



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2024-0110603
(43) 공개일자 2024년07월15일

(51) 국제특허분류(Int. Cl.)
C12N 9/22 (2006.01) C12N 15/10 (2017.01)
C12N 15/113 (2010.01)
(52) CPC특허분류
C12N 9/22 (2013.01)
C12N 15/102 (2013.01)
(21) 출원번호 10-2024-7019181
(22) 출원일자(국제) 2022년11월23일
심사청구일자 없음
(85) 번역문제출일자 2024년06월07일
(86) 국제출원번호 PCT/US2022/080437
(87) 국제공개번호 WO 2023/097282
국제공개일자 2023년06월01일
(30) 우선권주장
63/282,999 2021년11월24일 미국(US)
(뒷면에 계속)

(71) 출원인
메타지노미, 인크.
미국 94608 캘리포니아주 에머리빌 파크 애비뉴
1545
(72) 발명자
토마스 브라이언 씨
미국 94608 캘리포니아 에머리빌 파크 애비뉴
1545
브라운 크리스토퍼
미국 94608 캘리포니아 에머리빌 파크 애비뉴
1545
(뒷면에 계속)
(74) 대리인
김진희, 김태홍

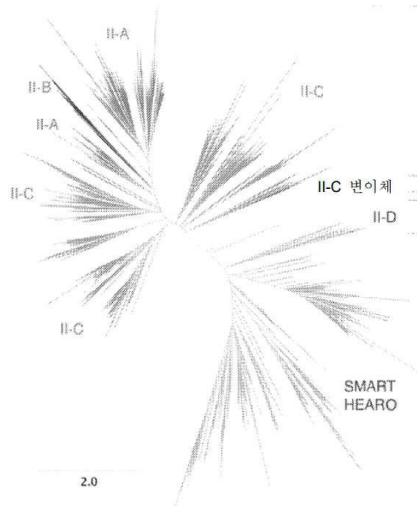
전체 청구항 수 : 총 170 항

(54) 발명의 명칭 엔도뉴클레아제 시스템

(57) 요약

본 개시는 SMART(SMall ARchaeal-associATed)(소형 고세균 연관) 뉴클레아제 시스템을 포함하는, 미배양 미생물로부터 유도된 엔도뉴클레아제 효소 시스템 및 이러한 효소 및 이의 변이체를 사용하는 방법을 제공한다.

대표도 - 도21e



(52) CPC특허분류

C12N 15/113 (2013.01)

C12N 2310/20 (2017.05)

(72) 발명자

데보토 오드라

미국 94608 캘리포니아 에머리빌 파크 애비뉴 1545

버터필드 크리스티나

미국 94608 캘리포니아 에머리빌 파크 애비뉴 1545

알렉산더 리사

미국 94608 캘리포니아 에머리빌 파크 애비뉴 1545

골츠만 다니엘라 에스에이

미국 94608 캘리포니아 에머리빌 파크 애비뉴 1545

라모테 레베카

미국 94608 캘리포니아 에머리빌 파크 애비뉴 1545

(30) 우선권주장

63/289,981 2021년12월15일 미국(US)

63/356,908 2022년06월29일 미국(US)

명세서

청구범위

청구항 1

조작된 뉴클레아제 시스템으로서,

(a) RuvC 도메인 및 HNH 도메인을 포함하는 엔도뉴클레아제이되, 상기 엔도뉴클레아제는 미배양 미생물로부터 유래되는, 엔도뉴클레아제; 및

(b) 상기 엔도뉴클레아제와 복합체를 형성하도록 구성된 조작된 가이드 리보핵산 구조체로서:

(i) 표적 데옥시리보핵산 서열과 혼성화되도록 구성된 가이드 리보핵산 서열; 및

(ii) 상기 엔도뉴클레아제에 결합하도록 구성된 리보핵산 서열을 포함하는, 조작된 가이드 리보핵산 구조체를 포함하되;

상기 엔도뉴클레아제는 약 96 kDa 이하, 약 80 kDa 이하, 약 70 kDa 이하, 또는 약 60 kDa 이하의 분자량을 가지고, 여기에서:

(1) 상기 엔도뉴클레아제는 서열번호 1-198, 221-459, 463-612, 617-668, 674-675, 975-1002, 1260-1321 중 어느 하나로부터의 아르기닌 풍부 영역 또는 PF14239 상동성을 갖는 도메인과 적어도 70%, 적어도 75%, 적어도 80%, 또는 적어도 90%의 서열 동일성을 갖는, 아르기닌 풍부 영역 또는 PF14239 상동성을 갖는 도메인, 또는 이의 변이체를 포함하거나;

(2) 상기 엔도뉴클레아제는 서열번호 1-198, 221-459, 463-612, 617-668, 674-675, 975-1002, 1260-1321 중 어느 하나로부터의 REC 도메인과 적어도 70%, 적어도 75%, 적어도 80%, 또는 적어도 90%의 서열 동일성을 갖는, REC 도메인, 또는 이의 변이체를 포함하거나;

(3) 상기 엔도뉴클레아제는 서열번호 1-198, 221-459, 463-612, 617-668, 674-675, 975-1002, 1260-1321 중 어느 하나와 적어도 70%, 적어도 75%, 적어도 80%, 또는 적어도 90%의 서열 동일성을 갖는, 서열, 또는 이의 변이체를 포함하는, 조작된 뉴클레아제 시스템.

청구항 2

제1항에 있어서, 상기 엔도뉴클레아제는 고세균 엔도뉴클레아제인, 조작된 뉴클레아제 시스템.

청구항 3

제1항 또는 제2항에 있어서, 상기 엔도뉴클레아제는 서열번호 1-198, 221-459, 463-612, 617-668, 674-675, 975-1002, 1260-1321 중 어느 하나와 적어도 70%, 적어도 75%, 적어도 80%, 또는 적어도 90%의 서열 동일성을 갖는, 서열을 포함하는, 조작된 뉴클레아제 시스템.

청구항 4

제1항 내지 제3항 중 어느 한 항에 있어서, 상기 엔도뉴클레아제는 RRxRR 모티프 또는 PF14239 상동성을 갖는 도메인을 포함하는 아르기닌-풍부 영역을 추가로 포함하는, 조작된 뉴클레아제 시스템.

청구항 5

제4항에 있어서, 상기 아르기닌 풍부 영역 또는 상기 PF14239 상동성을 갖는 도메인은 서열 번호 1-198, 221-459, 463-612, 617-668, 674-675, 975-1002, 1260-1321 중 어느 하나의 상기 아르기닌 풍부 영역 또는 상기 PF14239 상동성을 갖는 도메인과 적어도 85%, 적어도 90%, 또는 적어도 95%의 동일성을 갖는, 조작된 뉴클레아제 시스템.

청구항 6

제1항 내지 제5항 중 어느 한 항에 있어서, 상기 엔도뉴클레아제는 REC(인식) 도메인을 추가로 포함하는, 조작

된 뉴클레아제 시스템.

청구항 7

제6항에 있어서, 상기 REC 도메인은 서열번호 1-198, 221-459, 463-612, 617-668, 674-675, 975-1002, 1260-1321 중 어느 하나의 REC 도메인과 적어도 85%, 적어도 90%, 또는 적어도 95%의 동일성을 갖는, 조작된 뉴클레아제 시스템.

청구항 8

제1항 내지 제7항 중 어느 한 항에 있어서, 상기 엔도뉴클레아제는 BH(가교 나선) 도메인, WED(웨이) 도메인, 또는 PI(PAM 상호작용) 또는 TI(TAM 상호작용) 도메인을 추가로 포함하는, 조작된 뉴클레아제 시스템.

청구항 9

제8항에 있어서, 상기 BH 도메인, 상기 WED 도메인, 또는 상기 PI 도메인은 서열번호 1-198, 221-459, 463-612, 617-668, 674-675, 975-1002, 1260-1321 중 어느 하나의 BH 도메인, WED 도메인, 또는 PI 도메인과 적어도 85%, 적어도 90%, 또는 적어도 95%의 동일성을 갖는, 조작된 뉴클레아제 시스템.

청구항 10

조작된 뉴클레아제 시스템으로서:

(a) RuvC-I 도메인 및 HNH 도메인을 포함하는 엔도뉴클레아제; 및

(b) 상기 엔도뉴클레아제와 복합체를 형성하도록 구성된 조작된 가이드 리보핵산 구조체로서:

(i) 표적 테옥시리보핵산 서열과 혼성화되도록 구성된 가이드 리보핵산 서열; 및

(ii) 상기 엔도뉴클레아제에 결합하도록 구성된 리보핵산 서열을 포함하는, 조작된 가이드 리보핵산 구조체를 포함하되,

상기 엔도뉴클레아제는 서열번호 674-675, 975-1002, 1260-1321 중 어느 하나와 적어도 70%, 적어도 75%, 적어도 80%, 또는 적어도 90%의 서열 동일성을 갖는 서열, 또는 이의 변이체를 포함하는, 조작된 뉴클레아제 시스템.

청구항 11

제10항에 있어서, 상기 엔도뉴클레아제는 고세균 엔도뉴클레아제인, 조작된 뉴클레아제 시스템.

청구항 12

제10항 또는 제11항 중 어느 한 항에 있어서, 상기 엔도뉴클레아제는 RRxRR 모티프 또는 PF14239 상동성을 갖는 도메인을 포함하는 아르기닌-풍부 영역을 추가로 포함하는, 조작된 뉴클레아제 시스템.

청구항 13

제12항에 있어서, 상기 아르기닌 풍부 영역 또는 상기 PF14239 상동성을 갖는 도메인은 서열 번호 674-675, 975-1002, 1260-1321 중 어느 하나의 아르기닌 풍부 영역과 적어도 85%, 적어도 90%, 또는 적어도 95%의 동일성을 갖는, 조작된 뉴클레아제 시스템.

청구항 14

제10항 내지 제13항 중 어느 한 항에 있어서, 상기 엔도뉴클레아제는 REC(인식) 도메인을 추가로 포함하는, 조작된 뉴클레아제 시스템.

청구항 15

제14항에 있어서, 상기 REC 도메인은 서열번호 674-675, 975-1002, 1260-1321 중 어느 하나의 REC 도메인과 적어도 85%, 적어도 90%, 또는 적어도 95%의 동일성을 갖는, 조작된 뉴클레아제 시스템.

청구항 16

제10항 내지 제15항 중 어느 한 항에 있어서, 상기 엔도뉴클레아제는 BH 도메인, WED 도메인, 및 PI 도메인을 추가로 포함하는, 조작된 뉴클레아제 시스템.

청구항 17

제16항에 있어서, 상기 BH 도메인, 상기 WED 도메인, 또는 상기 PI 도메인은 서열번호 674-675, 975-1002, 1260-1321 중 어느 하나의 BH 도메인, WED 도메인, 또는 PI 도메인과 적어도 85%, 적어도 90%, 또는 적어도 95%의 동일성을 갖는, 조작된 뉴클레아제 시스템.

청구항 18

제10항 내지 제17항 중 어느 한 항에 있어서, 상기 엔도뉴클레아제는 미배양 미생물로부터 유래되는, 조작된 뉴클레아제 시스템.

청구항 19

제1항 내지 제18항 중 어느 한 항에 있어서, 상기 엔도뉴클레아제에 결합하도록 구성된 상기 리보핵산 서열은 서열번호 199-200, 460-461, 또는 669-673 중 어느 하나와 적어도 80%의 서열 동일성을 갖는 서열, 또는 서열번호 201-203, 613-616, 677-686, 1003-1022, 또는 1231-1259 중 어느 하나의 비축퇴 뉴클레오티드와 적어도 80%의 서열 동일성을 갖는 서열을 포함하는, 조작된 뉴클레아제 시스템.

청구항 20

제19항에 있어서, 상기 가이드 핵산 구조체는 서열번호 201-203, 613-616, 677-686, 1003-1022, 또는 1231-1259 중 어느 하나의 비축퇴 뉴클레오티드와 적어도 80%의 동일성을 갖는 서열을 포함하는, 조작된 뉴클레아제 시스템.

청구항 21

조작된 뉴클레아제 시스템으로서,

(a) 조작된 가이드 리보핵산 구조체이되;

(i) 표적 테옥시리보핵산 서열과 혼성화되도록 구성된 가이드 리보핵산 서열; 및

(ii) 엔도뉴클레아제에 결합하도록 구성된 리보핵산 서열을 포함하며,

여기에서, 상기 리보핵산 서열은 서열번호 199-200, 460-461, 또는 669-673 중 어느 하나와 적어도 80%의 서열 동일성을 갖는 서열, 또는 서열번호 677-686, 1006-1012, 또는 1231-1259 중 어느 하나의 비가변 뉴클레오티드와 적어도 80%의 서열 동일성을 갖는 서열을 포함하는, 조작된 가이드 리보핵산 구조체; 및

(b) 상기 조작된 가이드 리보핵산에 결합하도록 구성된 RNA 가이드 엔도뉴클레아제를 포함하는, 조작된 뉴클레아제 시스템.

청구항 22

제21항에 있어서, 상기 RNA 가이드 엔도뉴클레아제는 고세균 엔도뉴클레아제인, 조작된 뉴클레아제 시스템.

청구항 23

제21항 또는 제22항에 있어서, 상기 엔도뉴클레아제는 약 120 kDa 이하, 100 kDa 이하, 90 kDa 이하, 또는 60 kDa 이하의 분자량을 갖는, 조작된 뉴클레아제 시스템.

청구항 24

제1항 내지 제23항 중 어느 한 항에 있어서, 상기 조작된 가이드 리보핵산 구조체는 적어도 2개의 리보핵산 폴리뉴클레오티드를 포함하는, 조작된 뉴클레아제 시스템.

청구항 25

제1항 내지 제24항 중 어느 한 항에 있어서, 상기 조작된 가이드 리보핵산 구조체는 상기 가이드 리보핵산 서열

및 상기 엔도뉴클레아제와 결합하도록 구성된 상기 리보핵산 서열을 포함하는 하나의 리보핵산 폴리뉴클레오티드를 포함하는, 조작된 뉴클레아제 시스템.

청구항 26

제1항 내지 제25항 중 어느 한 항에 있어서, 상기 가이드 리보핵산 서열은 원핵, 박테리아, 고세균, 진핵, 진균, 식물, 포유류, 또는 인간 게놈 서열에 상보적인, 조작된 뉴클레아제 시스템.

청구항 27

제1항 내지 제26항 중 어느 한 항에 있어서, 상기 가이드 리보핵산 서열은 약 14 내지 약 28개 뉴클레오티드 길이, 약 18 내지 약 26개 뉴클레오티드 길이, 약 22 내지 약 26개 뉴클레오티드 길이, 또는 약 24개 뉴클레오티드 길이인, 조작된 뉴클레아제 시스템.

청구항 28

제1항 내지 제27항 중 어느 한 항에 있어서, 상기 가이드 리보핵산 서열은 서열번호 462, 676, 또는 1229-1230 중 어느 하나와 적어도 80%의 서열 동일성을 갖는 서열을 포함하는, 조작된 뉴클레아제 시스템.

청구항 29

제1항 내지 제28항 중 어느 한 항에 있어서, 상기 엔도뉴클레아제가 상기 엔도뉴클레아제의 N-말단 또는 C-말단에 근접한 하나 이상의 핵 국소화 서열(NLS)을 포함하는, 조작된 뉴클레아제 시스템.

청구항 30

제1항 내지 제29항 중 어느 한 항에 있어서, 상기 NLS는 서열번호 205-220 중 어느 하나로부터 선택되는 서열을 포함하는, 조작된 뉴클레아제 시스템.

청구항 31

제1항 내지 제30항 중 어느 한 항에 있어서,

단일 또는 이중 가닥 DNA 복구 템플릿으로서, 5'에서 3'까지: 상기 표적 데옥시리보핵산 서열에 대해 5'에 있는 적어도 20개 뉴클레오티드의 서열을 포함하는 제1 상동 아암, 적어도 10개 뉴클레오티드의 합성 DNA 서열, 및 상기 표적 서열에 대해 3'에 있는 적어도 20개 뉴클레오티드의 서열을 포함하는 제2 상동 아암을 포함하는, 단일 또는 이중 가닥 DNA 복구 템플릿을 추가로 포함하는, 조작된 뉴클레아제 시스템.

청구항 32

제31항에 있어서, 상기 제1 또는 제2 상동 아암은 적어도 40, 80, 120, 150, 200, 300, 500, 또는 1,000개 뉴클레오티드의 서열을 포함하는, 조작된 뉴클레아제 시스템.

청구항 33

제1항 내지 제32항 중 어느 한 항에 있어서, 상기 시스템은 Mg^{2+} 의 공급원을 추가로 포함하는, 조작된 뉴클레아제 시스템.

청구항 34

제1항 내지 제33항 중 어느 한 항에 있어서, 상기 엔도뉴클레아제 및 상기 엔도뉴클레아제에 결합하도록 구성된 상기 리보핵산 서열은 동일한 문 내의 구별되는 종으로부터 유래되는, 조작된 뉴클레아제 시스템.

청구항 35

제1항 내지 제34항 중 어느 한 항에 있어서, 상기 엔도뉴클레아제는 서열번호 2-24 중 어느 하나와 적어도 70%의 서열 동일성을 갖는 서열을 포함하고, 상기 가이드 RNA 구조체는 줄기 및 루프를 포함하는 헤어핀을 포함하도록 예측된 RNA 서열을 포함하며, 여기에서 상기 줄기는 적어도 10 쌍의 리보뉴클레오티드 및 개재 다중루프를 포함하는, 조작된 뉴클레아제 시스템.

청구항 36

제35항에 있어서, 상기 가이드 RNA 구조체는 제2 줄기 및 제2 루프를 추가로 포함하되, 상기 제2 줄기는 적어도 5 쌍의 리보뉴클레오타이드를 포함하는, 조작된 엔도뉴클레아제 시스템.

청구항 37

제35항에 있어서, 상기 가이드 RNA 구조체는 적어도 2개의 헤어핀을 포함하는 RNA 구조체를 추가로 포함하는, 조작된 엔도뉴클레아제 시스템.

청구항 38

제1항 내지 제37항 중 어느 한 항에 있어서:

- a) 상기 엔도뉴클레아제는 서열번호 1-198, 221-459, 463-612, 617-668, 674-675, 975-1002, 1260-1321 중 어느 하나와 적어도 70%, 적어도 80%, 또는 적어도 90% 동일한 서열, 또는 이의 변이체를 포함하고;
- b) 상기 가이드 RNA 구조체는 서열번호 199-200, 460-461, 또는 669-673 중 어느 하나, 또는 서열번호 201-203, 613-616, 677-686, 1006-1012, 또는 1231-1259 중 어느 하나의 비가변 뉴클레오타이드와 적어도 70%, 적어도 80%, 또는 적어도 90% 동일한 서열을 포함하는, 조작된 뉴클레아제 시스템.

청구항 39

제1항 내지 제37항 중 어느 한 항에 있어서:

- a) 상기 엔도뉴클레아제는 서열번호 1-198, 221-459, 463-612, 617-668, 674-675, 975-1002, 1260-1321 중 어느 하나와 적어도 70%, 적어도 80%, 또는 적어도 90% 동일한 서열을 포함하고;
- b) 상기 가이드 RNA 구조체는 클래스 2, II형 sgRNA 또는 tracr 서열과 적어도 70%, 적어도 80%, 또는 적어도 90% 동일한 서열을 포함하는, 조작된 뉴클레아제 시스템.

청구항 40

제1항 내지 제39항 중 어느 한 항에 있어서, 상기 서열 동일성은 BLASTP, CLUSTALW, MUSCLE, MAFFT, 또는 Smith-Waterman 상동성 검색 알고리즘의 파라미터를 사용하는 CLUSTALW에 의해 결정되는, 조작된 뉴클레아제 시스템.

청구항 41

제40항에 있어서, 상기 서열 동일성은, 단어 길이(W) 3, 기대치(E) 10의 파라미터, 및 BLOSUM62 스코어링 매트릭스(존재 11, 연장 1의 갭 비용 설정)를 사용하고, 조건부 조성 스코어 매트릭스 조정을 사용하는, 상기 BLASTP 상동성 검색 알고리즘에 의해 결정되는, 조작된 뉴클레아제 시스템.

청구항 42

제1항 내지 제41항 중 어느 한 항에 있어서, 상기 엔도뉴클레아제는 Cas9 엔도뉴클레아제, Cas14 엔도뉴클레아제, Cas12a 엔도뉴클레아제, Cas12b 엔도뉴클레아제, Cas 12c 엔도뉴클레아제, Cas12d 엔도뉴클레아제, Cas12e 엔도뉴클레아제, Cas13a 엔도뉴클레아제, Cas13b 엔도뉴클레아제, Cas13c 엔도뉴클레아제, 또는 Cas13d 엔도뉴클레아제가 아닌, 조작된 뉴클레아제 시스템.

청구항 43

제42항에 있어서, 상기 엔도뉴클레아제는 Cas9 엔도뉴클레아제에 대해 80% 미만의 동일성을 갖는, 조작된 뉴클레아제 시스템.

청구항 44

조작된 뉴클레아제 시스템으로서:

- (a) ANGG(서열번호 1029), NARAA(서열번호 1030), ATGAAA(서열번호 1031), ATGA(서열번호 1032), 또는 WTGG(서열번호 1033) 중 어느 하나를 포함하는 표적 인접 모티프(TAM) 서열에 대해 선택성을 갖도록 구성된 엔도뉴클레

아제이되, 상기 엔도뉴클레아제는 서열번호 1-198, 221-459, 463-612, 617-668, 674-675, 975-1002, 1260-1321 중 어느 하나의 TAM 상호작용 도메인과 적어도 75%, 80%, 또는 90%의 동일성을 갖는 TAM 상호작용 도메인을 포함하는, 엔도뉴클레아제; 및

(b) 조작된 가이드 RNA이되, 상기 조작된 가이드 RNA는 상기 엔도뉴클레아제와 복합체를 형성하도록 구성되고, 상기 조작된 가이드 RNA는 표적 핵산 서열에 혼성화되도록 구성된 스페이서 서열을 포함하는, 조작된 가이드 RNA를 포함하는, 조작된 뉴클레아제 시스템.

청구항 45

제44항에 있어서, 상기 TAM 상호작용 도메인은 서열번호 674의 TAM 상호작용 도메인과 적어도 80%의 서열 동일성을 갖는 서열 또는 이의 변이체, 또는 서열번호 675의 TAM 상호작용 도메인과 적어도 80%의 서열 동일성을 갖는 서열 또는 이의 변이체를 포함하는, 조작된 뉴클레아제 시스템.

청구항 46

제44항 또는 제45항 중 어느 한 항에 있어서, 상기 엔도뉴클레아제 시스템은 진핵생물, 진균류, 식물, 포유류, 또는 인간 게놈 폴리뉴클레오티드 서열에 상보적인 서열을 포함하는, 조작된 뉴클레아제 시스템.

청구항 47

제44항 내지 제46항 중 어느 한 항에 있어서, 상기 가이드 RNA는 30-280개 뉴클레오티드 길이인, 조작된 뉴클레아제 시스템.

청구항 48

제44항 내지 제47항 중 어느 한 항에 있어서,

단일 또는 이중 가닥 DNA 복구 템플릿으로서, 5'에서 3'까지: 상기 표적 테옥시리보핵산 서열에 대해 5'에 있는 적어도 20개 뉴클레오티드의 서열을 포함하는 제1 상동 아암, 적어도 10개 뉴클레오티드의 합성 DNA 서열, 및 상기 표적 서열에 대해 3'에 있는 적어도 20개 뉴클레오티드의 서열을 포함하는 제2 상동 아암을 포함하는, 단일 또는 이중 가닥 DNA 복구 템플릿을 추가로 포함하는, 조작된 뉴클레아제 시스템.

청구항 49

제48항에 있어서, 상기 제1 또는 제2 상동 아암은 적어도 40개 뉴클레오티드의 서열을 포함하는, 조작된 뉴클레아제 시스템.

청구항 50

제48항 또는 제49항에 있어서, 상기 제1 및 제2 상동 아암은 진핵생물의 게놈 서열과 상동인, 조작된 뉴클레아제 시스템.

청구항 51

제48항 내지 제50항 중 어느 한 항에 있어서, 상기 단일-가닥 또는 이중-가닥 DNA 복구 템플릿은 이식유전자 공여자를 포함하는, 조작된 뉴클레아제 시스템.

청구항 52

제44항 내지 제51항 중 어느 한 항에 있어서, 1개 또는 2개의 단일-가닥 DNA 분절이 측면에 위치한 이중-가닥 DNA 분절을 포함하는 DNA 복구 템플릿을 추가로 포함하는, 조작된 뉴클레아제 시스템.

청구항 53

제52항에 있어서, 상기 단일-가닥 DNA 분절이 상기 이중-가닥 DNA 분절의 5' 말단에 접합되는, 조작된 뉴클레아제 시스템.

청구항 54

제52항에 있어서, 상기 단일-가닥 DNA 분절이 상기 이중-가닥 DNA 분절의 3' 말단에 접합되는, 조작된 뉴클레아제 시스템.

제 시스템.

청구항 55

제52항 내지 제54항 중 어느 한 항에 있어서, 상기 단일-가닥 DNA 분절이 4개 내지 10개 뉴클레오티드 염기의 길이를 갖는, 조작된 뉴클레아제 시스템.

청구항 56

제52항 내지 제55항 중 어느 한 항에 있어서, 상기 단일-가닥 DNA 분절이 상기 스페이서 서열 내의 서열에 상보적인 뉴클레오티드 서열을 갖는, 조작된 뉴클레아제 시스템.

청구항 57

제52항 내지 제56항 중 어느 한 항에 있어서, 상기 이중-가닥 DNA 서열이 바코드, 개방 해독 프레임, 인핸서, 프로모터, 단백질-코딩 서열, miRNA 코딩 서열, RNA 코딩 서열, 또는 이식유전자를 포함하는, 조작된 뉴클레아제 시스템.

청구항 58

제52항 내지 제57항 중 어느 한 항에 있어서, 상기 이중-가닥 DNA 서열이 뉴클레아제 절단 부위가 측면에 위치하는, 조작된 뉴클레아제 시스템.

청구항 59

조작된 뉴클레아제 시스템으로서:

- (a) NRR을 포함하는 프로토스페이서 인접 모티프(PAM) 서열에 대해 선택성을 갖도록 구성된 엔도뉴클레아제이되, 상기 엔도뉴클레아제는 서열번호 1313-1318 중 어느 하나의 PAM 상호작용 도메인과 적어도 75%, 80%, 또는 90%의 동일성을 갖는 PAM 상호작용 도메인을 포함하는, 엔도뉴클레아제; 및
- (b) 조작된 가이드 RNA이되, 상기 조작된 가이드 RNA는 상기 엔도뉴클레아제와 복합체를 형성하도록 구성되고, 상기 조작된 가이드 RNA는 표적 핵산 서열에 혼성화되도록 구성된 스페이서 서열을 포함하는, 조작된 가이드 RNA를 포함하는, 조작된 뉴클레아제 시스템.

청구항 60

제59항에 있어서, 상기 TAM 상호작용 도메인은 서열번호 674의 TAM 상호작용 도메인과 적어도 80%의 서열 동일성을 갖는 서열 또는 이의 변이체, 또는 서열번호 675의 TAM 상호작용 도메인과 적어도 80%의 서열 동일성을 갖는 서열 또는 이의 변이체를 포함하는, 조작된 뉴클레아제 시스템.

청구항 61

제59항 또는 제60항 중 어느 한 항에 있어서, 상기 엔도뉴클레아제 시스템은 진핵생물, 진균류, 식물, 포유류, 또는 인간 계놈 폴리뉴클레오티드 서열에 상보적인 서열을 포함하는, 조작된 뉴클레아제 시스템.

청구항 62

제59항 내지 제61항 중 어느 한 항에 있어서, 상기 가이드 RNA는 30-280개 뉴클레오티드 길이인, 조작된 뉴클레아제 시스템.

청구항 63

제59항 내지 제62항 중 어느 한 항에 있어서,

단일 또는 이중 가닥 DNA 복구 템플릿으로서, 5'에서 3'까지: 상기 표적 데옥시리보핵산 서열에 대해 5'에 있는 적어도 20개 뉴클레오티드의 서열을 포함하는 제1 상동 아암, 적어도 10개 뉴클레오티드의 합성 DNA 서열, 및 상기 표적 서열에 대해 3'에 있는 적어도 20개 뉴클레오티드의 서열을 포함하는 제2 상동 아암을 포함하는, 단일 또는 이중 가닥 DNA 복구 템플릿을 추가로 포함하는, 조작된 뉴클레아제 시스템.

청구항 64

제63항에 있어서, 상기 제1 또는 제2 상동 아암은 적어도 40개 뉴클레오타이드의 서열을 포함하는, 조작된 뉴클레아제 시스템.

청구항 65

제63항 또는 제64항에 있어서, 상기 제1 및 제2 상동 아암은 진핵생물의 게놈 서열과 상동인, 조작된 뉴클레아제 시스템.

청구항 66

제59항 내지 제65항 중 어느 한 항에 있어서, 상기 단일-가닥 또는 이중-가닥 DNA 복구 템플릿은 이식유전자 공여자를 포함하는, 조작된 뉴클레아제 시스템.

청구항 67

제59항 내지 제66항 중 어느 한 항에 있어서, 1개 또는 2개의 단일-가닥 DNA 분절이 측면에 위치한 이중-가닥 DNA 분절을 포함하는 DNA 복구 템플릿을 추가로 포함하는, 조작된 뉴클레아제 시스템.

청구항 68

제67항에 있어서, 상기 단일-가닥 DNA 분절은 상기 이중-가닥 DNA 분절의 5' 말단에 접합되는, 조작된 뉴클레아제 시스템.

청구항 69

제67항에 있어서, 상기 단일-가닥 DNA 분절은 상기 이중-가닥 DNA 분절의 3' 말단에 접합되는, 조작된 뉴클레아제 시스템.

청구항 70

제67항 내지 제69항 중 어느 한 항에 있어서, 상기 단일-가닥 DNA 분절은 4 내지 10개 뉴클레오타이드 염기의 길이를 갖는, 조작된 뉴클레아제 시스템.

청구항 71

제67항 내지 제70항 중 어느 한 항에 있어서, 상기 단일-가닥 DNA 분절은 상기 스페이서 서열 내의 서열에 상보적인 뉴클레오타이드 서열을 갖는, 조작된 뉴클레아제 시스템.

청구항 72

제67항 내지 제71항 중 어느 한 항에 있어서, 상기 이중-가닥 DNA 서열은 바코드, 개방 해독 프레임, 인핸서, 프로모터, 단백질-코딩 서열, miRNA 코딩 서열, RNA 코딩 서열, 또는 이식유전자를 포함하는, 조작된 뉴클레아제 시스템.

청구항 73

제67항 내지 제72항 중 어느 한 항에 있어서, 상기 이중-가닥 DNA 서열은 뉴클레아제 절단 부위가 측면에 위치하는, 조작된 뉴클레아제 시스템.

청구항 74

조작된 단일 가이드 리보핵산 폴리뉴클레오타이드로서:

- a) 표적 DNA 분자 내의 표적 서열에 상보적인 뉴클레오타이드 서열을 포함하는, DNA-표적화 분절; 및
 - b) 이중-가닥 RNA(dsRNA) 이중체를 형성하도록 혼성화되는 뉴클레오타이드의 2개의 상보적 신장부를 포함하는, 단백질-결합 분절을 포함하되,
- 상기 뉴클레오타이드의 2개의 상보적 신장부는 개재 뉴클레오타이드와 서로 공유적으로 연결되고,

상기 조작된 가이드 리보핵산 폴리뉴클레오타이드는 서열번호 674-675, 975-1002, 1260-1321 중 어느 하나와 적어도 75%의 서열 동일성을 갖는 변이체를 포함하는 엔도뉴클레아제, 또는 이의 변이체와 복합체를 형성하도록 구성되는, 조작된 단일 가이드 리보핵산 폴리뉴클레오타이드.

청구항 75

제74항에 있어서, 상기 DNA-표적화 분절은 뉴클레오타이드의 상기 2개의 상보적 신장부 둘 모두의 5'에 위치되는, 조작된 가이드 리보핵산 폴리뉴클레오타이드.

청구항 76

제74항 또는 제75항 중 어느 한 항에 있어서:

- a) 상기 단백질 결합 분절은 서열번호 199-200, 460-461, 또는 669-673 중 어느 하나와 적어도 70%, 적어도 80%, 또는 적어도 90%의 동일성을 갖는 서열을 포함하거나;
- b) 상기 단백질 결합 분절은 서열번호 201-203, 613-616, 677-686, 1003-1022, 또는 1231-1259 중 어느 하나의 비가변 뉴클레오타이드와 적어도 70%, 적어도 80%, 또는 적어도 90%의 동일성을 갖는 서열을 포함하는, 조작된 가이드 리보핵산 폴리뉴클레오타이드.

청구항 77

제74항 또는 제75항 중 어느 한 항에 있어서:

- a) 상기 엔도뉴클레아제는 서열번호 674-675, 975-1002, 1260-1321 중 어느 하나와 적어도 70%, 적어도 80%, 또는 적어도 90% 동일한 서열, 또는 이의 변이체를 포함하고;
- b) 상기 가이드 RNA 구조체는 클래스 2, II형 sgRNA와 적어도 70%, 적어도 80%, 또는 적어도 90% 동일한 서열을 포함하는, 가이드 리보핵산 폴리뉴클레오타이드.

청구항 78

제1항 내지 제77항 중 어느 한 항에 있어서, 상기 엔도뉴클레아제는 상기 엔도뉴클레아제에 커플링된 염기 편집기 또는 히스톤 편집기를 추가로 포함하는, 조작된 가이드 리보핵산 폴리뉴클레오타이드.

청구항 79

제78항에 있어서, 상기 염기 편집기는 아데노신 탈아미노효소인, 조작된 가이드 리보핵산 폴리뉴클레오타이드.

청구항 80

제79항에 있어서, 상기 아데노신 탈아미노효소는 ADAR1 또는 ADAR2를 포함하는, 조작된 가이드 리보핵산 폴리뉴클레오타이드.

청구항 81

제78항에 있어서, 상기 염기 편집기는 시토신 탈아미노효소인, 조작된 가이드 리보핵산 폴리뉴클레오타이드.

청구항 82

제81항에 있어서, 상기 시토신 탈아미노효소는 APOBEC1, APOBEC2, APOBEC3A, APOBEC3B, APOBEC3C, APOBEC3D, APOBEC3F, APOBEC3G, APOBEC3H, 또는 APOBEC4를 포함하는, 조작된 가이드 리보핵산 폴리뉴클레오타이드.

청구항 83

제74항 내지 제82항 중 어느 한 항의 조작된 가이드 리보핵산 폴리뉴클레오타이드를 암호화하는, 테옥시리보핵산 폴리뉴클레오타이드.

청구항 84

유기체에서의 발현에 최적화된 조작된 핵산 서열을 포함하는 핵산으로서, 상기 핵산은 RuvC 도메인 및 HNH 도메인을 포함하는 엔도뉴클레아제를 암호화되, 상기 엔도뉴클레아제는 미배양 미생물로부터 유래되고, 상기 엔도

뉴클레아제는 약 120 kDa 이하, 100 kDa 이하, 90 kDa 이하, 60 kDa 이하, 또는 30 kDa 이하의 분자량을 가지며,

상기 엔도뉴클레아제는 서열번호 674-675, 975-1002, 1260-1321을 포함하거나 이에 대해 적어도 70%의 서열 동일성을 갖는 이의 변이체를 포함하는, 핵산.

청구항 85

제84항에 있어서, 상기 엔도뉴클레아제는 상기 엔도뉴클레아제의 N-말단 또는 C-말단에 근접한 하나 이상의 핵국소화 서열(NLS)을 암호화하는 서열을 추가로 포함하는, 핵산.

청구항 86

제85항에 있어서, 상기 NLS는 서열번호 205-220으로부터 선택되는 서열을 포함하는, 핵산.

청구항 87

제84항 내지 제86항 중 어느 한 항에 있어서, 상기 유기체는 원핵생물, 박테리아, 진핵생물, 진균류, 식물, 포유류, 설치류, 또는 인간인, 핵산.

청구항 88

제87항에 있어서, 상기 유기체는 원핵생물 또는 박테리아이고, 상기 유기체는 상기 엔도뉴클레아제가 유래되는 유기체와 상이한 유기체인, 핵산.

청구항 89

제87항에 있어서, 상기 유기체는 미배양 미생물이 아닌, 핵산.

청구항 90

RuvC-I 도메인 및 HNH 도메인을 포함하는 RNA 가이드 엔도뉴클레아제를 암호화하는 핵산을 포함하는 벡터로서, 상기 엔도뉴클레아제는 미배양 미생물로부터 유래되고, 상기 엔도뉴클레아제는 약 120 kDa 이하, 100 kDa 이하, 90 kDa 이하, 또는 60 kDa 이하의 분자량을 가지며, 상기 RNA 가이드 엔도뉴클레아제는 선택적으로 고세균이며, 여기에서 상기 RNA 가이드 엔도뉴클레아제는 서열번호 674-675, 975-1002, 1260-1321을 포함하거나 이에 대해 적어도 70%의 서열 동일성을 갖는 이의 변이체를 포함하는, 벡터.

청구항 91

제90항에 있어서, 상기 엔도뉴클레아제는 RRxRR 모티프 또는 PF14239 상동성을 갖는 도메인을 포함하는 아르기닌-풍부 영역을 추가로 포함하는, 벡터.

청구항 92

제90항 또는 제91항에 있어서, 상기 엔도뉴클레아제는 REC(인식) 도메인을 추가로 포함하는, 벡터.

청구항 93

제91항 또는 제92항에 있어서, 상기 엔도뉴클레아제는 BH 도메인, WED 도메인, 및 표적 인접 모티프(TAM)-상호작용(TI) 도메인을 추가로 포함하는, 벡터.

청구항 94

제93항에 있어서, 상기 TI 도메인은 서열 번호 1-198, 221-459, 463-612, 617-668, 674-675, 975-1002, 1260-1321 중 어느 하나의 TI 도메인을 포함하는, 벡터.

청구항 95

제84항 내지 제89항 중 어느 한 항의 핵산을 포함하는, 벡터.

청구항 96

제90항 내지 제95항 중 어느 한 항에 있어서, 상기 엔도뉴클레아제와 복합체를 형성하도록 구성된 조작된 가이드 리보핵산 구조체를 암호화하는 핵산을 추가로 포함하되, 상기 조작된 가이드 리보핵산 구조체는:

- a) 표적 테옥시리보핵산 서열과 혼성화되도록 구성된 가이드 리보핵산 서열; 및
- b) 상기 엔도뉴클레아제에 결합하도록 구성된 tracr 리보핵산 서열을 포함하는, 벡터.

청구항 97

제90항 내지 제96항 중 어느 한 항에 있어서, 상기 벡터는 플라스미드, 미니서클, CELiD, 아데노-연관 바이러스(AAV) 유래 비리온, 또는 렌티바이러스인, 벡터.

청구항 98

제90항 내지 제97항 중 어느 한 항의 벡터를 포함하는, 세포.

청구항 99

제98항에 있어서, 상기 세포는 박테리아, 고세균, 진균, 진핵생물, 포유류, 또는 식물 세포인, 세포.

청구항 100

제99항에 있어서, 상기 세포는 박테리아 세포인, 세포.

청구항 101

엔도뉴클레아제를 제조하는 방법으로서, 제98항 내지 제100항 중 어느 한 항의 상기 세포를 배양하는 단계를 포함하는, 방법.

청구항 102

이중-가닥 테옥시리보핵산 폴리뉴클레오티드를 결합, 절단, 마킹, 또는 변형시키는 방법으로서, 상기 방법은:

(a) 상기 이중-가닥 테옥시리보핵산 폴리뉴클레오티드를, 상기 엔도뉴클레아제 및 상기 이중-가닥 테옥시리보핵산 폴리뉴클레오티드에 결합하도록 구성된 조작된 가이드 리보핵산 구조로 복합체 중 엔도뉴클레아제와 접촉시키는 단계를 포함하되;

(b) 여기에서 상기 이중-가닥 테옥시리보핵산 폴리뉴클레오티드는 표적 인접 모티프(TAM)를 포함하고;

상기 엔도뉴클레아제는 약 120 kDa 이하, 100 kDa 이하, 90 kDa 이하, 또는 60 kDa 이하의 분자량을 갖는, 방법.

청구항 103

제102항에 있어서, 상기 엔도뉴클레아제는 상기 이중-가닥 테옥시리보핵산 폴리뉴클레오티드를 절단하며, 여기에서 상기 TAM은 서열번호 1023-1044 중 어느 하나를 포함하는, 방법.

청구항 104

제102항 또는 제103항에 있어서, 상기 엔도뉴클레아제는 상기 TAM으로부터 상기 이중-가닥 테옥시리보핵산 폴리뉴클레오티드 5-7개 뉴클레오티드, 5개 뉴클레오티드, 6개 뉴클레오티드, 또는 7개 뉴클레오티드를 절단하는, 방법.

청구항 105

제102항에 있어서, 상기 엔도뉴클레아제는 서열번호 1-198, 221-459, 463-612, 617-668, 674-675, 975-1002, 1260-1321 중 어느 하나와 적어도 70%, 적어도 75%, 적어도 80%, 또는 적어도 90%의 서열 동일성을 갖는 변이체를 포함하는, 방법.

청구항 106

이중-가닥 테옥시리보핵산 폴리뉴클레오티드를 결합, 절단, 마킹, 또는 변형시키는 방법으로서, 상기 방법은

(a) 상기 이중-가닥 데옥시리보핵산 폴리뉴클레오티드를, 상기 엔도뉴클레아제 및 상기 이중-가닥 데옥시리보핵산 폴리뉴클레오티드에 결합하도록 구성된 조작된 가이드 리보핵산 구조로 복합체 중 RNA 가이드 고세균 엔도뉴클레아제와 접촉시키는 단계를 포함하되;

(b) 여기에서 상기 이중-가닥 데옥시리보핵산 폴리뉴클레오티드는 프로토스페이서 인접 모티프(PAM)를 포함하고;

상기 엔도뉴클레아제는 서열번호 674-675, 975-1002, 1260-1321 중 어느 하나와 적어도 70%, 적어도 75%, 적어도 80%, 또는 적어도 90%의 서열 동일성을 갖는 변이체를 포함하는, 방법.

청구항 107

제106항에 있어서, 상기 엔도뉴클레아제는 상기 이중-가닥 데옥시리보핵산 폴리뉴클레오티드를 절단하며, 여기에서 상기 PAM은 NGG를 포함하는, 방법.

청구항 108

제106항 또는 제107항에 있어서, 상기 엔도뉴클레아제는 상기 PAM으로부터 상기 이중-가닥 데옥시리보핵산 폴리뉴클레오티드 6-9개 뉴클레오티드, 또는 7개 뉴클레오티드를 절단하는, 방법.

청구항 109

제102항 내지 제108항 중 어느 한 항에 있어서, 상기 엔도뉴클레아제는 Cas9 엔도뉴클레아제, Cas14 엔도뉴클레아제, Cas12a 엔도뉴클레아제, Cas12b 엔도뉴클레아제, Cas 12c 엔도뉴클레아제, Cas12d 엔도뉴클레아제, Cas12e 엔도뉴클레아제, Cas13a 엔도뉴클레아제, Cas13b 엔도뉴클레아제, Cas13c 엔도뉴클레아제, 또는 Cas13d 엔도뉴클레아제가 아닌, 조작된 뉴클레아제 시스템.

청구항 110

제102항 내지 제109항 중 어느 한 항에 있어서, 상기 엔도뉴클레아제는 미배양 미생물로부터 유래되는, 방법.

청구항 111

제102항 내지 제110항 중 어느 한 항에 있어서, 상기 이중-가닥 데옥시리보핵산 폴리뉴클레오티드는 원핵생물, 고세균, 박테리아, 진핵생물, 식물, 진균류, 포유류, 설치류, 또는 인간 이중-가닥 데옥시리보핵산 폴리뉴클레오티드인, 방법.

청구항 112

제111항에 있어서, 상기 이중-가닥 데옥시리보핵산 폴리뉴클레오티드는 상기 엔도뉴클레아제가 유래된 종 이외의 종으로부터의 원핵생물, 고세균, 또는 박테리아 이중-가닥 데옥시리보핵산 폴리뉴클레오티드인, 방법.

청구항 113

표적 핵산 유전자좌를 변형시키는 방법으로서, 상기 방법은 제-244--73항 중 어느 한 항의 상기 조작된 뉴클레아제 시스템을 상기 표적 핵산 유전자좌에 전달하는 단계를 포함하고, 여기에서, 상기 엔도뉴클레아제는 상기 조작된 가이드 리보핵산 구조와 복합체를 형성하도록 구성되며, 여기에서 상기 복합체는 상기 복합체가 상기 표적 핵산 유전자좌에 결합할 때 상기 복합체가 상기 표적 핵 유전자좌를 변형시키도록 구성되는, 방법.

청구항 114

제113항에 있어서, 상기 표적 핵산 유전자좌를 변형시키는 단계는 상기 표적 핵산 유전자좌를 결합, 니킹, 절단, 또는 마킹하는 단계를 포함하는, 방법.

청구항 115

제113항 또는 제114항에 있어서, 상기 표적 핵산 유전자좌는 데옥시리보핵산(DNA) 또는 리보핵산(RNA)을 포함하는, 방법.

청구항 116

제115항에 있어서, 상기 표적 핵산 유전자좌는 게놈 진핵생물 DNA, 고세균 DNA, 바이러스 DNA, 또는 박테리아 DNA를 포함하는, 방법.

청구항 117

제116항에 있어서, 상기 표적 핵산은 박테리아 DNA를 포함하되, 상기 박테리아 DNA는 상기 엔도뉴클레아제가 유래된 종과 상이한 박테리아 또는 고세균으로부터 유래되는, 방법.

청구항 118

제113항 내지 제117항 중 어느 한 항에 있어서, 상기 표적 핵산 유전자좌는 시험관 내에 있는, 방법.

청구항 119

제113항 내지 제117항 중 어느 한 항에 있어서, 상기 표적 핵산 유전자좌는 세포 내에 있는, 방법.

청구항 120

제119항에 있어서, 상기 엔도뉴클레아제 및 상기 조작된 가이드 핵산 구조체는 별도의 핵산 분자에 의해 암호화되는, 방법.

청구항 121

제119항 또는 제120항에 있어서, 상기 세포는 원핵 세포, 박테리아 세포, 고세균 세포, 진핵 세포, 진균 세포, 식물 세포, 동물 세포, 포유류 세포, 설치류 세포, 영장류 세포, 또는 인간 세포인, 방법.

청구항 122

제119항 또는 제120항에 있어서, 상기 세포는 상기 엔도뉴클레아제가 유래된 종과 상이한 종으로부터 유래되는, 방법.

청구항 123

제119항 내지 제122항 중 어느 한 항에 있어서, 상기 조작된 뉴클레아제 시스템을 상기 표적 핵산 유전자좌에 전달하는 단계는 제84항 내지 제89항 중 어느 한 항의 핵산 또는 제90항 내지 제97항 중 어느 한 항의 벡터를 전달하는 단계를 포함하는, 방법.

청구항 124

제113항 내지 제123항 중 어느 한 항에 있어서, 상기 조작된 뉴클레아제 시스템을 상기 표적 핵산 유전자좌에 전달하는 단계는 상기 엔도뉴클레아제를 암호화하는 개방 해독 프레임을 포함하는 핵산을 전달하는 단계를 포함하는, 방법.

청구항 125

제124항에 있어서, 상기 핵산은 상기 엔도뉴클레아제를 암호화하는 상기 개방 해독 프레임이 작동 가능하게 연결되는 프로모터를 포함하는, 방법.

청구항 126

제122항 내지 제125항 중 어느 한 항에 있어서, 상기 조작된 뉴클레아제 시스템을 상기 표적 핵산 유전자좌에 전달하는 단계는 상기 엔도뉴클레아제를 암호화하는 상기 개방 해독 프레임을 함유하는 캡핑된 mRNA를 전달하는 단계를 포함하는, 방법.

청구항 127

제113항 내지 제120항 중 어느 한 항에 있어서, 상기 조작된 뉴클레아제 시스템을 상기 표적 핵산 유전자좌에 전달하는 단계는 번역된 폴리펩티드를 전달하는 단계를 포함하는, 방법.

청구항 128

제113항 내지 제126항 중 어느 한 항에 있어서, 상기 조작된 뉴클레아제 시스템을 상기 표적 핵산 유전자좌에 전달하는 단계는 리보핵산(RNA) pol III 프로모터에 작동 가능하게 연결된 상기 조작된 가이드 리보핵산 구조체를 암호화하는 데옥시리보핵산(DNA)을 전달하는 단계를 포함하는, 방법.

청구항 129

제113항 내지 제128항 중 어느 한 항에 있어서, 상기 엔도뉴클레아제는 상기 표적 유전자좌에서, 또는 이에 근접하여 단일-가닥 파단 또는 이중-가닥 파단을 유도하는, 방법.

청구항 130

제129항에 있어서, 상기 엔도뉴클레아제는 프로토스페이서 인접 모티프(PAM)로부터 상기 표적 유전자좌 5'에 근접한 이중 가닥 파단을 유도하는, 방법.

청구항 131

제130항에 있어서, 상기 엔도뉴클레아제는 상기 PAM으로부터 5'에 있는 6-8개의 뉴클레오티드 또는 7개의 뉴클레오티드의 이중-가닥 파단을 유도하는, 방법.

청구항 132

제113항 내지 제128항 중 어느 한 항에 있어서, 상기 조작된 뉴클레아제 시스템은 상기 표적 유전자좌 내에서 또는 이에 근접하여 뉴클레오티드 염기의 화학적 변형을 유도하는, 방법.

청구항 133

제132항에 있어서, 상기 화학적 변형은 아데노신 또는 시토신 뉴클레오티드의 탈아민화인, 방법.

청구항 134

제113항 내지 제128항 중 어느 한 항에 있어서, 상기 엔도뉴클레아제는 상기 엔도뉴클레아제에 커플링된 염기 편집기를 추가로 포함하는, 조작된 뉴클레아제 시스템.

청구항 135

제134항에 있어서, 상기 염기 편집기는 아데노신 탈아미노효소인, 조작된 뉴클레아제 시스템.

청구항 136

제135항에 있어서, 상기 아데노신 탈아미노효소는 ADAR1 또는 ADAR2를 포함하는, 조작된 뉴클레아제 시스템.

청구항 137

제134항에 있어서, 상기 염기 편집기는 시토신 탈아미노효소인, 조작된 뉴클레아제 시스템.

청구항 138

제137항에 있어서, 상기 시토신 탈아미노효소는 APOBEC1, APOBEC2, APOBEC3A, APOBEC3B, APOBEC3C, APOBEC3D, APOBEC3F, APOBEC3G, APOBEC3H, 또는 APOBEC4를 포함하는, 조작된 뉴클레아제 시스템.

청구항 139

세포에서 TRAC 유전자좌를 파괴하는 방법으로서, 상기 방법은 상기 세포를 조성물에 접촉시키는 단계를 포함하되, 상기 조성물은:

(a) 서열번호 1-198, 221-459, 463-612, 617-668, 674-675, 975-1002, 1260-1321 중 어느 하나와 적어도 75%의 동일성을 갖는 엔도뉴클레아제, 또는 이의 변이체, 및

(b) 조작된 가이드 RNA이되, 상기 조작된 가이드 RNA는 상기 엔도뉴클레아제와 복합체를 형성하도록 구성되고, 상기 조작된 가이드 RNA는 상기 유전자좌의 영역에 혼성화되도록 구성된 스페이서 서열을 포함하는, 조작된 가이드 RNA를 포함하고,

여기에서, 상기 조작된 가이드 RNA는 서열번호 1079-1082, 1145-1166, 및 1169-1170 중 어느 하나에 혼성화되도록 구성되는, 방법.

청구항 140

제139항에 있어서, 상기 조작된 가이드 RNA는 서열번호 1123-1144 또는 1167-1168 중 어느 하나와 적어도 약 80%의 동일성을 갖는 서열을 포함하는, 방법.

청구항 141

제139항 또는 제140항에 있어서, 상기 조작된 가이드 RNA는 서열번호 1123-1144 또는 1167-1168 중 어느 하나의 변형된 뉴클레오티드를 포함하는, 방법.

청구항 142

제139항 내지 제141항 중 어느 한 항에 있어서, 상기 조작된 가이드 RNA는 서열번호 1145-1166 또는 1169-1170 중 어느 하나와 상보적인 서열과 적어도 약 80%의 동일성을 갖는 서열을 포함하는, 방법.

청구항 143

제139항 내지 제142항 중 어느 한 항에 있어서, 상기 엔도뉴클레아제는 서열번호 582, 988, 990, 993, 996, 999, 또는 1002와 적어도 약 75%, 80%, 또는 90%의 서열 동일성을 갖는, 방법.

청구항 144

제139항 내지 제143항 중 어느 한 항에 있어서, 상기 영역은 서열번호 1023-1044 중 어느 하나를 포함하는 프로토스페이스 인접 모티프(PAM)에 대해 5'인, 방법.

청구항 145

단리된 RNA 분자로서, 서열번호 1123-1144 또는 1167-1168 중 어느 하나와 적어도 약 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 서열을 포함하는, 단리된 RNA 분자.

청구항 146

제145항에 있어서, 서열번호 1123-1144 또는 1167-1168 중 어느 하나에 인용된 화학적 변형의 패턴을 추가로 포함하는, 단리된 RNA 분자.

청구항 147

세포의 TRAC 유전자좌를 변형시키기 위한, 제145항 또는 제146항의 단리된 RNA 분자의, 용도.

청구항 148

세포에서 AAVS1 유전자좌를 파괴하는 방법으로서, 상기 방법은 상기 세포를 조성물에 접촉시키는 단계를 포함하되, 상기 조성물은:

(a) 서열번호 1-198, 221-459, 463-612, 617-668, 674-675, 975-1002, 1260-1321 중 어느 하나와 적어도 75%의 동일성을 갖는 엔도뉴클레아제, 또는 이의 변이체, 및

(b) 조작된 가이드 RNA이되, 상기 조작된 가이드 RNA는 상기 엔도뉴클레아제와 복합체를 형성하도록 구성되고, 상기 조작된 가이드 RNA는 상기 유전자좌의 영역에 혼성화되도록 구성된 스페이스 서열을 포함하는, 조작된 가이드 RNA를 포함하고,

여기에서, 상기 조작된 가이드 RNA는 서열번호 1105-1122 중 어느 하나에 혼성화되도록 구성되는, 방법.

청구항 149

제148항에 있어서, 상기 조작된 가이드 RNA는 서열번호 1087-1104 중 어느 하나와 적어도 약 80%의 동일성을 갖는 서열을 포함하는, 방법.

청구항 150

제148항 또는 제149항에 있어서, 상기 조작된 가이드 RNA는 서열번호 1087-1104 중 어느 하나의 변형된 뉴클레오타이드를 포함하는, 방법.

청구항 151

제148항 내지 제141항 중 어느 한 항에 있어서, 상기 조작된 가이드 RNA는 서열번호 1105-1122 중 어느 하나와 상보적인 서열과 적어도 약 80%의 동일성을 갖는 서열을 포함하는, 방법.

청구항 152

제148항 내지 제151항 중 어느 한 항에 있어서, 상기 엔도뉴클레아제는 서열번호 582, 988, 990, 993, 996, 999, 또는 1002와 적어도 약 75%, 80%, 또는 90%의 서열 동일성을 갖는, 방법.

청구항 153

제148항 내지 제151항 중 어느 한 항에 있어서, 상기 엔도뉴클레아제는 서열번호 582와 적어도 약 75%, 80%, 또는 90%의 서열 동일성을 갖는, 방법.

청구항 154

제148항 내지 제152항 중 어느 한 항에 있어서, 상기 영역은 서열번호 1023-1044 중 어느 하나를 포함하는 프로토스페이스 인접 모티프(PAM)에 대해 5'인, 방법.

청구항 155

단리된 RNA 분자로서, 서열번호 1087-1104 중 어느 하나와 적어도 약 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 서열을 포함하는, 단리된 RNA 분자.

청구항 156

제155항에 있어서, 서열번호 1087-1104 중 어느 하나에 인용된 화학적 변형의 패턴을 추가로 포함하는, 단리된 RNA 분자.

청구항 157

세포의 TRAC 유전자좌를 변형시키기 위한, 제155항 또는 제156항의 단리된 RNA 분자의, 용도.

청구항 158

조작된 뉴클레아제 시스템으로서:

(a) RuvC 도메인 및 HNH 도메인을 포함하는 엔도뉴클레아제이되, 상기 엔도뉴클레아제는 서열번호 25-27, 30, 126, 582, 594, 118, 128, 396, 530, 618, 620, 621, 653, 656, 657, 656 중 어느 하나와 적어도 80%의 서열 동일성을 갖는 서열, 또는 이의 변이체를 포함하는, 엔도뉴클레아제, 및

(b) 상기 엔도뉴클레아제와 복합체를 형성하도록 구성된 조작된 가이드 리보핵산 구조체로서:

(i) 표적 데옥시리보핵산 서열과 혼성화되도록 구성된 가이드 리보핵산 서열; 및

(ii) 상기 엔도뉴클레아제에 결합하도록 구성된 리보핵산 서열을 포함하는, 조작된 가이드 리보핵산 구조체를 포함하되;

상기 엔도뉴클레아제에 결합하도록 구성된 상기 리보핵산 서열은 서열번호 677-681, 686, 1006-1008, 1011-1014, 또는 1231-1259 중 어느 하나의 비측위 뉴클레오타이드와 적어도 80%의 서열 동일성을 갖는 서열을 포함하는, 조작된 뉴클레아제 시스템.

청구항 159

제158항에 있어서, 상기 조작된 가이드 리보핵산 구조체는 상기 가이드 리보핵산 서열 및 상기 엔도뉴클레아제

와 결합하도록 구성된 상기 리보핵산 서열을 포함하는 하나의 리보핵산 폴리뉴클레오티드를 포함하는, 조작된 뉴클레아제 시스템.

청구항 160

제158항 또는 제159항 중 어느 한 항에 있어서, 상기 가이드 리보핵산 서열은 원핵, 박테리아, 고세균, 진핵, 진균, 식물, 포유류, 또는 인간 게놈 서열에 상보적인, 조작된 뉴클레아제 시스템.

청구항 161

제158항 내지 제160항 중 어느 한 항에 있어서, 상기 엔도뉴클레아제가 상기 엔도뉴클레아제의 N-말단 또는 C-말단에 근접한 하나 이상의 핵 국소화 서열(NLS)을 포함하는, 조작된 뉴클레아제 시스템.

청구항 162

제161항에 있어서, 상기 NLS는 서열번호 205-220 중 어느 하나로부터 선택되는 서열을 포함하는, 조작된 뉴클레아제 시스템.

청구항 163

제158항 내지 제162항 중 어느 한 항에 있어서, 5'에서 3'으로, 상기 표적 데옥시리보핵산 서열의 5'에 있는 적어도 20개 뉴클레오티드의 서열을 포함하는 제1 상동 아암, 적어도 10개 뉴클레오티드의 합성 DNA 서열, 및 상기 표적 서열의 3'에 있는 적어도 20개 뉴클레오티드의 서열을 포함하는 제2 상동 아암을 포함하는 단일-가닥 또는 이중-가닥 DNA 복구 템플릿을 추가로 포함하는, 조작된 뉴클레아제 시스템.

청구항 164

제158항 내지 제163항 중 어느 한 항에 있어서, 상기 엔도뉴클레아제 및 상기 tracr 리보핵산 서열은 동일한 문내의 구별되는 박테리아 종으로부터 유래되는, 조작된 뉴클레아제 시스템.

청구항 165

제158항 내지 제164항 중 어느 한 항에 있어서, 상기 엔도뉴클레아제는 Cas9 엔도뉴클레아제, Cas14 엔도뉴클레아제, Cas12a 엔도뉴클레아제, Cas12b 엔도뉴클레아제, Cas 12c 엔도뉴클레아제, Cas12d 엔도뉴클레아제, Cas12e 엔도뉴클레아제, Cas13a 엔도뉴클레아제, Cas13b 엔도뉴클레아제, Cas13c 엔도뉴클레아제, 또는 Cas13d 엔도뉴클레아제가 아닌, 조작된 뉴클레아제 시스템.

청구항 166

제158항 내지 제165항 중 어느 한 항에 있어서, 상기 엔도뉴클레아제는 부수적인 ssDNA 절단 활성을 나타내지 않는, 조작된 뉴클레아제 시스템.

청구항 167

조작된 뉴클레아제 시스템으로서:

(a) RuvC 도메인 및 HNH 도메인을 포함하는 엔도뉴클레아제로서; 상기 엔도뉴클레아제는 본원에 기술된 엔도뉴클레아제 효과기 서열 중 어느 하나와 적어도 80%의 서열 동일성을 갖는 서열, 또는 이의 변이체, 및

(b) 상기 엔도뉴클레아제와 복합체를 형성하도록 구성된 조작된 가이드 리보핵산 구조체로서:

(i) 표적 데옥시리보핵산 서열과 혼성화되도록 구성된 가이드 리보핵산 서열; 및

(ii) 상기 엔도뉴클레아제에 결합하도록 구성된 리보핵산 서열을 포함하는, 조작된 가이드 리보핵산 구조체를 포함하되;

상기 엔도뉴클레아제는 본원에 기술된 sgRNA 서열 중 어느 하나의 비측위 뉴클레오티드와 적어도 80%의 서열 동일성을 갖는 서열, 또는 이의 변이체를 포함하는, 조작된 뉴클레아제 시스템.

청구항 168

본원에 기술된 sgRNA 서열 중 어느 하나의 비측위 뉴클레오티드와 적어도 약 80%, 81%, 82%, 83%, 84%, 85%,

86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 서열을 포함하는, 단리된 RNA 분자.

청구항 169

본원에 기술된 서열 중 어느 하나를 포함하는, 핵산.

청구항 170

본원에 기술된 핵산 서열 중 어느 하나를 포함하는, 벡터.

발명의 설명

기술 분야

[0001] 상호 참조

[0002] 본 출원은 2021년 11월 24일에 출원된 미국 특허 가출원 제63/282,999호; 2021년 12월 15일에 출원된 미국 특허 가출원 제63/289,981호; 및 2022년 6월 29일에 출원된 미국 특허 가출원 제63/356,908호의 이익을 주장하며, 이들 각각은 그 전체가 참조로서 본원에 통합된다.

[0003] 본 출원은, 그 전체가 참조로서 본원에 통합되는 PCT 출원 번호 PCT/US21/24945에 관한 것이다.

[0004] 서열 목록

[0005] 본 출원은 XML 포맷으로 전자적으로 제출된 서열 목록을 포함하며, 동 서열 목록은 그 전체가 참조로서 본원에 통합된다. 2022년 11월 23일에 생성된 상기 XML 사본의 명칭은 55921-741_601_SL.xml이며, 크기는 1,897,990 바이트이다.

배경 기술

[0006] Cas 효소는 이와 연관된 일정한 간격을 두고 주기적으로 분포하는 짧은 회문 반복서열(CRISPR) 가이드 리보핵산(RNA)과 함께 원핵 면역 체계의 만연한(약 45%의 박테리아, 약 84%의 고세균) 구성요소인 것으로 보이며, 이들은 CRISPR-RNA 가이드된 핵산 절단에 의해 비자기 핵산, 예컨대 감염성 바이러스 및 플라스미드에 대해 이러한 미생물을 보호하는 역할을 한다. CRISPR RNA 요소를 암호화하는 데옥시리보핵산(DNA) 요소는 구조 및 길이가 비교적 보존될 수 있지만, 이들의 CRISPR-연관(Cas) 단백질은 매우 다양하며, 매우 다양한 핵산-상호 작용 도메인을 함유한다. CRISPR DNA 요소는 1987년 초에 관찰되었지만, CRISPR/Cas 복합체의 프로그램 가능한 엔도뉴클레아제 절단 능력은 비교적 최근에 인식되었고, 이는 다양한 DNA 조작 및 유전자 편집 응용에서 재조합 CRISPR/Cas 시스템의 사용으로 이어지고 있다. 이들 효소의 유용성으로 인해, 이들은 광범위한 생명공학, 유전자 편집 및 치료적 응용 분야로 용도가 변경되고 있다. 이들의 단일 효과기 아키텍처로 인해, 현재 게놈 조작 분야로 용도가 변경되는 대부분의 시스템은 CRISPR 클래스 2 카테고리에 속한다.

발명의 내용

[0007] 많은 클래스 2 Cas 효과기의 큰 크기(약 1200개 아미노산 초과)는 치료 용도를 위한 전달을 어렵게 한다. 이에 따라, SMART(SMall ARchaeal-associATed)(소형 고세균 연관) 뉴클레아제 시스템으로 지칭되는 신규 추정 가이드 dsDNA 뉴클레아제에 관한 방법, 조성물, 및 시스템이 본원에 기술된다. 이들 엔도뉴클레아제 효과기는 이들의 작은 크기(약 400 aa 내지 약 1050 aa), RuvC 및 HNH 촉매 도메인의 존재, 및 신규 생화학적 메커니즘을 함께 제시하는 다른 예측된 단백질 특징으로 정의된다.

[0008] 일부 양태에서, 본 개시는 조작된 뉴클레아제 시스템을 제공하며, 이는: (a) RuvC 도메인 및 HNH 도메인을 포함하는 엔도뉴클레아제이되, 엔도뉴클레아제는 미배양 미생물로부터 유래되는, 엔도뉴클레아제; 및 (b) 엔도뉴클레아제와 복합체를 형성하도록 구성된 조작된 가이드 리보핵산 구조체로서: (i) 표적 데옥시리보핵산 서열과 혼성화되도록 구성된 가이드 리보핵산 서열; 및 (ii) 엔도뉴클레아제에 결합하도록 구성된 리보핵산 서열을 포함하는, 조작된 가이드 리보핵산 구조체를 포함하되; 엔도뉴클레아제는 약 96 kDa 이하, 약 80 kDa 이하, 약 70 kDa 이하, 또는 약 60 kDa 이하의 분자량을 가지고, 여기에서: (1) 엔도뉴클레아제는 서열번호 1-198, 221-459, 463-612, 617-668, 674-675, 975-1002, 1260-1321 중 어느 하나로부터의 아르기닌 풍부 영역 또는 PF14239 상동성을 갖는 도메인과 적어도 70%, 적어도 75%, 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%,

적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100%의 서열 동일성을 갖는, 아르기닌 풍부 영역 또는 PF14239 상동성을 갖는 도메인, 또는 이의 변이체를 포함하거나; (2) 엔도뉴클레아제는 서열번호 1-198, 221-459, 463-612, 617-668, 674-675, 975-1002, 1260-1321 중 어느 하나로부터의 REC 도메인과 적어도 70%, 적어도 75%, 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100%의 서열 동일성을 갖는, REC 도메인, 또는 이의 변이체를 포함하거나; (3) 상기 엔도뉴클레아제는 서열번호 1-198, 221-459, 463-612, 617-668, 674-675, 975-1002, 1260-1321 중 어느 하나와 적어도 70%, 적어도 75%, 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100%의 서열 동일성을 갖는, 서열, 또는 이의 변이체를 포함한다. 일부 구현예에서, (1) 엔도뉴클레아제는 서열번호 674-675, 975-1002, 또는 1260-1321 중 어느 하나로부터의 아르기닌 풍부 영역 또는 PF14239 상동성을 갖는 도메인과 적어도 70%, 적어도 75%, 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100%의 서열 동일성을 갖는 아르기닌 풍부 영역 또는 PF14239 상동성을 갖는 도메인, 또는 이의 변이체를 포함하거나; (2) 엔도뉴클레아제는 서열번호 674-675, 975-1002, 또는 1260-1321 중 어느 하나로부터의 REC 도메인과 적어도 70%, 적어도 75%, 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100%의 서열 동일성을 갖는 REC 도메인, 또는 이의 변이체를 포함하거나; (3) 엔도뉴클레아제는 서열번호 674-675, 975-1002, 또는 1260-1321 중 어느 하나와 적어도 70%, 적어도 75%, 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100%의 서열 동일성을 갖는 서열, 또는 이의 변이체를 포함한다. 일부 구현예에서, 엔도뉴클레아제는 고세균 엔도뉴클레아제이다. 일부 구현예에서, 엔도뉴클레아제는 서열번호 1-198, 221-459, 463-612, 617-668, 674-675, 975-1002, 1260-1321 중 어느 하나와 적어도 70%, 적어도 75%, 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100%의 서열 동일성을 갖는 서열을 포함한다. 일부 구현예에서, 엔도뉴클레아제는 RRxRR 모티프 또는 PF14239 상동성을 갖는 도메인을 포함하는 아르기닌-풍부 영역을 추가로 포함한다. 일부 구현예에서, 아르기닌 풍부 영역 또는 PF14239 상동성을 갖는 도메인은 서열번호 1-198, 221-459, 463-612, 617-668, 674-675, 975-1002, 1260-1321 중 어느 하나의 아르기닌 풍부 영역 또는 PF14239 상동성을 갖는 도메인과 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100%의 서열 동일성을 갖는다. 일부 구현예에서, 엔도뉴클레아제는 REC(인식) 도메인을 추가로 포함한다. 일부 구현예에서, REC 도메인은 서열번호 1-198, 221-459, 463-612, 617-668, 674-675, 975-1002, 1260-1321 중 어느 하나의 REC 도메인과 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100%의 서열 동일성을 갖는다. 일부 구현예에서, 엔도뉴클레아제는 BH(가교 나선) 도메인, WED(웨이) 도메인, 또는 PI(PAM 상호작용) 또는 TI(TAM 상호작용) 도메인을 추가로 포함한다. 일부 구현예에서, WED 도메인, 또는 PI 도메인은 서열번호 1-198, 221-459, 463-612, 617-668, 674-675, 975-1002, 1260-1321 중 어느 하나의 BH 도메인, WED 도메인, 또는 PI 도메인과 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100%의 서열 동일성을 갖는다.

[0009] 일부 양태에서, 본 개시는 조작된 뉴클레아제 시스템을 제공하며, 이는: (a) RuvC 도메인 및 HNH 도메인을 포함하는 엔도뉴클레아제; 및 (b) 엔도뉴클레아제와 복합체를 형성하도록 구성된 조작된 가이드 리보핵산 구조체로서: (i) 표적 데옥시리보핵산 서열과 혼성화되도록 구성된 가이드 리보핵산 서열; 및 (ii) 엔도뉴클레아제에 결

합하도록 구성된 리보핵산 서열을 포함하는, 조작된 가이드 리보핵산 구조체를 포함하되, 엔도뉴클레아제는 서열번호 674-675, 975-1002, 1260-1321 중 어느 하나와 적어도 70%, 적어도 75%, 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100%의 서열 동일성을 갖는 서열, 또는 이의 변이체를 포함한다. 일부 구현예에서, 엔도뉴클레아제는 고세균 엔도뉴클레아제이다. 일부 구현예에서, 엔도뉴클레아제는 RRxRR 모티프 또는 PF14239 상동성을 갖는 도메인을 포함하는 아르기닌-풍부 영역을 추가로 포함한다. 일부 구현예에서, 아르기닌 풍부 영역 또는 PF14239 상동성을 갖는 도메인은 서열번호 674-675, 975-1002, 1260-1321 중 어느 하나의 아르기닌 풍부 영역과 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100%의 서열 동일성을 갖는다. 일부 구현예에서, 엔도뉴클레아제는 REC(인식) 도메인을 추가로 포함한다. 일부 구현예에서, REC 도메인은 서열번호 674-675, 975-1002, 1260-1321 중 어느 하나의 REC 도메인과 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100%의 서열 동일성을 갖는다. 일부 구현예에서, 엔도뉴클레아제는 BH 도메인, WED 도메인, 및 PI 도메인을 추가로 포함한다. 일부 구현예에서, BH 도메인, WED 도메인, 또는 PI 도메인은 서열번호 674-675, 975-1002, 1260-1321 중 어느 하나의 BH 도메인, WED 도메인, 또는 PI 도메인과 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100%의 서열 동일성을 갖는다. 일부 구현예에서, 엔도뉴클레아제는 미배양 미생물로부터 유래된다. 일부 구현예에서, 엔도뉴클레아제에 결합하도록 구성된 리보핵산 서열은 서열번호 199-200, 460-461, 또는 669-673 중 어느 하나와 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100%의 서열 동일성을 갖는 서열, 또는 서열번호 201-203, 613-616, 677-686, 1003-1022, 또는 1231-1259 중 어느 하나의 비측퇴 뉴클레오티드와 적어도 80%의 서열 동일성을 갖는 서열을 포함한다. 일부 구현예에서, 가이드 핵산 구조체는 서열번호 201-203, 613-616, 677-686, 1003-1022, 또는 1231-1259 중 어느 하나의 비측퇴 뉴클레오티드와 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100%의 서열 동일성을 갖는 서열을 포함한다.

[0010]

일부 양태에서, 본 개시는 조작된 뉴클레아제 시스템을 제공하며, 이는: (a) 조작된 가이드 리보핵산 구조체로서: (i) 표적 데옥시리보핵산 서열과 혼성화되도록 구성된 가이드 리보핵산 서열; 및 (ii) 엔도뉴클레아제에 결합하도록 구성된 리보핵산 서열을 포함하되, 리보핵산 서열은 서열번호 199-200, 460-461, 또는 669-673 중 어느 하나와 적어도 80%의 서열 동일성을 갖는 서열, 또는 서열번호 677-686, 1006-1012, 또는 1231-1259 중 어느 하나의 비가변 뉴클레오티드와 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100%의 서열 동일성을 갖는 서열을 포함하는, 조작된 가이드 리보핵산 구조체; 및 (b) 조작된 가이드 리보핵산 서열에 결합하도록 구성된 RNA 가이드 엔도뉴클레아제를 포함한다. 일부 구현예에서, RNA 가이드 엔도뉴클레아제는 고세균 엔도뉴클레아제이다. 일부 구현예에서, 엔도뉴클레아제는 약 120 kDa 이하, 100 kDa 이하, 90 kDa 이하, 또는 60 kDa 이하의 분자량을 갖는다. 일부 구현예에서, 조작된 가이드 리보핵산 구조체는 적어도 2개의 리보핵산 폴리뉴클레오티드를 포함한다. 일부 구현예에서, 조작된 가이드 리보핵산 구조체는 가이드 리보핵산 서열 및 엔도뉴클레아제에 결합하도록 구성된 리보핵산 서열을 포함하는 단일 리보핵산 폴리뉴클레오티드를 포함한다. 일부 구현예에서, 가이드 리보핵산 서열은 원핵, 박테리아, 고세균, 진핵, 진균, 식물, 포유류, 또는 인간 게놈 서열에 상보적이다. 일부 구현예에서, 가이드 리보핵산 서열은 약 14 내지 약 28개 뉴클레오티드 길이, 약 18 내지 약 26개 뉴클레오티드 길이, 약 22 내지 약 26개 뉴클레오티드 길이, 또는 약 24개 뉴클레오티드 길이이다. 일부 구현예에서, 가이드 리보핵산 서열은 서열번호 462, 676, 또는 1229-1230 중 어느 하나와 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100%의 서열 동일성을 갖는 서열을 포함한다.

는 100%의 서열 동일성을 갖는 서열을 포함한다. 일부 구현예에서, 엔도뉴클레아제는 엔도뉴클레아제의 N-말단 또는 C-말단에 근접한 하나 이상의 핵 국소화 서열(NLS)을 포함한다. 일부 구현예에서, NLS는 서열번호 205-220 중 어느 하나로부터 선택되는 서열을 포함한다. 일부 구현예에서, 시스템은, 5'에서 3'까지 다음을 포함하는 단일 또는 이중 가닥 DNA 복구 템플릿을 추가로 포함한다: 표적 테옥시리보핵산 서열에 대해 5'에 있는 적어도 20개 뉴클레오타이드의 서열을 포함하는 제1 상동 아암, 적어도 10개 뉴클레오타이드의 합성 DNA 서열, 및 표적 서열에 대해 3'에 있는 적어도 20개 뉴클레오타이드의 서열을 포함하는 제2 상동 아암. 일부 구현예에서, 제1 또는 제2 상동 아암은 적어도 40, 80, 120, 150, 200, 300, 500, 또는 1,000개의 뉴클레오타이드의 서열을 포함한다. 일부 구현예에서, 시스템은 Mg^{2+} 의 공급원을 추가로 포함한다. 일부 구현예에서, 엔도뉴클레아제 및 엔도뉴클레아제에 결합하도록 구성된 리보핵산 서열은 동일한 문(phylum) 내의 구별되는 종으로부터 유래된다. 일부 구현예에서, 엔도뉴클레아제는 서열번호 2-24 중 어느 하나와 적어도 70%의 서열 동일성을 갖는 서열을 포함하고, 가이드 RNA 구조체는 줄기 및 루프를 포함하는 헤어핀을 포함하도록 예측된 RNA 서열을 포함하며, 여기에서 줄기는 적어도 10 쌍의 리보뉴클레오타이드 및 개재 다중루프를 포함한다. 일부 구현예에서, 가이드 RNA 구조체는 제2 줄기 및 제2 루프를 추가로 포함하되, 제2 줄기는 적어도 5 쌍의 리보뉴클레오타이드를 포함한다. 일부 구현예에서, 가이드 RNA 구조체는 적어도 2개의 헤어핀을 포함하는 RNA 구조체를 추가로 포함한다. 일부 구현예에서, a) 엔도뉴클레아제는 서열번호 1-198, 221-459, 463-612, 617-668, 674-675, 975-1002, 1260-1321 중 어느 하나와 적어도 70%, 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100%의 서열 동일성을 갖는 서열, 또는 이의 변이체를 포함하고; b) 가이드 RNA 구조체는 서열번호 199-200, 460-461, 또는 669-673 중 어느 하나와 적어도 70%, 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100%의 서열 동일성을 갖는 서열, 또는 서열번호 201-203, 613-616, 677-686, 1006-1012, 또는 1231-1259 중 어느 하나의 비가변 뉴클레오타이드를 포함한다. 일부 구현예에서, a) 엔도뉴클레아제는 서열번호 1-198, 221-459, 463-612, 617-668, 674-675, 975-1002, 1260-1321 중 어느 하나와 적어도 70%, 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100%의 서열 동일성을 갖는 서열을 포함하고; b) 가이드 RNA 구조체는 클래스 2, II형 sgRNA 또는 tracr 서열과 적어도 70%, 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100%의 동일성을 갖는 서열을 포함한다. 일부 구현예에서, 서열 동일성은 BLASTP, CLUSTALW, MUSCLE, MAFFT, 또는 Smith-Waterman 상동성 검색 알고리즘 파라미터를 사용하는 CLUSTALW에 의해 결정된다. 일부 구현예에서, 서열 동일성은, 단어 길이(W) 3, 기대치(E) 10의 파라미터, 및 BLOSUM62 스코어링 매트릭스(존재 11, 연장 1의 갭 비용 설정)를 사용하고, 조건부 조성 스코어 매트릭스 조정을 사용하는, BLASTP 상동성 검색 알고리즘에 의해 결정된다. 일부 구현예에서, 엔도뉴클레아제는 Cas9 엔도뉴클레아제, Cas14 엔도뉴클레아제, Cas12a 엔도뉴클레아제, Cas12b 엔도뉴클레아제, Cas 12c 엔도뉴클레아제, Cas12d 엔도뉴클레아제, Cas12e 엔도뉴클레아제, Cas13a 엔도뉴클레아제, Cas13b 엔도뉴클레아제, Cas13c 엔도뉴클레아제, 또는 Cas 13d 엔도뉴클레아제가 아니다. 일부 구현예에서, 엔도뉴클레아제는 Cas9 엔도뉴클레아제와 80% 미만의 동일성을 갖는다.

[0011]

일부 양태에서, 본 개시는 조작된 뉴클레아제 시스템을 제공하며, 이는: (a) ANG(서열번호 1029), NARAA(서열번호 1030), ATGAAA(서열번호 1031), ATGA(서열번호 1032), 또는 WTGG(서열번호 1033) 중 어느 하나를 포함하는 표적 인접 모티프(TAM) 서열에 대해 선택성을 갖도록 구성된 엔도뉴클레아제이되, 엔도뉴클레아제는 서열번호 1-198, 221-459, 463-612, 617-668, 674-675, 975-1002, 1260-1321 중 어느 하나의 TAM 상호작용 도메인과 적어도 75%, 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100%의 서열 동일성을 갖는 TAM 상호작용 도메인을 포함하는, 엔도뉴클레아제; 및 (b) 조작된 가이드 RNA이되, 조작된 가이드 RNA는 엔도뉴클레아제와 복합체를 형성하도록 구성되고, 조작된 가이드 RNA는 표적 핵산 서열에 혼성화되도록 구성된 스페이서 서열을 포함하는, 조작된 가이드 RNA를 포함한다. 일부 구현예에서, TAM 상호작용 도메인은 서열번호 674의 TAM 상호작용 도메인과 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도

97%, 적어도 98%, 적어도 99%, 또는 100%의 서열 동일성을 갖는 서열 또는 이의 변이체, 또는 서열번호 675의 TAM 상호작용 도메인과 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100%의 서열 동일성을 갖는 서열 또는, 이의 변이체를 포함한다. 일부 구현예에서, 엔도뉴클레아제 시스템은 진핵생물, 진균류, 식물, 포유류, 또는 인간 게놈 폴리뉴클레오타이드 서열에 상보적인 서열을 포함한다. 일부 구현예에서, 가이드 RNA는 30-280개 뉴클레오타이드 길이이다. 일부 구현예에서, 시스템은, 5'에서 3'까지 다음을 포함하는 단일 또는 이중 가닥 DNA 복구 템플릿을 추가로 포함한다: 표적 데옥시리보핵산 서열에 대해 5'에 있는 적어도 20개 뉴클레오타이드의 서열을 포함하는 제1 상동 아암, 적어도 10개 뉴클레오타이드의 합성 DNA 서열, 및 표적 서열에 대해 3'에 있는 적어도 20개 뉴클레오타이드의 서열을 포함하는 제2 상동 아암. 일부 구현예에서, 제1 또는 제2 상동 아암은 적어도 40개 뉴클레오타이드의 서열을 포함한다. 일부 구현예에서, 제1 및 제2 상동 아암은 진핵생물의 게놈 서열과 상동이다. 일부 구현예에서, 단일-가닥 또는 이중-가닥 DNA 복구 템플릿은 이식유전자 공여자를 포함한다. 일부 구현예에서, 시스템은 1개 또는 2개의 단일-가닥 DNA 분절이 측면에 위치한 이중-가닥 DNA 분절을 포함하는 DNA 복구 템플릿을 추가로 포함한다. 일부 구현예에서, 단일-가닥 DNA 분절은 이중-가닥 DNA 분절의 5' 말단에 접합된다. 일부 구현예에서, 단일-가닥 DNA 분절은 이중-가닥 DNA 분절의 3' 말단에 접합된다. 일부 구현예에서, 단일-가닥 DNA 분절은 4 내지 10개 뉴클레오타이드 염기의 길이를 갖는다. 일부 구현예에서, 단일-가닥 DNA 분절은 스페이서 서열 내의 서열에 상보적인 뉴클레오타이드 서열을 갖는다. 일부 구현예에서, 이중-가닥 DNA 서열은 바코드, 개방 해독 프레임, 인핸서, 프로모터, 단백질-코딩 서열, miRNA 코딩 서열, RNA 코딩 서열, 또는 이식유전자를 포함한다. 일부 구현예에서, 이중-가닥 DNA 서열은 뉴클레아제 절단 부위의 측면에 위치한다.

[0012] 일부 양태에서, 본 개시는 조작된 뉴클레아제 시스템을 제공하며, 이는: (a) NRR을 포함하는 프로토스페이서 인접 모티프(PAM) 서열에 대해 선택성을 갖도록 구성된 엔도뉴클레아제이되, 엔도뉴클레아제는 서열번호 1313-1318 중 어느 하나의 PAM 상호작용 도메인과 적어도 75%, 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100%의 서열 동일성을 갖는 PAM 상호작용 도메인을 포함하는, 엔도뉴클레아제; 및 (b) 조작된 가이드 RNA이되, 조작된 가이드 RNA는 엔도뉴클레아제와 복합체를 형성하도록 구성되고, 조작된 가이드 RNA는 표적 핵산 서열에 혼성화되도록 구성된 스페이서 서열을 포함하는, 조작된 가이드 RNA를 포함한다. 일부 구현예에서, TAM 상호작용 도메인은 서열번호 674의 TAM 상호작용 도메인과 적어도 80%의 서열 동일성을 갖는 서열 또는 이의 변이체, 또는 서열번호 675의 TAM 상호작용 도메인과 적어도 80%의 서열 동일성을 갖는 서열 또는 이의 변이체를 포함한다. 일부 구현예에서, 엔도뉴클레아제 시스템은 진핵생물, 진균류, 식물, 포유류, 또는 인간 게놈 폴리뉴클레오타이드 서열에 상보적인 서열을 포함한다. 일부 구현예에서, 가이드 RNA는 30-280개 뉴클레오타이드 길이이다. 일부 구현예에서, 시스템은, 5'에서 3'까지 다음을 포함하는 단일 또는 이중 가닥 DNA 복구 템플릿을 추가로 포함한다: 표적 데옥시리보핵산 서열에 대해 5'에 있는 적어도 20개 뉴클레오타이드의 서열을 포함하는 제1 상동 아암, 적어도 10개 뉴클레오타이드의 합성 DNA 서열, 및 표적 서열에 대해 3'에 있는 적어도 20개 뉴클레오타이드의 서열을 포함하는 제2 상동 아암. 일부 구현예에서, 제1 또는 제2 상동 아암은 적어도 40개 뉴클레오타이드의 서열을 포함한다. 일부 구현예에서, 제1 및 제2 상동 아암은 진핵생물의 게놈 서열과 상동이다. 일부 구현예에서, 단일-가닥 또는 이중-가닥 DNA 복구 템플릿은 이식유전자 공여자를 포함한다. 일부 구현예에서, 시스템은 1개 또는 2개의 단일-가닥 DNA 분절이 측면에 위치한 이중-가닥 DNA 분절을 포함하는 DNA 복구 템플릿을 추가로 포함한다. 일부 구현예에서, 단일-가닥 DNA 분절은 이중-가닥 DNA 분절의 5' 말단에 접합된다. 일부 구현예에서, 단일-가닥 DNA 분절은 이중-가닥 DNA 분절의 3' 말단에 접합된다. 일부 구현예에서, 단일-가닥 DNA 분절은 4 내지 10개 뉴클레오타이드 염기의 길이를 갖는다. 일부 구현예에서, 단일-가닥 DNA 분절은 스페이서 서열 내의 서열에 상보적인 뉴클레오타이드 서열을 갖는다. 일부 구현예에서, 이중-가닥 DNA 서열은 바코드, 개방 해독 프레임, 인핸서, 프로모터, 단백질-코딩 서열, miRNA 코딩 서열, RNA 코딩 서열, 또는 이식유전자를 포함한다. 일부 구현예에서, 이중-가닥 DNA 서열은 뉴클레아제 절단 부위의 측면에 위치한다.

[0013] 일부 양태에서, 본 개시는 조작된 단일 가이드 리보핵산 폴리뉴클레오타이드를 제공하며, 이는: a) 표적 DNA 분자의 표적 서열에 상보적인 뉴클레오타이드 서열을 포함하는 DNA-표적화 분절; 및 b) 이중-가닥 RNA(dsRNA) 이중체를 형성하도록 혼성화되는 뉴클레오타이드의 2개의 상보적 신장부를 포함하는 단백질-결합 분절로서, 뉴클레오타이드의 2개의 상보적 신장부는 개재 뉴클레오타이드와 서로 공유 연결되는, 단백질-결합 분절을 포함하고, 여기에서 조작된 가이드 리보핵산 폴리뉴클레오타이드는 서열번호 674-675, 975-1002, 1260-1321 중 어느 하나와 적어도 75%, 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적

어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100%의 서열 동일성을 갖는 변이체를 포함하는 엔도뉴클레아제, 또는 이의 변이체와 복합체를 형성하도록 구성된다. 일부 구현예에서, DNA-표적화 분절은 뉴클레오타이드의 2개의 상보적 신장부 둘 모두의 5'에 위치한다. 일부 구현예에서, a) 단백질 결합 분절은 서열번호 199-200, 460-461, 또는 669-673 중 어느 하나와 적어도 70%, 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100%의 서열 동일성을 갖는 서열을 포함하거나, b) 단백질 결합 분절은 서열번호 201-203, 613-616, 677-686, 1003-1022, 또는 1231-1259 중 어느 하나의 비가변 뉴클레오타이드와 적어도 70%, 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100%의 서열 동일성을 갖는 서열을 포함한다. 일부 구현예에서, a) 엔도뉴클레아제는 서열번호 674-675, 975-1002, 1260-1321 중 어느 하나와 적어도 70%, 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100%의 서열 동일성을 갖는 서열, 또는 이의 변이체를 포함하거나, b) 가이드 RNA 구조체는 클래스 2, II형 sgRNA와 적어도 70%, 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100%의 서열 동일성을 갖는 서열을 포함한다. 일부 구현예에서, 엔도뉴클레아제는 엔도뉴클레아제에 커플링된 염기 편집기(editor)에디터 또는 히스톤 편집기를 추가로 포함한다. 일부 구현예에서, 염기 편집기는 아데노신 탈아미노효소이다. 일부 구현예에서, 아데노신 탈아미노효소는 ADAR1 또는 ADAR2를 포함한다. 일부 구현예에서, 염기 편집기는 시토신 탈아미노효소이다. 일부 구현예에서, 시토신 탈아미노효소는 APOBEC1, APOBEC2, APOBEC3A, APOBEC3B, APOBEC3C, APOBEC3D, APOBEC3F, APOBEC3G, APOBEC3H, 또는 APOBEC4를 포함한다.

[0014] 일부 양태에서, 본 개시는 본원에 기술된 조작된 가이드 리보핵산 폴리뉴클레오타이드 중 어느 하나를 암호화하는 테옥시리보핵산 폴리뉴클레오타이드를 제공한다.

[0015] 일부 양태에서, 본 개시는 유기체에서의 발현에 대해 최적화된 조작된 핵산 서열을 포함하는 핵산을 제공하며, 여기에서, 핵산은 RuvC 도메인 및 HNH 도메인을 포함하는 엔도뉴클레아제를 암호화하고, 엔도뉴클레아제는 미배양 미생물로부터 유래되고, 엔도뉴클레아제는 약 120 kDa 이하, 100 kDa 이하, 90 kDa 이하, 60 kDa 이하, 또는 30 kDa 이하의 분자량을 가지며,

[0016] 엔도뉴클레아제는 서열번호 674-675, 975-1002, 1260-1321를 포함하거나, 이에 대해 적어도 70%, 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100%의 서열 동일성을 갖는 이의 변이체를 포함한다. 일부 구현예에서, 엔도뉴클레아제는 엔도뉴클레아제의 N-말단 또는 C-말단에 근접한 하나 이상의 핵 국소화 서열(NLS)을 암호화하는 서열을 추가로 포함한다. 일부 구현예에서, NLS는 서열번호 205-220으로부터 선택되는 서열을 포함한다. 일부 구현예에서, 유기체는 원핵생물, 박테리아, 진핵생물, 진균류, 식물, 포유류, 설치류 또는 인간이다. 일부 구현예에서, 유기체는 원핵생물 또는 박테리아이고, 유기체는 엔도뉴클레아제가 유래되는 유기체와 상이한 유기체이다. 일부 구현예에서, 유기체는 미배양 유기체가 아니다.

[0017] 일부 양태에서, 본 개시는 RuvC-I 도메인 및 HNH 도메인을 포함하는 RNA 가이드 엔도뉴클레아제를 암호화하는 핵산 서열을 포함하는 벡터를 제공하며, 엔도뉴클레아제는 미배양 미생물로부터 유래되고, 엔도뉴클레아제는 약 120 kDa 이하, 100 kDa 이하, 90 kDa 이하, 또는 60 kDa 이하의 분자량을 가지며, RNA 가이드 엔도뉴클레아제는 선택적으로 고세균이며, RNA 가이드 엔도뉴클레아제는 서열번호 674-675, 975-1002, 1260-1321를 포함하거나, 이에 대해 적어도 70%, 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100%의 서열 동일성을 갖는 이의 변이체를 포함한다. 일부 구현예에서, 엔도뉴클레아제는 RRxRR 모티프 또는 PF14239 상동성을 갖는 도메인을 포함하는 아르기닌-풍부 영역을 추가로 포함한다. 일부 구현예에서, 엔도뉴클레아제는 REC(인식) 도메인을 추가로 포함한다. 일부 구현예에서, 엔도뉴클레아제는 BH 도메인, WED 도메인, 및 표적 인접 모티프(TAM)-상호작용

(TI) 도메인을 추가로 포함한다. 일부 구현예에서, TI 도메인은 서열번호 1-198, 221-459, 463-612, 617-668, 674-675, 975-1002, 1260-1321 중 어느 하나의 TI 도메인을 포함한다.

[0018] 일부 양태에서, 본 개시는 본원에 기술된 핵산 중 어느 하나를 포함하는 벡터를 제공한다. 일부 구현예에서, 벡터는 엔도뉴클레아제와 복합체를 형성하도록 구성된 조작된 가이드 리보핵산 구조체를 암호화하는 핵산을 추가로 포함하며, 조작된 가이드 리보핵산 구조체는: a) 표적 데옥시리보핵산 서열과 혼성화되도록 구성된 가이드 리보핵산 서열; 및 b) 엔도뉴클레아제에 결합하도록 구성된 리보핵산 서열을 포함한다. 일부 구현예에서, 벡터는 플라스미드, 미니서클, CELiD, 아테노-연관 바이러스(AAV) 유래 비리온, 또는 렌티바이러스이다.

[0019] 일부 양태에서, 본 개시는 본원에 기술된 벡터 중 어느 하나를 포함하는 세포를 제공한다. 일부 구현예에서, 세포는 박테리아, 고세균, 진균, 진핵생물, 포유류, 또는 식물 세포이다. 일부 구현예에서, 세포는 박테리아 세포이다.

[0020] 일부 양태에서, 본 개시는 엔도뉴클레아제를 제조하는 방법을 제공하며, 방법은 본원에 기술된 세포 중 어느 하나를 배양하는 단계를 포함한다.

[0021] 일부 양태에서, 본 개시는 이중-가닥 데옥시리보핵산 폴리뉴클레오티드를 결합, 절단, 마킹, 또는 변형하는 방법을 제공하며, 방법은: (a) 이중-가닥 데옥시리보핵산 폴리뉴클레오티드를, 엔도뉴클레아제 및 이중-가닥 데옥시리보핵산 폴리뉴클레오티드에 결합하도록 구성된 조작된 가이드 리보핵산 구조체로 복합체 중 엔도뉴클레아제와 접촉시키는 단계를 포함하며; (b) 여기에서, 이중-가닥 데옥시리보핵산 폴리뉴클레오티드는 표적 인접 모티프(TAM)를 포함하고; 여기에서, 엔도뉴클레아제는 120 kDa 이하, 100 kDa 이하, 90 kDa 이하, 또는 60 kDa 이하의 분자량을 갖는다. 일부 구현예에서, 엔도뉴클레아제는 이중-가닥 데옥시리보핵산 폴리뉴클레오티드를 절단하며, 여기에서 TAM은 서열번호 1023-1044 중 어느 하나를 포함한다. 일부 구현예에서, 엔도뉴클레아제는 TAM으로부터 이중-가닥 데옥시리보핵산 폴리뉴클레오티드 5-7개 뉴클레오티드, 5개 뉴클레오티드, 6개 뉴클레오티드, 또는 7개 뉴클레오티드를 절단한다. 일부 구현예에서, 엔도뉴클레아제는 서열번호 1-198, 221-459, 463-612, 617-668, 674-675, 975-1002, 1260-1321 중 어느 하나와 적어도 70%, 적어도 75%, 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100%의 서열 동일성을 갖는 변이체를 포함한다.

[0022] 일부 양태에서, 본 개시는 이중-가닥 데옥시리보핵산 폴리뉴클레오티드를 결합, 절단, 마킹, 또는 변형하는 방법을 제공하며, 방법은: (a) 이중-가닥 데옥시리보핵산 폴리뉴클레오티드를, 엔도뉴클레아제 및 이중-가닥 데옥시리보핵산 폴리뉴클레오티드에 결합하도록 구성된 조작된 가이드 리보핵산 구조체로 복합체 중 RNA 가이드 고세균 엔도뉴클레아제와 접촉시키는 단계를 포함하며; (b) 여기에서, 이중-가닥 데옥시리보핵산 폴리뉴클레오티드는 프로토스페이서 인접 모티프(PAM)를 포함하고; 여기에서, 엔도뉴클레아제는 서열번호 674-675, 975-1002, 1260-1321 중 어느 하나와 적어도 70%, 적어도 75%, 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100%의 서열 동일성을 갖는 변이체를 포함한다. 일부 구현예에서, 엔도뉴클레아제는 이중-가닥 데옥시리보핵산 폴리뉴클레오티드를 절단하며, 여기에서 PAM은 NGG를 포함한다. 일부 구현예에서, 엔도뉴클레아제는 PAM으로부터 이중-가닥 데옥시리보핵산 폴리뉴클레오티드 6-9개 또는 7개의 뉴클레오티드를 절단한다. 일부 구현예에서, 엔도뉴클레아제는 Cas9 엔도뉴클레아제, Cas14 엔도뉴클레아제, Cas12a 엔도뉴클레아제, Cas12b 엔도뉴클레아제, Cas 12c 엔도뉴클레아제, Cas12d 엔도뉴클레아제, Cas12e 엔도뉴클레아제, Cas13a 엔도뉴클레아제, Cas13b 엔도뉴클레아제, Cas13c 엔도뉴클레아제, 또는 Cas 13d 엔도뉴클레아제가 아니다. 일부 구현예에서, 엔도뉴클레아제는 미배양 미생물로부터 유래된다. 일부 구현예에서, 이중-가닥 데옥시리보핵산 폴리뉴클레오티드는 원핵생물, 고세균, 박테리아, 진핵생물, 식물, 진균류, 포유류, 설치류, 또는 인간 이중-가닥 데옥시리보핵산 폴리뉴클레오티드이다. 일부 구현예에서, 이중-가닥 데옥시리보핵산 폴리뉴클레오티드는 엔도뉴클레아제가 유래된 종 이외의 종으로부터의 원핵생물, 고세균, 또는 박테리아 이중-가닥 데옥시리보핵산 폴리뉴클레오티드이다.

[0023] 일부 구현예에서, 본 개시는 표적 핵산 유전자좌를 변형시키는 방법을 제공하며, 방법은 본원에 기술된 조작된 뉴클레아제 시스템 중 어느 하나를 표적 핵산 유전자좌에 전달하는 단계를 포함하되, 엔도뉴클레아제는 조작된 가이드 리보핵산 구조와 복합체를 형성하도록 구성되고, 복합체는 해당 복합체가 표적 핵산 유전자좌에 결합할 때 복합체가 표적 핵산 유전자좌를 변형시키도록 구성된다. 일부 구현예에서, 표적 핵산 유전자좌는 표적 핵산 유전자좌의 결합, 니킹, 절단, 또는 마킹을 포함한다. 일부 구현예에서, 표적 핵산 유전자좌는 데옥시리보핵산

(DNA) 또는 리보핵산(RNA)을 포함한다. 일부 구현예에서, 표적 핵산은 게놈 진핵생물 DNA, 고세균 DNA, 바이러스 DNA, 또는 박테리아 DNA를 포함한다. 일부 구현예에서, 표적 핵산은 박테리아 DNA를 포함하되, 박테리아 DNA는 엔도뉴클레아제가 유래된 종과 상이한 박테리아 또는 고세균으로부터 유래된다. 일부 구현예에서, 표적 핵산 유전자좌는 시험관 내에 있다. 일부 구현예에서, 표적 핵산 유전자좌는 세포 내에 있다. 일부 구현예에서, 엔도뉴클레아제 및 조작된 가이드 핵산 구조체는 별도의 핵산 분자에 의해 암호화된다. 일부 구현예에서, 세포는 원핵생물 세포, 박테리아 세포, 고세균 세포, 진핵생물 세포, 진균류 세포, 식물 세포, 동물 세포, 포유류 세포, 설치류 세포, 영장류 세포, 또는 인간 세포이다. 일부 구현예에서, 세포는 엔도뉴클레아제가 유래된 종과 상이한 종으로부터 유래된다. 일부 구현예에서, 조작된 뉴클레아제 시스템을 표적 핵산 유전자좌에 전달하는 단계는 본원에 기술된 핵산 중 어느 하나 또는 본원에 기술된 벡터 중 어느 하나를 전달하는 단계를 포함한다. 일부 구현예에서, 조작된 뉴클레아제 시스템을 표적 핵산 유전자좌에 전달하는 단계는 엔도뉴클레아제를 암호화하는 개방 해독 프레임에 포함하는 핵산을 전달하는 단계를 포함한다. 일부 구현예에서, 핵산은 엔도뉴클레아제를 암호화하는 개방 해독 프레임이 작동 가능하게 연결되는 프로모터를 포함한다. 일부 구현예에서, 조작된 뉴클레아제 시스템을 표적 핵산 유전자좌에 전달하는 단계는 엔도뉴클레아제를 암호화하는 개방 해독 프레임을 함유하는 캡핑된 mRNA를 전달하는 단계를 포함한다. 일부 구현예에서, 조작된 뉴클레아제 시스템을 표적 핵산 유전자좌에 전달하는 단계는 번역된 폴리펩티드를 전달하는 단계를 포함한다. 일부 구현예에서, 조작된 뉴클레아제 시스템을 표적 핵산 유전자좌에 전달하는 단계는 리보핵산(RNA) pol III 프로모터에 작동 가능하게 연결된 조작된 가이드 리보핵산 구조를 암호화하는 데옥시리보핵산(DNA)을 전달하는 단계를 포함한다. 일부 구현예에서, 엔도뉴클레아제는 표적 유전자좌에서 또는 이에 근접하여 단일-가닥 파단 또는 이중-가닥 파단을 유도한다. 일부 구현예에서, 엔도뉴클레아제는 프로토스페이스 인접 모티프(PAM)로부터 표적 유전자좌 5'에 근접한 이중 가닥 파단을 유도한다. 일부 구현예에서, 엔도뉴클레아제는 PAM으로부터 5'에 있는 6-8개의 뉴클레오티드 또는 7개의 뉴클레오티드의 이중-가닥 파단을 유도한다. 일부 구현예에서, 조작된 뉴클레아제 시스템은 표적 유전자좌 내에서 또는 이에 근접한 뉴클레오티드 염기의 화학적 변형을 유도한다. 일부 구현예에서, 화학적 변형은 아데노신 또는 시토신 뉴클레오티드의 탈아미노화이다. 일부 구현예에서, 엔도뉴클레아제는 엔도뉴클레아제에 커플링된 염기 편집기를 추가로 포함한다. 일부 구현예에서, 염기 편집기는 아데노신 탈아미노효소이다. 일부 구현예에서, 아데노신 탈아미노효소는 ADAR1 또는 ADAR2를 포함한다. 일부 구현예에서, 염기 편집기는 시토신 탈아미노효소이다. 일부 구현예에서, 시토신 탈아미노효소는 APOBEC1, APOBEC2, APOBEC3A, APOBEC3B, APOBEC3C, APOBEC3D, APOBEC3F, APOBEC3G, APOBEC3H, 또는 APOBEC4를 포함한다.

[0024] 일부 양태에서, 본 개시는 세포에서 TRAC 유전자좌를 파괴하는 방법을 제공하며, 이는 조성물을 세포에 접촉시키는 단계를 포함하되, 조성물은: (a) 서열번호 1-198, 221-459, 463-612, 617-668, 674-675, 975-1002, 1260-1321 중 어느 하나와 적어도 75%, 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100%의 서열 동일성을 갖는 엔도뉴클레아제, 또는 이의 변이체, 및 (b) 조작된 가이드 RNA로서, 조작된 가이드 RNA는 엔도뉴클레아제와 복합체를 형성하도록 구성되고, 조작된 가이드 RNA는 유전자좌의 영역에 혼성화되도록 구성된 스페이스 서열을 포함하며, 여기에서 조작된 가이드 RNA는 서열번호 1079-1082, 1145-1166, 및 1169-1170 중 어느 하나에 혼성화되도록 구성되는, 조작된 가이드 RNA를 포함한다. 일부 구현예에서, 조작된 가이드 RNA는 서열번호 1123-1144 또는 1167-1168 중 어느 하나와 적어도 약 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100%의 서열 동일성을 갖는 서열을 포함한다. 일부 구현예에서, 조작된 가이드 RNA는 서열번호 1123-1144 또는 1167-1168 중 어느 하나의 변형된 뉴클레오티드를 포함한다. 일부 구현예에서, 조작된 가이드 RNA는 서열번호 1145-1166 또는 1169-1170 중 어느 하나와 상보적인 서열과 적어도 약 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100%의 서열 동일성을 갖는다. 일부 구현예에서, 영역은 서열번호 1023-1044 중 어느 하나를 포함하는 프로토스페이스 인접 모티프(PAM)에 대해 5'이다.

[0025] 일부 양태에서, 본 개시는 서열번호 1123-1144 또는 1167-1168 중 어느 하나와 적어도 약 80%, 81%, 82%, 83%,

84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 서열을 포함하는 단리된 RNA 분자를 제공한다. 일부 구현예에서, 단리된 RNA 분자는 서열번호 1123-1144 또는 1167-1168 중 어느 하나에 인용된 화학적 변형의 패턴을 포함한다.

[0026] 일부 양태에서, 본 개시는 세포의 TRAC 유전자좌를 변형시키기 위한 본원에 개시된 단리된 RNA 분자 중 어느 하나의 용도를 제공한다.

[0027] 일부 양태에서, 본 개시는 세포에서 AAVS1 유전자좌를 파괴하는 방법을 제공하며, 이는 조성물을 세포에 접촉시키는 단계를 포함하되, 조성물은: (a) 서열번호 1-198, 221-459, 463-612, 617-668, 674-675, 975-1002, 1260-1321 중 어느 하나와 적어도 75%, 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100%의 서열 동일성을 갖는 엔도뉴클레아제, 또는 이의 변이체, 및 (b) 조작된 가이드 RNA로서, 조작된 가이드 RNA는 엔도뉴클레아제와 복합체를 형성하도록 구성되고, 조작된 가이드 RNA는 유전자좌의 영역에 혼성화되도록 구성된 스페이서 서열을 포함하며, 여기에서 조작된 가이드 RNA는 서열번호 1105-1122 중 어느 하나에 혼성화되도록 구성되는, 조작된 가이드 RNA를 포함한다. 일부 구현예에서, 조작된 가이드 RNA는 서열번호 1087-1104 중 어느 하나와 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100%의 서열 동일성을 갖는 서열을 포함한다. 일부 구현예에서, 조작된 가이드 RNA는 서열번호 1087-1104 중 어느 하나의 변형된 뉴클레오타이드를 포함한다. 일부 구현예에서, 조작된 가이드 RNA는 서열번호 1105-1122 중 어느 하나와 상보적인 서열과 적어도 약 80% 동일성을 갖는 서열을 포함한다. 일부 구현예에서, 엔도뉴클레아제는 서열번호 582, 988, 990, 993, 996, 999, 또는 1002 중 어느 하나와 적어도 약 75%, 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100%의 서열 동일성을 갖는다. 일부 구현예에서, 엔도뉴클레아제는 서열번호 582와 적어도 약 75%, 80%, 또는 90%의 서열 동일성을 갖는다. 일부 구현예에서, 영역은, 서열번호 1023-1044 중 어느 하나를 포함하는 프로토스페이서 인접 모티프(PAM)에 대해 5'이다.

[0028] 일부 양태에서, 본 개시는 서열번호 1087-1104 중 어느 하나와 적어도 약 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 서열을 포함하는 단리된 RNA 분자를 제공한다. 일부 구현예에서, RNA 분자는 서열번호 1087-1104 중 어느 하나에 인용된 화학적 변형의 패턴을 포함한다.

[0029] 일부 양태에서, 본 개시는 조작된 뉴클레아제 시스템을 제공하며, 이는: (a) RuvC 도메인 및 HNH 도메인을 포함하는 엔도뉴클레아제로서, 엔도뉴클레아제는 서열번호 25-27, 30, 126, 582, 594, 118, 128, 396, 530, 618, 620, 621, 653, 656, 657, 656 중 어느 하나와 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100%의 서열 동일성을 갖는 서열, 또는 이의 변이체를 포함하는, 엔도뉴클레아제, 및 (b) 엔도뉴클레아제와 복합체를 형성하도록 구성된 조작된 가이드 리보핵산 구조체로서: (i) 표적 데옥시리보핵산 서열과 혼성화되도록 구성된 가이드 리보핵산 서열; 및 (ii) 엔도뉴클레아제에 결합하도록 구성된 리보핵산 서열을 포함하는, 조작된 가이드 리보핵산 구조체를 포함하되, 엔도뉴클레아제에 결합하도록 구성된 리보핵산 서열은 서열번호 677-681, 686, 1006-1008, 1011-1014, 또는 1231-1259 중 어느 하나의 비촉퇴 뉴클레오타이드와 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100%의 서열 동일성을 갖는 서열을 포함한다. 일부 구현예에서, 조작된 가이드 리보핵산 구조체는 가이드 리보핵산 서열 및 엔도뉴클레아제에 결합하도록 구성된 리보핵산 서열을 포함하는 단일 리보핵산 폴리뉴클레오타이드를 포함한다. 일부 구현예에서, 가이드 리보핵산 서열은 원핵, 박테리아, 고세균, 진핵, 진균, 식물, 포유류, 또는 인간 게놈 서열에 상보적이다. 일부 구현예에서, 엔도뉴클레아제는 엔도뉴클레아제의 N-말단 또는 C-말단에 근접한 하나 이상의 핵 국소화 서열(NLS)을 포함한다. 일부 구현예에서, NLS는 서열번호 205-220 중 어느 하나로부터 선택되는 서열을 포함한다. 일부 구현예에서, 시스템은, 5'에서 3'까지 다음을 포함하는 단일 또는 이중 가닥 DNA 복구 템플릿을 추가로 포함한다: 표적 데옥시리보핵산 서열에 대해 5'에 있는 적어도 20개 뉴클레오타이드의 서열을 포함하는 제1 상동 아암, 적어도 10개 뉴클레오타이드의 합성 DNA 서열, 및 표적 서열에 대해 3'에

있는 적어도 20개 뉴클레오타이드의 서열을 포함하는 제2 상동 아암. 일부 구현예에서, 엔도뉴클레아제 및 엔도뉴클레아제에 결합하도록 구성된 리보핵산 서열은 동일한 문(phylum) 내의 구별되는 종으로부터 유래된다. 일부 구현예에서, 엔도뉴클레아제는 Cas9 엔도뉴클레아제, Cas14 엔도뉴클레아제, Cas12a 엔도뉴클레아제, Cas12b 엔도뉴클레아제, Cas 12c 엔도뉴클레아제, Cas12d 엔도뉴클레아제, Cas12e 엔도뉴클레아제, Cas13a 엔도뉴클레아제, Cas13b 엔도뉴클레아제, Cas13c 엔도뉴클레아제, 또는 Cas 13d 엔도뉴클레아제가 아니다. 일부 구현예에서, 엔도뉴클레아제는 부수적인 ssDNA 절단 활성을 나타내지 않는다.

[0030] 일부 양태에서, 본 개시는 조작된 뉴클레아제 시스템을 제공하며, 이는: (a) RuvC 도메인 및 HNH 도메인을 포함하는 엔도뉴클레아제로서; 엔도뉴클레아제는 본원에 기술된 엔도뉴클레아제 효과기 서열 중 어느 하나와 적어도 80%의 서열 동일성을 갖는 서열, 또는 이의 변이체, 및 (b) 엔도뉴클레아제와 복합체를 형성하도록 구성된 조작된 가이드 리보핵산 구조체로서: (i) 표적 데옥시리보핵산 서열과 혼성화되도록 구성된 가이드 리보핵산 서열; 및 (ii) 엔도뉴클레아제에 결합하도록 구성된 리보핵산 서열을 포함하는, 조작된 가이드 리보핵산 구조체를 포함하되; 엔도뉴클레아제는 본원에 기술된 sgRNA 서열 중 어느 하나의 비축된 뉴클레오타이드와 적어도 80%의 서열 동일성을 갖는 서열, 또는 이의 변이체를 포함한다.

[0031] 일부 양태에서, 본 개시는 본원에 기술된 sgRNA 서열 중 어느 하나의 비축된 뉴클레오타이드와 적어도 약 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일한 서열을 포함하는 단리된 RNA 분자를 제공한다.

[0032] 일부 양태에서, 본 개시는 본원에 기술된 서열 중 어느 하나를 포함하는 핵산을 제공한다.

[0033] 일부 양태에서, 본 개시는 본원에 기술된 핵산 서열 중 어느 하나를 포함하는 벡터를 제공한다.

[0034] 일부 양태에서, 본 개시는 조작된 뉴클레아제 시스템을 제공하며, 이는: (a) RuvC 도메인 및 HNH 도메인을 포함하는 엔도뉴클레아제로서, 전술한 엔도뉴클레아제는 미생물로부터 유래되는, 엔도뉴클레아제; 및 (b) 전술한 엔도뉴클레아제와 복합체를 형성하도록 구성된 조작된 가이드 리보핵산 구조체로서 (i) 표적 데옥시리보핵산 서열과 혼성화되도록 구성된 가이드 리보핵산 서열; 및 (ii) 전술한 엔도뉴클레아제에 결합하도록 구성된 tracr 리보핵산 서열을 포함하는, 조작된 가이드 리보핵산 구조를 포함하며, 여기에서 전술한 엔도뉴클레아제는 약 96 kDa 이하의 분자량을 갖는다. 일부 구현예에서, 전술한 엔도뉴클레아제는 고세균 엔도뉴클레아제이다. 일부 구현예에서, 전술한 엔도뉴클레아제는 클래스 2, II형 Cas 엔도뉴클레아제이다. 일부 구현예에서, 전술한 엔도뉴클레아제는 서열번호 1-198, 221-459, 463-612, 617-668, 674-675, 975-1002, 1260-1321 중 어느 하나와 적어도 70%, 적어도 75%, 적어도 80%, 또는 적어도 90%의 서열 동일성을 갖는 서열, 또는 이의 변이체를 포함한다. 일부 구현예에서, 전술한 엔도뉴클레아제는 RRxRR 모티프 또는 PF14239 상동성을 갖는 도메인을 포함하는 아르기닌-풍부 영역을 추가로 포함한다. 일부 구현예에서, 전술한 아르기닌 풍부 영역 또는 전술한 PF14239 상동성을 갖는 도메인은 서열 번호 1-198, 221-459, 463-612, 617-668, 674-675, 975-1002, 1260-1321 중 어느 하나의 아르기닌 풍부 영역 또는 PF14239 상동성을 갖는 도메인, 또는 이의 변이체와 적어도 85%, 적어도 90%, 또는 적어도 95%의 동일성을 갖는다. 일부 구현예에서, 전술한 엔도뉴클레아제는 REC(인식) 도메인을 추가로 포함한다. 일부 구현예에서, 전술한 REC 도메인은 서열번호 1-198, 221-459, 463-612, 617-668, 674-675, 975-1002, 1260-1321 중 어느 하나의 REC 도메인, 또는 이의 변이체와 적어도 85%, 적어도 90%, 또는 적어도 95%의 동일성을 갖는다. 일부 구현예에서, 전술한 엔도뉴클레아제는 BH(가교 나선) 도메인, WED(웨이) 도메인, 및 PI(PAM 상호작용) 도메인을 추가로 포함한다. 일부 구현예에서, 전술한 BH 도메인, 전술한 WED 도메인, 또는 전술한 PI 도메인은 서열번호 1-198, 221-459, 463-612, 617-668, 674-675, 975-1002, 1260-1321 중 어느 하나의 BH 도메인, WED 도메인, 또는 PI 도메인, 또는 이의 변이체와 적어도 85%, 적어도 90%, 또는 적어도 95%의 동일성을 갖는다.

[0035] 일부 양태에서, 본 개시는 조작된 뉴클레아제 시스템을 제공하며, 이는: (a) RuvC_I 도메인 및 HNH 도메인을 포함하는 엔도뉴클레아제; 및 (b) 전술한 엔도뉴클레아제와 복합체를 형성하도록 구성된 조작된 가이드 리보핵산 구조체로서: (i) 표적 데옥시리보핵산 서열과 혼성화되도록 구성된 가이드 리보핵산 서열; 및 (ii) 전술한 엔도뉴클레아제에 결합하도록 구성된 리보핵산 서열을 포함하는, 조작된 가이드 리보핵산 구조체를 포함하되, 전술한 엔도뉴클레아제는 서열번호 1-198, 221-459, 463-612, 617-668, 674-675, 975-1002, 1260-1321 중 어느 하나와 적어도 70%, 적어도 75%, 적어도 80%, 또는 적어도 90%의 서열 동일성을 갖는 서열, 또는 이의 변이체를 포함한다. 일부 구현예에서, 전술한 엔도뉴클레아제는 고세균 엔도뉴클레아제이다. 일부 구현예에서, 전술한 엔도뉴클레아제는 클래스 2, II형 Cas 엔도뉴클레아제이다. 일부 구현예에서, 전술한 엔도뉴클레아제는 RRxRR 모티프 또는 PF14239 상동성을 갖는 도메인을 포함하는 아르기닌-풍부 영역을 추가로 포함한다. 일부 구현예에서,

전술한 아르기닌 풍부 영역 또는 전술한 PF14239 상동성을 갖는 도메인은 서열 번호 1-198, 221-459, 463-612, 617-668, 674-675, 975-1002, 1260-1321 중 어느 하나의 아르기닌 풍부 영역, 또는 이의 변이체와 적어도 85%, 적어도 90%, 또는 적어도 95%의 동일성을 갖는다. 일부 구현예에서, 전술한 엔도뉴클레아제는 REC(인식) 도메인을 추가로 포함한다. 일부 구현예에서, 전술한 REC 도메인은 서열번호 1-198, 221-459, 463-612, 617-668, 674-675, 975-1002, 1260-1321 중 어느 하나의 REC 도메인, 또는 이의 변이체와 적어도 85%, 적어도 90%, 또는 적어도 95%의 동일성을 갖는다. 일부 구현예에서, 전술한 엔도뉴클레아제는 BH 도메인, WED 도메인, 및 PI 도메인을 추가로 포함한다. 일부 구현예에서, 전술한 BH 도메인, 전술한 WED 도메인, 또는 전술한 PI 도메인은 서열 번호 1-198, 221-459, 463-612, 617-668, 674-675, 975-1002, 1260-1321 중 어느 하나의 BH 도메인, WED 도메인, 또는 PI 도메인, 또는 이의 변이체와 적어도 85%, 적어도 90%, 또는 적어도 95%의 동일성을 갖는다. 일부 구현예에서, 전술한 엔도뉴클레아제는 미배양 미생물로부터 유래된다. 일부 구현예에서, 전술한 엔도뉴클레아제에 결합하도록 구성된 전술한 리보핵산 서열은 서열번호 199-200, 460-461, 또는 669-673 중 어느 하나와 적어도 80%의 서열 동일성을 갖는 서열, 또는 서열번호 201-203, 613-616, 677-686, 1003-1022, 또는 1231-1259 중 어느 하나의 비축퇴 뉴클레오티드와 적어도 80%의 서열 동일성을 갖는 서열을 포함한다. 일부 구현예에서, 전술한 가이드 핵산 구조체는 서열번호 201-203, 613-616, 677-686, 1003-1022, 또는 1231-1259 중 어느 하나의 비축퇴 뉴클레오티드와 적어도 80%의 동일성을 갖는 서열을 포함한다.

[0036]

일부 양태에서, 본 개시는 조작된 뉴클레아제 시스템을 제공하며, 이는: (a) 조작된 가이드 리보핵산 구조체로서: (i) 표적 테옥시리보핵산 서열과 혼성화되도록 구성된 가이드 리보핵산 서열; 및 (ii) 엔도뉴클레아제에 결합하도록 구성된 리보핵산 서열을 포함하되, 전술한 리보핵산 서열은 서열번호 199-200, 460-461, 또는 669-673 중 어느 하나와 적어도 80%의 서열 동일성을 갖는 서열, 또는 서열번호 201-203, 613-616, 677-686, 1003-1022, 또는 1231-1259 중 어느 하나의 비가변 뉴클레오티드와 적어도 80%의 서열 동일성을 갖는 서열을 포함하는, 조작된 가이드 리보핵산 구조체; 및 (b) 전술한 조작된 가이드 리보핵산 서열에 결합하도록 구성된 RNA 가이드 엔도뉴클레아제를 포함한다. 일부 구현예에서, 전술한 RNA 가이드 엔도뉴클레아제는 고세균 엔도뉴클레아제이다. 일부 구현예에서, 전술한 엔도뉴클레아제는 약 120 kDa 이하, 100 kDa 이하, 90 kDa 이하, 또는 60 kDa 이하의 분자량을 갖는다. 일부 구현예에서, 전술한 조작된 가이드 리보핵산 구조는 적어도 2개의 리보핵산 폴리뉴클레오티드를 포함한다. 일부 구현예에서, 전술한 조작된 가이드 리보핵산 구조는 전술한 가이드 리보핵산 서열 및 전술한 tracr 리보핵산 서열을 포함하는 단일 리보핵산 폴리뉴클레오티드를 포함한다. 일부 구현예에서, 전술한 가이드 리보핵산 서열은 원핵, 박테리아, 고세균, 진핵, 진균, 식물, 포유류, 또는 인간 게놈 서열에 상보적이다. 일부 구현예에서, 전술한 가이드 리보핵산 서열은 15-24개 뉴클레오티드 길이이다. 일부 구현예에서, 전술한 엔도뉴클레아제는 전술한 엔도뉴클레아제의 N-말단 또는 C-말단에 근접한 하나 이상의 핵 국소화 서열(NLS)을 포함한다. 일부 구현예에서, 전술한 NLS는 서열번호 205-220로부터 선택되는 서열을 포함한다. 일부 구현예에서, 시스템은, 5'에서 3'으로 다음을 포함하는 단일 또는 이중-가닥 DNA 복구 템플릿을 추가로 포함한다: 전술한 표적 테옥시리보핵산 서열에 대해 5'에 있는 적어도 20개 뉴클레오티드의 서열을 포함하는 제1 상동 아암, 적어도 10개 뉴클레오티드의 합성 DNA 서열, 및 전술한 표적 서열에 대해 3'에 있는 적어도 20개 뉴클레오티드의 서열을 포함하는 제2 상동 아암. 일부 구현예에서, 전술한 제1 또는 제2 상동 아암은 적어도 40, 80, 120, 150, 200, 300, 500, 또는 1,000개의 뉴클레오티드의 서열을 포함한다. 일부 구현예에서, 전술한 시스템은 Mg^{2+} 의 공급원을 추가로 포함한다. 일부 구현예에서, 전술한 엔도뉴클레아제 및 전술한 tracr 리보핵산 서열은 동일한 문 내의 구별되는 박테리아 종으로부터 유래된다. 일부 구현예에서, 전술한 엔도뉴클레아제는 서열번호 2-24 중 어느 하나와 적어도 70%의 서열 동일성을 갖는 서열을 포함하고, 전술한 가이드 RNA 구조체는 줄기 및 루프를 포함하는 헤어핀을 포함하도록 예측된 RNA 서열을 포함하며, 여기에서 전술한 줄기는 적어도 12 쌍의 리보뉴클레오티드를 포함한다. 일부 구현예에서, 전술한 가이드 RNA 구조체는 제2 줄기 및 제2 루프를 추가로 포함하되, 제2 줄기는 적어도 5 쌍의 리보뉴클레오티드를 포함한다. 일부 구현예에서, 전술한 가이드 RNA 구조체는 적어도 2개의 헤어핀을 포함하는 RNA 구조체를 추가로 포함한다. 일부 구현예에서, 전술한 엔도뉴클레아제는 서열번호 1과 적어도 70%의 서열 동일성을 갖는 서열을 포함하고, 전술한 가이드 RNA 구조체는 줄기 및 루프를 포함하는 적어도 4개의 헤어핀을 포함하도록 예측된 RNA 서열을 포함한다. 일부 구현예에서, a) 전술한 엔도뉴클레아제는 서열번호 1, 2, 10, 17, 또는 613-616 중 어느 하나와 적어도 70%, 적어도 80%, 또는 적어도 90% 동일한 서열을 포함하고; (b) 전술한 가이드 RNA 구조체는 서열번호 199-200 또는 669-673 중 어느 하나 또는 서열번호 201-203, 613-616 중 어느 하나의 비가변 뉴클레오티드와 적어도 70%, 적어도 80%, 또는 적어도 90% 동일한 서열을 포함한다. 일부 구현예에서, a) 전술한 엔도뉴클레아제는 서열번호 1-24, 462-488, 또는 501-612 중 어느 하나와 적어도 70%, 적어도 80%, 또는 적어도 90% 동일한 서열을 포함하고; (b) 전술한 가이드 RNA 구조체는 서열번호 199-200 또는 669-673 중 어느 하나 또는 서열번호 201-203 또는 613-616 중 어느 하나의 비가

변 뉴클레오티드와 적어도 70%, 적어도 80%, 또는 적어도 90% 동일한 서열을 포함한다. 일부 구현예에서, a) 전술한 엔도뉴클레아제는 서열번호 2, 10, 또는 17 중 어느 하나와 적어도 70%, 적어도 80%, 또는 적어도 90% 동일한 서열을 포함하고; (b) 전술한 가이드 RNA 구조체는 서열번호 202-203 또는 613-614 중 어느 하나의 비가변 뉴클레오티드와 적어도 70%, 적어도 80%, 또는 적어도 90% 동일한 서열을 포함한다. 일부 구현예에서, a) 전술한 엔도뉴클레아제는 서열번호 25-198, 221-459, 또는 489-580 중 어느 하나와 적어도 70%, 적어도 80%, 또는 적어도 90% 동일한 서열을 포함하고; (b) 전술한 가이드 RNA 구조체는 클래스 2, II형 sgRNA 또는 tracr 서열과 적어도 70%, 적어도 80%, 또는 적어도 90% 동일한 서열을 포함한다. 일부 구현예에서, 전술한 서열 동일성은 BLASTP, CLUSTALW, MUSCLE, MAFFT, 또는 Smith-Waterman 상동성 검색 알고리즘 파라미터를 사용하는 CLUSTALW에 의해 결정된다. 일부 구현예에서, 전술한 서열 동일성은, 단어 길이(W) 3, 기대치(E) 10의 파라미터, 및 BLOSUM62 스코어링 매트릭스(존재 11, 연장 1의 갭 비용 설정)를 사용하고, 조건부 조성 스코어 매트릭스 조정을 사용하는, BLASTP 상동성 검색 알고리즘에 의해 결정된다. 일부 구현예에서, 전술한 엔도뉴클레아제는 Cas9 엔도뉴클레아제, Cas14 엔도뉴클레아제, Cas12a 엔도뉴클레아제, Cas12b 엔도뉴클레아제, Cas 12c 엔도뉴클레아제, Cas12d 엔도뉴클레아제, Cas12e 엔도뉴클레아제, Cas13a 엔도뉴클레아제, Cas13b 엔도뉴클레아제, Cas13c 엔도뉴클레아제, 또는 Cas13d 엔도뉴클레아제가 아니다. 일부 구현예에서, 전술한 엔도뉴클레아제는 Cas9 엔도뉴클레아제와 80% 미만의 동일성을 갖는다.

[0037] 일부 양태에서, 본 개시는 조작된 단일 가이드 리보핵산 폴리뉴클레오티드를 제공하며, 이는: a) 표적 DNA 분자의 표적 서열에 상보적인 뉴클레오티드 서열을 포함하는 DNA-표적화 분절; 및 b) 이중-가닥 RNA(dsRNA) 이중체를 형성하도록 혼성화되는 뉴클레오티드의 2개의 상보적 신장부를 포함하는 단백질-결합 분절로서, 상기 뉴클레오티드의 2개의 상보적 신장부는 개재 뉴클레오티드와 서로 공유 연결되는, 단백질-결합 분절을 포함하고, 여기에서 상기 조작된 가이드 리보핵산 폴리뉴클레오티드는 서열번호 1-198, 221-459, 463-612, 617-668, 674-675, 975-1002, 1260-1321 중 어느 하나와 적어도 75%의 서열 동일성을 갖는 변이체를 포함하는 엔도뉴클레아제, 또는 이의 변이체와 복합체를 형성하도록 구성된다. 일부 구현예에서, 전술한 DNA-표적화 분절은 뉴클레오티드의 전술한 2개의 상보적 신장부 둘 모두의 5'에 위치한다. 일부 구현예에서, a) 전술한 단백질 결합 분절은 서열번호 199-200 또는 669-673 중 어느 하나와 적어도 70%, 적어도 80%, 또는 적어도 90%의 동일성을 갖는 서열을 포함하고; (b) 전술한 단백질 결합 분절은 서열번호 201-203 또는 613-616 중 어느 하나의 비가변 뉴클레오티드와 적어도 70%, 적어도 80%, 또는 적어도 90%의 동일성을 갖는 서열을 포함한다. 일부 구현예에서, a) 전술한 엔도뉴클레아제는 서열번호 2, 10, 또는 17 중 어느 하나와 적어도 70%, 적어도 80%, 또는 적어도 90% 동일한 서열을 포함하고; (b) 전술한 가이드 RNA 구조체는 서열번호 200, 또는 서열번호 202-203 또는 613-614 중 어느 하나의 비가변 뉴클레오티드 중 적어도 하나와 적어도 70%, 적어도 80%, 또는 적어도 90% 동일한 서열을 포함한다. 일부 구현예에서, a) 전술한 엔도뉴클레아제는 서열번호 25-198, 221-459, 또는 489-580 중 어느 하나와 적어도 70%, 적어도 80%, 또는 적어도 90% 동일한 서열을 포함하고; (b) 전술한 가이드 RNA 구조체는 클래스 2, II형 sgRNA와 적어도 70%, 적어도 80%, 또는 적어도 90% 동일한 서열을 포함한다. 일부 구현예에서, 전술한 엔도뉴클레아제는 전술한 엔도뉴클레아제에 커플링된 염기 편집기 또는 히스톤 편집기를 추가로 포함한다. 일부 구현예에서, 전술한 염기 편집기는 아데노신 탈아미노효소이다. 일부 구현예에서, 전술한 아데노신 탈아미노효소는 ADAR1 또는 ADAR2를 포함한다. 일부 구현예에서, 전술한 염기 편집기는 시토신 탈아미노효소이다. 일부 구현예에서, 전술한 시토신 탈아미노효소는 APOBEC1, APOBEC2, APOBEC3A, APOBEC3B, APOBEC3C, APOBEC3D, APOBEC3F, APOBEC3G, APOBEC3H, 또는 APOBEC4를 포함한다.

[0038] 일부 양태에서, 본 개시는 본원에 기술된 조작된 가이드 리보핵산 폴리뉴클레오티드 중 어느 하나를 암호화하는 데옥시리보핵산 폴리뉴클레오티드를 제공한다.

[0039] 일부 양태에서, 본 개시는 유기체에서의 발현에 최적화된 조작된 핵산 서열을 포함하는 핵산을 제공하되, 전술한 핵산은 RuvC 도메인 및 HNH 도메인을 포함하는 클래스 2, II형 Cas 엔도뉴클레아제를 암호화하고, 여기에서 전술한 엔도뉴클레아제는 미배양 미생물로부터 유래되고, 전술한 엔도뉴클레아제는 약 120 kDa 이하, 100 kDa 이하, 90 kDa 이하, 60 kDa 이하, 또는 30 kDa 이하의 분자량을 갖는다. 일부 구현예에서, 전술한 엔도뉴클레아제는 서열번호 1-198, 221-459, 463-612, 617-668, 674-675, 975-1002, 1260-1321, 또는 이에 대해 적어도 70%의 서열 동일성을 갖는 변이체를 포함한다. 일부 구현예에서, 전술한 엔도뉴클레아제는 전술한 엔도뉴클레아제의 N-말단 또는 C-말단에 근접한 하나 이상의 핵 국소화 서열(NLS)을 암호화하는 서열을 포함한다. 일부 구현예에서, 전술한 NLS는 서열번호 205-220로부터 선택되는 서열을 포함한다. 일부 구현예에서, 전술한 유기체는 원핵생물, 박테리아, 진핵생물, 진균류, 식물, 포유류, 설치류, 또는 인간이다. 일부 구현예에서, 전술한 유기체는 원핵생물 또는 박테리아이고, 전술한 유기체는 전술한 엔도뉴클레아제가 유래되는 유기체와 상이한 유기체이

다. 일부 구현예에서, 전술한 유기체는 전술한 미배양 유기체가 아니다.

- [0040] 일부 양태에서, 본 개시는 RuvC-I 도메인 및 HNH 도메인을 포함하는 RNA 가이드 엔도뉴클레아제를 암호화하는 핵산을 포함하는 벡터를 제공하되, 전술한 엔도뉴클레아제는 미배양 미생물로부터 유래되고, 전술한 엔도뉴클레아제는 약 120 kDa 이하, 100 kDa 이하, 90 kDa 이하, 또는 60 kDa 이하의 분자량을 가지며, RNA 가이드 엔도뉴클레아제는 선택적으로 고세균이다. 일부 구현예에서, 전술한 엔도뉴클레아제는 RRxRR 모티프 또는 PF14239 상동성을 갖는 도메인을 포함하는 아르기닌-풍부 영역을 추가로 포함한다. 일부 구현예에서, 전술한 엔도뉴클레아제는 REC(인식) 도메인을 추가로 포함한다. 일부 구현예에서, 전술한 엔도뉴클레아제는 BH 도메인, WED 도메인, 및 PI 도메인을 추가로 포함한다.
- [0041] 일부 양태에서, 본 개시는 본원에 기술된 핵산 중 어느 하나를 포함하는 벡터를 제공한다. 일부 구현예에서, 벡터는 전술한 엔도뉴클레아제와 복합체를 형성하도록 구성된 조작된 가이드 리보핵산 구조체를 암호화하는 핵산을 추가로 포함하며, 전술한 조작된 가이드 리보핵산 구조체는: a) 표적 데옥시리보핵산 서열과 혼성화되도록 구성된 가이드 리보핵산 서열; 및 b) 전술한 엔도뉴클레아제에 결합하도록 구성된 tracr 리보핵산 서열을 포함한다. 일부 구현예에서, 벡터는 플라스미드, 미니서클, CELiD, 아테노-연관 바이러스(AAV) 유래 비리온, 또는 렌티바이러스이다.
- [0042] 일부 양태에서, 본 개시는 본원에 기술된 벡터 중 어느 하나를 포함하는 세포를 제공한다. 일부 구현예에서, 전술한 세포는 박테리아, 고세균, 진균, 진핵생물, 포유류, 또는 식물 세포이다. 일부 구현예에서, 전술한 세포는 박테리아 세포이다.
- [0043] 일부 양태에서, 본 개시는 엔도뉴클레아제를 제조하는 방법을 제공하며, 방법은 본원에 기술된 세포 중 어느 하나를 배양하는 단계를 포함한다.
- [0044] 일부 양태에서, 본 개시는 이중-가닥 데옥시리보핵산 폴리뉴클레오티드를 결합, 절단, 마킹, 또는 변형하는 방법을 제공하며, 방법은: (a) 전술한 이중-가닥 데옥시리보핵산 폴리뉴클레오티드를, 전술한 엔도뉴클레아제 및 전술한 이중-가닥 데옥시리보핵산 폴리뉴클레오티드에 결합하도록 구성된 조작된 가이드 리보핵산 구조체로 복합체 중 클래스 2, II형 Cas 엔도뉴클레아제와 접촉시키는 단계를 포함하며; (b) 여기에서, 전술한 이중-가닥 데옥시리보핵산 폴리뉴클레오티드는 프로토스페이서 인접 모티프(PAM)를 포함하고; 여기에서, 전술한 엔도뉴클레아제는 120 kDa 이하, 100 kDa 이하, 90 kDa 이하, 또는 60 kDa 이하의 분자량을 갖는다. 일부 구현예에서, 전술한 엔도뉴클레아제는 전술한 이중-가닥 데옥시리보핵산 폴리뉴클레오티드를 절단하며, 여기에서 전술한 PAM은 NGG를 포함한다. 일부 구현예에서, 전술한 엔도뉴클레아제는 전술한 PAM으로부터 전술한 이중-가닥 데옥시리보핵산 폴리뉴클레오티드 6-8개 뉴클레오티드 또는 7개 뉴클레오티드를 절단한다. 일부 구현예에서, 전술한 엔도뉴클레아제는 서열번호 1-198, 221-459, 463-612, 617-668, 674-675, 975-1002, 1260-1321 중 어느 하나와 적어도 70%, 적어도 75%, 적어도 80%, 또는 적어도 90%의 서열 동일성을 갖는 변이체, 또는 이의 변이체를 포함한다.
- [0045] 일부 양태에서, 본 개시는 이중-가닥 데옥시리보핵산 폴리뉴클레오티드를 결합, 절단, 마킹, 또는 변형하는 방법을 제공하며, 방법은: (a) 전술한 이중-가닥 데옥시리보핵산 폴리뉴클레오티드를, 전술한 엔도뉴클레아제 및 전술한 이중-가닥 데옥시리보핵산 폴리뉴클레오티드에 결합하도록 구성된 조작된 가이드 리보핵산 구조체로 복합체 중 RNA 가이드 고세균 엔도뉴클레아제와 접촉시키는 단계를 포함하며; 여기에서, 전술한 이중-가닥 데옥시리보핵산 폴리뉴클레오티드는 프로토스페이서 인접 모티프(PAM)를 포함하고; 여기에서, 전술한 엔도뉴클레아제는 서열번호 1-198, 221-459, 463-612, 617-668, 674-675, 975-1002, 1260-1321 중 어느 하나와 적어도 70%, 적어도 75%, 적어도 80%, 또는 적어도 90%의 서열 동일성을 갖는 변이체, 또는 이의 변이체를 포함한다. 일부 구현예에서, 전술한 엔도뉴클레아제는 전술한 이중-가닥 데옥시리보핵산 폴리뉴클레오티드를 절단하며, 여기에서 전술한 PAM은 NGG를 포함한다. 일부 구현예에서, 전술한 엔도뉴클레아제는 전술한 PAM으로부터 전술한 이중-가닥 데옥시리보핵산 폴리뉴클레오티드 6-8개 또는 7개의 뉴클레오티드를 절단한다. 일부 구현예에서, 전술한 클래스 2, II형 Cas 엔도뉴클레아제는 Cas9 엔도뉴클레아제, Cas14 엔도뉴클레아제, Cas12a 엔도뉴클레아제, Cas12b 엔도뉴클레아제, Cas 12c 엔도뉴클레아제, Cas12d 엔도뉴클레아제, Cas12e 엔도뉴클레아제, Cas13a 엔도뉴클레아제, Cas13b 엔도뉴클레아제, Cas13c 엔도뉴클레아제, 또는 Cas 13d 엔도뉴클레아제가 아니다. 일부 실시예에서, 상기 클래스 2, II형 Cas 엔도뉴클레아제는 미배양 미생물로부터 유래된다. 일부 구현예에서, 전술한 이중-가닥 데옥시리보핵산 폴리뉴클레오티드는 원핵생물, 고세균, 박테리아, 진핵생물, 식물, 진균류, 포유류, 설치류, 또는 인간 이중-가닥 데옥시리보핵산 폴리뉴클레오티드이다. 일부 구현예에서, 전술한 이중-가닥 데옥시리보핵산 폴리뉴클레오티드는 전술한 엔도뉴클레아제가 유래된 종 이외의 종으로부터의 원핵생물, 고세균, 또

는 박테리아 이중-가닥 데옥시리보핵산 폴리뉴클레오티드이다.

[0046]

일부 양태에서, 본 개시는 표적 핵산 유전자좌를 변형시키는 방법을 제공하며, 전술한 방법은 본원에 기술된 조작된 뉴클레아제 시스템 중 어느 하나를 전술한 표적 핵산 유전자좌에 전달하는 단계를 포함하며, 여기에서, 전술한 엔도뉴클레아제는 전술한 조작된 가이드 리보핵산 구조와 복합체를 형성하도록 구성되고, 전술한 복합체는 전술한 복합체가 전술한 표적 핵산 유전자좌에 결합할 때 전술한 복합체가 전술한 표적 핵산 유전자좌를 변형시키도록 구성된다. 일부 구현예에서, 전술한 표적 핵산 유전자좌를 변형시키는 단계는 전술한 표적 핵산 유전자좌에 대한 결합, 니킹, 절단, 또는 마킹하는 단계를 포함한다. 일부 구현예에서, 전술한 표적 핵산 유전자좌는 데옥시리보핵산(DNA) 또는 리보핵산(RNA)을 포함한다. 일부 구현예에서, 전술한 표적 핵산은 게놈 진핵생물 DNA, 고세균 DNA, 바이러스 DNA, 또는 박테리아 DNA를 포함한다. 일부 구현예에서, 전술한 표적 핵산은 박테리아 DNA를 포함하되, 전술한 박테리아 DNA는 전술한 엔도뉴클레아제가 유래된 종과 상이한 박테리아 또는 고세균으로부터 유래된다. 일부 구현예에서, 전술한 표적 핵산 유전자좌는 시험관 내에 있다. 일부 구현예에서, 전술한 표적 핵산 유전자좌는 세포 내에 있다. 일부 구현예에서, 전술한 엔도뉴클레아제 및 전술한 조작된 가이드 핵산 구조체는 별도의 핵산 분자에 의해 암호화된다. 일부 구현예에서, 전술한 세포는 원핵생물 세포, 박테리아 세포, 고세균 세포, 진핵생물 세포, 진균류 세포, 식물 세포, 동물 세포, 포유류 세포, 설치류 세포, 영장류 세포, 또는 인간 세포이다. 일부 구현예에서, 전술한 세포는 전술한 엔도뉴클레아제가 유래된 종과 상이한 종으로부터 유래된다. 일부 구현예에서, 전술한 조작된 뉴클레아제 시스템을 전술한 표적 핵산 유전자좌에 전달하는 단계는 본원에 기술된 핵산 중 어느 하나 또는 본원에 기술된 벡터 중 어느 하나를 전달하는 단계를 포함한다. 일부 구현예에서, 전술한 조작된 뉴클레아제 시스템을 전술한 표적 핵산 유전자좌에 전달하는 단계는 전술한 엔도뉴클레아제를 암호화하는 개방 해독 프레임워크를 포함하는 핵산을 전달하는 단계를 포함한다. 일부 구현예에서, 전술한 핵산은 전술한 엔도뉴클레아제를 암호화하는 전술한 개방 해독 프레임워크 작동 가능하게 연결되는 프로모터를 포함한다. 일부 구현예에서, 전술한 조작된 뉴클레아제 시스템을 전술한 표적 핵산 유전자좌에 전달하는 단계는 전술한 엔도뉴클레아제를 암호화하는 전술한 개방 해독 프레임워크 함유하는 캡핑된 mRNA를 전달하는 단계를 포함한다. 일부 구현예에서, 전술한 조작된 뉴클레아제 시스템을 전술한 표적 핵산 유전자좌에 전달하는 단계는 번역된 폴리펩티드를 전달하는 단계를 포함한다. 일부 구현예에서, 전술한 조작된 뉴클레아제 시스템을 전술한 표적 핵산 유전자좌에 전달하는 단계는 리보핵산(RNA) pol III 프로모터에 작동 가능하게 연결된 전술한 조작된 가이드 리보핵산 구조를 암호화하는 데옥시리보핵산(DNA)을 전달하는 단계를 포함한다. 일부 구현예에서, 전술한 엔도뉴클레아제는 전술한 표적 유전자좌에서 또는 이에 근접하여 단일-가닥 파단 또는 이중-가닥 파단을 유도한다. 일부 구현예에서, 전술한 엔도뉴클레아제는 프로토스페이서 인접 모티프(PAM)로부터 전술한 표적 유전자좌 5'에 근접한 이중 가닥 파단을 유도한다. 일부 구현예에서, 전술한 엔도뉴클레아제는 전술한 PAM으로부터 5'에 있는 6-8개의 뉴클레오티드 또는 7개의 뉴클레오티드의 이중-가닥 파단을 유도한다. 일부 구현예에서, 전술한 조작된 뉴클레아제 시스템은 전술한 표적 유전자좌 내 또는 이의 근위의 뉴클레오티드 염기의 화학적 변형 또는 전술한 표적 유전자좌 내 또는 근위의 히스톤의 화학적 변형을 유도한다. 일부 구현예에서, 전술한 화학적 변형은 아데노신 또는 시토신 뉴클레오티드의 탈아미노화이다. 일부 구현예에서, 전술한 엔도뉴클레아제는 전술한 엔도뉴클레아제에 커플링된 염기 편집기를 추가로 포함한다. 일부 구현예에서, 전술한 염기 편집기는 아데노신 탈아미노효소이다. 일부 구현예에서, 전술한 아데노신 탈아미노효소는 ADAR1 또는 ADAR2를 포함한다. 일부 구현예에서, 전술한 염기 편집기는 시토신 탈아미노효소이다. 일부 구현예에서, 전술한 시토신 탈아미노효소는 APOBEC1, APOBEC2, APOBEC3A, APOBEC3B, APOBEC3C, APOBEC3D, APOBEC3F, APOBEC3G, APOBEC3H, 또는 APOBEC4를 포함한다.

[0047]

본 개시의 추가 양태 및 이점은, 본 개시의 예시적인 구현예만이 도시되고 설명되는, 다음의 상세한 설명으로부터 당업자에게 쉽게 명백해질 것이다. 인지하게 되겠지만, 본 개시는 다른 구현예 및 상이한 구현예가 가능하고, 본 개시의 몇몇 세부 사항은 다양한 명백한 측면에서 본 개시를 벗어나지 않고도 변형될 수 있다. 따라서, 도면 및 본 발명을 실시하기 위한 구체적인 내용은 본질적으로 예시적인 것으로 간주되어야 하며, 제한적인 것으로 간주되지 않아야 한다.

[0048]

참조에 의한 통합

[0049]

본 명세서에 언급된 모든 간행물, 특허, 및 특허 출원은 마치 각각의 개별 간행물, 특허, 또는 특허 출원이 참조에 의해 구체적으로 및 개별적으로 통합된 것으로 표시된 것과 동일한 정도로 참조에 의해 본원에 통합된다.

도면의 간단한 설명

[0050]

본 발명의 신규한 특징은 특허 첨부된 청구범위에 명시되어 있다. 본 발명의 특징 및 장점은 본 발명의 원리가

활용되는 예시적인 구현예가 제시되는, 하기 발명을 실시하기 위한 구체적인 내용 및 첨부 도면(본원에서의 "도(Figure/FIG)")를 참조함으로써 보다 잘 이해될 것이다.

도 1a - 도 1b는 상이한 부류 및 유형의 CRISPR/Cas 유전자와의 상동성 관계를 나타내는 계통분석도를 도시한다. 클래스 2, II-A, II-B, 및 II-C형 Cas 시스템과 관련하여 본원에 기술된 SMART I 및 II Cas 효소 클래스가 도시되며, 이는 이들 시스템이 II-A, II-B, 및 II-C와는 별도의 클래스로 그룹화됨을 입증한다. (**도 1a**)는 Cas9 참조 서열의 맥락에서 SMART 계통유전학 트리를 도시하고(여기에서 SMART 효과기는 Cas9 참조 서열(II-A, II-B, 및 II-C형)로부터 거리를 두고 클러스터링됨); (**도 1b**)는 SMART 효소의 하위 그룹을 예시하는 SMART 계통유전학 트리를 도시한다.

도 2는 본원에 기술된 SMART 효과기에 대한 길이 분포를 도시하며, 이는 SMART I 및 II 효소가 Cas9-유사 효소보다 낮은 분자량으로 클러스터링됨을 나타낸다. SMART 뉴클레아제는 약 400 aa의 하나의 피크(SMART II) 및 약 750 aa의 제2 피크(SMART I)를 갖는 이중모드 분포를 나타낸다. Cas9 뉴클레아제는 또한 약 1,100 aa(예를 들어, SaCas9) 및 1,300 aa(예를 들어, SpCas9)의 피크를 갖는 이중모드 분포를 나타낸다.

도 3a - 도 3d는 '소형' II형 뉴클레아제 MG33-1, MG35-236의 게놈 컨텍스트를 도시한다. SMART 뉴클레아제 및 CRISPR 부속 단백질은 짙은 회색 화살표로 도시되어 있고, 다른 유전자는 밝은 회색 화살표로 도시되어 있다. 게놈 단편의 모든 유전자에 대해 예측된 도메인은 화살표 아래에 회색 박스로 도시되어 있다. 이는 다음을 나타낸다: (**도 3a**) SMART II 뉴클레아제 MG35-236으로부터 상류에 암호화된 SMART I MG33-1 뉴클레아제 및 CRISPR 유전자와의 게놈 컨텍스트로서, SMART II로부터 하류에 트랜스포사아제 TnpA 및 TnpB를 운반하는 예측된 삽입 서열을 나타낸다; (**도 3b**) SMART I 뉴클레아제 MG34-1의 게놈 컨텍스트로서, 환경 발현 시퀀싱 관독은 CRISPR 어레이 및 예측된 tracrRNA 아래에 정렬되어 제시되고, 해당 영역에 대한 전사체 커버리지는 콘티그 서열 위에 제시된다; (**도 3c**) SMART I 뉴클레아제 MG34-16의 게놈 컨텍스트로서, 환경 발현 시퀀싱 관독은 CRISPR 어레이 및 예측된 tracrRNA 아래에 정렬되어 제시되고, 해당 영역에 대한 전사체 커버리지는 콘티그 서열 위에 제시된다; 그리고 (**도 3d**) (**도 3d**)의 MG34-16 CRISPR 어레이로부터의 스페이스 7에 의해 표적화된 게놈 단편으로서, 해당 게놈 단편은 바이러스 특이적 유전자 주석 터미나아제 및 포털에 기반으로 하는 파지로부터 유래된 것으로 식별되었다; 삽입도는 알려지지 않은 기능의 바이러스 유전자의 C-말단을 표적화하는 MG34-16 스페이스 7의 위치를 나타낸다 - MG34-16에 대한 추정 NGG PAM은 스페이스 일치로부터 하류에 회색 박스로 강조된다.

도 4a - 도 4c는 예시적인 SMART 엔도뉴클레아제(MG33-1(서열번호 1), MG33-2(서열번호 463), MG33-3(서열번호 464), MG34-1(서열번호 2), MG 34-9(서열번호 10), MG34-16(서열번호 17), MG 102-1(서열번호 581), MG102-2(서열번호 582), MG35-1(서열번호 25), MG 35-2(서열번호 26), MG 35-3(서열번호 27), MG 35-102(서열번호 126), MG35-236(서열번호 284), MG35-419(서열번호 222), MG35-420(서열번호 223), 및 MG 35-421(서열번호 224))의 다중 서열 정렬을 나타내며, 여기에서 참조 도메인으로 사용된 SaCas9의 서열은 참조 서열 아래에 직사각형으로 표시되고, 촉매 잔기는 각각의 서열 위의 사각형으로 표시된다. 이는 다음을 나타낸다: (**도 4a**) RuvC-I 및 가교 나선 도메인을 함유하는 엔도뉴클레아제 영역의 정렬; (**도 4b**) RuvC-III 도메인을 함유하는 영역의 정렬; 및 (**도 4c**) RuvC-II 및 HNH 도메인을 함유하는 영역의 정렬.

도 5a - 도 5c는 MG34-1을 예로서 사용한, SMART I 엔도뉴클레아제에 대한 예시적인 도메인 조직을 도시한다. 이는 다음을 나타낸다: (**도 5a**)는 3개의 RuvC 도메인, 가교 나선("BH"), 인식 도메인("REC")을 방해하는 Pfam PF14239와 상동성인 도메인, HNH 엔도뉴클레아제 도메인("HNH"), 웨지 도메인("WED"), 및 PAM 상호작용 도메인(P1)을 포함하는 SMART I 뉴클레아제의 예측된 도메인 아키텍처를 나타내는 다이어그램이다; (**도 5b**) 참조 Cas9 뉴클레아제 서열에 대한 2개의 SMART I 뉴클레아제의 다중 서열 정렬 요약으로서, RuvC 및 HNH 촉매 잔기는 각각의 서열 위의 흑색 막대로 표시되고, SaCas의 결정 구조와 3D 공간에서 정렬되는 영역은 둥근 박스로 표시되고, 파선은 SMART 및 SaCas9의 3D 구조 예측 사이의 3D 공간에서 정렬이 불량하거나 정렬이 없는 영역을 나타낸다.

도 6a - 도 6b는 MG35 계열 효소(MG35-3, MG35-4)를 예로서 사용하여, SMART II 엔도뉴클레아제에 대한 예시적인 도메인 조직을 나타낸다. 이는 다음을 나타낸다: (**도 6a**)는 3개의 RuvC 도메인, Pfam PF14239와 상동성을 갖는 도메인, HNH 엔도뉴클레아제 도메인, 알려지지 않은 도메인, 및 인식 도메인(REC)을 포함하는 SMART II 뉴클레아제의 예측된 도메인 아키텍처를 나타내는 다이어그램이다; 그리고 (**도 6b**)는 참조 Cas9 뉴클레아제 서열에 대한 2개의 SMART II 뉴클레아제의 다중 서열 정렬 요약으로서, RuvC 및 HNH 촉매 잔기는 각각의 서열 위의 흑색 막대로 표시되고, SaCas의 결정 구조와 3D 공간에서 정렬되는 영역은 둥근 박스로 표시되고, 가이드/표적/PAM 서열의 인식에 관여할 수 있는 3D 구조 예측으로부터 식별된 잔기는 (RRXRR 및 REC 도메인 내의) MG35-419

서열 위의 짙은 회색 박스로 표시된다.

도 7a - 도 7b는 SMART 효소의 다양한 특징부를 나타낸다. 이는 다음을 나타낸다: (**도 7a**)는 본원에 도식된 다양한 효소의 SMART I 도메인 내 spCas9의 SMART I 도메인의 동일성을 나타내는 점 도표로서, 이는 최대 약 35%의 서열 동일성을 갖는다는 것을 나타낸다; (**도 7b**)는 본원에 기술된 효소의 개별 SMART I 도메인의 길이의 점 도표이다.

도 8a - 도 8b는 다양한 SMART-특이적 모티프 대 Cas9 뉴클레아제 서열에서 예측된 모티프의 계수 분포를 나타내며, 이는 이들 모티프가 SMART 효소에서 보다 흔히 발생함을 나타낸다; 모티프는 803개의 참조 Cas9 서열(II-A, II-B, 및 II-C형), 84개의 SMART I 서열, 및 471개의 SMART II 서열에 대해 예측되었다. 이는 다음을 나타낸다: (**도 8a**)는 다양한 유형의 클래스 2 Cas 효소에서의 Zn-결합 리본 모티프(CX_[2-4]C 및 CX_[2-4]H)의 계수 빈도의 박스 플롯이고; (**도 8b**)는 다양한 유형의 클래스 2 Cas 효소에서의 RRXRR 모티프의 계수 빈도의 히스토그램이다. (**도 8a**) 및 (**도 8b**)에서의 선은 평균 계수 값을 따르는 한편, 이상치는 점으로 표시된다.

도 9a - 도 9d는 SMART I 엔도뉴클레아제를 이용한 절단 활성에 대한, 설계된 단일 가이드 RNA(sgRNA)의 예측된 가이드 RNA 구조체를 도시한다. 나타난 것은 다음과 같다: (**도 9a**) MG34-1 sgRNA 1; (**도 9b**) MG34-1 sgRNA 2; (**도 9c**) MG34-9 sgRNA 1, 및 (**도 9d**) MG34-16 sgRNA 1.

도 10a - 도 10b는 실시예 1에 기술된 바와 같은 SMART I 뉴클레아제의 절단 특성화를 도시한다. (**도 10a**)는 2가지 sgRNA 설계를 갖는 MG34-1에 대한 절단 검정의 결과 산물의 Agilent TapeStation 겔을 음성 대조군과 비교하여 나타낸다. 레인 L3: 래더. 레인 A4: Apo, sgRNA 없음. 레인 B4 및 C4: 시험된 MG34-1 sgRNA(sg1: 서열번호 612, sg2: 613). 절단 산물 밴드는 화살표로 표시된다. 레인 G3 및 H3: 회색으로 표시됨, 본 실험과 관련이 없음. (**도 10b**)는 MG34-1, 34-9 및 34-16의 활성을 나타내는 결과 산물의 PCR 겔을 나타낸다. 레인 L1: 래더. 레인 2-7: MG34-1에 대해 6개 스페이서 길이를 갖는 sgRNA 설계. 레인 8 및 9: 34-9 및 34-16에 대한 각각의 sgRNA 설계. 화살표는 절단 확인 밴드를 나타낸다.

도 11a - 도 11c는 MG34 뉴클레아제에 대한 서열 절단 선호도를 나타낸다. (**도 11a**)는 sgRNA 1(상단, 서열번호 612) 및 sgRNA 2(하단, 서열번호 613)를 갖는 MG34-1에 대한 컨센서스 PAM 서열(NGGN)의 SeqLogo 표현을 나타낸다. (**도 11b**)는 MG34-1에 대한 절단 부위의 위치를 나타내는 히스토그램을 나타내며, 이는 MG34-1이 PAM으로부터 대략 위치 7에서의 절단을 선호함을 입증한다. (**도 11c**)는 MG34-9(박스로 강조 표시됨)에 대한 바람직한 NGG PAM을 나타내는 생거 시퀀싱 크로마토그램을 나타낸다. 화살표는 PAM으로부터 위치 7에서의 절단 부위를 나타낸다.

도 12a - 도 12c는 MG 34-1에 대한 대장균에서의 플라스미드 표적화 실험의 결과를 나타낸다. (**도 12a**)는 플라스미드 절단을 입증하는 대장균 균주의 복제 플레이팅을 나타낸다; MG34-1 및 sgRNA를 발현하는 대장균을 sgRNA(+sp)에 대한 표적을 함유하는 카나마이신 내성 플라스미드로 형질전환시켰다. 음성 대조군(표적 및 PAM(-sp) 없음) 대비 성장 손상(+sp)을 나타내는 플레이트 사분면은 효소에 의한 성공적인 표적화 및 절단을 나타낸다. 실험은 2회 반복으로 3회 수행되었다. (**도 12b**)는 비표적 대조군(-sp) 대비 표적 조건(+sp)에서의 성장 억제제를 나타내는 도 12a의 복제 플레이팅 실험으로부터의 콜로니 형성 유닛(cfu) 측정 도표를 나타내며, 이는 플라스미드가 절단되었음을 입증한다. (**도 12c**)는 콜로니 형성 유닛(cfu) 측정(로그 스케일)의 막대 도표를 나타내며, 이는 비표적 대조군(녹색 막대) 대비 표적 조건(백색 막대)에서의 대장균 성장 억제제를 나타낸다. 각각의 뉴클레아제에 대한 플라스미드 간섭 검정을 SpCas9 양성 대조군과 함께 3회 수행하였다.

도 13a - 도 13b는 MG35-419에 대한 SMART 시스템의 예시적인 게놈 컨텍스트를 나타낸다. SMART 뉴클레아제는 짙은 회색 화살표로 도시되어 있고, 다른 유전자는 이보다 밝은 회색 화살표로 도시되어 있다. 게놈 단편의 모든 유전자에 대해 예측된 도메인은 화살표 아래에 회색 박스로 도시되어 있다. 환경 발현 시퀀싱 판독은 (**도 13a**)의 CRISPR 어레이 아래, 그리고 (**도 13b**)의 효과기로부터의 상류에 정렬되어 제시된다. 발현을 나타내는 영역에 대한 전사체 커버리지는 콘티그 서열 위에 제시되어 있다. (**도 13a**)는 SMART II MG35-419 효과기 및 그 부근에 암호화된 CRISPR 유전자좌의 게놈 컨텍스트를 나타낸다. (**도 13b**)는 전사된 5' UTR을 나타내는 SMART II 효과기 MG35-3의 게놈 컨텍스트를 나타낸다.

도 14는 SMART II MG35-419에 대한 3D 구조 예측을 나타낸다. 이러한 3D 모델은 SaCas9 결정 구조의 영역 크기의 절반 미만임에도 불구하고, 이의 결정 구조의 영역과 양호하게 정렬된다. SaCas9 템플릿과 정렬된 영역은 측매 로브(RuvC-I, HNH 및 RuvC-III 도메인) 및 짧은 인식 영역(REC) 로브를 포함한다. SMART II-특이적 도메인은 RRXRR 모티프 및 Pfam PF14239에 대한 상동성을 함유하는 도메인, 및 알려지지 않은 기능의 도메인을 포함한

다.

도 15는 SMART II 효과기에 대한 예비 절단 검정의 결과를 도시한다. MG35-420(서열번호 223) 단백질 분취물을 전체 유전자좌가 발현된 TXTL 추출물 중에서의 절단 활성에 대해 시험하였다. 실험은 PAM 라이브러리(dsDNA 표적), 순방향 및 역방향 배향(fw 및 rv) 둘 모두에서 유전자좌에서 예측된 반복 영역(cr1), 및 잠재적으로 관련 보조인자를 암호화하는 유전자간 영역을 갖는 단백질 분취물을 인큐베이션하였다. 레인 2-9(cr 어레이 없음): 반복 영역이 없는 대조군 실험. Apo: 표적 PAM 라이브러리를 갖는 단백질 분취물 단독. 표지 1 - 2.5는 7개의 상이한 유전자간 영역을 나타낸다. -IG: 대조군으로서 포함된 유전자간 영역 없음. 결합 산물의 PCR 겔은 dsDNA 절단을 시사하는 추정 절단 밴드(화살표)를 나타낸다.

도 16a - 도 16b는 SMART 시스템의 게놈 컨텍스트를 도시한다. SMART 뉴클레아제는 짙은 회색 화살표로 도시되어 있고, 다른 유전자는 이보다 밝은 회색 화살표로 도시되어 있다. 게놈 단편의 모든 유전자에 대해 예측된 도메인은 화살표 아래에 회색 박스로 도시되어 있다. 환경 발현 시퀀싱 관독은 효과기로부터의 상류에 정렬되어 제시된다. **도 16a**는 SMART II MG35-419 효과기의 게놈 컨텍스트를 도시한다. **도 16b**는 SMART II MG35-102 효과기의 게놈 컨텍스트를 도시한다.

도 17a - 도 17b는 MG35-420이 활성 dsDNA 뉴클레아제임을 입증하는 데이터를 도시한다. **도 17a**는 MG35-420 효과기의 게놈 컨텍스트를 도시한다. 효과기는 역방향 배향의 어두운 화살표로 표시되고, 예측된 PFAM 도메인은 화살표 아래의 직사각형으로 표시되며, 가이드 RNA를 암호화할 수 있는 유전자간 영역은 흑색 선 상의 "IG"로 주석이 달린다. CRISPR-유사 반복 영역이 콘티그 내에 존재한다. **도 17b**는 TXTL에서 절단 활성에 대해 시험된 정제된 단백질 분취물의 결과를 도시한다. 실험은 PAM 라이브러리(dsDNA 표적), 순방향 및 역방향 배향(fw 및 rv) 둘 모두에서 유전자좌에서 예측된 CRISPR-유사 반복 영역(cr1), 및 잠재적으로 관련 보조인자를 암호화하는 유전자간 영역을 갖는 정제된 단백질 분취물을 인큐베이션하였다. 레인 2-9(cr 어레이 없음): 반복 영역이 없는 대조군 실험. Apo: 표적 PAM 라이브러리를 갖는 단백질 분취물 단독. 표지 1 - 2.5는 7개의 상이한 유전자간 영역을 나타낸다. -IG: 대조군으로서 포함된 유전자간 영역 없음. 결합 산물의 PCR 겔은 dsDNA 절단을 시사하는 추정 절단 밴드(화살표)를 나타낸다. "4"로 표시된 레인 상에서 회수된 밴드는 CRISPR-유사 영역 및 SMART II 5' UTR와 함께 효소를 인큐베이션한 것으로부터의 절단 밴드를 나타낸다.

도 18a - 도 18b는 MG35-420에 대한 예측된 가이드 RNA를 도시한다. **도 18a**는 5' UTR을 갖는 SMART II 효과기의 시험관 내 전사 반응으로부터 서열화된 RNASeq 관독을 나타내는 MG35-420 효과기의 게놈 컨텍스트를 도시한다. 효과기는 역방향 배향의 어두운 화살표로 표시되고, 예측된 PFAM 도메인은 화살표 아래의 직사각형으로 표시되고, 예측된 가이드 RNA는 흑색 선 상에 주석으로 달린다. **도 18b**는 SMART II MG35-420 추정 가이드 RNA의 이차 구조체 표현을 도시한다.

도 19a - 도 19b는 SMART II 효과기와 연관된 보존된 UTR 영역의 다중 서열 정렬(MSA)을 도시한다. **도 19a**는 SMART II 효과기의 시작 코돈으로부터 바로 상류에 있는 영역의 전장 MSA를 도시한다. 해당 정렬 위의 동일성 백분율 히스토그램은 보존 영역을 나타낸다(5' UTR 가이드 RNA로 주석이 달림, 회색 화살표). **도 19b**는 추정 가이드 RNA 암호화 서열 내의 고도로 보존된 영역을 도시한다. 동일성 백분율 히스토그램 및 서열 로고 표현이 정렬 위에 도시된다, 동일한 염기는 흑색 박스로 강조 표시된다.

도 20a - 도 20b는 MG35 효과기가 sgRNA를 사용하는 활성 dsDNA 뉴클레아제임을 입증하는 데이터를 도시한다. **도 20a**는 시험관 내 절단 검정의 결과를 도시한다. sgRNA 갖는 효과기(sg) 및 이를 갖지 않는 효과기(Apo)를 PAM 라이브러리(dsDNA 표적)와 함께 인큐베이션한 시험관 내 전사/번역 반응으로 분석하였다. 절단 산물은 PCR(예상되는 크기에서의 뉴클레아제 생성 밴드에 의한 성공적인 RNA 가이드 절단; 화살표)을 통해 증폭되었다. **도 20b**는 표적 인접 모티프(TAM)를 도시한다.

도 21a - 도 21f는 SMART 효소가 다양한 표적화 능력을 갖는 신규 뉴클레아제임을 입증하는 데이터를 도시한다. **도 21a**는 SpCas9 대비 SMART 뉴클레아제의 예측된 도메인 아키텍처를 도시한다. **도 21b**는 SMART MG102-2 시스템의 게놈 컨텍스트를 도시한다. tracrRNA 및 CRISPR 어레이 배향을 효과기를 이용한 시험관 내 절단 활성으로 확인하였다. **도 21c**는 SMART MG34-1 시스템의 게놈 컨텍스트를 도시한다. 적응 모듈 유전자(Cas1, Cas2, Cas4 및 추정 Csn2)가 식별되었다, 환경 RNASeq 관독은 tracrRNA를 암호화하는 어레이 및 유전자간 영역에 대한 순방향 배향으로 맵핑되었다. 유전자좌에 암호화된 다른 유전자는 황색 화살표로 표시된다. tracrRNA 및 CRISPR 어레이 배향을 효과기를 이용한 시험관 내 절단 활성으로 확인하였다. **도 21d**는 2개의 활성 SMART HEARO 뉴클레아제에 대한 HEARO RNA 이차 구조를 도시한다. 컨센서스 표적 모티프 서열의 SeqLogo 표현이 제시된다. **도 21e**는 Cas9 및 IscB 참조 서열 대비 SMART 뉴클레아제의 계통유전학 단백질 트리를 도시한다. SMART 효과기 및 고세균 Cas9

서열(청록색 및 보라색 분지)은 보고된 Cas9 참조 서열(II-A, II-B, 및 II-C형, 회색 분지)과 거리를 두고 관련되었다. 해당 트리는 공유된 RuvC-II/HNH/RuvC-III 도메인의 다중 서열 정렬로부터 추론되었다. SMART MG33 계열의 뉴클레아제(와인색 분지) 클러스터는 CRISPR II-C형 변이체 시스템을 갖는 한편, 다른 CRISPR-연관 SMART 뉴클레아제(청록색 분지) 클러스터는 최근 II-D형으로 분류된 서열을 가짐. SMART HEARO 뉴클레아제(라일락 분지)는 HEARO ORF 및 IscB 서열을 갖는다. **도 21f**는 SMART CRISPR II형 패밀리의 계통유전학 계통군(clade)을 도시한다. 해당 계통군은 **도 21e**에 도시된 계통유전학 트리를 확대하여 표현한 것이다. 내부 계통 분할 노드에 대한 로컬 지원 값이 제시되며, 이의 범위는 0 내지 1이다. 활성 SMART 뉴클레아제에 대한 생화학적 절단 활성 검정으로부터의 컨센서스 표적 모티프 서열 및 sgRNA 설계의 SeqLogo 표현이 제시된다.

도 22a - 도 22d는 SMART I이 dsDNA 뉴클레아제임을 입증하는 데이터를 도시한다. **도 22a**는 MG34-1이 PAM으로부터의 위치 7에서 dsDNA를 우선적으로 절단하는 것을 나타내는 절단 위치 선호도의 히스토그램을 도시한다. 삽입도는 MG34-1이 엇갈린 절단부를 생성하며, 여기에서 위치 3에서의 절단은 표적 가닥(TS) 상에서 발생하는 한편, 위치 6-7에서의 절단은 비표적 가닥(NTS) 상에서 발생한다는 것을 나타낸다. **도 22b**는 MG34-1에 대해 18 bp 스페이서에 대한 선호도를 나타내는, 다양한 스페이서 길이의 DNA 절단 백분율의 분포를 도시한다. **도 22c**는 MG34-1에 대한 시계열 절단 검정을 도시하며, 이는 SpCas9 대비 느린 동역학을 시사한다. **도 22d**는 플라스미드 표적화 검정을 도시한다. 좌측: 해당 방법의 다이어그램은 효과적 뉴클레아제(MG34-1 또는 MG34-9) 및 sgRNA 보조인자를 발현하는 조작된 대장균 균주를 나타낸다. 표적 또는 비표적 스페이서(음성 대조군)를 갖는 항생제 내성 유전자를 함유하는 플라스미드로 형질전환될 경우, 표적 플라스미드에 대한 성장 장애가 발생하였다. 중간 및 우측: MG34-1(중간) 또는 MG34-9(우측) 효소 및 sgRNA를 암호화하는 플라스미드에 대한 대략 2배의 성장 억제제를 나타내는 막대 도표.

도 23은 SMART HNH 엔도뉴클레아제-연관 RNA 및 ORF(HEARO)(35-1, 35-2, 35-3, 35-6, 35-102, 및 IscB) 및 SMART(34-1, 102-2, 102-14, 102-35, 102-45) 뉴클레아제의 군에 대한 전체 단백질 길이에 걸친 아미노산 함량 백분율을 도시한다. 높은 백분율의 아르기닌(R) 및 리신(K) 함량은 녹색으로 강조되는 한편, 낮은 메티오닌(M) 함량은 주황색으로 강조된다. Uniref50 데이터베이스(Carugo, vol. 17,12 (2008): 2187-91)의 대부분의 단백질의 아미노산 함량 백분율을 비교에 사용하였다.

도 24a는 SMART 단백질 중 아미노산 함량의 백분율(Y축) 대비 Uniref50 데이터베이스에서의 단백질의 평균 아미노산 함량(X축)의 산포도를 도시한다. 아르기닌(R) 및 리신(K) 함량은 선형 추세로부터 벗어난다. **도 24b**는 Uniref50 데이터베이스에서의 아미노산 백분율에 대한 SMART 단백질 중 아미노산 백분율의 비율을 나타내는 도표를 도시한다. 모든 비율의 평균은 0.99이고, SD는 0.22이다. 정상성을 가정할 경우, 녹색 선은 평균으로부터 2개의 표준 편차를 나타낸다.

도 25a - 도 25d는 SMART 효소가 dsDNA 뉴클레아제임을 입증하는 데이터를 도시한다. **도 25a**는 차세대 시퀀싱(NGS)으로부터의 비표적 가닥(NTS) 상의 3개의 SMART 뉴클레아제에 대한 절단 위치 선호도의 히스토그램을 도시한다. 삽입도는 SMART 뉴클레아제가 엇갈린 절단부를 생성하며, 여기에서 위치 3에서의 절단은 표적 가닥(TS) 상에서 발생하는 한편, PAM으로부터의 위치 5-7에서의 절단은 NTS 상에서 발생한다는 것을 나타낸다. TS 절단 부위는 생거 런-오프 시퀀싱을 통해 결정되었다. **도 25b**는 콜로니 형성 유닛(cfu) 측정(로그 스케일)의 막대 도표를 도시하며, 이는 비표적 대조군 대비 표적 조건에서의 대장균 성장 억제를 나타낸다. 각각의 뉴클레아제에 대한 플라스미드 간섭 검정을 SpCas9 양성 대조군과 함께 3회 수행하였다. **도 25c**는 다양한 스페이서 길이에 대한 시험관 내 DNA 절단 효율의 측정치를 도시하며, 이는 SMART 뉴클레아제에 대해 18-20 bp 스페이서에 대한 선호도를 나타내는 한편, SMART HEARO 35-1은 24 bp 스페이서를 선호한다. (*) 스페이서 길이 14 bp(34-1) 및 30 bp(35-1 및 102-2)는 평가되지 않았다. **도 25d**는 PAM으로부터 위치 -1 내지 -13에서 표적 스페이서에 대한 높은 특이성을 나타내는 불일치 사멸 검정을 도시한다. 좌측: 콜로니 형성 유닛(cfu) 측정(로그 스케일)의 막대 도표로서, 이는 스페이서 함유 불일치 및 비표적 대조군 대비 표적 조건에서의 대장균 성장 억제를 나타낸다. 우측 상단: 불일치 사멸 검정의 다이어그램. 뉴클레아제 발현 및 가이드 발현을 위한 2개의 플라스미드를 함유하는 대장균은 프로토스페이서에서 불일치를 갖는 표적 플라스미드 라이브러리로 형질전환되었다. 우측 하단: 표적 스페이서의 각각의 위치에서의 불일치 허용도를 나타내는 히트맵. 허용 가능한 불일치를 갖는 표적 스페이서 및 스페이서의 경우, 성장이 억제될 것으로 예상되었다(자주색). 필요한 염기 페어링을 갖는 위치는 효율적으로 절단되지 않을 것이고 출력 라이브러리(황색)에서 비교적 풍부할 것이다. 각각의 뉴클레아제에 대한 라이브러리를 이용한 플라스미드 간섭(사멸) 검정을 2회 수행하였다.

도 26은 MG102-2가 인간 세포에서 고도의 활성 뉴클레아제임을 입증하는 데이터를 도시한다. sgRNA의 농도를 증가시키면서(150, 300 및 450 pmol/반응), TRAC 유전자좌(가이드 A1 및 B1)에서 MG102-2 mRNA 및 2개의 sgRNA

표적화 부위를 핵감염시킴으로써 뉴클레아제 활성을 시험하였다. 모의 대조군은 mRNA 및 가이드 부재시 표적 영역에서의 백그라운드 편집 수준을 나타낸다.

도 27은 불일치 사멸 검정을 도시하며, 이는 MG102-2 및 MG35-1에 대해 시험된 스페이서의 각각의 위치에서 불일치를 갖는 스페이서에 대한 로그 배수 변화 절단 활성을 나타낸다.

도 28은 SMART 뉴클레아제가 ssDNA 상에서 활성을 나타내지 않음을 입증하는 데이터를 도시한다.

도 29는 SMART 뉴클레아제에 대한 가이드 및 염 농도 적정을 도시한다. MG102-2(라인 1-6) 및 SMART HEARO 35-1(라인 7-18)에 대한 시험관 내 절단 검정은 표적 플라스미드 DNA(약 3500 bp)가 선형 DNA 산물(2500 bp 미만)로 절단됨을 나타낸다.

도 30a - 도 30g는 인간 세포에서의 SMART I 편집 효율을 입증하는 데이터를 도시한다. 유전자좌의 다중 부위를 표적화하는 SMART I mRNA 및 sgRNA(450 pmol/반응)를 핵감염시킴으로써 뉴클레아제 활성을 시험하였다. 각각의 막대는 특정 스페이서(가이드)에 의해 표적화된 부위에서의 편집 효율을 나타낸다. **도 30a**는 AAVS1 유전자좌를 표적화하는 MG102-2에 대한 데이터를 도시한다. **도 30b, 30c, 30d, 30e, 30f, 및 30g**는 TRAC 유전자좌를 각각 표적화하는 MG102-39, MG102-42, MG102-48, MG33-34, MG102-26, 및 MG102-45에 대한 데이터를 도시한다.

도 31은 4개의 SMART HEARO 뉴클레아제의 5' UTR 뉴클레오타이드 서열의 다중 서열 정렬을 도시한다. HEARO RNA의 시작 전 영역(박스)은 불량한 유사성을 나타내는 한편, 제1 구조적 헤어핀 주위의 강한 보존이 관찰된다(삽입도).

도 32a - 도 32g는 SMART HEARO 뉴클레아제의 게놈 컨텍스트를 도시한다. SMART HEARO 뉴클레아제의 대부분은 CRISPR-연관은 아니지만(예를 들어, MG35-104, **도 32a**), 소수의 SMART HEARO 뉴클레아제는 CRISPR 어레이와 연관된다(예를 들어, **도 32b** 및 **32c**의 MG35-463 및 MG35-556). SMART HEARO 뉴클레아제는 해당 유전자 아래에 주석이 달린 RRR 및 HNH PAM 도메인과 함께 짙은 회색 화살표로 표시된다. 공분산 모델(CM)로부터 예측된 HEARO RNA는 SMART HEARO 효과기 유전자(CM HEARO RNA)로부터 상류에 나타났다. RAR: 반복-항반복. **도 32d-32g**는 다음과 같은 3개의 활성 뉴클레아제에 대한 HEARO RNA 이차 구조를 도시한다: 각각, MG35-104 sg1, MG35-463 sg2(CRISPR-독립적), MG35-463 sg3(CRISPR-연관), 및 MG35-556 이중 가이드 HEARO RNA(CRISPR-연관).

도 33a - 도 33c는 시험관 내 SMART HEARO 절단 활성을 도시한다. SMART II 효과기를 이의 단일 가이드 RNA 및 PAM 라이브러리(dsDNA 표적)와 함께 인큐베이션한 시험관 내 전사/번역 반응으로 분석하였다. 절단 산물은 절단 부위에 대한 결찰 및 후속 PCR(예상되는 크기에서의 뉴클레아제 생성 밴드에 의한 성공적인 RNA 가이드 절단: 화살표)을 통해 증폭되었다. **도 33a**의 경우, 라인 라벨은 다음과 같다: L: 래더; PC: 양성 대조군(PC)으로서의 MG35-1 뉴클레아제; 1: MG35-94; 2: MG35-104; 3: MG35-346; 4: MG35-350; 5: MG35-423; 6: MG35-422; 7: MG35-461; 8: MG35-465; 9: MG35-515. **도 33b**의 경우, 라인 라벨은 다음과 같다: L: 래더; PC: 양성 대조군(PC)으로서의 MG35-1 뉴클레아제; 10: MG35-517; 11: sgRNA 설계 1을 갖는 MG35-518; 12: sgRNA 설계 2를 갖는 MG35-518; 13: MG35-519; 14: sgRNA 설계 1을 갖는 MG35-550; 15: sgRNA 설계 2를 갖는 MG35-550; 16: MG35-553; 17: sgRNA 설계 1을 갖는 MG35-554; 18: sgRNA 설계 2를 갖는 MG35-554; 19: MG35-555; 및 20: MG35-556. **도 33c**의 경우, SMART II 효과기를 TAM/PAM 농축 프로토콜을 통한 절단 활성에 대해 검정하였다. 효과기를 이의 단일 가이드 RNA의 존재 하에 시험관 내 전사/번역(IVTT) 반응으로 발현시키고, 이어서 PAM 라이브러리(dsDNA 표적)에 첨가하였다. 절단 산물은 절단 부위에 대한 결찰 및 후속 PCR(예상되는 크기에서의 뉴클레아제 생성 밴드에 의한 성공적인 RNA 가이드 절단: 화살표)을 통해 증폭되었다. 나타난 반응은 PCR 세정 전이므로, 프라이머 및 어댑터-이량체 밴드는 < 100 bp 크기로 관찰되었다.

도 34는 활성 SMART HEARO 뉴클레아제에 대한 TAM 인식 모티프를 도시한다. **도 33a-33c**에서 식별된 밴드의 NGS 시퀀싱을 사용하여 각각의 뉴클레아제에 대한 TAM 및 바람직한 절단 위치를 생성하였다. Geneious(Andronescu 2007)에 의해 예측된 작업 가이드의 구조가 그 안에 제시된다. 절단은 비표적 가닥 상의 위치 5-10 사이에서 발생한다.

도 35a - 도 35b는 활성 SMART HEARO 뉴클레아제에 대한 시험관 내 절단 효율을 도시한다. **도 35a**의 경우, 절단은 반응 생성물의 수퍼코일링된(절단되지 않은) 선형(절단된) 전이에 의해 측정되었고 Agilent TapeStation 상에서 시각화되었다. 화살표는 효소에 의해 성공적으로 표적화된 절단(선형화) 후의 초기 dsDNA 산물(수퍼코일링됨) 및 dsDNA 산물을 나타낸다. PE: PURExpress; sgRNA, 단일 가이드 RNA. **도 35b**는 **도 35a**로부터의 정량화의 막대 도표 표현을 나타낸다. DNA: RNP 반응 없는 DNA 단독 대조군(음성 대조군); Apo: sgRNA가 첨가되지 않은 RNP 반응; Ho1o: sgRNA와의 RNP 반응.

도 36a - 도 36b는 SMART HEARO 가이드 조작용 도시한다. 5개의 활성 SMART HEARO sgRNA는 이의 서열 중 하나 이상의 폴리T 관을 가졌다. 원래의 가이드 대비 활성을 비교하기 위해 후보 당 3개의 폴리T 돌연변이체 sgRNA를 설계하였다. 가이드를 시험관 내에서 전사하고 동일한 농도로 정규화한 다음, 시험관 내 절단 효율 반응에 사용하였다. **도 36a**는 MG35-518에 대한 폴리-T 영역 및 조작된 가이드 서열을 갖는 예시적인 가이드 RNA를 도시한다. **도 36b**는 고유 가이드 대비 조작된 SMART HEARO 가이드 RNA의 절단 효율을 도시한다. Apo: 가이드가 첨가되지 않음(음성 대조군); WT: 고유 가이드 RNA.

도 37a - 도 37d는 SMART I 뉴클레아제의 계통유전학 분석을 도시한다. 전역(g-ins-i) 또는 국소(l-ins-i) 다중 서열 정렬로부터의 FastTree 또는 RAxML로 계통유전학 트리를 추론하였다. 계통유전학적 불확실성을 설명하기 위해, 6개의 재구성된 서열을 다수의 트리로부터 수득하였다(폐쇄 원으로 강조된 노드: MG34-26, MG34-27, MG34-28, MG34-29, MG34-30 및 MG34-31).

도 38은 활성 MG34-1 뉴클레아제의 예측된 구조 대비 재구성된 SMART I MG34-30의 3D 구조 예측을 도시한다. 전체적인 단백질의 양호한 구조적 정렬이 두 구조 사이의 중첩 및 낮은 RMSD 값에 의해 관찰되었다.

도 39는 재구성된 SMART I 효과기가 활성 뉴클레아제임을 입증하는 데이터를 도시한다. 신규 SMART I 효과기를 PAM 농축 프로토콜을 통한 절단 활성에 대해 검정하였다. 효과기를 다른 활성 MG34 뉴클레아제로부터의 단일 가이드 RNA의 존재 하에 시험관 내 전사/번역(IVTT) 반응으로 발현시키고, PAM 라이브러리(dsDNA 표적)에 첨가하였다. 절단 산물은 절단 부위에 대한 결찰 및 후속 PCR 증폭(예상되는 180 bp 크기에서의 뉴클레아제 생성 밴드에 의한 성공적인 RNA 가이드 절단: 화살표)을 통해 증폭되었다. MG34-27 및 MG34-29는 3개의 시험된 가이드 RNA에서 명확한 활성을 나타냈다.

도 40은 연산 재구성으로부터의 활성 SMART I 뉴클레아제에 대한 PAM 인식 모티프를 도시한다. **도 39**에서 식별된 밴드의 NGS 시퀀싱을 사용하여 각각의 뉴클레아제에 대한 PAM 및 바람직한 절단 위치를 생성하였다. 절단은 비표적 가닥 상의 PAM으로부터 위치 6과 8 사이에서 발생한다.

서열 목록에 대한 간단한 설명

본원과 함께 제출된 서열 목록은 본 개시내용에 따른 방법, 조성물, 및 시스템에 사용하기 위한 예시적인 폴리뉴클레오티드 및 폴리펩티드 서열을 제공한다. 서열 목록 내 서열에 대한 예시적인 설명이 아래에 제시되어 있다.

MG33 뉴클레아제

서열번호 1, 463-486, 981-988, 및 1289-1312은 MG33 뉴클레아제의 전장 펩티드 서열을 나타낸다.

서열번호 199 및 669-670은 MG33 뉴클레아제와 함께 기능하도록 예측된 tracrRNA의 뉴클레오티드 서열을 나타낸다.

서열번호 201 및 1003-1005는 MG33 뉴클레아제와 함께 기능할 것으로 예측된, 예측된 단일 가이드 RNA(sgRNA) 서열의 뉴클레오티드 서열을 나타낸다. "N"은 가변 잔기를 나타내고, 비-N-잔기는 스캐폴드 서열을 나타낸다.

서열번호 1023-1028은 MG33 뉴클레아제와 호환 가능한 PAM 서열을 나타낸다.

서열번호 1045-1054는 본원에 기술된 MG33 뉴클레아제의 CRISPR 반복을 나타낸다.

MG34 뉴클레아제

서열번호 2-24, 487-488, 및 1313-1321은 MG34 뉴클레아제의 전장 펩티드 서열을 나타낸다.

서열번호 200은 MG34 뉴클레아제와 함께 기능하도록 예측된 tracrRNA의 뉴클레오티드 서열을 나타낸다.

서열번호 202, 203, 및 613-616은 MG34 뉴클레아제와 함께 기능할 것으로 예측된, 예측된 단일 가이드 RNA(sgRNA) 서열의 뉴클레오티드 서열을 나타낸다. "N"은 가변 잔기를 나타내고, 비-N-잔기는 스캐폴드 서열을 나타낸다.

서열번호 1023-1028은 MG34 뉴클레아제와 호환 가능한 PAM 서열을 나타낸다.

서열번호 1055-1057은 본원에 기술된 MG34 뉴클레아제의 CRISPR 반복을 나타낸다.

MG35 뉴클레아제

서열번호 25-198, 221-459, 489-580, 617-668, 및 674-675는 MG35 뉴클레아제의 전장 펩티드 서열을 나타낸다.

서열번호 460-461은 MG35 뉴클레아제와 동일한 유전자좌로부터 유래된 MG35 tracrRNA의 뉴클레오티드 서열을 나타낸다.

서열번호 462, 676, 및 1229-1230은 본원에 기술된 MG35 뉴클레아제의 CRISPR 반복을 나타낸다.

서열번호 677-686, 1006-1012, 및 1231-1259는 MG35 단일 가이드 RNA의 뉴클레오티드 서열을 나타낸다.

서열번호 687-974는 MG35 단일 가이드 RNA 암호화 서열의 뉴클레오티드 서열을 나타낸다.

서열번호 1029-1034는 MG35 뉴클레아제와 호환 가능한 PAM 서열을 나타낸다.

서열번호 1172-1228은 본원에 기술된 MG35 뉴클레아제를 암호화하는 유전자좌의 뉴클레오티드 서열을 나타낸다.

MG102 뉴클레아제

서열번호 581-612, 989-1002, 및 1260-1273은 MG102 뉴클레아제의 전장 펩티드 서열을 나타낸다.

서열번호 672-673은 MG102 뉴클레아제와 동일한 유전자좌로부터 유래된 MG102 tracrRNA의 뉴클레오티드 서열을 나타낸다.

서열번호 205-220은 본 개시에 따른 뉴클레아제에 첨부될 수 있는 예시적인 핵 국소화 서열(NLS)의 서열을 나타낸다.

서열번호 1013-1022는 MG102 단일 가이드 RNA의 뉴클레오티드 서열을 나타낸다.

서열번호 1035-1044는 MG102 뉴클레아제와 호환 가능한 PAM 서열을 나타낸다.

서열번호 1058-1072는 본원에 기술된 MG102 뉴클레아제의 CRISPR 반복을 나타낸다.

서열번호 1171은 본원에 기술된 MG102 뉴클레아제를 암호화하는 유전자좌의 뉴클레오티드 서열을 나타낸다.

MG143 뉴클레아제

서열번호 975는 MG143 뉴클레아제의 전장 펩티드 서열을 나타낸다.

서열번호 1073은 본원에 기술된 MG143 뉴클레아제의 CRISPR 반복을 나타낸다.

MG144 뉴클레아제

서열번호 976-979 및 1274-1288은 MG144 뉴클레아제의 전장 펩티드 서열을 나타낸다.

서열번호 1074-1077은 본원에 기술된 MG144 뉴클레아제의 CRISPR 반복을 나타낸다.

MG145 뉴클레아제

서열번호 980는 MG145 뉴클레아제의 전장 펩티드 서열을 나타낸다.

서열번호 1078은 본원에 기술된 MG145 뉴클레아제의 CRISPR 반복을 나타낸다.

MG102 TRAC 표적화

서열번호 1079-1082 및 1145-1166은 TRAC 표적 부위의 DNA 서열을 나타낸다.

서열번호 1083-1086 및 1123-1144는 TRAC를 표적화하기 위해 MG102 뉴클레아제와 기능하도록 조작된 sgRNA의 뉴클레오티드 서열을 나타낸다.

MG33 TRAC 표적화

서열번호 1167-1168은 TRAC를 표적화하기 위해 MG33 뉴클레아제와 기능하도록 조작된 sgRNA의 뉴클레오티드 서열을 나타낸다.

서열번호 1169-1170은 TRAC 표적 부위의 DNA 서열을 나타낸다.

AAVS1 표적화

서열번호 1087-1104는 AAVS1을 표적화하기 위해 MG102 뉴클레아제와 기능하도록 조작된 sgRNA의 뉴클레오티드

서열을 나타낸다.

서열번호 1105-1122은 AAVS1 표적 부위의 DNA 서열을 나타낸다.

발명을 실시하기 위한 구체적인 내용

- [0051] 본 발명의 다양한 구현예가 본원에 도시되고 기술되었지만, 이러한 구현예는 단지 예시로서 제공된다는 것은 당업자에게 명백할 것이다. 본 발명의 범위 내에서 다양한 변이, 변형, 및 치환이 당업자에게 발생할 수 있다. 본원에 기술된 본 발명의 구현예에 대한 다양한 대안이 사용될 수 있음을 이해해야 한다.
- [0052] 달리 명시되지 않는 한, 본원에 개시된 일부 방법을 실시하는 데에는 면역학, 생화학, 화학, 분자 생물학, 미생물학, 세포 생물학, 계통, 및 재조합 DNA의 기술이 사용된다. 예를 들어 Sambrook 및 Green, Molecular Cloning: A Laboratory Manual, 제4판(2012); the series Current Protocols in Molecular Biology (F. M. Ausubel 등(편); the series Methods In Enzymology (Academic Press, Inc.), PCR 2: A Practical Approach (M.J. MacPherson, B.D. Hames 및 G.R. Taylor(편) (1995)), Harlow and Lane(편) (1988) Antibodies, A Laboratory Manual, and Culture of Animal Cells: A Manual of Basic Technique and Specialized Applications, 제6판(R.I. Freshney(편) (2010))을 참조한다(이들은 그 전체가 참조로서 본원에 통합됨).
- [0053] 본원에서 사용되는 바와 같이, 단수 형태("a", "an" 및 "the")는 문맥상 달리 명시되지 않는 한, 복수 형태도 포함하도록 의도된다. 또한, 용어 "포함하는(including, includes, having, has, with)" 또는 이의 변형된 표현이 발명을 실시하기 위한 구체적인 내용 또는 청구범위에 사용되는 정도까지, 이러한 용어는 용어 "포함하는(comprising)"과 유사한 방식으로 포괄적인 것으로 의도된다.
- [0054] 용어 "약" 또는 "대략"은 당업자에 의해 결정되는 특정 값에 대한 허용 가능한 오차 범위 내의 것을 의미하며, 이는 값이 측정되거나 결정되는 방법, 즉, 측정 시스템의 한계에 부분적으로 좌우될 것이다. 예를 들어, "약"은 당 기술분야의 관행에 따라 하나 또는 둘 이상의 표준 편차 이내를 의미할 수 있다. 대안적으로, "약"은 주어진 값의 최대 20%, 최대 15%, 최대 10%, 최대 5%, 또는 최대 1%의 범위를 의미할 수 있다.
- [0055] 본원에서 사용되는 바와 같이, "세포"는 일반적으로 생물학적 세포를 지칭한다. 세포는 살아있는 유기체의 기본 구조, 기능, 또는 생물학적 단위일 수 있다. 세포는 하나 이상의 세포를 갖는 임의의 유기체로부터 유래될 수 있다. 일부 비제한적인 예는 다음을 포함한다: 원핵 세포, 진핵 세포, 박테리아 세포, 고세균 세포, 단세포 진핵생물의 세포, 원생동물 세포, 식물 유래의 세포(예를 들어, 식물 작물, 과일, 야채, 곡물, 대두, 옥수수(corn), 옥수수(maize), 밀, 쌀, 토마토, 쌀, 카사바, 사탕수수, 호박, 건초, 감자, 면, 대마, 담배, 개화 식물, 침엽수, 겉씨식물, 양치류, 석송, 뿔이끼류, 우산이끼, 이끼 유래의 세포), 해조류 세포(예를 들어, 보트리오크스 브라우니(*Botryococcus braunii*), 녹조류(*Chlamydomonas reinhardtii*), 클라미도모나스 라인하르트티, 나노클로롭시스 가디타나(*Nannochloropsis gaditana*), 클로렐라 피레노이도사(*Chlorella pyrenoidosa*), 쌍발이모자반(*Sargassum patens* C. Agardh 등), 해초(예를 들어, 켈프), 진균 세포(예를 들어, 효모 세포, 버섯 유래의 세포), 동물 세포, 무척추 동물(예를 들어, 초파리, 자포류, 극피동물, 선충 등) 유래의 세포, 척추동물(예를 들어, 생선, 양서류, 파충류, 새, 포유동물) 유래의 세포, 포유동물(예를 들어, 돼지, 젖소, 염소, 양, 설치류, 랫트, 마우스, 비인간 영장류, 인간 등) 유래의 세포, 등. 때로는, 세포는 천연 유기체로부터 유래되지 않는다(예를 들어, 세포는 합성으로 만들어질 수 있고, 이는 가끔 인공 세포라 불린다).
- [0056] 본원에서 사용되는 용어 "뉴클레오티드"는 일반적으로 염기-당-인산염의 조합을 지칭한다. 뉴클레오티드는 합성 뉴클레오티드를 포함할 수 있다. 뉴클레오티드는 합성 뉴클레오티드 유사체를 포함할 수 있다. 뉴클레오티드는 핵산 서열(예를 들어, 데옥시리보핵산(DNA) 및 리보핵산(RNA))의 단량체 단위일 수 있다. 뉴클레오티드라는 용어는 리보뉴클레오시드 삼인산, 아데노신 삼인산(ATP), 우리딘 삼인산(UTP), 시토신 삼인산(CTP), 구아노신 삼인산(GTP), 및 데옥시리보뉴클레오시드 삼인산, 예컨대 dATP, dCTP, dITP, dUTP, dGTP, dTTP, 또는 이들의 유도체를 포함할 수 있다. 이러한 유도체는, 예를 들어, [α S]dATP, 7-데아자-dGTP 및 7-데아자-dATP, 및 이를 함유하는 핵산 분자에 뉴클레아제 저항성을 부여하는 뉴클레오티드 유도체를 포함할 수 있다. 본원에서 사용되는 바와 같이, 용어 뉴클레오티드는 디데옥시리보뉴클레오시드 삼인산(ddNTP) 및 이들의 유도체를 지칭할 수 있다. 디데옥시리보뉴클레오시드 삼인산의 예시적인 예는 ddATP, ddCTP, ddGTP, ddITP, 및 ddTTP를 포함할 수 있지만, 이에 한정되지는 않는다. 뉴클레오티드는 표지되지 않거나, 예컨대 광학적으로 검출 가능한 모이어티(예를 들어, 형광단)를 포함하는 모이어티를 사용하여 검출 가능하게 표지될 수 있다. 표지화는 양자점(quantum dots)으로 수행될 수도 있다. 검출 가능한 표지는, 예를 들어, 방사성 동위원소, 형광 표지, 화학발광 표지, 생물발광 표지, 및 효소 표지를 포함할 수 있다. 뉴클레오티드의 형광 표지는 플루오레세인, 5-카복시플루오레세

인(FAM), 2'7'-디메톡시-4'5'-디클로로-6-카르복시플루오레세인(JOE), 로다민, 6-카르복시로다민(R6G), N,N,N',N'-테트라메틸-6-카르복시로다민(TAMRA), 6-카르복시-X-로다민(ROX), 4-(4'디메틸아미노페닐아조) 벤조산(DABCYL), 캐스케이드 블루(Cascade Blue), 오레곤 그린(Oregon Green), 텍사스 레드(Texas Red), 사아닌, 및 5-(2'-아미노에틸)아미노나프탈렌-1-설폰산(EDANS)를 포함할 수 있지만 이에 한정되지는 않는다. 형광 표지된 뉴클레오타이드의 특정 예는 다음을 포함할 수 있다: Perkin Elmer(Foster City, Calif)로부터 입수할 수 있는 [R6G]dUTP, [TAMRA]dUTP, [R110]dCTP, [R6G]dCTP, [TAMRA]dCTP, [JOE]ddATP, [R6G]ddATP, [FAM]ddCTP, [R110]ddCTP, [TAMRA]ddGTP, [ROX]ddTTP, [dR6G]ddATP, [dR110]ddCTP, [dTAMRA]ddGTP, 및 [dROX]ddTTP; Amersham(Arlington Heights, Ill.)으로부터 입수할 수 있는 FluoroLink 테옥시뉴클레오타이드, FluoroLink Cy3-dCTP, FluoroLink Cy5-dCTP, FluoroLink 플루오르 X-dCTP, FluoroLink Cy3-dUTP, 및 FluoroLink Cy5-dUTP; Boehringer Mannheim(Indianapolis, Ind.)으로부터 입수할 수 있는 플루오레세인-15-dATP, 플루오레세인-12-dUTP, 테트라메틸-로다민-6-dUTP, IR770-9-dATP, 플루오레세인-12-ddUTP, 플루오레세인-12-UTP, 및 플루오레세인-15-2'-dATP; 및 Molecular Probes(Eugene, Oreg.)로부터 입수할 수 있는 염색체 표지된 뉴클레오타이드, BODIPY-FL-14-UTP, BODIPY-FL-4-UTP, BODIPY-TMR-14-UTP, BODIPY-TMR-14-dUTP, BODIPY-TR-14-UTP, BODIPY-TR-14-dUTP, 캐스케이드 블루-7-UTP, 캐스케이드 블루-7-dUTP, 플루오레세인-12-UTP, 플루오레세인-12-dUTP, 오레곤 그린 488-5-dUTP, 로다민 그린-5-UTP, 로다민 그린-5-dUTP, 테트라메틸로다민-6-UTP, 테트라메틸로다민-6-dUTP, 텍사스 레드-5-UTP, 텍사스 레드-5-dUTP, 및 텍사스 레드-12-dUTP. 뉴클레오타이드는 화학적 변형에 의해 표지되거나 표시될 수도 있다. 화학적으로 변형된 단일 뉴클레오타이드는 비오틴-dNTP일 수 있다. 비오틴화된 dNTP의 일부 비제한적인 예는 다음을 포함할 수 있다: 비오틴-dATP(예를 들어, 비오-N6-ddATP, 비오틴-14-dATP), 비오틴-dCTP(예를 들어, 비오틴-11-dCTP, 비오틴-14-dCTP), 및 비오틴-dUTP(예를 들어, 비오틴-11-dUTP, 비오틴-16-dUTP, 비오틴-20-dUTP). 뉴클레오타이드는 뉴클레오타이드 유사체를 포함할 수 있다. 일부 구현예에서, 뉴클레오타이드 유사체는 뉴클레오타이드의 특정 화학적 특성을 변경시키지만 뉴클레오타이드 유사체가 의도된 기능을 수행하는 능력(예를 들어, RNA 또는 DNA 내의 다른 뉴클레오타이드에 혼성화)을 유지하도록 임의의 위치에서 변형되는 천연 뉴클레오타이드의 구조를 포함할 수 있다. 유도체화될 수 있는 뉴클레오타이드의 위치의 예는, 위치 5, 예를 들어, 5-(2-아미노)프로필 우리딘, 5-브로모 우리딘, 5-프로핀 우리딘, 5-프로페닐 우리딘 등; 위치 6, 예를 들어, 6-(2-아미노)프로필 우리딘; 아데노신 또는 구아노신의 경우 위치 8, 예를 들어, 8-브로모 구아노신, 8-클로로 구아노신, 8-플루오로구아노신 등을 포함한다. 뉴클레오타이드 유사체는 또한 데아자 뉴클레오타이드, 예를 들어, 7-데아자-아데노신: O- 및 N-변형된(예를 들어, 알킬화된, 예를 들어, N6-메틸 아데노신) 뉴클레오타이드; 및 Herdewijn의 문헌[Antisense Nucleic Acid Drug Dev., 2000 Aug. 10(4):297-310]에 기술된 것과 같은 다른 헤테로시클릭적으로 변형된 뉴클레오타이드 유사체를 포함한다. 뉴클레오타이드 유사체는 또한 뉴클레오타이드의 당 부분에 대한 변형을 포함할 수 있다. 예를 들어, 2' OH-기는 H, OR, R, F, Cl, Br, I, SH, SR, NH₂, NHR, NR₂, COOR, 또는 OR로부터 선택되는 기에 의해 치환될 수 있다(여기에서 R은 치환되거나 치환되지 않은 C1-C6 알킬, 알케닐, 알키닐, 아릴 등임). 다른 가능한 변형은 미국 특허 제5,858,988호 및 제6,291,438호에 기술되어 있다. 유도체화될 수 있는 뉴클레오타이드의 위치의 예는, 위치 5, 예를 들어, 5-(2-아미노)프로필 우리딘, 5-브로모 우리딘, 5-프로핀 우리딘, 5-프로페닐 우리딘 등; 위치 6, 예를 들어, 6-(2-아미노)프로필 우리딘; 아데노신 또는 구아노신의 경우 위치 8, 예를 들어, 8-브로모 구아노신, 8-클로로 구아노신, 8-플루오로구아노신 등을 포함한다. 뉴클레오타이드 유사체는 또한 데아자 뉴클레오타이드, 예를 들어, 7-데아자-아데노신: O- 및 N-변형된(예를 들어, 알킬화된, 예를 들어, N6-메틸 아데노신) 뉴클레오타이드; 및 Herdewijn의 문헌[Antisense Nucleic Acid Drug Dev., 2000 Aug. 10(4):297-310]에 기술된 것과 같은 다른 헤테로시클릭적으로 변형된 뉴클레오타이드 유사체를 포함한다.

[0057] 뉴클레오타이드 유사체는 또한 뉴클레오타이드의 당 부분에 대한 변형을 포함할 수 있다. 예를 들어, 2' OH-기는 H, OR, R, F, Cl, Br, I, SH, SR, NH₂, NHR, NR₂, COOR, 또는 OR로부터 선택되는 기에 의해 치환될 수 있다(여기에서 R은 치환되거나 치환되지 않은 C1-C6 알킬, 알케닐, 알키닐, 아릴 등임). 다른 가능한 변형은 미국 특허 제5,858,988호 및 제6,291,438호에 기술되어 있다.

[0058] 용어 "폴리뉴클레오타이드", "올리고뉴클레오타이드", 및 "핵산"은 일반적으로 임의의 길이를 가진 뉴클레오타이드의 중합체 형태를 지칭하도록 사용 교환적으로 사용되며, 상기 뉴클레오타이드는 단일 가닥, 이중 가닥, 또는 다중 가닥의 테옥시리보뉴클레오타이드 또는 리보뉴클레오타이드이거나, 이의 유사체일 수 있다. 폴리뉴클레오타이드는 세포에 대해 외인성이거나 내인성일 수 있다. 폴리뉴클레오타이드는 무세포 환경에서 존재할 수 있다. 폴리뉴클레오타이드는 유전자이거나 이의 단편일 수 있다. 폴리뉴클레오타이드는 DNA일 수 있다. 폴리뉴클레오타이드는 RNA일 수 있다. 폴리뉴클레오타이드는 임의의 3차원 구조를 가질 수 있고 임의의 기능을 수행할 수 있다. 폴리뉴클레오타이드는 하나 이상의 유사체(예를 들어, 백본, 당, 또는 핵염기가 변형된 유사체)를 포함할 수 있다. 존재하는 경우,

뉴클레오타이드 구조에 대한 변형은 중합체의 조립 전 또는 후에 부여될 수 있다. 유사체의 일부 비제한적인 예는 다음을 포함한다: 5-브로모우라실, 펩티드 핵산, 제노 핵산(xeno nucleic acid), 모르폴리노, 잠금 핵산, 글리콜 핵산, 트레오스 핵산, 디데옥시뉴클레오타이드, 코르디세핀, 7-테아자-GTP, 형광단 (예를 들어, 당류에 연결된 로다민 또는 플루오레세인), 티올 함유 뉴클레오타이드, 비오틴 연결된 뉴클레오타이드, 형광 염기 유사체, CpG 섬, 메틸-7-구아노신, 메틸화된 뉴클레오타이드, 이노신, 티오우리딘, 슈도우리딘, 디하이드로우리딘, 큐오신, 및 와이오신. 폴리뉴클레오타이드의 비제한적인 예는 다음을 포함한다: 유전자 또는 유전자 단편의 코딩 또는 비-코딩 영역, 연결 분석으로부터 정의된 유전자좌/유전자좌들, 엑손, 인트론, 메신저 RNA(mRNA), 전달 RNA (tRNA), 리보솜 RNA (rRNA), 짧은 간섭 RNA(siRNA), 짧은 헤어핀 RNA(shRNA), 마이크로-RNA (miRNA), 리보자임, cDNA, 재조합 폴리뉴클레오타이드, 분지형 폴리뉴클레오타이드, 플라스미드, 벡터, 임의의 서열의 단리된 DNA, 임의의 서열의 단리된 RNA, 무세포 DNA(cfDNA) 및 무세포 RNA(cfRNA)를 포함하는 무세포 폴리뉴클레오타이드, 핵산 프로브, 및 프라이머. 뉴클레오타이드의 서열은 비-뉴클레오타이드 성분에 의해 중단될 수 있다.

[0059] 용어 "형질감염(transfection 또는 transfected)"은 일반적으로 비-바이러스적인 방법 또는 바이러스-기반 방법에 의해 핵산을 세포 내로 도입하는 것을 지칭한다. 핵산 분자는 완전한 단백질 또는 이의 기능적 부분을 암호화하는 유전자 서열일 수 있다. 예를 들어, Sambrook 등의 문헌[1989, Molecular Cloning: A Laboratory Manual, 18.1-18.88(이는 본원에 참조로서 전체적으로 통합됨)]을 참조한다.

[0060] 용어 "펩티드", "폴리펩티드", 및 "단백질"은 일반적으로 펩티드 결합에 의해 결합된 적어도 2개의 아미노산 잔기로 이루어진 중합체를 지칭하도록 본원에서 상호 교환적으로 사용된다. 이 용어는 중합체의 특정 길이를 의미하지 않으며, 펩티드가 재조합 기술, 화학적 또는 효소적 합성을 사용해 생산되는지 또는 자연적으로 발생하는지를 암시하거나 구별하도록 의도되지도 않는다. 상기 용어는 자연적으로 발생하는 아미노산 중합체뿐만 아니라 적어도 하나의 변형된 아미노산을 포함하는 아미노산 중합체에도 적용된다. 일부 경우에, 중합체는 비-아미노산에 의해 중단될 수 있다. 상기 용어는 전장 단백질을 포함하는 임의의 길이의 아미노산 사슬, 및 2차 또는 3차 구조(예를 들어, 도메인)가 있거나 없는 단백질을 포함한다. 상기 용어는 예를 들어, 이황화 결합 형성, 당질화, 지질화, 아세틸화, 인산화, 산화, 및 표지 성분과의 접합과 같은 임의의 다른 조작에 의해 변형된 아미노산 중합체도 포함한다. 본원에서 사용되는 용어 "아미노산"은, 변형된 아미노산 및 아미노산 유사체를 포함하되 이에 한정되지 않는, 천연 및 비-천연 아미노산을 일반적으로 지칭한다. 변형된 아미노산은, 아미노산 상에는 자연적으로 존재하지 않는 기 또는 화학적 모이어티를 포함하도록 화학적으로 변형된 천연 아미노산 및 비-천연 아미노산을 포함할 수 있다. 아미노산 유사체는 아미노산 유도체를 지칭할 수 있다. 용어 "아미노산"은 D-아미노산 및 L-아미노산 둘 모두를 포함한다.

[0061] 본원에서 사용되는 바와 같이, 용어 "비-고유(non-native)"는 고유 핵산 또는 단백질에서 발견되지 않는 핵산 또는 폴리펩티드 서열을 일반적으로 지칭할 수 있다. 비-고유는 친화도 태그를 지칭할 수 있다. 비-고유는 융합을 지칭할 수 있다. 비-고유는 돌연변이, 삽입, 또는 결실을 포함하는 자연 발생 핵산 또는 폴리펩티드 서열을 지칭할 수 있다. 비-고유 서열은 비-고유 서열이 융합되는 핵산 또는 폴리펩티드 서열에 의해서도 나타날 수 있는 활성(예를 들어, 효소 활성, 금속전이효소 활성, 아세틸전이효소 활성, 키나아제 활성, 유비퀴틴화 활성 등)을 나타내거나 이를 암호화할 수 있다. 비-고유 핵산 또는 폴리펩티드 서열은 유전자 조작에 의해 자연 발생 핵산 또는 폴리펩티드 서열(또는 이의 변이체)에 연결되어 키메라 핵산 또는 폴리펩티드를 암호화하는 키메라 핵산 또는 폴리펩티드 서열을 생성할 수 있다.

[0062] 본원에서 사용되는 바와 같이, 용어 "프로모터"는, 유전자의 전사 또는 발현을 조절하고 RNA 전사가 개시되는 뉴클레오타이드 또는 뉴클레오타이드의 영역에 인접하게 위치하거나 이와 중첩될 수 있는 조절 DNA 영역을 일반적으로 지칭한다. 프로모터는 종종 전사 인자로서 지칭되는 단백질 인자에 결합하는 특이적 DNA 서열을 함유할 수 있는데, 상기 인자는 RNA 중합효소가 DNA에 결합하는 것을 용이하게 하여 유전자 전사를 유도한다. '코어 프로모터'로도 지칭되는 '기저 프로모터(basal promoter)'는 작동 가능하게 연결된 폴리뉴클레오타이드의 전사 발현을 촉진하는 모든 기본 요소를 함유하는 프로모터를 일반적으로 지칭할 수 있다. 진핵생물 기저 프로모터는, 반드시 그런 것은 아니지만, 통상적으로, TATA-박스 또는 CAAT 박스를 함유한다.

[0063] 본원에서 사용되는 바와 같이, 용어 "발현"은 핵산 서열 또는 폴리뉴클레오타이드가 DNA 주형으로부터 (예컨대 mRNA 또는 다른 RNA 전사체로) 전사되는 공정 또는 전사된 mRNA가 후속하여 펩티드, 폴리펩티드, 또는 단백질로 번역되는 공정을 일반적으로 지칭한다. 전사체 및 암호화된 폴리펩티드는 "유전자 산물"로서 통칭될 수 있다. 폴리뉴클레오타이드가 게놈 DNA로부터 유래되는 경우, 발현은 진핵 세포에서 mRNA의 스플라이싱을 포함할 수 있다.

- [0064] 본원에서 사용되는 바와 같이, "작동 가능하게 연결된(operably linked, operable linkage, operatively linked)" 또는 이와 문법적으로 동등한 표현은 유전자 요소, 예를 들어, 프로모터, 인핸서, 폴리아데닐화 서열 등의 병치를 일반적으로 지칭하는데, 여기서 요소들은 이들이 예상된 방식으로 작동하도록 허용하는 관계에 있다. 예를 들어, 프로모터 또는 인핸서 서열을 포함할 수 있는 조절 요소가 코딩 서열의 전사 개시에 도움을 주는 경우, 조절 요소는 코딩 영역에 작동 가능하게 연결된다. 이러한 기능적 관계가 유지되는 한, 조절 요소와 코딩 영역 사이에 개재 잔기가 있을 수 있다.
- [0065] 본원에서 사용되는 바와 같이, "벡터"는 일반적으로 폴리뉴클레오티드를 포함하거나 폴리뉴클레오티드와 결합하고 폴리뉴클레오티드를 세포로 전달하는 것을 매개하는데 사용될 수 있는 거대분자 또는 거대분자의 연관을 지칭한다. 벡터의 예는 플라스미드, 바이러스 벡터, 리포좀, 및 기타 유전자 전달 비히클을 포함한다. 벡터는 유전자에 작동 가능하게 연결되어 표적에서 유전자의 발현을 용이하게 하는 유전자 요소, 예를 들어 조절 요소를 일반적으로 포함한다.
- [0066] 본원에서 사용되는 바와 같이, "발현 카세트" 및 "핵산 카세트"는 함께 발현되거나 발현을 위해 작동 가능하게 연결된 핵산 서열 또는 요소의 조합을 지칭하기 위해 일반적으로 상호 교환적으로 사용된다. 일부 경우에, 발현 카세트는 조절 요소와 발현을 위해 조절 요소가 작동 가능하게 연결되는 유전자의 조합을 지칭한다.
- [0067] DNA 또는 단백질 서열의 "기능적 단편"은 전장 DNA 또는 단백질 서열의 생물학적 활성과 실질적으로 유사한 생물학적 활성(기능적 또는 구조적 활성)을 보유하는 단편을 일반적으로 지칭한다. DNA 서열의 생물학적 활성은 전장 서열에 기인한 방식으로 발현에 영향을 미치는 이의 능력일 수 있다.
- [0068] 본원에서 사용되는 바와 같이, "조작된" 객체란 객체가 인간 개입에 의해 변형되었음을 일반적으로 나타낸다. 비제한적인 실시예에 따르면: 핵산은 이의 서열을 자연에서 발생하지 않는 서열로 변경함으로써 변형될 수 있고; 핵산은 이 핵산을 자연에서 연관되지 않는 핵산과 결합시키되, 결합 산물이 원래 핵산에 존재하지 않는 기능을 갖도록 결합시킴으로써 변형될 수 있고; 조작된 핵산은 자연에서 존재하지 않는 서열을 이용해 시험관 내에서 합성될 수 있고; 단백질은 이의 아미노산 서열을 자연에서 존재하지 않는 서열과 치환함으로써 변형될 수 있고; 조작된 단백질은 새로운 기능 또는 특성을 획득할 수 있다. "조작된" 시스템은 적어도 하나의 조작된 구성요소를 포함한다.
- [0069] 본원에서 사용되는 바와 같이, 용어 "최적으로 정렬된"은 대체적으로 가장 높은 동일성 백분율 점수를 제공하거나 일치하는 잔기의 수를 최대화하는 2개의 아미노산 서열의 정렬을 지칭한다.
- [0070] 본원에서 사용되는 바와 같이, "합성" 및 "인공"은 자연 발생 인간 단백질과 낮은 서열 동일성(예를 들어, 50% 미만의 서열 동일성, 25% 미만의 서열 동일성, 10% 미만의 서열 동일성, 5% 미만의 서열 동일성, 1% 미만의 서열 동일성)을 갖는 단백질 또는 이의 도메인을 지칭하도록 상호 교환적으로 사용된다. 예를 들어, VPR 및 VP64 도메인은 합성 전사 활성화 도메인이다.
- [0071] 본원에서 사용되는 바와 같이, 용어 "tracrRNA" 또는 "tracr 서열"은 예시적인 야생형 tracrRNA 서열(예를 들어 화농성연쇄상구균, 황색포도상구균 등 유래의 tracrRNA)과 적어도 약 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 또는 100% 서열 동일성 또는 서열 유사성을 갖는 핵산을 일반적으로 지칭할 수 있다. tracrRNA는 예시적인 야생형 tracrRNA 서열(예를 들어 화농성연쇄상구균, 황색포도상구균 등 유래의 tracrRNA)과 최대 약 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 또는 100% 서열 동일성 또는 서열 유사성을 갖는 핵산을 일반적으로 지칭할 수 있다. tracrRNA는 tracrRNA의 변형된 형태로서, 결실, 삽입, 또는 치환과 같은 뉴클레오티드 변화, 변이체, 돌연변이, 또는 키메라를 포함할 수 있는 형태를 지칭할 수 있다. tracrRNA는 적어도 6개의 연속 뉴클레오티드의 길이에 걸쳐 야생형 예시적인 tracrRNA(예를 들어, 화농성연쇄상구균, 황색포도상구균 등의 tracrRNA) 서열과 적어도 약 60% 동일할 수 있는 핵산을 지칭할 수 있다. 예를 들어, tracrRNA 서열은 적어도 6개의 연속 뉴클레오티드의 길이에 걸쳐, 예시적인 야생형 tracrRNA(예를 들어, 화농성연쇄상구균, 황색포도상구균 등으로부터의 tracrRNA)과 적어도 약 60% 동일하거나, 적어도 약 65% 동일하거나, 적어도 약 70% 동일하거나, 적어도 약 75% 동일하거나, 적어도 약 80% 동일하거나, 적어도 약 85% 동일하거나, 적어도 약 90% 동일하거나, 적어도 약 95% 동일하거나, 적어도 약 98% 동일하거나, 적어도 약 99% 동일하거나, 100% 동일할 수 있다. II형 tracrRNA 서열은 인접한 CRISPR 어레이에서 반복 서열의 일부와 상보성을 갖는 영역을 식별함으로써 게놈 서열 상에서 예측될 수 있다.
- [0072] 본원에서 사용되는 바와 같이, "가이드 핵산"은 다른 핵산에 혼성화될 수 있는 핵산을 일반적으로 지칭할 수 있다. 가이드 핵산은 RNA일 수 있다. 가이드 핵산은 DNA일 수 있다. 가이드 핵산은 핵산의 서열에 부위 특이적

로 결합하도록 프로그래밍될 수 있다. 표적화될 핵산 또는 표적 핵산은 뉴클레오티드를 포함할 수 있다. 가이드 핵산은 뉴클레오티드를 포함할 수 있다. 표적 핵산의 일부는 가이드 핵산의 일부에 상보적일 수 있다. 가이드 핵산에 상보적이고 가이드 핵산과 혼성화되는 이중-가닥 표적 폴리뉴클레오티드의 가닥은 상보적 가닥으로 지칭될 수 있다. 상보적 가닥에 상보적이고, 따라서 가이드 핵산에 상보적이 아닐 수 있는 이중 가닥 표적 폴리뉴클레오티드의 가닥은 비상보적 가닥으로 지칭될 수 있다. 가이드 핵산은 하나의 폴리뉴클레오티드 사슬을 포함할 수 있고 "단일 가이드 핵산"으로 지칭될 수 있다. 가이드 핵산은 2개의 폴리뉴클레오티드 사슬을 포함할 수 있고 "이중 가이드 핵산"으로 지칭될 수 있다. 달리 명시되지 않는 경우, 용어 "가이드 핵산"은 단일 가이드 핵산 및 이중 가이드 핵산 둘 다를 포함할 수 있고, 둘 다를 지칭할 수 있다. 가이드 핵산은 "핵산-표적화 분절" 또는 "핵산-표적화 서열"로서 지칭될 수 있는 분절을 포함할 수 있다. 핵산-표적화 분절은 "단백질-결합 분절" 또는 "단백질-결합 서열" 또는 "Cas 단백질-결합 분절"로서 지칭될 수 있는 하위 분절을 포함할 수 있다.

[0073] 2개 이상의 핵산 또는 폴리펩티드 서열의 맥락에서의 용어 "서열 동일성" 또는 "동일성 백분율"은, 서열 비교 알고리즘을 사용해 측정했을 때, 부분적 또는 전체 비교 윈도우에 걸쳐 비교하고 최대 상응에 정렬했을 때 동일한 2개(예를 들어, 쌍으로 정렬했을 때) 또는 그 이상(예를 들어, 다수의 서열을 정렬했을 때)의 서열; 또는 동일한 아미노산 잔기 또는 뉴클레오티드의 특정 백분율을 갖는 2개(예를 들어, 쌍으로 정렬했을 때) 또는 그 이상(예를 들어, 다수의 서열을 정렬했을 때)의 서열을 일반적으로 지칭한다. 폴리펩티드 서열에 대한 적절한 서열 비교 알고리즘은 예를 들어 다음을 포함한다: 단어 길이(W) 3, 기대치(E) 10의 파라미터, 및 BLOSUM62 스코어링 매트릭스(존재 11, 연장 1의 갭 비용 설정)를 사용하고, 30개 잔기를 초과하는 길이의 폴리펩티드 서열에 대해서는 조건부 조성 스코어 매트릭스 조정(conditional compositional score matrix)을 사용하는 BLASTP; 단어 길이(W) 2, 기대치(E) 1000000의 파라미터, 및 30개 잔기 미만의 서열에 대해서는 PAM30 스코어링 매트릭스(개방 갭 9, 연장 갭 1의 갭 비용 설정 - 이들은 <https://blast.ncbi.nlm.nih.gov>에서 이용할 수 있는 BLAST 세트 중 BLASTP에 대한 디폴트 파라미터임)를 사용하는 BLASTP; 또는 CLUSTALW; 일치 2, 불일치 -1, 및 갭 -1의 파라미터를 사용하는 Smith-Waterman 상동성 검색 알고리즘; 디폴트 파라미터를 사용하는 MUSCLE; retree 2 및 최대 반복 1000의 파라미터를 사용하는 MAFFT; 디폴트 파라미터를 사용하는 Novafold; 디폴트 파라미터를 사용하는 HMMER hmalign.

[0074] 본원에서 사용되는 바와 같이, 용어 "RuvC_III 도메인"은 RuvC 엔도뉴클레아제 도메인의 세 번째 불연속 분절을 일반적으로 지칭한다(RuvC 뉴클레아제 도메인은 3개의 불연속 분절 RuvC_I, RuvC_II, 및 RuvC_III으로 구성됨). RuvC 도메인 또는 이의 분절(예를 들어, RuvC_I, RuvC_II, 또는 RuvC_III)은 문서화된 도메인 서열에 대한 정렬, 주석이 달린 도메인을 가진 단백질에 대한 구조 정렬, 또는 문서화된 도메인 서열(예를 들어, RuvC_III의 경우 Pfam HMM PF18541)에 기초하여 구축된 히든 마르코프 모델(Hidden Markov Models, HMM)과의 비교에 의해 일반적으로 식별될 수 있다.

[0075] 본원에서 사용되는 바와 같이, 용어 "HNH 도메인"은 특징적인 히스티딘 및 아스파라긴 잔기를 갖는 엔도뉴클레아제 도메인을 일반적으로 지칭한다. HNH 도메인은 문서화된 도메인 서열에 대한 정렬, 주석이 달린 도메인을 가진 단백질에 대한 구조 정렬, 또는 문서화된 도메인 서열(예를 들어, 도메인 HNH의 경우 Pfam HMM PF01844)에 기초하여 구축된 히든 마르코프 모델(HMM)과의 비교에 의해 일반적으로 식별될 수 있다.

[0076] 본원에서 사용되는 바와 같이, 용어 "가교 나선 도메인" 또는 "BH 도메인"은 대체적으로 표적 DNA의 결합 시 절단 활성을 개시하는 데 중요한 역할을 하는 Cas 효소에 존재하는 아르기닌-풍부 나선 도메인을 지칭한다.

[0077] 본원에서 사용되는 바와 같이, 용어 "인식 도메인" 또는 "REC 도메인"은 대체적으로 gRNA의 반복:항-반복 이중체와 상호작용하고 Cas 엔도뉴클레아제/gRNA 복합체의 형성을 매개하는 것으로 여겨지는 도메인을 지칭한다.

[0078] 본원에서 사용되는 바와 같이, 용어 "웨지 도메인" 또는 "WED 도메인"은 대체적으로 4개의 알파 나선이 측면에 위치한 꼬인 5-가닥 베타 시트를 포함하는 폴딩부를 지칭하며, 이는 대체적으로 Cas 효소에 대한 뒤틀린 반복:항-반복 이중체의 인식을 담당한다. WED 도메인은 단일 가이드 RNA 스캐폴드의 인식을 담당할 수 있다.

[0079] 본원에서 사용되는 바와 같이, 용어 "PAM 상호작용 도메인" 또는 "PI 도메인"은 대체적으로 가이드 RNA의 비-상보성 DNA 가닥 상의 PAM 서열을 인식하기 위해 엔도뉴클레아제-DNA-복합체에 위치한 Cas 효소에서 발견되는 도메인을 지칭한다.

[0080] 개요

[0081] 독특한 기능 및 구조를 갖는 새로운 Cas 효소의 발견은 데옥시리보핵산(DNA)을 추가로 파괴할 수 있는 편집 기술을 제공함으로써, 속도, 특이성, 기능, 및 사용 편의성을 개선할 수 있다. 미생물에서 CRISPR 시스템의 예측

유병률 및 미생물 종의 순수한 다양성과 관련하여, 기능적으로 특성화된 CRISPR/Cas 효소는 문헌에 상대적으로 거의 존재하지 않는다. 이는 부분적으로는 많은 수의 미생물 종이 실험실 조건에서 쉽게 배양되지 않을 수 있기 때문이다. 많은 수의 미생물 종을 나타내는 자연 환경 적소로부터의 메타게놈 시퀀싱은 문서화된 새로운 CRISPR/Cas 시스템의 수를 극적으로 증가시키고 새로운 올리고뉴클레오타이드 편집 기능의 발견을 가속화할 수 있는 가능성을 제공할 수 있다. 이러한 접근법의 결실에 대한 최근의 예는 2016년에 천연 미생물 군집의 메타게놈 분석에서 CasX/CasY CRISPR 시스템을 발견한 것에 의해 입증된다.

[0082] CRISPR/Cas 시스템은 미생물에서 적응성 면역 체계로서 기능하는 것으로 기술된 RNA-지향성 뉴클레아제 복합체이다. 이들의 자연적인 맥락에서, CRISPR/Cas 시스템은 CRISPR(일정한 간격을 두고 주기적으로 분포하는 짧은 회문 반복서열) 오페론 또는 유전자좌에서 발생하며, 이는 일반적으로 2개의 부분을 포함한다: (i) RNA-기반 표적화 요소를 암호화하는, 동일하게 짧은 스페이서 서열에 의해 분리된 짧은 반복 서열(30-40 bp)의 어레이; 및 (ii) 부속 단백질/효소와 함께 RNA-기반 표적화 요소가 지향하는 뉴클레아제 폴리펩티드를 암호화하는 Cas를 암호화하는 ORF. 특정 표적 핵산 서열의 효율적인 뉴클레아제 표적화는 일반적으로 다음 두 가지 모두를 필요로 한다: (i) 표적의 첫 6-8개의 핵산(표적 시드)과 crRNA 가이드 사이의 상보적 혼성화; 및 (ii) 표적 시드의 정의된 근위 이내에 프로토스페이서-인접 모티프(PAM) 서열의 존재(PAM은 일반적으로 숙주 게놈 내에서 흔히 나타나지 않는 서열임). 시스템의 정확한 기능 및 구성에 따라, CRISPR-Cas 시스템은 공통의 기능적 특성 및 진화적 유사성을 기반으로 일반적으로 2개의 클래스, 5개의 유형, 및 16개의 하위 유형으로 구성된다.

[0083] 클래스 I CRISPR-Cas 시스템은 큰 다중 서브유닛 작동자 복합체를 가지며, I형, III형, 및 IV형을 포함한다.

[0084] I형 CRISPR-Cas 시스템은 구성요소의 측면에서 복잡성이 보통인 것으로 간주된다. I형 CRISPR-Cas 시스템에서, RNA-표적화 요소의 어레이는 긴 전구체 crRNA(pre-crRNA)로서 전사되고, 이는 반복 요소에서 가공되어 짧고 성숙한 crRNA를 유리시키는데, 이는 프로토스페이서-인접 모티프(PAM)로 불리는 적절한 짧은 컨센서스 서열이 뒤에 이어질 때 뉴클레아제 복합체를 핵산 표적에 대해 유도한다. 이러한 가공은 캐스케이드(Cascade)로 불리는 큰 엔도뉴클레아제 복합체의 엔도리보뉴클레아제 서브유닛(Cas6)을 통해 이루어지는데, 이는 crRNA-지향성 뉴클레아제 복합체의 뉴클레아제(Cas3) 단백질 성분도 포함한다. Cas I 뉴클레아제는 주로 DNA 뉴클레아제로서 기능한다.

[0085] III형 CRISPR 시스템은 Csm 또는 Cmr 단백질 서브유닛을 포함하는 반복체-결합된 신비한 단백질(RAMP)과 함께 Cas10으로 알려진 중심 뉴클레아제의 존재를 특징으로 할 수 있다. I형 시스템에서와 같이, 성숙한 crRNA는 Cas6-유사 효소를 사용하여 pre-crRNA로부터 가공된다. I형 및 II형 시스템과 달리, III형 시스템은 DNA-RNA 이중체(예컨대, RNA 중합효소에 대한 탭플릿으로서 사용되는 DNA 가닥)를 표적화하고 절단하는 것으로 보인다.

[0086] IV형 CRISPR-Cas 시스템은 고도로 감소된 큰 서브유닛 뉴클레아제(csf1), Cas5(csf3) 및 Cas7(csf2)군의 RAMP 단백질에 대한 2개의 유전자, 및 일부 경우, 예측된 작은 서브유닛에 대한 유전자를 포함하는 효과기 복합체를 가지며; 이러한 시스템은 내인성 플라스미드 상에서 흔히 발견된다.

[0087] 클래스 II CRISPR-Cas 시스템은 단일-폴리펩티드 다중도메인 뉴클레아제 효과기를 일반적으로 가지며, II형, V형, 및 VI형을 포함한다.

[0088] II형 CRISPR-Cas 시스템은 구성요소 측면에서 가장 단순한 것으로 간주된다. II형 CRISPR-Cas 시스템에서, CRISPR 어레이를 성숙한 crRNA로 가공하는 데에는 특별한 엔도뉴클레아제 서브유닛의 존재가 필요하지 않고, 오히려 어레이 반복 서열에 상보적인 영역을 갖는 작은 트랜스-암호화된 crRNA(tracrRNA)가 필요하며; 여기서 tracrRNA는 이의 상응하는 효과기 뉴클레아제(예를 들어, Cas9) 및 반복 서열 둘 다와 상호작용하여 전구체 dsRNA 구조를 형성하는데, 이는 내인성 RNase III에 의해 절단되어 tracrRNA 및 crRNA 둘 다와 함께 로딩되는 성숙한 효과기 효소를 생성한다. Cas II 뉴클레아제는 DNA 뉴클레아제이다. II형 작동자는 대체적으로 RuvC-유사 뉴클레아제 도메인의 폴딩부 내에 삽입된 무관한 HNH 뉴클레아제 도메인과 함께 RNase H 접힘부를 입양하는 RuvC-유사 엔도뉴클레아제 도메인을 포함하는 구조를 나타낸다. RuvC-유사 도메인은 표적 (예를 들어, crRNA 상보적인) DNA 가닥의 절단을 담당하는 반면, HNH 도메인은 변위된 DNA 가닥의 절단을 담당한다.

[0089] V형 CRISPR-Cas 시스템은 II형 효과기의 구조와 유사하고, RuvC-유사 도메인을 포함하는 뉴클레아제 효과기(예를 들어, Cas 12) 구조를 특징으로 한다. II형과 유사하게, (전부는 아니지만) 대부분의 V형 CRISPR 시스템은 tracrRNA를 사용해 pre-crRNA를 성숙한 crRNA로 처리하지만; pre-crRNA를 다수의 crRNA로 절단하기 위해 RNase III을 필요로 하는 II형 시스템과 달리, V형 시스템은 효과기 뉴클레아제 자체를 사용해 pre-crRNA를 절단할 수 있다. II형 CRISPR-Cas 시스템과 마찬가지로, V형 CRISPR-Cas 시스템은 DNA 뉴클레아제이다. II형 CRISPR-Cas

시스템과 달리, 일부 V형 효소(예를 들어, Cas12a)는 이중-가닥 표적 서열의 제1 crRNA 가이드 절단에 의해 활성화되는 강력한 단일-가닥 비특이적 데옥시리보뉴클레아제 활성을 갖는 것으로 보인다.

[0090] VI형 CRISPR-Cas 시스템은 RNA-가이드된 RNA 엔도뉴클레아제를 갖는다. RuvC-유사 도메인 대신에, VI형 시스템의 단일 폴리펩티드 작동자(예를 들어, Cas13)는 2개의 HEPN 리보뉴클레아제 도메인을 포함한다. II형 및 V형 시스템 둘 모두와 상이하게, VI형 시스템 또한 pre-crRNA를 crRNA로 가공하는 데 tracrRNA를 필요로 하지 않는 것으로 보인다. 그러나, V형 시스템과 유사하게, 일부 VI형 시스템(예를 들어, C2C2)은 표적 RNA의 제1 crRNA 유도 절단에 의해 활성화되는 강력한 단일-가닥 비특이적 뉴클레아제(리보뉴클레아제) 활성을 갖는 것으로 보인다.

[0091] 아키텍처가 더 단순하기 때문에, 클래스 II CRISPR-Cas는 설계자 뉴클레아제/게놈 편집 애플리케이션으로서 조작 및 개발에 가장 널리 채택되어 왔다.

[0092] 이러한 시스템을 시험관 내 용도로 초기에 도입한 것 중 하나는 Jinek 등의 문헌에서 확인할 수 있다 (문헌 [Science. 2012 Aug 17;337(6096):816-21], 그 전체가 참조로서 본원에 통합됨). Jinek의 연구에서 다음을 포함하는 시스템이 처음 기술되었다: (i) 재조합적으로 발현되고, 화농성 연쇄상구균(*S. pyogenes*) SF370으로부터 단리된, 정제된 전장 Cas9(예를 들어, 클래스 II, II형 Cas 효소); (ii) 절단될 표적 DNA 서열에 상보적이고, 3' tracr-결합 서열이 이어지는, 약 20 nt의 5' 서열을 갖는 정제된 성숙한 약 42 nt의 crRNA(T7 프로모터 서열을 갖는 합성 DNA 템플릿으로부터 시험관 내 전사된 전체 crRNA); (iii) T7 프로모터 서열을 갖는 합성 DNA 템플릿으로부터 시험관 내 전사된 정제된 tracrRNA; 및 (iv) Mg²⁺. 이후에, Jinek은 (ii)의 crRNA가 링커(예를 들어, GAAA)에 의해 (iii)의 5' 말단에 결합되어 Cas9를 저절로 표적을 향하게 할 수 있는 단일 융합된 합성 가이드 RNA(sgRNA)를 형성하는 개선된, 조작된 시스템을 기술하였다(도 2의 상단 패널 및 하단 패널을 비교함).

[0093] 이후에, Mali 등은 다음을 암호화하는 DNA를 제공함으로써, 포유류 세포에 사용하기 위해 이 시스템을 채택하였다: (i) C-말단 핵 국소화 서열(예를 들어, SV40 NLS)을 갖는 적절한 포유류 프로모터 및 적절한 폴리아데닐화 신호(예를 들어, TK pA 신호) 하에 코돈 최적화된 Cas9(예를 들어, 클래스 II, II형 Cas 효소)를 암호화하는 ORF; 및 (ii) 적절한 중합효소 III 프로모터(예를 들어, U6 프로모터) 하에 sgRNA(G로 시작하는 5' 서열에 이어서, 3' tracr-결합 서열에 결합된 20 nt의 상보성 표적화 핵산 서열, 링커, 및 tracrRNA 서열을 갖는)를 암호화하는 ORF (Mali 등의 문헌[Science. 2013 Feb 15; 339(6121): 823-826]을 참조하고, 동 문헌은 그 전체가 참조로서 본원에 통합됨).

[0094] MG 효소

[0095] 일 양태에서, 본 개시는 조작된 뉴클레아제 시스템을 제공한다. 조작된 뉴클레아제 시스템은 (a) 엔도뉴클레아제를 포함할 수 있다. 일부 경우, 엔도뉴클레아제는 RuvC 도메인 및 HNH 도메인을 포함한다. 엔도뉴클레아제는 미배양 미생물로부터 유래될 수 있다. 엔도뉴클레아제는 Cas 엔도뉴클레아제일 수 있다. 엔도뉴클레아제는 클래스 2 엔도뉴클레아제일 수 있다. 엔도뉴클레아제는 클래스 2, II형 Cas 엔도뉴클레아제일 수 있다. 조작된 뉴클레아제 시스템은 (b) 조작된 가이드 리보핵산 구조체를 포함할 수 있다. 조작된 가이드 리보핵산 구조체는 엔도뉴클레아제와 복합체를 형성하도록 구성될 수 있다. 일부 경우, 엔도뉴클레아제와 복합체를 형성하도록 구성된 조작된 가이드 리보핵산 구조체는 가이드 리보핵산 서열을 포함한다. 가이드 리보핵산 서열은 표적 데옥시리보핵산 서열과 혼성화되도록 구성될 수 있다. 일부 경우, 엔도뉴클레아제와 복합체를 형성하도록 구성된 조작된 가이드 리보핵산 구조체는 tracr 리보핵산 서열을 포함한다. tracr 리보핵산 서열은 엔도뉴클레아제에 결합하도록 구성될 수 있다. 일부 경우, 엔도뉴클레아제는 약 120 kDa 이하, 약 110 kDa 이하, 약 100 kDa 이하, 약 90 kDa 이하, 약 80 kDa 이하, 약 70 kDa 이하, 약 60 kDa 이하, 약 50 kDa 이하, 약 40 kDa 이하, 약 30 kDa 이하, 약 20 kDa 이하, 또는 약 10 kDa 이하의 분자량을 갖는다.

[0096] 일부 경우, 엔도뉴클레아제는 서열번호 1-198, 221-459, 463-612, 617-668, 674-675, 975-1002, 1260-1321 중 어느 하나와 적어도 50%, 적어도 55%, 적어도 50%, 적어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99%의 서열 동일성을 갖는 서열, 또는 이의 변이체를 포함한다.

[0097] 일 양태에서, 본 개시는 조작된 뉴클레아제 시스템을 제공한다. 조작된 뉴클레아제 시스템은 (a) 엔도뉴클레아제를 포함할 수 있다. 엔도뉴클레아제는 RuvC-1 도메인 또는 RuvC 도메인을 포함할 수 있다. 엔도뉴클레아제는 HNH 도메인을 포함할 수 있다. 엔도뉴클레아제는 RuvC-1 도메인 및 HNH 도메인을 포함할 수 있다. 엔도뉴클레아

제는 Cas 엔도뉴클레아제일 수 있다. 엔도뉴클레아제는 클래스 2 엔도뉴클레아제일 수 있다. 엔도뉴클레아제는 클래스 2, II형 Cas 엔도뉴클레아제일 수 있다. 조작된 뉴클레아제 시스템은 (b) 조작된 가이드 리보핵산을 포함할 수 있다. 조작된 가이드 리보핵산 구조체는 엔도뉴클레아제와 복합체를 형성하도록 구성될 수 있다. 엔도뉴클레아제와 복합체를 형성하도록 구성된 가이드 리보핵산 구조체는 가이드 리보핵산 서열을 포함할 수 있다. 가이드 리보핵산 서열은 표적 데옥시리보핵산 서열과 혼성화되도록 구성될 수 있다. 엔도뉴클레아제와 복합체를 형성하도록 구성된 조작된 가이드 리보핵산 구조체는 tracr 리보핵산 서열을 포함할 수 있다. tracr 리보핵산 서열은 엔도뉴클레아제에 결합하도록 구성될 수 있다. 엔도뉴클레아제는 서열번호 1-198, 221-459, 463-612, 617-668, 674-675, 975-1002, 1260-1321 중 어느 하나와 적어도 50%, 적어도 55%, 적어도 50%, 적어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99%의 서열 동일성을 갖는 서열을 포함할 수 있다. 엔도뉴클레아제는 고세균 엔도뉴클레아제일 수 있다. 엔도뉴클레아제는 클래스 2, II형 Cas 엔도뉴클레아제일 수 있다. 엔도뉴클레아제는 RRxRR 모티프 또는 PF14239 상동성을 갖는 도메인을 포함하는 아르기닌 풍부 영역을 포함할 수 있다. 아르기닌 풍부 영역 또는 PF14239 상동성을 갖는 도메인은 서열번호 1-198, 221-459, 463-612, 617-668, 674-675, 975-1002, 1260-1321 중 어느 하나의 아르기닌 풍부 영역 또는 PF14239 상동성을 갖는 도메인과 적어도 50%, 적어도 55%, 적어도 50%, 적어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99%의 서열 동일성을 갖는 서열, 또는 이의 변이체를 포함할 수 있다. 아르기닌 풍부 도메인 또는 PF14239 상동성을 갖는 도메인의 도메인 경계는 MG34-1 또는 MG34-9에 대한 최적 정렬에 의해 식별될 수 있다. 엔도뉴클레아제는 REC 도메인을 포함할 수 있다. REC 도메인은 서열번호 1-198, 221-459, 463-612, 617-668, 674-675, 975-1002, 1260-1321 중 어느 하나의 REC 도메인과 적어도 50%, 적어도 55%, 적어도 50%, 적어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99%의 서열 동일성을 갖는 서열, 또는 이의 변이체를 포함할 수 있다. REC 도메인의 도메인 경계는 MG34-1 또는 MG34-9에 대한 최적 정렬에 의해 식별될 수 있다. 엔도뉴클레아제는 BH(가교 나선(Bridge Helix)) 도메인을 포함할 수 있다. BH 도메인은 서열번호 1-198, 221-459, 463-612, 617-668, 674-675, 975-1002, 1260-1321 중 어느 하나의 BH 도메인과 적어도 50%, 적어도 55%, 적어도 50%, 적어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99%의 서열 동일성을 갖는 서열, 또는 이의 변이체를 포함할 수 있다. BH 도메인의 도메인 경계는 MG34-1 또는 MG34-9에 대한 최적 정렬에 의해 식별될 수 있다.

[0098] 엔도뉴클레아제는 WED(웨이) 도메인을 포함할 수 있다. WED 도메인은 서열번호 1-198, 221-459, 463-612, 617-668, 674-675, 975-1002, 1260-1321 중 어느 하나의 WED 도메인과 적어도 50%, 적어도 55%, 적어도 50%, 적어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99%의 서열 동일성을 갖는 서열, 또는 이의 변이체를 포함할 수 있다. WED 도메인의 도메인 경계는 MG34-1 또는 MG34-9에 대한 최적 정렬에 의해 식별될 수 있다. 엔도뉴클레아제는 PI(PAM-상호작용) 도메인을 포함할 수 있다. PI 도메인은 서열번호 1-198, 221-459, 463-612, 617-668, 674-675, 975-1002, 1260-1321 중 어느 하나의 PI 도메인과 적어도 50%, 적어도 55%, 적어도 50%, 적어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99%의 서열 동일성을 갖는 서열, 또는 이의 변이체를 포함할 수 있다. PI 도메인의 도메인 경계는 MG34-1 또는 MG34-9에 대한 최적 정렬에 의해 식별될 수 있다.

[0099] 일부 경우, 엔도뉴클레아제는 미배양 미생물로부터 유래된다. 일부 경우, tracr 리보핵산 서열은 서열번호 199-200, 460-461, 또는 669-673 중 어느 하나로부터의 적어도 50개, 적어도 60개, 적어도 70개, 적어도 80개의 연속 뉴클레오타이드와 50%, 적어도 55%, 적어도 50%, 적어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도

75%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99%의 서열 동일성을 갖는 서열, 또는 서열번호 201-203, 613-616, 677-686, 1003-1022, 또는 1231-1259 중 어느 하나의 비가변 뉴클레오타이드의 적어도 50개, 적어도 60개, 적어도 70개, 적어도 80개의 연속 뉴클레오타이드와 50%, 적어도 55%, 적어도 50%, 적어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99%의 서열 동일성을 갖는 서열을 포함한다.

[0100] 일부 경우, 가이드 핵산 구조체는 서열번호 201을 포함한다. 일부 경우, 가이드 핵산 구조체는 서열번호 202를 포함한다. 일부 경우, 가이드 핵산 구조체는 서열번호 203을 포함한다. 일부 경우, 가이드 핵산 구조체는 서열번호 201-203을 포함한다. 일부 경우, 가이드 핵산 구조체는 서열번호 613을 포함한다. 일부 경우, 가이드 핵산 구조체는 서열번호 614를 포함한다. 일부 경우, 가이드 핵산 구조체는 서열번호 615를 포함한다. 일부 경우, 가이드 핵산 구조체는 서열번호 616을 포함한다.

[0101] 일 양태에서, 본 개시는 조작된 뉴클레아제 시스템을 제공한다. 조작된 뉴클레아제 시스템은 (a) 조작된 가이드 리보핵산 구조체를 포함할 수 있다. 조작된 가이드 리보핵산 구조체는 가이드 리보핵산 서열을 포함할 수 있다. 가이드 리보핵산 서열은 표적 테옥시리보핵산 서열과 혼성화되도록 구성될 수 있다. 조작된 가이드 리보핵산 구조체는 tracr 리보핵산 서열을 포함할 수 있다. tracr 리보핵산 서열은 엔도뉴클레아제에 결합하도록 구성될 수 있다. 일부 경우, tracr 리보핵산 서열은 서열번호 199-200, 460-461, 또는 669-673 중 어느 하나로부터의 적어도 50개, 적어도 60개, 적어도 70개, 적어도 80개의 연속 뉴클레오타이드와 50%, 적어도 55%, 적어도 50%, 적어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99%의 서열 동일성을 갖는 서열, 또는 서열번호 201-203, 613-616, 677-686, 1003-1022, 또는 1231-1259 중 어느 하나의 비가변 뉴클레오타이드의 적어도 15개, 적어도 20개, 적어도 25개, 적어도 30개, 적어도 35개, 적어도 40개, 적어도 45개, 적어도 50개, 적어도 60개, 적어도 70개, 적어도 80개의 연속 뉴클레오타이드와 50%, 적어도 55%, 적어도 50%, 적어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99%의 서열 동일성을 갖는 서열을 포함한다.

[0102] 일부 경우, 조작된 뉴클레아제 시스템은 엔도뉴클레아제를 포함한다. 엔도뉴클레아제는 클래스 2 엔도뉴클레아제일 수 있다. 엔도뉴클레아제는 Cas 엔도뉴클레아제일 수 있다. 엔도뉴클레아제는 클래스 2, II형 Cas 엔도뉴클레아제일 수 있다.

[0103] 일부 경우, 엔도뉴클레아제는 특정 분자량 범위를 갖는다. 일부 구현예에서, 엔도뉴클레아제는 약 120 kDa 이하, 약 110 kDa 이하, 약 105 kDa 이하, 약 100 kDa 이하, 약 95 kDa 이하, 약 90 kDa 이하, 약 85 kDa 이하, 약 80 kDa 이하, 약 75 kDa 이하, 약 70 kDa 이하, 약 65 kDa 이하, 약 60 kDa 이하, 약 55 kDa 이하, 약 50 kDa 이하, 약 45 kDa 이하, 약 40 kDa 이하, 약 35 kDa 이하, 약 30 kDa 이하, 약 25 kDa 이하, 약 20 kDa 이하, 약 15 kDa 이하, 또는 약 10 kDa 이하의 분자량을 갖는다. 일부 경우, 조작된 가이드 리보핵산 구조체는 적어도 2개의 리보핵산 폴리뉴클레오타이드를 포함한다. 일부 경우, 엔도뉴클레아제는 특정 수의 잔기를 포함한다. 엔도뉴클레아제는 약 1,100개 이하의 잔기, 약 1,000개 이하의 잔기, 약 950개 이하의 잔기, 약 900개 이하의 잔기, 약 850개 이하의 잔기, 약 800개 이하의 잔기, 약 750개 이하의 잔기, 약 700개 이하의 잔기, 약 650개 이하의 잔기, 약 600개 이하의 잔기, 약 550개 이하의 잔기, 약 500개 이하의 잔기, 약 450개 이하의 잔기, 약 400개 이하의 잔기, 또는 약 350개 이하의 잔기를 포함할 수 있다. 엔도뉴클레아제는 약 700 내지 약 1,100개의 잔기를 포함할 수 있다. 엔도뉴클레아제는 약 400 내지 약 600개의 잔기를 포함할 수 있다. 일부 경우, 조작된 가이드 리보핵산 구조체는 단일 리보핵산 폴리뉴클레오타이드를 포함한다. 단일 리보핵산 폴리뉴클레오타이드는 가이드 리보핵산 서열 및 tracr 리보핵산 서열을 포함할 수 있다.

[0104] 일부 경우, 가이드 리보핵산 서열은 원핵, 박테리아, 고세균, 진핵, 진균, 식물, 포유류, 또는 인간 게놈 서열에 상보적이다. 일부 경우, 가이드 리보핵산 서열은 원핵 게놈 서열에 상보적이다. 일부 경우, 가이드 리보핵산 서열은 박테리아 게놈 서열에 상보적이다. 일부 경우, 가이드 리보핵산 서열은 고세균 게놈 서열에 상보적이다. 일부 경우, 가이드 리보핵산 서열은 진핵 게놈 서열에 상보적이다. 일부 경우, 가이드 리보핵산 서열은 식물 게놈 서열에 상보적이다. 일부 경우, 가이드 리보핵산 서열은 포유류 게놈 서열에 상보적이다. 일부 경우, 가이드 리보핵산 서열은 인간 게놈 서열에 상보적

이다.

[0105]

일부 경우, 가이드 리보핵산 표적화 서열 또는 스페이서는 10-30개 뉴클레오타이드 길이, 또는 12-28개 뉴클레오타이드 길이, 또는 15-24개 뉴클레오타이드 길이이다. 일부 경우, 엔도뉴클레아제는 엔도뉴클레아제의 N-말단 또는 C-말단에 근접한 하나 이상의 핵 국소화 서열(NLS)을 포함한다. 일부 경우, NLS는 서열번호 205-220으로부터 선택되는 서열을 포함한다.

[표 1] 본 개시에 따라 Cas 효과기와 함께 사용될 수 있는 NLS 서열의 예.

공급원	NLS 아미노산 서열	서열번호
SV40 NLS	PKKKRKV	205
뉴클레오플라스민 이분	KRPAATKKAGQAKKKK	206
c-myc	PAAKRVKLD	207
c-myc	RQRRNELKRSP	208
hnRNPA1 M9	NQSSNFGPMKGGNFGGRSSGPYGGGGQYFAKPRNQGGY	209
임포르틴-알파 IBB 도메인	RMRIZFKNKGKDTAELRRRRVEVSVELRKAKKDEQILKRRNV	210
Myoma T 단백질	VSRKRPRP	211
Myoma T 단백질	PPKKARED	212
p53	PQPKKKPL	213
마우스 c-abl IV	SALIKKKKKMAP	214
인플루엔자 바이러스 NS1	DRLRR	215
인플루엔자 바이러스 NS1	PKQKKRK	216
간염 바이러스 델타 항원	RKLKKKIKKL	217
마우스 Mx1 단백질	REKKKFLKRR	218
인간 폴리(ADP-리보오스) 중합효소	KRKGDEVVDGVDEVAKKKSKK	219
스테로이드 호르몬 수용체 글루코코르티코이드	RKCLQAGMNLEARKTKK	220

[0106]

[0107]

하나 이상의 보존적 아미노산 치환을 갖는 본원에 기술된 효소 중 어느 하나의 변이체가 본 개시에 포함된다. 이러한 보존적 치환은 폴리펩티드의 3차원 구조 또는 기능을 파괴하지 않고도 폴리펩티드의 아미노산 서열에서 이루어질 수 있다. 보존적 치환은 소수성, 극성, 및 R 사슬 길이가 서로 유사한 아미노산들을 치환함으로써 달성될 수 있다. 추가적으로 또는 대안적으로, 상이한 종의 상동성 단백질의 정렬된 서열을 비교함으로써, 종들 간에 돌연변이된 아미노산 잔기(예를 들어, 암호화된 단백질의 기본 기능을 변경시키지 않은 비보존적 잔기)를 위치시킴으로써 보존적 치환을 식별할 수 있다. 이러한 보존적으로 치환된 변이체는 본원에 기술된 엔도뉴클레아제 단백질 서열 중 어느 하나와 적어도 약 20%, 적어도 약 25%, 적어도 약 30%, 적어도 약 35%, 적어도 약 40%, 적어도 약 45%, 적어도 약 50%, 적어도 약 55%, 적어도 약 60%, 적어도 약 65%, 적어도 약 70%, 적어도 약 75%, 적어도 약 80%, 적어도 약 85%, 적어도 약 86%, 적어도 약 87%, 적어도 약 88%, 적어도 약 89%, 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98%, 또는 적어도 약 99% 동일성을 갖는 변이체를 포함할 수 있다. 일부 구현예에서, 이러한 보존적으로 치환된 변이체는 기능적 변이체이다. 이러한 기능적 변이체는 엔도뉴클레아제의 하나 이상의 중요한 활성 부위 잔기 또는 가이드 RNA 결합 잔기의 활성이 파괴되지 않도록 치환된 서열을 포함할 수 있다. 일부 구현예에서, 본원에 기술된 단백질 중 어느 하나의 기능적 변이체에는 도 4에 언급된 보존된 또는 기능적 잔기 중 적어도 하나의 치환이 결여되어 있다. 일부 구현예에서, 본원에 기술된 단백질 중 어느 하나의 기능적 변이체에는 도 4에 언급된 보존된 또는 기능적 잔기 모두의 치환이 결여되어 있다. 또한, 본원에 기술된 뉴클레아제 중 어느 하나의 변경된 활성 변이체가 본원에 제공된다. 이러한 변경된 활성 변이체는 본원에서 (예를 들어, 도 4에서) 식별되거나 RuvC 도메인에 대해 일반적으로 기술된 하나 이상의 촉매 잔기에서 불활성화 돌연변이를 포함할 수 있다. 이러한 변경된 활성 변이체는 RuvCI, RuvCII, 또는 RuvCIII 도메인의 촉매 잔기에서 변화-스위치 돌연변이를 포함할 수 있다.

[0108] 기능적으로 유사한 아미노산을 제공하는 보존적 치환 표는 다양한 참조 문헌을 통해 이용할 수 있다(예를 들어, Creighton의 문헌[Proteins: Structures and Molecular Properties (W H Freeman & Co.; 제2판 (1993년 12월)) 참조]. 다음의 8개의 기는 서로에 대해 보존적 치환인 아미노산을 각각 함유한다:

[0109] 1) 알라닌 (A), 글리신 (G);

[0110] 2) 아스파르트산 (D), 글루탐산 (E);

[0111] 3) 아스파라긴 (N), 글루타민 (Q);

[0112] 4) 아르기닌 (R), 리신 (K);

[0113] 5) 이소류신 (I), 류신 (L), 메티오닌 (M), 발린 (V);

[0114] 6) 페닐알라닌 (F), 티로신 (Y), 트립토판 (W);

[0115] 7) 세린 (S), 트레오닌 (T); 및

[0116] 8) 시스테인 (C), 메티오닌 (M)

[0117] 특정 도메인과 서열 동일성을 갖는 본원에 기술된 엔도뉴클레아제 중 어느 하나의 변이체가 본 개시에 포함된다. 도메인은 아르기닌 풍부 도메인(예를 들어, PF14239 상동성을 갖는 도메인), REC(인식) 도메인, BH(가교 나선) 도메인, WED(웨이) 도메인, PI(PAM-상호작용) 도메인, PF14239 상동성 도메인, 또는 본원에 기술된 임의의 다른 도메인일 수 있다. 일부 구현예에서, 이들 도메인 중 하나 이상을 포함하는 잔기는 (예를 들어, 아래의 단백질 및 관심 단백질 중 하나가 최적으로 정렬될 때) 아래의 단백질 중 하나에 대한 정렬에 의해 단백질에서 식별되며, 여기에서 잔기 경계, 예를 들어 도메인이 기술된다.

[표 2] 본원에 기술된 엔도뉴클레아제에 대한 예시적인 도메인 경계

	RuvC -I	BH	REC	PF14239 상동성을 가진 도메인	RuvC -II	HNH	RuvC -III	WED 및 PI
MG34- 1 효과기	1-41	42-76	77- 281	4-65; 123-339	282- 323	324- 459	460- 551	552- 747
MG34- 9 효과기	1-41	42-76	77- 280	4-65; 123-338	281- 322	323- 490	491- 582	583- 778

[0118]

[0119] 일부 경우, 조작된 뉴클레아제 시스템은 단일-가닥 DNA 복구 템플릿을 추가로 포함한다. 일부 경우, 조작된 뉴클레아제 시스템은 이중-가닥 DNA 복구 템플릿을 추가로 포함한다. 일부 경우, 단일-가닥 또는 이중-가닥 DNA 복구 템플릿은, 5'에서 3'으로, 표적 데옥시리보핵산 서열에 대해 5'에 있는 적어도 20개 뉴클레오타이드의 서열을 포함하는 제1 상동 아암을 포함한다. 일부 경우, 단일-가닥 또는 이중-가닥 DNA 복구 템플릿은, 5'에서 3'으로, 적어도 10개 뉴클레오타이드의 합성 DNA 서열을 포함한다. 일부 경우, 단일-가닥 또는 이중-가닥 DNA 복구 템플릿은, 5'에서 3'으로, 표적 서열에 대해 3'에 있는 적어도 20개 뉴클레오타이드의 서열을 포함하는 제2 상동 아암을 포함한다. 일부 경우, 단일 또는 이중 가닥 DNA 복구 템플릿은, 5'에서 3'으로, 표적 데옥시리보핵산 서열에 대해 5'에 있는 적어도 20개 뉴클레오타이드의 서열을 포함하는 제1 상동 아암, 적어도 10개 뉴클레오타이드의 합성 DNA 서열, 또는 표적 서열에 대해 3'에 있는 적어도 20개 뉴클레오타이드의 서열을 포함하는 제2 상동 아암을 포함한다.

[0120] 일부 경우, 제1 상동 아암은 적어도 10개, 적어도 20개, 적어도 30개, 적어도 40개, 적어도 50개, 적어도 60개, 적어도 70개, 적어도 80개, 적어도 90개, 적어도 100개, 적어도 110개, 적어도 120개, 적어도 130개, 적어도 140개, 적어도 150개, 적어도 175개, 적어도 200개, 적어도 250개, 적어도 300개, 적어도 400개, 적어도 500개, 적어도 750개, 또는 적어도 1000개 뉴클레오타이드의 서열을 포함한다. 일부 경우, 조작된 뉴클레아제 시스템은 Mg^{2+} 의 공급원을 추가로 포함한다. 일부 경우, 엔도뉴클레아제 및 tracr 리보핵산 서열은 구별되는 박테리아 종으로부터 유래된다. 일부 경우, 엔도뉴클레아제 및 tract 리보핵산 서열은 동일한 문 내의 구별되는 박테리아 종으로부터 유래된다.

[0121] 일부 경우, 엔도뉴클레아제는 서열번호 1-24 또는 462-488 중 어느 하나와 적어도 50%, 적어도 55%, 적어도 50%, 적어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99%의 서열 동일성을 갖는 서열을 포함한다. 일부 경우, 가이드 RNA 구조체는 헤어핀을 포함하도록 예측된

RNA 서열을 포함한다. 일부 경우, 헤어핀은 줄기 및 루프를 포함한다. 일부 경우, 줄기는 적어도 12 쌍, 적어도 14 쌍, 적어도 16 쌍 또는 적어도 18 쌍의 리보뉴클레오타이드를 포함한다.

[0122] 일부 경우, 가이드 RNA 구조체는 제2 줄기 및 제2 루프를 추가로 포함한다. 일부 경우, 제2 줄기는 적어도 5 쌍, 적어도 6 쌍, 적어도 7 쌍, 적어도 8 쌍, 적어도 9 쌍 또는 적어도 10 쌍의 리보뉴클레오타이드를 포함한다. 일부 경우, 가이드 RNA 구조체는 RNA 구조체를 추가로 포함하며, 이러한 RNA 구조체는 적어도 2개의 헤어핀을 포함한다. 일부 구현예에서, 엔도뉴클레아제는 서열번호 1과 적어도 50%, 적어도 55%, 적어도 50%, 적어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99%의 서열 동일성을 갖는 서열을 포함하고, 가이드 RNA 구조체는 적어도 4개의 헤어핀을 포함하도록 예측된 RNA 서열을 포함한다. 일부 경우, 이들 4개의 헤어핀 각각은 줄기 및 루프를 포함한다.

[0123] 일부 경우, 조작된 뉴클레아제 시스템은 서열번호 1과 적어도 50%, 적어도 55%, 적어도 50%, 적어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99% 동일한 서열을 포함한다. 일부 구현예에서, 조작된 뉴클레아제 시스템은 서열번호 199, 또는 서열번호 201의 비가변 뉴클레오타이드 중 어느 하나와 적어도 50%, 적어도 55%, 적어도 50%, 적어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99% 동일한 서열을 포함하는 가이드 RNA 구조체를 포함한다.

[0124] 일부 경우, 조작된 뉴클레아제 시스템은 서열번호 1-24 또는 462-488 중 어느 하나와 적어도 50%, 적어도 55%, 적어도 50%, 적어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99% 동일한 서열을 포함한다. 일부 경우, 조작된 뉴클레아제 시스템은 서열번호 199-200, 460-461, 또는 669-673 중 어느 하나, 또는 서열번호 201-203, 613-616, 677-686, 1003-1022, 또는 1231-1259 중 어느 하나의 비가변 뉴클레오타이드와 적어도 50%, 적어도 55%, 적어도 50%, 적어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99% 동일한 서열을 포함한다.

[0125] 일부 경우, 서열 동일성은 BLASTP, CLUSTALW, MUSCLE, MAFFT, 또는 Smith-Waterman 상동성 검색 알고리즘 파라미터를 사용하는 CLUSTALW에 의해 결정된다. 일부 경우, 서열 동일성은, 단어 길이(W) 3, 기대치(E) 10의 파라미터, 및 BLOSUM62 스코어링 매트릭스(존재 11, 연장 1의 갭 비용 설정)를 사용하고, 조건부 조성 스코어 매트릭스 조정을 사용하는, 전술한 BLASTP 상동성 검색 알고리즘에 의해 결정된다.

[0126] 일부 경우, 엔도뉴클레아제는 Cas9 엔도뉴클레아제, Cas14 엔도뉴클레아제, Cas12a 엔도뉴클레아제, Cas12b 엔도뉴클레아제, Cas 12c 엔도뉴클레아제, Cas12d 엔도뉴클레아제, Cas12e 엔도뉴클레아제, Cas13a 엔도뉴클레아제, Cas13b 엔도뉴클레아제, Cas13c 엔도뉴클레아제, 또는 Cas 13d 엔도뉴클레아제가 아니다. 일부 경우, 엔도뉴클레아제는 Cas9 엔도뉴클레아제와 80% 미만의 동일성, 75% 미만의 동일성, 70% 미만의 동일성, 65% 미만의 동일성, 60% 미만의 동일성, 55% 미만의 동일성, 또는 50% 미만의 동일성을 갖는다.

[0127] 일 양태에서, 본 개시는 (a) DNA-표적화 분절을 포함하는 조작된 가이드 RNA를 제공한다. 일부 경우, DNA-표적화 분절은 표적 DNA 서열의 표적 서열에 상보적인 뉴클레오타이드 서열을 포함한다. 일부 경우, 조작된 단일 가이드 리보핵산 폴리뉴클레오타이드는 단백질-결합 분절을 포함한다. 단백질-결합 분절은 이중-가닥 RNA(dsRNA) 이중체를 형성하도록 혼성화되는 뉴클레오타이드의 2개의 상보적 신장부를 포함한다. 일부 경우, 뉴클레오타이드의 2개의 상보적 신장부는 개재 뉴클레오타이드와 서로 공유적으로 연결된다. 일부 경우, 조작된 가이드 리보핵산 폴리뉴클레오타이드는 서열번호 1-198, 221-459, 463-612, 617-668, 674-675, 975-1002, 1260-1321 중 어느 하나와 적어도 50%, 적어도 55%, 적어도 50%, 적어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99%의 서열 동일성을 갖는 서열, 또는 이의 변이체를 포함하는 엔도뉴클레아제와 복합체를 형성하도록 구성된다.

[0128] 일부 경우, DNA-표적화 분절은 뉴클레오타이드의 2개의 상보적 신장부 둘 모두의 5'에 위치한다. 일부 경우, 단백질 결합 단편은 서열번호 199-200, 460-461, 669-673 중 어느 하나, 또는 서열번호 201-203, 613-616, 677-686, 1003-1022, 또는 1231-1259 중 어느 하나의 비가변 뉴클레오타이드와 적어도 50%, 적어도 55%, 적어도 50%,

적어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99% 동일한 서열을 포함한다. 일부 경우, 테옥시리보핵산 폴리뉴클레오티드는 본원에 기술된 조작된 가이드 리보핵산 폴리뉴클레오티드를 암호화한다.

[0129] 일 양태에서, 본 개시는 조작된 핵산 서열을 포함하는 핵산을 제공한다. 일부 경우, 조작된 핵산 서열은 유기체에서의 발현에 최적화된다. 일부 경우, 핵산은 엔도뉴클레아제를 암호화한다. 엔도뉴클레아제는 Cas 엔도뉴클레아제일 수 있다. 엔도뉴클레아제는 클래스 2 엔도뉴클레아제일 수 있다. 엔도뉴클레아제는 클래스 2, II형 Cas 엔도뉴클레아제일 수 있다. 일부 경우, 엔도뉴클레아제는 RuvC 도메인 및 HNH 도메인을 포함한다. 일부 경우, 엔도뉴클레아제는 미배양 미생물로부터 유래된다. 일부 경우, 엔도뉴클레아제는 특정 분자량 범위를 갖는다. 일부 구현예에서, 엔도뉴클레아제는 약 120 kDa 이하, 약 110 kDa 이하, 약 105 kDa 이하, 약 100 kDa 이하, 약 95 kDa 이하, 약 90 kDa 이하, 약 85 kDa 이하, 약 80 kDa 이하, 약 75 kDa 이하, 약 70 kDa 이하, 약 65 kDa 이하, 약 60 kDa 이하, 약 55 kDa 이하, 약 50 kDa 이하, 약 45 kDa 이하, 약 40 kDa 이하, 약 35 kDa 이하, 약 30 kDa 이하, 약 25 kDa 이하, 약 20 kDa 이하, 약 15 kDa 이하, 또는 약 10 kDa 이하의 분자량을 갖는다. 일부 경우, 조작된 가이드 리보핵산 구조체는 적어도 2개의 리보핵산 폴리뉴클레오티드를 포함한다. 일부 경우, 엔도뉴클레아제는 특정 수의 잔기를 포함한다. 엔도뉴클레아제는 약 1,100개 이하의 잔기, 약 1,000개 이하의 잔기, 약 950개 이하의 잔기, 약 900개 이하의 잔기, 약 850개 이하의 잔기, 약 800개 이하의 잔기, 약 750개 이하의 잔기, 약 700개 이하의 잔기, 약 650개 이하의 잔기, 약 600개 이하의 잔기, 약 550개 이하의 잔기, 약 500개 이하의 잔기, 약 450개 이하의 잔기, 약 400개 이하의 잔기, 또는 약 350개 이하의 잔기를 포함할 수 있다. 엔도뉴클레아제는 약 700 내지 약 1,100개의 잔기를 포함할 수 있다. 엔도뉴클레아제는 약 400 내지 약 600개의 잔기를 포함할 수 있다. 일부 경우, 엔도뉴클레아제는 서열번호 1-198, 221-459, 463-612, 617-668, 674-675, 975-1002, 1260-1321, 또는 이에 대해 적어도 50%, 적어도 55%, 적어도 50%, 적어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99%의 서열 동일성을 갖는 이의 변이체를 포함한다. 일부 경우, 엔도뉴클레아제는 전술한 엔도뉴클레아제의 N-말단 또는 C-말단에 근접한 하나 이상의 핵 국소화 서열(NLS)을 암호화하는 서열을 추가로 포함한다. 일부 경우, NLS는 서열번호 205-220으로부터 선택되는 서열을 포함한다.

[0130] 일부 경우, 유기체는 원핵생물, 박테리아, 진핵생물, 진균류, 식물, 포유류, 설치류, 또는 인간이다. 일부 경우, 유기체는 원핵생물이다. 일부 경우, 유기체는 박테리아이다. 일부 경우, 유기체는 진핵생물이다. 일부 경우, 유기체는 진균류이다. 일부 경우, 유기체는 식물이다. 일부 경우, 유기체는 포유류이다. 일부 경우, 유기체는 설치류이다. 일부 경우, 유기체는 인간이다. 유기체는 원핵생물 또는 박테리아인 경우, 유기체는 엔도뉴클레아제가 유래되는 유기체와 상이한 유기체일 수 있다. 일부 경우, 유기체는 미배양 유기체가 아니다.

[0131] 일 양태에서, 본 개시는 핵산 서열을 포함하는 벡터를 제공한다. 일부 경우, 핵산 서열은 엔도뉴클레아제를 암호화한다. 일부 경우, 엔도뉴클레아제는 Cas 엔도뉴클레아제이다. 일부 경우, 엔도뉴클레아제는 클래스 2 엔도뉴클레아제이다. 일부 경우, 엔도뉴클레아제는 클래스 2, II형 Cas 엔도뉴클레아제이다. 엔도뉴클레아제는 RuvC-I 도메인 및 HNH 도메인을 포함할 수 있다. 일부 경우, 엔도뉴클레아제는 미배양 미생물로부터 유래된다. 일부 경우, 엔도뉴클레아제는 특정 분자량 범위를 갖는다. 일부 구현예에서, 엔도뉴클레아제는 약 120 kDa 이하, 약 110 kDa 이하, 약 105 kDa 이하, 약 100 kDa 이하, 약 95 kDa 이하, 약 90 kDa 이하, 약 85 kDa 이하, 약 80 kDa 이하, 약 75 kDa 이하, 약 70 kDa 이하, 약 65 kDa 이하, 약 60 kDa 이하, 약 55 kDa 이하, 약 50 kDa 이하, 약 45 kDa 이하, 약 40 kDa 이하, 약 35 kDa 이하, 약 30 kDa 이하, 약 25 kDa 이하, 약 20 kDa 이하, 약 15 kDa 이하, 또는 약 10 kDa 이하의 분자량을 갖는다. 일부 경우, 조작된 가이드 리보핵산 구조체는 적어도 2개의 리보핵산 폴리뉴클레오티드를 포함한다. 일부 경우, 엔도뉴클레아제는 특정 수의 잔기를 포함한다. 엔도뉴클레아제는 약 1,100개 이하의 잔기, 약 1,000개 이하의 잔기, 약 950개 이하의 잔기, 약 900개 이하의 잔기, 약 850개 이하의 잔기, 약 800개 이하의 잔기, 약 750개 이하의 잔기, 약 700개 이하의 잔기, 약 650개 이하의 잔기, 약 600개 이하의 잔기, 약 550개 이하의 잔기, 약 500개 이하의 잔기, 약 450개 이하의 잔기, 약 400개 이하의 잔기, 또는 약 350개 이하의 잔기를 포함할 수 있다. 엔도뉴클레아제는 약 700 내지 약 1,100개의 잔기를 포함할 수 있다. 엔도뉴클레아제는 약 400 내지 약 600개의 잔기를 포함할 수 있다.

[0132] 일부 양태에서, 본 개시는 전술한 표적 유전자좌 5'에 근위에서 프로토스페이서 인접 모티프(PAM)까지 이중 가닥 절단을 유도하도록 구성된 본원에 기술된 엔도뉴클레아제를 제공한다. 엔도뉴클레아제는 PAM으로부터 6-8개 뉴클레오티드 또는 PAM으로부터 7개 뉴클레오티드의 이중 가닥 절단을 유도할 수 있다. 일부 양태에서, 본 개시

는 전술한 표적 유전자좌 5'에 근위에서 프로토스페이서 인접 모티프(PAM)까지 단일 가닥 절단을 유도하도록 구성된 본원에 기술된 엔도뉴클레아제를 제공한다. 엔도뉴클레아제는 PAM으로부터 6-8개 뉴클레오타이드 또는 PAM으로부터 7개 뉴클레오타이드의 단일 가닥 절단을 유도할 수 있다. 일부 경우, 단일 가닥 절단을 유도하도록 구성된 엔도뉴클레아제는 본원에 기술된 엔도뉴클레아제의 하나 이상의 촉매 잔기에서 비활성화 돌연변이를 포함한다.

[0133] 일부 양태에서, 본 개시는 엔도뉴클레아제 시스템에 의해 표적화된 표적 유전자좌 내에서 또는 이에 근접하여 뉴클레오타이드 염기의 화학적 변형을 야기하도록 구성된 본원에 기술된 엔도뉴클레아제 시스템을 제공한다. 이러한 경우, 뉴클레오타이드 염기의 화학적 변형은 대체적으로 뉴클레오타이드의 당 또는 인산염 부분의 변형보다는 염기 페어링에 관여된 화학적 모이어티의 변형을 지칭한다. 화학적 변형은 아데노신 또는 시토신 뉴클레오타이드의 탈아미노화를 포함할 수 있다. 일부 경우, 화학적 변형을 야기하도록 구성된 엔도뉴클레아제 시스템은 전술한 엔도뉴클레아제에 대해 프레임 내에 커플링 또는 융합된 염기 편집기를 갖는 엔도뉴클레아제를 포함한다. 염기 편집기가 융합 또는 커플링되는 엔도뉴클레아제는 엔도뉴클레아제의 적어도 하나의 촉매 잔기(예를 들어, RuvC 도메인)에서 비활성화 돌연변이를 포함할 수 있다. 염기 편집기는 N-말단 또는 C-말단에서 전술한 엔도뉴클레아제에 융합되거나, 화학적 접합을 통해 연결될 수 있다. 염기 편집기는 다음을 포함하나 이에 한정되지 않는 임의의 아데노신 또는 시토신 탈아미노효소를 포함할 수 있다: 아데노신 탈아미노효소 RNA 특이적 1(ADAR1), 아데노신 탈아미노효소 RNA 특이적 2(ADAR2), 아포지질단백질 B mRNA 편집 효소 촉매 서브유닛 1(APOBEC1), 아포지질단백질 B mRNA 편집 효소 촉매 서브유닛 2(APOBEC2), 아포지질단백질 B mRNA 편집 효소 촉매 서브유닛 3A(APOBEC3A), 아포지질단백질 B mRNA 편집 효소 촉매 서브유닛 3B(APOBEC3B), 아포지질단백질 B mRNA 편집 효소 촉매 서브유닛 3C(APOBEC3C), 아포지질단백질 B mRNA 편집 효소 촉매 서브유닛 3D(APOBEC3D), 아포지질단백질 B mRNA 편집 효소 촉매 서브유닛 3F(APOBEC3F), 아포지질단백질 B mRNA 편집 효소 촉매 서브유닛 3G(APOBEC3G), 아포지질단백질 B mRNA 편집 효소 촉매 서브유닛 3H(APOBEC3H), 또는 아포지질단백질 B mRNA 편집 효소 촉매 서브유닛 4(APOBEC4), 또는 이의 기능적 단편. 염기 편집기는 효모, 진핵생물, 포유류, 또는 인간 염기 편집기를 포함할 수 있다.

[0134] 일부 양태에서, 본 개시는 엔도뉴클레아제 시스템에 의해 표적화된 표적 유전자좌 내에서 또는 이에 근접하여 히스톤의 화학적 변형을 야기하도록 구성된 본원에 기술된 엔도뉴클레아제 시스템을 제공한다. 일부 경우, 히스톤의 화학적 변형을 야기하도록 구성된 엔도뉴클레아제 시스템은 전술한 엔도뉴클레아제에 대해 프레임 내에 커플링 또는 융합된 히스톤 편집기를 갖는 엔도뉴클레아제를 포함한다. 히스톤 편집기는 N-말단 또는 C-말단에서 엔도뉴클레아제에 커플링 또는 융합될 수 있다. 일부 구현예에서, 화학적 변형은 메틸화, 아세틸화, 탈메틸화, 또는 탈아세틸화를 포함할 수 있다. 히스톤 편집기가 융합 또는 커플링되는 엔도뉴클레아제는 엔도뉴클레아제의 적어도 하나의 촉매 잔기(예를 들어, RuvC 도메인)에서 비활성화 돌연변이를 포함할 수 있다. 히스톤 에디터는 히스톤 메틸트랜스퍼라아제(예를 들어, ASH1L, DOT1L, EHMT1, EHMT2, EZH1, EZH2, MLL, MLL2, MLL3, MLL4, MLL5, NSD1, PRDM2, SET, SETBP1, SETD1A, SETD1B, SETD2, SETD3, SETD4, SETD5, SETD6, SETD7, SETD8, SETD9, SETDB1, SETDB2, SETMAR, SMYD1, SMYD2, SMYD3, SMYD4, SMYD5, SUV39H1, SUV39H2, SUV420H1, 또는 SUV420H2), 히스톤 탈메틸효소(예를 들어, KDM1, KDM2, KDM3, KDM4, KDM5, 또는 KDM6 계열), 히스톤 아세틸트랜스퍼라아제(예를 들어, GNAT 또는 HAT 계열 아세틸트랜스퍼라아제), 또는 히스톤 탈아세틸화효소(예를 들어, HDAC1, HDAC2, HDAC 3, HDAC4, HDAC5, HDAC6, HDAC7, HDAC8, HDAC9, HDAC10, HDAC11, SIRT1, SIRT2, SIRT3, SIRT4, SIRT5, SIRT6, 또는 SIRT7)를 포함할 수 있다. 히스톤 편집기는 효모, 진핵생물, 포유류, 또는 인간 히스톤 편집기를 포함할 수 있다.

[0135] 일 양태에서, 본 개시는 본원에 기술된 핵산을 포함하는 벡터를 제공한다. 일부 경우, 벡터는 조작된 가이드 리보핵산 구조체를 암호화하는 핵산을 추가로 포함한다. 조작된 가이드 리보핵산 구조체는 엔도뉴클레아제와 복합체를 형성하도록 구성될 수 있다. 일부 경우, 조작된 가이드 리보핵산 구조체는 가이드 리보핵산 서열을 포함한다. 일부 경우, 가이드 리보핵산 서열은 표적 데옥시리보핵산 서열과 혼성화되도록 구성된다. 일부 경우, 조작된 가이드 리보핵산 구조체는 tracr 리보핵산 서열을 포함한다. 일부 경우, tracr 리보핵산 서열은 엔도뉴클레아제에 결합하도록 구성된다. 일부 경우, 벡터는 플라스미드, 미니서클, CELiD, 아데노-연관 바이러스(AAV) 유래 비리온, 렌티바이러스이다.

[0136] 일 양태에서, 본 개시는 본원에 기술된 벡터 중 어느 하나를 포함하는 세포를 제공한다.

[0137] 일 양태에서, 본 개시는 엔도뉴클레아제를 제조하는 방법을 제공한다. 방법은 본원에 기술된 세포 중 어느 하나를 배양하는 단계를 포함할 수 있다.

[0138] 일 양태에서, 본 개시는 이중-가닥 데옥시리보핵산 폴리뉴클레오타이드를 결합, 절단, 마킹, 또는 변형시키는 방

법을 제공한다. 방법은 이중-가닥 데옥시리보핵산 폴리뉴클레오티드를 엔도뉴클레아제와 접촉시키는 단계를 포함할 수 있다. 일부 경우, 엔도뉴클레아제는 Cas 엔도뉴클레아제이다. 일부 경우, 엔도뉴클레아제는 클래스 2 엔도뉴클레아제이다. 일부 경우, 엔도뉴클레아제는 클래스 2, II형 Cas 엔도뉴클레아제이다. 엔도뉴클레아제는 조작된 가이드 리보핵산 구조체와 복합체를 이룰 수 있다. 일부 경우, 조작된 가이드 리보핵산 구조체는 엔도뉴클레아제 및 이중-가닥 데옥시리보핵산 폴리뉴클레오티드에 결합하도록 구성된다. 일부 경우, 이중-가닥 데옥시리보핵산 폴리뉴클레오티드는 프로토스페이서 인접 모티프(PAM)를 포함한다. 일부 경우, 엔도뉴클레아제는 약 120 kDa 이하, 약 110 kDa 이하, 약 100 kDa 이하, 약 90 kDa 이하, 약 80 kDa 이하, 약 70 kDa 이하, 약 60 kDa 이하, 약 50 kDa 이하, 약 40 kDa 이하, 약 30 kDa 이하, 약 20 kDa 이하, 또는 약 10 kDa 이하의 분자량을 갖는다. 일부 경우, 엔도뉴클레아제는 서열번호 1-198, 221-459, 463-612, 617-668, 674-675, 975-1002, 1260-1321 중 어느 하나와 적어도 50%, 적어도 55%, 적어도 50%, 적어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99%의 서열 동일성을 갖는 변이체, 또는 이의 변이체를 포함한다.

[0139] 일 양태에서, 본 개시는 이중-가닥 데옥시리보핵산 폴리뉴클레오티드를 결합, 절단, 마킹, 또는 변형시키는 방법을 제공한다. 방법은 이중-가닥 데옥시리보핵산 폴리뉴클레오티드를 엔도뉴클레아제와 접촉시키는 단계를 포함할 수 있다. 일부 경우, 엔도뉴클레아제는 Cas 엔도뉴클레아제이다. 일부 경우, 엔도뉴클레아제는 클래스 2 엔도뉴클레아제이다. 일부 경우, 엔도뉴클레아제는 클래스 2, II형 Cas 엔도뉴클레아제이다. 엔도뉴클레아제는 조작된 가이드 리보핵산 구조체와 복합체를 이룰 수 있다. 일부 경우, 조작된 가이드 리보핵산 구조체는 엔도뉴클레아제 및 이중-가닥 데옥시리보핵산 폴리뉴클레오티드에 결합하도록 구성될 수 있다. 일부 경우, 이중-가닥 데옥시리보핵산 폴리뉴클레오티드는 프로토스페이서 인접 모티프(PAM)를 포함한다. 일부 경우, PAM은 NGG이다. 일부 경우, 엔도뉴클레아제는 서열번호 1-198, 221-459, 463-612, 617-668, 674-675, 975-1002, 1260-1321 중 어느 하나와 적어도 50%, 적어도 55%, 적어도 50%, 적어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99%의 서열 동일성을 갖는 변이체, 또는 이의 변이체를 포함한다.

[0140] 일부 경우, 엔도뉴클레아제는 Cas9 엔도뉴클레아제, Cas14 엔도뉴클레아제, Cas12a 엔도뉴클레아제, Cas12b 엔도뉴클레아제, Cas 12c 엔도뉴클레아제, Cas12d 엔도뉴클레아제, Cas12e 엔도뉴클레아제, Cas13a 엔도뉴클레아제, Cas13b 엔도뉴클레아제, Cas13c 엔도뉴클레아제, 또는 Cas 13d 엔도뉴클레아제가 아니다. 일부 경우, 엔도뉴클레아제는 미배양 미생물로부터 유래된다. 일부 경우, 이중-가닥 데옥시리보핵산 폴리뉴클레오티드는 원핵생물, 고세균, 박테리아, 진핵생물, 식물, 진균류, 포유류, 설치류, 또는 인간 이중-가닥 데옥시리보핵산 폴리뉴클레오티드이다. 일부 경우, 이중-가닥 데옥시리보핵산 폴리뉴클레오티드는 엔도뉴클레아제가 유래된 종 이외의 종으로부터의 원핵생물, 고세균, 또는 박테리아 이중-가닥 데옥시리보핵산 폴리뉴클레오티드이다.

[0141] 일 양태에서, 본 개시는 표적 핵산 유전자좌를 변형시키는 방법을 제공한다. 방법은 본원에 기술된 조작된 뉴클레아제 시스템을 표적 핵산 유전자좌에 전달하는 단계를 포함할 수 있다. 일부 경우, 엔도뉴클레아제는 조작된 가이드 리보핵산 구조와 복합체를 형성하도록 구성된다. 일부 경우, 복합체는 해당 복합체가 표적 핵산 유전자좌에 결합할 때, 해당 복합체가 표적 핵산 유전자좌를 변형시키도록 구성된다. 일부 경우, 표적 핵산 유전자좌를 변형시키는 단계는 표적 핵산 유전자좌에 대한 결합, 니킹, 절단, 또는 마킹 단계를 포함한다.

[0142] 일부 경우, 표적 핵산 유전자좌는 데옥시리보핵산(DNA) 또는 리보핵산(RNA)을 포함한다. 일부 경우, 표적 핵산은 계놈 진핵생물 DNA, 바이러스 DNA, 또는 박테리아 DNA를 포함한다. 일부 경우, 표적 핵산은 박테리아 DNA를 포함한다. 박테리아 DNA는 엔도뉴클레아제가 유래된 종과 상이한 박테리아 종으로부터 유래될 수 있다. 일부 경우, 표적 핵산 유전자좌는 시험관 내에 있다. 일부 경우, 핵산 유전자좌는 세포 내에 있다. 일부 경우, 별도의 핵산 분자에 대해 암호화된 엔도뉴클레아제 및 조작된 가이드 핵산 구조체가 제공된다. 일부 경우, 세포는 원핵생물 세포, 박테리아 세포, 진핵생물 세포, 진균류 세포, 식물 세포, 동물 세포, 포유류 세포, 설치류 세포, 영장류 세포, 또는 인간 세포이다. 일부 경우, 세포는 엔도뉴클레아제가 유래된 종과 상이한 종으로부터 유래된다.

[0143] 일부 경우, 조작된 뉴클레아제 시스템을 표적 핵산 유전자좌에 전달하는 단계는 본원에 기술된 핵산 또는 본원에 기술된 벡터를 전달하는 단계를 포함한다. 일부 경우, 조작된 뉴클레아제 시스템을 표적 핵산 유전자좌에 전달하는 단계는 엔도뉴클레아제를 암호화하는 개방 해독 프레임을 포함하는 핵산을 전달하는 단계를 포함한다. 일부 경우에, 핵산은 엔도뉴클레아제를 암호화하는 개방 해독 프레임이 작동 가능하게 연결되는 프로모터를 포

함한다. 일부 경우, 조작된 뉴클레아제 시스템을 표적 핵산 유전자좌에 전달하는 단계는 전술한 엔도뉴클레아제를 암호화하는 개방 해독 프레임에 함유하는 캡핑된 mRNA를 전달하는 단계를 포함한다. 일부 경우, 조작된 뉴클레아제 시스템을 전술한 표적 핵산 유전자좌에 전달하는 단계는 번역된 폴리펩티드를 전달하는 단계를 포함한다.

[0144] 일부 경우, 조작된 뉴클레아제 시스템을 표적 핵산 유전자좌에 전달하는 단계는 리보핵산(RNA) pol III 프로모터에 작동 가능하게 연결된 조작된 가이드 리보핵산 구조를 암호화하는 데옥시리보핵산(DNA)을 전달하는 단계를 포함한다. 일부 경우, 엔도뉴클레아제는 표적 유전자좌에서 또는 이에 근접하여 단일-가닥 파단 또는 이중-가닥 파단을 유도한다.

[0145] 본 개시의 시스템은, 예를 들어 핵산 편집(예를 들어, 유전자 편집), 핵산 분자에 대한 결합(예를 들어, 서열-특이적 결합)과 같은 다양한 응용에 사용될 수 있다. 이러한 시스템은, 예를 들어 대상체에서 질환을 유발할 수 있는 유전적으로 물려받은 돌연변이를 처리(예를 들어 제거 또는 치환)하는 데 사용될 수 있고, 세포에서 유전자의 기능을 확실하게 하기 위해 유전자를 불활성화시키는 데 사용될 수 있고, (예를 들어, 역-전사된 바이러스 RNA를 절단하거나 질환-유발 돌연변이를 암호화하는 증폭된 DNA 서열을 절단함으로써) 질환을 유발하는 유전적 요소를 검출하기 위한 진단 도구로서 사용될 수 있고, 특정 뉴클레오티드 서열(예를 들어, 박테리아에서 항생제 내 박테리아를 암호화하는 서열)을 표적화하고 검출하기 위한 프로브와 조합된 비활성화된 효소로서 사용될 수 있고, 바이러스 계통을 표적화함으로써 바이러스를 불활성화시키거나 바이러스가 숙주 세포를 감염시킬 수 없게 하는 데 사용될 수 있고, 유전자를 추가하거나 대사 경로를 변경하여 유기체가 귀중한 소분자, 거대분자, 또는 이차 대사물을 생산하도록 이를 조작하는 데 사용될 수 있고, 진화적 선택을 위한 유전자 구동 요소를 확립하는 데 사용될 수 있고, 바이오센서로서 외래 소분자 및 뉴클레오티드에 의한 세포 섭동을 검출하는 데 사용될 수 있다.

[0146] 실시예

[0147] 실시예 1 - 메타게놈에 의한 신규 Cas 효과기의 발견

[0148] 메타게놈 마이닝

[0149] 퇴적물, 토양 및 동물로부터 메타게놈 샘플을 수집하였다. 데옥시리보핵산(DNA)을 Zymobiomics DNA 미니-분취 키트로 추출하고 Illumina HiSeq[®] 2500 상에서 시퀀싱하였다. 재산 소유자의 동의 하에 샘플을 수집하였다. Qiagen DNeasy PowerSoil 키트 또는 ZymoBIOMICS DNA Miniprep Kit를 사용하여 샘플로부터 DNA를 추출하였다. 시퀀싱 라이브러리 제조(Illumina TruSeq) 및 Illumina HiSeq 4000 또는 Novaseq 상에서의 시퀀싱을 위해, DNA를 UC Berkeley의 Vincent J. Coates Genomics Sequencing Laboratory에 보냈다(400-800 bp 표적 삽입 크기를 갖는 쌍을 이룬 150개 염기쌍(bp) 판독). 또한, 공개적으로 이용 가능한 고온, 그리고 토양 및 해양 메타게놈 시퀀싱 데이터를 NCBI SRA로부터 다운로드하였다. BMap(Bushnell B., sourceforge.net/projects/bmap/)을 사용하여 시퀀싱 판독을 트리밍하고 Megahit(<https://paperpile.com/c/QSZG6K/cIMrh>)로 조립하였다. Prodigal(<https://paperpile.com/c/QSZG6K/BJ6oW>)로 단백질 서열을 예측하였다. 보고된 II형 CRISPR 뉴클레아제의 HMM 프로파일을 구축하고 HMMER3(hmmer.org)을 사용하여 모든 예측된 단백질에 대해 검색하였다. Mincd(<https://github.com/ctSkenner/ctSkenner/minced> 또는 <https://paperpile.com/c/QSZG6K/OPC44>)로 조립된 콘티그에 대해 CRISPR 어레이를 예측하였다. Kaiju(<https://paperpile.com/c/QSZG6K/nMi6k>)를 사용하여 단백질에 대한 분류법을 할당하고, 모든 암호화된 단백질의 컨센서스를 발견하여 콘티그 분류법을 결정하였다.

[0150] 예측 및 참조(예를 들어, SpCas9, SaCas9, 및 AsCas9) II형 효과기 단백질을 MAFFT(<https://paperpile.com/c/QSZG6K/sVHNH>)로 정렬시키고, FastTree2(<https://paperpile.com/c/QSZG6K/osZNM>)를 사용하여 계통유전학 트리를 추론하였다. 본 연구에서 회수된 서열로 구성된 계통군으로부터 신규 계열을 식별하였다. 계열 내로부터, 실험실 분석을 위한 모든 성분을 함유하는 경우(즉, CRISPR 어레이 및 예측된 tracrRNA를 갖는 양호하게 조립되고 주석이 달린 콘티그 상에서 발견됨), 이는 후보로 선택되었다. 선택된 대표적인 서열 및 참조 서열을 MUSCLE(<https://paperpile.com/c/QSZG6K/IT01a>)를 사용하여 정렬하여 측매 및 PAM 상호작용 잔기를 식별하였다.

[0151] 이러한 메타게놈 워크플로우는 본원에 기술된 SMART(SMall ARchaeal-associaTed) 엔도뉴클레아제 시스템에 대한 설명을 제공하였다.

- [0152] 활성 잔기 시그니처를 함유하는 SMART 엔도뉴클레아제의 발견
- [0153] 메타게놈 데이터로부터 조립된 수만 개의 고품질 CRISPR Cas 시스템의 마이닝으로부터, RuvC 및 HNH 도메인 둘 모두를 포함하지만 비정상적으로 작은 크기(< 900 aa)를 갖는 신규 효과기를 발견하였다(도 21a). 이들 효과기 뉴클레아제는 참조 포인트로서의 고세균 Cas9 엔도뉴클레아제에 대해 낮은 서열 유사성(< 20% 아미노산 동일성)을 나타냈다. 효과기 단백질 서열의 계통유전학 분석은 SMART 시스템이 하위유형 A, B, 또는 C로부터의 양호하게 연구된 II형 시스템에 비해 다양성을 갖는 군임을 나타냈다(도 1a 및 21b).
- [0154] 이들 콤팩트 "SMART" 효과기(약 400 - 1000개 아미노산, 도 2)는 CRISPR 어레이에 인접한 게놈의 유전자좌에 나타났다. 이들 인접한 SMART 유전자좌 중 일부는 또한 동일한 오페론(operon) 내에서의 tracrRNA 및 CRISPR 적응 유전자(예를 들어, 스페이서 획득에 관여하는 유전자) cas1, cas2, 또는 cas4를 암호화하는 것으로 예측된 서열을 포함하였다(도 3 및 도 21a). 이들의 콤팩트한 크기에도 불구하고, SMART 효과기는 참조 SaCas9 서열과 정렬될 때 6개의 추정 HNH 및 RuvC 촉매 잔기를 함유한다(도 4). 또한, 3D 구조 예측은 가이드 및 표적 결합뿐만 아니라 PAM의 인식에 관여하는 잔기를 식별하였으며, 이는 SMART 효과기가 활성 dsDNA 엔도뉴클레아제임을 시사한다.
- [0155] SMART 엔도뉴클레아제의 다수의 군
- [0156] 주요 촉매 및 결합 잔기의 위치에 기초하여, SMART 뉴클레아제는 3개의 RuvC 도메인, 일반적으로 RRxRR 모티프(예를 들어, PF14239 상동성을 갖는 도메인)를 함유하는 아르기닌 풍부 영역, HNH 엔도뉴클레아제 도메인, 및 추정 인식 도메인을 포함한다(도 5 및 도 6). 이들 도메인은 참조 서열과 낮은 서열 유사성을 공유한다(도 7). 또한, SMART 효과기뿐만 아니라 참조 고세균 서열은 Cas9 뉴클레아제보다 유의하게 보다 빈번하게 RRxRR 및 아연 결합 리본 모티프(CX[2-4]C 또는 CX[2-4]H)를 함유한다(도 8). 또한, Cas9 효과기 서열과 달리, 대부분의 SMART 효과기는 종종 다양한 엔도뉴클레아제와 연관된 Pfam 도메인 PF14239에 대한 유의한 히트를 함유한다. SMART 효과기 크기, 계통유전학 관계, 및 오페론 및 도메인 아키텍처 둘 모두의 차이에 기초하여, 이들 시스템을 2개의 주요 군으로 분류하였다: SMART I 및 SMART II. 이들 군의 중요한 특징은 아래의 표 3에 요약되어 있으며, 이는 또한 클래스 2, II형 A/B/C Cas 효소와 비교된 상이점을 예시한다.

[표 3] 본원에 기술된 SMART I 및 II 그룹 효소의 속성

속성	SMART I	SMART II	II 형: A, B, C
Zn-결합 잔기	예	예	아니오
가교 나선	예	아니오	예
PAM 상호작용 및 WED 도메인	예	아니오	예
RRxRR 모티프	예	예	아니오
REC 도메인	PF14239 와 상동성을 갖는 신규 도메인	C-말단에서의 신규 도메인	Cas9 REC 도메인
PF14239 상동성을 갖는 도메인	예	예	아니오
단일계통 계통군	예	아니오	예
TnpB 에 관련됨	예	예	예
오페론은 IS605 Tns 반복을 포함함	아니오	일부	아니오
<900 aa	일부	예	아니오
CRISPR-연관	예	일부	예
RuvC 및 HNH 함유	예	예	예

[0157]

[0158]

SMART 뉴클레아제는 Cas9에서와 같이 RuvC 및 HNH 도메인을 함유하지만, RuvC-I, 가교 나선 및 인식 도메인은 양호하게 정렬되지 않는다. SMART 뉴클레아제와 참조 서열 사이의 진화적 관계를 최대한으로 이해하기 위해, 전장 SMART, 보고되고 분류된 참조 II형 서열(예를 들어, Burstein, D. 등의 문헌[New CRISPR-Cas systems from uncultivated microbes. *Nature* **2017**, *542*, 237-241]; 및 Gasiunas, G. 등의 문헌[A catalogue of biochemically diverse CRISPR-Cas9 orthologs. *Nat Commun* **2020**, *11*, 5512] 참조, 이들 각각의 전체 내용은 본원에 참조로서 통합됨)뿐만 아니라, > 10,300개의 최근 보고된 Cas9 상동체 및 IscB 서열(예를 들어, Altae-Tran, H. 등의 문헌[The widespread IS200/605 transposon family encodes diverse programmable RNA-guided endonucleases. *Science* **2021**, *374*, 57-65] 참조, 이의 전체 내용은 본원에 참조로서 통합됨)의 다중 서열 정렬을 생성하였다. RuvC-II/HNH/RuvC-III 도메인을 포함하는 트리밍되고 양호하게 정렬된 영역을 유지하였다. 이러한 최종 정렬로부터 추론된 계통유전학 분석은 II-A, II-B, 및 II-C로서 현재 분류된 보고된 Cas9 효과기로부터 멀리 클러스터링하는 효과기의 다양한 계통군을 나타냈다(도 21e). 분류된 II형 효과기에 계통유전학적으로 더 가까운 것으로 밝혀진 2개의 SMART 계통군은 CRISPR 어레이에 인접하여 암호화될 가능성이 보다 높았다(도 21b, 21c, 및 21e). SMART 뉴클레아제의 MG33 계열은 II-C2형 효과기와 클러스터링되어, 해당 계통군을 크게 확장시킨다(도 21e 및 21f, 연보라색 분지). 이러한 계열은 SMART 효소 중 가장 큰 900-1050 aa의 대표를 포함하며, 이들의 길이 분포는 가장 작은 분류된 II-C형 효소와 중첩된다. 보다 먼 SMART 계통군(도 21e 및 21f, 청록색, 녹색 및 황색 분지)는 최근 II-D형으로 분류된 "조기 Cas9" 서열을 함유한다(도 21e 및 21f, 연회색 분지). 이들 CRISPR 시스템은 대체적으로 SMART로 총칭될 수 있다.

[0159]

SMART I 엔도뉴클레아제

[0160]

SMART I 효과기는 대략 600개 아미노산 내지 1,050개 아미노산 크기의 범위이다. 이들의 계통 컨텍스트에서의 공통적인 특징은 적응 모듈 유전자(예를 들어, 스페이서 획득에 관여하는 유전자) 및 CRISPR 어레이 근처에서의 예측된 tracrRNA였으며, 이의 구성은 II형 및 V형 CRISPR 시스템과 유사하였다(도 3a, 3b 및 3c). SMART I 효과기의 RRxRR 모티프 함유 영역은 고유하지만, 이는 Cas9 뉴클레아제의 아르기닌 풍부 가교 나선과 유사한 기능적 역할을 할 수 있다. SaCas9 결정 구조에 대한 모델링에서, SMART I 효과기의 예측된 3D 구조는 인식 로브(중

중 Pfam 도메인 PF14239를 함유함) 및 RuvC-II 도메인 내에서 미정렬 영역을 나타냈다(도 5). 해당 결과는 이들 도메인이 다른 II형 효과기에 비해 상이한 기원을 갖는다는 것을 나타냈다. II형 효과기 계통유전학 트리에서의 분산된 위치 및 보고된 II형 효과기와와 낮은 서열 유사성과 함께 고려하면(도 1a 및 도 21b), 이들 결과는 SMART I 엔도뉴클레아제가 II형 CRISPR 시스템의 새로운 군에 속한다는 것을 나타낸다. CRISPR 시스템에 대한 공인된 분류에 따라, 이들 SMART I 시스템은 II-D형으로 분류되었다.

[0161] SMART I MG34-1 시스템에 대한 환경 RNA 발현 데이터를 사용하여 추정 단일 가이드 RNA(sgRNA)를 조작하였다. 또한, SMART I 반복 및 tracrRNA 예측으로부터 설계된 다수의 sgRNA를 시험관 내 PAM 농축 검정으로 시험하였다. SMART I 효소의 경우, PAM 서열의 최적 식별은 해당 단계에서 말단 복구 및 무딘 말단 절찰을 사용하여 수행하였으며, 이는 이들 효소가 엇갈린 이중 가닥 DNA 절단을 생성할 수 있음을 시사한다. 검정을 통해 다중 sgRNA 설계를 사용한 MG34-1(서열번호 2), MG34-9(서열번호 9), 및 MG34-16(서열번호 17)에 대한 dsDNA 절단을 확인하였다(도 7, 서열번호 612-615의 사용을 도시함). MG34-1은 표적 인식 및 절단에 대한 NGGN PAM에 대한 선호도를 입증한 반면(도 8a), MG102 계열의 구성원은 표적 인식 및 절단에 대한 3' NRC PAM을 인식한다(도 21c). 절단 부위의 분석은 위치 7에서의 우선적인 절단을 나타냈다(도 8b 및 도 22a). 이러한 결과는, PAM으로부터의 위치 2-3에서 우선적으로 절단하는 다른 II형 효소로부터의 절단 메커니즘과 비교 시, 신규 생화학적 메커니즘을 제시하며, 이는 SMART I CRISPR 시스템에 대한 새로운 분류를 뒷받침한다.

[0162] 일부 SMART I 시스템에 대한 환경 발현 데이터는 CRISPR 어레이의 인 시츄 전사 및 예측된 tracrRNA를 암호화하는 유전자간 영역을 확인하였다(도 3b 및 3c). 또한, 활성 CRISPR 표적화의 사례를, 동일하거나 관련된 메타게놈으로부터 조립된 다른 게놈 서열과 일치하는 스페이서 서열을 검색함으로써 평가하였다. 이들 선을 따라, SMART I CRISPR 어레이에서의 암호화된 스페이서 중 하나에 의해 표적화되는 파지 게놈(도 3c 및 3d)이 식별되었다. 표적 서열에 인접한 영역의 분석은 GG 모티프를 함유하는 3' PAM 서열을 제시한다(도 3d). 이들 결과는 SMART I CRISPR 시스템이 파지 방어에 관여하는 RNA 유도 효과기로서 이의 자연 환경에서 활성이며, 표적화된 DNA 또는 RNA를 절단하거나 분해하는 뉴클레아제로서 기능할 가능성이 있음을 나타낸다.

[0163] SMART I 효과기는 활성, RNA 가이드 dsDNA CRISPR 엔도뉴클레아제임

[0164] SMART I MG34-1 및 MG34-16 시스템에 대한 환경 RNA 발현 데이터를 사용하여 추정 단일 가이드 RNA(sgRNA)를 조작하였다(도 3b 및 3c, 및 도 9). 또한, SMART I 반복 및 tracrRNA 예측으로부터 설계된 다수의 sgRNA를 시험관 내 PAM 농축 검정으로 시험하였다(도 10). 이에 대한 검정은 다중 sgRNA 설계를 사용한 MG34-1, MG34-9, 및 MG34-16에 대한 프로그램 가능한 dsDNA 절단을 확인하였다(도 10). MG34-1 및 MG34-9는 표적 인식 및 절단을 위해 NGGN PAM을 필요로 한다(도 11a 및 11c). 절단 부위의 분석은 위치 7에서 우선적인 절단을 나타낸다(도 11b 및 11c). 이러한 결과는, PAM으로부터의 위치 3에서 우선적으로 절단하는 Cas9 효소와 비교 시, 신규 생화학적 메커니즘을 제시하며, 이는 SMART I CRISPR 시스템에 대한 새로운 분류를 추가로 뒷받침한다.

[0165] 말단 복구 절차를 갖지 않는 PAM 농축 검정은 SMART I 뉴클레아제에 대한 활성을 나타내지 않았다. PAM 농축 프로토콜에서의 절찰 전에 무딘 말단 단편을 생성하기 위한 말단 복구에 대한 요건은 이들 효소가 엇갈린 이중 가닥 DNA 절단을 생성함을 나타낸다. MG34-1 뉴클레아제의 절단 산물의 시퀀싱으로 엇갈린 이중 가닥 절단을 확인하였다(도 22a). 이러한 결과는, PAM으로부터의 위치 2-3에서 우선적으로 절단하는 가장 많이 보고된 II형 효소로부터의 메커니즘과 비교 시, 신규 생화학적 메커니즘을 제시한다. 정제된 단백질을 이용한 시험관 내 절단 검정은 표적 가이드 18 bp 길이의 표적화된 DNA 절단에서 MG34-1이 보다 효율적임을 나타내고, 시계열 절단 검정은 동일한 가이드로 시험했을 경우 MG34-1이 참조 SpCas9와 비교하여 보다 느린 속도로 절단함을 나타낸다(도 22b 및 22c).

[0166] 대장균에서 수행된 실험은 해당 시스템이 세포에서 뉴클레아제로서 기능하는 데 요구되는 활성을 갖는 것을 나타냈다. MG34-1 및 MG34-9 sgRNA를 발현하는 대장균 균주를 sgRNA에 대한 표적을 함유하는 카나마이신 내성 플라스미드로 형질전환시켰다. 항생제의 존재 하에서, 항생제 내성 플라스미드의 성공적인 표적화 및 절단은 성장 결함을 야기할 것이다. 이에 대한 검정은 sgRNA에 대한 표적을 함유하지 않은 카나마이신 내성 플라스미드로 수행된 대조군 실험과 비교하여 약 2배 내지 10배의 성장 억제를 나타냈다(도 12 및 도 22d).

[0167] SMART II 엔도뉴클레아제

[0168] SMART II 효과기는 SMART I 효과기 대비 보다 적은 크기 분포(약 400개 아미노산 내지 600개 아미노산)를 갖는다. 이들의 게놈 컨텍스트는 비정상적인 반복 영역 또는 CRISPR 어레이를 제시하였다. 비-CRISPR 반복 영역은 약 10 내지 30 bp 초과 크기 범위의 직접 반복을 함유한다. 일부 경우, 이들은 다수의 구별되는 반복 유닛을 포

함한다. 때때로, 공통 CRISPR 식별 알고리즘은 이들 영역을 CRISPR 시스템으로 플래그 표시할 것이지만; 보다 면밀한 조사는 스페이스 서열로서 식별된 영역이 해당 어레이에서 반복된다는 것을 나타낼 것이다. 해당 어레이는 효과기에 바로 인접하지 않지만, 이들은 동일한 게놈 영역 내에 있다(도 3a, MG35-236 및 도 13a, 예를 들어, 효과기 유전자로부터 > 20 kb). SMART II 시스템 오페론에는 대체적으로 적응 모듈 유전자(예를 들어, 스페이스 획득에 관여하는 유전자)가 없었다.

[0169] 구조 예측은 클래스 2, II형 Cas 효과기에서 종종 발견되는 6개의 RuvC 및 HNH 뉴클레아제 촉매 잔기(도 6) 모두에 더하여, 가이드 RNA 결합, 표적 절단, 및 PAM의 인식 및 PAM과의 상호작용에 관여하는 Cas 효소의 특징적인 잔기를 식별하였다. 또한, SMART II 효과기는 표적 핵산 모티프의 인식 및 결합에 관여할 가능성이 있는 다수의 RRXRR 및 아연 결합 리본 모티프(CX_[2-4]C 또는 CX_[2-4]H)를 함유하였다. 주요 잔기의 위치에 기초하면, SMART II 뉴클레아제의 예측된 도메인 아키텍처는 3개의 RuvC 하위도메인, RRXRR 모티프를 함유하는 아르기닌-풍부 영역(예를 들어, PF14239 상동성을 갖는 도메인), HNH 엔도뉴클레아제 도메인, 알려지지 않은 도메인, 및 인식 도메인(REC)을 포함하였다(도 6). SMART II 효과기의 도메인 아키텍처는 II형 Cas9 뉴클레아제에 대한 보고된 도메인 아키텍처와 상이하였다(도 6 및 도 14).

[0170] 일부 SMART II 시스템에 대한 환경 전사체 데이터는 자연 환경에서 CRISPR 어레이 및 다른 반복 영역의 인 시츄 발현을 확인하였다(도 13a). 일부 SMART II 효과기의 5' 미번역 영역(UTR)의 전사 또한 환경 발현 데이터로부터 관찰되었으며(도 13b 및 도 16), 이는 해당 영역이 뉴클레아제 활성 또는 SMART 시스템의 조절에 중요할 수 있음을 시사한다.

[0171] SMART II 효과기 단백질, 반복 영역, 및 연관 유전자간 영역으로 수행된 예비 시험관 내 실험은 이들 효소가 가능하게는 프로그래밍 가능한 방식으로 dsDNA를 절단하는 능력을 갖는다는 것을 나타낸다(도 15 및 도 17). 해당 결과는 SMART II 뉴클레아제 활성이 RNA 또는 DNA 가이드될 수 있음을 시사하며, 이는 CRISPR 어레이와 같은 반복 영역의 사용, 또는 TIR 또는 5' UTR과 같은 유전자와 내에 암호화된 특징부의 인식을 통하는 것이 요구될 수 있다. SMART II 효과기의 5' UTR은 시험관 내 전사 검정에서 능동적으로 전사되며 고도의 이차 구조를 나타낸다(도 18). SMART II 효과기의 시작 코돈으로부터 바로 상류에 있는 영역의 다중 서열 정렬은 보존 블록을 나타내며(도 19), 이는 SMART II 효과기와 연관된 5' UTR이 절단 활성을 위한 DNA를 표적화하기 위한 효과기에 대한 RNA 가이드를 암호화함을 시사한다.

[0172] 최근, 짧은 Cas9 상동체는 효과기의 5' UTR 영역에 암호화된 가이드 RNA를 사용하여 프로그래밍 가능한 dsDNA 뉴클레아제인 것으로 보고되었다(Altae-Tran, Kannan 등의 문헌[*Science* 2021]). 이들 시스템에서, 효과기의 전사된 5' UTR로부터의 상류에서 표적화 "스페이스"가 식별되었으며, 이는 SMART II 효소가 이들의 5' UTR에 암호화된 예측된 가이드 RNA의 5' 말단에 "표적 스페이스"를 첨가함으로써 특이적 DNA 부위를 표적화하고 절단하도록 재프로그래밍될 수 있음을 시사한다. SMART II 효과기의 5' UTR 영역에 암호화된 가이드 RNA의 5' 말단에 대한 표적 스페이스의 첨가는 표적 dsDNA 절단을 위한 효과기를 다양한 표적 인접 모티프(TAM)로 활성화시켰다(도 20).

[0173] 일부 SMART II 효과기는 트랜스포자아제 TnpA 및 TnpB를 암호화하는 추정 삽입 서열(IS) 옆에서 관찰되었다(도 3a). IS의 말단은 예측된 헤어핀 구조를 갖는 말단 반전 반복(TIR)을 함유하는 것으로 식별되었고, IS가 그 안에 통합될 가능성이 가장 높은 표적 부위 복제 또한 식별되었다. 또한, 일부 SMART II 유전자좌는 SMART II 효과기의 측면에 위치하는 추정 TIR을 암호화하였다(예를 들어, 도 3).

[0174] SMART HEARO 계통군은 바이러스-연관 RNA-가이드 dsDNA 뉴클레아제를 함유함

[0175] 계통유전학 분석에 따르면 600 aa 길이 미만의 SMART 뉴클레아제(도 21e, 라일락색 분지)는 2개의 주요 계통군을 형성하는 보고된 IscB 서열("Cas9-유사 삽입 서열"(예를 들어, Kapitonov, V. V., Makarova, K. S. 및 Koonin, E. V.의 문헌[ISC, a Novel Group of Bacterial and Archaeal DNA Transposons That Encode Cas9 Homologs. *J Bacteriol* 2016, 198, 797-807] 참조, 이의 전체 내용은 본원에 참조로서 통합됨)(도 21e, 짙은 회색 분지)과 함께 클러스터링된다. Kapitonov와 연구진은 RuvC 및 HNH 도메인의 존재에 기초하여 Cas9와의 IscB 상동성을 보고하였으며, 이어서 이러한 동일한 효소의 군의 PLMP 도메인을 기술하였다(예를 들어, Altae-Tran, H. 등의 문헌[The widespread IS200/605 transposon family encodes diverse programmable RNA-guided endonucleases. *Science* 2021, 374, 57-65] 참조, 이의 전체 내용은 본원에 참조로서 통합됨). 3D 구조 예측을 사용하여 이들 단백질이 대체적으로 RRXRR 모티프를 함유하는 아르기닌 풍부 영역을 함유함을 입증하였다. 아르기닌 풍부 영역은 Cas9의 가교 나선과 유사한 것으로 제안되었지만; 해당 영역 또는 RuvC-I 도메인 둘 모두 참조 3D 구조의 가교 나선 및 RuvC-I 도메인과 3D 공간에서 양호하게 정렬되는 것으로 발견되지 않았다. 이러한

IscB/SMART 효소는 PAM 상호작용 도메인이 결여되어 있다. 대신, Zn-결합 리본 모티프를 함유하는 C-말단 "WED/REC" 도메인은 표적 모티프 인식에 관여할 수 있다. 단백질 도메인, 촉매 잔기, 및 3D 모델은 Cas9와의 진화적 관계를 나타내지만, 대부분의 IscB/SMART 효과기는 CRISPR와 연관되지 않는다(예를 들어, 게놈 컨텍스트에서 CRISPR 반복에 근접하여 발견되지 않음). IscB/SMART 시스템을 포함하는 군은 대체적으로 크기가 컴팩트하고(약 400 내지 600 aa) 박테리아 및 고세균 게놈에 광범위하게 분포된다. 이들 효과기를 암호화하는 게놈 단편의 16% 이상이 바이러스 또는 프로파지(prophage)-유래일 가능성이 있는 것으로 분류되었으며, 이러한 시스템의 진화에 바이러스가 연루된다는 것으로 밝혀졌다.

[0176] SMART 시스템과 연관된 비암호화 RNA(ncRNA)를 검색한 결과, IscB/SMART 5' 미번역 영역(UTR)의 65%가 RFam 데이터베이스(RF02033)로부터의 HNH 엔도뉴클레아제-연관 RNA 및 ORF(HEARO) RNA에 대한 히트를 함유하는 것으로 밝혀졌다. 이러한 ncRNA는 생물정보학 분석으로부터 고도로 구조화된 RNA로 처음 설명되었지만(예를 들어, Weinberg, Z., Perreault, J., Meyer, M. M. 및 Breaker, R. R.의 문헌[Exceptional structured noncoding RNAs revealed by bacterial metagenome analysis. *Nature* **2009**, *462*, 656-659] 참조, 이의 전체 내용은 본원에 참조로서 통합됨), 관련 HEARO ORF의 기능은 보고되지 않았다(예를 들어, Harris, K. A. 및 Breaker, R. R.의 문헌[Large Noncoding RNAs in Bacteria. *Microbiol Spectr* **2018**, *6*] 참조, 이의 전체 내용은 본원에 참조로서 통합됨). 추정 HEARO HNH 엔도뉴클레아제 ORF는 또한 RuvC 및 HNH 촉매 도메인을 함유하고 IscB/SMART 효과기와 함께 클러스터링된다는 것이 확인되었다. 따라서, IscB, 소형 SMART, 및 HEARO ORF는 비-Cas 엔도뉴클레아제의 하나의 큰 군을 대표한다. 최근, IscB의 5' UTR은 dsDNA 뉴클레아제 활성화에 필요한 단일 가이드 RNA를 암호화하는 것으로 보고되었으며, 해당 저자는 이를 오메가(Omega) RNA로서 지칭한다(예를 들어, Altae-Tran, H. 등의 문헌[The widespread IS200/605 transposon family encodes diverse programmable RNA-guided endonucleases. *Science* **2021**, *374*, 57-65] 참조, 이의 전체 내용은 본원에 참조로서 통합됨). 기능에 대한 가이드 RNA의 요건 확인에서, IscB/SMART/HEARO 시스템의 5' UTR의 인 시츄 자연 발현을 관찰하였으며, 이는 시험관 내 전사 검정에 의해 재현되었다. 오메가 RNA 구조는 HEARO RNA와 높은 구조 유사성을 공유한다. IscB/SMART/HEARO 시스템을 통합하는 특징(광범위한 분류학적 기원 및 아르기닌 잔기의 농축)뿐만 아니라 이들 효소와 연관된 가이드 RNA의 연대기적 발견을 고려하여, 본 연구진은 IscB/SMART/HEARO 시스템을 SMART HEARO로서 광범위하게 기능적으로 분류하는 것을 지지한다(도 21e). 본 연구진은, Altae-Tran 및 Kannan 등(예를 들어, Altae-Tran, H. 등의 문헌[The widespread IS200/605 transposon family encodes diverse programmable RNA-guided endonucleases. *Science* **2021**, *374*, 57-65])이 기술한 바와 같이, 시험관 내에서 SMART HEARO 절단 활성을 평가하고 HEARO RNA의 5' "스페이스" 영역을 재프로그래밍하여 필요한 표적화 모티프를 식별하였다(도 21d). 또한, 대장균에서의 플라스미드 간접 검정은 SMART HEARO 뉴클레아제가 SpCas9에 비해 고도로 활성임을 나타내고(SpCas9가 나타난 약 98배 억제 대비 MG35-1의 경우 > 570배 억제, 도 25b), 특이성 실험은 프로토스펙트에서의 불일치에 대한 낮은 허용도를 나타낸다(도 25d).

[0177] 실시예 2 - 본원에 기술된 엔도뉴클레아제에 대한 PAM 서열 식별/확인

[0178] 추정 SMART 엔도뉴클레아제를 대장균 용해물-기반 발현 시스템(PURExpress, New England Biolabs)에서 발현시켰다. 이러한 시스템에서, 엔도뉴클레아제를 대장균에 대해 코돈 최적화하고, T7 프로모터 및 C-말단 His 태그를 갖는 벡터 내로 클로닝하였다. T7 프로모터 및 종결자 서열로부터 각각 프라이머 결합 부위 150 bp 상류 및 하류에서 유전자를 PCR 증폭시켰다. 해당 PCR 산물을 5 nM 농도의 NEB PURExpress에 첨가하고 37°C에서 2시간 동안 발현시켜 PAM 검정을 위한 엔도뉴클레아제를 생성하였다.

[0179] 본원에 기술된 각각의 SMART Cas 효소와 호환 가능한 추정 sgRNA는 시퀀싱 데이터로부터 조립된 콘티그 CRISPR 유전자좌에 조립된 RNAseq 관독으로부터 확인되었으며: 이차 구조는 Geneious 소프트웨어 패키지(<https://www.geneious.com>)의 CRISPR 어레이로부터의 반복 서열과 함께 RNAseq 데이터로부터의 tracr 영역에 대해 결정되었으며, 생성된 나선은 트리밍되고 GAAA 테트라 루프와 연결되었다. 다양한 길이의 반복-항-반복 나선 트리밍뿐만 아니라 상이한 스페이스 길이 및 상이한 tracr 종결 지점 또한 시험하였다(도 12, 서열번호 612-615를 나타냄). 그런 다음, 각각의 sgRNA를 조립 PCR을 통해 조립하고, SPRI 비드로 정제하고, 짧은 RNA 전사물에 대한 제조사의 권장 프로토콜(HiScribe T7 키트, NEB)에 따라 시험관 내 전사(IVT)하였다. RNA 전사 반응물을 Monarch RNA 키트로 세척하고, Tapestation(Agilent)을 통해 순도를 확인하였다.

[0180] PAM 서열은 추정 뉴클레아제에 의해 절단될 수 있는 무작위로 생성된 잠재적 PAM 서열을 함유하는 플라스미드를 시퀀싱하여 결정하였다. 이러한 시스템에서, 추정 뉴클레아제를 암호화하는 대장균 코돈 최적화된 뉴클레오티드 서열을 T7 프로모터의 조절 하에 PCR 단편으로부터 시험관 내에서 전사하고 번역하였다. T7 프로모터로 이루어진 최소 CRISPR 어레이를 갖는 제2 PCR 단편에 이어서 반복-스페이스-반복 서열을 동일한 반응으로 전사시켰다.

TXTL 시스템에서의 엔도뉴클레아제 및 반복-스페이서-반복 서열의 성공적인 발현에 이어지는 CRISPR 어레이 프로세싱은 활성 시험관 내 CRISPR 뉴클레아제 복합체를 제공하였다.

[0181] 8N 혼합 축퇴 염기(잠재적 PAM 서열)에 선행하는 최소 어레이와 일치하는 스페이서 서열을 함유하는 표적 플라스미드 라이브러리를 TXTL 반응(변역된 Cas 효소의 5배 희석물, 8 N PAM 플라스미드 라이브러리의 5 nM, 및 PAM를 표적화하는 sgRNA의 50 nM을 포함하는, 10 mM Tris pH 7.5, 100 mM NaCl, 및 10 mM MgCl₂)의 출력물과 함께 배양하였다. 1 내지 3시간 후, 반응을 중단시키고, DNA 클리닝 키트를 통해 DNA를 회수하였다. 어댑터 서열은 엔도뉴클레아제에 의해 절단된 활성 PAM 서열을 갖는 DNA에 무딘-말단으로 결합된 반면, 절단되지 않은 DNA는 결합을 위해 접근할 수 없었다. 이에 이어, 활성 PAM 서열을 포함하는 DNA 분절을 PCR에 의해 라이브러리 및 어댑터 서열에 특이적인 프라이머로 증폭시켰다. PCR 증폭 생성물을 겔 상에서 분해하여 절단 이벤트에 해당하는 앰플리콘을 식별하였다. 절단 반응물의 증폭된 분절은 NGS 라이브러리 제조를 위한 템플릿 또는 생거 시퀀싱을 위한 기질로도 사용되었다. 8N에서 시작하는 라이브러리의 하위 집합인 결과적인 라이브러리를 시퀀싱하여 CRISPR 복합체와 호환 가능한 PAM 활성을 갖는 서열을 밝혀냈다. 처리된 RNA 작제물을 사용한 PAM 시험의 경우, 시험관 내 전사된 RNA를 플라스미드 라이브러리와 함께 첨가하고 최소 CRISPR 어레이/tracr 템플릿을 생략한 것을 제외하고는 동일한 절차를 반복하였다. 이들 검정에서 다음의 스페이서 서열을 표적으로 사용하였다(5'-CGUGAGCCACCACGUCGCAAGCCUCGAC-3').

[0182] PAM 검정으로부터 원시 시퀀싱 판독을 수득한 후, 해당 판독을 Phred 품질 점수 > 20으로 필터링하였다. PAM에 인접한 백본으로부터의 보고된 DNA 서열을 나타내는 24 bp를 PAM-근위 영역을 찾기 위한 기준으로서 사용하였고 인접한 8 bp를 추정 PAM으로서 식별하였다. PAM과 연결 어댑터 사이의 거리 또한 각각의 리드에 대해 측정하였다. 기준 서열 또는 어댑터 서열과 정확하게 일치하지 않은 판독치를 제외시켰다. 가장 빈번한 절단 부위 ±2 bp를 갖는 PAM이 선택적으로 분석에 포함되도록 PAM 서열을 절단 부위 빈도로 필터링하였다. Logomaker(Tareen A, Kinney JB.의 문헌[Logomaker: beautiful sequence logos in Python. Bioinformatics. 2020;36(7):2272-2274], 이는 본원에 참조로서 통합됨)를 사용하여 서열 로고를 생성하기 위해 PAM의 필터링된 목록을 사용하였다.

[0183] **실시예 3 - 예측된 RNA 폴딩에 대한 프로토콜**

[0184] 활성 단일 RNA 서열의 예측된 RNA 폴딩을 Andronescu 2007의 방법을 사용하여 37° 에서 연산한다. 염기의 색상은 해당 염기의 염기 페어링 확률에 상응하며, 여기에서 적색은 높은 확률이고 청색은 낮은 확률이다.

[0185] **실시예 4 -시험관 내 절단 효율**

[0186] 엔도뉴클레아제는 프로테아제 결핍 대장균 B 균주에서의 유도성 T7 프로모터로부터의 His-태그된 융합 단백질로서 발현된다. 엔도뉴클레아제를 2개의 핵 국소화 신호(N-말단 NLS 뉴클레오플라스민 이분 및 C-말단 시미안 바이러스 40 T-항원 NLS PPKKKRK), 말토오스 결합 단백질(MBP) 태그, 담배 에칭 바이러스(TEV) 프로테아제 절단 부위, 및 6XHis 태그에 N으로부터 C 말단의 다음 순서로 융합하였다: 6XHis-MBP-TEV-NLS-유전자-NLS-STOP. 이 단백질을 자가유도 배지(MagicMedia ThermoFisher)로 NEB Iq 대장균에서 pTac 프로모터 하에 발현시키고, 30°C 에서 성장시키고, 16°C에서 유도하였다.

[0187] His-태그된 단백질을 발현하는 세포를 초음파처리로 용해시키고, His-태그된 단백질을 AKTA Avant FPLC(GE Lifescience) 상의 HisTrap FF 컬럼(GE Lifescience) 상의 Ni-NTA 친화도 크로마토그래피로 정제하였다. 용리액을 아크릴아미드 겔(Bio-Rad) 상의 SDS-PAGE로 분해하고, InstantBlue Ultrafast Coomassie(Sigma-Aldrich)로 염색하였다. ImageLab 소프트웨어(Bio-Rad)를 사용하는 단백질 밴드의 밀도계를 사용하여 순도를 결정하였다. 정제된 엔도뉴클레아제를 50 mM Tris-HCl, 300 mM NaCl, 1 mM TCEP, 5% 글리세롤로 이루어진 보관 완충액(pH 7.5) 내로 투석하고; -80°C에서 보관하였다.

[0188] 스페이서 서열 및 PAM 서열을 함유하는 표적 DNA(예를 들어, 실시예 2에서와 같이 결정됨)를 DNA 합성에 의해 작제하였다. PAM이 퇴행성 염기를 가질 경우 단일 대표 PAM이 시험을 위해 선택된다. 표적 DNA는, 일 단부로부터 700 bp에 위치한 PAM 및 스페이서를 갖는 PCR 증폭을 통해 플라스미드로부터 유래된 2200 bp의 선형 DNA로 구성된다. 성공적인 절단은 700 및 1500 bp의 단편을 생성한다. 표적 DNA, 시험관 내 전사된 단일 RNA, 및 정제된 재조합 단백질을 과량의 단백질 및 RNA와 함께 절단 완충액(10 mM Tris, 100 mM NaCl, 10 mM MgCl₂) 중에 합치고, 5분 내지 3시간, 통상적으로 1시간 동안 인큐베이션한다. RNase A를 첨가하고 60분의 인큐베이션에서 반응을 중단시킨다. 이에 이어, 반응물을 1.2% TAE 아가로오스 겔 상에서 분리하고, 절단된 표적 DNA의 분획을 ImageLab 소프트웨어에서 정량화한다.

[0189] **실시예 5 - 대장균 내 활성**

[0190] 대장균은 이중 가닥 DNA 절단을 효율적으로 복구할 능력이 없다. 따라서, 게놈 DNA의 절단은 치명적인 이벤트일 수 있다. 이러한 현상을 활용하여, 엔도뉴클레아제 활성을 대장균에서 이의 게놈 DNA에 통합된 스페이서/표적 및 PAM 서열을 갖는 표적 균주에서 엔도뉴클레아제 및 가이드 RNA를 재조합적으로 발현시킴으로써 시험하였다.

[0191] 박테리아 세포에서의 뉴클레아제 활성을 시험하기 위해, BL21(DE3) 균주(NEB)를 T7-유도 효과기 및 sgRNA를 함유하는 플라스미드(각각의 플라스미드에 대해 10 ng)로 형질전환시키고, 도말하고 밤새 성장시켰다. 생성된 콜로니를 3회 밤새 배양한 다음, SOB에서 계대배양하고 OD 0.4-0.6으로 성장시켰다. 표준 키트 프로토콜(Zymo Mix 및 Go 키트)에 따라 0.5 OD 등가의 세포 배양물을 화학적합성으로 만들고, 백본에 스페이서 및 PAM를 포함하거나 포함하지 않는 130 ng의 카나마이신 플라스미드로 형질전환시켰다. 열충격 후, 37°C에서 1시간 동안 SOC에서 형질전환체를 회복시키고, 뉴클레아제 효율을 유도 배지(항생제 및 0.05 mM IPTG를 포함하는 LB 한천 플레이트) 상에서 성장시킨 5배 희석 시리즈로 측정하였다. 희석 시리즈로부터 콜로니를 정량화하여 뉴클레아제 유도 플라스미드 절단으로 인한 전체 억제율 측정하였다.

[0192] 도 12는 해당 검정에 대한 결과를 나타낸다. 도 12에서, 패널(a)는 플라스미드 절단을 입증하는 대장균 균주의 복제 플레이팅을 나타낸다; MG34-1 및 sgRNA를 발현하는 대장균을 sgRNA(+sp)에 대한 표적을 함유하는 카나마이신 내성 플라스미드로 형질전환시켰다. 음성 대조군(표적 및 PAM(-sp) 없음) 대비 성장 손상(+sp)을 나타내는 플레이트 사본면은 효소에 의한 성공적인 표적화 및 절단을 나타낸다. 실험은 2회 반복으로 3회 수행되었다. 도 12에서, 패널 b는 비표적 대조군(-sp) 대비 표적 조건(+sp)에서의 성장 억제를 나타내는 패널 a의 복제 플레이팅 실험으로부터의 콜로니 형성 유닛(cfu) 측정 도표를 나타내며, 이는 플라스미드가 절단되었음을 입증한다. 도 12에서, 패널 c는 콜로니 형성 유닛(cfu) 측정(로그 스케일)의 막대 도표를 나타내며, 이는 다양한 SMART 뉴클레아제에 대한 비표적 대조군(녹색 막대) 대비 표적 조건(백색 막대)에서의 대장균 성장 억제를 나타낸다. 각각의 뉴클레아제에 대한 플라스미드 간접 검정을 SpCas9 양성 대조군과 함께 3회 수행하였다.

[0193] 게놈 DNA에 통합된 PAM 서열을 갖는 조작된 균주(예를 들어, 실시예 2에서와 같이 결정됨)를 엔도뉴클레아제를 암호화하는 DNA로 형질전환시킨다. 이에 이어서, 형질전환체를 화학적합성으로 만들고, 표적 서열에 특이적이거나("표적화") 표적에 비특이적인("비표적화") 50 ng의 가이드 RNA(예를 들어, crRNAs)로 형질전환시킨다. 열충격 후, 37°C에서 2시간 동안 SOC에서 형질전환을 회복시킨다. 이에 이어서, 뉴클레아제 효율은 유도 배지 상에서 성장시킨 5배 희석 시리즈로 측정된다. 콜로니를 희석 시리즈로부터 3회 정량화한다.

[0194] **실시예 6 - 포유류 세포에서의 MG CRISPR 복합체의 게놈 절단 활성 시험**

[0195] 포유류 세포에서의 표적화 및 절단 활성을 규명하기 위해, MG Cas 효과기 단백질 서열을 다음 2개의 포유류 발현 벡터에서 시험한다: (a) C-말단 SV40 NLS 및 2A-GFP 태그를 갖는 하나의 벡터, 및 (b) GFP 태그가 없고 2개의 SV40 NLS 서열을 갖는 하나의 벡터, 즉 N-말단 상의 하나의 벡터 및 C-말단 상의 하나의 벡터. NLS 서열은 본원에 기술된 NLS 서열 중 어느 하나를 포함한다. 일부 경우, 엔도뉴클레아제를 암호화하는 뉴클레오티드 서열은 포유류 세포에서의 발현을 위해 코돈 최적화된다.

[0196] 표적화 서열이 부착된 상응하는 cRNA 서열은 제2 포유류 발현 벡터 내로 클로닝된다. 2개의 플라스미드를 HEK293T 세포 내로 공동 형질감염시킨다. 발현 플라스미드 및 gRNA 표적화 플라스미드를 HEK293T 세포 내로 공동 형질감염시킨 후 72시간차에, DNA를 추출하고 이를 NGS-라이브러리의 제조에 사용한다. NHEJ 백분율은 포유류 세포에서의 효소의 표적화 효율을 입증하기 위해 표적 부위의 시퀀싱에서의 인델을 통해 측정된다. 각각의 단백질의 활성을 시험하기 위해 적어도 10개의 상이한 표적 부위를 선택하였다.

[0197] **실시예 7 - 본원에 기술된 MG 계열의 예측된 활성**

[0198] 인 시츄 발현 및 단백질 서열 분석은 이들 효소가 활성 뉴클레아제임을 나타냈다. 이들은 예측된 엔도뉴클레아제-연관 도메인(RRXRR 및 HNH-엔도뉴클레아제 Pfam 도메인과 일치함, 도 2, 도 3a 및 3b)를 함유하고, 예측된 HNH 및 RuvC 촉매 잔기를 함유한다(예를 들어, 도 2, 도 3a 및 3b의 직사각형). 또한, 리보뉴클레아제 H-유사 단백질 계열에서 발견되는 RRXRR 모티프의 존재는 잠재적 RNA 표적화 또는 뉴클레아제 활성을 나타낸다(도 2 참조).

[0199] 발현 데이터는 후보 MG34-1 뉴클레아제, tracrRNA 및 CRISPR 어레이에 대한 인 시츄 천연 활성을 입증한다(도 4).

[0200] **실시예 8 - mRNA 전달을 이용한 포유류 세포에서의 활성**

- [0201] mRNA를 이용한 세포 형질감염/형질변환을 사용하는 게놈 편집의 경우, 코딩 서열은 Twist Bioscience 또는 Thermo Fisher Scientific(GeneArt)으로부터의 알고리즘을 사용하여 최적화된 마우스 또는 인간 코돈이다. 카세트는 코딩 엔도뉴클레아제 서열에 첨부된 다음 2개의 핵 국소화 신호로 구성된다: 각각 N 및 C 말단에서의 SV40 및 뉴클레오플라스민. 또한, 인간 보체 3(C3)으로부터의 미번역 영역은 카세트 내의 코딩 서열에 5' 및 3' 둘 모두에 첨부된다.
- [0202] 그런 다음, 해당 카세트를 긴 폴리 A 신장부의 상류에 있는 mRNA 생산 벡터 내로 클로닝한다. mRNA 작제물 조직은 다음과 같을 수 있다: C3로부터의 5'UTR - SV40 NLS - 코돈 최적화 SMART 유전자 - 뉴클레오플라스민 NLS - C3로부터의 3'UTR - 107 폴리A 테일. 이어서, mRNA의 전사 실행은 조작된 T7 RNA 중합효소(Hi-T7: New England Biolabs)를 사용하는 T7 프로모터에 의해 유도된다. mRNA의 5' 캡핑은 CleanCap AG(Trilink Biolabs)를 사용하는 공동 전사에 의해 발생한다. 이어서, MEGAclean Transcription Clean-Up 키트(Thermo Fisher Scientific)를 사용하여 mRNA를 정제한다.
- [0203] 포유류 세포는 전사된 mRNA 및 관심 게놈 영역을 표적화하는 적어도 10개의 가이드 세트와, Lipofectamine Messenger Max(Thermo Fisher Scientific)를 사용하여 공동 형질감염된다. 일정 기간(예를 들어, 48시간) 동안 세포를 인큐베이션한 다음, Purelink Genomic DNA 추출 키트(Fisher Scientific)를 사용하여 게놈 DNA를 분리한다. 관심 영역은 특정 프라이머를 사용하여 증폭된다. 이어서, 편집 결과의 철저한 분석을 위해 CRISPR 편집 및 NGS의 추론(Inference of CRISPR Edits and NGS)을 사용하는 생거(Sanger) 시퀀싱으로 편집을 평가한다.
- [0204] **실시예 9 - SMART II 가이드 RNA 예측**
- [0205] SMART II 효과기 서열의 시작 코돈으로부터 바로 상류의 400 bp를 포함하는 영역을 활성화에 필요한 가이드 RNA(UTR)를 잠재적으로 암호화함으로써 추출하였다. UTR 서열을 MAFFT(mafft-ginsi 알고리즘)와 정렬시키고, 보존 블록을 나타내는 영역을 추정 가이드 RNA로서 주석을 달았다.
- [0206] **실시예 10 - 활성화 및 PAM 결정 검증**
- [0207] RNASeq 또는 UTR 정렬로부터 예측된 추정 가이드 RNA를 Geneious에서 폴딩시켰다. 표적 스페이서를 가이드 RNA의 5' 또는 3' 말단에 첨부하여 단일 가이드 RNA(sgRNA)를 설계하였다. sgRNA를 조합 PCR을 통해 조합하고, SPRI 비드로 정제하고, 짧은 RNA 전사물에 대한 제조사의 권장 프로토콜(HiScribe T7 키트, NEB)에 따라 시험관 내 전사(IVT)하였다. RNA 반응물을 Monarch RNA 키트로 세척하고, TapeStation(Agilent)을 통해 순도를 확인하였다.
- [0208] PURExpress(New England Biolabs)로 절단 및 PAM 결정 검정을 수행하였다. 요약하면, 단백질을 대장균에 대해 코돈 최적화하고, T7 프로모터 및 C-말단 His 태그를 갖는 벡터 내로 클로닝하였다. T7 프로모터 및 종결자 서열로부터 각각 프라이머 결합 부위 150 bp 상류 및 하류에서 유전자를 PCR 증폭시켰다. 해당 PCR 산물을 NEB PURExpress에 5 nM 농도로 첨가하고 37°C에서 2시간 동안 발현시켰다. 이 시점 후, 절단 반응물을 PURExpress의 5배 희석물, 8 N PAM 플라스미드 라이브러리의 5 nM, 및 PAM 라이브러리를 표적화하는 50 nM의 sgRNA와 함께, 10 mM Tris pH 7.5, 100 mM NaCl, 및 10 mM MgCl₂ 중에서 조합하였다.
- [0209] PURExpress 반응으로부터의 절단 생성물을 AMPure SPRI 비드(Beckman Coulter)로 세정하여 회수하였다. Klenow 단편 및 dNTP(New England Biolabs)의 첨가를 통해 DNA를 둔화시켰다. 무단-말단 생성물을 100배 초과하는 이중 가닥 어댑터 서열과 연결시키고, NGS 라이브러리의 제조를 위한 템플릿으로서 사용하였으며, 이로부터의 서열 분석으로부터 PAM 요건을 결정하였다.
- [0210] 원시 NGS 판독치를 Phred 품질 점수 > 20으로 필터링하였다. PAM에 인접한 백본으로부터의 보고된 DNA 서열을 나타내는 24 bp를 PAM-근위 영역을 찾기 위한 기준으로서 사용하였고 인접한 8 bp를 추정 PAM으로서 식별하였다. PAM과 연결 어댑터 사이의 거리 또한 각각의 리드에 대해 측정하였다. 기준 서열 또는 어댑터 서열과 정확하게 일치하지 않은 판독치를 제외시켰다. 가장 빈번한 절단 부위 ±2 bp를 갖는 PAM이 선택적으로 분석에 포함되도록 PAM 서열을 절단 부위 빈도로 필터링하였다. Logomaker를 사용하여, 필터링된 PAM 목록을 사용하여 시퀀스 로고를 생성하였다.
- [0211] **실시예 11 - SMART 아미노산 조성**
- [0212] SMART 단백질 서열의 아미노산 조성을 설명하기 위해, 각각의 잔기가 관찰된 횟수를 총 단백질 길이로 나누고, 여기에, 100을 곱하여, SMART 서열의 군에 대한 아미노산 함량 백분율을 계산하였다. 이어서, 해당 아미노산 조성을 Uniprot50 데이터베이스(Carugo의 문헌[Protein Sci. 2008])로부터의 방대한 단백질 서열 세트에 대해 보

고된 함량 백분율과 비교하였다. SMART HEARO 및 SMART(II-D형)의 두 단백질 군은 Uniref50 단백질 서열에서 관찰된 함량에 비해 비정상적으로 높은 아르기닌 및 리신 아미노산 함량을 함유한다(도 23).

[0213] 평균적으로, SMART의 아르기닌 및 리신 조성 백분율은 SMART 서열의 다른 잔기에 대해 관찰된 선형 추세뿐만 아니라 Uniref50 데이터베이스의 단백질의 잔기 조성으로부터 벗어난다(도 24a). 또한, SMART의 메티오닌 함량은 Uniref50 데이터베이스의 단백질에서 관찰된 함량보다 통계적으로 낮은 것으로 관찰되었다(도 24b).

[0214] SMART의 물리화학적 특성을 설명하기 위해, R의 "분자" 및 "펩티드" 패키지를 갖는 서열로부터 등전점, 분자량 및 전하를 결정하였다. SMART 서열에서 관찰된 높은 아르기닌 및 리신 함량은 중성 pH에서의 높은 등전점 및 전하에 기여할 수 있다(표 4).

[표 4] SMART 계열 구성원의 이론적 특성

뉴클레아제	길이(a.a.)	MW(Da)	pI	pH 7.2에서의 전하
MG35-1	428	48300.1	11.1	52.6
MG35-2	524	59310.4	10.4	34.3
MG35-3	423	47899.5	10.8	39.8
MG35-6	428	48373.1	10.9	43.6
MG35-102	424	47544.1	10.8	38.7
IscB (Altae-Tran, 2021)	439	49447.5	11.7	47.2
MG34-1	747	86518.4	10.2	46.9
MG102-2	946	107544.0	9.9	38.8
MG102-14	949	108596.1	10.2	50.6
MG102-35	954	108186.4	9.9	38.3
MG102-45	952	107614.8	10.0	40.5

특성은 R 패키지 펩티드(The R Journal. 7(1), 4-14(2015)) 및 protr(Bioinformatics, 2015년 6월 1:31(11):1857-9)를 사용하여 계산함. 세포내 pH가 7.0 내지 7.4의 범위인 경향이 있기 때문에 pH 7.2를 선택함(Biochemical Journal, 1988, 250(1): 1-8).

[0215]

[0216] SMART 뉴클레아제의 높은 아르기닌 및 Zn-결합 리본 모티프 함량은 이들 효소가 본질적으로 무질서한 영역을 함유할 수 있음을 시사하며, 이는 단백질이 대형 가이드 RNA 및 표적 DNA와 상호작용할 수 있는 유연성을 제공할 수 있다. 본질적으로 무질서한 영역은 원래의 결합되지 않은 상태에서 안정적인 삼차 구조가 결여된 단백질 분절이고(예를 들어, Bitard-Feildel, T., Lamiabie, A., Mornon, J.-P. 및 Callebaut, I.의 문헌[Order in Disorder as Observed by the "Hydrophobic Cluster Analysis" of Protein Sequences. *Proteomics* **2018**, 18, e1800054] 참조, 이의 전체 내용은 본원에 참조로서 통합됨), 이는 다중음이온(예컨대 RNA)과 상호작용하는 양 전하를 띤 아르기닌이 풍부할 수 있고(예를 들어, Murthy, A. C. 등의 문헌[Molecular interactions underlying liquid-liquid phase separation of the FUS low complexity domain. *Nat Struct Mol Biol* **2019**, 26, 637-648] 참조, 이의 전체 내용은 본원에 참조로서 통합됨), "검색 기능"을 보조하는 Zn 결합 리본 사이의 링커로 발견될 수 있으며(예를 들어, Dyson, H. J.의 문헌[Roles of intrinsic disorder in protein-nucleic acid interactions. *Mol Biosyst* **2011**, 8, 97-104] 참조, 이의 전체 내용은 본원에 참조로서 통합됨), 이들 모두는 SMART 뉴클레아제에서 관찰되는 특징이다.

[0217] 실시예 12 - 불일치 사멸 검정

[0218] 다양한 SMART 효소의 특이성을 결정하기 위해, 대장균 BL21(DE3) 균주(NEB)를 T7 유도 효과기(암피실린 내성) 및 T7-유도 sgRNA(클로람페니콜 내성)를 함유하는 플라스미드로 형질전환시키고, 도말하고, 밤새 성장시키는 불일치 사멸 검정을 개발하였다. 생성된 콜로니를 다음과 같은 세 가지 조건에서 100 ng의 카나마이신 플라스미드로 적격화시키고 형질전환시켰다: 백본 내의 표적 스페이서 및 PAM, 각각 24 nt 스페이서 및 일정한 PAM을 따라 단일 불일치를 함유하는 25개의 플라스미드 라이브러리, 또는 스페이서 또는 PAM이 없는 대조군 플라스미드(도 25d). 열 충격 후, 형질전환물을 37°C에서 2시간 동안 SOC 배지에서 회복시켰다. 배양물을 도말하고 유도 배지(항생제 및 0.05 mM IPTG를 포함하는 LB 한천 플레이트) 상에서 37°C에서 밤새 성장시켰다. 미니프랩 키트

(Qiagen)를 통해 생존하는 불일치 콜로니로부터 플라스미드를 추출하였다. PCR을 통해 표적 영역을 증폭시키고 NGS를 통해 분석하였다. 미처리 라이브러리에 비해 풍부한 스페이서는 뉴클레아제에 의해 인식되고 절단될 수 없었으므로, 효과기가 불일치를 허용하지 않는 영역인 것으로 간주된다. 불일치가 허용되는 경우, 효소는 항생제 내성 플라스미드를 절단할 것으로 예상되고 성장 장애가 관찰될 것이다. MG102-2 뉴클레아제는 PAM으로부터 표적 플라스미드의 처음 13개의 위치를 따라 불일치를 허용하지 않는 것으로 관찰된 한편, 위치 14로부터는 가변적인 불일치 허용이 관찰되었다(도 25d 및 도 27). 이러한 결과는 SMART 뉴클레아제가 고도로 특이적일 수 있으며 측부 ssDNA 절단을 나타내지 않음을 시사한다(도 28).

[0219] 실시예 13 - SMART 뉴클레아제 MG102-2를 이용한 인간 세포 편집

[0220] ATCC로부터의 K562 세포를 ATCC 프로토콜에 따라 배양하였다. TRAC 유전자좌를 표적화하는 2개의 sgRNA를 MG102-2 PAM에 기초하여 설계하고 IDT로 화학적으로 합성하였다. 유전자 편집 실험을 위해, 500 ng의 시험관 내 합성 MG102-2 mRNA 및 150, 300, 또는 450 pmol의 표지된 sgRNA 중 하나를 Lonza 4D Nucleofector(프로그램 FF-120)를 사용하여 1.5×10^5 세포 중에 공동-핵감염시켰다. 동시에, TRAC 가이드에 의해 표적화되는 부위의 백그라운드를 평가하기 위해 mRNA나 가이드 둘 모두 없이 세포를 핵감염시켰다. QuickExtract(Lucigen #09050)를 사용하는 게놈 DNA 추출을 위해 전기천공 후 72시간차에 세포를 수확하고, Illumina Miseq 상에서 차세대 시퀀싱에 대해 처리하였다. 생성된 데이터를 인텔 계산기 스크립트로 분석하였다.

[0221] SMART 뉴클레아제를 T 세포 수용체 알파 불변 유전자좌(TRAC)를 표적화하는 인간 세포에 mRNA를 통한 전달은 MG102-2 뉴클레아제를 사용한 2개의 TRAC 표적 부위 중 하나에서 90%를 초과하는 편집 활성을 야기하였다(도 26). 시험관 내 실험에서 관찰된 바와 같이(도 29), sgRNA 양의 증가는 표적 유전자좌 둘 모두에서의 편집 효율을 향상시켰다(도 26). 인간 세포의 핵에 대한 MG34-1 시스템의 국소화(핵 국소화 신호, NLS와 융합됨)가 확인되었지만, 뉴클레아제-유도 InDel 형성은 해당 뉴클레아제에 대해 검출되지 않았다.

[0222] 실시예 14 - SMART 뉴클레아제의 절단 선호도

[0223] MG34-1 및 MG102-2 뉴클레아제의 절단 산물의 시퀀싱은 이들 효소가 엇갈린 이중 가닥 DNA 절단을 생성함을 나타낸다(도 25a). 절단 부위의 분석은 PAM으로부터 위치 5 내지 7에서의 선택적 절단을 나타낸다(도 25a). 이러한 결과는, 무딘 말단뿐만 아니라, PAM으로부터 위치 3 내지 5에 우선적으로 있는 엇갈린 절단을 생성하는 대부분의 Cas9 효소와 비교하여, 드물게 관찰되는 생화학적 절단 메커니즘을 제시한다. 시험관 내 전사/번역 반응 및 정제된 단백질을 이용한 시험관 내 절단 검정은 MG34-1 및 MG102-2가 18 및 20개 뉴클레오티드 스페이서인 경우 가장 효율적임을 나타낸다(도 25c). 또한, 대장균 플라스미드 간접 검정으로 생체 내 활성을 확인하였으며, 이는 특정 표적화 스페이서를 갖는 5개의 SMART 뉴클레아제에 대해 2배(MG34-9) 내지 > 500배(MG102-2)의 성장 억제를 나타낸다(도 25b).

[0224] 실시예 15 - SMART I 효소는 인간 세포에서 활성 뉴클레아제임

[0225] ATCC로부터 구매한 K562 세포를 ATCC 프로토콜에 따라 배양하였다. TRAC 또는 AAVS1 유전자좌를 표적화하는 sgRNA를 MG102-2, MG102-36, MG102-39, MG102-42, MG102-45, 및 MG33-34에 의해 인식된 PAM에 기초하여 설계하고, IDT로 화학적으로 합성하였다. 유전자 편집 실험을 위해, 500 ng의 시험관 내 합성 뉴클레아제 RNA 및 450 pmol의 표지된 sgRNA를 Lonza 4D Nucleofector(프로그램 FF-120)를 사용하여 1.5×10^5 세포 중에 공동-핵감염시켰다. QuickExtract(Lucigen #09050)를 사용하는 게놈 DNA 추출을 위해 전기천공 후 72시간차에 세포를 수확하고, Illumina Miseq 상에서 앰플리콘 차세대 시퀀싱에 대해 처리하였다. 생성된 데이터를 자체개발 인텔 계산기 스크립트로 분석하였다.

[0226] 본원의 다른 곳에서 기술된 바와 같이, SMART I 뉴클레아제 MG102-2는 mRNA를 통해 전달될 경우 인간 게놈의 TRAC 유전자좌의 2개의 표적 부위에서 활성이다. 또한, MG102-2(서열번호 582)는 또한 인간 게놈의 AAVS1 유전자좌(안전 하버 유전자좌)에서 활성이며, 여기에서 효소의 절단 효율은 8개의 상이한 표적 부위에서 82.6%만큼 높고, > 50% 편집 효율을 가짐을 확인하였다(도 30a). 또한, MG102-39(서열번호 993), MG102-42(서열번호 996), 및 MG102-48(서열번호 1002)은 mRNA에 의해 전달될 경우 인간 게놈의 TRAC 유전자좌에서 > 40%의 절단 활성을 나타낸 한편(도 30b-30d), MG33-34(서열번호 988), MG102-36(서열번호 990), 및 MG102-45(서열번호 999)는 TRAC 유전자좌에서 백그라운드(10%)를 초과하는 절단 효율을 나타냈다(도 30e-30g).

[표 5] 실시예 15 에서 시험된 가이드 RNA 및 표적화 서열

서열번호	명칭	서열
1087	MG102-2 AAVS1 A5	mU ⁵ mU ⁵ mC ³ rUrGrGrGrArGrArGrGrUrArGrGrGrArGrGrGrUrUrUrCrArArUrCrArArArCrUrGrArArArGrUrUrCrGrGrUrUrGrArArArArGrArGrArUrCrCrGrUrCrCrGrGrArGrGrUrGrCrArCrUrCrGrGrGrArUrGrGrGrGrArGrUrCrCrGrGrCrArCrUrUrGrGrGrUrUrUrCrCrCrGrGrGrCrUrUrArCrGrCrUrUrCrGrGrArArArArGrGrGrArUrGrUrGrArGrCrCrArA ⁵ mU ⁵ mU ⁵ mU
1088	MG102-2 AAVS1 H8	mG ³ mC ³ mC ³ rCrUrGrGrGrArArUrArUrArArGrGrUrGrGrUrCrCrCrGrUrUrUrCrArArUrCrArArArCrUrGrArArArGrUrUrCrGrGrUrUrGrArArArGrArGrArGrCrArUrCrCrGrUrCrCrGrGrArGrGrUrGrCrArCrUrUrGrCrGrUrUrUrCrCrCrCrGrGrCrUrUrArCrGrCrUrUrCrGrGrArArArArArGrGrCrCrCrUrUrCrGrGrCrArCrGrUrCrGrArArGrArCrArGrGrArUrGrUrGrArGrCrCrCrArA ⁵ mU ⁵ mU ⁵ mU
1089	MG102-2 AAVS1 H9	mA ⁵ mU ⁵ mG ³ rCrUrGrUrCrCrUrGrArArGrUrGrGrArCrArUrArGrGrGrUrUrUrCrArArUrCrArArArCrUrGrArArArGrUrUrCrCrGrGrUrUrGrArArArGrArGrArCrUrCrCrGrUrCrCrGrGrArGrGrUrGrCrArCrUrCrGrGrGrArUrGrGrGrGrArGrUrCrCrGrGrCrArCrUrUrGrCrGrUrUrUrUrCrCrCrCrGrGrCrUrUrArCrGrCrUrUrCrGrGrArArArArArGrGrGrCrCrUrUrCrGrGrCrArCrGrUrCrGrArArGrArCrArGrGrArUrGrUrGrArGrCrCrCrArA ⁵ mU ⁵ mU ⁵ mU
1090	MG102-2 AAVS1 D11	mC ³ mU ⁵ mA ⁵ rGrArGrArGrGrUrArArGrGrGrGrGrUrArGrGrGrGrGrUrUrUrCrArArUrCrArArArCrUrGrArArArGrUrUrCrCrGrGrUrUrGrArArArGrArGrArCrUrCrCrGrUrCrCrGrGrArGrGrUrGrCrArCrUrCrGrGrGrArCrUrUrGrCrGrUrUrUrUrCrCrCrCrGrGrCrUrUrArCrGrCrUrUrCrGrGrArArArArArGrGrGrCrCrUrUrCrGrGrCrArCrGrUrCrGrArArGrArCrArGrGrArUrGrUrGrArGrCrCrCrArA ⁵ mU ⁵ mU ⁵ mU
1091	MG102-2 AAVS1 E7	mA ⁵ mG ³ mG ³ rArArGrGrArGrGrArGrGrCrCrUrArArGrGrArUrGrGrGrUrUrUrCrArArUrCrArArArCrUrGrArArArGrUrUrCrCrGrGrUrUrGrArArArGrArGrArCrUrCrCrGrUrCrCrGrGrArGrGrUrGrCrArCrUrUrGrCrGrUrUrUrGrGrGrCrArGrUrCrCrGrGrCrArCrUrUrGrCrGrUrUrUrUrCrCrCrCrGrGrCrUrUrArCrGrCrUrUrCrGrGrArArArArArGrGrGrCrCrUrUrCrGrGrCrArCrGrUrCrGrArArGrArCrArGrGrArUrGrUrGrArGrCrCrCrArA ⁵ mU ⁵ mU ⁵ mU
1092	MG102-2 AAVS1 D7	mA ⁵ mU ⁵ mA ⁵ rUrCrArGrGrArGrArCrUrArGrGrArArGrGrArGrGrGrUrUrUrCrArArUrCrArArArCrUrGrArArArGrUrUrCrCrGrGrUrUrGrArArArGrArGrArCrUrCrCrGrUrCrCrGrGrArGrGrUrGrCrArCrUrCrGrGrGrArUrGrGrGrGrCrArGrUrCrCrGrGrCrArCrUrUrGrCrGrUrUrUrUrCrCrCrCrGrGrCrUrUrArCrGrCrUrUrCrGrGrArArArArArGrGrGrCrCrUrUrCrGrGrCrArCrGrUrCrGrArArGrArCrArGrGrArUrGrUrGrArGrCrCrCrArA ⁵ mU ⁵ mU ⁵ mU
1093	MG102-2 AAVS1 B7	mC ³ mU ⁵ mG ³ rCrCrUrArArCrArGrGrArGrGrUrGrGrGrGrGrUrUrArGrUrUrUrCrArArUrCrArArArCrUrGrArArArGrUrUrCrCrGrGrUrUrGrArArArGrArGrArCrUrCrCrGrUrCrCrGrGrArGrGrUrGrCrArCrUrUrGrCrGrUrUrUrGrGrGrCrArGrUrCrCrGrGrCrArCrUrUrGrCrGrUrUrUrUrCrCrCrCrGrGrCrUrUrArCrGrCrUrUrCrGrGrArArArArArGrGrGrCrCrUrUrCrGrGrCrArCrGrUrCrGrArArGrArCrArGrGrArUrGrUrGrArGrCrCrCrArA ⁵ mU ⁵ mU ⁵ mU
1094	MG102-2 AAVS1 D12	mG ³ mC ³ mA ⁵ rArGrArGrGrArUrGrGrArGrArGrGrUrGrGrCrUrArArGrUrUrUrCrArArUrCrArArArCrUrGrArArArGrUrUrCrCrGrGrUrUrGrArArArGrArGrArCrUrCrCrGrUrCrCrGrGrArArGrArCrArGrGrArUrGrUrGrArGrCrCrCrArA ⁵ mU ⁵ mU ⁵ mU

[0227]

서열번호	명칭	서열
		UrCrCrGrGrArGrGrGrUrGrCrArCrUrCrCrGrGrGrArUrGrGrGrCrArGrUrCrCrCrGrGrCrArCrUrUrGrCrGrUrUrUrCrCrCrCrGrGrCrUrUrArCrGrCrUrUrCrGrGrArArArArGrGrCrCrUrUrCrGrGrCrArCrGrUrCrGrArArGrArCrArGrGrArUrGrUrGrArGrCrCrCrArA*ml ¹⁰ ml ¹⁰ ml ¹⁰
1095	MG102-2 AAVS1 C8	mG ² mA ² mG ² rGrGrGrArCrArGrArUrArArArArGrUrArCrCrCrArGrGrUrUrUrCrArArUrCrArArCrUrGrArArArGrUrUrCrCrGrGrUrUrGrArArArArGrArGrCrArUrCrCrGrUrCrCrGrGrGrUrGrCrArCrUrCrCrGrGrGrArUrGrGrGrCrArGrUrCrCrGrGrCrArCrUrGrGrArArArArGrGrGrCrCrCrUrUrArCrGrCrUrUrCrGrGrArArArArGrGrCrCrCrUrUrCrGrGrCrArCrGrUrCrGrArArGrArCrArGrGrArUrGrUrGrArGrCrCrCrArA*ml ¹⁰ ml ¹⁰ ml ¹⁰
1096	MG102-2 AAVS1 A8	mG ² ml ¹⁰ mG ² rGrCrCrCrCrArCrUrGrUrGrGrGrUrGrGrArGrGrGrGrUrUrUrCrArArUrCrArArCrUrGrArArArGrUrUrCrCrGrGrUrUrGrArArArGrArGrCrArUrCrCrGrUrCrCrGrGrArGrGrUrGrCrArCrUrCrCrGrGrGrArUrGrGrGrCrArGrUrCrCrCrGrGrCrArCrUrUrGrGrArArArArGrGrGrCrCrCrUrUrArCrGrCrUrUrCrGrGrArArArArGrGrCrCrCrUrUrCrGrGrCrArCrGrUrCrGrArArGrArCrArGrGrArUrGrUrGrArGrCrCrCrArA*ml ¹⁰ ml ¹⁰ ml ¹⁰
1097	MG102-2 AAVS1 G6	ml ¹⁰ mG ² mG ² rCrUrCrCrArGrGrArArArUrGrGrGrGrUrGrUrGrUrGrUrUrUrCrArArUrCrArArCrUrGrArArArGrUrUrCrCrGrGrUrUrGrArArArGrArGrCrArUrCrCrGrUrCrCrGrGrGrUrGrCrArCrUrCrCrGrGrGrArUrGrGrGrUrUrUrCrCrCrCrGrGrCrUrUrArCrGrCrUrUrCrGrGrArArArArGrGrCrCrCrUrUrCrGrGrCrArCrGrUrCrGrArArGrArCrArGrGrArUrGrUrGrArGrCrCrCrArA*ml ¹⁰ ml ¹⁰ ml ¹⁰
1098	MG102-2 AAVS1 E5	mG ² ml ¹⁰ mG ² rGrCrCrArCrUrGrArGrArArCrCrGrGrGrCrArGrGrUrGrUrUrUrCrArArUrCrArArCrUrGrArArArGrUrUrCrCrGrGrUrUrGrArArArGrArGrCrArUrCrCrGrUrCrCrGrGrArGrGrUrGrCrArCrUrCrCrGrGrGrArUrGrGrGrGrCrArGrUrCrCrCrGrGrCrArCrUrUrGrGrArArArArGrGrGrCrCrCrUrUrArCrGrCrUrUrCrGrGrArArArArGrGrGrCrCrCrUrUrCrGrGrCrArCrGrUrCrGrArArGrArCrArGrGrArUrGrUrGrArGrCrCrCrArA*ml ¹⁰ ml ¹⁰ ml ¹⁰
1099	MG102-2 AAVS1 G7	ml ¹⁰ mC ² ml ¹⁰ rGrUrCrArCrArArUrCrCrUrGrUrCrCrCrUrArGrUrGrUrUrUrCrArArUrCrArArCrUrGrArArArGrUrUrCrCrGrGrUrUrGrArArArGrArGrCrArUrCrCrGrUrCrCrGrGrArGrGrUrGrCrArCrUrCrCrGrGrGrArUrGrGrGrGrCrArGrUrCrCrCrGrGrCrArCrUrUrGrGrUrUrUrUrCrCrCrCrGrGrCrUrUrArCrGrCrUrUrCrGrGrArArArArGrGrCrCrCrUrUrCrGrGrCrArCrGrUrCrGrArArGrArCrArGrGrArUrGrUrGrArGrCrCrCrArA*ml ¹⁰ ml ¹⁰ ml ¹⁰
1100	MG102-2 AAVS1 C3	ml ¹⁰ ml ¹⁰ mC ² rUrCrCrUrCrUrGrGrGrArArGrUrGrUrArArGrGrArGrUrUrUrCrArArUrCrArArCrUrGrArArArGrUrUrCrCrGrGrUrUrGrArArArGrArGrCrArUrCrCrGrUrCrCrGrGrArGrGrUrGrCrArCrUrCrCrGrGrGrArUrGrGrGrGrGrGrGrGrCrCrCrUrUrArCrGrCrUrUrCrGrGrArArArArGrGrGrCrCrCrUrUrCrGrGrCrArCrGrUrCrGrArArGrArCrArGrGrArUrGrUrGrArGrCrCrCrArA*ml ¹⁰ ml ¹⁰ ml ¹⁰
1101	MG102-2 AAVS1 E1	mC ² mC ² ml ¹⁰ rGrCrCrArGrGrArCrGrGrGrCrUrGrGrCrUrArCrUrGrUrUrUrCrArArUrCrArArCrUrGrArArArGrUrUrCrCrGrGrUrUrGrArArArGrArGrCrArUrCrCrGrUrCrCrGrGrArGrGrUrGrCrArCrUrCrCrGrGrGrArUrGrGrGrGrCrArGrUrCrCrCrGrGrCrArCrUrUrGrGrArArArArGrGrGrCrCrCrUrUrArCrGrCrUrUrCrGrGrArArArArGrGrGrCrCrCrUrUrCrGrGrCrArCrGrUrCrGrArArGrArCrArGrGrArUrGrUrGrArGrCrCrCrArA*ml ¹⁰ ml ¹⁰ ml ¹⁰

[0228]

서열번호	명칭	서열
1102	MG102-2 AAVS1 E2	mA*mA*mA*rUrUrGrGrGrGrArCrUrArGrArArGrGrUr GrArArGrGrUrUrUrCrArArUrCrArArArCrUrGrArArAr GrUrUrCrCrGrGrUrUrGrArArArGrArGrCrArUrCrCrGr UrCrCrGrGrArGrGrGrUrGrCrArCrUrCrGrGrGrArUrGr GrGrGrCrArGrUrCrCrGrGrCrArCrUrUrGrCrGrUrUrUr UrCrCrCrGrGrCrUrUrArCrGrCrUrUrCrGrArArArAr ArGrGrCrCrUrUrCrGrGrCrArCrGrUrCrGrArArGrAr CrArGrGrArUrGrUrGrArGrCrCrCrArA*mU*mU
1103	MG102-2 AAVS1 H6	mG*mG*mG*rUrGrUrGrUrCrArCrArGrArUrArGrGr ArArUrCrGrUrUrUrCrArArUrCrArArCrUrGrArArAr GrUrUrCrCrGrGrUrUrGrArArArGrArGrCrArUrCrCrGr UrCrCrGrGrArGrGrUrGrCrArCrUrCrGrGrGrArUrGr GrGrGrCrArGrUrCrCrGrGrCrArCrUrUrGrCrGrUrUrUr UrCrCrCrGrGrCrUrUrArCrGrCrUrUrCrGrArArArAr ArGrGrCrCrUrUrGrGrGrCrArCrGrUrCrGrArArGrAr CrArGrGrArUrGrUrGrArGrCrCrCrArA*mU*mU
1104	MG102-2 AAVS1 H11	mA*mG*mA*rGrGrUrGrArCrCrGrArArUrCrCrArCrAr GrGrArGrGrUrUrUrCrArArUrCrArArArCrUrGrArArAr GrUrUrCrCrGrGrUrUrGrArArArGrArGrCrArUrCrCrGr UrCrCrGrGrArGrGrGrUrGrCrArCrUrCrGrGrGrArUrGr GrGrGrCrArGrUrCrCrGrGrCrArCrUrUrCrGrGrUrUrUr UrCrCrCrGrGrCrUrUrArCrGrCrUrUrCrGrArArArAr ArGrGrCrCrUrUrCrGrGrCrArCrGrUrCrGrArArGrAr CrArGrGrArUrGrUrGrArGrCrCrCrArA*mU*mU
1105	MG102-2 AAVS1 A5	TTCTGGGAGAGGGTAGCGCAGGGT
1106	MG102-2 AAVS1 H8	GCCTTGGGAATATAAGGTGGTCCC
1107	MG102-2 AAVS1 H9	ATGCTGTCTCTGAAGTGGACATAGG
1108	MG102-2 AAVS1 D11	CTAGAGAGGTAAGGGGGGTAGGGG
1109	MG102-2 AAVS1 E7	AGGAAGGAGGAGGCCAAGGATGG
1110	MG102-2 AAVS1 D7	ATATCAGGAGACTAGGAAGGAGGA
1111	MG102-2 AAVS1 B7	CTGCCAACAGGAGGTGGGGTTA
1112	MG102-2 AAVS1 D12	GCAAGAGGATGGAGAGGTGGCTAA
1113	MG102-2 AAVS1 C8	GAGGGGACAGATAAAAGTACCCAG
1114	MG102-2 AAVS1 A8	GTGGCCCCACTGTGGGGTGGAGGG
1115	MG102-2 AAVS1 G6	TGGCTCCAGGAAATGGGGGTGTGT
1116	MG102-2 AAVS1 E5	GTTGCCACTGAGAACCAGGGCAGGT
1117	MG102-2 AAVS1 G7	TCTGTCACCAATCTGTTCCTTAGT
1118	MG102-2 AAVS1 C3	TTCCTCTTGGGAAGTGTAAAGGA
1119	MG102-2 AAVS1 E1	CCTGCCAGGACGGGGCTGGCTACT
1120	MG102-2 AAVS1 E2	AAATTGGGGACTAGAAAGGTGAAG
1121	MG102-2 AAVS1 H6	GGGTGTGTACACAGATAAGGAATC
1122	MG102-2 AAVS1 H11	AGAGGTGACCCGAATCCACAGGAG
1123	MG102-36 TRAC D12	mG*mC*mC*rArCrUrUrUrCrArGrGrArGrGrGrArUr UrCrGrGrGrUrUrUrCrArArUrCrArArArCrUrGrArArAr GrUrUrCrCrGrGrUrUrGrArArArGrArGrCrArUrCrCrGr UrCrUrGrUrCrArGrGrCrArCrUrCrCrGrGrArUrGrGrGr GrCrArGrUrCrCrGrGrCrUrUrUrGrCrGrGrUrUrArCr CrGrArUrGrGrGrCrArArCrGrUrGrUrCrGrArUrGrUrAr GrCrCrArArCrUrGrCrArGrArCrArCrGrUrCrUrUrUr GrArCrArGrGrArUrGrUrGrArGrCrCrCrArU*mU*mU
1124	MG102-36 TRAC F1	mG*mA*mC*rCrUrGrCrCrGrUrGrUrArCrArGrCrUr GrArGrArGrUrUrUrCrArArUrCrArArArCrUrGrArArAr GrUrUrCrCrGrGrUrUrGrArArArGrArGrCrArUrCrCrGr UrCrUrGrUrCrArGrGrCrArCrUrCrCrGrGrArUrGrGrGr GrCrArGrUrCrCrGrGrCrUrUrUrGrCrGrGrUrUrArCr CrGrArUrGrCrGrGrCrArArCrGrUrGrUrCrGrArUrGrUrAr

[0229]

서열번호	명칭	서열
1125	MG102-36 TRAC H6	GrCrCrArArCrUrGrCrArGrArCrArCrGrUrCrUrUrUr GrArCrArGrGrArUrGrUrGrArGrCrCrArU ¹⁰ mU ¹⁰ mU mU ¹⁰ mU ¹⁰ mG ² rArArGrUrCrCrArUrArGrArCrUrCrArU rUrCrUrGrUrUrCrArArUrCrArArArCrUrGrArArArGr UrUrCrCrGrGrUrUrGrArArArGrArGrArUrCrCrGrUr CrUrGrUrCrArGrGrCrArCrUrCrCrGrGrArUrGrGrGrGr CrArGrUrCrCrGrGrCrUrCrUrGrCrGrGrUrUrArCrCr GrArUrGrCrGrCrArArCrGrUrGrUrCrGrArUrGrUrArGr CrCrArArCrUrGrCrCrArGrArCrArCrGrUrCrUrUrUrGr ArCrArGrGrArUrGrArGrCrCrCrArU ¹⁰ mU ¹⁰ mU
1126	MG102-39 TRAC F4	mG ² mC ¹⁰ mU ¹⁰ rGrCrCrUrUrArCrCrUrGrGrCrUrGrGr GrGrArGrUrUrUrCrArGrUrUrArCrCrCrUrGrArGrAr ArUrCrArGrGrCrUrGrArArGrCrUrGrArArArGrArGrCr ArUrCrCrGrUrCrCrGrGrArArGrUrCrArCrUrCrCrGr GrGrUrArGrGrCrArGrArUrCrCrGrCrUrCrUrUrGr GrUrCrUrCrUrCrUrGrGrCrCrUrUrUrUrCrGrGrGr CrUrCrCrGrArGrArGrArGrCrUrUrCrCrGrGrCrAr UrGrUrCrUrCrGrGrArCrArGrArUrGrUrGrArGrCrCr UrUrU ¹⁰ mU ¹⁰ mU ¹⁰ mU
1127	MG102-39 TRAC A9	mU ¹⁰ mC ¹⁰ mU ¹⁰ rUrGrGrUrUrUrArCrArGrArUrArCrGrArA rCrCrUrGrUrUrCrArGrUrUrArCrCrCrUrGrArGrArAr rUrCrArGrGrCrUrGrArArGrCrUrGrArArArGrArGrCrA rUrCrCrGrUrCrCrGrGrArArGrUrCrCrArCrUrCrCrGrG rGrUrUrArGrGrCrArGrArUrCrCrGrGrCrUrCrUrGrG rUrCrCrUrCrUrCrUrGrCrCrCrUrUrUrUrCrGrGrCr rUrCrCrGrArGrArGrArGrCrUrUrCrCrGrGrCrArU rGrUrCrUrUrCrGrGrArCrArGrGrArUrGrUrGrArGrCrUr rUrU ¹⁰ mU ¹⁰ mU ¹⁰ mU
1128	MG102-39 TRAC G11	mG ² mG ² mC ¹⁰ rCrArCrUrUrCrArGrGrArGrGrArGrAr UrUrCrGrGrUrUrCrArGrUrUrArCrCrCrUrGrArGrArAr ArUrCrArGrGrCrUrGrArArGrCrUrGrArArArGrArGrCr ArUrCrCrGrUrCrCrGrGrArArGrUrCrCrArCrUrCrCrGr GrGrUrUrArGrGrCrArGrArUrCrCrGrGrCrUrCrUrUrGr GrUrCrCrUrCrUrCrUrGrGrCrCrUrUrUrUrCrGrGrGr CrUrCrCrGrArGrArGrArGrCrCrUrUrCrCrGrGrCrAr UrGrUrCrUrCrGrGrArCrArGrArUrGrUrGrArGrCrCr UrUrU ¹⁰ mU ¹⁰ mU ¹⁰ mU
1129	MG102-39 TRAC C11	mC ¹⁰ mA ¹⁰ mG ² rCrCrGrCrArGrCrGrUrCrArUrGrArGrCrAr GrArUrUrGrUrUrCrArGrUrUrArCrCrCrUrGrArGrArAr ArUrCrArGrGrCrUrGrArArGrCrUrGrArArArGrArGrCr ArUrCrCrGrUrCrCrGrGrArArGrUrCrCrArCrUrCrCrGr GrGrUrUrArGrGrCrArGrArUrCrCrGrGrCrUrCrUrUrGr GrUrCrCrUrCrUrCrUrGrGrCrCrUrUrUrUrCrGrGrGr CrUrCrCrGrArGrArGrArGrCrCrUrUrCrCrGrGrCrAr UrGrUrCrUrCrGrGrArCrArGrArUrGrUrGrArGrCrCr UrUrU ¹⁰ mU ¹⁰ mU ¹⁰ mU
1130	MG102-39 TRAC B6	mC ¹⁰ mC ¹⁰ mA ¹⁰ rGrGrCrCrArCrArGrCrArCrUrGrUrUrGrCrU rCrUrUrGrUrUrCrArGrUrUrArCrCrCrUrGrArGrArArA rUrCrArGrGrCrUrGrArArGrUrGrArArArGrArGrCrA rUrCrCrGrUrCrCrGrGrArArGrGrUrCrCrArCrUrCrCrGrG rGrUrUrArGrGrCrArGrArUrCrCrGrGrCrUrCrUrGrG rUrCrCrUrCrUrCrUrGrGrCrCrCrUrUrUrUrCrGrGrCr rUrCrCrGrArGrArGrArGrCrUrUrCrCrGrGrCrArU rGrUrCrUrUrCrGrGrArCrArGrGrArUrGrUrGrArGrCrUr rUrU ¹⁰ mU ¹⁰ mU ¹⁰ mU
1131	MG102-39 TRAC B5	mG ² mU ¹⁰ mC ¹⁰ rUrUrCrUrGrGrArUrArArUrGrCrUrGrUrU rGrUrUrGrUrUrCrArGrUrUrArCrCrCrUrGrArGrArArA rUrCrArGrGrCrUrGrArArGrCrUrGrArArArGrArGrCrA rUrCrCrGrUrCrCrGrGrArArGrGrUrCrCrArCrUrCrCrGrG rGrUrUrArGrGrGrCrArGrArUrCrCrGrGrCrUrCrUrGrG rUrCrCrUrCrUrCrUrGrGrCrCrCrUrUrUrUrCrGrGrGrC

[0230]

서열번호	명칭	서열
		rU rC rC rG rA rG rA rG rG rA rA rG rC rC rU rU rC rC rG rG rC rA rU rG rU rC rU rU rC rG rG rA rC rA rG rC rA rU rG rU rG rA rG rC rU rU rU rU mU mU mU mU
1132	MG102-39 TRAC G9	mG mA mU mU rG rG rG rU rU rU rC rC rG rA rA rU rC rU rU rC rU rC rC rU rG rU rU rU rU rC rA rG rU rU rA rC rC rC rU rG rA rG rA rA rU rC rA rG rG rC rU rG rA rG rG rC rU rG rA rA rA rG rA rG rC rA rU rC rC rG rU rC rC rG rG rA rG rU rC rC rA rC rU rC rC rG rG rG rU rU rA rG rG rC rA rG rA rU rC rC rG rG rC rU rC rU rU rG rG rU rC rU rU rU rC rU rG rG rC rC rU rU rU rU rU rC rG rG rG rC rU rC rC rG rA rG rA rG rA rA rG rC rC rU rU rC rC rG rG rC rA rG rU rC rU rU rC rG rG rA rC rA rG rA rU rG rU rG rA rG rC rU rU rU rU mU mU mU mU
1133	MG102-39 TRAC D1	mA mU mU mU rC rU rG rA rU rG rU rG rU rA rU rU rC rA rC rA rG rA rC rA rG rU rU rU rC rA rG rU rU rA rC rC rC rU rG rA rG rA rA rU rC rA rG rC rU rG rA rA rG rC rU rG rA rA rA rG rA rG rC rA rU rC rC rG rU rC rC rG rG rA rA rG rU rC rC rA rC rU rC rC rG rG rG rU rU rA rG rG rC rA rG rA rU rC rC rG rG rC rU rC rU rU rG rG rU rC rC rC rU rC rC rU rG rG rC rC rC rU rU rU rU rC rG rG rC rU rC rC rG rA rG rA rG rA rA rG rC rC rU rU rC rC rG rG rC rU rG rU rC rU rU rC rG rG rA rC rA rG rA rU rG rU rG rA rG rC rU rU rU rU mU mU mU mU
1134	MG102-39 TRAC B11	mA mU mU mU rC rC rC rC rG rC rA rG rC rG rU rC rA rU rG rA rG rC rA rG rA rU rG rU rU rU rC rA rG rU rU rA rC rC rC rU rG rA rG rA r A rU rC rA rG rC rU rG rA rA rG rC rU rG rA rA rA rG rA rG rC r A rU rC rC rG rU rC rC rG rG rA rA rG rU rC rC rA rC rU rC rC rG rG rU rU rA rG rG rC rA rG rA rU rC rC rG rC rU rC rU rU rG rG rU rC rC rU rC rU rC rU rG rG rC rC rU rU rU rU rC rG rG rC rU rC rC rG rA rG rA rG rA rA rG rC rC rU rU rC rC rG rC rA rU rG rU rC rU rC rG rG rA rC rA rG rA rU rG rU rG rA rG rC rC rU rU rU mU mU mU mU
1135	MG102-39 TRAC D4	mA mU mU mU rC rU rG rU rC rC rU rU rA rC rU rG rG rC rU rG rG rG rU rU rU rC rA rG rU rU rA rC rC rC rU rG rA rG rA rA rU rC rA rG rC rU rG rA rA rG rC rU rG rA rA rA rG rA rG rC rA rU rC rC rG rU rC rC rG rG rA rA rG rU rC rC rA rC rU rC rC rG rG rG rU rU rA rG rG rC rA rG rA rU rC rC rG rC rU rC rU rU rG rG rU rC rC rU rC rU rC rU rG rG rC rC rC rU rU rU rU rC rG rG rC rU rC rC rG rA rG rA rG rA rA rG rC rC rU rU rC rC rG rG rC rU rG rU rC rU rU rC rG rG rA rC rA rG rA rU rG rU rG rA rG rC rU rU rU rU mU mU mU mU
1136	MG102-39 TRAC F2	mC mU mU mU rC rA rG rU rG rC rU rG rU rG rG rC rU rG rG rA rG rC rA rA rG rU rU rC rA rG rU rA rC rC rC rU rG rA rG rA r A rU rC rA rG rC rU rG rA rA rG rU rG rA rC rA rA rA rG rA rC A rU rC rC rG rU rC rC rG rG rA rA rG rU rC rC rA rC rU rC rG r G rU rU rA rG rG rC rA rG rA rU rC rC rG rC rU rC rU rU rG rG rU rC rC rU rC rU rC rU rG rG rC rC rU rU rU rU rC rG rG rC rU rC rC rG rA rG rA rG rA rA rG rC rC rU rU rC rC rG rC rA rU rG rU rC rU rC rG rG rA rC rA rG rA rU rG rU rG rA rG rC rC rU rU rU mU mU mU mU
1137	MG102-39 TRAC G1	mG mU mU mU rC rA rG rA rC rU rG rA rG rG rU rC rU rA rU rG rA rC rU rU rG rU rU rC rA rG rU rA rC rA rC rC rU rG rA rA r A rU rC rA rG rG rC rU rG rA rA rG rU rG rA rA rA rG rA rC r A rU rC rC rG rU rC rC rG rG rA rA rG rU rC rA rC rU rC rG r G rG rU rU rA rG rG rC rA rG rA rU rC rC rG rC rU rC rU rU rG rU rC rC rU rC rU rC rU rG rG rC rC rU rU rU rU rC rG rG rC rU rC rC rG rA rG rA rG rA rA rG rC rC rU rU rC rC rG rC rA rU rG rU rC rU rC rG rG rA rC rA rG rA rU rG rU rG rA rG rC rC rU rU rU mU mU mU mU
1138	MG102-42 TRAC D10	mG mU mU mU rC rC rG rA rA rU rC rC rU rC rC rU rC rU rG rA rA rG rU rG rU rU rC rA rG rU rC rA rC rA rC rC rU rG rA rA rA rG rG rU rG rG rU rG rA rC rU rG rA rA rA rA rG rA rG rC rA rA rG rC rC rG rC rG rC rC rG rA rC rA rG rC rC rC rC rG rG rA rA rU rU rU rU mU mU mU mU

서열번호	명칭	서열
		rGrGrArCrArGrUUrCrCrCrGrGrCrCrCrUUrGrCrArArGrG rCrArGrCrArCrGrArGrArGrArGrCrGrUUrGrCrCrGrArArA rUUrGrGrCrGrCrCrGrGrCrUUrUArUUrGrUUrGrGrUUrGrArGrU
1139	MG102-42 TRAC D12	mG ^m C ^m C ^m *ArCrUUrUUrCrArGrArGrGrArGrArGrArU UrCrGrGrGrUUrUUrCrArGrCrArGrArCrUUrGrArArArAr GrUUrGrUUrGrArCrUUrGrArArArArGrArGrCrArArAr GrCrGrUUrGrArCrGrCrCrGrArGrCrArCrCrGrGrGrArArU GrGrGrArCrArGrUUrUUrCrCrGrGrCrCrCrUUrGrCrArArG GrCrArGrCrArCrArGrArGrArArGrCrGrUUrGrCrGrArAr ArUUrGrGrCrGrCrCrGrGrCrUUrArUUrGrUUrGrUUrGrArGr UrCrCrArUUrUArUUrU ^m U ^m U ^m U ^m
1140	MG102-42 TRAC E12	mC ^m A ^m *mG ^r GrArGrGrArGrArGrArUUrCrGrGrArArCrC CrArArUUrGrUUrUUrCrArGrCrCrArArCrUUrGrArArArAr GrUUrGrUUrGrArCrUUrGrArArArArGrArGrCrArArAr GrCrGrGrGrCrArGrCrCrArGrCrArCrCrGrGrGrArArU GrGrGrArCrArGrUUrUUrCrCrGrGrGrCrCrUUrGrCrArArG GrCrArGrCrArCrArGrArGrArArGrCrGrUUrGrCrCrGrArAr ArUUrGrGrCrGrCrCrGrGrCrUUrArUUrGrUUrGrUUrGrArGr UrCrCrArUUrUArUUrU ^m U ^m U ^m U ^m U ^m
1141	MG102-45 TRAC B1	ml ^u mG ^m U ^m *rCrArCrArCrArGrArUUrUUrCrCrArGrArArC rCrCrUUrGrUUrUUrCrArArUUrCrArArGrCrUUrGrArArArArG rUUrUUrCrGrUUrUUrGrArArArArGrArGrCrArUUrCrGrU rUUrGrArUUrArGrCrCrArUUrGrCrArCrUUrCrGrGrArArU rGrGrGrGrCrArGrUUrUUrCrCrGrGrCrUUrUUrGrCrGrArC rUUrCrArArUUrGrUUrGrUUrGrUUrGrUUrCrArUUrUUrGrArG rCrCrArArCrUUrGrUUrGrArGrArCrArGrUUrUUrUUrUUrU GrArGrArCrArGrGrArUUrGrUUrGrArGrCrCrUUrUAr ^m U [*] mU ^u mU ^u
1142	MG102-45 TRAC C11	mC ^m U ^m U ^m *rCrArArGrGrCrCrCrUUrCrArCrCrUUrCrArG rCrUUrGrGrUUrUUrCrArArUUrCrArArGrCrUUrGrArArArArG rUUrUUrCrGrGrUUrUUrGrArArArArGrArGrCrArUUrCrGrU rUUrUUrGrArUUrArGrCrArUUrGrCrArCrUUrCrGrGrArArU rGrGrGrGrCrArGrUUrUUrCrCrGrGrCrUUrUUrUUrGrCrGrArC rUUrCrArArUUrGrGrUUrGrUUrUUrGrCrUUrCrUUrUUrGrArG rCrCrArArUUrUUrUUrCrArGrArCrArGrUUrUUrUUrUUrU GrArGrArCrArGrGrArUUrGrUUrGrArGrCrCrCrUUrUAr ^m U [*] mU ^u mU ^u
1143	MG102-48 TRAC A1	ml ^u mC ^m C ^m *rUUrCrUUrGrUUrCrCrCrArCrArGrArUArUUrC rCrArGrGrUUrUUrCrArArUUrCrArArCrGrGrArArArCrG rGrUUrCrGrGrUUrGrArArArArGrArGrArUUrCrGrGrU rUUrUUrGrArArGrArUUrGrCrArUUrUUrCrGrGrGrArUUrArG rGrGrCrArGrUUrUUrCrCrGrGrCrUUrUUrUUrGrCrUUrUUrU rCrCrCrCrGrGrUUrArArGrArCrUUrCrGrGrArArGrCrArA rGrUUrCrUUrUUrCrArGrCrArGrUUrUUrGrArArGrArCrAr rCrGrArUUrUUrGrArGrCrCrUUrArU ^m U ^m U ^m U ^m U ^m
1144	MG102-48 TRAC D12	mG ^m C ^m C ^m *rArCrUUrUUrCrArUUrGrArGrGrGrArGrArU UrCrGrGrGrUUrUUrCrArUUrCrArArCrCrGrGrArArGrArU GrUUrUUrCrGrUUrUUrGrArArArArGrArGrCrArUUrCrGrG UrUUrGrArArGrGrArUUrGrGrArUUrCrCrGrGrGrArUUrAr GrGrGrCrArGrUUrUUrCrCrGrGrCrUUrUUrGrCrUUrGrUUrU UrUUrCrCrGrGrUUrArArGrArCrUUrCrGrGrArArGrCrArA rGrUUrCrUUrUUrCrArGrCrArGrUUrUUrGrArArGrArCrA rCrGrArUUrUUrGrArGrCrCrUUrArU ^m U ^m U ^m U ^m U ^m
1145	MG102-36 TRAC D12	GCCACTTTCAGGAGGAGGATTCGG
1146	MG102-36 TRAC F1	GACCCITGCCGTGTACCACTGTAGA
1147	MG102-36 TRAC H6	TTGAGCTCCATAGACTCATCTAFTGCT
1148	MG102-39 TRAC F4	CTGGCCCTTACTTGGGCTTGGGGAA
1149	MG102-39 TRAC A9	TCCTTGTTTACAGATACCAACCT

[0232]

서열번호	명칭	서열
1150	MG102-39 TRAC G11	GGCCACTTTCAGGAGGAGGATTCG
1151	MG102-39 TRAC C11	CAGCCGCAGCGTCATGAGCAGATT
1152	MG102-39 TRAC B6	CCAGGCCACAGCACTGTGTCTCTT
1153	MG102-39 TRAC B5	GCTTCTGGAATAATGCTGTGTGTT
1154	MG102-39 TRAC G9	GATTGGGTTCGGAATCCTCTCTCT
1155	MG102-39 TRAC D1	ATTCTGATGTGTATATCACAGACA
1156	MG102-39 TRAC B11	ACAGCCGCAGCGTCATGAGCAGAT
1157	MG102-39 TRAC D4	AAAGCTGCCCTTACCTGGGCTGGG
1158	MG102-39 TRAC F2	CAACAGTGTGTGGCTGGAGCAA
1159	MG102-39 TRAC G1	GCTAGACATGAGGTCTATGGACTT
1160	MG102-42 TRAC D10	GTTCCGAATCCTCTCTCTGAAAGT
1161	MG102-42 TRAC D12	GCCACTTTCAGGAGGAGGATTCGG
1162	MG102-42 TRAC E12	CAGGAGGAGGATTCGGAACCCAAT
1163	MG102-45 TRAC B1	TGTCACACAGATATCCAGAACCCT
1164	MG102-45 TRAC C11	CTCAAGGCCCTCACCTCAGCTG
1165	MG102-48 TRAC A1	TCCTCTTGTCACAGATAATCCAG
1166	MG102-48 TRAC D12	GCCACTTTCAGGAGGAGGATTCGG
1167	MG33-34 TRAC F6	mA*mC*mC*rCrGrCrCrArCrUrUrCrArGrGrArGrGrCrUrUrCrArCrUrCrUrArGrGrArArArGrCrUrArGrArGrUrGrArArGrArGrCrCrCrArGrGrCrGrCrUrGrCrUrCrCrArGrUrCrGrGrCrArCrUrArGrGrUrArGrGrCrArGrCrCrCrGrCrArGrGrCrGrGrUrArCrUrCrGrGrArCrCrCrGrGrCrArArArGrGrGrCrArArGrGrUrU*mG*mG*mG*ml
1168	MG33-34 TRAC E6	mU*mA*mA*ArCrCrCrGrGrCrArCrUrUrUrCrArGrGrCrUrUrCrArCrUrArGrCrGrArArArGrCrUrArGrGrUrUrArGrArArArGrCrCrArGrGrCrGrCrUrCrArGrUrCrUrCrGrGrCrGrArUrGrUrArCrCrArGrGrArUrCrGrGrCrArCrUrArGrGrUrArGrGrArGrGrCrCrCrGrGrCrArGrGrGrGrUrArCrUrCrGrGrArCrCrCrGrGrCrArArArGrGrGrCrArArGrGrUrU*mG*mG*mG*ml
1169	MG33-34 TRAC F6	ACCCGGCCACTTTCAGGAGG
1170	MG33-34 TRAC E6	TAAACCCGGCCACTTTCAGG

r = 고투 리보오스 염기, m = 2'-O 메틸 변형된 염기, F = 2' 플루오로 변형된 염기, * = 포스포로티오에이트 결합

실시예 16 - SMART HEARO 효소는 활성 뉴클레아제임

SMART HEARO 가이드 RNA의 인 실리코 예측

신규 SMART HEARO 뉴클레아제와 연관된 가이드(HEARO) RNA를 식별하기 위해, 305개의 추정 효과기의 5' UTR 영역에 상응하는 뉴클레오티드 서열을 추출하였다. 이들 5' UTR 뉴클레오티드 서열은, 파라미터 mafft-xinsi로 MAFFT(Katoh K, Standley DM.의 문헌[MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol.* 2013, 30(4), 772-780], 이의 전체 내용은 본원에 참조로서 통합됨)(<https://mafft.cbrc.jp/alignment/software/>)를 사용하여 정렬되었으며, 보존 영역은 HEARO RNA 경계를 기술하는 데 사용되었다(도 31). 또한, 활성 SMART HEARO 뉴클레아제의 HEARO RNA 서열을 사용하여 공분산 모델을 생성하여, 신규 SMART HEARO 뉴클레아제를 암호화하는 게놈 단편에서 추가 HEARO RNA를 예측하였다. 공분산 모델은 mafft-xinsi를 사용하는 활성 HEARO RNA 서열의 다중 서열 정렬(MSA)로부터 구축된다(<https://mafft.cbrc.jp/alignment/software/>). MSA의 이차 구조는 RNAalifold(Vienna Package, <https://www.tbi.univie.ac.at/RNA/>)로 결정하였고, 공분산 모델은 Infernal package(<http://eddylab.org/infernal/>)로 구축되었다. 후보 SMART HEARO 뉴클레아제 및 305 5' UTR 영역을 함유하는 콘티그를 Infernal 명령 'cmsearch'로 공분산 모델을 사용하여 검색하였다. 신규 후보에 대한 5' UTR 정렬 및 공분산 모델로부터 예측된 HEARO RNA를 시험관 내에서 시험하였다.

시험관 내 TAM 결정 검증

5' 말단에 표적화 스페이서를 갖는 sgRNA(HEARO RNA)를 조립 PCR을 통해 작제하고, SPRI 비드로 정제하거나 유전자 단편(IDT)으로 정렬한 다음, 짧은 RNA 전사물에 대한 제조사의 권장 프로토콜(IVT, HiScribe T7 키트, New England Biolabs)에 따라 시험관 내 전사를 수행하였다. RNA 반응물을 Monarch RNA 키트로 세척하고, Tapestation(Agilent)을 통해 순도를 확인하였다. PURExpress(New England Biolabs)로 절단 및 TAM 결정 검정을 수행하였다. 요약하면, 단백질을 대장균에 대해 코돈 최적화하고, T7 프로모터 및 C-말단 His 태그를 갖는 벡터 내로 클로닝하였다. T7 프로모터 및 종결자 서열로부터 각각 프라이머 결합 부위 150 bp 상류 및 하류에서 유전자를 PCR 증폭시켰다. 해당 PCR 산물을 PURExpress(New England Biolabs)에 5 nM의 최종 농도로 첨가하고 37°C에서 2시간 동안 발현시켰다. 절단 반응물을 PURExpress의 5배 희석물, 8 N PAM 플라스미드 라이브러리의 5 nM, 및 PAM 라이브러리를 표적화하는 50 nM의 sgRNA와 함께, 10 mM Tris pH 7.5, 100 mM NaCl, 및 10 mM MgCl₂ 중에서 조립하였다. TXTL 반응으로부터의 절단 생성물을 SPRI 비드(AMPure Beckman Coulter 또는 HighPrep

Sigma-Aldrich)로 세정하여 회수하였다. Klenow 단편 및 dNTP(New England Biolabs)의 첨가를 통해 DNA를 둔화시켰다. 무딘-말단 생성물을 100배 초과하는 이중 가닥 어댑터 서열과 연결시키고, NGS 라이브러리의 제조를 위한 템플릿으로서 사용하였으며, 이로부터의 서열 분석으로부터 PAM 요건을 결정하였다. 원시 NGS 판독치를 Phred 품질 점수 > 20으로 필터링하였다. PAM에 인접한 백본으로부터 식별된 DNA 서열을 나타내는 14-24 bp를 PAM-근위 영역을 찾기 위한 기준으로 사용하였으며, 인접한 8 bp를 추정 표적 인접 모티프(TAM)으로서 식별하였다. TAM과 연결 어댑터 사이의 거리 또한 각각의 판독치에 대해 측정하였다. 기준 서열 또는 어댑터 서열과 정확하게 일치하지 않은 판독치를 제외시켰다. 가장 빈번한 절단 부위 ± 2 bp를 갖는 TAM만이 분석에 포함되도록 TAM 서열을 절단 부위 빈도로 필터링하였다. Logomaker(Tareen, A. 및 Kinney, J. B.의 문헌[Logomaker: beautiful sequence logos in Python. *Bioinformatics* 2020, 36, 2272-2274])를 사용하여, 필터링된 TAM 목록을 사용하여 시퀀스 로고를 생성하였다..

[0239] SMART II(HEARO) 효과기는 표적화된 dsDNA 절단을 위해 이의 5' UTR 영역에 암호화된 가이드(HEARO) RNA와 상호작용하는 짧은(약 400 내지 600 aa 길이) 뉴클레아제이다(도 32a 및 32d). 대부분의 경우, SMART HEARO 시스템은 CRISPR과 연관되지 않지만, 소수의 SMART HEARO 뉴클레아제는 CRISPR과 연관될 수 있다. 예를 들어, SMART HEARO MG35-463(서열번호 530)은 CRISPR 어레이로부터 하류로 암호화된다(도 32b). 공분산 모델로부터 예측된 HEARO 가이드 RNA의 5' 말단은 해당 어레이의 최종 CRISPR 반복과 중첩되며(도 32b 및 32f, sg3), 이는 전체 표적화 단일 가이드 RNA가 해당 어레이의 최종 스페이스 및 최종 반복뿐만 아니라 HEARO RNA를 포함함을 시사한다(도 32f, sg3). 또한, 해당 후보에 대한 공분산 모델은 CRISPR 어레이로부터 상류에 있고 이와 관련이 없는 제2 HEARO RNA를 예측하였다(도 32b 및 32e, sg2). CRISPR-연관 SMART HEARO 시스템의 다른 예는 MG35-556(서열번호 659)(도 32c)이며, 여기에서 HEARO RNA는 효과기의 5' UTR 영역에서 암호화되고, 이는 CRISPR 반복 중 하나에 상보적인 항반복을 함유한다(도 32c). 이는 하나의 CRISPR 반복(이의 5' 말단에서 표적화 스페이스를 운반할 가능성이 있음)이 HEARO RNA의 5' 말단에 어닐링되어 다른 단일 가이드 HEARO RNA와 유사한 구조로 폴딩되는 이중 가이드 RNA 가이드 HEARO 시스템의 예를 제시한다(도 32g).

[0240] 절단 활성에 대한 시험의 경우, 다수의 SMART HEARO 뉴클레아제는 시험관 내 TAM 결정 검정에서 활성이었으며, 이들 중 일부는 다중 sgRNA 설계를 사용한 것이었다(도 33a-33c). MG35-104(서열번호 128), HEARO MG35-463(서열번호 530), 및 MG35-518(서열번호 621)은 강한 밴드 강도 판독으로 나타난 바와 같이, 가장 활성인 뉴클레아제였다(도 33). 또한, SMART HEARO MG35-463(서열번호 530)은, 가이드 RNA가 65%의 쌍을 이룬 뉴클레오티드 동일성만을 공유함에도 불구하고, 이의 CRISPR 연관(서열번호 1237) 및 CRISPR 독립적(서열번호 1236) HEARO RNA 둘 모두에 대해 기능적이다(도 32d, 32e, 및 도 33c). 활성 MG35 후보는 다양한 TAM을 인식하고 TAM 모티프로부터의 위치 5 또는 7에 대한 절단 선택성을 나타낸다(도 34).

[0241] 실시예 17 - SMART HEARO 효소는 효율적인 뉴클레아제임

[0242] 시험관 내 절단 검정

[0243] MG35 뉴클레아제를 시험관 내 전사/번역(IVTT)(New England Biolabs)을 사용하여 37°C에서 2시간 동안 발현시켰다. 전사는 뉴클레아제를 코딩하는 선형 DNA 템플릿 상의 T7 프로모터에 의해 유도되었다. 가이드 RNA를 시험관 내에서 별도로 전사하고, 선택된 농도, 대체적으로 0.4 내지 4 μ M로 IVTT 혼합물 중에 첨가하였다. 1 x 효과기 완충액(10 mM Tris-HCl pH 7.5, 100 mM NaCl, 10 mM MgCl₂) 또는 1 x New England Biolabs 2.1 완충액(10 mM Tris-HCl pH 7.9, 50 mM NaCl, 10 mM MgCl₂, 100 μ g/ml BSA) 중의 10 μ L 반응물 부피 중의 5 nM의 수퍼코일링된 DNA에 3 μ L의 RNP 샘플을 첨가하여 시험관 내 절단 반응을 수행하였다. 반응물을 37°C에서 1시간 동안 인큐베이션한 다음, 0.2 μ g의 RNase A(New England Biolabs)를 첨가하여 퀵칭시키고, 37°C에서 20분 동안 인큐베이션하였다. 이어서, 4 유닛의 단백질분해효소 K(New England Biolabs)를 첨가한 다음 55°C에서 30분 동안 인큐베이션하였다. 분석 및 시각화에 대해 제조사 권장 지침에 따라 D5000 Tapestation 키트(Agilent)를 사용하는 모세관 전기영동으로 반응물을 분석하였다. 성공적인 절단은 선형 dsDNA로 절단된 수퍼코일링된 2200 bp DNA를 제공한다.

[0244] 활성 가이드 RNA 및 TAM 인식 모티프를 식별한 후, SMART HEARO 뉴클레아제를 이의 가이드 RNA와 함께 뉴클레아제의 시험관 내 전사/번역 공발현을 통해 시험관 내 절단 효율에 대해 시험하고, 후속하여 가이드 RNA 및 TAM/PAM 농축 스크린에서 식별된 TAM에 의해 표적화된 스페이스를 함유하는 표적 플라스미드와 함께 인큐베이션하였다. 절단은 비절단 산물(수퍼코일링됨)으로부터 절단된 선형 DNA로의 전이에 의해 측정된다(도 35a 및 35b). 결과는 MG35-104(서열번호 128)가 다른 활성 SMART HEARO 뉴클레아제에 비해 dsDNA 절단에서 매우 효율

적임을 나타낸다(도 35a 및 35b).

[0245] 실시예 18 - SMART HEARO 가이드 조작

[0246] 일부 활성 SMART II 뉴클레아제의 가이드 RNA는 하나 이상의 폴리-T 영역을 함유하며(4개 이상의 T 염기를 순차적으로 함유함)(도 36a), 이는 전사 효율을 제한할 수 있다. 후보 당 3개의 폴리T 돌연변이체 sgRNA를 설계하고 시험관 내에서 절단 활성에 대해 시험하고, 이들의 활성을 이들의 천연 가이드 RNA를 갖는 후보의 활성과 비교하였다(도 36a 및 36b). 결과는 MG35-94가 돌연변이체 가이드 M2 및 M3과 함께 활성인 한편, MG35-104는 3개의 가이드 돌연변이 M1-M3 모두와 함께 활성임을 나타내며, 여기에서 가이드 M3은 다른 가이드와 비교하여 가장 높은 활성을 보유함을 나타낸다. MG35-518은 시험된 3개의 돌연변이체 모두에서 활성이지만, M1은 가장 높은 활성을 나타낸다(도 36b).

[표 6] 실시예 18에서 시험된 변이체 가이드 RNA

서열번호	설명	서열
1258	MG35-94_M2 단일 가이드 RNA	(N20)GUAUUCGUCCAUAAAUAACUUAAGGCAACUAAGUAGUUAAGGUUACCCGCUUUGGUUUCUUCGGAACUCCGUUAGGGGCGAAAAUAUAGGUACUUCUGGAUGCAUUCUCCAGUCCGAGACUCUACGGGGAACGAUUAACAGGUCUGAUGGAAAGGCCAGUGUGUUCUUAUUAACCGCUUUCUAAAUUAGCUAGGAAACCAUACUCGCGCAAGCGAAGAUUGUAACAAUUU
1259	MG35-94_M3 단일 가이드 RNA	(N20)GUAUUCGUCCAUAAAUAACUUAAGGCAACUAAGUAGAUUAAGGUUACCCGCUUUGGUUUCUUCGGAACUCCGUUAGGGGCGAAAAUAUAGGUACUUCUGGAUGCAUUCUCCAGUCCGAGACUCUACGGGGAACGAUUAACAGGUCUGAUGGAAAGGCCAGUGUGUUCUUAUUAACCGCUUUCUAAAUUAGCUAGGAAACCAUACUCGCGCAAGCGAAGAUUGUAACAAUUU
1249	MG35-104_M1 단일 가이드 RNA	(N20)GUAAGGAACCCCGUAGCUAAAGCUAGGGGCUAUUCAUCCCGUCCCUUCGGGCGGGCUUAGAUAGCCGAACCUUACCAGCCUAAGACCUUCGAGGUCUACGUUAUUAAGGUCACGAUACCUAUCAAUGCGUCGUAGUUGUUAAGCUCUACUCGUGGUUGUUAACAUCUGUAUUGGUUUAAGGAAGUGCAAUACAGCCCAACAAGCCUUGAAUACAUUGGCGAAGCGAACAUACCCAGCAUUGGAGUCCUUCAAUCA
1250	MG35-104_M2 단일 가이드 RNA	(N20)GUAAGGAACCCCGUAGCUAAAGCUAGGGGCUUAUCAUCCCGUCCCUUCGGGCGGGCUUAGUAAGCCGAACCUUACCAGCCUAAGACCUUCGAGGUCUACGUUAUUAAGGUCACGAUACCUAUAUUGCGUCGUAGUUGUUAAGCUCUACUCGUGGUUGUUAACAUCUGUAUUGGUUUAAGGAAGUGCAAUACAGCCCAACAAGCCUUGAAUACAUUGGCGAAGCGAACAUACCCAGCAUUGGAGUCCUUCAAUCA
1251	MG35-104_M3 단일 가이드 RNA	(N20)GUAAGGAACCCCGUAGCUAAAGCUAGGGGCUAUUCAUCCCGUCCCUUCGGGCGGGCUUAGAUAGCCGAACCUUACCAGCCUAAGACCUUCGAGGUCUACGUUCUUAAGGUCACGAUACCUAUAUUGCGUCGUAGUUGUUAAGCUCUACUCGUGGUUGUUAACAUCUGUAUUGGUUUAAGGAAGUGCAAUACAGCCCAACAAGCCUUGAGAACAUUGGCGAAGCGAACAUACCCAGCAUUGGAGUCCUUCAAUCA
1252	MG35-518_M1 단일 가이드 RNA	(N20)AUCAAUAACCAACCCACUAAGUGGGCGGAUUGCUGACUCUUAUACAAUGAGUUGAGAAACCGUGAUUGAUUAGCCUCAGUUAUAAACUACGUUAUUGUAAAUUAUAGGUACCGUGGAUGUCCGCCUAGUCCUACGCGCUACGCUUUAUUAUAAACAGUUCUGAUUGGUAGGAAACAGUGUAAUAAAGAUUAUAAACUACAAGAUAAAUUGGCGAAGGCAUAAAGGGUUGUUAUACCCGCUUACCGCAUUAUUAUAAACAU
1253	MG35-518_M2 단일 가이드 RNA	(N20)AUCAAUAACCAACCCACUAAGUGGGCGGAUUGCUGACUCUUAUACAAUGAGUUGAGAAACCGUGAUUGAUUAGCCUCAGUUAUAAACUACGUUAUUGUAAAUUAUAGGUACCGUGGAUGUCCGCCUAGUCCUACGCGCUACGCUUUAUUAUAAACAGUUCUGAUUGGUAGGAAACAGUGUAAUAAAGAUUAUAAACUACAAGAUAAAUUGGCGAAGGCAUAAAGGGUUGUUAUACCCGCUUACCGCAUUAUUAUAAACAU
1254	MG35-518_M3 단일 가이드 RNA	(N20)AUCAAUAACCAACCCACUAAGUGGGCGGAUUGCUGACUCUGUUAUAAUGAGUUGAGAAACCGUGAUUGAUUAGCCUCAGUUAUAAACUACGUUAUUGUAAAUUAUAGGUACCGUGGAUGUCCGCCUAGUCCUACGCGCUACGCUUUAUUAUAAACAGUUCUGAUUGGUAGGAA

[0247]

서열번호	설명	서열
		CAGUGUAAUAAAAGAUUAAAACUACAAGAUAAACAUUG GCGAAGGCAUAAAGGGUUGUUAUACCCGCUUAC CGCAUAAAUAACAU
1255	MG35-553_M1 단일 가이드 RNA	(N20)GUCACUACCCACGACUAAAGUCGCGGGCUUG AAUUAAGGAUAGUGCUAUGUACUAGCCUUAUUCAGCC CGGUUGACUAGCCUAAAGACCAAUUGUGCUACGUUA UGCAGGAAAUAGGUACCCGGAUGUACAGCCUAGU CCCGGGCUCUACGGUAGAGGUUAAACAGCUCUGAC GGUAGGAGCAGUGCUUCAUGCGUUAACCCUGCAA UAAAUUGGCGAAGGCUAACUACGGAUGCGCAUC CGGCUUACAGCAUAAUUGCAGCAGAAAA
1256	MG35-553_M2 단일 가이드 RNA	(N20)GUCACUACCCACGACUAAAGUCGCGGGCUUG UAAUUAAGGAUAGUGCUAUGUACUAGCCUUAUUCAGC CGGUUGACUAGCCUAAAGACCAAUUGUGCUACGUUA AUGCAGGAAAUAGGUACCCGGAUGUACAGCCUAG UCCCGGGCUCUACGGUAGAGGUUAAACAGCUCUGA CGGUAGGAGCAGUGCUUCAUGCGUUAACCCUGCA AUAAAUUGGCGAAGGCUAACUACGGAUGCGCAU CCGGCUUACAGCAUAAUUGCAGCAGAAAA
1257	MG35-553_M3 단일 가이드 RNA	(N20)GUCACUACCCACGACUAAAGUCGCGGGCUUG UAAUUAAGGAUAGUGCUAUGUACUAGCCUUAUUCAGC CGGUUGACUAGCCUAAAGACCAAUUGUGCUACGUUA AUGCAGGAAAUAGGUACCCGGAUGUACAGCCUAG UCCCGGGCUCUACGGUAGAGGUUAAACAGCUCUGA CGGUAGGAGCAGUGCUUCAUGCGUUAACCCUGCA AUAAAUUGGCGAAGGCUAACUACGGAUGCGCAU CCGGCUUACAGCAUAAUUGCAGCAGAAAA

실시예 19 - 신규 SMART I 뉴클레아제의 연산 재구성

신규 서열의 인 실리코 재구성

SMART I 뉴클레아제의 추가적인 다양성을 생성하기 위한 노력의 일환으로, 조상 서열 재구성 알고리즘을 사용하여 이중 뉴클레아제 서열을 재구성하였다. 조상 서열 재구성은 기존의 단백질 서열 및 이들 사이에서 추론된 관계를 사용하여 현재 소멸된, 고대의 단백질의 서열을 재구성하는 연산 기술이다(Harms, M. 및 Thornton J.W.의 문헌[Analyzing protein structure and function using ancestral gene reconstruction. *Current Opinion in Structural Biology* 2010, 20, 360-366]). 이 기술은 SMART I MG34 계열의 신규 서열을 연산적으로 재구성하는데 사용되었다. 해당 분석을 위해, L-INS-i 또는 G-INS-i 매개변수를 사용하는MAFFT(Katoh K, Standley DM.의 문헌[MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol.* 2013, 30(4), 772-780])를 사용하여 190개의 SMART I 단백질 서열을 정렬하고, Fasttree(Price, M.N., Dehal, P.S., 및 Arkin, A.P.의 문헌[FastTree 2 -- Approximately Maximum-Likelihood Trees for Large Alignments. *PLoS ONE* 2010, 5(3), e9490]) 또는 RAXML(Stamatakis, A.의 문헌[RAXML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 2014, 30(9), 1312-1313])을 사용하여 계통유전학 트리를 구축하였다(도 37). 해당 트리는 SpCas9 및 SaCas9를 사용하여 루팅되었다. 서열 재구성은 PAML 4.8의 codeml 패키지(Yang, Z.의 문헌[PAML 4: a program package for phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution* 2007, 24, 1586-1591])를 사용하여 수행되었으며, 이는 계통유전학의 불확실성을 설명하기 위해, 정렬 및 트리 구축 방법의 4가지 조합 모두에 적용되었다. 각각의 재구성된 노드에 대해 삽입 및 결실을 수작업으로 식별하였다.

시험관 내 PAM 결정 검증

후보 단백질을 대장균에 대해 코돈 최적화하고, T7 프로모터 및 C-말단 His 태그를 갖는 벡터 내로 클로닝하였다. T7 프로모터 및 종결자 서열로부터 각각 프라이머 결합 부위 150 bp 상류 및 하류에서 유전자를 PCR 증폭시켰다. 해당 PCR 산물을 PURExpress(New England Biolabs)에 5 nM의 최종 농도로 첨가하고 37°C에서 2시간 동안 발현시켰다. 절단 반응물을 PURExpress의 5배 희석물, 8 N PAM 플라스미드 라이브러리의 5 nM, 및 PAM 라이브러리를 표적화하는 50 nM의 sgRNA와 함께, 10 mM Tris pH 7.5, 100 mM NaCl, 및 10 mM MgCl₂ 중에서 조립하였다. TXTL 반응으로부터의 절단 생성물을 SPRI 비드(AMPure Beckman Coulter 또는 HighPrep Sigma-Aldrich)로 세정하여 회수하였다. Klenow 단편 및 dNTP(New England Biolabs)의 첨가를 통해 DNA를 둔화시켰다. 무딘-말단 생성물을 100배 초과하는 이중 가닥 어댑터 서열과 연결시키고, NGS 라이브러리의 제조를 위한 템플릿으로서 사용하였으며, 이로부터의 서열 분석으로부터 PAM 요건을 결정하였다. 원시 NGS 판독치를 Phred 품질 점수 > 20으로 필터링하였다. PAM에 인접한 백본으로부터의 보고된 DNA 서열을 나타내는 14-24 bp를 PAM-근위 영역을 찾기 위한 기준으로서 사용하였으며, 인접한 8 bp를 추정 PAM으로서 식별하였다. PAM과 연결 어댑터 사이의 거리 또한 각각의 리드에 대해 측정하였다. 기준 서열 또는 어댑터 서열과 정확하게 일치하지 않은 판독치를 제외시켰다.

가장 빈번한 절단 부위 ± 2 bp를 갖는 PAM만이 분석에 포함되도록 PAM 서열을 절단 부위 빈도로 필터링하였다. Logomaker(Tareen, A. 및 Kinney, J. B.의 문헌[Logomaker: beautiful sequence logos in Python. *Bioinformatics* 2020, 36, 2272-2274])를 사용하여, 필터링된 PAM 목록을 사용하여 시퀀스 로고를 생성하였다..

[0254] MG34 계열의 6개의 서열을 높은 신뢰도 기준으로 재구성하고(표 5 및 도 37), 다중 서열 정렬 및 3D 구조 예측으로부터 촉매 및 결합 도메인을 확인하였다(도 38).

[표 7] SMART I 뉴클레아제 MG34-1 에 대한 연산식 유도 MG34 후보의 비교

후보	서열번호	평균 지지	MG34-1 동일성(%)	길이 (아미노산)
MG34-26	1313	0.73	66	768
MG34-27	1314	0.93	78	745
MG34-28	1315	0.75	70	765
MG34-29	1316	0.92	79	748
MG34-30	1317	0.74	79	766
MG34-31	1318	0.74	72	768

평균 지지 값은 재구성된 서열에 대한 평균 확률을 0 내지 1 의 척도로 나타냄.
> 0.7 인 지지 값은 재구성된 서열에서의 높은 신뢰성을 나타냄.

[0255]

[0256] 구조체 간의 주요 차이는 인식 로브에 있으며, 이는 이들 재구성된 효소가 MG34-1과 유사한 뉴클레아제 활성을 나타낼 수 있음을 시사한다. 신규로 재구성된 후보에 대한 강력한 지지를 고려하여, 다음과 같은 3개의 활성 MG34 뉴클레아제로부터의 가이드 RNA를 사용하는 PAM 농축 검정에서 6개의 신규 뉴클레아제를 시험관 내 절단 활성에 대해 시험하였다: MG34-1 sgRNA 1(서열번호 613), MG34-9 sgRNA 1(서열번호 615), 및 MG34-16 sgRNA 1(서열번호 616). 신규 뉴클레아제 MG34-27(서열번호 1314) 및 MG34-29(서열번호 1316)는 대략 180 bp에서의 예상된 절단 밴드에 의해 나타낸 바와 같이, 3개의 시험된 sgRNA 모두에 대해 활성이었다(도 39). 이들 신규 뉴클레아제에 의해 표적화된 PAM은 3' nRR일 가능성이 높으며, 여기에서 nGG는 가장 흔히 인식되는 PAM이다(도 40). 해당 결과는 신규로 재구성된 뉴클레아제가 다른 활성 MG34 뉴클레아제 대비 보다 이완된 PAM 인식을 가지며(예를 들어, MG34-1은 3' nGG PAM을 인식함), PAM으로부터 위치 6-9에서의 유연적인 절단 선호도를 가짐을 나타낸다(도 40).

[0257]

[표 8] 서열 목록에 포함되지 않은, 본원에서 참고되는 추가 단백질 및 핵산 서열의 목록

1.

번호	서열번호	설명	유형	유기체	기타 정보	표본 또는 서열
MG33 활성 효과기	1	MG33-1 효과기	단백질	알려지지 않음	미배양 유기체	
MG34 활성 효과기	2	MG34-1 효과기	단백질	알려지지 않음	미배양 유기체	
MG34 효과기	3	MG34-2 효과기	단백질	알려지지 않음	미배양 유기체	
MG34 효과기	4	MG34-3 효과기	단백질	알려지지 않음	미배양 유기체	
MG34 효과기	5	MG34-4 효과기	단백질	알려지지 않음	미배양 유기체	
MG34 효과기	6	MG34-5 효과기	단백질	알려지지 않음	미배양 유기체	
MG34 효과기	7	MG34-6 효과기	단백질	알려지지 않음	미배양 유기체	
MG34 효과기	8	MG34-7 효과기	단백질	알려지지 않음	미배양 유기체	
MG34 효과기	9	MG34-8 효과기	단백질	알려지지 않음	미배양 유기체	
MG34 효과기	10	MG34-9 효과기	단백질	알려지지 않음	미배양 유기체	
MG34 효과기	11	MG34-10 효과기	단백질	알려지지 않음	미배양 유기체	
MG34 효과기	12	MG34-11 효과기	단백질	알려지지 않음	미배양 유기체	
MG34 효과기	13	MG34-12 효과기	단백질	알려지지 않음	미배양 유기체	
MG34 효과기	14	MG34-13 효과기	단백질	알려지지 않음	미배양 유기체	
MG34 효과기	15	MG34-14 효과기	단백질	알려지지 않음	미배양 유기체	
MG34 효과기	16	MG34-15 효과기	단백질	알려지지 않음	미배양 유기체	
MG34 효과기	17	MG34-16 효과기	단백질	알려지지 않음	미배양 유기체	
MG34 효과기	18	MG34-17 효과기	단백질	알려지지 않음	미배양 유기체	
MG34 효과기	19	MG34-18 효과기	단백질	알려지지 않음	미배양 유기체	
MG34 효과기	20	MG34-19 효과기	단백질	알려지지 않음	미배양 유기체	
MG34 효과기	21	MG34-20 효과기	단백질	알려지지 않음	미배양 유기체	
MG34 효과기	22	MG34-21 효과기	단백질	알려지지 않음	미배양 유기체	
MG34 효과기	23	MG34-22 효과기	단백질	알려지지 않음	미배양 유기체	
MG34 효과기	24	MG34-23 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	25	MG35-1 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	26	MG35-2 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	27	MG35-3 효과기	단백질	알려지지 않음	미배양 유기체	

[0258]

별주	서열번호	설명	유형	유기체	기타 정보	로벨트 또는 서열
MG35 효과기	28	MG35-4 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	29	MG35-5 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	30	MG35-6 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	31	MG35-7 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	32	MG35-8 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	33	MG35-9 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	34	MG35-10 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	35	MG35-11 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	36	MG35-12 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	37	MG35-13 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	38	MG35-14 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	39	MG35-15 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	40	MG35-16 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	41	MG35-17 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	42	MG35-18 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	43	MG35-19 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	44	MG35-20 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	45	MG35-21 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	46	MG35-22 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	47	MG35-23 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	48	MG35-24 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	49	MG35-25 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	50	MG35-26 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	51	MG35-27 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	52	MG35-28 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	53	MG35-29 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	54	MG35-30 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	55	MG35-31 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	56	MG35-32 효과기	단백질	알려지지 않음	미배양 유기체	

번호	서열번호	설명	유형	유기체	기타 정보	모델트 또는 서열
MG35 효과기	57	MG35-33 효과기	단백질	알러지치 않음	미배양 유기체	
MG35 효과기	58	MG35-34 효과기	단백질	알러지치 않음	미배양 유기체	
MG35 효과기	59	MG35-35 효과기	단백질	알러지치 않음	미배양 유기체	
MG35 효과기	60	MG35-36 효과기	단백질	알러지치 않음	미배양 유기체	
MG35 효과기	61	MG35-37 효과기	단백질	알러지치 않음	미배양 유기체	
MG35 효과기	62	MG35-38 효과기	단백질	알러지치 않음	미배양 유기체	
MG35 효과기	63	MG35-39 효과기	단백질	알러지치 않음	미배양 유기체	
MG35 효과기	64	MG35-40 효과기	단백질	알러지치 않음	미배양 유기체	
MG35 효과기	65	MG35-41 효과기	단백질	알러지치 않음	미배양 유기체	
MG35 효과기	66	MG35-42 효과기	단백질	알러지치 않음	미배양 유기체	
MG35 효과기	67	MG35-43 효과기	단백질	알러지치 않음	미배양 유기체	
MG35 효과기	68	MG35-44 효과기	단백질	알러지치 않음	미배양 유기체	
MG35 효과기	69	MG35-45 효과기	단백질	알러지치 않음	미배양 유기체	
MG35 효과기	70	MG35-46 효과기	단백질	알러지치 않음	미배양 유기체	
MG35 효과기	71	MG35-47 효과기	단백질	알러지치 않음	미배양 유기체	
MG35 효과기	72	MG35-48 효과기	단백질	알러지치 않음	미배양 유기체	
MG35 효과기	73	MG35-49 효과기	단백질	알러지치 않음	미배양 유기체	
MG35 효과기	74	MG35-50 효과기	단백질	알러지치 않음	미배양 유기체	
MG35 효과기	75	MG35-51 효과기	단백질	알러지치 않음	미배양 유기체	
MG35 효과기	76	MG35-52 효과기	단백질	알러지치 않음	미배양 유기체	
MG35 효과기	77	MG35-53 효과기	단백질	알러지치 않음	미배양 유기체	
MG35 효과기	78	MG35-54 효과기	단백질	알러지치 않음	미배양 유기체	
MG35 효과기	79	MG35-55 효과기	단백질	알러지치 않음	미배양 유기체	
MG35 효과기	80	MG35-56 효과기	단백질	알러지치 않음	미배양 유기체	
MG35 효과기	81	MG35-57 효과기	단백질	알러지치 않음	미배양 유기체	
MG35 효과기	82	MG35-58 효과기	단백질	알러지치 않음	미배양 유기체	
MG35 효과기	83	MG35-59 효과기	단백질	알러지치 않음	미배양 유기체	
MG35 효과기	84	MG35-60 효과기	단백질	알러지치 않음	미배양 유기체	
MG35 효과기	85	MG35-61 효과기	단백질	알러지치 않음	미배양 유기체	

[0259]

번호	서열번호	설명	유형	유기체	기타 정보	코멘트 또는 서열
MG35 효과기	86	MG35-62 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	87	MG35-63 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	88	MG35-64 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	89	MG35-65 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	90	MG35-66 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	91	MG35-67 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	92	MG35-68 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	93	MG35-69 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	94	MG35-70 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	95	MG35-71 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	96	MG35-72 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	97	MG35-73 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	98	MG35-74 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	99	MG35-75 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	100	MG35-76 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	101	MG35-77 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	102	MG35-78 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	103	MG35-79 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	104	MG35-80 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	105	MG35-81 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	106	MG35-82 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	107	MG35-83 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	108	MG35-84 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	109	MG35-85 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	110	MG35-86 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	111	MG35-87 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	112	MG35-88 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	113	MG35-89 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	114	MG35-90 효과기	단백질	알려지지 않음	미배양 유기체	

별주	서열번호	설명	유형	유기체	기타 정보	코멘트 또는 서열
MG35 효과기	115	MG35-91 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	116	MG35-92 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	117	MG35-93 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	118	MG35-94 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	119	MG35-95 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	120	MG35-96 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	121	MG35-97 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	122	MG35-98 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	123	MG35-99 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	124	MG35-100 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	125	MG35-101 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	126	MG35-102 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	127	MG35-103 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	128	MG35-104 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	129	MG35-105 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	130	MG35-106 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	131	MG35-107 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	132	MG35-108 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	133	MG35-109 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	134	MG35-110 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	135	MG35-111 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	136	MG35-112 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	137	MG35-113 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	138	MG35-114 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	139	MG35-115 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	140	MG35-116 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	141	MG35-117 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	142	MG35-118 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	143	MG35-119 효과기	단백질	알려지지 않음	미배양 유기체	

번호	서열번호	설명	유형	유기체	기타 정보	코멘트 또는 서열
MG35 효과기	144	MG35-120 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	145	MG35-121 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	146	MG35-122 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	147	MG35-123 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	148	MG35-124 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	149	MG35-125 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	150	MG35-126 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	151	MG35-127 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	152	MG35-128 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	153	MG35-129 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	154	MG35-130 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	155	MG35-131 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	156	MG35-132 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	157	MG35-133 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	158	MG35-134 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	159	MG35-135 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	160	MG35-136 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	161	MG35-137 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	162	MG35-138 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	163	MG35-139 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	164	MG35-140 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	165	MG35-141 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	166	MG35-142 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	167	MG35-143 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	168	MG35-144 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	169	MG35-146 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	170	MG35-147 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	171	MG35-148 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	172	MG35-149 효과기	단백질	알려지지 않음	미배양 유기체	

별주	서열번호	설명	유형	유기체	기타 정보	코멘트 또는 서열
MG35 효과기	173	MG35-150 효과기	단백질	인러지지 않음	미배양 유기체	
MG35 효과기	174	MG35-151 효과기	단백질	인러지지 않음	미배양 유기체	
MG35 효과기	175	MG35-152 효과기	단백질	인러지지 않음	미배양 유기체	
MG35 효과기	176	MG35-153 효과기	단백질	인러지지 않음	미배양 유기체	
MG35 효과기	177	MG35-154 효과기	단백질	인러지지 않음	미배양 유기체	
MG35 효과기	178	MG35-155 효과기	단백질	인러지지 않음	미배양 유기체	
MG35 효과기	179	MG35-156 효과기	단백질	인러지지 않음	미배양 유기체	
MG35 효과기	180	MG35-157 효과기	단백질	인러지지 않음	미배양 유기체	
MG35 효과기	181	MG35-158 효과기	단백질	인러지지 않음	미배양 유기체	
MG35 효과기	182	MG35-159 효과기	단백질	인러지지 않음	미배양 유기체	
MG35 효과기	183	MG35-160 효과기	단백질	인러지지 않음	미배양 유기체	
MG35 효과기	184	MG35-161 효과기	단백질	인러지지 않음	미배양 유기체	
MG35 효과기	185	MG35-162 효과기	단백질	인러지지 않음	미배양 유기체	
MG35 효과기	186	MG35-163 효과기	단백질	인러지지 않음	미배양 유기체	
MG35 효과기	187	MG35-164 효과기	단백질	인러지지 않음	미배양 유기체	
MG35 효과기	188	MG35-165 효과기	단백질	인러지지 않음	미배양 유기체	
MG35 효과기	189	MG35-166 효과기	단백질	인러지지 않음	미배양 유기체	
MG35 효과기	190	MG35-167 효과기	단백질	인러지지 않음	미배양 유기체	
MG35 효과기	191	MG35-168 효과기	단백질	인러지지 않음	미배양 유기체	
MG35 효과기	192	MG35-169 효과기	단백질	인러지지 않음	미배양 유기체	
MG35 효과기	193	MG35-170 효과기	단백질	인러지지 않음	미배양 유기체	
MG35 효과기	194	MG35-171 효과기	단백질	인러지지 않음	미배양 유기체	
MG35 효과기	195	MG35-172 효과기	단백질	인러지지 않음	미배양 유기체	
MG35 효과기	196	MG35-173 효과기	단백질	인러지지 않음	미배양 유기체	
MG35 효과기	197	MG35-174 효과기	단백질	인러지지 않음	미배양 유기체	
MG35 효과기	198	MG35-175 효과기	단백질	인러지지 않음	미배양 유기체	
MG33 활성 효과기 tracrRNA 서열	199	MG33-1 tracrRNA 서열	뉴클레오티드	인공 서열		
MG34 활성 효과기 tracrRNA 서열	200	MG34-1 tracrRNA 서열	뉴클레오티드	인공 서열		

별주	서열번호	설명	유형	유기체	기타 정보	코멘트 또는 서열
추정 sgRNA	201	추정 MG33-1 sgRNA	뉴클레오타이드	인공 서열		
추정 sgRNA	202	추정 MG34-1 sgRNA	뉴클레오타이드	인공 서열		
추정 sgRNA	203	추정 MG34-1 sgRNA	뉴클레오타이드	인공 서열		
표지	204	시험 표지 서열	뉴클레오타이드	인공 서열		
NLS	205	SV40 NLS	단백질	원숭이 비콜로화 바이러스 40 T		
NLS	206	뉴클레오타이드 인공 서열	단백질	인간		
NLS	207	e-myc NLS	단백질	인간		
NLS	208	e-myc NLS	단백질	인간		
NLS	209	hnRNP A1 M9 NLS	단백질	마우스		
NLS	210	임프린트-알파 B2B 도메인 NLS	단백질	인간		
NLS	211	Myoma T 단백질 NLS	단백질	쥐과 폴리오마바이러스		
NLS	212	Myoma T 단백질 NLS	단백질	쥐과 폴리오마바이러스		
NLS	213	p53 NLS	단백질	인간		
NLS	214	마우스 c-abl IV NLS	단백질	마우스		
NLS	215	인플루엔자 바이러스 NS1 NLS	단백질	인플루엔자 바이러스 NS1		
NLS	216	인플루엔자 바이러스 NS1 NLS	단백질	인플루엔자 바이러스 NS1		
NLS	217	간염 바이러스 델타 항원 NLS	단백질	간염 바이러스 델타		
NLS	218	마우스 Mx1 단백질 NLS	단백질	마우스		
NLS	219	인간 폴리(ADP-리보솜) 분해 효소 NLS	단백질	인간		
NLS	220	스테인리스 스틸 수용체 단백질 코팅 폴리(ADP-리보솜) NLS	단백질	인간		
MG35 효과기	221	MG35-1 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	222	MG35-119 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	223	MG35-120 효과기	단백질	알려지지 않음	미배양 유기체	

[0264]

번호	서열번호	설명	유형	유기체	기타 정보	코멘트 또는 서열
MG35 효과기	224	MG35-421 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	225	MG35-176 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	226	MG35-177 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	227	MG35-178 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	228	MG35-179 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	229	MG35-180 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	230	MG35-181 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	231	MG35-183 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	232	MG35-184 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	233	MG35-185 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	234	MG35-186 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	235	MG35-187 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	236	MG35-188 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	237	MG35-189 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	238	MG35-190 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	239	MG35-191 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	240	MG35-192 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	241	MG35-193 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	242	MG35-194 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	243	MG35-195 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	244	MG35-196 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	245	MG35-197 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	246	MG35-198 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	247	MG35-199 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	248	MG35-200 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	249	MG35-201 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	250	MG35-202 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	251	MG35-203 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	252	MG35-204 효과기	단백질	알려지지 않음	미배양 유기체	

[0265]

별주	서열번호	설명	유형	유기체	기타 정보	코멘트 또는 서열
MG35 효파기	253	MG35-205 효파기	단백질	알려지지 않음	미배양 유기체	
MG35 효파기	254	MG35-206 효파기	단백질	알려지지 않음	미배양 유기체	
MG35 효파기	255	MG35-207 효파기	단백질	알려지지 않음	미배양 유기체	
MG35 효파기	256	MG35-208 효파기	단백질	알려지지 않음	미배양 유기체	
MG35 효파기	257	MG35-209 효파기	단백질	알려지지 않음	미배양 유기체	
MG35 효파기	258	MG35-210 효파기	단백질	알려지지 않음	미배양 유기체	
MG35 효파기	259	MG35-211 효파기	단백질	알려지지 않음	미배양 유기체	
MG35 효파기	260	MG35-212 효파기	단백질	알려지지 않음	미배양 유기체	
MG35 효파기	261	MG35-213 효파기	단백질	알려지지 않음	미배양 유기체	
MG35 효파기	262	MG35-214 효파기	단백질	알려지지 않음	미배양 유기체	
MG35 효파기	263	MG35-215 효파기	단백질	알려지지 않음	미배양 유기체	
MG35 효파기	264	MG35-216 효파기	단백질	알려지지 않음	미배양 유기체	
MG35 효파기	265	MG35-217 효파기	단백질	알려지지 않음	미배양 유기체	
MG35 효파기	266	MG35-218 효파기	단백질	알려지지 않음	미배양 유기체	
MG35 효파기	267	MG35-219 효파기	단백질	알려지지 않음	미배양 유기체	
MG35 효파기	268	MG35-220 효파기	단백질	알려지지 않음	미배양 유기체	
MG35 효파기	269	MG35-221 효파기	단백질	알려지지 않음	미배양 유기체	
MG35 효파기	270	MG35-222 효파기	단백질	알려지지 않음	미배양 유기체	
MG35 효파기	271	MG35-223 효파기	단백질	알려지지 않음	미배양 유기체	
MG35 효파기	272	MG35-224 효파기	단백질	알려지지 않음	미배양 유기체	
MG35 효파기	273	MG35-225 효파기	단백질	알려지지 않음	미배양 유기체	
MG35 효파기	274	MG35-226 효파기	단백질	알려지지 않음	미배양 유기체	
MG35 효파기	275	MG35-227 효파기	단백질	알려지지 않음	미배양 유기체	
MG35 효파기	276	MG35-228 효파기	단백질	알려지지 않음	미배양 유기체	
MG35 효파기	277	MG35-229 효파기	단백질	알려지지 않음	미배양 유기체	
MG35 효파기	278	MG35-230 효파기	단백질	알려지지 않음	미배양 유기체	
MG35 효파기	279	MG35-231 효파기	단백질	알려지지 않음	미배양 유기체	
MG35 효파기	280	MG35-232 효파기	단백질	알려지지 않음	미배양 유기체	
MG35 효파기	281	MG35-233 효파기	단백질	알려지지 않음	미배양 유기체	

번호	서열번호	설명	유형	유기체	기타 정보	코멘트 또는 서열
MG35 효과기	282	MG35-234 효과기	단백질	알러지지 않음	미배양 유기체	
MG35 효과기	283	MG35-235 효과기	단백질	알러지지 않음	미배양 유기체	
MG35 효과기	284	MG35-236 효과기	단백질	알러지지 않음	미배양 유기체	
MG35 효과기	285	MG35-237 효과기	단백질	알러지지 않음	미배양 유기체	
MG35 효과기	286	MG35-238 효과기	단백질	알러지지 않음	미배양 유기체	
MG35 효과기	287	MG35-239 효과기	단백질	알러지지 않음	미배양 유기체	
MG35 효과기	288	MG35-240 효과기	단백질	알러지지 않음	미배양 유기체	
MG35 효과기	289	MG35-241 효과기	단백질	알러지지 않음	미배양 유기체	
MG35 효과기	290	MG35-242 효과기	단백질	알러지지 않음	미배양 유기체	
MG35 효과기	291	MG35-243 효과기	단백질	알러지지 않음	미배양 유기체	
MG35 효과기	292	MG35-244 효과기	단백질	알러지지 않음	미배양 유기체	
MG35 효과기	293	MG35-245 효과기	단백질	알러지지 않음	미배양 유기체	
MG35 효과기	294	MG35-246 효과기	단백질	알러지지 않음	미배양 유기체	
MG35 효과기	295	MG35-247 효과기	단백질	알러지지 않음	미배양 유기체	
MG35 효과기	296	MG35-248 효과기	단백질	알러지지 않음	미배양 유기체	
MG35 효과기	297	MG35-249 효과기	단백질	알러지지 않음	미배양 유기체	
MG35 효과기	298	MG35-250 효과기	단백질	알러지지 않음	미배양 유기체	
MG35 효과기	299	MG35-251 효과기	단백질	알러지지 않음	미배양 유기체	
MG35 효과기	300	MG35-252 효과기	단백질	알러지지 않음	미배양 유기체	
MG35 효과기	301	MG35-253 효과기	단백질	알러지지 않음	미배양 유기체	
MG35 효과기	302	MG35-254 효과기	단백질	알러지지 않음	미배양 유기체	
MG35 효과기	303	MG35-255 효과기	단백질	알러지지 않음	미배양 유기체	
MG35 효과기	304	MG35-256 효과기	단백질	알러지지 않음	미배양 유기체	
MG35 효과기	305	MG35-257 효과기	단백질	알러지지 않음	미배양 유기체	
MG35 효과기	306	MG35-258 효과기	단백질	알러지지 않음	미배양 유기체	
MG35 효과기	307	MG35-259 효과기	단백질	알러지지 않음	미배양 유기체	
MG35 효과기	308	MG35-260 효과기	단백질	알러지지 않음	미배양 유기체	
MG35 효과기	309	MG35-261 효과기	단백질	알러지지 않음	미배양 유기체	
MG35 효과기	310	MG35-262 효과기	단백질	알러지지 않음	미배양 유기체	

번호	서열번호	설명	유형	유기체	기타 정보	코멘트 또는 서열
MG35 효과기	311	MG35-263 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	312	MG35-264 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	313	MG35-265 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	314	MG35-266 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	315	MG35-267 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	316	MG35-268 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	317	MG35-269 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	318	MG35-270 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	319	MG35-271 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	320	MG35-272 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	321	MG35-273 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	322	MG35-274 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	323	MG35-275 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	324	MG35-276 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	325	MG35-277 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	326	MG35-278 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	327	MG35-279 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	328	MG35-280 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	329	MG35-281 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	330	MG35-282 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	331	MG35-283 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	332	MG35-284 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	333	MG35-285 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	334	MG35-286 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	335	MG35-287 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	336	MG35-288 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	337	MG35-289 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	338	MG35-290 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	339	MG35-291 효과기	단백질	알려지지 않음	미배양 유기체	

별주	서열번호	설명	유형	유기체	기타 정보	표본 또는 서열
MG35 효과기	340	MG35-292 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	341	MG35-293 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	342	MG35-294 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	343	MG35-295 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	344	MG35-296 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	345	MG35-297 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	346	MG35-298 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	347	MG35-299 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	348	MG35-300 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	349	MG35-301 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	350	MG35-302 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	351	MG35-303 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	352	MG35-304 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	353	MG35-305 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	354	MG35-307 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	355	MG35-308 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	356	MG35-309 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	357	MG35-310 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	358	MG35-311 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	359	MG35-312 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	360	MG35-313 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	361	MG35-314 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	362	MG35-315 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	363	MG35-316 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	364	MG35-317 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	365	MG35-318 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	366	MG35-319 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	367	MG35-320 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	368	MG35-321 효과기	단백질	알려지지 않음	미배양 유기체	

별주	서열번호	설명	유형	유기체	기타 정보	포렌트 또는 서열
MG35 효과기	369	MG35-322 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	370	MG35-323 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	371	MG35-324 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	372	MG35-325 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	373	MG35-326 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	374	MG35-327 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	375	MG35-328 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	376	MG35-329 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	377	MG35-330 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	378	MG35-331 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	379	MG35-333 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	380	MG35-334 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	381	MG35-335 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	382	MG35-336 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	383	MG35-337 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	384	MG35-338 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	385	MG35-339 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	386	MG35-340 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	387	MG35-341 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	388	MG35-342 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	389	MG35-343 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	390	MG35-344 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	391	MG35-345 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	392	MG35-346 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	393	MG35-347 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	394	MG35-348 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	395	MG35-349 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	396	MG35-350 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	397	MG35-351 효과기	단백질	알려지지 않음	미배양 유기체	

별주	서열번호	설명	유형	유기체	기타 정보	코멘트 또는 서열
MG35 효과기	398	MG35-352 효과기	단백질	알러지지 않음	미배양 유기체	
MG35 효과기	399	MG35-353 효과기	단백질	알러지지 않음	미배양 유기체	
MG35 효과기	400	MG35-354 효과기	단백질	알러지지 않음	미배양 유기체	
MG35 효과기	401	MG35-355 효과기	단백질	알러지지 않음	미배양 유기체	
MG35 효과기	402	MG35-356 효과기	단백질	알러지지 않음	미배양 유기체	
MG35 효과기	403	MG35-357 효과기	단백질	알러지지 않음	미배양 유기체	
MG35 효과기	404	MG35-358 효과기	단백질	알러지지 않음	미배양 유기체	
MG35 효과기	405	MG35-359 효과기	단백질	알러지지 않음	미배양 유기체	
MG35 효과기	406	MG35-360 효과기	단백질	알러지지 않음	미배양 유기체	
MG35 효과기	407	MG35-361 효과기	단백질	알러지지 않음	미배양 유기체	
MG35 효과기	408	MG35-362 효과기	단백질	알러지지 않음	미배양 유기체	
MG35 효과기	409	MG35-363 효과기	단백질	알러지지 않음	미배양 유기체	
MG35 효과기	410	MG35-364 효과기	단백질	알러지지 않음	미배양 유기체	
MG35 효과기	411	MG35-365 효과기	단백질	알러지지 않음	미배양 유기체	
MG35 효과기	412	MG35-366 효과기	단백질	알러지지 않음	미배양 유기체	
MG35 효과기	413	MG35-367 효과기	단백질	알러지지 않음	미배양 유기체	
MG35 효과기	414	MG35-368 효과기	단백질	알러지지 않음	미배양 유기체	
MG35 효과기	415	MG35-369 효과기	단백질	알러지지 않음	미배양 유기체	
MG35 효과기	416	MG35-370 효과기	단백질	알러지지 않음	미배양 유기체	
MG35 효과기	417	MG35-371 효과기	단백질	알러지지 않음	미배양 유기체	
MG35 효과기	418	MG35-372 효과기	단백질	알러지지 않음	미배양 유기체	
MG35 효과기	419	MG35-373 효과기	단백질	알러지지 않음	미배양 유기체	
MG35 효과기	420	MG35-374 효과기	단백질	알러지지 않음	미배양 유기체	
MG35 효과기	421	MG35-375 효과기	단백질	알러지지 않음	미배양 유기체	
MG35 효과기	422	MG35-376 효과기	단백질	알러지지 않음	미배양 유기체	
MG35 효과기	423	MG35-377 효과기	단백질	알러지지 않음	미배양 유기체	
MG35 효과기	424	MG35-378 효과기	단백질	알러지지 않음	미배양 유기체	
MG35 효과기	425	MG35-379 효과기	단백질	알러지지 않음	미배양 유기체	
MG35 효과기	426	MG35-384 효과기	단백질	알러지지 않음	미배양 유기체	

[0271]

별주	서열번호	설명	유형	유기체	기타 정보	코멘트 또는 서열
MG35 효과기	427	MG35-385 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	428	MG35-386 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	429	MG35-387 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	430	MG35-388 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	431	MG35-389 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	432	MG35-390 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	433	MG35-391 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	434	MG35-392 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	435	MG35-393 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	436	MG35-394 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	437	MG35-395 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	438	MG35-396 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	439	MG35-397 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	440	MG35-398 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	441	MG35-399 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	442	MG35-400 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	443	MG35-401 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	444	MG35-402 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	445	MG35-403 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	446	MG35-404 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	447	MG35-405 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	448	MG35-406 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	449	MG35-408 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	450	MG35-409 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	451	MG35-410 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	452	MG35-411 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	453	MG35-412 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	454	MG35-413 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	455	MG35-414 효과기	단백질	알려지지 않음	미배양 유기체	

[0272]

번호	서열번호	설명	유형	유기체	기타 정보	포렌트 또는 서열
MG35 효과기	466	MG35-415 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	467	MG35-416 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	468	MG35-417 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	469	MG35-418 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	460	MG35-4 tracrRNA 서열	뉴클레오타이드	인공 서열		
tracrRNA 서열						
추정 tracrRNA	461	추정 MG35-3 tracrRNA	뉴클레오타이드	인공 서열		
번복	462	MG35-3 반복	뉴클레오타이드	인공 서열		
MG33 효과기	463	MG33-2 효과기	단백질	알려지지 않음	미배양 유기체	
MG33 효과기	464	MG33-3 효과기	단백질	알려지지 않음	미배양 유기체	
MG33 효과기	465	MG33-4 효과기	단백질	알려지지 않음	미배양 유기체	
MG33 효과기	466	MG33-5 효과기	단백질	알려지지 않음	미배양 유기체	
MG33 효과기	467	MG33-6 효과기	단백질	알려지지 않음	미배양 유기체	
MG33 효과기	468	MG33-7 효과기	단백질	알려지지 않음	미배양 유기체	
MG33 효과기	469	MG33-8 효과기	단백질	알려지지 않음	미배양 유기체	
MG33 효과기	470	MG33-9 효과기	단백질	알려지지 않음	미배양 유기체	
MG33 효과기	471	MG33-10 효과기	단백질	알려지지 않음	미배양 유기체	
MG33 효과기	472	MG33-11 효과기	단백질	알려지지 않음	미배양 유기체	
MG33 효과기	473	MG33-12 효과기	단백질	알려지지 않음	미배양 유기체	
MG33 효과기	474	MG33-13 효과기	단백질	알려지지 않음	미배양 유기체	
MG33 효과기	475	MG33-14 효과기	단백질	알려지지 않음	미배양 유기체	
MG33 효과기	476	MG33-15 효과기	단백질	알려지지 않음	미배양 유기체	
MG33 효과기	477	MG33-16 효과기	단백질	알려지지 않음	미배양 유기체	
MG33 효과기	478	MG33-17 효과기	단백질	알려지지 않음	미배양 유기체	
MG33 효과기	479	MG33-18 효과기	단백질	알려지지 않음	미배양 유기체	
MG33 효과기	480	MG33-19 효과기	단백질	알려지지 않음	미배양 유기체	
MG33 효과기	481	MG33-20 효과기	단백질	알려지지 않음	미배양 유기체	
MG33 효과기	482	MG33-21 효과기	단백질	알려지지 않음	미배양 유기체	
MG33 효과기	483	MG33-22 효과기	단백질	알려지지 않음	미배양 유기체	
MG33 효과기	484	MG33-23 효과기	단백질	알려지지 않음	미배양 유기체	

[0273]

별주	서열번호	설명	유형	유기체	기타 정보	코멘트 또는 서열
MG33 효과기	485	MG33-24 효과기	단백질	인러지치 없음	미배양 유기체	
MG33 효과기	486	MG33-26 효과기	단백질	인러지치 없음	미배양 유기체	
MG34 효과기	487	MG34-23 효과기	단백질	인러지치 없음	미배양 유기체	
MG34 효과기	488	MG34-24 효과기	단백질	인러지치 없음	미배양 유기체	
MG35 효과기	489	MG35-422 효과기	단백질	인러지치 없음	미배양 유기체	
MG35 효과기	490	MG35-423 효과기	단백질	인러지치 없음	미배양 유기체	
MG35 효과기	491	MG35-424 효과기	단백질	인러지치 없음	미배양 유기체	
MG35 효과기	492	MG35-425 효과기	단백질	인러지치 없음	미배양 유기체	
MG35 효과기	493	MG35-426 효과기	단백질	인러지치 없음	미배양 유기체	
MG35 효과기	494	MG35-427 효과기	단백질	인러지치 없음	미배양 유기체	
MG35 효과기	495	MG35-428 효과기	단백질	인러지치 없음	미배양 유기체	
MG35 효과기	496	MG35-429 효과기	단백질	인러지치 없음	미배양 유기체	
MG35 효과기	497	MG35-430 효과기	단백질	인러지치 없음	미배양 유기체	
MG35 효과기	498	MG35-431 효과기	단백질	인러지치 없음	미배양 유기체	
MG35 효과기	499	MG35-432 효과기	단백질	인러지치 없음	미배양 유기체	
MG35 효과기	500	MG35-433 효과기	단백질	인러지치 없음	미배양 유기체	
MG35 효과기	501	MG35-434 효과기	단백질	인러지치 없음	미배양 유기체	
MG35 효과기	502	MG35-435 효과기	단백질	인러지치 없음	미배양 유기체	
MG35 효과기	503	MG35-436 효과기	단백질	인러지치 없음	미배양 유기체	
MG35 효과기	504	MG35-437 효과기	단백질	인러지치 없음	미배양 유기체	
MG35 효과기	505	MG35-438 효과기	단백질	인러지치 없음	미배양 유기체	
MG35 효과기	506	MG35-439 효과기	단백질	인러지치 없음	미배양 유기체	
MG35 효과기	507	MG35-440 효과기	단백질	인러지치 없음	미배양 유기체	
MG35 효과기	508	MG35-441 효과기	단백질	인러지치 없음	미배양 유기체	
MG35 효과기	509	MG35-442 효과기	단백질	인러지치 없음	미배양 유기체	
MG35 효과기	510	MG35-443 효과기	단백질	인러지치 없음	미배양 유기체	
MG35 효과기	511	MG35-444 효과기	단백질	인러지치 없음	미배양 유기체	
MG35 효과기	512	MG35-445 효과기	단백질	인러지치 없음	미배양 유기체	
MG35 효과기	513	MG35-446 효과기	단백질	인러지치 없음	미배양 유기체	

[0274]

번호	서열번호	설명	유형	유기체	기타 정보	프렌트 또는 서열
MG35 효과기	514	MG35-447 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	515	MG35-448 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	516	MG35-449 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	517	MG35-450 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	518	MG35-451 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	519	MG35-452 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	520	MG35-453 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	521	MG35-454 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	522	MG35-455 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	523	MG35-456 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	524	MG35-457 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	525	MG35-458 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	526	MG35-459 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	527	MG35-460 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	528	MG35-461 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	529	MG35-462 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	530	MG35-463 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	531	MG35-464 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	532	MG35-465 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	533	MG35-466 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	534	MG35-467 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	535	MG35-468 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	536	MG35-469 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	537	MG35-470 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	538	MG35-471 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	539	MG35-472 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	540	MG35-473 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	541	MG35-474 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	542	MG35-475 효과기	단백질	알려지지 않음	미배양 유기체	

[0275]

별주	서열번호	설명	유형	유기체	기타 정보	모델트 또는 서열
MG35 효과기	543	MG35-476 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	544	MG35-477 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	545	MG35-478 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	546	MG35-479 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	547	MG35-480 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	548	MG35-481 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	549	MG35-482 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	550	MG35-483 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	551	MG35-484 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	552	MG35-485 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	553	MG35-486 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	554	MG35-487 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	555	MG35-488 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	556	MG35-489 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	557	MG35-490 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	558	MG35-491 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	559	MG35-492 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	560	MG35-493 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	561	MG35-494 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	562	MG35-495 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	563	MG35-496 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	564	MG35-497 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	565	MG35-498 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	566	MG35-499 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	567	MG35-500 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	568	MG35-501 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	569	MG35-502 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	570	MG35-503 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	571	MG35-504 효과기	단백질	알려지지 않음	미배양 유기체	

[0276]

별주	서열번호	설명	유형	유기체	기타 정보	모델트 또는 서열
MG35 효과기	572	MG35-505 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	573	MG35-506 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	574	MG35-507 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	575	MG35-508 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	576	MG35-509 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	577	MG35-510 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	578	MG35-511 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	579	MG35-512 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	580	MG35-513 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	581	MG102-1 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	582	MG102-2 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	583	MG102-3 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	584	MG102-4 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	585	MG102-5 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	586	MG102-6 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	587	MG102-7 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	588	MG102-8 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	589	MG102-9 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	590	MG102-10 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	591	MG102-11 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	592	MG102-12 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	593	MG102-13 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	594	MG102-14 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	595	MG102-15 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	596	MG102-16 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	597	MG102-17 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	598	MG102-18 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	599	MG102-19 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	600	MG102-20 효과기	단백질	알려지지 않음	미배양 유기체	

[0277]

별주	서열번호	설명	유형	유기체	기타 정보	표본 또는 서열
MG102 효과기	601	MG102-21 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	602	MG102-22 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	603	MG102-23 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	604	MG102-24 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	605	MG102-25 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	606	MG102-27 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	607	MG102-28 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	608	MG102-29 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	609	MG102-30 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	610	MG102-31 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	611	MG102-32 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	612	MG102-33 효과기	단백질	알려지지 않음	미배양 유기체	
MG34 sgRNA	613	MG34-1 활성 효과기 sgRNA 1	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG34 sgRNA	614	MG34-1 활성 효과기 sgRNA 2	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG34 sgRNA	615	MG34-9 활성 효과기 sgRNA 1	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG34 sgRNA	616	MG34-16 활성 효과기 sgRNA 1	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 효과기	617	MG35-514 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	618	MG35-515 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	619	MG35-516 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	620	MG35-517 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	621	MG35-518 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	622	MG35-519 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	623	MG35-520 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	624	MG35-521 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	625	MG35-522 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	626	MG35-523 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	627	MG35-524 효과기	단백질	알려지지 않음	미배양 유기체	

[0278]

별주	서열번호	설명	유형	유기체	기타 정보	코멘트 또는 서열
MG35 효과기	628	MG35-525 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	629	MG35-526 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	630	MG35-527 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	631	MG35-528 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	632	MG35-529 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	633	MG35-530 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	634	MG35-531 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	635	MG35-532 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	636	MG35-533 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	637	MG35-534 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	638	MG35-535 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	639	MG35-536 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	640	MG35-537 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	641	MG35-538 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	642	MG35-539 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	643	MG35-540 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	644	MG35-541 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	645	MG35-542 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	646	MG35-543 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	647	MG35-544 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	648	MG35-545 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	649	MG35-546 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	650	MG35-547 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	651	MG35-548 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	652	MG35-549 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	653	MG35-550 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	654	MG35-551 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	655	MG35-552 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	656	MG35-553 효과기	단백질	알려지지 않음	미배양 유기체	

번호	서열번호	설명	유형	유기체	기타 정보	코멘트 또는 서열
MG35 효과기	657	MG35-554 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	658	MG35-555 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	659	MG35-556 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	660	MG35-557 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	661	MG35-558 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	662	MG35-559 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	663	MG35-560 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	664	MG35-561 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	665	MG35-562 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	666	MG35-563 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	667	MG35-564 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	668	MG35-565 효과기	단백질	알려지지 않음	미배양 유기체	
MG33 tracrRNA	669	MG33-2 tracrRNA 1	뉴클레오타이드	인공 서열	MG33 tracrRNA	
MG33 tracrRNA	670	MG33-2 tracrRNA 2	뉴클레오타이드	인공 서열	MG33 tracrRNA	
MG33 tracrRNA	671	MG33-3 tracrRNA 1	뉴클레오타이드	인공 서열	MG33 tracrRNA	
MG102 tracrRNA	672	MG102-1 tracrRNA 1	뉴클레오타이드	인공 서열	MG102 tracrRNA	
MG102 tracrRNA	673	MG102-2 tracrRNA 1	뉴클레오타이드	인공 서열	MG102 tracrRNA	
MG35 효과기	674	MG35-566 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 효과기	675	MG35-567 효과기	단백질	알려지지 않음	미배양 유기체	
MG35 예측 CRISPR 반복	676	MG35-420 예측 CRISPR 반복	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA	677	MG35-1 효과기 추정 단일 가이드 RNA 1	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA	678	MG35-1 효과기 추정 단일 가이드 RNA 2	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA	679	MG35-2 효과기 추정 단일 가이드 RNA 1	뉴클레오타이드	알려지지 않음	미배양 유기체	

[0280]

별주	서열번호	설명	유형	유기체	기타 정보	표본 또는 서열
MG35-37 유전자 단위	680	MG35-37 유전자 단위	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-37 유전자 단위	681	MG35-37 유전자 단위	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-37 유전자 단위	682	MG35-37 유전자 단위	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-37 유전자 단위	683	MG35-37 유전자 단위	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-37 유전자 단위	684	MG35-37 유전자 단위	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-37 유전자 단위	685	MG35-37 유전자 단위	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-37 유전자 단위	686	MG35-37 유전자 단위	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-37 유전자 단위	687	MG35-37 유전자 단위	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-37 유전자 단위	688	MG35-37 유전자 단위	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-37 유전자 단위	689	MG35-37 유전자 단위	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-37 유전자 단위	690	MG35-37 유전자 단위	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-37 유전자 단위	691	MG35-37 유전자 단위	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-37 유전자 단위	692	MG35-37 유전자 단위	뉴클레오타이드	알려지지 않음	미배양 유기체	

[0281]

별주	서열번호	설명	유형	유기체	기법 정보	코멘트 또는 서열
MG35 추정 단일 가이드 RNA 암호화 서열	693	MG35-38 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	694	MG35-40, MG35-42, MG35-43 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	695	MG35-41 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	696	MG35-44 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	697	MG35-45, MG35-39, MG35-116, MG35-219 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	698	MG35-46 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	699	MG35-48 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	700	MG35-49 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	701	MG35-52 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	702	MG35-53, MG35-54 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	703	MG35-55 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	

[0282]

별주	서열번호	설명	유형	유기체	기타 정보	모델트 또는 서열
MG35-56, MG35-287 효과기 추정 단일 가이드 서열	704	MG35-56, MG35-287 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-57, MG35-59 효과기 추정 단일 가이드 서열	705	MG35-57, MG35-59 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-60, MG35-62 효과기 추정 단일 가이드 서열	706	MG35-60, MG35-62 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-63, MG35-65 효과기 추정 단일 가이드 서열	707	MG35-63, MG35-65 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-64, MG35-66 효과기 추정 단일 가이드 서열	708	MG35-64, MG35-66 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-67, MG35-69 효과기 추정 단일 가이드 서열	709	MG35-67, MG35-69 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-70, MG35-72 효과기 추정 단일 가이드 서열	710	MG35-70, MG35-72 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-71, MG35-73 효과기 추정 단일 가이드 서열	711	MG35-71, MG35-73 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-74, MG35-76 효과기 추정 단일 가이드 서열	712	MG35-74, MG35-76 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-75, MG35-77 효과기 추정 단일 가이드 서열	713	MG35-75, MG35-77 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-78, MG35-80 효과기 추정 단일 가이드 서열	714	MG35-78, MG35-80 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	

[0283]

별주	서열번호	설명	유형	유기체	기타 정보	표본 또는 서열
MG35-73 추적 단일 가이드 RNA 암호화 서열	715	MG35-73 효과기 추적 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-74 추적 단일 가이드 RNA 암호화 서열	716	MG35-74 효과기 추적 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-75 추적 단일 가이드 RNA 암호화 서열	717	MG35-75 효과기 추적 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-76 추적 단일 가이드 RNA 암호화 서열	718	MG35-76 효과기 추적 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-77 추적 단일 가이드 RNA 암호화 서열	719	MG35-77, MG35-97, MG35-98 효과기 추적 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-78 추적 단일 가이드 RNA 암호화 서열	720	MG35-80, MG35-81 효과기 추적 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-79 추적 단일 가이드 RNA 암호화 서열	721	MG35-82, MG35-95, MG35-96 효과기 추적 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-80 추적 단일 가이드 RNA 암호화 서열	722	MG35-86 효과기 추적 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-81 추적 단일 가이드 RNA 암호화 서열	723	MG35-87, MG35-88, MG35-89 효과기 추적 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-82 추적 단일 가이드 RNA 암호화 서열	724	MG35-91, MG35-92 효과기 추적 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-83 추적 단일 가이드 RNA 암호화 서열	725	MG35-93, MG35-94 효과기 추적 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	

[0284]

별주	서열번호	설명	유형	유기체	기타 정보	코멘트 또는 서열
MG35-99-99 효과기 측정 단일 가이드 RNA 암호화 서열	726	MG35-99 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-101 효과기 측정 단일 가이드 RNA 암호화 서열	727	MG35-101 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-102 효과기 측정 단일 가이드 RNA 암호화 서열	728	MG35-102 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-103, MG35-104 효과기 측정 단일 가이드 RNA 암호화 서열	729	MG35-103, MG35-104 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-105 효과기 측정 단일 가이드 RNA 암호화 서열	730	MG35-105 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-106 효과기 측정 단일 가이드 RNA 암호화 서열	731	MG35-106 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-107 효과기 측정 단일 가이드 RNA 암호화 서열	732	MG35-107 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-108 효과기 측정 단일 가이드 RNA 암호화 서열	733	MG35-108 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-109 효과기 측정 단일 가이드 RNA 암호화 서열	734	MG35-109 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-110, MG35-112 효과기 측정 단일 가이드 RNA 암호화 서열	735	MG35-110, MG35-112 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-111 효과기 측정 단일 가이드 RNA 암호화 서열	736	MG35-111 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	

[0285]

별주	서열번호	설명	유형	유기체	기타 정보	포렌트 또는 서열
MG35 추정 단일 가이드 RNA 암호화 서열	737	MG35-113 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	738	MG35-114 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	739	MG35-115 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	740	MG35-116 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	741	MG35-117, MG35-118, MG35-119 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	742	MG35-120 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	743	MG35-121 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	744	MG35-122 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	745	MG35-123, MG35-124 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	746	MG35-125 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	747	MG35-126, MG35-377 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	

[0286]

번호	시험번호	설명	유형	유기체	기타 정보	코멘트 또는 서열
MG35 추적 단일 가이드 RNA 암호화 서열	748	MG35-127 효과기 추적 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추적 단일 가이드 RNA 암호화 서열	749	MG35-128 효과기 추적 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추적 단일 가이드 RNA 암호화 서열	750	MG35-129 효과기 추적 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추적 단일 가이드 RNA 암호화 서열	751	MG35-130 효과기 추적 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추적 단일 가이드 RNA 암호화 서열	752	MG35-131 효과기 추적 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추적 단일 가이드 RNA 암호화 서열	753	MG35-147 효과기 추적 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추적 단일 가이드 RNA 암호화 서열	754	MG35-148 효과기 추적 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추적 단일 가이드 RNA 암호화 서열	755	MG35-149 효과기 추적 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추적 단일 가이드 RNA 암호화 서열	756	MG35-150 효과기 추적 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추적 단일 가이드 RNA 암호화 서열	757	MG35-151 효과기 추적 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추적 단일 가이드 RNA 암호화 서열	758	MG35-152 효과기 추적 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	

[0287]

번호	서열번호	설명	유형	유기체	기타 정보	코멘트 또는 서열
MG35-153 효과기 측정 가이드 RNA 암호화 서열	759	MG35-153 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-156 효과기 측정 가이드 RNA 암호화 서열	760	MG35-156 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-157 효과기 측정 가이드 RNA 암호화 서열	761	MG35-157 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-159 효과기 측정 가이드 RNA 암호화 서열	762	MG35-159 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-160 효과기 측정 가이드 RNA 암호화 서열	763	MG35-160 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-165 효과기 측정 가이드 RNA 암호화 서열	764	MG35-165 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-171 효과기 측정 가이드 RNA 암호화 서열	765	MG35-171 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-214 효과기 측정 가이드 RNA 암호화 서열	766	MG35-214 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-217 효과기 측정 가이드 RNA 암호화 서열	767	MG35-217 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-218 효과기 측정 가이드 RNA 암호화 서열	768	MG35-218 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-220 효과기 측정 가이드 RNA 암호화 서열	769	MG35-220 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	

[0288]

번호	서열번호	설명	유형	유기체	기타 정보	코멘트 또는 서열
MG35 추정 단일 가이드 RNA 암호화 서열	770	MG35-221, MG35-222, MG35-351, MG35-352, MG35-353 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	771	MG35-223 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	772	MG35-224 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	773	MG35-225 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	774	MG35-226 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	775	MG35-227 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	776	MG35-228 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	777	MG35-229 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	778	MG35-230 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	779	MG35-231 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	780	MG35-232 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	

[0289]

별주	서열번호	설명	유형	유기체	기타 정보	코멘트 또는 서열
MG35 추정 단일 가이드 RNA 암호화 서열	781	MG35-233 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	782	MG35-235 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	783	MG35-236 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	784	MG35-238, MG35-237 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	785	MG35-239, MG35-240, MG35-241 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	786	MG35-242 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	787	MG35-243 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	788	MG35-244 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	789	MG35-245 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	790	MG35-246 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	791	MG35-247 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	

[0290]

별주	서열번호	설명	유형	유기체	기타 정보	표본 또는 서열
MG35-248 단일 가이드 RNA 암호화 서열	792	MG35-248 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-249 단일 가이드 RNA 암호화 서열	793	MG35-249 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-250 단일 가이드 RNA 암호화 서열	794	MG35-250, MG35-251 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-252 단일 가이드 RNA 암호화 서열	795	MG35-252 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-253 단일 가이드 RNA 암호화 서열	796	MG35-253 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-255 단일 가이드 RNA 암호화 서열	797	MG35-255 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-256 단일 가이드 RNA 암호화 서열	798	MG35-256 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-257 단일 가이드 RNA 암호화 서열	799	MG35-257 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-258 단일 가이드 RNA 암호화 서열	800	MG35-258 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-259 단일 가이드 RNA 암호화 서열	801	MG35-259 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-260 단일 가이드 RNA 암호화 서열	802	MG35-260 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	

[0291]

별주	서열번호	설명	유형	유기체	기타 정보	코멘트 또는 서열
MG35 추정 단일 가이드 RNA 암호화 서열	803	MG35-262, MG35-263 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	804	MG35-266, MG35-270 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	805	MG35-267 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	806	MG35-268 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	807	MG35-269, MG35-264, MG35-265 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	808	MG35-271 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	809	MG35-272 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	810	MG35-273 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	811	MG35-274 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	812	MG35-275 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	813	MG35-276 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	

[0292]

번호	서열번호	설명	유형	유기체	기타 정보	코멘트 또는 서열
MG35-277 단일 가이드 RNA 암호화 서열	814	MG35-277 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-278 단일 가이드 RNA 암호화 서열	815	MG35-278 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-279 단일 가이드 RNA 암호화 서열	816	MG35-279 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-280 단일 가이드 RNA 암호화 서열	817	MG35-280 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-281 단일 가이드 RNA 암호화 서열	818	MG35-281 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-282 단일 가이드 RNA 암호화 서열	819	MG35-282, MG35-283 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-284 단일 가이드 RNA 암호화 서열	820	MG35-284 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-285 단일 가이드 RNA 암호화 서열	821	MG35-285 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-286 단일 가이드 RNA 암호화 서열	822	MG35-286 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-289 단일 가이드 RNA 암호화 서열	823	MG35-289, MG35-293 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-296 단일 가이드 RNA 암호화 서열	824	MG35-296 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	

[0293]

별주	서열번호	설명	유형	유기체	기타 정보	코멘트 또는 서열
MG35 추정 단일 가이드 RNA 암호화 서열	825	MG35-298, MG35-299 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오티드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	826	MG35-300 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오티드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	827	MG35-302 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오티드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	828	MG35-303 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오티드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	829	MG35-305 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오티드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	830	MG35-307 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오티드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	831	MG35-308 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오티드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	832	MG35-309 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오티드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	833	MG35-310 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오티드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	834	MG35-311 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오티드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	835	MG35-312 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오티드	알려지지 않음	미배양 유기체	

[0294]

별주	서열번호	설명	유형	유기체	기타 정보	코멘트 또는 서열
MG35 추정 단일 가이드 RNA 암호화 서열	836	MG35-313 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	837	MG35-314, MG35-261 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	838	MG35-315 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	839	MG35-316 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	840	MG35-317 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	841	MG35-318 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	842	MG35-319 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	843	MG35-321 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	844	MG35-322 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	845	MG35-325 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	846	MG35-326 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	

[0295]

별주	서열번호	설명	유형	유기체	기타 정보	코멘트 또는 서열
MG35-327 단일 가이드 RNA 암호화 서열	847	MG35-327 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-328 단일 가이드 RNA 암호화 서열	848	MG35-328 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-329 단일 가이드 RNA 암호화 서열	849	MG35-329 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-330 단일 가이드 RNA 암호화 서열	850	MG35-330 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-331 단일 가이드 RNA 암호화 서열	851	MG35-331 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-332 단일 가이드 RNA 암호화 서열	852	MG35-332, MG35-333, MG35-335 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-334 단일 가이드 RNA 암호화 서열	853	MG35-334 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-336 단일 가이드 RNA 암호화 서열	854	MG35-336 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-340 단일 가이드 RNA 암호화 서열	855	MG35-340 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-341 단일 가이드 RNA 암호화 서열	856	MG35-341 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-342 단일 가이드 RNA 암호화 서열	857	MG35-342 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	

[0296]

별주	서열번호	설명	유형	유기체	기타 정보	코멘트 또는 서열
MG35 추정 단일 가이드 RNA 암호화 서열	858	MG35-343 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	859	MG35-344, MG35-345 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	860	MG35-346, MG35-347 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	861	MG35-348 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	862	MG35-349 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	863	MG35-350 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	864	MG35-351 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	865	MG35-355, MG35-356 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	866	MG35-357 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	867	MG35-358 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	868	MG35-359 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	

[0297]

번호	시험번호	설명	유형	유기체	기타 정보	코멘트 또는 서열
MG35-360, MG35-361 효과기 측정 단일 가이드 서열	869	MG35-360, MG35-361 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-362, MG35-363 효과기 측정 단일 가이드 RNA 암호화 서열	870	MG35-362, MG35-363 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-364, MG35-365 효과기 측정 단일 가이드 RNA 암호화 서열	871	MG35-364, MG35-365 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-366, MG35-367 효과기 측정 단일 가이드 RNA 암호화 서열	872	MG35-366, MG35-367 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-368, MG35-369 효과기 측정 단일 가이드 RNA 암호화 서열	873	MG35-368, MG35-369 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-370, MG35-371 효과기 측정 단일 가이드 RNA 암호화 서열	874	MG35-370, MG35-371 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-372, MG35-373 효과기 측정 단일 가이드 RNA 암호화 서열	875	MG35-372, MG35-373 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-374, MG35-375 효과기 측정 단일 가이드 RNA 암호화 서열	876	MG35-374, MG35-375 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-376, MG35-377 효과기 측정 단일 가이드 RNA 암호화 서열	877	MG35-376, MG35-377 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-378, MG35-379 효과기 측정 단일 가이드 RNA 암호화 서열	878	MG35-378, MG35-379 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35-380, MG35-381 효과기 측정 단일 가이드 RNA 암호화 서열	879	MG35-380, MG35-381 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	

[0298]

번호	서열번호	설명	유형	유기체	기타 정보	코멘트 또는 서열
MG35 추정 단일 가이드 RNA 암호화 서열	880	MG35-374 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	881	MG35-375 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	882	MG35-376 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	883	MG35-378 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	884	MG35-379 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	885	MG35-384 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	886	MG35-386 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	887	MG35-388 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	888	MG35-419, MG35-330 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	889	MG35-420, MG35-337 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	890	MG35-421, MG35-338 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	

[0299]

별주	서열번호	설명	유형	유기체	기타 정보	코멘트 또는 서열
MG35 추정 단일 가이드 RNA 암호화 서열	891	MG35-422 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	892	MG35-423 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	893	MG35-424 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	894	MG35-426 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	895	MG35-427 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	896	MG35-428, MG35-436, MG35-437, MG35-457 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	897	MG35-429, MG35-449 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	898	MG35-430 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	899	MG35-431, MG35-442 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	900	MG35-432 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	901	MG35-433, MG35-425 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	

[0300]

별주	서열번호	설명	유형	유기체	기타 정보	코멘트 또는 서열
MG35 측정 단일 가이드 RNA 암호화 서열	902	MG35-434, MG35-455 표과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 측정 단일 가이드 RNA 암호화 서열	903	MG35-435 표과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 측정 단일 가이드 RNA 암호화 서열	904	MG35-438 표과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 측정 단일 가이드 RNA 암호화 서열	905	MG35-439 표과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 측정 단일 가이드 RNA 암호화 서열	906	MG35-440 표과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 측정 단일 가이드 RNA 암호화 서열	907	MG35-441, MG35-443 표과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 측정 단일 가이드 RNA 암호화 서열	908	MG35-444 표과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 측정 단일 가이드 RNA 암호화 서열	909	MG35-445 표과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 측정 단일 가이드 RNA 암호화 서열	910	MG35-446, MG35-448, MG35-456 표과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 측정 단일 가이드 RNA 암호화 서열	911	MG35-447 표과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 측정 단일 가이드 RNA 암호화 서열	912	MG35-450 표과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	

[0301]

[0302]

별주	서열번호	설명	유형	유기체	기타 정보	코멘트 또는 서열
MG35 추정 단일 가이드 RNA 암호화 서열	913	MG35-451 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	914	MG35-452 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	915	MG35-453 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	916	MG35-454 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	917	MG35-458, MG35-523 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	918	MG35-459 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	919	MG35-460 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	920	MG35-461 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	921	MG35-462 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	922	MG35-463 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	923	MG35-464 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	

번호	서열번호	설명	유형	유기체	기타 정보	코멘트 또는 서열
MG35 측정 단일 가이드 RNA 암호화 서열	924	MG35-465 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 측정 단일 가이드 RNA 암호화 서열	925	MG35-466 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 측정 단일 가이드 RNA 암호화 서열	926	MG35-510 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 측정 단일 가이드 RNA 암호화 서열	927	MG35-511, MG35-512 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 측정 단일 가이드 RNA 암호화 서열	928	MG35-513 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 측정 단일 가이드 RNA 암호화 서열	929	MG35-514 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 측정 단일 가이드 RNA 암호화 서열	930	MG35-515 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 측정 단일 가이드 RNA 암호화 서열	931	MG35-516 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 측정 단일 가이드 RNA 암호화 서열	932	MG35-517 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 측정 단일 가이드 RNA 암호화 서열	933	MG35-518 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 측정 단일 가이드 RNA 암호화 서열	934	MG35-519 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	

[0303]

별주	서열번호	설명	유형	유기체	기타 정보	코멘트 또는 서열
MG35 추출 단일 가이드 RNA 암호화 서열	935	MG35-520 효과기 추출 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추출 단일 가이드 RNA 암호화 서열	936	MG35-521 효과기 추출 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추출 단일 가이드 RNA 암호화 서열	937	MG35-522, MG35-526, MG35-546, MG35-548 효과기 추출 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추출 단일 가이드 RNA 암호화 서열	938	MG35-524 효과기 추출 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추출 단일 가이드 RNA 암호화 서열	939	MG35-525, MG35-537 효과기 추출 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추출 단일 가이드 RNA 암호화 서열	940	MG35-527 효과기 추출 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추출 단일 가이드 RNA 암호화 서열	941	MG35-528 효과기 추출 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추출 단일 가이드 RNA 암호화 서열	942	MG35-529 효과기 추출 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추출 단일 가이드 RNA 암호화 서열	943	MG35-530 효과기 추출 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추출 단일 가이드 RNA 암호화 서열	944	MG35-531 효과기 추출 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추출 단일 가이드 RNA 암호화 서열	945	MG35-532 효과기 추출 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	

[0304]

번호	서열번호	설명	유형	유기체	기타 정보	코멘트 또는 서열
MG35 추정 단일 가이드 RNA 암호화 서열	946	MG35-533 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	947	MG35-534 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	948	MG35-535 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	949	MG35-536 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	950	MG35-538 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	951	MG35-539 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	952	MG35-540 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	953	MG35-541 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	954	MG35-542 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	955	MG35-543 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	956	MG35-544 효과기 측정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	

[0305]

[0306]

별주	서열번호	설명	유형	유기체	기타 정보	표본 또는 서열
MG35 추정 단일 가이드 RNA 암호화 서열	957	MG35-545 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	958	MG35-547 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	959	MG35-549 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	960	MG35-550 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	961	MG35-552 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	962	MG35-553 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	963	MG35-554 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	964	MG35-555 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	965	MG35-556 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	966	MG35-557 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG35 추정 단일 가이드 RNA 암호화 서열	967	MG35-558 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	알려지지 않음	미배양 유기체	

별주	서열번호	설명	유형	유기체	기타 정보	코멘트 또는 서열
MG35-558-559 효과기 추정 단일 가이드 RNA 암호화 서열	968	MG35-558-559 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	일려지지 않음	미배양 유기체	
MG35-560-561 효과기 추정 단일 가이드 RNA 암호화 서열	969	MG35-560-561 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	일려지지 않음	미배양 유기체	
MG35-562-563 효과기 추정 단일 가이드 RNA 암호화 서열	970	MG35-562-563 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	일려지지 않음	미배양 유기체	
MG35-564-565 효과기 추정 단일 가이드 RNA 암호화 서열	971	MG35-564-565 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	일려지지 않음	미배양 유기체	
MG35-566-567 효과기 추정 단일 가이드 RNA 암호화 서열	972	MG35-566-567 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	일려지지 않음	미배양 유기체	
MG35-568-569 효과기 추정 단일 가이드 RNA 암호화 서열	973	MG35-568-569 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	일려지지 않음	미배양 유기체	
MG35-570-571 효과기 추정 단일 가이드 RNA 암호화 서열	974	MG35-570-571 효과기 추정 단일 가이드 RNA 암호화 서열	뉴클레오타이드	일려지지 않음	미배양 유기체	
MG35-572-573 효과기 추정 단일 가이드 RNA 암호화 서열	975	MG35-572-573 효과기 추정 단일 가이드 RNA 암호화 서열	단백질	일려지지 않음	미배양 유기체	
MG35-574-575 효과기 추정 단일 가이드 RNA 암호화 서열	976	MG35-574-575 효과기 추정 단일 가이드 RNA 암호화 서열	단백질	일려지지 않음	미배양 유기체	
MG35-576-577 효과기 추정 단일 가이드 RNA 암호화 서열	977	MG35-576-577 효과기 추정 단일 가이드 RNA 암호화 서열	단백질	일려지지 않음	미배양 유기체	
MG35-578-579 효과기 추정 단일 가이드 RNA 암호화 서열	978	MG35-578-579 효과기 추정 단일 가이드 RNA 암호화 서열	단백질	일려지지 않음	미배양 유기체	
MG35-580-581 효과기 추정 단일 가이드 RNA 암호화 서열	979	MG35-580-581 효과기 추정 단일 가이드 RNA 암호화 서열	단백질	일려지지 않음	미배양 유기체	
MG35-582-583 효과기 추정 단일 가이드 RNA 암호화 서열	980	MG35-582-583 효과기 추정 단일 가이드 RNA 암호화 서열	단백질	일려지지 않음	미배양 유기체	
MG35-584-585 효과기 추정 단일 가이드 RNA 암호화 서열	981	MG35-584-585 효과기 추정 단일 가이드 RNA 암호화 서열	단백질	일려지지 않음	미배양 유기체	
MG35-586-587 효과기 추정 단일 가이드 RNA 암호화 서열	982	MG35-586-587 효과기 추정 단일 가이드 RNA 암호화 서열	단백질	일려지지 않음	미배양 유기체	
MG35-588-589 효과기 추정 단일 가이드 RNA 암호화 서열	983	MG35-588-589 효과기 추정 단일 가이드 RNA 암호화 서열	단백질	일려지지 않음	미배양 유기체	
MG35-590-591 효과기 추정 단일 가이드 RNA 암호화 서열	984	MG35-590-591 효과기 추정 단일 가이드 RNA 암호화 서열	단백질	일려지지 않음	미배양 유기체	
MG35-592-593 효과기 추정 단일 가이드 RNA 암호화 서열	985	MG35-592-593 효과기 추정 단일 가이드 RNA 암호화 서열	단백질	일려지지 않음	미배양 유기체	
MG35-594-595 효과기 추정 단일 가이드 RNA 암호화 서열	986	MG35-594-595 효과기 추정 단일 가이드 RNA 암호화 서열	단백질	일려지지 않음	미배양 유기체	

[0307]

번호	서열번호	설명	유형	유기체	기타 정보	코멘트 또는 서열
MG33 효과기	987	MG33-33 효과기	단백질	알려지지 않음	미배양 유기체	
MG33 효과기	988	MG33-34 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	989	MG102-35 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	990	MG102-36 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	991	MG102-37 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	992	MG102-38 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	993	MG102-39 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	994	MG102-40 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	995	MG102-41 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	996	MG102-42 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	997	MG102-43 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	998	MG102-44 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	999	MG102-45 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	1000	MG102-46 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	1001	MG102-47 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	1002	MG102-48 효과기	단백질	알려지지 않음	미배양 유기체	
MG33 활성 효과기 sgRNA	1003	MG33-3 활성 효과기 sgRNA	뉴클레오타이드	인공 서열	N/A	
MG33 활성 효과기 sgRNA	1004	MG33-31 활성 효과기 sgRNA	뉴클레오타이드	인공 서열	N/A	
MG33 활성 효과기 sgRNA	1005	MG33-34 활성 효과기 sgRNA	뉴클레오타이드	인공 서열	N/A	
MG35 활성 효과기 sgRNA	1006	MG35-1 활성 효과기 sgRNA	뉴클레오타이드	인공 서열	N/A	
MG35 활성 효과기 sgRNA	1007	MG35-2 활성 효과기 sgRNA	뉴클레오타이드	인공 서열	N/A	
MG35 활성 효과기 sgRNA	1008	MG35-3 활성 효과기 sgRNA	뉴클레오타이드	인공 서열	N/A	
MG35 활성 효과기 sgRNA	1009	MG35-4 활성 효과기 sgRNA	뉴클레오타이드	인공 서열	N/A	
MG35 활성 효과기 sgRNA	1010	MG35-5 활성 효과기 sgRNA	뉴클레오타이드	인공 서열	N/A	

[0308]

별구	서열번호	설명	유형	유기체	기타 정보	코멘트 또는 서열
MG35 활성 효과기 sgRNA	1011	MG35-6 활성 효과기 sgRNA	뉴클레오타이드	인공 서열	N/A	
MG35 활성 효과기 sgRNA	1012	MG35-102 활성 효과기 sgRNA	뉴클레오타이드	인공 서열	N/A	
MG102 활성 효과기 sgRNA	1013	MG102-2 활성 효과기 sgRNA	뉴클레오타이드	인공 서열	N/A	
MG102 활성 효과기 sgRNA	1014	MG102-14 활성 효과기 sgRNA	뉴클레오타이드	인공 서열	N/A	
MG102 활성 효과기 sgRNA	1015	MG102-35 활성 효과기 sgRNA	뉴클레오타이드	인공 서열	N/A	
MG102 활성 효과기 sgRNA	1016	MG102-38 활성 효과기 sgRNA	뉴클레오타이드	인공 서열	N/A	
MG102 활성 효과기 sgRNA	1017	MG102-39 활성 효과기 sgRNA	뉴클레오타이드	인공 서열	N/A	
MG102 활성 효과기 sgRNA	1018	MG102-42 활성 효과기 sgRNA	뉴클레오타이드	인공 서열	N/A	
MG102 활성 효과기 sgRNA	1019	MG102-43 활성 효과기 sgRNA	뉴클레오타이드	인공 서열	N/A	
MG102 활성 효과기 sgRNA	1020	MG102-45 활성 효과기 sgRNA	뉴클레오타이드	인공 서열	N/A	
MG102 활성 효과기 sgRNA	1021	MG102-47 활성 효과기 sgRNA	뉴클레오타이드	인공 서열	N/A	
MG102 활성 효과기 sgRNA	1022	MG102-48 활성 효과기 sgRNA	뉴클레오타이드	인공 서열	N/A	
MG33 활성 효과기 PAM	1023	MG34-1, MG34-9, MG34-16, MG33-3, MG33-31, MG33-34 활성 효과기 PAM	뉴클레오타이드	인공 서열	N/A	n3G
MG33 활성 효과기 PAM	1024	MG33-31, MG33-34 활성 효과기 PAM	뉴클레오타이드	인공 서열	N/A	n3Gmmmm
MG33 활성 효과기 PAM	1025	MG33-31, MG33-34 활성 효과기 PAM	뉴클레오타이드	인공 서열	N/A	n3Gmmmm
MG34 활성 효과기 PAM	1026	MG34-1, MG34-9, MG34-16, MG33-3 활성 효과기 PAM	뉴클레오타이드	인공 서열	N/A	n3G

[0309]

별주	서열번호	설명	유형	유기체	기타 정보	코덱스 또는 서열
MG34 활성 효과기 PAM	1027	MG34-1, MG34-9, MG34-16, MG33-8 활성 효과기 PAM	뉴클레오타이드	인공 서열	N/A	nGG
MG34 활성 효과기 PAM	1028	MG34-1, MG34-9, MG34-16, MG33-8 활성 효과기 PAM	뉴클레오타이드	인공 서열	N/A	nGG
MG35 활성 효과기 PAM	1029	MG35-1 활성 효과기 PAM	뉴클레오타이드	인공 서열	N/A	AnGg
MG35 활성 효과기 PAM	1030	MG35-2 활성 효과기 PAM	뉴클레오타이드	인공 서열	N/A	nARAA
MG35 활성 효과기 PAM	1031	MG35-3 활성 효과기 PAM	뉴클레오타이드	인공 서열	N/A	ATGaaa
MG35 활성 효과기 PAM	1032	MG35-4 활성 효과기 PAM	뉴클레오타이드	인공 서열	N/A	ATGA
MG35 활성 효과기 PAM	1033	MG35-5 활성 효과기 PAM	뉴클레오타이드	인공 서열	N/A	WTGG
MG35 활성 효과기 PAM	1034	MG35-102 활성 효과기 PAM	뉴클레오타이드	인공 서열	N/A	RTGA
MG102 활성 효과기 PAM	1035	MG102-2 활성 효과기 PAM	뉴클레오타이드	인공 서열	N/A	nRC
MG102 활성 효과기 PAM	1036	MG102-14, MG102-35, MG102-36, MG102-42, MG102-43, MG102-45, MG102-47, MG102-48 활성 효과기 PAM	뉴클레오타이드	인공 서열	N/A	nRCnnnn
MG102 활성 효과기 PAM	1037	MG102-14, MG102-35, MG102-36, MG102-42, MG102-43, MG102-45, MG102-47, MG102-48 활성 효과기 PAM	뉴클레오타이드	인공 서열	N/A	nRCnnnn
MG102 활성 효과기 PAM	1038	MG102-14, MG102-35, MG102-36, MG102-42, MG102-43, MG102-45, MG102-47, MG102-48 활성 효과기 PAM	뉴클레오타이드	인공 서열	N/A	nRCnnnn

[0310]

행주	서열번호	설명	유형	유기체	기타 정보	코덱트 또는 서열
MG102 활성 효과기 PAM	1039	MG102-39 활성 효과기 PAM	뉴클레오타이드	인공 서열	N/A	naRnnnn
MG102 활성 효과기 PAM	1040	MG102-14, MG102-35, MG102-36, MG102-42, MG102-43, MG102-45, MG102-47, MG102-48 활성 효과기 PAM	뉴클레오타이드	인공 서열	N/A	nrCnnnn
MG102 활성 효과기 PAM	1041	MG102-14, MG102-35, MG102-36, MG102-42, MG102-43, MG102-45, MG102-47, MG102-48 활성 효과기 PAM	뉴클레오타이드	인공 서열	N/A	nrCnnnn
MG102 활성 효과기 PAM	1042	MG102-14, MG102-35, MG102-36, MG102-42, MG102-43, MG102-45, MG102-47, MG102-48 활성 효과기 PAM	뉴클레오타이드	인공 서열	N/A	nrCnnnn
MG102 활성 효과기 PAM	1043	MG102-14, MG102-35, MG102-36, MG102-42, MG102-43, MG102-45, MG102-47, MG102-48 활성 효과기 PAM	뉴클레오타이드	인공 서열	N/A	nrCnnnn
MG102 활성 효과기 PAM	1044	MG102-14, MG102-35, MG102-36, MG102-42, MG102-43, MG102-45, MG102-47, MG102-48 활성 효과기 PAM	뉴클레오타이드	인공 서열	N/A	nrCnnnn
MG33 CRISPR 반복	1045	MG33-1 CRISPR 반복	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG33 CRISPR 반복	1046	MG33-2 CRISPR 반복	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG33 CRISPR 반복	1047	MG33-3 CRISPR 반복	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG33 CRISPR 반복	1048	MG33-27 CRISPR 반복	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG33 CRISPR 반복	1049	MG33-28 CRISPR 반복	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG33 CRISPR 반복	1050	MG33-29 CRISPR 반복	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG33 CRISPR 반복	1051	MG33-30 CRISPR 반복	뉴클레오타이드	알려지지 않음	미배양 유기체	

[0311]

[0312]

범주	서열번호	설명	유형	유기체	기타 정보	표본 또는 서열
MG33 CRISPR 반복	1052	MG33-31 CRISPR 반복	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG33 CRISPR 반복	1053	MG33-33 CRISPR 반복	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG33 CRISPR 반복	1054	MG33-34 CRISPR 반복	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG34 CRISPR 반복	1055	MG34-1, MG34-9 CRISPR 반복	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG34 CRISPR 반복	1056	MG34-1, MG34-9 CRISPR 반복	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG34 CRISPR 반복	1057	MG34-25 CRISPR 반복	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG102 CRISPR 반복	1058	MG102-1 CRISPR 반복	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG102 CRISPR 반복	1059	MG102-2 CRISPR 반복	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG102 CRISPR 반복	1060	MG102-3 CRISPR 반복	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG102 CRISPR 반복	1061	MG102-10 CRISPR 반복	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG102 CRISPR 반복	1062	MG102-14 CRISPR 반복	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG102 CRISPR 반복	1063	MG102-35 CRISPR 반복	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG102 CRISPR 반복	1064	MG102-36 CRISPR 반복	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG102 CRISPR 반복	1065	MG102-38 CRISPR 반복	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG102 CRISPR 반복	1066	MG102-39 CRISPR 반복	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG102 CRISPR 반복	1067	MG102-42 CRISPR 반복	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG102 CRISPR 반복	1068	MG102-43 CRISPR 반복	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG102 CRISPR 반복	1069	MG102-44 CRISPR 반복	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG102 CRISPR 반복	1070	MG102-45 CRISPR 반복	뉴클레오타이드	알려지지 않음	미배양 유기체	

별주	서열번호	설명	유형	유기체	기타 정보	표본 또는 서열
MG102 CRISPR 반복	1071	MG102-47 CRISPR 반복	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG102 CRISPR 반복	1072	MG102-48 CRISPR 반복	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG143 CRISPR 반복	1073	MG143-1 CRISPR 반복	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG144 CRISPR 반복	1074	MG144-1 CRISPR 반복	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG144 CRISPR 반복	1075	MG144-2 CRISPR 반복	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG144 CRISPR 반복	1076	MG144-3, MG144-4 CRISPR 반복	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG144 CRISPR 반복	1077	MG144-3, MG144-4 CRISPR 반복	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG145 CRISPR 반복	1078	MG145-1 CRISPR 반복	뉴클레오타이드	알려지지 않음	미배양 유기체	
MG102-2 인간 TRAC 표지 부위	1079	MG102-2 TRAC A1 24nt	뉴클레오타이드	인공 서열	N/A	
MG102-2 인간 TRAC 표지 부위	1080	MG102-2 TRAC B1 24nt	뉴클레오타이드	인공 서열	N/A	
MG102-2 인간 TRAC 표지 부위	1081	MG102-2 TRAC A1 20nt	뉴클레오타이드	인공 서열	N/A	
MG102-2 인간 TRAC 표지 부위	1082	MG102-2 TRAC B1 20nt	뉴클레오타이드	인공 서열	N/A	
MG102-2 인간 TRAC 표지 부위	1083	MG102-2 TRAC A1 24nt sgRNA	뉴클레오타이드	인공 서열	N/A	
MG102-2 인간 TRAC 표지 부위	1084	MG102-2 TRAC B1 24nt sgRNA	뉴클레오타이드	인공 서열	N/A	
MG102-2 인간 TRAC 표지 부위	1085	MG102-2 TRAC A1 20nt sgRNA	뉴클레오타이드	인공 서열	N/A	
MG102-2 인간 TRAC 표지 부위	1086	MG102-2 TRAC B1 20nt sgRNA	뉴클레오타이드	인공 서열	N/A	
AAVS1 유전체 표지화하는 MG102-2 가이드	1087	MG102-2 AAVS1 A5	뉴클레오타이드	인공 서열		

[0313]

별주	서열번호	설명	유형	유기체	기타 정보	모델트 또는 서열
AAVS1 표적화 MG102-2 카이드	1088	MG102-2 AAVS1 H8	뉴클레오타이드	인공 서열		
AAVS1 표적화 MG102-2 카이드	1089	MG102-2 AAVS1 H9	뉴클레오타이드	인공 서열		
AAVS1 표적화 MG102-2 카이드	1090	MG102-2 AAVS1 D11	뉴클레오타이드	인공 서열		
AAVS1 표적화 MG102-2 카이드	1091	MG102-2 AAVS1 E7	뉴클레오타이드	인공 서열		
AAVS1 표적화 MG102-2 카이드	1092	MG102-2 AAVS1 D7	뉴클레오타이드	인공 서열		
AAVS1 표적화 MG102-2 카이드	1093	MG102-2 AAVS1 B7	뉴클레오타이드	인공 서열		
AAVS1 표적화 MG102-2 카이드	1094	MG102-2 AAVS1 D12	뉴클레오타이드	인공 서열		
AAVS1 표적화 MG102-2 카이드	1095	MG102-2 AAVS1 C8	뉴클레오타이드	인공 서열		
AAVS1 표적화 MG102-2 카이드	1096	MG102-2 AAVS1 A8	뉴클레오타이드	인공 서열		
AAVS1 표적화 MG102-2 카이드	1097	MG102-2 AAVS1 G6	뉴클레오타이드	인공 서열		
AAVS1 표적화 MG102-2 카이드	1098	MG102-2 AAVS1 E5	뉴클레오타이드	인공 서열		
AAVS1 표적화 MG102-2 카이드	1099	MG102-2 AAVS1 G7	뉴클레오타이드	인공 서열		
AAVS1 표적화 MG102-2 카이드	1100	MG102-2 AAVS1 C3	뉴클레오타이드	인공 서열		
AAVS1 표적화 MG102-2 카이드	1101	MG102-2 AAVS1 E1	뉴클레오타이드	인공 서열		
AAVS1 표적화 MG102-2 카이드	1102	MG102-2 AAVS1 E2	뉴클레오타이드	인공 서열		
AAVS1 표적화 MG102-2 카이드	1103	MG102-2 AAVS1 H6	뉴클레오타이드	인공 서열		
AAVS1 표적화 MG102-2 카이드	1104	MG102-2 AAVS1 H11	뉴클레오타이드	인공 서열		

[0314]

별주	서열번호	설명	유형	유기체	기타 정보	모델트 또는 서열
AAVS1 표적 부위의 DNA 서열	1105	MG102-2 AAVS1 A5	뉴클레오타이드	인공 서열		
AAVS1 표적 부위의 DNA 서열	1106	MG102-2 AAVS1 H8	뉴클레오타이드	인공 서열		
AAVS1 표적 부위의 DNA 서열	1107	MG102-2 AAVS1 H9	뉴클레오타이드	인공 서열		
AAVS1 표적 부위의 DNA 서열	1108	MG102-2 AAVS1 D11	뉴클레오타이드	인공 서열		
AAVS1 표적 부위의 DNA 서열	1109	MG102-2 AAVS1 E7	뉴클레오타이드	인공 서열		
AAVS1 표적 부위의 DNA 서열	1110	MG102-2 AAVS1 D7	뉴클레오타이드	인공 서열		
AAVS1 표적 부위의 DNA 서열	1111	MG102-2 AAVS1 B7	뉴클레오타이드	인공 서열		
AAVS1 표적 부위의 DNA 서열	1112	MG102-2 AAVS1 D12	뉴클레오타이드	인공 서열		
AAVS1 표적 부위의 DNA 서열	1113	MG102-2 AAVS1 C8	뉴클레오타이드	인공 서열		
AAVS1 표적 부위의 DNA 서열	1114	MG102-2 AAVS1 A8	뉴클레오타이드	인공 서열		
AAVS1 표적 부위의 DNA 서열	1115	MG102-2 AAVS1 G6	뉴클레오타이드	인공 서열		
AAVS1 표적 부위의 DNA 서열	1116	MG102-2 AAVS1 E5	뉴클레오타이드	인공 서열		
AAVS1 표적 부위의 DNA 서열	1117	MG102-2 AAVS1 G7	뉴클레오타이드	인공 서열		
AAVS1 표적 부위의 DNA 서열	1118	MG102-2 AAVS1 C3	뉴클레오타이드	인공 서열		
AAVS1 표적 부위의 DNA 서열	1119	MG102-2 AAVS1 E1	뉴클레오타이드	인공 서열		
AAVS1 표적 부위의 DNA 서열	1120	MG102-2 AAVS1 E2	뉴클레오타이드	인공 서열		
AAVS1 표적 부위의 DNA 서열	1121	MG102-2 AAVS1 H6	뉴클레오타이드	인공 서열		

[0315]

별주	서열번호	설명	유형	유기체	기타 정보	포맷 또는 서열
AAVS1 표적 부위의 DNA 서열	1122	MG102-2 AAVS1 H11	뉴클레오타이드	인공 서열		
TRAC 표적화 MG102-36 가이드	1123	MG102-36 TRAC D12	뉴클레오타이드	인공 서열		
TRAC 표적화 MG102-36 가이드	1124	MG102-36 TRAC F1	뉴클레오타이드	인공 서열		
TRAC 표적화 MG102-36 가이드	1125	MG102-36 TRAC H6	뉴클레오타이드	인공 서열		
TRAC 표적화 MG102-39 가이드	1126	MG102-39 TRAC F4	뉴클레오타이드	인공 서열		
TRAC 표적화 MG102-39 가이드	1127	MG102-39 TRAC A9	뉴클레오타이드	인공 서열		
TRAC 표적화 MG102-39 가이드	1128	MG102-39 TRAC G11	뉴클레오타이드	인공 서열		
TRAC 표적화 MG102-39 가이드	1129	MG102-39 TRAC C11	뉴클레오타이드	인공 서열		
TRAC 표적화 MG102-39 가이드	1130	MG102-39 TRAC B6	뉴클레오타이드	인공 서열		
TRAC 표적화 MG102-39 가이드	1131	MG102-39 TRAC B5	뉴클레오타이드	인공 서열		
TRAC 표적화 MG102-39 가이드	1132	MG102-39 TRAC G9	뉴클레오타이드	인공 서열		
TRAC 표적화 MG102-39 가이드	1133	MG102-39 TRAC D1	뉴클레오타이드	인공 서열		
TRAC 표적화 MG102-39 가이드	1134	MG102-39 TRAC B11	뉴클레오타이드	인공 서열		
TRAC 표적화 MG102-39 가이드	1135	MG102-39 TRAC D4	뉴클레오타이드	인공 서열		
TRAC 표적화 MG102-39 가이드	1136	MG102-39 TRAC F2	뉴클레오타이드	인공 서열		
TRAC 표적화 MG102-39 가이드	1137	MG102-39 TRAC G1	뉴클레오타이드	인공 서열		
TRAC 표적화 MG102-42 가이드	1138	MG102-42 TRAC D10	뉴클레오타이드	인공 서열		

[0316]

첨주	서열번호	설명	유형	유기체	기타 정보	코멘트 또는 서열
TRAC 표적화 MG102-42 가이드	1139	MG102-42 TRAC D12	뉴클레오타이드	인공 서열		
TRAC 표적화 MG102-42 가이드	1140	MG102-42 TRAC E12	뉴클레오타이드	인공 서열		
TRAC 표적화 MG102-45 가이드	1141	MG102-45 TRAC B1	뉴클레오타이드	인공 서열		
TRAC 표적화 MG102-45 가이드	1142	MG102-45 TRAC C11	뉴클레오타이드	인공 서열		
TRAC 표적화 MG102-48 가이드	1143	MG102-48 TRAC A1	뉴클레오타이드	인공 서열		
TRAC 표적화 MG102-48 가이드	1144	MG102-48 TRAC D12	뉴클레오타이드	인공 서열		
TRAC 표적 부위의 DNA 서열	1145	MG102-36 TRAC D12	뉴클레오타이드	인공 서열		
TRAC 표적 부위의 DNA 서열	1146	MG102-36 TRAC F1	뉴클레오타이드	인공 서열		
TRAC 표적 부위의 DNA 서열	1147	MG102-36 TRAC H6	뉴클레오타이드	인공 서열		
TRAC 표적 부위의 DNA 서열	1148	MG102-39 TRAC F4	뉴클레오타이드	인공 서열		
TRAC 표적 부위의 DNA 서열	1149	MG102-39 TRAC A9	뉴클레오타이드	인공 서열		
TRAC 표적 부위의 DNA 서열	1150	MG102-39 TRAC G11	뉴클레오타이드	인공 서열		
TRAC 표적 부위의 DNA 서열	1151	MG102-39 TRAC C11	뉴클레오타이드	인공 서열		
TRAC 표적 부위의 DNA 서열	1152	MG102-39 TRAC B6	뉴클레오타이드	인공 서열		
TRAC 표적 부위의 DNA 서열	1153	MG102-39 TRAC B5	뉴클레오타이드	인공 서열		
TRAC 표적 부위의 DNA 서열	1154	MG102-39 TRAC G9	뉴클레오타이드	인공 서열		
TRAC 표적 부위의 DNA 서열	1155	MG102-39 TRAC D1	뉴클레오타이드	인공 서열		

[0317]

별주	시험번호	설명	유형	유기체	기타 정보	표본 또는 서열
TRAC 표적 부위의 DNA 서열	1156	MG102-39 TRAC B11	뉴클레오타이드	인공 서열		
TRAC 표적 부위의 DNA 서열	1157	MG102-39 TRAC D4	뉴클레오타이드	인공 서열		
TRAC 표적 부위의 DNA 서열	1158	MG102-39 TRAC F2	뉴클레오타이드	인공 서열		
TRAC 표적 부위의 DNA 서열	1159	MG102-39 TRAC G1	뉴클레오타이드	인공 서열		
TRAC 표적 부위의 DNA 서열	1160	MG102-42 TRAC D10	뉴클레오타이드	인공 서열		
TRAC 표적 부위의 DNA 서열	1161	MG102-42 TRAC D12	뉴클레오타이드	인공 서열		
TRAC 표적 부위의 DNA 서열	1162	MG102-42 TRAC E12	뉴클레오타이드	인공 서열		
TRAC 표적 부위의 DNA 서열	1163	MG102-45 TRAC B1	뉴클레오타이드	인공 서열		
TRAC 표적 부위의 DNA 서열	1164	MG102-45 TRAC C11	뉴클레오타이드	인공 서열		
TRAC 표적 부위의 DNA 서열	1165	MG102-48 TRAC A1	뉴클레오타이드	인공 서열		
TRAC 표적 부위의 DNA 서열	1166	MG102-48 TRAC D12	뉴클레오타이드	인공 서열		
TRAC 표적화 MG33-34 가이드	1167	MG33-34 TRAC F6	뉴클레오타이드	인공 서열		
TRAC 표적화 MG33-34 가이드	1168	MG33-34 TRAC E6	뉴클레오타이드	인공 서열		
TRAC 표적 부위의 DNA 서열	1169	MG33-34 TRAC F6	뉴클레오타이드	인공 서열		
TRAC 표적 부위의 DNA 서열	1170	MG33-34 TRAC E6	뉴클레오타이드	인공 서열		
표과기를 암호화하는 MG102 유전자좌	1171	표과기를 암호화하는 MG102-33 유전자좌	뉴클레오타이드	알려지지 않음	미배양 유기체	
표과기를 암호화하는 MG35 유전자좌	1172	표과기를 암호화하는 MG35-3 유전자좌	뉴클레오타이드	알려지지 않음	미배양 유기체	

[0318]

별주	서열번호	설명	유형	유기체	기타 정보	표본 또는 서열
효과기를 암호화하는 MG35 유전자와	1173	효과기를 암호화하는 MG35-7 유전자와	뉴클레오티드	알려지지 않음	미배양 유기체	
효과기를 암호화하는 MG35 유전자와	1174	효과기를 암호화하는 MG35-15 유전자와	뉴클레오티드	알려지지 않음	미배양 유기체	
효과기를 암호화하는 MG35 유전자와	1175	효과기를 암호화하는 MG35-20 유전자와	뉴클레오티드	알려지지 않음	미배양 유기체	
효과기를 암호화하는 MG35 유전자와	1176	효과기를 암호화하는 MG35-46 유전자와	뉴클레오티드	알려지지 않음	미배양 유기체	
효과기를 암호화하는 MG35 유전자와	1177	효과기를 암호화하는 MG35-58 유전자와	뉴클레오티드	알려지지 않음	미배양 유기체	
효과기를 암호화하는 MG35 유전자와	1178	효과기를 암호화하는 MG35-59 유전자와	뉴클레오티드	알려지지 않음	미배양 유기체	
효과기를 암호화하는 MG35 유전자와	1179	효과기를 암호화하는 MG35-76 유전자와	뉴클레오티드	알려지지 않음	미배양 유기체	
효과기를 암호화하는 MG35 유전자와	1180	효과기를 암호화하는 MG35-99 유전자와	뉴클레오티드	알려지지 않음	미배양 유기체	
효과기를 암호화하는 MG35 유전자와	1181	효과기를 암호화하는 MG35-100 유전자와	뉴클레오티드	알려지지 않음	미배양 유기체	
효과기를 암호화하는 MG35 유전자와	1182	효과기를 암호화하는 MG35-102 유전자와	뉴클레오티드	알려지지 않음	미배양 유기체	
효과기를 암호화하는 MG35 유전자와	1183	효과기를 암호화하는 MG35-103 유전자와	뉴클레오티드	알려지지 않음	미배양 유기체	
효과기를 암호화하는 MG35 유전자와	1184	효과기를 암호화하는 MG35-104 유전자와	뉴클레오티드	알려지지 않음	미배양 유기체	
효과기를 암호화하는 MG35 유전자와	1185	효과기를 암호화하는 MG35-114 유전자와	뉴클레오티드	알려지지 않음	미배양 유기체	
효과기를 암호화하는 MG35 유전자와	1186	효과기를 암호화하는 MG35-132 유전자와	뉴클레오티드	알려지지 않음	미배양 유기체	
효과기를 암호화하는 MG35 유전자와	1187	효과기를 암호화하는 MG35-168 유전자와	뉴클레오티드	알려지지 않음	미배양 유기체	
효과기를 암호화하는 MG35 유전자와	1188	효과기를 암호화하는 MG35-176 유전자와	뉴클레오티드	알려지지 않음	미배양 유기체	
효과기를 암호화하는 MG35 유전자와	1189	효과기를 암호화하는 MG35-177 유전자와	뉴클레오티드	알려지지 않음	미배양 유기체	

[0319]

별주	서열번호	설명	유형	유기체	기타 정보	코멘트 또는 서열
효과기를 암호화하는 MG35 유전자와	1190	효과기를 암호화하는 MG35-179 유전자와	뉴클레오타이드	알려지지 않음	미배양 유기체	
효과기를 암호화하는 MG35 유전자와	1191	효과기를 암호화하는 MG35-201 유전자와	뉴클레오타이드	알려지지 않음	미배양 유기체	
효과기를 암호화하는 MG35 유전자와	1192	효과기를 암호화하는 MG35-231 유전자와	뉴클레오타이드	알려지지 않음	미배양 유기체	
효과기를 암호화하는 MG35 유전자와	1193	효과기를 암호화하는 MG35-232 유전자와	뉴클레오타이드	알려지지 않음	미배양 유기체	
효과기를 암호화하는 MG35 유전자와	1194	효과기를 암호화하는 MG35-233 유전자와	뉴클레오타이드	알려지지 않음	미배양 유기체	
효과기를 암호화하는 MG35 유전자와	1195	효과기를 암호화하는 MG35-237 유전자와	뉴클레오타이드	알려지지 않음	미배양 유기체	
효과기를 암호화하는 MG35 유전자와	1196	효과기를 암호화하는 MG35-238 유전자와	뉴클레오타이드	알려지지 않음	미배양 유기체	
효과기를 암호화하는 MG35 유전자와	1197	효과기를 암호화하는 MG35-240 유전자와	뉴클레오타이드	알려지지 않음	미배양 유기체	
효과기를 암호화하는 MG35 유전자와	1198	효과기를 암호화하는 MG35-291 유전자와	뉴클레오타이드	알려지지 않음	미배양 유기체	
효과기를 암호화하는 MG35 유전자와	1199	효과기를 암호화하는 MG35-296 유전자와	뉴클레오타이드	알려지지 않음	미배양 유기체	
효과기를 암호화하는 MG35 유전자와	1200	효과기를 암호화하는 MG35-298 유전자와	뉴클레오타이드	알려지지 않음	미배양 유기체	
효과기를 암호화하는 MG35 유전자와	1201	효과기를 암호화하는 MG35-299 유전자와	뉴클레오타이드	알려지지 않음	미배양 유기체	
효과기를 암호화하는 MG35 유전자와	1202	효과기를 암호화하는 MG35-302 유전자와	뉴클레오타이드	알려지지 않음	미배양 유기체	
효과기를 암호화하는 MG35 유전자와	1203	효과기를 암호화하는 MG35-309 유전자와	뉴클레오타이드	알려지지 않음	미배양 유기체	
효과기를 암호화하는 MG35 유전자와	1204	효과기를 암호화하는 MG35-323 유전자와	뉴클레오타이드	알려지지 않음	미배양 유기체	
효과기를 암호화하는 MG35 유전자와	1205	효과기를 암호화하는 MG35-326 유전자와	뉴클레오타이드	알려지지 않음	미배양 유기체	
효과기를 암호화하는 MG35 유전자와	1206	효과기를 암호화하는 MG35-337 유전자와	뉴클레오타이드	알려지지 않음	미배양 유기체	

[0320]

별주	서열번호	설명	유형	유기체	기타 정보	코멘트 또는 서열
효과기를 암호화하는 MG35 유전자와	1207	효과기를 암호화하는 MG35-339 유전자와	뉴클레오타이드	알려지지 않음	미배양 유기체	
효과기를 암호화하는 MG35 유전자와	1208	효과기를 암호화하는 MG35-339 유전자와	뉴클레오타이드	알려지지 않음	미배양 유기체	
효과기를 암호화하는 MG35 유전자와	1209	효과기를 암호화하는 MG35-344 유전자와	뉴클레오타이드	알려지지 않음	미배양 유기체	
효과기를 암호화하는 MG35 유전자와	1210	효과기를 암호화하는 MG35-345 유전자와	뉴클레오타이드	알려지지 않음	미배양 유기체	
효과기를 암호화하는 MG35 유전자와	1211	효과기를 암호화하는 MG35-346 유전자와	뉴클레오타이드	알려지지 않음	미배양 유기체	
효과기를 암호화하는 MG35 유전자와	1212	효과기를 암호화하는 MG35-347 유전자와	뉴클레오타이드	알려지지 않음	미배양 유기체	
효과기를 암호화하는 MG35 유전자와	1213	효과기를 암호화하는 MG35-348 유전자와	뉴클레오타이드	알려지지 않음	미배양 유기체	
효과기를 암호화하는 MG35 유전자와	1214	효과기를 암호화하는 MG35-349 유전자와	뉴클레오타이드	알려지지 않음	미배양 유기체	
효과기를 암호화하는 MG35 유전자와	1215	효과기를 암호화하는 MG35-350 유전자와	뉴클레오타이드	알려지지 않음	미배양 유기체	
효과기를 암호화하는 MG35 유전자와	1216	효과기를 암호화하는 MG35-354 유전자와	뉴클레오타이드	알려지지 않음	미배양 유기체	
효과기를 암호화하는 MG35 유전자와	1217	효과기를 암호화하는 MG35-357 유전자와	뉴클레오타이드	알려지지 않음	미배양 유기체	
효과기를 암호화하는 MG35 유전자와	1218	효과기를 암호화하는 MG35-358 유전자와	뉴클레오타이드	알려지지 않음	미배양 유기체	
효과기를 암호화하는 MG35 유전자와	1219	효과기를 암호화하는 MG35-359 유전자와	뉴클레오타이드	알려지지 않음	미배양 유기체	
효과기를 암호화하는 MG35 유전자와	1220	효과기를 암호화하는 MG35-364 유전자와	뉴클레오타이드	알려지지 않음	미배양 유기체	
효과기를 암호화하는 MG35 유전자와	1221	효과기를 암호화하는 MG35-366 유전자와	뉴클레오타이드	알려지지 않음	미배양 유기체	
효과기를 암호화하는 MG35 유전자와	1222	효과기를 암호화하는 MG35-393 유전자와	뉴클레오타이드	알려지지 않음	미배양 유기체	
효과기를 암호화하는 MG35 유전자와	1223	효과기를 암호화하는 MG35-404 유전자와	뉴클레오타이드	알려지지 않음	미배양 유기체	
효과기를 암호화하는 MG35 유전자와		효과기를 암호화하는 MG35-411 유전자와	뉴클레오타이드	알려지지 않음	미배양 유기체	

[0321]

별주	서열번호	설명	유형	유기체	기타 정보	모델 또는 서열
효과기를 암호화하는 MG35 유전자와	1224	효과기를 암호화하는 MG35-418 유전자와	뉴클레오타이드	일려지지 않음	미배양 유기체	
효과기를 암호화하는 MG35 유전자와	1225	효과기를 암호화하는 MG35-419 유전자와	뉴클레오타이드	일려지지 않음	미배양 유기체	
효과기를 암호화하는 MG35 유전자와	1226	효과기를 암호화하는 MG35-420 유전자와	뉴클레오타이드	일려지지 않음	미배양 유기체	
효과기를 암호화하는 MG35 유전자와	1227	효과기를 암호화하는 MG35-421 유전자와	뉴클레오타이드	일려지지 않음	미배양 유기체	
효과기를 암호화하는 MG35 유전자와	1228	효과기를 암호화하는 MG35-423 유전자와	뉴클레오타이드	일려지지 않음	미배양 유기체	
MG35 예측 CRISPR 반복	1229	MG35-463 CRISPR 반복	뉴클레오타이드	인공 서열		
MG35 예측 CRISPR 반복	1230	MG35-556 CRISPR 반복	뉴클레오타이드	인공 서열		
MG35 활성 효과기 sgRNA	1231	MG35-94 sg1 단일 가이드 RNA	뉴클레오타이드	인공 서열		
MG35 활성 효과기 sgRNA	1232	MG35-94 sg2 단일 가이드 RNA	뉴클레오타이드	인공 서열		
MG35 활성 효과기 sgRNA	1233	MG35-94 sg3 단일 가이드 RNA	뉴클레오타이드	인공 서열		
MG35 활성 효과기 sgRNA	1234	MG35-104 sg1 단일 가이드 RNA	뉴클레오타이드	인공 서열		
MG35 활성 효과기 sgRNA	1235	MG35-350 sg3 단일 가이드 RNA	뉴클레오타이드	인공 서열		
MG35 활성 효과기 sgRNA	1236	MG35-463 sg2 단일 가이드 RNA	뉴클레오타이드	인공 서열		
MG35 활성 효과기 sgRNA	1237	MG35-463 sg3 단일 가이드 RNA	뉴클레오타이드	인공 서열		
MG35 활성 효과기 sgRNA	1238	MG35-515 sg2 단일 가이드 RNA	뉴클레오타이드	인공 서열		
MG35 활성 효과기 sgRNA	1239	MG35-515 sg3 단일 가이드 RNA	뉴클레오타이드	인공 서열		
MG35 활성 효과기 sgRNA	1240	MG35-517 sg2 단일 가이드 RNA	뉴클레오타이드	인공 서열		

[0322]

번호	서열번호	설명	유형	유기체	기타 정보	표본 또는 서열
MG35 활성 효과기 sgRNA	1241	MG35-518 sg1 단일 가이드 RNA	뉴클레오타이드	인공 서열		
MG35 활성 효과기 sgRNA	1242	MG35-519 sg1 단일 가이드 RNA	뉴클레오타이드	인공 서열		
MG35 활성 효과기 sgRNA	1243	MG35-519 sg2 단일 가이드 RNA	뉴클레오타이드	인공 서열		
MG35 활성 효과기 sgRNA	1244	MG35-519 sg3 단일 가이드 RNA	뉴클레오타이드	인공 서열		
MG35 활성 효과기 sgRNA	1245	MG35-550 sg1 단일 가이드 RNA	뉴클레오타이드	인공 서열		
MG35 활성 효과기 sgRNA	1246	MG35-553 sg1 단일 가이드 RNA	뉴클레오타이드	인공 서열		
MG35 활성 효과기 sgRNA	1247	MG35-554 sg3 단일 가이드 RNA	뉴클레오타이드	인공 서열		
MG35 활성 효과기 sgRNA	1248	MG35-554 sg4 단일 가이드 RNA	뉴클레오타이드	인공 서열		
MG35 활성 효과기 sgRNA	1249	MG35-104_M1 단일 가이드 RNA	뉴클레오타이드	인공 서열		
MG35 활성 효과기 sgRNA	1250	MG35-104_M2 단일 가이드 RNA	뉴클레오타이드	인공 서열		
MG35 활성 효과기 sgRNA	1251	MG35-104_M3 단일 가이드 RNA	뉴클레오타이드	인공 서열		
MG35 활성 효과기 sgRNA	1252	MG35-518_M1 단일 가이드 RNA	뉴클레오타이드	인공 서열		
MG35 활성 효과기 sgRNA	1253	MG35-518_M2 단일 가이드 RNA	뉴클레오타이드	인공 서열		
MG35 활성 효과기 sgRNA	1254	MG35-518_M3 단일 가이드 RNA	뉴클레오타이드	인공 서열		
MG35 활성 효과기 sgRNA	1255	MG35-553_M1 단일 가이드 RNA	뉴클레오타이드	인공 서열		
MG35 활성 효과기 sgRNA	1256	MG35-553_M2 단일 가이드 RNA	뉴클레오타이드	인공 서열		
MG35 활성 효과기 sgRNA	1257	MG35-553_M3 단일 가이드 RNA	뉴클레오타이드	인공 서열		

[0323]

별주	서열번호	설명	유형	유기체	기타 정보	표본 또는 서열
MG35 활성 효과기 sgRNA	1258	MG35-94_M2 단일 가이드 RNA	뉴클레오타이드	인공 서열		
MG35 활성 효과기 sgRNA	1259	MG35-94_M3 단일 가이드 RNA	뉴클레오타이드	인공 서열		
MG102 효과기	1260	MG102-49 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	1261	MG102-50 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	1262	MG102-51 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	1263	MG102-52 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	1264	MG102-53 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	1265	MG102-54 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	1266	MG102-55 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	1267	MG102-56 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	1268	MG102-57 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	1269	MG102-58 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	1270	MG102-59 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	1271	MG102-60 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	1272	MG102-61 효과기	단백질	알려지지 않음	미배양 유기체	
MG102 효과기	1273	MG102-62 효과기	단백질	알려지지 않음	미배양 유기체	
MG144 효과기	1274	MG144-5 효과기	단백질	알려지지 않음	미배양 유기체	
MG144 효과기	1275	MG144-6 효과기	단백질	알려지지 않음	미배양 유기체	
MG144 효과기	1276	MG144-7 효과기	단백질	알려지지 않음	미배양 유기체	
MG144 효과기	1277	MG144-8 효과기	단백질	알려지지 않음	미배양 유기체	
MG144 효과기	1278	MG144-9 효과기	단백질	알려지지 않음	미배양 유기체	
MG144 효과기	1279	MG144-10 효과기	단백질	알려지지 않음	미배양 유기체	
MG144 효과기	1280	MG144-11 효과기	단백질	알려지지 않음	미배양 유기체	
MG144 효과기	1281	MG144-12 효과기	단백질	알려지지 않음	미배양 유기체	
MG144 효과기	1282	MG144-13 효과기	단백질	알려지지 않음	미배양 유기체	
MG144 효과기	1283	MG144-14 효과기	단백질	알려지지 않음	미배양 유기체	
MG144 효과기	1284	MG144-15 효과기	단백질	알려지지 않음	미배양 유기체	
MG144 효과기	1285	MG144-16 효과기	단백질	알려지지 않음	미배양 유기체	

[0324]

별주	서열번호	설명	유형	유기체	기타 정보	코멘트 또는 서열
MG144 효과기	1286	MG144-17 효과기	단백질	알려지지 않음	미배양 유기체	
MG144 효과기	1287	MG144-18 효과기	단백질	알려지지 않음	미배양 유기체	
MG144 효과기	1288	MG144-19 효과기	단백질	알려지지 않음	미배양 유기체	
MG33 효과기	1289	MG33-36 효과기	단백질	알려지지 않음	미배양 유기체	
MG33 효과기	1290	MG33-37 효과기	단백질	알려지지 않음	미배양 유기체	
MG33 효과기	1291	MG33-38 효과기	단백질	알려지지 않음	미배양 유기체	
MG33 효과기	1292	MG33-39 효과기	단백질	알려지지 않음	미배양 유기체	
MG33 효과기	1293	MG33-40 효과기	단백질	알려지지 않음	미배양 유기체	
MG33 효과기	1294	MG33-41 효과기	단백질	알려지지 않음	미배양 유기체	
MG33 효과기	1295	MG33-42 효과기	단백질	알려지지 않음	미배양 유기체	
MG33 효과기	1296	MG33-43 효과기	단백질	알려지지 않음	미배양 유기체	
MG33 효과기	1297	MG33-44 효과기	단백질	알려지지 않음	미배양 유기체	
MG33 효과기	1298	MG33-45 효과기	단백질	알려지지 않음	미배양 유기체	
MG33 효과기	1299	MG33-46 효과기	단백질	알려지지 않음	미배양 유기체	
MG33 효과기	1300	MG33-47 효과기	단백질	알려지지 않음	미배양 유기체	
MG33 효과기	1301	MG33-48 효과기	단백질	알려지지 않음	미배양 유기체	
MG33 효과기	1302	MG33-49 효과기	단백질	알려지지 않음	미배양 유기체	
MG33 효과기	1303	MG33-50 효과기	단백질	알려지지 않음	미배양 유기체	
MG33 효과기	1304	MG33-51 효과기	단백질	알려지지 않음	미배양 유기체	
MG33 효과기	1305	MG33-52 효과기	단백질	알려지지 않음	미배양 유기체	
MG33 효과기	1306	MG33-53 효과기	단백질	알려지지 않음	미배양 유기체	
MG33 효과기	1307	MG33-54 효과기	단백질	알려지지 않음	미배양 유기체	
MG33 효과기	1308	MG33-55 효과기	단백질	알려지지 않음	미배양 유기체	
MG33 효과기	1309	MG33-56 효과기	단백질	알려지지 않음	미배양 유기체	
MG33 효과기	1310	MG33-57 효과기	단백질	알려지지 않음	미배양 유기체	
MG33 효과기	1311	MG33-58 효과기	단백질	알려지지 않음	미배양 유기체	
MG33 효과기	1312	MG33-59 효과기	단백질	알려지지 않음	미배양 유기체	
MG34 효과기	1313	MG34-26 효과기	단백질	알려지지 않음	미배양 유기체	
MG34 효과기	1314	MG34-27 효과기	단백질	알려지지 않음	미배양 유기체	

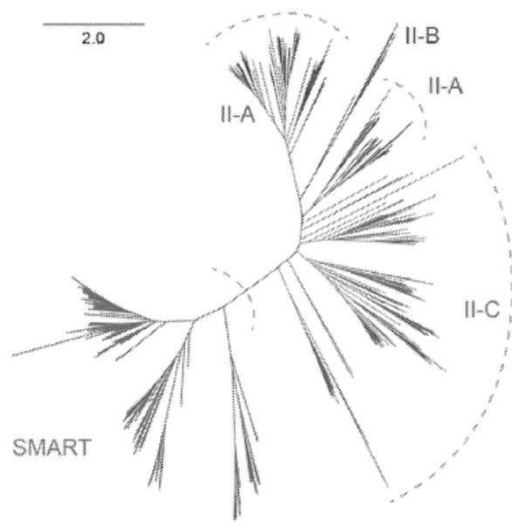
[0325]

발주	서열번호	설명	유형	유기체	기타 정보	코멘트 또는 서열
MG34 효과기	1315	MG34-28 효과기	단백질	알려지지 않음	미배양 유기체	
MG34 효과기	1316	MG34-29 효과기	단백질	알려지지 않음	미배양 유기체	
MG34 효과기	1317	MG34-30 효과기	단백질	알려지지 않음	미배양 유기체	
MG34 효과기	1318	MG34-31 효과기	단백질	알려지지 않음	미배양 유기체	
MG34 효과기	1319	MG34-32 효과기	단백질	알려지지 않음	미배양 유기체	
MG34 효과기	1320	MG34-33 효과기	단백질	알려지지 않음	미배양 유기체	
MG34 효과기	1321	MG34-34 효과기	단백질	알려지지 않음	미배양 유기체	

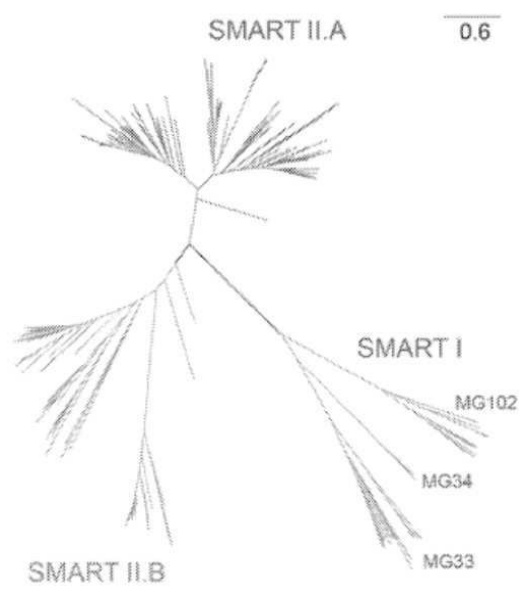
본 발명의 바람직한 실시예가 본원에 도시되고 기술되었지만, 이러한 실시예는 단지 예시로서 제공된다는 것은 당업자에게 명백할 것이다. 본 발명은 본 명세서 내에 제공된 특정 예에 의해 한정되는 것으로 의도되지 않는다. 본 발명은 전술한 명세서를 참조하여 기술되었지만, 본원의 실시예의 설명 및 예시는 한정적인 의미로 해석되는 것을 의미하지는 않는다. 이제 본 발명을 벗어나지 않고도 많은 변이, 변화, 및 치환이 당업자에게 일어날 것이다. 또한, 본 발명의 모든 측면은 다양한 조건 및 변수에 따라 달라지는 본원에 제시된 특정 도시, 구성, 또는 상대 비율로 한정되지 않음을 이해할 것이다. 본원에 기술된 본 발명의 실시예에 대한 다양한 대안이 본 발명을 실시하는 데 사용될 수 있음을 이해해야 한다. 따라서, 본 발명은 임의의 이러한 대안, 변형, 변이, 또는 균등물도 포괄하는 것으로 고려된다. 다음의 청구범위는 본 발명의 범위를 정의하고, 이들 청구범위의 범위에 속하는 방법 및 구조와 이들의 등가물이 이에 의해 포괄되는 것으로 의도된다.

도면

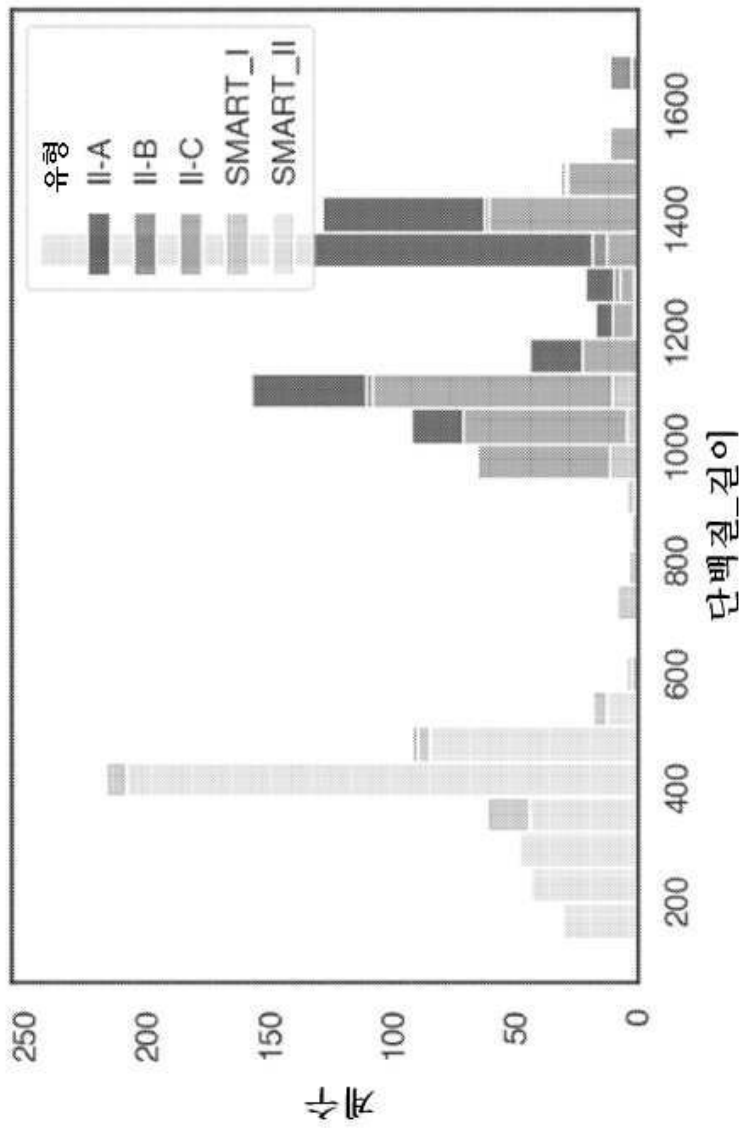
도면1a



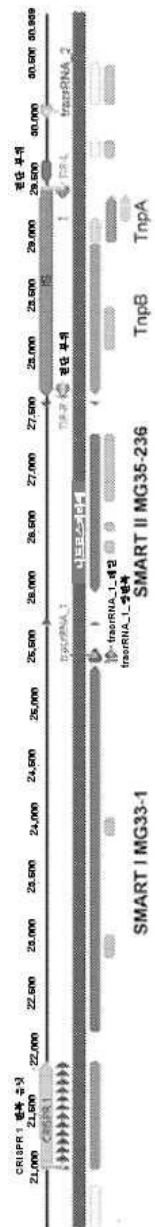
도면1b



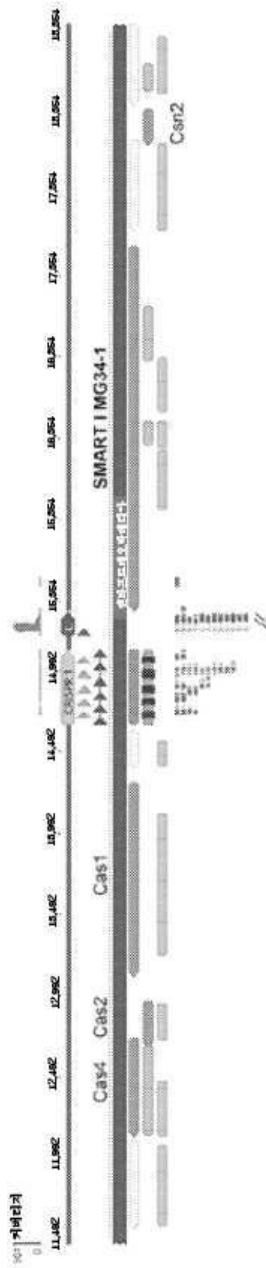
도면2



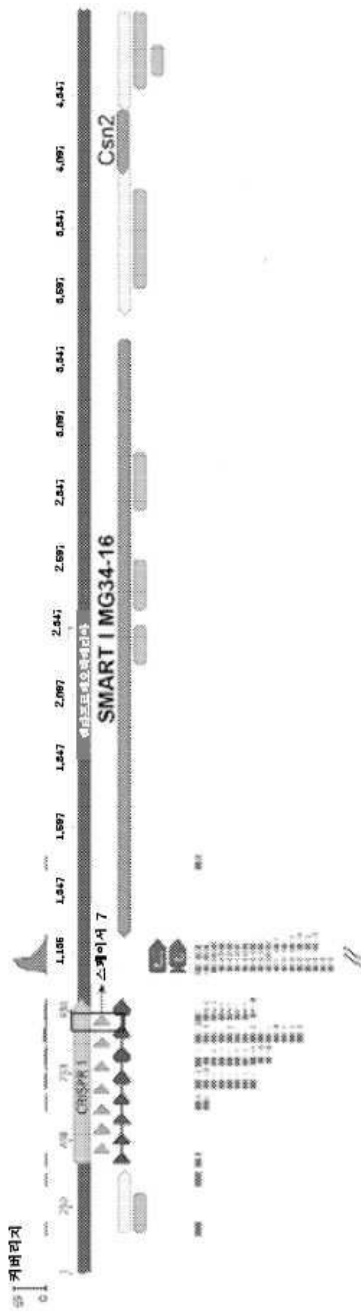
도면3a



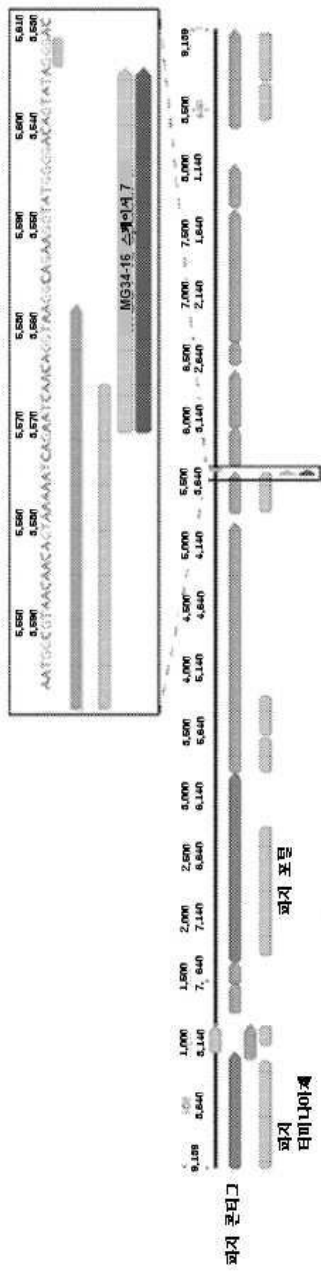
도면 3b



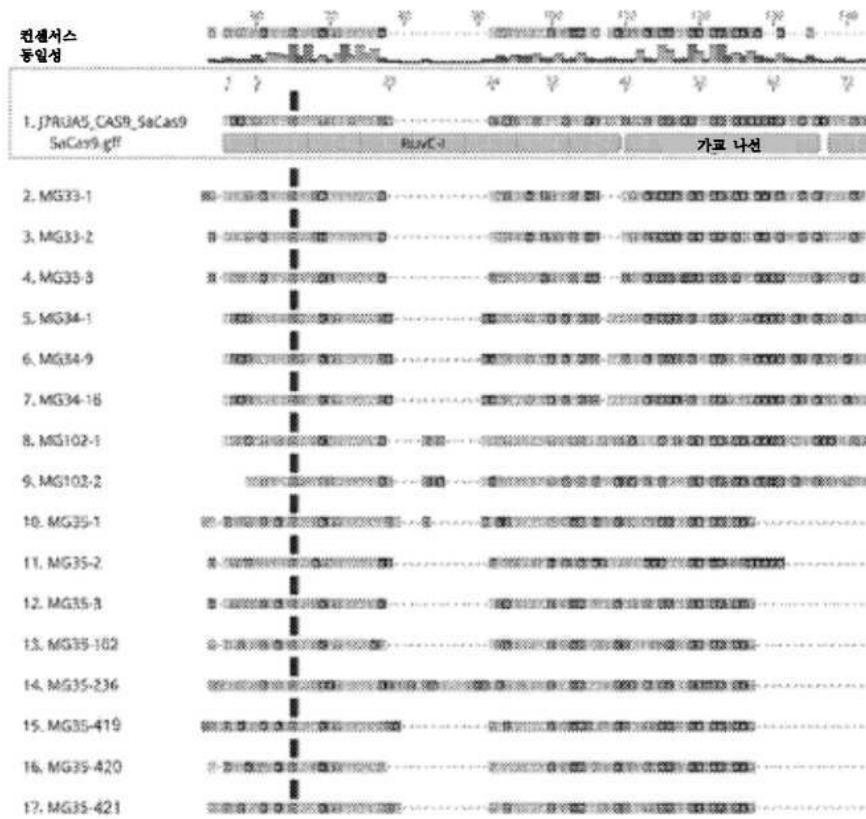
도면3c



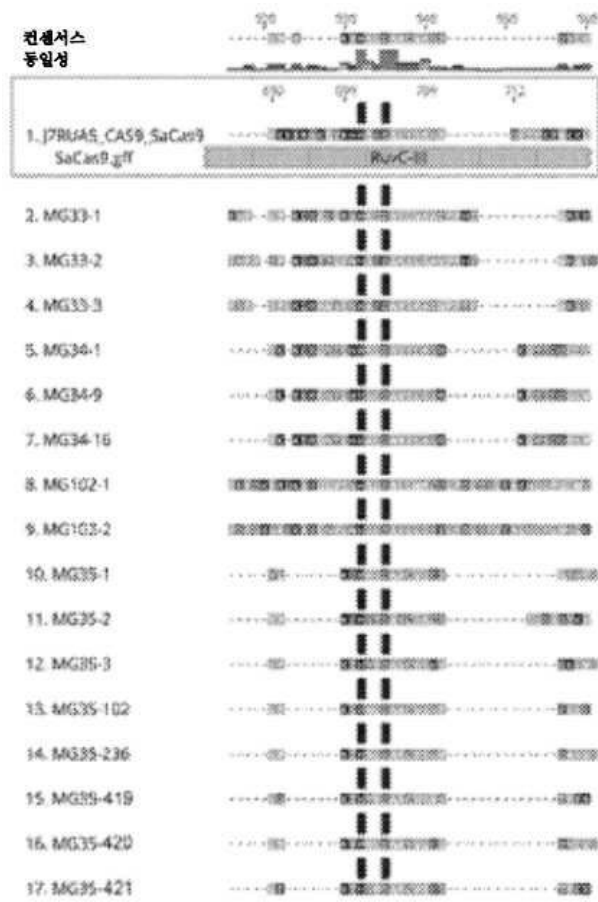
도면 3d



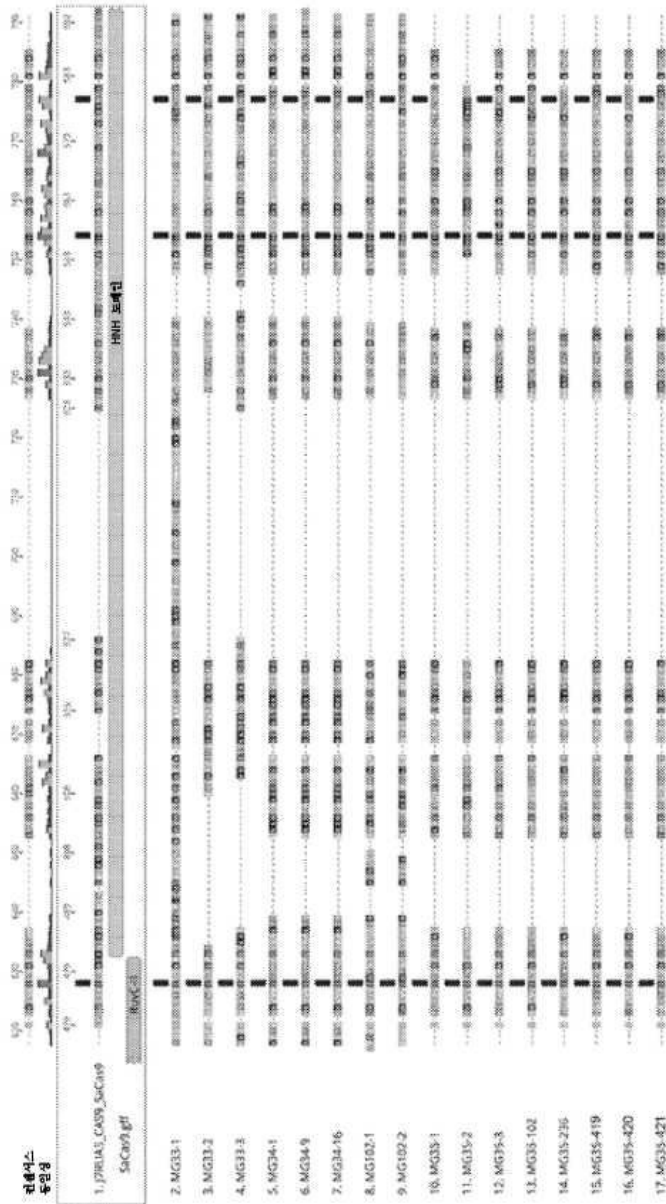
도면4a



도면4b



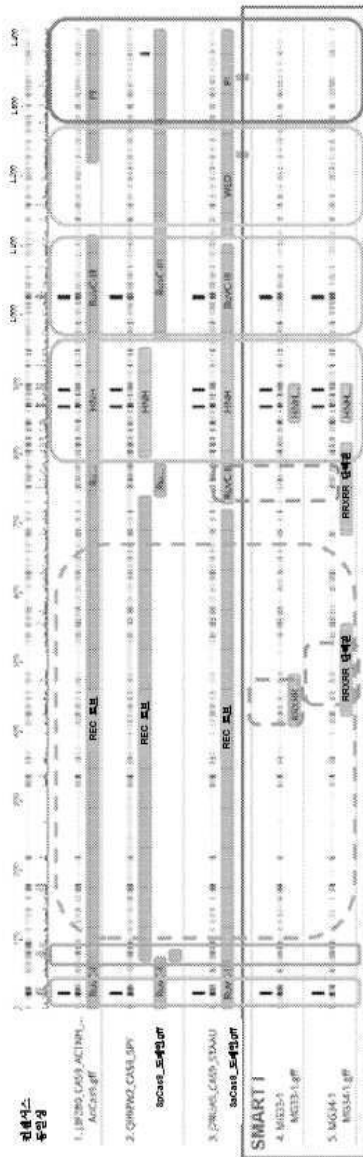
도면4c



도면5a

중대형 중형 소형 미니형	MG34-1 도메인								
	RuCo-I	BH	REC	F1025가 AMB와 연동 가능	REC	RuCo-II (외부)	HNH	RuCo-III	WED

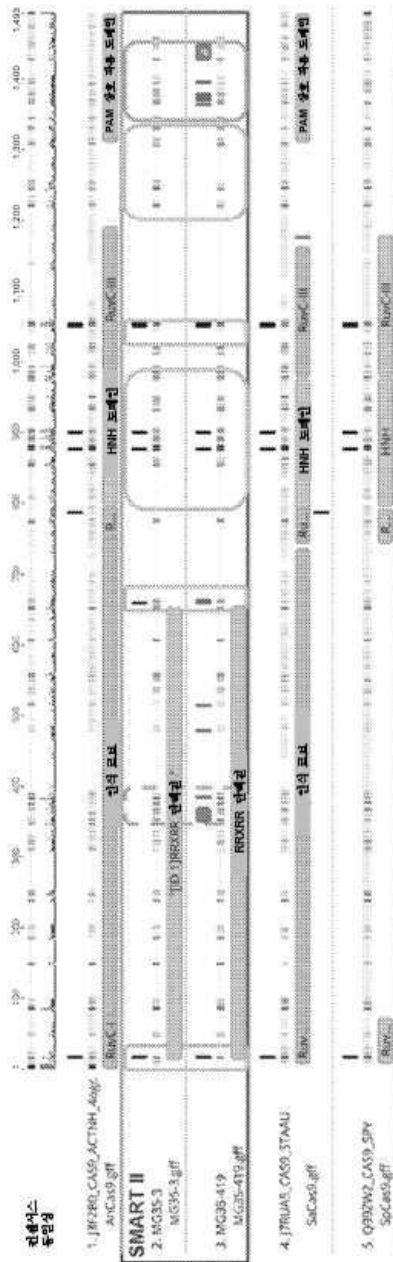
도면 5b



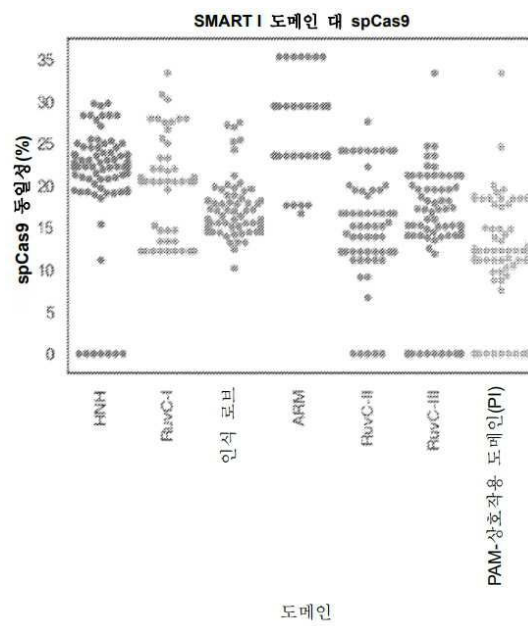
도면6a



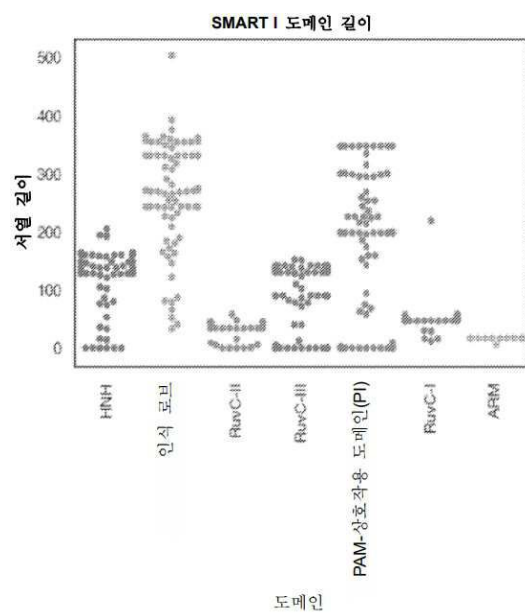
도면 6b



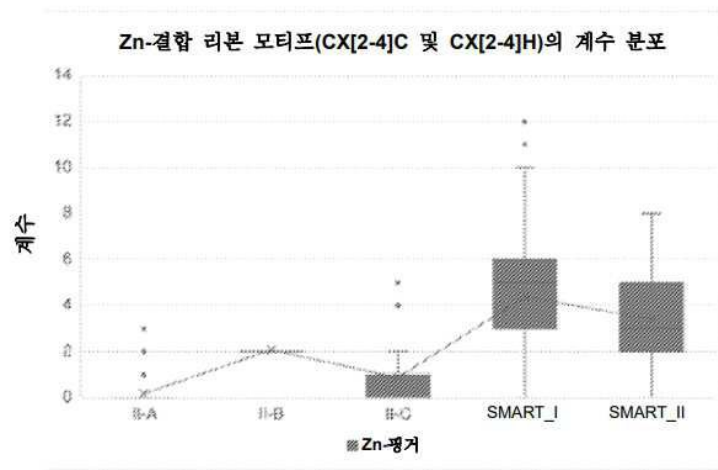
도면7a



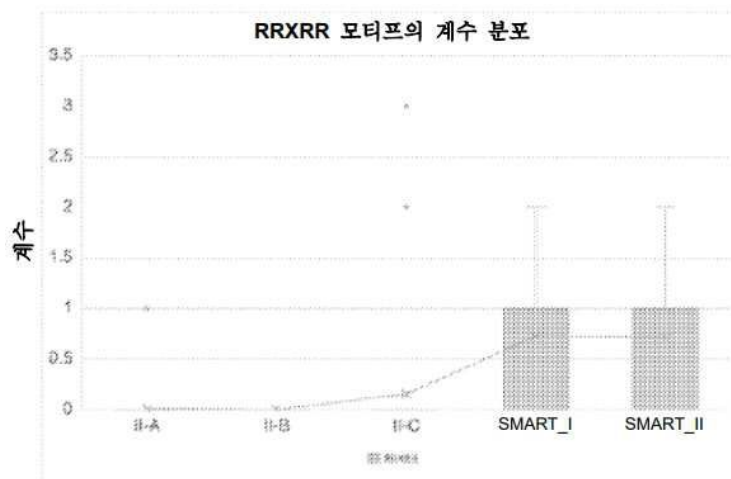
도면7b



도면8a



도면8b



도면9a



도면9b



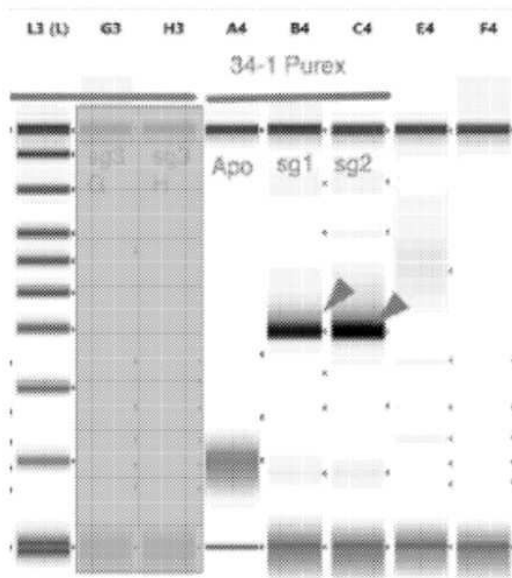
도면9c



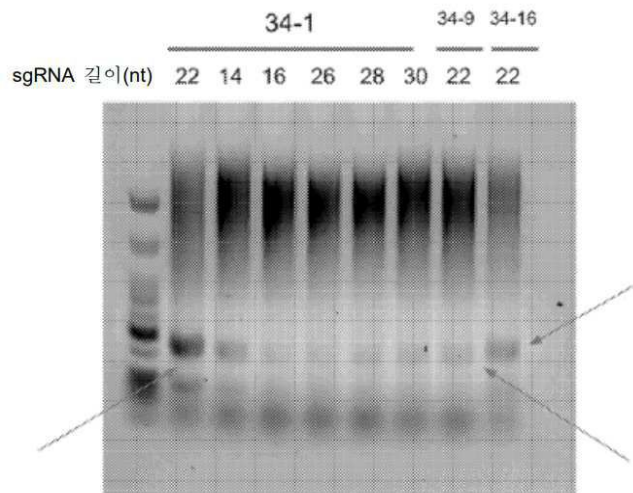
도면9d



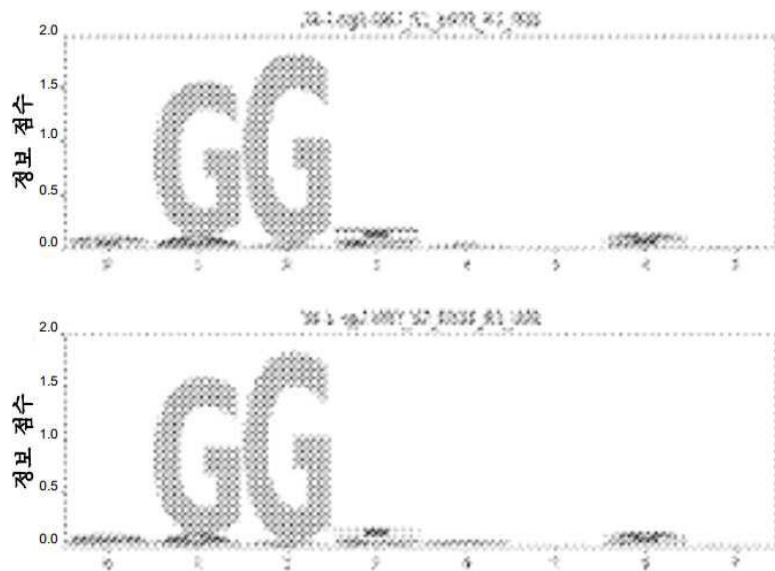
도면10a



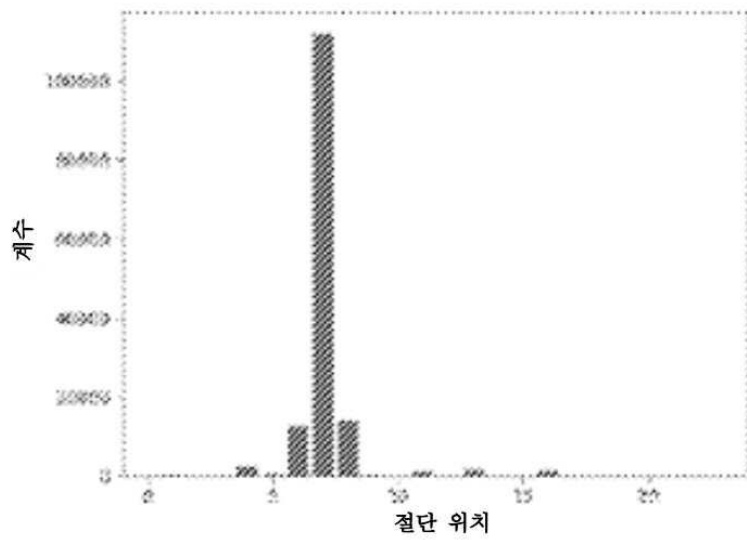
도면10b



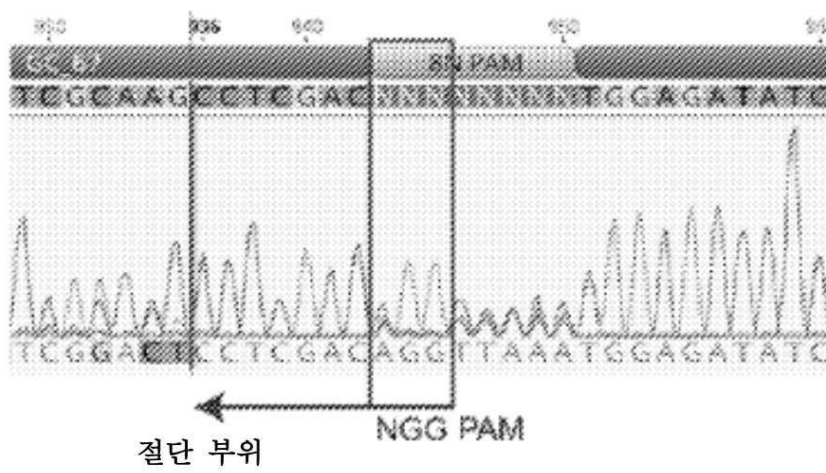
도면11a



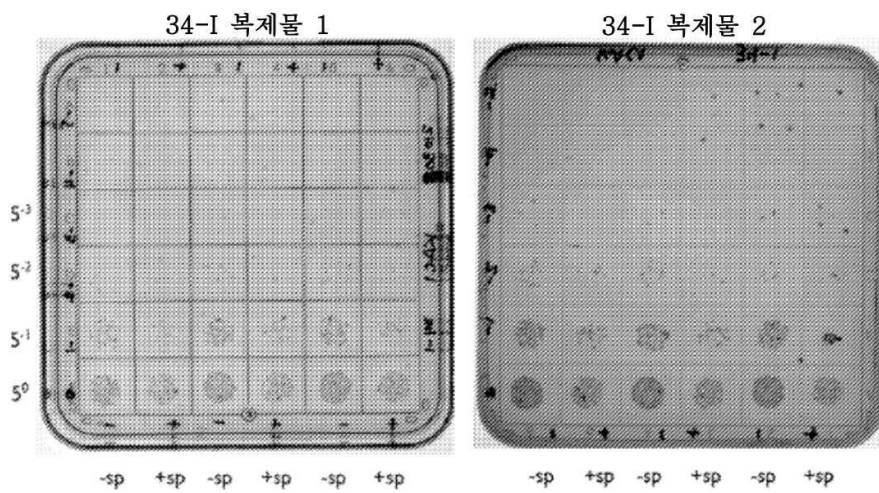
도면11b



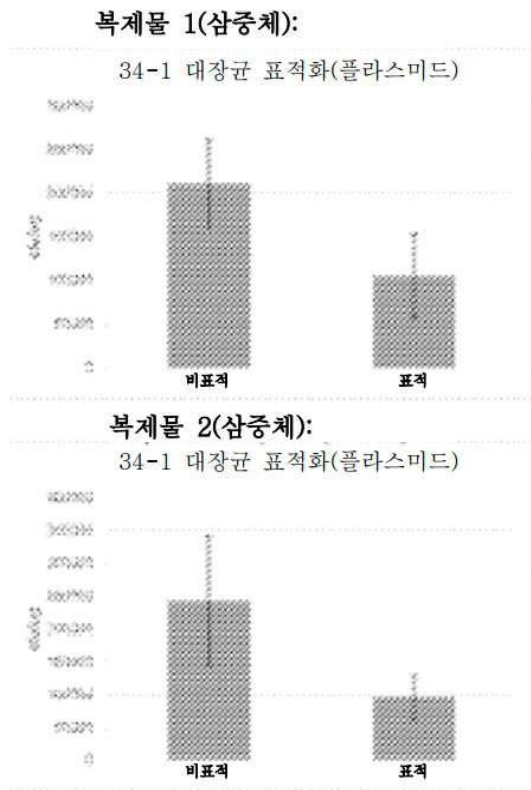
도면11c



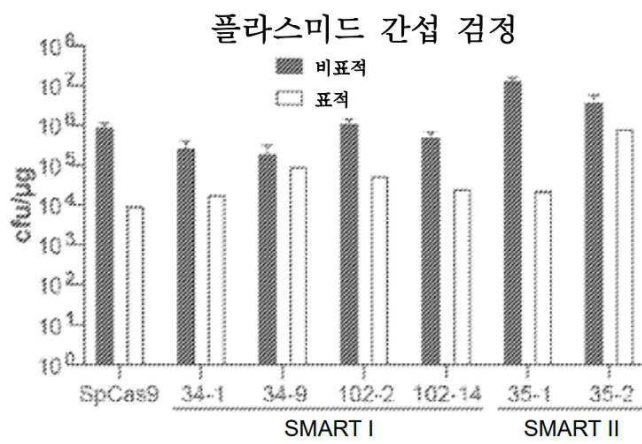
도면12a



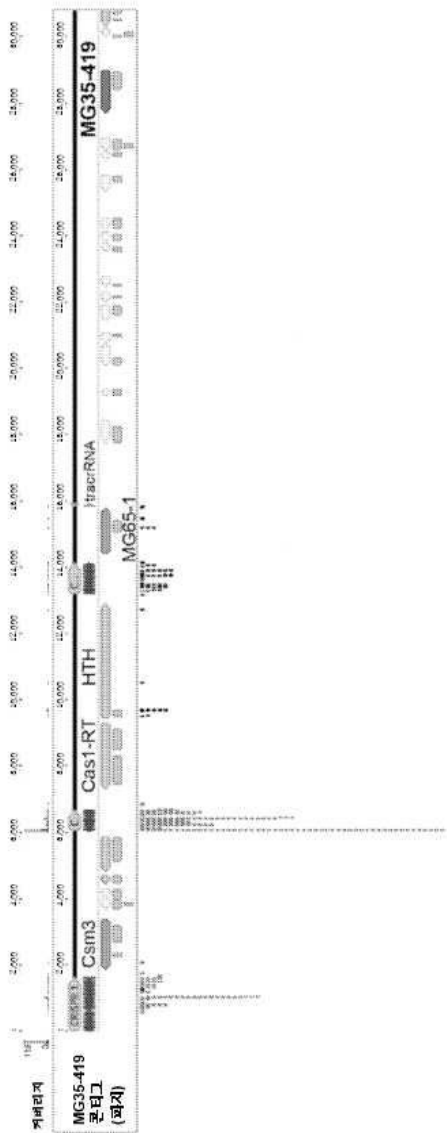
도면12b



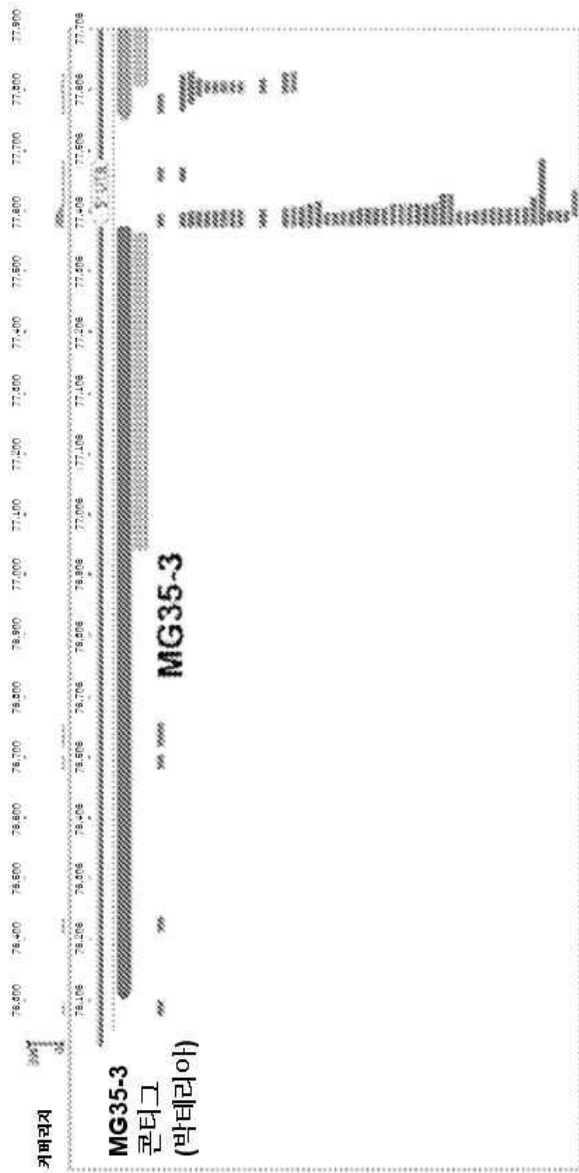
도면12c



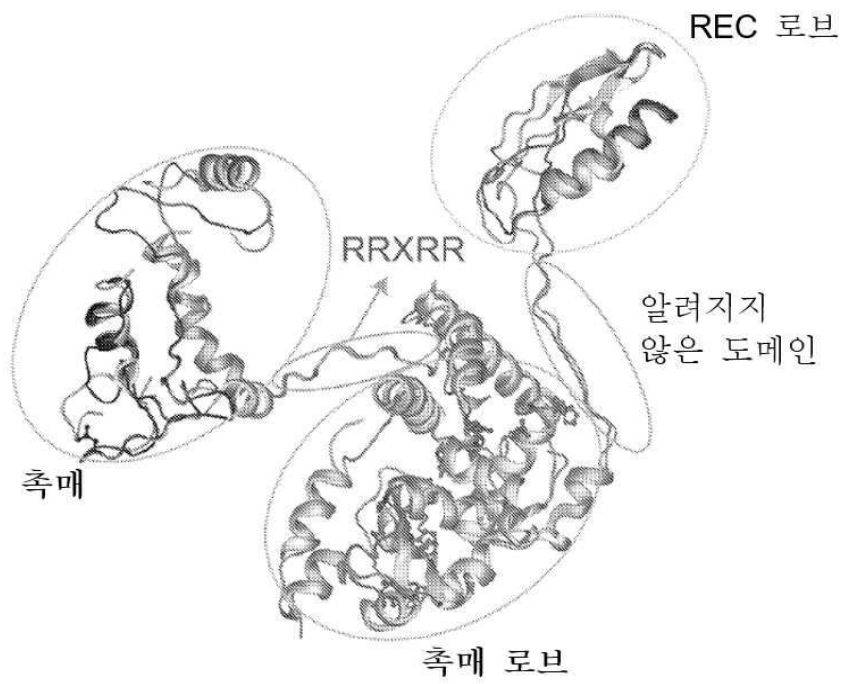
도면13a



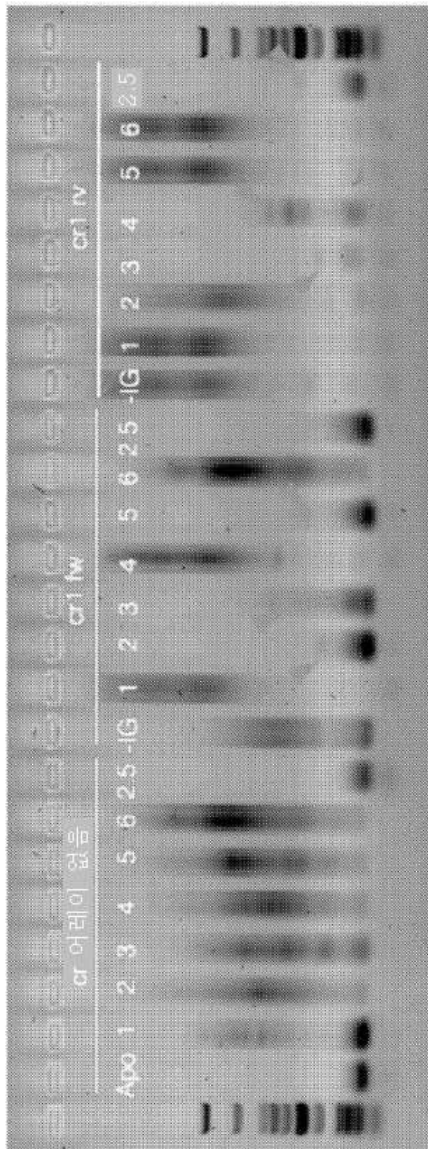
도면 13b



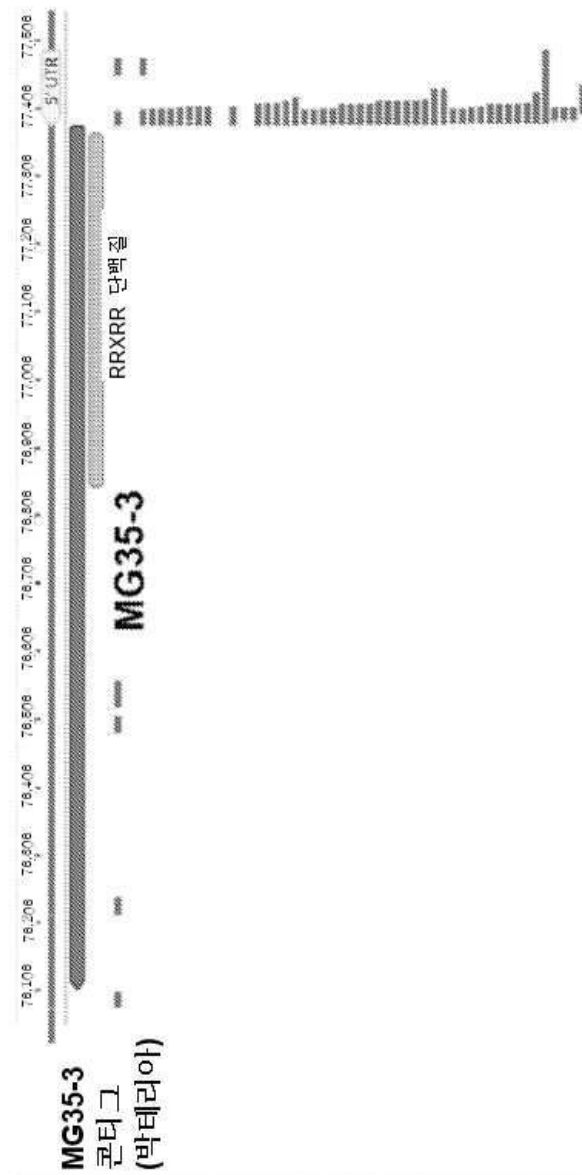
도면14



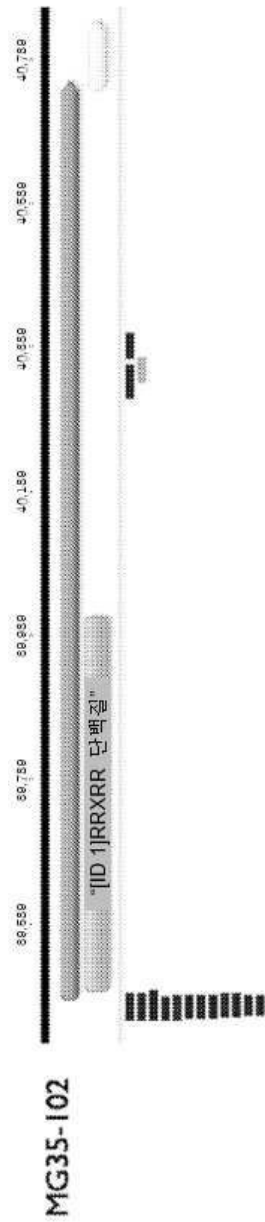
도면 15



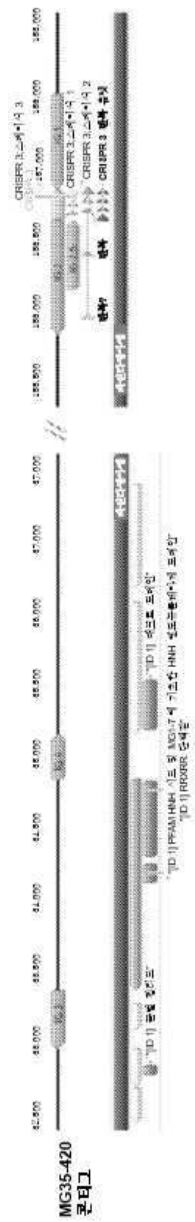
도면16a



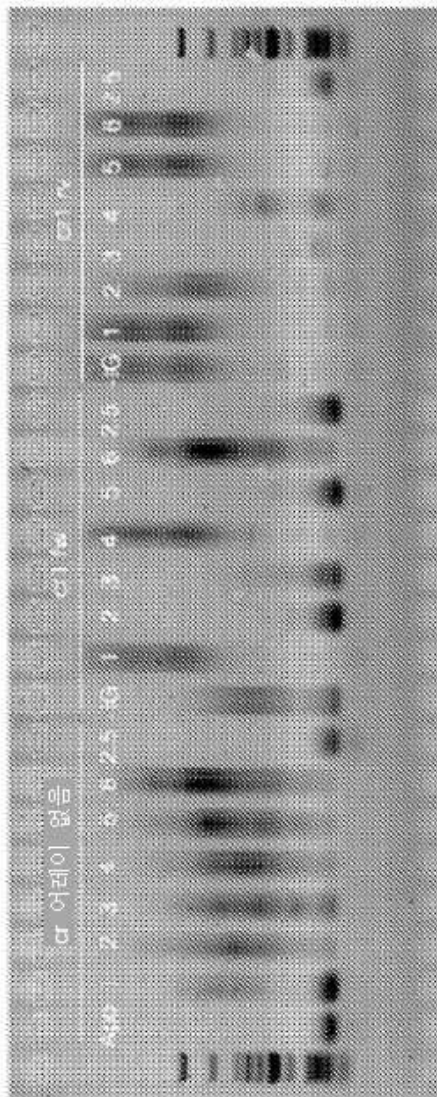
도면16b



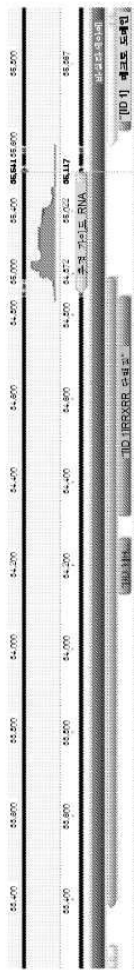
도면17a



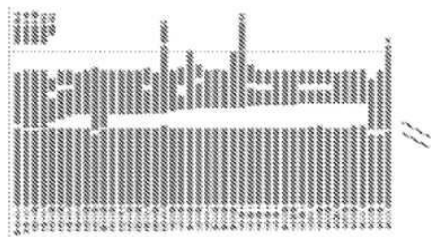
도면17b



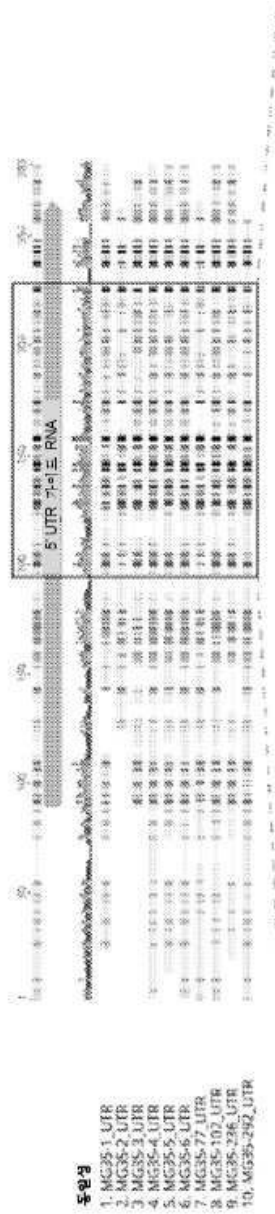
도면18a



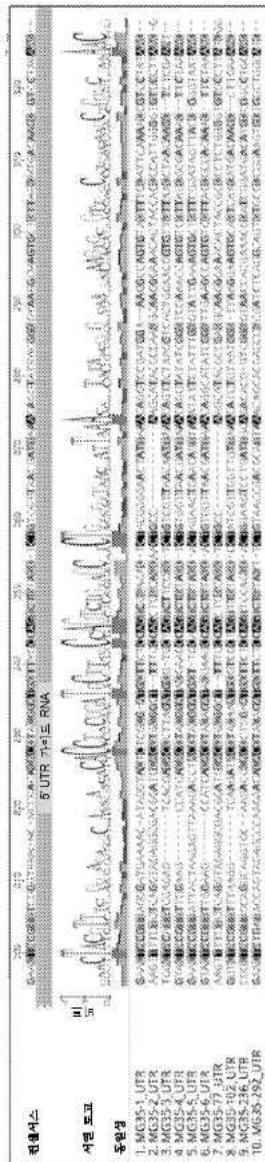
도면18b



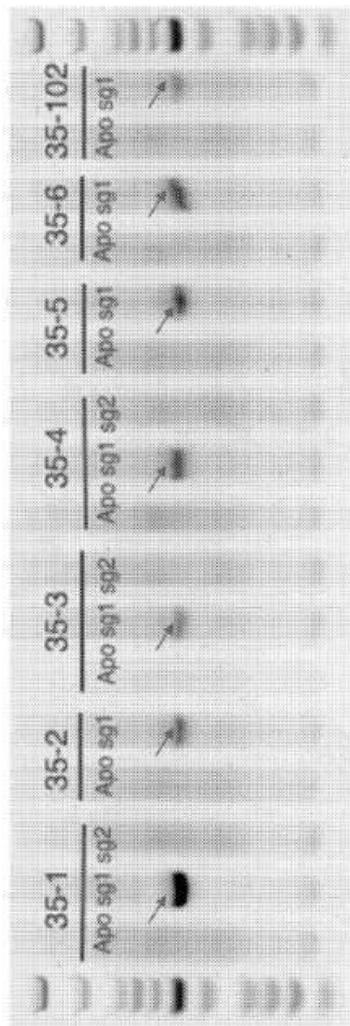
도면 19a



도면 19b



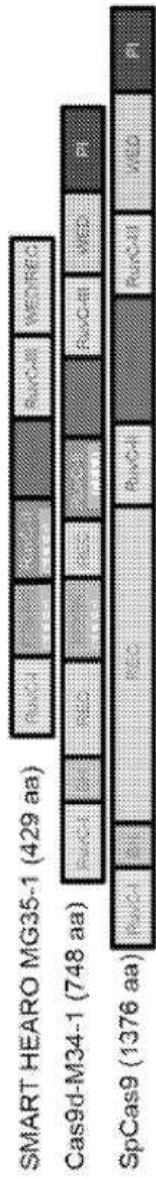
도면20a



도면20b

후보	단백질 크기(aa)	gRNA 크기(nt)	3' TAM(생거)	절단 부위
35-1	429	270	AaGG	5 및 7
35-2	495	234	nARAA	7
35-3	424	278	ATGAnA	7
35-4	429	277	ATGA	7 및 13
35-5	494	287	ATGG	5 및 7
35-6	428	277	A-종부, NGS 필요	13
35-102	424	292	nTGA	7

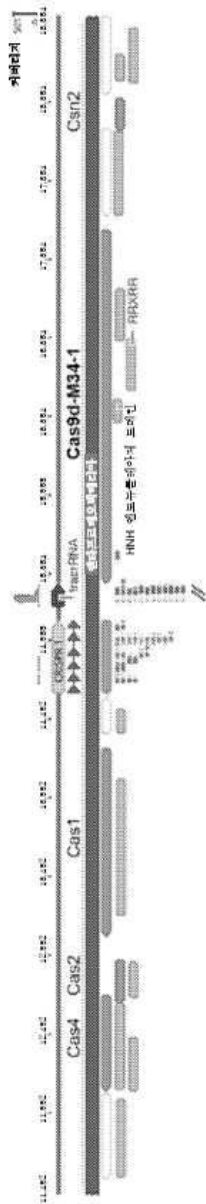
도면21a



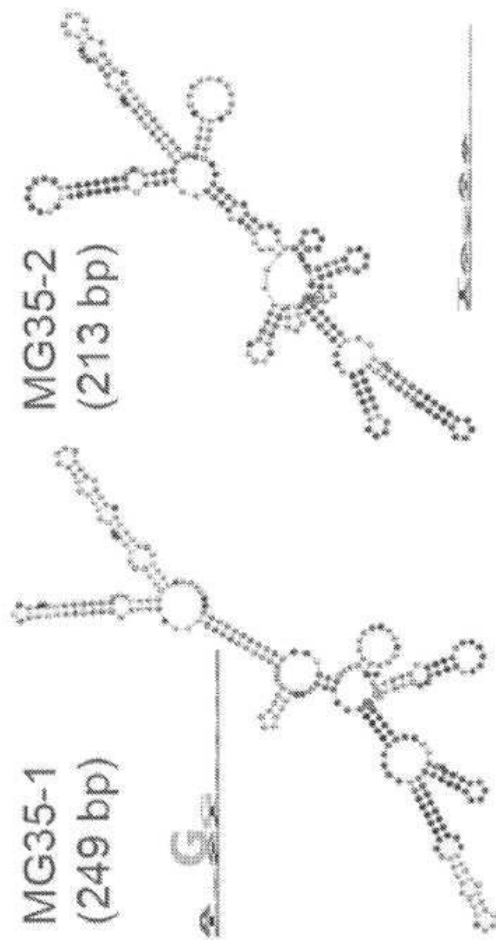
도면21b



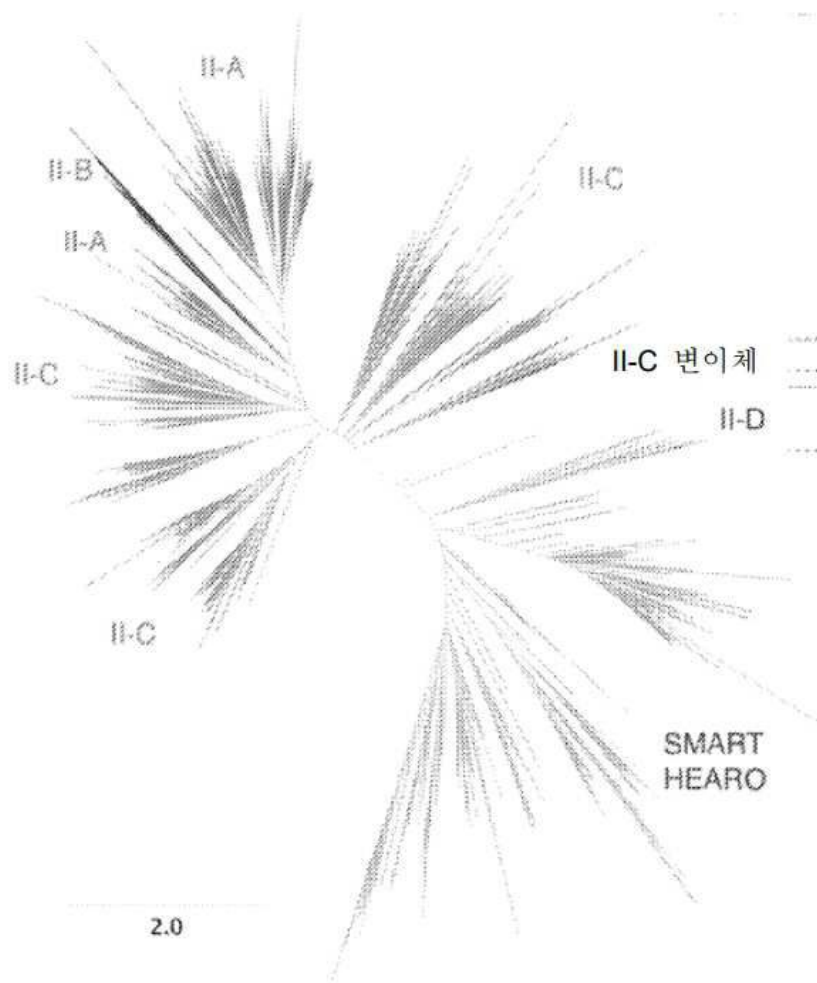
도면21c



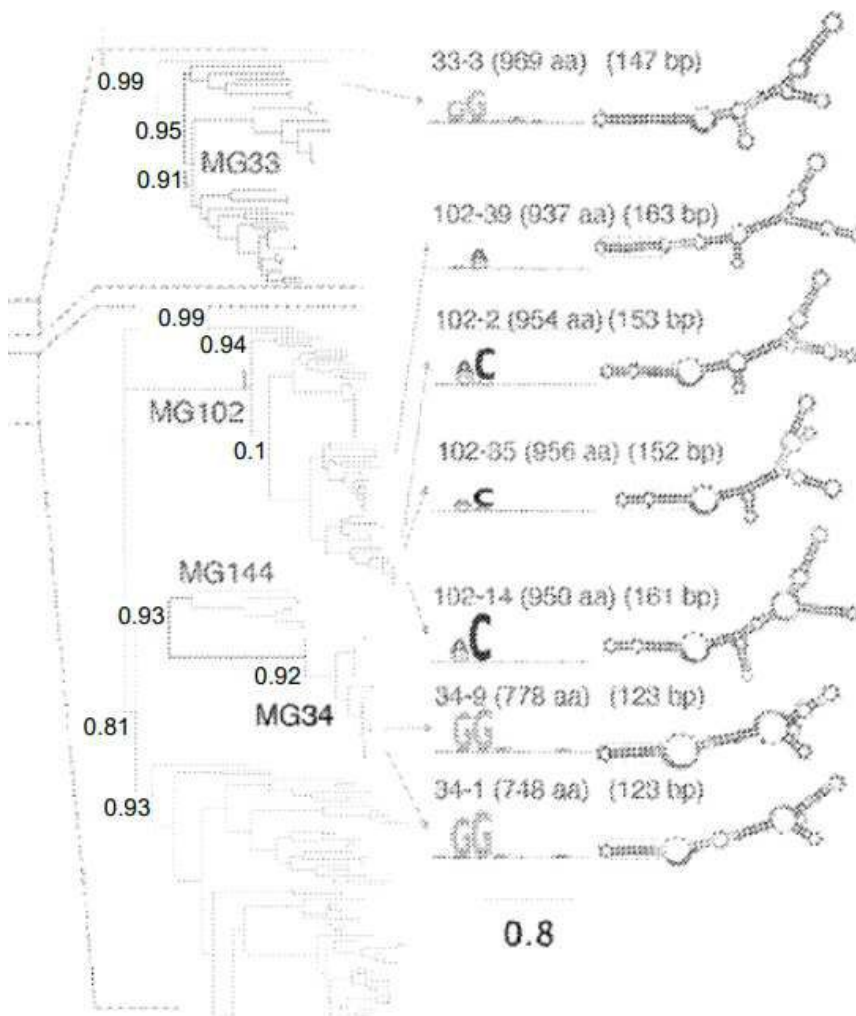
도면21d



도면21e

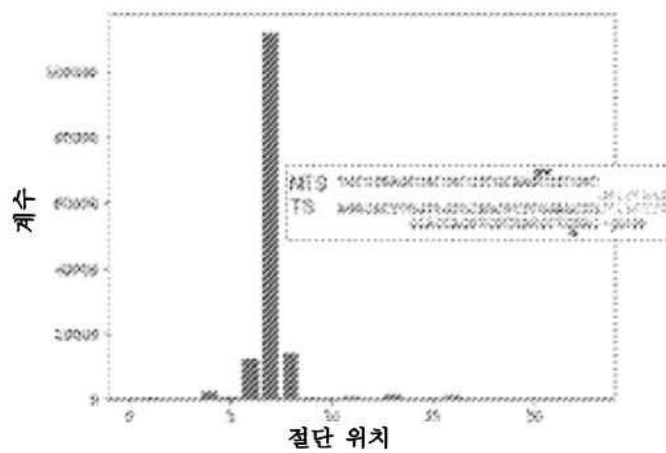


도면21f

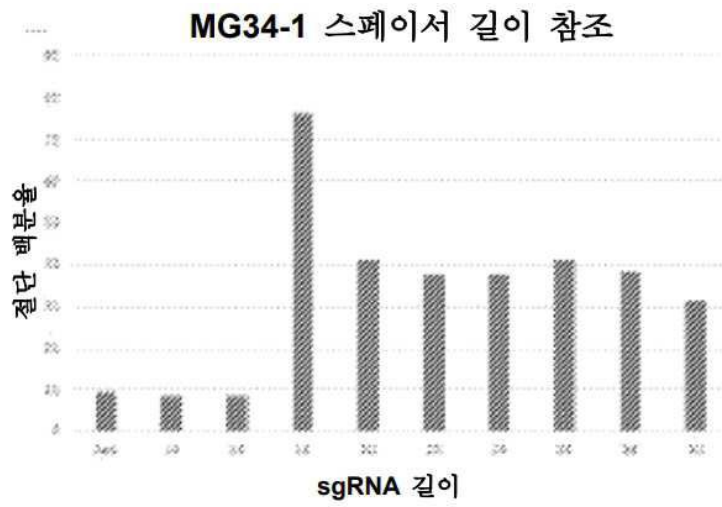


도면22a

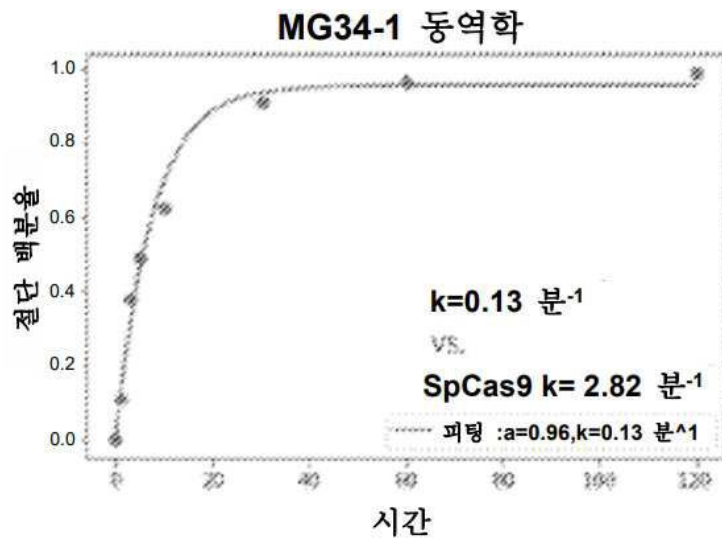
MG34-1 의 절단 위치



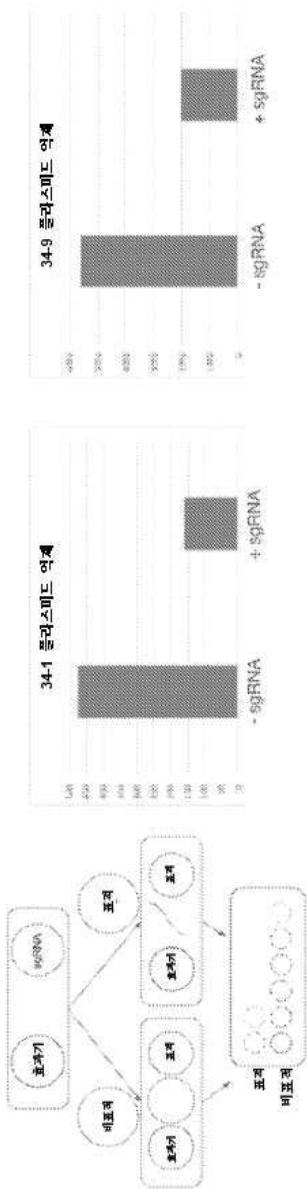
도면22b



도면22c



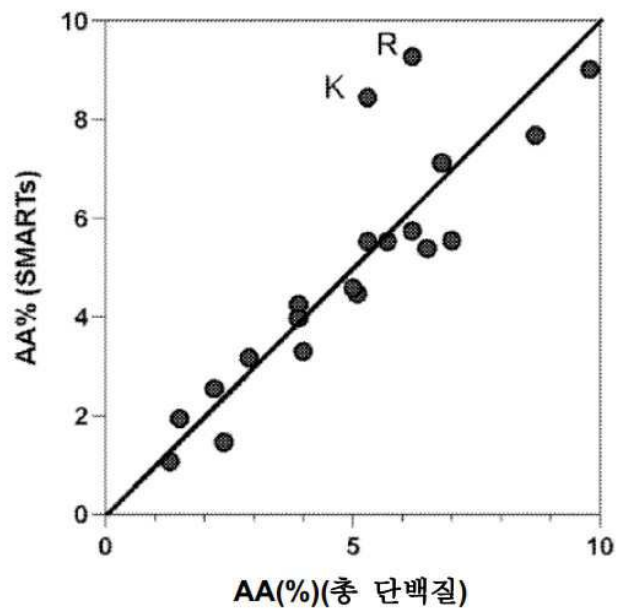
도면22d



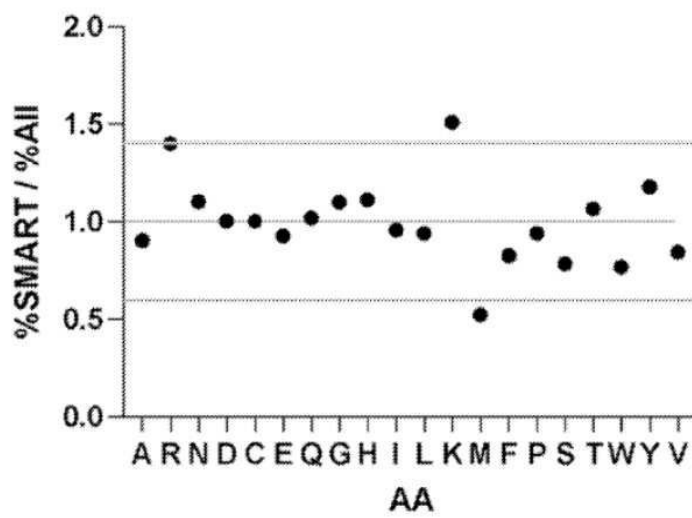
도면23

	A	R	N	D	C	E	Q	G	H	I	L	K	M	F	P	S	T	W	Y	V
MG35-1	7.2	9.8	4.2	3.7	2.1	3.7	4.7	8.9	1.9	6.3	7.5	10.0	1.2	3.0	4.4	5.1	5.6	1.2	3.5	5.8
MG35-2	8.4	8.0	5.0	4.6	2.1	4.8	3.6	6.1	4.4	4.8	8.4	7.8	2.7	3.1	4.8	4.6	6.5	0.8	3.2	6.5
MG35-3	7.6	8.5	4.5	3.5	2.1	4.3	4.5	7.1	2.8	5.7	9.7	6.7	1.2	4.3	4.3	5.2	6.9	0.9	3.1	5.2
MG35-6	6.8	9.3	4.7	3.7	2.3	4.0	4.7	7.7	2.3	5.4	9.6	6.6	0.9	3.0	4.7	4.9	7.0	1.4	2.8	9.1
MG35-102	8.0	8.0	4.5	3.3	2.4	4.0	5.4	8.0	2.6	5.7	9.7	8.5	0.7	3.8	5.0	4.7	6.4	1.2	2.8	5.4
IscB (Aliae-Tran, 2021)	10.0	11.4	3.2	2.1	1.6	4.6	6.2	7.7	3.4	5.0	10.9	5.8	1.4	2.5	4.3	4.6	5.5	1.4	2.3	6.2
MG34-1	5.5	8.6	5.4	5.8	1.6	9.4	1.7	6.2	2.4	6.2	6.6	12.9	2.0	3.5	3.2	4.4	5.2	0.7	3.1	5.9
MG102-2	6.8	8.7	3.0	5.9	1.7	6.0	4.3	6.7	2.4	5.8	9.3	7.4	1.7	3.4	5.1	7.3	5.0	1.1	3.7	4.9
MG102-14	8.1	11.1	3.7	5.0	1.9	7.7	2.8	5.0	2.3	5.6	9.0	7.0	1.2	3.2	5.0	6.7	4.6	1.1	3.6	4.6
MG102-35	8.0	9.0	4.4	6.1	1.8	7.3	2.8	6.4	1.7	4.9	9.4	6.5	1.6	3.2	4.7	6.9	4.5	0.9	3.4	4.4
MG102-45	8.1	9.6	4.2	5.5	1.9	7.4	3.0	7.6	1.9	5.5	9.1	7.6	1.5	3.3	5.0	6.7	3.7	1.1	3.1	4.3
평균	8.7	8.2	3.9	5.1	1.5	6.2	3.9	6.6	2.2	5.7	9.8	6.3	2.4	4.0	5.0	7.0	5.3	1.3	2.9	6.5

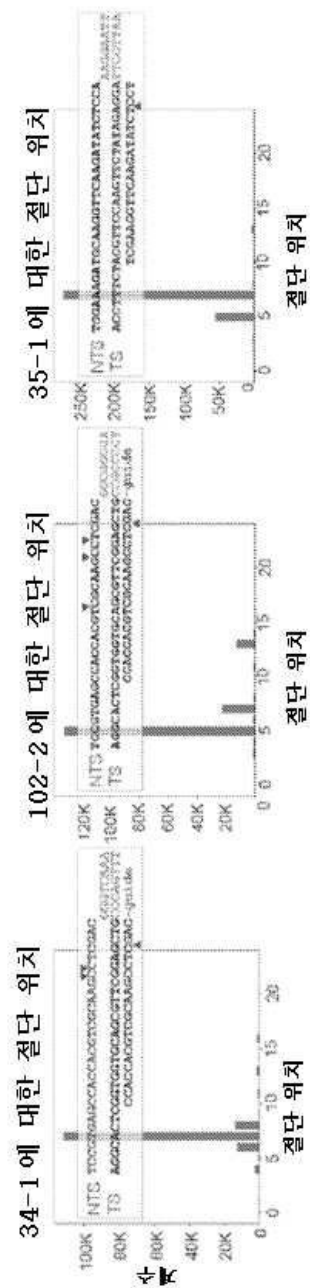
도면24a



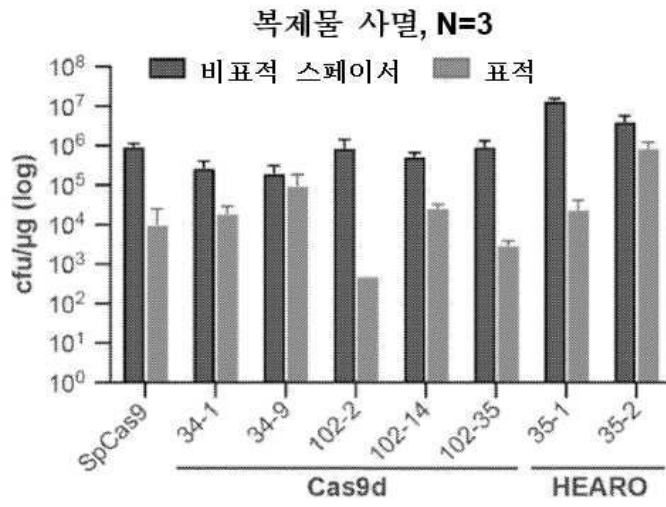
도면24b



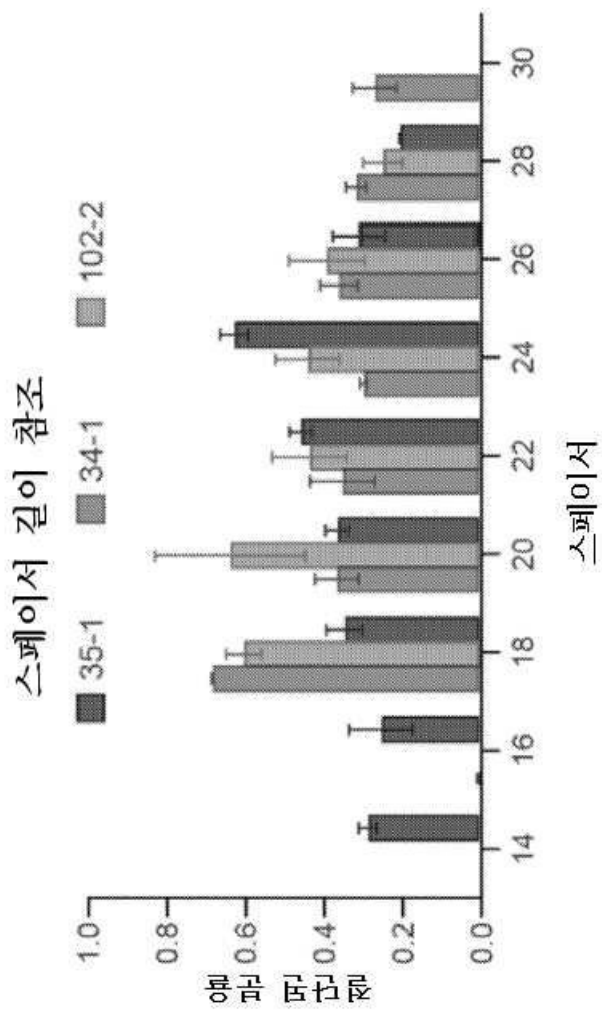
도면25a

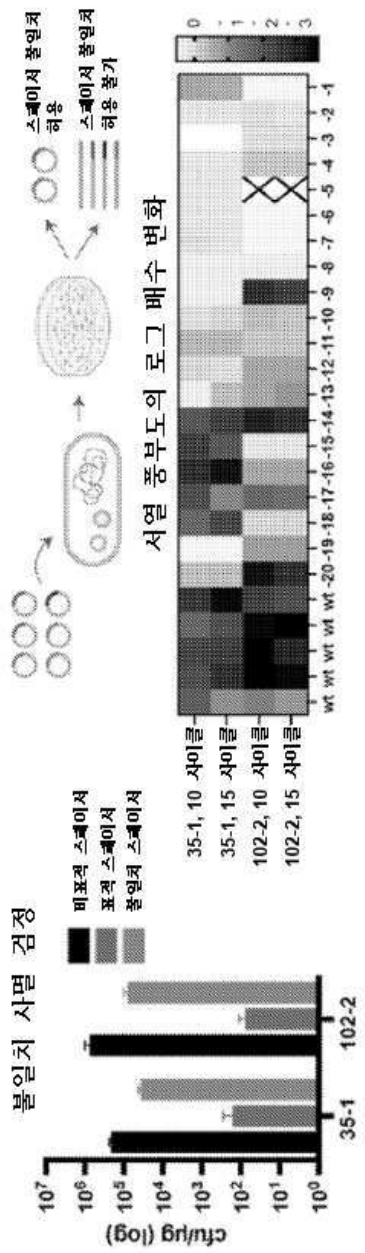


도면25b

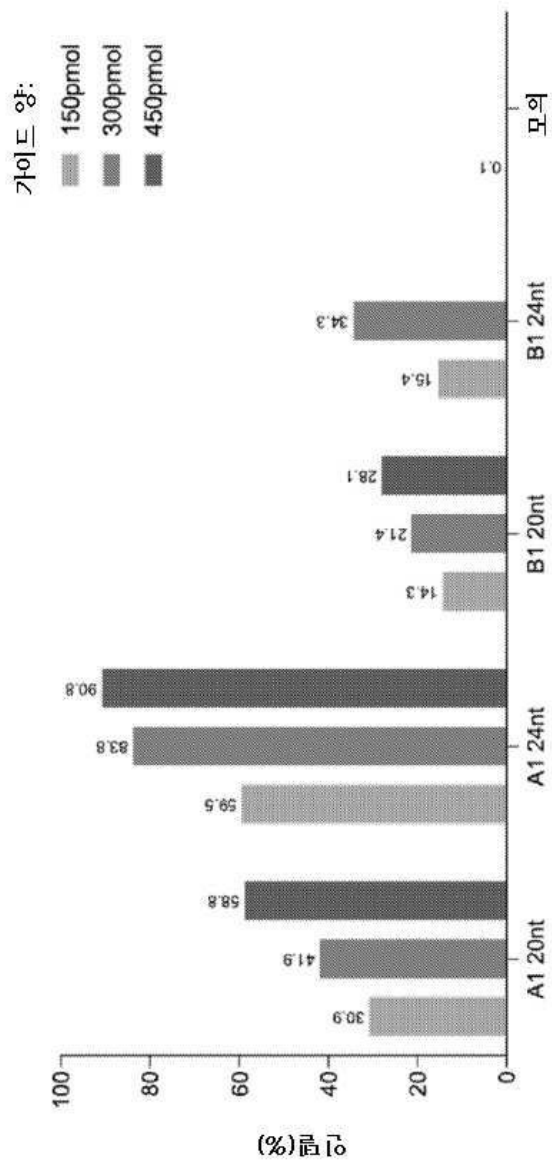


도면25c

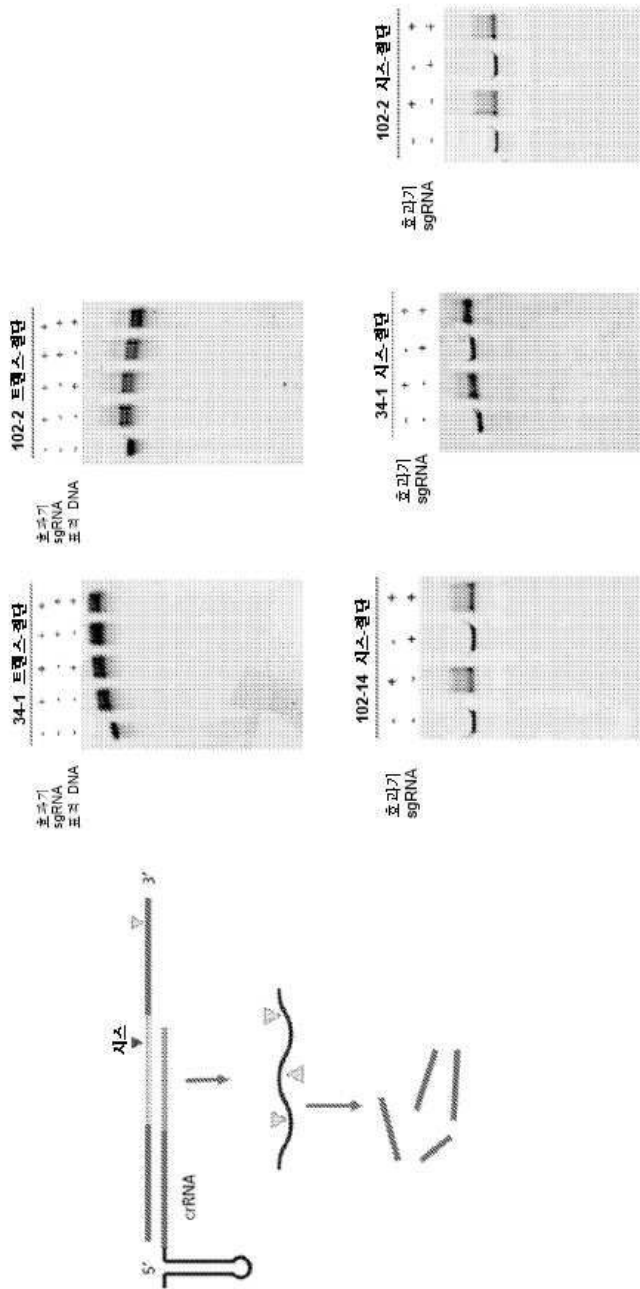




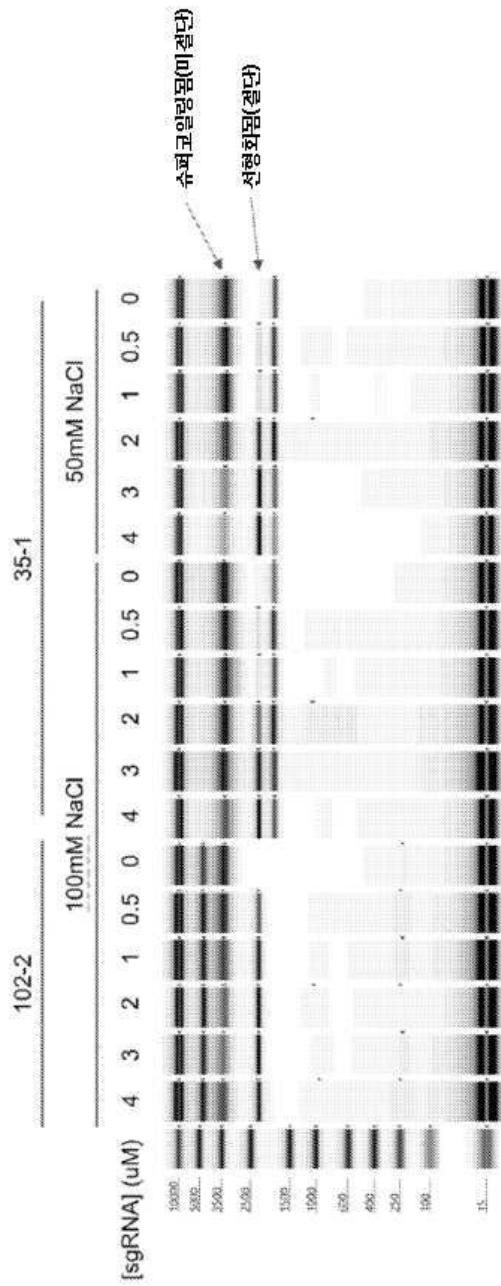
도면26



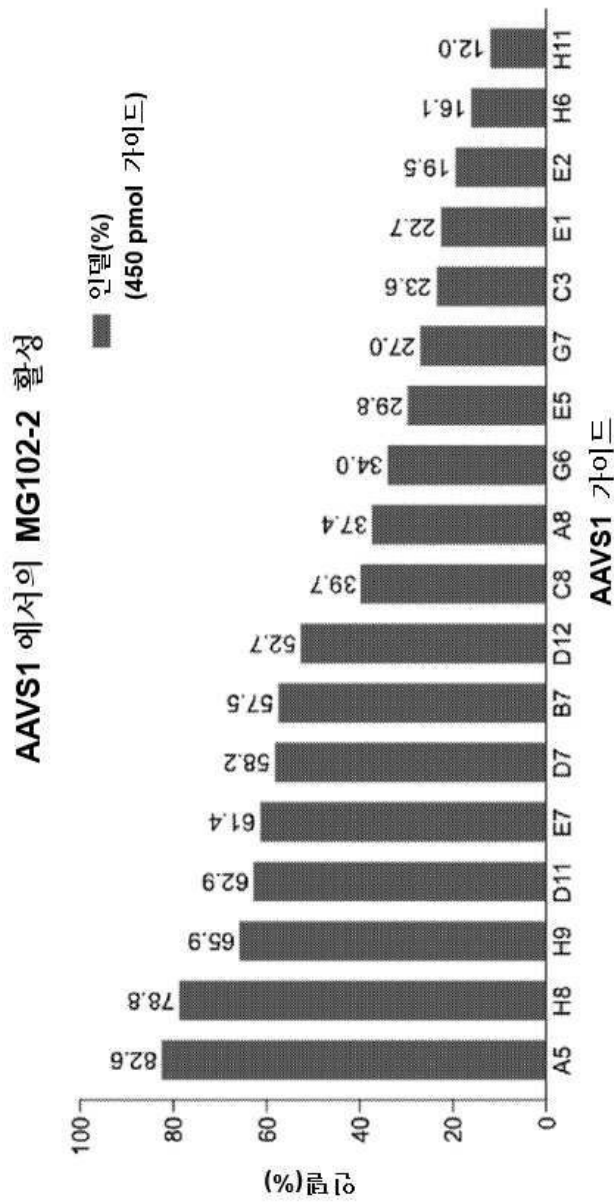
도면28



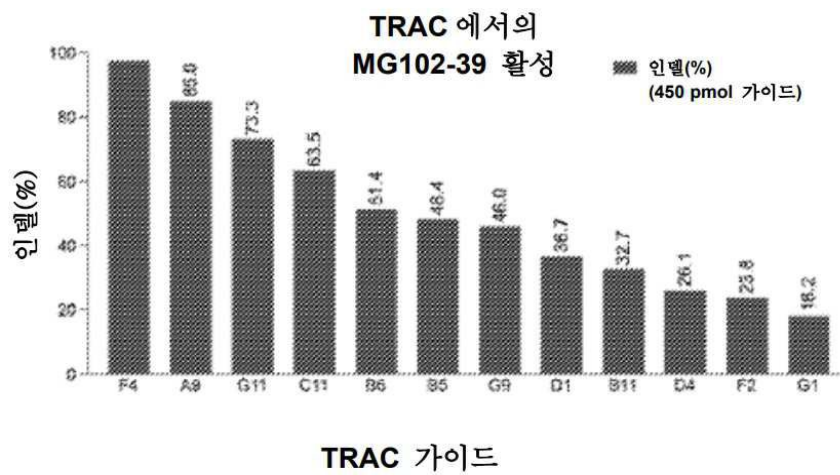
도면29



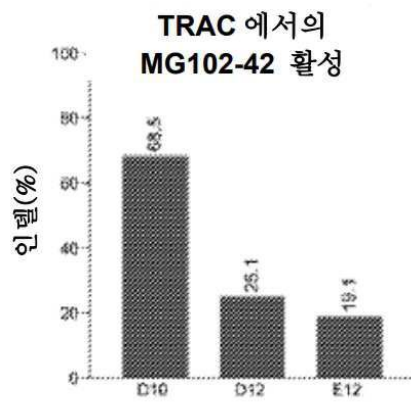
도면30a



도면30b

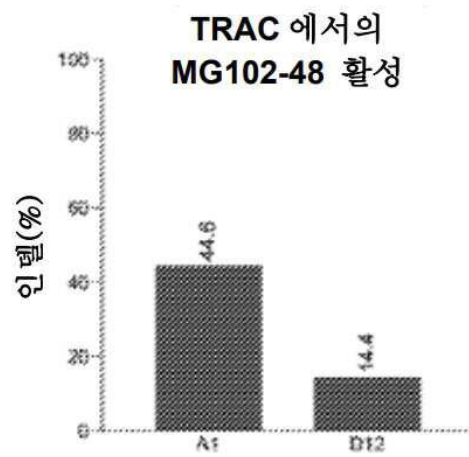


도면30c



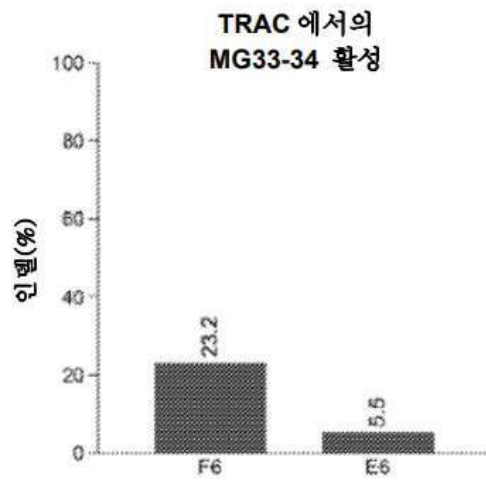
TRAC 가이드

도면30d



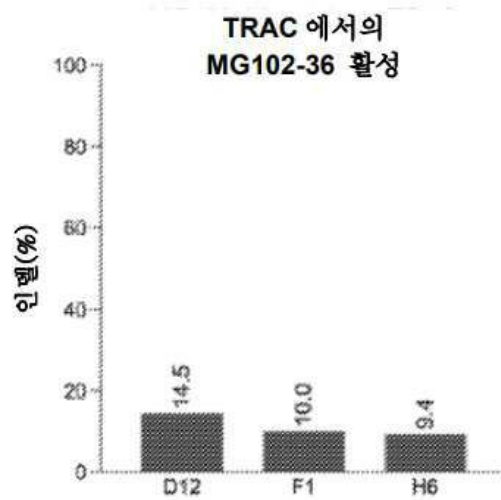
TRAC 가이드

도면30e



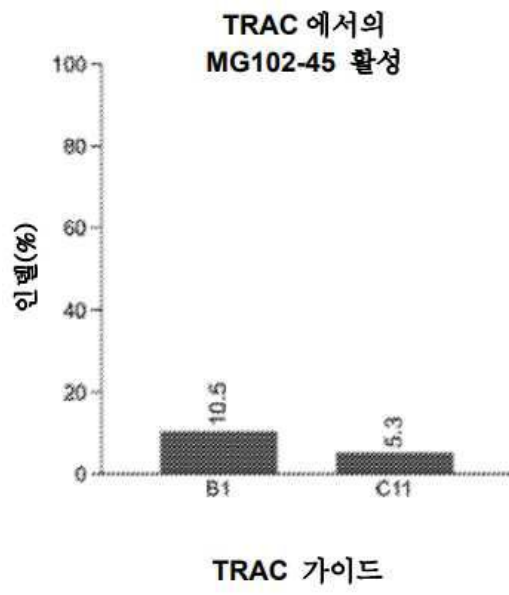
TRAC 가이드

도면30f

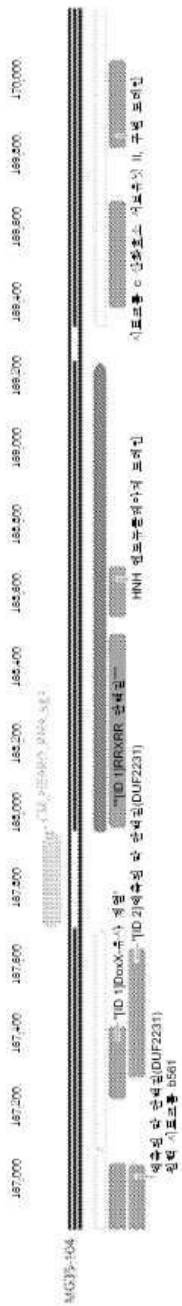


TRAC 가이드

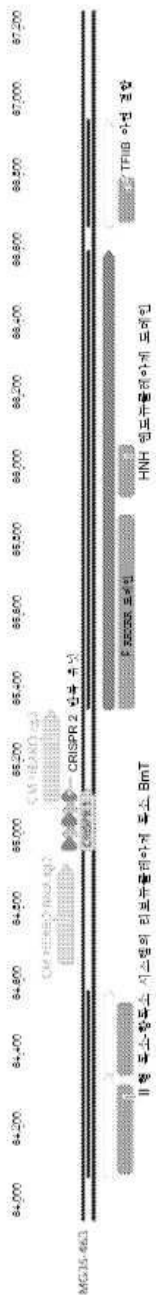
도면 30g



도면32a



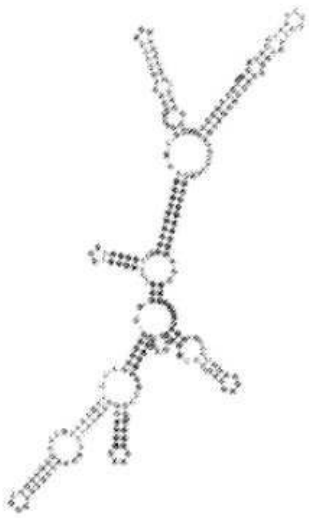
도면 32b



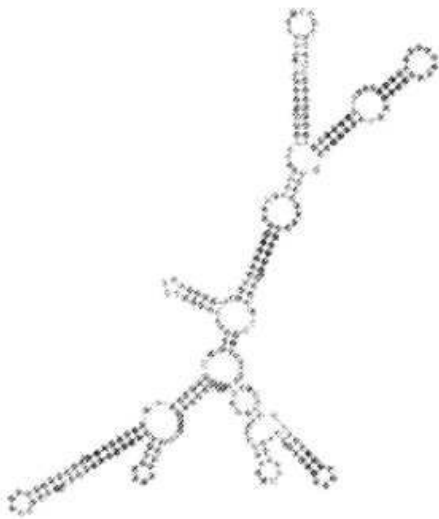
도면32c



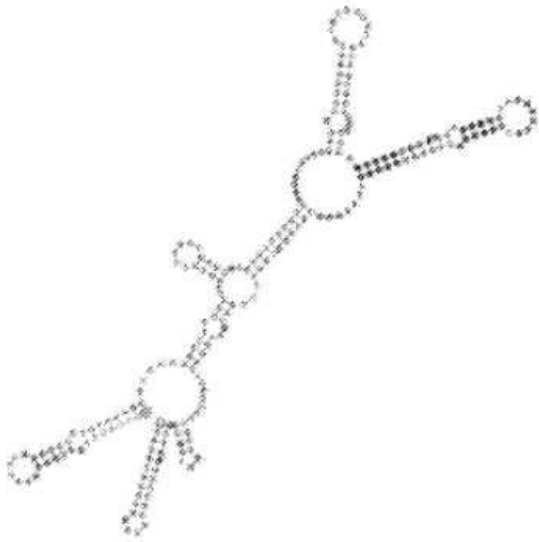
도면32d



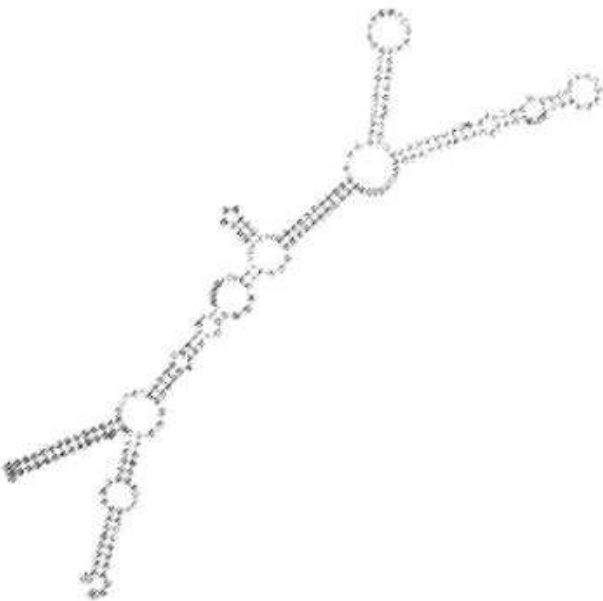
도면32e



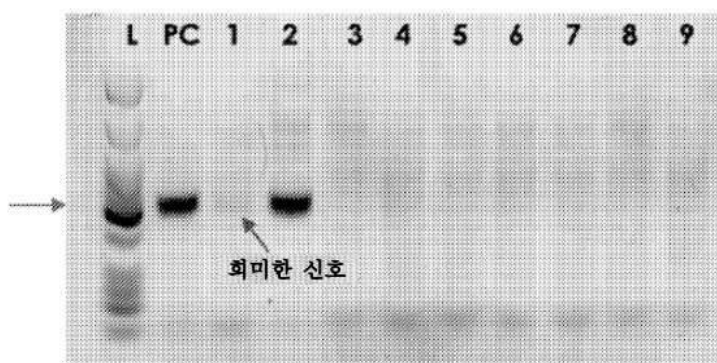
도면32f



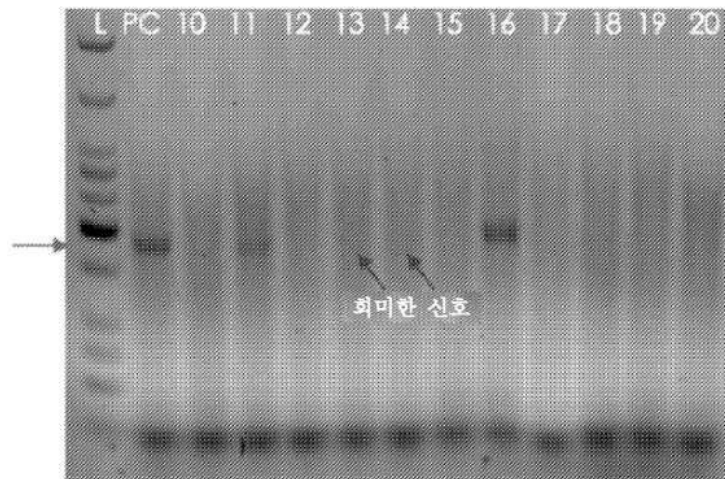
도면32g



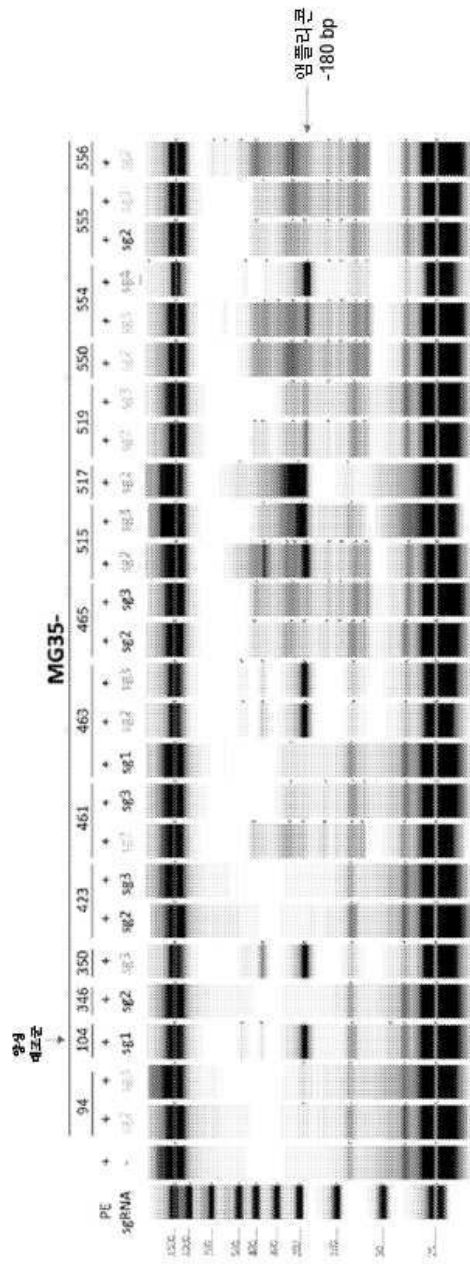
도면33a



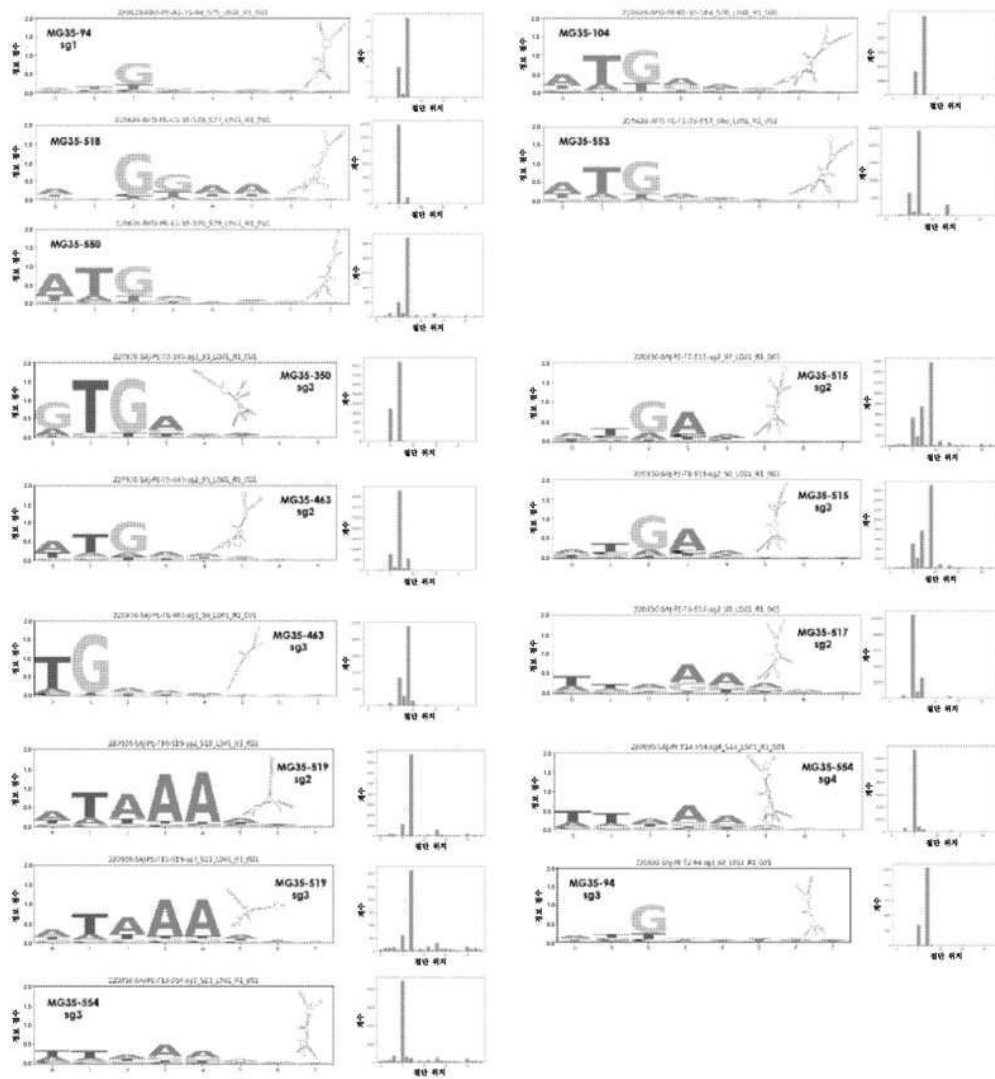
도면33b



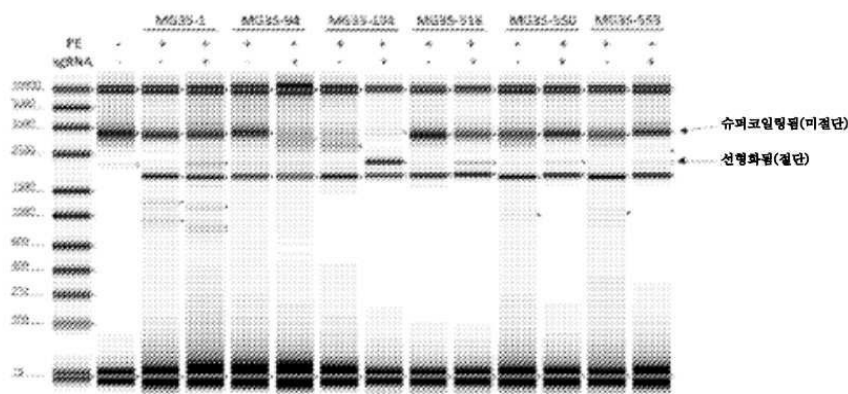
도면33c



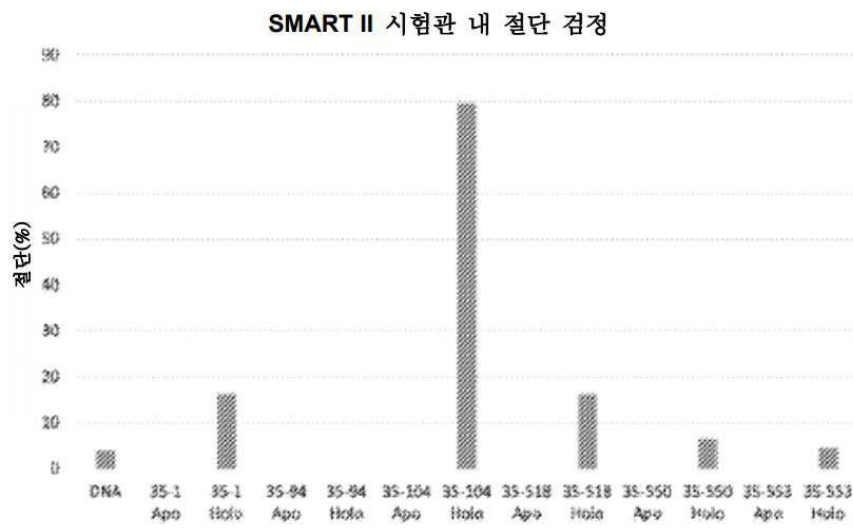
도면34



도면35a



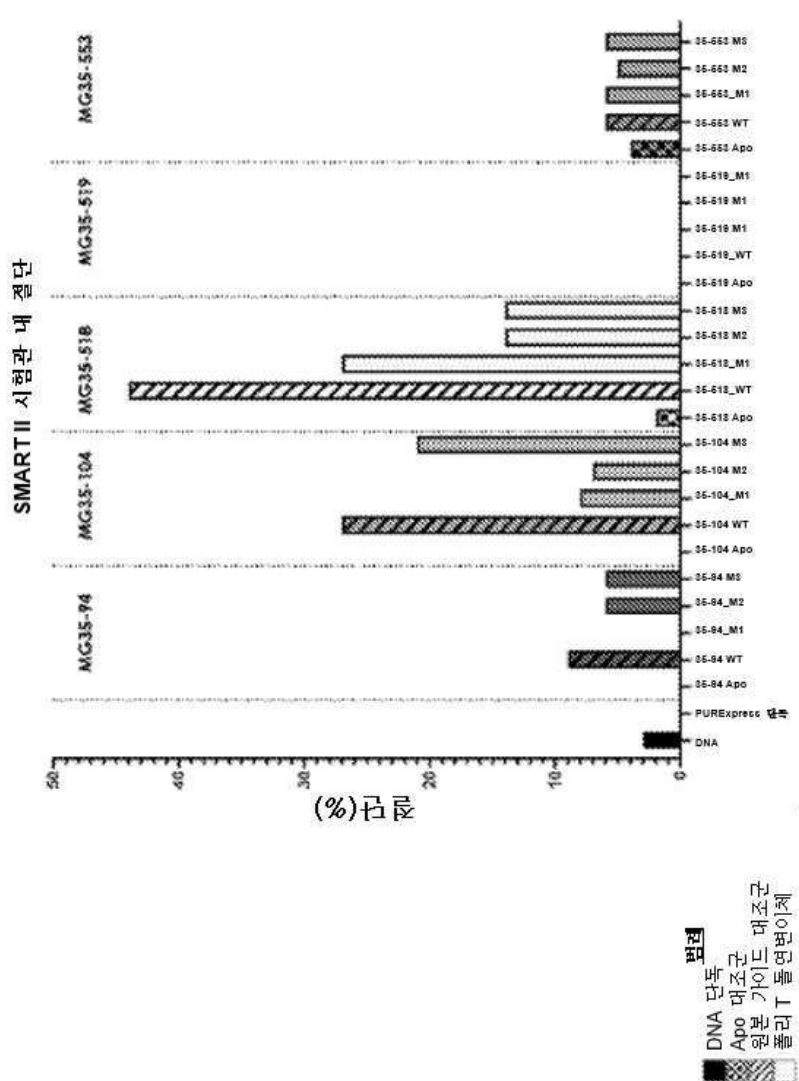
도면35b



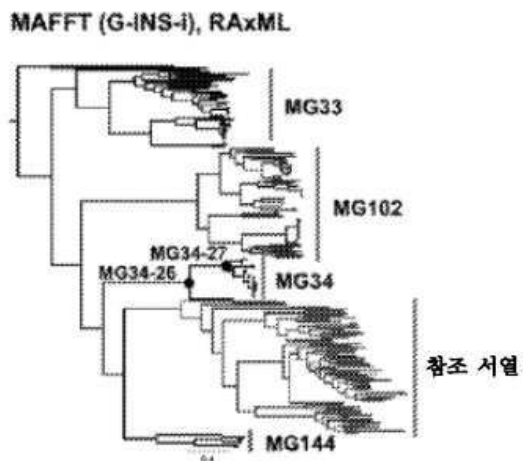
도면36a



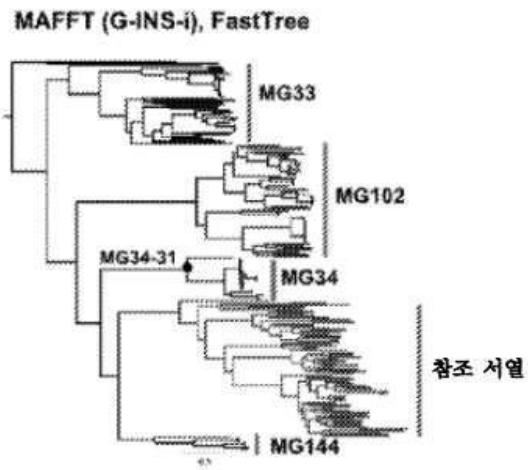
도면36b



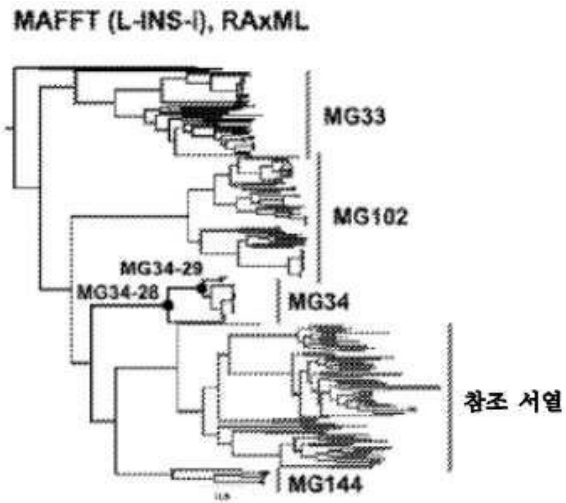
도면37a



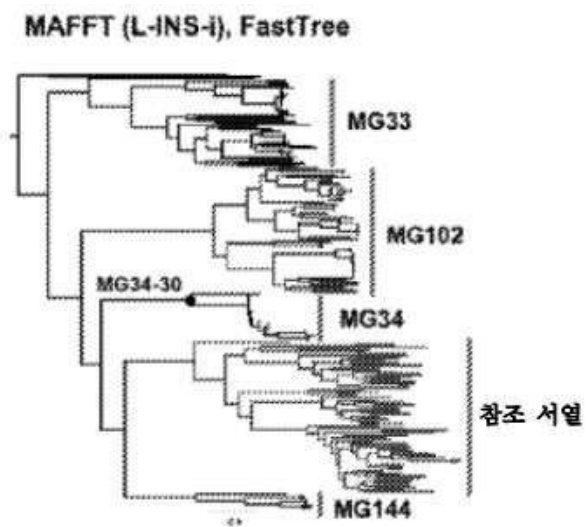
도면37b



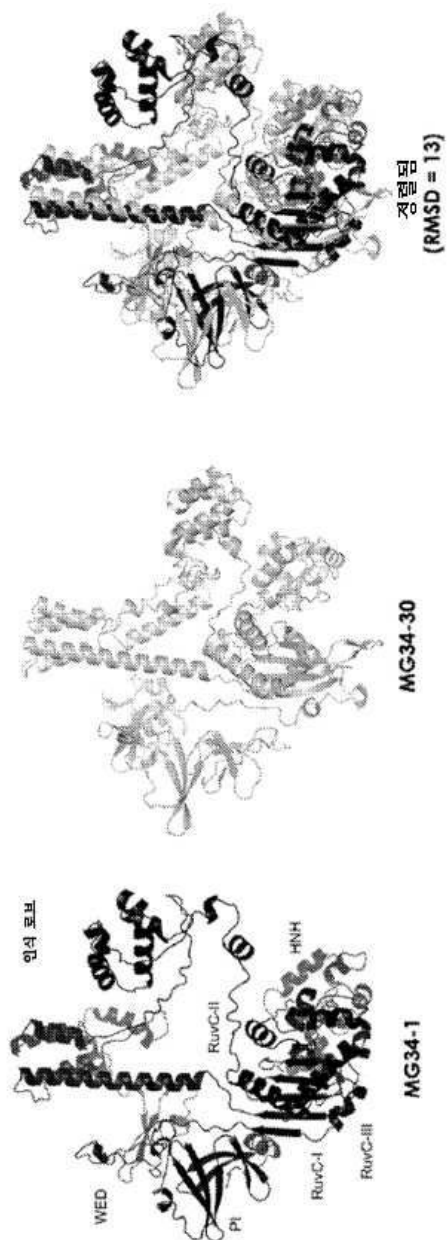
도면37c



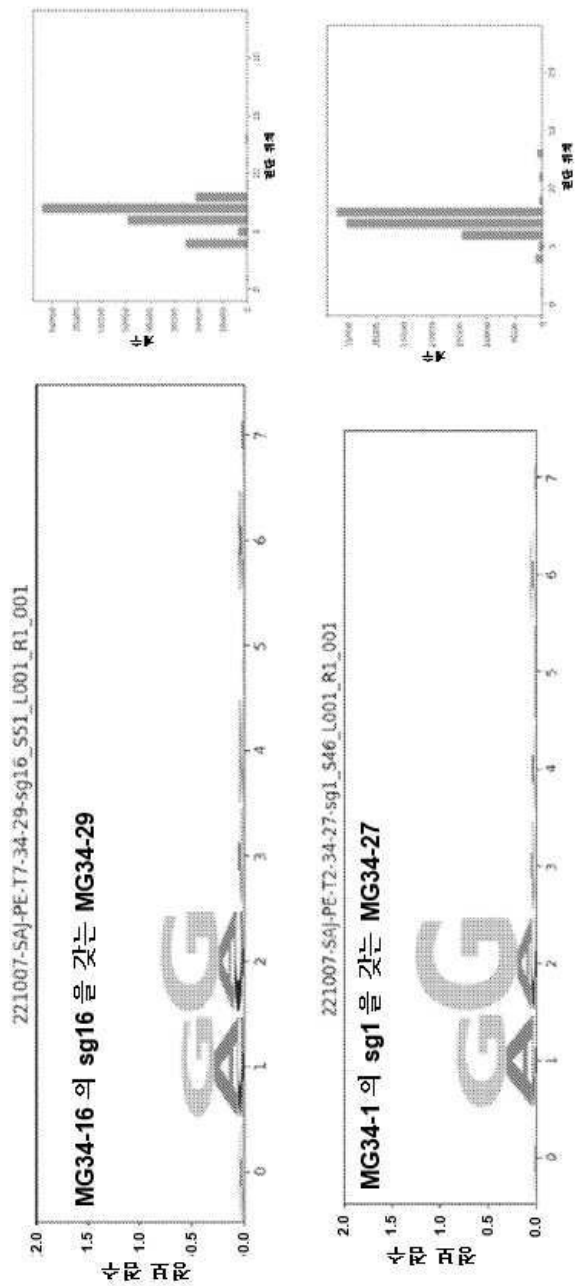
도면37d



도면38



도면40



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.