(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2023-0158476
(43) 공개일자 2023년11월20일

(51) 국제특허분류(Int. Cl.)
C12N 15/86 (2006.01) *A61K 31/7088* (2006.01)
C12N 15/113 (2010.01) *C12N 9/22* (2006.01)
C12N 9/78 (2006.01)
(52) CPC특허분류
C12N 15/86 (2013.01)
A61K 31/7088 (2013.01)
(21) 출원번호 10-2023-7029781
(22) 출원일자(국제) 2022년02월18일
심사청구일자 없음
(85) 번역문제출일자 2023년08월31일
(86) 국제출원번호 PCT/US2022/017075
(87) 국제공개번호 WO 2022/178307
국제공개일자 2022년08월25일
(30) 우선권주장
63/151,542 2021년02월19일 미국(US)
63/241,989 2021년09월08일 미국(US)

(71) 출원인
빔 테라퓨틱스, 인크.
미국 02142 매사추세츠주 캠브리지 메인 스트리트 238 나인쓰 플로어
(72) 발명자
란 페이
미국, 메사추세츠 02139, 캠브리지, 2 플로어, 26 랜즈다운 스트리트, 씨/오 빔 테라퓨틱스, 인크.
린 치에유
미국, 메사추세츠 02139, 캠브리지, 2 플로어, 26 랜즈다운 스트리트, 씨/오 빔 테라퓨틱스, 인크.
(74) 대리인
특허법인한얼

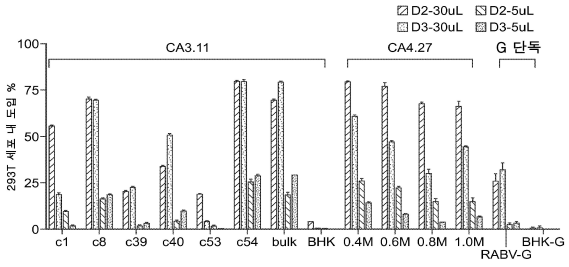
전체 청구항 수 : 총 165 항

(54) 발명의 명칭 **유전자 요법을 위한 재조합 광견병 바이러스**

(57) 요 약

재조합 광견병 바이러스 게놈과 재조합 광견병 바이러스 및 이식유전자를 표적 세포 안으로 전달하는데 있어서 그 사용에 대한 방법이 본 명세서에 제공된다. 또한, 재조합 광견병 바이러스를 생산하기 위해 패키징 시스템 및 패키징 시스템을 사용하는 방법이 제공된다.

대 표 도



(52) CPC특허분류
C12N 15/113 (2013.01)
C12N 9/22 (2013.01)
C12N 9/78 (2013.01)
C12Y 305/04004 (2013.01)
C12Y 305/04005 (2013.01)
C12N 2760/20121 (2013.01)
C12N 2760/20122 (2013.01)
C12N 2760/20143 (2013.01)
C12N 2760/20152 (2013.01)

명 세 서

청구범위

청구항 1

치료적 이식유전자를 인코딩하는 핵산을 포함하는 재조합 광견병 바이러스 게놈으로서,

게놈은 광견병 바이러스 당단백질 또는 이의 기능적 변이체를 인코딩하는 G 유전자를 결하고/결하거나;

게놈은 광견병 바이러스 폴리머라제 또는 이의 기능적 변이체를 인코딩하는 L 유전자를 결하는, 재조합 광견병 바이러스 게놈.

청구항 2

제1항에 있어서, 게놈은 광견병 바이러스 당단백질 또는 이의 기능적 변이체를 인코딩하는 G 유전자를 결하는, 재조합 게놈.

청구항 3

제1항에 있어서, 게놈은 광견병 바이러스 당단백질 또는 이의 기능적 변이체를 인코딩하는 G 유전자를 결하고, 게놈은 광견병 바이러스 폴리머라제 또는 이의 기능적 변이체를 인코딩하는 L 유전자를 결하는, 재조합 게놈.

청구항 4

제1항 내지 제3항 중 어느 한 항에 있어서, 게놈은,

광견병 바이러스 핵단백질 또는 이의 기능적 변이체를 인코딩하는 N 유전자;

광견병 바이러스 인단백질 또는 이의 기능적 변이체를 인코딩하는 P 유전자; 및

광견병 바이러스 매트릭스 단백질 또는 이의 기능적 변이체를 인코딩하는 M 유전자를 포함하는, 재조합 게놈.

청구항 5

제1항 내지 제4항 중 어느 한 항에 있어서, 게놈은,

광견병 바이러스 핵단백질 또는 이의 기능적 변이체를 인코딩하는 N 유전자;

광견병 바이러스 인단백질 또는 이의 기능적 변이체를 인코딩하는 P 유전자; 및/또는

광견병 바이러스 매트릭스 단백질 또는 이의 기능적 변이체를 인코딩하는 M 유전자를 결하는, 재조합 게놈.

청구항 6

제1항 내지 제5항 중 어느 한 항의 재조합 광견병 바이러스 게놈 및 광견병 바이러스 당단백질을 포함하는, 재조합 광견병 바이러스 입자.

청구항 7

재조합 광견병 바이러스 입자로서,

광견병 바이러스 당단백질; 및

치료적 이식유전자를 인코딩하는 핵산을 포함하는 재조합 광견병 바이러스 게놈을 포함하며,

게놈은 광견병 바이러스 당단백질 또는 이의 기능적 변이체를 인코딩하는 G 유전자를 결하고/결하거나;

게놈은 광견병 바이러스 폴리머라제 또는 이의 기능적 변이체를 인코딩하는 L 유전자를 결하는, 재조합 광견병 바이러스 입자.

청구항 8

제7항에 있어서, 게놈은 광견병 바이러스 당단백질 또는 이의 기능적 변이체를 인코딩하는 G 유전자를 결하는, 재조합 바이러스 입자.

청구항 9

제7항에 있어서, 게놈은 광견병 바이러스 당단백질 또는 이의 기능적 변이체를 인코딩하는 G 유전자를 결하고, 게놈은 광견병 바이러스 폴리머라제 또는 이의 기능적 변이체를 인코딩하는 L 유전자를 결하는, 재조합 바이러스 입자.

청구항 10

제7항 내지 제9항 중 어느 한 항에 있어서, 게놈은,

광견병 바이러스 핵단백질 또는 이의 기능적 변이체를 인코딩하는 N 유전자;

광견병 바이러스 인단백질 또는 이의 기능적 변이체를 인코딩하는 P 유전자; 및

광견병 바이러스 매트릭스 단백질 또는 이의 기능적 변이체를 인코딩하는 M 유전자를 포함하는, 재조합 바이러스 입자.

청구항 11

제7항 내지 제9항 중 어느 한 항에 있어서, 게놈은,

광견병 바이러스 핵단백질 또는 이의 기능적 변이체를 인코딩하는 N 유전자;

광견병 바이러스 인단백질 또는 이의 기능적 변이체를 인코딩하는 P 유전자; 및/또는

광견병 바이러스 매트릭스 단백질 또는 이의 기능적 변이체를 인코딩하는 M 유전자를 결하는, 재조합 바이러스 입자.

청구항 12

제7항 내지 제11항 중 어느 한 항에 있어서, 재조합 광견병 바이러스는 복제 불능인, 재조합 바이러스 입자.

청구항 13

제7항 내지 제11항 중 어느 한 항에 있어서, 재조합 광견병 바이러스는 복제 결핍성인, 재조합 바이러스 입자.

청구항 14

제1항 내지 제13항 중 어느 한 항에 있어서, 각각의 유전자는 전사 조절 요소에 작동가능하게 연결되는, 재조합 게놈 또는 바이러스 입자.

청구항 15

제14항에 있어서, 전사 조절 요소는 전사 개시 신호를 포함하는, 재조합 게놈 또는 바이러스 입자.

청구항 16

제15항에 있어서, 전사 개시 신호는 광견병 바이러스에 대해 외인성인, 재조합 게놈 또는 바이러스 입자.

청구항 17

제15항에 있어서, 전사 개시 신호는 광견병 바이러스에 대해 내인성인, 재조합 게놈 또는 바이러스 입자.

청구항 18

제1항 내지 제17항 중 어느 한 항에 있어서, 각각의 유전자는 전사 종결 폴리아데닐화 신호에 작동가능하게 연결되는, 재조합 게놈 또는 바이러스 입자.

청구항 19

제1항 내지 제18항 중 어느 한 항에 있어서, 치료적 이식유전자는 핵산 편집 시스템 또는 이의 구성요소를 인코

딩하는 서열을 포함하는, 재조합 게놈 또는 바이러스 입자.

청구항 20

제19항에 있어서, 핵산 편집 시스템 또는 이의 구성요소는 클러스터링된 규칙적으로 간격을 둔 짧은 회문 반복(CRISPR) 시스템, 아연 핑거 뉴클레아제(ZFN), 메가뉴클레아제, 및 전사 활성제-유사 효과기-기반 뉴클레아제(TALEN)를 포함하는, 재조합 게놈 또는 바이러스 입자.

청구항 21

제19항 또는 제20항에 있어서, 핵산 편집 시스템은 CRISPR 시스템을 포함하는, 재조합 게놈 또는 바이러스 입자.

청구항 22

제21항에 있어서, CRISPR-시스템은 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인 및 핵염기 편집 도메인을 포함하는 핵염기 편집기를 포함하는, 재조합 게놈 또는 바이러스 입자.

청구항 23

제22항에 있어서, 핵염기 편집 도메인은 아데노신 데아미나제, 시티딘 데아미나제 또는 이의 기능적 변이체를 포함하는, 재조합 게놈 또는 바이러스 입자.

청구항 24

제22항 또는 제23항에 있어서, 핵염기 편집 도메인은 아데노신 데아미나제를 포함하는, 재조합 게놈 또는 바이러스 입자.

청구항 25

제24항에 있어서, 아데노신 데아미나제는 표 10, 표 11, 표 12, 표 13, 표 14 또는 표 15에 인용된 임의의 아데노신 염기 편집기로부터의 TadA 데아미나제를 포함하는, 재조합 게놈 또는 바이러스 입자.

청구항 26

제24항에 있어서, 핵염기 편집 도메인은 ABE 0.1, ABE 0.2, ABE 1.1, ABE 1.2, ABE2.1, ABE2.2, ABE2.3, ABE2.4, ABE2.5, ABE2.6, ABE2.7, ABE2.8, ABE2.9, ABE2.10, ABE2.11, ABE2.12, ABE3.1, ABE3.2, ABE3.3, ABE3.4, ABE3.5, ABE3.6, ABE3.7, ABE3.8, ABE4.1, ABE4.2, ABE4.3, ABE5.1, ABE5.2, ABE5.3, ABE5.4, ABE5.5, ABE5.6, ABE5.7, ABE5.8, ABE5.9, ABE5.10, ABE5.11, ABE5.12, ABE5.13, ABE5.14, ABE6.1, ABE6.2, ABE6.3, ABE6.4, ABE6.5, ABE6.6, ABE7.1, ABE7.2, ABE7.3, ABE7.4, ABE7.5, ABE7.6, ABE7.7, ABE7.8, ABE 7.9, ABE7.10, ABE8.1-m, ABE8.2-m, ABE8.3-m, ABE8.4-m, ABE8.5-m, ABE8.6-m, ABE8.7-m, ABE8.8-m, ABE8.9-m, ABE8.10-m, ABE8.11-m, ABE8.12-m, ABE8.13-m, ABE8.14-m, ABE8.15-m, ABE8.16-m, ABE8.17-m, ABE8.18-m, ABE8.19-m, ABE8.20-m, ABE8.21-m, ABE8.22-m, ABE8.23-m, ABE8.24-m, ABE8.1-d, ABE8.2-d, ABE8.3-d, ABE8.4-d, ABE8.5-d, ABE8.6-d, ABE8.7-d, ABE8.8-d, ABE8.9-d, ABE8.10-d, ABE8.11-d, ABE8.12-d, ABE8.13-d, ABE8.14-d, ABE8.15-d, ABE8.16-d, ABE8.17-d, ABE8.18-d, ABE8.19-d, ABE8.20-d, ABE8.21-d, ABE8.22-d, ABE8.23-d, ABE8.24-d, ABE8a-m, ABE8b-m, ABE8c-m, ABE8d-m, ABE8e-m, ABE8a-d, ABE8b-d, ABE8c-d, ABE8d-d, ABE8e-d, ABE9.1, ABE9.2, ABE9.3, ABE9.4, ABE9.5, ABE9.6, ABE9.7, ABE9.8, ABE9.9, ABE9.10, ABE9.11, ABE9.12, ABE9.13, ABE9.14, ABE9.15, ABE9.16, ABE9.17, ABE9.18, ABE9.19, ABE9.2, ABE9.21, ABE9.22, ABE9.23, ABE9.24, ABE9.25, ABE9.26, ABE9.27, ABE9.28, ABE9.29, ABE9.30, ABE9.31, ABE9.32, ABE9.33, ABE9.34, ABE9.35, ABE9.36, ABE9.37, ABE9.38, ABE9.39, ABE9.40, ABE9.41, ABE9.42, ABE9.43, ABE9.44, ABE9.45, ABE9.46, ABE9.47, ABE9.48, ABE9.49, ABE9.50, ABE9.51, ABE9.52, ABE9.53, ABE9.54, ABE9.55, ABE9.56, ABE9.57, 및 ABE9.58의 아데노신 염기 편집기 중 임의의 하나로부터의 아데노신 데아미나제를 포함하는, 재조합 게놈 또는 바이러스 입자.

청구항 27

제24항에 있어서, 핵염기 편집 도메인은 ABE 0.1, ABE 0.2, ABE 1.1, ABE 1.2, ABE2.1, ABE2.2, ABE2.3, ABE2.4, ABE2.5, ABE2.6, ABE2.7, ABE2.8, ABE2.9, ABE2.10, ABE2.11, ABE2.12, ABE3.1, ABE3.2, ABE3.3,

ABE3.4, ABE3.5, ABE3.6, ABE3.7, ABE3.8, ABE4.1, ABE4.2, ABE4.3, ABE5.1, ABE5.2, ABE5.3, ABE5.4, ABE5.5, ABE5.6, ABE5.7, ABE5.8, ABE5.9, ABE5.10, ABE5.11, ABE5.12, ABE5.13, ABE5.14, ABE6.1, ABE6.2, ABE6.3, ABE6.4, ABE6.5, ABE6.6, ABE7.1, ABE7.2, ABE7.3, ABE7.4, ABE7.5, ABE7.6, ABE7.7, ABE7.8, ABE 7.9, ABE7.10, ABE8.1-m, ABE8.2-m, ABE8.3-m, ABE8.4-m, ABE8.5-m, ABE8.6-m, ABE8.7-m, ABE8.8-m, ABE8.9-m, ABE8.10-m, ABE8.11-m, ABE8.12-m, ABE8.13-m, ABE8.14-m, ABE8.15-m, ABE8.16-m, ABE8.17-m, ABE8.18-m, ABE8.19-m, ABE8.20-m, ABE8.21-m, ABE8.22-m, ABE8.23-m, ABE8.24-m, ABE8.1-d, ABE8.2-d, ABE8.3-d, ABE8.4-d, ABE8.5-d, ABE8.6-d, ABE8.7-d, ABE8.8-d, ABE8.9-d, ABE8.10-d, ABE8.11-d, ABE8.12-d, ABE8.13-d, ABE8.14-d, ABE8.15-d, ABE8.16-d, ABE8.17-d, ABE8.18-d, ABE8.19-d, ABE8.20-d, ABE8.21-d, ABE8.22-d, ABE8.23-d, ABE8.24-d, ABE8a-m, ABE8b-m, ABE8c-m, ABE8d-m, ABE8e-m, ABE8a-d, ABE8b-d, ABE8c-d, ABE8d-d, ABE8e-d, ABE9.1, ABE9.2, ABE9.3, ABE9.4, ABE9.5, ABE9.6, ABE9.7, ABE9.8, ABE9.9, ABE9.10, ABE9.11, ABE9.12, ABE9.13, ABE9.14, ABE9.15, ABE9.16, ABE9.17, ABE9.18, ABE9.19, ABE9.2, ABE9.21, ABE9.22, ABE9.23, ABE9.24, ABE9.25, ABE9.26, ABE9.27, ABE9.28, ABE9.29, ABE9.30, ABE9.31, ABE9.32, ABE9.33, ABE9.34, ABE9.35, ABE9.36, ABE9.37, ABE9.38, ABE9.39, ABE9.40, ABE9.41, ABE9.42, ABE9.43, ABE9.44, ABE9.45, ABE9.46, ABE9.47, ABE9.48, ABE9.49, ABE9.50, ABE9.51, ABE9.52, ABE9.53, ABE9.54, ABE9.55, ABE9.56, ABE9.57, 및 ABE9.58로 구성된 군으로부터 선택된 아데노신 염기 편집기를 포함하는, 재조합 게놈 또는 바이러스 입자.

청구항 28

제24항에 있어서, 핵염기 편집 도메인은 ABE7.10 또는 ABE8.20의 아데노신 염기 편집기를 포함하는, 재조합 게놈 또는 바이러스 입자.

청구항 29

제22항 또는 제23항에 있어서, 핵염기 편집 도메인은 시티딘 데아미나제를 포함하는, 재조합 게놈 또는 바이러스 입자.

청구항 30

제29항에 있어서, 시티딘 데아미나제는 *페트로미존 마리누스* 시토신 데아미나제 1(PmCDA1), 활성화-유도된 시티딘 데아미나제(AID) 및 APOBEC로 구성된 군으로부터 선택되는, 재조합 게놈 또는 바이러스 입자.

청구항 31

제22항 내지 제30항 중 어느 한 항에 있어서, 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인은 Cas9 폴리펩티드, Cas12 폴리펩티드 또는 이의 기능적 변이체를 포함하는, 재조합 게놈 또는 바이러스 입자.

청구항 32

제22항 내지 제31항 중 어느 한 항에 있어서, CRISPR-시스템은 가이드 RNA(gRNA) 또는 gRNA를 인코딩하는 핵산 서열을 추가로 포함하는, 재조합 게놈 또는 바이러스 입자.

청구항 33

제1항 내지 제32항 중 어느 한 항에 있어서, 치료적 이식유전자는 우라실 글리코실라제 억제제(UGI)를 인코딩하는 서열을 포함하는, 재조합 게놈 또는 바이러스 입자.

청구항 34

제1항 내지 제33항 중 어느 한 항에 있어서, 치료적 이식유전자는 핵 국소화 신호(NLS)를 인코딩하는 서열을 포함하는, 재조합 게놈 또는 바이러스 입자.

청구항 35

제22항 내지 제32항 중 어느 한 항에 있어서, 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인 및/또는 핵염기 편집 도메인은 핵 국소화 신호(NLS)를 추가로 포함하는, 재조합 게놈 또는 바이러스 입자.

청구항 36

제22항 내지 제35항 중 어느 한 항에 있어서, 시티딘 데아미나제는 우라실 글리코실라제 억제제(UGI)를 추가로 포함하는, 재조합 게놈 또는 바이러스 입자.

청구항 37

제1항 내지 제36항 중 어느 한 항에 있어서, 치료적 이식유전자는 치료적 폴리펩티드 및/또는 치료적 핵산을 인코딩하는 서열을 포함하는, 재조합 게놈 또는 바이러스 입자.

청구항 38

제37항에 있어서, 치료적 폴리펩티드 및/또는 치료적 핵산은 세포로부터 분비되는, 재조합 게놈 또는 바이러스 입자.

청구항 39

제1항 내지 제38항 중 어느 한 항에 있어서, 치료적 이식유전자를 인코딩하는 핵산은 약 300bp, 약 400bp, 약 500bp, 약 600bp, 약 700bp, 약 800bp, 약 900bp, 약 1,000bp, 약 1,100bp, 약 1,200bp, 약 1,300bp, 약 1,400bp, 약 1,500bp, 약 1,600bp, 약 1,700bp, 약 1,800bp, 약 1,900bp, 약 2,000bp, 약 2,100bp, 약 2,200bp, 약 2,300bp, 약 2,400bp, 약 2,500bp, 약 2,600bp, 약 2,700bp, 약 2,800bp, 약 2,900bp, 또는 약 3,000bp보다 큰, 재조합 게놈 또는 바이러스 입자.

청구항 40

제1항 내지 제39항 중 어느 한 항에 있어서, 치료적 이식유전자를 인코딩하는 핵산은 약 300bp보다 큰, 재조합 게놈 또는 바이러스 입자.

청구항 41

제1항 내지 제40항 중 어느 한 항에 있어서, 치료적 이식유전자를 인코딩하는 핵산은 약 650bp보다 큰, 재조합 게놈 또는 바이러스 입자.

청구항 42

제1항 내지 제41항 중 어느 한 항에 있어서, 치료적 이식유전자를 인코딩하는 핵산은 약 1,000bp보다 큰, 재조합 게놈 또는 바이러스 입자.

청구항 43

제1항 내지 제42항 중 어느 한 항에 있어서, 치료적 이식유전자를 인코딩하는 핵산은 약 3,000bp보다 큰, 재조합 게놈 또는 바이러스 입자.

청구항 44

제1항 내지 제43항 중 어느 한 항에 있어서, 치료적 이식유전자를 인코딩하는 핵산은 약 4,500bp보다 큰, 재조합 게놈 또는 바이러스 입자.

청구항 45

제1항 내지 제44항 중 어느 한 항에 있어서, 치료적 이식유전자를 인코딩하는 핵산은 약 8,500bp보다 큰, 재조합 게놈 또는 바이러스 입자.

청구항 46

제1항 내지 제45항 중 어느 한 항에 있어서, 치료적 이식유전자를 인코딩하는 핵산은 약 10,000bp보다 큰, 재조합 게놈 또는 바이러스 입자.

청구항 47

제1항 내지 제46항 중 어느 한 항에 있어서, 치료적 이식유전자는 전사 조절 요소에 작동가능하게 연결된, 재조합 게놈 또는 바이러스 입자.

청구항 48

제47항에 있어서, 전사 조절 요소는 전사 개시 신호를 포함하는, 재조합 게놈 또는 바이러스 입자.

청구항 49

제48항에 있어서, 전사 개시 신호는 광견병 바이러스에 대해 외인성인, 재조합 게놈 또는 바이러스 입자.

청구항 50

제48항에 있어서, 전사 개시 신호는 광견병 바이러스에 대해 내인성인, 재조합 게놈 또는 바이러스 입자.

청구항 51

제1항 내지 제50항 중 어느 한 항에 있어서, 치료적 이식유전자는 전사 종결 폴리아데닐화 신호에 작동가능하게 연결된, 재조합 게놈 또는 바이러스 입자.

청구항 52

제6항 내지 제51항 중 어느 한 항의 재조합 바이러스 입자를 포함하는 약학적 조성물.

청구항 53

제6항 내지 제51항 중 어느 한 항의 재조합 바이러스 입자로 표적 세포를 형질도입시키는 단계를 포함하는, 표적 세포에서 치료적 이식유전자를 발현시키는 방법.

청구항 54

표적 세포에서 핵염기 편집기를 발현시키는 방법으로서, 표적 세포를 재조합 광견병 바이러스 입자로 형질도입시키는 단계를 포함하며, 재조합 바이러스 입자는,

광견병 바이러스 당단백질 또는 이의 변이체; 및

폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인 및 핵염기 편집 도메인을 포함하는 핵염기 편집기를 인코딩하는 핵산을 포함하는 재조합 광견병 바이러스 게놈을 포함하며,

게놈은 광견병 바이러스 당단백질 또는 이의 기능적 변이체를 인코딩하는 G 유전자를 결하고/결하거나;

게놈은 광견병 바이러스 폴리머라제 또는 이의 기능적 변이체를 인코딩하는 L 유전자를 결하는, 방법.

청구항 55

제54항에 있어서, 게놈은,

광견병 바이러스 핵단백질 또는 이의 기능적 변이체를 인코딩하는 N 유전자;

광견병 바이러스 인단백질 또는 이의 기능적 변이체를 인코딩하는 P 유전자; 및

광견병 바이러스 매트릭스 단백질 또는 이의 기능적 변이체를 인코딩하는 M 유전자를 포함하는, 방법.

청구항 56

제54항에 있어서, 게놈은,

광견병 바이러스 핵단백질 또는 이의 기능적 변이체를 인코딩하는 N 유전자;

광견병 바이러스 인단백질 또는 이의 기능적 변이체를 인코딩하는 P 유전자; 및/또는

광견병 바이러스 매트릭스 단백질 또는 이의 기능적 변이체를 인코딩하는 M 유전자를 결하는, 방법.

청구항 57

제54항 내지 제56항 중 어느 한 항에 있어서, 각각의 유전자는 전사 조절 요소에 작동가능하게 연결되는, 방법.

청구항 58

제54항 내지 제57항 중 어느 한 항에 있어서, 전사 조절 요소는 전사 개시 신호를 포함하는, 방법.

청구항 59

제58항에 있어서, 전사 개시 신호는 광견병 바이러스에 대해 외인성인, 방법.

청구항 60

제58항에 있어서, 전사 개시 신호는 광견병 바이러스에 대해 내인성인, 방법.

청구항 61

제54항 내지 제60항 중 어느 한 항에 있어서, 각각의 유전자는 전사 종결 폴리아데닐화 신호에 작동가능하게 연결되는, 방법.

청구항 62

제54항 내지 제61항 중 어느 한 항에 있어서, 핵염기 편집 도메인은 아데노신 데아미나제, 시티딘 데아미나제, 또는 이의 기능적 변이체를 포함하는, 방법.

청구항 63

제54항 내지 제62항 중 어느 한 항에 있어서, 핵염기 편집 도메인은 아데노신 데아미나제를 포함하는, 방법.

청구항 64

제63항에 있어서, 아데노신 데아미나제는 표 10, 표 11, 표 12, 표 13, 표 14 또는 표 15에 인용된 임의의 아데노신 염기 편집기로부터의 TadA 데아미나제를 포함하는, 방법.

청구항 65

제63항에 있어서, 핵염기 편집 도메인은 ABE 0.1, ABE 0.2, ABE 1.1, ABE 1.2, ABE2.1, ABE2.2, ABE2.3, ABE2.4, ABE2.5, ABE2.6, ABE2.7, ABE2.8, ABE2.9, ABE2.10, ABE2.11, ABE2.12, ABE3.1, ABE3.2, ABE3.3, ABE3.4, ABE3.5, ABE3.6, ABE3.7, ABE3.8, ABE4.1, ABE4.2, ABE4.3, ABE5.1, ABE5.2, ABE5.3, ABE5.4, ABE5.5, ABE5.6, ABE5.7, ABE5.8, ABE5.9, ABE5.10, ABE5.11, ABE5.12, ABE5.13, ABE5.14, ABE6.1, ABE6.2, ABE6.3, ABE6.4, ABE6.5, ABE6.6, ABE7.1, ABE7.2, ABE7.3, ABE7.4, ABE7.5, ABE7.6, ABE7.7, ABE7.8, ABE 7.9, ABE7.10, ABE8.1-m, ABE8.2-m, ABE8.3-m, ABE8.4-m, ABE8.5-m, ABE8.6-m, ABE8.7-m, ABE8.8-m, ABE8.9-m, ABE8.10-m, ABE8.11-m, ABE8.12-m, ABE8.13-m, ABE8.14-m, ABE8.15-m, ABE8.16-m, ABE8.17-m, ABE8.18-m, ABE8.19-m, ABE8.20-m, ABE8.21-m, ABE8.22-m, ABE8.23-m, ABE8.24-m, ABE8.1-d, ABE8.2-d, ABE8.3-d, ABE8.4-d, ABE8.5-d, ABE8.6-d, ABE8.7-d, ABE8.8-d, ABE8.9-d, ABE8.10-d, ABE8.11-d, ABE8.12-d, ABE8.13-d, ABE8.14-d, ABE8.15-d, ABE8.16-d, ABE8.17-d, ABE8.18-d, ABE8.19-d, ABE8.20-d, ABE8.21-d, ABE8.22-d, ABE8.23-d, ABE8.24-d, ABE8a-m, ABE8b-m, ABE8c-m, ABE8d-m, ABE8e-m, ABE8a-d, ABE8b-d, ABE8c-d, ABE8d-d, ABE8e-d, ABE9.1, ABE9.2, ABE9.3, ABE9.4, ABE9.5, ABE9.6, ABE9.7, ABE9.8, ABE9.9, ABE9.10, ABE9.11, ABE9.12, ABE9.13, ABE9.14, ABE9.15, ABE9.16, ABE9.17, ABE9.18, ABE9.19, ABE9.2, ABE9.21, ABE9.22, ABE9.23, ABE9.24, ABE9.25, ABE9.26, ABE9.27, ABE9.28, ABE9.29, ABE9.30, ABE9.31, ABE9.32, ABE9.33, ABE9.34, ABE9.35, ABE9.36, ABE9.37, ABE9.38, ABE9.39, ABE9.40, ABE9.41, ABE9.42, ABE9.43, ABE9.44, ABE9.45, ABE9.46, ABE9.47, ABE9.48, ABE9.49, ABE9.50, ABE9.51, ABE9.52, ABE9.53, ABE9.54, ABE9.55, ABE9.56, ABE9.57, 및 ABE9.58의 아데노신 염기 편집기 중 임의의 하나로부터의 아데노신 데아미나제를 포함하는, 방법.

청구항 66

제54항 내지 제65항 중 어느 한 항에 있어서, 핵염기 편집 도메인은 ABE 0.1, ABE 0.2, ABE 1.1, ABE 1.2, ABE2.1, ABE2.2, ABE2.3, ABE2.4, ABE2.5, ABE2.6, ABE2.7, ABE2.8, ABE2.9, ABE2.10, ABE2.11, ABE2.12, ABE3.1, ABE3.2, ABE3.3, ABE3.4, ABE3.5, ABE3.6, ABE3.7, ABE3.8, ABE4.1, ABE4.2, ABE4.3, ABE5.1, ABE5.2, ABE5.3, ABE5.4, ABE5.5, ABE5.6, ABE5.7, ABE5.8, ABE5.9, ABE5.10, ABE5.11, ABE5.12, ABE5.13, ABE5.14, ABE6.1, ABE6.2, ABE6.3, ABE6.4, ABE6.5, ABE6.6, ABE7.1, ABE7.2, ABE7.3, ABE7.4, ABE7.5, ABE7.6, ABE7.7, ABE7.8, ABE 7.9, ABE7.10, ABE8.1-m, ABE8.2-m, ABE8.3-m, ABE8.4-m, ABE8.5-m, ABE8.6-m,

ABE8.7-m, ABE8.8-m, ABE8.9-m, ABE8.10-m, ABE8.11-m, ABE8.12-m, ABE8.13-m, ABE8.14-m, ABE8.15-m, ABE8.16-m, ABE8.17-m, ABE8.18-m, ABE8.19-m, ABE8.20-m, ABE8.21-m, ABE8.22-m, ABE8.23-m, ABE8.24-m, ABE8.1-d, ABE8.2-d, ABE8.3-d, ABE8.4-d, ABE8.5-d, ABE8.6-d, ABE8.7-d, ABE8.8-d, ABE8.9-d, ABE8.10-d, ABE8.11-d, ABE8.12-d, ABE8.13-d, ABE8.14-d, ABE8.15-d, ABE8.16-d, ABE8.17-d, ABE8.18-d, ABE8.19-d, ABE8.20-d, ABE8.21-d, ABE8.22-d, ABE8.23-d, ABE8.24-d, ABE8a-m, ABE8b-m, ABE8c-m, ABE8d-m, ABE8e-m, ABE8a-d, ABE8b-d, ABE8c-d, ABE8d-d, ABE8e-d, ABE9.1, ABE9.2, ABE9.3, ABE9.4, ABE9.5, ABE9.6, ABE9.7, ABE9.8, ABE9.9, ABE9.10, ABE9.11, ABE9.12, ABE9.13, ABE9.14, ABE9.15, ABE9.16, ABE9.17, ABE9.18, ABE9.19, ABE9.2, ABE9.21, ABE9.22, ABE9.23, ABE9.24, ABE9.25, ABE9.26, ABE9.27, ABE9.28, ABE9.29, ABE9.30, ABE9.31, ABE9.32, ABE9.33, ABE9.34, ABE9.35, ABE9.36, ABE9.37, ABE9.38, ABE9.39, ABE9.40, ABE9.41, ABE9.42, ABE9.43, ABE9.44, ABE9.45, ABE9.46, ABE9.47, ABE9.48, ABE9.49, ABE9.50, ABE9.51, ABE9.52, ABE9.53, ABE9.54, ABE9.55, ABE9.56, ABE9.57, 및 ABE9.58로 구성된 군으로부터 선택된 아데노신 염기 편집기를 포함하는, 방법.

청구항 67

제66항에 있어서, 핵염기 편집 도메인은 ABE7.10 또는 ABE8.20의 아데노신 염기 편집기를 포함하는, 방법.

청구항 68

제62항에 있어서, 핵염기 편집 도메인은 시티딘 데아미나제인, 방법.

청구항 69

제68항에 있어서, 시티딘 데아미나제는 *페트로미존 마리누스* 시토신 데아미나제 1(PmCDA1), 활성화-유도된 시티딘 데아미나제(AID) 및 APOBEC로 구성된 군으로부터 선택되는, 방법.

청구항 70

제54항 내지 제69항 중 어느 한 항에 있어서, 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인은 Cas9 폴리펩티드, Cas12 폴리펩티드, 또는 이의 기능적 변이체를 포함하는, 방법.

청구항 71

제54항 내지 제70항 중 어느 한 항에 있어서, 재조합 게놈은 가이드 RNA(gRNA)를 인코딩하는 핵산 서열을 추가로 포함하는, 방법.

청구항 72

제71항에 있어서, 가이드 RNA(gRNA)는 표적 세포에 외인적으로 제공되는, 방법.

청구항 73

제71항 또는 제72항에 있어서, gRNA는 표적 세포의 게놈 좌위를 표적화할 수 있는, 방법.

청구항 74

제54항 내지 제73항 중 어느 한 항에 있어서, 표적 세포는 *생체외에서* 형질도입되는, 방법.

청구항 75

제74항에 있어서, 표적 세포는 인간 세포인, 방법.

청구항 76

제74항 또는 제75항에 있어서, 표적 세포는 인간으로부터 획득되는, 방법.

청구항 77

제74항 내지 제76항 중 어느 한 항에 있어서, 표적 세포는 인간에 대해 자가조직인, 방법.

청구항 78

제74항 내지 제76항 중 어느 한 항에 있어서, 표적 세포는 인간에 대해 동종이계인, 방법.

청구항 79

제54항 내지 제73항 중 어느 한 항에 있어서, 표적 세포는 *생체내에서* 형질도입되는, 방법.

청구항 80

제79항에 있어서, 표적 세포는 인간 세포인, 방법.

청구항 81

제79항 또는 제80항에 있어서, 표적 세포는 신경 세포, 상피 세포 또는 간세포인, 방법.

청구항 82

제79항 내지 제81항 중 어느 한 항에 있어서, 표적 세포는 인간에 있는, 방법.

청구항 83

제6항 내지 제51항 중 어느 한 항의 재조합 바이러스 입자 또는 제52항의 약학적 조성물을 대상체에게 투여하는 단계를 포함하는, 대상체에게 치료적 이식유전자를 전달하는 방법.

청구항 84

대상체에게 핵염기 편집기를 전달하는 방법으로서, 대상체에게 재조합 광견병 바이러스 입자를 투여하는 단계를 포함하며, 여기서 재조합 바이러스 입자는,

광견병 바이러스 당단백질 또는 이의 변이체; 및

폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인 및 핵염기 편집 도메인을 포함하는 핵염기 편집기를 인코딩하는 핵산을 포함하는 재조합 광견병 바이러스 게놈을 포함하며,

게놈은 광견병 바이러스 당단백질 또는 이의 기능적 변이체를 인코딩하는 G 유전자를 결하고/결하거나;

게놈은 광견병 바이러스 폴리머라제 또는 이의 기능적 변이체를 인코딩하는 L 유전자를 결하는, 방법.

청구항 85

제84항에 있어서, 게놈은,

광견병 바이러스 핵단백질 또는 이의 기능적 변이체를 인코딩하는 N 유전자;

광견병 바이러스 인단백질 또는 이의 기능적 변이체를 인코딩하는 P 유전자; 및

광견병 바이러스 매트릭스 단백질 또는 이의 기능적 변이체를 인코딩하는 M 유전자를 포함하는, 방법.

청구항 86

제84항에 있어서, 게놈은,

광견병 바이러스 핵단백질 또는 이의 기능적 변이체를 인코딩하는 N 유전자;

광견병 바이러스 인단백질 또는 이의 기능적 변이체를 인코딩하는 P 유전자; 및/또는

광견병 바이러스 매트릭스 단백질 또는 이의 기능적 변이체를 인코딩하는 M 유전자를 결하는, 방법.

청구항 87

제84항 내지 제86항 중 어느 한 항에 있어서, 핵염기 편집 도메인은 아데노신 데아미나제, 시티딘 데아미나제, 또는 이의 기능적 변이체를 포함하는, 방법.

청구항 88

제84항 내지 제87항 중 어느 한 항에 있어서, 핵염기 편집 도메인은 아데노신 데아미나제를 포함하는, 방법.

청구항 89

제88항에 있어서, 아데노신 데아미나제는 표 10, 표 11, 표 12, 표 13, 표 14 또는 표 15에 인용된 임의의 아데노신 염기 편집기로부터의 TadA 데아미나제를 포함하는, 방법.

청구항 90

제84항 내지 제88항 중 어느 한 항에 있어서, 핵염기 편집 도메인은 ABE 0.1, ABE 0.2, ABE 1.1, ABE 1.2, ABE2.1, ABE2.2, ABE2.3, ABE2.4, ABE2.5, ABE2.6, ABE2.7, ABE2.8, ABE2.9, ABE2.10, ABE2.11, ABE2.12, ABE3.1, ABE3.2, ABE3.3, ABE3.4, ABE3.5, ABE3.6, ABE3.7, ABE3.8, ABE4.1, ABE4.2, ABE4.3, ABE5.1, ABE5.2, ABE5.3, ABE5.4, ABE5.5, ABE5.6, ABE5.7, ABE5.8, ABE5.9, ABE5.10, ABE5.11, ABE5.12, ABE5.13, ABE5.14, ABE6.1, ABE6.2, ABE6.3, ABE6.4, ABE6.5, ABE6.6, ABE7.1, ABE7.2, ABE7.3, ABE7.4, ABE7.5, ABE7.6, ABE7.7, ABE7.8, ABE 7.9, ABE7.10, ABE8.1-m, ABE8.2-m, ABE8.3-m, ABE8.4-m, ABE8.5-m, ABE8.6-m, ABE8.7-m, ABE8.8-m, ABE8.9-m, ABE8.10-m, ABE8.11-m, ABE8.12-m, ABE8.13-m, ABE8.14-m, ABE8.15-m, ABE8.16-m, ABE8.17-m, ABE8.18-m, ABE8.19-m, ABE8.20-m, ABE8.21-m, ABE8.22-m, ABE8.23-m, ABE8.24-m, ABE8.1-d, ABE8.2-d, ABE8.3-d, ABE8.4-d, ABE8.5-d, ABE8.6-d, ABE8.7-d, ABE8.8-d, ABE8.9-d, ABE8.10-d, ABE8.11-d, ABE8.12-d, ABE8.13-d, ABE8.14-d, ABE8.15-d, ABE8.16-d, ABE8.17-d, ABE8.18-d, ABE8.19-d, ABE8.20-d, ABE8.21-d, ABE8.22-d, ABE8.23-d, ABE8.24-d, ABE8a-m, ABE8b-m, ABE8c-m, ABE8d-m, ABE8e-m, ABE8a-d, ABE8b-d, ABE8c-d, ABE8d-d, ABE8e-d, ABE9.1, ABE9.2, ABE9.3, ABE9.4, ABE9.5, ABE9.6, ABE9.7, ABE9.8, ABE9.9, ABE9.10, ABE9.11, ABE9.12, ABE9.13, ABE9.14, ABE9.15, ABE9.16, ABE9.17, ABE9.18, ABE9.19, ABE9.2, ABE9.21, ABE9.22, ABE9.23, ABE9.24, ABE9.25, ABE9.26, ABE9.27, ABE9.28, ABE9.29, ABE9.30, ABE9.31, ABE9.32, ABE9.33, ABE9.34, ABE9.35, ABE9.36, ABE9.37, ABE9.38, ABE9.39, ABE9.40, ABE9.41, ABE9.42, ABE9.43, ABE9.44, ABE9.45, ABE9.46, ABE9.47, ABE9.48, ABE9.49, ABE9.50, ABE9.51, ABE9.52, ABE9.53, ABE9.54, ABE9.55, ABE9.56, ABE9.57, 및 ABE9.58의 아데노신 염기 편집기 중 임의의 하나로부터의 아데노신 데아미나제를 포함하는, 방법.

청구항 91

제84항 내지 제88항 중 어느 한 항에 있어서, 핵염기 편집 도메인은 ABE 0.1, ABE 0.2, ABE 1.1, ABE 1.2, ABE2.1, ABE2.2, ABE2.3, ABE2.4, ABE2.5, ABE2.6, ABE2.7, ABE2.8, ABE2.9, ABE2.10, ABE2.11, ABE2.12, ABE3.1, ABE3.2, ABE3.3, ABE3.4, ABE3.5, ABE3.6, ABE3.7, ABE3.8, ABE4.1, ABE4.2, ABE4.3, ABE5.1, ABE5.2, ABE5.3, ABE5.4, ABE5.5, ABE5.6, ABE5.7, ABE5.8, ABE5.9, ABE5.10, ABE5.11, ABE5.12, ABE5.13, ABE5.14, ABE6.1, ABE6.2, ABE6.3, ABE6.4, ABE6.5, ABE6.6, ABE7.1, ABE7.2, ABE7.3, ABE7.4, ABE7.5, ABE7.6, ABE7.7, ABE7.8, ABE 7.9, ABE7.10, ABE8.1-m, ABE8.2-m, ABE8.3-m, ABE8.4-m, ABE8.5-m, ABE8.6-m, ABE8.7-m, ABE8.8-m, ABE8.9-m, ABE8.10-m, ABE8.11-m, ABE8.12-m, ABE8.13-m, ABE8.14-m, ABE8.15-m, ABE8.16-m, ABE8.17-m, ABE8.18-m, ABE8.19-m, ABE8.20-m, ABE8.21-m, ABE8.22-m, ABE8.23-m, ABE8.24-m, ABE8.1-d, ABE8.2-d, ABE8.3-d, ABE8.4-d, ABE8.5-d, ABE8.6-d, ABE8.7-d, ABE8.8-d, ABE8.9-d, ABE8.10-d, ABE8.11-d, ABE8.12-d, ABE8.13-d, ABE8.14-d, ABE8.15-d, ABE8.16-d, ABE8.17-d, ABE8.18-d, ABE8.19-d, ABE8.20-d, ABE8.21-d, ABE8.22-d, ABE8.23-d, ABE8.24-d, ABE8a-m, ABE8b-m, ABE8c-m, ABE8d-m, ABE8e-m, ABE8a-d, ABE8b-d, ABE8c-d, ABE8d-d, ABE8e-d, ABE9.1, ABE9.2, ABE9.3, ABE9.4, ABE9.5, ABE9.6, ABE9.7, ABE9.8, ABE9.9, ABE9.10, ABE9.11, ABE9.12, ABE9.13, ABE9.14, ABE9.15, ABE9.16, ABE9.17, ABE9.18, ABE9.19, ABE9.2, ABE9.21, ABE9.22, ABE9.23, ABE9.24, ABE9.25, ABE9.26, ABE9.27, ABE9.28, ABE9.29, ABE9.30, ABE9.31, ABE9.32, ABE9.33, ABE9.34, ABE9.35, ABE9.36, ABE9.37, ABE9.38, ABE9.39, ABE9.40, ABE9.41, ABE9.42, ABE9.43, ABE9.44, ABE9.45, ABE9.46, ABE9.47, ABE9.48, ABE9.49, ABE9.50, ABE9.51, ABE9.52, ABE9.53, ABE9.54, ABE9.55, ABE9.56, ABE9.57, 및 ABE9.58로 구성된 군으로부터 선택된 아데노신 염기 편집기를 포함하는, 방법.

청구항 92

제91항에 있어서, 핵염기 편집 도메인은 ABE7.10 또는 ABE8.20의 아데노신 염기 편집기를 포함하는, 방법.

청구항 93

제87항에 있어서, 핵염기 편집 도메인은 시티딘 데아미나제인, 방법.

청구항 94

제93항에 있어서, 시티딘 데아미나제는 *페트로미존 마리누스* 시토신 데아미나제 1(PmCDA1), 활성화-유도된 시티딘 데아미나제(AID) 및 APOBEC로 구성된 군으로부터 선택되는, 방법.

청구항 95

제84항 내지 제94항 중 어느 한 항에 있어서, 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인은 Cas9 폴리펩티드, Cas12 폴리펩티드, 또는 이의 기능적 변이체를 포함하는, 방법.

청구항 96

제84항 내지 제95항 중 어느 한 항에 있어서, 재조합 게놈은 가이드 RNA(gRNA)를 추가로 포함하는, 방법.

청구항 97

제96항에 있어서, 가이드 RNA(gRNA)는 표적 세포에 외인적으로 제공되는, 방법.

청구항 98

제96항 또는 제97항에 있어서, gRNA는 표적 세포의 게놈 좌위를 표적화할 수 있는, 방법.

청구항 99

제83항 내지 제98항 중 어느 한 항에 있어서, 대상체는 포유동물인, 방법.

청구항 100

제83항 내지 제99항 중 어느 한 항에 있어서, 대상체는, 인간인, 방법.

청구항 101

광견병 바이러스 입자의 재조합 제조를 위한 패키징 시스템으로서, 상기 패키징 시스템은,

광견병 바이러스 핵단백질 또는 이의 기능적 변이체를 인코딩하는 N 유전자;

광견병 바이러스 인단백질 또는 이의 기능적 변이체를 인코딩하는 P 유전자;

광견병 바이러스 폴리머라제 또는 이의 기능적 변이체를 인코딩하는 L 유전자; 및

재조합 광견병 바이러스 게놈을 포함하며,

게놈은 광견병 바이러스 당단백질 또는 이의 기능적 변이체를 인코딩하는 G 유전자를 결하고/결하거나;

게놈은 광견병 바이러스 폴리머라제 또는 이의 기능적 변이체를 인코딩하는 L 유전자를 결하는, 패키징 시스템.

청구항 102

제101항에 있어서, 재조합 광견병 바이러스 게놈은 이식유전자 또는 치료적 이식유전자를 인코딩하는 핵산을 포함하는, 패키징 시스템.

청구항 103

제101항 또는 제102항에 있어서, 재조합 광견병 바이러스 게놈은 바이러스 게놈 벡터 내에 포함되는, 패키징 시스템.

청구항 104

제101항 내지 제103항 중 어느 한 항에 있어서, N, P 및 L 유전자는 각각 별도의 벡터 내에 포함되는, 패키징 시스템.

청구항 105

제104항에 있어서, N, P 및 L 유전자 각각은 전사 조절 요소에 작동가능하게 연결되는, 패키징 시스템.

청구항 106

제105항에 있어서, 전사 조절 요소는 프로모터 및/또는 인핸서를 포함하는, 패키징 시스템.

청구항 107

제106항에 있어서, 프로모터는 구성적 프로모터인, 패키징 시스템.

청구항 108

제106항 또는 제107항에 있어서, 프로모터는 신장 인자 1α 프로모터인, 패키징 시스템.

청구항 109

제104항 내지 제108항 중 어느 한 항에 있어서, 별도의 벡터는 각각 별도의 형질감염 플라스미드 내에 함유되어 있는, 패키징 시스템.

청구항 110

제101항 내지 제103항 중 어느 한 항에 있어서, N, P 및 L 유전자는 단일 벡터 내에 포함되는, 패키징 시스템.

청구항 111

제110항에 있어서, 단일 벡터는 N 및 P 유전자를 포함하는 제1 발현 카세트, 및 L 유전자를 포함하는 제2 발현 카세트를 포함하는, 패키징 시스템.

청구항 112

제111항에 있어서, 제1 발현 카세트는 5'에서 3'으로:

전사 조절 요소;

P 유전자; 및

N 유전자를 포함하는, 패키징 시스템.

청구항 113

제111항 또는 제112항에 있어서, 제1 발현 카세트는 5'에서 3'으로:

전사 조절 요소;

P 유전자;

리보솜 스키핑 요소; 및

N 유전자를 포함하는, 패키징 시스템.

청구항 114

제113항에 있어서, 리보솜 스키핑 요소는 IRES 요소인, 패키징 시스템.

청구항 115

제113항에 있어서, 리보솜 스키핑 요소는 2A 요소인, 패키징 시스템.

청구항 116

제111항 내지 제115항 중 어느 한 항에 있어서, 제2 발현 카세트는 5'에서 3'으로:

전사 조절 요소; 및

L 유전자를 포함하는, 패키징 시스템.

청구항 117

제112항 내지 제116항 중 어느 한 항에 있어서, 전사 조절 요소는 프로모터 및/또는 인핸서를 포함하는, 패키징 시스템.

청구항 118

제117항에 있어서, 프로모터는 구성적 프로모터인, 패키징 시스템.

청구항 119

제117항 또는 제118항에 있어서, 프로모터는 신장 인자 1α 프로모터인, 패키징 시스템.

청구항 120

제111항 내지 제119항 중 어느 한 항에 있어서, 제1 발현 카세트와 제2 발현 카세트는 벡터에서 반대 방향으로 있는, 패키징 시스템.

청구항 121

제110항 내지 제120항 중 어느 한 항에 있어서, 단일 벡터는 단일 형질감염 플라스미드 내에 함유되는, 패키징 시스템.

청구항 122

제101항 내지 제121항 중 어느 한 항에 있어서, 광견병 바이러스 매트릭스 단백질 또는 이의 기능적 변이체를 인코딩하는 M 유전자를 추가로 포함하는, 패키징 시스템.

청구항 123

제122항에 있어서, M 유전자는 벡터 내에 포함되는, 패키징 시스템.

청구항 124

제122항 또는 제123항에 있어서, M 유전자는 전사 조절 요소에 작동가능하게 연결된, 패키징 시스템.

청구항 125

제124항에 있어서, 전사 조절 요소는 프로모터 및/또는 인핸서를 포함하는, 패키징 시스템.

청구항 126

제123항 내지 제125항 중 어느 한 항에 있어서, M 유전자를 포함하는 벡터는 형질감염 플라스미드 내에 함유되는, 패키징 시스템.

청구항 127

제101항 내지 제126항 중 어느 한 항에 있어서, 광견병 바이러스 당단백질 또는 이의 기능적 변이체를 인코딩하는 G 유전자를 추가로 포함하는, 패키징 시스템.

청구항 128

제127항에 있어서, G 유전자는 벡터 내에 포함되는, 패키징 시스템.

청구항 129

제127항 또는 제128항에 있어서, G 유전자는 전사 조절 요소에 작동가능하게 연결된, 패키징 시스템.

청구항 130

제129항에 있어서, 전사 조절 요소는 프로모터 및/또는 인핸서를 포함하는, 패키징 시스템.

청구항 131

제127항 내지 제130항 중 어느 한 항에 있어서, G 유전자를 포함하는 벡터는 형질감염 플라스미드 내에 함유되는, 패키징 시스템.

청구항 132

재조합 광견병 바이러스 입자를 생산하기 위한 방법으로서, 상기 방법은 재조합 광견병 바이러스 입자를 형성하기 위해 재조합 광견병 바이러스 게놈을 봉입하도록 작동하는 조건 하에서 제101항 내지 제131항 중 어느 한 항의 패키징 시스템을 세포 안으로 도입하는 단계를 포함하는, 방법.

청구항 133

제132항에 있어서, 도입하는 단계는 전기천공, 뉴클레오펙션 또는 리포펙션에 의해 매개되는, 방법.

청구항 134

제101항 내지 제131항 중 어느 한 항의 패키징 시스템을 포함하는 재조합 광견병 바이러스 입자 패키징 세포.

청구항 135

재조합 광견병 바이러스 입자 패키징 세포로서,

광견병 바이러스 핵단백질 또는 이의 기능적 변이체를 인코딩하는 N 유전자;

광견병 바이러스 인단백질 또는 이의 기능적 변이체를 인코딩하는 P 유전자; 및

광견병 바이러스 폴리머라제 또는 이의 기능적 변이체를 인코딩하는 L 유전자를 포함하는, 패키징 세포.

청구항 136

제135항에 있어서, 제1 발현 카세트는 N 및 P 유전자를 포함하는, 패키징 세포.

청구항 137

제136항에 있어서, 제1 발현 카세트는 5'에서 3'으로:

전사 조절 요소;

P 유전자; 및

N 유전자를 포함하는, 패키징 세포.

청구항 138

제136항 또는 제137항에 있어서, 제1 발현 카세트는 5'에서 3'으로:

전사 조절 요소;

P 유전자;

리보솜 스키핑 요소; 및

N 유전자를 포함하는, 패키징 세포.

청구항 139

제138항에 있어서, 리보솜 스키핑 요소는 IRES 요소인, 패키징 세포.

청구항 140

제138항에 있어서, 리보솜 스키핑 요소는 2A 요소인, 패키징 세포.

청구항 141

제135항 내지 제140항 중 어느 한 항에 있어서, 제2 발현 카세트는 L 유전자를 포함하는, 패키징 세포.

청구항 142

제141항에 있어서, 제2 발현 카세트는 5'에서 3'으로:

전사 조절 요소; 및

L 유전자를 포함하는, 패키징 세포.

청구항 143

제137항 내지 제142항 중 어느 한 항에 있어서, 전사 조절 요소는 프로모터 및/또는 인핸서를 포함하는, 패키징 세포.

청구항 144

제143항에 있어서, 프로모터는 구성적 프로모터인, 패키징 세포.

청구항 145

제143항 또는 제144항에 있어서, 프로모터는 신장 인자 1α 프로모터인, 패키징 세포.

청구항 146

제142항 내지 제145항 중 어느 한 항에 있어서, 제1 발현 카세트와 제2 발현 카세트는 벡터에서 반대 방향으로 있는, 패키징 세포.

청구항 147

제137항 내지 제146항 중 어느 한 항에 있어서, 광견병 바이러스 매트릭스 단백질을 인코딩하는 M 유전자 및/또는 광견병 바이러스 당단백질을 인코딩하는 G 유전자를 추가로 포함하는, 패키징 세포.

청구항 148

제147항에 있어서, M 유전자 및/또는 G 유전자는 전사 조절 요소에 작동가능하게 연결된, 패키징 세포.

청구항 149

제148항에 있어서, 전사 조절 요소는 프로모터 및/또는 인핸서를 포함하는, 패키징 세포.

청구항 150

제147항 내지 제149항 중 어느 한 항에 있어서, 제3 발현 및/또는 제4 카세트는 광견병 바이러스 G 유전자 및/또는 광견병 바이러스 M 유전자를 포함하는, 패키징 세포.

청구항 151

제134항 내지 제150항 중 어느 한 항에 있어서, 패키징 세포는 포유동물, 박테리아 또는 곤충 기원의 것인, 패키징 세포.

청구항 152

제134항 내지 제151항 중 어느 한 항에 있어서, 패키징 세포는 HEK293 세포, VERO 세포, BHK 세포 및 BSR 세포로 구성된 군으로부터 선택되는, 패키징 세포.

청구항 153

재조합 광견병 바이러스를 생산하는 방법으로서, 상기 방법은 재조합 광견병 바이러스를 형성하도록 당단백질에 재조합 광견병 바이러스 게놈을 봉입하도록 작동하는 조건 하에서, 치료적 이식유전자를 포함하는 핵산을 포함하는 재조합 광견병 바이러스 게놈을 포함하는 핵산을 제134항 내지 제152항 중 어느 한 항의 패키징 세포 안으로 도입하는 단계를 포함하며: 여기서

게놈은 광견병 바이러스 당단백질을 인코딩하는 내인성 G 유전자를 결하고; 그리고

게놈은 광견병 바이러스 폴리머라제를 인코딩하는 내인성 L 유전자를 결하는, 방법.

청구항 154

제153항에 있어서, 재조합 게놈은,

광견병 바이러스 핵단백질 또는 이의 기능적 변이체를 인코딩하는 N 유전자;

광견병 바이러스 인단백질 또는 이의 기능적 변이체를 인코딩하는 P 유전자; 및

광견병 바이러스 매트릭스 단백질 또는 이의 기능적 변이체를 인코딩하는 M 유전자를 포함하는, 방법.

청구항 155

제153항 또는 제154항에 있어서, 각각의 유전자는 전사 조절 요소에 작동가능하게 연결되는, 방법.

청구항 156

제155항에 있어서, 전사 조절 요소는 전사 개시 신호를 포함하는, 방법.

청구항 157

제156항에 있어서, 전사 개시 신호는 광견병 바이러스에 대해 외인성인, 방법.

청구항 158

제156항에 있어서, 전사 개시 신호는 광견병 바이러스에 대해 내인성인, 방법.

청구항 159

제153항 내지 제158항 중 어느 한 항에 있어서, 각각의 유전자는 전사 종결 폴리아데닐화 신호에 작동가능하게 연결되는, 방법.

청구항 160

제153항 내지 제159항 중 어느 한 항에 있어서, 재조합 광견병 바이러스 역가는 약 1E8 TU/mL보다 큰, 방법.

청구항 161

제153항 내지 제159항 중 어느 한 항에 있어서, 재조합 광견병 바이러스 역가는 약 1E8 TU/mL 내지 약 1E9 TU/mL인, 방법.

청구항 162

대상체에서 질환 또는 장애를 치료하는 방법으로서, 상기 방법은 제6항 내지 제51항 중 어느 한 항의 재조합 광견병 바이러스 입자 또는 제32항의 약학 조성물을 대상체에게 투여하는 단계를 포함하는, 방법.

청구항 163

제162항에 있어서, 질환 또는 장애는 신경학적 질환 또는 장애인, 방법.

청구항 164

제162항에 있어서, 질환 또는 장애는 안과 질환 또는 장애인, 방법.

청구항 165

대상체에서의 질환 또는 장애를 치료하기 위한 약제의 제조에 있어서, 제6항 내지 제51항 중 어느 한 항의 재조합 광견병 바이러스 또는 제52항의 약학적 조성물의 용도.

발명의 설명

기 술 분 야

[0001] **관련된 출원**

[0002] 본 출원은 2021년 2월 19일에 출원된 미국 가특허출원 일련번호 63/151,542 및 2021년 9월 8일에 출원된 미국 가특허출원 일련번호 63/241,989의 이익을 주장하며, 그 전체 개시내용은 본 명세서에 참조로 포함된다.

배 경 기 술

[0003] 유전자 요법은 관심있는 세포를 형질도입하고 이식유전자를 발현하기 위해 주로 바이러스 유전자 전달 시스템의 사용을 포함한다. 유전자 요법에 일반적으로 사용되는 바이러스 시스템은 바이러스에서 유래되고 유의한 단점을 안고 있다. 현재 바이러스 시스템을 사용함에 있어 어려움은: 제한된 작은 패키징 용량; 게놈 통합과 같은 의도되지 않은 결과적인 귀결; 제한된 조직 친화성; 표적 모집단에서 기존 면역 및/또는 면역 반응, 재-복용에 대한 제한된 능력; 및 게놈 불안정성, 면역 제거 및 세포 독성에 기인한 제한된 내구성을 포함한다. 예를 들어, 아데노바이러스 벡터-매개된 유전자 요법은 높은 형질도입 효율을 입증하고 많은 상이한 세포 유형을 감염시키는데 사용될 수 있는 반면, 특정한 단점은 비-통합 및 높은 면역원성을 포함한다. 아데노-연관된 바이러스 벡터-매개된 유전자 요법의 단점은 높은 면역원성과 제한된 패키징 용량을 포함한다. 또 다른 예로서, 레트로바이러스 벡터-매개된 유전자 요법은 낮은 형질도입 효율 및 보체에 의한 불활성화 문제를 겪고 있다.

[0004] 따라서, 현재의 바이러스 시스템에 비해 장점이 있는 신규한 바이러스 유전자 전달 시스템에 대한 필요성이 있다.

발명의 내용

[0005] 표적 세포를 형질도입하고 그 안에서 이식유전자를 발현하는데 사용될 수 있는, 광견병 바이러스로부터 유래된 재조합 바이러스 벡터 및 재조합 바이러스가 본 명세서에 제공된다. 본 개시내용에 의해 제공된 재조합 광견병 벡터 및 바이러스는 효과적인 바이러스 유전자 전달 시스템으로서의 용도를 발견한다. 본 명세서에 기재된 재조합 바이러스를 생산하는 바이러스 패키징 시스템 및 방법이 또한 제공된다.

[0006] 일 양태에서, 치료적 이식유전자를 인코딩하는 핵산을 포함하는 재조합 광견병 바이러스 게놈으로서, 여기서 게놈은 광견병 바이러스 당단백질 또는 이의 기능적 변이체를 인코딩하는 G 유전자를 결하고/결하거나; 게놈은 광견병 바이러스 폴리머라제 또는 이의 기능적 변이체를 인코딩하는 L 유전자를 결하는, 재조합 광견병 바이러스 게놈이 제공된다.

[0007] 특정한 예시적 실시형태에서, 게놈은 광견병 바이러스 당단백질 또는 이의 기능적 변이체를 인코딩하는 G 유전자를 결여한다. 특정한 예시적 실시형태에서, 게놈은 광견병 바이러스 당단백질 또는 이의 기능적 변이체를 인코딩하는 G 유전자를 결하고, 게놈은 광견병 바이러스 폴리머라제 또는 이의 기능적 변이체를 인코딩하는 L 유전자를 결한다.

[0008] 특정한 예시적 실시형태에서, 게놈은 광견병 바이러스 핵단백질 또는 이의 기능적 변이체를 인코딩하는 N 유전자; 광견병 바이러스 인단백질 또는 이의 기능적 변이체를 인코딩하는 P 유전자; 및 광견병 바이러스 매트릭스 단백질 또는 이의 기능적 변이체를 인코딩하는 M 유전자를 포함한다.

[0009] 또 다른 양태에서, 본 명세서에 기재된 바와 같은 광견병 바이러스 당단백질 및 재조합 광견병 바이러스 게놈을 포함하는 재조합 광견병 바이러스 입자가 제공된다.

[0010] 또 다른 양태에서, 광견병 바이러스 당단백질; 및 치료적 이식유전자를 인코딩하는 핵산을 포함하는 재조합 광견병 바이러스 게놈을 포함하는, 재조합 광견병 바이러스 입자가 제공되며, 여기서: 게놈은 광견병 바이러스 당단백질 또는 이의 기능적 변이체를 인코딩하는 G 유전자를 결하고/결하거나; 게놈은 광견병 바이러스 폴리머라제 또는 이의 기능적 변이체를 인코딩하는 L 유전자를 결한다.

[0011] 특정한 예시적 실시형태에서, 게놈은 광견병 바이러스 당단백질 또는 이의 기능적 변이체를 인코딩하는 G 유전자를 결한다. 특정한 예시적 실시형태에서, 게놈은 광견병 바이러스 당단백질 또는 이의 기능적 변이체를 인코딩하는 G 유전자를 결하고, 그리고 여기서 게놈은 광견병 바이러스 폴리머라제 또는 이의 기능적 변이체를 인코딩하는 L 유전자를 결한다.

[0012] 특정한 예시적 실시형태에서, 게놈은 다음을 결한다:

[0013] 광견병 바이러스 당단백질 또는 이의 기능적 변이체를 인코딩하는 G 유전자;

[0014] 광견병 바이러스 폴리머라제 또는 이의 기능적 변이체를 인코딩하는 L 유전자; 및

[0015] 광견병 바이러스 매트릭스 단백질 또는 이의 기능적 변이체를 인코딩하는 M 유전자.

[0016] 특정한 예시적 실시형태에서, 게놈은 다음을 결한다:

[0017] 광견병 바이러스 당단백질 또는 이의 기능적 변이체를 인코딩하는 G 유전자;

[0018] 광견병 바이러스 폴리머라제 또는 이의 기능적 변이체를 인코딩하는 L 유전자;

[0019] 광견병 바이러스 매트릭스 단백질 또는 이의 기능적 변이체를 인코딩하는 M 유전자; 및

[0020] 광견병 바이러스 인단백질 또는 이의 기능적 변이체를 인코딩하는 P 유전자.

[0021] 특정한 예시적 실시형태에서, 게놈은 다음을 결한다:

[0022] 광견병 바이러스 당단백질 또는 이의 기능적 변이체를 인코딩하는 G 유전자;

[0023] 광견병 바이러스 폴리머라제 또는 이의 기능적 변이체를 인코딩하는 L 유전자;

[0024] 광견병 바이러스 매트릭스 단백질 또는 이의 기능적 변이체를 인코딩하는 M 유전자;

[0025] 광견병 바이러스 인단백질 또는 이의 기능적 변이체를 인코딩하는 P 유전자; 및

[0026] 광견병 바이러스 핵단백질 또는 이의 기능적 변이체를 인코딩하는 N 유전자.

[0027] 특정한 예시적 실시형태에서, 게놈은 치료적 이식유전자만을 인코딩한다.

[0028] 특정한 예시적 실시형태에서, 게놈은 광견병 바이러스 핵단백질 또는 이의 기능적 변이체를 인코딩하는 N 유전자; 광견병 바이러스 인단백질 또는 이의 기능적 변이체를 인코딩하는 P 유전자; 및 광견병 바이러스 매트릭스 단백질 또는 이의 기능적 변이체를 인코딩하는 M 유전자를 포함한다.

[0029] 다른 예시적 실시형태에서, 게놈은 광견병 바이러스 핵단백질 또는 이의 기능적 변이체를 인코딩하는 N 유전자; 광견병 바이러스 인단백질 또는 이의 기능적 변이체를 인코딩하는 P 유전자; 및/또는 광견병 바이러스 매트릭스 단백질 또는 이의 기능적 변이체를 인코딩하는 M 유전자를 결한다.

[0030] 특정한 예시적 실시형태에서, 재조합 광견병 바이러스는 복제 불능이다. 다른 예시적인 실시형태에서, 재조합 광견병 바이러스는 복제 결핍성이다.

[0031] 특정한 예시적 실시형태에서, 각각의 유전자는 전사 조절 요소에 작동가능하게 연결된다. 특정한 예시적 실시형태에서, 전사 조절 요소는 전사 개시 신호를 포함한다. 특정한 예시적 실시형태에서, 전사 개시 신호는 광견병 바이러스에 대해 외인성이다. 특정한 예시적 실시형태에서, 전사 개시 신호는 광견병 바이러스에 내인성이다.

[0032] 특정한 예시적 실시형태에서, 각각의 유전자는 전사 종결 폴리아데닐화 신호에 작동가능하게 연결된다.

[0033] 특정한 예시적 실시형태에서, 치료적 이식유전자는 핵산 편집 시스템 또는 이의 성분을 인코딩하는 서열을 포함한다. 특정한 예시적 실시형태에서, 핵산 편집 시스템은 클러스터링된 규칙적으로 간격을 둔 짧은 회문 반복(CRISPR) 시스템, 아연 핑거 뉴클레아제(ZFN), 메가뉴클레아제, 및 전사 활성제-유사 효과기-기반 뉴클레아제(TALEN)를 포함한다. 특정한 예시적 실시형태에서, 핵산 편집 시스템은 CRISPR 시스템을 포함한다. 특정한 예시적 실시형태에서, CRISPR-시스템은 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인 및 핵염기 편집 도메인을 포함하는 핵염기 편집기를 포함한다.

[0034] 특정한 예시적 실시형태에서, 핵염기 편집 도메인은 아데노신 데아미나제, 시티딘 데아미나제, 또는 이의 기능적 변이체이다. 특정한 예시적 실시형태에서, 핵염기 편집 도메인은 아데노신 데아미나제이다. 특정한 예시적 실시형태에서, 아데노신 데아미나제는 표 10, 표 11, 표 12, 표 13, 표 14 또는 표 15에 인용된 임의의 아데노신 염기 편집기로부터의 TadA 데아미나제를 포함한다.

[0035] 특정한 예시적 실시형태에서 핵염기 편집 도메인은 다음의 아데노신 염기 편집기 중 임의의 하나로부터의 아데노신 데아미나제를 포함한다: ABE 0.1, ABE 0.2, ABE 1.1, ABE 1.2, ABE2.1, ABE2.2, ABE2.3, ABE2.4, ABE2.5, ABE2.6, ABE2.7, ABE2.8, ABE2.9, ABE2.10, ABE2.11, ABE2.12, ABE3.1, ABE3.2, ABE3.3, ABE3.4, ABE3.5, ABE3.6, ABE3.7, ABE3.8, ABE4.1, ABE4.2, ABE4.3, ABE5.1, ABE5.2, ABE5.3, ABE5.4, ABE5.5, ABE5.6,

ABE5.7, ABE5.8, ABE5.9, ABE5.10, ABE5.11, ABE5.12, ABE5.13, ABE5.14, ABE6.1, ABE6.2, ABE6.3, ABE6.4, ABE6.5, ABE6.6, ABE7.1, ABE7.2, ABE7.3, ABE7.4, ABE7.5, ABE7.6, ABE7.7, ABE7.8, ABE 7.9, ABE7.10, ABE8.1-m, ABE8.2-m, ABE8.3-m, ABE8.4-m, ABE8.5-m, ABE8.6-m, ABE8.7-m, ABE8.8-m, ABE8.9-m, ABE8.10-m, ABE8.11-m, ABE8.12-m, ABE8.13-m, ABE8.14-m, ABE8.15-m, ABE8.16-m, ABE8.17-m, ABE8.18-m, ABE8.19-m, ABE8.20-m, ABE8.21-m, ABE8.22-m, ABE8.23-m, ABE8.24-m, ABE8.1-d, ABE8.2-d, ABE8.3-d, ABE8.4-d, ABE8.5-d, ABE8.6-d, ABE8.7-d, ABE8.8-d, ABE8.9-d, ABE8.10-d, ABE8.11-d, ABE8.12-d, ABE8.13-d, ABE8.14-d, ABE8.15-d, ABE8.16-d, ABE8.17-d, ABE8.18-d, ABE8.19-d, ABE8.20-d, ABE8.21-d, ABE8.22-d, ABE8.23-d, ABE8.24-d, ABE8a-m, ABE8b-m, ABE8c-m, ABE8d-m, ABE8e-m, ABE8a-d, ABE8b-d, ABE8c-d, ABE8d-d, ABE8e-d, ABE9.1, ABE9.2, ABE9.3, ABE9.4, ABE9.5, ABE9.6, ABE9.7, ABE9.8, ABE9.9, ABE9.10, ABE9.11, ABE9.12, ABE9.13, ABE9.14, ABE9.15, ABE9.16, ABE9.17, ABE9.18, ABE9.19, ABE9.2, ABE9.21, ABE9.22, ABE9.23, ABE9.24, ABE9.25, ABE9.26, ABE9.27, ABE9.28, ABE9.29, ABE9.30, ABE9.31, ABE9.32, ABE9.33, ABE9.34, ABE9.35, ABE9.36, ABE9.37, ABE9.38, ABE9.39, ABE9.40, ABE9.41, ABE9.42, ABE9.43, ABE9.44, ABE9.45, ABE9.46, ABE9.47, ABE9.48, ABE9.49, ABE9.50, ABE9.51, ABE9.52, ABE9.53, ABE9.54, ABE9.55, ABE9.56, ABE9.57, 및 ABE9.58.

[0036] 특정한 예시적 실시형태에서 핵염기 편집 도메인은 다음으로 구성된 군으로부터 선택되는 아데노신 염기 편집기를 포함한다: ABE 0.1, ABE 0.2, ABE 1.1, ABE 1.2, ABE2.1, ABE2.2, ABE2.3, ABE2.4, ABE2.5, ABE2.6, ABE2.7, ABE2.8, ABE2.9, ABE2.10, ABE2.11, ABE2.12, ABE3.1, ABE3.2, ABE3.3, ABE3.4, ABE3.5, ABE3.6, ABE3.7, ABE3.8, ABE4.1, ABE4.2, ABE4.3, ABE5.1, ABE5.2, ABE5.3, ABE5.4, ABE5.5, ABE5.6, ABE5.7, ABE5.8, ABE5.9, ABE5.10, ABE5.11, ABE5.12, ABE5.13, ABE5.14, ABE6.1, ABE6.2, ABE6.3, ABE6.4, ABE6.5, ABE6.6, ABE7.1, ABE7.2, ABE7.3, ABE7.4, ABE7.5, ABE7.6, ABE7.7, ABE7.8, ABE 7.9, ABE7.10, ABE8.1-m, ABE8.2-m, ABE8.3-m, ABE8.4-m, ABE8.5-m, ABE8.6-m, ABE8.7-m, ABE8.8-m, ABE8.9-m, ABE8.10-m, ABE8.11-m, ABE8.12-m, ABE8.13-m, ABE8.14-m, ABE8.15-m, ABE8.16-m, ABE8.17-m, ABE8.18-m, ABE8.19-m, ABE8.20-m, ABE8.21-m, ABE8.22-m, ABE8.23-m, ABE8.24-m, ABE8.1-d, ABE8.2-d, ABE8.3-d, ABE8.4-d, ABE8.5-d, ABE8.6-d, ABE8.7-d, ABE8.8-d, ABE8.9-d, ABE8.10-d, ABE8.11-d, ABE8.12-d, ABE8.13-d, ABE8.14-d, ABE8.15-d, ABE8.16-d, ABE8.17-d, ABE8.18-d, ABE8.19-d, ABE8.20-d, ABE8.21-d, ABE8.22-d, ABE8.23-d, ABE8.24-d, ABE8a-m, ABE8b-m, ABE8c-m, ABE8d-m, ABE8e-m, ABE8a-d, ABE8b-d, ABE8c-d, ABE8d-d, ABE8e-d, ABE9.1, ABE9.2, ABE9.3, ABE9.4, ABE9.5, ABE9.6, ABE9.7, ABE9.8, ABE9.9, ABE9.10, ABE9.11, ABE9.12, ABE9.13, ABE9.14, ABE9.15, ABE9.16, ABE9.17, ABE9.18, ABE9.19, ABE9.2, ABE9.21, ABE9.22, ABE9.23, ABE9.24, ABE9.25, ABE9.26, ABE9.27, ABE9.28, ABE9.29, ABE9.30, ABE9.31, ABE9.32, ABE9.33, ABE9.34, ABE9.35, ABE9.36, ABE9.37, ABE9.38, ABE9.39, ABE9.40, ABE9.41, ABE9.42, ABE9.43, ABE9.44, ABE9.45, ABE9.46, ABE9.47, ABE9.48, ABE9.49, ABE9.50, ABE9.51, ABE9.52, ABE9.53, ABE9.54, ABE9.55, ABE9.56, ABE9.57, 및 ABE9.58. 특정한 예시적 실시형태에서, 아데노신 데아미나제는 ABE7.10 또는 ABE8.20이다.

[0037] 특정한 예시적 실시형태에서 핵염기 편집 도메인은 시티딘 데아미나제를 포함한다. 특정한 예시적 실시형태에서 시티딘 데아미나제는 *페트로미존 마리누스* 시토신 데아미나제 1(PmCDA1), 활성화-유도된 시티딘 데아미나제(AID) 및 APOBEC로 구성된 군으로부터 선택된다.

[0038] 특정한 예시적 실시형태에서, 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인은 Cas9 폴리펩티드, Cas12 폴리펩티드, 또는 이의 기능적 변이체를 포함한다.

[0039] 특정한 예시적 실시형태에서, CRISPR-시스템은 가이드 RNA(gRNA) 또는 gRNA를 인코딩하는 핵산 서열을 추가로 포함한다.

[0040] 특정한 예시적 실시형태에서, 치료적 이식유전자는 우라실 글리코실라제 억제제(UGI)를 인코딩하는 서열을 포함한다.

[0041] 특정한 예시적 실시형태에서, 치료적 이식유전자는 핵 국소화 신호(NLS)를 인코딩하는 서열을 포함한다.

[0042] 특정한 예시적 실시형태에서, 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인 및/또는 핵염기 편집 도메인은 핵 국소화 신호(NLS)를 추가로 포함한다.

[0043] 특정한 예시적 실시형태에서, 시티딘 데아미나제는 우라실 글리코실라제 억제제(UGI)를 추가로 포함한다.

[0044] 특정한 예시적 실시형태에서, 치료적 이식유전자는 치료적 폴리펩티드 및/또는 치료적 핵산을 포함한다. 특정한 예시적 실시형태에서, 치료적 폴리펩티드 및/또는 치료적 핵산은 세포로부터 분비된다.

[0045] 특정한 예시적 실시형태에서, 치료적 이식유전자를 인코딩하는 핵산은 약 300bp, 약 400bp, 약 500bp, 약 600bp, 약 700bp, 약 800bp, 약 900bp, 약 1,000bp, 약 1,100bp, 약 1,200bp, 약 1,300bp, 약 1,400bp, 약 1,500bp, 약 1,600bp, 약 1,700bp, 약 1,800bp, 약 1,900bp, 약 2,000bp, 약 2,100bp, 약 2,200bp, 약 2,300bp , 약 2,400bp, 약 2,500bp, 약 2,600bp, 약 2,700bp, 약 2,800bp, 약 2,900bp, 또는 약 3,000bp보다 크다. 특정한 예시적 실시형태에서, 치료적 이식유전자를 인코딩하는 핵산은 약 300bp보다 크다. 특정한 예시적 실시형태에서, 치료적 이식유전자를 인코딩하는 핵산은 약 650bp보다 크다. 특정한 예시적 실시형태에서, 치료적 이식유전자를 인코딩하는 핵산은 약 1,000bp보다 크다. 특정한 예시적 실시형태에서, 치료적 이식유전자를 인코딩하는 핵산은 약 3,000bp보다 크다. 특정한 예시적 실시형태에서, 치료적 이식유전자를 인코딩하는 핵산은 약 4,500bp보다 크다. 특정한 예시적 실시형태에서, 치료적 이식유전자를 인코딩하는 핵산은 약 8,500bp보다 크다. 특정한 예시적 실시형태에서, 치료적 이식유전자를 인코딩하는 핵산은 약 10,000bp보다 크다.

[0046] 특정한 예시적 실시형태에서, 치료적 이식유전자는 전사 조절 요소에 작동가능하게 연결된다. 특정한 예시적 실시형태에서, 전사 조절 요소는 전사 개시 신호를 포함한다. 특정한 예시적 실시형태에서, 전사 개시 신호는 광견병 바이러스에 대해 외인성이다. 특정한 예시적 실시형태에서, 전사 개시 신호는 광견병 바이러스에 내인성이다.

[0047] 특정한 예시적 실시형태에서, 치료적 이식유전자는 전사 종결 폴리아데닐화 신호에 작동가능하게 연결된다.

[0048] 또 다른 양태에서, 본 명세서에 기재된 바와 같은 재조합 바이러스 입자를 포함하는 약학적 조성물이 제공된다.

[0049] 또 다른 양태에서, 본 명세서에 기재된 바와 같은 재조합 바이러스 입자로 표적 세포를 형질도입시키는 것을 포함하는, 표적 세포에서 치료적 이식유전자를 발현시키는 방법이 제공된다.

[0050] 또 다른 양태에서, 표적 세포에 재조합 광견병 바이러스 입자를 형질도입시키는 것을 포함하는 표적 세포에서 핵염기 편집기를 발현시키는 방법이 제공되며, 여기서 재조합 바이러스 입자는 광견병 바이러스 당단백질; 및 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인 및 핵염기 편집 도메인을 포함하는 핵염기 편집기를 인코딩하는 핵산을 포함하는 재조합 광견병 바이러스 게놈을 포함하며, 여기서: 게놈은 광견병 바이러스 당단백질 또는 이의 기능적 변이체를 인코딩하는 G 유전자를 결하고/결하거나; 게놈은 광견병 바이러스 폴리머라제 또는 이의 기능적 변이체를 인코딩하는 L 유전자를 결한다.

[0051] 특정한 예시적 실시형태에서, 게놈은 다음을 추가로 결한다:

[0052] 광견병 바이러스 매트릭스 단백질 또는 이의 기능적 변이체를 인코딩하는 M 유전자.

[0053] 특정한 예시적 실시형태에서, 게놈은 다음을 추가로 결한다:

[0054] 광견병 바이러스 매트릭스 단백질 또는 이의 기능적 변이체를 인코딩하는 M 유전자; 및

[0055] 광견병 바이러스 인단백질 또는 이의 기능적 변이체를 인코딩하는 P 유전자.

[0056] 특정한 예시적 실시형태에서, 게놈은 다음을 추가로 결한다:

[0057] 광견병 바이러스 매트릭스 단백질 또는 이의 기능적 변이체를 인코딩하는 M 유전자;

[0058] 광견병 바이러스 인단백질 또는 이의 기능적 변이체를 인코딩하는 P 유전자; 및

[0059] 광견병 바이러스 핵단백질 또는 이의 기능적 변이체를 인코딩하는 N 유전자.

[0060] 특정한 예시적 실시형태에서, 게놈은 핵염기 편집기만을 인코딩한다.

[0061] 특정한 예시적 실시형태에서, 게놈은 광견병 바이러스 핵단백질 또는 이의 기능적 변이체를 인코딩하는 N 유전자; 광견병 바이러스 인단백질 또는 이의 기능적 변이체를 인코딩하는 P 유전자; 및 광견병 바이러스 매트릭스 단백질 또는 이의 기능적 변이체를 인코딩하는 M 유전자를 포함한다.

[0062] 특정한 예시적 실시형태에서, 각각의 유전자는 전사 조절 요소에 작동가능하게 연결된다. 특정한 예시적 실시형태에서, 전사 조절 요소는 전사 개시 신호를 포함한다. 특정한 예시적 실시형태에서, 전사 개시 신호는 광견병 바이러스에 대해 외인성이다. 특정한 예시적 실시형태에서, 전사 개시 신호는 광견병 바이러스에 내인성이다.

[0063] 특정한 예시적 실시형태에서, 각각의 유전자는 전사 종결 폴리아데닐화 신호에 작동가능하게 연결된다.

[0064] 특정한 예시적 실시형태에서, 핵염기 편집 도메인은 아데노신 데아미나제, 시티딘 데아미나제, 또는 이의 기능

적 변이체를 포함한다.

[0065] 특정한 예시적 실시형태에서, 핵염기 편집 도메인은 아데노신 데아미나제를 포함한다. 특정한 예시적 실시형태에서, 아데노신 데아미나제는 표 10, 표 11, 표 12, 표 13, 표 14 또는 표 15에 인용된 임의의 아데노신 염기 편집기로부터의 TadA 데아미나제를 포함한다.

[0066] 특정한 예시적 실시형태에서, 핵염기 편집 도메인은 다음의 아데노신 염기 편집기 중 임의의 하나로부터의 아데노신 데아미나제를 포함한다: ABE 0.1, ABE 0.2, ABE 1.1, ABE 1.2, ABE2.1, ABE2.2, ABE2.3, ABE2.4, ABE2.5, ABE2.6, ABE2.7, ABE2.8, ABE2.9, ABE2.10, ABE2.11, ABE2.12, ABE3.1, ABE3.2, ABE3.3, ABE3.4, ABE3.5, ABE3.6, ABE3.7, ABE3.8, ABE4.1, ABE4.2, ABE4.3, ABE5.1, ABE5.2, ABE5.3, ABE5.4, ABE5.5, ABE5.6, ABE5.7, ABE5.8, ABE5.9, ABE5.10, ABE5.11, ABE5.12, ABE5.13, ABE5.14, ABE6.1, ABE6.2, ABE6.3, ABE6.4, ABE6.5, ABE6.6, ABE7.1, ABE7.2, ABE7.3, ABE7.4, ABE7.5, ABE7.6, ABE7.7, ABE7.8, ABE 7.9, ABE7.10, ABE8.1-m, ABE8.2-m, ABE8.3-m, ABE8.4-m, ABE8.5-m, ABE8.6-m, ABE8.7-m, ABE8.8-m, ABE8.9-m, ABE8.10-m, ABE8.11-m, ABE8.12-m, ABE8.13-m, ABE8.14-m, ABE8.15-m, ABE8.16-m, ABE8.17-m, ABE8.18-m, ABE8.19-m, ABE8.20-m, ABE8.21-m, ABE8.22-m, ABE8.23-m, ABE8.24-m, ABE8.1-d, ABE8.2-d, ABE8.3-d, ABE8.4-d, ABE8.5-d, ABE8.6-d, ABE8.7-d, ABE8.8-d, ABE8.9-d, ABE8.10-d, ABE8.11-d, ABE8.12-d, ABE8.13-d, ABE8.14-d, ABE8.15-d, ABE8.16-d, ABE8.17-d, ABE8.18-d, ABE8.19-d, ABE8.20-d, ABE8.21-d, ABE8.22-d, ABE8.23-d, ABE8.24-d, ABE8a-m, ABE8b-m, ABE8c-m, ABE8d-m, ABE8e-m, ABE8a-d, ABE8b-d, ABE8c-d, ABE8d-d, ABE8e-d, ABE9.1, ABE9.2, ABE9.3, ABE9.4, ABE9.5, ABE9.6, ABE9.7, ABE9.8, ABE9.9, ABE9.10, ABE9.11, ABE9.12, ABE9.13, ABE9.14, ABE9.15, ABE9.16, ABE9.17, ABE9.18, ABE9.19, ABE9.2, ABE9.21, ABE9.22, ABE9.23, ABE9.24, ABE9.25, ABE9.26, ABE9.27, ABE9.28, ABE9.29, ABE9.30, ABE9.31, ABE9.32, ABE9.33, ABE9.34, ABE9.35, ABE9.36, ABE9.37, ABE9.38, ABE9.39, ABE9.40, ABE9.41, ABE9.42, ABE9.43, ABE9.44, ABE9.45, ABE9.46, ABE9.47, ABE9.48, ABE9.49, ABE9.50, ABE9.51, ABE9.52, ABE9.53, ABE9.54, ABE9.55, ABE9.56, ABE9.57, 및 ABE9.58.

[0067] 특정한 예시적 실시형태에서, 핵염기 편집 도메인은 다음으로 구성된 군으로부터 선택된 아데노신 염기 편집기를 포함한다: ABE 0.1, ABE 0.2, ABE 1.1, ABE 1.2, ABE2.1, ABE2.2, ABE2.3, ABE2.4, ABE2.5, ABE2.6, ABE2.7, ABE2.8, ABE2.9, ABE2.10, ABE2.11, ABE2.12, ABE3.1, ABE3.2, ABE3.3, ABE3.4, ABE3.5, ABE3.6, ABE3.7, ABE3.8, ABE4.1, ABE4.2, ABE4.3, ABE5.1, ABE5.2, ABE5.3, ABE5.4, ABE5.5, ABE5.6, ABE5.7, ABE5.8, ABE5.9, ABE5.10, ABE5.11, ABE5.12, ABE5.13, ABE5.14, ABE6.1, ABE6.2, ABE6.3, ABE6.4, ABE6.5, ABE6.6, ABE7.1, ABE7.2, ABE7.3, ABE7.4, ABE7.5, ABE7.6, ABE7.7, ABE7.8, ABE 7.9, ABE7.10, ABE8.1-m, ABE8.2-m, ABE8.3-m, ABE8.4-m, ABE8.5-m, ABE8.6-m, ABE8.7-m, ABE8.8-m, ABE8.9-m, ABE8.10-m, ABE8.11-m, ABE8.12-m, ABE8.13-m, ABE8.14-m, ABE8.15-m, ABE8.16-m, ABE8.17-m, ABE8.18-m, ABE8.19-m, ABE8.20-m, ABE8.21-m, ABE8.22-m, ABE8.23-m, ABE8.24-m, ABE8.1-d, ABE8.2-d, ABE8.3-d, ABE8.4-d, ABE8.5-d, ABE8.6-d, ABE8.7-d, ABE8.8-d, ABE8.9-d, ABE8.10-d, ABE8.11-d, ABE8.12-d, ABE8.13-d, ABE8.14-d, ABE8.15-d, ABE8.16-d, ABE8.17-d, ABE8.18-d, ABE8.19-d, ABE8.20-d, ABE8.21-d, ABE8.22-d, ABE8.23-d, ABE8.24-d, ABE8a-m, ABE8b-m, ABE8c-m, ABE8d-m, ABE8e-m, ABE8a-d, ABE8b-d, ABE8c-d, ABE8d-d, ABE8e-d, ABE9.1, ABE9.2, ABE9.3, ABE9.4, ABE9.5, ABE9.6, ABE9.7, ABE9.8, ABE9.9, ABE9.10, ABE9.11, ABE9.12, ABE9.13, ABE9.14, ABE9.15, ABE9.16, ABE9.17, ABE9.18, ABE9.19, ABE9.2, ABE9.21, ABE9.22, ABE9.23, ABE9.24, ABE9.25, ABE9.26, ABE9.27, ABE9.28, ABE9.29, ABE9.30, ABE9.31, ABE9.32, ABE9.33, ABE9.34, ABE9.35, ABE9.36, ABE9.37, ABE9.38, ABE9.39, ABE9.40, ABE9.41, ABE9.42, ABE9.43, ABE9.44, ABE9.45, ABE9.46, ABE9.47, ABE9.48, ABE9.49, ABE9.50, ABE9.51, ABE9.52, ABE9.53, ABE9.54, ABE9.55, ABE9.56, ABE9.57, 및 ABE9.58.

[0068] 특정한 예시적 실시형태에서, 핵염기 편집 도메인은 ABE7.10 또는 ABE8.20의 아데노신 염기 편집기를 포함한다.

[0069] 특정한 예시적 실시형태에서, 핵염기 편집 도메인은 시티딘 데아미나제이다. 특정한 예시적 실시형태에서, 시티딘 데아미나제는 *페트로미존 마리누스* 시토신 데아미나제 1(PmCDA1), 활성화-유도된 시티딘 데아미나제(AID) 및 APOBEC로 구성된 군으로부터 선택된다.

[0070] 특정한 예시적 실시형태에서, 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인은 Cas9 폴리펩티드, Cas12 폴리펩티드, 또는 이의 기능적 변이체를 포함한다.

[0071] 특정한 예시적 실시형태에서, 재조합 게놈은 가이드 RNA(gRNA)를 추가로 포함한다. 특정한 예시적 실시형태에서, 가이드 RNA(gRNA)는 표적 세포에 외인적으로 제공된다. 특정한 예시적 실시형태에서, gRNA는 표

적 세포의 게놈 좌위를 표적화할 수 있다.

[0072] 특정한 예시적 실시형태에서, 표적 세포는 *생체외에서* 형질도입된다. 특정한 예시적 실시형태에서, 표적 세포는 인간 세포이다. 특정한 예시적 실시형태에서, 표적 세포는 인간으로부터 획득된다. 특정한 예시적 실시형태에서, 표적 세포는 인간에 대해 자가조직이다. 특정한 예시적 실시형태에서, 표적 세포는 인간에 동종이계이다.

[0073] 특정한 예시적 실시형태에서, 표적 세포는 *생체내에서* 형질도입된다. 특정한 예시적 실시형태에서, 표적 세포는 인간 세포이다. 특정한 예시적 실시형태에서, 표적 세포는 뉴런 세포, 상피 세포 또는 간세포이다. 특정한 예시적 실시형태에서, 표적 세포는 인간에 있다.

[0074] 또 다른 양태에서, 대상체에게 본 명세서에 기재된 바와 같은 재조합 바이러스 입자 또는 본 명세서에 기재된 바와 같은 약학적 조성물을 투여하는 것을 포함하는, 대상체에게 치료적 이식유전자를 전달하는 방법이 제공된다.

[0075] 또 다른 양태에서, 대상체에게 재조합 광견병 바이러스 입자를 투여하는 것을 포함하는, 대상체에게 핵염기 편집기를 전달하는 방법이 제공되며, 여기서 재조합 바이러스 입자는 광견병 바이러스 당단백질; 및 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인 및 핵염기 편집 도메인을 포함하는 핵염기 편집기를 인코딩하는 핵산을 포함하는 재조합 광견병 바이러스 게놈을 포함하며, 여기서: 게놈은 광견병 바이러스 당단백질 또는 이의 기능적 변이체를 인코딩하는 G 유전자를 결하고/결하거나; 게놈은 광견병 바이러스 폴리머라제 또는 이의 기능적 변이체를 인코딩하는 L 유전자를 결한다.

[0076] 특정한 예시적 실시형태에서, 여기서 게놈은 광견병 바이러스 핵단백질 또는 이의 기능적 변이체를 인코딩하는 N 유전자; 광견병 바이러스 인단백질 또는 이의 기능적 변이체를 인코딩하는 P 유전자; 및 광견병 바이러스 매트릭스 단백질 또는 이의 기능적 변이체를 인코딩하는 M 유전자를 포함한다.

[0077] 특정한 예시적 실시형태에서, 여기서 핵염기 편집 도메인은 아데노신 데아미나제, 시티딘 데아미나제, 또는 이의 기능적 변이체이다. 특정한 예시적 실시형태에서, 염기 편집기는 아데노신 데아미나제이다. 특정한 예시적 실시형태에서, 아데노신 데아미나제는 ABE7.10이다.

[0078] 특정한 예시적 실시형태에서, 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인은 Cas9 폴리펩티드, Cas12 폴리펩티드, 또는 이의 기능적 변이체를 포함한다.

[0079] 특정한 예시적 실시형태에서, 재조합 게놈은 가이드 RNA(gRNA)를 인코딩하는 핵산 서열을 추가로 포함한다. 특정한 예시적 실시형태에서, 가이드 RNA(gRNA)는 표적 세포에 외인적으로 제공된다. 특정한 예시적 실시형태에서, gRNA는 표적 세포의 게놈 좌위를 표적화할 수 있다.

[0080] 특정한 예시적 실시형태에서, 대상체는 포유동물이다. 특정한 예시적 실시형태에서, 대상체는 인간이다.

[0081] 또 다른 양태에서, 광견병 바이러스 입자의 재조합 제조를 위한 패키징 시스템이 제공되며, 여기서 패키징 시스템은 광견병 바이러스 핵단백질 또는 이의 기능적 변이체를 인코딩하는 N 유전자; 광견병 바이러스 인단백질 또는 이의 기능적 변이체를 인코딩하는 P 유전자; 광견병 바이러스 폴리머라제 또는 이의 기능적 변이체를 인코딩하는 L 유전자; 및 재조합 광견병 바이러스 게놈을 포함하며, 여기서: 게놈은 광견병 바이러스 당단백질 또는 이의 기능적 변이체를 인코딩하는 G 유전자를 결하고/결하거나; 게놈은 광견병 바이러스 폴리머라제 또는 이의 기능적 변이체를 인코딩하는 L 유전자를 결여한다.

[0082] 특정한 예시적 실시형태에서, 재조합 광견병 바이러스 게놈은 이식유전자 또는 치료적 이식유전자를 인코딩하는 핵산을 포함한다.

[0083] 특정한 예시적 실시형태에서, 재조합 광견병 바이러스 게놈은 바이러스 게놈 벡터 내에 포함된다.

[0084] 특정한 예시적 실시형태에서, N, P 및 L 유전자는 각각 별도의 벡터 내에 포함된다. 특정한 예시적 실시형태에서, 각각의 N, P 및 L 유전자는 전사 조절 요소에 작동가능하게 연결된다. 특정한 예시적 실시형태에서, 전사 조절 요소는 프로모터 및/또는 인핸서를 포함한다. 특정한 예시적 실시형태에서, 프로모터는 구성적 프로모터이다. 특정한 예시적 실시형태에서, 프로모터는 신장 인자 1α 프로모터이다. 특정한 예시적 실시형태에서, 별도의 벡터는 각각 별도의 형질감염 플라스미드 내에 함유된다.

[0085] 특정한 예시적 실시형태에서, N, P 및 L 유전자는 단일 벡터 내에 포함된다. 특정한 예시적 실시형태에서, 단일 벡터는 N 및 P 유전자를 포함하는 제1 발현 카세트, 및 L 유전자를 포함하는 제2 발현 카세트를 포함한다. 특정

한 예시적 실시형태에서, 제1 발현 카세트는 5'에서 3'으로: 전사 조절 요소; P 유전자; 그리고 N 유전자를 포함한다. 특정한 예시적 실시형태에서, 제1 발현 카세트는 5'에서 3'으로: 전사 조절 요소; P 유전자; 리보솜 스키핑 요소; 그리고 N 유전자를 포함한다. 특정한 예시적 실시형태에서, 리보솜 스키핑 요소는 IRES 요소이다. 특정한 예시적 실시형태에서, 리보솜 스키핑 요소는 2A 요소이다. 특정한 예시적 실시형태에서, 제2 발현 카세트는 5'에서 3'으로: 전사 조절 요소; 그리고 L 유전자를 포함한다. 특정한 예시적 실시형태에서, 전사 조절 요소는 프로모터 및/또는 인핸서를 포함한다. 특정한 예시적 실시형태에서, 프로모터는 구성적 프로모터이다. 특정한 예시적 실시형태에서, 프로모터는 신장 인자 1α 프로모터이다. 특정한 예시적 실시형태에서, 제1 발현 카세트와 제2 발현 카세트는 벡터에서 반대 방향에 있다. 특정한 예시적 실시형태에서, 단일 벡터는 단일 형질감염 플라스미드 내에 함유된다.

[0086] 특정한 예시적 실시형태에서, 패키징 시스템은 광견병 바이러스 매트릭스 단백질 또는 이의 기능적 변이체를 인코딩하는 M 유전자를 추가로 포함한다. 특정한 예시적 실시형태에서, M 유전자는 벡터 내에 포함된다. 특정한 예시적 실시형태에서, M 유전자는 전사 조절 요소에 작동가능하게 연결된다. 특정한 예시적 실시형태에서, 전사 조절 요소는 프로모터 및/또는 인핸서를 포함한다. 특정한 예시적 실시형태에서, M 유전자를 포함하는 벡터는 형질감염 플라스미드 내에 함유된다.

[0087] 특정한 예시적 실시형태에서, 패키징 시스템은 광견병 바이러스 당단백질 또는 이의 기능적 변이체를 인코딩하는 G 유전자를 추가로 포함한다. 특정한 예시적 실시형태에서, G 유전자는 벡터 내에 포함된다. 특정한 예시적 실시형태에서, G 유전자는 전사 조절 요소에 작동가능하게 연결된다. 특정한 예시적 실시형태에서, 전사 조절 요소는 프로모터 및/또는 인핸서를 포함한다. 특정한 예시적 실시형태에서, G 유전자를 포함하는 벡터는 형질감염 플라스미드 내에 함유된다.

[0088] 또 다른 양태에서, 재조합 광견병 바이러스 입자를 생산하는 방법이 제공되며, 상기 방법은 재조합 광견병 바이러스 입자를 형성하기 위해 재조합 광견병 바이러스 게놈을 봉입하도록 작동하는 조건 하에서 본 명세서에 기재된 바와 같은 패키징 시스템을 세포 안으로 도입하는 것을 포함한다.

[0089] 특정한 예시적 실시형태에서, 도입하는 것은 전기천공, 뉴클레오펙션 또는 리포펙션에 의해 매개된다.

[0090] 다른 양태에서, 본 명세서에 기재된 바와 같은 패키징 시스템을 포함하는 재조합 광견병 바이러스 입자 패키징 세포가 제공된다.

[0091] 또 다른 양태에서, 광견병 바이러스 핵단백질 또는 이의 기능적 변이체를 인코딩하는 N 유전자; 광견병 바이러스 인단백질 또는 이의 기능적 변이체를 인코딩하는 P 유전자; 및 광견병 바이러스 폴리머라제 또는 이의 기능적 변이체를 인코딩하는 L 유전자를 포함하는 재조합 광견병 바이러스 입자 패키징 세포가 제공된다.

[0092] 특정한 예시적 실시형태에서, 제1 발현 카세트는 N 및 P 유전자를 포함한다. 특정한 예시적 실시형태에서, 제1 발현 카세트는 5'에서 3'으로: 전사 조절 요소; P 유전자; 그리고 N 유전자를 포함한다. 특정한 예시적 실시형태에서, 제1 발현 카세트는 5'에서 3'으로: 전사 조절 요소; P 유전자; 리보솜 스키핑 요소; 그리고 N 유전자를 포함한다. 특정한 예시적 실시형태에서, 리보솜 스키핑 요소는 IRES 요소이다. 특정한 예시적 실시형태에서, 리보솜 스키핑 요소는 2A 요소이다.

[0093] 특정한 예시적 실시형태에서, 제2 발현 카세트는 L 유전자를 포함한다. 특정한 예시적 실시형태에서, 제2 발현 카세트는 5'에서 3'으로: 전사 조절 요소; 그리고 L 유전자를 포함한다. 특정한 예시적 실시형태에서, 전사 조절 요소는 프로모터 및/또는 인핸서를 포함한다. 특정한 예시적 실시형태에서, 프로모터는 구성적 프로모터이다. 특정한 예시적 실시형태에서, 프로모터는 신장 인자 1α 프로모터이다.

[0094] 특정한 예시적 실시형태에서, 제1 및 제2 발현 카세트는 벡터에서 반대 방향에 있다.

[0095] 특정한 예시적 실시형태에서, 재조합 광견병 바이러스 입자 패키징 세포는 광견병 바이러스 매트릭스 단백질을 인코딩하는 M 유전자 및/또는 광견병 바이러스 당단백질을 인코딩하는 G 유전자를 추가로 포함한다. 특정한 예시적 실시형태에서, M 유전자 및/또는 G 유전자는 전사 조절 요소에 작동가능하게 연결된다. 특정한 예시적 실시형태에서, 전사 조절 요소는 프로모터 및/또는 인핸서를 포함한다. 특정한 예시적 실시형태에서, 제3 발현 및/또는 제4 카세트는 광견병 바이러스 G 유전자 및/또는 광견병 바이러스 M 유전자를 포함한다.

[0096] 특정한 예시적 실시형태에서, 패키징 세포는 포유동물, 박테리아 또는 곤충 기원의 것이다. 특정한 예시적 실시형태에서, 패키징 세포는 HEK293 세포, VERO 세포, BHK 세포 및 BSR 세포로 구성된 군으로부터 선택된다.

[0097] 또 다른 양태에서, 재조합 광견병 바이러스를 생산하는 방법이 제공되며, 상기 방법은 재조합 광견병 바이러스

를 형성하기 위해 당단백질에 재조합 광견병 바이러스 게놈을 봉입하기 위해 작동하는 조건 하에서 치료적 이식 유전자를 포함하는 핵산을 포함하는 재조합 광견병 바이러스 게놈을 포함하는 핵산을 본 명세서에 기술된 바와 같은 패키징 세포 안으로 도입하는 것을 포함하며, 여기서: 게놈은 광견병 바이러스 당단백질을 인코딩하는 내인성 G 유전자를 결하고; 게놈은 광견병 바이러스 폴리머라제를 인코딩하는 내인성 L 유전자를 결한다.

[0098] 특정한 예시적 실시형태에서, 재조합 게놈은 광견병 바이러스 핵단백질 또는 이의 기능적 변이체를 인코딩하는 N 유전자; 광견병 바이러스 인단백질 또는 이의 기능적 변이체를 인코딩하는 P 유전자; 및 광견병 바이러스 매트릭스 단백질 또는 이의 기능적 변이체를 인코딩하는 M 유전자를 포함한다.

[0099] 특정한 예시적 실시형태에서, 각각의 유전자는 전사 조절 요소에 작동가능하게 연결된다. 특정한 예시적 실시형태에서, 전사 조절 요소는 전사 개시 신호를 포함한다. 특정한 예시적 실시형태에서, 전사 개시 신호는 광견병 바이러스에 대해 외인성이다. 특정한 예시적 실시형태에서, 전사 개시 신호는 광견병 바이러스에 대해 내인성이다.

[0100] 특정한 예시적 실시형태에서, 각각의 유전자는 전사 종결 폴리아데닐화 신호에 작동가능하게 연결된다.

[0101] 특정한 예시적 실시형태에서, 재조합 광견병 바이러스 역가는 약 1E8 TU/mL보다 크다. 특정한 예시적 실시형태에서, 재조합 광견병 바이러스 역가는 약 1E8 TU/mL 내지 약 1E9 TU/mL이다.

[0102] 또 다른 양태에서, 대상체에서 질환 또는 장애를 치료하는 방법이 제공되며, 상기 방법은 본 명세서에 기재된 바와 같은 재조합 광견병 바이러스 입자 또는 본 명세서에 기재된 바와 같은 약학적 조성물을 대상체에게 투여하는 것을 포함한다.

[0103] 특정한 예시적 실시형태에서, 질환 또는 장애는 신경학적 질환 또는 장애이다. 특정한 예시적 실시형태에서, 질환 또는 장애는 안과 질환 또는 장애이다.

[0104] 또 다른 양태에서, 대상체에서 질환 또는 장애를 치료하기 위한 약제의 제조에 있어서, 본 명세서에 기술된 바와 같은 재조합 광견병 바이러스 또는 본 명세서에 기술된 바와 같은 약학적 조성물의 용도가 제공된다.

도면의 간단한 설명

[0105] **도 1**은 다양한 안정한 세포주로부터 지시된 날짜에 수확된 동등 부피의 바이러스-함유 상등액으로부터 293T 세포에 대한 상대적 감염성을 보여주는 차트이다.

도 2a는 VIR218 레플리콘을 도시하는 개략도이다.

도 2b는 재조합 광견병 바이러스 입자 매개된 유전자 전달을 위한 생산 및 감염 계획을 묘사하는 개략도이다.

도 2c는 핵염기 편집기를 인코딩하는 재조합 광견병 바이러스 게놈을 포함하는 재조합 광견병 바이러스 입자가 표적 서열의 유전자 편집에 영향을 미칠 수 있음을 묘사하는 차트이다.

도 3a는 gRNA, 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인 및 핵염기 편집기를 포함하는 재조합 광견병 바이러스 게놈의 조직화를 묘사하는 개략도이다.

도 3b는 RNA의 절단 부위를 나타내는 화살표로 2개의 tRNA 서열 사이의 gRNA를 인코딩하는 gRNA - tRNA 발현 카세트를 묘사하는 개략도이다.

도 3c는 gRNA(제1 gRNA 및 제2 gRNA)를 인코딩하는 gRNA - tRNA 발현 카세트를 묘사하는 개략도이며, 여기서 제1 gRNA는 제1 tRNA와 제2 tRNA 사이에 있고 제2 gRNA가 이어진다.

도 3d는 gRNA(제1 gRNA 및 제2 gRNA)를 인코딩하는 gRNA - tRNA 발현 카세트를 묘사하는 개략도이며, 여기서 제1 gRNA는 제1 tRNA와 제2 tRNA 사이에 있고, 제2 gRNA는 제2 tRNA와 제3 tRNA 사이에 있다.

도 3e는 핵염기 편집기를 인코딩하는 재조합 광견병 바이러스 게놈 및 다중 tRNA 사이에 인코딩된 gRNA를 포함하는 재조합 광견병 바이러스 입자로 형질도입된 HEK 세포에서 % 감염 및 % A>G 염기 편집을 묘사하는 차트이다. % 염기 편집은 Hek2-표적화 gRNA 및 IEDG-표적화 gRNA에 의해 표적화된 Hek2 부위 및 IEDG 부위에서 측정되었다.

도 4a는 Cre 재조합효소-2A-GFP 리포터 유전자를 인코딩하는 ΔG, ΔGL 및 ΔALL(즉, ΔGLNPM) 광견병 바이러스 레플리콘을 묘사하는 개략도이다.

도 4b는 ΔG, ΔGL 또는 ΔALL(즉, ΔGLNPM) 배경에서 광견병 바이러스 벡터로 형질감염된 세포의 형광 이미지를 묘사하며, 각 벡터는 Cre 재조합효소-2A-GFP 리포터 유전자를 인코딩한다. 이미지는 도 2b에 묘사된 바와 같이 표준 형질감염 및 최적화된 형질감염 하에서 형질감염-후 6일에 촬영되었다.

도 4c는 GFP 형광에 의해 측정된 바와 같은 리포터 세포 안으로 % 바이러스 도입을 나타내는 막대 그래프를 묘사한다.

도 5는 ABE 8.20 염기 편집기를 인코딩하는 RABV 게놈을 갖는 RABV로 형질도입된 293T 세포에서 A에서 G 염기 편집 퍼센트를 묘사하는 차트이다. SynV로 지칭되는 RABV SAD B19 ΔG 바이러스의 코돈-최적화된 버전이 RABV 게놈으로 사용되었다. 염기 편집기 유전자는 RABV 게놈의 G 유전자를 대체했다. HEK2 및 ABCA4 좌위를 표적화하는 U6-구동 가이드 RNA(gRNA)는 감염 시에 표적 세포 안으로 트랜스로 형질감염되었다. 기능적 바이러스 역가에 기반한 감염의 다중도(MOI)가 표시된다.

도 6은 여러 가지 다른 염기 편집기 중 하나를 인코딩하는 RABV 게놈을 갖는 RABV로 형질도입된 293T 세포에서 C에서 T 염기 편집 퍼센트를 묘사하는 2개의 차트를 묘사한다. SynV RABV 게놈이 사용되었다. 염기 편집기 유전자는 RABV 게놈에서 G 유전자를 대체했다. HEK2(상단 차트) 및 ABCA4(하단 차트) 좌위를 표적화하는 U6-구동 gRNA는 감염 시에 표적 세포 안으로 트랜스로 형질감염되었다. 기능적 바이러스 역가에 기반한 MOI가 표시된다.

도 7은 ABE7.10 염기를 인코딩하는 RABV 게놈을 갖는 RABV로 형질도입된 293T 세포에서 A에서 G 염기 편집 퍼센트를 묘사하는 2개의 차트를 묘사한다. SAD B19 ΔGL RABV 게놈이 염기 편집기를 인코딩하는데 사용되었으며, 여기서 염기 편집기는 RABV 게놈 내의 G 단백질을 대체했다. 한 세트의 세포가 RABV L 단백질(바이러스 + L)을 인코딩하는 플라스미드에 의해 추가로 형질감염되었다. HEK2 및 ABCA4 좌위를 표적화하는 U6-구동 gRNA는 감염 시에 표적 세포 안으로 트랜스로 형질감염되었다. 기능적 바이러스 역가에 기반한 MOI가 표시된다.

도 8은 ΔN, ΔP, ΔM 및 ΔL 배경에서 광견병 바이러스 벡터로 형질감염된 세포의 형광 이미지를 묘사하며, 각 벡터는 Cre 재조합효소-2A-GFP 리포터 유전자를 인코딩한다. 누락된 유전자의 보완 여부에 관계없이 293T 세포 또는 CE1.30 세포가 형질감염되었다. 이미지는 감염-후 4일에 촬영되었다.

도 9는 GFP 형광에 의해 측정된 바와 같은 리포터 세포(293T 또는 CE1.30) 안으로 % 바이러스 도입을 나타내는 막대 그래프를 묘사한다. **도 8**에서 형질감염된 세포로부터의 상등액이 사용되었다.

도 10은 GFP 및 mScarlet 형광에 의해 측정된, CE1.30 세포 안으로 % 바이러스 도입을 나타내는 그래프를 묘사한다. Cre-2a-GFP를 이식유전자로 인코딩하는 ΔGL, ΔMGL, ΔPMGL 및 ΔALL 광견병 바이러스 레플리콘은 CE1.30 세포에서 성장되었고 Cre와 GFP mRNA 둘 모두의 기능적 전달을 나타내도록 리포터 세포에서 역가측정되었다.

도 11은 ΔGL, ΔMGL, ΔPMGL 및 ΔALL 배경에서 광견병 바이러스 벡터로 형질감염된 세포의 형광 이미지를 묘사하며, 각 벡터는 Cre 재조합효소-2A-GFP 리포터 유전자를 인코딩한다. 293T 세포 또는 CE1.30 세포가 형질감염되었다. 이미지는 감염-후 3일에 촬영되었다.

도 12는 RABV 유전자를 발현하는데 사용되는 발현 벡터의 도식을 묘사한다. 벡터 VIR069는 EF1α 프로모터, RABV N 유전자, 2a 리보솜 스키핑 요소, RABV P 유전자, 2a 리보솜 스키핑 요소, 그리고 RABV M 유전자의 5'에서 3'까지의 제1 발현 카세트를 함유한다. VIR069는 RPBSA 프로모터 그리고 BFP-제오신 선택 마커 유전자의 5'에서 3'까지의 제1 발현 카세트에 역방향으로 제2 발현 카세트를 함유한다. 벡터 VIR071은 EF1α 프로모터, RABV N 유전자, IRES, RABV P 유전자, 2a 리보솜 스키핑 요소, 그리고 RABV M 유전자의 5'에서 3'까지의 제1 발현 카세트를 함유한다. VIR071은 RPBSA 프로모터, RABV L 유전자, 2a 리보솜 스키핑 요소, 그리고 제오신 선택 마커 유전자의 5'에서 3'까지의 제1 발현 카세트에 역방향으로 제2 발현 카세트를 함유한다. 벡터 VIR112는 EF1α 프로모터, RABV N 유전자, IRES, RABV P 유전자, IRES, 그리고 RABV M 유전자의 5'에서 3'까지의 제1 발현 카세트를 함유한다. VIR071은 RPBSA 프로모터, RABV L 유전자, 2a 리보솜 스키핑 요소, 그리고 제오신 선택 마커 유전자의 5'에서 3'까지의 제1 발현 카세트에 역방향으로 제2 발현 카세트를 함유한다.

발명을 실시하기 위한 구체적인 내용

[0106] 이식유전자(예를 들어, 치료적 이식유전자)를 포함하는 핵산을 포함하는 재조합 광견병 바이러스 게놈이 본 명세서에 제공된다. 특정 실시형태에서, 재조합 광견병 바이러스 게놈은 광견병 바이러스 당단백질을 인코딩하는 내인성 G 유전자를 결하고, 광견병 바이러스 폴리머라제를 인코딩하는 내인성 L 유전자를 결한다. 또한 재조합 광견병 바이러스 입자를 생산하는 방법이 제공된다. 이러한 방법은 일반적으로 본 명세서에 기재된 패키징 시스

템을 적합한 패키징 세포 안으로 도입하는 것을 포함한다.

[0107] 본 명세서에 기재된 방법은 본 명세서에 개시된 특정 방법 및 실험 조건에 제한되지 않고 그러한 방법 및 조건은 다양할 수 있다는 것이 이해되어야 한다. 또한 본 명세서에 사용된 용어는 단지 특정 실시형태를 기술하기 위한 것이고 제한하려는 의도가 아니라는 것도 이해되어야 한다. 본 명세서에 기재된 방법은 당업자의 기술 내에 있는 통상적인 분자 및 세포의 생물학적 및 면역학적 기술을 사용한다. 이러한 기술은 숙련된 기술자에게 잘 알려져 있고 과학 문헌에 설명되어 있다.

[0108] **A.정의**

[0109] 달리 정의되지 않는 한, 본 명세서에 사용된 모든 기술 및 과학 용어는 본 발명이 속하는 기술분야의 숙련가에 의해 일반적으로 이해되는 의미를 갖는다. 다음 참고문헌은 본 발명에 사용된 다수의 용어에 대한 일반적인 정의를 당업자에게 제공한다: Singleton 등, Dictionary of Microbiology and Molecular Biology (2nd ed. 1994); The Cambridge Dictionary of Science and Technology (Walker ed., 1988); The Glossary of Genetics, 5th Ed., R. Rieger 등 (eds.), Springer Verlag (1991); 및 Hale & Marham, The Harper Collins Dictionary of Biology (1991). 본 명세서에서 사용된 바와 같은 다음 용어는 달리 명시하지 않는 한 아래에 부여된 의미를 갖는다.

[0110] "아데노신 데아미나제"는 아데닌 또는 아데노신의 가수분해적 탈아미노화를 촉매할 수 있는 폴리펩티드 또는 이의 단편을 의미한다. 일부 실시형태에서, 데아미나제 또는 데아미나제 도메인은 아데노신에서 이노신으로 또는 데옥시 아데노신에서 데옥시이노신으로의 가수분해적 탈아미노화를 촉매하는 아데노신 데아미나제이다. 일부 실시형태에서, 아데노신 데아미나제는 데옥시리보핵산(DNA)에서 아데닌 또는 아데노신의 가수분해적 탈아미노화를 촉매한다. 본 명세서에서 제공되는 아데노신 데아미나제(예를 들어 조작된 아데노신 데아미나제, 진화된 아데노신 데아미나제)는 박테리아와 같은 임의의 유기체로부터의 것일 수 있다.

[0111] "아데노신 데아미나제 염기 편집기 8(ABE8) 폴리펩티드" 또는 "ABE8"은 다음 참조 서열의 아미노산 위치 82 및/또는 166에서 변경을 포함하는 아데노신 데아미나제 변이체를 포함하는 본 명세서에 정의된 염기 편집기를 의미한다:
MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRV
VFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTD(서열번호: 8).

[0112] 일부 실시형태에서, ABE8은 참조 서열에 비해 본 명세서에 기재된 바와 같은 추가 변경을 포함한다.

[0113] "아데노신 데아미나제 염기 편집기 8(ABE8) 폴리뉴클레오티드"는 ABE8을 인코딩하는 폴리뉴클레오티드를 의미한다.

[0114] "투여하는 것"은 본 명세서에 기재된 하나 이상의 조성물을 환자 또는 대상체에게 제공하는 것으로 본 명세서에서 지칭된다.

[0115] "작용제"는 임의의 소분자 화학적 화합물, 항체, 핵산 분자, 또는 폴리펩티드 또는 이의 단편을 지칭한다.

[0116] "변경"은 본 명세서에 기재된 것과 같은 표준 기술 공지 방법에 의해 검출되는 분석물, 유전자 또는 폴리펩티드의 수준, 구조 또는 활성에서의 변화(증가 또는 감소)를 의미한다. 본 명세서에 사용된 바와 같이, 변경은 발현 수준의 10% 변화, 25% 변화, 40% 변화, 및 발현 수준의 50% 이상의 변화를 포함한다. 일부 실시형태에서, 변경은 핵염기 또는 아미노산의 삽입, 결실 또는 치환을 포함한다.

[0117] "개선하다"는 질환의 발달 또는 진행을 감소, 억제, 약화, 감쇠, 정지 또는 안정화시키는 것을 의미한다.

[0118] "유사체"는 동일하지 않지만 유사한 기능적 또는 구조적 특징을 갖는 분자를 의미한다. 예를 들어, 폴리펩티드 유사체는 상응하는 자연적으로-발생하는 폴리펩티드의 생물학적 활성을 유지하면서 자연적으로 발생하는 폴리펩티드에 비해 유사체의 기능을 향상시키는 특정 생화학적 변형을 갖는다. 이러한 생화학적 변형은 예를 들어 리간드 결합을 변경하지 않고도 유사체의 단백질분해효소 저항성, 막 투과성 또는 반감기를 증가시킬 수 있다. 유사체는 비천연 아미노산을 포함할 수 있다.

[0119] "염기 편집기(BE)" 또는 "핵염기 편집기 폴리펩티드(NBE)"는 폴리뉴클레오티드에 결합하고 핵염기 변형 활성을 갖는 작용제를 의미한다. 다양한 실시형태에서, 염기 편집기는 가이드 폴리뉴클레오티드(*예를 들어,* 가이드 RNA(gRNA))와 연계한 핵염기 변형 폴리펩티드(*예를 들어,* 데아미나제) 및 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인(*예를 들어,* Cas9 또는 Cpf1)을 포함한다. 염기 편집기의 대표적인 핵산 및 단백질 서열은

서열 목록에 서열번호: 274-283으로 제공된다.

[0120] "염기 편집 활성"은 폴리뉴클레오티드 내의 염기를 화학적으로 변경시키는 작용을 의미한다. 일 실시형태에서, 제1 염기는 제2 염기로 전환된다. 일 실시형태에서, 염기 편집 활성은 시티딘 데아미나제 활성, *예를 들어* 표적 C·G를 T·A로 전환시키는 활성이다. 다른 실시형태에서, 염기 편집 활성은 아데노신 또는 아데닌 데아미나제 활성, *예를 들어* A·T를 G·C로 전환하는 활성이다.

[0121] 용어 "염기 편집기 시스템"은 표적 뉴클레오티드 서열의 핵염기를 편집하기 위한 분자간 복합체를 지칭한다. 다양한 실시형태에서, 염기 편집기(BE) 시스템은 (1) 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인, 표적 뉴클레오티드 서열에서 핵염기를 탈아미노화하기 위한 데아미나제 도메인(*예를 들어*, 시티딘 데아미나제 또는 아데노신 데아미나제); 및 (2) 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인과 연계한 하나 이상의 가이드 폴리뉴클레오티드(*예를 들어*, 가이드 RNA)를 포함한다. 다양한 실시형태에서, 염기 편집기(BE) 시스템은 아데노신 데아미나제 또는 시티딘 데아미나제로부터 선택된 핵염기 편집기 도메인, 및 핵산 서열 특이적 결합 활성을 갖는 도메인을 포함한다. 일부 실시형태에서, 염기 편집기 시스템은 (1) 표적 뉴클레오티드 서열에서 하나 이상의 핵염기를 탈아미노화하기 위한 폴리뉴클레오티드 프로그램가능 DNA 결합 도메인 및 데아미나제 도메인을 포함하는 염기 편집기(BE); 및 (2) 폴리뉴클레오티드 프로그램가능 DNA 결합 도메인과 연계한 하나 이상의 가이드 RNA를 포함한다. 일부 실시형태에서, 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인은 폴리뉴클레오티드 프로그램가능 DNA 결합 도메인이다. 일부 실시형태에서, 염기 편집기는 시티딘 염기 편집기(CBE)이다. 일부 실시형태에서, 염기 편집기는 아데닌 또는 아데노신 염기 편집기(ABE)이다. 일부 실시형태에서, 염기 편집기는 아데닌 또는 아데노신 염기 편집기(ABE) 또는 시티딘 염기 편집기(CBE)이다.

[0122] "염기 편집 활성"은 폴리뉴클레오티드 내의 염기를 화학적으로 변경시키는 작용을 의미한다. 일 실시형태에서, 제1 염기는 제2 염기로 전환된다. 일 실시형태에서, 염기 편집 활성은 시티딘 데아미나제 활성, 예를 들어 표적 C·G를 T·A로 전환시키는 활성이다. 다른 실시형태에서, 염기 편집 활성은 아데노신 데아미나제 활성, 예를 들어 A·T를 G·C로 전환시키는 활성이다.

[0123] 용어 "Cas9" 또는 "Cas9 도메인"은 Cas9 단백질 또는 이의 단편(*예를 들어*, Cas9의 활성, 비활성 또는 부분적으로 활성 DNA 절단 도메인 및/또는 Cas9의 gRNA 결합 도메인을 포함하는 단백질)을 포함하는 RNA 가이드된 뉴클레아제를 지칭한다. Cas9 뉴클레아제는 또한 때때로 casnl 뉴클레아제 또는 CRISPR(클러스터링된 규칙적으로 간격을 둔 짧은 회문 반복) 연관된 뉴클레아제로도 지칭된다.

[0124] 용어 "보존적 아미노산 치환" 또는 "보존적 돌연변이"는 공통 특성을 갖는 다른 아미노산에 의한 하나의 아미노산의 대체를 지칭한다. 개별 아미노산 사이의 공통 특성을 정의하는 기능적 방법은 상동 유기체의 상응하는 단백질 사이의 아미노산 변화의 표준화된 빈도를 분석하는 것이다(Schulz, G. E. 및 Schirmer, R. H., Principles of Protein Structure, Springer-Verlag, New York (1979)). 그러한 분석에 따르면, 그룹 내의 아미노산이 서로 우선적으로 교환되고 따라서 전반적인 단백질 구조에 미치는 그 영향이 가장 서로 유사한 아미노산의 그룹이 정의될 수 있다(Schulz, G.E. 및 Schirmer, R.H., 상기 참조). 보존적 돌연변이의 비-제한적 예는 아미노산의 아미노산 치환, 예를 들어 양전하가 유지될 수 있도록 아르기닌을 리신으로 그리고 그 반대로; 음전하가 유지될 수 있도록 아스파르트산을 글루탐산으로 그리고 그 반대로; 유리 -OH가 유지될 수 있도록 트레오닌을 세린으로; 및 유리 $-NH_2$가 유지될 수 있도록 아스파라긴을 글루타민으로의 치환을 포함한다.

[0125] 본 명세서에서 상호교환적으로 사용된 바와 같은 용어 "코딩 서열" 또는 "단백질 코딩 서열"은 단백질을 코딩하는 폴리뉴클레오티드의 세그먼트를 지칭한다. 코딩 서열은 개방 판독 프레임으로도 지칭될 수 있다. 영역 또는 서열은 시작 코돈에 의해 5' 말단에 더 근접하고 정지 코돈으로 3' 말단에 더 근접하여 결합된다. 본 명세서에 기술된 염기 편집기에 유용한 정지 코돈은 다음을 포함한다:

[0126] 글루타민 **C**AG → **T**AG 정지 코돈

[0127] **C**AA → **T**AA

[0128] 아르기닌 CGA → **T**GA

[0129] 트립토판 TG**G** → TG**A**

[0130] **T**GG → **T**AG

[0131] **T**GG → **TA**A

[0132] "시티딘 데아미나제"는 아미노기를 카르보닐기로 전환시키는 탈아미노화 반응을 촉매할 수 있는 폴리펩티드 또는 이의 단편을 의미한다. 일 실시형태에서, 시티딘 데아미나제는 시토신을 우라실로 또는 5-메틸시토신을 티민으로 전환시킨다. *페트로미존 마리누스*(*페트로미존 마리누스* 시토신 데아미나제 1, "PmCDA1")로부터 유래된 PmCDA1(서열번호: 41-42), 포유동물(*예를 들어*, 인간, 돼지, 소, 말, 원숭이 등)로부터 유래된 AID(활성화-유도된 시티딘 데아미나제; AICDA)(예시적인 AID 폴리펩티드 서열은 서열 목록에 서열번호: 43-44, 1372 및 1374-1377로 제공되어 있음) 및 APOBEC는 예시적인 시티딘 데아미나제이다(예시적인 APOBEC 폴리펩티드 서열은 서열 목록에 서열번호: 1378-1416, 1421, 및 1422로 제공된다. 추가의 예시적인 시티딘 데아미나제(CDA) 서열은 서열 목록에 서열번호: 1373, 1417-1420으로 제공된다. APOBEC 폴리펩티드 서열을 포함하는, 추가의 예시적인 시티딘 데아미나제 서열은 서열 목록에 서열번호 1378-1422로 제공된다. 개시내용의 일부 양태는 핵산 분자(예를 들어, DNA 또는 RNA)에서 시토신을 탈아미노화할 수 있는 염기 편집기 단백질 및 염기 편집기 시스템을 제공한다. 일부 실시형태에서 염기 편집기 단백질 또는 염기 편집기 시스템은 핵산 분자(예를 들어, DNA 또는 RNA)에서 시티딘을 우리딘으로 탈아미노화할 수 있는 시티딘 데아미나제 또는 시토신 데아미나제를 포함한다. 본 명세서에 제공된 임의의 시티딘 또는 시토신 데아미나제는 자연적으로-발생하는 시티딘 및 시토신 데아미나제의 변이체를 포함한다는 것을 이해해야 한다. 이러한 변이체는 핵산 분자(예를 들어, DNA 또는 RNA)에서 염기 편집기 단백질 또는 시스템의 표적-상 데아미나제 활성의 효율성을 증가시키도록 조작될 수 있다.

[0133] 본 명세서에 사용된 바와 같은 용어 "데아미나제" 또는 "데아미나제 도메인"은 탈아미노화 반응을 촉매하는 단백질 또는 효소를 지칭한다.

[0134] "검출"은 검출되어 지는 분석물의 존재, 부재 또는 양을 식별하는 것을 지칭한다. 일 실시형태에서, 폴리뉴클레오티드 또는 폴리펩티드에서 서열 변경이 검출된다. 다른 실시형태에서, 삽입결실의 존재가 검출된다.

[0135] "검출가능한 표지"는 관심있는 분자에 연결될 때 분광학, 광화학적, 생화학적, 면역화학적 또는 화학적 수단을 통해 후자를 검출가능하게 만드는 조성물을 의미한다. 예를 들어, 유용한 표지는 방사성 동위원소, 자성 비드, 금속 비드, 콜로이드 입자, 형광 염료, 전자-밀집 시약, 효소(예를 들어, 효소 연결된 면역흡착 검정(ELISA)에서 일반적으로 사용되는 것), 비오틴, 디곡시게닌 또는 합텐을 포함한다.

[0136] "질환"은 세포, 조직 또는 기관의 정상적인 기능을 손상시키거나 방해하는 임의의 병태 또는 장애를 의미한다. 예시적인 질환은 신경계 질환 및 안과 질환을 포함한다.

[0137] "유효량"은 치료되지 않은 환자 또는 질환이 없는 개체, 즉 건강한 개체에 비해 질환의 증상을 개선하는데 필요한 작용제 또는 활성 화합물, *예를 들어*, 본 명세서에 기재된 바와 같은 염기 편집기의 양을 의미하거나, 또는 원하는 생물학적 반응을 이끌어내기에 충분한 작용제 또는 활성 화합물의 양이다. 질환의 치료적 치료를 위해 본 발명을 실시하는데 사용되는 활성 화합물(들)의 유효량은 투여의 방식, 대상체의 연령, 체중 및 전반적인 건강에 따라 다양하다. 궁극적으로, 주치의나 수의사가 적절한 양과 복용량 섭생을 결정할 것이다. 이러한 양이 "유효한" 양으로 지칭된다. 일 실시형태에서, 유효량은 세포(*예를 들어*, *시험관내* 또는 *생체내* 세포)에서 관심있는 유전자에서의 변경을 도입하기에 충분한 발명의 염기 편집기의 양이다. 일 실시형태에서, 유효량은 치료적 효과를 달성하는데 필요한 염기 편집기의 양이다. 그러한 치료적 효과는 대상체, 조직 또는 기관의 모든 세포에서 병원성 유전자를 변경하기에 충분할 필요는 없지만, 대상체, 조직 또는 기관에 존재하는 세포의 약 1%, 5%, 10%, 25%, 50%, 75% 또는 75% 이상에서 병원성 유전자를 변경하기에 충분하면 된다. 일 실시형태에서, 유효량은 질환의 하나 이상의 증상을 개선하는데 충분하다.

[0138] 용어 "엑소뉴클레아제"는 자유 말단으로부터 핵산(*예를 들어*, RNA 또는 DNA)을 소화할 수 있는 단백질 또는 폴리펩티드를 지칭한다.

[0139] 용어 "엔도뉴클레아제"는 핵산(*예를 들어*, DNA 또는 RNA)에서의 내부 영역을 촉매작용(*예를 들어*, 절단)할 수 있는 단백질 또는 폴리펩티드를 지칭한다.

[0140] "단편"은 폴리펩티드 또는 핵산 분자의 부분을 의미한다. 이 부분은 참조 핵산 분자 또는 폴리펩티드의 전체 길이의 적어도 약 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% 또는 90%를 함유한다. 단편은 10, 20, 30, 40, 50, 60, 70, 80, 90 또는 100, 200, 300, 400, 500, 600, 700, 800, 900 또는 1000개의 뉴클레오티드 또는 아미노산을 함유할 수 있다.

[0141] "가이드 RNA" 또는 "gRNA"는 표적 서열에 특이적이고 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인 단백질(*예를 들어*, Cas9 또는 Cpf1)과 복합체를 형성할 수 있는 폴리뉴클레오티드 또는 폴리뉴클레오티드 복합체를 의미한다. 실시형태에서, 가이드 폴리뉴클레오티드는 가이드 RNA(gRNA)이다. gRNA는 2개 이상의 RNA의 복

합체로 존재할 수 있거나 단일 RNA 분자로 존재할 수 있다.

[0142] "tRNA" 또는 "전달 RNA"는 세포에서 절단될 수 있는 2차 및/또는 3차 구조를 포함하는 RNA 분자를 의미한다. 절단은 예를 들어, RNase P 또는 RNase Z와 같은 RNase를 통해 발생할 수 있다.

[0143] "혼성화"는 상보성인 핵염기 사이의, 왓슨-크릭, 후그스틴 또는 역 후그스틴 수소 결합일 수 있는, 수소 결합을 의미한다. 예를 들어, 아데닌과 티민은 수소 결합 형성을 통해 쌍을 이루는 상보성인 핵염기이다.

[0144] "증가"는 적어도 10%, 25%, 50%, 75% 또는 100%의 긍정적인 변경을 의미한다.

[0145] 용어 "염기 수선의 억제제", "염기 수선 억제제", "IBR" 또는 이의 문법적 동등어는 핵산 수선 효소, 예를 들어 염기 절제 수선 효소의 활성을 억제할 수 있는 단백질을 지칭한다.

[0146] 용어 "단리된", "정제된" 또는 "생물학적으로 순수한"은 그 원래 상태에서 발견되는 대로 일반적으로 수반되는 구성요소가 다양한 정도로 없는 물질을 지칭한다. "단리하다"는 원래 공급원 또는 주변환경으로부터 분리의 정도를 나타낸다. "정제하다"는 단리보다 높은 분리의 정도를 나타낸다. "정제된" 또는 "생물학적으로 순수한" 단백질은 임의의 불순물이 단백질의 생물학적 특성에 실질적으로 영향을 미치거나 다른 부정적인 결과를 야기하지 않도록 다른 물질이 충분하게 없다. 즉, 본 발명의 핵산 또는 펩티드는 재조합 DNA 기술에 의해 생산될 때 세포 물질, 바이러스 물질 또는 배양 배지가 실질적으로 없거나 화학적으로 합성될 때 화학적 전구체 또는 기타 화학 물질이 실질적으로 없는 경우 정제된다. 순도와 균질성은 전형적으로 분석 화학 기술, 예를 들어, 폴리아크릴아미드 겔 전기영동 또는 고성능 액체 크로마토그래피를 사용하여 결정된다. 용어 "정제된"은 핵산 또는 단백질이 전기영동 겔에서 본질적으로 하나의 밴드를 초래한다는 것을 나타낼 수 있다. 변형, 예를 들어, 인산화 또는 글리코실화가 적용될 수 있는 단백질의 경우, 다양한 변형으로 인해 별도로 정제될 수 있는 다양한 단리된 단백질을 초래할 수 있다.

[0147] "단리된 폴리뉴클레오티드"는 발명의 핵산 분자가 유래된 유기체의 자연적으로-발생하는 게놈에서 유전자에 측접하는 유전자가 없는 핵산(*예를 들어,* DNA)을 의미한다. 따라서 용어는, 예를 들어, 벡터; 자율적으로 복제되는 플라스미드 또는 바이러스; 또는 원핵생물 또는 진핵생물의 게놈 DNA 안으로 통합된; 또는 다른 서열과 독립적으로 별도의 분자(예를 들어, cDNA 또는 PCR 또는 제한 엔도뉴클레아제 소화에 의해 생성된 게놈 또는 cDNA 단편)으로 존재하는 재조합 DNA를 포함한다. 부가하여, 용어는 DNA 분자로부터 전사되는 RNA 분자뿐만 아니라 추가 폴리펩티드 서열을 인코딩하는 하이브리드 유전자의 일부인 재조합 DNA를 포함한다.

[0148] "단리된 폴리펩티드"는 자연적으로 수반하는 구성요소로부터 분리된 발명의 폴리펩티드를 의미한다. 전형적으로, 폴리펩티드는 그것이 자연적으로 연관되어 있는 단백질 및 자연적으로-발생하는 유기 분자가 중량 기준으로 적어도 60% 없을 때 단리된다. 바람직하게는, 제제는 중량 기준으로 적어도 75%, 더욱 바람직하게는 적어도 90%, 가장 바람직하게는 적어도 99%가 발명의 폴리펩티드이다. 발명의 단리된 폴리펩티드는, 예를 들어 천연 공급원으로부터의 추출, 그러한 폴리펩티드를 인코딩하는 재조합 핵산의 발현; 또는 단백질을 화학적으로 합성함에 의해 획득될 수 있다. 순도는 임의의 적절한 방법, 예를 들어 컬럼 크로마토그래피, 폴리아크릴아미드 겔 전기영동 또는 HPLC 분석에 의해 측정될 수 있다.

[0149] 본 명세서에 사용된 바와 같은 용어 "돌연변이"는 서열, *예를 들어* 핵산 또는 아미노산 서열 내의 잔기의 또 다른 잔기로의 치환, 또는 서열 내의 하나 이상의 잔기의 결실 또는 삽입을 지칭한다. 돌연변이는 전형적으로 원래의 잔기를 식별하고 이어서 서열 내의 잔기의 위치를 식별하고 새로이 치환된 잔기의 식별에 의해 본 명세서에 기술된다. 본 명세서에 제공된 아미노산 치환(돌연변이)을 만드는 다양한 방법이 당업계에 잘 알려져 있고, 예를 들어 Green 및 Sambrook, *Molecular Cloning: A Laboratory Manual* (4th ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (2012))에 의해 제공된다.

[0150] 본 명세서에 사용된 바와 같은 용어 "핵산" 및 "핵산 분자"는 핵염기와 산성 모이어티, 예를 들어 뉴클레오시드, 뉴클레오티드 또는 뉴클레오티드의 중합체를 포함하는 화합물을 지칭한다. 전형적으로, 중합체성 핵산, 예를 들어 3개 이상의 뉴클레오티드를 포함하는 핵산 분자는 선형 분자이며, 여기서 인접한 뉴클레오티드는 포스포디에스테르 연결을 통해 서로 연결된다. 일부 실시형태에서, "핵산"은 개별 핵산 잔기(*예를 들어* 뉴클레오티드 및/또는 뉴클레오시드)를 지칭한다. 일부 실시형태에서, "핵산"은 3개 이상의 개별 뉴클레오티드 잔기를 포함하는 올리고뉴클레오티드 사슬을 지칭한다. 본 명세서에 사용된', 용어 "올리고뉴클레오티드" 및 "폴리뉴클레오티드"는 뉴클레오티드의 중합체(예를 들어, 적어도 3개 뉴클레오티드의 스트링)를 지칭하기 위해 상호교환적으로 사용될 수 있다. 일부 실시형태에서, "핵산"은 단일 및/또는 이중-가닥 DNA뿐만 아니라 RNA를 포괄한다. 핵산은 예를 들어 게놈, 전사체, mRNA, tRNA, rRNA, siRNA, snRNA, 플라스미드, 코스미드, 염색체, 염색

분체 또는 기타 자연적으로 발생하는 핵산 분자의 맥락에서 자연적으로 발생할 수 있다. 반면, 핵산 분자는 비-자연적으로 발생하는 분자, 예를 들어 재조합 DNA 또는 RNA, 인공 염색체, 조작된 게놈 또는 이의 단편, 또는 합성 DNA, RNA, DNA/RNA 하이브리드, 또는 비-자연적으로 발생하는 뉴클레오티드 또는 뉴클레오시드를 포함하는 것일 수 있다. 더욱이, 용어 "핵산", "DNA", "RNA" 및/또는 유사한 용어는 핵산 유사체, *예를 들어,* 포스포디에스테르 골격 이외의 다른 유사체를 포함한다. 핵산은 천연 공급원으로부터 정제될 수 있고, 재조합 발현 시스템을 사용하여 생성되고 선택적으로 정제되고, 화학적으로 합성 등이 될 수 있다. 적절한 경우, 예를 들어 화학적으로 합성된 분자의 경우, 핵산은 화학적으로 변형된 염기 또는 당 및 골격 변형을 갖는 유사체와 같은 뉴클레오시드 유사체를 포함할 수 있다. 달리 명시되지 않는 한, 핵산 서열은 5'에서 3' 방향으로 제시된다. 일부 실시형태에서, 핵산은 천연 뉴클레오시드(*예를 들어,* 아데노신, 티미딘, 구아노신, 시티딘, 우리딘, 데옥시아데노신, 데옥시티미딘, 데옥시구아노신 및 데옥시시티딘); 뉴클레오시드 유사체(예를 들어, 2-아미노아데노신, 2-티오티미딘, 이노신, 피롤로-피리미딘, 3-메틸 아데노신, 5-메틸시티딘, 2-아미노아데노신, C5-브로모우리딘, C5-플루오로우리딘, C5-요오도우리딘, C5-프로피닐-우리딘, C5-프로피닐-시티딘, C5-메틸시티딘, 2-아미노아데노신, 7-데아자아데노신, 7-데아자구아노신, 8-옥소아데노신, 8-옥소구아노신, O(6)-메틸구아닌 및 2-티오시티딘); 화학적으로 변형된 염기; 생물학적으로 변형된 염기(예를 들어, 메틸화된 염기); 개재된 염기; 변형된 당(2'-예를 들어, 플루오로리보스, 리보스, 2'-데옥시리보스, 아라비노스 및 헥소스); 및/또는 변형된 포스페이트 기(예를 들어, 포스포로티오에이트 및 5'-*N*-포스포르아미다이트 연결)이거나 이를 포함한다.

[0151] 용어 "핵 국소화 서열", "핵 국소화 신호" 또는 "NLS"는 단백질이 세포 핵 안으로 유입되는 것을 촉진하는 아미노산 서열을 지칭한다. 핵 국소화 서열은 당업계에 공지되어 있고, 예를 들어 2001년 5월 31일에 WO/2001/038547로 공개된, Plank 등의 2000년 11월 23일 출원된 국제 PCT 출원인 PCT/EP2000/011690에 기재되어 있으며, 그 내용은 예시적인 핵 국소화 서열의 개시내용에 대해 참조로 본 명세서에 포함된다. 다른 실시형태에서, NLS는 예를 들어 Koblan 등, Nature Biotech. 2018 doi:10.1038/nbt.4172에 개시된 최적화된 NLS이다. 일부 실시형태에서, NLS는 다음 아미노산 서열 KRTADGSEFESPKKKRKV(서열번호: 84), KRPAATKKAGQAKKKK(서열번호: 85), KKTELQTTNAENKTKKL(서열번호: 86), KRGINDRNFWRGENGRKTR(서열번호: 87), RKSGKIAAIVVKRPRK(서열번호: 88), PKKKRKV(서열번호: 89), 또는 MDSLLMNRRKFLYQFKNVRWAKGRRETYLC(서열번호: 90)를 포함한다.

[0152] 본 명세서에서 상호교환적으로 사용되는 용어 "핵염기", "질소성 염기" 또는 "염기"는 뉴클레오시드를 형성하는 질소-함유 생물학적 화합물을 지칭하며, 이는 차례로 뉴클레오티드의 구성요소이다. 염기쌍을 형성하고 서로 쌓이는 핵염기의 능력은 리보핵산(RNA) 및 디옥시리보핵산(DNA)과 같은 장쇄 나선형 구조로 직접적으로 이어진다. 5개의 핵염기 - 아데닌(A), 시토신(C), 구아닌(G), 티민(T), 우라실(U) -은 일차 또는 규범적인 것으로 불린다. 아데닌과 구아닌은 퓨린 및 시토신에서 유래되고, 우라실 및 티민은 피리미딘에서 유래된다. DNA와 RNA는 또한 변형된 다른(비-일차) 염기를 함유할 수도 있다. 비-제한적인 예시적인 변형된 핵염기는 하이포크산틴, 크산틴, 7-메틸구아닌, 5,6-디하이드로우라실, 5-메틸시토신(m5C) 및 5-하이드로메틸시토신을 포함할 수 있다. 하이포크산틴과 크산틴은 돌연변이 유발물질 존재를 통해 생성될 수 있으며 둘 모두 탈아미노화(아민기의 카르보닐기로의 대체)를 통해 생성될 수 있다. 하이포크산틴은 아데닌에서 변형될 수 있다. 크산틴은 구아닌에서 변형될 수 있다. 우라실은 시토신의 탈아미노화로부터 발생할 수 있다. "뉴클레오시드"는 핵염기와 5탄당(리보스 또는 디옥시리보스)으로 구성된다. 뉴클레오시드의 예는 아데노신, 구아노신, 우리딘, 시티딘, 5-메틸우리딘(m5U), 데옥시아데노신, 데옥시구아노신, 티미딘, 데옥시우리딘 및 데옥시시티딘을 포함한다. 변형된 핵염기를 갖는 뉴클레오시드의 예는 이노신(I), 크산토신(X), 7-메틸구아노신(m7G), 디하이드로우리딘(D), 5-메틸시티딘(m5C) 및 슈도우리딘(Ψ)을 포함한다. "뉴클레오티드"는 핵염기, 5탄당(리보스 또는 디옥시리보스) 및 적어도 하나의 포스페이트기로 구성된다.

[0153] 본 명세서에 사용된 바와 같은 용어 "핵산" 및 "핵산 분자"는 핵염기와 산성 모이어티, *예를 들어*, 뉴클레오시드, 뉴클레오티드 또는 뉴클레오티드의 중합체를 포함하는 화합물을 지칭한다.

[0154] 본 명세서에 사용된 바와 같은 용어 "올리고뉴클레오티드" 및 "폴리뉴클레오티드"는 뉴클레오티드의 중합체를 지칭하기 위해 상호교환적으로 사용될 수 있다.

[0155] 용어 "핵산 프로그램가능 DNA 결합 단백질" 또는 "napDNAbp"는 napDNAbp를 특정 핵산 서열로 안내하는 핵산(*예를 들어*, DNA 또는 RNA), 예컨대 가이드 핵산 또는 가이드 폴리뉴클레오티드(*예를 들어*, gRNA)와 회합하는 단백질을 지칭하기 위해 "폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인"과 상호교환적으로 사용될 수 있다. 일부 실시형태에서, 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인은 폴리뉴클레오티드 프로그램가능 DNA 결합 도메인이다. 일부 실시형태에서, 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인

은 폴리뉴클레오티드 프로그램가능 RNA 결합 도메인이다. 일부 실시형태에서, 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인은 Cas9 단백질이다. Cas9 단백질은 Cas9 단백질을 가이드 RNA에 상보성인 특정 DNA 서열로 안내하는 가이드 RNA와 회합할 수 있다. 일부 실시형태에서, napDNAbp는 Cas9 도메인, 예를 들어 뉴클레아제 활성 Cas9, Cas9 닉카제(nCas9) 또는 뉴클레아제 불활성 Cas9(dCas9)이다. 핵산 프로그램가능 DNA 결합 단백질의 비-제한적인 예는 Cas9(*예를 들어*, dCas9 및 nCas9), Cas12a/Cpf1, Cas12b/C2c1, Cas12c/C2c3, Cas12d/CasY, Cas12e/CasX, Cas12g, Cas12h, Cas12i, 및 Cas12j/CasΦ(Cas12j/Casphi)를 포함한다. Cas 효소의 비-제한적인 예는 Cas1, Cas1B, Cas2, Cas3, Cas4, Cas5, Cas5d, Cas5t, Cas5h, Cas5a, Cas6, Cas7, Cas8, Cas8a, Cas8b, Cas8c, Cas9(Csn1 또는 Csx12로도 알려짐), Cas10, Cas10d, Cas12a/Cpf1, Cas12b/C2c1, Cas12c/C2c3, Cas12d/CasY, Cas12e/CasX, Cas12g, Cas12h, Cas12i, Cas12j/CasΦ, Cpf1, Csy1, Csy2, Csy3, Csy4, Cse1, Cse2, Cse3, Cse4, Cse5e, Csc1, Csc2, Csa5, Csn1, Csn2, Csm1, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Csb1, Csb2, Csb3, Csx17, Csx14, Csx10, Csx16, CsaX, Csx3, Csx1, Csx1S, Csx11, Csf1, Csf2, CsO, Csf4, Csd1, Csd2, Cst1, Cst2, Csh1, Csh2, Csa1, Csa2, Csa3, Csa4, Csa5, 유형 II Cas 효과기 단백질, 유형 V Cas 효과기 단백질, 유형 VI Cas 효과기 단백질, CARF, DinG, 이의 동족체, 또는 이의 변형 또는 조작된 버전을 포함한다. 다른 핵산 프로그램가능 DNA 결합 단백질도 본 개시내용의 범주 내에 있지만, 본 개시내용에 구체적으로 나열되지는 않을 수도 있다. *예를 들어,* Makarova 등 "Classification and Nomenclature of CRISPR-Cas Systems: Where from Here?" *CRISPR J.* 2018 Oct;1:325-336. doi: 10.1089/crispr.2018.0033; Yan 등, "Functionally diverse type V CRISPR-Cas systems" *Science*. 2019 Jan 4;363(6422):88-91. doi: 10.1126/science.aav7271을 참조하며, 각각의 전체 내용은 본 명세서에 참조로 포함된다. 예시적인 핵산 프로그램가능 DNA 결합 단백질 및 핵산 프로그램가능 DNA 결합 단백질을 인코딩하는 핵산 서열은 서열 목록에 서열번호: 223, 230-232, 235-242, 246-256 및 285-294로 제공된다.

[0156] 본 명세서에 사용된 바와 같은 용어 "핵염기 편집 도메인" 또는 "핵염기 편집 단백질"은, 시토신(또는 시티딘)에서 우라실(또는 우리딘) 또는 티민(또는 티미딘), 아데닌(또는 아데노신)에서 하이포크산틴(또는 이노신)으로 탈아미노화뿐만 아니라 비-주형 뉴클레오티드 부가 및 삽입과 같은, RNA 또는 DNA에서 핵염기 변형을 촉매할 수 있는 단백질 또는 효소를 지칭한다. 일부 실시형태에서, 핵염기 편집 도메인은 데아미나제 도메인(예를 들어, 아데닌 데아미나제 또는 아데노신 데아미나제; 또는 시티딘 데아미나제 또는 시토신 데아미나제)이다.

[0157] 본 명세서에 사용된 바와 같이, "작용제를 획득하는"에서와 같이 "획득하는"은 작용제를 합성하거나, 구매하거나, 달리 수득하는 것을 포함한다.

[0158] 본 명세서에 사용된 바와 같은 "환자" 또는 "대상체"는 질환 또는 장애를 갖는 것으로 진단되거나, 전개하거나 전개할 위험이 있거나, 전개하거나 전개할 것으로 의심되는 포유동물 대상체 또는 개체를 지칭한다. 일부 실시형태에서, 용어 "환자"는 질환 또는 장애를 전개할 가능성이 평균보다 높은 포유동물 대상체를 지칭한다. 예시적인 환자는 본 명세서에 개시된 요법으로부터 이익을 얻을 수 있는 인간, 비-인간 영장류, 고양이, 개, 돼지, 소, 고양이, 말, 낙타, 라마, 염소, 양, 설치류(*예를 들어*, 마우스, 토끼, 랫트, 또는 기니피그) 및 기타 포유동물일 수 있다. 예시적인 인간 환자는 남성 및/또는 여성일 수 있다.

[0159] "이를 필요로 하는 환자" 또는 "이를 필요로 하는 대상체"는 본 명세서에서 질환 또는 장애를 갖는 것으로 진단되거나, 가질 위험이 있거나, 갖는 것으로 예정되거나, 갖는 것으로 의심되는 환자로 지칭된다.

[0160] 용어 "병원성 돌연변이", "병원성 변이체", "질환 케이싱 돌연변이", "질환 야기 변이체", "유해한 돌연변이" 또는 "소인 돌연변이"는 특정 질환이나 장애에 대한 개체의 감수성 또는 소인을 증가시키는 유전적 변형 또는 돌연변이를 지칭한다. 일부 실시형태에서, 병원성 돌연변이는 유전자에 의해 인코딩되는 단백질에서 적어도 하나의 병원성 아미노산에 의해 치환된 적어도 하나의 야생형 아미노산을 포함한다.

[0161] 용어 "단백질", "펩티드", "폴리펩티드" 및 이의 문법적 동등어는 본 명세서에서 상호교환적으로 사용되고, 펩티드(아미드) 결합에 의해 함께 연결된 아미노산 잔기의 중합체를 지칭한다. 단백질, 펩티드 또는 폴리펩티드는 자연적으로 발생하는, 재조합, 또는 합성, 또는 이의 조합일 수 있다.

[0162] 본 명세서에 사용된 바와 같은 용어 "융합 단백질"은 적어도 2개의 서로 다른 단백질로부터의 단백질 도메인을 포함하는 하이브리드 폴리펩티드를 지칭한다.

[0163] 단백질 또는 핵산과 관련하여 본 명세서에서 사용된 바와 같은 용어 "재조합"은 자연에서는 발생하지 않지만 인간 조작의 생성물인 단백질 또는 핵산을 지칭한다. 예를 들어, 일부 실시형태에서, 재조합 단백질 또는 핵산 분자는 임의의 자연적으로 발생하는 서열과 비교하여 적어도 1개, 적어도 2개, 적어도 3개, 적어도 4개, 적어도 5

개, 적어도 6개 또는 적어도 7개의 돌연변이를 포함하는 아미노산 또는 뉴클레오티드 서열을 포함한다.

[0164] "감소한다"는 적어도 10%, 25%, 50%, 75% 또는 100%의 음성적 변화를 의미한다.

[0165] "참조 서열"은 서열 비교를 위한 기반으로 사용되는 정의된 서열이다. 참조 서열은 특정화된 서열; 예를 들어, 전체-길이의 cDNA 또는 유전자 서열의 세그먼트, 또는 완전한 cDNA 또는 유전자 서열의 서브세트 또는 전체일 수 있다. 폴리펩티드의 경우, 참조 폴리펩티드 서열의 길이는 일반적으로 적어도 약 16개 아미노산, 적어도 약 20개 아미노산, 적어도 약 25개 아미노산, 약 35개 아미노산, 약 50개 아미노산, 또는 약 100개 아미노산일 것이다. 핵산의 경우, 참조 핵산 서열의 길이는 일반적으로 적어도 약 50개 뉴클레오티드, 적어도 약 60개 뉴클레오티드, 적어도 약 75개 뉴클레오티드, 약 100개 뉴클레오티드 또는 약 300개 뉴클레오티드 또는 그 주위 또는 그 사이의 임의의 정수일 것이다. 일부 실시형태에서, 참조 서열은 관심있는 단백질의 야생형 서열이다. 다른 실시형태에서, 참조 서열은 야생형 단백질을 인코딩하는 폴리뉴클레오티드 서열이다.

[0166] 용어 "RNA-프로그램가능 뉴클레아제" 및 "RNA-가이드된 뉴클레아제"는 절단에 대한 표적이 아닌 하나 이상의 RNA(들)와 함께 사용된다(*예를 들어*, 이와 결합하거나 회합한다). 일부 실시형태에서, RNA-프로그램가능 뉴클레아제는 RNA와 복합체일 때 뉴클레아제:RNA 복합체로 지칭될 수 있다. 전형적으로 결합된 RNA(들)는 가이드 RNA(gRNA)로 지칭된다. 일부 실시형태에서, RNA-프로그램가능 뉴클레아제는 (CRISPR-연관된 시스템) Cas9 엔도뉴클레아제, 예를 들어, *스트렙토코커스 피오게네스*로부터의 Cas9(Csn1)이다.

[0167] 용어 "단일 뉴클레오티드 다형성(SNP)"은 게놈에서의 특정 위치에서 발생하는 단일 뉴클레오티드의 변이이며, 여기서 각 변이는 모집단 내에서 어느 정도 인지가능한 정도(*예를 들어*, >1%)로 존재한다.

[0168] "특이적으로 결합한다"는 발명의 폴리펩티드 및/또는 핵산 분자를 인식하고 결합하지만, 샘플, 예를 들어 생물학적 샘플에서 다른 분자는 실질적으로 인식하고 결합하지 않는 핵산 분자, 폴리펩티드, 폴리펩티드/폴리뉴클레오티드 복합체, 화합물 또는 분자를 의미한다.

[0169] "실질적으로 동일한"은 참조 아미노산 서열과 적어도 50% 동일성을 나타내는 폴리펩티드 또는 핵산 분자를 의미한다. 일 실시형태에서, 참조 서열은 야생형 아미노산 또는 핵산 서열이다. 다른 실시형태에서, 참조 서열은 본 명세서에 기재된 아미노산 또는 핵산 서열 중 임의의 하나이다. 일 실시형태에서, 이러한 서열은 비교를 위해 사용된 서열과 아미노산 수준 또는 핵산 수준에서 적어도 60%, 80%, 85%, 90%, 95% 또는 심지어 99% 동일하다.

[0170] 서열 동일성은 전형적으로 서열 분석 소프트웨어(예를 들어, Genetics Computer Group, University of Wisconsin Biotechnology Center, 1710 University Avenue, Madison, Wis. 53705의 서열 분석 소프트웨어 패키지인, BLAST, BESTFIT, GAP 또는 PILEUP/PRETTYBOX 프로그램)를 사용하여 측정된다. 이러한 소프트웨어는 다양한 치환, 결실 및/또는 기타 변형에 대한 상동성의 정도를 할당함에 의해 동일하거나 유사한 서열을 일치시킨다. 보존적 치환은 전형적으로 다음 그룹 내의 치환을 포함한다: 글리신, 알라닌; 발린, 이소류신, 류신; 아스파르트산, 글루탐산, 아스파라긴, 글루타민; 세린, 트레오닌; 리신, 아르기닌; 및 페닐알라닌, 티로신. 동일성의 정도를 결정하기 위한 예시적인 접근법에서, e^{-3}과 e^{-100} 사이의 확률 점수가 밀접하게 관련된 서열을 나타내는 BLAST 프로그램이 사용될 수 있다.

[0171] 예를 들어, 다음 매개변수를 갖는 COBALT가 사용된다:

[0172] a) 정렬 매개변수: 갭 페널티-11,-1 및 엔드-갭 페널티-5,-1,

[0173] b) CDD 매개변수: RPS BLAST 온 사용; Blast E-값 0.003; 보존된 열을 찾고 다시 계산함, 및

[0174] c) 쿼리 클러스터링 매개변수: 쿼리 클러스터 온 사용; 단어 크기 4; 최대 클러스터 거리 0.8; 알파벳 레귤러.

[0175] 예를 들어, 다음 매개변수를 갖는 EMBOSS 니들이 사용된다:

[0176] a) 매트릭스: BLOSUM62;

[0177] b) 갭 오픈: 10;

[0178] c) 갭 연장: 0.5;

[0179] d) 출력 형식: 쌍;

[0180] e) 엔드 갭 페널티: 가성;

[0181] f) 엔드 갭 오픈: 10; 및

[0182] g) 엔드 갭 연장: 0.5.

[0183] 발명의 방법에 유용한 핵산 분자는 발명의 폴리펩티드 또는 이의 단편을 인코딩하는 임의의 핵산 분자를 포함한다. 이러한 핵산 분자는 내인성 핵산 서열과 100% 동일할 필요는 없지만 전형적으로 실질적인 동일성을 나타낼 것이다. 내인성 서열에 대해 "실질적인 동일성"을 갖는 폴리뉴클레오티드는 전형적으로 이중-가닥 핵산 분자의 적어도 하나의 가닥과 혼성화할 수 있다. 발명의 방법에 유용한 핵산 분자는 발명의 폴리펩티드 또는 이의 단편을 인코딩하는 임의의 핵산 분자를 포함한다. 이러한 핵산 분자는 내인성 핵산 서열과 100% 동일할 필요는 없지만 전형적으로 실질적인 동일성을 나타낼 것이다. 내인성 서열에 대해 "실질적인 동일성"을 갖는 폴리뉴클레오티드는 전형적으로 이중-가닥 핵산 분자의 적어도 하나의 가닥과 혼성화할 수 있다. "혼성화한다"는 다양한 엄격 조건 하에서 상보성인 폴리뉴클레오티드 서열(*예를 들어*, 본 명세서에 기재된 유전자) 또는 이의 부분 사이에 이중-가닥 분자를 형성하는 쌍을 의미한다. (*예를 들어*, Wahl, G. M. 및 S. L. Berger (1987) Methods Enzymol. 152:399; Kimmel, A. R. (1987) Methods Enzymol. 152:507 참조).

[0184] 예를 들어, 엄격한 염 농도는 통상적으로 약 750mM NaCl 및 75mM 구연산삼나트륨 미만, 바람직하게는 약 500mM NaCl 및 50mM 구연산삼나트륨 미만, 보다 바람직하게는 약 250mM NaCl 및 25mM 구연산삼나트륨 미만일 것이다. 낮은 엄격도 혼성화는 유기 용매, *예를 들어* 포름아미드의 부재에서 얻어질 수 있는 반면, 높은 엄격도 혼성화는 적어도 약 35% 포름아미드, 보다 바람직하게는 적어도 약 50% 포름아미드의 존재에서 얻어질 수 있다. 엄격한 온도 조건은 통상적으로 적어도 약 30℃, 보다 바람직하게는 적어도 약 37℃, 가장 바람직하게는 적어도 약 42℃의 온도를 포함할 것이다. 혼성화 시간, 세제, *예를 들어*, 나트륨 도데실 설페이트(SDS)의 농도, 및 담체 DNA의 함입 또는 배제와 같은 다양한 추가 매개변수는 당업자에게 잘 알려져 있다. 필요에 따라 이러한 다양한 조건을 조합하여 다양한 수준의 엄격도를 달성한다. 바람직한: 실시형태에서, 혼성화는 750mM NaCl, 75mM 구연산삼나트륨 및 1% SDS 중에서 30℃에서 일어날 것이다. 보다 바람직한 실시형태에서, 혼성화는 500mM NaCl, 50mM 구연산삼나트륨, 1% SDS, 35% 포름아미드 및 100㎍/ml 변성된 연어 정자 DNA(ssDNA) 중에서 37℃에서 일어날 것이다. 가장 바람직한 실시형태에서, 혼성화는 250mM NaCl, 25mM 구연산삼나트륨, 1% SDS, 50% 포름아미드 및 200㎍/ml ssDNA 중에서 42℃에서 일어날 것이다. 이들 조건에 대한 유용한 변이는 당업자에게 쉽게 명백할 것이다.

[0185] 대부분의 응용분야에서 혼성화에 이어지는 세정 단계도 엄격하게 다를 것이다. 세정 엄격도 조건은 염 농도와 온도에 의해 정의될 수 있다. 상기와 같이 염 농도를 줄이거나 온도를 증가시킴에 의해 세정 엄격도가 증가될 수 있다. 예를 들어, 세정 단계를 위한 엄격한 염 농도는 바람직하게는 약 30mM NaCl 및 3mM 구연산삼나트륨 미만, 가장 바람직하게는 약 15mM NaCl 및 1.5mM 구연산삼나트륨 미만일 것이다. 세정 단계를 위한 엄격한 온도 조건은 통상적으로 적어도 약 25℃, 더 바람직하게는 적어도 약 42℃, 더욱 더 바람직하게는 적어도 약 68℃의 온도를 포함할 것이다. 실시형태에서, 세정 단계는 30mM NaCl, 3mM 구연산삼나트륨 및 0.1% SDS 중에서 25℃에서 일어날 것이다. 다른 실시형태에서, 세정 단계는 15mM NaCl, 1.5mM 구연산삼나트륨 및 0.1% SDS 중에서 42℃에서 일어날 것이다. 더욱 바람직한 실시형태에서, 세정 단계는 15mM NaCl, 1.5mM 구연산삼나트륨 및 0.1% SDS 중에서 68℃에서 일어날 것이다. 이들 조건에 대한 추가적인 변이는 당업자에게 쉽게 명백할 것이다. 혼성화 기술은 당업자에게 잘 알려져 있고, 예를 들어 Benton 및 Davis (Science 196:180, 1977); Grunstein 및 Hogness (Proc. Natl. Acad. Sci., USA 72:3961, 1975); Ausubel 등 (Current Protocols in Molecular Biology, Wiley Interscience, New York, 2001); Berger 및 Kimmel (Guide to Molecular Cloning Techniques, 1987, Academic Press, New York); 및 Sambrook 등, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, New York에 기술되어 있다.

[0186] "분할"은 둘 이상의 단편으로 나누어지는 것을 의미한다.

[0187] "분할 Cas9 단백질" 또는 "분할 Cas9"는 2개의 별도의 뉴클레오티드 서열에 의해 인코딩되는 N-말단 단편 및 C-말단 단편으로 제공되는 Cas9 단백질을 지칭한다. Cas9 단백질의 N-말단 부분 및 C-말단 부분에 상응하는 폴리펩티드는 스플라이싱되어 "재구성된" Cas9 단백질을 형성할 수 있다.

[0188] 용어 "표적 부위"는 데아미나제(*예를 들어*, 시티딘 또는 아데닌 데아미나제) 또는 데아미나제를 포함하는 융합 단백질(*예를 들어*, 본 명세서에 개시된 dCas9-아데노신 데아미나제 융합 단백질 또는 염기 편집기)에 의해 탈아미노화되는 핵산 분자 내의 서열을 지칭한다.

[0189] 본 명세서에 사용된 바와 같은 용어 "치료하다", 치료하는, "치료" 등은 장애 및/또는 이와 연관된 증상을 감소 또는 개선하거나 원하는 약리학적 및/또는 생리학적 효과를 얻는 것을 지칭한다. 배제되지는 않지만, 장애 또는 병태를 치료하는 것이 그와 연관된 장애, 병태 또는 증상을 완전히 제거할 것을 요구하지 않는다는 것이 이해될

것이다. 일부 실시형태에서, 효과는 치료적이며, 즉 제한 없이, 효과는 질환 및/또는 질환에 기인할 수 있는 유해 증상을 부분적으로 또는 완전히 감소, 감쇄, 폐기, 경감, 완화, 강도 감소 또는 치유한다. 일부 실시형태에서, 효과는 예방적이며, 즉 효과는 질환 또는 병태의 발생 또는 재발을 보호하거나 예방한다. 이를 위해, 현재 개시된 방법은 본 명세서에 기재된 조성물의 치료적으로 유효한 양을 투여하는 것을 포함한다.

[0190] "우라실 글리코실라제 억제제" 또는 "UGI"는 우라실-절제 수선 시스템을 억제하는 작용제를 의미한다. 시티딘 데아미나제를 포함하는 염기 편집기는 시토신을 우라실로 전환한 다음 DNA 복제 또는 수선을 통해 티민으로 전환한다. 염기 편집기에 우라실 DNA 글리코실라제의 억제제(UGI)를 포함하는 것은 U를 다시 C로 변경하는 염기 절단 수선을 방지한다. 예시적인 UGI는 다음과 같은 아미노산 서열을 포함한다:

[0191] >sp|P14739|UNGI_BPPB2 우라실-DNA 글리코실라제 억제제

[0192] MTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMLLTSD APEYKPWALVIQDSNGENKIKML(서열번호: 106).

[0193] 본 문서에 제공된 범주는 그 범위 내의 모든 값에 대해 줄여서 표현된 것으로 이해된다. 예를 들어, 1 내지 50의 범위는 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 또는 50으로 구성된 군으로부터 임의의 수, 수의 조합, 또는 하위-범위를 포함하는 것으로 이해된다.

[0194] 본 명세서의 변수의 임의의 정의에서 화학기의 목록의 인용은 그 변수의 정의를 나열된 기의 임의의 단일 기 또는 조합으로 포함한다. 본 명세서의 변수 또는 양태에 대한 실시형태의 언급은 임의의 단일 실시형태로서 또는 임의의 다른 실시형태 또는 그 부분과 조합한 그 실시형태를 포함한다.

[0195] 모든 용어는 당업자가 이해하는 대로 이해되도록 의도되었다. 다르게 정의되지 않는 한, 본 명세서에 사용된 모든 기술 및 과학 용어는 개시내용이 속하는 기술 분야의 통상인에 의해 일반적으로 이해되는 것과 동일한 의미를 갖는다.

[0196] 본 출원에서, 단수형의 사용은 달리 구체적으로 언급되지 않는 한 복수형을 포함한다. 명세서에서 사용된 바와 같은 단수형 "a", "an" 및 "the"는 문맥상 명백하게 다르게 지시하지 않는 한 복수형을 포함한다는 점에 유의해야 한다. 본 출원에서 "또는"의 사용은 달리 언급되지 않는 한 "및/또는"을 의미한다. 더욱이, 용어 "포함하는"뿐만 아니라 "포함하다", "포함한다", "포함된"과 같은 다른 형태의 사용은 제한되지 않는다.

[0197] 본 명세서 및 청구항(들)에 사용된 바와 같이, 단어 "포함하는" (및 "포함하다" 및 "포함한다"와 같은, 포함하는의 임의의 형태), "갖는" (및 "가지다" 및 "갖는다"와 같은, 갖는의 임의의 모든 형태), "포괄하는" (및 "포괄하다" 및 "포괄한다"와 같은, 포괄하는의 임의의 형태), 또는 "함유하는" (및 "함유하다" 및 "함유한다"와 같은, 함유하는의 임의의 형태)은 포괄적이거나 공개-목적이고, 추가의 인용되지 않은 요소나 방법 단계를 제외하지 않는다. 본 명세서에서 논의된 임의의 실시형태는 본 개시내용의 임의의 방법 또는 조성물에 관해 구현될 수 있고, 그 반대도 가능하다는 것이 고려된다. 더욱이, 본 개시내용의 조성물은 본 개시내용의 방법을 달성하는데 사용될 수 있다.

[0198] 용어 "약" 또는 "대략적으로"는 당업자에 의해 결정된 바와 같은 특정 값에 대해 허용가능한 오차 범위 내를 의미하며, 이는 부분적으로 값이 측정되거나 결정되는 방식, 즉 측정 시스템의 한계에 부분적으로 의존할 것이다. 예를 들어, "약"은 당업계의 실시에 따라 1 또는 1 초과의 표준 편차 이내를 의미할 수 있다. 대안적으로, "약"은 주어진 값의 최대 20%, 최대 10%, 최대 5% 또는 최대 1%의 범위를 의미할 수 있다. 대안적으로, 특히 생물학적 시스템 또는 과정과 관련하여, 용어는 값의 크기의 정도 이내, *예를 들어*, 5-배 이내, 2-배 이내를 의미할 수 있다. 출원 및 청구범위에 특정 값이 기술되어 있는 경우, 달리 언급되지 않는 한, 용어 "약"은 특정 값에 대해 허용가능한 오류 범위 내에 있는 것으로 가정되어야 함을 의미한다.

[0199] 명세서에서 "특정 실시형태", "일부 실시형태", "실시형태", "일 실시형태", 또는 "다른 실시형태"에 대한 언급은 실시형태와 연계하여 기술된 특정 특징, 구조 또는 특성이 본 개시내용의, 반드시 모든 실시형태는 아니지만, 적어도 일부 실시형태에 포함된다는 것을 의미한다.

[0200] **B. <u>재조합 광견병 바이러스</u>**

[0201] 표적 세포를 형질도입하는데 유용한 재조합 광견병 바이러스가 본 명세서에 제공된다. 일 양태에서, 본 개시내용의 재조합 광견병 바이러스는 광견병 바이러스 당단백질 및 재조합 광견병 바이러스 게놈을 포함한다. 특정 실시형태에서, 재조합 광견병 바이러스 게놈은 이식유전자를 포함하는 핵산을 인코딩한다. 특정 실시형태에서, 재조합 광견병 바이러스 게놈은 치료적 이식유전자를 포함하는 핵산을 인코딩한다. 이와 같이, 본 개시내용의

재조합 광견병 바이러스는 표적 세포를 형질도입하는 방법에 이용될 수 있으며, 여기서 재조합 광견병 바이러스는 광견병 바이러스 당단백질 및 이식유전자(예를 들어, 치료적 이식유전자)를 포함하는 핵산을 포함하는 재조합 광견병 바이러스 게놈을 포함한다. 표적 세포의 형질도입 시, 재조합 광견병 바이러스 게놈 내에 포함된 이식유전자가 발현되고 이식유전자 생성물이 생산된다.

[0202] *광견병 리사바이러스*로도 알려진 광견병 바이러스는 *랍도비리대* 계통의 *리사바이러스* 속의 음성 센스 단일 가닥 RNA 바이러스이다. 광견병 바이러스는 원통형 형태를 갖고 있으며 구조는 나선형 리보핵단백질 코어를 둘러싸는 당단백질 G로 구성된 지단백질 외피를 포함한다. 광견병 바이러스 게놈은 게놈의 전사 및 복제를 촉진하는 단백질과 바이러스의 구조적 구성요소를 구성하는 단백질을 인코딩하는 5개의 유전자를 함유한다. 5개의 유전자는: 광견병 바이러스 핵단백질을 인코딩하는 N 유전자; 광견병 바이러스 인단백질을 인코딩하는 P 유전자; 광견병 바이러스 매트릭스 단백질을 인코딩하는 M 유전자; 광견병 바이러스 당단백질을 인코딩하는 G 유전자; 및 광견병 바이러스 폴리머라제를 인코딩하는 L 유전자이다. 바이러스 게놈 RNA와 핵단백질은 함께 광견병 바이러스 폴리머라제(RNA-의존성 RNA 폴리머라제)에 의한 복제 및 전사를 위한 주형으로 기능하는 리보핵단백질을 형성한다.

[0203] 특정 실시형태에서, 본 개시내용의 재조합 광견병 바이러스 게놈은 하나 이상의 광견병 바이러스 유전자가 제거되어 있다. 예를 들어, N 유전자, P 유전자, M 유전자, L 유전자 및/또는 G 유전자는 재조합 광견병 바이러스 게놈에 없을 수 있다. 특정 실시형태에서, 재조합 광견병 바이러스 게놈은 광견병 바이러스 당단백질 또는 이의 기능적 변이체를 인코딩하는 G 유전자를 결한다. 광견병 바이러스 당단백질을 인코딩하는 G 유전자를 결하는 재조합 광견병 바이러스 게놈은 바이러스가 내인적으로 당단백질을 생성할 수 있는 것을 방지한다. 당단백질은 바이러스 생활 주기의 마지막 단계에서만 필요하기 때문에 이 결실은 바이러스가 처음으로 감염된 세포를 넘어 퍼지는 것을 방지하지만 바이러스가 해당 시점까지 그 복제 주기의 전체를 완료하는 것을 방지하지는 않는다. 특정 실시형태에서, 재조합 광견병 바이러스 게놈은 광견병 바이러스 폴리머라제 또는 이의 기능적 변이체를 인코딩하는 L 유전자를 결한다. L 유전자 생성물은 바이러스 유전자의 전사와 바이러스 게놈의 복제 둘 모두에 필요하고, L 유전자의 결실은 표적 형질도입된 세포의 낮은 세포독성을 초래할 수 있다. 예를 들어, Chatterjee 등, *Nat. Neurosci.* (2018) 21(4): 638-646을 참조하며, 그 개시내용은 그 전체가 본 명세서에 참조로 포함된다. 특정 실시형태에서, 재조합 광견병 바이러스 게놈은 광견병 바이러스 당단백질 또는 이의 기능적 변이체를 인코딩하는 G 유전자를 결하고, 광견병 바이러스 폴리머라제 또는 이의 기능적 변이체를 인코딩하는 L 유전자를 결한다.

[0204] 본 명세서에 기재된 바와 같은 광견병 바이러스 유전자를 결하는 재조합 광견병 바이러스 게놈은 광견병 바이러스 유전자의 전부 또는 일부를 결하는 광견병 바이러스 게놈을 지칭한다는 것은 당업자에 의해 쉽게 인식된다. 예를 들어, G 유전자를 결하는 재조합 광견병 바이러스 게놈은 G 유전자의 전부 또는 일부를 결할 수 있으며, 여기서 G 유전자의 일부는 G 유전자 생성물의 기능에 필요하다. 특정 실시형태에서, G 유전자 생성물의 기능에 필요한 G 유전자의 일부가 결여되면 절삭된 비-기능성 당단백질의 생산을 초래할 수 있다. 특정 실시형태에서, L 유전자를 결하는 재조합 광견병 바이러스 게놈은 L 유전자의 전부 또는 일부를 결할 수 있으며, 여기서 L 유전자의 일부는 L 유전자 생성물의 기능에 필요하다. 특정 실시형태에서, L 유전자 생성물의 기능에 필요한 L 유전자의 일부가 결여되면 절삭된 비-기능성 RNA-의존성 RNA 폴리머라제의 생산을 초래할 수 있다.

[0205] 특정 실시형태에서, 본 개시내용의 재조합 광견병 바이러스 게놈은 이식유전자를 포함하는 핵산을 인코딩한다. 특정 실시형태에서, 이식유전자를 포함하는 핵산은 본 명세서에 기재된 바와 같이 제거된 하나 이상의 광견병 바이러스 유전자를 대체한다. 예를 들어, 이식유전자를 포함하는 핵산은 광견병 바이러스 유전자의 전부 또는 일부를 대체할 수 있다. 특정 실시형태에서, 이식유전자를 포함하는 핵산은 G 유전자의 전부 또는 일부를 대체하며, 여기서 G 유전자의 일부는 G 유전자 생성물의 기능에 필요하다. 특정 실시형태에서, 이식유전자를 포함하는 핵산은 L 유전자의 전부 또는 일부를 대체하며, 여기서 L 유전자의 일부는 L 유전자 생성물의 기능에 필요하다. 특정 실시형태에서, 이식유전자를 포함하는 핵산은 L 유전자의 전부 또는 일부를 대체하며, 여기서 L 유전자의 일부는 L 유전자 생성물의 기능에 필요하며; 그리고 G 유전자의 전부 또는 일부로, 여기서 G 유전자의 일부는 G 유전자 생성물의 기능에 필요하다.

[0206] 특정 실시형태에서, 본 개시내용의 재조합 광견병 바이러스 게놈은 이식유전자를 포함하는 핵산을 인코딩하며, 여기서 이식유전자는 본 명세서에 기재된 바와 같이 제거된 하나 이상의 광견병 바이러스 유전자를 대체한다. 특정 실시형태에서, 재조합 광견병 바이러스 게놈은 광견병 바이러스 핵단백질 또는 이의 기능적 변이체를 인코딩하는 N 유전자, 광견병 바이러스 인단백질 또는 이의 기능적 변이체를 인코딩하는 P 유전자, 및/또는 광견병

바이러스 매트릭스 단백질 또는 이의 기능적 변이체를 인코딩하는 M 유전자를 포함한다.

[0207] N, P, M, L 및 G 유전자의 예시적인 핵산 서열, 및 이의 유전자 생성물의 아미노산 서열이 표 1에 제공되어 있다.

[0208] 표 1: N, P, M, L 및 G의 예시적인 서열

서열번호:	서열
서열번호: 4001 N 유전자 (핵산)	atggatgccgacaagattgtattcaaagtcaataatcaggtggtctctttgaagcctgagattatcgtggatcaatatgagta caagtaccctgccatcaaagatttgaaaaagccctgtataaccctagggaaaggctcccgatttaaataaagcatacaagt cagttttgtcaggcatgagcgccgccaaacttaatcctgacgatgtatgttcctatttggcagcggcaatgcagttttttgagg ggacatgtccggaagactggaccagctatggaattgtgattgcacgaaaaggagataagatcaccccaggttctctggtg gagataaaacgtactgatgtagaagggaattgggctctgacaggaggcatggaactgacaagagaccccactgtccct gagcatgcgtccttagtcggtcttctcttgaagtctgtataggttgagcaaaatatccgggcaaaacactggtaactataagac aaacattgcagacaggatagagcagatttttgagacagccccttttgttaaaatcgtggaacaccatactctaatgacaact cacaaaatgtgtgctaattggagtactatacccaaacttcagatttttggccggaacctatgacatgtttttctcccggattgagc atctatattcagcaatcagagtgggcacagttgtcactgcttatgaagactgttcaggactggtatcatttactgggttcataaa acaaatcaatctcaccgctagagaggcaatactatatttcttccacaagaactttgaggaagagataagaagaatgtttga gccagggcaggagacagctgttcctcactcttatttcatccacttccgttcactaggcttgagtgggaaatctccttattcatca aatgctgttggtcacgtgttcaatctcattcactttgtaggatgctatatgggtcaagtcagatccctaaatgcaacggttattgc tgcatgtgctcctcatgaaatgtctgttctaggggctatctgggagaggaattcttcgggaaagggacatttgaaagaaga ttcttcagagatgagaaagaacttcaagaatacgaggcggctgaactgacaaagactgacgtagcactggcagatgat ggaactgtcaactctgacgacgaggactacttttcaggtgaaaccagaagtccggaggctgtttatactcgaatcatgatg aatggaggtcgactaaagagatctcacatacggagatatgtctcagtcagttccaatcatcaagcccgtccaaactcattc gccgagtttctaaacaagacatattcgagtgactca
서열번호: 4013 최적화된 N 유전자 코돈 (핵산)	ATGGATGCCGACAAGATTGTATTCAAAGTCAACAACCAGGTGGTGTCCCTGAAGCC TGAGATCATCGTGGACCAGTACGAGTACAAGTACCCCGCCATCAAGGACCTGAAG AAGCCCTGTATCACCCTGGGCAAAGCCCCTGACCTGAACAAAGCCTACAAGAGCG TGCTGAGCGGCATGTCTGCCGCCAAGCTGAACCCTGATGACGTGTGCTCTTATCT GGCCGCTGCCATGCAGTTCTTCGAGGGCACATGTCCCGAGGACTGGACCAGCTAT GGAATCGTGATCGCCCGGAAGGGCGATAAGATCACACCTGGCAGCCTGGTGGAA ATCAAGAGAACCGACGTGGAAGGCAACTGGGCCCTGACAGGTGGCATGGAACTG ACCAGAGATCCCACCGTGCCTGAGCACGCTTCTCTTGTTGGACTGCTGCTGAGCC TGTACCGGCTGTCTAAGATCAGCGGACAGAACACCGGCAACTACAAGACCAATAT CGCCGACCGGATCGAGCAGATTTTCGAGACAGCCCCTTTCGTGAAGATCGTGGAA CACCACACACTGATGACCACACACAAGATGTGCGCCAACTGGTCTACAATCCCCAA CTTCAGATTCCTGGCCGGCACCTACGACATGTTCTTCAGCAGAATCGAGCACCTGT ACTCTGCCATCAGAGTGGGCACAGTCGTGACCGCCTACGAGGATTGCTCTGGCCT GGTGTCCTTCACCGGCTTCATCAAGCAGATCAACCTGACCGCCAGAGAGGCCATC CTGTACTTTTTCCACAAGAACTTCGAGGAAGAGATCCGGCGGATGTTCGAGCCCG GACAAGAAACAGCCGTGCCTCACAGCTACTTCATCCACTTCAGAAGCCTGGGCCT GTCCGGCAAGAGCCCCTACTCTTCTAATGCCGTGGGCCACGTGTTCAACCTGATC CACTTCGTGGGCTGCTACATGGGCCAAGTGCGGAGCCTGAATGCCACAGTGATTG CCGCCTGTGCTCCCCACGAAATGTCTGTGCTCGGAGGCTATCTGGGCGAAGAGTT CTTTGGCAAGGGCACCTTCGAGCGGAGATTCTTCCGGGACGAGAAAGAGCTGCAA GAGTACGAGGCCGCCGAGCTGACCAAAACAGATGTGGCCCTGGCCGATGACGGC ACCGTGAACTCTGACGACGAGGACTACTTCAGCGGCGAGACTAGAAGCCCCGAG GCCGTGTATACCCGGATTATGATGAATGGCGGCAGGCTGAAGCGGAGCCACATTC GGAGATACGTGTCCGTGTCCAGCAACCATCAGGCCAGACCTAACAGCTTCGCCGA GTTCCTGAACAAGACCTACTCGAGTGACTCATAA
서열번호: 4002 N 유전자 (아미노산)	MDADKIVFKVNNQVVSLKPEIIVDQYEYKYPAIKDLKKPCITLGKAPDLNKAYKSVLSGM SAAKLNPDDVCSYLAAAMQFFEGTCPEDWTSYGIVIARKGDKITPGSLVEIKRTDVEGN WALTGGMELTRDPTVPEHASLVGLLLSLYRLSKISGQNTGNYKTNIADRIEQIFETAPFV KIVEHHTLMTTHKMCANWSTIPNFRFLAGTYDMFFSRIEHLYSAIRVGTVVTAYEDCSG LVSFTGFIKQINLTAREAILYFFHKNFEEEIRRMFEPGQETAVPHSYFIHFRSLGLSGKSP YSSNAVGHVFNLIHFVGCYMGQVRSLNATVIAACAPHEMSVLGGYLGEEFFGKGTFER RFFRDEKELQEYEAAELTKTDVALADDGTVNSDDEDYFSGETRSPEAVYTRIMMNGG RLKRSHIRRYVSVSSNHQARPNSFAEFLNKTYSSDS
서열번호: 4003	ctcgatcctggagaggtctatgatgaccctattgacccaatcgagttagaggctgaacccagaggaacccccattgtcccc aacatcttgaggaactctgactacaatctcaactctcctttgatagaagatcctgctagactaatgttagaatggttaaaaac agggaatagacctatcggatgactctaacagacaattgctccaggtctttcagagttttgaaagattatttcaagaaggtag atttgggttctctcaaggtgggcggaatggctgcacagtcaatgatttctctctggttatatggtgcccactctgaatccaacag

[0209]

L 유전자 (핵산)	gagccggagatgtataacagacttggcccatttctattccaagtcgtcccccatagagaagctgttgaatctcacgctagga aatagagggctgagaatccccccagagggagtgttaagttgccttgagagggttgattatgataatgcatttggaaggtatc ttgccaacacgtattcctcttacttgttcttccatgtaatcaccttatacatgaacgccctagactgggatgaagaaaagacca tcctagcattatggaaagatttaacctcagtggacatcgggaaggacttggtaaagttcaaagaccaaatatggggactg ctgatcgtgacaaaggactttgtttactcccaaagttccaattgtctttttgacagaaactacacacttatgctaaaagatcttttc ttgtctcgcttcaactccttaatggtcttgctctctcccccagagccccgatactcagatgacttgatatctcaactatgccagct gtacattgctggggatcaagtcttgtctatgtgtggaaactccggctatgaagtcatcaaaatattggagccatatgtcgtga atagtttagtccagagagcagaaaagtttaggcctctcattcattccttgggagactttcctgtatttataaaagacaaggtaa gtcaacttgaagagacgttcggtccctgtgcaagaaggttctttagggctctggatcaattcgacaacatacatgacttggttt ttgtgtttggctgttacaggcattgggggcacccatatatagattatcgaaagggtctgtcaaaactatatgatcaggttcacct taaaaaaatgatagataagtcctaccaggagtgcttagcaagcgacctagccaggaggatccttagatgggggttttgata agtactccaagtggtatctggattcaagattcctagcccgagaccacccctgactcctatatacaaaacccaaacatggcc acccaaacatattgtagacttggtgggggatacatggcacaagctcccgatcacgcagatctttgagattcctgaatcaat ggatccgtcagaaatattggatgacaaatcacattctttcaccagaacgagactagcttcttggctgtcagaaaaccgagg ggggcctgttcctagcgaaaaagttattatcacggccctgtctaagccgcctgtcaatccccgagagtttctgaggtctatag acctcggaggattgccagatgaagacttgataattggcctcaagccaaaggaacgggaattgaagattgaaggtcgatt ctttgctctaatgtcatggaatctaagattgtattttgtcatcactgaaaaactcttggccaactacatcttgccactttttgacgcg ctgactatgacagacaacctgaacaaggtgtttaaaaagctgatcgacagggtcaccgggcaagggcttttggactattc aagggtcacatatgcatttcacctggactatgaaaagtggaacaaccatcaaagattagagtcaacagaggatgtattttc tgtcctagatcaagtgtttggattgaagagagtgttttctagaacacacgagtttttttcaaaaggcctggatctattattcagac agatcagacctcatcgggttacgggaggatcaaatatactgcttagatgcgtccaacggcccaacctgttggaatggcca ggatggcgggctagaaggcttacggcagaagggctggagtctagtcagcttattgatgatagatagagaatctcaaatca ggaacacaagaaccaaaatactagctcaaggagacaaccaggttttatgtccgacatacatgttgtcgccagggctatct caagaggggctcctctatgaattggagagaatatcaaggaatgcactttcgatatacagagccgtcgaggaaggggcat ctaagctagggctgatcatcaagaaagaagagaccatgtgtagttatgacttcctcatctatggaaaaacccctttgtttaga ggtaacatattggtgcctgagtccaaaagatgggccagagtctcttgcgtctctaatgaccaaatagtcaacctcgccaata taatgtcgacagtgtccaccaatgcgctaacagtggcacaacactctcaatctttgatcaaaccgatgagggattttctgctc atgtcagtacaggcagtctttcactacctgctatttagcccaatcttaaagggaagagtttacaagattctgagcgctgaagg ggagagctttctcctagccatgtcaaggataatctatctagatccttctttgggagggatatctggaatgtccctcggaagatt ccatatacgacagttctcagaccctgtctctgaagggttatccttctggagagagatctggttaagctcccaagagtcctgga ttcacgcgttgtgtcaagaggctggaaacccagatcttggagagagaacactcgagagcttcactcgccttctagaagatc cgaccaccttaaatatcagaggaggggccagtcctaccattctactcaaggatgcaatcagaaaggctttatatgacgag gtggacaaggtggaaaattcagagtttcgagaggcaatcctgttgtccaagacccatagagataattttatactcttcttaat atctgttgagcctctgtttcctcgatttctcagtgagctattcagttcgtctttttgggaatccccgagtcaatcattggattgatac aaaactcccgaacgataagaaggcagtttagaaagagtctctcaaaaactttagaagaatccttctacaactcagagatc cacgggattagtcggatgacccagacacctcagagggttgggggggtgtggccttgctcttcagagagggcagatctact tagggagatctcttggggaagaaaagtggtaggcacgacagttcctcacccttctgagatgttgggattacttcccaagtcc tctatttcttgcacttgtggagcaacaggaggaggcaatcctagagtttctgtatcagtactcccgtcctttgatcagtcatttttttt cacgaggccccctaaagggatacttgggctcgtccacctctatgtcgacccagctattccatgcatgggaaaaagtcacta atgttcatgtggtgaagagagctctatcgttaaaagaatctataaactggttcattactagagattccaacttggctcaagctct aattaggaacattatgtctctgacaggccctgatttccctctagaggaggcccctgtcttcaaaaggacggggtcagccttg cataggttcaagtctgccagatacagcgaaggagggtattcttctgtctgcccgaacctcctctctcatatttctgttagtacag acaccatgtctgatttgacccaagacgggaagaactacgatttcatgttccagccattgatgctttatgcacagacatggac atcagagctggtacagagagacacaaggctaagagactctacgtttcattggcacctccgatgcaacaggtgtgtgaga cccattgacgacgtgaccctggagacctctcagatcttcgagtttccggatgtgtcgaaaagaatatccagaatggtttctgg ggctgtgcctcacttccagaggcttcccgatatccgtctgagaccaggagattttgaatctctaagcggtagagaaaagtct caccatatcggatcagctcaggggctcttatactcaatcttagtggcaattcacgactcaggatacaatgatggaaccatctt ccctgtcaacatatacggcaaggtttcccctagagactatttgagagggctcgcaaggggagtattgataggatcctcgatt tgcttcttgacaagaatgacaaatatcaatattaatagacctcttgaattggtctcaggggtaatctcatatattctcctgaggct agataaccatccctccttgtacataatgctcagagaaccgtctcttagaggagagatattttctatccctcagaaaatccccg ccgcttatccaaccactatgaaagaaggcaacagatcaatcttgtgttatctccaacatgtgctacgctatgagcgagagat aatcacggcgtctccagagaatgactggctatggatcttttcagactttagaagtgccaaaatgacgtacctatccctcatta cttaccagtctcatcttctactccagagggttgagagaaacctatctaagagtatgagagataacctgcgacaattgagttct ttgatgaggcaggtgctgggcgggcacggagaagataccttagagtcagacgacaacattcaacgactgctaaaagac tctttacgaaggacaagatgggtggatcaagaggtgcgccatgcagctagaaccatgactggagattacagccccaac aagaaggtgtcccgtaaggtaggatgttcagaatgggtctgctctgctcaacaggttgcagtctctacctcagcaaacccg gcccctgtctcggagcttgacataagggccctctctaagaggttccagaaccctttgatctcgggcttgagagtggttcagtg ggcaaccggtgctcattataagcttaagcctattctagatgatctcaatgttttcccatctctctgccttgtagttggggacgggt caggggggatatcaagggcagtcctcaacatgtttccagatgccaagcttgtgttcaacagtcttttagaggtgaatgacct gatggcttccggaacacatccactgcctccttcagcaatcatgaggggaggaaatgatatcgtctccagagtgatagatctt

[0210]

	gactcaatctgggaaaaaccgtccgacttgagaaacttggcaacctggaaatacttccagtcagtccaaaagcaggtca acatgtcctatgacctcattatttgcgatgcagaagttactgacattgcatctatcaaccggatcaccctgttaatgtccgatttt gcattgtctatagatggaccactctatttggtcttcaaaacttatgggactatgctagtaaatccaaactacaaggctattcaa cacctgtcaagagcgttcccctcggtcacagggtttatcacccaagtaacttcgtctttttcatctgagctctacctccgattctc caaacgagggaagtttttcagagatgctgagtacttgacctcttccacccttcgagaaatgagccttgtgttattcaattgtag cagccccaagagtgagatgcagagagctcgttccttgaactatcaggatcttgtgagaggatttcctgaagaaatcatatc aaatccttacaatgagatgatcataactctgattgacagtgatgtagaatcttttctagtccacaagatggttgatgatcttgag ttacagaggggaactctgtctaaagtggctatcattatagccatcatgatagttttctccaacagagtcttcaacgtttccaaac ccctaactgacccctcgttctatccaccgtctgatcccaaaatcctgaggcacttcaacatatgttgcagtactatgatgtatct atctactgctttaggtgacgtccctagcttcgcaagacttcacgacctgtataacagacctataacttattacttcagaaagca agtcattcgagggaacgtttatctatcttggagttggtccaacgacacctcagtgttcaaaagggtagcctgtaattctagcct gagtctgtcatctcactggatcaggttgatttacaagatagtgaagactaccagactcgttggcagcatcaaggatctatcc agagaagtggaaagacaccttcataggtacaacaggtggatcaccctagaggatatcagatctagatcatccctactag actacagttgcctg
서열번호: 4014 최적화된 L 유전자 코돈 (핵산)	ATGCTCGATCCTGGAGAGGTCTATGATGACCCCATCGATCCTATCGAGCTGGAAG CCGAGCCTAGAGGCACACCCATCGTGCCCAACATCCTGCGGAACAGCGACTACAA CCTGAACAGCCCTCTGATCGAGGACCCCGCCAGACTGATGCTGGAATGGCTGAAA ACCGGCAACAGACCCTACCGGATGACCCTGACCGATAACTGCTCCCGGTCCTTCA GGGTGCTGAAGGACTACTTCAAGAAGGTGGACCTGGGCAGCCTGAAAGTCGGAG GAATGGCCGCTCAGAGCATGATCAGCCTGTGGCTGTATGGCGCCCACAGCGAGA GCAACAGATCCAGAAGATGCATCACCGATCTGGCCCACTTCTACAGCAAGAGCAG CCCCATCGAGAAGCTGCTGAATCTGACCCTGGGCAACCGCGGCCTGAGAATACCT CCTGAAGGCGTGCTGAGCTGCCTGGAAAGAGTGGACTACGACAACGCCTTCGGC AGATACCTGGCCAACACCTACAGCAGCTACCTGTTCTTCCACGTGATCACCCTGTA TATGAACGCCCTGGACTGGGACGAAGAGAAAACCATTCTGGCCCTGTGGAAGGAC CTGACCTCTGTGGACATCGGCAAGGACCTGGTCAAGTTCAAGGACCAGATTTGGG GCCTGCTGATCGTGACCAAGGACTTCGTGTACTCCCAGAGCAGCAACTGCCTGTT CGACCGGAACTACACCCTGATGCTGAAAGACCTGTTCCTGAGCCGGTTCAACAGC CTGATGGTGCTGCTGTCTCCTCCAGAGCCTAGATACAGCGACGACCTGATCTCCC AGCTGTGCCAGCTGTATATTGCCGGCGATCAGGTGCTGAGCATGTGCGGCAATAG CGGCTACGAAGTGATCAAGATCCTGGAACCTTACGTCGTGAACAGCCTGGTGCAG CGGGCCGAGAAGTTCAGACCACTGATTCACAGCCTGGGCGACTTCCCCGTGTTCA TCAAGGACAAGGTGTCCCAGCTGGAAGAGACATTCGGCCCCTGCGCCAGAAGATT CTTCAGAGCCCTGGACCAGTTCGACAACATCCACGACCTGGTGTTCGTGTTCGGC TGCTACAGACACTGGGGACACCCCTACATCGACTACAGAAAGGGCCTGAGCAAGC TGTACGACCAGGTTCACCTGAAGAAGATGATCGACAAGAGCTACCAAGAGTGCCT GGCCAGCGACCTGGCCAGACGTATTCTGAGATGGGGCTTCGACAAGTACAGCAAG TGGTATCTGGACAGCCGGTTCCTGGCTCGGGATCACCCTCTGACTCCCTACATCA AGACCCAGACCTGGCCTCCTAAGCACATCGTGGATCTCGTGGGCGACACCTGGCA CAAGCTGCCCATCACACAGATTTTCGAGATCCCCGAGAGCATGGACCCCAGCGAG ATTCTGGACGACAAGTCCCACAGCTTCACCCGGACAAGACTGGCCTCTTGGCTGA GCGAGAATAGAGGCGGACCTGTGCCTAGCGAGAAAGTGATCATCACCGCTCTGTC CAAGCCTCCAGTGAACCCCAGAGAGTTCCTGCGGTCTATCGATCTCGGCGGCCTG CCTGATGAGGACCTGATCATTGGCCTGAAGCCTAAAGAGCGCGAGCTGAAGATCG AGGGCAGATTCTTCGCCCTGATGAGCTGGAACCTGCGGCTGTACTTTGTGATCAC CGAGAAACTGCTGGCCAACTACATCCTGCCTCTGTTCGACGCCCTGACCATGACC GACAATCTGAACAAGGTGTTCAAGAAACTGATCGACAGAGTGACCGGCCAGGGAC TGCTGGACTACAGCAGAGTGACATACGCCTTCCACCTGGATTACGAGAAGTGGAA CAACCACCAGCGGCTGGAAAGCACCGAGGACGTGTTCTCTGTGCTGGACCAGGT GTTCGGCCTGAAGAGAGTGTTCAGCAGAACCCACGAGTTCTTCCAGAAAGCCTGG ATCTACTACAGCGACCGCAGCGATCTGATCGGACTGAGAGAGGACCAAATCTACT GCCTGGACGCCAGCAATGGCCCTACCTGTTGGAATGGACAGGACGGCGGACTGG AAGGACTGAGACAGAAAGGCTGGTCCCTGGTGTCCCTGCTGATGATTGACAGAGA GAGCCAGATCAGAAACACGCGGACCAAGATTCTGGCTCAGGGCGACAACCAGGT GCTGTGCCCTACCTATATGCTGAGCCCTGGCCTGTCTCAAGAGGGGCTGCTGTAC GAACTGGAACGGATCAGCAGAAACGCCCTGTCCATCTATAGAGCCGTGGAAGAGG GCGCCTCTAAGCTGGGCCTGATCATCAAGAAAGAAGAGACAATGTGCAGCTACGA CTTCCTGATCTACGGCAAGACCCCTCTGTTCCGGGGCAACATTCTGGTGCCCGAG TCCAAGAGATGGGCCAGAGTGTCCTGCGTGTCCAACGACCAGATCGTGAATCTGG CCAATATCATGTCCACCGTGTCCACCAACGCTCTGACAGTGGCCCAGCACAGCCA

[0211]

[0212]

	GTCTCTGATCAAGCCTATGCGGGACTTCCTGCTCATGTCCGTGCAGGCCGTGTTC CACTACCTGCTGTTCAGCCCTATCCTGAAGGGCAGAGTGTATAAGATCCTGAGCG CCGAGGGCGAGAGCTTCCTGCTTGCCATGAGCCGGATCATCTATCTGGACCCTAG CCTCGGCGGCATCAGCGGCATGTCTCTGGGCAGATTTCACATCCGGCAGTTCAGC GACCCTGTGTCCGAGGGCCTGTCCTTTTGGAGAGAAATCTGGCTGTCTAGCCAAG AGAGCTGGATTCACGCCCTGTGCCAAGAAGCCGGCAATCCCGATCTGGGCGAGA GAACCCTGGAATCCTTCACCAGACTGCTTGAGGACCCCACCACTCTGAACATCAG AGGCGGAGCCTCTCCAACCATCCTGCTCAAGGACGCCATCCGCAAGGCCCTGTAT GACGAGGTGGACAAAGTGGAAAACAGCGAGTTCAGAGAGGCCATTCTGCTGAGCA AGACCCACCGGGACAACTTCATCCTGTTCCTCATCTCCGTGGAACCCCTGTTTCCT AGATTCCTGTCCGAGCTGTTCTCCAGCTCCTTCCTGGGCATCCCTGAGTCCATCAT CGGCCTGATCCAGAACTCCCGGACCATCCGCAGACAGTTCAGAAAGAGCCTGTCT AAGACCCTGGAAGAGTCCTTCTACAACAGCGAAATCCACGGCATCTCCCGGATGA CACAGACCCCTCAAAGAGTCGGCGGCGTCTGGCCTTGTTCTAGCGAAAGAGCCGA CCTGCTGAGAGAGATCAGCTGGGGCAGAAAGGTCGTGGGCACCACAGTGCCTCA TCCAAGCGAAATGCTGGGCCTCCTGCCTAAGAGCAGCATCAGCTGTACCTGTGGC GCTACCGGCGGAGGCAATCCTAGAGTGTCTGTGTCTGTGCTGCCCAGCTTCGACC AGAGCTTCTTCAGCAGAGGACCTCTGAAGGGCTACCTGGGCTCTAGCACCAGCAT GAGCACCCAGCTGTTTCACGCCTGGGAAAAAGTGACCAATGTGCACGTGGTCAAG AGAGCCCTGTCTCTGAAAGAGAGCATCAACTGGTTCATCACGCGGGACAGCAATC TGGCACAGGCCCTGATTCGGAACATCATGTCCCTGACTGGCCCTGACTTTCCACT GGAAGAGGCCCCTGTGTTCAAGCGCACAGGATCTGCCCTGCACAGATTCAAGAGC GCCAGATACTCCGAAGGCGGCTACAGCTCCGTGTGTCCCAATCTGCTGTCCCACA TCTCTGTGTCCACCGACACCATGTCCGATCTGACCCAGGACGGCAAGAACTACGA TTTCATGTTCCAGCCTCTCATGCTGTACGCCCAGACATGGACAAGCGAGCTGGTG CAGAGAGACACCCGGCTGAGAGATAGCACCTTCCACTGGCACCTGAGGTGCAACA GATGCGTGCGGCCCATCGACGACGTGACACTGGAAACCTCTCAGATATTCGAGTT CCCAGACGTGTCCAAGCGGATCAGCCGAATGGTGTCTGGCGCCGTGCCTCACTTC CAAAGACTGCCCGACATCAGACTGCGGCCAGGCGATTTTGAGAGCCTGAGCGGC AGAGAGAAGTCCCACCACATCGGATCTGCACAGGGCCTGCTCTACTCTATCCTGG TGGCCATCCACGATTCCGGCTACAACGACGGCACCATCTTTCCCGTGAACATCTAC GGAAAAGTGTCCCCACGGGACTACCTGAGAGGACTGGCTAGAGGGGTGCTGATC GGCAGCAGCATCTGCTTTCTGACCAGAATGACCAACATCAACATCAATAGGCCCCT GGAACTGGTGTCCGGCGTGATCAGCTATATCCTGCTGCGGCTGGACAATCACCCC AGCCTGTATATCATGCTGAGGGAACCCAGCCTGCGGGGCGAGATTTTTAGCATCC CTCAGAAGATCCCTGCCGCCTATCCTACCACCATGAAGGAAGGCAATCGGAGCAT CCTGTGCTACCTCCAGCATGTGCTGAGATACGAGCGGGAAATCATTACCGCCTCT CCAGAGAACGATTGGCTGTGGATCTTTAGCGACTTCCGCAGCGCCAAGATGACCT ACCTGAGCCTGATCACCTACCAGAGCCATCTGCTGCTCCAGAGAGTGGAACGGAA CCTGTCCAAGAGCATGAGGGACAACCTGAGGCAGCTGTCCTCTCTGATGAGACAG GTGCTCGGAGGACACGGCGAGGATACACTGGAATCTGACGACAATATCCAGCGG CTCCTGAAAGACAGCCTGAGAAGAACCAGATGGGTTGACCAAGAAGTGCGCCACG CCGCCAGAACAATGACCGGCGATTACAGCCCCAACAAGAAAGTGTCCAGAAAAGT GGGCTGCTCCGAGTGGGTCTGCTCTGCTCAGCAAGTTGCCGTGTCTACCAGCGCC AATCCTGCACCAGTTTCCGAGCTGGACATTAGAGCCCTGAGCAAACGGTTCCAGA ATCCTCTGATCTCTGGCCTGAGAGTGGTGCAGTGGGCTACAGGCGCCCACTACAA GCTGAAGCCCATCCTGGACGATCTGAACGTGTTCCCTAGCCTGTGTCTGGTCGTC GGAGATGGATCTGGCGGAATCAGCAGAGCCGTGCTGAATATGTTCCCCGACGCCA AGCTGGTGTTCAATAGCCTGCTGGAAGTGAACGATCTGATGGCCAGCGGCACACA CCCTCTGCCACCAAGCGCAATTATGAGAGGCGGCAACGACATCGTGTCCAGAGTG ATCGACCTGGACTCCATCTGGGAGAAGCCCTCCGACCTGAGAAACCTGGCCACCT GGAAGTACTTTCAGAGCGTGCAGAAACAAGTGAACATGAGCTACGACCTCATCATC TGCGACGCCGAAGTGACCGACATTGCCTCCATCAACAGAATCACACTGCTGATGT CCGACTTCGCTCTGAGCATCGACGGCCCTCTGTATCTGGTGTTTAAGACCTACGG CACAATGCTCGTGAACCCTAACTACAAGGCCATCCAGCACCTCAGCAGGGCCTTT CCAAGCGTGACCGGCTTCATCACCCAAGTGACCTCCAGCTTCAGCAGCGAGCTGT ATCTGCGGTTCAGCAAGCGGGGCAAGTTCTTCAGGGACGCCGAGTACCTGACCAG CAGCACACTGAGAGAAATGAGTCTGGTGCTGTTCAACTGCTCTAGCCCCAAGAGC GAGATGCAGAGAGCTAGAAGCCTGAACTACCAGGACCTCGTGCGGGGCTTCCCC GAAGAGATCATCTCTAACCCCTACAACGAGATGATCATTACCCTGATCGACTCCGA

	CGTCGAGAGCTTTCTGGTGCACAAGATGGTGGACGACCTGGAACTTCAGAGGGGC ACACTGTCCAAGGTGGCCATTATCATTGCCATTATGATCGTGTTCTCCAACCGGGT TTTCAATGTCTCCAAGCCACTGACAGACCCCAGCTTCTACCCTCCTAGCGACCCAA AGATCCTGCGGCACTTCAACATCTGTTGTAGCACCATGATGTACCTGAGCACCGCA CTGGGAGATGTGCCATCCTTTGCCAGACTGCACGACCTGTATAACAGACCCATCA CCTACTACTTTCGGAAGCAAGTGATCCGCGGCAACGTGTACCTGTCCTGGTCTTG GAGCAACGACACCAGCGTGTTCAAAAGAGTGGCCTGCAACAGCTCTCTGTCCCTG AGCAGCCACTGGATCAGACTGATCTACAAGATCGTCAAGACCACACGGCTCGTGG GCAGCATCAAGGATCTGAGTAGAGAGGTGGAAAGGCATCTGCATCGGTACAATCG GTGGATCACACTTGAGGACATCCGGTCCAGATCATCCCTACTAGACTACAGTTGCC TGTGA
서열번호: 4004 L 유전자 (아미노산)	LDPGEVYDDPIDPIELEAEPRGTPIVPNILRNSDYNLNSPLIEDPARLMLEWLKTGNRPY RMTLTDNCSRSFRVLKDYFKKVDLGSLKVGGMAAQSMISLWLYGAHSESNRSRRCIT DLAHFYSKSSPIEKLLNLTLGNRGLRIPPEGVLSCLERVDYDNAFGRYLANTYSSYLFF HVITLYMNALDWDEEKTILALWKDLTSVDIGKDLVKFKDQIWGLLIVTKDFVYSQSSNCL FDRNYTLMLKDLFLSRFNSLMVLLSPPEPRYSDDLISQLCQLYIAGDQVLSMCGNSGY EVIKILEPYVVNSLVQRAEKFRPLIHSLGDFPVFIKDKVSQLEETFGPCARRFFRALDQF DNIHDLVFVFGCYRHWGHPYIDYRKGLSKLYDQVHLKKMIDKSYQECLASDLARRILR WGFDKYSKWYLDSRFLARDHPLTPYIKTQTWPPKHIVDLVGDTWHKLPITQIFEIPESM DPSEILDDKSHSFTRTRLASWLSENRGGPVPSEKVIITALSKPPVNPREFLRSIDLGGLP DEDLIIGLKPKERELKIEGRFFALMSWNLRLYFVITEKLLANYILPLFDALTMTDNLNKVF KKLIDRVTGQGLLDYSRVTYAFHLDYEKWNNHQRLESTEDVFSVLDQVFGLKRVFSRT HEFFQKAWIYYSDRSDLIGLREDQIYCLDASNGPTCWNGQDGGLEGLRQKGWSLVSL LMIDRESQIRNTRTKILAQGDNQVLCPTYMLSPGLSQEGLLYELERISRNALSIYRAVEE GASKLGLIIKKEETMCSYDFLIYGKTPLFRGNILVPESKRWARVSCVSNDQIVNLANIMS TVSTNALTVAQHSQSLIKPMRDFLLMSVQAVFHYLLFSPILKGRVYKILSAEGESFLLAM SRIIYLDPSLGGISGMSLGRFHIRQFSDPVSEGLSFWREIWLSSQESWIHALCQEAGNP DLGERTLESFTRLLEDPTTLNIRGGASPTILLKDAIRKALYDEVDKVENSEFREAILLSKT HRDNFILFLISVEPLFPRFLSELFSSSFLGIPESIIGLIQNSRTIRRQFRKSLSKTLEESFYN SEIHGISRMTQTPQRVGGVWPCSSERADLLREISWGRKVVGTTVPHPSEMLGLLPKS SISCTCGATGGGNPRVSVSVLPSFDQSFFSRGPLKGYLGSSTSMSTQLFHAWEKVTN VHVVKRALSLKESINWFITRDSNLAQALIRNIMSLTGPDFPLEEAPVFKRTGSALHRFKS ARYSEGGYSSVCPNLLSHISVSTDTMSDLTQDGKNYDFMFQPLMLYAQTWTSELVQR DTRLRDSTFHWHLRCNRCVRPIDDVTLETSQIFEFPDVSKRISRMVSGAVPHFQRLPDI RLRPGDFESLSGREKSHHIGSAQGLLYSILVAIHDSGYNDGTIFPVNIYGKVSPRDYLR GLARGVLIGSSICFLTRMTNININRPLELVSGVISYILLRLDNHPSLYIMLREPSLRGEIFSI PQKIPAAYPTTMKEGNRSILCYLQHVLRYEREIITASPENDWLWIFSDFRSAKMTYLSLI TYQSHLLLQRVERNLSKSMRDNLRQLSSLMRQVLGGHGEDTLESDDNIQRLLKDSLR RTRWVDQEVRHAARTMTGDYSPNKKVSRKVGCSEWVCSAQQVAVSTSANPAPVSEL DIRALSKRFQNPLISGLRVVQWATGAHYKLKPILDDLNVFPSLCLVVGDGSGGISRAVL NMFPDAKLVFNSLLEVNDLMASGTHPLPPSAIMRGGNDIVSRVIDLDSIWEKPSDLRNL ATWKYFQSVQKQVNMSYDLIICDAEVTDIASINRITLLMSDFALSIDGPLYLVFKTYGTML VNPNYKAIQHLSRAFPSVTGFITQVTSSFSSELYLRFSKRGKFFRDAEYLTSSTLREMS LVLFNCSSPKSEMQRARSLNYQDLVRGFPEEIISNPYNEMIITLIDSDVESFLVHKMVDD LELQRGTLSKVAIIIAIMIVFSNRVFNVSKPLTDPSFYPPSDPKILRHFNICCSTMMYLSTA LGDVPSFARLHDLYNRPITYYFRKQVIRGNVYLSWSWSNDTSVFKRVACNSSLSLSSH WIRLIYKIVKTTRLVGSIKDLSREVERHLHRYNRWITLEDIRSRSSLLDYSCL
서열번호: 4005 M 유전자 (핵산)	ttctagaagcagagaggaatctttgtcctcttcggacctttgtgtctgaagagacatgtcagaccatagttgacatgctctcgg gttcatgttgatacaccagactctgccctggatatgacactgttttgcaatcactcttatttgcaatccgacgaactcagtatcat catcccaagtgatctcctgagagtattccaactcctcccctcaagagggcccctggaatcagcccactggaagataaag gttctcctcaatttgtatacccagttcaggccctcagggactggagatcctgacaaagccagtccaataaccactttgactaa cccgatcatcctatgattcccagaatatatctcgtcgaatgatttcagaatgtgccgcaggatcctgaacgagtaaccattcg ggctacacactttaacccttccgttgatacaaaagttcctcatgttcttcttgcctgtaagttctttcagcgggacgtattcaggg ggtggaagccacaagtcatcgtcatccagaggggctgacgcgggagaggatttttgagtgtcctcgtccctgcggtttttca ctatcttacgtaggaggtt
서열번호: 4014	ATGAACCTGCTGCGGAAGATCGTGAAGAATCGGCGCGACGAGGACACCCAGAAG TCTAGCCCTGCTTCTGCCCCTCTGGACGACGATGATCTGTGGCTGCCTCCTCCAG AGTACGTGCCCCTGAAAGAGCTGACCGGCAAAAAGAACATGCGGAATTTCTGCAT CAACGGCCGCGTGAAAGTGTGCAGCCCCAACGGCTACAGCTTCAGAATCCTGCG GCACATCCTGAAGTCCTTCGACGAAATCTACAGCGGCAACCACCGGATGATCGGC

[0213]

최적화된 M 유전자 코돈 (핵산)	CTGGTCAAGGTTGTGATTGGACTGGCCCTGAGCGGCTCTCCTGTTCCTGAAGGCC TGAACTGGGTCTACAAGCTGCGGCGGACATTCATTTTTCAGTGGGCCGACAGCAG AGGCCCTCTGGAAGGCGAGGAACTGGAATACTCCCAAGAGATCACCTGGGACGA CGACACCGAGTTTGTGGGCCTGCAGATCAGAGTGATCGCCAAGCAGTGCCACATC CAGGGCAGAGTGTGGTGCATCAACATGAACCCCAGAGCCTGCCAGCTTTGGAGC GACATGTCTCTGCAGACCCAGCGGAGCGAAGAGGACAAGGATtcctctctgcttctagaataa
서열번호: 4006 M 유전자 (아미노산)	NLLRKIVKNRRDEDTQKSSPASAPLDDDDLWLPPPEYVPLKELTGKKNMRNFCINGRV KVCSPNGYSFRILRHILKSFDEIYSGNHRMIGLVKVVIGLALSGSPVPEGLNWVYKLRRT FIFQWADSRGPLEGEELEYSQEITWDDDTEFVGLQIRVIAKQCHIQGRVWCINMNPRA CQLWSDMSLQTQRSEEDKDSSLLLE
서열번호: 4007 P 유전자 (핵산)	agcaagatctttgtcaatcctagtgctattagagccggtctggccgatcttgagatggctgaagaaactgttgatctgatcaat agaaatatcgaagacaatcaggctcatctccaaggggaacccatagaggtggacaatctccctgaggatatggggcga cttcacctggatgatggaaaatcgcccaaccatggtgagatagccaaggtgggagaaggcaagtatcgagaggactttc agatggatgaaggagaggatcctagcttcctgttccagtcatacctggaaaatgttggagtccaaatagtcagacaaatg aggtcaggagagagatttctcaagatatggtcacagaccgtagaagagattatatcctatgtcgcggtcaactttcccaac cctccaggaaagtcttcagaggataaatcaacccagactactggccgagagctcaagaaggagacaacacccactcc ttctcagagagaaagccaatcatcgaaagccaggatggcggctcaaattgcttctggccctccagcccttgaatggtcgg ctaccaatgaagaggatgatctatcagtggaggctgagatcgctcaccagattgcagaaagtttctccaaaaaatataagt ttccctctcgatcctcagggatactcttgtataattttgagcaattgaaaatgaaccttgatgatatagttaaagaggcaaaaa atgtaccaggtgtgacccgtttagcccatgacgggtccaaactcccctaagatgtgtactgggatgggtcgctttggccaa ctctaagaaattccagttgttagtcgaatccgacaagctgagtaaaatcatgcaagatgacttgaatcgctatacatcttgc
서열번호: 4015 최적화된 P 유전자 코돈 (핵산)	ATGAGCAAGATCTTTGTCAATCCTAGTGCTATCAGAGCCGGCCTGGCTGATCTTGA GATGGCCGAGGAAACCGTGGACCTGATCAACCGGAACATCGAGGACAATCAGGC CCATCTGCAGGGCGAGCCTATCGAGGTTGACAATCTGCCCGAGGACATGGGCAG ACTGCACCTGGATGATGGCAAGAGCCCTAACCACGGCGAGATCGCCAAAGTTGGC GAGGGCAAGTACCGCGAGGACTTCCAGATGGACGAGGGCGAAGATCCCAGCTTC CTGTTCCAGTCCTACCTGGAAAACGTGGGCGTGCAGATCGTGCGGCAGATGAGAA GCGGCGAGCGGTTCCTGAAGATTTGGAGCCAGACCGTGGAAGAGATCATCAGCTA CGTGGCCGTGAACTTCCCCAATCCTCCAGGCAAGAGCAGCGAGGACAAGAGCAC ACAGACCACCGGCAGAGAGCTGAAGAAAGAGACAACCCCTACACCTAGCCAGAGA GAGAGCCAGAGCAGCAAGGCCAGAATGGCCGCTCAGATTGCCTCTGGACCTCCA GCTCTTGAGTGGAGCGCCACCAACGAAGAGGACGACCTGTCTGTGGAAGCCGAG ATTGCCCACCAGATCGCCGAGAGCTTCAGCAAGAAGTACAAGTTCCCCAGCAGAA GCAGCGGCATCCTGCTGTATAACTTCGAGCAGCTGAAGATGAACCTGGACGACAT CGTGAAAGAGGCCAAGAACGTCCCCGGCGTGACAAGACTGGCCCACGATGGATC TAAGCTGCCCCTGAGATGTGTGCTCGGATGGGTTGCCCTGGCCAATAGCAAGAAA TTCCAGCTGCTGGTGGAAAGCGACAAGCTGTCCAAGATCATGCAGGATGACTTGA ATCGCTATACATCTTGCTAA
서열번호: 4008 P 유전자 (아미노산)	SKIFVNPSAIRAGLADLEMAEETVDLINRNIEDNQAHLQGEPIEVDNLPEDMGRLHLDD GKSPNHGEIAKVGEGKYREDFQMDEGEDPSFLFQSYLENVGVQIVRQMRSGERFLKI WSQTVEEIISYVAVNFPNPPGKSSEDKSTQTTGRELKKETTPTPSQRESQSSKARMAA QIASGPPALEWSATNEEDDLSVEAEIAHQIAESFSKKYKFPSRSSGILLYNFEQLKMNL DDIVKEAKNVPGVTRLAHDGSKLPLRCVLGWVALANSKKFQLLVESDKLSKIMQDDLN RYTSC
서열번호: 4009 G 유전자 (핵산)	atggttcctcaggctctcctgtttgtacccctctggtttttccattgtgttttgggaaattccctatttacacgataccagacaagct tggtccctggagtccgattgacatacatcacctcagctgcccaaacaatttggtagtggaggacgaaggatgcaccaacc tgtcagggttctcctacatggaacttaaagttggatacatcttagccataaaagtgaacgggttcacttgcacaggcgttgtg acggaggctgaaacctacactaacttcgttggttatgtcacaaccacgttcaaaagaaagcatttccgcccaacaccaga tgcatgtagagccgcgtacaactggaagatggccggtgaccccagatatgaagagtctctacacaatccgtaccctgact accgctggcttcgaactgtaaaaaccaccaaggagtctctcgttatcatatctccaagtgtggcagatttggacccatatga cagatcccttcactcgagggtcttccctagcgggaagtgctcaggagtagcggtgtcttctacctactgctccactaaccac gattacaccatttggatgcccgagaatccgagactagggatgtcttgtgacatttttaccaatagtagagggaagagagca

[0214]

	tccaaagggagtgagacttgcggctttgtagatgaaagaggcctatataagtctttaaaaggagcatgcaaactcaagtta tgtggagttctaggacttagacttatggatggaacatgggtctcgatgcaaacatcaaatgaaaccaaatggtgccctccc gataagttggtgaacctgcacgacttfcgctcagacgaaattgagcaccttgttgtagaggagttggtcaggaagagaga ggagtgtctggatgcactagagtccatcatgacaaccaagtcagtgagtttcagacgtctcagtcatttaagaaaacttgtc cctgggtttggaaaagcatataccatattcaacaagaccttgatggaagccgatgctcactacaagtcagtcagaacttgg aatgagatcctcccttcaaaagggtgtttaagagttgggggaggtgtcatcctcatgtgaacggggtgtttttcaatggtata atattaggacctgacggcaatgtcttaatcccagagatgcaatcatccctcctccagcaacatatggagttgttggaatcctc ggttatccccctfgtgcacccccfggcagacccgtctaccgttttcaaggacggtgacgaggctgaggattttgttgaagttc accttcccgatgtgcacaatcaggtctcaggagttgacttgggtctcccgaactgggggaagtatgtattactgagtgcagg ggccctgactgccttgatgttgataattttcctgatgacatgttgtagaagagtcaatcgatcagaacctacgcaacacaatc tcagagggacagggagggaggtgtcagtcactccccaaagcgggaagatcatatcttcatgggaatcacacaagagtg ggggtgagaccagactg
서열번호: 4016 최적화된 G 유전자 코돈 (핵산)	ATGGTTCCTCAGGCTCTGCTGTTCGTGCCCCTGCTGGTGTTCCCTCTGTGCTTCG GCAAGTTCCCCATCTACACAATCCCCGACAAGCTCGGCCCTTGGAGCCCTATCGA TATTCACCACCTGAGCTGCCCCAACAACCTGGTGGTGGAAGATGAGGGCTGCACC AACCTGAGCGGCTTCAGCTACATGGAACTGAAAGTGGGCTACATCCTGGCCATCA AAGTGAACGGCTTCACCTGTACCGGCGTCGTGACAGAGGCCGAGACATACACCAA CTTCGTGGGCTACGTGACCACCACCTTCAAGCGGAAGCACTTCAGACCCACACCT GACGCCTGTAGAGCCGCCTACAACTGGAAAATGGCCGGCGATCCCAGATACGAG GAAAGCCTGCACAACCCCTATCCTGACTACAGATGGCTGAGAACCGTGAAAACCA CCAAAGAGAGCCTGGTCATCATCAGCCCCAGCGTGGCCGATCTGGACCCCTATGA TAGATCCCTGCACAGCAGAGTGTTCCCCAGCGGCAAATGTTCTGGCGTGGCCGTG TCTAGCACCTACTGCTCCACCAACCACGACTACACCATCTGGATGCCCGAGAATCC CAGACTGGGCATGAGCTGCGACATCTTCACCAACAGCCGGGGCAAGAGAGCCAG CAAGGGCTCTGAGACATGCGGCTTCGTGGACGAGAGGGGCCTGTATAAGTCTCTG AAGGGCGCCTGCAAGCTGAAGCTGTGCGGAGTTCTGGGCCTGAGACTGATGGAT GGCACCTGGGTGTCCATGCAGACCAGCAACGAGACAAAGTGGTGCCCTCCTGACA AGCTGGTCAACCTGCACGACTTCAGAAGCGACGAGATCGAGCATCTGGTGGTCGA GGAACTCGTGCGGAAGAGGGAAGAATGCCTGGACGCCCTGGAATCCATTATGACC ACCAAGAGCGTGTCCTTCCGGCGGCTGTCTCACCTGAGAAAACTGGTGCCCGGCT TTGGCAAGGCTTACACAATCTTCAACAAGACCCTGATGGAAGCCGACGCTCACTAC AAGTCTGTGCGGACCTGGAACGAGATCCTGCCTTCCAAGGGCTGCCTGAGAGTTG GCGGAAGATGTCACCCTCACGTGAACGGCGTGTTCTTCAACGGCATCATCCTGGG ACCTGACGGCAACGTGCTGATCCCTGAGATGCAGTCTAGCCTGCTCCAGCAACAC ATGGAATTGCTGGAAAGCAGCGTGATCCCTCTGGTGCACCCTCTGGCCGATCCTA GCACCGTGTTTAAGGATGGCGACGAGGCCGAGGACTTCGTGGAAGTGCATCTGC CCGACGTGCACAATCAGGTGTCAGGCGTTGACCTGGGCCTGCCTAACTGGGGCA AATACGTGCTGCTTTCTGCCGGCGCTCTGACAGCCCTGATGCTGATCATCTTCCTG ATGACCTGCTGTCGGAGAGTGAACAGAAGCGAGCCCACACAGCACAACCTGAGAG GCACAGGCAGAGAAGTGTCCGTGACACCTCAGAGCGGCAAGATCATCAGCAGCT GGGAGAGCCACAAGAGCGGCGGAGAAACAAGACTgTAA
서열번호: 4010 G 유전자 (아미노산)	MVPQALLFVPLLVFPLCFGKFPIYTIPDKLGPWSPIDIHHLSCPNNLVVEDEGCTNLSGF SYMELKVGYILAIKVNGFTCTGVVTEAETYTNFVGYVTTTFKRKHFRPTPDACRAAYN WKMAGDPRYEESLHNPYPDYRWLRTVKTTKESLVIISPSVADLDPYDRSLHSRVFPSG KCSGVAVSSTYCSTNHDYTIWMPENPRLGMSCDIFTNSRGKRASKGSETCGFVDERG LYKSLKGACKLKLCGVLGLRLMDGTWVSMQTSNETKWCPPDKLVNLHDFRSDEIEHL VVEELVRKREECLDALESIMTTKSVSFRRLSHLRKLVPGFGKAYTIFNKTLMEADAHYK SVRTWNEILPSKGCLRVGGRCHPHVNGVFFNGIILGPDGNVLIPEMQSSLLQQHMELL ESSVIPLVHPLADPSTVFKDGDEAEDFVEVHLPDVHNQVSGVDLGLPNWGKYVLLSAG ALTALMLIIFLMTCCRRVNRSEPTQHNLRGTGREVSVTPQSGKIISSWESHKSGGETRL

[0215]

[0216] 특정 실시형태에서, 재조합 광견병 바이러스 게놈은 서열번호: 4001에 제시된 핵산 서열과 약 60%, 약 65%, 약 70%, 약 75%, 약 80%, 약 81%, 약 82%, 약 83%, 약 84%, 약 85%, 약 86%, 약 87%, 약 88%, 약 89%, 약 90%, 약 91%, 약 92%, 약 93%, 약 94%, 약 95%, 약 96%, 약 97%, 약 98%, 약 99% 동일한 핵산 서열을 갖는 N 유전자를 포함한다. 특정 실시형태에서, 재조합 광견병 바이러스 게놈은 서열번호: 4001에 제시된 핵산 서열과 적어도 60%, 적어도 65%, 약 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% 동일한 핵산 서열을 갖는 N 유전자를 포함한다. 특정 실시형태에서, 재조합 광견병 바이러스 게놈은 서열번호: 4001에 제시된 핵산 서열을 포함하는 N 유전자를 포함한다. 특정 실시형태에서, 재조합 광견병 바이러스 게놈은 서열번호: 4001에 제시된 핵산 서열로 구성된 N 유전자를 포함한다. 특정 실시형태에서, N 유전자는 서열번호: 4002에 제시된 아미노산 서열과 약 80%, 약 81%, 약 82%, 약 83%, 약 84%, 약 85%, 약 86%, 약 87%, 약 88%, 약 89%, 약 90%, 약 91%, 약 92%, 약 93%, 약 94%, 약 95%, 약 96%, 약 97%, 약 98%, 약 99% 동일한 아미노산을 인코딩한다. 특정 실시형태에서, N 유전자는 서열번호: 4002에 제시된 아미노산 서열과 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% 동일한 아미노산 서열을 인코딩한다. 특정 실시형태에서, N 유전자는 서열번호: 4002에 제시된 아미노산 서열을 포함하는 아미노산 서열을 인코딩한다. 특정 실시형태에서, N 유전자는 서열번호: 4002에 제시된 아미노산 서열로 구성된 아미노산 서열을 인코딩한다.

[0217] 특정 실시형태에서, 재조합 광견병 바이러스 게놈은 서열번호: 4003에 제시된 핵산 서열과 약 60%, 약 65%, 약 70%, 약 75%, 약 80%, 약 81%, 약 82%, 약 83%, 약 84%, 약 85%, 약 86%, 약 87%, 약 88%, 약 89%, 약 90%, 약 91%, 약 92%, 약 93%, 약 94%, 약 95%, 약 96%, 약 97%, 약 98%, 약 99% 동일한 핵산 서열을 갖는 L 유전자를 포함한다. 특정 실시형태에서, 재조합 광견병 바이러스 게놈은 서열번호: 4003에 제시된 핵산 서열과 적어도 60%, 적어도 65%, 약 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%, 적

어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99% 동일한 핵산 서열을 갖는 L 유전자를 포함한다. 특정 실시형태에서, 재조합 광견병 바이러스 게놈은 서열번호: 4003에 제시된 핵산 서열을 포함하는 L 유전자를 포함한다. 특정 실시형태에서, 재조합 광견병 바이러스 게놈은 서열번호: 4003에 제시된 핵산 서열로 구성된 L 유전자를 포함한다. 특정 실시형태에서, L 유전자는 서열번호: 4004에 제시된 아미노산 서열과 약 80%, 약 81%, 약 82%, 약 83%, 약 84%, 약 85%, 약 86%, 약 87%, 약 88%, 약 89%, 약 90%, 약 91%, 약 92%, 약 93%, 약 94%, 약 95%, 약 96%, 약 97%, 약 98%, 약 99% 동일한 아미노산을 인코딩한다. 특정 실시형태에서, L 유전자는 서열번호: 4004에 제시된 아미노산 서열과 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% 동일한 아미노산 서열을 인코딩한다. 특정 실시형태에서, L 유전자는 서열번호: 4004에 제시된 아미노산 서열을 포함하는 아미노산 서열을 인코딩한다. 특정 실시형태에서, L 유전자는 서열번호: 4004에 제시된 아미노산 서열로 구성된 아미노산 서열을 인코딩한다.

[0218] 특정 실시형태에서, 재조합 광견병 바이러스 게놈은 서열번호: 4005에 제시된 핵산 서열과 약 60%, 약 65%, 약 70%, 약 75%, 약 80%, 약 81%, 약 82%, 약 83%, 약 84%, 약 85%, 약 86%, 약 87%, 약 88%, 약 89%, 약 90%, 약 91%, 약 92%, 약 93%, 약 94%, 약 95%, 약 96%, 약 97%, 약 98%, 약 99% 동일한 핵산 서열을 갖는 M 유전자를 포함한다. 특정 실시형태에서, 재조합 광견병 바이러스 게놈은 서열번호: 4005에 제시된 핵산 서열과 적어도 60%, 적어도 65%, 약 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% 동일한 핵산 서열을 갖는 M 유전자를 포함한다. 특정 실시형태에서, 재조합 광견병 바이러스 게놈은 서열번호: 4005에 제시된 핵산 서열을 포함하는 M 유전자를 포함한다. 특정 실시형태에서, 재조합 광견병 바이러스 게놈은 서열번호: 4005에 제시된 핵산 서열로 구성된 M 유전자를 포함한다. 특정 실시형태에서, M 유전자는 서열번호: 4006에 제시된 아미노산 서열과 약 80%, 약 81%, 약 82%, 약 83%, 약 84%, 약 85%, 약 86%, 약 87%, 약 88%, 약 89%, 약 90%, 약 91%, 약 92%, 약 93%, 약 94%, 약 95%, 약 96%, 약 97%, 약 98%, 약 99% 동일한 아미노산을 인코딩한다. 특정 실시형태에서, M 유전자는 서열번호: 4006에 제시된 아미노산 서열과 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% 동일한 아미노산 서열을 인코딩한다. 특정 실시형태에서, M 유전자는 서열번호: 4006에 제시된 아미노산 서열을 포함하는 아미노산 서열을 인코딩한다. 특정 실시형태에서, M 유전자는 서열번호: 4006에 제시된 아미노산 서열로 구성된 아미노산 서열을 인코딩한다.

[0219] 특정 실시형태에서, 재조합 광견병 바이러스 게놈은 서열번호: 4007에 제시된 핵산 서열과 약 60%, 약 65%, 약 70%, 약 75%, 약 80%, 약 81%, 약 82%, 약 83%, 약 84%, 약 85%, 약 86%, 약 87%, 약 88%, 약 89%, 약 90%, 약 91%, 약 92%, 약 93%, 약 94%, 약 95%, 약 96%, 약 97%, 약 98%, 약 99% 동일한 핵산 서열을 갖는 P 유전자를 포함한다. 특정 실시형태에서, 재조합 광견병 바이러스 게놈은 서열번호: 4007에 제시된 핵산 서열과 적어도 60%, 적어도 65%, 약 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% 동일한 핵산 서열을 갖는 P 유전자를 포함한다. 특정 실시형태에서, 재조합 광견병 바이러스 게놈은 서열번호: 4007에 제시된 핵산 서열을 포함하는 P 유전자를 포함한다. 특정 실시형태에서, 재조합 광견병 바이러스 게놈은 서열번호: 4007에 제시된 핵산 서열로 구성된 P 유전자를 포함한다. 특정 실시형태에서, P 유전자는 서열번호: 4008에 제시된 아미노산 서열과 약 80%, 약 81%, 약 82%, 약 83%, 약 84%, 약 85%, 약 86%, 약 87%, 약 88%, 약 89%, 약 90%, 약 91%, 약 92%, 약 93%, 약 94%, 약 95%, 약 96%, 약 97%, 약 98%, 약 99% 동일한 아미노산을 인코딩한다. 특정 실시형태에서, P 유전자는 서열번호: 4008에 제시된 아미노산 서열과 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% 동일한 아미노산 서열을 인코딩한다. 특정 실시형태에서, P 유전자는 서열번호: 4008에 제시된 아미노산 서열을 포함하는 아미노산 서열을 인코딩한다. 특정 실시형태에서, P 유전자는 서열번호: 4008에 제시된 아미노산 서열로 구성된 아미노산 서열을 인코딩한다.

[0220] 특정 실시형태에서, 재조합 광견병 바이러스 게놈은 서열번호: 4009에 제시된 핵산 서열과 약 60%, 약 65%, 약 70%, 약 75%, 약 80%, 약 81%, 약 82%, 약 83%, 약 84%, 약 85%, 약 86%, 약 87%, 약 88%, 약 89%, 약 90%, 약

91%, 약 92%, 약 93%, 약 94%, 약 95%, 약 96%, 약 97%, 약 98%, 약 99% 동일한 핵산 서열을 갖는 G 유전자를 포함한다. 특정 실시형태에서, 재조합 광견병 바이러스 게놈은 서열번호: 4009에 제시된 핵산 서열과 적어도 60%, 적어도 65%, 약 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% 동일한 핵산 서열을 갖는 G 유전자를 포함한다. 특정 실시형태에서, 재조합 광견병 바이러스 게놈은 서열번호: 4009에 제시된 핵산 서열을 포함하는 G 유전자를 포함한다. 특정 실시형태에서, 재조합 광견병 바이러스 게놈은 서열번호: 4009에 제시된 핵산 서열로 구성된 G 유전자를 포함한다. 특정 실시형태에서, G 유전자는 서열번호: 4010에 제시된 아미노산 서열과 약 80%, 약 81%, 약 82%, 약 83%, 약 84%, 약 85%, 약 86%, 약 87%, 약 88%, 약 89%, 약 90%, 약 91%, 약 92%, 약 93%, 약 94%, 약 95%, 약 96%, 약 97%, 약 98%, 약 99% 동일한 아미노산을 인코딩한다. 특정 실시형태에서, G 유전자는 서열번호: 4010에 제시된 아미노산 서열과 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% 동일한 아미노산 서열을 인코딩한다. 특정 실시형태에서, G 유전자는 서열번호: 4010에 제시된 아미노산 서열을 포함하는 아미노산 서열을 인코딩한다. 특정 실시형태에서, G 유전자는 서열번호: 4010에 제시된 아미노산 서열로 구성된 아미노산 서열을 인코딩한다.

[0221] 본 개시내용의 재조합 광견병 바이러스 게놈 내에 포함된 각각의 유전자는 전사 조절 요소에 작동가능하게 연결될 수 있다. 유전자가 단일 발현 카세트 상에 연결된 특정 실시형태에서, 단일 전사 조절 요소는 유전자의 발현을 제어할 수 있다. 특정 실시형태에서, 각각의 유전자는 별개의 전사 조절 요소에 작동가능하게 연결된다. 특정 실시형태에서, 각각의 유전자에 대한 전사 조절 요소는 동일할 수 있다. 특정 실시형태에서, 각각의 유전자에 대한 전사 조절 요소는 상이할 수 있다.

[0222] 특정 실시형태에서, 각각의 유전자는 전사 조절 요소에 작동가능하게 연결되어 있으며, 여기서 전사 조절 요소는 그에 작동가능하게 연결된 유전자의 발현을 제어할 수 있다. 특정 실시형태에서, 전사 조절 요소는 전사 개시 신호를 포함한다. 전사 개시 신호는 광견병 바이러스에 대해 내인성 또는 외인성일 수 있다. 특정 실시형태에서, 전사 개시 신호는 합성 전사 개시 신호이다. 특정 실시형태에서, 이식유전자를 인코딩하는 핵산은 전사 종결 폴리아데닐화 신호에 추가로 작동가능하게 연결된다. 전사 종결 폴리아데닐화 신호는 광견병 바이러스에 대해 내인성 또는 외인성일 수 있다. 특정 실시형태에서, 전사 종결 폴리아데닐화 신호는 합성 전사 종결 폴리아데닐화 신호이다. 적합한 전사 개시 신호 및 전사 종결 폴리아데닐라톤 신호의 예는 당업자에게 공지되어 있으며, 예를 들어, Albertini 등, *Adv. Virus. Res.* (2011) 79: 1-22; Ogino 및 Green, *Viruses* (2019) 11(6): 504; Ogino 등, *Nucl. Acids. Res.* (2019) 47(1): 299-309; 및 Ogino 및 Green, *Front. Microbiol.* (2019) 10: 1490에 기술되어 있으며, 그 개시내용은 그 전체가 참조로 본 명세서에 포함된다.

[0223] 특정 실시형태에서, 재조합 광견병 바이러스는 복제 불능이다. 본 명세서에 사용된 바와 같은 용어 "복제 불능"은 감염 후 숙주 세포에서 전체 복제 주기를 완료할 수 없는 바이러스를 지칭한다. 일반적으로 바이러스는 바이러스 복제에 필요한 바이러스 유전자를 결실 또는 비활성화를 통해 복제 불능으로 될 수 있다. 특정 실시형태에서, 재조합 광견병 바이러스 게놈에서 P 유전자, L 유전자, M 유전자, N 유전자 및 G 유전자 중 임의의 하나 이상의 결실 또는 불활성화를 통해 재조합 광견병 바이러스가 복제 불능으로 된다. 다른 실시형태에서, 재조합 광견병 바이러스는 복제 결핍이다. 본 명세서에 사용된 바와 같은 용어 "복제 결핍"은 야생형 또는 변형되지 않은 바이러스에 비해 덜 효율적으로 감염 후 숙주 세포에서 전체 복제 주기를 완료할 수 있는 변형된 바이러스를 지칭한다.

[0224] **C. 가이드 RNA & 이를 인코딩하는 재조합 재조합 광견병 바이러스 게놈**

[0225] 일 양태에서, 개시내용은 5' 말단 및 3' 말단을 포함하는 제1 가이드 RNA(gRNA)를 인코딩하는 핵산; 및 제1 gRNA를 인코딩하는 핵산의 5' 말단의 상류 또는 제1 gRNA를 인코딩하는 핵산의 3' 말단의 하류의 둘 중 하나에 위치된 제1 전달 RNA(tRNA)를 인코딩하는 핵산을 포함하는, 재조합 광견병 바이러스 게놈을 제공한다.

[0226] 특정 실시형태에서, 재조합 광견병 바이러스 게놈은 제2 tRNA를 인코딩하는 핵산을 포함한다. 특정 실시형태에서, 재조합 광견병 바이러스 게놈은 제3 tRNA를 인코딩하는 핵산을 포함한다. 특정 실시형태에서, 재조합 광견병 바이러스 게놈은 제4 tRNA를 인코딩하는 핵산을 포함한다. 특정 실시형태에서, 재조합 광견병 바이러스 게놈은 제5 tRNA를 인코딩하는 핵산을 포함한다.

[0227] 특정 실시형태에서, 제1 tRNA를 인코딩하는 핵산은 제1 gRNA를 인코딩하는 핵산의 5' 말단의 상류에 위치되고;

제2 tRNA를 인코딩하는 핵산은 제1 gRNA를 인코딩하는 핵산의 3' 말단 하류에 위치된다.

[0228] 특정 실시형태에서, 제1 tRNA의 뉴클레오티드 서열과 제2 tRNA, 제3 tRNA, 제4 tRNA 및/또는 제5 tRNA의 뉴클레오티드 서열은 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 100% 동일하다.

[0229] 특정 실시형태에서, 제1 tRNA와 제2 tRNA, 제3 tRNA, 제4 tRNA 및/또는 제5 tRNA는 동일한 아미노산을 특정한다. 예를 들어, 제1 tRNA와 제2 tRNA는 서로 다른 안티-코돈 루프 서열을 보유하며, 각각의 안티-코돈 루프 서열은 동일한 아미노산에 상응한다(예를 들어, 5' GGC 3'을 포함하는 안티-코돈 루프 서열을 갖는 제1 tRNA는 Ala를 특정하고, 5' AGC 3'을 포함하는 안티-코돈 루프 서열을 갖는 제2 tRNA 또한 Ala를 특정함).

[0230] 특정 실시형태에서, 제1 tRNA와 제2 tRNA, 제3 tRNA, 제4 tRNA 및/또는 제5 tRNA는 서로 다른 아미노산을 특정한다. 예를 들어, 제1 tRNA와 제2 tRNA는 서로 다른 안티-코돈 루프 서열을 보유하며, 각각의 안티-코돈 루프 서열은 서로 다른 아미노산에 상응한다(예를 들어, 5' GGC 3'을 포함하는 안티-코돈 루프 서열을 갖는 제1 tRNA는 Ala를 특정하고, 5' AAA 3'을 포함하는 안티-코돈 루프 서열을 갖는 제2 tRNA는 Phe를 특정함).

[0231] 특정 실시형태에서, 재조합 광견병 바이러스 게놈은 제1 tRNA, 제2 tRNA, 제3 tRNA, 제4 tRNA 및/또는 제5 tRNA를 인코딩하는 2개 이상의 핵산을 포함한다. 특정 실시형태에서, 재조합 광견병 바이러스 게놈은 제1 tRNA, 제2 tRNA, 제3 tRNA, 제4 tRNA 및/또는 제5 tRNA를 인코딩하는 2개의 핵산을 포함한다. 특정 실시형태에서, 재조합 광견병 바이러스 게놈은 제1 tRNA, 제2 tRNA, 제3 tRNA, 제4 tRNA 및/또는 제5 tRNA를 인코딩하는 3개의 핵산을 포함한다. 특정 실시형태에서, 재조합 광견병 바이러스 게놈은 제1 tRNA, 제2 tRNA, 제3 tRNA, 제4 tRNA 및/또는 제5 tRNA를 인코딩하는 4개의 핵산을 포함한다. 특정 실시형태에서, 재조합 광견병 바이러스 게놈은 제1 tRNA, 제2 tRNA, 제3 tRNA, 제4 tRNA 및/또는 제5 tRNA를 인코딩하는 5개의 핵산을 포함한다.

[0232] 특정 실시형태에서, 재조합 광견병 바이러스 게놈은 제2 gRNA, 제3 gRNA, 제4 gRNA 및/또는 제5 gRNA를 인코딩하는 핵산을 포함한다.

[0233] 특정 실시형태에서, 2개 이상의 핵산은 동일한 gRNA를 인코딩한다. 특정 실시형태에서, 2개 이상의 핵산은 적어도 하나의 상이한 gRNA를 인코딩한다. 특정 실시형태에서, 제1 gRNA의 뉴클레오티드 서열 및 제2 gRNA, 제3 gRNA, 제4 gRNA 및/또는 제5 gRNA의 뉴클레오티드 서열은 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 100% 동일하다.

[0234] 특정 실시형태에서, 제1 gRNA와 제2 gRNA, 제3 gRNA, 제4 gRNA 및/또는 제5 gRNA는 동일한 표적 핵산 서열에 특이적으로 혼성화된다. 특정 실시형태에서, 제1 gRNA와 제2 gRNA, 제3 gRNA, 제4 gRNA 및/또는 제5 gRNA는 상이한 표적 핵산 서열에 특이적으로 혼성화된다.

[0235] 특정 실시형태에서, 제1 tRNA, 제2 tRNA, 제3 tRNA, 제4 tRNA 및/또는 제5 tRNA는 각각 tRNA-ala, tRNA-arg, tRNA-asn, tRNA-asp, tRNA-cys, tRNA-gln, tRNA-gly, tRNA-his, tRNA-ile, tRNA-leu, tRNA-lys, tRNA-met, tRNA-phe, tRNA-pro, tRNA-pyl, tRNA-sec, tRNA-ser, tRNA-thr, tRNA-trp, tRNA-tyr, 및 tRNA-val로 구성된 군으로부터 선택된다.

[0236] 특정 실시형태에서, 재조합 광견병 바이러스 게놈은 음성-가닥 RNA 바이러스 유전자를 인코딩하는 핵산을 포함한다.

[0237] 특정 실시형태에서, 재조합 광견병 바이러스 게놈은 이식유전자(예를 들어, 핵염기 편집기)를 인코딩하는 핵산을 포함한다.

[0238] 특정 실시형태에서, 제1 gRNA를 인코딩하는 핵산 및 제1 tRNA를 인코딩하는 핵산은 각각 음성-가닥 RNA 바이러스 유전자를 인코딩하는 2개의 핵산 사이에 위치된다.

[0239] 특정 실시형태에서, 제1 gRNA를 인코딩하는 핵산 및 제1 tRNA를 인코딩하는 핵산은 각각 이식유전자를 인코딩하는 2개의 핵산 사이에 위치된다.

[0240] 특정 실시형태에서, 제1 gRNA를 인코딩하는 핵산 및 제1 tRNA를 인코딩하는 핵산은 음성-가닥 RNA 바이러스 유전자를 인코딩하는 핵산과 이식유전자를 인코딩하는 핵산 사이에 위치된다.

[0241] 특정 실시형태에서, 재조합 광견병 바이러스 게놈은 5'에서 3'으로 음성-가닥 RNA 바이러스 전사 개시 신호, tRNA를 인코딩하는 핵산, gRNA를 인코딩하는 핵산, 그리고 전사 종료 폴리아데닐화 신호를 포함하는 gRNA 발현 카세트를 포함한다.

[0242] 특정 실시형태에서, 재조합 광견병 바이러스 게놈은 5'에서 3'으로 음성-가닥 RNA 바이러스 전사 개시 신호, 제1 tRNA를 인코딩하는 핵산, 제1 gRNA를 인코딩하는 핵산, 제2 tRNA를 인코딩하는 핵산, 그리고 전사 종결 폴리아데닐화 신호를 포함하는 gRNA 발현 카세트를 포함한다.

[0243] 특정 실시형태에서, 재조합 광견병 바이러스 게놈은 5'에서 3'으로 음성-가닥 RNA 바이러스 전사 개시 신호, 제1 tRNA를 인코딩하는 핵산, 제1 gRNA를 인코딩하는 핵산, 제2 tRNA를 인코딩하는 핵산, 제2 gRNA를 인코딩하는 핵산, 그리고 전사 종결 폴리아데닐화 신호를 포함하는 gRNA 발현 카세트를 포함한다.

[0244] 특정 실시형태에서, 재조합 광견병 바이러스 게놈은 5'에서 3'으로 음성-가닥 RNA 바이러스 전사 개시 신호, 제1 tRNA를 인코딩하는 핵산, 제1 gRNA를 인코딩하는 핵산, 제2 tRNA를 인코딩하는 핵산, 제2 gRNA를 인코딩하는 핵산, 그리고 전사 종결 폴리아데닐화 신호를 포함하는 gRNA 발현 카세트를 포함한다.

[0245] 특정 실시형태에서, 재조합 광견병 바이러스 게놈은 5'에서 3'으로 음성-가닥 RNA 바이러스 전사 개시 신호, 제1 tRNA를 인코딩하는 핵산, 제1 gRNA를 인코딩하는 핵산, 제2 tRNA를 인코딩하는 핵산, 제2 gRNA를 인코딩하는 핵산, 제3 tRNA를 인코딩하는 핵산, 및 전사 종결 폴리아데닐화 신호를 포함하는 gRNA 발현 카세트를 포함한다.

[0246] gRNA 발현 카세트의 특정 실시형태에서, 제1 tRNA, 제2 tRNA 및/또는 제3 tRNA를 인코딩하는 핵산은 동일하다. gRNA 발현 카세트의 특정 실시형태에서, 제1 tRNA, 제2 tRNA 및/또는 제3 tRNA를 인코딩하는 핵산은 상이하다. gRNA 발현 카세트의 특정 실시형태에서, 제1 tRNA의 뉴클레오티드 서열과 제2 tRNA의 뉴클레오티드 서열은 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 100% 동일하다. gRNA 발현 카세트의 특정 실시형태에서, 제1 tRNA와 제2 tRNA는 동일한 아미노산을 특정한다. gRNA 발현 카세트의 특정 실시형태에서, 제1 tRNA와 제2 tRNA는 상이한 아미노산을 특정한다. gRNA 발현 카세트의 특정 실시형태에서, 제1 gRNA 및/또는 제2 gRNA를 인코딩하는 핵산은 동일하다. gRNA 발현 카세트의 특정 실시형태에서, 제1 gRNA 및/또는 제2 gRNA를 인코딩하는 핵산은 상이하다. gRNA 발현 카세트의 특정 실시형태에서, 제1 gRNA의 뉴클레오티드 서열과 제2 gRNA의 뉴클레오티드 서열은 적어도 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 100% 동일하다. gRNA 발현 카세트의 특정 실시형태에서, 제1 gRNA와 제2 gRNA는 동일한 표적 핵산 서열에 특이적으로 혼성화된다. gRNA 발현 카세트의 특정 실시형태에서, 제1 gRNA와 제2 gRNA는 상이한 표적 핵산 서열에 특이적으로 혼성화된다. gRNA 발현 카세트의 특정 실시형태에서, 전사 종결 폴리아데닐화 신호는 내인성 전사 종결 폴리아데닐화 신호를 포함한다. gRNA 발현 카세트의 특정 실시형태에서, 전사 종결 폴리아데닐화 신호는 이종 전사 종결 폴리아데닐화 신호를 포함한다.

[0247] **D. 치료적 이식유전자**

[0248] 특정 실시형태에서, 본 개시내용의 재조합 광견병 바이러스 게놈은 치료적 이식유전자를 포함하는 핵산을 인코딩한다. 본 명세서에서 사용된 바와 같은 용어 "치료적"은 치료 및/또는 예방을 지칭한다. 본 명세서에 사용된 바와 같은 용어 "치료적 이식유전자"은 필요한 대상체에 대한 치료 및/또는 예방을 유효하게 할 수 있는 이식유전자 생성물을 인코딩하는 이식유전자를 지칭한다. 특정 실시형태에서, 치료적 효과는 대상체가 겪고 있는 질환 상태의 억제, 완화 또는 근절에 의해 달성된다. 치료적 이식유전자는 필요한 대상체에서 치료 및/또는 예방을 유효하게 하여, 대상체에서 질환 상태의 억제, 완화 또는 근절을 초래할 수 있는 임의의 치료제를 인코딩할 수 있다. 특정 실시형태에서, 치료적 이식유전자는 일단 처리되면, 예를 들어 세포에서 처리되면 이를 필요로 하는 대상체에 대한 치료 및/또는 예방을 유효하게 할 수 있는 이식유전자 생성물의 전구체를 인코딩한다.

[0249] 특정 실시형태에서, 치료적 이식유전자를 인코딩하는 핵산은 약 300bp, 약 400bp, 약 500bp, 약 600bp, 약 700bp, 약 800bp, 약 900bp, 약 1,000bp, 약 1,100bp, 약 1,200bp, 약 1,300bp, 약 1,400bp, 약 1,500bp, 약 1,600bp, 약 1,700bp, 약 1,800bp, 약 1,900bp, 약 2,000bp, 약 2,100bp, 약 2,200bp, 약 2,300bp, 약 2,400bp, 약 2,500bp, 약 2,600bp, 약 2,700bp, 약 2,800bp, 약 2,900bp, 또는 약 3,000bp보다 크다.

[0250] 특정 실시형태에서, 치료적 이식유전자를 인코딩하는 핵산은 약 300bp보다 크다(예를 들어, 치료적 이식유전자는 약 350bp, 약 400bp, 약 450bp, 약 500bp, 약 550bp, 약 600bp, 또는 약 650bp이다). 특정 실시형태에서, 치료적 이식유전자를 인코딩하는 핵산은 약 650bp보다 크다(예를 들어, 치료적 이식유전자는 약 700bp, 약 750bp, 약 800bp, 약 850bp, 약 900bp, 약 950bp, 또는 약 1,000bp이다). 특정 실시형태에서, 치료적 이식유전자를 인코딩하는 핵산은 약 1,000bp보다 크다(예를 들어, 치료적 이식유전자는 약 1,500bp, 약 2,000bp, 약 2,500bp, 또는 약 3,000bp이다). 특정 실시형태에서, 치료적 이식유전자를 인코딩하는 핵산은 약 3,000bp보다 크다(예를 들어, 치료적 이식유전자는 약 3,500bp, 약 4,000bp, 또는 약 4,500bp이다).

[0251] 특정 실시형태에서, 치료적 이식유전자를 인코딩하는 핵산은 약 4,500bp보다 크다(예를 들어, 치료적 이식유전

자는 약 5,000bp, 약 5,500bp, 약 6,000bp, 약 6,500bp, 약 7,000bp, 약 7,500bp, 약 8,000bp, 또는 약 8,500bp이다).

[0252] 특정 실시형태에서, 치료적 이식유전자를 인코딩하는 핵산은 약 8,500bp보다 크다(예를 들어, 치료적 이식유전자는 약 9,000bp, 약 9,500bp, 또는 약 10,000bp이다).

[0253] 특정 실시형태에서, 치료적 이식유전자를 인코딩하는 핵산은 약 10,000bp보다 크다(예를 들어, 치료적 이식유전자는 약 10,500bp, 약 11,000bp, 약 11,500bp, 약 12,000bp, 약 12,500bp, 약 13,000bp, 약 13,500bp, 약 14,000bp, 약 14,500bp 또는 약 15,000bp이다).

[0254] 특정 실시형태에서, 치료적 이식유전자를 인코딩하는 핵산은 약 4,000bp 내지 약 6,000bp이다(예를 들어, 치료적 이식유전자는 약 4,000bp, 약 4,500bp, 약 5,000bp, 약 5,500bp 또는 약 6,000bp이다).

[0255] 특정 실시형태에서, 치료적 이식유전자는 치료적 핵산을 인코딩한다. 치료적 이식유전자는 당업계에 공지된 임의의 치료적 핵산, 예를 들어 제한 없이 임의의 안티센스 RNA(단일-가닥 RNA), 임의의 작은 간섭 RNA(이중-가닥 RNA), 임의의 RNA 압타머 및/또는 임의의 메신저 RNA(mRNA)를 인코딩할 수 있다. 예를 들어, 치료적 이식유전자는 제한 없이, miRNA, miRNA 모방체, siRNA, shRNA, gRNA, 긴 비코딩 RNA, 인핸서 RNA, RNA 압타자임, RNA 압타머, 안타고미R 및/또는 또는 합성 RNA를 인코딩할 수 있다. 특정 실시형태에서, 치료적 핵산은 RNA 결합 부위, 예를 들어 miRNA 결합 부위일 수 있다. 다양한 다른 유형의 치료적 핵산이 당업자에게 공지되어 있다.

[0256] 특정 실시형태에서, 치료적 이식유전자는 치료적 폴리펩티드를 인코딩한다. 치료적 이식유전자는 당업계에 공지된 임의의 치료적 폴리펩티드, 예를 들어 제한 없이 결핍되거나 비정상적인 단백질을 대체할 수 있는 치료적 폴리펩티드; 기존 경로를 증대시킬 수 있는 치료적 폴리펩티드; 신규한 기능 또는 활성(예를 들어, 그 결핍으로부터 고통을 받는 대상체에게 유익한 신규한 기능 또는 활성)을 제공할 수 있는 치료적 폴리펩티드; 분자 또는 유기체(예를 들어, 표적 세포를 수용하는 유기체와 다른 유기체)를 방해하는 치료적 폴리펩티드; 및/또는 다른 화합물 또는 단백질(예를 들어, 방사성핵종, 세포독성 약물 및/또는 효과기 단백질)을 전달하는 치료적 폴리펩티드를 인코딩할 수 있다. 예를 들어, 치료적 이식유전자는 제한 없이, 핵산 변형 단백질(예를 들어, 아데닌 또는 시티딘 염기 편집기) 또는 시스템, 항체 또는 항체-기반 약물, 항응고제, 혈액 인자, 뼈 형태발생 단백질, 조작된 단백질 스캐폴드, 효소, Fc 융합 단백질, 성장 인자, 호르몬, 인터페론, 인터루킨 및/또는 혈전용해제를 인코딩할 수 있다. 다양한 다른 유형의 치료적 폴리펩티드가 당업자에게 공지되어 있다.

[0257] 특정 실시형태에서, 치료적 이식유전자는 핵산 편집 시스템 또는 이의 구성요소를 인코딩한다. 일부 실시형태에서 치료적 이식유전자는 핵산 결합 단백질(예를 들어, 아연 핑거, TALE, 또는 핵산 프로그램가능 핵산 결합 단백질, 예컨대 Cas9)을 포함하는 단백질을 인코딩한다. 일부 실시형태에서, 핵산 편집 시스템 구성요소는 핵산 프로그램가능 핵산 결합 단백질(예를 들어, Cas9)이다. 일부 실시형태에서, 핵산 편집 시스템 구성요소는 가이드 RNA(gRNA)이다.

[0258] 일부 실시형태에서, 치료적 이식유전자는 CRISPR 시스템을 인코딩한다. 일부 실시형태에서, CRISPR 시스템은 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인 및 핵염기 편집 도메인을 포함하는 핵염기 편집기를 포함한다. 일부 실시형태에서, 핵염기 편집 도메인은 아데노신 데아미나제, 시티딘 데아미나제, 시토신 데아미나제, 또는 이의 기능적 변이체(예를 들어, DNA 또는 RNA와 같은 핵산 분자에서 핵염기를 탈아미노화할 수 있는 기능적 변이체)이다. 일부 실시형태에서, 핵염기 편집 도메인은 아데노신 데아미나제이다. 일부 실시형태에서, 아데노신 데아미나제는 ABE7.10이다. 일부 실시형태에서, 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인은 Cas9 폴리펩티드, Cas12 폴리펩티드, 또는 이의 기능적 변이체이다. 일부 실시형태에서, CRISPR 시스템은 가이드 RNA(gRNA) 또는 gRNA를 인코딩하는 핵산을 추가로 포함한다.

[0259] 일부 실시형태에서, 치료적 이식유전자는 핵염기 변형 단백질(예를 들어, 염기 편집기 단백질)을 인코딩한다. 일부 실시형태에서, 치료적 이식유전자는 아데노신 염기 편집기(예를 들어, ABE7.10)를 인코딩한다. 일부 실시형태에서 치료적 이식유전자는 시티딘 염기 편집기를 인코딩한다. 일부 실시형태에서 치료적 이식유전자는 DNA 또는 RNA에서 시토신을 탈아미노화할 수 있는 시토신 염기 편집기를 인코딩한다.

[0260] 특정 실시형태에서, 치료적 이식유전자는 핵산 편집 시스템, 예를 들어, 본 명세서에 추가로 기술된 염기 편집기 시스템을 인코딩한다.

[0261] 본 명세서에 기재된 본 개시내용의 재조합 광견병 바이러스 게놈이 치료적 이식유전자를 포함하는 핵산을 인코딩하며, 여기서 치료적 이식유전자는 치료적 폴리펩티드 및/또는 치료적 핵산을 인코딩하며, 예를 들어, 특정 실시형태에서, 치료적 이식유전자는 치료적 폴리펩티드와 치료적 핵산의 조합을 인코딩한다는 것이 당업자에게

쉽게 명백할 것이다. 특정 실시형태에서, 치료적 이식유전자는 하나 이상의 치료적 폴리펩티드를 인코딩한다. 특정 실시형태에서, 치료적 이식유전자는 하나 이상의 치료적 핵산을 인코딩한다. 특정 실시형태에서, 치료적 이식유전자는 하나 이상의 치료 폴리펩티드와 하나 이상의 치료 핵산의 조합을 인코딩한다. 표적 세포 안으로 치료적 폴리펩티드와 치료적 핵산의 조합의 전달은 당업자에게 공지된 다양한 목적을 제공할 수 있다. 특정 실시형태에서, 치료적 폴리펩티드는 표적 세포로 전달될 수 있으며, 여기서 전달은 특정 다른 세포 유형에 대해 탈표적화된다. 예를 들어, 치료적 이식유전자는 치료적 폴리펩티드 및/또는 치료적 핵산을 인코딩할 수 있고, miRNA 결합 부위도 포함할 수 있다. miRNA 결합 부위는 세포 유형 탈표적화를 위해 기능할 수 있다. 예를 들어, 간에서 배타적으로 발현되는 miRNA122a는 간세포 탈표적화를 위해 이용될 수 있다. 예를 들어, Dhungel 등, *Molecules* (2018) 23(7): 1500을 참조한다.

[0262] 특정 실시형태에서, 치료적 이식유전자는 하나 이상의 리포터 서열을 추가로 인코딩한다. 리포터 서열은 표적 세포에서 발현될 때 직접적으로 또는 간접적으로 검출가능한 신호를 생성한다. 적합한 리포터 서열의 예는 제한 없이, 형광 단백질(예를 들어, GFP, RFP, YFP), 알칼리성 포스파타제, 티미딘 키나제, 클로람페니콜 아세틸트랜스퍼라제(CAT), 루시퍼라제, β-갈락토시다제(LacZ) 및 β-락타마제를 인코딩하는 서열을 포함한다. 세포 표면 막-결합 단백질, 예를 들어 고친화성 항체가 결합하는 막-결합 단백질, 예를 들어 인플루엔자 혈구응집소 단백질(HA), CD2, CD4, CD8 및 예를 들어, 항원 도메인으로 태그된 막-결합 단백질(예를 들어, HA 태그, FLAG 태그, Myc 태그, 폴리히스티딘 태그)을 포함하여 당업자에게 공지된 기타를 인코딩하는 서열이 리포터 서열로서 적합할 수도 있다.

[0263] 특정 실시형태에서, 치료적 이식유전자는 리포터 유전자 및/또는 선택가능한 마커를 인코딩하지 않는다. 특정 실시형태에서, 치료적 이식유전자는 형광 리포터 단백질(예를 들어, GFP, YFP, RFP, tdTomato)을 인코딩하지 않는다. 특정 실시형태에서, 치료적 이식유전자는 β-갈락토시다제(LacZ)를 인코딩하지 않는다. 특정 실시형태에서, 치료적 이식유전자는 클로람페니콜 아세틸트랜스퍼라제(CAT)를 인코딩하지 않는다.

[0264] 특정 실시형태에서, 치료적 이식유전자는 폴리머라제(예를 들어, DNA 폴리머라제, DNA-지향된 RNA 폴리머라제, RNA-지향된 DNA 폴루머라제(RT), 텔로머라제)를 인코딩하지 않는다.

[0265] 특정 실시형태에서, 치료적 이식유전자는 부위-특이적 재조합효소(예를 들어, Cre, FLP, Hin 또는 Tre 재조합효소)를 인코딩하지 않는다.

[0266] 특정 실시형태에서, 치료적 이식유전자는 바이러스 항원을 인코딩하지 않는다.

[0267] 특정 실시형태에서, 치료적 이식유전자는 향-세포자멸사 단백질(예를 들어 시토크롬 c)을 인코딩하지 않는다.

[0268] 특정 실시형태에서, 치료적 이식유전자는 면역글로불린(예를 들어, 면역글로불린 중쇄 및/또는 경쇄)을 인코딩하지 않는다.

[0269] 특정 실시형태에서, 치료적 이식유전자는 신경전달물질, 신경펩티드, 수용체, 신경 성장 인자, 또는 신경세포 유전자를 인코딩하지 않는다.

[0270] 특정 실시형태에서, 치료적 이식유전자는 치료적 폴리펩티드 및/또는 치료적 핵산을 인코딩하며, 여기서 치료적 폴리펩티드 및/또는 치료적 핵산은 분비된다(예를 들어, 세포로부터 분비된다). 예를 들어, 본 명세서에 기술된 본 개시내용의 재조합 광견병 바이러스 게놈은 표적 세포 안으로 도입될 수 있으며, 여기서 재조합 광견병 바이러스 게놈은 치료적 이식유전자를 포함하는 핵산을 인코딩하고, 치료적 이식유전자는 분비되는 치료적 폴리펩티드 및/또는 치료적 핵산(예를 들어, 분비가능한 치료적 이식유전자 및/또는 분비가능한 치료적 핵산)을 인코딩한다. 발현 시 치료적 폴리펩티드 및/또는 핵산은 표적 세포의 외부로 분비될 수 있다. 특정 실시형태에서, 발현 시 치료적 폴리펩티드 및/또는 핵산은 치료적 폴리펩티드 및/또는 핵산 상에 존재하는 내인성 요소(예를 들어, 세포외 분비를 지시하는 내인성 신호 펩티드)의 덕분에 분비된다. 특정 실시형태에서, 발현 시 치료적 폴리펩티드 및/또는 핵산은 치료적 폴리펩티드 및/또는 핵산 상에 존재하는 외인성 요소(예를 들어, 세포외 분비를 지시하는 외인성 신호 펩티드)의 덕분에 분비된다. 분비가능한 치료적 폴리펩티드 및/또는 핵산의 전달은 특정 질환의 치료에 유용하다. 예를 들어, 리소좀의 대사 기능장애로 인해 발생하는 리소좀 축적 장애(LSD)는 특정 세포외 LSD 효소가 흡수되어 효소-결핍 또는 효소-비정상 세포의 리소좀을 표적화하도록 하는 독특한 교차-교정 특성을 포함한다. 특정 효소의 교차-교정 특성은 효소 대체 요법으로 알려진 승인된 요법의 기초를 형성한다. 예를 들어, Rastall 및 Amalfitano, *Appl. Clin. Genet.* (2015) 8: 157-169를 참조한다.

[0271] 특정 실시형태에서, 본 개시내용의 재조합 광견병 바이러스 게놈은 이식유전자를 인코딩하는 핵산에 작동가능하게 연결된 전사 조절 요소를 포함한다. 전사 조절 요소는 그에 작동가능하게 연결된 이식유전자의 발현(예를 들

어, 인코딩된 치료적 폴리펩티드 및/또는 핵산의 발현)을 제어할 수 있다. 특정 실시형태에서, 전사 조절 요소는 전사 개시 신호를 포함한다. 전사 개시 신호는 광견병 바이러스에 대해 내인성 또는 외인성일 수 있다. 특정 실시형태에서, 전사 개시 신호는 합성 전사 개시 신호이다. 특정 실시형태에서, 이식유전자를 인코딩하는 핵산은 전사 종결 폴리아데닐화 신호에 추가로 작동가능하게 연결된다. 전사 종결 폴리아데닐화 신호는 광견병 바이러스에 대해 내인성 또는 외인성일 수 있다. 특정 실시형태에서, 전사 종결 폴리아데닐화 신호는 합성 전사 종결 폴리아데닐화 신호이다. 적합한 전사 개시 신호 및 전사 종결 폴리아데닐라톤 신호의 예는 당업자에게 공지되어 있고, 예를 들어, Albertini 등, *Adv. Virus. Res.* (2011) 79: 1-22; Ogino 및 Green, *Viruses* (2019) 11(6): 504; 및 Ogino 및 Green, *Front. Microbiol.* (2019) 10: 1490에 기술되어 있으며, 그 개시내용은 그 전체가 본 명세서에 참조로 포함된다.

[0272] 치료적 이식유전자를 포함하는 핵산을 포함하는 본 개시내용의 재조합 광견병 바이러스 게놈은 치료적 이식유전자의 발현을 보조 및/또는 향상시키는 당업자에게 공지된 임의의 요소를 추가로 포함할 수 있다.

[0273] 본 개시내용의 재조합 광견병 바이러스 게놈은 본 명세서에 기재된 방법에 의해 재조합 광견병 바이러스 입자 안으로 통합된다. 특정 실시형태에서, 본 개시내용의 재조합 광견병 바이러스 입자는 광견병 바이러스 당단백질 및 본 명세서에 기재된 치료적 이식유전자를 포함하는 핵산을 포함하는 재조합 광견병 바이러스 게놈을 포함한다. 특정 실시형태에서, 재조합 광견병 바이러스 입자는 광견병 바이러스 당단백질; 및 치료적 이식유전자를 포함하는 핵산을 포함하는 재조합 광견병 바이러스 게놈을 포함하며, 여기서, 게놈은 광견병 바이러스 당단백질을 인코딩하는 내인성 G 유전자를 결한다. 특정 실시형태에서, 재조합 광견병 바이러스 입자는 광견병 바이러스 당단백질; 및 치료적 이식유전자를 포함하는 핵산을 포함하는 재조합 광견병 바이러스 게놈을 포함하며, 여기서, 게놈은 광견병 바이러스 당단백질을 인코딩하는 내인성 G 유전자를 결하고; 여기서 게놈은 광견병 바이러스 폴리머라제를 인코딩하는 내인성 L 유전자를 결한다.

[0274] **E. 핵염기 편집기**

[0275] 특정한 예시적 실시형태에서, 본 명세서에 기재된 방법 및 조성물에 유용한 치료적 이식유전자는 폴리뉴클레오티드의 표적 뉴클레오티드 서열을 편집, 변형 또는 변경하는 핵염기 편집기이다. 본 명세서에 기재된 핵염기 편집기는 전형적으로 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인 및 핵염기 편집 도메인(예를 들어, 아데노신 데아미나제 또는 시티딘 데아미나제)을 포함한다. 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인은 결합된 가이드 폴리뉴클레오티드(*예를 들어*, gRNA)와 연계할 때 표적 폴리뉴클레오티드 서열에 특이적으로 결합할 수 있고 이에 의해 편집하고자 하는 표적 핵산 서열에 염기 편집기를 국소화시킬 수 있다.

[0276] ***폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인***

[0277] 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인은 폴리뉴클레오티드(예를 들어, RNA, DNA)에 결합한다. 염기 편집기의 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인은 그 자체로 하나 이상의 도메인(예를 들어, 하나 이상의 뉴클레아제 도메인)을 포함할 수 있다. 일부 실시형태에서, 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인의 뉴클레아제 도메인은 엔도뉴클레아제 또는 엑소뉴클레아제를 포함할 수 있다. 엔도뉴클레아제는 이중-가닥 핵산의 단일 가닥 또는 이중-가닥 핵산 분자의 두 가닥 모두를 절단할 수 있다. 일부 실시형태에서, 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인의 뉴클레아제 도메인은 표적 폴리뉴클레오티드의 0개, 1개 또는 2개 가닥을 자를 수 있다.

[0278] 염기 편집기 안으로 통합될 수 있는 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인의 비-제한적 예는 CRISPR 단백질-유래된 도메인, 제한 뉴클레아제, 메가뉴클레아제, TAL 뉴클레아제(TALEN) 및 아연 핑거 뉴클레아제(ZFN)를 포함한다. 일부 실시형태에서, 염기 편집기는 결합된 가이드 핵산을 통해 핵산의 CRISPR(즉, 클러스터링된 규칙적으로 간격을 둔 짧은 회문 반복)-매개된 변형 동안 핵산 서열에 결합할 수 있는 천연 또는 변형된 단백질 또는 이의 일부를 포함하는 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인을 포함한다. 이러한 단백질은 본 명세서에서 "CRISPR 단백질"로 지칭된다. 따라서, CRISPR 단백질의 전부 또는 일부를 포함하는 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인을 포함하는 염기 편집기(즉, 염기 편집기의 "CRISPR 단백질-유래된 도메인"으로도 지칭되는, CRISPR 단백질의 전부 또는 일부를 도메인으로 포함하는 염기 편집기)가 본 명세서에 개시되어 있다. 염기 편집기 안으로 통합된 CRISPR 단백질-유래된 도메인은 CRISPR 단백질의 야생형 또는 자연 버전과 비교하여 변형될 수 있다. 예를 들어, 아래에 기술된 바와 같이 CRISPR 단백질-유래된 도메인은 CRISPR 단백질의 야생형 또는 자연 버전에 비해 하나 이상의 돌연변이, 삽입, 결실, 재배열 및/또는 재조합을 포함할 수 있다.

[0279] 본 명세서에서 사용될 수 있는 Cas 단백질은 클래스 1 및 클래스 2를 포함한다. Cas 단백질의 비-제한적인 예는 Cas1, Cas1B, Cas2, Cas3, Cas4, Cas5, Cas5d, Cas5t, Cas5h, Cas5a, Cas6, Cas7, Cas8, Cas9(Csn1 또는 Csx12로도 알려짐), Cas10, Csy1, Csy2, Csy3, Csy4, Cse1, Cse2, Cse3, Cse4, Cse5e, Csc1, Csc2, Csa5, Csn1, Csn2, Csm1, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Csb1, Csb2, Csb3, Csx17, Csx14, Csx10, Csx16, CsaX, Csx3, Csx1, Csx1S, Csf1, Csf2, Cs0, Csf4, Csd1, Csd2, Cst1, Cst2, Csh1, Csh2, Csa1, Csa2, Csa3, Csa4, Csa5, Cas12a/Cpf1, Cas12b/C2c1(예를 들어, 서열번호: 258), Cas12c/C2c3, Cas12d/CasY, Cas12e/CasX, Cas12g, Cas12h, Cas12i, 및 Cas12j/CasΦ, CARF, DinG, 이의 동족체, 또는 이의 변형된 버전을 포함한다. CRISPR 효소는 표적 서열 내 및/또는 표적 서열의 보체 내와 같은 표적 서열에서 가닥 중 하나 또는 둘 모두의 절단을 지시할 수 있다. 예를 들어, CRISPR 효소는 표적 서열의 제1 또는 마지막 뉴클레오티드로부터의 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 50, 100, 200, 500 이상의 염기쌍 내에서 가닥 중 하나 또는 둘 모두의 절단을 지시할 수 있다.

[0280] 돌연변이된 CRISPR 효소가 표적 서열을 함유하는 표적 폴리뉴클레오티드의 가닥 중 하나 또는 둘 모두를 절단하는 능력을 결하도록 상응하는 야생형 효소에 대해 돌연변이된 CRISPR 효소를 인코딩하는 벡터가 사용될 수 있다. Cas 단백질(예를 들어, Cas9, Cas12) 또는 Cas 도메인(예를 들어, Cas9, Cas12)은 야생형 예시적인 Cas 폴리펩티드 또는 Cas 도메인에 적어도 또는 적어도 약 50%, 60%, 70%, 80%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 100% 서열 동일성 및/또는 서열 상동성을 갖는 폴리펩티드 또는 도메인을 지칭할 수 있다. Cas(예를 들어, Cas9, Cas12)는 결실, 삽입, 치환, 변이체, 돌연변이, 융합, 키메라 또는 이의 임의의 조합과 같은 아미노산 변화를 포함할 수 있는 야생형 또는 변형된 형태의 Cas 단백질을 지칭할 수 있다.

[0281] 일부 실시형태에서, 염기 편집기의 CRISPR 단백질-유래된 도메인은 *코리네박테리움 울세란스*(NCBI 참조번호: NC_015683.1, NC_017317.1); *코리네박테리움 디프테리아*(NCBI 참조번호: NC_016782.1, NC_016786.1); *스피로플라스마 시르피디콜라*(NCBI 참조번호: NC_021284.1); *프레보텔라 인터메디아*(NCBI 참조번호: NC_017861.1); *스피로플라스마 타이완넨세*(NCBI 참조번호: NC_021846.1); *스트렙토코커스 이니애*(NCBI 참조번호: NC_021314.1); *벨리엘라 발티카*(NCBI 참조번호: NC_018010.1); *사이크로플렉서스 토퀴스*(NCBI 참조번호: NC_018721.1); *스트렙토코커스 써모필루스*(NCBI 참조번호: YP_820832.1); *리스테리아 이노쿠아*(NCBI 참조번호: NP_472073.1); *캄필로박터 제주니*(NCBI 참조번호: YP_002344900.1); *나이세리아 메닝기티디스*(NCBI 참조번호: YP_002342100.1), *스트렙토코커스 피오게네스* 또는 *스타필로코커스 아우레우스*로부터의 Cas9의 전부 또는 일부를 포함할 수 있다.

[0282] Cas9 뉴클레아제 서열 및 구조는 당업자에게 잘 알려져 있다(예를 들어, "Complete genome sequence of an M1 strain of Streptococcus pyogenes". Ferretti 등, *Proc. Natl. Acad. Sci. U.S.A.* 98:4658-4663(2001); "CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III." Deltcheva E., 등, Nature 471:602-607(2011); 및 "A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity." Jinek M., 등, *Science* 337:816-821(2012)을 참조하며, 이들 각각의 전체 내용은 본 명세서에 참조로 포함된다). Cas9 오솔로그는 S. 피오게네스 및 S. 써모필루스를 포함하지만 이에 제한되지 않는 다양한 종에 기술되었다. 추가의 적합한 Cas9 뉴클레아제 및 서열은 본 개시내용에 기반하여 당업자에게 명백할 것이고, 이러한 Cas9 뉴클레아제 및 서열은, 그 전체 내용이 본 명세서에 참조로 포함된, Chylinski, Rhun, 및 Charpentier, "The tracrRNA and Cas9 families of type II CRISPR-Cas immunity systems" (2013) *RNA Biology* 10:5, 726-737에 개시된 유기체 및 좌위로부터의 Cas9 서열을 포함한다.

[0283] *높은 충실도 Cas9 도메인*

[0284] 개시내용의 일부 양태는 높은 충실도 Cas9 도메인을 제공한다. 높은 충실도 Cas9 도메인은 당업계에 알려져 있고, 예를 들어 Kleinstiver, B.P., 등 "High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects." *Nature* 529, 490-495 (2016); 및 Slaymaker, I.M., 등 "Rationally engineered Cas9 nucleases with improved specificity." *Science* 351, 84-88 (2015)에 기술되어 있으며; 그 각각의 전체 내용은 본 명세서에 참조로 포함된다. 예시적인 높은 충실도 Cas9 도메인은 서열 목록에 서열번호: 1423으로 제공된다. 일부 실시형태에서, 높은 충실도 Cas9 도메인은 상응하는 야생형 Cas9 도메인에 비해 Cas9 도메인과 DNA의 당-인산염 골격 사이의 정전기적 상호작용을 감소시키는 하나 이상의 돌연변이를 포함하는 조작된 Cas9 도메인이다. DNA의 당-인산염 골격과의 정전기적 상호작용이 감소된 높은 충실도 Cas9 도메인은 더 적은 표적-외 효과를 갖는다. 일부 실시형태에서, Cas9 도메인(*예를 들어*, 야생형 Cas9 도메인(서열번호: 223 및 233))은 Cas9 도메인과 DNA의 당-인산염 골격 사이의 회합을 감소시키는 하나 이상의 돌연변이를 포함한다. 일부 실시형태에서, Cas9 도메인은 Cas9 도메인과 DNA의 당-인산염 골격 사이의 회합을 적어도 1%, 적어도 2%, 적어도 3%, 적어도 4%, 적어도 5%, 적어도 10%, 적어도 15%, 적어도 20%, 적어도 25%, 적어도 30%, 적어도 35%, 적어도

40%, 적어도 45%, 적어도 50%, 적어도 55%, 적어도 60%, 적어도 65% 또는 적어도 70% 감소시키는 하나 이상의 돌연변이를 포함한다.

[0285] 일부 실시형태에서, 본 명세서에 제공된 임의의 Cas9 융합 단백질은 D10A, N497X, R661X, Q695X 및/또는 Q926X 돌연변이 중 하나 이상, 또는 본 명세서에 제공된 임의의 아미노산 서열에서 상응하는 돌연변이를 포함하며, 여기서 X는 임의의 아미노산이다. 일부 실시형태에서, 높은 충실도 Cas9 효소는 SpCas9(K855A), eSpCas9(1.1), SpCas9-HF1, 또는 초정밀 Cas9 변이체(HypaCas9)이다. 일부 실시형태에서, 변형된 Cas9 eSpCas9(1.1)은 HNH/RuvC 홈과 비-표적 DNA 가닥 사이의 상호작용을 약화시켜, 가닥 분리를 방지하고 표적-외 부위에서 절단하는 알라닌 치환을 함유한다. 유사하게, SpCas9-HF1은 DNA 인산염 골격과 Cas9의 상호작용을 파괴하는 알라닌 치환을 통해 표적-외 편집을 낮춘다. HypaCas9는 Cas9 교정 및 표적 분별을 증가시키는 REC3 도메인에서의 돌연변이(SpCas9 N692A/M694A/Q695A/H698A)를 함유한다. 3가지 높은 충실도 효소는 모두 야생형 Cas9보다 낮은 표적-외 편집을 생성한다.

[0286] *감소된 배타성을 갖는 Cas9 도메인*

[0287] 전형적으로 *S. 피오게네스*로부터의 Cas9(spCas9)와 같은 Cas9 단백질은 "프로토스페이서 인접 모티프(PAM)" 또는 PAM-유사 모티프가 필요하며, 이는 CRISPR 박테리아 적응성 면역 시스템에서 Cas9 뉴클레아제에 의해 표적화된 DNA 서열이 바로 이어지는 2-6 염기쌍 DNA 서열이다. 특정 핵산 영역을 결합하기 위해 NGG PAM 서열의 존재가 필요하며, 여기서 "NGG"에서의 "N"은 아데노신(A), 티미딘(T) 또는 시토신(C)이고 G는 구아노신이다. 이는 게놈 내에서 원하는 염기를 편집하는 능력을 제한할 수 있다. 일부 실시형태에서, 본 명세서에 제공된 염기 편집 융합 단백질은 정확한 장소, 예를 들어 PAM의 상류에 있는 표적 염기를 포함하는 영역에 위치되어야 할 수 있다. 예를 들어 Komor, A.C., 등, "Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage" *Nature* 533, 420-424 (2016)을 참조하며, 그 전체 내용은 본 명세서에 참고로 포함된다. PAM 서열에 결합할 수 있는 spCas9 단백질에 대한 예시적인 폴리펩티드 서열은 서열 목록에 서열번호: 223, 234, 및 1304-1307로 제공된다. 따라서, 일부 실시형태에서, 본 명세서에 제공된 임의의 융합 단백질은 규범적(*예를 들어*, NGG) PAM 서열을 함유하지 않는 뉴클레오티드 서열에 결합할 수 있는 Cas9 도메인을 함유할 수 있다. 비-규범적 PAM 서열에 결합하는 Cas9 도메인은 당업계에 기술되어 있고 숙련가에게 명백할 것이다. 예를 들어, 비-규범적 PAM 서열에 결합하는 Cas9 도메인은 Kleinstiver, B. P., 등, "Engineered CRISPR-Cas9 nucleases with altered PAM specificities" *Nature* 523, 481-485 (2015); 및 Kleinstiver, B. P., 등, "Broadening the targeting range of *Staphylococcus aureus* CRISPR-Cas9 by modifying PAM recognition" *Nature Biotechnology* 33, 1293-1298 (2015)에 기술되어 있으며; 각각의 전체 내용은 참조로 본 명세서에 포함된다.

[0288] *닉카제*

[0289] 일부 실시형태에서, 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인은 닉카제 도메인을 포함할 수 있다. 본 명세서에서 용어 "닉카제"는 이중나선의 핵산 분자(예를 들어, DNA)에서 두 가닥 중 한 가닥만 절단할 수 있는 뉴클레아제 도메인을 포함하는 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인을 지칭한다. 일부 실시형태에서, 닉카제는 활성 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인 안으로 하나 이상의 돌연변이를 도입함에 의해 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인의 완전히 촉매적으로 활성인(예를 들어, 자연) 형태로부터 유래될 수 있다. 예를 들어, 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인이 Cas9로부터 유래된 닉카제 도메인을 포함하는 경우, Cas9-유래된 닉카제 도메인은 위치 840에 D10A 돌연변이 및 히스티딘을 포함할 수 있다. 이러한 실시형태에서, 잔기 H840은 촉매 활성을 유지하고 이에 의해 핵산 이중나선의 단일 가닥을 절단될 수 있다. 또 다른 예에서, Cas9-유래된 닉카제 도메인은 H840A 돌연변이를 포함할 수 있는 반면, 위치 10에서 아미노산 잔기는 D를 유지한다. 일부 실시형태에서, 닉카제는 닉카제 활성에 필요하지 않은 뉴클레아제 도메인의 전부 또는 일부를 제거함에 의해 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인의 완전히 촉매적으로 활성인(예를 들어, 자연) 형태로부터 유래될 수 있다. 예를 들어, 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인이 Cas9로부터 유래된 닉카제 도메인을 포함하는 경우, Cas9-유래된 닉카제 도메인은 RuvC 도메인 또는 HNH 도메인의 전부 또는 일부의 결실을 포함할 수 있다.

[0290] 일부 실시형태에서, 야생형 Cas9는 다음 아미노산 서열에 상응하거나 이를 포함한다:

[0291] MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLE
ESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINAS
GVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDIL

RVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG
SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPK
HSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEEN
EDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVS
GQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYL
QNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIK
RQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRK
MIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIAR
KKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGN
ELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKY
FDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD(서열번호:223)(단일 밑줄: HNH 도메인, 이중 밑줄: RuvC 도메인).

[0292] 일부 실시형태에서, 닉카제 도메인(예를 들어, Cas9-유래된 닉카제 도메인, Cas12-유래된 닉카제 도메인)을 포함하는 염기 편집기에 의해 절단되는 핵산 이중나선 표적 폴리뉴클레오티드 서열의 가닥은 염기 편집기에 의해 편집되지 않는 가닥이다(즉, 염기 편집기에 의해 절단되는 가닥은 편집되어 지는 염기를 포함하는 가닥의 반대편에 있음). 다른 실시형태에서, 닉카제 도메인(예를 들어, Cas9-유래된 닉카제 도메인, Cas12-유래된 닉카제 도메인)을 포함하는 염기 편집기는 편집을 위해 표적화되어 지는 DNA 분자의 가닥을 절단할 수 있다. 이러한 실시형태에서, 비-표적화된 가닥은 절단되지 않는다.

[0293] 일부 실시형태에서, Cas9 뉴클레아제는 불활성(예를 들어, 불활성화된) DNA 절단 도메인을 가지며, 즉 Cas9는 "nCas9" 단백질("닉카제" Cas9의 경우)로 지칭되는 닉카제이다. Cas9 니카아제는 이중나선의 핵산 분자(예를 들어, 이중나선의 DNA 분자)의 한 가닥만 절단할 수 있는 Cas9 단백질일 수 있다. 일부 실시형태에서 Cas9 닉카제는 이중나선의 핵산 분자의 표적 가닥을 절단하는데, 이는 Cas9 닉카제가 Cas9에 결합된 gRNA(예를 들어, sgRNA)에 염기쌍을 이루는(상보성인) 가닥을 절단한다는 것을 의미한다. 일부 실시형태에서, Cas9 닉카제는 D10A 돌연변이를 포함하고 위치 840에 히스티딘을 갖는다. 일부 실시형태에서 Cas9 닉카제는 이중나선의 핵산 분자의 비-표적, 비-염기-편집된 가닥을 절단하는데, 이는 Cas9 닉카제가 Cas9에 결합된 gRNA(예를 들어, sgRNA)와 염기쌍을 이루지 않은 가닥을 절단한다는 것을 의미한다. 일부 실시형태에서, Cas9 닉카아제는 H840A 돌연변이를 포함하고 위치 10에 아스파르트산 잔기 또는 상응하는 돌연변이를 갖는다. 일부 실시형태에서 Cas9 닉카제는 본 명세서에 제공된 Cas9 닉카제 중 임의의 하나와 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99% 또는 적어도 99.5% 동일한 아미노산 서열을 포함한다. 추가의 적합한 Cas9 니카아제는 본 개시내용 및 해당 분야의 지식을 기반으로 당업자에게 명백할 것이며, 본 개시내용의 범주 내에 있다.

[0294] 예시적인 촉매적으로 Cas9 닉카제(nCas9)의 아미노산 서열은 다음과 같다:

[0295] MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLE
ESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINAS
GVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDIL
RVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG
SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPK
HSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEEN
EDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVS
GQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYL
QNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIK
RQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRK
MIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIAR
KKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGN
ELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKY
FDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD(서열번호: 234)

[0296] Cas9 뉴클레아제는 2가지 기능성 엔도뉴클레아제 도메인: RuvC 및 HNH를 갖는다. Cas9는 표적 DNA의 반대 가닥을 절단하기 위해 뉴클레아제 도메인을 위치시키는 표적 결합 시 형태적 변화를 겪는다. Cas9-매개된 DNA 절단의 최종 결과는 표적 DNA(PAM 서열의 상류의 ~3-4 뉴클레오티드) 내의 이중-가닥 파손(DSB)이다. 그 다음 생성

된 DSB는 다음 2가지 일반적인 수선 경로 중 하나에 의해 수선된다: (1) 효율적이지만 오류-경향의 비-상동성 말단 접합(NHEJ) 경로; 또는 (2) 덜 효율적이지만 높은-충실도 상동체 지향된 수선(HDR) 경로.

[0297] 비-상동성 말단 접합(NHEJ) 및/또는 상동체 지향된 수선(HDR)의 "효율성"은 임의의 편리한 방법에 의해 계산될 수 있다. 예를 들어, 일부 실시형태에서 효율성은 성공적인 HDR의 백분율의 관점에서 표현될 수 있다. 예를 들어, 서베이어 뉴클레아제 검정이 절단 생성물을 생성하기 위해 사용될 수 있고 생성물 대 기질의 비율이 사용되어 백분율을 계산할 수 있다. 예를 들어, 성공적인 HDR의 결과로 새로 통합된 제한 서열을 함유하는 DNA를 직접적으로 절단하는 서베이어 뉴클레아제 효소가 사용될 수 있다. 더 많이 절단된 기질은 더 높은 HDR 퍼센트(HDR의 더 높은 효율)를 나타낸다. 예시적인 예로서, HDR의 분율(백분율)은 다음 방정식 [(절단 생성물)/(기질 플러스 절단 생성물)](*예를 들어*, (b+c)/(a+b+c), 여기서 "a"는 DNA 기질의 밴드 강도이고 "b"와 "c"는 절단 생성물임)을 사용하여 계산될 수 있다.

[0298] 일부 실시형태에서, 효율성은 성공적인 NHEJ의 백분율의 관점에서 표현될 수 있다. 예를 들어, 절단 생성물을 생성하기 위해 T7 엔도뉴클레아제 I 검정이 사용될 수 있고 NHEJ 백분율을 계산하기 위해 생성물 대 기질의 비율이 사용될 수 있다. T7 엔도뉴클레아제 I은 야생형 및 돌연변이체 DNA 가닥의 혼성화로 인해 발생하는 불일치된 이종이중나선 DNA를 절단한다(NHEJ는 원래 절단의 부위에서 작은 무작위 삽입 또는 결실(삽입결실)를 생성한다). 더 많은 절단은 더 큰 NHEJ 퍼센트(NHEJ의 더 큰 효율성)를 나타낸다. 예시적인 예로서, NHEJ의 분율(백분율)은 다음 방정식을 사용하여 계산될 수 있다: $(1-(1-(b+c)/(a+b+c))^{1/2})\times 100$, 여기서 "a"는 DNA 기질의 밴드 강도이고 "b"와 "c"는 절단 생성물이다(Ran 등, *Cell*. 2013 Sep. 12; 154(6):1380-9; 및 Ran 등, *Nat Protoc*. 2013 Nov.; 8(11): 2281-2308).

[0299] NHEJ 수선 경로는 가장 활성인 수선 메커니즘이고, DSB 부위에서 작은 뉴클레오티드 삽입 또는 결실(삽입결실)을 빈번하게 야기한다. NHEJ-매개된 DSB 수선의 무작위성은 Cas9 및 gRNA 또는 가이드 폴리뉴클레오티드를 발현하는 세포의 모집단이 다양한 돌연변이 배열을 초래할 수 있기 때문에 중요한 실제적 의미를 갖는다. 대부분의 실시형태에서, NHEJ는 표적화된 유전자의 개방 판독 프레임(ORF) 내의 조기 정지 코돈으로 이어지는 아미노산 결실, 삽입 또는 프레임시프트 돌연변이를 초래하는 표적 DNA에서 작은 삽입결실을 발생시킨다. 이상적인 최종 결과는 표적 유전자 내의 기능-상실 돌연변이이다.

[0300] NHEJ-매개된 DSB 수선은 종종 유전자의 개방 판독 프레임을 파괴하는 반면, 상동체 지향된 수선(HDR)은 단일 뉴클레오티드 변화부터 형광단 또는 태그의 추가와 같은 대규모 삽입에 이르기까지 특정 뉴클레오티드 변화를 생성하는데 사용될 수 있다.

[0301] 유전자 편집을 위해 HDR을 이용하기 위해 원하는 서열을 함유하는 DNA 수선 주형을 gRNA(들) 및 Cas9 또는 Cas9 닉카제와 함께 관심있는 세포 유형 안으로 전달할 수 있다. 수선 주형은 원하는 편집뿐만 아니라 표적의 바로 상류 및 하류에 추가 상동성 서열(왼쪽 & 오른쪽 상동성 아암이라 함)을 함유할 수 있다. 각 상동성 아암의 길이는 도입되어 지는 변화의 크기에 따라 달라질 수 있으며, 더 큰 삽입에는 더 긴 상동성 아암이 필요하다. 수선 주형은 단일-가닥 올리고뉴클레오티드, 이중-가닥 올리고뉴클레오티드 또는 이중-가닥 DNA 플라스미드일 수 있다. HDR의 효율성은 Cas9, gRNA 및 외인성 수선 주형을 발현하는 세포에서도 일반적으로 낮다(변형된 대립유전자의 <10%). HDR은 세포 주기의 S 및 G2 단계에서 발생하므로 세포를 동기화함에 의해 HDR의 효율성이 향상될 수 있다. NHEJ에 수반된 유전자를 화학적으로 또는 유전적으로 억제하는 것도 HDR 빈도를 증가시킬 수 있다.

[0302] 일부 실시형태에서, Cas9는 변형된 Cas9이다. 주어진 gRNA 표적화 서열은 게놈 전반에 걸쳐 부분적인 상동성이 존재하는 추가 부위를 가질 수 있다. 이들 부위는 표적-외로 지칭되고 gRNA를 설계할 때 고려되어야 한다. gRNA 설계를 최적화하는 것 외에도 Cas9에 대한 변형을 통해 CRISPR 특이성이 증가될 수도 있다. Cas9는 두 뉴클레아제 도메인인, RuvC와 HNH의 조합된 활성을 통해 이중 가닥 파손(DSB)을 생성한다. SpCas9의 D10A 돌연변이체인 Cas9 닉카제는 하나의 뉴클레아제 도메인을 유지하고 DSB 이외의 DNA 닉을 생성한다. 닉카제 시스템은 특정 유전자 편집을 위해 편집한 HDR-매개된 유전자와 조합될 수도 있다.

[0303] *촉매적으로 사멸된 뉴클레아제*

[0304] 촉매적으로 사멸된(즉, 표적 폴리뉴클레오티드 서열을 절단할 수 없는) 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인을 포함하는 염기 편집기가 또한 본 명세서에 제공된다. 본 명세서에서 용어 "촉매적으로 사멸" 및 "뉴클레아제 사멸"은 핵산의 가닥을 절단하는데 있어 그의 불능을 초래하는 하나 이상의 돌연변이 및/또는 결실을 갖는 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인을 지칭하기 위해 상호교환적으로 사용된다. 일부 실시형태에서, 촉매적으로 사멸된 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인 염기

편집기는 하나 이상의 뉴클레아제 도메인에서 특정 점 돌연변이의 결과로 뉴클레아제 활성을 결할 수 있다. 예를 들어, Cas9 도메인을 포함하는 염기 편집기의 경우, Cas9는 D10A 돌연변이와 H840A 돌연변이 둘 모두를 포함할 수 있다. 이러한 돌연변이는 두 뉴클레아제 도메인 모두를 불활성화하여 뉴클레아제 활성의 손실을 초래한다. 다른 실시형태에서, 촉매적으로 사멸된 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인은 촉매적 도메인(*예를 들어*, RuvC1 및/또는 HNH 도메인)의 전부 또는 일부의 하나 이상의 결실을 포함할 수 있다. 추가 실시형태에서, 촉매적으로 사멸된 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인은 점 돌연변이(*예를 들어*, D10A 또는 H840A)뿐만 아니라 뉴클레아제 도메인의 전체 또는 일부의 결실을 포함한다. dCas9 도메인은 당업계에 공지되어 있고, 예를 들어 Qi 등, "Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression." *Cell*. 2013; 152(5):1173-83에 기재되어 있으며, 그 전체 내용은 본 명세서에 참고로 포함된다.

[0305] 추가의 적합한 뉴클레아제-불활성 dCas9 도메인은 본 개시내용 및 해당 분야의 지식에 기반하여 당업자에게 명백할 것이고, 본 개시내용의 범주 내에 있다. 이러한 추가의 예시적인 적합한 뉴클레아제-불활성 Cas9 도메인은 D10A/H840A, D10A/D839A/H840A 및 D10A/D839A/H840A/N863A 돌연변이체 도메인을 포함하지만 이에 제한되지 않는다(*예를 들어*, Prashant 등, CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nature Biotechnology*. 2013; 31(9): 833-838을 참조하며, 이의 전체 내용은 본 명세서에 참조로 포함된다).

[0306] 일부 실시형태에서, dCas9는 Cas9 뉴클레아제 활성을 불활성화하는 하나 이상의 돌연변이를 갖는 Cas9 아미노산 서열에 상응하거나, 부분적으로 또는 전체적으로 이를 포함한다. 일부 실시형태에서, 뉴클레아제-불활성 dCas9 도메인은 본 명세서에 제시된 아미노산 서열의 D10X 돌연변이 및 H840X 돌연변이, 또는 본 명세서에 제공된 임의의 아미노산 서열에 상응하는 돌연변이를 포함하며, 여기서 X는 임의의 아미노산 변화이다. 일부 실시형태에서, 뉴클레아제-불활성 dCas9 도메인은 본 명세서에 제시된 아미노산 서열의 D10A 돌연변이 및 H840A 돌연변이, 또는 본 명세서에 제공된 임의의 아미노산 서열에 상응하는 돌연변이를 포함한다. 일부 실시형태에서, 뉴클레아제-불활성 Cas9 도메인은 클로닝 벡터 pPlatTET-gRNA2(수탁 번호 BAV54124)에 제시된 아미노산 서열을 포함한다.

[0307] 일부 실시형태에서, 변이체 Cas9 단백질은 가이드 표적 서열의 상보성 가닥을 절단할 수 있지만 이중 가닥 가이드 표적 서열의 비-상보성 가닥을 절단하는 능력이 감소된다. 예를 들어, 변이체 Cas9 단백질은 RuvC 도메인의 기능을 감소시키는 돌연변이(아미노산 치환)를 가질 수 있다. 비-제한적인 예로서, 일부 실시형태에서, 변이체 Cas9 단백질은 D10A(아미노산 위치 10에서 아스파르테이트에서 알라닌으로)를 가지고 따라서 이중 가닥 가이드 표적 서열의 상보성 가닥을 절단할 수 있지만 이중 가닥 가이드 표적 서열의 비-상보성 가닥을 절단하는 능력이 감소된다(따라서 변이체 Cas9 단백질이 이중 가닥 표적 핵산을 절단할 때 이중 가닥 파손(DSB) 대신 단일 가닥 파손(SSB)을 초래함)(예를 들어, Jinek 등, *Science*. 2012 Aug. 17; 337(6096):816-21 참조).

[0308] 일부 실시형태에서, 변이체 Cas9 단백질은 이중 가닥 가이드 표적 서열의 비-상보성 가닥을 절단할 수 있지만 가이드 표적 서열의 상보성 가닥을 절단하는 능력이 감소된다. 예를 들어, 변이체 Cas9 단백질은 HNH 도메인(RuvC/HNH/RuvC 도메인 모티프)의 기능을 감소시키는 돌연변이(아미노산 치환)를 가질 수 있다. 비-제한적인 예로서, 일부 실시형태에서, 변이체 Cas9 단백질은 H840A(아미노산 위치 840에서 히스티딘에서 알라닌으로) 돌연변이를 가지고 따라서 가이드 표적 서열의 비-상보성 가닥을 절단할 수 있지만 가이드 표적 서열의 상보성 가닥을 절단하는 능력이 감소된다 (따라서 변이체 Cas9 단백질이 이중 가닥 가이드 표적 서열을 절단할 때 DSB 대신 SSB를 초래함). 이러한 Cas9 단백질은 가이드 표적 서열(*예를 들어*, 단일 가닥 가이드 표적 서열)을 절단하는 감소된 능력을 갖지만, 가이드 표적 서열(*예를 들어*, 단일 가닥 가이드 표적 서열)에 결합하는 능력은 유지한다.

[0309] 또 다른 비-제한적인 예로서, 일부 실시형태에서, 변이체 Cas9 단백질은 폴리펩티드가 표적 DNA를 절단하는 감소된 능력을 갖도록 W476A 및 W1126A 돌연변이를 담지한다. 이러한 Cas9 단백질은 표적 DNA(*예를 들어*, 단일 가닥 표적 DNA)를 절단하는 감소된 능력을 갖지만 표적 DNA(*예를 들어*, 단일 가닥 표적 DNA)에 결합하는 능력을 유지한다.

[0310] 또 다른 비-제한적인 예로서, 일부 실시형태에서, 변이체 Cas9 단백질은 폴리펩티드가 표적 DNA를 절단하는 감소된 능력을 갖도록 P475A, W476A, N477A, D1125A, W1126A 및 D1127A 돌연변이를 담지한다. 이러한 Cas9 단백질은 표적 DNA(*예를 들어*, 단일 가닥 표적 DNA)를 절단하는 감소된 능력을 갖지만 표적 DNA(*예를 들어*, 단일 가닥 표적 DNA)에 결합하는 능력을 유지한다.

[0311] 또 다른 비-제한적인 예로서, 일부 실시형태에서, 변이체 Cas9 단백질은 폴리펩티드가 표적 DNA를 절단하는 감소된 능력을 갖도록 H840A, W476A 및 W1126A 돌연변이를 담지한다. 이러한 Cas9 단백질은 표적 DNA(*예를 들어*, 단일 가닥 표적 DNA)를 절단하는 감소된 능력을 갖지만 표적 DNA(*예를 들어*, 단일 가닥 표적 DNA)에 결합하는 능력을 유지한다. 또 다른 비-제한적인 예로서, 일부 실시형태에서, 변이체 Cas9 단백질은 폴리펩티드가 표적 DNA를 절단하는 감소 능력을 갖도록 H840A, D10A, W476A 및 W1126A 돌연변이를 담지한다. 이러한 Cas9 단백질은 표적 DNA(*예를 들어*, 단일 가닥 표적 DNA)를 절단하는 감소된 능력을 갖지만 표적 DNA(*예를 들어*, 단일 가닥 표적 DNA)에 결합하는 능력을 유지한다. 일부 실시형태에서, 변이체 Cas9는 Cas9 HNH 도메인에서의 위치 840에서 촉매적 His 잔기를 복원했다(A840H).

[0312] 또 다른 비-제한적인 예로서, 일부 실시형태에서, 변이체 Cas9 단백질은 폴리펩티드는 표적 DNA를 절단하는 감소된 능력을 갖도록 H840A, P475A, W476A, N477A, D1125A, W1126A 및 D1127A 돌연변이를 담지한다. 이러한 Cas9 단백질은 표적 DNA(*예를 들어*, 단일 가닥 표적 DNA)를 절단하는 감소된 능력을 갖지만 표적 DNA(*예를 들어*, 단일 가닥 표적 DNA)에 결합하는 능력을 유지한다. 또 다른 비-제한적인 예로서, 일부 실시형태에서, 변이 Cas9 단백질은 폴리펩티드가 표적 DNA를 절단하는 감소된 능력을 갖도록 D10A, H840A, P475A, W476A, N477A, D1125A, W1126A 및 D1127A 돌연변이를 담지한다. 이러한 Cas9 단백질은 표적 DNA(*예를 들어*, 단일 가닥 표적 DNA)를 절단하는 감소된 능력을 갖지만 표적 DNA(*예를 들어*, 단일 가닥 표적 DNA)에 결합하는 능력을 유지한다. 일부 실시형태에서, 변이체 Cas9 단백질이 W476A 및 W1126A 돌연변이를 담지하는 경우 또는 변이체 Cas9 단백질이 P475A, W476A, N477A, D1125A, W1126A 및 D1127A 돌연변이를 담지하는 경우, 변이체 Cas9 단백질은 PAM 서열에 효율적으로 결합하지 않는다. 따라서, 일부 그러한 실시형태에서, 그러한 변이체 Cas9 단백질이 결합의 방법에 사용될 때, 방법은 PAM 서열을 필요로 하지 않는다. 환언하면, 일부 실시형태에서 그러한 변이체 Cas9 단백질이 결합의 방법에 사용될 때, 방법은 가이드 RNA를 포함할 수 있지만, 방법은 PAM 서열의 부재에서 수행될 수 있다(그리고 결합의 특이성은 따라서 가이드 RNA의 표적화 세그먼트에 의해 제공된다). 상기 효과를 달성(*즉*, 하나 또는 다른 뉴클레아제 부분을 불활성화)하기 위해 다른 잔기가 돌연변이될 수 있다. 비-제한적인 예로서, 잔기 D10, G12, G17, E762, H840, N854, N863, H982, H983, A984, D986 및/또는 A987이 변경(*즉*, 치환)될 수 있다. 또한, 알라닌 치환 이외의 돌연변이가 적합하다.

[0313] 일부 실시형태에서, 감소된 촉매적 활성을 갖는 변이체 Cas9 단백질(*예를 들어*, Cas9 단백질이 D10, G12, G17, E762, H840, N854, N863, H982, H983, A984, D986 및/또는 A987 돌연변이, *예를 들어* D10A, G12A, G17A, E762A, H840A, N854A, N863A, H982A, H983A, A984A 및/또는 D986A를 갖는 경우), 변이체 Cas9 단백질은 그것이 가이드 RNA와 상호작용하는 능력을 유지하는 한 (그것이 여전히 가이드 RNA에 의해 표적 DNA 서열로 가이드되기 때문에) 여전히 부위-특이적 방식으로 표적 DNA에 결합할 수 있다.

[0314] 일부 실시형태에서, 변이체 Cas 단백질은 spCas9, spCas9-VRQR, spCas9-VRER, xCas9(sp), saCas9, saCas9-KKH, spCas9-MQKSER, spCas9-LRKIQK, 또는 spCas9-LRVSQL일 수 있다.

[0315] 일부 실시형태에서, Cas9 도메인은 *스타필로코커스 아우레우스*(SaCas9)로부터의 Cas9 도메인이다. 일부 실시형태에서, SaCas9 도메인은 뉴클레아제 활성 SaCas9, 뉴클레아제 불활성 SaCas9(SaCas9d), 또는 SaCas9 닉카제(SaCas9n)이다. 일부 실시형태에서, SaCas9는 N579A 돌연변이, 또는 본 명세서와 함께 제출된 서열 목록에 제공된 임의의 아미노산 서열에서 상응하는 돌연변이를 포함한다.

[0316] 일부 실시형태에서, SaCas9 도메인, SaCas9d 도메인 또는 SaCas9n 도메인은 비-규범적 PAM을 갖는 핵산 서열에 결합할 수 있다. 일부 실시형태에서, SaCas9 도메인, SaCas9d 도메인, 또는 SaCas9n 도메인은 NNGRRT 또는 NNGRRV PAM 서열을 갖는 핵산 서열에 결합할 수 있다. 일부 실시형태에서, SaCas9 도메인은 E781X, N967X 및 R1014X 돌연변이 중 하나 이상, 또는 본 명세서에 제공된 임의의 아미노산 서열에서의 상응하는 돌연변이를 포함하며, 여기서 X는 임의의 아미노산이다. 일부 실시형태에서, SaCas9 도메인은 E781K, N967K 및 R1014H 돌연변이 중 하나 이상, 또는 본 명세서에 제공된 임의의 아미노산 서열에서 하나 이상의 상응하는 돌연변이를 포함한다. 일부 실시형태에서, SaCas9 도메인은 E781K, N967K 또는 R1014H 돌연변이, 또는 본 명세서에 제공된 임의의 아미노산 서열에서 상응하는 돌연변이를 포함한다.

[0317] 일부 실시형태에서, 융합 단백질에 존재하는 Cas9 도메인 중 하나는 PAM 서열에 대한 필요성이 없는 가이드 뉴클레오티드 서열-프로그램가능 DNA-결합 단백질 도메인으로 대체될 수 있다. 일부 실시형태에서 Cas9는 SaCas9이다. SaCas9의 잔기 A579는 N579로부터 돌연변이되어 SaCas9 닉카제를 생성할 수 있다. 잔기 K781, K967 및 H1014는 E781, N967 및 R1014로부터 돌연변이되어 SaKKH Cas9를 생성할 수 있다.

[0318] 일부 실시형태에서, 아미노산 치환 D1135M, S1136Q, G1218K, E1219F, A1322R, D1332A, R1335E 및

T1337R(SpCas9-MQKFRAER)을 포함하고 변경된 PAM 5'-NGC-3'에 대한 특이성을 갖는 변형된 SpCas9가 사용되었다.

[0319] *S. 피오게네스* Cas9에 대한 대안은 포유동물 세포에서 절단 활성을 나타내는 Cpf1 계열로부터의 RNA-유도된 엔도뉴클레아제를 포함할 수 있다. *프레보텔라* 및 *프란시셀라 1*로부터의 CRISPR(CRISPR/Cpf1)은 CRISPR/Cas9 시스템에 유사한 DNA-편집 기술이다. Cpf1은 클래스 II CRISPR/Cas 시스템의 RNA-가이드된 엔도뉴클레아제이다. 이 후천적 면역 메커니즘은 *프레보텔라* 및 *프란시셀라* 박테리아에서 발견된다. Cpf1 유전자는 가이드 RNA를 사용하여 바이러스 DNA를 찾아 절단하는 엔도뉴클레아제를 코딩하는, CRISPR 좌위와 회합된다. Cpf1은 Cas9보다 더 작고 단순한 엔도뉴클레아제로 CRISPR/Cas9 시스템 제한 중 일부를 극복한다. Cas9 뉴클레아제와 달리, Cpf1-매개된 DNA 절단의 결과는 짧은 3' 오버행을 갖는 이중-가닥 파손이다. Cpf1의 엇갈린 절단 패턴은 전통적인 제한 효소 클로닝에 유사한 방향성 유전자 전달의 가능성을 열어, 유전자 편집의 효율성을 증가시킬 수 있다. 상기에 기술된 Cas9 변이체 및 오솔로그와 마찬가지로, Cpf1은 CRISPR에 의해 표적화 될 수 있는 부위의 수를 SpCas9가 선호하는 NGG PAM 부위를 결하는 AT-풍부 영역 또는 AT-풍부 게놈으로 확장할 수도 있다. Cpf1 좌위는 혼합된 알파/베타 도메인, 나선형 영역이 이어지는 RuvC-I, RuvC-II 및 아연 핑거-유사 도메인을 함유한다. Cpf1 단백질은 Cas9의 RuvC 도메인에 유사한 RuvC-유사 엔도뉴클레아제 도메인을 갖는다.

[0320] 더욱이, Cpf1은 Cas9와 달리 HNH 엔도뉴클레아제 도메인을 갖지 않고, Cpf1의 N-말단은 Cas9의 알파-나선 인식 로브를 갖지 않는다. Cpf1 CRISPR-Cas 도메인 아키텍처는 Cpf1이 기능적으로 고유하며, 클래스 2, 유형 V CRISPR 시스템으로 분류됨을 보여준다. Cpf1 좌위는 유형 II 시스템보다 유형 I 및 III과 더 유사한 Cas1, Cas2 및 Cas4 단백질을 인코딩한다. 기능성 Cpf1은 트랜스-활성화 CRISPR RNA(tracrRNA)를 필요로 하지 않고, 따라서 CRISPR(crRNA)만을 필요로 한다. Cpf1은 Cas9보다 작을 뿐만 아니라 더 작은 sgRNA 분자(대략적으로 많게는 Cas9 뉴클레오티드의 절반)를 갖기 때문에 이것은 게놈 편집에 이점이 있다. Cpf1-crRNA 복합체는 Cas9에 의해 표적화된 G-풍부 PAM과 달리 프로토스페이서 인접 모티프 5'-YTN-3' 또는 5'-TTN-3'의 식별에 의해 표적 DNA 또는 RNA를 절단한다. PAM의 식별 후, Cpf1은 4개 또는 5개 뉴클레오티드의 오버행을 갖는 점성-말단-유사 DNA 이중-가닥 파손을 도입한다.

[0321] 일부 실시형태에서, Cas9는 변경된 PAM 서열에 대한 특이성을 갖는 Cas9 변이체이다. 일부 실시형태에서, 추가의 Cas9 변이체 및 PAM 서열은 Miller, S.M., 등 Continuous evolution of SpCas9 variants compatible with non-G PAMs, Nat. Biotechnol. (2020)에 기술되어 있으며, 그 전체 내용은 참조로 본 명세서에 포함된다. 일부 실시형태에서, Cas9 변이체는 특정한 PAM 필요성을 갖지 않는다. 일부 실시형태에서, Cas9 변이체, *예를 들어* SpCas9 변이체는 NRNH PAM에 대한 특이성을 가지며, 여기서 R은 A 또는 G이고 H는 A, C 또는 T이다. 일부 실시형태에서, SpCas9 변이체는 PAM 서열 AAA, TAA, CAA, GAA, TAT, GAT, 또는 CAC에 대한 특이성을 갖는다. 일부 실시형태에서, SpCas9 변이체는 위치 1114, 1134, 1135, 1137, 1139, 1151, 1180, 1188, 1211, 1218, 1219, 1221, 1249, 1256, 1264, 1290, 1318, 1317, 1320, 1321, 1323, 1332, 1333, 1335, 1337, 또는 1339 또는 이의 상응하는 위치에서 아미노산 치환을 포함한다. 일부 실시형태에서, SpCas9 변이체는 위치 1114, 1135, 1218, 1219, 1221, 1249, 1320, 1321, 1323, 1332, 1333, 1335, 또는 1337 또는 이의 상응하는 위치에서 아미노산 치환을 포함한다. 일부 실시형태에서, SpCas9 변이체는 위치 1114, 1134, 1135, 1137, 1139, 1151, 1180, 1188, 1211, 1219, 1221, 1256, 1264, 1290, 1318, 1317, 1320, 1323, 1333 또는 이의 상응하는 위치에서 아미노산 치환을 포함한다. 일부 실시형태에서, SpCas9 변이체는 위치 1114, 1131, 1135, 1150, 1156, 1180, 1191, 1218, 1219, 1221, 1227, 1249, 1253, 1286, 1293, 1320, 1321, 1332, 1335, 1339 또는 이의 상응하는 위치에서 아미노산 치환을 포함한다. 일부 실시형태에서, SpCas9 변이체는 위치 1114, 1127, 1135, 1180, 1207, 1219, 1234, 1286, 1301, 1332, 1335, 1337, 1338, 1349 또는 이의 상응하는 위치에서 아미노산 치환을 포함한다. SpCas9 변이체의 예시적인 아미노산 치환 및 PAM 특이성은 **표 2A-2D**에 나타나 있다.

[0322] 표 2A SpCas9 변이체

SpCas9	SpCas9 아미노산 위치												
	1114	1135	1218	1219	1221	1249	1320	1321	1323	1332	1333	1335	1337
	R	D	G	E	Q	P	A	P	A	D	R	R	T
AAA		N		V	H						G		
AAA		N		V	H						G		
AAA				V							G		
TAA	G	N		V							I		
TAA		N		V							I		A
TAA	G	N		V							I		A
CAA				V							K		
CAA		N		V							K		
CAA		N		V							K		
GAA				V	H		V				K		
GAA		N		V			V				K		
GAA				V	H		V				K		
TAT			S	V	H	S		S				L	
TAT			S	V	H	S		S				L	
TAT			S	V	H	S		S				L	
GAT				V							I		
GAT				V					D			Q	
GAT				V					D			Q	
CAC				V						N		Q	N
CAC		N		V								Q	N
CAC				V						N		Q	N

[0323]

[0324]

표 2B

SpCas9	SpCas9 아미노산 위치																			
	1114	1134	1135	1137	1139	1151	1180	1188	1211	1219	1221	1256	1264	1290	1318	1317	1320	1323	1333
	R	F	D	P	V	K	D	K	K	E	Q	Q	H	V	L	N	A	A	R
GAA										V	H						V		K
GAA			N	S						V							V	D	K
GAA			N							V	H		Y				V		K
CAA			N							V	H		Y				V		K
CAA	G		N	S						V	H		Y				V		K
CAA			N					R		V	H						V		K
CAA			N				G		R	V	H		Y				V		K
CAA			N							V	H		Y				V		K
AAA			N				G			V	H	R	Y				V	D	K
CAA	G		N				G			V	H		Y				V	D	K
CAA		L	N				G			V	H		Y			T	V	D	K
TAA	G		N				G			V	H		Y	G	S		V	D	K
TAA	G		N			E	G			V	H		Y		S		V		K
TAA	G		N				G			V	H		Y		S		V	D	K
TAA	G		N				G		R	V	H						V		K
TAA			N				G		R	V	H		Y				V		K
TAA	G		N		A		G			V	H						V		K
TAA	G		N							V	H						V		K

[0325]

표 2C

SpCas9	SpCas9 아미노산 위치																				
	1114	1131	1135	1150	1156	1180	1191	1218	1219	1221	1227	1249	1253	1286	1303	1320	1321	1325	1332	1335	1339
	R	Y	D	E	K	D	K	G	E	Q	A	P	E	N	A	A	P	D	R	L	T
SacB.TAT			N				N		V	H						V	S		L		
SacB.TAT			N					S	V	H		S					S	G	L		
AAT			N					S	V	H	V	S		K	T		S	G	L	I	
TAT	G		N			G		S	V	H		S	K				S	G	L		
TAT	G		N			G		S	V	H		S					S	G	L		
TAT	G	C	N			G		S	V	H		S					S	G	L		
TAT	G	C	N			G		S	V	H		S					S	G	L		
TAT	G	C	N			G		S	V	H		S					S	G	L		
TAT	G	C	N		E	G		S	V	H		S					S	G	L		
TAT	G	C	N	V		G		S	V	H		S					S	G	L		
TAT		C	N			G		S	V	H		S					S	G	L		
TAT	G	C	N			G		S	V	H		S					S	G	L		

[0326] 표 2D

SpCas9	SpCas9 아미노산 위치																	
	1114	1127	1135	1180	1207	1219	1234	1286	1301	1332	1335	1337	1338	1349
	R	D	D	D	E	E	N	N	P	D	R	T	S	H
SacB.CAC			N			V				N	Q	N		
AAC	G		N			V				N	Q	N		
AAC	G		N			V				N	Q	N		
TAC	G		N			V				N	Q	N		
TAC	G		N			V		H		N	Q	N		
TAC	G		N		G	V	D	H		N	Q	N		
TAC	G		N			V				N	Q	N		
TAC	G	G	N	E		V		H		N	Q	N		
TAC	G		N			V		H		N	Q	N	T	R
TAC	G		N			V				N	Q	N	T	R

[0327] 일부 실시형태에서, 핵산 프로그램가능 DNA 결합 단백질(napDNAbp)은 미생물 CRISPR-Cas 시스템의 단일 효과기이다. 미생물 CRISPR-Cas 시스템의 단일 효과기는 제한 없이, Cas9, Cpf1, Cas12b/C2c1 및 Cas12c/C2c3을 포함한다. 전형적으로, 미생물 CRISPR-Cas 시스템은 클래스 1과 클래스 2 시스템으로 분할된다. 클래스 1 시스템에는 다중 하위단위 효과기 복합체가 있는 반면, 클래스 2 시스템에는 단일 단백질 효과기가 있다. 예를 들어, Cas9 및 Cpf1은 클래스 2 효과기이다. Cas9 및 Cpf1에 부가하여, 3가지 별개의 클래스 2 CRISPR-Cas 시스템(Cas12b/C2c1 및 Cas12c/C2c3)이 Shmakov 등, "Discovery and Functional Characterization of Diverse Class 2 CRISPR Cas Systems", *Mol. Cell,* 2015 Nov. 5; 60(3): 385-397에 기술되어 있으며, 그 전체 내용은 참조로 본 명세서에 포함된다. 2개의 시스템의 효과기인, Cas12b/C2c1 및 Cas12c/C2c3은 Cpf1과 관련된 RuvC-유사 엔도뉴클레아제 도메인을 함유한다. 제3 시스템은 2개의 예측된 HEPN RNase 도메인을 갖는 효과기를 함유한다. 성숙 CRISPR RNA의 생산은 Cas12b/C2c1에 의한 CRISPR RNA의 생산과 달리 tracrRNA-독립적이다. Cas12b/C2c1은 DNA 절단을 위해 CRISPR RNA와 tracrRNA 둘 모두에 의존한다.

[0328] 일부 실시형태에서, napDNAbp는 원형 순열체(예를 들어, 서열번호: 257)이다.

[0329] *알리시클로바실러스 에시도테라스트리스* Cas12b/C2c1(AacC2c1)의 결정 구조는 키메라 단일-분자 가이드 RNA(sgRNA)와 복합체로 보고되었다. 예를 들어, Liu 등, "C2c1-sgRNA Complex Structure Reveals RNA-Guided DNA Cleavage Mechanism", *Mol. Cell,* 2017 Jan. 19; 65(2):310-322를 참조하며, 그 전체 내용은 참조로 본 명세서에 포함된다. 결정 구조는 또한 삼원 복합체로서 표적 DNA에 결합된 *알리시클로바실러스 에시도테레스트리스* C2c1에서도 보고되었다. 예를 들어, Yang 등, "PAM-dependent Target DNA Recognition and Cleavage by C2C1 CRISPR-Cas endonuclease", *Cell,* 2016 Dec. 15; 167(7):1814-1828을 참조하며, 그 전체 내용은 참조로 본 명세서에 포함된다. 표적 및 비-표적 DNA 가닥 모두를 갖는 AacC2c1의 촉매적으로 적격인 형태는 단일 RuvC 촉매적 포켓 내에 독립적으로 위치되어 포획되어, Cas12b/C2c1-매개된 절단으로 인해 표적 DNA의 엇갈린 7개-뉴클레오티드 파손을 초래한다. Cas12b/C2c1 삼원 복합체와 이전에 식별된 Cas9 및 Cpf1 대응체 간의 구조적 비교는 CRISPR-Cas9 시스템에 의해 사용된 메커니즘의 다양성을 입증한다.

[0330] 일부 실시형태에서, 본 명세서에 제공된 임의의 융합 단백질의 핵산 프로그램가능 DNA 결합 단백질(napDNAbp)은 Cas12b/C2c1 또는 Cas12c/C2c3 단백질일 수 있다. 일부 실시형태에서, napDNAbp는 Cas12b/C2c1 단백질이다. 일부 실시형태에서, napDNAbp는 Cas12c/C2c3 단백질이다. 일부 실시형태에서, napDNAbp는 자연적으로-발생하는 Cas12b/C2c1 또는 Cas12c/C2c3 단백질과 적어도 85%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99% 또는 적어도 99.5% 동일한 아미노산 서열을 포함한다. 일부 실시형태에서, napDNAbp는 자연적으로-발생하는 Cas12b/C2c1 또는 Cas12c/C2c3 단백질이다. 일부 실시형태에서, napDNAbp는 본 명세서에 제공된 napDNAbp 서열 중 임의의 하나와 적어도 85%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 적어도 99.5% 동일한 아미노산 서열을 포함한다. 다른 박테리아 종으로부터의 Cas12b/C2c1 또는 Cas12c/C2c3도 본 개시내용에 따라 사용될 수 있다는 것이 이해되어야 한다.

[0331] 일부 실시형태에서, napDNAbp는 Cas12c를 지칭한다. 일부 실시형태에서, Cas12c 단백질은 Cas12c1(서열번호: 266) 또는 Cas12c1의 변이체이다. 일부 실시형태에서, Cas12 단백질은 Cas12c2(서열번호: 267) 또는 Cas12c2의 변이체이다. 일부 실시형태에서, Cas12 단백질은 *올레이필루스* sp. HI0009로부터의 Cas12c 단백질(즉, OspCas12c; 서열번호: 268) 또는 OspCas12c의 변이체이다. 이들 Cas12c 분자는 Yan 등, "Functionally Diverse Type V CRISPR-Cas Systems," Science, 2019 Jan. 4; 363: 88-91에 기술되어 있으며; 그 전체 내용은 참조로 본 명세서에 포함된다. 일부 실시형태에서, napDNAbp는 자연적으로-발생하는 Cas12c1, Cas12c2 또는 OspCas12c 단백질과 적어도 85%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99% 또는 적어도 99.5% 동일한 아미노산 서열을 포함한다. 일부 실시형태에서, napDNAbp는 자연적으로-발생하는 Cas12c1, Cas12c2, 또는 OspCas12c 단백질이다. 일부 실시형태에서, napDNAbp는 본 명세서에 기재된 임의의 Cas12c1, Cas12c2 또는 OspCas12c 단백질과 적어도 85%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 99.5% 동일한 아미노산 서열을 포함한다. 다른 박테리아 종으로부터의 Cas12c1, Cas12c2 또는 OspCas12c도 본 개시내용에 따라 사용될 수 있다는 것이 이해되어야 한다.

[0332] 일부 실시형태에서, napDNAbp는 예를 들어, Yan 등, "Functionally Diverse Type V CRISPR-Cas Systems," Science, 2019 Jan. 4; 363: 88-91에 기술된 Cas12g, Cas12h, 또는 Cas12i를 지칭하며; 각각의 전체 내용은 참조로 본 명세서에 포함된다. 예시적인 Cas12g, Cas12h 및 Cas12i 폴리펩티드 서열은 서열 목록에 서열번호: 269-272로 제공된다. 10테라바이트 초과의 서열 데이터를 집계함에 의해, Cas12g, Cas12h 및 Cas12i를 포함하여 이전에 특성화된 클래스 V 단백질과 약한 유사성을 보이는 유형 V Cas 단백질의 새로운 분류가 식별되었다. 일부 실시형태에서, Cas12 단백질은 Cas12g 또는 Cas12g의 변이체이다. 일부 실시형태에서, Cas12 단백질은 Cas12h 또는 Cas12h의 변이체이다. 일부 실시형태에서, Cas12 단백질은 Cas12i 또는 Cas12i의 변이체이다. 다른 RNA-가이드된 DNA 결합 단백질이 napDNAbp로서 사용될 수 있고 본 개시내용의 범주 내에 있다는 것이 이해되어야 한다. 일부 실시형태에서, napDNAbp는 자연적으로-발생하는 Cas12g, Cas12h 또는 Cas12i 단백질과 적어도 85%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99% 또는 적어도 99.5% 동일한 아미노산 서열을 포함한다. 일부 실시형태에서, napDNAbp는 자연적으로-발생하는 Cas12g, Cas12h 또는 Cas12i 단백질이다. 일부 실시형태에서, napDNAbp는 본 명세서에 기재된 임의의 Cas12g, Cas12h 또는 Cas12i 단백질과 적어도 85%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 99.5% 동일한 아미노산 서열을 포함한다. 다른 박테리아 종으로부터의 Cas12g, Cas12h 또는 Cas12i도 본 개시내용에 따라 사용될 수 있다는 것이 이해되어야 한다. 일부 실시형태에서, Cas12i는 Cas12i1 또는 Cas12i2이다.

[0333] 일부 실시형태에서, 본 명세서에 제공된 임의의 융합 단백질의 핵산 프로그램가능 DNA 결합 단백질(napDNAbp)은 Cas12j/CasΦ 단백질일 수 있다. Cas12j/CasΦ는 Pausch 등, "CRISPR-CasΦ from huge phages is a hypercompact genome editor," *Science*, 17 July 2020, Vol. 369, Issue 6501, pp. 333-337에 기술되어 있으며, 이는 그 전체가 본 명세서에 참조로 포함된다. 일부 실시형태에서, napDNAbp는 자연적으로-발생하는 Cas12j/CasΦ 단백질과 적어도 85%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99% 또는 99.5% 동일한 아미노산 서열을 포함한다. 일부 실시형태에서, napDNAbp는 자연적으로-발생하는 Cas12j/CasΦ 단백질이다. 일부 실시형태에서, napDNAbp는 뉴클레아제 불활성("사멸") Cas12j/CasΦ 단백질이다. 다른 종으로부터의 Cas12j/CasΦ도 본 개시내용에 따라 사용될 수 있다는 것이 이해되어야 한다.

[0334] ***내부 삽입을 갖는 융합 단백질***

[0335] 핵산 프로그램가능 핵산 결합 단백질, 예를 들어 napDNAbp에 융합된 이종 폴리펩티드를 포함하는 융합 단백질이 본 명세서에 제공된다. 이종 폴리펩티드는 자연 또는 야생형 napDNAbp 폴리펩티드 서열에서는 발견되지 않는 폴리펩티드일 수 있다. 이종 폴리펩티드는 napDNAbp의 C-말단, napDNAbp의 N-말단에서 napDNAbp에 융합될 수 있거나, napDNAbp의 내부 장소에서 삽입될 수 있다 일부 실시형태에서, 이종 폴리펩티드는 데아미나제(예를 들어, 아데노신 데아미나제의 시티딘) 또는 이의 기능적 단편이다. 예를 들어, 융합 단백질은 Cas9 또는 Cas12(예를 들어, Cas12b/C2c1) 폴리펩티드의 N-말단 단편 및 C-말단 단편에 의해 측접된 데아미나제를 포함할 수 있다. 일부 실시형태에서, 시티딘 데아미나제는 APOBEC 데아미나제(예를 들어, APOBEC1)이다. 일부 실시형태에서, 아데노신 데아미나제는 TadA(예를 들어, TadA*7.10 또는 TadA*8)이다. 일부 실시형태에서, TadA는 TadA*8 또는 TadA*9이다. 본 명세서에 기재된 바와 같은 TadA 서열(예를 들어, TadA7.10 또는 TadA*8)은 상기-기재된 융합 단백질에 적합한 데아미나제이다.

[0336] 일부 실시형태에서, 융합 단백질은 다음 구조를 포함한다:

[0337] NH2-[napDNAbp의 N-말단 단편]-[데아미나제]-[napDNAbp의 C-말단 단편]-COOH;

[0338] NH2-[Cas9의 N-말단 단편]-[아데노신 데아미나제]-[Cas9의 C-말단 단편]-COOH;

[0339] NH2-[Cas12의 N-말단 단편]-[아데노신 데아미나제]-[Cas12의 C-말단 단편]-COOH;

[0340] NH2-[Cas9의 N-말단 단편]-[시티딘 데아미나제]-[Cas9의 C-말단 단편]-COOH;

[0341] NH2-[Cas12의 N-말단 단편]-[시티딘 데아미나제]-[Cas12의 C-말단 단편]-COOH;

[0342] 여기서 "]-["의 각 경우는 선택적 링커이다.

[0343] 데아미나제는 원형 순열체 데아미나제일 수 있다. 예를 들어, 데아미나제는 원형 순열체 아데노신 데아미나제일 수 있다. 일부 실시형태에서, 데아미나제는 TadA 참조 서열에서 번호가 매겨진 아미노산 잔기 116, 136 또는 65에서 원형으로 바꾸어진 원형 순열체 TadA이다.

[0344] 융합 단백질은 하나 초과의 데아미나제를 포함할 수 있다. 융합 단백질은 예를 들어 1, 2, 3, 4, 5개 이상의 데아미나제를 포함할 수 있다. 일부 실시형태에서, 융합 단백질은 1개 또는 2개의 데아미나제를 포함한다. 융합 단백질에서 2개 이상의 데아미나제는 아데노신 데아미나제, 시티딘 데아미나제, 또는 이의 조합일 수 있다. 2개 이상의 데아미나제는 동종이량체 또는 이종이량체일 수 있다. 2개 이상의 데아미나제가 napDNAbp에 나란히 삽입될 수 있다. 일부 실시형태에서, 2개 이상의 데아미나제는 napDNAbp에서 나란히 있지 않을 수 있다.

[0345] 일부 실시형태에서, 융합 단백질에서의 napDNAbp는 Cas9 폴리펩티드 또는 이의 단편이다. Cas9 폴리펩티드는 변이체 Cas9 폴리펩티드일 수 있다. 일부 실시형태에서, Cas9 폴리펩티드는 Cas9 닉카제(nCas9) 폴리펩티드 또는 이의 단편이다. 일부 실시형태에서, Cas9 폴리펩티드는 뉴클레아제 사멸 Cas9(dCas9) 폴리펩티드 또는 이의 단편이다. 융합 단백질에서의 Cas9 폴리펩티드는 전장 Cas9 폴리펩티드일 수 있다. 일부 경우에, 융합 단백질에서의 Cas9 폴리펩티드는 전장 Cas9 폴리펩티드가 아닐 수 있다. Cas9 폴리펩티드는 예를 들어 자연적으로-발생하는 Cas9 단백질에 비해 N-말단 또는 C-말단 종단에서 절삭될 수 있다. Cas9 폴리펩티드는 원형으로 바꾸어진 Cas9 단백질일 수 있다. Cas9 폴리펩티드는 여전히 표적 폴리뉴클레오티드 및 가이드 핵산 서열에 결합할 수 있는 Cas9 폴리펩티드의 단편, 부분 또는 도메인일 수 있다.

[0346] 일부 실시형태에서, Cas9 폴리펩티드는 스트렙토코커스 피오게네스 Cas9(SpCas9), 스타필로코커스 아우레우스 Cas9(SaCas9), 스트렙토코커스 써모필루스 1 Cas9(St1Cas9), 또는 본 명세서에 기재된 임의의 Cas9 폴리펩티드의 단편 또는 변이체이다.

[0347] 일부 실시형태에서, 융합 단백질은 Cas9 내에 삽입된 아데노신 데아미나제 도메인 및 시티딘 데아미나제 도메인을 포함한다. 일부 실시형태에서, 아데노신 데아미나제는 Cas9 내에 융합되고 시티딘 데아미나제는 C-말단에 융합된다. 일부 실시형태에서, 아데노신 데아미나제는 Cas9 내에 융합되고 시티딘 데아미나제는 N-말단에 융합된다. 일부 실시형태에서 시티딘 데아미나제는 Cas9 내에 융합되고 아데노신 데아미나제는 C-말단에 융합된다. 일부 실시형태에서, 시티딘 데아미나제는 Cas9 내에 융합되고 아데노신 데아미나제는 N-말단에 융합된다.

[0348] 아데노신 데아미나제와 시티딘 데아미나제 및 Cas9를 갖는 융합 단백질의 예시적인 구조는 다음과 같이 제공된다:

[0349] NH2-[Cas9(아데노신 데아미나제)]-[시티딘 데아미나제]-COOH;

[0350] NH2-[시티딘 데아미나제]-[Cas9(아데노신 데아미나제)]-COOH;

[0351] NH2-[Cas9(시티딘 데아미나제)]-[아데노신 데아미나제]-COOH; 또는

[0352] NH2-[아데노신 데아미나제]-[Cas9(시티딘 데아미나제)]-COOH.

[0353] 일부 실시형태에서, 상기 일반 아키텍처에서 사용된 "-"는 선택적인 링커의 존재를 나타낸다.

[0354] 다양한 실시형태에서, 촉매 도메인은 아데노신 데아미나제 활성과 같은 DNA 변형 활성(예를 들어, 데아미나제 활성)을 갖는다. 일부 실시형태에서, 아데노신 데아미나제는 TadA(예를 들어, TadA*7.10)이다. 일부 실시형태에서, TadA는 TadA*8이다. 일부 실시형태에서, TadA*8은 Cas9 내에 융합되고 시티딘 데아미나제는 C-말단에 융합된다. 일부 실시형태에서, TadA*8은 Cas9 내에 융합되고 시티딘 데아미나제는 N-말단에 융합된다. 일부 실시형태에서, 시티딘 데아미나제는 Cas9 내에 융합되고 TadA*8은 C-말단에 융합된다. 일부 실시형태에서, 시티딘 데아미나제는 Cas9 내에 융합되고 TadA*8은 N-말단에 융합된다. TadA*8 및 시티딘 데아미나제 및 Cas9를 갖는 융합 단백질의 예시적인 구조는 다음과 같이 제공된다:

[0355] NH2-[Cas9(TadA*8)]-[시티딘 데아미나제]-COOH;

[0356] NH2-[시티딘 데아미나제]-[Cas9(TadA*8)]-COOH;

[0357] NH2-[Cas9(시티딘 데아미나제)]-[TadA*8]-COOH; 또는

[0358] NH2-[TadA*8]-[Cas9(시티딘 데아미나제)]-COOH.

[0359] 일부 실시형태에서, 상기 일반 아키텍처에서 사용된 "-"는 선택적인 링커의 존재를 나타낸다.

[0360] 이종 폴리펩티드(예를 들어, 데아미나제)는, 예를 들어, napDNAbp가 표적 폴리뉴클레오티드 및 가이드 핵산에 결합하는 그 능력을 유지하도록 적절한 장소에서 napDNAbp(예를 들어, Cas9 또는 Cas12(예를 들어, Cas12b/C2c1))에 삽입될 수 있다. 데아미나제(예를 들어, 아데노신 데아미나제, 시티딘 데아미나제 또는 아데노신 데아미나제 및 시티딘 데아미나제)는 데아미나제의 기능(예를 들어, 염기 편집 활성) 또는 napDNAbp(예를 들어, 표적 핵산 및 가이드 핵산에 결합하는 능력)를 손상함에 없이 napDNAbp 안으로 삽입될 수 있다. 데아미나제(예를 들어, 아데노신 데아미나제, 시티딘 데아미나제, 또는 아데노신 데아미나제 및 시티딘 데아미나제)는 예를 들어 결정학 연구에 의해 나타난 바와 같이 무질서한 영역 또는 고온 인자 또는 B-인자를 포함하는 영역에서 napDNAbp에 삽입될 수 있다. 예를 들어 용매에 노출된 영역 및 루프 같이, 덜 정돈되거나, 무질서하거나 또는 구조화되지 않은 단백질의 영역이 구조나 기능을 손상함에 없이 삽입에 사용될 수 있다. 데아미나제(예를 들어, 아데노신 데아미나제, 시티딘 데아미나제, 또는 아데노신 데아미나제 및 시티딘 데아미나제)는 유연한 루프 영역 또는 용매-노출된 영역에서의 napDNAbp에 삽입될 수 있다. 일부 실시형태에서, 데아미나제(예를 들어, 아데노신 데아미나제, 시티딘 데아미나제, 또는 아데노신 데아미나제 및 시티딘 데아미나제)는 Cas9 또는 Cas12b/C2c1 폴리펩티드의 유연한 루프에 삽입된다.

[0361] 일부 실시형태에서, 데아미나제(예를 들어, 아데노신 데아미나제, 시티딘 데아미나제, 또는 아데노신 데아미나제 및 시티딘 데아미나제)의 삽입 위치는 Cas9 폴리펩티드의 결정 구조의 B-인자 분석에 의해 결정된다. 일부 실시형태에서, 데아미나제(예를 들어, 아데노신 데아미나제, 시티딘 데아미나제, 또는 아데노신 데아미나제 및 시티딘 데아미나제)는 평균보다 높은 B-인자(예를 들어, 무질서한 영역을 포함하는 전체 단백질 또는 단백질 도메인과 비교하여 더 높은 B 인자)를 포함하는 Cas9 폴리펩티드의 영역에 삽입된다. B-인자 또는 온도 인자는 (예를 들어, 온도-의존성 원자 진동 또는 결정 격자에서 정적 무질서의 결과로서) 그의 평균 위치로부터 원자의 변동을 나타낼 수 있다. 골격 원자에 대한 높은 B-인자(예를 들어, 평균 B-인자보다 높음)는 상대적으로 높은 국소 이동성을 갖는 영역을 나타낼 수 있다. 이러한 영역은 구조나 기능을 손상함에 없이 데아미나제를 삽입하는데 사용될 수 있다. 데아미나제(예를 들어, 아데노신 데아미나제, 시티딘 데아미나제, 또는 아데노신 데아미나제 및 시티딘 데아미나제)는 총 단백질에 대한 평균 B-인자보다 50%, 60%, 70%, 80%, 90%, 100%, 110%, 120%, 130%, 140%, 150%, 160%, 170%, 180%, 190%, 200% 또는 200% 초과로 더 많은 B-인자가 있는 Cα 원자를 갖는 잔기가 있는 위치에서 삽입될 수 있다. 데아미나제(예를 들어, 아데노신 데아미나제, 시티딘 데아미나제, 또는 아데노신 데아미나제 및 시티딘 데아미나제)는 잔기를 포함하는 Cas9 단백질 도메인에 대한 평균 B-인자보다 50%, 60%, 70%, 80%, 90%, 100%, 110%, 120%, 130%, 140%, 150%, 160%, 170%, 180%, 190%, 200% 또는 200% 초과로 더 많은 B-인자가 있는 Cα 원자를 갖는 잔기가 있는 위치에서 삽입될 수 있다. 평균보다 높은 B-인자를 포함하는 Cas9 폴리펩티드 위치는 예를 들어, 상기 Cas9 참조 서열에서 번호가 매겨진 잔기 768, 792, 1052, 1015, 1022, 1026, 1029, 1067, 1040, 1054, 1068, 1246, 1247 및 1248을 포함할 수 있다. 평균보다 높은 B-인

자를 포함하는 Cas9 폴리펩티드 영역은 예를 들어, 상기 Cas9 참조 서열에서 번호가 매겨진 잔기 792-872, 792-906 및 2-791을 포함할 수 있다.

[0362] 이종 폴리펩티드(예를 들어, 데아미나제)는 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 768, 791, 792, 1015, 1016, 1022, 1023, 1026, 1029, 1040, 1052, 1054, 1067, 1068, 1069, 1246, 1247 및 1248 또는 기타 Cas9 폴리펩티드에서의 상응하는 아미노산 잔기로 구성된 군으로부터 선택된 아미노산 잔기에서 napDNAbp에 삽입될 수 있다. 일부 실시형태에서, 이종 폴리펩티드는 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 아미노산 위치 768-769, 791-792, 792-793, 1015-1016, 1022-1023, 1026-1027, 1029-1030, 1040-1041, 1052-1053, 1054-1055, 1067-1068, 1068-1069, 1247-1248, 또는 1248-1249 또는 이의 상응하는 아미노산 위치 사이에 삽입된다. 일부 실시형태에서, 이종 폴리펩티드는 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 아미노산 위치 769-770, 792-793, 793-794, 1016-1017, 1023-1024, 1027-1028, 1030-1031, 1041-1042, 1053-1054, 1055-1056, 1068-1069, 1069-1070, 1248-1249, 또는 1249-1250 또는 이의 상응하는 아미노산 위치 사이에 삽입된다. 일부 실시형태에서, 이종 폴리펩티드는 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 768, 791, 792, 1015, 1016, 1022, 1023, 1026, 1029, 1040, 1052, 1054, 1067, 1068, 1069, 1246, 1247, 및 1248로 구성된 군으로부터 선택된 아미노산 잔기, 또는 기타 Cas9 폴리펩티드에서의 상응하는 아미노산 잔기를 대체한다. 삽입 위치와 관련하여 상기 Cas9 참조 서열에 대한 참조는 예시적인 목적을 위한 것임을 이해해야 한다. 본 명세서에 논의된 바와 같은 삽입은 상기 Cas9 참조 서열의 Cas9 폴리펩티드 서열에 제한되지 않지만, 변이체 Cas9 폴리펩티드, 예를 들어 Cas9 닉카제(nCas9), 뉴클레아제 사멸 Cas9(dCas9), 뉴클레아제 도메인을 결하는 Cas9 변이체, 절삭된 Cas9, 또는 부분 또는 전체 HNH 도메인을 결하는 Cas9 도메인에서 상응하는 위치에서의 삽입을 포함한다.

[0363] 이종 폴리펩티드(예를 들어, 데아미나제)는 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 768, 792, 1022, 1026, 1040, 1068 및 1247로 구성된 군으로부터 선택된 아미노산 잔기, 또 기타 Cas9 폴리펩티드에서 상응하는 아미노산 잔기에서의 napDNAbp에 삽입될 수 있다. 일부 실시형태에서, 이종 폴리펩티드는 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 아미노산 위치 768-769, 792-793, 1022-1023, 1026-1027, 1029-1030, 1040-1041, 1068-1069, 또는 1247-1248, 또는 이의 상응하는 아미노산 위치 사이에 삽입된다. 일부 실시형태에서, 이종 폴리펩티드는 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 아미노산 위치 769-770, 793-794, 1023-1024, 1027-1028, 1030-1031, 1041-1042, 1069-1070, 또는 1248-1249, 또는 이의 상응하는 아미노산 위치 사이에 삽입된다. 일부 실시형태에서, 이종 폴리펩티드는 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 768, 792, 1022, 1026, 1040, 1068 및 1247로 구성된 군으로부터 선택된 아미노산 잔기, 또는 기타 Cas9 폴리펩티드에서 상응하는 아미노산 잔기를 대체한다.

[0364] 이종 폴리펩티드(예를 들어, 데아미나제)는 본 명세서에 기재된 바와 같은 아미노산 잔기, 또는 기타 Cas9 폴리펩티드에서 상응하는 아미노산 잔기에서의 napDNAbp에 삽입될 수 있다. 실시형태에서, 이종 폴리펩티드(예를 들어, 데아미나제)는 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 1002, 1003, 1025, 1052-1056, 1242-1247, 1061-1077, 943-947, 686-691, 569-578, 530-539, 및 1060-1077로 구성된 군으로부터 선택된 아미노산 잔기, 또는 기타 Cas9 폴리펩티드에서 상응하는 아미노산 잔기에서의 napDNAbp에 삽입될 수 있다. 데아미나제(예를 들어, 아데노신 데아미나제, 시티딘 데아미나제, 또는 아데노신 데아미나제 및 시티딘 데아미나제)는 잔기의 N-말단 또는 C-말단에서 삽입될 수 있거나 잔기를 대체할 수 있다. 일부 실시형태에서, 데아미나제(예를 들어, 아데노신 데아미나제, 시티딘 데아미나제, 또는 아데노신 데아미나제 및 시티딘 데아미나제)는 잔기의 C-말단에서 삽입된다.

[0365] 일부 실시형태에서, 아데노신 데아미나제(예를 들어, TadA)는 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 1015, 1022, 1029, 1040, 1068, 1247, 1054, 1026, 768, 1067, 1248, 1052, 및 1246으로 구성된 군으로부터 선택된 아미노산 잔기, 또는 기타 Cas9 폴리펩티드에서 상응하는 아미노산 잔기에서 삽입된다. 일부 실시형태에서, 아데노신 데아미나제(예를 들어, TadA)는 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 잔기 792-872, 792-906 또는 2-791, 또는 기타 Cas9 폴리펩티드에서 상응하는 아미노산 잔기의 장소에 삽입된다. 일부 실시형태에서, 아데노신 데아미나제는 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 1015, 1022, 1029, 1040, 1068, 1247, 1054, 1026, 768, 1067, 1248, 1052, 및 1246으로 구성된 군으로부터 선택된 아미노산, 또는 기타 Cas9 폴리펩티드에서 상응하는 아미노산 잔기의 N-말단에서 삽입된다. 일부 실시형태에서, 아데노신 데아미나제는 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 1015, 1022, 1029, 1040, 1068, 1247, 1054, 1026, 768, 1067, 1248, 1052, 및 1246으로 구성된 군으로부터 선택된 아미노산, 또는 기타 Cas9 폴리펩티드에서 상응하는 아미노산 잔기의 C-말단에서 삽입된다. 일부 실시형태에서, 아데노신 데아미나제는 상기 Cas9 참조 서열에서 번

호가 매겨진 바와 같은 1015, 1022, 1029, 1040, 1068, 1247, 1054, 1026, 768, 1067, 1248, 1052 및 1246으로 구성된 군으로부터 선택된 아미노산, 또는 기타 Cas9 폴리펩티드에서 상응하는 아미노산 잔기를 대체하기 위해 삽입된다.

[0366] 일부 실시형태에서, 시티딘 데아미나제(예를 들어, APOBEC1)는 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 1016, 1023, 1029, 1040, 1069 및 1247로 구성된 군으로부터 선택된 아미노산 잔기, 또는 기타 Cas9 폴리펩티드에서 상응하는 아미노산 잔기에서 삽입된다. 일부 실시형태에서, 시티딘 데아미나제는 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 1016, 1023, 1029, 1040, 1069 및 1247로 구성된 군으로부터 선택된 아미노산, 또는 기타 Cas9 폴리펩티드에서 상응하는 아미노산 잔기의 N-말단에서 삽입된다. 일부 실시형태에서 시티딘 데아미나제는 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 1016, 1023, 1029, 1040, 1069 및 1247로 구성된 군으로부터 선택된 아미노산, 또는 기타 Cas9 폴리펩티드에서 상응하는 아미노산 잔기의 C-말단에 삽입된다. 일부 실시형태에서 시티딘 데아미나제는 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 1016, 1023, 1029, 1040, 1069 및 1247로 구성된 군으로부터 선택된 아미노산, 또는 기타 Cas9 폴리펩티드에서 상응하는 아미노산 잔기를 대체하기 위해 삽입된다.

[0367] 일부 실시형태에서, 데아미나제(예를 들어, 아데노신 데아미나제, 시티딘 데아미나제, 또는 아데노신 데아미나제 및 시티딘 데아미나제)는 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 아미노산 잔기 768 또는 기타 Cas9 폴리펩티드에서 상응하는 아미노산 잔기에서 삽입된다. 일부 실시형태에서, 데아미나제(예를 들어, 아데노신 데아미나제, 시티딘 데아미나제, 또는 아데노신 데아미나제 및 시티딘 데아미나제)는 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 아미노산 잔기 768, 또는 기타 Cas9 폴리펩티드에서 상응하는 아미노산 잔기의 N-말단에서 삽입된다. 일부 실시형태에서, 데아미나제(예를 들어, 아데노신 데아미나제, 시티딘 데아미나제, 또는 아데노신 데아미나제 및 시티딘 데아미나제)는 상기 Cas9 참조 서열에서 번호가 바와 같은 아미노산 잔기 768, 또는 기타 Cas9 폴리펩티드에서 상응하는 아미노산 잔기의 C-말단에서 삽입된다. 일부 실시형태에서, 데아미나제(예를 들어, 아데노신 데아미나제, 시티딘 데아미나제, 또는 아데노신 데아미나제 및 시티딘 데아미나제)는 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 아미노산 잔기 768, 또는 기타 Cas9 폴리펩티드에서 상응하는 아미노산 잔기를 대체하기 위해 삽입된다.

[0368] 일부 실시형태에서, 데아미나제(예를 들어, 아데노신 데아미나제, 시티딘 데아미나제, 또는 아데노신 데아미나제 및 시티딘 데아미나제)는 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 아미노산 잔기 791에서 삽입되거나, 아미노산 잔기 792에서 삽입되거나, 또는 기타 Cas9 폴리펩티드에서 상응하는 아미노산 잔기에서 삽입된다. 일부 실시형태에서, 데아미나제(예를 들어, 아데노신 데아미나제, 시티딘 데아미나제, 또는 아데노신 데아미나제 및 시티딘 데아미나제)는 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 아미노산 잔기 791의 N-말단에서 삽입되거나, 아미노산 792의 N-말단에서 삽입되거나, 또는 기타 Cas9 폴리펩티드에서 상응하는 아미노산 잔기의 N-말단에서 삽입된다. 일부 실시형태에서, 데아미나제(예를 들어, 아데노신 데아미나제, 시티딘 데아미나제, 또는 아데노신 데아미나제 및 시티딘 데아미나제)는 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 아미노산 791의 C-말단에서 삽입되거나, 아미노산 792의 N-말단에서 삽입되거나, 또는 기타 Cas9 폴리펩티드에서 상응하는 아미노산 잔기에서 삽입된다. 일부 실시형태에서, 데아미나제(예를 들어, 아데노신 데아미나제, 시티딘 데아미나제, 또는 아데노신 데아미나제 및 시티딘 데아미나제)는 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 아미노산 791을 대체하기 위해 삽입되거나, 아미노산 잔기 792를 대체하기 위해 삽입되거나, 또는 기타 Cas9 폴리펩티드에서 상응하는 아미노산을 대체하기 위해 삽입된다.

[0369] 일부 실시형태에서, 데아미나제(예를 들어, 아데노신 데아미나제, 시티딘 데아미나제, 또는 아데노신 데아미나제 및 시티딘 데아미나제)는 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 아미노산 잔기 1016, 또는 기타 Cas9 폴리펩티드에서의 상응하는 아미노산 잔기에서 삽입된다. 일부 실시형태에서, 데아미나제(예를 들어, 아데노신 데아미나제, 시티딘 데아미나제, 또는 아데노신 데아미나제 및 시티딘 데아미나제)는 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 아미노산 잔기 1016, 또는 기타 Cas9 폴리펩티드에서의 상응하는 아미노산 잔기의 N-말단에서 삽입된다. 일부 실시형태에서, 데아미나제(예를 들어, 아데노신 데아미나제, 시티딘 데아미나제, 또는 아데노신 데아미나제 및 시티딘 데아미나제)는 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 아미노산 잔기 1016, 또는 기타 Cas9 폴리펩티드에서의 상응하는 아미노산 잔기의 C-말단에서 삽입된다. 일부 실시형태에서, 데아미나제(예를 들어, 아데노신 데아미나제, 시티딘 데아미나제, 또는 아데노신 데아미나제 및 시티딘 데아미나제)는 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 아미노산 잔기 1016, 또는 기타 Cas9 폴리펩티드에서의 상응하는 아미노산 잔기를 대체하기 위해 삽입된다.

[0370] 일부 실시형태에서, 데아미나제(예를 들어, 아데노신 데아미나제, 시티딘 데아미나제, 또는 아데노신 데아미나

제 및 시티딘 데아미나제)는 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 아미노산 잔기 1022에서 삽입되거나, 아미노산 잔기 1023에서 삽입되거나, 또는 기타 Cas9 폴리펩티드에서의 상응하는 아미노산 잔기에서 삽입된다. 일부 실시형태에서, 데아미나제(예를 들어, 아데노신 데아미나제, 시티딘 데아미나제, 또는 아데노신 데아미나제 및 시티딘 데아미나제)는 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 아미노산 잔기 1022의 N-말단에서 삽입되거나, 아미노산 잔기 1023의 N-말단에 삽입되거나, 또는 기타 Cas9 폴리펩티드에서의 상응하는 아미노산 잔기에서 삽입된다. 일부 실시형태에서, 데아미나제(예를 들어, 아데노신 데아미나제, 시티딘 데아미나제, 또는 아데노신 데아미나제 및 시티딘 데아미나제)는 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 아미노산 잔기 1022의 C-말단에서 삽입되거나, 아미노산 잔기 1023의 C-말단에 삽입되거나, 또는 기타 Cas9 폴리펩티드에서의 상응하는 아미노산 잔기에서 삽입된다. 일부 실시형태에서, 데아미나제(예를 들어, 아데노신 데아미나제, 시티딘 데아미나제, 또는 아데노신 데아미나제 및 시티딘 데아미나제)는 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 아미노산 잔기 1022를 대체하기 위해 삽입되거나, 아미노산 잔기 1023을 대체하기 위해 삽입되거나, 또는 기타 Cas9 폴리펩티드에서의 상응하는 아미노산 잔기를 대체하기 위해 삽입된다.

[0371] 일부 실시형태에서, 데아미나제(예를 들어, 아데노신 데아미나제, 시티딘 데아미나제, 또는 아데노신 데아미나제 및 시티딘 데아미나제)는 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 아미노산 잔기 1026에서 삽입되거나, 아미노산 잔기 1029에서 삽입되거나, 또는 기타 Cas9 폴리펩티드에서의 상응하는 아미노산 잔기에서 삽입된다. 일부 실시형태에서, 데아미나제(예를 들어, 아데노신 데아미나제, 시티딘 데아미나제, 또는 아데노신 데아미나제 및 시티딘 데아미나제)는 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 아미노산 잔기 1026의 N-말단에서 삽입되거나, 아미노산 잔기 1029의 N-말단에 삽입되거나, 또는 기타 Cas9 폴리펩티드에서의 상응하는 아미노산 잔기에서 삽입된다. 일부 실시형태에서, 데아미나제(예를 들어, 아데노신 데아미나제, 시티딘 데아미나제, 또는 아데노신 데아미나제 및 시티딘 데아미나제)는 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 아미노산 잔기 1026의 C-말단에서 삽입되거나, 아미노산 잔기 1029의 C-말단에 삽입되거나, 또는 기타 Cas9 폴리펩티드에서의 상응하는 아미노산 잔기에서 삽입된다. 일부 실시형태에서, 데아미나제(예를 들어, 아데노신 데아미나제, 시티딘 데아미나제, 또는 아데노신 데아미나제 및 시티딘 데아미나제)는 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 아미노산 잔기 1026을 대체하기 위해 삽입되거나, 아미노산 잔기 1029를 대체하기 위해 삽입되거나, 또는 기타 Cas9 폴리펩티드에서의 상응하는 아미노산 잔기를 대체하기 위해 삽입된다.

[0372] 일부 실시형태에서, 데아미나제(예를 들어, 아데노신 데아미나제, 시티딘 데아미나제, 또는 아데노신 데아미나제 및 시티딘 데아미나제)는 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 아미노산 잔기 1040, 또는 기타 Cas9 폴리펩티드에서의 상응하는 아미노산 잔기에서 삽입된다. 일부 실시형태에서, 데아미나제(예를 들어, 아데노신 데아미나제, 시티딘 데아미나제, 또는 아데노신 데아미나제 및 시티딘 데아미나제)는 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 아미노산 잔기 1040, 또는 기타 Cas9 폴리펩티드에서의 상응하는 아미노산 잔기의 N-말단에서 삽입된다. 일부 실시형태에서, 데아미나제(예를 들어, 아데노신 데아미나제, 시티딘 데아미나제, 또는 아데노신 데아미나제 및 시티딘 데아미나제)는 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 아미노산 잔기 1040, 또는 기타 Cas9 폴리펩티드에서의 상응하는 아미노산 잔기의 C-말단에서 삽입된다. 일부 실시형태에서, 데아미나제(예를 들어, 아데노신 데아미나제, 시티딘 데아미나제, 또는 아데노신 데아미나제 및 시티딘 데아미나제)는 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 아미노산 잔기 1040, 또는 기타 Cas9 폴리펩티드에서의 상응하는 아미노산 잔기를 대체하기 위해 삽입된다.

[0373] 일부 실시형태에서, 데아미나제(예를 들어, 아데노신 데아미나제, 시티딘 데아미나제, 또는 아데노신 데아미나제 및 시티딘 데아미나제)는 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 아미노산 잔기 1052에서 삽입되거나, 아미노산 잔기 1054에서 삽입되거나, 또는 기타 Cas9 폴리펩티드에서의 상응하는 아미노산 잔기에서 삽입된다. 일부 실시형태에서, 데아미나제(예를 들어, 아데노신 데아미나제, 시티딘 데아미나제, 또는 아데노신 데아미나제 및 시티딘 데아미나제)는 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 아미노산 잔기 1052의 N-말단에서 삽입되거나, 아미노산 잔기 1054의 N-말단에 삽입되거나, 또는 기타 Cas9 폴리펩티드에서의 상응하는 아미노산 잔기에서 삽입된다. 일부 실시형태에서, 데아미나제(예를 들어, 아데노신 데아미나제, 시티딘 데아미나제, 또는 아데노신 데아미나제 및 시티딘 데아미나제)는 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 아미노산 잔기 1052의 C-말단에서 삽입되거나, 아미노산 잔기 1054의 C-말단에 삽입되거나, 또는 기타 Cas9 폴리펩티드에서의 상응하는 아미노산 잔기에서 삽입된다. 일부 실시형태에서, 데아미나제(예를 들어, 아데노신 데아미나제, 시티딘 데아미나제, 또는 아데노신 데아미나제 및 시티딘 데아미나제)는 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 아미노산 잔기 1052를 대체하기 위해 삽입되거나, 아미노산 잔기 1054를 대체하기 위해 삽입되거나, 또는 기타 Cas9 폴리펩티드에서의 상응하는 아미노산 잔기를 대체하기 위해 삽입된다.

[0374] 일부 실시형태에서, 데아미나제(예를 들어, 아데노신 데아미나제, 시티딘 데아미나제, 또는 아데노신 데아미나제 및 시티딘 데아미나제)는 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 아미노산 잔기 1067에서 삽입되거나, 아미노산 잔기 1068에서 삽입되거나, 아미노산 잔기 1069에서 삽입되거나, 또는 기타 Cas9 폴리펩티드에서의 상응하는 아미노산 잔기에서 삽입된다. 일부 실시형태에서, 데아미나제(예를 들어, 아데노신 데아미나제, 시티딘 데아미나제, 또는 아데노신 데아미나제 및 시티딘 데아미나제)는 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 아미노산 잔기 1067의 N-말단에서 삽입되거나, 아미노산 잔기 1068의 N-말단에 삽입되거나, 아미노산 잔기 1069의 N-말단에 삽입되거나, 또는 기타 Cas9 폴리펩티드에서의 상응하는 아미노산 잔기에서 삽입된다. 일부 실시형태에서, 데아미나제(예를 들어, 아데노신 데아미나제, 시티딘 데아미나제, 또는 아데노신 데아미나제 및 시티딘 데아미나제)는 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 아미노산 잔기 1067의 C-말단에서 삽입되거나, 아미노산 잔기 1068의 C-말단에 삽입되거나, 아미노산 잔기 1069의 C-말단에 삽입되거나, 또는 기타 Cas9 폴리펩티드에서의 상응하는 아미노산 잔기에서 삽입된다. 일부 실시형태에서, 데아미나제(예를 들어, 아데노신 데아미나제, 시티딘 데아미나제, 또는 아데노신 데아미나제 및 시티딘 데아미나제)는 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 아미노산 잔기 1067을 대체하기 위해 삽입되거나, 아미노산 잔기 1068을 대체하기 위해 삽입되거나, 아미노산 잔기 1069를 대체하기 위해 삽입되거나, 또는 기타 Cas9 폴리펩티드에서의 상응하는 아미노산 잔기를 대체하기 위해 삽입된다.

[0375] 일부 실시형태에서, 데아미나제(예를 들어, 아데노신 데아미나제, 시티딘 데아미나제, 또는 아데노신 데아미나제 및 시티딘 데아미나제)는 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 아미노산 잔기 1246에서 삽입되거나, 아미노산 잔기 1247에서 삽입되거나, 아미노산 잔기 1248에서 삽입되거나, 또는 기타 Cas9 폴리펩티드에서의 상응하는 아미노산 잔기에서 삽입된다. 일부 실시형태에서, 데아미나제(예를 들어, 아데노신 데아미나제, 시티딘 데아미나제, 또는 아데노신 데아미나제 및 시티딘 데아미나제)는 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 아미노산 잔기 1246의 N-말단에서 삽입되거나, 아미노산 잔기 1247의 N-말단에 삽입되거나, 아미노산 잔기 1248의 N-말단에 삽입되거나, 또는 기타 Cas9 폴리펩티드에서의 상응하는 아미노산 잔기에서 삽입된다. 일부 실시형태에서, 데아미나제(예를 들어, 아데노신 데아미나제, 시티딘 데아미나제, 또는 아데노신 데아미나제 및 시티딘 데아미나제)는 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 아미노산 잔기 1246의 C-말단에서 삽입되거나, 아미노산 잔기 1247의 C-말단에 삽입되거나, 아미노산 잔기 1248의 C-말단에 삽입되거나, 또는 기타 Cas9 폴리펩티드에서의 상응하는 아미노산 잔기에서 삽입된다. 일부 실시형태에서, 데아미나제(예를 들어, 아데노신 데아미나제, 시티딘 데아미나제, 또는 아데노신 데아미나제 및 시티딘 데아미나제)는 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 아미노산 잔기 1246을 대체하기 위해 삽입되거나, 아미노산 잔기 1247을 대체하기 위해 삽입되거나, 아미노산 잔기 1248을 대체하기 위해 삽입되거나, 또는 기타 Cas9 폴리펩티드에서의 상응하는 아미노산 잔기를 대체하기 위해 삽입된다.

[0376] 일부 실시형태에서, 이종 폴리펩티드(예를 들어, 데아미나제)는 Cas9 폴리펩티드의 유연한 루프에 삽입된다. 유연한 루프 부분은 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 530-537, 569-570, 686-691, 943-947, 1002-1025, 1052-1077, 1232-1247 또는 1298-1300, 또는 기타 Cas9 폴리펩티드에서의 상응하는 아미노산 잔기로 구성된 군에서 선택될 수 있다. 유연한 루프 부분은 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 1-529, 538-568, 580-685, 692-942, 948-1001, 1026-1051, 1078-1231 또는 1248-1297, 또는 기타 Cas9 폴리펩티드에서의 상응하는 아미노산 잔기로 구성된 군에서 선택될 수 있다.

[0377] 이종 폴리펩티드(예를 들어, 아데닌 데아미나제)는 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 아미노산 잔기: 1017-1069, 1242-1247, 1052-1056, 1060-1077, 1002 - 1003, 943-947, 530-537, 568-579, 686-691, 1242-1247, 1298 - 1300, 1066-1077, 1052-1056, 또는 1060-1077, 또는 기타 Cas9 폴리펩티드에서의 상응하는 아미노산 잔기에 상응하는 Cas9 폴리펩티드 영역 안으로 삽입될 수 있다.

[0378] 이종 폴리펩티드(예를 들어, 아데닌 데아미나제)는 Cas9 폴리펩티드의 결실된 영역 대신 삽입될 수 있다. 결실된 영역은 Cas9 폴리펩티드의 N-말단 또는 C-말단 부분에 상응할 수 있다. 일부 실시형태에서, 결실된 영역은 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 잔기 792-872, 또는 기타 Cas9 폴리펩티드에서의 상응하는 아미노산 잔기에 상응한다. 일부 실시형태에서, 결실된 영역은 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 잔기 792-906, 또는 기타 Cas9 폴리펩티드에서의 상응하는 아미노산 잔기에 상응한다. 일부 실시형태에서, 결실된 영역은 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 잔기 2-791, 또는 기타 Cas9 폴리펩티드에서의 상응하는 아미노산 잔기에 상응한다. 일부 실시형태에서, 결실된 영역은 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 잔기 1017-1069, 또는 이의 상응하는 아미노산 잔기에 상응한다.

[0379] 예시적인 내부 융합 염기 편집기가 아래 표 3에 제공된다:

[0380] 표 3: Cas9 단백질에서 삽입 좌위

BE ID	변형	기타 ID
IBE001	Cas9 TadA ins 1015	ISLAY01
IBE002	Cas9 TadA ins 1022	ISLAY02
IBE003	Cas9 TadA ins 1029	ISLAY03
IBE004	Cas9 TadA ins 1040	ISLAY04
IBE005	Cas9 TadA ins 1068	ISLAY05
IBE006	Cas9 TadA ins 1247	ISLAY06
IBE007	Cas9 TadA ins 1054	ISLAY07
IBE008	Cas9 TadA ins 1026	ISLAY08
IBE009	Cas9 TadA ins 768	ISLAY09
IBE020	델타 HNH TadA 792	ISLAY20
IBE021	N-term 융합 단일 TadA 나선 절삭된 165-말단	ISLAY21
IBE029	TadA-원형 순열체 116 ins1067	ISLAY29
IBE031	TadA- 원형 순열체 136 ins1248	ISLAY31
IBE032	TadA- 원형 순열체 136ins 1052	ISLAY32
IBE035	델타 792-872 TadA ins	ISLAY35
IBE036	델타 792-906 TadA ins	ISLAY36
IBE043	TadA-원형 순열체 65 ins1246	ISLAY43
IBE044	TadA ins C-term 절삭 2 791	ISLAY44

[0381]

[0382] 이종 폴리펩티드(예를 들어, 데아미나제)는 Cas9 폴리펩티드의 구조적 또는 기능적 도메인 내에 삽입될 수 있다. 이종 폴리펩티드(예를 들어, 데아미나제)는 Cas9 폴리펩티드의 2개의 구조적 또는 기능적 도메인 사이에 삽입될 수 있다. 이종 폴리펩티드(예를 들어, 데아미나제)는, 예를 들어, Cas9 폴리펩티드로부터 도메인을 결실한 후 Cas9 폴리펩티드의 구조적 또는 기능적 도메인 대신 삽입될 수 있다. Cas9 폴리펩티드의 구조적 또는 기능적 도메인은, 예를 들어, RuvC I, RuvC II, RuvC III, Rec1, Rec2, PI 또는 HNH를 포함할 수 있다.

[0383] 일부 실시형태에서, Cas9 폴리펩티드는 RuvC I, RuvC II, RuvC III, Rec1, Rec2, PI, 또는 HNH 도메인으로 구성된 군으로부터 선택된 하나 이상의 도메인을 결한다. 일부 실시형태에서 Cas9 폴리펩티드는 뉴클레아제 도메인을 결한다. 일부 실시형태에서, Cas9 폴리펩티드는 HNH 도메인을 결한다. 일부 실시형태에서, Cas9 폴리펩티드는 HNH 도메인의 일부를 결하여 Cas9 폴리펩티드는 감소되거나 폐지된 HNH 활성을 갖는다. 일부 실시형태에서 Cas9 폴리펩티드는 뉴클레아제 도메인의 결실을 포함하고, 데아미나아제가 뉴클레아제 도메인을 대체하기 위해 삽입된다. 일부 실시형태에서, HNH 도메인은 결실되고 데아미나제가 그 자리에 삽입된다. 일부 실시형태에서, 하나 이상의 RuvC 도메인이 결실되고 데아미나제가 그 자리에 삽입된다.

[0384] 이종 폴리펩티드를 포함하는 융합 단백질은 napDNAbp의 N-말단 및 C-말단 단편에 의해 측접될 수 있다. 일부 실시형태에서, 융합 단백질은 Cas9 폴리펩티드의 N-말단 단편 및 C-말단 단편에 의해 측접된 데아미나제를 포함한다. N 말단 단편 또는 C 말단 단편은 표적 폴리뉴클레오티드 서열에 결합할 수 있다. N 말단 단편의 C-말단 또는 C 말단 단편의 N-말단은 Cas9 폴리펩티드의 유연한 루프의 일부를 포함할 수 있다. N 말단 단편의 C-말단 또는 C 말단 단편의 N-말단은 Cas9 폴리펩티드의 알파-나선 구조의 일부를 포함할 수 있다. N-말단 단편 또는 C-말단 단편은 DNA 결합 도메인을 포함할 수 있다. N-말단 단편 또는 C-말단 단편은 RuvC 도메인을 포함할 수 있다. N-말단 단편 또는 C-말단 단편은 HNH 도메인을 포함할 수 있다. 일부 실시형태에서, N-말단 단편 및 C-말단 단편 중 어느 것도 HNH 도메인을 포함하지 않는다.

[0385] 일부 실시형태에서, N 말단 Cas9 단편의 C-말단은 융합 단백질이 표적 핵염기를 탈아미노화할 때 표적 핵염기에

근접한 아미노산을 포함한다. 일부 실시형태에서, C 말단 Cas9 단편의 N-말단은 융합 단백질이 표적 핵염기를 탈아미노화할 때 표적 핵염기에 근접한 아미노산을 포함한다. N 말단 Cas9 단편의 C-말단 또는 C 말단 Cas9 단편의 N-말단에서의 아미노산과 표적 핵염기 사이에 근접성을 갖기 위해 다양한 데아미나제의 삽입 위치가 다를 수 있다. 예를 들어, 데아미나제의 삽입 위치는 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 1015, 1022, 1029, 1040, 1068, 1247, 1054, 1026, 768, 1067, 1248, 1052 및 1246으로 구성된 군으로부터 선택된 아미노산 잔기, 또는 기타 Cas9 폴리펩티드에서의 상응하는 아미노산 잔기에 있을 수 있다.

[0386] 융합 단백질의 N-말단 Cas9 단편(즉, 융합 단백질에서 데아미나제에 측접하는 N-말단 Cas9 단편)은 Cas9 폴리펩티드의 N-말단을 포함할 수 있다. 융합 단백질의 N-말단 Cas9 단편은 적어도 약: 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200 또는 1300개 아미노산의 길이를 포함할 수 있다. 융합 단백질의 N-말단 Cas9 단편은 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 아미노산 잔기: 1-56, 1-95, 1-200, 1-300, 1-400, 1-500, 1-600, 1-700, 1-718, 1-765, 1-780, 1-906, 1-918 또는 1-1100, 또는 기타 Cas9 폴리펩티드에서의 상응하는 아미노산 잔기에 상응하는 서열을 포함할 수 있다. N-말단 Cas9 단편은 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 아미노산 잔기: 1-56, 1-95, 1-200, 1-300, 1-400, 1-500, 1-600, 1-700, 1-718, 1-765, 1-780, 1-906, 1-918 또는 1-1100, 또는 기타 Cas9 폴리펩티드에서의 상응하는 아미노산 잔기와 적어도 85%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 적어도 99.5% 서열 동일성을 포함하는 서열을 포함할 수 있다.

[0387] 융합 단백질의 C-말단 Cas9 단편(즉, 융합 단백질에서 데아미나제에 측접하는 C-말단 Cas9 단편)은 Cas9 폴리펩티드의 C-말단을 포함할 수 있다. 융합 단백질의 C-말단 Cas9 단편은 적어도 약: 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200 또는 1300개 아미노산의 길이를 포함할 수 있다. 융합 단백질의 C-말단 Cas9 단편은 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 아미노산 잔기: 1099-1368, 918-1368, 906-1368, 780-1368, 765-1368, 718-1368, 94-1368, 또는 56-1368, 또는 기타 Cas9 폴리펩티드에서의 상응하는 아미노산 잔기에 상응하는 서열을 포함할 수 있다. N-말단 Cas9 단편은 상기 Cas9 참조 서열에서 번호가 매겨진 바와 같은 아미노산 잔기: 1099-1368, 918-1368, 906-1368, 780-1368, 765-1368, 718-1368, 94-1368, 또는 56-1368, 또는 기타 Cas9 폴리펩티드에서의 상응하는 아미노산 잔기와 적어도 85%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 적어도 99.5% 서열 동일성을 포함하는 서열을 포함할 수 있다.

[0388] 함께 취해진 융합 단백질의 N-말단 Cas9 단편 및 C-말단 Cas9 단편은, 예를 들어, 상기 Cas9 참조 서열에 제시된 바와 같은 전장 자연적으로 발생하는 Cas9 폴리펩티드 서열에 상응하지 않을 수 있다.

[0389] 본 명세서에 기재된 융합 단백질은 비-표적 부위(예를 들어, 표적-외 부위)에서 감소된 탈아미노화, 예컨대 감소된 게놈 전체 의사의 탈아미노화로 표적화된 탈아미노화를 일으킬 수 있다. 본 명세서에 기재된 융합 단백질은 비-표적 부위에서 감소 방관자 탈아미노화로 표적화된 탈아민화를 일으킬 수 있다. 바람직하지 않은 탈아미노화 또는 표적-외 탈아미노화는, 예를 들어, Cas9 폴리펩티드의 N 말단 또는 C 말단에 융합된 데아미나제를 포함하는 끝 말단 융합 단백질과 비교하여 적어도 30%, 적어도 40%, 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 90%, 적어도 95%, 또는 적어도 99% 만큼 감소될 수 있다. 바람직하지 않은 탈아미노화 또는 표적-외 탈아미노화는, 예를 들어, Cas9 폴리펩티드의 N 말단 또는 C 말단에 융합된 데아미나제를 포함하는 끝 말단 융합 단백질과 비교하여 적어도 1-배, 적어도 2-배, 적어도 3-배, 적어도 4-배, 적어도 5-배, 적어도 10배, 적어도 15배, 적어도 20배, 적어도 30배, 적어도 40배, 적어도 50배, 적어도 60배, 적어도 70배, 적어도 80배, 적어도 90배, 또는 적어도 100배 만큼 감소될 수 있다.

[0390] 일부 실시형태에서, 융합 단백질의 데아미나제(예를 들어, 아데노신 데아미나제, 시티딘 데아미나제, 또는 아데노신 데아미나제 및 시티딘 데아미나제)는 R-루프의 범위 내에서 2개 이하의 핵염기를 탈아미노화한다. 일부 실시형태에서, 융합 단백질의 데아미나제는 R-루프의 범위 내에서 3개 이하의 핵염기를 탈아미노화한다. 일부 실시형태에서, 융합 단백질의 데아미나제는 R-루프의 범위 내에서 2, 3, 4, 5, 6, 7, 8, 9 또는 10개 이하의 핵염기를 탈아미노화한다. R-루프는 DNA:RNA 하이브리드, DNA:DNA 또는 RNA:RNA 상보적 구조를 포함하고 단일-가닥 DNA와 연관된 3-가닥 핵산 구조이다. 본 명세서에 사용된 바와 같이, R-루프는 표적 폴리뉴클레오티드가 CRISPR 복합체 또는 염기 편집 복합체와 접촉될 때 형성될 수 있으며, 여기서 가이드 폴리뉴클레오티드, 예를 들어, 가이드 RNA의 일부는 표적 폴리뉴클레오티드, 예를 들어 표적 DNA의 부분과 혼성화하고 치환된다. 일부 실시형태에서, R-루프는 스페이서 서열 및 표적 DNA 상보적 서열의 혼성화된 영역을 포함한다. R-루프 영역은 길이가 약 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 또는 50개의 핵염기 쌍의 것일 수

있다. 일부 실시형태에서, R-루프 영역은 길이가 약 20개의 핵염기 쌍이다. 본 명세서에 사용된 바와 같이, R-루프 영역은 가이드 폴리뉴클레오티드와 혼성화하는 표적 DNA 가닥으로 제한되지 않는다는 것이 이해되어야 한다. 예를 들어, R-루프 영역 내 표적 핵염기의 편집은 가이드 RNA에 대한 상보적 가닥을 포함하는 DNA 가닥에 대한 것일 수도 있고, 가이드 RNA에 상보성인 가닥의 반대 가닥인 DNA 가닥에 대한 것일 수도 있다. 일부 실시형태에서, R-루프의 영역에서 편집은 표적 DNA 서열에서 가이드 RNA에 대한 비-상보성 가닥(프로토스페이서 가닥) 상의 핵염기를 편집하는 것을 포함한다.

[0391] 본 명세서에 기재된 융합 단백질은 규범적 염기 편집과 상이한 편집 윈도우에서 표적 탈아미노화를 일으킬 수 있다. 일부 실시형태에서, 표적 핵염기는 표적 폴리뉴클레오티드 서열에서 PAM 서열 상류에 있는 약 1 내지 약 20개 염기이다. 일부 실시형태에서, 표적 핵염기는 표적 폴리뉴클레오티드 서열에서 PAM 서열 상류에 있는 약 2 내지 약 12개 염기이다. 일부 실시형태에서, 표적 핵염기는 PAM 서열의 상류 또는 떨어져 있는 약 1 내지 9 염기쌍, 약 2 내지 10 염기쌍, 약 3 내지 11 염기쌍, 약 4 내지 12 염기쌍, 약 5 내지 13 염기쌍, 약 6 내지 14 염기쌍, 약 7 내지 15 염기쌍, 약 8 내지 16 염기쌍, 약 9 내지 17 염기쌍, 약 10 내지 18 염기쌍, 약 11 내지 19 염기쌍, 약 12 내지 20 염기쌍, 약 1 내지 7 염기쌍, 약 2 내지 8 염기쌍, 약 3 내지 9 염기쌍, 약 4 내지 10 염기쌍, 약 5 내지 11 염기쌍, 약 6 내지 12 염기쌍, 약 7 내지 13 염기쌍, 약 8 내지 14 염기쌍, 약 9 내지 15 염기쌍, 약 10 내지 16 염기쌍, 약 11 내지 17 염기쌍, 약 12 내지 18 염기쌍, 약 13 내지 19 염기쌍, 약 14 내지 20 염기쌍, 약 1 내지 5 염기쌍, 약 2 내지 6 염기쌍, 약 3 내지 7 염기쌍, 약 4 내지 8 염기쌍, 약 5 내지 9 염기쌍, 약 6 내지 10 염기쌍, 약 7 내지 11 염기쌍, 약 8 내지 12 염기쌍, 약 9 내지 13 염기쌍, 약 10 내지 14 염기쌍, 약 11 내지 15 염기쌍, 약 12 내지 16 염기쌍, 약 13 내지 17 염기쌍, 약 14 내지 18 염기쌍, 약 15 내지 19 염기쌍, 약 16 내지 20 염기쌍, 약 1 내지 3 염기쌍, 약 2 내지 4 염기쌍, 약 3 내지 5 염기쌍, 약 4 내지 6 염기쌍, 약 5 내지 7 염기쌍, 약 6 내지 8 염기쌍, 약 7 내지 9 염기쌍, 약 8 내지 10 염기쌍, 약 9 내지 11 염기쌍, 약 10 내지 12 염기쌍, 약 11 내지 13 염기쌍, 약 12 내지 14 염기쌍, 약 13 내지 15 염기쌍, 약 14 내지 16 염기쌍, 약 15 내지 17 염기쌍, 약 16 내지 18 염기쌍, 약 17 내지 19 염기쌍, 약 18 내지 20 염기쌍이다. 일부 실시형태에서, 표적 핵염기는 PAM 서열의 상류 또는 이로부터 떨어져 있는 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 이상의 염기쌍이다. 일부 실시형태에서, 표적 핵염기는 PAM 서열의 상류에 있는 약 1, 2, 3, 4, 5, 6, 7, 8 또는 9 염기쌍이다. 일부 실시형태에서, 표적 핵염기는 PAM 서열의 상류에 있는 약 2, 3, 4 또는 6 염기쌍이다.

[0392] 융합 단백질은 하나 초과의 이종 폴리펩티드를 포함할 수 있다. 예를 들어, 융합 단백질은 하나 초과의 UGI 도메인 및/또는 하나 초과의 핵 국소화 신호를 추가로 포함할 수 있다. 2개 이상의 이종 도메인이 나란히 삽입될 수 있다. 2개 이상의 이종 도메인은 NapDNAbp에서 나란히 되지 않는 장소에서 삽입될 수 있다.

[0393] 융합 단백질은 데아미나제와 napDNAbp 폴리펩티드 사이에 링커를 포함할 수 있다. 링커는 펩티드 또는 비-펩티드 링커일 수 있다. 예를 들어, 링커는 XTEN, (GGGS)n(서열번호: 1308), (GGGGS)n(서열번호: 109), (G)n, (EAAAK)n(서열번호: 1309), (GGS)n, SGSETPGTSESATPES(서열번호 56)일 수 있다. 일부 실시형태에서, 융합 단백질은 N-말단 Cas9 단편과 데아미나제 사이에 링커를 포함한다. 일부 실시형태에서, 융합 단백질은 C-말단 Cas9 단편과 데아미나제 사이에 링커를 포함한다. 일부 실시형태에서, napDNAbp의 N-말단 및 C-말단 단편은 링커로 데아미나제에 연결된다. 일부 실시형태에서, N-말단 및 C-말단 단편은 링커 없이 데아미나제 도메인에 조인된다. 일부 실시형태에서, 융합 단백질은 N-말단 Cas9 단편과 데아미나제 사이에 링커를 포함하지만, C-말단 Cas9 단편과 데아미나제 사이에 링커를 포함하지 않는다. 일부 실시형태에서, 융합 단백질은 C-말단 Cas9 단편과 데아미나제 사이에 링커를 포함하지만, N-말단 Cas9 단편과 데아미나제 사이에 링커를 포함하지 않는다.

[0394] 일부 실시형태에서, 융합 단백질에서의 napDNAbp는 Cas12 폴리펩티드, 예를 들어 Cas12b/C2c1 또는 이의 단편이다. Cas12 폴리펩티드는 변이체 Cas12 폴리펩티드일 수 있다. 다른 실시형태에서, Cas12 폴리펩티드의 N- 또는 C-말단 단편은 핵산 프로그램가능 DNA 결합 도메인 또는 RuvC 도메인을 포함한다. 다른 실시형태에서, 융합 단백질은 Cas12 폴리펩티드와 촉매 도메인 사이에 링커를 함유한다. 다른 실시형태에서, 링커의 아미노산 서열은 GGSGGS(서열번호: 273) 또는 GSSGSETPGTSESATPESSG(서열번호: 1310)이다. 다른 실시형태에서, 링커는 강성 링커이다. 상기 양태의 다른 실시형태에서, 링커는 GGAGGCTCTGGAGGAAGC(서열번호: 1311) 또는 GGCTCTTCTGGATCTGAAACACCTGGCACAAGCGAGAGCGCCACCCCTGAGAGCTCTGGC(서열번호: 1312)에 의해 인코딩된다.

[0395] Cas12 폴리펩티드의 N- 및 C-말단 단편에 의해 측접된 이종 촉매 도메인을 포함하는 융합 단백질은 또한 본 명세서에 기재된 방법에서의 염기 편집에 유용하다. Cas12 및 하나 이상의 데아미나제 도메인, 예를 들어 아데노신 데아미나제를 포함하거나 Cas12 서열에 의해 측접된 아데노신 데아미나제 도메인을 포함하는 융합 단백질은 또한 표적 서열의 고도로 특이적이고 효율적인 염기 편집에 유용하다. 실시형태에서, 키메라 Cas12 융합 단백질

은 Cas12 폴리펩티드 내에 삽입된 이종 촉매 도메인(예를 들어, 아데노신 데아미나제, 시티딘 데아미나제, 또는 아데노신 데아미나제 및 시티딘 데아미나제)을 함유한다. 일부 실시형태에서, 융합 단백질은 Cas12 내에 삽입된 아데노신 데아미나제 도메인 및 시티딘 데아미나제 도메인을 포함한다. 일부 실시형태에서, 아데노신 데아미나제는 Cas12 내에 융합되고 시티딘 데아미나제는 C-말단에 융합된다. 일부 실시형태에서, 아데노신 데아미나제는 Cas12 내에 융합되고 시티딘 데아미나제는 N-말단에 융합된다. 일부 실시형태에서 시티딘 데아미나제는 Cas12 내에 융합되고 아데노신 데아미나제는 C-말단에 융합된다. 일부 실시형태에서, 시티딘 데아미나제는 Cas12 내에 융합되고 아데노신 데아미나제는 N-말단에 융합된다. 아데노신 데아미나제와 시티딘 데아미나제 및 Cas12를 갖는 융합 단백질의 예시적인 구조는 다음과 같이 제공된다:

[0396] NH2-[Cas12(아데노신 데아미나제)]-[시티딘 데아미나제]-COOH;

[0397] NH2-[시티딘 데아미나제]-[Cas12(아데노신 데아미나제)]-COOH;

[0398] NH2-[Cas12(시티딘 데아미나제)]-[아데노신 데아미나제]-COOH; 또는

[0399] NH2-[아데노신 데아미나제]-[Cas12(시티딘 데아미나제)]-COOH;

[0400] 일부 실시형태에서, 상기의 일반 아키텍처에서 사용된 "-"는 선택적인 링커의 존재를 나타낸다.

[0401] 다양한 실시형태에서, 촉매 도메인은 아데노신 데아미나제 활성과 같은 DNA 변형 활성(예를 들어, 데아미나제 활성)을 갖는다. 일부 실시형태에서, 아데노신 데아미나제는 TadA(예를 들어, TadA*7.10)이다. 일부 실시형태에서, TadA는 TadA*8이다. 일부 실시형태에서, TadA*8은 Cas12 내에 융합되고 시티딘 데아미나제는 C-말단에 융합된다. 일부 실시형태에서, TadA*8은 Cas12 내에 융합되고 시티딘 데아미나제는 N-말단에 융합된다. 일부 실시형태에서 시티딘 데아미나제는 Cas12 내에 융합되고 TadA*8은 C-말단에 융합된다. 일부 실시형태에서, 시티딘 데아미나제는 Cas12 내에 융합되고 TadA*8은 N-말단에 융합된다. TadA*8과 시티딘 데아미나제 및 Cas12를 갖는 융합 단백질의 예시적인 구조는 다음과 같이 제공된다:

[0402] N-[Cas12(TadA*8)]-[시티딘 데아미나제]-C;

[0403] N-[시티딘 데아미나제]-[Cas12(TadA*8)]-C;

[0404] N-[Cas12(시티딘 데아미나제)]-[TadA*8]-C; 또는

[0405] N-[TadA*8]-[Cas12(시티딘 데아미나제)]-C.

[0406] 일부 실시형태에서, 상기의 일반 아키텍처에서 사용된 "-"는 선택적인 링커의 존재를 나타낸다.

[0407] 다른 실시형태에서, 융합 단백질은 하나 이상의 촉매 도메인을 함유한다. 다른 실시형태에서, 하나 이상의 촉매 도메인 중 적어도 하나는 Cas12 폴리펩티드 내에 삽입되거나 Cas12 N-말단 또는 C-말단에 융합된다. 다른 실시형태에서, 하나 이상의 촉매 도메인 중 적어도 하나는 Cas12 폴리펩티드의 루프, 알파 나선 영역, 구조화되지 않은 부분, 또는 용매 접근가능한 부분 내에 삽입된다. 다른 실시형태에서, Cas12 폴리펩티드는 Cas12a, Cas12b, Cas12c, Cas12d, Cas12e, Cas12g, Cas12h, Cas12i, 또는 Cas12j/CasΦ이다. 다른 실시형태에서, Cas12 폴리펩티드는 바실러스 히사시이 Cas12b, 바실러스 써모아밀로보란스 Cas12b, 바실러스 sp. V3-13 Cas12b, 또는 알리시클로바실러스 아시디필루스 Cas12b(서열번호: 259)에 대해 적어도 약 85% 아미노산 서열 동일성을 갖는다. 다른 실시형태에서, Cas12 폴리펩티드는 바실러스 히사시이 Cas12b(서열번호: 260), 바실러스 써모아밀로보란스 Cas12b, 바실러스 sp. V3-13 Cas12b, 또는 알리시클로바실러스 아시디필루스 Cas12b에 대해 적어도 약 90% 아미노산 서열 동일성을 갖는다. 다른 실시형태에서, Cas12 폴리펩티드는 바실러스 히사시이 Cas12b, 바실러스 써모아밀로보란스 Cas12b(서열번호: 265), 바실러스 sp. V3-13 Cas12b(서열번호: 264), 또는 알리시클로바실러스 아시디필루스 Cas12b에 대해 적어도 약 95% 아미노산 서열 동일성을 갖는다. 다른 실시형태에서, Cas12 폴리펩티드는 바실러스 히사시이 Cas12b, 바실러스 써모아밀로보란스 Cas12b, 바실러스 sp. V3-13 Cas12b, 또는 알리시클로바실러스 아시디필루스 Cas12b의 단편을 함유하거나 본질적으로 이로 구성된다. 실시형태에서, Cas12 폴리펩티드는 BvCas12b(V4)를 함유하며, 이는 일부 실시형태에서 5' mRNA Cap---5' UTR---bhCas12b---정지 서열 --- 3' UTR --- 120폴리A 꼬리(서열번호: 261-263)로 발현된다.

[0408] 다른 실시형태에서, 촉매적 도메인은 BhCas12b의 아미노산 위치 153-154, 255-256, 306-307, 980-981, 1019-1020, 534-535, 604-605, 또는 344-345 또는 Cas12a, Cas12c, Cas12d, Cas12e, Cas12g, Cas12h, Cas12i 또는 Cas12j/CasΦ의 상응하는 아미노산 잔기 사이에 삽입된다. 다른 실시형태에서, 촉매적 도메인은 BhCas12b의 아미노산 P153과 S154 사이에 삽입된다. 다른 실시형태에서, 촉매적 도메인은 BhCas12b의 아미노산 K255와 E256

사이에 삽입된다. 다른 실시형태에서, 촉매적 도메인은 BhCas12b의 아미노산 D980과 G981 사이에 삽입된다. 다른 실시형태에서, 촉매적 도메인은 BhCas12b의 아미노산 K1019와 L1020 사이에 삽입된다. 다른 실시형태에서, 촉매적 도메인은 BhCas12b의 아미노산 F534와 P535 사이에 삽입된다. 다른 실시형태에서, 촉매적 도메인은 BhCas12b의 아미노산 K604와 G605 사이에 삽입된다. 다른 실시형태에서, 촉매적 도메인은 BhCas12b의 아미노산 H344와 F345 사이에 삽입된다. 다른 실시형태에서, 촉매적 도메인은 BvCas12b의 아미노산 위치 147과 148, 248과 249, 299와 300, 991과 992, 또는 1031과 1032 또는 Cas12a, Cas12c, Cas12d, Cas12e, Cas12g, Cas12h, Cas12i 또는 Cas12j/CasΦ의 상응하는 아미노산 잔기 사이에 삽입된다. 다른 실시형태에서, 촉매적 도메인은 BvCas12b의 아미노산 P147과 D148 사이에 삽입된다. 다른 실시형태에서, 촉매적 도메인은 BvCas12b의 아미노산 G248과 G249 사이에 삽입된다. 다른 실시형태에서, 촉매적 도메인은 BvCas12b의 아미노산 P299와 E300 사이에 삽입된다. 다른 실시형태에서, 촉매적 도메인은 BvCas12b의 아미노산 G991과 E992 사이에 삽입된다. 다른 실시형태에서, 촉매적 도메인은 BvCas12b의 아미노산 K1031과 M1032 사이에 삽입된다. 다른 실시형태에서, 촉매적 도메인은 AaCas12b의 아미노산 위치 157과 158, 258과 259, 310과 311, 1008과 1009, 또는 1044와 1045 사이 또는 Cas12a, Cas12c, Cas12d, Cas12e, Cas12g, Cas12h, Cas12i 또는 Cas12j/CasΦ의 상응하는 아미노산 잔기 사이에 삽입된다. 다른 실시형태에서, 촉매적 도메인은 AaCas12b의 아미노산 P157과 G158 사이에 삽입된다. 다른 실시형태에서, 촉매적 도메인은 AaCas12b의 아미노산 V258과 G259 사이에 삽입된다. 다른 실시형태에서, 촉매적 도메인은 AaCas12b의 아미노산 D310과 P311 사이에 삽입된다. 다른 실시형태에서, 촉매적 도메인은 AaCas12b의 아미노산 G1008과 E1009 사이에 삽입된다. 다른 실시형태에서, 촉매적 도메인은 AaCas12b에서의 아미노산 G1044와 K1045 사이에 삽입된다.

[0409] 다른 실시형태에서, 융합 단백질은 핵 국소화 신호(예를 들어, 이분 핵 국소화 신호)를 함유한다. 다른 실시형태에서, 핵 국소화 신호의 아미노산 서열은 MAPKKKRKVGIHGVPAA(서열번호: 1313)이다. 상기 양태의 다른 실시형태에서, 핵 국소화 신호는 다음 서열에 의해 인코딩된다:

[0410] ATGGCCCCAAAGAAGAAGCGGAAGGTCGGTATCCACGGAGTCCCAGCAGCC(서열번호: 1314). 다른 실시형태에서, Cas12b 폴리펩티드는 RuvC 도메인의 촉매적 활성을 침묵시키는 돌연변이를 함유한다. 다른 실시형태에서 Cas12b 폴리펩티드는 D574A, D829A 및/또는 D952A 돌연변이를 함유한다. 다른 실시형태에서, 융합 단백질은 태그(예를 들어, 인플루엔자 혈구응집소 태그)를 추가로 함유한다.

[0411] 일부 실시형태에서, 융합 단백질은 내부로 융합된 핵염기 편집 도메인(예를 들어, 데아미나제 도메인, 예를 들어 아데노신 데아미나제 도메인의 전부 또는 일부)을 갖는 napDNAbp 도메인(예를 들어, Cas12-유래된 도메인)을 포함한다. 일부 실시형태에서, napDNAbp는 Cas12b이다. 일부 실시형태에서, 염기 편집기는 아래 표 4에 제공된 좌위에서 삽입된 내부로 융합된 TadA*8 도메인을 갖는 BhCas12b 도메인을 포함한다.

[0412] 표 4: Cas12b 단백질에서 삽입 좌위

BhCas12b	삽입 부위	aa 사이 삽입됨
위치 1	153	PS
위치 2	255	KE
위치 3	306	DE
위치 4	980	DG
위치 5	1019	KL
위치 6	534	FP
위치 7	604	KG
위치 8	344	HF
BvCas12b	삽입 부위	aa 사이 삽입됨
위치 1	147	PD
위치 2	248	GG
위치 3	299	PE
위치 4	991	GE
위치 5	1031	KM
AaCas12b	삽입 부위	aa 사이 삽입됨
위치 1	157	PG
위치 2	258	VG
위치 3	310	DP
위치 4	1008	GE
위치 5	1044	GK

[0413]

[0414] 비제한적인 예로서, 아데노신 데아미나제(예를 들어, TadA*8.13)는 BhCas12b 안으로 삽입되어 핵산 서열을 효과적으로 편집하는 융합 단백질(예를 들어, TadA*8.13-BhCas12b)을 생성할 수 있다.

[0415] 일부 실시형태에서, 본 명세서에 기재된 염기 편집 시스템은 Cas9 안으로 삽입된 TadA를 갖는 ABE이다. Cas9 안으로 삽입된 TadA를 갖는 관련 ABE의 폴리펩티드 서열은 첨부된 서열 목록에 서열번호: 1315-1360으로 제공된다.

[0416] 일부 실시형태에서, TadA 또는 이의 변이체를 식별된 위치에서 Cas9 폴리펩티드 안으로 삽입하기 위해 아데노신 데아미나제 염기 편집기가 생성되었다.

[0417] 예시적이지만 비제한적인 융합 단백질은 국제 PCT 출원 번호 PCT/US2020/016285 및 미국 가출원 번호 62/852,228 및 62/852,224에 기재되어 있으며, 이의 내용은 그 전체가 본 명세서에 참고로 포함된다.

[0418] ***A에서 G로 편집***

[0419] 일부 실시형태에서, 본 명세서에 기재된 염기 편집기는 아데노신 데아미나제 도메인을 포함한다. 염기 편집기의 이러한 아데노신 데아미나제 도메인은 G의 염기페어링 특성을 나타내는 이노신(I)을 형성하도록 A를 탈아미노화함에 의해 아데닌(A) 핵염기의 구아닌(G) 핵염기로의 편집을 촉진할 수 있다. 아데노신 데아미나제는 데옥시리

보핵산(DNA)에서의 데옥시아데노신 잔기의 아데닌을 탈아미노화(즉, 아민기를 제거)할 수 있다. 일부 실시형태에서, A-대-G 염기 편집기는 이노신 염기 삭제 수선의 억제제, 예를 들어 우라실 글리코실라제 억제제(UGI) 도메인 또는 촉매적으로 불활성인 이노신 특이적 뉴클레아제를 추가로 포함한다. 임의의 특정 이론에 얽매이기를 바라지 않고, UGI 도메인 또는 촉매적으로 불활성인 이노신 특이적 뉴클레아제는 탈아미노화된 아데노신 잔기(예를 들어, 이노신)의 염기 삭제 수선를 억제하거나 방지할 수 있으며, 이는 염기 편집기의 활성 또는 효율성을 향상시킬 수 있다.

[0420] 아데노신 데아미나제를 포함하는 염기 편집기는 DNA, RNA 및 DNA-RNA 하이브리드를 포함한 임의의 폴리뉴클레오티드에 작용할 수 있다. 특정 실시형태에서, 아데노신 데아미나제를 포함하는 염기 편집기는 RNA를 포함하는 폴리뉴클레오티드의 표적 A를 탈아미노화할 수 있다. 예를 들어, 염기 편집기는 RNA 폴리뉴클레오티드 및/또는 DNA-RNA 하이브리드 폴리뉴클레오티드의 표적 A를 탈아미노화할 수 있는 아데노신 데아미나제 도메인을 포함할 수 있다. 실시형태에서, 염기 편집기 안에 합체된 아데노신 데아미나제는 RNA(ADAR, 예를 들어 ADAR1 또는 ADAR2) 또는 tRNA(ADAT)에 작용하는 아데노신 데아미나제의 전부 또는 일부를 포함한다. 아데노신 데아미나제 도메인을 포함하는 염기 편집기는 또한 DNA 폴리뉴클레오티드의 A 핵염기를 탈아미노화할 수 있다. 실시형태에서 염기 편집기의 아데노신 데아미나제 도메인은 ADAT가 DNA에서 표적 A를 탈아미노화하도록 허용하는 하나 이상의 돌연변이를 포함하는 ADAT의 전부 또는 일부를 포함한다. 예를 들어, 염기 편집기는 다음 돌연변이 중 하나 이상을 포함하는 대장균(EcTadA)으로부터 ADAT의 전부 또는 일부를 포함할 수 있다: D108N, A106V, D147Y, E155V, L84F, H123Y, I156F 또는 다른 아데노신 데아미나제에서 상응하는 돌연변이. 예시적인 ADAT 상동체 폴리펩티드 서열은 서열 목록에 서열번호: 1363-1370으로 제공된다.

[0421] 아데노신 데아미나제는 임의의 적합한 유기체(예를 들어, *E. coli*)로부터 유래될 수 있다. 일부 실시형태에서, 아데노신 데아미나제는 원핵생물로부터의 것이다. 일부 실시형태에서, 아데노신 데아미나제는 박테리아로부터의 것이다. 일부 실시형태에서, 아데노신 데아미나제는 *대장균, 스타필로코커스 아우레우스, 살모넬라 타이피, 쉬와넬라 푸트리파시엔스, 헤모필루스 인플루엔자, 카울로박터 크레센투스* 또는 *바실러스 서브틸리스*로부터의 것이다. 일부 실시형태에서, 아데노신 데아미나제는 *대장균*으로부터의 것이다. 일부 실시형태에서, 아데닌 데아미나제는 본 명세서에 제공된 돌연변이 중 임의의 것에 상응하는 하나 이상의 돌연변이(예를 들어, ecTadA에서의 돌연변이)를 포함하는 자연적으로-발생하는 아데노신 데아미나제이다. 임의의 상동성 단백질에서의 상응하는 잔기는 예를 들어 서열 정렬 및 상동체 잔기의 결정에 의해 식별될 수 있다. 본 명세서에 기재된 임의의 돌연변이(예를 들어, ecTadA에서 식별된 임의의 돌연변이)에 상응하는 임의의 자연적으로-발생하는 아데노신 데아미나제(예를 들어, ecTadA와 상동성을 가짐)에서의 돌연변이가 따라서 생성될 수 있다.

[0422] 일부 실시형태에서, 아데노신 데아미나제는 본 명세서에 제공된 임의의 아데노신 데아미나제에 제시된 아미노산 서열 중 임의의 하나와 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 적어도 99.5% 동일한 아미노산 서열을 포함한다. 본 명세서에 제공된 아데노신 데아미나제는 하나 이상의 돌연변이(*예를 들어*, 본 명세서에 제공된 임의의 돌연변이)를 포함할 수 있음을 이해해야 한다. 개시내용은 본 명세서에 기재된 임의의 돌연변이 또는 이의 조합 플러스 특정 퍼센트 동일성을 갖는 임의의 데아미나제 도메인을 제공한다. 일부 실시형태에서, 아데노신 데아미나제는 본 명세서에 제공된 임의의 아데노신 데아미나제 또는 참조 서열과 비교하여 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 21, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50개 이상의 돌연변이를 갖는 아미노산 서열을 포함한다. 일부 실시형태에서, 아데노신 데아미나제는 당업계에 공지되어 있거나 본 명세서에 기재된 아미노산 서열 중 임의의 하나와 비교하여 적어도 5개, 적어도 10개, 적어도 15개, 적어도 20개, 적어도 25개, 적어도 30개, 적어도 35개, 적어도 40개, 적어도 45개, 적어도 50개, 적어도 60개, 적어도 70개, 적어도 80개, 적어도 90개, 적어도 100개, 적어도 110개, 적어도 120개, 적어도 130개, 적어도 140개, 적어도 150개, 적어도 160개 또는 적어도 170개의 동일한 연속 아미노산 잔기를 갖는 아미노산 서열을 포함한다.

[0423] 본 명세서에 제공된 임의의 돌연변이(*예를 들어*, TadA 참조 서열에 기반함)은 *대장균* TadA(ecTadA), *S. 아우레우스* TadA(saTadA) 또는 다른 아데노신 데아미나제(*예를 들어*, 박테리아 아데노신 데아미나제)와 같은 다른 아데노신 데아미나제 안으로 도입될 수 있음을 이해해야 한다. 추가의 데아미나제가 유사하게 정렬되어 본 명세서에 제공된 바와 같이 돌연변이될 수 있는 상동성 아미노산 잔기를 식별할 수 있다는 것이 당업자에게 명백할 것이다. 따라서, TadA 참조 서열에서 식별된 임의의 돌연변이는 상동성 아미노산 잔기를 갖는 다른 아데노신 데아미나제(*예를 들어*, ecTada)에서 만들어질 수 있다. 또한 본 명세서에 제공된 임의의 돌연변이는 TadA 참조 서열 또는 다른 아데노신 데아미나제에서 개별적으로 또는 임의의 조합으로 만들어질 수 있다는 것을 이해해야 한다.

[0424] 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 D108X 돌연변이, 또는 다른 아데노신 데아미나제에서의 상응하는 돌연변이를 포함하며, 여기서 X는 야생형 아데노신 데아미나제에서의 상응하는 아미노산 이외의 임의의 아미노산을 나타낸다. 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 D108G, D108N, D108V, D108A 또는 D108Y 돌연변이, 또는 다른 아데노신 데아미나제에서의 상응하는 돌연변이를 포함한다. 그러나, 추가의 데아미나제가 유사하게 정렬되어 본 명세서에 제공된 바와 같이 돌연변이될 수 있는 상동성 아미노산 잔기를 식별할 수 있다는 것이 이해되어야 한다.

[0425] 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 A106X 돌연변이, 또는 다른 아데노신 데아미나제에서의 상응하는 돌연변이를 포함하며, 여기서 X는 야생형 아데노신 데아미나제에서의 상응하는 아미노산 이외의 임의의 아미노산을 나타낸다. 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 A106V 돌연변이, 또는 다른 아데노신 데아미나제(*예를 들어*, ecTadA)에서의 상응하는 돌연변이를 포함한다.

[0426] 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 E155X 돌연변이, 또는 다른 아데노신 데아미나제에서의 상응하는 돌연변이를 포함하며, 여기서 X의 존재는 야생형 아데노신 데아미나제에서의 상응하는 아미노산 이외의 임의의 아미노산을 나타낸다. 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 E155D, E155G 또는 E155V 돌연변이, 또는 다른 아데노신 데아미나제(*예를 들어*, ecTadA)에서의 상응하는 돌연변이를 포함한다.

[0427] 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 D147X 돌연변이, 또는 다른 아데노신 데아미나제에서의 상응하는 돌연변이를 포함하며, 여기서 X의 존재는 야생형 아데노신 데아미나제에서의 상응하는 아미노산 이외의 임의의 아미노산을 나타낸다. 일부 실시형태에서, 아데노신 데아미나제는 D147Y인, TadA 참조 서열에서의 돌연변이, 또는 다른 아데노신 데아미나제(*예를 들어*, ecTadA)에서의 상응하는 돌연변이를 포함한다.

[0428] 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 돌연변이인, A106X, E155X 또는 D147X, 또는 다른 아데노신 데아미나제(*예를 들어*, ecTadA)에서의 상응하는 돌연변이를 포함하며, 여기서 X는 야생형 아데노신 데아미나제에서의 상응하는 아미노산 이외의 임의의 아미노산을 나타낸다. 일부 실시형태에서, 아데노신 데아미나제는 E155D, E155G 또는 E155V 돌연변이를 포함한다. 일부 실시형태에서, 아데노신 데아미나제는 D147Y를 포함한다.

[0429] 또한 본 명세서에 제공된 임의의 돌연변이는 ecTadA 또는 다른 아데노신 데아미나제에서 개별적으로 또는 임의의 조합으로 만들어질 수 있음을 이해해야 한다. 예를 들어, 아데노신 데아미나제는 TadA 참조 서열에서 D108N, A106V, E155V 및/또는 D147Y 돌연변이 또는 다른 아데노신 데아미나제(*예를 들어*, ecTadA)에서의 상응하는 돌연변이를 함유할 수 있다. 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 다음 돌연변이의 그룹(돌연변이의 그룹은 ";"로 구분됨), 또는 다른 아데노신 데아미나제에서의 상응하는 돌연변이를 포함한다: D108N 및 A106V; D108N 및 E155V; D108N 및 D147Y; A106V 및 E155V; A106V 및 D147Y; E155V 및 D147Y; D108N, A106V 및 E155V; D108N, A106V 및 D147Y; D108N, E155V 및 D147Y; A106V, E155V 및 D147Y; 및 D108N, A106V, E155V 및 D147Y. 그러나, 본 명세서에 제공된 상응하는 돌연변이의 임의의 조합은 아데노신 데아미나제(*예를 들어*, ecTadA)에서 만들어질 수 있음을 이해해야 한다.

[0430] 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 H8X, T17X, L18X, W23X, L34X, W45X, R51X, A56X, E59X, E85X, M94X, I95X, V102X, F104X, A106X, R107X, D108X, K110X, M118X, N127X, A138X, F149X, M151X, R153X, Q154X, I156X, 및/또는 K157X 돌연변이 중 하나 이상, 또는 다른 아데노신 데아미나제에서의 하나 이상의 상응하는 돌연변이를 포함하며, 여기서 X의 존재는 야생형 아데노신 데아미나제에서의 상응하는 아미노산 이외의 임의의 아미노산을 나타낸다. 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 H8Y, T17S, L18E, W23L, L34S, W45L, R51H, A56E, 또는 A56S, E59G, E85K, 또는 E85G, M94L, I95L, V102A, F104L, A106V, R107C, 또는 R107H, 또는 R107P, D108G, 또는 D108N, 또는 D108V, 또는 D108A, 또는 D108Y, K110I, M118K, N127S, A138V, F149Y, M151V, R153C, Q154L, I156D 및/또는 K157R 돌연변이 중 하나 이상, 또는 다른 아데노신 데아미나제에서의 하나 이상의 상응하는 돌연변이를 포함한다.

[0431] 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 H8X, D108X 및/또는 N127X 돌연변이 중 하나 이상, 또는 다른 아데노신 데아미나제에서 하나 이상의 상응하는 돌연변이를 포함하며, 여기서 X는 임의의 아미노산의 존재를 나타낸다. 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 H8Y, D108N 및/또는 N127S 돌연변이 중 하나 이상, 또는 다른 아데노신 데아미나제에서 하나 이상의 상응하는 돌연변이를 포함한다.

[0432] 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 H8X, R26X, M61X, L68X, M70X, A106X, D108X,

A109X, N127X, D147X, R152X, Q154X, E155X, K161X, Q163X 및/또는 T166X 돌연변이 중 하나 이상, 또는 다른 아데노신 데아미나제에서 하나 이상의 상응하는 돌연변이를 포함하며, 여기서 X는 야생형 아데노신 데아미나제에서의 상응하는 아미노산 이외의 임의의 아미노산의 존재를 나타낸다. 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 H8Y, R26W, M61I, L68Q, M70V, A106T, D108N, A109T, N127S, D147Y, R152C, Q154H or Q154R, E155G or E155V or E155D, K161Q, Q163H, 및/또는 T166P 돌연변이 중 하나 이상, 또는 다른 아데노신 데아미나제에서 하나 이상의 상응하는 돌연변이를 포함한다.

[0433] 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 H8X, D108X, N127X, D147X, R152X 및 Q154X로 구성된 군으로부터 선택된 1개, 2개, 3개, 4개, 5개 또는 6개의 돌연변이, 또는 다른 아데노신 데아미나제(*예를 들어*, ecTadA)에서 상응하는 돌연변이 또는 돌연변이들을 포함하며, 여기서 X는 야생형 아데노신 데아미나제에서의 상응하는 아미노산 이외의 임의의 아미노산의 존재를 나타낸다. 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 H8X, M61X, M70X, D108X, N127X, Q154X, E155X 및 Q163X로 구성된 군으로부터 선택된 1개, 2개, 3개, 4개, 5개, 6개, 7개 또는 8개의 돌연변이, 또는 다른 아데노신 데아미나제(*예를 들어*, ecTadA)에서의 상응하는 돌연변이 또는 돌연변이들을 포함하며, 여기서 X는 야생형 아데노신 데아미나제에서의 상응하는 아미노산 이외의 임의의 아미노산의 존재를 나타낸다. 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 H8X, D108X, N127X, E155X 및 T166X로 구성된 군으로부터 선택된 1개, 2개, 3개, 4개, 또는 5개의 돌연변이, 또는 다른 아데노신 데아미나제(*예를 들어*, ecTadA)에서의 상응하는 돌연변이 또는 돌연변이들을 포함하며, 여기서 X는 야생형 아데노신 데아미나제에서의 상응하는 아미노산 이외의 임의의 아미노산의 존재를 나타낸다.

[0434] 일부 실시형태에서, 아데노신 데아미나제는 H8X, A106X 및 D108X로 구성된 군으로부터 선택된 1개, 2개, 3개, 4개, 5개 또는 6개의 돌연변이, 또는 다른 아데노신 데아미나제에서의 상응하는 돌연변이 또는 돌연변이들을 포함하며, 여기서 X는 야생형 아데노신 데아미나제에서의 상응하는 아미노산 이외의 임의의 아미노산의 존재를 나타낸다. 일부 실시형태에서, 아데노신 데아미나제는 H8X, R26X, L68X, D108X, N127X, D147X 및 E155X로 구성된 군으로부터 선택된 1개, 2개, 3개, 4개, 5개, 6개, 7개 또는 8개의 돌연변이, 또는 다른 아데노신 데아미나제에서의 상응하는 돌연변이 또는 돌연변이들을 포함하며, 여기서 X는 야생형 아데노신 데아미나제에서의 상응하는 아미노산 이외의 임의의 아미노산의 존재를 나타낸다.

[0435] 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 H8X, R126X, L68X, D108X, N127X, D147X 및 E155X로 구성된 군으로부터 선택된 1개, 2개, 3개, 4개, 5개, 6개 또는 7개의 돌연변이, 또는 다른 아데노신 데아미나제에서의 상응하는 돌연변이 또는 돌연변이들을 포함하며, 여기서 X는 야생형 아데노신 데아미나제에서의 상응하는 아미노산 이외의 임의의 아미노산의 존재를 나타낸다. 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 H8X, D108X, A109X, N127X 및 E155X로 구성된 군으로부터 선택된 1개, 2개, 3개, 4개 또는 5개의 돌연변이, 또는 다른 아데노신 데아미나제에서의 상응하는 돌연변이 또는 돌연변이들을 포함하며, 여기서 X는 야생형 아데노신 데아미나제에서의 상응하는 아미노산 이외의 임의의 아미노산의 존재를 나타낸다.

[0436] 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 H8Y, D108N, N127S, D147Y, R152C 및 Q154H로 구성된 군으로부터 선택된 1개, 2개, 3개, 4개, 5개 또는 6개의 돌연변이, 또는 다른 아데노신 데아미나제(*예를 들어*, ecTadA)에서 상응하는 돌연변이 또는 돌연변이들을 포함한다. 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 H8Y, M61I, M70V, D108N, N127S, Q154R, E155G 및 Q163H로 구성된 군으로부터 선택된 1개, 2개, 3개, 4개, 5개, 6개, 7개 또는 8개의 돌연변이, 또는 다른 아데노신 데아미나제(*예를 들어*, ecTadA)에서의 상응하는 돌연변이 또는 돌연변이들을 포함한다. 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 H8Y, D108N, N127S, E155V 및 T166P로 구성된 군으로부터 선택된 1개, 2개, 3개, 4개, 또는 5개의 돌연변이, 또는 다른 아데노신 데아미나제(*예를 들어*, ecTadA)에서의 상응하는 돌연변이 또는 돌연변이들을 포함한다. 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 H8Y, A106T, D108N, N127S, E155D 및 K161Q로 구성된 군으로부터 선택된 1개, 2개, 3개, 4개, 5개 또는 6개의 돌연변이, 또는 다른 아데노신 데아미나제(*예를 들어*, ecTadA)에서의 상응하는 돌연변이 또는 돌연변이들을 포함한다. 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 H8Y, R26W, L68Q, D108N, N127S, D147Y 및 E155V로 구성된 군으로부터 선택된 1개, 2개, 3개, 4개, 5개, 6개, 7개 또는 8개의 돌연변이, 또는 다른 아데노신 데아미나제(*예를 들어*, ecTadA)에서의 상응하는 돌연변이 또는 돌연변이들을 포함한다. 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 H8Y, D108N, A109T, N127S 및 E155G로 구성된 군으로부터 선택된 1개, 2개, 3개, 4개, 또는 5개의 돌연변이, 또는 다른 아데노신 데아미나제(*예를 들어*, ecTadA)에서의 상응하는 돌연변이 또는 돌연변이들을 포함한다.

[0437] 일부 실시형태에서, 아데노신 데아미나제는 다른 아데노신 데아미나제에서 하나 이상의 또는 하나 이상의 상응하는 돌연변이를 포함한다. 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에 D108N, D108G 또는 D108V 돌연변이, 또는 다른 아데노신 데아미나제에서의 상응하는 돌연변이를 포함한다. 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 A106V 및 D108N 돌연변이, 또는 다른 아데노신 데아미나제에서의 상응하는 돌연변이를 포함한다. 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 R107C 및 D108N 돌연변이, 또는 다른 아데노신 데아미나제에서의 상응하는 돌연변이를 포함한다. 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 H8Y, D108N, N127S, D147Y 및 Q154H 돌연변이, 또는 다른 아데노신 데아미나제에서의 상응하는 돌연변이를 포함한다. 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 H8Y, D108N, N127S, D147Y 및 E155V 돌연변이, 또는 다른 아데노신 데아미나제에서의 상응하는 돌연변이를 포함한다. 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 D108N, D147Y 및 E155V 돌연변이, 또는 다른 아데노신 데아미나제에서의 상응하는 돌연변이를 포함한다. 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 H8Y, D108N 및 N127S 돌연변이, 또는 다른 아데노신 데아미나제에서의 상응하는 돌연변이를 포함한다. 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 A106V, D108N, D147Y 및 E155V 돌연변이, 또는 다른 아데노신 데아미나제(*예를 들어*, ecTadA)에서의 상응하는 돌연변이를 포함한다.

[0438] 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 S2X, H8X, I49X, L84X, H123X, N127X, I156X 및/또는 K160X 돌연변이 중 하나 이상, 또는 다른 아데노신 데아미나제에서 하나 이상의 상응하는 돌연변이를 포함하며, 여기서 X의 존재는 야생형 아데노신 데아미나제에서의 상응하는 아미노산 이외의 임의의 아미노산을 나타낸다. 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 S2A, H8Y, I49F, L84F, H123Y, N127S, I156F 및/또는 K160S 돌연변이 중 하나 이상, 또는 다른 아데노신 데아미나제(*예를 들어*, ecTadA)에서 하나 이상의 상응하는 돌연변이를 포함한다.

[0439] 일부 실시형태에서, 아데노신 데아미나제는 L84X 돌연변이 아데노신 데아미나제를 포함하며, 여기서 X는 야생형 아데노신 데아미나제에서의 상응하는 아미노산 이외의 임의의 아미노산을 나타낸다. 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 L84F 돌연변이, 또는 다른 아데노신 데아미나제(*예를 들어*, ecTadA)에서의 상응하는 돌연변이를 포함한다.

[0440] 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 H123X 돌연변이, 또는 다른 아데노신 데아미나제에서의 상응하는 돌연변이를 포함하며, 여기서 X는 야생형 아데노신 데아미나제에서의 상응하는 아미노산 이외의 임의의 아미노산을 나타낸다. 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 H123Y 돌연변이, 또는 다른 아데노신 데아미나제에서의 상응하는 돌연변이를 포함한다.

[0441] 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 I156X 돌연변이, 또는 다른 아데노신 데아미나제에서의 상응하는 돌연변이를 포함하며, 여기서 X는 야생형 아데노신 데아미나제에서의 상응하는 아미노산 이외의 임의의 아미노산을 나타낸다. 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 I156F 돌연변이, 또는 다른 아데노신 데아미나제에서의 상응하는 돌연변이를 포함한다.

[0442] 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 L84X, A106X, D108X, H123X, D147X, E155X 및 I156X로 구성된 군으로부터 선택된 1개, 2개, 3개, 4개, 5개, 6개 또는 7개의 돌연변이, 또는 다른 아데노신 데아미나제에서의 상응하는 돌연변이 또는 돌연변이들을 포함하며, 여기서 X는 야생형 아데노신 데아미나제에서의 상응하는 아미노산 이외의 임의의 아미노산의 존재를 나타낸다. 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 S2X, I49X, A106X, D108X, D147X 및 E155X로 구성된 군으로부터 선택된 1개, 2개, 3개, 4개, 5개 또는 6개의 돌연변이, 또는 다른 아데노신 데아미나제에서의 상응하는 돌연변이 또는 돌연변이들을 포함하며, 여기서 X는 야생형 아데노신 데아미나제에서의 상응하는 아미노산 이외의 임의의 아미노산의 존재를 나타낸다. 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 H8X, A106X, D108X, N127X 및 K160X로 구성된 군으로부터 선택된 1개, 2개, 3개, 4개 또는 5개의 돌연변이, 또는 다른 아데노신 데아미나제에서의 상응하는 돌연변이 또는 돌연변이들을 포함하며, 여기서 X는 야생형 아데노신 데아미나제에서의 상응하는 아미노산 이외의 임의의 아미노산의 존재를 나타낸다.

[0443] 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 L84F, A106V, D108N, H123Y, D147Y, E155V 및 I156F로 구성된 군으로부터 선택된 1개, 2개, 3개, 4개, 5개, 6개 또는 7개의 돌연변이, 또는 다른 아데노신 데아미나제에서의 상응하는 돌연변이 또는 돌연변이들을 포함한다. 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 S2A, I49F, A106V, D108N, D147Y 및 E155V로 구성된 군으로부터 선택된 1개, 2개, 3개, 4개, 5개 또는 6개의 돌연변이를 포함한다.

[0444] 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 H8Y, A106T, D108N, N127S 및 K160S로 구성된 군으로부터 선택된 1개, 2개, 3개, 4개 또는 5개의 돌연변이, 또는 다른 아데노신 데아미나제에서의 상응하는 돌연변이 또는 돌연변이들을 포함한다.

[0445] 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 E25X, R26X, R107X, A142X 및/또는 A143X 돌연변이 중 하나 이상, 또는 다른 아데노신 데아미나제에서 하나 이상의 상응하는 돌연변이를 포함하며, 여기서 X의 존재는 야생형 아데노신 데아미나제에서의 상응하는 아미노산 이외의 임의의 아미노산을 나타낸다. 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 E25M, E25D, E25A, E25R, E25V, E25S, E25Y, R26G, R26N, R26Q, R26C, R26L, R26K, R107P, R107K, R107A, R107N, R107W, R107H, R107S, A142N, A142D, A142G, A143D, A143G, A143E, A143L, A143W, A143M, A143S, A143Q 및/또는 A143R 돌연변이 중 하나 이상, 또는 다른 아데노신 데아미나제에서 하나 이상의 상응하는 돌연변이를 포함한다. 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에 상응하는 본 명세서에 기재된 돌연변이 중 하나 이상, 또는 다른 아데노신 데아미나제에서 하나 이상의 상응하는 돌연변이를 포함한다.

[0446] 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 E25X 돌연변이, 또는 다른 아데노신 데아미나제에서의 상응하는 돌연변이를 포함하며, 여기서 X는 야생형 아데노신 데아미나제에서의 상응하는 아미노산 이외의 임의의 아미노산을 나타낸다. 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 E25M, E25D, E25A, E25R, E25V, E25S 또는 E25Y 돌연변이, 또는 다른 아데노신 데아미나제(*예를 들어*, ecTadA)에서의 상응하는 돌연변이를 포함한다.

[0447] 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 R26X 돌연변이, 또는 다른 아데노신 데아미나제에서의 상응하는 돌연변이를 포함하며, 여기서 X는 야생형 아데노신 데아미나제에서의 상응하는 아미노산 이외의 임의의 아미노산을 나타낸다. 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 R26G, R26N, R26Q, R26C, R26L 또는 R26K 돌연변이, 또는 다른 아데노신 데아미나제(*예를 들어*, ecTadA)에서의 상응하는 돌연변이를 포함한다.

[0448] 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 R107X 돌연변이, 또는 다른 아데노신 데아미나제에서의 상응하는 돌연변이를 포함하며, 여기서 X는 야생형 아데노신 데아미나제에서의 상응하는 아미노산 이외의 임의의 아미노산을 나타낸다. 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 R107P, R107K, R107A, R107N, R107W, R107H 또는 R107S 돌연변이, 또는 다른 아데노신 데아미나제(*예를 들어*, ecTadA)에서의 상응하는 돌연변이를 포함한다.

[0449] 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 A142X 돌연변이, 또는 다른 아데노신 데아미나제에서의 상응하는 돌연변이를 포함하며, 여기서 X는 야생형 아데노신 데아미나제에서의 상응하는 아미노산 이외의 임의의 아미노산을 나타낸다. 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 돌연변이인, A142N, A142D, A142G, 또는 다른 아데노신 데아미나제(*예를 들어*, ecTadA)에서의 상응하는 돌연변이를 포함한다.

[0450] 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 A143X 돌연변이, 또는 다른 아데노신 데아미나제에서의 상응하는 돌연변이를 포함하며, 여기서 X는 야생형 아데노신 데아미나제에서의 상응하는 아미노산 이외의 임의의 아미노산을 나타낸다. 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 A143D, A143G, A143E, A143L, A143W, A143M, A143S, A143Q 및/또는 A143R 돌연변이, 또는 다른 아데노신 데아미나제(*예를 들어*, ecTadA)에서의 상응하는 돌연변이를 포함한다.

[0451] 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 H36X, N37X, P48X, I49X, R51X, M70X, N72X, D77X, E134X, S146X, Q154X, K157X 및/또는 K161X 돌연변이 중 하나 이상, 또는 다른 아데노신 데아미나제에서 하나 이상의 상응하는 돌연변이를 포함하며, 여기서 X의 존재는 야생형 아데노신 데아미나제에서의 상응하는 아미노산 이외의 임의의 아미노산을 나타낸다. 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 H36L, N37T, N37S, P48T, P48L, I49V, R51H, R51L, M70L, N72S, D77G, E134G, S146R, S146C, Q154H, K157N 및/또는 K161T 돌연변이 중 하나 이상, 또는 다른 아데노신 데아미나제(*예를 들어*, ecTadA)에서 하나 이상의 상응하는 돌연변이를 포함한다.

[0452] 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 H36X 돌연변이, 또는 다른 아데노신 데아미나제에서의 상응하는 돌연변이를 포함하며, 여기서 X는 야생형 아데노신 데아미나제에서의 상응하는 아미노산 이외의 임의의 아미노산을 나타낸다. 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 H36L

돌연변이, 또는 다른 아데노신 데아미나제에서의 상응하는 돌연변이를 포함한다.

[0453] 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 N37X 돌연변이, 또는 다른 아데노신 데아미나제에서의 상응하는 돌연변이를 포함하며, 여기서 X는 야생형 아데노신 데아미나제에서의 상응하는 아미노산 이외의 임의의 아미노산을 나타낸다. 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 N37T 또는 N37S 돌연변이, 또는 다른 아데노신 데아미나제에서의 상응하는 돌연변이를 포함한다.

[0454] 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 P48X 돌연변이, 또는 다른 아데노신 데아미나제에서의 상응하는 돌연변이를 포함하며, 여기서 X는 야생형 아데노신 데아미나제에서의 상응하는 아미노산 이외의 임의의 아미노산을 나타낸다. 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 P48T 또는 P48L 돌연변이, 또는 다른 아데노신 데아미나제에서 상응하는 돌연변이를 포함한다.

[0455] 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 R51X 돌연변이, 또는 다른 아데노신 데아미나제에서의 상응하는 돌연변이를 포함하며, 여기서 X는 야생형 아데노신 데아미나제에서의 상응하는 아미노산 이외의 임의의 아미노산을 나타낸다. 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 R51H 또는 R51L 돌연변이, 또는 다른 아데노신 데아미나제에서의 상응하는 돌연변이를 포함한다.

[0456] 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 S146X 돌연변이, 또는 다른 아데노신 데아미나제에서의 상응하는 돌연변이를 포함하며, 여기서 X는 야생형 아데노신 데아미나제에서의 상응하는 아미노산 이외의 임의의 아미노산을 나타낸다. 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 S146R 또는 S146C 돌연변이, 또는 다른 아데노신 데아미나제에서의 상응하는 돌연변이를 포함한다.

[0457] 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 K157X 돌연변이, 또는 다른 아데노신 데아미나제에서의 상응하는 돌연변이를 포함하며, 여기서 X는 야생형 아데노신 데아미나제에서의 상응하는 아미노산 이외의 임의의 아미노산을 나타낸다. 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 K157N 돌연변이, 또는 다른 아데노신 데아미나제에서의 상응하는 돌연변이를 포함한다.

[0458] 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 P48X 돌연변이, 또는 다른 아데노신 데아미나제에서의 상응하는 돌연변이를 포함하며, 여기서 X는 야생형 아데노신 데아미나제에서의 상응하는 아미노산 이외의 임의의 아미노산을 나타낸다. 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 P48S, P48T, 또는 P48A 돌연변이, 또는 다른 아데노신 데아미나제에서의 상응하는 돌연변이를 포함한다.

[0459] 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 A142X 돌연변이, 또는 다른 아데노신 데아미나제에서의 상응하는 돌연변이를 포함하며, 여기서 X는 야생형 아데노신 데아미나제에서의 상응하는 아미노산 이외의 임의의 아미노산을 나타낸다. 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 A142N 돌연변이, 또는 다른 아데노신 데아미나제에서의 상응하는 돌연변이를 포함한다.

[0460] 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 W23X 돌연변이, 또는 다른 아데노신 데아미나제에서의 상응하는 돌연변이를 포함하며, 여기서 X는 야생형 아데노신 데아미나제에서의 상응하는 아미노산 이외의 임의의 아미노산을 나타낸다. 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 W23R 또는 W23L 돌연변이, 또는 다른 아데노신 데아미나제에서의 상응하는 돌연변이를 포함한다.

[0461] 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 R152X 돌연변이, 또는 다른 아데노신 데아미나제에서의 상응하는 돌연변이를 포함하며, 여기서 X는 야생형 아데노신 데아미나제에서의 상응하는 아미노산 이외의 임의의 아미노산을 나타낸다. 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열에서 R152P 또는 R52H 돌연변이, 또는 다른 아데노신 데아미나제에서의 상응하는 돌연변이를 포함한다.

[0462] 일 실시형태에서, 아데노신 데아미나제는 돌연변이 H36L, R51L, L84F, A106V, D108N, H123Y, S146C, D147Y, E155V, I156F 및 K157N을 포함할 수 있다. 일부 실시형태에서, 아데노신 데아미나제는 TadA 참조 서열과 관련하여 다음 돌연변이의 조합을 포함하며, 여기서 조합의 각 돌연변이는 "_"에 의해 구분되고, 돌연변이의 각 조합은 괄호 사이에 있다:

[0463] (A106V_D108N),

[0464] (R107C_D108N),

[0465] (H8Y_D108N_N127S_D147Y_Q154H),

[0466] (H8Y _D108N_N127S_D147Y_E155V),

[0467] (D108N_D147Y_E155V),

[0468] (H8Y_D108N_N127S),

[0469] (H8Y_D108N_N127S_D147Y_Q154H),

[0470] (A106V_D108N_D147Y_E155V),

[0471] (D108Q_D147Y_E155V),

[0472] (D108M_D147Y_E155V),

[0473] (D108L_D147Y_E155V),

[0474] (D108K_D147Y_E155V),

[0475] (D108I_D147Y_E155V),

[0476] (D108F_D147Y_E155V),

[0477] (A106V_D108N_D147Y),

[0478] (A106V_D108M_D147Y_E155V),

[0479] (E59A_A106V_D108N_D147Y_E155V),

[0480] (E59A cat dead_A106V_D108N_D147Y_E155V),

[0481] (L84F_A106V_D108N_H123Y_D147Y_E155V_I156Y),

[0482] (L84F_A106V_D108N_H123Y_D147Y_E155V_I156F),

[0483] (D103A_D104N),

[0484] (G22P_D103A_D104N),

[0485] (D103A_D104N_S138A),

[0486] (R26G_L84F_A106V_R107H_D108N_H123Y_A142N_A143D_D147Y_E155V_I156F),

[0487] (E25G_R26G_L84F_A106V_R107H_D108N_H123Y_A142N_A143D_D147Y_E155V_I156F),

[0488] (E25D_R26G_L84F_A106V_R107K_D108N_H123Y_A142N_A143G_D147Y_E155V_I156F),
(R26Q_L84F_A106V_D108N_H123Y_A142N_D147Y_E155V_I156F),

[0489] (E25M_R26G_L84F_A106V_R107P_D108N_H123Y_A142N_A143D_D147Y_E155V_I156F),
(R26C_L84F_A106V_R107H_D108N_H123Y_A142N_D147Y_E155V_I156F),
(L84F_A106V_D108N_H123Y_A142N_A143L_D147Y_E155V_I156F),

[0490] (R26G_L84F_A106V_D108N_H123Y_A142N_D147Y_E155V_I156F),

[0491] (E25A_R26G_L84F_A106V_R107N_D108N_H123Y_A142N_A143E_D147Y_E155V_I156F),
(R26G_L84F_A106V_R107H_D108N_H123Y_A142N_A143D_D147Y_E155V_I156F),

[0492] (A106V_D108N_A142N_D147Y_E155V),

[0493] (R26G_A106V_D108N_A142N_D147Y_E155V),

[0494] (E25D_R26G_A106V_R107K_D108N_A142N_A143G_D147Y_E155V),
(R26G_A106V_D108N_R107H_A142N_A143D_D147Y_E155V),

[0495] (E25D_R26G_A106V_D108N_A142N_D147Y_E155V),

[0496] (A106V_R107K_D108N_A142N_D147Y_E155V),

[0497] (A106V_D108N_A142N_A143G_D147Y_E155V),

[0498] (A106V_D108N_A142N_A143L_D147Y_E155V),

[0499] (H36L_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F _K157N),

[0500] (N37T_P48T_M70L_L84F_A106V_D108N_H123Y_D147Y_I49V_E155V_I156F),

[0501] (N37S_L84F_A106V_D108N_H123Y_D147Y_E155V_I156F_K161T),

[0502] (H36L_L84F_A106V_D108N_H123Y_D147Y_Q154H_E155V_I156F),

[0503] (N72S_L84F_A106V_D108N_H123Y_S146R_D147Y_E155V_I156F),

[0504] (H36L_P48L_L84F_A106V_D108N_H123Y_E134G_D147Y_E155V_I156F),

[0505] (H36L_L84F_A106V_D108N_H123Y_ D147Y_E155V_I156F_K157N)
(H36L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F),

[0506] (L84F_A106V_D108N_H123Y_S146R_D147Y_E155V_I156F_K161T),

[0507] (N37S_R51H_D77G_L84F_A106V_D108N_H123Y_D147Y_E155V_I156F),

[0508] (R51L_L84F_A106V_D108N_H123Y_D147Y_E155V_I156F_K157N),

[0509] (D24G_Q71R_L84F_H96L_A106V_D108N_H123Y_D147Y_E155V_I156F_K160E),

[0510] (H36L_G67V_L84F_A106V_D108N_H123Y_S146T_D147Y_E155V_I156F),

[0511] (Q71L_L84F_A106V_D108N_H123Y_L137M_A143E_D147Y_E155V_I156F),

[0512] (E25G_L84F_A106V_D108N_H123Y_D147Y_E155V_I156F_Q159L),

[0513] (L84F_A91T_F104I_A106V_D108N_H123Y_D147Y_E155V_I156F),

[0514] (N72D_L84F_A106V_D108N_H123Y_G125A_D147Y_E155V_I156F),

[0515] (P48S_L84F_S97C_A106V_D108N_H123Y_D147Y_E155V_I156F),

[0516] (W23G_L84F_A106V_D108N_H123Y_D147Y_E155V_I156F),

[0517] (D24G_P48L_Q71R_L84F_A106V_D108N_H123Y_D147Y_E155V_I156F_Q159L),

[0518] (L84F_A106V_D108N_H123Y_A142N_D147Y_E155V_I156F),

[0519] (H36L_R51L_L84F_A106V_D108N_H123Y_A142N_S146C_D147Y_E155V_I156F_K157N),
(N37S_L84F_A106V_D108N_H123Y_A142N_D147Y_E155V_I156F_K161T),

[0520] (L84F_A106V_D108N_D147Y_E155V_I156F),

[0521] (R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N_K161T),

[0522] (L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K161T),

[0523] (L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N_K160E_K161T),

[0524] (L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N_K160E),
(R74Q_L84F_A106V_D108N_H123Y_D147Y_E155V_I156F),

[0525] (R74A_L84F_A106V_D108N_H123Y_D147Y_E155V_I156F),

[0526] (L84F_A106V_D108N_H123Y_D147Y_E155V_I156F),

[0527] (R74Q_L84F_A106V_D108N_H123Y_D147Y_E155V_I156F),

[0528] (L84F_R98Q_A106V_D108N_H123Y_D147Y_E155V_I156F),

[0529] (L84F_A106V_D108N_H123Y_R129Q_D147Y_E155V_I156F),

[0530] (P48S_L84F_A106V_D108N_H123Y_A142N_D147Y_E155V_I156F),

[0531] (P48S_A142N),

[0532] (P48T_I49V_L84F_A106V_D108N_H123Y_A142N_D147Y_E155V_I156F_L157N),

[0533] (P48T_I49V_A142N),

[0534] (H36L_P48S_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F _K157N),
(H36L_P48S_R51L_L84F_A106V_D108N_H123Y_S146C_A142N_D147Y_E155V_I156F
(H36L_P48T_I49V_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F _K157N),

[0535] (H36L_P48T_I49V_R51L_L84F_A106V_D108N_H123Y_A142N_S146C_D147Y_E155V_ I156F _K157N),

[0536] (H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F _K157N),

[0537] (H36L_P48A_R51L_L84F_A106V_D108N_H123Y_A142N_S146C_D147Y_E155V_I156F _K157N),
(H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_A142N_D147Y_E155V_I156F _K157N),

[0538] (W23L_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F _K157N),
(W23R_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F _K157N),
(W23L_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146R_D147Y_E155V_I156F _K161T),

[0539] (H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_R152H_E155V_I156F _K157N),

[0540] (H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_R152P_E155V_I156F _K157N),

[0541] (W23L_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_R152P_E155V _I156F _K157N),

[0542] (W23L_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_A142A_S146C_D147Y_E155V_I156F _K157N),

[0543] (W23L_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_A142A_S146C_D147Y_R152P _E155V_I156F _K157N),

[0544] (W23L_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146R_D147Y_E155V_I156F _K161T),

[0545] (W23R_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_R152P_E155V _I156F _K157N),
(H36L_P48A_R51L_L84F_A106V_D108N_H123Y_A142N_S146C_D147Y_R152P_E155 V_I156F _K157N).

[0546] 일부 실시형태에서, TadA 데아미나제는 TadA 변이체이다. 일부 실시형태에서, TadA 변이체는 TadA*7.10이다. 특정 실시형태에서, 융합 단백질은 단일 TadA*7.10 도메인(*예를 들어*, 단량체로서 제공됨)을 포함한다. 다른 실시형태에서, 융합 단백질은 이종이량체를 형성할 수 있는 TadA*7.10 및 TadA(wt)를 포함한다. 일 실시형태에서, 발명의 융합 단백질은 Cas9 닉카제에 연결된, TadA*7.10에 연결된 야생형 TadA를 포함한다.

[0547] 일부 실시형태에서, TadA*7.10은 적어도 하나의 변경을 포함한다. 일부 실시형태에서, 아데노신 데아미나제는 다음 서열에서의 변경을 포함한다:

[0548] TadA*7.10
MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRV
VFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTD(서열번호: 8)

[0549] 일부 실시형태에서, TadA*7.10은 아미노산 82 및/또는 166에서 변경을 포함한다. 특정 실시형태에서, TadA*7.10은 다음 변경 중 하나 이상을 포함한다: Y147T, Y147R, Q154S, Y123H, V82S, T166R 및/또는 Q154R. 다른 실시형태에서, TadA*7.10의 변이체는 Y147T + Q154R; Y147T + Q154S; Y147R + Q154S; V82S + Q154S; V82S + Y147R; V82S + Q154R; V82S + Y123H; I76Y + V82S; V82S + Y123H + Y147T; V82S + Y123H + Y147R; V82S + Y123H + Q154R; Y147R + Q154R +Y123H; Y147R + Q154R + I76Y; Y147R + Q154R + T166R; Y123H + Y147R + Q154R + I76Y; V82S + Y123H + Y147R + Q154R; 및 I76Y + V82S + Y123H + Y147R + Q154R의 군으로부터 선택된 변경의 조합을 포함한다.

[0550] 일부 실시형태에서, 아데노신 데아미나제 변이체(*예를 들어*, TadA*8)는 결실을 포함한다. 일부 실시형태에서, 아데노신 데아미나제 변이체는 C 말단의 결실을 포함한다. 특정 실시형태에서, 아데노신 데아미나제 변이체는 TadA 참조 서열인, TadA*7.10에 비해 잔기 149, 150, 151, 152, 153, 154, 155, 156, 및 157에서 시작하는 C 말단의 결실, 또는 다른 TadA에서의 해당 돌연변이를 포함한다.

[0551] 다른 실시형태에서, 아데노신 데아미나제 변이체(*예를 들어*, TadA*8)는 TadA 참조 서열인, TadA*7.10에 비해 다음 변경: Y147T, Y147R, Q154S, Y123H, V82S, T166R 및/또는 Q154R, 또는 다른 TadA에서의 상응하는 돌연변이 중 하나 이상을 포함하는 단량체이다. 다른 실시형태에서, 아데노신 데아미나제 변이체(TadA*8)는 TadA 참조 서

열인 TadA*7.10에 비해 Y147T + Q154R; Y147T + Q154S; Y147R + Q154S; V82S + Q154S; V82S + Y147R; V82S + Q154R; V82S + Y123H; I76Y + V82S; V82S + Y123H + Y147T; V82S + Y123H + Y147R; V82S + Y123H + Q154R; Y147R + Q154R +Y123H; Y147R + Q154R + I76Y; Y147R + Q154R + T166R; Y123H + Y147R + Q154R + I76Y; V82S + Y123H + Y147R + Q154R; 및 I76Y + V82S + Y123H + Y147R + Q154R의 군으로부터 선택된 변경의 조합, 또는 다른 TadA에서의 상응하는 돌연변이를 포함하는 단량체이다.

[0552] 다른 실시형태에서, 아데노신 데아미나제 변이체는 각각 TadA 참조 서열인, TadA*7.10에 비해 다음 변경 Y147T, Y147R, Q154S, Y123H, V82S, T166R 및/또는 Q154R 중 하나 이상, 또는 다른 TadA에서의 상응하는 돌연변이를 갖는 2개의 아데노신 데아미나제 도메인(*예를 들어*, TadA*8)을 포함하는 동종이량체이다. 다른 실시형태에서, 아데노신 데아미나제 변이체는 각각 TadA 참조 서열인, TadA*7.10에 비해 Y147T + Q154R; Y147T + Q154S; Y147R + Q154S; V82S + Q154S; V82S + Y147R; V82S + Q154R; V82S + Y123H; I76Y + V82S; V82S + Y123H + Y147T; V82S + Y123H + Y147R; V82S + Y123H + Q154R; Y147R + Q154R +Y123H; Y147R + Q154R + I76Y; Y147R + Q154R + T166R; Y123H + Y147R + Q154R + I76Y; V82S + Y123H + Y147R + Q154R; 및 I76Y + V82S + Y123H + Y147R + Q154R의 군으로부터 선택된 변경의 조합, 또는 다른 TadA에서의 상응하는 돌연변이를 갖는 2개의 아데노신 데아미나제 도메인(*예를 들어*, TadA*8)을 포함하는 동종이량체이다.

[0553] 다른 실시형태에서, 아데노신 데아미나제 변이체는 야생형 아데노신 데아미나제 도메인과 TadA 참조 서열인, TadA*7.10에 비해 다음 변경 Y147T, Y147R, Q154S, Y123H, V82S, T166R 및/또는 Q154R 중 하나 이상, 또는 다른 TadA에서의 상응하는 돌연변이를 포함하는 아데노신 데아미나제 변이체 도메인(*예를 들어*, TadA*8)의 이종이량체이다. 다른 실시형태에서, 아데노신 데아미나제 변이체는 야생형 아데노신 데아미나제 도메인과 TadA 참조 서열인, TadA*7.10에 비해 Y147T + Q154R; Y147T + Q154S; Y147R + Q154S; V82S + Q154S; V82S + Y147R; V82S + Q154R; V82S + Y123H; I76Y + V82S; V82S + Y123H + Y147T; V82S + Y123H + Y147R; V82S + Y123H + Q154R; Y147R + Q154R +Y123H; Y147R + Q154R + I76Y; Y147R + Q154R + T166R; Y123H + Y147R + Q154R + I76Y; V82S + Y123H + Y147R + Q154R; 및 I76Y + V82S + Y123H + Y147R + Q154R의 군으로부터 선택된 변경의 조합, 또는 다른 TadA에서의 상응하는 돌연변이를 포함하는 아데노신 데아미나제 변이체 도메인(*예를 들어*, TadA*8)의 이종이량체이다.

[0554] 다른 실시형태에서, 아데노신 데아미나제 변이체는 TadA*7.10 도메인과 TadA 참조 서열인, TadA*7.10에 비해 다음 변경 Y147T, Y147R, Q154S, Y123H, V82S, T166R 및/또는 Q154R 중 하나 이상, 또는 다른 TadA에서의 상응하는 돌연변이를 포함하는 아데노신 데아미나제 변이체 도메인(*예를 들어*, TadA*8)의 이종이량체이다. 다른 실시형태에서, 아데노신 데아미나제 변이체는 TadA*7.10 도메인과 TadA 참조 서열인, TadA*7.10에 비해 Y147T + Q154R; Y147T + Q154S; Y147R + Q154S; V82S + Q154S; V82S + Y147R; V82S + Q154R; V82S + Y123H; I76Y + V82S; V82S + Y123H + Y147T; V82S + Y123H + Y147R; V82S + Y123H + Q154R; Y147R + Q154R +Y123H; Y147R + Q154R + I76Y; Y147R + Q154R + T166R; Y123H + Y147R + Q154R + I76Y; V82S + Y123H + Y147R + Q154R; 및 I76Y + V82S + Y123H + Y147R + Q154R의 군으로부터 선택된 변경의 조합, 또는 다른 TadA에서의 상응하는 돌연변이를 포함하는 아데노신 데아미나제 변이체 도메인(*예를 들어*, TadA*8)의 이종이량체이다.

[0555] 특정 실시형태에서, 아데노신 데아미나제 이종이량체는 TadA*8 도메인과 스타필로코커스 아우레우스(S. 아우레우스) TadA, 바실루스 서브틸리스(B. 서브틸리스) TadA, 살모넬라 티피무리움(S. 티피무리움) TadA, 쉬와넬라 푸트리파시엔스(S. 푸트리파시엔스) TadA, 헤모필루스 인플루엔자 F3031(H. 인플루엔자) TadA, 카울로박터 크레센투스(C. 크레센투스) TadA, 게오박터 술푸레두센스(G. 술푸레두센스) TadA 또는 TadA*7.10으로부터 선택된 아데노신 데아미나제 도메인을 포함한다.

[0556] 일부 실시형태에서, 아데노신 데아미나제는 TadA*8이다. 일 실시형태에서, 아데노신 데아미나아제는 아데노신 데아미나아제 활성을 갖는 다음 서열 또는 이의 단편을 포함하거나 본질적으로 구성되는 TadA*8이다:

[0557] MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRV VFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCTFFRMPRQVFNAQKKAQSSTD(서열번호: 12)

[0558] 일부 실시형태에서, TadA*8은 절삭된다. 일부 실시형태에서, 절삭된 TadA*8에는 전장 TadA*8에 비해 1, 2, 3, 4, 5,6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 6, 17, 18, 19, 또는 20개 N-말단 아미노산 잔기가 없다. 일부 실시형태에서, 절삭된 TadA*8에는 전장 TadA*8에 비해 1, 2, 3, 4, 5,6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 6, 17, 18, 19, 또는 20개 C-말단 아미노산 잔기가 없다. 일부 실시형태에서 아데노신 데아미나제 변이체는 전장 TadA*8이다.

[0559] 일부 실시형태에서 TadA*8은 TadA*8.1, TadA*8.2, TadA*8.3, TadA*8.4, TadA*8.5, TadA*8.6, TadA*8.7, TadA*8.8, TadA*8.9, TadA*8.10, TadA*8.11, TadA*8.12, TadA*8.13, TadA*8.14, TadA*8.15, TadA*8.16, TadA*8.17, TadA*8.18, TadA*8.19, TadA*8.20, TadA*8.21, TadA*8.22, TadA*8.23, 또는 TadA*8.24이다.

[0560] 다른 실시형태에서, 개시내용의 염기 편집기는 TadA 참조 서열인, TadA*7.10에 비해 다음 변경: R26C, V88A, A109S, T111R, D119N, H122N, Y147D, F149Y, T166I 및/또는 D167N 중 하나 이상, 또는 다른 TadA에서의 상응하는 돌연변이를 포함하는 아데노신 데아미나제 변이체(*예를 들어*, TadA*8) 단량체를 포함한다. 다른 실시형태에서, 아데노신 데아미나제 변이체(TadA*8) 단량체는 TadA 참조 서열인, TadA*7.10에 비해 R26C + A109S + T111R + D119N + H122N + Y147D + F149Y + T166I + D167N; V88A + A109S + T111R + D119N + H122N + F149Y + T166I + D167N; R26C + A109S + T111R + D119N + H122N + F149Y + T166I + D167N; V88A + T111R + D119N + F149Y; 및 A109S + T111R + D119N + H122N + Y147D + F149Y + T166I + D167N의 군으로부터 선택된 변경의 조합, 또는 다른 TadA에서의 상응하는 돌연변이를 포함한다.

[0561] 다른 실시형태에서, 염기 편집기는 야생형 아데노신 데아미나제 도메인과 TadA 참조 서열인, TadA*7.10에 비해 다음 변경 R26C, V88A, A109S, T111R, D119N, H122N, Y147D, F149Y, T166I 및/또는 D167N 중 하나 이상, 또는 다른 TadA에서의 상응하는 돌연변이를 포함하는 아데노신 데아미나제 변이체 도메인(*예를 들어*, TadA*8)의 이종이량체를 포함한다. 다른 실시형태에서, 염기 편집기는 야생형 아데노신 데아미나제 도메인과 TadA 참조 서열인, TadA*7.10에 비해 R26C + A109S + T111R + D119N + H122N + Y147D + F149Y + T166I + D167N; V88A + A109S + T111R + D119N + H122N + F149Y + T166I + D167N; R26C + A109S + T111R + D119N + H122N + F149Y + T166I + D167N; V88A + T111R + D119N + F149Y; 및 A109S + T111R + D119N + H122N + Y147D + F149Y + T166I + D167N의 군으로부터 선택된 변경의 조합, 또는 다른 TadA에서의 상응하는 돌연변이를 포함하는 아데노신 데아미나제 변이체 도메인(*예를 들어*, TadA*8)의 이종이량체를 포함한다.

[0562] 다른 실시형태에서, 염기 편집기는 TadA*7.10 도메인과 TadA 참조 서열인, TadA*7.10에 비해 다음 변경 R26C, V88A, A109S, T111R, D119N, H122N, Y147D, F149Y, T166I 및/또는 D167N 중 하나 이상, 또는 다른 TadA에서의 상응하는 돌연변이를 포함하는 아데노신 데아미나제 변이체 도메인(*예를 들어*, TadA*8)의 이종이량체를 포함한다. 다른 실시형태에서, 염기 편집기는 TadA*7.10 도메인과 TadA 참조 서열인, TadA*7.10에 비해 R26C + A109S + T111R + D119N + H122N + Y147D + F149Y + T166I + D167N; V88A + A109S + T111R + D119N + H122N + F149Y + T166I + D167N; R26C + A109S + T111R + D119N + H122N + F149Y + T166I + D167N; V88A + T111R + D119N + F149Y; 및 A109S + T111R + D119N + H122N + Y147D + F149Y + T166I + D167N의 군으로부터 선택된 변경의 조합, 또는 다른 TadA에서의 상응하는 돌연변이를 포함하는 아데노신 데아미나제 변이체 도메인(*예를 들어*, TadA*8)의 이종이량체를 포함한다.

[0563] 일부 실시형태에서, TadA*8은 표 5에 나타난 바와 같은 변이체이다. 표 5는 TadA 아미노산 서열에서의 특정 아미노산 위치 번호 및 TadA-7.10 아데노신 데아미나제에서의 이들 위치에 존재하는 아미노산을 나타낸다. 표 5는 또한, 그 전체 내용이 본 명세서에 참조로 포함된, M. Richter 등, 2020, Nature Biotechnology, doi.org/10.1038/s41587-020-0453-z에 기술된 바와 같이 파지-보조된 비-연속 진화(PANCE) 및 파지-보조된 연속 진화(PACE)가 이어지는 TadA-7.10에 비해 TadA 변이체에서의 아미노산 변화를 나타낸다. 일부 실시형태에서, TadA*8은 TadA*8a, TadA*8b, TadA*8c, TadA*8d 또는 TadA*8e이다. 일부 실시형태에서, TadA*8은 TadA*8e이다.

[0564] **표 5. TadA*8 변이체 선택**

		TadA 아미노산 번호									
	TadA	26	88	109	111	119	122	147	149	166	167
	TadA-7.10	R	V	A	T	D	H	Y	F	T	D
PANCE 1					R						
PANCE 2				S/T	R						
	TadA-8a	C		S	R	N	N	D	Y	I	N
	TadA-8b		A	S	R	N	N		Y	I	N
PACE	TadA-8c	C		S	R	N	N		Y	I	N
	TadA-8d		A		R	N			Y		
	TadA-8e			S	R	N	N	D	Y	I	N

[0565]

[0566] 일 실시형태에서, 발명의 융합 단백질은 Cas9 닉카제에 연결된, 본 명세서에 기재된 아데노신 데아미나제 변이체(*예를 들어*, TadA*8)에 연결된 야생형 TadA를 포함한다. 특정 실시형태에서, 융합 단백질은 단일 TadA*8 도메인(*예를 들어*, 단량체로서 제공됨)을 포함한다. 다른 실시형태에서, 융합 단백질은 이종이량체를 형성할 수 있는 TadA*8 및 TadA(wt)를 포함한다.

[0567] 일부 실시형태에서, 아데노신 데아미나제는 본 명세서에 제공된 임의의 아데노신 데아미나제에 제시된 아미노산 서열 중 임의의 하나와 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 적어도 99.5% 동일한 아미노산 서열을 포함한다. 본 명세서에 제공된 아데노신 데아미나제는 하나 이상의 돌연변이(*예를 들어*, 본 명세서에 제공된 임의의 돌연변이)를 포함할 수 있음을 이해해야 한다. 개시내용은 특정 퍼센트 동일성을 갖는 임의의 데아미나제 도메인 플러스 본 명세서에 기재된 임의의 돌연변이 또는 이의 조합을 제공한다. 일부 실시형태에서, 아데노신 데아미나제는 참조 서열, 또는 본 명세서에 제공된 임의의 아데노신 데아미나제와 비교하여 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 21, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50 이상의 돌연변이를 갖는 아미노산 서열을 포함한다. 일부 실시형태에서, 아데노신 데아미나제는 당업계에 공지되어 있거나 본 명세서에 기재된 아미노산 서열 중 임의의 하나와 비교하여 적어도 5개, 적어도 10개, 적어도 15개, 적어도 20개, 적어도 25개, 적어도 30개, 적어도 35개, 적어도 40개, 적어도 45개, 적어도 50개, 적어도 60개, 적어도 70개, 적어도 80개, 적어도 90개, 적어도 100개, 적어도 110개, 적어도 120개, 적어도 130개, 적어도 140개, 적어도 150개, 적어도 160개 또는 적어도 170개의 동일한 연속 아미노산 잔기를 갖는 아미노산 서열을 포함한다.

[0568] 특정 실시형태에서, TadA*8은 굵게 도시된 다음 위치 중 임의의 것에서 하나 이상의 돌연변이를 포함한다. 다른 실시형태에서, TadA*8은 밑줄로 도시된 위치 중 임의의 것에서 하나 이상의 돌연변이를 포함한다:

[0569] MSEVEFSHEY WMRHALTLAK RARDEREVPV GAVLVLNNRV IGEGWNRAIG [50] LHDPTAHAEI MALRQGGLVM QNYRLIDATL YVTFEPCVMC AGAMIHSRIG [100] RVVFGVRNAK TGAAGSLMDV LHYPGMNHRV EITEGILADE CAALLCYFFR [150] MPRQVFNAQK KAQSSTD(서열번호: 8)

[0570] 예를 들어, TadA*8은 아미노산 위치 82 및/또는 166(*예를 들어*, V82S, T166R)에서의 변경을 단독으로 또는 TadA 참조 서열인, TadA*7.10에 비하여 다음 Y147T, Y147R, Q154S, Y123H 및/또는 Q154R 중 임의의 하나 이상, 또는 다른 TadA에서의 상응하는 돌연변이와 조합하여 포함한다. 특정 실시형태에서, 변경의 조합은 TadA 참조 서열인, TadA*7.10에 비하여 Y147T + Q154R; Y147T + Q154S; Y147R + Q154S; V82S + Q154S; V82S + Y147R; V82S + Q154R; V82S + Y123H; I76Y + V82S; V82S + Y123H + Y147T; V82S + Y123H + Y147R; V82S + Y123H + Q154R; Y147R + Q154R +Y123H; Y147R + Q154R + I76Y; Y147R + Q154R + T166R; Y123H + Y147R + Q154R + I76Y; V82S + Y123H + Y147R + Q154R; 및 I76Y + V82S + Y123H + Y147R + Q154R의 군, 또는 다른 TadA에서의 상응하는 돌연변이로부터 선택된다.

[0571] 일부 실시형태에서, TadA*8은 절삭된다. 일부 실시형태에서, 절삭된 TadA*8에는 전장 TadA*8에 비해 1, 2, 3, 4, 5,6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 6, 17, 18, 19, 또는 20개의 N-말단 아미노산 잔기가 없다. 일부 실시형태에서, 절삭된 TadA*8에는 전장 TadA*8에 비해 1, 2, 3, 4, 5,6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 6, 17, 18, 19, 또는 20개의 C-말단 아미노산 잔기가 없다. 일부 실시형태에서 아데노신 데아미나제 변이체는 전장 TadA*8이다.

[0572] 일 실시형태에서, 발명의 융합 단백질은 Cas9 닉카제에 연결된, 본 명세서에 기재된 아데노신 데아미나제 변이체(*예를 들어*, TadA*8)에 연결된 야생형 TadA를 포함한다. 특정 실시형태에서, 융합 단백질은 단일 TadA*8 도메인(*예를 들어*, 단량체로서 제공됨)을 포함한다. 다른 실시형태에서, 염기 편집기는 이종이량체를 형성할 수 있는 TadA*8 및 TadA(wt)를 포함한다.

[0573] 특정 실시형태에서, 융합 단백질은 단일(*예를 들어*, 단량체로서 제공됨) TadA*8을 포함한다. 일부 실시형태에서, TadA*8은 Cas9 닉카제에 연결된다. 일부 실시형태에서, 발명의 융합 단백질은 TadA*8에 연결된 야생형 TadA(TadA(wt))의 이종이량체로서 포함한다. 다른 실시형태에서, 발명의 융합 단백질은 TadA*8에 연결된 TadA*7.10의 이종이량체로서 포함한다. 일부 실시형태에서, 염기 편집기는 TadA*8 변이체 단량체를 포함하는 ABE8이다. 일부 실시형태에서, 염기 편집기는 TadA*8 및 TadA(wt)의 이종이량체를 포함하는 ABE8이다. 일부 실시형태에서, 염기 편집기는 TadA*8 및 TadA*7.10의 이종이량체를 포함하는 ABE8이다. 일부 실시형태에서, 염기 편집기는 TadA*8의 이종이량체를 포함하는 ABE8이다. 일부 실시형태에서, TadA*8은 표 11, 13 또는 14로부터 선택된다. 일부 실시형태에서, ABE8은 표 13, 14 또는 16으로부터 선택된다.

[0574] 일부 실시형태에서, 아데노신 데아미나제는 TadA*9 변이체이다. 일부 실시형태에서, 아데노신 데아미나제는 다음 서열을 참조하여 하기 기재된 변이체로부터 선택된 TadA*9 변이체(TadA*7.10으로 지칭됨)이다:

[0575] MSEVEFSHEY WMRHALTLAK RARDEREVPV GAVLVLNNRV IGEGWNRAIG

[0576] LHDPTAHAEI MALRQGGLVM QNYRLIDATL YVTFEPCVMC AGAMIHSRIG

[0577] RVVFGVRNAK TGAAGSLMDV LHYPGMNHRV EITEGILADE CAALLCYFFR

[0578] MPRQVFNAQK KAQSSTD(서열번호: 8).

[0579] 일부 실시형태에서, 아데노신 데아미나제는 다음 변경 중 하나 이상을 포함한다: R21N, R23H, E25F, N38G, L51W, P54C, M70V, Q71M, N72K, Y73S, V82T, M94V, P124W, T133K, D139L, D139M, C146R, 및 A158K. 하나 이상의 변경은 상기 서열에서 밑줄과 굵은 글꼴로 도시된다.

[0580] 일부 실시형태에서, 아데노신 데아미나제는 다음 변경의 조합 중 하나 이상을 포함한다: V82S + Q154R + Y147R; V82S + Q154R + Y123H; V82S + Q154R + Y147R+ Y123H; Q154R + Y147R + Y123H + I76Y+ V82S; V82S + I76Y; V82S + Y147R; V82S + Y147R + Y123H; V82S + Q154R + Y123H; Q154R + Y147R + Y123H + I76Y; V82S + Y147R; V82S + Y147R + Y123H; V82S + Q154R + Y123H; V82S + Q154R + Y147R; V82S + Q154R + Y147R; Q154R + Y147R + Y123H + I76Y; Q154R + Y147R + Y123H + I76Y + V82S; I76Y_V82S_Y123H_Y147R_Q154R; Y147R + Q154R + H123H; 및 V82S + Q154R.

[0581] 일부 실시형태에서, 아데노신 데아미나제는 다음 변경의 조합 중 하나 이상을 포함한다: E25F + V82S + Y123H, T133K + Y147R + Q154R; E25F + V82S + Y123H + Y147R + Q154R; L51W + V82S + Y123H + C146R + Y147R + Q154R; Y73S + V82S + Y123H + Y147R + Q154R; P54C + V82S + Y123H + Y147R + Q154R; N38G + V82T + Y123H + Y147R + Q154R; N72K + V82S + Y123H + D139L + Y147R + Q154R; E25F + V82S + Y123H + D139M + Y147R + Q154R; Q71M + V82S + Y123H + Y147R + Q154R; E25F + V82S + Y123H + T133K + Y147R + Q154R; E25F + V82S + Y123H + Y147R + Q154R; V82S + Y123H + P124W + Y147R + Q154R; L51W + V82S + Y123H + C146R + Y147R + Q154R; P54C + V82S + Y123H + Y147R + Q154R; Y73S + V82S + Y123H + Y147R + Q154R; N38G + V82T + Y123H + Y147R + Q154R; R23H + V82S + Y123H + Y147R + Q154R; R21N + V82S + Y123H + Y147R + Q154R; V82S + Y123H + Y147R + Q154R + A158K; N72K + V82S + Y123H + D139L + Y147R + Q154R; E25F + V82S + Y123H + D139M + Y147R + Q154R; 및 M70V + V82S + M94V + Y123H + Y147R + Q154R

[0582] 일부 실시형태에서, 아데노신 데아미나제는 다음 변경의 조합 중 하나 이상을 포함한다: Q71M + V82S + Y123H + Y147R + Q154R; E25F + I76Y+ V82S + Y123H + Y147R + Q154R; I76Y + V82T + Y123H + Y147R + Q154R; N38G + I76Y + V82S + Y123H + Y147R + Q154R; R23H + I76Y + V82S + Y123H + Y147R + Q154R; P54C + I76Y + V82S + Y123H + Y147R + Q154R; R21N + I76Y + V82S + Y123H + Y147R + Q154R; I76Y + V82S + Y123H + D139M + Y147R + Q154R; Y73S + I76Y + V82S + Y123H + Y147R + Q154R; E25F + I76Y + V82S + Y123H + Y147R + Q154R; I76Y + V82T + Y123H + Y147R + Q154R; N38G + I76Y + V82S + Y123H + Y147R + Q154R; R23H + I76Y + V82S + Y123H + Y147R + Q154R; P54C + I76Y + V82S + Y123H + Y147R + Q154R; R21N + I76Y + V82S + Y123H + Y147R + Q154R; I76Y + V82S + Y123H + D139M + Y147R + Q154R; Y73S + I76Y + V82S + Y123H + Y147R + Q154R; 및 V82S + Q154R; N72K_V82S + Y123H + Y147R + Q154R; Q71M_V82S + Y123H + Y147R + Q154R; V82S + Y123H + T133K + Y147R + Q154R; V82S + Y123H + T133K + Y147R + Q154R + A158K; M70V +Q71M +N72K +V82S + Y123H + Y147R + Q154R; N72K_V82S + Y123H + Y147R + Q154R; Q71M_V82S + Y123H + Y147R + Q154R; M70V +V82S + M94V + Y123H + Y147R + Q154R; V82S + Y123H + T133K + Y147R + Q154R; V82S + Y123H + T133K + Y147R + Q154R + A158K; 및 M70V +Q71M +N72K +V82S + Y123H + Y147R + Q154R. 일부 실시형태에서, 아데노신 데아미나제는 단량체로 발현된다. 다른 실시형태에서, 아데노신 데아미나제는 이종이량체로 발현된다. 일부 실시형태에서, 데아미나제 또는 다른 폴리펩티드 서열은, 예를 들어 융합 단백질의 성분으로 포함될 때 메티오닌을 결한다. 이는 위치의 번호지정을 변경할 수 있다. 그러나, 당업자는 이러한 상응하는 돌연변이가 동일한 돌연변이, *예를 들어* Y73S와 Y72S 및 D139M과 D138M을 지칭한다는 것을 이해할 것이다.

[0583] 일부 실시형태에서, TadA*9 변이체는 본 명세서에 기재된 바와 같을 표 17에 기재된 변경을 포함한다. 일부 실시형태에서, TadA*9 변이체는 단량체이다. 일부 실시형태에서, TadA*9 변이체는 야생형 TadA 아데노신 데아미나제를 갖는 이종이량체이다. 일부 실시형태에서, TadA*9 변이체는 또 다른 TadA 변이체(*예를 들어*, TadA*8, TadA*9)를 갖는 이종이량체이다. TadA*9 아데노신 데아미나제의 추가 세부사항은 국제 PCT 출원 번호 PCT/2020/049975에 기재되어 있으며, 이는 그 전체가 본 명세서에 참고로 포함된다.

[0584] 본 명세서에 제공된 임의의 돌연변이 및 임의의 추가 돌연변이(*예를 들어*, ecTadA 아미노산 서열을 기반으로 함)는 임의의 다른 아데노신 데아미나제 안으로 도입될 수 있다. 본 명세서에 제공된 임의의 돌연변이는 개별적

으로 또는 TadA 참조 서열 또는 다른 아데노신 데아미나제(*예를 들어*, ecTadA)에서 임의의 조합으로 만들어 질 수 있다.

[0585] A에서 G 핵염기 편집 단백질의 세부사항은 국제 PCT 출원 번호 PCT/2017/045381(WO2018/027078) 및 Gaudelli, N.M., 등, "Programmable base edit of A·T to G·C in genomic DNA without DNA cleavage" Nature, 551, 464-471 (2017)"에 기술되어 있으며, 그 전체 내용은 본 명세서에 참조로 포함된다.

[0586] ***C에서 T 편집***

[0587] 일부 실시형태에서, 본 명세서에 개시된 염기 편집기는 폴리뉴클레오티드의 표적 시티딘(C) 염기를 탈아민화하여 티민의 염기 페어링 특성을 갖는 우리딘(U)을 생성할 수 있는 시티딘 데아미나제를 포함하는 융합 단백질을 포함한다. 예를 들어 폴리뉴클레오티드가 이중-가닥(*예를 들어*, DNA)인 일부 실시형태에서, 우리딘 염기는 그 다음 (*예를 들어*, 세포 수선 기계에 의해) 티미딘 염기로 치환되어 C:G에서 T:A 전환을 일으킬 수 있다. 다른 실시형태에서, 염기 편집기에 의한 핵산에서의 C에서 U로의 탈아미노화는 U에서 T로의 치환을 수반할 수 없다.

[0588] U를 발생시키기 위한 폴리뉴클레오티드에서의 표적 C의 탈아미노화는 본 명세서에 기재된 염기 편집기에 의해 실행될 수 있는 염기 편집 유형의 비-제한적인 예이다. 또 다른 예에서, 시티딘 데아미나제 도메인을 포함하는 염기 편집기는 시토신(C) 염기의 구아닌(G) 염기로의 전환을 매개할 수 있다. 예를 들어, 염기 편집기의 시티딘 데아미나제 도메인에 의한 시티딘의 탈아미노화에 의해 생성된 폴리뉴클레오티드의 U는 염기 (*예를 들어*, 우라실 DNA 글리코실라제(UDG) 도메인에 의해) 절단 수선 메커니즘에 의해 폴리뉴클레오티드로부터 절단되어 비염기성 부위를 생성할 수 있다. 비염기성 부위의 반대편 핵염기는 그 다음 예를 들어 트랜스레시온 폴리머라제에 의해 C와 같은 또 다른 염기로 (*예를 들어*, 염기 수선 기계에 의해) 대체될 수 있다. 비염기성 부위 반대편의 핵염기가 C로 대체되는 것이 전형적이지만, 다른 치환(*예를 들어*, A, G 또는 T)도 발생할 수 있다.

[0589] 따라서, 일부 실시형태에서 본 명세서에 기재된 염기 편집기는 폴리뉴클레오티드에서 표적 C를 U로 탈아미노화할 수 있는 탈아미노화 도메인(*예를 들어*, 시티딘 데아미나제 도메인)을 포함한다. 더욱이, 아래에 기술된 바와 같이, 염기 편집기는 탈아미노화로 인한 U의, 일부 실시형태에서, T 또는 G로의 전환을 촉진하는 추가 도메인을 포함할 수 있다. 예를 들어, 시티딘 데아미나제 도메인을 포함하는 염기 편집기는 우라실 글리코실라제 억제제(UGI) 도메인을 추가로 포함하여 U의 T에 의한 치환을 매개하여 C-대-T 염기 편집 이벤트를 완료할 수 있다. 또 다른 예에서, 염기 편집기는 트랜스레시온 폴리머라제가 비염기성 부위 반대편에 C의 통합을 촉진할 수 있기 때문에(즉, 비염기성 부위에서 G의 통합을 초래하여, C-대-G 염기 편집 이벤트를 완료함), C-대-G 염기 편집의 효율성을 향상시키기 위해 트랜스레시온 폴리머라제를 통합할 수 있다.

[0590] 시티딘 데아미나제를 도메인으로 포함하는 염기 편집기는 DNA, RNA 및 DNA-RNA 하이브리드를 포함한 임의의 폴리뉴클레오티드에서 표적 C를 탈아미노화할 수 있다. 전형적으로, 시티딘 데아미나제는 폴리뉴클레오티드의 단일-가닥 부분의 맥락에서 위치된 C 핵염기를 촉매한다. 일부 실시형태에서, 표적 C를 포함하는 전체 폴리뉴클레오티드는 단일-가닥일 수 있다. 예를 들어, 염기 편집기 안으로 통합된 시티딘 데아미나제는 단일-가닥 RNA 폴리뉴클레오티드에서의 표적 C를 탈아미노화할 수 있다. 다른 실시형태에서, 시티딘 데아미나제 도메인을 포함하는 염기 편집기는 이중-가닥 폴리뉴클레오티드에 작용할 수 있지만, 표적 C는 탈아미노화 반응 시 단일-가닥 상태에 있는 폴리뉴클레오티드의 부분에 위치될 수 있다. 예를 들어, NAGPB 도메인이 Cas9 도메인을 포함하는 실시형태에서, 여러 뉴클레오티드는 Cas9-gRNA-표적 DNA 복합체의 형성 동안 쌍을 이루지 않은 상태로 남아 Cas9 "R-루프 복합체"의 형성을 초래할 수 있다. 이들 쌍을 이루지 않은 뉴클레오티드는 단일-가닥 특정 뉴클레오티드 데아미나제 효소(예를 들어, 시티딘 데아미나제)에 대한 기질 역할을 할 수 있는 단일-가닥 DNA의 버블을 형성할 수 있다.

[0591] 일부 실시형태에서, 염기 편집기의 시티딘 데아미나제는 아포지단백질 B mRNA 편집 복합체(APOBEC) 계열 데아미나제의 전부 또는 일부를 포함할 수 있다. APOBEC은 진화적으로 보존된 시티딘 데아미나제의 계열이다. 이 계열의 구성원은 C-대-U 편집 효소이다. APOBEC 유사 단백질의 N-말단 도메인은 촉매적 도메인인 반면, C-말단 도메인은 의사촉매적 도메인이다. 보다 구체적으로, 촉매적 도메인은 아연 의존성 시티딘 데아미나제 도메인이고 시티딘 탈아미노화에 중요하다. APOBEC 계열 구성원은 APOBEC1, APOBEC2, APOBEC3A, APOBEC3B, APOBEC3C, APOBEC3D(현재 "APOBEC3E"는 이를 지칭함), APOBEC3F, APOBEC3G, APOBEC3H, APOBEC4 및 활성화-유도된(시티딘) 데아미나제를 포함한다. 일부 실시형태에서, 염기 편집기 안으로 합체된 데아미나제는 APOBEC1 데아미나제의 전부 또는 일부를 포함한다. 일부 실시형태에서, 염기 편집기 안으로 합체된 데아미나제는 APOBEC2 데아미나제의 전부 또는 일부를 포함한다. 일부 실시형태에서, 염기 편집기 안으로 합체된 데아미나제는 APOBEC3 데아미나제의 전부 또는 일부를 포함한다. 일부 실시형태에서, 염기 편집기 안으로 합체된 데아미나제는 APOBEC3A 데아미

나제의 전부 또는 일부를 포함한다. 일부 실시형태에서, 염기 편집기 안으로 합체된 데아미나제는 APOBEC3B 데아미나제의 전부 또는 일부를 포함한다. 일부 실시형태에서, 염기 편집기 안으로 합체된 데아미나제는 APOBEC3C 데아미나제의 전부 또는 일부를 포함한다. 일부 실시형태에서, 염기 편집기 안으로 합체된 데아미나제는 APOBEC3D 데아미나제의 전부 또는 일부를 포함한다. 일부 실시형태에서, 염기 편집기 안으로 합체된 데아미나제는 APOBEC3E 데아미나제의 전부 또는 일부를 포함한다. 일부 실시형태에서, 염기 편집기 안으로 합체된 데아미나제는 APOBEC3F 데아미나제의 전부 또는 일부를 포함한다. 일부 실시형태에서, 염기 편집기 안으로 합체된 데아미나제는 APOBEC3G 데아미나제의 전부 또는 일부를 포함한다. 일부 실시형태에서, 염기 편집기 안으로 합체된 데아미나제는 APOBEC3H 데아미나제의 전부 또는 일부를 포함한다. 일부 실시형태에서, 염기 편집기 안으로 합체된 데아미나제는 APOBEC4 데아미나제의 전부 또는 일부를 포함한다. 일부 실시형태에서, 염기 편집기 안으로 합체된 데아미나제는 활성화-유도된 데아미나제(AID)의 전부 또는 일부를 포함한다. 일부 실시형태에서 염기 편집기 안으로 합체된 데아미나제는 시티딘 데아미나제 1(CDA1)의 전부 또는 일부를 포함한다. 염기 편집기는 임의의 적합한 유기체(예를 들어, 인간 또는 랫트)로부터의 데아미나제를 포함할 수 있음을 이해해야 한다. 일부 실시형태에서, 염기 편집기의 데아미나제 도메인은 인간, 침팬지, 고릴라, 원숭이, 소, 개, 랫트 또는 마우스로부터의 것이다. 일부 실시형태에서, 염기 편집기의 데아미나제 도메인은 랫트(예를 들어, 랫트 APOBEC1)로부터 유래된다. 일부 실시형태에서, 염기 편집기의 데아미나제 도메인은 인간 APOBEC1이다. 일부 실시형태에서, 염기 편집기의 데아미나제 도메인은 pmCDA1이다.

[0592] 본 개시내용의 양태에 따라 Cas9에 융합될 수 있는 다른 예시적인 데아미나제가 아래에 제공된다. 실시형태에서, 데아미나제는 활성화-유도된 데아미나제(AID)이다. 일부 실시형태에서, 각각의 서열의 활성 도메인, *예를 들어* 국소화 신호를 갖지 않는 도메인(핵 방출 신호, 세포질 국소화 신호를 갖지 않는 핵 국소화 서열)이 사용될 수 있다는 것이 이해되어야 한다.

[0593] 본 개시내용의 일부 양태는, 예를 들어, 데아미나제 도메인에서 점 돌연변이를 만듦에 의해 본 명세서에 기재된 임의의 융합 단백질의 데아미나제 도메인 촉매적 활성을 조절하는 것이 융합 단백질의 진행성(*예를 들어*, 염기 편집기)에 영향을 미친다는 인식에 기반한다. 예를 들어, 염기 편집 융합 단백질 내에서 데아미나제 도메인의 촉매적 활성을 감소시키지만 제거하지는 않는 돌연변이는 데아미나제 도메인이 표적 잔기에 인접한 잔기의 탈아미노화를 촉매할 가능성을 줄여서 탈아미노화 윈도우를 좁힐 수 있다. 탈아미노화 윈도우를 좁힐 수 있는 능력은 특정 표적 잔기에 인접한 잔기의 원치 않는 탈아미노화를 방지할 수 있으며, 이는 표적-외 효과를 감소시키거나 방지할 수 있다.

[0594] 예를 들어, 일부 실시형태에서, 염기 편집기 안으로 합체된 APOBEC 데아미나제는 rAPOBEC1의 H121X, H122X, R126X, R126X, R118X, W90X, W90X 및 R132X로 구성된 군으로부터 선택된 하나 이상의 돌연변이, 또는 다른 APOBEC 데아미나제에서 하나 이상의 상응하는 돌연변이를 포함할 수 있으며, 여기서 X는 임의의 아미노산이다. 일부 실시형태에서, 염기 편집기 안으로 합체된 APOBEC 데아미나제는 rAPOBEC1의 H121R, H122R, R126A, R126E, R118A, W90A, W90Y 및 R132E로 구성된 군으로부터 선택된 하나 이상의 돌연변이, 또는 다른 APOBEC 데아미나제에서 하나 이상의 상응하는 돌연변이를 포함할 수 있다.

[0595] 일부 실시형태에서, 염기 편집기 안으로 합체된 APOBEC 데아미나제는 hAPOBEC3G의 D316X, D317X, R320X, R320X, R313X, W285X, W285X, R326X로 구성된 군으로부터 선택된 하나 이상의 돌연변이, 또는 다른 APOBEC 데아미나제에서 하나 이상의 상응하는 돌연변이를 포함할 수 있으며, 여기서 X는 임의의 아미노산이다. 일부 실시형태에서, 본 명세서에 제공된 임의의 융합 단백질은 hAPOBEC3G의 D316R, D317R, R320A, R320E, R313A, W285A, W285Y, R326E로 구성된 군으로부터 선택된 하나 이상의 돌연변이, 또는 다른 APOBEC 데아미나제에서 하나 이상의 상응하는 돌연변이를 포함한다.

[0596] 일부 실시형태에서, 염기 편집기 안으로 합체된 APOBEC 데아미나제는 rAPOBEC1의 H121R 및 H122R 돌연변이, 또는 다른 APOBEC 데아미나제에서 하나 이상의 상응하는 돌연변이를 포함할 수 있다. 일부 실시형태에서, 염기 편집기 안으로 합체된 APOBEC 데아미나제는 rAPOBEC1의 R126A 돌연변이, 또는 다른 APOBEC 데아미나제에서 하나 이상의 상응하는 돌연변이를 포함하는 APOBEC 데아미나제를 포함할 수 있다. 일부 실시형태에서, 염기 편집기 안으로 합체된 APOBEC 데아미나제는 rAPOBEC1의 R126E 돌연변이, 또는 다른 APOBEC 데아미나제에서 하나 이상의 상응하는 돌연변이를 포함하는 APOBEC 데아미나제를 포함할 수 있다. 일부 실시형태에서, 염기 편집기 안으로 합체된 APOBEC 데아미나제는 rAPOBEC1의 R118A 돌연변이, 또는 다른 APOBEC 데아미나제에서 하나 이상의 상응하는 돌연변이를 포함하는 APOBEC 데아미나제를 포함할 수 있다. 일부 실시형태에서, 염기 편집기 안으로 합체된 APOBEC 데아미나제는 rAPOBEC1의 W90A 돌연변이, 또는 다른 APOBEC 데아미나제에서 하나 이상의 상응하는 돌연변이를 포함하는 APOBEC 데아미나제를 포함할 수 있다. 일부 실시형태에서, 염기 편집기 안으로 합체된 APOBEC

데아미나제는 rAPOBEC1의 W90Y 돌연변이, 또는 다른 APOBEC 데아미나제에서 하나 이상의 상응하는 돌연변이를 포함하는 APOBEC 데아미나제를 포함할 수 있다. 일부 실시형태에서, 염기 편집기 안으로 합체된 APOBEC 데아미나제는 rAPOBEC1의 R132E 돌연변이, 또는 다른 APOBEC 데아미나제에서 하나 이상의 상응하는 돌연변이를 포함하는 APOBEC 데아미나제를 포함할 수 있다. 일부 실시형태에서, 염기 편집기 안으로 합체된 APOBEC 데아미나제는 rAPOBEC1의 W90Y 및 R126E 돌연변이, 또는 다른 APOBEC 데아미나제에서 하나 이상의 상응하는 돌연변이를 포함하는 APOBEC 데아미나제를 포함할 수 있다. 일부 실시형태에서, 염기 편집기 안으로 합체된 APOBEC 데아미나제는 rAPOBEC1의 R126E 및 R132E 돌연변이, 또는 다른 APOBEC 데아미나제에서 하나 이상의 상응하는 돌연변이를 포함하는 APOBEC 데아미나제를 포함할 수 있다. 일부 실시형태에서, 염기 편집기 안으로 합체된 APOBEC 데아미나제는 rAPOBEC1의 W90Y 및 R132E 돌연변이, 또는 다른 APOBEC 데아미나제에서 하나 이상의 상응하는 돌연변이를 포함하는 APOBEC 데아미나제를 포함할 수 있다. 일부 실시형태에서, 염기 편집기 안으로 합체된 APOBEC 데아미나제는 rAPOBEC1의 W90Y, R126E 및 R132E 돌연변이, 또는 다른 APOBEC 데아미나제에서 하나 이상의 상응하는 돌연변이를 포함하는 APOBEC 데아미나제를 포함할 수 있다.

[0597] 일부 실시형태에서, 염기 편집기 안으로 합체된 APOBEC 데아미나제는 hAPOBEC3G의 D316R 및 D317R 돌연변이, 또는 다른 APOBEC 데아미나제에서 하나 이상의 상응하는 돌연변이를 포함하는 APOBEC 데아미나제를 포함할 수 있다. 일부 실시형태에서, 본 명세서에 제공된 임의의 융합 단백질은 hAPOBEC3G의 R320A 돌연변이, 또는 다른 APOBEC 데아미나제에서 하나 이상의 상응하는 돌연변이를 포함하는 APOBEC 데아미나제를 포함한다. 일부 실시형태에서, 염기 편집기 안으로 합체된 APOBEC 데아미나제는 hAPOBEC3G의 R320E 돌연변이, 또는 다른 APOBEC 데아미나제에서 하나 이상의 상응하는 돌연변이를 포함하는 APOBEC 데아미나제를 포함할 수 있다. 일부 실시형태에서, 염기 편집기 안으로 합체된 APOBEC 데아미나제는 hAPOBEC3G의 R313A 돌연변이, 또는 다른 APOBEC 데아미나제에서 하나 이상의 상응하는 돌연변이를 포함하는 APOBEC 데아미나제를 포함할 수 있다. 일부 실시형태에서, 염기 편집기 안으로 합체된 APOBEC 데아미나제는 hAPOBEC3G의 W285A 돌연변이, 또는 다른 APOBEC 데아미나제에서 하나 이상의 상응하는 돌연변이를 포함하는 APOBEC 데아미나제를 포함할 수 있다. 일부 실시형태에서, 염기 편집기 안으로 합체된 APOBEC 데아미나제는 hAPOBEC3G의 W285Y 돌연변이, 또는 다른 APOBEC 데아미나제에서 하나 이상의 상응하는 돌연변이를 포함하는 APOBEC 데아미나제를 포함할 수 있다. 일부 실시형태에서, 염기 편집기 안으로 합체된 APOBEC 데아미나제는 hAPOBEC3G의 R326E 돌연변이, 또는 다른 APOBEC 데아미나제에서 하나 이상의 상응하는 돌연변이를 포함하는 APOBEC 데아미나제를 포함할 수 있다. 일부 실시형태에서, 염기 편집기 안으로 합체된 APOBEC 데아미나제는 hAPOBEC3G의 W285Y 및 R320E 돌연변이, 또는 다른 APOBEC 데아미나제에서 하나 이상의 상응하는 돌연변이를 포함하는 APOBEC 데아미나제를 포함할 수 있다. 일부 실시형태에서, 염기 편집기 안으로 합체된 APOBEC 데아미나제는 hAPOBEC3G의 R320E 및 R326E 돌연변이, 또는 다른 APOBEC 데아미나제에서 하나 이상의 상응하는 돌연변이를 포함하는 APOBEC 데아미나제를 포함할 수 있다. 일부 실시형태에서, 염기 편집기 안으로 합체된 APOBEC 데아미나제는 hAPOBEC3G의 W285Y 및 R326E 돌연변이, 또는 다른 APOBEC 데아미나제에서 하나 이상의 상응하는 돌연변이를 포함하는 APOBEC 데아미나제를 포함할 수 있다. 일부 실시형태에서, 염기 편집기 안으로 합체된 APOBEC 데아미나제는 hAPOBEC3G의 W285Y, R320E 및 R326E 돌연변이, 또는 다른 APOBEC 데아미나제에서 하나 이상의 상응하는 돌연변이를 포함하는 APOBEC 데아미나제를 포함할 수 있다.

[0598] SaBE3, SaKKH-BE3, VQR-BE3, EQR-BE3, VRER-BE3, YE1-BE3, EE-BE3, YE2-BE3 및 YEE-BE3을 포함하지만 이에 제한되지 않는 다수의 변형된 시티딘 데아미나제가 상업적으로 이용가능하며, 이는 Addgene으로부터 이용가능하다(플라스미드 85169, 85170, 85171, 85172, 85173, 85174, 85175, 85176, 85177). 일부 실시형태에서, 염기 편집기 안으로 합체된 데아미나제는 APOBEC1 데아미나제의 전부 또는 일부를 포함한다.

[0599] C에서 T 핵염기 편집 단백질의 세부사항은 국제 PCT 출원 번호 PCT/US2016/058344(WO2017/070632) 및 Komor, A.C., 등, "Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage" Nature 533, 420-424 (2016)에 기술되어 있으며, 그 전체 내용은 본 명세서에 참고로 포함된다.

[0600] ***시티딘 데아미나제***

[0601] 일부 실시형태에서, 발명의 융합 단백질은 하나 이상의 시티딘 데아미나제 도메인을 포함한다. 일부 실시형태에서, 본 명세서에 제공된 시티딘 데아미나제는 시토신 또는 5-메틸시토신을 우라실 또는 티민으로 탈아미노화할 수 있다. 일부 실시형태에서, 본 명세서에 제공된 시티딘 데아미나제는 DNA에서 시토신을 탈아미노화할 수 있다. 시티딘 데아미나제는 임의의 적합한 유기체로부터 유래될 수 있다. 일부 실시형태에서, 시티딘 데아미나제는 본 명세서에 제공된 임의의 돌연변이에 상응하는 하나 이상의 돌연변이를 포함하는 자연적으로-발생하는 시티딘 데아미나제이다. 당업자는, *예를 들어* 서열 정렬 및 상동체 잔기의 결정에 의해 임의의 상동 단백질에서 상응하는 잔기를 식별할 수 있을 것이다. 따라서, 당업자는 본 명세서에 기재된 임의의 돌연변이에 상응하는 임

의의 자연적으로-발생하는 시티딘 데아미나제에서 돌연변이를 생성할 수 있을 것이다. 일부 실시형태에서 시티딘 데아미나제는 원핵생물로부터의 것이다. 일부 실시형태에서, 시티딘 데아미나제는 박테리아로부터의 것이다. 일부 실시형태에서, 시티딘 데아미나제는 포유동물(예를 들어, 인간)로부터의 것이다.

[0602] 일부 실시형태에서, 시티딘 데아미나제는 본 명세서에 제시된 시티딘 데아미나제 아미노산 서열 중 임의의 하나와 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 적어도 99.5% 동일한 아미노산 서열을 포함한다. 본 명세서에 제공된 시티딘 데아미나제는 하나 이상의 돌연변이(예를 들어, 본 명세서에 제공된 임의의 돌연변이)를 포함할 수 있음을 이해해야 한다. 일부 실시형태는 임의의 이전 양태 또는 본 명세서에 기재된 바와 같은 시티딘 데아미나제 핵염기 편집기 폴리펩티드를 인코딩하는 폴리뉴클레오티드 분자를 제공한다. 일부 실시형태에서, 폴리뉴클레오티드는 코돈 최적화된다.

[0603] 개시내용은 특정 퍼센트 동일성을 갖는 임의의 데아미나제 도메인 플러스 본 명세서에 기재된 임의의 돌연변이 또는 이의 조합을 제공한다. 일부 실시형태에서, 시티딘 데아미나제는 참조 서열 또는 본 명세서에 제공된 임의의 시티딘 데아미나제와 비교하여 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 21, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50 이상의 돌연변이를 갖는 아미노산을 포함한다. 일부 실시형태에서, 시티딘 데아미나제는 당업계에 공지되어 있거나 본 명세서에 기재된 아미노산 서열 중 임의의 하나와 비교하여 적어도 5개, 적어도 10개, 적어도 15개, 적어도 20개, 적어도 25개, 적어도 30개, 적어도 35개, 적어도 40개, 적어도 45개, 적어도 50개, 적어도 60개, 적어도 70개, 적어도 80개, 적어도 90개, 적어도 100개, 적어도 110개, 적어도 120개, 적어도 130개, 적어도 140개, 적어도 150개, 적어도 160개, 또는 적어도 170개의 동일한 연속 아미노산 잔기를 갖는 아미노산을 포함한다.

[0604] 발명의 융합 단백질인 제2 단백질은 2개 이상의 핵산 편집 도메인을 포함한다.

[0605] ***가이드 폴리뉴클레오티드***

[0606] 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인은 결합된 가이드 폴리뉴클레오티드(*예를 들어*, gRNA)와 연합할 때 표적 폴리뉴클레오티드 서열에 특이적으로 결합할 수 있고(*즉*, 결합된 가이드 핵산의 염기와 표적 폴리뉴클레오티드 서열의 염기 사이의 상보성 염기 페어링을 통함) 이에 의해 편집하고자 하는 표적 핵산 서열에 염기 편집기를 국지화할 수 있다. 일부 실시형태에서, 표적 폴리뉴클레오티드 서열은 단일-가닥 DNA 또는 이중-가닥 DNA를 포함한다. 일부 실시형태에서, 표적 폴리뉴클레오티드 서열은 RNA를 포함한다. 일부 실시형태에서, 표적 폴리뉴클레오티드 서열은 DNA-RNA 하이브리드를 포함한다.

[0607] CRISPR는 이동성 유전 요소(바이러스, 전이 요소 및 접합성 플라스미드)에 대해 보호를 제공하는 적응형 면역 시스템이다. CRISPR 클러스터는 스페이서, 선행하는 이동성 요소에 상보성인 서열 및 표적 침입 핵산을 함유한다. CRISPR 클러스터는 CRISPR RNA(crRNA)로 전사되고 처리된다. 유형 II CRISPR 시스템에서 pre-crRNA의 올바른 처리에는 트랜스-인코딩된 소형 RNA(tracrRNA), 내인성 리보뉴클레아제 3(rnc) 및 Cas9 단백질이 필요하다. tracrRNA는 pre-crRNA의 리보뉴클레아제 3-보조된 처리를 위한 가이드로 역할을 한다. 이어서, Cas9/crRNA/tracrRNA는 스페이서에 상보성인 선형 또는 원형 dsDNA 표적을 내핵용리적으로 절단한다. crRNA에 상보성이 아닌 표적 가닥은 먼저 내핵용리적으로 절단된 다음, 외핵용리적으로 3'-5' 트리밍된다. 자연적으로, DNA-결합과 절단은 전형적으로 단백질과 두 RNA 모두를 필요로 한다. 그러나, 단일 가이드 RNA("sgRNA" 또는 간단히 "gRNA")는 crRNA 및 tracrRNA 둘 모두의 측면을 단일 RNA 종 안으로 합체하도록 조작될 수 있다. *예를 들어,* Jinek M., Chylinski K., Fonfara I., Hauer M., Doudna J. A., Charpentier E. Science 337:816-821(2012)*을 참조하며*, 그 전체 내용은 본 명세서에 참고로 포함된다. Cas9는 CRISPR 반복 서열에서 짧은 모티프(PAM 또는 프로토스페이서 인접 모티프)를 인식하여 자기 대 비-자기를 구별하는 데 도움을 준다. *예를 들어* "Complete genome sequence of an M1 strain of *Streptococcus pyogenes*." Ferretti, J.J. 등, Natl. Acad. Sci. U.S.A. 98:4658-4663(2001); "CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III." Deltcheva E. 등, Nature 471:602-607(2011); 및 "Programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity." Jinek M.*et al*, Science 337:816-821(2012)*을 참조하며*, 각각의 그 전체 내용은 본 명세서에 참고로 포함된다).

[0608] PAM 서열은 당업계에 공지된 임의의 PAM 서열일 수 있다. 적합한 PAM 서열은 NGG, NGA, NGC, NGN, NGT, NGCG, NGAG, NGAN, NGNG, NGCN, NGCG, NGTN, NNGRRT, NNNRRT, NNGRR(N), TTTV, TYCV, TYCV, TATV, NNNNGATT, NNAGAAW 또는 NAAAAC를 포함하지만 이에 제한되지 않는다. Y는 피리미딘이고; N은 임의의 뉴클레오티드 염기이

고; W는 A 또는 T이다.

[0609] 실시형태에서, 본 명세서에 기재된 가이드 폴리뉴클레오티드는 RNA 또는 DNA일 수 있다. 일 실시형태에서, 가이드 폴리뉴클레오티드는 gRNA이다. RNA/Cas 복합체는 Cas 단백질을 표적 DNA로 "가이드"하는데 도움을 줄 수 있다. Cas9/crRNA/tracrRNA는 스페이서에 상보성인 선형 또는 원형 dsDNA 표적을 내핵용리적으로 절단한다. crRNA에 상보성이 아닌 표적 가닥은 먼저 내핵용리적으로 절단된 다음, 외핵용리적으로 3'-5' 트리밍된다. 자연적으로, DNA-결합과 절단은 전형적으로 단백질과 두 RNA 모두를 필요로 한다. 그러나, 단일 가이드 RNA("sgRNA" 또는 간단히 "gRNA")는 crRNA 및 tracrRNA 둘 모두의 측면을 단일 RNA 종 안으로 합체하도록 조작될 수 있다. *예를 들어,* Jinek M. 등, Science 337:816-821(2012)을 참조하며, 그 전체 내용은 본 명세서에 참고로 포함된다.

[0610] 일부 실시형태에서, 가이드 폴리뉴클레오티드는 적어도 하나의 단일 가이드 RNA("sgRNA" 또는 "gRNA")이다. 일부 실시형태에서, 가이드 폴리뉴클레오티드는 예를 들어 상보성 염기 페어링(*예를 들어*, 이중 가이드 폴리뉴클레오티드, 이중 gRNA)을 통해 서로 상호작용할 수 있는 2개 이상의 개별 폴리뉴클레오티드를 포함한다. 예를 들어, 가이드 폴리뉴클레오티드는 CRISPR RNA(crRNA) 및 트랜스-활성화 CRISPR RNA(tracrRNA)를 포함할 수 있거나 하나 이상의 트랜스-활성화 CRISPR RNA(tracrRNA)를 포함할 수 있다.

[0611] 일부 실시형태에서, 가이드 폴리뉴클레오티드는 적어도 하나의 tracrRNA이다. 일부 실시형태에서, 가이드 폴리뉴클레오티드는 폴리뉴클레오티드-프로그램가능 DNA-결합 도메인(예를 들어, Cas9 또는 Cpf1)을 표적 뉴클레오티드 서열로 가이드하기 위해 PAM 서열을 필요로 하지 않는다.

[0612] 가이드 폴리뉴클레오티드는 천연 또는 비-천연(또는 천연이 아닌) 뉴클레오티드(*예를 들어*, 펩티드 핵산 또는 뉴클레오티드 유사체)를 포함할 수 있다. 일부 경우에, 가이드 핵산 서열의 표적화 영역은 길이가 적어도 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29 또는 30개 뉴클레오티드일 수 있다. 가이드 핵산의 표적화 영역은 길이가 10-30개 사이의 뉴클레오티드, 또는 길이가 15-25개 사이의 뉴클레오티드, 또는 길이가 15-20개 사이의 뉴클레오티드일 수 있다.

[0613] 일부 실시형태에서, 본 명세서에 제공된 염기 편집기는 하나 이상의 가이드 폴리뉴클레오티드(*예를 들어,* 다중 gRNA)를 활용한다. 일부 실시형태에서, 단일 가이드 폴리뉴클레오티드는 본 명세서에 기재된 상이한 염기 편집기에 활용된다. 예를 들어, 단일 가이드 폴리뉴클레오티드는 시티딘 염기 편집기 및 아데노신 염기 편집기에 활용될 수 있다.

[0614] 일부 실시형태에서, 본 명세서에 기재된 방법은 조작된 Cas 단백질을 활용할 수 있다. 가이드 RNA(gRNA)는 Cas-결합에 필요한 스캐폴드 서열과 변형되어 지는 게놈 표적을 정의하는 사용자-정의된 ~20 뉴클레오티드 스페이서로 구성된 짧은 합성 RNA이다. 예시적인 gRNA 스캐폴드 서열은 서열 목록에 서열번호: 224-230, 223, 3000 및 243-245로 제공된다. 따라서, 숙련된 기술자는 gRNA 표적화 서열이 게놈의 잔여부와 비교하여 게놈 표적에 대해 얼마나 특이적인가에 의해 부분적으로 결정되는 Cas 단백질 특이성의 게놈 표적을 변경할 수 있다.

[0615] 다른 실시형태에서, 가이드 폴리뉴클레오티드는 단일 분자에서의 핵산(즉, 단일-분자 가이드 핵산)의 폴리뉴클레오티드 표적화 부분과 핵산의 스캐폴드 부분 둘 모두를 포함할 수 있다. 예를 들어, 단일-분자 가이드 폴리뉴클레오티드는 단일 가이드 RNA(sgRNA 또는 gRNA)일 수 있다. 본 명세서에서 가이드 폴리뉴클레오티드 서열이라는 용어는 염기 편집기와 상호작용하고 이를 표적 폴리뉴클레오티드 서열로 지향하도록 할 수 있는 임의의 단일, 이중 또는 다중-분자 핵산을 고려한다.

[0616] 전형적으로, 가이드 폴리뉴클레오티드(*예를 들어*, crRNA/trRNA 복합체 또는 gRNA)는 표적 폴리뉴클레오티드 서열을 인식하고 이에 결합할 수 있는 서열을 포함하는 "폴리뉴클레오티드-표적화 세그먼트" 및 염기 편집기의 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인 구성요소 내의 가이드 폴리뉴클레오티드를 안정화시키는 "단백질-결합 세그먼트"를 포함한다. 일부 실시형태에서, 가이드 폴리뉴클레오티드의 폴리뉴클레오티드 표적화 세그먼트는 DNA 폴리뉴클레오티드를 인식하고 이에 결합하여, DNA에서 염기의 편집을 촉진한다. 다른 경우에, 가이드 폴리뉴클레오티드의 폴리뉴클레오티드 표적화 세그먼트는 RNA 폴리뉴클레오티드를 인식하고 이에 결합하여, RNA에서 염기의 편집을 촉진한다. 본 명세서에서 "세그먼트"는 분자의 섹션 또는 영역, *예를 들어* 가이드 폴리뉴클레오티드에서 뉴클레오티드의 연속하는 스트레치를 지칭한다. 세그먼트는 또한 세그먼트가 하나 초과 분자의 영역을 포함할 수 있도록 복합체의 영역/섹션을 지칭할 수 있다. 예를 들어, 가이드 폴리뉴클레오티드가 다중 핵산 분자를 포함하는 경우, 단백질-결합 세그먼트는 예로서 상보성의 영역을 따라 혼성화되는 다중 개별 분자의 전부 또는 일부를 포함할 수 있다. 일부 실시형태에서, 2개의 별도의 분자를 포함하는 DNA-표적화 RNA의 단백질-결합 세그먼트는 (i) 길이가 100개 염기쌍인 제1 RNA 분자의 염기쌍 40-75 및 (ii) 길이가 50

개 염기쌍인 제2 RNA 분자의 염기쌍 10-25를 포함할 수 있다. "세그먼트"의 정의는 특별한 맥락에서 달리 구체적으로 정의되지 않는 한 총 염기쌍의 특정 수에 제한되지 않고, 주어진 RNA 분자로부터의 염기쌍의 임의의 특정 수에 제한되지 않고, 복합체 내의 별도 분자의 특정 수에 제한되지 않고, 임의의 총 길이의 것인 RNA 분자의 영역을 포함할 수 있고 다른 분자에 상보성을 갖는 영역을 포함할 수 있다.

[0617] 가이드 폴리뉴클레오티드는 화학적으로 합성되거나, 효소적으로 합성되거나, 이의 조합일 수 있다. 예를 들어, gRNA는 표준 포스포르아미다이트-기반 고체-상 합성 방법을 사용하여 합성될 수 있다. 대안적으로, gRNA는 gRNA를 인코딩하는 DNA를 파지 RNA 폴리머라제에 의해 인식되는 프로모터 제어 서열에 작동가능하게 연결함에 의해 *시험관내에서* 합성될 수 있다. 적합한 파지 프로모터 서열의 예는 T7, T3, SP6 프로모터 서열 또는 이의 변이를 포함한다. gRNA가 2개의 별도의 분자(*예를 들어*, crRNA 및 tracrRNA)를 포함하는 실시형태에서, crRNA는 화학적으로 합성될 수 있고 tracrRNA는 효소적으로 합성될 수 있다. gRNA 분자는 *시험관내에서* 전사될 수 있다.

[0618] 가이드 폴리뉴클레오티드는, 예를 들어, gRNA를 인코딩하는 DNA, 예를 들어 gRNA를 인코딩하는 서열을 포함하는 DNA 벡터에 의해 발현될 수 있다. gRNA는 단독으로 또는 인코딩된 염기 편집기와 함께 인코딩될 수 있다. 이러한 DNA 서열은 발현 시스템, *예를 들어* 세포 내로 함께 또는 별도로 도입될 수 있다. 예를 들어, 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인 및 gRNA를 인코딩하는 DNA 서열은 세포 내로 도입될 수 있으며, 각 DNA 서열은 별도의 분자(*예를 들어*, 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인 코딩 서열을 함유하는 하나의 벡터 및 gRNA 코딩 서열을 함유하는 제2 벡터)의 일부일 수 있거나 또는 둘 모두 동일한 분자(*예를 들어*, 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인과 gRNA 둘 모두에 대한 코딩 (및 조절) 서열을 함유하는 하나의 벡터)의 일부일 수 있다. RNA는 합성 DNA 분자, *예를 들어*, gBlocks® 유전자 단편으로부터 전사될 수 있다.

[0619] gRNA 또는 가이드 폴리뉴클레오티드는 3개의 영역: 염색체 서열에서의 표적 부위에 상보성일 수 있는 5' 말단에서 제1 영역, 스템 루프 구조를 형성할 수 있는 제2 내부 영역, 및 단일-가닥이 될 수 있는 제3 3' 영역을 포함할 수 있다. 각 gRNA의 제1 영역은 또한 각 gRNA가 융합 단백질을 특정 표적 부위로 가이드하도록 다를 수 있다. 더욱이, 각 gRNA의 제2 및 제3 영역은 모든 gRNA에서 동일할 수 있다.

[0620] gRNA의 제1 영역 또는 가이드 폴리뉴클레오티드는 gRNA의 제1 영역이 표적 부위와 염기쌍을 이룰 수 있도록 염색체 서열에서의 표적 부위에서 서열에 상보성일 수 있다. 일부 경우에, gRNA의 제1 영역은 10개 또는 약 10개의 뉴클레오티드 내지 25개의 뉴클레오티드(즉, 10개의 뉴클레오티드 내지 뉴클레오티드; 또는 약 10개의 뉴클레오티드 내지 약 25개의 뉴클레오티드; 또는 10개의 뉴클레오티드 내지 약 25개의 뉴클레오티드; 또는 약 10개 뉴클레오티드 내지 25개 뉴클레오티드) 이상을 포함할 수 있다. 예를 들어, 염색체 서열에서 표적 부위와 gRNA의 제1 영역 사이의 염기 페어링의 영역은 길이가 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 22, 23, 24, 25 이상의 뉴클레오티드일 수 있거나, 또는 약 그 수일 수 있다. 때때로, gRNA의 제1 영역은 길이가 19, 20 또는 21개 뉴클레오티드일 수 있거나, 또는 약 그 수일 수 있다.

[0621] gRNA 또는 가이드 폴리뉴클레오티드는 또한 2차 구조를 형성하는 제2 영역을 포함할 수 있다. 예를 들어, gRNA에 의해 형성된 2차 구조는 줄기 (또는 헤어핀) 및 루프를 포함할 수 있다. 루프와 줄기의 길이는 다양할 수 있다. 예를 들어, 루프는 길이가 약 3 내지 10개 뉴클레오티드의 범위일 수 있고, 줄기는 길이가 약 6 내지 20개 염기쌍의 범위일 수 있다. 줄기는 1개 내지 10개 또는 약 10개의 뉴클레오티드의 하나 이상의 돌출부를 포함할 수 있다. 제2 영역의 전반적인 길이는 약 16 내지 60개의 뉴클레오티드 길이의 범위일 수 있다. 예를 들어, 루프는 길이가 4개 또는 약 4개의 뉴클레오티드일 수 있고 줄기는 12개 또는 약 12개의 염기쌍일 수 있다.

[0622] gRNA 또는 가이드 폴리뉴클레오티드는 또한 본질적으로 단일-가닥일 수 있는 3' 말단에서 제3 영역을 포함할 수 있다. 예를 들어, 제3 영역은 때로는 관심있는 세포에서 임의의 염색체 서열에 대해 상보성이 아니고 때로는 gRNA의 나머지에 대해 상보성이 아니다. 더욱이, 제3 영역의 길이는 다양할 수 있다. 제3 영역은 길이가 4개 초과 또는 약 4개 초과 뉴클레오티드일 수 있다. 예를 들어, 제3 영역은 길이가 5 내지 60개 또는 약 5 내지 60개 뉴클레오티드의 범위일 수 있다.

[0623] gRNA 또는 가이드 폴리뉴클레오티드는 유전자 표적의 임의의 엑손 또는 인트론을 표적화할 수 있다. 일부 경우에, 가이드는 유전자의 엑손 1 또는 2를 표적화할 수 있고, 다른 경우에; 가이드는 유전자의 엑손 3 또는 4를 표적화할 수 있다. 일부 실시형태에서, 조성물은 모두 동일한 엑손을 표적화하는 다중 gRNA 또는 상이한 엑손을 표적화하는 다중 gRNA를 포함한다. 유전자의 엑손 및/또는 인트론이 표적화될 수 있다.

[0624] gRNA 또는 가이드 폴리뉴클레오티드는 약 20개의 뉴클레오티드 또는 약 20개 미만의 뉴클레오티드(예를 들어,

적어도 약 5, 10, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30개의 뉴클레오티드) 또는 어떤 경우는 약 1-100개 사이의 뉴클레오티드(예를 들어, 5, 10, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 40, 50, 60, 70, 80, 90, 100)의 핵산 서열을 표적화될 수 있다. 표적 핵산 서열은 PAM의 제1 뉴클레오티드의 5' 바로 옆에 20개의 염기 또는 약 20개의 염기일 수 있다. gRNA는 핵산 서열을 표적화할 수 있다. 표적 핵산은 적어도 또는 적어도 약 1-10, 1-20, 1-30, 1-40, 1-50, 1-60, 1-70, 1-80, 1-90 또는 1-100개의 뉴클레오티드일 수 있다.

[0625] 가이드 폴리뉴클레오티드, *예를 들어* gRNA 및 표적화 서열을 선택, 설계 및 검증하는 방법은 본 명세서에 기재되어 있으며 당업자에게 공지되어 있다. 예를 들어, 핵염기 편집기 시스템(*예를 들어*, AID 도메인)에서 데아미나제 도메인의 잠재적 기질 난잡성의 영향을 최소화하기 위해, 의도하지 않게 탈아미노화에 대해 표적화될 수 있는 잔기(*예를 들어*, 잠재적으로 표적 핵산 좌위 내의 단일 가닥 DNA 상에 잔류할 수 있는 표적-외 C 잔기)의 수가 최소화될 수 있다. 부가하여, *예를 들어*, 게놈 전반에 걸쳐 총 표적-외 활성을 최소화하기 위해, 표적 핵산 서열에 상응하는 gRNA를 최적화하기 위해 소프트웨어 도구가 사용될 수 있다. 예를 들어 *S. 피오게네스* Cas9를 사용하여 각각의 가능한 표적화 도메인 선정에 대해 모든 표적-외 서열(선행하여 선택된 PAM, *예를 들어*, NAG 또는 NGG)은 일치하지 않는 염기-쌍의 최대 특정 수(*예를 들어*, 1, 2, 3, 4, 5, 6, 7, 8, 9 또는 10)를 함유하는 게놈 전반에 걸쳐 식별될 수 있다. 표적 부위에 상보성인 gRNA의 제1 영역이 식별될 수 있고, 모든 제1 영역(*예를 들어*, crRNA)이 그의 전체 예측된 표적-외 점수에 따라 순위가 매겨길 수 있으며; 최상-순위의 표적화 도메인은 가장 큰 표적-상 활성과 가장 적은 표적-외 활성을 가질 가능성이 있는 도메인을 나타낸다. gRNA를 표적화하는 후보는 당업계에 공지된 방법 및/또는 본 명세서에 제시된 바와 같은 방법을 사용하여 기능적으로 평가될 수 있다.

[0626] 비-제한적인 예로서, Cas9와 함께 사용하기 위한 gRNA의 crRNA에서 표적 DNA 혼성화 서열은 DNA 서열 검색 알고리즘을 사용하여 식별될 수 있다. gRNA 설계는 Bae S., Park J., & Kim J.-S. Cas-OFFinder: A fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. Bioinformatics 30, 1473-1475 (2014)에 기재된 바와 같은 공개 도구 cas-offinder에 기반한 맞춤형 gRNA 설계 소프트웨어를 사용하여 수행된다. 이 소프트웨어는 그 게놈-전체의 표적-외 성향을 계산한 후 가이드를 채점한다. 전형적으로 길이가 17 내지 24에 이르는 가이드에 대해 완벽한 일치부터 7 불일치에 이르는 일치가 고려된다. 표적-외 부위가 계산적으로-결정되면, 각 가이드에 대해 집계 점수가 계산되고 웹-인터페이스를 사용하여 표 형식의 출력으로 요약된다. PAM 서열에 인접한 잠재적인 표적 부위를 식별하는 것 외에도, 소프트웨어는 선택된 표적 부위와 1, 2, 3 또는 3개 초과의 뉴클레오티드가 다른 모든 PAM 인접 서열도 식별한다. 표적 핵산 서열, *예를 들어* 표적 유전자에 대한 게놈 DNA 서열이 획득될 수 있고 반복 요소는 공개적으로 이용가능한 도구, 예를 들어 RepeatMasker 프로그램을 사용하여 스크리닝될 수 있다. RepeatMasker는 반복되는 요소와 낮은 복잡성의 영역에 대해 입력 DNA 서열을 검색한다. 출력은 주어진 쿼리 서열에 존재하는 반복의 자세한 주석이다.

[0627] 식별에 이어서, gRNA, *예를 들어*, crRNA의 제1 영역은 표적 부위까지의 그 거리, 직교성 및 관련 PAM 서열과 밀접한 일치에 대한 5' 뉴클레오티드(예를 들어, 관련 PAM, *예를 들어*, S. 피오게네스의 경우 NGG PAM, S. 아우레우스의 경우 NNGRRT 또는 NNGRRV PAM을 함유하는 인간 게놈에서 밀접한 일치의 식별에 기반한 5' G)의 존재에 기반하여 역가로 순위가 지정된다. 본 명세서에 사용된 바와 같이, 직교성은 표적 서열에 대한 최소 수의 불일치를 함유하는 인간 게놈에서 서열의 수를 지칭한다. "높은 수준의 직교성" 또는 "우수한 직교성"은 예를 들어, 의도된 표적 외에 인간 게놈에서 동등한 서열이 없거나 표적 서열에서 하나 또는 두 개의 불일치를 함유하는 임의의 서열이 없는 20-량체 표적 도메인을 지칭할 수 있다. 표적-외 DNA 절단을 최소화하기 위해 우수한 직교성을 갖는 표적화 도메인이 선택될 수 있다.

[0628] 그런 다음 gRNA는 RNA 분자 또는 비-RNA 핵산 분자, *예를 들어* DNA 분자로서 세포 또는 배아 안으로 도입될 수 있다. 일 실시형태에서, gRNA를 인코딩하는 DNA는 관심있는 세포 또는 배아에서 gRNA의 발현을 위한 프로모터 제어 서열에 작동가능하게 연결된다. RNA 코딩 서열은 RNA 폴리머라제 III(Pol III)에 의해 인식되는 프로모터 서열에 작동가능하게 연결될 수 있다. gRNA를 발현하는데 사용될 수 있는 플라스미드 벡터는 px330 벡터 및 px333 벡터를 포함하지만 이에 제한되지 않는다. 일부 경우에, 플라스미드 벡터(*예를 들어*, px333 벡터)는 적어도 2개의 gRNA-인코딩 DNA 서열을 포함할 수 있다. 더욱이, 벡터는 추가적인 발현 제어 서열(*예를 들어*, 인핸서 서열, Kozak 서열, 폴리아데닐화 서열, 전사 종결 서열 등), 선택가능한 마커 서열(*예를 들어*, GFP 또는 항생제 내성 유전자 예컨대 퓨로마이신), 복제의 기원, 등을 포함할 수 있다. gRNA를 인코딩하는 DNA 분자는 또한 선형일 수 있다. gRNA 또는 가이드 폴리뉴클레오티드를 인코딩하는 DNA 분자는 또한 원형일 수 있다.

[0629] 일부 실시형태에서, 염기-편집 활성을 검출하고 후보 가이드 폴리뉴클레오티드를 테스트하기 위해 리포터 시스

템이 사용된다. 일부 실시형태에서, 리포터 시스템은 염기 편집 활성이 리포터 유전자의 발현을 유도하는 리포터 유전자 기반 검정을 포함한다. 예를 들어, 리포터 시스템은 비활성화된 시작 코돈, *예를 들어* 주형 가닥 상에 3'-TAC-5'에서 3'-CAC-5'로의 돌연변이를 포함하는 리포터 유전자를 포함할 수 있다. 표적 C의 성공적인 탈아미노화 시, 상응하는 mRNA가 5'-GUG-3' 대신 5'-AUG-3'으로 전사되어 리포터 유전자의 번역이 가능해질 것이다. 적합한 리포터 유전자는 당업자에게 명백할 것이다. 리포터 유전자의 비-제한적인 예는 녹색 형광 단백질(GFP), 적색 형광 단백질(RFP), 루시퍼라제, 분비된 알칼리성 포스파타제(SEAP)를 인코딩하는 유전자, 또는 그 발현이 검출가능하고 당업자에게 명백한 임의의 기타 유전자를 포함한다. 리포터 시스템은, *예를 들어* 각각의 데아미나제가 표적화할 표적 DNA 서열과 관련하여 어느 잔기(들)를 결정하기 위해 많은 다양한 gRNA를 테스트하는데 사용될 수 있다. 비-주형 가닥을 표적화하는 sgRNA는 또한 특정 염기 편집 단백질, *예를 들어*, Cas9 데아미나제 융합 단백질의 표적-외 효과를 평가하기 위해 테스트될 수 있다. 일부 실시형태에서, 이러한 gRNA는 돌연변이된 시작 코돈이 gRNA와 염기-페어링되지 않도록 설계될 수 있다. 가이드 폴리뉴클레오티드는 표준 리보뉴클레오티드, 변형된 리보뉴클레오티드(*예를 들어*, 슈도우리딘), 리보뉴클레오티드 이성질체, 및/또는 리보뉴클레오티드 유사체를 포함할 수 있다. 일부 실시형태에서, 가이드 폴리뉴클레오티드는 적어도 하나의 검출가능한 표지를 포함할 수 있다. 검출가능한 표지는 형광단(*예를 들어*, FAM, TMR, Cy3, Cy5, 텍사스 레드, 오레곤 그린, 알렉사 플루오르, 할로 태그 또는 적합한 형광 염료), 검출 태그(*예를 들어*, 비오틴, 디곡시게닌 등), 양자점 또는 금 입자일 수 있다.

[0630] 일부 실시형태에서, 염기 편집기 시스템은 다중 가이드 폴리뉴클레오티드, *예를 들어* gRNA를 포함할 수 있다. 예를 들어, gRNA는 염기 편집기 시스템에 포함된 하나 이상의 표적 좌위(*예를 들어*, 적어도 1 gRNA, 적어도 2 gRNA, 적어도 5 gRNA, 적어도 10 gRNA, 적어도 20 gRNA, 적어도 30 gRNA, 적어도 50 gRNA)를 표적화할 수 있다. 다중 gRNA 서열은 일렬로 배열될 수 있고 바람직하게는 직접 반복부에 의해 분리된다.

[0631] 가이드 폴리뉴클레오티드는 새롭거나 증강된 특징을 가진 핵산을 제공하기 위해 하나 이상의 변형을 포함할 수 있다. 가이드 폴리뉴클레오티드는 핵산 친화성 태그를 포함할 수 있다. 가이드 폴리뉴클레오티드는 합성 뉴클레오티드, 합성 뉴클레오티드 유사체, 뉴클레오티드 유도체 및/또는 변형된 뉴클레오티드를 포함할 수 있다.

[0632] 일부 경우에, gRNA 또는 가이드 폴리뉴클레오티드는 변형을 포함할 수 있다. 변형은 gRNA 또는 가이드 폴리뉴클레오티드의 임의 위치에서 이루어질 수 있다. 단일 gRNA 또는 가이드 폴리뉴클레오티드에 대해 하나 초과의 변형이 이루어질 수 있다. gRNA 또는 가이드 폴리뉴클레오티드는 변형 후 품질 관리를 받을 수 있다. 일부 경우에, 품질 관리는 PAGE, HPLC, MS 또는 이의 임의의 조합을 포함할 수 있다.

[0633] gRNA 또는 가이드 폴리뉴클레오티드의 변형은 치환, 삽입, 결실, 화학적 변형, 물리적 변형, 안정화, 정제 또는 이의 임의의 조합일 수 있다.

[0634] gRNA 또는 가이드 폴리뉴클레오티드는 또한 5' 아데닐레이트, 5' 구아노신-트리포스페이트 캡, 5' N7-메틸구아노신-트리포스페이트 캡, 5' 트리포스페이트 캡, 3' 포스페이트, 3' 티오포스페이트, 5' 포스페이트, 5' 티오포스페이트, Cis-Syn 티미딘 이량체, 삼량체, C12 스페이서, C3 스페이서, C6 스페이서, dSpacer, PC 스페이서, rSpacer, 스페이서 18, 스페이서 9, 3'-3' 변형, 5'-5' 변형, 비염기성, 아크리딘, 아조벤젠, 비오틴, 비오틴 BB, 비오틴 TEG, 콜레스테롤 TEG, 데스티오비오틴 TEG, DNP TEG, DNP-X, DOTA, dT-비오틴, 이중 비오틴, PC 비오틴, psoralen C2, psoralen C6, TINA, 3' DABCYL, 블랙 홀 소광제 1, 블랙 홀 소광제 2, DABCYL SE, dT-DABCYL, IRDye QC-1, QSY-21, QSY-35, QSY-7, QSY-9, 카르복실 링커, 티올 링커, 2'-데옥시리보뉴클레오시드 유사체 퓨린, 2' -데옥시리보뉴클레오시드 유사체 피리미딘, 리보뉴클레오시드 유사체, 2'-O-메틸 리보뉴클레오시드 유사체, 당 변형 유사체, 워블/범용 염기, 형광 염료 표지, 2'-플루오로 RNA, 2'-O-메틸 RNA, 메틸포스포네이트, 포스포디에스테르 DNA, 포스포디에스테르 RNA, 포스포티오에이트 DNA, 포스포로티오에이트 RNA, UNA, 슈도우리딘-5'-트리포스페이트, 5'-메틸시티딘-5'-트리포스페이트, 또는 이의 임의의 조합에 의해 변형될 수 있다.

[0635] 일부 경우에, 변형은 영구적이다. 다른 경우에, 변형은 일시적이다. 일부 경우에, gRNA 또는 가이드 폴리뉴클레오티드에 대한 다중 변형이 이루어진다. gRNA 또는 가이드 폴리뉴클레오티드 변형은 그 형태, 극성, 소수성, 화학적 반응성, 염기-페어링 상호작용 또는 이의 임의의 조합과 같은 뉴클레오티드의 물리화학적 특성을 변경할 수 있다.

[0636] 가이드 폴리뉴클레오티드는 프로모터 및 가이드 RNA를 인코딩하는 서열을 포함하는 단리된 gRNA 또는 플라스미드 DNA로 세포를 형질감염시킴에 의해 세포 내로 전달될 수 있다. gRNA 또는 가이드 폴리뉴클레오티드는 바이러스-매개된 유전자 전달을 사용하는 것과 같은 다른 방식으로 세포 내로 전달될 수도 있다. gRNA 또는 가이드 폴

리뉴클레오티드는 단리될 수 있다. 예를 들어, gRNA는 단리된 RNA의 형태로 세포 또는 유기체 안으로 형질감염될 수 있다. gRNA는 당업계에 공지된 임의의 *시험관내* 전사 시스템을 사용하여 *시험관내* 전사에 의해 제조될 수 있다. gRNA는 gRNA에 대한 인코딩 서열을 포함하는 플라스미드의 형태 이외의 단리된 RNA의 형태로 세포에 전달될 수 있다.

[0637] 변형은 또한 포스포로티오에이트 대체물일 수 있다. 일부 경우에, 천연 포스포디에스테르 결합은 세포 뉴클레아제에 의해 빠른 분해를 받기 쉬울 수 있고; 포스포로티오에이트(PS) 결합 대체물을 사용한 뉴클레오티드 간 연결의 변형은 세포 분해에 의한 가수분해에 대해 더 안정적일 수 있다. 변형은 gRNA 또는 가이드 폴리뉴클레오티드에서의 안정성을 증가시킬 수 있다. 변형은 또한 생물학적 활성을 증강시킬 수 있다. 일부 경우에, 포스포로티오에이트 증강된 RNA gRNA는 RNase A, RNase T1, 송아지 혈청 뉴클레아제 또는 이의 임의의 조합을 억제할 수 있다. 이들 특성으로 인해 *생체내* 또는 *시험관내에서* 뉴클레아제에 노출될 가능성이 높은 적용 분야에서 PS-RNA gRNA를 사용할 수 있다. 예를 들어, 포스포로티오에이트(PS) 결합은 엑소뉴클레아제 분해를 억제할 수 있는 gRNA의 5'- 또는 ''-말단에서 마지막 3-5개의 뉴클레오티드 사이에 도입될 수 있다. 일부 경우에, 엔도뉴클레아제에 의한 공격을 감소시키기 위해 전체 gRNA 전반에 걸쳐서 포스포로티오에이트 결합이 추가될 수 있다.

[0638] 일부 실시형태에서, 가이드 RNA는 스플라이스 부위(즉, 스플라이스 수용체(SA) 또는 스플라이스 공여체(SD)를 파괴하도록 설계된다. 일부 실시형태에서, 가이드 RNA는 염기 편집이 조기 STOP 코돈을 초래하도록 설계된다.

[0639] ***프로토스페이서 인접 모티프***

[0640] 용어 "프로토스페이서 인접 모티프(PAM)" 또는 PAM-유사 모티프는 CRISPR 박테리아 적응성 면역 시스템에서 Cas9 뉴클레아제에 의해 표적화된 DNA 서열에 바로 이어지는 2-6개의 염기쌍 DNA 서열을 지칭한다. 일부 실시형태에서, PAM은 5' PAM(즉, 프로토스페이서의 5' 말단의 상류에 위치됨)일 수 있다. 다른 실시형태에서, PAM은 3' PAM(즉, 프로토스페이서의 5' 말단의 하류에 위치됨)일 수 있다. PAM 서열은 표적 결합에 필수적이지만 정확한 서열은 Cas 단백질의 유형에 따라 다르다. PAM 서열은 당업계에 공지된 임의의 PAM 서열일 수 있다. 적합한 PAM 서열은 NGG, NGA, NGC, NGN, NGT, NGTT, NGCG, NGAG, NGAN, NGNG, NGCN, NGCG, NGTN, NNGRRT, NNNRRT, NNGRR(N), TTTV, TYCV, TYCV, TATV, NNNNGATT, NNAGAAW, 또는 NAAAAC를 포함하지만 이에 제한되지 않는다. Y는 피리미딘이고; N은 임의의 뉴클레오티드 염기이고; W는 A 또는 T이다.

[0641] 본 명세서에 제공된 염기 편집기는 규범적 또는 비-규범적 프로토스페이서 인접 모티프(PAM) 서열을 함유하는 뉴클레오티드 서열에 결합할 수 있는 CRISPR 단백질-유래된 도메인을 포함할 수 있다. PAM 부위는 표적 폴리뉴클레오티드 서열에 근접한 뉴클레오티드 서열이다. 개시내용의 일부 양태는 상이한 PAM 특이성을 갖는 CRISPR 단백질의 전부 또는 일부를 포함하는 염기 편집기를 제공한다.

[0642] 예를 들어, 전형적으로 *S. 피오게네스*로부터의 Cas9(spCas9)와 같은 Cas9 단백질은 특정 핵산 영역에 결합하기 위해 규범적 NGG PAM 서열이 필요하며, 여기서 "NGG"에서의 "N"은 아데닌(A), 티민(T), 구아닌(G) 또는 시토신(C)이고, G는 구아닌이다. PAM은 CRISPR 단백질-특이적일 수 있으며 상이한 CRISPR 단백질-유래된 도메인을 포함하는 상이한 염기 편집기 간에 다를 수 있다. PAM은 표적 서열의 5' 또는 3'일 수 있다. PAM은 표적 서열의 상류 또는 하류일 수 있다. PAM은 길이가 1, 2, 3, 4, 5, 6, 7, 8, 9, 10개 이상의 뉴클레오티드일 수 있다. 종종 PAM을 길이가 2-6개 뉴클레오티드 사이이다.

[0643] 일부 실시형태에서, PAM은 "NRN"에서 "N"이 아데닌(A), 티민(T), 구아닌(G), 또는 시토신(C)이고 R이 아데닌(A) 또는 구아닌(G)인 "NRN" PAM이거나; 또는 PAM은 NYN에서 "N"이 아데닌(A), 티민(T), 구아닌(G), 또는 시토신(C)이고, Y가 시티딘(C) 또는 티민(T)인 "NYN" PAM이며, 예를 들어, 그 전체 내용이 본 명세서에 참조로 포함된, R.T. Walton 등, 2020, *Science*, 10.1126/science.aba8853 (2020)에 기재된 바와 같다.

[0644] 여러 PAM 변이체가 아래 **표 6**에 기술되어 있다.

[0645] **표 6. Cas9 단백질 및 상응하는 PAM 서열**

변이체	PAM
spCas9	NGG
spCas9-VRQR	NGA
spCas9-VRER	NGCG
xCas9 (sp)	NGN
saCas9	NNGRRT
saCas9-KKH	NNNRRT
spCas9-MQKSER	NGCG
spCas9-MQKSER	NGCN
spCas9-LRKIQK	NGTN
spCas9-LRVSQK	NGTN
spCas9-LRVSQL	NGTN
spCas9-MQKFRAER	NGC
Cpf1	5' (TTTV)
SpyMac	5'-NAA-3'

[0646]

[0647] 일부 실시형태에서, PAM은 NGC이다. 일부 실시형태에서, NGC PAM은 Cas9 변이체에 의해 인식된다. 일부 실시형태에서, NGC PAM 변이체는 D1135M, S1136Q, G1218K, E1219F, A1322R, D1332A, R1335E 및 T1337R(총칭하여 "MQKFRAER"로 지칭됨)로부터 선택된 하나 이상의 아미노산 치환을 포함한다.

[0648] 일부 실시형태에서, PAM은 NGT이다. 일부 실시형태에서 NGT PAM은 Cas9 변이체에 의해 인식된다. 일부 실시형태에서, NGT PAM 변이체는 하나 이상의 잔기 1335, 1337, 1135, 1136, 1218 및/또는 1219에서 표적화된 돌연변이를 통해 생성된다. 일부 실시형태에서, NGT PAM 변이체는 하나 이상의 잔기 1219, 1335, 1337, 1218에서 표적화된 돌연변이를 통해 창출된다. 일부 실시형태에서, NGT PAM 변이체는 하나 이상의 잔기 1135, 1136, 1218, 1219 및 1335에서 표적화된 돌연변이를 통해 창출된다. 일부 실시형태에서, NGT PAM 변이체는 아래 **표 7A** 및 **7B**에 제공된 표적화된 돌연변이의 세트로부터 선택된다.

[0649] **표 7A: 잔기 1219, 1335, 1337, 1218에서 NGT PAM 변이체 돌연변이**

변이체	E1219V	R1335Q	T1337	G1218
1	F	V	T	
2	F	V	R	
3	F	V	Q	
4	F	V	L	
5	F	V	T	R
6	F	V	R	R
7	F	V	Q	R
8	F	V	L	R
9	L	L	T	
10	L	L	R	
11	L	L	Q	
12	L	L	L	
13	F	I	T	
14	F	I	R	
15	F	I	Q	
16	F	I	L	
17	F	G	C	
18	H	L	N	
19	F	G	C	A
20	H	L	N	V
21	L	A	W	
22	L	A	F	
23	L	A	Y	
24	I	A	W	
25	I	A	F	
26	I	A	Y	

[0650]

[0651] **표 7B: 잔기 1135, 1136, 1218, 1219, 및 1335에서 NGT PAM 변이체 돌연변이**

변이체	D1135L	S1136R	G1218S	E1219V	R1335Q
27	G				
28	V				
29	I				
30		A			
31		W			
32		H			
33		K			
34			K		
35			R		
36			Q		
37			T		
38			N		
39				I	
40				A	
41				N	
42				Q	
43				G	
44				L	

[0652]

45				S	
46				T	
47					L
48					I
49					V
50					N
51					S
52					T
53					F
54					Y
55	N1286Q	I1331F			

[0653]

[0654] 일부 실시형태에서, NGT PAM 변이체는 **표 7A** 및 **표 7B**에서의 변이체 5, 7, 28, 31 또는 36으로부터 선택된다. 일부 실시형태에서, 변이체는 개선된 NGT PAM 인식을 갖는다.

[0655] 일부 실시형태에서, NGT PAM 변이체는 잔기 1219, 1335, 1337 및/또는 1218에서 돌연변이를 갖는다. 일부 실시형태에서, 아래 **표 8**에 제공된 변이체로부터 개선된 인식을 위한 돌연변이를 갖는 NGT PAM 변이체가 선택된다.

[0656] **표 8: 잔기 1219, 1335, 1337 및 1218에서 NGT PAM 변이체 돌연변이**

변이체	E1219V	R1335Q	T1337	G1218
1	F	V	T	
2	F	V	R	
3	F	V	Q	
4	F	V	L	
5	F	V	T	R
6	F	V	R	R
7	F	V	Q	R
8	F	V	L	R

[0657]

[0658] 일부 실시형태에서, NGT PAM은 아래 **표 9**에 제공된 변이체로부터 선택된다.

[0659] **표 9. NGT PAM 변이체**

	NGTN 변이체	D1135	S1136	G1218	E1219	A1322R	R1335	T1337
변이체 1	LRKIQK	L	R	K	I	-	Q	K
변이체 2	LRSVQK	L	R	S	V	-	Q	K
변이체 3	LRSVQL	L	R	S	V	-	Q	L
변이체 4	LRKIRQK	L	R	K	I	R	Q	K
변이체 5	LRSVRQK	L	R	S	V	R	Q	K
변이체 6	LRSVRQL	L	R	S	V	R	Q	L

[0660]

[0661] 일부 실시형태에서 NGTN 변이체는 변이체 1이다. 일부 실시형태에서, NGTN 변이체는 변이체 2이다. 일부 실시형태에서, NGTN 변이체는 변이체 3이다. 일부 실시형태에서, NGTN 변이체는 변이체 4이다. 일부 실시형태에서,

NGTN 변이체는 변이체 5이다. 일부 실시형태에서, NGTN 변이체는 변이체 6이다.

[0662] 일부 실시형태에서, Cas9 도메인은 *스트렙토코커스 피오게네스*(SpCas9)로부터의 Cas9 도메인이다. 일부 실시형태에서, SpCas9 도메인은 뉴클레아제 활성 SpCas9, 뉴클레아제 불활성 SpCas9(SpCas9d), 또는 SpCas9 닉카제(SpCas9n)이다. 일부 실시형태에서, SpCas9는 D9X 돌연변이, 또는 본 명세서에 제공된 임의의 아미노산 서열에서의 상응하는 돌연변이를 포함하며, 여기서 X는 D를 제외한 임의의 아미노산이다. 일부 실시형태에서, SpCas9는 D9A 돌연변이 또는 본 명세서에 제공된 임의의 아미노산 서열에서의 상응하는 돌연변이를 포함한다. 일부 실시형태에서, SpCas9 도메인, SpCas9d 도메인, 또는 SpCas9n 도메인은 비-규범적 PAM을 갖는 핵산 서열에 결합할 수 있다. 일부 실시형태에서, SpCas9 도메인, SpCas9d 도메인 또는 SpCas9n 도메인은 NGG, NGA 또는 NGCG PAM 서열을 갖는 핵산 서열에 결합할 수 있다.

[0663] 일부 실시형태에서, SpCas9 도메인은 D1135X, R1335X 및 T1337X 돌연변이 중 하나 이상, 또는 본 명세서에 제공된 임의의 아미노산 서열에서의 상응하는 돌연변이를 포함하며, 여기서 X는 임의의 아미노산이다. 일부 실시형태에서, SpCas9 도메인은 D1135E, R1335Q 및 T1337R 돌연변이 중 하나 이상, 또는 본 명세서에 제공된 임의의 아미노산 서열에서의 상응하는 돌연변이를 포함한다. 일부 실시형태에서, SpCas9 도메인은 D1135E, R1335Q 및 T1337R 돌연변이, 또는 본 명세서에 제공된 임의의 아미노산 서열에서의 상응하는 돌연변이를 포함한다. 일부 실시형태에서, SpCas9 도메인은 D1135X, R1335X 및 T1337X 돌연변이 중 하나 이상, 또는 본 명세서에 제공된 임의의 아미노산 서열에서의 상응하는 돌연변이를 포함하며, 여기서 X는 임의의 아미노산이다. 일부 실시형태에서, SpCas9 도메인은 D1135V, R1335Q 및 T1337R 돌연변이 중 하나 이상, 또는 본 명세서에 제공된 임의의 아미노산 서열에서의 상응하는 돌연변이를 포함한다. 일부 실시형태에서, SpCas9 도메인은 D1135V, R1335Q 및 T1337R 돌연변이, 또는 본 명세서에 제공된 임의의 아미노산 서열에서의 상응하는 돌연변이를 포함한다. 일부 실시형태에서, SpCas9 도메인은 D1135X, G1218X, R1335X 및 T1337X 돌연변이 중 하나 이상, 또는 본 명세서에 제공된 임의의 아미노산 서열에서의 상응하는 돌연변이를 포함하며, 여기서 X는 임의의 아미노산이다. 일부 실시형태에서, SpCas9 도메인은 D1135V, G1218R, R1335Q 및 T1337R 돌연변이 중 하나 이상, 또는 본 명세서에 제공된 임의의 아미노산 서열에서의 상응하는 돌연변이를 포함한다. 일부 실시형태에서, SpCas9 도메인은 D1135V, G1218R, R1335Q 및 T1337R 돌연변이, 또는 본 명세서에 제공된 임의의 아미노산 서열에서의 상응하는 돌연변이를 포함한다.

[0664] 일부 예에서, 본 명세서에 개시된 염기 편집기의 CRISPR 단백질-유래된 도메인에 의해 인식되는 PAM은 염기 편집기를 인코딩하는 삽입물(*예를 들어*, AAV 삽입물)에 대한 별도의 올리고뉴클레오티드 상의 세포에 제공될 수 있다. 이러한 실시형태에서, 별도의 올리고뉴클레오티드 상에 PAM을 제공하는 것은 표적 서열과 동일한 폴리뉴클레오티드 상에 인접한 PAM이 존재하지 않기 때문에 그렇지 않으면 절단될 수 없는 표적 서열의 절단을 허용할 수 있다.

[0665] 실시형태에서, *S. 피오게네스* Cas9(SpCas9)는 게놈 조작을 위한 CRISPR 엔도뉴클레아제로 사용될 수 있다. 그러나, 다른 것들도 사용될 수 있다. 일부 실시형태에서, 특정 게놈 표적을 표적화하기 위해 다른 엔도뉴클레아제가 사용될 수 있다. 일부 실시형태에서, 비-NGG PAM 서열을 갖는 합성 SpCas9-유래된 변이체가 사용될 수 있다. 추가로, 다양한 종으로부터의 다른 Cas9 이종상동체가 식별되었고 이들 "비-SpCas9"는 본 개시내용에 또한 유용할 수 있는 다양한 PAM 서열에 결합할 수 있다. 예를 들어, SpCas9의 상대적으로 큰 크기(대략적으로 4kb 코딩 서열)는 세포에서 효율적으로 발현될 수 없는 SpCas9 cDNA를 수반하는 플라스미드로 이어질 수 있다. 역으로, *스타필로코커스 아우레우스* Cas9(SaCas9)에 대한 코딩 서열은 SpCas9보다 대략적으로 1킬로베이스 짧아, 가능하기로는 세포에서 그것이 효율적으로 발현되도록 허용한다. SpCas9와 유사하게 SaCas9 엔도뉴클레아제는 *시험관내* 포유동물 세포와 *생체내* 마우스에서 표적 유전자를 변형할 수 있다. 일부 실시형태에서, Cas 단백질은 다른 PAM 서열을 표적화할 수 있다. 일부 실시형태에서, 표적 유전자는 예를 들어 Cas9 PAM, 5'-NGG에 인접할 수 있다. 다른 실시형태에서, 다른 Cas9 오솔로그는 다른 PAM 요구사항을 가질 수 있다. 예를 들어 *S. 써모필루스*(CRISPR1의 경우 5'-NNAGAA, CRISPR3의 경우 5'-NGGNG) 및 *나이세리아 메닝기티디스*(5'-NNNNGATT)의 것과 같은 다른 PAM이 또한 표적 유전자에 인접하여 발견될 수 있다.

[0666] 일부 실시형태에서, *S. 피오게네스* 시스템의 경우, 표적 유전자 서열은 5'-NGG PAM보다 앞에 있을 수 있고(즉, 이에 5'일 수 있고), 20-nt 가이드 RNA 서열은 반대 가닥과 염기 쌍을 이루어 PAM에 인접한 Cas9 절단을 매개할 수 있다. 일부 실시형태에서, 인접한 컷팅은 PAM의 상류 3개 또는 약 3개 염기쌍일 수 있다. 일부 실시형태에서, 인접한 컷팅은 PAM의 상류 10개 또는 약 10개 염기쌍일 수 있다. 일부 실시형태에서, 인접한 컷팅은 PAM의 상류 0-20개 또는 약 0-20개 염기쌍일 수 있다. 예를 들어 인접한 컷팅은 PAM의 상류 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29 또는 30

개 염기쌍 옆에 있을 수 있다. 인접한 컷팅은 또한 1 내지 30개 염기쌍만큼 PAM의 하류에 있을 수 있다. PAM 서열에 결합할 수 있는 예시적인 SpCas9 단백질의 서열은 다음과 같다:

[0667] 일부 실시형태에서, 조작된 SpCas9 변이체는 3' H(비-G PAM)에 의해 측접된 프로토스페이서 인접 모티프(PAM) 서열을 인식할 수 있다(**표 2A-2D** *참조*). 일부 실시형태에서, SpCas9 변이체는 NRNH PAM(여기서 R은 A 또는 G이고 H는 A, C 또는 T임)을 인식한다. 일부 실시형태에서, 비-G PAM은 NRRH, NRTH 또는 NRCH이다(*예를 들어*, Miller, S.M., 등 Continuous evolution of SpCas9 variants compatible with non-G PAMs, Nat. Biotechnol. (2020)을 참조하며, 그 내용은 그 전체가 참고로 본 명세서에 포함된다).

[0668] 일부 실시형태에서 Cas9 도메인은 재조합 Cas9 도메인이다. 일부 실시형태에서, 재조합 Cas9 도메인은 SpyMacCas9 도메인이다. 일부 실시형태에서, SpyMacCas9 도메인은 뉴클레아제 활성 SpyMacCas9, 뉴클레아제 불활성 SpyMacCas9(SpyMacCas9d), 또는 SpyMacCas9 닉카제(SpyMacCas9n)이다. 일부 실시형태에서, SaCas9 도메인, SaCas9d 도메인 또는 SaCas9n 도메인은 비-규범적 PAM을 갖는 핵산 서열에 결합할 수 있다. 일부 실시형태에서, SpyMacCas9 도메인, SpCas9d 도메인, 또는 SpCas9n 도메인은 NAA PAM 서열을 갖는 핵산 서열에 결합할 수 있다.

[0669] 천연 5'-NAAN-3' PAM 특이성을 갖는 *스트렙토코커스 마카카에*에서 Spy Cas9의 예시적인 Cas9 A 상동체의 서열은 당업계에 공지되어 있고, 예를 들어 Jakimo 등, (www.biorxiv.org/content/biorxiv/early/2018/09/27/429654.full.pdf)에 기재되어 있고, 서열 목록에서 서열 번호: 1307에 있다.

[0670] 일부 실시형태에서, 변이체 Cas9 단백질은 H840A, P475A, W476A, N477A, D1125A, W1126A 및 D1218A 돌연변이를 담지하여 폴리펩티드가 표적 DNA 또는 RNA를 절단하는 감소된 능력을 갖는다. 일부 실시형태에서, 변이체 Cas9 단백질은 서열번호: 234에 비해 H840A, P475A, W476A, N477A, D1125A, W1126A 및 D1218A 돌연변이를 담지한다. 이러한 Cas9 단백질은 표적 DNA(*예를 들어*, 단일 가닥 표적 DNA)를 절단하는 감소된 능력을 갖지만, 표적 DNA (*예를 들어*, 단일 가닥 표적 DNA)에 결합하는 능력을 유지한다. 또 다른 비-제한적인 예로서, 일부 실시형태에서 변이체 Cas9 단백질은 D10A, H840A, P475A, W476A, N477A, D1125A, W1126A 및 D1218A 돌연변이를 담지하여 폴리펩티드가 표적 DNA를 절단하는 감소된 능력을 갖는다. 이러한 Cas9 단백질은 표적 DNA(*예를 들어*, 단일 가닥 표적 DNA)를 절단하는 감소된 능력을 갖지만, 표적 DNA(*예를 들어*, 단일 가닥 표적 DNA)에 결합하는 능력을 유지한다. 일부 실시형태에서, 변이 Cas9 단백질은 서열번호: 234에 비해 D10A, H840A, P475A, W476A, N477A, D1125A, W1126A 및 D1218A 돌연변이를 담지한다. 일부 실시형태에서, 변이체 Cas9 단백질이 W476A 및 W1126A 돌연변이를 담지하는 경우 또는 변이체 Cas9 단백질은 P475A, W476A, N477A, D1125A, W1126A 및 D1218A 돌연변이를 담지하는 경우, 변이체 Cas9 단백질은 PAM 서열에 효율적으로 결합하지 않는다. 따라서, 일부 이러한 경우에, 이러한 변이체 Cas9 단백질이 결합의 방법에 사용되는 경우, 방법은 PAM 서열을 필요로 하지 않는다. 환언하면, 일부 실시형태에서 이러한 변이체 Cas9 단백질이 결합의 방법에 사용되는 경우, 방법은 가이드 RNA를 포함할 수 있지만, 방법은 PAM 서열의 부재에서 수행될 수 있다(그리고 결합의 특이성은 따라서 가이드 RNA의 표적화 세그먼트에 의해 제공된다). 상기의 효과를 달성하기 위해 다른 잔기가 돌연변이될 수 있다(즉, 하나 또는 다른 뉴클레아제 부분을 불활성화함). 비-제한적인 예로서, 잔기 D10, G12, G17, E762, H840, N854, N863, H982, H983, A984, D986 및/또는 A987은 변경(즉, 치환)될 수 있다. 또한, 알라닌 치환 이외의 돌연변이도 적합하다.

[0671] 일부 실시형태에서, 염기 편집기의 CRISPR 단백질-유래된 도메인은 규범적 PAM 서열(NGG)을 갖는 Cas9 단백질의 전부 또는 일부를 포함할 수 있다. 다른 실시형태에서, 염기 편집기의 Cas9-유래된 도메인은 비-규범적 PAM 서열을 활용할 수 있다. 이러한 서열은 당업계에 기술되어 있고 당업자에게 명백할 것이다. 예를 들어, 비-규범적 PAM 서열에 결합하는 Cas9 도메인은 Kleinstiver, B. P., 등, "Engineered CRISPR-Cas9 nucleases with altered PAM specificities" Nature 523, 481-485 (2015); 및 Kleinstiver, B. P., 등, "Broadening the targeting range of Staphylococcus aureus CRISPR-Cas9 by modifying PAM recognition" Nature Biotechnology 33, 1293-1298 (2015); R.T. Walton 등 "Unconstrained genome targeting with near-PAMless engineered CRISPR-Cas9 variants" *Science* 10.1126/science.aba8853 (2020); Hu 등 "Evolved Cas9 variants with broad PAM compatibility and high DNA specificity," *Nature*, 2018 Apr. 5, 556(7699), 57-63; Miller 등, "Continuous evolution of SpCas9 variants compatible with non-G PAMs" *Nat. Biotechnol.*, 2020 Apr;38(4):471-481에 기술되어 있으며; 각각의 전체 내용은 참조로 본 명세서에 포함된다.

[0672] ***NapDNAbp 및 시티딘 데아미나제 및/또는 아데노신 데아미나제를 포함하는 융합 단백질***

[0673] 개시내용의 일부 양태는 Cas9 도메인 또는 다른 핵산 프로그램가능 DNA 결합 단백질(*예를 들어*, Cas12) 및 하나 이상의 시티딘 데아미나제 또는 아데노신 데아미나제 도메인을 포함하는 융합 단백질을 제공한다. Cas9 도메인은 본 명세서에 제공된 임의의 Cas9 도메인 또는 Cas9 단백질(*예를 들어*, dCas9 또는 nCas9)일 수 있음을 이해해야 한다. 일부 실시형태에서, 본 명세서에 제공된 임의의 Cas9 도메인 또는 Cas9 단백질(*예를 들어*, dCas9 또는 nCas9)은 본 명세서에 제공된 임의의 시티딘 데아미나제 및/또는 아데노신 데아미나제와 융합될 수 있다. 본 명세서에 개시된 염기 편집기의 도메인은 임의의 순서로 배열될 수 있다.

[0674] 일부 실시형태에서, 융합 단백질은 다음 도메인 A-C, A-D, 또는 A-E를 포함한다:

[0675] NH_2-[A-B-C]-COOH;

[0676] NH_2-[A-B-C-D]-COOH; 또는

[0677] NH_2-[A-B-C-D-E]-COOH;

[0678] 여기서 A 및 C 또는 A, C 및 E는 각각 다음 중 하나 이상을 포함한다:

[0679] 아데노신 데아미나제 도메인 또는 이의 활성 단편,

[0680] 시티딘 데아미나제 도메인 또는 이의 활성 단편, 그리고

[0681] 여기서 B 또는 B 및 D는 각각 핵산 서열 특이적 결합 활성을 갖는 하나 이상의 도메인을 포함한다.

[0682] 일부 실시형태에서, 융합 단백질은 다음 구조를 포함한다:

[0683] NH_2-[A_n-B_o-C_n]-COOH;

[0684] NH_2-[A_n-B_o-C_n-D_o]-COOH; 또는

[0685] NH_2-[A_n-B_o-C_p-D_o-E_q]-COOH;

[0686] 여기서 A 및 C 또는 A, C 및 E는 각각 다음 중 하나 이상을 포함한다:

[0687] 아데노신 데아미나제 도메인 또는 이의 활성 단편,

[0688] 시티딘 데아미나제 도메인 또는 이의 활성 단편, 그리고

[0689] 여기서 n은 정수: 1, 2, 3, 4 또는 5이고, p는 정수: 0, 1, 2, 3, 4 또는 5이고; q는 정수 0, 1, 2, 3, 4 또는 5이고; 그리고 B 또는 B와 D는 각각 핵산 서열 특이적 결합 활성을 갖는 도메인을 포함하고; 그리고 o는 정수: 1, 2, 3, 4 또는 5이다.

[0690] 예를 들어, 제한 없이, 일부 실시형태에서, 융합 단백질은 다음 구조를 포함한다:

[0691] NH2-[아데노신 데아미나제]-[Cas9 도메인]-COOH;

[0692] NH2-[Cas9 도메인]-[아데노신 데아미나제]-COOH;

[0693] NH2-[시티딘 데아미나제]-[Cas9 도메인]-COOH;

[0694] NH2-[Cas9 도메인]-[시티딘 데아미나제]-COOH;

[0695] NH2-[시티딘 데아미나제]-[Cas9 도메인]-[아데노신 데아미나제]-COOH;

[0696] NH2-[아데노신 데아미나제]-[Cas9 도메인]-[시티딘 데아미나제]-COOH;

[0697] NH2-[아데노신 데아미나제]-[시티딘 데아미나제]-[Cas9 도메인]-COOH;

[0698] NH2-[시티딘 데아미나제]-[아데노신 데아미나제]-[Cas9 도메인]-COOH;

[0699] NH2-[Cas9 도메인]-[아데노신 데아미나제]-[시티딘 데아미나제]-COOH; 또는

[0700] NH2-[Cas9 도메인]-[시티딘 데아미나제]-[아데노신 데아미나제]-COOH.

[0701] 일부 실시형태에서, 본 명세서에 제공된 임의의 Cas12 도메인 또는 Cas12 단백질은 본 명세서에 제공된 임의의 시티딘 또는 아데노신 데아미나제와 융합될 수 있다. 예를 들어, 제한 없이, 일부 실시형태에서, 융합 단백질은

다음 구조를 포함한다:

[0702] NH2-[아데노신 데아미나제]-[Cas12 도메인]-COOH;

[0703] NH2-[Cas12 도메인]-[아데노신 데아미나제]-COOH;

[0704] NH2-[시티딘 데아미나제]-[Cas12 도메인]-COOH;

[0705] NH2-[Cas12 도메인]-[시티딘 데아미나제]-COOH;

[0706] NH2-[시티딘 데아미나제]-[Cas12 도메인]-[아데노신 데아미나제]-COOH;

[0707] NH2-[아데노신 데아미나제]-[Cas12 도메인]-[시티딘 데아미나제]-COOH;

[0708] NH2-[아데노신 데아미나제]-[시티딘 데아미나제]-[Cas12 도메인]-COOH;

[0709] NH2-[시티딘 데아미나제]-[아데노신 데아미나제]-[Cas12 도메인]-COOH;

[0710] NH2-[Cas12 도메인]-[아데노신 데아미나제]-[시티딘 데아미나제]-COOH; 또는

[0711] NH2-[Cas12 도메인]-[시티딘 데아미나제]-[아데노신 데아미나제]-COOH.

[0712] 일부 실시형태에서, 아데노신 데아미나제는 TadA*8이다. 예시적인 융합 단백질 구조는 다음을 포함한다:

[0713] NH2-[TadA*8]-[Cas9 도메인]-COOH;

[0714] NH2-[Cas9 도메인]-[TadA*8]-COOH;

[0715] NH2-[TadA*8]-[Cas12 도메인]-COOH; 또는

[0716] NH2-[Cas12 도메인]-[TadA*8]-COOH.

[0717] 일부 실시형태에서, 융합 단백질의 아데노신 데아미나제는 TadA*8 및 시티딘 데아미나제 및/또는 아데노신 데아미나제를 포함한다. 일부 실시형태에서, TadA*8은 TadA*8.1, TadA*8.2, TadA*8.3, TadA*8.4, TadA*8.5, TadA*8.6, TadA*8.7, TadA*8.8, TadA*8.9, TadA*8.10, TadA*8.11, TadA*8.12, TadA*8.13, TadA*8.14, TadA*8.15, TadA*8.16, TadA*8.17, TadA*8.18, TadA*8.19, TadA*8.20, TadA*8.21, TadA*8.22, TadA*8.23, 또는 TadA*8.24이다.

[0718] 예시적인 융합 단백질 구조는 다음을 포함한다:

[0719] NH2-[TadA*8]-[Cas9/Cas12]-[아데노신 데아미나제]-COOH;

[0720] NH2-[아데노신 데아미나제]-[Cas9/Cas12]-[TadA*8]-COOH;

[0721] NH2-[TadA*8]-[Cas9/Cas12]-[시티딘 데아미나제]-COOH; 또는

[0722] NH2-[시티딘 데아미나제]-[Cas9/Cas12]-[TadA*8]-COOH.

[0723] 일부 실시형태에서, 융합 단백질의 아데노신 데아미나제는 TadA*9 및 시티딘 데아미나제 및/또는 아데노신 데아미나제를 포함한다. 예시적인 융합 단백질 구조는 다음을 포함한다:

[0724] NH2-[TadA*9]-[Cas9/Cas12]-[아데노신 데아미나제]-COOH;

[0725] NH2-[아데노신 데아미나제]-[Cas9/Cas12]-[TadA*9]-COOH;

[0726] NH2-[TadA*9]-[Cas9/Cas12]-[시티딘 데아미나제]-COOH; 또는

[0727] NH2-[시티딘 데아미나제]-[Cas9/Cas12]-[TadA*9]-COOH.

[0728] 일부 실시형태에서, 융합 단백질은 Cas9 또는 Cas12 폴리펩티드의 N-말단 단편 및 C-말단 단편에 의해 측접된 데아미나제를 포함할 수 있다. 일부 실시형태에서, 융합 단백질은 Cas9 또는 Cas12 폴리펩티드의 N-말단 단편 및 C-말단 단편에 의해 측접된 시티딘 데아미나제를 포함한다. 일부 실시형태에서, 융합 단백질은 Cas9 또는 Cas 12 폴리펩티드의 N-말단 단편 및 C-말단 단편에 의해 측접된 아데노신 데아미나제를 포함한다.

[0729] 일부 실시형태에서, 시티딘 데아미나제 또는 아데노신 데아미나제 및 napDNAbp(*예를 들어*, Cas9 또는 Cas12 도메인)를 포함하는 융합 단백질은 링커 서열을 포함하지 않는다. 일부 실시형태에서, 링커는 시티딘 또는 아데노신 데아미나제와 napDNAbp 사이에 존재한다. 일부 실시형태에서, 상기의 일반 아키텍처에서 사용된 "-"는 선택

적인 링커의 존재를 나타낸다. 일부 실시형태에서 시티딘 또는 아데노신 데아미나제와 napDNAbp는 본 명세서에 제공된 임의의 링커를 통해 융합된다. 예를 들어, 일부 실시형태에서 시티딘 또는 아데노신 데아미나제 및 napDNAbp는 본 명세서에 제공된 임의의 링커를 통해 융합된다.

[0730] 본 개시내용의 융합 단백질은 하나 이상의 추가 특징을 포함할 수 있음을 이해해야 한다. 예를 들어, 일부 실시형태에서, 융합 단백질은 억제제, 세포질 국소화 서열, 핵 유출 서열과 같은 유출 서열, 또는 다른 국소화 서열뿐만 아니라 융합 단백질의 가용화, 정제 또는 검출에 유용한 서열 태그를 포함할 수 있다. 본 명세서에서 제공된 적합한 단백질 태그는 비오틴 카르복실라제 담체 단백질(BCCP) 태그, myc-태그, 칼모듈린-태그, FLAG-태그, 헤마글루티닌(HA)-태그, 또한 히스티딘 태그 또는 His-태그로도 지칭되는, 폴리히스티딘 태그, 말토스 결합 단백질(MBP)-태그, nus-태그, 글루타티온-S-트랜스퍼라제(GST)-태그, 녹색 형광 단백질(GFP)-태그, 티오레독신-태그, S-태그, Softag(예를 들어, Softag 1, Softag 3), 스트렙-태그, 비오틴 리가제 태그, FlAsH 태그, V5 태그 및 SBP-태그를 포함하지만 이에 제한되지 않는다. 추가적인 적합한 서열은 당업자에게 명백할 것이다. 일부 실시형태에서, 융합 단백질은 하나 이상의 His 태그를 포함한다.

[0731] 예시적이지만 비제한적인 융합 단백질은 국제 PCT 출원 번호 PCT/2017/044935, PCT/US2019/044935, 및 PCT/US2020/016288에 기재되어 있으며, 이들 각각은 그 전체가 본 명세서에 참고로 포함된다.

[0732] ***핵 국소화 서열(NLS)을 포함하는 융합 단백질***

[0733] 일부 실시형태에서, 본 명세서에 제공된 융합 단백질은 하나 이상(*예를 들어*, 2, 3, 4, 5)의 핵 표적화 서열, 예를 들어 핵 국소화 서열(NLS)을 추가로 포함한다. 일 실시형태에서, 이분 NLS가 사용된다. 일부 실시형태에서, NLS는 세포 핵 안으로(*예를 들어*, 핵 수송에 의해) NLS를 포함하는 단백질의 유입을 촉진하는 아미노산 서열을 포함한다. 일부 실시형태에서, NLS는 융합 단백질의 N-말단 또는 C-말단에 융합된다. 일부 실시형태에서, NLS는 nCas9 도메인 또는 dCas9 도메인의 C-말단 또는 N-말단에 융합된다. 일부 실시형태에서, NLS는 Cas12 도메인의 N-말단 또는 C-말단에 융합된다. 일부 실시형태에서, NLS는 시티딘 또는 아데노신 데아미나제의 N-말단 또는 C-말단에 융합된다. 일부 실시형태에서, NLS는 하나 이상의 링커를 통해 융합 단백질에 융합된다. 일부 실시형태에서, NLS는 링커 없이 융합 단백질에 융합된다. 일부 실시형태에서, NLS는 본 명세서에 제공되거나 참조된 NLS 서열 중 임의의 하나의 아미노산 서열을 포함한다. 추가적인 핵 국소화 서열은 당업계에 공지되어 있으며 당업자에게 명백할 것이다. 예를 들어, NLS 서열은 Plank 등의 PCT/EP2000/011690에 기재되어 있으며, 그 내용은 예시적인 핵 국소화 서열의 개시에 대해 본 명세서에 참고로 포함된다. 일부 실시형태에서, NLS는 아미노산 서열 PKKKRKVEGADKRTADGSEFESPKKKRKV(서열번호: 83), KRTADGSEFESPKKKRKV(서열번호: 84), KRPAATKKAGQAKKKK(서열번호: 85), KKTELQTTNAENKTKKL(서열번호: 86), KRGINDRNFWRGENGRKTR(서열번호: 87), RKSGKIAAIVVKRPRKPKKKRKV(서열번호: 1424), 또는 MDSLLMNRRKFLYQFKNVRWAKGRRETYLC(서열번호: 90)를 포함한다.

[0734] 일부 실시형태에서, 시티딘 또는 아데노신 데아미나제, Cas9 도메인 및 NLS를 포함하는 융합 단백질은 링커 서열을 포함하지 않는다. 일부 실시형태에서, 하나 이상의 도메인 또는 단백질(*예를 들어*, 시티딘 또는 아데노신 데아미나제, Cas9 도메인 또는 NLS) 사이에 링커 서열이 존재한다. 일부 실시형태에서, 링커는 시티딘 데아미나제 및 아데노신 데아미나제 도메인과 napDNAbp 사이에 존재한다. 일부 실시형태에서, 아래의 일반 아키텍처에서 사용된 "-"는 선택적인 링커의 존재를 나타낸다. 일부 실시형태에서, 시티딘 데아미나제 및 아데노신 데아미나제와 napDNAbp는 본 명세서에 제공된 임의의 링커를 통해 융합된다. 예를 들어, 일부 실시형태에서 시티딘 데아미나제 및 아데노신 데아미나제와 napDNAbp는 본 명세서에 제공된 임의의 링커를 통해 융합된다.

[0735] 일부 실시형태에서, 시티딘 또는 아데노신 데아미나제 및 napDNAbp(*예를 들어*, Cas9 또는 Cas12) 도메인을 갖는 예시적인 napDNAbp(*예를 들어*, Cas9 또는 Cas12) 융합 단백질의 일반 아키텍처는 다음 구조 중 임의의 하나를 포함하며, 여기서 NLS는 핵 국소화 서열(*예를 들어*, 본 명세서에 제공된 임의의 NLS)이고, NH_2는 융합 단백질의 N-말단이고, COOH는 융합 단백질의 C-말단이다:

[0736] NH_2-NLS-[시티딘 데아미나제]-[napDNAbp 도메인]-COOH;

[0737] NH_2-NLS [napDNAbp 도메인]-[시티딘 데아미나제]-COOH;

[0738] NH_2-[시티딘 데아미나제]-[napDNAbp 도메인]-NLS-COOH;

[0739] NH_2-[napDNAbp 도메인]-[시티딘 데아미나제]-NLS-COOH;

[0740] NH_2-NLS-[아데노신 데아미나제]-[napDNAbp 도메인]-COOH;

[0741] NH_2-NLS [napDNAbp 도메인]-[아데노신 데아미나제]-COOH;

[0742] NH_2-[아데노신 데아미나제]-[napDNAbp 도메인]-NLS-COOH;

[0743] NH_2-[napDNAbp 도메인]-[아데노신 데아미나제]-NLS-COOH;

[0744] NH_2-NLS-[시티딘 데아미나제]-[napDNAbp 도메인]-[아데노신 데아미나제]-COOH;

[0745] NH_2-NLS-[아데노신 데아미나제]-[napDNAbp 도메인]-[시티딘 데아미나제]-COOH;

[0746] NH_2-NLS-[아데노신 데아미나제] [시티딘 데아미나제]-[napDNAbp 도메인]-COOH;

[0747] NH_2-NLS-[시티딘 데아미나제]-[아데노신 데아미나제]-[napDNAbp 도메인]-COOH;

[0748] NH_2-NLS-[napDNAbp 도메인]-[아데노신 데아미나제]-[시티딘 데아미나제]-COOH;

[0749] NH_2-NLS-[napDNAbp 도메인]-[시티딘 데아미나제]-[아데노신 데아미나제]-COOH;

[0750] NH_2-[시티딘 데아미나제]-[napDNAbp 도메인]-[아데노신 데아미나제]-NLS-COOH;

[0751] NH_2-[아데노신 데아미나제]-[napDNAbp 도메인]-[시티딘 데아미나제]-NLS-COOH;

[0752] NH_2-[아데노신 데아미나제] [시티딘 데아미나제]-[napDNAbp 도메인]-NLS-COOH;

[0753] NH_2-[시티딘 데아미나제]-[아데노신 데아미나제]-[napDNAbp 도메인]-NLS-COOH;

[0754] NH_2-[napDNAbp 도메인]-[아데노신 데아미나제]-[시티딘 데아미나제]-NLS-COOH; 또는

[0755] NH_2-[napDNAbp 도메인]-[시티딘 데아미나제]-[아데노신 데아미나제]-NLS-COOH. 일부 실시형태에서, NLS는 링커에 존재하거나 NLS는, 예를 들어, 본 명세서에 기재된 링커에 의해 측접된다. 이분 NLS는 상대적으로 짧은 스페이서 서열에 의해 분리된 2개의 염기성 아미노산 클러스터를 포함한다(따라서 단분 NLS는 그렇지 않지만, 이분 - 2개 부분임). 뉴클레오플라스민의 NLS인 KR[PAATKKAGQA]KKKK(서열번호: 85)는 유비퀴터스 이분 신호: 약 10개 아미노산의 스페이서에 의해 분리된, 염기성 아미노산의 두 클러스터의 원형이다. 예시적인 이분 NLS의 서열은 다음과 같다:

[0756] PKKKRKVEGADKRTADGSEFESPKKKRKV(서열번호: 83)

[0757] 하나 이상의 핵 국소화 서열(NLS)을 포함하는 CRISPR 효소를 인코딩하는 벡터가 사용될 수 있다. 예를 들어, 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10개의 NLS가 사용될 수 있다. CRISPR 효소는 아미노-말단 또는 그 근처에서 NLS, 카르복시-말단 또는 그 근처에서 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10개 또는 그 이상의 NLS, 또는 이의 임의의 조합(*예를 들어*, 아미노-말단에서 하나 이상의 NLS 및 카르복시 말단에서 하나 이상의 NLS)을 포함할 수 있다. 하나 초과의 NLS가 존재할 때, 각각은 서로 독립적으로 선택될 수 있어, 단일 NLS는 하나 초과의 카피에 존재할 수 있고/있거나 하나 이상의 카피에 존재하는 하나 이상의 다른 NLS와 조합되어 존재할 수 있다.

[0758] 방법에 사용된 CRISPR 효소는 약 6개의 NLS를 포함할 수 있다. NLS에 가장 가까운 아미노산이 N- 또는 C-말단으로부터 폴리펩티드 사슬을 따라 약 50개 아미노산 이내, *예를 들어* 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 40 또는 50개 아미노산 이내일 때 NLS는 N- 또는 C-말단 근처로 간주된다.

[0759] ***추가 도메인***

[0760] 본 명세서에 기재된 염기 편집기는 폴리뉴클레오티드의 핵염기의 핵염기 편집, 변형 또는 변경을 촉진하는데 도움이 되는 임의의 도메인을 포함할 수 있다. 일부 실시형태에서, 염기 편집기는 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인(*예를 들어*, Cas9), 핵염기 편집 도메인(*예를 들어*, 데아미나제 도메인) 및 하나 이상의 추가 도메인을 포함한다. 일부 실시형태에서, 추가 도메인은 염기 편집기의 효소적 또는 촉매적 기능, 염기 편집기의 결합 기능을 촉진할 수 있거나, 원하는 염기 편집 결과를 방해할 수 있는 세포 기계(*예를 들어*, 효소)의 억제제일 수 있다. 일부 실시형태에서, 염기 편집기는 뉴클레아제, 닉카제, 재조합효소, 데아미나제, 메

틸트랜스퍼라제, 메틸라제, 아세틸라제, 아세틸트랜스퍼라제, 전사 활성제, 또는 전사 억제인자 도메인을 포함할 수 있다.

[0761] 일부 실시형태에서, 염기 편집기는 우라실 글리코실라제 억제제(UGI) 도메인을 포함할 수 있다. 일부 실시형태에서, U:G 이종이중나선 DNA의 존재에 대한 세포 DNA 수선 반응은 세포에서 핵염기 편집 효율성에서의 감소를 담당할 수 있다. 이러한 실시형태에서, 우라실 DNA 글리코실라제(UDG)는 세포에서 DNA로부터 U의 제거를 촉매할 수 있으며, 이는 염기 절제 수선(BER)을 개시할 수 있고, 대부분 C:G 쌍으로 U:G 쌍의 전환을 초래한다. 이러한 실시형태에서, BER은 단일 가닥에 결합하고, 편집된 염기를 차단하고, UGI를 억제하고, BER을 억제하고, 편집된 염기를 보호하고, 및/또는 비-편집된 가닥의 수선을 촉진하는 하나 이상의 도메인을 포함하는 염기 편집기에서 억제될 수 있다. 따라서, 본 개시내용은 UGI 도메인을 포함하는 염기 편집기 융합 단백질을 고려한다.

[0762] 일부 실시형태에서, 염기 편집기는 이중-가닥 파괴(DSB) 결합 단백질의 전부 또는 일부를 도메인으로 포함한다. 예를 들어, DSB 결합 단백질은 DSB의 말단에 결합할 수 있고 분해로부터 이를 보호할 수 있는 박테리오파지 Mu의 Gam 단백질을 포함할 수 있다. Komor, A.C., 등, "Improved base excision repair inhibition and bacteriophage Mu Gam protein yields C:G-to-T:A base editors with higher efficiency and product purity" Science Advances 3:eaao4774 (2017)를 참조하며, 그 전체 그 내용은 참조로 본 명세서에 포함된다.

[0763] 추가적으로, 일부 실시형태에서, Gam 단백질은 염기 편집기의 N 말단에 융합될 수 있다. 일부 실시형태에서, Gam 단백질은 염기 편집기의 C 말단에 융합될 수 있다. 박테리오파지 Mu의 Gam 단백질은 이중 가닥 파괴(DSB)의 종단에 결합하고 분해로부터 이를 보호할 수 있다. 일부 실시형태에서, DSB의 유리 종단에 결합하기 위해 Gam을 사용하는 것은 염기 편집의 과정 동안 삽입결실 형성을 감소시킬 수 있다. 일부 실시형태에서, 174-잔기 Gam 단백질은 염기 편집기의 N 말단에 융합된다. Komor, A.C., 등, "Improved base excision repair inhibition and bacteriophage Mu Gam protein yields C:G-to-T:A base editors with higher efficiency and product purity" Science Advances 3:eaao4774 (2017)를 참조한다. 일부 실시형태에서, 돌연변이 또는 돌연변이들은 야생형 도메인에 비해 염기 편집기 도메인의 길이를 변경할 수 있다. 예를 들어, 적어도 하나의 도메인에서 적어도 하나의 아미노산의 삭제는 염기 편집기의 길이를 감소시킬 수 있다. 다른 경우에, 돌연변이 또는 돌연변이들은 야생형 도메인에 비해 도메인의 길이를 변경하지 않는다. 예를 들어, 임의의 도메인에서 치환은 염기 편집기의 길이를 변경하지 않는다.

[0764] 모든 도메인의 길이가 야생형 도메인과 동일한 이러한 염기 편집기의 비-제한적인 예는 다음을 포함할 수 있다:

[0765] NH2-[핵염기 편집 도메인]-링커1-[APOBEC1]-링커2-[핵염기 편집 도메인]-COOH;

[0766] NH2-[핵염기 편집 도메인]-링커1-[APOBEC1]-[핵염기 편집 도메인]-COOH;

[0767] NH2-[핵염기 편집 도메인]-[APOBEC1]-링커2-[핵염기 편집 도메인]-COOH;

[0768] NH2-[핵염기 편집 도메인]-[APOBEC1]-[핵염기 편집 도메인]-COOH;

[0769] NH2-[핵염기 편집 도메인]-링커1-[APOBEC1]-링커2-[핵염기 편집 도메인]-[UGI]-COOH;

[0770] NH2-[핵염기 편집 도메인]-링커1-[APOBEC1]-[핵염기 편집 도메인]-[UGI]-COOH;

[0771] NH2-[핵염기 편집 도메인]-[APOBEC1]-링커2-[핵염기 편집 도메인]-[UGI]-COOH;

[0772] NH2-[핵염기 편집 도메인]-[APOBEC1]-[핵염기 편집 도메인]-[UGI]-COOH;

[0773] NH2-[UGI]-[핵염기 편집 도메인]-링커1-[APOBEC1]-링커2-[핵염기 편집 도메인]-COOH;

[0774] NH2-[UGI]-[핵염기 편집 도메인]-링커1-[APOBEC1]-[핵염기 편집 도메인]-COOH;

[0775] NH2-[UGI]-[핵염기 편집 도메인]-[APOBEC1]-링커2-[핵염기 편집 도메인]-COOH; 또는

[0776] NH2-[UGI]-[핵염기 편집 도메인]-[APOBEC1]-[핵염기 편집 도메인]-COOH.

[0777] **F. 염기 편집기 시스템**

[0778] 염기 편집기 시스템을 사용하여 핵염기를 편집하기 위한 시스템, 조성물 및 방법이 본 명세서에 제공된다. 일부 실시형태에서, 염기 편집기 시스템은 (1) 핵염기를 편집하기 위한 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인 및 핵염기 편집 도메인(예를 들어, 데아미나제 도메인)을 포함하는 염기 편집기(BE); 및 (2) 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인과 조합한 가이드 폴리뉴클레오티드(예를 들어, 가이드

RNA)를 포함한다. 일부 실시형태에서, 염기 편집기 시스템은 시티딘 염기 편집기(CBE) 또는 아데노신 염기 편집기(ABE)이다. 일부 실시형태에서, 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인은 폴리뉴클레오티드 프로그램가능 DNA 또는 RNA 결합 도메인이다. 일부 실시형태에서, 핵염기 편집 도메인은 데아미나제 도메인이다. 일부 실시형태에서, 데아미나제 도메인은 시티딘 데아미나제 또는 시토신 데아미나제일 수 있다. 일부 실시형태에서, 데아미나제 도메인은 아데닌 데아미나제 또는 아데노신 데아미나제일 수 있다. 일부 실시형태에서, 아데노신 염기 편집기는 DNA에서 아데닌을 탈아미노화할 수 있다. 일부 실시형태에서, 염기 편집기는 DNA에서 시티딘을 탈아미노화할 수 있다.

[0779] 일부 실시형태에서, 본 명세서에 제공된 바와 같은 염기 편집 시스템은 이중-가닥 DNA 파괴를 생성함이 없이, 공여자 DNA 주형을 요함이 없이, 그리고 과도한 확률론적 삽입 및 결실을 유도함이 없이, DNA에서 프로그램가능 단일 뉴클레오티드(C→T 또는 A→G) 변화를 유도하기 위해 촉매적으로 결함이 있는 스트렙토코커스 피오게네스 Cas9, 데아미나제(예를 들어, 시티딘 또는 아데노신 데아미나제) 및 염기 절제 수선의 억제제를 함유하는 융합 단백질을 사용하는 게놈 편집에 대한 새로운 접근법을 제공한다.

[0780] 핵염기 편집 단백질의 세부사항은 국제 PCT 출원 번호 PCT/2017/045381(WO2018/027078) 및 PCT/US2016/058344(WO2017/070632)에 기재되어 있으며, 이들 각각은 그 전체가 본 명세서에 참조로 포함된다. 또한 Komor, A.C., 등, "Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage" Nature 533, 420-424 (2016); Gaudelli, N.M., 등, "Programmable base editing of A·T to G·C in genomic DNA without DNA cleavage" Nature 551, 464-471 (2017); 및 Komor, A.C., 등, "Improved base excision repair inhibition and bacteriophage Mu Gam protein yields C:G-to-T:A base editors with higher efficiency and product purity" Science Advances 3:eaao4774 (2017)를 참조하며, 그 전체 그 내용은 참조로 본 명세서에 포함된다.

[0781] 본 명세서에 제공된 염기 편집기 시스템의 사용은 다음의 단계를 포함한다: (a) 대상체의 폴리뉴클레오티드(예를 들어, 이중- 또는 단일 가닥 DNA 또는 RNA)의 표적 뉴클레오티드 서열을 핵염기 편집기(예를 들어, 아데노신 염기 편집기 또는 시티딘 염기 편집기) 및 가이드 다중핵산(예를 들어, gRNA)과 접촉시키는 단계로서, 여기서 표적 뉴클레오티드 서열은 표적화된 핵염기 쌍을 포함하는, 단계; (b) 상기 표적 영역의 가닥 분리를 유도하는 단계; (c) 표적 영역의 단일 가닥에서의 상기 표적 핵염기 쌍의 제1 핵염기를 제2 핵염기로 전환시키는 단계; 및 (d) 상기 표적 영역의 한 가닥 이하를 절단하는 단계로서, 여기서 제1 핵염기에 상보성인 제3 핵염기는 제2 핵염기에 상보성인 제4 핵염기에 의해 대체되는, 단계를 포함한다. 일부 실시형태에서 단계(b)가 생략된다는 것이 이해되어야 한다. 일부 실시형태에서, 상기 표적화된 핵염기 쌍은 하나 이상의 유전자에서의 복수의 핵염기 쌍이다. 일부 실시형태에서, 본 명세서에 제공된 염기 편집기 시스템은 하나 이상의 유전자에서의 복수의 핵염기 쌍의 다중 편집이 가능하다. 일부 실시형태에서, 복수의 핵염기 쌍은 동일한 유전자에 위치된다. 일부 실시형태에서, 복수의 핵염기 쌍은 하나 이상의 유전자에 위치되며, 여기서 적어도 하나의 유전자는 상이한 좌위에 위치된다.

[0782] 일부 실시형태에서, 컷팅된 단일 가닥(닉이 있는 가닥)은 가이드 핵산에 혼성화된다. 일부 실시형태에서, 컷팅된 단일 가닥은 제1 핵염기를 포함하는 가닥의 반대편에 있다. 일부 실시형태에서, 염기 편집기는 Cas9 도메인을 포함한다. 일부 실시형태에서, 제1 염기는 아데닌이고, 제2 염기는 G, C, A 또는 T가 아니다. 일부 실시형태에서, 제2 염기는 이노신이다.

[0783] 일부 실시형태에서, 단일 가이드 폴리뉴클레오티드는 데아미나제를 표적 핵산 서열로 표적화하는데 이용될 수 있다. 일부 실시형태에서, 단일 쌍의 가이드 폴리뉴클레오티드는 상이한 데아미나제를 표적 핵산 서열로 표적화하는데 이용될 수 있다.

[0784] 염기 편집기 시스템의 핵염기 구성요소와 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 구성요소는 공유적으로 또는 비-공유적으로 서로 연관될 수 있다. 예를 들어, 일부 실시형태에서, 데아미나제 도메인은 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인에 의해 표적 뉴클레오티드 서열에 표적화될 수 있다. 일부 실시형태에서, 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인은 데아미나제 도메인에 융합되거나 연결될 수 있다. 일부 실시형태에서, 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인은 데아미나제 도메인과 비-공유적으로 상호작용하거나 회합함에 의해 데아미나제 도메인을 표적 뉴클레오티드 서열에 표적화할 수 있다. 예를 들어, 일부 실시형태에서, 핵염기 편집 구성요소, 예를 들어 데아미나제 구성요소는 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인의 일부인 추가의 이종 부분 또는 도메인과 상호작용할 수 있거나, 회합할 수 있거나, 또는 복합체를 형성할 수 있는 추가의 이종 부분 또는 도메인을 포함할 수 있다. 일부 실시

형태에서, 추가의 이종 부분은 폴리펩티드에 결합하거나, 상호작용하거나, 회합하거나, 또는 복합체를 형성할 수 있다. 일부 실시형태에서, 추가의 이종 부분은 폴리뉴클레오티드에 결합하거나, 상호작용하거나, 회합하거나, 또는 복합체를 형성할 수 있다. 일부 실시형태에서, 추가의 이종 부분은 가이드 폴리뉴클레오티드에 결합할 수 있다. 일부 실시형태에서, 추가의 이종 부분은 폴리펩티드 링커에 결합할 수 있다. 일부 실시형태에서, 추가의 이종 부분은 폴리뉴클레오티드 링커에 결합할 수 있다. 추가의 이종 부분은 단백질 도메인일 수 있다. 일부 실시형태에서, 추가의 이종 부분은 K 상동체(KH) 도메인, MS2 코트 단백질 도메인, PP7 코트 단백질 도메인, SfMu Com 코트 단백질 도메인, 스테릴 알파 모티프, 텔로머라제 Ku 결합 모티프 및 Ku 단백질, 텔로머라제 Sm7 결합 모티프 및 Sm7 단백질, 또는 RNA 인식 모티프일 수 있다.

[0785] 염기 편집기 시스템은 가이드 폴리뉴클레오티드 구성요소를 추가로 포함할 수 있다. 염기 편집기 시스템의 구성요소는 공유 결합, 비공유 상호작용, 또는 이의 회합 및 상호작용의 임의의 조합을 통해 서로 회합될 수 있음이 이해되어야 한다. 일부 실시형태에서, 데아미나제 도메인은 가이드 폴리뉴클레오티드에 의해 표적 뉴클레오티드 서열로 표적화될 수 있다. 예를 들어, 일부 실시형태에서, 염기 편집기 시스템의 핵염기 편집 구성요소, 예를 들어 데아미나제 구성요소는 가이드 폴리뉴클레오티드의 부분 또는 세그먼트(예를 들어, 폴리뉴클레오티드 모티프)와 상호작용하거나, 회합하거나, 복합체를 형성할 수 있는 추가의 이종 부분 또는 도메인(예를 들어, RNA 또는 DNA 결합 단백질과 같은 폴리뉴클레오티드 결합 도메인)을 포함할 수 있다. 일부 실시형태에서, 추가의 이종 부분 또는 도메인(예를 들어, RNA 또는 DNA 결합 단백질과 같은 폴리뉴클레오티드 결합 도메인)은 데아미나제 도메인에 융합되거나 연결될 수 있다. 일부 실시형태에서, 추가의 이종 부분은 폴리펩티드에 결합하거나, 상호작용하거나, 회합하거나, 또는 복합체를 형성할 수 있다. 일부 실시형태에서, 추가의 이종 부분은 폴리뉴클레오티드에 결합하거나, 상호작용하거나, 회합하거나, 또는 복합체를 형성할 수 있다. 일부 실시형태에서, 추가의 이종 부분은 가이드 폴리뉴클레오티드에 결합할 수 있다. 일부 실시형태에서, 추가의 이종 부분은 폴리펩티드 링커에 결합할 수 있다. 일부 실시형태에서, 추가의 이종 부분은 폴리뉴클레오티드 링커에 결합할 수 있다. 추가의 이종 부분은 단백질 도메인일 수 있다. 일부 실시형태에서, 추가의 이종 부분은 K 상동체(KH) 도메인, MS2 코트 단백질 도메인, PP7 코트 단백질 도메인, SfMu Com 코트 단백질 도메인, 멸균 알파 모티프, 텔로머라제 Ku 결합 모티프 및 Ku 단백질, 텔로머라제 Sm7 결합 모티프 및 Sm7 단백질, 또는 RNA 인식 모티프일 수 있다.

[0786] 일부 실시형태에서, 염기 편집기 시스템은 염기 절제 수선(BER) 구성요소의 억제제를 추가로 포함할 수 있다. 염기 편집기 시스템의 구성요소는 공유 결합, 비공유 상호작용, 또는 이의 회합 및 상호작용의 임의의 조합을 통해 서로 회합될 수 있음이 이해되어야 한다. BER 구성요소의 억제제는 염기 절제 수선 억제제를 포함할 수 있다. 일부 실시형태에서, 염기 절제 수선의 억제제는 우라실 DNA 글리코실라제 억제제(UGI)일 수 있다. 일부 실시형태에서, 염기 절제 수선의 억제제는 이노신 염기 절제 수선 억제제일 수 있다. 일부 실시형태에서, 염기 절제 수선의 억제제는 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인에 의해 표적 뉴클레오티드 서열에 표적화될 수 있다. 일부 실시형태에서, 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인은 염기 절제 수선의 억제제에 융합되거나 연결될 수 있다. 일부 실시형태에서, 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인은 데아미나제 도메인 및 염기 절제 수선의 억제제에 융합되거나 연결될 수 있다. 일부 실시형태에서, 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인은 염기 절제 수선의 억제제와 비-공유적으로 상호작용하거나 회합함에 의해 염기 절제 수선의 억제제를 표적 뉴클레오티드 서열로 표적화할 수 있다. 예를 들어, 일부 실시형태에서, 염기 절제 수선의 억제제 구성요소는 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인의 일부인 추가의 이종 부분 또는 도메인과 상호작용할 수 있거나, 회합할 수 있거나, 또는 복합체를 형성할 수 있는 추가의 이종 부분 또는 도메인을 포함할 수 있다. 일부 실시형태에서, 염기 절제 수선의 억제제는 가이드 폴리뉴클레오티드에 의해 표적 뉴클레오티드 서열로 표적화될 수 있다. 예를 들어, 일부 실시형태에서, 염기 절제 수선의 억제제는 가이드 폴리뉴클레오티드의 일부 또는 세그먼트(*예를 들어*, 폴리뉴클레오티드 모티프)와 상호작용할 수 있거나, 회합할 수 있거나, 또는 복합체를 형성할 수 있는 추가의 이종 부분 또는 도메인(*예를 들어*, RNA 또는 DNA 결합 단백질과 같은 폴리뉴클레오티드 결합 도메인)을 포함할 수 있다. 일부 실시형태에서, 가이드 폴리뉴클레오티드의 추가의 이종 부분 또는 도메인(*예를 들어*, RNA 또는 DNA 결합 단백질과 같은 폴리뉴클레오티드 결합 도메인)은 염기 절제 수선의 억제제에 융합되거나 연결될 수 있다. 일부 실시형태에서, 추가의 이종 부분은 폴리뉴클레오티드에 결합하거나, 상호작용하거나, 회합하거나, 복합체를 형성할 수 있다. 일부 실시형태에서, 추가의 이종 부분은 가이드 폴리뉴클레오티드에 결합할 수 있다. 일부 실시형태에서, 추가의 이종 부분은 폴리펩티드 링커에 결합할 수 있다. 일부 실시형태에서, 추가의 이종 부분은 폴리뉴클레오티드 링커에 결합할 수 있다. 추가의 이종 부분은 단백질 도메인일 수 있다. 일부 실시형태에서, 추가의 이종 부분은 K 상동체(KH) 도메인, MS2 코트 단백질 도메인, PP7 코트 단백질 도메인, SfMu Com 코트 단백질 도메인, 멸균 알파 모티프, 텔로머라제 Ku 결합 모티프 및 Ku 단백질, 텔로머라제 Sm7 결합 모티프 및 Sm7 단백질, 또는

RNA 인식 모티프일 수 있다.

[0787] 일부 실시형태에서, 염기 편집기는 편집된 가닥의 염기 절제 수선(BER)을 억제한다. 일부 실시형태에서, 염기 편집기는 비-편집된 가닥을 보호하거나 결합한다. 일부 실시형태에서, 염기 편집기는 UGI 활성을 포함한다. 일부 실시형태에서, 염기 편집기는 촉매적으로 불활성인 이노신-특이적 뉴클레아제를 포함한다. 일부 실시형태에서, 염기 편집기는 닉카제 활성을 포함한다. 일부 실시형태에서, 염기쌍의 의도된 편집은 PAM 부위의 상류이다. 일부 실시형태에서, 염기쌍의 의도된 편집은 PAM 부위의 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 또는 20개 뉴클레오티드 상류이다. 일부 실시형태에서, 염기쌍의 의도된 편집은 PAM 부위의 하류이다. 일부 실시형태에서, 의도된 편집된 염기쌍은 PAM 부위의 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 또는 20개 뉴클레오티드 하류 스트림이다.

[0788] 일부 실시형태에서, 방법은 규범적(*예를 들어*, NGG) PAM 부위를 필요로 하지 않는다. 일부 실시형태에서, 핵염기 편집기는 링커 또는 스페이서를 포함한다. 일부 실시형태에서, 링커 또는 스페이서는 길이가 1-25개 아미노산이다. 일부 실시형태에서, 링커 또는 스페이서는 길이가 5-20개 아미노산이다. 일부 실시형태에서, 링커 또는 스페이서는 길이가 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 또는 20개 아미노산이다.

[0789] 일부 실시형태에서, 본 명세서에 제공된 염기 편집 융합 단백질은 정확한 장소, 예를 들어 표적 염기가 정의된 영역(예를 들어, "탈아미노화 윈도우") 내에 배치되는 곳에 위치되어야할 필요가 있다. 일부 실시형태에서, 표적은 4개 염기 영역 내에 있을 수 있다. 일부 실시형태에서, 이러한 정의된 표적 영역은 PAM 상류의 대략적으로 15개 염기일 수 있다. Komor, A.C., 등, "Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage" Nature 533, 420-424 (2016); Gaudelli, N.M., 등, "Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage" Nature 551, 464-471 (2017); 및 Komor, A.C., 등, "Improved base excision repair inhibition and bacteriophage Mu Gam protein yields C:G-to-T:A base editors with higher efficiency and product purity" Science Advances 3:eaao4774 (2017)를 참조하며, 그 전체 내용은 참조로 본 명세서에 포함된다.

[0790] 일부 실시형태에서, 표적 영역은 표적 윈도우를 포함하며, 여기서 표적 윈도우는 표적 핵염기 쌍을 포함한다. 일부 실시형태에서, 표적 윈도우는 1-10개의 뉴클레오티드를 포함한다. 일부 실시형태에서, 표적 윈도우는 길이가 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 또는 20개의 뉴클레오티드이다. 일부 실시형태에서, 염기쌍의 의도된 편집은 표적 윈도우 내에 있다. 일부 실시형태에서, 표적 윈도우는 염기쌍의 의도된 편집을 포함한다. 일부 실시형태에서, 방법은 본 명세서에 제공된 임의의 염기 편집기를 사용하여 수행된다. 일부 실시형태에서, 표적 윈도우는 탈아미노화 윈도우이다. 탈아미노화 윈도우는 염기 편집기가 표적 뉴클레오티드에 작용하고 탈아미노화하는 정의된 영역일 수 있다. 일부 실시형태에서, 탈아미노화 윈도우는 2, 3, 4, 5, 6, 7, 8, 9 또는 10개의 염기 영역 내에 있다. 일부 실시형태에서, 탈아미노화 윈도우는 PAM 상류의 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24 또는 25개의 염기이다.

[0791] 본 개시내용의 염기 편집기는 표적 폴리뉴클레오티드 서열의 편집을 촉진하는 임의의 도메인, 특징 또는 아미노산 서열을 포함할 수 있다. 예를 들어, 일부 실시형태에서, 염기 편집기는 핵 국소화 서열(NLS)을 포함한다. 일부 실시형태에서, 염기 편집기의 NLS는 데아미나제 도메인과 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인 사이에 국소화된다. 일부 실시형태에서, 염기 편집기의 NLS는 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인에 대해 C-말단에 국소화된다.

[0792] 본 명세서에 개시된 바와 같은 염기 편집기에 존재할 수 있는 다른 예시적인 특징은 세포질 국소화 서열과 같은 국소화 서열, 핵 유출 서열과 같은 유출 서열, 또는 다른 국소화 서열뿐만 아니라 융합 단백질의 가용화, 정제, 또는 검출에 유용한 서열 태그이다. 본 명세서에서 제공된 적합한 단백질 태그는 비오틴 카르복실라제 담체 단백질(BCCP) 태그, myc-태그, 칼모듈린-태그, FLAG-태그, 헤마글루티닌(HA)-태그, 또한 히스티딘 태그 또는 His-태그로도 지칭되는, 폴리히스티딘 태그, 말토스 결합 단백질(MBP)-태그, nus-태그, 글루타티온-S-트랜스퍼라제(GST)-태그, 녹색 형광 단백질(GFP)-태그, 티오레독신-태그, S-태그, Softag(예를 들어, Softag 1, Softag 3), 스트렙-태그, 비오틴 리가제 태그, FlAsH 태그, V5 태그 및 SBP-태그를 포함하지만 이에 제한되지 않는다. 추가의 적합한 서열은 당업자에게 명백할 것이다. 일부 실시형태에서, 융합 단백질은 하나 이상의 His 태그를 포함한다.

[0793] 일부 실시형태에서, 비-제한적인 예시적인 시티딘 염기 편집기(CBE)는 BE1(APOBEC1-XTEN-dCas9), BE2(APOBEC1-XTEN-dCas9-UGI), BE3(APOBEC1-XTEN-dCas9(A840H)-UGI), BE3-Gam, saBE3, saBE4-Gam, BE4, BE4-Gam, saBE4, 또는 saB4E-Gam을 포함한다. BE4는 APOBEC1-Cas9n(D10A) 링커를 32개 아미노산으로 확장하고 Cas9n-UGI 링커를

9개 아미노산으로 확장하고, UGI의 제2 카피를 단일 염기 편집기 작제물 안으로 다른 9-아미노산 링커를 갖는 작제물의 C 말단에 부가한다. 염기 편집기 saBE3 및 saBE4는 더 작은 *S. 아우레우스* Cas9n(D10A)으로 대체된 *S. 피오게네스* Cas9n(D10A)을 갖는다. BE3-Gam, saBE3-Gam, BE4-Gam 및 saBE4-Gam은 16개 아미노산 XTEN 링커를 통해 BE3, saBE3, BE4 및 saBE4의 N-말단에 융합된 Gam 단백질의 174개 잔기를 갖는다.

[0794] 일부 실시형태에서, 아데노신 염기 편집기(ABE)는 DNA에서 아데닌을 탈아미노화할 수 있다. 일부 실시형태에서, ABE는 BE3의 APOBEC1 구성요소를 천연 또는 조작된 *대장균* TadA, 인간 ADAR2, 마우스 ADA, 또는 인간 ADAT2로 대체함에 의해 생성된다. 일부 실시형태에서, ABE는 진화된 TadA 변이체를 포함한다. 일부 실시형태에서, ABE는 ABE 1.2(TadA*-XTEN-nCas9-NLS)이다. 일부 실시형태에서, TadA*는 A106V 및 D108N 돌연변이를 포함한다.

[0795] 일부 실시형태에서, ABE는 2-세대 ABE이다. 일부 실시형태에서, ABE는 TadA*(TadA*2.1)에 추가의 돌연변이 D147Y 및 E155V를 포함하는 ABE2.1이다. 일부 실시형태에서, ABE는 인간 알킬 아데닌 DNA 글리코실라제(E125Q 돌연변이를 갖는 AAG)의 촉매적으로 불활성화된 버전에 융합된 ABE2.2, ABE2.1이다. 일부 실시형태에서, ABE는 *대장균* Endo V(D35A 돌연변이로 불활성화됨)의 촉매적으로 불활성화된 버전에 융합된 ABE2.3, ABE2.1이다. 일부 실시형태에서, ABE는 ABE2.1에서의 링커보다 2배 더 긴(32개 아미노산, $(SGGS)_2$(서열번호: 1425)-XTEN-$(SGGS)_2$(서열번호: 1425)) 링커를 갖는 ABE2.6이다. 일부 실시형태에서, ABE는 ABE2.7이며, 이는 추가의 야생형 TadA 단량체로 묶인 ABE2.1이다. 일부 실시형태에서, ABE는 ABE2.8이며, 이는 추가의 TadA*2.1 단량체로 묶인 ABE2.1이다. 일부 실시형태에서, ABE는 ABE2.9이며, 이는 ABE2.1의 N-말단에 대한 진화된 TadA(TadA*2.1)의 직접적인 융합이다. 일부 실시형태에서, ABE는 ABE2.10이며, 이는 ABE2.1의 N-말단에 대한 야생형 TadA의 직접적인 융합이다. 일부 실시형태에서, ABE는 ABE2.11이며, 이는 TadA* 단량체의 N-말단에서 불활성화 E59A 돌연변이를 갖는 ABE2.9이다. 일부 실시형태에서, ABE는 ABE2.12이며, 이는 내부 TadA* 단량체에서 불활성화 E59A 돌연변이를 갖는 ABE2.9이다.

[0796] 일부 실시형태에서, ABE는 3세대 ABE이다. 일부 실시형태에서, ABE는 ABE3.1이며, 이는 3개의 추가의 TadA 돌연변이(L84F, H123Y 및 I156F)를 갖는 ABE2.3이다.

[0797] 일부 실시형태에서, ABE는 4세대 ABE이다. 일부 실시형태에서, ABE는 ABE4.3이며, 이는 추가의 TadA 돌연변이 A142N(TadA*4.3)을 갖는 ABE3.1이다.

[0798] 일부 실시형태에서, ABE는 5세대 ABE이다. 일부 실시형태에서, ABE는 ABE5.1이며, 이는 생존하는 클론(H36L, R51L, S146C 및 K157N)으로부터 돌연변이의 컨센서스 세트를 ABE3.1 안으로 이입함에 의해 생성된다. 일부 실시형태에서, ABE는 ABE5.3이며, 이는 내부 진화된 TadA*에 융합된 야생형 *대장균* TadA를 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE는 하기 표 10에 나타난 바와 같은, ABE5.2, ABE5.4, ABE5.5, ABE5.6, ABE5.7, ABE5.8, ABE5.9, ABE5.10, ABE5.11, ABE5.12, ABE5.13, 또는 ABE5.14이다. 일부 실시형태에서, ABE는 6세대 ABE이다. 일부 실시형태에서, ABE는 하기 표 10에 나타난 바와 같이 ABE6.1, ABE6.2, ABE6.3, ABE6.4, ABE6.5, 또는 ABE6.6이다. 일부 실시형태에서, ABE는 7세대 ABE이다. 일부 실시형태에서, ABE는 하기 표 10에 나타난 바와 같이 ABE7.1, ABE7.2, ABE7.3, ABE7.4, ABE7.5, ABE7.6, ABE7.7, ABE7.8, ABE 7.9, 또는 ABE7.10이다.

[0799] **표 10. ABE의 유전자형**

	23	26	36	37	48	49	51	72	84	87	106	108	123	125	142	146	147	152	155	156	157	161
ABE0.1	W	R	H	N	P		R	N	L	S	A	D	H	G	A	S	D	R	E	I	K	K
ABE0.2	W	R	H	N	P		R	N	L	S	A	D	H	G	A	S	D	R	E	I	K	K
ABE1.1	W	R	H	N	P		R	N	L	S	A	N	H	G	A	S	D	R	E	I	K	K
ABE1.2	W	R	H	N	P		R	N	L	S	V	N	H	G	A	S	D	R	E	I	K	K
ABE2.1	W	R	H	N	P		R	N	L	S	V	N	H	G	A	S	Y	R	V	I	K	K
ABE2.2	W	R	H	N	P		R	N	L	S	V	N	H	G	A	S	Y	R	V	I	K	K
ABE2.3	W	R	H	N	P		R	N	L	S	V	N	H	G	A	S	Y	R	V	I	K	K
ABE2.4	W	R	H	N	P		R	N	L	S	V	N	H	G	A	S	Y	R	V	I	K	K
ABE2.5	W	R	H	N	P		R	N	L	S	V	N	H	G	A	S	Y	R	V	I	K	K
ABE2.6	W	R	H	N	P		R	N	L	S	V	N	H	G	A	S	Y	R	V	I	K	K
ABE2.7	W	R	H	N	P		R	N	L	S	V	N	H	G	A	S	Y	R	V	I	K	K
ABE2.8	W	R	H	N	P		R	N	L	S	V	N	H	G	A	S	Y	R	V	I	K	K
ABE2.9	W	R	H	N	P		R	N	L	S	V	N	H	G	A	S	Y	R	V	I	K	K
ABE2.10	W	R	H	N	P		R	N	L	S	V	N	H	G	A	S	Y	R	V	I	K	K
ABE2.11	W	R	H	N	P		R	N	L	S	V	N	H	G	A	S	Y	R	V	I	K	K
ABE2.12	W	R	H	N	P		R	N	L	S	V	N	H	G	A	S	Y	R	V	I	K	K
ABE3.1	W	R	H	N	P		R	N	F	S	V	N	Y	G	A	S	Y	R	V	F	K	K
ABE3.2	W	R	H	N	P		R	N	F	S	V	N	Y	G	A	S	Y	R	V	F	K	K
ABE3.3	W	R	H	N	P		R	N	F	S	V	N	Y	G	A	S	Y	R	V	F	K	K
ABE3.4	W	R	H	N	P		R	N	F	S	V	N	Y	G	A	S	Y	R	V	F	K	K
ABE3.5	W	R	H	N	P		R	N	F	S	V	N	Y	G	A	S	Y	R	V	F	K	K
ABE3.6	W	R	H	N	P		R	N	F	S	V	N	Y	G	A	S	Y	R	V	F	K	K
ABE3.7	W	R	H	N	P		R	N	F	S	V	N	Y	G	A	S	Y	R	V	F	K	K
ABE3.8	W	R	H	N	P		R	N	F	S	V	N	Y	G	A	S	Y	R	V	F	K	K
ABE4.1	W	R	H	N	P		R	N	L	S	V	N	H	G	N	S	Y	R	V	I	K	K
ABE4.2	W	G	H	N	P		R	N	L	S	V	N	H	G	N	S	Y	R	V	I	K	K
ABE4.3	W	R	H	N	P		R	N	F	S	V	N	Y	G	N	S	Y	R	V	F	K	K
ABE5.1	W	R	L	N	P		L	N	F	S	V	N	Y	G	A	C	Y	R	V	F	N	K
ABE5.2	W	R	H	S	P		R	N	F	S	V	N	Y	G	A	S	Y	R	V	F	K	T
ABE5.3	W	R	L	N	P		L	N	I	S	V	N	Y	G	A	C	Y	R	V	F	N	K
ABE5.4	W	R	H	S	P		R	N	F	S	V	N	Y	G	A	S	Y	R	V	F	K	T
ABE5.5	W	R	L	N	P		L	N	F	S	V	N	Y	G	A	C	Y	R	V	F	N	K
ABE5.6	W	R	L	N	P		L	N	F	S	V	N	Y	G	A	C	Y	R	V	F	N	K
ABE5.7	W	R	L	N	P		L	N	F	S	V	N	Y	G	A	C	Y	R	V	F	N	K
ABE5.8	W	R	L	N	P		L	N	F	S	V	N	Y	G	A	C	Y	R	V	F	N	K
ABE5.9	W	R	L	N	P		L	N	F	S	V	N	Y	G	A	C	Y	R	V	F	N	K
ABE5.10	W	R	L	N	P		L	N	F	S	V	N	Y	G	A	C	Y	R	V	F	N	K
ABE5.11	W	R	L	N	P		L	N	F	S	V	N	Y	G	A	C	Y	R	V	F	N	K
ABE5.12	W	R	L	N	P		L	N	F	S	V	N	Y	G	A	C	Y	R	V	F	N	K
ABE5.13	W	R	H	N	P		L	D	F	S	V	N	Y	A	A	S	Y	R	V	F	K	K
ABE5.14	W	R	H	N	S		L	N	F	C	V	N	Y	G	A	S	Y	R	V	F	K	K
ABE6.1	W	R	H	N	S		L	N	F	S	V	N	Y	G	N	S	Y	R	V	F	K	K
ABE6.2	W	R	H	N	T	V	L	N	F	S	V	N	Y	G	N	S	Y	R	V	F	N	K
ABE6.3	W	R	L	N	S		L	N	F	S	V	N	Y	G	A	C	Y	R	V	F	N	K
ABE6.4	W	R	L	N	S		L	N	F	S	V	N	Y	G	N	C	Y	R	V	F	N	K
ABE6.5	W	R	L	N	T	V	L	N	F	S	V	N	Y	G	A	C	Y	R	V	F	N	K
ABE6.6	W	R	L	N	T	V	L	N	F	S	V	N	Y	G	N	C	Y	R	V	F	N	K

[0800]

	23	26	36	37	48	49	51	72	84	87	106	108	123	125	142	146	147	152	155	156	157	161
ABE7.1	W	R	L	N	A		L	N	F	S	V	N	Y	G	A	C	Y	R	V	F	N	K
ABE7.2	W	R	L	N	A		L	N	F	S	V	N	Y	G	N	C	Y	R	V	F	N	K
ABE7.3	L	R	L	N	A		L	N	F	S	V	N	Y	G	A	C	Y	R	V	F	N	K
ABE7.4	R	R	L	N	A		L	N	F	S	V	N	Y	G	A	C	Y	R	V	F	N	K
ABE7.5	W	R	L	N	A		L	N	F	S	V	N	Y	G	A	C	Y	H	V	F	N	K
ABE7.6	W	R	L	N	A		L	N	I	S	V	N	Y	G	A	C	Y	P	V	F	N	K
ABE7.7	L	R	L	N	A		L	N	F	S	V	N	Y	G	A	C	Y	P	V	F	N	K
ABE7.8	L	R	L	N	A		L	N	F	S	V	N	Y	G	N	C	Y	R	V	F	N	K
ABE7.9	L	R	L	N	A		L	N	F	S	V	N	Y	G	N	C	Y	P	V	F	N	K
ABE7.10	R	R	L	N	A		L	N	F	S	V	N	Y	G	A	C	Y	P	V	F	N	K

[0801]

[0802] 일부 실시형태에서, 염기 편집기는 8세대 ABE(ABE8)이다. 일부 실시형태에서, ABE8은 TadA*8 변이체를 함유한다. 일부 실시형태에서, ABE8은 TadA*8 변이체("ABE8.x-m")를 함유하는 단량체성 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.1-m이며, 이는 Y147T 돌연변이(TadA*8.1)를 갖는 TadA*7.10을 함유하는 단량체성 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.2-m이며, 이는 Y147R 돌연변이(TadA*8.2)를 갖는 TadA*7.10을 함유하는 단량체성 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.3-m이며, 이는 Q154S 돌연변이(TadA*8.3)를 갖는 TadA*7.10을 함유하는 단량체성 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.4-m이며, 이는 Y123H 돌연변이(TadA*8.4)를 갖는 TadA*7.10을 함유하는 단량체성 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.5-m이며, 이는 V82S 돌연변이(TadA*8.5)를 갖는 TadA*7.10을 함유하는 단량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.6-m이며, 이는 T166R 돌연변이(TadA*8.6)를 갖는 TadA*7.10을 함유하는 단량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.7-m이며, 이는 Q154R 돌연변이(TadA*8.7)를 갖는 TadA*7.10을 함유하는 단량체성 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.8-m이며, 이는 Y147R, Q154R 및 Y123H 돌연변이(TadA*8.8)를 갖는 TadA*7.10을 함유하는 단량체성 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.9-m이며, 이는 Y147R, Q154R 및 I76Y 돌연변이(TadA*8.9)를 갖는 TadA*7.10을 함유하는 단량

체성 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.10-m이며, 이는 Y147R, Q154R 및 T166R 돌연변이(TadA*8.10)를 갖는 TadA*7.10을 함유하는 단량체성 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.11-m이며, 이는 Y147T 및 Q154R 돌연변이(TadA*8.11)를 갖는 TadA*7.10을 함유하는 단량체성 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.12-m이며, 이는 Y147T 및 Q154S 돌연변이(TadA*8.12)를 갖는 TadA*7.10을 함유하는 단량체 작제물을 갖는다.

[0803] 일부 실시형태에서, ABE8은 ABE8.13-m이며, 이는 Y123H(H123Y로부터 복귀된 Y123H), Y147R, Q154R 및 I76Y 돌연변이(TadA*8.13)를 갖는 TadA*7.10을 함유하는 단량체성 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.14-m이며, 이는 I76Y 및 V82S 돌연변이(TadA*8.14)를 갖는 TadA*7.10을 함유하는 단량체성 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.15-m이며, 이는 V82S 및 Y147R 돌연변이(TadA*8.15)를 갖는 TadA*7.10을 함유하는 단량체성 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.16-m이며, 이는 V82S, Y123H(H123Y로부터 복귀된 Y123H) 및 Y147R 돌연변이(TadA*8.16)를 갖는 TadA*7.10을 함유하는 단량체성 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.17-m이며, 이는 V82S 및 Q154R 돌연변이(TadA*8.17)를 갖는 TadA*7.10을 함유하는 단량체성 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.18-m이며, 이는 V82S, Y123H(H123Y에서 복귀된 Y123H) 및 Q154R 돌연변이(TadA*8.18)를 갖는 TadA*7.10을 함유하는 단량체성 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.19-m이며, 이는 V82S, Y123H(H123Y로부터 복귀된 Y123H), Y147R 및 Q154R 돌연변이(TadA*8.19)를 갖는 TadA*7.10을 함유하는 단량체성 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.20-m이며, 이는 I76Y, V82S, Y123H(H123Y로부터 복귀된 Y123H), Y147R 및 Q154R 돌연변이(TadA*8.20)를 갖는 TadA*7.10을 함유하는 단량체성 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.21-m이며, 이는 Y147R 및 Q154S 돌연변이(TadA*8.21)를 갖는 TadA*7.10을 함유하는 단량체성 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.22-m이며, 이는 V82S 및 Q154S 돌연변이(TadA*8.22)를 갖는 TadA*7.10을 함유하는 단량체성 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.23-m이며, 이는 V82S 및 Y123H(H123Y로부터 복귀된 Y123H) 돌연변이(TadA*8.23)를 갖는 TadA*7.10을 함유하는 단량체성 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.24-m이며, 이는 V82S, Y123H(H123Y로부터 복귀된 Y123H) 및 Y147T 돌연변이(TadA*8.24)를 갖는 TadA*7.10을 함유하는 단량체성 작제물을 갖는다.

[0804] 일부 실시형태에서, ABE8은 TadA*8 변이체("ABE8.x-d")에 융합된 야생형 *대장균* TadA를 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.1-d이며, 이는 Y147T 돌연변이(TadA*8.1)를 갖는 TadA*7.10에 융합된 야생형 *대장균* TadA를 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.2-d이며, 이는 Y147R 돌연변이(TadA*8.2)를 갖는 TadA*7.10에 융합된 야생형 *대장균* TadA를 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.3-d이며, 이는 Q154S 돌연변이(TadA*8.3)를 갖는 TadA*7.10에 융합된 야생형 *대장균* TadA를 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.4-d이며, 이는 Y123H 돌연변이(TadA*8.4)를 갖는 TadA*7.10에 융합된 야생형 *대장균* TadA를 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.5-d이며, 이는 V82S 돌연변이(TadA*8.5)를 갖는 TadA*7.10에 융합된 야생형 *대장균* TadA를 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.6-d이며, 이는 T166R 돌연변이(TadA*8.6)를 갖는 TadA*7.10에 융합된 야생형 *대장균* TadA를 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.7-d이며, 이는 Q154R 돌연변이(TadA*8.7)를 갖는 TadA*7.10에 융합된 야생형 *대장균* TadA를 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.8-d이며, 이는 Y147R, Q154R 및 Y123H 돌연변이(TadA*8.8)를 갖는 TadA*7.10에 융합된 야생형 *대장균* TadA를 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.9-d이며, 이는 Y147R, Q154R 및 I76Y 돌연변이(TadA*8.9)를 갖는 TadA*7.10에 융합된 야생형 *대장균* TadA를 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.10-d이며, 이는 Y147R, Q154R 및 T166R 돌연변이(TadA*8.10)를 갖는 TadA*7.10에 융합된 야생형 *대장균* TadA를 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.11-d이며, 이는 Y147T 및 Q154R 돌연변이(TadA*8.11)를 갖는 TadA*7.10에 융합된 야생형 *대장균* TadA를 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.12-d이며, 이는 Y147T 및 Q154S 돌연변이(TadA*8.12)를 갖는 TadA*7.10에 융합된 야생형 *대장균* TadA를 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.13-d이며, 이는 Y123H(H123Y로부터 복귀된 Y123H), Y147R, Q154R 및 I76Y 돌연변이(TadA*8.13)를 갖는 TadA*7.10에 융합된 야생형 *대장균* TadA를 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.14-d이며, 이는 I76Y 및 V82S 돌연변이(TadA*8.14)를 갖는 TadA*7.10에 융합된 야생형 *대장균* TadA를 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.15-d이며, 이는 V82S 및 Y147R 돌연변이(TadA*8.15)를 갖는 TadA*7.10에 융합된 야생형 *대장균* TadA를 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.16-d이며, 이는 V82S, Y123H(H123Y에서 복귀된 Y123H) 및 Y147R 돌연변이

(TadA*8.16)를 갖는 TadA*7.10에 융합된 야생형 *대장균* TadA를 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.17-d이며, 이는 V82S 및 Q154R 돌연변이(TadA*8.17)를 갖는 TadA*7.10에 융합된 야생형 *대장균* TadA를 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.18-d이며, 이는 V82S, Y123H(H123Y에서 복귀된 Y123H) 및 Q154R 돌연변이(TadA*8.18)를 갖는 TadA*7.10에 융합된 야생형 *대장균* TadA를 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.19-d이며, 이는 V82S, Y123H(H123Y로부터 복귀된 Y123H), Y147R 및 Q154R 돌연변이(TadA*8.19)를 갖는 TadA*7.10에 융합된 야생형 *대장균* TadA를 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.20-d이며, 이는 I76Y, V82S, Y123H(H123Y로부터 복귀된 Y123H), Y147R 및 Q154R 돌연변이(TadA*8.20)를 갖는 TadA*7.10에 융합된 야생형 *대장균* TadA를 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.21-d이며, 이는 Y147R 및 Q154S 돌연변이(TadA*8.21)를 갖는 TadA*7.10에 융합된 야생형 *대장균* TadA를 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.22-d이며, 이는 V82S 및 Q154S 돌연변이(TadA*8.22)를 갖는 TadA*7.10에 융합된 야생형 *대장균* TadA를 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.23-d이며, 이는 V82S 및 Y123H(H123Y로부터 복귀된 Y123H) 돌연변이(TadA*8.23)를 갖는 TadA*7.10에 융합된 야생형 *대장균* TadA를 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.24-d이며, 이는 V82S, Y123H(H123Y로부터 복귀된 Y123H) 및 Y147T 돌연변이(TadA*8.24)를 갖는 TadA*7.10에 융합된 야생형 *대장균* TadA를 함유하는 이종이량체 작제물을 갖는다.

[0805] 일부 실시형태에서, ABE8은 TadA*8 변이체("ABE8.x-7")에 융합된 TadA*7.10을 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.1-7이며, 이는 Y147T 돌연변이(TadA*8.1)를 갖는 TadA*7.10에 융합된 TadA*7.10을 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.2-7이며, 이는 Y147R 돌연변이(TadA*8.2)를 갖는 TadA*7.10에 융합된 TadA*7.10을 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.3-7이며, 이는 Q154S 돌연변이(TadA*8.3)를 갖는 TadA*7.10에 융합된 TadA*7.10을 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.4-7이며, 이는 Y123H 돌연변이(TadA*8.4)를 갖는 TadA*7.10에 융합된 TadA*7.10을 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.5-7이며, 이는 V82S 돌연변이(TadA*8.5)를 갖는 TadA*7.10에 융합된 TadA*7.10을 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.6-7이며, 이는 T166R 돌연변이(TadA*8.6)를 갖는 TadA*7.10에 융합된 TadA*7.10을 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.7-7이며, 이는 Q154R 돌연변이(TadA*8.7)를 갖는 TadA*7.10에 융합된 TadA*7.10을 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.8-7이며, 이는 Y147R, Q154R 및 Y123H 돌연변이(TadA*8.8)를 갖는 TadA*7.10에 융합된 TadA*7.10을 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.9-7이며, 이는 Y147R, Q154R 및 I76Y 돌연변이(TadA*8.9)를 갖는 TadA*7.10에 융합된 TadA*7.10을 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.10-7이며, 이는 Y147R, Q154R 및 T166R 돌연변이(TadA*8.10)를 갖는 TadA*7.10에 융합된 TadA*7.10을 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.11-7이며, 이는 Y147T 및 Q154R 돌연변이(TadA*8.11)를 갖는 TadA*7.10에 융합된 TadA*7.10을 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.12-7이며, 이는 Y147T 및 Q154S 돌연변이(TadA*8.12)를 갖는 TadA*7.10에 융합된 TadA*7.10을 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.13-7이며, 이는 Y123H(H123Y로부터 복귀된 Y123H), Y147R, Q154R 및 I76Y 돌연변이(TadA*8.13)를 갖는 TadA*7.10에 융합된 TadA*7.10을 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.14-7이며, 이는 I76Y 및 V82S 돌연변이(TadA*8.14)를 갖는 TadA*7.10에 융합된 TadA*7.10을 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.15-7이며, 이는 V82S 및 Y147R 돌연변이(TadA*8.15)를 갖는 TadA*7.10에 융합된 TadA*7.10을 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.16-7이며, 이는 V82S, Y123H(H123Y로부터 복귀된 Y123H) 및 Y147R 돌연변이(TadA*8.16)를 갖는 TadA*7.10에 융합된 TadA*7.10을 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.17-7이며, 이는 V82S 및 Q154R 돌연변이(TadA*8.17)를 갖는 TadA*7.10에 융합된 TadA*7.10을 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.18-7이며, 이는 V82S, Y123H(H123Y로부터 복귀된 Y123H) 및 Q154R 돌연변이(TadA*8.18)를 갖는 TadA*7.10에 융합된 TadA*7.10을 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.19-7이며, 이는 V82S, Y123H(H123Y로부터 복귀된 Y123H), Y147R 및 Q154R 돌연변이(TadA*8.19)를 갖는 TadA*7.10에 융합된 TadA*7.10을 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.20-7이며, 이는 I76Y, V82S, Y123H(H123Y로부터 복귀된 Y123H), Y147R 및 Q154R 돌연변이(TadA*8.20)를 갖는 TadA*7.10에 융합된 TadA*7.10을 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.21-7이며, 이는 Y147R 및 Q154S 돌연변이(TadA*8.21)를 갖는 TadA*7.10에 융합된

TadA*7.10을 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.22-7이며, 이는 V82S 및 Q154S 돌연변이(TadA*8.22)를 갖는 TadA*7.10에 융합된 TadA*7.10을 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.23-7이며, 이는 V82S 및 Y123H(H123Y로부터 복귀된 Y123H) 돌연변이(TadA*8.23)를 갖는 TadA*7.10에 융합된 TadA*7.10을 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8.24-7이며, 이는 V82S, Y123H(H123Y로부터 복귀된 Y123H) 및 Y147T 돌연변이(TadA*8.24)를 갖는 TadA*7.10에 융합된 TadA*7.10을 함유하는 이종이량체 작제물을 갖는다.

[0806] 일부 실시형태에서, ABE는 아래 표 11에 나타난 바와 같은 ABE8.1-m, ABE8.2-m, ABE8.3-m, ABE8.4-m, ABE8.5-m, ABE8.6-m, ABE8.7-m, ABE8.8-m, ABE8.9-m, ABE8.10-m, ABE8.11-m, ABE8.12-m, ABE8.13-m, ABE8.14-m, ABE8.15-m, ABE8.16-m, ABE8.17-m, ABE8.18-m, ABE8.19-m, ABE8.20-m, ABE8.21-m, ABE8.22-m, ABE8.23-m, ABE8.24-m, ABE8.1-d, ABE8.2-d, ABE8.3-d, ABE8.4-d, ABE8.5-d, ABE8.6-d, ABE8.7-d, ABE8.8-d, ABE8.9-d, ABE8.10-d, ABE8.11-d, ABE8.12-d, ABE8.13-d, ABE8.14-d, ABE8.15-d, ABE8.16-d, ABE8.17-d, ABE8.18-d, ABE8.19-d, ABE8.20-d, ABE8.21-d, ABE8.22-d, ABE8.23-d, 또는 ABE8.24-d이다.

[0807] **표 11: 아데노신 데아미나제 염기 편집기 8(ABE8) 변이체**

ABE8	아데노신 데아미나제	아데노신 데아미나제 설명
ABE8.1-m	TadA*8.1	단량체_TadA*7.10 + Y147T
ABE8.2-m	TadA*8.2	단량체_TadA*7.10 + Y147R
ABE8.3-m	TadA*8.3	단량체_TadA*7.10 + Q154S
ABE8.4-m	TadA*8.4	단량체_TadA*7.10 + Y123H
ABE8.5-m	TadA*8.5	단량체_TadA*7.10 + V82S
ABE8.6-m	TadA*8.6	단량체_TadA*7.10 + T166R
ABE8.7-m	TadA*8.7	단량체_TadA*7.10 + Q154R
ABE8.8-m	TadA*8.8	단량체_TadA*7.10 + Y147R_Q154R_Y123H
ABE8.9-m	TadA*8.9	단량체_TadA*7.10 + Y147R_Q154R_I76Y
ABE8.10-m	TadA*8.10	단량체_TadA*7.10 + Y147R_Q154R_T166R
ABE8.11-m	TadA*8.11	단량체_TadA*7.10 + Y147T_Q154R
ABE8.12-m	TadA*8.12	단량체_TadA*7.10 + Y147T_Q154S
ABE8.13-m	TadA*8.13	단량체_TadA*7.10 + Y123H_Y147R_Q154R_I76Y
ABE8.14-m	TadA*8.14	단량체_TadA*7.10 + I76Y_V82S
ABE8.15-m	TadA*8.15	단량체_TadA*7.10 + V82S_Y147R
ABE8.16-m	TadA*8.16	단량체_TadA*7.10 + V82S_Y123H_Y147R
ABE8.17-m	TadA*8.17	단량체_TadA*7.10 + V82S_Q154R
ABE8.18-m	TadA*8.18	단량체_TadA*7.10 + V82S_Y123H_Q154R
ABE8.19-m	TadA*8.19	단량체_TadA*7.10 + V82S_Y123H_Y147R_Q154R
ABE8.20-m	TadA*8.20	단량체_TadA*7.10 + I76Y_V82S_Y123H_Y147R_Q154R
ABE8.21-m	TadA*8.21	단량체_TadA*7.10 + Y147R_Q154S
ABE8.22-m	TadA*8.22	단량체_TadA*7.10 + V82S_Q154S
ABE8.23-m	TadA*8.23	단량체_TadA*7.10 + V82S_Y123H
ABE8.24-m	TadA*8.24	단량체_TadA*7.10 + V82S_Y123H_Y147T
ABE8.1-d	TadA*8.1	이종이량체_(WT) + (TadA*7.10 + Y147T)
ABE8.2-d	TadA*8.2	이종이량체_(WT) + (TadA*7.10 + Y147R)
ABE8.3-d	TadA*8.3	이종이량체_(WT) + (TadA*7.10 + Q154S)
ABE8.4-d	TadA*8.4	이종이량체_(WT) + (TadA*7.10 + Y123H)
ABE8.5-d	TadA*8.5	이종이량체_(WT) + (TadA*7.10 + V82S)
ABE8.6-d	TadA*8.6	이종이량체_(WT) + (TadA*7.10 + T166R)

[0808]

ABE8.7-d	TadA*8.7	이종이량체_(WT) + (TadA*7.10 + Q154R)
ABE8.8-d	TadA*8.8	이종이량체_(WT) + (TadA*7.10 + Y147R_Q154R_Y123H)
ABE8.9-d	TadA*8.9	이종이량체_(WT) + (TadA*7.10 + Y147R_Q154R_I76Y)
ABE8.10-d	TadA*8.10	이종이량체_(WT) + (TadA*7.10 + Y147R_Q154R_T166R)
ABE8.11-d	TadA*8.11	이종이량체_(WT) + (TadA*7.10 + Y147T_Q154R)
ABE8.12-d	TadA*8.12	이종이량체_(WT) + (TadA*7.10 + Y147T_Q154S)
ABE8.13-d	TadA*8.13	이종이량체_(WT) + (TadA*7.10 + Y123H_Y147T_Q154R_I76Y)
ABE8.14-d	TadA*8.14	이종이량체_(WT) + (TadA*7.10 + I76Y_V82S)
ABE8.15-d	TadA*8.15	이종이량체_(WT) + (TadA*7.10 + V82S_ Y147R)
ABE8.16-d	TadA*8.16	이종이량체_(WT) + (TadA*7.10 + V82S_Y123H_Y147R)
ABE8.17-d	TadA*8.17	이종이량체_(WT) + (TadA*7.10 + V82S_Q154R)
ABE8.18-d	TadA*8.18	이종이량체_(WT) + (TadA*7.10 + V82S_Y123H_Q154R)
ABE8.19-d	TadA*8.19	이종이량체_(WT) + (TadA*7.10 + V82S_Y123H_Y147R_Q154R)
ABE8.20-d	TadA*8.20	이종이량체_(WT) + (TadA*7.10 + I76Y_V82S_Y123H_Y147R_Q154R)
ABE8.21-d	TadA*8.21	이종이량체_(WT) + (TadA*7.10 + Y147R_Q154S)
ABE8.22-d	TadA*8.22	이종이량체_(WT) + (TadA*7.10 + V82S_Q154S)
ABE8.23-d	TadA*8.23	이종이량체_(WT) + (TadA*7.10 + V82S_Y123H)
ABE8.24-d	TadA*8.24	이종이량체_(WT) + (TadA*7.10 + V82S_Y123H_Y147T)

[0809]

[0810] 일부 실시형태에서, ABE8은 ABE8a-m이며, 이는 R26C, A109S, T111R, D119N, H122N, Y147D, F149Y, T166I 및 D167N 돌연변이(TadA*8a)를 갖는 TadA*7.10을 함유하는 단량체성 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8b-m이며, 이는 V88A, A109S, T111R, D119N, H122N, F149Y, T166I 및 D167N 돌연변이(TadA*8b)를 갖는 TadA*7.10을 함유하는 단량체성 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8c-m이며, 이는 R26C, A109S, T111R, D119N, H122N, F149Y, T166I 및 D167N 돌연변이(TadA*8c)를 갖는 TadA*7.10을 함유하는 단량체성 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8d-m이며, 이는 V88A, T111R, D119N 및 F149Y 돌연변이(TadA*8d)를 갖는 TadA*7.10을 함유하는 단량체성 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8e-m이며, 이는 A109S, T111R, D119N, H122N, Y147D, F149Y, T166I 및 D167N 돌연변이(TadA*8e)를 갖는 TadA*7.10을 함유하는 단량체성 작제물을 갖는다.

[0811] 일부 실시형태에서, ABE8은 ABE8a-d이며, 이는 R26C, A109S, T111R, D119, H122N, Y147D, F149Y, T166I 및 D167N 돌연변이(TadA*8a)를 갖는 TadA*7.10에 융합된 야생형 *대장균* TadA를 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8b-d이며, 이는 V88A, A109S, T111R, D119N, H122N, F149Y, T166I 및 D167N 돌연변이(TadA*8b)를 갖는 TadA*7.10에 융합된 야생형 *대장균* TadA를 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8c-d이며, 이는 R26C, A109S, T111R, D119N, H122N, F149Y, T166I 및 D167N 돌연변이(TadA*8c)를 갖는 TadA*7.10에 융합된 야생형 *대장균* TadA를 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8d-d이며, 이는 V88A, T111R, D119N 및 F149Y 돌연변이(TadA*8d)를 갖는 TadA*7.10에 융합된 야생형 *대장균* TadA를 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8e-d이며, 이는 A109S, T111R, D119N, H122N, Y147D, F149Y, T166I 및 D167N 돌연변이(TadA*8e)를 갖는 TadA*7.10에 융합된 야생형 *대장균* TadA를 함유하는 이종이량체 작제물을 갖는다.

[0812] 일부 실시형태에서, ABE8은 ABE8a-7이며, 이는 R26C, A109S, T111R, D119, H122N, Y147D, F149Y, T166I 및 D167N 돌연변이(TadA*8a)를 갖는 TadA*7.10에 융합된 TadA*7.10을 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8b-7이며, 이는 V88A, A109S, T111R, D119N, H122N, F149Y, T166I 및 D167N 돌연변이(TadA*8b)를 갖는 TadA*7.10에 융합된 TadA*7.10을 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8c-7이며, 이는 R26C, A109S, T111R, D119N, H122N, F149Y, T166I 및 D167N 돌연변이(TadA*8c)를 갖는 TadA*7.10에 융합된 TadA*7.10을 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8d-7

이며, 이는 V88A, T111R, D119N 및 F149Y 돌연변이(TadA*8d)를 갖는 TadA*7.10에 융합된 TadA*7.10을 함유하는 이종이량체 작제물을 갖는다. 일부 실시형태에서, ABE8은 ABE8e-7이며, 이는 A109S, T111R, D119N, H122N, Y147D, F149Y, T166I 및 D167N 돌연변이(TadA*8e)를 갖는 TadA*7.10에 융합된 TadA*7.10을 함유하는 이종이량체 작제물을 갖는다.

[0813] 일부 실시형태에서, ABE는 하기 표 12에 나타난 바와 같은 ABE8a-m, ABE8b-m, ABE8c-m, ABE8d-m, ABE8e-m, ABE8a-d, ABE8b-d, ABE8c-d, ABE8d-d 또는 ABE8e-d이다. 일부 실시형태에서, ABE는 ABE8e-m 또는 ABE8e-d이다. ABE8e는 SpCas9 이외의 Cas 동족체, 예를 들어 SaCas9, SaCas9-KKH, Cas12a 동족체, *예를 들어* LbCas12a, enAs-Cas12a, SpCas9-NG 및 원형으로 바꾸어진 CP1028-SpCas9 및 CP1041-SpCas9와 함께 사용될 때 효율적인 아데닌 염기 편집 활성과 낮은 삽입결실 형성을 나타낸다. 표 12에서 ABE8e에 대해 도시된 돌연변이에 부가하여, TadA 도메인 안으로 V106W 치환을 도입함에 의해 표적-외 RNA 및 DNA 편집이 감소되었다(그 전체 내용이 본 명세서에 참조로 포함된, M. Richter 등, 2020, Nature Biotechnology, doi.org/10.1038/s41587-020-0453-z에 기술된 바와 같음).

[0814] **표 12: 추가의 아데노신 데아미나제 염기 편집기 8 변이체. 표에서, "단량체"는 표시된 변경을 포함하는 단일 TadA*7.10을 포함하는 ABE를 나타내고 "이종이량체"는 *대장균* TadA 아데노신 데아미나제에 융합된 표시된 변경을 포함하는 TadA*7.10을 포함하는 ABE를 나타낸다.**

ABE8 염기 편집기	아데노신 데아미나제	아데노신 데아미나제 설명
ABE8a-m	TadA*8a	단량체_TadA*7.10 + R26C + A109S + T111R + D119N + H122N + Y147D + F149Y + T166I + D167N
ABE8b-m	TadA*8b	단량체_TadA*7.10 + V88A + A109S + T111R + D119N + H122N + F149Y + T166I + D167N
ABE8c-m	TadA*8c	단량체_TadA*7.10 + R26C + A109S + T111R + D119N + H122N + F149Y + T166I + D167N
ABE8d-m	TadA*8d	단량체_TadA*7.10 + V88A + T111R + D119N + F149Y
ABE8e-m	TadA*8e	단량체_TadA*7.10 + A109S + T111R + D119N + H122N + Y147D + F149Y + T166I + D167N
ABE8a-d	TadA*8a	이종이량체_(WT) + (TadA*7.10 + R26C + A109S + T111R + D119N + H122N + Y147D + F149Y + T166I + D167N)
ABE8b-d	TadA*8b	이종이량체_(WT) + (TadA*7.10 + V88A + A109S + T111R + D119N + H122N + F149Y + T166I + D167N)
ABE8c-d	TadA*8c	이종이량체_(WT) + (TadA*7.10 + R26C + A109S + T111R + D119N + H122N + F149Y + T166I + D167N)
ABE8d-d	TadA*8d	이종이량체_(WT) + (TadA*7.10 + V88A + T111R + D119N + F149Y)
ABE8e-d	TadA*8e	이종이량체_(WT) + (TadA*7.10 + A109S + T111R + D119N + H122N + Y147D + F149Y + T166I + D167N)

[0815]

[0816] 일부 실시형태에서, 염기 편집기(*예를 들어*, ABE8)는 원형 순열체 Cas9(*예를 들어*, CP5 또는 CP6) 및 이분 핵국소화 서열을 포함하는 스캐폴드 안으로 아데노신 데아미나제 변이체(*예를 들어*, TadA*8)를 클로닝함에 의해 생성된다. 일부 실시형태에서, 염기 편집기(*예를 들어*, ABE7.9, ABE7.10 또는 ABE8)는 NGC PAM CP5 변이체(*S. 피오게네스* Cas9 또는 spVRQR Cas9)이다. 일부 실시형태에서, 염기 편집기(*예를 들어*, ABE7.9, ABE7.10 또는 ABE8)는 AGA PAM CP5 변이체(*S. 피오게네스* Cas9 또는 spVRQR Cas9)이다. 일부 실시형태에서, 염기 편집기(*예를 들어*, ABE7.9, ABE7.10 또는 ABE8)는 NGC PAM CP6 변이체(*S. 피오게네스* Cas9 또는 spVRQR Cas9)이다. 일부 실시형태에서, 염기 편집기(*예를 들어* ABE7.9, ABE7.10 또는 ABE8)는 AGA PAM CP6 변이체(*S. 피오게네스* Cas9 또는 spVRQR Cas9)이다.

[0817] 일부 실시형태에서, ABE는 하기 표 13에 나타난 바와 같은 유전자형을 갖는다.

[0818] **표 13. ABE의 유전자형**

	23	26	36	37	48	49	51	72	84	87	105	108	123	125	142	145	147	152	155	156	157	161
ABE7.9	L	R	L	N	A		L	N	F	S	V	N	Y	G	N	C	Y	P	V	F	N	K
ABE7.10	R	R	L	N	A		L	N	F	S	V	N	Y	G	A	C	Y	P	V	F	N	K

[0819]

[0820] 하기 표 14에 나타난 바와 같이, 40개 ABE8의 유전자형이 기재되어 있다. ABE의 진화된 *대장균* TadA 부분에서의

잔기 위치가 표시되어 있다. ABE8에서의 돌연변이 변화는 ABE7.10 돌연변이와 구별되는 경우 도시된다. 일부 실시형태에서, ABE는 하기 표 14에 나타난 바와 같은 ABE 중 하나의 유전자형을 갖는다.

[0821] **표 14. 진화된 TadA에서의 잔기 동일성**

	23	36	48	51	76	82	84	106	108	123	146	147	152	154	155	156	157	166
ABE7.10	R	L	A	L	I	V	F	V	N	Y	C	Y	P	Q	V	F	N	T
ABE8.1-m												T						
ABE8.2-m												R						
ABE8.3-m														S				
ABE8.4-m										H								
ABE8.5-m						S												
ABE8.6-m																		R
ABE8.7-m														R				
ABE8.8-m										H		R		R				

[0822]

ABE8.9-m					Y							R		R				
ABE8.10-m												R		R				R
ABE8.11-m												T		R				
ABE8.12-m												T		S				
ABE8.13-m					Y					H		R		R				
ABE8.14-m					Y	S												
ABE8.15-m						S						R						
ABE8.16-m						S				H		R						
ABE8.17-m						S								R				
ABE8.18-m						S				H				R				
ABE8.19-m						S				H		R		R				
ABE8.20-m					Y	S				H		R		R				
ABE8.21-m												R		S				
ABE8.22-m						S								S				
ABE8.23-m						S				H								
ABE8.24-m						S				H		T						
ABE8.1-d												T						
ABE8.2-d												R						
ABE8.3-d														S				
ABE8.4-d										H								
ABE8.5-d						S												
ABE8.6-d																		R
ABE8.7-d														R				
ABE8.8-d										H		R		R				
ABE8.9-d					Y							R		R				
ABE8.10-d												R		R				R
ABE8.11-d												T		R				
ABE8.12-d												T		S				
ABE8.13-d					Y					H		R		R				

[0823]

ABE8.14-d					Y	S											
ABE8.15-d						S						R					
ABE8.16-d						S				H		R					
ABE8.17-d						S								R			
ABE8.18-d						S				H				R			
ABE8.19-d						S				H		R		R			
ABE8.20-d					Y	S				H		R		R			
ABE8.21-d												R		S			
ABE8.22-d						S								S			
ABE8.23-d						S				H							
ABE8.24-d						S				H		T					

[0824]

[0825] 일부 실시형태에서, 염기 편집기는 ABE8.1이며, 이는 아데노신 데아미나제 활성을 갖는 다음 서열 또는 이의 단편을 포함하거나 본질적으로 이로 구성된다:

[0826] ABE8.1_Y147T_CP5_NGC PAM_단량체

[0827] MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRV VFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCTFFRMPRQVFNAQKKAQSSTD*SGGSSGGSSGSETPGTSESATPESSGGSSGGS*EIGKA TAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGF MQPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAKFLQKGNELALPSKYVNFL YLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPRAFKYFDTTIARKEYRS TKEVLDATLIHQSITGLYETRIDLSQLGGD*GGSGGSGGSGGSGGSGGSGGSGGM*DKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALL FDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLI YLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNF DLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNG YAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNS RFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVK QLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSR KLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARE NQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSD NVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDF QFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEGADKRTADGSEFESPKKKRKV(서열번호: 1426)

[0828] 상기 서열에서, 단순한 텍스트는 아데노신 데아미나제 서열을 나타내고, 굵은 서열은 Cas9로부터 유래된 서열을 나타내고, 이탤릭체 서열은 링커 서열을 나타내고, 밑줄 친 서열은 이분 핵 국소화 서열을 나타낸다. 다른 ABE8 서열은 첨부된 서열 목록(서열번호: 1427-1449)에 제공되어 있다.

[0829] 일부 실시형태에서, 염기 편집기는 9세대 ABE(ABE9)이다. 일부 실시형태에서, ABE9는 TadA*9 변이체를 함유한다. ABE9 염기 편집기는 본 명세서에 기재된 바와 같은 ABE 7*10 참조 서열에 대한 변경을 함유하는 아미노산 서열을 포함하는 아데노신 데아미나제 변이체를 포함한다. 예시적인 ABE9 변이체는 표 15에 나열되어 있다. ABE9 염기 편집기의 세부사항은 국제 PCT 출원 번호 PCT/2020/049975에 기재되어 있으며, 이는 그 전체가 본 명세서에 참고로 포함된다.

[0830] **표 15. 아데노신 데아미나제 염기 편집기 9(ABE9) 변이체. 표에서, "단량체"는 표시된 변경을 포함하는 단일 TadA*7.10을 포함하는 ABE를 나타내고 "이종이량체"는 *대장균* TadA 아데노신 데아미나제에 융합된 표시된 변경을 포함하는 TadA*7.10을 포함하는 ABE를 나타낸다.**

ABE9 설명	변경
ABE9.1_단량체	E25F, V82S, Y123H, T133K, Y147R, Q154R
ABE9.2_단량체	E25F, V82S, Y123H, Y147R, Q154R
ABE9.3 _단량체	V82S, Y123H, P124W, Y147R, Q154R
ABE9.4_단량체	L51W, V82S, Y123H, C146R, Y147R, Q154R
ABE9.5_단량체	P54C, V82S, Y123H, Y147R, Q154R
ABE9.6_단량체	Y73S, V82S, Y123H, Y147R, Q154R
ABE9.7_단량체	N38G, V82T, Y123H, Y147R, Q154R
ABE9.8_단량체	R23H, V82S, Y123H, Y147R, Q154R
ABE9.9_단량체	R21N, V82S, Y123H, Y147R, Q154R
ABE9.10_단량체	V82S, Y123H, Y147R, Q154R, A158K
ABE9.11_단량체	N72K, V82S, Y123H, D139L, Y147R, Q154R,
ABE9.12_단량체	E25F, V82S, Y123H, D139M, Y147R, Q154R
ABE9.13_단량체	M70V, V82S, M94V, Y123H, Y147R, Q154R
ABE9.14_단량체	Q71M, V82S, Y123H, Y147R, Q154R

[0831]

ABE9.15_이종이량체	E25F, V82S, Y123H, T133K, Y147R, Q154R
ABE9.16_이종이량체	E25F, V82S, Y123H, Y147R, Q154R
ABE9.17_이종이량체	V82S, Y123H, P124W, Y147R, Q154R
ABE9.18_이종이량체	L51W, V82S, Y123H, C146R, Y147R, Q154R
ABE9.19_이종이량체	P54C, V82S, Y123H, Y147R, Q154R
ABE9.2 _이종이량체	Y73S, V82S, Y123H, Y147R, Q154R
ABE9.21_이종이량체	N38G, V82T, Y123H, Y147R, Q154R
ABE9.22_이종이량체	R23H, V82S, Y123H, Y147R, Q154R
ABE9.23_이종이량체	R21N, V82S, Y123H, Y147R, Q154R
ABE9.24_이종이량체	V82S, Y123H, Y147R, Q154R, A158K
ABE9.25_이종이량체	N72K, V82S, Y123H, D139L, Y147R, Q154R,
ABE9.26_이종이량체	E25F, V82S, Y123H, D139M, Y147R, Q154R
ABE9.27 _이종이량체	M70V, V82S, M94V, Y123H, Y147R, Q154R
ABE9.28_이종이량체	Q71M, V82S, Y123H, Y147R, Q154R
ABE9.29_단량체	E25F_I76Y_V82S_Y123H_Y147R_Q154R
ABE9.30_단량체	I76Y_V82T_Y123H_Y147R_Q154R
ABE9.31_단량체	N38G_I76Y_V82S_Y123H_Y147R_Q154R
ABE9.32_단량체	N38G_I76Y_V82T_Y123H_Y147R_Q154R
ABE9.33_단량체	R23H_I76Y_V82S_Y123H_Y147R_Q154R
ABE9.34_단량체	P54C_I76Y_V82S_Y123H_Y147R_Q154R
ABE9.35_단량체	R21N_I76Y_V82S_Y123H_Y147R_Q154R
ABE9.36_단량체	I76Y_V82S_Y123H_D138M_Y147R_Q154R
ABE9.37_단량체	Y72S_I76Y_V82S_Y123H_Y147R_Q154R
ABE9.38_이종이량체	E25F_I76Y_V82S_Y123H_Y147R_Q154R
ABE9.39_이종이량체	I76Y_V82T_Y123H_Y147R_Q154R
ABE9.40_이종이량체	N38G_I76Y_V82S_Y123H_Y147R_Q154R
ABE9.41_이종이량체	N38G_I76Y_V82T_Y123H_Y147R_Q154R
ABE9.42_이종이량체	R23H_I76Y_V82S_Y123H_Y147R_Q154R
ABE9.43_이종이량체	P54C_I76Y_V82S_Y123H_Y147R_Q154R
ABE9.44_이종이량체	R21N_I76Y_V82S_Y123H_Y147R_Q154R
ABE9.45_이종이량체	I76Y_V82S_Y123H_D138M_Y147R_Q154R
ABE9.46_이종이량체	Y72S_I76Y_V82S_Y123H_Y147R_Q154R
ABE9.47_단량체	N72K_V82S, Y123H, Y147R, Q154R

[0832]

ABE9.48_단량체	Q71M_V82S, Y123H, Y147R, Q154R
ABE9.49_단량체	M70V,V82S, M94V, Y123H, Y147R, Q154R
ABE9.50_단량체	V82S, Y123H, T133K, Y147R, Q154R
ABE9.51_단량체	V82S, Y123H, T133K, Y147R, Q154R, A158K
ABE9.52_단량체	M70V,Q71M,N72K,V82S, Y123H, Y147R, Q154R
ABE9.53_이종이량체	N72K_V82S, Y123H, Y147R, Q154R
ABE9.54_이종이량체	Q71M_V82S, Y123H, Y147R, Q154R
ABE9.55_이종이량체	M70V,V82S, M94V, Y123H, Y147R, Q154R
ABE9.56_이종이량체	V82S, Y123H, T133K, Y147R, Q154R
ABE9.57_이종이량체	V82S, Y123H, T133K, Y147R, Q154R, A158K
ABE9.58_이종이량체	M70V, Q71M, N72K, V82S, Y123H, Y147R, Q154R

[0833]

[0834] 일부 실시형태에서, 염기 편집기는 우라실 글리코실라제 억제제(UGI)의 전부 또는 일부를 포함하는 도메인을 포함한다. 일부 실시형태에서, 염기 편집기는 핵산 폴리머라제의 전부 또는 일부를 포함하는 도메인을 포함한다. 일부 실시형태에서, 염기 편집기는 핵산 폴리머라제(NAP)의 전부 또는 일부를 도메인으로 포함할 수 있다. 예를 들어, 염기 편집기는 진핵생물 NAP의 전부 또는 일부를 포함할 수 있다. 일부 실시형태에서, 염기 편집기 안으로 합체된 NAP 또는 이의 일부는 DNA 폴리머라제이다. 일부 실시형태에서, 염기 편집기 안으로 합체된 NAP 또는 이의 일부는 트랜스레시온 폴리머라제 활성을 갖는다. 일부 실시형태에서, 염기 편집기 안으로 합체된 NAP 또는 이의 일부는 트랜스레시온 DNA 폴리머라제이다. 일부 실시형태에서, 염기 편집기 안으로 합체된 NAP 또는 이의 일부는 Rev7, Rev1 복합체, 폴리머라제 이오타, 폴리머라제 카파 또는 폴리머라제 에타이다. 일부 실시형태에서, 염기 편집기 안으로 합체된 NAP 또는 이의 일부는 진핵생물 폴리머라제 알파, 베타, 감마, 델타, 엡실론, 감마, 에타, 이오타, 카파, 람다, 뮤 또는 누 구성요소이다. 일부 실시형태에서, 염기 편집기 안으로 합체된 NAP 또는 이의 일부는 핵산 폴리머라제(*예를 들어*, 트랜스레시온 DNA 폴리머라제)와 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 99.5% 동일한 아미노산 서열을 포함한다. 일부 실시형태에서, 염기 편집기 안으로 합체된 핵산 폴리머라제 또는 이의 일부는 트랜스레시온 DNA 폴리머라제이다.

[0835] 일부 실시형태에서, 염기 편집기의 도메인은 다중 도메인을 포함할 수 있다. 예를 들어, Cas9로부터 유래된 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인을 포함하는 염기 편집기는 야생형 또는 천연 Cas9의 REC 로브 및 NUC 로브에 상응하는 REC 로브 및 NUC 로브를 포함할 수 있다. 다른 예에서, 염기 편집기는 RuvCI 도메인, BH 도메인, REC1 도메인, REC2 도메인, RuvCII 도메인, L1 도메인, HNH 도메인, L2 도메인, RuvCIII 도메인, WED 도메인, TOPO 도메인 또는 CTD 도메인 중 하나 이상을 포함할 수 있다. 일부 실시형태에서, 염기 편집기의 하나 이상의 도메인은 도메인을 포함하는 폴리펩티드의 야생형 버전에 비해 돌연변이(*예를 들어*, 치환, 삽입, 결실)를 포함한다. 예를 들어, 폴리뉴클레오티드 프로그램가능 DNA 결합 도메인의 HNH 도메인은 H840A 치환을 포함할 수 있다. 또 다른 예에서, 폴리뉴클레오티드 프로그램가능 DNA 결합 도메인의 RuvCI 도메인은 D10A 치환을 포함할 수 있다.

[0836] 본 명세서에 개시된 염기 편집기의 상이한 도메인(*예를 들어*, 인접한 도메인)은 하나 이상의 링커 도메인(*예를 들어*, XTEN 링커 도메인)을 사용하거나 사용하지 않고 서로 연결될 수 있다. 일부 실시형태에서, 링커 도메인은 결합(*예를 들어*, 공유 결합), 화학 기, 또는 2개의 분자 또는 모이어티, *예를 들어* 융합 단백질의 2개의 도메인, 예컨대 예를 들어 제1 도메인(*예를 들어*, Cas9-유래된 도메인) 및 제2 도메인(*예를 들어*, 아데노신 데아미나제 도메인 또는 시티딘 데아미나제 도메인)을 연결하는 분자일 수 있다. 일부 실시형태에서, 링커는 공유 결합(*예를 들어*, 탄소-탄소 결합, 이황화 결합, 탄소-헤테로 원자 결합 등)이다. 특정 실시형태에서, 링커는 아미드 연결의 탄소 질소 결합이다. 특정 실시형태에서, 링커는 환형 또는 비환형, 치환 또는 비치환, 분지형 또는 비분지형 지방족 또는 헤테로지방족 링커이다. 특정 실시형태에서, 링커는 중합체성(*예를 들어,* 폴리에틸렌, 폴리에틸렌 글리콜, 폴리아미드, 폴리에스테르 등)이다. 특정 실시형태에서, 링커는 아미노알칸산의 단량체, 이량체 또는 중합체를 포함한다. 일부 실시형태에서, 링커는 아미노알칸산(*예를 들어*, 글리신, 에탄산, 알라닌, 베타-알라닌, 3-아미노프로판산, 4-아미노부탄산, 5-펜탄산 등)을 포함한다. 일부 실시형태에서, 링커는 아미노헥산산(Ahx)의 단량체, 이량체 또는 중합체를 포함한다. 특정 실시형태에서, 링커는 탄소환식 모이어티(*예를 들어*, 시클로펜탄, 시클로헥산)를 기반으로 한다. 다른 실시형태에서, 링커는 폴리에틸렌 글리콜 모이어티(PEG)를 포함한다. 특정 실시형태에서, 링커는 아릴 또는 헤테로아릴 모이어티를 포함한다. 특정 실시형태에서, 링커는 페닐 고리를 기반으로 한다. 링커는 펩티드로부터 링커로의 친핵체(*예를 들어*, 티올, 아미노)의 부착을 촉진하기 위해 기능화된 모이어티를 포함할 수 있다. 임의의 친전자체는 링커의 일부로 사용될 수 있다. 예시적인 친전자체는 활성화된 에스테르, 활성화된 아미드, 마이클 수용체, 알킬 할라이드, 아릴 할라이드, 아실 할라이드 및 이소티오시아네이트를 포함하지만 이에 제한되지 않는다. 일부 실시형태에서, 링커는 Cas9 뉴클레아제 도메인을 포함하는 RNA-프로그램가능 뉴클레아제의 gRNA 결합 도메인과 핵산 편집 단백질의 촉매적 도메인을 연결시킨다. 일부 실시형태에서, 링커는 dCas9와 제2 도메인(*예를 들어*, UGI 등)을 연결시킨다.

[0837] ***링커***

[0838] 특정 실시형태에서, 링커는 발명의 임의의 펩티드 또는 펩티드 도메인을 연결하는데 사용될 수 있다. 링커는 공유 결합만큼 단순할 수 있거나, 길이에서 많은 원자를 갖는 중합체성 링커일 수 있다. 특정 실시형태에서, 링커는 폴리펩티드이거나 아미노산을 기반으로 한다. 다른 실시형태에서, 링커는 펩티드-유사가 아니다. 특정 실시형태에서, 링커는 공유 결합(*예를 들어*, 탄소-탄소 결합, 이황화 결합, 탄소-헤테로원자 결합 등)이다. 특정 실시형태에서, 링커는 아미드 연결의 탄소-질소 결합이다. 특정 실시형태에서, 링커는 환형 또는 비환형, 치환 또는 비치환, 분지형 또는 비분지형 지방족 또는 헤테로지방족 링커이다. 특정 실시형태에서, 링커는 중합체성(*예*

를 들어, 폴리에틸렌, 폴리에틸렌 글리콜, 폴리아미드, 폴리에스테르 *등*)이다. 특정 실시형태에서, 링커는 아미노알칸산의 단량체, 이량체 또는 중합체를 포함한다. 특정 실시형태에서, 링커는 아미노알칸산(*예를 들어*, 글리신, 에탄산, 알라닌, 베타-알라닌, 3-아미노프로판산, 4-아미노부탄산, 5-펜탄산 *등*)을 포함한다. 특정 실시형태에서, 링커는 아미노헥산산(Ahx)의 단량체, 이량체 또는 중합체를 포함한다. 특정 실시형태에서, 링커는 탄소환식 모이어티(*예를 들어*, 시클로펜탄, 시클로헥산)를 기반으로 한다. 다른 실시형태에서, 링커는 폴리에틸렌 글리콜 모이어티(PEG)를 포함한다. 다른 실시형태에서, 링커는 아미노산을 포함한다. 특정 실시형태에서, 링커는 펩티드를 포함한다. 특정 실시형태에서, 링커는 아릴 또는 헤테로아릴 모이어티를 포함한다. 특정 실시형태에서, 링커는 페닐 고리를 기반으로 한다. 링커는 펩티드로부터 링커로의 친핵체(*예를 들어*, 티올, 아미노)의 부착을 촉진하기 위해 기능화된 모이어티를 포함할 수 있다. 임의의 친전자체가 링커의 일부로 사용될 수 있다. 예시적인 친전자체는 활성화된 에스테르, 활성화된 아미드, 마이클 수용체, 알킬 할라이드, 아릴 할라이드, 아실 할라이드 및 이소티오시아네이트를 포함하지만 이에 제한되지 않는다.

[0839] 전형적으로, 링커는 2개의 기, 분자 또는 다른 모이어티 사이에 위치되거나 이에 의해 측접되고, 공유 결합을 통해 각각에 연결되어 둘을 연결한다. 일부 실시형태에서, 링커는 아미노산 또는 복수의 아미노산(*예를 들어*, 펩티드 또는 단백질)이다. 일부 실시형태에서, 링커는 유기 분자, 기, 중합체 또는 화학적 모이어티이다. 일부 실시형태에서, 링커는 길이가 2-100개 아미노산, 예를 들어 길이가 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 30-35, 35-40, 40-45, 45-50, 50-60, 60-70, 70-80, 80-90, 90-100, 100-150, 또는 150-200개 아미노산이다. 일부 실시형태에서, 링커는 길이가 약 3 내지 약 104(*예를 들어*, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95 또는 100)개 아미노산 이다. 더 길거나 더 짧은 링커도 고려된다.

[0840] 일부 실시형태에서, 본 명세서에 제공된 임의의 융합 단백질은 링커를 통해 서로 융합된 시티딘 또는 아데노신 데아미나제 및 Cas9 도메인을 포함한다. 시티딘 또는 아데노신 데아미나제와 Cas9 도메인 사이의 다양한 링커 길이 및 유연성이 시티딘 또는 아데노신 데아미나제 핵염기 편집기에 대한 활성에 대해 최적의 길이를 달성하기 위해 이용될 수 있다(예를 들어, 형태 (GGGS)n(서열번호: 1308), (GGGGS)n(서열번호: 109), 및 (G)n의 매우 유연한 링커로부터 형태 (EAAAK)n(서열번호: 1309), (SGGS)n(서열번호: 57), SGSETPGTSESATPES(서열번호: 56)(예를 들어, Guilinger JP, 등 Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. Nat. Biotechnol. 2014; 32(6): 577-82를 참조하며; 그 전체 내용은 본 명세서에 참고로 포함됨) 및 (XP)n의 보다 단단한 링커에 이르기까지 다양함). 일부 실시형태에서, n은 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 또는 15이다. 일부 실시형태에서, 링커는 (GGS)n 모티프를 포함하며, 여기서 n은 1, 3 또는 7이다. 일부 실시형태에서, 본 명세서에 제공된 임의의 융합 단백질의 Cas9 도메인과 시티딘 데아미나제 또는 아데노신 데아미나제는, 또한 XTEN 링커로도 지칭될 수 있는 아미노산 서열 SGSETPGTSESATPES(서열번호: 56)를 포함하는 링커를 통해 융합된다. 일부 실시형태에서, 링커는 복수의 프롤린 잔기를 포함하고 길이가 5-21, 5-14, 5-9, 5-7 아미노산, 예를 들어 PAPAP(서열번호: 65), PAPAPA(서열번호: 66), PAPAPAP(서열번호 67), PAPAPAPA(서열번호: 68), P(AP)4(서열번호: 69), P(AP)7(서열번호: 70), P(AP)10(서열번호: 71)이다(예를 들어, Tan J, Zhang F, Karcher D, Bock R. Engineering of high-precision base editors for site-specific single nucleotide replacement. Nat Commun. 2019 Jan 25;10(1):439를 참조하며; 전체 내용은 본 명세서에 참조로 포함된다). 이러한 프롤린-풍부 링커는 "단단한" 링커로도 명명된다.

[0841] 다른 실시형태에서, 염기 편집기 시스템은 데아미나제(DNA 데아미나제), 예를 들어 아데노신 또는 시티딘 데아미나제와 비-공유적으로 상호작용하고 일시적으로 아데노신 또는 시티딘 데아미나제를 최소의 또는 감소된 방관자 또는 표적-인접 효과로, 특정 편집을 위한 표적 폴리뉴클레오티드 서열에서의 표적 핵염기로 유인하는 구성요소(단백질)를 포함한다. 데아미나제-상호작용 단백질을 포함하는 이러한 비-공유 시스템 및 방법은 DNA 데아미나제를 특정 게놈 표적 핵염기에 유인하고 표적-상 및 표적-인접 편집의 이벤트를 디커플링하여 보다 정확한 단일 염기 치환 돌연변이의 달성을 증강시키는 역할을 한다. 실시형태에서, 데아미나제-상호작용 단백질은 데아미나제의 활성(촉매적) 부위가 표적 핵염기(예를 들어, 각각 아데노신 또는 시티딘)와 계합하는 것을 차단하거나 방해함에 없이 데아미나제(예를 들어, 아데노신 데아미나제 또는 시티딘 데아미나제)에 결합한다. "MagnEdit"로 명명되는 이러한 시스템은 Cas9 및 gRNA 복합체에 묶인 상호작용하는 단백질을 포함하고 공동-발현된 아데노신 또는 시티딘 데아미나제(외인성 또는 내인성)를 유인하여 특정 게놈 표적 부위를 편집할 수 있고, McCann, J. 등, 2020, "MagnEdit - interacting factors that recruit DNA-editing enzymes to single base targets," Life-Science-Alliance, Vol. 3, No. 4 (e201900606), (doi 10.26508/lsa.201900606)에 기술되어 있으며, 그 내용은 그 전체가 본 명세서에 참조로 포함된다. 실시형태에서, DNA 데아미나제는 본 명세서에

기재된 바와 같은 아데노신 데아미나제 변이체(예를 들어, TadA*8)이다.

[0842] 또 다른 실시형태에서, "Suntag"로 불리는 시스템은 염기 편집기의 단백질(예를 들어, 아데노신 데아미나제 또는 시티딘 데아미나제) 구성요소 또는 이의 다중 카피를 폴리뉴클레오티드 표적 부위로 모집하여 감소된 인접한 표적 편집을 갖는 부위에서 염기 편집을 달성하는데 사용되는 비-공유적으로 상호작용하는 구성요소를 포함하며, 예를 들어 Tanenbaum, M.E. 등, "A protein tagging system for signal amplification in gene expression and fluorescence imaging," Cell. 2014 October 23; 159(3): 635-646. doi:10.1016/j.cell.2014.09.039; 및 Huang, Y.-H. 등, 2017, "DNA epigenome editing using CRISPR-Cas SunTag-directed DNMT3A," Genome Biol 18: 176. doi:10.1186/s13059-017-1306-z에 기재된 바와 같으며, 그 각각의 내용은 본 명세서에 그 전체가 참조로 포함된다. 실시형태에서, DNA 데아미나제는 본 명세서에 기재된 바와 같은 아데노신 데아미나제 변이체(예를 들어, TadA*8)이다.

[0843] ***가이드 RNA를 갖는 핵산 프로그램가능 DNA 결합 단백질***

[0844] 세포에서의 염기 편집을 위한 조성물 및 방법이 본 명세서에 제공된다. 본 명세서에 제공된 바와 같은 가이드 다중핵산 서열, 예를 들어 가이드 RNA 서열, 또는 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 이상의 가이드 RNA 서열의 조합을 포함하는 조성물이 본 명세서에 추가로 제공된다. 일부 실시형태에서, 본 명세서에 제공된 바와 같은 염기 편집을 위한 조성물은 염기 편집기, 예를 들어 C-염기 편집기 또는 A-염기 편집기를 인코딩하는 폴리뉴클레오티드를 추가로 포함한다. 예를 들어, 염기 편집을 위한 조성물은 BE, BE4, ABE, 및 제공된 바와 같은 하나 이상의 가이드 RNA의 조합을 인코딩하는 mRNA 서열을 포함할 수 있다. 염기 편집을 위한 조성물은 염기 편집기 폴리펩티드 및 본 명세서에 제공된 임의의 가이드 RNA 중 하나 이상의 조합을 포함할 수 있다. 이러한 조성물은 상이한 전달 접근법, 예를 들어, 전기천공, 뉴클레오펙션, 바이러스 형질도입 또는 형질감염을 통해 세포에서의 염기 편집을 수행하는데 사용될 수 있다. 일부 실시형태에서, 염기 편집을 위한 조성물은 염기 편집기를 인코딩하는 mRNA 서열 및 전기천공을 위해 본 명세서에 제공된 하나 이상의 가이드 RNA 서열의 조합을 포함한다.

[0845] 본 개시내용의 일부 양태는 본 명세서에 제공된 임의의 융합 단백질, 및 융합 단백질의 핵산 프로그램가능 DNA 결합 단백질(napDNAbp) 도메인(*예를 들어*, Cas9(*예를 들어*, dCas9, 뉴클레아제 활성 Cas9, 또는 Cas9 닉카제) 또는 Cas12)에 결합된 가이드 RNA를 포함하는 복합체를 제공한다. 이들 복합체는 리보핵단백질(RNP)로도 명명된다. 일부 실시형태에서, 가이드 핵산(*예를 들어*, 가이드 RNA)은 15-100개의 뉴클레오티드 길이이고 표적 서열에 상보성인 적어도 10개의 연속 뉴클레오티드의 서열을 포함한다. 일부 실시형태에서, 가이드 RNA는 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 또는 50개의 뉴클레오티드 길이이다. 일부 실시형태에서, 가이드 RNA는 표적 서열에 상보성인 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 또는 40개의 연속 뉴클레오티드의 서열을 포함한다. 일부 실시형태에서, 표적 서열은 DNA 서열이다. 일부 실시형태에서, 표적 서열은 RNA 서열이다. 일부 실시형태에서, 표적 서열은 박테리아, 효모, 진균, 곤충, 식물 또는 동물의 게놈에서의 서열이다. 일부 실시형태에서, 표적 서열은 인간의 게놈에서의 서열이다. 일부 실시형태에서, 표적 서열의 3' 말단은 규범적 PAM 서열(NGG)에 바로 인접해 있다. 일부 실시형태에서, 표적 서열의 3' 말단은 비-규범적 PAM 서열(*예를 들어*, 표 6에 나열된 서열 또는 5'-NAA-3')에 바로 인접해 있다. 일부 실시형태에서, 가이드 핵산(*예를 들어*, 가이드 RNA)은 관심있는 유전자(*예를 들어*, 질환 또는 장애와 연관된 유전자)에서의 서열에 상보성이다.

[0846] 본 개시내용의 일부 양태는 본 명세서에 제공된 융합 단백질 또는 복합체를 사용하는 방법을 제공한다. 예를 들어, 본 개시내용의 일부 양태는 DNA 분자를 본 명세서에 제공된 임의의 융합 단백질 및 적어도 하나의 가이드 RNA와 접촉시키는 것을 포함하는 방법을 제공하며, 여기서 가이드 RNA는 약 15-100개의 뉴클레오티드 길이이고 표적 서열에 상보성인 적어도 10개 연속 뉴클레오티드의 서열을 포함한다. 일부 실시형태에서, 표적 서열의 3' 말단은 AGC, GAG, TTT, GTG 또는 CAA 서열에 바로 인접해 있다. 일부 실시형태에서, 표적 서열의 3' 말단은 NGA, NGCG, NGN, NNGRRT, NNNRRT, NGCG, NGCN, NGTN, NGTN, NGTN, 또는 5'(TTTV) 서열에 바로 인접해 있다. 일부 실시형태에서, 표적 서열의 3' 말단은 *예를 들어*, TTN, DTTN, GTTN, ATTN, ATTC, DTTNT, WTTN, HATY, TTTN, TTTV, TTTC, TG, RTR 또는 YTN PAM 부위에 바로 인접해 있다.

[0847] 각 서열에서 특정 위치 또는 잔기의 번호지정은 사용된 특정 단백질 및 번호 지정 계획에 의존한다는 것이 이해될 것이다. 번호지정은, 예를 들어, 성숙한 단백질의 전구체와 성숙한 단백질 자체에서 다를 수 있고, 종마다 서열에서의 차이가 번호지정에 영향을 미칠 수 있다. 당업자는 당업계에 널리 공지된 방법, 예를 들어 서열 정

렬 및 상동체 잔기의 결정에 의해 임의의 상동체 단백질 및 각각의 인코딩 핵산에서의 각각의 잔기를 식별할 수 있을 것이다.

[0848] 본 명세서에 개시된 임의의 융합 단백질을 표적 부위, 예를 들어 편집되어 지는 돌연변이를 포함하는 부위로 표적화하기 위해, 전형적으로 가이드 RNA와 함께 융합 단백질을 공동-발현하는 것이 필요하다는 것이 당업자에게 명백할 것이다. 본 명세서의 다른 곳에서 더 자세히 설명한 바와 같이, 가이드 RNA는 전형적으로 napDNAbp(예를 들어, Cas9 또는 Cas12) 결합을 허용하는 tracrRNA 프레임워크, 및 napDNAbp:핵산 편집 효소/도메인 융합 단백질에 서열 특이성을 부여하는 가이드 서열을 포함한다. 대안적으로, 가이드 RNA 및 tracrRNA는 2개의 핵산 분자로서 별도로 제공될 수 있다. 일부 실시형태에서, 가이드 RNA는 가이드 서열이 표적 서열에 상보성인 서열을 포함하는 구조를 포함한다. 가이드 서열은 전형적으로 20개 뉴클레오티드 길이이다. napDNAbp:핵산 편집 효소/도메인 융합 단백질을 특정 게놈의 표적 부위에 표적화하기 위한 적합한 가이드 RNA의 서열은 본 개시내용에 기반하여 당업자에게 명백할 것이다. 이러한 적합한 가이드 RNA 서열은 전형적으로 편집되어 지는 표적 뉴클레오티드의 상류 또는 하류 50개 뉴클레오티드 내의 핵산 서열에 상보성인 가이드 서열을 포함한다. 제공된 임의의 융합 단백질을 특정 표적 서열에 표적화하는데 적합한 일부 예시적인 가이드 RNA 서열이 본 명세서에 제공된다.

[0849] sgRNA의 별개의 부분은 Cas9(예를 들어, SpyCas9) 및/또는 DNA 표적과 상호작용하는 다양한 특징을 형성할 것으로 예측된다. Cas9 엔도뉴클레아제 활성을 지시하는 천연 crRNA:tracrRNA 이중나선 및 단일 가이드 RNA(sgRNA) 내에서 6개의 보존된 모듈이 식별되었다(Briner 등, Guide RNA Functional Modules Direct Cas9 Activity and Orthogonality Mol Cell. 2014 Oct 23;56(2):333-339 참조). 6개의 모듈은 DNA 표적화를 담당하는 스페이서, CRISPR 반복:tracrRNA 이중나선에 의해 형성된 상부 줄기, 돌출부, 하부 줄기, 넥서스 및 tracrRNA의 3' 말단으로부터 헤어핀을 포함한다. 상부 및 하부 줄기는 주로 인산염 골격과의 서열-독립적 상호작용을 통해 Cas9와 상호작용한다. 일부 실시형태에서, 상부 줄기는 불필요하다. 일부 실시형태에서, 하부 줄기의 염기에서 보존된 우라실 뉴클레오티드 서열은 불필요하다. 돌출부는 Cas9의 Rec1 도메인과의 특정한 측쇄 상호작용에 참여한다. U44의 핵염기는 Tyr 325 및 His 328의 측쇄와 상호작용하는 반면, G43은 Tyr 329와 상호작용한다. 넥서스는 sgRNA:Cas9 상호작용의 핵심을 형성하고 sgRNA와 Cas9 및 표적 DNA 둘 모두 사이의 교차점에 있다. A51과 A52의 핵염기는 Phe 1105의 측쇄와 상호작용하고; U56은 Arg 457 및 Asn 459와 상호작용하고; U59의 핵염기는 Arg 74, Asn 77, Pro 475, Leu 455, Phe 446 및 Ile 448의 측쇄에 의해 정의된 소수성 포켓 안으로 삽입되고; C60은 Leu 455, Ala 456 및 Asn 459와 상호작용하고, C61은 Arg 70의 측쇄와 상호작용하며, 이는 차례로 C15와 상호작용한다. 일부 실시형태에서, 이들 돌연변이 중 하나 이상은 sgRNA:Cas9 상호작용을 최적화하기 위해 Cas9(예를 들어, spyCas9)에 대한 sgRNA의 돌출부 및/또는 넥서스에서 이루어진다.

[0850] 더욱이, tracrRNA 넥서스와 헤어핀은 Cas9 페어링에 중요하고 서로 다른 Cas9 단백질을 분리하는 직교성 장벽을 교차하도록 교체될 수 있으며, 이는 직교 Cas9 단백질의 추가의 활용에 대해 주된 역할을 한다. 일부 실시형태에서, 넥서스와 헤어핀은 직교 Cas9 단백질을 표적화하기 위해 교체된다. 일부 실시형태에서, sgRNA에는 상부 줄기, 헤어핀 1 및/또는 하부 줄기의 서열 유연성이 분배되어 보다 컴팩트하고 구조적으로 안정한 가이드 RNA를 설계한다. 일부 실시형태에서, 모듈은 다양한 키메라 가이드가 있는 단일 Cas9를 사용하거나 키메라 sgRNA의 상이한 조합이 있는 직교 시스템을 동시적으로 사용함에 의해 다중 편집을 최적화하도록 변형된다. 가이드 기능 모듈 및 이의 방법에 관한 세부사항은 예를 들어 Briner 등, Guide RNA Functional Modules Direct Cas9 Activity and Orthogonality Mol Cell. 2014 Oct 23;56(2):333-339에 기술되어 있으며, 그 내용은 그 전체가 본 명세서에 참조로 포함된다.

[0851] 본 명세서에 개시된 염기 편집기의 도메인은 임의의 순서로 배열될 수 있다. 예를 들어, 폴리뉴클레오티드-프로그램가능 뉴클레오티드-결합 도메인(예를 들어, Cas9 또는 Cas12) 및 데아미나제 도메인(예를 들어, 시티딘 또는 아데노신 데아미나제)을 포함하는 융합 단백질을 포함하는 염기 편집기의 비-제한적 예는 다음과 같이 배열될 수 있다.

[0852] NH2-[핵염기 편집 도메인]-링커1-[핵염기 편집 도메인]-COOH;

[0853] NH2-[데아미나제]-링커1-[핵염기 편집 도메인]-COOH;

[0854] NH2-[데아미나제]-링커1-[핵염기 편집 도메인]-링커2-[UGI]-COOH;

[0855] NH2-[데아미나제]-링커1-[핵염기 편집 도메인]-COOH;

[0856] NH2-[아데노신 데아미나제]-링커1-[핵염기 편집 도메인]-COOH;

[0857] NH2-[핵염기 편집 도메인]-[데아미나제]-COOH;

[0858] NH2-[데아미나제]-[핵염기 편집 도메인]-[이노신 BER 억제제]-COOH;

[0859] NH2-[데아미나제]-[이노신 BER 억제제]-[핵염기 편집 도메인]-COOH;

[0860] NH2-[이노신 BER 억제제]-[데아미나제]-[핵염기 편집 도메인]-COOH;

[0861] NH2-[핵염기 편집 도메인]-[데아미나제]-[이노신 BER 억제제]-COOH;

[0862] NH2-[핵염기 편집 도메인]-[이노신 BER 억제제]-[데아미나제]-COOH;

[0863] NH2-[이노신 BER 억제제]-[핵염기 편집 도메인]-[데아미나제]-COOH;

[0864] NH2-[핵염기 편집 도메인]-링커1-[데아미나제]-링커2-[핵염기 편집 도메인]-COOH;

[0865] NH2-[핵염기 편집 도메인]-링커1-[데아미나제]-[핵염기 편집 도메인]-COOH;

[0866] NH2-[핵염기 편집 도메인]-[데아미나제]-링커2-[핵염기 편집 도메인]-COOH;

[0867] NH2-[핵염기 편집 도메인]-[데아미나제]-[핵염기 편집 도메인]-COOH;

[0868] NH2-[핵염기 편집 도메인]-링커1-[데아미나제]-링커2-[핵염기 편집 도메인]-[이노신 BER 억제제]-COOH;

[0869] NH2-[핵염기 편집 도메인]-링커1-[데아미나제]-[핵염기 편집 도메인]-[이노신 BER 억제제]-COOH;

[0870] NH2-[핵염기 편집 도메인]-[데아미나제]-링커2-[핵염기 편집 도메인]-[이노신 BER 억제제]-COOH;

[0871] NH2-[핵염기 편집 도메인]-[데아미나제]-[핵염기 편집 도메인]-[이노신 BER 억제제]-COOH;

[0872] NH2-[이노신 BER 억제제]-[핵염기 편집 도메인]-링커1-[데아미나제]-링커2-[핵염기 편집 도메인]-COOH;

[0873] NH2-[이노신 BER 억제제]-[핵염기 편집 도메인]-링커1-[데아미나제]-[핵염기 편집 도메인]-COOH;

[0874] NH2-[이노신 BER 억제제]-[핵염기 편집 도메인]-[데아미나제]-링커2-[핵염기 편집 도메인]-COOH; 또는

[0875] NH2-[이노신 BER 억제제]NH2-[핵염기 편집 도메인]-[데아미나제]-[핵염기 편집 도메인]-COOH.

[0876] 일부 실시형태에서, 본 명세서에 제공된 염기 편집 융합 단백질은 정확한 장소, 예를 들어 표적 염기가 정의된 영역(예를 들어, "탈아미노화 윈도우") 내에 배치되는 곳에 위치되어야 할 필요가 있다. 일부 실시형태에서, 표적은 4-염기 영역 내에 있을 수 있다. 일부 실시형태에서, 이러한 정의된 표적 영역은 PAM 상류의 대략적으로 15개 염기일 수 있다. Komor, A.C., 등, "Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage" Nature 533, 420-424 (2016); Gaudelli, N.M., 등, "Programmable base editing of A·T to G·C in genomic DNA without DNA cleavage" Nature 551, 464-471 (2017); 및 Komor, A.C., 등, "Improved base excision repair inhibition and bacteriophage Mu Gam protein yields C:G-to-T:A base editors with higher efficiency and product purity" Science Advances 3:eaao4774 (2017)를 참조하며, 그 전체 내용은 참조로 본 명세서에 포함된다.

[0877] 정의된 표적 영역은 탈아미노화 윈도우일 수 있다. 탈아미노화 윈도우는 염기 편집기가 표적 뉴클레오티드에 작용하고 탈아미노화하는 정의된 영역일 수 있다. 일부 실시형태에서, 탈아미노화 윈도우는 2, 3, 4, 5, 6, 7, 8, 9 또는 10개의 염기 영역 내에 있다. 일부 실시형태에서, 탈아미노화 윈도우는 PAM 상류에 있는 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 또는 25개 염기이다.

[0878] 본 개시내용의 염기 편집기는 표적 폴리뉴클레오티드 서열의 편집을 촉진하는 임의의 도메인, 특징 또는 아미노산 서열을 포함할 수 있다. 예를 들어, 일부 실시형태에서, 염기 편집기는 핵 국소화 서열(NLS)을 포함한다. 일부 실시형태에서, 염기 편집기의 NLS는 데아미나제 도메인과 napDNAbp 도메인 사이에 국소화된다. 일부 실시형태에서, 염기 편집기의 NLS는 napDNAbp 도메인의 C-말단에 국소화된다.

[0879] 융합 단백질에 포함될 수 있는 단백질 도메인의 비-제한적인 예는 데아미나제 도메인(예를 들어, 아데노신 데아미나제 또는 시티딘 데아미나제), 우라실 글리코실라제 억제제(UGI) 도메인, 에피토프 태그, 리포터 유전자 서열 및/또는 본 명세서에 기재된 활성 중 하나 이상을 갖는 단백질 도메인을 포함한다.

[0880] 도메인은 에피토프 태그, 리포터 단백질, 기타 결합 도메인으로 검출되거나 표지될 수 있다. 에피토프 태그의 비-제한적인 예는 히스티딘(His) 태그, V5 태그, FLAG 태그, 인플루엔자 헤마글루티닌(HA) 태그, Myc 태그,

VSV-G 태그 및 티오레독신(Trx) 태그를 포함한다. 리포터 유전자의 예는 글루타티온-5-트랜스퍼라제(GST), 양고추냉이 퍼옥시다제(HRP), 클로람페니콜 아세틸트랜스퍼라제(CAT) 베타-갈락토시다제, 베타-글루쿠로니다제, 루시퍼라제, 녹색 형광 단백질(GFP), HcRed, DsRed, 청록색 형광 단백질(CFP), 황색 형광 단백질(YFP) 및 청색 형광 단백질(BFP)을 포함한 자가형광 단백질을 포함하지만 이에 제한되지 않는다. 추가의 단백질 서열은 DNA 분자에 결합하거나 다른 세포 분자에 결합하는 아미노산 서열을 포함할 수 있으며, 여기에는 말토스 결합 단백질(MBP), S-태그, Lex A DNA 결합 도메인(DBD) 융합, GAL4 DNA 결합 도메인 융합, 및 단순 포진 바이러스(HSV) BP16 단백질 융합이 포함되지만 이에 제한되지 않는다.

[0881] ***시티딘 또는 아데노신 데아미나제 및 Cas9 도메인을 포함하는 융합 단백질을 사용하는 방법***

[0882] 본 개시내용의 일부 양태는 본 명세서에 제공된 융합 단백질 또는 복합체를 사용하는 방법을 제공한다. 예를 들어, 본 개시내용의 일부 양태는 DNA 분자를 본 명세서에 제공된 임의의 융합 단백질 및 본 명세서에 기재된 적어도 하나의 가이드 RNA와 접촉시키는 것을 포함하는 방법을 제공한다.

[0883] 일부 실시형태에서, 발명의 융합 단백질은 관심있는 표적 유전자를 편집하는데 사용된다. 특히, 본 명세서에 기재된 시티딘 데아미나제 또는 아데노신 데아미나제 핵염기 편집기는 표적 서열 내에서 다중 돌연변이를 만들 수 있다. 이들 돌연변이는 표적의 기능에 영향을 미칠 수 있다. 예를 들어, 시티딘 데아미나제 또는 아데노신 데아미나제 핵염기 편집기를 사용하여 조절 영역을 표적화하는 경우 조절 영역의 기능이 변경되고 하류 단백질의 발현이 감소되거나 제거된다.

[0884] 각 서열에서의 특정 위치 또는 잔기의 번호지정은 사용된 특정 단백질 및 번호지정 계획에 의존한다는 것이 이해될 것이다. 번호지정은, 예를 들어, 성숙 단백질의 전구체와 성숙 단백질 자체에서 다를 수 있고 종마다 서열에서의 차이가 번호지정에 영향을 줄 수 있다. 당업자는 당업계에 널리 공지된 방법, 예를 들어 서열 정렬 및 상동체 잔기의 결정에 의해 임의의 상동체 단백질 및 각각의 인코딩 핵산에서의 각각의 잔기를 식별할 수 있을 것이다.

[0885] 본 명세서에 개시된 바와 같이, Cas9 도메인 및 시티딘 또는 아데노신 데아미나제를 포함하는 임의의 융합 단백질을 표적 부위, 예를 들어 편집되어 지는 돌연변이를 포함하는 부위에 표적화하기 위해서는 가이드 RNA, 예를 들어 sgRNA가 공동-발현될 수 있다는 것이 당업자에게 명백할 것이다. 본 명세서의 다른 곳에서 더 자세히 설명된 바와 같이, 가이드 RNA는 전형적으로 Cas9 결합을 허용하는 tracrRNA 프레임워크, 및 Cas9:핵산 편집 효소/도메인 융합 단백질에 서열 특이성을 부여하는 가이드 서열을 포함한다. 대안적으로, 가이드 RNA 및 tracrRNA는 2개의 핵산 분자로서 별도로 제공될 수 있다. 일부 실시형태에서, 가이드 RNA는 가이드 서열이 표적 서열에 상보성인 서열을 포함하는 구조를 포함한다. 가이드 서열은 전형적으로 20개 뉴클레오티드 길이이다. Cas9:핵산 편집 효소/도메인 융합 단백질을 특정 게놈 표적 부위에 표적화하기 위한 적합한 가이드 RNA의 서열은 본 개시내용에 기반하여 당업자에게 명백할 것이다. 이러한 적합한 가이드 RNA 서열은 전형적으로 편집되어 지는 표적 뉴클레오티드 상류 또는 하류의 50개 뉴클레오티드 내의 핵산 서열에 상보성인 가이드 서열을 포함한다. 제공된 임의의 융합 단백질을 특정 표적 서열에 표적화하는데 적합한 일부 예시적인 가이드 RNA 서열이 본 명세서에 제공된다.

[0886] ***염기 편집기 효율성***

[0887] 일부 실시형태에서, 본 명세서에 제공된 방법의 목적은 유전자 편집을 통해 유전자 및/또는 유전자 생성물을 변경하는 것이다. 본 명세서에 제공된 핵염기 편집 단백질은 시험관 또는 생체내에서 유전자 편집-기반 인간 요법에 사용될 수 있다. 당업자는 본 명세서에 제공된 핵염기 편집 단백질, 예를 들어 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인(예를 들어, Cas9) 및 핵염기 편집 도메인(예를 들어, 아데노신 데아미나제 도메인 또는 시티딘 데아미나제 도메인)을 포함하는 융합 단백질이 A에서 G로 또는 C에서 T로 뉴클레오티드를 편집하는데 사용할 수 있다는 것이 이해될 것이다.

[0888] 유리하게도, 본 명세서에 제공된 바와 같은 염기 편집 시스템은 이중 가닥 DNA 파괴를 생성함이 없이, 공여자 DNA 주형을 요구함이 없이, CRISPR처럼 과도한 확률론적 삽입 및 결실을 유도함이 없이 게놈 편집을 제공한다. 일부 실시형태에서, 본 개시내용은 유의미한 수의 의도되지 않은 돌연변이, 예컨대 의도되지 않은 점 돌연변이를 생성함에 없이, 핵산(*예를 들어,* 대상체의 게놈 내의 핵산)에서 의도된 돌연변이, 예컨대 STOP 코돈을 효율적으로 생성하는 염기 편집기를 제공한다. 일부 실시형태에서, 의도된 돌연변이는 의도된 돌연변이를 생성하도록 특별히 설계된 가이드 폴리뉴클레오티드(*예를 들어,* gRNA)에 결합된 특정 염기 편집기(*예를 들어,* 아데노신 염기 편집기 또는 시티딘 염기 편집기)에 의해 생성되는 돌연변이이다. 일부 실시형태에서, 의도된 돌연변이는

질환 또는 장애, *예를 들어* 신경학적 또는 안과학적 질환 또는 장애와 연관된 표적 항원과 연관된 유전자에 있다. 일부 실시형태에서, 의도된 돌연변이는 질환 또는 장애, *예를 들어* 신경학적 또는 안과학적 질환 또는 장애와 연관된 표적 항원과 연관된 유전자에서의 아데닌(A)에서 구아닌(G)으로의 점 돌연변이(*예를 들어*, SNP)이다. 일부 실시형태에서, 의도된 돌연변이는 유전자의 코딩 영역 또는 비-코딩 영역(*예를 들어*, 조절 영역 또는 요소) 내의 아데닌(A)에서 구아닌(G)으로의 점 돌연변이이다. 일부 실시형태에서, 의도된 돌연변이는 질환 또는 장애, *예를 들어* 신경학적 또는 안과학적 질환 또는 장애와 연관된 표적 항원과 연관된 유전자에서의 시토신(C)에서 티민(T)으로의 점 돌연변이(*예를 들어*, SNP)이다. 일부 실시형태에서, 의도된 돌연변이는 유전자의 코딩 영역 또는 비-코딩 영역(*예를 들어*, 조절 영역 또는 요소) 내의 시토신(C)에서 티민(T)으로의 점 돌연변이이다. 일부 실시형태에서, 의도된 돌연변이는 STOP 코돈, 예를 들어 유전자의 코딩 영역 내의 조기 STOP 코돈을 생성하는 점 돌연변이이다. 일부 실시형태에서, 의도된 돌연변이는 정지 코돈을 제거하는 돌연변이이다.

[0889] 발명의 염기 편집기는 상당한 비율의 삽입결실을 생성함이 없이 단백질을 인코딩하는 특정한 뉴클레오티드 염기를 유리하게 변형시킨다. 본 명세서에 사용된 바와 같은 "삽입결실"은 핵산 내의 뉴클레오티드 염기의 삽입 또는 결실을 지칭한다. 이러한 삽입 또는 결실은 유전자의 코딩 영역 내에서 프레임 시프트 돌연변이를 유발할 수 있다. 일부 실시형태에서, 핵산에서 다수의 삽입 또는 결실(즉, 삽입결실)을 생성함이 없이, 핵산 내의 특정한 뉴클레오티드를 효율적으로 변형(예를 들어, 돌연변이)시키는 염기 편집기를 생성하는 것이 바람직하다. 일부 실시형태에서, 핵산에서 다수의 삽입 또는 결실(즉, 삽입결실)을 생성함이 없이, 핵산 내의 특정한 뉴클레오티드를 효율적으로 변형(예를 들어, 돌연변이 또는 메틸화)시키는 염기 편집기를 생성하는 것이 바람직하다. 특정 실시형태에서, 본 명세서에 제공된 임의의 염기 편집기는 삽입결실 대비 더 큰 비율의 의도된 변형(예를 들어, 메틸화)을 생성할 수 있다. 특정 실시형태에서, 본 명세서에 제공된 임의의 염기 편집기는 삽입결실 대비 더 큰 비율의 의도된 변형(예를 들어, 돌연변이)을 생성할 수 있다.

[0890] 일부 실시형태에서, 본 명세서에 제공된 염기 편집기는 1:1보다 큰 의도된 돌연변이 대 삽입결실(즉, 의도된 점 돌연변이:의도되지 않은 점 돌연변이)의 비율을 생성할 수 있다. 일부 실시형태에서, 본 명세서에 제공된 염기 편집기는 적어도 1.5:1, 적어도 2:1, 적어도 2.5:1, 적어도 3:1, 적어도 3.5:1, 적어도 4:1, 적어도 4.5:1, 적어도 5:1, 적어도 5.5:1, 적어도 6:1, 적어도 6.5:1, 적어도 7:1, 적어도 7.5:1, 적어도 8:1, 적어도 10:1, 적어도 12:1, 적어도 15:1, 적어도 20:1, 적어도 25:1, 적어도 30:1, 적어도 40:1, 적어도 50:1, 적어도 100:1, 적어도 200:1, 적어도 300:1, 적어도 400:1, 적어도 500:1, 적어도 600:1, 적어도 700:1, 적어도 800: 1, 적어도 900:1, 또는 적어도 1000:1 이상인 의도된 돌연변이 대 삽입결실의 비율을 생성할 수 있다. 의도된 돌연변이 및 삽입결실의 수는 임의의 적합한 방법을 사용하여 결정될 수 있다.

[0891] 일부 실시형태에서, 본 명세서에 제공된 염기 편집기는 핵산의 영역에서 삽입결실의 형성을 제한할 수 있다. 일부 실시형태에서, 영역은 염기 편집기에 의해 표적화된 뉴클레오티드에 있거나, 염기 편집기에 의해 표적화된 뉴클레오티드의 2, 3, 4, 5, 6, 7, 8, 9 또는 10개 뉴클레오티드 내의 영역에 있다. 일부 실시형태에서, 본 명세서에 제공된 임의의 염기 편집기는 1% 미만, 1.5% 미만, 2% 미만, 2.5% 미만, 3% 미만. 3.5% 미만, 4% 미만, 4.5% 미만, 5% 미만, 6% 미만, 7% 미만, 8% 미만, 9% 미만, 10% 미만, 12% 미만, 15% 미만, 또는 20% 미만으로 핵산의 영역에서 삽입결실의 형성을 제한할 수 있다. 핵산 영역에서 형성된 삽입결실의 수는 핵산(예를 들어, 세포의 게놈 내의 핵산)이 염기 편집기에 노출되는 시간의 양에 의존할 수 있다. 일부 실시형태에서, 삽입결실의 수 또는 비율은 염기 편집기에 핵산(예를 들어, 세포의 게놈 내의 핵산)을 노출시킨 지 적어도 1시간, 적어도 2시간, 적어도 6시간, 적어도 12시간, 적어도 24시간, 적어도 36시간, 적어도 48시간, 적어도 3일, 적어도 4일, 적어도 5일, 적어도 7일, 적어도 10일, 또는 적어도 14일 후에 결정된다.

[0892] 개시내용의 일부 양태는 본 명세서에 제공된 임의의 염기 편집기가 상당한 수의 의도되지 않은 돌연변이(예를 들어, 의사의 표적-외 편집 또는 방관자 편집)를 생성함에 없이 핵산(예를 들어 대상체의 게놈 내의 핵산)에서 의도된 돌연변이를 효율적으로 생성할 수 있다는 인식에 기반한다. 일부 실시형태에서, 의도된 돌연변이는 의도된 돌연변이를 생성하도록 특별히 설계된, gRNA에 결합된 특정한 염기 편집기에 의해 생성된 돌연변이이다. 일부 실시형태에서, 의도된 돌연변이는 정지 코돈, 예를 들어 유전자의 코딩 영역 내의 조기 정지 코돈을 생성하는 돌연변이이다. 일부 실시형태에서, 의도된 돌연변이는 정지 코돈을 제거하는 돌연변이이다. 일부 실시형태에서, 의도된 돌연변이는 유전자의 스플라이싱을 변경하는 돌연변이이다. 일부 실시형태에서, 의도된 돌연변이는 유전자(예를 들어, 유전자 프로모터 또는 유전자 억제인자)의 조절 서열을 변경하는 돌연변이이다. 일부 실시형태에서, 본 명세서에 제공된 임의의 염기 편집기는 1:1보다 큰 의도된 돌연변이 대 의도되지 않은 돌연변이(예를 들어, 의도된 돌연변이:의도되지 않은 돌연변이)의 비율을 생성할 수 있다. 일부 실시형태에서, 본 명세서에 제공된 임의의 염기 편집기는 적어도 1.5:1, 적어도 2:1, 적어도 2.5:1, 적어도 3:1, 적어도 3.5:1, 적어도

4:1, 적어도 4.5:1, 적어도 5:1, 적어도 5.5:1, 적어도 6:1, 적어도 6.5:1, 적어도 7:1, 적어도 7.5:1, 적어도 8:1, 적어도 10:1, 적어도 12:1, 적어도 15:1, 적어도 20:1, 적어도 25:1, 적어도 30:1, 적어도 40:1, 적어도 50:1, 적어도 100:1, 적어도 150:1, 적어도 200:1, 적어도 250:1, 적어도 500:1, 또는 적어도 1000:1 이상인 의도된 돌연변이 대 의도되지 않은 돌연변이의 비율을 생성할 수 있다. 본 명세서에 기재된 염기 편집기의 특징은 본 명세서에 제공된 임의의 융합 단백질, 또는 융합 단백질을 사용하는 방법에 적용될 수 있다는 것이 이해되어야 한다.

[0893] 염기 편집은 종종 유전적 변형, 유전자 변형 및 핵산 서열의 변형과 같은 "변형"으로 지칭되고, 변형이 염기 편집 변형이라는 맥락에 기반하여 명확하게 이해될 수 있다. 따라서 염기 편집 변형은, 예를 들어 개시내용 전반에 걸쳐 논의된 데아미나제 활성의 결과로서 뉴클레오티드 염기 수준에서의 변형이며, 이는 그 다음 유전자 서열에서의 변화를 초래하고 유전자 생성물에 영향을 미칠 수 있다. 따라서 본질적으로, 본 명세서에 기재된 유전자 편집 변형은 구조적으로 및/또는 기능적으로 유전자의 변형을 초래할 수 있으며, 여기서 유전자 생성물의 발현은 변형될 수 있으며, 예를 들어 유전자의 발현은 녹아웃되거나; 또는 반대로 강화되거나, 또는 어떤 상황에서는 유전자 기능이나 활성이 변형될 수 있다. 본 명세서에 개시된 방법을 사용하여, 염기 편집 효율은 염기 편집이 수행되는 유전자의 녹다운 효율로서 결정될 수 있으며, 여기서 염기 편집은 유전자의 발현을 녹다운시키도록 의도된다. 녹다운 수준은, 예를 들어, 유세포분석; 정량적 RT-PCR, 노던 블롯 분석 또는 파이로시퀀싱과 같은 임의의 기타 적합한 검정과 같은 RNA 발현을 검출하기 위한 검정에 의한 단백질 발현 수준에 대한 검정과 같은 임의의 검출 검정에 의해 발현 수준을 결정함에 의해 정량적으로 검증될 수 있고; 뉴클레오티드 시퀀싱 반응에 의해 정성적으로 검증될 수 있다.

[0894] 일부 실시형태에서, 변형, 예를 들어 단일 염기 편집은 유전자 표적화된 발현의 적어도 10% 감소를 초래한다. 일부 실시형태에서, 염기 편집 효율은 유전자 표적화된 발현의 적어도 10% 감소를 초래할 수 있다. 일부 실시형태에서, 염기 편집 효율은 유전자 표적화된 발현의 적어도 20% 감소를 초래할 수 있다. 일부 실시형태에서, 염기 편집 효율은 유전자 표적화된 발현의 적어도 30% 감소를 초래할 수 있다. 일부 실시형태에서, 염기 편집 효율은 유전자 표적화된 발현의 적어도 40% 감소를 초래할 수 있다. 일부 실시형태에서, 염기 편집 효율은 유전자 표적화된 발현의 적어도 50% 감소를 초래할 수 있다. 일부 실시형태에서, 염기 편집 효율은 표적화된 유전자 발현의 적어도 60% 감소를 초래할 수 있다. 일부 실시형태에서, 염기 편집 효율은 표적화된 유전자 발현의 적어도 70% 감소를 초래할 수 있다. 일부 실시형태에서, 염기 편집 효율은 표적화된 유전자 발현의 적어도 80% 감소를 초래할 수 있다. 일부 실시형태에서, 염기 편집 효율은 표적화된 유전자 발현의 적어도 90% 감소를 초래할 수 있다. 일부 실시형태에서, 염기 편집 효율은 표적화된 유전자 발현의 적어도 91% 감소를 초래할 수 있다. 일부 실시형태에서, 염기 편집 효율은 표적화된 유전자 발현의 적어도 92% 감소를 초래할 수 있다. 일부 실시형태에서, 염기 편집 효율은 표적화된 유전자 발현의 적어도 93% 감소를 초래할 수 있다. 일부 실시형태에서, 염기 편집 효율은 표적화된 유전자 발현의 적어도 94% 감소를 초래할 수 있다. 일부 실시형태에서, 염기 편집 효율은 표적화된 유전자 발현의 적어도 95% 감소를 초래할 수 있다. 일부 실시형태에서, 염기 편집 효율은 표적화된 유전자 발현의 적어도 96% 감소를 초래할 수 있다. 일부 실시형태에서, 염기 편집 효율은 표적화된 유전자 발현의 적어도 97% 감소를 초래할 수 있다. 일부 실시형태에서, 염기 편집 효율은 표적화된 유전자 발현의 적어도 98% 감소를 초래할 수 있다. 일부 실시형태에서, 염기 편집 효율은 표적화된 유전자 발현의 적어도 99% 감소를 초래할 수 있다. 일부 실시형태에서, 염기 편집 효율은 표적화되는 유전자의 녹아웃(유전자 발현의 100% 녹다운)을 초래할 수 있다.

[0895] 일부 실시형태에서, 본 명세서에 제공된 임의의 염기 편집기 시스템은 표적 폴리뉴클레오티드 서열에서 50% 미만, 40% 미만, 30% 미만, 20% 미만, 19% 미만, 18% 미만, 17% 미만, 16% 미만, 15% 미만, 14% 미만, 13% 미만, 12% 미만, 11% 미만, 10% 미만, 9% 미만, 8% 미만, 7% 미만, 6% 미만, 5% 미만, 4% 미만, 3% 미만, 2% 미만, 1% 미만, 0.9% 미만, 0.8% 미만, 0.7% 미만, 0.6% 미만, 0.5% 미만, 0.4% 미만, 0.3% 미만, 0.2% 미만, 0.1% 미만, 0.09% 미만, 0.08% 미만, 0.07% 미만, 0.06% 미만, 0.05% 미만, 0.04% 미만, 0.03% 미만, 0.02% 미만, 또는 0.01% 미만의 삽입결실 형성을 초래한다.

[0896] 일부 실시형태에서, 표적화된 변형, 예를 들어 단일 염기 편집은 상이한 가이드 RNA로 염기 편집을 위해 적어도 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49 또는 50개의 상이한 내인성 서열을 동시적으로 표적화하기 위해 사용된다. 일부 실시형태에서, 표적화된 변형, 예를 들어 단일 염기 편집은 상이한 가이드 RNA로 염기 편집을 위해 적어도 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44,

45, 46, 47, 48, 49, 50개 이상의 상이한 내인성 유전자 서열을 순차적으로 표적화하기 위해 사용된다.

[0897] 개시내용의 일부 양태는 본 명세서에 제공된 임의의 염기 편집기는, 의도되지 않은 점 돌연변이(즉, 방관자에 의한 돌연변이)와 같은 의도되지 않은 돌연변이의 유의한 수를 생성함이 없이, 핵산(예를 들어, 대상체의 게놈 내의 핵산)에서 점 돌연변이와 같은 의도된 돌연변이를 효율적으로 생성할 수 있다는 인식에 기반한다. 일부 실시형태에서, 본 명세서에 제공된 임의의 염기 편집기는 의도된 돌연변이의 적어도 0.01%(즉, 적어도 0.01% 염기 편집 효율)를 생성할 수 있다. 일부 실시형태에서, 본 명세서에 제공된 임의의 염기 편집기는 의도된 돌연변이의 적어도 0.01%, 1%, 2%, 3%, 4%, 5%, 10%, 15%, 20%, 25%, 30%, 40%, 45%, 50%, 60%, 70%, 80%, 90%, 95%, 또는 99%를 생성할 수 있다.

[0898] 일부 실시형태에서, 본 명세서에 기재된 ABE8 염기 편집기 변이체 중 하나를 포함하는 임의의 염기 편집기 시스템은 표적 폴리뉴클레오티드 서열에서 50% 미만, 40% 미만, 30% 미만, 20% 미만, 19% 미만, 18% 미만, 17% 미만, 16% 미만, 15% 미만, 14% 미만, 13% 미만, 12% 미만, 11% 미만, 10% 미만, 9% 미만, 8% 미만, 7% 미만, 6% 미만, 5% 미만, 4% 미만, 3% 미만, 2% 미만, 1% 미만, 0.9% 미만, 0.8% 미만, 0.7% 미만, 0.6% 미만, 0.5% 미만, 0.4% 미만, 0.3% 미만, 0.2% 미만, 0.1% 미만, 0.09% 미만, 0.08% 미만, 0.07% 미만, 0.06% 미만, 0.05% 미만, 0.04% 미만, 0.03% 미만, 0.02% 미만, 또는 0.01% 미만의 삽입결실 형성을 초래한다. 일부 실시형태에서, 본 명세서에 기재된 ABE8 염기 편집기 변이체 중 하나를 포함하는 임의의 염기 편집기 시스템은 표적 폴리뉴클레오티드 서열에서 0.8% 미만의 삽입결실 형성을 초래한다. 일부 실시형태에서, 본 명세서에 기재된 ABE8 염기 편집기 변이체 중 하나를 포함하는 임의의 염기 편집기 시스템은 표적 폴리뉴클레오티드 서열에서 최대 0.8%의 삽입결실 형성을 초래한다. 일부 실시형태에서, 본 명세서에 기재된 ABE8 염기 편집기 변이체 중 하나를 포함하는 임의의 염기 편집기 시스템은 표적 폴리뉴클레오티드 서열에서 0.3% 미만의 삽입결실 형성을 초래한다. 일부 실시형태에서, 기재된 ABE8 염기 편집기 변이체 중 하나를 포함하는 임의의 염기 편집기 시스템은 ABE7 염기 편집기 중 하나를 포함하는 염기 편집기 시스템과 비교하여 표적 폴리뉴클레오티드 서열에서 더 낮은 삽입결실 형성을 초래한다. 일부 실시형태에서, 본 명세서에 기재된 ABE8 염기 편집기 변이체 중 하나를 포함하는 임의의 염기 편집기 시스템은 ABE7.10을 포함하는 염기 편집기 시스템과 비교하여 표적 폴리뉴클레오티드 서열에서 더 낮은 삽입결실 형성을 초래한다.

[0899] 일부 실시형태에서, 본 명세서에 기재된 ABE8 염기 편집기 변이체 중 하나를 포함하는 임의의 염기 편집기 시스템은 ABE7 염기 편집기 중 하나를 포함하는 염기 편집기 시스템에 비해 삽입결실 빈도에서 감소를 갖는다. 일부 실시형태에서, 본 명세서에 기재된 ABE8 염기 편집기 변이체 중 하나를 포함하는 임의의 염기 편집기 시스템은 ABE7 염기 편집기 중 하나를 포함하는 염기 편집기 시스템에 비해 삽입결실 빈도에서 적어도 0.01%, 적어도 1%, 적어도 2%, 적어도 3%, 적어도 4%, 적어도 5%, 적어도 10%, 적어도 15%, 적어도 20%, 적어도 25%, 적어도 30%, 적어도 35%, 적어도 40%, 적어도 45%, 적어도 50%, 적어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 또는 적어도 95% 감소를 갖는다. 일부 실시형태에서, 본 명세서에 기재된 ABE8 염기 편집기 변이체 중 하나를 포함하는 염기 편집기 시스템은 ABE7.10을 포함하는 염기 편집기 시스템에 비해 삽입결실 빈도에서 적어도 0.01%, 적어도 1%, 적어도 2%, 적어도 3%, 적어도 4%, 적어도 5%, 적어도 10%, 적어도 15%, 적어도 20%, 적어도 25%, 적어도 30%, 적어도 35%, 적어도 40%, 적어도 45%, 적어도 50%, 적어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 또는 적어도 95% 감소를 갖는다.

[0900] 발명은 효율성과 특이성이 증가된 아데노신 데아미나제 변이체(예를 들어, ABE8 변이체)를 제공한다. 특히, 본 명세서에 기재된 아데노신 데아미나제 변이체는 폴리뉴클레오티드 내의 원하는 염기를 편집할 가능성이 더 높고, 변경되어 지도록 의도되지 않은 염기(예를 들어, "방관자")를 편집할 가능성은 적다.

[0901] 일부 실시형태에서, 본 명세서에 기재된 ABE8 염기 편집기 변이체 중 하나를 포함하는 임의의 염기 편집 시스템은 감소된 방관자 편집 또는 돌연변이를 갖는다. 일부 실시형태에서, 의도되지 않은 편집 또는 돌연변이는 방관자 돌연변이 또는 방관자 편집, 예를 들어 표적 뉴클레오티드 서열의 표적 윈도우에서의 의도되지 않은 또는 비-표적 위치에서 표적 염기(예를 들어, A 또는 C)의 염기 편집이다. 일부 실시형태에서, 본 명세서에 기재된 ABE8 염기 편집기 변이체 중 하나를 포함하는 임의의 염기 편집 시스템은 ABE7 염기 편집기, 예를 들어 ABE7.10을 포함하는 염기 편집기 시스템에 비해 감소된 방관자 편집 또는 돌연변이를 갖는다. 일부 실시형태에서, 본 명세서에 기재된 ABE8 염기 편집기 변이체 중 하나를 포함하는 임의의 염기 편집 시스템은 ABE7 염기 편집기, 예를 들어 ABE7.10을 포함하는 염기 편집기 시스템에 비해 적어도 1%, 적어도 2%, 적어도 3%, 적어도 4%, 적어도 5%, 적어도 10%, 적어도 15%, 적어도 20%, 적어도 25%, 적어도 30%, 적어도 35%, 적어도 40%, 적어도 45%, 적어도 50%, 적어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도

90%, 적어도 95%, 또는 적어도 99% 만큼 감소된 방관자 편집 또는 돌연변이를 갖는다. 일부 실시형태에서, 본 명세서에 기재된 ABE8 염기 편집기 변이체 중 하나를 포함하는 임의의 염기 편집 시스템은 ABE7 염기 편집기, 예를 들어 ABE7.10을 포함하는 염기 편집기 시스템에 비해 적어도 1.1배, 적어도 1.2배, 적어도 1.3배, 적어도 1.4배, 적어도 1.5배, 적어도 1.6배, 적어도 1.7배, 적어도 1.8배, 적어도 1.9배, 적어도 2.0배, 적어도 2.1배, 적어도 2.2배, 적어도 2.3배, 적어도 2.4배, 적어도 2.5배, 적어도 2.6배, 적어도 2.7배, 적어도 2.8배, 적어도 2.9배, 또는 적어도 3.0배 만큼 감소된 방관자 편집 또는 돌연변이를 갖는다.

[0902] 일부 실시형태에서, 본 명세서에 기재된 ABE8 염기 편집기 변이체 중 하나를 포함하는 임의의 염기 편집 시스템은 감소된 의사의 편집을 갖는다. 일부 실시형태에서, 의도되지 않은 편집 또는 돌연변이는 의사의 돌연변이 또는 의사의 편집, 예를 들어 게놈의 의도되지 않거나 비-표적 영역에서 표적 염기(예를 들어 A 또는 C)의 비-특이적 편집 또는 가이드 독립적 편집이다. 일부 실시형태에서, 본 명세서에 기재된 ABE8 염기 편집기 변이체 중 하나를 포함하는 임의의 염기 편집 시스템은 ABE7 염기 편집기, 예를 들어 ABE7.10을 포함하는 염기 편집기 시스템에 비해 감소된 의사의 편집을 갖는다. 일부 실시형태에서, 본 명세서에 기재된 ABE8 염기 편집기 변이체 중 하나를 포함하는 임의의 염기 편집 시스템은 ABE7 염기 편집기, 예를 들어, ABE7.10을 포함하는 염기 편집기 시스템과 비교하여 적어도 1%, 적어도 2%, 적어도 3%, 적어도 4%, 적어도 5%, 적어도 10%, 적어도 15%, 적어도 20%, 적어도 25%, 적어도 30%, 적어도 35%, 적어도 40%, 적어도 45%, 적어도 50%, 적어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 또는 적어도 99% 만큼 감소된 의사의 편집을 갖는다. 일부 실시형태에서, 본 명세서에 기재된 ABE8 염기 편집기 변이체 중 하나를 포함하는 임의의 염기 편집 시스템은 ABE7 염기 편집기, 예를 들어, ABE7.10을 포함하는 염기 편집기 시스템과 비교하여 적어도 1.1배, 적어도 1.2배, 적어도 1.3배, 적어도 1.4배, 적어도 1.5배, 적어도 1.6배, 적어도 1.7배, 적어도 1.8배, 적어도 1.9배, 적어도 2.0배, 적어도 2.1배, 적어도 2.2배, 적어도 2.3배, 적어도 2.4배, 적어도 2.5배, 적어도 2.6배, 적어도 2.7배, 적어도 2.8배, 적어도 2.9배, 또는 적어도 3.0배 만큼 감소된 의사의 편집을 갖는다.

[0903] 일부 실시형태에서, 본 명세서에 기재된 임의의 ABE8 염기 편집기 변이체는 적어도 0.01%, 적어도 1%, 적어도 2%, 적어도 3%, 적어도 4%, 적어도 5%, 적어도 10%, 적어도 15%, 적어도 20%, 적어도 25%, 적어도 30%, 적어도 35%, 적어도 40%, 적어도 45%, 적어도 50%, 적어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 또는 적어도 99% 염기 편집 효율성을 갖는다. 일부 실시형태에서, 염기 편집 효율성은 세포의 모집단에서 편집된 핵염기의 백분율을 계산함에 의해 측정될 수 있다. 일부 실시형태에서, 본 명세서에 기재된 임의의 ABE8 염기 편집기 변이체는 세포의 모집단에서 편집된 핵염기에 의해 측정된 바와 같은 적어도 0.01%, 적어도 1%, 적어도 2%, 적어도 3%, 적어도 4%, 적어도 5%, 적어도 10%, 적어도 15%, 적어도 20%, 적어도 25%, 적어도 30%, 적어도 35%, 적어도 40%, 적어도 45%, 적어도 50%, 적어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 또는 적어도 99%의 염기 편집 효율성을 갖는다.

[0904] 일부 실시형태에서, 본 명세서에 기재된 임의의 ABE8 염기 편집기 변이체는 ABE7 염기 편집기에 비해 더 높은 기본 편집 효율성을 갖는다. 일부 실시형태에서, 본 명세서에 기재된 임의의 ABE8 염기 편집기 변이체는 ABE7 염기 편집기, 예를 들어, ABE7.10과 비교하여 적어도 1%, 적어도 2%, 적어도 3%, 적어도 4%, 적어도 5%, 적어도 10%, 적어도 15%, 적어도 20%, 적어도 25%, 적어도 30%, 적어도 35%, 적어도 40%, 적어도 45%, 적어도 50%, 적어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 99%, 적어도 100%, 적어도 105%, 적어도 110%, 적어도 115%, 적어도 120%, 적어도 125%, 적어도 130%, 적어도 135%, 적어도 140%, 적어도 145%, 적어도 150%, 적어도 155%, 적어도 160%, 적어도 165%, 적어도 170%, 적어도 175%, 적어도 180%, 적어도 185%, 적어도 190%, 적어도 195%, 적어도 200%, 적어도 210%, 적어도 220%, 적어도 230%, 적어도 240%, 적어도 250%, 적어도 260%, 적어도 270%, 적어도 280%, 적어도 290%, 적어도 300%, 적어도 310%, 적어도 320%, 적어도 330%, 적어도 340%, 적어도 350%, 적어도 360%, 적어도 370%, 적어도 380%, 적어도 390%, 적어도 400%, 적어도 450%, 또는 적어도 500% 더 높은 염기 편집 효율성을 갖는다.

[0905] 일부 실시형태에서, 본 명세서에 기재된 임의의 ABE8 염기 편집기 변이체는 ABE7 염기 편집기, 예를 들어, ABE7.10과 비교하여 적어도 1.1배, 적어도 1.2배, 적어도 1.3배, 적어도 1.4배, 적어도 1.5배, 적어도 1.6배, 적어도 1.7배, 적어도 1.8배, 적어도 1.9배, 적어도 2.0배, 적어도 2.1배, 적어도 2.2배, 적어도 2.3배, 적어도 2.4배, 적어도 2.5배, 적어도 2.6배, 적어도 2.7배, 적어도 2.8배, 적어도 2.9배, 적어도 3.0배, 적어도 3.1배, 적어도 3.2, 적어도 3.3배, 적어도 3.4배, 적어도 3.5배, 적어도 3.6배, 적어도 3.7배, 적어도 3.8배, 적어도 3.9배, 적어도 4.0배, 적어도 4.1배, 적어도 4.2배, 적어도 4.3배, 적어도 4.4배, 적어도 4.5배, 적어도

4.6배, 적어도 4.7배, 적어도 4.8배, 적어도 4.9배, 또는 적어도 5.0배 더 높은 염기 편집 효율성을 갖는다.

[0906] 일부 실시형태에서, 본 명세서에 기재된 임의의 ABE8 염기 편집기 변이체는 적어도 0.01%, 적어도 1%, 적어도 2%, 적어도 3%, 적어도 4%, 적어도 5%, 적어도 10%, 적어도 15%, 적어도 20%, 적어도 25%, 적어도 30%, 적어도 35%, 적어도 40%, 적어도 45%, 적어도 50%, 적어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 또는 적어도 99% 표적-상 염기 편집 효율성을 갖는다. 일부 실시형태에서, 본 명세서에 기재된 임의의 ABE8 염기 편집기 변이체는 세포의 모집단에서 편집된 표적 핵염기에 의해 측정된 바와 같은 적어도 0.01%, 적어도 1%, 적어도 2%, 적어도 3%, 적어도 4%, 적어도 5%, 적어도 10%, 적어도 15%, 적어도 20%, 적어도 25%, 적어도 30%, 적어도 35%, 적어도 40%, 적어도 45%, 적어도 50%, 적어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 또는 적어도 99%의 표적-상 염기 편집 효율성을 갖는다.

[0907] 일부 실시형태에서, 본 명세서에 기재된 임의의 ABE8 염기 편집기 변이체는 ABE7 염기 편집기에 비해 더 높은 표적-상 염기 편집 효율성을 갖는다. 일부 실시형태에서, 본 명세서에 기재된 임의의 ABE8 염기 편집기 변이체는 ABE7 염기 편집기, 예를 들어, ABE7.10에 비해 적어도 1%, 적어도 2%, 적어도 3%, 적어도 4%, 적어도 5%, 적어도 10%, 적어도 15%, 적어도 20%, 적어도 25%, 적어도 30%, 적어도 35%, 적어도 40%, 적어도 45%, 적어도 50%, 적어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 99%, 적어도 100%, 적어도 105%, 적어도 110%, 적어도 115%, 적어도 120%, 적어도 125%, 적어도 130%, 적어도 135%, 적어도 140%, 적어도 145%, 적어도 150%, 적어도 155%, 적어도 160%, 적어도 165%, 적어도 170%, 적어도 175%, 적어도 180%, 적어도 185%, 적어도 190%, 적어도 195%, 적어도 200%, 적어도 210%, 적어도 220%, 적어도 230%, 적어도 240%, 적어도 250%, 적어도 260%, 적어도 270%, 적어도 280%, 적어도 290%, 적어도 300%, 적어도 310%, 적어도 320%, 적어도 330%, 적어도 340%, 적어도 350%, 적어도 360%, 적어도 370%, 적어도 380%, 적어도 390%, 적어도 400%, 적어도 450%, 또는 적어도 500% 더 높은 표적-상 염기 편집 효율성을 갖는다.

[0908] 일부 실시형태에서, 본 명세서에 기재된 임의의 ABE8 염기 편집기 변이체는 ABE7 염기 편집기, 예를 들어, ABE7.10에 비해 적어도 1.1배, 적어도 1.2배, 적어도 1.3배, 적어도 1.4배, 적어도 1.5배, 적어도 1.6배, 적어도 1.7배, 적어도 1.8배, 적어도 1.9배, 적어도 2.0배, 적어도 2.1배, 적어도 2.2배, 적어도 2.3배, 적어도 2.4배, 적어도 2.5배, 적어도 2.6배, 적어도 2.7배, 적어도 2.8배, 적어도 2.9배, 적어도 3.0배, 적어도 3.1배, 적어도 3.2배, 적어도 3.3배, 적어도 3.4배, 적어도 3.5배, 적어도 3.6배, 적어도 3.7배, 적어도 3.8배, 적어도 3.9배, 적어도 4.0배, 적어도 4.1배, 적어도 4.2배, 적어도 4.3배, 적어도 4.4배, 적어도 4.5배, 적어도 4.6배, 적어도 4.7배, 적어도 4.8배, 적어도 4.9배, 또는 적어도 5.0배 더 높은 표적-상 염기 편집 효율성을 갖는다.

[0909] 본 명세서에 기재된 ABE8 염기 편집기 변이체는 플라스미드, 벡터, LNP 복합체 또는 mRNA를 통해 숙주 세포로 전달될 수 있다. 일부 실시형태에서, 본 명세서에 기재된 임의의 ABE8 염기 편집기 변이체는 mRNA로서 숙주 세포에 전달된다. 일부 실시형태에서, 핵산 기반 전달 시스템, 예를 들어 mRNA를 통해 전달된 ABE8 염기 편집기는 편집된 핵염기에 의해 측정된 바와 같은 적어도 1%, 적어도 2%, 적어도 3%, 적어도 4%, 적어도 5%, 적어도 10%, 적어도 15%, 적어도 20%, 적어도 25%, 적어도 30%, 적어도 35%, 적어도 40%, 적어도 45%, 적어도 50%, 적어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 또는 적어도 99%의 표적-상 편집 효율성을 갖는다. 일부 실시형태에서, mRNA 시스템에 의해 전달된 ABE8 염기 편집기는 플라스미드 또는 벡터 시스템에 의해 전달된 ABE8 염기 편집기에 비해 더 높은 염기 편집 효율성을 갖는다. 일부 실시형태에서, 본 명세서에 기재된 임의의 ABE8 염기 편집기 변이체는 플라스미드 또는 벡터 시스템에 의해 전달될 때와 비교하여 mRNA 시스템에 의해 전달될 때 적어도 1%, 적어도 2%, 적어도 3%, 적어도 4%, 적어도 5%, 적어도 10%, 적어도 15%, 적어도 20%, 적어도 25%, 적어도 30%, 적어도 35%, 적어도 40%, 적어도 45%, 적어도 50%, 적어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 99%, 적어도 100%, 적어도 105%, 적어도 110%, 적어도 115%, 적어도 120%, 적어도 125%, 적어도 130%, 적어도 135%, 적어도 140%, 적어도 145%, 적어도 150%, 적어도 155%, 적어도 160%, 적어도 165%, 적어도 170%, 적어도 175%, 적어도 180%, 적어도 185%, 적어도 190%, 적어도 195%, 적어도 200%, 적어도 210%, 적어도 220%, 적어도 230%, 적어도 240%, 적어도 250%, 적어도 260%, 적어도 270%, 적어도 280%, 적어도 290%, 적어도 300%, 적어도 310%, 적어도 320%, 적어도 330%, 적어도 340%, 적어도 350%, 적어도 360%, 적어도 370%, 적어도 380%, 적어도 390%, 적어도 400%, 적어도 450%, 또는 적어도 500%의 표적-상 편집 효율성을 갖는다. 일부 실시형태에서, 본 명세서에 기재된 임의의 ABE8 염기 편집기 변이체는 플라스미드 또는 벡터 시스템에

의해 전달될 때와 비교하여 mRNA 시스템에 의해 전달될 때 적어도 1.1배, 적어도 1.2배, 적어도 1.3배, 적어도 1.4배, 적어도 1.5배, 적어도 1.6배, 적어도 1.7배, 적어도 1.8배, 적어도 1.9배, 적어도 2.0배, 적어도 2.1배, 적어도 2.2배, 적어도 2.3배, 적어도 2.4배, 적어도 2.5배, 적어도 2.6배, 적어도 2.7배, 적어도 2.8배, 적어도 2.9배, 적어도 3.0배, 적어도 3.1배, 적어도 3.2배, 적어도 3.3배, 적어도 3.4배, 적어도 3.5배, 적어도 3.6배, 적어도 3.7배, 적어도 3.8배, 적어도 3.9배, 적어도 4.0배, 적어도 4.1배, 적어도 4.2배, 적어도 4.3배, 적어도 4.4배, 적어도 4.5배, 적어도 4.6배, 적어도 4.7배, 적어도 4.8배, 적어도 4.9배, 또는 적어도 5.0배 더 높은 표적-상 편집 효율성을 갖는다.

[0910] 일부 실시형태에서, 본 명세서에 기재된 ABE8 염기 편집기 변이체 중 하나를 포함하는 임의의 염기 편집기 시스템은 표적 폴리뉴클레오티드 서열에서의 50% 미만, 40% 미만, 30% 미만, 20% 미만, 19% 미만, 18% 미만, 17% 미만, 16% 미만, 15% 미만, 14% 미만, 13% 미만, 12% 미만, 11% 미만, 10% 미만, 9% 미만, 8% 미만, 7% 미만, 6% 미만, 5% 미만, 4% 미만, 3% 미만, 2% 미만, 1% 미만, 0.9% 미만, 0.8% 미만, 0.7% 미만, 0.6% 미만, 0.5% 미만, 0.4% 미만, 0.3% 미만, 0.2% 미만, 0.1% 미만, 0.09% 미만, 0.08% 미만, 0.07% 미만, 0.06% 미만, 0.05% 미만, 0.04% 미만, 0.03% 미만, 0.02% 미만, 또는 0.01% 미만의 표적-외 편집을 초래한다.

[0911] 일부 실시형태에서, 본 명세서에 기재된 임의의 ABE8 염기 편집기 변이체는 플라스미드 또는 벡터 시스템에 의해 전달될 때와 비교하여 mRNA 시스템에 의해 전달될 때 더 낮은 가이드된 표적-외 편집 효율성을 갖는다. 일부 실시형태에서, 본 명세서에 기재된 임의의 ABE8 염기 편집기 변이체는 플라스미드 또는 벡터 시스템에 의해 전달될 때와 비교하여 mRNA 시스템에 의해 전달될 때 적어도 1%, 적어도 2%, 적어도 3%, 적어도 4%, 적어도 5%, 적어도 10%, 적어도 15%, 적어도 20%, 적어도 25%, 적어도 30%, 적어도 35%, 적어도 40%, 적어도 45%, 적어도 50%, 적어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 또는 적어도 99% 더 낮은 가이드된 표적-외 편집 효율성을 갖는다. 일부 실시형태에서, 본 명세서에 기재된 임의의 ABE8 염기 편집기 변이체는 플라스미드 또는 벡터 시스템에 의해 전달될 때와 비교하여 mRNA 시스템에 의해 전달될 때 적어도 1.1배, 적어도 1.2배, 적어도 1.3배, 적어도 1.4배, 적어도 1.5배, 적어도 1.6배, 적어도 1.7배, 적어도 1.8배, 적어도 1.9배, 적어도 2.0배, 적어도 2.1배, 적어도 2.2배, 적어도 2.3배, 적어도 2.4배, 적어도 2.5배, 적어도 2.6배, 적어도 2.7배, 적어도 2.8배, 적어도 2.9배, 또는 적어도 3.0배 더 낮은 가이드된 표적-외 편집 효율성을 갖는다. 일부 실시형태에서, 본 명세서에 기재된 임의의 ABE8 염기 편집기 변이체는 플라스미드 또는 벡터 시스템에 의해 전달될 때와 비교하여 mRNA 시스템에 의해 전달될 때 가이드된 표적-외 편집 효율성에서 적어도 약 2.2배 감소를 갖는다.

[0912] 일부 실시형태에서, 본 명세서에 기재된 임의의 ABE8 염기 편집기 변이체는 플라스미드 또는 벡터 시스템에 의해 전달될 때와 비교하여 mRNA 시스템에 의해 전달될 때 더 낮은 가이드-독립적 표적-외 편집 효율성을 갖는다. 일부 실시형태에서, 본 명세서에 기재된 임의의 ABE8 염기 편집기 변이체는 플라스미드 또는 벡터 시스템에 의해 전달될 때와 비교하여 mRNA 시스템에 의해 전달될 때 적어도 1%, 적어도 2%, 적어도 3%, 적어도 4%, 적어도 5%, 적어도 10%, 적어도 15%, 적어도 20%, 적어도 25%, 적어도 30%, 적어도 35%, 적어도 40%, 적어도 45%, 적어도 50%, 적어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 또는 적어도 99% 더 낮은 가이드-독립적 표적-외 편집 효율성을 갖는다. 일부 실시형태에서, 본 명세서에 기재된 임의의 ABE8 염기 편집기 변이체는 플라스미드 또는 벡터 시스템에 의해 전달될 때와 비교하여 mRNA 시스템에 의해 전달될 때 적어도 1.1배, 적어도 1.2배, 적어도 1.3배, 적어도 1.4배, 적어도 1.5배, 적어도 1.6배, 적어도 1.7배, 적어도 1.8배, 적어도 1.9배, 적어도 2.0배, 적어도 2.1배, 적어도 2.2배, 적어도 2.3배, 적어도 2.4배, 적어도 2.5배, 적어도 2.6배, 적어도 2.7배, 적어도 2.8배, 적어도 2.9배, 적어도 3.0배, 적어도 5.0배, 적어도 10.0배, 적어도 20.0배, 적어도 50.0배, 적어도 70.0배, 적어도 100.0배, 적어도 120.0배, 적어도 130.0배, 또는 적어도 150.0배 더 낮은 가이드-독립적 표적-외 편집 효율성을 갖는다. 일부 실시형태에서, 본 명세서에 기재된 ABE8 염기 편집기 변이체는 플라스미드 또는 벡터 시스템에 의해 전달될 때와 비교하여 mRNA 시스템에 의해 전달될 때 가이드-독립적 표적-외 편집 효율성(예를 들어, 의사의 RNA 탈아미노화)에서 134.0배 감소를 갖는다. 일부 실시형태에서, 본 명세서에 기재된 ABE8 염기 편집기 변이체는 게놈 전반에 걸쳐 가이드-독립적 돌연변이 비율을 증가시키지 않는다.

[0913] 일부 실시형태에서, 단일 유전자 전달 이벤트(예를 들어, 형질도입, 형질감염, 전기천공 또는 임의의 다른 방법에 의함)는 세포의 게놈 내의 5개 서열의 염기 편집을 표적화하기 위해 사용될 수 있다. 일부 실시형태에서, 단일 유전자 전달 이벤트는 세포의 게놈 내의 6개 서열의 염기 편집을 표적화하기 위해 사용될 수 있다. 일부 실시형태에서, 단일 유전자 전달 이벤트는 세포의 게놈 내의 7개 서열의 염기 편집을 표적화하기 위해 사용될 수 있다. 일부 실시형태에서, 단일 전기천공 이벤트는 세포의 게놈 내의 8개 서열의 염기 편집을 표적화하기 위해

사용될 수 있다. 일부 실시형태에서, 단일 유전자 전달 이벤트는 세포의 게놈 내의 9개 서열의 염기 편집을 표적화하기 위해 사용될 수 있다. 일부 실시형태에서, 단일 유전자 전달 이벤트는 세포의 게놈 내의 10개 서열의 염기 편집을 표적화하기 위해 사용될 수 있다. 일부 실시형태에서, 단일 유전자 전달 이벤트는 세포의 게놈 내의 20개 서열의 염기 편집을 표적화하기 위해 사용될 수 있다. 일부 실시형태에서, 단일 유전자 전달 이벤트는 세포의 게놈 내의 30개 서열의 염기 편집을 표적화하기 위해 사용될 수 있다. 일부 실시형태에서, 단일 유전자 전달 이벤트는 세포의 게놈 내의 40개 서열의 염기 편집을 표적화하기 위해 사용될 수 있다. 일부 실시형태에서, 단일 유전자 전달 이벤트는 세포의 게놈 내의 50개 서열의 염기 편집을 표적화하기 위해 사용될 수 있다.

[0914] 일부 실시형태에서, 본 명세서에 기재된 방법, 예를 들어 염기 편집 방법은 표적-외 효과를 최소화 내지 전혀 갖지 않는다.

[0915] 일부 실시형태에서, 본 명세서에 기재된 염기 편집 방법은 적어도 50%의 성공적으로 편집된 세포 모집단(즉, 성공적으로 조작된 세포)을 초래한다. 일부 실시형태에서, 본 명세서에 기재된 염기 편집 방법은 적어도 55%의 성공적으로 편집된 세포 모집단을 초래한다. 일부 실시형태에서, 본 명세서에 기재된 염기 편집 방법은 적어도 60%의 성공적으로 편집된 세포 모집단을 초래한다. 일부 실시형태에서, 본 명세서에 기재된 염기 편집 방법은 적어도 65%의 성공적으로 편집된 세포 모집단을 초래한다. 일부 실시형태에서, 본 명세서에 기재된 염기 편집 방법은 적어도 70%의 성공적으로 편집된 세포 모집단을 초래한다. 일부 실시형태에서, 본 명세서에 기재된 염기 편집 방법은 적어도 75%의 성공적으로 편집된 세포 모집단을 초래한다. 일부 실시형태에서, 본 명세서에 기재된 염기 편집 방법은 적어도 80%의 성공적으로 편집된 세포 모집단을 초래한다. 일부 실시형태에서, 본 명세서에 기재된 염기 편집 방법은 적어도 85%의 성공적으로 편집된 세포 모집단을 초래한다. 일부 실시형태에서, 본 명세서에 기재된 염기 편집 방법은 적어도 90%의 성공적으로 편집된 세포 모집단을 초래한다. 일부 실시형태에서, 본 명세서에 기재된 염기 편집 방법은 적어도 95%의 성공적으로 편집된 세포 모집단을 초래한다. 일부 실시형태에서, 본 명세서에 기재된 염기 편집 방법은 약 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 100%의 성공적으로 편집된 세포 모집단을 초래한다.

[0916] 일부 실시형태에서, 염기 편집 개입에 이어 살아있는 세포 회복은 염기 편집 이벤트 당시에서 시작하는 세포 모집단의 적어도 60%, 70%, 80%, 90%보다 크다. 일부 실시형태에서, 상기 기재된 바와 같은 살아있는 세포 회복은 약 70%이다. 일부 실시형태에서, 상기 기재된 바와 같은 살아있는 세포 회복은 약 75%이다. 일부 실시형태에서, 상기 기재된 바와 같은 살아있는 세포 회복은 약 80%이다. 일부 실시형태에서, 상기 기재된 바와 같은 살아있는 세포 회복은 약 85%이다. 일부 실시형태에서, 상기 기재된 바와 같은 살아있는 세포 회복은 염기 편집 이벤트 당시에 모집단에서의 세포 중 약 90%, 또는 약 91%, 92%, 93%, 94% 95%, 96%, 97%, 98%, 99% 또는 100%이다.

[0917] 일부 실시형태에서 조작된 세포 모집단은 시험관내에서 약 2-배, 약 3-배, 약 4-배, 약 5-배, 약 6-배, 약 7-배, 약 8-배, 약 9-배, 약 10-배, 약 15-배, 약 20-배, 약 25-배, 약 30-배, 약 35-배, 약 40-배, 약 45-배, 약 50-배, 또는 약 100-배로 추가로 확장될 수 있다.

[0918] 의도된 돌연변이 및 삽입결실의 수는 임의의 적합한 방법, 예를 들어 국제 PCT 출원 번호 PCT/2017/045381(WO2018/027078) 및 PCT/US2016/058344(WO2017/070632); Komor, A.C., 등, "Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage" Nature 533, 420-424 (2016); Gaudelli, N.M., 등, "Programmable base editing of A·T to G·C in genomic DNA without DNA cleavage" Nature 551, 464-471 (2017); 및 Komor, A.C., 등, "Improved base excision repair inhibition and bacteriophage Mu Gam protein yields C:G-to-T:A base editors with higher efficiency and product purity" Science Advances 3:eaao4774 (2017)에 기재된 바와 같은 방법을 사용하여 결정될 수 있으며; 그 전체 내용은 참조로 본 명세서에 포함된다.

[0919] 일부 실시형태에서, 삽입결실 빈도를 계산하기 위해, 삽입결실이 발생할 수 있는 윈도우의 양 측면에 측접하는 2개의 10-bp 서열에 대한 정확한 일치에 대해 시퀀싱 판독을 스캐닝한다. 정확한 일치가 위치되지 않으면, 판독은 분석에서 제외된다. 이 삽입결실 윈도우의 길이가 참조 서열과 정확하게 일치하는 경우 판독은 삽입결실을 함유하지 않는 것으로 분류된다. 삽입결실 윈도우가 참조 서열보다 길거나 짧은 2개 이상의 염기인 경우, 시퀀싱 판독은 각각 삽입 또는 결실로 분류된다. 일부 실시형태에서, 본 명세서에 제공된 염기 편집기는 핵산의 영역에서 삽입결실의 형성을 제한할 수 있다. 일부 실시형태에서, 영역은 염기 편집기에 의해 표적화된 뉴클레오티드에 있거나, 염기 편집기에 의해 표적화된 뉴클레오티드의 2, 3, 4, 5, 6, 7, 8, 9 또는 10개 뉴클레오티드 내의 영역이다.

[0920] 표적 뉴클레오티드 영역에서 형성된 삽입결실의 수는 핵산(예를 들어, 세포의 게놈 내의 핵산)이 염기 편집기에 노출되는 시간의 양에 따라 달라질 수 있다. 일부 실시형태에서, 삽입결실의 수 또는 비율은 표적 뉴클레오티드 서열(예를 들어, 세포의 게놈 내의 핵산)을 염기 편집기에 노출시킨지 적어도 1시간, 적어도 2시간, 적어도 6시간, 적어도 12시간, 적어도 24시간, 적어도 36시간, 적어도 48시간, 적어도 3일, 적어도 4일, 적어도 5일, 적어도 7일, 적어도 10일, 또는 적어도 14일 후에 결정된다. 본 명세서에 기재된 바와 같은 염기 편집기의 특징은 본 명세서에 제공된 임의의 융합 단백질, 또는 융합 단백질을 사용하는 방법에 적용될 수 있음을 이해해야 한다.

[0921] 염기 편집기 효율성에 대한 자세한 내용은 국제 PCT 출원 번호 PCT/2017/045381(WO 2018/027078) 및 PCT/US2016/058344(WO 2017/070632)에 기술되어 있으며, 이들 각각은 그 전체가 참조로 본 명세서에 포함된다. 또한 Komor, A.C., 등, "Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage" Nature 533, 420-424 (2016); Gaudelli, N.M., 등, "Programmable base editing of A·T to G·C in genomic DNA without DNA cleavage" Nature 551, 464-471 (2017); 및 Komor, A.C., 등, "Improved base excision repair inhibition and bacteriophage Mu Gam protein yields C:G-to-T:A base editors with higher efficiency and product purity" Science Advances 3:eaao4774 (2017)를 참조하며; 그 전체 내용은 참조로 본 명세서에 포함된다. 일부 실시형태에서, 본 명세서에 제공된 방법을 사용하여 하나 이상의 유전자에서 복수의 핵염기 쌍을 편집하는 것은 적어도 하나의 의도된 돌연변이의 형성을 초래한다. 일부 실시형태에서, 상기 적어도 하나의 의도된 돌연변이의 상기 형성은 유전자의 정상적인 기능의 파괴를 초래한다. 일부 실시형태에서, 상기 적어도 하나의 의도된 돌연변이의 상기 형성은 유전자에 의해 인코딩되는 단백질의 발현을 감소시키거나 제거한다. 다중 편집은 본 명세서에 제공된 임의의 방법 또는 방법의 조합을 사용하여 수행될 수 있음을 이해해야 한다.

[0922] ***조작된 뉴클레아제***

[0923] 일부 실시형태에서, 유전자 편집 시스템은 조작된 뉴클레아제(예를 들어, 메가뉴클레아제, 아연 핑거 뉴클레아제(ZFN), 전사 활성제-유사 효과기 뉴클레아제(TALEN) 또는 Cas 뉴클레아제를 포함한다. 일부 실시형태에서, 유전자 편집 시스템은 ZFN을 포함한다. ZFN은 아연-핑거 DNA 결합 도메인("ZF")과 뉴클레아제 도메인을 포함하는 융합 단백질이다. 각각의 자연적으로-발생하는 ZF는 3개의 연속 염기쌍(DNA 삼중항)에 결합할 수 있고, ZF 반복이 조합되어 DNA 표적 서열을 인식하고 충분한 친화성을 제공한다. 따라서, 조작된 ZF 반복은 조합되어 더 긴 DNA 서열, 예컨대, 예를 들어 9개 염기쌍, 12개 염기쌍, 15개 염기쌍, 18개 염기쌍 등을 인식한다. 일부 실시형태에서, ZFN은 제한 엔도뉴클레아제(예를 들어, FokI)로부터 뉴클레아제 도메인에 융합된 ZF를 포함한다. 일부 실시형태에서, 뉴클레아제 도메인은 뉴클레아제가 이량체화하여 활성이 되는 경우에서와 같이, 이량체화 도메인, 및 ZF 반복을 포함하는 한 쌍의 ZFN을 포함하고 뉴클레아제 도메인은 DNA 분자의 반대편 가닥에 있는 각 ZF 반복에 의해 인식되는 2개의 절반 표적 서열을 포함하는 표적 서열을, 그 사이에서 상호연결하는 서열(때로는 문헌에서 스페이서라고도 함)로 표적화하도록 설계되었다. 예를 들어, 상호연결하는 서열은 길이가 5 내지 7개의 염기쌍이다. 쌍의 양 ZFN 모두가 결합할 때 뉴클레아제 도메인은 이량체화되고 상호연결하는 서열 내에 DSB를 도입할 수 있다. 일부 실시형태에서, 뉴클레아제 도메인의 이량체화 도메인은 이량체화를 촉진하기 위한 노브-인투-홀 모티프를 포함한다.

[0924] 일부 실시형태에서, 유전자 편집 시스템은 TALEN을 포함한다. TALEN의 DNA 결합 도메인은 일반적으로 34개 또는 35개 아미노산 반복("모듈" 또는 "TAL 모듈")의 가변 수를 포함하며, 각 모듈은 단일 DNA 염기쌍 A, T, G 또는 C에 결합한다. 각 모듈의 위치 12 및 13에서 인접 잔기("반복-가변 이중-잔기" 또는 RVD)는 모듈이 결합하는 단일 DNA 염기쌍을 특정한다. 일부 실시형태에서, TALEN은 제한 엔도뉴클레아제(예를 들어, FokI)로부터의 뉴클레아제 도메인을 포함할 수 있다. 일부 실시형태에서, 뉴클레아제 도메인은 이량체화되어 활성이 될 수 있고, 한 쌍의 TALENS는 DNA 분자의 반대편 가닥에 있는 각 DNA 결합 도메인에 의해 인식되는 2개의 절반 표적 서열을 포함하는 표적 서열을, 그 사이에서 상호연결하는 서열로 표적화하도록 설계되었다. 예를 들어, 각각의 절반 표적 서열은 10 내지 20개의 염기쌍의 범위에 있고, 상호연결하는 서열은 길이가 12 내지 19개의 염기쌍이다. 쌍의 양 TALEN 모두가 결합할 때, 뉴클레아제 도메인은 이량체화되고 상호연결하는 서열 내에 이중 가닥 파괴를 도입할 수 있다. 일부 실시형태에서, 뉴클레아제 도메인의 이량체화 도메인은 이량체화를 촉진하기 위한 노브-인투-홀 모티프를 포함할 수 있다.

[0925] 일부 실시형태에서, 유전자 편집 시스템은 메가뉴클레아제를 포함한다. 자연적으로-발생하는 메가뉴클레아제는 약 12 내지 40개 염기쌍의 이중-가닥 DNA 서열을 인식하고 절단하고 일반적으로 5개의 계열로 그룹화된다. 일부 실시형태에서, 메가뉴클레아제는 LAGLIDADG 계열, GIY-YIG 계열, HNH 계열, His-Cys 박스 계열 및 PD-(D/E)XK

계열로부터 선택된다. 일부 실시형태에서, 메가뉴클레아제의 DNA 결합 도메인은 그의 동족 표적 서열 이외의 서열을 인식하고 이에 결합하도록 조작된다. 일부 실시형태에서, 메가뉴클레아제의 DNA 결합 도메인은 이종 뉴클레아제 도메인에 융합된다. 일부 실시형태에서, 귀소 엔도뉴클레아제와 같은 메가뉴클레아제는 TAL 모듈에 융합되어 "megaTAL" 단백질과 같은 하이브리드 단백질을 창출한다. megaTAL 단백질은 메가뉴클레아제의 DNA 결합 도메인과 TAL 모듈 둘 모두의 표적 서열을 인식함에 의해 개선된 DNA 표적화 특이성을 가질 수 있다.

[0926] **G. 약학적 조성물 및 제형**

[0927] 본 명세서에 기재된 임의의 재조합 광견병 바이러스 게놈 및 재조합 광견병 바이러스를 포함하는 조성물(예를 들어, 약학적 조성물)이 본 명세서에 제공된다. 본 명세서에서 사용된 바와 같은 용어 "약학적 조성물"은 약학적 용도로 제형화된 조성물을 지칭한다. 특정 실시형태에서, 약학적 조성물은 약학적으로 허용가능한 담체를 추가로 포함한다. 특정 실시형태에서, 약학적 조성물은 (예를 들어, 특정 전달, 반감기 증가 또는 기타 치료적 화합물을 위해) 추가 작용제를 포함한다.

[0928] 본 명세서에 사용된 바와 같은 용어 "약학적으로 허용가능한 담체"는 신체의 한 부위(예를 들어, 전달 부위)로부터 다른 부위(예를 들어, 신체의 표적 기관, 조직 또는 일부)로 화합물(예를 들어, 본 명세서에 기재된 재조합 광견병 바이러스 게놈 또는 재조합 광견병 바이러스)을 수송하거나 운반하는데 관여하는, 액체 또는 고체 충전제, 희석제, 부형제, 제조 보조제(예를 들어, 윤활제, 활석 마그네슘, 칼슘 또는 아연 스테아레이트, 또는 스테릭산), 또는 용매 캡슐화 물질과 같은 약학적으로-허용가능한 물질, 조성물 또는 비히클을 지칭한다. 약학적으로 허용가능한 담체는 제형의 다른 성분과 상용성이 있고 대상체의 조직에 해를 끼치지 않는다(예를 들어, 생리학적으로 적합하고, 멸균되고, 생리학적 pH 등)는 의미에서 "허용가능"하다.

[0929] 약학적으로-허용가능한 담체로 작용할 수 있는 물질의 일부 비제한적인 예는: (1) 당류, 예컨대 락토스, 글루코스 및 수크로스; (2) 전분류, 예컨대 옥수수 전분, 감자 전분; (3) 셀룰로오스 및 그 유도체, 예컨대 나트륨 카르복시메틸 셀룰로오스, 메틸셀룰로오스, 에틸 셀룰로오스, 미세결정질 셀룰로오스 및 셀룰로오스 아세테이트; (4) 분말 트라가칸스; (5) 맥아; (6) 젤라틴; (7) 윤활제, 예컨대 스테아르산마그네슘, 라우릴황산나트륨, 활석; (8) 부형제, 예컨대 코코아 버터 및 좌약 왁스; (9) 오일류, 예컨대 땅콩유, 면실유, 홍화유, 참기름, 올리브유, 옥수수유 및 대두유; (10) 글리콜, 예컨대 프로필렌 글리콜; (11) 폴리올, 예컨대 글리세린, 소르비톨, 만니톨 및 폴리에틸렌 글리콜(PEG); (12) 에스테르, 예컨대 에틸 올레이트 및 에틸 라우레이트; (13) 한천; (14) 완충제, 예컨대 수산화마그네슘, 및 수산화알루미늄; (15) 알긴산; (16) 발열원이 없는 물; (17) 등장성 식염수; (18) 링거 용액; (19) 에틸알코올; (20) pH 완충 용액; (21) 폴리에스테르, 폴리카보네이트 및/또는 폴리무수물; (22) 증량제, 예컨대 폴리펩티드 및 아미노산 (23) 혈청 알코올, 예컨대 에탄올; 및 (23) 약학적 제형에 이용되는 기타 비-독성 호환성 물질을 포함한다. 습윤제, 착색제, 이형제, 코팅제, 감미제, 향미제, 향료, 방부제 및 항산화제도 제형에 존재할 수 있다. "부형제", "담체", "약학적으로 허용가능한 담체", "비히클" 등과 같은 용어는 본 명세서에서 상호교환적으로 사용된다.

[0930] 약학적 조성물은 약 5.0 내지 약 8.0 범위에서와 같이, 생리학적 pH를 반영하는 미리결정된 수준에서 제형의 pH를 유지하기 위해 하나 이상의 pH 완충 화합물을 포함할 수 있다. 수성 액체 제형에 사용되는 pH 완충 화합물은 히스티딘과 같은 아미노산, 또는 히스티딘 및 글리신과 같은 아미노산의 혼합물일 수 있다. 대안적으로, pH 완충 화합물은 바람직하게는 제형의 pH를 미리결정된 수준, 예컨대 약 5.0 내지 약 8.0 범위로 유지하고 칼슘 이온을 킬레이트화하지 않는 작용제이다. 이러한 pH 완충 화합물의 예시적인 예는 이미다졸 및 아세테이트 이온을 포함하지만 이에 제한되지 않는다. pH 완충 화합물은 제형의 pH를 미리결정된 수준으로 유지하는데 적합한 임의의 양으로 존재할 수 있다.

[0931] 약학적 조성물은 또한 하나 이상의 삼투 조절제, 즉 제형의 삼투 특성(예를 들어, 긴장성, 오스몰농도 및/또는 삼투압)을 수령 개체의 혈류 및 혈액 세포에 허용가능한 수준으로 조절하는 화합물을 함유할 수 있다. 삼투압 조절제는 칼슘 이온을 킬레이트화하지 않는 작용제일 수 있다. 삼투압 조절제는 제형의 삼투 특성을 조절하는 당업자에게 공지되거나 이용가능한 임의의 화합물일 수 있다. 당업자는 발명의 제형에 사용하기 위한 주어진 삼투압 조절제의 적합성을 경험적으로 결정할 수 있다. 적합한 유형의 삼투압 조절제의 예시적인 예는 염화나트륨 및 아세트산나트륨과 같은 염; 수크로스, 덱스트로스, 만니톨과 같은 당류; 글리신과 같은 아미노산; 및 이들 작용제 및/또는 작용제의 유형 중 하나 이상의 혼합물을 포함하지만 이에 제한되지 않는다. 삼투 조절제(들)는 제형의 삼투 특성을 조절하기에 충분한 임의의 농도로 존재할 수 있다.

[0932] 특정 실시형태에서, 약학적 조성물은, 예를 들어 유전자 요법을 위해, 대상체에게 전달하기 위해 제형화된다. 본 명세서에 기재된 약학적 조성물을 투여하는 적합한 경로는, 제한 없이: 국소, 피하, 경피, 피내, 병변내, 관

절내, 복강내, 방광내, 경점막, 치은, 치아내, 와우내, 고실경유, 기관내, 경막외, 척수강내, 근육내, 정맥내, 혈관내, 골내, 눈주위, 종양내, 뇌내 및 뇌실내 투여를 포함한다.

[0933] 특정 실시형태에서, 본 명세서에 기재된 약학적 조성물은 질환이 있는 부위(예를 들어, 종양 부위)에 국소적으로 투여된다. 특정 실시형태에서, 본 명세서에 기재된 약학 조성물은 주사, 카테터의 수단, 좌약의 수단, 또는 실라스틱 막과 같은 막 또는 섬유를 포함한, 다공성, 비-다공성 또는 젤라틴성 물질로 되는 임플란트에 의해 대상체에게 투여된다.

[0934] 특정 실시형태에서, 본 명세서에 기재된 약학적 조성물은 제어된 방출 시스템에서 전달된다. 특정 실시형태에서, 펌프가 사용될 수 있다(예를 들어, Langer, 1990, Science 249: 1527-1533; Sefton, 1989, CRC Crit. Ref. Biomed. Eng. 14:201; Buchwald et al, 1980, Surgery 88:507; Saudek 등, 1989, N. Engl. J. Med. 321:574 참조). 특정 실시형태에서, 중합체 물질이 사용될 수 있다. 예를 들어, Medical Applications of Controlled Release (Langer and Wise eds., CRC Press, Boca Raton, Fla., 1974); Controlled Drug Bioavailability, Drug Product Design and Performance (Smolen and Ball eds., Wiley, New York, 1984); Ranger 및 Peppas, 1983, Macromol. Sci. Rev. Macromol. Chem. 23:61 참조. 또한, Levy et al, 1985, Science 228: 190; During et al, 1989, Ann. Neurol. 25:351; Howard et ah, 1989, J. Neurosurg. 71: 105 참조. 기타 제어된 방출 시스템이, 예를 들어, 상기의 Langer에 논의되어 있다.

[0935] 특정 실시형태에서, 약학적 조성물은 대상체, 예를 들어 인간에 대한 정맥내 또는 피하 투여를 위해 적응된 조성물로서 일상적인 절차에 따라 제형화된다. 특정 실시형태에서, 주사에 의한 투여를 위한 약학적 조성물은 가용화제로서 사용되는 멸균 등장액 및 주사의 부위에서 통증을 완화하기 위한 리그노카인과 같은 국소 마취제에서의 용액이다.

[0936] 일반적으로, 성분은 개별적으로 공급되거나, 예를 들어 활성제의 양을 표시하는 앰플 또는 봉지와 같이 기밀하게 밀봉된 용기에서의 건조 동결건조 분말 또는 물이 없는 농축물로서, 단위 복용량 형태로 함께 혼합되어 공급된다.

[0937] 주입에 의해 약품이 투여되어 지는 경우, 그것은 멸균 약품 등급의 물이나 식염수를 함유하는 주입병에 분배될 수 있다. 약학적 조성물이 주사에 의해 투여되는 경우, 투여 이전에 성분을 혼합할 수 있도록 멸균 주사용수 또는 식염수의 앰플을 제공할 수 있다.

[0938] 전신 투여용 약학적 조성물은 액체, 예를 들어 멸균 식염수, 락테이트의 링거액 또는 행크액일 수 있다. 부가하여, 약학적 조성물은 고체 형태로 될 수 있고, 사용 직전에 재-용해되거나 현탁될 수 있다. 동결건조된 형태도 고려된다. 약학적 조성물은 리포솜 또는 미세결정과 같은 지질 입자 또는 소포 내에 함유될 수 있으며, 이는 또한 비경구 투여에도 적합하다. 입자는 조성물이 그 안에 함유되어 있는 한 단일층모양 또는 복수층모양과 같은 임의의 적합한 구조일 수 있다. 화합물은 융합원성 지질 디올레오일포스파티딜에탄올아민(DOPE), 낮은 수준(5-10 mol%)의 양이온성 지질을 함유하고 폴리에틸렌글리콜(PEG) 코팅에 의해 안정화된 "안정화된 플라스미드-지질 입자"(SPLP)에 포획될 수 있다(예를 들어, Zhang Y. P. 등, Gene Ther. 1999, 6: 1438-47 참조). 1,2-디올레오일-3-트리메틸암모늄-프로판 또는 "DOTAP"와 같은 양성으로 하전된 지질이 이러한 입자 및 소포에 특히 바람직하다. 이러한 지질 입자의 제조는 잘 알려져 있다. 예를 들어, 미국 특허 번호 4,880,635; 4,906,477; 4,911,928; 4,917,951; 4,920,016; 및 4,921,757을 참조하며; 이들 각각은 참조로 본 명세서에 포함된다.

[0939] 본 명세서에 기재된 약학적 조성물은 단위 용량으로 투여되거나 포장될 수 있다. 본 개시내용의 약학적 조성물에 관하여 사용될 때 용어 "단위 용량"은 대상체에 대한 단위의 복용량으로 적합한 물리적으로 분리된 단위를 지칭하며, 각 단위는 필요한 희석제; 즉, 담체 또는 비히클과 회합하여 원하는 치료적 효과를 생성하도록 계산된 미리결정된 양의 활성 물질을 함유한다.

[0940] 더욱이, 약학적 조성물은 (a) 동결건조된 형태에서의 발명의 화합물을 함유하는 용기 및 (b) 약학적으로 허용가능한 희석제(예를 들어, 멸균, 발명의 동결건조된 화합물의 재구성 또는 희석에 사용됨)를 함유하는 제2 용기를 포함하는 약학적 키트로서 제공될 수 있다. 선택적으로 그러한 용기와 관련된 것은) 의약품 또는 생물학적 생성물의 제조, 사용 또는 판매를 규제하는 정부 기관이 규정한 형식의 통지일 수 있으며, 이 통지는 인간 투여를 위한 제조, 사용 또는 판매 기관의 승인을 반영한다.

[0941] 또 다른 양태에서, 상기 기재된 질환의 치료에 유용한 물질을 함유하는 제조의 물품이 포함된다. 특정 실시형태에서, 제조의 물품은 용기와 라벨을 포함한다. 적합한 용기는 예를 들어 병, 바이알, 주사기 및 시험관을 포함한다. 용기는 유리나 플라스틱과 같은 다양한 재료로 만들어질 수 있다. 특정 실시형태에서, 용기는 질환을 치

료하는데 효과적이며 멸균 접근 포트를 가질 수 있는 조성물(예를 들어, 본 명세서에 기재된 재조합 광견병 바이러스 게놈 또는 재조합 광견병 바이러스)을 담는다. 예를 들어, 용기는 피하 주사 바늘에 의해 뚫을 수 있는 정맥주사 용액 백 또는 마개를 갖는 바이알일 수 있다. 조성물에서 활성제는 개시내용의 화합물(예를 들어, 재조합 광견병 바이러스 게놈 또는 재조합 광견병 바이러스)이다. 특정 실시형태에서, 용기 상의 또는 용기와 연관된 표지는 조성물이 선택된 질환을 치료하는데 사용된다는 것을 나타낸다. 제조의 물품은 인산염-완충 식염수, 링거액 또는 덱스트로스 용액과 같은 약학적으로-허용가능한 완충액을 포함하는 제2 용기를 추가로 포함할 수 있다. 이는 다른 완충제, 희석제, 필터, 바늘, 주사기 및 사용 지침을 갖는 패키지 삽입물을 포함하여 상업적 및 사용자 관점에서 바람직한 다른 물질을 추가로 포함할 수 있다.

[0942] 일부 실시형태에서, 본 명세서에 기재된 임의의 재조합 광견병 바이러스 게놈 또는 재조합 광견병 바이러스는 약학적 조성물의 일부로 제공된다. 일부 실시형태에서, 약학적 조성물은 본 명세서에 기재된 임의의 재조합 광견병 바이러스 게놈 또는 재조합 광견병 바이러스를 포함한다. 일부 실시형태에서, 약학적 조성물은 본 명세서에 제공된 임의의 복합체를 포함한다.

[0943] 일부 실시형태에서, 본 명세서에 제공된 조성물은 대상체 내에서 표적화된 게놈 변형을 수행하기 위해 대상체, 예를 들어 인간 대상체에게 투여된다. 일부 실시형태에서, 세포는 대상체로부터 얻어지고 본 명세서에 제공된 임의의 약학적 조성물과 접촉된다. 일부 실시형태에서, 대상체로부터 제거되고 생체외에서 약학적 조성물과 접촉된 세포는 선택적으로 원하는 게놈 변형이 세포에서 수행되거나 검출된 후에 대상체에 재-도입된다. 뉴클레아제를 포함하는 약학적 조성물을 전달하는 방법은 공지되어 있고, 예를 들어 미국 특허 번호 6,453,242; 6,503,717; 6,534,261; 6,599,692; 6,607,882; 6,689,558; 6,824,978; 6,933,113; 6,979,539; 7,013,219; 및 7,163,824에 기재되어 있으며; 이들 모두의 개시내용은 그 전체가 본 명세서에 참고로 포함된다. 본 명세서에 제공된 약학적 조성물의 설명은 주로 인간에게 투여하기에 적합한 약학적 조성물에 대한 것이지만, 당업자는 이러한 조성물이 일반적으로 모든 종류의 동물 또는 유기체에 투여하기에 적합하다는 것을 이해할 것이다. 조성물을 다양한 동물에 대한 투여에 적합하도록 하기 위해 인간에 대한 투여에 적합한 약학적 조성물의 변형은 잘 이해되고, 숙련된 수의 약리학자는 필요하면 단지 일반적인 실험을 통해 이러한 변형을 설계 및/또는 수행할 수 있다. 약학적 조성물의 투여가 고려되는 대상체는 인간 및/또는 다른 영장류; 포유동물, 가축, 애완동물 및 소, 돼지, 말, 양, 고양이, 개, 마우스 및/또는 랫트와 같은 상업적으로 관련된 포유동물; 및/또는 닭, 오리, 거위 및/또는 칠면조와 같은 상업적으로 관련된 새를 포함한 새를 포함하지만 이에 제한되지 않는다.

[0944] 본 명세서에 기재된 약학적 조성물의 제형은 약리학의 분야에 공지되어 있거나 향후 개발될 임의의 방법에 의해 제조될 수 있다. 일반적으로, 이러한 제조의 방법은 활성 성분(들)을 부형제 및/또는 하나 이상의 다른 보조 성분과 회합하여 결합시키는 단계, 및 그 다음 필요한 경우 및/또는 바람직한 경우 생성물을 원하는 단일- 또는 다중-용량 단위 안으로 성형 및/또는 포장하는 단계를 포함한다.

[0945] 약학적 제형은 추가로 약학적으로 허용가능한 부형제를 포함할 수 있으며, 이는 본 명세서에 사용된 바와 같이 임의의 및 모든 용매, 분산 매질, 희석제 또는 기타 액체 비히클, 분산 또는 현탁 보조제, 표면 활성제, 등장화제, 증점제 또는 유화제, 방부제, 고체 결합제, 윤활제 등을, 원하는 특정 복용량 형태에 적합하게 포함한다. Remington's The Science and Practice of Pharmacy, 21st Edition, A. R. Gennaro(Lippincott, Williams & Wilkins, Baltimore, MD, 2006; 그 전체가 참고로 본 명세서에 포함됨)는 약학적 조성물을 제형화하는데 사용되는 다양한 부형제 및 이의 제조를 위한 공지된 기술을 개시한다. 뉴클레아제를 포함하는 약학적 조성물을 생성하기 위한 추가의 적합한 방법, 시약, 부형제 및 용매에 대해서는 그 전체가 참고로 본 명세서에 포함된 PCT 출원 PCT/US2010/055131(공개 번호 WO2011053982 A8, 2010년 11월 2일 출원됨)을 또한 참조한다. 임의의 바람직하지 않은 생물학적 효과를 생성하거나 그렇지 않으면 약학 조성물의 임의의 다른 성분(들)과 유해한 방식으로 상호작용함에 의한 것과 같이 임의의 통상적인 부형제 매질이 물질 또는 그 유도체와 양립할 수 없는 경우를 제외하고, 그 사용은 본 개시내용의 범주 내에 있는 것으로 고려된다. 특정 실시형태에서, 본 발명에 따른 조성물은 임의의 다양한 질환, 장애 및/또는 병태의 치료에 사용될 수 있다.

[0946] 본 개시내용의 다양한 양태는 달리 명시되지 않는 한, 당업자의 범위 내에 있는 분자 생물학(재조합 기술 포함), 미생물학, 세포 생물학, 생화학 및 면역학의 통상적인 기술을 이용한다. 이러한 기술은 문헌, 예컨대 "Molecular Cloning: A Laboratory Manual", second edition (Sambrook, 1989); "Oligonucleotide Synthesis" (Gait, 1984);"Animal Cell Culture" (Freshney, 1987); "Methods in Enzymology," 및 "Handbook of Experimental Immunology" (Weir, 1996); "Gene Transfer Vectors for Mammalian Cells" (Miller 및 Calos, 1987); "Current Protocols in Molecular Biology" (Ausubel, 1987); "PCR: The Polymerase Chain Reaction", (Mullis, 1994); "Current Protocols in Immunology" (Coligan, 1991)에 완전하게 설명되어 있다. 이들 기술은

본 개시내용의 다양한 양태의 생산에 적용가능하고, 따라서 이를 제조하고 실시하는 데 고려될 수 있다.

[0947] **H. 폴리뉴클레오티드, 벡터 및 세포**

[0948] (i) 본 명세서에 기재된 재조합 광견병 바이러스 게놈; (ii) 광견병 바이러스 핵단백질 또는 이의 기능적 변이체를 인코딩하는 N 유전자; (iii) 광견병 바이러스 인단백질 또는 이의 기능적 변이체를 인코딩하는 P 유전자; (iv) 광견병 바이러스 폴리머라제(예를 들어, RNA-의존성 RNA 폴리머라제) 또는 이의 기능적 변이체를 인코딩하는 L 유전자; (v) 광견병 바이러스 당단백질 또는 이의 기능적 변이체를 인코딩하는 G 유전자; 및/또는 (vi) 광견병 바이러스 매트릭스 단백질 또는 이의 기능적 변이체를 인코딩하는 M 유전자를 포함하는 폴리뉴클레오티드가 본 명세서에 제공된다.

[0949] 본 명세서에 기재된 폴리뉴클레오티드는 당업계에 공지된 임의의 방법에 의해, 예컨대 DNA 사슬을 화학적 합성함에 의해, PCR에 의해 또는 깁슨 어셈블리 방법에 의해 얻어질 수 있다. 화학적 합성 또는 PCR 방법 또는 깁슨 어셈블리 방법의 조합에 의해 전장 DNA를 구축하는 이점은 융합 단백질이 숙주 세포에서 높은 수준으로 발현되는 것을 보장하도록 코돈이 최적화될 수 있다는 것이다. 최적화된 코돈은 카즈사 DNA 연구소의 홈페이지에 공개된 유전적 코드 사용 빈도 데이터베이스(http://www.kazusa.or.jp/codon/index.html)를 사용하여 선택할 수 있다. 특정 실시형태에서, 폴리뉴클레오티드는 코돈 최적화된다. 특정 실시형태에서, 폴리뉴클레오티드는 RNA 최적화에 의해 최적화될 수 있다. 예를 들어 외인성 신호 서열로 신호 서열의 대체, 불안정성 요소의 제거, 억제 영역의 제거, 잠재적인 스플라이스 부위의 제거, 및 당업계의 통상인에게 공지된 기타 최적화 방법을 포함하여, 재조합 발현에 대한 안정성을 증가시키기 위해 추가적인 최적화 방법이 포함될 수 있다. 예를 들어, 미국 특허 번호 6,794,498을 참조하며, 그 개시내용은 그 전체가 본 명세서에 참고로 포함된다.

[0950] 일단 획득되면, 본 개시내용의 폴리뉴클레오티드는 적합한 발현 벡터 안으로 합체될 수 있다. 따라서, 본 개시내용은 또한 본 명세서에 개시된 임의의 폴리뉴클레오티드를 별도로 또는 조합하여 포함하는 벡터를 제공한다. 적합한 벡터는 플라스미드, 바이러스, 인공 염색체, 백미드, 코스미드 및 당업계의 통상인에게 공지된 기타를 포함한다. 특정 실시형태에서, 벡터는 발현 벡터이다.

[0951] 적합한 발현 벡터는 *대장균*-유래된 플라스미드(예를 들어, pBR322, pBR325, pUC12, pUC13); *바실러스 서브틸리스*-유래된 플라스미드(예를 들어, pUB110, pTP5, pC194); 효모-유래된 플라스미드(예를 들어, pSH19, pSH15); 곤충 세포에서의 발현에 적합한 플라스미드(예를 들어, pFast-Bac); 포유동물 세포에서의 발현에 적합한 플라스미드(예를 들어, pXT1, pRc/CMV, pRc/RSV, pcDNA1/Neo); 또한 람다 파지 등과 같은 박테리오파지를 포함하며; 사용될 수 있는 다른 벡터는 곤충 바이러스 벡터, 예컨대 바큘로바이러스 등(예를 들어, BmNPV, AcNPV); 및 레트로바이러스, 백시니아 바이러스, 아데노바이러스 등과 같은 포유동물 세포에서의 발현에 적합한 바이러스 벡터를 포함한다.

[0952] 폴리뉴클레오티드 및 벡터를 포함하는 유전자 및/또는 이식유전자는 전형적으로 전사 조절 요소의 제어 하에 발현된다. 특정 실시형태에서, 전사 조절 요소는 하나 이상의 인핸서 요소, 인트론 요소, 및/또는 프로모터 요소를 포함할 수 있다. 특정 실시형태에서, 전사 조절 요소는 구성적 프로모터를 포함한다. 전사 조절 요소의 예는 선택적으로 CMV 인핸서, EF1α 프로모터(인간 신장 인자 1 알파로부터의 프로모터), CBA 프로모터(CMV 초기 인핸서 및 치킨 β-액틴 프로모터를 포함함), CAG 프로모터(CBA 프로모터 및 토끼 β-글로빈 인트론을 포함함), CAGGS 프로모터(CMV 인핸서, CBA 프로모터, 및 치킨 β-액틴 인트론 1/엑손 1을 포함함), PGK 프로모터 (포스포글리세레이트 키나제로부터의 프로모터), U6 프로모터(U6 핵 프로모터), Ubc 프로모터(인간 유비퀴틴 C로부터의 프로모터), CASI 프로모터(CMV 인핸서, 유비퀴틴 C 인핸서, 및 치킨 β-액틴 프로모터를 포함함), 및 CALM1 프로모터(칼모듈린 1로부터의 프로모터)를 포함하는 CMV 프로모터(인간 사이토메갈로바이러스로부터의 프로모터)를 포함하는 것들을 포함한다. 다양한 구성적 전사 조절 요소가 당업계의 통상인에게 공지되어 있다.

[0953] 특정 실시형태에서, 전사 조절 요소는 유도성 프로모터를 포함한다. 예를 들어, 전사 조절 요소는 유도성 TRE 프로모터(테트라사이클린 반응 요소 프로모터)를 포함할 수 있다. 이러한 유도성 프로모터는 활성제 단백질이 프로모터에 결합할 수 없기 때문에 프로모터가 불활성인 경우 양성 유도성일 수 있거나, 또는 억제인자 단백질이 전사를 방지하는 프로모터에 결합하는 음성 유도성일 수 있다. 유도성 프로모터의 예는 화학적으로 유도성인 것, 예를 들어 테트라사이클린 ON(Tet-On) 프로모터 시스템, lac 억제인자 시스템, pBad 원핵생물 프로모터 및 기타 예컨대 알코올 또는 스테로이드 조절된 프로모터를 포함한다. 유도성 프로모터는 온도 유도성, 예를 들어 열 또는 냉 유도된 프로모터(예를 들어, Hsp70 또는 Hsp90-유래된 프로모터) 및 광이 전사를 조절하는데 사용될 수 있는 광 유도성일 수 있다. 특정 실시형태에서, 전사 조절 요소는 억제성 프로모터를 포함한다. 다양한 유도성 전사 조절 요소가 당업계의 통상인에게 알려져 있다.

[0954] 특정 실시형태에서, 전사 조절 요소는 유전자 또는 이식유전자에 외인성인 프로모터를 포함한다. 특정 실시형태에서, 전사 조절 요소는 합성 프로모터를 포함한다.

[0955] 적합한 프로모터는 원하는 숙주 세포에서의 발현을 위한 그 용도에 기반하여 선택될 수 있다. 예를 들어, 숙주가 동물세포인 경우에는, 다음 프로모터 중 임의의 하나가 사용된다: SR-알파 프로모터, SV40 프로모터, LTR 프로모터, CMV(사이토메갈로바이러스) 프로모터, RSV(라우스 육종 바이러스) 프로모터, MoMuLV(몰로니 마우스 백혈병 바이러스) LTR, HSV-TK(단순 헤르페스 바이러스 티미딘 키나제) 프로모터 등이 사용된다. 특정 실시형태에서, 프로모터는 CMV 프로모터 또는 SR 알파 프로모터이다. 특정 실시형태에서, 프로모터는 신장 인자 1-알파(EF1α) 프로모터이다. 숙주 세포가 대장균인 경우, 임의의 다음 프로모터가 사용될 수 있다: trp 프로모터, lac 프로모터, recA 프로모터, 람다PL 프로모터, lpp 프로모터, T7 프로모터 등. 숙주가 바실러스속인 경우, 임의의 다음 프로모터가 사용될 수 있다: SPO1 프로모터, SPO2 프로모터, penP 프로모터 등. 숙주가 효모인 경우, 임의의 다음 프로모터가 사용될 수 있다: Gal1/10 프로모터, PHO5 프로모터, PGK 프로모터, GAP 프로모터, ADH 프로모터 등. 숙주가 곤충 세포인 경우, 임의의 다음 프로모터가 사용될 수 있다: 폴리헤드린 프로모터, P10 프로모터 등. 숙주가 식물 세포인 경우, 임의의 다음 프로모터가 사용될 수 있다: CaMV35S 프로모터, CaMV19S 프로모터, NOS 프로모터 등.

[0956] 원하는 경우, 발현 벡터는 또한 인핸서, 스플라이싱 신호, 터미네이터, 폴리아데닐화 신호, 선택 마커(예를 들어, 약물 저항성 유전자, 영양요구성 상보성 유전자 등) 또는 복제 기원 중 하나 이상을 포함한다.

[0957] 본 개시내용의 폴리뉴클레오티드는 당업자에게 공지된 일상적인 방법을 사용하여 박테리아, 효모, 진균, 곤충, 식물 및 동물 세포를 포함하지만 이에 제한되지 않는, 사실상 임의의 관심있는 숙주 세포 안으로 도입될 수 있다.

[0958] *에스케리키아* 속은 *대장균* K12/DH1, *대장균* JM103, *대장균* JA221, *대장균* HB101, *대장균* C600 등을 포함한다. *바실러스* 속은 *바실러스 서브틸리스* M114, *바실러스 서브틸리스* 207-21 등을 포함한다.

[0959] 본 개시내용의 폴리뉴클레오티드를 숙주시키는데 유용한 효모는 *사카로마이세스 세레비지애* AH22, AH22 R$^-$, NA87-11A, DKD-5D, 20B-12, *스키조사카로마이세스 폼베* NCYC1913, NCYC2036, *피치아 파스토리스* KM71 등을 포함한다.

[0960] 본 개시내용의 폴리뉴클레오티드는, 예를 들어, AcNPV와 같은 바이러스 벡터를 사용하여 곤충 세포 안으로 도입될 수 있다. 곤충 숙주 세포는 임의의 다음 세포주를 포함한다: 양배추 거염벌레 유충-유래된 정착주(*스포도프테라 프루기페르다* 세포; Sf 세포), *트리코플루시아니*의 중장으로부터 유래된 MG1 세포, High Five, *트리코플루시아니*의 난으로부터 유래된 세포, *마메스트라 브라시카에*-유래된 세포, *에스티그메나 아크라즈*-유래된 세포 등. 바이러스가 BmNPV인 경우에는 *봄빅스 모리*-유래된 세포주(*봄빅스 모리* N 세포; BmN 세포) 등이 사용된다. Sf 세포는 예를 들어 Sf9 세포(ATCC CRL1711), Sf21s 세포 등을 포함한다.

[0961] 제한 없이, 원숭이 COS-7 세포, 원숭이 Vero 세포, 차이니즈 햄스터 난소(CHO) 세포, dhfr 유전자-결핍 CHO 세포, 마우스 L 세포, 마우스 AtT-20 세포, 마우스 골수종 세포, 랫트 GH3 세포, 인간 FL 세포, 인간 배아 신장(HEK) 세포(예를 들어, HEK293, HEK293T), COS 세포(*예를 들어*, COS1 또는 COS), BHK 세포, MDCK 세포, NS0 세포, PER.C6 세포, CRL7030 세포, HsS78Bst 세포, HeLa 세포, NIH 3T3 세포, HepG2 세포, SP210 세포, R1.1 세포, B-W 세포, L-M 세포, BSC1 세포, BSC40 세포, YB/20 세포 및 BMT10 세포 등을 포함하는 포유동물 세포주가 사용될 수 있다.

[0962] 특정 실시형태에서, 적합한 세포는 포유동물, 박테리아 또는 곤충 기원의 것이다. 특정 실시형태에서, 세포는 HEK293 세포, HEK293T 세포, VERO 세포, BHK 세포 및 BSR 세포로 구성된 군으로부터 선택된다.

[0963] 상기-언급된 모든 숙주 세포는 반수체(단배체) 또는 배수체(예를 들어, 이배체, 삼배체, 사배체 등)일 수 있다.

[0964] 본 명세서에 기재된 숙주 세포 안으로 개시내용의 폴리뉴클레오티드를 도입하는 다양한 방법은 당업계의 통상인에게 공지되어 있다. 예를 들어, 이러한 방법은 당업계에 공지된 임의의 형질감염 방법의 사용(예를 들어, 리소자임, PEG, CaCl2 공침, 전기천공, 미세주입, 입자 총, 리포펙션, 아그로박테리움 등을 사용함)을 포함할 수 있다. 형질감염 방법은 형질감염되어 지는 숙주 세포에 기반하여 선택된다. *대장균*은 예를 들어 Proc. Natl. Acad. Sci. USA, 69, 2110 (1972), Gene, 17, 107 (1982) 등에 기재된 방법에 따라 형질전환된다. *바실러스* 속을 형질도입하는 방법은 예를 들어 Molecular & General Genetics, 168, 111 (1979)에 기재되어 있다. 효모 세포는 예를 들어 Methods in Enzymology, 194, 182-187 (1991), Proc. Natl. Acad. Sci. USA, 75, 1929 (1978)

등에 기재된 방법을 사용하여 형질도입된다. 곤충 세포는 예를 들어 Bio/Technology, 6, 47-55(1988) 등에 기재된 방법을 사용하여 형질감염된다. 포유동물 세포는 예를 들어 Cell Engineering additional volume 8, New Cell Engineering Experiment Protocol, 263-267 (1995) (published by Shujunsha), 및 Virology, 52, 456 (1973)에 기재된 방법을 사용하여 형질감염된다.

[0965] 본 개시내용의 발현 벡터를 포함하는 세포는 숙주에 따라 달라지는 공지된 방법에 따라 배양된다. 예를 들어, *대장균* 또는 *바실러스* 속 세포를 배양하는 경우에는 액체 배지를 사용한다. 배지는 바람직하게는 형질전환체의 성장에 필요한 탄소원, 질소원, 무기 물질 및 기타 성분을 함유한다. 탄소원의 예는 글루코스, 덱스트린, 가용성 전분, 수크로스 등을 포함하고; 질소원의 예는 암모늄염, 질산염, 옥수수 침지액, 펩톤, 카세인, 고기 추출물, 콩박, 감자 추출물 등과 같은 무기 또는 유기 물질을 포함하고; 무기 물질의 예는 염화칼슘, 인산이수소나트륨, 염화마그네슘 등을 포함한다. 배지는 또한 효모 추출물, 비타민, 성장 촉진 인자 등을 함유할 수 있다. 배지의 pH는 바람직하게는 약 5 내지 약 8이다. *대장균*, 예를 들어 M9를 배양하기 위한 배지로서 글루코스 및 카사미노산을 함유하는 배지(예를 들어, Journal of Experiments in Molecular Genetics, 431-433, Cold Spring Harbor Laboratory, New York 1972 참조)가 사용된다. *대장균*은 일반적으로 약 15 내지 43℃에서 배양된다. 필요한 경우 통기 및 교반이 수행될 수 있다. *바실러스* 속은 일반적으로 약 30 내지 40℃에서 배양된다. 필요한 경우 통기 및 교반이 수행된다.

[0966] 효모를 배양하는데 적합한 배지의 예는 Burkholder 최소 배지, 0.5% 카사미노산을 함유하는 SD 배지 등을 포함한다. 배지의 pH는 바람직하게는 약 5 - 약 8이다. 배양은 일반적으로 약 20℃ 내지 약 35℃에서 수행된다. 필요한 경우 통기 및 교반이 수행될 수 있다.

[0967] 곤충 세포나 곤충을 배양하기 위한 배지로는 불활성화된 10% 소혈청과 같은 첨가제를 함유하는 Grace's 곤충 배지 등이 사용된다. 배지의 pH는 바람직하게는 약 6.2 내지 약 6.4이다. 세포는 약 27℃에서 배양된다. 필요한 경우 통기 및 교반이 수행될 수 있다.

[0968] 포유동물 세포는, 예를 들어 소 태아 혈청 약 5 내지 약 20%를 함유하는 최소 필수 배지(MEM), 둘베코 변형 이글 배지(DMEM), RPMI 1640 배지, 199 배지 등 중 임의의 하나에서 배양된다. 배지의 pH는 바람직하게는 약 6 내지 약 8이다. 배양은 약 30℃ 내지 약 40℃에서 수행된다. 필요한 경우 통기 및 교반이 수행될 수 있다.

[0969] **I. 패키징 시스템 및 그 방법**

[0970] 본 개시내용은 본 명세서에 기재된 광견병 바이러스 입자의 재조합 제조에 유용한 패키징 시스템을 제공한다. 특히, 패키징 시스템은 본 명세서에 기재된 광견병 바이러스 입자의 제조에 필요한 필수 구성요소를 제공한다. 특정 실시형태에서, 패키징 시스템은 재조합 광견병 바이러스 게놈을 포함하는 광견병 바이러스 입자의 재조합 제조에 유용하며, 여기서 게놈은 광견병 바이러스 당단백질 또는 이의 기능적 변이체를 인코딩하는 G 유전자를 결하고/결하거나; 게놈은 광견병 바이러스 폴리머라제 또는 이의 기능적 변이체를 인코딩하는 L 유전자를 결한다. 특정 실시형태에서, 패키징 시스템은 재조합 광견병 바이러스 게놈을 포함하는 광견병 바이러스 입자의 재조합 제조에 유용하며, 여기서 게놈은 광견병 바이러스 당단백질 또는 이의 기능적 변이체를 인코딩하는 G 유전자를 결한다. 특정 실시형태에서, 패키징 시스템은 재조합 광견병 바이러스 게놈을 포함하는 광견병 바이러스 입자의 재조합 제조에 유용하며, 여기서 게놈은 광견병 바이러스 당단백질 또는 이의 기능적 변이체를 인코딩하는 G 유전자를 결하고; 게놈은 광견병 바이러스 폴리머라제 또는 이의 기능적 변이체를 인코딩하는 L 유전자를 결한다.

[0971] 본 명세서에 기재된 패키징 시스템은 일반적으로 (i) 광견병 바이러스 핵단백질 또는 이의 기능적 변이체를 인코딩하는 N 유전자; (ii) 광견병 바이러스 인단백질 또는 이의 기능적 변이체를 인코딩하는 P 유전자; 및 (iii) 광견병 바이러스 폴리머라제 또는 이의 기능적 변이체를 인코딩하는 L 유전자를 포함하거나 이로 구성된다. 특정 실시형태에서, 패키징 시스템은 광견병 바이러스 매트릭스 단백질 또는 이의 기능적 변이체를 인코딩하는 M 유전자를 추가로 포함한다. 특정 실시형태에서, 패키징 시스템은 광견병 바이러스 당단백질 또는 이의 기능적 변이체를 인코딩하는 G 유전자를 추가로 포함한다.

[0972] 패키징 시스템의 N, P 및 L 유전자는 하나 이상의 벡터(예를 들어, 형질감염 플라스미드)에 제공될 수 있다. 예를 들어, 패키징 시스템은 N, P 및 L 유전자 각각에 대한 별도의 형질감염 플라스미드, 예를 들어 광견병 바이러스 핵단백질 또는 이의 기능적 변이체를 인코딩하는 N 유전자를 포함하는 제1 형질감염 플라스미드; 광견병 바이러스 인단백질 또는 이의 기능적 변이체를 인코딩하는 P 유전자를 포함하는 제2 형질감염 플라스미드; 및 광견병 바이러스 폴리머라제 또는 이의 기능적 변이체를 인코딩하는 L 유전자를 포함하는 제3 형질감염 플라스

미드를 포함한다. 특정 실시형태에서, 단일 형질감염 플라스미드는 N, P 및 L 유전자의 2개 이상을 포함한다. 예를 들어, 패키징 시스템은 광견병 바이러스 핵단백질 또는 이의 기능적 변이체를 인코딩하는 N 유전자 및 광견병 바이러스 인단백질 또는 이의 기능적 변이체를 인코딩하는 P 유전자를 포함하는 형질감염 플라스미드를 포함할 수 있으며; 패키징 시스템은 광견병 바이러스 핵단백질 또는 이의 기능적 변이체를 인코딩하는 N 유전자 및 광견병 바이러스 폴리머라제 또는 이의 기능적 변이체를 인코딩하는 L 유전자를 포함하는 형질감염 플라스미드를 포함할 수 있으며; 패키징 시스템은 광견병 바이러스 인단백질 또는 이의 기능적 변이체를 인코딩하는 P 유전자, 및 광견병 바이러스 폴리머라제 또는 이의 기능적 변이체를 인코딩하는 L 유전자를 포함하는 형질감염 플라스미드를 포함할 수 있다. 특정 실시형태에서, 패키징 시스템은 광견병 바이러스 핵단백질 또는 이의 기능적 변이체를 인코딩하는 N 유전자, 광견병 바이러스 인단백질 또는 이의 기능적 변이체를 인코딩하는 P 유전자, 및 광견병 바이러스 폴리머라제 또는 이의 기능적 변이체를 인코딩하는 L 유전자를 포함하는 형질감염 플라스미드를 포함할 수 있다.

[0973] 패키징 시스템의 M 및 G 유전자는 하나 이상의 형질감염 플라스미드에 제공될 수 있다. 특정 실시형태에서, 패키징 시스템은 M 및 G 유전자에 대한 별도의 형질감염 플라스미드를 포함한다. 예를 들어, 특정 실시형태에서, 패키징 시스템은 광견병 바이러스 매트릭스 단백질 또는 이의 기능적 변이체를 인코딩하는 M 유전자를 포함하는 형질감염 플라스미드를 추가로 포함할 수 있다. 특정 실시형태에서, 패키징 시스템은 광견병 바이러스 당단백질 또는 이의 기능적 변이체를 인코딩하는 G 유전자를 포함하는 형질감염 플라스미드를 추가로 포함할 수 있다. M 및/또는 G 유전자는 또한 본 명세서에 기재된 바와 같은 N, P 및/또는 L 유전자를 포함하는 형질감염 플라스미드에 조합될 수 있다. 예를 들어, 단일 형질감염 플라스미드는 광견병 바이러스 핵단백질 또는 이의 기능적 변이체를 인코딩하는 N 유전자, 광견병 바이러스 인단백질 또는 이의 기능적 변이체를 인코딩하는 P 유전자, 광견병 바이러스 폴리머라제 또는 이의 기능적 변이체를 인코딩하는 L 유전자, 광견병 바이러스 매트릭스 단백질 또는 이의 기능적 변이체을 인코딩하는 M 유전자, 및 광견병 바이러스 당단백질 또는 이의 기능적 변이체를 인코딩하는 G 유전자를 포함할 수 있다. 다른 조합이 당업계의 통상인에 의해 다양한 쉽게 이해될 수 있다.

[0974] N, P, L, M 및/또는 G 유전자는 모두 하나 이상의 전사 조절 요소의 제어하에 있을 수 있다. 특정 실시형태에서, 전사 조절 요소는 프로모터 및/또는 인핸서 서열을 포함한다. 특정 실시형태에서, 전사 조절 요소는 EF1α 프로모터를 포함한다. 다양한 프로모터 및/또는 인핸서 서열이 당업계에 공지되어 있고 본 명세서에서 예로서 기재되어 있으며, 당업계의 통상인은 그 필요에 적합한 프로모터 및/또는 인핸서 서열을 선택할 수 있을 것이다.

[0975] N, P, L, M 및/또는 G 유전자의 2개 이상이 동일한 벡터 상에 존재하는 경우, 2개 이상의 유전자는 하나 이상의 발현 카세트에 존재할 수 있다. 예를 들어, N, P, L, M 및/또는 G 유전자 각각은 전사 조절 요소 및/또는 전사 종결 요소를 포함하는 각각 그 자체의 발현 카세트 내에 있을 수 있다.

[0976] 2개 이상의 유전자가 동일한 발현 카세트에 존재하는 경우, 유전자는 링커 서열에 의해 분리될 수 있다. 특정 실시형태에서, 링커 서열은 내부 리보솜 도입 부위(IRES)를 인코딩하는 핵산 서열을 포함하는 리보솜 스키핑 요소이다. 본 명세서에 사용된', "내부 리보솜 도입 부위" 또는 "IRES"는 단백질 코딩 영역의 ATG와 같은 개시 코돈으로의 직접적인 내부 리보솜 도입을 촉진하여 유전자의 캡-독립적 번역을 유도하는 요소를 지칭한다. 제한 없이, 바이러스 또는 세포성 mRNA 공급원으로부터 얻을 수 있는 IRES, 예를 들어, 면역글로블린 중쇄 결합 단백질(BiP); 혈관 내피 성장 인자(VEGF); 섬유아세포 성장 인자 2; 인슐린-유사 성장 인자; 번역 개시 인자 eIF4G; 효모 전사 인자 TFIID 및 HAP4; 및 예를 들어 카디오바이러스, 라이노바이러스, 아프토바이러스, HCV, 프렌드 뮤어라인 백혈병 바이러스(FrMLV) 및 몰로니 뮤어라인 백혈병 바이러스(MoMLV)로부터 얻을 수 있는 IRES를 포함하는, 다양한 내부 리보솜 도입 부위가 당업계의 통상인에게 공지되어 있다. 특정 실시형태에서, 링커 서열은 자가-절단 펩티드를 인코딩하는 핵산 서열을 포함하는 리보솜 스키핑 요소이다. 본 명세서에 사용된', "자가-절단 펩티드" 또는 "2A 펩티드"는 다중 단백질이 번역 시 성분 단백질로 해리하는, 다단백질로서 인코딩되도록 하는 올리고펩티드를 지칭한다. 용어 "자가-절단"의 사용은 단백질분해 절단 반응을 암시하는 것으로 의도되지 않는다. 제한 없이, 피코르나바이러스과 바이러스 계열의 구성원에서 발견되는 것들, 예를 들어 구제역 바이러스(FMDV), 말 비염 A 바이러스(ERAV0, 테세아 아시그나 바이러스(TaV) 및 돼지 테스코 바이러스-1(PTV-1); 및 테일로바이러스 및 뇌심근염 바이러스와 같은 카리오바이러스를 포함한 다양한 자가-절단 또는 2A 펩티드가 당업계의 통상인에게 공지되어 있다. FMDV, ERAV, PTV-1 및 TaV로부터 유래된 2A 펩티드는 본 명세서에서 각각 "F2A", "E2A", "P2A" 및 "T2A"로 지칭된다. 당업계의 통상인은 그 필요에 적절한 링커 서열을 선택할 수 있을 것이다.

[0977] 특정 실시형태에서, 단일 벡터(예를 들어, 형질감염 플라스미드)는 N 및 P 유전자를 포함하는 제1 발현 카세트,

및 L 유전자를 포함하는 제2 발현 카세트를 포함한다. 특정 실시형태에서, 제1 발현 카세트는 5'에서 3'으로: 전사 조절 요소; P 유전자; 그리고 N 유전자를 포함한다. 특정 실시형태에서, 제1 발현 카세트는 5'에서 3'으로: 전사 조절 요소; P 유전자; 리보솜 스키핑 요소; 그리고 N 유전자를 포함한다. 특정 실시형태에서, 제2 발현 카세트는 5'에서 3'으로: 전사 조절 요소; 그리고 L 유전자를 포함한다. 특정 실시형태에서, 제1 발현 카세트 및 제2 발현 카세트는 벡터 내에서 동일한 방향일 수 있다. 특정 실시형태에서, 제1 발현 카세트와 제2 발현 카세트는 벡터 내에서 반대 방향에 있을 수 있다.

[0978] 따라서, 본 개시내용의 패키징 시스템은 (i) 재조합 광견병 바이러스 게놈 벡터(예를 들어, 바이러스 게놈 형질감염 플라스미드); 및 (ii) N, P, L, M 및/또는 G 유전자를 포함하는 하나 이상의 형질감염 플라스미드를 포함한다. N, P, L, M 및/또는 G 유전자를 포함하는 하나 이상의 형질감염 플라스미드는 당업계의 통상인에게 공지된 다양한 방법을 사용하여 숙주 세포(예를 들어, 재조합 광견병 바이러스 입자 패키징 세포) 안으로 도입될 수 있다. 예를 들어, 하나 이상의 형질감염 플라스미드는 전기천공, 뉴클레오펙션 또는 리포펙션에 의해 적합한 숙주 세포 안으로 도입될 수 있다.

[0979] 본 개시내용은 또한 광견병 바이러스 입자의 재조합 제조에 대한 방법을 제공하며, 여기서 방법은 재조합 광견병 바이러스 입자를 형성하기 위해 재조합 광견병 바이러스 게놈을 둘러싸도록 작동하는 조건 하에서 본 명세서에 기재된 패키징 시스템을 세포 안으로 도입하는 것을 포함한다. 특정 실시형태에서, 숙주 패키징 세포는 N, P, L, M 및/또는 G 유전자를 포함하는 하나 이상의 형질감염 플라스미드로 일시적으로 형질감염될 수 있다. 특정 실시형태에서, 숙주 패키징 세포는 N, P, L, M 및/또는 G 유전자를 포함하는 하나 이상의 형질감염 플라스미드로 형질감염될 수 있으며, 여기서 숙주 패키징 세포는 안정한 세포주로 추가로 만들어진다. 안정한 세포주를 생산하는 다양한 방법은 당업계의 통상인에게 공지되어 있다. 일반적으로, 관심있는 유전자(예를 들어, N, P, L, M 및/또는 G 유전자)는 세포 안으로 도입된 다음 세포의 핵 안으로 도입되고 최종적으로 세포의 게놈 안으로 합체된다. 염색체 통합 이벤트는 드물고 안정적으로-합체된 세포주가 선택되고 배양되어지도록 해야 한다. G418, 디하이드로폴레이트 환원효소(DHFR) 또는 글루타민 합성효소에 대한 내성을 부여하는, 네오마이신 포스포트랜스퍼라제와 같은 항생제에 대한 내성을 포함하는 다양한 선택 시스템이 당업계에 공지되어 있다. 안정한 세포주를 생산하기 위한 다른 방법은 실험예에 기술된 바와 같은 슬리핑 뷰티(SB) 시스템의 사용을 포함한다. 간략하게, 관심있는 통합체를 포함하는 트랜스포손은 TA 디뉴클레오티드로의 정확한 합체를 초래하는 역전된 반복/직접 반복 서열을 측접함으로 설계된다. SB 트랜스포존 기반 안정 세포주 생성에 대한 방법은 당업계에 공지되어 있으며, 예를 들어, Davidson 등, *Cold Spring Harb Protoc.* (2009) 4(8): 1018-1023을 참조함다. 안정한 세포주는 또한 렌티바이러스 벡터의 사용을 통해 생성될 수 있으며, 예를 들어, Tandon 등, *Bio Protoc.* (2018) 8(21): e3073을 참조함다.

[0980] 그런 다음 재조합 광견병 바이러스 게놈 벡터(예를 들어, 바이러스 게놈 형질감염 플라스미드)는 N, P, L, M 및/또는 G 유전자가 안정적으로-합체되었거나 일시적으로 그 안에 형질감염된 숙주 패키징 세포 안으로 도입된다.

[0981] 따라서, 특정 실시형태에서, 광견병 바이러스 입자의 재조합 제조를 위한 방법은 (i) 재조합 광견병 바이러스 게놈 벡터(예를 들어, 바이러스 게놈 형질감염 플라스미드); 및 (ii) N, P, L, M 및/또는 G 유전자를 포함하는 하나 이상의 형질감염 플라스미드를 숙주 패키징 세포 안으로 도입하는 것을 포함한다. 특정 실시형태에서, 광견병 바이러스 입자의 재조합 제조를 위한 방법은 재조합 광견병 바이러스 게놈 벡터(예를 들어, 바이러스 게놈 형질감염 플라스미드)를 숙주 패키징 세포 안으로 도입하는 것을 포함하며, 여기서 숙주 패키징 세포는 그 안에 안정적으로 합체된 N, P, L, M 및/또는 G 유전자를 포함한다. 재조합 광견병 바이러스 입자의 제조에 대한 방법은 당업계에 공지되어 있으며, 예를 들어 Trabelsi 등, *Vaccine* (2019) 37(47): 7052-7060; Wickersham 등, *Nature Protoc.* (2010) 5(3): 595-606; Ghanem 등, *Eur. J. Cell Biol.* (2012) 91: 10-16; Osakada 및 Wickersham, *Nature Protoc.* (2013) 8(8): 1583-1601; 및 Sullivan 및 Wickersham, *Cold Spring Harb Protoc.* (2015) 4: 386-91를 참조하며, 이의 개시내용은 그 전체가 본 명세서에 참조로 포함된다.

[0982] 특정 실시형태에서, 본 명세서에 기재된 생산의 방법을 사용하여 획득된 재조합 광견병 바이러스 입자 역가는 약 1E8 형질도입 단위(TU)/mL보다 크다. 예를 들어, 특정 실시형태에서, 획득된 재조합 광견병 바이러스 입자 역가는 약 8E7 TU/mL, 약 9E7 TU/mL, 약 1E8 TU/mL, 약 1.1E8 TU/mL, 약 1.2E8 TU/mL, 약 1.3E8 TU/mL, 약 1.4E8 TU/mL, 약 1.5E8 TU/mL, 약 1.6E8 TU/mL, 약 1.7E8 TU/mL, 약 1.8E8 TU/mL, 약 1.9E8 TU/mL, 약 2E8 TU/mL, 약 2.5E8 TU/mL, 약 3E8 TU/mL, 약 3.5E8 TU/mL, 약 4E8 TU/mL, 약 4.5E8 TU/mL, 약 5E8 TU/mL, 약 5.5E8 TU/mL, 약 6E8 TU/mL, 약 6.5E8 TU/mL, 약 7E8 TU/mL, 약 7.5E8 TU/mL, 약 8E8 TU/mL, 약 8.5E8 TU/mL, 약 9E8 TU/mL, 약 9.1E8 TU/mL, 약 9.2E8 TU/mL, 약 9.3E8 TU/mL, 약 9.4E8 TU/mL, 약 9.5E8 TU/mL, 약 9.6E8 TU/mL, 약 9.7E8 TU/mL, 약 9.8E8 TU/mL, 약 9.9E8 TU/mL, 약 1E9 TU/mL, 약 1.1E9 TU/mL, 약 1.2E9 TU/mL,

또는 앞서 언급된 역가 사이의 임의의 값이다. 특정 실시형태에서, 획득된 재조합 광견병 바이러스 입자 역가는 약 1E8 TU/mL 내지 약 1E9 TU/mL, 예를 들어 8E7 TU/mL 내지 1.2E9 TU/mL 및 그 사이의 임의 범위이다.

[0983] **J. 유전자 요법의 방법**

[0984] 본 명세서에 기재된 재조합 광견병 바이러스 입자를 사용한 유전자 요법의 방법이 본 명세서에 제공된다. 특정 실시형태에서, 표적 세포에서 치료적 이식유전자를 발현시키기 위한 방법이 제공된다. 특정 실시형태에서, 표적 세포에서 염기 편집기를 발현시키기 위한 방법이 제공된다.

[0985] 특정 실시형태에서, 표적 세포에서 치료적 이식유전자를 발현시키기 위한 방법은 본 명세서에 기재된 바와 같은 재조합 광견병 바이러스 입자로 표적 세포를 형질도입하는 것을 포함한다. 예를 들어, 표적 세포에서 치료적 이식유전자를 발현시키기 위한 방법은 표적 세포를 광견병 바이러스 당단백질을 포함하는 재조합 광견병 바이러스 입자; 및 치료적 이식유전자를 인코딩하는 핵산을 포함하는 재조합 광견병 바이러스 게놈으로 형질도입하는 것을 포함하며, 여기서: 게놈은 광견병 바이러스 당단백질 또는 이의 기능적 변이체를 인코딩하는 G 유전자를 결하고/결하거나; 게놈은 광견병 바이러스 폴리머라제 또는 이의 기능적 변이체를 인코딩하는 L 유전자를 결한다. 특정 실시형태에서, 방법은 표적 세포를 광견병 바이러스 당단백질을 포함하는 재조합 광견병 바이러스 입자; 및 치료적 이식유전자를 인코딩하는 핵산을 포함하는 재조합 광견병 바이러스 게놈으로 형질도입하는 것을 포함하며, 여기서: 게놈은 광견병 바이러스 당단백질 또는 이의 기능적 변이체를 인코딩하는 G 유전자를 결한다. 특정 실시형태에서, 방법은 표적 세포를 광견병 바이러스 당단백질을 포함하는 재조합 광견병 바이러스 입자; 및 치료적 이식유전자를 인코딩하는 핵산을 포함하는 재조합 광견병 바이러스 게놈으로 형질도입하는 것을 포함하며, 여기서: 게놈은 광견병 바이러스 당단백질 또는 이의 기능적 변이체를 인코딩하는 G 유전자를 결하고; 게놈은 광견병 바이러스 폴리머라제 또는 이의 기능적 변이체를 인코딩하는 L 유전자를 결한다.

[0986] 표적 세포를 재조합 바이러스 입자로 형질도입하는 다양한 방법은 당업계의 통상인에게 공지되어 있다. 예를 들어, 표적 세포는 재조합 바이러스 입자와 접촉하여 바이러스 입자의 수용체-매개된 부착을 초래하고, 이어서 바이러스 입자의 세포 안으로의 클라트린-의존적 세포내이입을 초래할 수 있다.

[0987] 특정 실시형태에서, 표적 세포에서 핵염기 편집기를 발현시키기 위한 방법이 제공된다. 예를 들어, 이러한 방법은 표적 세포를 재조합 광견병 바이러스 입자로 형질도입하는 것을 포함하며, 여기서 재조합 바이러스 입자는 광견병 바이러스 당단백질; 및 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인 및 핵염기 편집 도메인을 포함하는 핵염기 편집기를 인코딩하는 핵산을 포함하는 재조합 광견병 바이러스 게놈을 포함하며, 여기서: 게놈은 광견병 바이러스 당단백질 또는 이의 기능적 변이체를 인코딩하는 G 유전자를 결하고/결하거나; 게놈은 광견병 바이러스 폴리머라제 또는 이의 기능적 변이체를 인코딩하는 L 유전자를 결한다. 특정 실시형태에서, 방법은 표적 세포를 재조합 광견병 바이러스 입자로 형질도입하는 것을 포함하며, 여기서 재조합 바이러스 입자는 광견병 바이러스 당단백질; 및 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인 및 핵염기 편집 도메인을 포함하는 핵염기 편집기를 인코딩하는 핵산을 포함하는 재조합 광견병 바이러스 게놈을 포함하며, 여기서 게놈은 광견병 바이러스 당단백질 또는 이의 기능적 변이체를 인코딩하는 G 유전자를 결한다. 특정 실시형태에서, 방법은 표적 세포를 재조합 광견병 바이러스 입자로 형질도입하는 것을 포함하며, 여기서 재조합 바이러스 입자는 광견병 바이러스 당단백질; 및 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인 및 핵염기 편집 도메인을 포함하는 핵염기 편집기를 인코딩하는 핵산을 포함하는 재조합 광견병 바이러스 게놈을 포함하며, 여기서: 게놈은 광견병 바이러스 당단백질 또는 이의 기능적 변이체를 인코딩하는 G 유전자를 결하고; 게놈은 광견병 바이러스 폴리머라제 또는 이의 기능적 변이체를 인코딩하는 L 유전자를 결한다.

[0988] 방법이 표적 세포에서 핵염기 편집기를 발현하기 위한 것인 경우, 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인은 결합된 가이드 폴리뉴클레오티드(예를 들어, gRNA)와 조합될 때 표적 폴리뉴클레오티드 서열에 특이적으로 결합할 수 있고 이에 의해 염기 편집기를 편집하고자 하는 표적 핵산 서열에 국소화할 수 있다.

[0989] 특정 실시형태에서, gRNA는 *시스로* 표적 세포에 제공된다. 예를 들어, gRNA는 재조합 광견병 바이러스 게놈 내에 포함될 수 있다. gRNA는 재조합 광견병 바이러스 게놈 내의 임의의 위치, 예를 들어, 하나 이상의 광견병 바이러스 유전자(예를 들어, N 유전자 또는 P 유전자)와 핵염기 편집기를 인코딩하는 핵산 사이, 또는 두 광견병 바이러스 유전자 사이에, 또는 재조합 광견병 바이러스 게놈의 말단 종단(예를 들어, 5' 말단 또는 3' 말단)에 포함될 수 있다.

[0990] 특정 실시형태에서, gRNA는 표적 세포에 *트랜스로* 제공된다(예를 들어, 외인적으로 제공됨). 예를 들어, gRNA는 재조합 광견병 바이러스 입자 외부의 별도 벡터 내에 포함될 수 있다. 적합한 벡터는 바이러스 벡터, 플라스미

드 및 기타 당업계의 통상인에게 공지된 벡터를 포함하지만 이에 제한되지 않는다. gRNA가 표적 세포에 *트랜스로* 제공되는 실시형태에서, gRNA 벡터는 당업계의 통상인에게 공지된 다양한 방법, 예를 들어, 제한 없이 전기천공을 통해 표적 세포 안으로 도입된다.

[0991] 대상체에게 치료적 이식유전자(예를 들어, 핵염기 편집기)를 전달하는 방법이 또한 제공된다. 특정 실시형태에서, 방법은 재조합 광견병 바이러스 입자를 대상체에게 투여하는 것을 포함하며, 여기서 재조합 바이러스 입자는 광견병 바이러스 당단백질; 및 치료적 이식유전자(예를 들어, 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인 및 핵염기 편집 도메인을 포함하는 핵염기 편집기)를 인코딩하는 핵산을 포함하는 재조합 광견병 바이러스 게놈을 포함하며, 여기서: 게놈은 광견병 바이러스 당단백질 또는 그것의 기능적 변이체를 인코딩하는 G 유전자를 결하고/결하거나; 게놈은 광견병 바이러스 폴리머라제 또는 이의 기능적 변이체를 인코딩하는 L 유전자를 결한다. 특정 실시형태에서, 방법은 재조합 광견병 바이러스 입자를 대상체에게 투여하는 것을 포함하며, 여기서 재조합 바이러스 입자는 광견병 바이러스 당단백질; 및 치료적 이식유전자(예를 들어, 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인 및 핵염기 편집 도메인을 포함하는 핵염기 편집기)를 인코딩하는 핵산을 포함하는 재조합 광견병 바이러스 게놈을 포함하며, 여기서: 게놈은 광견병 바이러스 당단백질 또는 이의 기능적 변이체를 인코딩하는 G 유전자를 결한다. 특정 실시형태에서, 방법은 재조합 광견병 바이러스 입자를 대상체에게 투여하는 것을 포함하며, 여기서 재조합 바이러스 입자는 광견병 바이러스 당단백질; 및 치료적 이식유전자(예를 들어, 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인 및 핵염기 편집 도메인을 포함하는 핵염기 편집기)를 인코딩하는 핵산을 포함하는 재조합 광견병 바이러스 게놈을 포함하며, 여기서: 게놈은 광견병 바이러스 당단백질 또는 이의 기능적 변이체를 인코딩하는 G 유전자를 결하고; 게놈은 광견병 바이러스 폴리머라제 또는 이의 기능적 변이체를 인코딩하는 L 유전자를 결한다.

[0992] 치료적 이식유전자(예를 들어, 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인 및 핵염기 편집 도메인을 포함하는 핵염기 편집기)를 전달 및/또는 발현하는 방법은 질환 또는 장애의 치료에서 용도를 발견한다. 특정 실시형태에서, 대상체에서 질환 또는 장애를 치료하는 방법은 본 명세서에 기재된 재조합 광견병 바이러스 입자 또는 본 명세서에 기재된 약학적 조성물을 투여하는 것을 포함한다. 특정 실시형태에서, 질환 또는 장애는 신경학적 질환 또는 장애이다. 특정 실시형태에서, 질환 또는 장애는 안과 질환 또는 장애이다.

[0993] 본 명세서에서 고려되는 약학적 조성물의 투여는 주입, 수혈 또는 비경구로를 포함하나 이에 제한되지 않는 통상적인 기술을 사용하여 수행될 수 있다. 일부 실시형태에서, 비경구 투여는 혈관내, 정맥내, 근육내, 동맥내, 척수강내, 종양내, 피내, 복강내, 경기관, 피하, 표피하, 관절내, 피막하, 거미막하 및 흉골내로 주입 또는 주사하는 것을 포함한다.

[0994] 본 발명의 실시는 달리 나타내지 않는 한, 당업자의 범위 내에 있는 분자 생물학(재조합 기술 포함), 미생물학, 세포 생물학, 생화학 및 면역학의 통상적인 기술을 이용한다. 이러한 기술은 문헌, 예컨대 "Molecular Cloning: A Laboratory Manual", second edition (Sambrook, 1989); "Oligonucleotide Synthesis" (Gait, 1984); "Animal Cell Culture" (Freshney, 1987); "Methods in Enzymology" "Handbook of Experimental Immunology" (Weir, 1996); "Gene Transfer Vectors for Mammalian Cells" (Miller 및 Calos, 1987); "Current Protocols in Molecular Biology" (Ausubel, 1987); "PCR: The Polymerase Chain Reaction", (Mullis, 1994); "Current Protocols in Immunology" (Coligan, 1991)에 완전하게 설명되어 있다. 이들 기술은 발명의 폴리뉴클레오티드 및 폴리펩티드의 생산에 적용가능하고, 따라서 발명의 제조 및 실시에 고려될 수 있다. 특정 실시형태에 특히 유용한 기술은 다음 섹션에서 논의될 것이다.

[0995] 다음 실시예는 당업계의 통상인에게 발명의 검정, 스크리닝 및 치료적 방법을 만들고 사용하는 방법에 대한 완전한 개시 및 설명을 제공하기 위해 제시되고, 발명자가 그의 발명품으로 간주하는 것의 범주를 제한하는 것으로 의도되지는 않는다.

[0996] **K. 실험예**

[0997] 실시예 1: 안정한 세포주의 생성

[0998] 아래 표 16에 기술된 안정한 세포주가 생성되었다:

[0999] 표 16: 안정한 세포주

세포주 명칭	설명	합체하는 벡터	선택 마커
CA1.9 (RABV-G)	광견병 바이러스 G 유전자를 안정적으로 발현하는 HEK293T 세포주	VIR120	블라스티시딘
CE1.13 (RABV-G)	코돈-최적화된 광견병 바이러스 G 유전자를 안정적으로 발현하는 HEK293T 세포주	VIR035	블라스티시딘
CA3.11	광견병 바이러스 N, P 및 L 유전자를 안정적으로 발현하는 HEK293T 세포주	VIR069	블라스티시딘
CA4.27	광견병 바이러스 N, P 및 L 유전자를 안정적으로 발현하는 HEK293T 세포주	VIR071	제오신
CA3.10 and CE1.10	광견병 바이러스 M 유전자를 안정적으로 발현하는 HEK293T 세포주	VIR121	블라스티시딘

[1000]

CE1.30	광견병 바이러스 N, P, M 및 L 유전자를 안정적으로 발현하는 HEK293T 세포주	VIR112	제오신

[1001]

[1002] VIR120, VIR035, VIR069, VIR071, VIR121 및 VIR112로부터 선택된 슬리핑 뷰티 트랜스포사제 시스템-호환 합체 벡터 중 하나를 슬리핑 뷰티 트랜스포사제 SB100X를 갖는 HEK293T 세포 안으로 공동-형질감염시켰다. VIR120은 EF1-알파 프로모터의 제어 하에 광견병 바이러스 G 유전자를 포함하는 발현 카세트를 함유하고; VIR035는 EF1-알파 프로모터의 제어 하에 코돈-최적화된 광견병 바이러스 G 유전자를 포함하는 발현 카세트를 함유하고; VIR069는 5'에서 3'으로 EF1-알파 프로모터, 광견병 바이러스 N 유전자, T2A 펩티드, 광견병 바이러스 P 유전자, P2A 펩티드 및 광견병 바이러스 L 유전자를 포함하는 발현 카세트를 함유하고; VIR071은 5'에서 3'으로 EF1-알파 프로모터, 광견병 바이러스 M 유전자, P2A 펩티드, 광견병 바이러스 P 유전자, IRES 및 광견병 바이러스 N 유전자를 포함하는 제1 발현 카세트와, 5'에서 3'으로 RPBSA 프로모터, 및 광견병 바이러스 L 유전자를 포함하는 제2 발현 카세트를 함유하며, 여기서 제1 발현 카세트와 제2 발현 카세트는 반대 방향으로 되고; VIR121은 EF1-알파 프로모터의 제어 하에 광견병 바이러스 M 유전자를 포함하는 발현 카세트를 함유하고; VIR112는

5'에서 3'으로 광견병 N 유전자, IRES, 광견병 P 유전자, IRES 및 광견병 M 유전자를 EF1-알파의 제어 하에 포함하는 발현 카세트와, 5'에서 3'으로 제오신, T2A 펩티드 및 광견병 L 유전자를 RPBSA 프로모터의 제어 하에 포함하는 제2 발현 카세트를 함유하며, 여기서 제1 발현 카세트와 제2 발현 카세트는 반대 방향으로 된다. VIR069, VIR071 및 VIR112 벡터의 도식은 **도 12**에 도시되어 있다.

[1003] 동시-형질감염 1일 후, 사용된 합체하는 벡터에 의존하여 블라스티시딘 또는 제오신을 사용하여 선택이 시작되었다. 필요에 따라 동시-형질감염 후 2일차부터 21일차까지 선택은 계속되었다. 28일차까지, 모든 생존 세포는 안정적으로 합체된 이식유전자를 가졌다.

[1004] 실시예 2: 재조합 광견병 바이러스 입자의 생산

[1005] 일차 생산을 위해, 0일차에 리포펙타민 3000을 사용하여 (i) 발현 벡터의 보체 플라스미드 혼합물 1.5μg 및 (ii) 광견병 레플리콘을 인코딩하는 플라스미드 0.5μg을 안정한 세포주 안으로 형질감염시켰다. 형질감염은 표 17에 따라 수행하였다:

[1006] 표 17: 형질감염 혼합물

안정한 세포주	보체 플라스미드 혼합물	레플리콘
RABV-G	DNA52 + VIR8, VIR9, VIR10, VIR11, VIR12	VIR045 "G-결실됨"
RABV-G	DNA52 + VIR8, VIR9, VIR10, VIR11, VIR12	VIR092 "G/L-결실됨"
CA3.11	DNA52 + VIR8, VIR9, VIR10, VIR11, VIR12	VIR045 "G-결실됨"
CA3.11	DNA52 + VIR8, VIR9, VIR10, VIR11, VIR12	VIR092 "G/L-결실됨"
CA4.27	DNA52 + VIR8, VIR9, VIR10, VIR11, VIR12	VIR045 "G-결실됨"
CA4.27	DNA52 + VIR8, VIR9, VIR10, VIR11, VIR12	VIR092 "G/L-결실됨"

[1007]

[1008] VIR045 레플리콘은 G 유전자가 결실된 광견병 SAD L16 전체 레플리콘을 함유한다. VIR092 레플리콘은 L 유전자가 추가로 결실된 VIR045로부터 유래되었다. VIR045와 VIR092 둘 모두는 GFP를 인코딩하는 서열을 함유한다. DNA52는 T7 RNA 폴리머라제를 인코딩하는 서열을 포함하는 발현 벡터이다. VIR8-12는 광견병 바이러스 N, P, M, G 및 L 유전자를 포함하는 발현 벡터이다.

[1009] 1일차에 배지를 OptiMem + 5% FBS("O5")로 변경했다. 1일차 배지는 폐기되었다. 3일차에 시작하여 바이러스 상등액을 수집하고 배지를 매일 신선한 O5 배지로 교체했다. 3-7일차로부터 바이러스 상등액을 풀링하여 4℃에 보관했다.

[1010] 풀링된 바이러스 상등액을 정화하여 4000rpm에서 15분 동안 원심분리에 의해 세포 잔해를 제거했다. 바이러스 입자는 Lenti-X 농축기(Takara Bio)에 대한 프로토콜에 따라 침전시키고 농축하였다. 상등액을 제거하고 펠렛을 O5 배지에 재현탁하여 농축된 바이러스 스톡을 생성했다. 농축된 바이러스 스톡을 사용하여 후속 증폭 계대를 접종했다.

[1011] 2차 바이러스 증폭은 다음과 같이 수행하였다. 0일차에 바이러스 스톡을 안정한 세포주에 첨가했다. 필요한 경우, 형질도입 시 추가 플라스미드를 안정한 세포주 안으로 공동-형질감염시켰다. VIR045 레플리콘을 사용하여 생산된 바이러스 스톡의 경우, RABV-G 안정 세포주에서 증폭될 때 추가로 필요한 것은 없었다. VIR092 레플리콘을 사용하여 생산된 바이러스 스톡의 경우, 증폭은 다음의 방식으로 수행되었으며 효율성은 괄호 안에 표시되어 있으며 – "+"가 많을수록 효율성이 더 높음을 나타낸다: (1) N, P 및 L 유전자를 함유하는 플라스미드로 공동-형질감염된 RABV-G 안정 세포주(+); (2) G 유전자를 함유하는 플라스미드로 공동-형질감염된 CA4.27 안정 세포주(++); 및 (3) G 유전자가 더욱 안정적으로 합체된 CA4.27 안정 세포주(+++).

[1012] 1일차에 배지는 O5 배지로 변경되었다. 1일차 배지는 폐기되었다. 2 내지 7일차에 바이러스 상등액을 수확하고 풀링하였다.

[1013] 또 다른 실험에서, VIR045 또는 VIR092 레플리콘으로 형질감염된 일차 형질감염 세포주 HEK293T 대조군 세포, RABV-G, CA3.11 및 CA4.27 사이에서 GFP 발현을 비교했다. 완전 보체 플라스미드 혼합물을 각 세포주 안으로 공동-형질감염시켰다. 표 18은 일차 형질감염 후 8일에 촬영한 이미지를 기반으로 한 GFP 발현의 정성적 수준을 보여주며, 여기서 "+"가 많을수록 GFP 발현이 더 높다는 것을 나타낸다.

[1014] 표 18: 일차 형질감염 세포주에서의 GFP 발현

[1015]

레플리콘	HEK293T	RABV-G	CA3.11	CA4.27
VIR045	+	+++++	+	+++
VIR092	-	-	-	++

[1016] 바이러스 상등액을 2 내지 4일차에 매일 수집하고 풀링하고 Lenti-X 농축기로 농축했다. 농축된 VIR045 바이러스 상등액을 RABV-G 세포에 첨가하고, 농축 VIR092 바이러스 상등액을 N, P 및 L 유전자를 함유하는 플라스미드로 형질감염된 RABV-G 세포에 첨가했다. 형질감염 2일 후 촬영한 이미지에 기반하여 재조합 광견병 바이러스 입자의 생성을 나타내는, GFP 발현의 정성적 수준을 표 19에 나타내었으며, 여기서 "+"가 많을수록 GFP 발현이 더 높다는 것을 나타낸다.

[1017] 표 19: 일차 증폭에서의 GFP 발현

[1018]

레플리콘	HEK293T	RABV-G	CA3.11	CA4.27
VIR045	+	+++++	+	+
VIR092	+++	+++	+++	+++

[1019] 또 다른 실험에서, 재조합 광견병 바이러스 상대적 감염성이 다양한 안정한 세포주를 사용하여 결정된 것으로부터 수득된 바이러스 상등액에 대해 결정되었다(**도 1**).

[1020] 안정한 세포주 c1, c8, c39, c40, c53 및 c54는 CA3.11 안정한 세포주("벌크")로부터 유래된 클론 세포주였다. 합체하는 벡터 VIR069("BHK") 및 합체하는 벡터 VIR120("BHK-G")을 사용하는 BHK 세포주도 생성되었다. CA4.27 세포를 웰당 0.4, 0.6, 0.8 또는 1백만 개의 세포로 도말했다.

[1021] 바이러스 상등액을 서로 다른 날(D2 또는 D3)에 수확하고 이어서 **도 1**에 표시된 부피(5uL 또는 30uL)에서 순수 HEK293T 세포를 감염시키는데 사용했다. GFP의 발현에 의해 결정된 바와 같이 감염된 세포의 백분율을 나타내는 유세포측정법에 의해 역가측정을 수행하였다.

[1022] 실시예 3: 재조합 광견병 바이러스 입자 유전자 전달

[1023] 재조합 광견병 바이러스 입자가 유전자 전달에 사용될 수 있는지 여부를 조사하기 위해, 레플리콘 VIR218이 생성되었다. VIR218은 아데노신 데아미나제 ABE7.10을 인코딩하는 서열의 추가로 VIR092로부터 유래되었다; **도 2a**는 VIR218의 도식이다. **도 2b**는 이어지는 생산 및 증폭 계획을 보여주는 도식이다. 완전 보체 플라스미드 혼합물로 VIR218을 순수 HEK293T 세포 안으로 공동-형질감염시켜 일차 생산을 수행했다. RABV-G 세포주 상에 N, P 및 L 유전자를 함유하는 플라스미드의 추가 형질감염으로 2차 및 3차 증폭을 수행했다. 바이러스 상등액을 수집하고 상기에서 기술된 바와 같이 농축하여 바이러스 스톡을 생성했다. 그런 다음 바이러스 스톡을 HEK2-2를 표적화하는 gRNA(gaacacaaagcatagactgc; 서열번호: 4011)를 포함하는 플라스미드를 리포펙션을 통한 형질감염과 함께, 그리고 선택적으로 L 유전자("보충 L")를 포함하는 플라스미드로 공동-형질감염으로 순수 293T 세포에 첨가하였다. 게놈 DNA를 추출하고 표준 PCR/라이브러리 준비를 수행하여 게놈 표적을 증폭시키고 편집을 평가했다(**도 2c**). **도 2c**에 도시된 바와 같이, A>G 편집은 감염된 HEK293T 세포에서 검출되었다.

[1024] 실시예 4: 재조합 광견병 바이러스 게놈

[1025] 실시예 1-3에서는 ΔGL 배경에서 재조합 광견병 바이러스 게놈의 사용을 기술했다. 바이러스의 생산은 다음에 ΔALL(즉, ΔGLNPM) 배경에서의 재조합 광견병 바이러스 게놈에서 테스트되었다(**도 4a**에서의 레플리콘 묘사 참조). 테스트된 광견병 바이러스 레플리콘은 Cre 재조합효소-2A-GFP 리포터 유전자를 인코딩했다.

[1026] **도 4b**에 도시된 바와 같이, ΔG, ΔGL 또는 ΔALL(즉, ΔGLNPM) 배경에서 광견병 바이러스 벡터로 형질감염된 세포는 Cre 재조합효소-2A-GFP 리포터 유전자로부터 형광을 방출했다. 이미지는 **도 2b**에 묘사된 바와 같이 표준

형질감염 및 최적화된 형질감염 하에서 형질감염-후 6일에 촬영되었다.

[1027] **도 4c**에 도시된 바와 같이, **도 4b**에서 생성된 광견병 바이러스의 % 바이러스 도입은 GFP 형광에 의해 측정된 바와 같이, 리포터 세포 안으로 형질도입에 이어서 결정되었다.

[1028] 이 연구는 모든 바이러스 유전자(즉, ΔALL 또는 ΔGLNPM)를 결하는 조작된 재조합 광견병 바이러스 게놈이 이식유전자를 발현하는데 사용될 수 있음을 입증한다.

[1029] 실시예 5: tRNA를 절단함으로 gRNA를 광견병 게놈 안으로 인코딩하는 것

[1030] 광견병 바이러스 게놈에서 gRNA가 인코딩될 수 있는지 여부를 조사하기 위해, 레플리콘 VIR621이 **도 3a**에 도시된 조직화에서 생성되었다. VIR621은 ABE8에 함유된 폴리뉴클레오티드 프로그램가능 뉴클레오티드 결합 도메인 및 아데노신 데아미나제와 G 유전자를 결하는 바이러스 게놈을 인코딩하는 서열(**도 3a**)이 추가된 2개의 측접하는 절단 tRNA 및 개재 gRNA(**도 3b**)를 인코딩하는 DNA538로부터 유래되었다. 다중 표적 tRNA는 또한 다중화를 허용하는 상이한 tRNA 조합 사이 또는 그 후에 인코딩되었다(**도 3c, 도 3d**). 표 20에 나열된 바와 같은 tRNA와 gRNA의 여러 조합을 **도 3e**에서 편집 효율성에 대해 테스트했다. **도 3e**에 도시된 바와 같이, HEK2 및 IEDG 유전자의 A>G 편집은 무 gRNA(VIR596), HEK2를 표적화하는 단일 gRNA(VIR621, VIR622), IEDG를 표적화하는 단일 gRNA(VIR712, VIR713) 또는 동일한 바이러스 레플리콘에서 HEK2 및 IEDG를 표적화하는 다중화된 다중 gRNA(VIR714, VIR715, VIR717, VIR718, VIR719, VIR720, VIR627, VIR628, VIR629)를 함유하는 바이러스 레플리콘을 갖는 감염된 HEK293T 세포에서 검출되었다.

[1031] 표 20: tRNA 및 gRNA 레플리콘

벡터 명칭	벡터 설명	삽입 명칭	삽입 서열
VIR621	SynV ΔG ABE8-20-2a-GFP tRNA Pro-Hek2 gRNA	tRNA Pro 3' 방출	gtacaagTAAGAAGTTGAATAACAAAATGC CGGAAATCTACGGATTGTGTATATCCAT CATGAAAAAAACTAACACCCCTCCTTTC GAACCATCCCAAACggctcgttggtctaggggtat gattctcgcttagggtgcgagaggtcccgggttcaaatccc ggacgagcccGGAACACAAAGCATAGACTG Cgttttagagcta GAAAtagcaagttaaaataaggcta gtccgttatcaacttgaaaaagtggcaccgagtcggtgctt ttCGAGGAAGGAGGTCTGAGGAGGTCAC TGcgaaccagtttgtgtcggctcgttggtctaggggtatga ttctcgcttagggtgcgagaggtcccgggttcaaatcccgg acgagccctctagaagtgctgggtcatcta
VIR622	SynV ΔG ABE8-20-2a-GFP tRNA Ile-Hek2 gRNA	tRNA Ile 3' 방출	gtacaagTAAGAAGTTGAATAACAAAATGC CGGAAATCTACGGATTGTGTATATCCAT CATGAAAAAAACTAACACCCCTCCTTTC GAACCATCCCAAACgctccagtggcgcaatcggt tagcgcgcggtacttataagacagtgcacctgtgagcaat gccgaggttgtgagttcaagcctcacctggagcaGGAA CACAAAGCATAGACTGCgttttagagctaGAA Atagcaagttaaaataaggctagtccgttatcaacttgaa aaagtggcaccgagtcggtgcttCACACACACAAg ctccagtggcgcaatcggttagcgcgcggtacttataaga cagtgcaGCCgCGAGGAAGGAGGTCTGAG GAGGTCACTGcGGCcctgtgagcaatgccgagg ttgtgagttcaagcctcacctggagcata
VIR623	SynV ΔG ABE8-20-2a-GFP tRNA Ile-정점 Hek2 gRNA	tRNA 정점 방출	gtacaagTAAGAAGTTGAATAACAAAATGC CGGAAATCTACGGATTGTGTATATCCAT CATGAAAAAAACTAACACCCCTCCTTTC GAACCATCCCAAACgctccagtggcgcaatcggt tagcgcgcggtacttataagacagtgcaGAACACAA AGCATAGACTGCgttttagagctaCCGAAAGG tagcaagttaaaataaggctagtccgttatcaacttgaaaa agtggcaccgagtcggtgcttcacacacacacaCGAG

[1032]

[1033]

			GAAGGAGGTCTGAGGAGGTCACTGcgcctgtgagcaatgccgaggttgtgagttcaagcctcacctggagcata
VIR624	긴 링커를 갖는 SynV ΔG ABE8-20-2a-GFP tRNA Ile-정점 Hek2 gRNA	긴 링커를 갖는 tRNA 정점 방출	gtacaagTAAGAAGTTGAATAACAAAATGCCGGAAATCTACGGATTGTGTATATCCATCATGAAAAAAACTAACACCCCTCCTTTCGAACCATCCCAAACgctccagtggcgcaatcggttagcgcgcggtacttataagacagtgcagGAACACAAAGCATAGACTGCgttttagagctaCCGAAAGGtagcaagttaaaaCaaggctagtccgttatcaacttgaaaaagtggcaccgagtcggtgctttGGCCCGAGGAAGGAGGTCTGAGGAGGTCACTGGGCCAAAACAACAACCCAACCAACAAACCAACACCAAACAACAAACCAAACCCCAACAAACAACCACCAACCCAAACAAcctgtgagcaatgccgaggttgtgagttcaagcctcacctggagcata
VIR625	긴 링커를 갖는 SynV ΔG ABE8-20-2a-GFP tRNA Ile-정점 Hek2 gRNA	안정화된 tRNA 정점 방출	gtacaagTAAGAAGTTGAATAACAAAATGCCGGAAATCTACGGATTGTGTATATCCATCATGAAAAAAACTAACACCCCTCCTTTCGAACCATCCCAAACgctccagtggcgcaatcggtagcgcgcggtacttataagacagtgcaGGAGCCCGAACACAAAGCATAGACTGCgttttagagctaGGCCCGAGGAAGGAGGTCTGAGGAGGTCACTGGGCCtagcaagttaaaataaggctagtccgttatcaacttgaaaaagtggcaccgagtcggtgcttAAAACAACAACCCAACCAACAAACCAACACCAAACAACAAACCAAACCCCAACAAACAACCACCAACCCAAACAAGGGCTCCcctgtgagcaatgccgaggttgtgagttcaagcctcacctggagcata
VIR626	SynV ΔG ABE8-20-2a-GFP tRNA Ile	바꾸어진 tRNA Ile	gtacaagTAAGAAGTTGAATAACAAAATGCCGGAAATCTACGGATTGTGTATATCCATCATGAAAAAAACTAACACCCCTCCTTTCGAACCATCCCAAACGGGCTCCcctgtgagcaatgccgaggttgtgagttcaagcctcacctggagcaGAAAgctccagtggcgcaatcggttagcgcgcggtacttata

	바꾸어진 Hek2 gRNA		agacagtgcaGGAGCCCGAACACAAAGCATAGACTGCgttttagagctaGGCCCGAGGAAGGAGGTCTGAGGAGGTCACTGGGCCtagcaagttaaaataaggctagtccgttatcaacttgaaaaagtggcaccgagtcggtgcttAAAACAACAACCCAACCAACAAACCAACACCAAACAACAAACCAAAACCCCAACAAACAACCACCAACCCAAACAAta
VIR627	SynRV tRNA-Pro-Thr IEDG Hek2 ΔG Abe820m-T2a-mScarlet	P-IEDG-T-Hek2	AACATCCCTCAAAagactcaaggaaagggctcgttggtctaggggtatgattctcgcttagggtgcgagaggtccgggttcaaatcccggacgagcccGcgtGtAgggTaaccatgaacGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTCACACACACAAggctccatagctcaggggttagagcactggtcttgtaaaccaggggtcgcgagttcaattctcgctggggcttGGAACACAAAGCATAGACTGCgttttagagctaGCCgCGAGGAAGGAGGTCTGAGGAGGTCACTGcGGCtagcaagttaaaataaggctagtccgttatcaacttgaaaaagtggcaccgagtcggtgcttttttaaTTAAccgagaaaaaaa
VIR628	SynRV tRNA-Val-Lys IEDG Hek2 ΔG Abe820m-T2a-mScarlet	V-IEDG-K-Hek2	AACATCCCTCAAAagactcaaggaaaggtttccgtagtgtagtggttatcacgttcgcctcacacgcgaaaggtcccggttcgaaaccgggcggaaacaGcgtGtAgggTaaccatgaacGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTCACACACACAAgcccggctagctcagtcggtagagcatgagactcttaatctcagggtcgtgggttcgagccccacgttgggcgGGAACACAAAGCATAGACTGCgttttagagctaGCCgCGAGGAAGGAGGTCTGAGGAGGTCACTGcGGCtagcaagttaaaataaggctagtccgttatcaacttgaaaaagtggcaccgagtcggtgcttttttaaTTAAccgagaaaaaaaa

[1034]

[1035]

VIR629	SynRV tRNA-Asp-Gly-Glu IEDG Hek2 ΔG Abe820m-T2a-mScarlet	D-IEDG-G-Hek2-Q	AACATCCCTCAAAagactcaaggaaagtcctcgtt agtatagtggtgagtatccccgcctgtcacgcgggagacc ggggttcgattccccgacggggagGcgtGtAgggTaa ccatgaacGTTTTAGAGCTAGAAATAGCAA GTTAAAATAAGGCTAGTCCGTTATCAAC TTGAAAAAGTGGCACCGAGTCGGTGCT TTTTTCACACACACAAgcgttggtggtatagtggt gagcatagctgccttccaagcagttgacccgggttcgattc ccggccaacgcaGGAACACAAAGCATAGAC TGCgttttagagctaGCCgCGAGGAAGGAGG TCTGAGGAGGTCACTGcGGCtagcaagttaa aataaggctagtccgttatcaacttgaaaaagtggcaccg agtcggtgctttCACACACACAAtccttggtggtctag tggttaggattcggcgctctcaccgccgcggcccgggttc gattcccggtcagggaattaaTTAAccgagaaaaaaa
VIR712	SynRV tRNA-Ile-Ile(com) IEDG PacI ΔG Abe820m-T2a-mScarlet	VIR622 삽입	AACATCCCTCAAAagactcaaggaaaggctcca gtggcgcaatcggttagcgcgcggtacttataagacagtg cacctgtgagcaatgccgaggttgtgagttcaagcctcac ctggagcaGcgtGtAgggTaaccatgaacGTTTTA GAGCTAGAAATAGCAAGTTAAAATAAGG CTAGTCCGTTATCAACTTGAAAAAGTGG CACCGAGTCGGTGCTTTTTTCACACACA CAAgctccagtggcgcaatcggttagcgcgcggtactta taagacagtgcaGCCgCGAGGAAGGAGGTC TGAGGAGGTCACTGcGGCcctgtgagcaatg ccgaggttgtgagttcaagcctcacctggagcaTTAAT TAAtccgagaaaaaaa
VIR713	SynRV tRNA-Ile IEDG PacI ΔG Abe820m-T2a-mScarlet	5'Ile to PacI	AACATCCCTCAAAagactcaaggaaaggctcca gtggcgcaatcggttagcgcgcggtacttataagacagtg cacctgtgagcaatgccgaggttgtgagttcaagcctcac ctggagcaGcgtGtAgggTaaccatgaacGTTTTA GAGCTAGAAATAGCAAGTTAAAATAAGG CTAGTCCGTTATCAACTTGAAAAAGTGG CACCGAGTCGGTGCTTTTTTaattaacgaga aaaaaa

[1036]

VIR714	SynRV tRNA-Ile-Ile(corn) IEDG Hek2 ΔG Abe820m-T2a-mScarlet	I-IEDG-I-Hek2	AACATCCCTCAAAagactcaaggaaaggctcca gtggcgcaatcggttagcgcgcggtacttataagacagtg cacctgtgagcaatgccgaggttgtgagttcaagcctcac ctggagcaGcgtGtAgggTaaccatgaacGTTTTA GAGCTAGAAATAGCAAGTTAAAATAAGG CTAGTCCGTTATCAACTTGAAAAAGTGG CACCGAGTCGGTGCTTTTTTCACACACA CAAgctccagtggcgcaatcggttagcgcgcggtactta taagacagtgcaGCCgCGAGGAAGGAGGTC TGAGGAGGTCACTGcGGCcctgtgagcaatg ccgaggttgtgagttcaagcctcacctggagcaGGAA CACAAAGCATAGACTGCgttttagagctaGCC gCGAGGAAGGAGGTCTGAGGAGGTCAC TGcGGCtagcaagttaaaataaggctagtccgttatca acttgaaaaagtggcaccgagtcggtgctttttccgagaa aaaaa
VIR715	SynRV tRNA-Ile-Gly IEDG Hek2 ΔG Abe820m-T2a-mScarlet	I-IEDG-G-Hek2	AACATCCCTCAAAagactcaaggaaaggctcca gtggcgcaatcggttagcgcgcggtacttataagacagtg cacctgtgagcaatgccgaggttgtgagttcaagcctcac ctggagcaGcgtGtAgggTaaccatgaacGTTTTA GAGCTAGAAATAGCAAGTTAAAATAAGG CTAGTCCGTTATCAACTTGAAAAAGTGG CACCGAGTCGGTGCTTTTTTCACACACA CAAgcgttggtggtatagtggtgagcatagctgccttcca agcagttgacccgggttcgattcccggccaacgcaGGA ACACAAAGCATAGACTGCgttttagagctaGC CgCGAGGAAGGAGGTCTGAGGAGGTCA CTGcGGCtagcaagttaaaataaggctagtccgttatc aacttgaaaaagtggcaccgagtcggtgctttttaaTTA Accgagaaaaaaa
VIR716	SynRV tRNA-Ile-Lys IEDG Hek2 ΔG Abe820m-	I-IEDG-K-Hek2	AACATCCCTCAAAagactcaaggaaaggctcca gtggcgcaatcggttagcgcgcggtacttataagacagtg cacctgtgagcaatgccgaggttgtgagttcaagcctcac ctggagcaGcgtGtAgggTaaccatgaacGTTTTA GAGCTAGAAATAGCAAGTTAAAATAAGG CTAGTCCGTTATCAACTTGAAAAAGTGG

	T2a-mScarlet		CACCGAGTCGGTGCTTTTTTCACACACACAAgcccggctagctcagtcggtagagcatgagactcttaatctcagggtcgtgggttcgagccccacgttgggcgGGAACACAAAGCATAGACTGCgttttagagctaGCCgCGAGGAAGGAGGTCTGAGGAGGTCACTGcGGCtagcaagttaaaataaggctagtccgttatcaacttgaaaaagtggcaccgagtcggtgcttttttaaTTAAccgagaaaaaaa
VIR717	SynRV tRNA-Ile-Leu IEDG Hek2 ΔG Abe820m-T2a-mScarlet	I-IEDG-L-Hek2	AACATCCCTCAAAagactcaaggaaaggctccagtggcgcaatcggttagcgcgcggtacttataagacagtgcacctgtgagcaatgccgaggttgtgagttcaagcctcacctggagcaGcgtGtAgggTaaccatgaacGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTCACACACACAAggtagcgtggccgagcggtctaaggcgctggattaaggctccagtctcttcgggggcgtgggttcgaatcccaccgctgccaGGAACACAAAGCATAGACTGCgtttagagctaGCCgCGAGGAAGGAGGTCTGAGGAGGTCACTGcGGCtagcaagttaaaataaggctagtccgttatcaacttgaaaaagtggcaccgagtcggtgcttttttaaTTAAccgagaaaaaaa
VIR718	SynRV tRNA-Ile-Pro IEDG Hek2 ΔG Abe820m-T2a-mScarlet	I-IEDG-P-Hek2	AACATCCCTCAAAagactcaaggaaaggctccagtggcgcaatcggttagcgcgcggtacttataagacagtgcacctgtgagcaatgccgaggttgtgagttcaagcctcacctggagcaGcgtGtAgggTaaccatgaacGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTCACACACACAAggctcgttggtctaggggtatgattctcgcttagggtgcgagaggtcccgggttcaaatcccggacgagcccGGAACACAAAGCATAGACTGCgttttagagctaGCCgCGAGGAAGGAGGTCTGAGGAGGTCACTGcGGCtagcaagttaaaataaggctagtccgttatcaacttgaaaaagtggcaccgagtcggtgcttttttaaTTAA

[1037]

VIR719	SynRV tRNA-Ile-Thr IEDG Hek2 ΔG Abe820m-T2a-mScarlet	I-IEDG-T-Hek2	AACATCCCTCAAAagactcaaggaaaggctccagtggcgcaatcggttagcgcgcggtacttataagacagtgcacctgtgagcaatgccgaggttgtgagttcaagcctcacctggagcaGcgtGtAgggTaaccatgaacGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTCACACACACAAggctccatagctcaggggttagagcactggtcttgtaaaccaggggtcgcgagttcaattctcgctggggcttGGAACACAAAGCATAGACTGCgttttagagctaGCCgCGAGGAAGGAGGTCTGAGGAGGTCACTGcGGCtagcaagttaaaataaggctagtccgttatcaacttgaaaaagtggcaccgagtcggtgctttttaaTTAAccgagaaaaaa
VIR720	SynRV tRNA-Ile-Val IEDG Hek2 ΔG Abe820m-T2a-mScarlet	I-IEDG-V-Hek2	AACATCCCTCAAAagactcaaggaaaggctccagtggcgcaatcggttagcgcgcggtacttataagacagtgcacctgtgagcaatgccgaggttgtgagttcaagcctcacctggagcaGcgtGtAgggTaaccatgaacGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTCACACACACAAgtttccgtagtgtagtggttatcacgttcgcctcacacgcgaaaggtccccggttcgaaaccgggcggaaacaGGAACACAAAGCATAGACTGCgttttagagctaGCCgCGAGGAAGGAGGTCTGAGGAGGTCACTGcGGCtagcaagttaaaataaggctagtccgttatcaacttgaaaaagtggcaccgagtcggtgctttttaaTTAAccgagaaaaaaa

DNA538 서열:

DNA538	EFS-tRNA-Pro-HEK2 gRNA	tRNA-Pro-HEK2 gRNA	ggctcgttggtctaggggtatgattctcgcttagggtgcgaggaggtcccgggttcaaatcccggacgagcccGAACACAAAGCATAGACTGCgtCttagagctaGGCCCGAGGAAGGAGGTCTGAGGAGGTCACTGGGCCtagcaagttaaGataaggctagtccgttatcaacttgaaaaagtggcaccgagtcggtgcttaaccagtttgtgt

[1038]

			cggctcgttggtctaggggtatgattctcgcttagggtgcga gaggtcccgggttcaaatcccggacgagccc

VIR622 서열:

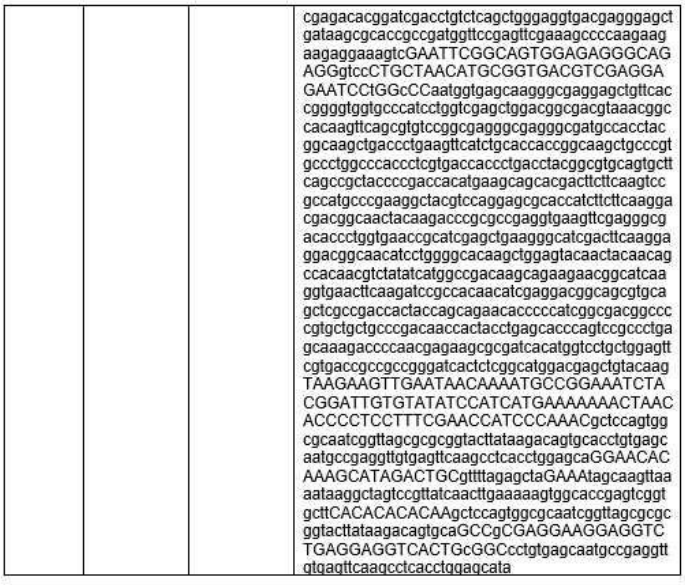
VIR622	VIR622 SynV delG ABE8-20-2a-GFP tRNA Ile-Hek2 gRNA	ABE8-20-2a-GFP tRNA Ile-Hek2 gRNA	ATGtccgaagtcgagttttcccatgagtactggatgagacacgcatt gactctcgcaaagagggctcgagatgaacgcgaggtgcccgtggg ggcagtactcgtgctcaacaatcgcgtaatcggcgaaggttggaata gggcaatcggactccacgaccccactgcacatgcggaaatcatggc ccttcgacagggagggcttgtgatgcagaattatcgacttatcgatgcg acgctgtacgtcacgtttgaaccttgcgtaatgtgcgcgggagctatga ttcactcccgcattggacgagttgtattcggtgttcgcaacgccaagac gggtgccgcaggttcactgatggacgtgctgcattacccaggcatga accaccgggtagaaatcacagaaggcatattggcggacgaatgtgc ggcgctgttgtgttacttttttcgcatgcccaggcgtgtctttaacgccca gaaaaaagcacaatcctctactgactctggtggttcttctggtggttcta gcggcagcgagactcccgggacctcagagtccgccacacccgaa agttctggtggttcttctggtggttctgacaagaagtacagcatcggcct ggccatcggcaccaactctgtgggctgggccgtgatcaccgacgagt acaaggtgcccagcaagaaattcaaggtgctgggcaacaccgacc ggcacagcatcaagaagaacctgatcggagccctgctgttcgacag cggcgaaacagccgaggccacccggctgaagagaaccgccaga agaagatacaccagacggaagaaccggatctgctatctgcaagag atcttcagcaacgagatggccaaggtggacgacagcttcttccacag actggaagagtccttcctggtggaagaggataagaagcacgagcg gcaccccatcttcggcaacatcgtggacgaggtggcctaccacgag aagtaccccaccatctaccacctgagaaagaaactggtggacagc accgacaaggccgacctgcggctgatctatctggccctggcccacat gatcaagttccggggccacttcctgatcgagggcgacctgaaccccg acaacagcgacgtggacaagctgttcatccagctggtgcagaccta caaccagctgttcgaggaaaaccccatcaacgccagcggcgtgga cgccaaggccatcctgtctgccagactgagcaagagcagacggctg gaaaatctgatcgcccagctgcccggcgagaagaagaatggcctgt tcggaaacctgattgccctgagcctgggcctgaccccccaacttcaag agcaacttcgacctggccgaggatgccaaactgcagctgagcaag gacacctacgacgacgacctggacaacctgctggcccagatcggc gaccagtacgccgacctgtttctggccgccaagaacctgtccgacgc catcctgctgagcgacatcctgagagtgaacaccgagatcaccaag gccccctgagcgcctctatgatcaagagatacgacgagcaccacc aggacctgaccctgctgaaagctctcgtgcggcagcagctgcctgag aagtacaaagagattttcttcgaccagagcaagaacggctacgccg gctacattgacggcggagccagccaggaagagttctacaagttcatc aagcccatcctggaaaagatggacggcaccgaggaactgctcgtg aagctgaacagagaggacctgctgcggaagcagcggaccttcgac aacggcagcatcccccaccagatccacctgggagagctgcacgcc attctgcggcggcaggaagatttttacccattcctgaaggacaaccgg gaaaagatcgagaagatcctgaccttccgcatcccctactacgtggg ccctctggccaggggaaacagcagattcgcctggatgaccagaaa gagcgaggaaaccatcacccctggaacttcgaggaagtggtgga caagggcgcttccgcccagagcttcatcgagcggatgaccaacttcg ataagaacctgcccaacgagaaggtgctgcccaagcacagcctgct gtacgagtacttcaccgtgtataacgagctgaccaaagtgaaatacgt

[1039]

[1040]



[1041]



[1042] tRNA-gRNA-tRNA 카셋트(VIR622 내):

[1043] gctccagtggcgcaatcggttagcgcgcggtacttataagacagtgcacctgtgagcaatgccgaggttgtgagttcaagcctcacctggagcaGGAACACA
AAGCATAGACTGCgttttagagctaGAAAtagcaagttaaaataaggctagtccgttatcaacttgaaaaagtggcaccgagtcggtgcttCACACACACAA
gctccagtggcgcaatcggttagcgcgcggtacttataagacagtgcaGCCgCGAGGAAGGAGGTCTGAGGAGGTCACTGcGGCcctgtgagcaatgccgag
gttgtgagttcaagcctcacctggagca

[1044] 실시예 6: 재조합 광견병 바이러스 입자 유전자 전달 - 추가 이식유전자

[1045] SynV로 지칭되는 SAD B19 ΔG RABV 게놈의 코돈-최적화된 버전은 염기 편집기 ABE8.20을 인코딩하는데 사용되었으며, 여기서 치료적 화물(예를 들어, 염기 편집기)은 바이러스 게놈 내의 G 단백질을 대체했다. HEK2 및 ABCA4 좌위를 표적화하는 U6-구동 gRNA는 감염 시에 표적 293T 세포 안으로 트랜스로 형질감염되었다. 기능적 바이러스 역가를 기준으로 32, 8, 1의 감염의 다중도(MOI)가 사용되었다. **도 5**에 도시된 바와 같이, ABE 8.20 염기 편집기를 인코딩하는 RABV 게놈을 갖는 RABV를 형질도입한 293T 세포에서 A에서 G 염기 편집이 관찰되었다.

[1046] SynV ΔG RABV 게놈은 여러 다른 염기 편집기를 293T 세포 안으로 전달하는데 사용되었다. **도 6**에 도시된 바와 같이, 여러 BE4 염기 편집기 중 하나를 인코딩하는 RABV 게놈을 갖는 RABV로 형질도입된 293T 세포에서 C에서 T 염기 편집이 관찰되었다.

[1047] 상기에 기술된 ΔG RABV 게놈에 부가하여, ΔGL RABV 게놈을 사용하여 염기 편집기 ABE7.10을 전달했다. **도 7**에 도시된 바와 같이, ABE7.10을 인코딩하는 RABV 게놈을 갖는 RABV로 형질도입된 293T 세포에서 A에서 G 염기 편집이 관찰되었다. 한 세트의 세포가 RABV L 단백질(바이러스 + L)을 인코딩하는 플라스미드에 의해 추가로 형질감염되었다.

[1048] **도 5** 내지 **도 7**에 의해 나타난 각각의 상기 실험에서, 강력한 염기 편집이 여러 가지 상이한 염기 편집기 중 하나를 발현하는 RABV로 형질도입된 세포에서 관찰되었다. 이 강력한 활성은 ΔG 및 ΔGL RABV 게놈 둘 모두에서, 수많은 MOI 및 세포 게놈에서의 2개의 별개의 부위(HEK2 및 ABCA4 좌위)에서 관찰되었다.

[1049] 실시예 7: 재조합 광견병 바이러스 게놈 — 추가 게놈 아키텍처

[1050] 실시예 4는 ΔALL(즉, ΔGLNPM) 배경에서 재조합 광견병 바이러스 게놈의 사용을 기술했다. 추가 RABV 게놈 아키텍처가 이식유전자 발현이 가능한지를 결정하기 위해 다음으로 테스트되었다. 상기에서 기술된 293T 세포와 세포주 CE1.30이 바이러스 생산에 사용되었다. 세포는 ΔN, ΔP, ΔM 및 ΔL 배경에서 광견병 바이러스 벡터로 형질감염되었으며, 각 벡터는 Cre 재조합효소-2A-GFP 리포터 유전자를 인코딩한다. 293T 세포 또는 CE1.30 세포는 누락된 유전자의 보완으로 또는 보완 없이 형질감염되었다(예를 들어, ΔN 벡터는 N 유전자 발현 벡터와 공동-형질감염되었고, ΔP 벡터는 P 유전자 발현 벡터와 공동-형질감염되었고, 등등임). 형질감염-후 6일에 형광 이미지를 촬영했다. **도 8**에 도시된 바와 같이, CE1.30 세포주는 누락된 유전자의 추가적인 보완으로 또는 보완 없이 ΔN, ΔP, ΔM 및 ΔL 레플리콘의 성장을 지원했다.

[1051] 상기 언급된 세포 형질감염의 세포 상등액을 다음으로 형질감염되지 않은 293T 세포에 적용하여 광견병 바이러스의 % 바이러스 도입을 결정했다. % 바이러스 도입은 GFP 형광에 의해 측정되었다. **도 9**에 도시된 바와 같이, 세포 안으로 % 바이러스 도입은 CE1.30 세포주에서 테스트된 각 광견병 바이러스 벡터에 대해 검출되었다.

[1052] ΔMGL 및 ΔPMGL의 추가 RABV 게놈 아키텍처는 CE1.30 세포주에서 수행되는 바이러스 생산과 유사한 방식으로 테스트되었다. Cre-2a-GFP를 이식유전자로 인코딩하는 ΔGL, ΔMGL, ΔPMGL 및 ΔALL 벡터는 CE1.30 세포에서 성장되었고 리포터 세포에서 역가측정되어 Cre 및 GFP mRNA 둘 모두의 기능적 전달을 보여주었다. **도 10**에 도시된 바와 같이, GFP 및 mScarlet 형광에 의해 측정된 바와 같이, 모든 RABV 게놈 배경에서 % 바이러스 도입이 검출되었다. GFP 이식유전자 발현은 또한 **도 11**에 도시된 바와 같이 형광에 의해 검출되었다.

[1053] 다른 실시형태

[1054] 전술한 설명으로부터, 본 명세서에 기술된 발명을 다양한 용도 및 조건에 적용하기 위해 수정 및 변형이 이루어질 수 있다는 것이 명백해질 것이다. 이러한 실시형태는 또한 다음 청구범위의 범주 내에 있다.

[1055] 본 명세서에서 변수의 임의의 정의에서의 요소 목록의 언급은 나열된 요소의 임의의 단일 요소 또는 조합(또는 하위조합)으로서의 그 변수의 정의를 포함한다. 본 명세서의 실시형태의 언급은 임의의 단일 실시형태로서 또는 임의의 다른 실시형태 또는 그 부분과 결합된 형태로서 그 실시형태를 포함한다.

[1056] 본 명세서에 언급된 모든 특허 및 간행물은 각각의 독립적인 특허 및 간행물이 참조로 포함되도록 구체적이고 개별적으로 표시된 것과 동일한 정도로 참조로 본 명세서에 포함된다.

[1057] L. 서열목록

설명	서열번호:	서열
아데노신 데아미나제 참조 서열	8	MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGW NRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCA GAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGIL ADECAALLCYFFRMPRQVFNAQKKAQSSTD
P153 폴리뉴클레오티드 에서 BhCas12b GGSGGS-ABE8-Xten20	274	GCCACCATGGCCCCAAAGAAGAAGCGGAAGGTCGGTATCCACG GAGTCCCAGCAGCCGCCACCAGATCCTTCATCCTGAAGATCGA GCCCAACGAGGAAGTGAAGAAAGGCCTCTGGAAAACCCACGAG GTGCTGAACCACGGAATCGCCTACTACATGAATATCCTGAAGCT GATCCGGCAAGAGGCCATCTACGAGCACCACGAGCAGGACCCC AAGAATCCCAAGAAGGTGTCCAAGGCCGAGATCCAGGCCGAGC TGTGGGATTTCGTGCTGAAGATGCAGAAGTGCAACAGCTTCACA CACGAGGTGGACAAGGACGAGGTGTTCAACATCCTGAGAGAGC TGTACGAGGAACTGGTGCCCAGCAGCGTGGAAAAGAAGGGCGA AGCCAACCAGCTGAGCAACAAGTTTCTGTACCCTCTGGTGGACC CCAACAGCCAGTCTGGAAAGGGAACAGCCAGCAGCGGCAGAAA GCCCAGATGGTACAACCTGAAGATTGCCGGCGATCCCGGAGGC TCTGGAGGAAGCTCCGAAGTCGAGTTTTCCCATGAGTACTGGAT GAGACACGCATTGACTCTCGCAAAGAGGGCTCGAGATGAACGC GAGGTGCCCGTGGGGGCAGTACTCGTGCTCAACAATCGCGTAA TCGGCGAAGGTTGGAATAGGGCAATCGGACTCCACGACCCCAC TGCACATGCGGAAATCATGGCCCTTCGACAGGGAGGGCTTGTG ATGCAGAATTATCGACTTTATGATGCGACGCTGTACGTCACGTTT GAACCTTGCGTAATGTGCGCGGGAGCTATGATTCACTCCCGCAT TGGACGAGTTGTATTCGGTGTTCGCAACGCCAAGACGGGTGCC GCAGGTTCACTGATGGACGTGCTGCATCATCCAGGCATGAACCA CCGGGTAGAAATCACAGAAGGCATATTGGCGGACGAATGTGCG GCGCTGTTGTGTCGTTTTTTTCGCATGCCCAGGCGGGTCTTTAA CGCCCAGAAAAAAGCACAATCCTCTACTGACGGCTCTTCTGGAT CTGAAACACCTGGCACAAGCGAGAGCGCCACCCCTGAGAGCTC TGGCTCCTGGGAAGAAGAGAAGAAGAAGTGGGAAGAAGATAAG AAAAAGGACCCGCTGGCCAAGATCCTGGGCAAGCTGGCTGAGT ACGGACTGATCCCTCTGTTCATCCCCTACACCGACAGCAACGAG CCCATCGTGAAAGAAATCAAGTGGATGGAAAAGTCCCGGAACCA GAGCGTGCGGCGGCTGGATAAGGACATGTTCATTCAGGCCCTG GAACGGTTCCTGAGCTGGGAGAGCTGGAACCTGAAAGTGAAAG AGGAATACGAGAAGGTCGAGAAAGAGTACAAGACCCTGGAAGA GAGGATCAAAGAGGACATCCAGGCTCTGAAGGCTCTGGAACAG TATGAGAAAGAGCGGCAAGAACAGCTGCTGCGGGACACCCTGA ACACCAACGAGTACCGGCTGAGCAAGAGAGGCCTTAGAGGCTG GCGGGAAATCATCCAGAAATGGCTGAAAATGGACGAGAACGAG CCCTCCGAGAAGTACCTGGAAGTGTTCAAGGACTACCAGCGGA AGCACCCTAGAGAGGCCGGCGATTACAGCGTGTACGAGTTCCT GTCCAAGAAAGAGAACCACTTCATCTGGCGGAATCACCCTGAGT ACCCCTACCTGTACGCCACCTTCTGCGAGATCGACAAGAAAAAG AAGGACGCCAAGCAGCAGGCCACCTTCACACTGGCCGATCCTA TCAATCACCCTCTGTGGGTCCGATTCGAGGAAAGAAGCGGCAG CAACCTGAACAAGTACAGAATCCTGACCGAGCAGCTGCACACC GAGAAGCTGAAGAAAAAGCTGACAGTGCAGCTGGACCGGCTGA TCTACCCTACAGAATCTGGCGGCTGGGAAGAGAAGGGCAAAGT GGACATTGTGCTGCTGCCCAGCCGGCAGTTCTACAACCAGATCT TCCTGGACATCGAGGAAAAGGGCAAGCACGCCTTCACCTACAA GGATGAGAGCATCAAGTTCCCTCTGAAGGGCACACTCGGCGGA GCCAGAGTGCAGTTCGACAGAGATCACCTGAGAAGATACCCTCA CAAGGTGGAAAGCGGCAACGTGGGCAGAATCTACTTCAACATG ACCGTGAACATCGAGCCTACAGAGTCCCCAGTGTCCAAGTCTCT GAAGATCCACCGGGACGACTTCCCCAAGGTGGTCAACTTCAAG

[1058]

		CCCAAAGAACTGACCGAGTGGATCAAGGACAGCAAGGGCAAGA AACTGAAGTCCGGCATCGAGTCCCTGGAAATCGGCCTGAGAGT GATGAGCATCGACCTGGGACAGAGACAGGCCGCTGCCGCCTCT ATTTTCGAGGTGGTGGATCAGAAGCCCGACATCGAAGGCAAGC TGTTTTTCCCAATCAAGGGCACCGAGCTGTATGCCGTGCACAGA GCCAGCTTCAACATCAAGCTGCCCGGCGAGACACTGGTCAAGA GCAGAGAAGTGCTGCGGAAGGCCAGAGAGGACAATCTGAAACT GATGAACCAGAAGCTCAACTTCCTGCGGAACGTGCTGCACTTCC AGCAGTTCGAGGACATCACCGAGAGAGAGAAGCGGGTCACCAA GTGGATCAGCAGACAAGAGAACAGCGACGTGCCCCTGGTGTAC CAGGATGAGCTGATCCAGATCCGCGAGCTGATGTACAAGCCTTA CAAGGACTGGGTCGCCTTCCTGAAGCAGCTCCACAAGAGACTG GAAGTCGAGATCGGCAAAGAAGTGAAGCACTGGCGGAAGTCCC TGAGCGACGGAAGAAAGGGCCTGTACGGCATCTCCCTGAAGAA CATCGACGAGATCGATCGGACCCGGAAGTTCCTGCTGAGATGG TCCCTGAGGCCTACCGAACCTGGCGAAGTGCGTAGACTGGAAC CCGGCCAGAGATTCGCCATCGACCAGCTGAATCACCTGAACGC CCTGAAAGAAGATCGGCTGAAGAAGATGGCCAACACCATCATCA TGCACGCCCTGGGCTACTGCTACGACGTGCGGAAGAAGAAATG GCAGGCTAAGAACCCCGCCTGCCAGATCATCCTGTTCGAGGAT CTGAGCAACTACAACCCCTACGAGGAAAGGTCCCGCTTCGAGA ACAGCAAGCTCATGAAGTGGTCCAGACGCGAGATCCCCAGACA GGTTGCACTGCAGGGCGAGATCTATGGCCTGCAAGTGGGAGAA GTGGGCGCTCAGTTCAGCAGCAGATTCCACGCCAAGACAGGCA GCCCTGGCATCAGATGTAGCGTCGTGACCAAAGAGAAGCTGCA GGACAATCGGTTCTTCAAGAATCTGCAGAGAGAGGGCAGACTG ACCCTGGACAAAATCGCCGTGCTGAAAGAGGGCGATCTGTACC CAGACAAAGGCGGCGAGAAGTTCATCAGCCTGAGCAAGGATCG GAAGTGCGTGACCACACACGCCGACATCAACGCCGCTCAGAAC CTGCAGAAGCGGTTCTGGACAAGAACCCACGGCTTCTACAAGG TGTACTGCAAGGCCTACCAGGTGGACGGCCAGACCGTGTACAT CCCTGAGAGCAAGGACCAGAAGCAGAAGATCATCGAAGAGTTC GGCGAGGGCTACTTCATTCTGAAGGACGGGGTGTACGAATGGG TCAACGCCGGCAAGCTGAAAATCAAGAAGGGCAGCTCCAAGCA GAGCAGCAGCGAGCTGGTGGATAGCGACATCCTGAAAGACAGC TTCGACCTGGCCTCCGAGCTGAAAGGCGAAAAGCTGATGCTGT ACAGGGACCCCAGCGGCAATGTGTTCCCCAGCGACAAATGGAT GGCCGCTGGCGTGTTCTTCGGAAAGCTGGAACGCATCCTGATC AGCAAGCTGACCAACCAGTACTCCATCAGCACCATCGAGGACG ACAGCAGCAAGCAGTCTATGAAAAGGCCGGCGGCCACGAAAAA GGCCGGCCAGGCAAAAAAGAAAAAGGGATCCTACCCATACGAT GTTCCAGATTACGCTTATCCCTACGACGTGCCTGATTATGCATAC CCATATGATGTCCCCGACTATGCCTAA
P153 폴리펩티드에서 BhCas12b GGSGGS-ABE8- Xten20	275	MAPKKKRKVGIHGVPAAATRSFILKIEPNEEVKKGLWKTHEVLNHGI AYYMNILKLIRQEAIYEHHEQDPKNPKKVSKAEIQAELWDFVLKMQK CNSFTHEVDKDEVFNILRELYEELVPSSVEKKGEANQLSNKFLYPLV DPNSQSGKGTASSGRKPRWYNLKIAGDPGGSGGSSEVEFSHEYW MRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTA HAEIMALRQGGLVMQNYRLYDATLYVTFEPCVMCAGAMIHSRIGRV VFGVRNAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLCRF FRMPRRVFNAQKKAQSSTDGSSGSETPGTSESATPESSGSWEEE KKKWEEDKKKDPLAKILGKLAEYGLIPLFIPYTDSNEPIVKEIKWMEK SRNQSVRRLDKDMFIQALERFLSWESWNLKVKEEYEKVEKEYKTL EERIKEDIQALKALEQYEKERQEQLLRDTLNTNEYRLSKRGLRGWR EIIQKWLKMDENEPSEKYLEVFKDYQRKHPREAGDYSVYEFLSKKE NHFIWRNHPEYPYLYATFCEIDKKKKDAKQQATFTLADPINHPLWV RFEERSGSNLNKYRILTEQLHTEKLKKKLTVQLDRLIYPTESGGWEE KGKVDIVLLPSRQFYNQIFLDIEEKGKHAFTYKDESIKFPLKGTLGGA RVQFDRDHLRRYPHKVESGNVGRIYFNMTVNIEPTESPVSKSLKIH

[1059]

		RDDFPKVVNFKPKELTEWIKDSKGKKLKSGIESLEIGLRVMSIDLGQ RQAAAASIFEVVDQKPDIEGKLFFPIKGTELYAVHRASFNIKLPGETL VKSREVLRKAREDNLKLMNQKLNFLRNVLHFQQFEDITEREKRVTK WISRQENSDVPLVYQDELIQIRELMYKPYKDWVAFLKQLHKRLEVEI GKEVKHWRKSLSDGRKGLYGISLKNIDEIDRTRKFLLRWSLRPTEP GEVRRLEPGQRFAIDQLNHLNALKEDRLKKMANTIIMHALGYCYDV RKKKWQAKNPACQIILFEDLSNYNPYEERSRFENSKLMKWSRREIP RQVALQGEIYGLQVGEVGAQFSSRFHAKTGSPGIRCSVVTKEKLQ DNRFFKNLQREGRLTLDKIAVLKEGDLYPDKGGEKFISLSKDRKCVT THADINAAQNLQKRFWTRTHGFYKVYCKAYQVDGQTVYIPESKDQ KQKIIEEFGEGYFILKDGVYEWVNAGKLKIKKGSSKQSSSELVDSDIL KDSFDLASELKGEKLMLYRDPSGNVFPSDKWMAAGVFFGKLERILI SKLTNQYSISTIEDDSSKQSMKRPAATKKAGQAKKKKGSYPYDVPD YAYPYDVPDYAYPYDVPDYA
K255 폴리뉴클레오티드 에서 BhCas12b GGSGGS-ABE8- Xten20	276	GCCACCATGGCCCCAAAGAAGAAGCGGAAGGTCGGTATCCACG GAGTCCCAGCAGCCGCCACCAGATCCTTCATCCTGAAGATCGA GCCCAACGAGGAAGTGAAGAAAGGCCTCTGGAAAACCCACGAG GTGCTGAACCACGGAATCGCCTACTACATGAATATCCTGAAGCT GATCCGGCAAGAGGCCATCTACGAGCACCACGAGCAGGACCCC AAGAATCCCAAGAAGGTGTCCAAGGCCGAGATCCAGGCCGAGC TGTGGGATTTCGTGCTGAAGATGCAGAAGTGCAACAGCTTCACA CACGAGGTGGACAAGGACGAGGTGTTCAACATCCTGAGAGAGC TGTACGAGGAACTGGTGCCCAGCAGCGTGGAAAAGAAGGGCGA AGCCAACCAGCTGAGCAACAAGTTTCTGTACCCTCTGGTGGACC CCAACAGCCAGTCTGGAAAGGGAACAGCCAGCAGCGGCAGAAA GCCCAGATGGTACAACCTGAAGATTGCCGGCGATCCCTCCTGG GAAGAAGAGAAGAAGAAGTGGGAAGAAGATAAGAAAAAGGACC CGCTGGCCAAGATCCTGGGCAAGCTGGCTGAGTACGGACTGAT CCCTCTGTTCATCCCCTACACCGACAGCAACGAGCCCATCGTGA AAGAAATCAAGTGGATGGAAAAGTCCCGGAACCAGAGCGTGCG GCGGCTGGATAAGGACATGTTCATTCAGGCCCTGGAACGGTTC CTGAGCTGGGAGAGCTGGAACCTGAAAGTGAAAGAGGAATACG AGAAGGTCGAGAAAGAGTACAAGACCCTGGAAGAGAGGATCAA AGGAGGCTCTGGAGGAAGCTCCGAAGTCGAGTTTTCCCATGAG TACTGGATGAGACACGCATTGACTCTCGCAAAGAGGGCTCGAG ATGAACGCGAGGTGCCCGTGGGGGCAGTACTCGTGCTCAACAA TCGCGTAATCGGCGAAGGTTGGAATAGGGCAATCGGACTCCAC GACCCCACTGCACATGCGGAAATCATGGCCCTTCGACAGGGAG GGCTTGTGATGCAGAATTATCGACTTTATGATGCGACGCTGTAC GTCACGTTTGAACCTTGCGTAATGTGCGCGGGAGCTATGATTCA CTCCCGCATTGGACGAGTTGTATTCGGTGTTCGCAACGCCAAGA CGGGTGCCGCAGGTTCACTGATGGACGTGCTGCATCATCCAGG CATGAACCACCGGGTAGAAATCACAGAAGGCATATTGGCGGAC GAATGTGCGGCGCTGTTGTGTCGTTTTTTTCGCATGCCCAGGCG GGTCTTTAACGCCCAGAAAAAAGCACAATCCTCTACTGACGGCT CTTCTGGATCTGAAACACCTGGCACAAGCGAGAGCGCCACCCC TGAGAGCTCTGGCGAGGACATCCAGGCTCTGAAGGCTCTGGAA CAGTATGAGAAAGAGCGGCAAGAACAGCTGCTGCGGGACACCC TGAACACCAACGAGTACCGGCTGAGCAAGAGAGGCCTTAGAGG CTGGCGGGAAATCATCCAGAAATGGCTGAAAATGGACGAGAAC GAGCCCTCCGAGAAGTACCTGGAAGTGTTCAAGGACTACCAGC GGAAGCACCCTAGAGAGGCCGGCGATTACAGCGTGTACGAGTT CCTGTCCAAGAAAGAGAACCACTTCATCTGGCGGAATCACCCTG AGTACCCCTACCTGTACGCCACCTTCTGCGAGATCGACAAGAAA AAGAAGGACGCCAAGCAGCAGGCCACCTTCACACTGGCCGATC CTATCAATCACCCTCTGTGGGTCCGATTCGAGGAAAGAAGCGGC AGCAACCTGAACAAGTACAGAATCCTGACCGAGCAGCTGCACAC CGAGAAGCTGAAGAAAAAGCTGACAGTGCAGCTGGACCGGCTG ATCTACCCTACAGAATCTGGCGGCTGGGAAGAGAAGGGCAAAG

[1060]

[1061]

		TGGACATTGTGCTGCTGCCCAGCCGGCAGTTCTACAACCAGATC TTCCTGGACATCGAGGAAAAGGGCAAGCACGCCTTCACCTACAA GGATGAGAGCATCAAGTTCCCTCTGAAGGGCACACTCGGCGGA GCCAGAGTGCAGTTCGACAGAGATCACCTGAGAAGATACCCTCA CAAGGTGGAAAGCGGCAACGTGGGCAGAATCTACTTCAACATG ACCGTGAACATCGAGCCTACAGAGTCCCCAGTGTCCAAGTCTCT GAAGATCCACCGGGACGACTTCCCCAAGGTGGTCAACTTCAAG CCCAAAGAACTGACCGAGTGGATCAAGGACAGCAAGGGCAAGA AACTGAAGTCCGGCATCGAGTCCCTGGAAATCGGCCTGAGAGT GATGAGCATCGACCTGGGACAGAGACAGGCCGCTGCCGCCTCT ATTTTCGAGGTGGTGGATCAGAAGCCCGACATCGAAGGCAAGC TGTTTTTCCCAATCAAGGGCACCGAGCTGTATGCCGTGCACAGA GCCAGCTTCAACATCAAGCTGCCCGGCGAGACACTGGTCAAGA GCAGAGAAGTGCTGCGGAAGGCCAGAGAGGACAATCTGAAACT GATGAACCAGAAGCTCAACTTCCTGCGGAACGTGCTGCACTTCC AGCAGTTCGAGGACATCACCGAGAGAGAGAAGCGGGTCACCAA GTGGATCAGCAGACAAGAGAACAGCGACGTGCCCCTGGTGTAC CAGGATGAGCTGATCCAGATCCGCGAGCTGATGTACAAGCCTTA CAAGGACTGGGTCGCCTTCCTGAAGCAGCTCCACAAGAGACTG GAAGTCGAGATCGGCAAAGAAGTGAAGCACTGGCGGAAGTCCC TGAGCGACGGAAGAAAGGGCCTGTACGGCATCTCCCTGAAGAA CATCGACGAGATCGATCGGACCCGGAAGTTCCTGCTGAGATGG TCCCTGAGGCCTACCGAACCTGGCGAAGTGCGTAGACTGGAAC CCGGCCAGAGATTCGCCATCGACCAGCTGAATCACCTGAACGC CCTGAAAGAAGATCGGCTGAAGAAGATGGCCAACACCATCATCA TGCACGCCCTGGGCTACTGCTACGACGTGCGGAAGAAGAAATG GCAGGCTAAGAACCCCGCCTGCCAGATCATCCTGTTCGAGGAT CTGAGCAACTACAACCCCTACGAGGAAAGGTCCCGCTTCGAGA ACAGCAAGCTCATGAAGTGGTCCAGACGCGAGATCCCCAGACA GGTTGCACTGCAGGGCGAGATCTATGGCCTGCAAGTGGGAGAA GTGGGCGCTCAGTTCAGCAGCAGATTCCACGCCAAGACAGGCA GCCCTGGCATCAGATGTAGCGTCGTGACCAAAGAGAAGCTGCA GGACAATCGGTTCTTCAAGAATCTGCAGAGAGAGGGCAGACTG ACCCTGGACAAAATCGCCGTGCTGAAAGAGGGCGATCTGTACC CAGACAAAGGCGGCGAGAAGTTCATCAGCCTGAGCAAGGATCG GAAGTGCGTGACCACACACGCCGACATCAACGCCGCTCAGAAC CTGCAGAAGCGGTTCTGGACAAGAACCCACGGCTTCTACAAGG TGTACTGCAAGGCCTACCAGGTGGACGGCCAGACCGTGTACAT CCCTGAGAGCAAGGACCAGAAGCAGAAGATCATCGAAGAGTTC GGCGAGGGCTACTTCATTCTGAAGGACGGGGTGTACGAATGGG TCAACGCCGGCAAGCTGAAAATCAAGAAGGGCAGCTCCAAGCA GAGCAGCAGCGAGCTGGTGGATAGCGACATCCTGAAAGACAGC TTCGACCTGGCCTCCGAGCTGAAAGGCGAAAAGCTGATGCTGT ACAGGGACCCCAGCGGCAATGTGTTCCCCAGCGACAAATGGAT GGCCGCTGGCGTGTTCTTCGGAAAGCTGGAACGCATCCTGATC AGCAAGCTGACCAACCAGTACTCCATCAGCACCATCGAGGACG ACAGCAGCAAGCAGTCTATGAAAAGGCCGGCGGCCACGAAAAA GGCCGGCCAGGCAAAAAAGAAAAAGGGATCCTACCCATACGAT GTTCCAGATTACGCTTATCCCTACGACGTGCCTGATTATGCATAC CCATATGATGTCCCCGACTATGCCTAA
K255 폴리펩티드에서 BhCas12b GGSGGS-ABE8- Xten20	277	MAPKKKRKVGIHGVPAAATRSFILKIEPNEEVKKGLWKTHEVLNHGI AYYMNILKLIRQEAIYEHHEQDPKNPKKVSKAEIQAELWDFVLKMQK CNSFTHEVDKDEVFNILRELYEELVPSSVEKKGEANQLSNKFLYPLV DPNSQSGKGTASSGRKPRWYNLKIAGDPSWEEEKKKWEEDKKKD PLAKILGKLAEYGLIPLFIPYTDSNEPIVKEIKWMEKSRNQSVRRLDK DMFIQALERFLSWESWNLKVKEEYEKVEKEYKTLEERIKGGSGGSS EVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNR AIGLHDPTAHAEIMALRQGGLVMQNYRLYDATLYVTFEPCVMCAGA MIHSRIGRVVFGVRNAKTGAAGSLMDVLHHPGMNHRVEITEGILAD

[1062]

		ECAALLCRFFRMPRRVFNAQKKAQSSTDGSSGSETPGTSESATPE SSGEDIQALKALEQYEKERQEQLLRDTLNTNEYRLSKRGLRGWREI IQKWLKMDENEPSEKYLEVFKDYQRKHPREAGDYSVYEFLSKKEN HFIWRNHPEYPYLYATFCEIDKKKKDAKQQATFTLADPINHPLWVRF EERSGSNLNKYRILTEQLHTEKLKKKLTVQLDRLIYPTESGGWEEK GKVDIVLLPSRQFYNQIFLDIEEKGKHAFTYKDESIKFPLKGTLGGAR VQFDRDHLRRYPHKVESGNVGRIYFNMTVNIEPTESPVSKSLKIHR DDFPKVVNFKPKELTEWIKDSKGKKLKSGIESLEIGLRVMSIDLGQR QAAAASIFEVVDQKPDIEGKLFFPIKGTELYAVHRASFNIKLPGETLV KSREVLRKAREDNLKLMNQKLNFLRNVLHFQQFEDITEREKRVTKW ISRQENSDVPLVYQDELIQIRELMYKPYKDWVAFLKQLHKRLEVEIG KEVKHWRKSLSDGRKGLYGISLKNIDEIDRTRKFLLRWSLRPTEPG EVRRLEPGQRFAIDQLNHLNALKEDRLKKMANTIIMHALGYCYDVR KKKWQAKNPACQIILFEDLSNYNPYEERSRFENSKLMKWSRREIPR QVALQGEIYGLQVGEVGAQFSSRFHAKTGSPGIRCSVVTKEKLQD NRFFKNLQREGRLTLDKIAVLKEGDLYPDKGGEKFISLSKDRKCVTT HADINAAQNLQKRFWTRTHGFYKVYCKAYQVDGQTVYIPESKDQK QKIIEEFGEGYFILKDGVYEWVNAGKLKIKKGSSKQSSSELVDSDILK DSFDLASELKGEKLMLYRDPSGNVFPSDKWMAAGVFFGKLERILIS KLTNQYSISTIEDDSSKQSMKRPAATKKAGQAKKKKGSYPYDVPDY AYPYDVPDYAYPYDVPDYA
D306 폴리뉴클레오티드 에서 BhCas12b GGSGGS-ABE8- Xten20	278	GCCACCATGGCCCCAAAGAAGAAGCGGAAGGTCGGTATCCACG GAGTCCCAGCAGCCGCCACCAGATCCTTCATCCTGAAGATCGA GCCCAACGAGGAAGTGAAGAAAGGCCTCTGGAAAACCCACGAG GTGCTGAACCACGGAATCGCCTACTACATGAATATCCTGAAGCT GATCCGGCAAGAGGCCATCTACGAGCACCACGAGCAGGACCCC AAGAATCCCAAGAAGGTGTCCAAGGCCGAGATCCAGGCCGAGC TGTGGGATTTCGTGCTGAAGATGCAGAAGTGCAACAGCTTCACA CACGAGGTGGACAAGGACGAGGTGTTCAACATCCTGAGAGAGC TGTACGAGGAACTGGTGCCCAGCAGCGTGGAAAAGAAGGGCGA AGCCAACCAGCTGAGCAACAAGTTTCTGTACCCTCTGGTGGACC CCAACAGCCAGTCTGGAAAGGGAACAGCCAGCAGCGGCAGAAA GCCCAGATGGTACAACCTGAAGATTGCCGGCGATCCCTCCTGG GAAGAAGAGAAGAAGAAGTGGGAAGAAGATAAGAAAAAGGACC CGCTGGCCAAGATCCTGGGCAAGCTGGCTGAGTACGGACTGAT CCCTCTGTTCATCCCCTACACCGACAGCAACGAGCCCATCGTGA AAGAAATCAAGTGGATGGAAAAGTCCCGGAACCAGAGCGTGCG GCGGCTGGATAAGGACATGTTCATTCAGGCCCTGGAACGGTTC CTGAGCTGGGAGAGCTGGAACCTGAAAGTGAAAGAGGAATACG AGAAGGTCGAGAAAGAGTACAAGACCCTGGAAGAGAGGATCAA AGAGGACATCCAGGCTCTGAAGGCTCTGGAACAGTATGAGAAA GAGCGGCAAGAACAGCTGCTGCGGGACACCCTGAACACCAACG AGTACCGGCTGAGCAAGAGAGGCCTTAGAGGCTGGCGGGAAAT CATCCAGAAATGGCTGAAAATGGACGGAGGCTCTGGAGGAAGC TCCGAAGTCGAGTTTTCCCATGAGTACTGGATGAGACACGCATT GACTCTCGCAAAGAGGGCTCGAGATGAACGCGAGGTGCCCGTG GGGGCAGTACTCGTGCTCAACAATCGCGTAATCGGCGAAGGTT GGAATAGGGCAATCGGACTCCACGACCCCACTGCACATGCGGA AATCATGGCCCTTCGACAGGGAGGGCTTGTGATGCAGAATTATC GACTTTATGATGCGACGCTGTACGTCACGTTTGAACCTTGCGTA ATGTGCGCGGGAGCTATGATTCACTCCCGCATTGGACGAGTTGT ATTCGGTGTTCGCAACGCCAAGACGGGTGCCGCAGGTTCACTG ATGGACGTGCTGCATCATCCAGGCATGAACCACCGGGTAGAAAT CACAGAAGGCATATTGGCGGACGAATGTGCGGCGCTGTTGTGT CGTTTTTTTCGCATGCCCAGGCGGGTCTTTAACGCCCAGAAAAA AGCACAATCCTCTACTGACGGCTCTTCTGGATCTGAAACACCTG GCACAAGCGAGAGCGCCACCCCTGAGAGCTCTGGCGAGAACGA GCCCTCCGAGAAGTACCTGGAAGTGTTCAAGGACTACCAGCGG AAGCACCCTAGAGAGGCCGGCGATTACAGCGTGTACGAGTTCC

		TGTCCAAGAAAGAGAACCACTTCATCTGGCGGAATCACCCTGAG TACCCCTACCTGTACGCCACCTTCTGCGAGATCGACAAGAAAAA GAAGGACGCCAAGCAGCAGGCCACCTTCACACTGGCCGATCCT ATCAATCACCCTCTGTGGGTCCGATTCGAGGAAAGAAGCGGCA GCAACCTGAACAAGTACAGAATCCTGACCGAGCAGCTGCACAC CGAGAAGCTGAAGAAAAAGCTGACAGTGCAGCTGGACCGGCTG ATCTACCCTACAGAATCTGGCGGCTGGGAAGAGAAGGGCAAAG TGGACATTGTGCTGCTGCCCAGCCGGCAGTTCTACAACCAGATC TTCCTGGACATCGAGGAAAAGGGCAAGCACGCCTTCACCTACAA GGATGAGAGCATCAAGTTCCCTCTGAAGGGCACACTCGGCGGA GCCAGAGTGCAGTTCGACAGAGATCACCTGAGAAGATACCCTCA CAAGGTGGAAAGCGGCAACGTGGGCAGAATCTACTTCAACATG ACCGTGAACATCGAGCCTACAGAGTCCCCAGTGTCCAAGTCTCT GAAGATCCACCGGGACGACTTCCCCAAGGTGGTCAACTTCAAG CCCAAAGAACTGACCGAGTGGATCAAGGACAGCAAGGGCAAGA AACTGAAGTCCGGCATCGAGTCCCTGGAAATCGGCCTGAGAGT GATGAGCATCGACCTGGGACAGAGACAGGCCGCTGCCGCCTCT ATTTTCGAGGTGGTGGATCAGAAGCCCGACATCGAAGGCAAGC TGTTTTTCCCAATCAAGGGCACCGAGCTGTATGCCGTGCACAGA GCCAGCTTCAACATCAAGCTGCCCGGCGAGACACTGGTCAAGA GCAGAGAAGTGCTGCGGAAGGCCAGAGAGGACAATCTGAAACT GATGAACCAGAAGCTCAACTTCCTGCGGAACGTGCTGCACTTCC AGCAGTTCGAGGACATCACCGAGAGAGAGAAGCGGGTCACCAA GTGGATCAGCAGACAAGAGAACAGCGACGTGCCCCTGGTGTAC CAGGATGAGCTGATCCAGATCCGCGAGCTGATGTACAAGCCTTA CAAGGACTGGGTCGCCTTCCTGAAGCAGCTCCACAAGAGACTG GAAGTCGAGATCGGCAAAGAAGTGAAGCACTGGCGGAAGTCCC TGAGCGACGGAAGAAAGGGCCTGTACGGCATCTCCCTGAAGAA CATCGACGAGATCGATCGGACCCGGAAGTTCCTGCTGAGATGG TCCCTGAGGCCTACCGAACCTGGCGAAGTGCGTAGACTGGAAC CCGGCCAGAGATTCGCCATCGACCAGCTGAATCACCTGAACGC CCTGAAAGAAGATCGGCTGAAGAAGATGGCCAACACCATCATCA TGCACGCCCTGGGCTACTGCTACGACGTGCGGAAGAAGAAATG GCAGGCTAAGAACCCCGCCTGCCAGATCATCCTGTTCGAGGAT CTGAGCAACTACAACCCCTACGAGGAAAGGTCCCGCTTCGAGA ACAGCAAGCTCATGAAGTGGTCCAGACGCGAGATCCCCAGACA GGTTGCACTGCAGGGCGAGATCTATGGCCTGCAAGTGGGAGAA GTGGGCGCTCAGTTCAGCAGCAGATTCCACGCCAAGACAGGCA GCCCTGGCATCAGATGTAGCGTCGTGACCAAAGAGAAGCTGCA GGACAATCGGTTCTTCAAGAATCTGCAGAGAGAGGGCAGACTG ACCCTGGACAAAATCGCCGTGCTGAAAGAGGGCGATCTGTACC CAGACAAAGGCGGCGAGAAGTTCATCAGCCTGAGCAAGGATCG GAAGTGCGTGACCACACACGCCGACATCAACGCCGCTCAGAAC CTGCAGAAGCGGTTCTGGACAAGAACCCACGGCTTCTACAAGG TGTACTGCAAGGCCTACCAGGTGGACGGCCAGACCGTGTACAT CCCTGAGAGCAAGGACCAGAAGCAGAAGATCATCGAAGAGTTC GGCGAGGGCTACTTCATTCTGAAGGACGGGGTGTACGAATGGG TCAACGCCGGCAAGCTGAAAATCAAGAAGGGCAGCTCCAAGCA GAGCAGCAGCGAGCTGGTGGATAGCGACATCCTGAAAGACAGC TTCGACCTGGCCTCCGAGCTGAAAGGCGAAAAGCTGATGCTGT ACAGGGACCCCAGCGGCAATGTGTTCCCCAGCGACAAATGGAT GGCCGCTGGCGTGTTCTTCGGAAAGCTGGAACGCATCCTGATC AGCAAGCTGACCAACCAGTACTCCATCAGCACCATCGAGGACG ACAGCAGCAAGCAGTCTATGAAAAGGCCGGCGGCCACGAAAAA GGCCGGCCAGGCAAAAAAGAAAAAGGGATCCTACCCATACGAT GTTCCAGATTACGCTTATCCCTACGACGTGCCTGATTATGCATAC CCATATGATGTCCCCGACTATGCCTAA
D306 폴리펩티드에서	279	MAPKKKRKVGIHGVPAAATRSFILKIEPNEEVKKGLWKTHEVLNHGI AYYMNILKLIRQEAIYEHHEQDPKNPKKVSKAEIQAELWDFVLKMQK

[1063]

[1064]

BhCas12b GGSGGS-ABE8-Xten20		CNSFTHEVDKDEVFNILRELYEELVPSSVEKKGEANQLSNKFLYPLV DPNSQSGKGTASSGRKPRWYNLKIAGDPSWEEEKKKWEEDKKKD PLAKILGKLAEYGLIPLFIPYTDSNEPIVKEIKWMEKSRNQSVRRLDK DMFIQALERFLSWESWNLKVKEEYEKVEKEYKTLEERIKEDIQALKA LEQYEKERQEQLLRDTLNTNEYRLSKRGLRGWREIIQKWLKMDGG SGGSSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIG EGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLYDATLYVTFEPC VMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHHPGMNHRVEI TEGILADECAALLCRFFRMPRRVFNAQKKAQSSTDGSSGSETPGT SESATPESS유전자PSEKYLEVFKDYQRKHPREAGDYSVYEFLSKK ENHFIWRNHPEYPYLYATFCEIDKKKKDAKQQATFTLADPINHPLW VRFEERSGSNLNKYRILTEQLHTEKLKKKLTVQLDRLIYPTESGGWE EKGKVDIVLLPSRQFYNQIFLDIEEKGKHAFTYKDESIKFPLKGTLGG ARVQFDRDHLRRYPHKVESGNVGRIYFNMTVNIEPTESPVSKSLKI HRDDFPKVVNFKPKELTEWIKDSKGKKLKSGIESLEIGLRVMSIDLG QRQAAAASIFEVVDQKPDIEGKLFFPIKGTELYAVHRASFNIKLPGET LVKSREVLRKAREDNLKLMNQKLNFLRNVLHFQQFEDITEREKRVT KWISRQENSDVPLVYQDELIQIRELMYKPYKDWVAFLKQLHKRLEV EIGKEVKHWRKSLSDGRKGLYGISLKNIDEIDRTRKFLLRWSLRPTE PGEVRRLEPGQRFAIDQLNHLNALKEDRLKKMANTIIMHALGYCYD VRKKKWQAKNPACQIILFEDLSNYNPYEERSRFENSKLMKWSRREI PRQVALQGEIYGLQVGEVGAQFSSRFHAKTGSPGIRCSVVTKEKLQ DNRFFKNLQREGRLTLDKIAVLKEGDLYPDKGGEKFISLSKDRKCVT THADINAAQNLQKRFWTRTHGFYKVYCKAYQVDGQTVYIPESKDQ KQKIIEEFGEGYFILKDGVYEWVNAGKLKIKKGSSKQSSSELVDSDIL KDSFDLASELKGEKLMLYRDPSGNVFPSDKWMAAGVFFGKLERILI SKLTNQYSISTIEDDSSKQSMKRPAATKKAGQAKKKKGSYPYDVPD YAYPYDVPDYAYPYDVPDYA
D980 폴리뉴클레오티드 에서 BhCas12b GGSGGS-ABE8-Xten20	280	GCCACCATGGCCCCAAAGAAGAAGCGGAAGGTCGGTATCCACG GAGTCCCAGCAGCCGCCACCAGATCCTTCATCCTGAAGATCGA GCCCAACGAGGAAGTGAAGAAAGGCCTCTGGAAAACCCACGAG GTGCTGAACCACGGAATCGCCTACTACATGAATATCCTGAAGCT GATCCGGCAAGAGGCCATCTACGAGCACCACGAGCAGGACCCC AAGAATCCCAAGAAGGTGTCCAAGGCCGAGATCCAGGCCGAGC TGTGGGATTTCGTGCTGAAGATGCAGAAGTGCAACAGCTTCACA CACGAGGTGGACAAGGACGAGGTGTTCAACATCCTGAGAGAGC TGTACGAGGAACTGGTGCCCAGCAGCGTGGAAAAGAAGGGCGA AGCCAACCAGCTGAGCAACAAGTTTCTGTACCCTCTGGTGGACC CCAACAGCCAGTCTGGAAAGGGAACAGCCAGCAGCGGCAGAAA GCCCAGATGGTACAACCTGAAGATTGCCGGCGATCCCTCCTGG GAAGAAGAGAAGAAGAAGTGGGAAGAAGATAAGAAAAAGGACC CGCTGGCCAAGATCCTGGGCAAGCTGGCTGAGTACGGACTGAT CCCTCTGTTCATCCCCTACACCGACAGCAACGAGCCCATCGTGA AAGAAATCAAGTGGATGGAAAAGTCCCGGAACCAGAGCGTGCG GCGGCTGGATAAGGACATGTTCATTCAGGCCCTGGAACGGTTC CTGAGCTGGGAGAGCTGGAACCTGAAAGTGAAAGAGGAATACG AGAAGGTCGAGAAAGAGTACAAGACCCTGGAAGAGAGGATCAA AGAGGACATCCAGGCTCTGAAGGCTCTGGAACAGTATGAGAAA GAGCGGCAAGAACAGCTGCTGCGGGACACCCTGAACACCAACG AGTACCGGCTGAGCAAGAGAGGCCTTAGAGGCTGGCGGGAAAT CATCCAGAAATGGCTGAAAATGGACGAGAACGAGCCCTCCGAG AAGTACCTGGAAGTGTTCAAGGACTACCAGCGGAAGCACCCTA GAGAGGCCGGCGATTACAGCGTGTACGAGTTCCTGTCCAAGAA AGAGAACCACTTCATCTGGCGGAATCACCCTGAGTACCCCTACC TGTACGCCACCTTCTGCGAGATCGACAAGAAAAAGAAGGACGC CAAGCAGCAGGCCACCTTCACACTGGCCGATCCTATCAATCACC CTCTGTGGGTCCGATTCGAGGAAAGAAGCGGCAGCAACCTGAA CAAGTACAGAATCCTGACCGAGCAGCTGCACACCGAGAAGCTG AAGAAAAAGCTGACAGTGCAGCTGGACCGGCTGATCTACCCTAC

[1065]

		AGAATCTGGCGGCTGGGAAGAGAAGGGCAAAGTGGACATTGTG CTGCTGCCCAGCCGGCAGTTCTACAACCAGATCTTCCTGGACAT CGAGGAAAAGGGCAAGCACGCCTTCACCTACAAGGATGAGAGC ATCAAGTTCCCTCTGAAGGGCACACTCGGCGGAGCCAGAGTGC AGTTCGACAGAGATCACCTGAGAAGATACCCTCACAAGGTGGAA AGCGGCAACGTGGGCAGAATCTACTTCAACATGACCGTGAACAT CGAGCCTACAGAGTCCCCAGTGTCCAAGTCTCTGAAGATCCACC GGGACGACTTCCCCAAGGTGGTCAACTTCAAGCCCAAAGAACT GACCGAGTGGATCAAGGACAGCAAGGGCAAGAAACTGAAGTCC GGCATCGAGTCCCTGGAAATCGGCCTGAGAGTGATGAGCATCG ACCTGGGACAGAGACAGGCCGCTGCCGCCTCTATTTTCGAGGT GGTGGATCAGAAGCCCGACATCGAAGGCAAGCTGTTTTTCCCAA TCAAGGGCACCGAGCTGTATGCCGTGCACAGAGCCAGCTTCAA CATCAAGCTGCCCGGCGAGACACTGGTCAAGAGCAGAGAAGTG CTGCGGAAGGCCAGAGAGGACAATCTGAAACTGATGAACCAGA AGCTCAACTTCCTGCGGAACGTGCTGCACTTCCAGCAGTTCGAG GACATCACCGAGAGAGAGAAGCGGGTCACCAAGTGGATCAGCA GACAAGAGAACAGCGACGTGCCCCTGGTGTACCAGGATGAGCT GATCCAGATCCGCGAGCTGATGTACAAGCCTTACAAGGACTGG GTCGCCTTCCTGAAGCAGCTCCACAAGAGACTGGAAGTCGAGA TCGGCAAAGAAGTGAAGCACTGGCGGAAGTCCCTGAGCGACGG AAGAAAGGGCCTGTACGGCATCTCCCTGAAGAACATCGACGAG ATCGATCGGACCCGGAAGTTCCTGCTGAGATGGTCCCTGAGGC CTACCGAACCTGGCGAAGTGCGTAGACTGGAACCCGGCCAGAG ATTCGCCATCGACCAGCTGAATCACCTGAACGCCCTGAAAGAAG ATCGGCTGAAGAAGATGGCCAACACCATCATCATGCACGCCCTG GGCTACTGCTACGACGTGCGGAAGAAGAAATGGCAGGCTAAGA ACCCCGCCTGCCAGATCATCCTGTTCGAGGATCTGAGCAACTAC AACCCCTACGAGGAAAGGTCCCGCTTCGAGAACAGCAAGCTCA TGAAGTGGTCCAGACGCGAGATCCCCAGACAGGTTGCACTGCA GGGCGAGATCTATGGCCTGCAAGTGGGAGAAGTGGGCGCTCAG TTCAGCAGCAGATTCCACGCCAAGACAGGCAGCCCTGGCATCA GATGTAGCGTCGTGACCAAAGAGAAGCTGCAGGACAATCGGTT CTTCAAGAATCTGCAGAGAGAGGGCAGACTGACCCTGGACAAA ATCGCCGTGCTGAAAGAGGGCGATCTGTACCCAGACAAAGGCG GCGAGAAGTTCATCAGCCTGAGCAAGGATCGGAAGTGCGTGAC CACACACGCCGACATCAACGCCGCTCAGAACCTGCAGAAGCGG TTCTGGACAAGAACCCACGGCTTCTACAAGGTGTACTGCAAGGC CTACCAGGTGGACGGAGGCTCTGGAGGAAGCTCCGAAGTCGAG TTTTCCCATGAGTACTGGATGAGACACGCATTGACTCTCGCAAA GAGGGCTCGAGATGAACGCGAGGTGCCCGTGGGGGCAGTACT CGTGCTCAACAATCGCGTAATCGGCGAAGGTTGGAATAGGGCA ATCGGACTCCACGACCCCACTGCACATGCGGAAATCATGGCCC TTCGACAGGGAGGGCTTGTGATGCAGAATTATCGACTTTATGAT GCGACGCTGTACGTCACGTTTGAACCTTGCGTAATGTGCGCGG GAGCTATGATTCACTCCCGCATTGGACGAGTTGTATTCGGTGTT CGCAACGCCAAGACGGGTGCCGCAGGTTCACTGATGGACGTGC TGCATCATCCAGGCATGAACCACCGGGTAGAAATCACAGAAGG CATATTGGCGGACGAATGTGCGGCGCTGTTGTGTCGTTTTTTTC GCATGCCCAGGCGGGTCTTTAACGCCCAGAAAAAAGCACAATC CTCTACTGACGGCTCTTCTGGATCTGAAACACCTGGCACAAGCG AGAGCGCCACCCCTGAGAGCTCTGGCGGCCAGACCGTGTACAT CCCTGAGAGCAAGGACCAGAAGCAGAAGATCATCGAAGAGTTC GGCGAGGGCTACTTCATTCTGAAGGACGGGGTGTACGAATGGG TCAACGCCGGCAAGCTGAAAATCAAGAAGGGCAGCTCCAAGCA GAGCAGCAGCGAGCTGGTGGATAGCGACATCCTGAAAGACAGC TTCGACCTGGCCTCCGAGCTGAAAGGCGAAAAGCTGATGCTGT ACAGGGACCCCAGCGGCAATGTGTTCCCCAGCGACAAATGGAT GGCCGCTGGCGTGTTCTTCGGAAAGCTGGAACGCATCCTGATC

		AGCAAGCTGACCAACCAGTACTCCATCAGCACCATCGAGGACG ACAGCAGCAAGCAGTCTATGAAAAGGCCGGCGGCCACGAAAAA GGCCGGCCAGGCAAAAAAGAAAAAGGGATCCTACCCATACGAT GTTCCAGATTACGCTTATCCCTACGACGTGCCTGATTATGCATAC CCATATGATGTCCCCGACTATGCCTAA
D980 폴리렙티드에서 BhCas12b GGSGGS-ABE8- Xten20	281	MAPKKKRKVGIHGVPAAATRSFILKIEPNEEVKKGLWKTHEVLNHGI AYYMNILKLIRQEAIYEHHEQDPKNPKKVSKAEIQAELWDFVLKMQK CNSFTHEVDKDEVFNILRELYEELVPSSVEKKGEANQLSNKFLYPLV DPNSQSGKGTASSGRKPRWYNLKIAGDPSWEEEKKKWEEDKKKD PLAKILGKLAEYGLIPLFIPYTDSNEPIVKEIKWMEKSRNQSVRRLDK DMFIQALERFLSWESWNLKVKEEYEKVEKEYKTLEERIKEDIQALKA LEQYEKERQEQLLRDTLNTNEYRLSKRGLRGWREIIQKWLKMDEN EPSEKYLEVFKDYQRKHPREAGDYSVYEFLSKKENHFIWRNHPEY PYLYATFCEIDKKKKDAKQQATFTLADPINHPLWVRFEERSGSNLN KYRILTEQLHTEKLKKKLTVQLDRLIYPTESGGWEEKGKVDIVLLPS RQFYNQIFLDIEEKGKHAFTYKDESIKFPLKGTLGGARVQFDRDHLR RYPHKVESGNVGRIYFNMTVNIEPTESPVSKSLKIHRDDFPKVVNFK PKELTEWIKDSKGKKLKSGIESLEIGLRVMSIDLGQRQAAAASIFEVV DQKPDIEGKLFFPIKGTELYAVHRASFNIKLPGETLVKSREVLRKAR EDNLKLMNQKLNFLRNVLHFQQFEDITEREKRVTKWISRQENSDVP LVYQDELIQIRELMYKPYKDWVAFLKQLHKRLEVEIGKEVKHWRKS LSDGRKGLYGISLKNIDEIDRTRKFLLRWSLRPTEPGEVRRLEPGQR FAIDQLNHLNALKEDRLKKMANTIIMHALGYCYDVRKKKWQAKNPA CQIILFEDLSNYNPYEERSRFENSKLMKWSRREIPRQVALQGEIYGL QVGEVGAQFSSRFHAKTGSPGIRCSVVTKEKLQDNRFFKNLQREG RLTLDKIAVLKEGDLYPDKGGEKFISLSKDRKCVTTHADINAAQNLQ KRFWTRTHGFYKVYCKAYQVDGGSGGSSEVEFSHEYWMRHALTL AKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAEIMALR QGGLVMQNYRLYDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNA KTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLCRFFRMPRRV FNAQKKAQSSTDGSSGSETPGTSESATPESSGGQTVYIPESKDQK QKIIEEFGEGYFILKDGVYEWVNAGKLKIKKGSSKQSSSELVDSDILK DSFDLASELKGEKLMLYRDPSGNVFPSDKWMAAGVFFGKLERILIS KLTNQYSISTIEDDSSKQSMKRPAATKKAGQAKKKKGSYPYDVPDY AYPYDVPDYAYPYDVPDYA
K1019 폴리뉴클레오티드 에서 BhCas12b GGSGGS-ABE8- Xten20	282	GCCACCATGGCCCCAAAGAAGAAGCGGAAGGTCGGTATCCACG GAGTCCCAGCAGCCGCCACCAGATCCTTCATCCTGAAGATCGA GCCCAACGAGGAAGTGAAGAAAGGCCTCTGGAAAACCCACGAG GTGCTGAACCACGGAATCGCCTACTACATGAATATCCTGAAGCT GATCCGGCAAGAGGCCATCTACGAGCACCACGAGCAGGACCCC AAGAATCCCAAGAAGGTGTCCAAGGCCGAGATCCAGGCCGAGC TGTGGGATTTCGTGCTGAAGATGCAGAAGTGCAACAGCTTCACA CACGAGGTGGACAAGGACGAGGTGTTCAACATCCTGAGAGAGC TGTACGAGGAACTGGTGCCCAGCAGCGTGGAAAAGAAGGGCGA AGCCAACCAGCTGAGCAACAAGTTTCTGTACCCTCTGGTGGACC CCAACAGCCAGTCTGGAAAGGGAACAGCCAGCAGCGGCAGAAA GCCCAGATGGTACAACCTGAAGATTGCCGGCGATCCCTCCTGG GAAGAAGAGAAGAAGAAGTGGGAAGAAGATAAGAAAAAGGACC CGCTGGCCAAGATCCTGGGCAAGCTGGCTGAGTACGGACTGAT CCCTCTGTTCATCCCCTACACCGACAGCAACGAGCCCATCGTGA AAGAAATCAAGTGGATGGAAAAGTCCCGGAACCAGAGCGTGCG GCGGCTGGATAAGGACATGTTCATTCAGGCCCTGGAACGGTTC CTGAGCTGGGAGAGCTGGAACCTGAAAGTGAAAGAGGAATACG AGAAGGTCGAGAAAGAGTACAAGACCCTGGAAGAGAGGATCAA AGAGGACATCCAGGCTCTGAAGGCTCTGGAACAGTATGAGAAA GAGCGGCAAGAACAGCTGCTGCGGGACACCCTGAACACCAACG AGTACCGGCTGAGCAAGAGAGGCCTTAGAGGCTGGCGGGAAAT CATCCAGAAATGGCTGAAAATGGACGAGAACGAGCCCTCCGAG AAGTACCTGGAAGTGTTCAAGGACTACCAGCGGAAGCACCCTA

[1066]

[1067]

		GAGAGGCCGGCGATTACAGCGTGTACGAGTTCCTGTCCAAGAA AGAGAACCACTTCATCTGGCGGAATCACCCTGAGTACCCCTACC TGTACGCCACCTTCTGCGAGATCGACAAGAAAAAGAAGGACGC CAAGCAGCAGGCCACCTTCACACTGGCCGATCCTATCAATCACC CTCTGTGGGTCCGATTCGAGGAAAGAAGCGGCAGCAACCTGAA CAAGTACAGAATCCTGACCGAGCAGCTGCACACCGAGAAGCTG AAGAAAAAGCTGACAGTGCAGCTGGACCGGCTGATCTACCCTAC AGAATCTGGCGGCTGGGAAGAGAAGGGCAAAGTGGACATTGTG CTGCTGCCCAGCCGGCAGTTCTACAACCAGATCTTCCTGGACAT CGAGGAAAAGGGCAAGCACGCCTTCACCTACAAGGATGAGAGC ATCAAGTTCCCTCTGAAGGGCACACTCGGCGGAGCCAGAGTGC AGTTCGACAGAGATCACCTGAGAAGATACCCTCACAAGGTGGAA AGCGGCAACGTGGGCAGAATCTACTTCAACATGACCGTGAACAT CGAGCCTACAGAGTCCCCAGTGTCCAAGTCTCTGAAGATCCACC GGGACGACTTCCCCAAGGTGGTCAACTTCAAGCCCAAAGAACT GACCGAGTGGATCAAGGACAGCAAGGGCAAGAAACTGAAGTCC GGCATCGAGTCCCTGGAAATCGGCCTGAGAGTGATGAGCATCG ACCTGGGACAGAGACAGGCCGCTGCCGCCTCTATTTTCGAGGT GGTGGATCAGAAGCCCGACATCGAAGGCAAGCTGTTTTTCCCAA TCAAGGGCACCGAGCTGTATGCCGTGCACAGAGCCAGCTTCAA CATCAAGCTGCCCGGCGAGACACTGGTCAAGAGCAGAGAAGTG CTGCGGAAGGCCAGAGAGGACAATCTGAAACTGATGAACCAGA AGCTCAACTTCCTGCGGAACGTGCTGCACTTCCAGCAGTTCGAG GACATCACCGAGAGAGAGAAGCGGGTCACCAAGTGGATCAGCA GACAAGAGAACAGCGACGTGCCCCTGGTGTACCAGGATGAGCT GATCCAGATCCGCGAGCTGATGTACAAGCCTTACAAGGACTGG GTCGCCTTCCTGAAGCAGCTCCACAAGAGACTGGAAGTCGAGA TCGGCAAAGAAGTGAAGCACTGGCGGAAGTCCCTGAGCGACGG AAGAAAGGGCCTGTACGGCATCTCCCTGAAGAACATCGACGAG ATCGATCGGACCCGGAAGTTCCTGCTGAGATGGTCCCTGAGGC CTACCGAACCTGGCGAAGTGCGTAGACTGGAACCCGGCCAGAG ATTCGCCATCGACCAGCTGAATCACCTGAACGCCCTGAAAGAAG ATCGGCTGAAGAAGATGGCCAACACCATCATCATGCACGCCCTG GGCTACTGCTACGACGTGCGGAAGAAGAAATGGCAGGCTAAGA ACCCCGCCTGCCAGATCATCCTGTTCGAGGATCTGAGCAACTAC AACCCCTACGAGGAAAGGTCCCGCTTCGAGAACAGCAAGCTCA TGAAGTGGTCCAGACGCGAGATCCCCAGACAGGTTGCACTGCA GGGCGAGATCTATGGCCTGCAAGTGGGAGAAGTGGGCGCTCAG TTCAGCAGCAGATTCCACGCCAAGACAGGCAGCCCTGGCATCA GATGTAGCGTCGTGACCAAAGAGAAGCTGCAGGACAATCGGTT CTTCAAGAATCTGCAGAGAGAGGGCAGACTGACCCTGGACAAA ATCGCCGTGCTGAAAGAGGGCGATCTGTACCCAGACAAAGGCG GCGAGAAGTTCATCAGCCTGAGCAAGGATCGGAAGTGCGTGAC CACACACGCCGACATCAACGCCGCTCAGAACCTGCAGAAGCGG TTCTGGACAAGAACCCACGGCTTCTACAAGGTGTACTGCAAGGC CTACCAGGTGGACGGCCAGACCGTGTACATCCCTGAGAGCAAG GACCAGAAGCAGAAGATCATCGAAGAGTTCGGCGAGGGCTACT TCATTCTGAAGGACGGGGTGTACGAATGGGTCAACGCCGGCAA GGGAGGCTCTGGAGGAAGCTCCGAAGTCGAGTTTTCCCATGAG TACTGGATGAGACACGCATTGACTCTCGCAAAGAGGGCTCGAG ATGAACGCGAGGTGCCCGTGGGGGCAGTACTCGTGCTCAACAA TCGCGTAATCGGCGAAGGTTGGAATAGGGCAATCGGACTCCAC GACCCCACTGCACATGCGGAAATCATGGCCCTTCGACAGGGAG GGCTTGTGATGCAGAATTATCGACTTTATGATGCGACGCTGTAC GTCACGTTTGAACCTTGCGTAATGTGCGCGGGAGCTATGATTCA CTCCCGCATTGGACGAGTTGTATTCGGTGTTCGCAACGCCAAGA CGGGTGCCGCAGGTTCACTGATGGACGTGCTGCATCATCCAGG CATGAACCACCGGGTAGAAATCACAGAAGGCATATTGGCGGAC GAATGTGCGGCGCTGTTGTGTCGTTTTTTTCGCATGCCCAGGCG

		GGTCTTTAACGCCCAGAAAAAAGCACAATCCTCTACTGACGGCT CTTCTGGATCTGAAACACCTGGCACAAGCGAGAGCGCCACCCC TGAGAGCTCTGGCCTGAAAATCAAGAAGGGCAGCTCCAAGCAG AGCAGCAGCGAGCTGGTGGATAGCGACATCCTGAAAGACAGCT TCGACCTGGCCTCCGAGCTGAAAGGCGAAAAGCTGATGCTGTA CAGGGACCCCAGCGGCAATGTGTTCCCCAGCGACAAATGGATG GCCGCTGGCGTGTTCTTCGGAAAGCTGGAACGCATCCTGATCA GCAAGCTGACCAACCAGTACTCCATCAGCACCATCGAGGACGA CAGCAGCAAGCAGTCTATGAAAAGGCCGGCGGCCACGAAAAAG GCCGGCCAGGCAAAAAAGAAAAAGGGATCCTACCCATACGATG TTCCAGATTACGCTTATCCCTACGACGTGCCTGATTATGCATACC CATATGATGTCCCCGACTATGCCTAA
K1019 폴리펩티드에서 BhCas12b GGSGGS-ABE8- Xten20	283	MAPKKKRKVGIHGVPAAATRSFILKIEPNEEVKKGLWKTHEVLNHGI AYYMNILKLIRQEAIYEHHEQDPKNPKKVSKAEIQAELWDFVLKMQK CNSFTHEVDKDEVFNILRELYEELVPSSVEKKGEANQLSNKFLYPLV DPNSQSGKGTASSGRKPRWYNLKIAGDPSWEEEKKKWEEDKKKD PLAKILGKLAEYGLIPLFIPYTDSNEPIVKEIKWMEKSRNQSVRRLDK DMFIQALERFLSWESWNLKVKEEYEKVEKEYKTLEERIKEDIQALKA LEQYEKERQEQLLRDTLNTNEYRLSKRGLRGWREIIQKWLKMDEN EPSEKYLEVFKDYQRKHPREAGDYSVYEFLSKKENHFIWRNHPEY PYLYATFCEIDKKKKDAKQQATFTLADPINHPLWVRFEERSGSNLN KYRILTEQLHTEKLKKKLTVQLDRLIYPTESGGWEEKGKVDIVLLPS RQFYNQIFLDIEEKGKHAFTYKDESIKFPLKGTLGGARVQFDRDHLR RYPHKVESGNVGRIYFNMTVNIEPTESPVSKSLKIHRDDFPKVVNFK PKELTEWIKDSKGKKLKSGIESLEIGLRVMSIDLGQRQAAAASIFEVV DQKPDIEGKLFFPIKGTELYAVHRASFNIKLPGETLVKSREVLRKAR EDNLKLMNQKLNFLRNVLHFQQFEDITEREKRVTKWISRQENSDVP LVYQDELIQIRELMYKPYKDWVAFLKQLHKRLEVEIGKEVKHWRKS LSDGRKGLYGISLKNIDEIDRTRKFLLRWSLRPTEPGEVRRLEPGQR FAIDQLNHLNALKEDRLKKMANTIIMHALGYCYDVRKKKWQAKNPA CQIILFEDLSNYNPYEERSRFENSKLMKWSRREIPRQVALQGEIYGL QVGEVGAQFSSRFHAKTGSPGIRCSVVTKEKLQDNRFFKNLQREG RLTLDKIAVLKEGDLYPDKGGEKFISLSKDRKCVTTHADINAAQNLQ KRFWTRTHGFYKVYCKAYQVDGQTVYIPESKDQKQKIIEEFGEGYF ILKDGVYEWVNAGKGGSGGSSEVEFSHEYWMRHALTLAKRARDE REVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVM QNYRLYDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAG SLMDVLHHPGMNHRVEITEGILADECAALLCRFFRMPRRVFNAQKK AQSSTDGSSGSETPGTSESATPESSGLKIKKGSSKQSSSELVDSDI LKDSFDLASELKGEKLMLYRDPSGNVFPSDKWMAAGVFFGKLERIL ISKLTNQYSISTIEDDSSKQSMKRPAATKKAGQAKKKKGSYPYDVP DYAYPYDVPDYAYPYDVPDYA
tr|A5H718|A5H718 _PETMA 시토신 데아미나제 OS=페트로미존 마리누스 OX=7757 PE=2 SV=1 아미노산 서열; PmCDA1 아미노산 서열	41	MTDAEYVRIHEKLDIYTFKKQFFNNKKSVSHRCYVLFELKRRGERR ACFWGYAVNKPQSGTERGIHAEIFSIRKVEEYLRDNPGQFTINWYS SWSPCADCAEKILEWYNQELRGNGHTLKIWACKLYYEKNARNQIG LWNLRDNGVGLNVMVSEHYQCCRKIFIQSSHNQLNENRWLEKTLK RAEKRRSELSIMIQVKILHTTKSPAV
EF094822.1 페트로미존 마리누스 단리물	42	TGACACGACACAGCCGTGTATATGAGGAAGGGTAGCTGGATGG GGGGGGGGGGAATACGTTCAGAGAGGACATTAGCGAGCGTCTT GTTGGTGGCCTTGAGTCTAGACACCTGCAGACATGACCGACGC TGAGTACGTGAGAATCCATGAGAAGTTGGACATCTACACGTTTA AGAAACAGTTTTTCAACAACAAAAAATCCGTGTCGCATAGATGCT

[1068]

PmCDA.21 시토신 데아미나제 mRNA, 완전한 cds; PmCDA1 아미노산 서열		ACGTTCTCTTTGAATTAAAACGACGGGGTGAACGTAGAGCGTGT TTTTGGGGCTATGCTGTGAATAAACCACAGAGCGGGACAGAACG TGGAATTCACGCCGAAATCTTTAGCATTAGAAAAGTCGAAGAATA CCTGCGCGACAACCCCGGACAATTCACGATAAATTGGTACTCAT CCTGGAGTCCTTGTGCAGATTGCGCTGAAAAGATCTTAGAATGG TATAACCAGGAGCTGCGGGGGAACGGCCACACTTTGAAAATCT GGGCTTGCAAACTCTATTACGAGAAAAATGCGAGGAATCAAATT GGGCTGTGGAACCTCAGAGATAACGGGGTTGGGTTGAATGTAA TGGTAAGTGAACACTACCAATGTTGCAGGAAAATATTCATCCAAT CGTCGCACAATCAATTGAATGAGAATAGATGGCTTGAGAAGACT TTGAAGCGAGCTGAAAAACGACGGAGCGAGTTGTCCATTATGAT TCAGGTAAAAATACTCCACACCACTAAGAGTCCTGCTGTTTAAGA GGCTATGCGGATGGTTTTC
tr|Q6QJ80|Q6QJ80_인간 활성화-유도된 시티딘 데아미나제 OS=호모 사피엔스 OX=9606 GN=AICDA PE=2 SV=1; AID 아미노산 서열	43	MDSLLMNRRKFLYQFKNVRWAKGRRETYLCYVVKRRDSATSFSLD FGYLRNKNGCHVELLFLRYISDWDLDPGRCYRVTWFTSWSPCYDC ARHVADFLRGNPNLSLRIFTARLYFCEDRKAEPEGLRRLHRAGVQI AIMTFKAPV
NG_011588.1:5001-15681 호모 사피엔스 활성화 유도된 시티딘 데아미나제 (AICDA), 염색체 12 상의 RefSeq유전자 (LRG_17); 인간 AID의 CDS의 핵산 서열	44	AGAGAACCATCATTAATTGAAGTGAGATTTTTCTGGCCTGAGACT TGCAGGGAGGCAAGAAGACACTCTGGACACCACTATGGACAGG TAAAGAGGCAGTCTTCTCGTGGGTGATTGCACTGGCCTTCCTCT CAGAGCAAATCTGAGTAATGAGACTGGTAGCTATCCCTTTCTCT CATGTAACTGTCTGACTGATAAGATCAGCTTGATCAATATGCATA TATATTTTTTGATCTGTCTCCTTTTCTTCTATTCAGATCTTATACG CTGTCAGCCCAATTCTTTCTGTTTCAGACTTCTCTTGATTTCCCT CTTTTTCATGTGGCAAAAGAAGTAGTGCGTACAATGTACTGATTC GTCCTGAGATTTGTACCATGGTTGAAACTAATTTATGGTAATAAT ATTAACATAGCAAATCTTTAGAGACTCAAATCATGAAAAGGTAAT AGCAGTACTGTACTAAAAACGGTAGTGCTAATTTTCGTAATAATT TTGTAAATATTCAACAGTAAAACAACTTGAAGACACACTTTCCTA GGGAGGCGTTACTGAAATAATTTAGCTATAGTAAGAAAATTTGTA ATTTTAGAAATGCCAAGCATTCTAAATTAATTGCTTGAAAGTCACT ATGATTGTGTCCATTATAAGGAGACAAATTCATTCAAGCAAGTTA TTTAATGTTAAAGGCCCAATTGTTAGGCAGTTAATGGCACTTTTA CTATTAACTAATCTTTCCATTTGTTCAGACGTAGCTTAACTTACCT CTTAGGTGTGAATTTGGTTAAGGTCCTCATAATGTCTTTATGTGC AGTTTTTGATAGGTTATTGTCATAGAACTTATTCTATTCCTACATT TATGATTACTATGGATGTATGAGAATAACACCTAATCCTTATACTT TACCTCAATTTAACTCCTTTATAAAGAACTTACATTACAGAATAAA GATTTTTTAAAAATATATTTTTTTGTAGAGACAGGGTCTTAGCCCA GCCGAGGCTGGTCTCTAAGTCCTGGCCCAAGCGATCCTCCTGC CTGGGCCTCCTAAAGTGCTGGAATTATAGACATGAGCCATCACA TCCAATATACAGAATAAAGATTTTTAATGGAGGATTTAATGTTCTT CAGAAAATTTTCTTGAGGTCAGACAATGTCAAATGTCTCCTCAGT TTACACTGAGATTTTGAAAACAAGTCTGAGCTATAGGTCCTTGTG AAGGGTCCATTGGAAATACTTGTTCAAAGTAAAATGGAAAGCAAA GGTAAAATCAGCAGTTGAAATTCAGAGAAAGACAGAAAAGGAGA AAAGATGAAATTCAACAGGACAGAAGGGAAATATATTATCATTAA GGAGGACAGTATCTGTAGAGCTCATTAGTGATGGCAAAATGACT TGGTCAGGATTATTTTTAACCCGCTTGTTTCTGGTTTGCACGGCT GGGGATGCAGCTAGGGTTCTGCCTCAGGGAGCACAGCTGTCCA GAGCAGCTGTCAGCCTGCAAGCCTGAAACACTCCCTCGGTAAA GTCCTTCCTACTCAGGACAGAAATGACGAGAACAGGGAGCTGG

[1069]

		AAACAGGCCCCTAACCAGAGAAGGGAAGTAATGGATCAACAAAG TTAACTAGCAGGTCAGGATCACGCAATTCATTTCACTCTGACTG GTAACATGTGACAGAAACAGTGTAGGCTTATTGTATTTTCATGTA GAGTAGGACCCAAAAATCCACCCAAAGTCCTTTATCTATGCCAC ATCCTTCTTATCTATACTTCCAGGACACTTTTTCTTCCTTATGATA AGGCTCTCTCTCTCTCCACACACACACACACACACACACACACA CACACACACACACACACAAACACACACCCCGCCAACCAAGGTG CATGTAAAAAGATGTAGATTCCTCTGCCTTTCTCATCTACACAGC CCAGGAGGGTAAGTTAATATAAGAGGGATTTATTGGTAAGAGAT GATGCTTAATCTGTTTAACACTGGGCCTCAAAGAGAGAATTTCTT TTCTTCTGTACTTATTAAGCACCTATTATGTGTTGAGCTTATATAT ACAAAGGGTTATTATATGCTAATATAGTAATAGTAATGGTGGTTG GTACTATGGTAATTACCATAAAAATTATTATCCTTTTAAAATAAAG CTAATTATTATTGGATCTTTTTTAGTATTCATTTTATGTTTTTTATG TTTTTGATTTTTTAAAAGACAATCTCACCCTGTTACCCAGGCTGG AGTGCAGTGGTGCAATCATAGCTTTCTGCAGTCTTGAACTCCTG GGCTCAAGCAATCCTCCTGCCTTGGCCTCCCAAAGTGTTGGGAT ACAGTCATGAGCCACTGCATCTGGCCTAGGATCCATTTAGATTA AAATATGCATTTTAAATTTTAAAATAATATGGCTAATTTTTACCTTA TGTAATGTGTATACTGGCAATAAATCTAGTTTGCTGCCTAAAGTT TAAAGTGCTTTCCAGTAAGCTTCATGTACGTGAGGGGAGACATT TAAAGTGAAACAGACAGCCAGGTGTGGTGGCTCACGCCTGTAAT CCCAGCACTCTGGGAGGCTGAGGTGGGTGGATCGCTTGAGCCC TGGAGTTCAAGACCAGCCTGAGCAACATGGCAAAACGCTGTTTC TATAACAAAAATTAGCCGGGCATGGTGGCATGTGCCTGTGGTCC CAGCTACTAGGGGGCTGAGGCAGGAGAATCGTTGGAGCCCAGG AGGTCAAGGCTGCACTGAGCAGTGCTTGCGCCACTGCACTCCA GCCTGGGTGACAGGACCAGACCTTGCCTCAAAAAAATAAGAAGA AAAATTAAAAATAAATGGAAACAACTACAAAGAGCTGTTGTCCTA GATGAGCTACTTAGTTAGGCTGATATTTTGGTATTTAACTTTTAAA GTCAGGGTCTGTCACCTGCACTACATTATTAAAATATCAATTCTC AATGTATATCCACACAAAGACTGGTACGTGAATGTTCATAGTACC TTTATTCACAAAACCCCAAAGTAGAGACTATCCAAATATCCATCA ACAAGTGAACAAATAAACAAAATGTGCTATATCCATGCAATGGAA TACCACCCTGCAGTACAAAGAAGCTACTTGGGGATGAATCCCAA AGTCATGACGCTAAATGAAAGAGTCAGACATGAAGGAGGAGATA ATGTATGCCATACGAAATTCTAGAAAATGAAAGTAACTTATAGTT ACAGAAAGCAAATCAGGGCAGGCATAGAGGCTCACACCTGTAAT CCCAGCACTTTGAGAGGCCACGTGGGAAGATTGCTAGAACTCA GGAGTTCAAGACCAGCCTGGGCAACACAGTGAAACTCCATTCTC CACAAAAATGGGAAAAAAAGAAAGCAAATCAGTGGTTGTCCTGT GGGGAGGGGAAGGACTGCAAAGAGGGAAGAAGCTCTGGTGGG GTGAGGGTGGTGATTCAGGTTCTGTATCCTGACTGTGGTAGCAG TTTGGGGTGTTTACATCCAAAAATATTCGTAGAATTATGCATCTT AAATGGGTGGAGTTTACTGTATGTAAATTATACCTCAATGTAAGA AAAAATAATGTGTAAGAAAACTTTCAATTCTCTTGCCAGCAAACG TTATTCAAATTCCTGAGCCCTTTACTTCGCAAATTCTCTGCACTT CTGCCCCGTACCATTAGGTGACAGCACTAGCTCCACAAATTGGA TAAATGCATTTCTGGAAAAGACTAGGGACAAAATCCAGGCATCA CTTGTGCTTTCATATCAACCATGCTGTACAGCTTGTGTTGCTGTC TGCAGCTGCAATGGGGACTCTTGATTTCTTTAAGGAAACTTGGG TTACCAGAGTATTTCCACAAATGCTATTCAAATTAGTGCTTATGAT ATGCAAGACACTGTGCTAGGAGCCAGAAAACAAAGAGGAGGAG AAATCAGTCATTATGTGGGAACAACATAGCAAGATATTTAGATCA TTTTGACTAGTTAAAAAAGCAGCAGAGTACAAAATCACACATGCA ATCAGTATAATCCAAATCATGTAAATATGTGCCTGTAGAAAGACT AGAGGAATAAACACAAGAATCTTAACAGTCATTGTCATTAGACAC TAAGTCTAATTATTATTATTAGACACTATGATATTTGAGATTTAAA AAATCTTTAATATTTTAAAATTTAGAGCTCTTCTATTTTTCCATAGT

[1070]

[1071]

		ATTCAAGTTTGACAATGATCAAGTATTACTCTTTCTTTTTTTTTTTT TTTTTTTTTTTTGAGATGGAGTTTTGGTCTTGTTGCCCATGCTG GAGTGGAATGGCATGACCATAGCTCACTGCAACCTCCACCTCCT GGGTTCAAGCAAAGCTGTCGCCTCAGCCTCCCGGGTAGATGGG ATTACAGGCGCCCACCACCACACTCGGCTAATGTTTGTATTTTTA GTAGAGATGGGGTTTCACCATGTTGGCCAGGCTGGTCTCAAACT CCTGACCTCAGAGGATCCACCTGCCTCAGCCTCCCAAAGTGCT GGGATTACAGATGTAGGCCACTGCGCCCGGCCAAGTATTGCTC TTATACATTAAAAAACAGGTGTGAGCCACTGCGCCCAGCCAGGT ATTGCTCTTATACATTAAAAAATAGGCCGGTGCAGTGGCTCACG CCTGTAATCCCAGCACTTTGGGAAGCCAAGGCGGGCAGAACAC CCGAGGTCAGGAGTCCAAGGCCAGCCTGGCCAAGATGGTGAAA CCCCGTCTCTATTAAAAATACAAACATTACCTGGGCATGATGGTG GGCGCCTGTAATCCCAGCTACTCAGGAGGCTGAGGCAGGAGGA TCCGCGGAGCCTGGCAGATCTGCCTGAGCCTGGGAGGTTGAGG CTACAGTAAGCCAAGATCATGCCAGTATACTTCAGCCTGGGCGA CAAAGTGAGACCGTAACAAAAAAAAAAAAATTTAAAAAAAGAAAT TTAGATCAAGATCCAACTGTAAAAAGTGGCCTAAACACCACATTA AAGAGTTTGGAGTTTATTCTGCAGGCAGAAGAGAACCATCAGGG GGTCTTCAGCATGGGAATGGCATGGTGCACCTGGTTTTTGTGAG ATCATGGTGGTGACAGTGTGGGGAATGTTATTTTGGAGGGACTG GAGGCAGACAGACCGGTTAAAAGGCCAGCACAACAGATAAGGA GGAAGAAGATGAGGGCTTGGACCGAAGCAGAGAAGAGCAAACA GGGAAGGTACAAATTCAAGAAATATTGGGGGGTTTGAATCAACA CATTTAGATGATTAATTAAATATGAGGACTGAGGAATAAGAAATG AGTCAAGGATGGTTCCAGGCTGCTAGGCTGCTTACCTGAGGTG GCAAAGTCGGGAGGAGTGGCAGTTTAGGACAGGGGGCAGTTGA GGAATATTGTTTTGATCATTTTGAGTTTGAGGTACAAGTTGGACA CTTAGGTAAAGACTGGAGGGGAAATCTGAATATACAATTATGGG ACTGAGGAACAAGTTTATTTTATTTTTTGTTTCGTTTTCTTGTTGA AGAACAAATTTAATTGTAATCCCAAGTCATCAGCATCTAGAAGAC AGTGGCAGGAGGTGACTGTCTTGTGGGTAAGGGTTTGGGGTCC TTGATGAGTATCTCTCAATTGGCCTTAAATATAAGCAGGAAAAGG AGTTTATGATGGATTCCAGGCTCAGCAGGGCTCAGGAGGGCTC AGGCAGCCAGCAGAGGAAGTCAGAGCATCTTCTTTGGTTTAGCC CAAGTAATGACTTCCTTAAAAAGCTGAAGGAAAATCCAGAGTGA CCAGATTATAAACTGTACTCTTGCATTTTCTCTCCCTCCTCTCAC CCACAGCCTCTTGATGAACCGGAGGAAGTTTCTTTACCAATTCA AAAATGTCCGCTGGGCTAAGGGTCGGCGTGAGACCTACCTGTG CTACGTAGTGAAGAGGCGTGACAGTGCTACATCCTTTTCACTGG ACTTTGGTTATCTTCGCAATAAGGTATCAATTAAAGTCGGCTTTG CAAGCAGTTTAATGGTCAACTGTGAGTGCTTTTAGAGCCACCTG CTGATGGTATTACTTCCATCCTTTTTTGGCATTTGTGTCTCTATCA CATTCCTCAAATCCTTTTTTTTATTTCTTTTTCCATGTCCATGCAC CCATATTAGACATGGCCCAAAATATGTGATTTAATTCCTCCCCAG TAATGCTGGGCACCCTAATACCACTCCTTCCTTCAGTGCCAAGA ACAACTGCTCCCAAACTGTTTACCAGCTTTCCTCAGCATCTGAAT TGCCTTTGAGATTAATTAAGCTAAAAGCATTTTTATATGGGAGAA TATTATCAGCTTGTCCAAGCAAAAATTTTAAATGTGAAAAACAAAT TGTGTCTTAAGCATTTTTGAAAATTAAGGAAGAAGAATTTGGGAA AAAATTAACGGTGGCTCAATTCTGTCTTCCAAATGATTTCTTTTC CCTCCTACTCACATGGGTCGTAGGCCAGTGAATACATTCAACAT GGTGATCCCCAGAAAACTCAGAGAAGCCTCGGCTGATGATTAAT TAAATTGATCTTTCGGCTACCCGAGAGAATTACATTTCCAAGAGA CTTCTTCACCAAAATCCAGATGGGTTTACATAAACTTCTGCCCAC GGGTATCTCCTCTCTCCTAACACGCTGTGACGTCTGGGCTTGGT GGAATCTCAGGGAAGCATCCGTGGGGTGGAAGGTCATCGTCTG GCTCGTTGTTTGATGGTTATATTACCATGCAATTTTCTTTGCCTA CATTTGTATTGAATACATCCCAATCTCCTTCCTATTCGGTGACAT

[1072]

		GACACATTCTATTTCAGAAGGCTTTGATTTTATCAAGCACTTTCAT TTACTTCTCATGGCAGTGCCTATTACTTCTCTTACAATACCCATC TGTCTGCTTTACCAAAATCTATTTCCCCTTTTCAGATCCTCCCAA ATGGTCCTCATAAACTGTCCTGCCTCCACCTAGTGGTCCAGGTA TATTTCCACAATGTTACATCAACAGGCACTTCTAGCCATTTTCCT TCTCAAAAGGTGCAAAAAGCAACTTCATAAACACAAATTAAATCT TCGGTGAGGTAGTGTGATGCTGCTTCCTCCCAACTCAGCGCACT TCGTCTTCCTCATTCCACAAAAACCCATAGCCTTCCTTCACTCTG CAGGACTAGTGCTGCCAAGGGTTCAGCTCTACCTACTGGTGTGC TCTTTTGAGCAAGTTGCTTAGCCTCTCTGTAACACAAGGACAATA GCTGCAAGCATCCCCAAAGATCATTGCAGGAGACAATGACTAAG GCTACCAGAGCCGCAATAAAAGTCAGTGAATTTTAGCGTGGTCC TCTCTGTCTCTCCAGAACGGCTGCCACGTGGAATTGCTCTTCCT CCGCTACATCTCGGACTGGGACCTAGACCCTGGCCGCTGCTAC CGCGTCACCTGGTTCACCTCCTGGAGCCCCTGCTACGACTGTG CCCGACATGTGGCCGACTTTCTGCGAGGGAACCCCAACCTCAG TCTGAGGATCTTCACCGCGCGCCTCTACTTCTGTGAGGACCGCA AGGCTGAGCCCGAGGGGCTGCGGCGGCTGCACCGCGCCGGG GTGCAAATAGCCATCATGACCTTCAAAGGTGCGAAAGGGCCTTC CGCGCAGGCGCAGTGCAGCAGCCCGCATTCGGGATTGCGATG CGGAATGAATGAGTTAGTGGGGAAGCTCGAGGGGAAGAAGTGG GCGGGGATTCTGGTTCACCTCTGGAGCCGAAATTAAAGATTAGA AGCAGAGAAAAGAGTGAATGGCTCAGAGACAAGGCCCCGAGGA AATGAGAAAATGGGGCCAGGGTTGCTTCTTTCCCCTCGATTTGG AACCTGAACTGTCTTCTACCCCCATATCCCCGCCTTTTTTTCCTT TTTTTTTTTTTGAAGATTATTTTTACTGCTGGAATACTTTTGTAGAA AACCACGAAAGAACTTTCAAAGCCTGGGAAGGGCTGCATGAAAA TTCAGTTCGTCTCTCCAGACAGCTTCGGCGCATCCTTTTGGTAA GGGGCTTCCTCGCTTTTTAAATTTTCTTTCTTTCTCTACAGTCTTT TTTGGAGTTTCGTATATTTCTTATATTTTCTTATTGTTCAATCACTC TCAGTTTTCATCTGATGAAAACTTTATTTCTCCTCCACATCAGCTT TTTCTTCTGCTGTTTCACCATTCAGAGCCCTCTGCTAAGGTTCCT TTTCCCTCCCTTTTCTTTCTTTTGTTGTTTCACATCTTTAAATTTCT GTCTCTCCCCAGGGTTGCGTTTCCTTCCTGGTCAGAATTCTTTTC TCCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAAACAAACAAACAAAAAA CCCAAAAAAACTCTTTCCCAATTTACTTTCTTCCAACATGTTACAA AGCCATCCACTCAGTTTAGAAGACTCTCCGGCCCCACCGACCCC CAACCTCGTTTTGAAGCCATTCACTCAATTTGCTTCTCTCTTTCT CTACAGCCCCTGTATGAGGTTGATGACTTACGAGACGCATTTCG TACTTTGGGACTTTGATAGCAACTTCCAGGAATGTCACACACGAT GAAATATCTCTGCTGAAGACAGTGGATAAAAAACAGTCCTTCAA GTCTTCTCTGTTTTTATTCTTCAACTCTCACTTTCTTAGAGTTTAC AGAAAAAATATTTATATACGACTCTTTAAAAAGATCTATGTCTTGA AAATAGAGAAGGAACACAGGTCTGGCCAGGGACGTGCTGCAAT TGGTGCAGTTTTGAATGCAACATTGTCCCCTACTGGGAATAACA GAACTGCAGGACCTGGGAGCATCCTAAAGTGTCAACGTTTTTCT ATGACTTTTAGGTAGGATGAGAGCAGAAGGTAGATCCTAAAAAG CATGGTGAGAGGATCAAATGTTTTTATATCAACATCCTTTATTATT TGATTCATTTGAGTTAACAGTGGTGTTAGTGATAGATTTTTCTATT CTTTTCCCTTGACGTTTACTTTCAAGTAACACAAACTCTTCCATCA GGCCATGATCTATAGGACCTCCTAATGAGAGTATCTGGGTGATT GTGACCCCAAACCATCTCTCCAAAGCATTAATATCCAATCATGCG CTGTATGTTTTAATCAGCAGAAGCATGTTTTTATGTTTGTACAAAA GAAGATTGTTATGGGTGGGGATGGAGGTATAGACCATGCATGGT CACCTTCAAGCTACTTTAATAAAGGATCTTAAAATGGGCAGGAG GACTGTGAACAAGACACCCTAATAATGGGTTGATGTCTGAAGTA GCAAATCTTCTGGAAACGCAAACTCTTTTAAGGAAGTCCCTAATT TAGAAACACCCACAAACTTCACATATCATAATTAGCAAACAATTG GAAGGAAGTTGCTTGAATGTTGGGGAGAGGAAAATCTATTGGCT

		CTCGTGGGTCTCTTCATCTCAGAAATGCCAATCAGGTCAAGGTT TGCTACATTTTGTATGTGTGTGATGCTTCTCCCAAAGGTATATTA ACTATATAAGAGAGTTGTGACAAAACAGAATGATAAAGCTGCGA ACCGTGGCACACGCTCATAGTTCTAGCTGCTTGGGAGGTTGAG GAGGGAGGATGGCTTGAACACAGGTGTTCAAGGCCAGCCTGGG CAACATAACAAGATCCTGTCTCTCAAAAAAAAAAAAAAAAAAAAG AAAGAGAGAGGGCCGGGCGTGGTGGCTCACGCCTGTAATCCCA GCACTTTGGGAGGCCGAGCCGGGCGGATCACCTGTGGTCAGG AGTTTGAGACCAGCCTGGCCAACATGGCAAAACCCCGTCTGTAC TCAAAATGCAAAAATTAGCCAGGCGTGGTAGCAGGCACCTGTAA TCCCAGCTACTTGGGAGGCTGAGGCAGGAGAATCGCTTGAACC CAGGAGGTGGAGGTTGCAGTAAGCTGAGATCGTGCCGTTGCAC TCCAGCCTGGGCGACAAGAGCAAGACTCTGTCTCAGAAAAAAAA AAAAAAAAGAGAGAGAGAGAAAGAGAACAATATTTGGGAGAG AAGGATGGGGAAGCATTGCAAGGAAATTGTGCTTTATCCAACAA AATGTAAGGAGCCAATAAGGGATCCCTATTTGTCTCTTTTGGTGT CTATTTGTCCCTAACAACTGTCTTTGACAGTGAGAAAAATATTCA GAATAACCATATCCCTGTGCCGTTATTACCTAGCAACCCTTGCAA TGAAGATGAGCAGATCCACAGGAAAACTTGAATGCACAACTGTC TTATTTTAATCTTATTGTACATAAGTTTGTAAAAGAGTTAAAAATT GTTACTTCATGTATTCATTTATATTTTATATTATTTTGCGTCTAATG ATTTTTTATTAACATGATTTCCTTTTCTGATATATTGAAATGGAGT CTCAAAGCTTCATAAATTTATAACTTTAGAAATGATTCTAATAACA ACGTATGTAATTGTAACATTGCAGTAATGGTGCTACGAAGCCATT TCTCTTGATTTTTAGTAAACTTTTATGACAGCAAATTTGCTTCTGG CTCACTTTCAATCAGTTAAATAAATGATAAATAATTTTGGAAGCTG TGAAGATAAAATACCAAATAAAATAATATAAAAGTGATTTATATGA AGTTAAAATAAAAAATCAGTATGATGGAATAAACTTG
개 AID (clAID) 폴리펩티드 서열	1374	MDSLLMKQRKFLYHFKNVRWAKGRHETYLCYVVKRRDSATSFSLD FGHLRNKSGCHVELLFLRYISDWDLDPGRCYRVTWFTSWSPCYDC ARHVADFLRGYPNLSLRIFAARLYFCEDRKAEPEGLRRLHRAGVQI AIMTFKDYFYCWNTFVENREKTFKAWEGLHENSVRLSRQLRRILLP LYEVDDLRDAFRTLGL
소 AID (btAID) 폴리펩티드 서열	1375	MDSLLKKQRQFLYQFKNVRWAKGRHETYLCYVVKRRDSPTSFSLD FGHLRNKAGCHVELLFLRYISDWDLDPGRCYRVTWFTSWSPCYDC ARHVADFLRGYPNLSLRIFTARLYFCDKERKAEPEGLRRLHRAGVQ IAIMTFKDYFYCWNTFVENHERTFKAWEGLHENSVRLSRQLRRILLP LYEVDDLRDAFRTLGL
랫트 AID 폴리펩티드 서열	1376	MAVGSKPKAALVGPHWERERIWCFLCSTGLGTQQTGQTSRWLRP AATQDPVSPPRSLLMKQRKFLYHFKNVRWAKGRHETYLCYVVKRR DSATSFSLDFGYLRNKSGCHVELLFLRYISDWDLDPGRCYRVTWFT SWSPCYDCARHVADFLRGNPNLSLRIFTARLTGWGALPAGLMSPA RPSDYFYCWNTFVENHERTFKAWEGLHENSVRLSRRLRRILLPLYE VDDLRDAFRTLGL
마우스 (mAID) AID 폴리펩티드 서열	1377	MDSLLMNRRKFLYQFKNVRWAKGRRETYLCYVVKRRDSATSFSLD FGYLRNKNGCHVELLFLRYISDWDLDPGRCYRVTWFTSWSPCYDC ARHVADFLRGNPNLSLRIFTARLYFCEDRKAEPEGLRRLHRAGVQI AIMTFKDYFYCWNTFVENHERTFKAWEGLHENSVRLSRQLRRILLP LYEVDDLRDAFRTLGL
rAPOBEC-1 폴리펩티드 서열	1378	MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELRKETCLLYEINWGG RHSIWRHTSQNTNKHVEVNFIEKFTTERYFCPNTRCSITWFLSWSP CGECSRAITEFLSRYPHVTLFIYIARLYHHADPRNRQGLRDLISSGVT IQIMTEQESGYCWRNFVNYSPSNEAHWPRYPHLWVRLYVLELYCII LGLPPCLNILRRKQPQLTFFTIALQSCHYQRLPPHILWATGLK
maAPOBEC-1 폴리펩티드 서열	1379	MSSETGPVVVDPTLRRRIEPHEFDAFFDQGELRKETCLLYEIRWGG RHNIWRHTGQNTSRHVEINFIEKFTSERYFYPSTRCSIVWFLSWSP CGECSKAITEFLSGHPNVTLFIYAARLYHHTDQRNRQGLRDLISRGV TIRIMTEQEYCYCWRNFVNYPPSNEVYWPRYPNLWMRLYALELYCI HLGLPPCLKIKRRHQYPLTFFRLNLQSCHYQRIPPHILWATGFI

[1073]

ppAPOBEC-1 폴리펩티드 서열	1380	MTSEKGPSTGDPTLRRRIESWEFDVFYDPRELRKETCLLYEIKWGM SRKIWRSSGKNTTNHVEVNFIKKFTSERRFHSSISCSITWFLSWSPC WECSQAIREFLSQHPGVTLVIYVARLFWHMDQRNRQGLRDLVNSG VTIQIMRASEYYHCWRNFVNYPPGDEAHWPQYPPLWMMLYALELH CIILSLPPCLKISRRWQNHLAFFRLHLQNCHYQTIPPHILLATGLIHPS VTWR
ocAPOBEC1 폴리펩티드 서열	1381	MASEKGPSNKDYTLRRRIEPWEFEVFFDPQELRKEACLLYEIKWGA SSKTWRSSGKNTTNHVEVNFLEKLTSEGRLGPSTCCSITWFLSWS PCWECSMAIREFLSQHPGVTLIIFVARLFQHMDRRNRQGLKDLVTS GVTVRVMSVSEYCYCWENFVNYPPGKAAQWPRYPPRWMLMYAL ELYCIILGLPPCLKISRRHQKQLTFFSLTPQYCHYKMIPPYILLATGLL QPSVPWR
mdAPOBEC-1 폴리펩티드 서열	1382	MNSKTGPSVGDATLRRRIKPWEFVAFFNPQELRKETCLLYEIKWGN QNIWRHSNQNTSQHAEINFMEKFTAERHFNSSVRCSITWFLSWSP CWECSKAIRKFLDHYPNVTLAIFISRLYWHMDQQHRQGLKELVHSG VTIQIMSYSEYHYCWRNFVDYPQGEEDYWPKYPYLWIMLYVLELH CIILGLPPCLKISGSHSNQLALFSLDLQDCHYQKIPYNVLVATGLVQP FVTWR
ppAPOBEC-2 폴리펩티드 서열	1383	MAQKEEAAAATEAASQNGEDLENLDDPEKLKELIELPPFEIVTGERL PANFFKFQFRNVEYSSGRNKTFLCYVVEAQGKGGQVQASRGYLE DEHAAAHAEEAFFNTILPAFDPALRYNVTWYVSSSPCAACADRIIKT LSKTKNLRLLILVGRLFMWEELEIQDALKKLKEAGCKLRIMKPQDFE YVWQNFVEQEEGESKAFQPWEDIQENFLYYEEKLADILK
btAPOBEC-2 폴리펩티드 서열	1384	MAQKEEAAAAAEPASQNGEEVENLEDPEKLKELIELPPFEIVTGERL PAHYFKFQFRNVEYSSGRNKTFLCYVVEAQSKGGQVQASRGYLED EHATNHAEEAFFNSIMPTFDPALRYMVTWYVSSSPCAACADRIVKT LNKTKNLRLLILVGRLFMWEEPEIQAALRKLKEAGCRLRIMKPQDFE YIWQNFVEQEEGESKAFEPWEDIQENFLYYEEKLADILK
mAPOBEC-3-(1) 폴리펩티드 서열	1385	MQPQRLGPRAGMGPFCLGCSHRKCYSPIRNLISQETFKFHFKNLG YAKGRKDTFLCYEVTRKDCDSPVSLHHGVFKNKDNIHAEICFLYWF HDKVLKVLSPREEFKITWYMSWSPCFECAEQIVRFLATHHNLSLDIF SSRLYNVQDPETQQNLCRLVQEGAQVAAMDLYEFKKCWKKFVDN GGRRFRPWKRLLTNFRYQDSKLQEILRPCYISVPSSSSSTLSNICLT KGLPETRFWVEGRRMDPLSEEEFYSQFYNQRVKHLCYYHRMKPY LCYQLEQFNGQAPLKGCLLSEKGKQHAEILFLDKIRSMELSQVTITC YLTWSPCPNCAWQLAAFKRDRPDLILHIYTSRLYFHWKRPFQKGLC SLWQSGILVDVMDLPQFTDCWTNFVNPKRPFWPWKGLEIISRRTQ RRLRRIKESWGLQDLVNDFGNLQLGPPMS
APOBEC-3-(2) (마우스 APOBEC-3) 폴리펩티드 서열	1386	MGPFCLGCSHRKCYSPIRNLISQETFKFHFKNLGYAKGRKDTFLCY EVTRKDCDSPVSLHHGVFKNKDNIHAEICFLYWFHDKVLKVLSPRE EFKITWYMSWSPCFECAEQIVRFLATHHNLSLDIFSSRLYNVQDPET QQNLCRLVQEGAQVAAMDLYEFKKCWKKFVDNGGRRFRPWKRLL TNFRYQDSKLQEILRPCYIPVPSSSSSTLSNICLTKGLPETRFCVEG RRMDPLSEEEFYSQFYNQRVKHLCYYHRMKPYLCYQLEQFNGQA PLKGCLLSEKGKQHAEILFLDKIRSMELSQVTITCYLTWSPCPNCAW QLAAFKRDRPDLILHIYTSRLYFHWKRPFQKGLCSLWQSGILVDVM DLPQFTDCWTNFVNPKRPFWPWKGLEIISRRTQRRLRRIKESWGL QDLVNDFGNLQLGPPMS
APOBEC-3 폴리펩티드 서열	1387	MGPFCLGCSHRKCYSPIRNLISQETFKFHFKNRLRYAIDRKDTFLCY EVTRKDCDSPVSLHHGVFKNKDNIHAEICFLYWFHDKVLKVLSPRE EFKITWYMSWSPCFECAEQVLRFLATHHNLSLDIFSSRLYNIRDPEN QQNLCRLVQEGAQVAAMDLYEFKKCWKKFVDNGGRRFRPWKKLL TNFRYQDSKLQEILRPCYIPVPSSSSSTLSNICLTKGLPETRFCVER RRVHLLSEEEFYSQFYNQRVKHLCYYHGVKPYLCYQLEQFNGQAP LKGCLLSEKGKQHAEILFLDKIRSMELSQVIITCYLTWSPCPNCAWQ LAAFKRDRPDLILHIYTSRLYFHWKRPFQKGLCSLWQSGILVDVMDL PQFTDCWTNFVNPKRPFWPWKGLEIISRRTQRRLHRIKESWGLQD LVNDFGNLQLGPPMS

[1074]

hAPOBEC-3A 폴리펩티드 서열	1388	MEASPASGPRHLMDPHIFTSNFNNGIGRHKTYLCYEVERLDNGTSV KMDQHRGFLHNQAKNLLCGFYGRHAELRFLDLVPSLQLDPAQIYR VTWFISWSPCFSWGCAGEVRAFLQENTHVRLRIFAARIYDYDPLYK EALQMLRDAGAQVSIMTYDEFKHCWDTFVDHQGCPFQPWDGLDE HSQALSGRLRAILQNQGN
hAPOBEC-3F 폴리펩티드 서열	1389	MKPHFRNTVERMYRDTFSYNFYNRPILSRRNTVWLCYEVKTKGPS RPRLDAKIFRGQVYSQPEHHAEMCFLSWFCGNQLPAYKCFQITWF VSWTPCPDCVAKLAEFLAEHPNVTLTISAARLYYYWERDYRRALCR LSQAGARVKIMDDEEFAYCWENFVYSEGQPFMPWYKFDDNYAFL HRTLKEILRNPMEAMYPHIFYFHFKNLRKAYGRNESWLCFTMEVVK HHSPVSWKRGVFRNQVDPETHCHAERCFLSWFCDDILSPNTNYEV TWYTSWSPCPECAGEVAEFLARHSNVNLTIFTARLYYFWDTDYQE GLRSLSQEGASVEIMGYKDFKYCWENFVYNDDEPFKPWKGLKYNF LFLDSKLQEILE
붉은털 원숭이 APOBEC-3G 폴리펩티드 서열	1390	MVEPMDPRTFVSNFNNRPILSGLNTVWLCCEVKTKDPSGPPLDAKI FQGKVYSKAKYHPEMRFLRWFHKWRQLHHDQEYKVTWYVSWSP CTRCANSVATFLAKDPKVTLTIFVARLYYFWKPDYQQALRILCQKR GGPHATMKIMNYNEFQDCWNKFVDGRGKPFKPRNNLPKHYTLLQ ATLGELLRHLMDPGTFTSNFNNKPWVSGQHETYLCYKVERLHNDT WVPLNQHRGFLRNQAPNIHGFPKGRHAELCFLDLIPFWKLDGQQY RVTCFTSWSPCFSCAQEMAKFISNNEHVSLCIFAARIYDDQGRYQE GLRALHRDGAKIAMMNYSEFEYCWDTFVDRQGRPFQPWDGLDEH SQALSGRLRAI
침팬지 APOBEC-3G 폴리펩티드 서열	1391	MKPHFRNPVERMYQDTFSDNFYNRPILSHRNTVWLCYEVKTKGPS RPPLDAKIFRGQVYSKLKYHPEMRFFHWFSKWRKLHRDQEYEVT WYISWSPCTKCTRDVATFLAEDPKVTLTIFVARLYYFWDPDYQEAL RSLCQKRDGPRATMKIMNYDEFQHCWSKFVYSQRELFEPWNNLP KYYILLHIMLGEILRHSMDPPTFTSNFNNELWVRGRHETYLCYEVER LHNDTWVLLNQRRGFLCNQAPHKHGFLEGRHAELCFLDVIPFWKL DLHQDYRVTCFTSWSPCFSCAQEMAKFISNNKHVSLCIFAARIYDD QGRCQEGLRTLAKAGAKISIMTYSEFKHCWDTFVDHQGCPFQPWD GLEEHSQALSGRLRAILQNQGN
사바나 원숭이 APOBEC-3G 폴리펩티드 서열	1392	MNPQIRNMVEQMEPDIFVYYFNNRPILSGRNTVWLCYEVKTKDPS GPPLDANIFQGKLYPEAKDHPEMKFLHWFRKWRQLHRDQEYEVT WYVSWSPCTRCANSVATFLAEDPKVTLTIFVARLYYFWKPDYQQA LRILCQERGGPHATMKIMNYNEFQHCWNEFVDGQGKPFKPRKNLP KHYTLLHATLGELLRHVMDPGTFTSNFNNKPWVSGQRETYLCYKV ERSHNDTWVLLNQHRGFLRNQAPDRHGFPKGRHAELCFLDLIPFW KLDDQQYRVTCFTSWSPCFSCAQKMAKFISNNKHVSLCIFAARIYD DQGRCQEGLRTLHRDGAKIAVMNYSEFEYCWDTFVDRQGRPFQP WDGLDEHSQALSGRLRAI
인간 APOBEC-3G 폴리펩티드 서열	1393	MKPHFRNTVERMYRDTFSYNFYNRPILSRRNTVWLCYEVKTKGPS RPPLDAKIFRGQVYSELKYHPEMRFFHWFSKWRKLHRDQEYEVT WYISWSPCTKCTRDMATFLAEDPKVTLTIFVARLYYFWDPDYQEAL RSLCQKRDGPRATMKIMNYDEFQHCWSKFVYSQRELFEPWNNLP KYYILLHIMLGEILRHSMDPPTFTFNFNNEPWVRGRHETYLCYEVER MHNDTWVLLNQRRGFLCNQAPHKHGFLEGRHAELCFLDVIPFWKL DLDQDYRVTCFTSWSPCFSCAQEMAKFISKNKHVSLCIFTARIYDD QGRCQEGLRTLAEAGAKISIMTYSEFKHCWDTFVDHQGCPFQPWD GLDEHSQDLSGRLRAILQNQEN
인간 APOBEC-3B 폴리펩티드 서열	1394	MNPQIRNPMERMYRDTFYDNFENEPILYGRSYTWLCYEVKIKRGR SNLLWDTGVFRGQVYFKPQYHAEMCFLSWFCGNQLPAYKCFQIT WFVSWTPCPDCVAKLAEFLSEHPNVTLTISAARLYYYWERDYRRAL CRLSQAGARVTIMDYEEFAYCWENFVYNEGQQFMPWYKFDENYA FLHRTLKEILRYLMDPDTFTFNFNNDPLVLRRRQTYLCYEVERLDN GTWVLMDQHMGFLCNEAKNLLCGFYGRHAELRFLDLVPSLQLDPA QIYRVTWFISWSPCFSWGCAGEVRAFLQENTHVRLRIFAARIYDYD PLYKEALQMLRDAGAQVSIMTYDEFEYCWDTFVYRQGCPFQPWD GLEEHSQALSGRLRAILQNQGN

[1075]

[1076]

랫트 APOBEC-3B 폴리펩티드 서열	1395	MQPQGLGPNAGMGPVCLGCSHRRPYSPIRNPLKKLYQQTFYFHFK NVRYAWGRKNNFLCYEVNGMDCALPVPLRQGVFRKQGHIHAELC FIYWFHDKVLRVLSPMEEFKVTWYMSWSPCSKCAEQVARFLAAHR NLSLAIFSSRLYYYLRNPNYQQKLCRLIQEGVHVAAMDLPEFKKCW NKFVDNDGQPFRPWMRLRINFSFYDCKLQEIFSRMNLLREDVFYLQ FNNSHRVKPVQNRYYRRKSYLCYQLERANGQEPLKGYLLYKKGEQ HVEILFLEKMRSMELSQVRITCYLTWSPCPNCARQLAAFKKDHPDLI LRIYTSRLYFWRKKFQKGLCTLWRSGIHVDVMDLPQFADCWTNFV NPQRPFRPWNELEKNSWRIQRRLRRIKESWGL
소 APOBEC-3B 폴리펩티드 서열	1396	DGWEVAFRSGTVLKAGVLGVSMTEGWAGSGHPGQGACVWTPGT RNTMNLLREVLFKQQFGNQPRVPAPYYRRKTYLCYQLKQRNDLTL DRGCFRNKKQRHAERFIDKINSLDLNPSQSYKIICYITWSPCPNCAN ELVNFITRNNHLKLEIFASRLYFHWIKSFKMGLQDLQNAGISVAVMT HTEFEDCWEQFVDNQSRPFQPWDKLEQYSASIRRRLQRILTAPI
침팬지 APOBEC-3B 폴리펩티드 서열	1397	MNPQIRNPMEWMYQRTFYYNFENEPILYGRSYTWLCYEVKIRRGH SNLLWDTGVFRGQMYSQPEHHAEMCFLSWFCGNQLSAYKCFQIT WFVSWTPCPDCVAKLAKFLAEHPNVTLTISAARLYYYWERDYRRAL CRLSQAGARVKIMDDEEFAYCWENFVYNEGQPFMPWYKFDDNYA FLHRTLKEIIRHLMDPDTFTFNFNNDPLVLRRHQTYLCYEVERLDNG TWVLMDQHMGFLCNEAKNLLCGFYGRHAELRFLDLVPSLQLDPAQ IYRVTWFISWSPCFSWGCAGQVRAFLQENTHVRLRIFAARIYDYDP LYKEALQMLRDAGAQVSIMTYDEFEYCWDTFVYRQGCPFQPWDG LEEHSQALSGRLRAILQVRASSLCMVPHRPPPPPQSPGPCLPLCSE PPLGSLLPTGRPAPSLPFLLTASFSFPPPASLPPLPSLSLSPGHLPV PSFHSLTSCSIQPPCSSRIRETEGWASVSKEGRDLG
인간 APOBEC-3C 폴리펩티드 서열	1398	MNPQIRNPMKAMYPGTFYFQFKNLWEANDRNETWLCFTVEGIKRR SVVSWKTGVFRNQVDSETHCHAERCFLSWFCDDILSPNTKYQVTW YTSWSPCPDCAGEVAEFLARHSNVNLTIFTARLYYFQYPCYQEGLR SLSQEGVAVEIMDYEDFKYCWENFVYNDNEPFKPWKGLKTNFRLL KRRLRESLQ
고릴라 APOBEC-3C 폴리펩티드 서열	1399	MNPQIRNPMKAMYPGTFYFQFKNLWEANDRNETWLCFTVEGIKRR SVVSWKTGVFRNQVDSETHCHAERCFLSWECDDILSPNTNYQVT WYTSWSPCPECAGEVAEFLARHSNVNLTIFTARLYYFQDTDYQEG LRSLSQEGVAVKIMDYKDFKYCWENFVYNDDEPFKPWKGLKYNFR FLKRRLQEILE
붉은털 원숭이 APOBEC-3A 폴리펩티드 서열	1400	MDGSPASRPRHLMDPNTFTFNFNNDLSVRGRHQTYLCYEVERLDN GTWVPMDERRGFLCNKAKNVPCGDYGCHVELRFLCEVPSWQLDP AQTYRVTWFISWSPCFRRGCAGQVRVFLQENKHVRLRIFAARIYDY DPLYQEALRTLRDAGAQVSIMTYEEFKHCWDTFVDRQGRPFQPW DGLDEHSQALSGRLRAILQNQGN
소 APOBEC-3A 폴리펩티드 서열	1401	MDEYTFTENFNNQGWPSKTYLCYEMERLDGDATIPLDEYKGFVRN KGLDQPEKPCHAELYFLGKIHSWNLDRNQHYRLTCFISWSPCYDC AQKLTTFLKENHHISLHILASRIYTHNRFGCHQSGLCELQAAGARITI MTFEDFKHCWETFVDHKGKPFQPWEGLNVKSQALCTELQAILKTQ QN
인간 APOBEC-3H 폴리펩티드 서열	1402	MALLTAETFRLQFNNKRRLRRPYYPRKALLCYQLTPQNGSTPTRGY FENKKKCHAEICFINEIKSMGLDETQCYQVTCYLTWSPCSSCAWEL VDFIKAHDHLNLGIFASRLYYHWCKPQQKGLRLLCGSQVPVEVMGF PKFADCWENFVDHEKPLSFNPYKMLEELDKNSRAIKRRLERIKIPGV RAQGRYMDILCDAEV
붉은털 원숭이 APOBEC-3H 폴리펩티드 서열	1403	MALLTAKTFSLQFNNKRRVNKPYYPRKALLCYQLTPQNGSTPTRG HLKNKKKDHAEIRFINKIKSMGLDETQCYQVTCYLTWSPCPSCAGE LVDFIKAHRHLNLRIFASRLYYHWRPNYQEGLLLLCGSQVPVEVMG LPEFTDCWENFVDHKEPPSFNPSEKLEELDKNSQAIKRRLERIKSR SVDVLENGLRSLQLGPVTPSSSIRNSR
인간 APOBEC-3D 폴리펩티드 서열	1404	MNPQIRNPMERMYRDTFYDNFENEPILYGRSYTWLCYEVKIKRGR SNLLWDTGVFRGPVLPKRQSNHRQEVYFRFENHAEMCFLSWFCG NRLPANRRFQITWFVSWNPCLPCVVKVTKFLAEHPNVTLTISAARL YYYRDRDWRWVLLRLHKAGARVKIMDYEDFAYCWENFVCNEGQP

		FMPWYKFDDNYASLHRTLKEILRNPMEAMYPHIFYFHFKNLLKACG RNESWLCFTMEVTKHHSAVFRKRGVFRNQVDPETHCHAERCFLS WFCDDILSPNTNYEVTWYTSWSPCPECAGEVAEFLARHSNVNLTIF TARLCYFWDTDYQEGLCSLSQEGASVKIMGYKDFVSCWKNFVYSD DEPFKPWKGLQTNFRLLKRRLREILQ
인간 APOBEC-1 폴리펩티드 서열	1405	MTSEKGPSTGDPTLRRRIEPWEFDVFYDPRELRKEACLLYEIKWG MSRKIWRSSGKNTTNHVEVNFIKKFTSERDFHPSMSCSITWFLSWS PCWECSQAIREFLSRHPGVTLVIYVARLFWHMDQQNRQGLRDLVN SGVTIQIMRASEYYHCWRNFVNYPPGDEAHWPQYPPLWMMLYAL ELHCIILSLPPCLKISRRWQNHLTFFRLHLQNCHYQTIPPHILLATGLI HPSVAWR
마우스 APOBEC-1 폴리펩티드 서열	1406	MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELRKETCLLYEINWGG RHSVWRHTSQNTSNHVEVNFLEKFTTERYFRPNTRCSITWFLSWS PCGECSRAITEFLSRHPYVTLFIYIARLYHHTDQRNRQGLRDLISSG VTIQIMTEQEYCYCWRNFVNYPPSNEAYWPRYPHLWVKLYVLELY CIILGLPPCLKILRRKQPQLTFFTITLQTCHYQRIPPHLLWATGLK
인간 APOBEC-2 폴리펩티드 서열	1407	MAQKEEAAVATEAASQNGEDLENLDDPEKLKELIELPPFEIVTGERL PANFFKFQFRNVEYSSGRNKTFLCYVVEAQGKGGQVQASRGYLE DEHAAAHAEEAFFNTILPAFDPALRYNVTWYVSSSPCAACADRIIKT LSKTKNLRLLILVGRLFMWEEPEIQAALKKLKEAGCKLRIMKPQDFE YVWQNFVEQEEGESKAFQPWEDIQENFLYYEEKLADILK
마우스 APOBEC-2 폴리펩티드 서열	1408	MAQKEEAAEAAAPASQNGDDLENLEDPEKLKELIDLPPFEIVTGVRL PVNFFKFQFRNVEYSSGRNKTFLCYVVEVQSKGGQAQATQGYLED EHAGAHAEEAFFNTILPAFDPALKYNVTWYVSSSPCAACADRILKTL SKTKNLRLLILVSRLFMWEEPEVQAALKKLKEAGCKLRIMKPQDFEY IWQNFVEQEEGESKAFEPWEDIQENFLYYEEKLADILK
랫트 APOBEC-2 폴리펩티드 서열	1409	MAQKEEAAEAAAPASQNGDDLENLEDPEKLKELIDLPPFEIVTGVRL PVNFFKFQFRNVEYSSGRNKTFLCYVVEAQSKGGQVQATQGYLED EHAGAHAEEAFFNTILPAFDPALKYNVTWYVSSSPCAACADRILKTL SKTKNLRLLILVSRLFMWEEPEVQAALKKLKEAGCKLRIMKPQDFEY LWQNFVEQEEGESKAFEPWEDIQENFLYYEEKLADILK
페트로미존 마리누스 CDA1 (pmCDAI) 폴리펩티드 서열	1410	MTDAEYVRIHEKLDIYTFKKQFFNNKKSVSHRCYVLFELKRRGERR ACFWGYAVNKPQSGTERGIHAEIFSIRKVEEYLRDNPGQFTINWYS SWSPCADCAEKILEWYNQELRGNGHTLKIWACKLYYEKNARNQIG LWNLRDNGVGLNVMVSEHYQCCRKIFIQSSHNQLNENRWLEKTLK RAEKRRSELSFMIQVKILHTTKSPAV
인간 APOBEC3G D316R D317R 폴리펩티드 서열	1411	MKPHFRNTVERMYRDTFSYNFYNRPILSRRNTVWLCYEVKTKGPS RPPLDAKIFRGQVYSELKYHPEMRFFHWFSKWRKLHRDQEYEVT WYISWSPCTKCTRDMATFLAEDPKVTLTIFVARLYYFWDPDYQEAL RSLCQKRDGPRATMKFNYDEFQHCWSKFVYSQRELFEPWNNLPK YYILLHFMLGEILRHSMDPPTFTFNFNNEPWVRGRHETYLCYEVER MHNDTWVLLNQRRGFLCNQAPHKHGFLEGRHAELCFLDVIPFWKL DLDQDYRVTCFTSWSPCFSCAQEMAKFISKKHVSLCIFTARIYRRQ GRCQEGLRTLAEAGAKISFTYSEFKHCWDTFVDHQGCPFQPWDG LDEHSQDLSGRLRAILQNQEN
인간 APOBEC3G 사슬 A 폴리펩티드 서열	1412	MDPPTFTFNFNNEPWWGRHETYLCYEVERMHNDTWVLLNQRRGF LCNQAPHKHGFLEGRHAELCFLDVIPFWKLDLDQDYRVTCFTSWS PCFSCAQEMAKFISKNKHVSLCIFTARIYDDQGRCQEGLRTLAEAG AKISFTYSEFKHCWDTFVDHQGCPFQPWDGLDEHSQDLSGRLRAI LQ
인간 APOBEC3G 사슬 A D120R D121R 폴리펩티드 서열	1414	MDPPTFTFNFNNEPWVRGRHETYLCYEVERMHNDTWVLLNQRRG FLCNQAPHKHGFLEGRHAELCFLDVIPFWKLDLDQDYRVTCFTSW SPCFSCAQEMAKFISKNKHVSLCIFTARIYRRQGRCQEGLRTLAEA GAKISFMTYSEFKHCWDTFVDHQGCPFQPWDGLDEHSQDLSGRL RAILQ
hAPOBEC-4 폴리펩티드 서열	1415	MEPIYEEYLANHGTIVKPYYWLSFSLDCSNCPYHIRTGEEARVSLTE FCQIFGFPYGTTFPQTKHLTFYELKTSSGSLVQKGHASSCTGNYIHP

[1077]

[1078]

		ESMLFEMNGYLDSAIYNNDSIRHIILYSNNSPCNEANHCCISKMYNF LITYPGITLSIYFSQLYHTEMDFPASAWNREALRSLASLWPRVVLSPI SGGIWHSVLHSFISGVSGSHVFQPILTGRALADRHNAYEINAITGVK PYFTDVLLQTKRNPNTKAQEALESYPLNNAFPGQFFQMPSGQLQP NLPPDLRAPVVFVLVPLRDLPPMHMGQNPNKPRNIVRHLNMPQMS FQETKDLGRLPTGRSVEIVEITEQFASSKEADEKKKKKGKK
rAPOBEC-4 폴리펩티드 서열	1416	MEPLYEEYLTHSGTIVKPYYWLSVSLNCTNCPYHIRTGEEARVPYT EFHQTFGFPWSTYPQTKHLTFYELRSSSGNLIQKGLASNCTGSHTH PESMLFERDGYLDSLIFHDSNIRHIILYSNNSPCDEANHCCISKMYNF LMNYPEVTLSVFFSQLYHTENQFPTSAWNREALRGLASLWPQVTL SAISGGIWQSILETFVSGISEGLTAVRPFTAGRTLTDRYNAYEINCIT EVKPYFTDALHSWQKENQDQKVWAASENQPLHNTTPAQWQPDM SQDCRTPAVFMLVPYRDLPPIHVNPSPQKPRTVVRHLNTLQLSASK VKALRKSPSGRPVKKEEARKGSTRSQEANETNKSKWKKQTLFIKS NICHLLEREQKKIGILSSWSV
rAPOBEC-1 (엘타 177-186) 폴리펩티드 서열	1421	MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELRKETCLLYEINWGG RHSIWRHTSQNTNKHVEVNFIEKFTTERYFCPNTRCSITWFLSWSP CGECSRAITEFLSRYPHVTLFIYIARLYHHADPRNRQGLRDLISSGVT IQIMTEQESGYCWRNFVNYSPSNEAHWPRYPHLWVRGLPPCLNIL RRKQPQLTFFTIALQSCHYQRLPPHILWATGLK
rAPOBEC-1 (엘타 202-213) 폴리펩티드 서열	1422	MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELRKETCLLYEINWGG RHSIWRHTSQNTNKHVEVNFIEKFTTERYFCPNTRCSITWFLSWSP CGECSRAITEFLSRYPHVTLFIYIARLYHHADPRNRQGLRDLISSGVT IQIMTEQESGYCWRNFVNYSPSNEAHWPRYPHLWVRLYVLELYCII LGLPPCLNILRRKQPQHYQRLPPHILWATGLK
마우스 AID (mAPOBEC-4) 폴리펩티드 서열	1373	MDSLLMKQKKFLYHFKNVRWAKGRHETYLCYVVKRRDSATSCSLD FGHLRNKSGCHVELLFLRYISDWDLDPGRCYRVTWFTSWSPCYDC ARHVAEFLRWNPNLSLRIFTARLYFCEDRKAEPEGLRRLHRAGVQI GIMTFKDYFYCWNTFVENRERTFKAWEGLHENSVRLTRQLRRILLP LYEVDDLRDAFRMLGF
pmCDA-1 폴리펩티드 서열	1417	MAGYECVRVSEKLDFDTFEFQFENLHYATERHRTYVIFDVKPQSAG GRSRRLWGYIINNPNVCHAELILMSMIDRHLESNPGVYAMTWYMS WSPCANCSSKLNPWLKNLLEEQGHTLTMHFSRIYDRDREGDHRGL RGLKHVSNSFRMGVVGRAEVKECLAEYVEASRRTLTWLDTTESMA AKMRRKLFCILVRCAGMRESGIPLHLFTLQTPLLSGRVVWWRV
pmCDA-2 폴리펩티드 서열	1418	MELREVVDCALASCVRHEPLSRVAFLRCFAAPSQKPRGTVILFYVE GAGRGVTGGHAVNYNKQGTSIHAEVLLLSAVRAALLRRRRCEDGE EATRGCTLHCYSTYSPCRDCVEYIQEFGASTGVRVVIHCCRLYELD VNRRRSEAEGVLRSLSRLGRDFRLMGPRDAIALLLGGRLANTADG ESGASGNAWVTETNVVEPLVDMTGFGDEDLHAQVQRNKQIREAY ANYASAVSLMLGELHVDPDKFPFLAEFLAQTSVEPSGTPRETRGRP RGASSRGPEIGRQRPADFERALGAYGLFLHPRIVSREADREEIKRD LIVVMRKHNYQGP
pmCDA-5 폴리펩티드 서열	1419	MAGDENVRVSEKLDFDTFEFQFENLHYATERHRTYVIFDVKPQSAG GRSRRLWGYIINNPNVCHAELILMSMIDRHLESNPGVYAMTWYMS WSPCANCSSKLNPWLKNLLEEQGHTLMMHFSRIYDRDREGDHRG LRGLKHVSNSFRMGVVGRAEVKECLAEYVEASRRTLTWLDTTESM AAKMRRKLFCILVRCAGMRESGMPLHLFT
yCD 폴리펩티드 서열	1420	MVTGGMASKWDQKGMDIAYEEAALGYKEGGVPIGGCLINNKDGSV LGRGHNMRFQKGSATLHGEISTLENCGRLEGKVYKDTTLYTTLSPC DMCTGAIIMYGIPRCVVGENVNFKSKGEKYLQTRGHEVVVVDDER CKKIMKQFIDERPQDWFEDIGE
NLS	84	KRTADGSEFESPKKKRKV
NLS	85	KRPAATKKAGQAKKKK
NLS	86	KKTELQTTNAENKTKKL
NLS	87	KRGINDRNFWRGENGRKTR
NLS	88	RKSGKIAAIVVKRPRK
NLS	89	PKKKRKV
NLS	90	MDSLLMNRRKFLYQFKNVRWAKGRRETYLC

[1079]

WT cas9 도메인	223	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNL IGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED ILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHK PENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVE NTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQ RKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNT KYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAY LNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAK YFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATV RKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK KYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEK NPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKG NELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEII EQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLG APAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGG D
gRNA 스캐폴드 뉴클레오티드 서열	230	GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGU UAUCAACUUGAAAAAGUGGGACCGAGUCGGUGCUUUU
야생형 spCas9 폴리뉴클레오티드 서열	231	ATGGATAAGAAATACTCAATAGGCTTAGATATCGGCACAAATAGC GTCGGATGGGCGGTGATCACTGATGATTATAAGGTTCCGTCTAA AAAGTTCAAGGTTCTGGGAAATACAGACCGCCACAGTATCAAAA AAAATCTTATAGGGGCTCTTTTATTTGGCAGTGGAGAGACAGCG GAAGCGACTCGTCTCAAACGGACAGCTCGTAGAAGGTATACAC GTCGGAAGAATCGTATTTGTTATCTACAGGAGATTTTTTCAAATG AGATGGCGAAAGTAGATGATAGTTTCTTTCATCGACTTGAAGAGT CTTTTTTGGTGGAAGAAGACAAGAAGCATGAACGTCATCCTATTT TTGGAAATATAGTAGATGAAGTTGCTTATCATGAGAAATATCCAA CTATCTATCATCTGCGAAAAAAATTGGCAGATTCTACTGATAAAG CGGATTTGCGCTTAATCTATTTGGCCTTAGCGCATATGATTAAGT TTCGTGGTCATTTTTTGATTGAGGGAGATTTAAATCCTGATAATA GTGATGTGGACAAACTATTTATCCAGTTGGTACAAATCTACAATC AATTATTTGAAGAAAACCCTATTAACGCAAGTAGAGTAGATGCTA AAGCGATTCTTTCTGCACGATTGAGTAAATCAAGACGATTAGAAA ATCTCATTGCTCAGCTCCCCGGTGAGAAGAGAAATGGCTTGTTT GGGAATCTCATTGCTTTGTCATTGGGATTGACCCCTAATTTTAAA TCAAATTTTGATTTGGCAGAAGATGCTAAATTACAGCTTTCAAAA GATACTTACGATGATGATTTAGATAATTTATTGGCGCAAATTGGA GATCAATATGCTGATTTGTTTTTGGCAGCTAAGAATTTATCAGAT GCTATTTTACTTTCAGATATCCTAAGAGTAAATAGTGAAATAACTA AGGCTCCCCTATCAGCTTCAATGATTAAGCGCTACGATGAACAT CATCAAGACTTGACTCTTTTAAAAGCTTTAGTTCGACAACAACTT CCAGAAAAGTATAAAGAAATCTTTTTTGATCAATCAAAAAACGGA TATGCAGGTTATATTGATGGGGGAGCTAGCCAAGAAGAATTTTA TAAATTTATCAAACCAATTTTAGAAAAAATGGATGGTACTGAGGA

[1080]

		ATTATTGGTGAAACTAAATCGTGAAGATTTGCTGCGCAAGCAAC GGACCTTTGACAACGGCTCTATTCCCCATCAAATTCACTTGGGT GAGCTGCATGCTATTTTGAGAAGACAAGAAGACTTTTATCCATTT TTAAAAGACAATCGTGAGAAGATTGAAAAAATCTTGACTTTTCGA ATTCCTTATTATGTTGGTCCATTGGCGCGTGGCAATAGTCGTTTT GCATGGATGACTCGGAAGTCTGAAGAAACAATTACCCCATGGAA TTTTGAAGAAGTTGTCGATAAAGGTGCTTCAGCTCAATCATTTAT TGAACGCATGACAAACTTTGATAAAAATCTTCCAAATGAAAAAGT ACTACCAAAACATAGTTTGCTTTATGAGTATTTTACGGTTTATAAC GAATTGACAAAGGTCAAATATGTTACTGAGGGAATGCGAAAACC AGCATTTCTTTCAGGTGAACAGAAGAAAGCCATTGTTGATTTACT CTTCAAAACAAATCGAAAAGTAACCGTTAAGCAATTAAAAGAAGA TTATTTCAAAAAAATAGAATGTTTTGATAGTGTTGAAATTTCAGGA GTTGAAGATAGATTTAATGCTTCATTAGGCGCCTACCATGATTTG CTAAAAATTATTAAAGATAAAGATTTTTTGGATAATGAAGAAAATG AAGATATCTTAGAGGATATTGTTTTAACATTGACCTTATTTGAAGA TAGGGGGATGATTGAGGAAAGACTTAAAACATATGCTCACCTCT TTGATGATAAGGTGATGAAACAGCTTAAACGTCGCCGTTATACT GGTTGGGGACGTTTGTCTCGAAAATTGATTAATGGTATTAGGGA TAAGCAATCTGGCAAAACAATATTAGATTTTTTGAAATCAGATGG TTTTGCCAATCGCAATTTTATGCAGCTGATCCATGATGATAGTTT GACATTTAAAGAAGATATTCAAAAAGCACAGGTGTCTGGACAAG GCCATAGTTTACATGAACAGATTGCTAACTTAGCTGGCAGTCCT GCTATTAAAAAAGGTATTTTACAGACTGTAAAAATTGTTGATGAA CTGGTCAAAGTAATGGGGCATAAGCCAGAAAATATCGTTATTGA AATGGCACGTGAAAATCAGACAACTCAAAAGGGCCAGAAAAATT CGCGAGAGCGTATGAAACGAATCGAAGAAGGTATCAAAGAATTA GGAAGTCAGATTCTTAAAGAGCATCCTGTTGAAAATACTCAATTG CAAAATGAAAAGCTCTATCTCTATTATCTACAAAATGGAAGAGAC ATGTATGTGGACCAAGAATTAGATATTAATCGTTTAAGTGATTAT GATGTCGATCACATTGTTCCACAAAGTTTCATTAAAGACGATTCA ATAGACAATAAGGTACTAACGCGTTCTGATAAAAATCGTGGTAAA TCGGATAACGTTCCAAGTGAAGAAGTAGTCAAAAAGATGAAAAA CTATTGGAGACAACTTCTAAACGCCAAGTTAATCACTCAACGTAA GTTTGATAATTTAACGAAAGCTGAACGTGGAGGTTTGAGTGAAC TTGATAAAGCTGGTTTTATCAAACGCCAATTGGTTGAAACTCGCC AAATCACTAAGCATGTGGCACAAATTTTGGATAGTCGCATGAATA CTAAATACGATGAAAATGATAAACTTATTCGAGAGGTTAAAGTGA TTACCTTAAAATCTAAATTAGTTTCTGACTTCCGAAAAGATTTCCA ATTCTATAAAGTACGTGAGATTAACAATTACCATCATGCCCATGA TGCGTATCTAAATGCCGTCGTTGGAACTGCTTTGATTAAGAAATA TCCAAAACTTGAATCGGAGTTTGTCTATGGTGATTATAAAGTTTA TGATGTTCGTAAAATGATTGCTAAGTCTGAGCAAGAAATAGGCAA AGCAACCGCAAAATATTTCTTTTACTCTAATATCATGAACTTCTTC AAAACAGAAATTACACTTGCAAATGGAGAGATTCGCAAACGCCC TCTAATCGAAACTAATGGGGAAACTGGAGAAATTGTCTGGGATA AAGGGCGAGATTTTGCCACAGTGCGCAAAGTATTGTCCATGCCC CAAGTCAATATTGTCAAGAAAACAGAAGTACAGACAGGCGGATT CTCCAAGGAGTCAATTTTACCAAAAAGAAATTCGGACAAGCTTAT TGCTCGTAAAAAAGACTGGGATCCAAAAAAATATGGTGGTTTTGA TAGTCCAACGGTAGCTTATTCAGTCCTAGTGGTTGCTAAGGTGG AAAAAGGGAAATCGAAGAAGTTAAAATCCGTTAAAGAGTTACTAG GGATCACAATTATGGAAAGAAGTTCCTTTGAAAAAAATCCGATTG ACTTTTTAGAAGCTAAAGGATATAAGGAAGTTAAAAAAGACTTAA TCATTAAACTACCTAAATATAGTCTTTTTGAGTTAGAAAACGGTC GTAAACGGATGCTGGCTAGTGCCGGAGAATTACAAAAAGGAAAT GAGCTGGCTCTGCCAAGCAAATATGTGAATTTTTTATATTTAGCT AGTCATTATGAAAAGTTGAAGGGTAGTCCAGAAGATAACGAACA AAAACAATTGTTTGTGGAGCAGCATAAGCATTATTTAGATGAGAT

		TATTGAGCAAATCAGTGAATTTTCTAAGCGTGTTATTTTAGCAGA TGCCAATTTAGATAAAGTTCTTAGTGCATATAACAAACATAGAGA CAAACCAATACGTGAACAAGCAGAAAATATTATTCATTTATTTAC GTTGACGAATCTTGGAGCTCCCGCTGCTTTTAAATATTTTGATAC AACAATTGATCGTAAACGATATACGTCTACAAAAGAAGTTTTAGA TGCCACTCTTATCCATCAATCCATCACTGGTCTTTATGAAACACG CATTGATTTGAGTCAGCTAGGAGGTGACTGA
spCas9 폴리펩티드 서열	232	MDKKYSIGLDIGTNSVGWAVITDDYKVPSKKFKVLGNTDRHSIKKNL IGALLFGSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LADSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV QIYNQLFEENPINASRVDAKAILSARLSKSRRLENLIAQLPGEKRNGL FGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGD QYADLFLAAKNLSDAILLSDILRVNSEITKAPLSASMIKRYDEHHQDL TLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPIL EKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQE DFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITP WNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVY NELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKED YFKKIECFDSVEISGVEDRFNASLGAYHDLLKIIKDKDFLDNEENEDI LEDIVLTLTLFEDRGMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGHSLHEQIANLAGSPAIKKGILQTVKIVDELVKVMGHKPE NIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENT QLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFIKDD SIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRK FDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKY DENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLN AVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYF FYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRK VLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKY GGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNP IDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNE LALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQ ISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAP AAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD
야생형 Cas9 폴리뉴클레오티드 서열	235	ATGGATAAAAAGTATTCTATTGGTTTAGACATCGGCACTAATTCC GTTGGATGGGCTGTCATAACCGATGAATACAAAGTACCTTCAAA GAAATTTAAGGTGTTGGGGAACACAGACCGTCATTCGATTAAAA AGAATCTTATCGGTGCCCTCCTATTCGATAGTGGCGAAACGGCA GAGGCGACTCGCCTGAAACGAACCGCTCGGAGAAGGTATACAC GTCGCAAGAACCGAATATGTTACTTACAAGAAATTTTTAGCAATG AGATGGCCAAAGTTGACGATTCTTTCTTTCACCGTTTGGAAGAGT CCTTCCTTGTCGAAGAGGACAAGAAACATGAACGGCACCCCATC TTTGGAAACATAGTAGATGAGGTGGCATATCATGAAAAGTACCC AACGATTTATCACCTCAGAAAAAAGCTAGTTGACTCAACTGATAA AGCGGACCTGAGGTTAATCTACTTGGCTCTTGCCCATATGATAA AGTTCCGTGGGCACTTTCTCATTGAGGGTGATCTAAATCCGGAC AACTCGGATGTCGACAAACTGTTCATCCAGTTAGTACAAACCTAT AATCAGTTGTTTGAAGAGAACCCTATAAATGCAAGTGGCGTGGA TGCGAAGGCTATTCTTAGCGCCCGCCTCTCTAAATCCCGACGGC TAGAAAACCTGATCGCACAATTACCCGGAGAGAAGAAAAATGGG TTGTTCGGTAACCTTATAGCGCTCTCACTAGGCCTGACACCAAA TTTTAAGTCGAACTTCGACTTAGCTGAAGATGCCAAATTGCAGCT TAGTAAGGACACGTACGATGACGATCTCGACAATCTACTGGCAC AAATTGGAGATCAGTATGCGGACTTATTTTTGGCTGCCAAAAACC TTAGCGATGCAATCCTCCTATCTGACATACTGAGAGTTAATACTG AGATTACCAAGGCGCCGTTATCCGCTTCAATGATCAAAAGGTAC GATGAACATCACCAAGACTTGACACTTCTCAAGGCCCTAGTCCG

[1081]

[1082]

		TCAGCAACTGCCTGAGAAATATAAGGAAATATTCTTTGATCAGTC GAAAAACGGGTACGCAGGTTATATTGACGGCGGAGCGAGTCAA GAGGAATTCTACAAGTTTATCAAACCCATATTAGAGAAGATGGAT GGGACGGAAGAGTTGCTTGTAAAACTCAATCGCGAAGATCTACT GCGAAAGCAGCGGACTTTCGACAACGGTAGCATTCCACATCAAA TCCACTTAGGCGAATTGCATGCTATACTTAGAAGGCAGGAGGAT TTTTATCCGTTCCTCAAAGACAATCGTGAAAAGATTGAGAAAATC CTAACCTTTCGCATACCTTACTATGTGGGACCCCTGGCCCGAGG GAACTCTCGGTTCGCATGGATGACAAGAAAGTCCGAAGAAACGA TTACTCCATGGAATTTTGAGGAAGTTGTCGATAAAGGTGCGTCA GCTCAATCGTTCATCGAGAGGATGACCAACTTTGACAAGAATTTA CCGAACGAAAAAGTATTGCCTAAGCACAGTTTACTTTACGAGTAT TTCACAGTGTACAATGAACTCACGAAAGTTAAGTATGTCACTGAG GGCATGCGTAAACCCGCCTTTCTAAGCGGAGAACAGAAGAAAG CAATAGTAGATCTGTTATTCAAGACCAACCGCAAAGTGACAGTTA AGCAATTGAAAGAGGACTACTTTAAGAAAATTGAATGCTTCGATT CTGTCGAGATCTCCGGGGTAGAAGATCGATTTAATGCGTCACTT GGTACGTATCATGACCTCCTAAAGATAATTAAAGATAAGGACTTC CTGGATAACGAAGAGAATGAAGATATCTTAGAAGATATAGTGTTG ACTCTTACCCTCTTTGAAGATCGGGAAATGATTGAGGAAAGACT AAAAACATACGCTCACCTGTTCGACGATAAGGTTATGAAACAGTT AAAGAGGCGTCGCTATACGGGCTGGGGACGATTGTCGCGGAAA CTTATCAACGGGATAAGAGACAAGCAAAGTGGTAAAACTATTCT CGATTTTCTAAAGAGCGACGGCTTCGCCAATAGGAACTTTATGC AGCTGATCCATGATGACTCTTTAACCTTCAAAGAGGATATACAAA AGGCACAGGTTTCCGGACAAGGGGACTCATTGCACGAACATATT GCGAATCTTGCTGGTTCGCCAGCCATCAAAAAGGGCATACTCCA GACAGTCAAAGTAGTGGATGAGCTAGTTAAGGTCATGGGACGTC ACAAACCGGAAAACATTGTAATCGAGATGGCACGCGAAAATCAA ACGACTCAGAAGGGGCAAAAAAACAGTCGAGAGCGGATGAAGA GAATAGAAGAGGGTATTAAAGAACTGGGCAGCCAGATCTTAAAG GAGCATCCTGTGGAAAATACCCAATTGCAGAACGAGAAACTTTA CCTCTATTACCTACAAAATGGAAGGGACATGTATGTTGATCAGG AACTGGACATAAACCGTTTATCTGATTACGACGTCGATCACATTG TACCCCAATCCTTTTTGAAGGACGATTCAATCGACAATAAAGTGC TTACACGCTCGGATAAGAACCGAGGGAAAAGTGACAATGTTCCA AGCGAGGAAGTCGTAAAGAAAATGAAGAACTATTGGCGGCAGCT CCTAAATGCGAAACTGATAACGCAAAGAAAGTTCGATAACTTAAC TAAAGCTGAGAGGGGTGGCTTGTCTGAACTTGACAAGGCCGGA TTTATTAAACGTCAGCTCGTGGAAACCCGCCAAATCACAAAGCA TGTTGCACAGATACTAGATTCCCGAATGAATACGAAATACGACG AGAACGATAAGCTGATTCGGGAAGTCAAAGTAATCACTTTAAAGT CAAAATTGGTGTCGGACTTCAGAAAGGATTTTCAATTCTATAAAG TTAGGGAGATAAATAACTACCACCATGCGCACGACGCTTATCTT AATGCCGTCGTAGGGACCGCACTCATTAAGAAATACCCGAAGCT AGAAAGTGAGTTTGTGTATGGTGATTACAAAGTTTATGACGTCCG TAAGATGATCGCGAAAAGCGAACAGGAGATAGGCAAGGCTACA GCCAAATACTTCTTTTATTCTAACATTATGAATTTCTTTAAGACGG AAATCACTCTGGCAAACGGAGAGATACGCAAACGACCTTTAATT GAAACCAATGGGGAGACAGGTGAAATCGTATGGGATAAGGGCC GGGACTTCGCGACGGTGAGAAAAGTTTTGTCCATGCCCCAAGTC AACATAGTAAAGAAAACTGAGGTGCAGACCGGAGGGTTTTCAAA GGAATCGATTCTTCCAAAAAGGAATAGTGATAAGCTCATCGCTC GTAAAAAGGACTGGGACCCGAAAAAGTACGGTGGCTTCGATAG CCCTACAGTTGCCTATTCTGTCCTAGTAGTGGCAAAAGTTGAGA AGGGAAAATCCAAGAAACTGAAGTCAGTCAAAGAATTATTGGGG ATAACGATTATGGAGCGCTCGTCTTTTGAAAAGAACCCCATCGA CTTCCTTGAGGCGAAAGGTTACAAGGAAGTAAAAAAGGATCTCA TAATTAAACTACCAAAGTATAGTCTGTTTGAGTTAGAAAATGGCC

		GAAAACGGATGTTGGCTAGCGCCGGAGAGCTTCAAAAGGGGAA CGAACTCGCACTACCGTCTAAATACGTGAATTTCCTGTATTTAGC GTCCCATTACGAGAAGTTGAAAGGTTCACCTGAAGATAACGAAC AGAAGCAACTTTTTGTTGAGCAGCACAAACATTATCTCGACGAAA TCATAGAGCAAATTTCGGAATTCAGTAAGAGAGTCATCCTAGCT GATGCCAATCTGGACAAAGTATTAAGCGCATACAACAAGCACAG GGATAAACCCATACGTGAGCAGGCGGAAAATATTATCCATTTGT TTACTCTTACCAACCTCGGCGCTCCAGCCGCATTCAAGTATTTTG ACACAACGATAGATCGCAAACGATACACTTCTACCAAGGAGGTG CTAGACGCGACACTGATTCACCAATCCATCACGGGATTATATGA AACTCGGATAGATTTGTCACAGCTTGGGGGTGACGGATCCCCCA AGAAGAAGAGGAAAGTCTCGAGCGACTACAAAGACCATGACGG TGATTATAAAGATCATGACATCGATTACAAGGATGACGATGACAA GGCTGCAGGA
야생형 Cas9 폴리펩티드 서열	236	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED ILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHK PENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVE NTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQ RKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNT KYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAY LNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAK YFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATV RKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK KYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEK NPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKG NELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEII EQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLG APAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGG D
스트렙토코커스 피오게네스로부터 Cas9 (NCBI Ref. Seq.: NC_002737.2) 폴리뉴클레오티드 서열	237	ATGGATAAGAAATACTCAATAGGCTTAGATATCGGCACAAATAGC GTCGGATGGGCGGTGATCACTGATGAATATAAGGTTCCGTCTAA AAAGTTCAAGGTTCTGGGAAATACAGACCGCCACAGTATCAAAA AAAATCTTATAGGGGCTCTTTTATTTGACAGTGGAGAGACAGCG GAAGCGACTCGTCTCAAACGGACAGCTCGTAGAAGGTATACAC GTCGGAAGAATCGTATTTGTTATCTACAGGAGATTTTTTCAAATG AGATGGCGAAAGTAGATGATAGTTTCTTTCATCGACTTGAAGAGT CTTTTTTGGTGGAAGAAGACAAGAAGCATGAACGTCATCCTATTT TTGGAAATATAGTAGATGAAGTTGCTTATCATGAGAAATATCCAA CTATCTATCATCTGCGAAAAAAATTGGTAGATTCTACTGATAAAG CGGATTTGCGCTTAATCTATTTGGCCTTAGCGCATATGATTAAGT TTCGTGGTCATTTTTTGATTGAGGGAGATTTAAATCCTGATAATA GTGATGTGGACAAACTATTTATCCAGTTGGTACAAACCTACAATC AATTATTTGAAGAAAACCCTATTAACGCAAGTGGAGTAGATGCTA AAGCGATTCTTTCTGCACGATTGAGTAAATCAAGACGATTAGAAA

[1083]

[1084]

		ATCTCATTGCTCAGCTCCCCGGTGAGAAGAAAAATGGCTTATTT GGGAATCTCATTGCTTTGTCATTGGGTTTGACCCCTAATTTTAAA TCAAATTTTGATTTGGCAGAAGATGCTAAATTACAGCTTTCAAAA GATACTTACGATGATGATTTAGATAATTTATTGGCGCAAATTGGA GATCAATATGCTGATTTGTTTTTGGCAGCTAAGAATTTATCAGAT GCTATTTTACTTTCAGATATCCTAAGAGTAAATACTGAAATAACTA AGGCTCCCCTATCAGCTTCAATGATTAAACGCTACGATGAACAT CATCAAGACTTGACTCTTTTAAAAGCTTTAGTTCGACAACAACTT CCAGAAAAGTATAAAGAAATCTTTTTTGATCAATCAAAAAACGGA TATGCAGGTTATATTGATGGGGGAGCTAGCCAAGAAGAATTTTA TAAATTTATCAAACCAATTTTAGAAAAAATGGATGGTACTGAGGA ATTATTGGTGAAACTAAATCGTGAAGATTTGCTGCGCAAGCAAC GGACCTTTGACAACGGCTCTATTCCCCATCAAATTCACTTGGGT GAGCTGCATGCTATTTTGAGAAGACAAGAAGACTTTTATCCATTT TTAAAAGACAATCGTGAGAAGATTGAAAAAATCTTGACTTTTCGA ATTCCTTATTATGTTGGTCCATTGGCGCGTGGCAATAGTCGTTTT GCATGGATGACTCGGAAGTCTGAAGAAACAATTACCCCATGGAA TTTTGAAGAAGTTGTCGATAAAGGTGCTTCAGCTCAATCATTTAT TGAACGCATGACAAACTTTGATAAAAATCTTCCAAATGAAAAAGT ACTACCAAAACATAGTTTGCTTTATGAGTATTTTACGGTTTATAAC GAATTGACAAAGGTCAAATATGTTACTGAAGGAATGCGAAAACC AGCATTTCTTTCAGGTGAACAGAAGAAAGCCATTGTTGATTTACT CTTCAAAACAAATCGAAAAGTAACCGTTAAGCAATTAAAAGAAGA TTATTTCAAAAAAATAGAATGTTTTGATAGTGTTGAAATTTCAGGA GTTGAAGATAGATTTAATGCTTCATTAGGTACCTACCATGATTTG CTAAAAATTATTAAAGATAAAGATTTTTTGGATAATGAAGAAAATG AAGATATCTTAGAGGATATTGTTTTAACATTGACCTTATTTGAAGA TAGGGAGATGATTGAGGAAAGACTTAAAACATATGCTCACCTCTT TGATGATAAGGTGATGAAACAGCTTAAACGTCGCCGTTATACTG GTTGGGGACGTTTGTCTCGAAAATTGATTAATGGTATTAGGGATA AGCAATCTGGCAAAACAATATTAGATTTTTTGAAATCAGATGGTT TTGCCAATCGCAATTTTATGCAGCTGATCCATGATGATAGTTTGA CATTTAAAGAAGACATTCAAAAAGCACAAGTGTCTGGACAAGGC GATAGTTTACATGAACATATTGCAAATTTAGCTGGTAGCCCTGCT ATTAAAAAAGGTATTTTACAGACTGTAAAAGTTGTTGATGAATTG GTCAAAGTAATGGGGCGGCATAAGCCAGAAAATATCGTTATTGA AATGGCACGTGAAAATCAGACAACTCAAAAGGGCCAGAAAAATT CGCGAGAGCGTATGAAACGAATCGAAGAAGGTATCAAAGAATTA GGAAGTCAGATTCTTAAAGAGCATCCTGTTGAAAATACTCAATTG CAAAATGAAAAGCTCTATCTCTATTATCTCCAAAATGGAAGAGAC ATGTATGTGGACCAAGAATTAGATATTAATCGTTTAAGTGATTAT GATGTCGATCACATTGTTCCACAAAGTTTCCTTAAAGACGATTCA ATAGACAATAAGGTCTTAACGCGTTCTGATAAAAATCGTGGTAAA TCGGATAACGTTCCAAGTGAAGAAGTAGTCAAAAAGATGAAAAA CTATTGGAGACAACTTCTAAACGCCAAGTTAATCACTCAACGTAA GTTTGATAATTTAACGAAAGCTGAACGTGGAGGTTTGAGTGAAC TTGATAAAGCTGGTTTTATCAAACGCCAATTGGTTGAAACTCGCC AAATCACTAAGCATGTGGCACAAATTTTGGATAGTCGCATGAATA CTAAATACGATGAAAATGATAAACTTATTCGAGAGGTTAAAGTGA TTACCTTAAAATCTAAATTAGTTTCTGACTTCCGAAAAGATTTCCA ATTCTATAAAGTACGTGAGATTAACAATTACCATCATGCCCATGA TGCGTATCTAAATGCCGTCGTTGGAACTGCTTTGATTAAGAAATA TCCAAAACTTGAATCGGAGTTTGTCTATGGTGATTATAAAGTTTA TGATGTTCGTAAAATGATTGCTAAGTCTGAGCAAGAAATAGGCAA AGCAACCGCAAAATATTTCTTTTACTCTAATATCATGAACTTCTTC AAAACAGAAATTACACTTGCAAATGGAGAGATTCGCAAACGCCC TCTAATCGAAACTAATGGGGAAACTGGAGAAATTGTCTGGGATA AAGGGCGAGATTTTGCCACAGTGCGCAAAGTATTGTCCATGCCC CAAGTCAATATTGTCAAGAAAACAGAAGTACAGACAGGCGGATT

		CTCCAAGGAGTCAATTTTACCAAAAAGAAATTCGGACAAGCTTAT TGCTCGTAAAAAAGACTGGGATCCAAAAAAATATGGTGGTTTTGA TAGTCCAACGGTAGCTTATTCAGTCCTAGTGGTTGCTAAGGTGG AAAAAGGGAAATCGAAGAAGTTAAAATCCGTTAAAGAGTTACTAG GGATCACAATTATGGAAAGAAGTTCCTTTGAAAAAAATCCGATTG ACTTTTTAGAAGCTAAAGGATATAAGGAAGTTAAAAAAGACTTAA TCATTAAACTACCTAAATATAGTCTTTTTGAGTTAGAAAACGGTC GTAAACGGATGCTGGCTAGTGCCGGAGAATTACAAAAAGGAAAT GAGCTGGCTCTGCCAAGCAAATATGTGAATTTTTTATATTTAGCT AGTCATTATGAAAAGTTGAAGGGTAGTCCAGAAGATAACGAACA AAAACAATTGTTTGTGGAGCAGCATAAGCATTATTTAGATGAGAT TATTGAGCAAATCAGTGAATTTTCTAAGCGTGTTATTTTAGCAGA TGCCAATTTAGATAAAGTTCTTAGTGCATATAACAAACATAGAGA CAAACCAATACGTGAACAAGCAGAAAATATTATTCATTTATTTAC GTTGACGAATCTTGGAGCTCCCGCTGCTTTTAAATATTTTGATAC AACAATTGATCGTAAACGATATACGTCTACAAAAGAAGTTTTAGA TGCCACTCTTATCCATCAATCCATCACTGGTCTTTATGAAACACG CATTGATTTGAGTCAGCTAGGAGGTGACTGA
촉매적으로 불활성 Cas9 (dCas9) 폴리펩티드 서열	238	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED ILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHK PENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVE NTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLK DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQ RKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNT KYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAY LNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAK YFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATV RKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK KYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEK NPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKG NELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEII EQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLG APAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGG D
tr|F0NN87|F0NN87_SULIHCRISPR-회합된Casx 단백질 OS = 술폴로버스 아이스랜디커스 (균주 HVE10/4) GN = SiH_0402 PE=4 SV=1); CasX 폴리펩티드 서열	239	MEVPLYNIFGDNYIIQVATEAENSTIYNNKVEIDDEELRNVLNLAYKIA KNNEDAAAERRGKAKKKKGEEGETTTSNIILPLSGNDKNPWTETLK CYNFPTTVALSEVFKNFSQVKECEEVSAPSFVKPEFYEFGRSPGM VERTRRVKLEVEPHYLIIAAAGWVLTRLGKAKVSEGDYVGVNVFTP TRGILYSLIQNVNGIVPGIKPETAFGLWIARKVVSSVTNPNVSVVRIY TISDAVGQNPTTINGGFSIDLTKLLEKRYLLSERLEAIARNALSISSNM RERYIVLANYIYEYLTG SKRLEDLLYFANRDLIMNLNSDDGKVRDLK LISAYVNGELIRGEG

[1085]

tr|F0NH53|F0NH53_SULIR CRISPR 회합된 단백질, Casx OS = 술폴로버스 아이스랜디커스 (균주 REY15A) GN=SiRe_0771 PE=4 SV=1); CasX 폴리펩티드 서열	240	MEVPLYNIFGDNYIIQVATEAENSTIYNNKVEIDDEELRNVLNLAYKIA KNNEDAAAERRGKAKKKKGEEGETTTSNIILPLSGNDKNPWTETLK CYNFPTTVALSEVFKNFSQVKECEEVSAPSFVKPEFYKFGRSPGM VERTRRVKLEVEPHYLIMAAAGWVLTRLGKAKVSEGDYVGVNVFT PTRGILYSLIQNVNGIVPGIKPETAFGLWIARKVVSSVTNPNVSVVSI YTISDAVGQNPTTINGGFSIDLTKLLEKRDLLSERLEAIARNALSISSN MRERYIVLANYIYEYLTGSKRLEDLLYFANRDLIMNLNSDDGKVRDL KLISAYVNGELIRGEG
CasX 폴리펩티드 서열	241	MEKRINKIRKKLSADNATKPVSRSGPMKTLLVRVMTDDLKKRLEKR RKKPEVMPQVISNNAANNLRMLLDDYTKMKEAILQVYWQEFKDDH VGLMCKFAQPASKKIDQNKLKPEMDEKGNLTTAGFACSQCGQPLF VYKLEQVSEKGKAYTNYFGRCNVAEHEKLILLAQLKPVKDSDEAVT YSLGKFGQRALDFYSIHVTKESTHPVKPLAQIAGNRYASGPVGKAL SDACMGTIASFLSKYQDIIIEHQKVVKGNQKRLESLRELAGKENLEY PSVTLPPQPHTKEGVDFAYNEVIARVRMWVNLNLWQKLKLSRDDA KPLLRLKGFPSFPVVERRENEVDWWNTINEVKKLIDAKRDMGRVF WSGVTAEKRNTILEGYNYLPNENDHKKREGSLENPKKPAKRQFGD LLLYLEKKYAGDWGKVFDEAWERIDKKIAGLTSHIEREEARNAEDA QSKAVLTDWLRAKASFVLERLKEMDEKEFYACEIQLQKWYGDLRG NPFAVEAENRVVDISGFSIGSDGHSIQYRNLLAWKYLENGKREFYLL MNYGKKGRIRFTDGTDIKKSGKWQGLLYGGGKAKVIDLTFDPDDE QLIILPLAFGTRQGREFIWNDLLSLETGLIKLANGRVIEKTIYNKKIGR DEPALFVALTFERREVVDPSNIKPVNLIGVARGENIPAVIALTDPEGC PLPEFKDSSGGPTDILRIGEGYKEKQRAIQAAKEVEQRRAGGYSRK FASKSRNLADDMVRNSARDLFYHAVTHDAVLVFANLSRGFGRQGK RTFMTERQYTKMEDWLTAKLAYEGLTSKTYLSKTLAQYTSKTCSNC GFTITYADMDVMLVRLKKTSDGWATTLNNKELKAEYQITYYNRYKR QTVEKELSAELDRLSEESGNNDISKWTKGRRDEALFLLKKRFSHRP VQEQFVCLDCGHEVHAAEQAALNIARSWLFLNSNSTEFKSYKSGK QPFVGAWQAFYKRRLKEVWKPNA
APG80656.1 CRISPR-회합된 단백질 CasY [배양되지 않은 파르쿠박테리아 그룹 박테리움]; CasY 폴리펩티드 서열	**242**	MSKRHPRISGVKGYRLHAQRLEYTGKSGAMRTIKYPLYSSPSGGR TVPREIVSAINDDYVGLYGLSNFDDLYNAEKRNEEKVYSVLDFWYD CVQYGAVFSYTAPGLLKNVAEVRGGSYELTKTLKGSHLYDELQIDK VIKFLNKKEISRANGSLDKLKKDIIDCFKAEYRERHKDQCNKLADDIK NAKKDAGASLGERQKKLFRDFFGISEQSENDKPSFTNPLNLTCCLL PFDTVNNNRNRGEVLFNKLKEYAQKLDKNEGSLEMWEYIGIGNSG TAFSNFLGEGFLGRLRENKITELKKAMMDITDAWRGQEQEEELEKR LRILAALTIKLREPKFDNHWGGYRSDINGKLSSWLQNYINQTVKIKE DLKGHKKDLKKAKEMINRFGESDTKEEAVVSSLLESIEKIVPDDSAD DEKPDIPAIAIYRRFLSDGRLTLNRFVQREDVQEALIKERLEAEKKKK PKKRKKKSDAEDEKETIDFKELFPHLAKPLKLVPNFYGDSKRELYKK YKNAAIYTDALWKAVEKIYKSAFSSSLKNSFFDTDFDKDFFIKRLQKI FSVYRRFNTDKWKPIVKNSFAPYCDIVSLAENEVLYKPKQSRSRKS AAIDKNRVRLPSTENIAKAGIALARELSVAGFDWKDLLKKEEHEEYID LIELHKTALALLLAVTETQLDISALDFVENGTVKDFMKTRDGNLVLEG RFLEMFSQSIVFSELRGLAGLMSRKEFITRSAIQTMNGKQAELLYIP HEFQSAKITTPKEMSRAFLDLAPAEFATSLEPESLSEKSLLKLKQMR YYPHYFGYELTRTGQGIDGGVAENALRLEKSPVKKREIKCKQYKTL GRGQNKIVLYVRSSYYQTQFLEWFLHRPKNVQTDVAVSGSFLIDEK KVKTRWNYDALTVALEPVSGSERVFVSQPFTIFPEKSAEEEGQRYL GIDIGEYGIAYTALEITGDSAKILDQNFISDPQLKTLREEVKGLKLDQR RGTFAMPSTKIARIRESLVHSLRNRIHHLALKHKAKIVYELEVSRFEE GKQKIKKVYATLKKADVYSEIDADKNLQTTVWGKLAVASEISASYTS QFCGACKKLWRAEMQVDETITTQELIGTVRVIKGGTLIDAIKDFMRP

[1086]

		PIFDENDTPFPKYRDFCDKHHISKKMRGNSCLFICPFCRANADADIQ ASQTIALLRYVKEEKKVEDYFERFRKLKNIKVLGQMKKI
야생형 Cpf1 폴리펩티드 서열	246	MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYK KAKQIIDKYHQFFIEEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQK DFKSAKDTIKKQISEYIKDSEKFKNLFNQNLIDAKKGQESDLILWLKQ SKDNGIELFKANSDITDIDEALEIIKSFKGWTTYFKGFHENRKNVYSS NDIPTSIIYRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIKKDLAEE LTFDIDYKTSEVNQRVFSLDEVFEIANFNNYLNQSGITKFNTIIGGKF VNGENTKRKGINEYINLYSQQINDKTLKKYKMSVLFKQILSDTESKS FVIDKLEDDSDVVTTMQSFYEQIAAFKTVEEKSIKETLSLLFDDLKAQ KLDLSKIYFKNDKSLTDLSQQVFDDYSVIGTAVLEYITQQIAPKNLDN PSKKEQELIAKKTEKAKYLSLETIKLALEEFNKHRDIDKQCRFEEILA NFAAIPMIFDEIAQNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIKD LLDQTNNLLHKLKIFHISQSEDKANILDKDEHFYLVFEECYFELANIVP LYNKIRNYITQKPYSDEKFKLNFENSTLANGWDKNKEPDNTAILFIKD DKYYLGVMNKKNNKIFDDKAIKENKGEGYKKIVYKLLPGANKMLPK VFFSAKSIKFYNPSEDILRIRNHSTHTKNGSPQKGYEKFEFNIEDCR KFIDFYKQSISKHPEWKDFGFRFSDTQRYNSIDEFYREVENQGYKL TFENISESYIDSVVNQGKLYLFQIYNKDFSAYSKGRPNLHTLYWKAL FDERNLQDVVYKLNGEAELFYRKQSIPKKITHPAKEAIANKNKDNPK KESVFEYDLIKDKRFTEDKFFFHCPITINFKSSGANKFNDEINLLLKE KANDVHILSIDRGERHLAYYTLVDGKGNIIKQDTFNIIGNDRMKTNYH DKLAAIEKDRDSARKDWKKINNIKEMKEGYLSQVVHEIAKLVIEYNAI VVFEDLNFGFKRGRFKVEKQVYQKLEKMLIEKLNYLVFKDNEFDKT GGVLRAYQLTAPFETFKKMGKQTGIIYYVPAGFTSKICPVTGFVNQL YPKYESVSKSQEFFSKFDKICYNLDKGYFEFSFDYKNFGDKAAKGK WTIASFGSRLINFRNSDKNHNWDTREVYPTKELEKLLKDYSIEYGH GECIKAAICGESDKKFFAKLTSVLNTILQMRNSKTGTELDYLISPVAD VNGNFFDSRQAPKNMPQDADANGAYHIGLKGLMLLGRIKNNQEGK KLNLVIKNEEYFEFVQNRNN
Cpf1 D917A 폴리펩티드 서열	247	MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYK KAKQIIDKYHQFFIEEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQK DFKSAKDTIKKQISEYIKDSEKFKNLFNQNLIDAKKGQESDLILWLKQ SKDNGIELFKANSDITDIDEALEIIKSFKGWTTYFKGFHENRKNVYSS NDIPTSIIYRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIKKDLAEE LTFDIDYKTSEVNQRVFSLDEVFEIANFNNYLNQSGITKFNTIIGGKF VNGENTKRKGINEYINLYSQQINDKTLKKYKMSVLFKQILSDTESKS FVIDKLEDDSDVVTTMQSFYEQIAAFKTVEEKSIKETLSLLFDDLKAQ KLDLSKIYFKNDKSLTDLSQQVFDDYSVIGTAVLEYITQQIAPKNLDN PSKKEQELIAKKTEKAKYLSLETIKLALEEFNKHRDIDKQCRFEEILA NFAAIPMIFDEIAQNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIKD LLDQTNNLLHKLKIFHISQSEDKANILDKDEHFYLVFEECYFELANIVP LYNKIRNYITQKPYSDEKFKLNFENSTLANGWDKNKEPDNTAILFIKD DKYYLGVMNKKNNKIFDDKAIKENKGEGYKKIVYKLLPGANKMLPK VFFSAKSIKFYNPSEDILRIRNHSTHTKNGSPQKGYEKFEFNIEDCR KFIDFYKQSISKHPEWKDFGFRFSDTQRYNSIDEFYREVENQGYKL TFENISESYIDSVVNQGKLYLFQIYNKDFSAYSKGRPNLHTLYWKAL FDERNLQDVVYKLNGEAELFYRKQSIPKKITHPAKEAIANKNKDNPK KESVFEYDLIKDKRFTEDKFFFHCPITINFKSSGANKFNDEINLLLKE KANDVHILSIARGERHLAYYTLVDGKGNIIKQDTFNIIGNDRMKTNYH DKLAAIEKDRDSARKDWKKINNIKEMKEGYLSQVVHEIAKLVIEYNAI VVFEDLNFGFKRGRFKVEKQVYQKLEKMLIEKLNYLVFKDNEFDKT GGVLRAYQLTAPFETFKKMGKQTGIIYYVPAGFTSKICPVTGFVNQL YPKYESVSKSQEFFSKFDKICYNLDKGYFEFSFDYKNFGDKAAKGK WTIASFGSRLINFRNSDKNHNWDTREVYPTKELEKLLKDYSIEYGH GECIKAAICGESDKKFFAKLTSVLNTILQMRNSKTGTELDYLISPVAD VNGNFFDSRQAPKNMPQDADANGAYHIGLKGLMLLGRIKNNQEGK KLNLVIKNEEYFEFVQNRNN

[1087]

Cpf1 E1006A 폴리펩티드 서열	248	MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYK KAKQIIDKYHQFFIEEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQK DFKSAKDTIKKQISEYIKDSEKFKNLFNQNLIDAKKGQESDLILWLKQ SKDNGIELFKANSDITDIDEALEIIKSFKGWTTYFKGFHENRKNVYSS NDIPTSIIYRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIKKDLAEE LTFDIDYKTSEVNQRVFSLDEVFEIANFNNYLNQSGITKFNTIIGGKF VNGENTKRKGINEYINLYSQQINDKTLKKYKMSVLFKQILSDTESKS FVIDKLEDDSDVVTTMQSFYEQIAAFKTVEEKSIKETLSLLFDDLKAQ KLDLSKIYFKNDKSLTDLSQQVFDDYSVIGTAVLEYITQQIAPKNLDN PSKKEQELIAKKTEKAKYLSLETIKLALEEFNKHRDIDKQCRFEEILA NFAAIPMIFDEIAQNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIKD LLDQTNNLLHKLKIFHISQSEDKANILDKDEHFYLVFEECYFELANIVP LYNKIRNYITQKPYSDEKFKLNFENSTLANGWDKNKEPDNTAILFIKD DKYYLGVMNKKNNKIFDDKAIKENKGEGYKKIVYKLLPGANKMLPK VFFSAKSIKFYNPSEDILRIRNHSTHTKNGSPQKGYEKFEFNIEDCR KFIDFYKQSISKHPEWKDFGFRFSDTQRYNSIDEFYREVENQGYKL TFENISESYIDSVVNQGKLYLFQIYNKDFSAYSKGRPNLHTLYWKAL FDERNLQDVVYKLNGEAELFYRKQSIPKKITHPAKEAIANKNKDNPK KESVFEYDLIKDKRFTEDKFFFHCPITINFKSSGANKFNDEINLLLKE KANDVHILSIDRGERHLAYYTLVDGKGNIIKQDTFNIIGNDRMKTNYH DKLAAIEKDRDSARKDWKKINNIKEMKEGYLSQVVHEIAKLVIEYNAI VVFADLNFGFKRGRFKVEKQVYQKLEKMLIEKLNYLVFKDNEFDKT GGVLRAYQLTAPFETFKKMGKQTGIIYYVPAGFTSKICPVTGFVNQL YPKYESVSKSQEFFSKFDKICYNLDKGYFEFSFDYKNFGDKAAKGK WTIASFGSRLINFRNSDKNHNWDTREVYPTKELEKLLKDYSIEYGH GECIKAAICGESDKKFFAKLTSVLNTILQMRNSKTGTELDYLISPVAD VNGNFFDSRQAPKNMPQDADANGAYHIGLKGLMLLGRIKNNQEGK KLNLVIKNEEYFEFVQNRNN
Cpf1 D1255A 폴리펩티드 서열	249	MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYK KAKQIIDKYHQFFIEEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQK DFKSAKDTIKKQISEYIKDSEKFKNLFNQNLIDAKKGQESDLILWLKQ SKDNGIELFKANSDITDIDEALEIIKSFKGWTTYFKGFHENRKNVYSS NDIPTSIIYRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIKKDLAEE LTFDIDYKTSEVNQRVFSLDEVFEIANFNNYLNQSGITKFNTIIGGKF VNGENTKRKGINEYINLYSQQINDKTLKKYKMSVLFKQILSDTESKS FVIDKLEDDSDVVTTMQSFYEQIAAFKTVEEKSIKETLSLLFDDLKAQ KLDLSKIYFKNDKSLTDLSQQVFDDYSVIGTAVLEYITQQIAPKNLDN PSKKEQELIAKKTEKAKYLSLETIKLALEEFNKHRDIDKQCRFEEILA NFAAIPMIFDEIAQNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIKD LLDQTNNLLHKLKIFHISQSEDKANILDKDEHFYLVFEECYFELANIVP LYNKIRNYITQKPYSDEKFKLNFENSTLANGWDKNKEPDNTAILFIKD DKYYLGVMNKKNNKIFDDKAIKENKGEGYKKIVYKLLPGANKMLPK VFFSAKSIKFYNPSEDILRIRNHSTHTKNGSPQKGYEKFEFNIEDCR KFIDFYKQSISKHPEWKDFGFRFSDTQRYNSIDEFYREVENQGYKL TFENISESYIDSVVNQGKLYLFQIYNKDFSAYSKGRPNLHTLYWKAL FDERNLQDVVYKLNGEAELFYRKQSIPKKITHPAKEAIANKNKDNPK KESVFEYDLIKDKRFTEDKFFFHCPITINFKSSGANKFNDEINLLLKE KANDVHILSIDRGERHLAYYTLVDGKGNIIKQDTFNIIGNDRMKTNYH DKLAAIEKDRDSARKDWKKINNIKEMKEGYLSQVVHEIAKLVIEYNAI VVFEDLNFGFKRGRFKVEKQVYQKLEKMLIEKLNYLVFKDNEFDKT GGVLRAYQLTAPFETFKKMGKQTGIIYYVPAGFTSKICPVTGFVNQL YPKYESVSKSQEFFSKFDKICYNLDKGYFEFSFDYKNFGDKAAKGK WTIASFGSRLINFRNSDKNHNWDTREVYPTKELEKLLKDYSIEYGH GECIKAAICGESDKKFFAKLTSVLNTILQMRNSKTGTELDYLISPVAD VNGNFFDSRQAPKNMPQDAAANGAYHIGLKGLMLLGRIKNNQEGK KLNLVIKNEEYFEFVQNRNN
Cpf1 D917A/E1006A 폴리펩티드 서열	250	MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYK KAKQIIDKYHQFFIEEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQK DFKSAKDTIKKQISEYIKDSEKFKNLFNQNLIDAKKGQESDLILWLKQ

[1088]

		SKDNGIELFKANSDITDIDEALEIIKSFKGWTTYFKGFHENRKNVYSS NDIPTSIIYRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIKKDLAEE LTFDIDYKTSEVNQRVFSLDEVFEIANFNNYLNQSGITKFNTIIGGKF VNGENTKRKGINEYINLYSQQINDKTLKKYKMSVLFKQILSDTESKS FVIDKLEDDSDVVTTMQSFYEQIAAFKTVEEKSIKETLSLLFDDLKAQ KLDLSKIYFKNDKSLTDLSQQVFDDYSVIGTAVLEYITQQIAPKNLDN PSKKEQELIAKKTEKAKYLSLETIKLALEEFNKHRDIDKQCRFEEILA NFAAIPMIFDEIAQNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIKD LLDQTNNLLHKLKIFHISQSEDKANILDKDEHFYLVFEECYFELANIVP LYNKIRNYITQKPYSDEKFKLNFENSTLANGWDKNKEPDNTAILFIKD DKYYLGVMNKKNNKIFDDKAIKENKGEGYKKIVYKLLPGANKMLPK VFFSAKSIKFYNPSEDILRIRNHSTHTKNGSPQKGYEKFEFNIEDCR KFIDFYKQSISKHPEWKDFGFRFSDTQRYNSIDEFYREVENQGYKL TFENISESYIDSVVNQGKLYLFQIYNKDFSAYSKGRPNLHTLYWKAL FDERNLQDVVYKLNGEAELFYRKQSIPKKITHPAKEAIANKNKDNPK KESVFEYDLIKDKRFTEDKFFFHCPITINFKSSGANKFNDEINLLLKE KANDVHILSIARGERHLAYYTLVDGKGNIIKQDTFNIIGNDRMKTNYH DKLAAIEKDRDSARKDWKKINNIKEMKEGYLSQVVHEIAKLVIEYNAI VVFADLNFGFKRGRFKVEKQVYQKLEKMLIEKLNYLVFKDNEFDKT GGVLRAYQLTAPFETFKKMGKQTGIIYYVPAGFTSKICPVTGFVNQL YPKYESVSKSQEFFSKFDKICYNLDKGYFEFSFDYKNFGDKAAKGK WTIASFGSRLINFRNSDKNHNWDTREVYPTKELEKLLKDYSIEYGH GECIKAAICGESDKKFFAKLTSVLNTILQMRNSKTGTELDYLISPVAD VNGNFFDSRQAPKNMPQDADANGAYHIGLKGLMLLGRIKNNQEGK KLNLVIKNEEYFEFVQNRNN
Cpf1 D917A/D1255A 폴리펩티드 서열	251	MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYK KAKQIIDKYHQFFIEEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQK DFKSAKDTIKKQISEYIKDSEKFKNLFNQNLIDAKKGQESDLILWLKQ SKDNGIELFKANSDITDIDEALEIIKSFKGWTTYFKGFHENRKNVYSS NDIPTSIIYRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIKKDLAEE LTFDIDYKTSEVNQRVFSLDEVFEIANFNNYLNQSGITKFNTIIGGKF VNGENTKRKGINEYINLYSQQINDKTLKKYKMSVLFKQILSDTESKS FVIDKLEDDSDVVTTMQSFYEQIAAFKTVEEKSIKETLSLLFDDLKAQ KLDLSKIYFKNDKSLTDLSQQVFDDYSVIGTAVLEYITQQIAPKNLDN PSKKEQELIAKKTEKAKYLSLETIKLALEEFNKHRDIDKQCRFEEILA NFAAIPMIFDEIAQNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIKD LLDQTNNLLHKLKIFHISQSEDKANILDKDEHFYLVFEECYFELANIVP LYNKIRNYITQKPYSDEKFKLNFENSTLANGWDKNKEPDNTAILFIKD DKYYLGVMNKKNNKIFDDKAIKENKGEGYKKIVYKLLPGANKMLPK VFFSAKSIKFYNPSEDILRIRNHSTHTKNGSPQKGYEKFEFNIEDCR KFIDFYKQSISKHPEWKDFGFRFSDTQRYNSIDEFYREVENQGYKL TFENISESYIDSVVNQGKLYLFQIYNKDFSAYSKGRPNLHTLYWKAL FDERNLQDVVYKLNGEAELFYRKQSIPKKITHPAKEAIANKNKDNPK KESVFEYDLIKDKRFTEDKFFFHCPITINFKSSGANKFNDEINLLLKE KANDVHILSIARGERHLAYYTLVDGKGNIIKQDTFNIIGNDRMKTNYH DKLAAIEKDRDSARKDWKKINNIKEMKEGYLSQVVHEIAKLVIEYNAI VVFEDLNFGFKRGRFKVEKQVYQKLEKMLIEKLNYLVFKDNEFDKT GGVLRAYQLTAPFETFKKMGKQTGIIYYVPAGFTSKICPVTGFVNQL YPKYESVSKSQEFFSKFDKICYNLDKGYFEFSFDYKNFGDKAAKGK WTIASFGSRLINFRNSDKNHNWDTREVYPTKELEKLLKDYSIEYGH GECIKAAICGESDKKFFAKLTSVLNTILQMRNSKTGTELDYLISPVAD VNGNFFDSRQAPKNMPQDAAANGAYHIGLKGLMLLGRIKNNQEGK KLNLVIKNEEYFEFVQNRNN
Cpf1 E1006A/D1255A 폴리펩티드 서열	252	MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYK KAKQIIDKYHQFFIEEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQK DFKSAKDTIKKQISEYIKDSEKFKNLFNQNLIDAKKGQESDLILWLKQ SKDNGIELFKANSDITDIDEALEIIKSFKGWTTYFKGFHENRKNVYSS NDIPTSIIYRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIKKDLAEE LTFDIDYKTSEVNQRVFSLDEVFEIANFNNYLNQSGITKFNTIIGGKF

[1089]

		VNGENTKRKGINEYINLYSQQINDKTLKKYKMSVLFKQILSDTESKS FVIDKLEDDSDVVTTMQSFYEQIAAFKTVEEKSIKETLSLLFDDLKAQ KLDLSKIYFKNDKSLTDLSQQVFDDYSVIGTAVLEYITQQIAPKNLDN PSKKEQELIAKKTEKAKYLSLETIKLALEEFNKHRDIDKQCRFEEILA NFAAIPMIFDEIAQNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIKD LLDQTNNLLHKLKIFHISQSEDKANILDKDEHFYLVFEECYFELANIVP LYNKIRNYITQKPYSDEKFKLNFENSTLANGWDKNKEPDNTAILFIKD DKYYLGVMNKKNNKIFDDKAIKENKGEGYKKIVYKLLPGANKMLPK VFFSAKSIKFYNPSEDILRIRNHSTHTKNGSPQKGYEKFEFNIEDCR KFIDFYKQSISKHPEWKDFGFRFSDTQRYNSIDEFYREVENQGYKL TFENISESYIDSVVNQGKLYLFQIYNKDFSAYSKGRPNLHTLYWKAL FDERNLQDVVYKLNGEAELFYRKQSIPKKITHPAKEAIANKNKDNPK KESVFEYDLIKDKRFTEDKFFFHCPITINFKSSGANKFNDEINLLLKE KANDVHILSIDRGERHLAYYTLVDGKGNIIKQDTFNIIGNDRMKTNYH DKLAAIEKDRDSARKDWKKINNIKEMKEGYLSQVVHEIAKLVIEYNAI VVFADLNFGFKRGRFKVEKQVYQKLEKMLIEKLNYLVFKDNEFDKT GGVLRAYQLTAPFETFKKMGKQTGIIYYVPAGFTSKICPVTGFVNQL YPKYESVSKSQEFFSKFDKICYNLDKGYFEFSFDYKNFGDKAAKGK WTIASFGSRLINFRNSDKNHNWDTREVYPTKELEKLLKDYSIEYGH GECIKAAICGESDKKFFAKLTSVLNTILQMRNSKTGTELDYLISPVAD VNGNFFDSRQAPKNMPQDAAANGAYHIGLKGLMLLGRIKNNQEGK KLNLVIKNEEYFEFVQNRNN
Cpf1 D917A/E1006A/D1 255A 폴리펩티드 서열	253	MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYK KAKQIIDKYHQFFIEEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQK DFKSAKDTIKKQISEYIKDSEKFKNLFNQNLIDAKKGQESDLILWLKQ SKDNGIELFKANSDITDIDEALEIIKSFKGWTTYFKGFHENRKNVYSS NDIPTSIIYRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIKKDLAEE LTFDIDYKTSEVNQRVFSLDEVFEIANFNNYLNQSGITKFNTIIGGKF VNGENTKRKGINEYINLYSQQINDKTLKKYKMSVLFKQILSDTESKS FVIDKLEDDSDVVTTMQSFYEQIAAFKTVEEKSIKETLSLLFDDLKAQ KLDLSKIYFKNDKSLTDLSQQVFDDYSVIGTAVLEYITQQIAPKNLDN PSKKEQELIAKKTEKAKYLSLETIKLALEEFNKHRDIDKQCRFEEILA NFAAIPMIFDEIAQNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIKD LLDQTNNLLHKLKIFHISQSEDKANILDKDEHFYLVFEECYFELANIVP LYNKIRNYITQKPYSDEKFKLNFENSTLANGWDKNKEPDNTAILFIKD DKYYLGVMNKKNNKIFDDKAIKENKGEGYKKIVYKLLPGANKMLPK VFFSAKSIKFYNPSEDILRIRNHSTHTKNGSPQKGYEKFEFNIEDCR KFIDFYKQSISKHPEWKDFGFRFSDTQRYNSIDEFYREVENQGYKL TFENISESYIDSVVNQGKLYLFQIYNKDFSAYSKGRPNLHTLYWKAL FDERNLQDVVYKLNGEAELFYRKQSIPKKITHPAKEAIANKNKDNPK KESVFEYDLIKDKRFTEDKFFFHCPITINFKSSGANKFNDEINLLLKE KANDVHILSIARGERHLAYYTLVDGKGNIIKQDTFNIIGNDRMKTNYH DKLAAIEKDRDSARKDWKKINNIKEMKEGYLSQVVHEIAKLVIEYNAI VVFADLNFGFKRGRFKVEKQVYQKLEKMLIEKLNYLVFKDNEFDKT GGVLRAYQLTAPFETFKKMGKQTGIIYYVPAGFTSKICPVTGFVNQL YPKYESVSKSQEFFSKFDKICYNLDKGYFEFSFDYKNFGDKAAKGK WTIASFGSRLINFRNSDKNHNWDTREVYPTKELEKLLKDYSIEYGH GECIKAAICGESDKKFFAKLTSVLNTILQMRNSKTGTELDYLISPVAD VNGNFFDSRQAPKNMPQDAAANGAYHIGLKGLMLLGRIKNNQEGK KLNLVIKNEEYFEFVQNRNN
합성 폴리펩티드	254	KRNYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRS KRGARRLKRRRRHRIQRVKKLLFDYNLLTDHSELSGINPYEARVKG LSQKLSEEEFSAALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRN SKALEEKYVAELQLERLKKDGEVRGSINRFKTSDYVKEAKQLLKVQ KAYHQLDQSFIDTYIDLLETRRTYYEGPGEGSPFGWKDIKEWYEML MGHCTYFPEELRSVKYAYNADLYNALNDLNNLVITRDENEKLEYYE KFQIIENVFKQKKKPTLKQIAKEILVNEEDIKGYRVTSTGKPEFTNLKV YHDIKDITARKEIIENAELLDQIAKILTIYQSSEDIQEELTNLNSELTQE EIEQISNLKGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKLVPKK

[1090]

		VDLSQQKEIPTTLVDDFILSPVVKRSFIQSIKVINAIIKKYGLPNDIIIELA REKNSKDAQKMINEMQKRNRQTNERIEEIIRTTGKENAKYLIEKIKLH DMQEGKCLYSLEAIPLEDLLNNPFNYEVDHIIPRSVSFDNSFNNKVL VKQEENSKKGNRTPFQYLSSSDSKISYETFKKHILNLAKGKGRISKT KKEYLLEERDINRFSVQKDFINRNLVDTRYATRGLMNLLRSYFRVN NLDVKVKSINGGFTSFLRRKWKFKKERNKGYKHHAEDALIIANADFI FKEWKKLDKAKKVMENQMFEEKQAESMPEIETEQEYKEIFITPHQIK HIKDFKDYKYSHRVDKKPNRELINDTLYSTRKDDKGNTLIVNNLNGL YDKDNDKLKKLINKSPEKLLMYHHDPQTYQKLKLIMEQYGDEKNPL YKYYEETGNYLTKYSKKDNGPVIKKIKYYGNKLNAHLDITDDYPNSR NKVVKLSLKPYRFDVYLDNGVYKFVTVKNLDVIKKENYYEVNSKCY EEAKKLKKISNQAEFIASFYNNDLIKINGELYRVIGVNNDLLNRIEVNM IDITYREYLENMNDKRPPRIIKTIASKTQSIKKYSTDILGNLYEVKSKK HPQIIKKG
SaCas9n 폴리펩티드 서열	255	KRNYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRS KRGARRLKRRRRHRIQRVKKLLFDYNLLTDHSELSGINPYEARVKG LSQKLSEEEFSAALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRN SKALEEKYVAELQLERLKKDGEVRGSINRFKTSDYVKEAKQLLKVQ KAYHQLDQSFIDTYIDLLETRRTYYEGPGEGSPFGWKDIKEWYEML MGHCTYFPEELRSVKYAYNADLYNALNDLNNLVITRDENEKLEYYE KFQIIENVFKQKKKPTLKQIAKEILVNEEDIKGYRVTSTGKPEFTNLKV YHDIKDITARKEIIENAELLDQIAKILTIYQSSEDIQEELTNLNSELTQE EIEQISNLKGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKLVPKK VDLSQQKEIPTTLVDDFILSPVVKRSFIQSIKVINAIIKKYGLPNDIIIELA REKNSKDAQKMINEMQKRNRQTNERIEEIIRTTGKENAKYLIEKIKLH DMQEGKCLYSLEAIPLEDLLNNPFNYEVDHIIPRSVSFDNSFNNKVL VKQEEASKKGNRTPFQYLSSSDSKISYETFKKHILNLAKGKGRISKT KKEYLLEERDINRFSVQKDFINRNLVDTRYATRGLMNLLRSYFRVN NLDVKVKSINGGFTSFLRRKWKFKKERNKGYKHHAEDALIIANADFI FKEWKKLDKAKKVMENQMFEEKQAESMPEIETEQEYKEIFITPHQIK HIKDFKDYKYSHRVDKKPNRELINDTLYSTRKDDKGNTLIVNNLNGL YDKDNDKLKKLINKSPEKLLMYHHDPQTYQKLKLIMEQYGDEKNPL YKYYEETGNYLTKYSKKDNGPVIKKIKYYGNKLNAHLDITDDYPNSR NKVVKLSLKPYRFDVYLDNGVYKFVTVKNLDVIKKENYYEVNSKCY EEAKKLKKISNQAEFIASFYNNDLIKINGELYRVIGVNNDLLNRIEVNM IDITYREYLENMNDKRPPRIIKTIASKTQSIKKYSTDILGNLYEVKSKK HPQIIKKG
SaKKH Cas9 폴리펩티드 서열	**256**	KRNYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRS KRGARRLKRRRRHRIQRVKKLLFDYNLLTDHSELSGINPYEARVKG LSQKLSEEEFSAALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRN SKALEEKYVAELQLERLKKDGEVRGSINRFKTSDYVKEAKQLLKVQ KAYHQLDQSFIDTYIDLLETRRTYYEGPGEGSPFGWKDIKEWYEML MGHCTYFPEELRSVKYAYNADLYNALNDLNNLVITRDENEKLEYYE KFQIIENVFKQKKKPTLKQIAKEILVNEEDIKGYRVTSTGKPEFTNLKV YHDIKDITARKEIIENAELLDQIAKILTIYQSSEDIQEELTNLNSELTQE EIEQISNLKGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKLVPKK VDLSQQKEIPTTLVDDFILSPVVKRSFIQSIKVINAIIKKYGLPNDIIIELA REKNSKDAQKMINEMQKRNRQTNERIEEIIRTTGKENAKYLIEKIKLH DMQEGKCLYSLEAIPLEDLLNNPFNYEVDHIIPRSVSFDNSFNNKVL VKQEEASKKGNRTPFQYLSSSDSKISYETFKKHILNLAKGKGRISKT KKEYLLEERDINRFSVQKDFINRNLVDTRYATRGLMNLLRSYFRVN NLDVKVKSINGGFTSFLRRKWKFKKERNKGYKHHAEDALIIANADFI FKEWKKLDKAKKVMENQMFEEKQAESMPEIETEQEYKEIFITPHQIK HIKDFKDYKYSHRVDKKPNRKLINDTLYSTRKDDKGNTLIVNNLNGL YDKDNDKLKKLINKSPEKLLMYHHDPQTYQKLKLIMEQYGDEKNPL YKYYEETGNYLTKYSKKDNGPVIKKIKYYGNKLNAHLDITDDYPNSR NKVVKLSLKPYRFDVYLDNGVYKFVTVKNLDVIKKENYYEVNSKCY EEAKKLKKISNQAEFIASFYKNDLIKINGELYRVIGVNNDLLNRIEVNM

[1091]

		IDITYREYLENMNDKRPPHIIKTIASKTQSIKKYSTDILGNLYEVKSKK HPQIIKKG
Casphi-1 폴리펩티드 서열	285	MADTPTLFTQFLRHHLPGQRFRKDILKQAGRILANKGEDATIAFLRG KSEESPPDFQPPVKCPIIACSRPLTEWPIYQASVAIQGYVYGQSLAE FEASDPGCSKDGLLGWFDKTGVCTDYFSVQGLNLIFQNARKRYIGV QTKVTNRNEKRHKKLKRINAKRIAEGLPELTSDEPESALDETGHLID PPGLNTNIYCYQQVSPKPLALSEVNQLPTAYAGYSTSGDDPIQPMV TKDRLSISKGQPGYIPEHQRALLSQKKHRRMRGYGLKARALLVIVRI QDDWAVIDLRSLLRNAYWRRIVQTKEPSTITKLLKLVTGDPVLDATR MVATFTYKPGIVQVRSAKCLKNKQGSKLFSERYLNETVSVTSIDLGS NNLVAVATYRLVNGNTPELLQRFTLPSHLVKDFERYKQAHDTLEDSI QKTAVASLPQGQQTEIRMWSMYGFREAQERVCQELGLADGSIPW NVMTATSTILTDLFLARGGDPKKCMFTSEPKKKKNSKQVLYKIRDR AWAKMYRTLLSKETREAWNKALWGLKRGSPDYARLSKRKEELAR RCVNYTISTAEKRAQCGRTIVALEDLNIGFFHGRGKQEPGWVGLFT RKKENRWLMQALHKAFLELAHHRGYHVIEVNPAYTSQTCPVCRHC DPDNRDQHNREAFHCIGCGFRGNADLDVATHNIAMVAITGESLKRA RGSVASKTPQPLAAE
Casphi-2 폴리펩티드 서열	286	MPKPAVESEFSKVLKKHFPGERFRSSYMKRGGKILAAQGEEAVVA YLQGKSEEEPPNFQPPAKCHVVTKSRDFAEWPIMKASEAIQRYIYA LSTTERAACKPGKSSESHAAWFAATGVSNHGYSHVQGLNLIFDHTL GRYDGVLKKVQLRNEKARARLESINASRADEGLPEIKAEEEEVATN ETGHLLQPPGINPSFYVYQTISPQAYRPRDEIVLPPEYAGYVRDPNA PIPLGVVRNRCDIQKGCPGYIPEWQREAGTAISPKTGKAVTVPGLS PKKNKRMRRYWRSEKEKAQDALLVTVRIGTDWVVIDVRGLLRNAR WRTIAPKDISLNALLDLFTGDPVIDVRRNIVTFTYTLDACGTYARKWT LKGKQTKATLDKLTATQTVALVAIDLGQTNPISAGISRVTQENGALQ CEPLDRFTLPDDLLKDISAYRIAWDRNEEELRARSVEALPEAQQAE VRALDGVSKETARTQLCADFGLDPKRLPWDKMSSNTTFISEALLSN SVSRDQVFFTPAPKKGAKKKAPVEVMRKDRTWARAYKPRLSVEAQ KLKNEALWALKRTSPEYLKLSRRKEELCRRSINYVIEKTRRRTQCQI VIPVIEDLNVRFFHGSGKRLPGWDNFFTAKKENRWFIQGLHKAFSD LRTHRSFYVFEVRPERTSITCPKCGHCEVGNRDGEAFQCLSCGKT CNADLDVATHNLTQVALTGKTMPKREEPRDAQGTAPARKTKKASK SKAPPAEREDQTPAQEPSQTS
Casphi-3 폴리펩티드 서열	287	MEKEITELTKIRREFPNKKFSSTDMKKAGKLLKAEGPDAVRDFLNSC QEIIGDFKPPVKTNIVSISRPFEEWPVSMVGRAIQEYYFSLTKEELES VHPGTSSEDHKSFFNITGLSNYNYTSVQGLNLIFKNAKAIYDGTLVK ANNKNKKLEKKFNEINHKRSLEGLPIITPDFEEPFDENGHLNNPPGI NRNIYGYQGCAAKVFVPSKHKMVSLPKEYEGYNRDPNLSLAGFRN RLEIPEGEPGHVPWFQRMDIPEGQIGHVNKIQRFNFVHGKNSGKVK FSDKTGRVKRYHHSKYKDATKPYKFLEESKKVSALDSILAIITIGDDW VVFDIRGLYRNVFYRELAQKGLTAVQLLDLFTGDPVIDPKKGVVTFS YKEGVVPVFSQKIVPRFKSRDTLEKLTSQGPVALLSVDLGQNEPVA ARVCSLKNINDKITLDNSCRISFLDDYKKQIKDYRDSLDELEIKIRLEAI NSLETNQQVEIRDLDVFSADRAKANTVDMFDIDPNLISWDSMSDAR VSTQISDLYLKNGGDESRVYFEINNKRIKRSDYNISQLVRPKLSDST RKNLNDSIWKLKRTSEEYLKLSKRKLELSRAVVNYTIRQSKLLSGIN DIVIILEDLDVKKKFNGRGIRDIGWDNFFSSRKENRWFIPAFHKAFSE LSSNRGLCVIEVNPAWTSATCPDCGFCSKENRDGINFTCRKCGVS YHADIDVATLNIARVAVLGKPMSGPADRERLGDTKKPRVARSRKTM KRKDISNSTVEAMVTA
Casphi-4 폴리펩티드 서열	288	MYSLEMADLKSEPSLLAKLLRDRFPGKYWLPKYWKLAEKKRLTGG EEAACEYMADKQLDSPPPNFRPPARCVILAKSRPFEDWPVHRVAS KAQSFVIGLSEQGFAALRAAPPSTADARRDWLRSHGASEDDLMAL EAQLLETIMGNAISLHGGVLKKIDNANVKAAKRLSGRNEARLNKGLQ ELPPEQEGSAYGADGLLVNPPGLNLNIYCRKSCCPKPVKNTARFVG HYPGYLRDSDSILISGTMDRLTIIEGMPGHIPAWQREQGLVKPGGR RRRLSGSESNMRQKVDPSTGPRRSTRSGTVNRSNQRTGRNGDPL

[1092]

		LVEIRMKEDWVLLDARGLLRNLRWRESKRGLSCDHEDLSLSGLLAL FSGDPVIDPVRNEVVFLYGEGIIPVRSTKPVGTRQSKKLLERQASM GPLTLISCDLGQTNLIAGRASAISLTHGSLGVRSSVRIELDPEIIKSFE RLRKDADRLETEILTAAKETLSDEQRGEVNSHEKDSPQTAKASLCR ELGLHPPSLPWGQMGPSTTFIADMLISHGRDDDAFLSHGEFPTLEK RKKFDKRFCLESRPLLSSETRKALNESLWEVKRTSSEYARLSQRKK EMARRAVNFVVEISRRKTGLSNVIVNIEDLNVRIFHGGGKQAPGWD GFFRPKSENRWFIQAIHKAFSDLAAHHGIPVIESDPQRTSMTCPEC GHCDSKNRNGVRFLCKGCGASMDADFDAACRNLERVALTGKPMP KPSTSCERLLSATTGKVCSDHSLSHDAIEKAS
Casphi-5 폴리펩티드 서열	289	MSSLPTPLELLKQKHADLFKGLQFSSKDNKMAGKVLKKDGEEAALA FLSERGVSRGELPNFRPPAKTLVVAQSRPFEEFPIYRVSEAIQLYVY SLSVKELETVPSGSSTKKEHQRFFQDSSVPDFGYTSVQGLNKIFGL ARGIYLGVITRGENQLQKAKSKHEALNKKRRASGEAETEFDPTPYE YMTPERKLAKPPGVNHSIMCYVDISVDEFDFRNPDGIVLPSEYAGY CREINTAIEKGTVDRLGHLKGGPGYIPGHQRKESTTEGPKINFRKG RIRRSYTALYAKRDSRRVRQGKLALPSYRHHMMRLNSNAESAILAV IFFGKDWVVFDLRGLLRNVRWRNLFVDGSTPSTLLGMFGDPVIDPK RGVVAFCYKEQIVPVVSKSITKMVKAPELLNKLYLKSEDPLVLVAIDL GQTNPVGVGVYRVMNASLDYEVVTRFALESELLREIESYRQRTNAF EAQIRAETFDAMTSEEQEEITRVRAFSASKAKENVCHRFGMPVDAV DWATMGSNTIHIAKWVMRHGDPSLVEVLEYRKDNEIKLDKNGVPK KVKLTDKRIANLTSIRLRFSQETSKHYNDTMWELRRKHPVYQKLSK SKADFSRRVVNSIIRRVNHLVPRARIVFIIEDLKNLGKVFHGSGKREL GWDSYFEPKSENRWFIQVLHKAFSETGKHKGYYIIECWPNWTSCT CPKCSCCDSENRHGEVFRCLACGYTCNTDFGTAPDNLVKIATTGK GLPGPKKRCKGSSKGKNPKIARSSETGVSVTESGAPKVKKSSPTQ TSQSSSQSAP
Casphi-6 폴리펩티드 서열	290	MNKIEKEKTPLAKLMNENFAGLRFPFAIIKQAGKKLLKEGELKTIEYM TGKGSIEPLPNFKPPVKCLIVAKRRDLKYFPICKASCEIQSYVYSLNY KDFMDYFSTPMTSQKQHEEFFKKSGLNIEYQNVAGLNLIFNNVKNT YNGVILKVKNRNEKLKKKAIKNNYEFEEIKTFNDDGCLINKPGINNVI YCFQSISPKILKNITHLPKEYNDYDCSVDRNIIQKYVSRLDIPESQPG HVPEWQRKLPEFNNTNNPRRRRKWYSNGRNISKGYSVDQVNQAK IEDSLLAQIKIGEDWIILDIRGLLRDLNRRELISYKNKLTIKDVLGFFSD YPIIDIKKNLVTFCYKEGVIQVVSQKSIGNKKSKQLLEKLIENKPIALVS IDLGQTNPVSVKISKLNKINNKISIESFTYRFLNEEILKEIEKYRKDYDK LELKLINEA
Casphi-7 폴리펩티드 서열	291	MSNTAVSTREHMSNKTTPPSPLSLLLRAHFPGLKFESQDYKIAGKK LRDGGPEAVISYLTGKGQAKLKDVKPPAKAFVIAQSRPFIEWDLVRV SRQIQEKIFGIPATKGRPKQDGLSETAFNEAVASLEVDGKSKLNEET RAAFYEVLGLDAPSLHAQAQNALIKSAISIREGVLKKVENRNEKNLS KTKRRKEAGEEATFVEEKAHDERGYLIHPPGVNQTIPGYQAVVIKS CPSDFIGLPSGCLAKESAEALTDYLPHDRMTIPKGQPGYVPEWQHP LLNRRKNRRRRDWYSASLNKPKATCSKRSGTPNRKNSRTDQIQSG RFKGAIPVLMRFQDEWVIIDIRGLLRNARYRKLLKEKSTIPDLLSLFT GDPSIDMRQGVCTFIYKAGQACSAKMVKTKNAPEILSELTKSGPVV LVSIDLGQTNPIAAKVSRVTQLSDGQLSHETLLRELLSNDSSDGKEI ARYRVASDRLRDKLANLAVERLSPEHKSEILRAKNDTPALCKARVC AALGLNPEMIAWDKMTPYTEFLATAYLEKGGDRKVATLKPKNRPE MLRRDIKFKGTEGVRIEVSPEAAEAYREAQWDLQRTSPEYLRLSTW KQELTKRILNQLRHKAAKSSQCEVVVMAFEDLNIKMMHGNGKWAD GGWDAFFIKKRENRWFMQAFHKSLTELGAHKGVPTIEVTPHRTSIT CTKCGHCDKANRDGERFACQKCGFVAHADLEIATDNIERVALTGKP MPKPESERSGDAKKSVGARKAAFKPEEDAEAAE
Casphi-8 폴리펩티드 서열	292	MIKPTVSQFLTPGFKLIRNHSRTAGLKLKNEGEEACKKFVRENEIPK DECPNFQGGPAIANIIAKSREFTEWEIYQSSLAIQEVIFTLPKDKLPE PILKEEWRAQWLSEHGLDTVPYKEAAGLNLIIKNAVNTYKGVQVKV DNKNKNNLAKINRKNEIAKLNGEQEISFEEIKAFDDKGYLLQKPSPN

[1093]

		KSIYCYQSVSPKPFITSKYHNVNLPEEYIGYYRKSNEPIVSPYQFDRL RIPIGEPGYVPKWQYTFLSKKENKRRKLSKRIKNVSPILGIICIKKDW CVFDMRGLLRTNHWKKYHKPTDSINDLFDYFTGDPVIDTKANVVRF RYKMENGIVNYKPVREKKGKELLENICDQNGSCKLATVDVGQNNP VAIGLFELKKVNGELTKTLISRHPTPIDFCNKITAYRERYDKLESSIKL DAIKQLTSEQKIEVDNYNNNFTPQNTKQIVCSKLNINPNDLPWDKMI SGTHFISEKAQVSNKSEIYFTSTDKGKTKDVMKSDYKWFQDYKPKL SKEVRDALSDIEWRLRRESLEFNKLSKSREQDARQLANWISSMCD VIGIENLVKKNNFFGGSGKREPGWDNFYKPKKENRWWINAIHKALT ELSQNKGKRVILLPAMRTSITCPKCKYCDSKNRNGEKFNCLKCGIEL NADIDVATENLATVAITAQSMPKPTCERSGDAKKPVRARKAKAPEF HDKLAPSYTVVLREAV
Casphi-9 폴리펩티드 서열	293	MRSSREIGDKILMRQPAEKTAFQVFRQEVIGTQKLSGGDAKTAGRL YKQGKMEAAREWLLKGARDDVPPNFQPPAKCLVVAVSHPFEEWDI SKTNHDVQAYIYAQPLQAEGHLNGLSEKWEDTSADQHKLWFEKTG VPDRGLPVQAINKIAKAAVNRAFGVVRKVENRNEKRRSRDNRIAEH NRENGLTEVVREAPEVATNADGFLLHPPGIDPSILSYASVSPVPYNS SKHSFVRLPEEYQAYNVEPDAPIPQFVVEDRFAIPPGQPGYVPEW QRLKCSTNKHRRMRQWSNQDYKPKAGRRAKPLEFQAHLTRERAK GALLVVMRIKEDWVVFDVRGLLRNVEWRKVLSEEAREKLTLKGLLD LFTGDPVIDTKRGIVTFLYKAEITKILSKRTVKTKNARDLLLRLTEPGE DGLRREVGLVAVDLGQTHPIAAAIYRIGRTSAGALESTVLHRQGLRE DQKEKLKEYRKRHTALDSRLRKEAFETLSVEQQKEIVTVSGSGAQI TKDKVCNYLGVDPSTLPWEKMGSYTHFISDDFLRRGGDPNIVHFD RQPKKGKVSKKSQRIKRSDSQWVGRMRPRLSQETAKARMEADWA AQNENEEYKRLARSKQELARWCVNTLLQNTRCITQCDEIVVVIEDL NVKSLHGKGAREPGWDNFFTPKTENRWFIQILHKTFSELPKHRGE HVIEGCPLRTSITCPACSYCDKNSRNGEKFVCVACGATFHADFEVA TYNLVRLATTGMPMPKSLERQGGGEKAGGARKARKKAKQVEKIVV QANANVTMNGASLHSP
Casphi-10 폴리펩티드 서열	294	MDMLDTETNYATETPAQQQDYSPKPPKKAQRAPKGFSKKARPEKK PPKPITLFTQKHFSGVRFLKRVIRDASKILKLSESRTITFLEQAIERDG SAPPDVTPPVHNTIMAVTRPFEEWPEVILSKALQKHCYALTKKIKIKT WPKKGPGKKCLAAWSARTKIPLIPGQVQATNGLFDRIGSIYDGVEK KVTNRNANKKLEYDEAIKEGRNPAVPEYETAYNIDGTLINKPGYNPN LYITQSRTPRLITEADRPLVEKILWQMVEKKTQSRNQARRARLEKAA HLQGLPVPKFVPEKVDRSQKIEIRIIDPLDKIEPYMPQDRMAIKASQD GHVPYWQRPFLSKRRNRRVRAGWGKQVSSIQAWLTGALLVIVRLG NEAFLADIRGALRNAQWRKLLKPDATYQSLFNLFTGDPVVNTRTNH LTMAYREGVVNIVKSRSFKGRQTREHLLTLLGQGKTVAGVSFDLGQ KHAAGLLAAHFGLGEDGNPVFTPIQACFLPQRYLDSLTNYRNRYDA LTLDMRRQSLLALTPAQQQEFADAQRDPGGQAKRACCLKLNLNPD EIRWDLVSGISTMISDLYIERGGDPRDVHQQVETKPKGKRKSEIRIL KIRDGKWAYDFRPKIADETRKAQREQLWKLQKASSEFERLSRYKINI ARAIANWALQWGRELSGCDIVIPVLEDLNVGSKFFDGKGKWLLGW DNRFTPKKENRWFIKVLHKAVAELAPHRGVPVYEVMPHRTSMTCP ACHYCHPTNREGDRFECQSCHVVKNTDRDVAPYNILRVAVEGKTL DRWQAEKKPQAEPDRPMILIDNQES
>splP14739IUNGI_ BPPB2 우라실- DNA 글리코실라제 억제제	106	MTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDE STDENVMLLTSD APEYKPWALVIQDSNGENKIKML
Cas12b/C2c1	258	MAVKSIKVKLRLDDMPEIRAGLWKLHKEVNAGVRYYTEWLSLLRQE NLYRRSPNGDGEQECDKTAEECKAELLERLRARQVENGHRGPAG SDDELLQLARQLYELLVPQAIGAKGDAQQIARKFLSPLADKDAVGG LGIAKAGNKPRWVRMREAGEPGWEEEKEKAETRKSADRTADVLR ALADFGLKPLMRVYTDSEMSSVEWKPLRKGQAVRTWDRDMFQQA IERMMSWESWNQRVGQEYAKLVEQKNRFEQKNFVGQEHLVHLVN

[1094]

		QLQQDMKEASPGLESKEQTAHYVTGRALRGSDKVFEKWGKLAPD APFDLYDAEIKNVQRRNTRRFGSHDLFAKLAEPEYQALWREDASFL TRYAVYNSILRKLNHAKMFATFTLPDATAHPIWTRFDKLGGNLHQY TFLFNEFGERRHAIRFHKLLKVENGVAREVDDVTVPISMSEQLDNLL PRDPNEPIALYFRDYGAEQHFTGEFGGAKIQCRRDQLAHMHRRRG ARDVYLNVSVRVQSQSEARGERRPPYAAVFRLVGDNHRAFVHFDK LSDYLAEHPDDGKLGSEGLLSGLRVMSVDLGLRTSASISVFRVARK DELKPNSKGRVPFFFPIKGNDNLVAVHERSQLLKLPGETESKDLRAI REERQRTLRQLRTQLAYLRLLVRCGSEDVGRRERSWAKLIEQPVD AANHMTPDWREAFENELQKLKSLHGICSDKEWMDAVYESVRRVW RHMGKQVRDWRKDVRSGERPKIRGYAKDVVGGNSIEQIEYLERQY KFLKSWSFFGKVSGQVIRAEKGSRFAITLREHIDHAKEDRLKKLADR IIMEALGYVYALDERGKGKWVAKYPPCQLILEELSEYQFNNDRPPS ENNQLMQWSHRGVFQELINQAQVHDLLVGTMYAAFSSRFDARTG APGIRCRRVPARCTQEHNPEPFPWWLNKFVVEHTLDACPLRADDLI PTGEGEIFVSPFSAEEGDFHQIHADLNAAQNLQQRLWSDFDISQIRL RCDWGEVDGELVLIPRLTGKRTADSYSNKVFYTNTGVTYYERERG KKRRKVFAQEKLSEEEAELLVEADEAREKSVVLMRDPSGIINRGNW TRQKEFWSMV NQRIEGYLVKQIRSRVPLQDSACENTGDI
높은 충실도 Cas9 폴리펩티드 서열	1423	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT PWNFEEVVDKGASAQSFIERMTAFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED ILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGA LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMALIHDDSLTFKEDIQ KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHK PENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVE NTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQ RKFDNLTKAERGGLSELDKAGFIKRQLVETRAITKHVAQILDSRMNT KYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAY LNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAK YFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATV RKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK KYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEK NPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKG NELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEII EQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLG APAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGG D
Wt Cas9 도메인	233	ATGGATAAAAAGTATTCTATTGGTTTAGACATCGGCACTAATTCC GTTGGATGGGCTGTCATAACCGATGAATACAAAGTACCTTCAAA GAAATTTAAGGTGTTGGGGAACACAGACCGTCATTCGATTAAAA AGAATCTTATCGGTGCCCTCCTATTCGATAGTGGCGAAACGGCA GAGGCGACTCGCCTGAAACGAACCGCTCGGAGAAGGTATACAC GTCGCAAGAACCGAATATGTTACTTACAAGAAATTTTTAGCAATG AGATGGCCAAAGTTGACGATTCTTTCTTTCACCGTTTGGAAGAGT CCTTCCTTGTCGAAGAGGACAAGAAACATGAACGGCACCCCATC TTTGGAAACATAGTAGATGAGGTGGCATATCATGAAAAGTACCC AACGATTTATCACCTCAGAAAAAAGCTAGTTGACTCAACTGATAA

[1095]

[1096]

		AGCGGACCTGAGGTTAATCTACTTGGCTCTTGCCCATATGATAA AGTTCCGTGGGCACTTTCTCATTGAGGGTGATCTAAATCCGGAC AACTCGGATGTCGACAAACTGTTCATCCAGTTAGTACAAACCTAT AATCAGTTGTTTGAAGAGAACCCTATAAATGCAAGTGGCGTGGA TGCGAAGGCTATTCTTAGCGCCCGCCTCTCTAAATCCCGACGGC TAGAAAACCTGATCGCACAATTACCCGGAGAGAAGAAAAATGGG TTGTTCGGTAACCTTATAGCGCTCTCACTAGGCCTGACACCAAA TTTTAAGTCGAACTTCGACTTAGCTGAAGATGCCAAATTGCAGCT TAGTAAGGACACGTACGATGACGATCTCGACAATCTACTGGCAC AAATTGGAGATCAGTATGCGGACTTATTTTTGGCTGCCAAAAACC TTAGCGATGCAATCCTCCTATCTGACATACTGAGAGTTAATACTG AGATTACCAAGGCGCCGTTATCCGCTTCAATGATCAAAAGGTAC GATGAACATCACCAAGACTTGACACTTCTCAAGGCCCTAGTCCG TCAGCAACTGCCTGAGAAATATAAGGAAATATTCTTTGATCAGTC GAAAAACGGGTACGCAGGTTATATTGACGGCGGAGCGAGTCAA GAGGAATTCTACAAGTTTATCAAACCCATATTAGAGAAGATGGAT GGGACGGAAGAGTTGCTTGTAAAACTCAATCGCGAAGATCTACT GCGAAAGCAGCGGACTTTCGACAACGGTAGCATTCCACATCAAA TCCACTTAGGCGAATTGCATGCTATACTTAGAAGGCAGGAGGAT TTTTATCCGTTCCTCAAAGACAATCGTGAAAAGATTGAGAAAATC CTAACCTTTCGCATACCTTACTATGTGGGACCCCTGGCCCGAGG GAACTCTCGGTTCGCATGGATGACAAGAAAGTCCGAAGAAACGA TTACTCCATGGAATTTTGAGGAAGTTGTCGATAAAGGTGCGTCA GCTCAATCGTTCATCGAGAGGATGACCAACTTTGACAAGAATTTA CCGAACGAAAAAGTATTGCCTAAGCACAGTTTACTTTACGAGTAT TTCACAGTGTACAATGAACTCACGAAAGTTAAGTATGTCACTGAG GGCATGCGTAAACCCGCCTTTCTAAGCGGAGAACAGAAGAAAG CAATAGTAGATCTGTTATTCAAGACCAACCGCAAAGTGACAGTTA AGCAATTGAAAGAGGACTACTTTAAGAAAATTGAATGCTTCGATT CTGTCGAGATCTCCGGGGTAGAAGATCGATTTAATGCGTCACTT GGTACGTATCATGACCTCCTAAAGATAATTAAAGATAAGGACTTC CTGGATAACGAAGAGAATGAAGATATCTTAGAAGATATAGTGTTG ACTCTTACCCTCTTTGAAGATCGGGAAATGATTGAGGAAAGACT AAAAACATACGCTCACCTGTTCGACGATAAGGTTATGAAACAGTT AAAGAGGCGTCGCTATACGGGCTGGGGACGATTGTCGCGGAAA CTTATCAACGGGATAAGAGACAAGCAAAGTGGTAAAACTATTCT CGATTTTCTAAAGAGCGACGGCTTCGCCAATAGGAACTTTATGC AGCTGATCCATGATGACTCTTTAACCTTCAAAGAGGATATACAAA AGGCACAGGTTTCCGGACAAGGGGACTCATTGCACGAACATATT GCGAATCTTGCTGGTTCGCCAGCCATCAAAAAGGGCATACTCCA GACAGTCAAAGTAGTGGATGAGCTAGTTAAGGTCATGGGACGTC ACAAACCGGAAAACATTGTAATCGAGATGGCACGCGAAAATCAA ACGACTCAGAAGGGGCAAAAAAACAGTCGAGAGCGGATGAAGA GAATAGAAGAGGGTATTAAAGAACTGGGCAGCCAGATCTTAAAG GAGCATCCTGTGGAAAATACCCAATTGCAGAACGAGAAACTTTA CCTCTATTACCTACAAAATGGAAGGGACATGTATGTTGATCAGG AACTGGACATAAACCGTTTATCTGATTACGACGTCGATCACATTG TACCCCAATCCTTTTTGAAGGACGATTCAATCGACAATAAAGTGC TTACACGCTCGGATAAGAACCGAGGGAAAAGTGACAATGTTCCA AGCGAGGAAGTCGTAAAGAAAATGAAGAACTATTGGCGGCAGCT CCTAAATGCGAAACTGATAACGCAAAGAAAGTTCGATAACTTAAC TAAAGCTGAGAGGGGTGGCTTGTCTGAACTTGACAAGGCCGGA TTTATTAAACGTCAGCTCGTGGAAACCCGCCAAATCACAAAGCA TGTTGCACAGATACTAGATTCCCGAATGAATACGAAATACGACG AGAACGATAAGCTGATTCGGGAAGTCAAAGTAATCACTTTAAAGT CAAAATTGGTGTCGGACTTCAGAAAGGATTTTCAATTCTATAAAG TTAGGGAGATAAATAACTACCACCATGCGCACGACGCTTATCTT AATGCCGTCGTAGGGACCGCACTCATTAAGAAATACCCGAAGCT AGAAAGTGAGTTTGTGTATGGTGATTACAAAGTTTATGACGTCCG

		TAAGATGATCGCGAAAAGCGAACAGGAGATAGGCAAGGCTACA GCCAAATACTTCTTTTATTCTAACATTATGAATTTCTTTAAGACGG AAATCACTCTGGCAAACGGAGAGATACGCAAACGACCTTTAATT GAAACCAATGGGGAGACAGGTGAAATCGTATGGGATAAGGGCC GGGACTTCGCGACGGTGAGAAAAGTTTTGTCCATGCCCCAAGTC AACATAGTAAAGAAAACTGAGGTGCAGACCGGAGGGTTTTCAAA GGAATCGATTCTTCCAAAAAGGAATAGTGATAAGCTCATCGCTC GTAAAAAGGACTGGGACCCGAAAAAGTACGGTGGCTTCGATAG CCCTACAGTTGCCTATTCTGTCCTAGTAGTGGCAAAAGTTGAGA AGGGAAAATCCAAGAAACTGAAGTCAGTCAAAGAATTATTGGGG ATAACGATTATGGAGCGCTCGTCTTTTGAAAAGAACCCCATCGA CTTCCTTGAGGCGAAAGGTTACAAGGAAGTAAAAAAGGATCTCA TAATTAAACTACCAAAGTATAGTCTGTTTGAGTTAGAAAATGGCC GAAAACGGATGTTGGCTAGCGCCGGAGAGCTTCAAAAGGGGAA CGAACTCGCACTACCGTCTAAATACGTGAATTTCCTGTATTTAGC GTCCCATTACGAGAAGTTGAAAGGTTCACCTGAAGATAACGAAC AGAAGCAACTTTTTGTTGAGCAGCACAAACATTATCTCGACGAAA TCATAGAGCAAATTTCGGAATTCAGTAAGAGAGTCATCCTAGCT GATGCCAATCTGGACAAAGTATTAAGCGCATACAACAAGCACAG GGATAAACCCATACGTGAGCAGGCGGAAAATATTATCCATTTGT TTACTCTTACCAACCTCGGCGCTCCAGCCGCATTCAAGTATTTTG ACACAACGATAGATCGCAAACGATACACTTCTACCAAGGAGGTG CTAGACGCGACACTGATTCACCAATCCATCACGGGATTATATGA AACTCGGATAGATTTGTCACAGCTTGGGGGTGACGGATCCCCCA AGAAGAAGAGGAAAGTCTCGAGCGACTACAAAGACCATGACGG TGATTATAAAGATCATGACATCGATTACAAGGATGACGATGACAA GGCTGCAGGA
야생형 Cas9 폴리펩티드 서열	234	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED ILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHK PENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVE NTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQ RKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNT KYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAY LNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAK YFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATV RKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK KYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEK NPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKG NELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEII EQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLG APAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGG D

[1097]

PAM-결합 SpEQR Cas9 폴리펩티드 서열	1304	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED ILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHK PENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVE NTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQ RKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNT KYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAY LNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAK YFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATV RKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK KYGGFESPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEK NPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKG NELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEII EQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLG APAAFKYFDTTIDRKQYRSTKEVLDATLIHQSITGLYETRIDLSQLGG D
PAM-결합 SpVQR Cas9 폴리펩티드 서열	1305	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED ILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHK PENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVE NTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQ RKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNT KYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAY LNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAK YFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATV RKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK KYGGFVSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEK NPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKG NELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEII EQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLG

[1098]

		APAAFKYFDTTIDRKQYRSTKEVLDATLIHQSITGLYETRIDLSQLGG D
SpVQR Cas9 폴리펩티드 서열	1306	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED ILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHK PENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVE NTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQ RKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNT KYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAY LNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAK YFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATV RKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK KYGGFVSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEK NPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASARELQKG NELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEII EQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLG APAAFKYFDTTIDRKEYRSTKEVLDATLIHQSITGLYETRIDLSQLGG D
SpyMacCas9 폴리펩티드 서열	1307	MDKKYSIGLDIGTNSVGWAVITDDYKVPSKKFKVLGNTDRHSIKKNL IGALLFGSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LADSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV QIYNQLFEENPINASRVDAKAILSARLSKSRRLENLIAQLPGEKRNGL FGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGD QYADLFLAAKNLSDAILLSDILRVNSEITKAPLSASMIKRYDEHHQDL TLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPIL EKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQE DFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITP WNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVY NELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKED YFKKIECFDSVEISGVEDRFNASLGAYHDLLKIIKDKDFLDNEENEDI LEDIVLTLTLFEDRGMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGHSLHEQIANLAGSPAIKKGILQTVKIVDELVKVMGHKPE NIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENT QLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFIKDD SIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRK FDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKY DENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLN AVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYF FYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRK VLSMPQVNIVKKTEIQTVGQNGGLFDDNPKSPLEVTPSKLVPLKKEL NPKKYGGYQKPTTAYPVLLITDTKQLIPISVMNKKQFEQNPVKFLRD RGYQQVGKNDFIKLPKYTLVDIGDGIKRLWASSKEIHKGNQLVVSKK SQILLYHAHHLDSDLSNDYLQNHNQQFDVLFNEIISFSKKCKLGKEHI

[1099]

		QKIENVYSNKKNSASIEELAESFIKLLGFTQLGATSPFNFLGVKLNQK QYKGKKDYILPCTEGTLIRQSITGLYETRVDLSKIGED
CP5 폴리펩티드 서열	257	EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWD KGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIAR KKDWDPKKYGGFMQPTVAYSVLVVAKVEKGKSKKLKSVKELLGITI MERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLA SAKFLQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQ HKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENII HLFTLTNLGAPRAFKYFDTTIARKEYRSTKEVLDATLIHQSITGLYET RIDLSQLGGDGGSGGSGGSGGSGGSGGSGGMDKKYSIGLAIGTN SVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEA TRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVE EDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYL ALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINA SGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPN FKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLS DAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPE KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVK LNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREK IEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGA SAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEG MRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEI SGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFED REMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQ SGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSL HEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMAREN QTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYL YYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTR SDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAE RGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIR EVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALI KKYPKLESEFVYGDYKVYDVRKMIAKSEQEGADKRTADGSEFESP KKKRKV
Cas12c1 폴리펩티드 서열	266	MQTKKTHLHLISAKASRKYRRTIACLSDTAKKDLERRKQSGAADPA QELSCLKTIKFKLEVPEGSKLPSFDRISQIYNALETIEKGSLSYLLFALI LSGFRIFPNSSAAKTFASSSCYKNDQFASQIKEIFGEMVKNFIPSELE SILKKGRRKNNKDWTEENIKRVLNSEFGRKNSEGSSALFDSFLSKF SQELFRKFDSWNEVNKKYLEAAELLDSMLASYGPFDSVCKMIGDS DSRNSLPDKSTIAFTNNAEITVDIESSVMPYMAIAALLREYRQSKSKA APVAYVQSHLTTTNGNGLSWFFKFGLDLIRKAPVSSKQSTSDGSKS LQELFSVPDDKLDGLKFIKEACEALPEASLLCGEKGELLGYQDFRTS FAGHIDSWVANYVNRLFELIELVNQLPESIKLPSILTQKNHNLVASLG LQEAEVSHSLELFEGLVKNVRQTLKKLAGIDISSSPNEQDIKEFYAFS DVLNRLGSIRNQIENAVQTAKKDKIDLESAIEWKEWKKLKKLPKLNG LGGGVPKQQELLDKALESVKQIRHYQRIDFERVIQWAVNEHCLETV PKFLVDAEKKKINKESSTDFAAKENAVRFLLEGIGAAARGKTDSVSK AAYNWFVVNNFLAKKDLNRYFINCQGCIYKPPYSKRRSLAFALRSD NKDTIEVVWEKFETFYKEISKEIEKFNIFSQEFQTFLHLENLRMKLLL RRIQKPIPAEIAFFSLPQEYYDSLPPNVAFLALNQEITPSEYITQFNLY SSFLNGNLILLRRSRSYLRAKFSWVGNSKLIYAAKEARLWKIPNAYW KSDEWKMILDSNVLVFDKAGNVLPAPTLKKVCEREGDLRLFYPLLR QLPHDWCYRNPFVKSVGREKNVIEVNKEGEPKVASALPGSLFRLIG PAPFKSLLDDCFFNPLDKDLRECMLIVDQEISQKVEAQKVEASLESC TYSIAVPIRYHLEEPKVSNQFENVLAIDQGEAGLAYAVFSLKSIGEAE TKPIAVGTIRIPSIRRLIHSVSTYRKKKQRLQNFKQNYDSTAFIMREN VTGDVCAKIVGLMKEFNAFPVLEYDVKNLESGSRQLSAVYKAVNSH FLYFKEPGRDALRKQLWYGGDSWTIDGIEIVTRERKEDGKEGVEKI VPLKVFPGRSVSARFTSKTCSCCGRNVFDWLFTEKKAKTNKKFNV NSKGELTTADGVIQLFEADRSKGPKFYARRKERTPLTKPIAKGSYSL

[1100]

		EEIERRVRTNLRRAPKSKQSRDTSQSQYFCVYKDCALHFSGMQAD ENAAINIGRRFLTALRKNRRSDFPSNVKISDRLLDN
Cas12c2 폴리펩티드 서열	267	MTKHSIPLHAFRNSGADARKWKGRIALLAKRGKETMRTLQFPLEMS EPEAAAINTTPFAVAYNAIEGTGKGTLFDYWAKLHLAGFRFFPSGG AATIFRQQAVFEDASWNAAFCQQSGKDWPWLVPSKLYERFTKAPR EVAKKDGSKKSIEFTQENVANESHVSLVGASITDKTPEDQKEFFLK MAGALAEKFDSWKSANEDRIVAMKVIDEFLKSEGLHLPSLENIAVKC SVETKPDNATVAWHDAPMSGVQNLAIGVFATCASRIDNIYDLNGGK LSKLIQESATTPNVTALSWLFGKGLEYFRTTDIDTIMQDFNIPASAKE SIKPLVESAQAIPTMTVLGKKNYAPFRPNFGGKIDSWIANYASRLML LNDILEQIEPGFELPQALLDNETLMSGIDMTGDELKELIEAVYAWVD AAKQGLATLLGRGGNVDDAVQTFEQFSAMMDTLNGTLNTISARYV RAVEMAGKDEARLEKLIECKFDIPKWCKSVPKLVGISGGLPKVEEEI KVMNAAFKDVRARMFVRFEEIAAYVASKGAGMDVYDALEKRELEQI KKLKSAVPERAHIQAYRAVLHRIGRAVQNCSEKTKQLFSSKVIEMG VFKNPSHLNNFIFNQKGAIYRSPFDRSRHAPYQLHADKLLKNDWLE LLAEISATLMASESTEQMEDALRLERTRLQLQLSGLPDWEYPASLA KPDIEVEIQTALKMQLAKDTVTSDVLQRAFNLYSSVLSGLTFKLLRR SFSLKMRFSVADTTQLIYVPKVCDWAIPKQYLQAEGEIGIAARVVTE SSPAKMVTEVEMKEPKALGHFMQQAPHDWYFDASLGGTQVAGRI VEKGKEVGKERKLVGYRMRGNSAYKTVLDKSLVGNTELSQCSMIIE IPYTQTVDADFRAQVQAGLPKVSINLPVKETITASNKDEQMLFDRFV AIDLGERGLGYAVFDAKTLELQESGHRPIKAITNLLNRTHHYEQRPN QRQKFQAKFNVNLSELRENTVGDVCHQINRICAYYNAFPVLEYMVP DRLDKQLKSVYESVTNRYIWSSTDAHKSARVQFWLGGETWEHPYL KSAKDKKPLVLSPGRGASGKGTSQTCSCCGRNPFDLIKDMKPRAKI AVVDGKAKLENSELKLFERNLESKDDMLARRHRNERAGMEQPLTP GNYTVDEIKALLRANLRRAPKNRRTKDTTVSEYHCVFSDCGKTMHA DENAAVNIGGKFIADIEK
OspCas12c 폴리펩티드 서열	268	MTKLRHRQKKLTHDWAGSKKREVLGSNGKLQNPLLMPVKKGQVT EFRKAFSAYARATKGEMTDGRKNMFTHSFEPFKTKPSLHQCELAD KAYQSLHSYLPGSLAHFLLSAHALGFRIFSKSGEATAFQASSKIEAY ESKLASELACVDLSIQNLTISTLFNALTTSVRGKGEETSADPLIARFY TLLTGKPLSRDTQGPERDLAEVISRKIASSFGTWKEMTANPLQSLQ FFEEELHALDANVSLSPAFDVLIKMNDLQGDLKNRTIVFDPDAPVFE YNAEDPADIIIKLTARYAKEAVIKNQNVGNYVKNAITTTNANGLGWLL NKGLSLLPVSTDDELLEFIGVERSHPSCHALIELIAQLEAPELFEKNV FSDTRSEVQGMIDSAVSNHIARLSSSRNSLSMDSEELERLIKSFQIH TPHCSLFIGAQSLSQQLESLPEALQSGVNSADILLGSTQYMLTNSLV EESIATYQRTLNRINYLSGVAGQINGAIKRKAIDGEKIHLPAAWSELIS LPFIGQPVIDVESDLAHLKNQYQTLSNEFDTLISALQKNFDLNFNKAL LNRTQHFEAMCRSTKKNALSKPEIVSYRDLLARLTSCLYRGSLVLR RAGIEVLKKHKIFESNSELREHVHERKHFVFVSPLDRKAKKLLRLTD SRPDLLHVIDEILQHDNLENKDRESLWLVRSGYLLAGLPDQLSSSFI NLPIITQKGDRRLIDLIQYDQINRDAFVMLVTSAFKSNLSGLQYRANK QSFVVTRTLSPYLGSKLVYVPKDKDWLVPSQMFEGRFADILQSDY MVWKDAGRLCVIDTAKHLSNIKKSVFSSEEVLAFLRELPHRTFIQTE VRGLGVNVDGIAFNNGDIPSLKTFSNCVQVKVSRTNTSLVQTLNRW FEGGKVSPPSIQFERAYYKKDDQIHEDAAKRKIRFQMPATELVHAS DDAGWTPSYLLGIDPGEYGMGLSLVSINNGEVLDSGFIHINSLINFA SKKSNHQTKVVPRQQYKSPYANYLEQSKDSAAGDIAHILDRLIYKLN ALPVFEALSGNSQSAADQVWTKVLSFYTWGDNDAQNSIRKQHWF GASHWDIKGMLRQPPTEKKPKPYIAFPGSQVSSYGNSQRCSCCGR NPIEQLREMAKDTSIKELKIRNSEIQLFDGTIKLFNPDPSTVIERRRHN LGPSRIPVADRTFKNISPSSLEFKELITIVSRSIRHSPEFIAKKRGIGSE YFCAYSDCNSSLNSEANAAANVAQKFQKQLFFEL
Cas12g1 폴리펩티드 서열	269	MAQASSTPAVSPRPRPRYREERTLVRKLLPRPGQSKQEFRENVKK LRKAFLQFNADVSGVCQWAIQFRPRYGKPAEPTETFWKFFLEPET SLPPNDSRSPEFRRLQAFEAAAGINGAAALDDPAFTNELRDSILAVA

[1101]

		SRPKTKEAQRLFSRLKDYQPAHRMILAKVAAEWIESRYRRAHQNW ERNYEEWKKEKQEWEQNHPELTPEIREAFNQIFQQLEVKEKRVRIC PAARLLQNKDNCQYAGKNKHSVLCNQFNEFKKNHLQGKAIKFFYK DAEKYLRCGLQSLKPNVQGPFREDWNKYLRYMNLKEETLRGKNG GRLPHCKNLGQECEFNPHTALCKQYQQQLSSRPDLVQHDELYRK WRREYWREPRKPVFRYPSVKRHSIAKIFGENYFQADFKNSVVGLR LDSMPAGQYLEFAFAPWPRNYRPQPGETEISSVHLHFVGTRPRIGF RFRVPHKRSRFDCTQEELDELRSRTFPRKAQDQKFLEAARKRLLET FPGNAEQELRLLAVDLGTDSARAAFFIGKTFQQAFPLKIVKIEKLYEQ WPNQKQAGDRRDASSKQPRPGLSRDHVGRHLQKMRAQASEIAQK RQELTGTPAPETTTDQAAKKATLQPFDLRGLTVHTARMIRDWARLN ARQIIQLAEENQVDLIVLESLRGFRPPGYENLDQEKKRRVAFFAHGR IRRKVTEKAVERGMRVVTVPYLASSKVCAECRKKQKDNKQWEKNK KRGLFKCEGCGSQAQVDENAARVLGRVFWGEIELPTAIP
Cas12h1 폴리펩티드 서열	270	MKVHEIPRSQLLKIKQYEGSFVEWYRDLQEDRKKFASLLFRWAAFG YAAREDDGATYISPSQALLERRLLLGDAEDVAIKFLDVLFKGGAPSS SCYSLFYEDFALRDKAKYSGAKREFIEGLATMPLDKIIERIRQDEQLS KIPAEEWLILGAEYSPEEIWEQVAPRIVNVDRSLGKQLRERLGIKCR RPHDAGYCKILMEVVARQLRSHNETYHEYLNQTHEMKTKVANNLT NEFDLVCEFAEVLEEKNYGLGWYVLWQGVKQALKEQKKPTKIQIAV DQLRQPKFAGLLTAKWRALKGAYDTWKLKKRLEKRKAFPYMPNW DNDYQIPVGLTGLGVFTLEVKRTEVVVDLKEHGKLFCSHSHYFGDL TAEKHPSRYHLKFRHKLKLRKRDSRVEPTIGPWIEAALREITIQKKP NGVFYLGLPYALSHGIDNFQIAKRFFSAAKPDKEVINGLPSEMVVGA ADLNLSNIVAPVKARIGKGLEGPLHALDYGYGELIDGPKILTPDGPR CGELISLKRDIVEIKSAIKEFKACQREGLTMSEETTTWLSEVESPSDS PRCMIQSRIADTSRRLNSFKYQMNKEGYQDLAEALRLLDAMDSYN SLLESYQRMHLSPGEQSPKEAKFDTKRASFRDLLRRRVAHTIVEYF DDCDIVFFEDLDGPSDSDSRNNALVKLLSPRTLLLYIRQALEKRGIG MVEVAKDGTSQNNPISGHVGWRNKQNKSEIYFYEDKELLVMDADE VGAMNILCRGLNHSVCPYSFVTKAPEKKNDEKKEGDYGKRVKRFL KDRYGSSNVRFLVASMGFVTVTTKRPKDALVGKRLYYHGGELVTH DLHNRMKDEIKYLVEKEVLARRVSLSDSTIKSYKSFAHV
Cas12i1 폴리펩티드 서열	271	MSNKEKNASETRKAYTTKMIPRSHDRMKLLGNFMDYLMDGTPIFFE LWNQFGGGIDRDIISGTANKDKISDDLLLAVNWFKVMPINSKPQGVS PSNLANLFQQYSGSEPDIQAQEYFASNFDTEKHQWKDMRVEYERL LAELQLSRSDMHHDLKLMYKEKCIGLSLSTAHYITSVMFGTGAKNN RQTKHQFYSKVIQLLEESTQINSVEQLASIILKAGDCDSYRKLRIRCS RKGATPSILKIVQDYELGTNHDDEVNVPSLIANLKEKLGRFEYECEW KCMEKIKAFLASKVGPYYLGSYSAMLENALSPIKGMTTKNCKFVLK QIDAKNDIKYENEPFGKIVEGFFDSPYFESDTNVKWVLHPHHIGESN IKTLWEDLNAIHSKYEEDIASLSEDKKEKRIKVYQGDVCQTINTYCEE VGKEAKTPLVQLLRYLYSRKDDIAVDKIIDGITFLSKKHKVEKQKINP VIQKYPSFNFGNNSKLLGKIISPKDKLKHNLKCNRNQVDNYIWIEIKV LNTKTMRWEKHHYALSSTRFLEEVYYPATSENPPDALAARFRTKTN GYEGKPALSAEQIEQIRSAPVGLRKVKKRQMRLEAARQQNLLPRYT WGKDFNINICKRGNNFEVTLATKVKKKKEKNYKVVLGYDANIVRKN TYAAIEAHANGDGVIDYNDLPVKPIESGFVTVESQVRDKSYDQLSY NGVKLLYCKPHVESRRSFLEKYRNGTMKDNRGNNIQIDFMKDFEAI ADDETSLYYFNMKYCKLLQSSIRNHSSQAKEYREEIFELLRDGKLSV LKLSSLSNLSFVMFKVAKSLIGTYFGHLLKKPKNSKSDVKAPPITDE DKQKADPEMFALRLALEEKRLNKVKSKKEVIANKIVAKALELRDKYG PVLIKGENISDTTKKGKKSSTNSFLMDWLARGVANKVKEMVMMHQ GLEFVEVNPNFTSHQDPFVHKNPENTFRARYSRCTPSELTEKNRK EILSFLSDKPSKRPTNAYYNEGAMAFLATYGLKKNDVLGVSLEKFK QIMANILHQRSEDQLLFPSRGGMFYLATYKLDADATSVNWNGKQF WVCNADLVAAYNVGLVDIQKDFKKK
Cas12i2 폴리펩티드 서열	272	MSSAIKSYKSVLRPNERKNQLLKSTIQCLEDGSAFFFKMLQGLFGGI TPEIVRFSTEQEKQQQDIALWCAVNWFRPVSQDSLTHTIASDNLVE

[1102]

		KFEEYYGGTASDAIKQYFSASIGESYYWNDCRQQYYDLCRELGVE VSDLTHDLEILCREKCLAVATESNQNNSIISVLFGTGEKEDRSVKLRI TKKILEAISNLKEIPKNVAPIQEIILNVAKATKETFRQVYAGNLGAPST LEKFIAKDGQKEFDLKKLQTDLKKVIRGKSKERDWCCQEELRSYVE QNTIQYDLWAWGEMFNKAHTALKIKSTRNYNFAKQRLEQFKEIQSL NNLLVVKKLNDFFDSEFFSGEETYTICVHHLGGKDLSKLYKAWEDD PADPENAIVVLCDDLKNNFKKEPIRNILRYIFTIRQECSAQDILAAAKY NQQLDRYKSQKANPSVLGNQGFTWTNAVILPEKAQRNDRPNSLDL RIWLYLKLRHPDGRWKKHHIPFYDTRFFQEIYAAGNSPVDTCQFRT PRFGYHLPKLTDQTAIRVNKKHVKAAKTEARIRLAIQQGTLPVSNLKI TEISATINSKGQVRIPVKFDVGRQKGTLQIGDRFCGYDQNQTASHA YSLWEVVKEGQYHKELGCFVRFISSGDIVSITENRGNQFDQLSYEG LAYPQYADWRKKASKFVSLWQITKKNKKKEIVTVEAKEKFDAICKY QPRLYKFNKEYAYLLRDIVRGKSLVELQQIRQEIFRFIEQDCGVTRL GSLSLSTLETVKAVKGIIYSYFSTALNASKNNPISDEQRKEFDPELFA LLEKLELIRTRKKKQKVERIANSLIQTCLENNIKFIRGEGDLSTTNNAT KKKANSRSMDWLARGVFNKIRQLAPMHNITLFGCGSLYTSHQDPL VHRNPDKAMKCRWAAIPVKDIGDWVLRKLSQNLRAKNIGTGEYYH QGVKEFLSHYELQDLEEELLKWRSDRKSNIPCWVLQNRLAEKLGN KEAVVYIPVRGGRIYFATHKVATGAVSIVFDQKQVWVCNADHVAAA NIALTVKGIGEQSSDEENPDGSRIKLQLTS
링커	1308	(GGGS)N
링커	109	(GGGGS)N
링커	1309	(EAAAK)N
링커	56	SGSETPGTSESATPES
	57	(SGGS)N
링커	273	GGSGGS
링커	1310	GSSGSETPGTSESATPESSG
링커	1311	GGAGGCTCTGGAGGAAGC
링커	1312	GGCTCTTCTGGATCTGAAACACCTGGCACAAGCGAGAGCGCCA CCCCTGAGAGCTCTGGC
AacCas12b 폴리펩티드 서열	259	MAVKSMKVKLRLDNMPEIRAGLWKLHTEVNAGVRYYTEWLSLLRQ ENLYRRSPNGDGEQECYKTAEECKAELLERLRARQVENGHCGPA GSDDELLQLARQLYELLVPQAIGAKGDAQQIARKFLSPLADKDAVG GLGIAKAGNKPRWVRMREAGEPGWEEEKAKAEARKSTDRTADVL RALADFGLKPLMRVYTDSDMSSVQWKPLRKGQAVRTWDRDMFQ QAIERMMSWESWNQRVGEAYAKLVEQKSRFEQKNFVGQEHLVQL VNQLQQDMKEASHGLESKEQTAHYLTGRALRGSDKVFEKWEKLD PDAPFDLYDTEIKNVQRRNTRRFGSHDLFAKLAEPKYQALWREDA SFLTRYAVYNSIVRKLNHAKMFATFTLPDATAHPIWTRFDKLGGNLH QYTFLFNEFGEGRHAIRFQKLLTVEDGVAKEVDDVTVPISMSAQLD DLLPRDPHELVALYFQDYGAEQHLAGEFGGAKIQYRRDQLNHLHA RRGARDVYLNLSVRVQSQSEARGERRPPYAAVFRLVGDNHRAFV HFDKLSDYLAEHPDDGKLGSEGLLSGLRVMSVDLGLRTSASISVFR VARKDELKPNSEGRVPFCFPIEGNENLVAVHERSQLLKLPGETESK DLRAIREERQRTLRQLRTQLAYLRLLVRCGSEDVGRRERSWAKLIE QPMDANQMTPDWREAFEDELQKLKSLYGICGDREWTEAVYESVR RVWRHMGKQVRDWRKDVRSGERPKIRGYQKDVVGGNSIEQIEYL ERQYKFLKSWSFFGKVSGQVIRAEKGSRFAITLREHIDHAKEDRLK KLADRIIMEALGYVYALDDERGKGKWVAKYPPCQLILLEELSEYQFN NDRPPSENNQLMQWSHRGVFQELLNQAQVHDLLVGTMYAAFSSR FDARTGAPGIRCRRVPARCAREQNPEPFPWWLNKFVAEHKLDGC PLRADDLIPTGEGEFFVSPFSAEEGDFHQIHADLNAAQNLQRRLWS DFDISQIRLRCDWGEVDGEPVLIPRTTGKRTADSYGNKVFYTKTGV TYYERERGKKRRKVFAQEELSEEEAELLVEADEAREKSVVLMRDP

[1103]

		SGIINRGDWTRQKEFWSMVNQRIEGYLVKQIRSRVRLQESACENT GDI
BhCas12b 폴리펩티드 서열	260	MAPKKKRKVGIHGVPAAATRSFILKIEPNEEVKKGLWKTHEVLNHGI AYYMNILKLIRQEAIYEHHEQDPKNPKKVSKAEIQAELWDFVLKMQK CNSFTHEVDKDEVFNILRELYEELVPSSVEKKGEANQLSNKFLYPLV DPNSQSGKGTASSGRKPRWYNLKIAGDPSWEEEKKKWEEDKKKD PLAKILGKLAEYGLIPLFIPYTDSNEPIVKEIKWMEKSRNQSVRRLDK DMFIQALERFLSWESWNLKVKEEYEKVEKEYKTLEERIKEDIQALKA LEQYEKERQEQLLRDTLNTNEYRLSKRGLRGWREIIQKWLKMDEN EPSEKYLEVFKDYQRKHPREAGDYSVYEFLSKKENHFIWRNHPEY PYLYATFCEIDKKKKDAKQQATFTLADPINHPLWVRFEERSGSNLN KYRILTEQLHTEKLKKKLTVQLDRLIYPTESGGWEEKGKVDIVLLPS RQFYNQIFLDIEEKGKHAFTYKDESIKFPLKGTLGGARVQFDRDHLR RYPHKVESGNVGRIYFNMTVNIEPTESPVSKSLKIHRDDFPKVVNFK PKELTEWIKDSKGKKLKSGIESLEIGLRVMSIDLGQRQAAAASIFEVV DQKPDIEGKLFFPIKGTELYAVHRASFNIKLPGETLVKSREVLRKAR EDNLKLMNQKLNFLRNVLHFQQFEDITEREKRVTKWISRQENSDVP LVYQDELIQIRELMYKPYKDWVAFLKQLHKRLEVEIGKEVKHWRKS LSDGRKGLYGISLKNIDEIDRTRKFLLRWSLRPTEPGEVRRLEPGQR FAIDQLNHLNALKEDRLKKMANTIIMHALGYCYDVRKKKWQAKNPA CQIILFEDLSNYNPYEERSRFENSKLMKWSRREIPRQVALQGEIYGL QVGEVGAQFSSRFHAKTGSPGIRCSVVTKEKLQDNRFFKNLQREG RLTLDKIAVLKEGDLYPDKGGEKFISLSKDRKCVTTHADINAAQNLQ KRFWTRTHGFYKVYCKAYQVDGQTVYIPESKDQKQKIIEEFGEGYF ILKDGVYEWVNAGKLKIKKGSSKQSSSELVDSDILKDSFDLASELKG EKLMLYRDPSGNVFPSDKWMAAGVFFGKLERILISKLTNQYSISTIE DDSSKQSMKRPAATKKAGQAKKKK
BvCas12b (바실러스 sp. V3-13) 폴리뉴클레오티드 서열	264	MAIRSIKLKMKTNSGTDSIYLRKALWRTHQLINEGIAYYMNLLTLYRQ EAIGDKTKEAYQAELINIIRNQQRNNGSSEEHGSDQEILALLRQLYEL IIPSSIGESGDANQLGNKFLYPLVDPNSQSGKGTSNAGRKPRWKRL KEEGNPDWELEKKKDEERKAKDPTVKIFDNLNKYGLLPLFPLFTNIQ KDIEWLPLGKRQSVRKWDKDMFIQAIERLLSWESWNRRVADEYKQ LKEKTESYYKEHLTGGEEWIEKIRKFEKERNMELEKNAFAPNDGYFI TSRQIRGWDRVYEKWSKLPESASPEELWKVVAEQQNKMSEGFGD PKVFSFLANRENRDIWRGHSERIYHIAAYNGLQKKLSRTKEQATFTL PDAIEHPLWIRYESPGGTNLNLFKLEEKQKKNYYVTLSKIIWPSEEK WIEKENIEIPLAPSIQFNRQIKLKQHVKGKQEISFSDYSSRISLDGVLG GSRIQFNRKYIKNHKELLGEGDIGPVFFNLVVDVAPLQETRNGRLQ SPIGKALKVISSDFSKVIDYKPKELMDWMNTGSASNSFGVASLLEG MRVMSIDMGQRTSASVSIFEVVKELPKDQEQKLFYSINDTELFAIHK RSFLLNLPGEVVTKNNKQQRQERRKKRQFVRSQIRMLANVLRLET KKTPDERKKAIHKLMEIVQSYDSWTASQKEVWEKELNLLTNMAAFN DEIWKESLVELHHRIEPYVGQIVSKWRKGLSEGRKNLAGISMWNID ELEDTRRLLISWSKRSRTPGEANRIETDEPFGSSLLQHIQNVKDDRL KQMANLIIMTALGFKYDKEEKDRYKRWKETYPACQIILFENLNRYLF NLDRSRRENSRLMKWAHRSIPRTVSMQGEMFGLQVGDVRSEYSS RFHAKTGAPGIRCHALTEEDLKAGSNTLKRLIEDGFINESELAYLKK GDIIPSQGGELFVTLSKRYKKDSDNNELTVIHADINAAQNLQKRFWQ QNSEVYRVPCQLARMGEDKLYIPKSQTETIKKYFGKGSFVKNNTEQ EVYKWEKSEKMKIKTDTTFDLQDLDGFEDISKTIELAQEQQKKYLTM FRDPSGYFFNNETWRPQKEYWSIVNNIIKSCLKKKILSNKVEL
BTCas12b.BTCas12b 폴리펩티드 서열	265	MATRSFILKIEPNEEVKKGLWKTHEVLNHGIAYYMNILKLIRQEAIYE HHEQDPKNPKKVSKAEIQAELWDFVLKMQKCNSFTHEVDKDVVFNI LRELYEELVPSSVEKKGEANQLSNKFLYPLVDPNSQSGKGTASSG RKPRWYNLKIAGDPSWEEEKKKWEEDKKKDPLAKILGKLAEYGLIP LFIPFTDSNEPIVKEIKWMEKSRNQSVRRLDKDMFIQALERFLSWES WNLKVKEEYEKVEKEHKTLEERIKEDIQAFKSLEQYEKERQEQLLR DTLNTNEYRLSKRGLRGWREIIQKWLKMDENEPSEKYLEVFKDYQ RKHPREAGDYSVYEFLSKKENHFIWRNHPEYPYLYATFCEIDKKKK

[1104]

[1105]

		DAKQQATFTLADPINHPLWVRFEERSGSNLNKYRILTEQLHTEKLKK KLTVQLDRLIYPTESGGWEEKGKVDIVLLPSRQFYNQIFLDIEEKGK HAFTYKDESIKFPLKGTLGGARVQFDRDHLRRYPHKVESGNVGRIY FNMTVNIEPTESPVSKSLKIHRDDFPKFVNFKPKELTEWIKDSKGKK LKSGIESLEIGLRVMSIDLGQRQAAAASIFEVVDQKPDIEGKLFFPIK GTELYAVHRASFNIKLPGETLVKSREVLRKAREDNLKLMNQKLNFL RNVLHFQQFEDITEREKRVTKWISRQENSDVPLVYQDELIQIRELMY KPYKDWVAFLKQLHKRLEVEIGKEVKHWRKSLSDGRKGLYGISLKN IDEIDRTRKFLLRWSLRPTEPGEVRRLEPGQRFAIDQLNHLNALKED RLKKMANTIIMHALGYCYDVRKKKWQAKNPACQIILFEDLSNYNPYE ERSRFENSKLMKWSRREIPRQVALQGEIYGLQVGEVGAQFSSRFH AKTGSPGIRCSVVTKEKLQDNRFFKNLQREGRLTLDKIAVLKEGDLY PDKGGEKFISLSKDRKLVTTHADINAAQNLQKRFWTRTHGFYKVYC KAYQVDGQTVYIPESKDQKQKIIEEFGEGYFILKDGVYEWGNAGKL KIKKGSSKQSSSELVDSDILKDSFDLASELKGEKLMLYRDPSGNVFP SDKWMAAGVFFGKLERILISKLTNQYSISTIEDDSSKQSM
5'UTR	261	GGGAAATAAGAGAGAAAAGAAGAGTAAGAAGAAATATAAGAGCC ACC
3'UTR (TriLink 표준 UTR)	262	GCTGGAGCCTCGGTGGCCATGCTTCTTGCCCCTTGGGCCTCCC CCCAGCCCCTCCTCCCCTTCCTGCACCCGTACCCCCGTGGTCTT TGAATAAAGTCTGA
bhCas12b (V4) 폴리뉴클레오티드 서열	263	ATGGCCCCAAAGAAGAAGCGGAAGGTCGGTATCCACGGAGTCC CAGCAGCCGCCACCAGATCCTTCATCCTGAAGATCGAGCCCAA CGAGGAAGTGAAGAAAGGCCTCTGGAAAACCCACGAGGTGCTG AACCACGGAATCGCCTACTACATGAATATCCTGAAGCTGATCCG GCAAGAGGCCATCTACGAGCACCACGAGCAGGACCCCAAGAAT CCCAAGAAGGTGTCCAAGGCCGAGATCCAGGCCGAGCTGTGGG ATTTCGTGCTGAAGATGCAGAAGTGCAACAGCTTCACACACGAG GTGGACAAGGACGAGGTGTTCAACATCCTGAGAGAGCTGTACG AGGAACTGGTGCCCAGCAGCGTGGAAAAGAAGGGCGAAGCCAA CCAGCTGAGCAACAAGTTTCTGTACCCTCTGGTGGACCCCAACA GCCAGTCTGGAAAGGGAACAGCCAGCAGCGGCAGAAAGCCCA GATGGTACAACCTGAAGATTGCCGGCGATCCCTCCTGGGAAGA AGAGAAGAAGAAGTGGGAAGAAGATAAGAAAAAGGACCCGCTG GCCAAGATCCTGGGCAAGCTGGCTGAGTACGGACTGATCCCTC TGTTCATCCCCTACACCGACAGCAACGAGCCCATCGTGAAAGAA ATCAAGTGGATGGAAAAGTCCCGGAACCAGAGCGTGCGGCGGC TGGATAAGGACATGTTCATTCAGGCCCTGGAACGGTTCCTGAGC TGGGAGAGCTGGAACCTGAAAGTGAAAGAGGAATACGAGAAGG TCGAGAAAGAGTACAAGACCCTGGAAGAGAGGATCAAAGAGGA CATCCAGGCTCTGAAGGCTCTGGAACAGTATGAGAAAGAGCGG CAAGAACAGCTGCTGCGGGACACCCTGAACACCAACGAGTACC GGCTGAGCAAGAGAGGCCTTAGAGGCTGGCGGGAAATCATCCA GAAATGGCTGAAAATGGACGAGAACGAGCCCTCCGAGAAGTAC CTGGAAGTGTTCAAGGACTACCAGCGGAAGCACCCTAGAGAGG CCGGCGATTACAGCGTGTACGAGTTCCTGTCCAAGAAAGAGAAC CACTTCATCTGGCGGAATCACCCTGAGTACCCCTACCTGTACGC CACCTTCTGCGAGATCGACAAGAAAAAGAAGGACGCCAAGCAG CAGGCCACCTTCACACTGGCCGATCCTATCAATCACCCTCTGTG GGTCCGATTCGAGGAAAGAAGCGGCAGCAACCTGAACAAGTAC AGAATCCTGACCGAGCAGCTGCACACCGAGAAGCTGAAGAAAA AGCTGACAGTGCAGCTGGACCGGCTGATCTACCCTACAGAATCT GGCGGCTGGGAAGAGAAGGGCAAAGTGGACATTGTGCTGCTGC CCAGCCGGCAGTTCTACAACCAGATCTTCCTGGACATCGAGGAA AAGGGCAAGCACGCCTTCACCTACAAGGATGAGAGCATCAAGTT CCCTCTGAAGGGCACACTCGGCGGAGCCAGAGTGCAGTTCGAC AGAGATCACCTGAGAAGATACCCTCACAAGGTGGAAAGCGGCA ACGTGGGCAGAATCTACTTCAACATGACCGTGAACATCGAGCCT ACAGAGTCCCCAGTGTCCAAGTCTCTGAAGATCCACCGGGACG

		ACTTCCCCAAGGTGGTCAACTTCAAGCCCAAAGAACTGACCGAG TGGATCAAGGACAGCAAGGGCAAGAAACTGAAGTCCGGCATCG AGTCCCTGGAAATCGGCCTGAGAGTGATGAGCATCGACCTGGG ACAGAGACAGGCCGCTGCCGCCTCTATTTTCGAGGTGGTGGAT CAGAAGCCCGACATCGAAGGCAAGCTGTTTTTCCCAATCAAGGG CACCGAGCTGTATGCCGTGCACAGAGCCAGCTTCAACATCAAG CTGCCCGGCGAGACACTGGTCAAGAGCAGAGAAGTGCTGCGGA AGGCCAGAGAGGACAATCTGAAACTGATGAACCAGAAGCTCAAC TTCCTGCGGAACGTGCTGCACTTCCAGCAGTTCGAGGACATCAC CGAGAGAGAGAAGCGGGTCACCAAGTGGATCAGCAGACAAGAG AACAGCGACGTGCCCCTGGTGTACCAGGATGAGCTGATCCAGA TCCGCGAGCTGATGTACAAGCCTTACAAGGACTGGGTCGCCTTC CTGAAGCAGCTCCACAAGAGACTGGAAGTCGAGATCGGCAAAG AAGTGAAGCACTGGCGGAAGTCCCTGAGCGACGGAAGAAAGGG CCTGTACGGCATCTCCCTGAAGAACATCGACGAGATCGATCGGA CCCGGAAGTTCCTGCTGAGATGGTCCCTGAGGCCTACCGAACC TGGCGAAGTGCGTAGACTGGAACCCGGCCAGAGATTCGCCATC GACCAGCTGAATCACCTGAACGCCCTGAAAGAAGATCGGCTGA AGAAGATGGCCAACACCATCATCATGCACGCCCTGGGCTACTG CTACGACGTGCGGAAGAAGAAATGGCAGGCTAAGAACCCCGCC TGCCAGATCATCCTGTTCGAGGATCTGAGCAACTACAACCCCTA CGAGGAAAGGTCCCGCTTCGAGAACAGCAAGCTCATGAAGTGG TCCAGACGCGAGATCCCCAGACAGGTTGCACTGCAGGGCGAGA TCTATGGCCTGCAAGTGGGAGAAGTGGGCGCTCAGTTCAGCAG CAGATTCCACGCCAAGACAGGCAGCCCTGGCATCAGATGTAGC GTCGTGACCAAAGAGAAGCTGCAGGACAATCGGTTCTTCAAGAA TCTGCAGAGAGAGGGCAGACTGACCCTGGACAAAATCGCCGTG CTGAAAGAGGGCGATCTGTACCCAGACAAAGGCGGCGAGAAGT TCATCAGCCTGAGCAAGGATCGGAAGTGCGTGACCACACACGC CGACATCAACGCCGCTCAGAACCTGCAGAAGCGGTTCTGGACA AGAACCCACGGCTTCTACAAGGTGTACTGCAAGGCCTACCAGGT GGACGGCCAGACCGTGTACATCCCTGAGAGCAAGGACCAGAAG CAGAAGATCATCGAAGAGTTCGGCGAGGGCTACTTCATTCTGAA GGACGGGGTGTACGAATGGGTCAACGCCGGCAAGCTGAAAATC AAGAAGGGCAGCTCCAAGCAGAGCAGCAGCGAGCTGGTGGATA GCGACATCCTGAAAGACAGCTTCGACCTGGCCTCCGAGCTGAA AGGCGAAAAGCTGATGCTGTACAGGGACCCCAGCGGCAATGTG TTCCCCAGCGACAAATGGATGGCCGCTGGCGTGTTCTTCGGAA AGCTGGAACGCATCCTGATCAGCAAGCTGACCAACCAGTACTCC ATCAGCACCATCGAGGACGACAGCAGCAAGCAGTCTATGAAAA GGCCGGCGGCCACGAAAAAGGCCGGCCAGGCAAAAAAGAAAAA G
NLS	1313	MAPKKKRKVGIHGVPAA
NLS	1314	ATGGCCCCAAAGAAGAAGCGGAAGGTCGGTATCCACGGAGTCC CAGCAGCC
101 Cas9 TadAins 1015 폴리펩티드 서열	1315	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED ILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR

[1106]

		LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHK PENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVE NTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQ RKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNT KYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAY LNAVVGTALIKKYPKLESEFVYGDYKVGSSGSETPGTSESATPESS GSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGW NRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCA GAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGIL ADECAALLCYFFRMPRQVFNAQKKAQSSTDYDVRKMIAKSEQEIGK ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRD FATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDW DPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSS FEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQ KGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLD EIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTN LGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQL GGD
102 Cas9 TadAins 1022 폴리펩티드 서열	1316	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED ILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHK PENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVE NTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQ RKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNT KYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAY LNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIGSSGSETPGTSE SATPESSGSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNN RVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTF EPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNH RVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTDAKSEQEIG KATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGR DFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKD WDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMER SSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGE LQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHY LDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTL TNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLS QLGGD
103 Cas9 TadAins 1029 폴리펩티드 서열	1317	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG

[1107]

		LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED ILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHK PENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVE NTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQ RKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNT KYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAY LNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGSSGSE TPGTSESATPESSGSEVEFSHEYWMRHALTLAKRARDEREVPVGA VLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDA TLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYP GMNHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTDGKA TAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDF ATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDW DPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSS FEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQ KGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLD EIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTN LGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQL GGD
103 Cas9 TadAins 1040 폴리펩티드 서열	1318	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED ILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHK PENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVE NTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQ RKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNT KYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAY LNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAK YFFYSGSSGSETPGTSESATPESSGSEVEFSHEYWMRHALTLAKR ARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQG GLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTG AAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVFNA QKKAQSSTDNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGR DFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKD WDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMER SSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGE

[1108]

		LQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHY LDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTL TNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLS QLGGD
105 Cas9 TadAins 1068 폴리펩티드 서열	1319	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED ILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHK PENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVE NTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQ RKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNT KYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAY LNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAK YFFYSNIMNFFKTEITLANGEIRKRPLIETNGEGSSGSETPGTSESAT PESSGSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVI GEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPC VMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEI TEGILADECAALLCYFFRMPRQVFNAQKKAQSSTDTGEIVWDKGR DFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKD WDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMER SSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGE LQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHY LDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTL TNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLS QLGGD
106 Cas9 TadAins 1247 폴리펩티드 서열	1320	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED ILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHK PENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVE NTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQ RKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNT KYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAY

[1109]

		LNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAK YFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATV RKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK KYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEK NPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKG NELALPSKYVNFLYLASHYEKLKGSSSGSETPGTSESATPESSGSE VEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRA IGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMI HSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADEC AALLCYFFRMPRQVFNAQKKAQSSTDSPEDNEQKQLFVEQHKHYL DEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLT NLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQ LGGD
107 Cas9 TadAins 1054 폴리펩티드 서열	1321	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED ILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHK PENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVE NTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQ RKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNT KYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAY LNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAK YFFYSNIMNFFKTEITLANGSSGSETPGTSESATPESSGSEVEFSHE YWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDP TAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIG RVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLC YFFRMPRQVFNAQKKAQSSTDGEIRKRPLIETNGETGEIVWDKGRD FATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDW DPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSS FEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQ KGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLD EIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTN LGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQL GGD
108 Cas9 TadAins 1026 폴리펩티드 서열	1322	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI GALLFDSGETAEATRLKRTARRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE

[1110]

		DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED ILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHK PENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVE NTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQ RKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNT KYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAY LNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEGSSGSETP GTSESATPESSGSEVEFSHEYWMRHALTLAKRARDEREVPVGAVL VLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATL YVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPG MNHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTDQEIG KATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGR DFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKD WDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMER SSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGE LQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHY LDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTL TNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLS QLGGD
109 Cas9 TadAins 768 폴리펩티드 서열	1323	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED ILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHK PENIVIEMARENQGSSGSETPGTSESATPESSGSEVEFSHEYWMR HALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAE IMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFG VRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRM PRTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLY LYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLT RSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKA ERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLI REVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTAL IKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMN FFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQ VNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSP TVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAK GYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKY VNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKR VILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYF DTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD
110.1 Cas9 TadAins 1250 폴리펩티드 서열	1324	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV

[1111]

		QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED ILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHK PENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVE NTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQ RKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNT KYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAY LNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAK YFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATV RKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK KYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEK NPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKG NELALPSKYVNFLYLASHYEKLKGSPGSSGSETPGTSESATPESSG SEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWN RAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAG AMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILA DECAALLCYFFRMPREDNEQKQLFVEQHKHYLDEIIEQISEFSKRVI LADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDT TIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD
110.2 Cas9 TadAins 1250 폴리펩티드 서열	1325	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED ILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHK PENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVE NTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQ RKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNT KYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAY LNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAK YFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATV RKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK KYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEK NPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKG NELALPSKYVNFLYLASHYEKLKGSPGSSGSSGSETPGTSESATPE SSGSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGE GWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVM CAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITE

[1112]

		GILADECAALLCYFFRMPREDNEQKQLFVEQHKHYLDEIIEQISEFS KRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFK YFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD
110.3 Cas9 TadAins 1250 폴리펩티드 서열	1326	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED ILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHK PENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVE NTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQ RKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNT KYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAY LNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAK YFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATV
110.4 Cas9 TadAins 1250 폴리펩티드 서열	1327	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED ILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHK PENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVE NTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQ RKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNT KYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAY LNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAK YFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATV

[1113]

		RKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK KYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEK NPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKG NELALPSKYVNFLYLASHYEKLKGSPGSSGSSGSETPGTSESATPE SGSSSGSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRV IGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEP CVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRV EITEGILADECAALLCYFFRMRREDNEQKQLFVEQHKHYLDEIIEQIS EFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPA AFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD
110.5 Cas9 TadAins 1249 폴리펩티드 서열	1328	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED ILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHK PENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVE NTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQ RKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNT KYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAY LNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAK YFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATV RKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK KYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEK NPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKG NELALPSKYVNFLYLASHYEKLKGSGSSGSSGSETPGTSESATPES GSSSGSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVI GEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPC VMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEI TEGILADECAALLCYFFRMRRPEDNEQKQLFVEQHKHYLDEIIEQIS EFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPA AFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD
110.5 Cas9 TadAins 멀타 59-66 1250 폴리펩티드 서열	1329	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED ILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHK

[1114]

[1115]

		PENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVE NTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQ RKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNT KYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAY LNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAK YFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATV RKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK KYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEK NPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKG NELALPSKYVNFLYLASHYEKLKGSPGSSGSSGSETPGTSESATPE SGSSGSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVI GEGWNRAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAM IHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADE CAALLCYFFRMPRQVFNAQKKAQSSTDEDNEQKQLFVEQHKHYLD EIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTN LGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQL GGD
110.6 Cas9 TadAins 1251 폴리펩티드 서열	1330	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED ILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHK PENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVE NTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQ RKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNT KYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAY LNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAK YFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATV RKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK KYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEK NPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKG NELALPSKYVNFLYLASHYEKLKGSPEGSSGSSGSETPGTSESATP ESGSSSGSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNR VIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFE PCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHR VEITEGILADECAALLCYFFRMRRDNEQKQLFVEQHKHYLDEIIEQIS EFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPA AFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD
110.7 Cas9 TadAins 1252 폴리펩티드 서열	1331	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI

[1116]

		LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED ILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHK PENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVE NTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQ RKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNT KYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAY LNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAK YFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATV RKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK KYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEK NPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKG NELALPSKYVNFLYLASHYEKLKGSPEDGSSGSSGSETPGTSESAT PESGSSSGSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNN RVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTF EPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNH RVEITEGILADECAALLCYFFRMRRNEQKQLFVEQHKHYLDEIIEQIS EFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPA AFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD
110.8 Cas9 TadAins 멀타 59-66 C-절삭 1250 폴리펩티드 서열	1332	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED ILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHK PENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVE NTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQ RKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNT KYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAY LNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAK YFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATV RKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK KYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEK NPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKG NELALPSKYVNFLYLASHYEKLKGSPGSSGSETPGTSESATPESSG SEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWN RAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIG RVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLC YFFRMPRQEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLD KVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRY TSTKEVLDATLIHQSITGLYETRIDLSQLGGD

111.1 Cas9 TadAins 997 폴리펩티드 서열	1333	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED ILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHK PENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVE NTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQ RKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNT KYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAY LNAVVGTALSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIG EGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCV MCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEIT EGILADECAALLCYFFRMPRQVFNAQKKAQSSTDGSSGSETPGTS ESATPESSGIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAK YFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATV RKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK KYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEK NPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKG NELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEII EQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLG APAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGG D
111.2 Cas9 TadAins 997 폴리펩티드 서열	1334	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED ILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHK PENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVE NTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQ RKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNT KYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAY LNAVVGTALSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIG EGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCV MCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEIT EGILADECAALLCYFFRMPRQVFNAQKKAQSSTDGSSGSSGSETP

[1117]

		GTSESATPESSGGSSIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEI GKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKG RDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKK DWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIME RSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASA GELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQH KHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIH LFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRI DLSQLGGD
112 멀타 HNH TadA 폴리펩티드 서열	1335	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED ILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHK PENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSEVEFSHEY WMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPT AHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGR VVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCY FFRMPRQVFNAQKKAQSSTDGGLSELDKAGFIKRQLVETRQITKHV AQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREIN NYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKS EQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIV WDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKL IARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLG ITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRML ASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVE QHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAEN IIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYET RIDLSQLGGD
113 N-term 단일 TadA 나선 trunc 165-말단 폴리펩티드 서열	1336	MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGW NRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCA GAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGIL ADECAALLCYFFRMPRSGGSSGGSSGSETPGTSESATPESSGGSS GGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKK NLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAK VDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLR KKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQ LVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKK NGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQ IGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHH QDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFI KPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILR RQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEE TITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYF TVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQL KEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEEN EDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGW GRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKE

[1118]

		DIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGR HKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHP VENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSF LKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLI TQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRM NTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHD AYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKAT AKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFA TVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWD PKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSF EKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQ KGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLD EIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTN LGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQL GGD
114 N-term 단일 TadA 나선 trunc 165-말단 멜타 59-65 폴리펩티드 서열	1337	MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGW NRTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSR IGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAAL LCYFFRMPRSGGSSGGSSGSETPGTSESATPESSGGSSGGSDKK YSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALL FDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFF HRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDS TDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYN QLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGN LIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYA DLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLK ALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKM DGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFY PFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWN FEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNE LTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYF KKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILE DIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLS RKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKA QVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPE NIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENT QLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDD SIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRK FDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKY DENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLN AVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYF FYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRK VLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKY GGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNP IDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNE LALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQ ISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAP AAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD
115.1 Cas9 TadAins1004 폴리펩티드 서열	1338	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV

[1119]

		YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED ILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHK PENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVE NTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQ RKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNT KYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAY LNAVVGTALIKKYPKGSSGSETPGTSESATPESSGSEVEFSHEYW MRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTA HAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRV VFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYF FRMPRQLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIM NFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMP QVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDS PTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEA KGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSK YVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSK RVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKY FDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD
115.2 Cas9 TadAins1005 폴리펩티드 서열	1339	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED ILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHK PENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVE NTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQ RKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNT KYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAY LNAVVGTALIKKYPKLGSSGSETPGTSESATPESSGSEVEFSHEYW MRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTA HAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRV VFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYF FRMPRQESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMN FFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQ VNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSP TVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAK GYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKY VNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKR VILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYF DTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD
115.3 Cas9 TadAins1006 폴리펩티드 서열	1340	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV

[1120]

		QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED ILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHK PENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVE NTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQ RKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNT KYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAY LNAVVGTALIKKYPKLEGSSGSETPGTSESATPESSGSEVEFSHEY WMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPT AHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGR VVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCY FFRMPRQSEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMN FFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQ VNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSP TVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAK GYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKY VNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKR VILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYF DTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD
115.4 Cas9 TadAins1007 폴리펩티드 서열	1341	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED ILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHK PENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVE NTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQ RKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNT KYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAY LNAVVGTALIKKYPKLESGSSGSETPGTSESATPESSGSEVEFSHE YWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDP TAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIG RVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLC YFFRMPRQEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMN FFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQ VNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSP TVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAK GYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKY

[1121]

		VNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKR VILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYF DTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD
116.1 Cas9 TadAins C-term 절삭2 792 폴리펩티드 서열	1342	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED ILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHK PENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVE NTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLT RSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKA ERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLI REVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTAL IKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMN FFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQ VNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSP TVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAK GYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKY VNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKR VILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYF DTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD
116.2 Cas9 TadAins C-term 절삭2 791 폴리펩티드 서열	1343	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED ILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHK PENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSSGSETPGT SESATPESSGSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVL NNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYV TFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMN HRVEITEGILADECAALLCYFFRMPRQGSQILKEHPVENTQLQNEKL YLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLT RSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKA

[1122]

		ERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLI REVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTAL IKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMN FFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQ VNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSP TVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAK GYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKY VNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKR VILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYF DTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD
116.3 Cas9 TadAins C-term 절삭2 790 폴리펩티드 서열	1344	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED ILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHK PENIVIEMARENQTTQKGQKNSRERMKRIEEGIKEGSSGSETPGTS ESATPESSGSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLN NRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVT FEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNH RVEITEGILADECAALLCYFFRMPRQLGSQILKEHPVENTQLQNEKL YLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLT RSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKA ERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLI REVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTAL IKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMN FFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQ VNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSP TVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAK GYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKY VNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKR VILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYF DTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD
117 Cas9 델타 1017-1069 폴리펩티드 서열	1345	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED ILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHK

[1123]

[1124]

		PENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVE NTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQ RKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNT KYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAY LNAVVGTALIKKYPKLESEFVYGDYKVYSSGSEVEFSHEYWMRHAL TLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAEIMA LRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRN AKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQ VFNAQKKAQSSTDGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQ TGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVA KVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLII KLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKV LSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTS TKEVLDATLIHQSITGLYETRIDLSQLGGD
118 Cas9 TadA-CP116ins 1067 폴리펩티드 서열	1346	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI GALLFDSGETAEATRLKRTARRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED ILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHK PENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVE NTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQ RKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNT KYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAY LNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAK YFFYSNIMNFFKTEITLANGEIRKRPLIETNMNHRVEITEGILADECAA LLCYFFRMPRQVFNAQKKAQSSTDGSSGSETPGTSESATPESSGS EVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNR AIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGA MIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGGETGEIVWDKGRD FATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDW DPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSS FEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQ KGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLD EIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTN LGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQL GGD
119 Cas9 TadAins 701 폴리펩티드 서열	1347	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI GALLFDSGETAEATRLKRTARRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ

		EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED ILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSGSSGSET PGTSESATPESSGSEVEFSHEYWMRHALTLAKRARDEREVPVGAV LVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDAT LYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYP GMNHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTDLTF KEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVM GRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKE HPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQ SFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNA KLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDS RMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHA HDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRD FATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDW DPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSS FEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQ KGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLD EIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTN LGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQL GGD
120 Cas9 TadACP136ins 1248 폴리펩티드 서열	1348	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED ILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHK PENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVE NTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQ RKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNT KYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAY LNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAK YFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATV RKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK KYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEK NPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKG NELALPSKYVNFLYLASHYEKLKGSMNHRVEITEGILADECAALLCY FFRMPRQVFNAQKKAQSSTDGSSGSETPGTSESATPESSGSEVEF SHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGL HDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHS RIGRVVFGVRNAKTGAAGSLMDVLHYPGPEDNEQKQLFVEQHKHY LDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTL TNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLS QLGGD

[1125]

121 Cas9 TadACP136ins 1052 폴리펩티드 서열	1349	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED ILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHK PENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVE NTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQ RKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNT KYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAY LNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAK YFFYSNIMNFFKTEITLAMNHRVEITEGILADECAALLCYFFRMPRQV FNAQKKAQSSTDGSSGSETPGTSESATPESSGSEVEFSHEYWMR HALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAE IMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFG VRNAKTGAAGSLMDVLHYPGNGEIRKRPLIETNGETGEIVWDKGRD FATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDW DPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSS FEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQ KGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLD EIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTN LGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQL GGD
122 Cas9 TadACP136ins 1041 폴리펩티드 서열	1350	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED ILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHK PENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVE NTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQ RKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNT KYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAY LNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAK YFFYSMNHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSST DGSSGSETPGTSESATPESSGSEVEFSHEYWMRHALTLAKRARDE REVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVM

[1126]

		QNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGS LMDVLHYPGNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGR DFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKD WDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMER SSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGE LQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHY LDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTL TNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLS QLGGD
123 Cas9 TadACP139ins 1299 폴리펩티드 서열	1351	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED ILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHK PENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVE NTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQ RKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNT KYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAY LNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAK YFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATV RKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK KYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEK NPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKG NELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEII EQISEFSKRVILADANLDKVLSAYNKHRMNHRVEITEGILADECAALL CYFFRMPRQVFNAQKKAQSSTDGSSGSETPGTSESATPESSGSEV EFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAI GLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMI HSRIGRVVFGVRNAKTGAAGSLMDVLHYPGDKPIREQAENIIHLFTL TNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLS QLGGD
124 Cas9 멀타 792-872 TadAins 폴리펩티드 서열	1352	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED ILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHK

[1127]

[1128]

		PENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSEVEFSHEY WMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPT AHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGR VVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCY FFRMPRQVFNAQKKAQSSTDEEVVKKMKNYWRQLLNAKLITQRKF DNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYD ENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNA VVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFF YSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKV LSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYG GFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPI DFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNEL ALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQI SEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAP AAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD
125 Cas9 멀타 792-906 TadAins 폴리펩티드 서열	1353	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED ILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHK PENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSEVEFSHEY WMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPT AHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGR VVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCY FFRMPRQVFNAQKKAQSSTDGLSELDKAGFIKRQLVETRQITKHVA QILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINN YHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSE QEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVW DKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIA RKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGIT IMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLA SAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVE QHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAEN IIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYET RIDLSQLGGD
126 TadA CP65ins 1003 폴리펩티드 서열	1354	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED

		ILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHK PENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVE NTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQ RKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNT KYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAY LNAVVGTALIKKYPKTAHAEIMALRQGGLVMQNYRLIDATLYVTFEP CVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRV EITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTDGSSGSETPG TSESATPESSGSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLV LNNRVIGEGWNRAIGLHDPLESEFVYGDYKVYDVRKMIAKSEQEIG KATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGR DFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKD WDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMER SSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGE LQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHY LDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTL TNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLS QLGGD
127 TadA CP65ins 1016 폴리펩티드 서열	1355	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED ILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHK PENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVE NTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQ RKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNT KYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAY LNAVVGTALIKKYPKLESEFVYGDYKVTAHAEIMALRQGGLVMQNY RLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMD VLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSS TDGSSGSETPGTSESATPESSGSEVEFSHEYWMRHALTLAKRARD EREVPVGAVLVLNNRVIGEGWNRAIGLHDPYDVRKMIAKSEQEIGK ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRD FATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDW DPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSS FEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQ KGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLD EIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTN LGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQL GGD
128 TadA CP65ins 1022 폴리펩티드 서열	1356	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV

[1129]

		QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED ILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHK PENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVE NTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQ RKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNT KYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAY LNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMITAHAEIMALRQGG LVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGA AGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVFNAQ KKAQSSTDGSSGSETPGTSESATPESSGSEVEFSHEYWMRHALTL AKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPAKSEQEIGKA TAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDF ATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDW DPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSS FEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQ KGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLD EIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTN LGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQL GGD
129 TadA CP65ins 1029 폴리펩티드 서열	1357	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED ILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHK PENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVE NTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQ RKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNT KYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAY LNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEITAHAEI MALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGV RNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMP RQVFNAQKKAQSSTDGSSGSETPGTSESATPESSGSEVEFSHEY WMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPG KATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGR DFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKD WDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMER

[1130]

		SSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGE LQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHY LDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTL TNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLS QLGGD
130 TadA CP65ins 1041 폴리랩티드 서열	1358	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED ILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHK PENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVE NTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQ RKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNT KYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAY LNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAK YFFYSTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMI HSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADEC AALLCYFFRMPRQVFNAQKKAQSSTDGSSGSETPGTSESATPESS GSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGW NRAIGLHDPNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRD FATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDW DPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSS FEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQ KGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLD EIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTN LGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQL GGD
131 TadA CP65ins 1054 폴리랩티드 서열	1359	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED ILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHK PENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVE NTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQ RKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNT

[1131]

		KYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAY LNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAK YFFYSNIMNFFKTEITLANTAHAEIMALRQGGLVMQNYRLIDATLYVT FEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNH RVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTDGSSGSET PGTSESATPESSGSEVEFSHEYWMRHALTLAKRARDEREVPVGAV LVLNNRVIGEGWNRAIGLHDPGEIRKRPLIETNGETGEIVWDKGRDF ATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDW DPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSS FEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQ KGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLD EIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTN LGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQL GGD
132 TadA CP65ins 1246 폴리펩티드 서열	1360	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNG LFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQD LTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT PWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED ILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHK PENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVE NTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQ RKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNT KYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAY LNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAK YFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATV RKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK KYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEK NPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKG NELALPSKYVNFLYLASHYEKLKGTAHAEIMALRQGGLVMQNYRLI DATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVL HYPGMNHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTD GSSGSETPGTSESATPESSGSEVEFSHEYWMRHALTLAKRARDER EVPVGAVLVLNNRVIGEGWNRAIGLHDPSPEDNEQKQLFVEQHKH YLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFT LTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLS QLGGD
TadA 폴리펩티드 서열	1363	MGSHMTNDIYFMTLAIEEAKKAAQLGEVPIGAIITKDDEVIARAHNLR ETLQQPTAHAEHIAIERAAKVLGSWRLEGCTLYVTLEPCVMCAGTIV MSRIPRVVYGADDPKGGCSGSLMNLLQQSNFNHRAIVDKGVLKEA CSTLLTTFFKNLRANKKSTN
TadA 폴리펩티드 서열	1364	MTQDELYMKEAIKEAKKAEEKGEVPIGAVLVINGEIIARAHNLRETEQ RSIAHAEMLVIDEACKALGTWRLEGATLYVTLEPCPMCAGAVVLSR VEKVVFGAFDPKGGCSGTLMNLLQEERFNHQAEVVSGVLEEECG GMLSAFFRELRKKKKAARKNLSE
TadA 폴리펩티드 서열	1365	MPPAFITGVTSLSDVELDHEYWMRHALTLAKRAWDEREVPVGAVL VHNHRVIGEGWNRPIGRHDPTAHAEIMALRQGGLVLQNYRLLDTTL YVTLEPCVMCAGAMVHSRIGRVVFGARDAKTGAAGSLIDVLHHPG

[1132]

		MNHRVEIIEGVLRDECATLLSDFFRMRRQEIKALKKADRAEGAGPA V
TadA 폴리펩티드 서열	1366	MDEYWMQVAMQMAEKAEAAGEVPVGAVLVKDGQQIATGYNLSIS QHDPTAHAEILCLRSAGKKLENYRLLDATLYITLEPCAMCAGAMVH SRIARVVYGARDEKTGAAGTVVNLLQHPAFNHQVEVTSGVLAEAC SAQLSRFFKRRRDEKKALKLAQRAQQGIE
TadA 폴리펩티드 서열	1367	MDAAKVRSEFDEKMMRYALELADKAEALGEIPVGAVLVDDARNIIG EGWNLSIVQSDPTAHAEIIALRNGAKNIQNYRLLNSTLYVTLEPCTM CAGAILHSRIKRLVFGASDYKTGAIGSRFHFFDDYKMNHTLEITSGV LAEECSQKLSTFFQKRREEKKIEKALLKSLSDK
TadA 폴리펩티드 서열	1368	MRTDESEDQDHRMMRLALDAARAAAEAGETPVGAVILDPSTGEVI ATAGNGPIAAHDPTAHAEIAAMRAAAAKLGNYRLTDLTLVVTLEPCA MCAGAISHARIGRVVFGADDPKGGAVVHGPKFFAQPTCHWRPEVT GGVLADESADLLRGFFRARRKAKI
TadA 폴리펩티드 서열	1369	MSSLKKTPIRDDAYWMGKAIREAAKAAARDEVPIGAVIVRDGAVIGR GHNLREGSNDPSAHAEMIAIRQAARRSANWRLTGATLYVTLEPCL MCMGAIILARLERVVFGCYDPKGGAAGSLYDLSADPRLNHQVRLSP GVCQEECGTMLSDFFRDLRRRKKAKATPALFIDERKVPPEP
ecTadA 폴리펩티드 서열	1370	MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGW NRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCA GAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGIL ADECAALLCYFFRMPRQVFNAQKKAQSSTD
TadA*7.10	8	MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGW NRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCA GAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGIL ADECAALLCYFFRMPRQVFNAQKKAQSSTD
TadA*8	12	MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGW NRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCA GAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGIL ADECAALLCTFFRMPRQVFNAQKKAQSSTD
gRNA 스캐폴드 뉴클레오티드 서열	224	GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGU UAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU
gRNA 스캐폴드 뉴클레오티드 서열	225	GUUUUUGUACUCUCAAGAUUUAAGUAACUGUACAACGAAACUU ACACAGUUACUUAAAUCUUGCAGAAGCUACAAAGAUAAGGCUU CAUGCCGAAAUCAACACCCUGUCAUUUUAUGGCAGGGUG
S. 피오게네스 gRNA 스캐폴드 뉴클레오티드 서열	226	GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGU UAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC
S. 아우레우스 gRNA 스캐폴드 뉴클레오티드 서열	227	GUUUUAGUACUCUGUAAUGAAAAUUACAGAAUCUACUAAAACA AGGCAAAAUGCCGUGUUUAUCUCGUCAACUUGUUGGCGAGA
BhCas12b gRNA 스캐폴드 뉴클레오티드 서열	228	GUUCUGUCUUUUGGUCAGGACAACCGUCUAGCUAUAAGUGCU GCAGGGUGUGAGAAACUCCUAUUGCUGGACGAUGUCUCUUAC GAGGCAUUAGCAC
BvCas12b gRNA 스캐폴드 뉴클레오티드 서열	229	GACCUAUAGGGUCAAUGAAUCUGUGCGUGUGCCAUAAGUAAU UAAAAAUUACCCACCACAGGAGCACCUGAAAACAGGUGCUUGG CAC
gRNA 스캐폴드 뉴클레오티드 서열	230	GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGU UAUCAACUUGAAAAAGUGGGACCGAGUCGGUGCUUUU
gRNA 스캐폴드 뉴클레오티드 서열	3000	GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGU UAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU

[1133]

BhCas12b gRNA 스캐폴드 + 가이드 서열	243	GUUCUGUCUUUUGGUCAGGACAACCGUCUAGCUAUAAGUGCU GCAGGGUGUGAGAAACUCCUAUUGCUGGACGAUGUCUCUUAC GAGGCAUUAGCACNNNNNNNNNNNNNNNNNNNN
BvCas12b gRNA 스캐폴드 + 가이드 서열	244	GACCUAUAGGGUCAAUGAAUCUGUGCGUGUGCCAUAAGUAAU UAAAAAUUACCCACCACAGGAGCACCUGAAAACAGGUGCUUGG CACNNNNNNNNNNNNNNNNNNN
AaCas12b gRNA 스캐폴드 + 가이드 서열	245	GUCUAAAGGACAGAAUUUUUCAACGGGUGUGCCAAUGGCCAC UUUCCAGGUGGCAAAGCCCGUUGAACUUCUCAAAAAGAACGAU CUGAGAAGUGGCACNNNNNNNNNNNNNNNNNNNN
SpyMacCas9 폴리펩티드 서열	1307	MDKKYSIGLDIGTNSVGWAVITDDYKVPSKKFKVLGNTDRHSIKKNL IGALLFGSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LADSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV QIYNQLFEENPINASRVDAKAILSARLSKSRRLENLIAQLPGEKRNGL FGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGD QYADLFLAAKNLSDAILLSDILRVNSEITKAPLSASMIKRYDEHHQDL TLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPIL EKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQE DFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITP WNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVY NELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKED YFKKIECFDSVEISGVEDRFNASLGAYHDLLKIIKDKDFLDNEENEDI LEDIVLTLTLFEDRGMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGHSLHEQIANLAGSPAIKKGILQTVKIVDELVKVMGHKPE NIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENT QLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFIKDD SIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRK FDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKY DENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLN AVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYF FYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRK VLSMPQVNIVKKTEIQTVGQNGGLFDDNPKSPLEVTPSKLVPLKKEL NPKKYGGYQKPTTAYPVLLITDTKQLIPISVMNKKQFEQNPVKFLRD RGYQQVGKNDFIKLPKYTLVDIGDGIKRLWASSKEIHKGNQLVVSKK SQILLYHAHHLDSDLSNDYLQNHNQQFDVLFNEIISFSKKCKLGKEHI QKIENVYSNKKNSASIEELAESFIKLLGFTQLGATSPFNFLGVKLNQK QYKGKKDYILPCTEGTLIRQSITGLYETRVDLSKIGED
NLS	83	PKKKRKVEGADKRTADGSEFESPKKKRKV
NLS	84	KRTADGSEFESPKKKRKV
NLS	85	KRPAATKKAGQAKKKK
NLS	86	KKTELQTTNAENKTKKL
NLS	87	KRGINDRNFWRGENGRKTR
NLS	1424	RKSGKIAAIVVKRPRKPKKKRKV
NLS	90	MDSLLMNRRKFLYQFKNVRWAKGRRETYLC
링커	1425	(SGGS)2
pNMG-B335 ABE8.1_Y147T_CP5_NGC PAM_단량체 폴리펩티드 서열	1426	MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGW NRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCA GAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGIL ADECAALLCTFFRMPRQVFNAQKKAQSSTDSGGSSGGSSGSETP GTSESATPESSGGSSGGSEIGKATAKYFFYSNIMNFFKTEITLANGE IRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTG GFSKESILPKRNSDKLIARKKDWDPKKYGGFMQPTVAYSVLVVAKV EKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKL PKYSLFELENGRKRMLASAKFLQKGNELALPSKYVNFLYLASHYEK

[1134]

		LKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLS AYNKHRDKPIREQAENIIHLFTLTNLGAPRAFKYFDTTIARKEYRSTK EVLDATLIHQSITGLYETRIDLSQLGGDGGSGGSGGSGGSGGSGG SGGMDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIK KNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMA KVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHL RKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFI QLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEK KNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLA QIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEH HQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYK FIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAIL RRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKS EETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLY EYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVK QLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNE ENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYT GWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTF KEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVM GRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKE HPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQ SFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNA KLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDS RMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHA HDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEGA DKRTADGSEFESPKKKRKV
NGC PAM CP5 폴리펩티드 서열을 갖는 pNMG-357_ABE8.14	1427	MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEG WNRPIGRHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMC AGAMIHSRIGRVVFGARDAKTGAAGSLMDVLHHPGMNHRVEITEGI LADECAALLSDFFRMRRQEIKAQKKAQSSTDGGSSGGSSGSETPG TSESATPESSGGSSGGSMSEVEFSHEYWMRHALTLAKRARDERE VPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQN YRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLM DVLHYPGMNHRVEITEGILADECAALLCTFFRMPRQVFNAQKKAQS STDSGGSSGGSSGSETPGTSESATPESSGGSSGGSEIGKATAKYF FYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRK VLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKY GGFMQPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKN PIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAKFLQKGNE LALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQ ISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAP RAFKYFDTTIARKEYRSTKEVLDATLIHQSITGLYETRIDLSQLGGDG GSGGSGGSGGSGGSGGSGGMDKKYSIGLAIGTNSVGWAVITDEY KVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRY TRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIF GNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGH FLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSAR LSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDA KLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNT EITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKN GYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRT FDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVG PLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNF DKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQ KKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASL GTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYA HLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSD GFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAI KKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRE

[1135]

		RMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVD QELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVP SEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFI KRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSD FRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYG DYKVYDVRKMIAKSEQEGADKRTADGSEFESPKKKRKV
ABE8.8-m 폴리펩티드 서열	1428	MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGW NRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCA GAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHHPGMNHRVEITEGIL ADECAALLCRFFRMPRRVFNAQKKAQSSTDSGGSSGGSSGSETP GTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPS KKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRK NRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIV DEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIE GDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSK SRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQ LSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITK APLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYA GYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDN GSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLA RGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKN LPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAI VDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYH DLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFD DKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFAN RNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGI LQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMK RIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEE VVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQ LVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRK DFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYK VYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLI ETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKE SILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKS KKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSL FELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSP EDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDA TLIHQSITGLYETRIDLSQLGGDEGADKRTADGSEFESPKKKRKV
ABE8.8-d 폴리펩티드 서열	1429	MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEG WNRPIGRHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMC AGAMIHSRIGRVVFGARDAKTGAAGSLMDVLHHPGMNHRVEITEGI LADECAALLSDFFRMRRQEIKAQKKAQSSTDSGGSSGGSSGSETP GTSESATPESSGGSSGGSSEVEFSHEYWMRHALTLAKRARDERE VPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQN YRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLM DVLHHPGMNHRVEITEGILADECAALLCRFFRMPRRVFNAQKKAQS STDSGGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGLAI GTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGET AEATRLKRTARRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEES FLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADL RLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEEN PINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLG LTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAK NLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQ LPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEEL LVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDN REKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVD

[1136]

		KGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYV TEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDS VEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLF EDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRD KQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGD SLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARE NQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLY LYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLT RSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKA ERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLI REVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTAL IKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMN FFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQ VNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSP TVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAK GYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKY VNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKR VILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYF DTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDEGADKRT ADGSEFESPKKKRKV
ABE8.13-m 폴리펩티드 서열	1430	MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGW NRAIGLHDPTAHAEIMALRQGGLVMQNYRLYDATLYVTFEPCVMCA GAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHHPGMNHRVEITEGIL ADECAALLCRFFRMPRRVFNAQKKAQSSTDSGGSSGGSSGSETP GTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPS KKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRK NRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIV DEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIE GDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSK SRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQ LSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITK APLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYA GYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDN GSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLA RGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKN LPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAI VDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYH DLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFD DKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFAN RNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGI LQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMK RIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEE VVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQ LVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRK DFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYK VYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLI ETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKE SILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKS KKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSL FELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSP EDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDA TLIHQSITGLYETRIDLSQLGGDEGADKRTADGSEFESPKKKRKV
ABE8.13-d 폴리펩티드 서열	1431	MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEG WNRPIGRHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMC AGAMIHSRIGRVVFGARDAKTGAAGSLMDVLHHPGMNHRVEITEGI LADECAALLSDFFRMRRQEIKAQKKAQSSTDSGGSSGGSSGSETP GTSESATPESSGGSSGGSSEVEFSHEYWMRHALTLAKRARDERE

[1137]

		VPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQN YRLYDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLM DVLHHPGMNHRVEITEGILADECAALLCRFFRMPRRVFNAQKKAQS STDSGGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGLAI GTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGET AEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEES FLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADL RLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEEN PINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLG LTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAK NLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQ LPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEEL LVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDN REKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVD KGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYV TEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDS VEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLF EDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRD KQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGD SLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARE NQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLY LYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLT RSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKA ERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLI REVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTAL IKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMN FFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQ VNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSP TVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAK GYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKY VNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKR VILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYF DTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDEGADKRT ADGSEFESPKKKRKV
ABE8.17-m 폴리펩티드 서열	1432	MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGW NRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYSTFEPCVMCA GAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGIL ADECAALLCYFFRMPRRVFNAQKKAQSSTDSGGSSGGSSGSETP GTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPS KKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRK NRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIV DEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIE GDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSK SRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQ LSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITK APLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYA GYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDN GSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLA RGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKN LPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAI VDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYH DLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFD DKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFAN RNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGI LQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMK RIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEE VVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQ LVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRK

[1138]

		DFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYK VYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLI ETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKE SILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKS KKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSL FELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSP EDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDA TLIHQSITGLYETRIDLSQLGGDEGADKRTADGSEFESPKKKRKV
ABE8.17-d 폴리펩티드 서열	1433	MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEG WNRPIGRHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMC AGAMIHSRIGRVVFGARDAKTGAAGSLMDVLHHPGMNHRVEITEGI LADECAALLSDFFRMRRQEIKAQKKAQSSTDSGGSSGGSSGSETP GTSESATPESSGGSSGGSSEVEFSHEYWMRHALTLAKRARDERE VPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQN YRLIDATLYSTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLM DVLHYPGMNHRVEITEGILADECAALLCYFFRMPRRVFNAQKKAQS STDSGGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGLAI GTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGET AEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEES FLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADL RLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEEN PINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLG LTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAK NLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQ LPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEEL LVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDN REKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVD KGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYV TEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDS VEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLF EDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRD KQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGD SLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARE NQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLY LYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLT RSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKA ERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLI REVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTAL IKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMN FFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQ VNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSP TVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAK GYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKY VNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKR VILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYF DTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDEGADKRT ADGSEFESPKKKRKV
ABE8.20-m 폴리펩티드 서열	1434	MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGW NRAIGLHDPTAHAEIMALRQGGLVMQNYRLYDATLYSTFEPCVMCA GAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHHPGMNHRVEITEGIL ADECAALLCRFFRMPRRVFNAQKKAQSSTDSGGSSGGSSGSETP GTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPS KKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRK NRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIV DEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIE GDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSK SRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQ LSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITK

[1139]

		APLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYA GYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDN GSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLA RGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKN LPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAI VDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYH DLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFD DKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFAN RNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGI LQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMK RIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEE VVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQ LVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRK DFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYK VYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLI ETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKE SILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKS KKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSL FELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSP EDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDA TLIHQSITGLYETRIDLSQLGGDEGADKRTADGSEFESPKKKRKV
ABE8.20-d 폴리펩티드 서열	1435	MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEG WNRPIGRHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMC AGAMIHSRIGRVVFGARDAKTGAAGSLMDVLHHPGMNHRVEITEGI LADECAALLSDFFRMRRQEIKAQKKAQSSTDSGGSSGGSSGSETP GTSESATPESSGGSSGGSSEVEFSHEYWMRHALTLAKRARDERE VPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQN YRLYDATLYSTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLM DVLHHPGMNHRVEITEGILADECAALLCRFFRMPRRVFNAQKKAQS STDSGGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGLAI GTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGET AEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEES FLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADL RLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEEN PINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLG LTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAK NLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQ LPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEEL LVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDN REKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVD KGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYV TEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDS VEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLF EDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRD KQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGD SLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARE NQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLY LYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLT RSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKA ERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLI REVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTAL IKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMN FFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQ VNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSP TVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAK GYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKY VNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKR

[1140]

		VILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYF DTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDEGADKRT ADGSEFESPKKKRKV
01. monoABE8.1_bpN LS + Y147T 폴리펩티드 서열	1436	MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGW NRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCA GAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGIL ADECAALLCTFFRMPRQVFNAQKKAQSSTDSGGSSGGSSGSETP GTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPS KKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRK NRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIV DEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIE GDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSK SRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQ LSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITK APLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYA GYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDN GSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLA RGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKN LPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAI VDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYH DLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFD DKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFAN RNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGI LQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMK RIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEE VVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQ LVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRK DFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYK VYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLI ETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKE SILPKRNSDKLIARKKDWDPKKYGGFVSPTVAYSVLVVAKVEKGKS KKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSL FELENGRKRMLASARELQKGNELALPSKYVNFLYLASHYEKLKGSP EDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKQYRSTKEVLDA TLIHQSITGLYETRIDLSQLGGDEGADKRTADGSEFESPKKKRKV
02. monoABE8.1_bpN LS + Y147R 폴리펩티드 서열	1437	MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGW NRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCA GAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGIL ADECAALLCRFFRMPRQVFNAQKKAQSSTDSGGSSGGSSGSETP GTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPS KKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRK NRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIV DEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIE GDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSK SRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQ LSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITK APLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYA GYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDN GSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLA RGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKN LPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAI VDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYH DLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFD DKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFAN RNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGI LQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMK RIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL

[1141]

		DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEE VVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQ LVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRK DFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYK VYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLI ETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKE SILPKRNSDKLIARKKDWDPKKYGGFVSPTVAYSVLVVAKVEKGKS KKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSL FELENGRKRMLASARELQKGNELALPSKYVNFLYLASHYEKLKGSP EDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKQYRSTKEVLDA TLIHQSITGLYETRIDLSQLGGDEGADKRTADGSEFESPKKKRKV
03. monoABE8.1_bpNLS + Q154S 폴리펩티드 서열	1438	MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGW NRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCA GAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGIL ADECAALLCYFFRMPRSVFNAQKKAQSSTDSGGSSGGSSGSETP GTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPS KKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRK NRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIV DEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIE GDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSK SRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQ LSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITK APLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYA GYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDN GSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLA RGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKN LPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAI VDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYH DLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFD DKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFAN RNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGI LQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMK RIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEE VVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQ LVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRK DFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYK VYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLI ETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKE SILPKRNSDKLIARKKDWDPKKYGGFVSPTVAYSVLVVAKVEKGKS KKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSL FELENGRKRMLASARELQKGNELALPSKYVNFLYLASHYEKLKGSP EDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKQYRSTKEVLDA TLIHQSITGLYETRIDLSQLGGDEGADKRTADGSEFESPKKKRKV
04. monoABE8.1_bpNLS + Y123H 폴리펩티드 서열	1439	MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGW NRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCA GAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHHPGMNHRVEITEGIL ADECAALLCYFFRMPRQVFNAQKKAQSSTDSGGSSGGSSGSETP GTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPS KKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRK NRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIV DEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIE GDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSK SRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQ LSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITK APLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYA GYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDN

[1142]

[1143]

		GSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLA RGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKN LPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAI VDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYH DLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFD DKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFAN RNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGI LQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMK RIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEE VVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQ LVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRK DFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYK VYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLI ETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKE SILPKRNSDKLIARKKDWDPKKYGGFVSPTVAYSVLVVAKVEKGKS KKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSL FELENGRKRMLASARELQKGNELALPSKYVNFLYLASHYEKLKGSP EDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKQYRSTKEVLDA TLIHQSITGLYETRIDLSQLGGDEGADKRTADGSEFESPKKKRKV
05. monoABE8.1_bpNLS + V82S 폴리펩티드 서열	1440	MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGW NRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYSTFEPCVMCA GAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGIL ADECAALLCYFFRMPRQVFNAQKKAQSSTDSGGSSGGSSGSETP GTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPS KKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRK NRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIV DEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIE GDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSK SRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQ LSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITK APLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYA GYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDN GSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLA RGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKN LPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAI VDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYH DLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFD DKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFAN RNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGI LQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMK RIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEE VVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQ LVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRK DFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYK VYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLI ETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKE SILPKRNSDKLIARKKDWDPKKYGGFVSPTVAYSVLVVAKVEKGKS KKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSL FELENGRKRMLASARELQKGNELALPSKYVNFLYLASHYEKLKGSP EDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKQYRSTKEVLDA TLIHQSITGLYETRIDLSQLGGDEGADKRTADGSEFESPKKKRKV
06. monoABE8.1_bpNLS + T166R 폴리펩티드 서열	1441	MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGW NRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCA GAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGIL ADECAALLCYFFRMPRQVFNAQKKAQSSRDSGGSSGGSSGSETP

[1144]

		GTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPS KKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRK NRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIV DEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIE GDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSK SRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQ LSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITK APLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYA GYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDN GSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLA RGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKN LPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAI VDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYH DLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFD DKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFAN RNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGI LQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMK RIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEE VVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQ LVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRK DFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYK VYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLI ETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKE SILPKRNSDKLIARKKDWDPKKYGGFVSPTVAYSVLVVAKVEKGKS KKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSL FELENGRKRMLASARELQKGNELALPSKYVNFLYLASHYEKLKGSP EDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKQYRSTKEVLDA TLIHQSITGLYETRIDLSQLGGDEGADKRTADGSEFESPKKKRKV
07. monoABE8.1_bpN LS + Q154R 폴리랩티드 서열	1442	MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGW NRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCA GAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGIL ADECAALLCYFFRMPRRVFNAQKKAQSSTDSGGSSGGSSGSETP GTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPS KKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRK NRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIV DEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIE GDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSK SRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQ LSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITK APLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYA GYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDN GSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLA RGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKN LPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAI VDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYH DLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFD DKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFAN RNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGI LQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMK RIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEE VVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQ LVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRK DFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYK VYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLI ETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKE SILPKRNSDKLIARKKDWDPKKYGGFVSPTVAYSVLVVAKVEKGKS

		KKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSL FELENGRKRMLASARELQKGNELALPSKYVNFLYLASHYEKLKGSP EDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKQYRSTKEVLDA TLIHQSITGLYETRIDLSQLGGDEGADKRTADGSEFESPKKKRKV
08. monoABE8.1_bpNLS + Y147R_Q154R_Y123H 폴리펩티드 서열	1443	MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGW NRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCA GAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHHPGMNHRVEITEGIL ADECAALLCRFFRMPRRVFNAQKKAQSSTDSGGSSGGSSGSETP GTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPS KKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRK NRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIV DEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIE GDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSK SRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQ LSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITK APLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYA GYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDN GSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLA RGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKN LPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAI VDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYH DLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFD DKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFAN RNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGI LQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMK RIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEE VVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQ LVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRK DFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYK VYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLI ETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKE SILPKRNSDKLIARKKDWDPKKYGGFVSPTVAYSVLVVAKVEKGKS KKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSL FELENGRKRMLASARELQKGNELALPSKYVNFLYLASHYEKLKGSP EDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKQYRSTKEVLDA TLIHQSITGLYETRIDLSQLGGDEGADKRTADGSEFESPKKKRKV
09. monoABE8.1_bpNLS + Y147R_Q154R_I76Y 폴리펩티드 서열	1444	MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGW NRAIGLHDPTAHAEIMALRQGGLVMQNYRLYDATLYVTFEPCVMCA GAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGIL ADECAALLCRFFRMPRRVFNAQKKAQSSTDSGGSSGGSSGSETP GTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPS KKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRK NRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIV DEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIE GDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSK SRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQ LSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITK APLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYA GYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDN GSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLA RGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKN LPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAI VDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYH DLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFD DKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFAN RNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGI

[1145]

		LQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMK RIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEE VVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQ LVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRK DFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYK VYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLI ETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKE SILPKRNSDKLIARKKDWDPKKYGGFVSPTVAYSVLVVAKVEKGKS KKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSL FELENGRKRMLASARELQKGNELALPSKYVNFLYLASHYEKLKGSP EDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKQYRSTKEVLDA TLIHQSITGLYETRIDLSQLGGDEGADKRTADGSEFESPKKKRKV
10. monoABE8.1_bpNLS + Y147R_Q154R_T166R 폴리펩티드 서열	1445	MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGW NRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCA GAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGIL ADECAALLCRFFRMPRRVFNAQKKAQSSRDSGGSSGGSSGSETP GTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPS KKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRK NRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIV DEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIE GDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSK SRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQ LSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITK APLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYA GYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDN GSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLA RGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKN LPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAI VDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYH DLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFD DKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFAN RNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGI LQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMK RIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEE VVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQ LVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRK DFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYK VYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLI ETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKE SILPKRNSDKLIARKKDWDPKKYGGFVSPTVAYSVLVVAKVEKGKS KKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSL FELENGRKRMLASARELQKGNELALPSKYVNFLYLASHYEKLKGSP EDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKQYRSTKEVLDA TLIHQSITGLYETRIDLSQLGGDEGADKRTADGSEFESPKKKRKV
11. monoABE8.1_bpNLS + Y147T_Q154R 폴리펩티드 서열	1446	MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGW NRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCA GAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGIL ADECAALLCTFFRMPRRVFNAQKKAQSSTDSGGSSGGSSGSETP GTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPS KKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRK NRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIV DEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIE GDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSK SRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQ LSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITK

[1146]

		APLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYA GYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDN GSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLA RGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKN LPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAI VDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYH DLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFD DKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFAN RNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGI LQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMK RIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEE VVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQ LVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRK DFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYK VYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLI ETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKE SILPKRNSDKLIARKKDWDPKKYGGFVSPTVAYSVLVVAKVEKGKS KKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSL FELENGRKRMLASARELQKGNELALPSKYVNFLYLASHYEKLKGSP EDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKQYRSTKEVLDA TLIHQSITGLYETRIDLSQLGGDEGADKRTADGSEFESPKKKRKV
12. monoABE8.1_bpN LS + Y147T_Q154S 폴리펩티드 서열	1447	MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGW NRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCA GAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGIL ADECAALLCTFFRMPRSVFNAQKKAQSSTDSGGSSGGSSGSETP GTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPS KKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRK NRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIV DEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIE GDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSK SRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQ LSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITK APLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYA GYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDN GSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLA RGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKN LPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAI VDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYH DLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFD DKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFAN RNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGI LQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMK RIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEE VVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQ LVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRK DFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYK VYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLI ETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKE SILPKRNSDKLIARKKDWDPKKYGGFVSPTVAYSVLVVAKVEKGKS KKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSL FELENGRKRMLASARELQKGNELALPSKYVNFLYLASHYEKLKGSP EDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKQYRSTKEVLDA TLIHQSITGLYETRIDLSQLGGDEGADKRTADGSEFESPKKKRKV
13. monoABE8.1_bpN	1448	MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGW NRAIGLHDPTAHAEIMALRQGGLVMQNYRLYDATLYVTFEPCVMCA

[1147]

[1148]

LS + H123Y123H_Y147R_Q154R_I76Y 폴리펩티드 서열		GAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHHPGMNHRVEITEGIL ADECAALLCRFFRMPRRVFNAQKKAQSSTDSGGSSGGSSGSETP GTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPS KKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRK NRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIV DEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIE GDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSK SRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQ LSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITK APLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYA GYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDN GSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLA RGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKN LPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAI VDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYH DLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFD DKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFAN RNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGI LQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMK RIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEE VVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQ LVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRK DFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYK VYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLI ETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKE SILPKRNSDKLIARKKDWDPKKYGGFVSPTVAYSVLVVAKVEKGKS KKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSL FELENGRKRMLASARELQKGNELALPSKYVNFLYLASHYEKLKGSP EDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKQYRSTKEVLDA TLIHQSITGLYETRIDLSQLGGDEGADKRTADGSEFESPKKKRKV
14. monoABE8.1_bpNLS + V82S + Q154R 폴리펩티드 서열	1449	MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGW NRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYSTFEPCVMCA GAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGIL ADECAALLCYFFRMPRRVFNAQKKAQSSTDSGGSSGGSSGSETP GTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPS KKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRK NRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIV DEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIE GDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSK SRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQ LSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITK APLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYA GYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDN GSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLA RGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKN LPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAI VDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYH DLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFD DKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFAN RNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGI LQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMK RIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEE VVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQ LVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRK DFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYK VYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLI

		ETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKE SILPKRNSDKLIARKKDWDPKKYGGFVSPTVAYSVLVVAKVEKGKS KKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSL FELENGRKRMLASARELQKGNELALPSKYVNFLYLASHYEKLKGSP EDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKQYRSTKEVLDA TLIHQSITGLYETRIDLSQLGGDEGADKRTADGSEFESPKKKRKV
링커	65	PAPAP
링커	66	PAPAPA
링커	67	PAPAPAP
링커	68	PAPAPAPA
링커	69	P(AP)4
링커	70	P(AP)7
링커	71	P(AP)10
N 유전자 (핵산)	4001	atggatgccgacaagattgtattcaaagtcaataatcaggtggtctctttgaagcctgagattatc gtggatcaatatgagtacaagtaccctgccatcaaagatttgaaaaagccctgtataaccctag gaaaggctcccgatttaaataaagcatacaagtcagttttgtcaggcatgagcgccgccaaac ttaatcctgacgatgtatgttcctatttggcagcggcaatgcagttttttgaggggacatgtccgga agactggaccagctatggaattgtgattgcacgaaaaggagataagatcaccccaggttctct ggtggagataaaacgtactgatgtagaagggaattgggctctgacaggaggcatggaactga caagagaccccactgtccctgagcatgcgtccttagtcggtcttctcttgagtctgtataggttgag caaaatatccgggcaaaacactggtaactataagacaaacattgcagacaggatagagcag attttgagacagccccttttgttaaaatcgtggaacaccatactctaatgacaactcacaaaatg tgtgctaattggagtactataccaaacttcagatttttggccggaacctatgacatgtttttctcccg gattgagcatctatattcagcaatcagagtgggcacagttgtcactgcttatgaagactgttcagg actggtatcatttactgggttcataaaacaaatcaatctcaccgctagagaggcaatactatattt cttccacaagaactttgaggaagagataagaagaatgtttgagccagggcaggagacagct gttcctcactcttatttcatccacttccgttcactaggcttgagtgggaaatctccttattcatcaaatg ctgttggtcacgtgttcaatctcattcactttgtaggatgctatatgggtcaagtcagatccctaaat gcaacggttattgctgcatgtgctcctcatgaaatgtctgttctagggggctatctgggagaggaa ttcttcgggaaagggacatttgaaagaagattcttcagagatgagaaagaacttcaagaatac gaggcggctgaactgacaaagactgacgtagcactggcagatgatggaactgtcaactctga cgacgaggactacttttcaggtgaaaccagaagtccggaggctgtttatactcgaatcatgatg aatggaggtcgactaaagagatctcacatacggagatatgtctcagtcagttccaatcatcaag cccgtccaaactcattcgccgagtttctaaacaagacatattcgagtgactca
N 유전자 (아미노산)	4002	MDADKIVFKVNNQVVSLKPEIIVDQYEYKYPAIKDLKKPCITLGKAPD LNKAYKSVLSGMSAAKLNPDDVCSYLAAAMQFFEGTCPEDWTSYG IVIARKGDKITPGSLVEIKRTDVEGNWALTGGMELTRDPTVPEHASL VGLLLSLYRLSKISGQNTGNYKTNIADRIEQIFETAPFVKIVEHHTLMT THKMCANWSTIPNFRFLAGTYDMFFSRIEHLYSAIRVGTVVTAYED CSGLVSFTGFIKQINLTAREAILYFFHKNFEEEIRRMFEPGQETAVPH SYFIHFRSLGLSGKSPYSSNAVGHVFNLIHFVGCYMGQVRSLNATVI AACAPHEMSVLGGYLGEEFFGKGTFERRFFRDEKELQEYEAAELT KTDVALADDGTVNSDDEDYFSGETRSPEAVYTRIMMNGGRLKRSH IRRYVSVSSNHQARPNSFAEFLNKTYSSDS
L 유전자 (핵산)	4003	ctcgatcctggagaggtctatgatgaccctattgacccaatcgagttagaggctgaacccagag gaacccccattgtccccaacatcttgaggaactctgactacaatctcaactctcctttgatagaag atcctgctagactaatgttagaatggttaaaaacagggaatagaccttatcggatgactctaaca gacaattgctccaggtctttcagagttttgaaagattatttcaagaaggtagatttgggttctctcaa ggtgggcggaatggctgcacagtcaatgatttctctctggttatatggtgcccactctgaatccaa caggagccggagatgtataacagacttggcccatttctattccaagtcgtcccccatagagaag ctgttgaatctcacgctaggaaatagagggctgagaatccccccagagggagtgttaagttgc cttgagagggttgattatgataatgcatttggaaggtatcttgccaacacgtattcctcttacttgttct tccatgtaatcaccttatacatgaacgccctagactgggatgaagaaaagaccatcctagcatt atggaaagatttaacctcagtggacatcgggaaggacttggtaaagttcaaagaccaaatatg

[1149]

[1150]

		gggactgctgatcgtgacaaaggactttgtttactcccaaagttccaattgtctttttgacagaaac tacacacttatgctaaaagatcttttcttgtctcgcttcaactccttaatggtcttgctctctcccccag agccccgatactcagatgacttgatatctcaactatgccagctgtacattgctggggatcaagtct tgtctatgtgtggaaactccggctatgaagtcatcaaaatattggagccatatgtcgtgaatagttt agtccagagagcagaaaagtttaggcctctcattcattccttgggagactttcctgtatttataaaa gacaaggtaagtcaacttgaagagacgttcggtccctgtgcaagaaggttctttagggctctgg atcaattcgacaacatacatgacttggtttttgtgtttggctgttacaggcattgggggcacccatat atagattatcgaaagggtctgtcaaaactatatgatcaggttcaccttaaaaaaatgatagataa gtcctaccaggagtgcttagcaagcgacctagccaggaggatccttagatggggttttgataag tactccaagtggtatctggattcaagattcctagcccgagaccacccctgactccttatatcaaa acccaaacatggccacccaaacatattgtagacttggtgggggatacatggcacaagctccc gatcacgcagatctttgagattcctgaatcaatggatccgtcagaaatattggatgacaaatcac attctttcaccagaacgagactagcttcttggctgtcagaaaaccgaggggggcctgttcctagc gaaaaagttattatcacggccctgtctaagccgcctgtcaatccccgagagtttctgaggtctata gacctcggaggattgccagatgaagacttgataattggcctcaagccaaaggaacgggaatt gaagattgaaggtcgattctttgctctaatgtcatggaatctaagattgtattttgtcatcactgaaa aactcttggccaactacatcttgccacttttgacgcgctgactatgacagacaacctgaacaag gtgtttaaaaagctgatcgacagggtcaccgggcaagggcttttggactattcaagggtcacat atgcatttcacctggactatgaaaagtggaacaaccatcaaagattagagtcaacagaggatg tattttctgtcctagatcaagtgtttggattgaagagagtgttttctagaacacacgagttttttcaaa aggcctggatctattattcagacagatcagacctcatcgggttacgggaggatcaaatatactg cttagatgcgtccaacggcccaacctgttggaatggccaggatggcgggctagaaggcttacg gcagaagggctggagtctagtcagcttattgatgatagatagagaatctcaaatcaggaacac aagaaccaaaatactagctcaaggagacaaccaggttttatgtccgacatacatgttgtcgcc agggctatctcaagaggggctcctctatgaattggagagaatatcaaggaatgcactttcgatat acagagccgtcgaggaaggggcatctaagctagggctgatcatcaagaaagaagagacca tgtgtagttatgacttcctcatctatggaaaaaccccttgtttagaggtaacatattggtgcctgagt ccaaaagatgggccagagtctcttgcgtctctaatgaccaaatagtcaacctcgccaatataat gtcgacagtgtccaccaatgcgctaacagtggcacaacactctcaatctttgatcaaaccgatg agggattttctgctcatgtcagtacaggcagtctttcactacctgctatttagcccaatcttaaaggg aagagtttacaagattctgagcgctgaaggggagagctttctcctagccatgtcaaggataatct atctagatccttctttgggagggatatctggaatgtccctcggaagattccatatacgacagttctc agaccctgtctctgaagggttatccttctggagagagatctggttaagctcccaagagtcctgga ttcacgcgttgtgtcaagaggctggaaacccagatcttggagagagaacactcgagagcttca ctcgccttctagaagatccgaccaccttaaatatcagaggaggggccagtcctaccattctactc aaggatgcaatcagaaaggctttatatgacgaggtggacaaggtggaaaattcagagtttcga gaggcaatcctgttgtccaagacccatagagataattttatactcttcttaatatctgttgagcctct gtttcctcgatttctcagtgagctattcagttcgtctttttgggaatccccgagtcaatcattggattg atacaaaactcccgaacgataagaaggcagtttagaaagagtctctcaaaaactttagaaga atccttctacaactcagagatccacgggattagtcggatgacccagacacctcagagggttgg gggggtgtggccttgctcttcagagagggcagatctacttagggagatctcttggggaagaaa agtggtaggcacgacagttcctcaccttctgagatgttgggattacttcccaagtcctctatttctt gcacttgtggagcaacaggaggaggcaatcctagagtttctgtatcagtactcccgtcctttgatc agtcatttttttcacgaggccccctaaagggatacttgggctcgtccacctctatgtcgacccagct attccatgcatgggaaaaagtcactaatgttcatgtggtgaagagagctctatcgttaaaagaat ctataaactggttcattactagagattccaacttggctcaagctctaattaggaacattatgtctctg acaggccctgatttccctctagaggaggcccctgtcttcaaaaggacggggtcagccttgcata ggttcaagtctgccagatacagcgaaggagggtattcttctgtctgcccgaacctcctctctcata tttctgttagtacagacaccatgtctgatttgacccaagacgggaagaactacgatttcatgttcc agccattgatgctttatgcacagacatggacatcagagctggtacagagagacacaaggcta agagactctacgtttcattggcacctccgatgcaacaggtgtgtgagacccattgacgacgtga ccctggagacctctcagatcttcgagtttccggatgtgtcgaaaagaatatccagaatggtttctg gggctgtgcctcacttccagaggcttcccgatatccgtctgagaccaggagattttgaatctctaa gcggtagagaaaagtctcaccatatcggatcagctcaggggctcttatactcaatcttagtggc aattcacgactcaggatacaatgatggaaccatcttccctgtcaacatatacggcaaggtttccc ctagagactatttgagagggctcgcaaggggagtattgataggatcctcgatttgcttcttgacaa gaatgacaaatatcaatattaatagacctcttgaattggtctcaggggtaatctcatatattctcctg aggctagataaccatccctccttgtacataatgctcagagaaccgtctcttagaggagagatatt ttctatcctcagaaaatccccgccgcttatccaaccactatgaaagaagcaacagatcaatc

		ttgtgttatctccaacatgtgctacgctatgagcgagagataatcacggcgtctccagagaatga ctggctatggatcttttcagactttagaagtgccaaaatgacgtacctatccctcattacttaccagt ctcatcttctactccagagggttgagagaaacctatctaagagtatgagagataacctgcgaca attgagttctttgatgaggcaggtgctgggcgggcacggagaagatacctagagtcagacga caacattcaacgactgctaaaagactctttacgaaggacaagatgggtggatcaagaggtgc gccatgcagctagaaccatgactggagattacagccccaacaagaaggtgtcccgtaaggta ggatgttcagaatgggtctgctctgctcaacaggttgcagtctctacctcagcaaacccggcccc tgtctcggagcttgacataagggccctctctaagaggttccagaacccttttgatctcgggcttgag agtggttcagtgggcaaccggtgctcattataagcttaagcctattctagatgatctcaatgttttcc catctctctgccttgtagttggggacgggtcaggggggatatcaagggcagtcctcaacatgttt ccagatgccaagcttgtgttcaacagtctttagaggtgaatgacctgatggcttccggaacaca tccactgcctccttcagcaatcatgaggggaggaaatgatatcgtctccagagtgatagatcttg actcaatctgggaaaaaccgtccgacttgagaaacttggcaacctggaaatacttccagtcagt ccaaaagcaggtcaacatgtcctatgacctcattatttgcgatgcagaagttactgacattgcatc tatcaaccggatcaccctgttaatgtccgattttgcattgtctatagatggaccactctatttggtcttc aaaacttatgggactatgctagtaaatccaaactacaaggcttattcaacacctgtcaagagcgt tcccctcggtcacagggtttatcacccaagtaacttcgtctttttcatctgagctctacctccgattct ccaaacgagggaagtttttcagagatgctgagtacttgacctcttccacccttcgagaaatgagc cttgtgttattcaattgtagcagccccaagagtgagatgcagagagctcgttccttgaactatcag gatcttgtgagaggatttcctgaagaaatcatatcaaatccttacaatgagatgatcataactctg attgacagtgatgtagaatctttttctagtccacaagatggttgatgatcttgagttacagaggggga actctgtctaaagtggctatcattatagccatcatgatagttttctccaacagagtcttcaacgtttcc aaaccccctaactgaccctcgttctatccaccgtctgatcccaaaatcctgaggcacttcaacat atgttgcagtactatgatgtatctatctactgctttaggtgacgtccctagcttcgcaagacttcacg acctgtataacagacctataacttattacttcagaaagcaagtcattcgagggaacgtttatctat cttggagttggtccaacgacacctcagtgttcaaaagggtagcctgtaattctagcctgagtctgt catctcactggatcaggttgatttacaagatagtgaagactaccagactcgttggcagcatcaa ggatctatccagagaagtggaaagacaccttcataggtacaacaggtggatcaccctagagg atatcagatctagatcatccctactagactacagttgcctg
L 유전자 (아미노산)	4004	LDPGEVYDDPIDPIELEAEPRGTPIVPNILRNSDYNLNSPLIEDPARL MLEWLKTGNRPYRMTLTDNCSRSFRVLKDYFKKVDLGSLKVGGMA AQSMISLWLYGAHSESNRSRRCITDLAHFYSKSSPIEKLLNLTLGNR GLRIPPEGVLSCLERVDYDNAFGRYLANTYSSYLFFHVITLYMNALD WDEEKTILALWKDLTSVDIGKDLVKFKDQIWGLLIVTKDFVYSQSSN CLFDRNYTLMLKDLFLSRFNSLMVLLSPPEPRYSDDLISQLCQLYIA GDQVLSMCGNSGYEVIKILEPYVVNSLVQRAEKFRPLIHSLGDFPVF IKDKVSQLEETFGPCARRFFRALDQFDNIHDLVFVFGCYRHWGHPY IDYRKGLSKLYDQVHLKKMIDKSYQECLASDLARRILRWGFDKYSK WYLDSRFLARDHPLTPYIKTQTWPPKHIVDLVGDTWHKLPITQIFEIP ESMDPSEILDDKSHSFTRTRLASWLSENRGGPVPSEKVIITALSKPP VNPREFLRSIDLGGLPDEDLIIGLKPKERELKIEGRFFALMSWNLRLY FVITEKLLANYILPLFDALTMTDNLNKVFKKLIDRVTGQGLLDYSRVT YAFHLDYEKWNNHQRLESTEDVFSVLDQVFGLKRVFSRTHEFFQK AWIYYSDRSDLIGLREDQIYCLDASNGPTCWNGQDGGLEGLRQKG WSLVSLLMIDRESQIRNTRTKILAQGDNQVLCPTYMLSPGLSQEGLL YELERISRNALSIYRAVEEGASKLGLIIKKEETMCSYDFLIYGKTPLFR GNILVPESKRWARVSCVSNDQIVNLANIMSTVSTNALTVAQHSQSLI KPMRDFLLMSVQAVFHYLLFSPILKGRVYKILSAEGESFLLAMSRIIY LDPSLGGISGMSLGRFHIRQFSDPVSEGLSFWREIWLSSQESWIHA LCQEAGNPDLGERTLESFTRLLEDPTTLNIRGGASPTILLKDAIRKAL YDEVDKVENSEFREAILLSKTHRDNFILFLISVEPLFPRFLSELFSSSF LGIPESIIGLIQNSRTIRRQFRKSLSKTLEESFYNSEIHGISRMTQTPQ RVGGVWPCSSERADLLREISWGRKVVGTTVPHPSEMLGLLPKSSI SCTCGATGGGNPRVSVSVLPSFDQSFFSRGPLKGYLGSSTSMSTQ LFHAWEKVTNVHVVKRALSLKESINWFITRDSNLAQALIRNIMSLTG PDFPLEEAPVFKRTGSALHRFKSARYSEGGYSSVCPNLLSHISVST DTMSDLTQDGKNYDFMFQPLMLYAQTWTSELVQRDTRLRDSTFH WHLRCNRCVRPIDDVTLETSQIFEFPDVSKRISRMVSGAVPHFQRL PDIRLRPGDFESLSGREKSHHIGSAQGLLYSILVAIHDSGYNDGTIFP

[1151]

		VNIYGKVSPRDYLRGLARGVLIGSSICFLTRMTNININRPLELVSGVIS YILLRLDNHPSLYIMLREPSLRGEIFSIPQKIPAAYPTTMKEGNRSILC YLQHVLRYEREIITASPENDWLWIFSDFRSAKMTYLSLITYQSHLLLQ RVERNLSKSMRDNLRQLSSLMRQVLGGHGEDTLESDDNIQRLLKD SLRRTRWVDQEVRHAARTMTGDYSPNKKVSRKVGCSEWVCSAQ QVAVSTSANPAPVSELDIRALSKRFQNPLISGLRVVQWATGAHYKL KPILDDLNVFPSLCLVVGDGSGGISRAVLNMFPDAKLVFNSLLEVND LMASGTHPLPPSAIMRGGNDIVSRVIDLDSIWEKPSDLRNLATWKYF QSVQKQVNMSYDLIICDAEVTDIASINRITLLMSDFALSIDGPLYLVFK TYGTMLVNPNYKAIQHLSRAFPSVTGFITQVTSSFSSELYLRFSKRG KFFRDAEYLTSSTLREMSLVLFNCSSPKSEMQRARSLNYQDLVRG FPEEIISNPYNEMIITLIDSDVESFLVHKMVDDLELQRGTLSKVAIIIAI MIVFSNRVFNVSKPLTDPSFYPPSDPKILRHFNICCSTMMYLSTALG DVPSFARLHDLYNRPITYYFRKQVIRGNVYLSWSWSNDTSVFKRVA CNSSLSLSSHWIRLIYKIVKTTRLVGSIKDLSREVERHLHRYNRWITL EDIRSRSSLLDYSCL
M 유전자 (핵산)	4005	ttctagaagcagagaggaatctttgtcctcttcggacctttgtgtctgaagagacatgtcagacca tagttgacatgctctcgggttcatgttgatacaccagactctgccctggatatgacactgttttgcaa tcactcttatttgcaatccgacgaactcagtatcatcatcccaagtgatctcctgagagtattccaa ctcctcccctttcaagagggcccctggaatcagcccactggaagataaaggttctcctcaatttgt atacccagttcaggccctcagggactggagatcctgacaaagccagtccaataaccactttga ctaacccgatcatcctatgattcccagaatatatctcgtcgaatgatttcagaatgtgccgcagga tcctgaacgagtaaccattcgggctacacacttttaacccttccgttgatacaaaagttcctcatgtt cttcttgcctgtaagttctttcagcgggacgtattcaggggtggaagccacaagtcatcgtcatc cagaggggctgacgcgggagaggatttttgagtgtcctcgtcccgcggtttttcactatcttacgt aggaggtt
M 유전자 (아미노산)	4006	NLLRKIVKNRRDEDTQKSSPASAPLDDDDLWLPPPEYVPLKELTGK KNMRNFCINGRVKVCSPNGYSFRILRHILKSFDEIYSGNHRMIGLVK VVIGLALSGSPVPEGLNWVYKLRRTFIFQWADSRGPLEGEELEYSQ EITWDDDTEFVGLQIRVIAKQCHIQGRVWCINMNPRACQLWSDMSL QTQRSEEDKDSSLLLE
P 유전자 (핵산)	4007	agcaagatctttgtcaatcctagtgctattagagccggtctggccgatcttgagatggctgaaga aactgttgatctgatcaatagaaatatcgaagacaatcaggctcatctccaaggggaacccat agaggtggacaatctccctgaggatatggggcgacttcacctggatgatggaaaatcgcccca accatggtgagatagccaaggtgggagaaggcaagtatcgagaggactttcagatggatga aggagaggatcctagcttcctgttccagtcatacctggaaaatgttggagtccaaatagtcaga caaatgaggtcaggagagagatttctcaagatatggtcacagaccgtagaagagattatatcc tatgtcgcggtcaactttcccaaccctccaggaaagtcttcagaggataaatcaacccagacta ctggccgagagctcaagaaggagacaacacccactccttctcagagagaaagccaatcatc gaaagccaggatggcggctcaaattgcttctggccctccagcccttgaatggtcggctaccaat gaagaggatgatctatcagtggaggctgagatcgctcaccagattgcagaaagtttctccaaa aaatataagtttccctctcgatcctcagggatactcttgtataattttgagcaattgaaaatgaacct tgatgatatagttaaagaggcaaaaaatgtaccaggtgtgacccgtttagcccatgacgggtcc aaactccccctaagatgtgtactgggatgggtcgctttggccaactctaagaaattccagttgtta gtcgaatccgacaagctgagtaaaatcatgcaagatgacttgaatcgctatacatcttgc
P 유전자 (아미노산)	4008	SKIFVNPSAIRAGLADLEMAEETVDLINRNIEDNQAHLQGEPIEVDNL PEDMGRLHLDDGKSPNHGEIAKVGEGKYREDFQMDEGEDPSFLF QSYLENVGVQIVRQMRSGERFLKIWSQTVEEIISYVAVNFPNPPGK SSEDKSTQTTGRELKKETTPTPSQRESQSSKARMAAQIASGPPALE WSATNEEDDLSVEAEIAHQIAESFSKKYKFPSRSSGILLYNFEQLKM NLDDIVKEAKNVPGVTRLAHDGSKLPLRCVLGWVALANSKKFQLLV ESDKLSKIMQDDLNRYTSC
G 유전자 (핵산)	4009	atggttcctcaggctctcctgtttgtaccccttctggttttccattgtgttttgggaaattccctatttaca cgataccagacaagcttggtccctggagtccgattgacatacatcacctcagctgcccaaaca atttggtagtggaggacgaaggatgcaccaacctgtcagggttctcctacatggaacttaaagtt ggatacatcttagccataaaagtgaacgggttcacttgcacaggcgttgtgacggaggctgaa acctacactaacttcgttggttatgtcacaaccacgttcaaaagaaagcatttccgcccaacacc agatgcatgtagagccgcgtacaactggaagatggccggtgaccccagatatgaagagtctc tacacaatccgtaccctgactaccgctggcttcgaactgtaaaaaccaccaaggaqtctctcqtt

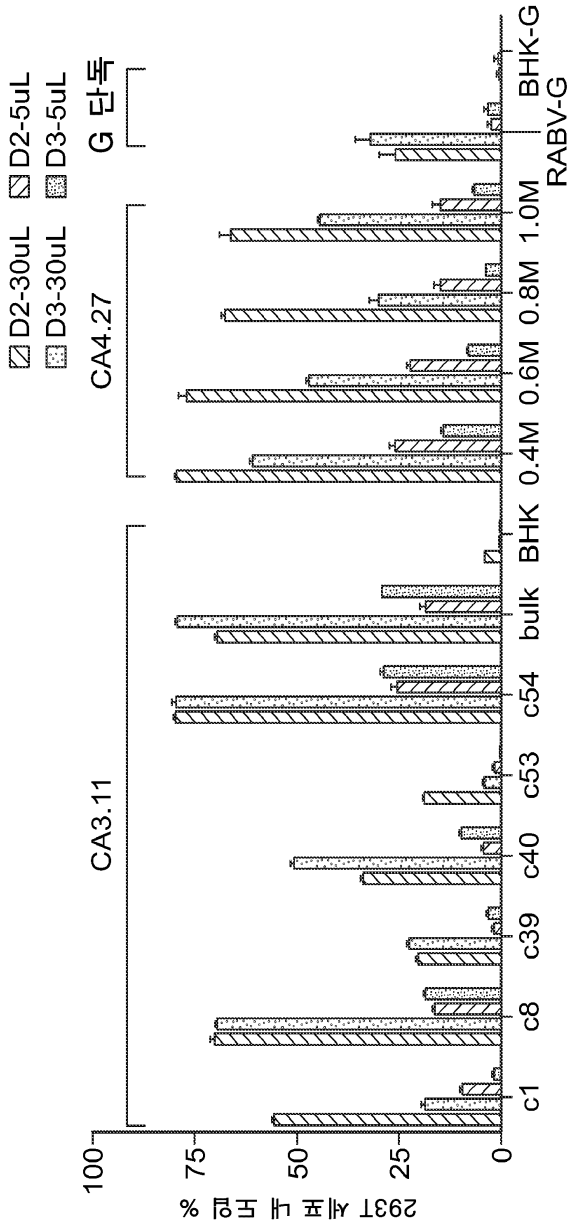
[1152]

		atcatatctccaagtgtggcagatttggacccatatgacagatcccttcactcgagggtcttccct agcgggaagtgctcaggagtagcggtgtcttctacctactgctccactaaccacgattacacca tttggatgcccgagaatccgagactagggatgtcttgtgacatttttaccaatagtagagggaag agagcatccaaagggagtgagacttgcggctttgtagatgaaagaggcctatataagtctttaa aaggagcatgcaaactcaagttatgtggagttctaggacttagacttatggatggaacatgggt ctcgatgcaaacatcaaatgaaaccaaatggtgccctcccgataagttggtgaacctgcacga ctttcgctcagacgaaattgagcaccttgttgtagaggagttggtcaggaagagaggaggagtgt ctggatgcactagagtccatcatgacaaccaagtcagtgagtttcagacgtctcagtcatttaag aaaacttgtccctgggtttggaaaagcatatacctatattcaacaagaccttgatggaagccgat gctcactacaagtcagtcagaacttggaatgagatcctcccttcaaaagggtgtttaagagttgg ggggaggtgtcatcctcatgtgaacgggggtgtttttcaatggtataatattaggacctgacggca atgtcttaatcccagagatgcaatcatccctcctccagcaacatatggagttgttggaatcctcgg ttatccccctgtgcaccccctggcagacccgtctaccgttttcaaggacggtgacgaggctgag gattttgttgaagttcaccttcccgatgtgcacaatcaggtctcaggagttgacttgggtctcccga actgggggaagtatgtattactgagtgcaggggccctgactgccttgatgttgataattttcctgat gacatgttgtagaagagtcaatcgatcagaacctacgcaacacaatctcagagggacaggg agggaggtgtcagtcactccccaaagcgggaagatcatatcttcatgggaatcacacaagag tgggggtgagaccagactg
G 유전자 (아미노산)	4010	MVPQALLFVPLLVFPLCFGKFPIYTIPDKLGPWSPIDIHHLSCPNNLV VEDEGCTNLSGFSYMELKVGYILAIKVNGFTCTGVVTEAETYTNFV GYVTTTFKRKHFRPTPDACRAAYNWKMAGDPRYEESLHNPYPDY RWLRTVKTTKESLVIISPSVADLDPYDRSLHSRVFPSGKCSGVAVS STYCSTNHDYTIWMPENPRLGMSCDIFTNSRGKRASKGSETCGFV DERGLYKSLKGACKLKLCGVLGLRLMDGTWVSMQTSNETKWCPP DKLVNLHDFRSDEIEHLVVEELVRKREECLDALESIMTTKSVSFRRL SHLRKLVPGFGKAYTIFNKTLMEADAHYKSVRTWNEILPSKGCLRV GGRCHPHVNGVFFNGIILGPDGNVLIPEMQSSLLQQHMELLESSVIP LVHPLADPSTVFKDGDEAEDFVEVHLPDVHNQVSGVDLGLPNWGK YVLLSAGALTALMLIIFLMTCCRRVNRSEPTQHNLRGTGREVSVTP QSGKIISSWESHKSGGETRL
HEK2-2 표적	4011	gaacacaaagcatagactgc
ABCA4 표적	4012	tgtcggagttcgccctggag

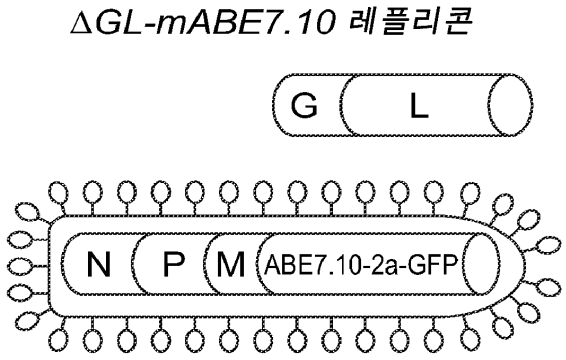
[1153]

도면

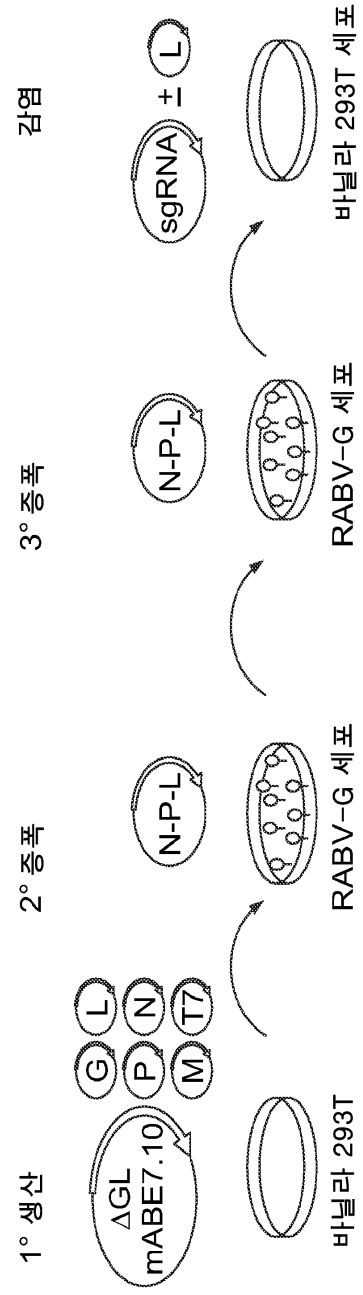
도면1


293T 세포 내 도입 %
100
75
50
25
0
CA3.11
CA4.27
G 단독
D2-30uL
D3-30uL
D2-5uL
D3-5uL
c1
c8
c39
c40
c53
c54
bulk
BHK
0.4M
0.6M
0.8M
1.0M
RABV-G
BHK-G

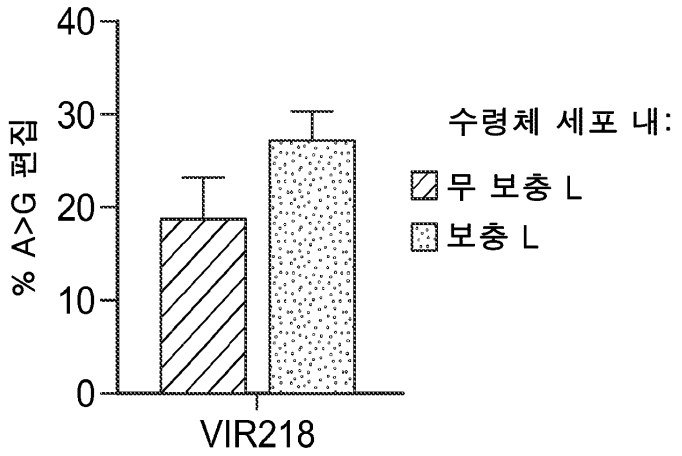
도면2a


ΔGL-mABE7.10 레플리콘
G
L
N
P
M
ABE7.10-2a-GFP

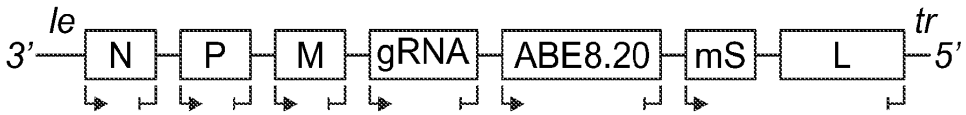
도면2b


1° 생산
2° 증폭
3° 증폭
감염
ΔGL
mABE7.10
G
L
P
N
M
T7
N-P-L
N-P-L
sgRNA
± L
바닐라 293T
RABV-G 세포
RABV-G 세포
바닐라 293T 세포

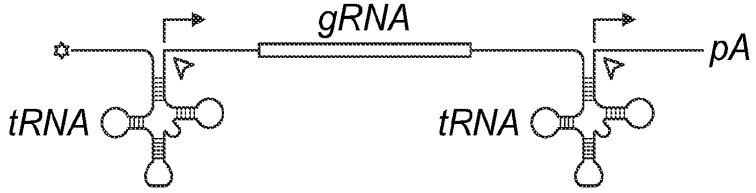
도면2c



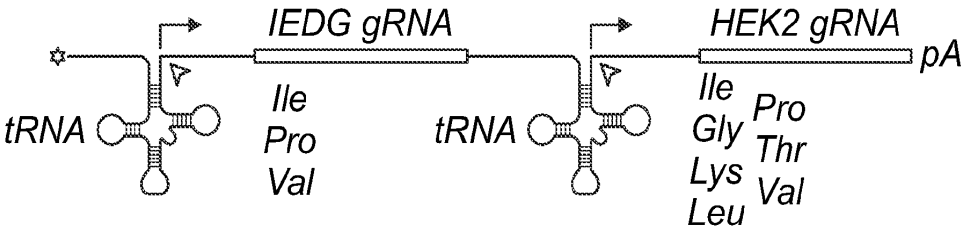
도면3a



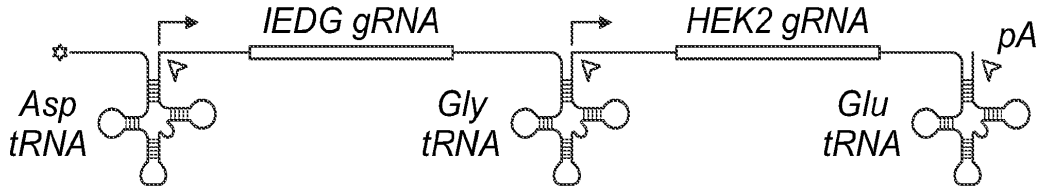
도면3b



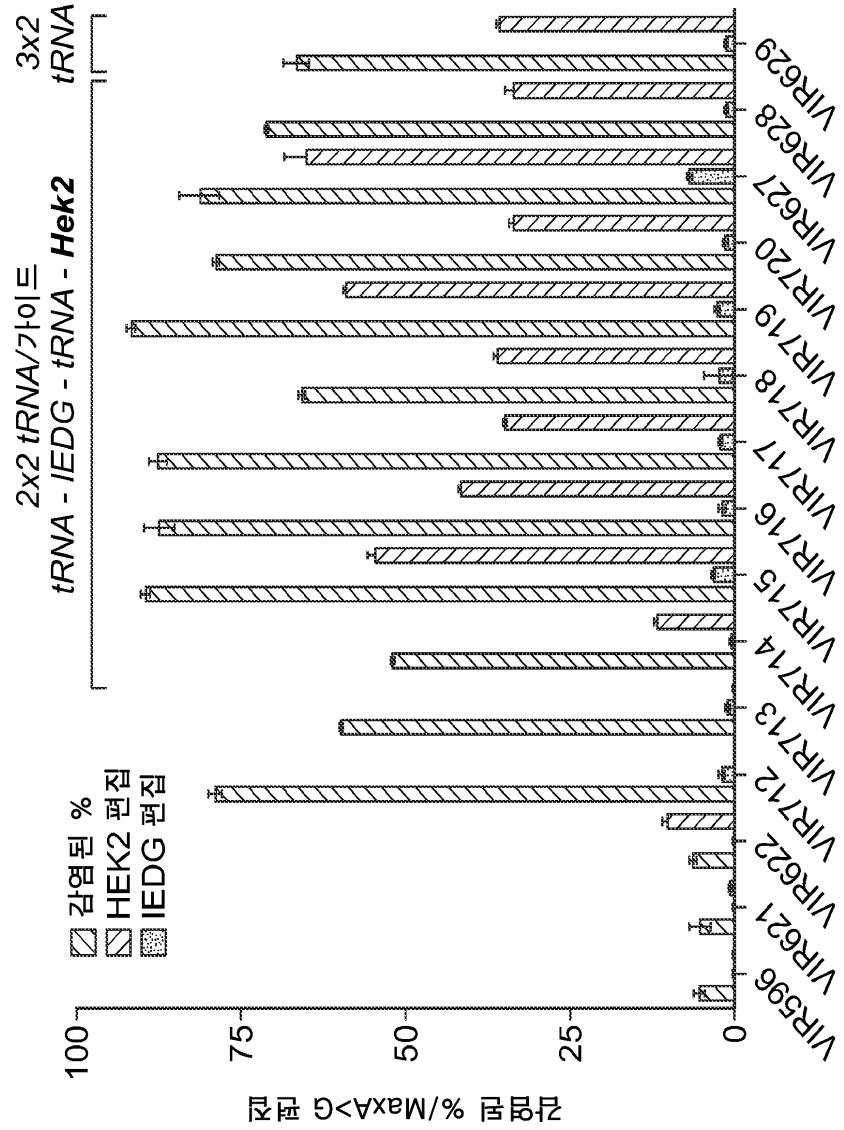
도면3c



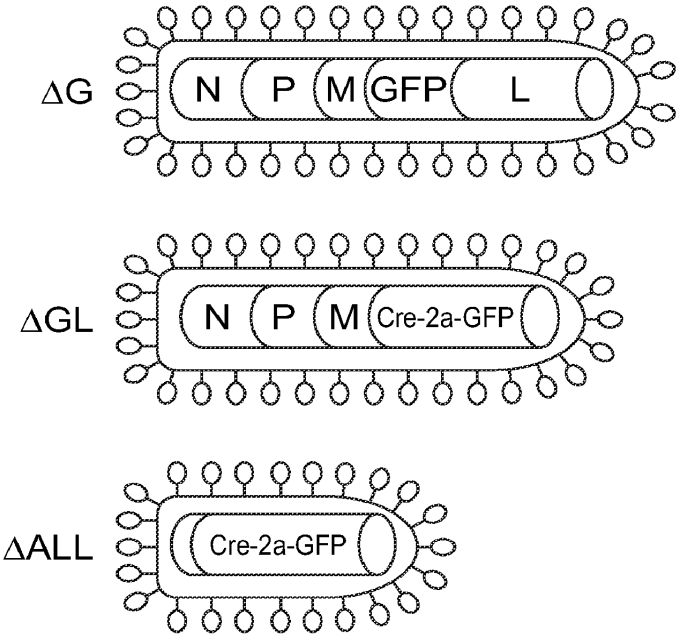
도면3d



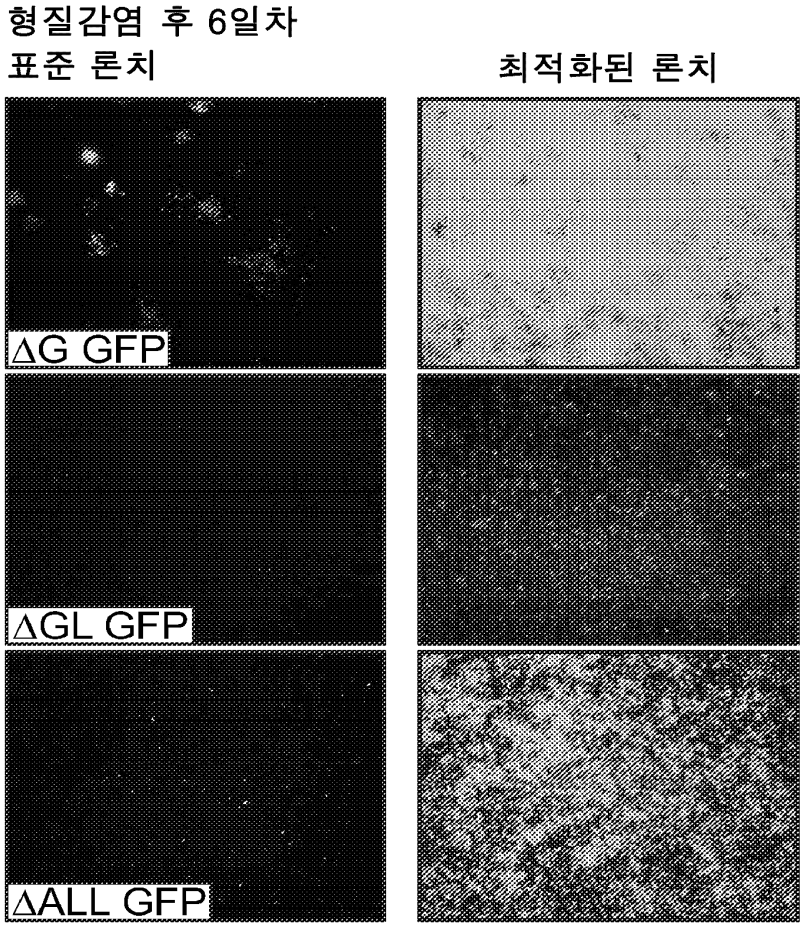
도면3e


2x2 tRNA/가이드
tRNA - IEDG - tRNA - Hek2
3x2
tRNA
감염된 %
HEK2 편집
IEDG 편집
감염된 %/MaxA>G 편집
100
75
50
25
0
VIR596
VIR621
VIR622
VIR712
VIR713
VIR714
VIR715
VIR716
VIR717
VIR718
VIR719
VIR720
VIR627
VIR628
VIR629

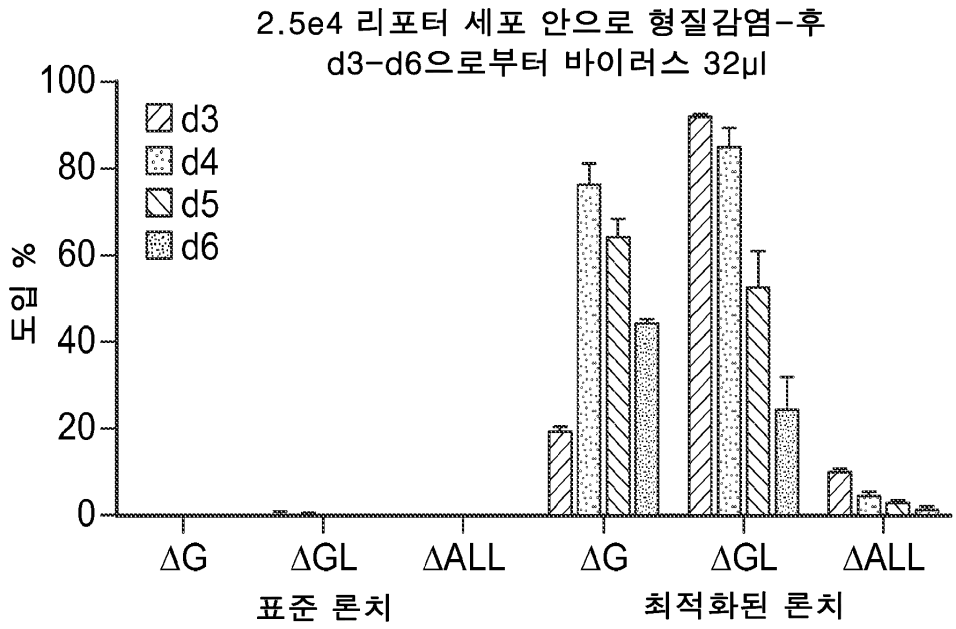
도면4a


ΔG
N
P
M
GFP
L
ΔGL
N
P
M
Cre-2a-GFP
ΔALL
Cre-2a-GFP

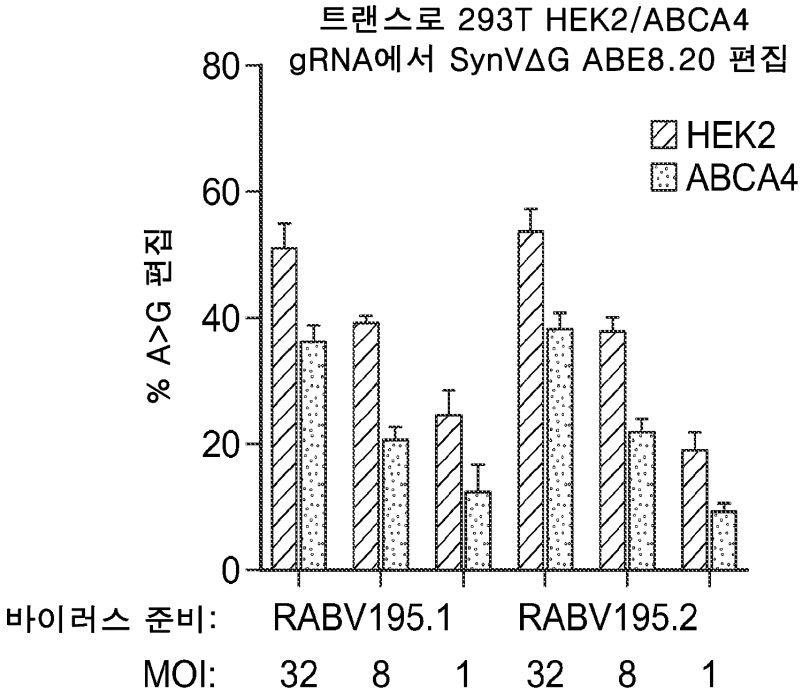
도면4b


형질감염 후 6일차
표준 론치
최적화된 론치
ΔG GFP
ΔGL GFP
ΔALL GFP

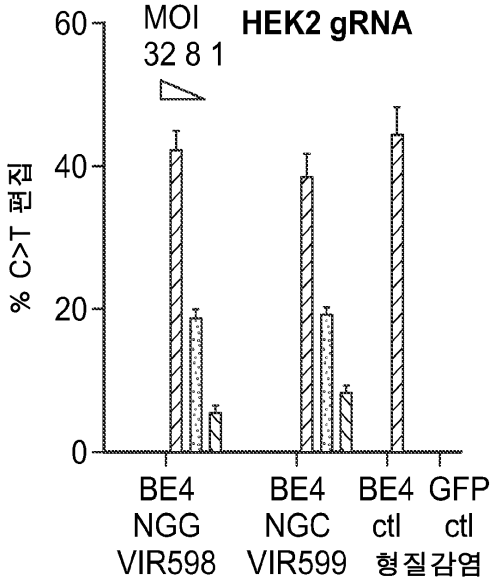
도면4c

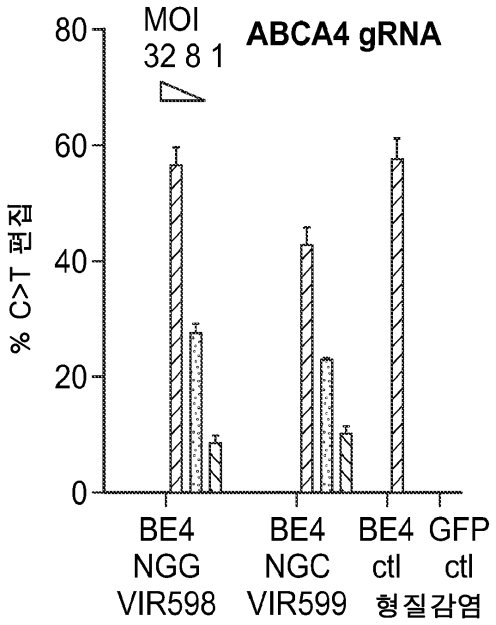

2.5e4 리포터 세포 안으로 형질감염-후
d3-d6으로부터 바이러스 32μl
100
80
60
40
20
0
% 바이러스
d3
d4
d5
d6
ΔG
ΔGL
ΔALL
ΔG
ΔGL
ΔALL
표준 론치
최적화된 론치

도면5

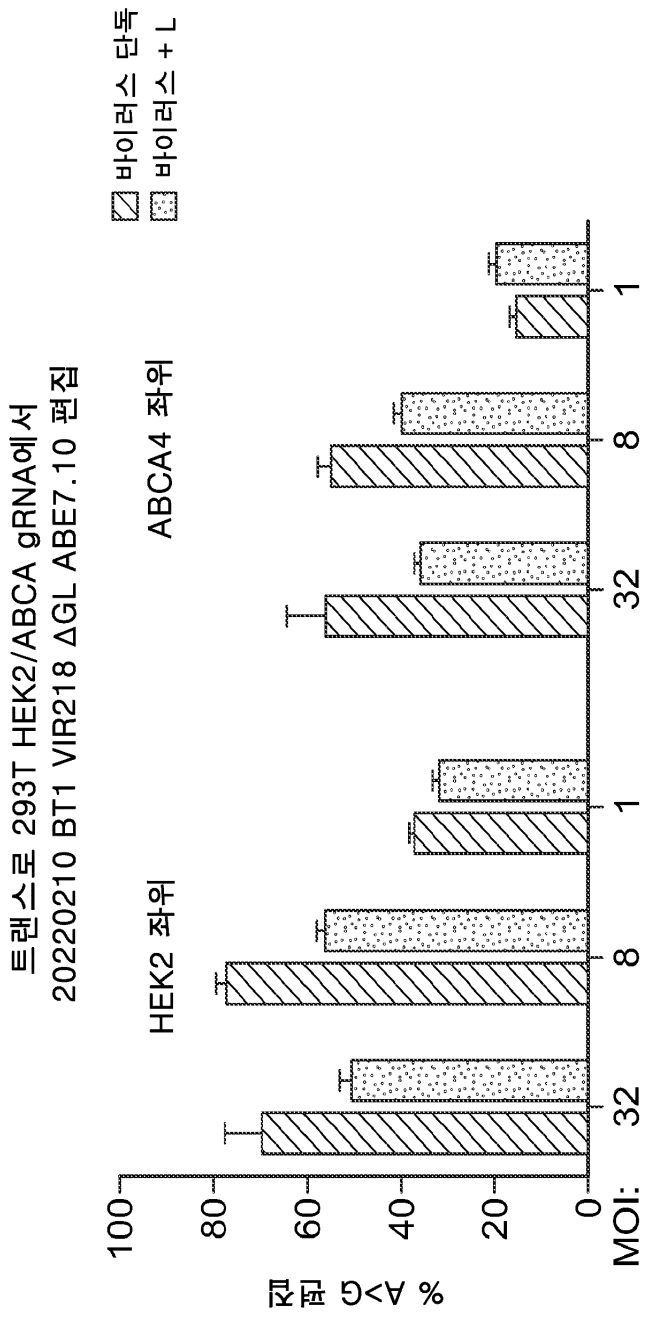

트랜스로 293T HEK2/ABCA4
gRNA에서 SynVΔG ABE8.20 편집
80
60
40
20
0
% A>G 편집
HEK2
ABCA4
바이러스 준비:
RABV195.1
RABV195.2
MOI:
32
8
1
32
8
1

도면6

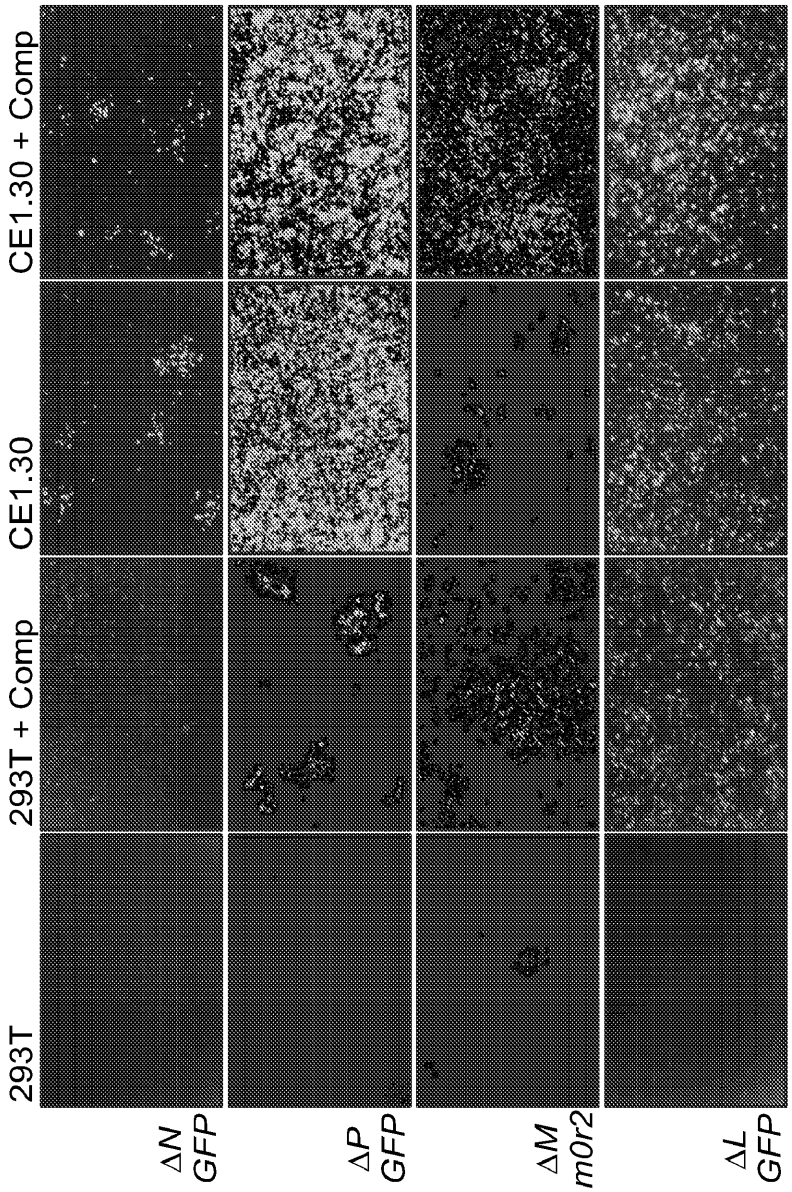

MOI
32 8 1
HEK2 gRNA
% C>T 편집
60
40
20
0
BE4
NGG
VIR598
BE4
NGC
VIR599
BE4
ctl
형질
GFP
ctl
감염


MOI
32 8 1
ABCA4 gRNA
% C>T 편집
80
60
40
20
0
BE4
NGG
VIR598
BE4
NGC
VIR599
BE4
ctl
형질
GFP
ctl
감염

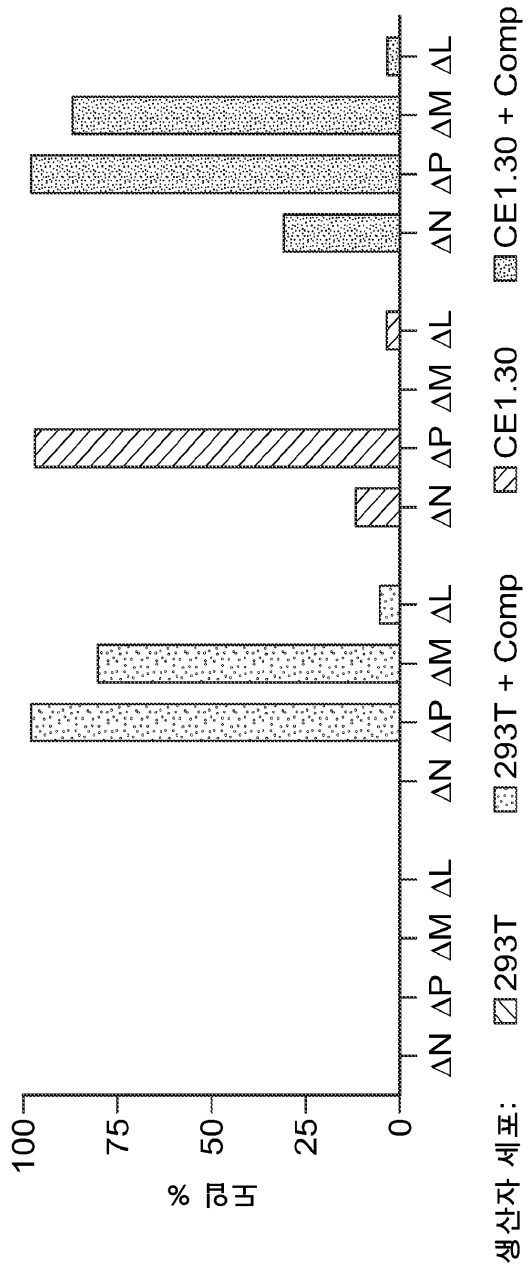
도면7


트랜스로 293T HEK2/ABCA gRNA에서
20220210 BT1 VIR218 ΔGL ABE7.10 편집
바이러스 단독
바이러스 + L
HEK2 좌위
ABCA4 좌위
% A>G 편집
100
80
60
40
20
0
MOI:
32
8
1
32
8
1

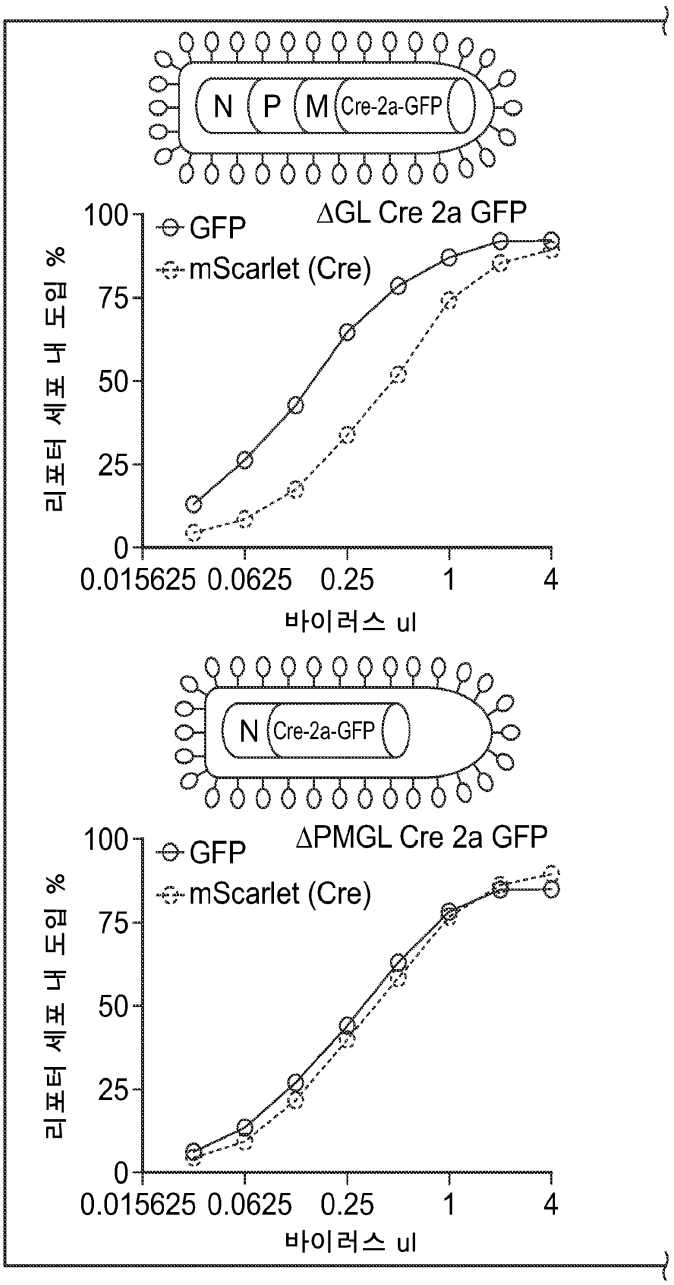
도면8



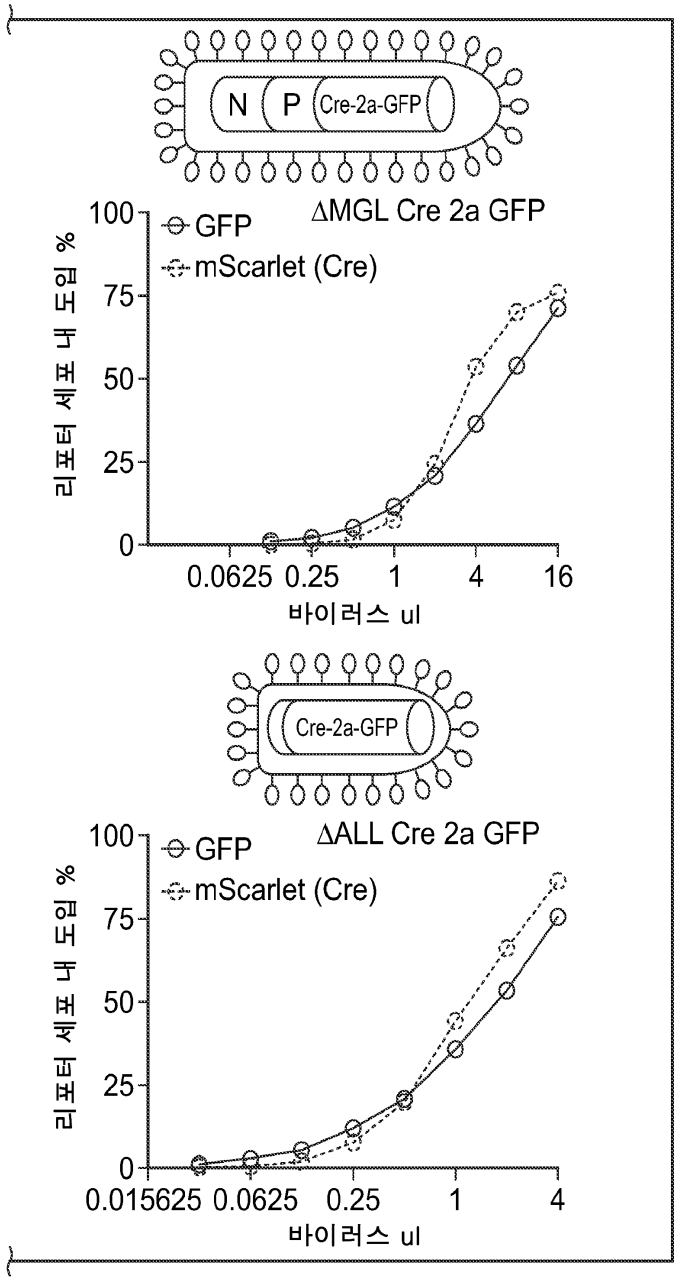
도면9


100
75
50
25
0
% 억제
ΔN ΔP ΔM ΔL
ΔN ΔP ΔM ΔL
ΔN ΔP ΔM ΔL
ΔN ΔP ΔM ΔL
생산자 세포:
293T
293T + Comp
CE1.30
CE1.30 + Comp

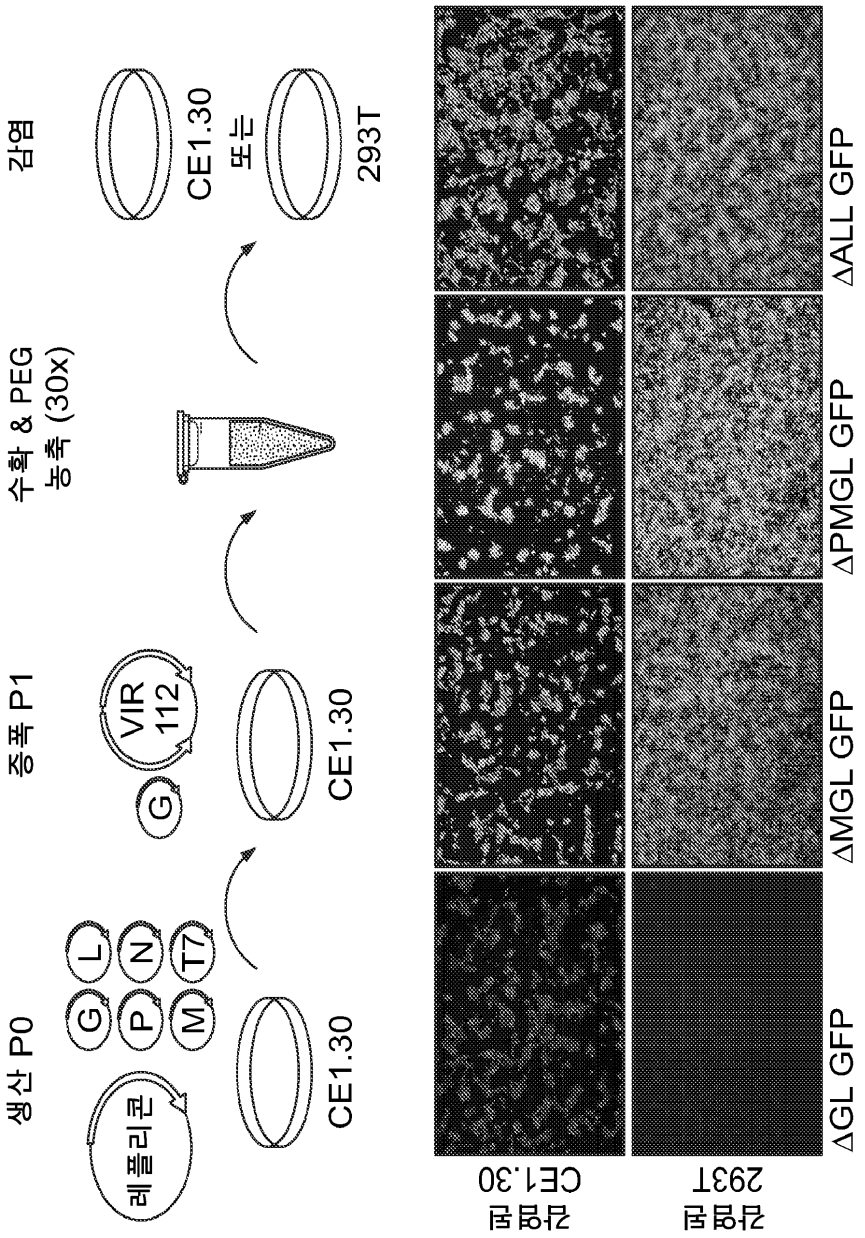
도면10a


N P M Cre-2a-GFP
ΔGL Cre 2a GFP
GFP
mScarlet (Cre)
리포터 세포 내 도입 %
100
75
50
25
0
0.015625 0.0625 0.25 1 4
바이러스 ul
N Cre-2a-GFP
ΔPMGL Cre 2a GFP
GFP
mScarlet (Cre)
리포터 세포 내 도입 %
100
75
50
25
0
0.015625 0.0625 0.25 1 4
바이러스 ul

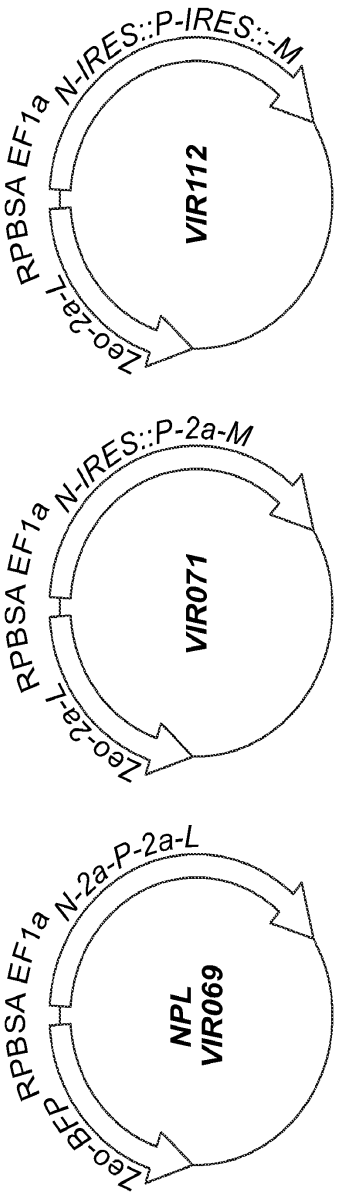
도면10b


N
P
Cre-2a-GFP
ΔMGL Cre 2a GFP
GFP
mScarlet (Cre)
리포터 세포 내 도입 %
100
75
50
25
0
0.0625
0.25
1
4
16
바이러스 ul
Cre-2a-GFP
ΔALL Cre 2a GFP
GFP
mScarlet (Cre)
리포터 세포 내 도입 %
100
75
50
25
0
0.015625
0.0625
0.25
1
4
바이러스 ul

도면11


생산 P0
레플리콘
G
P
M
L
N
T7
CE1.30
증폭 P1
G
VIR 112
CE1.30
수확 & PEG 농축 (30x)
감염
CE1.30 또는 293T
ΔGL GFP
ΔMGL GFP
ΔPMGL GFP
ΔALL GFP
293T 감염됨
CE1.30 감염됨

도면12


RPBSA EF1a N-IRES::P-IRES::-M
VIR112
Zeo-2a-L
RPBSA EF1a N-IRES::P-2a-M
VIR071
Zeo-2a-L
RPBSA EF1a N-2a-P-2a-L
NPL
VIR069
Zeo-BFP