백질을 유전자또는 이의 조절 서열의 조합의 표적 부위 서열의 조합으로 가이드하고, 여기서 효과기 도메인은 표적 유전자 조합의 전사를 감소시키거나 억제하기 위해 표적화된 후성적 변형을 매개한다. 전사를 감소시키거나 억제하는 다양한 효과기 도메인 중 임의의 것이 아래에서 추가로 설명되는 바와 같이 사용될 수 있다.

[0231] *1. CRISPR 기반 DNA-결합 도메인*

[0232] [0253] 일부 측면에서, DNA-결합 도메인은 CRISPR 관련(Cas) 단백질 또는 이의 변이체를 포함하거나 Cas 단백질 또는 이의 변이체에서 유래되며 뉴클레아제가 비활성이다(즉, dCas 단백질이다).

[0233] [0254] 일부 구현예에서, Cas 단백질은 클래스 1 CRISPR 시스템(즉, 다중 Cas 단백질 시스템), 예컨대 I형, III형, 또는 IV형 CRISPR 시스템에서 유래한다. 일부 구현예에서, Cas 단백질은 클래스 2 CRISPR 시스템(즉, 다중 Cas 단백질 시스템), 예컨대 II형, V형, 또는 VI형 CRISPR 시스템에서 유래한다. 일부 구현예에서, Cas 단백질은 V형 CRISPR 시스템에서 유래한다. 일부 구현예에서, Cas 단백질은 예를 들어, WO 2017/189308 A1 및 문헌[Zetsche, B. et al. Cell. 163(3):759-71 (2015)]에 기재된 Cas12 단백질(즉, Cpf1) 또는 이의 변이체에서 유래한다. 일부 구현예에서, Cas 단백질은 II형 CRISPR 시스템에서 유래한다. 일부 구현예에서, Cas 단백질은 예를 들어 WO 2013/176772 A1, WO 2014/152432 A2, WO 2014/093661 A2, WO 2014/093655 A2, 문헌[Jinek, M. et al. Science 337(6096):816-21 (2012), Mali, P. et al. Science 339(6121):823-6 (2013), Cong, L. et al. Science 339(6121):819-23 (2013), Perez-Pinera, P. et al. Nat. Methods 10, 973-976 (2013), 또는 Mali, P. et al. Nat. Biotechnol. 31, 833-838 (2013)]에 기재된 Cas9 단백질 또는 이의 변이체에서 유래한다. 유전자 편집 및 조절에 사용하기 위한 다양한 CRISPR/Cas 시스템 및 연관된 Cas 단백질은 예를 들어 문헌[Moon, S.B. et al. Exp. Mol. Med. 51, 1-11 (2019), Zhang, F. Q. Rev. Biophys. 52, E6 (2019), 및 Makarova K.S. et al. Methods Mol. Biol. 1311:47-75 (2015)]에 기재되어 있다.

[0234] [0255] 일부 구현예에서, dCas9 단백질은 천연 발생 Cas9 분자 또는 이의 변이체에서 유래한 서열을 포함할 수 있다. 일부 구현예에서, dCas9 단백질은 화농성연쇄상구균[*S. pyogenes*], 스트렙토코커스 써모필러스(*S. thermophilus*), 황색포도상구균[*S. aureus*], 캠필로박터 제주니(*C. jejuni*), 수막구균[*N. meningitidis*], 프란시셀라 노비시다(*F. novicida*), 연쇄상구균 카니스[*S. canis*], 스타필로코커스 아우리쿨라리스(*S. auricularis*), 또는 이의 변이체의 천연 발생 Cas9 분자에서 유래한 서열을 포함할 수 있다. 일부 구현예에서, dCas9 단백질은 황색포도상구균의 천연 발생 Cas9 분자에서 유래한 서열을 포함한다. 일부 구현예에서, dCas9 단백질은 화농성연쇄상구균의 천연 발생 Cas9 분자에서 유래한 서열을 포함한다.

[0235] [0256] 다른 박테리아 균에서의 Cas9 상동체의 비제한적 예는 아카리오클로리스 마리나(*Acaryochloris marina*) MBIC11017; 아세토할로비움 아라바티쿰(*Acetohalobium arabaticum*) DSM 5501; 아시디티오바실러스 칼두스(*Acidithiobacillus caldus*); 아시디티오바실러스 페로옥시단스(*Acidithiobacillus ferrooxidans*) ATCC 23270; 알리사이클로바실러스 악시도칼다리우스(*Alicyclobacillus acidocaldarius*) LAA1; 알리사이클로바실러스 악시도칼다리우스 아종 악시도칼다리우스[*Alicyclobacillus acidocaldarius subsp. acidocaldarius*] DSM 446; 알로크로마티움 비노섬(*Allochromatium vinosum*) DSM 180; 암모니펙스 데겐시(*Ammonifex degensii*) KC4; 아나바에나

바리알비리스(*Anabaena variabilis*) ATCC 29413; 아트로스피라 맥시마(*Arthrospira maxima*) CS-328; 아르트로스피라 플라텐시스 계통 파라카[*Arthrospira platensis str. Paraca*]; 아르트로스피라 종[*Arthrospira sp.*] PCC 8005; 바실러스 슈도미코이데스(*Bacillus pseudomycoides*) DSM 12442; 바실러스 셀레니티레두센스(*Bacillus selenitireducens*) MLS10; 부르크홀데리알레스 박테리움(*Burkholderiales bacterium*) 1_1_47; 칼디셀룰로시룹토르 벡스키이(*Caldicelulosiruptor becscii*) DSM 6725; 칸디다투스 데설포루디스 아우닥스비아토르(*Candidatus Desulforudis audaxviator*) MP104C; 칼디셀룰로시럽토 하이드로써머리스(*Caldicellulosiruptor hydrothermalis*) 108; 클로스트리디움 파지(Clostridium phage) c-st; 클로스트리디움 보툴리눔 A3 계통 로크 마리[*Clostridium botulinum A3 str. Loch Maree*]; 클로스트리디움 보툴리눔 Ba4 계통 657; 클로스트리듐 디피실레(*Clostridium difficile*) QCD-63q42; 크로코스파이라 왓소니이(*Crocosphaera watsonii*) WH 8501; 시아노테세 종[*Cyanothece sp.*] ATCC 51142; 시아노테세 종 CCY0110; 시아노테세 종 PCC 7424; 시아노테세 종 PCC 7822; 엑시구오박테리움 시비리쿰(*Exiguobacterium sibiricum*) 255-15; 피네골디아 마그나(*Finegoldia magna*) ATCC 29328; 크테노도박데르 라세미페르(*Ktedonobacter racemifer*) DSM 44963; 락토바실러스 델브루키 아종 불가리쿠스[*Lactobacillus delbrueckii subsp. bulgaricus*] PB2003/044-T3-4; 락토바실러스 살리바리우스(*Lactobacillus salivarius*) ATCC 11741; 리스테리아 이노큐아(*Listeria innocua*); 링비아 종[*Lyngbya sp.*] PCC 8106; 마리노박데르 종[*Marinobacter sp.*] ELB17; 메타노할로비움 에베스티가튬(*Methanohalobium evestigatum*) Z-7303; 마이크로시스티스 파지 Ma-LMM01; 마이크로시스티스 아이루기노사(*Microcystis aeruginosa*) NIES-843; 마이크로스실라 마리나(*Microscilla marina*) ATCC 23134; 마이크로콜러스 크토노플라스테스(*Microcoleus chthonoplastes*) PCC 7420; 수막구균; 나이트로소코커스 할로필러스(*Nitrosococcus halophilus*) Nc4; 노카르디옵시스 다손빌레이 아종 다손빌레이[*Nocardiopsis dassonvillei subsp. dassonvillei*] DSM 43111; 노둘라리아 스푸미제나(*Nodularia spumigena*) CCY9414; 노스톡 종[*Nostoc sp.*] PCC 7120; 오실라토리아 종[*Oscillatoria sp.*] PCC 6506; 펠로토마쿨룸 써모프로피오니쿰[*Pelotomaculum thermopropionicum*] SI; 페트로토가 모빌리스[*Petrotoga mobilis*] SJ95; 폴라로모나스 나프탈레니보란스(*Polaromonas naphthalenivorans*) CJ2; 폴라로모나스 종[*Polaromonas sp.*] JS666; 슈도알테로모나스 할로플란크티스(*Pseudoalteromonas haloplanktis*) TAC125; 스트렙토미세스 프리스티나에스피랄리스[*Streptomyces pristinaespiralis*] ATCC 25486; 스트렙토미세스 프리스티나에스피랄리스 ATCC 25486; 스트렙토코커스 써모필러스(*Streptococcus thermophilus*); 스트렙토미세스 비리도크로모게네스(*Streptomyces viridochromogenes*) DSM 40736; 스트렙토스포란기움 로세움(*Streptosporangium roseum*) DSM 43021; 시네코코커스 종[*Synechococcus sp.*] PCC 7335; 및 써모시포 아프리카누스(*Thermosipho africanus*) TCF52B[문헌(Chylinski et al., RNA Biol., 2013; 10(5): 726-737)]에서 확인되는 Cas 단백질을 포함하지만 이에 제한되지 않는다.

[0236] [0257] 일부 측면에서, Cas 단백질은 뉴클레아제 활성이 결여된 변이체(즉, dCas 단백질)이다. 일부 구현예에서, Cas 단백질은 뉴클레아제 활성이 감소되거나 제거되도록 돌연변이가 된다. 이러한 Cas 단백질은 본원에서 상호교환적으로 지칭되는 비활성화 Cas 또는 죽은 Cas(dCas) 또는 뉴클레아제 비활성 Cas(iCas) 단백질로 지칭된다. 일부 구현예에서, 변이체 Cas 단백질은 뉴클레아제 활성이 부족하거나 비활성화된 Cas9(dCas9, 또는 iCas9) 단백질인 변이체 Cas9 단백질이다.

[0237] [0258] 일부 구현예에서, Cas9 단백질 또는 이의 변이체는 황색포도상구균 Cas9[Staphylococcus aureus Cas9, SaCas9] 단백질 또는 이의 변이체에서 유래한다. 일부 구현예에서, 변이체 Cas9는 서열번호 596의 위치 번호를 참조하여 D10A 및 N580A에서 선택되는 적어도 하나의 아미노산 돌연변이를 포함하는 황색포도상구균 dCas9 단백질[dSaCas9]이다. 일부 구현예에서, 변이체 Cas9 단백질은 서열번호 597에 기재된 서열 또는 이에 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99% 서열 동일성을 갖는 아미노산 서열을 포함한다.

[0238] [0259] 일부 구현예에서, Cas9 단백질 또는 이의 변이체는 화농성연쇄상구균 Cas9[Staphylococcus aureus Cas9, SpCas9] 단백질 또는 이의 변이체에서 유래한다. 일부 구현예에서, 변이체 Cas9는 서열번호 598의 위치 번호를 참조하여 D10A 및 H840A에서 선택되는 적어도 하나의 아미노산 돌연변이를 포함하는 화농성연쇄상구균 dCas9 단백질[dSpCas9]이다. 일부 구현예에서, 변이체 Cas9 단백질은 서열번호 599에 기재된 서열 또는 이에 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99% 서열 동일성을 갖는 아미노산 서열을 포함한다.

[0239] ***2. 가이드 RNA***

[0240] [0260] 일부 구현예에서, Cas 단백질(예를 들어, dCas9)은 하나 이상의 가이드 RNA[gRNA]와 조합되어 제공되거나 복합체를 형성하여 제공된다. 일부 측면에서, gRNA는 임의의 전술한 바와 같이, 표적 유전자 및/또는 그의 조절 요소의 표적 부위에 gRNA/Cas RNP 복합체의 특이적 표적화 또는 호밍을 촉진하는 핵산이다. 일부 구현예에

서, gRNA의 표적 부위는 프로토스페이서로 지칭될 수 있다.

[0241] [0261] 섹션 I.A.에서 전술한 임의의 부분과 같이, gRNA 예컨대 표적 유전자 또는 이의 DNA 조절 요소를 표적화하거나 이에 결합하는 gRNA가 본원에 제공된다. 일부 구현예에서, gRNA는 Cas 단백질 또는 이의 변이체와 복합체를 형성할 수 있다. 일부 구현예에서, gRNA는 표적 부위에 혼성화할 수 있거나 표적 부위에 상보적인 gRNA 스페이서 서열(즉, 스페이서 서열 또는 가이드 서열), 예컨대 섹션 I.A 또는 아래에 추가로 기재된 임의의 표적 부위를 포함한다. 일부 구현예에서, gRNA는 Cas 단백질과 복합체를 형성하거나 이에 결합하는 스캐폴드 서열을 포함한다.

[0242] [0262] 일부 구현예에서, 본원에 제공된 gRNA는 키메라 gRNA이다. 일반적으로, gRNA는 단분자(즉, 단일 RNA 분자로 구성됨) 또는 (하나보다는 많고 전형적으로 두 개의 개별 RNA 분자를 포함하는) 모듈일 수 있다. 모듈 gRNA는 단분자로 조작될 수 있되, 개별 모듈 RNA 분자에서의 서열은 때때로 키메라 gRNA, 합성 gRNA, 또는 단일 gRNA로 지칭되는 단일 gRNA 분자에 포함된다. 일부 구현예에서, 키메라 gRNA는 두 개의 비코딩 RNA 서열인 crRNA 서열 및 tracrRNA 서열, 예를 들어 WO 2013/176772 A1 또는 문헌[Jinek, M. et al. Science 337(6096):816-21 (2012)]에 기재된 서열의 융합체이다. 일부 구현예에서, 키메라 gRNA는 II형 효과기 시스템에 관여하는 천연 발생 crRNA:tracrRNA 이중가닥을 모방하되, 천연 발생 crRNA:tracrRNA 이중가닥은 Cas9 단백질에 대한 가이드로서 작용한다.

[0243] [0263] 일부 측면에서, gRNA의 스페이서 서열은 표적 유전자 및/또는 그 조절 요소의 표적 부위와 혼성화하고 표적 부위의 서열에 CRISPR 복합체의 서열 특이적 결합을 지시하기 위해 표적 부위 또는 그 DNA 조절 요소(예컨대 섹션 I.A.에 설명된 것)와 충분한 상보성을 갖는 적어도 일부를 포함하는 폴리뉴클레오타이드 서열이다. 혼성화를 일으키고 CRISPR 복합체의 형성을 촉진하기에 충분한 상보성이 있는 한, 완전한 상보성은 반드시 필요한 것은 아니다. 일부 구현예에서, gRNA는 표적 부위에 상보적인, 예를 들어 적어도 80%, 85%, 90%, 95%, 98%, 99% 또는 100%(예컨대 완전히 상보적)인 스페이서 서열을 포함한다. 표적 부위 서열을 포함하는 표적 핵산의 가닥은 표적 핵산의 "상보적 가닥"으로 지칭될 수 있다.

[0244] [0264] 일부 구현예에서, gRNA 스페이서 서열은 길이가 약 14개 뉴클레오타이드(nt) 내지 약 26개 nt, 또는 16개 nt 내지 22개 nt이다. 일부 구현예에서, gRNA 스페이서 서열은 길이가 14개 nt, 15개 nt, 16개 nt, 17개 nt, 18개 nt, 19개 nt, 20개 nt, 21개 nt 또는 22개 nt, 23개 nt, 24개 nt, 25개 nt, 또는 26개 nt이다. 일부 구현예에서, gRNA 스페이서 서열은 길이가 18개 nt, 19개 nt, 20개 nt, 21개 nt 또는 22개 nt이다. 일부 구현예에서, gRNA 스페이서 서열은 길이가 19개 nt이다.

[0245] [0265] gRNA의 표적 부위는 프로토스페이서로 지칭될 수 있다. 일부 측면에서, gRNA 스페이서는 특정 프로토스페이서 인접 모티프[protospacer-adjacent motif, PAM], 즉 Cas 결합 특이성에 기여하고/하거나 이에 요구되는 프로토스페이서에 바로 인접한 서열을 갖는 프로토스페이서를 표적화하도록 설계된다. CRISPR/Cas 시스템마다 표적화를 위한 PAM 요건이 상이하다. 예를 들어, 일부 구현예에서, 화농성연쇄상구균 Cas9는 PAM 5'-NGG-3'(서열번호 629)을 사용하고, 여기서 N은 임의의 뉴클레오타이드이다. 황색포도상구균 Cas9는 PAM 5'-NNGRRT-3'(서열번호 630)을 사용하고, 여기서 N은 임의의 뉴클레오타이드이고, R은 G 또는 A이다. 수막구균 Cas9는 PAM 5'-NNNNGATT-3'(서열번호 210)을 사용하고, 여기서 N은 임의의 뉴클레오타이드이다. C. 제주니 Cas9는 PAM 5'-NNNNRYAC-3'(서열번호 631)을 사용하고, 여기서 N은 임의의 뉴클레오타이드이고, R은 G 또는 A이고, Y는 C 또는 T이다. S. 써모필러스는 PAM 5'-NNAGAAW-3'(서열번호 632)을 사용하되, N은 임의의 뉴클레오타이드이고, W는 A 또는 T이다. F. 노비시다 Cas9는 PAM 5'-NGG-3'(서열번호 634)을 사용하되, N은 임의의 뉴클레오타이드이다. T. 덴티콜라 Cas9는 PAM 5'-NAAAAC-3'(서열번호 635)을 사용하되, N은 임의의 뉴클레오타이드이다. 다양한 종의 Cas12a(Cpf1으로도 알려짐)는 PAM 5'-TTTV-3'(서열번호 636)을 사용한다. Cas 단백질은 위에 열거된 것과 상이한 PAM을 사용하거나 사용하도록 조작될 수 있다. 예를 들어, 돌연변이된 SpCas9 단백질은 PAM 5'-NGG-3'(서열번호 629), 5'-NGAN-3'(서열번호 637), 5'-NGNG-3'(서열번호 638), 5'-NGAG-3'(서열번호 639), 또는 5'-NGCG-3'(서열번호 640)을 사용할 수 있다. 일부 구현예에서, S. pyogenes Cas9 또는 이의 변이체와 복합체를 이루기 위한 gRNA의 프로토스페이서는 서열번호 588에 제시되어 있다. 일부 구현예에서, S. aureus Cas9 또는 이의 변이체와 복합체를 이루기 위한 gRNA의 프로토스페이서는 서열번호 589에 제시되어 있다.

[0246] [0266] 스페이서 서열은 스페이서 서열 내의 2차 구조의 정도를 감소시키도록 선택될 수 있다. 2차 구조는 임의의 적합한 폴리뉴클레오타이드 폴딩 알고리즘으로 결정될 수 있다.

[0247] [0267] 일부 구현예에서, (가이드 서열을 포함한) gRNA는 염기 우라실(uracil, U)을 포함할 것인 반면, gRNA 분자를 인코딩하는 DNA는 염기 티민(thymine, T)을 포함할 것이다. 이론에 얽매이고자 하는 것은 아니지만, 일부

구현예에서, 가이드 서열과 표적 서열의 상보성은 gRNA 분자/Cas 분자 복합체와 표적 핵산의 상호작용의 특이성에 기여하는 것으로 여겨진다. 가이드 서열 및 표적 서열 쌍에서, 가이드 서열 내의 우라실 염기는 표적 서열 내의 아데닌 염기와 쌍을 이룰 것으로 이해된다.

[0248] [0268] 일부 구현예에서, gRNA의 뉴클레오타이드 중 하나 이상 또는 전부는 예를 들어, gRNA가 분해에 덜 민감하게 하고/하거나 생체 적합성을 개선하도록 변형을 가질 수 있다. 비제한적인 예로서, gRNA의 골격은 포스포로싸이오에이트 또는 기타 변형(들)으로 변형될 수 있다. 일부 경우에, gRNA의 뉴클레오타이드는 2' 변형, 예를 들어 2-아세틸화, 예를 들어 2' 메틸화, 또는 기타 변형(들)을 포함할 수 있다.

[0249] [0269] gRNA 및 예시적인 표적화 도메인을 설계하는 방법은 예를 들어, 국제 PCT 공개 번호 WO 2014/197748 A2, WO 2016/130600 A2, WO 2017/180915 A2, WO 2021/226555 A2, WO 2013/176772 A1, WO 2014/152432 A2, WO 2014/093661 A2, WO 2014/093655 A2, WO 2015/089427 A1, WO 2016/049258 A2, WO 2016/123578 A1, WO 2021/076744 A1, WO 2014/191128 A1, WO 2015/161276 A2, WO 2017/193107 A2 및 WO 2017/093969 A1에 기재된 것을 포함할 수 있다.

[0250] [0270] 일부 구현예에서, 본원에 제공된 gRNA는 공유적으로 폐쇄된 원형 DNA(cccDNA) 형태, 이완된 원형 DNA(rcDNA)로 존재하는 표적 부위를 표적으로 한다. 일부 측면에서, 표적 부위는 HBV RNA 전사의 침묵을 초래하는 DNA 메틸화를 증착하는 데 최적의 HBV 게놈 서열일 수 있다.

[0251] [0271] 일부 구현예에서, 표적 부위는 HBV 복제 및/또는 HBV 전사를 제어하는 데 관여하는 유전자 또는 이의 조절 요소에 또는 그 근처에 있다. 일부 측면에서, 표적 부위는 프로모터에 또는 그 근처에 있다. 일부 측면에서, 표적 부위는 인핸서 영역에 또는 그 근처에 있다. 일부 측면에서, 표적 부위는 전사 처리 제어 영역에 또는 그 근처에 있다. 일부 측면에서, 표적 부위는 HBV 전사를 침묵시키기 위한 DNA 메틸화를 축적하는 데 최적의 유전자일 수 있다.

[0252] [0272] 일부 구현예에서, gRNA는 표 3에 나타낸 바와 같이 서열번호 1-34 중 어느 하나에서 선택된 서열, 적어도 14개 뉴클레오타이드의 그의 연속 부분, 상기 중 어느 하나의 상보적 서열, 또는 상기 중 임의의 것과 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 이상의 서열 동일성을 갖는 서열을 포함하는 표적 부위를 표적으로 한다.

[0253] [0273] 일부 구현예에서, gRNA는 표 4에 나타낸 바와 같이 서열번호 35-100 중 하나에서 선택된 서열을 포함하는 표적 부위 또는 적어도 14nt의 그의 연속 부분, 또는 앞서 언급된 서열 중 하나와 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 이상의 서열 동일성을 갖는 서열을 표적으로 한다.

[0254] [0274] 일부 구현예에서, 표 5에 나타낸 바와 같이 서열번호 101-195 중 하나에서 선택된 서열, 적어도 14개 뉴클레오타이드의 이의 연속 부분, 상기 중 어느 하나의 상보적 서열, 또는 상기 중 어느 하나와 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 이상의 서열 동일성을 갖는 서열을 포함하는 표적 부위를 표적으로 한다.

[0255] [0275] 일부 구현예에서, gRNA는 추가로 서열번호 587에 제시된 스캐폴드 서열을 포함한다. 일부 구현예에서, gRNA는 추가로 스캐폴드 서열을 포함한다. 일부 구현예에서, 스캐폴드 서열은 서열번호 587에 제시된 서열(GUUUAAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC), 또는 그의 전체 또는 일부와 적어도 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100%의 서열 동일성을 갖는 서열을 포함한다. 일부 구현예에서, 스캐폴드 서열은 서열번호 587에 제시된다.

[0256] [0276] 일부 구현예에서, gRNA는 표 6에 나타낸 바와 같이 서열번호 391-585 중 어느 하나에서 선택된 서열, 또는 서열번호 391-585 중 어느 하나와 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 이상의 서열 동일성을 갖는 서열을 포함한다. 일부 구현예에서, gRNA는 서열번호 391-585 중 하나에 제시된다.

[0257] [0277] 일부 구현예에서, gRNA는 서열번호 391-424 중 어느 하나로부터 선택된 서열, 또는 앞서 언급된 서열 중 어느 하나와 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 이상의 서열 동일성을 갖는 서열을 포함한다.

[0258] [0278] 일부 구현예에서, gRNA는 서열번호 425-490 중 어느 하나로부터 선택된 서열, 또는 앞서 언급된 서열 중 어느 하나와 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 이상의

서열 동일성을 갖는 서열을 포함한다.

[0259] [0279] 일부 구현예에서, gRNA는 서열번호 490-585 중 어느 하나로부터 선택된 서열, 또는 앞서 언급된 서열 중 어느 하나와 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 이상의 서열 동일성을 갖는 서열을 포함한다.

[0260] [0280] 일부 구현예에서, 제공된 gRNA 서열 중 어느 것이든 Cas9와 복합되거나 Cas9와 함께 제공된다. 일부 구현예에서, Cas9는 dCas9이다. 일부 구현예에서, dCas9는 dSpCas9, 예를 들어 서열번호 599에 제시된 dSpCas9이다.

표 3. HBV 게놈 보존도가 >90%이고 불일치가 없는 표적 부위 서열 및 gRNA 스페이서(HBV0)

[0261]

표적 부위 (프로토스페이서) 서열	표적 SEQ ID	서열명	중앙 HBV 위치	RNA 스페이서 서열	RNA 스페이서 SEQ ID
CCCTATCTTATCAACACTTC	1	HBVg_1	2310	CCCUAUCUUAUCAACACUUC	196
GCAGAGGTGAAAAAGTTGCA	2	HBVg_2	1834	GCAGAGGUGAAAAAGUUGCA	197
TGGACTTCTCTCAATTTTCT	3	HBVg_3	259	UGGACUUCUCUCAAUUUUCU	198
ACCCCGCCTGTAACACGAGC	4	HBVg_4	209	ACCCCGCCUGUAACACGAGC	199
CCCGCCTGTAACACGAGCAG	5	HBVg_5	207	CCCGCCUGUAACACGAGCAG	200
CACCACGAGTCTAGACTCTG	6	HBVg_6	261	CACCACGAGUCUAGACUCUG	201
GAGGTGAAGCGAAGTGCACA	7	HBVg_7	1597	GAGGUGAAGCGAAGUGCACA	202
CCGGAAGTGTTGATAAGATA	8	HBVg_8	2333	CCGGAAGUGUUGAUAAGAUA	203
AGAAGATGAGGCATAGCAGC	9	HBVg_9	434	AGAAGAUGAGGCAUAGCAGC	204
TCCGCAGTATGGATCGGCAG	10	HBVg_10	1277	UCCGCAGUAUGGAUCGGCAG	205
GGACTTCTCTCAATTTTCTA	11	HBVg_11	260	GGACUUCUCUCAAUUUUCUA	206
CCACCCAAGGCACAGCTTGG	12	HBVg_12	1891	CCACCCAAGGCACAGCUUGG	207
AGAGAGGTGCGCCCCGTGGT	13	HBVg_13	1535	AGAGAGGUGCGCCCCGUGGU	208
GATTGAGATCTTCTGCGACG	14	HBVg_14	2433	GAUUGAGAUCUUCUGCGACG	209
CAAGCCTCCAAGCTGTGCCT	15	HBVg_15	1864	CAAGCCUCCAAGCUGUGCCU	210
GGCGAGGGAGTTCTTCTTCT	16	HBVg_16	2388	GGCGAGGGAGUUCUUCUUCU	211
TCCGGAAGTGTTGATAAGAT	17	HBVg_17	2334	UCCGGAAGUGUUGAUAAGAU	212
AAGCCACCCAAGGCACAGCT	18	HBVg_18	1894	AAGCCACCCAAGGCACAGCU	213
CCTCCAAGCTGTGCCTTGGG	19	HBVg_19	1868	CCUCCAAGCUGUGCCUUGGG	214
GTAAAGAGAGGTGCGCCCCG	20	HBVg_20	1539	GUAAAGAGAGGUGCGCCCCG	215
GGCAGATGAGAAGGCACAGA	21	HBVg_21	1571	GGCAGAUGAGAAGGCACAGA	216
AGGAGTTCCGCAGTATGGAT	22	HBVg_22	1283	AGGAGUUCCGCAGUAUGGAU	217
GCTGTGCCTTGGGTGGCTTT	23	HBVg_23	1875	GCUGUGCCUUGGGUGGCUUU	218
ACCCCTGCTCGTGTTACAGG	24	HBVg_24	183	ACCCCUGCUCGUGUUACAGG	219
CGGAAGTGTTGATAAGATAG	25	HBVg_25	2332	CGGAAGUGUUGAUAAGAUAG	220

CCTGCTGGTGGCTCCAGTTC	26	HBVg_26	55	CCUGCUGGUGGCUCCAGUUC	221
CGAGGGAGTTCTTCTTCTAG	27	HBVg_27	2386	CGAGGGAGUUCUUCUUCUAG	222
GGGGCGCACCTCTCTTTACG	28	HBVg_28	1520	GGGGCGCACCUCUCUUUACG	223
AGCTTGGAGGCTTGAACAGT	29	HBVg_29	1878	AGCUUGGAGGCUUGAACAGU	224
AAGCCTCCAAGCTGTGCCTT	30	HBVg_30	1865	AAGCCUCCAAGCUGUGCCUU	225
CCCCTGCTCGTGTTACAGGC	31	HBVg_31	184	CCCCUGCUCGUGUUACAGGC	226
GCGAGGGAGTTCTTCTTCTA	32	HBVg_32	2387	GCGAGGGAGUUCUUCUUCUA	227
TACTAGTGCCATTTGTTCAG	33	HBVg_33	677	UACUAGUGCCAUUUGUUCAG	228
GACTTCTCTCAATTTTCTAG	34	HBVg_34	261	GACUUCUCUCAAUUUUCUAG	229

표 4. HBV 게놈 보존도가 >90%이고 1-2 개의 불일치를 갖는 표적 부위 서열 및 gRNA 스페이서(HBV1)

표적 부위 (프로토스페이서) 서열	표적 SEQ ID	서열명	중앙 HBV 위치	RNA 스페이서 서열	RNA 스페이서 SEQ ID
ATTGACCCGTATAAAGAATT	35	HBVg_35	1906	AUUGACCCGUAUAAAGAAUU	230
ACCCAAAGACAAAAGAAAAT	36	HBVg_36	825	ACCCAAAGACAAAAGAAAAU	231
GTCCTCTTATGTAAGACCTT	37	HBVg_37	1663	GUCCUCUUAUGUAAGACCUU	232
TGATCGGGAAAGAATCCCAG	38	HBVg_38	2930	UGAUCGGGAAAGAAUCCCAG	233
TTTGCTGACGCAACCCCCAC	39	HBVg_39	1182	UUUGCUGACGCAACCCCCAC	234
ATGAATCTAGCCACCTGGGT	40	HBVg_40	2095	AUGAAUCUAGCCACCUGGGU	235
GGTCTCCATGCGACGTGCAG	41	HBVg_41	1616	GGUCUCCAUGCGACGUGCAG	236
GGACTGAGGCCCACTCCCAT	42	HBVg_42	659	GGACUGAGGCCCACUCCCAU	237
CACAGAGTCTAGACTCGTGG	43	HBVg_43	239	CACAGAGUCUAGACUCGUGG	238
GAAGAACCAACAAGAAGATG	44	HBVg_44	446	GAAGAACCAACAAGAAGAUG	239
ACACGGTCCGGCAGATGAGA	45	HBVg_45	1580	ACACGGUCCGGCAGAUGAGA	240
GACATGAACATGAGATGATT	46	HBVg_46	1856	GACAUGAACAUGAGAUGAUU	241
GCAGCACAGCCTAGCAGCCA	47	HBVg_47	1394	GCAGCACAGCCUAGCAGCCA	242
TCCTGGAATTAGAGGACAAA	48	HBVg_48	490	UCCUGGAAUUAGAGGACAAA	243

[0262]

GTCTTACATAAGAGGACTCT	49	HBVg_49	1646	GUCUUACAUAAGAGGACUCU	244
TTGTGGGTCACCATATTCTT	50	HBVg_50	2809	UUGUGGGUCACCAUAUUCUU	245
CGCAAAATACCTATGGGAGT	51	HBVg_51	627	CGCAAAAUACCUAUGGGAGU	246
GGGTTGCGTCAGCAAACACT	52	HBVg_52	1198	GGGUUGCGUCAGCAAACACU	247
AGCTCTTGTTCCCAAGAATA	53	HBVg_53	2842	AGCUCUUGUUCCCAAGAAUA	248
TGACATACTTTCCAATCAAT	54	HBVg_54	992	UGACAUACUUUCCAAUCAAU	249
CAGATGAGAAGGCACAGACG	55	HBVg_55	1569	CAGAUGAGAAGGCACAGACG	250
CCCCGCCTGTAACACGAGCA	56	HBVg_56	208	CCCCGCCUGUAACACGAGCA	251
GGGTGGAGCCCTCAGGCTCA	57	HBVg_57	3072	GGGUGGAGCCCUCAGGCUCA	252
ATTCCTTGGACTCATAAGGT	58	HBVg_58	2453	AUUCCUUGGACUCAUAAGGU	253
TTTGTGGGTCACCATATTCT	59	HBVg_59	2808	UUUGUGGGUCACCAUAUUCU	254
GTGAAAAAGTTGCATGGTGC	60	HBVg_60	1828	GUGAAAAAGUUGCAUGGUGC	255
CCTGAACTGGAGCCACCAGC	61	HBVg_61	78	CCUGAACUGGAGCCACCAGC	256
TCCTCTGCCGATCCATACTG	62	HBVg_62	1253	UCCUCUGCCGAUCCAUACUG	257
CGGCTAGGAGTTCCGCAGTA	63	HBVg_63	1288	CGGCUAGGAGUUCCGCAGUA	258
AATGTCAACGACCGACCTTG	64	HBVg_64	1678	AAUGUCAACGACCGACCUUG	259
GACCTTCGTCTGCGAGGCGA	65	HBVg_65	2403	GACCUUCGUCUGCGAGGCGA	260
GTTGCCGGGCAACGGGGTAA	66	HBVg_66	1162	GUUGCCGGGCAACGGGGUAA	261
GATTGAGACCTTCGTCTGCG	67	HBVg_67	2409	GAUUGAGACCUUCGUCUGCG	262
AGGACCCCTGCTCGTGTTAC	68	HBVg_68	180	AGGACCCCUGCUCGUGUUAC	263
TTTGAAGTATGCCTCAAGGT	69	HBVg_69	1712	UUUGAAGUAUGCCUCAAGGU	264
CCGCTTGTTTTGCTCGCAGC	70	HBVg_70	1285	CCGCUUGUUUUGCUCGCAGC	265
TGCTAGGCTGTGCTGCCAAC	71	HBVg_71	1378	UGCUAGGCUGUGCUGCCAAC	266
TGCCGATTGGTGGAGGCAGG	72	HBVg_72	3146	UGCCGAUUGGUGGAGGCAGG	267
TCTTTGTACTAGGAGGCTGT	73	HBVg_73	1764	UCUUUGUACUAGGAGGCUGU	268
CGTCCCGCGCAGGATCCAGT	74	HBVg_74	1416	CGUCCCGCGCAGGAUCCAGU	269
AAAGCCCAAGATGATGGGAT	75	HBVg_75	627	AAAGCCCAAGAUGAUGGGAU	270
GCAGATGAGAAGGCACAGAC	76	HBVg_76	1570	GCAGAUGAGAAGGCACAGAC	271
CGATTGGTGGAGGCAGGAGG	77	HBVg_77	3143	CGAUUGGUGGAGGCAGGAGG	272

[0263]

AGGAGGCTGTAGGCATAAAT	78	HBVg_78	1774	AGGAGGCUGUAGGCAUAAAU	273
CCATGCCCCAAAGCCACCCA	79	HBVg_79	1904	CCAUGCCCCAAAGCCACCCA	274
AGGTTGGGGACTGCGAATTT	80	HBVg_80	326	AGGUUGGGGACUGCGAAUUU	275
AGACCTTCGTCTGCGAGGCG	81	HBVg_81	2404	AGACCUUCGUCUGCGAGGCG	276
CCTGGAATTAGAGGACAAAC	82	HBVg_82	489	CCUGGAAUUAGAGGACAAAC	277
TTTCAGTTATATGGATGATG	83	HBVg_83	728	UUUCAGUUAUAUGGAUGAUG	278
GTAACACGAGCAGGGGTCCT	84	HBVg_84	200	GUAACACGAGCAGGGGUCCU	279
CATCTTCTTGTTGGTTCTTC	85	HBVg_85	426	CAUCUUCUUGUUGGUUCUUC	280
CGGGGAGACCGCGTAAAGAG	86	HBVg_86	1551	CGGGGAGACCGCGUAAAGAG	281
CTAGACTCTGTGGTATTGTG	87	HBVg_87	251	CUAGACUCUGUGGUAUUGUG	282
CCCTGCTCGTGTTACAGGCG	88	HBVg_88	185	CCCUGCUCGUGUUACAGGCG	283
TACCACAGAGTCTAGACTCG	89	HBVg_89	236	UACCACAGAGUCUAGACUCG	284
TCGCAAAATACCTATGGGAG	90	HBVg_90	626	UCGCAAAAUACCUAUGGGAG	285
GTCTGTGCCTTCTCATCTGC	91	HBVg_91	1550	GUCUGUGCCUUCUCAUCUGC	286
ACACGTAGCGCCTCATTTTG	92	HBVg_92	2792	ACACGUAGCGCCUCAUUUUG	287
TTGGGGTTGAGGTCCCAATC	93	HBVg_93	2990	UUGGGGUUGAGGUCCCAAUC	288
CCCCGAGACGGGTCGTCCGC	94	HBVg_94	1469	CCCCGAGACGGGUCGUCCGC	289
CCTACGAACCACTGAACAAA	95	HBVg_95	708	CCUACGAACCACUGAACAAA	290
TTACATACTCTGTGGAAGGC	96	HBVg_96	2747	UUACAUACUCUGUGGAAGGC	291
ACCTCCTTTCCATGGCTGCT	97	HBVg_97	1362	ACCUCCUUUCCAUGGCUGCU	292
GTTATCGCTGGATGTGTCTG	98	HBVg_98	365	GUUAUCGCUGGAUGUGUCUG	293
AACATGAGATGATTAGGCAG	99	HBVg_99	1850	AACAUGAGAUGAUUAGGCAG	294
ACTTCTCTCAATTTTCTAGG	100	HBVg_100	262	ACUUCUCUCAAUUUUCUAGG	295

표 5. HBV 게놈 보존도가 70-90%이고 최대 2 개의 불일치를 갖는 표적 부위 서열 및 gRNA 스페이서(HBV2)

표적 부위 (프로토스페이서) 서열	표적 SEQ ID	서열명	중앙 HBV 위치	gRNA 스페이서 서열	RNA 스페이서 SEQ ID

[0264]

CACTTTCTCGCCAACTTACA	101	HBVg_101	1091	CACUUUCUCGCCAACUUACA	296
CATAAGGTGGGAAACTTTAC	102	HBVg_102	2465	CAUAAGGUGGGAAACUUUAC	297
CCAAACCTCGAAAAGGCATG	103	HBVg_103	2863	CCAAACCUCGAAAAGGCAUG	298
ATAGAAGGAAAGAAGTCAGA	104	HBVg_104	1980	AUAGAAGGAAAGAAGUCAGA	299
GCTGCTCCTTTTACACAATG	105	HBVg_105	1017	GCUGCUCCUUUUACACAAUG	300
GAAGCGAAGTGCACACGGTC	106	HBVg_106	1592	GAAGCGAAGUGCACACGGUC	301
GGATCATCAACCACCAGCAC	107	HBVg_107	487	GGAUCAUCAACCACCAGCAC	302
GAGCCAAGAGAAACGGACTG	108	HBVg_108	673	GAGCCAAGAGAAACGGACUG	303
CTTCACCTCTGCACGTCGCA	109	HBVg_109	1588	CUUCACCUCUGCACGUCGCA	304
ACAATGTTCCGGAGACTCTA	110	HBVg_110	2044	ACAAUGUUCCGGAGACUCUA	305
TCCGCGGGATTCAGCGCCGA	111	HBVg_111	1454	UCCGCGGGAUUCAGCGCCGA	306
TTAATGAGTGGGAGGAGTTG	112	HBVg_112	1722	UUAAUGAGUGGGAGGAGUUG	307
CCAACTCAAACAATCCAGAT	113	HBVg_113	2953	CCAACUCAAACAAUCCAGAU	308
GCTGCCAACTGGATCCTGCG	114	HBVg_114	1389	GCUGCCAACUGGAUCCUGCG	309
AGTCTTTGAAGTATGCCTCA	115	HBVg_115	1716	AGUCUUUGAAGUAUGCCUCA	310
TCCTGACTGCCGATTGGTGG	116	HBVg_116	3153	UCCUGACUGCCGAUUGGUGG	311
TCTTGTCCTCCAATTTGTCC	117	HBVg_117	343	UCUUGUCCUCCAAUUUGUCC	312
CGATAACCAGGACAAATTGG	118	HBVg_118	372	CGAUAACCAGGACAAAUUGG	313
ATTTGGAAGATCCAGCATCC	119	HBVg_119	2123	AUUUGGAAGAUCCAGCAUCC	314
CTGTTTGGCTTTCAGTTATA	120	HBVg_120	719	CUGUUUGGCUUUCAGUUAUA	315
CTCCTCCTGCCTCCACCAAT	121	HBVg_121	3121	CUCCUCCUGCCUCCACCAAU	316
GTCATCCTCAGGCCATGCAG	122	HBVg_122	3190	GUCAUCCUCAGGCCAUGCAG	317
CTGCCGTTCCGGCCGACCAC	123	HBVg_123	1500	CUGCCGUUCCGGCCGACCAC	318
CCTTCCTGACTGCCGATTGG	124	HBVg_124	3156	CCUUCCUGACUGCCGAUUGG	319
ACCTGCACGACTCCTGCTCA	125	HBVg_125	520	ACCUGCACGACUCCUGCUCA	320
GGCCTGTATTTTCCTGCTGG	126	HBVg_126	43	GGCCUGUAUUUUCCUGCUGG	321
AACATAGAGGTTCCTTGAGC	127	HBVg_127	555	AACAUAGAGGUUCCUUGAGC	322
TGCCGTTCCGGCCGACCACG	128	HBVg_128	1501	UGCCGUUCCGGCCGACCACG	323
ATAGGCCATCAGCGCATGCG	129	HBVg_129	1218	AUAGGCCAUCAGCGCAUGCG	324

[0265]

GCCTCCACCAATCGGCAGTC	130	HBVg_130	3129	GCCUCCACCAAUCGGCAGUC	325
GTCCTTTGTTTACGTCCCGT	131	HBVg_131	1415	GUCCUUUGUUUACGUCCCGU	326
TGGGAACAAGAGCTACAGCA	132	HBVg_132	2828	UGGGAACAAGAGCUACAGCA	327
CTGTAAACAGGCCTATTGAT	133	HBVg_133	958	CUGUAAACAGGCCUAUUGAU	328
GTCGCAGAAGATCTCAATCT	134	HBVg_134	2414	GUCGCAGAAGAUCUCAAUCU	329
CTGCCTTCCTGACTGCCGAT	135	HBVg_135	3159	CUGCCUUCCUGACUGCCGAU	330
ACTACTAATTCCCTGGATGC	136	HBVg_136	2157	ACUACUAAUUCCCUGGAUGC	331
CACATTTCTTGCCTTACTTT	137	HBVg_137	2209	CACAUUUCUUGCCUUACUUU	332
GGATGACTGTCTCTTAGAGG	138	HBVg_138	3197	GGAUGACUGUCUCUUAGAGG	333
GCTATGCCTCATCTTCTTGT	139	HBVg_139	417	GCUAUGCCUCAUCUUCUUGU	334
CCCGTCGGCGCTGAATCCCG	140	HBVg_140	1430	CCCGUCGGCGCUGAAUCCCG	335
TAGTATTCCTTGGACTCATA	141	HBVg_141	2449	UAGUAUUCCUUGGACUCAUA	336
AGGTAGGAGCGGGAGCATTC	142	HBVg_142	3016	AGGUAGGAGCGGGAGCAUUC	337
TCTTTTGGGGTGGAGCCCTC	143	HBVg_143	3065	UCUUUUGGGGUGGAGCCCUC	338
TCAGTATGCCCTGAGCCTGA	144	HBVg_144	3103	UCAGUAUGCCCUGAGCCUGA	339
TTTAATGAGTGGGAGGAGTT	145	HBVg_145	1721	UUUAAUGAGUGGGAGGAGUU	340
CTCCCTCGCCTCGCAGACGA	146	HBVg_146	2378	CUCCCUCGCCUCGCAGACGA	341
GATAAGATAGGGGCATTTGG	147	HBVg_147	2322	GAUAAGAUAGGGGCAUUUGG	342
CCGCGGGATTCAGCGCCGAC	148	HBVg_148	1453	CCGCGGGAUUCAGCGCCGAC	343
CAGCGATAACCAGGACAAAT	149	HBVg_149	375	CAGCGAUAACCAGGACAAAU	344
CAGGTAGGAGTGGGAGCATT	150	HBVg_150	3021	CAGGUAGGAGUGGGAGCAUU	345
GTTTAATGAGTGGGAGGAGT	151	HBVg_151	1720	GUUUAAUGAGUGGGAGGAGU	346
TGGTGAGTGATTGGAGGTTG	152	HBVg_152	340	UGGUGAGUGAUUGGAGGUUG	347
TAATGAGTGGGAGGAGTTGG	153	HBVg_153	1723	UAAUGAGUGGGAGGAGUUGG	348
ACTACATGTTCTGGATAATA	154	HBVg_154	2719	ACUACAUGUUCUGGAUAAUA	349
GGCATAGCAGCAGGATGAAG	155	HBVg_155	425	GGCAUAGCAGCAGGAUGAAG	350
GTTGATAAGATAGGGGCATT	156	HBVg_156	2325	GUUGAUAAGAUAGGGGCAUU	351
TCAACGAATTGTGGGTCTTT	157	HBVg_157	989	UCAACGAAUUGUGGGUCUUU	352
TATGGATGATGTGGTATTGG	158	HBVg_158	737	UAUGGAUGAUGUGGUAUUGG	353

[0266]

CAACGAATTGTGGGTCTTTT	159	HBVg_159	990	CAACGAAUUGUGGGUCUUUU	354
CATTTGTTCAGTGGTTCGTA	160	HBVg_160	686	CAUUUGUUCAGUGGUUCGUA	355
CGTCTAACAACAGTAGTTTC	161	HBVg_161	2352	CGUCUAACAACAGUAGUUUC	356
TGCCTGAGTGCTGTATGGTG	162	HBVg_162	2070	UGCCUGAGUGCUGUAUGGUG	357
GCCCCGAGACGGGTCGTCCG	163	HBVg_163	1470	GCCCCGAGACGGGUCGUCCG	358
GACTGCCGATTGGTGGAGGC	164	HBVg_164	3149	GACUGCCGAUUGGUGGAGGC	359
TATATGGATGATGTGGTATT	165	HBVg_165	735	UAUAUGGAUGAUGUGGUAUU	360
CTTGAGTATTTGGTGTCTTT	166	HBVg_166	2245	CUUGAGUAUUUGGUGUCUUU	361
GATCTGGTGGGCGTTCACGG	167	HBVg_167	1637	GAUCUGGUGGGCGUUCACGG	362
AGACTGGGAGGAGTTGGGGG	168	HBVg_168	1726	AGACUGGGAGGAGUUGGGGG	363
AGTCCTCTTATGTAAGACCT	169	HBVg_169	1664	AGUCCUCUUAUGUAAGACCU	364
CTCAAGATGTTGTACAGACT	170	HBVg_170	783	CUCAAGAUGUUGUACAGACU	365
GGGAACAAGAGCTACAGCAT	171	HBVg_171	2829	GGGAACAAGAGCUACAGCAU	366
TCGCAGAAGATCTCAATCTC	172	HBVg_172	2415	UCGCAGAAGAUCUCAAUCUC	367
GGGGTGGAGCCCTCAGGCTC	173	HBVg_173	3071	GGGGUGGAGCCCUCAGGCUC	368
TATTCCTTGGACTCATAAGG	174	HBVg_174	2452	UAUUCCUUGGACUCAUAAGG	369
TCTAAGAGACAGTCATCCTC	175	HBVg_175	3179	UCUAAGAGACAGUCAUCCUC	370
CAACTCAAACAATCCAGATT	176	HBVg_176	2954	CAACUCAAACAAUCCAGAUU	371
AAACAAAGGACGTCCCGCGC	177	HBVg_177	1426	AAACAAAGGACGUCCCGCGC	372
CTGCCAACTGGATCCTGCGC	178	HBVg_178	1390	CUGCCAACUGGAUCCUGCGC	373
GAAGCTCCAAATTCTTTATA	179	HBVg_179	1935	GAAGCUCCAAAUUCUUUAUA	374
GTCAGTATGCCCTGAGCCTG	180	HBVg_180	3104	GUCAGUAUGCCCUGAGCCUG	375
TTTCCCACCTTATGAGTCCA	181	HBVg_181	2479	UUUCCCACCUUAUGAGUCCA	376
ACAAGAGGTTGGTGAGTGAT	182	HBVg_182	349	ACAAGAGGUUGGUGAGUGAU	377
GAAAGCCCAAGATGATGGGA	183	HBVg_183	628	GAAAGCCCAAGAUGAUGGGA	378
AGGTTCCACGCATGCGCTGA	184	HBVg_184	1246	AGGUUCCACGCAUGCGCUGA	379
TTGGTGAGTGATTGGAGGTT	185	HBVg_185	341	UUGGUGAGUGAUUGGAGGUU	380
AGAGCTACAGCATGGGAGGT	186	HBVg_186	2836	AGAGCUACAGCAUGGGAGGU	381
ATATGGATGATGTGGTATTG	187	HBVg_187	736	AUAUGGAUGAUGUGGUAUUG	382

[0267]

CCATTTGTTCAGTGGTTCGT	188	HBVg_188	685	CCAUUUGUUCAGUGGUUCGU	383
TTATATGGATGATGTGGTAT	189	HBVg_189	734	UUAUAUGGAUGAUGUGGUAU	384
GGAGTGGGAGCATTCGGGCC	190	HBVg_190	3021	GGAGUGGGAGCAUUCGGGCC	385
ATTTGGTGTCTTTTGGAGTG	191	HBVg_191	2252	AUUUGGUGUCUUUUGGAGUG	386
CACAGAAAGGCCTTGTAAGT	192	HBVg_192	1124	CACAGAAAGGCCUUGUAAGU	387
ACCAATTTTCTTTTGTCTTT	193	HBVg_193	801	ACCAAUUUUCUUUUGUCUUU	388
AGGTTAATGGTCTTTGTACT	194	HBVg_194	1754	AGGUUAAUGGUCUUUGUACU	389
TACCAATTTTCTTTTGTCTT	195	HBVg_195	800	UACCAAUUUUCUUUUGUCUU	390

표 6. 유전자-표적화 gRNA

서열명	유전자-표적화 gRNA	SEQ ID
HBVg_1	CCCUAUCUUAUCAACACUUCGUUUAAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC	391
HBVg_2	GCAGAGGUGAAAAAGUUGCAGUUUAAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC	392
HBVg_3	UGGACUUCUCUCAAUUUUCUGUUUAAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC	393
HBVg_4	ACCCCGCCUGUAACACGAGCGUUUAAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC	394
HBVg_5	CCCGCCUGUAACACGAGCAGGUUUAAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC	395
HBVg_6	CACCACGAGUCUAGACUCUGGUUUAAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC	396
HBVg_7	GAGGUGAAGCGAAGUGCACAGUUUAAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC	397
HBVg_8	CCGGAAGUGUUGAUAAGAUAGUUUAAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC	398
HBVg_9	AGAAGAUGAGGCAUAGCAGCGUUUAAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC	399
HBVg_10	UCCGCAGUAUGGAUCGGCAGGUUUAAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC	400
HBVg_11	GGACUUCUCUCAAUUUUCUAGUUUAAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC	401
HBVg_12	CCACCCAAGGCACAGCUUGGGUUUAAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC	402

[0268]

HBVg_13	AGAGAGGUGCGCCCCGUGGUGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	403
HBVg_14	GAUUGAGAUCUUCUGCGACGGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	404
HBVg_15	CAAGCCUCCAAGCUGUGCCUGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	405
HBVg_16	GGCGAGGGAGUUCUUCUUCUGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	406
HBVg_17	UCCGGAAGUGUUGAUAAGAUGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	407
HBVg_18	AAGCCACCCAAGGCACAGCUGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	408
HBVg_19	CCUCCAAGCUGUGCCUUGGGGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	409
HBVg_20	GUAAAGAGAGGUGCGCCCCGGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	410
HBVg_21	GGCAGAUGAGAAGGCACAGAGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	411
HBVg_22	AGGAGUUCCGCAGUAUGGAUGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	412
HBVg_23	GCUGUGCCUUGGGUGGCUUUGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	413
HBVg_24	ACCCCUGCUCGUGUUACAGGGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	414
HBVg_25	CGGAAGUGUUGAUAAGAUAGGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	415
HBVg_26	CCUGCUGGUGGCUCCAGUUCGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	416
HBVg_27	CGAGGGAGUUCUUCUUCUAGGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	417
HBVg_28	GGGGCGCACCUCUCUUUACGGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	418
HBVg_29	AGCUUGGAGGCUUGAACAGUGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	419
HBVg_30	AAGCCUCCAAGCUGUGCCUUGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	420
HBVg_31	CCCCUGCUCGUGUUACAGGCGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	421

[0269]

HBVg_32	GCGAGGGAGUUCUUCUUCUAGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	422
HBVg_33	UACUAGUGCCAUUUGUUCAGGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	423
HBVg_34	GACUUCUCUCAAUUUUCUAGGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	424
HBVg_35	AUUGACCCGUAUAAAGAAUUGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	425
HBVg_36	ACCCAAAGACAAAAGAAAAUGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	426
HBVg_37	GUCCUCUUAUGUAAGACCUUGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	427
HBVg_38	UGAUCGGGAAAGAAUCCCAGGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	428
HBVg_39	UUUGCUGACGCAACCCCCACGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	429
HBVg_40	AUGAAUCUAGCCACCUGGGUGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	430
HBVg_41	GGUCUCCAUGCGACGUGCAGGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	431
HBVg_42	GGACUGAGGCCCACUCCCAUGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	432
HBVg_43	CACAGAGUCUAGACUCGUGGGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	433
HBVg_44	GAAGAACCAACAAGAAGAUGGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	434
HBVg_45	ACACGGUCCGGCAGAUGAGAGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	435
HBVg_46	GACAUGAACAUGAGAUGAUUGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	436
HBVg_47	GCAGCACAGCCUAGCAGCCAGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	437
HBVg_48	UCCUGGAAUUAGAGGACAAAGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	438
HBVg_49	GUCUUACAUAAGAGGACUCUGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	439
HBVg_50	UUGUGGGUCACCAUAUUCUUGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	440

[0270]

HBVg_51	CGCAAAAUACCUAUGGGAGUGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	441
HBVg_52	GGGUUGCGUCAGCAAACACUGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	442
HBVg_53	AGCUCUUGUUCCCAAGAAUAGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	443
HBVg_54	UGACAUACUUUCCAAUCAAUGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	444
HBVg_55	CAGAUGAGAAGGCACAGACGGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	445
HBVg_56	CCCCGCCUGUAACACGAGCAGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	446
HBVg_57	GGGUGGAGCCCUCAGGCUCAGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	447
HBVg_58	AUUCCUUGGACUCAUAAGGUGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	448
HBVg_59	UUUGUGGGUCACCAUAUUCUGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	449
HBVg_60	GUGAAAAAGUUGCAUGGUGCGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	450
HBVg_61	CCUGAACUGGAGCCACCAGCGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	451
HBVg_62	UCCUCUGCCGAUCCAUACUGGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	452
HBVg_63	CGGCUAGGAGUUCCGCAGUAGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	453
HBVg_64	AAUGUCAACGACCGACCUUGGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	454
HBVg_65	GACCUUCGUCUGCGAGGCGAGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	455
HBVg_66	GUUGCCGGGCAACGGGGUAAGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	456
HBVg_67	GAUUGAGACCUUCGUCUGCGGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	457
HBVg_68	AGGACCCCUGCUCGUGUUACGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	458
HBVg_69	UUUGAAGUAUGCCUCAAGGUGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	459

[0271]

HBVg_70	CCGCUUGUUUUGCUCGCAGCGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	460
HBVg_71	UGCUAGGCUGUGCUGCCAACGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	461
HBVg_72	UGCCGAUUGGUGGAGGCAGGGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	462
HBVg_73	UCUUUGUACUAGGAGGCUGUGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	463
HBVg_74	CGUCCCGCGCAGGAUCCAGUGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	464
HBVg_75	AAAGCCCAAGAUGAUGGGAUGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	465
HBVg_76	GCAGAUGAGAAGGCACAGACGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	466
HBVg_77	CGAUUGGUGGAGGCAGGAGGGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	467
HBVg_78	AGGAGGCUGUAGGCAUAAAUGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	468
HBVg_79	CCAUGCCCCAAAGCCACCCAGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	469
HBVg_80	AGGUUGGGGACUGCGAAUUUGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	470
HBVg_81	AGACCUUCGUCUGCGAGGCGGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	471
HBVg_82	CCUGGAAUUAGAGGACAAACGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	472
HBVg_83	UUUCAGUUAUAUGGAUGAUGGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	473
HBVg_84	GUAACACGAGCAGGGGUCCUGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	474
HBVg_85	CAUCUUCUUGUUGGUUCUUCGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	475
HBVg_86	CGGGGAGACCGCGUAAAGAGGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	476
HBVg_87	CUAGACUCUGUGGUAUUGUGGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	477
HBVg_88	CCCUGCUCGUGUUACAGGCGGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	478

[0272]

HBVg_89	UACCACAGAGUCUAGACUCGGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	479
HBVg_90	UCGCAAAAUACCUAUGGGAGGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	480
HBVg_91	GUCUGUGCCUUCUCAUCUGCGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	481
HBVg_92	ACACGUAGCGCCUCAUUUUGGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	482
HBVg_93	UUGGGGUUGAGGUCCCAAUCGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	483
HBVg_94	CCCCGAGACGGGUCGUCCGCGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	484
HBVg_95	CCUACGAACCACUGAACAAAGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	485
HBVg_96	UUACAUACUCUGUGGAAGGCGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	486
HBVg_97	ACCUCCUUUCCAUGGCUGCUGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	487
HBVg_98	GUUAUCGCUGGAUGUGUCUGGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	488
HBVg_99	AACAUGAGAUGAUUAGGCAGGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	489
HBVg_100	ACUUCUCUCAAUUUUCUAGGGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	490
HBVg_101	CACUUUCUCGCCAACUUACAGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	491
HBVg_102	CAUAAGGUGGGAAACUUUACGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	492
HBVg_103	CCAAACCUCGAAAAGGCAUGGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	493
HBVg_104	AUAGAAGGAAAGAAGUCAGAGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	494
HBVg_105	GCUGCUCCUUUUACACAAUGGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	495
HBVg_106	GAAGCGAAGUGCACACGGUCGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	496
HBVg_107	GGAUCAUCAACCACCAGCACGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	497

[0273]

HBVg_108	GAGCCAAGAGAAACGGACUGGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	498
HBVg_109	CUUCACCUCUGCACGUCGCAGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	499
HBVg_110	ACAAUGUUCCGGAGACUCUAGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	500
HBVg_111	UCCGCGGGAUUCAGCGCCGAGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	501
HBVg_112	UUAAUGAGUGGGAGGAGUUGGUUUAAGAGCUAUGCUGGAAACAGCAUAG CAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAG UCGGUGC	502
HBVg_113	CCAACUCAAACAAUCCAGAUGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	503
HBVg_114	GCUGCCAACUGGAUCCUGCGGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	504
HBVg_115	AGUCUUUGAAGUAUGCCUCAGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	505
HBVg_116	UCCUGACUGCCGAUUGGUGGGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	506
HBVg_117	UCUUGUCCUCCAAUUUGUCCGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	507
HBVg_118	CGAUAACCAGGACAAAUUGGGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	508
HBVg_119	AUUUGGAAGAUCCAGCAUCCGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	509
HBVg_120	CUGUUUGGCUUUCAGUUAUAGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	510
HBVg_121	CUCCUCCUGCCUCCACCAAUGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	511
HBVg_122	GUCAUCCUCAGGCCAUGCAGGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	512
HBVg_123	CUGCCGUUCCGGCCGACCACGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	513
HBVg_124	CCUUCCUGACUGCCGAUUGGGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	514
HBVg_125	ACCUGCACGACUCCUGCUCAGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	515
HBVg_126	GGCCUGUAUUUUCCUGCUGGGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	516

[0274]

HBVg_127	AACAUAGAGGUUCCUUGAGCGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	517
HBVg_128	UGCCGUUCCGGCCGACCACGGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	518
HBVg_129	AUAGGCCAUCAGCGCAUGCGGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	519
HBVg_130	GCCUCCACCAAUCGGCAGUCGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	520
HBVg_131	GUCCUUUGUUUACGUCCCGUGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	521
HBVg_132	UGGGAACAAGAGCUACAGCAGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	522
HBVg_133	CUGUAAACAGGCCUAUUGAUGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	523
HBVg_134	GUCGCAGAAGAUCUCAAUCUGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	524
HBVg_135	CUGCCUUCCUGACUGCCGAUGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	525
HBVg_136	ACUACUAAUUCCCUGGAUGCGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	526
HBVg_137	CACAUUUCUUGCCUUACUUUGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	527
HBVg_138	GGAUGACUGUCUCUUAGAGGGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	528
HBVg_139	GCUAUGCCUCAUCUUCUUGUGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	529
HBVg_140	CCCGUCGGCGCUGAAUCCCGGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	530
HBVg_141	UAGUAUUCCUUGGACUCAUAGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	531
HBVg_142	AGGUAGGAGCGGGAGCAUUCGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	532
HBVg_143	UCUUUUGGGGUGGAGCCCUCGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	533
HBVg_144	UCAGUAUGCCCUGAGCCUGAGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	534
HBVg_145	UUUAAUGAGUGGGAGGAGUUGUUUAAGAGCUAUGCUGGAAACAGCAUAG CAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAG UCGGUGC	535

[0275]

HBVg_146	CUCCCUCGCCUCGCAGACGAGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	536
HBVg_147	GAUAAGAUAGGGGCAUUUGGGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	537
HBVg_148	CCGCGGGAUUCAGCGCCGACGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	538
HBVg_149	CAGCGAUAACCAGGACAAAUGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	539
HBVg_150	CAGGUAGGAGUGGGAGCAUUGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	540
HBVg_151	GUUUAAUGAGUGGGAGGAGUGUUUAAGAGCUAUGCUGGAAACAGCAUAG CAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAG UCGGUGC	541
HBVg_152	UGGUGAGUGAUUGGAGGUUGGUUUAAGAGCUAUGCUGGAAACAGCAUAG CAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAG UCGGUGC	542
HBVg_153	UAAUGAGUGGGAGGAGUUGGGUUUAAGAGCUAUGCUGGAAACAGCAUAG CAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAG UCGGUGC	543
HBVg_154	ACUACAUGUUCUGGAUAAUAGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	544
HBVg_155	GGCAUAGCAGCAGGAUGAAGGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	545
HBVg_156	GUUGAUAAGAUAGGGGCAUUGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	546
HBVg_157	UCAACGAAUUGUGGGUCUUUGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	547
HBVg_158	UAUGGAUGAUGUGGUAUUGGGUUUAAGAGCUAUGCUGGAAACAGCAUAG CAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAG UCGGUGC	548
HBVg_159	CAACGAAUUGUGGGUCUUUUGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	549
HBVg_160	CAUUUGUUCAGUGGUUCGUAGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	550
HBVg_161	CGUCUAACAACAGUAGUUUCGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	551
HBVg_162	UGCCUGAGUGCUGUAUGGUGGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	552
HBVg_163	GCCCCGAGACGGGUCGUCCGGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	553
HBVg_164	GACUGCCGAUUGGUGGAGGCGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	554

[0276]

HBVg_165	UAUAUGGAUGAUGUGGUAUUGUUUAAGAGCUAUGCUGGAAACAGCAUAG CAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAG UCGGUGC	555
HBVg_166	CUUGAGUAUUUGGUGUCUUUGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	556
HBVg_167	GAUCUGGUGGGCGUUCACGGGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	557
HBVg_168	AGACUGGGAGGAGUUGGGGGGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	558
HBVg_169	AGUCCUCUUAUGUAAGACCUGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	559
HBVg_170	CUCAAGAUGUUGUACAGACUGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	560
HBVg_171	GGGAACAAGAGCUACAGCAUGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	561
HBVg_172	UCGCAGAAGAUCUCAAUCUCGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	562
HBVg_173	GGGGUGGAGCCCUCAGGCUCGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	563
HBVg_174	UAUUCCUUGGACUCAUAAGGGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	564
HBVg_175	UCUAAGAGACAGUCAUCCUCGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	565
HBVg_176	CAACUCAAACAAUCCAGAUUGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	566
HBVg_177	AAACAAAGGACGUCCCGCGCGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	567
HBVg_178	CUGCCAACUGGAUCCUGCGCGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	568
HBVg_179	GAAGCUCCAAAUUCUUUAUAGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	569
HBVg_180	GUCAGUAUGCCCUGAGCCUGGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	570
HBVg_181	UUUCCCACCUUAUGAGUCCAGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	571
HBVg_182	ACAAGAGGUUGGUGAGUGAUGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	572
HBVg_183	GAAAGCCCAAGAUGAUGGGAGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	573

[0277]

HBVg_184	AGGUUCCACGCAUGCGCUGAGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	574
HBVg_185	UUGGUGAGUGAUUGGAGGUUGUUUAAGAGCUAUGCUGGAAACAGCAUAG CAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAG UCGGUGC	575
HBVg_186	AGAGCUACAGCAUGGGAGGUGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	576
HBVg_187	AUAUGGAUGAUGUGGUAUUGGUUUAAGAGCUAUGCUGGAAACAGCAUAG CAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAG UCGGUGC	577
HBVg_188	CCAUUUGUUCAGUGGUUCGUGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	578
HBVg_189	UUAUAUGGAUGAUGUGGUAUGUUUAAGAGCUAUGCUGGAAACAGCAUAG CAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAG UCGGUGC	579
HBVg_190	GGAGUGGGAGCAUUCGGGCCGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	580
HBVg_191	AUUUGGUGUCUUUUGGAGUGGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	581
HBVg_192	CACAGAAAGGCCUUGUAAGUGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	582
HBVg_193	ACCAAUUUUCUUUUGUCUUUGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	583
HBVg_194	AGGUUAAUGGUCUUUGUACUGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	584
HBVg_195	UACCAAUUUUCUUUUGUCUUGUUUAAGAGCUAUGCUGGAAACAGCAUAGC AAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGC	585

[0278]

[0279] [0281] 일부 구현예에서, 본원에 제공된 gRNA는 B형 간염 바이러스 게놈의 표적 부위를 표적으로 삼는다. 일부 구현예에서, gRNA는 HBV 게놈의 0-3300bp 사이에 위치한 부위를 표적으로 삼는다. 일부 구현예에서, gRNA는 B형 간염 바이러스 게놈(B형 간염 바이러스 아형 ayw, 완전한 게놈, GenBank: U95551.1), 서열번호 650을 참조로 하는 위치에 상응하는, HBV 게놈의 43bp - 490bp, 1033bp - 1749bp, 1800bp - 1950bp, 또는 2953bp - 3182 사이에 위치한 부위를 표적으로 삼는다. 일부 구현예에서, gRNA는 B형 간염 바이러스 게놈(B형 간염 바이러스 아형 ayw, 완전한 게놈, GenBank: U95551.1), 서열번호 650을 참조로 하는 위치에 상응하는, HBV 게놈의 1bp - 42bp, 491bp - 1032bp, 1750bp - 1799bp, 또는 1951bp - 2952bp 사이에 위치한 부위를 표적으로 삼는다. 일부 구현예에서, gRNA는 HBV 복제 및/또는 전사에 관여하는 조절 요소 또는 그 근처의 부위를 표적으로 한다. 일부 구현예에서, gRNA는 폴리머라제 유전자, S-패밀리 유전자, X-유전자 또는 코어 패밀리 유전자를 표적으로 한다. 일부 구현예에서, gRNA는 M/S-HBs, X, 기저 코어, L-HBs 프로모터 영역을 표적으로 한다. 일부 구현예에서, gRNA는 Enh1 또는 Enh2 인핸서 영역을 표적으로 한다. 일부 구현예에서, gRNA는 HBV 코딩 영역을 표적으로 한다.

[0280] [0282] 일부 구현예에서, 본 명세서에서 제공되는 gRNA는 서열번호 196-229 중 어느 하나로부터 선택된 서열, 적어도 14개 뉴클레오타이드(예컨대 14, 15, 16, 17, 18 또는 19개 뉴클레오타이드)의 이의 연속 부분, 상기 중 어느 하나의 상보적 서열, 또는 상기 중 임의의 것과 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 이상의 서열 동일성을 갖는 서열을 포함한다. 일부 구현예에서, gRNA는 서열번호 230-295 중 어느 하나로부터 선택된 서열을 포함하는 스페이서 서열, 적어도 14nt(예컨대 14, 15, 16, 17, 18 또는 19개 뉴클레오타이드)의 이의 연속 부분, 또는 앞서 언급된 서열 중 하나와 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 이상의 서열 동일성을 갖는 서열을 포함한다. 일부 구현예에서, gRNA는 서열번호 296-390 중 어느 하나로부터 선택된 서열을 포함하는 스페이서 서열, 적어도 14nt(예컨대 14, 15, 16, 17, 18 또는 19개 뉴클레오타이드)의 이의 연속 부분, 또는 앞서 언급한 것과 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9% 또는 100% 이상의 서열 동일성을 갖는 서열을 포함한다. 일부 구현예에서, gRNA는 추가로 스캐폴드 서열을 포함한다. 일부 구현예에서, 스캐폴드 서열은 서열번호 587에 제시된 서열, 또는 서열번호 587과 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 이상의 서열 동일성을 갖는 서열을 포함한다. 일부 구현예에서, 스페이서 서열 및 스캐폴드 서열을 포함하는 gRNA는 서열번호 391-585 중 어느 하나에서 선택된 서열, 또는 그 전체

또는 일부와 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 이상의 서열 동일성을 갖는 서열을 포함한다.

[0281] [0283] 일부 구현예에서, gRNA는 서열번호 370, 333, 387, 347, 313, 320, 380, 256, 258, 311, 319, 230, 272 중 어느 하나로부터 선택된 서열을 포함하는 스페이서 서열, 적어도 14개 뉴클레오타이드(예컨대 14, 15, 16, 17, 18 또는 19개 뉴클레오타이드)의 이의 연속 부분, 상기 중 어느 하나의 상보적 서열 또는 상기 중 임의의 것과 적어도 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9% 또는 100%의 서열 동일성을 갖는 서열을 포함한다. 일부 구현예에서, gRNA는 추가로 스캐폴드 서열을 포함한다. 일부 구현예에서, 스캐폴드 서열은 서열번호 587에 제시된 서열, 또는 서열번호 587과 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 이상의 서열 동일성을 갖는 서열을 포함한다. 일부 구현예에서, 스페이서 서열 및 스캐폴드 서열을 포함하는 gRNA는 서열번호 565, 528, 542, 508, 515, 575, 515, 453, 506, 514, 425 또는 472 중 어느 하나로부터 선택된 서열, 또는 그 전체 또는 일부와 적어도 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9% 또는 100%의 서열 동일성을 갖는 서열을 포함한다. 일부 구현예에서, gRNA는 서열번호 565, 528, 542, 508, 515, 575, 515, 453, 506, 514, 425 또는 472에 제시된다.

[0282] [0284] 일부 구현예에서, gRNA는 서열번호 200, 205, 207, 213, 217, 221, 224, 233, 251, 256, 257, 258, 263, 267, 274, 275, 277, 279, 293, 294, 311, 313, 316, 319, 320, 330, 333, 338, 345, 347, 353, 359, 370, 371, 377, 380, 384, 385, 387 중 어느 하나로부터 선택된 서열을 포함하는 스페이서 서열, 적어도 14개 뉴클레오타이드(예컨대 14, 15, 16, 17, 18 또는 19개의 뉴클레오타이드)의 이의 연속 부분, 상기 중 어느 하나의 상보적 서열, 또는 상기 중 임의의 것과 적어도 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100%의 서열 동일성을 갖는 서열을 포함한다. 일부 구현예에서, gRNA는 추가로 스캐폴드 서열을 포함한다. 일부 구현예에서, 스캐폴드 서열은 서열번호 587에 제시된 서열, 또는 서열번호 587과 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 이상의 서열 동일성을 갖는 서열을 포함한다. 일부 구현예에서, 스페이서 서열 및 스캐폴드 서열을 포함하는 gRNA는 서열번호 395, 400, 402, 408, 412, 416, 419, 428, 446, 451, 452, 453, 458, 462, 465, 469, 470, 472, 474, 488, 489, 506, 508, 511, 514, 515, 525, 528, 533, 540, 542, 548, 554, 565, 566, 572, 575, 579, 580, 582 중 어느 하나에서 선택된 서열 또는 이의 전부 또는 일부와 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100%의 서열 동일성을 갖는 서열을 포함한다. 일부 구현예에서, gRNA는 서열번호 200, 201, 207, 217, 221, 224, 233, 237, 238, 246, 251, 256, 258, 263, 267, 274, 270, 277, 279, 283, 284, 293, 294, 308, 311, 313, 316, 319, 320, 325, 328, 330, 333, 338, 345, 347, 350, 353, 359, 360, 370, 371, 377, 380, 384, 385, 387 중 어느 하나로부터 선택된 서열을 포함하는 스페이서 서열, 적어도 14개 뉴클레오타이드(예컨대 14, 15, 16, 17, 18 또는 19개 뉴클레오타이드)의 이의 연속 부분, 상기 중 어느 하나의 상보적 서열, 또는 상기 중 임의의 것과 적어도 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9% 또는 100% 서열 동일성을 갖는 서열을 포함한다. 일부 구현예에서, gRNA는 추가로 스캐폴드 서열을 포함한다. 일부 구현예에서, 스캐폴드 서열은 서열번호 587에 제시된 서열, 또는 서열번호 587과 적어도 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 서열 동일성을 갖는 서열을 포함한다. 일부 구현예에서, 스페이서 서열 및 스캐폴드 서열을 포함하는 gRNA는 서열번호 369, 395, 402, 408, 412, 416, 419, 428, 432, 433, 441, 446, 451, 453, 458, 462, 465, 469, 472, 474, 478, 479, 488, 489, 503, 506, 508, 511, 514, 515, 520, 523, 525, 575, 528, 533, 540, 542, 545, 548, 554, 555, 565, 566, 572, 579, 580, 582 중 하나에서 선택된 서열, 또는 그 전체 또는 일부와 적어도 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100%의 서열 동일성을 갖는 서열을 포함한다. 일부 구현예에서, gRNA는 서열번호 369, 395, 402, 408, 412, 416, 419, 428, 432, 433, 441, 446, 451, 453, 458, 462, 465, 469, 472, 474, 478, 479, 488, 489, 503, 506, 508, 511, 514, 515, 520, 523, 525, 575, 528, 533, 540, 542, 545, 548, 554, 555, 565, 566, 572, 579, 580 또는 582에 제시된다.

[0283] [0285] 일부 구현예에서, gRNA는 서열번호 207, 213, 215, 217, 221, 222, 241, 245, 258, 261, 268, 274, 380, 387 중 어느 하나로부터 선택된 서열을 포함하는 스페이서 서열, 적어도 14개 뉴클레오타이드(예컨대 14, 15, 16, 17, 18 또는 19개 뉴클레오타이드)의 이의 연속 부분, 또는 상기 중 임의의 것과의 상보적 서열 또는 상기 중 임의의 것과 적어도 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9% 또는 100%의 서열 동일성을 갖는 서열을 포함한다. 일부 구현예에서, gRNA는 서열 서열번호 217을 포함하는 스페이서 서열, 적어도 14개 뉴클레오타이드(예컨대 14, 15, 16, 17, 18 또는 19개 뉴클레오타이드)의 이의 연속 부분, 상기 중 임의의 것과의 상보적 서열, 또는 서열번호 217과 적어도 80%, 85%, 90%, 91%, 92%, 93%, 94%,

95%, 96%, 97%, 98%, 99%, 99.5%, 99.9% 또는 100%의 서열 동일성을 갖는 서열을 포함한다. 일부 구현예에서, gRNA는 추가로 스캐폴드 서열을 포함한다. 일부 구현예에서, 스캐폴드 서열은 서열번호 587에 제시된 서열, 또는 서열번호 587과 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 또는 100% 이상의 서열 동일성을 갖는 서열을 포함한다. 일부 구현예에서, 스페이서 서열 및 스캐폴드 서열을 포함하는 gRNA는 서열번호 402, 408, 410, 412, 416, 417, 436, 440, 453, 456, 463, 469, 575, 582 중 하나에서 선택된 서열 또는 그의 전체 또는 일부와 적어도 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9% 또는 100%의 서열 동일성을 갖는 서열을 포함한다. 일부 구현예에서, 적어도 두 개의 유전자 및/또는 그 조절 요소의 후성적 변형을 위한 제공된 다중화된 후성적 변형 DNA-표적화 시스템은 Cas9 단백질과 같은 Cas 단백질과 복합체를 이룬 상기 언급된 gRNA를 포함한다. 일부 구현예에서, Cas9는 dCas9이다. 일부 구현예에서, dCas9는 서열번호 599에 제시된 dSpCas9와 같은 dSpCas9 또는 이의 변이체 및/또는 융합체이다.

[0284] [0286] 일부 구현예에서, 본 발명에 따라 gRNA의 조합이 제공된다. 일부 구현예에서, 본 발명에 따라 gRNA의 조합을 포함하는 다중화된 후성적 변형 DNA-표적화 시스템이 제공된다.

[0285] [0287] 일부 구현예에서, gRNA의 조합은 적어도 두 개의 다른 유전자 또는 이의 조절 요소를 표적으로 하는 적어도 두 개의 gRNA를 포함한다. 일부 구현예에서, gRNA의 조합은 제1 유전자 또는 이의 조절 요소를 표적으로 하는 제1 gRNA와 제2 유전자 또는 이의 조절 요소를 표적으로 하는 제2 gRNA를 포함한다. 일부 구현예에서, 제1 gRNA는 폴리머라제 유전자, S-패밀리 유전자, X-유전자, 코어 패밀리 유전자, pre-S1 프로모터, pre-S2 프로모터, X 프로모터, 기저 코어 프로모터, Enh1 인핸서, Enh2 인핸서, 전사 처리 제어 영역 및 HBV 게놈 내의 모든 코딩 영역으로 구성된 목록에서 선택된 유전자 또는 이의 조절 요소를 표적으로 하고, 제2 gRNA는 폴리머라제 유전자, S-패밀리 유전자, X-유전자, 코어 패밀리 유전자, pre-S1 프로모터, pre-S2 프로모터, X 프로모터, 기저 코어 프로모터, Enh1 인핸서, Enh2 인핸서, 전사 처리 제어 영역 및 HBV 게놈 내의 모든 코딩 영역으로 구성된 목록에서 선택된 유전자 또는 이의 조절 요소를 표적으로 하며, 제1 및 제2 gRNA는 서로 다른 유전자 또는 이의 조절 요소를 표적으로 한다. 일부 구현예에서, 제1 gRNA는 Enh1 인핸서를 표적으로 하고, 제2 gRNA는 L-HBs 프로모터, M-HBs 프로모터, S-HBs 프로모터, X-프로모터, 기저 코어 프로모터, S-유전자 프로모터 및 Enh2 인핸서로 구성된 목록에서 선택된 유전자를 표적으로 한다. 일부 구현예에서, 제1 gRNA는 Enh1 인핸서를 표적으로 하고, 제2 gRNA는 L-HBs를 표적으로 한다. 일부 구현예에서, 제1 gRNA 및 제2 gRNA는 표 1에 나열된 유전자 조합에서 선택된 두 유전자의 조합을 표적으로 한다. 일부 구현예에서, 제1 gRNA 및 제2 gRNA는 각각 독립적으로 본원에서 설명된 gRNA 중 하나에서 선택된다.

[0286] [0288] 일부 구현예에서, gRNA의 조합은 적어도 3개의 다른 유전자 또는 이의 조절 요소를 표적으로 하는 적어도 3개의 gRNA를 포함한다. 일부 구현예에서, gRNA의 조합은 제1 유전자를 표적으로 하는 제1 gRNA, 제2 유전자를 표적으로 하는 제2 gRNA, 제3 유전자를 표적으로 하는 제3 gRNA를 포함한다. 일부 구현예에서, gRNA의 조합은 적어도 3개의 다른 유전자 또는 이의 조절 요소를 표적으로 하는 적어도 3개의 gRNA를 포함한다. 일부 구현예에서, gRNA의 조합은 제1 유전자 또는 이의 조절 요소를 표적으로 하는 제1 gRNA, 제2 유전자 또는 이의 조절 요소를 표적으로 하는 제2 gRNA, 제3 유전자 또는 이의 조절 요소를 표적으로 하는 제3 gRNA를 포함한다. 일부 구현예에서,

[0287] 제1 gRNA는 폴리머라제 유전자, S-패밀리 유전자, X-유전자, 코어 패밀리 유전자, pre-S1 프로모터, pre-S2 프로모터, X 프로모터, 기저 코어 프로모터, Enh1 인핸서, Enh2 인핸서, 전사체 처리 제어 영역 및 HBV 게놈 내의 모든 코딩 영역으로 구성된 목록에서 선택된 유전자 또는 이의 조절 요소를 표적으로 하고, 제2 gRNA는 폴리머라제 유전자, S-패밀리 유전자, X-유전자, 코어 패밀리 유전자, pre-S1 프로모터, pre-S2 프로모터, X 프로모터, 기저 코어 프로모터, Enh1 인핸서, Enh2 인핸서, 전사체 처리 제어 영역 및 HBV 게놈 내의 모든 코딩 영역으로 구성된 목록에서 선택된 유전자 또는 이의 조절 요소를 표적으로 하며, 제3 gRNA는 폴리머라제 유전자, S-패밀리 유전자, X-유전자, 코어 패밀리 유전자, pre-S1 프로모터, pre-S2 프로모터, X 프로모터, 기저 코어 프로모터, Enh1 인핸서, Enh2 인핸서, 전사 처리 제어 영역 및 HBV 게놈 내의 모든 코딩 영역으로 구성된 목록에서 선택된 유전자 또는 이의 조절 요소를 표적으로 하고, 제1 제2 및 제3 gRNA는 서로 다른 유전자 또는 이의 조절 요소를 표적으로 한다. 일부 구현예에서, 제1 gRNA는 Enh1 인핸서를 표적으로 하고, 제2 gRNA는 L-HBs 프로모터, M-HBs 프로모터, S-HBs 프로모터, X-프로모터, 기저 코어 프로모터, S-유전자 프로모터, Enh1 인핸서 및 Enh2 인핸서로 구성된 목록에서 선택된 유전자를 표적으로 하며, 제3 gRNA는 L-HBs 프로모터, M-HBs 프로모터, S-HBs 프로모터, X-프로모터, 기저 코어 프로모터, S-유전자 프로모터, Enh1 인핸서 및 Enh2 인핸서로 구성된 목록에서 선택된 유전자를 표적으로 하고, 제2 gRNA와 제3 gRNA는 서로 다른 유전자를 표적으로 한다. 일부 구현예에서, 제1 gRNA는 Enh1 인핸서를 표적으로 하고, 제2 gRNA는 L-HBs 프로모터를 표적으로

하며, 제3 gRNA는 M-HBs 프로모터, S-HBs 프로모터, X-프로모터, 기저 코어 프로모터, S-유전자 프로모터 및 Enh2 인핸서로 구성된 목록에서 선택된 유전자를 표적으로 한다. 일부 구현예에서, 제1 gRNA, 제2 gRNA 및 제3 gRNA는 표 2에 나열된 유전자 조합에서 선택된 세 유전자의 조합을 표적으로 한다. 일부 구현예에서, 제1 gRNA, 제2 gRNA 및 제3 gRNA는 각각 독립적으로 본원에서 설명된 gRNA 중 어느 하나로부터 선택된다.

[0288] [0289] 일부 구현예에서, gRNA의 조합은 적어도 4개의 다른 유전자 또는 이의 조절 요소를 표적으로 하는 적어도 4개의 gRNA를 포함한다. 일부 구현예에서, gRNA의 조합은 제1 유전자를 표적으로 하는 제1 gRNA, 제2 유전자를 표적으로 하는 제2 gRNA, 제3 유전자를 표적으로 하는 제3 gRNA, 제4 유전자를 표적으로 하는 제4 gRNA를 포함한다. 일부 구현예에서, gRNA의 조합은 적어도 4개의 다른 유전자 또는 이의 조절 요소를 표적으로 하는 적어도 4개의 gRNA를 포함한다. 일부 구현예에서, gRNA의 조합은 제1 유전자 또는 이의 조절 요소를 표적으로 하는 제1 gRNA, 제2 유전자 또는 이의 조절 요소를 표적으로 하는 제2 gRNA, 제3 유전자 또는 이의 조절 요소를 표적으로 하는 제3 gRNA, 제3 유전자 또는 이의 조절 요소를 표적으로 하는 제4 gRNA를 포함한다. 일부 구현예에서, 제1 gRNA는 폴리머라제 유전자, S-패밀리 유전자, X-유전자, 코어 패밀리 유전자, pre-S1 프로모터, pre-S2 프로모터, X 프로모터, 기저 코어 프로모터, Enh1 인핸서, Enh2 인핸서, 전사체 처리 제어 영역 및 HBV 게놈 내의 모든 코딩 영역으로 구성된 목록에서 선택된 유전자 또는 이의 조절 요소를 표적으로 하고, 제2 gRNA는 폴리머라제 유전자, S-패밀리 유전자, X-유전자, 코어 패밀리 유전자, pre-S1 프로모터, pre-S2 프로모터, X 프로모터, 기저 코어 프로모터, Enh1 인핸서, Enh2 인핸서, 전사체 처리 제어 영역 및 HBV 게놈 내의 모든 코딩 영역으로 구성된 목록에서 선택된 유전자 또는 이의 조절 요소를 표적으로 하며, 제3 gRNA는 폴리머라제 유전자, S-패밀리 유전자, X-유전자, 코어 패밀리 유전자, pre-S1 프로모터, pre-S2 프로모터, X 프로모터, 기저 코어 프로모터, Enh1 인핸서, Enh2 인핸서, 전사체 처리 제어 영역 및 HBV 게놈 내의 모든 코딩 영역으로 구성된 목록에서 선택된 유전자 또는 이의 조절 요소를 표적으로 하고, 제4 gRNA는 폴리머라제 유전자, S-패밀리 유전자, X-유전자, 코어 패밀리 유전자, pre-S1 프로모터, pre-S2 프로모터, X 프로모터, 기저 코어 프로모터, Enh1 인핸서, Enh2 인핸서, 전사체 처리 제어 영역 및 HBV 게놈 내의 모든 코딩 영역으로 구성된 목록에서 선택된 유전자 또는 이의 조절 요소를 표적으로 하며, 제1, 제2, 제3 및 제4 gRNA는 서로 다른 유전자 또는 이의 조절 요소를 표적으로 한다. 일부 구현예에서, 제1 gRNA는 Enh1 인핸서를 표적으로 하고, 제2 gRNA는 L-HBs 프로모터, M-HBs 프로모터, S-HBs 프로모터, X-프로모터, 기저 코어 프로모터, S-유전자 프로모터, Enh1 인핸서 및 Enh2 인핸서로 구성된 목록에서 선택된 유전자를 표적으로 하며, 제3 gRNA는 L-HBs 프로모터, M-HBs 프로모터, S-HBs 프로모터, X-프로모터, 기저 코어 프로모터, S-유전자 프로모터, Enh1 인핸서 및 Enh2 인핸서로 구성된 목록에서 선택된 유전자를 표적으로 하고, 제4 gRNA는 L-HBs 프로모터, M-HBs 프로모터, S-HBs 프로모터, X-프로모터, 기저 코어 프로모터, S-유전자 프로모터, Enh1 인핸서 및 Enh2 인핸서로 구성된 목록에서 선택된 유전자를 표적으로 하며, 제1, 제2, 제3 및 제4 gRNA는 서로 다른 유전자 또는 이의 조절 요소를 표적으로 한다. 일부 구현예에서, 제1 gRNA, 제2 gRNA, 제3 gRNA 및 제4 gRNA는 표 2에 나열된 유전자 또는 조절 요소의 조합에서 선택된 네 가지 유전자 또는 조절 요소의 조합을 표적으로 한다. 일부 구현예에서, 제1 gRNA, 제2 gRNA, 제3 gRNA 및 제4 gRNA는 각각 여기에 설명된 gRNA 중 어느 하나로부터 독립적으로 선택된다.

[0289] [0290] 일부 구현예에서, gRNA의 조합은 적어도 5개의 다른 유전자 또는 이의 조절 요소를 표적으로 하는 적어도 5개의 gRNA를 포함한다. 일부 구현예에서, gRNA의 조합은 적어도 6개의 다른 유전자 및/또는 그 조절 요소를 표적으로 하는 적어도 6개의 gRNA를 포함한다. 일부 구현예에서, 제1, 제2, 제3, 제4, 제5, 및/또는 제6 유전자 또는 이의 조절 요소는 서로 상이하다.

[0290] **C. 조작된 아연 집게 단백질(eZFP)**

[0291] [0291] 일부 측면에서, 본 발명에서는 아연 집게 단백질(ZFP), 예를 들어, 조작된 아연 집게 단백질(eZFP)이 제공된다. 일부 구현예에서, eZFP는 HBV 유전자의 표적 부위 또는 B형 간염 바이러스 서열의 유전자의 조절 요소에 결합하거나 결합할 수 있다. 일부 측면에서, eZFP는 유전자 또는 조절 요소의 전사 억제를 위한 효과기 도메인의 특정 표적화를 용이하게 할 수 있다. 일부 구현예에서, 본 발명에 따라 eZFP와 하나 이상의 다른 요소(예컨대 전사 억제를 위한 효과기 도메인)를 포함하는 융합 단백질을 포함하는 후성 유전적 변형 DNA-표적화 시스템이 제공된다. 따라서, 일부 측면에서 eZFP는 예를 들어 HBV 바이러스 감염, 간 질환 또는 암과 같은 HBV와 관련된 질병 또는 장애를 치료하기 위한 조성물 및 방법과 관련하여 HBV 유전자 또는 조절 요소의 발현 감소를 용이하게 한다.

[0292] [0292] 일부 구현예에서, 아연 집게 단백질(ZFP), 아연 집게 DNA 결합 단백질, 또는 아연 집게 DNA 결합 도메인은 구조가 아연 이온의 배위를 통해 안정화되는 결합 도메인 내의 아미노산 서열의 영역인, 하나 이상의 아연 집게를 통해 서열 특이적 방식으로 DNA를 결합시키는, 단백질, 또는 더 큰 단백질 내의 도메인이다. ZFP 중에는

인공 또는 조작된 ZFP(eZFP)가 있으며, 이는 일반적으로 개별 아연 집게의 조립에 의해 생성된 9-18개 뉴클레오타이드 길이의 특정 DNA 서열을 표적으로 하는 ZFP 도메인을 포함한다. ZFP에는 단일 집게 도메인의 길이가 약 30개 아미노산이고 단일 베타 회전의 2개의 시스테인과 아연을 통해 배위된 2개의 불변 히스티딘 잔기를 포함하는 알파 나선을 포함하고 두 개, 세 개, 네 개, 다섯 개 또는 여섯 개의 집게를 갖는 알파 나선을 함유하는 것이 포함된다. 일반적으로 ZFP의 서열 특이성은 아연 집게 인식 나선(아연 집게 인식 영역이라고도 함)의 네 개의 나선 위치(-1, 2, 3 및 6)에서 아미노산 치환을 함으로써 변경될 수 있다. 따라서 예를 들어, 융합 단백질과 같은 ZFP 또는 ZFP를 포함하는 분자는 비천연적으로 발생할 수 있다. 예를 들어, 선택한 표적 부위에 결합하도록 조작된다.

[0293] [0293] 일부 구현예에서, 아연 집게는 맞춤 설계(즉, 사용자에 의해 설계)되거나 상업적 공급원에서 획득된다. 아연 집게 단백질을 설계하는 다양한 방법이 이용 가능하다. 예를 들어, 목적 표적 DNA 서열에 결합하는 아연 집게 단백질의 설계 방법은, 예를 들어 문헌[Liu, Q. et al., PNAS, 94(11):5525-30 (1997); Wright, D.A. et al., Nat. Protoc., 1(3):1637-52 (2006); Gersbach, C.A. et al., Acc. Chem. Res., 47(8):2309-18 (2014); Bhakta M.S. et al., Methods Mol. Biol., 649:3-30 (2010); and Gaj et al., Trends Biotechnol, 31(7):397-405 (2013)]에 기재되어 있다. 또한, 목적 DNA 표적 서열에 결합하는 아연 집게 단백질을 설계하기 위한 다양한 웹 기반 도구가 공개적으로 이용 가능하다. 예를 들어, scripps.edu/barbas/zfdesign/zfdesignhome.php의 월드 와이드 웹에서 이용 가능한 Scripps에서의 Zinc Finger Tools 설계 웹 사이트를 참조한다. 목적 DNA 표적 서열에 결합하도록 아연 집게 단백질을 설계하기 위한 다양한 상업적 서비스도 이용 가능하다. 예를 들어, Creative Biolabs(creative-biolabs.com/Design-and-Synthesis-of-Artificial-Zinc-Finger-Proteins.html의 월드 와이드 웹), Addgene(addgene.org/kits/zfc-modular-assembly/의 월드 와이드 웹)의 Zinc Finger Consortium Modular Assembly Kit, 또는 Sigma Aldrich(sigmaaldrich.com/life-science/zinc-finger-nuclease-technology/custom-zfn.html의 월드 와이드 웹)의 CompoZr Custom ZFN Service에 의해 제공되는 상업적으로 이용 가능한 서비스 또는 키트를 참조한다.

[0294] [0294] 일부 구현예에서, 본원에 제공된 것은 eZFP와 하나 이상의 다른 요소, 예를 들어 전사 억제를 위한 효과기 도메인을 포함하는 융합 단백질을 포함하는 후성적 변형 DNA-표적화 시스템이다. 일부 구현예에서, 후성적 변형 DNA-표적화 시스템의 적어도 하나의 DNA-결합 도메인은 조작된 아연 집게 단백질(eZFP)을 포함한다. 일부 구현예에서, 후성적 변형 DNA-표적화 시스템은 하나 이상의 HBV 유전자 또는 이의 조절 요소의 표적 부위에 결합하는 조작된 아연 집게 단백질(eZFP)을 포함한다. 제공된 eZFP 중 어느 하나가 표적화하는 표적 부위는 본 명세서에 설명된 바와 같은 임의의 것일 수 있으며, 예를 들어 섹션 IA에 설명된 임의의 표적 부위일 수 있다.

[0295] [0295] 일부 구현예에서, 본원에서 제공된 eZFP(예컨대 후성적 변형 DNA-표적화 시스템의 융합 단백질에 포함된 eZFP)에 대한 표적 부위는 하나 이상의 HBV 유전자 또는 이의 조절 요소에 있다. 일부 구현예에서, 표적 부위는 HBV 게놈의 CpG 섬(예컨대 CpG 섬 1, CpG 섬 2, CpG 섬 3)에 있다. 일부 구현예에서, 표적 부위는 HBV 게놈의 CpG 섬 2에 있다. 일부 구현예에서, 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조하는 위치에 상응하는 HBV 게놈의 1033bp-1749bp에 걸쳐 있는 표적 영역 내에 있다. 일부 구현예에서, 표적 부위는 HBx 시작 코돈 상류 300개 염기쌍(bp) 이내, 250bp 이내, 200bp 이내, 150bp 이내, 140bp 이내, 130bp 이내, 120bp 이내, 110bp 이내 또는 100bp 이내에 걸쳐 있는 표적 영역 내에 있다. 일부 구현예에서, 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로1250-1374 bp에 걸친 서열에 상응하는 표적 영역 서열 내에 있다. 일부 구현예에서, 표적 영역은 서열번호 1068에 제시된 서열을 갖는다. 일부 구현예에서, 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로1255-1302 bp에 걸친 서열에 상응하는 표적 영역 서열 내에 있다. 일부 구현예에서, 표적 영역은 서열번호 1069에 제시된 서열을 갖는다. 일부 구현예에서, 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 1260-1300 bp에 걸친 서열에 상응하는 표적 영역 서열 내에 있다. 일부 구현예에서, 표적 영역은 서열번호 1070에 제시된 서열을 갖는다.

[0296] [0296] 일부 구현예에서, 본원에서 제공되는 eZFP(예컨대 후성적 변형 DNA-표적 시스템의 융합 단백질에 포함된 eZFP)에 대한 표적 부위는 서열번호 1028-1055 중 어느 하나에 제시된 뉴클레오타이드 서열, 적어도 12nt의 이의 연속 부분, 또는 상기 중 어느 하나의 상보적 서열을 포함한다. 일부 구현예에서, 본원에서 제공되는 eZFP에 대한 표적 부위는 서열번호 1028-1055 중 어느 하나에 제시된 뉴클레오타이드 서열을 포함한다. 일부 구현예에서, 본원에서 제공되는 eZFP에 대한 표적 부위는 서열번호 1045, 1046 또는 1052 중 어느 하나에 제시된 뉴클레오타이드 서열, 적어도 12nt의 이의 연속 부분, 또는 상기 중 어느 하나의 상보적 서열을 포함한다. 일부 구현예에서, 표적 부위는 서열번호 1045, 1046 또는 1052 중 하나에 명시된 뉴클레오타이드 서열을 포함한다.

[0297] [0297] 일부 구현예에서, 표적 부위는 인간 게놈 DNA에 통합된 HBV 서열과 같은 이중 가닥 DNA에 포함된다. 일

부 구현예에서, 표적 부위는 공유적으로 폐쇄된 원형(cccDNA) HBV 서열에 포함된다. 일부 구현예에서, 표적 부위는 이완된 원형 DNA(rcNDA) HBV 서열에 포함된다. 일부 구현예에서, eZFP는 표적 부위에 결합할 수 있다. 일부 구현예에서, eZFP는 표적 부위에 결합한다. 일부 구현예에서, 결합은 표적 특이적이다. 예를 들어, 일부 구현예에서, eZFP는 표적 부위에 결합하고, 다른 서열을 포함하는 다른 부위에는 결합하지 않는다. 예를 들어, 일부 구현예에서, 본원에서 공개된 개별 eZFP는 서열번호 1028-1055 중 하나에 명시된 표적 부위에 결합하고, 다른 표적 부위에는 결합하지 않는다. 일부 구현예에서, 본원에 공개된 개별 eZFP는 서열번호 1045, 1046 또는 1052 중 하나에 명시된 표적 부위에 결합하고 다른 표적 부위에는 결합하지 않는다. 일부 구현예에서, 본원에 제공된 eZFP의 표적 부위는 표 7에 제시된 서열을 포함한다.

표 7. eZFP 표적 서열

eZFP 표적 부위 서열	SEQ ID NO:
CAAGTGTTTGCTGACGCA	1028
GGCTGGGGCTTGGTCATG	1029
GCTGGGGCTTGGTCATGG	1030
TTGGTCATGGGCCATCAG	1031
TGCGTGGAACCTTTTCGG	1032
GCAGCAGGTCTGGAGCAA	1033
CAGCAGGTCTGGAGCAAA	1034
AGCAGGTCTGGAGCAAAC	1035
ATCGTATCCATGGCTGCT	1036
CCAGTGGGGGTTGCGTCA	1037
CCCCAGCCAGTGGGGGTT	1038
GCGCTGATGGCCCATGAC	1039
CACGCACGCGCTGATGGC	1040
AGAGGAGCCGAAAAGGTT	1041
GGATCGGCAGAGGAGCCG	1042
CAGTATGGATCGGCAGAG	1043
GTTCCGCAGTATGGATCG	1044
GGAGTTCCGCAGTATGGA	1045
GCAAAACAAGCGGCTAGG	1046
TTTGCTCCAGACCTGCTG	1047
GATGTATATTTGCGGGAG	1048
CAGCCAGTGGGGGTTGCG	1049
GATGGCCCATGACCAAGC	1050
GGGGTAAAGGTTCAGGTA	1051
GCTAGGAGTTCCGCAGTA	1052
GCGGCTAGGAGTTCCGCA	1053
GTATGGATCGGCAGAGGA	1054
CCGATCCATACTGCGGAA	1055

[0298]

[0299] [0298] 일부 구현예에서, 본원에서 제공된 eZFP의 표적 부위(예컨대 후성적 변형 DNA-표적화 시스템의 융합 단백질에 포함된 eZFP)는 서열번호 1045에 제시된 뉴클레오타이드 서열을 포함한다. 일부 구현예에서, 본원에서 제공된 eZFP의 표적 부위는 적어도 12nt의 이의 연속된 부분 또는 상기 중 어느 하나의 상보적 서열. 일부 구현예에서, 본원에서 제공된 eZFP의 표적 부위는 서열번호 1045에 제시된 서열을 포함한다.

[0300] [0299] 일부 구현예에서, 본원에서 제공된 eZFP의 표적 부위(예컨대 후성적 변형 DNA-표적화 시스템의 융합 단백질에 포함된 eZFP)는 서열번호 1046에 제시된 뉴클레오타이드 서열을 포함한다. 일부 구현예에서, 본원에서 제공된 eZFP의 표적 부위는 적어도 12nt의 이의 연속된 부분 또는 상기 중 어느 하나의 상보적 서열. 일부 구현예에서, 본원에서 제공된 eZFP의 표적 부위는 서열번호 1046에 제시된 서열을 포함한다.

[0301] [0300] 일부 구현예에서, 본원에서 제공된 eZFP의 표적 부위(예컨대 후성적 변형 DNA-표적화 시스템의 융합 단백질에 포함된 eZFP)는 서열번호 1052에 제시된 뉴클레오타이드 서열을 포함한다. 일부 구현예에서, 본원에서 제공된 eZFP의 표적 부위는 적어도 12nt의 이의 연속된 부분 또는 상기 중 어느 하나의 상보적 서열. 일부 구현

예에서, 본원에서 제공된 eZFP의 표적 부위는 서열번호 1052에 제시된 서열을 포함한다.

[0302] [0301] 일부 구현예에서, eZFP는 복수의 아연 집게를 포함한다. 일부 구현예에서, 각 아연 집게는 인식 영역을 포함한다. 일부 구현예에서, 인식 영역들은 함께 eZFP의 서열 특이적 결합, 예를 들어 특정 표적 부위에 대한 결합을 용이하게 한다. 일부 구현예에서, eZFP는 N-말단에서 C-말단의 순서로 F1에서 F6으로 표시된 6개의 아연 집게를 포함하며, 이들 각각은 특정 표적 부위에 대한 서열 특이적 결합을 용이하게 하는 상응하는 인식 영역 F1 내지 F6을 포함한다.

[0303] [0302] 일부 구현예에서, 본원에 제공된 특이적 표적 부위를 표적화하는 ㄷZFP의 특징들을 표 E4에 나타내었다. 일부 구현예에서, eZFP는 N-말단에서 C-말단 순서로 F1에서 F6까지 표시된 6개의 아연 집게를 포함하며, 이들 각각은 표 E4에 나와 있는 바와 같이 상응하는 인식 영역 F1-F6을 포함한다. 일부 구현예에서, 인식 영역 F1-F6은 표 E4에 나와 있는 바와 같이 지정된 표적 부위 서열에 대한 특정 결합을 용이하게 한다. 일부 구현예에서, eZFP는 표 E4에 나와 있는 바와 같이 인식 영역을 포함하는 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 표 8에 나와 있는 바와 같이 DNA 서열에 의해 인코딩될 수 있다.

표 8. eZFP DNA 서열

SEQ IDNOs	서열	Description
888	GCTGCTATGGCAGAGCGGCCATTTCAGTGTCGAATTTGCATGCGAAATTTCTCTAGCGA AGCGGATCGAAGTAGACATATCCGGACCCATACGGGTGAGAAACCGTTCGCGTGTGAT ATATGCGGTCGCAAATTCGCAGACCGCTCCAACCTGACAAGGCACACAAAAATTCATA CTGGAAGCCAGAAACCGTTTCAGTGCCGGATTTGTATGCGCAATTTCTCTCAATCCTCC GATCTCTCCCGCCACATCAGGACTCATACCGGGGAGAAACCCTTTGCTTGTGACATTTG TGGCAGAAAGTTTGCCTATCACTGGTACCTTAAGAAGCACACCAAGATCCATACTGGCT CTCAAAAGCCTTTCCAATGCCGAATATGTATGCGAAACTTTTCTAGGTCTGACTCCCTC TCTGTACATATCCGCACTCACACGGGTGAAAAACCATTTGCCTGTGACATATGTGGCAG AAAATTTGCTCAAAATGCGAACCGAAAAACGCATACGAAAATCCATTTGCGACAAAAA GATGCAGCTCGG	ZFP DNA 서열
889	GCGGCAATGGCAGAACGCCCGTTTCAGTGTAGAATCTGCATGCGCAATTTCTCCAGAA GTGACGTACTCAGCACGCATATCAGAACACATACTGGTGAAAAACCGTTTGCTTGTGAT ATCTGTGGTAAGAAATTCGCGGATAACTCCTCAAGGACGCGGCATACGAAGATCCACA CTGGTTCTCAAAAGCCGTTCCAGTGCCGGATATGCATGCGGAACTTTAGTCGGCCATAT ACGCTTCGACTTCATATTAGGACTCACACCGGCGAAAAGCCGTTCGCCTGTGACATTTG CGGGAGAAAATTCGCTGACTCTAGTCACCGCACGAGGCACACTAAAATACATACTGGT TCACAAAAACCATTCCAGTGCCGAATTTGCATGCGAAATTTTTCCCGAAGTGACCATCT CTCACAGCACATCCGCACCCATACAGGGGAGAAGCCCTTCGCTTGTGATATATGCGGA CGCAAGTTCGCGGACAGCTCACACCGGACCCGCCATACAAAGATCCACTTGAGACAGA AAGATGCAGCGCGG	ZFP DNA 서열
890	GCCGCAATGGCCGAAAGACCATTTCAGTGCAGGATATGTATGCGCAACTTCTCTCGCA GTGACCACCTGAGTCAACATATCAGGACACACACGGGTGAAAAGCCTTTTGCATGCGA TATTTGTGGTCGAAAATTTGCTCAGTCTGCGGACCGAACCAAGCACACTAAAATTCATA CCGGCTCACAGAAACCGTTTCAATGCCGCATCTGTATGAGGAATTTCTCTAGATCAGAC CACTTGTCCCAACACATCCGGACTCATACTGGAGAAAAGCCGTTTGCATGTGACATTTG TGGCAGGAAGTTTGCTAGAAGGTCTGACCTTAAAAGGCACACAAAAATTCATACGGGT TCCCAGAAACCATTTCAGTGCCGGATATGCATGCGGAACTTTTCACGAAGCGACCACCT CTCCCGACATATTCGAACGCACACTGGTGAGAAGCCGTTTGCTTGCGACATTTGCGGAC GCAAGTTCGCTCAGAGCTCCGACTTGAGGAGGCATACCAAGATTCATCTCCGGCAGAA AGATGCCGCGCGG	ZFP DNA 서열
891	GCCGCGATGGCTGAGAGACCATTTCAGTGTCGAATCTGCATGAGAAATTTTTCAAGGA GTGACAATCTGTCTGAGCACATACGAACACATACTGGGGAAAAACCCTTTGCATGTGA CATTTGTGGAAGAAAGTTTGCTACCAGCTCAAATCGCAAAACACATACAAAGATACAT ACCGGCTCCCAAAAGCCATTCCAGTGCCGCATCTGCATGAGGAACTTTTCCGATCGCTC ACATCTTACCCGCCACATAAGAACTCACACAGGCGAAAAGCCCTTTGCCTGCGATATAT GCGGACGGAAGTTCGCCCGCTCCGACGCTTTGACCCAGCATACCAAAATCCATACTGG GTCTCAAAAGCCATTTCAGTGCCGAATCTGTATGAGGAATTTCTCCGACAGGTCAGCAT TGGCACGGCATATCCGCACCCATACCGGTGAGAAGCCTTTTGCTTGCGATATCTGTGGA CGAAAATTTGCCCGGAGGTTCACTCTCTCCAAACACACAAAGATACATCTGCGCCAAA AGGATGCAGCCCGG	ZFP DNA 서열
892	GCCGCAATGGCTGAGCGCCCGTTCCAGTGCAGAATATGCATGCGGAATTTTTCTAGGTC AGATCATTTGTCTGAGCATATTCGCACACACACGGGAGAGAAGCCCTTTGCTTGCGATA TATGTGGAAGGAAATTCGCGCAATACAGTGGGCGCTACTACCATACAAAGATCCATAC GGGCTCCCAGAAGCCCTTCCAATGTCGAATATGTATGAGGAATTTTAGTCACGGACAA ACATTGAATGAACATATACGCACTCACACTGGTGAAAAACCATTTGCGTGCGATATTTG CGGAAGGAAGTTTGCTCAGTCTGGGAATTTGGCGCGACACACCAAGATCCACACAGGA TCCCAGAAACCATTTCAGTGCAGAATTTGTATGAGAAACTTTAGCCGCAGTGACAGTCT	ZFP DNA 서열

[0304]

	CTTGAGGCACATACGGACTCATACTGGGGAGAAACCATTCGCCTGCGATATTTGTGGA CGAAAGTTCGCCTGTCGCGAGTACAGAGGCAAGCACACTAAGATACATCTTAGGCAAA AGGACGCTGCACGG	
893	GCCGCGATGGCTGAGAGGCCTTTTCAATGTCGAATCTGTATGAGGAACTTCTCTCAATC TGCTAATCGCACGACGCACATTCGAACGCATACCGGTGAGAAGCCATTCGCGTGCGAT ATCTGCGGACGGAAATTCGCGAGGTCAGCTAATCTTACACGGCACACGAAGATCCACA CAGGGTCACAGAAACCTTTTCAGTGTCGCATTTGCATGAGGAATTTCTCCCGATCTGAC GTCCTTAGCGAACATATACGAACTCACACGGGCGAGAAGCCATTTGCGTGCGATATAT GCGGGAGGAAGTTTGCCACCTCTGGACATCTGAGTCGACATACCAAAATTCATACCGG TAGTCAGAAGCCGTTCCAATGCAGAATATGTATGCGAAATTTCTCTCAAAGCTCAGACT TGTCTAGGCACATAAGAACGCACACGGGTGAAAAACCTTTCGCGTGTGATATCTGCGG CAGAAAGTTCGCACAATGGTCCACCCGAAAGCGGCATACGAAGATTCACCTCAGACAG AAAGACGCTGCCCGG	ZFP DNA 서열
894	GCGGCGATGGCAGAACGCCCGTTCCAATGCAGAATATGTATGAGAAACTTCTCCCAGA GCGGAAATCTGGCACGCCACATCCGGACACACACGGGAGAGAAGCCATTTGCTTGTGA CATTTGTGGTCGCAAATTTGCCGCCACCTGTTGTCTGGCACATCATACTAAGATACATA CGGGGTCACAGAAACCATTCCAATGTAGGATCTGCATGCGGAATTTTTCTCGGTGGCAG TATTTGCCTACGCATATTAGAACCCACGCCGGTGAGAAACCGTTTGCATGTGACATCTG CGGACGAAAGTTTGCCGATAGATCTGCGCTTGCTAGGCATACTAAAATCCACACGGGG TCCCAGAAGCCTTTTCAGTGTCGGATATGTATGAGGAACTTCAGTCGATCAGACAACCT TAGCGAGCATATTCGGACGCATACTGGAGAAAAACCTTTTGCTTGTGATATATGCGGTA GGAAGTTCGCCAAACGGTGTAACCTTCGCTGTCACACCAAAATACATCTTCGCCAGAA AGATGCGGCCCGG	ZFP DNA 서열
895	GCCGCTATGGCTGAAAGACCATTCCAGTGCAGAATATGTATGAGGAATTTTTCTAATCC CGCGAACCTTACGCGCCATATCAGGACGCACACGGGCGAAAAGCCCTTCGCCTGCGAC ATTTGTGGGAGAAAGTTTGCTCAAAACGCGACCAGGACAAAGCACACGAAAATTCACA CTGGTAGCCAGAAGCCGTTCCAGTGTAGGATCTGTATGCGCAATTTCTCTCAGTCCGGG CACCTCGCGCGACACATAAGAACTCATACGGGGGAGAAGCCGTTTGCATGTGACATCT GCGGCCGCAAGTTTGCGAATAGGCATGACAGGGCAAAACATACGAAGATCCATACAG GTTCTCAAAAACCTTTCCAATGTCGAATATGCATGCGCAACTTTAGTCGGTCAGACCAC CTTTCTGAACACATCAGGACACACACTGGCGAAAAGCCGTTCGCATGTGACATTTGCG GCAGAAAGTTCGCACAAAGACGGTCCCGCTATAAGCACACCAAAATTCACCTTAGGCA AAAGGATGCAGCTCGG	ZFP DNA 서열
896	GCGGCAATGGCAGAACGACCCTTCCAATGCCGCATATGTATGCGAAACTTCAGCCAGA GCTCAGATCTTTCCAGACACATCAGGACTCACACTGGCGAAAAACCATTTGCATGCGAT ATATGCGGGAGAAAATTCGCGCACCGCAGTACGCGAAACAGGCATACAAAGATACAT ACTGGCAGTCAAAAGCCATTTCAATGTCGAATATGCATGAGGAACTTTAGTCGATCTGA CGTGCTGAGCGCTCACATACGGACCCATACCGGAGAGAAACCATTCGCTTGTGACATC TGTGGTAGGAAGTTCGCGGATTCCCGGACCCGCAAAAATCATACTAAAATTCACACTG GGTCTCAGAAGCCCTTTCAGTGTAGGATATGTATGCGCAATTTTAGCCAGAGTGGTTCA TTGACTCGGCATATCAGAACACATACTGGAGAGAAACCTTTCGCGTGTGATATTTGCGG TCGAAAGTTCGCAGATCAGAGTGGACTTGCGCACCATACTAAGATCCACCTGAGACAG AAGGACGCTGCGCGG	ZFP DNA 서열
897	GCTGCCATGGCGGAGCGCCCTTTCCAGTGTAGGATATGTATGCGCAACTTCAGTCAGAA CCCAGCCCAGTGGCGGCACATACGGACGCATACTGGAGAAAAGCCATTTGCATGTGAT ATCTGCGGGCGAAAATTCGCGCGGTCAGCAGATTTGAGCCGGCATACGAAGATCCATA CAGGTTCACAAAAGCCATTTCAATGTCGGATATGTATGCGGAACTTCAGCACGTCCGGC TCATTGTCAAGACATATACGAACTCATACCGGAGAGAAACCCTTCGCGTGCGACATTTG CGGTCGGAAGTTCGCGCGATCCGACCATCTGTCACGACATACGAAAATACACACTGGC TCTCAAAAGCCGTTTCAGTGCAGAATTTGCATGAGAAATTTTAGCAGGAGCGACTCACT CCTTCGGCATATACGAACACACACTGGTGAGAAGCCATTTGCCTGTGATATTTGTGGAC GAAAGTTTGCGCAATCTTACGATAGGTTTCAGCATACAAAAATCCACCTTCGGCAAAA GGACGCGGCACGG	ZFP DNA 서열
898	GCTGCCATGGCTGAACGACCGTTTCAATGTCGAATTTGCATGCGCAACTTCTCCACGTC CGGGTCTCTCAGTAGACACATCAGAACGCATACTGGTGAAAAACCATTCGCTTGTGAC ATATGCGGCCGAAAATTCGCGCGGAGCGACCACCTGTCACGGCATACCAAAATTCACA CCGGGAGTCAAAAACCGTTCCAGTGTAGGATATGTATGCGCAACTTCAGCCGGTCTGA CAGTCTGCTTCGACATATTCGGACGCACACTGGTGAAAAGCCGTTTGCGTGCGACATTT GTGGTCGAAAGTTCGCTCAATCTTATGATAGGTTTCAACACACCAAAATACATACGGGC	ZFP DNA 서열

[0305]

	TCCCAGAAGCCGTTCCAGTGCAGAATATGCATGAGAAATTTCTCTCGCAGTGACAATTT GTCCACCCATATTCGAACGCACACCGGCGAGAAACCCTTCGCCTGCGATATTTGCGGTC GCAAGTTCGCAGACAACAGGGATAGGATAAAACATACGAAGATCCATCTGAGGCAAA AAGACGCCGCCCGG	
899	GCAGCCATGGCAGAGCGGCCATTCCAGTGCAGAATCTGCATGCGGAACTTTTCCGATA GGTCCAATCTGTCACGCCATATTAGGACACACACGGGTGAAAAACCGTTCGCGTGTGA CATATGCGGTCGCAAATTCGCCCTGAGACAGAACCTGATTATGCACACAAAAATACAT ACGGGAAGCCAGAAACCGTTCCAGTGTCGGATATGCATGAGGAACTTCAGTGAGAGGG GGACTTTGGCGAGGCACATCAGGACTCACACTGGGGAGAAGCCCTTTGCATGTGATAT CTGTGGCCGAAAATTTGCTCGATCAGATGCTCTCACCCAACATACAAAGATCCATACTG GCTCTCAAAAACCGTTTCAATGTAGAATTTGTATGCGCAACTTCTCTCGGTCAGATAGC CTGTCCCAGCATATCCGAACTCATACAGGTGAGAAACCCTTCGCATGCGACATCTGTGG GCGAAAATTTGCTAGAAAAGCAGACCGGACCCGACACACAAAGATTCATCTGCGACAA AAAGACGCCGCCCGG	ZFP DNA 서열
900	GCGGCCATGGCTGAGAGGCCTTTTCAATGTAGAATATGTATGCGAAATTTTTCACAGTA CTGTTGTCTCACGAACCACATAAGGACTCATACAGGGGAGAAACCATTTGCCTGTGAC ATTTGCGGTCGCAAATTTGCTACTTCTGGAAACCTGACTCGGCACACTAAGATTCACAC AGGGTCCCAGAAGCCCTTCCAGTGTCGCATTTGCATGAGGAATTTTAGTCAAAGCTCTG ACTTGTCAAGGCATATTCGCACGCACACGGGCGAAAAGCCGTTCGCTTGCGACATATG CGGGCGGAAATTTGCCTTCCGCTATTATTTGAAGAGACACACCAAGATACATACGGGC TCTCAGAAGCCCTTTCAGTGTAGGATTTGCATGCGCAATTTTTCACAATCTGGTGATCTC ACGCGACACATCCGGACTCACACAGGTGAAAAGCCTTTCGCGTGCGACATTTGCGGCC GGAAGTTTGCTGACAAGGGCAACCTCACAAAGCATACGAAGATTCACTTGAGGCAGAA AGATGCTGCTCGG	ZFP DNA 서열
901	GCCGCCATGGCCGAACGACCATTCCAGTGCAGGATATGTATGCGCAATTTTTCAACCAG TGGTTCATTGTCACGACATATTAGAACACACACCGGTGAGAAACCCTTTGCGTGTGACA TCTGTGGGAGGAAATTCGCAAGATCTGACAACCTTACGACACATACAAAGATTCACAC AGGCTCTCAAAAGCCCTTCCAGTGCCGAATTTGCATGCGAAACTTTTCCCAGTCTGGTA ATCTCGCTCGACATATCAGAACCCACACGGGGGAAAAACCATTCGCTTGTGATATTTGC GGACGAAAGTTCGCCGACAGAACCACACTCATGAGACACACTAAAATCCATACTGGTA GTCAGAAGCCGTTTCAGTGTAGAATCTGCATGAGGAACTTTTCCCAGTCAGGCCACCTT GCAAGACATATACGAACTCACACTGGAGAAAAGCCGTTCGCCTGTGACATTTGTGGGC GCAAGTTCGCGCAACTCACCCATCTGAATAGCCATACGAAGATTCACTTGAGACAGAA AGATGCGGCTCGG	ZFP DNA 서열
902	GCAGCTATGGCTGAACGCCCATTCCAGTGTCGGATCTGCATGCGCAACTTTTCTATAAA ACACGATCTTCACCGACACATTCGGACACATACTGGGGAGAAGCCCTTTGCGTGTGAC ATCTGTGGCCGAAAGTTCGCTAGATCCGCAAACTTGACTCGGCATACGAAAATTCACA CTGGAAGCCAGAAACCTTTCCAATGTCGAATCTGTATGAGGAACTTTAGCAGAAGTGA TAATCTCGCCAGGCATATCCGAACGCACACAGGCGAGAAGCCATTCGCATGTGATATT TGTGGTAGAAAGTTCGCCCAAAATGTCTCTCGCCCACGCCATACTAAGATCCACACGG GCTCCCAGAAGCCGTTCCAATGCCGCATTTGCATGCGAAACTTTTCCAGATCAGACGAT CTGAGCAAGCATATTAGGACGCATACAGGGGAGAAGCCTTTTGCTTGCGACATTTGCG GCCGGAAATTTGCTGACTCAAGTCACAGAACACGGCATACCAAGATACACCTTCGACA AAAAGATGCCGCACGG	ZFP DNA 서열
903	GCGGCCATGGCGGAACGACCCTTTCAGTGCCGAATTTGCATGAGGAACTTTTCACGATC TGATAACCTGGCGAGGCACATCCGAACACATACGGGCGAGAAGCCATTCGCATGTGAT ATCTGCGGGCGAAAGTTCGCCCAAAATGTCAGTAGACCGCGACATACTAAAATACACA CTGGCTCACAGAAGCCGTTCCAATGCCGCATCTGTATGCGCAATTTTTCCCGAAGCGAC GATCTGTCTAAACATATTCGGACGCACACTGGGGAAAAGCCTTTCGCTTGTGACATCTG TGGGAGGAAGTTCGCTGACAGCTCTCATAGGACACGCCATACTAAGATTCATACCGGA AGCCAGAAGCCTTTCCAGTGTCGGATTTGCATGAGAAACTTTAGCACTTCTAGCAACAG AAAGACACATATACGAACCCATACGGGTGAGAAACCGTTCGCATGCGATATCTGTGGG CGAAAATTTGCAGCCCAATGGACCAGAGCTTGCCATACCAAGATACACCTTCGGCAGA AGGACGCTGCACGG	ZFP DNA 서열
904	GCTGCGATGGCAGAACGACCTTTTCAATGCCGAATTTGTATGAGGAACTTTTCCCGGTC AGACGACCTTTCCAAGCACATCAGAACTCATACCGGAGAAAAACCGTTCGCCTGTGAC ATTTGTGGACGGAAGTTTGCTGACTCCTCTCACAGGACTCGCCACACTAAGATACACAC CGGAAGTCAGAAGCCCTTCCAATGTAGGATATGCATGAGAAACTTCAGTACGTCATCA AACCGAAAAACGCATATCAGGACACATACCGGCGAAAAGCCGTTTGCATGTGATATCT	ZFP DNA

[0306]

	GCGGCAGGAAATTTGCAGCTCAGTGGACACGGGCATGTCACACAAAAATCCATACCGG TAGTCAAAAACCGTTTCAGTGTCGAATCTGCATGAGGAACTTTAGCCGGAAGCAGACG AGAACCACGCATATAAGAACTCACACAGGTGAGAAACCCTTTGCGTGCGATATCTGCG GTCGCAAATTTGCTCACCGATCCTCCCTGAGGCGACATACTAAAATACATCTGCGACAG AAAGACGCGGCTCGG	서열
905	GCCGCCATGGCAGAACGGCCTTTTCAGTGTCGGATCTGCATGAGAAACTTTAGTCAGA GTGCCCATCGCAAGAATCATATTCGAACTCATACCGGTGAAAAACCGTTCGCGTGCGA CATCTGTGGTCGAAAGTTTGCCACATCATCCAATAGAAAAACGCATACTAAGATTCATA CCGGTAGCCAGAAACCATTCCAATGTAGAATCTGCATGCGAAATTTCAGCAGGAGTGA CAATTTGTCCGCACATATACGGACACACACGGGCGAAAAACCCTTTGCTTGCGATATAT GCGGTAGAAAGTTCGCTAGGAACAACGACCGAAAAACACACACGAAGATTCATACAG GTAGTCAGAAGCCATTTCAATGTCGGATCTGTATGCGAAATTTCTCTACTTCTGGCAGC CTGTCCCGGCACATCAGAACACATACCGGTGAGAAGCCATTTGCATGTGACATATGTG GGAGAAAATTTGCCCAGGCGGGTCACCTTGCGAAGCATACAAAGATCCACCTCCGCCA GAAGGACGCCGCACGG	ZFP DNA 서열
906	GCGGCTATGGCAGAGCGGCCATTCCAATGTAGGATATGTATGAGGAACTTCTCCCGGA GCGATCACCTGTCCCAACACATCCGCACGCATACGGGTGAGAAGCCGTTTGCTTGTGAT ATCTGCGGAAGAAAATTTGCAGCATCCAGTACACGCACAAAGCATACGAAGATTCATA CGGGATCCCAAAAGCCCTTTCAATGCAGGATTTGTATGAGGAACTTCAGTCGGTCCGAC GATCTGACACGACATATTAGAACTCATACTGGAGAGAAGCCATTCGCATGTGACATCT GCGGTAGGAAGTTCGCGCAGAAATCTAACCTGTCATCTCACACCAAGATACATACAGG CTCACAGAAGCCGTTTCAATGCCGCATCTGCATGAGGAATTTCAGCCAGTCCGCAAAC AGAACTACGCATATTCGGACGCATACCGGCGAGAAGCCGTTTGCCTGCGACATTTGCG GGAGGAAATTCGCACAGAACGCGACCAGAACCAAACACACCAAAATCCATCTTAGGC AAAAGGATGCGGCCCGG	ZFP DNA 서열
907	GCGGCCATGGCAGAACGACCCTTTCAGTGCCGAATTTGCATGCGGAACTTTAGTCGCA GTGACACCCTGAGCGAGCATATTCGCACGCATACGGGAGAGAAGCCATTTGCATGCGA CATCTGCGGTAGAAAGTTTGCGAGGCGCTGGACGTTGGTAGGCCACACGAAAATCCAT ACAGGCTCCCAGAAACCCTTCCAGTGCAGAATTTGTATGCGCAATTTTAGTGACAGAA GTAACTTGTCCCGACATATAAGGACGCACACCGGCGAAAAACCGTTTGCCTGTGATAT CTGTGGTCGGAAATTCGCCCAGTCCGGTGACTTGACACGGCATACCAAAATACACACT GGAAGCCAAAAGCCTTTTCAGTGTCGCATATGTATGCGCAACTTCAGCCAGAGTAGTG ACCTTTCACGGCATATACGGACGCATACGGGTGAGAAACCCTTCGCCTGTGACATTTGC GGGCGAAAGTTTGCATATCATTGGTACCTGAAAAAACATACGAAAATACATTTGAGAC AAAAAGATGCAGCCCGG	ZFP DNA 서열
908	GCTGCCATGGCCGAGCGCCCGTTTCAATGCAGAATATGTATGAGGAACTTCTCTAGAA GCGCCAATCTTGCGCGACACATTAGAACTCACACAGGCGAGAAACCTTTCGCCTGTGA CATCTGCGGCAGAAAATTTGCGCGAAGTGATAACTTGCGCGAGCACACAAAAATCCAT ACCGGTTCCCAAAAGCCCTTTCAATGTAGGATTTGCATGAGGAACTTTTCAAGACCATA CACGCTGAGACTCCACATTCGCACGCATACGGGAGAAAAACCATTTGCTTGTGACATA TGCGGCCGAAAGTTTGCACACCGATCCAACTTGAATAAGCATACCAAAATCCATACGG GGTCTCAGAAACCCTTCCAGTGTCGGATTTGTATGCGAAACTTCTCTCAGAGTGGTTCT CTTACGCGCCACATTAGAACCCACACGGGGGAAAAGCCATTCGCGTGCGATATCTGTG GCCGGAAATTTGCTACTTCCGCAAATCTTTCTCGACATACAAAGATACATCTTAGACAG AAGGATGCGGCACGG	ZFP DNA 서열
909	TTGGAGCCCGGTGAAAAACCATATGCCTGTCCAGAGTGTGGAAAAAGTTTTAGCAGAA GCGACGACCTGGTTAGGCATCAACGAACTCACACAGGGGAAAAGCCGTACAAATGTCC TGAATGCGGAAAGTCATTCTCCACTTCAGGTAGCCTCGTGCGACATCAGCGCACACATA CTGGCGAAAAACCTTATAAGTGTCCGGAATGCGGCAAGAGTTTTTCTAGAAGTGATAA GCTCGTGCGACACCAGCGGACGCACACGGGTGAGAAGCCCTACGCCTGCCCAGAGTGC GGAAAGTCCTTTAGTAGGTCTGACGAACTGGTCCGGCACCAACGAACCCATACAGGGG AGAAACCCTACAAATGTCCAGAGTGCGGGAAATCATTTTCAACGTCTCACAGCTTGAC GGAACATCAGAGGACCCATACAGGCGAAAAGCCTTACAAATGTCCGGAGTGCGGAAA GTCTTTCAGCCGGGCTGATAATCTCACGGAGCACCAAAGAACCCACACGGGTAAGAAA ACCTCT	ZFP DNA 서열
910	TTGGAGCCGGGAGAAAAGCCATATGCCTGTCCGGAATGCGGCAAAAGCTTTTCAGAAC GGTCCCATTTGCGGGAACATCAGCGCACCCATACCGGAGAAAAGCCTTACAAGTGTCC CGAGTGTGGTAAGAGCTTTTCAACTTCCCACAGTCTCACTGAACATCAGCGAACTCACA CAGGCGAAAAACCATACAAATGCCCGGAGTGTGGAAAGAGTTTTTCTCAAGCTGGGCA	ZFP DNA

[0307]

	CTTGGCCAGCCACCAAAGGACTCATACGGGCGAGAAACCGTATGCCTGTCCAGAATGC GGGAAATCATTTTCTACTAGTCATTCCCTTACGGAACATCAGAGAACGCACACCGGGG AGAAGCCATACAAGTGTCCGGAGTGCGGAAAATCATTCTCCGACCCGGGTCATCTCGT TCGCCATCAGAGGACTCATACTGGAGAGAAACCGTATAAATGTCCTGAATGCGGGAAG AGCTTTTCTACATCTGGGAACCTTGTGCGGCACCAGCGAACCCACACGGGGAAAAAGA CTTCT	서열
911	CTGGAACCGGGTGAAAAACCCTACGCTTGTCCCGAGTGTGGGAAATCATTCTCAAGGG CAGATAATCTTACGGAACATCAGCGCACACACACCGGGGAAAAACCCTATAAGTGCCC TGAGTGTGGTAAATCATTCTCTACGTCAGGGTCATTGGTGCGCCATCAGAGGACCCATA CAGGCGAAAAACCGTATAAGTGCCCCGAATGCGGTAAAAGCTTCAGCAGAAAAGATA ACTTGAAAAATCACCAACGCACTCACACGGGAGAAAAACCGTACGCATGCCCGGAGTG CGGCAAGAGTTTCAGTCAGAGTAGCTCACTTGTTCGCCACCAAAGAACCCACACAGGC GAAAAGCCCTACAAGTGTCCTGAATGTGGAAAGAGTTTCAGTCGAAGTGATAAATTGG TTAGGCACCAAAGAACACACACTGGAGAGAAACCGTACAAGTGTCCAGAATGTGGCA AGTCTTTTTCTGATTCTGGAAATTTGCGCGTCCATCAAAGGACTCATACTGGGAAAAAG ACGTCC	ZFP DNA 서열
912	CTCGAACCAGGGGAAAAGCCATACGCTTGCCCCGAGTGTGGAAAATCTTTCTCTCAGTC AAGCTCCCTTGTCAGACACCAGAGAACCCATACGGGTGAGAAACCATACAAGTGCCCA GAGTGCGGCAAAAGCTTCAGTCAGAGTGGCGATCTGCGCCGGCATCAAAGAACTCACA CTGGAGAAAAGCCCTATAAGTGCCCCGAATGTGGAAAGAGTTTTTCCAGGAGTGACGA AAGAAAGAGACACCAGCGGACTCACACCGGCGAGAAACCATACGCATGCCCCGAGTG CGGAAAAAGTTTTTCCCACCGCACAACCCTTACCAATCATCAACGCACGCACACAGGG GAGAAACCCTACAAGTGCCCGGAGTGTGGCAAGTCATTCAGCCGAAGTGACCACCTCA CTAACCACCAAAGGACTCACACAGGAGAAAAACCCTACAAATGTCCGGAGTGCGGAA AATCTTTTTCCACGTCCGGTGAGCTGGTCCGCCATCAACGAACCCATACTGGTAAAAAA ACTAGC	ZFP DNA 서열
913	CTGGAACCCGGGGAGAAGCCGTACGCTTGCCCAGAGTGCGGAAAGAGCTTCTCTCAGA GCGGAGACCTTAGACGCCACCAGCGAACCCACACCGGCGAAAAACCGTATAAATGCCC GGAATGCGGCAAGAGTTTTAGTCGGTCCGATGAGCGAAAGAGGCATCAACGAACCCAT ACGGGAGAGAAACCCTACAAGTGCCCTGAGTGTGGTAAGTCATTTTCCCACAGAACGA CGTTGACGAATCACCAGAGAACCCATACGGGTGAGAAACCTTACGCTTGCCCGGAGTG CGGCAAAAGCTTCAGCCGGAGTGATCACTTGACCAATCATCAGAGGACACACACGGGT GAGAAGCCCTACAAATGTCCCGAATGCGGCAAGTCTTTCTCAACGTCAGGCGAACTCG TCCGGCACCAGCGAACACATACGGGAGAAAAGCCGTACAAATGTCCGGAATGCGGAA AGTCATTTTCACGGTCAGATGACTTGGTGCGACACCAGCGCACTCACACAGGCAAGAA GACCTCA	ZFP DNA 서열
914	CTGGAGCCTGGTGAGAAGCCGTATGCATGTCCTGAGTGTGGGAAGTCATTTAGTCAGA GGGCCCACTTGGAACGACACCAAAGGACCCACACTGGTGAAAAACCCTACAAATGCCC AGAGTGTGGTAAGTCTTTTTCACAGCTGGCCCACCTGAGAGCACACCAGCGAACTCAT ACGGGCGAGAAACCATACAAGTGTCCAGAGTGCGGAAAGTCATTCTCAGATCCCGGCC ACTTGGTGCGACATCAGAGAACGCACACAGGGGAGAAGCCTTATGCTTGCCCGGAATG CGGGAAGTCTTTCAGCCGCCGAAGTGCTTGTCGAAGGCACCAACGGACCCATACCGGT GAGAAACCATATAAGTGCCCAGAGTGTGGAAAGAGTTTTAGTCGATCCGATCACCTGA CTACGCACCAGCGGACGCACACAGGAGAGAAACCGTATAAGTGCCCTGAATGCGGTAA GAGCTTCTCTCAATCAAGCTCACTGGTTAGGCACCAACGCACTCATACCGGCAAGAAG ACGTCA	ZFP DNA 서열
915	CTTGAGCCGGGGGAAAAGCCTTATGCTTGCCCAGAGTGTGGCAAGAGCTTCTCCCAAA GTTCAAACCTCGTCCGACACCAAAGGACTCACACGGGCGAAAAACCGTATAAATGCCC CGAGTGCGGAAAGTCATTTTCCAGGTCCGACGATCTGGTCCGCCACCAGCGCACTCATA CGGGGGAAAAGCCCTATAAGTGCCCTGAGTGCGGCAAGAGCTTCTCAACTCACCTGGA TCTCATTCGCCATCAACGCACACATACAGGGGAGAAGCCTTACGCTTGTCCAGAGTGC GGCAAGTCTTTCAGTACGAGCGGAAACCTGACGGAACACCAGCGAACCCACACGGGCG AGAAGCCATATAAATGTCCCGAATGTGGAAAATCATTCTCTCGCCGATCTGCGTGCCGC CGGCATCAGAGGACACATACCGGAGAAAAGCCGTACAAATGCCCCGAGTGTGGAAAA TCCTTTAGCAGAAATGATACACTTACCGAGCATCAGAGGACGCACACTGGAAAAAAGA CATCT	ZFP DNA 서열

[0308]

[0309] [0303] 일부 구현예에서, eZFP(예컨대 후성적 변형 DNA-표적화 시스템의 융합 단백질에 포함된 eZFP)는 N-말단에서 C-말단 순서로 F1 내지 F6으로 표시되는 6개의 아연 집게를 포함하는 아연 집게 단백질을 포함한다. 일부 구현예에서, 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같다: 1) F1:SEADRSR(서열번호 720) F2:DRSNLTR(서열번호 721)F3:QSSDLSR(서열번호 722) F4:YHWYLKK(서열번호 723) F5:RSDSLSV(서열번호 724) F6:QNANRKT(서열번호 725); 2) F1:RSDVLST(서열번호 726) F2:DNSSRTR(서열번호 727) F3:RPYTLRL(서열번호 728) F4:DSSHRTR(서열번호 729) F5:RSDHLSQ(서열번호 730) F6:DSSHRTR(서열번호 731); 3) F1:RSDHLSQ(서열번호 732) F2:QSADRTK(서열번호 733) F3:RSDHLSQ(서열번호 734) F4:RRSDLKR(서열번호 735) F5:RSDHLSR(서열번호 736) F6:QSSDLRR(서열번호 737); 4) F1:RSDNLSE(서열번호 738) F2:TSSNRKT(서열번호 739) F3:DRSHLTR(서열번호 740) F4:RSDALTQ(서열번호 741) F5:DRSALAR(서열번호 742) F6:RRFTLSK(서열번호 743); 5) F1:RSDHLSE(서열번호 744) F2:QYSGRYY(서열번호 745) F3:HGQTLNE(서열번호 746) F4:QSGNLAR(서열번호 747) F5:RSDSLLR(서열번호 748) F6:CREYRGK(서열번호 749); 6) F1:QSANRTT(서열번호 750) F2:RSANLTR(서열번호 751) F3:RSDVLSE(서열번호 752) F4:TSGHLSR(서열번호 753) F5:QSSDLSR(서열번호 754), F6:QWSTRKR(서열번호 755); 7) F1:QSGNLAR(서열번호 756) F2:ATCCLAH(서열번호 757) F3:RWQYLPT(서열번호 758) F4:DRSALAR(서열번호 759) F5:RSDNLSE(서열번호 760)F6:KRCNLRC(서열번호 761); 8) F1:NPANLTR(서열번호 762) F2:QNATRTK(서열번호 763) F3:QSGHLAR(서열번호 764) F4:NRHDRAK(서열번호 765) F5:RSDHLSE(서열번호 766), F6:QRRSRYK(서열번호 767); 9) F1:QSSDLSR(서열번호 768) F2:HRSTRNR(서열번호 769) F3:RSDVLSA(서열번호 770) F4:DSRTRKN(서열번호 771) F5:QSGSLTR(서열번호

772) F6:DQSGLAH(서열번호 773); 10) F1:QNPAQWR(서열번호 774) F2:RSADLSR(서열번호 775) F3:TSGSLSR(서열번호 776) F4:RSDHLSR(서열번호 777) F5:RSDSLLR(서열번호 778) F6:QSYDRFQ(서열번호 779); 11) F1:TSGSLSR(서열번호 780) F2:RSDHLSR(서열번호 781) F3:RSDSLLR(서열번호 782) F4:QSYDRFQ(서열번호 783) F5:RSDNLST(서열번호 784) F6:DNRDRIK(서열번호 785);12) F1:DRSNLSR(서열번호 786) F2:LRQNLIM(서열번호 787) F3:ERGTLAR(서열번호 788) F4:RSDALTQ(서열번호 789) F5:RSDSLSQ(서열번호 790) F6:RKADRTR(서열번호 791); 13) F1:QYCCLTN(서열번호 792) F2:TSGNLTR(서열번호 793) F3:QSSDLSR(서열번호 794) F4:FRYYLKR(서열번호 795) F5:QSGDLTR(서열번호 796) F6:DKGNLTK(서열번호 797); 14) F1:TSGSLSR(서열번호 798) F2:RSDNLTT(서열번호 799) F3:QSGNLAR(서열번호 800) F4:DRTTLMR(서열번호 801) F5:QSGHLAR(서열번호 802) F6:QLTHLNS(서열번호 803); 15) F1:IKHDLHR(서열번호 804) F2:RSANLTR(서열번호 805) F3:RSDNLAR(서열번호 806) F4:QNVSRPR(서열번호 807) F5:RSDDLSK(서열번호 808) F6:DSSHRTR(서열번호 809); 16) F1:RSDNLAR(서열번호 810) F2:QNVSRPR(서열번호 811) F3:RSDDLSK(서열번호 812) F4:DSSHRTR(서열번호 813) F5:TSSNRKT(서열번호 814) F6:AQWTRAC(서열번호 815); 17) F1:RSDDLSK(서열번호 816) F2:DSSHRTR(서열번호 817) F3:TSSNRKT(서열번호 818) F4:AQWTRAC(서열번호 819) F5:RKQTRTT(서열번호 820) F6:HRSSLRR(서열번호 821); 18) F1:QSAHRKN(서열번호 822) F2:TSSNRKT(서열번호 823) F3:RSDNLSA(서열번호 824) F4:RNNDRKT(서열번호 825) F5:TSGSLSR(서열번호 826) F6:QAGHLAK(서열번호 827); 19) F1:RSDHLSQ(서열번호 828) F2:ASSTRTK(서열번호 829) F3:RSDDLTR(서열번호 830) F4:QKSNLSS(서열번호 831) F5:QSANRTT(서열번호 832) F6:QNATRTK(서열번호 833); 20) F1:RSDTLSE(서열번호 834) F2:RRWTLVG(서열번호 835) F3:DRSNLSR(서열번호 836) F4:QSGDLTR(서열번호 837) F5:QSSDLSR(서열번호 838) F6:YHWYLKK(서열번호 839); 21) F1:RSANLAR(서열번호 840) F2:RSDNLRE(서열번호 841) F3:RPYTLRL(서열번호 842) F4:HRSNLNK(서열번호 843) F5:QSGSLTR(서열번호 844) F6:TSANLSR(서열번호 845);m22) F1:RSDDLVR(서열번호 846) F2:TSGSLVR(서열번호 847) F3:RSDKLVR(서열번호 848) F4:RSDELVR(서열번호 849) F5:TSHSLTE(서열번호 850) F6:RADNLTE(서열번호 851); 23) F1:ERSHLRE(서열번호 852) F2:TSHSLTE(서열번호 853) F3:QAGHLAS(서열번호 854) F4:TSHSLTE(서열번호 855) F5:DPGHLVR(서열번호 856) F6:TSGNLVR(서열번호 857); 24) F1:RADNLTE(서열번호 858) F2:TSGSLVR(서열번호 859) F3:RKDNLKN(서열번호 860) F4:QSSSLVR(서열번호 861) F5:RSDKLVR(서열번호 862) F6:DSGNLRV(서열번호 863); 25) F1:QSSSLVR(서열번호 864) F2:QSGDLRR(서열번호 865) F3:RSDERKR(서열번호 866) F4:HRTTLTN(서열번호 867) F5:RSDHLTN(서열번호 868) F6:TSGELVR(서열번호 869); 26) F1:QSGDLRR(서열번호 870) F2:RSDERKR(서열번호 871) F3:HRTTLTN(서열번호 872) F4:RSDHLTN(서열번호 873) F5:TSGELVR(서열번호 874) F6:RSDDLVR(서열번호 875); 27) F1:QRAHLER(서열번호 876) F2:QLAHLRA(서열번호 877) F3:DPGHLVR(서열번호 878) F4:RRSACRR(서열번호 879) F5:RSDHLTT(서열번호 880) F6:QSSSLVR(서열번호 881); and 28) F1:QSSNLVR(서열번호 882) F2:RSDDLVR(서열번호 883) F3:THLDLIR(서열번호 884) F4:TSGNLTE(서열번호 885) F5:RRSACRR(서열번호 886) F6:RNDTLTE(서열번호 887).

[0310] [0304] 일부 구현예에서, eZFP(예컨대 후성적 변형 DNA-표적 시스템의 융합 단백질에 포함된 eZFP)는 서열번호 692-719 중 어느 하나에 제시된 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 888-915 중 어느 하나에 제시된 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩된다.

[0311] [0305] 일부 구현예에서, 본원에 따라 eZFP(예컨대 후성적 변형 DNA-표적화 시스템의 융합 단백질에 포함된 eZFP), 예를 들어 본원에 설명된 eZFP_1이다. 일부 구현예에서, eZFP는 서열번호 1028에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위, 적어도 12nt의 이의 연속 부분, 또는 상기 중 어느 하나의 상보적 서열을 표적화한다. 일부 구현예에서, eZFP는 서열번호 1028에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위를 표적으로 삼는다. 일부 구현예에서, eZFP는 N-말단에서 C-말단의 순서로 F1 내지 F6으로 표시되는 6개의 아연 집게를 포함하며, 각각은 상응하는 아연 집게 인식 영역 F1 내지 F6을 포함하고, 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같다: F1:SEADRSR(서열번호 720) F2:DRSNLTR(서열번호 721) F3:QSSDLSR(서열번호 722) F4:YHWYLKK(서열번호 723) F5:RSDSLSV(서열번호 724) F6:QNANRKT(서열번호 725): 일부 구현예에서, eZFP는 서열번호 692에 제시된 아미노산 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99% 서열 동일성을 갖는 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 692에 제시된 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 888에 제시된 뉴클레오타이드 서열, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99% 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩된다. 일부 구현예에서, eZFP는 서열번호 888에 제시된 뉴클레오타이드 서열에 의해 인코딩된다.

[0312] [0306] 일부 구현예에서, 본 발명에 따라 eZFP(예컨대 후성적 변형 DNA-표적화 시스템의 융합 단백질에 포함된 eZFP), 예를 들어 본원에 설명된 eZFP_2가 제공된다. 일부 구현예에서, eZFP는 서열번호 1029에 제시된 뉴클레

오타이드 서열을 포함하는 표적 부위, 적어도 12nt의 이의 연속된 부분 또는 상기 중 어느 하나의 상보적 서열을 표적으로 한다. 일부 구현예에서, eZFP는 서열번호 1029에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위를 표적으로 한다. 일부 구현예에서, eZFP는 N-말단에서 C-말단 순서로 F1 내지 F6으로 표시되는 6개의 아연 집게를 포함하며, 각각은 상응하는 아연 집게 인식 영역 F1 내지 F6을 포함하고, 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같다: F1:RSDVLST(서열번호 726) F2:DNSSRTR(서열번호 727) F3:RPYTLRL(서열번호 728) F4:DSSHRTR(서열번호 729) F5:RSDHLSQ(서열번호 730) F6:DSSHRTR(서열번호 731). 일부 구현예에서, eZFP는 서열번호 693에 제시된 아미노산 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 693에 제시된 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 889에 제시된 뉴클레오타이드 서열 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩된다. 일부 구현예에서, eZFP는 서열번호 889에 제시된 뉴클레오타이드 서열에 의해 인코딩된다.

[0313] [0307] 일부 구현예에서, 본 발명에 따라 eZFP(예컨대 후성적 변형 DNA-표적화 시스템의 융합 단백질에 포함된 eZFP), 예를 들어 본원에 설명된 eZFP_3이 제공된다. 일부 구현예에서, eZFP는 서열번호 1030에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위, 적어도 12nt의 이의 연속된 부분, 또는 상기 중 어느 하나의 상보적 서열을 표적으로 한다. 일부 구현예에서, eZFP는 서열번호 1030에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위를 표적으로 한다. 일부 구현예에서, eZFP는 N-말단에서 C-말단 순서로 F1 내지 F6으로 표시되는 6개의 아연 집게를 포함하며, 각각은 상응하는 아연 집게 인식 영역 F1 내지 F6을 포함하고, 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같다: F1:RSDHLSQ(서열번호 732) F2:QSADRTK(서열번호 733) F3:RSDHLSQ(서열번호 734) F4:RRSDLKR(서열번호 735) F5:RSDHLSR(서열번호 736) F6:QSSDLRR(서열번호 737). 일부 구현예에서, eZFP는 서열번호 694에 제시된 아미노산 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 694에 제시된 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 890에 제시된 뉴클레오타이드 서열 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩된다. 일부 구현예에서, eZFP는 서열번호 890에 제시된 뉴클레오타이드 서열에 의해 인코딩된다.

[0314] [0308] 일부 구현예에서, 본 발명에 따라 eZFP(예컨대 후성적 변형 DNA-표적화 시스템의 융합 단백질에 포함된 eZFP), 예를 들어 본원에 설명된 eZFP_4이 제공된다. 일부 구현예에서, eZFP는 서열번호 1031에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위, 적어도 12nt의 이의 연속된 부분, 또는 상기 중 어느 하나의 상보적 서열을 표적으로 한다. 일부 구현예에서, eZFP는 서열번호 1031에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위를 표적으로 한다. 일부 구현예에서, eZFP는 N-말단에서 C-말단 순서로 F1 내지 F6으로 표시되는 6개의 아연 집게를 포함하며, 각각은 상응하는 아연 집게 인식 영역 F1 내지 F6을 포함하고, 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같다: F1:RSDNLSE(서열번호 738) F2:TSSNRKT(서열번호 739) F3:DRSHLTR(서열번호 740) F4:RSDALTQ(서열번호 741) F5:DRSALAR(서열번호 742) F6:RRFTLSK(서열번호 743). 일부 구현예에서, eZFP는 서열번호 695에 제시된 아미노산 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 695에 제시된 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 891에 제시된 뉴클레오타이드 서열 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩된다. 일부 구현예에서, eZFP는 서열번호 891에 제시된 뉴클레오타이드 서열에 의해 인코딩된다.

[0315] [0309] 일부 구현예에서, 본 발명에 따라 eZFP(예컨대 후성적 변형 DNA-표적화 시스템의 융합 단백질에 포함된 eZFP), 예를 들어 본원에 설명된 eZFP_5가 제공된다. 일부 구현예에서, eZFP는 서열번호 1032에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위, 적어도 12nt의 이의 연속된 부분, 또는 상기 중 어느 하나의 상보적 서열을 표적으로 한다. 일부 구현예에서, eZFP는 서열번호 1032에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위를 표적으로 한다. 일부 구현예에서, eZFP는 N-말단에서 C-말단 순서로 F1 내지 F6으로 표시되는 6개의 아연 집게를 포함하며, 각각은 상응하는 아연 집게 인식 영역 F1 내지 F6을 포함하고, 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같다: F1:RSDHLSE(서열번호 744) F2:QYSGRYY(서열번호 745) F3:HGQTLNE(서열번호 746) F4:QSGNLAR(서열번호 747) F5:RSDSLLR(서열번호 748) F6:CREYRGK(서열번호 749). 일부 구현예에서, eZFP는 서열번호 696에 제시된 아미노산 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 696에

제시된 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 892에 제시된 뉴클레오타이드 서열 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩된다. 일부 구현예에서, eZFP는 서열번호 892에 제시된 뉴클레오타이드 서열에 의해 인코딩된다.

[0316] [0310] 일부 구현예에서, 본원에 따라 eZFP(예컨대 후성적 변형 DNA-표적화 시스템의 융합 단백질에 포함된 eZFP), 예를 들어 본원에 설명된 eZFP_6이 제공된다. 일부 구현예에서, eZFP는 서열번호 1033에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위, 적어도 12nt의 이의 연속된 부분, 또는 상기 중 어느 하나의 상보적 서열을 표적으로 한다. 일부 구현예에서, eZFP는 서열번호 1033에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위를 표적으로 한다. 일부 구현예에서, eZFP는 N-말단에서 C-말단 순서로 F1 내지 F6으로 표시되는 6개의 아연 집게를 포함하며, 각각은 상응하는 아연 집게 인식 영역 F1 내지 F6을 포함하고, 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같다: F1:QSANRTT(서열번호 750) F2:RSANLTR(서열번호 751) F3:RSDVLSE(서열번호 752) F4:TSGHLSR(서열번호 753) F5:QSSDLSR(서열번호 754) F6:QWSTRKR(서열번호 755). 일부 구현예에서, eZFP는 서열번호 697에 제시된 아미노산 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 697에 제시된 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 893에 제시된 뉴클레오타이드 서열 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩된다. 일부 구현예에서, eZFP는 서열번호 893에 제시된 뉴클레오타이드 서열에 의해 인코딩된다.

[0317] [0311] 일부 구현예에서, 본원에 따라 eZFP(예컨대 후성적 변형 DNA-표적화 시스템의 융합 단백질에 포함된 eZFP), 예를 들어 본원에 설명된 eZFP_7이 제공된다. 일부 구현예에서, eZFP는 서열번호 1034에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위, 적어도 12nt의 이의 연속된 부분, 또는 상기 중 어느 하나의 상보적 서열을 표적으로 한다. 일부 구현예에서, eZFP는 서열번호 1034에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위를 표적으로 한다. 일부 구현예에서, eZFP는 N-말단에서 C-말단 순서로 F1 내지 F6으로 표시되는 6개의 아연 집게를 포함하며, 각각은 상응하는 아연 집게 인식 영역 F1 내지 F6을 포함하고, 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같다: F1:QSGNLAR(서열번호 756) F2:ATCCLAH(서열번호 757) F3:RWQYLPT(서열번호 758) F4:DRSALAR(서열번호 759) F5:RSDNLSE(서열번호 760) F6:KRCNLRC(서열번호 761). 일부 구현예에서, eZFP는 서열번호 698에 제시된 아미노산 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 698에 제시된 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 894에 제시된 뉴클레오타이드 서열 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩된다. 일부 구현예에서, eZFP는 서열번호 894에 제시된 뉴클레오타이드 서열에 의해 인코딩된다.

[0318] [0312] 일부 구현예에서, 본원에 따라 eZFP(예컨대 후성적 변형 DNA-표적화 시스템의 융합 단백질에 포함된 eZFP), 예를 들어 본원에 설명된 eZFP_8이 제공된다. 일부 구현예에서, eZFP는 서열번호 1035에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위, 적어도 12nt의 이의 연속된 부분 또는 상기 중 어느 하나의 상보적 서열을 표적으로 한다. 일부 구현예에서, eZFP는 서열번호 1035에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위를 표적으로 한다. 일부 구현예에서, eZFP는 N-말단에서 C-말단 순서로 F1 내지 F6으로 표시되는 6개의 아연 집게를 포함하며, 각각은 상응하는 아연 집게 인식 영역 F1 내지 F6을 포함하고, 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같다: F1:NPANLTR(서열번호 762) F2:QNATRTK(서열번호 763) F3:QSGHLAR(서열번호 764) F4:NRHDRAK(서열번호 765) F5:RSDHLSE(서열번호 766) F6:QRRSRYK(서열번호 767). 일부 구현예에서, eZFP는 서열번호 699에 제시된 아미노산 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 699에 제시된 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 895에 제시된 뉴클레오타이드 서열 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩된다. 일부 구현예에서, eZFP는 서열번호 895에 제시된 뉴클레오타이드 서열에 의해 인코딩된다.

[0319] [0313] 일부 구현예에서, 본원에 따라 eZFP(예컨대 후성적 변형 DNA-표적화 시스템의 융합 단백질에 포함된 eZFP), 예를 들어 본원에 설명된 eZFP_9이 제공된다. 일부 구현예에서, eZFP는 서열번호 1036에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위, 적어도 12nt의 이의 연속된 부분, 또는 상기 중 어느 하나의 상보적 서열

을 표적으로 한다. 일부 구현예에서, eZFP는 서열번호 1036에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위를 표적으로 한다. 일부 구현예에서, eZFP는 N-말단에서 C-말단 순서로 F1 내지 F6으로 표시되는 6개의 아연 집게를 포함하며, 각각은 상응하는 아연 집게 인식 영역 F1 내지 F6을 포함하고, 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같다: F1:QSSDLSR(서열번호 768) F2:HRSTRNR(서열번호 769) F3:RSDVLSA(서열번호 770) F4:DSRTRKN(서열번호 771) F5:QSGSLTR(서열번호 772) F6:DQSGLAH(서열번호 773). 일부 구현예에서, eZFP는 서열번호 700에 제시된 아미노산 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 700에 제시된 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 896에 제시된 뉴클레오타이드 서열 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩된다. 일부 구현예에서, eZFP는 서열번호 896에 제시된 뉴클레오타이드 서열에 의해 인코딩된다.

[0320] [0314] 일부 구현예에서, 본원에 따라 eZFP(예컨대 후성적 변형 DNA-표적화 시스템의 융합 단백질에 포함된 eZFP), 예를 들어 본원에 설명된 eZFP_10이 제공된다. 일부 구현예에서, eZFP는 서열번호 1037에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위, 적어도 12nt의 이의 연속된 부분, 또는 상기 중 어느 하나의 상보적 서열을 표적으로 한다. 일부 구현예에서, eZFP는 서열번호 1037에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위를 표적으로 한다. 일부 구현예에서, eZFP는 N-말단에서 C-말단 순서로 F1 내지 F6으로 표시되는 6개의 아연 집게를 포함하며, 각각은 상응하는 아연 집게 인식 영역 F1 내지 F6을 포함하고, 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같다: F1:QNPAQWR(서열번호 774) F2:RSADLSR(서열번호 775) F3:TSGSLSR(서열번호 776) F4:RSDHLSR(서열번호 777) F5:RSDSLLR(서열번호 778) F6:QSYDRFQ(서열번호 779). 일부 구현예에서, eZFP는 서열번호 701에 제시된 아미노산 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 701에 제시된 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 897에 제시된 뉴클레오타이드 서열 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩된다. 일부 구현예에서, eZFP는 서열번호 897에 제시된 뉴클레오타이드 서열에 의해 인코딩된다.

[0321] [0315] 일부 구현예에서, 본원에 따라 eZFP(예컨대 후성적 변형 DNA-표적화 시스템의 융합 단백질에 포함된 eZFP), 예를 들어 본원에 설명된 eZFP_11이 제공된다. 일부 구현예에서, eZFP는 서열번호 1038에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위, 적어도 12nt의 이의 연속 부분, 또는 상기 중 어느 하나의 상보적 서열을 표적화한다. 일부 구현예에서, eZFP는 서열번호 1038에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위를 표적으로 한다. 일부 구현예에서, eZFP는 N-말단에서 C-말단 순서로 F1 내지 F6으로 표시되는 6개의 아연 집게를 포함하며, 각각은 상응하는 아연 집게 인식 영역 F1 내지 F6을 포함하고, 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같다: F1:TSGSLSR(서열번호 780) F2:RSDHLSR(서열번호 781) F3:RSDSLLR(서열번호 782) F4:QSYDRFQ(서열번호 783) F5:RSDNLST(서열번호 784) F6:DNRDRIK(서열번호 785). 일부 구현예에서, eZFP는 서열번호 702에 제시된 아미노산 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 702에 제시된 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 898에 제시된 뉴클레오타이드 서열 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩된다. 일부 구현예에서, eZFP는 서열번호 898에 제시된 뉴클레오타이드 서열에 의해 인코딩된다.

[0322] [0316] 일부 구현예에서, 본원에 따라 eZFP(예컨대 후성적 변형 DNA-표적화 시스템의 융합 단백질에 포함된 eZFP), 예를 들어 본원에 설명된 eZFP_12이 제공된다. 일부 구현예에서, eZFP는 서열번호 1039에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위, 적어도 12nt의 이의 연속된 부분, 또는 상기 중 어느 하나의 상보적 서열을 표적으로 한다. 일부 구현예에서, eZFP는 서열번호 1039에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위를 표적으로 한다. 일부 구현예에서, eZFP는 N-말단에서 C-말단 순서로 F1 내지 F6으로 표시되는 6개의 아연 집게를 포함하며, 각각은 상응하는 아연 집게 인식 영역 F1 내지 F6을 포함하고, 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같다: F1:DRSNLSR(서열번호 786) F2:LRQNLIM(서열번호 787) F3:ERGTLAR(서열번호 788) F4:RSDALTQ(서열번호 789) F5:RSDSLSQ(서열번호 790) F6:RKADRTR(서열번호 791). 일부 구현예에서, eZFP는 서열번호 703에 제시된 아미노산 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 703에 제시된 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 899에 제시된 뉴클레오타이드 서열 또는

이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩된다. 일부 구현예에서, eZFP는 서열번호 899에 제시된 뉴클레오타이드 서열에 의해 인코딩된다.

[0323] [0317] 일부 구현예에서, 본원에 따라 eZFP(예컨대 후성적 변형 DNA-표적화 시스템의 융합 단백질에 포함된 eZFP), 예를 들어 본원에 설명된 eZFP_13이 제공된다. 일부 구현예에서, eZFP는 서열번호 1040에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위, 적어도 12nt의 이의 연속된 부분, 또는 상기 중 어느 하나의 상보적 서열을 표적으로 한다. 일부 구현예에서, eZFP는 서열번호 1040에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위를 표적으로 한다. 일부 구현예에서, eZFP는 N-말단에서 C-말단 순서로 F1 내지 F6으로 표시되는 6개의 아연 집게를 포함하며, 각각은 상응하는 아연 집게 인식 영역 F1 내지 F6을 포함하고, 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같다: F1:QYCCLTN(서열번호 792) F2:TSGNLTR(서열번호 793) F3:QSSDLSR(서열번호 794) F4:FRYYLKR(서열번호 795) F5:QSGDLTR(서열번호 796) F6:DKGNLTK(서열번호 797). 일부 구현예에서, eZFP는 서열번호 704에 제시된 아미노산 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 704에 제시된 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 900에 제시된 뉴클레오타이드 서열 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩된다. 일부 구현예에서, eZFP는 서열번호 900에 제시된 뉴클레오타이드 서열에 의해 인코딩된다.

[0324] [0318] 일부 구현예에서, 본원에 따라 eZFP(예컨대 후성적 변형 DNA-표적화 시스템의 융합 단백질에 포함된 eZFP), 예를 들어 본원에 설명된 eZFP_14이 제공된다. 일부 구현예에서, eZFP는 서열번호 1041에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위, 적어도 12nt의 이의 연속된 부분, 또는 상기 중 어느 하나의 상보적 서열을 표적으로 한다. 일부 구현예에서, eZFP는 서열번호 1041에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위를 표적으로 한다. 일부 구현예에서, eZFP는 N-말단에서 C-말단 순서로 F1 내지 F6으로 표시되는 6개의 아연 집게를 포함하며, 각각은 상응하는 아연 집게 인식 영역 F1 내지 F6을 포함하고, 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같다: F1:TSGSLSR(서열번호 798) F2:RSDNLTT(서열번호 799) F3:QSGNLAR(서열번호 800) F4:DRTTLMR(서열번호 801) F5:QSGHLAR(서열번호 802) F6:QLTHLNS(서열번호 803). 일부 구현예에서, eZFP는 서열번호 705에 제시된 아미노산 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 705에 제시된 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 901에 제시된 뉴클레오타이드 서열 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩된다. 일부 구현예에서, eZFP는 서열번호 901에 제시된 뉴클레오타이드 서열에 의해 인코딩된다.

[0325] [0319] 일부 구현예에서, 본원에 따라 eZFP(예컨대 후성적 변형 DNA-표적화 시스템의 융합 단백질에 포함된 eZFP), 예를 들어 본원에 설명된 eZFP_15이 제공된다. 일부 구현예에서, eZFP는 서열번호 1042에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위, 적어도 12nt의 이의 연속된 부분, 또는 상기 중 어느 하나의 상보적 서열을 표적으로 한다. 일부 구현예에서, eZFP는 서열번호 1042에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위를 표적으로 한다. 일부 구현예에서, eZFP는 N-말단에서 C-말단 순서로 F1 내지 F6으로 표시되는 6개의 아연 집게를 포함하며, 각각은 상응하는 아연 집게 인식 영역 F1 내지 F6을 포함하고, 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같다: F1:IKHDLHR(서열번호 804) F2:RSANLTR(서열번호 805) F3:RSDNLAR(서열번호 806) F4:QNVSRPR(서열번호 807) F5:RSDDLSK(서열번호 808) F6:DSSHRTR(서열번호 809). 일부 구현예에서, eZFP는 서열번호 706에 제시된 아미노산 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 706에 제시된 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 902에 제시된 뉴클레오타이드 서열 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩된다. 일부 구현예에서, eZFP는 서열번호 902에 제시된 뉴클레오타이드 서열에 의해 인코딩된다.

[0326] [0320] 일부 구현예에서, 본원에 따라 eZFP(예컨대 후성적 변형 DNA-표적화 시스템의 융합 단백질에 포함된 eZFP), 예를 들어 본원에 설명된 eZFP_16이 제공된다. 일부 구현예에서, eZFP는 서열번호 1043에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위, 적어도 12nt의 이의 연속된 부분, 또는 상기 중 어느 하나의 상보적 서열을 표적으로 한다. 일부 구현예에서, eZFP는 서열번호 1043에 제시된 뉴클레오타이드 서열을 포함하는 표적 부

위를 표적으로 한다. 일부 구현예에서, eZFP는 N-말단에서 C-말단의 순서로 F1 내지 F6으로 표시되는 6개의 아연 집게를 포함하며, 각각은 상응하는 아연 집게 인식 영역 F1 내지 F6을 포함하고, 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같다: F1:RSDNLAR(서열번호 810) F2:QNVSRPR(서열번호 811) F3:RSDDLSK(서열번호 812) F4:DSSHRTR(서열번호 813) F5:TSSNRKT(서열번호 814) F6:AQWTRAC(서열번호 815). 일부 구현예에서, eZFP는 서열번호 707에 제시된 아미노산 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 707에 제시된 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 903에 제시된 뉴클레오타이드 서열 또는 적어도 이와 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩된다. 일부 구현예에서, eZFP는 서열번호 903에 제시된 뉴클레오타이드 서열에 의해 인코딩된다.

[0327] [0321] 일부 구현예에서, 본원에 따라 eZFP(예컨대 후성적 변형 DNA-표적화 시스템의 융합 단백질에 포함된 eZFP), 예를 들어 본원에 설명된 eZFP_17이 제공된다. 일부 구현예에서, eZFP는 서열번호 1044에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위, 적어도 12nt의 이의 연속된 부분, 또는 상기 중 어느 하나의 상보적 서열을 표적으로 한다. 일부 구현예에서, eZFP는 서열번호 1044에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위를 표적으로 한다. 일부 구현예에서, eZFP는 N-말단에서 C-말단 순서로 F1 내지 F6으로 표시되는 6개의 아연 집게를 포함하며, 각각은 상응하는 아연 집게 인식 영역 F1 내지 F6을 포함하고, 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같다: F1: RSDDLSK(서열번호 816) F2: DSSHRTR(서열번호 817) F3: TSSNRKT(서열번호 818) F4: AQWTRAC(서열번호 819) F5: RKQTRTT(서열번호 820) F6: HRSSLRR(서열번호 821). 일부 구현예에서, eZFP는 서열번호 708에 제시된 아미노산 서열 또는 이의 일부, 또는 적어도 이와 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 708에 제시된 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 904에 제시된 뉴클레오타이드 서열 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩된다. 일부 구현예에서, eZFP는 서열번호 904에 제시된 뉴클레오타이드 서열에 의해 인코딩된다.

[0328] [0322] 일부 구현예에서, 본원에 따라 eZFP(예컨대 후성적 변형 DNA-표적화 시스템의 융합 단백질에 포함된 eZFP), 예를 들어 본원에 설명된 eZFP_18이 제공된다. 일부 구현예에서, eZFP는 서열번호 1045에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위, 적어도 12nt의 이의 연속된 부분, 또는 상기 중 어느 하나의 상보적 서열을 표적으로 한다. 일부 구현예에서, eZFP는 서열번호 1045에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위를 표적으로 한다. 일부 구현예에서, eZFP는 N-말단에서 C-말단 순서로 F1 내지 F6으로 표시되는 6개의 아연 집게를 포함하며, 각각은 상응하는 아연 집게 인식 영역 F1 내지 F6을 포함하고, 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같다: F1:QSAHRKN(서열번호 822) F2:TSSNRKT(서열번호 823) F3:RSDNLSA(서열번호 824) F4:RNNDRKT(서열번호 825) F5:TSGSLSR(서열번호 826) F6:QAGHLAK(서열번호 827). 일부 구현예에서, eZFP는 서열번호 709에 제시된 아미노산 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 709에 제시된 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 905에 제시된 뉴클레오타이드 서열 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩된다. 일부 구현예에서, eZFP는 서열번호 905에 제시된 뉴클레오타이드 서열에 의해 인코딩된다.

[0329] [0323] 일부 구현예에서, 본원에 따라 eZFP(예컨대 후성적 변형 DNA-표적화 시스템의 융합 단백질에 포함된 eZFP), 예를 들어 본원에 설명된 eZFP_19이 제공된다. 일부 구현예에서, eZFP는 서열번호 1046에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위, 적어도 12nt의 이의 연속된 부분, 또는 상기 중 어느 하나의 상보적 서열을 표적으로 한다. 일부 구현예에서, eZFP는 서열번호 1046에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위를 표적으로 한다. 일부 구현예에서, eZFP는 N-말단에서 C-말단 순서로 F1 내지 F6으로 표시되는 6개의 아연 집게를 포함하며, 각각은 상응하는 아연 집게 인식 영역 F1 내지 F6을 포함하고, 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같다: F1:RSDHLSQ(서열번호 828) F2:ASSTRTK(서열번호 829) F3:RSDDLTR(서열번호 830) F4:QKSNLSS(서열번호 831) F5:QSANRTT(서열번호 832) F6:QNATRTK(서열번호 833). 일부 구현예에서, eZFP는 서열번호 710에 제시된 아미노산 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 710에 제시된 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 906에 제시된 뉴클레오타이드 서열 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드

서열에 의해 인코딩된다. 일부 구현예에서, eZFP는 서열번호 906에 제시된 뉴클레오타이드 서열에 의해 인코딩된다.

[0330] [0324] 일부 구현예에서, 본원에 따라 eZFP(예컨대 후성적 변형 DNA-표적화 시스템의 융합 단백질에 포함된 eZFP), 예를 들어 본원에 설명된 eZFP_20이 제공된다. 일부 구현예에서, eZFP는 서열번호 1047에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위, 적어도 12nt의 이의 연속된 부분, 또는 상기 중 어느 하나의 상보적 서열을 표적으로 한다. 일부 구현예에서, eZFP는 서열번호 1047에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위를 표적으로 한다. 일부 구현예에서, eZFP는 N-말단에서 C-말단 순서로 F1 내지 F6으로 표시되는 6개의 아연 집게를 포함하며, 각각은 상응하는 아연 집게 인식 영역 F1 내지 F6을 포함하고, 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같다: F1:RSDTLSE(서열번호 834) F2:RRWTLVG(서열번호 835) F3:DRSNLSR(서열번호 836) F4:QSGDLTR(서열번호 837) F5:QSSDLSR(서열번호 838) F6:YHWYLKK(서열번호 839). 일부 구현예에서, eZFP는 서열번호 711에 제시된 아미노산 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 711에 제시된 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 907에 제시된 뉴클레오타이드 서열 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩된다. 일부 구현예에서, eZFP는 서열번호 907에 제시된 뉴클레오타이드 서열에 의해 인코딩된다.

[0331] [0325] 일부 구현예에서, 본원에 따라 eZFP(예컨대 후성적 변형 DNA-표적화 시스템의 융합 단백질에 포함된 eZFP), 예를 들어 본원에 설명된 eZFP_21이 제공된다. 일부 구현예에서, eZFP는 서열번호 1048에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위, 적어도 12nt의 이의 연속된 부분, 또는 상기 중 어느 하나의 상보적 서열을 표적으로 한다. 일부 구현예에서, eZFP는 서열번호 1048에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위를 표적으로 한다. 일부 구현예에서, eZFP는 N-말단에서 C-말단의 순서로 F1 내지 F6으로 표시되는 6개의 아연 집게를 포함하며, 각각은 상응하는 아연 집게 인식 영역 F1 내지 F6을 포함하고, 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같다: F1:RSANLAR(서열번호 840) F2:RSDNLRE(서열번호 841) F3:RPYTLRL(서열번호 842) F4:HRSNLNK(서열번호 843) F5:QSGSLTR(서열번호 844) F6:TSANLSR(서열번호 845). 일부 구현예에서, eZFP는 서열번호 712에 제시된 아미노산 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 712에 제시된 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 908에 제시된 뉴클레오타이드 서열 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩된다. 일부 구현예에서, eZFP는 서열번호 908에 제시된 뉴클레오타이드 서열에 의해 인코딩된다.

[0332] [0326] 일부 구현예에서, 본원에 따라 eZFP(예컨대 후성적 변형 DNA-표적화 시스템의 융합 단백질에 포함된 eZFP), 예를 들어 본원에 설명된 eZFP_22이 제공된다. 일부 구현예에서, eZFP는 서열번호 1049에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위, 적어도 12nt의 이의 연속된 부분 또는 상기 중 어느 하나의 상보적 서열을 표적으로 한다. 일부 구현예에서, eZFP는 서열번호 1049에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위를 표적으로 한다. 일부 구현예에서, eZFP는 N-말단에서 C-말단 순서로 F1 내지 F6으로 표시되는 6개의 아연 집게를 포함하며, 각각은 상응하는 아연 집게 인식 영역 F1 내지 F6을 포함하고, 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같다: F1:RSDDLVR(서열번호 846) F2:TSGSLVR(서열번호 847) F3:RSDKLVR(서열번호 848) F4:RSDELVR(서열번호 849) F5:TSHSLTE(서열번호 850) F6:RADNLTE(서열번호 851). 일부 구현예에서, eZFP는 서열번호 713에 제시된 아미노산 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 713에 제시된 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 909에 제시된 뉴클레오타이드 서열 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩된다. 일부 구현예에서, eZFP는 서열번호 909에 제시된 뉴클레오타이드 서열에 의해 인코딩된다.

[0333] [0327] 일부 구현예에서, 본원에 따라 eZFP(예컨대 후성적 변형 DNA-표적화 시스템의 융합 단백질에 포함된 eZFP), 예를 들어 본원에 설명된 eZFP_23이 제공된다. 일부 구현예에서, eZFP는 서열번호 1050에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위, 적어도 12nt의 이의 연속된 부분, 또는 상기 중 어느 하나의 상보적 서열을 표적으로 한다. 일부 구현예에서, eZFP는 서열번호 1050에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위를 표적으로 한다. 일부 구현예에서, eZFP는 N-말단에서 C-말단 순서로 F1 내지 F6으로 표시되는 6개의 아연

집게를 포함하며, 각각은 상응하는 아연 집게 인식 영역 F1 내지 F6을 포함하고, 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같다: F1:ERSHLRE(서열번호 852) F2:TSHSLTE(서열번호 853) F3:QAGHLAS(서열번호 854) F4:TSHSLTE(서열번호 855) F5:DPGHLVR(서열번호 856) F6:TSGNLVR(서열번호 857). 일부 구현예에서, eZFP는 서열번호 714에 제시된 아미노산 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 714에 제시된 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 910에 제시된 뉴클레오타이드 서열 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩된다. 일부 구현예에서, eZFP는 서열번호 910에 제시된 뉴클레오타이드 서열에 의해 인코딩된다.

[0334] [0328] 일부 구현예에서, 본원에 따라 eZFP (예컨대 후성적 변형 DNA-표적화 시스템의 융합 단백질에 포함된 eZFP), 예를 들어 본원에 설명된 eZFP_24이 제공된다. 일부 구현예에서, eZFP는 서열번호 1051에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위, 적어도 12nt의 이의 연속된 부분, 또는 상기 중 어느 하나의 상보적 서열을 표적으로 한다. 일부 구현예에서, eZFP는 서열번호 1051에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위를 표적으로 한다. 일부 구현예에서, eZFP는 N-말단에서 C-말단의 순서로 F1 내지 F6으로 표시되는 6개의 아연 집게를 포함하며, 각각은 상응하는 아연 집게 인식 영역 F1 내지 F6을 포함하고, 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같다: F1:RADNLTE(서열번호 858) F2:TSGSLVR(서열번호 859) F3:RKDNLKN(서열번호 860) F4:QSSSLVR(서열번호 861) F5:RSDKLVR(서열번호 862) F6:DSGNLRV(서열번호 863). 일부 구현예에서, eZFP는 서열번호 715에 제시된 아미노산 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 715에 제시된 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 911에 제시된 뉴클레오타이드 서열 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩된다. 일부 구현예에서, eZFP는 서열번호 911에 제시된 뉴클레오타이드 서열에 의해 인코딩된다.

[0335] [0329] 일부 구현예에서, 본원에 따라 eZFP(예컨대 후성적 변형 DNA-표적화 시스템의 융합 단백질에 포함된 eZFP), 예를 들어 본원에 설명된 eZFP_25이 제공된다. 일부 구현예에서, eZFP는 서열번호 1052에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위, 적어도 12nt의 이의 연속된 부분, 또는 상기 중 어느 하나의 상보적 서열을 표적으로 한다. 일부 구현예에서, eZFP는 서열번호 1052에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위를 표적으로 한다. 일부 구현예에서, eZFP는 N-말단에서 C-말단 순서로 F1 내지 F6으로 표시되는 6개의 아연 집게를 포함하며, 각각은 상응하는 아연 집게 인식 영역 F1 내지 F6을 포함하고, 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같다: F1:QSSSLVR(서열번호 864) F2:QSGDLRR(서열번호 865) F3:RSDERKR(서열번호 866) F4:HRTTLTN(서열번호 867) F5:RSDHLTN(서열번호 868) F6:TSGELVR(서열번호 869). 일부 구현예에서, eZFP는 서열번호 716에 제시된 아미노산 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 716에 제시된 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 912에 제시된 뉴클레오타이드 서열 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩된다. 일부 구현예에서, eZFP는 서열번호 912에 제시된 뉴클레오타이드 서열에 의해 인코딩된다.

[0336] [0330] 일부 구현예에서, 본원에 따라 eZFP(예컨대 후성적 변형 DNA-표적화 시스템의 융합 단백질에 포함된 eZFP), 예를 들어 본원에 설명된 eZFP_26이 제공된다. 일부 구현예에서, eZFP는 서열번호 1053에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위, 적어도 12nt의 이의 연속된 부분, 또는 상기 중 어느 하나의 상보적 서열을 표적으로 한다. 일부 구현예에서, eZFP는 서열번호 1053에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위를 표적으로 한다. 일부 구현예에서, eZFP는 N-말단에서 C-말단 순서로 F1 내지 F6으로 표시되는 6개의 아연 집게를 포함하며, 각각은 상응하는 아연 집게 인식 영역 F1 내지 F6을 포함하고, 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같다: F1:QSGDLRR(서열번호 870) F2:RSDERKR(서열번호 871) F3:HRTTLTN(서열번호 872) F4:RSDHLTN(서열번호 873) F5:TSGELVR(서열번호 874) F6:RSDDLVR(서열번호 875). 일부 구현예에서, eZFP는 서열번호 717에 제시된 아미노산 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 717에 제시된 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 913에 제시된 뉴클레오타이드 서열 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩된다. 일부 구현예에서, eZFP는 서열번호 913에 제시된 뉴클레오타이드 서열에 의해 인코딩

된다.

[0337] [0331] 일부 구현예에서, 본원에 따라 eZFP(예컨대 후성적 변형 DNA-표적화 시스템의 융합 단백질에 포함된 eZFP), 예를 들어 본원에 설명된 eZFP_27이 제공된다. 일부 구현예에서, eZFP는 서열번호 1054에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위, 적어도 12nt의 이의 연속된 부분, 또는 상기 중 어느 하나의 상보적 서열을 표적으로 한다. 일부 구현예에서, eZFP는 서열번호 1054에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위를 표적으로 한다. 일부 구현예에서, eZFP는 N-말단에서 C-말단 순서로 F1 내지 F6으로 표시되는 6개의 아연 집게를 포함하며, 각각은 상응하는 아연 집게 인식 영역 F1 내지 F6을 포함하고, 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같다: F1:QRAHLER(서열번호 876) F2:QLAHLRA(서열번호 877) F3:DPGHLVR(서열번호 878) F4:RRSACRR(서열번호 879) F5:RSDHLTT(서열번호 880) F6:QSSSLVR(서열번호 881). 일부 구현예에서, eZFP는 서열번호 718에 제시된 아미노산 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 718에 제시된 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 914에 제시된 뉴클레오타이드 서열 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩된다. 일부 구현예에서, eZFP는 서열번호 914에 제시된 뉴클레오타이드 서열에 의해 인코딩된다.

[0338] [0332] 일부 구현예에서, 본원에 따라 eZFP(예컨대 후성적 변형 DNA-표적화 시스템의 융합 단백질에 포함된 eZFP), 예를 들어 본원에 설명된 eZFP_28이 제공된다. 일부 구현예에서, eZFP는 서열번호 1055에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위, 적어도 12nt의 이의 연속 부분, 또는 상기 중 어느 하나의 상보적 서열을 표적화한다. 일부 구현예에서, eZFP는 서열번호 1055에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위를 표적으로 한다. 일부 구현예에서, eZFP는 N-말단에서 C-말단 순서로 F1 내지 F6으로 표시되는 6개의 아연 집게를 포함하며, 각각은 상응하는 아연 집게 인식 영역 F1 내지 F6을 포함하고, 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같다: F1:QSSNLVR(서열번호 882) F2:RSDDLVR(서열번호 883) F3:THLDLIR(서열번호 884) F4:TSGNLTE(서열번호 885) F5:RRSACRR(서열번호 886) F6:RNDTLTE(서열번호 887). 일부 구현예에서, eZFP는 서열번호 719에 제시된 아미노산 서열 또는 이의 일부, 또는 적어도 이와 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 719에 제시된 아미노산 서열을 포함한다. 일부 구현예에서, eZFP는 서열번호 915에 제시된 뉴클레오타이드 서열 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩된다. 일부 구현예에서, eZFP는 서열번호 915에 제시된 뉴클레오타이드 서열에 의해 인코딩된다.

[0339] **D. 기타 DNA-결합 도메인**

[0340] [0333] 제공된 구현예 중 일부에서, DNA-결합 도메인은 전사 활성화제 유사 효과기(TALE), 메가뉴클레아제, 호밍 엔도뉴클레아제, 또는 I-SceI 효소 또는 이의 변이체를 포함한다. 일부 구현예에서, DNA-결합 도메인은 전술한 것 중 촉매적으로 비활성인 변이체를 포함한다.

[0341] [0334] 전사 활성인자 유사 효과기[TALE]는 잔토모나스(*Xanthomonas*) 박테리아에서 자연적으로 발견되는 단백질이다. TALE는 복수의 반복된 아미노산 서열을 포함하고, 각각의 반복은 표적 서열에서 하나의 염기에 대한 결합 특이성을 갖는다. 각각의 반복은 반복의 뉴클레오타이드 특이성을 결정하는 12번 및 13번 위치의 가변 잔기 쌍(반복 가변 이잔기[repeat variable diresidue, RVD])을 포함한다. 일부 구현예에서, 상이한 뉴클레오타이드의 인식과 연관된 RVD는 C를 인식하기 위한 HD, T를 인식하기 위한 NG, A를 인식하기 위한 NI, G 또는 A를 인식하기 위한 NN, A, C, G 또는 T를 인식하기 위한 NS, T를 인식하기 위한 HG, T를 인식하기 위한 IG, G를 인식하기 위한 NK, C를 인식하기 위한 HA, C를 인식하기 위한 ND, C를 인식하기 위한 HI, G를 인식하기 위한 HN, G를 인식하기 위한 NA, G 또는 A를 인식하기 위한 SN 및 T를 인식하기 위한 YG, A를 인식하기 위한 TL, A 또는 G를 인식하기 위한 VT 및 A를 인식하기 위한 SW이다. 일부 구현예에서, RVD는 뉴클레오타이드 A, T, C 및 G에 대한 특이성을 조절하고, 특히 이 특이성을 향상시키기 위해 다른 아미노산 잔기로 돌연변이될 수 있다. 유사한 모듈 염기당 핵산 결합 특성을 갖는 결합 도메인은 또한 상이한 박테리아 종에서 유래할 수 있다. 이들 대안적인 모듈 단백질은 TALE 반복보다 더 많은 서열 가변성을 나타낼 수 있다.

[0342] [0335] 일부 구현예에서, 'TALE DNA 결합 도메인' 또는 'TALE'는 하나 이상의 TALE 반복 도메인/단위를 포함하는 폴리펩타이드이다. 각각이 반복 가변 이잔기[RVD]를 포함하는 반복 도메인은 이의 동계 표적 DNA 서열에의 TALE의 결합에 연관되어 있다. 단일 '반복 단위'('반복'으로도 지칭됨)는 일반적으로 길이가 33-35개의 아미노

산이고, 천연 발생 TALE 단백질 내의 기타 TALE 반복 서열과 적어도 일부 서열 상동성을 나타낸다. TALE 단백질은 반복 단위 내에서 정준 또는 비정준 RVD를 사용하여 표적 부위에 결합하도록 설계될 수 있다. 예를 들어, 미국 특허 제8,586,526호 및 제9,458,205호를 참조한다.

[0343] [0336] 일부 구현예에서, TALE는 TALE에서 유래한 핵산 결합 도메인 및 효과기 도메인을 포함하는 융합 단백질이다. 일부 구현예에서, FXN 유전자좌의 하나 이상의 부위는 조작된 TALE에 의해 표적화될 수 있다.

[0344] [0337] 아연 집게 및 TALE DNA-결합 도메인은 예를 들어, 천연 발생 아연 집게 단백질의 인식 나선 영역의 조작(하나 이상의 아미노산을 변경)을 통해, DNA 결합에 관련된 TALE 반복의 아미노산('반복 가변 이잔기' 또는 RVD 영역)의 조작에 의해, 또는 모듈 DNA-결합 도메인, 예컨대 TALE 반복 또는 ZFP 도메인의 체계적인 정렬에 의해 미리 결정된 뉴클레오타이드 서열에 결합하도록 조작될 수 있다. 이에 따라, 조작된 아연 집게 단백질 또는 TALE 단백질은 비천연 발생 단백질이다. 아연 집게 단백질 및 TALE를 조작하기 위한 방법의 비제한적인 예는 설계 및 선별이다. 설계된 단백질은 자연에서 발생하지 않는 단백질이고, 이의 설계/조성은 주로 합리적인 기준에서 발생한다. 설계에 대한 합리적인 기준은 기존 ZFP 또는 TALE 설계(표준 및 비표준 RVD)와 결합 데이터에 관한 정보를 저장하는 데이터베이스에서 정보를 처리하기 위한 치환 규칙 및 컴퓨터화 알고리즘의 적용을 포함한다. 예를 들어 미국 특허 제9,458,205호, 제8,586,526호, 제6,140,081호, 제6,453,242호, 및 제6,534,261호를 참조하고, WO 98/53058, WO 98/53059, WO 98/53060, WO 02/016536 및 WO 03/016496 또한 참조한다.

[0345] **E. 효과기 도메인**

[0346] [0338] 일부 측면에서, 본원에 제공된 DNA-표적화 시스템은 하나 이상의 효과기 도메인을 포함한다. 일부 구현예에서, 본원에 제공된 DNA-표적화 시스템은 다음을 포함하는 융합 단백질을 포함한다: (a) B형 간염 바이러스 DNA 서열의 표적 부위, 예를 들어 유전자 또는 이의 조절 요소, 예를 들어 상기 설명된 것을 표적으로 하기 위한 DNA-결합 도메인, 및 (b) 적어도 하나의 효과기 도메인. 일부 측면에서, 효과기 도메인은 유전자의 전사를 감소시킬 수 있다. 즉, 전사 억제자 효과기 도메인을 포함한다. 일부 측면에서, 효과기 도메인은 전사 억제자 효과기 도메인을 포함한다.

[0347] [0339] 일부 측면에서, 효과기 도메인은 유전자 또는 이의 조절 요소에 이형적으로 동원될 때 유전자 및/또는 그 조절 요소의 전사를 억제, 유도, 촉매작용 또는 감소시킨다.

[0348] [0340] 일부 구현예에서, 효과기 도메인은 전사 억제, 전사 공동 억제, 전사 억제, 전사 인자 방출, 중합, 히스톤 변형, 히스톤 아세틸화, 히스톤 탈아세틸화, 뉴클레오솜 리모델링, 염색질 리모델링, 헤테로염색질 형성, 단백질분해, 유비퀴틴화, 탈유비퀴틴화, 인산화, 탈인산화, 스플라이싱, 핵산 결합, DNA 메틸화, DNA 탈메틸화, 히스톤 메틸화, 히스톤 탈메틸화, 또는 DNA 염기 산화를 억제, 유도하거나 촉매화하거나 이로 이어진다. 일부 구현예에서, 효과기 도메인은 전사 억제를 유도하거나 촉매화하거나 전사 억제 또는 전사 공동 억제로 이어진다. 일부 구현예에서, 효과기 도메인은 전사 억제를 유도한다. 일부 구현예에서, 효과기 도메인은 앞서 언급된 활동 자체(즉, 직접 작용함) 중 하나를 갖는다. 일부 구현예에서, 효과기 도메인은 앞서 언급된 활성 중 하나를 갖는(즉, 간접적으로 작용함) 폴리펩타이드 도메인을 동원하고/하거나 이와 상호작용한다.

[0349] [0341] 인간 유전자와 같은 내인성 포유류 유전자의 유전자 발현은 dCas9와 같은 DNA-결합 도메인과 효과기를 포함하는 융합 단백질을 표적으로 삼아 달성할 수 있다. 도메인은 하나 이상의 gRNA를 통해 포유류 유전자 또는 이의 조절 DNA 요소(예컨대 프로모터 또는 인핸서)에 전달된다. 다양한 효과기 도메인은 알려져 있으며 제공된 구현예에 따라 사용될 수 있다. 다양한 Cas 분자와 전사 억제 도메인을 갖는 Cas 융합 단백질과 같은 이러한 효과기 도메인에 의한 표적 유전자 및/또는 그 조절 요소의 억제는 예를 들어 WO2021226077, WO2017180915, WO2014197748, WO2014093655, US20190127713, WO2013176772, Adli, M. Nat. Commun. 9, 1911 (2018), Urrutia, R. Genome Biol. 4, 231 (2003), Groner, A. C. et al. PLoS Genet. 6, e1000869 (2010), Liu, X. S. et al. Cell 167, 233-247.e17 (2016), and Lei, Y. et al. Nat. Commun. 8, 16026 (2017)에 설명되어 있다.

[0350] [0342] 일부 구현예에서, 효과기 도메인은 KRAB 도메인, ERF 억제인자 도메인, MXI1 도메인, SID4X 도메인, MAD-SID 도메인, DNMT 패밀리 단백질 도메인(예를 들어, DNMT3A 또는 DNMT3B), 하나 이상의 DNMT 패밀리 단백질 또는 이의 도메인의 융합체(예를 들어, DNMT3A/L, DNMT3A 및 DNMT3L 도메인의 융합체를 포함함), EZH2 도메인, EZH2, SunTag 도메인, 전술한 것들 중 임의의 것의 부분 또는 완전 기능성 단편 또는 도메인, 또는 전술한 것들 중 임의의 것의 조합물에서 선택될 수 있다. 일부 구현예에서, 융합 단백질은 dCas9-KRAB일 수 있다. 일부 구현예에서, 융합 단백질은 DNMT3A/L-dCas9-KRAB일 수 있다. 일부 구현예에서, 융합 단백질은 KRAB-dCas9-

DNMT3A/L일 수 있다.

[0351] [0343] 일부 구현예에서, 효과기 도메인은 WO 2021/226077에 기재된 전사 억제자 도메인을 포함한다.

[0352] [0344] 일부 구현예에서, 효과기 도메인은 KRAB 도메인 또는 이의 변이체를 포함한다. KRAB-함유 아연 집게 단백질은 포유동물에서 전사 억제자의 가장 큰 패밀리을 구성한다. 크루펠 연관 박스[KRAB] 도메인은 많은 아연 집게 단백질 기반 전사 인자에 존재하는 전사 억제자 도메인이다. KRAB 도메인은 하전된 아미노산을 포함하고 하위 도메인 A 및 B로 나뉠 수 있다. KRAB 도메인은 보조억제자 KAP1(KRAB-연관 단백질-1), 후성적 판독기, 예컨대 헤테로염색질 단백질 1[heterochromatin protein 1, HP1], 및 기타 염색질 조절자를 동원하여 헤테로염색질 형성을 통해 전사 억제를 유도한다. KRAB-매개 유전자 억제는 억제된 유전자 프로모터에서 히스톤 H3-아세틸화의 손실 및 H3 라이신 9 트라이메틸화(H3K9me3)의 증가와 연관이 있다. dCas 융합 단백질을 포함하는 KRAB 도메인은, 예를 들어 WO 2017/180915, WO 2014/197748, US 2019/0127713, WO 2013/176772, 문헌[Urrutia R. et al. Genome Biol. 4, 231 (2003), Groner A. C. et al. PLoS Genet. 6, e1000869 (2010)]에 기재되어 있다. 일부 구현예에서, 효과기 도메인은 적어도 하나의 KRAB 도메인 또는 이의 변이체를 포함한다.

[0353] [0345] 일부 구현예에서 KRAB 도메인은 서열번호 590에 명시되어 있다. 일부 구현예에서, 효과기 도메인은 서열번호 590에 제시된 서열 또는 이의 일부를 포함하거나, 앞서 언급된 서열과 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다.

[0354] [0346] 일부 구현예에서 KRAB 도메인은 서열번호 669에 명시되어 있다. 일부 구현예에서, 효과기 도메인은 서열번호 669에 제시된 서열 또는 이의 일부를 포함하거나, 앞서 언급된 서열과 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다.

[0355] [0347] 일부 구현예에서, 효과기 도메인은 적어도 하나의 ERF 억제자 도메인 또는 이의 변이체를 포함한다. ETS2 억제자 도메인[ETS2 repressor factor, ERF]은 보존된 ets-DNA 결합 도메인을 포함하는 강력한 전사 억제자이고 단백질의 카복실 말단에서 별개의 도메인을 통해 전사를 억제한다. dCas 융합 단백질을 포함하는 ERF 억제자 도메인은, 예를 들어 WO 2017180915, WO 2014197748, WO 2013176772, 문헌[Mavrothalassitis, G., Ghysdael, J. Proteins of the ETS family with transcriptional repressor activity. Oncogene 19, 6524-6532 (2000)]에 기재되어 있다. 일부 구현예에서, 효과기 도메인은 적어도 하나의 ERF 억제자 도메인 또는 이의 변이체를 포함한다. 예시적인 ERF 억제자 도메인은 서열번호 600에 기재된다. 일부 구현예에서, 효과기 도메인은 서열번호 600에 기재된 서열, 또는 이의 부분, 또는 전술한 것 중 임의의 것과 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99% 서열 동일성을 갖는 아미노산 서열을 포함한다.

[0356] [0348] 일부 구현예에서, 효과기 도메인은 적어도 하나의 MXI1 도메인 또는 이의 변이체를 포함한다. MXI1 도메인은 myc-연관 인자 x[myc-associated factor x, MAX]에 대한 결합을 위해 경쟁함으로써 myc 전사 활성을 길항함으로써 기능한다. dCas 융합 단백질을 포함하는 MXI1 도메인은, 예를 들어 WO 2017180915, WO 2014197748, US 20190127713에 기재되어 있다. 일부 구현예에서, 효과기 도메인은 적어도 하나의 MXI1 도메인 또는 이의 변이체를 포함한다. 예시적인 MXI1 도메인은 서열번호 601에 기재된다. 일부 구현예에서, 효과기 도메인은 서열번호 601에 기재된 서열, 또는 이의 부분, 또는 전술한 것 중 임의의 것과 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99% 서열 동일성을 갖는 아미노산 서열을 포함한다.

[0357] [0349] 일부 구현예에서, 효과기 도메인은 적어도 하나의 SID4X 도메인 또는 이의 변이체를 포함한다. mSin3 상호작용 도메인[mSin3 interacting domain, SID]은 상이한 전사 억제자 단백질에 존재한다. 이는 전사 억제자 단백질, 예컨대 mSin3 A 보조억제자에 부착된 전사 억제자 도메인인 mSin3의 쌍을 이루는 양친매성 알파 나선 2[paired amphipathic alpha-helix 2, PAH2] 도메인과 상호작용한다.dCas9 분자는 네 개의 연결된 mSin3 상호작용 도메인(SID4X)에 융합될 수 있다. dCas 융합 단백질을 포함하는 SID 도메인은, 예를 들어 WO 2017180915, WO 2014197748, WO 2014093655에 기재되어 있다. 일부 구현예에서, 효과기 도메인은 적어도 하나의 SID 도메인 또는 이의 변이체를 포함한다. 예시적인 SID 도메인은 서열번호 602에 기재된다. 일부 구현예에서, 효과기 도메인은 서열번호 602에 기재된 서열, 또는 이의 부분, 또는 전술한 것 중 임의의 것과 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99% 서열 동일성을 갖는 아미노산 서열을 포함한다.

[0358] [0350] 일부 구현예에서, 효과기 도메인은 적어도 하나의 MAD 도메인 또는 이의 변이체를 포함한다. MAD 패밀리 단백질인 Mad1, Mxi1, Mad3 및 Mad4는 기본 나선-루프-나선-지퍼 클래스에 속하고 억제 활성에 필요한 보존된 N 말단 영역(Sin3 상호작용 도메인[SID]이라고 지칭함)을 포함한다. dCas 융합 단백질을 포함하는 MAD-SID 도메인은, 예를 들어 WO 2017180915, WO 2014197748, WO 2013176772에 기재되어 있다. 일부 구현예에서, 효과기 도메

인은 적어도 하나의 MAD-SID 도메인 또는 이의 변이체를 포함한다. 예시적인 MAD-SID 도메인은 서열번호 603에 기재된다. 일부 구현예에서, 효과기 도메인은 서열번호 603에 기재된 서열, 또는 이의 부분, 또는 전술한 것 중 임의의 것과 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99% 서열 동일성을 갖는 아미노산 서열을 포함한다.

[0359] [0351] 일부 구현예에서, 효과기 도메인은 DNMT3에서 유래하거나 또는 DNA 메틸트랜스퍼라제 활성을 갖는 이의 부분 또는 기능적 활성 변이체이다. DNMT3A 및 DNMT3B는 드노보 메틸화를 촉매하는 두 개의 DNA 메틸트랜스퍼라제고, 이는 부위에 따라 전사 억제와 연관될 수 있다. DNMT, 예컨대 DNMT3은 보편적인 메틸 공여자인 S-아데노실-L-메싸이오닌(S-adenosyl-L-methionine, SAM)에서 사이토신 잔기의 5번 위치로의 메틸기의 전달을 매개한다. 일부 측면에서, 이들 DNMT3 DNA 메틸트랜스퍼라제는 메틸화된 5-메틸사이토신으로의 사이토신 염기의 드노보 메틸화를 유도한다. dCas 융합 단백질을 포함하는 DNMT3은, 예를 들어 US20190127713, 문헌[Liu, X. S. et al. Cell 167, 233-247.e17 (2016), Lei, Y. et al. Nat. Commun. 8, 16026 (2017)]에 기재되어 있다. DNMT3 단백질, 예컨대 DNMT3A 및 DNMT3B는 조절 활성 및 표적화에 자연적으로 관여하는 N-말단 부분, 및 MTase C5-유형 도메인이라고 칭하는 C-말단 촉매 도메인을 함유한다. 일부 구현예에서, 본원에 제공된 구현예에서의 효과기 도메인은 촉매적 활성 C-말단 도메인을 함유하는 DNMT3A 또는 DNMT3B의 촉매적 활성 부분을 포함한다. 특히, DNMT3a 및 DNMT3b의 단리된 촉매 도메인은 촉매적 활성이다(예를 들어, 문헌[Gowher and Jeltsch (2002) J. Biol. Chem., 277:20409] 참조).

[0360] [0352] 일부 구현예에서는, DNMT3 도메인은 촉매적으로 활성인 DNMT3A 또는 DNMT3B의 효과기 도메인일 수 있다. 일부 구현예에서, 효과기 도메인은 DNMT3A 또는 DNMT3B의 전체 길이 또는 그 촉매적으로 활성인 부분일 수 있다. 일부 구현예에서, 효과기 도메인은 DNMT3A 또는 DNMT3B의 전체 길이 서열보다 짧은 촉매적으로 활성인 부분이다. 일부 구현예에서, 촉매적으로 활성인 부분은 DNA 메틸트랜스퍼라제 활성을 부여하는 아미노산의 연속된 서열이며, 예를 들어 시토신 염기의 메틸화를 매개하여 메틸화된 5-메틸시토신으로 메틸화하는 것과 같다. 일부 구현예에서, 아미노산의 연속된 서열은 길이가 280개 아미노산에서 330개 아미노산인 DNMT3A 또는 DNMT3B와 같은 DNMT3 단백질의 연속된 C-말단 부분이다. 일부 구현예에서, 연속 부분은 길이가 280개 아미노산, 290개 아미노산, 300개 아미노산, 310개 아미노산, 320개 아미노산 또는 330개 아미노산이거나, 상기 중 임의의 값 사이의 길이이다. 일부 구현예에서, DNMT3와 같은 DNMT의 촉매 활성 부분은 SAM-의존성 MTase C5-유형 도메인을 포함한다. 일부 구현예에서, DNMT3A 또는 DNMT3B의 도메인과 같은 DNMT3 도메인은 인간 기원이다.

[0361] [0353] 일부 구현예에서는, 효과기 도메인은 DNMT3A 또는 그 촉매 활성 부분 또는 변이체에서 유래한다. 대표적인 DNMT3A 도메인은 서열번호 604에 제시되어 있거나, 그 촉매 활성 부분이거나, 서열번호 604 또는 DNA 메틸트랜스퍼라제 활성을 나타내는 그 촉매 활성 부분과 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99%의 서열 동일성을 갖는 아미노산 서열이다. 일부 구현예에서, 촉매적으로 활성인 부분은 SAM-의존성 MTase C5-유형 도메인을 포함하는 서열번호 604의 아미노산의 연속된 부분이다(예컨대 서열번호 604의 아미노산 634-912에 해당). 일부 구현예에서, 서열번호 604의 아미노산의 연속된 서열은 적어도 250개 아미노산, 275개 아미노산, 300개 아미노산 또는 325개 아미노산, 또는 상기 중 임의의 값 사이의 임의의 값을 포함한다. 일부 구현예에서, 아미노산의 연속된 서열은 아미노산 634-912를 포함하고 길이가 280개 아미노산에서 330개 아미노산인 서열번호 604의 연속된 부분이다. 일부 구현예에서, 연속된 부분은 280개 아미노산, 290개 아미노산, 300개 아미노산, 310개 아미노산, 320개 아미노산 또는 길이는 330개 아미노산이거나, 앞서 언급된 것 사이의 임의 값의 길이이다.

[0362] [0354] 일부 구현예에서, DNMT3A 도메인은 서열번호 661에 제시되거나, 이의 촉매 활성 부분이거나, 서열번호 661 또는 DNA 메틸트랜스퍼라제 활성을 나타내는 이의 촉매 활성 부분과 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열이다. 일부 구현예에서, DNMT3A 도메인은 서열번호 661에 제시된다.

[0363] [0355] 일부 구현예에서, DNMT3A 도메인은 서열번호 665에 제시되거나, 이의 촉매 활성 부분이거나, 서열번호 665 또는 DNA 메틸트랜스퍼라제 활성을 나타내는 이의 촉매 활성 부분과 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열이다. 일부 구현예에서, DNMT3A 도메인은 서열번호 665에 제시된다.

[0364] [0356] 일부 구현예에서는, 효과기 도메인은 DNMT3B 또는 DNA 메틸트랜스퍼라제 활성을 나타내는 촉매 활성 부분 또는 이의 변이체에서 유래한다. 예시적인 DNMT3B 도메인은 서열번호 605에 제시되어 있거나, 그 촉매 활성 부분이거나, 서열번호 605 또는 DNA 메틸트랜스퍼라제 활성을 나타내는 그 촉매 활성 부분과 적어도 90%, 91%,

92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99%의 서열 동일성을 갖는 아미노산 서열이다. 일부 구현예에서, 촉매적 활성 부분은 SAM-의존성 MTase C5-유형 도메인을 포함하는 서열번호 605의 아미노산의 인접 부분이다(예를 들어, 서열번호 605의 아미노산 575-853에 상응함). 일부 구현예에서, 서열번호 605의 아미노산의 인접 서열은 적어도 250개 아미노산, 275개 아미노산, 300개 아미노산 또는 325개 아미노산, 또는 전술한 것 중 임의의 것 사이의 임의의 값을 포함한다. 일부 구현예에서, 아미노산의 인접 서열은 아미노산 575-853을 포함하는 서열번호 605의 인접 부분이고, 길이가 280개 아미노산 내지 330개 아미노산이다. 일부 구현예에서, 인접 부분은 길이가 280개 아미노산, 290개 아미노산, 300개 아미노산, 310개 아미노산, 320개 아미노산, 또는 330개 아미노산이거나 길이가 전술한 것 중 임의의 것 사이의 임의의 값이다.

[0365] [0357] 메틸트랜스퍼라제[methyltransferase, MTase] 활성을 평가하거나 모니터링하는 임의의 다양한 분석이 알려져 있다. 일부 구현예에서, DNA 메틸트랜스퍼라제 활성을 평가하기 위한 예시적인 검정은 방사성 DNA MTase 검정, 비색 DNA MTase 활성 검정, 형광 DNA MTase 활성 검정, 화학발광/생물발광 DNA MTase 활성 검정, 전기화학적 DNA MTase 활성 검정, 및 전기발생 화학발광[elctrogenerated chemiluminescence, ECL] DNA MTase 활성 검정이 포함되지만 이에 제한되지 않는다. 예시적인 검정은 문헌[Poh et al. Theranostics, 2016, 6:369-391; Li et al., Methods Appl. Fluoresc., 2017, 5:012002; Deng et al., Anal Chem., 2014, 86:2117-23; 및 Ma et al. J Mater Chem B., 2020, 8:3488-3501]에 기재되어 있다.

[0366] [0358] 일부 구현예에서는, 효과기 도메인에는 DNMT3L 또는 DNMT3L의 일부 또는 변이체 또는 그 일부가 포함된다. DNA(사이토신-5)-메틸트랜스퍼라제 3-유사[DNA(cytosine-5)-methyltransferase 3-like, DNMT3L]은 문맥에 따라 DNA 메틸화를 촉진하거나 억제할 수 있는 DNA 메틸트랜스퍼라제의 촉매적 비활성 조절 인자이다. DNMT3L은 DNMT3A 및 DNMT3B의 기능에 필수적이고, DNMT3L은 DNMT3A 및 DNMT3B와 상호작용하고 이들의 촉매 활성을 향상시킨다. 예를 들어, DNMT3L은 DNMT3A 또는 DNMT3B의 촉매 도메인과 상호작용하여 이종이량체를 형성하고, 이는 DNMT3L이 메틸화되지 않은 히스톤 꼬리에 결합하고 DNA 메틸트랜스퍼라제를 활성화하는 이중 기능을 갖는다는 것을 입증한다. 일부 구현예에서, 본원에서의 목적을 위한 DNMT3L의 부분 또는 변이체에 대한 언급은 DNMT3A 또는 DNMT3B의 촉매 도메인과 상호작용하고 DNMT3A 또는 DNMT3B의 DNA 메틸트랜스퍼라제 활성을 자극하거나 촉진할 수 있는 DNMT3L의 충분한 C-말단 서열 부분을 지칭한다(예를 들어, 문헌[Jia et al. Nature, 2007, 449:248-251; Gowher et al. J. Biol. Chem., 2005, 280: 13341-13348] 참조). 일부 구현예에서, DNMT3L 또는 이의 부분은 동물 기원이다. 일부 구현예에서, DNMT3L에서의 도메인은 쥐 기원이다. 일부 구현예에서, DNMT3L에서의 도메인은 인간 기원이다.

[0367] [0359] 일부 구현예에서, DNMT3L 도메인은 DNMT3A의 촉매 도메인과 상호작용하여 이종이량체를 형성하여 보다 활성인 DNA 메틸트랜스퍼라제를 제공하는 DNMT3L, 또는 이의 C-말단 부분 또는 변이체이다. 일부 구현예에서, 효과기 도메인은 DNMT3A 도메인 및 DNMT3L 도메인 (DNMT3A/3L)의 융합 도메인이다.

[0368] [0360] 일부 구현예에서, DNMT3L 도메인은 DNMT3B의 촉매 도메인과 상호작용하여 이종이량체를 형성하여 보다 활성인 DNA 메틸트랜스퍼라제를 제공하는 DNMT3L, 또는 이의 C-말단 부분 또는 변이체이다. 일부 구현예에서, 효과기 도메인은 DNMT3B 도메인 및 DNMT3L 도메인 (DNMT3B/3L)의 융합 도메인이다.

[0369] [0361] 일부 구현예에서, DNMT3L 도메인은 N-말단 시스테인-풍부 ATRX-Dnmt3-Dnmt3L(ADD) 도메인(예를 들어, 서열번호 607의 잔기 41-73 또는 서열번호 668에 기재된 서열의 75-207에 상응함)을 포함하지 않는, 전장 DNMT3L의 인접 C-말단 부분으로 이루어진 DNMT3L의 C-말단 부분이다. 일부 구현예에서, DNMT3L 도메인은 길이가 220개 아미노산 미만, 예컨대, 길이가 100 내지 215개 아미노산, 예컨대, 약 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210 또는 215개 아미노산, 또는 길이가 전술한 것 중 임의의 것 중 임의의 값 사이인 DNMT3L의 연속 C-말단 부분이다. 일부 구현예에서, DNMT3L 도메인은 205, 206, 207, 208, 209, 210, 211, 212, 213, 214 또는 215개의 아미노산 길이인 DNMT3L의 인접 C-말단 부분이다.

[0370] [0362] 일부 구현예에서, DNMT3L 도메인은 서열번호 668에 기재되어 있거나, 이의 촉매적 활성 부분이거나, 서열번호 668 또는 이의 부분에 대해 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99% 서열 동일성을 갖는 아미노산 서열이다. 일부 구현예에서, DNMT3L 도메인은 N-말단 시스테인-풍부 ATRX-Dnmt3-Dnmt3L(ADD) 도메인(서열번호 668에 기재된 서열의 잔기 75-207에 상응함)을 포함하지 않는, 서열번호 668에 기재된 전장 DNMT3L의 인접 C-말단 부분이다. 일부 구현예에서, DNMT3L 도메인은 길이가 220개 아미노산 미만, 예컨대, 길이가 100 내지 215개 아미노산, 예컨대, 약 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210 또는 215개 아미노산, 또는 길이가 전술한 것 중 임의의 것 중 임의의 값 사이인, 서열번호 668에 기재된 전장 DNMT3L의 연속 C-말단 부분이다. 일부 구현예에서, DNMT3L 도메인은 길이가 205, 206, 207, 208, 209,

210, 211, 212, 213, 214 또는 215개의 아미노산인, 서열번호 668에 기재된 전장 DNMT3L의 인접 C-말단 부분이다.

[0371] [0363] 일부 구현예에서, DNMT3L 도메인은 서열번호 664에 기재되어 있거나, 이의 부분이거나, 서열번호 664과 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99% 서열 동일성을 갖는 아미노산 서열이다. 일부 구현예에서, DNMT3L 도메인은 서열번호 664에 기재된다. 일부 구현예에서, DNMT3L 도메인은 서열번호 664에 기재된 바와 같은 N-말단 메싸이오닌을 함유하지 않는다.

[0372] [0364] 일부 구현예에서, DNMT3L 도메인은 인간 또는 인간화된 DNMT3L이다. 인간의 상응하는 서열은 마우스에서 유래한 Dnmt3L과 고도로 상동성이고, 쥐 서열과 적어도 90%의 서열 동일성을 갖는다. DNMT3L 도메인, 예컨대 쥐 DNMT3L의 도메인의 비인간 서열을 인간화하는 것은 당업자의 수준 내에 있다. 일부 구현예에서, 효과기 도메인은 서열번호 668에 기재된 쥐 DMT3L의 인간화 변이체인 DNMT3L 도메인 또는 DNMT3A 또는 DNMT3A와 상호작용할 수 있는 이의 부분을 포함한다. 일부 구현예에서, 효과기 도메인은 서열번호 664에 기재된 DNMT3L의 쥐 C-말단 부분의 인간화된 변이체인 DNMT3L 도메인을 포함한다.

[0373] [0365] 일부 구현예에서, 인간 기원 DNMT3L 도메인은 서열번호 607에 기재되어 있거나 이의 부분이거나, 서열번호 607 또는 이의 부분에 대해 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99% 서열 동일성을 갖는 아미노산 서열이다. 일부 구현예에서, DNMT3L 도메인은 N-말단 시스테인-풍부 ATRX-Dnmt3-Dnmt3L(ADD) 도메인(서열번호 607에 기재된 서열의 잔기 41-73에 상응함)을 포함하지 않는, 서열번호 607에 기재된 전장 DNMT3L의 인접 C-말단 부분이다. 일부 구현예에서, DNMT3L 도메인은 길이가 220개 아미노산 미만, 예컨대, 길이가 100 내지 215개 아미노산, 예컨대, 약 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210 또는 215개 아미노산, 또는 길이가 전술한 것 중 임의의 것 중 임의의 값 사이인, 서열번호 607에 기재된 전장 DNMT3L의 연속 C-말단 부분이다. 일부 구현예에서, DNMT3L 도메인은 길이가 205, 206, 207, 208, 209, 210, 211, 212, 213, 214 또는 215개의 아미노산인, 서열번호 197에 기재된 전장 DNMT3L의 인접 C-말단 부분이다.

[0374] [0366] 일부 구현예에서, DNMT3L 도메인은 서열번호 666에 기재된 서열을 포함하거나, 그 일부이거나, 서열번호 666과 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열이다. 일부 구현예에서, DNMT3L 도메인은 서열번호 666에 제시된다. 일부 구현예에서, DNMT3L 도메인은 N-말단 메티오닌을 함유한다.

[0375] [0367] 일부 구현예에서, 효과기 도메인은 DNMT3A와 DNMT3L(DNMT3A/L)의 융합을 포함한다. 융합 단백질은 상기 설명된 바와 같이 임의의 DNMT3A 및 DNMT3L 도메인을 포함한다. 일부 구현예에서, 융합 단백질은 서열번호 661에 제시된 DNMT3A 도메인과 서열번호 668에 제시된 DNMT3L 도메인을 임의의 순서로 배열하여 포함한다. 일부 구현예에서, 융합 단백질은 서열번호 661에 제시된 DNMT3A 도메인과 서열번호 664에 제시된 DNMT3L 도메인을 임의의 순서로 배열하여 포함한다. 일부 구현예에서, 융합 단백질은 서열번호 661에 제시된 DNMT3A 도메인과 서열번호 666에 제시된 DNMT3L 도메인을 임의의 순서로 배열하여 포함한다. 일부 구현예에서, 융합 단백질은 서열번호 665에 제시된 DNMT3A 도메인과 서열번호 668에 제시된 DNMT3L 도메인을 임의의 순서로 배열하여 포함한다. 일부 구현예에서, 융합 단백질은 서열번호 665에 제시된 DNMT3A 도메인과 서열번호 664에 제시된 DNMT3L 도메인을 임의의 순서로 배열하여 포함한다. 일부 구현예에서, 융합 단백질은 서열번호 665에 제시된 DNMT3A 도메인과 서열번호 666에 제시된 DNMT3L 도메인을 임의의 순서로 배열하여 포함한다. 일부 구현예에서, 제공된 융합 단백질에 존재하는 DNMT3A 및 DNMT3L 도메인은 DNA-결합 도메인, 다른 효과기 도메인 또는 링커와 같은 개입 서열에 의해 융합 단백질에서 서로 분리된다. 일부 구현예에서, 도메인은 서로 직접 연결되거나 펩타이드 링커와 같은 링커를 통해 연결된다. 일부 구현예에서, DNMT3A 및 DNMT3L 도메인은 DNMT3A 도메인과 DNMT3L 도메인을 연결하는 링커를 통해 융합 도메인으로 연결된다. 예시적인 링커가 여기에 설명되어 있다. 일부 구현예에서, 링커는 서열번호 667에 제시된 링커이다.

[0376] [0368] 예시적 DNMT3A/L 융합 도메인은 서열번호 651에 제시되어 있다. 일부 구현예에서, 효과기 도메인은 서열번호 651에 제시된 서열을 포함하거나, 서열번호 651과 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함하며 DNA 메틸트랜스퍼라제 활성을 나타낸다.

[0377] [0369] 일부 구현예에서, 효과기 도메인은 LSD1 도메인을 포함할 수 있다. (라이신 특이적 히스톤 탈메틸효소 1A로도 알려진) LSD1은 히스톤 H3의 라이신 잔기를 탈메틸화하여 맥락에 따라 보조활성인자 또는 보조억제자로서 작용할 수 있는 히스톤 탈메틸효소이다. dCas 융합 단백질을 포함한 LSD1은, 예를 들어, WO 2013/176772, WO 2014/152432, 및 문헌[Kearns, N. A. et al. Nat. Methods. 12(5):401-403 (2015)]에 기재되어 있다. 예시적인 LSD1 폴리펩타이드는 서열번호 606에 기재된다. 일부 구현예에서, 효과기 도메인은 서열번호 606에 기재된

서열, 또는 이의 부분, 또는 전술한 것 중 임의의 것과 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99% 서열 동일성을 갖는 아미노산 서열을 포함한다.

[0378] [0370] 일부 구현예에서, 효과기 도메인은 EZH2 도메인을 포함할 수 있다. (히스톤-라이신 N-메틸트랜스퍼라제 EZH2로도 알려진) EZH2는 PRC2/EED-EZH2 복합체의 촉매 소단위이고, 이는 일부 측면에서 히스톤 H3의 'Lys-9'(H3K9me) 및 'Lys-27'(H3K27me)을 메틸화하여 영향을 받는 표적 유전자의 전사 억제를 유도한다. dCas 융합 단백질을 포함한 EZH2는, 예를 들어 문헌[O'Geen, H. et al., Epigenetics Chromatin. 12(1):26 (2019)]에 기재되어 있다. 예시적인 EZH2 폴리펩타이드는 서열번호 608에 기재된다. 일부 구현예에서, 효과기 도메인은 서열번호 608에 기재된 서열, 또는 이의 부분, 또는 전술한 것 중 임의의 것과 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99% 서열 동일성을 갖는 아미노산 서열을 포함한다.

[0379] [0371] 일부 구현예에서, 효과기 도메인은 SunTag 도메인을 포함할 수 있다. SunTag는 반복 펩타이드 어레이로, 이는 반복 펩타이드에 결합하는 항체 융합 단백질의 여러 사본을 동원할 수 있다. 항체 융합 단백질은 표적 유전자의 전사를 감소시키기 위해 추가적인 효과기 도메인, 예컨대 전사 억제 도메인(예를 들어, KRAB)을 포함할 수 있다. 유전자 조절을 위한 dCas 융합 단백질을 포함하는 SunTag는, 예를 들어 WO 2016/011070 및 문헌[Tanenbaum, M. et al. Cell. 159(3):635-646 (2014)]에 기재되어 있다. 예시적인 SunTag 효과기 도메인은 아미노산 서열 GGSGG(서열번호 642)를 갖는 링커에 의해 분리된 아미노산 서열 LLPKNYHLENEVARLKKLVGER(서열번호 641)을 갖는 반복 GCN4 펩타이드를 포함한다. 일부 구현예에서, 링커는 서열번호 641에 기재된 서열의 적어도 하나의 사본, 또는 이의 부분, 또는 전술한 것 중 임의의 것과 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99% 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, SunTag 효과기 도메인은 KRAB를 포함하고 GCN4 펩타이드에 결합하는 항체 융합 단백질을 동원한다.

[0380] [0372] 일부 구현예에서, 후성적 변형 DNA-표적화 시스템은 DNA 메틸트랜스퍼라제(예컨대 DNMT3A 또는 DNMT3B)를 포함하는 적어도 하나의 효과기 도메인을 포함한다. 일부 구현예에서, 후성적 변형 DNA-표적화 시스템은 억제자 도메인(예컨대 KRAB)을 포함하는 적어도 하나의 효과기 도메인 을 포함한다. 일부 구현예에서, 적어도 하나의 효과기 도메인을 포함하는 후성적 변형 DNA-표적화 시스템은 DNA 메틸트랜스퍼라제와 히스톤 메틸트랜스퍼라제(예컨대 DNMT3A-KRAB 또는 DNMT3A/L-KRAB)를 포함한다.

[0381] **F. 융합 단백질**

[0382] [0373] 일부 측면에서, 본원에 제공된 DNA-표적화 시스템은 융합 단백질을 포함한다. 일부 구현예에서, 본원에 제공된 DNA-표적화 시스템은 적어도 하나의 DNA-표적화 모듈을 포함하는데, 여기서 적어도 하나의 DNA-표적화 모듈 각각은 다음을 포함하는 융합 단백질을 포함한다: (a) B형 간염 바이러스 DNA 서열의 표적 부위, 예를 들어 유전자 또는 이의 조절 요소, 예를 들어 본 명세서에 설명된 것과 같은 표적 부위를 표적화하기 위한 DNA-결합 도메인, 및 (b) 적어도 하나의 전사 억제자 효과기 도메인. 일부 측면에서, 융합 단백질은 본 명세서에 설명된 DNA-결합 도메인 중 적어도 하나와 본 명세서에 설명된 효과기 도메인 중 적어도 하나를 포함한다. 예를 들어, 일부 구현예에서, 융합 단백질은 섹션 I.B.에 설명된 것과 같은 CRISPR-Cas 또는 이의 변이체와 섹션 I.E.에 설명된 것과 같은 적어도 하나의 전사 억제자 효과기 도메인을 포함한다. 일부 측면에서, 융합 단백질은 유전자 또는 이의 조절 요소의 표적 부위를 표적화하고 유전자의 전사를 감소시키거나 억제한다.

[0383] [0374] 일부 구현예에서, 융합 단백질의 DNA-결합 도메인 및 효과기 도메인은 이종성이고, 즉, 도메인은 상이한 종에서 유래하거나 도메인 중 적어도 하나는 자연에서 발견되지 않는다. 일부 측면에서, 융합 단백질은 조작된 융합 단백질이고, 즉 융합 단백질은 자연에서 발견되지 않는다.

[0384] [0375] 일부 구현예에서, 적어도 하나의 효과기 도메인은 DNA-결합 도메인 또는 이의 성분의 N-말단, C-말단, 또는 N-말단과 C-말단 둘 다에 융합된다. 일부 구현예에서, 적어도 하나의 효과기 도메인은 DNA-결합 도메인에 직접, 또는 임의의 개입 아미노산 서열, 예컨대 링커 서열 또는 핵 국소화 서열[NLS]을 통해 융합될 수 있다.

[0385] [0376] 일부 구현예에서, 제공된 DNA-결합 시스템 또는 이의 DNA-표적화 모듈의 융합 단백질은 N-말단에서 C-말단 순서로 전사 억제자 도메인 및 DNA-결합 도메인을 포함한다. 일부 구현예에서, 제공된 DNA-결합 시스템 또는 이의 DNA-표적화 모듈의 융합 단백질은 N-말단에서 C-말단 순서로 DNA-결합 도메인 및 전사 억제자 효과기 도메인을 포함한다.

[0386] [0377] 일부 구현예에서, 융합 단백질의 적어도 하나의 효과기 도메인은 하나보다 많은 효과기 도메인을 포함한다. 일부 구현예에서, 융합 단백질은 2개, 3개 또는 4개의 효과기 도메인을 포함한다. 일부 구현예에서, 융합 단백질의 효과기 도메인 중 적어도 두 개는 상이하다. 일부 구현예에서, 융합 단백질의 효과기 도메인 각각은

상이하다. 일부 구현예에서, 적어도 하나의 효과기 도메인은 두 개의 효과기 도메인이 상이한 두 개의 효과기 도메인을 포함한다. 일부 구현예에서, 효과기 도메인 및 DNA-결합 도메인은 임의의 순서로 배열될 수 있다.

[0387] [0378] 일부 구현예에서, 융합 단백질의 적어도 하나의 효과기 도메인은 두 개의 상이한 효과기 도메인을 포함한다. 두 개의 상이한 효과기 도메인 및 DNA-결합 도메인은 임의의 순서로 배열될 수 있다. 일부 구현예에서, 각각의 효과기 도메인은 DNA-결합 도메인의 N-말단이되, 제1 효과기 도메인은 제2 효과기 도메인의 N-말단에 융합되고 제2 효과기 도메인은 DNA-결합 도메인의 N-말단에 융합된다. 일부 구현예에서, 제공된 DNA-결합 시스템 또는 이의 DNA-표적화 모듈의 융합 단백질은 N-말단 내지 C-말단 순서로 제1 전사 억제자 효과기 도메인, 제2 전사 억제자 효과기 도메인 및 DNA 결합 도메인을 포함한다. 일부 구현예에서, 효과기 도메인 각각은 DNA-결합 도메인의 C-말단이되, 제1 효과기 도메인은 DNA-결합 도메인의 C-말단에 융합되고 제2 효과기 도메인은 제1 효과기 도메인의 C-말단에 융합된다. 일부 구현예에서, 제공된 DNA-결합 시스템 또는 이의 DNA-표적화 모듈의 융합 단백질은 N-말단 내지 C-말단 순서로 DNA-결합 도메인, 제1 전사 억제자 효과기 도메인 및 제2 전사 억제자 효과기 도메인을 포함한다. 일부 구현예에서, DNA-결합 도메인은 효과기 도메인 사이에 있되, 하나의 효과기 도메인은 DNA-결합 도메인의 N-말단에 융합되고, 다른 효과기 도메인은 DNA-결합 도메인의 C-말단에 융합된다. 일부 구현예에서, 제공된 DNA-결합 시스템 또는 이의 DNA-표적화 모듈의 융합 단백질은 N-말단 내지 C-말단 순서로 제1 전사 억제자 효과기 도메인, DNA 결합 도메인 및 제2 전사 억제자 효과기 도메인을 포함한다. 일부 구현예에서, 하나 이상의 성분은 서로 직접, 또는 임의의 개입 아미노산 서열을 통해, 예컨대 링커 서열 또는 핵 국소화 서열[NLS]을 통해 융합될 수 있다.

[0388] [0379] 일부 구현예에서, 융합 단백질은 하나 이상의 링커를 포함한다. 일부 구현예에서, 하나 이상의 링커는 DNA-결합 도메인 또는 그 구성 요소를 융합 단백질의 적어도 하나의 효과기 도메인에 연결한다. 링커는 융합 단백질의 폴리펩타이드 서열의 어느 곳에나 포함될 수 있으며, 예를 들어, 효과기 도메인과 DNA-결합 도메인 또는 그 구성 요소 사이에 포함될 수 있다. 일부 구현예에서, 하나 이상의 링커는 융합 단백질의 두 효과기 도메인을 연결한다. 효과기 도메인의 수에 따라, 융합 단백질은 1, 2, 3, 4 또는 그 이상의 링커를 포함할 수 있다. 일부 구현예에서, 링커는 동일할 수도 있고 다를 수도 있다.

[0389] [0380] 링커는 임의의 길이일 수 있고, 융합 단백질 내 성분의 이동성을 촉진하거나 제한하도록 설계될 수 있다. 일부 구현예에서, 링커는 펩타이드 링커이다. 링커는 약 2개 내지 약 100개, 약 5개 내지 약 80개, 약 10개 내지 약 60개, 또는 약 20개 내지 약 50개 아미노산 중 임의의 아미노산 서열을 포함할 수 있다. 링커는 적어도 약 2, 3, 4, 5, 10, 15, 20, 25, 또는 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80 또는 85개 아미노산의 아미노산 서열을 포함할 수 있다. 링커는 약 100, 90, 80, 70, 60, 50, 또는 40개 미만 아미노산의 아미노산 서열을 포함할 수 있다. 당업자는 두 개의 도메인 연결을 위한 적절한 링커를 쉽게 선택할 수 있다. 일부 구현예에서, 펩타이드 링커는 유연한 링커이다. 가요성 링커는 일반적으로 글리신, 세린 또는 트레오닌과 같은 작은 비극성 또는 극성 잔기로 구성된다. 일부 구현예에서, 링커는 $Gly_4Ser_{(n)}$ 링커이되, n은 1 내지 10의 정수이다.

링커는 길이가 2 내지 20개의 아미노산인 아미노산 서열의 순차적 또는 일렬 반복을 포함할 수 있다. 링커는 아미노산 글리신(G), 세린(S), 및/또는 알라닌(A)이 풍부할 수 있다. 링커는 예를 들어, GS 링커를 포함할 수 있다. 예시적인 GS 링커는 서열 GGGGS(서열번호 609) 또는 식 (GGGGS)n으로 표시되는 바와 같이 서열의 반복을 포함할 수 있되, n은 GGGGS 서열이 반복되는 횟수(예를 들어, 1 내지 10회)를 나타내는 정수이다. 링커 서열이 반복되는 횟수는 링커 길이를 최적화하고 기능성 도메인의 적절한 분리를 달성하기 위해 조정될 수 있다. 링커의 다른 예는, 예를 들어, GGGGG(서열번호 610), GGAGG(서열번호 611), GGGGSSS(서열번호 231), 또는 GGGGAAA(서열번호 612)를 포함할 수 있다.

[0390] [0381] 일부 구현예에서, 인공 링커 서열이 사용될 수 있다. 일부 구현예에서, 링커는 EASGSGRASPGIPGSTR(서열번호 672)이다. 일부 구현예에서, 링커는 GIHGVPAA(서열번호 674)이다. 일부 구현예에서, 링커는 SSGNSNANSRGPSFSSGLVPLSLRGSH(서열번호 667)이다. 일부 구현예에서, 링커는 KRPAATKKAGQAKKKKASDAKSLTAWS(서열번호 677)이다.

[0391] [0382] 일부 구현예에서, 링커는 XTEN 링커이다. 일부 측면에서, XTEN 링커는 소수성 아미노산 잔기가 결여된 재조합 폴리펩타이드(예를 들어, 비구조화된 재조합 펩타이드)이다. 예시적인 XTEN 링커는 예를 들어, 문헌 [Schellenberger et al., Nature Biotechnology 27, 1186-1190 (2009)] 또는 WO 2021/247570에 기재된다. 일부 구현예에서, 융합 단백질에 링커를 포함시키면 표적 유전자의 억제가 강화된다. 일부 구현예에서, 링커는 서열번호 643에 기재된 서열, 또는 이의 부분, 또는 서열번호 643과 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99% 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 측면에서, 링커는 서열번호 643에 기

재된 서열, 또는 적어도 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70 또는 75개의 아미노산의 서열번호 643의 인접 부분을 포함한다. 일부 측면에서, 링커는 서열번호 643에 기재된 서열, 또는 적어도 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70 또는 75개의 아미노산의 서열번호 643의 인접 부분을 구성한다. 일부 구현예에서, 링커는 서열번호 643에 기재된 아미노산 서열을 포함한다. 일부 구현예에서, 링커는 서열번호 643에 기재된 아미노산 서열을 구성한다. 일부 구현예에서, 링커는 서열번호 658에 기재된 서열, 또는 이의 부분, 또는 전술한 것 중 임의의 것과 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99% 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 측면에서, 링커는 서열번호 658에 기재된 서열, 또는 적어도 5, 10, 또는 15개의 아미노산의 서열번호 658의 인접 부분을 포함한다. 일부 측면에서, 링커는 서열번호 658에 기재된 서열, 또는 적어도 5, 10 또는 15개의 아미노산의 서열번호 658의 인접 부분을 구성한다. 일부 구현예에서, 링커는 서열번호 658에 기재된 아미노산 서열을 포함한다. 일부 구현예에서, 링커는 서열번호 658에 기재된 아미노산 서열을 구성한다. 일부 구현예에서, 링커는 WO 2021/247570에 설명된 링커를 포함한다. 적절한 링커는 당업계에 알려진 합리적인 기준, 예를 들어 문헌 [Chen et al. Adv. Drug Deliv. Rev. 65(10):1357-1369 (2013)]에 기재된 바와 같은 기준에 기반하여 선택되거나 설계될 수 있다.

[0392] [0383] 일부 구현예에서, DNA-표적화 시스템의 융합 단백질, 또는 이의 DNA-표적화 모듈은 하나 이상의 핵 국소화 신호[nuclear localization signal, NLS]를 포함한다. 일부 구현예에서, 본원에 기재된 융합 단백질은 하나 이상의 핵 국소화 서열[NLS], 예컨대 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10개 이상의 NLS를 포함한다. 하나보다 많은 NLS가 존재하는 경우, 각각은 다른 것들에 독립적으로 선택될 수 있어서, 단일 NLS는 하나보다 많은 사본에 존재하고/하거나 하나 이상의 사본에 존재하는 하나 이상의 다른 NLS와 조합하여 존재할 수 있다. NLS의 비제한적 예는 아미노산 서열 PKKKRKV(서열번호 614)를 갖는 SV40 바이러스 거대 T-항원의 NLS, 뉴클레오플라스민에서의 NLS[예를 들어, 서열 KRPAATKKAGQAKKKK(서열번호 592)을 갖는 뉴클레오플라스민 이분 NLS], 아미노산 서열 PAAKRVKLD(서열번호 235) 또는 RQRRNELKRSP(서열번호 616)를 갖는 c-myc NLS, 서열 NQSSNFGPMKGGNFGGRSSGPYGGGGQYFAKPRNQGGY(서열번호 617)를 갖는 hRNPA1 M9 NLS, 임포틴-알파의 IBB 도메인의 서열 RMRIZFKNKGKDTAELRRRRVEVSVELRKAKKDEQILKRRNV(서열번호 618), 근종 T 단백질의 서열 VSRKRPRP(서열번호 619) 및 PPKKARED(서열번호 620), 인간 p53의 서열 PQPKKKPL(서열번호 621), 마우스 c-abl IV의 서열 SALIKKKKKMAP(서열번호 622), 인플루엔자 바이러스 NS1의 서열 DRLRR(서열번호 623) 및 PKQKKRK(서열번호 624), 간염 바이러스 델타 항원의 서열 RKLKKKIKKL(서열번호 625), 마우스 Mx1 단백질 서열 REKKKFLKRR(서열번호 626), 인간 폴리(ADP-리보스) 폴리머라제의 서열 KRKGDEVDGVDEVAKKKSKK(서열번호 627), 및 스테로이드 호르몬 수용체 (인간) 클루코코르티코이드의 서열 RKCLQAGMNLEARKTKK(서열번호 628)에서 유래하는 NLS 서열을 포함한다. 일반적으로, 하나 이상의 NLS는 진핵 세포의 핵에서 검출 가능한 양으로 융합 단백질의 축적을 유도하기에 강도가 충분하다. 일반적으로, 핵 위치 활성의 강도는 융합 단백질 내 NLS의 수, 사용된 특정 NLS(들), 또는 이들 인자의 조합물에서 유래할 수 있다. 핵 내의 축적의 검출은 임의의 적합한 기술로 수행될 수 있다. 예를 들어, 핵의 위치를 검출하기 위한 수단(예를 들어, DAPI와 같은 핵에 특이적인 염색)과 조합하는 것과 같이, 세포 내의 위치가 시각화될 수 있도록 검출 가능한 표지가 융합 단백질에 융합될 수 있다. 세포 핵은 또한 세포에서 단리될 수 있고, 이의 내용물은 이후 단백질을 검출하기 위한 임의의 적합한 과정, 예컨대 면역조직화학, 웨스턴 블롯, 또는 효소 활성 검정으로 분석될 수 있다. 핵에서의 축적은 또한 대조군 조건(예를 들어, 형질도입되지 않은 세포)에 비해, 예컨대 융합 단백질의 효과에 대한 검정(예를 들어, 융합 단백질을 포함하는 DNA-표적화 시스템으로 형질도입된 세포에서 변경된 유전자 발현 활성에 대한 검정)으로 간접적으로 결정될 수 있다. 일부 구현예에서, NLS는 서열번호 592 (KRPAATKKAGQAKKKK)에 제시된 서열 또는 이의 일부를 포함한다.

[0393] [0384] 일부 구현예에서, NLS는 링커를 통해 DNA-결합 도메인의 N-말단 또는 C-말단에 연결된다. 일부 구현예에서, NLS는 링커를 통해 효과기 도메인의 N-말단 또는 C-말단에 연결된다. 링커는 전술한 임의의 링커일 수 있다. 일부 구현예에서, 링커는 GIHGVPAA(서열번호 673)이다. 일부 구현예에서, NLS 및 링커는 서열 PKKKRKVGIHGVPAA(서열번호 670)를 갖는다.

[0394] [0385] 일부 구현예에서, 융합 단백질의 N-말단 또는 C-말단은 검출 및/또는 정제를 위한 모이어티에 연결될 수 있다. 일부 측면에서, 모이어티는 Flag 태그 DYKDDDDK(서열번호 671), 3xFlag 태그 MDYKDHDGDYKDHDI DYKDDDDK (서열번호 674), HA 태그 YPYDVPDYA(서열번호 675) 또는 His 태그, 예컨대 HHHHHH(서열번호 676)이거나 이를 포함할 수 있다.

[0395] ***1. 분할 융합 단백질***

[0396] [0386] 일부 구현예에서, 융합 단백질은 분할 단백질이고, 즉, 상호작용하거나 자가 조립하여 기능성 융합 단백질을 형성하는 두 개 이상의 개별 폴리펩타이드 도메인을 포함한다. 일부 측면에서, 분할 융합 단백질은 dCas9

및 효과기 도메인을 포함한다. 일부 측면에서, 융합 단백질은 분할 dCas9-효과기 도메인 융합 단백질을 포함한다.

[0397] [0387] 일부 구현예에서, 분할 융합 단백질은 트랜스-스플라이싱 인테인을 포함하는 개별 폴리펩타이드 도메인에서 조립된다. 인테인은 이의 숙주 단백질에서 자가 절제하고 펩타이드 결합과 측면 서열의 결찰을 촉매하는 내부 단백질 요소이다. 일부 구현예에서, 분할 융합 단백질은 N-말단 인테인을 포함하는 제1 폴리펩타이드 및 C-말단 인테인을 포함하는 제2 폴리펩타이드에서 조립된다. 일부 구현예에서, N 말단 인테인은 서열번호 653에 기재된 N 말단 Npu 인테인이다. 일부 구현예에서, C 말단 인테인은 서열번 호 655에 기재된 C 말단 Npu 인테인이다. 일부 구현예에서, N 말단 인테인은 서열번호 652에 기재된 N 말단 Npu 인테인이다. 일부 구현예에서, C 말단 인테인은 서열번호 654에 기재된 C 말단 Npu 인테인이다.

[0398] [0388] 일부 구현예에서, 분할 융합 단백질은 두 개의 폴리펩타이드로부터 조립된 분할 dCas9-효과기 도메인 융합 단백질을 포함한다. 예시적인 구현예에서, 제1 폴리펩타이드는 효과기 도메인 촉매 도메인과 dSpCas9의 N-말단 단편을 포함하고, 그 뒤에 N 말단 Npu 인테인(효과기 도메인-dSpCas9-573N)이 이어지고, 제2 폴리펩타이드는 C 말단 Npu 인테인과 그 뒤에 dSpCas9의 C-말단 단편(dSpCas9-573C; 서열번호 657)이 이어진다. 융합 단백질의 N- 및 C-말단 단편은 서열번호 598(서열번호 599에 제시된 dSpCas9 분자의 잔여물 572Glu에 해당)을 참조하여 dSpCas9 분자의 위치 573Glu에서 분할된다. 일부 측면에서, N-말단 Npu 인테인(서열번호 653) 및 C-말단 Npu 인테인(서열번호 655에 명시됨)은 두 단편을 자가 절제하고 연결하여 세포에서 발현될 때 전체 길이의 dSpCas9-효과기 도메인 융합 단백질을 형성할 수 있다.

[0399] [0389] 일부 구현예에서, 분할 단백질의 폴리펩타이드는 비공유적으로 상호작용하여 비분할 단백질의 활성을 반복하는 복합체를 형성할 수 있다. 예를 들어, 개별 폴리펩타이드로서 발현되는 Cas 효소의 두 개의 도메인은 예를 들어, 문헌 [Wright et al. PNAS 112(10):2984-2989 (2015)]에 기재된 바와 같이 gRNA에 의해 동원되어 gRNA와 복합체를 형성하는 전장 Cas 효소의 활성을 반복하는 삼원 복합체를 형성할 수 있다. 일부 구현예에서, 분할 단백질의 조립은 유도성(예를 들어, 광 유도성, 화학적 유도성, 소분자 유도성)이다.

[0400] [0390] 일부 측면에서, 분할 융합 단백질의 두 개의 폴리펩타이드는 개별 벡터, 예컨대 본원에 기재된 임의의 벡터에서 전달되고/되거나 발현될 수 있다. 일부 구현예에서, 분할 융합 단백질의 두 개의 폴리펩타이드는, 즉, 예를 들어 WO 2017/197238에 기재된 분할 AAV-기반 접근법을 사용하여, 세포에 전달되고/되거나 두 개의 개별 AAV 벡터에서 발현될 수 있다.

[0401] [0391] Cas 단백질 및 이의 융합체를 포함하는 분할 단백질 및 이의 전달의 합리적 설계에 대한 접근법은 예를 들어, WO 2016/114972, WO 2017/197238, 문헌[Zetsche. et al. Nat. Biotechnol. 33(2):139-42 (2015), Wright et al. PNAS 112(10):2984-2989 (2015), Truong. et al. Nucleic Acids Res. 43, 6450-6458 (2015), 및 Fine et al. Sci. Rep. 5, 10777 (2015)]에 기재되어 있다.

[0402] ***2. 예시적인 CRISPR 기반 융합 단백질***

[0403] [0392] 일부 구현예에서, 제공된 DNA-표적화 시스템 또는 그 DNA-표적화 모듈의 융합 단백질은 B형 간염 바이러스 DNA 서열의 표적 부위, 예를 들어 유전자 또는 이의 조절 요소, 및 적어도 하나의 전사 억제자 효과기 도메인을 표적화하는 DNA-결합 도메인으로 구성된다. DNA-결합 도메인 및 전사 억제자 도메인은 상기 설명한 바와 같은 것을 포함한다. 예시적인 융합 단백질은 아래에 추가로 설명된다.

[0404] [0393] 일부 구현예에서, 본원에서 명명된 융합 단백질은 명명된 융합 단백질의 요소를 임의의 구성 또는 순서로 포함한다. 예를 들어, dCas9-KRAB 융합 단백질은 dSpCas9 분자의 N- 또는 C-말단에 융합된 KRAB 도메인을 포함할 수 있다. 또 다른 예에서, dSpCas9-KRAB-DNMT3A/L 융합 단백질은 dSpCas9, KRAB 및 DNMT3A/L을 임의의 순서로 포함할 수 있다. 예를 들어, dSpCas9-KRAB-DNMT3A/L 융합 단백질은 N-말단에서 C-말단까지 DNMT3A/L, dSpCas9 및 KRAB를 포함할 수 있다. 본원에서 명명된 융합 단백질은 추가 요소를 포함할 수 있다. 예를 들어, dSpCas9-KRAB-DNMT3A/L 융합 단백질은 하나 이상의 링커, NLS 서열 또는 다른 서열을 임의의 조합 또는 순서로 포함할 수 있다.

[0405] [0394] 일부 구현예에서, DNA-표적화 시스템 또는 이의 DNA-표적화 모듈의 융합 단백질은 서열번호 599에 제시된 dSpCas9와 서열번호 600에 제시된 ERF 도메인을 함유한다. 일부 구현예에서, DNA-표적화 시스템 또는 이의 DNA-표적화 모듈의 융합 단백질은 서열번호 599에 제시된 dSpCas9와 서열번호 601에 제시된 MXI1 도메인을 함유한다. 일부 구현예에서, DNA-표적화 시스템 또는 이의 DNA-표적화 모듈의 융합 단백질은 서열번호 599에 제시된 dSpCas9와 서열번호 602에 제시된 SID4X 도메인을 함유한다. 일부 구현예에서, DNA-표적화 시스템 또는 이의

DNA-표적화 모듈의 융합 단백질은 서열번호 599에 제시된 dSpCas9와 서열번호 603에 제시된 MAD-SID 도메인을 함유한다. 일부 구현예에서, DNA-표적화 시스템 또는 이의 DNA-표적화 모듈의 융합 단백질은 서열번호 599에 제시된 dSpCas9와 서열번호 604에 제시된 DNMT3A 도메인을 함유한다. 일부 구현예에서, DNA-표적화 시스템 또는 이의 DNA-표적화 모듈의 융합 단백질은 서열번호 599에 제시된 dSpCas9와 서열번호 661에 제시된 DNMT3A 도메인을 함유한다. 일부 구현예에서, DNA-표적화 시스템 또는 이의 DNA-표적화 모듈의 융합 단백질은 서열번호 599에 제시된 dSpCas9와 서열번호 665에 제시된 DNMT3A 도메인을 함유한다. 일부 구현예에서, DNA-표적화 시스템 또는 이의 DNA-표적화 모듈의 융합 단백질은 서열번호 599에 제시된 dSpCas9와 서열번호 605에 제시된 DNMT3B 도메인을 함유한다. 일부 구현예에서, DNA-표적화 시스템 또는 이의 DNA-표적화 모듈의 융합 단백질은 서열번호 599에 제시된 dSpCas9와 서열번호 606에 제시된 LSD1 도메인을 함유한다. 일부 구현예에서, DNA-표적화 시스템 또는 이의 DNA-표적화 모듈의 융합 단백질은 서열번호 599에 제시된 dSpCas9와 서열번호 607에 제시된 DNMT3L 도메인을 함유한다. 일부 구현예에서, DNA-표적화 시스템 또는 이의 DNA-표적화 모듈의 융합 단백질은 서열번호 599에 제시된 dSpCas9와 서열번호 608에 제시된 EZH2 도메인을 함유한다. 일부 구현예에서, DNA-표적화 시스템 또는 이의 DNA-표적화 모듈의 융합 단백질은 서열번호 599에 제시된 dSpCas9와 서열번호 590에 제시된 KRAB 도메인을 함유한다. 일부 구현예에서, DNA-표적화 시스템 또는 이의 DNA-표적화 모듈의 융합 단백질은 서열번호 599에 제시된 dSpCas9와 서열번호 669에 제시된 KRAB 도메인을 포함한다. 일부 구현예에서, dSpCas9와 효과기 도메인 사이에 링커 및/또는 NLS가 존재한다. 일부 구현예에서, NLS는 융합 단백질의 N-말단 또는 C-말단에 존재할 수 있다.

[0406] [0395] 일부 구현예에서, 본원에서 제공되는 DNA-표적화 시스템 또는 이의 DNA-표적화 모듈의 융합 단백질은 DNA-결합 도메인과 KRAB 도메인을 포함한다. 일부 구현예에서, DNA-결합 도메인은 촉매적으로 비활성인 Cas 효소(dCas)이다. 일부 구현예에서, Cas는 dCas9, 예를 들어 dSaCas9 또는 dSpCas9이다. 일부 구현예에서, DNA-결합 도메인은 dSpCas9이다. 일부 구현예에서, DNA-결합 도메인은 서열번호 599에 제시된 dSpCas9이고 KRAB 도메인은 서열번호 590에 제시된다. 일부 구현예에서, DNA-결합 도메인은 서열번호 599에 제시된 dSpCas9이고 KRAB 도메인은 서열번호 669에 제시된다. 일부 구현예에서, 융합 단백질은 융합 단백질의 N- 또는 C-말단 또는 Cas와 KRAB 도메인 사이에 하나 이상의 링커 또는 NLS를 포함할 수 있다. 링커 또는 NLS는 본원에서 설명된 바와 같이 임의의 것일 수 있다.

[0407] [0396] 일부 구현예에서, DNA-표적화 시스템 또는 이의 DNA-표적화 모듈의 융합 단백질은 N-말단에서 C-말단 순서로: NLS 및/또는 링커, 서열번호 599에 제시된 dSpCas9, 링커 및/또는 NLS, 및 서열번호 590에 제시된 KRAB 도메인을 포함한다. 일부 구현예에서, DNA-표적화 시스템 또는 이의 DNA-표적화 모듈의 융합 단백질은 N-말단에서 C-말단 순서로: NLS 및/또는 링커, 서열번호 599에 제시된 dSpCas9, 링커 및/또는 NLS, 및 서열번호 669에 제시된 KRAB 도메인을 포함한다.

[0408] [0397] 일부 구현예에서, 본원에서 제공되는 융합 단백질은 NLS2-dSpCas9-NLS- KRAB-NLS2를 포함한다. 일부 구현예에서, 본원에서 제공되는 융합 단백질은 서열번호 595에 제시된 서열, 또는 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, 융합 단백질은 서열번호 595에 제시된 서열을 포함한다.

[0409] [0398] 일부 구현예에서, 본원에서 제공되는 DNA-표적화 시스템 또는 이의 DNA-표적화 모듈의 융합 단백질은 DNA-결합 도메인, DNMT3A 도메인 및 DNMT3L 도메인을 포함한다. 일부 구현예에서, 효과기 도메인은 DNMT3A와 DNMT3L(DNMT3A/L 도메인)의 융합 도메인이다. 일부 구현예에서, DNA-결합 도메인은 촉매적으로 비활성인 Cas 효소(dCas)이다. 일부 구현예에서, Cas는 dCas9, 예를 들어 dSaCas9 또는 dSpCas9이다. 일부 구현예에서, DNA-결합 도메인은 dSpCas9이다. 일부 구현예에서, DNA-결합 도메인은 서열번호 599에 제시된 dSpCas9이다. 일부 구현예에서, 링커 또는 NLS는 DNMT3A 도메인 또는 DNMT3L 도메인 중 하나를 DNA-결합 도메인과 연결한다. 일부 구현예에서, DNA-결합 도메인은 Cas이고, 링커 또는 NLS는 Cas와 DNMT3A 도메인 및 DNMT3L 도메인 중 하나 또는 둘 다 사이에 존재할 수 있다. 일부 구현예에서, 융합 단백질은 DNA-결합 도메인, DNMT3A 도메인 및 DNMT3L 도메인을 임의의 순서로 포함한다. 일부 구현예에서, 링커는 DNMT3A 도메인을 DNMT3L 도메인과 연결한다. 일부 구현예에서, DNMT3A 도메인 및 DNMT3B 도메인은 융합 도메인으로 연결된다. 일부 구현예에서, 융합 도메인은 서열번호 651에 제시된 DNMT3A/L 도메인이다.

[0410] [0399] 일부 구현예에서, DNA-표적화 시스템 또는 융합 단백질은 N-말단에서 C-말단 순서로: NLS 및/또는 링커, 서열번호 599에 제시된 dSpCas9, 링커 및/또는 NLS, 및 서열번호 651에 제시된 DNMT3A/L 도메인을 포함한다. 일부 구현예에서, DNA-표적화 시스템 또는 융합 단백질은 N-말단에서 C-말단 순서로: NLS 및/또는 링커, 서열번호 599에 제시된 dSpCas9, 링커 및/또는 NLS, 및 서열번호 651에 제시된 DNMT3A/L 도메인, 및 링커 및/또는 NLS를

포함한다.

[0411] [0400] 일부 구현예에서, 본원에 제공된 DNA-표적화 시스템 또는 이의 DNA-표적화 모듈의 융합 단백질은 DNA-결합 도메인 및 하나가 KRAB 도메인이고 다른 하나는 DNMT3A 도메인 및 DNMT3L 도메인인 두 개의 효과기 도메인을 포함한다. 일부 구현예에서, DNMT3A 도메인 및 DNMT3L 도메인으로 구성된 효과기 도메인은 DNMT3A 및 DNMT3L의 융합 도메인(DNMT3A/L 도메인)이다. 일부 구현예에서, DNA-결합 도메인은 촉매적으로 비활성인 Cas 효소(dCas)이다. 일부 구현예에서, Cas는 dCas9, 예컨대 dSaCas9 또는 dSpCas9이다. 일부 구현예에서, DNA-결합 도메인은 dSpCas9이다. 일부 구현예에서, 각각의 KRAB 도메인, DNMT3A 도메인 및 DNMT3L 도메인은 DNA-결합 도메인에 대해 N-말단이다. 일부 구현예에서, 각각의 KRAB 도메인, DNMT3A 도메인 및 DNMT3L 도메인은 DNA-결합 도메인에 대해 C-말단이다. 일부 구현예에서, DNA-결합 도메인은 KRAB 도메인 및 DNMT3A 또는 DNMT3L 도메인 중 하나 사이에 있다. 일부 구현예에서, 융합 도메인은 서열번호 651에 기재된 DNMT3A/L 도메인이다. 일부 구현예에서, KRAB 도메인은 서열번호 590에 기재된다. 일부 구현예에서, KRAB 도메인은 서열번호 669에 기재된다. 일부 구현예에서, DNA-결합 도메인은 서열번호 599에 기재된 dSpCas9이다. 일부 구현예에서, 융합 단백질은 예컨대 융합 단백질의 N-말단 또는 C-말단에 또는 Cas와 KRAB 도메인 또는 DNMT3A/L 도메인 사이에 있는 하나 이상의 링커 또는 NLS를 포함할 수 있다. 링커 또는 NLS는 본원에 기재된 임의의 것일 수 있다. 일부 구현예에서, 링커 또는 NLS는 DNMT3A/L 도메인 및 Cas 사이에 존재할 수 있다. 일부 구현예에서, 링커 또는 NLS는 Cas 및 KRAB 도메인 사이에 존재할 수 있다.

[0412] [0401] 일부 구현예에서, 융합 단백질은 DNA-결합 도메인, 및 제1 및 제2 효과기 도메인으로서 KRAB 도메인 및 DNMT3A/3L 융합 도메인을 함유한다. 일부 구현예에서, 제1 효과기 도메인은 제2 효과기 도메인의 N-말단에 융합되고, 제2 효과기 도메인은 DNA-결합 도메인의 N-말단에 융합된다. 일부 구현예에서, 본원에 제공된 융합 단백질은 DNMT3A/L-KRAB-dSpCas9를 순서대로 포함한다. 일부 구현예에서, 본원에 제공된 융합 단백질은 KRAB-DNMT3A/L-dSpCas9를 순서대로 포함한다. 일부 구현예에서, 각각의 KRAB 도메인 및 DNMT3A/3L 도메인은 DNA-결합 도메인의 C-말단이되, 제1 효과기 도메인은 DNA-결합 도메인의 C-말단에 융합되고, 제2 효과기 도메인은 제1 효과기 도메인의 C-말단에 융합된다. 일부 구현예에서, 본원에 제공된 융합 단백질은 dSpCas9-DNMT3A/L-KRAB를 순서대로 포함한다. 일부 구현예에서, 본원에 제공된 융합 단백질은 dSpCas9-KRAB-DNMT3A/L을 순서대로 포함한다. 일부 구현예에서, DNA-결합 도메인은 KRAB 도메인 및 DNMT3A/3L 도메인 사이에 있되, 하나의 효과기 도메인은 DNA-결합 도메인의 N-말단에 융합되고 다른 효과기 도메인은 DNA-결합 도메인의 C-말단에 융합된다. 일부 구현예에서, 본원에 제공된 융합 단백질은 KRAB-dSpCas9-DNMT3A/L을 순서대로 포함한다. 일부 구현예에서, 본원에 제공된 융합 단백질은 DNMT3A/L-dSpCas9-KRAB를 순서대로 포함한다.

[0413] [0402] 일부 구현예에서, DNA-표적화 시스템 또는 이의 DNA-표적화 모듈의 융합 단백질은 N-말단 내지 C-말단 순서로 서열번호 651에 기재된 DNMT3A/L 융합 도메인, NLS 및/또는 링커, 서열번호 599에 기재된 dSpCas9, 링커 및/또는 NLS, 및 서열번호 590에 기재된 KRAB 도메인을 포함한다. 일부 구현예에서, DNA-표적화 시스템 또는 이의 DNA-표적화 모듈의 융합 단백질은 N-말단에서 C-말단 순서로 다음을 포함한다: 서열번호 651에 제시된 DNMT3A/L 융합 도메인, NLS 및/또는 링커, 서열번호 599에 제시된 dSpCas9, 링커 및/또는 NLS, 및 서열번호 669에 제시된 KRAB 도메인.

[0414] [0403] 일부 구현예에서, DNA-표적화 시스템 또는 융합 단백질은 N-말단에서 C-말단 순서로 다음을 포함한다: 서열번호 651에 제시된 DNMT3A/L 융합 도메인, 제1 링커 및/또는 NLS, 서열번호 599에 제시된 dSpCas9, 제2 링커 및/또는 NLS, 및 서열번호 590에 제시된 KRAB 도메인. 일부 구현예에서, DNA-표적화 시스템 또는 융합 단백질은 N-말단에서 C-말단 순서로 다음을 포함한다: 서열번호 651에 제시된 DNMT3A/L 융합 도메인, 제1 링커 및/또는 NLS, 서열번호 599에 제시된 dSpCas9, 제2 링커 및/또는 NLS, 및 서열번호 669에 제시된 KRAB 도메인.

[0415] [0404] 일부 구현예에서, DNA-표적화 시스템 또는 이의 DNA-표적화 모듈의 융합 단백질은 서열번호 645에 제시된 서열, 또는 서열번호 645와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, 융합 단백질은 서열번호 645에 제시된 서열을 포함한다.

[0416] [0405] 일부 구현예에서, DNA-표적화 시스템 또는 이의 DNA-표적화 모듈의 융합 단백질은 서열번호 647에 제시된 서열, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, 융합 단백질은 서열번호 647에 제시된 서열을 포함한다.

[0417] [0406] 일부 구현예에서, DNA-표적화 시스템 또는 이의 DNA-표적화 모듈의 융합 단백질은 서열번호 649에 제시된 서열, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, 융합 단백질은 서열번호 649에 제시된 서열을 포함한다.

[0418] [0407] 일부 구현예에서, 폴리뉴클레오타이드는 DNMT3A/L-dCas9-KRAB 융합 단백질을 인코딩하는 서열, 예를 들어 서열번호 680에 제시된 서열, 또는 이와 적어도 약 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 또는 100%의 서열 동일성을 갖는 서열을 포함하는 mRNA 분자이다. 일부 구현예에서, 폴리뉴클레오타이드는 서열번호 680에 제시된다.

[0419] [0408] 일부 측면에서, 예시적인 링커 또는 NLS 서열은 여기에 설명된 모든 것일 수 있다. 일부 측면에서, 예시적인 링커 또는 NLS 서열은 DNA-결합 도메인(예컨대 dSpCas9)과 하나 이상의 효과기 도메인(예컨대 KRAB 및 DNMT3A/L) 사이에 임의의 순서로 배치될 수 있다. 상기 구현예의 모든 측면에서, NLS는 섹션 E에 명시된 것과 같이 설명된 모든 것일 수 있다. 상기 구현예의 모든 측면에서, 링커는 섹션 F에 명시된 것과 같이 설명된 모든 것일 수 있다.

[0420] *3. eZFP 융합 단백질*

[0421] [0409] 일부 측면에서, DNA-표적화 시스템은 eZFP 융합 단백질, 예를 들어 본원에 제공된 단백질을 포함한다. 일부 측면에서, eZFP 융합 단백질을 포함하는 DNA-표적화 시스템이 제공된다.

[0422] [0410] 일부 측면에서, 본원에 제공된 ZFP 융합 단백질은 HBV 유전자 또는 이의 조절 요소를 표적화하거나 표적화할 수 있다. 일부 구현예에서, 융합 단백질은 eZFP(즉, 융합 단백질은 eZFP 융합 단백질임)를 포함하며, 예를 들어, I-C 섹션에서 설명된 eZFP 중 하나이다. 일부 구현예에서, 융합 단백질은 추가로, 예를 들어, I-E 섹션에서 설명된 전사 억제를 위한 효과기 도메인 중 하나와 같은 후성적 효과기 도메인을 포함한다. 일부 구현예에서, 융합 단백질은 HBV 유전자 또는 이의 조절 요소의 발현을 감소시키는 적어도 하나의 후성적 효과기 도메인을 포함한다. 일부 구현예에서, 융합 단백질은 둘 이상의 효과기 도메인을 포함한다. 일부 구현예에서, 융합 단백질은 하나 이상의 추가 요소, 예를 들어, 핵 국소화 신호(NLS) 또는 링커, 예를 들어, 본 명세서에 설명된 NLS 또는 링커 중 하나이다. 일부 측면에서, 융합 단백질의 요소는 융합 단백질 내에서 N-말단에서 C-말단까지의 순서와 같이 임의의 적절한 순서로 배열될 수 있다. 일부 측면에서, 본원에서 제공된 eZFP를 포함하는 융합 단백질은 HBV 감염, 간 질환 또는 암과 관련된 질병 또는 장애를 치료하기 위한 조성물 및 방법과 관련하여 HBV 유전자 또는 이의 조절 요소의 감소된 발현을 촉진할 수 있다.

[0423] [0411] 일부 측면에서, eZFP를 포함하는 융합 단백질은 본원에서 제공된 임의의 표적 부위에 결합하거나 결합할 수 있다(즉, 표적화). 일부 구현예에서, 융합 단백질은 표적 부위에 결합한다. 일부 구현예에서, eZFP를 포함하는 융합 단백질은 융합 단백질의 다른 요소가 없는 경우 eZFP가 결합하는 표적 부위에 결합한다. 따라서 일부 구현예에서, 융합 단백질의 eZFP는 융합 단백질의 표적 특이적 결합을 용이하게 한다. 일부 측면에서, 융합 단백질은 본원에서 설명된 임의의 eZFP에 의해 표적화된 표적 부위, 예를 들어 섹션 I-C에 기재된 표적 부위를 표적으로 한다. 일부 구현예에서, 융합 단백질은 표 7의 표적 부위를 표적으로 한다. 일부 구현예에서, 융합 단백질은 표 E4의 표적 부위를 표적으로 한다. 일부 구현예에서, 융합 단백질은 서열번호 1028-1055 중 어느 하나에 제시된 뉴클레오타이드 서열, 적어도 12nt의 이의 연속 부분, 또는 상기 중 어느 하나의 상보적 서열을 포함하는 표적 부위를 표적으로 한다. 일부 구현예에서, 표적 부위는 서열번호 1028-1055 중 어느 하나에 제시된 뉴클레오타이드 서열을 포함한다. 일부 구현예에서, 표적 부위는 서열번호 1045, 1046 또는 1052 중 어느 하나에 제시된 뉴클레오타이드 서열, 적어도 12nt의 이의 연속 부분, 또는 상기 중 어느 하나의 상보적 서열을 포함한다. 일부 구현예에서, 표적 부위는 서열번호 1045, 1046 또는 1052 중 어느 하나에 제시된 뉴클레오타이드 서열을 포함한다.

[0424] [0412] 일부 측면에서, 융합 단백질은 표 E4에 제시된 eZFP 중 어느 하나를 포함한다. 일부 측면에서, 융합 단백질은 표 E4에 제시된 eZFP 중 하나에 대해 제시된 인식 영역 F1-F6을 포함하는 eZFP를 포함한다(즉, 다음을 포함함: 서열번호 720-725 각각; 서열번호 726-731 각각; 서열번호 732-737 각각; 서열번호 738-743 각각; 서열번호 744-749 각각; 서열번호 750-755 각각; 서열번호 756-761 각각; 서열번호 762-767 각각; 서열번호 768-773 각각; 서열번호 774-779 각각; 서열번호 780-785 각각; 서열번호 786-791 각각; 서열번호 792-797 각각; 서열번호 798-803 각각; 서열번호 804-809 각각; 서열번호 810-815 각각; 서열번호 816-821 각각; 서열번호 822-827 각각; 서열번호 828-833 각각; 서열번호 834-839 각각; 서열번호 840-845 각각; 서열번호 846-851 각각; 서열번호 852-857 각각; 서열번호 858-863 각각; 서열번호 864-869 각각; 서열번호 870-875 각각; 서열번호 876-881 각각; 또는 서열번호 882-887 각각). 일부 구현예에서, 융합 단백질은 서열번호 692-719 중 어느 하나에 제시된 아미노산 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, 융합 단백질은 서열번호 692-719 중 어느 하나에 제시된 아미노산 서열을 포함한다. 일부 구현예에서, 융합 단백질은 서열번호 709, 710, 716 중

어느 하나에 제시된 아미노산 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, 융합 단백질은 서열번호 709, 710 또는 716 중 어느 하나에 제시된 아미노산 서열을 포함한다. 일부 구현예에서, 융합 단백질의 eZFP는 서열번호 888-915 중 어느 하나에 제시된 뉴클레오타이드 서열, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩된다. 일부 구현예에서, 융합 단백질의 eZFP는 서열번호 888-915 중 어느 하나에 제시된 뉴클레오타이드 서열에 의해 인코딩된다. 일부 구현예에서, 융합 단백질의 eZFP는 서열번호 905, 906, 912 중 어느 하나에 제시된 뉴클레오타이드 서열, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩된다. 일부 구현예에서, 융합 단백질의 eZFP는 서열번호 905, 906, 또는 912 중 어느 하나에 제시된 뉴클레오타이드 서열에 의해 인코딩된다.

[0425] [0413] 일부 구현예에서, 본원에 따라 예컨대 표 E4에 기술된 eZPF 중 어느 하나와 같은 eZFP를 포함하는 융합 단백질이 제공도니다. 일부 구현예에서, 융합 단백질은 서열번호 1028-1055 중 어느 하나에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위, 적어도 12nt의 이의 연속 부분 또는 상기 중 어느 하나의 상보적 서열을 표적으로 한다. 일부 구현예에서, 융합 단백질은 서열번호 1028-1055 중 어느 하나에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위를 표적으로 한다. 일부 구현예에서, 융합 단백질의 eZFP는 N-말단에서 C-말단의 순서로 F1 내지 F6으로 표시되는 6개의 아연 집게를 포함하며, 각각은 상응하는 아연 집게 인식 영역 F1 내지 F6을 포함하고, 인식 영역 F1-F6의 아미노산 서열은 다음을 포함한다: 서열번호 720-725 각각; 서열번호 726-731 각각; 서열번호 732-737 각각; 서열번호 738-743 각각; 서열번호 744-749 각각; 서열번호 750-755 각각; 서열번호 756-761 각각; 서열번호 762-767 각각; 서열번호 768-773 각각; 서열번호 774-779 각각; 서열번호 780-785 각각; 서열번호 786-791 각각; 서열번호 792-797 각각; 서열번호 798-803 각각; 서열번호 804-809 각각; 서열번호 810-815 각각; 서열번호 816-821 각각; 서열번호 822-827 각각; 서열번호 828-833 각각; 서열번호 834-839 각각; 서열번호 840-845 각각; 서열번호 846-851 각각; 서열번호 852-857 각각; 서열번호 858-863 각각; 서열번호 864-869 각각; 서열번호 870-875 각각; 서열번호 876-881 각각; 또는 서열번호 882-887 각각.

[0426] [0414] 일부 구현예에서, 융합 단백질의 eZFP는 N-말단에서 C-말단 순서로 F1 내지 F6으로 표시되는 6개의 아연 집게를 포함하며, 각각은 상응하는 아연 집게 인식 영역 F1 내지 F6을 포함하고, 인식 영역 F1-F6의 아미노산 서열은 각각 다음을 포함한다: F1:QSAHRKN(서열번호 822) F2:TSSNRKT(서열번호 823) F3:RSDNLSA(서열번호 824) F4:RNNDRKT(서열번호 825) F5:TSGSLSR(서열번호 826) F6:QAGHLAK(서열번호 827).

[0427] [0415] 일부 측면에서, 본 발명에 따라 표 E5에 기술된 eZPF 중 하나와 같은 eZFP를 포함하는 융합 단백질이 제공된다. 일부 구현예에서, 융합 단백질은 서열번호 944-971에 제시된 아미노산 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, 융합 단백질은 서열번호 944-971에 제시된 아미노산 서열을 포함한다. 일부 구현예에서, 융합 단백질은 서열번호 916-943에 제시된 뉴클레오타이드 서열, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩된다. 일부 구현예에서, 융합 단백질은 서열번호 916-943에 제시된 뉴클레오타이드 서열에 의해 인코딩된다.

[0428] [0416] 일부 측면에서, 본 발명에 따라 표 E5에 기술된 eZPF 중 하나와 같은 eZFP를 포함하는 융합 단백질이 제공된다. 일부 구현예에서, 융합 단백질은 서열번호 961, 962, 968에 제시된 아미노산 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, 융합 단백질은 서열번호 961, 962, 968에 제시된 아미노산 서열을 포함한다. 일부 구현예에서, 융합 단백질은 서열번호 933, 934, 940에 제시된 뉴클레오타이드 서열, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩된다. 일부 구현예에서, 융합 단백질은 서열번호 933, 934, 940에 제시된 뉴클레오타이드 서열에 의해 인코딩된다.

[0429] [0417] 일부 측면에서, 본 발명에 따라 표 E6에 기술된 eZPF 중 하나와 같은 eZFP를 포함하는 융합 단백질이 제공된다. 일부 구현예에서, 융합 단백질은 서열번호 1000-1027에 제시된 아미노산 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, 융합 단백질은 서열번호 1000-1027에 제시된 아미노산 서열을 포함한다. 일부 구현예에서, 융합 단백질은 서열번호 972-999에 제시된 뉴클레오타이드 서열, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩된다. 일부 구현

예에서, 융합 단백질은 서열번호 916-943에 제시된 뉴클레오타이드 서열에 의해 인코딩된다.

[0430] [0418] 일부 측면에서, 본 발명에 따라 표 E6에 설명된 eZPF 중 하나와 같은 eZFP를 포함하는 융합 단백질이 제공된다. 일부 구현예에서, 융합 단백질은 서열번호 1017, 1018, 1024에 제시된 아미노산 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, 융합 단백질은 서열번호 1017, 1018, 1024에 제시된 아미노산 서열을 포함한다. 일부 구현예에서, 융합 단백질은 서열번호 989, 990, 996에 제시된 뉴클레오타이드 서열, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩된다. 일부 구현예에서, 융합 단백질은 서열번호 989, 990, 996에 제시된 뉴클레오타이드 서열에 의해 인코딩된다.

[0431] [0419] 일부 구현예에서, 융합 단백질의 eZFP는 N-말단에서 C-말단 순서로 F1 내지 F6으로 표시되는 6개의 아연 집게를 포함하며, 각각은 상응하는 아연 집게 인식 영역 F1 내지 F6을 포함하고, 인식 영역 F1-F6의 아미노산 서열은 각각 다음을 포함한다: F1:RSDHLSQ(서열번호 828) F2:ASSTRTK(서열번호 829) F3:RSDDLTR(서열번호 830) F4:QKSNLSS(서열번호 831) F5:QSANRTT(서열번호 832) F6:QNATRTK(서열번호 833).

[0432] [0420] 일부 구현예에서, 융합 단백질의 eZFP는 N-말단에서 C-말단 순서로 F1 내지 F6으로 표시되는 6개의 아연 집게를 포함하며, 각각은 상응하는 아연 집게 인식 영역 F1 내지 F6을 포함하고, 인식 영역 F1-F6의 아미노산 서열은 각각 다음을 포함한다: F1:QSSSLVR(서열번호 864) F2:QSGDLRR(서열번호 865) F3:RSDERKR(서열번호 866) F4:HRTTLTN(서열번호 867) F5:RSDHLTN(서열번호 868)F6:TSGELVR(서열번호 869).

[0433] [0421] 일부 구현예에서, 본 발명에 따라 본원에서 설명된 바와 같은 eZFP_18과 같은 eZFP를 포함하는 융합 단백질이 제공된다. 일부 구현예에서, 융합 단백질은 서열번호 1045에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위, 적어도 12nt의 이의 연속 부분 또는 상기 중 어느 하나의 상보적 서열을 표적으로 한다. 일부 구현예에서, 융합 단백질은 서열번호 1045에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위를 표적으로 한다. 일부 구현예에서, 융합 단백질은 N-말단에서 C-말단 순서로 F1 내지 F6으로 표시되는 6개의 아연 집게를 포함하며, 각각은 상응하는 아연 집게 인식 영역 F1 내지 F6을 포함하고, 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같다: F1:QSAHRKN(서열번호 822) F2:TSSNRKT(서열번호 823) F3:RSDNLSA(서열번호 824) F4:RNNDRKT(서열번호 825) F5:TSGSLSR(서열번호 826) F6:QAGHLAK(서열번호 827). 일부 구현예에서, 융합 단백질은 서열번호 709에 제시된 아미노산 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, 융합 단백질은 서열번호 709에 제시된 아미노산 서열을 포함한다. 일부 구현예에서, 융합 단백질은 서열번호 905에 제시된 뉴클레오타이드 서열 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩된다. 일부 구현예에서, 융합 단백질은 서열번호 905에 제시된 뉴클레오타이드 서열에 의해 인코딩된다.

[0434] [0422] 일부 구현예에서, 본 발명에 설명된 바와 같은 eZFP, 예를 들어 eZFP_19를 포함하는 융합 단백질이 제공된다. 일부 구현예에서, 융합 단백질은 서열번호 1046에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위, 적어도 12nt의 연속 부분, 또는 상기 중 어느 하나의 상보적 서열을 표적으로 한다. 일부 구현예에서, 융합 단백질은 서열번호 1046에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위를 표적으로 한다. 일부 구현예에서, 융합 단백질은 N-말단에서 C-말단 순서로 F1 내지 F6으로 표시되는 6개의 아연 집게를 포함하며, 각각은 상응하는 아연 집게 인식 영역 F1 내지 F6을 포함하고, 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같다: F1:RSDHLSQ(서열번호 828) F2:ASSTRTK(서열번호 829) F3:RSDDLTR(서열번호 830) F4:QKSNLSS(서열번호 831) F5:QSANRTT(서열번호 832) F6:QNATRTK(서열번호 833). 일부 구현예에서, 융합 단백질은 서열번호 710에 제시된 아미노산 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, 융합 단백질은 서열번호 710에 제시된 아미노산 서열을 포함한다. 일부 구현예에서, 융합 단백질은 서열번호 906에 제시된 뉴클레오타이드 서열 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩된다. 일부 구현예에서, 융합 단백질은 서열번호 906에 제시된 뉴클레오타이드 서열에 의해 인코딩된다.

[0435] [0423] 일부 구현예에서, 본 발명에 따라 본 발명에서 설명된 바와 같은 eZFP_25와 같은 eZFP를 포함하는 융합 단백질이 제공된다. 일부 구현예에서, 융합 단백질은 서열번호 1052에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위, 적어도 12nt의 연속 부분 또는 상기 중 어느 하나의 상보적 서열을 표적으로 한다. 일부

구현예에서, 융합 단백질은 서열번호 1052에 제시된 뉴클레오타이드 서열을 포함하는 표적 부위를 표적으로 한다. 일부 구현예에서, 융합 단백질은 N-말단에서 C-말단 순서로 F1 내지 F6으로 표시되는 6개의 아연 집게를 포함하며, 각각은 상응하는 아연 집게 인식 영역 F1 내지 F6을 포함하고, 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같다: F1:QSSSLVR(서열번호 864) F2:QSGDLRR(서열번호 865) F3:RSDERKR(서열번호 866) F4:HRTTLTN(서열번호 867) F5:RSDHLTN(서열번호 868) F6:TSGELVR(서열번호 869). 일부 구현예에서, 융합 단백질은 서열번호 716에 제시된 아미노산 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, 융합 단백질은 서열번호 716에 제시된 아미노산 서열을 포함한다. 일부 구현예에서, 융합 단백질은 서열번호 912에 제시된 뉴클레오타이드 서열 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩된다. 일부 구현예에서, 융합 단백질은 서열번호 912에 제시된 뉴클레오타이드 서열에 의해 인코딩된다.

[0436] [0424] 일부 측면에서, 융합 단백질은 또한 하나 이상의 핵 국소화 신호(NLS), 예를 들어 SV40 NLS(서열번호 614) 또는 여기에 설명된 모든 NLS를 추가로 포함한다. 일부 측면에서, NLS는 융합 단백질의 핵 국소화를 촉진할 수 있다.

[0437] **II. 폴리뉴클레오타이드 및 벡터 및 관련 전달 방법**

[0438] [0425] 일부 측면에서, 본원에서 기술된 DNA-표적화 시스템 중 하나 또는 상기의 일부 또는 구성요소를 인코딩하는 폴리뉴클레오타이드가 제공된다. 일부 측면에서, 폴리뉴클레오타이드는 후성적-변형 DNA-표적화 시스템의 구성요소 중 하나 및/또는 본 개시의 방법의 측면을 수행하는 데 필요한 핵산 또는 단백질 분자를 인코딩할 수 있다. 특정 구현예에서, 본원에서 기술된 융합 단백질 중 하나 및/또는 본원에서 기술된 gRNA 중 하나를 인코딩하는 폴리뉴클레오타이드가 제공된다.

[0439] [0426] 일부 구현예에서, 본원에 기재된 gRNA를 포함하는 폴리뉴클레오타이드가 제공된다. 일부 구현예에서, gRNA는 표적 세포 내의 유전자 작제물(즉, 벡터 또는 플라스미드)에서 전사된다. 일부 구현예에서, gRNA는 시험관 내 전사에 의해 생성되고 표적 세포에 전달된다. 일부 구현예에서, gRNA는 증가된 안정성을 위해 하나 이상의 변형된 뉴클레오타이드를 포함한다. 일부 구현예에서, gRNA는 융합 단백질과 RNP로서 미리 복합체를 형성한 표적 세포에 전달된다.

[0440] [0427] 일부 구현예에서, 폴리뉴클레오타이드는 RNA 또는 DNA이다. 일부 구현예에서, 제공된 융합 단백질을 인코딩하는 폴리뉴클레오타이드와 같은 폴리뉴클레오타이드는 mRNA이다. 일부 구현예에서, gRNA는 RNA로 제공되고 융합 단백질을 인코딩하는 폴리뉴클레오타이드는 mRNA이다. mRNA는 5' 캡핑되고/되거나 3' 폴리아데닐화될 수 있다. 다른 구현예에서, 제공된 융합 단백질을 인코딩하는 폴리뉴클레오타이드와 같은 본원에서 제공된 폴리뉴클레오타이드는 DNA이다. DNA는 벡터에 존재할 수 있다.

[0441] [0428] 일부 구현예에서, 폴리뉴클레오타이드는 DNMT3A/L-dCas9-KRAB 융합 단백질을 인코딩하는 서열, 예를 들어 서열번호 680에 제시된 서열, 또는 이와 적어도 약 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 또는 100%의 서열 동일성을 갖는 서열을 포함하는 mRNA 분자이다. 일부 구현예에서, 폴리뉴클레오타이드는 서열번호 680에 제시된다.

[0442] [0429] 일부 구현예에서, 제공된 폴리뉴클레오타이드는 (a) B형 간염 바이러스 DNA 서열의 표적 부위를 표적화하기 위한 DNA-결합 도메인, 예를 들어, 본원에 설명된 바와 같은 유전자 또는 이의 조절 요소; 및 (b) 유전자의 전사를 감소시킬 수 있는 적어도 하나의 효과기 도메인을 포함하는 본원에 설명된 바와 같은 융합 단백질을 인코딩한다. 일부 구현예에서, 적어도 하나의 효과기 도메인은 전사 억제자 효과기 도메인이다. 일부 구현예에서, 융합 단백질은 Cas 단백질 또는 이의 변이체와 유전자의 전사를 감소시킬 수 있는 적어도 하나의 전사 억제자 효과기 도메인의 융합 단백질을 포함한다. 일부 구현예에서, Cas는 dCas, 예를 들어 dCas9이다. 일부 구현예에서, dCas9는 dSpCas9, 예를 들어 서열번호 599에 명시된 dSpCas9를 인코딩하는 폴리뉴클레오타이드이다. 이러한 도메인 및 융합 단백질의 예로는 섹션 I에 설명된 바와 같은 것이 포함된다.

[0443] [0430] 일부 구현예에서, 본원에 제공된 DNA-표적화 시스템의 모든 구성 요소는 하나의 폴리뉴클레오타이드에 인코딩된다. 일부 구현예에서, 폴리뉴클레오타이드는 Cas DNA-표적화 도메인과 효과기 도메인을 포함하는 융합 단백질을 인코딩하고, 또한 HepB 바이러스 서열의 표적 부위, 예를 들어 유전자 또는 이의 조절 요소를 표적화하는 gRNA를 인코딩한다.

[0444] [0431] 일부 구현예에서, 본원에서 제공되는 다중화된 DNA-표적화 시스템의 모든 구성 요소는 하나의 폴리뉴클

레오타이드에 인코딩된다. 일부 측면에서, 다중화된 DNA-표적화 시스템은 적어도 제1 DNA-표적화 모듈과 제2 DNA-표적화 모듈을 포함하는데, 여기서 제1 DNA-표적화 모듈과 제2 DNA-표적화 모듈은 하나의 폴리뉴클레오타이드, 예를 들어 제1 폴리뉴클레오타이드에 인코딩된다. 일부 구현예에서, 제1 DNA-표적화 모듈과 제2 DNA-표적화 모듈은 하나의 폴리뉴클레오타이드, 예를 들어 제1 폴리뉴클레오타이드에 인코딩된다. 일부 구현예에서, 본원에서 제공되는 다중화된 DNA-표적화 시스템의 모든 구성 요소는 여러 개의 개별 폴리뉴클레오타이드, 예를 들어 제1 폴리뉴클레오타이드와 제2 폴리뉴클레오타이드에 인코딩된다.

[0445] [0432] 일부 구현예에서, 다중화된 DNA-표적화 시스템의 모듈은 Cas DNA-결합 도메인을 포함하는데, 여기서 제1 DNA-표적화 모듈은 제1 Cas 단백질을 포함하고 제2 DNA-표적화 모듈은 제2 Cas 단백질을 포함하며, 여기서 제1 Cas 단백질과 제2 Cas 단백질은 제1 폴리뉴클레오타이드에 인코딩된다. 이러한 일부 구현예에서, 제1 DNA-표적화 모듈의 제1 gRNA와 제2 DNA-표적화 모듈의 제2 gRNA는 제1 폴리뉴클레오타이드에 인코딩된다. 이러한 일부 구현예에서, 제1 DNA-표적화 모듈의 제1 gRNA와 제2 DNA-표적화 모듈의 제2 gRNA는 제2 폴리뉴클레오타이드에 인코딩된다.

[0446] [0433] 일부 구현예에서, 다중화된 DNA-표적화 시스템의 모듈은 Cas DNA-결합 도메인을 포함하는데, 여기서 제1 DNA-표적화 모듈은 제1 Cas 단백질을 포함하고 제2 DNA-표적화 모듈은 제2 Cas 단백질을 포함하며, 여기서 제1 Cas 단백질과 제2 Cas 단백질은 제1 폴리뉴클레오타이드와 제2 폴리뉴클레오타이드와 같이 별도의 폴리뉴클레오타이드에 의해 인코딩된다. 이러한 일부 구현예에서, 제1 DNA-표적화 모듈의 제1 gRNA는 제1 폴리뉴클레오타이드에 인코딩되고 제2 DNA-표적화 모듈의 제2 gRNA는 제2 폴리뉴클레오타이드에 인코딩된다. 일부 구현예에서, 제1 Cas 단백질과 제1 gRNA는 제1 폴리뉴클레오타이드에 인코딩되고 제2 Cas 단백질과 제2 gRNA는 제2 폴리뉴클레오타이드에 인코딩된다.

[0447] [0434] 일부 구현예에서, 다중화된 DNA-표적화 시스템의 모듈은 각각 Cas DNA-결합 도메인을 포함하는데, 여기서 제1 DNA-표적화 모듈과 제2 DNA-표적화 모듈의 Cas 단백질은 동일하고 동일한 뉴클레오타이드 서열에 의해 인코딩된다. 일부 구현예에서, Cas 단백질, 제1 gRNA 및 제2 gRNA는 제1 폴리뉴클레오타이드에 인코딩된다.

[0448] [0435] 일부 구현예에서, 다중화된 DNA-표적화 시스템은 제3 DNA-결합 모듈을 포함할 수 있다. 일부 구현예에서, 제3 DNA-결합 모듈의 구성 요소는 제1 폴리뉴클레오타이드, 제2 폴리뉴클레오타이드에 의해 인코딩되거나, 제3 폴리뉴클레오타이드에 의해 인코딩된다.

[0449] [0436] 일부 구현예에서, DNA-표적화 시스템 또는 다중화된 DNA-표적화 시스템 모듈의 DNA-결합 도메인을 인코딩하는 폴리뉴클레오타이드는 dCas9-KRAB 융합 단백질을 인코딩하는 서열을 포함한다. 일부 구현예에서, dSpCas9-KRAB 융합 단백질을 인코딩하는 폴리뉴클레오타이드는 서열번호 594에 제시된 서열, 또는 이와 적어도 약 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 또는 100%의 서열 동일성을 갖는 서열을 포함한다. 일부 구현예에서, dCas9-KRAB 융합 단백질을 인코딩하는 폴리뉴클레오타이드는 서열번호 594에 제시되어 있다. 일부 구현예에서, 폴리뉴클레오타이드는 서열번호 595를 포함하는 아미노산 서열, 또는 이와 적어도 약 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 또는 100%의 서열 동일성을 갖는 서열을 갖는 dSpCas9-KRAB 융합 단백질을 인코딩한다. 일부 구현예에서, 폴리뉴클레오타이드는 서열번호 595에 제시된 아미노산 서열을 갖는 dCas9-KRAB 융합 단백질을 인코딩한다. 일부 구현예에서, 폴리뉴클레오타이드는 RNA 또는 DNA이다. 일부 구현예에서, 제공된 융합 단백질을 인코딩하는 폴리뉴클레오타이드와 같은 폴리뉴클레오타이드는 mRNA이다. mRNA는 5' 캡핑 및/또는 3' 폴리아데닐화될 수 있다. 다른 구현예에서, 제공된 융합 단백질을 인코딩하는 폴리뉴클레오타이드와 같은 본원에서 제공된 폴리뉴클레오타이드는 DNA이다. DNA는 벡터에 존재할 수 있다.

[0450] [0437] 일부 구현예에서, DNA-표적화 시스템 또는 다중화된 DNA-표적화 시스템 모듈의 DNA-결합 도메인을 인코딩하는 폴리뉴클레오타이드는 DNMT3A/L-dCas9-KRAB 융합 단백질을 인코딩하는 서열을 포함한다. 일부 구현예에서, DNMT3A/L-dCas9-KRAB 융합 단백질을 인코딩하는 폴리뉴클레오타이드는 서열번호 644에 제시된 서열, 또는 이와 적어도 약 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 또는 100%의 서열 동일성을 갖는 서열을 포함한다. 일부 구현예에서, DNMT3A/L-dCas9-KRAB 융합 단백질을 인코딩하는 폴리뉴클레오타이드는 서열번호 644에 제시되어 있다. 일부 구현예에서, 폴리뉴클레오타이드는 서열번호 645를 포함하는 아미노산 서열, 또는 이와 적어도 약 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 또는 100%의 서열 동일성을 갖는 서열을 갖는 DNMT3A/L-dCas9-KRAB 융합 단백질을 인코딩한다. 일부 구현예에서, 폴리뉴클레오타이드는 서열번호 645에 제시된 아미노산 서열을 갖는 DNMT3A/L-dCas9-KRAB 융합 단백질을 인코딩한다. 일부 구현예에서, 폴리뉴클레오타이드는 RNA 또는 DNA이다. 일부 구현예에서, 제공된 융합 단백질을 인코딩하는 폴리뉴클

레오타이드와 같은 폴리뉴클레오타이드는 mRNA이다. mRNA는 5' 캡핑 및/또는 3' 폴리아데닐화될 수 있다. 다른 구현예에서, 제공된 융합 단백질을 인코딩하는 폴리뉴클레오타이드와 같은 본원에서 제공된 폴리뉴클레오타이드는 DNA이다. DNA는 벡터에 존재할 수 있다.

[0451] [0438] 일부 구현예에서, 폴리뉴클레오타이드는 DNMT3A/L-dCas9-KRAB 융합 단백질을 인코딩하는 서열을 포함한다. 일부 구현예에서, DNMT3A/L-dCas9-KRAB 융합 단백질을 인코딩하는 폴리뉴클레오타이드는 서열번호 646에 제시된 서열, 또는 이와 적어도 약 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 또는 100%의 서열 동일성을 갖는 서열을 포함한다. 일부 구현예에서, DNMT3A/L-dCas9-KRAB 융합 단백질을 인코딩하는 폴리뉴클레오타이드는 서열번호 646에 제시되어 있다. 일부 구현예에서, 폴리뉴클레오타이드는 서열번호 647을 포함하는 아미노산 서열, 또는 이와 적어도 약 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 또는 100%의 서열 동일성을 갖는 서열을 갖는 DNMT3A/L-dCas9-KRAB 융합 단백질을 인코딩한다. 일부 구현예에서, 폴리뉴클레오타이드는 서열번호 647에 제시된 아미노산 서열을 갖는 DNMT3A/L-dCas9-KRAB 융합 단백질을 인코딩한다. 일부 구현예에서, 폴리뉴클레오타이드는 RNA 또는 DNA이다. 일부 구현예에서, 제공된 융합 단백질을 인코딩하는 폴리뉴클레오타이드와 같은 폴리뉴클레오타이드는 mRNA이다. mRNA는 5' 캡핑 및/또는 3' 폴리아데닐화될 수 있다. 다른 구현예에서, 제공된 융합 단백질을 인코딩하는 폴리뉴클레오타이드와 같은 본원에서 제공된 폴리뉴클레오타이드는 DNA이다. DNA는 벡터에 존재할 수 있다.

[0452] [0439] 일부 구현예에서, 폴리뉴클레오타이드는 DNMT3A/L-dCas9-KRAB 융합 단백질을 인코딩하는 서열을 포함한다. 일부 구현예에서, DNMT3A/L-dCas9-KRAB 융합 단백질을 인코딩하는 폴리뉴클레오타이드는 서열번호 648에 제시된 서열, 또는 이와 적어도 약 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 또는 100%의 서열 동일성을 갖는 서열을 포함한다. 일부 구현예에서, DNMT3A/L-dCas9-KRAB 융합 단백질을 인코딩하는 폴리뉴클레오타이드는 서열번호 648에 제시되어 있다. 일부 구현예에서, 폴리뉴클레오타이드는 서열번호 649를 포함하는 아미노산 서열, 또는 이와 적어도 약 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 또는 100%의 서열 동일성을 갖는 서열을 갖는 DNMT3A/L-dCas9-KRAB-DNMT3A/L 융합 단백질을 인코딩한다. 일부 구현예에서, 폴리뉴클레오타이드는 서열번호 649에 제시된 아미노산 서열을 갖는 DNMT3A/L-dCas9-KRAB 융합 단백질을 인코딩한다. 일부 구현예에서, 폴리뉴클레오타이드는 RNA 또는 DNA이다. 일부 구현예에서, 제공된 융합 단백질을 인코딩하는 폴리뉴클레오타이드와 같은 폴리뉴클레오타이드는 mRNA이다. mRNA는 5' 캡핑 및/또는 3' 폴리아데닐화될 수 있다. 다른 구현예에서, 제공된 융합 단백질을 인코딩하는 폴리뉴클레오타이드와 같은 본원에서 제공된 폴리뉴클레오타이드는 DNA이다. DNA는 벡터에 존재할 수 있다.

[0453] [0440] 또한, 제공된 폴리뉴클레오타이드를 포함하는 벡터가 본원에 제공된다. 일부 구현예에서, 벡터는 플라스미드 또는 발현 벡터와 같은 유전적 구성을 포함한다.

[0454] [0441] 일부 구현예에서, 본원에서 제공되는 DNA-표적화 시스템의 융합 단백질을 인코딩하는 서열을 포함하는 발현 벡터는 또한 적어도 하나의 gRNA를 인코딩하는 폴리뉴클레오타이드 서열을 포함할 수 있다. gRNA를 인코딩하는 서열은 세포에서 gRNA의 발현을 위한 적어도 하나의 전사 제어 서열에 작동적으로 연결될 수 있다. 예를 들어, gRNA를 인코딩하는 DNA는 RNA 폴리머라제 III(Pol III)에 의해 인식되는 프로모터 서열에 작동적으로 연결될 수 있다. 적합한 Pol III 프로모터의 예로는 포유류 U6, U3, H1 및 7SL RNA 프로모터가 포함되지만 이에 국한되지 않는다.

[0455] [0442] 발현 벡터[예컨대, DNA 또는 RNA(예를 들어, mRNA) 발현 벡터]는 임의의 수의 적합한 전사 제어 서열을 포함할 수 있다. 예를 들어, 전사 제어 서열은 인핸서, 프로모터 또는 비번역 영역[untranslated region, UTR], 예컨대 3'UTR 또는 5'UTR을 포함할 수 있다. 일부 구현예에서, UTR은 발현 벡터(예를 들어, DNA 벡터)에 의해 인코딩되고/되거나 이에 존재한다. 일부 구현예에서, 융합 단백질을 인코딩하는 mRNA는 UTR을 포함한다. 일부 측면에서, 발현 벡터에 사용하기 위한 상이한 전사 조절 서열이 선택되어, 예를 들어 적절한 수준의 발현을 달성할 수 있다. 예를 들어, 일부 구현예에서, 특이적 조직 또는 세포 유형(예를 들어, 간 또는 간세포)에서의 발현을 용이하게 하는 UTR이 선택될 수 있다. 일부 구현예에서, 발현 벡터의 5'UTR은 서열번호 681에 기재된 서열을 인코딩하거나 포함한다. 일부 구현예에서, 발현 벡터의 3'UTR은 서열번호 682에 기재된 서열을 인코딩하거나 포함한다.

[0456] [0443] 일부 구현예에서, 유전자 및/또는 그의 조절 요소의 전사를 감소시킬 수 있는 하나 이상의 효과기 도메인 및 dCas를 포함하는 DNA-결합 도메인을 포함하는 융합 단백질을 인코딩하는 폴리뉴클레오타이드, 및 gRNA를 인코딩하는 폴리뉴클레오타이드 또는 폴리뉴클레오타이드의 조합물, 또는 gRNA의 조합물, 예컨대 두 개의 gRNA, 또는 세 개의 gRNA를 함유하는 벡터가 제공된다. 일부 구현예에서, dCas는 dCas9, 예컨대 dSpCas9이다. 일부 구

현예에서, 폴리뉴클레오타이드는 서열번호 599에 기재된 dSpCas9를 포함하는 융합 단백질을 인코딩한다. 일부 구현예에서, 폴리뉴클레오타이드(들)는 섹션 I.B.ii에 기재된 gRNA 또는 gRNA의 조합물을 인코딩한다. 예를 들어, 폴리뉴클레오타이드는 서열번호 196-229, 230-295 또는 296-390 중 어느 하나에서 선택되는 스페이서 서열, 또는 이의 적어도 14개 nt의 인접 부분을 각각 포함하는 gRNA의 조합물을 인코딩할 수 있다. 일부 구현예에서, 폴리뉴클레오타이드(들)는 각각 서열번호 391-424, 425-490 또는 491-585 중 어느 하나에 기재된 서열을 포함하는 gRNA의 조합물을 인코딩하거나 포함한다.

[0457] [0443] 일부 구현예에서는 dCas와 유전자 및/또는 그 조절 요소의 전사를 감소시킬 수 있는 적어도 하나의 효과기 도메인을 포함하는 DNA-결합 도메인을 포함하는 융합 단백질을 인코딩하는 폴리뉴클레오타이드를 포함하는 벡터와, gRNA 또는 gRNA의 조합, 예를 들어 두 개의 gRNA 또는 세 개의 gRNA를 인코딩하는 폴리뉴클레오타이드 또는 폴리뉴클레오타이드의 조합을 포함하는 벡터가 제공된다. 일부 구현예에서, dCas는 dCas9, 예를 들어 dSpCas9이다. 일부 구현예에서, 폴리뉴클레오타이드는 서열번호 599에 명시된 dSpCas9를 포함하는 융합 단백질을 인코딩한다. 일부 구현예에서, 폴리뉴클레오타이드는 섹션 IBii에 설명된 바와 같이 gRNA 또는 gRNA의 조합을 인코딩한다. 예를 들어, 폴리뉴클레오타이드는 gRNA의 조합을 인코딩할 수 있으며, 각각은 서열번호 196-229, 230-295 또는 296-390 중 하나에서 선택된 스페이서 서열 또는 적어도 14nt의 연속 부분을 포함한다. 일부 구현예에서 폴리뉴클레오타이드는 각각 서열번호 391-424, 425-490 또는 491-585 중 어느 하나에 제시된 서열을 포함하는 gRNA의 조합을 인코딩한다.

[0458] [0444] 일부 구현예에서, 효과기 도메인은 KRAB이다. 일부 구현예에서, 효과기 도메인은 DNMT3A/L이다. 일부 구현예에서, 벡터는 서열번호 590, 서열번호 595, 서열번호 645, 서열번호 647 또는 서열번호 649를 포함하는 아미노산 서열을 인코딩하는 폴리뉴클레오타이드, 또는 이와 적어도 약 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 100%의 서열 동일성을 갖는 서열, 및 섹션 I.B.ii에 기술된 것과 같은 gRNA 또는 gRNA의 조합을 인코딩하는 폴리뉴클레오타이드 또는 폴리뉴클레오타이드를 포함한다. 일부 구현예에서, 폴리뉴클레오타이드는 gRNA의 조합을 인코딩하는데, 각각은 서열번호 196-229, 230-295 또는 296-390 중 어느 하나에서 선택된 스페이서 서열 또는 적어도 14nt의 이의 연속 부분을 포함한다. 일부 구현예에서, 각 gRNA는 추가로 서열번호 587에 제시된 서열을 포함한다. 일부 구현예에서, 폴리뉴클레오타이드는 각각 서열번호 391-424, 425-490 또는 491-585 중 어느 하나에 제시된 서열을 포함하는 gRNA의 조합을 인코딩한다.

[0459] [0445] 일부 구현예에서, 폴리뉴클레오타이드는 융합 단백질과 적어도 하나의 gRNA를 인코딩한다.

[0460] [0446] 일부 구현예에서, 본원에 제공된 바와 같은 폴리뉴클레오타이드는 관심 진핵 세포 또는 동물에서 단백질로의 효율적인 번역을 위해 코돈 최적화될 수 있다. 예를 들어, 코돈은 인간, 마우스, 래트, 햄스터, 소, 돼지, 고양이, 개, 어류, 양서류, 식물, 효모, 곤충 등에서의 발현을 위해 최적화될 수 있다. 코돈 최적화에 대한 프로그램은 프리웨어로 이용 가능하다. 상업용 코돈 최적화 프로그램도 이용 가능하다.

[0461] [0447] 일부 구현예에서, 본원에 기재된 폴리뉴클레오타이드는 하나 이상의 전사 및/또는 번역 제어 요소를 포함할 수 있다. 이용되는 숙주/벡터 시스템에 따라, 구성적 및 유도성 프로모터, 전사 인핸서 요소, 전사 종결인자 등을 포함하는 임의의 수많은 적절한 전사 및 번역 제어 요소가 발현 벡터에서 사용될 수 있다.

[0462] [0448] 적절한 진핵세포 프로모터(즉, 진핵 세포에서 기능하는 프로모터)의 비제한적인 예는 거대세포바이러스[cytomegalovirus, CMV] 극초기, 단순포진 바이러스[herpes simplex virus, HSV] 티미딘 인산화효소, 초기 및 후기 SV40, 레트로바이러스로에서의 긴 말단 반복[long terminal repeat, LTR], 인간 연장 인자-1[elongation factor-1, EF1] 프로모터, 닭 베타-액틴[chicken beta-actin, CAG] 프로모터에 융합된 거대세포바이러스[CMV] 인핸서를 포함하는 하이브리드 작제물, 쥐 줄기 세포 바이러스 프로모터[murine stem cell virus promoter, MSCV], 포스포글리세레이트 인산화효소[phosphoglycerate kinase, PGK]-1 좌위 프로모터, 및 마우스 메탈로싸이오네인-I에서의 것을 포함한다.

[0463] [0449] DNA-표적화 시스템과 관련하여 사용되는 가이드 RNA를 포함하는 소형 RNA 발현에 있어서, 다양한 프로모터 예컨대, 예를 들어 U6 및 H1를 포함하는 RNA 폴리머라제 III 프로모터가 이로울 수 있다. 이러한 프로모터의 설명 및 이의 사용을 증진하기 위한 변수는 당업계에 알려져 있고, 추가적인 정보 및 접근법이 꾸준히 기재되고 있다. 예를 들어, 문헌[Ma, H. et al., Molecular Therapy—Nucleic Acids 3, e161 (2014) doi:10.1038/mtna.2014.12]을 참조한다.

[0464] [0450] 발현 벡터는 또한 번역 개시를 위한 리보솜 결합 부위 및 전사 종결인자를 함유할 수 있다. 발현 벡터는 또한 발현 증폭에 적절한 서열을 포함할 수 있다. 발현 벡터는 또한 부위 지정 폴리펩타이드에 융합되고, 따라

서 융합 단백질을 생성하는 비자연적인 태그(예컨대, 히스티딘 태그, 적혈구응집소 태그, 녹색 형광 단백질, 등)를 인코딩하는 뉴클레오타이드 서열을 포함할 수 있다.

[0465] [0451] 프로모터는 유도성 프로모터(예컨대, 열 충격 프로모터, 테트라사이클린-조절성 프로모터, 스테로이드-조절성 프로모터, 금속-조절성 프로모터, 에스트로겐 수용체-조절성 프로모터 등)일 수 있다. 프로모터는 구성적 프로모터(예컨대, CMV 프로모터, UBC 프로모터)일 수 있다. 일부 경우에, 프로모터는 공간적으로 제한되고/되거나 시간적으로 제한된 프로모터[예컨대, 조직 특이적 프로모터, 세포 유형 특이적 프로모터 등(예를 들어, T 세포 특이적 프로모터)]일 수 있다.

[0466] [0452] 고려되는 발현 벡터는, 백시니아 바이러스, 폴리오바이러스, 아데노바이러스, 아데노-연관 바이러스, SV40, 단순 포진 바이러스, 인간 면역결핍 바이러스, 레트로바이러스(예컨대, 쥐 백혈병 바이러스, 비장 괴사 바이러스, 및 레트로바이러스 예컨대 라우스 육종 바이러스, 하비 육종 바이러스, 조류 백혈병 바이러스, 렌티바이러스, 인간 면역결핍 바이러스, 골수증식성 육종 바이러스, 및 유방 종양 바이러스) 기반 바이러스 벡터 및 다른 재조합 벡터를 포함하지만 이에 제한되지 않는다. 진핵 표적 세포에 고려되는 기타 벡터는, 벡터 pXT1, pSG5, pSVK3, pBPV, pMSG, 및 pSVLSV40(Pharmacia)를 포함하지만 이에 제한되지 않는다. 숙주 세포와 양립가능한 한 다른 벡터가 사용될 수 있다.

[0467] [0453] 일부 구현예에서, 벡터는 바이러스 벡터, 예컨대 아데노 연관 바이러스[adeno-associated virus, AAV] 벡터, 레트로바이러스 벡터, 렌티바이러스 벡터 또는 감마레트로바이러스 벡터이다. 일부 구현예에서 일부 구현예에서, 바이러스 벡터는 아데노 연관 바이러스[AAV] 벡터이다. 일부 구현예에서, AAV 벡터는 AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8 또는 AAV9 벡터 중에서 선택된다. 일부 구현예에서, 벡터는 렌티바이러스 벡터이다. 일부 구현예에서, 벡터는 비바이러스 벡터, 예를 들어 지질 나노입자, 리포솜, 엑소솜, 또는 세포 투과성 펩타이드이다. 일부 구현예에서, 벡터는 하나의 벡터 또는 둘 이상의 벡터를 포함한다.

[0468] [0454] 일부 구현예에서, 본원에 기재된 벡터는 지질 나노입자[LNP]이거나 이를 포함한다. 일부 구현예에서, 본원에 기재된 후성적 변형 DNA-표적화 시스템, gRNA, Cas-gRNA 조합물, 폴리뉴클레오타이드, 융합 단백질, 또는 이의 성분 중 임의의 것은, 예컨대 전달을 위해 지질 나노입자[LNP]에 포함된다. 일부 구현예에서, 지질 나노입자는 전달을 위한 벡터이다. 일부 구현예에서, 나노입자는 적어도 하나의 나노입자를 포함할 수 있다. 지질은 DLin-DMA, DLin-K-DMA, 98N12- 5, C12-200, DLin-MC3-DMA, DLin-KC2-DMA, DODMA, PLGA, PEG, PEG-DMG 및 페길화된 지질에서 선택될 수 있지만 이에 제한되지 않는다. 또 다른 측면에서, 지질은 양이온성 지질, 예컨대 DLin-DMA, DLin-D-DMA, DLin-MC 3 -DMA, DLin-KC2-DMA 및 DODMA일 수 있지만 이에 제한되지 않는다.

[0469] [0455] 지질 나노입자는 핵산 및 단백질, 예컨대 CRISPR/Cas 시스템을 인코딩하고/하거나 포함하는 것을 포함하는 캡슐화되거나 연관된(예를 들어, 복합체화된) 치료제 전달용으로 사용될 수 있다. 예를 들어, 미국 특허 제10,723,692호, 미국 특허 제10,941,395호 및 WO 2015/035136호를 참조한다.

[0470] [0466] 일부 구현예에서, 제공된 방법은 mRNA, 예컨대 제공된 DNA-표적화 시스템 중 임의의 것, 예를 들어 본원에 제공된 융합 단백질 중 임의의 것의 단백질 성분을 인코딩하는 mRNA를 포함하는 지질 나노입자[LNP]의 용도를 포함한다. 일부 구현예에서, mRNA는 당업계에 알려진 방법, 예컨대 시험관 내 전사를 사용하여 생성될 수 있다. 방법의 일부 구현예에서, mRNA는 5' 캡을 포함한다. 일부 구현예에서, 5' 캡은 메신저 RNA와 같은 일차 전사체의 5' 말단에 있는 변경된 뉴클레오타이드이다. 일부 측면에서, mRNA의 5' 캡은 RNA 안정성 및 프로세싱, mRNA 대사, 핵에서 RNA 전사체의 프로세싱 및 성숙, 핵에서 세포질로의 mRNA의 수송, mRNA 안정성 및 mRNA의 단백질로의 효율적인 번역 중 하나 이상을 개선한다. 일부 구현예에서, 5' 캡은 천연 발생 5' 캡 또는 mRNA의 천연 발생 캡과 상이한 것일 수 있다. 5' 캡은 당업자에게 알려진 임의의 5' 캡일 수 있다. 특정 구현예에서, 5' 캡은 항역전 캡 유사체[Anti-Reverse Cap Analog, ARCA] 캡, 7-메틸-구아노신(7-methyl-guanosine, 7mG) 캡, CleanCap캡® 유사체, 백시니아 캡, 및 이의 유사체로 이루어진 군에서 선택된다. 예를 들어, 5' 캡은 비제한적으로 항 역전 캡 유사체[anti-reverse cap analog, ARCA](US7074596), 7-메틸-구아노신, CleanCap® 유사체, 예컨대 Cap 1 유사체(Trilink; San Diego, CA), 또는 예를 들어 백시니아 캡핑 효소 등을 사용하여 효소적으로 캡핑된 것을 포함할 수 있다. 일부 구현예에서, mRNA는 폴리아데닐화될 수 있다. mRNA는 인코딩된 단백질의 발현 및/또는 mRNA 자체의 안정성을 강화시키기 위해 다양한 5' 및 3' 비번역 서열 요소를 함유할 수 있다. 이러한 요소는, 예를 들어, 번역 후 조절 요소, 예컨대 우드척 간염 바이러스 전사후 조절 요소[woodchuck hepatitis virus post-transcriptional regulatory element, WPRE]를 포함할 수 있다. 일부 구현예에서, mRNA는 적어도 하나의 뉴클레오사이드 변형을 포함한다. mRNA는 뉴클레오사이드 유사체에 대한 천연 뉴클레오사이드의 변형을 함유할 수 있다. 당업계에 알려진 임의의 뉴클레오사이드 유사체가 구상된다. 이러한 뉴클레오사이드

유사체는, 예를 들어, US 8,278,036에 기재된 것을 포함할 수 있다. 방법의 특정 구현예에서, 뉴클레오사이드 변형은 우리딘에서 슈도우리딘으로의 변형 및 우리딘에서 N1-메틸 슈도우리딘으로의 변형으로 이루어진 군에서 선택된다. 방법의 특정 구현예에서, 뉴클레오사이드 변형은 우리딘에서 슈도우리딘으로 이루어진다.

[0471] [0457] 일부 구현예에서, 기술된 LNP 조성물은 PEG-지질(예를 들어, 폴리에틸렌 글리콜 성분을 포함하는 지질)을 포함한다. 일부 구현예에서, 기재된 LNP 조성물은 2종 이상의 PEG-지질을 포함한다. 예시적인 PEG-지질은 또한 PEG-변형 포스파티딜에탄올아민, PEG-변형 포스파티드산, PEG-변형 세라마이드, PEG-변형 다이알킬아민, PEG-변형 다이아실글리세롤, PEG-변형 다이알킬글리세롤, 및 이들의 혼합물이 포함되지만 이에 한정되지 않는다. 예를 들어, 하나 이상의 PEG-지질은 PEG- c-DOMG, PEG-DMG, PEG-DLPE, PEG-DMPE, PEG-DPPC, a PEG-DSPE 지질 또는 이들의 조합을 포함할 수 있다. 일부 구현예에서, PEG 모이어티는 임의로 치환된 에틸렌 글리콜 또는 에틸렌 옥사이드의 선형 또는 분지형 중합체이다. 일부 구현예에서, PEG 모이어티는, 예를 들어, 하나 이상의 알킬, 알콕시, 아실, 하이드록시, 또는 아릴기에 의해 치환된다. 일부 구현예에서, PEG 모이어티는 PEG 공중합체, 예컨대 PEG-폴리우레탄 또는 PEG-폴리프로필렌(예를 들어, 문헌[j. Milton Harris, Poly(ethylene glycol) chemistry: biotechnical and biomedical applications (1992)] 참조]을 포함한다. 일부 구현예에서, PEG-지질은 PEG-지질 접합체이다. 일부 구현예에서, PEG-지질은 상기 나노입자 조성물에 존재하는 총 지질 함량의 약 0.1 몰% 내지 약 6 몰%를 포함한다. 일부 구현예에서, PEG-지질의 수 평균 분자량은 약 200 Da 내지 약 5000 Da이다. 일부 구현예에서, LNP는 갈락토스의 아미노 당 유도체인 N-아세틸갈락토사민(Acetylgalactosamine, GalNAc)에 접합되고, GalNAc는 세포 표면 단백질인 아시알로당단백질 수용체[asialoglycoprotein receptor, ASGPR]를 인식하고 이에 결합할 수 있는 당 분자이다. ASGPR은 간세포(간세포)에서 풍부하게 발현된다. 일부 구현예에서, LNP에 대한 GalNAc 접합(예컨대 GalNAc-PEG 지질의 분자량 0.15%)은 마우스 모델(예컨대 FRG 마우스 모델)에서 간세포로의 LNP 전달을 개선한다. 일부 구현예에서, LNP에 대한 GalNAc 접합은 간세포로의 전달을 개선할 수 있다. 일부 구현예에서, 지질 나노입자는 GalNAC-접합 지질(예컨대 GalNAc-PEG 지질)을 포함한다. 일부 구현예에서, 지질 나노입자에서 GalNAC-접합된 지질의 몰 백분율은 약 0% 내지 약 2%이다. 일부 구현예에서, 지질 나노입자에서 GalNAC-접합 지질의 몰 백분율은 약 0.1%, 약 0.2%, 약 0.3%, 약 0.4%, 약 0.5%, 약 0.6%, 약 0.7%, 약 0.8%, 약 0.9%, 약 1.0%, 약 1.1%, 약 1.2%, 약 1.3%, 약 1.4%, 약 1.5%, 약 1.6%, 약 1.7%, 약 1.8%, 약 1.9%, 약 2.0% 이상이거나 전술한 것 중 임의의 것 사이의 값이다.

[0472] [0458] 일부 구현예에서, 본 방법에 유용한 LNP는 DLin-DMA(1,2-다이리놀레일옥시-3-다이메틸아미노프로페인), DLin-MC3-DMA(다이리놀레일메틸-4-다이메틸아미노뷰티레이트), DLin-KC2-DMA(2,2-다이리놀레일-4-(2-다이메틸아미노에틸)-[1,3]-다이올솔레인), DODMA(1,2-다이올레일옥시-N,N-다이메틸-3-아미노프로페인), SS-OP(비스[2-(4-{2-[4-(시스-9 옥타데세노일옥시)페닐아세톡시]에틸ethyl}피페리딘일)에틸]다이설파이드), 및 이의 유도체에서 선택되는 양이온성 지질을 포함한다. DLin-MC3-DMA 및 이의 유도체는, 예를 들어, WO 2010/144740에 기재되어 있다. DODMA 및 이의 유도체는, 예를 들어 US 7,745,651 및 문헌[Mok et al. (1999), Biochimica et Biophysica Acta, 1419(2): 137-150]에 기재되어 있다. DLin-DMA 및 이의 유도체는, 예를 들어, US 7,799,565에 기재되어 있다. DLin-KC2-DMA 및 이의 유도체는, 예를 들어, US 9,139,554에 기재되어 있다. SS-OP(NOF America Corporation, White Plains, NY)는, 예를 들어, www.nofamerica.com/store/index.php?dispatch=products.view&product_id=962에 기재되어 있다. 양이온성 지질의 추가 및 비제한적 예는 메틸피리디일-다이알킬산(methylpyridiyl-dialkyl acid, MPDACA), 팔미토일-올레오일-노어-아르기닌(palmitoyl-oleoyl-nor-arginine, PONA), 구아니딘오-다이알킬산(guanidino-dialkyl acid, GUADACA), 1,2-다이-O-옥타데센일-3-트라이메틸암모늄 프로페인(dioleoyl-3-trimethylammonium-propane, DOTMA), 1,2-다이올레일-3-트라이메틸암모늄-프로페인(dioleoyl-3-trimethylammonium-propane, DOTAP), 비스{2-[N-메틸-N-(a-D-토코페롤헤미석시네이트프로필)아미노]에틸}다이설파이드(SS-33/3AP05), 비스{2-[4-(a-D-토코페롤헤미석시네이트에틸)피페리딜]에틸}다이설파이드(SS33/4PE15), 비스{2-[4-(시스-9-옥타데세노에이트에틸)-1-피페리딘일]에틸} 다이설파이드(SS18/4PE16), 및 비스{2-[4-(시스,시스-9,12-옥타데카다이에노에이트에틸)-1-피페리딘일]에틸}다이설파이드(SS18/4PE13)를 포함한다. 추가의 구현예에서, 지질 나노입자는 또한 하나 이상의 비양이온성 지질 및 지질 접합체를 포함한다.

[0473] [0459] 일부 구현예에서, 양이온성 지질의 몰 농도는 총 지질 몰 농도의 약 20% 내지 약 80%, 약 30% 내지 약 70%, 약 40% 내지 약 60%, 약 45% 내지 약 55%, 또는 약 20%, 약 25%, 약 30%, 약 35%, 약 40%, 약 45%, 약 50%, 약 55%, 약 60%, 약 65%, 약 70%, 약 75%, 또는 약 80%이되, 총 지질 몰 농도는 양이온성 지질, 비양이온성 지질, 및 지질 접합체 몰 농도의 합이다. 특정 구현예에서, 지질 나노입자는 약 1 내지 약 20, 약 2 내지 약 16, 약 4 내지 약 12, 약 6 내지 약 10, 또는 약 1, 약 2, 약 3, 약 4, 약 5, 약 6, 약 7, 약 8, 약 9, 약 10,

약 11, 약 12, 약 13, 약 14, 약 15, 약 16, 약 17, 약 18, 약 19, 또는 약 20의 폴리뉴클레오타이드 중 임의의 것에 대한 양이온성 지질의 몰비를 포함한다.

[0474] [0460] 일부 구현예에서, 지질 나노입자는 적어도 하나의 비양이온성 지질을 포함할 수 있다. 특정 구현예에서, 비양이온성 지질의 몰 농도는 총 지질 몰 농도의 약 20% 내지 약 80%, 약 30% 내지 약 70%, 약 40% 내지 약 70%, 약 40% 내지 약 60%, 약 46% 내지 약 50%, 또는 약 20%, 약 25%, 약 30%, 약 35%, 약 40%, 약 45%, 약 48.5%, 약 50%, 약 55%, 약 60%, 약 65%, 약 70%, 약 75%, 또는 약 80%이다. 비양이온성 지질은, 일부 구현예에서, 인지질 및 스테로이드를 포함한다.

[0475] [0461] 일부 구현예에서, 본원에 기재된 지질 나노입자에 유용한 인지질은 1,2-다이스테아로일-sn-글리세로-3-포스포콜린(Distearoyl-sn-glycero-3-phosphocholine, DSPC), 1,2-다이데카노일-sn-글리세로-3-포스포콜린(Didecanoyl-sn-glycero-3- phosphocholine, DDPC), 1,2-다이에루코일-sn-글리세로-3-포스페이트(소듐 염)(DEPA-NA), 1,2-다이에루코일-sn-글리세로-3-포스포콜린(다이에루코일-sn-글리세로-3-포스포콜린, DEPC), 1,2-다이에루코일-sn-글리세로-3-포스포에탄올아민(Dierucoyl-sn-glycero-3- phosphoethanolamine, DEPE), 1,2-다이에루코일-sn-글리세로-3[포스포-라스-(1-글리세롤)(소듐 염) (DEPG-NA), 1,2-다이리놀레노일-sn-글리세로-3-포스포콜린(Dilinoleoyl-sn-glycero-3-phosphocholine, DLOPC), 1,2-다이라우로일-sn-글리세로-3-포스페이트(소듐 염)(DLPA-NA), 1,2-다이라우로일-sn-글리세로-3-포스포콜린(DLPC), 1,2-다이라우로일-sn-글리세로-3-포스포콜린(DLPC), 1,2-다이라우로일-sn-글리세로-3-포스포에탄올아민(1,2-Dilauroyl-sn-glycero-3-phosphoethanolamine, DLPE), 1,2-다이라우로일-sn-글리세로-3[포스포-라스-(1-글리세롤...)(소듐 염) (DLPG-NA), 1,2-다이라우로일-sn-글리세로-3[포스포-라스-(1-글리세롤)(암모늄 염) (DLPG-NH4), 1,2-다이라우로일-sn-글리세로-3-포스포세린(소듐 염)(DLPS-NA), 1,2-다이미리스토일-sn-글리세로-3-포스페이트(소듐 염)(DMPA-NA), 1,2-다이미리스토일-sn-글리세로-3-포스포콜린(Dimyristoyl-sn-glycero-3-phosphocholine, DMPC), 1,2-다이미리스토일-sn-글리세로-3-포스포에탄올아민(Dimyristoyl-sn-glycero-3-phosphoethanolamine, DMPE), 1,2-다이미리스토일-sn-글리세로-3[포스포-라스-(1-글리세롤)(소듐 염)(DMPG-NA), 1,2-다이미리스토일-sn-글리세로-3[포스포-라스-(1-글리세롤)(암모늄 염)(DMPG-NH4), 1,2-다이미리스토일-sn-글리세로-3[포스포-라스-(1-글리세롤)(소듐/암모늄 염)(DMPG-NH4/NA), 1,2-다이미리스토일-sn-글리세로-3-포스포세린(소듐 염)(DMPS-NA), 1,2-다이올레오일-sn-글리세로-3-포스페이트(소듐 염)(DOPA-NA), 1,2-다이올레오일-sn-글리세로-3-포스포콜린(Dioleoyl-sn- glycero-3-phosphocholine, DOPC), 1,2-다이올레오일-sn-글리세로-3-포스포에탄올아민(Dioleoyl-sn-glycero-3-phosphoethanolamine, DOPE), 1,2-다이올레오일-sn-글리세로-3[포스포-라스-(1-글리세롤)(소듐 염)(DOPG-NA), 1,2-다이올레오일-sn-글리세로-3-포스포세린(소듐 염)(DOPS-NA), 1,2-다이팔미토일-sn-글리세로-3-포스페이트(소듐 염)(DPPA-NA), 1,2-다이팔미토일-sn-글리세로-3-포스포콜린(Dipalmitoyl-sn-glycero-3-phosphocholine, DPPC), 1,2-다이팔미토일-sn-글리세로-3-포스포에탄올아민(Dipalmitoyl-sn-glycero-3-phosphoethanolamine, DPPE), 1,2-다이팔미토일-sn-글리세로-3[포스포-라스-(1-글리세롤)(소듐 염)(DPPG-NA), 1,2-다이팔미토일-sn-글리세로- 3[포스포-라스-(1-글리세롤)(암모늄 염)(DPPG-NH4), 1,2-다이팔미토일-sn-글리세로-3-포스포세린(소듐 염)(DPPS-NA), 1,2-다이스테아로일-sn-글리세로-3-포스페이트(소듐 염)(DSPA-NA), 1,2-다이스테아로일-sn-글리세로-3-포스포에탄올아민(Distearoyl-sn-glycero-3-phosphoethanolamine, DSPE), 1,2-다이스테아로일-sn-글리세로-3-[포스포-라스-(1-글리세롤)(소듐 염)(DSPG-NA), 1,2-다이스테아로일-sn-글리세로-3-[포스포-라스-(1-글리세롤)(암모늄 염)(DSPG-NH4), 1,2-다이스테아로일-sn-글리세로-3-포스포세린(소듐 염)(DSPS-NA), 난-PC[Egg-PC, EPC], 수소화된 난 PC[Hydrogenated Egg PC, HEPC], 수소화된 대두 PC[Hydrogenated Soy PC, HSPC], 1-미리스토일-sn-글리세로-3-포스포콜린(LYSOPCMYRISTIC) , 1-팔미토일-sn-글리세로-3-포스포콜린(LYSOPCPALMITIC), 1-스테아로일-sn-글리세로-3-포스포콜린(LYSOPC STEARIC), 1-미리스토일-2-팔미토일-sn-글리세로-3-포스포콜린(Myristoyl-2-palmitoyl-sn-glycero3-phosphocholine, MPPC), 1-미리스토일-2-스테아로일-sn-글리세로-3-포스포콜린(Myristoyl-2-stearoyl-sn-glycero-3-phosphocholine, MSPC), 1-팔미토일-2-미리스토일-sn-글리세로-3-포스포콜린(Palmitoyl-2-myristoyl-sn-glycero-3-phosphocholine, PMPC), 1-팔미토일-2-올레오일-sn-글리세로-3-포스포콜린(Palmitoyl-2- oleoyl-sn-glycero-3-phosphocholine, POPC), 1-팔미토일-2-올레오일-sn-글리세로-3-포스포에탄올아민(Palmitoyl-2-oleoyl-sn-glycero-3- phosphoethanolamine, POPE), 1-팔미토일-2-올레오일-sn-글리세로-3[포스포-라스-(1-글리세롤)](소듐 염)(POPG-NA), 1-팔미토일-2-스테아로일-sn-글리세로-3-포스포콜린(Palmitoyl-2-stearoyl-sn-glycero-3-phosphocholine, PSPC), 1-스테아로일-2-미리스토일-sn-글리세로-3-포스포콜린(Stearoyl-2-myristoyl-sn-glycero-3-phosphocholine, SMPC), 1-스테아로일-2-올레오일-sn-글리세로-3-포스포콜린(Stearoyl-2-oleoyl- sn-glycero-3-phosphocholine, SOPC), 및 1-스테아로일-2-팔미토일-sn-글리세로-3-포스포콜린(Stearoyl-2-palmitoyl-sn-glycero-3- phosphocholine, SPPC)을 포함하지만 이에 제한되지

않는다. 특정 구현예에서, 인지질은 DSPC이다. 특정 구현예에서, 인지질은 DOPE이다. 특정 구현예에서, 인지질은 DOPC이다.

[0476] [0462] 일부 구현예에서, 지질 나노입자에 포함되는 비양이온성 지질은 하나 이상의 스테로이드를 포함한다. 본원에 기재된 지질 나노입자에 유용한 스테로이드는 콜레스테롤과 같은 콜레스테인, 콜산과 같은 콜레인, 프로게스테론과 같은 프레그네인, 테스토스테론과 같은 안드로스테인, 및 에스트라다이올과 같은 에스트레인을 포함하지만 이에 제한되지 않는다. 추가적인 스테로이드는 콜레스테롤(양), 콜레스테롤 설페이트, 데스모스테롤-d6, 콜레스테롤-d7, 라토스테롤-d7, 데스모스테롤, 스티그마스테롤, 라노스테롤, 디하이드로콜레스테롤, 다이하이드로라노스테롤, 자이모스테롤, 라토스테롤, 자이모스테롤-d5, 14-디메틸-라노스테롤, 14-디메틸-라노스테롤-d6, 8(9)-디하이드로콜레스테롤, 8(14)-디하이드로콜레스테롤, 다이오스게닌, DHEA 설페이트, DHEA, 라노스테롤-d6, 다이하이드로라노스테롤-d7, 캄페스테롤-d6, 시토스테롤, 라노스테롤-95, 디하이드로 FF-MAS-d6, 자이모스테놀-d7, 자이모스테놀, 시토스탄올 캄페스탄올, 캄페스테롤, 7-디하이드로데스모스테롤, 프레그네놀론, 시토스테롤-d7, 디하이드로 T-MAS, 델타 5-아베나스테롤, 브라시카스테롤, 다이하이드로 FF-MAS, 24-메틸렌 콜레스테롤, 콜산 유도체, 콜레스테르일 에스터, 및 글리코실화된 스테롤을 포함하지만 이에 제한되지 않는다. 특정 구현예에서, 지질 나노입자는 콜레스테롤을 포함한다.

[0477] [0463] 일부 구현예에서, 지질 나노입자는 지질 접합체를 포함한다. 이러한 지질 접합체는 세라마이드 PEG 유도체, 예컨대 C8 PEG2000 세라마이드, C16 PEG2000 세라마이드, C8 PEG5000 세라마이드, C16 PEG5000 세라마이드, C8 PEG750 세라마이드, 및 C16 PEG750 세라마이드, 포스포에탄올아민 PEG 유도체, 예컨대 16:0 PEG5000PE, 14:0 PEG5000 PE, 18:0 PEG5000 PE, 18:1 PEG5000 PE, 16:0 PEG3000 PE, 14:0 PEG3000 PE, 18:0 PEG3000 PE, 18:1 PEG3000 PE, 16:0 PEG2000 PE, 14:0 PEG2000 PE, 18:0 PEG2000 PE, 18:1 PEG2000 PE 16:0 PEG1000 PE, 14:0 PEG1000 PE, 18:0 PEG1000 PE, 18:1 PEG 1000 PE, 16:0 PEG750 PE, 14:0 PEG750 PE, 18:0 PEG750 PE, 18:1 PEG750 PE, 16:0 PEG550 PE, 14:0 PEG550 PE, 18:0 PEG550 PE, 18:1 PEG550 PE, 16:0 PEG350 PE, 14:0 PEG350 PE, 18:0 PEG350 PE, 및 18:1 PEG350, 스테롤 PEG 유도체, 예컨대 콜-PEG600, 및 글리세롤 PEG 유도체, 예컨대 DMG-PEG5000, DSG-PEG5000, DPG- PEG5000, DMG-PEG3000, DSG-PEG3000, DPG-PEG3000, DMG-PEG2000, DSG- PEG2000, DPG-PEG2000, DMG-PEG1000, DSG-PEG1000, DPG-PEG1000, DMG- PEG750, DSG-PEG750, DPG-PEG750, DMG-PEG550, DSG-PEG550, DPG-PEG550, DMG-PEG350, DSG-PEG350, 및 DPG-PEG350를 포함하지만 이에 제한되지 않는다. 일부 구현예에서, 지질 접합체는 DMG-PEG이다. 일부 특정 구현예에서, 지질 접합체는 DMG-PEG2000이다. 일부 특정 구현예에서, 지질 접합체는 DMG-PEG5000이다.

[0478] [0464] 지질 나노입자를 포함하는 양이온성 지질, 비양이온성 지질 및/또는 지질 접합체를 선택하는 것뿐만 아니라, 예컨대 선택된 지질(들)의 특성, 의도된 표적 세포로의 전달 성질, 및 전달될 핵산 및/또는 단백질의 특성에 기초하여 이러한 지질 서로에 대한 상대적 몰비를 선택하는 것은 당업자의 수준 내에 있다. 추가적인 고려사항은, 예를 들어 알킬 사슬의 포화뿐만 아니라 선택된 지질(들)의 크기, 전하, pH, pKa, 융합성 및 독성을 포함한다. 따라서, 각 개별 성분의 몰비는 이에 따라 조절될 수 있다.

[0479] [0465] 방법에 사용하기 위한 지질 나노입자는 당업자에게 알려진 다양한 기술에 의해 제조될 수 있다. 핵산 지질 입자 및 제조 방법은, 예를 들어 미국 특허 공개 제20040142025호 및 제20070042031호에 개시되어 있다.

[0480] [0466] 일부 구현예에서, 지질 나노입자는 크기가 약 25 내지 약 500 nm 범위 내일 것이다. 일부 구현예에서, 지질 나노입자는 크기가 약 50 nm 내지 약 300 nm 또는 약 60 nm 내지 약 120 nm이다. 지질 나노입자의 크기는 문헌[Bloomfield, Ann. Rev. Biophys. Bioeng., 10:421A150 (1981)]에 기재된 준전기 광 산란[quasi-electric light scattering, QELS]에 의해 결정될 수 있다. 특정 크기 범위의 지질 나노입자 집단을 생성하기 위한 다양한 방법, 예를 들어 초음파 분해 또는 균질화가 당업계에 알려져 있다. 이러한 방법 중 하나가 미국 특허 제4,737,323호에 기재되어 있다.

[0481] [0467] 일부 구현예에서, 지질 나노입자는 지질 나노입자를 표적화된 세포, 예컨대 본원에 기재된 임의의 세포에 선택적으로 결합시키는 지질 나노입자의 표면에 고정된, 예를 들어 표적화 리간드와 같은 세포 표적화 분자(예를 들어, 항체, scFv 단백질, DART 분자, 펩타이드, 압타머 등)를 포함한다.

[0482] [0468] 일부 구현예에서, 벡터는 간 세포 및/또는 간세포 친화성을 나타낸다.

[0483] [0469] 일부 측면에서, 본원에 기재된 임의의 벡터 및 본원에 기재된 임의의 DNA-표적화 시스템, 본원에 기재된 임의의 gRNA, 본원에 기재된 임의의 융합 단백질, 또는 전술한 것들 중 임의의 것의 부분 또는 성분의 추가적인 부분 또는 추가적인 성분을 인코딩하는 하나 이상의 추가적인 폴리뉴클레오타이드를 포함하는 하나 이상의 추가

적인 벡터를 포함하는 복수의 벡터가 본원에 제공된다.

[0484] [0470] 본원에 기재된 폴리뉴클레오타이드 중 임의의 것을 포함하는 제1 벡터, 본원에 기재된 폴리뉴클레오타이드 중 임의의 것을 포함하는 제2 벡터, 및 임의로 본원에 기재된 폴리뉴클레오타이드 중 임의의 것을 포함하는 하나 이상의 추가적인 벡터를 포함하는 복수의 벡터가 제공된다.

[0485] [0471] 일부 측면에서, 본원에 제공된 벡터는 전달 비히클로 지칭될 수 있다. 일부 측면에서, 본원에 개시된 임의의 DNA-표적화 시스템, 이의 성분, 또는 폴리뉴클레오타이드는 세포로의 전달을 위해 전달 비히클 내로 또는 이의 표면에 패키징될 수 있다. 고려되는 전달 비히클은, 나노구, 리포솜, 양자점, 나노입자, 폴리에틸렌 글리콜 입자, 하이드로겔, 및 미셀을 포함하지만 이에 제한되지 않는다. 당업계에 기재되는 바와 같이, 다양한 표적화 모이어티가 이러한 비히클과 원하는 세포 유형 또는 위치의 우선적 상호작용을 증진하기 위해 사용될 수 있다.

[0486] [0472] 숙주 세포 내로 핵산을 도입하는 방법은 당업계에 알려져 있고, 임의의 알려진 방법이 핵산(예를 들어, 발현 작제물)을 세포 내로 도입하는데 사용될 수 있다. 적합한 방법은 바이러스 또는 박테리오파지 감염, 형질도입, 접합, 원형질체 융합, 리포펙션, 전기천공, 인산칼슘 침전, 폴리에틸렌이민(polyethyleneimine, PEI)-매개 형질도입, DEAE-덱스트란 매개 형질도입, 리포솜 매개 형질도입, 입자 건 기술, 인산칼슘 침전, 직접 마이크로-주사, 나노입자-매개 핵산 전달 등을 포함한다. 일부 구현예에서, 조성물은 mRNA 전달 및 리보핵단백질[ribonucleoprotein, RNP] 복합체 전달에 의해 전달될 수 있다. sgRNA와 복합된 DNA-결합 도메인을 포함하는 RNP 복합체의 직접 전달은 세포 내 전사 및 번역에 대한 필요성을 없앨 수 있고, 낮은 전사 및 번역 활성을 갖는 숙주 세포에 대한 견고한 플랫폼을 제공할 수 있다. RNP 복합체는 당업계에 알려진 임의의 방법으로 숙주 세포 내로 도입될 수 있다.

[0487] [0473] 본 개시의 핵산 또는 RNP는 바이러스 유사 입자[virus-like particle, VLP]를 사용하여 숙주에 포함될 수 있다. VLP는 정상 바이러스 벡터 성분, 예컨대 외피 및 캡시드를 함유하지만 바이러스 유전체가 없다. 예를 들어, Cas 및 sgRNA를 발현하는 핵산은 바이러스 벡터 성분, 예컨대 gag에 융합되고 생산자 세포에 도입될 수 있다. sgRNA 발현 벡터를 함유하는 생성된 바이러스 유사 입자는 효율적인 편집을 위해 숙주 세포를 감염시킬 수 있다.

[0488] [0474] 본 개시의 복합체, 폴리펩타이드 및 핵산의 도입은 단백질 형질도입 도메인[protein transduction domain, PTD]에 의해 발생할 수 있다. 인간 면역결핍 바이러스-1 TAT, 단순 포진 바이러스-1 VP22, 드르소필라 안테나피디아 Antp, 및 폴루아르지닌을 포함하는 PTD는 세포막을 가로지르고, 숙주 세포에 진입하고, 복합체, 폴리펩타이드 및 핵산을 세포 내로 전달할 수 있는 펩타이드 서열이다.

[0489] [0475] 본 개시의 복합체, 폴리펩타이드 및 핵산의 세포로의 도입은 예를 들어, WO 2017/193107 A2, WO 2016/123578 A1, WO 2014/152432 A2, WO 2014/093661 A2, WO 2014/093655 A2, 또는 WO 2021/226555 A2에 기재된 바와 같이 바이러스 또는 박테리오파지 감염, 형질도입, 접합, 원형질체 융합, 리포펙션, 전기천공, 뉴클레오펙션, 인산칼슘 침전, 폴리에틸렌이민(PEI) 매개 형질도입, DEAE-덱스트란 매개 형질도입, 리포솜 매개 형질도입, 입자 건 기술, 인산칼슘 침전, 직접 마이크로 주사, 나노입자 매개 핵산 전달 등에 의해 발생할 수 있다.

[0490] [0476] 폴리뉴클레오타이드의 도입을 위한 다양한 방법은 주지되어 있고 제공된 방법 및 조성물과 함께 사용될 수 있다. 예시적인 방법은 바이러스, 예를 들어 레트로바이러스 또는 렌티바이러스, 형질도입, 트랜스포존 및 전기천공을 통하는 것을 포함하는, 본원에 제공된 DNA 표적화 시스템을 인코딩하는 폴리뉴클레오타이드의 전달을 위한 것을 포함한다.

[0491] [0477] 일부 구현예에서, 폴리뉴클레오타이드는 적합한 벡터, 예컨대 발현 벡터 또는 벡터 내로 클로닝될 수 있다. 발현 벡터는 임의의 적합한 재조합 발현 벡터일 수 있고 임의의 적합한 세포를 형질도입 또는 형질도입시키는 데 사용될 수 있다. 적합한 벡터는 증식 및 확장 또는 발현 또는 둘 모두를 위해 설계된 것, 예컨대 플라스미드 및 바이러스를 포함한다.

[0492] [0478] 일부 구현예에서, 벡터는 pUC 시리즈(Fermentas Life Sciences), p블루스크립트 시리즈(Stratagene, LaJolla, Calif.), pET 시리즈(Novagen, Madison, Wis.), pGEX 시리즈(Pharmacia Biotech, Uppsala, Sweden), 또는 pEX 시리즈(Clontech, Palo Alto, Calif.)의 벡터일 수 있다. 일부 구현예에서, 동물 발현 벡터는 pEUK-C1, pMAM 및 pMAMneo(Clontech)를 포함한다. 일부 구현예에서, 바이러스 벡터, 예컨대 렌티바이러스 벡터 또는 레트로바이러스 벡터가 사용된다. 일부 구현예에서, 재조합 발현 벡터는 표준 재조합 DNA 기술을 사용하여 제조될 수 있다. 일부 구현예에서, 적절하다면, 벡터는 벡터가 DNA- 또는 RNA-기반인지 고려하여 벡터가 도입될 숙

주의 유형에 특이적인 조절 서열, 예컨대 전사 및 번역 개시 및 종결 코돈을 포함할 수 있다. 일부 구현예에서, 벡터는 재조합 수용체를 인코딩하는 뉴클레오타이드 서열에 작동가능한 상태로 연결된 비천연 프로모터를 함유할 수 있다. 일부 구현예에서, 프로모터는 비바이러스 프로모터 또는 바이러스 프로모터, 예컨대 거대세포바이러스[CMV] 프로모터, SV40 프로모터, RSV 프로모터, 및 쥐 줄기 세포 바이러스의 긴 말단 반복체에서 발견되는 프로모터일 수 있다. 당업자에게 알려진 다른 프로모터도 고려된다.

[0493] [0479] 일부 구현예에서, 재조합 핵산은 재조합 감염성 바이러스 입자, 예컨대, 예를 들어, 유인원 바이러스 40[SV40], 아데노바이러스, 또는 아데노 연관 바이러스[AAV]에서 유래한 벡터를 사용하여 세포 내로 전달된다. 일부 구현예에서, 재조합 핵산은 재조합 렌티바이러스 벡터 또는 레트로바이러스 벡터, 예컨대 감마 레트로바이러스 벡터를 사용하여 세포(예를 들어, T 세포) 내로 전달된다(예를 들어, 문헌[Koste et al. (2014) Gene Therapy 2014 Apr 3. doi: 10.1038/gt.2014.25; Carlens et al. (2000) Exp Hematol 28(10): 1137-46; Alonso-Camino et al. (2013) Mol Ther Nucl Acids 2, e93; Park et al., Trends Biotechnol. 2011 November 29(11): 550-557] 참조).

[0494] [0480] 일부 구현예에서, 레트로바이러스 벡터는 긴 말단 반복[LTR] 서열, 예를 들어 몰로니 쥐 백혈병 바이러스[Moloney murine leukemia virus, MoMLV], 골수증식성 육종 바이러스[myeloproliferative sarcoma virus, MPSV], 쥐 배아 줄기 세포 바이러스[murine embryonic stem cell virus, MESV], 쥐 줄기 세포 바이러스[murine stem cell virus, MSCV], 비장 병소 형성 바이러스[spleen focus forming virus, SFFV], 또는 아데노 연관 바이러스[AAV]에서 유래하는 레트로바이러스 벡터를 갖는다. 대부분의 레트로바이러스 벡터는 쥐 레트로바이러스에서 유래한다. 일부 구현예에서, 레트로바이러스는 임의의 조류 또는 포유류 세포 공급원에서 유래한 것을 포함한다. 레트로바이러스는 일반적으로 양쪽지향성이고, 이는 인간을 포함한 여러 종의 숙주 세포를 감염시킬 수 있음을 의미한다. 일 구현예에서, 발현될 유전자는 레트로바이러스 gag, pol 및/또는 env 서열을 대체한다. 다수의 예시적인 레트로바이러스 시스템이 기재되어 있다(예를 들어, 미국 특허 번호 5,219,740, 6,207,453, 5,219,740, 문헌[Miller and Rosman (1989) BioTechniques 7:980-990; Miller, A. D. (1990) Human Gene Therapy 1:5-14; Scarpa et al. (1991) Virology 180:849-852; Burns et al. (1993) Proc. Natl. Acad. Sci. USA 90:8033-8037; 및 Boris-Lawrie and Temin (1993) Cur. Opin. Genet. Develop. 3: 102-109]에 기재되어 있다.

[0495] [0481] 일부 구현예에서, 벡터는 렌티바이러스 벡터이다. 일부 구현예에서, 렌티바이러스 벡터는 통합효소 결핍 렌티바이러스 벡터이다. 일부 구현예에서, 렌티바이러스 벡터는 재조합 렌티바이러스 벡터이다. 일부 구현예에서, 렌티바이러스는 원하는 진화성을 위해(예를 들어, 간세포 향성(tropism)을 위해) 선택되거나 조작된다. 렌티바이러스 생성, 형질도입 및 조작 방법은, 예를 들어 문헌 [Kasaraneni, N. et al. Sci. Rep. 8(1):10990 (2018), Ghaleh, H.E.G. et al. Biomed. Pharmacother. 128:110276 (2020), 및 Milone, M.C. et al. Leukemia. 32(7):1529-1541 (2018)]에 기재된 바와 같이 알려져 있다. 렌티바이러스 형질도입을 위한 추가적인 방법이, 예를 들어 문헌[Wang et al. (2012) J. Immunother. 35(9): 689-701; Cooper et al. (2003) Blood. 101: 1637- 1644; Verhoeyen et al. (2009) Methods Mol Biol. 506: 97-114; 및 Cavalieri et al. (2003) Blood. 102(2): 497-505]에 기재되어 있다.

[0496] [0482] 일부 구현예에서, 재조합 핵산은 전기천공법을 통해 세포(예를 들어, T 세포) 내로 전달된다(예를 들어, 문헌[Chicaybam et al, (2013) PLoS ONE 8(3): e60298 및 Van Tedeloo et al. (2000) Gene Therapy 7(16): 1431-1437] 참조). 일부 구현예에서, 재조합 핵산은 전위를 통해 세포 내로 전달된다(예를 들어, 문헌[Manuri et al. (2010) Hum Gene Ther 21(4): 427-437; Sharma et al. (2013) Molec Ther Nucl Acids 2, e74; 및 Huang et al. (2009) Methods Mol Biol 506: 115-126] 참조). 세포 내로 유전 물질을 도입하고 발현시키는 다른 방법은 인산칼슘 형질도입(예를 들어, 문헌[Current Protocols in Molecular Biology, John Wiley& Sons, New York N.Y.]]에 기재된 바와 같음), 원형질체 융합, 양이온성 리포솜 매개 형질도입, 텅스텐 입자 촉진 미세입자 충격(문헌[Johnston, Nature, 346: 776-777 (1990)], 및 인산 스트론튬 DNA 공동 침전(문헌[Brash et al., Mol. Cell Biol., 7: 2031-2034 (1987)])을 포함한다.

[0497] **III. 약학 조성물 및 제제**

[0498] [0483] 일부 측면에서, 본원에 기재된 임의의 DNA-표적화 시스템 또는 이를 인코딩하는 폴리뉴클레오타이드 또는 벡터 중 임의의 것을 포함하는, 투여를 위한 조성물, 예컨대 약학 조성물 및 제제가 본원에 제공된다. 일부 측면에서, 약학 조성물은 본원에 제공된 하나 이상의 DNA-표적화 시스템 또는 이의 성분을 함유한다. 일부 측면에서, 약학 조성물은 본원에 제공된 DNA-표적화 시스템의 하나 이상의 성분을 인코딩하는 폴리뉴클레오타이드를

함유하는 하나 이상의 벡터, 예컨대 바이러스 벡터를 포함한다. 이러한 조성물은 제공된 방법 및/또는 제공된 제조 물품 또는 조성물, 예컨대 질환, 병태 및 장애의 예방 또는 치료 또는 검출, 진단 및 예후 방법에 따라 사용될 수 있다.

[0499] [0484] 용어 '약학적 제제'는 그 안에 함유된 유효 성분의 생물학적 활성이 유효하도록 허용하는 형태로서, 제제가 투여되는 대상에 허용불가능하게 독성인 추가적인 성분을 함유하지 않는 제제를 지칭한다.

[0500] [0485] 또 다른 구현예에서, 약학 조성물은 약학적으로 허용되는 부형제를 더 포함할 수 있다. 약학적으로 허용되는 부형제는 비히클, 보조제, 담체, 또는 희석제로서 기능적 분자일 수 있다.

[0501] [0486] '약학적으로 허용되는 담체'는 유효 성분 외에 대상에 비독성인 약학적 제제의 성분을 지칭한다. 약학적으로 허용되는 담체는 완충액, 부형제, 안정제 또는 방부제를 포함하지만 이에 제한되지 않는다.

[0502] [0487] 일부 측면에서, 담체의 선택은 특정 작용제 및/또는 투여 방법에 의해 부분적으로 결정된다. 따라서, 다양한 적합한 제제가 존재한다. 예를 들어, 약학 조성물은 보존제를 함유할 수 있다. 적합한 보존제는, 예를 들어, 메틸파라벤, 프로필파라벤, 벤조산소듐 및 염화벤잘코늄을 포함할 수 있다. 일부 측면에서, 2종 이상의 방부제의 혼합물이 사용된다. 보존제 또는 이의 혼합물은 일반적으로 총 조성물의 약 0.0001 중량% 내지 약 2 중량%의 양으로 존재한다. 담체는, 예를 들어 문헌[Remington's Pharmaceutical Sciences 16th edition, Osol, A. Ed. (1980)]에 기재된다. 약학적으로 허용되는 담체는 일반적으로 사용된 용량 및 농도에서 수혜자에게 비독성이고, 완충제, 예컨대 포스페이트, 시트레이트 및 기타 유기 산; 아스코르브산 및 메싸이오닌을 포함하는 항산화제; 보존제(예컨대, 옥타데실다이메틸벤질 암모늄 클로라이드; 헥사메토늄 클로라이드; 벤잘코늄 클로라이드; 벤제토늄 클로라이드; 페놀, 뷰틸 또는 벤질 알코올; 알킬 파라벤, 예컨대 메틸 또는 프로필 파라벤; 카테콜; 레조시놀; 사이클로헥산올; 3-펜탄올; 및 m-크레솔); 저분자량(약 10개 미만의 잔기) 폴리펩타이드; 단백질, 예컨대 혈청 알부민, 젤라틴 또는 면역글로불린; 친수성 중합체, 예컨대 폴리비닐피롤리돈; 아미노산, 예컨대 글라이신, 글루타민, 아스파라진, 히스티딘, 아르기닌 또는 라이신; 단당류, 이당류 및 기타 탄수화물, 예를 들어 글루코스, 만노스 또는 덱스트린; 킬레이트제, 예컨대 EDTA; 당, 예컨대 수크로스, 만니톨, 트레할로스 또는 소비톨; 염 형성 반대 이온, 예컨대 소듐; 금속 복합체(예를 들어, Zn-단백질 복합체); 및/또는 비이온성 계면활성제, 예컨대 폴리에틸렌 글리콜(polyethylene glycol, PEG)을 포함하지만 이에 제한되지 않는다.

[0503] [0488] 일부 구현예에서, 약학적으로 허용되는 부형제는 형질도입 촉진제일 수 있고, 이는 계면 활성제, 예컨대 면역 자극 복합체[immune-stimulating complex, ISCOM], 프로인트 불완전 보조제, 모노포스포릴 지질 A를 포함하는 LPS 유사체, 뮤라밀 펩타이드, 퀴논 유사체, 소포 예컨대 스쿠알렌 및 스쿠알렌, 하이알루론산, 지질, 리포솜, 칼슘 이온, 바이러스 단백질, 다가음이온, 다가양이온, 또는 나노입자, 또는 기타 알려진 형질도입 촉진제를 포함할 수 있다.

[0504] [0489] 일부 구현예에서, 형질도입 촉진제는 다가음이온, 폴리-L-글루타메이트(LGS)를 포함하는 다가양이온 또는 지질이다. 일부 구현예에서, 형질도입 촉진제는 폴리-L-글루타메이트이다. 일부 구현예에서, 형질도입 촉진제는 또한 계면 활성제, 예컨대 면역 자극 복합체[ISCOM], 프로인트 불완전 보조제, 모노포스포릴 지질 A를 포함하는 LPS 유사체, 뮤라밀 펩타이드, 퀴논 유사체 및 소포, 예컨대 스쿠알렌 및 스쿠알렌, 및 하이알루론산을 포함할 수 있고, 하이알루론산은 또한 유전적 작제물과 함께 투여되어 사용될 수 있다. 일부 구현예에서, DNA-표적화 시스템을 인코딩하는 DNA 벡터는 또한 형질도입 촉진제, 예컨대 지질, DNA-리포솜 혼합물(예를 들어, WO9324640 참조)로서 레시틴 리포솜 또는 당업계에 알려진 기타 리포솜을 포함하는 리포솜, 칼슘 이온, 바이러스 단백질, 다가음이온, 다가양이온 또는 나노입자, 또는 기타 알려진 형질도입 촉진제를 포함할 수 있다. 일부 구현예에서, 형질도입 촉진제는 다가음이온, 폴리-L-글루타메이트(LGS)를 포함하는 다가양이온 또는 지질이다.

[0505] [0490] 일부 구현예에서 조성물은 멸균 액체 제제, 예를 들어 등장성 수용액, 현탁액, 유제, 분산제 또는 점성 조성물로서 제공되고, 이는 일부 측면에서 선택된 pH로 완충될 수 있다. 액체 제제는 일반적으로 젤, 기타 점성 조성물 및 고체 조성물보다 제조하기가 더 용이하다. 또한, 액체 조성물은 특히 주사로 투여하기에 다소 더 편리하다. 한편, 점성 조성물은 특정 조직과의 더 긴 접촉 기간을 제공하기 위해 적절한 점도 범위 내에서 제제화될 수 있다. 액체 또는 점성 조성물은 담체를 포함할 수 있고, 이는 예를 들어 물, 식염수, 인산 완충 식염수, 폴리올(예를 들어 글리세롤, 프로필렌 글리콜, 액체 폴리에틸렌 글리콜) 및 이들의 적절한 혼합물을 함유하는 용매 또는 분산 매질일 수 있다.

[0506] [0491] 멸균 주사용 용액은 적절한 담체, 희석제, 또는 부형제, 예컨대 멸균수, 생리 식염수, 글루코스, 덱스트로스 등과의 혼합물과 같은 용매에 작용제를 포함함으로써 제조될 수 있다. 생체 내 또는 생체 외 투여용으로

사용되는 제제는 일반적으로 멸균성이다. 멸균은 예를 들어 멸균 여과막을 통한 여과에 의해 쉽게 이루어질 수 있다.

[0507] [0492] 일부 구현예에서, 약학 조성물은 질환 또는 병태를 치료 또는 예방하기에 효과적인 양, 예컨대 치료 유효량 또는 예방적 유효량의 성분을 함유한다. 일부 구현예에서 치료 또는 예방 효능은 치료 받은 대상의 주기적 평가로 모니터링된다. 며칠 또는 더 장기간에 걸쳐 반복된 투여를 위해, 병태에 따라 치료는 질환 증상의 원하는 억제가 발생할 때까지 반복된다. 그러나, 다른 용량 용법이 유용할 수 있다. 원하는 투여량은 조성물의 단일 볼루스 투여에 의해, 조성물의 다중 볼루스 투여에 의해, 또는 조성물의 연속 주입 투여에 의해 전달될 수 있다.

[0508] [0493] 일부 구현예에서, 조성물은 임의의 적합한 수단에 의해, 예를 들어 볼루스 주입에 의해 또는 주사에 의해, 예를 들어 정맥 내 또는 피하 주사에 의해 대상에게 투여될 수 있다. 일부 구현예에서, 주어진 용량은 조성물의 단일 볼루스 투여에 의해 투여된다. 일부 구현예에서, 조성물은 조성물의 다중 볼루스 투여에 의해, 예를 들어, 3일 이하의 기간에 걸쳐, 또는 조성물의 연속 주입 투여에 의해 투여된다. 일부 구현예에서, 조성물은 비경구, 예를 들어 정맥 내, 근육 내, 피하 또는 복강 내 투여에 의해 투여된다. 일부 구현예에서, 조성물은 정맥 내, 복강 내, 또는 피하 주사에 의한 말초 전신 전달을 사용하여 대상에게 투여된다.

[0509] [0494] 질환의 예방 또는 치료를 위해, 적절한 용량은 치료될 질환의 유형, 작용제 또는 작용제들의 유형, 세포 또는 재조합 수용체의 유형, 질환의 중증도 및 과정, 작용제 또는 세포가 예방적 또는 치료적 목적을 위해 투여되는지의 여부, 이전의 요법, 환자의 임상 이력 및 작용제 또는 세포에 대한 반응 및 담당의사의 재량에 의존할 것이다. 조성물은 일부 구현예에서 한 번에 또는 일련의 치료에 걸쳐 환자에게 적합하게 투여된다.

[0510] **IV. 치료 방법**

[0511] [0495] 예를 들어, 본원에 기재된 약학 조성물과 같은 임의의 조성물을 투여하는 것을 포함하는 치료 방법이 본원에 제공된다. 일부 측면에서, 본원에 기재된 임의의 조성물을 대상, 예컨대 질환 또는 장애를 갖는 대상에게 투여하는 방법이 또한 제공된다. 본원에 기재된 조성물, 예컨대 약학 조성물은 다양한 치료적, 진단적 및 예방적 징후에서 유용하다. 예를 들어, 조성물은 대상(예컨대 인간)의 다양한 질환 및 장애를 치료하는 데 유용하다. 이러한 방법 및 용도에는 예를 들어, HBV 바이러스 감염과 같은 질병, 병태 또는 장애가 있거나 HBV 바이러스 감염과 관련된 대상에게 조성물을 투여하는 것을 포함하는 치료 방법 및 용도가 포함된다. 특정 구현예에서, 대상은 B형 간염 바이러스 감염 또는 B형 간염 바이러스 감염으로 인한 간 질환으로 진단되었다. 일부 구현예에서, HBV 바이러스 감염 또는 관련 HBV 바이러스 감염은 D형 간염 바이러스 감염, 델타 간염, 급성 B형 간염, 급성 급성 B형 간염, 만성 B형 간염, 간 섬유증, 말기 간 질환 또는 간세포암과 같은 암에서 선택된다. 조성물은 질병 또는 장애의 치료를 위해 효과적인 양으로 투여된다. 용도에는 이러한 방법 및 치료에서 조성물의 용도와 이러한 치료 방법을 수행하기 위한 약물의 제조가 포함된다. 일부 구현예에서, 방법은 질병 또는 상태가 있거나 의심되는 대상에게 조성물을 투여하여 수행된다. 일부 구현예에서, 방법은 그에 따라 대상의 질병 또는 상태 또는 장애를 치료한다. 또한, 세포 및 조성물을 대상, 예를 들어 환자에게 투여하기 위한 치료 방법이 제공된다.

[0512] [0496] 일부 구현예에서, 조성물은 본원에 제공된 DNA-표적화 시스템 또는 이를 인코딩하는 폴리뉴클레오타이드 또는 벡터를 포함하는데, 여기서 조성물을 대상에게 전달하면 하나 이상의 HBV 유전자 및/또는 이의 조절 요소의 전사가 억제되어 대상에서 HBV 복제 및/또는 HBV 전사가 침묵하여 만성 B형 간염 감염을 포함한 질병 또는 상태를 치료한다. 일부 구현예에서, 본원에 제공된 DNA-표적화 시스템 또는 이를 인코딩하는 폴리뉴클레오타이드 또는 벡터를 포함하는 조성물을 투여하거나 사용하면 B형 간염 바이러스 복제 및/또는 HBV 전사와 관련된 하나 이상의 유전자 및/또는 이의 조절 요소의 발현이 감소한다. 일부 구현예에서, B형 간염 유전자 및/또는 이의 조절 요소는 공유 결합으로 폐쇄된 원형(cccDNA) 형태, 이완된 원형 DNA(rcNDA) 형태로 존재하거나 인간 게놈 DNA에 통합된다.

[0513] [0497] 일부 측면에서, 또한 본원에 제공된 모든 설명에 따라 대상의 만성 B형 간염 감염을 감소시키기 위해 B형 간염 바이러스 DNA 서열의 전사를 감소시키는 방법이 제공된다. 일부 측면에서, B형 간염 바이러스 DNA 서열의 전사를 감소시키면 B형 간염 바이러스 복제 및/또는 전사가 침묵된다. 예를 들어, B형 간염 바이러스 DNA 서열의 전사가 감소하면 HBsAg 전사체 및/또는 총 HBV RNA가 감소한다. B형 간염 바이러스 DNA 서열의 전사가 감소하면 통합 및 cccDNA 형태에서 HBsAg 수준이 감소하고/하거나 cccDNA 형태에서 HBcrAg 수준이 감소할 수도 있다.

[0514] [0498] 일부 구현예에서, 본원에서 제공된 바와 같이 DNA-표적화 시스템 또는 이를 인코딩하는 폴리뉴클레오타이드 또는 벡터를 함유하는 조성물을 대상에게 투여하는 방법은 생체내(즉, 대상 내에서)에서 수행된다.

[0515] [0499] 일부 구현예에서, DNA-표적화 시스템을 투여하는 방법 및 용도는 DNA-표적화 시스템을 간 세포(예컨대 간세포)에 전달하는 결과를 가져온다. 일부 구현예에서, 후성적 변형 DNA-표적화 시스템(또는 간 세포에 전달하기 위한 폴리뉴클레오타이드 또는 벡터 또는 상기 중 임의의 것의 조성물)을 대상에게 투여하는 방법은 DNA-표적화 시스템을 간 세포 또는 간 세포 집단과 접촉시키는 것이다. 일부 구현예에서, 접촉에 의해 후성적 변형 DNA-표적화 시스템(또는 간 세포에 전달하기 위한 폴리뉴클레오타이드 또는 벡터 또는 상기 중 임의의 것의 조성물)이 간 세포에 도입되며, 예를 들어 간 세포의 핵으로 이동하거나 국소화할 수 있는 곳에서 그러하다. 일부 구현예에서, 이 방법은 간 세포 또는 집단의 하나 이상의 간 세포에서 HBV 감염을 치료한다.

[0516] [0500] 일부 구현예에서, 간 세포로의 전달은 HBV 유전자 또는 조절 요소, 또는 DNA-표적화 시스템에 의해 표적되는 유전자 또는 조절 요소의 조합에서 후성적 변화를 초래한다. 일부 구현예에서, 후성적 변화는 다음 중 하나 이상의 변화를 포함한다: DNA 접근성, 히스톤 메틸화, 아세틸화, 인산화, 유비퀴틴화, 수모일화, 리보실화, 시트룰린화 및 DNA 메틸화. 일부 구현예에서, 후성적 변화는 본원에서 기술된 바와 같이 표적 유전자 또는 이의 조절 요소의 표적 부위의 변형된 DNA 메틸화이다. 일부 구현예에서, 후성적 변화는 본원에서 기술된 바와 같이 표적 유전자 또는 이의 조절 요소의 표적 부위의 히스톤 변형이다. 일부 구현예에서, DNA-표적화 시스템을 간 세포(예컨대 간세포)로 전달하면 대상의 간에서 HBV 유전자 및/또는 그 조절 요소의 전사가 감소한다.

[0517] [0501] 일부 구현예에서, HBV 에피게놈의 변형은 제공된 후성적 변형 DNA-표적화 시스템으로 본원에 설명된 바와 같이 유전자 및/또는 그 조절 요소의 조합을 표적화하여 HBV의 에피게놈을 변경함으로써 이루어진다. 일부 구현예에서, 변형된 HBV는 폴리머라제 유전자, S-패밀리 유전자, X-유전자 또는 코어 패밀리 유전자, 또는 프로모터, 인핸서 또는 전사 처리 제어 영역을 포함하는 HBC 복제 및/또는 HBV 전사를 제어하는 데 관여하는 유전자 또는 이의 조절 요소에서 또는 그 근처의 표적 부위에서 후성적 변화를 포함한다. 일부 구현예에서, 상기 표적 부위 중 임의의 것의 후성적 변화는 DNA 접근성, 히스톤 메틸화, 아세틸화, 인산화, 유비퀴틴화, 수모일화, 리보실화, 시트룰린화 및 DNA 메틸화 중 적어도 하나의 변화이며, 이는 본원에 설명된 DNA-표적화 시스템과 접촉 또는 도입되지 않은, 즉 방법에 적용되지 않은 비교 가능한 비수정 세포(예컨대 간 세포)와 비교된다.

[0518] [0502] 일부 구현예에서, DNA 결합 시스템의 투여는 간 세포에서 유전자 및/또는 그 조절 요소 또는 유전자 및/또는 그 조절 요소의 조합의 발현을 조절하는데, 예를 들어 섹션 I.A에 기술된 것들이다. 일부 구현예에서, 유전자, 전사체, 또는 유전자와 전사체의 조합, 예를 들어 섹션 I.A에 기술된 것들이 본 방법에 적용되지 않은, 즉 본 명세서에 기술된 DNA-표적화 시스템과 접촉 또는 도입되지 않은 비교 가능한 변형되지 않은 세포(예컨대 간)에 비해 감소된다. 일부 구현예에서, 하나 이상의 유전자의 전사는 적어도 약 1.2배, 1.25배, 1.3배, 1.4배, 1.5배, 1.6배, 1.7배, 1.75배, 1.8배, 1.9배, 2배, 2.5배, 3배, 4배, 5배, 10배, 20배, 30배, 40배, 50배, 100배, 200배, 300배, 400배, 500배, 1000배 이상 감소한다. 일부 구현예에서, 전사체 생성은 적어도 약 1.2배, 1.25배, 1.3배, 1.4배, 1.5배, 1.6배, 1.7 배, 1.75배, 1.8배, 1.9배, 2배, 2.5배, 3배, 4배, 5배, 10배, 20배, 30배, 40배, 50배, 100배, 200배, 300배, 400배, 500배, 1000배 이상 감소한다. 일부 구현예에서, 유전자 및/또는 그 조절 요소의 조합의 전사 감소는 대상에서 HBV 복제 및 단백질 수준을 감소시킨다. 일부 구현예에서, 대상의 간 세포는 폴리머라제 유전자, S-패밀리 유전자, X-유전자 또는 코어-패밀리 유전자의 전사가 감소되도록 조절되었다. 일부 구현예에서, 변형된 간 세포에서 유전자 조합의 각 유전자의 발현은 비교 가능한 변형되지 않은 간 세포에서 동일한 유전자의 발현과 비교하여 적어도 1.2배 이상 감소되며, 예를 들어 약 2배, 3배, 4배, 5배, 6배, 7배, 8배, 9배, 10배 이상 감소된다.

[0519] [0503] 일부 구현예에서, 치료할 질병, 병태 또는 장애는 간에 있다. 일부 구현예에서, 치료할 질병, 병태 또는 장애는 간 질환, 암 또는 바이러스 감염이다. 일부 구현예에서, 치료할 질병, 병태 또는 장애는 간염이다. 일부 구현예에서, 치료할 상태는 간부전이다. 일부 구현예에서, 간부전은 급성 간부전이다. 일부 구현예에서, 치료할 질병 또는 상태는 간경변이다. 일부 구현예에서, 치료할 질병 또는 상태는 암이다. 일부 구현예에서, 암은 고형 종양 암이다. 일부 구현예에서, 암은 간세포암이다. 일부 구현예에서, 간염은 급성 간염이다. 일부 구현예에서, 간염은 만성 간염이다.

[0520] [0504] 일부 구현예에서, 대상은 간 세포에서 HBV 전사체의 수치가 상승하거나, 간 세포 내부 및 표면에서 HBV 단백질의 수치가 상승하거나, HBV 단백질에 대한 순환 항체의 수치가 상승한 것으로 의심된다. 일부 구현예에서, 대상은 간 질환(예컨대 간염), 암(예컨대 간세포성 간암종), 또는 HBV 감염(급성 또는 만성 간염)이 있거나 의심된다.

[0521] [0505] 다중화된 후성적 변형 DNA-표적화 시스템이 대상(예컨대 인간)에게 투여되면, 일부 측면에서 변형된 세포 집단의 생물학적 활성은 여러 가지 알려진 방법 중 하나로 측정된다. 평가할 매개변수에는 간세포에서 감소된 HBV 전사체 수준 및/또는 HBV DNA 수준(예컨대 qRT-PCR/qPCR), 간세포 내부 및 표면에서 감소된 HBV 단백질(예컨대 HBsAg, HBeAg) 수준 (예컨대 형광 염색 및 유세포 분석으로 평가) 및/또는 ELISA를 사용하여 혈청을 분석하여 측정된 HBeAg가 포함된다. 일부 측면에서 생물학적 활성은 임상적 결과를 평가하여 측정된다. 매개변수에 대한 특정 임계값은 본원에 제공된 치료 방법의 효능을 결정하기 위해 설정될 수 있다.

[0522] [0506] 또한, B형 간염 바이러스 서열을 포함하는 세포에서 하나 이상의 유전자의 전사를 감소시키는 방법이 제공된다. 일부 구현예에서, 이 방법은 B형 간염 바이러스 서열에서 CpG 섬의 표적 CpG 메틸화를 유도하는 후성적 변형 DNA-표적화 시스템을 세포에 도입하는 것을 포함한다. 일부 구현예에서, 이 방법은 B형 간염 바이러스 서열에서 CpG 섬 2의 표적 CpG 메틸화를 유도하는 후성적 변형 DNA-표적화 시스템을 세포에 도입하는 것을 포함한다.

[0523] [0507] CpG 섬은 CG 디뉴클레오타이드의 높은 빈도를 포함하는 게놈 영역이다. 따라서 이러한 영역은 일반적으로 약 50%보다 큰 GC 백분율을 갖고 관찰/예상 CpG 비율은 약 60%보다 크다(Gardiner-Garden et al. "CpG islands in vertebrate genomes," J Mol Biol 196: 261-282 (1987)). CpG 섬은 종종 유전자 근처에 위치한다. 메틸화는 정상 세포에 일반적으로 존재하지 않는 부위에서 DNA의 시토신 잔기의 후성적 메틸화를 포함한다. CpG 섬의 메틸화 검출은 본 발명의 방법 및 예를 들어, 당업계에 잘 알려진 핵산 증폭 방법인 메틸화 특이적 폴리머라제 연쇄 반응(MS-PCR)을 포함하되 이에 국한되지 않는 당업계에 알려진 모든 방법에 따라 수행할 수 있다. 이 분석법에서는 DNA 서열의 중아황산염 변형을 통해 메틸화된 대립유전자와 메틸화되지 않은 대립유전자의 차이를 검출할 수 있다. DNA와 중아황산나트륨의 반응은 모든 메틸화되지 않은 시토신을 우라실로 전환하는데, 이는 Taq 폴리머라제에 의해 티민으로 인식되지만 메틸화된 시토신에는 영향을 미치지 않는다. 메틸화 또는 메틸화되지 않은 DNA에 특이적인 프라이머로 증폭하면 메틸화된 DNA와 메틸화되지 않은 DNA를 구별할 수 있다. 이 분석법은 관심 영역에서 시토신의 메틸화를 검출하기 위해 여러 샘플을 조사하는 간단하고 빠른 방법을 제공한다(Widschwendter et al., "Methylation and silencing of the retinoic acid receptor-beta2 gene in breast cancer" J Natl. Cancer Inst. 92(10):826-832 (2000)). DNA 메틸화의 검출 및/또는 정량 분석을 위한 당해 기술 분야에 알려진 다른 방법에는 크로마틴 면역침전 분석법(ChIP)(Mulero-Navarro et al., Carcinogenesis 27:1099-1104 (2006); Nakagawachi et al., Oncogene 22:8835-8844 (2003)); 중아황산염 변형 의존 형광 기반 실시간 PCR 분석법인 MethyLight®(Eads et al. Nucleic Acids Res. 28(8) e32(2000); Erhlich et al., Oncogene 21:6694-6702(2002)); 파이로시퀀싱(Lee et al. Clinical Cancer Research 14:2664-2672 (2008); Dejeux et al., J. Mol. Diagn. 9:510-520 (2007)); 및 RNA 염기 특이적 절단(MassCLEAVE™ 키트)(Sequenom, Inc., San Diego, Calif.)과 함께 MALDI-TOF 질량 분석법을 활용하는 Sequenom® MassARRAY® 시스템(Sequenom, Inc., San Diego, Calif.)이 있으나 이에 한정되지 않는다.

[0524] [0508] 또한, 본 발명에 따라 B형 간염 바이러스 서열을 포함하는 세포에서 하나 이상의 유전자의 전사를 감소시키는 방법이 제공된다. 일부 구현예에서, 이 방법은 서열번호 650의 뉴클레오타이드 위치를 참조로 B형 간염 바이러스 서열에서 67bp - 392bp, 1033bp - 1749bp 또는 2215bp - 2490bp에 해당하는 표적 영역 내의 영역의 표적 CpG 메틸화를 유도하는 후성적 변형 DNA-표적화 시스템을 세포에 도입하는 것을 포함한다.

[0525] [0509] 또한, 본 발명에 따라 B형 간염 바이러스 서열을 포함하는 세포에서 하나 이상의 유전자를 감소시키는 방법이 제공된다. 일부 구현예에서, 이 방법은 서열번호 650의 뉴클레오타이드 위치를 참조로 B형 간염 바이러스 서열에서 1033bp-1749bp에 해당하는 표적 영역 내의 영역의 표적 CpG 메틸화를 유도하는 후성적 변형 DNA-표적화 시스템을 세포에 도입하는 것을 포함한다.

[0526] [0510] 또한, 본 발명은 세포에서 B형 간염 바이러스 감염을 감소시키는 방법을 제공하는데, 이 방법은 B형 간염 바이러스 서열을 포함하는 세포에 B형 간염 바이러스 서열의 CpG 섬의 표적 CpG 메틸화를 유도하는 후성적 변형 DNA-표적화 시스템을 도입하는 단계를 포함한다.

[0527] [0511] 또한, 본 발명에서는 세포에서 B형 간염 바이러스 감염을 감소시키는 방법이 제공되는데, 이 방법은 B형 간염 바이러스 서열을 포함하는 세포에 B형 간염 바이러스 서열의 CpG 섬 2의 표적 CpG 메틸화를 유도하는 후성적 변형 DNA-표적화 시스템을 도입하는 단계를 포함한다.

[0528] [0512] 또한, 본 발명에서는 세포에서 B형 간염 바이러스 감염을 감소시키는 방법이 제공되는데, 이는 B형 간염 바이러스 서열을 포함하는 세포에 서열번호 650의 뉴클레오타이드 위치를 참조로 B형 간염 바이러스 서열의

67bp-392bp, 1033bp-1749bp 또는 2215bp-2490bp에 해당하는 표적 영역 내의 영역의 표적 CpG 메틸화를 유도하는 후성적 변형 DNA-표적화 시스템을 도입하는 것을 포함한다.

[0529] [0513] 또한, 본 발명은 세포에서 B형 간염 바이러스 감염을 감소시키는 방법을 제공하는데, 이 방법은 B형 간염 바이러스 서열을 포함하는 세포에 표적 영역 내의 서열번호 650의 뉴클레오타이드 위치를 참조로 B형 간염 바이러스 서열의 1033bp-1749bp에 상응하는 영역의 표적 CpG 메틸화를 유도하는 후성적 변형 DNA-표적화 시스템을 도입하는 단계를 포함한다.

[0530] [0514] 본원의 구현예 중 어느 하나에서, 후성적 변형 DNA-표적화 시스템은 (a) B형 간염 바이러스 DNA 서열에서 표적 부위를 표적으로 하기 위한 DNA-결합 도메인; 및 (b) DNA 메틸트랜스퍼라제 효과기 도메인을 포함하는 적어도 하나의 효과기 도메인을 포함하는 융합 단백질을 포함하는 적어도 하나의 DNA-표적화 모듈을 포함한다. 일부 구현예에 있어서, CpG 메틸화 영역은 표적 영역의 500개 염기쌍 내에 있다. 일부 구현예에 있어서, 도입은 대상의 생체 내 또는 생체 외에서 발생한다. 일부 구현예에 있어서, 세포는 포유류 세포(예컨대 인간 세포)이다. 일부 구현예에 있어서, 세포는 통합된 HBV DNA를 포함한다. 일부 구현예에 있어서, 세포는 에피솜 HBV cccDNA 풀을 포함하는 간세포이다. 일부 구현예에 있어서, 간세포는 HBV 단백질, 예를 들어 HBsAg 및/또는 HBeAg를 발현한다.

[0531] [0515] 또한, 본 발명은 B형 간염 바이러스 서열의 표적 영역 내에서 후성적 변형을 촉진하는 방법을 제공하는데, 이는 표적 영역 내의 표적 부위를 표적으로 하는 제공된 후성적 변형 DNA-표적화 시스템을 B형 간염 바이러스 서열을 포함하는 HBV에 감염된 세포에 도입하는 것을 포함한다.

[0532] [0516] 또한, 본 발명에서는 대상에서 B형 간염 바이러스 감염을 감소시키는 방법이 제공되는데, 이는 B형 간염 바이러스 서열의 CpG 섬의 CpG 메틸화를 증가시키는 후성적 변형 DNA-표적화 시스템을 B형 간염에 감염된 대상에게 투여하는 것을 포함한다. 일부 구현예에서, 후성적 변형 DNA-표적화 시스템은 (a) B형 간염 바이러스 DNA 서열의 표적 부위를 표적으로 삼기 위한 DNA-결합 도메인; 및 (b) DNA 메틸트랜스퍼라제 효과기 도메인을 포함하는 적어도 하나의 효과기 도메인을 포함한다.

[0533] [0517] 또한, 본 발명에서는 대상에서 간염 바이러스 감염을 감소시키는 방법이 제공되는데, 이는 B형 간염에 감염된 대상에게 B형 간염 바이러스 서열의 표적 영역 내의 영역의 CpG 메틸화를 증가시키는 후성적 변형 DNA-표적화 시스템을 투여하는 것을 포함한다. 일부 구현예에서, 표적 영역은 B형 간염 바이러스 서열의 CpG 섬 1, CpG 섬 2 또는 CpG 섬 3을 포함한다. 일부 구현예에서, 표적 영역은 서열번호 650의 뉴클레오타이드 위치를 참조로 B형 간염 바이러스 서열에서 67bp - 392bp, 1033bp - 1749bp 또는 2215bp - 2490bp에 해당하는 서열 내의 연속적인 뉴클레오타이드 서열을 포함한다. 일부 구현예에서, 표적 영역은 서열번호 650의 뉴클레오타이드 위치를 참조로 B형 간염 바이러스 서열에서 1255bp - 1290bp에 해당하는 서열 내의 연속적인 뉴클레오타이드 서열을 포함한다. 일부 구현예에서, 후성적 변형 DNA-표적 시스템은 (a) B형 간염 바이러스 DNA 서열 내의 표적 부위를 표적으로 하기 위한 DNA-결합 도메인; 및 (b) DNA 메틸트랜스퍼라제 효과기 도메인을 포함하는 적어도 하나의 효과기 도메인을 포함한다.

[0534] [0518] 본원의 임의의 구현예에서, CpG 메틸화 영역은 표적 영역의 약 1000bp, 약 750bp, 약 600bp, 약 500bp, 약 450bp, 약 400bp, 약 350bp, 약 300bp, 약 250bp, 약 200bp, 약 150bp, 약 100bp, 약 50bp 이내에 존재한다.

[0535] **V. 키트 및 제조 물품**

[0536] [0519] 또한, 제공된 구현예를 수행하기에 유용한 제조 물품, 시스템, 장치 및 키트가 제공된다. 일부 구현예에서, 제공된 제조 물품 또는 키트는 본원에 기재된 임의의 DNA-표적화 시스템, 본원에 기재된 임의의 gRNA, 본원에 기재된 임의의 융합 단백질, 본원에 기재된 임의의 폴리뉴클레오타이드, 본원에 기재된 임의의 복수의 폴리뉴클레오타이드, 본원에 기재된 임의의 벡터, 본원에 기재된 임의의 복수의 벡터 또는 전술한 것 중 임의의 일부 또는 성분, 또는 이들의 임의의 조합을 함유한다. 일부 구현예에서, 제조 물품 또는 키트는 제공된 방법을 수행하기에 유용한 폴리펩타이드, 폴리뉴클레오타이드, 핵산, 및/또는 벡터를 포함한다.

[0537] [0520] 일부 구현예에서, 제조 물품 또는 키트는 하나 이상의 용기, 일반적으로 복수의 용기, 포장 재료, 및 일반적으로 사용을 위한 설명서, 예를 들어 도입 또는 투여를 위한 설명서를 포함하는, 용기 또는 용기 및/또는 포장상의 또는 이와 연관된 표지 또는 포장 삽입물을 포함한다.

[0538] [0521] 또한, 예를 들어 요법 또는 치료에 사용하기 위한, 제공된 조성물, 예를 들어 약학 조성물을 투여하는데 유용한 제조 물품, 시스템, 장치, 및 키트가 제공된다. 일부 구현예에서, 본원에 제공된 제조 물품 또는 키

트는 벡터 및/또는 복수의 벡터, 예컨대 본원에 기재된 임의의 벡터 및/또는 복수의 벡터를 함유한다. 일부 측면에서, 본원에 제공된 제조 물품 또는 키트는 벡터 및/또는 복수의 벡터의 투여를 위해 사용될 수 있고, 사용을 위한 설명서를 포함할 수 있다.

[0539] [0522] 치료를 위한 세포 또는 세포 조성물을 함유하는 제조 물품 및/또는 키트는, 예를 들어 용기 및 용기상에 또는 이와 연관된 표지 또는 포장 삽입물을 포함할 수 있다. 적합한 용기는 예를 들어, 병, 바이알, 주사기, IV 용액 백 등을 포함한다. 용기는 유리 또는 플라스틱과 같은 다양한 재료로 형성될 수 있다. 일부 구현예에서 용기는 단독의 조성물 또는 병태의 치료, 예방 및/또는 진단에 효과적인 또 다른 조성물과 병용하는 조성물을 보유한다. 일부 구현예에서, 용기는 멸균 접근 포트를 갖는다. 예시적인 용기는 정맥 내 용액 백, 주사용 바늘로 관통 가능한 마개를 갖는 것을 포함하는 바이알, 또는 경구 투여용 작용제를 위한 병 또는 바이알을 포함한다. 표지 또는 포장 삽입물은 조성물이 질환 또는 병태를 치료하기 위해 사용됨을 나타낼 수 있다. 제조 물품은 조성물이 특정 병태를 치료하기 위해 사용될 수 있음을 나타내는 포장 삽입물을 더 포함할 수 있다. 또한, 또는 추가적으로, 제조 물품은 약학적으로 허용되는 완충제를 포함하는 또 다른 또는 동일한 용기를 더 포함할 수 있다. 이는 기타 재료, 예컨대 기타 완충제, 희석제, 필터, 바늘 및/또는 주사기를 더 포함할 수 있다.

[0540] **VI. 정의**

[0541] [0523] 달리 정의되지 않는 한, 본원에 사용되는 모든 기술 용어, 표기법 및 다른 기술적 및 과학적 용어 또는 전문 용어는 청구된 주제가 속하는 당업계의 숙련자에 의해 통상적으로 이해되는 것과 동일한 의미를 갖는 것으로 의도된다. 일부 경우에, 통상적으로 이해되는 의미를 갖는 용어는 명료성 및/또는 용이한 참조를 위해 본원에 정의되고, 본원에서 이러한 정의를 포함하는 것은 당업계에서 일반적으로 이해되는 것 이상의 실질적인 차이를 나타내는 것으로 반드시 해석되지 않아야 한다.

[0542] [0524] 본원에서 사용되는 단수 형태는 달리 명확하게 표시되지 않는 한, 복수의 대상을 포함한다. 예를 들어, 단수형은 '적어도 하나' 또는 '하나 이상'을 의미한다. 본원에 기재된 측면 및 변형은 이러한 측면 및 변형으로 '이루어진' 및/또는 '본질적으로 이루어진' 구현예를 포함하는 것으로 이해된다.

[0543] [0525] 본 개시 전반에 걸쳐, 청구된 주제의 다양한 측면이 범위 형식으로 제시된다. 범위 형식의 설명은 단지 편의성 및 간결성을 위한 것이고 청구된 주제의 범위에 대한 융통성 없는 제한으로 해석되어서는 안 된다는 것이 이해되어야 한다. 따라서, 범위의 설명은 모든 가능한 하위 범위뿐만 아니라 해당 범위 내의 개별 수치 값을 구체적으로 개시하는 것으로 간주되어야 한다. 예를 들어, 값의 범위가 제공되는 경우, 해당 범위의 상한과 하한 사이에 있는 각각의 중간 값 및 명시된 범위 내에 있는 임의의 다른 명시된 값 또는 중간 값은 청구된 주제에 포괄되는 것으로 이해된다. 이들 더 작은 범위의 상한과 하한은 독립적으로 더 작은 범위 내에 포함될 수 있고, 또한 명시된 범위에서 임의의 구체적으로 배제된 한계를 조건으로 청구된 주제의 범위 내에 포괄될 수 있다. 명시된 값이 제한치 중 하나 또는 둘 다를 포함하는 경우, 이들 포함된 한계의 중 어느 하나 또는 둘 다를 배제하는 범위가 또한 본 발명에 포함된다. 이것은 범위의 폭에 관계없이 적용한다.

[0544] [0526] 본원에 사용되는 용어 '약'은 용이하게 알려진 각각의 값에 대한 일반적인 오차 범위를 지칭한다. 본원에서 값 또는 매개변수의 '약'에 대한 참조는 이 값 또는 매개변수 자체에 대한 구현예를 포함(및 설명)한다. 예를 들어, '약 X'를 지칭하는 설명에는 'X'에 대한 설명이 포함된다. 일부 구현예에서, '약'은 ± 25%, ± 20%, ± 15%, ± 10%, ± 5%, 또는 ± 1%를 지칭할 수 있다.

[0545] [0527] 본 명세서에서 사용되는 용어 "대응하는"은 뉴클레오타이드 서열(또는 염기쌍) 또는 단백질 서열(아미노산)의 위치를 참조하는 것으로, 예를 들어, 서열 목록에 명시된 바와 같이 공개된 서열에서 염기쌍 또는 아미노산 위치에 "대응하는" 뉴클레오타이드 또는 아미노산 위치를 언급하는 것은 GAP 알고리즘과 같은 표준 정렬 알고리즘을 사용하여 동일성을 최대화하기 위해 공개된 서열과 정렬할 때 식별된 뉴클레오타이드 또는 아미노산 위치를 말한다. 서열을 정렬함으로써, 예를 들어 보존되고 동일한 아미노산 잔기를 안내로서 사용하여 상응하는 잔기를 식별할 수 있다. 일반적으로, 상응하는 위치를 확인하기 위해, 가장 높은 순서 일치를 얻도록 아미노산의 서열이 정렬된다(예를 들어, 문헌[Computational Molecular Biology, Lesk, A.M., ed., Oxford University Press, New York, 1988; Biocomputing: Informatics and Genome Projects, Smith, D.W., ed., Academic Press, New York, 1993; Computer Analysis of Sequence Data, Part I, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, 1994; Sequence Analysis in Molecular Biology, von Heinje, G., Academic Press, 1987; 및 Sequence Analysis Primer, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991; Carrillo et al. (1988) SIAM J Applied Math 48: 1073] 참조).

[0546] [0528] "유전자"에는 유전자 산물을 인코딩하는 DNA 영역과 유전자 산물의 생성을 조절하는 모든 DNA 영역이 포함되며, 이러한 조절 서열이 코딩 및/또는 전사된 서열에 인접해 있든 아니든 상관없다. 따라서 유전자에는 프로모터 서열, 종결자, 리보솜 결합 부위 및 내부 리보솜 진입 부위와 같은 번역 조절 서열, 인핸서, 사일런서, 인설레이터, 경계 요소, 복제 원점, 매트릭스 부착 부위 및 유전자좌 제어 영역이 포함되지만 이에 국한되지 않는다. 유전자의 서열은 일반적으로 세포의 염색체에서 고정된 염색체 위치 또는 유전자좌에 존재한다.

[0547] [0529] 유전자와 관련하여 본원에 상호교환적으로 사용되는 '조절 요소' 또는 'DNA 조절 요소'는 이러한 조절 서열이 코딩하고/하거나 전사된 서열에 인접하는지 여부에 관계없이 유전자 생성물의 생성을 조절하는 DNA 영역을 지칭한다. 이에 따라, 조절 요소는 프로모터 서열, 터미네이터, 번역 조절 서열, 예컨대, 리보솜 결합 부위 및 내부 리보솜 진입 부위, 인핸서, 사일런서, 인슐레이터, 경계 요소, 복제 개시점, 기질 부착 부위 및 유전자좌 조절 영역을 포함하지만 이로 반드시 제한되지는 않는다.

[0548] [0530] 본원에 사용되는 '표적 부위' 또는 '표적 핵산 서열'은 결합을 위한 충분한 조건이 존재한다면, 결합 분자(예를 들어, 본원에 개시된 DNA-결합 도메인)가 결합하는 핵산의 일부를 규정하는 핵산 서열이다.

[0549] [0531] 용어 유전자에 대한 '발현' 또는 '유전자 발현'은 유전자 산물 내로, 유전자에 함유된 정보의 전환을 지칭한다. 유전자 생성물은 유전자의 직접 전사 생성물(예를 들어, mRNA, tRNA, rRNA, 안티센스 RNA, 리보자임, 구조 RNA 또는 임의의 다른 유형의 RNA) 또는 mRNA의 번역에 의해 생성된 단백질일 수 있다. 유전자 생성물은 또한, 캡핑, 폴리아데닐화, 메틸화, 및 편집과 같은 과정에 의해 변형된 RNA, 및 예를 들어, 메틸화, 아세틸화, 인산화, 유비퀴틴화, ADP-라이보실화, 미리스토일화, 및 글리코실화에 의해 변형된 단백질을 포함한다. 따라서, 발현 또는 유전자 발현에 대한 언급은 단백질(또는 폴리펩타이드) 발현 또는 mRNA와 같은 유전자의 전사가능한 생성물의 발현을 포함한다. 단백질의 발현은 단백질의 세포 내 발현 또는 표면 발현을 포함할 수 있다. 일반적으로, 유전자 생성물, 예컨대 mRNA 또는 단백질의 발현은 세포에서 검출 가능한 수준이다.

[0550] [0532] 본원에서 사용되는 '검출 가능한' 발현 수준은 당업자에게 알려진 표준 기술에 의해 검출 가능한 수준을 의미하고, 예를 들어, 차등 디스플레이, 역전사효소[reverse transcriptase, RT]결합 폴리머라제 사슬 반응[RT-coupled polymerase chain reaction, PCR], 노던 블롯, 및/또는 RNA 분해효소 보호 분석뿐만 아니라 단백질 검출을 위한 면역친화도 기반 방법, 예컨대 유세포 분석, ELISA, 또는 웨스턴 블롯을 포함한다. 발현 수준의 정도는 표준 특성화 기술을 통해 시각화되거나 측정되기에 충분히 클 필요가 있다.

[0551] [0533] 본원에 사용되는 용어 'DNA-표적화 시스템' 또는 '후성적 변형 DNA-표적화 시스템'은 표적 유전자 및/또는 이의 조절 요소의 표적 부위를 표적화하는 DNA-결합 도메인(예컨대, 본원에 기재된 임의의 Cas, ZFN, 또는 TALE-기반 DNA-결합 도메인)을 포함하는 조성물을 지칭한다. 일부 구현예에서, DNA-표적화 시스템은 표적 부위를 표적화하도록 조작된다. 일부 구현예에서, DNA-표적화 시스템은 DNA-표적화 시스템에 의해 표적 부위로 동원되는 경우에 표적 유전자 및/또는 이의 조절 요소의 전사를 변형시키는 하나 이상의 효과기 도메인을 포함한다. 일부 구현예에서, DNA-표적화 시스템은 하나보다 많은 표적 부위(즉, 복수의 표적 부위), 예컨대 2, 3, 4, 5, 6개 이상의 표적 부위를 표적화할 수 있다.

[0552] [0534] 본원에 사용되는 용어 'DNA-표적화 모듈'은 하나의 표적 부위를 표적화하는 본원에 기재된 DNA-표적화 시스템의 임의의 조성물 또는 부분을 지칭한다. 예를 들어, 개별 DNA-표적화 모듈은 표적 유전자 및/또는 이의 조절 요소에 대한 표적 부위를 표적화하는 DNA-결합 도메인(예를 들어, ZFN 또는 TALE DNA-결합 도메인), 및 전사 억제자 도메인(예를 들어, KRAB)을 포함하는 융합 단백질을 포함할 수 있다. 다른 구현예에서, DNA-표적화 모듈은 (a) Cas 단백질 및 전사 억제자 도메인을 포함하는 융합 단백질, 및 (b) 표적 유전자 및/또는 이의 조절 요소의 표적 부위를 표적화하는 gRNA를 포함한다. 본원에 제공된 DNA-표적화 시스템은 하나 이상의(예컨대 두 개의) DNA-표적화 모듈을 포함할 수 있다. 본원에 제공된 DNA-표적화 시스템은 2 내지 10개의 DNA-표적화 모듈을 포함할 수 있다.

[0553] [0535] 일부 구현예에서, DNA-표적화 시스템의 두 개의 DNA-표적화 모듈은 개별(즉, 비중첩) 성분을 포함할 수 있다. 예를 들어, DNA-표적화 시스템은 두 개의 상이한 융합 단백질을 포함할 수 있고, 각각의 융합 단백질은 상이한 유전자 및/또는 이의 조절 요소를 표적화하고 억제한다. 예를 들어, DNA-표적화 시스템은 제1 유전자에 대한 표적 부위를 표적화하는 DNA-결합 도메인(예컨대, ZFN 또는 TALE DNA-결합 도메인) 및 전사 억제자 도메인을 포함하는 제1 융합 단백질을 포함하는 제1 DNA-표적화 모듈, 및 제2 유전자에 대한 표적 부위를 표적화하는 DNA-결합 도메인 및 전사 억제자 도메인을 포함하는 제2 융합 단백질을 포함하는 제2 DNA-표적화 모듈을 포함할 수 있다. 또 따른 예에서, DNA-표적화 시스템은 제1 유전자에 대한 표적 부위를 표적화하는 DNA-결합 도메인(예컨대, ZFN 또는 TALE DNA-결합 도메인) 및 전사 억제자 도메인을 포함하는 제1 융합 단백질을 포함하는 제1

DNA-표적화 모듈, 및 (a) Cas 단백질 및 전사 억제자 도메인을 포함하는 융합 단백질 및 (b) 제2 유전자에 대한 표적 부위를 표적화하는 gRNA를 포함하는 제2 DNA-표적화 모듈을 포함할 수 있다. 또 다른 예에서, DNA-표적화 시스템은 제1 Cas 단백질 및 전사 억제자 도메인을 포함하는 제1 융합 단백질 및 (b) 제1 Cas 단백질과 복합체를 형성하고 제1 표적 유전자에 대한 표적 부위를 표적화하는 제1 gRNA를 포함하는 제1 DNA-표적화 모듈, 및 제1 Cas 단백질과 상이한 제2 Cas 단백질 및 전사 억제자 도메인을 포함하는 제2 융합 단백질, 및 (b) 제2 Cas 단백질과 복합체를 형성하고 제2 Cas 단백질을 제2 표적 유전자에 대한 표적 부위로 표적화하는 제2 gRNA를 포함하는 제2 DNA-표적화 모듈을 포함할 수 있다. 상이한 Cas 단백질 변이체(예를 들어, SpCas9 및 SaCas9)는 본원에 기재된 바와 같이 상이한 gRNA 스캐폴드 서열 및 PAM과 양립가능하다는 것이 이해될 것이다. 따라서, 각각 상이한 표적 부위를 표적화하는, 다수의 비중첩 CRISPR/Cas-기반 DNA-표적화 모듈을 포함하는 단일 DNA-표적화 시스템을 조작하는 것이 가능하다.

[0554] [0536] 일부 구현예에서, DNA-표적화 시스템의 두 개의 DNA-표적화 모듈은 공유(즉, 중첩) 성분을 포함할 수 있다. 예를 들어, DNA-표적화 시스템은 (a) Cas 단백질 및 전사 억제자 도메인을 포함하는 융합 단백질 및 (b) 제1 유전자에 대한 표적 부위를 표적화하는 gRNA를 포함하는 제1 DNA-표적화 모듈, 및 (a) 제1 DNA-표적화 모듈의 융합 단백질, 및 (b) 제2 유전자에 대한 표적 부위를 표적화하는 gRNA를 포함하는 제2 DNA-표적화 모듈을 포함할 수 있다. 주어진 Cas 단백질에 대해 두 개 이상의 상이한 gRNA를 제공하는 것은 동일한 Cas 단백질의 상이한 분자가 두 개 이상의 gRNA의 표적 부위로 표적화될 수 있게 한다는 것이 이해될 것이다.

[0555] [0537] 본원에 사용되는 용어 '감소된(reduced) 발현' 또는 '감소된(decreased) 발현'은 DNA-표적화 시스템에 의해 특정 유전자 발현을 조절하기 위한 변형을 함유하지 않는 원래 또는 공급원 세포에서의 발현보다 낮은 임의의 형태의 발현, 예를 들어 (발현의 부재 또는 측정불가능한 발현일 수도 있는) 야생형 발현 수준을 의미한다. 본원의 '감소된[reduced] 발현' 또는 '감소된[decreased] 발현'에 대한 언급은 변형되지 않은 세포 또는 야생형 T 세포와 같은 T 세포 내로 DNA-결합 시스템을 도입하기 위해 접촉되거나 조작되기 전의 원래의 공급원 세포와 같은 변형을 함유하지 않은 세포에서의 수준에 대한 유전자 발현의 감소를 의미하는 것으로 간주된다. 발현 감소는 적어도 5%, 10%, 20%, 30%, 40% 또는 50%, 60%, 70%, 80%, 85%, 90%, 또는 100% 또는 그 이상일 수 있다. 일부 경우에, 발현 감소는 적어도 2배, 5배, 10배, 20배, 30배, 40배, 50배, 60배, 70배, 80배, 90배, 100배, 200배 이상일 수 있다.

[0556] [0538] 본원에 사용되는 용어 '감소된[reduced] 전사' 또는 '감소된[decreased] 전사'는 DNA-표적화 시스템에 의한 전사를 조절하기 위한 변형을 함유하지 않는 원래 또는 공급원 세포에서의 유전자의 전사보다 낮은 유전자의 전사 수준, 예를 들어 유전자의 야생형 전사 수준을 지칭한다. 감소된[reduced] 전사 또는 감소된[decreased] 전사에 대한 언급은 유전자의 전사가능한 생성물, 예컨대 mRNA의 수치의 감소를 지칭할 수 있다. 실시간 정량적 RT(역전사효소)-폴리머라제 연쇄 반응[quantitative RT-polymerase chain reaction, qRT-PCR], 노던 블롯, 마이크로어레이 분석 또는 RNA 서열분석[RNA sequencing, RNA-Seq]을 포함하지만 이에 제한되지 않는 임의의 다양한 방법이 전사가능한 생성물, 예컨대 mRNA의 수치를 모니터링 또는 정량화하는 데 사용될 수 있다. 전사 감소는 적어도 5%, 10%, 20%, 30%, 40% 또는 50%, 60%, 70%, 80%, 85%, 90%, 또는 100% 또는 그 이상일 수 있다. 일부 경우에, 전사 감소는 적어도 1배, 2배, 5배, 10배, 20배, 30배, 40배, 50배, 60배, 70배, 80배, 90배, 100배, 200배 이상일 수 있다.

[0557] [0539] 본 명세서에서 사용되는 "후성적 변형"은 DNA 서열의 변화가 아닌 DNA 메틸화, 히스톤 변형, miRNA 발현 조절과 같은 사건으로 인해 발생하는 유전자 발현의 변화를 의미한다.

[0558] [0540] 본원에 사용되는, 세포에 대한 용어 '변형' 또는 '변형된'은 세포 내의 유전자 발현에 영향을 미치는 세포 내의 임의의 변화 또는 변경을 지칭한다. 일부 구현예에서, 변형은 유전자 생성물의 발현을 변경(예를 들어, 감소)시키기 위해 유전자 또는 이의 조절 요소의 후성적 상태를 직접 변화시키는 후성적 변형이다. 일부 구현예에서, 본원에 기재된 변형은 표적 유전자 또는 선택된 폴리뉴클레오타이드 서열의 발현 감소를 일으킨다.

[0559] [0541] 본원에 사용되는 '융합' 분자는 둘 이상의 소단위 분자들이, 예컨대 공유적으로 연결된 분자이다. 융합 분자의 예는 융합 단백질(예를 들어, DNA-결합 도메인, 예컨대 ZFP, TALE DNA-결합 도메인 또는 CRISPR-Cas 단백질과 하나 이상의 효과기 도메인, 예컨대 전사활성화 도메인 사이의 융합)을 포함하지만 이에 제한되지는 않는다. 융합 분자는 또한 폴리뉴클레오타이드 성분이 폴리펩타이드 성분과 결합되어 기능성 분자를 형성하는 시스템(예를 들어, 단일 가이드 RNA가 기능성 도메인과 결합되어 유전자 발현을 조절하는 CRISPR/Cas 시스템)의 일부일 수 있다. 융합 분자는 또한 융합 핵산, 예를 들어 융합 단백질을 인코딩하는 핵산을 포함한다. 세포에서 융합 단백질의 발현은 융합 단백질의 세포로의 전달 또는 융합 단백질을 인코딩하는 폴리뉴클레오타이드의 세포

로의 전달에 의해 발생할 수 있고, 여기서 폴리뉴클레오타이드가 전사되고, 전사체가 번역되어 융합 단백질이 생성된다.

[0560] [0542] 본 명세서에서 사용되는 용어 "메틸화"는 일반적으로 정상 세포에는 존재하지 않는 부위에서 DNA의 시토신 잔기에 대한 후성적 메틸화가 존재하는 것을 의미한다.

[0561] [0543] 본원에 사용되는 용어 '벡터'는 이것이 연결된 또 다른 핵산을 증식시킬 수 있는 핵산 분자를 지칭한다. 용어는 자가 복제 핵산 구조로서의 벡터뿐만 아니라 이것이 도입된 숙주 세포의 게놈에 포함되는 벡터를 포함한다. 특정 벡터는 이들이 작동되게 연결되는 핵산의 발현을 지시할 수 있다. 이러한 벡터는 본원에 '발현 벡터'로서 지칭된다. 벡터 중에는 바이러스 벡터, 예컨대 아데노바이러스 벡터 또는 렌티바이러스 벡터가 있다.

[0562] [0544] 용어 '발현 벡터'는 발현하고자 하는 뉴클레오타이드 서열에 작동되게 연결되는 발현 조절 서열을 포함하는 재조합 폴리뉴클레오타이드를 포함하는 벡터를 지칭한다. 발현 벡터는 발현을 위한 충분한 시스 작용 요소를 포함하고, 발현을 위한 다른 요소는 숙주 세포에 의해 또는 시험관 내 발현 시스템에서 공급될 수 있다. 발현 벡터는 재조합 폴리뉴클레오타이드를 포함하는 코스미드, 플라스미드(예를 들어, 네이키드 또는 리포솜에 함유됨) 및 바이러스(예를 들어, 렌티바이러스, 레트로바이러스, 아데노바이러스, 및 아데노 연관 바이러스)를 포함하지만 이에 제한되지 않는다.

[0563] [0545] 용어 '단리된'은 자연 상태에서 변경되거나 제거되는 것을 의미한다. 예를 들어, 살아있는 동물에 자연적으로 존재하는 핵산 또는 펩타이드는 '단리된' 것이 아니지만 자연 상태의 공존하는 물질에서 부분적으로 또는 완전히 분리된 동일한 핵산 또는 펩타이드는 '단리된' 것이다. 단리된 핵산 또는 단백질은 실질적으로 정제된 형태로 존재할 수 있거나, 예를 들어 숙주 세포와 같은 비천연 환경에 존재할 수 있다.

[0564] [0546] 용어 '폴리뉴클레오타이드'는 뉴클레오타이드의 사슬을 의미한다. 또한, 핵산은 뉴클레오타이드의 중합체이다. 따라서, 본원에 사용되는 핵산 및 폴리뉴클레오타이드는 상호교환가능하다. 당업자는 핵산이 단일 '뉴클레오타이드'로 가수분해될 수 있는 폴리뉴클레오타이드라는 일반적인 지식을 갖는다. 단일 뉴클레오타이드는 뉴클레오사이드로 가수분해될 수 있다. 본원에 사용되는 폴리뉴클레오타이드는 제한없이 재조합 수단, 즉, 통상적인 클로닝 기술 및 PCR™ 등을 사용하여 재조합 라이브러리 또는 세포 게놈에서 핵산 서열을 클로닝하는 것을 포함하는, 당업계에서 이용 가능한 임의의 수단에 의해, 그리고 합성 수단에 의해 얻어지는 모든 핵산 서열을 포함하지만 이에 제한되지 않는다.

[0565] [0547] 본원에 사용되는 용어 '펩타이드', '폴리펩타이드', 및 '단백질'은 상호교환적으로 사용되고 펩타이드 결합에 의해 공유 연결된 아미노산 잔기로 구성된 화합물을 지칭한다. 단백질 또는 펩타이드는 적어도 두 개의 아미노산을 함유해야 하고, 단백질 또는 펩타이드 서열을 포함할 수 있는 최대 개수의 아미노산에는 제한을 두지 않는다. 폴리펩타이드는 펩타이드 결합에 의해 서로 연결된 두 개 이상의 아미노산을 포함하는 임의의 펩타이드 또는 단백질을 포함한다. 본원에 사용되는 용어는, 예를 들어 펩타이드, 올리고펩타이드 및 올리고머로서 당업계에서 또한 통상적으로 지칭되는 단쇄, 및 일반적으로 단백질로서 당업계에서 지칭되는, 많은 유형이 있는 보다 긴 쇄 둘 다를 지칭한다. '폴리펩타이드'는, 예를 들어, 특히, 생물학적으로 활성인 단편, 실질적으로 상동성 폴리펩타이드, 올리고펩타이드, 폴리펩타이드의 동종이량체, 이종이량체, 변이체, 변형된 폴리펩타이드, 유도체, 유사체, 융합 단백질을 포함한다. 폴리펩타이드는 천연 펩타이드, 재조합 펩타이드, 합성 펩타이드 또는 이의 조합물을 포함한다.

[0566] [0548] 본원에 사용되는 '아미노산 서열 동일성 백분율(%)' 및 '동일성 백분율'은, 아미노산 서열(참조 폴리펩타이드 서열)과 관련하여 사용되는 경우, 서열을 정렬하고 필요한 경우 최대 백분율 서열 동일성에 달하기 위해 갭을 도입한 후 표준 폴리펩타이드 서열 내 아미노산 잔기와 동일한 후보 서열(예를 들어, 대상 항체 또는 단편)에서의 아미노산 잔기의 백분율로서 정의되고 서열 동일성의 일부로서 임의의 보존적 치환을 고려하지 않는다. 아미노산 서열 동일성 백분율을 결정하기 위한 목적을 위한 정렬은 공개적으로 이용 가능한 컴퓨터 소프트웨어, 예컨대 BLAST, BLAST-2, ALIGN 또는 Megalign(DNASTAR) 소프트웨어를 사용하여, 일부 구현예에서, 다양한 알려진 방식으로 달성될 수 있다. 비교되는 서열의 전장에 걸쳐 최대 정렬을 달성하는 데 필요한 임의의 알고리즘을 포함하여, 서열을 정렬하기 위한 적절한 매개변수는 결정될 수 있다.

[0567] [0549] 일부 구현예에서, '작동가능한 상태로 연결된'은 적절한 분자(예를 들어, 전사 억제자 단백질)가 조절 서열에 결합되는 경우, 유전자 발현을 허용하는 방식으로 DNA 서열(예를 들어, 이종 핵산) 및 조절 서열(들)과 같은 성분의 회합을 포함할 수 있다. 따라서, 이는 기재된 성분은 이의 의도된 방식으로 작용하도록 하는 관계에 있다는 것을 의미한다.

[0568] [0550] 아미노산 치환은 폴리펩타이드에서 하나의 아미노산을 또 다른 아미노산으로 대체하는 것을 포함할 수 있다. 치환은 보존적 아미노산 치환 또는 비보존적 아미노산 치환일 수 있다. 아미노산 치환은 원하는 활성, 예를 들어, 보유된/개선된 항원 결합, 감소된 면역원성, 또는 개선된 ADCC 또는 CDC를 위해 스크리닝된 제품 및 관심 결합 분자, 예를 들어 항체에 도입될 수 있다.

[0569] [0551] 아미노산은 일반적으로 다음의 공통의 측쇄 특성에 따라 분류될 수 있다.

[0570] [0552] (1) 소수성: 노르류신, Met, Ala, Val, Leu, Ile;

[0571] [0553] (2) 중성 친수성: Cys, Ser, Thr, Asn, Gln;

[0572] [0554] (3) 산성: Asp, Glu;

[0573] [0555] (4) 염기성: His, Lys, Arg;

[0574] [0556] (5) 사슬 배향에 영향을 미치는 잔기: Gly, Pro;

[0575] [0557] (6) 방향족: Trp, Tyr, Phe.

[0576] [0558] 일부 구현예에서, 보존적 치환은 이들 클래스 중 하나의 구성원의 동일한 클래스의 다른 구성원으로의 교환을 수반할 수 있다. 일부 구현예에서, 비보존적 치환은 이들 클래스 중 하나의 구성원을 또 다른 클래스로 교환하는 것을 수반할 수 있다.

[0577] [0559] 본원에 사용되는 조성물은 세포를 포함하는 2종 이상의 생성물, 물질 또는 화합물의 임의의 혼합물을 지칭한다. 이는 용액, 현탁액, 액체, 분말, 페이스트, 수성, 비수성, 또는 이의 임의의 조합일 수 있다.

[0578] [0560] 본원에 사용되는 바와 같이, 상호교환적으로 사용되는 용어인 '대상' 또는 '개체'는 포유동물이다. 일부 구현예에서, '포유동물'은 인간, 비인간 영장류, 가축 및 농장 동물, 및 동물원, 스포츠, 또는 반려 동물, 예컨대 개, 말, 토끼, 소, 돼지, 햄스터, 게르빌루스쥐, 마우스, 흰족제비, 래트, 고양이, 원숭이 등을 포함한다. 일부 구현예에서, 대상 또는 개체는 인간이다. 일부 구현예에서, 대상은 질환, 장애 또는 병태를 갖는 것으로 알려지거나 의심되는 환자이다.

[0579] [0561] 본 명세서에서 사용되는 용어 "치료" 및 "치료"는 대상에게 효과적인 양의 세포(예컨대 T 세포), 예를 들어 DNA-표적화 시스템 또는 본 명세서에 기술된 DNA-표적화 시스템을 인코딩하는 폴리뉴클레오타이드에 의해 변형된 세포를 대상에게 투여하여, 대상에서 질병의 적어도 하나의 증상이 감소하거나 질병이 개선되는 것, 예를 들어 유익하거나 원하는 임상적 결과가 나타나도록 하는 것을 포함한다. 이 기술의 목적을 위해, 유익하거나 바람직한 임상 결과는 검출 가능하든 검출 가능하지 않든 하나 이상의 증상 완화, 질환 정도 감소, 안정화된(즉, 악화되지 않은) 질환 상태, 질환 진행 지연 또는 둔화, 질환 상태 개선 또는 완화, 및 관해(부분적이든 전체적이든)를 포함하나 이에 제한되지 않는다. 치료하는 것은 치료를 받지 않은 경우 예상되는 생존과 비교하여 생존을 연장하는 것을 지칭할 수 있다. 따라서, 치료는 질환 상태를 개선시키는 것일 수 있지만, 질환의 완전한 치료가 아닐 수 있다는 것을 당업자는 인지할 것이다. 일부 구현예에서, 질환 또는 장애의 하나 이상의 증상은 질환 치료 시 적어도 5%, 적어도 10%, 적어도 20%, 적어도 30%, 적어도 40%, 또는 적어도 50%로 완화된다.

[0580] [0562] 용어 '치료 유효량'은 연구원, 수의사, 의사 또는 기타 임상의에 의해 추구되는 조직, 시스템 또는 대상의 생물학적 또는 의학적 반응을 유발할 대상 화합물의 양을 지칭한다. 용어 '치료 유효량'은 투여되는 경우 치료되는 장애 또는 질환의 하나 이상의 징후 또는 증상을 발달을 방지하거나 어느 정도 완화시키기에 충분한 화합물 또는 세포와 같은 생물학적 분자의 양을 포함한다. 치료 유효량은 생물학적 분자, 질환 및 이의 심각도, 및 치료되는 대상의 연령, 체중 등에 따라 다를 것이다.

[0581] **VII. 예시적 구현예**

[0582] [0563] 제공된 구현예는 다음과 같다.

[0583] 1. 하나 이상의 B형 간염 바이러스(HBV) 유전자의 전사를 억제하기 위한 적어도 하나의 DNA-표적화 모듈을 포함하는 후성적 변형 DNA-표적화 시스템으로서, 여기서 적어도 하나의 DNA-표적화 모듈 각각은:

[0584] (a) B형 간염 바이러스 DNA 서열의 표적 부위를 표적화하기 위한 DNA-결합 도메인; 및

[0585] (b) 적어도 하나의 전사 억제자 효과기 도메인

[0586] 을 포함하는 융합 단백질을 포함하는 것인, 후성적 변형 DNA-표적화 시스템.

[0587] 2. 구현예 1에 있어서, 적어도 하나의 DNA-결합 도메인은 (a) Cas 단백질 또는 이의 변이체 및 (b) Cas 단백질과 결합하여 Cas 단백질을 표적 부위로 표적화할 수 있는 적어도 하나의 gRNA를 포함하는, 규칙적인 간격을 갖고 나타나는 짧은 회문 구조의 반복 서열(CRIPRE:Clustered Regularly Interspaced Short Palindromic Repeats) 관련 (Cas)-가이드(gRNA) 조합; 아연 집게 단백질(ZFP); 전사 활성화제 유사 효과기(TALE); 메가뉴클레아제; 호밍 엔도뉴클레아제; 또는 I-SceI 효소 또는 이의 변이체를 포함하며, 선택적으로 여기서 DNA-결합 도메인은 상기의 임의의 촉매적으로 비활성인 변이체를 포함하는 것인, 후성적 변형 DNA-표적화 시스템.

[0588] 3. 구현예 1 또는 2에 있어서, B형 간염 바이러스 DNA 서열은 HBV 유전자 또는 이의 조절 요소인, 후성적 변형 DNA-표적화 시스템.

[0589] 4. 구현예 1-3 중 어느 한 구현예에 있어서, 적어도 하나의 DNA-표적화 모듈은 하나 이상의 HBV 유전자 또는 이의 조절 요소의 복수의 표적 부위를 표적화하기 위한 복수의 DNA-표적화 모듈을 포함하는 것인, 후성적 변형 DNA-표적화 시스템.

[0590] 5. 구현예 4에 있어서, 복수의 DNA-표적화 모듈은 적어도 제1 DNA-표적화 모듈과 제2 DNA-표적화 모듈을 포함하고, 여기서:

[0591] (1) 제1 DNA-표적화 모듈은 제1 HBV 유전자의 전사를 억제하고, 여기서 제1 DNA-표적화 모듈은 (a) 제1 유전자 또는 이의 조절 DNA 요소의 표적 부위를 표적화하기 위한 DNA-결합 도메인; 및 (b) 적어도 하나의 전사 억제 도메인을 포함하는 제1 융합 단백질을 포함하고;

[0592] (2) 제2 DNA-표적화 모듈은 제2 HBV 유전자의 전사를 억제하고, 여기서 제2 DNA-표적화 모듈은 (a) 제2 유전자 또는 이의 조절 DNA 요소의 표적 부위를 표적화하기 위한 DNA-결합 도메인; 및 (b) 적어도 하나의 전사 억제 도메인을 포함하는 제2 융합 단백질을 포함하며, 선택적으로 여기서:

[0593] 제1 DNA-표적화 모듈과 제2 DNA-표적화 모듈은 동일한 융합 단백질을 공유하여 제1 및 제2 융합 단백질은 동일하고, 여기서 융합 단백질의 DNA-결합 도메인은 제1 가이드 RNA(gRNA) 및 제2 gRNA와 연관될 수 있는 CRISPER 관련(Cas) 단백질 또는 그 변이체이며, 여기서

[0594] 제1 DNA-표적화 모듈은 제1 HBV 유전자 또는 이의 조절 요소의 표적 부위를 표적으로 하는 제1 gRNA를 포함하고, 제2 DNA-표적화 모듈은 제2 HBV 유전자 또는 이의 조절 요소의 표적 부위를 표적으로 하는 제2 gRNA를 포함하는 것인, 후성적 변형 DNA-표적화 시스템.

[0595] 6. 하나 이상의 B형 간염 바이러스(HBV) 유전자의 전사를 억제하기 위한 후성적 변형 DNA-표적화 시스템으로서, 여기서 DNA-표적화 시스템은:

[0596] (a) CRISPR 관련 (Cas) 단백질 또는 이의 변이체 및 적어도 하나의 전사 억제자 효과기 도메인을 포함하는 융합 단백질; 및

[0597] (b) 적어도 제1 gRNA와 제2 gRNA를 포함하는 복수의 가이드 RNA(gRNA)로서, 여기서 제1 gRNA는 제1 HBV 유전자 또는 이의 조절 요소의 표적 부위를 표적으로 하고 제2 gRNA는 제2 HBV 유전자 또는 이의 조절 요소의 표적 부위를 표적으로 하는 것인 복수의 가이드 RNA(gRNA)

[0598] 를 포함하고, 여기서 제1 및 제2 유전자 또는 이의 조절 요소는 B형 간염 바이러스 복제 및/또는 HBV 전사를 조절하는 것인, 후성적 변형 DNA-표적화 시스템.

[0599] 7. 구현예 6에 있어서, DNA-표적화 시스템은 B형 간염 바이러스 복제 및/또는 HBV 전사를 조절하는 제3 유전자 또는 이의 조절 요소의 표적 부위를 표적화하는 제3 gRNA를 더 포함하고,

[0600] 선택적으로, 상기 시스템은 추가로 제4 유전자 또는 이의 조절 요소의 표적 부위를 표적으로 하는 제4 gRNA, 선택적으로 제5 유전자 또는 이의 조절 요소를 표적으로 하는 제5 gRNA, 및/또는 선택적으로 제6 유전자 또는 이의 조절 요소의 표적 부위를 표적으로 하는 제6 gRNA를 더 포함하고,

[0601] 여기서 유전자 또는 이의 조절 요소는 B형 간염 바이러스 복제 및/또는 HBV 전사를 조절하는 것인, 후성적 변형 DNA-표적화 시스템.

[0602] 8. 구현예 7에 있어서, 제1, 제2, 제3, 제4, 제5 및/또는 제6 유전자 또는 이의 조절 요소는 상이한 것인, 후성적 변형 DNA-표적화 시스템.

[0603] 9. 하나 이상의 B형 간염 바이러스(HBV) 유전자의 전사를 억제하기 위한 후성적 변형 DNA-표적화 시스템으로서, 상기 DNA-표적화 시스템은:

[0604] (a) CRSISPR(Clustered Regularly Interspaced Short Palindromic Repeats) 관련(CAS) 단백질 또는 이의 변이체와 적어도 하나의 전사 억제자 효과기 도메인을 포함하는 융합 단백질; 및

[0605] (b) 복수의 유전자 또는 이의 조절 요소의 복수의 표적 부위를 표적으로 하는 복수의 가이드 RNA(gRNA)

[0606] 를 포함하고, 여기서 상기 복수의 유전자 또는 이의 조절 요소는 B형 간염 바이러스의 복제 및/또는 HBV 전사를 조절하는 것인, 후성적 변형 DNA-표적화 시스템.

[0607] 10. 하나 초과의 B형 간염 바이러스(HBV) 유전자의 전사를 억제하기 위한 단일 DNA-표적화 모듈을 포함하는 후성적 변형 DNA-표적화 시스템으로서, 여기서 DNA-표적화 모듈은:

[0608] (a) CRSISPR(Clustered Regularly Interspaced Short Palindromic Repeats) 관련(CAS) 단백질 또는 이의 변이체와 적어도 하나의 전사 억제자 효과기 도메인을 포함하는 융합 단백질; 및

[0609] (b) 복수의 유전자 또는 이의 조절 요소의 복수의 표적 부위를 표적화하는 가이드 RNA(gRNA)

[0610] 를 포함하고, 여기서 상기 복수의 유전자 또는 이의 조절 요소는 B형 간염 바이러스 복제 및/또는 HBV 전사를 조절하는 것인, 후성적 변형 DNA-표적화 시스템.

[0611] 11. 구현예 1-10 중 어느 한 구현예에 있어서, 전사 억제는 HBV 복제 감소 및/또는 HBV 단백질 수치 감소가 초래하는 것인, 후성적 변형 DNA-표적화 시스템.

[0612] 12. 구현예 1-11 중 어느 한 구현예에 있어서, DNA-표적화 시스템은 유전적 파괴 또는 DNA 파손을 도입하지 않는, 후성적 변형 DNA-표적화 시스템.

[0613] 13. 구현예 1-9 및 11-12 중 어느 한 구현예에 있어서, 적어도 하나의 DNA-결합 모듈은 HBV DNA 서열 내 복수의 표적 부위를 함께 표적화하는 복수의 DNA-결합 모듈을 포함하고, 선택적으로 각 DNA-결합 모듈은 HBV DNA 서열에서 다른 표적 부위를 표적하는, 후성적 변형 DNA-표적화 시스템.

[0614] 14. 구현예 4, 5 및 9-13 중 어느 한 구현예에 있어서, 복수의 표적 부위는 2, 3, 4, 5 또는 6개의 다른 표적 부위인, 후성적 변형 DNA-표적화 시스템.

[0615] 15. 구현예 4, 5 및 9-14 중 어느 한 구현예에 있어서, 복수의 표적 부위는 각각 다른 HBV 유전자 또는 이의 조절 요소에 존재하는, 후성적 변형 DNA-표적화 시스템.

[0616] 16. 구현예 1-14 중 어느 한 구현예에 있어서, 각 표적 부위는 동일한 HBV 유전자 또는 이의 조절 요소에 존재하는, 후성적 변형 DNA-표적화 시스템.

[0617] 17. 구현예 1-5 및 11-16 중 어느 한 구현예에 있어서, 시스템은 2 내지 10개의 DNA-표적화 모듈을 포함하는, 후성적 변형 DNA-표적화 시스템.

[0618] 18. 구현예 1-5 및 11-17 중 어느 한 구현예에 있어서, DNA-표적화 모듈 중 두 개 이상은 동일한 융합 단백질을 공유하거나 또는 DNA-표적화 모듈 중 두 개 이상은 서로 다른 융합 단백질을 포함하는, 후성적 변형 DNA-표적화 시스템.

[0619] 19. 구현예 1-5 및 11-18 중 어느 한 구현예에 있어서, 각 DNA-표적화 모듈의 DNA-결합 도메인은 CRISPER 관련(Cas) 단백질 또는 이의 변이체와 적어도 하나의 전사 억제자 효과기 도메인을 포함하는 융합 단백질을 포함하고, 각 DNA-표적화 모듈은 고유한 gRNA를 포함하는, 후성적 변형 DNA-표적화 시스템.

[0620] 20. 구현예 1-19 중 어느 한 구현예에 있어서, 표적 부위 또는 각각의 표적 부위는 공유적으로 폐쇄된 원형 DNA(cccDNA) 형태, 이완된 원형 DNA(rcDNA) 형태 및/또는 인간 게놈 DNA에 통합된 HBV 바이러스 DNA로 존재하는, 후성적 변형 DNA-표적화 시스템.

[0621] 21. 구현예 1-20 중 어느 한 구현예에 있어서, 표적 부위 또는 각각의 표적 부위는 HBV 복제 및/또는 HBV 전사를 제어하는데 관여하는 유전자 또는 이의 조절 요소에 존재하거나 그 근처에 존재하는, 후성적 변형 DNA-표적화 시스템.

[0622] 22. 구현예 21에 있어서, HBV 복제 및/또는 HBV 전사를 제어하는데 관여하는 유전자는 폴리머라제, 외피

단백질, 캡시드 단백질, 전사 인자 또는 전사 전사 활성화제를 인코딩하는, 후성적 변형 DNA-표적화 시스템.

[0623] 23. 구현예 21 또는 구현예 22에 있어서, HBV 복제 및/또는 HBV 전사를 제어하는데 관여하는 유전자는 폴리머라제 유전자, S-패밀리 유전자, X-유전자 또는 코어 패밀리 유전자인, 후성적 변형 DNA-표적화 시스템.

[0624] 24. 구현예 1-23 중 어느 한 구현예에 있어서, 적어도 하나의 표적 부위는 B형 간염 바이러스 단백질 X(HBx)를 인코딩하는 X 유전자의 유전자 또는 이의 조절 요소에 존재하는, 후성적 변형 DNA-표적화 시스템.

[0625] 25. 구현예 1-24 중 어느 한 구현예에 있어서, 표적 부위 또는 각각의 표적 부위는 HBV 복제 및/또는 HBV 전사를 제어하는 HBV 유전자의 조절 요소에 있거나 그 근처에 있는, 후성적 변형 DNA-표적화 시스템.

[0626] 26. 구현예 25에 있어서, 조절 요소는 프로모터 영역인, 후성적 변형 DNA-표적화 시스템.

[0627] 27. 구현예 26에 있어서, 프로모터 영역은 pre-S1 프로모터, pre-S2 프로모터, X 프로모터 또는 기저 코어 프로모터인, 후성적 변형 DNA-표적화 시스템.

[0628] 28. 구현예 25에 있어서, 조절 요소는 인핸서 영역인, 후성적 변형 DNA-표적화 시스템.

[0629] 29. 구현예 28에 있어서, 인핸서 영역은 Enh1 또는 Enh2 인핸서 영역인, 후성적 변형 DNA-표적화 시스템.

[0630] 30. 구현예 25에 있어서, 조절 요소는 전사 처리 제어 영역인, 후성적 변형 DNA-표적화 시스템.

[0631] 31. 구현예 1-24 중 어느 한 구현예에 있어서, 표적 부위 또는 각각의 표적 부위는 HBV 유전자의 코딩 영역에 있는, 후성적 변형 DNA-표적화 시스템.

[0632] 32. 구현예 1-31 중 어느 한 구현예에 있어서, 표적 부위 또는 각각의 표적 부위는 전사 개시 부위로부터 500개 염기쌍(bp) 이내, 1000bp 이내, 1500bp 이내에 위치하는, 후성적 변형 DNA-표적화 시스템.

[0633] 33. 구현예 1-32 중 어느 한 구현예에 있어서, 표적 부위 또는 각각의 표적 부위는 HBV 게놈의 0-3300 염기쌍(bp) 사이의 염기쌍에 위치한 표적 영역 내에 위치하며, 선택적으로 서열번호 650에 제시된 HBV 게놈을 참조로 위치에 상응하는 0-3182 bp 사이에 위치하는, 후성적 변형 DNA-표적화 시스템.

[0634] 34. 구현예 1-33 중 어느 한 구현예에 있어서, 표적 부위 또는 각각의 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 HBV 게놈의 43bp - 490bp, 1033bp - 1749bp, 1800bp - 1950bp 또는 2953bp - 3182bp 사이의 염기쌍에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내에 위치하는, 후성적 변형 DNA-표적화 시스템.

[0635] 35. 구현예 1-33 중 어느 한 구현예에 있어서, 표적 부위 또는 각각의 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 HBV 게놈의 1 bp - 42 bp, 491 bp - 1032 bp, 1750 bp - 1799 bp, 또는 1951 bp - 2952 bp 사이의 염기쌍에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내에 위치하는, 후성적 변형 DNA-표적화 시스템.

[0636] 36. 구현예 1-33 중 어느 한 구현예에 있어서, 표적 부위 또는 각각의 표적 부위는 HBV 게놈의 CpG 섬에 있는, 후성적 변형 DNA-표적화 시스템.

[0637] 37. 구현예 1-33 및 36 중 어느 한 구현예에 있어서, 표적 부위 또는 각각의 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 HBV 게놈의 67bp-392bp, 1033bp-1749bp 또는 2215bp-2490bp 사이의 염기쌍에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내에 위치하는, 후성적 변형 DNA-표적화 시스템.

[0638] 38. 구현예 1-33, 36 및 37 중 어느 한 구현예에 있어서, 표적 부위 또는 각각의 표적 부위는 서열번호 650의 뉴클레오타이드 위치를 참조로 B형 간염 바이러스 서열의 1033bp-1749bp 사이의 염기쌍에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내에 위치하는, 후성적 변형 DNA-표적화 시스템.

[0639] 39. 구현예 1-37 중 어느 한 구현예에 있어서, 표적 부위 또는 각각의 표적 부위는 B형 간염 X 단백질(HBx) 시작 코돈 상류의 300개 염기쌍 내에 위치한 표적 영역 내에 있는, 후성적 변형 DNA-표적화 시스템.

[0640] 40. 하나 이상의 B형 간염 바이러스(HBV) 유전자의 전사를 억제하기 위한 적어도 하나의 DNA-표적화 모듈을 포함하는 후성적 변형 DNA-표적화 시스템으로서, 여기서 적어도 하나의 DNA-표적화 모듈 각각은 다음을 포함하는 융합 단백질을 포함하는 것인, 후성적 변형 DNA-표적화 시스템:

[0641] (a) B형 간염 X 단백질(HBx) 시작 코돈 상류의 300개 염기쌍 내에 있는 표적 영역 내의 표적 부위를 표적화하기 위한 DNA-결합 도메인; 및

[0642] (b) 적어도 하나의 전사 억제자 효과기 도메인.

[0643] 41. 구현예 1-40 중 어느 한 구현예에 있어서, 표적 부위 또는 각각의 표적 부위는 HBx 기저 코어 프로모터 영역에 위치하는, 후성적 변형 DNA-표적화 시스템.

[0644] 42. 구현예 1-40 중 어느 한 구현예에 있어서, 표적 부위 또는 각각의 표적 부위는 HBx 프로모터/인핸서 영역 내에 위치하는, 후성적 변형 DNA-표적화 시스템.

[0645] 43. 구현예 1-42 중 어느 한 구현예에 있어서, 표적 부위 또는 각각의 표적 부위는 B형 간염 X 단백질(HBx) 시작 코돈 상류의 250개 염기쌍 내에 걸쳐 있는 표적 영역 내에 있는, 후성적 변형 DNA-표적화 시스템.

[0646] 44. 구현예 1-43 중 어느 한 구현예에 있어서, 표적 부위 또는 각각의 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 HBV 게놈의 1060-1480bp 사이의 염기쌍에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내에 있는, 후성적 변형 DNA-표적화 시스템.

[0647] 45. 구현예 1-44 중 4, 표적 부위 또는 각각의 표적 부위는 B형 간염 X 단백질(HBx) 시작 코돈 상류의 150개 염기쌍 내에 걸쳐 있는 표적 영역 내에 있는, 후성적 변형 DNA-표적화 시스템.

[0648] 46. 구현예 1-45 중 어느 한 구현예에 있어서, 표적 부위 또는 각각의 표적 부위는 B형 간염 X 단백질(HBx) 시작 코돈 상류의 120개 염기쌍 내에 걸쳐 있는 표적 영역 내에 있는, 후성적 변형 DNA-표적화 시스템.

[0649] 47. 구현예 1-46 중 어느 한 구현예에 있어서, 표적 부위 또는 각각의 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 HBV 게놈의 1250-1374bp 사이의 염기쌍에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내에 있는, 후성적 변형 DNA-표적화 시스템.

[0650] 48. 구현예 1-47 중 어느 한 구현예에 있어서, 표적 부위 또는 각각의 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 HBV 게놈의 1255-1302bp 사이의 염기쌍에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내에 있는, 후성적 변형 DNA-표적화 시스템.

[0651] 49. 구현예 1-48 중 어느 한 구현예에 있어서, 표적 부위 또는 각각의 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 HBV 게놈의 1260-1300bp 사이의 염기쌍에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내에 있는, 후성적 변형 DNA-표적화 시스템.

[0652] 50. 구현예 1-49 중 어느 한 구현예에 있어서, 표적 부위 또는 각각의 표적 부위는 모든 B형 간염 바이러스 게놈과 최소 70% 상동성을 갖는, 후성적 변형 DNA-표적화 시스템.

[0653] 51. 구현예 1-50 중 어느 한 구현예에 있어서, 표적 부위 또는 각각의 표적 부위는 최소 1000개의 B형 간염 바이러스 게놈과 최소 70% 상동성을 갖는, 후성적 변형 DNA-표적화 시스템.

[0654] 52. 구현예 1-51 중 어느 한 구현예에 있어서, 표적 부위 또는 각각의 표적 부위는 최소 1000개의 B형 간염 바이러스 게놈과 최소 70% 상동적이며 최대 2개의 불일치를 포함하는, 후성적 변형 DNA-표적화 시스템.

[0655] 53. 구현예 1-52 중 어느 한 구현예에 있어서, 표적 부위 또는 각각의 표적 부위는 서열번호 1-195 중 어느 하나에 제시된 서열, 적어도 14개 뉴클레오타이드(nt)의 이의 연속 부분 또는 상기 서열 중 어느 하나의 상보적 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

[0656] 54. 구현예 1-53 어느 한 구현예에 있어서, 표적 부위 또는 각각의 표적 부위는 서열번호 175, 138, 192, 152, 118, 125, 185, 63, 116, 124, 35, 82 중 어느 하나에 제시된 서열, 적어도 14개 뉴클레오타이드(nt)의 이의 연속 부분, 또는 상기 중 어느 것의 상보적 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

[0657] 55. 구현예 1-54 중 어느 한 구현예에 있어서, 표적 부위 또는 각각의 표적 부위는 서열번호 175, 138, 192, 152, 118, 125, 185, 63, 116, 124, 35, 82 중 어느 하나에 제시된 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

[0658] 56. 구현예 1-53 중 어느 한 구현예에 있어서, 표적 부위 또는 각각의 표적 부위는 서열번호 5, 6, 12, 18, 22, 26, 29, 38, 42, 43, 51, 56, 61, 63, 68, 72, 75, 79, 82, 84, 88, 89, 98, 99, 113, 116, 121, 124, 125, 118, 130, 133, 135, 138, 143, 150, 152, 155, 158, 164, 165, 175, 176, 182, 185, 189, 190, 192 중 어느 하나에 제시된 서열, 적어도 14개 뉴클레오타이드(nt)의 이의 연속 부분 또는 상기 중 어느 하나의 상보적 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

[0659] 57. 구현예 1-53 및 56 중 어느 한 구현예에 있어서, 표적 부위 또는 각각의 표적 부위는 서열번호 5, 6, 12, 18, 22, 26, 29, 38, 42, 43, 51, 56, 61, 63, 68, 72, 75, 79, 82, 84, 88, 89, 98, 99, 113, 116, 121,

124, 125, 118, 130, 133, 135, 138, 143, 150, 152, 155, 158, 164, 165, 175, 176, 182, 185, 189, 190, 192 중 어느 하나에 제시된 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

[0660] 58. 구현예 1-53 및 58 중 어느 한 구현예에 있어서, 표적 부위 또는 각각의 표적 부위는 서열번호 12, 18, 20, 22, 26, 27, 46, 50, 63, 66, 73, 79, 185, 192 중 어느 하나에 제시된 서열, 적어도 14개 뉴클레오타이드의 이의 연속 부분 또는 상기 중 어느 하나의 상보적 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

[0661] 59. 구현예 1-53 및 58 중 어느 한 구현예에 있어서, 표적 부위 또는 각각의 표적 부위는 서열번호 12, 18, 20, 22, 26, 27, 46, 50, 63, 66, 73, 79, 185, 192 중 어느 하나에 제시된 것인, 후성적 변형 DNA-표적화 시스템.

[0662] 60. 구현예 1-53 및 56-59 중 어느 한 구현예에 있어서, 표적 부위 또는 각각의 표적 부위는 서열번호 22에 제시된 서열, 적어도 14개 뉴클레오타이드의 이의 연속 부분 또는 상기 중 어느 하나의 상보적 서열을 포함하고, 선택적으로 표적 부위는 서열번호 22에 제시된 것인, 후성적 변형 DNA-표적화 시스템.

[0663] 61. 구현예 1-53 및 56-59 중 어느 한 구현예에 있어서, 표적 부위 또는 각각의 표적 부위는 서열번호 63에 제시된 서열, 적어도 14개 뉴클레오타이드의 이의 연속 부분 또는 상기 중 어느 하나의 상보적 서열을 포함하고, 선택적으로 표적 부위는 서열번호 63에 제시된 것인, 후성적 변형 DNA-표적화 시스템.

[0664] 62. 구현예 2-53 중 어느 한 구현예에 있어서, gRNA 또는 각각의 gRNA는 서열번호 196-390 중 어느 하나에 제시된 서열을 포함하는 gRNA 스페이서 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

[0665] 63. 구현예 62에 있어서, gRNA 또는 각각의 gRNA는 서열번호 587에 제시된 서열을 추가로 포함하는, 후성적 변형 DNA-표적화 시스템.

[0666] 64. 구현예 2-53, 62 및 63 중 어느 한 구현예에 있어서, gRNA 또는 각각의 gRNA는 서열번호 391-585 중 어느 하나에 제시된 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

[0667] 65. 구현예 2-53 및 62-64 중 어느 한 구현예에 있어서, gRNA 또는 각각의 gRNA는 서열번호 391-585 중 어느 하나에 제시된 것인, 후성적 변형 DNA-표적화 시스템.

[0668] 66. 구현예 2-53 및 62-64 중 어느 한 구현예에 있어서, gRNA 또는 각각의 gRNA는 서열번호 370, 333, 387, 347, 313, 320, 380, 256, 258, 311, 319, 230, 272 중 어느 하나에 제시된 서열, 적어도 14개 뉴클레오타이드의 이의 연속 부분, 또는 상기 중 어느 하나의 상보적 서열을 포함하고, 선택적으로 gRNA 또는 각각의 gRNA는 서열번호 565, 528, 542, 508, 515, 575, 515, 453, 506, 514, 425 또는 472 중 어느 하나에 제시된 것인, 후성적 변형 DNA-표적화 시스템.

[0669] 67. 구현예 2-53 및 62-66 중 어느 한 구현예에 있어서, gRNA 또는 각각의 gRNA는 서열번호 370, 333, 387, 347, 313, 320, 380, 256, 258, 311, 319, 230, 272 중 어느 하나에 제시된 서열을 포함하고, 선택적으로 gRNA 또는 각각의 gRNA는 서열번호 565, 528, 542, 508, 515, 575, 515, 453, 506, 514, 425 또는 472 중 어느 하나에 제시된 것인, 후성적 변형 DNA-표적화 시스템.

[0670] 68. 구현예 2-53 및 62-65 중 어느 한 구현예에 있어서, gRNA 또는 각각의 gRNA는 서열번호 200, 201, 207, 217, 221, 224, 233, 237, 238, 246, 251, 256, 258, 263, 267, 274, 270, 277, 279, 283, 284, 293, 294, 308, 311, 313, 316, 319, 320, 325, 328, 330, 333, 338, 345, 347, 350, 353, 359, 360, 370, 371, 377, 380, 384, 385, 387 중 어느 하나에 제시된 서열. 적어도 14개 뉴클레오타이드의 이의 연속 부분 또는 상기 중 어느 하나의 상보적 서열을 포함하고, 선택적으로 여기서 gRNA 또는 각각의 gRNA는 서열번호 395, 402, 408, 412, 416, 419, 428, 432, 433, 441, 446, 451, 453, 458, 462, 465, 469, 472, 474, 478, 479, 488, 489, 503, 506, 508, 511, 514, 515, 520, 523, 525, 575, 528, 533, 540, 542, 545, 548, 554, 555, 565, 566, 572, 579, 580, 또는 582 중 어느 하나에 제시된 것인, 후성적 변형 DNA-표적화 시스템.

[0671] 69. 구현예 2-53, 62-65 및 68 중 어느 한 구현예에 있어서, gRNA 또는 각각의 gRNA는 서열번호 200, 201, 207, 217, 221, 224, 233, 237, 238, 246, 251, 256, 258, 263, 267, 274, 270, 277, 279, 283, 284, 293, 294, 308, 311, 313, 316, 319, 320, 325, 328, 330, 333, 338, 345, 347, 350, 353, 359, 360, 370, 371, 377, 380, 384, 385 또는 387 중 어느 하나에 제시된 서열을 포함하고, 선택적으로 gRNA 또는 각각의 gRNA는 서열번호 395, 402, 408, 412, 416, 419, 428, 432, 433, 441, 446, 451, 453, 458, 462, 465, 469, 472, 474, 478, 479, 488, 489, 503, 506, 508, 511, 514, 515, 520, 523, 525, 575, 528, 533, 540, 542, 545, 548, 554, 555, 565, 566, 572, 579, 580, 또는 582 중 어느 하나에 제시된 것인, 후성적 변형 DNA-표적화 시스템.

[0672] 70. 구현예 2-53 및 62-65 중 어느 한 구현예에 있어서, gRNA 또는 각각의 gRNA는 서열번호 207, 213, 215, 217, 221, 222, 241, 245, 258, 261, 268, 274, 380, 387 중 어느 하나에 제시된 서열, 적어도 14개 뉴클레오타이드의 이의 연속 부분, 또는 상기 중 어느 하나의 상보적 서열을 포함하고, 선택적으로 gRNA 또는 각각의 gRNA는 서열번호 402, 408, 410, 412, 416, 417, 436, 440, 453, 456, 463, 469, 575, 582 중 어느 하나에 제시된 것인, 후성적 변형 DNA-표적화 시스템.

[0673] 71. 구현예 2-53, 62-65 및 70 중 어느 한 구현예에 있어서, gRNA 또는 각각의 gRNA는 서열번호 207, 213, 215, 217, 221, 222, 241, 245, 258, 261, 268, 274, 380, 387 중 어느 하나에 제시된 서열을 포함하고, 선택적으로 gRNA 또는 각각의 gRNA는 서열번호 402, 408, 410, 412, 416, 417, 436, 440, 453, 456, 463, 469, 575, 582 중 어느 하나에 제시된 것인, 후성적 변형 DNA-표적화 시스템.

[0674] 72. 구현예 2-53, 62-65 및 68-71 중 어느 한 구현예에 있어서, gRNA 또는 각각의 gRNA는 서열번호 217에 제시된 서열, 적어도 14개 뉴클레오타이드의 이의 연속 부분 또는 상기 중 어느 하나의 상보적 서열을 포함하고, 선택적으로 gRNA 또는 각각의 gRNA는 서열번호 217에 제시된 서열을 포함하며, 선택적으로 gRNA 또는 각각의 gRNA는 서열번호 412 중 어느 하나에 제시된 것인, 후성적 변형 DNA-표적화 시스템.

[0675] 73. 구현예 2-53, 62-65 및 68-71 중 어느 한 구현예에 있어서, gRNA 또는 각각의 gRNA는 서열번호 258에 제시된 서열, 적어도 14개 뉴클레오타이드의 이의 연속 부분 또는 상기 중 어느 하나의 상보적 서열을 포함하고, 선택적으로 gRNA 또는 각각의 gRNA는 서열번호 258에 제시된 서열을 포함하며, 선택적으로 gRNA 또는 각각의 gRNA는 서열번호 453 중 어느 하나에 제시된 것인, 후성적 변형 DNA-표적화 시스템.

[0676] 74. 구현예 1-73 중 어느 한 구현예에 있어서, 표적 부위 또는 각각의 표적 부위는 모든 B형 간염 바이러스 게놈과 최소 90% 상동성인, 후성적 변형 DNA-표적화 시스템.

[0677] 75. 구현예 1-74 중 어느 한 구현예에 있어서, 표적 부위 또는 각각의 표적 부위는 최소 1000개의 B형 간염 바이러스 게놈과 최소 90% 상동성인, 후성적 변형 DNA-표적화 시스템.

[0678] 76. 구현예 1-75 중 어느 한 구현예에 있어서, 표적 부위 또는 각각의 표적 부위는 최소 1000개의 B형 간염 바이러스 게놈과 최소 90% 상동적이고 최대 2개의 불일치, 선택적으로 1개 또는 2개의 불일치를 포함하는, 후성적 변형 DNA-표적화 시스템.

[0679] 77. 구현예 1-53, 62-65 및 74-76 중 어느 한 구현예에 있어서, 표적 부위 또는 각각의 표적 부위는 서열번호 35-100 중 어느 하나에 제시된 서열, 적어도 14개 뉴클레오타이드(nt)의 이의 연속 부분 또는 상기 서열 중 어느 하나의 상보적 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

[0680] 78. 구현예 2-53, 62-65 및 74-77 중 어느 한 구현예에 있어서, gRNA 또는 각각의 gRNA는 서열번호 230-295 중 어느 하나에 제시된 서열을 포함하는 gRNA 스페이서 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

[0681] 79. 구현예 78에 있어서, gRNA 또는 각각의 gRNA는 추가로 서열번호 587에 제시된 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

[0682] 80. 구현예 2-53, 62-65 및 74-79 중 어느 한 구현예에 있어서, gRNA 또는 각각의 gRNA는 서열번호 230-295 중 어느 하나에 제시된 서열을 포함하고, 선택적으로 gRNA 또는 각각의 gRNA는 서열번호 425-490 중 어느 하나에 제시된 것인, 후성적 변형 DNA-표적화 시스템.

[0683] 81. 구현예 52-80 중 어느 한 구현예에 있어서, 최대 2개의 불일치는 gRNA 프로토스페이서의 5' 말단의 처음 12 nt에 위치하는, 후성적 변형 DNA-표적화 시스템.

[0684] 82. 구현예 1-51 및 53-73, 75 및 77-81 중 어느 한 구현예에 있어서, 표적 부위 또는 각각의 표적 부위는 최소 1000개의 B형 간염 바이러스 게놈과 최소 90% 상동적이며 불일치가 없는 것인, 후성적 변형 DNA-표적화 시스템.

[0685] 83. 구현예 1-53, 62-65, 74-76 및 82 중 어느 한 구현예에 있어서, 표적 부위는 서열번호 1-34 중 어느 하나에 제시된 서열, 적어도 14개 뉴클레오타이드(nt)의 이의 연속 부분 또는 상기 중 어느 하나의 상보적 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

[0686] 84. 구현예 2-53, 62-65, 74-76, 82 및 83 중 어느 한 구현예에 있어서, gRNA 또는 각각의 gRNA는 서열번호 196-229에 제시된 서열을 포함하는 gRNA 스페이서 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

[0687] 85. 구현예 62-84 중 어느 한 구현예에 있어서, gRNA 스페이서 서열은 길이가 14nt 내지 24nt 또는 16nt 내지

22nt인, 후성적 변형 DNA-표적화 시스템..

[0688] 86. 구현예 62-85 중 어느 한 구현예에 있어서, gRNA 스페이서 서열은 길이가 18 nt, 19 nt, 20 nt, 21 nt 또는 22 nt인, 후성적 변형 DNA-표적화 시스템.

[0689] 87. 구현예 62-86 중 어느 한 구현예에 있어서, gRNA 스페이서 서열은 안정성을 높이기 위해 변형된 뉴클레오타이드를 포함하는, 후성적 변형 DNA-표적화 시스템.

[0690] 88. 구현예 2-53, 62, 74-76 및 82-87 중 어느 한 구현예에 있어서, 상기 적어도 하나의 gRNA는 추가로 서열번호 587에 제시된 서열을 포함함.

[0691] 89. 구현예 2-53, 62, 74-76 및 82-88 중 어느 한 구현예에 있어서, gRNA 또는 각각의 gRNA는 서열번호 196-229 중 어느 하나에 제시된 서열을 포함하고, 선택적으로 gRNA 또는 각각의 gRNA는 서열번호 391-424 중 어느 하나에 제시된 것인, 후성적 변형 DNA-표적화 시스템.

[0692] 90. 구현예 2-53, 62, 74-76 및 82-89 중 어느 한 구현예에 있어서, gRNA 또는 각각의 gRNA는 서열번호 207, 213, 215, 217, 221, 222, 241, 245, 258, 261, 268, 274, 380, 387 중 어느 하나에 제시된 서열을 포함하고, 선택적으로 gRNA 또는 각각의 gRNA는 서열번호 402, 408, 410, 412, 416, 417, 436, 440, 453, 456, 463, 469, 575, 582 중 어느 하나에 제시된 것인, 후성적 변형 DNA-표적화 시스템.

[0693] 91. 구현예 2-53, 62, 74-76 및 82-90 중 어느 한 구현예에 있어서, gRNA는 서열번호 217에 제시된 서열을 포함하고, 선택적으로 gRNA는 서열번호 412에 제시된 것인, 후성적 변형 DNA-표적화 시스템.

[0694] 92. 구현예 2-91 중 후성적 변형 DNA-표적화 시스템, Cas 단백질 또는 이의 변이체는 Cas9 단백질 또는 이의 변이체인, 후성적 변형 DNA-표적화 시스템.

[0695] 93. 구현예 2-92 중 어느 한 구현예에 있어서, Cas 단백질 또는 이의 변이체는 Cas12 단백질 또는 이의 변이체인, 후성적 변형 DNA-표적화 시스템.

[0696] 94. 구현예 2-93 중 어느 한 구현예에 있어서, Cas 단백질 또는 이의 변이체는 변이체 Cas 단백질이고, 변이체 Cas 단백질은 뉴클레아제 활성이 없거나 비활성화된 Cas(dCas) 단백질인, 후성적 변형 DNA-표적화 시스템.

[0697] 95. 구현예 2-94 중 어느 한 구현예에 있어서. 변이체 Cas 단백질은 뉴클레아제 활성이 결여된 변이체 Cas9 단백질이거나 비활성화된 Cas9(dCas9) 단백질인, 후성적 변형 DNA-표적화 시스템

[0698] 96. 구현예 2-95 중 어느 한 구현예에 있어서, Cas9 단백질 또는 이의 변이체는 황색포도상구균(*Staphylococcus aureus*) Cas9 (SaCas9) 단백질 또는 이의 변이체인, 후성적 변형 DNA-표적화 시스템.

[0699] 97. 구현예 96에 있어서, 변이체 Cas9는 서열번호 596의 위치 번호를 참조로, D10A 및 N580A에서 선택된 적어도 하나의 아미노산 돌연변이를 포함하는 황색포도상구균 dCas9 단백질(dSaCas9)인, 후성적 변형 DNA-표적화 시스템.

[0700] 98. 구현예 96 또는 97에 이어서, 변이체 Cas9 단백질은 서열번호 597에 제시된 서열을 포함하거나, 서열번호 597과 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

[0701] 99. 구현예 2-95 중 어느 한 구현예에 있어서, Cas9 단백질 또는 이의 변이체는 화농성연쇄상구균(*Streptococcus pyogenes*) Cas9 (SpCas9) 단백질 또는 이의 변이체인, 후성적 변형 DNA-표적화 시스템

[0702] 100. 구현예 99에 있어서, 변이체 Cas9는 번호를 참조로 D10A 및 H840A 중에서 선택된 적어도 하나의 아미노산 돌연변이를 포함하는 화농성연쇄상구균(*Streptococcus pyogenes*) dCas9(dSpCas9) 단백질인, 후성적 변형 DNA-표적화 시스템.

[0703] 101. 구현예 99 또는 100에 있어서, 변이체 Cas9 단백질은 서열번호 599에 제시된 서열을 포함하거나, 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

[0704] 102. 구현예 1-5, 11-18, 20-33, 34-61 중 어느 한 구현예에 있어서, 상기 적어도 하나의 DNA-결합 도메인은 조작된 아연 집게 단백질(eZFP)을 포함하는, 후성적 변형 DNA-표적화 시스템.

[0705] 103. 구현예 1-5, 11-18, 20-33, 34-61 및 102 중 어느 한 구현예에 있어서, 적어도 하나의 DNA-결합 도메인은

eZFP인, 후성적 변형 DNA-표적화 시스템.

[0706] 104. 구현예 1-54, 11-18, 20-33, 34-61, 102 및 103 중 어느 한 구현예에 있어서, 표적 부위는 서열번호 1045, 1046, 1052 중 어느 하나에 제시된 뉴클레오타이드 서열, 최소 12nt의 이의 연속 부분, 또는 상기 중 어느 한 서열의 상보적 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

[0707] 105. 구현예 1-54, 11-18, 20-33, 34-61 및 102-104 중 어느 한 구현예에 있어서, 표적 부위는 서열번호 1045, 1046, 1052 중 어느 하나에 제시된 뉴클레오타이드 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

[0708] 106. 구현예 102-105 중 어느 한 구현예에 있어서, 아연 집게 단백질은 N-말단에서 C-말단 순서로 F1 내지 F6으로 표시되는 6개의 아연 집게를 포함하고, 여기서 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같은 것인 후성적 변형 DNA-표적화 시스템:

[0709] 1) F1:SEADRSR(서열번호 720) F2:DRSNLTR(서열번호 721)F3:QSSDLSR(서열번호 722) F4:YHWYLKK(서열번호 723) F5:RSDSLSV(서열번호 724) F6:QNAANRKT(서열번호 725);

[0710] 2) F1:RSDVLST(서열번호 726) F2:DNSSRTR(서열번호 727) F3:RPYTLRL(서열번호 728) F4:DSSHRTR(서열번호 729) F5:RSDHLSQ(서열번호 730) F6:DSSHRTR(서열번호 731);

[0711] 3) F1:RSDHLSQ(서열번호 732) F2:QSADRTK(서열번호 733) F3:RSDHLSQ(서열번호 734) F4:RRSDLKR(서열번호 735) F5:RSDHLSR(서열번호 736) F6:QSSDLRR(서열번호 737);

[0712] 4) F1:RSDNLSE(서열번호 738) F2:TSSNRKT(서열번호 739) F3:DRSHLTR(서열번호 740) F4:RSDALTQ(서열번호 741) F5:DRSALAR(서열번호 742) F6:RRFTLSK(서열번호 743);

[0713] 5) F1:RSDHLSE(서열번호 744) F2:QYSGRYY(서열번호 745) F3:HGQTLNE(서열번호 746) F4:QSGNLAR(서열번호 747) F5:RSDSLLR(서열번호 748)F6:CREYRGK(서열번호 749);

[0714] 6) F1:QSANRTT(서열번호 750) F2:RSANLTR(서열번호 751) F3:RSDVLSE(서열번호 752) F4:TSGHLSR(서열번호 753) F5:QSSDLSR(서열번호 754), F6:QWSTRKR(서열번호 755);

[0715] 7) F1:QSGNLAR(서열번호 756) F2:ATCCLAH(서열번호 757) F3:RWQYLPT(서열번호 758) F4:DRSALAR(서열번호 759) F5:RSDNLSE(서열번호 760)F6:KRCNLRC(서열번호 761);

[0716] 8) F1:NPANLTR(서열번호 762) F2:QNATRTK(서열번호 763) F3:QSGHLAR(서열번호 764) F4:NRHDRAK(서열번호 765) F5:RSDHLSE(서열번호 766), F6:QRRSRYK(서열번호 767);

[0717] 9) F1:QSSDLSR(서열번호 768) F2:HRSTRNR(서열번호 769) F3:RSDVLSA(서열번호 770) F4:DSRTRKN(서열번호 771) F5:QSGSLTR(서열번호 772) F6:DQSGLAH(서열번호 773);

[0718] 10) F1:QNPAQWR(서열번호 774) F2:RSADLSR(서열번호 775) F3:TSGSLSR(서열번호 776) F4:RSDHLSR(서열번호 777) F5:RSDSLLR(서열번호 778) F6:QSYDRFQ(서열번호 779);

[0719] 11) F1:TSGSLSR(서열번호 780) F2:RSDHLSR(서열번호 781) F3:RSDSLLR(서열번호 782) F4:QSYDRFQ(서열번호 783) F5:RSDNLST(서열번호 784) F6:DNRDRIK(서열번호 785);

[0720] 12) F1:DRSNLSR(서열번호 786) F2:LRQNLIM(서열번호 787) F3:ERGTLAR(서열번호 788) F4:RSDALTQ(서열번호 789) F5:RSDSLSQ(서열번호 790) F6:RKADRTR(서열번호 791);

[0721] 13) F1:QYCCLTN(서열번호 792) F2:TSGNLTR(서열번호 793) F3:QSSDLSR(서열번호 794) F4:FRYYLKR(서열번호 795) F5:QSGDLTR(서열번호 796) F6:DKGNLTK(서열번호 797);

[0722] 14) F1:TSGSLSR(서열번호 798) F2:RSDNLTT(서열번호 799) F3:QSGNLAR(서열번호 800) F4:DRTTLMR(서열번호 801) F5:QSGHLAR(서열번호 802) F6:QLTHLNS(서열번호 803);

[0723] 15) F1:IKHDLHR(서열번호 804) F2:RSANLTR(서열번호 805) F3:RSDNLAR(서열번호 806) F4:QNVSRPR(서열번호 807) F5:RSDDLSK(서열번호 808) F6:DSSHRTR(서열번호 809);

[0724] 16) F1:RSDNLAR(서열번호 810) F2:QNVSRPR(서열번호 811) F3:RSDDLSK(서열번호 812) F4:DSSHRTR(서열번호 813) F5:TSSNRKT(서열번호 814) F6:AQWTRAC(서열번호 815);

[0725] 17) F1:RSDDLSK(서열번호 816) F2:DSSHRTR(서열번호 817) F3:TSSNRKT(서열번호 818) F4:AQWTRAC(서열번호 819)

F5:RKQTRTT(서열번호 820) F6:HRSSLRR(서열번호 821);

[0726] 18) F1:QSAHRKN(서열번호 822) F2:TSSNRKT(서열번호 823) F3:RSDNLSA(서열번호 824) F4:RNNDRKT(서열번호 825) F5:TSGSLSR(서열번호 826) F6:QAGHLAK(서열번호 827);

[0727] 19) F1:RSDHLSQ(서열번호 828) F2:ASSTRTK(서열번호 829) F3:RSDDLTR(서열번호 830) F4:QKSNLSS(서열번호 831) F5:QSANRTT(서열번호 832) F6:QNATRTK(서열번호 833);

[0728] 20) F1:RSDTLSE(서열번호 834) F2:RRWTLVG(서열번호 835) F3:DRSNLSR(서열번호 836) F4:QSGDLTR(서열번호 837) F5:QSSDLSR(서열번호 838) F6:YHWYLKK(서열번호 839);

[0729] 21) F1:RSANLAR(서열번호 840) F2:RSDNLRE(서열번호 841) F3:RPYTLRL(서열번호 842) F4:HRSNLNK(서열번호 843) F5:QSGSLTR(서열번호 844) F6:TSANLSR(서열번호 845);

[0730] 22) F1:RSDDLVR(서열번호 846) F2:TSGSLVR(서열번호 847) F3:RSDKLVR(서열번호 848) F4:RSDELVR(서열번호 849) F5:TSHSLTE(서열번호 850) F6:RADNLTE(서열번호 851);

[0731] 23) F1:ERSHLRE(서열번호 852) F2:TSHSLTE(서열번호 853) F3:QAGHLAS(서열번호 854) F4:TSHSLTE(서열번호 855) F5:DPGHLVR(서열번호 856) F6:TSGNLVR(서열번호 857);

[0732] 24) F1:RADNLTE(서열번호 858) F2:TSGSLVR(서열번호 859) F3:RKDNLKN(서열번호 860) F4:QSSSLVR(서열번호 861) F5:RSDKLVR(서열번호 862) F6:DSGNLRV(서열번호 863);

[0733] 25) F1:QSSSLVR(서열번호 864) F2:QSGDLRR(서열번호 865) F3:RSDERKR(서열번호 866) F4:HRTTLTN(서열번호 867) F5:RSDHLTN(서열번호 868) F6:TSGELVR(서열번호 869);

[0734] 26) F1:QSGDLRR(서열번호 870) F2:RSDERKR(서열번호 871) F3:HRTTLTN(서열번호 872) F4:RSDHLTN(서열번호 873) F5:TSGELVR(서열번호 874) F6:RSDDLVR(서열번호 875);

[0735] 27) F1:QRAHLER(서열번호 876) F2:QLAHLRA(서열번호 877) F3:DPGHLVR(서열번호 878) F4:RRSACRR(서열번호 879) F5:RSDHLTT(서열번호 880) F6:QSSSLVR(서열번호 881); 및

[0736] 28) F1:QSSNLVR(서열번호 882) F2:RSDDLVR(서열번호 883) F3:THLDLIR(서열번호 884) F4:TSGNLTE(서열번호 885) F5:RRSACRR(서열번호 886) F6:RNDTLTE(서열번호 887).

[0737] 107. 구현예 102-106 중 어느 한 구현예에 있어서, 아연 집게 단백질은 N-말단에서 C-말단 순서로 F1에서 F6으로 표시되는 6개의 아연 집게를 포함하고,

[0738] 여기서 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같은 것인, 후성적 변형 DNA-표적화 시스템:

[0739] F1:QSAHRKN(서열번호 822) F2:TSSNRKT(서열번호 823) F3:RSDNLSA(서열번호 824) F4:RNNDRKT(서열번호 825) F5:TSGSLSR(서열번호 826) F6:QAGHLAK(서열번호 827).

[0740] 108. 구현예 102-106 중 어느 한 구현예에 있어서, 아연 집게 단백질은 N-말단에서 C-말단 순서로 F1에서 F6까지 표시된 6개의 아연 집게를 포함하고,

[0741] 여기서 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같은 것인, 후성적 변형 DNA-표적화 시스템:

[0742] F1:RSDHLSQ(서열번호 828) F2:ASSTRTK(서열번호 829) F3:RSDDLTR(서열번호 830) F4:QKSNLSS(서열번호 831) F5:QSANRTT(서열번호 832) F6:QNATRTK(서열번호 833),

[0743] 109. 구현예 102-106 중 어느 한 구현예에 있어서, 아연 집게 단백질은 N-말단에서 C-말단의 순서로 F1에서 F6까지 표시된 6개의 아연 집게를 포함하고,

[0744] 여기서 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같은 것인, 후성적 변형 DNA-표적화 시스템:

[0745] F1:QSSSLVR(서열번호 864) F2:QSGDLRR(서열번호 865) F3:RSDERKR(서열번호 866) F4:HRTTLTN(서열번호 867) F5:RSDHLTN(서열번호 868) F6:TSGELVR(서열번호 869).

[0746] 110. 후성적 변형 DNA-표적화 시스템으로서, 상기 후성적 변형 DNA-표적화 시스템으로서은 a) 하나 이상의 HBV 유전자 또는 이의 조절 요소의 표적 부위에 결합하는 eZFP 및 b) 하나 이상의 HBV 유전자의 전사를 억제하는 적어도 하나의 효과기 도메인을 포함하고, 여기서 아연 집게 단백질은 N-말단에서 C-말단 순서로 F1에서 F6까지 표시된 6개의 아연 집게를 포함하며, 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같은 것인, 후성적 변

형 DNA-표적화 시스템:

[0747] 1) F1:SEADRSR(서열번호 720) F2:DRSNLTR(서열번호 721)F3:QSSDLSR(서열번호 722) F4:YHWYLKK(서열번호 723) F5:RSDSLSV(서열번호 724) F6:QNAANRKT(서열번호 725);

[0748] 2) F1:RSDVLST(서열번호 726) F2:DNSSRTR(서열번호 727) F3:RPYTLRL(서열번호 728) F4:DSSHRTR(서열번호 729) F5:RSDHLSQ(서열번호 730) F6:DSSHRTR(서열번호 731);

[0749] 3) F1:RSDHLSQ(서열번호 732) F2:QSADRTK(서열번호 733) F3:RSDHLSQ(서열번호 734) F4:RRSDLKR(서열번호 735) F5:RSDHLSR(서열번호 736) F6:QSSDLRR(서열번호 737);

[0750] 4) F1:RSDNLSE(서열번호 738) F2:TSSNRKT(서열번호 739) F3:DRSHLTR(서열번호 740) F4:RSDALTQ(서열번호 741) F5:DRSALAR(서열번호 742) F6:RRFTLSK(서열번호 743);

[0751] 5) F1:RSDHLSE(서열번호 744) F2:QYSGRYY(서열번호 745) F3:HGQTLNE(서열번호 746) F4:QSGNLAR(서열번호 747) F5:RSDSLLR(서열번호 748)F6:CREYRGK(서열번호 749);

[0752] 6) F1:QSANRTT(서열번호 750) F2:RSANLTR(서열번호 751) F3:RSDVLSE(서열번호 752) F4:TSGHLSR(서열번호 753) F5:QSSDLSR(서열번호 754), F6:QWSTRKR(서열번호 755);

[0753] 7) F1:QSGNLAR(서열번호 756) F2:ATCCLAH(서열번호 757) F3:RWQYLPT(서열번호 758) F4:DRSALAR(서열번호 759) F5:RSDNLSE(서열번호 760)F6:KRCNLRC(서열번호 761);

[0754] 8) F1:NPANLTR(서열번호 762) F2:QNATRTK(서열번호 763) F3:QSGHLAR(서열번호 764) F4:NRHDRAK(서열번호 765) F5:RSDHLSE(서열번호 766), F6:QRRSRYK(서열번호 767);

[0755] 9) F1:QSSDLSR(서열번호 768) F2:HRSTRNR(서열번호 769) F3:RSDVLSA(서열번호 770) F4:DSRTRKN(서열번호 771) F5:QSGSLTR(서열번호 772) F6:DQSGLAH(서열번호 773);

[0756] 10) F1:QNPAQWR(서열번호 774) F2:RSADLSR(서열번호 775) F3:TSGSLSR(서열번호 776) F4:RSDHLSR(서열번호 777) F5:RSDSLLR(서열번호 778) F6:QSYDRFQ(서열번호 779);

[0757] 11) F1:TSGSLSR(서열번호 780) F2:RSDHLSR(서열번호 781) F3:RSDSLLR(서열번호 782) F4:QSYDRFQ(서열번호 783) F5:RSDNLST(서열번호 784) F6:DNRDRIK(서열번호 785);

[0758] 12) F1:DRSNLSR(서열번호 786) F2:LRQNLIM(서열번호 787) F3:ERGTLAR(서열번호 788) F4:RSDALTQ(서열번호 789) F5:RSDSLSQ(서열번호 790) F6:RKADRTR(서열번호 791);

[0759] 13) F1:QYCCLTN(서열번호 792) F2:TSGNLTR(서열번호 793) F3:QSSDLSR(서열번호 794) F4:FRYYLKR(서열번호 795) F5:QSGDLTR(서열번호 796) F6:DKGNLTK(서열번호 797);

[0760] 14) F1:TSGSLSR(서열번호 798) F2:RSDNLTT(서열번호 799) F3:QSGNLAR(서열번호 800) F4:DRTTLMR(서열번호 801) F5:QSGHLAR(서열번호 802) F6:QLTHLNS(서열번호 803);

[0761] 15) F1:IKHDLHR(서열번호 804) F2:RSANLTR(서열번호 805) F3:RSDNLAR(서열번호 806) F4:QNVSRPR(서열번호 807) F5:RSDDLSK(서열번호 808) F6:DSSHRTR(서열번호 809);

[0762] 16) F1:RSDNLAR(서열번호 810) F2:QNVSRPR(서열번호 811) F3:RSDDLSK(서열번호 812) F4:DSSHRTR(서열번호 813) F5:TSSNRKT(서열번호 814) F6:AQWTRAC(서열번호 815);

[0763] 17) F1:RSDDLSK(서열번호 816) F2:DSSHRTR(서열번호 817) F3:TSSNRKT(서열번호 818) F4:AQWTRAC(서열번호 819) F5:RKQTRTT(서열번호 820) F6:HRSSLRR(서열번호 821);

[0764] 18) F1:QSAHRKN(서열번호 822) F2:TSSNRKT(서열번호 823) F3:RSDNLSA(서열번호 824) F4:RNNDRKT(서열번호 825) F5:TSGSLSR(서열번호 826) F6:QAGHLAK(서열번호 827);

[0765] 19) F1:RSDHLSQ(서열번호 828) F2:ASSTRTK(서열번호 829) F3:RSDDLTR(서열번호 830) F4:QKSNLSS(서열번호 831) F5:QSANRTT(서열번호 832) F6:QNATRTK(서열번호 833);

[0766] 20) F1:RSDTLSE(서열번호 834) F2:RRWTLVG(서열번호 835) F3:DRSNLSR(서열번호 836) F4:QSGDLTR(서열번호 837) F5:QSSDLSR(서열번호 838) F6:YHWYLKK(서열번호 839);

[0767] 21) F1:RSANLAR(서열번호 840) F2:RSDNLRE(서열번호 841) F3:RPYTLRL(서열번호 842) F4:HRSNLNK(서열번호 843)

F5:QSGSLTR(서열번호 844) F6:TSANLSR(서열번호 845);

[0768] 22) F1:RSDDLVR(서열번호 846) F2:TSGSLVR(서열번호 847) F3:RSDKLVR(서열번호 848) F4:RSDELVR(서열번호 849) F5:TSHSLTE(서열번호 850) F6:RADNLTE(서열번호 851);

[0769] 23) F1:ERSHLRE(서열번호 852) F2:TSHSLTE(서열번호 853) F3:QAGHLAS(서열번호 854) F4:TSHSLTE(서열번호 855) F5:DPGHLVR(서열번호 856) F6:TSGNLVR(서열번호 857);

[0770] 24) F1:RADNLTE(서열번호 858) F2:TSGSLVR(서열번호 859) F3:RKDNLKN(서열번호 860) F4:QSSSLVR(서열번호 861) F5:RSDKLVR(서열번호 862) F6:DSGNLRV(서열번호 863);

[0771] 25) F1:QSSSLVR(서열번호 864) F2:QSGDLRR(서열번호 865) F3:RSDERKR(서열번호 866) F4:HRTTLTN(서열번호 867) F5:RSDHLTN(서열번호 868) F6:TSGELVR(서열번호 869);

[0772] 26) F1:QSGDLRR(서열번호 870) F2:RSDERKR(서열번호 871) F3:HRTTLTN(서열번호 872) F4:RSDHLTN(서열번호 873) F5:TSGELVR(서열번호 874) F6:RSDDLVR(서열번호 875);

[0773] 27) F1:QRAHLER(서열번호 876) F2:QLAHLRA(서열번호 877) F3:DPGHLVR(서열번호 878) F4:RRSACRR(서열번호 879) F5:RSDHLTT(서열번호 880) F6:QSSSLVR(서열번호 881); 및

[0774] 28) F1:QSSNLVR(서열번호 882) F2:RSDDLVR(서열번호 883) F3:THLDLIR(서열번호 884) F4:TSGNLTE(서열번호 885) F5:RRSACRR(서열번호 886) F6:RNDTLTE(서열번호 887).

[0775] 111. 구현예 102-110 중 어느 한 구현예에 있어서, 조작된 아연 집게 단백질은 서열번호 692-719 중 어느 하나에 제시된 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

[0776] 112. 구현예 102-111 중 어느 한 구현예에 있어서, 조작된 아연 집게 단백질은 서열번호 888-915 중 어느 하나에 제시된 서열 또는 이의 일부에 의해 인코딩되거나, 또는 상기 서열과 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩되는, 후성적 변형 DNA-표적화 시스템.

[0777] 113. 후성적 변형 DNA-표적화 시스템으로서, 상기 후성적 변형 DNA-표적화 시스템은 a) 하나 이상의 HBV 유전자 또는 이의 조절 요소의 표적 부위에 결합하는 조작된 아연 집게 단백질, 및 b) 하나 이상의 HBV 유전자의 전사를 억제하는 적어도 하나의 효과기 도메인을 포함하고, 여기서 아연 집게 단백질은 N-말단에서 C-말단 순서로 F1에서 F6까지 표시된 6개의 아연 집게를 포함하며,

[0778] 여기서 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같은 것인 후성적 변형 DNA-표적화 시스템:

[0779] F1:QSAHRKN(서열번호 822) F2:TSSNRKT(서열번호 823) F3:RSDNLSA(서열번호 824) F4:RNNDRKT(서열번호 825) F5:TSGSLSR(서열번호 826) F6:QAGHLAK(서열번호 827).

[0780] 114. 구현예 102-113 중 어느 한 구현예에 있어서, eZFP는 서열번호 709에 제시된 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

[0781] 115. 구현예 102-114 중 어느 한 구현예에 있어서, 조작된 아연 집게 단백질은 서열번호 709 중 어느 하나에 제시된 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

[0782] 116. 구현예 102-115 중 어느 한 구현예에 있어서, 조작된 아연 집게 단백질은 서열번호 905에 제시된 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩되는, 후성적 변형 DNA-표적화 시스템.

[0783] 117. 구현예 102-116 중 어느 한 구현예에 있어서, 조작된 아연 집게 단백질은 서열번호 905 중 어느 하나에 제시된 서열에 의해 인코딩되는, 후성적 변형 DNA-표적화 시스템.

[0784] 118. 후성적 변형 DNA-표적화 시스템으로서, 상기 후성적 변형 DNA-표적화 시스템은 a) 하나 이상의 HBV 유전자 또는 이의 조절 요소의 표적 부위에 결합하는 조작된 아연 집게 단백질, 및 b) 하나 이상의 HBV 유전자의 전사를 억제하는 적어도 하나의 효과기 도메인을 포함하고,

[0785] 여기서 아연 집게 단백질은 N-말단에서 C-말단의 순서로 F1에서 F6까지 표시된 6개의 아연 집게를 포함하며, 각

아연 집게 인식 영역의 아미노산 서열은 다음과 같은 것인, 후성적 변형 DNA-표적화 시스템:

[0786] F1:RSDHLSQ(서열번호 828) F2:ASSTRTK(서열번호 829) F3:RSDDLTR(서열번호 830) F4:QKSNLSS(서열번호 831) F5:QSANRTT(서열번호 832) F6:QNATRTK(서열번호 833).

[0787] 119. 구현예 102-112 및 118 중 어느 한 구현예에 있어서, 조작된 아연 집게 단백질은 서열번호 710에 제시된 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

[0788] 120. 구현예 102-112, 118 또는 119 중 어느 한 구현예에 있어서, 조작된 아연 집게 단백질은 서열번호 710 중 어느 하나에 제시된 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

[0789] 121. 구현예 102-112 및 118-120 중 어느 한 구현예에 있어서, 조작된 아연 집게 단백질은 서열번호 906에 제시된 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩되는, 후성적 변형 DNA-표적화 시스템.

[0790] 122. 구현예 102-112 및 118-121 중 어느 한 구현예에 있어서, 조작된 아연 집게 단백질은 서열번호 906 중 어느 하나에 제시된 서열에 의해 인코딩되는, 후성적 변형 DNA-표적화 시스템.

[0791] 123. 후성적 변형 DNA-표적화 시스템으로서, 상기 후성적 변형 DNA-표적화 시스템은 a) 하나 이상의 HBV 유전자 또는 이의 조절 요소의 표적 부위에 결합하는 조작된 아연 집게 단백질, 및 b) 하나 이상의 HBV 유전자의 전사를 억제하는 적어도 하나의 효과기 도메인을 포함하고,

[0792] 여기서 아연 집게 단백질은 N 말단에서 C 말단까지 순서대로 F1에서 F6까지 표시되는 6개의 아연 집게를 포함하며,

[0793] 여기서 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같은 것인, 후성적 변형 DNA-표적화 시스템:

[0794] F1:QSSSLVR(서열번호 864) F2:QSGDLRR(서열번호 865) F3:RSDERKR(서열번호 866) F4:HRTTLTN(서열번호 867) F5:RSDHLTN(서열번호 868) F6:TSGELVR(서열번호 869).

[0795] 124. 구현예 102-112 및 123 중 어느 한 구현예에 있어서, 조작된 아연 집게 단백질은 서열번호 716에 제시된 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

[0796] 125. 구현예 102-112, 123 또는 124 중 어느 한 구현예에 있어서, 상기 조작된 아연 집게 단백질은 서열번호 716 중 어느 하나에 제시된 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

[0797] 126. 구현예 102-112 및 123-125 중 어느 한 구현예에 있어서, 조작된 아연 집게 단백질은 서열번호 912에 제시된 서열 또는 이의 일부에 의해 인코딩되거나, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩되는, 후성적 변형 DNA-표적화 시스템.

[0798] 127. 구현예 102-112 및 123-126 중 어느 한 구현예에 있어서, 조작된 아연 집게 단백질은 서열번호 912 중 어느 하나에 제시된 서열에 의해 인코딩되는, 후성적 변형 DNA-표적화 시스템.

[0799] 128. 구현예 1-127 중 어느 한 구현예에 있어서, 적어도 하나의 효과기 도메인은 전사 억제를 유도하는, 후성적 변형 DNA-표적화 시스템.

[0800] 129. 구현예 1-128 중 어느 한 구현예에 있어서, 적어도 하나의 효과기 도메인은 DNA 메틸트랜스퍼라제인, 후성적 변형 DNA-표적화 시스템.

[0801] 130. 구현예 1-129 중 어느 한 구현예에 있어서, 상기 적어도 하나의 효과기 도메인은 DNA 메틸트랜스퍼라제와 헤테로크로마틴 유도 인자를 동원할 수 있는 억제자 도메인을 포함하거나, 선택적으로 헤테로크로마틴 유도 인자는 히스톤 메틸트랜스퍼라제를 포함하는, 후성적 변형 DNA-표적화 시스템.

[0802] 131. 구현예 1-130 중 어느 한 구현예에 있어서, 적어도 하나의 효과기 도메인은 DNA 메틸트랜스퍼라제와 히스톤 메틸트랜스퍼라제를 포함하는, 후성적 변형 DNA-표적화 시스템.

[0803] 132. 구현예 1-131 중 어느 한 구현예에 있어서, 적어도 하나의 효과기 도메인은 KRAB 억제자 도메인, ERF 억제자 도메인, Mxi1 억제자 도메인, SID4X 억제자 도메인, Mad-SID 억제자 도메인, LSD1 억제자 도메인, 또는 DNMT3A, DNMT3A-3L, DNMT3A/L-KRAB 융합 억제자 도메인, DNMT3B 도메인 결합 단백질, EZH2 억제자 도메인, 또

는 LSD1 억제자 도메인, 또는 상기 중 임의의 것의 변이체 중에서 선택되는, 후성적 변형 DNA-표적화 시스템.

[0804] 133. 구현예 1-132 중 어느 한 구현예에 있어서, 적어도 하나의 효과기 도메인은 서열번호 590, 또는 600-608, 651, 661, 664, 665, 666, 668 및 669 중 어느 하나로부터 선택된 서열 또는 이의 도메인, 이의 일부 또는 상기 중 어느 하나와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

[0805] 134. 구현예 1-133 중 어느 한 구현예에 있어서, 적어도 하나의 효과기 도메인은 KRAB 도메인 또는 그 변이체를 포함하는, 후성적 변형 DNA-표적화 시스템.

[0806] 135. 구현예 1-134 중 어느 한 구현예에 있어서, 적어도 하나의 효과기 도메인은 서열번호 590에 제시된 서열, 이의 일부 또는 상기 중 어느 것과 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

[0807] 136. 구현예 1-134 중 어느 한 구현예에 있어서, 적어도 하나의 효과기 도메인은 DNMT3A/L 도메인 또는 이의 변이체를 포함하는, 후성적 변형 DNA-표적화 시스템.

[0808] 137. 구현예 1-136 중 어느 한 구현예에 있어서;

[0809] 적어도 하나의 효과기 도메인은 서열번호 604 및 607에 제시된 서열, 이의 일부 또는 앞서 언급한 서열 중 어느 하나와 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 이상의 서열 동일성을 갖는 아미노산 서열을 포함하거나; 또는

[0810] 적어도 하나의 효과기 도메인은 서열번호 651에 제시된 서열, 이의 일부, 또는 서열번호 651과 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

[0811] 138. 구현예 1-137 중 어느 한 구현예에 있어서, 융합 단백질은 DNMT3A/3L-dCas9-KRAB 융합 단백질을 포함하는, 후성적 변형 DNA-표적화 시스템.

[0812] 139. 구현예 1-138 중 어느 한 구현예에 있어서, 융합 단백질은 서열번호 645에 제시된 서열, 이의 일부 또는 앞서 언급한 서열과 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 이상의 서열 동일성을 갖는 아미노산 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

[0813] 140. 구현예 1-139 중 어느 한 구현예에 있어서, 적어도 하나의 효과기 도메인은 DNA-결합 도메인 또는 이의 구성 요소의 N-말단, C-말단, 또는 N-말단과 C-말단 둘 다에 융합되는, 후성적 변형 DNA-표적화 시스템.

[0814] 141. 구현예 1-101, 또는 138-140 중 어느 한 구현예에 있어서, 융합 단백질은 서열번호 680에 제시된 서열, 이의 일부, 또는 앞서 언급한 서열과 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99% 이상의 서열 동일성을 갖는 아미노산 서열에 의해 인코딩되는, 후성적 변형 DNA-표적화 시스템.

[0815] 142. 구현예 1-101, 또는 138-140 중 어느 한 구현예에 있어서, 융합 단백질은 서열번호 680에 제시된 서열에 의해 인코딩되는, 후성적 변형 DNA-표적화 시스템.

[0816] 143. 구현예 1-5, 11-18, 20-33, 34-61 및 102-140 중 어느 한 구현예에 있어서, 융합 단백질은 서열번호 916-943 중 어느 하나에 제시된 서열, 이의 일부 또는 앞서 언급한 서열과 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 이상의 서열 동일성을 갖는 핵산 서열에 의해 인코딩되는, 후성적 변형 DNA-표적화 시스템.

[0817] 144. 구현예 1-5, 11-18, 20-33, 34-61, 102-140 및 143 중 어느 한 구현예에 있어서, 융합 단백질은 서열번호 944-971 중 어느 하나에 제시된 서열, 이의 일부 또는 앞서 언급한 서열과 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 이상의 서열 동일성을 갖는 아미노산 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

[0818] 145. 구현예 1-5, 11-18, 20-33, 34-61, 102-140 및 143-144 중 어느 한 구현예에 있어서, 융합 단백질은 서열번호 961, 962 또는 968 중 어느 하나에 제시된 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

[0819] 146. 구현예 1-5, 11-18, 20-33, 34-61, 102-140 및 143-145 중 어느 한 구현예에 있어서, 융합 단백질은 DNMT3A/3L-eZFP-KRAB 융합 단백질을 포함하는, 후성적 변형 DNA-표적화 시스템.

[0820] 147. 구현예 1-4, 11-18, 20-33, 34-61, 102-140 및 143-146 중 어느 한 구현예에 있어서, 융합 단백질은 서열번호 972-999 중 어느 하나에 제시된 서열, 이의 일부 또는 앞서 언급한 서열과 90%, 91%, 92%, 93%, 94%,

95%, 96%, 97%, 98% 또는 99% 이상의 서열 동일성을 갖는 핵산 서열에 의해 인코딩되는, 후성적 변형 DNA-표적화 시스템.

[0821] 148. 구현예 1-5, 11-18, 20-33, 34-61, 102-140 및 143-147 중 어느 한 구현예에 있어서, 융합 단백질은 서열 번호 933, 934 또는 940 중 어느 하나에 제시된 서열, 이의 일부 또는 앞서 언급한 서열과 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 이상의 서열 동일성을 갖는 핵산 서열에 의해 인코딩되는,

[0822] 149. 구현예 1-5, 11-18, 20-33, 34-61, 102-140 및 143-148 중 어느 한 구현예에 있어서, 융합 단백질은 서열 번호 1000-1027 중 어느 하나에 제시된 서열, 이의 일부 또는 앞서 언급한 서열과 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 이상의 서열 동일성을 갖는 아미노산 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

[0823] 150. 구현예 1-4, 11-18, 20-33, 34-61, 102-140 및 143-149 중 어느 한 구현예에 있어서, 융합 단백질은 서열 번호 1017, 1018 또는 1024 중 어느 하나에 제시된 서열, 이의 일부 또는 앞서 언급한 서열과 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 이상의 서열 동일성을 갖는 아미노산 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

[0824] 151. 구현예 1-150 중 어느 한 구현예에 있어서, 융합 단백질은 추가로 하나 이상의 핵 국소화 신호(NLS)를 포함하는, 후성적 변형 DNA-표적화 시스템.

[0825] 152. 구현예 1-151 중 어느 한 구현예에 있어서, 융합 단백질은 다음 중 두 개 이상을 연결하는 하나 이상의 링커를 더 포함하는, 후성적 변형 DNA-표적화 시스템: DNA-결합 도메인, 적어도 하나의 효과기 도메인 및 하나 이상의 핵 국소화 신호.

[0826] 153. 구현예 1-152 중 어느 한 구현예에 있어서, DNA-표적화 시스템은 모든 B형 간염 바이러스 게놈을 표적화하는, 후성적 변형 DNA-표적화 시스템.

[0827] 154. 구현예 1-153 중 어느 한 구현예에 있어서, DNA-표적화 시스템은 모든 B형 간염 바이러스 게놈의 최소 70%를 표적화하는, 후성적 변형 DNA-표적화 시스템.

[0828] 155. 구현예 1-154 중 어느 한 구현예에 있어서, DNA-표적화 시스템은 모든 B형 간염 바이러스 게놈의 최소 60%를 표적화하는, 후성적 변형 DNA-표적화 시스템.

[0829] 156. 구현예 1-155 중 어느 한 구현예에 있어서, DNA-표적화 시스템은 모든 B형 간염 바이러스 게놈의 최소 50%를 표적화하는, 후성적 변형 DNA-표적화 시스템.

[0830] 157. 구현예 1-156 중 어느 한 구현예에 있어서, DNA-표적화 시스템은 표적 부위 또는 그 근처에서 유전적 파괴 또는 DNA 파손을 도입할 수 없는, 후성적 변형 DNA-표적화 시스템

[0831] 158. 구현예 1-157 중 어느 한 구현예에 있어서, 하나 이상의 HBV 유전자의 전사 억제가 HBV DNA 서열의 RNA 수준 및/또는 단백질 수준의 감소를 초래하는, 후성적 변형 DNA-표적화 시스템.

[0832] 159. 구현예 1-158 중 어느 한 구현예에 있어서, 전사 억제는 총 B형 간염 바이러스 RNA 전사 수치의 감소를 포함하는, 후성적 변형 DNA-표적화 시스템.

[0833] 160. 구현예 1-159 중 어느 한 구현예에 있어서, 전사 억제는 B형 간염 pre-코어("preC"), pre-게놈("pgRNA"), preS1, preS2/S 및 HBx 수준의 감소를 포함하는, 후성적 변형 DNA-표적화 시스템.

[0834] 161. 구현예 1-160 중 어느 한 구현예에 있어서, 전사 억제는 HBx 수치의 감소를 포함하는, 후성적 변형 DNA-표적화 시스템.

[0835] 162. 구현예 1-161 중 어느 한 구현예에 있어서, 전사 억제는 B형 간염 표면 항원(HBsAg) 및/또는 B형 간염 바이러스 코어 관련 항원(HbcrAg) 단백질 수치의 감소를 포함하는, 후성적 변형 DNA-표적화 시스템.

[0836] 163. 구현예 1-162 중 어느 한 구현예에 있어서, 전사 억제는 HbsAg 전사체 및/또는 단백질 수치를 최소 90% 감소시키는 것을 포함하는, 후성적 변형 DNA-표적화 시스템.

[0837] 164. 구현예 1-163 중 어느 한 구현예에 있어서, 전사 억제는 cccDNA에서 HbcrAg 전사체 및/또는 단백질 수준이 최소 50% 감소하는 것을 포함하는, 후성적 변형 DNA-표적화 시스템.

[0838] 165. B형 간염 바이러스 DNA 서열의 표적 부위에 결합하는 가이드 RNA(gRNA).

[0839] 166. 구현예 165에 이Tdjj, B형 간염 바이러스 DNA 서열은 B형 간염(HBV) 유전자 또는 이의 조절 요소인, gRNA.

[0840] 167. 구현예 165 또는 구현예 166에 있어서, 표적 부위는 공유적으로 폐쇄된 원형, DNA(cccDNA) 형태, 이완된 원형 DNA(rcDNA) 형태로 존재하고/존재하거나 인간 게놈 DNA에 통합된, gRNA.

[0841] 168. 구현예 165-167 중 어느 한 구현예에 있어서, 표적 부위는 HBV 복제 및/또는 HBV 전사를 제어하는 데 관여하는 유전자 또는 이의 조절 요소에 있거나 그 근처에 있는, gRNA.

[0842] 169. 구현예 168에 있어서, HBV 복제 및/또는 HBV 전사를 제어하는데 관여하는 유전자는 폴리머라제, 외피 단백질, 캡시드 단백질, 전사 인자 또는 전사 전사 활성화제를 인코딩하는, gRNA.

[0843] 170. 구현예 168 또는 구현예 169에 있어서, HBV 복제 및/또는 HBV 전사를 제어하는 데 관여하는 유전자는 폴리머라제 유전자, S-패밀리 유전자, X-유전자 또는 코어-패밀리 유전자인, gRNA.

[0844] 171. 구현예 165-170 중 어느 한 구현예에 있어서, 표적 부위는 B형 간염 바이러스 단백질 X(HBx)를 인코딩하는 X 유전자의 유전자 또는 이의 조절 요소에 있는, gRNA.

[0845] 172. 구현예 165-171 중 어느 한 구현예에 있어서, 표적 부위는 HBV 복제 및/또는 HBV 전사를 제어하는 조절 요소에 있거나 그 근처에 있는, gRNA.

[0846] 173. 구현예 166-172 중 어느 한 구현예에 있어서, 조절 요소는 프로모터 영역인, gRNA.

[0847] 174. 구현예 173에 이어서, 프로모터 영역은 pre-S1 프로모터, pre-S2 프로모터, X 프로모터 또는 기저 코어 프로모터인, gRNA.

[0848] 175. 구현예 166-172 중 어느 한 구현예에 있어서, 조절 요소는 인핸서 영역인, gRNA.

[0849] 176 구현예 175에 있어서, 인핸서 영역은 Enh1 또는 Enh2 인핸서 영역인, gRNA.

[0850] 177. 구현예 166-172 중 어느 한 구현예에 있어서, 조절 요소는 전사 처리 제어 영역인, gRNA.

[0851] 178. 구현예 165-171 중 어느 한 구현예에 있어서, 표적 부위는 코딩 영역인, gRNA.

[0852] 179. 구현예 165-178 중 어느 한 구현예에 있어서, 표적 부위는 전사 개시 부위의 500bp 이내, 1000bp 이내, 1500bp 이내에 위치하는, gRNA.

[0853] 180. 구현예 165-179 중 어느 한 구현예에 있어서, 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 HBV 게놈의 0-3300 염기쌍(bp) 사이, 선택적으로 HBV 게놈의 0-3189 bp 사이에 위치한 염기쌍에 상응하는 서열을 갖는 표적 영역 내에 위치하는, gRNA.

[0854] 181. 구현예 165-180 중 어느 한 구현예에 있어서, 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 HBV 게놈의 43bp-490bp, 1033bp-1749bp, 1800bp-1950bp 또는 2953bp-3182bp 사이의 염기쌍에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내에 위치하는, gRNA.

[0855] 182. 구현예 165-181 중 어느 한 구현예에 있어서, 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 HBV 게놈의 1 bp-42 bp, 491 bp-1032 bp, 1750 bp-1799 bp 또는 1951 bp-2952 bp 사이의 염기쌍에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내에 위치하는, gRNA.

[0856] 183. 구현예 165-182 중 어느 한 구현예에 있어서, 표적 부위는 HBV 게놈의 CpG 섬에 있는, gRNA.

[0857] 184. 구현예 165-183 중 어느 한 구현예에 있어서, 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 HBV 게놈의 67bp-392bp, 1033bp-1749bp 또는 2215bp-2490bp 사이의 염기쌍에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내에 위치하는, gRNA.

[0858] 185. 구현예 165-184 중 어느 한 구현예에 있어서, 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 HBV 게놈의 1033bp-1749bp 사이의 염기쌍에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내에 위치하는, gRNA.

[0859] 186. 구현예 165-185 중 어느 한 구현예에 있어서, 표적 부위는 B형 간염 X 단백질(HBx) 시작 코돈 상류의 300개 염기쌍 내에 걸쳐 있는 표적 영역 내에 있는, gRNA.

[0860] 187. B형 간염 X 단백질(HBx) 시작 코돈 상류의 300개 염기쌍 내에 있는 표적 영역 내의 표적 부위에 결합하는 gRNA(gRNA).

[0861] 188. 구현예 165-187 중 어느 한 구현예에 있어서, 표적 부위는 HBx 기저 코어 프로모터 영역에 위치하는, gRNA.

[0862] 189. 구현예 165-187 중 어느 한 구현예에 있어서, 표적 부위는 HBx 프로모터/인핸서 영역 내에 위치하는, gRNA.

[0863] 190. 구현예 165-189 중 어느 한 구현예에 있어서, 표적 부위는 B형 간염 X 단백질(HBx) 시작 코돈 상류의 250개 염기쌍 내에 걸쳐 있는 표적 영역 내에 있는, gRNA.

[0864] 191. 구현예 165-190 중 어느 한 구현예에 있어서, 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 HBV 게놈의 1060-1480bp 사이의 염기쌍에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내에 있는, gRNA.

[0865] 192. 구현예 165-191 중 어느 한 구현예에 있어서, 표적 부위는 B형 간염 X 단백질(HBx) 시작 코돈 상류의 150개 염기쌍 내에 걸쳐 있는 표적 영역 내에 있는, gRNA.

[0866] 193. 구현예 165-192 중 어느 한 구현예에 있어서, 표적 부위는 B형 간염 X 단백질(HBx) 시작 코돈 상류의 120개 염기쌍 내에 걸쳐 있는 표적 영역 내에 있는, gRNA.

[0867] 194. 구현예 165-193 중 어느 한 구현예에 있어서, 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 HBV 게놈의 1250-1374bp 사이의 염기쌍에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내에 있는, gRNA.

[0868] 195. 구현예 165-194 중 어느 한 구현예에 있어서, 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 HBV 게놈의 1255-1302bp 사이의 염기쌍에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내에 있는, gRNA.

[0869] 196. 구현예 165-195 중 어느 한 구현예에 있어서, 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 HBV 게놈의 1260-1300bp 사이의 염기쌍에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내에 있는, gRNA.

[0870] 197. 구현예 165-196 중 어느 한 구현예에 있어서, gRNA는 서열번호 200, 201, 207, 217, 221, 224, 233, 237, 238, 246, 251, 256, 258, 263, 267, 274, 270, 277, 279, 283, 284, 293, 294, 308, 311, 313, 316, 319, 320, 325, 328, 330, 333, 338, 345, 347, 350, 353, 359, 360, 369, 370, 371, 377, 380, 384, 385 또는 387 중 어느 하나에 제시된 서열, 적어도 14개 뉴클레오타이드의 이의 연속 부분 또는 상기 중 어느 하나의 상보적 서열을 포함하고, 선택적으로 여기서 gRNA 또는 각각의 gRNA는 서열번호 395, 402, 408, 412, 416, 419, 428, 432, 433, 441, 446, 451, 453, 458, 462, 465, 469, 472, 474, 478, 479, 488, 489, 503, 506, 508, 511, 514, 515, 520, 523, 525, 575, 528, 533, 540, 542, 545, 548, 554, 555, 565, 566, 572, 579, 580, 또는 582 중 어느 하나에 제시된 것인, gRNA.

[0871] 198. 구현예 165-196 중 어느 한 구현예에 있어서, gRNA는 서열번호 395, 402, 408, 412, 416, 419, 428, 432, 433, 441, 446, 451, 453, 458, 462, 465, 469, 472, 474, 478, 479, 488, 489, 503, 506, 508, 511, 514, 515, 520, 523, 525, 575, 528, 533, 540, 542, 545, 548, 554, 555, 565, 566, 572, 579, 580, 또는 582 중 어느 하나에 제시된, gRNA.

[0872] 199. 구현예 165-198 중 어느 한 구현예에 있어서, gRNA는 서열번호 207, 213, 215, 217, 221, 222, 241, 245, 258, 261, 268, 274, 380, 387 중 어느 하나에 제시된 서열, 적어도 14개 뉴클레오타이드의 이의 연속 부분, 또는 상기 중 어느 하나의 상보적 서열을 포함하고, 선택적으로 gRNA 또는 각각의 gRNA는 서열번호 402, 408, 410, 412, 416, 417, 436, 440, 453, 456, 463, 469, 575 또는 582 중 어느 하나에 명시된 것인, gRNA.

[0873] 200. 구현예 165-199 중 어느 한 구현예에 있어서, gRNA는 서열번호 402, 408, 410, 412, 416, 417, 436, 440, 453, 456, 463, 469, 575 또는 582 중 어느 하나에 제시된 서열을 포함하는, gRNA.

[0874] 201. 구현예 165-200 중 어느 한 구현예에 있어서, gRNA는 서열번호 217에 제시된 서열, 적어도 14개 뉴클레오타이드의 이의 연속 부분, 또는 상기 중 어느 하나의 상보적 서열을 포함하고, 선택적으로 해당 gRNA 또는 각각의 gRNA는 서열번호 412 중 어느 하나에 제시된 것인, gRNA.

[0875] 202. 구현예 165-201 중 어느 한 구현예에 있어서, gRNA는 서열번호 412에 제시된 서열을 포함하는, gRNA.

[0876] 203. 다음을 포함하는 CRISPR Cas-가이드 RNA(gRNA) 조합:

[0877] (a) 규칙적인 간격을 갖고 나타나는 짧은 회문 구조의 반복 서열 관련(Cas: Clustered Regularly Interspaced Short Palindromic Repeats associated) 단백질 또는 이의 변이체; 및

[0878] (b) Cas 단백질 또는 이의 변이체를 B형 간염 바이러스 DNA 서열의 표적 부위에 표적화하는 구현예 165-202 중 어느 한 구현예의 적어도 하나의 gRNA.

[0879] 204. 구현예 203에 있어서, Cas 단백질 또는 이의 변이체는 Cas9 단백질 또는 이의 변이체인, CRISPR Cas-gRNA 조합.

[0880] 205. 구현예 203 또는 구현예 204에 있어서, Cas 단백질 또는 이의 변이체는 변이체 Cas 단백질이고, 여기서 변이체 Cas 단백질은 뉴클레아제 활성이 없거나 비활성화된 Cas(dCas) 단백질인, CRISPR Cas-gRNA 조합.

[0881] 206. 구현예 203-205 중 어느 한 구현예에 있어서, 변이체 Cas 단백질은 뉴클레아제 활성이 결여된 변이체 Cas9 단백질이거나 비활성화된 Cas9(dCas9) 단백질인, CRISPR Cas-gRNA 조합.

[0882] 207. 구현예 203-206 중 어느 한 구현예에 있어서, Cas9 단백질 또는 이의 변이체는 황색포도상구균(*Staphylococcus aureus*) Cas9(SaCas9) 단백질 또는 이의 변이체인, CRISPR Cas-gRNA 조합.

[0883] 208. 구현예 207에 있어서, 변이체 Cas9는 서열번호 596의 위치 번호를 참조로 D10A 및 N580A에서 선택된 적어도 하나의 아미노산 돌연변이를 포함하는 황색포도상구균(*Staphylococcus aureus*) dCas9 단백질(dSaCas9)인, CRISPR Cas-gRNA 조합.

[0884] 209. 구현예 207 또는 구현예 208에 있어서, 변이체 Cas9 단백질은 서열번호 597에 제시된 서열을 포함하거나, 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는, CRISPR Cas-gRNA 조합.

[0885] 210. 구현예 203-206 중 어느 한 구현예에 있어서, Cas9 단백질 또는 이의 변이체는 화농성연쇄상구균(*Streptococcus pyogenes*) Cas9(SpCas9) 단백질 또는 이의 변이체인, CRISPR Cas-gRNA 조합.

[0886] 211. 구현예 210에 있어서, 변이체 Cas9는 서열번호 598의 위치 번호를 참조하여 D10A 및 H840A 중에서 선택된 적어도 하나의 아미노산 돌연변이를 포함하는 화농성연쇄상구균(*Streptococcus pyogenes*) dCas9(dSpCas9) 단백질인, CRISPR Cas-gRNA 조합.

[0887] 212. 구현예 210 또는 구현예 211에 있어서, 변이체 Cas9 단백질은 서열번호 599에 제시된 서열, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는, CRISPR Cas-gRNA 조합.

[0888] 213. 구현예 1-164 중 어느 한 구현예의 후성적 변형 DNA-표적화 시스템, 또는 구현예 1-164 중 어느 한 구현예의 DNA-표적화 시스템의 융합 단백질, 구현예 165-202 중 어느 한 구현예의 gRNA, 구현예 203-212 중 어느 한 구현예의 CRISPR Cas-gRNA 조합, 또는 상기의 일부 또는 구성 요소를 인코딩하는, 폴리뉴클레오타이드.

[0889] 214. 구현예 1-164 중 어느 한 구현예의 후성적 변형 DNA-표적화 시스템, 또는 구현예 1-164 중 어느 한 구현예의 DNA-표적화 시스템의 융합 단백질, 구현예 165-202 중 어느 한 구현예의 gRNA, 구현예 203-212 중 어느 한 구현예의 CRISPR Cas-gRNA 조합, 또는 상기의 일부 또는 구성 요소를 인코딩하는, 복수의 폴리뉴클레오타이드.

[0890] 215. 구현예 213의 폴리뉴클레오타이드를 포함하는 벡터.

[0891] 216. 구현예 214의 복수의 폴리뉴클레오타이드를 포함하는 벡터.

[0892] 217. 구현예 215 또는 216에 있어서, 벡터는 바이러스 벡터인, 벡터.

[0893] 218. 구현예 217에 있어서, 벡터는 아데노 연관 바이러스(AAV) 벡터인, 벡터.

[0894] 219. 구현예 218에 있어서, 벡터는 AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8 또는 AAV9 중에서 선택되는, 벡터.

[0895] 220. 구현예 219에 있어서, 벡터는 렌티바이러스 벡터인, 벡터.

[0896] 221. 구현예 215 또는 구현예 216에 있어서, 벡터는 비바이러스 벡터인, 벡터.

[0897] 222. 구현예 221에 있어서, 비바이러스 벡터는 지질 나노입자, 리포좀, 엑소좀 또는 세포 침투 펩타이드 중에서 선택되는, 벡터.

[0898] 223. 구현예 215-222 중 어느 어느 한 구현예에 있어서, 벡터는 B형 간염 바이러스에 감염된 세포에 대해 향성을 나타내며, 선택적으로 감염된 세포는 간세포인, 벡터.

[0899] 224. 구현예 215-223 중 어느 한 구현예에 있어서, 벡터는 하나의 벡터, 또는 두 개 이상의 벡터를 포함하는, 벡터.

[0900] 225. B형 간염 바이러스 서열의 표적 영역 내에서 후성적 변형을 촉진하는 방법으로서, 상기 방법은 B형 간염 바이러스 서열을 포함하는 HBV 감염 세포 내로 표적 영역 내의 표적 부위를 표적으로 하는 후성적 변형 DNA-표적화 시스템을 도입하는 단계를 포함하는, 방법.

[0901] 226. B형 간염 바이러스 서열 내 표적 영역 내 CpG 메틸화를 증가시키는 방법으로서, 상기 방법은 B형 간염 바이러스 서열을 포함하는 HBV 감염 세포 내로 표적 영역 내 표적 부위를 표적으로 하는 후성적 변형 DNA-표적화 시스템을 도입하는 단계를 포함하는, 방법.

[0902] 227. B형 간염 바이러스 서열의 표적 영역의 후성적 변형을 촉진하는 방법으로서, 상기 방법은 B형 간염 바이러스 서열을 포함하는 HBV 감염 세포 내로 구현예 1-164 중 어느 한 구현예의 후성적 변형 DNA-표적 시스템, 구현예 165-202 중 어느 한 구현예의 gRNA, 구현예 203-212 중 어느 한 구현예의 CRISPR Cas-gRNA 조합, 구현예 213의 폴리뉴클레오타이드, 구현예 214의 복수의 폴리뉴클레오타이드, 구현예 215-224 중 어느 한 구현예의 벡터, 또는 상기의 일부 또는 구성요소를 도입하는 것을 포함하는, 방법.

[0903] 228. B형 간염 바이러스 서열의 표적 영역의 CpG 메틸화를 증가시키는 방법으로서, B형 간염 바이러스 서열을 포함하는 HBV 감염 세포에 구현예 1-164 중 어느 하나의 후성적 변형 DNA-표적화 시스템, 구현예 165-202 중 어느 하나의 gRNA, 구현예 203-212 중 어느 하나의 CRISPR Cas-gRNA 조합, 구현예 213의 폴리뉴클레오타이드, 구현예 214의 복수의 폴리뉴클레오타이드, 구현예 215-224 중 어느 하나의 벡터, 또는 상기의 일부 또는 구성요소를 도입하는 단계를 포함하는, 방법.

[0904] 229. 구현예 225-228 중 어느 한 구현예에 있어서, 표적 영역은 서열번호 650의 뉴클레오타이드 위치를 참조로 B형 간염 바이러스 서열의 1033bp-1749bp 사이의 염기쌍에 위치한 서열에 상응하는 서열 내의 연속된 뉴클레오타이드 서열을 포함하는, 방법.

[0905] 230. B형 간염 바이러스 서열을 포함하는 HBV에 감염된 세포에서 하나 이상의 유전자 전사를 감소시키는 방법으로서, 상기 방법은 서열번호 650의 뉴클레오타이드 위치를 참조로 B형 간염 바이러스 서열 내에서 표적 CpG 메틸화를 유도하는 후성적 변형 DNA-표적화 시스템을 세포에 도입하는 것을 포함하는, 방법.

[0906] 231. HBV에 감염된 세포에서 B형 간염 바이러스 감염을 감소시키는 방법으로서, B형 간염 바이러스 서열을 포함하는 세포 내로 서열번호 650의 뉴클레오타이드 위치를 참조로 B형 간염 바이러스 서열의 표적 영역 내에서 표적 CpG 메틸화를 유도하는 후성적 변형 DNA-표적화 시스템을 도입하는 단계를 포함하는, 방법.

[0907] 232. 구현예 225-231 중 어느 한 구현예에 있어서, 후성적 변형 DNA-표적화 시스템은 (a) B형 간염 바이러스 DNA 서열의 표적 부위를 표적화하기 위한 DNA-결합 도메인 및 (b) DNA 메틸트랜스퍼라제 효과기 도메인을 포함하는 적어도 하나의 효과기 도메인을 포함하는 융합 단백질을 포함하는 적어도 하나의 DNA-표적화 모듈을 포함하는, 방법.

[0908] 233. 구현예 226 및 228-232 중 어느 한 구현예에 있어서, CpG 메틸화 영역은 표적 영역의 500개 염기쌍 이내에 있는, 방법.

[0909] 234. 구현예 225-233 중 어느 한 구현예에 있어서, 도입은 대상의 생체 내 또는 생체 외에서 이루어지는, 방법.

[0910] 235. 구현예 225-234 중 어느 한 구현예에 있어서, 세포는 포유류 세포인, 방법.

[0911] 236. 구현예 225-235 중 어느 한 구현예에 있어서, 세포는 인간 세포인, 방법.

[0912] 237. 구현예 225-236 중 어느 한 구현예에 있어서, 세포는 통합된 HBV DNA를 포함하는, 방법.

[0913] 238. 구현예 225-237 중 어느 한 구현예에 있어서, 세포는 에피솜 HBV cccDNA 풀을 포함하는 간세포인, 방법.

[0914] 239. 구현예 238에 있어서, 간세포는 HBV 단백질을 발현하고, 여기서 HBV 단백질은 HBsAg, HBeAg 또는 HBcrAg이거나, 또는 상기의 조합인, 방법.

[0915] 240. 대상에서 B형 간염 바이러스 감염을 감소시키는 방법으로서, 상기 방법은 B형 간염 바이러스에 감염된 대상에게 B형 간염 바이러스 서열의 표적 영역 내에서 CpG 메틸화를 증가시키는 후성적 변형 DNA-표적화 시스템을 투여하는 단계를 포함하고, 여기서 후성적 변형 DNA-표적화 시스템은 (a) B형 간염 바이러스 DNA 서열의 표적 부위를 표적으로 삼기 위한 DNA-결합 도메인; 및 (b) DNA 메틸트랜스퍼라제 효과기 도메인을 포함하는 적어도

하나의 효과기 도메인을 포함하는 것인, 방법.

[0916] 241. 구현예 225-240 중 어느 한 구현예에 있어서, 표적 영역은 HBV 게놈의 CpG를 포함하는 영역인, 방법.

[0917] 242. 구현예 225-241 중 어느 한 구현예에 있어서, 표적 영역은 서열번호 650의 뉴클레오타이드 위치를 참조로 B형 간염 바이러스 서열에서 67bp-392bp, 1033bp-1749bp 또는 2215bp-2490bp 사이의 염기쌍에 위치한 서열에 상응하는 뉴클레오타이드의 연속 서열을 포함하는, 방법.

[0918] 243. 구현예 225-242 중 어느 한 구현예에 있어서, 표적 영역은 서열번호 650의 뉴클레오타이드 위치를 참조로 B형 간염 바이러스 서열의 1033bp-1749bp 사이의 염기쌍에 위치한 서열에 상응하는 뉴클레오타이드의 연속된 서열을 포함하는, 방법.

[0919] 244. 구현예 225-244 중 어느 한 구현예에 있어서, 표적 영역은 B형 간염 X 단백질(HBx) 시작 코돈 상류의 300개 염기쌍 이내에 위치하는, 방법.

[0920] 245. 구현예 225-244 중 어느 한 구현예에 있어서, 표적 영역은 HBx 기저 코어 프로모터 영역 내에 있는, 방법.

[0921] 246. 구현예 225-245 중 어느 한 구현예에 있어서, 표적 영역은 HBx 프로모터/인핸서 영역 내에 있는, 방법.

[0922] 247. 구현예 225-246 중 어느 한 구현예에 있어서, 표적 영역은 B형 간염 X 단백질(HBx) 시작 코돈 상류의 250개 염기쌍 이내에 있는, 방법.

[0923] 248. 구현예 225-247 중 어느 한 구현예에 있어서, 표적 영역은 서열번호 650의 뉴클레오타이드 위치를 참조로 B형 간염 바이러스 서열의 1060bp-1480bp 염기쌍 사이에 위치한 서열에 상응하는 뉴클레오타이드의 연속 서열을 포함하는, 방법.

[0924] 249. 구현예 225-248 중 어느 한 구현예에 있어서, 표적 영역은 B형 간염 X 단백질(HBx) 시작 코돈 상류의 150개 염기쌍 이내에 있는, 방법.

[0925] 250. 구현예 225-249 중 어느 한 구현예에 있어서, 표적 영역은 BX형 간염 단백질(HBx) 시작 코돈 상류의 120개 염기쌍 이내에 있는, 방법.

[0926] 251. 구현예 225-250 중 어느 한 구현예에 있어서, 표적 영역은 서열번호 650의 뉴클레오타이드 위치를 참조로 B형 간염 바이러스 서열의 염기쌍 1250bp-1374bp 사이에 위치한 서열에 상응하는 뉴클레오타이드의 연속된 서열을 포함하는, 방법.

[0927] 252. 구현예 225-251 중 어느 한 구현예에 있어서, 표적 영역은 서열번호 650의 뉴클레오타이드 위치를 참조로 B형 간염 바이러스 서열의 염기쌍 1260bp-1300bp 사이에 위치한 서열에 상응하는 뉴클레오타이드의 연속 서열을 포함하는, 방법.

[0928] 253. 구현예 225-252 중 어느 한 구현예에 있어서, DNA-결합 도메인은 (a) 규칙적인 간격을 갖고 나타나는 짧은 회문 구조의 반복 서열 관련(Cas) 단백질 또는 이의 변이체 및 (b) 적어도 하나의 gRNA를 포함하는 Cas-가이드 RNA(gRNA) 조합; 아연 집게 단백질(ZFP); 전사 활성화 인자 유사 효과기(TALE); 메가뉴클레아제; 호밍 엔도뉴클레아제; 또는 I-SceI 효소 또는 이의 변이체를 포함하고, 선택적으로 DNA-결합 도메인은 상기 중 임의의 것의 촉매적으로 비활성인 변이체를 포함하는, 방법.

[0929] 254. 구현예 225-254 중 어느 한 구현예에 있어서, CRISPR Cas-가이드 RNA(gRNA) 조합은:

[0930] (a) 규칙적인 간격을 갖고 나타나는 짧은 회문 구조의 반복 서열 관련(Cas) 단백질 또는 이의 변이체; 및

[0931] (b) B형 간염 바이러스 DNA 서열의 표적 부위에 Cas 단백질 또는 이의 변이체를 표적화하는 구현예 165-202 중 어느 한 구현예의 적어도 하나의 gRNA

[0932] 를 포함하는, 방법.

[0933] 255. 구현예 225-254 중 어느 한 구현예에 있어서, 표적 부위 또는 각각의 표적 부위는 서열번호 5, 6, 12, 18, 22, 26, 29, 38, 42, 43, 51, 56, 61, 63, 68, 72, 75, 79, 82, 84, 88, 89, 98, 99, 113, 116, 121, 124, 125, 118, 130, 133, 135, 138, 143, 150, 152, 155, 158, 164, 165, 175, 176, 182, 185, 189, 190, 192 중 어느 하나에 제시된 서열, 적어도 14개 뉴클레오타이드(nt)의 이의 연속 부분 또는 상기 중 어느 하나의 상보적 서열을 포함하는, 방법.

[0934] 256. 구현예 225-255 중 어느 한 구현예에 있어서, 표적 부위 또는 각각의 표적 부위는 서열번호 12, 18, 20,

22, 26, 27, 46, 50, 63, 66, 73, 79, 185, 192 중 어느 하나에 제시된 서열, 적어도 14개 뉴클레오타이드(nt)의 이의 연속 부분, 또는 상기 중 어느 하나의 상보적 서열을 포함하는, 방법.

[0935] 257. 구현예 225-256 중 어느 한 구현예에 있어서, 표적 부위 또는 각각의 표적 부위는 서열번호 22에 제시된 서열, 적어도 14개 뉴클레오타이드(nt)의 이의 연속 부분, 또는 앞서 언급한 서열의 상보적 서열을 포함하는, 방법.

[0936] 258. 구현예 225-257 중 어느 한 구현예에 있어서, gRNA 또는 각각의 gRNA는 서열번호 200, 201, 207, 217, 221, 224, 233, 237, 238, 246, 251, 256, 258, 263, 267, 274, 270, 277, 279, 283, 284, 293, 294, 308, 311, 313, 316, 319, 320, 325, 328, 330, 333, 338, 345, 347, 350, 353, 359, 360, 370, 371, 377, 380, 384, 385, 387 중 어느 하나에 제시된 서열을, 적어도 14개 뉴클레오타이드의 이의 연속 부분 또는 상기 중 어느 하나의 상보적 서열을 포함하고, 선택적으로 gRNA 또는 각각의 gRNA는 서열번호 395, 402, 408, 412, 416, 419, 428, 432, 433, 441, 446, 451, 453, 458, 462, 465, 469, 472, 474, 478, 479, 488, 489, 503, 506, 508, 511, 514, 515, 520, 523, 525, 575, 528, 533, 540, 542, 545, 548, 554, 555, 565, 566, 572, 579, 580, 또는 582 중 어느 하나에 제시된 것인, 방법.

[0937] 259. 구현예 225-258 중 어느 한 구현예에 있어서, gRNA 또는 각각의 gRNA는 서열번호 207, 213, 215, 217, 221, 222, 241, 245, 258, 261, 268, 274, 380, 387 중 어느 하나에 제시된 서열, 적어도 14개 뉴클레오타이드의 이의 연속 부분, 또는 상기 중 어느 하나의 상보적 서열을 포함하고, 선택적으로 gRNA 또는 각각의 gRNA는 서열번호 402, 408, 410, 412, 416, 417, 436, 440, 453, 456, 463, 469, 575, 582 중 어느 하나에 제시된 것인, 방법.

[0938] 260. 구현예 225-259 중 어느 한 구현예에 있어서, gRNA 또는 각각의 gRNA는 서열번호 217 중 어느 하나에 제시된 서열, 적어도 14개 뉴클레오타이드의 이의 연속 부분, 또는 상기 중 어느 하나의 상보적 서열을 포함하고, 선택적으로 gRNA 또는 각각의 gRNA는 서열번호 412 중 어느 하나에 제시된 것인, 방법.

[0939] 261. 구현예 225-253 중 어느 한 구현예에 있어서, 적어도 하나의 DNA-결합 도메인은 조작된 아연 집게 단백질(eZFP)을 포함하는, 방법.

[0940] 262. 구현예 225-253 또는 261 중 어느 한 구현예에 있어서, 표적 부위는 서열번호 1045, 1046, 1052 중 어느 하나에 제시된 뉴클레오타이드 서열, 최소 12nt의 이의 연속 부분, 또는 앞서 언급한 것의 상보적 서열을 포함하는, 방법.

[0941] 263. 구현예 225-253, 또는 261-262 중 어느 한 구현예에 있어서, 표적 부위는 서열번호 1045, 1046, 1052 중 어느 하나에 제시된 뉴클레오타이드 서열을 포함하는, 방법.

[0942] 264. 구현예 225-263 중 어느 한 구현예에 있어서, 적어도 하나의 효과기 도메인은 DNA 메틸트랜스퍼라제인, 방법.

[0943] 265. 구현예 225-264 중 어느 한 구현예에 있어서, 적어도 하나의 효과기 도메인은 DNA 메틸트랜스퍼라제와 헤테로크로마틴 유도 인자를 동원할 수 있는 억제자 도메인을 포함하거나, 선택적으로 헤테로크로마틴 유도 인자는 히스톤 메틸트랜스퍼라제를 포함하는, 방법.

[0944] 266. 구현예 225-265 중 어느 한 구현예에 있어서, 적어도 하나의 효과기 도메인은 DNA 메틸트랜스퍼라제와 히스톤 메틸트랜스퍼라제를 포함하는, 방법.

[0945] 267. 구현예 225-266 중 어느 한 구현예에 있어서, 적어도 하나의 효과기 도메인은 DNMT3A/L 도메인 또는 그 변이체를 포함하는, 방법.

[0946] 268. 구현예 225-267 중 어느 한 구현예에 있어서, 적어도 하나의 효과기 도메인은 서열번호 651에 제시된 서열, 이의 일부, 또는 상기 중 임의의 것과 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는, 방법.

[0947] 269. 구현예 225-268 중 어느 한 구현예에 있어서, 적어도 하나의 효과기 도메인은 KRAB 도메인 또는 이의 변이체를 더 포함하는, 방법.

[0948] 270. 구현예 225-269 중 어느 한 구현예에 있어서, 적어도 하나의 효과기 도메인은 서열번호 590에 제시된 서열, 이의 일부, 또는 상기 중 임의의 것과 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 더 포함하는, 방법.

[0949] 271. 구현예 225-270 중 어느 한 구현예에 있어서, DNA-표적화 시스템은 DNMT3A/3L-dCas9-KRAB 도메인 또는 그 변이체를 포함하는, 방법.

[0950] 272. 구현예 225-271 중 어느 한 구현예에 있어서, DNA-표적화 시스템은 서열번호 645에 제시된 서열, 이의 일부 또는 상기 중 어느 것과 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는, 방법.

[0951] 273. 구현예 225-260 또는 264-272 중 어느 한 구현예에 있어서, DNA-표적화 시스템은 서열번호 680에 제시된 서열, 이의 일부 또는 상기 중 어느 것과 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 핵산 서열을 포함하는, 방법.

[0952] 274. 구현예 225-260 또는 264-273 중 어느 한 구현예에 있어서, DNA-표적화 시스템은 서열번호 680에 제시된 서열을 포함하는, 방법.

[0953] 275. B형 간염 바이러스에 감염된 세포에서 하나 이상의 유전자의 전사를 억제하는 방법으로서, 상기 방법은 B형 간염 바이러스에 감염된 세포에 구현예 1-164 중 어느 한 구현예의 후성적 변형 DNA-표적화 시스템, 구현예 165-202 중 어느 하나의 gRNA, 구현예 203-212 중 어느 하나의 CRISPR Cas-gRNA 조합, 구현예 213의 폴리뉴클레오타이드, 구현예 214의 복수의 폴리뉴클레오타이드, 구현예 215-224 중 어느 하나의 벡터, 또는 상기의 일부 또는 구성요소를 도입하는 단계를 포함하는, 방법.

[0954] 276. 구현예 225-275 중 어느 한 구현예에 있어서, 하나 이상의 유전자는 DNA-표적화 시스템에 의해 후성적으로 변형되는. 방법.

[0955] 277. 구현예 225-275 중 어느 한 구현예에 있어서, 하나 이상의 유전자의 전사는 이 방법을 거치지 않은 비교 가능한 세포와 비교시 감소되는, 방법.

[0956] 278. 구현예 277에 있어서, 하나 이상의 유전자의 전사는 적어도 약 1.25배, 1.5배, 1.75배, 2.0배, 2.5배, 2.75배, 3.0배, 3.5배, 3.75배, 4.0배, 4.5배, 4.75배, 5.0배, 5.25배, 5.5배, 5.75배, 6배만큼 감소되는, 방법.

[0957] 279. 구현예 225-278 중 어느 한 구현예에 있어서, 하나 이상의 유전자의 전사의 억제는 HBV 복제 및/또는 HBV 전사의 감소를 초래하는, 방법.

[0958] 280. 구현예 225-279 중 어느 한 구현예에 있어서, HBV에 감염된 세포는 포유류 세포인, 방법.

[0959] 281. 구현예 225-280 중 어느 한 구현예에 있어서, HBV에 감염된 세포는 인간 세포인, 방법.

[0960] 282. 구현예 225-281 중 어느 한 구현예에 있어서, 세포는 통합된 HBV DNA를 포함하는, 방법.

[0961] 283. 구현예 225-282 중 어느 한 구현예에 있어서, 세포는 에피솜 HBV cccDNA 풀을 포함하는 간세포인, 방법.

[0962] 284. 구현예 225-283 중 어느 한 구현예에 있어서, 간세포는 HBV 단백질을 발현하고, 여기서 HBV 단백질은 HBsAg 및/또는 HBeAg인, 방법.

[0963] 285. 구현예 225-284 중 어느 한 구현예에 있어서, HBV에 감염된 세포는 대상에 존재하는, 방법.

[0964] 286. 구현예 225-285 중 어느 한 구현예에 있어서, 대상은 인간인, 방법.

[0965] 287. 구현예 225-286 중 어느 한 구현예에 있어서, 대상은 HBV 바이러스에 감염된, 방법.

[0966] 288. 구현예 225-288 중 어느 한 구현예에 있어서, 대상은 통합된 HBV DNA를 포함하는 간세포를 갖는, 방법.

[0967] 289. 구현예 225-288 중 어느 한 구현예에 있어서, 대상은 에피솜 HBV cccDNA 풀을 포함하는 간세포를 갖는, 방법.

[0968] 290. 구현예 225-289 중 어느 한 구현예에 있어서, 대상은 HBV 단백질을 발현하는 간세포를 갖고, 여기서 HBV 단백질은 HBsAg, HBeAg 또는 HBcrAg 및 이들의 조합인, 방법.

[0969] 291. 구현예 225-290 중 어느 한 구현예에 있어서, 대상은 HBV 바이러스 감염과 관련된 질병, 병태 또는 장애를 갖는, 방법.

[0970] 292. 구현예 291에 있어서, 질병, 병태 또는 장애는 간 질환 또는 암인, 방법.

[0971] 293. 구현예 291 또는 구현예 292에 있어서, 질병, 병태 또는 장애는 급성 간염, 만성 간염, 간부전 또는 간경변인, 방법.

[0972] 294. 구현예 291-293 중 어느 한 구현예에 있어서, 질병, 병태 또는 장애는 암이고, 선택적으로 암은 간세포암인, 방법.

[0973] 295. 구현예 215-224 중 어느 한 구현예의 벡터를 포함하는 약학 조성물.

[0974] 296. 구현예 295에 있어서, 벡터는 갈락토스의 아미노 당 유도체에 접합되고, 선택적으로 벡터는 N-아세틸갈락토사민(GalNAc) 부분에 접합된 것인, 약학 조성물.

[0975] 297. 구현예 1-164 중 어느 하나의 후성적 변형 DNA-표적화 시스템 또는 구현예 1-164 중 어느 하나의 융합 단백질, 구현예 165-202 중 어느 하나의 gRNA, 구현예 203-212 중 어느 하나의 CRISPR Cas-gRNA 조합, 구현예 213의 폴리뉴클레오타이드, 구현예 214의 복수의 폴리뉴클레오타이드, 구현예 215-224 중 어느 하나의 벡터, 또는 상기의 일부 또는 구성요소를 포함하는 약학 조성물.

[0976] 298. 구현예 295-297 중 어느 한 구현예에 있어서, 대상의 HBV 바이러스 감염을 치료하는 데 사용하기 위한, 약학 조성물.

[0977] 299. 구현예 295-298 중 어느 한 구현예에 있어서, 대상은 HBV 바이러스 감염과 관련된 질병, 병태 또는 장애를 갖는, 약학 조성물.

[0978] 300. 구현예 295-299 중 어느 한 구현예에 있어서, 대상에서 HBV 바이러스 감염과 관련된 질병, 장애 또는 상태를 치료하는 데 사용하기 위한, 약학 조성물.

[0979] 301. 구현예 295-300 중 어느 한 구현예에 있어서, 대상에서 HBV 바이러스 감염을 치료하기 위한 약물 제조에 사용하기 위한, 약학 조성물.

[0980] 302. 구현예 298-301 중 어느 한 구현예에 있어서, HBV 바이러스 감염은 질병, 장애 또는 병태와 연관된, 약학 조성물.

[0981] 303. 구현예 295-300 중 어느 한 구현예에 있어서, HBV 바이러스 감염과 관련된 대상에서 질병, 병태 또는 장애를 치료하기 위한 약물 제조에 사용하기 위한, 약학 조성물.

[0982] 304. 구현예 299, 300, 302 또는 303 중 어느 한 구현예에 있어서, 질병, 병태 또는 장애는 간 질환 또는 암인, 약학 조성물.

[0983] 305. 구현예 299, 300, 302 또는 303 중 어느 한 구현예에 있어서, 질병, 병태 또는 장애는 급성 간염, 만성 간염, 간부전 또는 간경변인, 약학 조성물.

[0984] 306. 구현예 299, 300, 302 303 또는 304 중 어느 한 구현예에 있어서, 질병, 병태 또는 장애는 암, 선택적으로 간세포암인, 약학 조성물.

[0985] 307. 구현예 295-306 중 어느 한 구현예에 있어서, 약학 조성물은 대상에게 생체 내로 투여되는, 약학 조성물.

[0986] 308. 구현예 295-307 중 어느 한 구현예에 있어서, 약학 조성물의 투여 후, 대상의 세포에서 하나 이상의 HBV 유전자의 전사가 억제되는, 약학 조성물.

[0987] 309. 구현예 308에 있어서, 하나 이상의 HBV 유전자는 HBV 복제 및/또는 HBV 전사를 제어하는 데 관여하는, 약학 조성물.

[0988] 310. 구현예 308 또는 309에 있어서, 하나 이상의 유전자는 폴리머라제 유전자, S-패밀리 유전자, X-유전자 또는 코어 패밀리 유전자인, 약학 조성물.

[0989] 311. 질병, 병태 또는 장애의 치료를 필요로 하는 대상에서 이를 치료하기 위한 방법으로서, 구현예 1-164 중 어느 하나의 후성적 변형 DNA-표적화 시스템, 구현예 165-202 중 어느 하나의 gRNA, 구현예 203-212 중 어느 하나의 CRISPR Cas-gRNA 조합, 구현예 213의 폴리뉴클레오타이드, 구현예 214의 복수의 폴리뉴클레오타이드, 구현예 215-224 중 어느 하나의 벡터, 구현예 297-310 중 어느 하나의 약학 조성물, 또는 상기의 일부 또는 구성요소를 대상에게 투여하는 단계를 포함하는, 방법.

[0990] 312. 대상에서 B형 간염 바이러스 감염을 감소시키는 방법으로서, B형 간염 바이러스 감염이 있는 대상에게 구현예 1-164 중 어느 하나의 후성적 변형 DNA-표적화 시스템, 구현예 165-202 중 어느 하나의 gRNA, 구현예 203-

212 중 어느 하나의 CRISPR Cas-gRNA 조합, 구현예 213의 폴리뉴클레오타이드, 구현예 214의 복수의 폴리뉴클레오타이드, 구현예 215-224 중 어느 하나의 벡터, 구현예 297-310 중 어느 하나의 약학 조성물, 또는 상기의 일부 또는 구성요소를 투여하는 단계를 포함하는, 방법

[0991] 313. 하나 이상의 HBV 유전자 또는 이의 조절 요소의 표적 부위에 결합하는 조작된 아연 집게 단백질(eZFP)로서, 여기서 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 HBV 게놈의 1033bp-1749bp 사이의 염기쌍에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내에 있는, 조작된 아연 집게 단백질(eZFP).

[0992] 314. 구현예 313에 있어서, 표적 부위는 B형 간염 X 단백질(HBx) 시작 코돈 상류의 300개 염기쌍 걸쳐 있는 표적 영역 내에 있는, eZFP.

[0993] 315. 구현예 313 또는 구현예 314에 있어서, 표적 부위는 HBx 기저 코어 프로모터 영역에 위치하는, eZFP.

[0994] 316. 구현예 313 또는 구현예 314에 있어서, 표적 부위는 HBx 프로모터/인핸서 영역 내에 위치하는, eZFP.

[0995] 317. 구현예 313-316 중 어느 한 구현예에 있어서, 표적 부위는 B형 간염 X 단백질(HBx) 시작 코돈 상류의 250개 염기쌍 내에 걸쳐 있는 표적 영역 내에 있는, eZFP.

[0996] 318. 구현예 313-317 중 어느 한 구현예에 있어서, 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 HBV 게놈의 1060-1480bp 사이의 염기쌍에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내에 있는, eZFP.

[0997] 319. 구현예 313-318 중 어느 한 구현예에 있어서, 표적 부위는 형 간염 X 단백질(HBx) 시작 코돈 상류의 150개 염기쌍 내에 걸쳐 있는 표적 영역 내에 있는, eZFP.

[0998] 320. 구현예 313-319 중 어느 한 구현예에 있어서, 표적 부위는 HBx(간염 BX) 단백질 시작 코돈 상류의 120개 염기쌍 내에 걸쳐 있는 표적 영역 내에 있는, eZFP.

[0999] 321. 구현예 313-320 중 어느 한 구현예에 있어서, 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 HBV 게놈의 1250-1374bp 사이의 염기쌍에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내에 있는, eZFP.

[1000] 322. 구현예 313-321 중 어느 한 구현예에 있어서, 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 HBV 게놈의 1255-1302bp 사이의 염기쌍에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내에 있는, eZFP.

[1001] 323. 구현예 313-322 중 어느 한 구현예에 있어서, 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 HBV 게놈의 1260-1300bp 사이의 염기쌍에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내에 있는, eZFP.

[1002] 324. 구현예 323에 있어서, 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 HBV 게놈의 1255bp-1290bp 사이의 염기쌍에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내에 있는, eZFP.

[1003] 325. 구현예 313-324 중 어느 한 구현예에 있어서, 표적 부위는 서열번호 1028-1055 중 어느 하나에 제시된 뉴클레오타이드 서열, 최소 12nt의 이의 연속 부분, 또는 앞서 언급한 것의 상보적 서열을 포함하는, eZFP.

[1004] 326. 구현예 313-325 중 어느 한 구현예에 있어서, 표적 부위는 서열번호 1028-1055 중 하나에 명시된 뉴클레오타이드 서열을 포함하는, eZFP.

[1005] 327. 구현예 313-326 중 어느 한 구현예에 있어서, 표적 부위는 서열번호 1045, 1046 또는 1052 중 어느 하나에 명시된 뉴클레오타이드 서열, 최소 12nt의 이의 연속 부분, 또는 앞서 언급한 것의 상보적 서열을 포함하는, eZFP.

[1006] 328. 구현예 313-326 중 어느 한 구현예에 있어서, 표적 부위는 서열번호 1045, 1046 또는 1052 중 하나에 제시된 뉴클레오타이드 서열을 포함하는, eZFP.

[1007] 329. 구현예 313-328 중 어느 한 구현예에 있어서, 아연 집게 단백질은 N-말단에서 C-말단 순서로 F1 내지 F6으로 표시되는 6개의 아연 집게를 포함하고,

[1008] 여기서 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같은 것인, eZFP:

[1009] 1) F1:SEADRSR(서열번호 720) F2:DRSNLTR(서열번호 721)F3:QSSDLSR(서열번호 722) F4:YHWYLKK(서열번호 723) F5:RSDSLSV(서열번호 724) F6:QNAANRKT(서열번호 725);

[1010] 2) F1:RSDVLST(서열번호 726) F2:DNSSRTR(서열번호 727) F3:RPYTLRL(서열번호 728) F4:DSSHRTR(서열번호 729) F5:RSDHLSQ(서열번호 730) F6:DSSHRTR(서열번호 731);

[1011] 3) F1:RSDHLSQ(서열번호 732) F2:QSADRTK(서열번호 733) F3:RSDHLSQ(서열번호 734) F4:RRSDLKR(서열번호 735) F5:RSDHLSR(서열번호 736) F6:QSSDLRR(서열번호 737);

[1012] 4) F1:RSDNLSE(서열번호 738) F2:TSSNRKT(서열번호 739) F3:DRSHLTR(서열번호 740) F4:RSDALTQ(서열번호 741) F5:DRSALAR(서열번호 742) F6:RRFTLSK(서열번호 743);

[1013] 5) F1:RSDHLSE(서열번호 744) F2:QYSGRYY(서열번호 745) F3:HGQTLNE(서열번호 746) F4:QSGNLAR(서열번호 747) F5:RSDSLLR(서열번호 748)F6:CREYRGK(서열번호 749);

[1014] 6) F1:QSANRTT(서열번호 750) F2:RSANLTR(서열번호 751) F3:RSDVLSE(서열번호 752) F4:TSGHLSR(서열번호 753) F5:QSSDLSR(서열번호 754), F6:QWSTRKR(서열번호 755);

[1015] 7) F1:QSGNLAR(서열번호 756) F2:ATCCLAH(서열번호 757) F3:RWQYLPT(서열번호 758) F4:DRSALAR(서열번호 759) F5:RSDNLSE(서열번호 760)F6:KRCNLRC(서열번호 761);

[1016] 8) F1:NPANLTR(서열번호 762) F2:QNATRTK(서열번호 763) F3:QSGHLAR(서열번호 764) F4:NRHDRAK(서열번호 765) F5:RSDHLSE(서열번호 766), F6:QRRSRYK(서열번호 767);

[1017] 9) F1:QSSDLSR(서열번호 768) F2:HRSTRNR(서열번호 769) F3:RSDVLSA(서열번호 770) F4:DSRTRKN(서열번호 771) F5:QSGSLTR(서열번호 772) F6:DQSGLAH(서열번호 773);

[1018] 10) F1:QNPAQWR(서열번호 774) F2:RSADLSR(서열번호 775) F3:TSGSLSR(서열번호 776) F4:RSDHLSR(서열번호 777) F5:RSDSLLR(서열번호 778) F6:QSYDRFQ(서열번호 779);

[1019] 11) F1:TSGSLSR(서열번호 780) F2:RSDHLSR(서열번호 781) F3:RSDSLLR(서열번호 782) F4:QSYDRFQ(서열번호 783) F5:RSDNLST(서열번호 784) F6:DNRDRIK(서열번호 785);

[1020] 12) F1:DRSNLSR(서열번호 786) F2:LRQNLIM(서열번호 787) F3:ERGTLAR(서열번호 788) F4:RSDALTQ(서열번호 789) F5:RSDSLSQ(서열번호 790) F6:RKADRTR(서열번호 791);

[1021] 13) F1:QYCCLTN(서열번호 792) F2:TSGNLTR(서열번호 793) F3:QSSDLSR(서열번호 794) F4:FRYYLKR(서열번호 795) F5:QSGDLTR(서열번호 796) F6:DKGNLTK(서열번호 797);

[1022] 14) F1:TSGSLSR(서열번호 798) F2:RSDNLTT(서열번호 799) F3:QSGNLAR(서열번호 800) F4:DRTTLMR(서열번호 801) F5:QSGHLAR(서열번호 802) F6:QLTHLNS(서열번호 803);

[1023] 15) F1:IKHDLHR(서열번호 804) F2:RSANLTR(서열번호 805) F3:RSDNLAR(서열번호 806) F4:QNVSRPR(서열번호 807) F5:RSDDLSK(서열번호 808) F6:DSSHRTR(서열번호 809);

[1024] 16) F1:RSDNLAR(서열번호 810) F2:QNVSRPR(서열번호 811) F3:RSDDLSK(서열번호 812) F4:DSSHRTR(서열번호 813) F5:TSSNRKT(서열번호 814) F6:AQWTRAC(서열번호 815);

[1025] 17) F1:RSDDLSK(서열번호 816) F2:DSSHRTR(서열번호 817) F3:TSSNRKT(서열번호 818) F4:AQWTRAC(서열번호 819) F5:RKQTRTT(서열번호 820) F6:HRSSLRR(서열번호 821);

[1026] 18) F1:QSAHRKN(서열번호 822) F2:TSSNRKT(서열번호 823) F3:RSDNLSA(서열번호 824) F4:RNNDRKT(서열번호 825) F5:TSGSLSR(서열번호 826) F6:QAGHLAK(서열번호 827);

[1027] 19) F1:RSDHLSQ(서열번호 828) F2:ASSTRTK(서열번호 829) F3:RSDDLTR(서열번호 830) F4:QKSNLSS(서열번호 831) F5:QSANRTT(서열번호 832) F6:QNATRTK(서열번호 833);

[1028] 20) F1:RSDTLSE(서열번호 834) F2:RRWTLVG(서열번호 835) F3:DRSNLSR(서열번호 836) F4:QSGDLTR(서열번호 837) F5:QSSDLSR(서열번호 838) F6:YHWYLKK(서열번호 839);

[1029] 21) F1:RSANLAR(서열번호 840) F2:RSDNLRE(서열번호 841) F3:RPYTLRL(서열번호 842) F4:HRSNLNK(서열번호 843) F5:QSGSLTR(서열번호 844) F6:TSANLSR(서열번호 845);

[1030] 22) F1:RSDDLVR(서열번호 846) F2:TSGSLVR(서열번호 847) F3:RSDKLVR(서열번호 848) F4:RSDELVR(서열번호 849) F5:TSHSLTE(서열번호 850) F6:RADNLTE(서열번호 851);

[1031] 23) F1:ERSHLRE(서열번호 852) F2:TSHSLTE(서열번호 853) F3:QAGHLAS(서열번호 854) F4:TSHSLTE(서열번호 855) F5:DPGHLVR(서열번호 856) F6:TSGNLVR(서열번호 857);

[1032] 24) F1:RADNLTE(서열번호 858) F2:TSGSLVR(서열번호 859) F3:RKDNLKN(서열번호 860) F4:QSSSLVR(서열번호 861) F5:RSDKLVR(서열번호 862) F6:DSGNLRV(서열번호 863);

[1033] 25) F1:QSSSLVR(서열번호 864) F2:QSGDLRR(서열번호 865) F3:RSDERKR(서열번호 866) F4:HRTTLTN(서열번호 867) F5:RSDHLTN(서열번호 868) F6:TSGELVR(서열번호 869);

[1034] 26) F1:QSGDLRR(서열번호 870) F2:RSDERKR(서열번호 871) F3:HRTTLTN(서열번호 872) F4:RSDHLTN(서열번호 873) F5:TSGELVR(서열번호 874) F6:RSDDLVR(서열번호 875);

[1035] 27) F1:QRAHLER(서열번호 876) F2:QLAHLRA(서열번호 877) F3:DPGHLVR(서열번호 878) F4:RRSACRR(서열번호 879) F5:RSDHLTT(서열번호 880) F6:QSSSLVR(서열번호 881); 및

[1036] 28) F1:QSSNLVR(서열번호 882) F2:RSDDLVR(서열번호 883) F3:THLDLIR(서열번호 884) F4:TSGNLTE(서열번호 885) F5:RRSACRR(서열번호 886) F6:RNDTLTE(서열번호 887).

[1037] 330. 구현예 313-329 중 어느 한 구현예에 있어서, 아연 집게 단백질은 N-말단에서 C-말단 순서로 F1 내지 F6으로 표시되는 6개의 아연 집게를 포함하고,

[1038] 여기서 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같은 것인, eZFP:

[1039] F1:QSAHRKN(서열번호 822) F2:TSSNRKT(서열번호 823) F3:RSDNLSA(서열번호 824) F4:RNNDRKT(서열번호 825) F5:TSGSLSR(서열번호 826) F6:QAGHLAK(서열번호 827).

[1040] 331. 구현예 313-329 중 어느 한 구현예에 있어서, 아연 집게 단백질은 N-말단에서 C-말단 순서로 F1 내지 F6으로 표시되는 6개의 아연 집게를 포함하고,

[1041] 여기서 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같은 것인, eZFP:

[1042] F1:RSDHLSQ(서열번호 828) F2:ASSTRTK(서열번호 829) F3:RSDDLTR(서열번호 830) F4:QKSNLSS(서열번호 831) F5:QSANRTT(서열번호 832) F6:QNATRTK(서열번호 833).

[1043] 332. 구현예 313-329 중 어느 한 구현예에 있어서, 아연 집게 단백질은 N-말단에서 C-말단 순서로 F1 내지 F6으로 표시되는 6개의 아연 집게를 포함하고,

[1044] 여기서 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같은 것인, eZFP:

[1045] F1:QSSSLVR(서열번호 864) F2:QSGDLRR(서열번호 865) F3:RSDERKR(서열번호 866) F4:HRTTLTN(서열번호 867) F5:RSDHLTN(서열번호 868) F6:TSGELVR(서열번호 869).

[1046] 333. 하나 이상의 HBV 유전자 또는 이의 조절 요소의 표적 부위에 결합하는 eZFP로서, 상기 아연 집게 단백질은 N-말단에서 C-말단 순서로 F1에서 F6까지 표시되는 6개의 아연 집게를 포함하고,

[1047] 여기서 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같은 것인, eZFP:

[1048] 1) F1:SEADRSR(서열번호 720) F2:DRSNLTR(서열번호 721)F3:QSSDLSR(서열번호 722) F4:YHWYLKK(서열번호 723) F5:RSDSLSV(서열번호 724) F6:QNAANRKT(서열번호 725);

[1049] 2) F1:RSDVLST(서열번호 726) F2:DNSSRTR(서열번호 727) F3:RPYTLRL(서열번호 728) F4:DSSHRTR(서열번호 729) F5:RSDHLSQ(서열번호 730) F6:DSSHRTR(서열번호 731);

[1050] 3) F1:RSDHLSQ(서열번호 732) F2:QSADRTK(서열번호 733) F3:RSDHLSQ(서열번호 734) F4:RRSDLKR(서열번호 735) F5:RSDHLSR(서열번호 736) F6:QSSDLRR(서열번호 737);

[1051] 4) F1:RSDNLSE(서열번호 738) F2:TSSNRKT(서열번호 739) F3:DRSHLTR(서열번호 740) F4:RSDALTQ(서열번호 741) F5:DRSALAR(서열번호 742) F6:RRFTLSK(서열번호 743);

[1052] 5) F1:RSDHLSE(서열번호 744) F2:QYSGRYY(서열번호 745) F3:HGQTLNE(서열번호 746) F4:QSGNLAR(서열번호 747) F5:RSDSLLR(서열번호 748)F6:CREYRGK(서열번호 749);

[1053] 6) F1:QSANRTT(서열번호 750) F2:RSANLTR(서열번호 751) F3:RSDVLSE(서열번호 752) F4:TSGHLSR(서열번호 753) F5:QSSDLSR(서열번호 754), F6:QWSTRKR(서열번호 755);

[1054] 7) F1:QSGNLAR(서열번호 756) F2:ATCCLAH(서열번호 757) F3:RWQYLPT(서열번호 758) F4:DRSALAR(서열번호 759)

F5:RSDNLSE(서열번호 760)F6:KRCNLRC(서열번호 761);

[1055] 8) F1:NPANLTR(서열번호 762) F2:QNATRTK(서열번호 763) F3:QSGHLAR(서열번호 764) F4:NRHDRAK(서열번호 765) F5:RSDHLSE(서열번호 766), F6:QRRSRYK(서열번호 767);

[1056] 9) F1:QSSDLSR(서열번호 768) F2:HRSTRNR(서열번호 769) F3:RSDVLSA(서열번호 770) F4:DSRTRKN(서열번호 771) F5:QSGSLTR(서열번호 772) F6:DQSGLAH(서열번호 773);

[1057] 10) F1:QNPAQWR(서열번호 774) F2:RSADLSR(서열번호 775) F3:TSGSLSR(서열번호 776) F4:RSDHLSR(서열번호 777) F5:RSDSLLR(서열번호 778) F6:QSYDRFQ(서열번호 779);

[1058] 11) F1:TSGSLSR(서열번호 780) F2:RSDHLSR(서열번호 781) F3:RSDSLLR(서열번호 782) F4:QSYDRFQ(서열번호 783) F5:RSDNLST(서열번호 784) F6:DNRDRIK(서열번호 785);

[1059] 12) F1:DRSNLSR(서열번호 786) F2:LRQNLIM(서열번호 787) F3:ERGTLAR(서열번호 788) F4:RSDALTQ(서열번호 789) F5:RSDSLSQ(서열번호 790) F6:RKADRTR(서열번호 791);

[1060] 13) F1:QYCCLTN(서열번호 792) F2:TSGNLTR(서열번호 793) F3:QSSDLSR(서열번호 794) F4:FRYYLKR(서열번호 795) F5:QSGDLTR(서열번호 796) F6:DKGNLTK(서열번호 797);

[1061] 14) F1:TSGSLSR(서열번호 798) F2:RSDNLTT(서열번호 799) F3:QSGNLAR(서열번호 800) F4:DRTTLMR(서열번호 801) F5:QSGHLAR(서열번호 802) F6:QLTHLNS(서열번호 803);

[1062] 15) F1:IKHDLHR(서열번호 804) F2:RSANLTR(서열번호 805) F3:RSDNLAR(서열번호 806) F4:QNVSRPR(서열번호 807) F5:RSDDLSK(서열번호 808) F6:DSSHRTR(서열번호 809);

[1063] 16) F1:RSDNLAR(서열번호 810) F2:QNVSRPR(서열번호 811) F3:RSDDLSK(서열번호 812) F4:DSSHRTR(서열번호 813) F5:TSSNRKT(서열번호 814) F6:AQWTRAC(서열번호 815);

[1064] 17) F1:RSDDLSK(서열번호 816) F2:DSSHRTR(서열번호 817) F3:TSSNRKT(서열번호 818) F4:AQWTRAC(서열번호 819) F5:RKQTRTT(서열번호 820) F6:HRSSLRR(서열번호 821);

[1065] 18) F1:QSAHRKN(서열번호 822) F2:TSSNRKT(서열번호 823) F3:RSDNLSA(서열번호 824) F4:RNNDRKT(서열번호 825) F5:TSGSLSR(서열번호 826) F6:QAGHLAK(서열번호 827);

[1066] 19) F1:RSDHLSQ(서열번호 828) F2:ASSTRTK(서열번호 829) F3:RSDDLTR(서열번호 830) F4:QKSNLSS(서열번호 831) F5:QSANRTT(서열번호 832) F6:QNATRTK(서열번호 833);

[1067] 20) F1:RSDTLSE(서열번호 834) F2:RRWTLVG(서열번호 835) F3:DRSNLSR(서열번호 836) F4:QSGDLTR(서열번호 837) F5:QSSDLSR(서열번호 838) F6:YHWYLKK(서열번호 839);

[1068] 21) F1:RSANLAR(서열번호 840) F2:RSDNLRE(서열번호 841) F3:RPYTLRL(서열번호 842) F4:HRSNLNK(서열번호 843) F5:QSGSLTR(서열번호 844) F6:TSANLSR(서열번호 845);

[1069] 22) F1:RSDDLVR(서열번호 846) F2:TSGSLVR(서열번호 847) F3:RSDKLVR(서열번호 848) F4:RSDELVR(서열번호 849) F5:TSHSLTE(서열번호 850) F6:RADNLTE(서열번호 851);

[1070] 23) F1:ERSHLRE(서열번호 852) F2:TSHSLTE(서열번호 853) F3:QAGHLAS(서열번호 854) F4:TSHSLTE(서열번호 855) F5:DPGHLVR(서열번호 856) F6:TSGNLVR(서열번호 857);

[1071] 24) F1:RADNLTE(서열번호 858) F2:TSGSLVR(서열번호 859) F3:RKDNLKN(서열번호 860) F4:QSSSLVR(서열번호 861) F5:RSDKLVR(서열번호 862) F6:DSGNLRV(서열번호 863);

[1072] 25) F1:QSSSLVR(서열번호 864) F2:QSGDLRR(서열번호 865) F3:RSDERKR(서열번호 866) F4:HRTTLTN(서열번호 867) F5:RSDHLTN(서열번호 868) F6:TSGELVR(서열번호 869);

[1073] 26) F1:QSGDLRR(서열번호 870) F2:RSDERKR(서열번호 871) F3:HRTTLTN(서열번호 872) F4:RSDHLTN(서열번호 873) F5:TSGELVR(서열번호 874) F6:RSDDLVR(서열번호 875);

[1074] 27) F1:QRAHLER(서열번호 876) F2:QLAHLRA(서열번호 877) F3:DPGHLVR(서열번호 878) F4:RRSACRR(서열번호 879) F5:RSDHLTT(서열번호 880) F6:QSSSLVR(서열번호 881); 및

[1075] 28) F1:QSSNLVR(서열번호 882) F2:RSDDLVR(서열번호 883) F3:THLDLIR(서열번호 884) F4:TSGNLTE(서열번호 885)

F5:RRSACRR(서열번호 886) F6:RNDTLTE(서열번호 887).

[1076] 334. 구현예 313-333 중 어느 한 구현예에 있어서, 조작된 아연 집게 단백질은 서열번호 692-719 중 어느 하나에 제시된 서열 또는 이의 일부를 포함하거나, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는, eZFP.

[1077] 335. 구현예 313-334 중 어느 한 구현예에 있어서, 조작된 아연 집게 단백질은 서열번호 888-915 중 어느 하나에 제시된 서열 또는 이의 일부에 의해 인코딩되거나, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩되는, eZFP.

[1078] 336. 하나 이상의 HBV 유전자 또는 이의 조절 요소의 표적 부위에 결합하는 조작된 아연 집게 단백질로서, 여기서 아연 집게 단백질은 N-말단에서 C-말단 순서로 F1에서 F6까지 표시된 6개의 아연 집게를 포함하고,

[1079] 여기서 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같은 것인, eZFP:

[1080] F1:QSAHRKN(서열번호 822) F2:TSSNRKT(서열번호 823) F3:RSDNLSA(서열번호 824) F4:RNNDRKT(서열번호 825) F5:TSGSLSR(서열번호 826) F6:QAGHLAK(서열번호 827).

[1081] 337. 구현예 313-336 중 어느 한 구현예에 있어서, 조작된 아연 집게 단백질은 서열번호 709에 제시된 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는, eZFP.

[1082] 338. 구현예 313-337 중 어느 한 구현예에 있어서, 조작된 아연 집게 단백질은 서열번호 709 중 어느 하나에 제시된 서열을 포함하는, eZFP.

[1083] 339. 구현예 313-338 중 어느 한 구현예에 있어서, 조작된 아연 집게 단백질은 서열번호 905에 제시된 서열 또는 이의 일부에 의해 인코딩되거나, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩되는, eZFP.

[1084] 340. 구현예 313-339 중 어느 한 구현예에 있어서, 조작된 아연 집게 단백질은 서열번호 905 중 어느 하나에 제시된 서열에 의해 인코딩되는, eZFP.

[1085] 341. 하나 이상의 HBV 유전자 또는 이의 조절 요소의 표적 부위에 결합하는 조작된 아연 집게 단백질로서, 상기 아연 집게 단백질은 N-말단에서 C-말단 순서로 F1에서 F6까지 표시된 6개의 아연 집게를 포함하고,

[1086] 여기서 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같은 것인, eZFP:

[1087] F1:RSDHLSQ(서열번호 828) F2:ASSTRTK(서열번호 829) F3:RSDDLTR(서열번호 830) F4:QKSNLSS(서열번호 831) F5:QSANRTT(서열번호 832) F6:QNATRTK(서열번호 833).

[1088] 342. 구현예 313-335 및 341 중 어느 한 구현예에 있어서, 조작된 아연 집게 단백질은 서열번호 710에 제시된 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는, eZFP.

[1089] 343. 구현예 313-335, 341 및 342 중 어느 한 구현예에 있어서, 조작된 아연 집게 단백질은 서열번호 710 중 어느 하나에 제시된 서열을 포함하는, eZFP.

[1090] 344. 구현예 313-335 및 341-343 중 어느 한 구현예에 있어서, 조작된 아연 집게 단백질은 서열번호 906에 제시된 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩되는, eZFP.

[1091] 345. 구현예 313-335 및 341-344 중 어느 한 구현예에 있어서, 조작된 아연 집게 단백질은 서열번호 906 중 어느 하나에 제시된 서열에 의해 인코딩되는, eZFP.

[1092] 346. 하나 이상의 HBV 유전자 또는 이의 조절 요소의 표적 부위에 결합하는 조작된 아연 집게 단백질로서, 여기서 아연 집게 단백질은 N-말단에서 C-말단 순서로 F1에서 F6까지 표시된 6개의 아연 집게를 포함하고,

[1093] 여기서 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같은 것인, eZFP:

[1094] F1:QSSSLVR(서열번호 864) F2:QSGDLRR(서열번호 865) F3:RSDERKR(서열번호 866) F4:HRTTLTN(서열번호 867) F5:RSDHLTN(서열번호 868) F6:TSGELVR(서열번호 869).

[1095] 347. 구현예 313-335 및 346 중 어느 한 구현예에 있어서, 조작된 아연 집게 단백질은 서열번호 716에 제시된 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 아미노산 서열을 포함하는, eZFP.

[1096] 348. 구현예 313-335, 346 및 347 중 어느 한 구현예에 있어서, 조작된 아연 집게 단백질은 서열번호 716 중 어느 하나에 제시된 서열을 포함하는, eZFP.

[1097] 349. 구현예 313-335 및 346-348 중 어느 한 구현예에 있어서, 조작된 아연 집게 단백질은 서열번호 912에 제시된 서열 또는 이의 일부, 또는 이와 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성을 갖는 뉴클레오타이드 서열에 의해 인코딩되는, eZFP.

[1098] 350. 구현예 313-335 및 346-349 중 어느 한 구현예에 있어서, 조작된 아연 집게 단백질은 서열번호 912 중 어느 하나에 제시된 서열에 의해 인코딩되는, eZFP.

[1099] [0564] 또한 다음의 구현예들도 제공된다:

[1100] 1. 하나 이상의 B형 간염 바이러스(HBV) 유전자의 전사를 억제하기 위한 적어도 하나의 DNA-표적화 모듈을 포함하는 후성적 변형 DNA-표적화 시스템으로서, 여기서 적어도 하나의 DNA-표적화 모듈 각각은 다음을 포함하는 융합 단백질을 포함하는 것인, 후성적 변형 DNA-표적화 시스템:

[1101] (a) B형 간염 바이러스 DNA 서열의 표적 부위를 표적화하기 위한 DNA-결합 도메인으로서, 여기서 표적 부위는 서열번호 650의 뉴클레오타이드 위치를 참조로 B형 간염 바이러스 서열의 1033bp-1749bp 사이의 염기쌍에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내에 있는 것인 DNA-결합 도메인; 및

[1102] (b) 적어도 하나의 전사 억제자 효과기 도메인.

[1103] 2. 구현예 1에 있어서, 표적 부위는 B형 간염 X 단백질(HBx) 시작 코돈 상류 300개 염기쌍 내에 위치한 표적 영역 내에 있는, 후성적 변형 DNA-표적화 시스템.

[1104] 3. B형 간염 바이러스(HBV) 유전자의 전사를 억제하기 위한 적어도 하나의 DNA-표적화 모듈을 포함하는 후성적 변형 DNA-표적화 시스템으로서, 여기서 DNA-표적화 모듈은 다음을 포함하는 융합 단백질을 포함하는 것인, 후성적 변형 DNA-표적화 시스템:

[1105] (a) B형 간염 X 단백질(HBx) 시작 코돈 상류의 300개 염기쌍 내에 걸쳐 있는 표적 영역 내의 표적 부위를 표적화하기 위한 DNA-결합 도메인; 및

[1106] (b) 적어도 하나의 전사 억제자 효과기 도메인.

[1107] 4. 구현예 1-3 중 어느 한 구현예에 있어서, 표적 부위는 HBx 기저 코어 프로모터 영역에 위치하는, 후성적 변형 DNA-표적화 시스템.

[1108] 5. 구현예 1-3 중 어느 한 구현예에 있어서, 표적 부위는 HBx 프로모터/인핸서 영역 내에 위치하는, 후성적 변형 DNA-표적화 시스템.

[1109] 6. 구현예 1-5 중 어느 한 구현예에 있어서, 표적 부위는 B형 간염 X 단백질(HBx) 시작 코돈 상류의 150개 염기쌍 내에 걸쳐 있는 표적 영역 내에 있는, 후성적 변형 DNA-표적화 시스템.

[1110] 7. 구현예 1-6 중 어느 한 구현예에 있어서, 표적 부위는 서열번호 650에 제시된 HBV 게놈을 기준으로 1255-1302 bp 사이의 염기쌍에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내에 있는, 후성적 변형 DNA-표적화 시스템.

[1111] 8. 구현예 1-7 중 어느 한 구현예에 있어서, 표적 부위는 서열번호 650에 제시된 HBV 게놈을 참조로 1260-1300bp 사이의 염기쌍에 위치한 서열에 상응하는 서열을 갖는 표적 영역 내에 있는, 후성적 변형 DNA-표적화 시스템.

[1112] 9. 구현예 1-8 중 어느 한 구현예에 있어서, 표적 부위는 서열번호 22 또는 서열번호 63에 제시된 서열, 적어도 14개 뉴클레오타이드의 이의 연속 부분 또는 상기 중 어느 하나의 상보적 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

[1113] 10. 구현예 1-9 중 어느 한 구현예에 있어서, 적어도 하나의 DNA-결합 도메인은 (a) 비활성화된 Cas(dCas) 및 (b) 적어도 하나의 gRNA를 포함하는 규칙적인 간격을 갖고 나타나는 짧은 회문 구조의 반복 서열(Clustered

Regularly Interspaced Short Palindromic Repeats) 관련(Cas)-가이드 RNA(gRNA) 조합을 포함하는, 후성적 변형 DNA-표적화 시스템.

[1114] 11. 구현예 10에 있어서, dCas는 dCas9 단백질인, 후성적 변형 DNA-표적화 시스템.

[1115] 12. 구현예 11에 있어서, dCas9는 서열번호 598의 위치 번호를 참조로 D10A 및 H840A 중에서 선택된 적어도 하나의 아미노산 돌연변이를 포함하는 화농성연쇄상구균(*Streptococcus pyogenes*) Cas9(SpCas9) 단백질인, 후성적 변형 DNA-표적화 시스템.

[1116] 13. 구현예 10-12 중 어느 한 구현예에 있어서, gRNA는 서열번호 217에 제시된 서열, 적어도 14개 뉴클레오타이드의 이의 연속 부분, 또는 앞서 언급한 서열의 상보적 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

[1117] 14. 구현예 10-13 중 어느 한 구현예에 있어서, gRNA는 서열번호 412에 제시된 것인, 후성적 변형 DNA-표적화 시스템.

[1118] 15. 구현예 10-12 중 어느 한 구현예에 있어서, gRNA는 서열번호 258에 제시된 서열, 적어도 14개 뉴클레오타이드의 이의 연속 부분, 또는 앞서 언급한 서열의 상보적 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

[1119] 16. 구현예 10-12 및 15 중 어느 한 구현예에 있어서, gRNA는 서열번호 453에 제시된 것인, 후성적 변형 DNA-표적화 시스템.

[1120] 17. 구현예 1-9 중 어느 한 구현예에 있어서, 적어도 하나의 DNA-결합 도메인은 조작된 아연 집게 단백질(eZFP)을 포함하는, 후성적 변형 DNA-표적화 시스템.

[1121] 18. 구현예 1-9 및 17 중 어느 한 구현예에 있어서, 표적 부위는 서열번호 1045, 1046, 1052 중 어느 하나에 명시된 뉴클레오타이드 서열, 적어도 12nt의 이의 연속 부분, 또는 앞서 언급한 서열의 상보적 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

[1122] 19. 구현예 17-18 중 어느 한 구현예에 있어서, 아연 집게 단백질은 N-말단에서 C-말단의 순서로 F1에서 F6까지 표시된 6개의 아연 집게를 포함하고, 각 아연 집게 인식 영역의 아미노산 서열은 다음과 같은 것인, 후성적 변형 DNA-표적화 시스템:

[1123] (A) F1:QSAHRKN(서열번호 822) F2:TSSNRKT(서열번호 823) F3:RSDNLSA(서열번호 824) F4:RNNDRKT(서열번호 825) F5:TSGSLSR(서열번호 826) F6:QAGHLAK(서열번호 827);

[1124] (B) F1:RSDHLSQ(서열번호 828) F2:ASSTRTK(서열번호 829) F3:RSDDLTR(서열번호 830) F4:QKSNLSS(서열번호 831) F5:QSANRTT(서열번호 832) F6:QNATRTK(서열번호 833); 또는

[1125] (C) F1:QSSSLVR(서열번호 864) F2:QSGDLRR(서열번호 865) F3:RSDERKR(서열번호 866) F4:HRTTLTN(서열번호 867) F5:RSDHLTN(서열번호 868) F6:TSGELVR(서열번호 869).

[1126] 20. 구현예 17-19 중 어느 한 구현예에 있어서, eZFP는 서열번호 709, 서열번호 710 또는 서열번호 716에 제시된 서열, 또는 이와 적어도 90%의 서열 동일성을 갖는 아미노산 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

[1127] 21. 구현예 1-20 중 어느 한 구현예에 있어서, 적어도 하나의 효과기 도메인은 DNA 메틸트랜스퍼라제인, 후성적 변형 DNA-표적화 시스템.

[1128] 22. 구현예 1-21 중 어느 한 구현예에 있어서, 적어도 하나의 효과기 도메인은 DNA 메틸트랜스퍼라제와 헤테로크로마틴 유도 인자를 동원할 수 있는 억제자 도메인을 포함하는, 후성적 변형 DNA-표적화 시스템.

[1129] 23. 구현예 1-22 중 어느 한 구현예에 있어서, 적어도 하나의 효과기 도메인은 DNA 메틸트랜스퍼라제와 히스톤 메틸트랜스퍼라제를 포함하는, 후성적 변형 DNA-표적화 시스템.

[1130] 24. 구현예 1-23 중 어느 한 구현예에 있어서, 적어도 하나의 효과기 도메인은 KRAB 억제자 도메인, ERF 억제자 도메인, Mxi1 억제자 도메인, SID4X 억제자 도메인, Mad-SID 억제자 도메인, LSD1 억제자 도메인, 또는 DNMT3A, DNMT3A-3L, DNMT3A/L-KRAB 융합 억제자 도메인, DNMT3B 도메인 결합 단백질 또는 LSD1 억제자 도메인, 또는 상기 중 어느 것의 변이체 중에서 선택되는, 후성적 변형 DNA-표적화 시스템.

[1131] 25. 구현예 1-24 중 어느 하나의 후성적 변형 DNA-표적화 시스템 또는 구현예 1-24 중 어느 하나의 DNA-표적화 시스템의 융합 단백질, 또는 상기 중 어느 하나의 일부 또는 구성요소를 인코딩하는, 폴리뉴클레오타이드.

[1132] 26. 구현예 25의 폴리뉴클레오타이드를 포함하는, 벡터.

[1133] [0565] 또한 다음의 구현예들도 제공된다:

[1134] 1. 하나 이상의 B형 간염 바이러스(HBV) 유전자의 전사를 억제하기 위한 적어도 하나의 DNA-표적화 모듈을 포함하는 후성적 변형 DNA-표적화 시스템으로서, 여기서 적어도 하나의 DNA-표적화 모듈 각각은:

[1135] (a) B형 간염 바이러스 DNA 서열의 표적 부위를 표적화하기 위한 DNA-결합 도메인; 및

[1136] (b) KRAB 도메인 및 DNMT3A/L 도메인을 포함하는 전사 억제자 효과기 도메인

[1137] 을 포함하는 융합 단백질을 포함하는 것인, 후성적 변형 DNA-표적화 시스템.

[1138] 2. 구현예 1에 있어서, KRAB 도메인은 서열번호 590에 제시된 서열을 포함하거나, 서열번호 590과 최소 90%의 서열 동일성을 갖는 아미노산 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

[1139] 3. 구현예 1 또는 2에 있어서, DNMT3A/L 도메인은 서열번호 651에 제시된 서열을 포함하거나, 서열번호 651과 적어도 90%의 서열 동일성을 갖는 아미노산 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

[1140] 4. 하나 이상의 B형 간염 바이러스(HBV) 유전자의 전사를 억제하기 위한 적어도 하나의 DNA-표적화 모듈을 포함하는 후성적 변형 DNA-표적화 시스템으로서, 여기서 적어도 하나의 DNA-표적화 모듈 각각은 다음을 포함하는 융합 단백질을 포함하는 것인, 후성적 변형 DNA-표적화 시스템:

[1141] (a) B형 간염 바이러스 DNA 서열의 표적 부위를 표적화하기 위한 DNA-결합 도메인; 및

[1142] (b) KRAB 도메인 및 DNMT3A/L 도메인을 포함하는 전사 억제자 효과기 도메인으로서, 여기서 KRAB 도메인은 서열번호 590에 제시된 서열, 또는 서열번호 590과 적어도 90%의 서열 동일성을 갖는 아미노산 서열을 포함하고, DNMT3A/L 도메인은 서열번호 651에 제시된 서열, 또는 서열번호 651과 적어도 90%의 서열 동일성을 갖는 아미노산 서열을 포함하는 것인, 전사 억제자 효과기 도메인.

[1143] 5. 구현예 1-4 중 어느 한 구현예에 있어서, KRAB 도메인과 DNMT3A/L 도메인은 DNA-결합 도메인의 N 말단, C 말단, 또는 N 말단과 C 말단 둘 다에 독립적으로 융합되는, 후성적 변형 DNA-표적화 시스템.

[1144] 6. 구현예 1-5 중 어느 한 구현예에 있어서, 융합 단백질은 하나 이상의 핵 국소화 신호(NLS)를 더 포함하는, 후성적 변형 DNA-표적화 시스템.

[1145] 7. 구현예 6에 있어서, 융합 단백질은 다음 중 두 개 이상을 연결하는 하나 이상의 링커를 더 포함하는, 후성적 변형 DNA-표적화 시스템: DNA-결합 도메인, KRAB 도메인, DNMT3A/L 도메인 및 하나 이상의 핵 국소화 신호.

[1146] 8. 구현예 1-7 중 어느 한 구현예에 있어서, B형 간염 바이러스 DNA 서열은 HBV 유전자 또는 이의 조절 요소인, 후성적 변형 DNA-표적화 시스템.

[1147] 9. 구현예 1-8 중 어느 한 구현예에 있어서, 표적 부위는 공유적으로 폐쇄된 원형 DNA(cccDNA) 형태, 이완된 원형 DNA(rcDNA) 형태로 존재하고/존재하거나 인간 게놈 DNA에 통합된 HBV 바이러스 DNA인, 후성적 변형 DNA-표적화 시스템.

[1148] 10. 구현예 1-9 중 어느 한 구현예에 있어서, 표적 부위는 HBV 복제 및/또는 HBV 전사를 제어하는 데 관여하는 유전자 또는 이의 조절 요소에 존재하거나 또는 그 근처에 존재하고, 유전자는 폴리머라제 유전자, S-패밀리 유전자, X-유전자 또는 코어 패밀리 유전자인, 후성적 변형 DNA-표적화 시스템.

[1149] 11. 구현예 1-10 중 어느 한 구현예에 있어서, 적어도 하나의 표적 부위는 B형 간염 바이러스 단백질 X(HBx)를 인코딩하는 X 유전자의 유전자 또는 이의 조절 요소에 있는, 후성적 변형 DNA-표적화 시스템.

[1150] 12. 구현예 10에 있어서, 조절 요소는 프로모터 영역인, 후성적 변형 DNA-표적화 시스템.

[1151] 13. 구현예 12에 있어서, 프로모터 영역은 pre-S1 프로모터, pre-S2 프로모터, X 프로모터 또는 기저 코어 프로모터인, 후성적 변형 DNA-표적화 시스템.

[1152] 14. 구현예 10에 있어서, 조절 요소는 인핸서 영역인, 후성적 변형 DNA-표적화 시스템.

[1153] 15. 구현예 14에 있어서, 인핸서 영역은 Enh1 또는 Enh2 인핸서 영역인, 후성적 변형 DNA-표적화 시스템.

[1154] 16. 구현예 1-15 중 어느 한 구현예에 있어서, 표적 부위는 최소 1000개의 B형 간염 바이러스 게놈과 최소 70%

상동성을 갖는, 후성적 변형 DNA-표적화 시스템.

[1155] 17. 구현예 1-16 중 어느 한 구현예에 있어서, 적어도 하나의 DNA-결합 도메인은 (a) 규칙적인 간격을 갖고 나타나는 짧은 회문 구조의 반복 서열 관련(Cas) 단백질 또는 이의 변이체 및 (b) 적어도 하나의 gRNA를 포함하는 Cas-가이드 RNA(gRNA) 조합; 아연 집게 단백질(ZFP); 전사 활성화 인자 유사 효과기(TALE); 메가뉴클레아제; 호밍d 엔도뉴클레아제; 또는 I-SceI 효소 또는 이의 변이체를 포함하는, 후성적 변형 DNA-표적화 시스템.

[1156] 18. 구현예 1-18 중 어느 한 구현예에 있어서, DNA-결합 도메인은 Cas-gRNA 조합이고, Cas는 비활성화된 Cas(dCas) 단백질인, 후성적 변형 DNA-표적화 시스템.

[1157] 19. 구현예 18에 있어서, dCas는 비활성화된 Cas9(dCas9) 단백질인, 후성적 변형 DNA-표적화 시스템.

[1158] 20. 구현예 19에 있어서, dCas9는 서열번호 596의 위치 번호를 참조하여 D10A 및 N580A에서 선택된 적어도 하나의 아미노산 돌연변이를 포함하는 황색포도상구균(*Staphylococcus aureus*) dCas9 단백질(dSaCas9)인, 후성적 변형 DNA-표적화 시스템.

[1159] 21. 구현예 19에 있어서, dCas9는 서열번호 598의 위치 번호를 참조로 D10A 및 H840A 중에서 선택된 적어도 하나의 아미노산 돌연변이를 포함하는 화농성연쇄상구균(*Streptococcus pyogenes*) dCas9 단백질(dSpCas9)인, 후성적 변형 DNA-표적화 시스템.

[1160] 22. 다음을 포함하는 후성적 변형 DNA-표적화 시스템:

[1161] (a) (i) 서열번호 596의 위치 번호를 참조로, D10A 및 N580A 중에서 선택된 적어도 하나의 아미노산 돌연변이를 포함하는 비활성화된 Cas9(dCas9)인 규칙적인 간격을 갖고 나타나는 짧은 회문 구조의 반복 서열 연관 단백질 9(Cas9) 및 (ii) KRAB 도메인 및 DNMT3A/L 도메인을 포함하는 전사 억제자 효과기 도메인을 포함하는 융합 단백질로서, 여기서 KRAB 도메인은 서열번호 590과 적어도 90%의 서열 동일성을 갖는 아미노산 서열을 포함하고 DNMT3A/L 도메인은 서열번호 651과 적어도 90%의 서열 동일성을 갖는 아미노산 서열을 포함하는 것인, 융합 단백질; 및

[1162] (b) B형 간염 바이러스 DNA 서열의 표적 부위를 표적으로 삼기 위한 gRNA.

[1163] 23. 구현예 22에 있어서, 표적 부위는 B형 간염 바이러스 단백질 X(HBx)를 인코딩하는 X 유전자의 유전자 또는 이의 조절 요소에 있는, 후성적 변형 DNA-표적화 시스템.

[1164] 24. 구현예 1-18 중 어느 한 구현예에 있어서, 적어도 하나의 DNA-결합 도메인은 조작된 아연 집게 단백질(eZFP)을 포함하는, 후성적 변형 DNA-표적화 시스템.

[1165] 25. (i) B형 간염 바이러스 DNA 서열의 표적 부위를 표적으로 하기 위한 조작된 아연 집게 단백질(eZFP) 및 (ii) KRAB 도메인과 DNMT3A/L 도메인을 포함하는 전사 억제자 효과기 도메인을 포함하는 융합 단백질을 포함하는 후성적 변형 DNA-표적화 시스템으로서, 여기서 KRAB 도메인은 서열번호 590과 적어도 90%의 서열 동일성을 갖는 아미노산 서열을 포함하고 DNMT3A/L 도메인은 서열번호 651과 적어도 90%의 서열 동일성을 갖는 아미노산 서열을 포함하는, 후성적 변형 DNA-표적화 시스템.

[1166] 26. 구현예 25에 있어서, 표적 부위는 B형 간염 바이러스 단백질 X(HBx)를 인코딩하는 X 유전자의 유전자 또는 이의 조절 요소에 있는, 후성적 변형 DNA-표적화 시스템.

[1167] 27. 구현예 1-26 중 어느 한 구현예에 있어서, DNA-표적화 시스템은 모든 B형 간염 바이러스 게놈의 적어도 70%를 표적화하는, 후성적 변형 DNA-표적화 시스템.

[1168] 28. 구현예 1-27 중 어느 하나의 후성적 변형 DNA-표적화 시스템 또는 구현예 1-27 중 어느 하나의 DNA-표적화 시스템의 융합 단백질, 또는 앞서 언급한 것의 일부 또는 구성요소를 인코딩하는, 폴리뉴클레오타이드.

[1169] 29. 구현예 28의 폴리뉴클레오타이드를 포함하는, 벡터.

[1170] **VIII. 실시예**

[1171] [0566] 다음 실시예는 설명 목적으로만 포함된 것으로 발명의 범위가 이에 한정되는 것은 아니다.

[1172] **실시예 1: B형 간염 바이러스 유전자 및 바이러스 복제 및 전사와 관련된 이의 조절 요소를 표적으로 하는 gRNA의 설계.**

[1173] [0567] B형 간염 바이러스(HBV) 유전자 및 바이러스 복제 및 전사와 관련된 조절 요소를 표적화하도록 gRNA를

설계하였다.

[1174] [0568] 컴퓨터 분석을 이용하여 6939개의 상이한 HBV 게놈(유전형 A-E)으로부터 마스터 gRNA 프로토스페이서 목록(표적 부위)을 생성하였다. 마스터 gRNA 목록에는 모든 게놈("[N](x20)-NGG"로 정의됨)에서 발견된 110, 438개의 고유한 프로토스페이서-PAM 서열(표적 부위)가 모두 포함되었다. 다음으로, 프로토스페이서 서열(표적 부위)를 모든 6939개의 HBV 게놈에 다시 정렬하였다. 그런 다음 프로토스페이서 서열을 프로토스페이서 인접 모티프(PAM)의 5' 말단에 있는 12개 염기의 상동성과 불일치를 기준으로, 'nnnnnnnnnnnnNNNNNN-NGG'에서 'n'으로 표시되는 바와 같이, 다음의 세 가지 범주로 오거나이즈하였다: 1) 불일치 없이 프로토스페이서가 존재함(HBV0), 2) 불일치가 하나 허용된 경우 프로토스페이서가 존재함(HBV1), 3) 최대 2개의 불일치가 허용된 경우 프로토스페이서가 존재함(HBV2).

[1175] [0569] 모든 프로토스페이서 서열에 대해이 세 가지 점수가 각각 생성되었다. 점수는 주어진 HBV 게놈에 대해 PAM의 5'-말단의 불일치에 대한 특정 허용 범위를 고려할 때 프로토스페이서가 존재하고 활성화될 확률을 나타낸다. 분석에서 동일한 코어 프로토스페이서-PAM 서열에 여러 변이체가 존재하는 것으로 밝혀졌다. 따라서 프로토스페이서 서열 목록은 코어 프로토스페이서-PAM 서열(NNNNNNNN-NGG)를 기준으로 추가로 그룹화되었다. 마지막으로 각 그룹에서 가장 높은 완벽한 게놈 정렬 점수(상동성)를 갖는 프로토스페이서 서열이 주어진 코어 프로토스페이서-PAM 서열에 대한 후보 gRNA로 선택되었다. HBV 게놈이 90% 이상 보존되고 불일치가 없는(HBV0) 스크리닝 후보 스페이서를 표 E1에 나타내었다. HBV 게놈이 90% 이상 보존되고 불일치가 1-2개인(HBV1) 스크리닝 후보 gRNA 스페이서는 표 E1에 나와 있다. 70-90% HBV 게놈 보존 및 최대 2개의 불일치(HBV2)를 갖는 스크리닝 후보 gRNA 스페이서는 표 E1에 나와 있다. 각 gRNA는 또한 SpCas9에 대한 스캐폴드 서열을 포함하며, 이는 다음 서열을 포함한다: GUUUAAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUCAA CUUGAAAAAGUGGCACCGAGUCGGUGC (서열번호 587).

[1176] [0570] 5'- (프로토스페이서) mN*mN*mN*rNrNrNrNrNrNrNrNrNrNrNrNrNrNrNrNrNrNrNrNrNrN rN

[1177] (스캐폴드)
rGrUrUrUrUrArGrArGrArArGrCrUrArArArArArArGrUrUrArArArArArGrGrUrUrArArArArGrGrCUrArGrUrCrCrGrUrUrArArArArArArGrUrGrUrGrArArArArArGrUrGrGrCrCrGrUrUrArArArArGrUrGrGrUrGrCrUrCrUrGrUrGrCrUrCrUrUrCrUrUrCrUrUrCrU*mU*mU*mU-3'

[1178] [0571] 여기서 m=2' OMethyl RNA이고 * = 포스포로티오화 결합이다.

표 E1. HBV 게놈 보존도가 >90%이고 불일치가 없는 후보 gRNA(HBV0)

표적 부위 (프로토스페이서) 서열	표적 SEQ ID	서열명	gRNA 스페이서 서열	RNA 스페이서 SEQ ID
CCCTATCTTATCAACACTTC	1	HBVg_1	CCCUAUCUUAUCAACACUUC	196
GCAGAGGTGAAAAAGTTGCA	2	HBVg_2	GCAGAGGUGAAAAAGUUGCA	197
TGGACTTCTCTCAATTTTCT	3	HBVg_3	UGGACUUCUCUCAAUUUUCU	198
ACCCCGCCTGTAACACGAGC	4	HBVg_4	ACCCCGCCUGUAACACGAGC	199
CCCGCCTGTAACACGAGCAG	5	HBVg_5	CCCGCCUGUAACACGAGCAG	200
CACCACGAGTCTAGACTCTG	6	HBVg_6	CACCACGAGUCUAGACUCUG	201
GAGGTGAAGCGAAGTGCACA	7	HBVg_7	GAGGUGAAGCGAAGUGCACA	202
CCGGAAGTGTTGATAAGATA	8	HBVg_8	CCGGAAGUGUUGAUAAGAUA	203
AGAAGATGAGGCATAGCAGC	9	HBVg_9	AGAAGAUGAGGCAUAGCAGC	204
TCCGCAGTATGGATCGGCAG	10	HBVg_10	UCCGCAGUAUGGAUCGGCAG	205
GGACTTCTCTCAATTTTCTA	11	HBVg_11	GGACUUCUCUCAAUUUUCUA	206
CCACCCAAGGCACAGCTTGG	12	HBVg_12	CCACCCAAGGCACAGCUUGG	207
AGAGAGGTGCGCCCCGTGGT	13	HBVg_13	AGAGAGGUGCGCCCCGUGGU	208
GATTGAGATCTTCTGCGACG	14	HBVg_14	GAUUGAGAUCUUCUGCGACG	209
CAAGCCTCCAAGCTGTGCCT	15	HBVg_15	CAAGCCUCCAAGCUGUGCCU	210
GGCGAGGGAGTTCTTCTTCT	16	HBVg_16	GGCGAGGGAGUUCUUCUUCU	211
TCCGGAAGTGTTGATAAGAT	17	HBVg_17	UCCGGAAGUGUUGAUAAGAU	212
AAGCCACCCAAGGCACAGCT	18	HBVg_18	AAGCCACCCAAGGCACAGCU	213
CCTCCAAGCTGTGCCTTGGG	19	HBVg_19	CCUCCAAGCUGUGCCUUGGG	214
GTAAAGAGAGGTGCGCCCCG	20	HBVg_20	GUAAAGAGAGGUGCGCCCCG	215
GGCAGATGAGAAGGCACAGA	21	HBVg_21	GGCAGAUGAGAAGGCACAGA	216
AGGAGTTCCGCAGTATGGAT	22	HBVg_22	AGGAGUUCCGCAGUAUGGAU	217
GCTGTGCCTTGGGTGGCTTT	23	HBVg_23	GCUGUGCCUUGGGUGGCUUU	218
ACCCCTGCTCGTGTTACAGG	24	HBVg_24	ACCCCUGCUCGUGUUACAGG	219
CGGAAGTGTTGATAAGATAG	25	HBVg_25	CGGAAGUGUUGAUAAGAUAG	220
CCTGCTGGTGGCTCCAGTTC	26	HBVg_26	CCUGCUGGUGGCUCCAGUUC	221
CGAGGGAGTTCTTCTTCTAG	27	HBVg_27	CGAGGGAGUUCUUCUUCUAG	222
GGGGCGCACCTCTCTTTACG	28	HBVg_28	GGGGCGCACCUCUCUUUACG	223
AGCTTGGAGGCTTGAACAGT	29	HBVg_29	AGCUUGGAGGCUUGAACAGU	224
AAGCCTCCAAGCTGTGCCTT	30	HBVg_30	AAGCCUCCAAGCUGUGCCUU	225

[1179]

CCCCTGCTCGTGTTACAGGC	31	HBVg_31	CCCCUGCUCGUGUUACAGGC	226
GCGAGGGAGTTCTTCTTCTA	32	HBVg_32	GCGAGGGAGUUCUUCUUCUA	227
TACTAGTGCCATTTGTTCAG	33	HBVg_33	UACUAGUGCCAUUUGUUCAG	228
GACTTCTCTCAATTTTCTAG	34	HBVg_34	GACUUCUCUCAAUUUUCUAG	229

표 E2. HBV 게놈 보존도가 >90%이고 1-2 개의 불일치를 갖는 후보 gRNA(HBV1)

표적 부위 (프로토스페이서) 서열	표적 SEQ ID	표적 서열	gRNA 스페이서 서열	RNA 스페이서 SEQ ID
ATTGACCCGTATAAAGAATT	35	HBVg_35	AUUGACCCGUAUAAAGAAU U	230
ACCCAAAGACAAAAGAAAAT	36	HBVg_36	ACCCAAAGACAAAAGAAAA U	231
GTCCTCTTATGTAAGACCTT	37	HBVg_37	GUCCUCUUAUGUAAGACCU U	232
TGATCGGGAAAGAATCCCAG	38	HBVg_38	UGAUCGGGAAAGAAUCCCA G	233
TTTGCTGACGCAACCCCCAC	39	HBVg_39	UUUGCUGACGCAACCCCCA C	234
ATGAATCTAGCCACCTGGGT	40	HBVg_40	AUGAAUCUAGCCACCUGGG U	235
GGTCTCCATGCGACGTGCAG	41	HBVg_41	GGUCUCCAUGCGACGUGCA G	236
GGACTGAGGCCCACTCCCAT	42	HBVg_42	GGACUGAGGCCCACUCCCA U	237
CACAGAGTCTAGACTCGTGG	43	HBVg_43	CACAGAGUCUAGACUCGUG G	238
GAAGAACCAACAAGAAGAT G	44	HBVg_44	GAAGAACCAACAAGAAGAU G	239
ACACGGTCCGGCAGATGAGA	45	HBVg_45	ACACGGUCCGGCAGAUGAG A	240
GACATGAACATGAGATGATT	46	HBVg_46	GACAUGAACAUGAGAUGAU U	241
GCAGCACAGCCTAGCAGCCA	47	HBVg_47	GCAGCACAGCCUAGCAGCC A	242
TCCTGGAATTAGAGGACAAA	48	HBVg_48	UCCUGGAAUUAGAGGACAA A	243
GTCTTACATAAGAGGACTCT	49	HBVg_49	GUCUUACAUAAGAGGACUC U	244
TTGTGGGTCACCATATTCTT	50	HBVg_50	UUGUGGGUCACCAUAUUCU U	245
CGCAAAATACCTATGGGAGT	51	HBVg_51	CGCAAAAUACCUAUGGGAG U	246
GGGTTGCGTCAGCAAACACT	52	HBVg_52	GGGUUGCGUCAGCAAACAC U	247
AGCTCTTGTTCCCAAGAATA	53	HBVg_53	AGCUCUUGUUCCCAAGAAU A	248
TGACATACTTTCCAATCAAT	54	HBVg_54	UGACAUACUUUCCAAUCAA U	249

[1180]

CAGATGAGAAGGCACAGACG	55	HBVg_55	G CAGAUGAGAAGGCACAGAC	250
CCCCGCCTGTAACACGAGCA	56	HBVg_56	A CCCCGCCUGUAACACGAGC	251
GGGTGGAGCCCTCAGGCTCA	57	HBVg_57	A GGGUGGAGCCCUCAGGCUC	252
ATTCCTTGGACTCATAAGGT	58	HBVg_58	U AUUCCUUGGACUCAUAAGG	253
TTTGTGGGTCACCATATTCT	59	HBVg_59	U UUUGUGGGUCACCAUAUUC	254
GTGAAAAAGTTGCATGGTGC	60	HBVg_60	C GUGAAAAAGUUGCAUGGUG	255
CCTGAACTGGAGCCACCAGC	61	HBVg_61	C CCUGAACUGGAGCCACCAG	256
TCCTCTGCCGATCCATACTG	62	HBVg_62	G UCCUCUGCCGAUCCAUACU	257
CGGCTAGGAGTTCCGCAGTA	63	HBVg_63	A CGGCUAGGAGUUCCGCAGU	258
AATGTCAACGACCGACCTTG	64	HBVg_64	G AAUGUCAACGACCGACCUU	259
GACCTTCGTCTGCGAGGCGA	65	HBVg_65	A GACCUUCGUCUGCGAGGCG	260
GTTGCCGGGCAACGGGGTAA	66	HBVg_66	A GUUGCCGGGCAACGGGGUA	261
GATTGAGACCTTCGTCTGCG	67	HBVg_67	G GAUUGAGACCUUCGUCUGC	262
AGGACCCCTGCTCGTGTTAC	68	HBVg_68	C AGGACCCCUGCUCGUGUUA	263
TTTGAAGTATGCCTCAAGGT	69	HBVg_69	U UUUGAAGUAUGCCUCAAGG	264
CCGCTTGTTTTGCTCGCAGC	70	HBVg_70	C CCGCUUGUUUUGCUCGCAG	265
TGCTAGGCTGTGCTGCCAAC	71	HBVg_71	C UGCUAGGCUGUGCUGCCAA	266
TGCCGATTGGTGGAGGCAGG	72	HBVg_72	G UGCCGAUUGGUGGAGGCAG	267
TCTTTGTACTAGGAGGCTGT	73	HBVg_73	U UCUUUGUACUAGGAGGCUG	268
CGTCCCGCGCAGGATCCAGT	74	HBVg_74	U CGUCCCGCGCAGGAUCCAG	269
AAAGCCCAAGATGATGGGAT	75	HBVg_75	U AAAGCCCAAGAUGAUGGGA	270
GCAGATGAGAAGGCACAGAC	76	HBVg_76	C GCAGAUGAGAAGGCACAGA	271
CGATTGGTGGAGGCAGGAGG	77	HBVg_77	G CGAUUGGUGGAGGCAGGAG	272
AGGAGGCTGTAGGCATAAAT	78	HBVg_78	U AGGAGGCUGUAGGCAUAAA	273
CCATGCCCCAAAGCCACCCA	79	HBVg_79	A CCAUGCCCCAAAGCCACCC	274
AGGTTGGGGACTGCGAATTT	80	HBVg_80	U AGGUUGGGGACUGCGAAUU	275
AGACCTTCGTCTGCGAGGCG	81	HBVg_81	G AGACCUUCGUCUGCGAGGC	276
CCTGGAATTAGAGGACAAAC	82	HBVg_82	C CCUGGAAUUAGAGGACAAA	277
TTTCAGTTATATGGATGATG	83	HBVg_83	G UUUCAGUUAUAUGGAUGAU	278

[1181]

GTAACACGAGCAGGGGTCCT	84	HBVg_84	GUAACACGAGCAGGGGUCCU	279
CATCTTCTTGTTGGTTCTTC	85	HBVg_85	CAUCUUCUUGUUGGUUCUUC	280
CGGGGAGACCGCGTAAAGAG	86	HBVg_86	CGGGGAGACCGCGUAAAGAG	281
CTAGACTCTGTGGTATTGTG	87	HBVg_87	CUAGACUCUGUGGUAUUGUG	282
CCCTGCTCGTGTTACAGGCG	88	HBVg_88	CCCUGCUCGUGUUACAGGCG	283
TACCACAGAGTCTAGACTCG	89	HBVg_89	UACCACAGAGUCUAGACUCG	284
TCGCAAAATACCTATGGGAG	90	HBVg_90	UCGCAAAAUACCUAUGGGAG	285
GTCTGTGCCTTCTCATCTGC	91	HBVg_91	GUCUGUGCCUUCUCAUCUGC	286
ACACGTAGCGCCTCATTTTG	92	HBVg_92	ACACGUAGCGCCUCAUUUUG	287
TTGGGGTTGAGGTCCCAATC	93	HBVg_93	UUGGGGUUGAGGUCCCAAUC	288
CCCCGAGACGGGTCGTCCGC	94	HBVg_94	CCCCGAGACGGGUCGUCCGC	289
CCTACGAACCACTGAACAAA	95	HBVg_95	CCUACGAACCACUGAACAAA	290
TTACATACTCTGTGGAAGGC	96	HBVg_96	UUACAUACUCUGUGGAAGGC	291
ACCTCCTTTCCATGGCTGCT	97	HBVg_97	ACCUCCUUUCCAUGGCUGCU	292
GTTATCGCTGGATGTGTCTG	98	HBVg_98	GUUAUCGCUGGAUGUGUCUG	293
AACATGAGATGATTAGGCAG	99	HBVg_99	AACAUGAGAUGAUUAGGCAG	294
ACTTCTCTCAATTTTCTAGG	100	HBVg_100	ACUUCUCUCAAUUUUCUAGG	295

표 E3. HBV 게놈 보존도가 79-90%이고 최대 2 개의 불일치를 갖는 후보 gRNA(HBV2)

표적 부위 (프로토스페이서) 서열	표적 SEQ ID	서열명	gRNA 스페이서 서열	RNA 스페이서 SEQ ID
CACTTTCTCGCCAACTTACA	101	HBVg_101	CACUUUCUCGCCAACUUACA	296
CATAAGGTGGGAAACTTTAC	102	HBVg_102	CAUAAGGUGGGAAACUUUAC	297
CCAAACCTCGAAAAGGCATG	103	HBVg_103	CCAAACCUCGAAAAGGCAUG	298
ATAGAAGGAAAGAAGTCAGA	104	HBVg_104	AUAGAAGGAAAGAAGUCAGA	299
GCTGCTCCTTTTACACAATG	105	HBVg_105	GCUGCUCCUUUUACACAAUG	300
GAAGCGAAGTGCACACGGTC	106	HBVg_106	GAAGCGAAGUGCACACGGUC	301
GGATCATCAACCACCAGCAC	107	HBVg_107	GGAUCAUCAACCACCAGCAC	302
GAGCCAAGAGAAACGGACTG	108	HBVg_108	GAGCCAAGAGAAACGGACUG	303
CTTCACCTCTGCACGTCGCA	109	HBVg_109	CUUCACCUCUGCACGUCGCA	304

[1182]

ACAATGTTCCGGAGACTCTA	110	HBVg_110	ACAAUGUUCCGGAGACUCUA	305
TCCGCGGGATTCAGCGCCGA	111	HBVg_111	UCCGCGGGAUUCAGCGCCGA	306
TTAATGAGTGGGAGGAGTTG	112	HBVg_112	UUAAUGAGUGGGAGGAGUUG	307
CCAACTCAAACAATCCAGAT	113	HBVg_113	CCAACUCAAACAAUCCAGAU	308
GCTGCCAACTGGATCCTGCG	114	HBVg_114	GCUGCCAACUGGAUCCUGCG	309
AGTCTTTGAAGTATGCCTCA	115	HBVg_115	AGUCUUUGAAGUAUGCCUCA	310
TCCTGACTGCCGATTGGTGG	116	HBVg_116	UCCUGACUGCCGAUUGGUGG	311
TCTTGTCCTCCAATTTGTCC	117	HBVg_117	UCUUGUCCUCCAAUUUGUCC	312
CGATAACCAGGACAAATTGG	118	HBVg_118	CGAUAACCAGGACAAAUUGG	313
ATTTGGAAGATCCAGCATCC	119	HBVg_119	AUUUGGAAGAUCCAGCAUCC	314
CTGTTTGGCTTTCAGTTATA	120	HBVg_120	CUGUUUGGCUUUCAGUUAUA	315
CTCCTCCTGCCTCCACCAAT	121	HBVg_121	CUCCUCCUGCCUCCACCAAU	316
GTCATCCTCAGGCCATGCAG	122	HBVg_122	GUCAUCCUCAGGCCAUGCAG	317
CTGCCGTTCCGGCCGACCAC	123	HBVg_123	CUGCCGUUCCGGCCGACCAC	318
CCTTCCTGACTGCCGATTGG	124	HBVg_124	CCUUCCUGACUGCCGAUUGG	319
ACCTGCACGACTCCTGCTCA	125	HBVg_125	ACCUGCACGACUCCUGCUCA	320
GGCCTGTATTTTCCTGCTGG	126	HBVg_126	GGCCUGUAUUUUCCUGCUGG	321
AACATAGAGGTTCCTTGAGC	127	HBVg_127	AACAUAGAGGUUCCUUGAGC	322
TGCCGTTCCGGCCGACCACG	128	HBVg_128	UGCCGUUCCGGCCGACCACG	323
ATAGGCCATCAGCGCATGCG	129	HBVg_129	AUAGGCCAUCAGCGCAUGCG	324
GCCTCCACCAATCGGCAGTC	130	HBVg_130	GCCUCCACCAAUCGGCAGUC	325
GTCCTTTGTTTACGTCCCGT	131	HBVg_131	GUCCUUUGUUUACGUCCCGU	326
TGGGAACAAGAGCTACAGCA	132	HBVg_132	UGGGAACAAGAGCUACAGCA	327
CTGTAAACAGGCCTATTGAT	133	HBVg_133	CUGUAAACAGGCCUAUUGAU	328
GTCGCAGAAGATCTCAATCT	134	HBVg_134	GUCGCAGAAGAUCUCAAUCU	329
CTGCCTTCCTGACTGCCGAT	135	HBVg_135	CUGCCUUCCUGACUGCCGAU	330
ACTACTAATTCCCTGGATGC	136	HBVg_136	ACUACUAAUUCCCUGGAUGC	331
CACATTTCTTGCCTTACTTT	137	HBVg_137	CACAUUUCUUGCCUUACUUU	332
GGATGACTGTCTCTTAGAGG	138	HBVg_138	GGAUGACUGUCUCUUAGAGG	333

[1183]

GCTATGCCTCATCTTCTTGT	139	HBVg_139	GCUAUGCCUCAUCUUCUU GU	334
CCCGTCGGCGCTGAATCCCG	140	HBVg_140	CCCGUCGGCGCUGAAUCCC G	335
TAGTATTCCTTGGACTCATA	141	HBVg_141	UAGUAUUCCUUGGACUCA UA	336
AGGTAGGAGCGGGAGCATTC	142	HBVg_142	AGGUAGGAGCGGGAGCAU UC	337
TCTTTTGGGGTGGAGCCCTC	143	HBVg_143	UCUUUUGGGGUGGAGCCC UC	338
TCAGTATGCCCTGAGCCTGA	144	HBVg_144	UCAGUAUGCCCUGAGCCU GA	339
TTTAATGAGTGGGAGGAGTT	145	HBVg_145	UUUAAUGAGUGGGAGGAG UU	340
CTCCCTCGCCTCGCAGACGA	146	HBVg_146	CUCCCUCGCCUCGCAGACG A	341
GATAAGATAGGGGCATTTGG	147	HBVg_147	GAUAAGAUAGGGGCAUUU GG	342
CCGCGGGATTCAGCGCCGAC	148	HBVg_148	CCGCGGGAUUCAGCGCCG AC	343
CAGCGATAACCAGGACAAAT	149	HBVg_149	CAGCGAUAACCAGGACAA AU	344
CAGGTAGGAGTGGGAGCATT	150	HBVg_150	CAGGUAGGAGUGGGAGCA UU	345
GTTTAATGAGTGGGAGGAGT	151	HBVg_151	GUUUAAUGAGUGGGAGGA GU	346
TGGTGAGTGATTGGAGGTTG	152	HBVg_152	UGGUGAGUGAUUGGAGGU UG	347
TAATGAGTGGGAGGAGTTGG	153	HBVg_153	UAAUGAGUGGGAGGAGUU GG	348
ACTACATGTTCTGGATAATA	154	HBVg_154	ACUACAUGUUCUGGAUAA UA	349
GGCATAGCAGCAGGATGAAG	155	HBVg_155	GGCAUAGCAGCAGGAUGA AG	350
GTTGATAAGATAGGGGCATT	156	HBVg_156	GUUGAUAAGAUAGGGGCA UU	351
TCAACGAATTGTGGGTCTTT	157	HBVg_157	UCAACGAAUUGUGGGUCU UU	352
TATGGATGATGTGGTATTGG	158	HBVg_158	UAUGGAUGAUGUGGUAUU GG	353
CAACGAATTGTGGGTCTTTT	159	HBVg_159	CAACGAAUUGUGGGUCUU UU	354
CATTTGTTCAGTGGTTCGTA	160	HBVg_160	CAUUUGUUCAGUGGUUCG UA	355
CGTCTAACAACAGTAGTTTC	161	HBVg_161	CGUCUAACAACAGUAGUU UC	356
TGCCTGAGTGCTGTATGGTG	162	HBVg_162	UGCCUGAGUGCUGUAUGG UG	357
GCCCCGAGACGGGTCGTCCG	163	HBVg_163	GCCCCGAGACGGGUCGUC CG	358
GACTGCCGATTGGTGGAGGC	164	HBVg_164	GACUGCCGAUUGGUGGAG GC	359
TATATGGATGATGTGGTATT	165	HBVg_165	UAUAUGGAUGAUGUGGUA UU	360
CTTGAGTATTTGGTGTCTTT	166	HBVg_166	CUUGAGUAUUUGGUGUCU UU	361
GATCTGGTGGGCGTTCACGG	167	HBVg_167	GAUCUGGUGGGCGUUCAC GG	362

[1184]

AGACTGGGAGGAGTTGGGGG	168	HBVg_168	AGACUGGGAGGAGUUGGGGG	363
AGTCCTCTTATGTAAGACCT	169	HBVg_169	AGUCCUCUUAUGUAAGACCU	364
CTCAAGATGTTGTACAGACT	170	HBVg_170	CUCAAGAUGUUGUACAGACU	365
GGGAACAAGAGCTACAGCAT	171	HBVg_171	GGGAACAAGAGCUACAGCAU	366
TCGCAGAAGATCTCAATCTC	172	HBVg_172	UCGCAGAAGAUCUCAAUCUC	367
GGGGTGGAGCCCTCAGGCTC	173	HBVg_173	GGGGUGGAGCCCUCAGGCUC	368
TATTCCTTGGACTCATAAGG	174	HBVg_174	UAUUCCUUGGACUCAUAAGG	369
TCTAAGAGACAGTCATCCTC	175	HBVg_175	UCUAAGAGACAGUCAUCCUC	370
CAACTCAAACAATCCAGATT	176	HBVg_176	CAACUCAAACAAUCCAGAUU	371
AAACAAAGGACGTCCCGCGC	177	HBVg_177	AAACAAAGGACGUCCCGCGC	372
CTGCCAACTGGATCCTGCGC	178	HBVg_178	CUGCCAACUGGAUCCUGCGC	373
GAAGCTCCAAATTCTTTATA	179	HBVg_179	GAAGCUCCAAAUUCUUUAUA	374
GTCAGTATGCCCTGAGCCTG	180	HBVg_180	GUCAGUAUGCCCUGAGCCUG	375
TTTCCCACCTTATGAGTCCA	181	HBVg_181	UUUCCCACCUUAUGAGUCCA	376
ACAAGAGGTTGGTGAGTGAT	182	HBVg_182	ACAAGAGGUUGGUGAGUGAU	377
GAAAGCCCAAGATGATGGGA	183	HBVg_183	GAAAGCCCAAGAUGAUGGGA	378
AGGTTCCACGCATGCGCTGA	184	HBVg_184	AGGUUCCACGCAUGCGCUGA	379
TTGGTGAGTGATTGGAGGTT	185	HBVg_185	UUGGUGAGUGAUUGGAGGUU	380
AGAGCTACAGCATGGGAGGT	186	HBVg_186	AGAGCUACAGCAUGGGAGGU	381
ATATGGATGATGTGGTATTG	187	HBVg_187	AUAUGGAUGAUGUGGUAUUG	382
CCATTTGTTCAGTGGTTCGT	188	HBVg_188	CCAUUUGUUCAGUGGUUCGU	383
TTATATGGATGATGTGGTAT	189	HBVg_189	UUAUAUGGAUGAUGUGGUAU	384
GGAGTGGGAGCATTCGGGCC	190	HBVg_190	GGAGUGGGAGCAUUCGGGCC	385
ATTTGGTGTCTTTTGGAGTG	191	HBVg_191	AUUUGGUGUCUUUUGGAGUG	386
CACAGAAAGGCCTTGTAAGT	192	HBVg_192	CACAGAAAGGCCUUGUAAGU	387
ACCAATTTTCTTTTGTCTTT	193	HBVg_193	ACCAAUUUUCUUUUGUCUUU	388
AGGTTAATGGTCTTTGTACT	194	HBVg_194	AGGUUAAUGGUCUUUGUACU	389
TACCAATTTTCTTTTGTCTT	195	HBVg_195	UACCAAUUUUCUUUUGUCUU	390

[1185]

[1186] **실시예 2: 바이러스 복제 및 전사와 관련된 유전자 및 이의 조절 요소의 표적화된 전사 억제를 위한 gRNA 식별.**

[1187] [0572] gRNA는 통합된 HBV DNA와 공유적으로 폐쇄된 원형 DNA(cccDNA) 또는 이완된 원형 DNA(rcDNA)에 존재하는 HBV 게놈 내의 표적 유전자의 전사 억제를 용이하게 하는 것으로 확인되었으며, 이는 gRNA와 dCas9-효과기 융합 단백질을 포함하는 DNA-표적화 시스템의 일부이다. 도 1 및 4는 HBV 게놈 전체에서 각 gRNA의 중간 표적화 위치를 나타낸다.

[1188] ***I) 통합 HBV 게놈의 조절자를 위한 일시적 배열 스크리닝***

[1189] A. dCas9-KRAB를 이용한 단기간(5일) 스크리닝

[1190] [0573] 표 E1-E3 및 표 6에 나열된 것들을 포함하여 HBV 전사와 관련된 유전자 및 이의 조절 요소를 표적으로 하는 가이드 RNA를, 통합된 서열(cccDNA 없음)에서만 HBV를 발현하는 Hep3B 세포에 형질도입하고, gRNA 표적 유전자의 전사 억제를 위해 dSpCas9-KRAB 융합 단백질(서열번호 594)을 인코딩하는 mRNA와 함께 형질도입하였다.

[1191] [0574] 불멸화된 HBV 감염 인간 간세포 세포주(Hep3B, PLC/PRF/5), HBV 게놈의 형질도입 카피를 포함하는 조작된 간세포 세포주(HepG2.2.15, HepAD38), 또는 HBV 진입에 필요한 동종 표면 수용체를 과발현하는 조작된 불멸화된 간세포주(HepG2.NTCP)와 같은 다른 세포주를 하나 이상의 gRNA 표적 유전자의 동시 전사 억제를 위한 dSpCas9-효과기 융합 단백질과 함께 고려할 수 있다. 또한 HBV 복제 및 전사 억제를 증가시키기 위해 동일하거나 다른 HBV 바이러스 유전자를 표적으로 하는 표 E1-E3 및 표 6의 여러 예시적 gRNA의 조합을 다중화된 접근

방식으로 전달할 수 있다.

[1192] [0575] 형질도입된 세포를 5일차에 HBsAg 전사체 또는 총 HBV RNA 전사체 수준의 qRT-PCR 측정을 통해 표적 HBV 전사체의 전사 억제에 대해 평가하였다(도 1 및 3). 형질도입된 세포는 통합 및 cccDNA에서 B형 간염 표면 항원(HBsAg) 수치 감소와 cccDNA에서만 B형 간염 바이러스 코어 관련 항원 단백질(HBcrAg) 수치 감소를 위해 ELISA를 통해 전사 억제를 평가할 수도 있다. HBsAg를 90%, HBeAg를 50% 까지 감소시키는 억제를 매개하는 gRNA를 식별하여 추가 분석을 수행할 수 있다.

[1193] [0576] 도 1은 dCas9-KRAB과 함께 전달 후 Hep3B 세포에서 각 gRNA에 대한 HBV RNA의 로그 배수 변화(log fold change)를 나타낸다. x축은 HBV 게놈에서 각 gRNA의 중간 표적화 위치를 나타낸다. 도 1에 도시된 바와 같이, 192개 gRNA 중 94개의 표적화된 부위(화살표로 표시된 "모든 억제")는 총 HBV RNA의 억제를 보였다. 식별된 "hit" gRNA는 폴리머라제, S, 짧은 X 또는 pre-코어 유전자를 표적화하였다. 192개 gRNA 중 48개가 표적화한 부위(화살표로 표시된 ">75% 억제")는 총 HBV RNA의 75% 이상 억제를 보였다. 서열번호 565 (HBVg_175), 528 (HBVg_138), 582 (HBVg_192), 542 (HBVg_152), 508 (HBVg_118), 515 (HBVg_125), 575 (HBVg_185), 453 (HBVg_63), 506 (HBVg_116), 514 (HBVg_124), 395 (HBVg_5), 472 (HBVg_82), 451 (HBVg_61), 488 (HBVg_98), 540 (HBVg_150), 533 (HBVg_143), 572 (HBVg_182), 566 (HBVg_176), 489 (HBVg_99), 469 (HBVg_79), 408 (HBVg_18), 465 (HBVg_75), 402 (HBVg_12), 474 (HBVg_84), 525 (HBVg_135), 416 (HBVg_26), 369 (HBVg_6), 554 (HBVg_164), 419 (HBVg_29), 545 (HBVg_155), 446 (HBVg_56), 580 (HBVg_190), 555 (HBVg_165), 412 (HBVg_22), 428 (HBVg_38), 458 (HBVg_68), 548 (HBVg_158), 511 (HBVg_121), 432 (HBVg_42), 441 (HBVg_51), 433 (HBVg_43), 579 (HBVg_189), 479 (HBVg_89), 478 (HBVg_88), 520 (HBVg_130), 462 (HBVg_72), 523 (HBVg_133), 및 503 (HBVg_113)에 제시된 가이드 RNA는 총 HBV RNA의 75% 이상 억제를 매개하였다. 또한 192개 gRNA 중 10개가 표적하는 부위(화살표로 표시한 ">90% 억제")는 총 HBV RNA의 90% 이상 억제를 보였다(최대 98.5%).

[1194] [0577] 구체적으로, M/S-HBs 프로모터, Enh1/X-프로모터, Enh2/Basal 코어 프로모터, L-HBs 프로모터 또는 M/S-유전자 프로모터 영역을 표적화하는 gRNA를 사용하자 총 HBV RNA의 90% 이상 억제가 관찰되었다. 서열번호 565, 528, 582, 542, 508, 515, 575, 453, 506, 514, 425 또는 472에 제시된 가이드 RNA는 총 HBV RNA의 90% 이상 억제를 매개하였다.

[1195] [0578] 총 HBV RNA의 가장 높은 전사 억제를 나타내는 gRNA(서열번호 528(HBVg_138), 582(HBVg_192), 565(HBVg_175), 453(HBVg_63))는 L-HBs 프로모터 및 Enh1/X-프로모터 영역(HBV 중간 위치: HBV 게놈의 3197bp(HBVg_138), 1124bp(HBVg_192), 3179bp(HBVg_175), 1288bp(HBVg_63))을 표적으로 삼았다. B형 간염 바이러스 게놈(B형 간염 바이러스 아형 ayw, 완전 게놈, GenBank: U95551.1), 서열번호 650)을 참조하는 위치에 해당한다 (도 3A). 중간 위치는 모든 HBV 유전자형에 걸친 평균 위치를 포함한다. 예를 들어, 중간 위치는 서열번호 650, 678 및 679에 제시된 서열을 포함하는 HBV 게놈의 평균 위치를 포함한다. HBsAg RNA의 가장 높은 전사 억제를 갖는 gRNA(서열번호 528(HBVg_138), 582(HBVg_192), 453(HBVg_63), 및 489(HBVg_99))는 B형 간염 바이러스 게놈(B형 간염 바이러스 아형 ayw, 완전 게놈, GenBank: U95551.1), 서열번호 650)을을 참조하는 위치에 상응하는 HBV 게놈의 Enh1/X-프로모터, Enh2/기저 코어 프로모터 및 L-HBs 프로모터 영역(HBV 중간 위치: 3197bp(HBVg_138), 1124bp(HBVg_192), 1288bp(HBVg_63,), 1850bp(HBVg_99), 및 1904bp(HBVg_79)를 표적으로 삼았다(도 3B).

[1196] B. DNMT3A/L-dSpCas9-KRAB를 이용한 장기 치료적 관련성 스크리닝

[1197] [0579] 표 E1-3 및 표 6에 나열된 것을 포함하여 HBV 전사와 관련된 유전자 및 그 조절 요소를 표적으로 하는 가이드 RNA를 통합된 서열(cccDNA 없음)에서만 HBV를 발현하는 Hep3B 세포에 형질도입하고, gRNA 표적 유전자의 전사 억제를 위해 DNMT3A/L-dSpCas9-KRAB 융합 단백질(서열번호 644)을 인코딩하는 mRNA와 함께 형질도입하였다.

[1198] [0580] 불멸화된 HBV 감염 인간 간세포 세포주(Hep3B, PLC/PRF/5), HBV 게놈의 형질도입 카피를 포함하는 조작된 간세포 세포주(HepG2.2.15, HepAD38), 또는 HBV 진입에 필요한 인지 표면 수용체를 과발현하는 조작된 불멸화된 간세포주(HepG2.NTCP)와 같은 다른 세포주를 하나 이상의 gRNA 표적 유전자의 동시 전사 억제를 위한 dSpCas9-효과기 융합 단백질을 인코딩하는 mRNA와 함께 고려할 수 있다. 또한, HBV 복제 및 전사 억제를 증가시키기 위해 동일하거나 다른 HBV 바이러스 유전자를 표적으로 하는 표 E1 및 표 6의 여러 모범적 gRNA의 조합을 다중화된 접근 방식으로 전달할 수 있다.

[1199] [0581] 형질도입된 세포를 qRT-PCR을 통한 총 HBV RNA 전사체 수치 측정을 통해 표적 HBV 전사체의 전사 억제에 대해 41일과 61일에 평가하였다(도 4 및 5). 형질도입된 세포는 통합 및 cccDNA에서 B형 간염 표면 항원(HBsAg) 수치 감소와 cccDNA에서만 B형 간염 바이러스 코어 관련 항원 단백질(HBcrAg) 수치 감소를 위한 ELISA를 통해 전사 억제에 대해 평가할 수도 있다. 총 HBV RNA 전사체 수치를 최소 80% 감소시키는 억제를 매개하는 gRNA는 추가 분석을 위해 식별되었다.

[1200] [0582] 도 4는 61일차에 각 gRNA에 대한 HBV RNA의 로그 배수 변화를 나타낸다. x축은 HBV 게놈 전체에서 각 gRNA의 중간 표적화 위치를 나타낸다. 확인된 gRNA는 폴리머라제, S, 짧은 X 또는 pre-코어 유전자를 표적화하였다. 구체적으로, pre-S1 프로모터, pre-S2 프로모터 및 폴리머라제 유전자를 표적화하는 gRNA에서 총 HBV RNA의 80% 이상 억제가 관찰되었다.

[1201] [0583] 총 HBV RNA의 전사 억제가 가장 높은 gRNA(HBVg_142; 서열번호 532, HBVg_138; 서열번호 528, HBVg_185: 서열번호 575, HBVg_26: 서열번호 416, 및 HBVg_152; 서열번호 542)는 HBV 게놈의 pre-S1 프로모터, pre-S2 프로모터, CpG 섬 1, 및 폴리머라제 유전자(HBV 게놈의 중간 위치: 3016bp(HBVg_142), 3197bp(HBVg_138), 341bp(HBVg_185), 55bp(HBVg_26), 및 340bp(HBVg_152)를 표적으로 삼았는데, 이는 B형 간염 바이러스 게놈(B형 간염 바이러스 아형 ayw, 완전 게놈, GenBank: U95551.1), 서열번호 650)을 참조하는 위치에 해당한다(도 5). 놀랍게도, 최소 80%의 전사 억제가 형질도입 후 41일 및 61일까지 지속되었는데, dCas9 융합 단백질이 이러한 늦은 시점에 더 이상 발현되지 않기 때문에 이는 후성적 편집의 효능이 오래 지속됨을 가리킨다.

[1202] [0584] 형질도입된 세포는 또한 총 HBV RNA 전사체 수치의 qRT-PCR 측정을 통해 더 긴 시간 지점에서 표적 HBV 전사체의 전사 억제에 대해 평가되었다(도 6A-6B). 서열번호 453(HBVg_63), 575(HBVg_185), 446(HBVg_56), 412(HBVg_22), 및 416(HBVg_26)에 제시된 여러 gRNA는 120회 이상의 세포 배가 및 248일 후 통합된 HBV RNA의 치료적으로 실행 가능한 억제 수치를 보였다. 단일 gRNA 투여 후 총 HBV RNA의 안정적인 억제가 달성되었다.

[1203] [0585] 연구 222일차의 형질도입된 세포는 동일한 gRNA 또는 비표적화 gRNA와 DNMT3A/L-dSpCas9-KRAB 융합 단백질(서열번호 644)을 인코딩하는 mRNA로 리포펙타민 MessengerMax를 사용하여 재형질도입되었다. 총 HBV RNA 수치는 형질도입 4일 후에 검사하여 재투여 시 억제 개선을 평가하였다. 가이드 RNA HBVg_22(서열번호 412), HBVg26(서열번호 416), HBVg_185(서열번호 575), HBVg_63(서열번호 453)은 1차 투여 후 상당한 억제를 보였고 재투여 후 HBV RNA 수치가 더욱 현저하게 감소하였다(도 7). 효과는 주어진 투여량에 대한 완전한 억제에 도달하면서 지속되었다. 전달이 보정되자 거의 모든 세포(~ 96%)에서 거의 완전한 억제(95% 벌크 RNA)가 달성되었다. 비표적 gRNA는 미처리 대조군 세포와 비교했을 때 억제 수준에 유의한 영향을 미치지 않았다. 더 낮은 기준 전달 수준에서 재투여를 평가하기 위한 추가 연구가 수행될 수 있다. 일차 인간 간세포 및 마우스 모델에서의 전달 및 재투여가 최적화될 수 있다.

[1204] ***II) 실제 감염 모델에서 에피솜 cccDNA 조절자에 대한 일시적 배열 스크리닝***

[1205] [0586] HBV를 100 GE/세포로 투여한 HepG2.NTCP 세포를 HBV 감염 3일 후에 대조군 또는 표준 치료(예컨대 엔테카비르(ETV)와 같은 뉴클레오사이드 유사체)로 처리하였다. HepG2.NTCP 세포는 네이티브 cccDNA의 생성 및 존재로 인해 HBV 감염이 지속되므로 HBV의 진정한 감염 모델을 가능하게 한다. HBV 감염 6일 후에 표 E1-E3 및 표 6에 나열된 것을 포함하여 HBV 전사와 관련된 유전자 및 조절 요소를 표적으로 하는 gRNA를 리포펙션을 통해 cccDNA에서 HBV를 발현하는 세포에 형질도입시키고 DNMT3A/L-dSpCas9-KRAB 융합 단백질(서열번호 644)을 인코딩하는 mRNA와 함께 형질도입시켜 gRNA 표적 HBV 유전자의 전사를 억제하였다.

[1206] [0587] 형질도입된 세포를 수확하여 5일차에 표적 HBV 전사체(총 HBV RNA 또는 3.5kB HBV RNA)의 전사 억제를 평가하였다. 3.5kb 총 HBV RNA 전사체 수준 측정은 qRT-PCR을 통해 수행되었으며, 대표적인 실행 하나를 나타냈다(도 8). 형질도입된 세포는 또한 ELISA를 통해 cccDNA에서 B형 간염 표면 항원(HBsAg) 수준과 B형 간염 바이러스 코어 관련 항원 단백질(HBcrAg) 수준이 감소했는지 전사 억제를 평가할 수 있다. 억제를 매개하여 3.5kb 총 HBV RNA 전사체 수준을 최소 80% 감소시킨 gRNA가 확인되었다.

[1207] [0588] 도 8은 스크린의 한 복제본에서 각 gRNA에 대한 3.5kb 총 HBV RNA의 로그 배수 변화를 나타낸다. x축은 HBV 게놈 전체에서 각 gRNA의 중간 표적화 위치를 나타낸다.

[1208] 최대 수준의 3.5 kb 총 HBV RNA 전사 억제를 보이는 gRNA(HBVg_45: 서열번호 435, HBVg_22: 서열번호 412, HBVg_456: 서열번호 66, HBVg_6: 서열번호 396, HBVg_34: 서열번호 424, HBVg_182: 서열번호 572)는 Enhancer 1/X 프로모터 영역 및 Enhancer II 조절 요소(HBV 중간 위치: 1580 bp (HBVg_45) (97.5%)), 폴리머라제 유전자

및 CpG 섬 2(HBV 중간 위치: 1283 bp (HBVg_22) (93.0%) 및 1162 bp (HBVg_66) (90.4%)), 그리고 폴리머라제 유전자 및 CpG 섬 1(HBV 중간 위치: 261 bp (HBVg_6 및/또는 HBVg_34) (89.8%) 및 349 bp (HBVg_182) (89.7%))을 표적으로 하였다.

[1209] [0589] cccDNA 스크린(rep #3)과 다른 cccDNA 스크린(rep #2)의 상관 관계 플롯은 감염 간 일관성을 보여주었다(도 9). 통합된 HBV, cccDNA 또는 통합된 것과 cccDNA를 모두 억제하는 gRNA가 식별되었다(도 10). 마지막으로, 단일 gRNA에서 다중화된 HBV 억제를 달성하는 gRNA가 식별되었다. 도 11은 여러 유전자와 조절 요소(예컨대 HBs/CpG 섬 1, HBx/Enh1 인핸서/CpG 섬 2)의 억제를 달성한 단일 gRNA가 표적으로 삼는 서열을 나열한다. 억제는 통합된 HBV DNA와 cccDNA HBV 게놈 내에서 높은 결합 효율로 달성되었다. 도 12는 여러 HBV 전사체(통합 및 cccDNA HBV 저장소 모두)에 걸쳐 지속적인 억제를 달성하는 gRNA(HBVg_22; 서열번호 412, HBVg_63; 서열번호 453, HBVg_185; 서열번호 575, HBVg_99: 서열번호 489)가 표적으로 삼는 서열을 나열한다. 억제는 mRNA/gRNA의 단일 일시적 전달 후에 효과적이다. 이러한 단일 gRNA는 다중화된 접근 방식을 위해 조합하여 사용될 수 있다.

[1210] [0590] 형질도입된 세포도 수확하여 이후 시간 지점인 14일에 표적 HBV 전사체의 전사 억제를 평가하였다. 3.5kb 총 HBV RNA 전사체 수치 측정은 qRT-PCR을 통해 수행하였다(도 13). 또한 형질도입된 세포에 대해 ELISA를 통해 cccDNA에서 B형 간염 표면 항원(HBsAg) 수치와 B형 간염 바이러스 코어 관련 항원 단백질(HbcrAg) 수치가 감소했는지 전사 억제를 평가할 수 있다. 도 13은 각 gRNA에 대한 3.5kb HBV RNA의 로그 배수 변화를 나타낸다. x축은 HBV 게놈 전체에서 각 gRNA의 중간 표적 위치를 나타낸다. 3.5kb HBV cccDNA의 80% 이상, 최대 97.2%를 억제하는 여러 gRNA(28/189)가 확인되었다. 식별된 gRNA는 HBVg_73(서열번호 463), HBVg_192(서열번호 582), HBVg_18(서열번호 408), HBVg_63(서열번호 453), HBVg_22(서열번호 412), HBVg_185(서열번호 575), HBVg_46(서열번호 436), HBVg_99(서열번호 489), HBVg_162(서열번호 552), HBVg_174(서열번호 568), HBVg_20(서열번호 410), HBVg_79(서열번호 469), HBVg_50(서열번호 440), HBVg_79(서열번호 469), HBVg_50(440), HBVg_27(서열번호 417), HBVg_182(서열번호 572), 및 HBVg_12(서열번호 402)이었다.

[1211] [0591] 도 14는 단일 gRNA HBVg_73(서열번호 463), HBVg_192(서열번호 582), HBVg_18(서열번호 408), HBVg_63(서열번호 456), HBVg_22(서열번호 412), HBVg_185(서열번호 575), HBVg_46(서열번호 436), HBVg_99(서열번호 489), HBVg_162(서열번호 552), HBVg_174(서열번호 564), HBVg_5(서열번호 410), HBVg_79(서열번호 469), HBVg_50(서열번호 440), HBVg_27(서열번호 417), HBVg_66(서열번호 456), 및 HBVg_182(서열번호 572)에 의해 표적 HBV cccDNA 전사물이 억제되는 것을 보여준다. gRNA가 표적으로 삼은 부위는 또한 HBV 유전자형에서 92% 보존되었다. 가이드 RNA HBVg_73(서열번호 463)(93%), HBVg_192(서열번호 582)(91.9%), 및 HBVg_18(서열번호 408)(91.8%)가 가장 높은 억제를 매개하였다.

[1212] **실시예 3: 1차 인간 간세포(PHH) 감염 모델에서 달성된 cccDNA의 강력한 억제**

[1213] [0592] 보다 자연스러운 맥락에서 gRNA 매개 에피-편집을 평가하기 위해, 1차 인간 간세포(PHH) 감염 모델이 사용되었다. 두 기증자(PHH 기증자 1 및 PHH 기증자 2)의 1차 인간 간세포가 이 실험에 사용되었으며, 두 기증자 모두 100 GE/세포의 HBV에 감염되었고 배양체에서 5일 동안 감염이 발전하도록 두었다. 그런 다음 세포에 DNMT3A/L-dSpCas9-KRAB(서열번호 644) 및 gRNA HBVg_22(서열번호 412)를 인코딩하는 mRNA를 형질도입시키고 RNA 분리 및 qPCR을 통해 억제를 평가하였다. 도 15A는 7일차와 13/14일차에 두 기증자 모두에서 HBV RNA가 억제되었음을 보여준다. PHH 기증자 2는 동일한 용량에서 PHH 기증자 1에 비해 기준선에서 약 5배 더 높은 수준의 감염을 지원하였다. 전달 효율성에 큰 차이가 있었음에도 불구하고 두 PHH 기증자 모두 HBVg_22(서열번호 412)에 대해 강력하고 지속적인 cccDNA 억제를 보였다.

[1214] [0593] 기증자 1의 1차 인간 간세포를 콜라겐 코팅된 24웰 플레이트 두 개에 $0.2x10^6$개 세포/cm^2로 접종하였다. 접종 24시간 후, 세포를 100GE/세포 및 200GE/세포의 HBV로 감염시키고 RNA 분리 및 3.5kb HBV 전사체의 qPCR을 통해 감염 여부를 평가하였다. 그런 다음 세포를 DNMT3A/L-dSpCas9-KRAB(서열번호 644) 및 gRNA(서열번호 412)를 인코딩하는 mRNA로 형질도입시키고 6일 후 RNA 분리 및 qPCR을 통해 억제 여부를 평가하였다. 도 15B는 저용량 HBV(기저부 1 x 3.5kb HBV 전사체) 및 고용량 HBV(기저부 12 x 3.5kb 전사체)로 감염된 PHH 감염 모델에서 강력한 억제를 보여준다. 결과는 저용량 및 고용량 그룹 모두에서 상당한 억제가 달성되었음을 보여준다. 억제는 HBV RNA 발현의 넓은 창(두 용량 간에 12배)에서 유지되었다.

[1215] **실시예 4: dSpCas9-KRAB 및 DNMT3A/L-dSpCas9-KRAB 융합 단백질에 의해 유도된 억제의 비교**

[1216] [0594] dSpCas9-KRAB 및 DNMT3A/L-dSpCas9-KRAB 융합 단백질에 의해 유도된 억제를 비교하기 위해, HepG2.NTCP 세포와 PxB PHH 세포에 HBVg_22(서열번호 412)를 인코딩하는 mRNA를 dSpCas9-KRAB 단독(서열번호 595) 또는

DNMT3A/L-dSpCas9-KRAB("D3AL-K"; 서열번호 645)와 함께 형질도입하였다. 세포를 RNA 분리 및 qPCR을 통해 20일 또는 14일 후에 억제에 대해 평가하였다. 도 16A-16B는 HepG2.NTCP 및 PxB PHH 모델에서 HBVg_22 (서열번호 412)가 dSpCas9-KRAB 단독(서열번호 595) 또는 DNMT3A/L-dSpCas9-KRAB("D3AL-K"; 서열번호 645) 융합 단백질과 함께 사용되었을 때 유도된 억제 효과를 비교하여 보여준다. dCas9-KRAB만을 일시적으로 전달하면 여러 세포 맥락에서 cccDNA의 장기 억제가 발생할 수 있다. 그러나 dCas9-KRAB만을 일시적으로 전달하면 통합된 HBV DNA의 장기 억제가 가능하지 않다. 이는 메틸화(예컨대 D3AL-K에 의한)가 HBV 통합체로부터 HBV RNA의 지속적인 전사 억제에 필요하며, DNMTR3A/3L 도메인 없이도 어느 정도의 지속적인 cccDNA 침묵이 달성될 수 있음을 보여준다.

[1217] **실시예 5: HBV 복제 및 발현을 억제하기 위한 표현형 촉진을 위한 다중화된 표적 전사 억제**

[1218] [0595] 실시예 2에서 확인된 gRNA의 다양한 조합을 포함하는 DNA-표적화 시스템을 선별하여 적어도 두 개의 유전자 또는 이의 조절 요소에 대한 다중화된 표적 전사 억제가 HBV 복제 및 발현을 억제하기 위한 표현형을 촉진하는 조건을 식별하였다.

[1219] [0596] gRNA와 dSpCas9-효과기를 인코딩하는 mRNA로 구성된 DNA-표적화 시스템은 통합된 서열과 cccDNA에서 HBV를 발현하는 세포주에 일시적으로 형질도입된다. gRNA는 개별적으로 또는 2개의 gRNA, 3개의 gRNA 또는 4개의 gRNA의 조합으로 형질도입되며, 각 gRNA는 다른 유전자 또는 조절 요소를 표적으로 한다. 또는 gRNA는 개별적으로 또는 2개의 gRNA, 3개의 gRNA 또는 4개의 gRNA의 조합으로 형질도입되며, 각 gRNA는 동일한 유전자 또는 조절 요소 내의 다른 표적 부위를 표적으로 한다.

[1220] [0597] 일례로, 2개의 gRNA의 조합은 Enh1/X-프로모터, Enh2/Basal 코어 프로모터, L-HBs 프로모터, M/S-HBs 프로모터 중 하나를 표적화하는 제1 gRNA와 Enh1/X-프로모터, Enh2/Basal 코어 프로모터, L-HBs 프로모터, M/S-HBs 프로모터 중 하나를 표적화하는 제2 gRNA를 포함하며, 여기서 제1 및 제2 gRNA는 서로 다른 조절 요소를 표적화한다. 다른 예에서, 2개의 gRNA의 조합은 Enh1/X-프로모터를 표적화하는 제1 gRNA와 Enh2/Basal 코어 프로모터, L-HBs 프로모터, M/S-HBs 프로모터 중 하나를 표적화하는 제2 gRNA를 포함한다. 또 다른 다른 예에서, 2개의 gRNA의 조합은 Enh1/X-프로모터를 표적화하는 제1 gRNA와 L-HBs 프로모터를 표적화하는 제2 gRNA를 포함한다.

[1221] [0598] 다른 예에서, 3개의 gRNA의 조합은 Enh1/X-프로모터, Enh2/Basal 코어 프로모터, L-HBs 프로모터 및 M/S-HBs 프로모터 중 하나를 표적화하는 제1 gRNA, Enh1/X-프로모터, Enh2/Basal 코어 프로모터, L-HBs 프로모터 및 M/S-HBs 프로모터 중 하나를 표적화하는 제2 gRNA, Enh1/X-프로모터, Enh2/Basal 코어 프로모터, L-HBs 프로모터 중 하나를 표적화하는 제3 gRNA를 포함하며, 여기서 제2 및 제3 gRNA는 각각 다른 조절 요소를 표적화한다. 다른 예에서, 3개의 gRNA의 조합은 Enh1/X-프로모터를 표적화하는 제1 gRNA, Enh2/Basal 코어 프로모터를 표적화하는 제2 gRNA, L-HBs 프로모터를 표적화하는 제3 gRNA를 포함한다.

[1222] [0599] 또 다른 예에서, 4개의 gRNA의 조합은 Enh1/X-프로모터, Enh2/Basal 코어 프로모터, L-HBs 프로모터 및 M/S-HBs 프로모터 중 하나를 표적으로 하는 제1 gRNA, Enh1/X-프로모터, Enh2/Basal 코어 프로모터, L-HBs 프로모터 및 M/S-HBs 프로모터 중 하나를 표적으로 하는 제2 gRNA, Enh1/X-프로모터, Enh2/Basal 코어 프로모터, L-HBs 프로모터 중 하나를 표적으로 하는 제3 gRNA, Enh1/X-프로모터, Enh2/Basal 코어 프로모터, L-HBs 프로모터 중 하나를 표적으로 하는 제4 gRNA를 포함하며, 여기서 제1, 제2, 제3 및 제4 gRNA는 다른 조절 요소를 표적으로 한다.

[1223] [0600] 또 다른 예에서, 3개의 gRNA의 조합은 L-HBs 프로모터 중 하나와 M/S-HBs 프로모터를 표적으로 하는 제1 gRNA, Enh2/Basal 코어 프로모터를 표적으로 하는 제2 gRNA, Enh1/X-프로모터를 표적으로 하는 제3 gRNA를 포함한다.

[1224] [0601] 또 다른 예로, 4개의 gRNA의 조합은 L-HBs와 M/S-HBs 프로모터를 각각 표적화하는 gRNA와 Enh2/Basal 코어 프로모터를 표적화하는 gRNA를 포함한다.

[1225] [0602] 형질도입 후, 세포는 HBV 복제 및 발현을 침묵시키기 위한 표현형을 촉진할 수 있는 표현형에 대해 평가된다. 구체적으로, 형질도입된 세포는 총 HBV RNA 전사체 수준 또는 3.5kb HBV RNA 전사체 수준의 qRT-PCR 측정을 통해 표적 HBV 전사체의 전사 억제에 대해 평가된다. 또한, 여러 유전자 및/또는 그 조절 요소(예컨대 Enh1/X 프로모터 및 L-HBs 프로모터)를 동시에 표적화하여 억제할 경우, 개별 유전자만을 표적하여 억제했을 때(예컨대 Enh1/X 프로모터 단독 또는 L-HBs 프로모터 단독)보다 HBV RNA 감소 효과가 더 큰 조건이 확인되었다.

[1226] [0603] HBV 복제와 발현을 억제하기 위해 다중화된 표적 전사 억제를 매개하는 gRNA의 조합을 이용한 DNA-표적

화 시스템이 확인되었다.

[1227] [0604] 도 17은 Hep3B 세포에서 HBV RNA 내의 다양한 영역에 대한 다중화된 표적 전사 억제를 매개하는 gRNA의 조합을 보여준다. 가장 많은 억제를 매개하는 여러 gRNA를 모아서 단일 gRNA(예컨대 HBVg_26, HBVg_63, HBVg_29 및 HBVg_121과 같은 "최상의 단일 gRNA") 및 비표적 gRNA에 의해 유도된 억제와 비교하였다. gRNA는 표준 용량(20ng)으로 리포펙션을 통해 형질도입시켜 높은 발현을 위해 1x 용량(예컨대 "풀링된 상위 gRNA" = 1x/4=5ng/개 또는 "모든 16개 gRNA" = 1x/16=1.25ng/개)을 유지하도록 모았다. 도 17에 표시된 바와 같이, 다양한 영역들을 표적으로 하는 gRNA는 비표적 대조군과 비교하여 억제를 매개하였다. 그러나 여러 gRNA의 조합이 단일 gRNA에 비해 억제를 항상 유의하게 개선한 것은 아니었다. 예를 들어, HBVg_26(서열번호 416) 단독 사용은 M/S-HBs 프로모터 영역에서 상위 gRNA 풀(HBVg_26(서열번호 416), HBVg_182(서열번호 572), HBVg_185(서열번호 575), HBVg_152(서열번호 542)의 조합)과 비교할 때 유사한 수준의 억제를 매개하였다. 반면, L-HBs 프로모터에서 상위 gRNA 풀 또는 각 프로모터의 상위 단일 gRNA를 하나씩 포함한 풀을 사용할 경우, L-HBs 프로모터의 최상의 단일 gRNA를 사용한 경우보다 억제를 증가시킬 수 있었다. 이는 함께 HBV 억제와 관련하여 gRNA의 다중 사용 효과가 맥락에 따라 달라질 수 있음을 시사한다.

[1228] [0605] gRNA의 조합들을 리포펙션 용량이 낮은 PLC/PRF/5(Alexander) 세포 모델에서 다중 전사 억제에 대해 시험하였다. PLC/PRF/5는 L-HB, M-HB, S-HB를 포함한 모든 동형의 B형 간염 바이러스 표면 항원(HbsAg)을 생성하는 인간 간암 세포주이다. 이 세포주는 Hep3B 세포에서 약 2개의 통합과 M/S-HB 동형에 비해 8-10개의 HBV 통합을 포함한다. D3AL-K를 인코딩하는 mRNA와 다양한 gRNA 조합으로 리포펙션한 지 14일 후에 RNA를 수확하고 총 HBV RNA에 대해 qPCR을 실행하였다. 도 18은 HBVg_22(서열번호 412) 단독이 14일째에 가장 큰 억제를 매개했고, 그 다음이 HBVg_63(서열번호 453)과 HBVg_20(서열번호 410)이었다는 것을 보여준다. 많은 gRNA가 이들 세포에서 기능적 표적화를 하지 않는다. gRNA의 다중 사용은 HBVg_22(서열번호 412)만을 사용한 경우에 비해 억제를 크게 개선(또는 방해)하지 않았다.

[1229] [0606] 또 다른 다중화 검정에서, PXB 1차 인간 간세포(PHH) 세포를 감염 다중도(MOI) 100으로 HBV로 감염시키고 5일 동안 배양하였다. 그런 다음 세포를 DNMT3A/L-dSpCas9-KRAB(서열번호 644) 및 다양한 다중화 gRNA 혼합물로 리포펙션하여 각 gRNA는 0.5배로 투여되도록 하였다. 리포펙션 후 7일차와 14일차에 pgRNA 억제를 검정하였다. 도 19는 HBVg_22(서열번호 412), HBVg_185(서열번호 575), HBVg_26(서열번호 416)에 의해 매개되는 다중화 표적 전사 억제가 비슷한 지속성을 보였음을 나타낸다.

[1230] **실시예 6. cccDNA 및 통합 HBV DNA의 표적 메틸화**

[1231] [0607] 표적 특이성을 평가하기 위해, HepG2.NTCP 세포에 지질 단독 대조군, DNMT3A/L-dSpCas9-KRAB 융합 구조물(서열번호 S644), 또는 dSpCas9-KRAB 구조물(서열번호 594)를 인코딩하는 mRNA를 갖는 비표적화 gRNA 대조군, gRNA(탈메틸화-비활성(demethylase-null) 대조군) 또는 DNMT3A/L-dSpCas9-KRAB 융합 구조물(서열번호 644)과 조합된 gRNA를 최대 용량으로 형질도입시켰다(도 20A 및 B). 맞춤형 HBV Methylome 라이브러리를 사용한 메틸 캡처 시퀀싱은 리포펙션 후 3일째에 수행하였다. 도 20A에서 볼 수 있듯이, 지질 단독, 비표적화 gRNA 대조군, 탈메틸화 비활성 대조군에서는 메틸화가 거의 또는 전혀 관찰되지 않았다. 그러나 DNMT3A/L-dSpCas9-KRAB 융합체(서열번호 644)와 다양한 gRNA 조합(도 20B)을 함께 사용한 경우, 농축되지 않은 총 HBV DNA 풀에서 HBV 게놈 전체에 걸쳐 표적화된 뚜렷한 메틸화 패턴이 나타났다. 메틸 시퀀싱 결과, DNMT3A/L-dSpCas9-KRAB 융합 단백질이 HBV 게놈 전체에 걸쳐 강력하고 특이적인 메틸화 패턴을 유도함이 확인되었다. 도 21은 Hep3B 세포에서 다양한 gRNA(HBVg_22(서열번호 412), HBVg_26(서열번호 416), HBVg_63(서열번호 453), HBVg_185(서열번호 575))와 함께 DNMT3A/L-dSpCas9-KRAB를 인코딩하는 mRNA를 전달한 후 통합된 HBV DNA에서 증가된 메틸화 패턴을 보여준다. 이러한 메틸화 패턴은 도 20B의 cccDNA에서 볼 수 있는 것과 유사한 HBV 게놈 전체의 경향을 보였다. 도 22는 HBVg_22 (서열번호 412)와 D3AL-K를 인코딩하는 mRNA로 일시적인 리포펙션을 한 후 HBV에 감염된 PHH 기증자 #1(100 Ge/세포)에서 핵을 분리한 후의 메틸화 패턴을 보여준다. 메틸-seq 이전에 핵을 분리하면 준비 과정에서 rcDNA 오염이 줄어들어 cccDNA 신호가 풍부해진다(포획된 HBV DNA의 55%). 이 메틸화 데이터는 HBVg_22(서열번호 412)에 대해 확인된 다른 패턴과 유사하며, 이러한 후성적 유전적 변화의 시간 경과에 따른 지속성을 강조하는 것이다.

[1232] [0608] 바이러스 감염이 없는 HepG2.NTCP 세포에서 최대 용량으로 다양한 HBV 표적 gRNA와 mRNA를 리포펙션한 후, RNA 시퀀싱을 수행하여 비표적 영역에서 유전자 발현의 차등적 변화를 평가하였다. 도 23A는 표적 gRNA가 있는 에피-에디터 DNMT3A/L-dSpCas9-KRAB 융합체(서열번호 645)와 비표적 gRNA가 있는 에피-에디터 DNMT3A/L-dSpCas9-KRAB 융합체(서열번호 645)를 비교한 것이다. 결과에 따르면 리포펙션 3일 후 gRNA HBVg_22(서열번호

412)의 결과로 대조군 비표적 gRNA에 비해 유전자 발현이 크게 변하지 않았다. 도 23B는 HBVg_22(서열번호 412)를 D3AL-K(서열번호 644)를 인코딩하는 mRNA 또는 DNMT3A/3L 도메인이 없는(dCas9-KRAB, 서열번호 594) mRNA와 함께 전달한 후 차별적으로 발현된 전사체를 비교하여 에피메틸화 매개 효과를 비교한 것으로, 유전자 발현에 D3AL 매개 변화가 효과적으로 나타나지 않음을 보여준다. 도 23C는 리포펙션 후 3일째 gRNA에 의해 매개된 차별적으로 발현된 유전자(DEG)의 최소한의 변화를 보여준다. 관찰된 변화의 대부분은 두 개 이상의 gRNA에서 공유되었으며("공유 DEG") 일반적으로 RNA 전달에 대한 일시적인 세포 반응과 선천 면역 반응 관련 유전자였다. 분석된 후보 gRNA의 대부분은 비표적화 gRNA와 비교하여 유사하게 낮은 수준의 오프-표적 유전자 발현으로 오프-표적 차등 발현을 최소한으로 또는 전혀 부여하지 않았다. 또한, DNMT3A/L-dCas9-KRAB를 인코딩하는 mRNA와 HBVg_22(서열번호 412) 또는 비표적화 gRNA를 캡슐화하는 LNP를 3가지 다른 수준(24웰 플레이트에서 웰당 50ng, 250ng, 500ng)으로 1차 인간 간세포에 전달하였다. 이들 LNP를 전달한 후, 리포펙션 후 5일차와 19일차에 PHH에서 RNA를 수확하고 RNA 시퀀싱을 수행하였다. 이러한 결과는 어떠한 용량이나 시점에서도 비표적화 gRNA와 HBVg_22(서열번호 412) 사이에 차별적으로 발현되는 유전자(*[FDR < 0.01, abs(L2FC)>1]*)가 없음을 보여주었다(도 23D). 운반체 전용 PHH에 비해 전사 변화가 거의 나타나지 않았다. HBVg_22(서열번호 412)는 비표적화 대조 gRNA와 유사한 성능을 보였으며, 최고 용량에서 19일차에 차등적으로 발현되는 유전자가 하나만 남았다.

[1233] **실시예 7: HepB에 대한 인간 키메라 간 마우스 연구에서 HBV 억제의 생체 내 평가.**

[1234] [0609] 이 실시예에서, 인간 세포가 마우스 간에 고도로 이식된 인간 키메라 간 마우스 모델(FRG)이 사용되었다. 이 마우스 모델에서, 마우스는 인간 간세포(70% 이상 인간화)로 이식된 다음 HBV로 감염되었다. 마우스 모델은 기본 구조를 가진 진정한 cccDNA의 존재, 에피-에디터와 상호 작용하는 인간 세포 내 후성 유전적 기계, 인간 유전체 배경, 임상 시험에 사용되는 생물학적으로 관련된 종말점(HBV DNA, HBsAg, HBeAg)의 재생산을 포함한 여러 가지 장점이 있다.

[1235] [0610] 8 HBV 게놈에 감염되었다. 도 24 에서 볼 수 있듯이, 바이러스 감염은 HBV DNA의 평탄기(꼬리 정맥 주사 후 1-2개월)에 도달할 때까지 계속되었다. 평탄기는 대부분 cccDNA로 구성되고 HBV 통합이 매우 희소한 만성 안정화 감염이 확립되었음을 나타낸다.

[1236] [0611] 만성 안정화 감염이 확립되면, 마우스에 표준 치료 요법(예컨대 엔테카비르, "ETV"와 같은 뉴클레오사이드 유사체)을 투여하고, 1μg/ml ETV를 음용수에 첨가하였다. 81일차에, 마우스에 a) 비표적화 gRNA와 함께 투여된 대표적인 에피-에디터 mRNA 조합 또는 b) HBV 게놈을 표적으로 하는 gRNA(들)와 함께 투여된 대표적인 에피-에디터(epi-editor) mRNA 조합을 투여하였다. 마우스에 다중화 방식(실시예 4에 설명된 대로)으로 동일하거나 다른 유전자를 표적화하는 단일 gRNA 또는 여러 gRNA를 투여하였다. 마우스를 코호트로 그룹화하고 1) 비표적 gRNA(nt1)(n=9), 2) HBx HBVg_22(서열번호 412)를 표적으로 하는 gRNA(n=9), 3) gRNA HBVg_22(서열번호 412)를 운반하는 GalNAc-접합 LNP(n=3), 4) HBs HBVg_185(서열번호 575)를 표적으로 하는 gRNA(n=7), 5) gRNA HBVg_22(서열번호 412) 및 HBVg_ 185(서열번호 575)(n=7), 6) gRNA HBVg_ 73(서열번호 463) 및 7) 비표적 gRNA(nt1)(n=7), 8) gRNA HBVg_22(서열번호 412)(n=7)을 투여하였다. 에피-에디터 mRNA(그룹 7/8의 경우 DNMT3A/L-dCas9-KRAB(서열번호 644) 또는 dSpCas9-KRAB 구조(서열번호 594))와 gRNA의 조합은 지질 나노입자 전달 시스템 및 RO 주입을 통해 킬로그램당 2밀리그램(mpk) 용량으로 전달되었으며, 이를 통해 평균적으로 인간 간세포의 50-60%에만 전달되었다(단백질 면역 표지에 의해 결정됨). 그룹 3의 N-아세틸갈락토사민(GalNAc) 접합-LNP는 FRG 마우스 모델에서 간세포로의 LNP 전달을 개선하기 위해 투여되었다. 모든 대상은 표시된 대로 혈청을 채취하여 연구 종료 시점(130일차)에 평가되었다. 마우스는 HBV 전유전체 RNA(pgRNA), 혈청 내 총 HBV DNA, 혈청 내 HBeAg 단백질, 혈청 내 HbsAg 단백질을 포함한 HBV 바이오마커의 수준을 모니터링하여 억제에 대해 평가되었다. 바이오마커 수준의 변화는 인간 키메라 간 마우스에 전달된 모든 세포에 대한 바이러스 복제 및 전사의 강력한 억제를 나타낸다. 마우스는 다중 억제가 효능 및 내구성, 생체 내 HBV 메틸화 패턴 또는 재투여 후 억제 강도에 미치는 영향에 대해 추가로 평가될 수 있다.

[1237] [0612] 생체분포 마우스 코호트(n=12)를 투여 후 18시간 후에 수확하여 여러 전달 방식에 걸친 전달 효율성을 평가하였다. 주요 코호트에서 억제를 평가하기 전에, 내성을 테스트하기 위해 센티넬 재투약 코호트에 1 mpk, 2 mpk 또는 3 mpk으로 재투약하였다. 마우스는 또한 21일 또는 24일(마우스 연구의 약 102일) 후에 재투약되었다.

[1238] [0613] 도 25에 나타난 바와 같이, HBsAg 및 HBV DNA의 약 50% 억제는 투여 후 5일(86일차)에 대표적인 gRNA HBVg_22(서열번호 412)에서 관찰되었다. 도 26은 억제 수준을 강조하는 투여 후(5일차)에서 투여 전(2일 전) 지표의 배수 변화를 보여준다. 이 데이터는 gRNA가 융합 단백질과 결합되어 네이티브 cccDNA로 생체 내에서 거의 완전한 억제를 유도하였다는 강력한 증거를 나타낸다. * = p<0.05, **= p<0.01, ***= p<0.001(대응표본

t-검정)

[1239] [0614] 도 24에 설명된 연구 시작 후 21일 후에 2 mpk gRNA로 마우스에 재투여한 후 재투여 코호트를 평가하였다. 도 27A-27D에 표시된 바와 같이 HBVg_22(서열번호 412)를 단독으로 또는 HBVg_185(서열번호 575)와 다중화하여 재투여하거나 GalNAc와 접합된 LNP를 통해 전달하거나 dSpCas9-KRAB 융합 단백질(서열번호 594)을 인코딩하는 mRNA와 함께 전달한 경우 억제가 개선되었다. (도 27A-C) 또는 DNMT3A/L-dSpCas9-KRAB(도 27D).

[1240] [0615] 추가적인 인간 키메라 FRG 마우스 연구는 유사한 만성 감염 단계로 실행되었고, LNP 투여 전에 유사한 ETV 투여 단계가 이어졌다. 이 연구에 사용된 모든 LNP는 이 키메라 마우스 컨텍스트에서 전달을 개선하기 위해 0.15% 분자량의 GalNAc-PEG 지질과 함께 접합되었다. HBVg_22(서열번호 412) 또는 DNMT3A/3L-dSpCas9-KRAB을 인코딩하는 mRNA와 함께 비표적화 gRNA를 포함하는 LNP의 단일 2mpk RO 주입 후. 이 단일 투여 후, 혈청 HBeAg 및 HBsAg 단백질 수치의 지속적인 억제가 있었고, 비표적화 gRNA와 비교했을 때 LNP 투여 후 33일까지 혈청 HBV DNA 수치가 유사하게 감소하였다(도 28A). 억제 정도에 대한 단일 마우스 추적 결과 이러한 혈청 HBV 단백질 지표에서 동물 전체에서 확인된 반응의 일관성이 강조된다(도 28B). LNP 투여 72시간 후에 마우스를 수확하고, 조직을 포르말린으로 밤새 고정한 다음, 샘플을 다중 채널 RNAscope에 적용하여 세포 내 개별 mRNA 분자를 확인하였다(도 28C). 이 시점에서 Cas9 mRNA 신호가 높은 영역에서 HBVg_22(서열번호 412)를 전달한 마우스에서 비표적화 gRNA를 받은 마우스에 비해 pgRNA 신호가 현저히 감소한 것을 확인할 수 있었으며, HBVg_22(서열번호 412) 그룹에서 대부분의 남아 있는 pgRNA 신호는 세포 밖에 있었다(도 28A).

[1241] **실시예 8: 바이러스 복제 및 전사와 관련된 유전자와 그 조절 요소의 표적 전사 억제를 위한 아연 집게 단백질의 식별.**

[1242] [0616] 조작된 아연 집게 단백질(eZFP)은 바이러스 복제 및 전사와 관련된 HBV의 유전자 및 조절 요소에 결합하도록 설계되었다. KRAB와 122개의 설계된 eZFP 중 하나를 포함하는 융합 단백질(eZFP-KRAB 융합 단백질)은 Hep3B 세포에 도입되어 전사 억제가 가장 높은 eZFP를 식별하였다.

[1243] [0617] eZFP는 섹션 I-C에서 설명한 대로 설계되었다. 융합 단백질을 인코딩하는 mRNA를 생성하여 두 뱃치(도 29A 및 29B)로 Hep3B 세포에 리포펙타민을 통해 형질도입하였다. eZFP-KRAB 융합 단백질(3xFLAG-SV40NLS-eZFP-SV40NLS-KRAB 융합 단백질)을 발현하는 세포를 3일 후 HBV RNA 수치에 대해 qPCR로 평가하였다. mRNA 수치를 HBVg_22 또는 HBVg_63과 함께 pTT183(dCas9-KRAB 대조군)으로 형질도입된 세포와 비교하였다. 음성 대조군에는 TriLink GFP 대조군, 인간 게놈의 고유한 18bp 서열에 결합하지 않는 비표적화 아연 집게 단백질 pZFP0131 및 비표적화 gRNA가 포함되었다. 스크리닝된 eZFP-KRAB 융합 단백질 중 28개는 3일차에 Hep3B 세포에서 HBV mRNA 수치 감소를 초래하였다(도 29A 및 29B; 또한 표 E4 참조). 여기에는 서열번호 944-971(표 E5)에 제시된 eZFP-KRAB 융합 단백질이 포함되었으며, 이는 서열번호 1028-1055에 제시된 표적 부위를 표적으로 삼았다.

표 E4. HBV 복제 및 전사와 관련된 유전자 및 조절 요소를 표적으로 하는 eZFP

eZFP 이름	표적 부위 서열	인식 영역 F1-F6	아미노산 (AA) 서열	AA 서열 SEQ ID NO:	예시적인 뉴클레오타이드 인코딩 서열 SEQ ID NO:
eZFP_1	CAAGTGTTTGCTGACGCA (SEQ ID NO:1028)	F1:SEADRSR(SEQ ID NO:720) F2:DRSNLTR(SEQ ID NO:721) F3:QSSDLSR(SEQ ID NO:722) F4:YHWYLKK(SEQ ID NO:723) F5:RSDSLSV(SEQ ID NO:724) F6:QNANRKT(SEQ ID NO:725)	AAMAERPFQCRICMRNFSSEADRSRHIRTHTGEKPFACDICGRKFADRSNLTRHTKIHTGSQKPFQCRICMRNFSQSSDLSRHIRTHTGEKPFACDICGRKFAYHWYLKKHTKIHTGSQKPFQCRICMRNFSRSDSLSVHIRTHTGEKPFACDICGRKFAQNANRKTHTKIHLRQKDAAR	692	888
eZFP_2	GGCTGGGGCTTGGTCATG (SEQ ID NO:1029)	F1:RSDVLST(SEQ ID NO:726) F2:DNSSRTR(SEQ ID NO:727) F3:RPYTLRL(SEQ ID NO:728) F4:DSSHRTR(SEQ ID NO:729) F5:RSDHLSQ(SEQ ID NO:730) F6:DSSHRTR(SEQ ID NO:731)	AAMAERPFQCRICMRNFSRSDVLSTHIRTHTGEKPFACDICGKKFADNSSRTRHTKIHTGSQKPFQCRICMRNFSRPYTLRLHIRTHTGEKPFACDICGRKFADSSHRTRHTKIHTGSQKPFQCRICMRNFSRSDHLSQHIRTHTGEKPFACDICGRKFADSSHRTRHTKIHLRQKDAAR	693	889
eZFP_3	GCTGGGGCTTGGTCATGG (SEQ ID NO:1030)	F1:RSDHLSQ(SEQ ID NO:732) F2:QSADRTK(SEQ ID NO:733) F3:RSDHLSQ(SEQ ID NO:734) F4:RRSDLKR(SEQ ID NO:735) F5:RSDHLSR(SEQ ID NO:736) F6:QSSDLRR(SEQ ID NO:737)	AAMAERPFQCRICMRNFSRSDHLSQHIRTHTGEKPFACDICGRKFAQSADRTKHTKIHTGSQKPFQCRICMRNFSRSDHLSQHIRTHTGEKPFACDICGRKFARRSDLKRHTKIHTGSQKPFQCRICMRNFSRSDHLSRHIRTHTGEKPFACDICGRKFAQSSDLRRHTKIHLRQKDAAR	694	890
eZFP_4	TTGGTCATGGGCCATCAG (SEQ ID NO:1031)	F1:RSDNLSE(SEQ ID NO:738) F2:TSSNRKT(SEQ ID NO:739) F3:DRSHLTR(SEQ ID NO:740) F4:RSDALTQ(SEQ ID NO:741) F5:DRSALAR(SEQ ID NO:742) F6:RRFTLSK(SEQ ID NO:743)	AAMAERPFQCRICMRNFSRSDNLSEHIRTHTGEKPFACDICGRKFATSSNRKTHTKIHTGSQKPFQCRICMRNFSDRSHLTRHIRTHTGEKPFACDICGRKFARSDALTQHTKIHTGSQKPFQCRICMRNFSDRSALARHIRTHTGEKPFACDICGRKFARRFTLSKHTKIHLRQKDAAR	695	891
eZFP_5	TGCGTGGAACCTTTTCGG (SEQ ID NO:1032)	F1:RSDHLSE(SEQ ID NO:744) F2:QYSGRYY(SEQ ID NO:745) F3:HGQTLNE(SEQ ID NO:746) F4:QSGNLAR(SEQ ID NO:747) F5:RSDSLLR(SEQ ID NO:748)	AAMAERPFQCRICMRNFSRSDHLSEHIRTHTGEKPFACDICGRKFAQYSGRYYHTKIHTGSQKPFQCRICMRNFSHGQTLNEHIRTHTGEKPFACDICGRKFAQSGNLARHTKIHTGSQKPFQCRICMRNFSRSDSLLRHIRTHTGEKPFACDICG	696	892

[1244]

		F6:CREYRGK(SEQ ID NO:749)	RKFACREYRGKHTKIHLRQKDAAR		
eZFP_6	GCAGCAGGTCTGGAGCAA (SEQ ID NO:1033)	F1:QSANRTT(SEQ ID NO:750) F2:RSANLTR(SEQ ID NO:751) F3:RSDVLSE(SEQ ID NO:752) F4:TSGHLSR(SEQ ID NO:753) F5:QSSDLSR(SEQ ID NO:754) F6:QWSTRKR(SEQ ID NO:755)	AAMAERPFQCRICMRNFSQSANRTTHIRTHTGEKPFACDICGRKFARSANLTRHTKIHTGSQKPFQCRICMRNFSRSDVLSEHIRTHTGEKPFACDICGRKFATSGHLSRHTKIHTGSQKPFQCRICMRNFSQSSDLSRHIRTHTGEKPFACDICGRKFAQWSTRKRHTKIHLRQKDAAR	697	893
eZFP_7	CAGCAGGTCTGGAGCAAA (SEQ ID NO:1034)	F1:QSGNLAR(SEQ ID NO:756) F2:ATCCLAH(SEQ ID NO:757) F3:RWQYLPT(SEQ ID NO:758) F4:DRSALAR(SEQ ID NO:759) F5:RSDNLSE(SEQ ID NO:760) F6:KRCNLRC(SEQ ID NO:761)	AAMAERPFQCRICMRNFSQSGNLARHIRTHTGEKPFACDICGRKFAATCCLAHHTKIHTGSQKPFQCRICMRNFSRWQYLPTHIRTHAGEKPFACDICGRKFADRSALARHTKIHTGSQKPFQCRICMRNFSRSDNLSEHIRTHTGEKPFACDICGRKFAKRCNLRCHTKIHLRQKDAAR	698	894
eZFP_8	AGCAGGTCTGGAGCAAC (SEQ ID NO:1035)	F1:NPANLTR(SEQ ID NO:762) F2:QNATRTK(SEQ ID NO:763) F3:QSGHLAR(SEQ ID NO:764) F4:NRHDRAK(SEQ ID NO:765) F5:RSDHLSE(SEQ ID NO:766) F6:QRRSRYK(SEQ ID NO:767)	AAMAERPFQCRICMRNFSNPANLTRHIRTHTGEKPFACDICGRKFAQNATRTKHTKIHTGSQKPFQCRICMRNFSQSGHLARHIRTHTGEKPFACDICGRKFANRHDRAKHTKIHTGSQKPFQCRICMRNFSRSDHLSEHIRTHTGEKPFACDICGRKFAQRRSRYKHTKIHLRQKDAAR	699	895
eZFP_9	ATCGTATCCATGGCTGCT (SEQ ID NO:1036)	F1:QSSDLSR(SEQ ID NO:768) F2:HRSTRNR(SEQ ID NO:769) F3:RSDVLSA(SEQ ID NO:770) F4:DSRTRKN(SEQ ID NO:771) F5:QSGSLTR(SEQ ID NO:772) F6:DQSGLAH(SEQ ID NO:773)	AAMAERPFQCRICMRNFSQSSDLSRHIRTHTGEKPFACDICGRKFAHRSTRNRHTKIHTGSQKPFQCRICMRNFSRSDVLSAHIRTHTGEKPFACDICGRKFADSRTRKNHTKIHTGSQKPFQCRICMRNFSQSGSLTRHIRTHTGEKPFACDICGRKFADQSGLAHHTKIHLRQKDAAR	700	896
eZFP_10	CCAGTGGGGGTTGCGTCA (SEQ ID NO:1037)	F1:QNPAQWR(SEQ ID NO:774) F2:RSADLSR(SEQ ID NO:775) F3:TSGSLSR(SEQ ID NO:776) F4:RSDHLSR(SEQ ID NO:777) F5:RSDSLLR(SEQ ID NO:778) F6:QSYDRFQ(SEQ ID NO:779)	AAMAERPFQCRICMRNFSQNPAQWRHIRTHTGEKPFACDICGRKFARSADLSRHTKIHTGSQKPFQCRICMRNFSTSGSLSRHIRTHTGEKPFACDICGRKFARSDHLSRHTKIHTGSQKPFQCRICMRNFSRSDSLLRHIRTHTGEKPFACDICGRKFAQSYDRFQHTKIHLRQKDAAR	701	897
eZFP_11	CCCCAGCCAGTGGGGGTT (SEQ ID	F1:TSGSLSR(SEQ ID NO:780) F2:RSDHLSR(SEQ ID NO:781) F3:RSDSLLR(SEQ ID NO:782) F4:QSYDRFQ(SEQ ID	AAMAERPFQCRICMRNFSTSGSLSRHIRTHTGEKPFACDICGRKFARSDHLSRHTKIHTGSQKPFQCRICMRNFSRSDSLLRHIRTHTGEKPFACDICG	702	898

[1245]

	NO:1038)	NO:783) F5:RSDNLST(SEQ ID NO:784) F6:DNRDRIK(SEQ ID NO:785)	RKFAQSYDRFQHTKIHTGSQKPFQCRICMRNFSRSDNLSTHIRTHTGEKPFACDICGRKFADNRDRIKHTKIHLRQKDAAR		
eZFP_12	GCGCTGATGGCCCATGAC (SEQ ID NO:1039)	F1:DRSNLSR(SEQ ID NO:786) F2:LRQNLIM(SEQ ID NO:787) F3:ERGTLAR(SEQ ID NO:788) F4:RSDALTQ(SEQ ID NO:789) F5:RSDSLSQ(SEQ ID NO:790) F6:RKADRTR(SEQ ID NO:791)	AAMAERPFQCRICMRNFSDRSNLSRHIRTHTGEKPFACDICGRKFALRQNLIMHTKIHTGSQKPFQCRICMRNFSERGTLARHIRTHTGEKPFACDICGRKFARSDALTQHTKIHTGSQKPFQCRICMRNFSRSDSLSQHIRTHTGEKPFACDICGRKFARKADRTRHTKIHLRQKDAAR	703	899
eZFP_13	CACGCACGCGCTGATGGC (SEQ ID NO:1040)	F1:QYCCLTN(SEQ ID NO:792) F2:TSGNLTR(SEQ ID NO:793) F3:QSSDLSR(SEQ ID NO:794) F4:FRYYLKR(SEQ ID NO:795) F5:QSGDLTR(SEQ ID NO:796) F6:DKGNLTK(SEQ ID NO:797)	AAMAERPFQCRICMRNFSQYCCLTNHIRTHTGEKPFACDICGRKFATSGNLTRHTKIHTGSQKPFQCRICMRNFSQSSDLSRHIRTHTGEKPFACDICGRKFAFRYYLKRHTKIHTGSQKPFQCRICMRNFSQSGDLTRHIRTHTGEKPFACDICGRKFADKGNLTKHTKIHLRQKDAAR	704	900
eZFP_14	AGAGGAGCCGAAAAGGTT (SEQ ID NO:1041)	F1:TSGSLSR(SEQ ID NO:798) F2:RSDNLTT(SEQ ID NO:799) F3:QSGNLAR(SEQ ID NO:800) F4:DRTTLMR(SEQ ID NO:801) F5:QSGHLAR(SEQ ID NO:802) F6:QLTHLNS(SEQ ID NO:803)	AAMAERPFQCRICMRNFSTSGSLSRHIRTHTGEKPFACDICGRKFARSDNLTTHTKIHTGSQKPFQCRICMRNFSQSGNLARHIRTHTGEKPFACDICGRKFADRTTLMRHTKIHTGSQKPFQCRICMRNFSQSGHLARHIRTHTGEKPFACDICGRKFAQLTHLNSHTKIHLRQKDAAR	705	901
eZFP_15	GGATCGGCAGAGGAGCCG (SEQ ID NO:1042)	F1:IKHDLHR(SEQ ID NO:804) F2:RSANLTR(SEQ ID NO:805) F3:RSDNLAR(SEQ ID NO:806) F4:QNVSRPR(SEQ ID NO:807) F5:RSDDLSK(SEQ ID NO:808) F6:DSSHRTR(SEQ ID NO:809)	AAMAERPFQCRICMRNFSIKHDLHRHIRTHTGEKPFACDICGRKFARSANLTRHTKIHTGSQKPFQCRICMRNFSRSDNLARHIRTHTGEKPFACDICGRKFAQNVSRPRHTKIHTGSQKPFQCRICMRNFSRSDDLSKHIRTHTGEKPFACDICGRKFADSSHRTRHTKIHLRQKDAAR	706	902
eZFP_16	CAGTATGGATCGGCAGAG (SEQ ID NO:1043)	F1:RSDNLAR(SEQ ID NO:810) F2:QNVSRPR(SEQ ID NO:811) F3:RSDDLSK(SEQ ID NO:812) F4:DSSHRTR(SEQ ID NO:813) F5:TSSNRKT(SEQ ID NO:814)	AAMAERPFQCRICMRNFSRSDNLARHIRTHTGEKPFACDICGRKFAQNVSRPRHTKIHTGSQKPFQCRICMRNFSRSDDLSKHIRTHTGEKPFACDICGRKFADSSHRTRHTKIHTGSQKPFQCRICMRNFSTSSNRKTHIRTHTGEKPFACDICGRKFAAQWTRACHTKIHLRQ	707	903

[1246]

		F6:AQWTRAC(SEQ ID NO:815)	KDAAR		
eZFP_17	GTTCCGCAGTATGGATCG (SEQ ID NO:1044)	F1:RSDDLSK(SEQ ID NO:816) F2:DSSHRTR(SEQ ID NO:817) F3:TSSNRKT(SEQ ID NO:818) F4:AQWTRAC(SEQ ID NO:819) F5:RKQTRTT(SEQ ID NO:820) F6:HRSSLRR(SEQ ID NO:821)	AAMAERPFQCRICMRNFSRSDDLSKHIRTHTGEKPFACDICGRKFADSSHRTRHTKIHTGSQKPFQCRICMRNFSTSSNRKTHIRTHTGEKPFACDICGRKFAAQWTRACHTKIHTGSQKPFQCRICMRNFSRKQTRTTHIRTHTGEKPFACDICGRKFAHRSSLRRHTKIHLRQKDAAR	708	904
eZFP_18	GGAGTTCCGCAGTATGGA (SEQ ID NO:1045)	F1:QSAHRKN(SEQ ID NO:822) F2:TSSNRKT(SEQ ID NO:823) F3:RSDNLSA(SEQ ID NO:824) F4:RNNDRKT(SEQ ID NO:825) F5:TSGSLSR(SEQ ID NO:826) F6:QAGHLAK(SEQ ID NO:827)	AAMAERPFQCRICMRNFSQSAHRKNHIRTHTGEKPFACDICGRKFATSSNRKTHTKIHTGSQKPFQCRICMRNFSRSDNLSAHIRTHTGEKPFACDICGRKFARNNDRKTHTKIHTGSQKPFQCRICMRNFSTSGSLSRHIRTHTGEKPFACDICGRKFAQAGHLAKHTKIHLRQKDAAR	709	905
eZFP_19	GCAAAACAAGCGGCTAGG (SEQ ID NO:1046)	F1:RSDHLSQ(SEQ ID NO:828) F2:ASSTRTK(SEQ ID NO:829) F3:RSDDLTR(SEQ ID NO:830) F4:QKSNLSS(SEQ ID NO:831) F5:QSANRTT(SEQ ID NO:832) F6:QNATRTK(SEQ ID NO:833)	AAMAERPFQCRICMRNFSRSDHLSQHIRTHTGEKPFACDICGRKFAASSTRTKHTKIHTGSQKPFQCRICMRNFSRSDDLTRHIRTHTGEKPFACDICGRKFAQKSNLSSHTKIHTGSQKPFQCRICMRNFSQSANRTTHIRTHTGEKPFACDICGRKFAQNATRTKHTKIHLRQKDAAR	710	906
eZFP_20	TTTGCTCCAGACCTGCTG (SEQ ID NO:1047)	F1:RSDTLSE(SEQ ID NO:834) F2:RRWTLVG(SEQ ID NO:835) F3:DRSNLSR(SEQ ID NO:836) F4:QSGDLTR(SEQ ID NO:837) F5:QSSDLSR(SEQ ID NO:838) F6:YHWYLKK(SEQ ID NO:839)	AAMAERPFQCRICMRNFSRSDTLSEHIRTHTGEKPFACDICGRKFARRWTLVGHTKIHTGSQKPFQCRICMRNFSDRSNLSRHIRTHTGEKPFACDICGRKFAQSGDLTRHTKIHTGSQKPFQCRICMRNFSQSSDLSRHIRTHTGEKPFACDICGRKFAYHWYLKKHTKIHLRQKDAAR	711	907
eZFP_21	GATGTATATTTGCGGGAG (SEQ ID NO:1048)	F1:RSANLAR(SEQ ID NO:840) F2:RSDNLRE(SEQ ID NO:841) F3:RPYTLRL(SEQ ID NO:842) F4:HRSNLNK(SEQ ID NO:843) F5:QSGSLTR(SEQ ID NO:844) F6:TSANLSR(SEQ ID NO:845)	AAMAERPFQCRICMRNFSRSANLARHIRTHTGEKPFACDICGRKFARSDNLREHTKIHTGSQKPFQCRICMRNFSRPYTLRLHIRTHTGEKPFACDICGRKFAHRSNLNKHTKIHTGSQKPFQCRICMRNFSQSGSLTRHIRTHTGEKPFACDICGRKFATSANLSRHTKIHLRQKDAAR	712	908
eZFP_22	CAGCCAGTGGGGGTTGCG (SEQ ID NO:1049)	F1:RSDDLVR(SEQ ID NO:846) F2:TSGSLVR(SEQ ID NO:847) F3:RSDKLVR(SEQ ID NO:848) F4:RSDELVR(SEQ ID NO:849)	LEPGEKPYACPECGKSFSRSDDLVRHQRTHTGEKPYKCPECGKSFSTSGSLVRHQRTHTGEKPYKCPECGKSFSRSDKLVRHQRTHTGEKPYACPECGKSFSRSDELVRHQRTHTGEKPYKCPECGKSFSTSHSL	713	909

[1247]

		F5:TSHSLTE(SEQ ID NO:850) F6:RADNLTE(SEQ ID NO:851)	TEHQRTHTGEKPYKCPECGKSFSRADNLTEHQRTHTGKKTS		
eZFP_23	GATGGCCCATGACCAAGC (SEQ ID NO:1050)	F1:ERSHLRE(SEQ ID NO:852) F2:TSHSLTE(SEQ ID NO:853) F3:QAGHLAS(SEQ ID NO:854) F4:TSHSLTE(SEQ ID NO:855) F5:DPGHLVR(SEQ ID NO:856) F6:TSGNLVR(SEQ ID NO:857)	LEPGEKPYACPECGKSFSERSHLREHQRTHTGEKPYKCPECGKSFSTSHSLTEHQRTHTGEKPYKCPECGKSFSQAGHLASHQRTHTGEKPYACPECGKSFSTSHSLTEHQRTHTGEKPYKCPECGKSFSDPGHLVRHQRTHTGEKPYKCPECGKSFSTSGNLVRHQRTHTGKKTS	714	910
eZFP_24	GGGGTAAAGGTTCAGGTA (SEQ ID NO:1051)	F1:RADNLTE(SEQ ID NO:858) F2:TSGSLVR(SEQ ID NO:859) F3:RKDNLKN(SEQ ID NO:860) F4:QSSSLVR(SEQ ID NO:861) F5:RSDKLVR(SEQ ID NO:862) F6:DSGNLRV(SEQ ID NO:863)	LEPGEKPYKCPECGKSFSRADNLTEHQRTHTGEKPYKCPECGKSFSTSGSLVRHQRTHTGEKPYKCPECGKSFSRKDNLKNHQRTHTGEKPYKCPECGKSFSQSSSLVRHQRTHTGEKPYKCPECGKSFSRSDKLVRHQRTHTGEKPYKCPECGKSFSDSGNLRVHQRTHTGKKTS	715	911
eZFP_25	GCTAGGAGTTCCGCAGTA (SEQ ID NO:1052)	F1:QSSSLVR(SEQ ID NO:864) F2:QSGDLRR(SEQ ID NO:865) F3:RSDERKR(SEQ ID NO:866) F4:HRTTLTN(SEQ ID NO:867) F5:RSDHLTN(SEQ ID NO:868) F6:TSGELVR(SEQ ID NO:869)	LEPGEKPYACPECGKSFSQSSSLVRHQRTHTGEKPYKCPECGKSFSQSGDLRRHQRTHTGEKPYKCPECGKSFSRSDERKRHQRTHTGEKPYACPECGKSFSHRTTLTNHQRTHTGEKPYKCPECGKSFSRSDHLTNHQRTHTGEKPYKCPECGKSFSTSGELVRHQRTHTGKKTS	716	912
eZFP_26	GCGGCTAGGAGTTCCGCA (SEQ ID NO:1053)	F1:QSGDLRR(SEQ ID NO:870) F2:RSDERKR(SEQ ID NO:871) F3:HRTTLTN(SEQ ID NO:872) F4:RSDHLTN(SEQ ID NO:873) F5:TSGELVR(SEQ ID NO:874) F6:RSDDLVR(SEQ ID NO:875)	LEPGEKPYACPECGKSFSQSGDLRRHQRTHTGEKPYKCPECGKSFSRSDERKRHQRTHTGEKPYKCPECGKSFSHRTTLTNHQRTHTGEKPYACPECGKSFSRSDHLTNHQRTHTGEKPYKCPECGKSFSTSGELVRHQRTHTGKKTS	717	913
eZFP_27	GTATGGATCGGCAGAGGA (SEQ ID NO:1054)	F1:QRAHLER(SEQ ID NO:876) F2:QLAHLRA(SEQ ID NO:877) F3:DPGHLVR(SEQ ID NO:878) F4:RRSACRR(SEQ ID NO:879) F5:RSDHLTT(SEQ ID NO:880) F6:QSSSLVR(SEQ ID NO:881)	LEPGEKPYACPECGKSFSQRAHLERHQRTHTGEKPYKCPECGKSFSQLAHLRAHQRTHTGEKPYKCPECGKSFSDPGHLVRHQRTHTGEKPYACPECGKSFSRRSACRRHQRTHTGEKPYKCPECGKSFSRSDHLTTHQRTHTGEKPYKCPECGKSFSQSSSLVRHQRTHTGKKTS	718	914
eZFP_28	CCGATCCATA	F1:QSSNLVR(SEQ ID NO:882)	LEPGEKPYACPECGKSFSQSSNLVRHQRTHTGEKPYKCP	719	915

[1248]

	CTGCGGAA (SEQ ID NO:1055)	F2:RSDDLVR(SEQ ID NO:883) F3:THLDLIR(SEQ ID NO:884) F4:TSGNLTE(SEQ ID NO:885) F5:RRSACRR(SEQ ID NO:886) F6:RNDTLTE(SEQ ID NO:887)	ECGKSFSRSDDLVRHQRTHTGEKPYKCPECGKSFSTHLDLIRHQRTHTGEKPYACPECGKSFSTSGNLTEHQRTHTGEKPYKCPECGKSFSRRSACRRHQRTHTGEKPYKCPECGKSFSRNDTLTEHQRTHTGKKTS		

표 E5. eZFP-KRAB 융합 단백질		
SEQ ID NOS	서열	설명
916	ATGGCACCGAAGAAAAAACGCAAAGTTGGTCGAGTACCCGCTGCTATGGCAGAGCGGCCATTTCAGTGTCGAATTTGCATGCGAAATTTCTCTAGCGAAGCGGATCGAAGTAGACATATCCGGACCCATACGGGTGAGAAACCGTTCGCGTGTGATATATGCGGTCGCAAATTCGCAGACCGCTCCAACCTGACAAGGCACACAAAAATTCATACTGGAAGCCAGAAACCGTTTCAGTGCCGGATTTGTATGCGCAATTTCTCTCAATCCTCCGATCTCTCCCGCCACATCAGGACTCATACCGGGGAGAAACCCTTTGCTTGTGACATTTGTGGCAGAAAGTTTGCCTATCACTGGTACCTTAAGAAGCACACCAAGATCCATACTGGCTCTCAAAAGCCTTTCCAATGCCGAATATGTATGCGAAACTTTTCTAGGTCTGACTCCCTCTCTGTACATATCCGCACTCACACGGGTGAAAAACCATTTGCCTGTGACATATGTGGCAGAAAATTTGCTCAAAATGCGAACCGAAAAACGCATACGAAAATCCATTTGCGACAAAAAGATGCAGCTCGGGGCTCCGGACCGAAAAAAAAGAGAAAGGTAAACGGCGGGGGAGGATCACGGACCCTGGTTACATTTAAGGACGTTTTTGTAGACTTCACGAGGGAGGAATGGAAGCTGTTGGACACCGCCCAGCAAATCGTGTATCGAAATGTAATGTTGGAAAATTATAAAAATCTGGTTTCCCTCGGCTACCAATTGACAAAACCAGACGTCATCCTCAGACTGGAAAAAGGGGAAGAACCTTGGCTTGTCTGA	3xFLAG-SV40 NLS--ZFP-SV40-NLS-KRAB (DNA 서열)
917	ATGGCACCGAAGAAAAAACGCAAAGTTGGTCGAGTACCCGCGGCAATGGCAGAACGCCCGTTTCAGTGTAGAATCTGCATGCGCAATTTCTCCAGAAGTGACGTACTCAGCACGCATATCAGAACACATACTGGTGAAAAACCGTTTGCTTGTGATATCTGTGGTAAGAAATTCGCGGATAACTCCTCAAGGACGCGGCATACGAAGATCCACACTGGTTCTCAAAAGCCGTTCCAGTGCCGGATATGCATGCGGAACTTTAGTCGGCCATATACGCTTCGACTTCATATTAGGACTCACACCGGCGAAAAGCCGTTCGCCTGTGACATTTGCGGGAGAAAATTCGCTGACTCTAGTCACCGCACGAGGCACACTAAAATACATACTGGTTCACAAAAACCATTCCAGTGCCGAATTTGCATGCGAAATTTTTCCCGAAGTGACCATCTCTCACAGCACATCCGCACCCATACAGGGGAGAAGCCCTTCGCTTGTGATATATGCGGACGCAAGTTCGCGGACAGCTCACACCGGACCCGCCATACAAAGATCCACTTGAGACAGAAAGATGCAGCGCGGGGCTCCGGACCGAAAAAAAAAGAGAAAGGTAAACGGCGGGGGAGGATCACGGACCCTGGTTACATTTAAGGACGTTTTTGTAGACTTCACGAGGGAGGAATGGAAGCTGTTGGACACCGCCCAGCAAATCGTGTATCGAAATGTAATGTTGGAAAATTATAAAAATCTGGTTTCCCTCGGCTACCAATTGACAAAACCAGACGTCATCCTCAGACTGGAAAAAGGGGAAGAACCTTGGCTTGTCTGA	3xFLAG-SV40 NLS--ZFP-SV40-NLS-KRAB (DNA 서열)
918	ATGGCACCGAAGAAAAAACGCAAAGTTGGTCGAGTACCCGCCGCAATGGCCGAAAGACCATTTCAGTGCAGGATATGTATGCGCAACTTCTCTCGCAGTGACCACCTGAGTCAACATATCAGGACACACACGGGTGAAAAGCCTTTTGCATGCGATATTTGTGGTCGAAAATTTGCTCAGTCTGCGGACCGAACCAAGCACACTAAAATTCATACCGGCTCACAGAAACCGTTTCAATGCCGCATCTGTATGAGGAATTTCTCTAGATCAGACCACTTGTCCCAACACATCCGGACTCATACTGGAGAAAAGCCGTTTGCATGTGACATTTGTGGCAGGAAGTTTGCTAGAAGGTCTGACCTTAAAAGGCACACAAAAATTCATACGGGTTCCCAGAAACCATTTCAGTGCCGGATATGCATGCGGAACTTTTCACGAAGCGACCACCTCTCCCGACATATTCGAACGCACACTGGTGAGAAGCCGTTTGCTTGCGACATTTGCGGACGCAAGTTCGCTCAGAGCTCCGACTTGAGG	3xFLAG-SV40 NLS--ZFP-SV40-NLS-KRAB (DNA 서열)

[1249]

	AGGCATACCAAGATTCATCTCCGGCAGAAAGATGCCGCGCGGGGCTCC GGACCGAAAAAAAAGAGAAAGGTAAACGGCGGGGGAGGATCACGGAC CCTGGTTACATTTAAGGACGTTTTTGTAGACTTCACGAGGGAGGAATGG AAGCTGTTGGACACCGCCCAGCAAATCGTGTATCGAAATGTAATGTTGG AAAATTATAAAAATCTGGTTTCCCTCGGCTACCAATTGACAAAACCAGA CGTCATCCTCAGACTGGAAAAAGGGGAAGAACCTTGGCTTGTCTGA	
919	ATGGCACCGAAGAAAAAACGCAAAGTTGGTCGAGTACCCGCCGCGATG GCTGAGAGACCATTTCAGTGTCGAATCTGCATGAGAAATTTTTCAAGGA GTGACAATCTGTCTGAGCACATACGAACACATACTGGGGAAAAACCCTT TGCATGTGACATTTGTGGAAGAAAGTTTGCTACCAGCTCAAATCGCAAA ACACATACAAAGATACATACCGGCTCCCAAAAGCCATTCCAGTGCCGC ATCTGCATGAGGAACTTTTCCGATCGCTCACATCTTACCCGCCACATAA GAACTCACACAGGCGAAAAGCCCTTTGCCTGCGATATATGCGGACGGA AGTTCGCCCGCTCCGACGCTTTGACCCAGCATACCAAAATCCATACTGG GTCTCAAAAGCCATTTCAGTGCCGAATCTGTATGAGGAATTTCTCCGAC AGGTCAGCATTGGCACGGCATATCCGCACCCATACCGGTGAGAAGCCTT TTGCTTGCGATATCTGTGGACGAAAATTTGCCCGGAGGTTCACTCTCTCC AAACACACAAAGATACATCTGCGCCAAAAGGATGCAGCCCGGGGCTCC GGACCGAAAAAAAAAGAGAAAGGTAAACGGCGGGGGAGGATCACGGAC CCTGGTTACATTTAAGGACGTTTTTGTAGACTTCACGAGGGAGGAATGG AAGCTGTTGGACACCGCCCAGCAAATCGTGTATCGAAATGTAATGTTGG AAAATTATAAAAATCTGGTTTCCCTCGGCTACCAATTGACAAAACCAGA CGTCATCCTCAGACTGGAAAAAGGGGAAGAACCTTGGCTTGTCTGA	3xFLAG-SV40 NLS--ZFP-SV40-NLS-KRAB (DNA 서열)
920	ATGGCACCGAAGAAAAAACGCAAAGTTGGTCGAGTACCCGCCGCAATG GCTGAGCGCCCGTTCCAGTGCAGAATATGCATGCGGAATTTTTCTAGGT CAGATCATTTGTCTGAGCATATTCGCACACACACGGGAGAGAAGCCCTT TGCTTGCGATATATGTGGAAGGAAATTCGCGCAATACAGTGGGCGCTAC TACCATACAAAGATCCATACGGGCTCCCAGAAGCCCTTCCAATGTCGAA TATGTATGAGGAATTTTAGTCACGGACAAACATTGAATGAACATATACG CACTCACACTGGTGAAAAACCATTTGCGTGCGATATTTGCGGAAGGAAG TTTGCTCAGTCTGGGAATTTGGCGCGACACACCAAGATCCACACAGGAT CCCAGAAACCATTTCAGTGCAGAATTTGTATGAGAAACTTTAGCCGCAG TGACAGTCTCTTGAGGCACATACGGACTCATACTGGGGAGAAACCATTC GCCTGCGATATTTGTGGACGAAAGTTCGCCTGTCGCGAGTACAGAGGCA AGCACACTAAGATACATCTTAGGCAAAAGGACGCTGCACGGGGCTCCG GACCGAAAAAAAAAGAGAAAGGTAAACGGCGGGGGAGGATCACGGACC CTGGTTACATTTAAGGACGTTTTTGTAGACTTCACGAGGGAGGAATGGA AGCTGTTGGACACCGCCCAGCAAATCGTGTATCGAAATGTAATGTTGGA AAATTATAAAAATCTGGTTTCCCTCGGCTACCAATTGACAAAACCAGAC GTCATCCTCAGACTGGAAAAAGGGGAAGAACCTTGGCTTGTCTGA	3xFLAG-SV40 NLS--ZFP-SV40-NLS-KRAB (DNA 서열)
921	ATGGCACCGAAGAAAAAACGCAAAGTTGGTCGAGTACCCGCCGCGATG GCTGAGAGGCCTTTTCAATGTCGAATCTGTATGAGGAACTTCTCTCAAT CTGCTAATCGCACGACGCACATTCGAACGCATACCGGTGAGAAGCCATT CGCGTGCGATATCTGCGGACGGAAATTCGCGAGGTCAGCTAATCTTACA CGGCACACGAAGATCCACACAGGGTCACAGAAACCTTTTCAGTGTCGC ATTTGCATGAGGAATTTCTCCCGATCTGACGTCCTTAGCGAACATATAC GAACTCACACGGGCGAGAAGCCATTTGCGTGCGATATATGCGGGAGGA AGTTTGCCACCTCTGGACATCTGAGTCGACATACCAAAATTCATACCGG TAGTCAGAAGCCGTTCCAATGCAGAATATGTATGCGAAATTTCTCTCAA AGCTCAGACTTGTCTAGGCACATAAGAACGCACACGGGTGAAAAACCT TTCGCGTGTGATATCTGCGGCAGAAAGTTCGCACAATGGTCCACCCGAA AGCGGCATACGAAGATTCACCTCAGACAGAAAGACGCTGCCCGGGGCT CCGGACCGAAAAAAAAAGAGAAAGGTAAACGGCGGGGGAGGATCACGG ACCCTGGTTACATTTAAGGACGTTTTTGTAGACTTCACGAGGGAGGAAT GGAAGCTGTTGGACACCGCCCAGCAAATCGTGTATCGAAATGTAATGTT GGAAAATTATAAAAATCTGGTTTCCCTCGGCTACCAATTGACAAAACCA GACGTCATCCTCAGACTGGAAAAAGGGGAAGAACCTTGGCTTGTCTGA	3xFLAG-SV40 NLS--ZFP-SV40-NLS-KRAB (DNA 서열)
922	ATGGCACCGAAGAAAAAACGCAAAGTTGGTCGAGTACCCGCGGCGATG GCAGAACGCCCGTTCCAATGCAGAATATGTATGAGAAACTTCTCCCAGA	3xFLAG-SV40 NLS--ZFP-

[1250]

	GCGGAAATCTGGCACGCCACATCCGGACACACACGGGAGAGAAGCCAT TTGCTTGTGACATTTGTGGTCGCAAATTTGCCGCCACCTGTTGTCTGGCA CATCATACTAAGATACATACGGGGTCACAGAAACCATTCCAATGTAGG ATCTGCATGCGGAATTTTTCTCGGTGGCAGTATTTGCCTACGCATATTAG AACCCACGCCGGTGAGAAACCGTTTGCATGTGACATCTGCGGACGAAA GTTTGCCGATAGATCTGCGCTTGCTAGGCATACTAAAATCCACACGGGG TCCCAGAAGCCTTTTCAGTGTCGGATATGTATGAGGAACTTCAGTCGAT CAGACAACCTTAGCGAGCATATTCGGACGCATACTGGAGAAAAACCTTT TGCTTGTGATATATGCGGTAGGAAGTTCGCCAAACGGTGTAACCTTCGC TGTCACACCAAAATACATCTTCGCCAGAAAGATGCGGCCCGGGGCTCCG GACCGAAAAAAAAAGAGAAAGGTAAACGGCGGGGGAGGATCACGGACC CTGGTTACATTTAAGGACGTTTTTGTAGACTTCACGAGGGAGGAATGGA AGCTGTTGGACACCGCCCAGCAAATCGTGTATCGAAATGTAATGTTGGA AAATTATAAAAATCTGGTTTCCCTCGGCTACCAATTGACAAAACCAGAC GTCATCCTCAGACTGGAAAAAGGGGAAGAACCTTGGCTTGTCTGA	SV40-NLS-KRAB (DNA 서열)
923	ATGGCACCGAAGAAAAAACGCAAAGTTGGTCGAGTACCCGCCGCTATG GCTGAAAGACCATTCCAGTGCAGAATATGTATGAGGAATTTTTCTAATC CCGCGAACCTTACGCGCCATATCAGGACGCACACGGGCGAAAAGCCCT TCGCCTGCGACATTTGTGGGAGAAAGTTTGCTCAAAACGCGACCAGGAC AAAGCACACGAAAATTCACACTGGTAGCCAGAAGCCGTTCCAGTGTAG GATCTGTATGCGCAATTTCTCTCAGTCCGGGCACCTCGCGCGACACATA AGAACTCATACGGGGGAGAAGCCGTTTGCATGTGACATCTGCGGCCGC AAGTTTGCGAATAGGCATGACAGGGCAAAACATACGAAGATCCATACA GGTTCTCAAAAACCTTTCCAATGTCGAATATGCATGCGCAACTTTAGTC GGTCAGACCACCTTTCTGAACACATCAGGACACACACTGGCGAAAAGC CGTTCGCATGTGACATTTGCGGCAGAAAGTTCGCACAAAGACGGTCCCG CTATAAGCACACCAAAATTCACCTTAGGCAAAAGGATGCAGCTCGGGG CTCCGGACCGAAAAAAAAAGAGAAAGGTAAACGGCGGGGGAGGATCAC GGACCCTGGTTACATTTAAGGACGTTTTTGTAGACTTCACGAGGGAGGA ATGGAAGCTGTTGGACACCGCCCAGCAAATCGTGTATCGAAATGTAATG TTGGAAAATTATAAAAATCTGGTTTCCCTCGGCTACCAATTGACAAAAC CAGACGTCATCCTCAGACTGGAAAAAGGGGAAGAACCTTGGCTTGTCT GA	3xFLAG-SV40 NLS--ZFP-SV40-NLS-KRAB (DNA 서열)
924	ATGGCACCGAAGAAAAAACGCAAAGTTGGTCGAGTACCCGCGGCAATG GCAGAACGACCCTTCCAATGCCGCATATGTATGCGAAACTTCAGCCAGA GCTCAGATCTTTCCAGACACATCAGGACTCACACTGGCGAAAAACCATT TGCATGCGATATATGCGGGAGAAAATTCGCGCACCGCAGTACGCGAAA CAGGCATACAAAGATACATACTGGCAGTCAAAAGCCATTTCAATGTCG AATATGCATGAGGAACTTTAGTCGATCTGACGTGCTGAGCGCTCACATA CGGACCCATACCGGAGAGAAACCATTCGCTTGTGACATCTGTGGTAGGA AGTTCGCGGATTCCCGGACCCGCAAAAATCATACTAAAATTCACACTGG GTCTCAGAAGCCCTTTCAGTGTAGGATATGTATGCGCAATTTTAGCCAG AGTGGTTCATTGACTCGGCATATCAGAACACATACTGGAGAGAAACCTT TCGCGTGTGATATTTGCGGTCGAAAGTTCGCAGATCAGAGTGGACTTGC GCACCATACTAAGATCCACCTGAGACAGAAGGACGCTGCGCGGGGCTC CGGACCGAAAAAAAAAGAGAAAGGTAAACGGCGGGGGAGGATCACGGA CCCTGGTTACATTTAAGGACGTTTTTGTAGACTTCACGAGGGAGGAATG GAAGCTGTTGGACACCGCCCAGCAAATCGTGTATCGAAATGTAATGTTG GAAAATTATAAAAATCTGGTTTCCCTCGGCTACCAATTGACAAAACCAG ACGTCATCCTCAGACTGGAAAAAGGGGAAGAACCTTGGCTTGTCTGA	3xFLAG-SV40 NLS--ZFP-SV40-NLS-KRAB (DNA 서열)
925	ATGGCACCGAAGAAAAAACGCAAAGTTGGTCGAGTACCCGCTGCCATG GCGGAGCGCCCTTTCCAGTGTAGGATATGTATGCGCAACTTCAGTCAGA ACCCAGCCCAGTGGCGGCACATACGGACGCATACTGGAGAAAAGCCAT TTGCATGTGATATCTGCGGGCGAAAATTCGCGCGGTCAGCAGATTTGAG CCGGCATACGAAGATCCATACAGGTTCACAAAAGCCATTTCAATGTCGG ATATGTATGCGGAACTTCAGCACGTCCGGCTCATTGTCAAGACATATAC GAACTCATACCGGAGAGAAACCCTTCGCGTGCGACATTTGCGGTCGGA AGTTCGCGCGATCCGACCATCTGTCACGACATACGAAAATACACACTGG CTCTCAAAAGCCGTTTCAGTGCAGAATTTGCATGAGAAATTTTAGCAGG	3xFLAG-SV40 NLS--ZFP-SV40-NLS-KRAB (DNA 서열)

[1251]

	AGCGACTCACTCCTTCGGCATATACGAACACACACTGGTGAGAAGCCAT TTGCCTGTGATATTTGTGGACGAAAGTTTGCGCAATCTTACGATAGGTTT CAGCATACAAAAATCCACCTTCGGCAAAAGGACGCGGCACGGGGCTCC GGACCGAAAAAAAAAGAGAAAGGTAAACGGCGGGGGAGGATCACGGAC CCTGGTTACATTTAAGGACGTTTTTGTAGACTTCACGAGGGAGGAATGG AAGCTGTTGGACACCGCCCAGCAAATCGTGTATCGAAATGTAATGTTGG AAAATTATAAAAATCTGGTTTCCCTCGGCTACCAATTGACAAAACCAGA CGTCATCCTCAGACTGGAAAAAGGGGAAGAACCTTGGCTTGTCTGA	
926	ATGGCACCGAAGAAAAAACGCAAAGTTGGTCGAGTACCCGCTGCCATG GCTGAACGACCGTTTCAATGTCGAATTTGCATGCGCAACTTCTCCACGT CCGGGTCTCTCAGTAGACACATCAGAACGCATACTGGTGAAAAACCATT CGCTTGTGACATATGCGGCCGAAAATTCGCGCGGAGCGACCACCTGTCA CGGCATACCAAAATTCACACCGGGAGTCAAAAACCGTTCCAGTGTAGG ATATGTATGCGCAACTTCAGCCGGTCTGACAGTCTGCTTCGACATATTC GGACGCACACTGGTGAAAAGCCGTTTGCGTGCGACATTTGTGGTCGAAA GTTCGCTCAATCTTATGATAGGTTTCAACACACCAAAATACATACGGGC TCCCAGAAGCCGTTCCAGTGCAGAATATGCATGAGAAATTTCTCTCGCA GTGACAATTTGTCCACCCATATTCGAACGCACACCGGCGAGAAACCCTT CGCCTGCGATATTTGCGGTCGCAAGTTCGCAGACAACAGGGATAGGAT AAAACATACGAAGATCCATCTGAGGCAAAAAGACGCCGCCCGGGGCTC CGGACCGAAAAAAAAAGAGAAAGGTAAACGGCGGGGGAGGATCACGGA CCCTGGTTACATTTAAGGACGTTTTTGTAGACTTCACGAGGGAGGAATG GAAGCTGTTGGACACCGCCCAGCAAATCGTGTATCGAAATGTAATGTTG GAAAATTATAAAAATCTGGTTTCCCTCGGCTACCAATTGACAAAACCAG ACGTCATCCTCAGACTGGAAAAAGGGGAAGAACCTTGGCTTGTCTGA	3xFLAG-SV40 NLS--ZFP-SV40-NLS-KRAB (DNA 서열)
927	ATGGCACCGAAGAAAAAACGCAAAGTTGGTCGAGTACCCGCAGCCATG GCAGAGCGGCCATTCCAGTGCAGAATCTGCATGCGGAACTTTTCCGATA GGTCCAATCTGTCACGCCATATTAGGACACACACGGGTGAAAAACCGTT CGCGTGTGACATATGCGGTCGCAAATTCGCCCTGAGACAGAACCTGATT ATGCACACAAAAATACATACGGGAAGCCAGAAACCGTTCCAGTGTCGG ATATGCATGAGGAACTTCAGTGAGAGGGGGACTTTGGCGAGGCACATC AGGACTCACACTGGGGAGAAGCCCTTTGCATGTGATATCTGTGGCCGAA AATTTGCTCGATCAGATGCTCTCACCCAACATACAAAGATCCATACTGG CTCTCAAAAACCGTTTCAATGTAGAATTTGTATGCGCAACTTCTCTCGGT CAGATAGCCTGTCCCAGCATATCCGAACTCATACAGGTGAGAAACCCTT CGCATGCGACATCTGTGGGCGAAAATTTGCTAGAAAAGCAGACCGGAC CCGACACACAAAGATTCATCTGCGACAAAAAGACGCCGCCCGGGGCTC CGGACCGAAAAAAAAAGAGAAAGGTAAACGGCGGGGGAGGATCACGGA CCCTGGTTACATTTAAGGACGTTTTTGTAGACTTCACGAGGGAGGAATG GAAGCTGTTGGACACCGCCCAGCAAATCGTGTATCGAAATGTAATGTTG GAAAATTATAAAAATCTGGTTTCCCTCGGCTACCAATTGACAAAACCAG ACGTCATCCTCAGACTGGAAAAAGGGGAAGAACCTTGGCTTGTCTGA	3xFLAG-SV40 NLS--ZFP-SV40-NLS-KRAB (DNA 서열)
928	ATGGCACCGAAGAAAAAACGCAAAGTTGGTCGAGTACCCGCGGCCATG GCTGAGAGGCCTTTTCAATGTAGAATATGTATGCGAAATTTTTCACAGT ACTGTTGTCTCACGAACCACATAAGGACTCATACAGGGGAGAAACCATT TGCCTGTGACATTTGCGGTCGCAAATTTGCTACTTCTGGAAACCTGACTC GGCACACTAAGATTCACACAGGGTCCCAGAAGCCCTTCCAGTGTCGCAT TTGCATGAGGAATTTTAGTCAAAGCTCTGACTTGTCAAGGCATATTCGC ACGCACACGGGCGAAAAGCCGTTCGCTTGCGACATATGCGGGCGGAAA TTTGCCTTCCGCTATTATTTGAAGAGACACACCAAGATACATACGGGCT CTCAGAAGCCCTTTCAGTGTAGGATTTGCATGCGCAATTTTTCACAATCT GGTGATCTCACGCGACACATCCGGACTCACACAGGTGAAAAGCCTTTCG CGTGCGACATTTGCGGCCGGAAGTTTGCTGACAAGGGCAACCTCACAA AGCATACGAAGATTCACTTGAGGCAGAAAGATGCTGCTCGGGGCTCCG GACCGAAAAAAAAAGAGAAAGGTAAACGGCGGGGGAGGATCACGGACC CTGGTTACATTTAAGGACGTTTTTGTAGACTTCACGAGGGAGGAATGGA AGCTGTTGGACACCGCCCAGCAAATCGTGTATCGAAATGTAATGTTGGA AAATTATAAAAATCTGGTTTCCCTCGGCTACCAATTGACAAAACCAGAC GTCATCCTCAGACTGGAAAAAGGGGAAGAACCTTGGCTTGTCTGA	3xFLAG-SV40 NLS--ZFP-SV40-NLS-KRAB (DNA 서열)

[1252]

929	ATGGCACCGAAGAAAAAACGCAAAGTTGGTCGAGTACCCGCCGCCATG GCCGAACGACCATTCCAGTGCAGGATATGTATGCGCAATTTTTCAACCA GTGGTTCATTGTCACGACATATTAGAACACACACCGGTGAGAAACCCTT TGCGTGTGACATCTGTGGGAGGAAATTCGCAAGATCTGACAACCTTACG ACACATACAAAGATTCACACAGGCTCTCAAAAGCCCTTCCAGTGCCGAA TTTGCATGCGAAACTTTTCCCAGTCTGGTAATCTCGCTCGACATATCAGA ACCCACACGGGGGAAAAACCATTCGCTTGTGATATTTGCGGACGAAAG TTCGCCGACAGAACCACACTCATGAGACACACTAAAATCCATACTGGTA GTCAGAAGCCGTTTCAGTGTAGAATCTGCATGAGGAACTTTTCCCAGTC AGGCCACCTTGCAAGACATATACGAACTCACACTGGAGAAAAGCCGTT CGCCTGTGACATTTGTGGGCGCAAGTTCGCGCAACTCACCCATCTGAAT AGCCATACGAAGATTCACTTGAGACAGAAAGATGCGGCTCGGGGCTCC GGACCGAAAAAAAAAGAGAAAGGTAAACGGCGGGGGAGGATCACGGAC CCTGGTTACATTTAAGGACGTTTTTGTAGACTTCACGAGGGAGGAATGG AAGCTGTTGGACACCGCCCAGCAAATCGTGTATCGAAATGTAATGTTGG AAAATTATAAAAATCTGGTTTCCCTCGGCTACCAATTGACAAAACCAGA CGTCATCCTCAGACTGGAAAAAGGGGAAGAACCTTGGCTTGTCTGA	3xFLAG-SV40 NLS--ZFP-SV40-NLS-KRAB (DNA 서열)
930	ATGGCACCGAAGAAAAAACGCAAAGTTGGTCGAGTACCCGCAGCTATG GCTGAACGCCCATTCCAGTGTCGGATCTGCATGCGCAACTTTTCTATAA AACACGATCTTCACCGACACATTCGGACACATACTGGGGAGAAGCCCTT TGCGTGTGACATCTGTGGCCGAAAGTTCGCTAGATCCGCAAACTTGACT CGGCATACGAAAATTCACACTGGAAGCCAGAAACCTTTCCAATGTCGA ATCTGTATGAGGAACTTTAGCAGAAGTGATAATCTCGCCAGGCATATCC GAACGCACACAGGCGAGAAGCCATTCGCATGTGATATTTGTGGTAGAA AGTTCGCCCAAAATGTCTCTCGCCCACGCCATACTAAGATCCACACGGG CTCCCAGAAGCCGTTCCAATGCCGCATTTGCATGCGAAACTTTTCCAGA TCAGACGATCTGAGCAAGCATATTAGGACGCATACAGGGGAGAAGCCT TTTGCTTGCGACATTTGCGGCCGGAAATTTGCTGACTCAAGTCACAGAA CACGGCATACCAAGATACACCTTCGACAAAAAGATGCCGCACGGGGCT CCGGACCGAAAAAAAAAGAGAAAGGTAAACGGCGGGGGAGGATCACGG ACCCTGGTTACATTTAAGGACGTTTTTGTAGACTTCACGAGGGAGGAAT GGAAGCTGTTGGACACCGCCCAGCAAATCGTGTATCGAAATGTAATGTT GGAAAATTATAAAAATCTGGTTTCCCTCGGCTACCAATTGACAAAACCA GACGTCATCCTCAGACTGGAAAAAGGGGAAGAACCTTGGCTTGTCTGA	3xFLAG-SV40 NLS--ZFP-SV40-NLS-KRAB (DNA 서열)
931	ATGGCACCGAAGAAAAAACGCAAAGTTGGTCGAGTACCCGCGGCCATG GCGGAACGACCCTTTCAGTGCCGAATTTGCATGAGGAACTTTTCACGAT CTGATAACCTGGCGAGGCACATCCGAACACATACGGGCGAGAAGCCAT TCGCATGTGATATCTGCGGGCGAAAGTTCGCCCAAAATGTCAGTAGACC GCGACATACTAAAATACACACTGGCTCACAGAAGCCGTTCCAATGCCGC ATCTGTATGCGCAATTTTTCCCGAAGCGACGATCTGTCTAAACATATTC GGACGCACACTGGGGAAAAGCCTTTCGCTTGTGACATCTGTGGGAGGA AGTTCGCTGACAGCTCTCATAGGACACGCCATACTAAGATTCATACCGG AAGCCAGAAGCCTTTCCAGTGTCGGATTTGCATGAGAAACTTTAGCACT TCTAGCAACAGAAAGACACATATACGAACCCATACGGGTGAGAAACCG TTCGCATGCGATATCTGTGGGCGAAAATTTGCAGCCCAATGGACCAGAG CTTGCCATACCAAGATACACCTTCGGCAGAAGGACGCTGCACGGGGCTC CGGACCGAAAAAAAAAGAGAAAGGTAAACGGCGGGGGAGGATCACGGA CCCTGGTTACATTTAAGGACGTTTTTGTAGACTTCACGAGGGAGGAATG GAAGCTGTTGGACACCGCCCAGCAAATCGTGTATCGAAATGTAATGTTG GAAAATTATAAAAATCTGGTTTCCCTCGGCTACCAATTGACAAAACCAG ACGTCATCCTCAGACTGGAAAAAGGGGAAGAACCTTGGCTTGTCTGA	3xFLAG-SV40 NLS--ZFP-SV40-NLS-KRAB (DNA 서열)
932	ATGGCACCGAAGAAAAAACGCAAAGTTGGTCGAGTACCCGCTGCGATG GCAGAACGACCTTTTCAATGCCGAATTTGTATGAGGAACTTTTCCCGGT CAGACGACCTTTCCAAGCACATCAGAACTCATACCGGAGAAAAACCGT TCGCCTGTGACATTTGTGGACGGAAGTTTGCTGACTCCTCTCACAGGAC TCGCCACACTAAGATACACACCGGAAGTCAGAAGCCCTTCCAATGTAG GATATGCATGAGAAACTTCAGTACGTCATCAAACCGAAAAACGCATAT CAGGACACATACCGGCGAAAAGCCGTTTGCATGTGATATCTGCGGCAG GAAATTTGCAGCTCAGTGGACACGGGCATGTCACACAAAAATCCATAC	3xFLAG-SV40 NLS--ZFP-SV40-NLS-KRAB (DNA 서열)

[1253]

	CGGTAGTCAAAAACCGTTTCAGTGTCGAATCTGCATGAGGAACTTTAGC CGGAAGCAGACGAGAACCACGCATATAAGAACTCACACAGGTGAGAAA CCCTTTGCGTGCGATATCTGCGGTCGCAAATTTGCTCACCGATCCTCCCT GAGGCGACATACTAAAATACATCTGCGACAGAAAGACGCGGCTCGGGG CTCCGGACCGAAAAAAAAAGAGAAAGGTAAACGGCGGGGGAGGATCAC GGACCCTGGTTACATTTAAGGACGTTTTTGTAGACTTCACGAGGGAGGA ATGGAAGCTGTTGGACACCGCCCAGCAAATCGTGTATCGAAATGTAATG TTGGAAAATTATAAAAATCTGGTTTCCCTCGGCTACCAATTGACAAAAC CAGACGTCATCCTCAGACTGGAAAAAGGGGAAGAACCTTGGCTTGTCT GA	
933	ATGGCACCGAAGAAAAAACGCAAAGTTGGTCGAGTACCCGCCGCCATG GCAGAACGGCCTTTTCAGTGTCGGATCTGCATGAGAAACTTTAGTCAGA GTGCCCATCGCAAGAATCATATTCGAACTCATACCGGTGAAAAACCGTT CGCGTGCGACATCTGTGGTCGAAAGTTTGCCACATCATCCAATAGAAAA ACGCATACTAAGATTCATACCGGTAGCCAGAAACCATTCCAATGTAGAA TCTGCATGCGAAATTTCAGCAGGAGTGACAATTTGTCCGCACATATACG GACACACACGGGCGAAAAACCCTTTGCTTGCGATATATGCGGTAGAAA GTTCGCTAGGAACAACGACCGAAAAACACACACGAAGATTCATACAGG TAGTCAGAAGCCATTTCAATGTCGGATCTGTATGCGAAATTTCTCTACTT CTGGCAGCCTGTCCCGGCACATCAGAACACATACCGGTGAGAAGCCATT TGCATGTGACATATGTGGGAGAAAATTTGCCCAGGCGGGTCACCTTGCG AAGCATACAAAGATCCACCTCCGCCAGAAGGACGCCGCACGGGGCTCC GGACCGAAAAAAAAAGAGAAAGGTAAACGGCGGGGGAGGATCACGGAC CCTGGTTACATTTAAGGACGTTTTTGTAGACTTCACGAGGGAGGAATGG AAGCTGTTGGACACCGCCCAGCAAATCGTGTATCGAAATGTAATGTTGG AAAATTATAAAAATCTGGTTTCCCTCGGCTACCAATTGACAAAACCAGA CGTCATCCTCAGACTGGAAAAAGGGGAAGAACCTTGGCTTGTCTGA	3xFLAG-SV40 NLS--ZFP-SV40-NLS-KRAB (DNA 서열)
934	ATGGCACCGAAGAAAAAACGCAAAGTTGGTCGAGTACCCGCGGCTATG GCAGAGCGGCCATTCCAATGTAGGATATGTATGAGGAACTTCTCCCGGA GCGATCACCTGTCCCAACACATCCGCACGCATACGGGTGAGAAGCCGTT TGCTTGTGATATCTGCGGAAGAAAATTTGCAGCATCCAGTACACGCACA AAGCATACGAAGATTCATACGGGATCCCAAAAGCCCTTTCAATGCAGG ATTTGTATGAGGAACTTCAGTCGGTCCGACGATCTGACACGACATATTA GAACTCATACTGGAGAGAAGCCATTCGCATGTGACATCTGCGGTAGGA AGTTCGCGCAGAAATCTAACCTGTCATCTCACACCAAGATACATACAGG CTCACAGAAGCCGTTTCAATGCCGCATCTGCATGAGGAATTTCAGCCAG TCCGCAAACAGAACTACGCATATTCGGACGCATACCGGCGAGAAGCCG TTTGCCTGCGACATTTGCGGGAGGAAATTCGCACAGAACGCGACCAGA ACCAAACACACCAAAATCCATCTTAGGCAAAAGGATGCGGCCCGGGGC TCCGGACCGAAAAAAAAAGAGAAAGGTAAACGGCGGGGGAGGATCACG GACCCTGGTTACATTTAAGGACGTTTTTGTAGACTTCACGAGGGAGGAA TGGAAGCTGTTGGACACCGCCCAGCAAATCGTGTATCGAAATGTAATGT TGGAAAATTATAAAAATCTGGTTTCCCTCGGCTACCAATTGACAAAACC AGACGTCATCCTCAGACTGGAAAAAGGGGAAGAACCTTGGCTTGTCTG A	3xFLAG-SV40 NLS--ZFP-SV40-NLS-KRAB (DNA 서열)
935	ATGGCACCGAAGAAAAAACGCAAAGTTGGTCGAGTACCCGCGGCCATG GCAGAACGACCCTTTCAGTGCCGAATTTGCATGCGGAACTTTAGTCGCA GTGACACCCTGAGCGAGCATATTCGCACGCATACGGGAGAGAAGCCAT TTGCATGCGACATCTGCGGTAGAAAGTTTGCGAGGCGCTGGACGTTGGT AGGCCACACGAAAATCCATACAGGCTCCCAGAAACCCTTCCAGTGCAG AATTTGTATGCGCAATTTTAGTGACAGAAGTAACTTGTCCCGACATATA AGGACGCACACCGGCGAAAAACCGTTTGCCTGTGATATCTGTGGTCGGA AATTCGCCCAGTCCGGTGACTTGACACGGCATACCAAAATACACACTGG AAGCCAAAAGCCTTTTCAGTGTCGCATATGTATGCGCAACTTCAGCCAG AGTAGTGACCTTTCACGGCATATACGGACGCATACGGGTGAGAAACCCT TCGCCTGTGACATTTGCGGGCGAAAGTTTGCATATCATTGGTACCTGAA AAAACATACGAAAATACATTTGAGACAAAAAGATGCAGCCCGGGGCTC CGGACCGAAAAAAAAAGAGAAAGGTAAACGGCGGGGGAGGATCACGGA CCCTGGTTACATTTAAGGACGTTTTTGTAGACTTCACGAGGGAGGAATG	3xFLAG-SV40 NLS--ZFP-SV40-NLS-KRAB (DNA 서열)

[1254]

	GAAGCTGTTGGACACCGCCCAGCAAATCGTGTATCGAAATGTAATGTTG GAAAATTATAAAAATCTGGTTTCCCTCGGCTACCAATTGACAAAACCAG ACGTCATCCTCAGACTGGAAAAAGGGGAAGAACCTTGGCTTGTCTGA	
936	ATGGCACCGAAGAAAAAACGCAAAGTTGGTCGAGTACCCGCTGCCATG GCCGAGCGCCCGTTTCAATGCAGAATATGTATGAGGAACTTCTCTAGAA GCGCCAATCTTGCGCGACACATTAGAACTCACACAGGCGAGAAACCTTT CGCCTGTGACATCTGCGGCAGAAAATTTGCGCGAAGTGATAACTTGCGC GAGCACACAAAAATCCATACCGGTTCCCAAAAGCCCTTTCAATGTAGGA TTTGCATGAGGAACTTTTCAAGACCATACACGCTGAGACTCCACATTCG CACGCATACGGGAGAAAAACCATTTGCTTGTGACATATGCGGCCGAAA GTTTGCACACCGATCCAACTTGAATAAGCATACCAAAATCCATACGGGG TCTCAGAAACCCTTCCAGTGTCGGATTTGTATGCGAAACTTCTCTCAGA GTGGTTCTCTTACGCGCCACATTAGAACCCACACGGGGGAAAAGCCATT CGCGTGCGATATCTGTGGCCGGAAATTTGCTACTTCCGCAAATCTTTCTC GACATACAAAGATACATCTTAGACAGAAGGATGCGGCACGGGGCTCCG GACCGAAAAAAAAGAGAAAGGTAAACGGCGGGGGAGGATCACGGACC CTGGTTACATTTAAGGACGTTTTTGTAGACTTCACGAGGGAGGAATGGA AGCTGTTGGACACCGCCCAGCAAATCGTGTATCGAAATGTAATGTTGGA AAATTATAAAAATCTGGTTTCCCTCGGCTACCAATTGACAAAACCAGAC GTCATCCTCAGACTGGAAAAAGGGGAAGAACCTTGGCTTGTCTGA	3xFLAG-SV40 NLS--ZFP-SV40-NLS-KRAB (DNA 서열)
937	ATGGCACCGAAGAAAAAACGCAAAGTTGGTCGAGTACCCTTGGAGCCC GGTGAAAAACCATATGCCTGTCCAGAGTGTGGAAAAAGTTTTAGCAGA AGCGACGACCTGGTTAGGCATCAACGAACTCACACAGGGGAAAAGCCG TACAAATGTCCTGAATGCGGAAAGTCATTCTCCACTTCAGGTAGCCTCG TGCGACATCAGCGCACACATACTGGCGAAAAACCTTATAAGTGTCCGG AATGCGGCAAGAGTTTTTCTAGAAGTGATAAGCTCGTGCGACACCAGCG GACGCACACGGGTGAGAAGCCCTACGCCTGCCCAGAGTGCGGAAAGTC CTTTAGTAGGTCTGACGAACTGGTCCGGCACCAACGAACCCATACAGGG GAGAAACCCTACAAATGTCCAGAGTGCGGGAAATCATTTTCAACGTCTC ACAGCTTGACGGAACATCAGAGGACCCATACAGGCGAAAAGCCTTACA AATGTCCGGAGTGCGGAAAGTCTTTCAGCCGGGCTGATAATCTCACGGA GCACCAAAGAACCCACACGGGTAAGAAAACCTCTGGCTCCGGACCGAA AAAAAAGAGAAAGGTAAACGGCGGGGGAGGATCACGGACCCTGGTTA CATTTAAGGACGTTTTTGTAGACTTCACGAGGGAGGAATGGAAGCTGTT GGACACCGCCCAGCAAATCGTGTATCGAAATGTAATGTTGGAAAATTAT AAAAATCTGGTTTCCCTCGGCTACCAATTGACAAAACCAGACGTCATCC TCAGACTGGAAAAAGGGGAAGAACCTTGGCTTGTCTGA	3xFLAG-SV40 NLS--ZFP-SV40-NLS-KRAB (DNA 서열)
938	ATGGCACCGAAGAAAAAACGCAAAGTTGGTCGAGTACCCTTGGAGCCG GGAGAAAAGCCATATGCCTGTCCGGAATGCGGCAAAAGCTTTTCAGAA CGGTCCCATTTGCGGGAACATCAGCGCACCCATACCGGAGAAAAGCCTT ACAAGTGTCCCGAGTGTGGTAAGAGCTTTTCAACTTCCCACAGTCTCAC TGAACATCAGCGAACTCACACAGGCGAAAAACCATACAAATGCCCGGA GTGTGGAAAGAGTTTTTCTCAAGCTGGGCACTTGGCCAGCCACCAAAGG ACTCATACGGGCGAGAAACCGTATGCCTGTCCAGAATGCGGGAAATCA TTTTCTACTAGTCATTCCCTTACGGAACATCAGAGAACGCACACCGGGG AGAAGCCATACAAGTGTCCGGAGTGCGGAAAATCATTCTCCGACCCGG GTCATCTCGTTCGCCATCAGAGGACTCATACTGGAGAGAAACCGTATAA ATGTCCTGAATGCGGGAAGAGCTTTTCTACATCTGGGAACCTTGTGCGG CACCAGCGAACCCACACGGGGAAAAAGACTTCTGGCTCCGGACCGAAA AAAAAGAGAAAGGTAAACGGCGGGGGAGGATCACGGACCCTGGTTAC ATTTAAGGACGTTTTTGTAGACTTCACGAGGGAGGAATGGAAGCTGTTG GACACCGCCCAGCAAATCGTGTATCGAAATGTAATGTTGGAAAATTATA AAAATCTGGTTTCCCTCGGCTACCAATTGACAAAACCAGACGTCATCCT CAGACTGGAAAAAGGGGAAGAACCTTGGCTTGTCTGA	3xFLAG-SV40 NLS--ZFP-SV40-NLS-KRAB (DNA 서열)
939	ATGGCACCGAAGAAAAAACGCAAAGTTGGTCGAGTACCCCTGGAACCG GGTGAAAAACCCTACGCTTGTCCCGAGTGTGGGAAATCATTCTCAAGGG CAGATAATCTTACGGAACATCAGCGCACACACACCGGGGAAAAACCCT ATAAGTGCCCTGAGTGTGGTAAATCATTCTCTACGTCAGGGTCATTGGT GCGCCATCAGAGGACCCATACAGGCGAAAAACCGTATAAGTGCCCCGA	3xFLAG-SV40 NLS--ZFP-SV40-NLS-KRAB (DNA 서열)

[1255]

	ATGCGGTAAAAGCTTCAGCAGAAAAGATAACTTGAAAAATCACCAACG CACTCACACGGGAGAAAAACCGTACGCATGCCCGGAGTGCGGCAAGAG TTTCAGTCAGAGTAGCTCACTTGTTCGCCACCAAAGAACCCACACAGGC GAAAAGCCCTACAAGTGTCCTGAATGTGGAAAGAGTTTCAGTCGAAGT GATAAATTGGTTAGGCACCAAAGAACACACACTGGAGAGAAACCGTAC AAGTGTCCAGAATGTGGCAAGTCTTTTTCTGATTCTGGAAATTTGCGCGT CCATCAAAGGACTCATACTGGGAAAAAGACGTCCGGCTCCGGACCGAA AAAAAAGAGAAAGGTAAACGGCGGGGGAGGATCACGGACCCTGGTTA CATTTAAGGACGTTTTTGTAGACTTCACGAGGGAGGAATGGAAGCTGTT GGACACCGCCCAGCAAATCGTGTATCGAAATGTAATGTTGGAAAATTAT AAAAATCTGGTTTCCCTCGGCTACCAATTGACAAAACCAGACGTCATCC TCAGACTGGAAAAAGGGGAAGAACCTTGGCTTGTCTGA	
940	ATGGCACCGAAGAAAAAACGCAAAGTTGGTCGAGTACCCCTCGAACCA GGGGAAAAGCCATACGCTTGCCCCGAGTGTGGAAAATCTTTCTCTCAGT CAAGCTCCCTTGTCAGACACCAGAGAACCCATACGGGTGAGAAACCAT ACAAGTGCCCAGAGTGCGGCAAAAGCTTCAGTCAGAGTGGCGATCTGC GCCGGCATCAAAGAACTCACACTGGAGAAAAGCCCTATAAGTGCCCCG AATGTGGAAAGAGTTTTTCCAGGAGTGACGAAAGAAAGAGACACCAGC GGACTCACACCGGCGAGAAACCATACGCATGCCCCGAGTGCGGAAAAA GTTTTTCCCACCGCACAACCCTTACCAATCATCAACGCACGCACACAGG GGAGAAACCCTACAAGTGCCCGGAGTGTGGCAAGTCATTCAGCCGAAG TGACCACCTCACTAACCACCAAAGGACTCACACAGGAGAAAAACCCTA CAAATGTCCGGAGTGCGGAAAATCTTTTTCCACGTCCGGTGAGCTGGTC CGCCATCAACGAACCCATACTGGTAAAAAAACTAGCGGCTCCGGACCG AAAAAAAAGAGAAAGGTAAACGGCGGGGGAGGATCACGGACCCTGGT TACATTTAAGGACGTTTTTGTAGACTTCACGAGGGAGGAATGGAAGCTG TTGGACACCGCCCAGCAAATCGTGTATCGAAATGTAATGTTGGAAAATT ATAAAAATCTGGTTTCCCTCGGCTACCAATTGACAAAACCAGACGTCAT CCTCAGACTGGAAAAAGGGGAAGAACCTTGGCTTGTCTGA	3xFLAG-SV40 NLS--ZFP-SV40-NLS-KRAB (DNA 서열)
941	ATGGCACCGAAGAAAAAACGCAAAGTTGGTCGAGTACCCCTGGAACCC GGGGAGAAGCCGTACGCTTGCCCAGAGTGCGGAAAGAGCTTCTCTCAG AGCGGAGACCTTAGACGCCACCAGCGAACCCACACCGGCGAAAAACCG TATAAATGCCCGGAATGCGGCAAGAGTTTTAGTCGGTCCGATGAGCGA AAGAGGCATCAACGAACCCATACGGGAGAGAAACCCTACAAGTGCCCT GAGTGTGGTAAGTCATTTTCCCACAGAACGACGTTGACGAATCACCAGA GAACCCATACGGGTGAGAAACCTTACGCTTGCCCGGAGTGCGGCAAAA GCTTCAGCCGGAGTGATCACTTGACCAATCATCAGAGGACACACACGG GTGAGAAGCCCTACAAATGTCCCGAATGCGGCAAGTCTTTCTCAACGTC AGGCGAACTCGTCCGGCACCAGCGAACACATACGGGAGAAAAGCCGTA CAAATGTCCGGAATGCGGAAAGTCATTTTCACGGTCAGATGACTTGGTG CGACACCAGCGCACTCACACAGGCAAGAAGACCTCAGGCTCCGGACCG AAAAAAAAGAGAAAGGTAAACGGCGGGGGAGGATCACGGACCCTGGT TACATTTAAGGACGTTTTTGTAGACTTCACGAGGGAGGAATGGAAGCTG TTGGACACCGCCCAGCAAATCGTGTATCGAAATGTAATGTTGGAAAATT ATAAAAATCTGGTTTCCCTCGGCTACCAATTGACAAAACCAGACGTCAT CCTCAGACTGGAAAAAGGGGAAGAACCTTGGCTTGTCTGA	3xFLAG-SV40 NLS--ZFP-SV40-NLS-KRAB (DNA 서열)
942	ATGGCACCGAAGAAAAAACGCAAAGTTGGTCGAGTACCCCTGGAGCCT GGTGAGAAGCCGTATGCATGTCCTGAGTGTGGGAAGTCATTTAGTCAGA GGGCCCACTTGGAACGACACCAAAGGACCCACACTGGTGAAAAACCCT ACAAATGCCCAGAGTGTGGTAAGTCTTTTTCACAGCTGGCCCACCTGAG AGCACACCAGCGAACTCATACGGGCGAGAAACCATACAAGTGTCCAGA GTGCGGAAAGTCATTCTCAGATCCCGGCCACTTGGTGCGACATCAGAGA ACGCACACAGGGGAGAAGCCTTATGCTTGCCCGGAATGCGGGAAGTCT TTCAGCCGCCGAAGTGCTTGTCGAAGGCACCAACGGACCCATACCGGTG AGAAACCATATAAGTGCCCAGAGTGTGGAAAGAGTTTTAGTCGATCCG ATCACCTGACTACGCACCAGCGGACGCACACAGGAGAGAAACCGTATA AGTGCCCTGAATGCGGTAAGAGCTTCTCTCAATCAAGCTCACTGGTTAG GCACCAACGCACTCATACCGGCAAGAAGACGTCAGGCTCCGGACCGAA AAAAAAGAGAAAGGTAAACGGCGGGGGAGGATCACGGACCCTGGTTA	3xFLAG-SV40 NLS--ZFP-SV40-NLS-KRAB (DNA 서열)

[1256]

	CATTTAAGGACGTTTTTGTAGACTTCACGAGGGAGGAATGGAAGCTGTT GGACACCGCCCAGCAAATCGTGTATCGAAATGTAATGTTGGAAAATTAT AAAAATCTGGTTTCCCTCGGCTACCAATTGACAAAACCAGACGTCATCC TCAGACTGGAAAAAGGGGAAGAACCTTGGCTTGTCTGA	
943	ATGGCACCGAAGAAAAAACGCAAAGTTGGTCGAGTACCCCTTGAGCCG GGGGAAAAGCCTTATGCTTGCCCAGAGTGTGGCAAGAGCTTCTCCCAAA GTTCAAACCTCGTCCGACACCAAAGGACTCACACGGGCGAAAAACCGT ATAAATGCCCCGAGTGCGGAAAGTCATTTTCCAGGTCCGACGATCTGGT CCGCCACCAGCGCACTCATACGGGGGAAAAGCCCTATAAGTGCCCTGA GTGCGGCAAGAGCTTCTCAACTCACCTGGATCTCATTCGCCATCAACGC ACACATACAGGGGAGAAGCCTTACGCTTGTCCAGAGTGCGGCAAGTCTT TCAGTACGAGCGGAAACCTGACGGAACACCAGCGAACCCACACGGGCG AGAAGCCATATAAATGTCCCGAATGTGGAAAATCATTCTCTCGCCGATC TGCGTGCCGCCGGCATCAGAGGACACATACCGGAGAAAAGCCGTACAA ATGCCCCGAGTGTGGAAAATCCTTTAGCAGAAATGATACACTTACCGAG CATCAGAGGACGCACACTGGAAAAAAGACATCTGGCTCCGGACCGAAA AAAAAGAGAAAGGTAAACGGCGGGGGAGGATCACGGACCCTGGTTAC ATTTAAGGACGTTTTTGTAGACTTCACGAGGGAGGAATGGAAGCTGTTG GACACCGCCCAGCAAATCGTGTATCGAAATGTAATGTTGGAAAATTATA AAAATCTGGTTTCCCTCGGCTACCAATTGACAAAACCAGACGTCATCCT CAGACTGGAAAAAGGGGAAGAACCTTGGCTTGTCTGA	3xFLAG-SV40 NLS--ZFP-SV40-NLS-KRAB (DNA 서열)
944	MAPKKKRKVGRVPAAMAERPFQCRICMRNFSSEADRSRHIRTHTGEKPFA CDICGRKFADRSNLTRHTKIHTGSQKPFQCRICMRNFSQSSDLSRHIRTHTG EKPFACDICGRKFAYHWYLKKHTKIHTGSQKPFQCRICMRNFSRSDSLSVHI RTHTGEKPFACDICGRKFAQNANRKTHTKIHLRQKDAARGSGPKKKRKVN GGGGSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSLG YQLTKPDVILRLEKGEEPWLV	3xFLAG-SV40 NLS--ZFP-SV40-NLS-KRAB 융합 (AA 서열)
945	MAPKKKRKVGRVPAAMAERPFQCRICMRNFSRSDVLSTHIRTHTGEKPFA CDICGKKFADNSSRTRHTKIHTGSQKPFQCRICMRNFSRPYTLRLHIRTHTG EKPFACDICGRKFADSSHRTRHTKIHTGSQKPFQCRICMRNFSRSDHLSQHI RTHTGEKPFACDICGRKFADSSHRTRHTKIHLRQKDAARGSGPKKKRKVN GGGGSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSLG YQLTKPDVILRLEKGEEPWLV	3xFLAG-SV40 NLS--ZFP-SV40-NLS-KRAB 융합 (AA 서열)
946	MAPKKKRKVGRVPAAMAERPFQCRICMRNFSRSDHLSQHIRTHTGEKPFA CDICGRKFAQSADRTKHTKIHTGSQKPFQCRICMRNFSRSDHLSQHIRTHTG EKPFACDICGRKFARRSDLKRHTKIHTGSQKPFQCRICMRNFSRSDHLSRHI RTHTGEKPFACDICGRKFAQSSDLRRHTKIHLRQKDAARGSGPKKKRKVN GGGGSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSLG YQLTKPDVILRLEKGEEPWLV	3xFLAG-SV40 NLS--ZFP-SV40-NLS-KRAB 융합 (AA 서열)
947	MAPKKKRKVGRVPAAMAERPFQCRICMRNFSRSDNLSEHIRTHTGEKPFA CDICGRKFATSSNRKTHTKIHTGSQKPFQCRICMRNFSDRSHLTRHIRTHTG EKPFACDICGRKFARSDALTQHTKIHTGSQKPFQCRICMRNFSDRSALARHI RTHTGEKPFACDICGRKFARRFTLSKHTKIHLRQKDAARGSGPKKKRKVNG GGGSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSLGY QLTKPDVILRLEKGEEPWLV	3xFLAG-SV40 NLS--ZFP-SV40-NLS-KRAB 융합 (AA 서열)
948	MAPKKKRKVGRVPAAMAERPFQCRICMRNFSRSDHLSEHIRTHTGEKPFA CDICGRKFAQYSGRYYHTKIHTGSQKPFQCRICMRNFSHGQTLNEHIRTHT GEKPFACDICGRKFAQSGNLARHTKIHTGSQKPFQCRICMRNFSRSDSLLRH IRTHTGEKPFACDICGRKFACREYRGKHTKIHLRQKDAARGSGPKKKRKVN GGGGSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSLG YQLTKPDVILRLEKGEEPWLV	3xFLAG-SV40 NLS--ZFP-SV40-NLS-KRAB 융합 (AA 서열)
949	MAPKKKRKVGRVPAAMAERPFQCRICMRNFSQSANRTTHIRTHTGEKPFA CDICGRKFARSANLTRHTKIHTGSQKPFQCRICMRNFSRSDVLSEHIRTHTG EKPFACDICGRKFATSGHLSRHTKIHTGSQKPFQCRICMRNFSQSSDLSRHIR THTGEKPFACDICGRKFAQWSTRKRHTKIHLRQKDAARGSGPKKKRKVNG GGGSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSLGY QLTKPDVILRLEKGEEPWLV	3xFLAG-SV40 NLS--ZFP-SV40-NLS-KRAB 융합 (AA 서열)
950	MAPKKKRKVGRVPAAMAERPFQCRICMRNFSQSGNLARHIRTHTGEKPFA CDICGRKFAATCCLAHHTKIHTGSQKPFQCRICMRNFSRWQYLPTHIRTHA	3xFLAG-SV40 NLS--ZFP-

[1257]

	GEKPFACDICGRKFADRSALARHTKIHTGSQKPFQCRICMRNFSRSDNLSEHIRTHTGEKPFACDICGRKFAKRCNLRCHTKIHLRQKDAARGSGPKKKRKVNGGGGSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSLGYQLTKPDVILRLEKGEEPWLV	SV40-NLS-KRAB 융합 (AA 서열)
951	MAPKKKRKVGRVPAAMAERPFQCRICMRNFSNPANLTRHIRTHTGEKPFACDICGRKFAQNATRTKHTKIHTGSQKPFQCRICMRNFSQSGHLARHIRTHTGEKPFACDICGRKFANRHDRAKHTKIHTGSQKPFQCRICMRNFSRSDHLSEHIRTHTGEKPFACDICGRKFAQRRSRYKHTKIHLRQKDAARGSGPKKKRKVNGGGGSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSLGYQLTKPDVILRLEKGEEPWLV	3xFLAG-SV40 NLS--ZFP-SV40-NLS-KRAB 융합 (AA 서열)
952	MAPKKKRKVGRVPAAMAERPFQCRICMRNFSQSSDLSRHIRTHTGEKPFACDICGRKFAHRSTRNRHTKIHTGSQKPFQCRICMRNFSRSDVLSAHIRTHTGEKPFACDICGRKFADSRTRKNHTKIHTGSQKPFQCRICMRNFSQSGSLTRHIRTHTGEKPFACDICGRKFADQSGLAHHTKIHLRQKDAARGSGPKKKRKVNGGGGSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSLGYQLTKPDVILRLEKGEEPWLV	3xFLAG-SV40 NLS--ZFP-SV40-NLS-KRAB 융합 (AA 서열)
953	MAPKKKRKVGRVPAAMAERPFQCRICMRNFSQNPAQWRHIRTHTGEKPFACDICGRKFARSADLSRHTKIHTGSQKPFQCRICMRNFSTSGSLSRHIRTHTGEKPFACDICGRKFARSDHLSRHTKIHTGSQKPFQCRICMRNFSRSDSLLRHIRTHTGEKPFACDICGRKFAQSYDRFQHTKIHLRQKDAARGSGPKKKRKVNGGGGSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSLGYQLTKPDVILRLEKGEEPWLV	3xFLAG-SV40 NLS--ZFP-SV40-NLS-KRAB 융합 (AA 서열)
954	MAPKKKRKVGRVPAAMAERPFQCRICMRNFSTSGSLSRHIRTHTGEKPFACDICGRKFARSDHLSRHTKIHTGSQKPFQCRICMRNFSRSDSLLRHIRTHTGEKPFACDICGRKFAQSYDRFQHTKIHTGSQKPFQCRICMRNFSRSDNLSTHIRTHTGEKPFACDICGRKFADNRDRIKHTKIHLRQKDAARGSGPKKKRKVNGGGGSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSLGYQLTKPDVILRLEKGEEPWLV	3xFLAG-SV40 NLS--ZFP-SV40-NLS-KRAB 융합 (AA 서열)
955	MAPKKKRKVGRVPAAMAERPFQCRICMRNFSDRSNLSRHIRTHTGEKPFACDICGRKFALRQNLIMHTKIHTGSQKPFQCRICMRNFSERGTLARHIRTHTGEKPFACDICGRKFARSDALTQHTKIHTGSQKPFQCRICMRNFSRSDSLSQHIRTHTGEKPFACDICGRKFARKADRTRHTKIHLRQKDAARGSGPKKKRKVNGGGGSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSLGYQLTKPDVILRLEKGEEPWLV	3xFLAG-SV40 NLS--ZFP-SV40-NLS-KRAB 융합 (AA 서열)
956	MAPKKKRKVGRVPAAMAERPFQCRICMRNFSQYCCLTNHIRTHTGEKPFACDICGRKFATSGNLTRHTKIHTGSQKPFQCRICMRNFSQSSDLSRHIRTHTGEKPFACDICGRKFAFRYYLKRHTKIHTGSQKPFQCRICMRNFSQSGDLTRHIRTHTGEKPFACDICGRKFADKGNLTKHTKIHLRQKDAARGSGPKKKRKVNGGGGSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSLGYQLTKPDVILRLEKGEEPWLV	3xFLAG-SV40 NLS--ZFP-SV40-NLS-KRAB 융합 (AA 서열)
957	MAPKKKRKVGRVPAAMAERPFQCRICMRNFSTSGSLSRHIRTHTGEKPFACDICGRKFARSDNLTTHTKIHTGSQKPFQCRICMRNFSQSGNLARHIRTHTGEKPFACDICGRKFADRTTLMRHTKIHTGSQKPFQCRICMRNFSQSGHLARHIRTHTGEKPFACDICGRKFAQLTHLNSHTKIHLRQKDAARGSGPKKKRKVNGGGGSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSLGYQLTKPDVILRLEKGEEPWLV	3xFLAG-SV40 NLS--ZFP-SV40-NLS-KRAB 융합 (AA 서열)
958	MAPKKKRKVGRVPAAMAERPFQCRICMRNFSIKHDLHRHIRTHTGEKPFACDICGRKFARSANLTRHTKIHTGSQKPFQCRICMRNFSRSDNLARHIRTHTGEKPFACDICGRKFAQNVSRPRHTKIHTGSQKPFQCRICMRNFSRSDDLSKHIRTHTGEKPFACDICGRKFADSSHRTRHTKIHLRQKDAARGSGPKKKRKVNGGGGSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSLGYQLTKPDVILRLEKGEEPWLV	3xFLAG-SV40 NLS--ZFP-SV40-NLS-KRAB 융합 (AA 서열)
959	MAPKKKRKVGRVPAAMAERPFQCRICMRNFSRSDNLARHIRTHTGEKPFACDICGRKFAQNVSRPRHTKIHTGSQKPFQCRICMRNFSRSDDLSKHIRTHTGEKPFACDICGRKFADSSHRTRHTKIHTGSQKPFQCRICMRNFSTSSNRKTHIRTHTGEKPFACDICGRKFAAQWTRACHTKIHLRQKDAARGSGPKKKRKVNGGGGSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSLGYQLTKPDVILRLEKGEEPWLV	3xFLAG-SV40 NLS--ZFP-SV40-NLS-KRAB 융합 (AA 서열)

[1258]

960	MAPKKKRKVGRVPAAMAERPFQCRICMRNFSRSDDLSKHIRTHTGEKPFA CDICGRKFADSSHRTRHTKIHTGSQKPFQCRICMRNFSTSSNRKTHIRTHTG EKPFACDICGRKFAAQWTRACHTKIHTGSQKPFQCRICMRNFSRKQTRTTH IRTHTGEKPFACDICGRKFAHRSSLRRHTKIHLRQKDAARGSGPKKKRKVN GGGGSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSLG YQLTKPDVILRLEKGEEPWLV	3xFLAG-SV40 NLS--ZFP- SV40-NLS- KRAB 융합 (AA 서열)
961	MAPKKKRKVGRVPAAMAERPFQCRICMRNFSQSAHRKNHIRTHTGEKPFA CDICGRKFATSSNRKTHTKIHTGSQKPFQCRICMRNFSRSDNLSAHIRTHTG EKPFACDICGRKFARNNDRKTHTKIHTGSQKPFQCRICMRNFSTSGSLSRHI RTHTGEKPFACDICGRKFAQAGHLAKHTKIHLRQKDAARGSGPKKKRKVN GGGGSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSLG YQLTKPDVILRLEKGEEPWLV	3xFLAG-SV40 NLS--ZFP- SV40-NLS- KRAB 융합 (AA 서열)
962	MAPKKKRKVGRVPAAMAERPFQCRICMRNFSRSDHLSQHIRTHTGEKPFA CDICGRKFAASSTRTKHTKIHTGSQKPFQCRICMRNFSRSDDLTRHIRTHTG EKPFACDICGRKFAQKSNLSSHTKIHTGSQKPFQCRICMRNFSQSANRTTHI RTHTGEKPFACDICGRKFAQNATRTKHTKIHLRQKDAARGSGPKKKRKVN GGGGSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSLG YQLTKPDVILRLEKGEEPWLV	3xFLAG-SV40 NLS--ZFP- SV40-NLS- KRAB 융합 (AA 서열)
963	MAPKKKRKVGRVPAAMAERPFQCRICMRNFSRSDTLSEHIRTHTGEKPFAC DICGRKFARRWTLVGHTKIHTGSQKPFQCRICMRNFSDRSNLSRHIRTHTGE KPFACDICGRKFAQSGDLTRHTKIHTGSQKPFQCRICMRNFSQSSDLSRHIR THTGEKPFACDICGRKFAYHWYLKKHTKIHLRQKDAARGSGPKKKRKVN GGGGSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSLG YQLTKPDVILRLEKGEEPWLV	3xFLAG-SV40 NLS--ZFP- SV40-NLS- KRAB 융합 (AA 서열)
964	MAPKKKRKVGRVPAAMAERPFQCRICMRNFSRSANLARHIRTHTGEKPFA CDICGRKFARSDNLREHTKIHTGSQKPFQCRICMRNFSRPYTLRLHIRTHTG EKPFACDICGRKFAHRSNLNKHTKIHTGSQKPFQCRICMRNFSQSGSLTRHI RTHTGEKPFACDICGRKFATSANLSRHTKIHLRQKDAARGSGPKKKRKVN GGGGSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSLG YQLTKPDVILRLEKGEEPWLV	3xFLAG-SV40 NLS--ZFP- SV40-NLS- KRAB 융합 (AA 서열)
965	MAPKKKRKVGRVPLEPGEKPYACPECGKSFSRSDDLVRHQRTHTGEKPYK CPECGKSFSTSGSLVRHQRTHTGEKPYKCPECGKSFSRSDKLVRHQRTHTG EKPYACPECGKSFSRSDELVRHQRTHTGEKPYKCPECGKSFSTSHSLTEHQR THTGEKPYKCPECGKSFSRADNLTEHQRTHTGKKTSGSGPKKKRKVNGGG GSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSLGYQLT KPDVILRLEKGEEPWLV	3xFLAG-SV40 NLS--ZFP- SV40-NLS- KRAB 융합 (AA 서열)
966	MAPKKKRKVGRVPLEPGEKPYACPECGKSFSERSHLREHQRTHTGEKPYK CPECGKSFSTSHSLTEHQRTHTGEKPYKCPECGKSFSQAGHLASHQRTHTG EKPYACPECGKSFSTSHSLTEHQRTHTGEKPYKCPECGKSFSDPGHLVRHQ RTHTGEKPYKCPECGKSFSTSGNLVRHQRTHTGKKTSGSGPKKKRKVNGG GGSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSLGYQL TKPDVILRLEKGEEPWLV	3xFLAG-SV40 NLS--ZFP- SV40-NLS- KRAB 융합 (AA 서열)
967	MAPKKKRKVGRVPLEPGEKPYKCPECGKSFSRADNLTEHQRTHTGEKPYK CPECGKSFSTSGSLVRHQRTHTGEKPYKCPECGKSFSRKDNLKNHQRTHTG EKPYKCPECGKSFSQSSSLVRHQRTHTGEKPYKCPECGKSFSRSDKLVRHQ RTHTGEKPYKCPECGKSFSDSGNLRVHQRTHTGKKTSGSGPKKKRKVNGG GGSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSLGYQL TKPDVILRLEKGEEPWLV	3xFLAG-SV40 NLS--ZFP- SV40-NLS- KRAB 융합 (AA 서열)
968	MAPKKKRKVGRVPLEPGEKPYACPECGKSFSQSSSLVRHQRTHTGEKPYK CPECGKSFSQSGDLRRHQRTHTGEKPYKCPECGKSFSRSDERKRHQRTHTG EKPYACPECGKSFSHRTTLTNHQRTHTGEKPYKCPECGKSFSRSDHLTNHQ RTHTGEKPYKCPECGKSFSTSGELVRHQRTHTGKKTSGSGPKKKRKVNGG GGSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSLGYQL TKPDVILRLEKGEEPWLV	3xFLAG-SV40 NLS--ZFP- SV40-NLS- KRAB 융합 (AA 서열)
969	MAPKKKRKVGRVPLEPGEKPYACPECGKSFSQSGDLRRHQRTHTGEKPYK CPECGKSFSRSDERKRHQRTHTGEKPYKCPECGKSFSHRTTLTNHQRTHTG EKPYACPECGKSFSRSDHLTNHQRTHTGEKPYKCPECGKSFSTSGELVRHQ RTHTGEKPYKCPECGKSFSRSDDLVRHQRTHTGKKTSGSGPKKKRKVNGG	3xFLAG-SV40 NLS--ZFP- SV40-NLS- KRAB 융합 (AA 서열)

[1259]

	GGSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSLGYQL TKPDVILRLEKGEEPWLV	
970	MAPKKKRKVGRVPLEPGEKPYACPECGKSFSQRAHLERHQRTHTGEKPYK CPECGKSFSQLAHLRAHQRTHTGEKPYKCPECGKSFSDPGHLVRHQRTHTG EKPYACPECGKSFSRRSACRRHQRTHTGEKPYKCPECGKSFSRSDHLTTHQ RTHTGEKPYKCPECGKSFSQSSSLVRHQRTHTGKKTSGSGPKKKRKVNGG GGSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSLGYQL TKPDVILRLEKGEEPWLV	3xFLAG-SV40 NLS--ZFP- SV40-NLS- KRAB 융합 (AA 서열)
971	MAPKKKRKVGRVPLEPGEKPYACPECGKSFSQSSNLVRHQRTHTGEKPYK CPECGKSFSRSDDLVRHQRTHTGEKPYKCPECGKSFSTHLDLIRHQRTHTG EKPYACPECGKSFSTSGNLTEHQRTHTGEKPYKCPECGKSFSRRSACRRHQ RTHTGEKPYKCPECGKSFSRNDTLTEHQRTHTGKKTSGSGPKKKRKVNGG GGSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSLGYQL TKPDVILRLEKGEEPWLV	3xFLAG-SV40 NLS--ZFP- SV40-NLS- KRAB 융합 (AA 서열)

[1260]

[1261] **실시예 9: dSpCas9-KRAB 및 eZFP-KRAB 융합 단백질에 의해 유도된 억제의 비교**

[1262] [0618] eZFP-KRAB 융합 단백질에 의해 유도된 전사 억제 수준을 dSpCas9-KRAB 및 gRNA 시스템과 비교하기 위해, Hep3B 세포를 서열번호 961, 962 또는 968에 제시된 가장 효과적인 eZFP-KRAB 융합 단백질을 인코딩하는 mRNA와 gRNA HBVg_22(서열번호 412) 또는 HBVg_63(서열번호 453)을 인코딩하는 mRNA로 리포펙타민을 통해 형질

도입하였다. 비표적화(NT) gRNA, GFP 및 ZFP-NT(KRAB 도메인을 포함하는 비표적화 eZFP)는 음성 대조군으로 포함되었다. eZFP-KRAB 융합체 또는 dSpCas9-KRAB 단백질을 발현하는 세포를 3일 후 HBV RNA 수준에 대해 qPCR로 평가하였다. 도 30은 i) dSpCas9-KRAB 융합 단백질(서열번호 595)과 조합한 HBVg_22(서열번호 412), ii) dSpCas9-KRAB 융합 단백질(서열번호 595)과 조합한 HBVg_63(서열번호 453), 및 iii) 서열번호 961, 962 및 968에 제시된 가장 효과적인 eZFP-KRAB 융합 단백질에 의해 유도된 억제 간의 예비 비교를 보여준다. 도 30에 도시된 바와 같이 ZFP-KRAB 융합 단백질은 dSpCas9-KRAB 융합 단백질에 의해 매개된 억제와 필적할만한 전사 억제를 유도하였다.

[1263] [0619] 도 31A 및 31B는 HB X-프로모터(HBx) 시작 부위를 따라 각 eZFP의 표적화 위치를 나타낸다. 도 31A-31B에 도시된 바와 같이, 가장 효과적인 eZFP(서열번호 961, 962 및 968에 제시됨)는 gRNA HBVg_22(서열번호 412) 및 HBVg_63(서열번호 453)이 표적화하는 것과 동일한 일반적인 영역을 표적화하였다. 이 데이터는 두 DNA-결합 도메인을 사용하여 상당한 억제가 달성되었음을 보여준다. 이 결과는 HBV 게놈의 1250개 염기쌍(bp)에서 1374개 bp 사이의 위치, 특히 1260bp에서 1300bp 사이의 영역을 표적화하여 HBV의 전사 억제를 유도하는 것이 견고함을 보여준다.

[1264] **실시예 10: HepG2.NTCP 진성(true) 감염 모델에서 eZFP-KRAB 융합체에 의해 달성된 cccDNA의 강력한 억제**

[1265] [0620] 실시예 8에서 확인된 28개의 eZFP-KRAB 융합 단백질을 HepG2.NTCP 진성 감염 모델에서 스크리닝하였다. HepG2.NTCP 세포는 음성 cccDNA의 생성 및 존재로 인해 HBV 감염이 지속되므로 HBV의 진성 감염 모델을 가능하게 한다.

[1266] [0621] 상위 28개 융합 단백질을 인코딩하는 mRNA를 생성하여 리포펙타민을 통해 두 배치로 HepG2.NTCP 세포에 형질도입시켰다. eZFP-KRAB 융합 단백질을 발현하는 세포는 리포펙션 후 6일째에 qPCR로 HBV RNA 수치를 평가하였다. 도 32A-32B는 HepG2.NTCP 세포에서 각 eZFP-KRAB 융합 단백질을 매개로 한 총 HBV RNA의 배수 변화를 보여준다. 도 32A-32B 에서 볼 수 있듯이, 여러 eZFP-KRAB 융합 단백질은 6일째까지 cccDNA에서 HBV 전유전체(pgRNA)를 효과적으로 억제하였다. 도 33은 HBx 프로모터 위치에서 각 eZFP-KRAB 융합 단백질을 매개로 한 총 HBV RNA의 배수 변화를 보여준다. eZFP 융합 단백질은 HBV 아형 전체에서 보존 비율에 따라 색상으로 구분된다. 도 33에서 볼 수 있듯이, 여러 eZFP-KRAB 융합 단백질은 cccDNA의 HBx 프로모터 영역 전반에 걸쳐 다양한 부위를 표적으로 삼고 전사 억제를 매개하였다. 구체적으로, 가장 우수한 eZFP-KRAB 억제자 단백질 중 하나(eZFP_18)는 HBV 아형 전체에서 약 91%의 보존율을 나타내었다.

[1267] [0622] 종합적으로, 이들 결과는 eZFP-KRAB 억제 단백질에 의한 HBV 발현의 강력한 억제를 보여준다.

[1268] **실시예 11: DNMT3A/L-eZFP-KRAB 융합 단백질에 의해 달성된 지속적인 HBV 억제.**

[1269] [0623] DNMT3A/L(서열번호 651), KRAB(서열번호 590) 및 가장 효율적인 12개의 eZFP 중 하나를 포함하는 융합 단백질을 전사 억제가 가장 높은 DNMT3A/L-eZFP-KRAB 융합 단백질을 식별하기 위해 설계하였다. 융합 단백질을 인코딩하는 mRNA를 생성하여 리포펙타민을 통해 Hep3B 세포로 형질도입시켰다. DNMT3A/L(서열번호 651) 및 KRAB(서열번호 590)에 연결된 eZFP를 포함하는 DNMT3A/L-eZFP-KRAB 융합 단백질(표 E6)을 발현하는 세포는 4일 또는 15일 후 HBV RNA 수치에 대해 qPCR로 평가되었다. 도 34는 각 DNMT3A/L-eZFP-KRAB 융합 단백질에 의해 매개되는 총 HBV RNA의 배수 변화를 보여준다. 도 34에 도시된 바와 같이, 몇몇 융합 단백질은 4일 및 15일에 전사 억제를 매개하였다. 구체적으로, eZFP_18(서열번호 1017) 및 eZFP_19(서열번호 1018)를 포함하는 DNMT3A/L-eZFP-KRAB 융합 단백질은 4일까지 가장 높은 전사 억제를 매개했으며, 이는 15일까지 지속되었다. 이들 융합 단백질에 의해 매개된 억제 수준은 HBVg_22(서열번호 412)와 조합된 DNMT3A/L-dSpCas9-KRAB(서열번호 645)에 의해 매개된 수준과 유사하였다.

[1270] [0624] DNMT3A/L-eZFP-KRAB 융합 단백질에 의해 매개되는 차별적으로 발현되는 유전자(DEG)를 평가하기 위해 HepG2.NTCP 세포에 DNMT3A/L-eZFP-KRAB 융합 단백질을 인코딩하는 mRNA를 형질도입시켰다. RNA를 72시간 후에 수확하고 RNA-시퀀싱을 수행하였다. 도 35는 리포펙션 후 72시간에 융합 단백질에 의해 매개되는 차별적으로 발현되는 유전자의 변화를 보여준다. 결과에 따르면 eZFP_18(서열번호 1017)과 eZFP_19(서열번호 1018)로 구성된 DNMT3A/L-eZFP-KRAB 융합 단백질은 지질 대조군에 비해 최소한의 오프-표적 차별 발현을 유도하였다. 또한, dSpCas9-KRAB(서열번호 594)와 함께 HBVg_22(서열번호 412)를 전달한 후에는 유전자 발현이 크게 변화하지 않았다. 도 36A-36B는 리포펙션 72시간 후 DNMT3A/L-eZFP-KRAB 융합 단백질을 매개로 한 차별적으로 발현된 유전자의 변화를 GFP(도 36A) 또는 비표적화(NT) 대조군(도 36B)과 비교한 것이다. eZFP_18(서열번호 1017)로 구성된 DNMT3A/L-eZFP-KRAB 융합 단백질은 GFP 단독 또는 비표적화 ZFP-KRAB_NT 융합 대조군과 비교했을 때 오프-표적

차등 발현이 최소이거나 전혀 없었다.

[1271] [0625] 이러한 결과를 종합해 보면, eZFP_18(서열번호 1017)과 eZFP_19(서열번호 1018)로 구성된 DNMT3A/L-eZFP-KRAB 억제자 융합 단백질에 의한 HBV RNA의 지속적인 전사 억제가 오프-표적 유전자 발현이 최소이거나 전혀 없는 가장 높은 전사 억제를 매개한다는 것을 알 수 있다. 일반적으로 이러한 데이터는 HBV 발현의 표적 억제를 위한 융합 단백질을 포함하여 eZFP의 치료적 유용성을 뒷받침한다.

표 E6. DNMT3A/L-eZFP-KRAB 융합 단백질

SEQ ID NOs	서열	설명
972	ATGCACGATCAGGAGTTTGACCCCCCTAAGGTGTACCCACCCGTGCCAGCCG AGAAGAGGAAGCCCATCCGCGTGCTGTCCCTGTTCGACGGCATCGCCACAGG CCTGCTGGTGCTGAAGGATCTGGGCATCCAGGTGGACAGATATATCGCCTCC GAGGTGTGCGAGGATTCTATCACCGTGGGCATGGTGAGGCACCAGGGCAAG ATCATGTACGTGGGCGACGTGCGCAGCGTGACACAGAAGCACATCCAGGAGT GGGGACCCTTCGACCTGGTCATCGGAGGCAGCCCCTGTAATGACCTGTCCAT CGTGAACCCTGCAAGGAAGGGCCTGTATGAGGGAACCGGCAGACTGTTCTTT GAGTTCTACAGGCTGCTGCACGACGCCCGCCCTAAGGAGGGCGATGACAGGC CATTCTTTTGGCTGTTTGAGAACGTGGTGGCCATGGGCGTGAGCGACAAGCG GGATATCTCCAGATTCCTGGAGTCTAATCCCGTGATGATCGATGCAAAGGAG GTGTCTGCCGCACACAGGGCAAGGTACTTTTGGGGAAATCTGCCTGGCATGA ACCGCCCACTGGCCAGCACCGTGAACGACAAGCTGGAGCTGCAGGAGTGCCT GGAGCACGGAAGGATCGCCAAGTTCTCCAAGGTGCGGACAATCACCACAAG ATCTAACAGCATCAAGCAGGGCAAGGATCAGCACTTCCCCGTGTTCATGAAT GAGAAGGAGGACATCCTGTGGTGTACCGAGATGGAGCGCGTGTTCGGCTTTC CAGTGCACTATACAGACGTGAGCAATATGAGCCGGCTGGCAAGGCAGAGAC TGCTGGGCCGGTCCTGGTCTGTGCCAGTGATCAGACACCTGTTCGCCCCCCTG AAGGAGTACTTTGCCTGCGTGTCTAGCGGCAACTCTAATGCCAACAGCAGAG GCCCTTCCTTTTCCTCTGGCCTGGTGCCACTGTCTCTGAGGGGCAGCCACATG GGCCCCATGGAGATCTACAAGACCGTGTCCGCCTGGAAGAGGCAGCCTGTGC GCGTGCTGTCTCTGTTCCGCAACATCGACAAGGTGCTGAAGAGCCTGGGCTTT CTGGAGAGCGGATCCGGATCTGGAGGAGGCACCCTGAAGTATGTGGAGGAT GTGACAAATGTGGTGCGGAGAGATGTGGAGAAGTGGGGCCCCTTCGATCTGG TGTACGGATCCACCCAGCCACTGGGAAGCTCCTGCGATAGGTGTCCAGGATG GTATATGTTCCAGTTTCACAGAATCCTGCAGTACGCACTGCCAAGGCAGGAG AGCCAGCGCCCTTTCTTTTGGATCTTTATGGACAACCTGCTGCTGACAGAGGA TGACCAGGAGACAACAACCCGCTTCCTGCAGACAGAGGCAGTGACCCTGCAG GATGTGAGGGGACGCGACTATCAGAATGCCATGCGGGTGTGGTCTAACATCC CTGGCCTGAAGAGCAAGCACGCCCCCCTGACCCCTAAGGAGGAGGAGTACCT GCAGGCCCAGGTGCGGAGCAGATCCAAGCTGGATGCCCCTAAGGTGGACCTG CTGGTGAAGAATTGTCTGCTGCCACTGCGGGAGTACTTCAAGTACTTTAGTCA GAATAGCCTGCCACTGGGTGGACCTAGCTCTGGAGCCCCGCCTCCGTCAGGT GGATCTCCTGCAGGTAGCCCAACTTCCACGGAAGAGGGCACGTCTGAGAGTG CGACTCCTGAGAGCGGGCCTGGCACAAGTACTGAGCCCAGCGAAGGTTCTGC ACCCGGGTCTCCAGCTGGGAGCCCTACCTCAACAGAGGAAGGTACCAGCACA GAGCCTTCTGAGGGTAGCGCTCCTGGCACGTCCACCGAACCGTCCGAGGGGA TGGCACCGAAGAAAAAACGCAAAGTTGGTCGAGCTGCTATGGCAGAGCGGC CATTTCAGTGTCGAATTTGCATGCGAAATTTCTCTAGCGAAGCGGATCGAAGT AGACATATCCGGACCCATACGGGTGAGAAACCGTTCGCGTGTGATATATGCG GTCGCAAATTCGCAGACCGCTCCAACCTGACAAGGCACACAAAAATTCATAC TGGAAGCCAGAAACCGTTTCAGTGCCGGATTTGTATGCGCAATTTCTCTCAAT CCTCCGATCTCTCCCGCCACATCAGGACTCATACCGGGGAGAAACCCTTTGCT TGTGACATTTGTGGCAGAAAGTTTGCCTATCACTGGTACCTTAAGAAGCACA CCAAGATCCATACTGGCTCTCAAAAGCCTTTCCAATGCCGAATATGTATGCG AAACTTTTCTAGGTCTGACTCCCTCTCTGTACATATCCGCACTCACACGGGTG AAAAACCATTTGCCTGTGACATATGTGGCAGAAAATTTGCTCAAAATGCGAA CCGAAAAACGCATACGAAAATCCATTTGCGACAAAAAGATGCAGCTCGGGCT CCGGACCGAAAAAAAAGAGAAAGGTAAACGGCGGGGGAGGATCACGGACC CTGGTTACATTTAAGGACGTTTTTGTAGACTTCACGAGGGAGGAATGGAAGC TGTTGGACACCGCCCAGCAAATCGTGTATCGAAATGTAATGTTGGAAAATTA TAAAAATCTGGTTTCCCTCGGCTACCAATTGACAAAACCAGACGTCATCCTCA GACTGGAAAAAGGGGAAGAACCTTGGCTTGTCTGA	3xFLAG-DNMT3AL-XTEN80-SV40 NLS-ZFP-SV40 NLS-KRAB (DNA 서열)

[1272]

973	ATGCACGATCAGGAGTTTGACCCCCCTAAGGTGTACCCACCCGTGCCAGCCG AGAAGAGGAAGCCCATCCGCGTGCTGTCCCTGTTCGACGGCATCGCCACAGG CCTGCTGGTGCTGAAGGATCTGGGCATCCAGGTGGACAGATATATCGCCTCC GAGGTGTGCGAGGATTCTATCACCGTGGGCATGGTGAGGCACCAGGGCAAG ATCATGTACGTGGGCGACGTGCGCAGCGTGACACAGAAGCACATCCAGGAGT GGGGACCCTTCGACCTGGTCATCGGAGGCAGCCCCTGTAATGACCTGTCCAT CGTGAACCCTGCAAGGAAGGGCCTGTATGAGGGAACCGGCAGACTGTTCTTT GAGTTCTACAGGCTGCTGCACGACGCCCGCCCTAAGGAGGGCGATGACAGGC CATTCTTTTGGCTGTTTGAGAACGTGGTGGCCATGGGCGTGAGCGACAAGCG GGATATCTCCAGATTCCTGGAGTCTAATCCCGTGATGATCGATGCAAAGGAG GTGTCTGCCGCACACAGGGCAAGGTACTTTTGGGGAAATCTGCCTGGCATGA ACCGCCCACTGGCCAGCACCGTGAACGACAAGCTGGAGCTGCAGGAGTGCCT GGAGCACGGAAGGATCGCCAAGTTCTCCAAGGTGCGGACAATCACCACAAG ATCTAACAGCATCAAGCAGGGCAAGGATCAGCACTTCCCCGTGTTCATGAAT GAGAAGGAGGACATCCTGTGGTGTACCGAGATGGAGCGCGTGTTCGGCTTTC CAGTGCACTATACAGACGTGAGCAATATGAGCCGGCTGGCAAGGCAGAGAC TGCTGGGCCGGTCCTGGTCTGTGCCAGTGATCAGACACCTGTTCGCCCCCCTG AAGGAGTACTTTGCCTGCGTGTCTAGCGGCAACTCTAATGCCAACAGCAGAG GCCCTTCCTTTTCCTCTGGCCTGGTGCCACTGTCTCTGAGGGGCAGCCACATG GGCCCCATGGAGATCTACAAGACCGTGTCCGCCTGGAAGAGGCAGCCTGTGC GCGTGCTGTCTCTGTTCCGCAACATCGACAAGGTGCTGAAGAGCCTGGGCTTT CTGGAGAGCGGATCCGGATCTGGAGGAGGCACCCTGAAGTATGTGGAGGAT GTGACAAATGTGGTGCGGAGAGATGTGGAGAAGTGGGGCCCCTTCGATCTGG TGTACGGATCCACCCAGCCACTGGGAAGCTCCTGCGATAGGTGTCCAGGATG GTATATGTTCCAGTTTCACAGAATCCTGCAGTACGCACTGCCAAGGCAGGAG AGCCAGCGCCCTTTCTTTTGGATCTTTATGGACAACCTGCTGCTGACAGAGGA TGACCAGGAGACAACAACCCGCTTCCTGCAGACAGAGGCAGTGACCCTGCAG GATGTGAGGGGACGCGACTATCAGAATGCCATGCGGGTGTGGTCTAACATCC CTGGCCTGAAGAGCAAGCACGCCCCCCTGACCCCTAAGGAGGAGGAGTACCT GCAGGCCCAGGTGCGGAGCAGATCCAAGCTGGATGCCCCTAAGGTGGACCTG CTGGTGAAGAATTGTCTGCTGCCACTGCGGGAGTACTTCAAGTACTTTAGTCA GAATAGCCTGCCACTGGGTGGACCTAGCTCTGGAGCCCCGCCTCCGTCAGGT GGATCTCCTGCAGGTAGCCCAACTTCCACGGAAGAGGGCACGTCTGAGAGTG CGACTCCTGAGAGCGGGCCTGGCACAAGTACTGAGCCCAGCGAAGGTTCTGC ACCCGGGTCTCCAGCTGGGAGCCCTACCTCAACAGAGGAAGGTACCAGCACA GAGCCTTCTGAGGGTAGCGCTCCTGGCACGTCCACCGAACCGTCCGAGGGGA TGGCACCGAAGAAAAAACGCAAAGTTGGTCGAGCGGCAATGGCAGAACGCC CGTTTCAGTGTAGAATCTGCATGCGCAATTTCTCCAGAAGTGACGTACTCAGC ACGCATATCAGAACACATACTGGTGAAAAACCGTTTGCTTGTGATATCTGTG GTAAGAAATTCGCGGATAACTCCTCAAGGACGCGGCATACGAAGATCCACAC TGGTTCTCAAAAGCCGTTCCAGTGCCGGATATGCATGCGGAACTTTAGTCGG CCATATACGCTTCGACTTCATATTAGGACTCACACCGGCGAAAAGCCGTTCG CCTGTGACATTTGCGGGAGAAAATTCGCTGACTCTAGTCACCGCACGAGGCA CACTAAAATACATACTGGTTCACAAAAACCATTCCAGTGCCGAATTTGCATG CGAAATTTTTCCCGAAGTGACCATCTCTCACAGCACATCCGCACCCATACAG GGGAGAAGCCCTTCGCTTGTGATATATGCGGACGCAAGTTCGCGGACAGCTC ACACCGGACCCGCCATACAAAGATCCACTTGAGACAGAAAGATGCAGCGCG GGCTCCGGACCGAAAAAAAAGAGAAAGGTAAACGGCGGGGGAGGATCACG GACCCTGGTTACATTTAAGGACGTTTTTGTAGACTTCACGAGGGAGGAATGG AAGCTGTTGGACACCGCCCAGCAAATCGTGTATCGAAATGTAATGTTGGAAA ATTATAAAAATCTGGTTTCCCTCGGCTACCAATTGACAAAACCAGACGTCATC CTCAGACTGGAAAAAGGGGAAGAACCTTGGCTTGTCTGA	3xFLAG- DNMT3AL- XTEN80-SV40 NLS-ZFP-SV40 NLS-KRAB (DNA 서열)
974	ATGCACGATCAGGAGTTTGACCCCCCTAAGGTGTACCCACCCGTGCCAGCCG AGAAGAGGAAGCCCATCCGCGTGCTGTCCCTGTTCGACGGCATCGCCACAGG CCTGCTGGTGCTGAAGGATCTGGGCATCCAGGTGGACAGATATATCGCCTCC GAGGTGTGCGAGGATTCTATCACCGTGGGCATGGTGAGGCACCAGGGCAAG ATCATGTACGTGGGCGACGTGCGCAGCGTGACACAGAAGCACATCCAGGAGT GGGGACCCTTCGACCTGGTCATCGGAGGCAGCCCCTGTAATGACCTGTCCAT CGTGAACCCTGCAAGGAAGGGCCTGTATGAGGGAACCGGCAGACTGTTCTTT	3xFLAG- DNMT3AL- XTEN80-SV40 NLS-ZFP-SV40 NLS-KRAB (DNA 서열)

[1273]

	GAGTTCTACAGGCTGCTGCACGACGCCCGCCCTAAGGAGGGCGATGACAGGC CATTCTTTTGGCTGTTTGAGAACGTGGTGGCCATGGGCGTGAGCGACAAGCG GGATATCTCCAGATTCCTGGAGTCTAATCCCGTGATGATCGATGCAAAGGAG GTGTCTGCCGCACACAGGGCAAGGTACTTTTGGGGAAATCTGCCTGGCATGA ACCGCCCACTGGCCAGCACCGTGAACGACAAGCTGGAGCTGCAGGAGTGCCT GGAGCACGGAAGGATCGCCAAGTTCTCCAAGGTGCGGACAATCACCACAAG ATCTAACAGCATCAAGCAGGGCAAGGATCAGCACTTCCCCGTGTTCATGAAT GAGAAGGAGGACATCCTGTGGTGTACCGAGATGGAGCGCGTGTTCGGCTTTC CAGTGCACTATACAGACGTGAGCAATATGAGCCGGCTGGCAAGGCAGAGAC TGCTGGGCCGGTCCTGGTCTGTGCCAGTGATCAGACACCTGTTCGCCCCCCTG AAGGAGTACTTTGCCTGCGTGTCTAGCGGCAACTCTAATGCCAACAGCAGAG GCCCTTCCTTTTCCTCTGGCCTGGTGCCACTGTCTCTGAGGGGCAGCCACATG GGCCCCATGGAGATCTACAAGACCGTGTCCGCCTGGAAGAGGCAGCCTGTGC GCGTGCTGTCTCTGTTCCGCAACATCGACAAGGTGCTGAAGAGCCTGGGCTTT CTGGAGAGCGGATCCGGATCTGGAGGAGGCACCCTGAAGTATGTGGAGGAT GTGACAAATGTGGTGCGGAGAGATGTGGAGAAGTGGGGCCCCTTCGATCTGG TGTACGGATCCACCCAGCCACTGGGAAGCTCCTGCGATAGGTGTCCAGGATG GTATATGTTCCAGTTTCACAGAATCCTGCAGTACGCACTGCCAAGGCAGGAG AGCCAGCGCCCTTTCTTTTGGATCTTTATGGACAACCTGCTGCTGACAGAGGA TGACCAGGAGACAACAACCCGCTTCCTGCAGACAGAGGCAGTGACCCTGCAG GATGTGAGGGGACGCGACTATCAGAATGCCATGCGGGTGTGGTCTAACATCC CTGGCCTGAAGAGCAAGCACGCCCCCCTGACCCCTAAGGAGGAGGAGTACCT GCAGGCCCAGGTGCGGAGCAGATCCAAGCTGGATGCCCCTAAGGTGGACCTG CTGGTGAAGAATTGTCTGCTGCCACTGCGGGAGTACTTCAAGTACTTTAGTCA GAATAGCCTGCCACTGGGTGGACCTAGCTCTGGAGCCCCGCCTCCGTCAGGT GGATCTCCTGCAGGTAGCCCAACTTCCACGGAAGAGGGCACGTCTGAGAGTG CGACTCCTGAGAGCGGGCCTGGCACAAGTACTGAGCCCAGCGAAGGTTCTGC ACCCGGGTCTCCAGCTGGGAGCCCTACCTCAACAGAGGAAGGTACCAGCACA GAGCCTTCTGAGGGTAGCGCTCCTGGCACGTCCACCGAACCGTCCGAGGGGA TGGCACCGAAGAAAAAACGCAAAGTTGGTCGAGCCGCAATGGCCGAAAGAC CATTTCAGTGCAGGATATGTATGCGCAACTTCTCTCGCAGTGACCACCTGAGT CAACATATCAGGACACACACGGGTGAAAAGCCTTTTGCATGCGATATTTGTG GTCGAAAATTTGCTCAGTCTGCGGACCGAACCAAGCACACTAAAATTCATAC CGGCTCACAGAAACCGTTTCAATGCCGCATCTGTATGAGGAATTTCTCTAGAT CAGACCACTTGTCCCAACACATCCGGACTCATACTGGAGAAAAGCCGTTTGC ATGTGACATTTGTGGCAGGAAGTTTGCTAGAAGGTCTGACCTTAAAAGGCAC ACAAAAATTCATACGGGTTCCCAGAAACCATTTCAGTGCCGGATATGCATGC GGAACTTTTCACGAAGCGACCACCTCTCCCGACATATTCGAACGCACACTGG TGAGAAGCCGTTTGCTTGCGACATTTGCGGACGCAAGTTCGCTCAGAGCTCC GACTTGAGGAGGCATACCAAGATTCATCTCCGGCAGAAAGATGCCGCGCGGG CTCCGGACCGAAAAAAAAAGAGAAAGGTAAACGGCGGGGGAGGATCACGGAC CCTGGTTACATTTAAGGACGTTTTTGTAGACTTCACGAGGGAGGAATGGAAG CTGTTGGACACCGCCCAGCAAATCGTGTATCGAAATGTAATGTTGGAAAATT ATAAAAATCTGGTTTCCCTCGGCTACCAATTGACAAAACCAGACGTCATCCTC AGACTGGAAAAAGGGGAAGAACCTTGGCTTGTCTGA	
975	ATGCACGATCAGGAGTTTGACCCCCCTAAGGTGTACCCACCCGTGCCAGCCG AGAAGAGGAAGCCCATCCGCGTGCTGTCCCTGTTCGACGGCATCGCCACAGG CCTGCTGGTGCTGAAGGATCTGGGCATCCAGGTGGACAGATATATCGCCTCC GAGGTGTGCGAGGATTCTATCACCGTGGGCATGGTGAGGCACCAGGGCAAG ATCATGTACGTGGGCGACGTGCGCAGCGTGACACAGAAGCACATCCAGGAGT GGGGACCCTTCGACCTGGTCATCGGAGGCAGCCCCTGTAATGACCTGTCCAT CGTGAACCCTGCAAGGAAGGGCCTGTATGAGGGAACCGGCAGACTGTTCTTT GAGTTCTACAGGCTGCTGCACGACGCCCGCCCTAAGGAGGGCGATGACAGGC CATTCTTTTGGCTGTTTGAGAACGTGGTGGCCATGGGCGTGAGCGACAAGCG GGATATCTCCAGATTCCTGGAGTCTAATCCCGTGATGATCGATGCAAAGGAG GTGTCTGCCGCACACAGGGCAAGGTACTTTTGGGGAAATCTGCCTGGCATGA ACCGCCCACTGGCCAGCACCGTGAACGACAAGCTGGAGCTGCAGGAGTGCCT GGAGCACGGAAGGATCGCCAAGTTCTCCAAGGTGCGGACAATCACCACAAG ATCTAACAGCATCAAGCAGGGCAAGGATCAGCACTTCCCCGTGTTCATGAAT	3xFLAG- DNMT3AL- XTEN80-SV40 NLS-ZFP-SV40 NLS-KRAB (DNA 서열)

[1274]

	GAGAAGGAGGACATCCTGTGGTGTACCGAGATGGAGCGCGTGTTCGGCTTTC CAGTGCACTATACAGACGTGAGCAATATGAGCCGGCTGGCAAGGCAGAGAC TGCTGGGCCGGTCCTGGTCTGTGCCAGTGATCAGACACCTGTTCGCCCCCCTG AAGGAGTACTTTGCCTGCGTGTCTAGCGGCAACTCTAATGCCAACAGCAGAG GCCCTTCCTTTTCCTCTGGCCTGGTGCCACTGTCTCTGAGGGGCAGCCACATG GGCCCCATGGAGATCTACAAGACCGTGTCCGCCTGGAAGAGGCAGCCTGTGC GCGTGCTGTCTCTGTTCCGCAACATCGACAAGGTGCTGAAGAGCCTGGGCTTT CTGGAGAGCGGATCCGGATCTGGAGGAGGCACCCTGAAGTATGTGGAGGAT GTGACAAATGTGGTGCGGAGAGATGTGGAGAAGTGGGGCCCCTTCGATCTGG TGTACGGATCCACCCAGCCACTGGGAAGCTCCTGCGATAGGTGTCCAGGATG GTATATGTTCCAGTTTCACAGAATCCTGCAGTACGCACTGCCAAGGCAGGAG AGCCAGCGCCCTTTCTTTTGGATCTTTATGGACAACCTGCTGCTGACAGAGGA TGACCAGGAGACAACAACCCGCTTCCTGCAGACAGAGGCAGTGACCCTGCAG GATGTGAGGGGACGCGACTATCAGAATGCCATGCGGGTGTGGTCTAACATCC CTGGCCTGAAGAGCAAGCACGCCCCCCTGACCCCTAAGGAGGAGGAGTACCT GCAGGCCCAGGTGCGGAGCAGATCCAAGCTGGATGCCCCTAAGGTGGACCTG CTGGTGAAGAATTGTCTGCTGCCACTGCGGGAGTACTTCAAGTACTTTAGTCA GAATAGCCTGCCACTGGGTGGACCTAGCTCTGGAGCCCCGCCTCCGTCAGGT GGATCTCCTGCAGGTAGCCCAACTTCCACGGAAGAGGGCACGTCTGAGAGTG CGACTCCTGAGAGCGGGCCTGGCACAAGTACTGAGCCCAGCGAAGGTTCTGC ACCCGGGTCTCCAGCTGGGAGCCCTACCTCAACAGAGGAAGGTACCAGCACA GAGCCTTCTGAGGGTAGCGCTCCTGGCACGTCCACCGAACCGTCCGAGGGGA TGGCACCGAAGAAAAAACGCAAAGTTGGTCGAGCCGCGATGGCTGAGAGAC CATTTCAGTGTCGAATCTGCATGAGAAATTTTTCAAGGAGTGACAATCTGTCT GAGCACATACGAACACATACTGGGGAAAAACCCTTTGCATGTGACATTTGTG GAAGAAAGTTTGCTACCAGCTCAAATCGCAAAACACATACAAAGATACATAC CGGCTCCCAAAAGCCATTCCAGTGCCGCATCTGCATGAGGAACTTTTCCGATC GCTCACATCTTACCCGCCACATAAGAACTCACACAGGCGAAAAGCCCTTTGC CTGCGATATATGCGGACGGAAGTTCGCCCGCTCCGACGCTTTGACCCAGCAT ACCAAAATCCATACTGGGTCTCAAAAGCCATTTCAGTGCCGAATCTGTATGA GGAATTTCTCCGACAGGTCAGCATTGGCACGGCATATCCGCACCCATACCGG TGAGAAGCCTTTTGCTTGCGATATCTGTGGACGAAAATTTGCCCGGAGGTTCA CTCTCTCCAAACACACAAAGATACATCTGCGCCAAAAGGATGCAGCCCGGGC TCCGGACCGAAAAAAAAAGAGAAAGGTAAACGGCGGGGGAGGATCACGGACC CTGGTTACATTTAAGGACGTTTTTGTAGACTTCACGAGGGAGGAATGGAAGC TGTTGGACACCGCCCAGCAAATCGTGTATCGAAATGTAATGTTGGAAAATTA TAAAAATCTGGTTTCCCTCGGCTACCAATTGACAAAACCAGACGTCATCCTCA GACTGGAAAAAGGGGAAGAACCTTGGCTTGTCTGA	
976	ATGCACGATCAGGAGTTTGACCCCCCTAAGGTGTACCCACCCGTGCCAGCCG AGAAGAGGAAGCCCATCCGCGTGCTGTCCCTGTTCGACGGCATCGCCACAGG CCTGCTGGTGCTGAAGGATCTGGGCATCCAGGTGGACAGATATATCGCCTCC GAGGTGTGCGAGGATTCTATCACCGTGGGCATGGTGAGGCACCAGGGCAAG ATCATGTACGTGGGCGACGTGCGCAGCGTGACACAGAAGCACATCCAGGAGT GGGGACCCTTCGACCTGGTCATCGGAGGCAGCCCCTGTAATGACCTGTCCAT CGTGAACCCTGCAAGGAAGGGCCTGTATGAGGGAACCGGCAGACTGTTCTTT GAGTTCTACAGGCTGCTGCACGACGCCCGCCCTAAGGAGGGCGATGACAGGC CATTCTTTTGGCTGTTTGAGAACGTGGTGGCCATGGGCGTGAGCGACAAGCG GGATATCTCCAGATTCCTGGAGTCTAATCCCGTGATGATCGATGCAAAGGAG GTGTCTGCCGCACACAGGGCAAGGTACTTTTGGGGAAATCTGCCTGGCATGA ACCGCCCACTGGCCAGCACCGTGAACGACAAGCTGGAGCTGCAGGAGTGCCT GGAGCACGGAAGGATCGCCAAGTTCTCCAAGGTGCGGACAATCACCACAAG ATCTAACAGCATCAAGCAGGGCAAGGATCAGCACTTCCCCGTGTTCATGAAT GAGAAGGAGGACATCCTGTGGTGTACCGAGATGGAGCGCGTGTTCGGCTTTC CAGTGCACTATACAGACGTGAGCAATATGAGCCGGCTGGCAAGGCAGAGAC TGCTGGGCCGGTCCTGGTCTGTGCCAGTGATCAGACACCTGTTCGCCCCCCTG AAGGAGTACTTTGCCTGCGTGTCTAGCGGCAACTCTAATGCCAACAGCAGAG GCCCTTCCTTTTCCTCTGGCCTGGTGCCACTGTCTCTGAGGGGCAGCCACATG GGCCCCATGGAGATCTACAAGACCGTGTCCGCCTGGAAGAGGCAGCCTGTGC GCGTGCTGTCTCTGTTCCGCAACATCGACAAGGTGCTGAAGAGCCTGGGCTTT	3xFLAG- DNMT3AL- XTEN80-SV40 NLS-ZFP-SV40 NLS-KRAB (DNA 서열)

[1275]

	CTGGAGAGCGGATCCGGATCTGGAGGAGGCACCCTGAAGTATGTGGAGGAT GTGACAAATGTGGTGCGGAGAGATGTGGAGAAGTGGGGCCCCTTCGATCTGG TGTACGGATCCACCCAGCCACTGGGAAGCTCCTGCGATAGGTGTCCAGGATG GTATATGTTCCAGTTTCACAGAATCCTGCAGTACGCACTGCCAAGGCAGGAG AGCCAGCGCCCTTTCTTTTGGATCTTTATGGACAACCTGCTGCTGACAGAGGA TGACCAGGAGACAACAACCCGCTTCCTGCAGACAGAGGCAGTGACCCTGCAG GATGTGAGGGGACGCGACTATCAGAATGCCATGCGGGTGTGGTCTAACATCC CTGGCCTGAAGAGCAAGCACGCCCCCCTGACCCCTAAGGAGGAGGAGTACCT GCAGGCCCAGGTGCGGAGCAGATCCAAGCTGGATGCCCCTAAGGTGGACCTG CTGGTGAAGAATTGTCTGCTGCCACTGCGGGAGTACTTCAAGTACTTTAGTCA GAATAGCCTGCCACTGGGTGGACCTAGCTCTGGAGCCCCGCCTCCGTCAGGT GGATCTCCTGCAGGTAGCCCAACTTCCACGGAAGAGGGCACGTCTGAGAGTG CGACTCCTGAGAGCGGGCCTGGCACAAGTACTGAGCCCAGCGAAGGTTCTGC ACCCGGGTCTCCAGCTGGGAGCCCTACCTCAACAGAGGAAGGTACCAGCACA GAGCCTTCTGAGGGTAGCGCTCCTGGCACGTCCACCGAACCGTCCGAGGGGA TGGCACCGAAGAAAAAACGCAAAGTTGGTCGAGCCGCAATGGCTGAGCGCC CGTTCCAGTGCAGAATATGCATGCGGAATTTTTCTAGGTCAGATCATTTGTCT GAGCATATTCGCACACACACGGGAGAGAAGCCCTTTGCTTGCGATATATGTG GAAGGAAATTCGCGCAATACAGTGGGCGCTACTACCATACAAAGATCCATAC GGGCTCCCAGAAGCCCTTCCAATGTCGAATATGTATGAGGAATTTTAGTCAC GGACAAACATTGAATGAACATATACGCACTCACACTGGTGAAAAACCATTTG CGTGCGATATTTGCGGAAGGAAGTTTGCTCAGTCTGGGAATTTGGCGCGACA CACCAAGATCCACACAGGATCCCAGAAACCATTTCAGTGCAGAATTTGTATG AGAAACTTTAGCCGCAGTGACAGTCTCTTGAGGCACATACGGACTCATACTG GGGAGAAACCATTCGCCTGCGATATTTGTGGACGAAAGTTCGCCTGTCGCGA GTACAGAGGCAAGCACACTAAGATACATCTTAGGCAAAAGGACGCTGCACG GGCTCCGGACCGAAAAAAAAGAGAAAGGTAAACGGCGGGGGAGGATCACG GACCCTGGTTACATTTAAGGACGTTTTTGTAGACTTCACGAGGGAGGAATGG AAGCTGTTGGACACCGCCCAGCAAATCGTGTATCGAAATGTAATGTTGGAAA ATTATAAAAATCTGGTTTCCCTCGGCTACCAATTGACAAAACCAGACGTCATC CTCAGACTGGAAAAAGGGGAAGAACCTTGGCTTGTCTGA	
977	ATGCACGATCAGGAGTTTGACCCCCCTAAGGTGTACCCACCCGTGCCAGCCG AGAAGAGGAAGCCCATCCGCGTGCTGTCCCTGTTCGACGGCATCGCCACAGG CCTGCTGGTGCTGAAGGATCTGGGCATCCAGGTGGACAGATATATCGCCTCC GAGGTGTGCGAGGATTCTATCACCGTGGGCATGGTGAGGCACCAGGGCAAG ATCATGTACGTGGGCGACGTGCGCAGCGTGACACAGAAGCACATCCAGGAGT GGGGACCCTTCGACCTGGTCATCGGAGGCAGCCCCTGTAATGACCTGTCCAT CGTGAACCCTGCAAGGAAGGGCCTGTATGAGGGAACCGGCAGACTGTTCTTT GAGTTCTACAGGCTGCTGCACGACGCCCGCCCTAAGGAGGGCGATGACAGGC CATTCTTTTGGCTGTTTGAGAACGTGGTGGCCATGGGCGTGAGCGACAAGCG GGATATCTCCAGATTCCTGGAGTCTAATCCCGTGATGATCGATGCAAAGGAG GTGTCTGCCGCACACAGGGCAAGGTACTTTTGGGGAAATCTGCCTGGCATGA ACCGCCCACTGGCCAGCACCGTGAACGACAAGCTGGAGCTGCAGGAGTGCCT GGAGCACGGAAGGATCGCCAAGTTCTCCAAGGTGCGGACAATCACCACAAG ATCTAACAGCATCAAGCAGGGCAAGGATCAGCACTTCCCCGTGTTCATGAAT GAGAAGGAGGACATCCTGTGGTGTACCGAGATGGAGCGCGTGTTCGGCTTTC CAGTGCACTATACAGACGTGAGCAATATGAGCCGGCTGGCAAGGCAGAGAC TGCTGGGCCGGTCCTGGTCTGTGCCAGTGATCAGACACCTGTTCGCCCCCCTG AAGGAGTACTTTGCCTGCGTGTCTAGCGGCAACTCTAATGCCAACAGCAGAG GCCCTTCCTTTCCTCTGGCCTGGTGCCACTGTCTCTGAGGGGCAGCCACATG GGCCCCATGGAGATCTACAAGACCGTGTCCGCCTGGAAGAGGCAGCCTGTGC GCGTGCTGTCTCTGTTCCGCAACATCGACAAGGTGCTGAAGAGCCTGGGCTTT CTGGAGAGCGGATCCGGATCTGGAGGAGGCACCCTGAAGTATGTGGAGGAT GTGACAAATGTGGTGCGGAGAGATGTGGAGAAGTGGGGCCCCTTCGATCTGG TGTACGGATCCACCCAGCCACTGGGAAGCTCCTGCGATAGGTGTCCAGGATG GTATATGTTCCAGTTTCACAGAATCCTGCAGTACGCACTGCCAAGGCAGGAG AGCCAGCGCCCTTTCTTTTGGATCTTTATGGACAACCTGCTGCTGACAGAGGA TGACCAGGAGACAACAACCCGCTTCCTGCAGACAGAGGCAGTGACCCTGCAG GATGTGAGGGGACGCGACTATCAGAATGCCATGCGGGTGTGGTCTAACATCC	3xFLAG- DNMT3AL- XTEN80-SV40 NLS-ZFP-SV40 NLS-KRAB (DNA 서열)

[1276]

	CTGGCCTGAAGAGCAAGCACGCCCCCCTGACCCCTAAGGAGGAGGAGTACCT GCAGGCCCAGGTGCGGAGCAGATCCAAGCTGGATGCCCCTAAGGTGGACCTG CTGGTGAAGAATTGTCTGCTGCCACTGCGGGAGTACTTCAAGTACTTTAGTCA GAATAGCCTGCCACTGGGTGGACCTAGCTCTGGAGCCCCGCCTCCGTCAGGT GGATCTCCTGCAGGTAGCCCAACTTCCACGGAAGAGGGCACGTCTGAGAGTG CGACTCCTGAGAGCGGGCCTGGCACAAGTACTGAGCCCAGCGAAGGTTCTGC ACCCGGGTCTCCAGCTGGGAGCCCTACCTCAACAGAGGAAGGTACCAGCACA GAGCCTTCTGAGGGTAGCGCTCCTGGCACGTCCACCGAACCGTCCGAGGGGA TGGCACCGAAGAAAAAACGCAAAGTTGGTCGAGCCGCGATGGCTGAGAGGC CTTTTCAATGTCGAATCTGTATGAGGAACTTCTCTCAATCTGCTAATCGCACG ACGCACATTCGAACGCATACCGGTGAGAAGCCATTCGCGTGCGATATCTGCG GACGGAAATTCGCGAGGTCAGCTAATCTTACACGGCACACGAAGATCCACAC AGGGTCACAGAAACCTTTTCAGTGTCGCATTTGCATGAGGAATTTCTCCCGAT CTGACGTCCTTAGCGAACATATACGAACTCACACGGGCGAGAAGCCATTTGC GTGCGATATATGCGGGAGGAAGTTTGCCACCTCTGGACATCTGAGTCGACAT ACCAAAATTCATACCGGTAGTCAGAAGCCGTTCCAATGCAGAATATGTATGC GAAATTTCTCTCAAAGCTCAGACTTGTCTAGGCACATAAGAACGCACACGGG TGAAAAACCTTTCGCGTGTGATATCTGCGGCAGAAAGTTCGCACAATGGTCC ACCCGAAAGCGGCATACGAAGATTCACCTCAGACAGAAAGACGCTGCCCGG GCTCCGGACCGAAAAAAAAAGAGAAAGGTAAACGGCGGGGGAGGATCACGG ACCCTGGTTACATTTAAGGACGTTTTTGTAGACTTCACGAGGGAGGAATGGA AGCTGTTGGACACCGCCCAGCAAATCGTGTATCGAAATGTAATGTTGGAAAA TTATAAAAATCTGGTTTCCCTCGGCTACCAATTGACAAAACCAGACGTCATCC TCAGACTGGAAAAAGGGGAAGAACCTTGGCTTGTCTGA	
978	ATGCACGATCAGGAGTTTGACCCCCCTAAGGTGTACCCACCCGTGCCAGCCG AGAAGAGGAAGCCCATCCGCGTGCTGTCCCTGTTCGACGGCATCGCCACAGG CCTGCTGGTGCTGAAGGATCTGGGCATCCAGGTGGACAGATATATCGCCTCC GAGGTGTGCGAGGATTCTATCACCGTGGGCATGGTGAGGCACCAGGGCAAG ATCATGTACGTGGGCGACGTGCGCAGCGTGACACAGAAGCACATCCAGGAGT GGGGACCCTTCGACCTGGTCATCGGAGGCAGCCCCTGTAATGACCTGTCCAT CGTGAACCCTGCAAGGAAGGGCCTGTATGAGGGAACCGGCAGACTGTTCTTT GAGTTCTACAGGCTGCTGCACGACGCCCGCCCTAAGGAGGGCGATGACAGGC CATTCTTTTGGCTGTTTGAGAACGTGGTGGCCATGGGCGTGAGCGACAAGCG GGATATCTCCAGATTCCTGGAGTCTAATCCCGTGATGATCGATGCAAAGGAG GTGTCTGCCGCACACAGGGCAAGGTACTTTTGGGGAAATCTGCCTGGCATGA ACCGCCCACTGGCCAGCACCGTGAACGACAAGCTGGAGCTGCAGGAGTGCCT GGAGCACGGAAGGATCGCCAAGTTCTCCAAGGTGCGGACAATCACCACAAG ATCTAACAGCATCAAGCAGGGCAAGGATCAGCACTTCCCCGTGTTCATGAAT GAGAAGGAGGACATCCTGTGGTGTACCGAGATGGAGCGCGTGTTCGGCTTTC CAGTGCACTATACAGACGTGAGCAATATGAGCCGGCTGGCAAGGCAGAGAC TGCTGGGCCGGTCCTGGTCTGTGCCAGTGATCAGACACCTGTTCGCCCCCCTG AAGGAGTACTTTGCCTGCGTGTCTAGCGGCAACTCTAATGCCAACAGCAGAG GCCCTTCCTTTTCCTCTGGCCTGGTGCCACTGTCTCTGAGGGGCAGCCACATG GGCCCCATGGAGATCTACAAGACCGTGTCCGCCTGGAAGAGGCAGCCTGTGC GCGTGCTGTCTCTGTTCCGCAACATCGACAAGGTGCTGAAGAGCCTGGGCTTT CTGGAGAGCGGATCCGGATCTGGAGGAGGCACCCTGAAGTATGTGGAGGAT GTGACAAATGTGGTGCGGAGAGATGTGGAGAAGTGGGGCCCCTTCGATCTGG TGTACGGATCCACCCAGCCACTGGGAAGCTCCTGCGATAGGTGTCCAGGATG GTATATGTTCCAGTTTCACAGAATCCTGCAGTACGCACTGCCAAGGCAGGAG AGCCAGCGCCCTTTCTTTTGGATCTTTATGGACAACCTGCTGCTGACAGAGGA TGACCAGGAGACAACAACCCGCTTCCTGCAGACAGAGGCAGTGACCCTGCAG GATGTGAGGGGACGCGACTATCAGAATGCCATGCGGGTGTGGTCTAACATCC CTGGCCTGAAGAGCAAGCACGCCCCCCTGACCCCTAAGGAGGAGGAGTACCT GCAGGCCCAGGTGCGGAGCAGATCCAAGCTGGATGCCCCTAAGGTGGACCTG CTGGTGAAGAATTGTCTGCTGCCACTGCGGGAGTACTTCAAGTACTTTAGTCA GAATAGCCTGCCACTGGGTGGACCTAGCTCTGGAGCCCCGCCTCCGTCAGGT GGATCTCCTGCAGGTAGCCCAACTTCCACGGAAGAGGGCACGTCTGAGAGTG CGACTCCTGAGAGCGGGCCTGGCACAAGTACTGAGCCCAGCGAAGGTTCTGC ACCCGGGTCTCCAGCTGGGAGCCCTACCTCAACAGAGGAAGGTACCAGCACA	3xFLAG- DNMT3AL- XTEN80-SV40 NLS-ZFP-SV40 NLS-KRAB (DNA 서열)

[1277]

	GAGCCTTCTGAGGGTAGCGCTCCTGGCACGTCCACCGAACCGTCCGAGGGGA TGGCACCGAAGAAAAAACGCAAAGTTGGTCGAGCGGCGATGGCAGAACGCC CGTTCCAATGCAGAATATGTATGAGAAACTTCTCCCAGAGCGGAAATCTGGC ACGCCACATCCGGACACACACGGGAGAGAAGCCATTTGCTTGTGACATTTGT GGTCGCAAATTTGCCGCCACCTGTTGTCTGGCACATCATACTAAGATACATAC GGGGTCACAGAAACCATTCCAATGTAGGATCTGCATGCGGAATTTTTCTCGG TGGCAGTATTTGCCTACGCATATTAGAACCCACGCCGGTGAGAAACCGTTTG CATGTGACATCTGCGGACGAAAGTTTGCCGATAGATCTGCGCTTGCTAGGCA TACTAAAATCCACACGGGGTCCCAGAAGCCTTTTCAGTGTCGGATATGTATG AGGAACTTCAGTCGATCAGACAACCTTAGCGAGCATATTCGGACGCATACTG GAGAAAAACCTTTTGCTTGTGATATATGCGGTAGGAAGTTCGCCAAACGGTG TAACCTTCGCTGTCACACCAAAATACATCTTCGCCAGAAAGATGCGGCCCGG GCTCCGGACCGAAAAAAAAGAGAAAGGTAAACGGCGGGGGAGGATCACGG ACCCTGGTTACATTTAAGGACGTTTTTGTAGACTTCACGAGGGAGGAATGGA AGCTGTTGGACACCGCCCAGCAAATCGTGTATCGAAATGTAATGTTGGAAAA TTATAAAAATCTGGTTTCCCTCGGCTACCAATTGACAAAACCAGACGTCATCC TCAGACTGGAAAAAGGGGAAGAACCTTGGCTTGTCTGA	
979	ATGCACGATCAGGAGTTTGACCCCCCTAAGGTGTACCCACCCGTGCCAGCCG AGAAGAGGAAGCCCATCCGCGTGCTGTCCCTGTTCGACGGCATCGCCACAGG CCTGCTGGTGCTGAAGGATCTGGGCATCCAGGTGGACAGATATATCGCCTCC GAGGTGTGCGAGGATTCTATCACCGTGGGCATGGTGAGGCACCAGGGCAAG ATCATGTACGTGGGCGACGTGCGCAGCGTGACACAGAAGCACATCCAGGAGT GGGGACCCTTCGACCTGGTCATCGGAGGCAGCCCCTGTAATGACCTGTCCAT CGTGAACCCTGCAAGGAAGGGCCTGTATGAGGGAACCGGCAGACTGTTCTTT GAGTTCTACAGGCTGCTGCACGACGCCCGCCCTAAGGAGGGCGATGACAGGC CATTCTTTTGGCTGTTTGAGAACGTGGTGGCCATGGGCGTGAGCGACAAGCG GGATATCTCCAGATTCCTGGAGTCTAATCCCGTGATGATCGATGCAAAGGAG GTGTCTGCCGCACACAGGGCAAGGTACTTTTGGGGAAATCTGCCTGGCATGA ACCGCCCACTGGCCAGCACCGTGAACGACAAGCTGGAGCTGCAGGAGTGCCT GGAGCACGGAAGGATCGCCAAGTTCTCCAAGGTGCGGACAATCACCACAAG ATCTAACAGCATCAAGCAGGGCAAGGATCAGCACTTCCCCGTGTTCATGAAT GAGAAGGAGGACATCCTGTGGTGTACCGAGATGGAGCGCGTGTTCGGCTTTC CAGTGCACTATACAGACGTGAGCAATATGAGCCGGCTGGCAAGGCAGAGAC TGCTGGGCCGGTCCTGGTCTGTGCCAGTGATCAGACACCTGTTCGCCCCCCTG AAGGAGTACTTTGCCTGCGTGTCTAGCGGCAACTCTAATGCCAACAGCAGAG GCCCTTCCTTTTCCTCTGGCCTGGTGCCACTGTCTCTGAGGGGCAGCCACATG GGCCCCATGGAGATCTACAAGACCGTGTCCGCCTGGAAGAGGCAGCCTGTGC GCGTGCTGTCTCTGTTCCGCAACATCGACAAGGTGCTGAAGAGCCTGGGCTTT CTGGAGAGCGGATCCGGATCTGGAGGAGGCACCCTGAAGTATGTGGAGGAT GTGACAAATGTGGTGCGGAGAGATGTGGAGAAGTGGGGCCCCTTCGATCTGG TGTACGGATCCACCCAGCCACTGGGAAGCTCCTGCGATAGGTGTCCAGGATG GTATATGTTCCAGTTTCACAGAATCCTGCAGTACGCACTGCCAAGGCAGGAG AGCCAGCGCCCTTTCTTTTGGATCTTTATGGACAACCTGCTGCTGACAGAGGA TGACCAGGAGACAACAACCCGCTTCCTGCAGACAGAGGCAGTGACCCTGCAG GATGTGAGGGGACGCGACTATCAGAATGCCATGCGGGTGTGGTCTAACATCC CTGGCCTGAAGAGCAAGCACGCCCCCCTGACCCCTAAGGAGGAGGAGTACCT GCAGGCCCAGGTGCGGAGCAGATCCAAGCTGGATGCCCCTAAGGTGGACCTG CTGGTGAAGAATTGTCTGCTGCCACTGCGGGAGTACTTCAAGTACTTTAGTCA GAATAGCCTGCCACTGGGTGGACCTAGCTCTGGAGCCCCGCCTCCGTCAGGT GGATCTCCTGCAGGTAGCCCAACTTCCACGGAAGAGGGCACGTCTGAGAGTG CGACTCCTGAGAGCGGGCCTGGCACAAGTACTGAGCCCAGCGAAGGTTCTGC ACCCGGGTCTCCAGCTGGGAGCCCTACCTCAACAGAGGAAGGTACCAGCACA GAGCCTTCTGAGGGTAGCGCTCCTGGCACGTCCACCGAACCGTCCGAGGGGA TGGCACCGAAGAAAAAACGCAAAGTTGGTCGAGCCGCTATGGCTGAAAGAC CATTCCAGTGCAGAATATGTATGAGGAATTTTTCTAATCCCGCGAACCTTACG CGCCATATCAGGACGCACACGGGCGAAAAGCCCTTCGCCTGCGACATTTGTG GGAGAAAGTTTGCTCAAAACGCGACCAGGACAAAGCACACGAAAATTCACA CTGGTAGCCAGAAGCCGTTCCAGTGTAGGATCTGTATGCGCAATTTCTCTCAG TCCGGGCACCTCGCGCGACACATAAGAACTCATACGGGGGAGAAGCCGTTTG	3xFLAG- DNMT3AL- XTEN80-SV40 NLS-ZFP-SV40 NLS-KRAB (DNA 서열)

[1278]

	CATGTGACATCTGCGGCCGCAAGTTTGCGAATAGGCATGACAGGGCAAAACA TACGAAGATCCATACAGGTTCTCAAAAACCTTTCCAATGTCGAATATGCATG CGCAACTTTAGTCGGTCAGACCACCTTTCTGAACACATCAGGACACACACTG GCGAAAAGCCGTTCGCATGTGACATTTGCGGCAGAAAGTTCGCACAAAGACG GTCCCGCTATAAGCACACCAAAATTCACCTTAGGCAAAAGGATGCAGCTCGG GCTCCGGACCGAAAAAAAAAGAGAAAGGTAAACGGCGGGGGAGGATCACGG ACCCTGGTTACATTTAAGGACGTTTTTGTAGACTTCACGAGGGAGGAATGGA AGCTGTTGGACACCGCCCAGCAAATCGTGTATCGAAATGTAATGTTGGAAAA TTATAAAAATCTGGTTTCCCTCGGCTACCAATTGACAAAACCAGACGTCATCC TCAGACTGGAAAAAGGGGAAGAACCTTGGCTTGTCTGA	
980	ATGCACGATCAGGAGTTTGACCCCCCTAAGGTGTACCCACCCGTGCCAGCCG AGAAGAGGAAGCCCATCCGCGTGCTGTCCCTGTTCGACGGCATCGCCACAGG CCTGCTGGTGCTGAAGGATCTGGGCATCCAGGTGGACAGATATATCGCCTCC GAGGTGTGCGAGGATTCTATCACCGTGGGCATGGTGAGGCACCAGGGCAAG ATCATGTACGTGGGCGACGTGCGCAGCGTGACACAGAAGCACATCCAGGAGT GGGGACCCTTCGACCTGGTCATCGGAGGCAGCCCCTGTAATGACCTGTCCAT CGTGAACCCTGCAAGGAAGGGCCTGTATGAGGGAACCGGCAGACTGTTCTTT GAGTTCTACAGGCTGCTGCACGACGCCCGCCCTAAGGAGGGCGATGACAGGC CATTCTTTTGGCTGTTTGAGAACGTGGTGGCCATGGGCGTGAGCGACAAGCG GGATATCTCCAGATTCCTGGAGTCTAATCCCGTGATGATCGATGCAAAGGAG GTGTCTGCCGCACACAGGGCAAGGTACTTTTGGGGAAATCTGCCTGGCATGA ACCGCCCACTGGCCAGCACCGTGAACGACAAGCTGGAGCTGCAGGAGTGCCT GGAGCACGGAAGGATCGCCAAGTTCTCCAAGGTGCGGACAATCACCACAAG ATCTAACAGCATCAAGCAGGGCAAGGATCAGCACTTCCCCGTGTTCATGAAT GAGAAGGAGGACATCCTGTGGTGTACCGAGATGGAGCGCGTGTTCGGCTTTC CAGTGCACTATACAGACGTGAGCAATATGAGCCGGCTGGCAAGGCAGAGAC TGCTGGGCCGGTCCTGGTCTGTGCCAGTGATCAGACACCTGTTCGCCCCCCTG AAGGAGTACTTTGCCTGCGTGTCTAGCGGCAACTCTAATGCCAACAGCAGAG GCCCTTCCTTTTCCTCTGGCCTGGTGCCACTGTCTCTGAGGGGCAGCCACATG GGCCCCATGGAGATCTACAAGACCGTGTCCGCCTGGAAGAGGCAGCCTGTGC GCGTGCTGTCTCTGTTCCGCAACATCGACAAGGTGCTGAAGAGCCTGGGCTTT CTGGAGAGCGGATCCGGATCTGGAGGAGGCACCCTGAAGTATGTGGAGGAT GTGACAAATGTGGTGCGGAGAGATGTGGAGAAGTGGGGCCCCTTCGATCTGG TGTACGGATCCACCCAGCCACTGGGAAGCTCCTGCGATAGGTGTCCAGGATG GTATATGTTCCAGTTTCACAGAATCCTGCAGTACGCACTGCCAAGGCAGGAG AGCCAGCGCCCTTTCTTTTGGATCTTTATGGACAACCTGCTGCTGACAGAGGA TGACCAGGAGACAACAACCCGCTTCCTGCAGACAGAGGCAGTGACCCTGCAG GATGTGAGGGGACGCGACTATCAGAATGCCATGCGGGTGTGGTCTAACATCC CTGGCCTGAAGAGCAAGCACGCCCCCCTGACCCCTAAGGAGGAGGAGTACCT GCAGGCCCAGGTGCGGAGCAGATCCAAGCTGGATGCCCCTAAGGTGGACCTG CTGGTGAAGAATTGTCTGCTGCCACTGCGGGAGTACTTCAAGTACTTTAGTCA GAATAGCCTGCCACTGGGTGGACCTAGCTCTGGAGCCCCGCCTCCGTCAGGT GGATCTCCTGCAGGTAGCCCAACTTCCACGGAAGAGGGCACGTCTGAGAGTG CGACTCCTGAGAGCGGGCCTGGCACAAGTACTGAGCCCAGCGAAGGTTCTGC ACCCGGGTCTCCAGCTGGGAGCCCTACCTCAACAGAGGAAGGTACCAGCACA GAGCCTTCTGAGGGTAGCGCTCCTGGCACGTCCACCGAACCGTCCGAGGGGA TGGCACCGAAGAAAAAACGCAAAGTTGGTCGAGCGGCAATGGCAGAACGAC CCTTCCAATGCCGCATATGTATGCGAAACTTCAGCCAGAGCTCAGATCTTTCC AGACACATCAGGACTCACACTGGCGAAAAACCATTTGCATGCGATATATGCG GGAGAAAATTCGCGCACCGCAGTACGCGAAACAGGCATACAAAGATACATA CTGGCAGTCAAAAGCCATTTCAATGTCGAATATGCATGAGGAACTTTAGTCG ATCTGACGTGCTGAGCGCTCACATACGGACCCATACCGGAGAGAAACCATTC GCTTGTGACATCTGTGGTAGGAAGTTCGCGGATTCCCGGACCCGCAAAAATC ATACTAAAATTCACACTGGGTCTCAGAAGCCCTTTCAGTGTAGGATATGTATG CGCAATTTTAGCCAGAGTGGTTCATTGACTCGGCATATCAGAACACATACTG GAGAGAAACCTTTCGCGTGTGATATTTGCGGTCGAAAGTTCGCAGATCAGAG TGGACTTGCGCACCATACTAAGATCCACCTGAGACAGAAGGACGCTGCGCGG GCTCCGGACCGAAAAAAAAAGAGAAAGGTAAACGGCGGGGGAGGATCACGG ACCCTGGTTACATTTAAGGACGTTTTTGTAGACTTCACGAGGGAGGAATGGA	3xFLAG- DNMT3AL- XTEN80-SV40 NLS-ZFP-SV40 NLS-KRAB (DNA 서열)

[1279]

	AGCTGTTGGACACCGCCCAGCAAATCGTGTATCGAAATGTAATGTTGGAAAA TTATAAAAATCTGGTTTCCCTCGGCTACCAATTGACAAAACCAGACGTCATCC TCAGACTGGAAAAAGGGGAAGAACCTTGGCTTGTCTGA	
981	ATGCACGATCAGGAGTTTGACCCCCCTAAGGTGTACCCACCCGTGCCAGCCG AGAAGAGGAAGCCCATCCGCGTGCTGTCCCTGTTCGACGGCATCGCCACAGG CCTGCTGGTGCTGAAGGATCTGGGCATCCAGGTGGACAGATATATCGCCTCC GAGGTGTGCGAGGATTCTATCACCGTGGGCATGGTGAGGCACCAGGGCAAG ATCATGTACGTGGGCGACGTGCGCAGCGTGACACAGAAGCACATCCAGGAGT GGGGACCCTTCGACCTGGTCATCGGAGGCAGCCCCTGTAATGACCTGTCCAT CGTGAACCCTGCAAGGAAGGGCCTGTATGAGGGAACCGGCAGACTGTTCTTT GAGTTCTACAGGCTGCTGCACGACGCCCGCCCTAAGGAGGGCGATGACAGGC CATTCTTTTGGCTGTTTGAGAACGTGGTGGCCATGGGCGTGAGCGACAAGCG GGATATCTCCAGATTCCTGGAGTCTAATCCCGTGATGATCGATGCAAAGGAG GTGTCTGCCGCACACAGGGCAAGGTACTTTTGGGGAAATCTGCCTGGCATGA ACCGCCCACTGGCCAGCACCGTGAACGACAAGCTGGAGCTGCAGGAGTGCCT GGAGCACGGAAGGATCGCCAAGTTCTCCAAGGTGCGGACAATCACCACAAG ATCTAACAGCATCAAGCAGGGCAAGGATCAGCACTTCCCCGTGTTCATGAAT GAGAAGGAGGACATCCTGTGGTGTACCGAGATGGAGCGCGTGTTCGGCTTTC CAGTGCACTATACAGACGTGAGCAATATGAGCCGGCTGGCAAGGCAGAGAC TGCTGGGCCGGTCCTGGTCTGTGCCAGTGATCAGACACCTGTTCGCCCCCCTG AAGGAGTACTTTGCCTGCGTGTCTAGCGGCAACTCTAATGCCAACAGCAGAG GCCCTTCCTTTTCCTCTGGCCTGGTGCCACTGTCTCTGAGGGGCAGCCACATG GGCCCCATGGAGATCTACAAGACCGTGTCCGCCTGGAAGAGGCAGCCTGTGC GCGTGCTGTCTCTGTTCCGCAACATCGACAAGGTGCTGAAGAGCCTGGGCTTT CTGGAGAGCGGATCCGGATCTGGAGGAGGCACCCTGAAGTATGTGGAGGAT GTGACAAATGTGGTGCGGAGAGATGTGGAGAAGTGGGGCCCCTTCGATCTGG TGTACGGATCCACCCAGCCACTGGGAAGCTCCTGCGATAGGTGTCCAGGATG GTATATGTTCCAGTTTCACAGAATCCTGCAGTACGCACTGCCAAGGCAGGAG AGCCAGCGCCCTTTCTTTTGGATCTTTATGGACAACCTGCTGCTGACAGAGGA TGACCAGGAGACAACAACCCGCTTCCTGCAGACAGAGGCAGTGACCCTGCAG GATGTGAGGGGACGCGACTATCAGAATGCCATGCGGGTGTGGTCTAACATCC CTGGCCTGAAGAGCAAGCACGCCCCCCTGACCCCTAAGGAGGAGGAGTACCT GCAGGCCCAGGTGCGGAGCAGATCCAAGCTGGATGCCCCTAAGGTGGACCTG CTGGTGAAGAATTGTCTGCTGCCACTGCGGGAGTACTTCAAGTACTTTAGTCA GAATAGCCTGCCACTGGGTGGACCTAGCTCTGGAGCCCCGCCTCCGTCAGGT GGATCTCCTGCAGGTAGCCCAACTTCCACGGAAGAGGGCACGTCTGAGAGTG CGACTCCTGAGAGCGGGCCTGGCACAAGTACTGAGCCCAGCGAAGGTTCTGC ACCCGGGTCTCCAGCTGGGAGCCCTACCTCAACAGAGGAAGGTACCAGCACA GAGCCTTCTGAGGGTAGCGCTCCTGGCACGTCCACCGAACCGTCCGAGGGGA TGGCACCGAAGAAAAAACGCAAAGTTGGTCGAGCTGCCATGGCGGAGCGCC CTTTCCAGTGTAGGATATGTATGCGCAACTTCAGTCAGAACCCAGCCCAGTG GCGGCACATACGGACGCATACTGGAGAAAAGCCATTTGCATGTGATATCTGC GGGCGAAAATTCGCGCGGTCAGCAGATTTGAGCCGGCATACGAAGATCCATA CAGGTTCACAAAAGCCATTTCAATGTCGGATATGTATGCGGAACTTCAGCAC GTCCGGCTCATTGTCAAGACATATACGAACTCATACCGGAGAGAAACCCTTC GCGTGCGACATTTGCGGTCGGAAGTTCGCGCGATCCGACCATCTGTCACGAC ATACGAAAATACACACTGGCTCTCAAAAGCCGTTTCAGTGCAGAATTTGCAT GAGAAATTTTAGCAGGAGCGACTCACTCCTTCGGCATATACGAACACACACT GGTGAGAAGCCATTTGCCTGTGATATTTGTGGACGAAAGTTTGCGCAATCTTA CGATAGGTTTCAGCATACAAAAATCCACCTTCGGCAAAAGGACGCGGCACGG GCTCCGGACCGAAAAAAAAGAGAAAGGTAAACGGCGGGGGAGGATCACGG ACCCTGGTTACATTTAAGGACGTTTTTGTAGACTTCACGAGGGAGGAATGGA AGCTGTTGGACACCGCCCAGCAAATCGTGTATCGAAATGTAATGTTGGAAAA TTATAAAAATCTGGTTTCCCTCGGCTACCAATTGACAAAACCAGACGTCATCC TCAGACTGGAAAAAGGGGAAGAACCTTGGCTTGTCTGA	3xFLAG- DNMT3AL- XTEN80-SV40 NLS-ZFP-SV40 NLS-KRAB (DNA 서열)
982	ATGCACGATCAGGAGTTTGACCCCCCTAAGGTGTACCCACCCGTGCCAGCCG AGAAGAGGAAGCCCATCCGCGTGCTGTCCCTGTTCGACGGCATCGCCACAGG CCTGCTGGTGCTGAAGGATCTGGGCATCCAGGTGGACAGATATATCGCCTCC GAGGTGTGCGAGGATTCTATCACCGTGGGCATGGTGAGGCACCAGGGCAAG	3xFLAG- DNMT3AL- XTEN80-SV40 NLS-ZFP-SV40

[1280]

	ATCATGTACGTGGGCGACGTGCGCAGCGTGACACAGAAGCACATCCAGGAGT GGGGACCCTTCGACCTGGTCATCGGAGGCAGCCCCTGTAATGACCTGTCCAT CGTGAACCCTGCAAGGAAGGGCCTGTATGAGGGAACCGGCAGACTGTTCTTT GAGTTCTACAGGCTGCTGCACGACGCCCGCCCTAAGGAGGGCGATGACAGGC CATTCTTTTGGCTGTTTGAGAACGTGGTGGCCATGGGCGTGAGCGACAAGCG GGATATCTCCAGATTCCTGGAGTCTAATCCCGTGATGATCGATGCAAAGGAG GTGTCTGCCGCACACAGGGCAAGGTACTTTTGGGGAAATCTGCCTGGCATGA ACCGCCCACTGGCCAGCACCGTGAACGACAAGCTGGAGCTGCAGGAGTGCCT GGAGCACGGAAGGATCGCCAAGTTCTCCAAGGTGCGGACAATCACCACAAG ATCTAACAGCATCAAGCAGGGCAAGGATCAGCACTTCCCCGTGTTCATGAAT GAGAAGGAGGACATCCTGTGGTGTACCGAGATGGAGCGCGTGTTCGGCTTTC CAGTGCACTATACAGACGTGAGCAATATGAGCCGGCTGGCAAGGCAGAGAC TGCTGGGCCGGTCCTGGTCTGTGCCAGTGATCAGACACCTGTTCGCCCCCCTG AAGGAGTACTTTGCCTGCGTGTCTAGCGGCAACTCTAATGCCAACAGCAGAG GCCCTTCCTTTTCCTCTGGCCTGGTGCCACTGTCTCTGAGGGGCAGCCACATG GGCCCCATGGAGATCTACAAGACCGTGTCCGCCTGGAAGAGGCAGCCTGTGC GCGTGCTGTCTCTGTTCCGCAACATCGACAAGGTGCTGAAGAGCCTGGGCTTT CTGGAGAGCGGATCCGGATCTGGAGGAGGCACCCTGAAGTATGTGGAGGAT GTGACAAATGTGGTGCGGAGAGATGTGGAGAAGTGGGGCCCCTTCGATCTGG TGTACGGATCCACCCAGCCACTGGGAAGCTCCTGCGATAGGTGTCCAGGATG GTATATGTTCCAGTTTCACAGAATCCTGCAGTACGCACTGCCAAGGCAGGAG AGCCAGCGCCCTTTCTTTTGGATCTTTATGGACAACCTGCTGCTGACAGAGGA TGACCAGGAGACAACAACCCGCTTCCTGCAGACAGAGGCAGTGACCCTGCAG GATGTGAGGGGACGCGACTATCAGAATGCCATGCGGGTGTGGTCTAACATCC CTGGCCTGAAGAGCAAGCACGCCCCCCTGACCCCTAAGGAGGAGGAGTACCT GCAGGCCCAGGTGCGGAGCAGATCCAAGCTGGATGCCCCTAAGGTGGACCTG CTGGTGAAGAATTGTCTGCTGCCACTGCGGGAGTACTTCAAGTACTTTAGTCA GAATAGCCTGCCACTGGGTGGACCTAGCTCTGGAGCCCCGCCTCCGTCAGGT GGATCTCCTGCAGGTAGCCCAACTTCCACGGAAGAGGGCACGTCTGAGAGTG CGACTCCTGAGAGCGGGCCTGGCACAAGTACTGAGCCCAGCGAAGGTTCTGC ACCCGGGTCTCCAGCTGGGAGCCCTACCTCAACAGAGGAAGGTACCAGCACA GAGCCTTCTGAGGGTAGCGCTCCTGGCACGTCCACCGAACCGTCCGAGGGGA TGGCACCGAAGAAAAAACGCAAAGTTGGTCGAGCTGCCATGGCTGAACGAC CGTTTCAATGTCGAATTTGCATGCGCAACTTCTCCACGTCCGGGTCTCTCAGT AGACACATCAGAACGCATACTGGTGAAAAACCATTCGCTTGTGACATATGCG GCCGAAAATTCGCGCGGAGCGACCACCTGTCACGGCATACCAAAATTCACAC CGGGAGTCAAAAACCGTTCCAGTGTAGGATATGTATGCGCAACTTCAGCCGG TCTGACAGTCTGCTTCGACATATTCGGACGCACACTGGTGAAAAGCCGTTTGC GTGCGACATTTGTGGTCGAAAGTTCGCTCAATCTTATGATAGGTTTCAACACA CCAAAATACATACGGGCTCCCAGAAGCCGTTCCAGTGCAGAATATGCATGAG AAATTTCTCTCGCAGTGACAATTTGTCCACCCATATTCGAACGCACACCGGCG AGAAACCCTTCGCCTGCGATATTTGCGGTCGCAAGTTCGCAGACAACAGGGA TAGGATAAAACATACGAAGATCCATCTGAGGCAAAAAGACGCCGCCCGGGC TCCGGACCGAAAAAAAAGAGAAAGGTAAACGGCGGGGGAGGATCACGGACC CTGGTTACATTTAAGGACGTTTTTGTAGACTTCACGAGGGAGGAATGGAAGC TGTTGGACACCGCCCAGCAAATCGTGTATCGAAATGTAATGTTGGAAAATTA TAAAAATCTGGTTTCCCTCGGCTACCAATTGACAAAACCAGACGTCATCCTCA GACTGGAAAAAGGGGAAGAACCTTGGCTTGTCTGA	NLS-KRAB (DNA 서열)
983	ATGCACGATCAGGAGTTTGACCCCCCTAAGGTGTACCCACCCGTGCCAGCCG AGAAGAGGAAGCCCATCCGCGTGCTGTCCCTGTTCGACGGCATCGCCACAGG CCTGCTGGTGCTGAAGGATCTGGGCATCCAGGTGGACAGATATATCGCCTCC GAGGTGTGCGAGGATTCTATCACCGTGGGCATGGTGAGGCACCAGGGCAAG ATCATGTACGTGGGCGACGTGCGCAGCGTGACACAGAAGCACATCCAGGAGT GGGGACCCTTCGACCTGGTCATCGGAGGCAGCCCCTGTAATGACCTGTCCAT CGTGAACCCTGCAAGGAAGGGCCTGTATGAGGGAACCGGCAGACTGTTCTTT GAGTTCTACAGGCTGCTGCACGACGCCCGCCCTAAGGAGGGCGATGACAGGC CATTCTTTTGGCTGTTTGAGAACGTGGTGGCCATGGGCGTGAGCGACAAGCG GGATATCTCCAGATTCCTGGAGTCTAATCCCGTGATGATCGATGCAAAGGAG GTGTCTGCCGCACACAGGGCAAGGTACTTTTGGGGAAATCTGCCTGGCATGA	3xFLAG- DNMT3AL- XTEN80-SV40 NLS-ZFP-SV40 NLS-KRAB (DNA 서열)

[1281]

	ACCGCCCACTGGCCAGCACCGTGAACGACAAGCTGGAGCTGCAGGAGTGCCT GGAGCACGGAAGGATCGCCAAGTTCTCCAAGGTGCGGACAATCACCACAAG ATCTAACAGCATCAAGCAGGGCAAGGATCAGCACTTCCCCGTGTTCATGAAT GAGAAGGAGGACATCCTGTGGTGTACCGAGATGGAGCGCGTGTTCGGCTTTC CAGTGCACTATACAGACGTGAGCAATATGAGCCGGCTGGCAAGGCAGAGAC TGCTGGGCCGGTCCTGGTCTGTGCCAGTGATCAGACACCTGTTCGCCCCCCTG AAGGAGTACTTTGCCTGCGTGTCTAGCGGCAACTCTAATGCCAACAGCAGAG GCCCTTCCTTTTCCTCTGGCCTGGTGCCACTGTCTCTGAGGGGCAGCCACATG GGCCCCATGGAGATCTACAAGACCGTGTCCGCCTGGAAGAGGCAGCCTGTGC GCGTGCTGTCTCTGTTCCGCAACATCGACAAGGTGCTGAAGAGCCTGGGCTTT CTGGAGAGCGGATCCGGATCTGGAGGAGGCACCCTGAAGTATGTGGAGGAT GTGACAAATGTGGTGCGGAGAGATGTGGAGAAGTGGGGCCCCTTCGATCTGG TGTACGGATCCACCCAGCCACTGGGAAGCTCCTGCGATAGGTGTCCAGGATG GTATATGTTCCAGTTTCACAGAATCCTGCAGTACGCACTGCCAAGGCAGGAG AGCCAGCGCCCTTTCTTTTGGATCTTTATGGACAACCTGCTGCTGACAGAGGA TGACCAGGAGACAACAACCCGCTTCCTGCAGACAGAGGCAGTGACCCTGCAG GATGTGAGGGGACGCGACTATCAGAATGCCATGCGGGTGTGGTCTAACATCC CTGGCCTGAAGAGCAAGCACGCCCCCCTGACCCCTAAGGAGGAGGAGTACCT GCAGGCCCAGGTGCGGAGCAGATCCAAGCTGGATGCCCCTAAGGTGGACCTG CTGGTGAAGAATTGTCTGCTGCCACTGCGGGAGTACTTCAAGTACTTTAGTCA GAATAGCCTGCCACTGGGTGGACCTAGCTCTGGAGCCCCGCCTCCGTCAGGT GGATCTCCTGCAGGTAGCCCAACTTCCACGGAAGAGGGCACGTCTGAGAGTG CGACTCCTGAGAGCGGGCCTGGCACAAGTACTGAGCCCAGCGAAGGTTCTGC ACCCGGGTCTCCAGCTGGGAGCCCTACCTCAACAGAGGAAGGTACCAGCACA GAGCCTTCTGAGGGTAGCGCTCCTGGCACGTCCACCGAACCGTCCGAGGGGA TGGCACCGAAGAAAAAACGCAAAGTTGGTCGAGCAGCCATGGCAGAGCGGC CATTCCAGTGCAGAATCTGCATGCGGAACTTTTCCGATAGGTCCAATCTGTCA CGCCATATTAGGACACACACGGGTGAAAAACCGTTCGCGTGTGACATATGCG GTCGCAAATTCGCCCTGAGACAGAACCTGATTATGCACACAAAAATACATAC GGGAAGCCAGAAACCGTTCCAGTGTCGGATATGCATGAGGAACTTCAGTGAG AGGGGGACTTTGGCGAGGCACATCAGGACTCACACTGGGGAGAAGCCCTTTG CATGTGATATCTGTGGCCGAAAATTTGCTCGATCAGATGCTCTCACCCAACAT ACAAAGATCCATACTGGCTCTCAAAAACCGTTTCAATGTAGAATTTGTATGC GCAACTTCTCTCGGTCAGATAGCCTGTCCCAGCATATCCGAACTCATACAGGT GAGAAACCCTTCGCATGCGACATCTGTGGGCGAAAATTTGCTAGAAAAGCAG ACCGGACCCGACACACAAAGATTCATCTGCGACAAAAAGACGCCGCCCGGG CTCCGGACCGAAAAAAAAGAGAAAGGTAAACGGCGGGGGAGGATCACGGAC CCTGGTTACATTTAAGGACGTTTTTGTAGACTTCACGAGGGAGGAATGGAAG CTGTTGGACACCGCCCAGCAAATCGTGTATCGAAATGTAATGTTGGAAAATT ATAAAAATCTGGTTTCCCTCGGCTACCAATTGACAAAACCAGACGTCATCCTC AGACTGGAAAAAGGGGAAGAACCTTGGCTTGTCTGA	
984	ATGCACGATCAGGAGTTTGACCCCCCTAAGGTGTACCCACCCGTGCCAGCCG AGAAGAGGAAGCCCATCCGCGTGCTGTCCCTGTTCGACGGCATCGCCACAGG CCTGCTGGTGCTGAAGGATCTGGGCATCCAGGTGGACAGATATATCGCCTCC GAGGTGTGCGAGGATTCTATCACCGTGGGCATGGTGAGGCACCAGGGCAAG ATCATGTACGTGGGCGACGTGCGCAGCGTGACACAGAAGCACATCCAGGAGT GGGGACCCTTCGACCTGGTCATCGGAGGCAGCCCCTGTAATGACCTGTCCAT CGTGAACCCTGCAAGGAAGGGCCTGTATGAGGGAACCGGCAGACTGTTCTTT GAGTTCTACAGGCTGCTGCACGACGCCCGCCCTAAGGAGGGCGATGACAGGC CATTCTTTTGGCTGTTTGAGAACGTGGTGGCCATGGGCGTGAGCGACAAGCG GGATATCTCCAGATTCCTGGAGTCTAATCCCGTGATGATCGATGCAAAGGAG GTGTCTGCCGCACACAGGGCAAGGTACTTTTGGGGAAATCTGCCTGGCATGA ACCGCCCACTGGCCAGCACCGTGAACGACAAGCTGGAGCTGCAGGAGTGCCT GGAGCACGGAAGGATCGCCAAGTTCTCCAAGGTGCGGACAATCACCACAAG ATCTAACAGCATCAAGCAGGGCAAGGATCAGCACTTCCCCGTGTTCATGAAT GAGAAGGAGGACATCCTGTGGTGTACCGAGATGGAGCGCGTGTTCGGCTTTC CAGTGCACTATACAGACGTGAGCAATATGAGCCGGCTGGCAAGGCAGAGAC TGCTGGGCCGGTCCTGGTCTGTGCCAGTGATCAGACACCTGTTCGCCCCCCTG AAGGAGTACTTTGCCTGCGTGTCTAGCGGCAACTCTAATGCCAACAGCAGAG	3xFLAG- DNMT3AL- XTEN80-SV40 NLS-ZFP-SV40 NLS-KRAB (DNA 서열)

[1282]

	GCCCTTCCTTTTCCTCTGGCCTGGTGCCACTGTCTCTGAGGGGCAGCCACATG GGCCCCATGGAGATCTACAAGACCGTGTCCGCCTGGAAGAGGCAGCCTGTGC GCGTGCTGTCTCTGTTCCGCAACATCGACAAGGTGCTGAAGAGCCTGGGCTTT CTGGAGAGCGGATCCGGATCTGGAGGAGGCACCCTGAAGTATGTGGAGGAT GTGACAAATGTGGTGCGGAGAGATGTGGAGAAGTGGGGCCCCTTCGATCTGG TGTACGGATCCACCCAGCCACTGGGAAGCTCCTGCGATAGGTGTCCAGGATG GTATATGTTCCAGTTTCACAGAATCCTGCAGTACGCACTGCCAAGGCAGGAG AGCCAGCGCCCTTTCTTTTGGATCTTTATGGACAACCTGCTGCTGACAGAGGA TGACCAGGAGACAACAACCCGCTTCCTGCAGACAGAGGCAGTGACCCTGCAG GATGTGAGGGGACGCGACTATCAGAATGCCATGCGGGTGTGGTCTAACATCC CTGGCCTGAAGAGCAAGCACGCCCCCCTGACCCCTAAGGAGGAGGAGTACCT GCAGGCCCAGGTGCGGAGCAGATCCAAGCTGGATGCCCCTAAGGTGGACCTG CTGGTGAAGAATTGTCTGCTGCCACTGCGGGAGTACTTCAAGTACTTTAGTCA GAATAGCCTGCCACTGGGTGGACCTAGCTCTGGAGCCCCGCCTCCGTCAGGT GGATCTCCTGCAGGTAGCCCAACTTCCACGGAAGAGGGCACGTCTGAGAGTG CGACTCCTGAGAGCGGGCCTGGCACAAGTACTGAGCCCAGCGAAGGTTCTGC ACCCGGGTCTCCAGCTGGGAGCCCTACCTCAACAGAGGAAGGTACCAGCACA GAGCCTTCTGAGGGTAGCGCTCCTGGCACGTCCACCGAACCGTCCGAGGGGA TGGCACCGAAGAAAAAACGCAAAGTTGGTCGAGCGGCCATGGCTGAGAGGC CTTTTCAATGTAGAATATGTATGCGAAATTTTTCACAGTACTGTTGTCTCACG AACCACATAAGGACTCATACAGGGGAGAAACCATTTGCCTGTGACATTTGCG GTCGCAAATTTGCTACTTCTGGAAACCTGACTCGGCACACTAAGATTCACAC AGGGTCCCAGAAGCCCTTCCAGTGTCGCATTTGCATGAGGAATTTTAGTCAA AGCTCTGACTTGTCAAGGCATATTCGCACGCACACGGGCGAAAAGCCGTTCG CTTGCGACATATGCGGGCGGAAATTTGCCTTCCGCTATTATTTGAAGAGACAC ACCAAGATACATACGGGCTCTCAGAAGCCCTTTCAGTGTAGGATTTGCATGC GCAATTTTTCACAATCTGGTGATCTCACGCGACACATCCGGACTCACACAGGT GAAAAGCCTTTCGCGTGCGACATTTGCGGCCGGAAGTTTGCTGACAAGGGCA ACCTCACAAAGCATACGAAGATTCACTTGAGGCAGAAAGATGCTGCTCGGGC TCCGGACCGAAAAAAAAGAGAAAGGTAAACGGCGGGGGAGGATCACGGACC CTGGTTACATTTAAGGACGTTTTTGTAGACTTCACGAGGGAGGAATGGAAGC TGTTGGACACCGCCCAGCAAATCGTGTATCGAAATGTAATGTTGGAAAATTA TAAAAATCTGGTTTCCCTCGGCTACCAATTGACAAAACCAGACGTCATCCTCA GACTGGAAAAAGGGGAAGAACCTTGGCTTGTCTGA	
985	ATGCACGATCAGGAGTTTGACCCCCCTAAGGTGTACCCACCCGTGCCAGCCG AGAAGAGGAAGCCCATCCGCGTGCTGTCCCTGTTCGACGGCATCGCCACAGG CCTGCTGGTGCTGAAGGATCTGGGCATCCAGGTGGACAGATATATCGCCTCC GAGGTGTGCGAGGATTCTATCACCGTGGGCATGGTGAGGCACCAGGGCAAG ATCATGTACGTGGGCGACGTGCGCAGCGTGACACAGAAGCACATCCAGGAGT GGGGACCCTTCGACCTGGTCATCGGAGGCAGCCCCTGTAATGACCTGTCCAT CGTGAACCCTGCAAGGAAGGGCCTGTATGAGGGAACCGGCAGACTGTTCTTT GAGTTCTACAGGCTGCTGCACGACGCCCGCCCTAAGGAGGGCGATGACAGGC CATTCTTTTGGCTGTTTGAGAACGTGGTGGCCATGGGCGTGAGCGACAAGCG GGATATCTCCAGATTCCTGGAGTCTAATCCCGTGATGATCGATGCAAAGGAG GTGTCTGCCGCACACAGGGCAAGGTACTTTTGGGGAAATCTGCCTGGCATGA ACCGCCCACTGGCCAGCACCGTGAACGACAAGCTGGAGCTGCAGGAGTGCCT GGAGCACGGAAGGATCGCCAAGTTCTCCAAGGTGCGGACAATCACCACAAG ATCTAACAGCATCAAGCAGGGCAAGGATCAGCACTTCCCCGTGTTCATGAAT GAGAAGGAGGACATCCTGTGGTGTACCGAGATGGAGCGCGTGTTCGGCTTTC CAGTGCACTATACAGACGTGAGCAATATGAGCCGGCTGGCAAGGCAGAGAC TGCTGGGCCGGTCCTGGTCTGTGCCAGTGATCAGACACCTGTTCGCCCCCCTG AAGGAGTACTTTGCCTGCGTGTCTAGCGGCAACTCTAATGCCAACAGCAGAG GCCCTTCCTTTTCCTCTGGCCTGGTGCCACTGTCTCTGAGGGGCAGCCACATG GGCCCCATGGAGATCTACAAGACCGTGTCCGCCTGGAAGAGGCAGCCTGTGC GCGTGCTGTCTCTGTTCCGCAACATCGACAAGGTGCTGAAGAGCCTGGGCTTT CTGGAGAGCGGATCCGGATCTGGAGGAGGCACCCTGAAGTATGTGGAGGAT GTGACAAATGTGGTGCGGAGAGATGTGGAGAAGTGGGGCCCCTTCGATCTGG TGTACGGATCCACCCAGCCACTGGGAAGCTCCTGCGATAGGTGTCCAGGATG GTATATGTTCCAGTTTCACAGAATCCTGCAGTACGCACTGCCAAGGCAGGAG	3xFLAG- DNMT3AL- XTEN80-SV40 NLS-ZFP-SV40 NLS-KRAB (DNA 서열)

[1283]

	AGCCAGCGCCCTTTCTTTTGGATCTTTATGGACAACCTGCTGCTGACAGAGGA TGACCAGGAGACAACAACCCGCTTCCTGCAGACAGAGGCAGTGACCCTGCAG GATGTGAGGGGACGCGACTATCAGAATGCCATGCGGGTGTGGTCTAACATCC CTGGCCTGAAGAGCAAGCACGCCCCCCTGACCCCTAAGGAGGAGGAGTACCT GCAGGCCCAGGTGCGGAGCAGATCCAAGCTGGATGCCCCTAAGGTGGACCTG CTGGTGAAGAATTGTCTGCTGCCACTGCGGGAGTACTTCAAGTACTTTAGTCA GAATAGCCTGCCACTGGGTGGACCTAGCTCTGGAGCCCCGCCTCCGTCAGGT GGATCTCCTGCAGGTAGCCCAACTTCCACGGAAGAGGGCACGTCTGAGAGTG CGACTCCTGAGAGCGGGCCTGGCACAAGTACTGAGCCCAGCGAAGGTTCTGC ACCCGGGTCTCCAGCTGGGAGCCCTACCTCAACAGAGGAAGGTACCAGCACA GAGCCTTCTGAGGGTAGCGCTCCTGGCACGTCCACCGAACCGTCCGAGGGGA TGGCACCGAAGAAAAAACGCAAAGTTGGTCGAGCCGCCATGGCCGAACGAC CATTCCAGTGCAGGATATGTATGCGCAATTTTTCAACCAGTGGTTCATTGTCA CGACATATTAGAACACACACCGGTGAGAAACCCTTTGCGTGTGACATCTGTG GGAGGAAATTCGCAAGATCTGACAACCTTACGACACATACAAAGATTCACAC AGGCTCTCAAAAGCCCTTCCAGTGCCGAATTTGCATGCGAAACTTTTCCCAGT CTGGTAATCTCGCTCGACATATCAGAACCCACACGGGGGAAAAACCATTCGC TTGTGATATTTGCGGACGAAAGTTCGCCGACAGAACCACACTCATGAGACAC ACTAAAATCCATACTGGTAGTCAGAAGCCGTTTCAGTGTAGAATCTGCATGA GGAACTTTTCCCAGTCAGGCCACCTTGCAAGACATATACGAACTCACACTGG AGAAAAGCCGTTCGCCTGTGACATTTGTGGGCGCAAGTTCGCGCAACTCACC CATCTGAATAGCCATACGAAGATTCACTTGAGACAGAAAGATGCGGCTCGGG CTCCGGACCGAAAAAAAAAGAGAAAGGTAAACGGCGGGGGAGGATCACGGAC CCTGGTTACATTTAAGGACGTTTTTGTAGACTTCACGAGGGAGGAATGGAAG CTGTTGGACACCGCCCAGCAAATCGTGTATCGAAATGTAATGTTGGAAAATT ATAAAAATCTGGTTTCCCTCGGCTACCAATTGACAAAACCAGACGTCATCCTC AGACTGGAAAAAGGGGAAGAACCTTGGCTTGTCTGA	
986	ATGCACGATCAGGAGTTTGACCCCCCTAAGGTGTACCCACCCGTGCCAGCCG AGAAGAGGAAGCCCATCCGCGTGCTGTCCCTGTTCGACGGCATCGCCACAGG CCTGCTGGTGCTGAAGGATCTGGGCATCCAGGTGGACAGATATATCGCCTCC GAGGTGTGCGAGGATTCTATCACCGTGGGCATGGTGAGGCACCAGGGCAAG ATCATGTACGTGGGCGACGTGCGCAGCGTGACACAGAAGCACATCCAGGAGT GGGGACCCTTCGACCTGGTCATCGGAGGCAGCCCCTGTAATGACCTGTCCAT CGTGAACCCTGCAAGGAAGGGCCTGTATGAGGGAACCGGCAGACTGTTCTTT GAGTTCTACAGGCTGCTGCACGACGCCCGCCCTAAGGAGGGCGATGACAGGC CATTCTTTTGGCTGTTTGAGAACGTGGTGGCCATGGGCGTGAGCGACAAGCG GGATATCTCCAGATTCCTGGAGTCTAATCCCGTGATGATCGATGCAAAGGAG GTGTCTGCCGCACACAGGGCAAGGTACTTTTGGGGAAATCTGCCTGGCATGA ACCGCCCACTGGCCAGCACCGTGAACGACAAGCTGGAGCTGCAGGAGTGCCT GGAGCACGGAAGGATCGCCAAGTTCTCCAAGGTGCGGACAATCACCACAAG ATCTAACAGCATCAAGCAGGGCAAGGATCAGCACTTCCCCGTGTTCATGAAT GAGAAGGAGGACATCCTGTGGTGTACCGAGATGGAGCGCGTGTTCGGCTTTC CAGTGCACTATACAGACGTGAGCAATATGAGCCGGCTGGCAAGGCAGAGAC TGCTGGGCCGGTCCTGGTCTGTGCCAGTGATCAGACACCTGTTCGCCCCCCTG AAGGAGTACTTTGCCTGCGTGTCTAGCGGCAACTCTAATGCCAACAGCAGAG GCCCTTCCTTTTCCTCTGGCCTGGTGCCACTGTCTCTGAGGGGCAGCCACATG GGCCCCATGGAGATCTACAAGACCGTGTCCGCCTGGAAGAGGCAGCCTGTGC GCGTGCTGTCTCTGTTCCGCAACATCGACAAGGTGCTGAAGAGCCTGGGCTTT CTGGAGAGCGGATCCGGATCTGGAGGAGGCACCCTGAAGTATGTGGAGGAT GTGACAAATGTGGTGCGGAGAGATGTGGAGAAGTGGGGCCCCTTCGATCTGG TGTACGGATCCACCCAGCCACTGGGAAGCTCCTGCGATAGGTGTCCAGGATG GTATATGTTCCAGTTTCACAGAATCCTGCAGTACGCACTGCCAAGGCAGGAG AGCCAGCGCCCTTTCTTTTGGATCTTTATGGACAACCTGCTGCTGACAGAGGA TGACCAGGAGACAACAACCCGCTTCCTGCAGACAGAGGCAGTGACCCTGCAG GATGTGAGGGGACGCGACTATCAGAATGCCATGCGGGTGTGGTCTAACATCC CTGGCCTGAAGAGCAAGCACGCCCCCCTGACCCCTAAGGAGGAGGAGTACCT GCAGGCCCAGGTGCGGAGCAGATCCAAGCTGGATGCCCCTAAGGTGGACCTG CTGGTGAAGAATTGTCTGCTGCCACTGCGGGAGTACTTCAAGTACTTTAGTCA GAATAGCCTGCCACTGGGTGGACCTAGCTCTGGAGCCCCGCCTCCGTCAGGT	3xFLAG- DNMT3AL- XTEN80-SV40 NLS-ZFP-SV40 NLS-KRAB (DNA 서열)

[1284]

	GGATCTCCTGCAGGTAGCCCAACTTCCACGGAAGAGGGCACGTCTGAGAGTG CGACTCCTGAGAGCGGGCCTGGCACAAGTACTGAGCCCAGCGAAGGTTCTGC ACCCGGGTCTCCAGCTGGGAGCCCTACCTCAACAGAGGAAGGTACCAGCACA GAGCCTTCTGAGGGTAGCGCTCCTGGCACGTCCACCGAACCGTCCGAGGGGA TGGCACCGAAGAAAAAACGCAAAGTTGGTCGAGCAGCTATGGCTGAACGCC CATTCCAGTGTCGGATCTGCATGCGCAACTTTTCTATAAAACACGATCTTCAC CGACACATTCGGACACATACTGGGGAGAAGCCCTTTGCGTGTGACATCTGTG GCCGAAAGTTCGCTAGATCCGCAAACTTGACTCGGCATACGAAAATTCACAC TGGAAGCCAGAAACCTTTCCAATGTCGAATCTGTATGAGGAACTTTAGCAGA AGTGATAATCTCGCCAGGCATATCCGAACGCACACAGGCGAGAAGCCATTCG CATGTGATATTTGTGGTAGAAAGTTCGCCCAAAATGTCTCTCGCCCACGCCAT ACTAAGATCCACACGGGCTCCCAGAAGCCGTTCCAATGCCGCATTTGCATGC GAAACTTTTCCAGATCAGACGATCTGAGCAAGCATATTAGGACGCATACAGG GGAGAAGCCTTTTGCTTGCGACATTTGCGGCCGGAAATTTGCTGACTCAAGTC ACAGAACACGGCATACCAAGATACACCTTCGACAAAAAGATGCCGCACGGG CTCCGGACCGAAAAAAAAGAGAAAGGTAAACGGCGGGGGAGGATCACGGAC CCTGGTTACATTTAAGGACGTTTTTGTAGACTTCACGAGGGAGGAATGGAAG CTGTTGGACACCGCCCAGCAAATCGTGTATCGAAATGTAATGTTGGAAAATT ATAAAAATCTGGTTTCCCTCGGCTACCAATTGACAAAACCAGACGTCATCCTC AGACTGGAAAAAGGGGAAGAACCTTGGCTTGTCTGA	
987	ATGCACGATCAGGAGTTTGACCCCCCTAAGGTGTACCCACCCGTGCCAGCCG AGAAGAGGAAGCCCATCCGCGTGCTGTCCCTGTTCGACGGCATCGCCACAGG CCTGCTGGTGCTGAAGGATCTGGGCATCCAGGTGGACAGATATATCGCCTCC GAGGTGTGCGAGGATTCTATCACCGTGGGCATGGTGAGGCACCAGGGCAAG ATCATGTACGTGGGCGACGTGCGCAGCGTGACACAGAAGCACATCCAGGAGT GGGGACCCTTCGACCTGGTCATCGGAGGCAGCCCCTGTAATGACCTGTCCAT CGTGAACCCTGCAAGGAAGGGCCTGTATGAGGGAACCGGCAGACTGTTCTTT GAGTTCTACAGGCTGCTGCACGACGCCCGCCCTAAGGAGGGCGATGACAGGC CATTCTTTTGGCTGTTTGAGAACGTGGTGGCCATGGGCGTGAGCGACAAGCG GGATATCTCCAGATTCCTGGAGTCTAATCCCGTGATGATCGATGCAAAGGAG GTGTCTGCCGCACACAGGGCAAGGTACTTTTGGGGAAATCTGCCTGGCATGA ACCGCCCACTGGCCAGCACCGTGAACGACAAGCTGGAGCTGCAGGAGTGCCT GGAGCACGGAAGGATCGCCAAGTTCTCCAAGGTGCGGACAATCACCACAAG ATCTAACAGCATCAAGCAGGGCAAGGATCAGCACTTCCCCGTGTTCATGAAT GAGAAGGAGGACATCCTGTGGTGTACCGAGATGGAGCGCGTGTTCGGCTTTC CAGTGCACTATACAGACGTGAGCAATATGAGCCGGCTGGCAAGGCAGAGAC TGCTGGGCCGGTCCTGGTCTGTGCCAGTGATCAGACACCTGTTCGCCCCCCTG AAGGAGTACTTTGCCTGCGTGTCTAGCGGCAACTCTAATGCCAACAGCAGAG GCCCTTCCTTTTCCTCTGGCCTGGTGCCACTGTCTCTGAGGGGCAGCCACATG GGCCCCATGGAGATCTACAAGACCGTGTCCGCCTGGAAGAGGCAGCCTGTGC GCGTGCTGTCTCTGTTCCGCAACATCGACAAGGTGCTGAAGAGCCTGGGCTTT CTGGAGAGCGGATCCGGATCTGGAGGAGGCACCCTGAAGTATGTGGAGGAT GTGACAAATGTGGTGCGGAGAGATGTGGAGAAGTGGGGCCCCTTCGATCTGG TGTACGGATCCACCCAGCCACTGGGAAGCTCCTGCGATAGGTGTCCAGGATG GTATATGTTCCAGTTTCACAGAATCCTGCAGTACGCACTGCCAAGGCAGGAG AGCCAGCGCCCTTTCTTTTGGATCTTTATGGACAACCTGCTGCTGACAGAGGA TGACCAGGAGACAACAACCCGCTTCCTGCAGACAGAGGCAGTGACCCTGCAG GATGTGAGGGGACGCGACTATCAGAATGCCATGCGGGTGTGGTCTAACATCC CTGGCCTGAAGAGCAAGCACGCCCCCCTGACCCCTAAGGAGGAGGAGTACCT GCAGGCCCAGGTGCGGAGCAGATCCAAGCTGGATGCCCCTAAGGTGGACCTG CTGGTGAAGAATTGTCTGCTGCCACTGCGGGAGTACTTCAAGTACTTTAGTCA GAATAGCCTGCCACTGGGTGGACCTAGCTCTGGAGCCCCGCCTCCGTCAGGT GGATCTCCTGCAGGTAGCCCAACTTCCACGGAAGAGGGCACGTCTGAGAGTG CGACTCCTGAGAGCGGGCCTGGCACAAGTACTGAGCCCAGCGAAGGTTCTGC ACCCGGGTCTCCAGCTGGGAGCCCTACCTCAACAGAGGAAGGTACCAGCACA GAGCCTTCTGAGGGTAGCGCTCCTGGCACGTCCACCGAACCGTCCGAGGGGA TGGCACCGAAGAAAAAACGCAAAGTTGGTCGAGCGGCCATGGCGGAACGAC CCTTTCAGTGCCGAATTTGCATGAGGAACTTTTCACGATCTGATAACCTGGCG AGGCACATCCGAACACATACGGGCGAGAAGCCATTCGCATGTGATATCTGCG	3xFLAG- DNMT3AL- XTEN80-SV40 NLS-ZFP-SV40 NLS-KRAB (DNA 서열)

[1285]

	GGCGAAAGTTCGCCCAAAATGTCAGTAGACCGCGACATACTAAAATACACAC TGGCTCACAGAAGCCGTTCCAATGCCGCATCTGTATGCGCAATTTTTCCCGAA GCGACGATCTGTCTAAACATATTCGGACGCACACTGGGGAAAAGCCTTTCGC TTGTGACATCTGTGGGAGGAAGTTCGCTGACAGCTCTCATAGGACACGCCAT ACTAAGATTCATACCGGAAGCCAGAAGCCTTTCCAGTGTCGGATTTGCATGA GAAACTTTAGCACTTCTAGCAACAGAAAGACACATATACGAACCCATACGGG TGAGAAACCGTTCGCATGCGATATCTGTGGGCGAAAATTTGCAGCCCAATGG ACCAGAGCTTGCCATACCAAGATACACCTTCGGCAGAAGGACGCTGCACGGG CTCCGGACCGAAAAAAAAGAGAAAGGTAAACGGCGGGGGAGGATCACGGAC CCTGGTTACATTTAAGGACGTTTTTGTAGACTTCACGAGGGAGGAATGGAAG CTGTTGGACACCGCCCAGCAAATCGTGTATCGAAATGTAATGTTGGAAAATT ATAAAAATCTGGTTTCCCTCGGCTACCAATTGACAAAACCAGACGTCATCCTC AGACTGGAAAAAGGGGAAGAACCTTGGCTTGTCTGA	
988	ATGCACGATCAGGAGTTTGACCCCCCTAAGGTGTACCCACCCGTGCCAGCCG AGAAGAGGAAGCCCATCCGCGTGCTGTCCCTGTTCGACGGCATCGCCACAGG CCTGCTGGTGCTGAAGGATCTGGGCATCCAGGTGGACAGATATATCGCCTCC GAGGTGTGCGAGGATTCTATCACCGTGGGCATGGTGAGGCACCAGGGCAAG ATCATGTACGTGGGCGACGTGCGCAGCGTGACACAGAAGCACATCCAGGAGT GGGGACCCTTCGACCTGGTCATCGGAGGCAGCCCCTGTAATGACCTGTCCAT CGTGAACCCTGCAAGGAAGGGCCTGTATGAGGGAACCGGCAGACTGTTCTTT GAGTTCTACAGGCTGCTGCACGACGCCCGCCCTAAGGAGGGCGATGACAGGC CATTCTTTTGGCTGTTTGAGAACGTGGTGGCCATGGGCGTGAGCGACAAGCG GGATATCTCCAGATTCCTGGAGTCTAATCCCGTGATGATCGATGCAAAGGAG GTGTCTGCCGCACACAGGGCAAGGTACTTTTGGGGAAATCTGCCTGGCATGA ACCGCCCACTGGCCAGCACCGTGAACGACAAGCTGGAGCTGCAGGAGTGCCT GGAGCACGGAAGGATCGCCAAGTTCTCCAAGGTGCGGACAATCACCACAAG ATCTAACAGCATCAAGCAGGGCAAGGATCAGCACTTCCCCGTGTTCATGAAT GAGAAGGAGGACATCCTGTGGTGTACCGAGATGGAGCGCGTGTTCGGCTTTC CAGTGCACTATACAGACGTGAGCAATATGAGCCGGCTGGCAAGGCAGAGAC TGCTGGGCCGGTCCTGGTCTGTGCCAGTGATCAGACACCTGTTCGCCCCCCTG AAGGAGTACTTTGCCTGCGTGTCTAGCGGCAACTCTAATGCCAACAGCAGAG GCCCTTCCTTTTCCTCTGGCCTGGTGCCACTGTCTCTGAGGGGCAGCCACATG GGCCCCATGGAGATCTACAAGACCGTGTCCGCCTGGAAGAGGCAGCCTGTGC GCGTGCTGTCTCTGTTCCGCAACATCGACAAGGTGCTGAAGAGCCTGGGCTTT CTGGAGAGCGGATCCGGATCTGGAGGAGGCACCCTGAAGTATGTGGAGGAT GTGACAAATGTGGTGCGGAGAGATGTGGAGAAGTGGGGCCCCTTCGATCTGG TGTACGGATCCACCCAGCCACTGGGAAGCTCCTGCGATAGGTGTCCAGGATG GTATATGTTCCAGTTTCACAGAATCCTGCAGTACGCACTGCCAAGGCAGGAG AGCCAGCGCCCTTTCTTTTGGATCTTTATGGACAACCTGCTGCTGACAGAGGA TGACCAGGAGACAACAACCCGCTTCCTGCAGACAGAGGGCAGTGACCCTGCAG GATGTGAGGGGACGCGACTATCAGAATGCCATGCGGGTGTGGTCTAACATCC CTGGCCTGAAGAGCAAGCACGCCCCCCTGACCCCTAAGGAGGAGGAGTACCT GCAGGCCCAGGTGCGGAGCAGATCCAAGCTGGATGCCCCTAAGGTGGACCTG CTGGTGAAGAATTGTCTGCTGCCACTGCGGGAGTACTTCAAGTACTTTAGTCA GAATAGCCTGCCACTGGGTGGACCTAGCTCTGGAGCCCCGCCTCCGTCAGGT GGATCTCCTGCAGGTAGCCCAACTTCCACGGAAGAGGGCACGTCTGAGAGTG CGACTCCTGAGAGCGGGCCTGGCACAAGTACTGAGCCCAGCGAAGGTTCTGC ACCCGGGTCTCCAGCTGGGAGCCCTACCTCAACAGAGGAAGGTACCAGCACA GAGCCTTCTGAGGGTAGCGCTCCTGGCACGTCCACCGAACCGTCCGAGGGGA TGGCACCGAAGAAAAAACGCAAAGTTGGTCGAGCTGCGATGGCAGAACGAC CTTTTCAATGCCGAATTTGTATGAGGAACTTTTCCCGGTCAGACGACCTTTCC AAGCACATCAGAACTCATACCGGAGAAAAACCGTTCGCCTGTGACATTTGTG GACGGAAGTTTGCTGACTCCTCTCACAGGACTCGCCACACTAAGATACACAC CGGAAGTCAGAAGCCCTTCCAATGTAGGATATGCATGAGAAACTTCAGTACG TCATCAAACCGAAAAACGCATATCAGGACACATACCGGCGAAAAGCCGTTTG CATGTGATATCTGCGGCAGGAAATTTGCAGCTCAGTGGACACGGGCATGTCA CACAAAAATCCATACCGGTAGTCAAAAACCGTTTCAGTGTCGAATCTGCATG AGGAACTTTAGCCGGAAGCAGACGAGAACCACGCATATAAGAACTCACACA GGTGAGAAACCCTTTGCGTGCGATATCTGCGGTCGCAAATTTGCTCACCGATC	3xFLAG- DNMT3AL- XTEN80-SV40 NLS-ZFP-SV40 NLS-KRAB (DNA 서열)

[1286]

	CTCCCTGAGGCGACATACTAAAATACATCTGCGACAGAAAGACGCGGCTCGG GCTCCGGACCGAAAAAAAAGAGAAAGGTAAACGGCGGGGGAGGATCACGG ACCCTGGTTACATTTAAGGACGTTTTTGTAGACTTCACGAGGGAGGAATGGA AGCTGTTGGACACCGCCCAGCAAATCGTGTATCGAAATGTAATGTTGGAAAA TTATAAAAATCTGGTTTCCCTCGGCTACCAATTGACAAAACCAGACGTCATCC TCAGACTGGAAAAAGGGGAAGAACCTTGGCTTGTCTGA	
989	ATGCACGATCAGGAGTTTGACCCCCCTAAGGTGTACCCACCCGTGCCAGCCG AGAAGAGGAAGCCCATCCGCGTGCTGTCCCTGTTCGACGGCATCGCCACAGG CCTGCTGGTGCTGAAGGATCTGGGCATCCAGGTGGACAGATATATCGCCTCC GAGGTGTGCGAGGATTCTATCACCGTGGGCATGGTGAGGCACCAGGGCAAG ATCATGTACGTGGGCGACGTGCGCAGCGTGACACAGAAGCACATCCAGGAGT GGGGACCCTTCGACCTGGTCATCGGAGGCAGCCCCTGTAATGACCTGTCCAT CGTGAACCCTGCAAGGAAGGGCCTGTATGAGGGAACCGGCAGACTGTTCTTT GAGTTCTACAGGCTGCTGCACGACGCCCGCCCTAAGGAGGGCGATGACAGGC CATTCTTTTGGCTGTTTGAGAACGTGGTGGCCATGGGCGTGAGCGACAAGCG GGATATCTCCAGATTCCTGGAGTCTAATCCCGTGATGATCGATGCAAAGGAG GTGTCTGCCGCACACAGGGCAAGGTACTTTTGGGGAAATCTGCCTGGCATGA ACCGCCCACTGGCCAGCACCGTGAACGACAAGCTGGAGCTGCAGGAGTGCCT GGAGCACGGAAGGATCGCCAAGTTCTCCAAGGTGCGGACAATCACCACAAG ATCTAACAGCATCAAGCAGGGCAAGGATCAGCACTTCCCCGTGTTCATGAAT GAGAAGGAGGACATCCTGTGGTGTACCGAGATGGAGCGCGTGTTCGGCTTTC CAGTGCACTATACAGACGTGAGCAATATGAGCCGGCTGGCAAGGCAGAGAC TGCTGGGCCGGTCCTGGTCTGTGCCAGTGATCAGACACCTGTTCGCCCCCCTG AAGGAGTACTTTGCCTGCGTGTCTAGCGGCAACTCTAATGCCAACAGCAGAG GCCCTTCCTTTTCCTCTGGCCTGGTGCCACTGTCTCTGAGGGGCAGCCACATG GGCCCCATGGAGATCTACAAGACCGTGTCCGCCTGGAAGAGGCAGCCTGTGC GCGTGCTGTCTCTGTTCCGCAACATCGACAAGGTGCTGAAGAGCCTGGGCTTT CTGGAGAGCGGATCCGGATCTGGAGGAGGCACCCTGAAGTATGTGGAGGAT GTGACAAATGTGGTGCGGAGAGATGTGGAGAAGTGGGGCCCCTTCGATCTGG TGTACGGATCCACCCAGCCACTGGGAAGCTCCTGCGATAGGTGTCCAGGATG GTATATGTTCCAGTTTCACAGAATCCTGCAGTACGCACTGCCAAGGCAGGAG AGCCAGCGCCCTTTCTTTTGGATCTTTATGGACAACCTGCTGCTGACAGAGGA TGACCAGGAGACAACAACCCGCTTCCTGCAGACAGAGGCAGTGACCCTGCAG GATGTGAGGGGACGCGACTATCAGAATGCCATGCGGGTGTGGTCTAACATCC CTGGCCTGAAGAGCAAGCACGCCCCCCTGACCCCTAAGGAGGAGGAGTACCT GCAGGCCCAGGTGCGGAGCAGATCCAAGCTGGATGCCCCTAAGGTGGACCTG CTGGTGAAGAATTGTCTGCTGCCACTGCGGGAGTACTTCAAGTACTTTAGTCA GAATAGCCTGCCACTGGGTGGACCTAGCTCTGGAGCCCCGCCTCCGTCAGGT GGATCTCCTGCAGGTAGCCCAACTTCCACGGAAGAGGGCACGTCTGAGAGTG CGACTCCTGAGAGCGGGCCTGGCACAAGTACTGAGCCCAGCGAAGGTTCTGC ACCCGGGTCTCCAGCTGGGAGCCCTACCTCAACAGAGGAAGGTACCAGCACA GAGCCTTCTGAGGGTAGCGCTCCTGGCACGTCCACCGAACCGTCCGAGGGGA TGGCACCGAAGAAAAAACGCAAAGTTGGTCGAGCCGCCATGGCAGAACGGC CTTTTCAGTGTCGGATCTGCATGAGAAACTTTAGTCAGAGTGCCCATCGCAAG AATCATATTCGAACTCATACCGGTGAAAAACCGTTCGCGTGCGACATCTGTG GTCGAAAGTTTGCCACATCATCCAATAGAAAAACGCATACTAAGATTCATAC CGGTAGCCAGAAACCATTCCAATGTAGAATCTGCATGCGAAATTTCAGCAGG AGTGACAATTTGTCCGCACATATACGGACACACACGGGCGAAAAACCCTTTG CTTGCGATATATGCGGTAGAAAGTTCGCTAGGAACAACGACCGAAAAACACA CACGAAGATTCATACAGGTAGTCAGAAGCCATTTCAATGTCGGATCTGTATG CGAAATTTCTCTACTTCTGGCAGCCTGTCCCGGCACATCAGAACACATACCGG TGAGAAGCCATTTGCATGTGACATATGTGGGAGAAAATTTGCCCAGGCGGGT CACCTTGCGAAGCATACAAAGATCCACCTCCGCCAGAAGGACGCCGCACGGG CTCCGGACCGAAAAAAAAGAGAAAGGTAAACGGCGGGGGAGGATCACGGAC CCTGGTTACATTTAAGGACGTTTTTGTAGACTTCACGAGGGAGGAATGGAAG CTGTTGGACACCGCCCAGCAAATCGTGTATCGAAATGTAATGTTGGAAAATT ATAAAAATCTGGTTTCCCTCGGCTACCAATTGACAAAACCAGACGTCATCCTC AGACTGGAAAAAGGGGAAGAACCTTGGCTTGTCTGA	3xFLAG- DNMT3AL- XTEN80-SV40 NLS-ZFP-SV40 NLS-KRAB (DNA 서열)

[1287]

990	ATGCACGATCAGGAGTTTGACCCCCCTAAGGTGTACCCACCCGTGCCAGCCG AGAAGAGGAAGCCCATCCGCGTGCTGTCCCTGTTCGACGGCATCGCCACAGG CCTGCTGGTGCTGAAGGATCTGGGCATCCAGGTGGACAGATATATCGCCTCC GAGGTGTGCGAGGATTCTATCACCGTGGGCATGGTGAGGCACCAGGGCAAG ATCATGTACGTGGGCGACGTGCGCAGCGTGACACAGAAGCACATCCAGGAGT GGGGACCCTTCGACCTGGTCATCGGAGGCAGCCCCTGTAATGACCTGTCCAT CGTGAACCCTGCAAGGAAGGGCCTGTATGAGGGAACCGGCAGACTGTTCTTT GAGTTCTACAGGCTGCTGCACGACGCCCGCCCTAAGGAGGGCGATGACAGGC CATTCTTTTGGCTGTTTGAGAACGTGGTGGCCATGGGCGTGAGCGACAAGCG GGATATCTCCAGATTCCTGGAGTCTAATCCCGTGATGATCGATGCAAAGGAG GTGTCTGCCGCACACAGGGCAAGGTACTTTTGGGGAAATCTGCCTGGCATGA ACCGCCCACTGGCCAGCACCGTGAACGACAAGCTGGAGCTGCAGGAGTGCCT GGAGCACGGAAGGATCGCCAAGTTCTCCAAGGTGCGGACAATCACCACAAG ATCTAACAGCATCAAGCAGGGCAAGGATCAGCACTTCCCCGTGTTCATGAAT GAGAAGGAGGACATCCTGTGGTGTACCGAGATGGAGCGCGTGTTCGGCTTTC CAGTGCACTATACAGACGTGAGCAATATGAGCCGGCTGGCAAGGCAGAGAC TGCTGGGCCGGTCCTGGTCTGTGCCAGTGATCAGACACCTGTTCGCCCCCCTG AAGGAGTACTTTGCCTGCGTGTCTAGCGGCAACTCTAATGCCAACAGCAGAG GCCCTTCCTTTTCCTCTGGCCTGGTGCCACTGTCTCTGAGGGGCAGCCACATG GGCCCCATGGAGATCTACAAGACCGTGTCCGCCTGGAAGAGGCAGCCTGTGC GCGTGCTGTCTCTGTTCCGCAACATCGACAAGGTGCTGAAGAGCCTGGGCTTT CTGGAGAGCGGATCCGGATCTGGAGGAGGCACCCTGAAGTATGTGGAGGAT GTGACAAATGTGGTGCGGAGAGATGTGGAGAAGTGGGGCCCCTTCGATCTGG TGTACGGATCCACCCAGCCACTGGGAAGCTCCTGCGATAGGTGTCCAGGATG GTATATGTTCCAGTTTCACAGAATCCTGCAGTACGCACTGCCAAGGCAGGAG AGCCAGCGCCCTTTCTTTTGGATCTTTATGGACAACCTGCTGCTGACAGAGGA TGACCAGGAGACAACAACCCGCTTCCTGCAGACAGAGGCAGTGACCCTGCAG GATGTGAGGGGACGCGACTATCAGAATGCCATGCGGGTGTGGTCTAACATCC CTGGCCTGAAGAGCAAGCACGCCCCCCTGACCCCTAAGGAGGAGGAGTACCT GCAGGCCCAGGTGCGGAGCAGATCCAAGCTGGATGCCCCTAAGGTGGACCTG CTGGTGAAGAATTGTCTGCTGCCACTGCGGGAGTACTTCAAGTACTTTAGTCA GAATAGCCTGCCACTGGGTGGACCTAGCTCTGGAGCCCCGCCTCCGTCAGGT GGATCTCCTGCAGGTAGCCCAACTTCCACGGAAGAGGGCACGTCTGAGAGTG CGACTCCTGAGAGCGGGCCTGGCACAAGTACTGAGCCCAGCGAAGGTTCTGC ACCCGGGTCTCCAGCTGGGAGCCCTACCTCAACAGAGGAAGGTACCAGCACA GAGCCTTCTGAGGGTAGCGCTCCTGGCACGTCCACCGAACCGTCCGAGGGGA TGGCACCGAAGAAAAAACGCAAAGTTGGTCGAGCGGCTATGGCAGAGCGGC CATTCCAATGTAGGATATGTATGAGGAACTTCTCCCGGAGCGATCACCTGTCC CAACACATCCGCACGCATACGGGTGAGAAGCCGTTTGCTTGTGATATCTGCG GAAGAAAATTTGCAGCATCCAGTACACGCACAAAGCATACGAAGATTCATAC GGGATCCCAAAAGCCCTTTCAATGCAGGATTTGTATGAGGAACTTCAGTCGG TCCGACGATCTGACACGACATATTAGAACTCATACTGGAGAGAAGCCATTCG CATGTGACATCTGCGGTAGGAAGTTCGCGCAGAAATCTAACCTGTCATCTCA CACCAAGATACATACAGGCTCACAGAAGCCGTTTCAATGCCGCATCTGCATG AGGAATTTCAGCCAGTCCGCAAACAGAACTACGCATATTCGGACGCATACCG GCGAGAAGCCGTTTGCCTGCGACATTTGCGGGAGGAAATTCGCACAGAACGC GACCAGAACCAAACACACCAAAATCCATCTTAGGCAAAAGGATGCGGCCCG GGCTCCGGACCGAAAAAAAAGAGAAAGGTAAACGGCGGGGGAGGATCACG GACCCTGGTTACATTTAAGGACGTTTTTGTAGACTTCACGAGGGAGGAATGG AAGCTGTTGGACACCGCCCAGCAAATCGTGTATCGAAATGTAATGTTGGAAA ATTATAAAAATCTGGTTTCCCTCGGCTACCAATTGACAAAACCAGACGTCATC CTCAGACTGGAAAAAGGGGAAGAACCTTGGCTTGTCTGA	3xFLAG- DNMT3AL- XTEN80-SV40 NLS-ZFP-SV40 NLS-KRAB (DNA 서열)
991	ATGCACGATCAGGAGTTTGACCCCCCTAAGGTGTACCCACCCGTGCCAGCCG AGAAGAGGAAGCCCATCCGCGTGCTGTCCCTGTTCGACGGCATCGCCACAGG CCTGCTGGTGCTGAAGGATCTGGGCATCCAGGTGGACAGATATATCGCCTCC GAGGTGTGCGAGGATTCTATCACCGTGGGCATGGTGAGGCACCAGGGCAAG ATCATGTACGTGGGCGACGTGCGCAGCGTGACACAGAAGCACATCCAGGAGT GGGGACCCTTCGACCTGGTCATCGGAGGCAGCCCCTGTAATGACCTGTCCAT CGTGAACCCTGCAAGGAAGGGCCTGTATGAGGGAACCGGCAGACTGTTCTTT	3xFLAG- DNMT3AL- XTEN80-SV40 NLS-ZFP-SV40 NLS-KRAB (DNA 서열)

[1288]

	GAGTTCTACAGGCTGCTGCACGACGCCCGCCCTAAGGAGGGCGATGACAGGC CATTCTTTTGGCTGTTTGAGAACGTGGTGGCCATGGGCGTGAGCGACAAGCG GGATATCTCCAGATTCCTGGAGTCTAATCCCGTGATGATCGATGCAAAGGAG GTGTCTGCCGCACACAGGGCAAGGTACTTTTGGGGAAATCTGCCTGGCATGA ACCGCCCACTGGCCAGCACCGTGAACGACAAGCTGGAGCTGCAGGAGTGCCT GGAGCACGGAAGGATCGCCAAGTTCTCCAAGGTGCGGACAATCACCACAAG ATCTAACAGCATCAAGCAGGGCAAGGATCAGCACTTCCCCGTGTTCATGAAT GAGAAGGAGGACATCCTGTGGTGTACCGAGATGGAGCGCGTGTTCGGCTTTC CAGTGCACTATACAGACGTGAGCAATATGAGCCGGCTGGCAAGGCAGAGAC TGCTGGGCCGGTCCTGGTCTGTGCCAGTGATCAGACACCTGTTCGCCCCCCTG AAGGAGTACTTTGCCTGCGTGTCTAGCGGCAACTCTAATGCCAACAGCAGAG GCCCTTCCTTTTCCTCTGGCCTGGTGCCACTGTCTCTGAGGGGCAGCCACATG GGCCCCATGGAGATCTACAAGACCGTGTCCGCCTGGAAGAGGCAGCCTGTGC GCGTGCTGTCTCTGTTCCGCAACATCGACAAGGTGCTGAAGAGCCTGGGCTTT CTGGAGAGCGGATCCGGATCTGGAGGAGGCACCCTGAAGTATGTGGAGGAT GTGACAAATGTGGTGCGGAGAGATGTGGAGAAGTGGGGCCCCTTCGATCTGG TGTACGGATCCACCCAGCCACTGGGAAGCTCCTGCGATAGGTGTCCAGGATG GTATATGTTCCAGTTTCACAGAATCCTGCAGTACGCACTGCCAAGGCAGGAG AGCCAGCGCCCTTTCTTTTGGATCTTTATGGACAACCTGCTGCTGACAGAGGA TGACCAGGAGACAACAACCCGCTTCCTGCAGACAGAGGCAGTGACCCTGCAG GATGTGAGGGGACGCGACTATCAGAATGCCATGCGGGTGTGGTCTAACATCC CTGGCCTGAAGAGCAAGCACGCCCCCCTGACCCCTAAGGAGGAGGAGTACCT GCAGGCCCAGGTGCGGAGCAGATCCAAGCTGGATGCCCCTAAGGTGGACCTG CTGGTGAAGAATTGTCTGCTGCCACTGCGGGAGTACTTCAAGTACTTTAGTCA GAATAGCCTGCCACTGGGTGGACCTAGCTCTGGAGCCCCGCCTCCGTCAGGT GGATCTCCTGCAGGTAGCCCAACTTCCACGGAAGAGGGCACGTCTGAGAGTG CGACTCCTGAGAGCGGGCCTGGCACAAGTACTGAGCCCAGCGAAGGTTCTGC ACCCGGGTCTCCAGCTGGGAGCCCTACCTCAACAGAGGAAGGTACCAGCACA GAGCCTTCTGAGGGTAGCGCTCCTGGCACGTCCACCGAACCGTCCGAGGGGA TGGCACCGAAGAAAAAACGCAAAGTTGGTCGAGCGGCCATGGCAGAACGAC CCTTTCAGTGCCGAATTTGCATGCGGAACTTTAGTCGCAGTGACACCCTGAGC GAGCATATTCGCACGCATACGGGAGAGAAGCCATTTGCATGCGACATCTGCG GTAGAAAGTTTGCGAGGCGCTGGACGTTGGTAGGCCACACGAAAATCCATAC AGGCTCCCAGAAACCCTTCCAGTGCAGAATTTGTATGCGCAATTTTAGTGAC AGAAGTAACTTGTCCCGACATATAAGGACGCACACCGGCGAAAAACCGTTTG CCTGTGATATCTGTGGTCGGAAATTCGCCCAGTCCGGTGACTTGACACGGCAT ACCAAAATACACACTGGAAGCCAAAAGCCTTTTCAGTGTCGCATATGTATGC GCAACTTCAGCCAGAGTAGTGACCTTTCACGGCATATACGGACGCATACGGG TGAGAAACCCTTCGCCTGTGACATTTGCGGGCGAAAGTTTGCATATCATTGGT ACCTGAAAAAACATACGAAAATACATTTGAGACAAAAAGATGCAGCCCGGG CTCCGGACCGAAAAAAAAGAGAAAGGTAAACGGCGGGGGAGGATCACGGAC CCTGGTTACATTTAAGGACGTTTTTGTAGACTTCACGAGGGAGGAATGGAAG CTGTTGGACACCGCCCAGCAAATCGTGTATCGAAATGTAATGTTGGAAAATT ATAAAAATCTGGTTTCCCTCGGCTACCAATTGACAAAACCAGACGTCATCCTC AGACTGGAAAAAGGGGAAGAACCTTGGCTTGTCTGA	
992	ATGCACGATCAGGAGTTTGACCCCCCTAAGGTGTACCCACCCGTGCCAGCCG AGAAGAGGAAGCCCATCCGCGTGCTGTCCCTGTTCGACGGCATCGCCACAGG CCTGCTGGTGCTGAAGGATCTGGGCATCCAGGTGGACAGATATATCGCCTCC GAGGTGTGCGAGGATTCTATCACCGTGGGCATGGTGAGGCACCAGGGCAAG ATCATGTACGTGGGCGACGTGCGCAGCGTGACACAGAAGCACATCCAGGAGT GGGGACCCTTCGACCTGGTCATCGGAGGCAGCCCCTGTAATGACCTGTCCAT CGTGAACCCTGCAAGGAAGGGCCTGTATGAGGGAACCGGCAGACTGTTCTTT GAGTTCTACAGGCTGCTGCACGACGCCCGCCCTAAGGAGGGCGATGACAGGC CATTCTTTTGGCTGTTTGAGAACGTGGTGGCCATGGGCGTGAGCGACAAGCG GGATATCTCCAGATTCCTGGAGTCTAATCCCGTGATGATCGATGCAAAGGAG GTGTCTGCCGCACACAGGGCAAGGTACTTTTGGGGAAATCTGCCTGGCATGA ACCGCCCACTGGCCAGCACCGTGAACGACAAGCTGGAGCTGCAGGAGTGCCT GGAGCACGGAAGGATCGCCAAGTTCTCCAAGGTGCGGACAATCACCACAAG ATCTAACAGCATCAAGCAGGGCAAGGATCAGCACTTCCCCGTGTTCATGAAT	3xFLAG- DNMT3AL- XTEN80-SV40 NLS-ZFP-SV40 NLS-KRAB (DNA 서열)

[1289]

	GAGAAGGAGGACATCCTGTGGTGTACCGAGATGGAGCGCGTGTTCGGCTTTC CAGTGCACTATACAGACGTGAGCAATATGAGCCGGCTGGCAAGGCAGAGAC TGCTGGGCCGGTCCTGGTCTGTGCCAGTGATCAGACACCTGTTCGCCCCCCTG AAGGAGTACTTTGCCTGCGTGTCTAGCGGCAACTCTAATGCCAACAGCAGAG GCCCTTCCTTTTCCTCTGGCCTGGTGCCACTGTCTCTGAGGGGCAGCCACATG GGCCCCATGGAGATCTACAAGACCGTGTCCGCCTGGAAGAGGCAGCCTGTGC GCGTGCTGTCTCTGTTCCGCAACATCGACAAGGTGCTGAAGAGCCTGGGCTTT CTGGAGAGCGGATCCGGATCTGGAGGAGGCACCCTGAAGTATGTGGAGGAT GTGACAAATGTGGTGCGGAGAGATGTGGAGAAGTGGGGCCCCTTCGATCTGG TGTACGGATCCACCCAGCCACTGGGAAGCTCCTGCGATAGGTGTCCAGGATG GTATATGTTCCAGTTTCACAGAATCCTGCAGTACGCACTGCCAAGGCAGGAG AGCCAGCGCCCTTTCTTTTGGATCTTTATGGACAACCTGCTGCTGACAGAGGA TGACCAGGAGACAACAACCCGCTTCCTGCAGACAGAGGCAGTGACCCTGCAG GATGTGAGGGGACGCGACTATCAGAATGCCATGCGGGTGTGGTCTAACATCC CTGGCCTGAAGAGCAAGCACGCCCCCCTGACCCCTAAGGAGGAGGAGTACCT GCAGGCCCAGGTGCGGAGCAGATCCAAGCTGGATGCCCCTAAGGTGGACCTG CTGGTGAAGAATTGTCTGCTGCCACTGCGGGAGTACTTCAAGTACTTTAGTCA GAATAGCCTGCCACTGGGTGGACCTAGCTCTGGAGCCCCGCCTCCGTCAGGT GGATCTCCTGCAGGTAGCCCAACTTCCACGGAAGAGGGCACGTCTGAGAGTG CGACTCCTGAGAGCGGGCCTGGCACAAGTACTGAGCCCAGCGAAGGTTCTGC ACCCGGGTCTCCAGCTGGGAGCCCTACCTCAACAGAGGAAGGTACCAGCACA GAGCCTTCTGAGGGTAGCGCTCCTGGCACGTCCACCGAACCGTCCGAGGGGA TGGCACCGAAGAAAAAACGCAAAGTTGGTCGAGCTGCCATGGCCGAGCGCC CGTTTCAATGCAGAATATGTATGAGGAACTTCTCTAGAAGCGCCAATCTTGC GCGACACATTAGAACTCACACAGGCGAGAAACCTTTCGCCTGTGACATCTGC GGCAGAAAATTTGCGCGAAGTGATAACTTGCGCGAGCACACAAAAATCCATA CCGGTTCCCAAAAGCCCTTTCAATGTAGGATTTGCATGAGGAACTTTTCAAGA CCATACACGCTGAGACTCCACATTCGCACGCATACGGGAGAAAAACCATTTG CTTGTGACATATGCGGCCGAAAGTTTGCACACCGATCCAACTTGAATAAGCA TACCAAAATCCATACGGGGTCTCAGAAACCCTTCCAGTGTCGGATTTGTATGC GAAACTTCTCTCAGAGTGGTTCTCTTACGCGCCACATTAGAACCCACACGGG GGAAAAGCCATTCGCGTGCGATATCTGTGGCCGGAAATTTGCTACTTCCGCA AATCTTTCTCGACATACAAAGATACATCTTAGACAGAAGGATGCGGCACGGG CTCCGGACCGAAAAAAAAAGAGAAAGGTAAACGGCGGGGGAGGATCACGGAC CCTGGTTACATTTAAGGACGTTTTTGTAGACTTCACGAGGGAGGAATGGAAG CTGTTGGACACCGCCCAGCAAATCGTGTATCGAAATGTAATGTTGGAAAATT ATAAAAATCTGGTTTCCCTCGGCTACCAATTGACAAAACCAGACGTCATCCTC AGACTGGAAAAAGGGGAAGAACCTTGGCTTGTCTGA	
993	ATGCACGATCAGGAGTTTGACCCCCCTAAGGTGTACCCACCCGTGCCAGCCG AGAAGAGGAAGCCCATCCGCGTGCTGTCCCTGTTCGACGGCATCGCCACAGG CCTGCTGGTGCTGAAGGATCTGGGCATCCAGGTGGACAGATATATCGCCTCC GAGGTGTGCGAGGATTCTATCACCGTGGGCATGGTGAGGCACCAGGGCAAG ATCATGTACGTGGGCGACGTGCGCAGCGTGACACAGAAGCACATCCAGGAGT GGGGACCCTTCGACCTGGTCATCGGAGGCAGCCCCTGTAATGACCTGTCCAT CGTGAACCCTGCAAGGAAGGGCCTGTATGAGGGAACCGGCAGACTGTTCTTT GAGTTCTACAGGCTGCTGCACGACGCCCGCCCTAAGGAGGGCGATGACAGGC CATTCTTTTGGCTGTTTGAGAACGTGGTGGCCATGGGCGTGAGCGACAAGCG GGATATCTCCAGATTCCTGGAGTCTAATCCCGTGATGATCGATGCAAAGGAG GTGTCTGCCGCACACAGGGCAAGGTACTTTTGGGGAAATCTGCCTGGCATGA ACCGCCCACTGGCCAGCACCGTGAACGACAAGCTGGAGCTGCAGGAGTGCCT GGAGCACGGAAGGATCGCCAAGTTCTCCAAGGTGCGGACAATCACCACAAG ATCTAACAGCATCAAGCAGGGCAAGGATCAGCACTTCCCCGTGTTCATGAAT GAGAAGGAGGACATCCTGTGGTGTACCGAGATGGAGCGCGTGTTCGGCTTTC CAGTGCACTATACAGACGTGAGCAATATGAGCCGGCTGGCAAGGCAGAGAC TGCTGGGCCGGTCCTGGTCTGTGCCAGTGATCAGACACCTGTTCGCCCCCCTG AAGGAGTACTTTGCCTGCGTGTCTAGCGGCAACTCTAATGCCAACAGCAGAG GCCCTTCCTTTTCCTCTGGCCTGGTGCCACTGTCTCTGAGGGGCAGCCACATG GGCCCCATGGAGATCTACAAGACCGTGTCCGCCTGGAAGAGGCAGCCTGTGC GCGTGCTGTCTCTGTTCCGCAACATCGACAAGGTGCTGAAGAGCCTGGGCTTT	3xFLAG- DNMT3AL- XTEN80-SV40 NLS-ZFP-SV40 NLS-KRAB (DNA 서열)

[1290]

	CTGGAGAGCGGATCCGGATCTGGAGGAGGCACCCTGAAGTATGTGGAGGAT GTGACAAATGTGGTGCGGAGAGATGTGGAGAAGTGGGGCCCCTTCGATCTGG TGTACGGATCCACCCAGCCACTGGGAAGCTCCTGCGATAGGTGTCCAGGATG GTATATGTTCCAGTTTCACAGAATCCTGCAGTACGCACTGCCAAGGCAGGAG AGCCAGCGCCCTTTCTTTTGGATCTTTATGGACAACCTGCTGCTGACAGAGGA TGACCAGGAGACAACAACCCGCTTCCTGCAGACAGAGGCAGTGACCCTGCAG GATGTGAGGGGACGCGACTATCAGAATGCCATGCGGGTGTGGTCTAACATCC CTGGCCTGAAGAGCAAGCACGCCCCCCTGACCCCTAAGGAGGAGGAGTACCT GCAGGCCCAGGTGCGGAGCAGATCCAAGCTGGATGCCCCTAAGGTGGACCTG CTGGTGAAGAATTGTCTGCTGCCACTGCGGGAGTACTTCAAGTACTTTAGTCA GAATAGCCTGCCACTGGGTGGACCTAGCTCTGGAGCCCCGCCTCCGTCAGGT GGATCTCCTGCAGGTAGCCCAACTTCCACGGAAGAGGGCACGTCTGAGAGTG CGACTCCTGAGAGCGGGCCTGGCACAAGTACTGAGCCCAGCGAAGGTTCTGC ACCCGGGTCTCCAGCTGGGAGCCCTACCTCAACAGAGGAAGGTACCAGCACA GAGCCTTCTGAGGGTAGCGCTCCTGGCACGTCCACCGAACCGTCCGAGGGGA TGGCACCGAAGAAAAAACGCAAAGTTGGTCGATTGGAGCCCGGTGAAAAAC CATATGCCTGTCCAGAGTGTGGAAAAAGTTTTAGCAGAAGCGACGACCTGGT TAGGCATCAACGAACTCACACAGGGGAAAAGCCGTACAAATGTCCTGAATGC GGAAAGTCATTCTCCACTTCAGGTAGCCTCGTGCGACATCAGCGCACACATA CTGGCGAAAAACCTTATAAGTGTCCGGAATGCGGCAAGAGTTTTTCTAGAAG TGATAAGCTCGTGCGACACCAGCGGACGCACACGGGTGAGAAGCCCTACGCC TGCCCAGAGTGCGGAAAGTCCTTTAGTAGGTCTGACGAACTGGTCCGGCACC AACGAACCCATACAGGGGAGAAACCCTACAAATGTCCAGAGTGCGGGAAAT CATTTTCAACGTCTCACAGCTTGACGGAACATCAGAGGACCCATACAGGCGA AAAGCCTTACAAATGTCCGGAGTGCGGAAAGTCTTTCAGCCGGGCTGATAAT CTCACGGAGCACCAAAGAACCCACACGGGTAAGAAAACCTCTGCTCCGGACC GAAAAAAAAGAGAAAGGTAAACGGCGGGGGAGGATCACGGACCCTGGTTAC ATTTAAGGACGTTTTTGTAGACTTCACGAGGGAGGAATGGAAGCTGTTGGAC ACCGCCCAGCAAATCGTGTATCGAAATGTAATGTTGGAAAATTATAAAAATC TGGTTTCCCTCGGCTACCAATTGACAAAACCAGACGTCATCCTCAGACTGGA AAAAGGGGAAGAACCTTGGCTTGTCTGA	
994	ATGCACGATCAGGAGTTTGACCCCCCTAAGGTGTACCCACCCGTGCCAGCCG AGAAGAGGAAGCCCATCCGCGTGCTGTCCCTGTTCGACGGCATCGCCACAGG CCTGCTGGTGCTGAAGGATCTGGGCATCCAGGTGGACAGATATATCGCCTCC GAGGTGTGCGAGGATTCTATCACCGTGGGCATGGTGAGGCACCAGGGCAAG ATCATGTACGTGGGCGACGTGCGCAGCGTGACACAGAAGCACATCCAGGAGT GGGGACCCTTCGACCTGGTCATCGGAGGCAGCCCCTGTAATGACCTGTCCAT CGTGAACCCTGCAAGGAAGGGCCTGTATGAGGGAACCGGCAGACTGTTCTTT GAGTTCTACAGGCTGCTGCACGACGCCCGCCCTAAGGAGGGCGATGACAGGC CATTCTTTTGGCTGTTTGAGAACGTGGTGGCCATGGGCGTGAGCGACAAGCG GGATATCTCCAGATTCCTGGAGTCTAATCCCGTGATGATCGATGCAAAGGAG GTGTCTGCCGCACACAGGGCAAGGTACTTTTGGGGAAATCTGCCTGGCATGA ACCGCCCACTGGCCAGCACCGTGAACGACAAGCTGGAGCTGCAGGAGTGCCT GGAGCACGGAAGGATCGCCAAGTTCTCCAAGGTGCGGACAATCACCACAAG ATCTAACAGCATCAAGCAGGGCAAGGATCAGCACTTCCCCGTGTTCATGAAT GAGAAGGAGGACATCCTGTGGTGTACCGAGATGGAGCGCGTGTTCGGCTTTC CAGTGCACTATACAGACGTGAGCAATATGAGCCGGCTGGCAAGGCAGAGAC TGCTGGGCCGGTCCTGGTCTGTGCCAGTGATCAGACACCTGTTCGCCCCCCTG AAGGAGTACTTTGCCTGCGTGTCTAGCGGCAACTCTAATGCCAACAGCAGAG GCCCTTCCTTTTCCTCTGGCCTGGTGCCACTGTCTCTGAGGGGCAGCCACATG GGCCCCATGGAGATCTACAAGACCGTGTCCGCCTGGAAGAGGCAGCCTGTGC GCGTGCTGTCTCTGTTCCGCAACATCGACAAGGTGCTGAAGAGCCTGGGCTTT CTGGAGAGCGGATCCGGATCTGGAGGAGGCACCCTGAAGTATGTGGAGGAT GTGACAAATGTGGTGCGGAGAGATGTGGAGAAGTGGGGCCCCTTCGATCTGG TGTACGGATCCACCCAGCCACTGGGAAGCTCCTGCGATAGGTGTCCAGGATG GTATATGTTCCAGTTTCACAGAATCCTGCAGTACGCACTGCCAAGGCAGGAG AGCCAGCGCCCTTTCTTTTGGATCTTTATGGACAACCTGCTGCTGACAGAGGA TGACCAGGAGACAACAACCCGCTTCCTGCAGACAGAGGCAGTGACCCTGCAG GATGTGAGGGGACGCGACTATCAGAATGCCATGCGGGTGTGGTCTAACATCC	3xFLAG- DNMT3AL- XTEN80-SV40 NLS-ZFP-SV40 NLS-KRAB (DNA 서열)

[1291]

	CTGGCCTGAAGAGCAAGCACGCCCCCCTGACCCCTAAGGAGGAGGAGTACCT GCAGGCCCAGGTGCGGAGCAGATCCAAGCTGGATGCCCCTAAGGTGGACCTG CTGGTGAAGAATTGTCTGCTGCCACTGCGGGAGTACTTCAAGTACTTTAGTCA GAATAGCCTGCCACTGGGTGGACCTAGCTCTGGAGCCCCGCCTCCGTCAGGT GGATCTCCTGCAGGTAGCCCAACTTCCACGGAAGAGGGCACGTCTGAGAGTG CGACTCCTGAGAGCGGGCCTGGCACAAGTACTGAGCCCAGCGAAGGTTCTGC ACCCGGGTCTCCAGCTGGGAGCCCTACCTCAACAGAGGAAGGTACCAGCACA GAGCCTTCTGAGGGTAGCGCTCCTGGCACGTCCACCGAACCGTCCGAGGGGA TGGCACCGAAGAAAAAACGCAAAGTTGGTCGATTGGAGCCGGGAGAAAAGC CATATGCCTGTCCGGAATGCGGCAAAAGCTTTTCAGAACGGTCCCATTTGCG GGAACATCAGCGCACCCATACCGGAGAAAAGCCTTACAAGTGTCCCGAGTGT GGTAAGAGCTTTTCAACTTCCCACAGTCTCACTGAACATCAGCGAACTCACA CAGGCGAAAAACCATACAAATGCCCGGAGTGTGGAAAGAGTTTTTCTCAAGC TGGGCACTTGGCCAGCCACCAAAGGACTCATACGGGCGAGAAACCGTATGCC TGTCCAGAATGCGGGAAATCATTTTCTACTAGTCATTCCCTTACGGAACATCA GAGAACGCACACCGGGGAGAAGCCATACAAGTGTCCGGAGTGCGGAAAATC ATTCTCCGACCCGGGTCATCTCGTTCGCCATCAGAGGACTCATACTGGAGAG AAACCGTATAAATGTCCTGAATGCGGGAAGAGCTTTTCTACATCTGGGAACC TTGTGCGGCACCAGCGAACCCACACGGGGAAAAAGACTTCTGCTCCGGACCG AAAAAAAAGAGAAAGGTAAACGGCGGGGGAGGATCACGGACCCTGGTTACA TTTAAGGACGTTTTTGTAGACTTCACGAGGGAGGAATGGAAGCTGTTGGACA CCGCCCAGCAAATCGTGTATCGAAATGTAATGTTGGAAAATTATAAAAATCT GGTTTCCCTCGGCTACCAATTGACAAAACCAGACGTCATCCTCAGACTGGAA AAAGGGGAAGAACCTTGGCTTGTCTGA	
995	ATGCACGATCAGGAGTTTGACCCCCCTAAGGTGTACCCACCCGTGCCAGCCG AGAAGAGGAAGCCCATCCGCGTGCTGTCCCTGTTCGACGGCATCGCCACAGG CCTGCTGGTGCTGAAGGATCTGGGCATCCAGGTGGACAGATATATCGCCTCC GAGGTGTGCGAGGATTCTATCACCGTGGGCATGGTGAGGCACCAGGGCAAG ATCATGTACGTGGGCGACGTGCGCAGCGTGACACAGAAGCACATCCAGGAGT GGGGACCCTTCGACCTGGTCATCGGAGGCAGCCCCTGTAATGACCTGTCCAT CGTGAACCCTGCAAGGAAGGGCCTGTATGAGGGAACCGGCAGACTGTTCTTT GAGTTCTACAGGCTGCTGCACGACGCCCGCCCTAAGGAGGGCGATGACAGGC CATTCTTTTGGCTGTTTGAGAACGTGGTGGCCATGGGCGTGAGCGACAAGCG GGATATCTCCAGATTCCTGGAGTCTAATCCCGTGATGATCGATGCAAAGGAG GTGTCTGCCGCACACAGGGCAAGGTACTTTTGGGGAAATCTGCCTGGCATGA ACCGCCCACTGGCCAGCACCGTGAACGACAAGCTGGAGCTGCAGGAGTGCCT GGAGCACGGAAGGATCGCCAAGTTCTCCAAGGTGCGGACAATCACCACAAG ATCTAACAGCATCAAGCAGGGCAAGGATCAGCACTTCCCCGTGTTCATGAAT GAGAAGGAGGACATCCTGTGGTGTACCGAGATGGAGCGCGTGTTCGGCTTTC CAGTGCACTATACAGACGTGAGCAATATGAGCCGGCTGGCAAGGCAGAGAC TGCTGGGCCGGTCCTGGTCTGTGCCAGTGATCAGACACCTGTTCGCCCCCCTG AAGGAGTACTTTGCCTGCGTGTCTAGCGGCAACTCTAATGCCAACAGCAGAG GCCCTTCCTTTTCCTCTGGCCTGGTGCCACTGTCTCTGAGGGGCAGCCACATG GGCCCCATGGAGATCTACAAGACCGTGTCCGCCTGGAAGAGGCAGCCTGTGC GCGTGCTGTCTCTGTTCCGCAACATCGACAAGGTGCTGAAGAGCCTGGGCTTT CTGGAGAGCGGATCCGGATCTGGAGGAGGCACCCTGAAGTATGTGGAGGAT GTGACAAATGTGGTGCGGAGAGATGTGGAGAAGTGGGGCCCCTTCGATCTGG TGTACGGATCCACCCAGCCACTGGGAAGCTCCTGCGATAGGTGTCCAGGATG GTATATGTTCCAGTTTCACAGAATCCTGCAGTACGCACTGCCAAGGCAGGAG AGCCAGCGCCCTTTCTTTTGGATCTTTATGGACAACCTGCTGCTGACAGAGGA TGACCAGGAGACAACAACCCGCTTCCTGCAGACAGAGGCAGTGACCCTGCAG GATGTGAGGGGACGCGACTATCAGAATGCCATGCGGGTGTGGTCTAACATCC CTGGCCTGAAGAGCAAGCACGCCCCCCTGACCCCTAAGGAGGAGGAGTACCT GCAGGCCCAGGTGCGGAGCAGATCCAAGCTGGATGCCCCTAAGGTGGACCTG CTGGTGAAGAATTGTCTGCTGCCACTGCGGGAGTACTTCAAGTACTTTAGTCA GAATAGCCTGCCACTGGGTGGACCTAGCTCTGGAGCCCCGCCTCCGTCAGGT GGATCTCCTGCAGGTAGCCCAACTTCCACGGAAGAGGGCACGTCTGAGAGTG CGACTCCTGAGAGCGGGCCTGGCACAAGTACTGAGCCCAGCGAAGGTTCTGC ACCCGGGTCTCCAGCTGGGAGCCCTACCTCAACAGAGGAAGGTACCAGCACA	3xFLAG- DNMT3AL- XTEN80-SV40 NLS-ZFP-SV40 NLS-KRAB (DNA 서열)

[1292]

	GAGCCTTCTGAGGGTAGCGCTCCTGGCACGTCCACCGAACCGTCCGAGGGGA TGGCACCGAAGAAAAAACGCAAAGTTGGTCGACTGGAACCGGGTGAAAAAC CCTACGCTTGTCCCGAGTGTGGGAAATCATTCTCAAGGGCAGATAATCTTAC GGAACATCAGCGCACACACACCGGGGAAAAACCCTATAAGTGCCCTGAGTGT GGTAAATCATTCTCTACGTCAGGGTCATTGGTGCGCCATCAGAGGACCCATA CAGGCGAAAAACCGTATAAGTGCCCCGAATGCGGTAAAAGCTTCAGCAGAA AAGATAACTTGAAAAATCACCAACGCACTCACACGGGAGAAAAACCGTACG CATGCCCGGAGTGCGGCAAGAGTTTCAGTCAGAGTAGCTCACTTGTTCGCCA CCAAAGAACCCACACAGGCGAAAAGCCCTACAAGTGTCCTGAATGTGGAAA GAGTTTCAGTCGAAGTGATAAATTGGTTAGGCACCAAAGAACACACACTGGA GAGAAACCGTACAAGTGTCCAGAATGTGGCAAGTCTTTTTCTGATTCTGGAA ATTTGCGCGTCCATCAAAGGACTCATACTGGGAAAAAGACGTCCGCTCCGGA CCGAAAAAAAAGAGAAAGGTAAACGGCGGGGGAGGATCACGGACCCTGGTT ACATTTAAGGACGTTTTTGTAGACTTCACGAGGGAGGAATGGAAGCTGTTGG ACACCGCCCAGCAAATCGTGTATCGAAATGTAATGTTGGAAAATTATAAAAA TCTGGTTTCCCTCGGCTACCAATTGACAAAACCAGACGTCATCCTCAGACTGG AAAAAGGGGAAGAACCTTGGCTTGTCTGA	
996	ATGCACGATCAGGAGTTTGACCCCCCTAAGGTGTACCCACCCGTGCCAGCCG AGAAGAGGAAGCCCATCCGCGTGCTGTCCCTGTTCGACGGCATCGCCACAGG CCTGCTGGTGCTGAAGGATCTGGGCATCCAGGTGGACAGATATATCGCCTCC GAGGTGTGCGAGGATTCTATCACCGTGGGCATGGTGAGGCACCAGGGCAAG ATCATGTACGTGGGCGACGTGCGCAGCGTGACACAGAAGCACATCCAGGAGT GGGGACCCTTCGACCTGGTCATCGGAGGCAGCCCCTGTAATGACCTGTCCAT CGTGAACCCTGCAAGGAAGGGCCTGTATGAGGGAACCGGCAGACTGTTCTTT GAGTTCTACAGGCTGCTGCACGACGCCCGCCCTAAGGAGGGCGATGACAGGC CATTCTTTTGGCTGTTTGAGAACGTGGTGGCCATGGGCGTGAGCGACAAGCG GGATATCTCCAGATTCCTGGAGTCTAATCCCGTGATGATCGATGCAAAGGAG GTGTCTGCCGCACACAGGGCAAGGTACTTTTGGGGAAATCTGCCTGGCATGA ACCGCCCACTGGCCAGCACCGTGAACGACAAGCTGGAGCTGCAGGAGTGCCT GGAGCACGGAAGGATCGCCAAGTTCTCCAAGGTGCGGACAATCACCACAAG ATCTAACAGCATCAAGCAGGGCAAGGATCAGCACTTCCCCGTGTTCATGAAT GAGAAGGAGGACATCCTGTGGTGTACCGAGATGGAGCGCGTGTTCGGCTTTC CAGTGCACTATACAGACGTGAGCAATATGAGCCGGCTGGCAAGGCAGAGAC TGCTGGGCCGGTCCTGGTCTGTGCCAGTGATCAGACACCTGTTCGCCCCCCTG AAGGAGTACTTTGCCTGCGTGTCTAGCGGCAACTCTAATGCCAACAGCAGAG GCCCTTCCTTTTCCTCTGGCCTGGTGCCACTGTCTCTGAGGGGCAGCCACATG GGCCCCATGGAGATCTACAAGACCGTGTCCGCCTGGAAGAGGCAGCCTGTGC GCGTGCTGTCTCTGTTCCGCAACATCGACAAGGTGCTGAAGAGCCTGGGCTTT CTGGAGAGCGGATCCGGATCTGGAGGAGGCACCCTGAAGTATGTGGAGGAT GTGACAAATGTGGTGCGGAGAGATGTGGAGAAGTGGGGCCCCTTCGATCTGG TGTACGGATCCACCCAGCCACTGGGAAGCTCCTGCGATAGGTGTCCAGGATG GTATATGTTCCAGTTTCACAGAATCCTGCAGTACGCACTGCCAAGGCAGGAG AGCCAGCGCCCTTTCTTTTGGATCTTTATGGACAACCTGCTGCTGACAGAGGA TGACCAGGAGACAACAACCCGCTTCCTGCAGACAGAGGCAGTGACCCTGCAG GATGTGAGGGGACGCGACTATCAGAATGCCATGCGGGTGTGGTCTAACATCC CTGGCCTGAAGAGCAAGCACGCCCCCCTGACCCCTAAGGAGGAGGAGTACCT GCAGGCCCAGGTGCGGAGCAGATCCAAGCTGGATGCCCCTAAGGTGGACCTG CTGGTGAAGAATTGTCTGCTGCCACTGCGGGAGTACTTCAAGTACTTTAGTCA GAATAGCCTGCCACTGGGTGGACCTAGCTCTGGAGCCCCGCCTCCGTCAGGT GGATCTCCTGCAGGTAGCCCAACTTCCACGGAAGAGGGCACGTCTGAGAGTG CGACTCCTGAGAGCGGGCCTGGCACAAGTACTGAGCCCAGCGAAGGTTCTGC ACCCGGGTCTCCAGCTGGGAGCCCTACCTCAACAGAGGAAGGTACCAGCACA GAGCCTTCTGAGGGTAGCGCTCCTGGCACGTCCACCGAACCGTCCGAGGGGA TGGCACCGAAGAAAAAACGCAAAGTTGGTCGACTCGAACCAGGGGAAAAGC CATACGCTTGCCCCGAGTGTGGAAAATCTTTCTCTCAGTCAAGCTCCCTTGTC AGACACCAGAGAACCCATACGGGTGAGAAACCATACAAGTGCCCAGAGTGC GGCAAAAGCTTCAGTCAGAGTGGCGATCTGCGCCGGCATCAAAGAACTCACA CTGGAGAAAAGCCCTATAAGTGCCCCGAATGTGGAAAGAGTTTTTCCAGGAG TGACGAAAGAAAGAGACACCAGCGGACTCACACCGGCGAGAAACCATACGC	3xFLAG- DNMT3AL- XTEN80-SV40 NLS-ZFP-SV40 NLS-KRAB (DNA 서열)

[1293]

	ATGCCCCGAGTGCGGAAAAAGTTTTTCCCACCGCACAACCCTTACCAATCAT CAACGCACGCACACAGGGGAGAAACCCTACAAGTGCCCGGAGTGTGGCAAG TCATTCAGCCGAAGTGACCACCTCACTAACCACCAAAGGACTCACACAGGAG AAAAACCCTACAAATGTCCGGAGTGCGGAAAATCTTTTTCCACGTCCGGTGA GCTGGTCCGCCATCAACGAACCCATACTGGTAAAAAAACTAGCGCTCCGGAC CGAAAAAAAAGAGAAAGGTAAACGGCGGGGGAGGATCACGGACCCTGGTTA CATTTAAGGACGTTTTTGTAGACTTCACGAGGGAGGAATGGAAGCTGTTGGA CACCGCCCAGCAAATCGTGTATCGAAATGTAATGTTGGAAAATTATAAAAAT CTGGTTTCCCTCGGCTACCAATTGACAAAACCAGACGTCATCCTCAGACTGG AAAAAGGGGAAGAACCTTGGCTTGTCTGA	
997	ATGCACGATCAGGAGTTTGACCCCCCTAAGGTGTACCCACCCGTGCCAGCCG AGAAGAGGAAGCCCATCCGCGTGCTGTCCCTGTTCGACGGCATCGCCACAGG CCTGCTGGTGCTGAAGGATCTGGGCATCCAGGTGGACAGATATATCGCCTCC GAGGTGTGCGAGGATTCTATCACCGTGGGCATGGTGAGGCACCAGGGCAAG ATCATGTACGTGGGCGACGTGCGCAGCGTGACACAGAAGCACATCCAGGAGT GGGGACCCTTCGACCTGGTCATCGGAGGCAGCCCCTGTAATGACCTGTCCAT CGTGAACCCTGCAAGGAAGGGCCTGTATGAGGGAACCGGCAGACTGTTCTTT GAGTTCTACAGGCTGCTGCACGACGCCCGCCCTAAGGAGGGCGATGACAGGC CATTCTTTTGGCTGTTTGAGAACGTGGTGGCCATGGGCGTGAGCGACAAGCG GGATATCTCCAGATTCCTGGAGTCTAATCCCGTGATGATCGATGCAAAGGAG GTGTCTGCCGCACACAGGGCAAGGTACTTTTGGGGAAATCTGCCTGGCATGA ACCGCCCACTGGCCAGCACCGTGAACGACAAGCTGGAGCTGCAGGAGTGCCT GGAGCACGGAAGGATCGCCAAGTTCTCCAAGGTGCGGACAATCACCACAAG ATCTAACAGCATCAAGCAGGGCAAGGATCAGCACTTCCCCGTGTTCATGAAT GAGAAGGAGGACATCCTGTGGTGTACCGAGATGGAGCGCGTGTTCGGCTTTC CAGTGCACTATACAGACGTGAGCAATATGAGCCGGCTGGCAAGGCAGAGAC TGCTGGGCCGGTCCTGGTCTGTGCCAGTGATCAGACACCTGTTCGCCCCCCTG AAGGAGTACTTTGCCTGCGTGTCTAGCGGCAACTCTAATGCCAACAGCAGAG GCCCTTCCTTTTCCTCTGGCCTGGTGCCACTGTCTCTGAGGGGCAGCCACATG GGCCCCATGGAGATCTACAAGACCGTGTCCGCCTGGAAGAGGCAGCCTGTGC GCGTGCTGTCTCTGTTCCGCAACATCGACAAGGTGCTGAAGAGCCTGGGCTTT CTGGAGAGCGGATCCGGATCTGGAGGAGGCACCCTGAAGTATGTGGAGGAT GTGACAAATGTGGTGCGGAGAGATGTGGAGAAGTGGGGCCCCTTCGATCTGG TGTACGGATCCACCCAGCCACTGGGAAGCTCCTGCGATAGGTGTCCAGGATG GTATATGTTCCAGTTTCACAGAATCCTGCAGTACGCACTGCCAAGGCAGGAG AGCCAGCGCCCTTTCTTTTGGATCTTTATGGACAACCTGCTGCTGACAGAGGA TGACCAGGAGACAACAACCCGCTTCCTGCAGACAGAGGCAGTGACCCTGCAG GATGTGAGGGGACGCGACTATCAGAATGCCATGCGGGTGTGGTCTAACATCC CTGGCCTGAAGAGCAAGCACGCCCCCCTGACCCCTAAGGAGGAGGAGTACCT GCAGGCCCAGGTGCGGAGCAGATCCAAGCTGGATGCCCCTAAGGTGGACCTG CTGGTGAAGAATTGTCTGCTGCCACTGCGGGAGTACTTCAAGTACTTTAGTCA GAATAGCCTGCCACTGGGTGGACCTAGCTCTGGAGCCCCGCCTCCGTCAGGT GGATCTCCTGCAGGTAGCCCAACTTCCACGGAAGAGGGCACGTCTGAGAGTG CGACTCCTGAGAGCGGGCCTGGCACAAGTACTGAGCCCAGCGAAGGTTCTGC ACCCGGGTCTCCAGCTGGGAGCCCTACCTCAACAGAGGAAGGTACCAGCACA GAGCCTTCTGAGGGTAGCGCTCCTGGCACGTCCACCGAACCGTCCGAGGGGA TGGCACCGAAGAAAAAACGCAAAGTTGGTCGACTGGAACCCGGGGAGAAGC CGTACGCTTGCCCAGAGTGCGGAAAGAGCTTCTCTCAGAGCGGAGACCTTAG ACGCCACCAGCGAACCCACACCGGCGAAAAACCGTATAAATGCCCGGAATG CGGCAAGAGTTTTAGTCGGTCCGATGAGCGAAAGAGGCATCAACGAACCCAT ACGGGAGAGAAACCCTACAAGTGCCCTGAGTGTGGTAAGTCATTTTCCCACA GAACGACGTTGACGAATCACCAGAGAACCCATACGGGTGAGAAACCTTACG CTTGCCCGGAGTGCGGCAAAAGCTTCAGCCGGAGTGATCACTTGACCAATCA TCAGAGGACACACACGGGTGAGAAGCCCTACAAATGTCCCGAATGCGGCAA GTCTTTCTCAACGTCAGGCGAACTCGTCCGGCACCAGCGAACACATACGGGA GAAAAGCCGTACAAATGTCCGGAATGCGGAAAGTCATTTTCACGGTCAGATG ACTTGGTGCGACACCAGCGCACTCACACAGGCAAGAAGACCTCAGCTCCGGA CCGAAAAAAAAGAGAAAGGTAAACGGCGGGGGAGGATCACGGACCCTGGTT ACATTTAAGGACGTTTTTGTAGACTTCACGAGGGAGGAATGGAAGCTGTTGG	3xFLAG- DNMT3AL- XTEN80-SV40 NLS-ZFP-SV40 NLS-KRAB (DNA 서열)

[1294]

[1295]

	ACACCGCCCAGCAAATCGTGTATCGAAATGTAATGTTGGAAAATTATAAAAA TCTGGTTTCCCTCGGCTACCAATTGACAAAACCAGACGTCATCCTCAGACTGG AAAAAGGGGAAGAACCTTGGCTTGTCTGA	
998	ATGCACGATCAGGAGTTTGACCCCCCTAAGGTGTACCCACCCGTGCCAGCCG AGAAGAGGAAGCCCATCCGCGTGCTGTCCCTGTTCGACGGCATCGCCACAGG CCTGCTGGTGCTGAAGGATCTGGGCATCCAGGTGGACAGATATATCGCCTCC GAGGTGTGCGAGGATTCTATCACCGTGGGCATGGTGAGGCACCAGGGCAAG ATCATGTACGTGGGCGACGTGCGCAGCGTGACACAGAAGCACATCCAGGAGT GGGGACCCTTCGACCTGGTCATCGGAGGCAGCCCCTGTAATGACCTGTCCAT CGTGAACCCTGCAAGGAAGGGCCTGTATGAGGGAACCGGCAGACTGTTCTTT GAGTTCTACAGGCTGCTGCACGACGCCCGCCCTAAGGAGGGCGATGACAGGC CATTCTTTTGGCTGTTTGAGAACGTGGTGGCCATGGGCGTGAGCGACAAGCG GGATATCTCCAGATTCCTGGAGTCTAATCCCGTGATGATCGATGCAAAGGAG GTGTCTGCCGCACACAGGGCAAGGTACTTTTGGGGAAATCTGCCTGGCATGA ACCGCCCACTGGCCAGCACCGTGAACGACAAGCTGGAGCTGCAGGAGTGCCT GGAGCACGGAAGGATCGCCAAGTTCTCCAAGGTGCGGACAATCACCACAAG ATCTAACAGCATCAAGCAGGGCAAGGATCAGCACTTCCCCGTGTTCATGAAT GAGAAGGAGGACATCCTGTGGTGTACCGAGATGGAGCGCGTGTTCGGCTTTC CAGTGCACTATACAGACGTGAGCAATATGAGCCGGCTGGCAAGGCAGAGAC TGCTGGGCCGGTCCTGGTCTGTGCCAGTGATCAGACACCTGTTCGCCCCCCTG AAGGAGTACTTTGCCTGCGTGTCTAGCGGCAACTCTAATGCCAACAGCAGAG GCCCTTCCTTTTCCTCTGGCCTGGTGCCACTGTCTCTGAGGGGCAGCCACATG GGCCCCATGGAGATCTACAAGACCGTGTCCGCCTGGAAGAGGCAGCCTGTGC GCGTGCTGTCTCTGTTCCGCAACATCGACAAGGTGCTGAAGAGCCTGGGCTTT CTGGAGAGCGGATCCGGATCTGGAGGAGGCACCCTGAAGTATGTGGAGGAT GTGACAAATGTGGTGCGGAGAGATGTGGAGAAGTGGGGCCCCTTCGATCTGG TGTACGGATCCACCCAGCCACTGGGAAGCTCCTGCGATAGGTGTCCAGGATG GTATATGTTCCAGTTTCACAGAATCCTGCAGTACGCACTGCCAAGGCAGGAG AGCCAGCGCCCTTTCTTTTGGATCTTTATGGACAACCTGCTGCTGACAGAGGA TGACCAGGAGACAACAACCCGCTTCCTGCAGACAGAGGCAGTGACCCTGCAG GATGTGAGGGGACGCGACTATCAGAATGCCATGCGGGTGTGGTCTAACATCC CTGGCCTGAAGAGCAAGCACGCCCCCCTGACCCCTAAGGAGGAGGAGTACCT GCAGGCCCAGGTGCGGAGCAGATCCAAGCTGGATGCCCCTAAGGTGGACCTG CTGGTGAAGAATTGTCTGCTGCCACTGCGGGAGTACTTCAAGTACTTTAGTCA GAATAGCCTGCCACTGGGTGGACCTAGCTCTGGAGCCCCGCCTCCGTCAGGT GGATCTCCTGCAGGTAGCCCAACTTCCACGGAAGAGGGCACGTCTGAGAGTG CGACTCCTGAGAGCGGGCCTGGCACAAGTACTGAGCCCAGCGAAGGTTCTGC ACCCGGGTCTCCAGCTGGGAGCCCTACCTCAACAGAGGAAGGTACCAGCACA GAGCCTTCTGAGGGTAGCGCTCCTGGCACGTCCACCGAACCGTCCGAGGGGA TGGCACCGAAGAAAAAACGCAAAGTTGGTCGACTGGAGCCTGGTGAGAAGC CGTATGCATGTCCTGAGTGTGGGAAGTCATTTAGTCAGAGGGCCCACTTGGA ACGACACCAAAGGACCCACACTGGTGAAAAACCCTACAAATGCCCAGAGTG TGGTAAGTCTTTTTCACAGCTGGCCCACCTGAGAGCACACCAGCGAACTCAT ACGGGCGAGAAACCATACAAGTGTCCAGAGTGCGGAAAGTCATTCTCAGATC CCGGCCACTTGGTGCGACATCAGAGAACGCACACAGGGGAGAAGCCTTATGC TTGCCCGGAATGCGGGAAGTCTTTCAGCCGCCGAAGTGCTTGTCGAAGGCAC CAACGGACCCATACCGGTGAGAAACCATATAAGTGCCCAGAGTGTGGAAAG AGTTTTAGTCGATCCGATCACCTGACTACGCACCAGCGGACGCACACAGGAG AGAAACCGTATAAGTGCCCTGAATGCGGTAAGAGCTTCTCTCAATCAAGCTC ACTGGTTAGGCACCAACGCACTCATACCGGCAAGAAGACGTCAGCTCCGGAC CGAAAAAAAAGAGAAAGGTAAACGGCGGGGGAGGATCACGGACCCTGGTTA CATTTAAGGACGTTTTTGTAGACTTCACGAGGGAGGAATGGAAGCTGTTGGA CACCGCCCAGCAAATCGTGTATCGAAATGTAATGTTGGAAAATTATAAAAAT CTGGTTTCCCTCGGCTACCAATTGACAAAACCAGACGTCATCCTCAGACTGG AAAAAGGGGAAGAACCTTGGCTTGTCTGA	3xFLAG- DNMT3AL- XTEN80-SV40 NLS-ZFP-SV40 NLS-KRAB (DNA 서열)
999	ATGCACGATCAGGAGTTTGACCCCCCTAAGGTGTACCCACCCGTGCCAGCCG AGAAGAGGAAGCCCATCCGCGTGCTGTCCCTGTTCGACGGCATCGCCACAGG CCTGCTGGTGCTGAAGGATCTGGGCATCCAGGTGGACAGATATATCGCCTCC GAGGTGTGCGAGGATTCTATCACCGTGGGCATGGTGAGGCACCAGGGCAAG	3xFLAG- DNMT3AL- XTEN80-SV40 NLS-ZFP-SV40

	ATCATGTACGTGGGCGACGTGCGCAGCGTGACACAGAAGCACATCCAGGAGT GGGGACCCTTCGACCTGGTCATCGGAGGCAGCCCCTGTAATGACCTGTCCAT CGTGAACCCTGCAAGGAAGGGCCTGTATGAGGGAACCGGCAGACTGTTCTTT GAGTTCTACAGGCTGCTGCACGACGCCCGCCCTAAGGAGGGCGATGACAGGC CATTCTTTTGGCTGTTTGAGAACGTGGTGGCCATGGGCGTGAGCGACAAGCG GGATATCTCCAGATTCCTGGAGTCTAATCCCGTGATGATCGATGCAAAGGAG GTGTCTGCCGCACACAGGGCAAGGTACTTTTGGGGAAATCTGCCTGGCATGA ACCGCCCACTGGCCAGCACCGTGAACGACAAGCTGGAGCTGCAGGAGTGCCT GGAGCACGGAAGGATCGCCAAGTTCTCCAAGGTGCGGACAATCACCACAAG ATCTAACAGCATCAAGCAGGGCAAGGATCAGCACTTCCCCGTGTTCATGAAT GAGAAGGAGGACATCCTGTGGTGTACCGAGATGGAGCGCGTGTTCGGCTTTC CAGTGCACTATACAGACGTGAGCAATATGAGCCGGCTGGCAAGGCAGAGAC TGCTGGGCCGGTCCTGGTCTGTGCCAGTGATCAGACACCTGTTCGCCCCCCTG AAGGAGTACTTTGCCTGCGTGTCTAGCGGCAACTCTAATGCCAACAGCAGAG GCCCTTCCTTTTCCTCTGGCCTGGTGCCACTGTCTCTGAGGGGCAGCCACATG GGCCCCATGGAGATCTACAAGACCGTGTCCGCCTGGAAGAGGCAGCCTGTGC GCGTGCTGTCTCTGTTCCGCAACATCGACAAGGTGCTGAAGAGCCTGGGCTTT CTGGAGAGCGGATCCGGATCTGGAGGAGGCACCCTGAAGTATGTGGAGGAT GTGACAAATGTGGTGCGGAGAGATGTGGAGAAGTGGGGCCCCTTCGATCTGG TGTACGGATCCACCCAGCCACTGGGAAGCTCCTGCGATAGGTGTCCAGGATG GTATATGTTCCAGTTTCACAGAATCCTGCAGTACGCACTGCCAAGGCAGGAG AGCCAGCGCCCTTTCTTTTGGATCTTTATGGACAACCTGCTGCTGACAGAGGA TGACCAGGAGACAACAACCCGCTTCCTGCAGACAGAGGCAGTGACCCTGCAG GATGTGAGGGGACGCGACTATCAGAATGCCATGCGGGTGTGGTCTAACATCC CTGGCCTGAAGAGCAAGCACGCCCCCCTGACCCCTAAGGAGGAGGAGTACCT GCAGGCCCAGGTGCGGAGCAGATCCAAGCTGGATGCCCCTAAGGTGGACCTG CTGGTGAAGAATTGTCTGCTGCCACTGCGGGAGTACTTCAAGTACTTTAGTCA GAATAGCCTGCCACTGGGTGGACCTAGCTCTGGAGCCCCGCCTCCGTCAGGT GGATCTCCTGCAGGTAGCCCAACTTCCACGGAAGAGGGCACGTCTGAGAGTG CGACTCCTGAGAGCGGGCCTGGCACAAGTACTGAGCCCAGCGAAGGTTCTGC ACCCGGGTCTCCAGCTGGGAGCCCTACCTCAACAGAGGAAGGTACCAGCACA GAGCCTTCTGAGGGTAGCGCTCCTGGCACGTCCACCGAACCGTCCGAGGGGA TGGCACCGAAGAAAAAACGCAAAGTTGGTCGACTTGAGCCGGGGGAAAAGC CTTATGCTTGCCCAGAGTGTGGCAAGAGCTTCTCCCAAAGTTCAAACCTCGTC CGACACCAAAGGACTCACACGGGCGAAAAACCGTATAAATGCCCCGAGTGC GGAAAGTCATTTTCCAGGTCCGACGATCTGGTCCGCCACCAGCGCACTCATA CGGGGGAAAAGCCCTATAAGTGCCCTGAGTGCGGCAAGAGCTTCTCAACTCA CCTGGATCTCATTCGCCATCAACGCACACATACAGGGGAGAAGCCTTACGCT TGTCCAGAGTGCGGCAAGTCTTTCAGTACGAGCGGAAACCTGACGGAACACC AGCGAACCCACACGGGCGAGAAGCCATATAAATGTCCCGAATGTGGAAAAT CATTCTCTCGCCGATCTGCGTGCCGCCGGCATCAGAGGACACATACCGGAGA AAAGCCGTACAAATGCCCCGAGTGTGGAAAATCCTTTAGCAGAAATGATACA CTTACCGAGCATCAGAGGACGCACACTGGAAAAAAGACATCTGCTCCGGACC GAAAAAAAAGAGAAAGGTAAACGGCGGGGGAGGATCACGGACCCTGGTTAC ATTTAAGGACGTTTTTGTAGACTTCACGAGGGAGGAATGGAAGCTGTTGGAC ACCGCCCAGCAAATCGTGTATCGAAATGTAATGTTGGAAAATTATAAAAATC TGGTTTCCCTCGGCTACCAATTGACAAAACCAGACGTCATCCTCAGACTGGA AAAAGGGGAAGAACCTTGGCTTGTCTGA	NLS-KRAB (DNA 서열)
10 00	MHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCE DSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARK GLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLES NPVMIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFS KVRTITTRSNSIKQGKDQHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSR LARQRLLGRSWSVPVIRHLFAPLKEYFACVSSGNSNANSRGPSFSSGLVPLSLRGS HMGPMEIYKTVSAWKRQPVRVLSLFRNIDKVLKSLGFLESGSGSGGGTLKYVED VTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRILQYALPRQES QRPFFWIFMDNLLLTEDDQETTTRFLQTEAVTLQDVRGRDYQNAMRVWSNIPGL KSKHAPLTPKEEEYLQAQVRSRSKLDAPKVDLLVKNCLLPLREYFKYFSQNSLPL GGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTST	3xFLAG- DNMT3AL- XTEN80-SV40 NLS-ZFP-SV40 NLS-KRAB (AA 서열)

	EEGTSTEPSEGSAPGTSTEPSEGMAPKKKRKVGRVPAAMAERPFQCRICMRNFSS EADRSRHIRTHTGEKPFACDICGRKFADRSNLTRHTKIHTGSQKPFQCRICMRNFS QSSDLSRHIRTHTGEKPFACDICGRKFAYHWYLKKHTKIHTGSQKPFQCRICMRN FSRSDSLSVHIRTHTGEKPFACDICGRKFAQNANRKTHTKIHLRQKDAARGSGPK KKRKVNGGGGSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVS LGYQLTKPDVILRLEKGEEPWLV	
1001	MHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCE DSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARK GLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLES NPVMIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFS KVRTITTRSNSIKQGKDQHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSR LARQRLLGRSWSVPVIRHLFAPLKEYFACVSSGNSNANSRGPSFSSGLVPLSLRGS HMGPMEIYKTVSAWKRQPVRVLSLFRNIDKVLKSLGFLESGSGSGGGTLKYVED VTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRILQYALPRQES QRPFFWIFMDNLLLTEDDQETTTRFLQTEAVTLQDVRGRDYQNAMRVWSNIPGL KSKHAPLTPKEEEYLQAQVRSRSKLDAPKVDLLVKNCLLPLREYFKYFSQNSLPL GGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTST EEGTSTEPSEGSAPGTSTEPSEGMAPKKKRKVGRVPAAMAERPFQCRICMRNFSR SDVLSTHIRTHTGEKPFACDICGKKFADNSSRTRHTKIHTGSQKPFQCRICMRNFS RPYTLRLHIRTHTGEKPFACDICGRKFADSSHRTRHTKIHTGSQKPFQCRICMRNF SRSDHLSQHIRTHTGEKPFACDICGRKFADSSHRTRHTKIHLRQKDAARGSGPKK KRKVNGGGGSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSL GYQLTKPDVILRLEKGEEPWLV	3xFLAG- DNMT3AL- XTEN80-SV40 NLS-ZFP-SV40 NLS-KRAB (AA 서열)
1002	MHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCE DSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARK GLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLES NPVMIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFS KVRTITTRSNSIKQGKDQHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSR LARQRLLGRSWSVPVIRHLFAPLKEYFACVSSGNSNANSRGPSFSSGLVPLSLRGS HMGPMEIYKTVSAWKRQPVRVLSLFRNIDKVLKSLGFLESGSGSGGGTLKYVED VTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRILQYALPRQES QRPFFWIFMDNLLLTEDDQETTTRFLQTEAVTLQDVRGRDYQNAMRVWSNIPGL KSKHAPLTPKEEEYLQAQVRSRSKLDAPKVDLLVKNCLLPLREYFKYFSQNSLPL GGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTST EEGTSTEPSEGSAPGTSTEPSEGMAPKKKRKVGRVPAAMAERPFQCRICMRNFSR SDHLSQHIRTHTGEKPFACDICGRKFAQSADRTKHTKIHTGSQKPFQCRICMRNFS RSDHLSQHIRTHTGEKPFACDICGRKFARRSDLKRHTKIHTGSQKPFQCRICMRNF SRSDHLSRHIRTHTGEKPFACDICGRKFAQSSDLRRHTKIHLRQKDAARGSGPKK KRKVNGGGGSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSL GYQLTKPDVILRLEKGEEPWLV	3xFLAG- DNMT3AL- XTEN80-SV40 NLS-ZFP-SV40 NLS-KRAB (AA 서열)
1003	MHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCE DSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARK GLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLES NPVMIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFS KVRTITTRSNSIKQGKDQHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSR LARQRLLGRSWSVPVIRHLFAPLKEYFACVSSGNSNANSRGPSFSSGLVPLSLRGS HMGPMEIYKTVSAWKRQPVRVLSLFRNIDKVLKSLGFLESGSGSGGGTLKYVED VTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRILQYALPRQES QRPFFWIFMDNLLLTEDDQETTTRFLQTEAVTLQDVRGRDYQNAMRVWSNIPGL KSKHAPLTPKEEEYLQAQVRSRSKLDAPKVDLLVKNCLLPLREYFKYFSQNSLPL GGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTST EEGTSTEPSEGSAPGTSTEPSEGMAPKKKRKVGRVPAAMAERPFQCRICMRNFSR SDNLSEHIRTHTGEKPFACDICGRKFATSSNRKTHTKIHTGSQKPFQCRICMRNFS DRSHLTRHIRTHTGEKPFACDICGRKFARSDALTQHTKIHTGSQKPFQCRICMRNF SDRSALARHIRTHTGEKPFACDICGRKFARRFTLSKHTKIHLRQKDAARGSGPKK KRKVNGGGGSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSL GYQLTKPDVILRLEKGEEPWLV	
1004	MHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCE DSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARK	3xFLAG- DNMT3AL-

[1296]

	GLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLES NPVMIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFS KVRTITTRSNSIKQGKDQHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSR LARQRLLGRSWSVPVIRHLFAPLKEYFACVSSGNSNANSRGPSFSSGLVPLSLRGS HMGPMEIYKTVSAWKRQPVRVLSLFRNIDKVLKSLGFLESGSGSGGGTLKYVED VTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRILQYALPRQES QRPFFWIFMDNLLLTEDDQETTTRFLQTEAVTLQDVRGRDYQNAMRVWSNIPGL KSKHAPLTPKEEEYLQAQVRSRSKLDAPKVDLLVKNCLLPLREYFKYFSQNSLPL GGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTST EEGTSTEPSEGSAPGTSTEPSEGMAPKKKRKVGRVPAAMAERPFQCRICMRNFSR SDHLSEHIRTHTGEKPFACDICGRKFAQYSGRYYHTKIHTGSQKPFQCRICMRNFS HGQTLNEHIRTHTGEKPFACDICGRKFAQSGNLARHTKIHTGSQKPFQCRICMRN FSRSDSLLRHIRTHTGEKPFACDICGRKFACREYRGKHTKIHLRQKDAARGSGPK KKRKVNGGGGSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVS LGYQLTKPDVILRLEKGEEPWLV	XTEN80-SV40 NLS-ZFP-SV40 NLS-KRAB (AA 서열)
10 05	MHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCE DSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARK GLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLES NPVMIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFS KVRTITTRSNSIKQGKDQHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSR LARQRLLGRSWSVPVIRHLFAPLKEYFACVSSGNSNANSRGPSFSSGLVPLSLRGS HMGPMEIYKTVSAWKRQPVRVLSLFRNIDKVLKSLGFLESGSGSGGGTLKYVED VTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRILQYALPRQES QRPFFWIFMDNLLLTEDDQETTTRFLQTEAVTLQDVRGRDYQNAMRVWSNIPGL KSKHAPLTPKEEEYLQAQVRSRSKLDAPKVDLLVKNCLLPLREYFKYFSQNSLPL GGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTST EEGTSTEPSEGSAPGTSTEPSEGMAPKKKRKVGRVPAAMAERPFQCRICMRNFSQ SANRTTHIRTHTGEKPFACDICGRKFARSANLTRHTKIHTGSQKPFQCRICMRNFS RSDVLSEHIRTHTGEKPFACDICGRKFATSGHLSRHTKIHTGSQKPFQCRICMRNF SQSSDLSRHIRTHTGEKPFACDICGRKFAQWSTRKRHTKIHLRQKDAARGSGPKK KRKVNGGGGSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSL GYQLTKPDVILRLEKGEEPWLV	3xFLAG- DNMT3AL- XTEN80-SV40 NLS-ZFP-SV40 NLS-KRAB (AA 서열)
10 06	MHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCE DSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARK GLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLES NPVMIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFS KVRTITTRSNSIKQGKDQHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSR LARQRLLGRSWSVPVIRHLFAPLKEYFACVSSGNSNANSRGPSFSSGLVPLSLRGS HMGPMEIYKTVSAWKRQPVRVLSLFRNIDKVLKSLGFLESGSGSGGGTLKYVED VTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRILQYALPRQES QRPFFWIFMDNLLLTEDDQETTTRFLQTEAVTLQDVRGRDYQNAMRVWSNIPGL KSKHAPLTPKEEEYLQAQVRSRSKLDAPKVDLLVKNCLLPLREYFKYFSQNSLPL GGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTST EEGTSTEPSEGSAPGTSTEPSEGMAPKKKRKVGRVPAAMAERPFQCRICMRNFSQ SGNLARHIRTHTGEKPFACDICGRKFAATCCLAHHTKIHTGSQKPFQCRICMRNF SRWQYLPTHIRTHAGEKPFACDICGRKFADRSALARHTKIHTGSQKPFQCRICMR NFSRSDNLSEHIRTHTGEKPFACDICGRKFAKRCNLRCHTKIHLRQKDAARGSGP KKKRKVNGGGGSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNL VSLGYQLTKPDVILRLEKGEEPWLV	3xFLAG- DNMT3AL- XTEN80-SV40 NLS-ZFP-SV40 NLS-KRAB (AA 서열)
10 07	MHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCE DSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARK GLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLES NPVMIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFS KVRTITTRSNSIKQGKDQHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSR LARQRLLGRSWSVPVIRHLFAPLKEYFACVSSGNSNANSRGPSFSSGLVPLSLRGS HMGPMEIYKTVSAWKRQPVRVLSLFRNIDKVLKSLGFLESGSGSGGGTLKYVED VTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRILQYALPRQES QRPFFWIFMDNLLLTEDDQETTTRFLQTEAVTLQDVRGRDYQNAMRVWSNIPGL KSKHAPLTPKEEEYLQAQVRSRSKLDAPKVDLLVKNCLLPLREYFKYFSQNSLPL	3xFLAG- DNMT3AL- XTEN80-SV40 NLS-ZFP-SV40 NLS-KRAB (AA 서열)

[1297]

	GGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTST EEGTSTEPSEGSAPGTSTEPSEGMAPKKKRKVGRVPAAMAERPFQCRICMRNFSN PANLTRHIRTHTGEKPFACDICGRKFAQNATRTKHTKIHTGSQKPFQCRICMRNF SQSGHLARHIRTHTGEKPFACDICGRKFANRHDRAKHTKIHTGSQKPFQCRICMR NFSRSDHLSEHIRTHTGEKPFACDICGRKFAQRRSRYKHTKIHLRQKDAARGSGP KKKRKVNGGGGSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNL VSLGYQLTKPDVILRLEKGEEPWLV	
1008	MHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCE DSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARK GLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLES NPVMIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFS KVRTITTRSNSIKQGKDQHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSR LARQRLLGRSWSVPVIRHLFAPLKEYFACVSSGNSNANSRGPSFSSGLVPLSLRGS HMGPMEIYKTVSAWKRQPVRVLSLFRNIDKVLKSLGFLESGSGSGGGTLKYVED VTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRILQYALPRQES QRPFFWIFMDNLLLTEDDQETTTRFLQTEAVTLQDVRGRDYQNAMRVWSNIPGL KSKHAPLTPKEEEYLQAQVRSRSKLDAPKVDLLVKNCLLPLREYFKYFSQNSLPL GGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTST EEGTSTEPSEGSAPGTSTEPSEGMAPKKKRKVGRVPAAMAERPFQCRICMRNFSQ SSDLSRHIRTHTGEKPFACDICGRKFAHRSTRNRHTKIHTGSQKPFQCRICMRNFS RSDVLSAHIRTHTGEKPFACDICGRKFADSRTRKNHTKIHTGSQKPFQCRICMRN FSQSGSLTRHIRTHTGEKPFACDICGRKFADQSGLAHHTKIHLRQKDAARGSGPK KKRKVNGGGGSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVS LGYQLTKPDVILRLEKGEEPWLV	3xFLAG- DNMT3AL- XTEN80-SV40 NLS-ZFP-SV40 NLS-KRAB (AA 서열)
1009	MHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCE DSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARK GLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLES NPVMIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFS KVRTITTRSNSIKQGKDQHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSR LARQRLLGRSWSVPVIRHLFAPLKEYFACVSSGNSNANSRGPSFSSGLVPLSLRGS HMGPMEIYKTVSAWKRQPVRVLSLFRNIDKVLKSLGFLESGSGSGGGTLKYVED VTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRILQYALPRQES QRPFFWIFMDNLLLTEDDQETTTRFLQTEAVTLQDVRGRDYQNAMRVWSNIPGL KSKHAPLTPKEEEYLQAQVRSRSKLDAPKVDLLVKNCLLPLREYFKYFSQNSLPL GGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTST EEGTSTEPSEGSAPGTSTEPSEGMAPKKKRKVGRVPAAMAERPFQCRICMRNFSQ NPAQWRHIRTHTGEKPFACDICGRKFARSADLSRHTKIHTGSQKPFQCRICMRNF STSGSLSRHIRTHTGEKPFACDICGRKFARSDHLSRHTKIHTGSQKPFQCRICMRN FSRSDSLLRHIRTHTGEKPFACDICGRKFAQSYDRFQHTKIHLRQKDAARGSGPK KKRKVNGGGGSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVS LGYQLTKPDVILRLEKGEEPWLV	3xFLAG- DNMT3AL- XTEN80-SV40 NLS-ZFP-SV40 NLS-KRAB (AA 서열)
1010	MHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCE DSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARK GLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLES NPVMIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFS KVRTITTRSNSIKQGKDQHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSR LARQRLLGRSWSVPVIRHLFAPLKEYFACVSSGNSNANSRGPSFSSGLVPLSLRGS HMGPMEIYKTVSAWKRQPVRVLSLFRNIDKVLKSLGFLESGSGSGGGTLKYVED VTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRILQYALPRQES QRPFFWIFMDNLLLTEDDQETTTRFLQTEAVTLQDVRGRDYQNAMRVWSNIPGL KSKHAPLTPKEEEYLQAQVRSRSKLDAPKVDLLVKNCLLPLREYFKYFSQNSLPL GGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTST EEGTSTEPSEGSAPGTSTEPSEGMAPKKKRKVGRVPAAMAERPFQCRICMRNFST SGSLSRHIRTHTGEKPFACDICGRKFARSDHLSRHTKIHTGSQKPFQCRICMRNFS RSDSLLRHIRTHTGEKPFACDICGRKFAQSYDRFQHTKIHTGSQKPFQCRICMRNF SRSDNLSTHIRTHTGEKPFACDICGRKFADNRDRIKHTKIHLRQKDAARGSGPKK KRKVNGGGGSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSL GYQLTKPDVILRLEKGEEPWLV	3xFLAG- DNMT3AL- XTEN80-SV40 NLS-ZFP-SV40 NLS-KRAB (AA 서열)

[1298]

1011	MHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCE DSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARK GLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLES NPVMIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFS KVRTITTRSNSIKQGKDQHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSR LARQRLLGRSWSVPVIRHLFAPLKEYFACVSSGNSNANSRGPSFSSGLVPLSLRGS HMGPMEIYKTVSAWKRQPVRVLSLFRNIDKVLKSLGFLESGSGSGGGTLKYVED VTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRILQYALPRQES QRPFFWIFMDNLLLTEDDQETTTRFLQTEAVTLQDVRGRDYQNAMRVWSNIPGL KSKHAPLTPKEEEYLQAQVRSRSKLDAPKVDLLVKNCLLPLREYFKYFSQNSLPL GGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTST EEGTSTEPSEGSAPGTSTEPSEGMAPKKKRKVGRVPAAMAERPFQCRICMRNFSD RSNLSRHIRTHTGEKPFACDICGRKFALRQNLIMHTKIHTGSQKPFQCRICMRNFS ERGTLARHIRTHTGEKPFACDICGRKFARSDALTQHTKIHTGSQKPFQCRICMRN FSRSDSLSQHIRTHTGEKPFACDICGRKFARKADRTRHTKIHLRQKDAARGSGPK KKRKVNGGGGSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVS LGYQLTKPDVILRLEKGEEPWLV	3xFLAG- DNMT3AL- XTEN80-SV40 NLS-ZFP-SV40 NLS-KRAB (AA 서열)
1012	MHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCE DSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARK GLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLES NPVMIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFS KVRTITTRSNSIKQGKDQHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSR LARQRLLGRSWSVPVIRHLFAPLKEYFACVSSGNSNANSRGPSFSSGLVPLSLRGS HMGPMEIYKTVSAWKRQPVRVLSLFRNIDKVLKSLGFLESGSGSGGGTLKYVED VTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRILQYALPRQES QRPFFWIFMDNLLLTEDDQETTTRFLQTEAVTLQDVRGRDYQNAMRVWSNIPGL KSKHAPLTPKEEEYLQAQVRSRSKLDAPKVDLLVKNCLLPLREYFKYFSQNSLPL GGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTST EEGTSTEPSEGSAPGTSTEPSEGMAPKKKRKVGRVPAAMAERPFQCRICMRNFSQ YCCLTNHIRTHTGEKPFACDICGRKFATSGNLTRHTKIHTGSQKPFQCRICMRNFS QSSDLSRHIRTHTGEKPFACDICGRKFAFRYYLKRHTKIHTGSQKPFQCRICMRNF SQSGDLTRHIRTHTGEKPFACDICGRKFADKGNLTKHTKIHLRQKDAARGSGPK KKRKVNGGGGSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVS LGYQLTKPDVILRLEKGEEPWLV	3xFLAG- DNMT3AL- XTEN80-SV40 NLS-ZFP-SV40 NLS-KRAB (AA 서열)
1013	MHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCE DSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARK GLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLES NPVMIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFS KVRTITTRSNSIKQGKDQHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSR LARQRLLGRSWSVPVIRHLFAPLKEYFACVSSGNSNANSRGPSFSSGLVPLSLRGS HMGPMEIYKTVSAWKRQPVRVLSLFRNIDKVLKSLGFLESGSGSGGGTLKYVED VTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRILQYALPRQES QRPFFWIFMDNLLLTEDDQETTTRFLQTEAVTLQDVRGRDYQNAMRVWSNIPGL KSKHAPLTPKEEEYLQAQVRSRSKLDAPKVDLLVKNCLLPLREYFKYFSQNSLPL GGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTST EEGTSTEPSEGSAPGTSTEPSEGMAPKKKRKVGRVPAAMAERPFQCRICMRNFST SGSLSRHIRTHTGEKPFACDICGRKFARSDNLTTHTKIHTGSQKPFQCRICMRNFS QSGNLARHIRTHTGEKPFACDICGRKFADRTTLMRHTKIHTGSQKPFQCRICMRN FSQSGHLARHIRTHTGEKPFACDICGRKFAQLTHLNSHTKIHLRQKDAARGSGPK KKRKVNGGGGSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVS LGYQLTKPDVILRLEKGEEPWLV	3xFLAG- DNMT3AL- XTEN80-SV40 NLS-ZFP-SV40 NLS-KRAB (AA 서열)
1014	MHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCE DSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARK GLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLES NPVMIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFS KVRTITTRSNSIKQGKDQHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSR LARQRLLGRSWSVPVIRHLFAPLKEYFACVSSGNSNANSRGPSFSSGLVPLSLRGS HMGPMEIYKTVSAWKRQPVRVLSLFRNIDKVLKSLGFLESGSGSGGGTLKYVED VTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRILQYALPRQES	3xFLAG- DNMT3AL- XTEN80-SV40 NLS-ZFP-SV40 NLS-KRAB (AA 서열)

[1299]

	QRPFFWIFMDNLLLTEDDQETTTRFLQTEAVTLQDVRGRDYQNAMRVWSNIPGL KSKHAPLTPKEEEYLQAQVRSRSKLDAPKVDLLVKNCLLPLREYFKYFSQNSLPL GGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTST EEGTSTEPSEGSAPGTSTEPSEGMAPKKKRKVGRVPAAMAERPFQCRICMRNFSI KHDLHRHIRTHTGEKPFACDICGRKFARSANLTRHTKIHTGSQKPFQCRICMRNF SRSDNLARHIRTHTGEKPFACDICGRKFAQNVSRPRHTKIHTGSQKPFQCRICMR NFSRSDDLSKHIRTHTGEKPFACDICGRKFADSSHRTRHTKIHLRQKDAARGSGP KKKRKVNGGGGSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNL VSLGYQLTKPDVILRLEKGEEPWLV	
10 15	MHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCE DSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARK GLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLES NPVMIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFS KVRTITTRSNSIKQGKDQHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSR LARQRLLGRSWSVPVIRHLFAPLKEYFACVSSGNSNANSRGPSFSSGLVPLSLRGS HMGPMEIYKTVSAWKRQPVRVLSLFRNIDKVLKSLGFLESGSGSGGGTLKYVED VTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRILQYALPRQES QRPFFWIFMDNLLLTEDDQETTTRFLQTEAVTLQDVRGRDYQNAMRVWSNIPGL KSKHAPLTPKEEEYLQAQVRSRSKLDAPKVDLLVKNCLLPLREYFKYFSQNSLPL GGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTST EEGTSTEPSEGSAPGTSTEPSEGMAPKKKRKVGRVPAAMAERPFQCRICMRNFSR SDNLARHIRTHTGEKPFACDICGRKFAQNVSRPRHTKIHTGSQKPFQCRICMRNFS RSDDLSKHIRTHTGEKPFACDICGRKFADSSHRTRHTKIHTGSQKPFQCRICMRNF STSSNRKTHIRTHTGEKPFACDICGRKFAAQWTRACHTKIHLRQKDAARGSGPK KKRKVNGGGGSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVS LGYQLTKPDVILRLEKGEEPWLV	3xFLAG- DNMT3AL- XTEN80-SV40 NLS-ZFP-SV40 NLS-KRAB (AA 서열)
10 16	MHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCE DSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARK GLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLES NPVMIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFS KVRTITTRSNSIKQGKDQHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSR LARQRLLGRSWSVPVIRHLFAPLKEYFACVSSGNSNANSRGPSFSSGLVPLSLRGS HMGPMEIYKTVSAWKRQPVRVLSLFRNIDKVLKSLGFLESGSGSGGGTLKYVED VTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRILQYALPRQES QRPFFWIFMDNLLLTEDDQETTTRFLQTEAVTLQDVRGRDYQNAMRVWSNIPGL KSKHAPLTPKEEEYLQAQVRSRSKLDAPKVDLLVKNCLLPLREYFKYFSQNSLPL GGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTST EEGTSTEPSEGSAPGTSTEPSEGMAPKKKRKVGRVPAAMAERPFQCRICMRNFSR SDDLSKHIRTHTGEKPFACDICGRKFADSSHRTRHTKIHTGSQKPFQCRICMRNFS TSSNRKTHIRTHTGEKPFACDICGRKFAAQWTRACHTKIHTGSQKPFQCRICMRN FSRKQTRTTHIRTHTGEKPFACDICGRKFAHRSSLRRHTKIHLRQKDAARGSGPK KKRKVNGGGGSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVS LGYQLTKPDVILRLEKGEEPWLV	3xFLAG- DNMT3AL- XTEN80-SV40 NLS-ZFP-SV40 NLS-KRAB (AA 서열)
10 17	MHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCE DSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARK GLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLES NPVMIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFS KVRTITTRSNSIKQGKDQHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSR LARQRLLGRSWSVPVIRHLFAPLKEYFACVSSGNSNANSRGPSFSSGLVPLSLRGS HMGPMEIYKTVSAWKRQPVRVLSLFRNIDKVLKSLGFLESGSGSGGGTLKYVED VTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRILQYALPRQES QRPFFWIFMDNLLLTEDDQETTTRFLQTEAVTLQDVRGRDYQNAMRVWSNIPGL KSKHAPLTPKEEEYLQAQVRSRSKLDAPKVDLLVKNCLLPLREYFKYFSQNSLPL GGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTST EEGTSTEPSEGSAPGTSTEPSEGMAPKKKRKVGRVPAAMAERPFQCRICMRNFSQ SAHRKNHIRTHTGEKPFACDICGRKFATSSNRKTHTKIHTGSQKPFQCRICMRNFS RSDNLSAHIRTHTGEKPFACDICGRKFARNNDRKTHTKIHTGSQKPFQCRICMRN FSTSGSLSRHIRTHTGEKPFACDICGRKFAQAGHLAKHTKIHLRQKDAARGSGPK	3xFLAG- DNMT3AL- XTEN80-SV40 NLS-ZFP-SV40 NLS-KRAB (AA 서열)

[1300]

	KKRKVNGGGGSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVS LGYQLTKPDVILRLEKGEEPWLV	
1018	MHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCE DSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARK GLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLES NPVMIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFS KVRTITTRSNSIKQGKDQHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSR LARQRLLGRSWSVPVIRHLFAPLKEYFACVSSGNSNANSRGPSFSSGLVPLSLRGS HMGPMEIYKTVSAWKRQPVRVLSLFRNIDKVLKSLGFLESGSGSGGGTLKYVED VTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRILQYALPRQES QRPFFWIFMDNLLLTEDDQETTTRFLQTEAVTLQDVRGRDYQNAMRVWSNIPGL KSKHAPLTPKEEEYLQAQVRSRSKLDAPKVDLLVKNCLLPLREYFKYFSQNSLPL GGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTST EEGTSTEPSEGSAPGTSTEPSEGMAPKKKRKVGRVPAAMAERPFQCRICMRNFSR SDHLSQHIRTHTGEKPFACDICGRKFAASSTRTKHTKIHTGSQKPFQCRICMRNFS RSDDLTRHIRTHTGEKPFACDICGRKFAQKSNLSSHTKIHTGSQKPFQCRICMRNF SQSANRTTHIRTHTGEKPFACDICGRKFAQNATRTKHTKIHLRQKDAARGSGPKK KRKVNGGGGSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSL GYQLTKPDVILRLEKGEEPWLV	3xFLAG- DNMT3AL- XTEN80-SV40 NLS-ZFP-SV40 NLS-KRAB (AA 서열)
1019	MHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCE DSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARK GLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLES NPVMIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFS KVRTITTRSNSIKQGKDQHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSR LARQRLLGRSWSVPVIRHLFAPLKEYFACVSSGNSNANSRGPSFSSGLVPLSLRGS HMGPMEIYKTVSAWKRQPVRVLSLFRNIDKVLKSLGFLESGSGSGGGTLKYVED VTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRILQYALPRQES QRPFFWIFMDNLLLTEDDQETTTRFLQTEAVTLQDVRGRDYQNAMRVWSNIPGL KSKHAPLTPKEEEYLQAQVRSRSKLDAPKVDLLVKNCLLPLREYFKYFSQNSLPL GGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTST EEGTSTEPSEGSAPGTSTEPSEGMAPKKKRKVGRVPAAMAERPFQCRICMRNFSR SDTLSEHIRTHTGEKPFACDICGRKFARRWTLVGHTKIHTGSQKPFQCRICMRNFS DRSNLSRHIRTHTGEKPFACDICGRKFAQSGDLTRHTKIHTGSQKPFQCRICMRNF SQSSDLSRHIRTHTGEKPFACDICGRKFAYHWYLKKHTKIHLRQKDAARGSGPK KKRKVNGGGGSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVS LGYQLTKPDVILRLEKGEEPWLV	3xFLAG- DNMT3AL- XTEN80-SV40 NLS-ZFP-SV40 NLS-KRAB (AA 서열)
1020	MHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCE DSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARK GLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLES NPVMIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFS KVRTITTRSNSIKQGKDQHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSR LARQRLLGRSWSVPVIRHLFAPLKEYFACVSSGNSNANSRGPSFSSGLVPLSLRGS HMGPMEIYKTVSAWKRQPVRVLSLFRNIDKVLKSLGFLESGSGSGGGTLKYVED VTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRILQYALPRQES QRPFFWIFMDNLLLTEDDQETTTRFLQTEAVTLQDVRGRDYQNAMRVWSNIPGL KSKHAPLTPKEEEYLQAQVRSRSKLDAPKVDLLVKNCLLPLREYFKYFSQNSLPL GGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTST EEGTSTEPSEGSAPGTSTEPSEGMAPKKKRKVGRVPAAMAERPFQCRICMRNFSR SANLARHIRTHTGEKPFACDICGRKFARSDNLREHTKIHTGSQKPFQCRICMRNFS RPYTLRLHIRTHTGEKPFACDICGRKFAHRSNLNKHTKIHTGSQKPFQCRICMRNF SQSGSLTRHIRTHTGEKPFACDICGRKFATSANLSRHTKIHLRQKDAARGSGPKK KRKVNGGGGSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSL GYQLTKPDVILRLEKGEEPWLV	3xFLAG- DNMT3AL- XTEN80-SV40 NLS-ZFP-SV40 NLS-KRAB (AA 서열)
1021	MHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCE DSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARK GLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLES NPVMIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFS KVRTITTRSNSIKQGKDQHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSR LARQRLLGRSWSVPVIRHLFAPLKEYFACVSSGNSNANSRGPSFSSGLVPLSLRGS	3xFLAG- DNMT3AL- XTEN80-SV40 NLS-ZFP-SV40 NLS-KRAB (AA 서열)

[1301]

	HMGPMEIYKTVSAWKRQPVRVLSLFRNIDKVLKSLGFLESGSGSGGGTLKYVED VTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRILQYALPRQES QRPFFWIFMDNLLLTEDDQETTTRFLQTEAVTLQDVRGRDYQNAMRVWSNIPGL KSKHAPLTPKEEEYLQAQVRSRSKLDAPKVDLLVKNCLLPLREYFKYFSQNSLPL GGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTST EEGTSTEPSEGSAPGTSTEPSEGMAPKKKRKVGRVPLEPGEKPYACPECGKSFSRS DDLVRHQRTHTGEKPYKCPECGKSFSTSGSLVRHQRTHTGEKPYKCPECGKSFS RSDKLVRHQRTHTGEKPYACPECGKSFSRSDELVRHQRTHTGEKPYKCPECGKS FSTSHSLTEHQRTHTGEKPYKCPECGKSFSRADNLTEHQRTHTGKKTSGSGPKKK RKVNGGGGSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSLG YQLTKPDVILRLEKGEEPWLV	
1022	MHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCE DSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARK GLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLES NPVMIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFS KVRTITTRSNSIKQGKDQHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSR LARQRLLGRSWSVPVIRHLFAPLKEYFACVSSGNSNANSRGPSFSSGLVPLSLRGS HMGPMEIYKTVSAWKRQPVRVLSLFRNIDKVLKSLGFLESGSGSGGGTLKYVED VTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRILQYALPRQES QRPFFWIFMDNLLLTEDDQETTTRFLQTEAVTLQDVRGRDYQNAMRVWSNIPGL KSKHAPLTPKEEEYLQAQVRSRSKLDAPKVDLLVKNCLLPLREYFKYFSQNSLPL GGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTST EEGTSTEPSEGSAPGTSTEPSEGMAPKKKRKVGRVPLEPGEKPYACPECGKSFSE RSHLREHQRTHTGEKPYKCPECGKSFSTSHSLTEHQRTHTGEKPYKCPECGKSFS QAGHLASHQRTHTGEKPYACPECGKSFSTSHSLTEHQRTHTGEKPYKCPECGKSF SDPGHLVRHQRTHTGEKPYKCPECGKSFSTSGNLVRHQRTHTGKKTSGSGPKKK RKVNGGGGSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSLG YQLTKPDVILRLEKGEEPWLV	3xFLAG- DNMT3AL- XTEN80-SV40 NLS-ZFP-SV40 NLS-KRAB (AA 서열)
1023	MHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCE DSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARK GLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLES NPVMIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFS KVRTITTRSNSIKQGKDQHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSR LARQRLLGRSWSVPVIRHLFAPLKEYFACVSSGNSNANSRGPSFSSGLVPLSLRGS HMGPMEIYKTVSAWKRQPVRVLSLFRNIDKVLKSLGFLESGSGSGGGTLKYVED VTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRILQYALPRQES QRPFFWIFMDNLLLTEDDQETTTRFLQTEAVTLQDVRGRDYQNAMRVWSNIPGL KSKHAPLTPKEEEYLQAQVRSRSKLDAPKVDLLVKNCLLPLREYFKYFSQNSLPL GGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTST EEGTSTEPSEGSAPGTSTEPSEGMAPKKKRKVGRVPLEPGEKPYKCPECGKSFSR ADNLTEHQRTHTGEKPYKCPECGKSFSTSGSLVRHQRTHTGEKPYKCPECGKSFS RKDNLKNHQRTHTGEKPYKCPECGKSFSQSSSLVRHQRTHTGEKPYKCPECGKS FSRSDKLVRHQRTHTGEKPYKCPECGKSFSDSGNLRVHQRTHTGKKTSGSGPKK KRKVNGGGGSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSL GYQLTKPDVILRLEKGEEPWLV	3xFLAG- DNMT3AL- XTEN80-SV40 NLS-ZFP-SV40 NLS-KRAB (AA 서열)
1024	MHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCE DSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARK GLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLES NPVMIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFS KVRTITTRSNSIKQGKDQHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSR LARQRLLGRSWSVPVIRHLFAPLKEYFACVSSGNSNANSRGPSFSSGLVPLSLRGS HMGPMEIYKTVSAWKRQPVRVLSLFRNIDKVLKSLGFLESGSGSGGGTLKYVED VTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRILQYALPRQES QRPFFWIFMDNLLLTEDDQETTTRFLQTEAVTLQDVRGRDYQNAMRVWSNIPGL KSKHAPLTPKEEEYLQAQVRSRSKLDAPKVDLLVKNCLLPLREYFKYFSQNSLPL GGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTST EEGTSTEPSEGSAPGTSTEPSEGMAPKKKRKVGRVPLEPGEKPYACPECGKSFSQ SSSLVRHQRTHTGEKPYKCPECGKSFSQSGDLRRHQRTHTGEKPYKCPECGKSFS RSDERKRHQRTHTGEKPYACPECGKSFSHRTTLTNHQRTHTGEKPYKCPECGKS	3xFLAG- DNMT3AL- XTEN80-SV40 NLS-ZFP-SV40 NLS-KRAB (AA 서열)

[1302]

	FSRSDHLTNHQRTHTGEKPYKCPECGKSFSTSGELVRHQRTHTGKKTSGSGPKK KRKVNGGGGSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSL GYQLTKPDVILRLEKGEEPWLV	
1025	MHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCE DSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARK GLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLES NPVMIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFS KVRTITTRSNSIKQGKDQHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSR LARQRLLGRSWSVPVIRHLFAPLKEYFACVSSGNSNANSRGPSFSSGLVPLSLRGS HMGPMEIYKTVSAWKRQPVRVLSLFRNIDKVLKSLGFLESGSGSGGGTLKYVED VTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRILQYALPRQES QRPFFWIFMDNLLLTEDDQETTTRFLQTEAVTLQDVRGRDYQNAMRVWSNIPGL KSKHAPLTPKEEEYLQAQVRSRSKLDAPKVDLLVKNCLLPLREYFKYFSQNSLPL GGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTST EEGTSTEPSEGSAPGTSTEPSEGMAPKKKRKVGRVPLEPGEKPYACPECGKSFSQ SGDLRRHQRTHTGEKPYKCPECGKSFSRSDERKRHQRTHTGEKPYKCPECGKSF SHRTTLTNHQRTHTGEKPYACPECGKSFSRSDHLTNHQRTHTGEKPYKCPECGK SFSTSGELVRHQRTHTGEKPYKCPECGKSFSRSDDLVRHQRTHTGKKTSGSGPKK KRKVNGGGGSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSL GYQLTKPDVILRLEKGEEPWLV	3xFLAG- DNMT3AL- XTEN80-SV40 NLS-ZFP-SV40 NLS-KRAB (AA 서열)
1026	MHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCE DSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARK GLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLES NPVMIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFS KVRTITTRSNSIKQGKDQHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSR LARQRLLGRSWSVPVIRHLFAPLKEYFACVSSGNSNANSRGPSFSSGLVPLSLRGS HMGPMEIYKTVSAWKRQPVRVLSLFRNIDKVLKSLGFLESGSGSGGGTLKYVED VTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRILQYALPRQES QRPFFWIFMDNLLLTEDDQETTTRFLQTEAVTLQDVRGRDYQNAMRVWSNIPGL KSKHAPLTPKEEEYLQAQVRSRSKLDAPKVDLLVKNCLLPLREYFKYFSQNSLPL GGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTST EEGTSTEPSEGSAPGTSTEPSEGMAPKKKRKVGRVPLEPGEKPYACPECGKSFSQ RAHLERHQRTHTGEKPYKCPECGKSFSQLAHLRAHQRTHTGEKPYKCPECGKSF SDPGHLVRHQRTHTGEKPYACPECGKSFSRRSACRRHQRTHTGEKPYKCPECGK SFSRSDHLTTHQRTHTGEKPYKCPECGKSFSQSSSLVRHQRTHTGKKTSGSGPKK KRKVNGGGGSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSL GYQLTKPDVILRLEKGEEPWLV	3xFLAG- DNMT3AL- XTEN80-SV40 NLS-ZFP-SV40 NLS-KRAB (AA 서열)
1027	MHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCE DSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARK GLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLES NPVMIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFS KVRTITTRSNSIKQGKDQHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSR LARQRLLGRSWSVPVIRHLFAPLKEYFACVSSGNSNANSRGPSFSSGLVPLSLRGS HMGPMEIYKTVSAWKRQPVRVLSLFRNIDKVLKSLGFLESGSGSGGGTLKYVED VTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRILQYALPRQES QRPFFWIFMDNLLLTEDDQETTTRFLQTEAVTLQDVRGRDYQNAMRVWSNIPGL KSKHAPLTPKEEEYLQAQVRSRSKLDAPKVDLLVKNCLLPLREYFKYFSQNSLPL GGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTST EEGTSTEPSEGSAPGTSTEPSEGMAPKKKRKVGRVPLEPGEKPYACPECGKSFSQ SSNLVRHQRTHTGEKPYKCPECGKSFSRSDDLVRHQRTHTGEKPYKCPECGKSFS THLDLIRHQRTHTGEKPYACPECGKSFSTSGNLTEHQRTHTGEKPYKCPECGKSF SRRSACRRHQRTHTGEKPYKCPECGKSFSRNDTLTEHQRTHTGKKTSGSGPKKK RKVNGGGGSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSLG YQLTKPDVILRLEKGEEPWLV	3xFLAG- DNMT3AL- XTEN80-SV40 NLS-ZFP-SV40 NLS-KRAB (AA 서열)

[1303]

[1304] [0626] 본 발명은 예를 들어, 발명의 다양한 측면을 설명하기 위해 제공된 특정한 공개된 구현예에 그 범위가 제한되도록 의도된 것이 아니다. 설명된 조성물 및 방법에 대한 다양한 변형은 본원에 제시된 설명 및 교시로부터 명백해질 것이다. 이러한 변형은 본 발명의 진정한 범위 및 정신에서 벗어나지 않고 실행될 수 있으며, 본 발명의 범위 내에 포함되도록 의도된다.

서열

SEQ ID NO	서열	설명
586	GTTTAAGAGCTATGCTGGAAACAGCATAGCAAGTTTAAATAAGGCTAG TCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC	SpCas9 gRNA 스캐폴드 서열 (DNA)
587	GUUUAAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAAUAAGGCU AGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC	SpCas9 gRNA 스캐폴드 서열 (RNA)
588	NGG	S. pyogenes Cas9 프로토스페이서 인접 모티프
589	NNGRRT	S. aureus Cas9 프로토스페이서 인접 모티프
590	RTLVTFKDVFVDFTREEWKLLDTAQQILYRNVMLENYKNLVSLGYQLTKP DVILRLEKGEEPWLV	KRAB (AA)
591	CGGACACTGGTGACCTTCAAGGATGTGTTTGTGGACTTCACCAGGGAGG AGTGGAAGCTGCTGGACACTGCTCAGCAGATCCTGTACAGAAATGTGA TGCTGGAGAACTATAAGAACCTGGTTTCCTTGGGTTATCAGCTTACTAA GCCAGATGTGATCCTCCGGTTGGAGAAGGGAGAAGAGCCCTGGCTGGT G	KRAB (nt)
592	KRPAATKKAGQAKKKK	NLS (AA)
593	PKKKRKV	NLS-2 (AA)
594	CCAAAGAAGAAGCGGAAGGTCGGTATCCACGGAGTCCCAGCAGCCGAC AAGAAGTACTCCATTGGGCTCGCCATCGGCACAAACAGCGTCGGCTGG GCCGTCATTACGGACGAGTACAAGGTGCCGAGCAAAAAATTCAAAGTT CTGGGCAATACCGATCGCCACAGCATAAAGAAGAACCTCATTGGCGCC CTCCTGTTCGACTCCGGGGAAACCGCCGAAGCCACGCGGCTCAAAAGA ACAGCACGGCGCAGATATACCCGCAGAAAGAATCGGATCTGCTACCtgca GGAGATCTTTAGTAATGAGATGGCTAAGGTGGATGACTCTTTCTTCCAT AGGCTGGAGGAGTCCTTTTTGGTGGAGGAGGATAAAAAGCACGAGCGC CACCCAATCTTTGGCAATATCGTGGACGAGGTGGCGTACCATGAAAAGT ACCCAACCATATATCATCTGAGGAAGAAGCTTGTAGACAGTACTGATA AGGCTGACTTGCGGTTGATCTATCTCGCGCTGGCGCATATGATCAAATT TCGGGGACACTTCCTCATCGAGGGGGACCTGAACCCAGACAACAGCGA TGTCGACAAACTCTTTATCCAACTGGTTCAGACTTACAATCAGCTTTTCG	NLS2-dSpCas9-NLS-KRAB-NLS2 (nt)

[1305]

	AAGAGAACCCGATCAACGCATCCGGAGTTGACGCCAAAGCAATCCTGA GCGCTAGGCTGTCCAAATCCCGGCGGCTCGAAAACCTCATCGCACAGCT CCCTGGGGAGAAGAAGAACGGCCTGTTTGGTAATCTTATCGCCCTGTCA CTCGGGCTGACCCCCAACTTTAAATCTAACTTCGACCTGGCCGAAGATG CCAAGCTTCAACTGAGCAAAGACACCTACGATGATGATCTCGACAATCT GCTGGCCCAGATCGGCGACCAGTACGCAGACCTTTTTTTGGCGGCAAAG AACCTGTCAGACGCCATTCTGCTGAGTGATATTCTGCGAGTGAACACGG AGATCACCAAAGCTCCGCTGAGCGCTAGTATGATCAAGCGCTATGATG AGCACCACCAAGACTTGACTTTGCTGAAGGCCCTTGTCAGACAGCAACT GCCTGAGAAGTACAAGGAAATTTTCTTCGATCAGTCTAAAAATGGCTAC GCCGGATACATTGACGGCGGAGCAAGCCAGGAGGAATTTTACAAATTT ATTAAGCCCATCTTGGAAAAAATGGACGGCACCGAGGAGCTGCTGGTA AAGCTTAACAGAGAAGATCTGTTGCGCAAACAGCGCACTTTCGACAAT GGAAGCATCCCCCACCAGATTCACCTGGGCGAACTGCACGCTATCCTCA GGCGGCAAGAGGATTTCTACCCCTTTTTGAAAGATAACAGGGAAAAGA TTGAGAAAATCCTCACATTTCGGATACCCTACTATGTAGGCCCCCTCGC CCGGGGAAATTCCAGATTCGCGTGGATGACTCGCAAATCAGAAGAGAC CATCACTCCCTGGAACTTCGAGGAAGTCGTGGATAAGGGGGCCTCTGCC CAGTCCTTCATCGAAAGGATGACTAACTTTGATAAAAATCTGCCTAACG AAAAGGTGCTTCCTAAACACTCTCTGCTGTACGAGTACTTCACAGTTTA TAACGAGCTCACCAAGGTCAAATACGTCACAGAAGGGATGAGAAAGCC AGCATTCCTGTCTGGAGAGCAGAAGAAAGCTATCGTGGACCTCCTCTTC AAGACGAACCGGAAAGTTACCGTGAAACAGCTCAAAGAAGACTATTTC AAAAAGATTGAATGTTTCGACTCTGTTGAAATCAGCGGAGTGGAGGAT CGCTTCAACGCATCCCTGGGAACGTATCACGATCTCCTGAAAATCATTA AAGACAAGGACTTCCTGGACAATGAGGAGAACGAGGACATTCTTGAGG ACATTGTCCTCACCCTTACGTTGTTTGAAGATAGGGAGATGATTGAAGA ACGCTTGAAAACTTACGCTCATCTCTTCGACGACAAAGTCATGAAACAG CTCAAGAGGCGCCGATATACAGGATGGGGGCGGCTGTCAAGAAAACTG ATCAATGGgatcCGAGACAAGCAGAGTGGAAAGACAATCCTGGATTTTCT TAAGTCCGATGGATTTGCCAACCGGAACTTCATGCAGTTGATCCATGAT GACTCTCTCACCTTTAAGGAGGACATCCAGAAAGCACAAGTTTCTGGCC AGGGGGACAGTCTTCACGAGCACATCGCTAATCTTGCAGGTAGCCCAG CTATCAAAAAGGGAATACTGCAGACCGTTAAGGTCGTGGATGAACTCG TCAAAGTAATGGGAAGGCATAAGCCCGAGAATATCGTTATCGAGATGG CCCGAGAGAACCAAACTACCCAGAAGGGACAGAAGAACAGTAGGGAA AGGATGAAGAGGATTGAAGAGGGTATAAAAGAACTGGGGTCCCAAATC CTTAAGGAACACCCAGTTGAAAACACCCAGCTTCAGAATGAGAAGCTC TACCTGTACTACCTGCAGAACGGCAGGGACATGTACGTGGATCAGGAA	

[1306]

	CTGGACATCAATCGGCTCTCCGACTACGACGTGGATGCCATCGTGCCCC AGTCTTTTCTCAAAGATGATTCTATTGATAATAAAGTGTTGACAAGATC CGATAAAAATAGAGGGAAGAGTGATAACGTCCCCTCAGAAGAAGTTGT CAAGAAAATGAAAAATTATTGGCGGCAGCTGCTGAACGCCAAACTGAT CACACAACGGAAGTTCGATAATCTGACTAAGGCTGAACGAGGTGGCCT GTCTGAGTTGGATAAAGCCGGCTTCATCAAAAGGCAGCTTGTTGAGACA CGCCAGATCACCAAgcacGTGGCCCAAATTCTCGATTCACGCATGAACAC CAAGTACGATGAAAATGACAAACTGATTCGAGAGGTGAAAGTTATTAC TCTGAAGTCTAAGCTGGTCTCAGATTTCAGAAAGGACTTTCAGTTTTAT AAGGTGAGAGAGATCAACAATTACCACCATGCGCATGATGCCTACCTG AATGCAGTGGTAGGCACTGCACTTATCAAAAAATATCCCAAGCTTGAAT CTGAATTTGTTTACGGAGACTATAAAGTGTACGATGTTAGGAAAATGAT CGCAAAGTCTGAGCAGGAAATAGGCAAGGCCACCGCTAAGTACTTCTT TTACAGCAATATTATGAATTTTTTCAAGACCGAGATTACACTGGCCAAT GGAGAGATTCGGAAGCGACCACTTATCGAAACAAACGGAGAAACAGG AGAAATCGTGTGGGACAAGGGTAGGGATTTCGCGACAGTCCGGAAGGT CCTGTCCATGCCGCAGGTGAACATCGTTAAAAAGACCGAAGTACAGAC CGGAGGCTTCTCCAAGGAAAGTATCCTCCCGAAAAGGAACAGCGACAA GCTGATCGCACGCAAAAAAGATTGGGACCCCAAGAAATACGGCGGATT CGATTCTCCTACAGTCGCTTACAGTGTACTGGTTGTGGCCAAAGTGGAG AAAGGGAAGTCTAAAAAACTCAAAAGCGTCAAGGAACTGCTGGGCATC ACAATCATGGAGCGATCAAGCTTCGAAAAAAACCCCATCGACTTTCTCG AGGCGAAAGGATATAAAGAGGTCAAAAAAGACCTCATCATTAAGCTTC CCAAGTACTCTCTCTTTGAGCTTGAAAACGGCCGGAAACGAATGCTCGC TAGTGCGGGCGAGCTGCAGAAAGGTAACGAGCTGGCACTGCCCTCTAA ATACGTTAATTTCTTGTATCTGGCCAGCCACTATGAAAAGCTCAAAGGG TCTCCCGAAGATAATGAGCAGAAGCAGCTGTTCGTGGAACAACACAAA CACTACCTTGATGAGATCATCGAGCAAATAAGCGAATTCTCCAAAAGA GTGATCCTCGCCGACGCTAACCTCGATAAGGTGCTTTCTGCTTACAATA AGCACAGGGATAAGCCCATCAGGGAGCAGGCAGAAAACATTATCCACT TGTTTACTCTGACCAACTTGGGCGCGCCTGCAGCCTTCAAGTACTTCGA CACCACCATAGACAGAAAGCGGTACACCTCTACAAAGGAGGTCCTGGA CGCCACACTGATTCATCAGTCAATTACGGGGCTCTATGAAACAAGAATC GACCTCTCTCAGCTCGGTGGAGACAAAAGGCCGGCGGCCACGAAAAAG GCCGGCCAGGCAAAAAAGAAAAAGGctagCgatgctaagtcactgactgcctggtcccgga cactggtgaccttcaaggatgtgtttgtggacttcaccagggaggagtggaagctgctggacactgctcagcagatcct gtacagaaatgtgatgctggagaactataagaacctggtttccttgggttatcagcttactaagccagatgtgatcctccg gttggagaagggagaagagccctggctggtggagagagaaattcaccaagagacccatcctgattcagagactgcat ttgaaatcaaatcatcagttCCGAAAAAGAAACGCAAAGTT	

[1307]

595	PKKKRKVGIHGVPAADKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLG NTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSN EMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLR KKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQ TYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLI ALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLA AKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLP EKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNRE DLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPY YVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDK NLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVD LLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIK DKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKR RRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTF KEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRH KPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQL QNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKV LTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAER GGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVI TLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLES EFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEI RKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKE SILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLK SVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRK RMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQ HKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFT LTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLG GDKRPAATKKAGQAKKKKASDAKSLTAWSRTLVTFKDVFVDFTREEWKL LDTAQQILYRNVMLENYKNLVSLGYQLTKPDVILRLEKGEEPWLVEREIHQ ETHPDSETAFEIKSSVPKKKRKV	NLS2-dSpCas9- NLS-KRAB-NLS2 (AA)
596	MKRNYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRG ARRLKRRRRHRIQRVKKLLFDYNLLTDHSELSGINPYEARVKGLSQKLSEE EFSAALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRNSKALEEKYVAELQ LERLKKDGEVRGSINRFKTSDYVKEAKQLLKVQKAYHQLDQSFIDTYIDLL ETRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFPEELRSVKYAYNAD LYNALNDLNNLVITRDENEKLEYYEKFQIIENVFKQKKKPTLKQIAKEILVN EEDIKGYRVTSTGKPEFTNLKVYHDIKDITARKEIIENAELLDQIAKILTIYQS SEDIQEELTNLNSELTQEEIEQISNLKGYTGTHNLSLKAINLILDELWHTNDN QIAIFNRLKLVPKKVDLSQQKEIPTTLVDDFILSPVVKRSFIQSIKVINAIIKKY	SaCas9 (AA)

[1308]

	GLPNDIIIELAREKNSKDAQKMINEMQKRNRQTNERIEEIIRTTGKENAKYLI EKIKLHDMQEGKCLYSLEAIPLEDLLNNPFNYEVDHIIPRSVSFDNSFNNKV LVKQEENSKKGNRTPFQYLSSSDSKISYETFKKHILNLAKGKGRISKTKKEY LLEERDINRFSVQKDFINRNLVDTRYATRGLMNLLRSYFRVNNLDVKVKSI NGGFTSFLRRKWKFKKERNKGYKHHAEDALIIANADFIFKEWKKLDKAKK VMENQMFEEKQAESMPEIETEQEYKEIFITPHQIKHIKDFKDYKYSHRVDK KPNRELINDTLYSTRKDDKGNTLIVNNLNGLYDKDNDKLKKLINKSPEKLL MYHHDPQTYQKLKLIMEQYGDEKNPLYKYYEETGNYLTKYSKKDNGPVI KKIKYYGNKLNAHLDITDDYPNSRNKVVKLSLKPYRFDVYLDNGVYKFVT VKNLDVIKKENYYEVNSKCYEEAKKLKKISNQAEFIASFYNNDLIKINGEL YRVIGVNNDLLNRIEVNMIDITYREYLENMNDKRPPRIIKTIASKTQSIKKYS TDILGNLYEVKSKKHPQIIKKG	
597	KRNYILGLAIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRGA RRLKRRRRHRIQRVKKLLFDYNLLTDHSELSGINPYEARVKGLSQKLSEEEF SAALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRNSKALEEKYVAELQL ERLKKDGEVRGSINRFKTSDYVKEAKQLLKVQKAYHQLDQSFIDTYIDLLE TRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFPEELRSVKYAYNADL YNALNDLNNLVITRDENEKLEYYEKFQIIENVFKQKKKPTLKQIAKEILVNE EDIKGYRVTSTGKPEFTNLKVYHDIKDITARKEIIENAELLDQIAKILTIYQSS EDIQEELTNLNSELTQEEIEQISNLKGYTGTHNLSLKAINLILDELWHTNDN QIAIFNRLKLVPKKVDLSQQKEIPTTLVDDFILSPVVKRSFIQSIKVINAIIKKY GLPNDIIIELAREKNSKDAQKMINEMQKRNRQTNERIEEIIRTTGKENAKYLI EKIKLHDMQEGKCLYSLEAIPLEDLLNNPFNYEVDHIIPRSVSFDNSFNNKV LVKQEEASKKGNRTPFQYLSSSDSKISYETFKKHILNLAKGKGRISKTKKEY LLEERDINRFSVQKDFINRNLVDTRYATRGLMNLLRSYFRVNNLDVKVKSI NGGFTSFLRRKWKFKKERNKGYKHHAEDALIIANADFIFKEWKKLDKAKK VMENQMFEEKQAESMPEIETEQEYKEIFITPHQIKHIKDFKDYKYSHRVDK KPNRELINDTLYSTRKDDKGNTLIVNNLNGLYDKDNDKLKKLINKSPEKLL MYHHDPQTYQKLKLIMEQYGDEKNPLYKYYEETGNYLTKYSKKDNGPVI KKIKYYGNKLNAHLDITDDYPNSRNKVVKLSLKPYRFDVYLDNGVYKFVT VKNLDVIKKENYYEVNSKCYEEAKKLKKISNQAEFIASFYNNDLIKINGEL YRVIGVNNDLLNRIEVNMIDITYREYLENMNDKRPPRIIKTIASKTQSIKKYS TDILGNLYEVKSKKHPQIIKKG	dSaCas9 (AA)
598	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGAL LFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLE ESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLI YLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGV DAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDL AEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVN	SpCas9 (AA)

[1309]

	TEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYA GYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPH QIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAW MTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLY EYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLK EDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILED IVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLING IRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSL HEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQ KGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRD MYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPS EEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLV ETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFY KVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMI AKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWD KGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKD WDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFE KNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNEL ALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSK RVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTT IDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD	
599	DKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALL FDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEE SFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIY LALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVD AKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLA EDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNT EITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAG YIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQ IHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWM TRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYE YFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDI VLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGI RDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLH EHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQK GQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDM YVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSE EVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVE	dSpCas9 (AA)

[1310]

	TRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYK VREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIA KSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDK GRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDW DPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKN PIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELAL PSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRV ILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTID RKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD	
600	MKTPADTGFAFPDWAYKPESSPGSRQIQLWHFILELLRKEEYQGVIAWQG DYGEFVIKDPDEVARLWGVRKCKPQMNYDKLSRALRYYYNKRILHKTKG KRFTYKFNFNKLVLVNYPFIDVGLAGGAVPQSAPPVPSGGSHFRFPPSTPSE VLSPTEDPRSPPACSSSSSSLFSAVVARRLGRGSVSDCSDGTSELEEPLGEDP RARPPGPPDLGAFRGPPLARLPHDPGVFRVYPRPRGGPEPLSPFPVSPLAGP GSLLPPQLSPALPMTPTHLAYTPSPTLSPMYPSGGGGPSGSGGGSHFSFSPED MKRYLQAHTQSVYNYHLSPRAFLHYPGLVVPQPQRPDKCPLPPMAPETPP VPSSASSSSSSSSSPFKFKLQPPPLGRRQRAAGEKAVAGADKSGGSAGGLAE GAGALAPPPPPPQIKVEPISEGESEEVEVTDISDEDEEDGEVFKTPRAPPAPP KPEPGEAPGASQCMPLKLRFKRRWSEDCRLEGGGGPAGGFEDEGEDKKVR GEGPGEAGGPLTPRRVSSDLQHATAQLSLEHRDS	ERF 도메인 (AA)
601	MERVKMINVQRLLEAAEFLERRERECEHGYASSFPSMPSPRLQHSKPPRRL SRAQKHSSGSSNTSTANRSTHNELEKNRRAHLRLCLERLKVLIPLGPDCTR HTTLGLLNKAKAHIKKLEEAERKSQHQLENLEREQRFLKWRLEQLQGPQE MERIRMDSIGSTISSDRSDSEREEIEVDVESTEFSHGEVDNISTTSISDIDDHSS LPSIGSDEGYSSASVKLSFTS	MXI1 도메인 (AA)
602	ASPKKKRKVEASGSGMNIQMLLEAADYLERREREAEHGYASMLPGSGMNI QMLLEAADYLERREREAEHGYASMLPGSGMNIQMLLEAADYLERREREA EHGYASMLPGSGMNIQMLLEAADYLERREREAEHGYASMLPSRSR	SID4X 도메인 (AA)
603	MAAAVRMNIQMLLEAADYLERREREAEHGYASMLPYNNKDRDALKRRN KSKKNNSSSRST HNEMEKNRRAHLRLCLEKLKGLVPLGPESSRHTTLSLLTKAKLHIKKLEDC DRKAVHQID QLQREQRHLKRQLEKLGIERIRMDSIGSTVSSERSDSDREEIDVDVESTDYL TGDLDWSS SSVSDSDERGSMQSLGSDEGYSSTSIKRIKLQDSHKACLG	MAD-SID 도메인 (AA)
604	MPAMPSSGPGDTSSSAAEREEDRKDGEEQEEPRGKEERQEPSTTARKVGRP GRKRKHPPVESGDTPKDPAVISKSPSMAQDSGASELLPNGDLEKRSEPQPEE GSPAGGQKGGAPAEGEGAAETLPEASRAVENGCCTPKEGRGAPAEAGKEQ KETNIESMKMEGSRGRLRGGLGWESSLRQRPMPRLTFQAGDPYYISKRKR	DNMT3A (AA)

[1311]

	DEWLARWKREAEKKAKVIAGMNAVEENQGPGESQKVEEASPPAVQQPTD PASPTVATTPEPVGSDAGDKNATKAGDDEPEYEDGRGFGIGELVWGKLRG FSWWPGRIVSWWMTGRSRAAEGTRWVMWFGDGKFSVVCVEKLMPLSSF CSAFHQATYNKQPMYRKAIYEVLQVASSRAGKLFPVCHDSDESDTAKAVE VQNKPMIEWALGGFQPSGPKGLEPPEEEKNPYKEVYTDMWVEPEAAAYA PPPPAKKPRKSTAEKPKVKEIIDERTRERLVYEVRQKCRNIEDICISCGSLNV TLEHPLFVGGMCQNCKNCFLECAYQYDDDGYQSYCTICCGGREVLMCGN NNCCRCFCVECVDLLVGPGAAQAAIKEDPWNCYMCGHKGTYGLLRRRED WPSRLQMFFANNHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLK DLGIQVDRYIASEVCEDSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPF DLVIGGSPCNDLSIVNPARKGLYEGTGRLFFEFYRLLHDARPKEGDDRPFF WLFENVVAMGVSDKRDISRFLESNPVMIDAKEVSAAHRARYFWGNLPGM NRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIKQGKDQHFPVFMN EKEDILWCTEMERVFGFPVHYTDVSNMSRLARQRLLGRSWSVPVIRHLFAP LKEYFA	
605	MKGDTRHLNGEEDAGGREDSILVNGACSDQSSDSPPILEAIRTPEIRGRRSSS RLSKREVSSLLSYTQDLTGDGDGEDGDGSDTPVMPKLFRETRTRSESPAVR TRNNNSVSSRERHRPSPRSTRGRQGRNHVDESPVEFPATRSLRRRATASAG TPWPSPPSSYLTIDLTDDTEDTHGTPQSSSTPYARLAQDSQQGGMESPQVEA DSGDGDSSEYQDGKEFGIGDLVWGKIKGFSWWPAMVVSWKATSKRQAM SGMRWVQWFGDGKFSEVSADKLVALGLFSQHFNLATFNKLVSYRKAMY HALEKARVRAGKTFPSSPGDSLEDQLKPMLEWAHGGFKPTGIEGLKPNNT QPVVNKSKVRRAGSRKLESRKYENKTRRRTADDSATSDYCPAPKRLKTNC YNNGKDRGDEDQSREQMASDVANNKSSLEDGCLSCGRKNPVSFHPLFEGG LCQTCRDRFLELFYMYDDDGYQSYCTVCCEGRELLLCSNTSCCRCFCVEC LEVLVGTGTAAEAKLQEPWSCYMCLPQRCHGVLRRRKDWNVRLQAFFTS DTGLEYEAPKLYPAIPAARRRPIRVLSLFDGIATGYLVLKELGIKVGKYVAS EVCEESIAVGTVKHEGNIKYVNDVRNITKKNIEEWGPFDLVIGGSPCNDLSN VNPARKGLYEGTGRLFFEFYHLLNYSRPKEGDDRPFFWMFENVVAMKVG DKRDISRFLECNPVMIDAIKVSAAHRARYFWGNLPGMNRPVIASKNDKLEL QDCLEYNRIAKLKKVQTITTKSNSIKQGKNQLFPVVMNGKEDVLWCTELE RIFGFPVHYTDVSNMGRGARQKLLGRSWSVPVIRHLFAPLKDYFACE	DNMT3B (AA)
606	MLSGKKAAAAAAAAAAAAATGTEAGPGTAGGSENGSEVAAQPAGLSGPAE VGPGAVGERTPRKKEPPRASPPGGLAEPPGSAGPQAGPTVVPGSATPMETG IAETPEGRRTSRRKRAKVEYREMDESLANLSEDEYYSEEERNAKAEKEKKL PPPPPQAPPEEENESEPEEPSGVEGAAFQSRLPHDRMTSQEAACFPDIISGPQ QTQKVFLFIRNRTLQLWLDNPKIQLTFEATLQQLEAPYNSDTVLVHRVHSY LERHGLINFGIYKRIKPLPTKKTGKVIIIGSGVSGLAAARQLQSFGMDVTLLE ARDRVGGRVATFRKGNYVADLGAMVVTGLGGNPMAVVSKQVNMELAKI	LSD1 (AA)

[1312]

	KQKCPLYEANGQAVPKEKDEMVEQEFNRLLEATSYLSHQLDFNVLNNKPV SLGQALEVVIQLQEKHVKDEQIEHWKKIVKTQEELKELLNKMVNLKEKIK ELHQQYKEASEVKPPRDITAEFLVKSKHRDLTALCKEYDELAETQGKLEEK LQELEANPPSDVYLSSRDRQILDWHFANLEFANATPLSTLSLKHWDQDDDF EFTGSHLTVRNGYSCVPVALAEGLDIKLNTAVRQVRYTASGCEVIAVNTRS TSQTFIYKCDAVLCTLPLGVLKQQPPAVQFVPPLPEWKTSAVQRMGFGNL NKVVLCFDRVFWDPSVNLFGHVGSTTASRGELFLFWNLYKAPILLALVAG EAAGIMENISDDVIVGRCLAILKGIFGSSAVPQPKETVVSRWRADPWARGS YSYVAAGSSGNDYDLMAQPITPGPSIPGAPQPIPRLFFAGEHTIRNYPATVH GALLSGLREAGRIADQFLGAMYTLPRQATPGVPAQQSPSM	
607	MAAIPALDPEAEPSMDVILVGSSELSSSVSPGTGRDLIAYEVKANQRNIEDIC ICCGSLQVHTQHPLFEGGICAPCKDKFLDALFLYDDDGYQSYCSICSGETLL ICGNPDCTRCYCFECVDSLVGPGTSGKVHAMSNWVCYLCLPSSSGLLQRR RKWRSQLKAFYDRESENPLEMFETVPVWRRQPVRVLSLFEDIKKELTSLGF LESGSDPGQLKHVVDVTDTVRKDVEEWGPFDLVYGATPPLGHTCDRPPSW YLFQFHRLLQYARPKPGSPRPFFWMFVDNLVLNKEDLDVASRFLEMEPVTI PDVHGGSLQNAVRVWSNIPAIRSRHWALVSEEELSLLAQNKQSSKLAAKW PTKLVKNCFLPLREYFKYFSTELTSSL	DNMT3L (AA)
608	MGQTGKKSEKGPVCWRKRVKSEYMRLRQLKRFRRADEVKTMFSSNRQKI LERTETLNQEWKQRRIQPVHIMTSVSSLRGTRECSVTSDLDFPAQVIPLKTL NAVASVPIMYSWSPLQQNFMVEDETVLHNIPYMGDEVLDQDGTFIEELIKN YDGKVHGDRECGFINDEIFVELVNALGQYNDDDDDDDGDDPDEREEKQK DLEDNRDDKETCPPRKFPADKIFEAISSMFPDKGTAEELKEKYKELTEQQLP GALPPECTPNIDGPNAKSVQREQSLHSFHTLFCRRCFKYDCFLHPFHATPNT YKRKNTETALDNKPCGPQCYQHLEGAKEFAAALTAERIKTPPKRPGGRRR GRLPNNSSRPSTPTISVLESKDTDSDREAGTETGGENNDKEEEEKKDETSSS SEANSRCQTPIKMKPNIEPPENVEWSGAEASMFRVLIGTYYDNFCAIARLIG TKTCRQVYEFRVKESSIIAPVPTEDVDTPPRKKKRKHRLWAAHCRKIQLKK DGSSNHVYNYQPCDHPRQPCDSSCPCVIAQNFCEKFCQCSSECQNRFPGCR CKAQCNTKQCPCYLAVRECDPDLCLTCGAADHWDSKNVSCKNCSIQRGS KKHLLLAPSDVAGWGIFIKDPVQKNEFISEYCGEIISQDEADRRGKVYDKY MCSFLFNLNNDFVVDATRKGNKIRFANHSVNPNCYAKVMMVNGDHRIGIF AKRAIQTGEELFFDYRYSQADALKYVGIEREMEIP	EZH2 (AA)
609	GGGGS	GGGGS 링커(AA)
610	GGGGG	GGGGG 링커(AA)
611	GGAGG	GGAGG 링커(AA)

[1313]

612	GGGGSSS	GGGGSSS 링커(AA)
613	GGGGAAA	GGGGAAA 링커(AA)
614	PKKKRKV	NLS - SV40 (AA)
615	PAAKRVKLD	NLS - cMyc (AA)
616	RQRRNELKRSP	NLS - cMyc (AA)
617	NQSSNFGPMKGGNFGGRSSGPYGGGGQYFAKPRNQGGY	NLS - hRNPA1 M9 (AA)
618	RMRIZFKNKGKDTAELRRRRVEVSVELRKAKKDEQILKRRNV	NLS - IBB 도메인 임포틴-알파(AA)
619	VSRKRPRP	NLS - 근종 T 단백질 n (AA)
620	PPKKARED	NLS - 근종 T 단백질 (AA)
621	PQPKKKPL	NLS - 인간 p53 (AA)
622	SALIKKKKKMAP	NLS - 마우스 c-abl IV (AA)
623	DRLRR	NLS - 인플루엔자 NS1 (AA)
624	PKQKKRK	NLS - 인플루엔자 NS1 (AA)
625	RKLKKKIKKL	NLS - 간염 델타 항원 (AA)
626	REKKKFLKRR	NLS - 마우스 Mx1 (AA)
627	KRKGDEVDGVDEVAKKKSKK	NLS - 인간 폴리(ADP-리보스) 폴리머라제 (AA)
628	RKCLQAGMNLEARKTKK	NLS - 글루코코르티코이드 (AA)
629	NGG	PAM - SpCas9
630	NNGRRT	PAM - SaCas9

[1314]

[1315]

631	NNNNGATT	PAM - N. meningitidis Cas9
632	NNNNRYAC	PAM - C. jejuni Cas9
633	NNAGAAW	PAM - S. thermophilus
634	NGG	PAM - F. Novicida
635	NAAAAC	PAM - T. denticola
636	TTTV	PAM - Cas12a/Cpf1
637	NGAN	변이체 PAM - SpCas9 변이체
638	NGNG	변이체 PAM - SpCas9 변이체
639	NGAG	변이체 PAM - SpCas9 변이체
640	NGCG	변이체 PAM - SpCas9 변이체
641	LLPKNYHLENEVARLKKLVGER	SunTag GCN4 펩타이드 (AA)
642	GGSGG	링커(AA)
643	GGPSSGAPPPSGGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSE	XTEN80 (aa)
644	ATGGACTACAAAGACCATGACGGTGATTATAAAGATCATGACATCGATTACAAGGATGACGATGACAAGCACGTTAACCACGATCAGGAGTTTGACCCCCCTAAGGTGTACCCACCCGTGCCAGCCGAGAAGAGGAAGCCCATCCGCGTGCTGTCCCTGTTCGACGGCATCGCCACAGGCCTGCTGGTGCTGAAGGATCTGGGCATCCAGGTGGACAGATATATCGCCTCCGAGGTGTGCGAGGATTCTATCACCGTGGGCATGGTGAGGCACCAGGGCAAGATCATGTACGTGGGCGACGTGCGCAGCGTGACACAGAAGCACATCCAGGAGTGGGGACCCTTCGACCTGGTCATCGGAGGCAGCCCCTGTAATGACCTGTCCATCGTGAACCCTGCAAGGAAGGGCCTGTATGAGGGAACCGGCAGACTGTTCTTTGAGTTCTACAGGCTGCTGCACGACGCCCGCCCTAAGGAGGGCGATGACAGGCCATTCTTTTGGCTGTTTGAGAACGTGGTGGCCATGGGCGTGAGCGACAAGCGGGATATCTCCAGATTCCTGGAGTCTAATCCCGTGATGATCGATGCAAAGGAGGTGTCTGCCGCACACAGGGCAAGGTACTTTTGGGGAAATCTGCCTGGCATGAACCGCCCACTGGCCAGCACCGTGAACGACAAGCTGGAGCTGCAGGAGTGCCTGGAGCACGGAAGGATCGCCAAGTTCTC	DNMT3A/L-dSpCas9-KRAB (nt)

	CAAGGTGCGGACAATCACCACAAGATCTAACAGCATCAAGCAGGGCAA GGATCAGCACTTCCCCGTGTTCATGAATGAGAAGGAGGACATCCTGTGG TGTACCGAGATGGAGCGCGTGTTCGGCTTTCCAGTGCACTATACAGACG TGAGCAATATGAGCCGGCTGGCAAGGCAGAGACTGCTGGGCCGGTCCT GGTCTGTGCCAGTGATCAGACACCTGTTCGCCCCCCTGAAGGAGTACTT TGCCTGCGTGTCTAGCGGCAACTCTAATGCCAACAGCAGAGGCCCTTCC TTTTCCTCTGGCCTGGTGCCACTGTCTCTGAGGGGCAGCCACATGGGCC CCATGGAGATCTACAAGACCGTGTCCGCCTGGAAGAGGCAGCCTGTGC GCGTGCTGTCTCTGTTCCGCAACATCGACAAGGTGCTGAAGAGCCTGGG CTTTCTGGAGAGCGGATCCGGATCTGGAGGAGGCACCCTGAAGTATGT GGAGGATGTGACAAATGTGGTGCGGAGAGATGTGGAGAAGTGGGGCCC CTTCGATCTGGTGTACGGATCCACCCAGCCACTGGGAAGCTCCTGCGAT AGGTGTCCAGGATGGTATATGTTCCAGTTTCACAGAATCCTGCAGTACG CACTGCCAAGGCAGGAGAGCCAGCGCCCTTTCTTTTGGATCTTTATGGA CAACCTGCTGCTGACAGAGGATGACCAGGAGACAACAACCCGCTTCCT GCAGACAGAGGCAGTGACCCTGCAGGATGTGAGGGGACGCGACTATCA GAATGCCATGCGGGTGTGGTCTAACATCCCTGGCCTGAAGAGCAAGCA CGCCCCCCTGACCCCTAAGGAGGAGGAGTACCTGCAGGCCCAGGTGCG GAGCAGATCCAAGCTGGATGCCCCTAAGGTGGACCTGCTGGTGAAGAA TTGTCTGCTGCCACTGCGGGAGTACTTCAAGTACTTTAGTCAGAATAGC CTGCCACTGgaggcaagcggatccggaagggcatctcctggaatcccaggaagcacccgcAACCCC AAGAAGAAGCGGAAGGTGGGCATCCACGGCGTGCCCGCCGCCGACAAG AAGTACAGCATCGGCCTGGCCATCGGCACCAACAGCGTGGGCTGGGCC GTGATCACCGACGAGTACAAGGTGCCCAGCAAGAAGTTCAAGGTGCTG GGCAACACCGACCGGCACAGCATCAAGAAGAACCTGATCGGCGCCCTG CTGTTCGACAGCGGCGAGACCGCCGAGGCCACCCGGCTGAAGCGGACC GCCCGGCGGCGGTACACCCGGCGGAAGAACCGGATCTGCTACCTGCAG GAGATCTTCAGCAACGAGATGGCCAAGGTGGACGACAGCTTCTTCCAC CGGCTGGAGGAGAGCTTCCTGGTGGAGGAGGACAAGAAGCACGAGCG GCACCCCATCTTCGGCAACATCGTGGACGAGGTGGCCTACCACGAGAA GTACCCCACCATCTACCACCTGCGGAAGAAGCTGGTGGACAGCACCGA CAAGGCCGACCTGCGGCTGATCTACCTGGCCCTGGCCCACATGATCAAG TTCCGGGGCCACTTCCTGATCGAGGGCGACCTGAACCCCGACAACAGC GACGTGGACAAGCTGTTCATCCAGCTGGTGCAGACCTACAACCAGCTGT TCGAGGAGAACCCCATCAACGCCAGCGGCGTGGACGCCAAGGCCATCC TGAGCGCCCGGCTGAGCAAGAGCCGGCGGCTGGAGAACCTGATCGCCC AGCTGCCCGGCGAGAAGAAGAACGGCCTGTTCGGCAACCTGATCGCCC TGAGCCTGGGCCTGACCCCCAACTTCAAGAGCAACTTCGACCTGGCCGA GGACGCCAAGCTGCAGCTGAGCAAGGACACCTACGACGACGACCTGGA	

[1316]

	CAACCTGCTGGCCCAGATCGGCGACCAGTACGCCGACCTGTTCCTGGCC GCCAAGAACCTGAGCGACGCCATCCTGCTGAGCGACATCCTGCGGGTG AACACCGAGATCACCAAGGCCCCCCTGAGCGCCAGCATGATCAAGCGG TACGACGAGCACCACCAGGACCTGACCCTGCTGAAGGCCCTGGTGCGG CAGCAGCTGCCCGAGAAGTACAAGGAGATCTTCTTCGACCAGAGCAAG AACGGCTACGCCGGCTACATCGACGGCGGCGCCAGCCAGGAGGAGTTC TACAAGTTCATCAAGCCCATCCTGGAGAAGATGGACGGCACCGAGGAG CTGCTGGTGAAGCTGAACCGGGAGGACCTGCTGCGGAAGCAGCGGACC TTCGACAACGGCAGCATCCCCCACCAGATCCACCTGGGCGAGCTGCAC GCCATCCTGCGGCGGCAGGAGGACTTCTACCCCTTCCTGAAGGACAACC GGGAGAAGATCGAGAAGATCCTGACCTTCCGGATCCCCTACTACGTGG GCCCCCTGGCCCGGGGCAACAGCCGGTTCGCCTGGATGACCCGGAAGA GCGAGGAGACCATCACCCCCTGGAACTTCGAGGAGGTGGTGGACAAGG GCGCCAGCGCCCAGAGCTTCATCGAGCGGATGACCAACTTCGACAAGA ACCTGCCCAACGAGAAGGTGCTGCCCAAGCACAGCCTGCTGTACGAGT ACTTCACCGTGTACAACGAGCTGACCAAGGTGAAGTACGTGACCGAGG GCATGCGGAAGCCCGCCTTCCTGAGCGGCGAGCAGAAGAAGGCCATCG TGGACCTGCTGTTCAAGACCAACCGGAAGGTGACCGTGAAGCAGCTGA AGGAGGACTACTTCAAGAAGATCGAGTGCTTCGACAGCGTGGAGATCA GCGGCGTGGAGGACCGGTTCAACGCCAGCCTGGGCACCTACCACGACC TGCTGAAGATCATCAAGGACAAGGACTTCCTGGACAACGAGGAGAACG AGGACATCCTGGAGGACATCGTGCTGACCCTGACCCTGTTCGAGGACCG GGAGATGATCGAGGAGCGGCTGAAGACCTACGCCCACCTGTTCGACGA CAAGGTGATGAAGCAGCTGAAGCGGCGGCGGTACACCGGCTGGGGCCG GCTGAGCCGGAAGCTGATCAACGGCATCCGGGACAAGCAGAGCGGCAA GACCATCCTGGACTTCCTGAAGAGCGACGGCTTCGCCAACCGGAACTTC ATGCAGCTGATCCACGACGACAGCCTGACCTTCAAGGAGGACATCCAG AAGGCCCAGGTGAGCGGCCAGGGCGACAGCCTGCACGAGCACATCGCC AACCTGGCCGGCAGCCCCGCCATCAAGAAGGGCATCCTGCAGACCGTG AAGGTGGTGGACGAGCTGGTGAAGGTGATGGGCCGGCACAAGCCCGAG AACATCGTGATCGAGATGGCCCGGGAGAACCAGACCACCCAGAAGGGC CAGAAGAACAGCCGGGAGCGGATGAAGCGGATCGAGGAGGGCATCAA GGAGCTGGGCAGCCAGATCCTGAAGGAGCACCCCGTGGAGAACACCCA GCTGCAGAACGAGAAGCTGTACCTGTACTACCTGCAGAACGGCCGGGA CATGTACGTGGACCAGGAGCTGGACATCAACCGGCTGAGCGACTACGA CGTGGACGCCATCGTGCCCCAGAGCTTCCTGAAGGACGACAGCATCGA CAACAAGGTGCTGACCCGGAGCGACAAGAACCGGGGCAAGAGCGACA ACGTGCCCAGCGAGGAGGTGGTGAAGAAGATGAAGAACTACTGGCGGC AGCTGCTGAACGCCAAGCTGATCACCCAGCGGAAGTTCGACAACCTGA	

[1317]

	CCAAGGCCGAGCGGGGCGGCCTGAGCGAGCTGGACAAGGCCGGCTTCA TCAAGCGGCAGCTGGTGGAGACCCGGCAGATCACCAAGCACGTGGCCC AGATCCTGGACAGCCGGATGAACACCAAGTACGACGAGAACGACAAGC TGATCCGGGAGGTGAAGGTGATCACCCTGAAGAGCAAGCTGGTGAGCG ACTTCCGGAAGGACTTCCAGTTCTACAAGGTGCGGGAGATCAACAACT ACCACCACGCCCACGACGCCTACCTGAACGCCGTGGTGGGCACCGCCCT GATCAAGAAGTACCCCAAGCTGGAGAGCGAGTTCGTGTACGGCGACTA CAAGGTGTACGACGTGCGGAAGATGATCGCCAAGAGCGAGCAGGAGAT CGGCAAGGCCACCGCCAAGTACTTCTTCTACAGCAACATCATGAACTTC TTCAAGACCGAGATCACCCTGGCCAACGGCGAGATCCGGAAGCGGCCC CTGATCGAGACCAACGGCGAGACCGGCGAGATCGTGTGGGACAAGGGC CGGGACTTCGCCACCGTGCGGAAGGTGCTGAGCATGCCCCAGGTGAAC ATCGTGAAGAAGACCGAGGTGCAGACCGGCGGCTTCAGCAAGGAGAGC ATCCTGCCCAAGCGGAACAGCGACAAGCTGATCGCCCGGAAGAAGGAC TGGGACCCCAAGAAGTACGGCGGCTTCGACAGCCCCACCGTGGCCTAC AGCGTGCTGGTGGTGGCCAAGGTGGAGAAGGGCAAGAGCAAGAAGCT GAAGAGCGTGAAGGAGCTGCTGGGCATCACCATCATGGAGCGGAGCAG CTTCGAGAAGAACCCCATCGACTTCCTGGAGGCCAAGGGCTACAAGGA GGTGAAGAAGGACCTGATCATCAAGCTGCCCAAGTACAGCCTGTTCGA GCTGGAGAACGGCCGGAAGCGGATGCTGGCCAGCGCCGGCGAGCTGCA GAAGGGCAACGAGCTGGCCCTGCCCAGCAAGTACGTGAACTTCCTGTA CCTGGCCAGCCACTACGAGAAGCTGAAGGGCAGCCCCGAGGACAACGA GCAGAAGCAGCTGTTCGTGGAGCAGCACAAGCACTACCTGGACGAGAT CATCGAGCAGATCAGCGAGTTCAGCAAGCGGGTGATCCTGGCCGACGC CAACCTGGACAAGGTGCTGAGCGCCTACAACAAGCACCGGGACAAGCC CATCCGGGAGCAGGCCGAGAACATCATCCACCTGTTCACCCTGACCAAC CTGGGCGCCCCCGCCGCCTTCAAGTACTTCGACACCACCATCGACCGGA AGCGGTACACCAGCACCAAGGAGGTGCTGGACGCCACCCTGATCCACC AGAGCATCACCGGCCTGTACGAGACCCGGATCGACCTGAGCCAGCTGG GCGGCGACAGCGGCGGCAAGCGGCCCGCCGCCACCAAGAAGGCCGGCC AGGCCAAGAAGAAGAAGGctagCgatgctaagtcactgactgcctggtcccggacactggtgacct tcaaggatgtgtttgtggacttcaccagggaggagtggaagctgctggacactgctcagcagatcctgtacagaaatgt gatgctggagaactataagaacctggtttccttgggttatcagcttactaagccagatgtgatcctccggttggagaagg gagaagagccctggctggtggagagagaaattcaccaagagacccatcctgattcagagactgcatttgaaatcaaat catcagttCCGAAAAAGAAACGCAAAGTT	

[1318]

645	MDYKDHDGDYKDHDIDYKDDDDKHVNHDQEFDPPKVYPPVPAEKRK PIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCEDSITVGMVRHQGKIMYV GDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARKGLYEGTGRLFFEFY RLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLESNPVMIDAKEV SAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTR SNSIKQGKDQHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSRLARQ RLLGRSWSVPVIRHLFAPLKEYFACVSSGNSNANSRGPSFSSGLVPLSLRGS HMGPMEIYKTVSAWKRQPVRVLSLFRNIDKVLKSLGFLESGSGSGGGTLK YVEDVTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRILQ YALPRQESQRPFFWIFMDNLLLTEDDQETTTRFLQTEAVTLQDVRGRDYQ NAMRVWSNIPGLKSKHAPLTPKEEEYLQAQVRSRSKLDAPKVDLLVKNCL LPLREYFKYFSQNSLPLEASGSGRASPGIPGSTRNPKKKRKVGIHGVPAADK KYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDS GETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFL VEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLA LAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDA KAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAE DAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTE ITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGY IDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQI HLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMT RKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEY FTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKED YFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIR DKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKG QKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMY VDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEE VVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVET RQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKV REINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKS EQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKG RDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDP KKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPI DFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPS KYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVIL ADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDR KRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGKRPAATKKAGQAK	DNMT3A/L- dSpCas9- KRAB(AA)

[1319]

	KKKASDAKSLTAWSRTLVTFKDVFVDFTREEWKLLDTAQQILYRNVMLE NYKNLVSLGYQLTKPDVILRLEKGEEPWLVEREIHQETHPDSETAFEIKSSV PKKKRKV	
646	ATGAACCACGATCAGGAGTTTGACCCCCCTAAGGTGTACCCACCCGTGC CAGCCGAGAAGAGGAAGCCCATCCGCGTGCTGTCCCTGTTCGACGGCA TCGCCACAGGCCTGCTGGTGCTGAAGGATCTGGGCATCCAGGTGGACA GATATATCGCCTCCGAGGTGTGCGAGGATTCTATCACCGTGGGCATGGT GAGGCACCAGGGCAAGATCATGTACGTGGGCGACGTGCGCAGCGTGAC ACAGAAGCACATCCAGGAGTGGGGACCCTTCGACCTGGTCATCGGAGG CAGCCCCTGTAATGACCTGTCCATCGTGAACCCTGCAAGGAAGGGCCTG TATGAGGGAACCGGCAGACTGTTCTTTGAGTTCTACAGGCTGCTGCACG ACGCCCGCCCTAAGGAGGGCGATGACAGGCCATTCTTTTGGCTGTTTGA GAACGTGGTGGCCATGGGCGTGAGCGACAAGCGGGATATCTCCAGATT CCTGGAGTCTAATCCCGTGATGATCGATGCAAAGGAGGTGTCTGCCGCA CACAGGGCAAGGTACTTTTGGGGAAATCTGCCTGGCATGAACCGCCCA CTGGCCAGCACCGTGAACGACAAGCTGGAGCTGCAGGAGTGCCTGGAG CACGGAAGGATCGCCAAGTTCTCCAAGGTGCGGACAATCACCACAAGA TCTAACAGCATCAAGCAGGGCAAGGATCAGCACTTCCCCGTGTTCATGA ATGAGAAGGAGGACATCCTGTGGTGTACCGAGATGGAGCGCGTGTTCG	DNMT3A/L- XTEN80-dSpCas9- KRAB (nt)

[1320]

	GCTTTCCAGTGCACTATACAGACGTGAGCAATATGAGCCGGCTGGCAA GGCAGAGACTGCTGGGCCGGTCCTGGTCTGTGCCAGTGATCAGACACCT GTTCGCCCCCCTGAAGGAGTACTTTGCCTGCGTGTCTAGCGGCAACTCT AATGCCAACAGCAGAGGCCCTTCCTTTTCCTCTGGCCTGGTGCCACTGT CTCTGAGGGGCAGCCACATGGGCCCCATGGAGATCTACAAGACCGTGT CCGCCTGGAAGAGGCAGCCTGTGCGCGTGCTGTCTCTGTTCCGCAACAT CGACAAGGTGCTGAAGAGCCTGGGCTTTCTGGAGAGCGGATCCGGATC TGGAGGAGGCACCCTGAAGTATGTGGAGGATGTGACAAATGTGGTGCG GAGAGATGTGGAGAAGTGGGGCCCCTTCGATCTGGTGTACGGATCCAC CCAGCCACTGGGAAGCTCCTGCGATAGGTGTCCAGGATGGTATATGTTC CAGTTTCACAGAATCCTGCAGTACGCACTGCCAAGGCAGGAGAGCCAG CGCCCTTTCTTTTGGATCTTTATGGACAACCTGCTGCTGACAGAGGATG ACCAGGAGACAACAACCCGCTTCCTGCAGACAGAGGCAGTGACCCTGC AGGATGTGAGGGGACGCGACTATCAGAATGCCATGCGGGTGTGGTCTA ACATCCCTGGCCTGAAGAGCAAGCACGCCCCCCTGACCCCTAAGGAGG AGGAGTACCTGCAGGCCCAGGTGCGGAGCAGATCCAAGCTGGATGCCC CTAAGGTGGACCTGCTGGTGAAGAATTGTCTGCTGCCACTGCGGGAGTA CTTCAAGTACTTTAGTCAGAATAGCCTGCCACTGGGAGGGCCGAGCTCT GGCGCACCCCCACCAAGTGGAGGGTCTCCTGCCGGGTCCCCAACATCTA CTGAAGAAGGCACCAGCGAATCCGCAACGCCCGAGTCAGGCCCTGGTA CCTCCACAGAACCATCTGAAGGTAGTGCGCCTGGTTCCCCAGCTGGAAG CCCTACTTCCACCGAAGAAGGCACGTCAACCGAACCAAGTGAAGGATC TGCCCCTGGGACCAGCACTGAACCATCTGAGGTTAACCCCAAGAAGAA GCGGAAGGTGGGCATCCACGGCGTGCCCGCCGCCGACAAGAAGTACAG CATCGGCCTGGCCATCGGCACCAACAGCGTGGGCTGGGCCGTGATCAC CGACGAGTACAAGGTGCCCAGCAAGAAGTTCAAGGTGCTGGGCAACAC CGACCGGCACAGCATCAAGAAGAACCTGATCGGCGCCCTGCTGTTCGA CAGCGGCGAGACCGCCGAGGCCACCCGGCTGAAGCGGACCGCCCGGCG GCGGTACACCCGGCGGAAGAACCGGATCTGCTACCTGCAGGAGATCTT CAGCAACGAGATGGCCAAGGTGGACGACAGCTTCTTCCACCGGCTGGA GGAGAGCTTCCTGGTGGAGGAGGACAAGAAGCACGAGCGGCACCCCAT CTTCGGCAACATCGTGGACGAGGTGGCCTACCACGAGAAGTACCCCAC CATCTACCACCTGCGGAAGAAGCTGGTGGACAGCACCGACAAGGCCGA CCTGCGGCTGATCTACCTGGCCCTGGCCCACATGATCAAGTTCCGGGGC CACTTCCTGATCGAGGGCGACCTGAACCCCGACAACAGCGACGTGGAC AAGCTGTTCATCCAGCTGGTGCAGACCTACAACCAGCTGTTCGAGGAGA ACCCCATCAACGCCAGCGGCGTGGACGCCAAGGCCATCCTGAGCGCCC GGCTGAGCAAGAGCCGGCGGCTGGAGAACCTGATCGCCCAGCTGCCCG GCGAGAAGAAGAACGGCCTGTTCGGCAACCTGATCGCCCTGAGCCTGG	

[1321]

	GCCTGACCCCCAACTTCAAGAGCAACTTCGACCTGGCCGAGGACGCCA AGCTGCAGCTGAGCAAGGACACCTACGACGACGACCTGGACAACCTGC TGGCCCAGATCGGCGACCAGTACGCCGACCTGTTCCTGGCCGCCAAGA ACCTGAGCGACGCCATCCTGCTGAGCGACATCCTGCGGGTGAACACCG AGATCACCAAGGCCCCCCTGAGCGCCAGCATGATCAAGCGGTACGACG AGCACCACCAGGACCTGACCCTGCTGAAGGCCCTGGTGCGGCAGCAGC TGCCCGAGAAGTACAAGGAGATCTTCTTCGACCAGAGCAAGAACGGCT ACGCCGGCTACATCGACGGCGGCGCCAGCCAGGAGGAGTTCTACAAGT TCATCAAGCCCATCCTGGAGAAGATGGACGGCACCGAGGAGCTGCTGG TGAAGCTGAACCGGGAGGACCTGCTGCGGAAGCAGCGGACCTTCGACA ACGGCAGCATCCCCCACCAGATCCACCTGGGCGAGCTGCACGCCATCCT GCGGCGGCAGGAGGACTTCTACCCCTTCCTGAAGGACAACCGGGAGAA GATCGAGAAGATCCTGACCTTCCGGATCCCCTACTACGTGGGCCCCCTG GCCCGGGGCAACAGCCGGTTCGCCTGGATGACCCGGAAGAGCGAGGAG ACCATCACCCCCTGGAACTTCGAGGAGGTGGTGGACAAGGGCGCCAGC GCCCAGAGCTTCATCGAGCGGATGACCAACTTCGACAAGAACCTGCCC AACGAGAAGGTGCTGCCCAAGCACAGCCTGCTGTACGAGTACTTCACC GTGTACAACGAGCTGACCAAGGTGAAGTACGTGACCGAGGGCATGCGG AAGCCCGCCTTCCTGAGCGGCGAGCAGAAGAAGGCCATCGTGGACCTG CTGTTCAAGACCAACCGGAAGGTGACCGTGAAGCAGCTGAAGGAGGAC TACTTCAAGAAGATCGAGTGCTTCGACAGCGTGGAGATCAGCGGCGTG GAGGACCGGTTCAACGCCAGCCTGGGCACCTACCACGACCTGCTGAAG ATCATCAAGGACAAGGACTTCCTGGACAACGAGGAGAACGAGGACATC CTGGAGGACATCGTGCTGACCCTGACCCTGTTCGAGGACCGGGAGATG ATCGAGGAGCGGCTGAAGACCTACGCCCACCTGTTCGACGACAAGGTG ATGAAGCAGCTGAAGCGGCGGCGGTACACCGGCTGGGGCCGGCTGAGC CGGAAGCTGATCAACGGCATCCGGGACAAGCAGAGCGGCAAGACCATC CTGGACTTCCTGAAGAGCGACGGCTTCGCCAACCGGAACTTCATGCAGC TGATCCACGACGACAGCCTGACCTTCAAGGAGGACATCCAGAAGGCCC AGGTGAGCGGCCAGGGCGACAGCCTGCACGAGCACATCGCCAACCTGG CCGGCAGCCCCGCCATCAAGAAGGGCATCCTGCAGACCGTGAAGGTGG TGGACGAGCTGGTGAAGGTGATGGGCCGGCACAAGCCCGAGAACATCG TGATCGAGATGGCCCGGGAGAACCAGACCACCCAGAAGGGCCAGAAG AACAGCCGGGAGCGGATGAAGCGGATCGAGGAGGGCATCAAGGAGCT GGGCAGCCAGATCCTGAAGGAGCACCCCGTGGAGAACACCCAGCTGCA GAACGAGAAGCTGTACCTGTACTACCTGCAGAACGGCCGGGACATGTA CGTGGACCAGGAGCTGGACATCAACCGGCTGAGCGACTACGACGTGGA CGCCATCGTGCCCCAGAGCTTCCTGAAGGACGACAGCATCGACAACAA GGTGCTGACCCGGAGCGACAAGAACCGGGGCAAGAGCGACAACGTGCC	

[1322]

	CAGCGAGGAGGTGGTGAAGAAGATGAAGAACTACTGGCGGCAGCTGCT GAACGCCAAGCTGATCACCCAGCGGAAGTTCGACAACCTGACCAAGGC CGAGCGGGGCGGCCTGAGCGAGCTGGACAAGGCCGGCTTCATCAAGCG GCAGCTGGTGGAGACCCGGCAGATCACCAAGCACGTGGCCCAGATCCT GGACAGCCGGATGAACACCAAGTACGACGAGAACGACAAGCTGATCCG GGAGGTGAAGGTGATCACCCTGAAGAGCAAGCTGGTGAGCGACTTCCG GAAGGACTTCCAGTTCTACAAGGTGCGGGAGATCAACAACTACCACCA CGCCCACGACGCCTACCTGAACGCCGTGGTGGGCACCGCCCTGATCAA GAAGTACCCCAAGCTGGAGAGCGAGTTCGTGTACGGCGACTACAAGGT GTACGACGTGCGGAAGATGATCGCCAAGAGCGAGCAGGAGATCGGCAA GGCCACCGCCAAGTACTTCTTCTACAGCAACATCATGAACTTCTTCAAG ACCGAGATCACCCTGGCCAACGGCGAGATCCGGAAGCGGCCCCTGATC GAGACCAACGGCGAGACCGGCGAGATCGTGTGGGACAAGGGCCGGGA CTTCGCCACCGTGCGGAAGGTGCTGAGCATGCCCCAGGTGAACATCGTG AAGAAGACCGAGGTGCAGACCGGCGGCTTCAGCAAGGAGAGCATCCTG CCCAAGCGGAACAGCGACAAGCTGATCGCCCGGAAGAAGGACTGGGAC CCCAAGAAGTACGGCGGCTTCGACAGCCCCACCGTGGCCTACAGCGTG CTGGTGGTGGCCAAGGTGGAGAAGGGCAAGAGCAAGAAGCTGAAGAG CGTGAAGGAGCTGCTGGGCATCACCATCATGGAGCGGAGCAGCTTCGA GAAGAACCCCATCGACTTCCTGGAGGCCAAGGGCTACAAGGAGGTGAA GAAGGACCTGATCATCAAGCTGCCCAAGTACAGCCTGTTCGAGCTGGA GAACGGCCGGAAGCGGATGCTGGCCAGCGCCGGCGAGCTGCAGAAGG GCAACGAGCTGGCCCTGCCCAGCAAGTACGTGAACTTCCTGTACCTGGC CAGCCACTACGAGAAGCTGAAGGGCAGCCCCGAGGACAACGAGCAGA AGCAGCTGTTCGTGGAGCAGCACAAGCACTACCTGGACGAGATCATCG AGCAGATCAGCGAGTTCAGCAAGCGGGTGATCCTGGCCGACGCCAACC TGGACAAGGTGCTGAGCGCCTACAACAAGCACCGGGACAAGCCCATCC GGGAGCAGGCCGAGAACATCATCCACCTGTTCACCCTGACCAACCTGG GCGCCCCCGCCGCCTTCAAGTACTTCGACACCACCATCGACCGGAAGCG GTACACCAGCACCAAGGAGGTGCTGGACGCCACCCTGATCCACCAGAG CATCACCGGCCTGTACGAGACCCGGATCGACCTGAGCCAGCTGGGCGG CGACAGCGGCGGCAAGCGGCCCGCCGCCACCAAGAAGGCCGGCCAGGC CAAGAAGAAGAAGGctagCgatgctaagtcactgactgcctggtcccggacactggtgaccttcaagga tgtgtttgtggacttcaccagggaggagtggaagctgctggacactgctcagcagatcctgtacagaaatgtgatgctg gagaactataagaacctggtttccttgggttatcagcttactaagccagatgtgatcctccggttggagaagggagaaga gccctggctggtggagagagaaattcaccaagagacccatcctgattcagagactgcatttgaaatcaaatcatcagtt CCGAAAAAGAAACGCAAAGTTTAG	

[1323]

647	MNHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIA SEVCEDSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCND LSIVNPARKGLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMG VSDKRDISRFLESNPVMIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDK LELQECLEHGRIAKFSKVRTITTRSNSIKQGKDQHFPVFMNEKEDILWCTEM ERVFGFPVHYTDVSNMSRLARQRLLGRSWSVPVIRHLFAPLKEYFACVSSG NSNANSRGPSFSSGLVPLSLRGSHMGPMEIYKTVSAWKRQPVRVLSLFRNI DKVLKSLGFLESGSGSGGGTLKYVEDVTNVVRRDVEKWGPFDLVYGSTQP LGSSCDRCPGWYMFQFHRILQYALPRQESQRPFFWIFMDNLLLTEDDQETT TRFLQTEAVTLQDVRGRDYQNAMRVWSNIPGLKSKHAPLTPKEEEYLQAQ VRSRSKLDAPKVDLLVKNCLLPLREYFKYFSQNSLPLGGPSSGAPPPSGGSP AGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEG SAPGTSTEPSEVNPKKKRKVGIHGVPAADKKYSIGLAIGTNSVGWAVITDE YKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRR KNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVA YHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDN SDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLP GEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLA QIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDL TLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMD GTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDN REKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASA QSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPA FLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNAS LGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLF DDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNF MQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVV DELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQI LKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQS FLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQR KFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEN DKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTA LIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFK TEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTE VQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKV EKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYS LFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDN EQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIRE QAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYE	DNMT3A/L- XTEN80-dSpCas9- KRAB (AA)

[1324]

	TRIDLSQLGGDSGGKRPAATKKAGQAKKKKASDAKSLTAWSRTLVTFKD VFVDFTREEWKLLDTAQQILYRNVMLENYKNLVSLGYQLTKPDVILRLEK GEEPWLVEREIHQETHPDSETAFEIKSSVPKKKRKV	
648	AGTAACCATGACCAGGAATTTGACCCCCCAAAGGTTTACCCACCTGTGC CAGCTGAGAAGAGGAAGCCCATCCGCGTGCTGTCTCTCTTTGATGGGAT TGCTACAGGGCTCCTGGTGCTGAAGGACCTGGGCATCCAAGTGGACCG CTACATTGCCTCCGAGGTGTGTGAGGACTCCATCACGGTGGGCATGGTG CGGCACCAGGGAAAGATCATGTACGTCGGGGACGTCCGCAGCGTCACA CAGAAGCATATCCAGGAGTGGGGCCCATTCGACCTGGTGATTGGAGGC AGTCCCTGCAATGACCTCTCCATTGTCAACCCTGCCCGCAAGGGACTTT ATGAGGGTACTGGCCGCCTCTTCTTTGAGTTCTACCGCCTCCTGCATGAT GCGCGGCCCAAGGAGGGAGATGATCGCCCCTTCTTCTGGCTCTTTGAGA ATGTGGTGGCCATGGGCGTTAGTGACAAGAGGGACATCTCGCGATTTCT TGAGTCTAACCCCGTGATGATTGACGCCAAAGAAGTGTCTGCTGCACAC AGGGCCCGTTACTTCTGGGGTAACCTTCCTGGCATGAACAGGCCTTTGG CATCCACTGTGAATGATAAGCTGGAGCTGCAAGAGTGTCTGGAGCACG GCAGAATAGCCAAGTTCAGCAAAGTGAGGACCATTACCACCAGGTCAA ACTCTATAAAGCAGGGCAAAGACCAGCATTTCCCCGTCTTCATGAACGA GAAGGAGGACATCCTGTGGTGCACTGAAATGGAAAGGGTGTTTGGCTT CCCCGTCCACTACACAGACGTCTCCAACATGAGCCGCTTGGCGAGGCAG AGACTGCTGGGCCGATCGTGGAGCGTGCCGGTCATCCGCCACCTCTTCG CTCCGCTGAAGGAATATTTTGCTTGTGTGTCTAGCGGCAATAGTAACGC TAACAGCCGCGGGCCGAGCTTCAGCAGCGGCCTGGTGCCGTTAAGCTTG CGCGGCAGCCATATGGGCCCTATGGAGATATACAAGACAGTGTCTGCA TGGAAGAGACAGCCAGTGCGGGTACTGAGCCTCTTCAGAAACATCGAC AAGGTACTAAAGAGTTTGGGCTTCTTGGAAAGCGGTTCTGGTTCTGGGG GAGGAACGCTGAAGTACGTGGAAGATGTCACAAATGTCGTGAGGAGAG ACGTGGAGAAATGGGGCCCCTTTGACCTGGTGTACGGCTCGACGCAGC CCCTAGGCAGCTCTTGTGATCGCTGTCCCGGCTGGTACATGTTCCAGTTC CACCGGATCCTGCAGTATGCGCTGCCTCGCCAGGAGAGTCAGCGGCCCT TCTTCTGGATATTCATGGACAATCTGCTGCTGACTGAGGATGACCAAGA GACAACTACCCGCTTCCTTCAGACAGAGGCTGTGACCCTCCAGGATGTC CGTGGCAGAGACTACCAGAATGCTATGCGGGTGTGGAGCAACATTCCA GGGCTGAAGAGCAAGCATGCGCCCCTGACCCCAAAGGAAGAAGAGTAT CTGCAAGCCCAAGTCAGAAGCAGGAGCAAGCTGGACGCCCCGAAAGTT GACCTCCTGGTGAAGAACTGCCTTCTCCCGCTGAGAGAGTACTTCAAGT ATTTTTCTCAAAACTCACTTCCTCTTGGAGGGCCGAGCTCTGGCGCACC CCCACCAAGTGGAGGGTCTCCTGCCGGGTCCCCAACATCTACTGAAGAA GGCACCAGCGAATCCGCAACGCCCGAGTCAGGCCCTGGTACCTCCACA	DNMT3A/L(CRISPROFF)-XTEN80-dSpCas9-KRAB (nt)

[1325]

	GAACCATCTGAAGGTAGTGCGCCTGGTTCCCCAGCTGGAAGCCCTACTT CCACCGAAGAAGGCACGTCAACCGAACCAAGTGAAGGATCTGCCCCTG GGACCAGCACTGAACCATCTGAGGTTAACCCCAAGAAGAAGCGGAAGG TGGGCATCCACGGCGTGCCCGCCGCCGACAAGAAGTACAGCATCGGCC TGGCCATCGGCACCAACAGCGTGGGCTGGGCCGTGATCACCGACGAGT ACAAGGTGCCCAGCAAGAAGTTCAAGGTGCTGGGCAACACCGACCGGC ACAGCATCAAGAAGAACCTGATCGGCGCCCTGCTGTTCGACAGCGGCG AGACCGCCGAGGCCACCCGGCTGAAGCGGACCGCCCGGCGGCGGTACA CCCGGCGGAAGAACCGGATCTGCTACCTGCAGGAGATCTTCAGCAACG AGATGGCCAAGGTGGACGACAGCTTCTTCCACCGGCTGGAGGAGAGCT TCCTGGTGGAGGAGGACAAGAAGCACGAGCGGCACCCCATCTTCGGCA ACATCGTGGACGAGGTGGCCTACCACGAGAAGTACCCCACCATCTACC ACCTGCGGAAGAAGCTGGTGGACAGCACCGACAAGGCCGACCTGCGGC TGATCTACCTGGCCCTGGCCCACATGATCAAGTTCCGGGGCCACTTCCT GATCGAGGGCGACCTGAACCCCGACAACAGCGACGTGGACAAGCTGTT CATCCAGCTGGTGCAGACCTACAACCAGCTGTTCGAGGAGAACCCCATC AACGCCAGCGGCGTGGACGCCAAGGCCATCCTGAGCGCCCGGCTGAGC AAGAGCCGGCGGCTGGAGAACCTGATCGCCCAGCTGCCCGGCGAGAAG AAGAACGGCCTGTTCGGCAACCTGATCGCCCTGAGCCTGGGCCTGACCC CCAACTTCAAGAGCAACTTCGACCTGGCCGAGGACGCCAAGCTGCAGC TGAGCAAGGACACCTACGACGACGACCTGGACAACCTGCTGGCCCAGA TCGGCGACCAGTACGCCGACCTGTTCCTGGCCGCCAAGAACCTGAGCG ACGCCATCCTGCTGAGCGACATCCTGCGGGTGAACACCGAGATCACCA AGGCCCCCCTGAGCGCCAGCATGATCAAGCGGTACGACGAGCACCACC AGGACCTGACCCTGCTGAAGGCCCTGGTGCGGCAGCAGCTGCCCGAGA AGTACAAGGAGATCTTCTTCGACCAGAGCAAGAACGGCTACGCCGGCT ACATCGACGGCGGCGCCAGCCAGGAGGAGTTCTACAAGTTCATCAAGC CCATCCTGGAGAAGATGGACGGCACCGAGGAGCTGCTGGTGAAGCTGA ACCGGGAGGACCTGCTGCGGAAGCAGCGGACCTTCGACAACGGCAGCA TCCCCCACCAGATCCACCTGGGCGAGCTGCACGCCATCCTGCGGCGGCA GGAGGACTTCTACCCCTTCCTGAAGGACAACCGGGAGAAGATCGAGAA GATCCTGACCTTCCGGATCCCCTACTACGTGGGCCCCCTGGCCCGGGGC AACAGCCGGTTCGCCTGGATGACCCGGAAGAGCGAGGAGACCATCACC CCCTGGAACTTCGAGGAGGTGGTGGACAAGGGCGCCAGCGCCCAGAGC TTCATCGAGCGGATGACCAACTTCGACAAGAACCTGCCCAACGAGAAG GTGCTGCCCAAGCACAGCCTGCTGTACGAGTACTTCACCGTGTACAACG AGCTGACCAAGGTGAAGTACGTGACCGAGGGCATGCGGAAGCCCGCCT TCCTGAGCGGCGAGCAGAAGAAGGCCATCGTGGACCTGCTGTTCAAGA CCAACCGGAAGGTGACCGTGAAGCAGCTGAAGGAGGACTACTTCAAGA	

[1326]

	AGATCGAGTGCTTCGACAGCGTGGAGATCAGCGGCGTGGAGGACCGGT TCAACGCCAGCCTGGGCACCTACCACGACCTGCTGAAGATCATCAAGG ACAAGGACTTCCTGGACAACGAGGAGAACGAGGACATCCTGGAGGACA TCGTGCTGACCCTGACCCTGTTCGAGGACCGGGAGATGATCGAGGAGC GGCTGAAGACCTACGCCCACCTGTTCGACGACAAGGTGATGAAGCAGC TGAAGCGGCGGCGGTACACCGGCTGGGGCCGGCTGAGCCGGAAGCTGA TCAACGGCATCCGGGACAAGCAGAGCGGCAAGACCATCCTGGACTTCC TGAAGAGCGACGGCTTCGCCAACCGGAACTTCATGCAGCTGATCCACG ACGACAGCCTGACCTTCAAGGAGGACATCCAGAAGGCCCAGGTGAGCG GCCAGGGCGACAGCCTGCACGAGCACATCGCCAACCTGGCCGGCAGCC CCGCCATCAAGAAGGGCATCCTGCAGACCGTGAAGGTGGTGGACGAGC TGGTGAAGGTGATGGGCCGGCACAAGCCCGAGAACATCGTGATCGAGA TGGCCCGGGAGAACCAGACCACCCAGAAGGGCCAGAAGAACAGCCGG GAGCGGATGAAGCGGATCGAGGAGGGCATCAAGGAGCTGGGCAGCCA GATCCTGAAGGAGCACCCCGTGGAGAACACCCAGCTGCAGAACGAGAA GCTGTACCTGTACTACCTGCAGAACGGCCGGGACATGTACGTGGACCA GGAGCTGGACATCAACCGGCTGAGCGACTACGACGTGGACGCCATCGT GCCCCAGAGCTTCCTGAAGGACGACAGCATCGACAACAAGGTGCTGAC CCGGAGCGACAAGAACCGGGGCAAGAGCGACAACGTGCCCAGCGAGG AGGTGGTGAAGAAGATGAAGAACTACTGGCGGCAGCTGCTGAACGCCA AGCTGATCACCCAGCGGAAGTTCGACAACCTGACCAAGGCCGAGCGGG GCGGCCTGAGCGAGCTGGACAAGGCCGGCTTCATCAAGCGGCAGCTGG TGGAGACCCGGCAGATCACCAAGCACGTGGCCCAGATCCTGGACAGCC GGATGAACACCAAGTACGACGAGAACGACAAGCTGATCCGGGAGGTGA AGGTGATCACCCTGAAGAGCAAGCTGGTGAGCGACTTCCGGAAGGACT TCCAGTTCTACAAGGTGCGGGAGATCAACAACTACCACCACGCCCACG ACGCCTACCTGAACGCCGTGGTGGGCACCGCCCTGATCAAGAAGTACC CCAAGCTGGAGAGCGAGTTCGTGTACGGCGACTACAAGGTGTACGACG TGCGGAAGATGATCGCCAAGAGCGAGCAGGAGATCGGCAAGGCCACCG CCAAGTACTTCTTCTACAGCAACATCATGAACTTCTTCAAGACCGAGAT CACCCTGGCCAACGGCGAGATCCGGAAGCGGCCCCTGATCGAGACCAA CGGCGAGACCGGCGAGATCGTGTGGGACAAGGGCCGGGACTTCGCCAC CGTGCGGAAGGTGCTGAGCATGCCCCAGGTGAACATCGTGAAGAAGAC CGAGGTGCAGACCGGCGGCTTCAGCAAGGAGAGCATCCTGCCCAAGCG GAACAGCGACAAGCTGATCGCCCGGAAGAAGGACTGGGACCCCAAGA AGTACGGCGGCTTCGACAGCCCCACCGTGGCCTACAGCGTGCTGGTGGT GGCCAAGGTGGAGAAGGGCAAGAGCAAGAAGCTGAAGAGCGTGAAGG AGCTGCTGGGCATCACCATCATGGAGCGGAGCAGCTTCGAGAAGAACC CCATCGACTTCCTGGAGGCCAAGGGCTACAAGGAGGTGAAGAAGGACC	

[1327]

	TGATCATCAAGCTGCCCAAGTACAGCCTGTTCGAGCTGGAGAACGGCC GGAAGCGGATGCTGGCCAGCGCCGGCGAGCTGCAGAAGGGCAACGAG CTGGCCCTGCCCAGCAAGTACGTGAACTTCCTGTACCTGGCCAGCCACT ACGAGAAGCTGAAGGGCAGCCCCGAGGACAACGAGCAGAAGCAGCTG TTCGTGGAGCAGCACAAGCACTACCTGGACGAGATCATCGAGCAGATC AGCGAGTTCAGCAAGCGGGTGATCCTGGCCGACGCCAACCTGGACAAG GTGCTGAGCGCCTACAACAAGCACCGGGACAAGCCCATCCGGGAGCAG GCCGAGAACATCATCCACCTGTTCACCCTGACCAACCTGGGCGCCCCCG CCGCCTTCAAGTACTTCGACACCACCATCGACCGGAAGCGGTACACCAG CACCAAGGAGGTGCTGGACGCCACCCTGATCCACCAGAGCATCACCGG CCTGTACGAGACCCGGATCGACCTGAGCCAGCTGGGCGGCGACAGCGG CGGCAAGCGGCCCGCCGCCACCAAGAAGGCCGGCCAGGCCAAGAAGA AGAAGGctagCgatgctaagtcactgactgcctggtcccggacactggtgaccttcaaggatgtgtttgtggactt caccagggaggagtggaagctgctggacactgctcagcagatcctgtacagaaatgtgatgctggagaactataaga acctggtttccttgggttatcagcttactaagccagatgtgatcctccggttggagaagggagaagagccctggctggtg gagagagaaattcaccaagagacccatcctgattcagagactgcatttgaaatcaaatcatcagttCCGAAAAA GAAACGCAAAGTTTAG	
649	SNHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIA SEVCEDSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCND LSIVNPARKGLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMG VSDKRDISRFLESNPVMIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDK LELQECLEHGRIAKFSKVRTITTRSNSIKQGKDQHFPVFMNEKEDILWCTEM ERVFGFPVHYTDVSNMSRLARQRLLGRSWSVPVIRHLFAPLKEYFACVSSG NSNANSRGPSFSSGLVPLSLRGSHMGPMEIYKTVSAWKRQPVRVLSLFRNI DKVLKSLGFLESGSGSGGGTLKYVEDVTNVVRRDVEKWGPFDLVYGSTQP LGSSCDRCPGWYMFQFHRILQYALPRQESQRPFFWIFMDNLLLTEDDQETT TRFLQTEAVTLQDVRGRDYQNAMRVWSNIPGLKSKHAPLTPKEEEYLQAQ VRSRSKLDAPKVDLLVKNCLLPLREYFKYFSQNSLPLGGPSSGAPPPSGGSP AGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEG SAPGTSTEPSEVNPKKKRKVGIHGVPAADKKYSIGLAIGTNSVGWAVITDE YKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRR KNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVA YHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDN SDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLP GEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLA QIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDL TLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMD GTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDN REKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASA	DNMT3A/L(CRISPROFF)-XTEN80-dSpCas9-KRAB(AA)

[1328]

	QSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPA FLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNAS LGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLF DDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNF MQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVV DELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQI LKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQS FLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQR KFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEN DKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTA LIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFK TEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTE VQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKV EKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYS LFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDN EQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIRE QAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYE TRIDLSQLGGDSGGKRPAATKKAGQAKKKKASDAKSLTAWSRTLVTFKD VFVDFTREEWKLLDTAQQILYRNVMLENYKNLVSLGYQLTKPDVILRLEK GEEPWLVEREIHQETHPDSETAFEIKSSVPKKKRKV	

[1329]

650	aattccacaaccttcaccaaactctgcaagatcccagagtgagaggcctgtatttccct gctggtggctccagttcaggagcagtaaaccctgttccgactactgcctctccctatcg tcaatcttctcgaggattggggaccctgcgctgaacatggagaacatcacatcaggattc ctaggacccctttctcgtgttacaggcgggggtttttcttgttgacaagaatcctcacaata ccgcagagtctagactcgtggtggacttctctcaattttctagggggaactaccgtgtgt cttggccaaaattcgcagtccccaacctccaatcactcaccaacctcctgtcctccaact tgtcctggttatcgctggatgtgtctgcggcgttttatcatcttcctcttcatcctgctg ctatgcctcatcttcttgttggttcttctggactatcaaggtatgttgcccgtttgtcct ctaattccaggatcctcaaccaccagcacgggaccatgccgaacctgcatgactactgct caaggaacctctatgtatccctcctgttgctgtaccaaaccttcggacggaaattgcacc tgtattcccatcccatcatcctgggctttcggaaaattcctatgggagtgggcctcagcc cgtttctcctggctcagtttactagtgccatttgttcagtggttcgtagggctttccccc actgtttggctttcagttatatggatgatgtggtattgggggccaagtctgtacagcatc ttgagtccctttttaccgctgttaccaattttcttttgtctttgggtatacatttaaacc ctaacaaaacaaagagatggggttactctctgaattttatgggttatgtcattggaagtt atgggtccttgccacaagaacacatcatacaaaaaatcaaagaatgttttagaaaacttc ctattaacaggcctattgattggaaagtatgtcaacgaattgtgggtcttttgggttttg ctgccccatttacacaatgtggttatcctgcgttaatgcccttgtatgcatgtattcaat ctaagcaggctttcactttctcgccaacttacaaggcctttctgtgtaaacaatacctga acctttaccccgttgcccggcaacggccaggtctgtgccaagtgtttgctgacgcaaccc ccactggctggggcttggtcatgggccatcagcgcgtgcgtggaacctttcggctcctc tgccgatccatactgcggaactcctagccgcttgttttgctcgcagcaggtctggagcaa acattatcgggactgataactctgttgtcctctcccgcaaatatacatcgtatccatggc tgctaggctgtgctgccaactggatcctgcgcgggacgtcctttgtttacgtcccgtcgg cgctgaatcctgcggacgacccttctcggggtcgcttgggactctctcgtccccttctcc gtctgccgttccgaccgaccacggggcgcacctctctttacgcggactccccgtctgtgc cttctcatctgccggaccgtgtgcacttcgcttcacctctgcacgtcgcatggagaccac cgtgaacgcccaccgaatgttgcccaaggtcttacataagaggactcttggactctctgc aatgtcaacgaccgaccttgaggcatacttcaaagactgtttgtttaaagactgggagga gttgggggaggagattagattaaaggtctttgtactaggaggctgtaggcataaattggt	B 형 간염 바이러스 게놈 (B 형 간염 바이러스 서브타입 ayw, 완전 게놈, GenBank: U95551.1),

[1330]

	ctgcgcaccagcaccatgcaacttttcacctctgcctaatcatctcttgttcat gtcct actgttcaagcctccaagctgtgccttgggtggctttggggcatggacatcgacc cttat aaagaatttggagctactgtggagttactctcgtttttgccttctgacttctttc cttca gtacgagatcttctagataccgcctcagctctgtatcgggaagccttagagtctc ctgag cattgttcacctcaccatactgcactcaggcaagcaattctttgctgggggggaac taatg actctagctacctgggtgggtgttaatttggaagatccagcatctagagacctag tagtc agttatgtcaacactaatatgggcctaaagttcaggcaactcttgtggtttcaca tttct tgtctcacttttggaagagaaaccgttatagagtatttggtgtctttcggagtgt ggatt cgcactcctccagcttatagaccaccaaatgcccctatcctatcaacacttccgg aaact actgttgttagacgacgaggcaggtcccctagaagaagaactccctcgcctcgca gacga aggtctcaatcgccgcgtcgcagaagatctcaatctcgggaacctcaatgttagt attcc ttggactcataaggtggggaactttactggtctttattcttctactgtacctgtc tttaa tcctcattggaaaacaccatcttttcctaatatacatttacaccaagacattatc aaaaa atgtgaacagtttgtaggcccacttacagttaatgagaaaagaagattgcaattg attat gcctgctaggttttatccaaaggttaccaaatatttaccattggataagggtatt aaacc ttattatccagaacatctagttaatcattacttccaaactagacactatttacac actct atggaaggcgggtatattatataagagagaaacaacacatagcgcctcattttgt gggtc accatattcttgggaacaagatctacagcatggggcagaatctttccaccagcaa tcctc tgggattctttcccgaccaccagttggatccagccttcagagcaaacacagcaaa tccag attgggacttcaatcccaacaaggacacctggccagacgccaacaaggtaggagc tggag cattcgggctgggtttcaccccaccgcacggaggccttttggggtggagccctca ggctc agggcatactacaaactttgccagcaaatccgcctcctgcctccaccaatcgcca gacag gaaggcagcctaccccgctgtctccacctttgagaaacactcatcctcaggccat gcagt gg	

[1331]

651	NHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIAS EVCEDSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDL SIVNPARKGLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGV SDKRDISRFLESNPVMIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKL ELQECLEHGRIAKFSKVRTITTRSNSIKQGKDQHFPVFMNEKEDILWCTEME RVFGFPVHYTDVSNMSRLARQRLLGRSWSVPVIRHLFAPLKEYFACVSSGN SNANSRGPSFSSGLVPLSLRGSHMGPMEIYKTVSAWKRQPVRVLSLFRNID KVLKSLGFLESGSGSGGGTLKYVEDVTNVVRRDVEKWGPFDLVYGSTQPL GSSCDRCPGWYMFQFHRILQYALPRQESQRPFFWIFMDNLLLTEDDQETTT RFLQTEAVTLQDVRGRDYQNAMRVWSNIPGLKSKHAPLTPKEEEYLQAQV RSRSKLDAPKVDLLVKNCLLPLREYFKYFSQNSLPL	DNMT3A/L (AA)
652	TGCCTGTCCTACGAGACAGAGATCCTGACAGTGGAGTATGGCCTGCTGC CAATCGGCAAGATCGTGGAGAAGAGGATCGAGTGTACCGTGTACTCTG TGGATAACAATGGCAACATCTATACACAGCCCGTGGCACAGTGGCACG ATAGGGGAGAGCAGGAGGTGTTCGAGTATTGCCTGGAGGACGGCAGCC TGATCAGGGCAACCAAGGACCACAAGTTCATGACAGTGGATGGCCAGA TGCTGCCCATCGAC	N term Npu Intein (nt)
653	CLSYETEILTVEYGLLPIGKIVEKRIECTVYSVDNNGNIYTQPVAQWHDRGE QEVFEYCLEDGSLIRATKDHKFMTVDGQMLPID	N term Npu Intein (aa)
654	ATCAAGATTGCTACACGGAAATACCTGGGAAAGCAGAACGTGTACGAC ATCGGCGTGGAGCGGGATCACAACTTCGCCCTGAAGAATGGCTTTATCG CCAGCAAT	C term Npu Intein (nt)
655	IKIATRKYLGKQNVYDIGVERDHNFALKNGFIASN	C term Npu Intein (aa)
656	ATGAAACGGACAGCCGACGGAAGCGAGTTCGAGTCACCAAAGAAGAA GCGGAAAGTCATCAAGATTGCTACACGGAAATACCTGGGAAAGCAGAA	dSpCas9-573-C Intein (nt)

[1332]

	CGTGTACGACATCGGCGTGGAGCGGGATCACAACTTCGCCCTGAAGAA TGGCTTTATCGCCAGCAATTGTTTCGACTCTGTTGAAATCAGCGGAGTG GAGGATCGCTTCAACGCATCCCTGGGAACGTATCACGATCTCCTGAAAA TCATTAAAGACAAGGACTTCCTGGACAATGAGGAGAACGAGGACATTC TTGAGGACATTGTCCTCACCCTTACGTTGTTTGAAGATAGGGAGATGAT TGAAGAACGCTTGAAAACTTACGCTCATCTCTTCGACGACAAAGTCATG AAACAGCTCAAGAGGCGCCGATATACAGGATGGGGGCGGCTGTCAAGA AAACTGATCAATGGgatcCGAGACAAGCAGAGTGGAAAGACAATCCTGG ATTTTCTTAAGTCCGATGGATTTGCCAACCGGAACTTCATGCAGTTGAT CCATGATGACTCTCTCACCTTTAAGGAGGACATCCAGAAAGCACAAGTT TCTGGCCAGGGGGACAGTCTTCACGAGCACATCGCTAATCTTGCAGGTA GCCCAGCTATCAAAAAGGGAATACTGCAGACCGTTAAGGTCGTGGATG AACTCGTCAAAGTAATGGGAAGGCATAAGCCCGAGAATATCGTTATCG AGATGGCCCGAGAGAACCAAACTACCCAGAAGGGACAGAAGAACAGT AGGGAAAGGATGAAGAGGATTGAAGAGGGTATAAAAGAACTGGGGTC CCAAATCCTTAAGGAACACCCAGTTGAAAACACCCAGCTTCAGAATGA GAAGCTCTACCTGTACTACCTGCAGAACGGCAGGGACATGTACGTGGA TCAGGAACTGGACATCAATCGGCTCTCCGACTACGACGTGGATGCCATC GTGCCCCAGTCTTTTCTCAAAGATGATTCTATTGATAATAAAGTGTTGA CAAGATCCGATAAAAATAGAGGGAAGAGTGATAACGTCCCCTCAGAAG AAGTTGTCAAGAAAATGAAAAATTATTGGCGGCAGCTGCTGAACGCCA AACTGATCACACAACGGAAGTTCGATAATCTGACTAAGGCTGAACGAG GTGGCCTGTCTGAGTTGGATAAAGCCGGCTTCATCAAAAGGCAGCTTGT TGAGACACGCCAGATCACCAAgcacGTGGCCCAAATTCTCGATTCACGCA TGAACACCAAGTACGATGAAAATGACAAACTGATTCGAGAGGTGAAAG TTATTACTCTGAAGTCTAAGCTGGTCTCAGATTTCAGAAAGGACTTTCA GTTTTATAAGGTGAGAGAGATCAACAATTACCACCATGCGCATGATGCC TACCTGAATGCAGTGGTAGGCACTGCACTTATCAAAAAATATCCCAAGC TTGAATCTGAATTTGTTTACGGAGACTATAAAGTGTACGATGTTAGGAA AATGATCGCAAAGTCTGAGCAGGAAATAGGCAAGGCCACCGCTAAGTA CTTCTTTTACAGCAATATTATGAATTTTTTCAAGACCGAGATTACACTGG CCAATGGAGAGATTCGGAAGCGACCACTTATCGAAACAAACGGAGAAA CAGGAGAAATCGTGTGGGACAAGGGTAGGGATTTCGCGACAGTCCGGA AGGTCCTGTCCATGCCGCAGGTGAACATCGTTAAAAAGACCGAAGTAC AGACCGGAGGCTTCTCCAAGGAAAGTATCCTCCCGAAAAGGAACAGCG ACAAGCTGATCGCACGCAAAAAAGATTGGGACCCCAAGAAATACGGCG GATTCGATTCTCCTACAGTCGCTTACAGTGTACTGGTTGTGGCCAAAGT GGAGAAAGGGAAGTCTAAAAAACTCAAAAGCGTCAAGGAACTGCTGG GCATCACAATCATGGAGCGATCAAGCTTCGAAAAAAACCCCATCGACT	

[1333]

	TTCTCGAGGCGAAAGGATATAAAGAGGTCAAAAAAGACCTCATCATTA AGCTTCCCAAGTACTCTCTCTTTGAGCTTGAAAACGGCCGGAAACGAAT GCTCGCTAGTGCGGGCGAGCTGCAGAAAGGTAACGAGCTGGCACTGCC CTCTAAATACGTTAATTTCTTGTATCTGGCCAGCCACTATGAAAAGCTC AAAGGGTCTCCCGAAGATAATGAGCAGAAGCAGCTGTTCGTGGAACAA CACAAACACTACCTTGATGAGATCATCGAGCAAATAAGCGAATTCTCCA AAAGAGTGATCCTCGCCGACGCTAACCTCGATAAGGTGCTTTCTGCTTA CAATAAGCACAGGGATAAGCCCATCAGGGAGCAGGCAGAAAACATTAT CCACTTGTTTACTCTGACCAACTTGGGCGCGCCTGCAGCCTTCAAGTAC TTCGACACCACCATAGACAGAAAGCGGTACACCTCTACAAAGGAGGTC CTGGACGCCACACTGATTCATCAGTCAATTACGGGGCTCTATGAAACAA GAATCGACCTCTCTCAGCTCGGTGGAGACAAAAGGCCGGCGGCCACGA AAAAGGCCGGCCAGGCAAAAAAGAAAAAGTGA	
657	MKRTADGSEFESPKKKRKVIKIATRKYLGKQNVYDIGVERDHNFALKNGFI ASNCFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTL TLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDK QSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHI ANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQ KNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV DQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEV VKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETR QITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVR EINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSE QEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGR DFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK KYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDF LEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKY VNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILAD ANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRY TSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRPAATKKAGQAKKKK	dSpCas9-573-C Intein (aa)
658	SGSETPGTSESATPES	링커(aa)
659	ATGGCGGCCATCCCAGCCCTGGACCCAGAGGCCGAGCCCAGCATGGAC GTGATTTTGGTGGGATCCAGTGAGCTCTCAAGCTCCGTTTCACCCGGGA CAGGCAGAGATCTTATTGCATATGAAGTCAAGGCTAACCAGCGAAATA TAGAAGACATCTGCATCTGCTGCGGAAGTCTCCAGGTTCACACACAGCA CCCTCTGTTTGAGGGAGGGATCTGCGCCCCATGTAAGGACAAGTTCCTG GATGCCCTCTTCCTGTACGACGATGACGGGTACCAATCCTACTGCTCCA TCTGCTGCTCCGGAGAAACGCTGCTCATCTGCGGAAACCCTGATTGCAC CCGATGCTACTGCTTCGAGTGTGTGGATAGCCTGGTCGGCCCCGGGACC	DNMT3L (nt)

[1334]

	TCGGGGAAGGTGCACGCCATGAGCAACTGGGTGTGCTACCTGTGCCTGC CGTCCTCCCGAAGCGGGCTGCTGCAGCGTCGGAGGAAGTGGCGCAGCC AGCTCAAGGCCTTCTACGACCGAGAGTCGGAGAATCCCCTTGAGATGTT CGAAACCGTGCCTGTGTGGAGGAGACAGCCAGTCCGGGTGCTGTCCCTT TTTGAAGACATCAAGAAAGAGCTGACGAGTTTGGGCTTTTTGGAAAGTG GTTCTGACCCGGGACAACTGAAGCATGTGGTTGATGTCACAGACACAGT GAGGAAGGATGTGGAGGAGTGGGGACCCTTCGATCTTGTGTACGGCGC CACACCTCCCCTGGGCCACACCTGTGACCGTCCTCCCAGCTGGTACCTG TTCCAGTTCCACCGGCTCCTGCAGTACGCACGGCCCAAGCCAGGCAGCC CCAGGCCCTTCTTCTGGATGTTCGTGGACAATCTGGTGCTGAACAAGGA AGACCTGGACGTCGCATCTCGCTTCCTGGAGATGGAGCCAGTCACCATC CCAGATGTCCACGGCGGATCCTTGCAGAATGCTGTCCGCGTGTGGAGCA ACATCCCAGCCATAAGGAGCAGGCACTGGGCTCTGGTTTCGGAAGAAG AATTGTCCCTGCTGGCCCAGAACAAGCAGAGCTCGAAGCTCGCGGCCA AGTGGCCCACCAAGCTGGTGAAGAACTGCTTTCTCCCCCTAAGAGAATA TTTCAAGTATTTTTCAACAGAACTCACTTCCTCTTTA	
660	ACCTACGGGCTGCTGCGGCGGCGAGAGGACTGGCCCTCCCGGCTCCAG ATGTTCTTCGCTAATAACCACGACCAGGAATTTGACCCTCCAAAGGTTT ACCCACCTGTCCCAGCTGAGAAGAGGAAGCCCATCCGGGTGCTGTCTCT CTTTGATGGAATCGCTACAGGGCTCCTGGTGCTGAAGGACTTGGGCATT CAGGTGGACCGCTACATTGCCTCGGAGGTGTGTGAGGACTCCATCACGG TGGGCATGGTGCGGCACCAGGGGAAGATCATGTACGTCGGGGACGTCC GCAGCGTCACACAGAAGCATATCCAGGAGTGGGGCCCATTCGATCTGG TGATTGGGGGCAGTCCCTGCAATGACCTCTCCATCGTCAACCCTGCTCG CAAGGGCCTCTACGAGGGCACTGGCCGGCTCTTCTTTGAGTTCTACCGC CTCCTGCATGATGCGCGGCCCAAGGAGGGAGATGATCGCCCCTTCTTCT GGCTCTTTGAGAATGTGGTGGCCATGGGCGTTAGTGACAAGAGGGACA TCTCGCGATTTCTCGAGTCCAACCCTGTGATGATTGATGCCAAAGAAGT GTCAGCTGCACACAGGGCCCGCTACTTCTGGGGTAACCTTCCCGGTATG AACAGGCCGTTGGCATCCACTGTGAATGATAAGCTGGAGCTGCAGGAG TGTCTGGAGCATGGCAGGATAGCCAAGTTCAGCAAAGTGAGGACCATT ACTACGAGGTCAAACTCCATAAAGCAGGGCAAAGACCAGCATTTTCCT GTCTTCATGAATGAGAAAGAGGACATCTTATGGTGCACTGAAATGGAA AGGGTATTTGGTTTCCCAGTCCACTATACTGACGTATCCAACATGAGCC GCTTGGCGAGGCAGAGACTGCTGGGCCGGTCATGGAGCGTGCCAGTCA TCCGCCACCTCTTCGCTCCGCTGAAGGAGTATTTTGCGTGTGTG	DNMT3A (nt)
661	TYGLLRRREDWPSRLQMFFANNHDQEFDPPKVYPPVPAEKRKPIRVLSLFD GIATGLLVLKDLGIQVDRYIASEVCEDSITVGMVRHQGKIMYVGDVRSVTQ KHIQEWGPFDLVIGGSPCNDLSIVNPARKGLYEGTGRLFFEFYRLLHDARPK	DNMT3A (AA)

[1335]

	EGDDRPFFWLFENVVAMGVSDKRDISRFLESNPVMIDAKEVSAAHRARYF WGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIKQGKD QHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSRLARQRLLGRSWSV PVIRHLFAPLKEYFACV	
662	AACCATGACCAGGAATTTGACCCCCCAAAGGTTTACCCACCTGTGCCAG CTGAGAAGAGGAAGCCCATCCGCGTGCTGTCTCTCTTTGATGGGATTGC TACAGGGCTCCTGGTGCTGAAGGACCTGGGCATCCAAGTGGACCGCTA CATTGCCTCCGAGGTGTGTGAGGACTCCATCACGGTGGGCATGGTGCGG CACCAGGGAAAGATCATGTACGTCGGGGACGTCCGCAGCGTCACACAG AAGCATATCCAGGAGTGGGGCCCATTCGACCTGGTGATTGGAGGCAGT CCCTGCAATGACCTCTCCATTGTCAACCCTGCCCGCAAGGGACTTTATG AGGGTACTGGCCGCCTCTTCTTTGAGTTCTACCGCCTCCTGCATGATGCG CGGCCCAAGGAGGGAGATGATCGCCCCTTCTTCTGGCTCTTTGAGAATG TGGTGGCCATGGGCGTTAGTGACAAGAGGGACATCTCGCGATTTCTTGA GTCTAACCCCGTGATGATTGACGCCAAAGAAGTGTCTGCTGCACACAGG GCCCGTTACTTCTGGGGTAACCTTCCTGGCATGAACAGGCCTTTGGCAT CCACTGTGAATGATAAGCTGGAGCTGCAAGAGTGTCTGGAGCACGGCA GAATAGCCAAGTTCAGCAAAGTGAGGACCATTACCACCAGGTCAAACT CTATAAAGCAGGGCAAAGACCAGCATTTCCCCGTCTTCATGAACGAGA AGGAGGACATCCTGTGGTGCACTGAAATGGAAAGGGTGTTTGGCTTCCC CGTCCACTACACAGACGTCTCCAACATGAGCCGCTTGGCGAGGCAGAG ACTGCTGGGCCGATCGTGGAGCGTGCCGGTCATCCGCCACCTCTTCGCT CCGCTGAAGGAATATTTTGCTTGTGTGTCTAGCGGCAATAGTAACGCTA ACAGCCGCGGGCCGAGCTTCAGCAGCGGCCTGGTGCCGTTAAGCTTGC GCGGCAGCCATATGGGCCCTATGGAGATATACAAGACAGTGTCTGCAT GGAAGAGACAGCCAGTGCGGGTACTGAGCCTCTTCAGAAACATCGACA AGGTACTAAAGAGTTTGGGCTTCTTGGAAAGCGGTTCTGGTTCTGGGGG AGGAACGCTGAAGTACGTGGAAGATGTCACAAATGTCGTGAGGAGAGA CGTGGAGAAATGGGGCCCCTTTGACCTGGTGTACGGCTCGACGCAGCCC CTAGGCAGCTCTTGTGATCGCTGTCCCGGCTGGTACATGTTCCAGTTCC ACCGGATCCTGCAGTATGCGCTGCCTCGCCAGGAGAGTCAGCGGCCCTT CTTCTGGATATTCATGGACAATCTGCTGCTGACTGAGGATGACCAAGAG ACAACTACCCGCTTCCTTCAGACAGAGGCTGTGACCCTCCAGGATGTCC GTGGCAGAGACTACCAGAATGCTATGCGGGTGTGGAGCAACATTCCAG GGCTGAAGAGCAAGCATGCGCCCCTGACCCCAAAGGAAGAAGAGTATC TGCAAGCCCAAGTCAGAAGCAGGAGCAAGCTGGACGCCCCGAAAGTTG ACCTCCTGGTGAAGAACTGCCTTCTCCCGCTGAGAGAGTACTTCAAGTA TTTTTCTCAAAACTCACTTCCTCTT	DNMT3A/L v1 (nt)

[1336]

663	AACCACGATCAGGAGTTTGACCCCCCTAAGGTGTACCCACCCGTGCCAG CCGAGAAGAGGAAGCCCATCCGCGTGCTGTCCCTGTTCGACGGCATCGC CACAGGCCTGCTGGTGCTGAAGGATCTGGGCATCCAGGTGGACAGATA TATCGCCTCCGAGGTGTGCGAGGATTCTATCACCGTGGGCATGGTGAGG CACCAGGGCAAGATCATGTACGTGGGCGACGTGCGCAGCGTGACACAG AAGCACATCCAGGAGTGGGGACCCTTCGACCTGGTCATCGGAGGCAGC CCCTGTAATGACCTGTCCATCGTGAACCCTGCAAGGAAGGGCCTGTATG AGGGAACCGGCAGACTGTTCTTTGAGTTCTACAGGCTGCTGCACGACGC CCGCCCTAAGGAGGGCGATGACAGGCCATTCTTTTGGCTGTTTGAGAAC GTGGTGGCCATGGGCGTGAGCGACAAGCGGGATATCTCCAGATTCCTG GAGTCTAATCCCGTGATGATCGATGCAAAGGAGGTGTCTGCCGCACAC AGGGCAAGGTACTTTTGGGGAAATCTGCCTGGCATGAACCGCCCACTG GCCAGCACCGTGAACGACAAGCTGGAGCTGCAGGAGTGCCTGGAGCAC GGAAGGATCGCCAAGTTCTCCAAGGTGCGGACAATCACCACAAGATCT AACAGCATCAAGCAGGGCAAGGATCAGCACTTCCCCGTGTTCATGAAT GAGAAGGAGGACATCCTGTGGTGTACCGAGATGGAGCGCGTGTTCGGC TTTCCAGTGCACTATACAGACGTGAGCAATATGAGCCGGCTGGCAAGG CAGAGACTGCTGGGCCGGTCCTGGTCTGTGCCAGTGATCAGACACCTGT TCGCCCCCCTGAAGGAGTACTTTGCCTGCGTGTCTAGCGGCAACTCTAA TGCCAACAGCAGAGGCCCTTCCTTTTCCTCTGGCCTGGTGCCACTGTCTC TGAGGGGCAGCCACATGGGCCCCATGGAGATCTACAAGACCGTGTCCG CCTGGAAGAGGCAGCCTGTGCGCGTGCTGTCTCTGTTCCGCAACATCGA CAAGGTGCTGAAGAGCCTGGGCTTTCTGGAGAGCGGATCCGGATCTGG AGGAGGCACCCTGAAGTATGTGGAGGATGTGACAAATGTGGTGCGGAG AGATGTGGAGAAGTGGGGCCCCTTCGATCTGGTGTACGGATCCACCCA GCCACTGGGAAGCTCCTGCGATAGGTGTCCAGGATGGTATATGTTCCAG TTTCACAGAATCCTGCAGTACGCACTGCCAAGGCAGGAGAGCCAGCGC CCTTTCTTTTGGATCTTTATGGACAACCTGCTGCTGACAGAGGATGACC AGGAGACAACAACCCGCTTCCTGCAGACAGAGGCAGTGACCCTGCAGG ATGTGAGGGGACGCGACTATCAGAATGCCATGCGGGTGTGGTCTAACA TCCCTGGCCTGAAGAGCAAGCACGCCCCCCTGACCCCTAAGGAGGAGG AGTACCTGCAGGCCCAGGTGCGGAGCAGATCCAAGCTGGATGCCCCTA AGGTGGACCTGCTGGTGAAGAATTGTCTGCTGCCACTGCGGGAGTACTT CAAGTACTTTAGTCAGAATAGCCTGCCACTG	DNMT3A/L v2 (nt)
664	MGPMEIYKTVSAWKRQPVRVLSLFRNIDKVLKSLGFLESGSGSGGGTLKY VEDVTNVVRRDVEKWGPFDLVYGSTQPLGSSCDRCPGWYMFQFHRILQY ALPRQESQRPFFWIFMDNLLLTEDDQETTTRFLQTEAVTLQDVRGRDYQN AMRVWSNIPGLKSKHAPLTPKEEEYLQAQVRSRSKLDAPKVDLLVKNCLL PLREYFKYFSQNSLPL	귀의 DNMT3L

[1337]

665	NHDQEFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIAS EVCEDSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDL SIVNPARKGLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGV SDKRDISRFLESNPVMIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKL ELQECLEHGRIAKFSKVRTITTRSNSIKQGKDQHFPVFMNEKEDILWCTEME RVFGFPVHYTDVSNMSRLARQRLLGRSWSVPVIRHLFAPLKEYFACV	인간 DNMT3A
666	NPLEMFETVPVWRRQPVRVLSLFEDIKKELTSLGFLESGSDPGQLKHVVDV TDTVRKDVEEWGPFDLVYGATPPLGHTCDRPPSWYLFQFHRLLQYARPKP GSPRPFFWMFVDNLVLNKEDLDVASRFLEMEPVTIPDVHGGSLQNAVRVW SNIPAIRSRHWALVSEEELSLLAQNKQSSKLAAKWPTKLVKNCFLPLREYF KYFSTELTSSL	C-말단 인간 DNM3L
667	SSGNSNANSRGPSFSSGLVPLSLRGSH	링커
668	MGSRETPSSCSKTLETLDLETSDSSSPDADSPLEEQWLKSSPALKEDSVDVV LEDCKEPLSPSSPPTGREMIRYEVKVNRRSIEDICLCCGTLQVYTRHPLFEGG LCAPCKDKFLESLFLYDDDGHQSYCTICCSGGTLFICESPDCTRCYCFECVD ILVGPGTSERINAMACWVCFLCLPFSRSGLLQRRKRWRHQLKAFHDQEGA GPMEIYKTVSAWKRQPVRVLSLFRNIDKVLKSL GFLESGSGSGGGTLKYVEDVTNVVRRDVEKWGPFDLVYGSTQPLGSSCDR CPGWYMFQFHRILQYALPRQESQRPFFWIFMDNLLLTEDDQETTTRFLQTE AVTLQDVRGRDYQNAMRVWSNIPGLKSKHAPLTPKEEEYLQAQVRSRSKL DAPKVDLLVKNCLLPLREYFKYFSQNSLPL	쥐의 DNMT3L
669	RTLVTFKDVFVDFTREEWKLLDTAQQILYRNVMLENYKNLVSLGYQLTKP DVILRLEKGEEPWLVEREIHQETHPDSETAFEIKSSV	KRAB
670	PKKKRKVGIHGVPAA	NLS 및 링커
671	DYKDDDDK	Flag tag
672	EASGSGRASPGIPGSTR	링커
673	GIHGVPAA	링커
674	MDYKDHDGDYKDHDI DYKDDDDK	3 x Flag 펩타이드
675	YPYDVPDYA	HA tag
676	HHHHHH	폴리-히스티딘 tag
677	KRPAATKKAGQAKKKKASDAKSLTAWS	링커
678	NTCCACAGCCTTCCACCAAGCTCTGCAGGATCCCAGAGTCAGGGGTCTG TATTTTCCTGCTGGTGGCTCCAGTTCAGGAACAGTAAACCCTGCTCCGA ATATTGCCTCTCACATCTCGTCAATCTCCGCGAGGACTGGGGACCCTGT GACGAACATGGAGAACATCACATCAGGATTCCTAGGACCCCTGCTCGT	Hep3B 컨센서스 서열 (3222bp)

[1338]

	GTTACAGGCGGGGTTTTTCTCGTTGACAAGAATCCTCACAATACCGCAG AGTCTAGACTCGTGGTGGACTTCTCTCAGTTTTCTAGGGGGTCCACCCG TGTGTCTTGGCCAAAATTCGCAGTCCCCAACCTCCAATCACTCACCAAC CTCCTGTCCTCCAATCTGTCCTGGTTATCGCTGGATGTGTCTGCGGCGTT TTATCATATTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTATTGGTTC TTCTGGATTATCAAGGTATGTTGCCCGTTTGTCCTCTAATTCCAGGATCA ACAACAACCAGTACGGGACCATGCAAAACCTGCACGACTCCTGCTCAA GGCAACTCTATGTTTCCCTCATGTTGCTGTACAAAACCTACGGATGGAA ATTGCACCTGTATTCCCATCCCATCGTCTTGGGCTTTCGCAAAATACCTA TGGGAGTGGGCCTCAGTCCGTTTCTCTTGGCTCAGTTTACTAGTGCCATT TGTTCAGTGGTTCGTAGGGCTTTCCCCCACTGTTTGGCTTTCAGCTATAT GGATGATGTGGTATTGGGGGCCAAGACTGTACAGCATCGTGAGTCCCTT TATACCGCTGTTACCAATTTTCTTTTGTCTCTGGGTATACATTTAAACCC TAACAAAACAAAAAGATGGGGTTATTCCCTAAACTTCATGGGTTACGTA ATTGGAAGTTGGGGGACATTGCCACAAGATCATATTGTACAAAAGATC AAACACTGTTTTAGAAAACTTCCTGTAAACAGGCCTATTGATTGGAAAG TATGTCAAAGGATTGTGGGTCTTTTGGGCTTTGCTGCTCCATTTACACAA TGTGGATATCCTGCCTTAATGCCTTTGTATGCATGTATACAAGCTAAAC AGGCTTTCACTTTCTCGCCAACTTACAAGGCCTTTCTAAGTAAACAGTA CCTGAACCTTTACCCCGTTGCTCGGCAACGGCCAGGTCTGTGCCAAGTG TTTGCTGACGCAACCCCCACTGGCTGGGGCTTAGCCATAGGCCATCAGC GCATGCGTGGAACCTTTGTGGCTCCTCTGCCGATCCATACTGCGGAACT CCTAGCCGCTTGTTTTGCTCGCAGCCGGTCTGGAGCAAAGCTCATCGGA ACTGACAATTCTGTCGTCCTCTCGCGGAAATATACATCGTTTCCATGGCT GCTAGGCTGTACTGCCAACTGGATCCTTCGCGGGACGTCCTTTGTTTAC GTCCCGTCGGCGCTGAATCCCGCGGACGACCCCTCTCGGGGCCGCTTGG GACTCTCTCGTCCCCTTCTCCGTCTGCCGTTCCAGCCGACCACGGGGCG CACCTCTCTTTACGCGGTCTCCCCGTCTGTGCCTTCTCATCTGCCGGTCC GTGTGCACTTCGCTTCACCTCTGCACGTTGCATGGAGACCACCGTGAAC GCCCATCAGAGCCTGCCCAAGGTCTTACATAAGAGGACTCTTGGACTCC CAGCAATGTCAACGACCGACCTTGAGGCCTACTTCAAAGACTGTGTGTT TAAGGACTGGGAGGAGCTGGGGGAGGAGATTAGGTTAATGATCTTTGT ATTAGGAGGCTGTAGGCATGAGCCTTTTATGCCTAAGAATGTGTCTGTC TCATAATGCAAGTGGTCAATCCAGCCAGCACCATGCGAACTTTTTCACC TCTGCCTAATCATCTCTTGTACATGTCCCACTGTTCAAGCCTCCAAGCTG TGCCTTGGGTGGCTTTGGGGCATGGACATTGACCCTTATAAAGAATTTG GAGCTACTGTGGAGTTACTCTCGTTTTTGCCTTCTGACTTTTTTCCTTCCG TCAGAGATCTCCTAGACACCGCCTCAGCTCTGTATCGGGAAGCCTTAGA GTCTCCTGAGCATTGCTCTCCTCACCATACTNNNNNNNNNNNNNNNN	

[1339]

	NNNN

	catgcgtggaacctttgtggctcctctgccgatccatactgcggaacttcttgcagcttgttttgctcgcagccggtctgga gcgaaactcatcgggactgataattctgtcgtcctttctcggaaatatacatcatttccatggctgctaggttgtgctgctaa ctggattcttcgcgggacgtcctttgtctacgtcccgtcggcgctgaatcctgcggacgacccctcccggggccgcttg ggactctatcgtccccttcttcgtctgccgtaccgtccgaccacggggcgcacctctctttacgcggtctcccgtctgtg ccttctcatctgccggtccgtgtgcacttcgcttcacctctgcacgttgcatggagaccaccgtgaacgcccatcggatc ctgcccaaggtcttacataagaggactcttggactcccagcaatgtcaacgaccgaccttgaggcttacttcaaagactg tgtgtttaaagactgggaggagttgggggaggagattaggttaaatattaggaggctgtaggcataaattggtctgcgca ccatcatcatgcaacttttcacctctgcctattcatctctcgttcatgtcctactgttcaagcctccaagctgtgccttgggtg gctttagggcatggacattgacccttataaagaatttggagctactgtggagttactctcgtttttgccttctgacttctttcctt cggtccgagatctcctagacaccgcctcagctctatatcgggaagccttagagtctcctgagcattgttcccctcatcata cagcactcaggcaagcaattctttgctgggggggaattaatgactctagctacctgggtgggtaataatttggaagatcca gcatccagggatctagtagtcaattatgtgaatactaacatgggcctaaagatcagacaactattgtggtttcatatttcttg ccttacttttggaagagaaacagtgcttgagtatttggtctccttcggagtgtggattcgcactcctccagcctatagacca ccaaatgcccctatcttatcaacacttccggaaactactgttgttagacgacgagaccgaggcaggtcccctagaagaa gaactccctcgcctcgcagacgaagatctcaatcgccgcgtcgcagaagatctcaatctcgggaatctcaatgttagtat tccttggactcataaggtgggaaattttactgggctttattcctctactgtccctatctttaatcctgaatggcaaacgccttc ctttcctaaaatccatttacacgaggacattattaataggtgtcagcaatttgtaggccctctcactgtaaatgaaaagaga agattgaacttaattatgcctgctaggttttatcctaactccactaaatatttgcctctagacaaaggaattaagccttattatc ctgaacatgtagttaatcattacttccagacccgacattatttacatactctttggaaggctgggattctatataagagggaa actacacgtagcgcctcattttgcgggtcaccatattcttgggaacaagagctacatcatgggaggttggttaccaaaac ctcgcaaaggcatggggacgaatctgtctgtccccaaccctctgggattctttcccgatcatcagttggacccagcattc ggagccaattcaaacaatccagactgggacttcaaccccacaaaggaccactggccacaagccaaccaggtaggag tgggagcattcggcccagggttcacccctccacacggaggtcttttggggtggagctctcaggctcaaggcacattgc atactgtgccagcagtgcctcctcctgcctccaccaatcggcagtcaggaaggcagcctactcccatctctccacctcta agagacagtcatcctcaggccatgcagtggaa	
680	ggagcaauugaccgggaagcucagaauaaacgcucaacuuuggccggaucuucuagagccaccaugaacca cgaucaggaguuugacccccccuaagguguacccacccgugccagccgagaagaggaagcccauccgcgugc ugucccuguucgacggcaucgccacaggccugcuggugcugaaggaucugggcauccagguggacagaua uaucgccuccgaggugugcgaggauucuaucaccgugggcauggugaggcaccagggcaagaucauguac gugggcgacgugcgcagcgugacacagaagcacauccaggaguggggacccuucgaccuggucaucggag gcagccccuguaaugaccuguccaucgugaacccugcaaggaagggccuguaugagggaaccggcagacug uucuuugaguucuacaggcugcugcacgacgcccgcccuaaggagggcgaugacaggccauucuuuuggc uguuugagaacgugguggccaugggcgugagcgacaagcgggauaucuccagauuccuggagucuaaucc cgugaugaucgaugcaaaggaggugucugccgcacacagggcaagguacuuuuggggaaaucugccuggc augaaccgcccacuggccagcaccgugaacgacaagcuggagcugcaggagugccuggagcacggaaggau cgccaaguucuccaaggugcggacaaucaccacaagaucuaacagcaucaagcagggcaaggaucagcacuu ccccguguucaugaaugagaaggaggacauccuguggugguaccgagauggagcgcguguucggcuuucca gugcacuauacagacgugagcaauaugagccggcuggcaaggcagagacugcugggccggcuccuggucug	HBVg_mRNA_1

[1341]

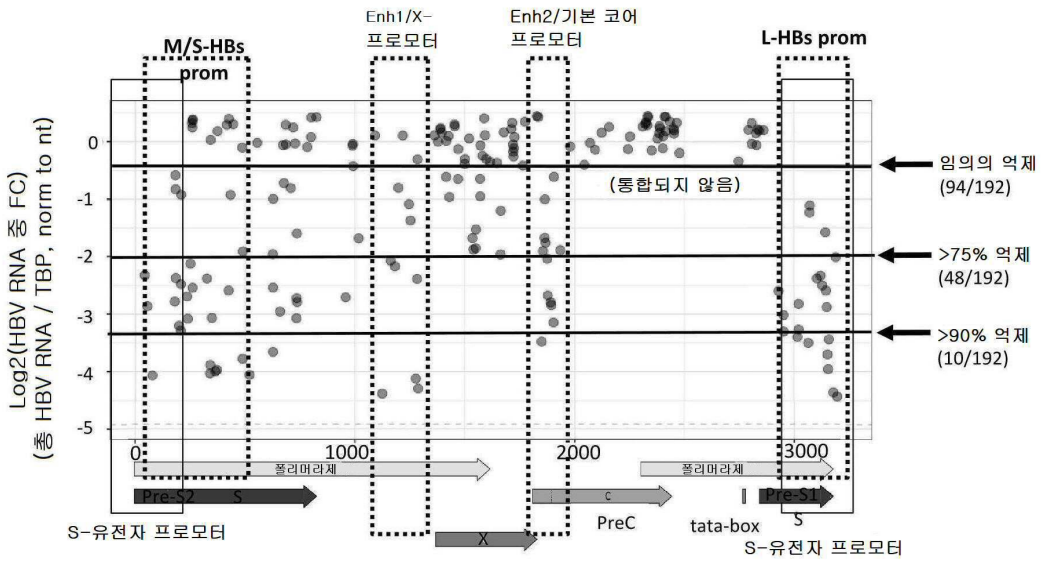
	ugccagugaucagacaccuguucgccccccugaaggaguacuuugccugcgugucuagcggcaacucuaau gccaacagcagaggcccuuccuuuuccucuggccuggugccacugucucugaggggcagccacaugggccc cauggagaucuacaagaccguguccgccuggaagaggcagccugugcgcgugcugucucuguuccgcaaca ucgacaaggugcucaagagccugggcuuucuggagagcggauccggaucuggaggaggcacccugaagua uguggaggaugugacaaauguggugcggagagauguggagaaguggggccccuucgaucugguguacgg auccacccagccacugggaagcuccugcgauagguguccaggaugguauauguuccaguuucacagaaucc ugcaguacgcacugccaaggcaggagagccagcgcccuuucuuuuggaucuuuauggacaaccugcugcu gacagaggaugaccaggagacaacaacccgcuuccugcagacagaggcagugacccugcaggaugugaggg gacgcgacuaucagaaugccaugcgggugguggucuaacaucccuggccugaaaagcaagcacgcccccug accccuaaggaggaggaguaccugcaggcccaggugcggagcagauccaagcuggaugccccuaaggugga ccugcuggugaagaauugucugcugccacugcgggaguacuucaaguacuuuagucagaauagccugcca cuggaggcaagcggauccggaagggcaucuccuggaaucccaggaagcacccgcaaccccaagaagaagcg gaaggugggcauccacggcgugcccgccgccgacaagaaguacagcaucggccuggccaucggcaccaaca gcgugggcugggccgugaucaccgacgaguacaaggugcccagcaagaaguucaaggugcugggcaacacc gaccggcacagcaucaagaagaaccugaucggcgcccugcuguucgacagcggcgagaccgccgaggccac ccggcugaagcggaccgcccggcggcgguacacccggcggaagaaccggaucugcuaccugcaggagaucu ucagcaacgagauggccaagguggacgacagcuucuuccaccggcuggaggagagcuuccugguggagga ggacaagaagcacgagcggcaccccaucuucggcaacaucguggacgagguggccuaccacgagaaguacc ccaccaucuaccaccugcggaagaagcugguggacagcaccgacaaggccgaccugcggcugaucuaccug gcccuggcccacaugaucaaguuccggggccacuuccugaucgagggcgaccugaaccccgacaacagcga cguggacaagcuguucauccagcuggugcagaccuacaaccagcuguucgaggagaaccccaucaacgcca gcggcguggacgccaaggccauccugagcgcccggcugagcaagagccggcggcuggagaaccugaucgcc cagcugcccggcgagaagaagaacggccuguucggcaaccugaucgcccugagccugggccugaccccaa cuucaagagcaacuucgaccuggccgaggacgccaagcugcagcugagcaaggacaccuacgacgacgaccu ggacaaccugcuggcccagaucggcgaccaguacgccgaccuguuccuggccgccaagaaccugagcgacg ccauccugcugagcgacauccugcgggugaacaccgagaucaccaaggccccccugagcgccagcaugauca agcgguacgacgagcaccaccaggaccugacccugcugaaggcccuggugcggcagcagcugcccgagaag uacaaggagaucuucuucgaccagagcaagaacggcuacgccggcuacaucgacggcggcgccagccagga ggaguucuacaaguucaucaagcccauccuggagaagauggacggcaccgaggagcugcuggugaagcuga accgggaggaccugcugcggaagcagcggaccuucgacaacggcagcaucccccaccagauccaccugggc gagcugcacgccauccugcggcggcaggaggacuucuacccccuuccugaaggacaaccgggagaagaucga gaagauccugaccuuccggauccccuacuacgugggccccccuggcccggggcaacagccggguucgccugga ugacccggaaaagcgaggagaccaucacccccuggaacuucgaggaggugguggacaagggcgccagcgcc cagagcuucaucgagcggaugaccaacuucgacaagaaccugcccaacgagaaggugcugcccaagcacagc cugcuguacgaguacuucaccguguacaacgagcugaccaaggugaaguacgugaccgagggcaugcggaa gcccgccuuccugagcggcgagcagaagaaggccaucguggaccugcuguucaagaccaaccggaaagguga ccgugaagcagcugaaggaggacuacuucaagaagaucgagugcuucgacagcguggagaucagcggcgu ggaggaccgguucaacgccagccugggcaccuaccacgaccugcugaagaucaucaaggacaaggacuucc uggacaacgaggagaacgaggacauccuggaggacaucgugcugacccugacccuguucgaggaccgggag	

[1342]

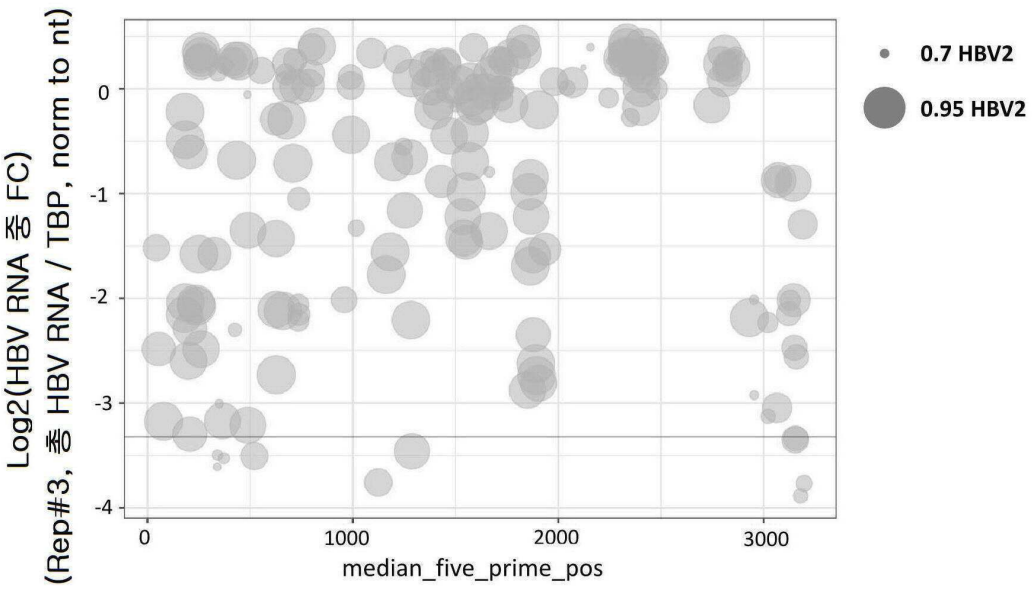
	augaucgaggagcggcugaagaccuacgcccaccuguucgacgacaaggugaugaagcagcugaagcggcg gcgguacaccggcuggggccggcugagccggaagcugaucaacggcaucccgggacaagcagagcggcaaga ccauccuggacuuccugaaaagcgacggcuucgccaaccggaacuucaugcagcugauccacgacgacagcc ugaccuucaaggaggacauccagaaggcccaggugagcggccagggcgacagccugcacgagcacaucgcc aaccuggccggcagccccgccaucaagaagggcauccugcagaccgugaaggugguggacgagcuggugaa ggugaugggccggcacaagcccgagaacaucgugaucgagauggcccgggagaaccagaccacccagaagg gccagaagaacagccgggagcggaugaagcggaucgaggagggcaucaaggagcugggcagccagauccug aaggagcaccccguggagaacacccagcugcagaacgagaagcuguaccuguacuaccugcagaacggccg ggacauguacguggaccaggagcuggacaucaaccggcugagcgacuacgacguggacgccaucgugcccc agagcuuccugaaggacgacagcaucgacaacaaggugcugacccggagcgacaagaaccggggcaagagc gacaacgugcccagcgaggagguggugaagaagaugaagaacuacuggcggcagcugcugaacgccaagcu gaucacccagcggaaguucgacaaccugaccaaggccgagcggggcggccugagcgagcuggacaaggccg gcuucaucaagcggcagcugguggagacccggcagaucaccaagcacguggcccagauccuggacagccgg augaacaccaaguacgacgagaacgacaagcugauccgggaggugaaggugaucacccugaaaagcaagcu ggugagcgacuuccggaaggacuuccaguucuacaaggugcgggagaucaacaacuaccaccacgcccacg acgccuaccugaacgccgugguggcaccgcccugaucaagaaguaccccaagcuggagagcgaguucgug uacggcgacuacaagguguacgacgugcggaagaugaucgccaagagcgagcaggagaucggcaaggccac cgccaaguacuucuucuacagcaacaucaugaacuucuucaagaccgagaucacccuggccaacggcgagau ccggaagcggccccugaucgagaccaacggcgagaccggcgagaucguguggggacaagggccgggacuucg ccaccgugcggaaggugcugagcaugccccaggugaacaucgugaagaagaccgaggugcagaccggcggc uucagcaaggagagcauccugcccaagcggaacagcgacaagcugaucgcccggaagaaggacugggaccc caagaaguacggcggcuucgacagccccaccguggccuacagcgugcuggugguggccaagguggagaag ggcaagagcaagaagcugaaaagcgugaaggagcugcugggcaucaccaucauggagcggagcagcuucga gaagaaccccaucgacuuccuggaggccaagggcuacaaggaggugaagaaggaccugaucaucaagcugc ccaaguacagccuguucgagcuggagaacggccggaagcggaugcuggccagcgccggcgagcugcagaag ggcaacgagcuggcccugcccagcaaguacgugaacuuccuguaccuggccagccacuacgagaagcugaa gggcagccccgaggacaacgagcagaagcagcuguucguggagcagcacaagcacuaccuggacgagauca ucgagcagaucagcgaguucagcaagcgggugauccuggccgacgccaaccuggacaaggugcugagcgcc uacaacaagcaccgggacaagcccauccgggagcaggccgagaacaucauccaccuguucacccugaccaac cugggcgccccgccgccuucaaguacuucgacaccaccaucgaccggaagcgguacaccagcaccaaggag gugcuggacgccacccugauccaccagagcaucaccggccuguacgagacccggaucgaccugagccagcu gggcggcgacagcggcggcaagcggcccgccgccaccaagaaggccggccaggccaagaagaagaaggcua gcgaugcuaagucacugacugccuggucccggacacuggugaccuucaaggauguguuuguggacuucac cagggaggaguggaagcugcuggacacugcucagcagauccuguacagaaaugugaugcuggagaacuau aagaaccuggguuuccuuggguuaucagcuuacuaagccagaugugauccucccgguuggagaagggagagg aacccuggcugguggagagagaaauucaccaagagacccauccugauucagagacugcauuugaaaucaaa ucaucaguuccgaaaaagaaacgcaaaguuuaguaagaauucaaagaaaguuucuucacauucucucgagc guacgaaaa

도면

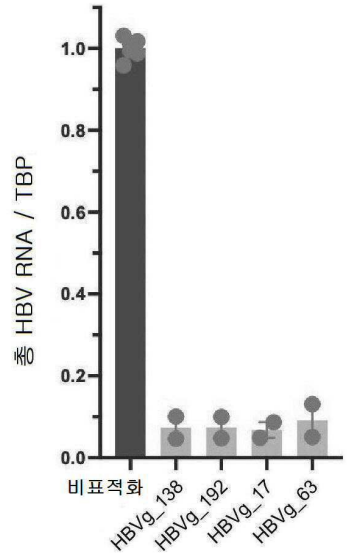
도면1


M/S-HBs prom
Enh1/X-프로모터
Enh2/기본 코어 프로모터
L-HBs prom
Log2(HBV RNA 중 FC)
(총 HBV RNA / TBP, norm to nt)
0
-1
-2
-3
-4
-5
(통합되지 않음)
임의의 억제 (94/192)
>75% 억제 (48/192)
>90% 억제 (10/192)
0
1000
2000
3000
폴리머라제
폴리머라제
Pre-S2
S
C
PreC
Pre-S1
S
tata-box
X
S-유전자 프로모터
S-유전자 프로모터

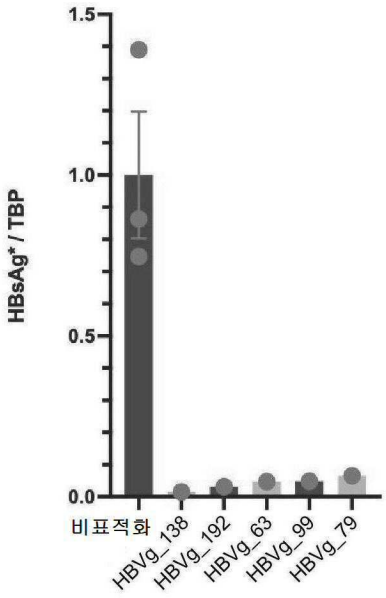
도면2


Log2(HBV RNA 중 FC)
(Rep#3, 총 HBV RNA / TBP, norm to nt)
0
-1
-2
-3
-4
0
1000
2000
3000
median_five_prime_pos
0.7 HBV2
0.95 HBV2

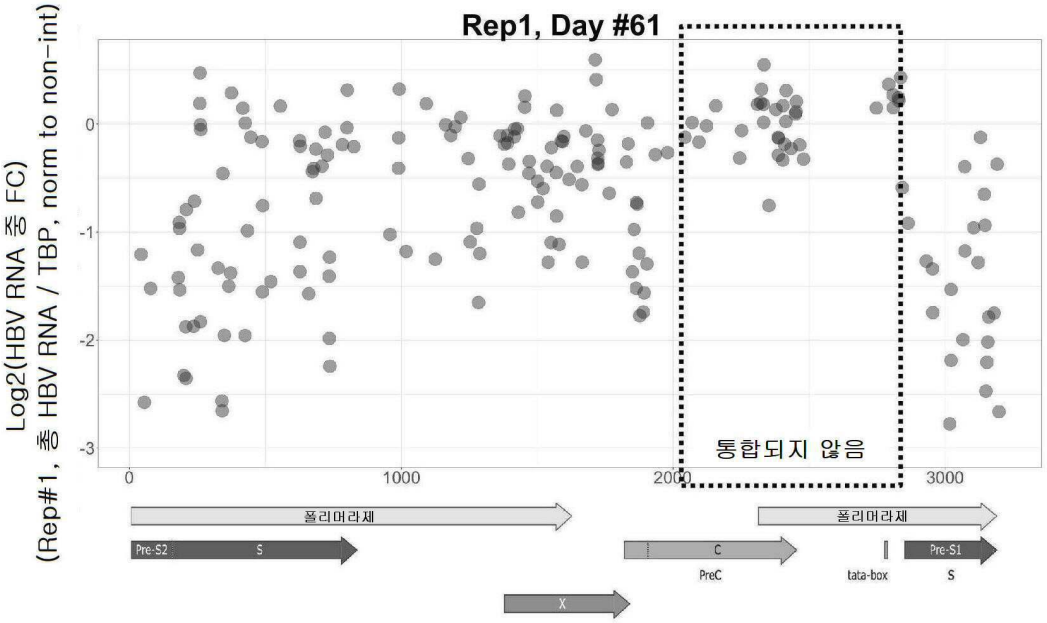
도면3a


총 HBV RNA / TBP
1.0
0.8
0.6
0.4
0.2
0.0
비표적화
HBVg_138
HBVg_192
HBVg_17
HBVg_63

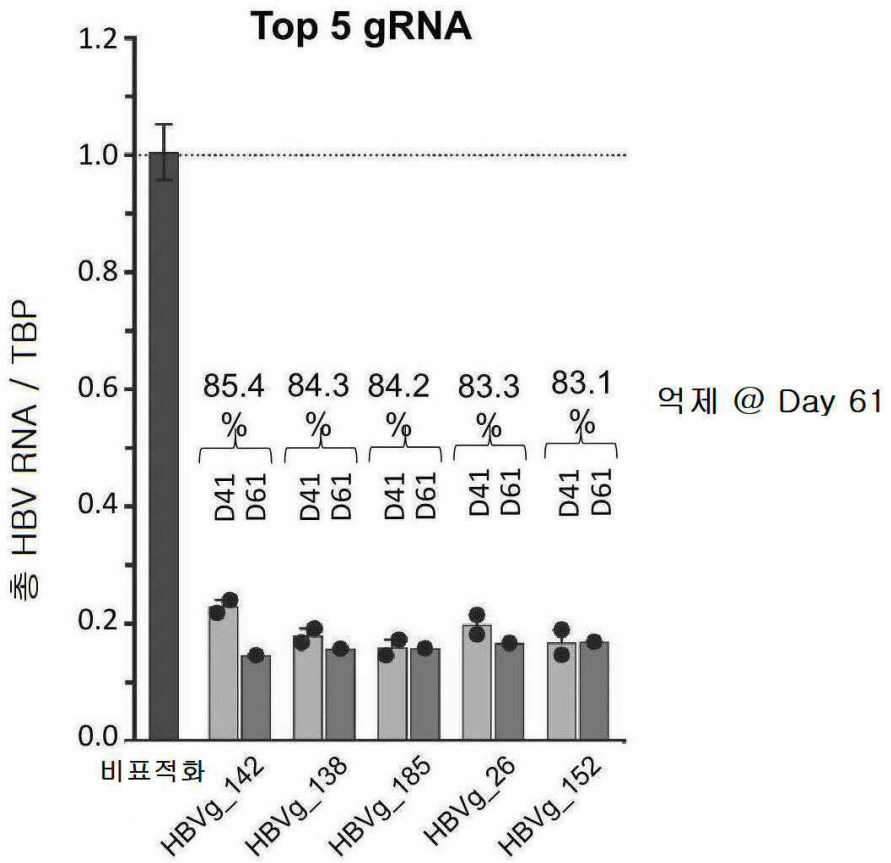
도면3b


HBsAg* / TBP
1.5
1.0
0.5
0.0
비표적화
HBVg_138
HBVg_192
HBVg_63
HBVg_99
HBVg_79

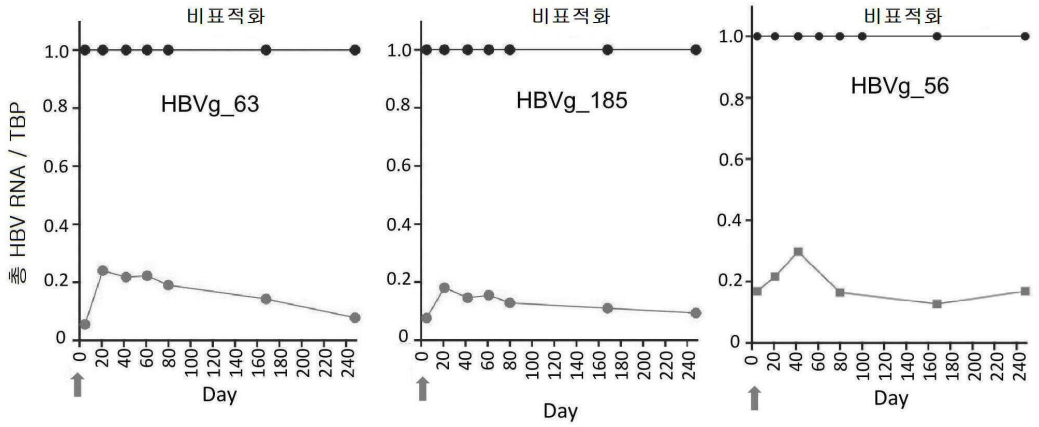
도면4


Rep1, Day #61
Log2(HBV RNA 중 FC)
(Rep#1, 총 HBV RNA / TBP, norm to non-int)
0
-1
-2
-3
0
1000
2000
3000
통합되지 않음
폴리머라제
폴리머라제
Pre-S2
S
C
PreC
tata-box
Pre-S1
S
X

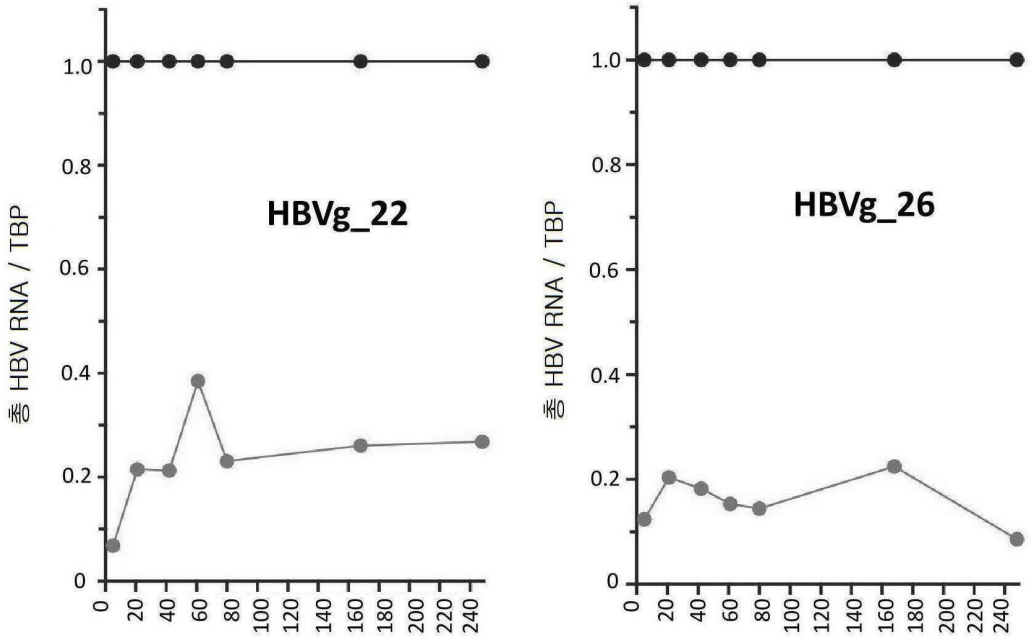
도면5


Top 5 gRNA
총 HBV RNA / TBP
1.2
1.0
0.8
0.6
0.4
0.2
0.0
85.4 %
84.3 %
84.2 %
83.3 %
83.1 %
D41
D61
억제 @ Day 61
비표적화
HBVg_142
HBVg_138
HBVg_185
HBVg_26
HBVg_152

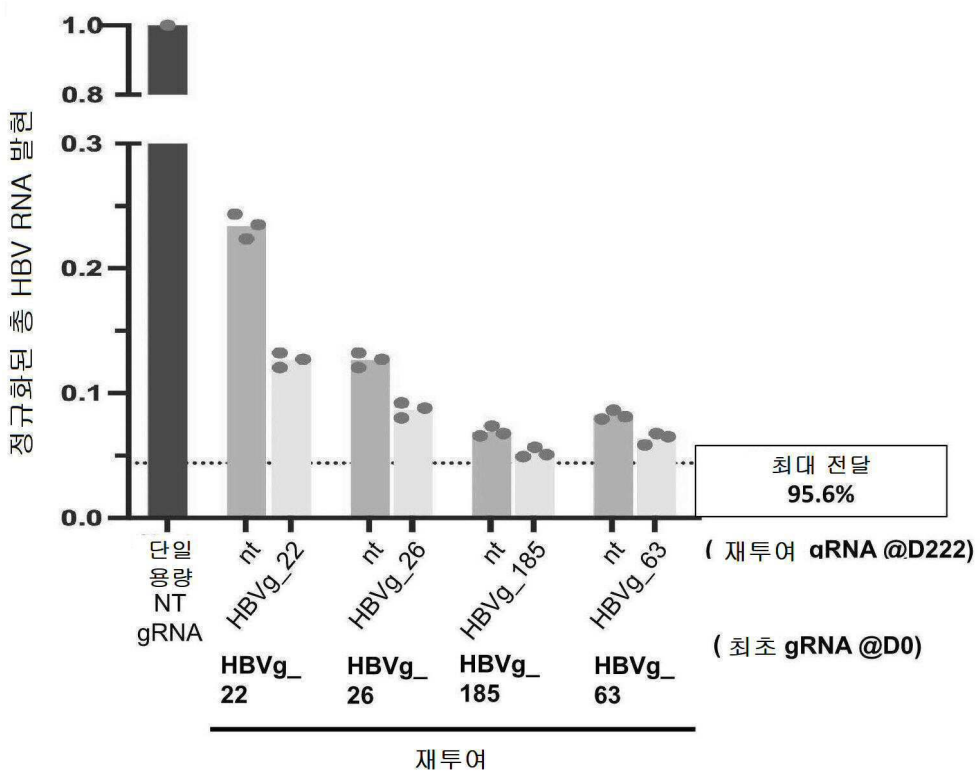
도면6a


비표적화
HBVg_63
HBVg_185
HBVg_56
총 HBV RNA / TBP
Day

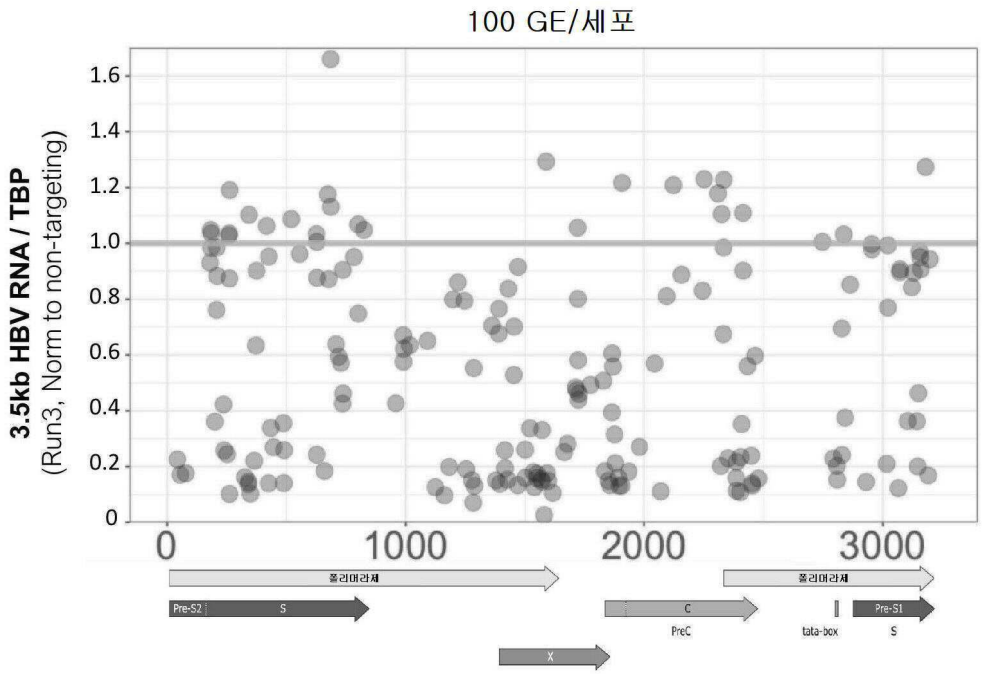
도면6b


HBVg_22
HBVg_26
총 HBV RNA / TBP

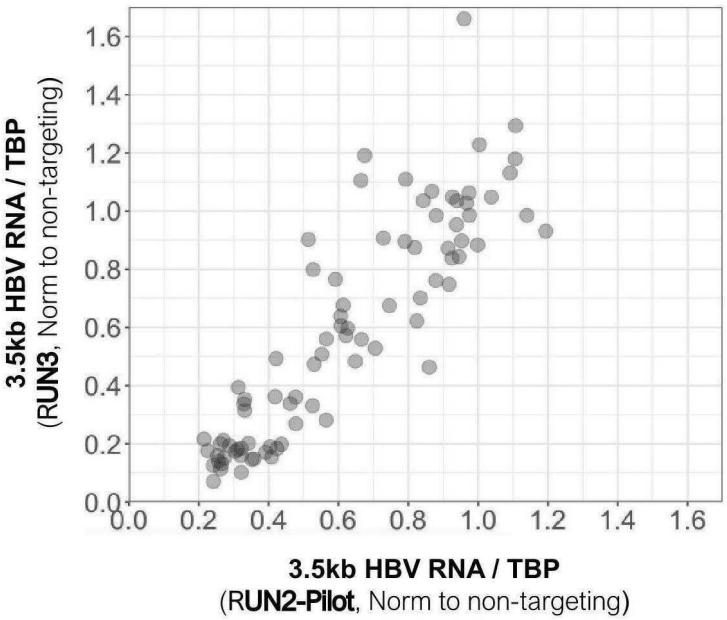
도면7


정규화된 총 HBV RNA 발현
1.0
0.8
0.3
0.2
0.1
0.0
최대 전달
95.6%
단일 용량 NT gRNA
nt HBVg_22
nt HBVg_26
nt HBVg_185
nt HBVg_63
(재투여 gRNA @D222)
(최초 gRNA @D0)
HBVg_ 22
HBVg_ 26
HBVg_ 185
HBVg_ 63
재투여

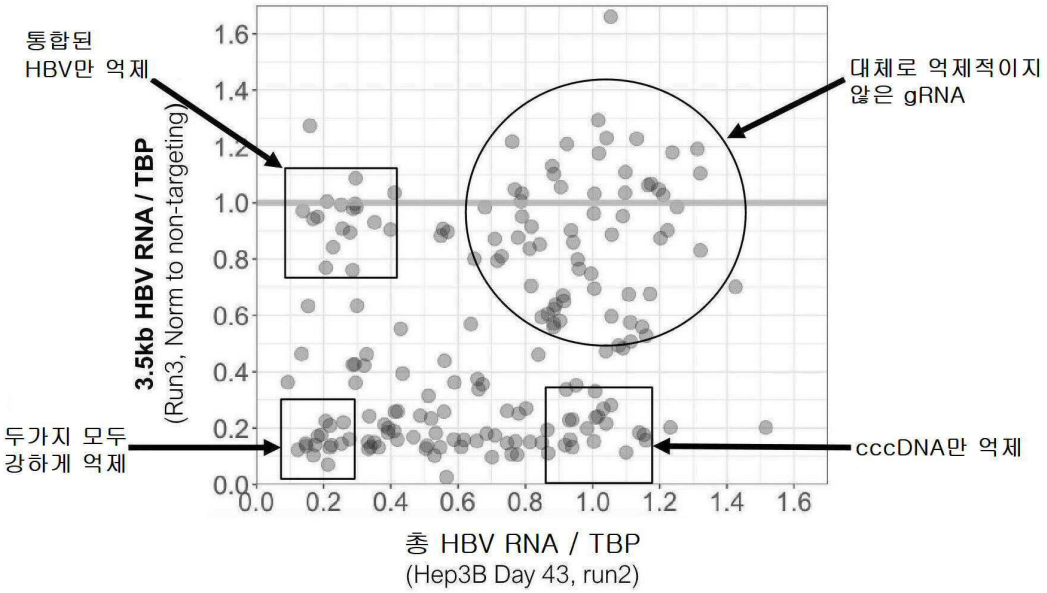
도면8


100 GE/세포
3.5kb HBV RNA / TBP
(Run3, Norm to non-targeting)
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0
0
1000
2000
3000
폴리머라제
Pre-S2
S
폴리머라제
C
PreC
tata-box
Pre-S1
S
X

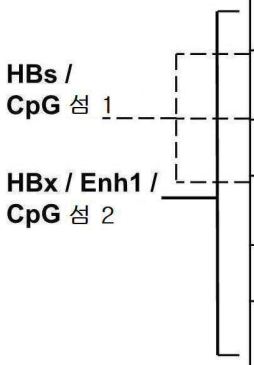
도면9


3.5kb HBV RNA / TBP
(RUN3, Norm to non-targeting)
3.5kb HBV RNA / TBP
(RUN2-Pilot, Norm to non-targeting)

도면10


통합된
HBV만 억제
대체로 억제적이지
않은 gRNA
두가지 모두
강하게 억제
cccDNA만 억제
3.5kb HBV RNA / TBP
(Run3, Norm to non-targeting)
총 HBV RNA / TBP
(Hep3B Day 43, run2)

도면11

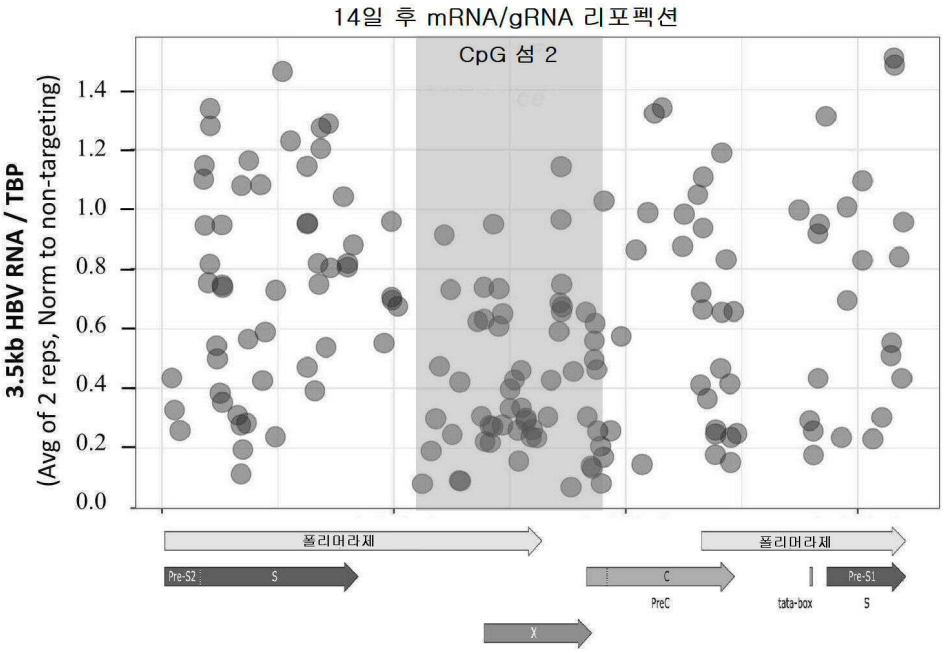


표적 서열	HBV 위치	cccDNA (D5)	Hep3B Int (D5)	Hep3B Int (D80)	HBV conserv % 0 MM (+1-2 MM)
HBVg_22	1283	93.0	89.4	77.0	92.1 (98.3)
HBVg_185	341	85.5	90.5	87.2	67.2 (70.9)
HBVg_23	340	86.4	91.5	79.4	66.9 (70.3)
HBVg_182	349	89.7	85.9	82.3	38.0 (70.5)
HBVg_82	489	86.0	86.9	87.1	55.1 (95.8)
HBVg_143	3065	87.8	86.6	89.4	30.7 (86.2)
HBVg_63	1288	87.0	93.8	81.0	54.7 (94.5)
HBVg_26	55	83.0	89.4	85.6	90.3 (91.5)

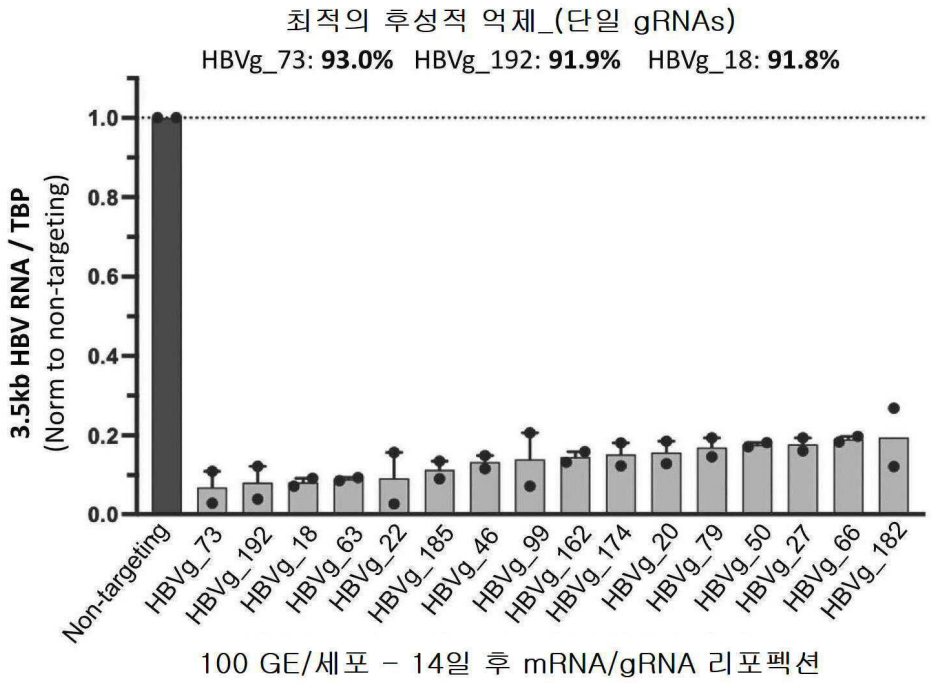
도면12

표적 서열	HBV 표적 위치	cccDNA 억제 (D14)	Int HBV 억제 (D168)	HBV conserv % 0 MM (+1/2 MM, Cas9-관용성)
HBVg_22	HBx / EN1	90.8%	74.0 %	92.1 (98.3)
HBVg_63	HBx / EN1	91.1%	85.7 %	54.7 (94.5)
HBVg_185	HBs	88.7%	88.9%	67.2 (70.9)
HBVg_99	BCP	86.1%	82.5%	51.6 (96.0)

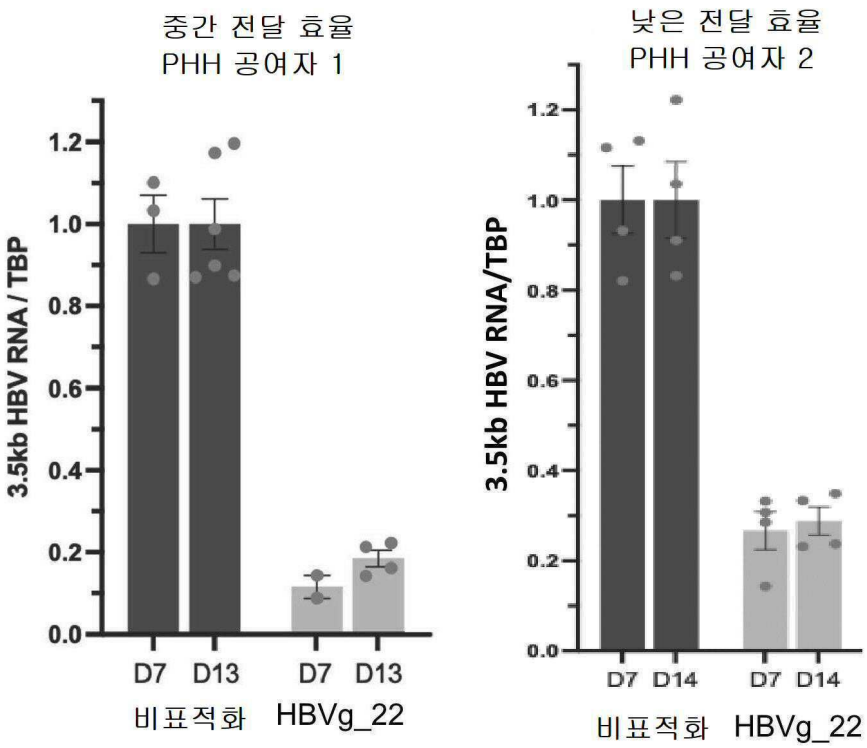
도면13


14일 후 mRNA/gRNA 리포펙션
CpG 섬 2
3.5kb HBV RNA / TBP
(Avg of 2 reps, Norm to non-targeting)
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
폴리머라제
폴리머라제
Pre-S2
S
C
PreC
tata-box
Pre-S1
S
X

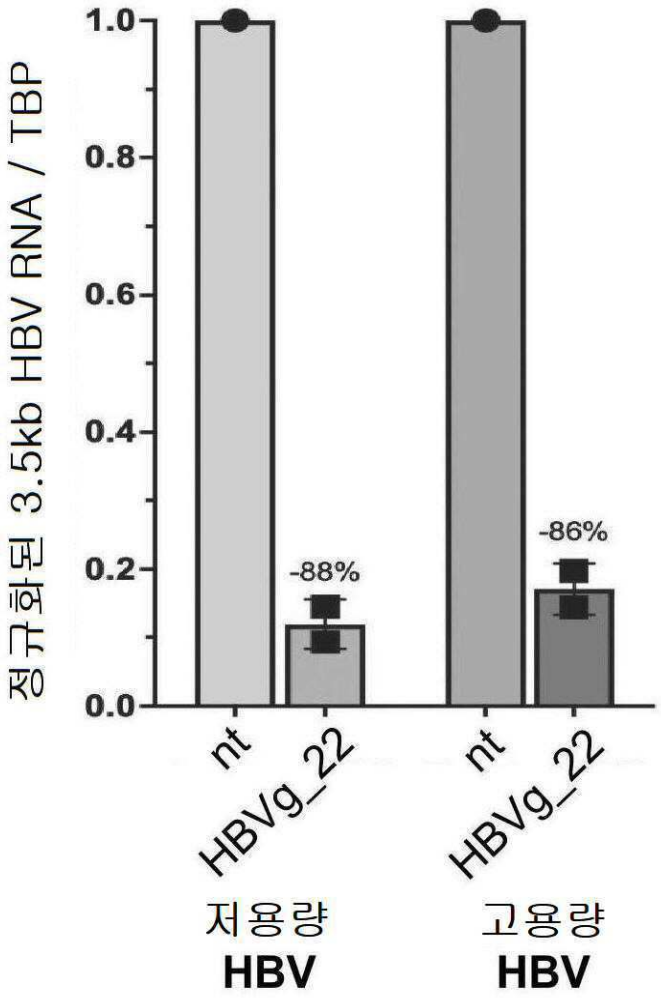
도면14


최적의 후성적 억제_(단일 gRNAs)
HBVg_73: 93.0% HBVg_192: 91.9% HBVg_18: 91.8%
3.5kb HBV RNA / TBP
(Norm to non-targeting)
1.0
0.8
0.6
0.4
0.2
0.0
Non-targeting
HBVg_73
HBVg_192
HBVg_18
HBVg_63
HBVg_22
HBVg_185
HBVg_46
HBVg_99
HBVg_162
HBVg_174
HBVg_20
HBVg_79
HBVg_50
HBVg_27
HBVg_66
HBVg_182
100 GE/세포 - 14일 후 mRNA/gRNA 리포펙션

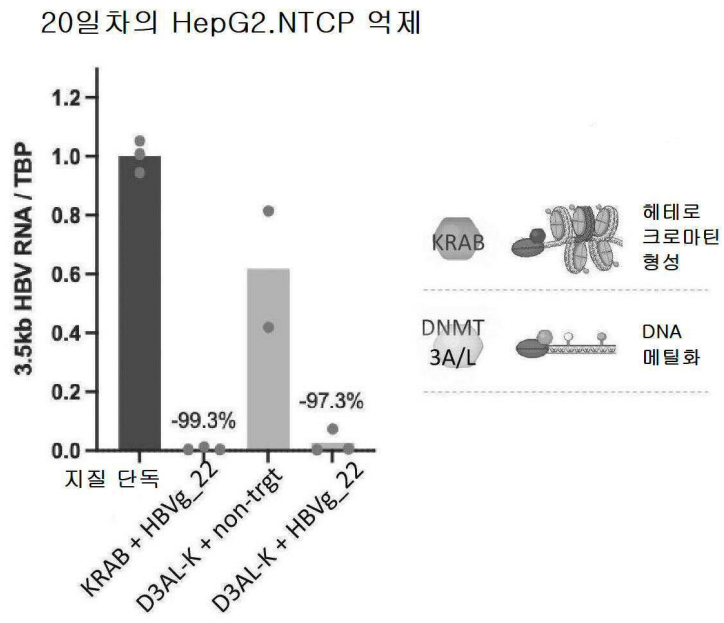
도면15a


중간 전달 효율
PHH 공여자 1
3.5kb HBV RNA / TBP
1.2
1.0
0.8
0.6
0.4
0.2
0.0
D7 D13 D7 D13
비표적화 HBVg_22
낮은 전달 효율
PHH 공여자 2
3.5kb HBV RNA/TBP
1.2
1.0
0.8
0.6
0.4
0.2
0.0
D7 D14 D7 D14
비표적화 HBVg_22

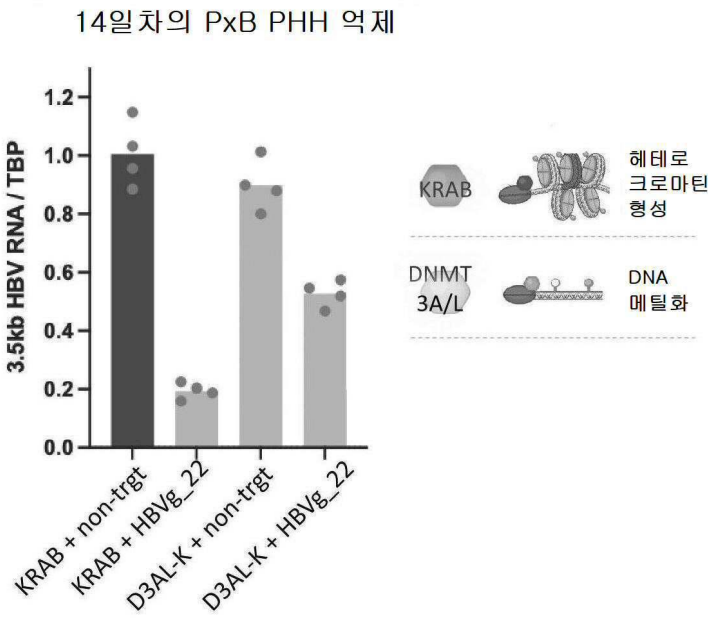
도면15b


정규화된 3.5kb HBV RNA / TBP
1.0
0.8
0.6
0.4
0.2
0.0
-88%
-86%
nt
HBVg_22
nt
HBVg_22
저용량
HBV
고용량
HBV

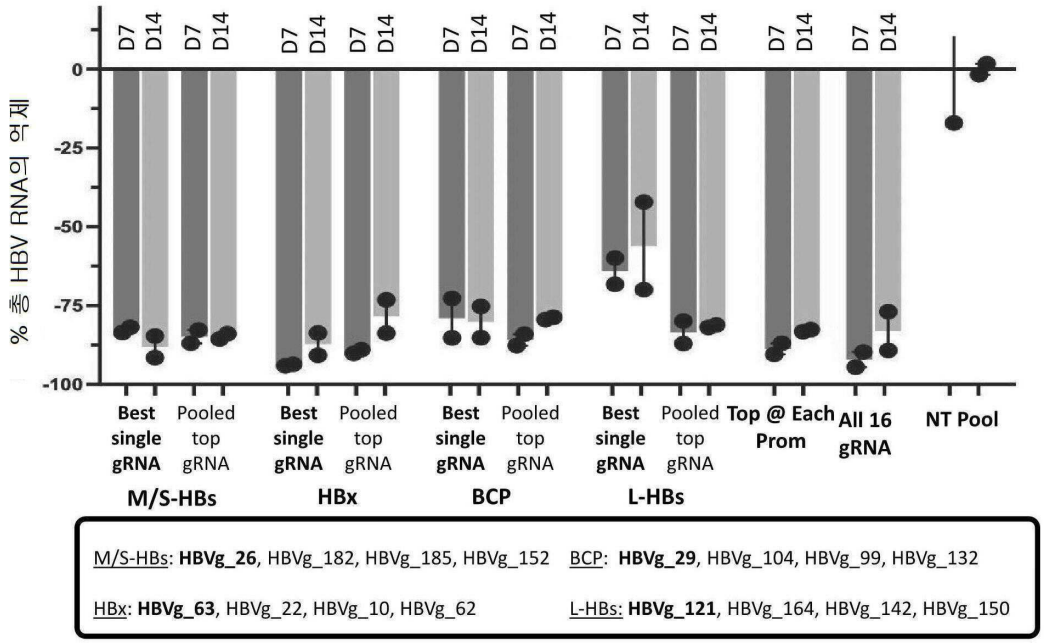
도면16a


20일차의 HepG2.NTCP 억제
3.5kb HBV RNA / TBP
1.2
1.0
0.8
0.6
0.4
0.2
0.0
-99.3%
-97.3%
지질 단독
KRAB + HBVg_22
D3AL-K + non-trgt
D3AL-K + HBVg_22
KRAB
헤테로
크로마틴
형성
DNMT
3A/L
DNA
메틸화

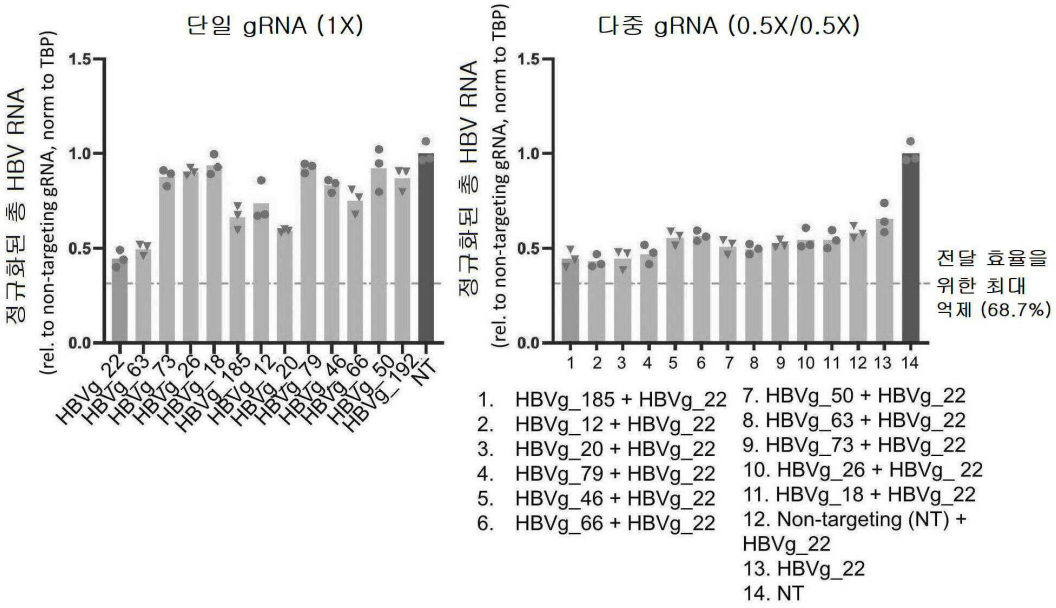
도면16b


14일차의 PxB PHH 억제
3.5kb HBV RNA / TBP
1.2
1.0
0.8
0.6
0.4
0.2
0.0
KRAB + non-trgt
KRAB + HBVg_22
D3AL-K + non-trgt
D3AL-K + HBVg_22
KRAB
헤테로
크로마틴
형성
DNMT
3A/L
DNA
메틸화

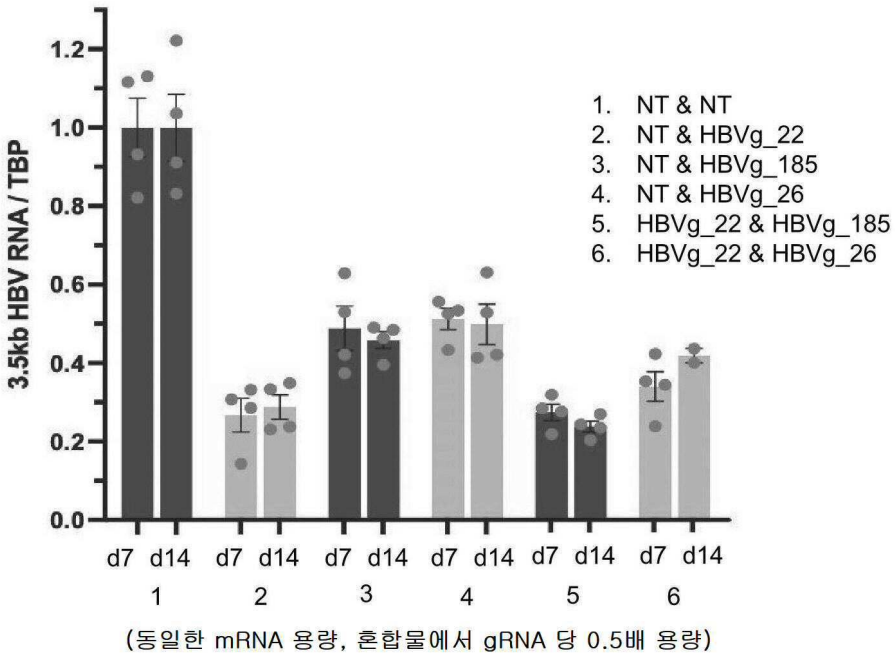
도면17


% 총 HBV RNA의 억제
0
-25
-50
-75
-100
D7
D14
Best single gRNA
Pooled top gRNA
Top @ Each Prom
All 16 gRNA
NT Pool
M/S-HBs
HBx
BCP
L-HBs
M/S-HBs: HBVg_26, HBVg_182, HBVg_185, HBVg_152
BCP: HBVg_29, HBVg_104, HBVg_99, HBVg_132
HBx: HBVg_63, HBVg_22, HBVg_10, HBVg_62
L-HBs: HBVg_121, HBVg_164, HBVg_142, HBVg_150

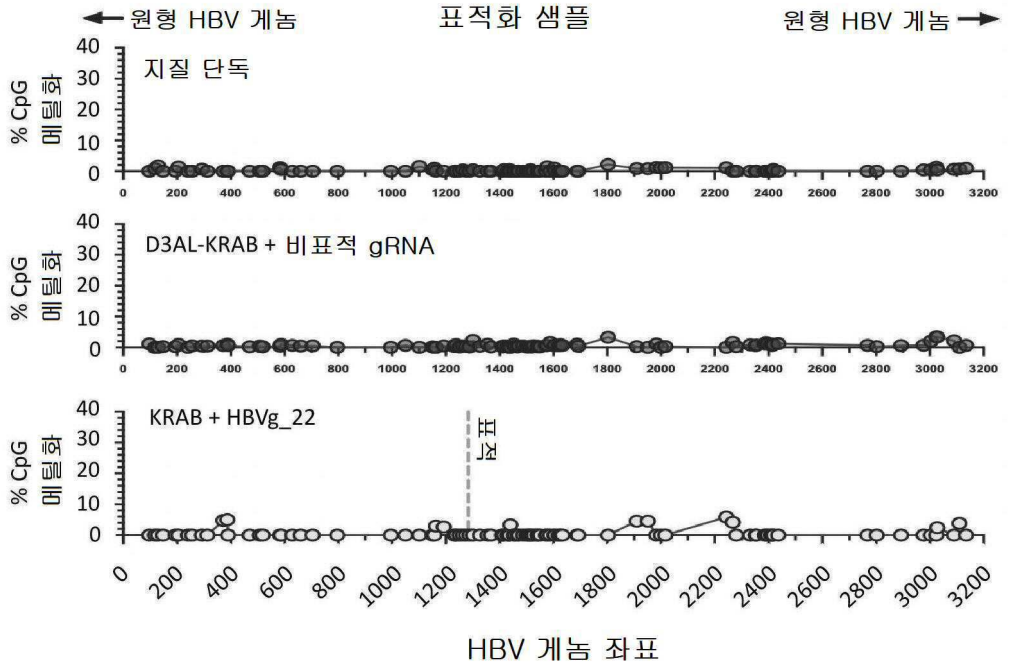
도면18


단일 gRNA (1X)
다중 gRNA (0.5X/0.5X)
정규화된 총 HBV RNA
(rel. to non-targeting gRNA, norm to TBP)
1.5
1.0
0.5
0.0
HBVg_22
HBVg_63
HBVg_73
HBVg_26
HBVg_18
HBVg_185
HBVg_12
HBVg_20
HBVg_79
HBVg_46
HBVg_66
HBVg_50
HBVg_192
NT
1 2 3 4 5 6 7 8 9 10 11 12 13 14
전달 효율을 위한 최대 억제 (68.7%)
1. HBVg_185 + HBVg_22
2. HBVg_12 + HBVg_22
3. HBVg_20 + HBVg_22
4. HBVg_79 + HBVg_22
5. HBVg_46 + HBVg_22
6. HBVg_66 + HBVg_22
7. HBVg_50 + HBVg_22
8. HBVg_63 + HBVg_22
9. HBVg_73 + HBVg_22
10. HBVg_26 + HBVg_ 22
11. HBVg_18 + HBVg_22
12. Non-targeting (NT) + HBVg_22
13. HBVg_22
14. NT

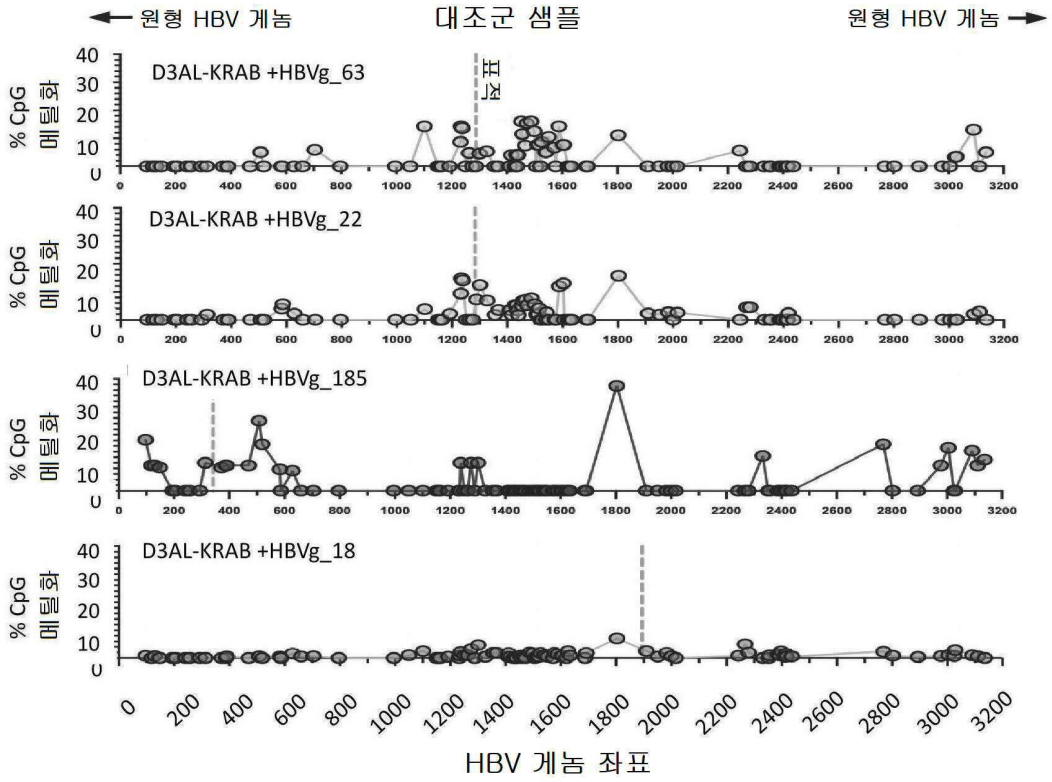
도면19


3.5kb HBV RNA / TBP
1.2
1.0
0.8
0.6
0.4
0.2
0.0
d7 d14 d7 d14 d7 d14 d7 d14 d7 d14 d7 d14
1 2 3 4 5 6
1. NT & NT
2. NT & HBVg_22
3. NT & HBVg_185
4. NT & HBVg_26
5. HBVg_22 & HBVg_185
6. HBVg_22 & HBVg_26
(동일한 mRNA 용량, 혼합물에서 gRNA 당 0.5배 용량)

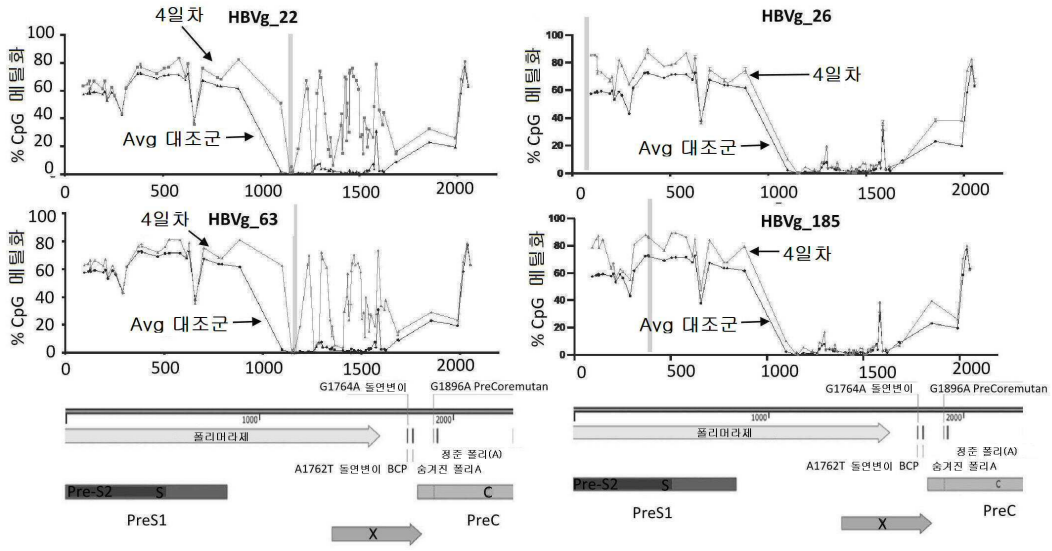
도면20a


원형 HBV 게놈
표적화 샘플
원형 HBV 게놈
지질 단독
D3AL-KRAB + 비표적 gRNA
KRAB + HBVg_22
표적
% CpG 메틸화
HBV 게놈 좌표

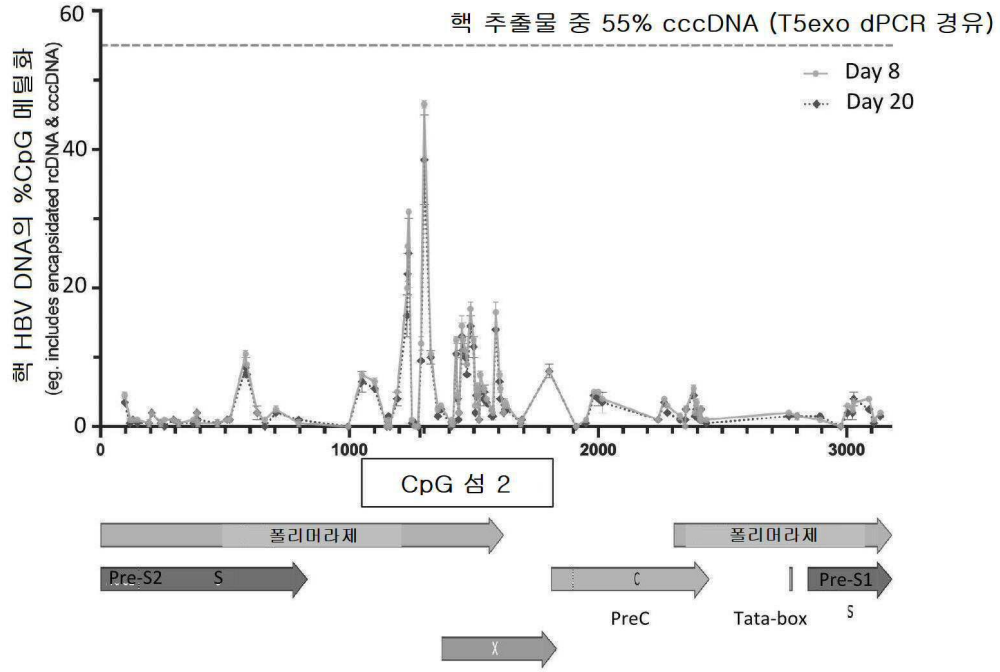
도면20b


원형 HBV 게놈
대조군 샘플
원형 HBV 게놈
D3AL-KRAB +HBVg_63
D3AL-KRAB +HBVg_22
D3AL-KRAB +HBVg_185
D3AL-KRAB +HBVg_18
표적
% CpG 메틸화
HBV 게놈 좌표

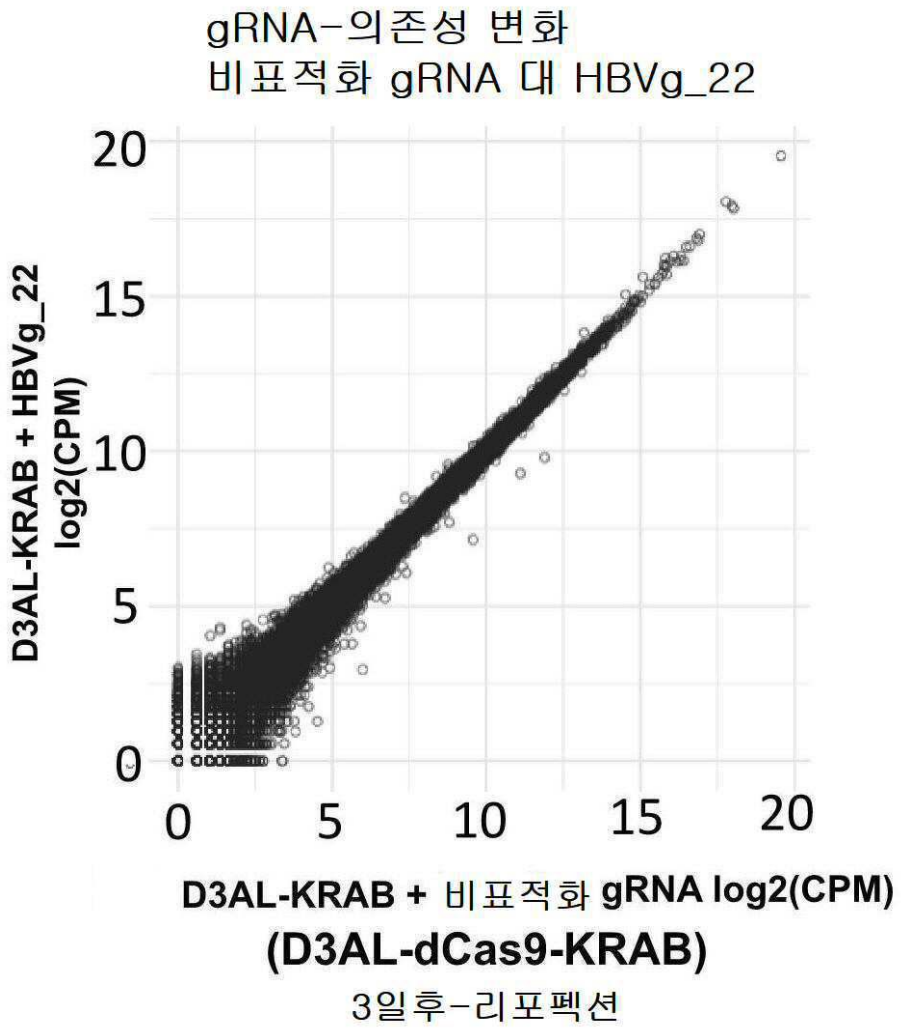
도면21


4일차
HBVg_22
Avg 대조군
% CpG 메틸화
HBVg_26
4일차
Avg 대조군
HBVg_63
4일차
Avg 대조군
HBVg_185
4일차
Avg 대조군
G1764A 돌연변이
G1896A PreCoremutan
폴리머라제
A1762T 돌연변이 BCP
정준 폴리(A)
송거진 폴리A
Pre-S2
S
C
PreS1
X
PreC

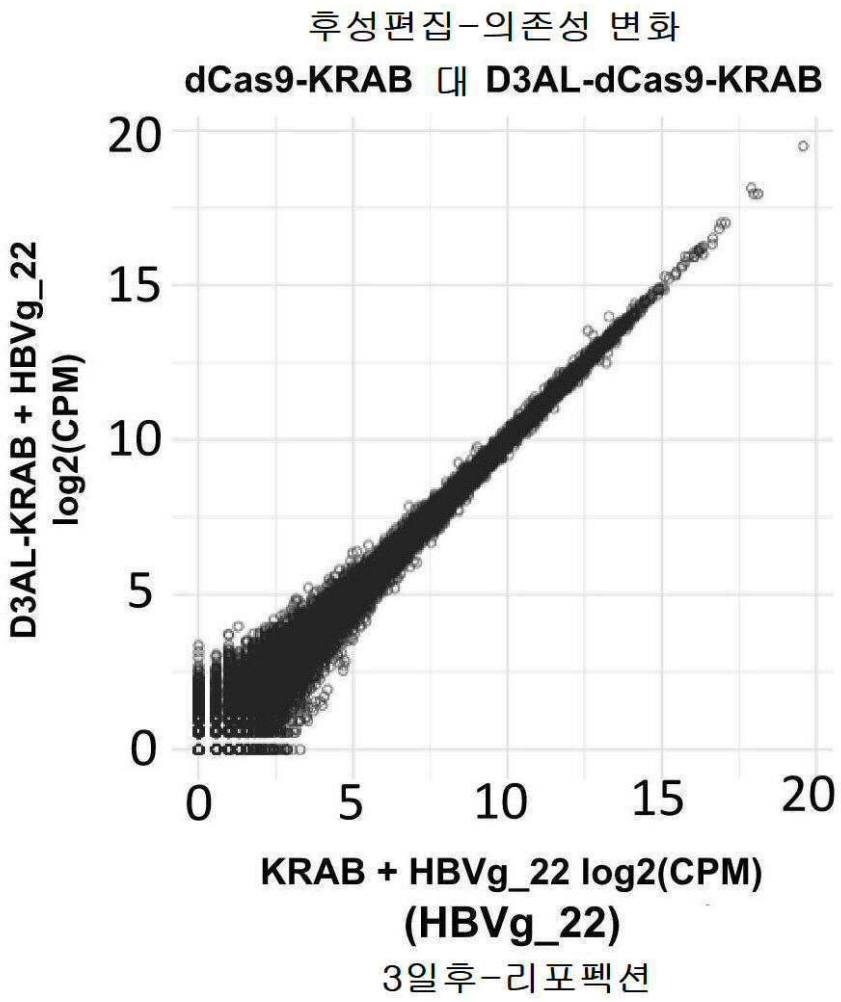
도면22


핵 추출물 중 55% cccDNA (T5exo dPCR 경유)
Day 8
Day 20
핵 HBV DNA의 %CpG 메틸화
(eg. includes encapsidated rcDNA & cccDNA)
CpG 섬 2
폴리머라제
Pre-S2
S
C
PreC
Tata-box
Pre-S1
X

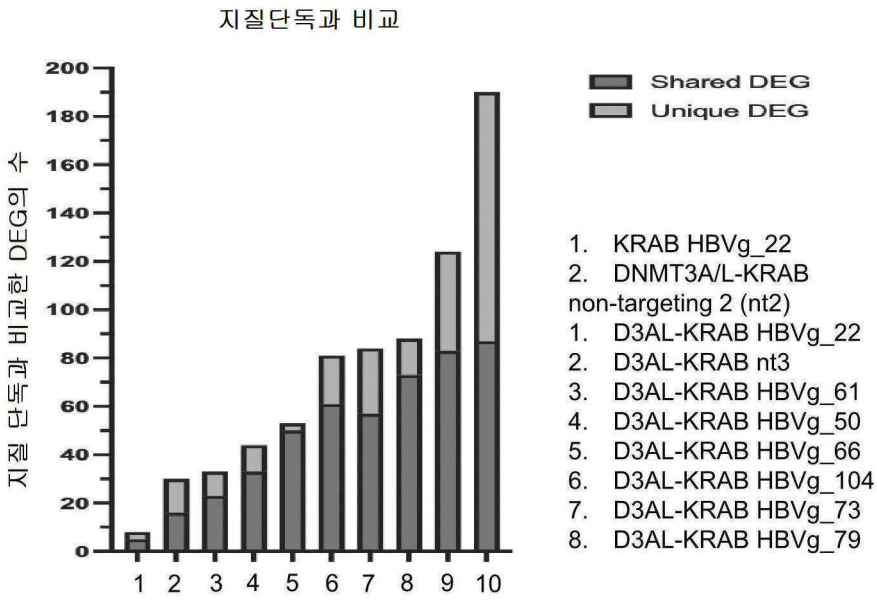
도면23a


gRNA-의존성 변화
비표적화 gRNA 대 HBVg_22
D3AL-KRAB + HBVg_22 log2(CPM)
0
5
10
15
20
D3AL-KRAB + 비표적화 gRNA log2(CPM)
(D3AL-dCas9-KRAB)
3일후-리포펙션

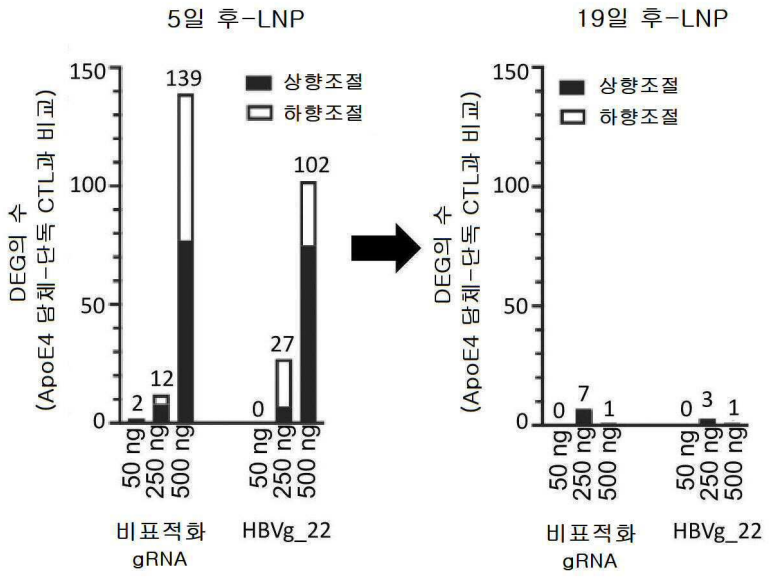
도면23b


후성편집-의존성 변화
dCas9-KRAB 대 D3AL-dCas9-KRAB
D3AL-KRAB + HBVg_22 log2(CPM)
KRAB + HBVg_22 log2(CPM)
(HBVg_22)
3일후-리포펙션

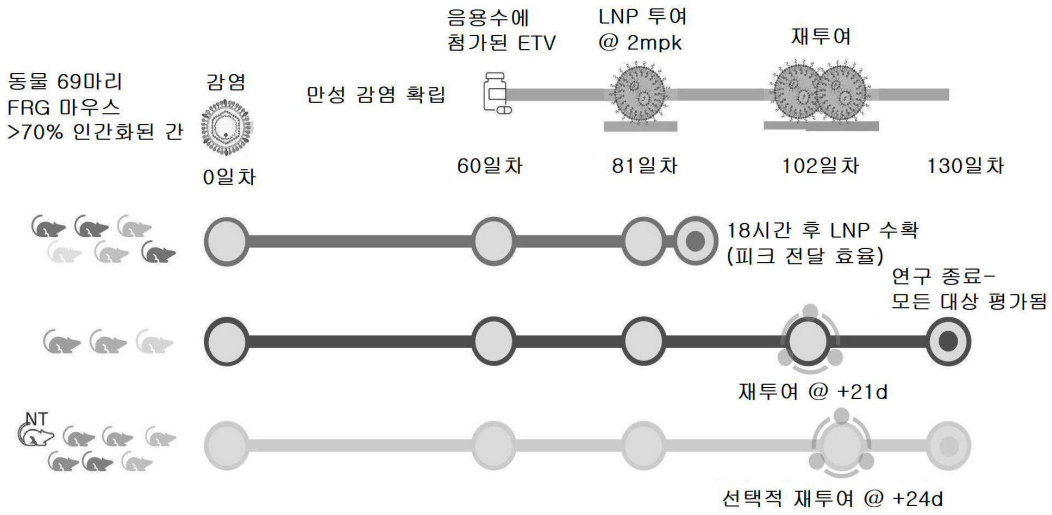
도면23c


지질단독과 비교
지질 단독과 비교한 DEG의 수
Shared DEG
Unique DEG
1. KRAB HBVg_22
2. DNMT3A/L-KRAB non-targeting 2 (nt2)
1. D3AL-KRAB HBVg_22
2. D3AL-KRAB nt3
3. D3AL-KRAB HBVg_61
4. D3AL-KRAB HBVg_50
5. D3AL-KRAB HBVg_66
6. D3AL-KRAB HBVg_104
7. D3AL-KRAB HBVg_73
8. D3AL-KRAB HBVg_79

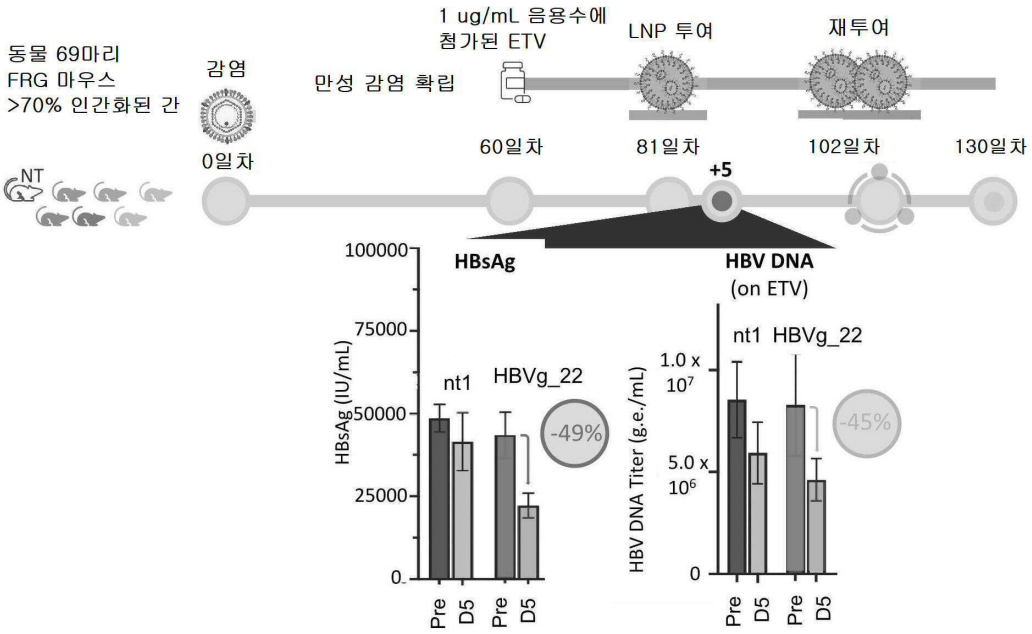
도면23d


5일 후-LNP
19일 후-LNP
DEG의 수
(ApoE4 담체-단독 CTL과 비교)
150
100
50
0
상향조절
하향조절
139
102
27
12
2
0
7
1
3
50 ng
250 ng
500 ng
비표적화 gRNA
HBVg_22

도면24


동물 69마리
FRG 마우스
>70% 인간화된 간
감염
0일차
만성 감염 확립
음용수에 첨가된 ETV
LNP 투여 @ 2mpk
재투여
60일차
81일차
102일차
130일차
18시간 후 LNP 수확 (피크 전달 효율)
연구 종료- 모든 대상 평가됨
재투여 @ +21d
NT
선택적 재투여 @ +24d

도면25



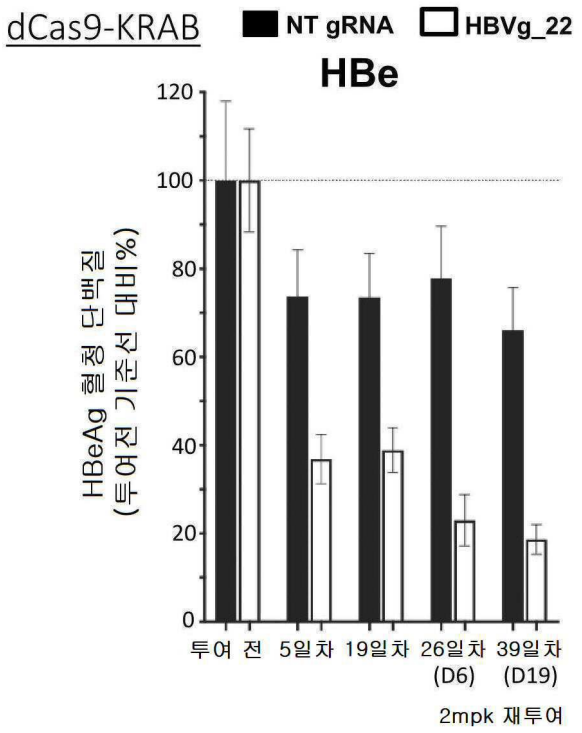
도면26

지표의 배수 변화 (투여 후/투여 전)

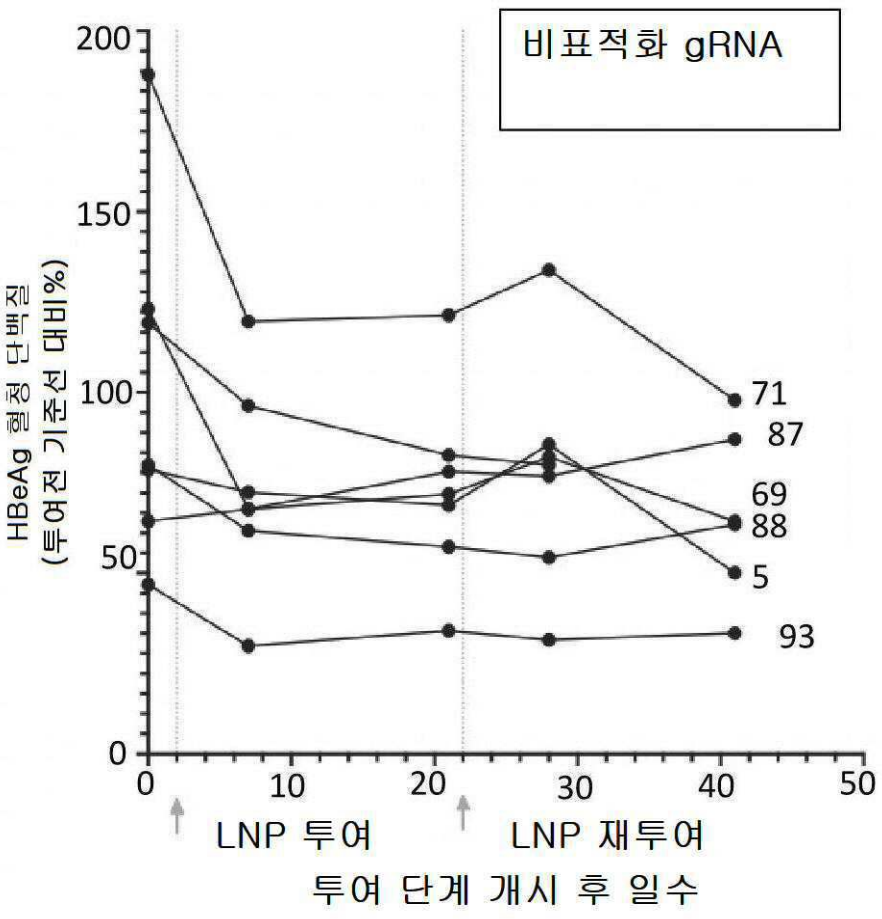
표적 서열	HBsAg	HBeAg
비표적화 gRNA #1 (nt1)	-15%	-28%
HBVg_22	- 49%***	-39%
HBVg185	+47%	-44%
HBVg_22 + HBVg_185	-31%**	-37%
KRAB + nt1	+32%	-27%
KRAB + HBVg_22	- 60%*	-64%

대응t-검정
* = p <0.05
** = p < 0.01
***= p <0.001

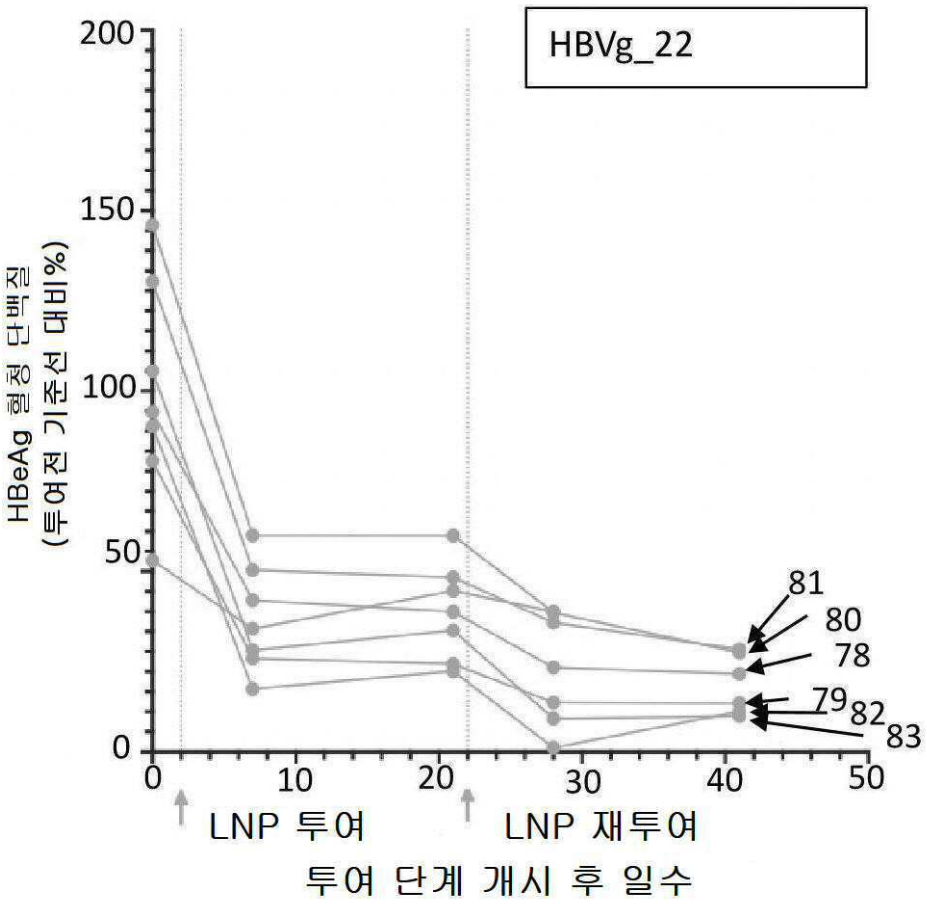
도면27a


dCas9-KRAB
NT gRNA
HBVg_22
HBe
HBeAg 혈청 단백질
(투여전 기준선 대비%)
120
100
80
60
40
20
0
투여 전
5일차
19일차
26일차
(D6)
39일차
(D19)
2mpk 재투여

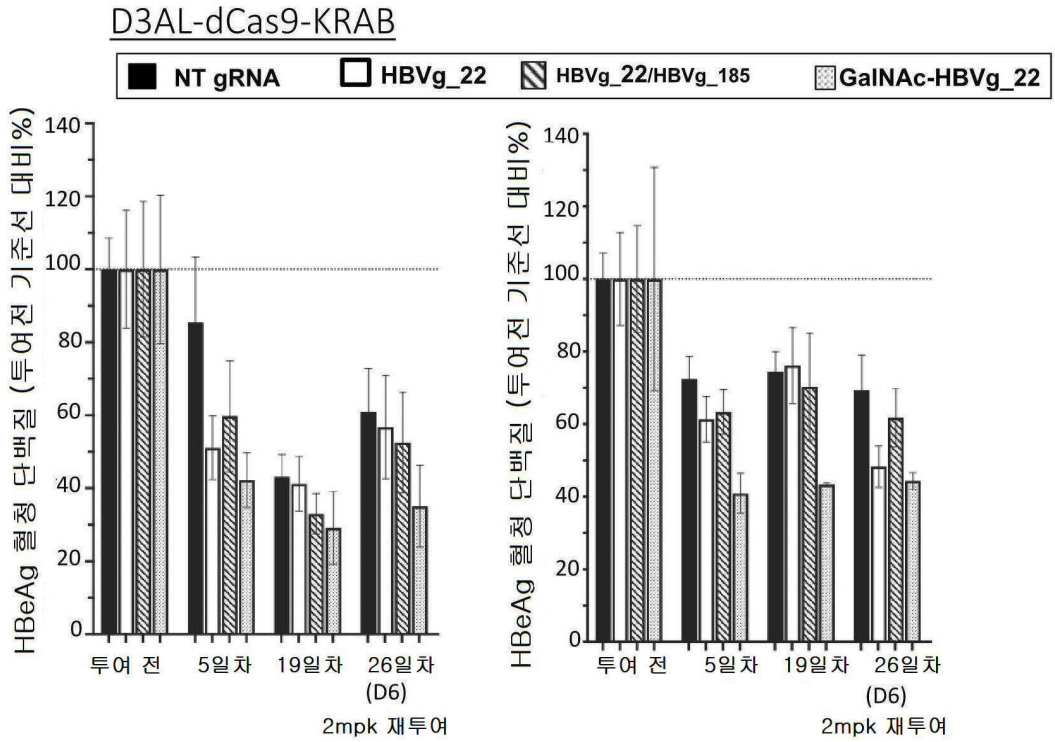
도면27b


비표적화 gRNA
HBeAg 혈청 단백질
(투여전 기준선 대비%)
200
150
100
50
0
71
87
69
88
5
93
0
10
20
30
40
50
LNP 투여
LNP 재투여
투여 단계 개시 후 일수

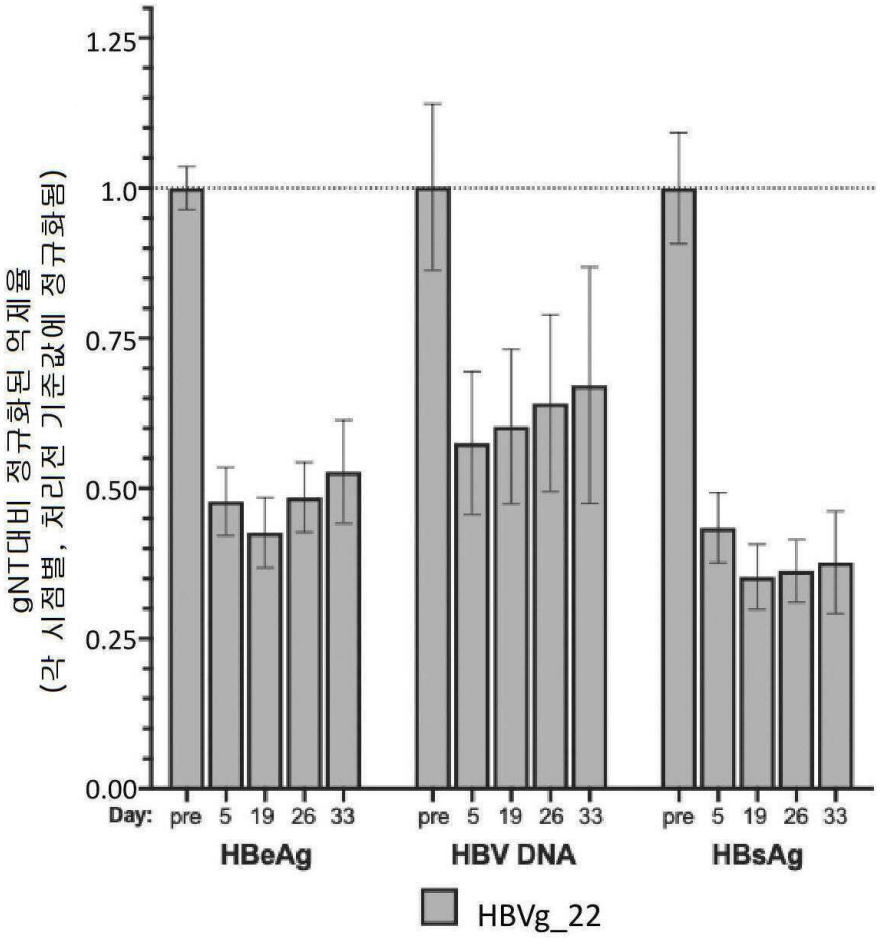
도면27c


HBVg_22
HBeAg 혈청 단백질
(투여전 기준선 대비%)
200
150
100
50
0
0
10
20
30
40
50
81
80
78
79
82
83
LNP 투여
LNP 재투여
투여 단계 개시 후 일수

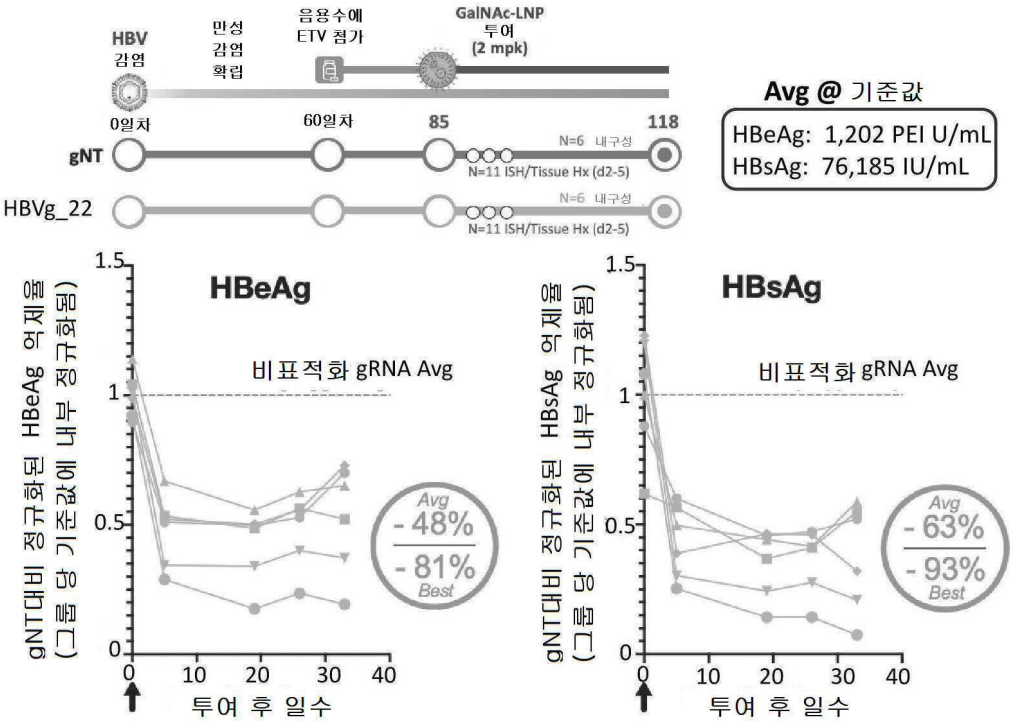
도면27d


D3AL-dCas9-KRAB
NT gRNA
HBVg_22
HBVg_22/HBVg_185
GalNAc-HBVg_22
HBeAg 혈청 단백질 (투여전 기준선 대비%)
140
120
100
80
60
40
20
0
투여 전
5일차
19일차
26일차
(D6)
2mpk 재투여
HBeAg 혈청 단백질 (투여전 기준선 대비%)
140
120
100
80
60
40
20
0
투여 전
5일차
19일차
26일차
(D6)
2mpk 재투여

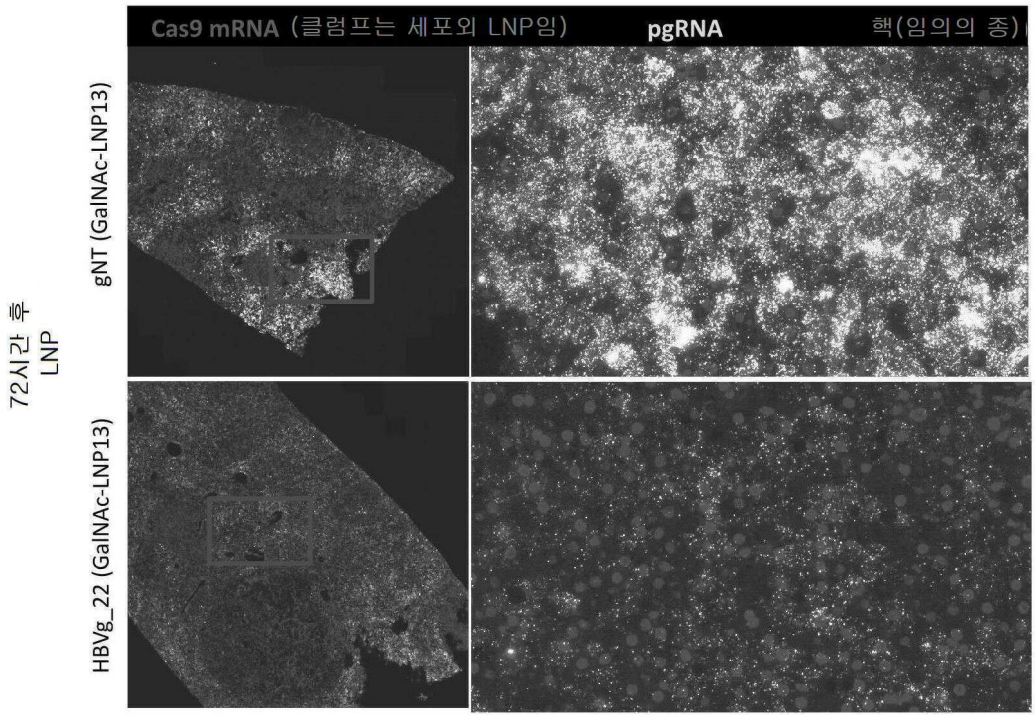
도면28a


gNT대비 정규화된 억제율
(각 시점별, 처리전 기준값에 정규화됨)
1.25
1.0
0.75
0.50
0.25
0.00
Day: pre 5 19 26 33
pre 5 19 26 33
pre 5 19 26 33
HBeAg
HBV DNA
HBsAg
HBVg_22

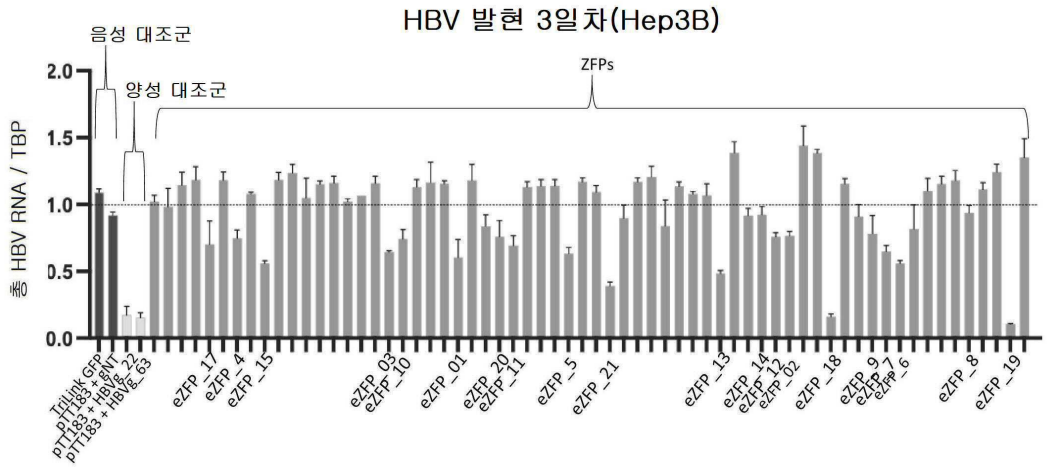
도면28b


HBV 감염
만성 감염 확립
음용수에 ETV 첨가
GalNAc-LNP 투여 (2 mpk)
0일차
60일차
85
118
gNT
HBVg_22
N=6 내구성
N=11 ISH/Tissue Hx (d2-5)
Avg @ 기준값
HBeAg: 1,202 PEI U/mL
HBsAg: 76,185 IU/mL
HBeAg
HBsAg
비표적화 gRNA Avg
gNT대비 정규화된 HBeAg 억제율 (그룹 당 기준값에 내부 정규화됨)
gNT대비 정규화된 HBsAg 억제율 (그룹 당 기준값에 내부 정규화됨)
Avg -48% -81% Best
Avg -63% -93% Best
투여 후 일수

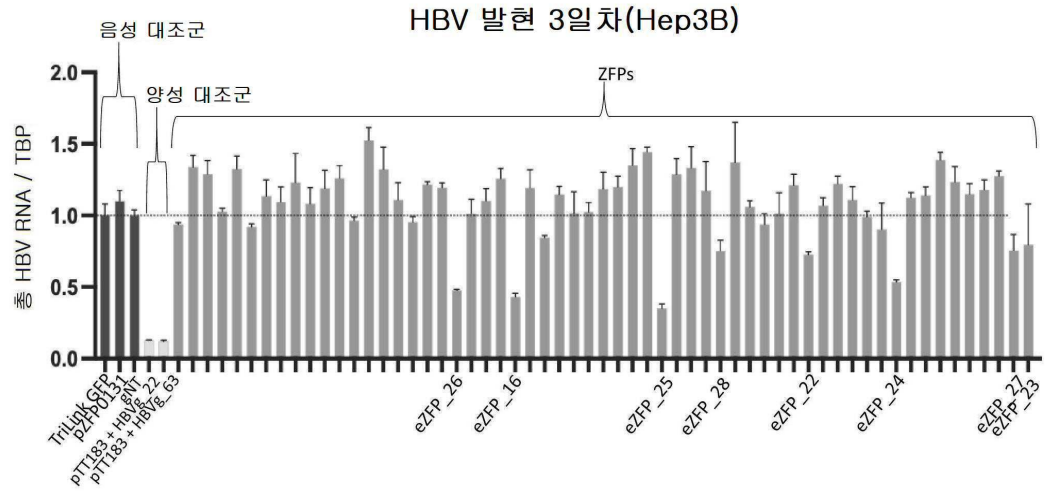
도면28c


Cas9 mRNA (클럼프는 세포외 LNP임)
pgRNA
핵(임의의 종)
72시간 후 LNP
gNT (GalNAc-LNP13)
HBVg_22 (GalNAc-LNP13)

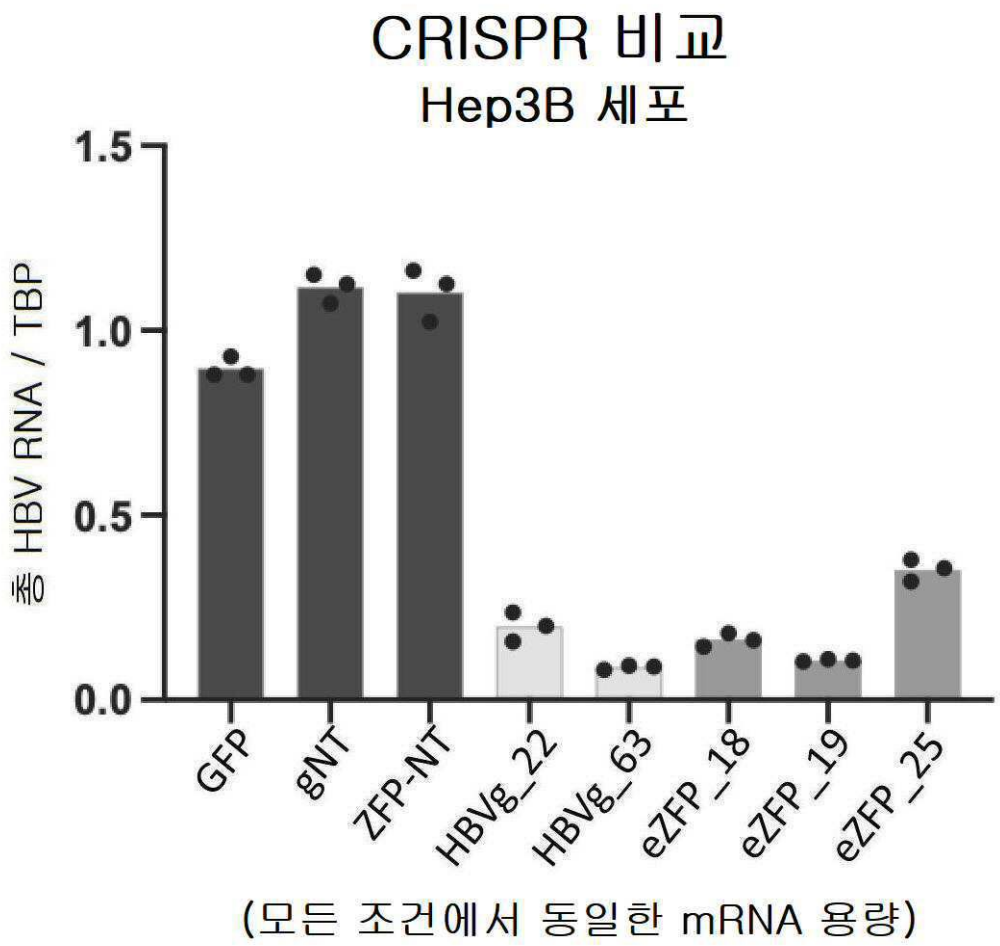
도면29a


HBV 발현 3일차(Hep3B)
음성 대조군
양성 대조군
ZFPs
총 HBV RNA / TBP
2.0
1.5
1.0
0.5
0.0
TriLink GFP
pTT183
pTT183 + gNT
pTT183 + HBVg_22
pTT183 + HBVg_63
eZFP_17
eZFP_4
eZFP_15
eZFP_03
eZFP_10
eZFP_01
eZFP_20
eZFP_11
eZFP_5
eZFP_21
eZFP_13
eZFP_14
eZFP_12
eZFP_02
eZFP_18
eZFP_9
eZFP_7
eZFP_6
eZFP_8
eZFP_19

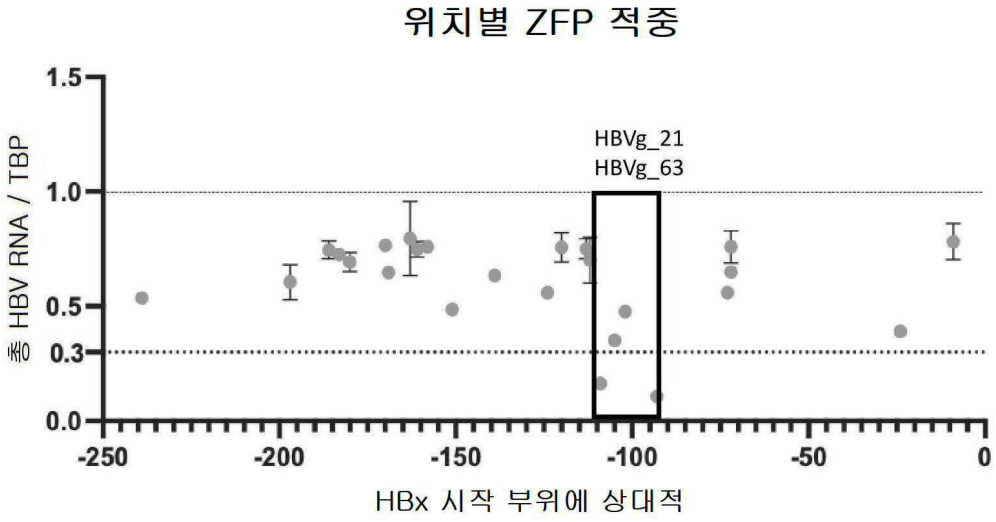
도면29b


HBV 발현 3일차(Hep3B)
음성 대조군
양성 대조군
ZFPs
총 HBV RNA / TBP
2.0
1.5
1.0
0.5
0.0
TriLink GFP
pZFpD131
pTT183 + gNT
pTT183 + HBVg_22
pTT183 + HBVg_63
eZFP_26
eZFP_16
eZFP_25
eZFP_28
eZFP_22
eZFP_24
eZFP_27
eZFP_23

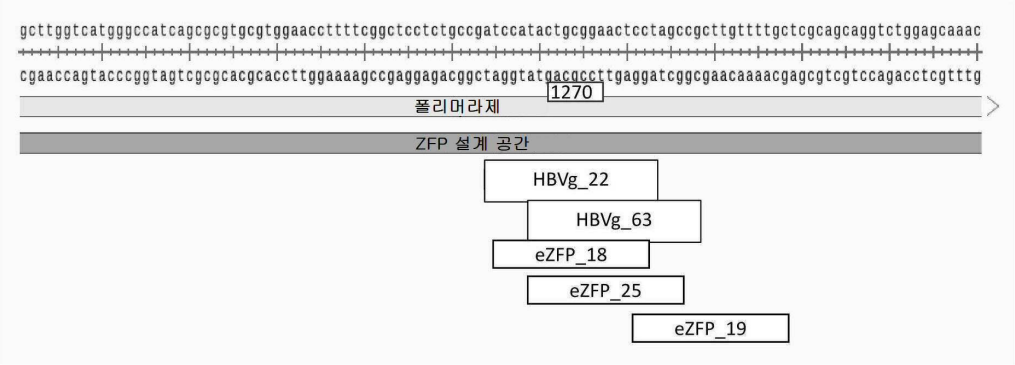
도면30


CRISPR 비교
Hep3B 세포
총 HBV RNA / TBP
1.5
1.0
0.5
0.0
GFP
gNT
ZFP-NT
HBVg_22
HBVg_63
eZFP_18
eZFP_19
eZFP_25
(모든 조건에서 동일한 mRNA 용량)

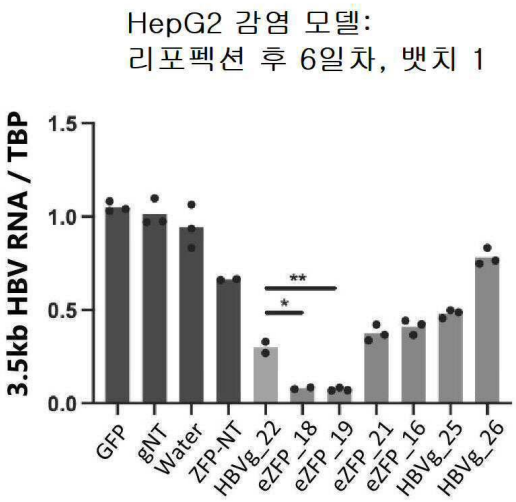
도면31a


위치별 ZFP 적중
총 HBV RNA / TBP
1.5
1.0
0.5
0.3
0.0
HBVg_21
HBVg_63
-250
-200
-150
-100
-50
0
HBx 시작 부위에 상대적

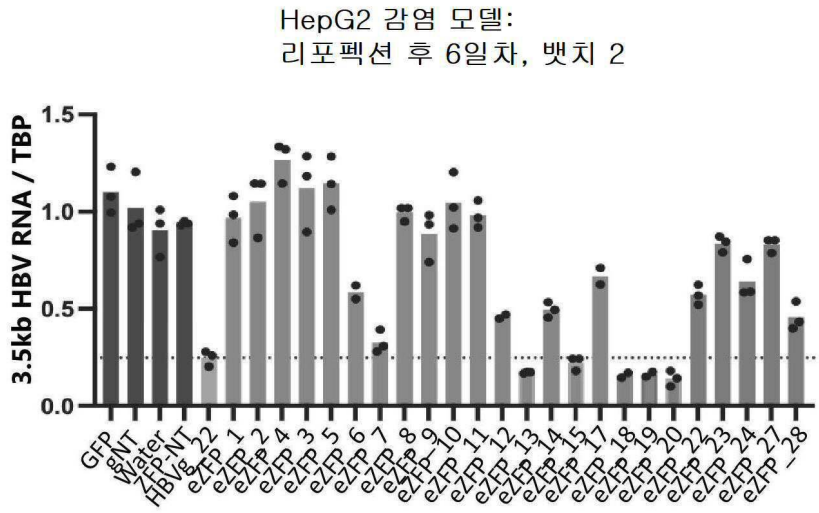
도면31b



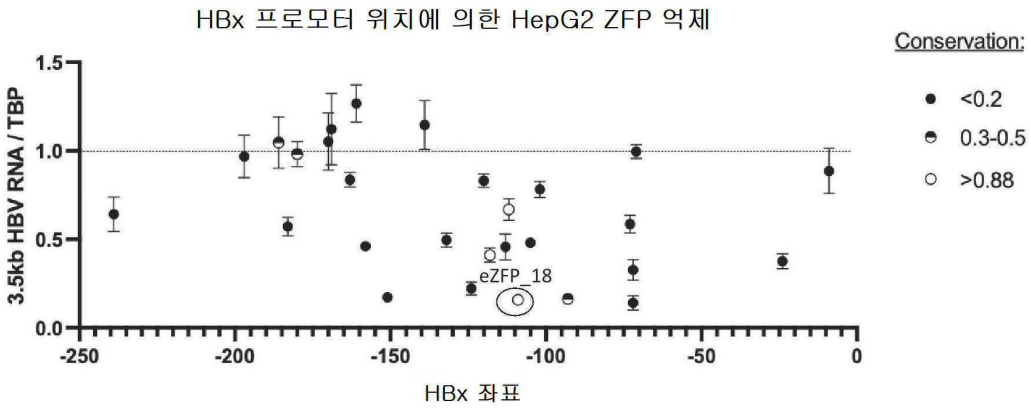
도면32a



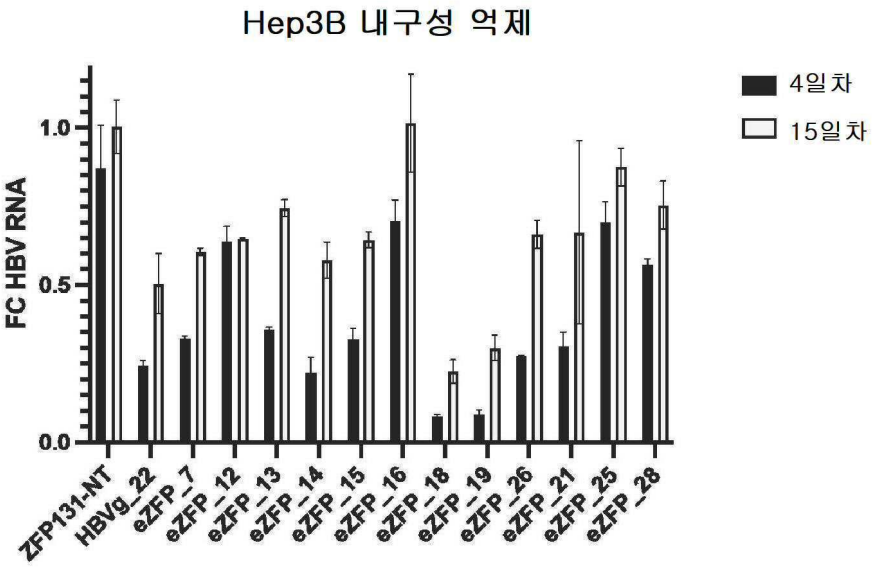
도면32b



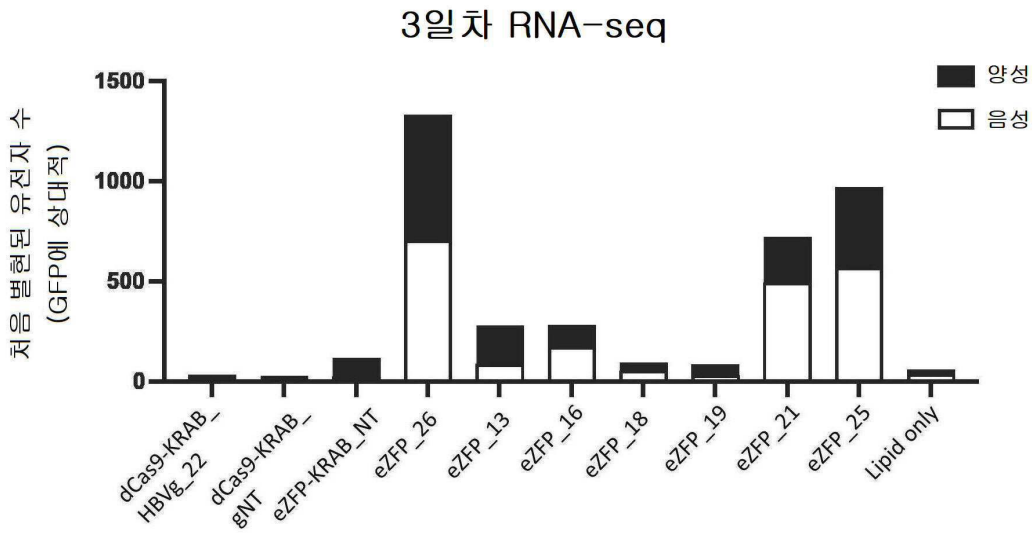
도면33


HBx 프로모터 위치에 의한 HepG2 ZFP 억제
3.5kb HBV RNA / TBP
1.5
1.0
0.5
0.0
-250
-200
-150
-100
-50
0
HBx 좌표
eZFP_18
Conservation:
<0.2
0.3-0.5
>0.88

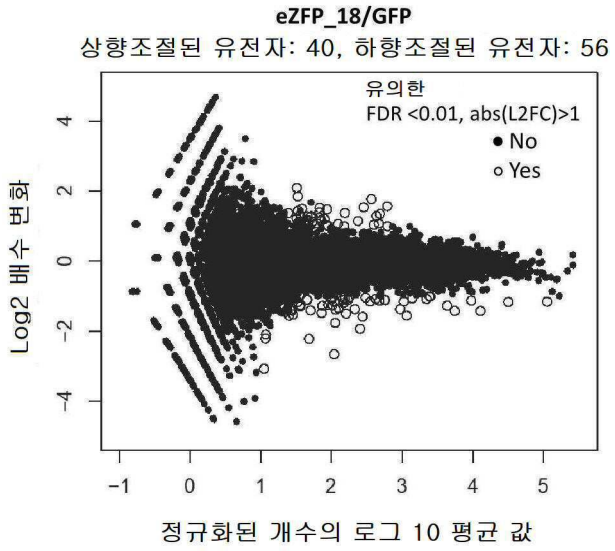
도면34


Hep3B 내구성 억제
FC HBV RNA
1.0
0.5
0.0
4일차
15일차
ZFP131-NT
HBVg_22
eZFP_7
eZFP_12
eZFP_13
eZFP_14
eZFP_15
eZFP_16
eZFP_18
eZFP_19
eZFP_26
eZFP_21
eZFP_25
eZFP_28

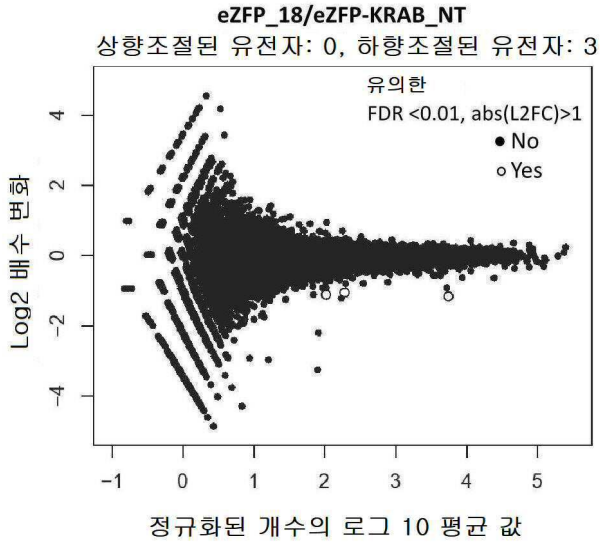
도면35


3일차 RNA-seq
양성
음성
처음 발현된 유전자 수
(GFP에 상대적)
1500
1000
500
0
dCas9-KRAB_
HBVg_22
dCas9-KRAB_
gNT
eZFP-KRAB_NT
eZFP_26
eZFP_13
eZFP_16
eZFP_18
eZFP_19
eZFP_21
eZFP_25
Lipid only

도면36a


eZFP_18/GFP
상향조절된 유전자: 40, 하향조절된 유전자: 56
유의한
FDR <0.01, abs(L2FC)>1
No
Yes
Log2 배수 변화
4
2
0
−2
−4
−1
0
1
2
3
4
5
정규화된 개수의 로그 10 평균 값

도면36b


eZFP_18/eZFP-KRAB_NT
상향조절된 유전자: 0, 하향조절된 유전자: 3
유의한
FDR <0.01, abs(L2FC)>1
No
Yes
Log2 배수 변화
4
2
0
−2
−4
−1
0
1
2
3
4
5
정규화된 개수의 로그 10 평균 값

서 열 목 록 (첨부)



아이콘을 클릭하시면 서열목록 파일이 열립니다.

본 공보 PDF는 첨부파일을 가지고 있습니다. Acrobat Reader PDF뷰어를 제공하지 않는 브라우저(크롬, 파이어폭스, 사파리 등)의 경우 첨부파일 열기가 제한되어 있으므로 Acrobat Reader PDF뷰어 설치 후 공보 PDF를 다운로드 받아 해당 뷰어에서 조회해주시기 바랍니다.