(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2025-0012194
(43) 공개일자 2025년01월23일

(51) 국제특허분류(Int. Cl.)
A61K 48/00 (2006.01) *A61K 31/7088* (2006.01)
A61P 25/00 (2006.01) *C12N 15/113* (2010.01)
C12N 15/86 (2006.01) *C12N 9/22* (2006.01)
(52) CPC특허분류
A61K 48/0058 (2013.01)
A61K 31/7088 (2013.01)
(21) 출원번호 10-2025-7000942(분할)
(22) 출원일자(국제) 2014년06월11일
심사청구일자 없음
(62) 원출원 특허 10-2016-7001296
원출원일자(국제) 2014년06월11일
심사청구일자 2019년05월16일
(85) 번역문제출일자 2025년01월10일
(86) 국제출원번호 PCT/US2014/041809
(87) 국제공개번호 WO 2014/204729
국제공개일자 2014년12월24일
(30) 우선권주장
61/836,123 2013년06월17일 미국(US)
(뒷면에 계속)

(71) 출원인
더 브로드 인스티튜트, 인코퍼레이티드
미국 02142 매사 추세츠주 캠브리지 메인 스트리트 415
매사추세츠 인스티튜트 오브 테크놀로지
미국 02139 매사추세츠주 캠브리지 매사추세츠 애비뉴 77
프레지던트 앤드 펠로우즈 오브 하바드 칼리지
미합중국, 메사추세츠 02138, 캠브리지, 퀸시스트리트17
(72) 발명자
장 펑
미국 02139 매사추세츠주 캠브리지 퍼시픽 스트리트 100 아파트먼트 11
콩 레
미국 02142 매사추세츠주 캠브리지 메모리얼 드라이브 100 아파트먼트 8-21비
(뒷면에 계속)
(74) 대리인
특허법인코리아나

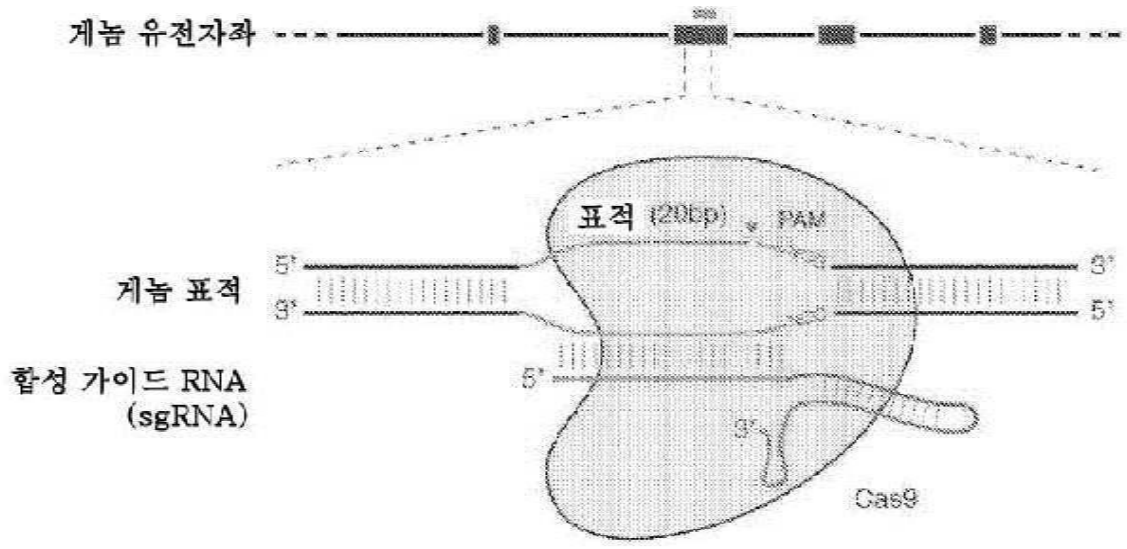
전체 청구항 수 : 총 19 항

(54) 발명의 명칭 **바이러스 구성성분을 사용하여 장애 및 질환을 표적화하기 위한 CRISPR-CAS 시스템 및 조성물의 전달, 용도 및 치료 적용**

(57) 요 약

본 발명은 서열의 조작 및/또는 표적 서열의 활성을 위한 시스템, 방법, 및 조성물의 전달, 유전자 조작 및 최적화를 제공한다. 전달 시스템, 및 전달을 위한 부위로서 표적화되는 조직 또는 기관이 제공된다. 또한, 일부가 CRISPR 복합체의 하나 이상의 구성성분을 암호화하는 벡터 및 벡터 시스템뿐만 아니라 상기 벡터의 설계 및 사용을 위한 방법이 제공된다. 표적 인지 및 독성의 회피를 위한 향상된 특이성을 보장하고, 질병 또는 질환의 상태를 변화시키거나 개선하기 위해 관심대상 게놈 유전자좌 내의 표적 부위를 편집하거나 변형시키기 위해 진핵 세포에서 CRISPR 복합체 형성을 유도하는 방법이 또한 제공된다.

대 표 도 - 도1



(52) CPC특허분류
A61P 25/00 (2018.01)
C12N 15/113 (2013.01)
C12N 15/86 (2013.01)
C12N 9/22 (2013.01)
C07K 2319/09 (2013.01)
C12N 2310/20 (2017.05)
C12N 2750/14143 (2013.01)
C12N 2830/008 (2013.01)

(72) 발명자
란 페이
미국 02116 매사추세츠주 보스턴 클레어든 스트리트 30
하이덴라이히 마티아스
미국 02144 매사추세츠주 소머빌 체스터 스트리트 53
스비에흐 루카시
미국 02120 매사추세츠주 보스턴 그린위치 스트리트 6 아파트먼트 3

(30) 우선권주장
61/847,537 2013년07월17일 미국(US)
61/862,355 2013년08월05일 미국(US)
61/871,301 2013년08월28일 미국(US)
PCT/US2013/074667 2013년12월12일 미국(US)
61/915,225 2013년12월12일 미국(US)
61/979,879 2014년04월15일 미국(US)

명 세 서

청구범위

청구항 1

CRISPR-Cas 시스템을 암호화하는 하나 이상의 바이러스 벡터를 포함하는 뇌 또는 신경학적 질환 또는 장애를 치료하기 위한 약학적 조성물로서; CRISPR-Cas 시스템은 적어도 하나의 핵 국소화 서열을 포함하는 Cas9, 뇌 또는 뉴런 세포에서 발현되는 표적 서열에 혼성화하고, 뇌 또는 신경학적 질환 또는 장애에 관여하는 유전자와 관련된 가이드 서열, tracr 메이트 서열, 및 tracr 서열을 포함하고; 상기 조성물은 진핵 유기체의 뇌 또는 뉴런 세포에 전달되어 생체 내에서 뇌 또는 뉴런 세포를 변형시키고, 바이러스 벡터는 Cas9를 암호화하는 폴리뉴클레오티드 서열에 작동 가능하게 연결된 뇌-특이적 프로모터 또는 뉴런-특이적 프로모터를 포함하는, 약학적 조성물.

청구항 2

제1항에 있어서, 유전자 산물의 발현은 뇌 또는 뉴런 세포에서 CRISPR-Cas 시스템에 의해 변경되어, 진핵 유기체에서 표현형 변화를 생성하는, 약학적 조성물.

청구항 3

제1항 또는 제2항에 있어서, Cas9는 적어도 2개의 핵 국소화 서열을 포함하는, 약학적 조성물.

청구항 4

제1항 또는 제2항에 있어서, CRISPR-Cas 시스템은 뇌 또는 뉴런 세포에서 하나 이상의 결함이 있는 유전자형을 교정하는 것인, 약학적 조성물.

청구항 5

제1항에 있어서, 하나 이상의 바이러스 벡터는:

(a) 가이드 서열, tracr 메이트 서열 및 tracr 서열을 암호화하는 폴리뉴클레오티드 서열에 작동 가능하게 연결된 제1 프로모터, 및

(b) 적어도 하나의 핵 국소화 서열 및 Cas9를 암호화하는 폴리뉴클레오티드 서열에 작동 가능하게 연결된 제2 프로모터를 포함하고,

구성성분 (a) 및 (b)는 동일 또는 상이한 바이러스 벡터 상에 위치하는, 약학적 조성물.

청구항 6

제1항에 있어서, 하나 이상의 바이러스 벡터는:

(a) 5' 에서 3' 으로, 가이드 서열, tracr 메이트 서열 및 tracr 서열을 포함하는 폴리뉴클레오티드 서열, 및

(b) 적어도 하나의 핵 국소화 서열 및 Cas9를 암호화하는 폴리뉴클레오티드 서열을 포함하고,

tracr 메이트 서열은 tracr 서열에 혼성화하고, 가이드 서열은 표적 서열에 대한 Cas9를 포함하는 CRISPR 복합체의 서열-특이적 결합을 지시하며, Cas9를 암호화하는 폴리뉴클레오티드 서열은 DNA 또는 RNA인, 약학적 조성물.

청구항 7

제1항에 있어서, 하나 이상의 바이러스 벡터는:

(a) 가이드 서열 및 tracr 메이트 서열을 포함하는 폴리뉴클레오티드 서열,

(b) 적어도 하나의 핵 국소화 서열 및 Cas9를 암호화하는 폴리뉴클레오티드 서열, 및

(c) tracr 서열을 포함하는 폴리뉴클레오티드 서열을 포함하고,

tracr 메이트 서열은 tracr 서열에 혼성화하고, 가이드 서열은 표적 서열에 대한 Cas9를 포함하는 CRISPR 복합체의 서열-특이적 결합을 지시하며, Cas9를 암호화하는 폴리뉴클레오티드 서열은 DNA 또는 RNA인, 약학적 조성물.

청구항 8

제1항 또는 제2항에 있어서, 2개 이상의 뉴런 유전자 생성물의 발현이 변경되는, 약학적 조성물.

청구항 9

제1항 또는 제2항에 있어서, 가이드 서열이 tracr 메이트 서열 및 tracr 서열에 융합된, 약학적 조성물.

청구항 10

제5항, 제6항 또는 제7항에 있어서, Cas9를 암호화하는 폴리뉴클레오티드 서열이 진핵 유기체에서의 발현을 위해 코돈 최적화된 것인, 약학적 조성물.

청구항 11

제1항 또는 제2항에 있어서, 진핵 유기체가 포유류인, 약학적 조성물.

청구항 12

제1항 또는 제2항에 있어서, Cas9는 S. 피오게네스 Cas9인, 약학적 조성물.

청구항 13

제1항 또는 제2항에 있어서, Cas9는 S. 아우레우스 Cas9인, 약학적 조성물.

청구항 14

제1항 또는 제2항에 있어서, 하나 이상의 바이러스 벡터는 아데노-연관 바이러스(AAV) 또는 렌티바이러스 벡터이거나; AAV1, AAV2, AAV4, AAV5, AAV8, AAV9, 또는 이들의 조합인, 약학적 조성물.

청구항 15

제5항에 있어서, (a) 및 (b)는 동일한 바이러스 벡터 상에 위치하는, 약학적 조성물.

청구항 16

제15항에 있어서, 상기 바이러스 벡터는 AAV 벡터이고, 상기 Cas9는 S. 아우레우스 Cas9인, 약학적 조성물.

청구항 17

제1항 또는 제2항에 있어서, 바이러스 벡터는 주사를 통해 전달되는, 약학적 조성물.

청구항 18

제5항, 제6항 또는 제7항에 있어서, Cas9를 암호화하는 폴리뉴클레오티드 서열이 CamkII, 파르브알부민(Parvalbumin), vGAT, DR1, DR2, GFAP 및 Arc로 이루어진 그룹으로부터 선택된 프로모터에 작동 가능하게 연결된, 약학적 조성물.

청구항 19

뇌 또는 신경학적 질환 또는 장애를 치료하기 위한 약제의 제조에 사용하기 위한 조성물로서, 조성물은 CRISPR-Cas 시스템을 암호화하는 하나 이상의 바이러스 벡터를 포함하고, CRISPR-Cas 시스템은 적어도 하나의 핵 국소화 서열을 포함하는 Cas9, 뇌 또는 뉴런 세포에서 발현되는 표적 서열에 혼성화하고, 뇌 또는 신경학적 질환 또는 장애에 관여하는 유전자와 관련된 가이드 서열, tracr 메이트 서열 및 tracr 서열을 포함하고, 바이러스 벡터는 Cas9를 암호화하는 폴리뉴클레오티드 서열에 작동 가능하게 연결된 뇌-특이적 프로모터 또는 뉴런-특이적 프로모터를 포함하는, 조성물.

발명의 설명

기 술 분 야

[0001] **관련 출원 및 참조로서의 포함**

[0002] 미국에 대해 2013년 6월 17일에 출원된 미국 가출원 61/836,123, 2013년 7월 17일에 출원된 61/847,537, 2013년 8월 5일에 출원된 61/862,355, 2013년 8월 28일에 출원된 61/871,301, 2013년 12월 12일에 출원된 61/915,225, 2014년 4월 15일에 출원된 61/979,879, 및 2013년 12월 12일에 출원된 PCT/US2013/074667이 우선권으로 주장되고, 본 출원은 또한 일부계속출원이며, 미국 특허법 하에서 허용될 수 있는 바와 같이, 상기 출원들에 대한 미국의 동등한 출원 또는 국내 단계가 추가로 주장될 수 있고, PCT/US2013/074667호 및 PCT/US2013/074667호가 우선권을 주장하는 출원에 대해 우선권이 주장될 수 있다.

[0003] 전술한 출원, 및 상기 출원에 또는 상기 출원의 절차 중에 인용된 모든 문헌("출원 인용 문헌") 및 상기 출원 인용 문헌에 인용되거나 참고된 모든 문헌, 및 본원에서 인용되거나 참고된 모든 문헌("본원 인용 문헌") 및 본원 인용 문헌에 인용되거나 참고된 모든 문헌은, 본원에 언급되거나 본원에 참고로 포함된 임의의 문헌에 언급된 임의의 제품에 대한 임의의 제조사의 지침서, 설명서, 제품 명세서 및 제품 시이트(sheet)와 함께, 본원에 참고로 포함되어 있으며, 그리고 본 발명의 실시에 사용될 수 있다. 더욱 구체적으로, 모든 참조된 문헌은 마치 각각의 개별 문헌을 참고로 포함하는 것으로 특정적으로 그리고 개별적으로 나타내는 것과 동일한 정도로 참고로 포함된다.

[0004] **기술 분야**

[0005] 본 발명은 일반적으로 클러스터링되고 규칙적으로 산재된 짧은 팔린드로믹 반복부(Clustered Regularly Interspaced Short Palindromic Repeats; CRISPR) 및 그의 구성성분과 관련된, 게놈 변동(genomic perturbation) 또는 유전자-편집(gene-editing)과 같이 서열 표적화를 수반하는 유전자 발현의 제어를 위해 사용되는 시스템, 방법 및 조성물의 전달, 유전자 조작 및 최적화 및 치료적 적용에 관한 것이다. 특히, 본 발명은 바이러스 벡터 전달, 바이러스 전달에 의한 유전자 요법, 및 유전자 기능의 이해 및 바이러스 벡터 전달을 통한 모델의 생성에 관한 양태에 관한 것이다.

[0006] **연방 정부가 후원하는 연구에 대한 성명**

[0007] 본 발명은 미국 국립 보건원(National Institutes of Health)에 의해 지급된 NIH 파이오니어 어워드(Pioneer Award)(1DP1MH100706)하의 정부 지원으로 수행되었다. 정부는 본 발명에 소정의 권리를 갖는다.

배 경 기 술

[0008] 게놈 시퀀싱(sequencing) 기법 및 분석 방법의 최근의 진전에 의해, 다양한 생물학적 기능 및 질병(disease)과 연관된 유전적 요인을 분류하고 발견하는 능력이 상당히 가속화되었다. 개별 유전 요소의 선택적 변동을 가능하게 함으로써 원인이 되는 유전 변이의 체계적인 역의 조작을 가능하게 할 뿐 아니라, 합성 생물학, 생명공학 및 의학 응용을 진전시키기 위하여, 정밀한 게놈 표적화 기술이 필요하다. 게놈-편집 기법, 예를 들어, 디자이너 징크 핑거, 전사 활성화제-유사 이펙터(effector)(TALE) 또는 귀소 메가뉴클레아제(homing meganuclease)가 표적화된 게놈 변동을 생성하는데 이용가능하지만, 가격이 알맞고, 설립하기 용이하며, 확대가능하고, 진핵 게놈 내의 다수의 위치를 표적화하는데 부합되는 새로운 게놈 조작 기술에 대한 필요성이 남아있다.

발명의 내용

과제의 해결 수단

[0009] 본 발명은 바이러스 구성성분을 사용하여 질환 및 장애를 처리하기 위하여 유전자 또는 게놈의 게놈 변동(genomic perturbation) 또는 유전자-편집(gene-editing)과 같이 서열 표적화를 위한 도구로서 CRISPR/Cas9 시스템의 개발 및 적용을 수반한다.

[0010] CRISPR/Cas 시스템은 특정 서열을 표적화하기 위해 맞춤형 단백질의 생성을 필요로 하지 않고, 오히려, 단일의 Cas 효소가 짧은 RNA 분자에 의해 프로그램화되어, 특정 DNA 표적을 인식할 수 있다. 게놈 시퀀싱(sequencing) 기법 및 분석 방법의 레퍼토리에 CRISPR-Cas 시스템을 부가하면, 방법을 상당히 단순화시킬 수 있으며, 다양한 생물학적 기능 및 질병과 연관된 유전적 요인을 분류하고 발견하는 능력을 가속화시킬 수 있다. 유해 영향 없이

게놈 교정을 위해 효율적으로 CRISPR-Cas 시스템을 사용하기 위하여, 청구된 발명의 양태인 이들 게놈 조작 도구의 최적화 및 조작의 양태를 이해하는 것이 중요하다.

[0011] 다수의 적용분야를 지니는 핵산 서열 표적화를 위한 대안의 및 강한 시스템 및 기법에 대한 절실한 필요가 존재한다. 본 발명의 양태는 이러한 필요를 처리하고 관련된 이점을 제공한다. 예시적인 CRISPR 복합체는 표적 폴리뉴클레오티드 내에서 표적 서열에 혼성화된 가이드 서열과 복합체화되는 CRISPR 효소 복합체를 포함한다. 가이드 서열은 tracr 메이트 서열에 연결되고, 이는 결국 tracr 서열에 혼성화된다.

[0012] 제1 양태에서, 본 발명은 조성물을 작동 가능하게 암호화하는 하나 이상의 바이러스 벡터를 포함할 수 있는 바이러스 벡터 시스템을 포함하는 비자연적으로 생기는 또는 유전자 조작된 조성물을 이의 발현을 위해 전달하는 단계를 포함할 수 있는 관심 대상의 게놈 유전자좌에서 표적 서열의 조작에 의해 유기체 또는 비인간 유기체를 변형시키는 방법을 제공하되, 여기서 상기 조성물은

[0013] (A) I. CRISPR-Cas 시스템 RNA 폴리뉴클레오티드 서열에 작동 가능하게 연결되는 제1 조절 요소로서, 폴리뉴클레오티드 서열은

[0014] (A) 진핵 세포 내 표적 서열에 혼성화할 수 있는 가이드 서열,

[0015] (b) tracr 메이트 서열, 및

[0016] (c) tracr 서열

[0017] 을 포함하는, 제1 조절 요소, 및

[0018] II. 선택적으로 적어도 하나 이상의 핵 국소화 서열을 포함할 수 있는, CRISPR 효소를 암호화하는 효소 암호화 서열에 작동 가능하게 연결되는 제2 조절 요소

[0019] 를 포함할 수 있는 하나 이상의 벡터를 포함할 수 있는 벡터 시스템을 포함하는 비자연적으로 생기는 또는 유전자 조작된 조성물

[0020] (여기서, (A), (b) 및 (c)는 5'에서 3' 방향으로 배열되고,

[0021] 구성성분 I 및 II는 시스템의 동일 또는 상이한 벡터 상에 위치되며,

[0022] 전사될 때, tracr 메이트 서열은 tracr 서열에 혼성화하고, 가이드 서열은 표적 서열에 대한 CRISPR 복합체의 서열-특이적 결합을 지시하며,

[0023] CRISPR 복합체는 (1) 표적 서열에 혼성화되는 가이드 서열, 및 (2) tracr 서열에 혼성화되는 tracr 메이트 서열과 복합체화된 CRISPR 효소를 포함함), 또는

[0024] (B) I. (A) 진핵 세포 내 표적 서열에 혼성화할 수 있는 가이드 서열, 및

[0025] (b) 적어도 하나 이상의 tracr 메이트 서열

[0026] 에, 작동 가능하게 연결되는 제1 조절 요소,

[0027] II. CRISPR 효소를 암호화하는 효소 암호화 서열에 작동 가능하게 연결되는 제2 조절 요소, 및

[0028] III. tracr 서열에 작동 가능하게 연결되는 제3 조절 요소

[0029] 를 포함할 수 있는 하나 이상의 벡터를 포함할 수 있는 벡터 시스템을 포함할 수 있는 비자연적으로 생기는 또는 유전자 조작된 조성물

[0030] (여기서, 구성성분 I, II 및 III은 시스템의 동일 또는 상이한 벡터 상에 위치되고,

[0031] 전사될 때, tracr 메이트 서열은 tracr 서열에 혼성화하고, 가이드 서열은 표적 서열에 대한 CRISPR 복합체의 서열-특이적 결합을 지시하며,

[0032] CRISPR 복합체는 (1) 표적 서열에 혼성화되는 가이드 서열, 및 (2) tracr 서열에 혼성화되는 tracr 메이트 서열과 복합체화된 CRISPR 효소를 포함함)

[0033] 을 포함할 수 있다.

[0034] 일 양태에서, 본 발명은 CRISPR-Cas 시스템 중 하나 이상의 요소를 사용하기 위한 방법을 제공한다. 본 발명의 CRISPR 복합체는 표적 폴리뉴클레오티드를 변형시키기 위한 효과적인 수단을 제공한다. 본 발명의 CRISPR 복합

체는 다양한 조직 및 기관 내 다수의 세포 유형에서 표적 폴리뉴클레오티드를 변형시키는 것(예를 들어, 결실, 삽입, 전좌, 불활성화, 활성화)을 포함하는 다양한 효용을 가진다. 이와 같이 본 발명의 CRISPR 복합체는, 예를 들어 유전자 또는 게놈 편집, 유전자 요법, 유전자 발견, 약물 선별, 질병 진단 및 예후에서 넓은 적용 범위를 가진다. 생체내, 시험관내 및 생체외 사용이 예견된다.

[0035] 본 발명의 양태는 야생형 Cas9 효소 및 이에 대한 핵산 분자 암호보다 길이가 더 짧은 최적의 활성을 갖는 가이드 RNA를 갖는 CRISPR-Cas9 시스템에서 개선된 표적화 특이성을 갖는 Cas9 효소, 및 키메라 Cas9 효소뿐만 아니라, Cas9 효소의 표적 특이성을 개선하는 방법, 또는 최적의 활성을 갖는 가이드 RNA를 설계 또는 준비하는 단계, 및/또는 야생형 Cas9보다 크기 또는 길이가 더 작은 Cas9를 선택 또는 준비함으로써, 야생형 Cas9에 대해서보다 전달 벡터 내에 더 적은 암호화 서열이 있기 때문에 전달 벡터 내로 핵산 서열 암호를 패키징하는 것이 더 진보되는 단계 및/또는 키메라 Cas9 효소를 생성하는 단계를 포함할 수 있는 CRISPR-Cas9 시스템을 설계하는 방법에 관한 것이다.

[0036] 또한 의학에서의 본 서열, 벡터, 효소 또는 시스템의 용도가 제공된다. 또한 유전자 또는 게놈 편집에서의 본 서열, 벡터, 효소 또는 시스템의 용도가 제공된다. 이는 생체내이거나 생체외이건 간에 유사분열 후 세포 조직 또는 세포와 관련된다.

[0037] 본 발명의 추가적인 양태에서, Cas9 효소는 하나 이상의 돌연변이를 포함할 수 있고, 기능성 도메인에 대한 융합이 있는 또는 융합이 없는 일반적 DNA 결합 단백질로서 사용될 수 있다. 돌연변이는 인공적으로 도입된 돌연변이 또는 기능획득 또는 기능상실 돌연변이일 수 있다. 돌연변이는 각각 RuvC 및 HNH 촉매적 도메인에서 촉매적 도메인(D10 및 H840) 중 하나에서의 돌연변이를 포함할 수 있지만, 이들로 제한되지 않는다. 추가 돌연변이가 특성규명되었다. 본 발명의 일 양태에서, 전사 활성화 도메인은 VP64일 수 있다. 본 발명의 다른 양태는, 전사 리프레서 도메인은 KRAB 또는 SID4X일 수 있다. 본 발명의 다른 양태는 전사 활성인자, 리프레서, 재조합효소, 전위효소, 히스톤 리모델러, 탈메틸화효소, DNA 메틸트랜스퍼라제, 크립토크롬, 광 유도성/제어가능 도메인 또는 화학적 유도성/제어가능 도메인을 포함하지만, 이것으로 제한되지 않는 도메인에 융합된 돌연변이된 Cas 9 효소에 관한 것이다.

[0038] 추가 구현예에서, 본 발명은 세포에서 이들 RNA의 성능을 향상시키는 돌연변이체 tracr RNA 및 직접 반복 서열 또는 돌연변이체 키메라 가이드 서열을 생성하는 방법을 제공한다. 본 발명의 양태는 또한 상기 서열의 선택을 제공한다.

[0039] 본 발명의 양태는 또한 CRISPR 복합체의 클로닝 및 전달을 단순화하는 방법을 제공한다. 본 발명의 바람직한 구현예에서, 적합한 프로모터, 예컨대 U6 프로모터는 DNA 올리고를 이용하여 증폭되고, 가이드 RNA 상에 첨가된다. 이어서 얻어진 PCR 산물은 세포 내로 트렌스펙션되어 가이드 RNA의 발현을 구동할 수 있다. 본 발명의 양태는 또한 시험관내 전사되거나 합성 회사로부터 주문되고 직접 트렌스펙션된 가이드 RNA에 관한 것이다.

[0040] 일 양태에서, 본 발명은 더 활성인 중합효소를 사용함으로써 활성을 개선하는 방법을 제공한다. 바람직한 구현예에서, T7 프로모터의 제어 하에서 가이드 RNA의 발현은 세포 내 T7의 발현에 의해 구동된다. 유리한 구현예에서, 세포는 진핵 세포이다. 바람직한 구현예에서, 진핵 세포는 인간 세포이다. 더 바람직한 구현예에서, 인간 세포는 환자 특이적 세포이다.

[0041] 일 양태에서, 본 발명은 Cas 효소의 독성을 감소시키는 방법을 제공한다. 특정 양태에서, Cas 효소는 본 명세서에 기재된 바와 같은 임의의 Cas9, 예를 들어 임의의 자연적으로 생기는 박테리아 Cas9뿐만 아니라 임의의 키메라, 돌연변이체, 상동체 또는 오솔로그이다. 바람직한 구현예에서, Cas9는 mRNA의 형태로 세포에 전달된다. 이는 효소를 일시적으로 발현시킴으로써 독성을 감소시킨다. 다른 바람직한 구현예에서, 본 발명은 또한 유도성 프로모터의 제어 하에서 Cas9를 발현시키는 방법 및 이에 사용되는 작제물을 제공한다.

[0042] 다른 양태에서, 본 발명은 CRISPR-Cas 시스템의 생체내 적용을 개선하는 방법을 제공한다. 바람직한 구현예에서, Cas 효소는 야생형 Cas9, 또는 임의의 자연적으로 생기는 박테리아 Cas9뿐만 아니라 임의의 키메라, 돌연변이체, 상동체 또는 오솔로그를 포함하는 본 명세서에 기재된 임의의 변형된 형태이다. 본 발명의 유리한 양태는 전달을 위해 바이러스 벡터 내로 용이하게 패키징되는 Cas9 상동체의 선택을 제공한다. Cas9 오솔로그는 통상적으로 3-4 RuvC 도메인 및 HNH 도메인의 일반적 조직화를 공유한다. 5' 대부분의 RuvC 도메인은 비상보적 가닥을 절단하고, HNH 도메인은 상보적 가닥을 절단한다. 모든 표기는 가이드 서열을 참조한다.

[0043] 5' RuvC 도메인 내 촉매적 잔기는 Cas9 오솔로그(스트렙토코커스 피오게네스 II형 CRISPR 유전자좌, 스트렙토코서스 써모필루스 CRISPR 유전자좌 1, 스트렙토코서스 써모필루스 CRISPR 유전자좌 3, 및 프란시스실라 노비시다

II형 CRISPR 유전자좌로부터의) 다른 Cas9 오솔로그와 관심 대상의 Cas9의 상동성 비교를 통해 동정되고, 보존된 Asp 잔기(D10)는 알라닌으로 돌연변이되어 상보성 가닥 닉카아제로 전환시킨다. 유사하게, HNH 도메인 내에 보존된 His 및 Asn 잔기는 알라닌으로 돌연변이되어 Cas9를 비상보성 가닥 닉킹 효소로 전환시킨다. 일부 구현예에서, 돌연변이 세트가 둘 다 만들어져서 Cas9를 비절단 효소로 전환시킬 수 있다.

[0044] 일부 구현예에서, CRISPR 효소는 I형 또는 III형 CRISPR 효소, 바람직하게는 II형 CRISPR 효소이다. 이 II형 CRISPR 효소는 임의의 Cas 효소일 수 있다. 바람직한 Cas 효소는, 이것이 II형 CRISPR 시스템으로부터의 다중 뉴클레아제 도메인을 지니는 가장 큰 뉴클레아제에 대해 상동성을 공유하는 일반적 효소 부류를 지칭할 수 있기 때문에 Cas9로서 동정될 수 있다. 가장 바람직하게는, Cas9 효소는 spCas9 또는 saCas9로부터 생기거나 또는 유래된다. 본 발명자들은 유래는, 유래된 효소가 야생형 효소와 높은 정도의 서열 동일성을 갖는 의미에서 야생형 효소에 크게 기반하지만, 이들은 본 명세서에 기재된 바와 같은 일부 방법에서 돌연변이(변형)되는 것을 의미한다.

[0045] 용어 Cas 및 CRISPR 효소는 달리 명확하지 않다면 일반적으로 본 명세서에 상호호환적으로 사용된다는 것이 인식될 것이다. 상기 언급한 바와 같이, 본 명세서에 사용되는 다수의 잔기 넘버링은 스트렙토코커스 피오게네스(*Streptotoccus pyogenes*) 내 II형 CRISPR 유전자좌로부터의 Cas9 효소를 지칭한다. 그러나, 본 발명은 SpCas9, SaCas9, St1Cas9 등과 같은 다른 미생물 종으로부터의 다수의 더 많은 Cas9를 포함하는 것으로 인식될 것이다. 본 명세서에서 추가 예가 제공된다. 당업자는 적절한 아미노산 서열의 비교에 의해 SpCas9 이외의 Cas9 효소에서 적절한 대응하는 잔기를 결정할 수 있을 것이다. 따라서, SpCas9 넘버링을 사용하기 위해 구체적 아미노산 대체가 언급되는 경우, 달리 내용에서 명확하지 않다면, 다른 Cas9 효소를 지칭하는 것으로 의도되며, 본 개시내용은 다른 Cas9 효소에서 대응하는 변형을 포함하는 것으로 의도된다. SpCas 또는 SaCas9가 특히 바람직한 Cas9 효소이다.

[0046] 이 예에서 인간에 대해 최적화된(즉, 인간에서의 발현을 위해 최적화됨), 코돈 최적화된 서열의 예가 본 명세서에 제공되며, SaCas9 인간 코돈 최적화된 서열을 참조한다. 이것이 바람직하지만, 다른 예가 가능하고 숙주종에 대한 코돈 최적화가 공지되어 있다는 것이 인식될 것이다.

[0047] 추가 구현예에서, 본 발명은 키메라 Cas9 단백질을 생성함으로써 Cas9의 기능을 향상시키는 방법을 제공한다. 키메라 Cas9 단백질 키메라 Cas9는 하나 이상의 자연적으로 생기는 Cas9로부터의 단편을 함유하는 새로운 Cas9일 수 있다. 이들 방법은 다른 Cas9 상동체의 C-말단 단편을 지니는 하나의 Cas9 상동체의 N-말단 단편을 융합시키는 단계를 포함할 수 있다. 이들 방법은 또한 키메라 Cas9 단백질에 의해 보여지는 새로운 특성의 선택을 고려한다.

[0048] 본 방법에서 유기체가 동물 또는 식물인 경우, 변형은 생체외에서 또는 시험관내에서, 예를 들어 세포 배양물에서 및 생체내가 아닌 일부 예에서 일어날 수 있다는 것이 인식될 것이다. 다른 구현예에서, 이는 생체내에서 일어날 수 있다.

[0049] 일 양태에서, 본 발명은 관심 대상의 게놈 유전자좌에서 표적 서열의 조작에 의해 유기체 또는 비인간 유기체를 변형시키는 방법으로서,

[0050] A) - I. CRISPR-Cas 시스템 RNA 폴리뉴클레오티드 서열, 선택적으로 키메라 RNA(chiRNA) 폴리뉴클레오티드 서열로서, 상기 폴리뉴클레오티드 서열은,

[0051] (a) 진핵 세포에서 표적 서열에 혼성화할 수 있는 가이드 서열,

[0052] (b) tracr 메이트 서열, 및

[0053] (c) tracr 서열을 포함할 수 있는, 폴리뉴클레오티드 서열, 및

[0054] II. 적어도 하나 이상의 핵 국소화 서열을 포함하는 CRISPR 효소를 암호화하는 폴리뉴클레오티드 서열

[0055] (여기서, (a), (b) 및 (c)는 5'에서 3' 방향으로 배열되고,

[0056] 전사될 때, tracr 메이트 서열은 tracr 서열에 혼성화하며, 가이드 서열은 표적 서열에 대한 CRISPR 복합체의 서열-특이적 결합을 지시하고,

[0057] CRISPR 복합체는 (1) 표적 서열에 혼성화되는 가이드 서열 및 (2) tracr 서열에 혼성화된 tracr 메이트 서열에 복합체화되는 CRISPR 효소를 포함하고, CRISPR 효소를 암호화하는 폴리뉴클레오티드 서열은 DNA 또는 RNA임),

[0058] 또는

[0059] (B) I. (a) 진핵세포 내 표적 서열에 혼성화할 수 있는 가이드 서열, 및

[0060] (b) 적어도 하나 이상의 tracr 메이트 서열을 포함할 수 있는, 폴리뉴클레오티드,

[0061] II. CRISPR 효소를 암호화하는 폴리뉴클레오티드 서열, 및

[0062] III. tracr 서열을 포함할 수 있는 폴리뉴클레오티드 서열,

[0063] (여기서, 전사될 때, tracr 메이트 서열은 tracr 서열에 혼성화하고, 가이드 서열은 표적 서열에 대한 CRISPR 복합체의 서열 특이적 결합을 지시하며,

[0064] CRISPR 복합체는 (1) 표적 서열에 혼성화된 가이드 서열 및 (2) tracr 서열에 혼성화된 tracr 메이트 서열과 복합체화된 CRISPR 효소를 포함하고, CRISPR 효소를 암호화하는 폴리뉴클레오티드 서열은 DNA 또는 RNA임)

[0065] 을 포함하는 비자연적으로 생기는 또는 유전자 조작된 조성물을 전달하는 단계를 포함하는 방법을 제공한다.

[0066] 일부 구현예에서, 본원에 제공된 임의의 또는 모든 양태에 적용가능하지만, 상기 (B)의 제2 대안이 바람직하다. 그러나, 제1 대안 (A)가 특히 바람직하다. 이는 2 개의 대안적 CRISPR 접근법의 특징을 이루는 본 발명의 모든 양태에 적용된다.

[0067] 본 출원은 기관 그 자체이거나 기관 내의 조직이거나 단순히 하나 이상의 표적 세포인지 아닌지 간에, 바이러스 벡터 전달에 관한 것이다. 표적 세포는 CRISPR-Cas 시스템의 전달을 위해 선택되는 것이다. 예를 들어, 간으로의 전달의 경우에, 상기 표적 세포는 헤파토사이트, 바람직하게는 일차 헤파토사이트일 수 있다. 표적 세포는 환자(CRISPR-지시된 유전자 요법을 필요로 하는 동물의 의미) 또는 모델 유기체인 척추동물 내에 존재할 수 있거나, 세포 배양물, 오가노이드 또는 다른 생체외 조직, 예컨대 헤파토사이트가 스캐폴드 상에 시딩되고 성장되는 "리버 온 어 칩(liver on a chip)"에 존재할 수 있다. 이식되지 않은 기관으로부터 회수된 헤파토사이트가 또한 유용한 표적이다. 생물학에 적용되는 3-D 프린팅 기법의 개발과 함께, 프린팅된 조직이 획득되며, 오가노이드를 생성시키기 위해 상기 방식으로 또는 칩 상에 프린팅된 간 세포 또는 조직이 또한 표적화될 수 있는 것이 전적으로 실행 가능하다. 본원에서 헤파토사이트의 논의는 다른 간 세포, 및 실제로 일반적으로 다른 세포 유형, 예컨대 뇌 또는 신장 세포에 동일하게 적용될 수 있으며, 이의 예는 본원에서 제공된다.

[0068] 따라서, 본 발명의 CRISPR-Cas 시스템이 전달된 간 세포, 예컨대 헤파토사이트를 포함할 수 있는 모델 유기체가 제공된다. 유사하게, 본 발명의 CRISPR-Cas 시스템이 전달된 2 개 이상의 간 세포, 예컨대 헤파토사이트의 생체외 집합물이 또한 제공된다. 이와 같은 집합물은 간 기관, 간 오가노이드, 간 세포 거주 스캐폴드('리버 온 어 칩')를 포함할 수 있다. 간이 바람직하지만, 본원에서 예로서 제공되어 있으므로, 또한, 물론 간이 아닌 대안, 예컨대 뇌 또는 신장이 예견된다. 상기 모델 또는 집합물을 생성시키는 방법이 또한 제공된다.

[0069] 특히, 상기 표적 세포는 Cas 효소를 발현시킬 수 있는 폴리뉴클레오티드를 발현하거나 포함할 수 있다. 본원에 논의된 바와 같이, 이는 녹다운을 포함하는 유전자 변동을 통해 유전자 기능을 질의하기 위한 즉시 이용가능한 모델을 제공하는 장점을 갖는다. 이는 간의 병태, 예컨대 본원에 나열된 아밀로이드증 등뿐만 아니라, 비만과 같은 더 광범위한 병태를 연구하는데 특히 유용하다.

[0070] 간 유전자 기능을 질의하는 방법이 또한 본원에 제공된다. 이들은 통상적으로 CRISPR-Cas 시스템을 생체내 또는 생체외에서 표적 세포에 전달하는 것을 포함한다. 그러나, 단백질로서 발현되거나 세포 내에 이미 포함된 폴리뉴클레오티드에 의해 암호화되건 간에 세포가 이미 Cas를 포함하는 경우, CRISPR 폴리뉴클레오티드만이 전달될 필요가 있다. 상기 방법은 본원에 논의된 바와 같은 표적 세포, 기관, 오가노이드, 칩 또는 세포 집합물로부터의 추출 및 선택적으로 다시 표적 세포, 기관, 오가노이드, 칩 또는 세포 집합물로의 재삽입을 포함할 수 있다. 전달은 세포의 핵으로의 폴리뉴클레오티드의 실제 물리적 전달뿐만 아니라 트랜스펙션을 의미한다.

[0071] 유전자 요법의 방법이 또한 예견된다. 예를 들어, 하나 이상의 결함이 있는 유전자형(예를 들어, 단일 점돌연변이)의 교정은 본원에 논의된 간 세포(모델을 포함함)에서의 본 발명의 CRISPR-Cas 시스템의 사용을 통해 달성가능하다. 간과 연관된 단일유전자 병태가 특히 바람직하고, 본원에 예시되어 있으며, CRISPR-Cas9 시스템 표적이 지질 대사 유전자인 ApoB인 실시예 38이 생체내에서 표현형 변화를 유도하는데 효과적이었던 것을 참조한다. 유전자 요법에서 사용하기 위한 조성물이 또한 제공된다.

[0072] 연구 및 유전자 요법에 대한 병태는, CRIPS-Cas 기술의 광범위한 적용으로 인하여 다수 존재하고, 다양하다. 적합한 예가 표 A, B 및 C를 포함하여, 본원에 제공되어 있다. 이들 중 임의의 것이 선택될 수 있으며, 각각은 바

람직한 것이다. 비제한적이지만, 소수의 특히 바람직한 예는 본원에 구체적으로 예시된 병태뿐만 아니라, 임의의 단일유전자 병태이며, 특히 낭포성 섬유증(CFTR)이다.

[0073] 다양한 Cas 효소가 예견되지만, Cas9가 특히 바람직하며, 출원인은 SaCa9에 대하여 간에서 특히 효능을 나타내었다. Cas 효소가 Sa Cas 효소라면, Sa로부터의 tracr 서열이 또한 바람직하다. 이와 같은 상황에서 적합한 PAM은 NNGRR이다.

[0074] 1 개의 가이드가 사용될 수 있으나, 2 개, 3 개, 4 개 또는 그 초과의 가이드를 이용한 소위 다중화가 유전자 기능의 질의 및 모델 생성(다수의 유전자 녹다운을 제공하기 위함)뿐만 아니라 다수의 결함이 있는 유전자형이 교정(단일 유전자 내의 다수의 에러 또는 더욱 가능하게는 여러 유전자에 걸쳐 분산된 다수의 에러)되어야 하는 유전자 요법에서 특히 유용하다. 대안적으로, 2 개의 가이드를 이용한 다중화가 Cas 동원을 보장하기 위해 1 개의 유전자 내의 다수의 표적의 간단한 선택 또는 표적외 효과를 감소시키기 위한 이중 닉카아제 접근법에 유용하다. 삼중 및 사중 가이드가 바람직하다. 본원에서 유전자는 게놈 유전자좌와 상호교환적으로 언급된다.

[0075] 본원에 기재된 인트론 접근법이 또한 이와 관련하여 유용하며, 여기서 가이드는 Cas 인트론 내에 위치된다.

[0076] 바람직한 전달 수단은 하기 Kanasty에 의해 기재된 방법, 예컨대 특히 가이드만이 전달되어야 하거나 단독으로 전달되어야 하는 LNP를 포함한다. 그러나, 렌티바이러스 및 AAV를 포함하는 바이러스 벡터가 일반적으로 바람직하다. 특히, 이들은 현재까지 성공적이었기 때문에 간에의 전달에 대해 바람직하다. 이들 중, AAV, 특히 혈청형 8이 바람직하며, AAV2/8이 효과적인 것으로 나타났다.

[0077] 일부 바람직한 표적 병태 및 유전자(이들이 간 또는 신장에 존재하거나 간 또는 신장의 병태인 한)는 대사 장애, 예컨대 아밀로이드 신경병증(TTR, PALB); 아밀로이드증(APOA1, APP, AAA, CVAP, AD1, GSN, FGA, LYZ, TTR, PALB); 간경화증(KRT18, KRT8, CIRH1A, NAIC, TEX292, KIAA1988); 낭포성 섬유증(CFTR, ABCC7, CF, MRP7); 글리코겐축적병(SLC2A2, GLUT2, G6PC, G6PT, G6PT1, GAA, LAMP2, LAMPB, AGL, GDE, GBE1, GYS2, PYGL, PFKM); 간샘종, 142330(TCF1, HNF1A, MODY3), 간부전, 조기 발병, 및 신경학적 장애(SCOD1, SCO1), 간리파제 결핍(LIPC), 간모세포종, 암 및 암종(CTNNB1, PDGFRL, PDGRL, PRLTS, AXIN1, AXIN, CTNNB1, TP53, P53, LFS1, IGF2R, MPRI, MET, CASP8, MCH5; 수질 낭성 신장병(UMOD, HNFJ, FJHN, MCKD2, ADMCKD2); 페닐케톤뇨증(PAH, PKU1, QDPR, DHPR, PTS); 다낭신장 및 간질환(FCYT, PKHD1, ARPKD, PKD1, PKD2, PKD4, PKDTS, PRKCSH, G19P1, PCLD, SEC63) 중 어느 하나이다. 다른 바람직한 표적은 PCSK9; Hmgcr; SERPINA1; ApoB; 및/또는 LDL 중 어느 하나 이상을 포함한다.

[0078] 표적 세포에서의 발현을 변화시키는 방법이 윤리적 입장에서 배제될 수 있는 생식계열의 변화를 포함하지 않는 것이 인식될 것이다. 사실, 줄기 세포의 트랜스펙션이 일부 구현예에서 예견되고 확실히 바람직하나, 비줄기 세포(즉, 유사분열 후 세포)가 특히 이들이 일부 재생을 나타내거나 헤파토사이트에서 나타내기 위해 자극될 수 있거나 이를 나타낼 수 있는 경우에 특히 바람직하다.

[0079] 타입 II CRISPRS가 특히 진핵생물에서의 사용에 특히 바람직한데, 이는 본 발명의 경우에서, 간은 어떠한 경우에서도 진핵생물, 특히 척추 동물에서만 발견되기 때문이다.

[0080] 표현형 변화를 발생시키기 위한 CRISPR-Cas 시스템의 사용이 특히 생체내에서 특히 유리하다.

[0081] 치료 적용이 예견되는 경우, 또는 표적 세포에서 다른 게놈 유전자 조작을 위해, 교정이 필요한 경우, 하기 게놈 DNA 표적의 닉킹 또는 절단 후, HDR 경로를 통한 교정이 바람직함이 인식될 것이다. 유전자 녹다운을 위해, NHEJ가 유리하나, HDR 경로를 통한 교정이 요법에 바람직하다. 상기 상황에서, 수선 주형을 전달하는 것이 바람직하다. 이는 ssDNA가 가장 바람직하나, 상응하는 DNA 주형을 제공하기 위한 레트로바이러스 벡터를 통한 RNA가 또한 가능하다. 당업자는 당 분야의 지식에 기여하는 본원의 교시내용으로부터 본 발명을 용이하게 실시할 수 있으며, 이와 관련하여 당 분야의 지식에 기여하는 본원의 교시내용으로부터 당업자가 상동성 암(arm) 길이에 관한 고려사항을 용이하게 인식하고 실행할 수 있는 것이 언급된다. 본원에 인용된 것을 포함하며, 본원의 발명자 Zhang을 포함하는 특허 출원 및 간행물이 언급된다. 수선 주형은 바람직하게는 CRISPR-Cas 시스템의 하나 이상의 요소와 공동 전달된다.

[0082] a) 표적 서열과 혼성화하는 CRISPR-Cas 시스템 가이드 RNA를 암호화하는 적어도 하나의 뉴클레오티드 서열에 작동 가능하게 연결되는 진핵 세포에서 작동 가능한 제1 조절 요소, 및

[0083] b) 타입-II Cas9 단백질을 암호화하는 뉴클레오티드 서열에 작동 가능하게 연결된 진핵 세포에서 작동 가능한 제2 조절 요소

[0084] 를 포함할 수 있는 하나 이상의 벡터를 포함할 수 있는 유전자 조작된 비자연적으로 생기는 클러스터링되고 규칙적으로 산재된 짧은 팔린드로믹 반복부(CRISPR)-CRISPR 회합된(Cas)(CRISPR-Cas) 시스템을, 세포 표적 서열을 갖고 유전자 산물을 암호화하는 DNA 분자를 함유하고 발현하는 진핵 세포로 도입시키는 것을 포함할 수 있는, 적어도 하나의 간 유전자 산물의 발현을 변화시키는 방법이 또한 제공되며,

[0085] 상기 구성성분 (a) 및 (b)는 시스템의 동일 또는 상이한 벡터 상에 위치됨으로써, 가이드 RNA가 표적 서열을 표적화하고, Cas9 단백질이 DNA 분자를 절단하며, 이에 의해 적어도 하나의 간 유전자 산물의 발현이 변화되고, Cas9 단백질 및 가이드 RNA는 함께 자연적으로 생기지 않는다.

[0086] 표적에 대한 하기의 언급은 달리 명백하지 않은 한 유전자 또는 세포, 통상적으로는 유전자를 언급하는 것으로 이해될 것이다.

[0087] 하기는 본 발명의 모든 양태에 동등하게 적용된다. 간이 본원에서 언급될 수 있는 경우, 이는 일반적으로 유사분열 후 세포, 특히 신장 또는 뇌 세포에 대한 언급인 것으로 이해된다. 표적 서열은 가장 바람직하게는 유사분열 후 세포 표적 서열이다. 유사분열 후 세포는 하기 기관 중 어느 하나에 속하거나 이로부터 유래될 수 있거나 (즉, 세포 또는 세포 유형의 기원), 이는 오가노이드 또는 생체외 모델 또는 하기의 세포를 포함하는 세포의 수집물일 수 있다:

[0088] 신장, 예를 들어, 사구체 세포;

[0089] 위, 췌장, 십이지장, 회장 및/또는 결장을 포함하는 소화계통;

[0090] 심장;

[0091] 폐;

[0092] 뇌, 특히 뉴런, 및/또는 일반적으로 CNS;

[0093] 망막 조직을 포함하는 눈;

[0094] 내이를 포함하는 귀;

[0095] 피부;

[0096] 근육;

[0097] 뼈; 및/또는

[0098] 일반적으로, 간(이는 일부 구현예에서는 배제되나, 이는 또한 별개의 출원의 주제이다)

[0099] 뇌 및 신장이 특히 바람직하다. 일부 구현예에서, 세포는 뇌 세포, 예를 들어, 뉴런이다. 일부 구현예에서, 세포는 신장 세포이다.

[0100] 바람직한 신장 세포는 하기 중 어느 하나 이상을 포함한다:

[0101] · 신장 사구체 벽세포;

[0102] · 신장 사구체 발세포;

[0103] · 신장 토리쪽세관 솔가장자리 세포;

[0104] · 헨레 고리 세부 세포;

[0105] · 두꺼운 상행각 세포;

[0106] · 신장 원위세관 세포;

[0107] · 신장 집합관 세포; 및

[0108] · 사이질 신장 세포.

[0109] 표적 세포의 바람직한 예는, 예를 들어 '신장' 또는 '간' 또는 '뼈' 또는 '귀'를 제목으로 하는 적절한 섹션(이들 중 임의의 것이 바람직함) 하에 하기 표뿐만 아니라 표 B에 제공된다. 이들 표적 중 어느 하나 이상이 바람직하다. 실시예 1 및 18은 또한 신장 세포(유사분열 후 세포가 아닌 줄기 세포이긴 함)를 표적으로 하나, 전달에 관한 교시내용이 적용될 수 있다.

[0110] 일부 특히 바람직한 구현예에서, 조작은 세포에서 표현형 변화를 발생시킨다.

[0111] 일부 구현예에서, 표현형 변화는 생체내에서 세포 내에서 발생되거나 유지될 수 있다. 세포가 생체내에서 트랜스펙션되거나, 추출되고, 생체외에서 트랜스펙션된 후, 동일하거나 상이한 숙주로 다시 재삽입(이식)된다.

[0112] CRISPR 효소, 및 선택적으로 가이드 서열의 발현은, 예를 들어, 유사분열 후 세포에서 상기 효소 및 선택적인 가이드를 발현시킬 수 있는 발현 카세트 내에 포함된 세포에 특이적인 프로모터의 조절하에 있을 수 있다. 즉, CRISPR 효소, 및 선택적으로 가이드 서열은 표적 세포에 특이적인 상기 프로모터에 작동 가능하게 연결된다.

[0113] 표적 세포는 유사분열 후 세포일 수 있다. 특히, 유사분열 후 세포가 뉴런인 경우 AAV 벡터 시스템이 특히 바람직하다. 체세포 또한 바람직하다.

[0114] CRISPR 효소에 대한 프로모터 및 가이드 서열에 대한 임의의 프로모터는 동일하거나 상이할 수 있다.

[0115] 본원의 논의, 특히 하기는 또한 본원에 기재된 임의의 방법, 용도 또는 조성물에 적용된다. CRISPR-Cas 시스템 RNA는 키메라 RNA(chiRNA)일 수 있다. CRISPR-Cas 시스템은 다중 키메라 및/또는 다중 멀티가이드 서열 및 단일 tracr 서열을 추가로 포함하는 다중화 CRISPR 효소 시스템일 수 있다. CRISPR 효소는 표적 서열의 위치에서 두 가닥 모두의 절단을 유도하는 뉴클레아제일 수 있다. CRISPR 효소는 하나 이상의 돌연변이를 포함할 수 있다. CRISPR 효소는 하나 이상의 돌연변이 D10A, E762A, H840A, N854A, N863A 또는 D986A를 포함할 수 있다. 하나 이상의 돌연변이는 CRISPR 효소의 RuvC1 도메인에 존재할 수 있다. CRISPR 효소는 표적 서열의 위치에서 절단을 유도하는 닉카아제(nickase)일 수 있다. 닉카아제는 이중 닉카아제일 수 있다. 적어도 2개 이상의 NLS가 바람직하다.

[0116] CRISPR 효소는 타입 II, 바람직하게는 Cas, 가장 바람직하게는 Cas9일 수 있다. Cas 또는 Cas9에 대한 언급(예를 들어, CRISPR-Cas 또는 CRISPR Cas9)는 임의의 Cas, 가장 바람직하게는 Cas9 및 특히 Sa 또는 Sp Cas9인 것으로 이해될 것이다(제공된 DSB, 닉카아제 또는 이중 닉카아제 기능에 대한 D10A와 같은 모든 돌연변이를 포함함).

[0117] CRISPR 효소는 촉매 도메인 내에 하나 이상의 돌연변이를 가질 수 있고, 전사될 때, tracr 메이트 서열은 tracr 서열에 혼성화하고, 가이드 서열은 표적 서열에 대한 CRISPR 복합체의 서열-특이적 결합을 지시하며, 효소는 기능성 도메인을 추가로 포함한다. 기능성 도메인은 전사 활성화 도메인일 수 있다. 전사 활성화 도메인은 VP64일 수 있다.

[0118] 상기 방법은 세포 내의 관심 대상의 게놈 유전자좌에서 DNA 듀플렉스의 마주하는 가닥 상의 제1 및 제2 표적 서열의 조작에 의해 표적외 변형을 최소화시키는 것을 추가로 포함할 수 있으며, 상기 방법은,

[0119] I. CRISPR-Cas 시스템 키메라 RNA(chiRNA) 폴리뉴클레오티드 서열로서, 상기 폴리뉴클레오티드 서열은,

[0120] (a) 제1 표적 서열에 혼성화할 수 있는 제1 가이드 서열,

[0121] (b) 제1 tracr 메이트 서열,

[0122] (c) 제1 tracr 서열,

[0123] (d) 제2 표적 서열에 혼성화할 수 있는 제2 가이드 서열,

[0124] (e) 제2 tracr 메이트 서열, 및

[0125] (f) 제2 tracr 서열을 포함하는, 폴리뉴클레오티드 서열,

[0126] (선택적으로, 링커 서열은 제1 tracr 서열과 제2 가이드 서열 사이에 존재함으로써, 제1 가이드 서열 및 제2 가이드 서열은 연쇄적으로 존재함),

[0127] II. 적어도 하나 이상의 핵 국소화 서열을 포함하는 CRISPR 효소를 암호화하는 폴리뉴클레오티드 서열

[0128] ((a), (b), (c), (d), (e) 및 (f)는 5'에서 3' 방향으로 배열되고, 폴리뉴클레오티드 서열은 제1 tracr 서열과 제2 가이드 서열 사이에 링커 서열을 포함함으로써, 제1 가이드 서열 및 제2 가이드 서열은 연쇄적으로 존재하고, 전사될 때, 제1 및 제2 tracr 메이트 서열은 제1 및 제2 tracr 서열에 각각 혼성화하고, 제1 및 제2 가이드 서열은 제1 및 제2 표적 서열에 대한 제1 및 제2 CRISPR 복합체의 서열-특이적 결합을 각각 지시함),

[0129] 또는

[0130] II. CRISPR 효소를 암호화하는 효소 암호화 서열에 작동 가능하게 연결된 제2 조절 구성요소

[0131] (구성성분 I 및 II는 시스템의 동일 또는 상이한 벡터 상에 위치되며, 전사될 때, 제1 tracr 메이트 서열은 제1 tracr 서열에 혼성화하고, 제1 및 제2 가이드 서열은 제1 및 제2 표적 서열에 대한 제1 및 제2 CRISPR 복합체의 서열-특이적 결합을 각각 지시하며,

[0132] 제1 CRISPR 복합체는 (1) 제1 표적 서열에 혼성화된 제1 가이드 서열, 및 (2) 제1 tracr 서열에 혼성화된 제1 tracr 메이트 서열과 복합체화된 CRISPR 효소를 포함하고,

[0133] 제2 CRISPR 복합체는 (1) 제2 표적 서열에 혼성화된 제2 가이드 서열, 및 (2) 제2 tracr 서열에 혼성화된 제2 tracr 메이트 서열과 복합체화된 CRISPR 효소를 포함하고,

[0134] CRISPR 효소를 암호화하는 폴리뉴클레오티드 서열은 DNA 또는 RNA이고,

[0135] 제1 가이드 서열은 제1 표적 서열 근처의 DNA 듀플렉스의 한 가닥의 절단을 지시하고, 제2 가이드 서열은 이중 가닥 파손을 유도하는 제2 표적 서열 근처의 다른 가닥의 절단을 지시함으로써 표적외 변형을 최소화하여 유기체 또는 비인간 유기체를 변형시킴)

[0136] 을 포함하는 비자연적으로 생기는 또는 유전자 조작된 조성물을 전달하는 단계를 포함한다.

[0137] 일부 구현예에서, 상기 (B)의 제2 대안이 바람직하다. 그러나, 제1 대안 (A)가 특히 바람직하다. 이는 2개의 대안적 CRISPR 접근법의 특징을 이루는 본 발명의 모든 양태에 적용된다.

[0138] 본 출원은 기관 그 자체이거나 기관 내의 조직이거나 단순히 하나 이상의 유사분열 후 세포, 예를 들어, 뉴런인지 아닌지 간에 유사분열 후 세포에 관한 것이다. 뉴런 및 신장 세포가 바람직하다. 유사분열 후 세포는 환자(CRISPR-지시된 유전자 요법을 필요로 하는 동물의 의미) 또는 모델 유기체인 척추동물 내에 존재할 수 있거나, 세포 배양물, 오가노이드 또는 다른 생체외 조직, 예를 들어, 간세포가 스캐폴드 상에 시딩되고 성장되는 "리버 온 어 칩(liver on a chip)"에 존재할 수 있다. 이식되지 않은 기관으로부터 수거된 간세포가 또한 유용한 표적이다. 생물학에 적용되는 3-D 프린팅 기술의 개발과 함께, 프린팅된 조직이 획득되며, 오가노이드를 생성시키기 위해 상기 방식으로 칩 상에 프린팅된 간 세포 또는 조직이 또한 표적화될 수 있는 것이 전적으로 실행 가능하다. 간이 아닌 대안, 특히 신장 조직 또는 다른 유사분열 후 세포/조직이 또한 예견된다.

[0139] 따라서, 본 발명의 CRISPR-Cas 시스템이 전달된 유사분열 후 세포, 예를 들어, 뉴런 또는 신장 세포를 포함하는 모델 유기체가 제공된다. 유사하게, 본 발명의 CRISPR-Cas 시스템이 전달된 2개 이상의 유사분열 후 세포, 예를 들어, 뉴런 또는 신장 세포의 생체외 집합물이 또한 제공된다. 이러한 집합물은 유사분열 후 기관, 오가노이드, 세포 거주 스캐폴드('키드니 온 어 칩(kidney on a chip)')를 포함할 수 있다. 상기 모델 또는 집합물을 생성시키는 방법이 또한 제공된다.

[0140] 특히, 상기 유사분열 후 세포는 Cas 효소를 발현시킬 수 있는 폴리뉴클레오티드를 발현하거나 포함할 수 있다. 본원에 논의된 바와 같이, 이는 녹다운을 포함하는 유전자 변동을 통해 유전자 기능을 질의하기 위한 즉시 이용 가능한 모델을 제공하는 장점을 갖는다. 이는 본원에 나열된 것과 같은 유사분열 후 세포, 예를 들어, 신장 또는 뇌 세포의 질환뿐만 아니라 비만과 같은 더 광범위한 질환을 연구하는데 특히 유용하다.

[0141] 유사분열 후 세포 유전자 기능을 질의하는 방법이 또한 본원에 제공된다. 이들은 통상적으로 CRISPR-Cas 시스템을 생체내 또는 생체외에서 유사분열 후 세포에 전달하는 것을 포함한다. 그러나, 단백질로서 발현되거나 세포 내에 이미 포함된 폴리뉴클레오티드에 의해 암호화되건 간에 세포가 이미 Cas를 포함하는 경우, CRISPR 폴리뉴클레오티드만이 전달될 필요가 있다. 상기 방법은 유사분열 후 세포로부터의 추출 및 선택적으로 다시 유사분열 후 세포로의 재삽입을 포함할 수 있다. 전달은 세포의 핵으로의 폴리뉴클레오티드의 실제 물리적 전달뿐만 아니라 트랜스펙션을 의미한다. 따라서, 전달은 달리 명백하지 않은 한 트랜스펙션을 포함하는 것으로 해석되어야 한다.

[0142] 세포의 제1 집단에 본 발명에 따른 CRISPR-Cas 시스템을 형질도입시킴으로써 세포의 제1 집단의 게놈을 변경시켜 세포의 제2 집단을 수득하는 것을 포함하는, 하나 이상의 동물 또는 식물 세포에서 유전자 변동을 유도하는 방법이 또한 제공된다. 상기 방법은 생체외 또는 생체내, 예를 들어, 세포 배양물 내 또는 생체외 또는 시험관내 모델(예를 들어, 오가노이드 또는 '칩 상의 동물 또는 식물 세포(animal or plant cell on a chip)')에서 수행될 수 있다. 대안적으로, 상기 방법은 생체내에서 수행될 수 있으며, 이 경우 이는 또한 대상체로부터 세포의 제1 집단을 분리시키고, 세포의 제2 집단을 대상체로 (다시) 이식하는 것을 포함할 수 있다. 유전자 변동은 1개 이상, 또는 2개 이상, 또는 3개 이상, 또는 4개 이상 유전자에 대한 변동일 수 있다. 유전자 변동은 유전자

기능(즉, 암호화된 유전자 생성물에서의 활성)의 감소일 수 있다. 이는, 예를 들어, 세포의 제1 집단의 게놈의 변화를 통해 유도되어 세포의 제2 집단이 수득될 수 있으며, 상기 세포의 제2 집단은 세포의 제1 집단에서 부재하는 단일유전자 질환과 같은 결함이 있는 유전자형을 갖는다. 이는 결함이 있는 서열을 제공하기 위해 본원에 논의된 바와 같이 상응하는 수선 주형을 필요로 할 수 있거나, 이는 DSB의 유도를 통할 수 있다. 특히, 유전자 변동은 유전자 녹다운이다. 일부 구현예에서, 동물 또는 식물 세포는 가장 바람직하게는 유사분열 후 세포, 예컨대 신장 또는 뇌(뉴런) 세포 또는 간 세포, 예를 들어, 일차 간세포이다.

[0143] 대안적으로, 유전자 변동은 유전자 기능(즉, 암호화된 유전자 생성물의 활성)의 증가일 수 있다. 이는, 예를 들어, 세포의 제2 집단을 수득하기 위해 세포의 제1 집단의 게놈의 변화를 통해 유도될 수 있으며, 세포의 제1 집단은 세포의 제2 집단에 부재하는(즉, 세포의 제2 집단에 대해 교정되는) 단일유전자 질환과 같은 결함이 있는 유전자형을 갖는다. 이는 교정된 서열을 제공하기 위해 본원에 논의된 바와 같은 상응하는 수선 주형을 필요로 할 수 있다.

[0144] 다중화가 이용되는 경위, 하나 이상의 유전자의 감소 및 하나 이상의 유전자의 증가의 혼합이 예견된다. 이는 (다중복합체 내) 가이드 중 하나 이상의 제공을 통해 달성될 수 있고, 기능을 감소시키기 위해 상응하는 수선 주형이 이용될 수 있는 한편, 기능을 증가시키기 위해 가이드 및 이의 상응하는 주형 중 하나 이상이 이용될 수 있다.

[0145] 동물 또는 식물 세포의 제1 집단에서 하나 이상의 유전자의 발현에서의 변화를 결정하고, 변화된 게놈(또는 유전자형)을 갖는 상기 제2 집단을 제공하기 위해 상기 제1 집단에서 상기 유전자 변동을 유도하고, 동물 또는 식물 세포의 제2 집단에서 하나 이상의 유전자의 발현에서의 변화를 결정함으로써 하나 이상의 유전자의 기능을 질의하는 것을 포함하는, 하나 이상의 동물 또는 식물 세포에서 하나 이상의 유전자의 기능을 질의하는 방법이 또한 제공된다. 일부 구현예에서, 동물 또는 식물 세포는 가장 바람직하게는 유사분열 후 세포, 예컨대 신장 또는 뇌(뉴런) 세포 또는 간 세포, 예를 들어, 일차 간세포이다.

[0146] 모델 및 상기 모델을 생성시키는 방법이 또한 제공된다. 상기 모델은 동물 또는 식물 세포를 포함하는 동물(생체내 모델)일 수 있거나, 이는 본원에 기재된 바와 같이 생체외 또는 시험관내 모델, 예를 들어, 동물 또는 식물 오가노이드 또는 '칩 상 동물 또는 식물 세포' 또는 동물 또는 식물 세포의 수집물, 예를 들어, 스캐폴드 상 동물 또는 식물 세포의 수집물일 수 있다. 어느 하나의 모델의 동물 또는 식물 세포는 바람직하게는 Cas9로 트랜스펙션될 것이다. 따라서, CRISPR 효소, 바람직하게는 Cas9, 예를 들어, Sa 또는 SpCas9를 포함하는 하나 이상의 동물 또는 식물 세포를 포함하는 모델이 특별히 제공된다. 모델 세포는 본원에 제공된 제2 조절 구성요소로 트랜스펙션되거나 형질도입될 수 있으며, 이는 적어도 하나 이상의 핵 국소화 서열(NLS)을 포함하는 CRISPR 효소를 암호화하는 효소 암호화 서열에 작동 가능하게 연결된 제2 조절 구성요소이다. 상기 모델은 상기 기재된 바와 같이 생체내 모델일 수 있거나, 이는 생체외 또는 시험관내 모델일 수 있다. 상기 모델은 하나 이상의 유전자의 기능의 신속한 질의를 가능케 하는데, 이는 상기 유전자의 기능을 교란시키기 위해 단지 CRISPR-Cas 시스템 폴리뉴클레오티드 서열(상기 하나 이상의 유전자를 표적화하는 하나 이상의 가이드 서열을 포함함)이 전달되는 것을 필요로 하기 때문이다. 즉, 상기 모델에서 유전자 기능을 질의하는 방법은 CRISPR-Cas 시스템 폴리뉴클레오티드 서열(하나 이상의 가이드 서열을 포함함)의 전달만을 포함할 수 있으며, Cas(CRISPR 효소)는 모델의 세포(들) 내에 이미 제공되어 있다. 본원에 기재된 바와 같은 적어도 하나 이상의 핵 국소화 서열(NLS)을 포함하는 CRISPR 효소를 암호화하는 효소 암호화 서열에 작동 가능하게 연결된 제2 조절 구성요소로 동물 또는 식물 세포의 제1 집단 내의 하나 이상의 동물 또는 식물 세포를 형질도입시키거나 트렌스펙션시킴으로써, CRISPR 효소를 포함하거나 이를 발현하는 동물 또는 식물 세포의 하나 이상의 제2 집단을 제공하는 것을 포함하는 상기 모델을 생성시키는 방법이 또한 제공된다. 일부 구현예에서, 동물 또는 식물 세포는 가장 바람직하게는 유사분열 후 세포, 예컨대 신장 또는 뇌(뉴런) 세포 또는 간 세포, 예를 들어, 일차 간세포이다.

[0147] 유전자 교란된 모델, 특히 유전자 녹다운 모델을 생성시키는 방법이 또한 제공된다. 이들 방법은 통상적으로 세포의 제1 집단에서 본원에 기재된 바와 같이 하나 이상의 유전자에서 유전자 교란을 유도함으로써 변화된 게놈(또는 유전자형)을 갖는 세포의 제2 집단을 제공하는 것을 포함할 수 있다. 이후, 세포의 제2 집단은, 예를 들어, 스캐폴드로 또는 칩 상으로 시딩됨으로써 생체외 또는 시험관내 모델을 제공할 수 있다. 대안적으로, 제2 집단은 생체내 동물 내에 포함될 수 있다.

[0148] 유전자 요법의 방법이 또한 예견된다. 예를 들어, 하나 이상의 결함이 있는 유전자형(예를 들어, 단일 점돌연변이)의 교정은 본원에 논의된 유사분열 후 세포(모델을 포함함)에서의 본 발명의 CRISPR-Cas 시스템의 이용을 통해 달성가능하다. 유사분열 후와 관련된 단일유전자 질환이 특히 바람직하고, 본원에 예시되며, CRISPR-Cas9 시

스템 표적이 지질 대사 유전자인 ApoB인 실시예 36이 생체내에서 표현형 변화를 유도하는데 효과적이었던 것을 참조하라. 실시예 38이 또한 본 발명의 시스템으로 형질도입된 마우스의 뇌에서 생체내에서 관찰된 표현형 행동 변화와 관련하여 도움이 된다. 유전자 요법에서 사용하기 위한 조성물이 또한 제공된다.

[0149] 다양한 Cas 효소가 예견되나, Cas9가 특히 바람직하고, 본 발명자는 SaCa9에 대해 간에서 특히 효능을 제시하였다. Cas 효소가 Sa Cas 효소인 경우 Sa로부터의 Tracr 서열이 또한 바람직하다. 상기 상황에서 적합한 PAM은 NNGRR이다.

[0150] 1개의 가이드가 이용될 수 있으나, 2개, 3개, 4개 또는 그 초과의 가이드를 이용한 소위 다중화가 유전자 기능의 교란 및 모델 생성(다수의 유전자 녹다운을 제공하기 위함)뿐만 아니라 다수의 결함이 있는 유전자형이 교정(단일 유전자 내의 다수의 에러 또는 더욱 가능하게는 여러 유전자에 걸쳐 분산된 다수의 에러)되는 유전자 요법에서 특히 유용하다. 대안적으로, 2개의 가이드를 이용한 다중화가 Cas 동원을 보장하기 위해 1개의 유전자 내의 다수의 표적의 간단한 선택 또는 표적외 효과를 감소시키기 위한 이중 닉카아제 접근법에 유용하다. 3중 및 4중 가이드가 바람직하다. 본원에서 유전자는 게놈 유전자좌와 상호교환적으로 언급된다.

[0151] 본원에 기재된 인트론 접근법이 또한 이와 관련하여 유용하며, 여기서 가이드는 Cas 인트론 내에 위치된다.

[0152] 바람직한 전달 수단은 하기 Kanasty에 의해 기재된 방법, 예를 들어, 특히 가이드만이 전달되어야 하거나 단독으로 전달되어야 하는 LNP를 포함한다. 그러나, 렌티바이러스 및 AAV를 포함하는 바이러스 벡터가 간에 대해 일반적으로 바람직한데, 이는 이들이 현재까지 성공적이었기 때문이다. 이들 중, AAV, 특히 혈청형 8이 바람직하며, AAV2/8이 효과적인 것으로 밝혀졌다. 일부 바람직한 표적(이들이 신장에 존재하거나 신장의 질환인 한)은 대사 장애, 예를 들어, 하기 중 어느 하나이다: 아밀로이드 신경병증(TTR, PALB); 아밀로이드증(APOA1, APP, AAA, CVAP, AD1, GSN, FGA, LYZ, TTR, PALB); 간경화증(KRT18, KRT8, CIRH1A, NAIC, TEX292, KIAA1988); 낭성 섬유증(CFTR, ABCC7, CF, MRP7); 글리코겐축적병(SLC2A2, GLUT2, G6PC, G6PT, G6PT1, GAA, LAMP2, LAMPB, AGL, GDE, GBE1, GYS2, PYGL, PFKM); 간샘종, 142330(TCF1, HNF1A, MODY3), 간부전, 조기 발병, 및 신경학적 장애(SCOD1, SCO1), 간 리파제 결핍(LIPC), 간모세포종, 암 및 암종(CTNNB1, PDGFRL, PDGRL, PRLTS, AXIN1, AXIN, CTNNB1, TP53, P53, LFS1, IGF2R, MPRI, MET, CASP8, MCH5; 수질 낭성 신장병(UMOD, HNFJ, FJHN, MCKD2, ADMCKD2); 페닐케톤뇨증(PAH, PKU1, QDPR, DHPR, PTS); 다낭신장 및 간질환(FCYT, PKHD1, ARPKD, PKD1, PKD2, PKD4, PKDTS, PRKCSH, G19P1, PCLD, SEC63). 다른 바람직한 표적은 하기 중 어느 하나 이상을 포함한다: PCSK9, HMGCR, APOB, LDLR, ANGPTL3, F8, F9/FIX, AAT, FAH, HPD, TAT, ATP7B, UGT1A1, OTC, ARH.

[0153] 유사분열 후 세포에서의 발현을 변화시키는 방법이 윤리적 입장에서 배제될 수 있는 생식계열의 변화를 포함하지 않는 것이 인지될 것이다. 사실, 줄기 세포의 트랜스펙션이 일부 구현예에서 예견되고 확실히 바람직하나, 뉴런 또는 신장 세포가 특히 이들이 일부 재생을 나타내거나 나타내기 위해 자극될 수 있는 경우에 특히 바람직하다.

[0154] 타입 II CRISPRS가 특히 진핵생물에서의 사용에 특히 바람직한데, 이는 본 발명의 경우에서, 간은 어떠한 경우에서도 진핵생물, 특히 척추동물에서만 발견되기 때문이다.

[0155] 표현형 변화를 발생시키기 위한 CRISPR-Cas 시스템의 이용이 특히 생체내에서 특히 유리하다. 본 발명자는 본 출원에서 상기를 나타내었다.

[0156] 치료 적용이 예견되는 경우, 또는 유사분열 후 세포에서 다른 게놈 유전자 조작을 위해, 교정이 필요한 경우, 하기 게놈 DNA 표적의 닉킹 또는 절단 후, HDR 경로를 통한 교정이 바람직함이 인지될 것이다. 유전자 녹다운을 위해, NHEJ가 유리하나, HDR 경로를 통한 교정이 요법에 바람직하다. 상기 상황에서, 수선 주형을 전달하는 것이 바람직하다. 이는 ssDNA가 가장 바람직하나, 상응하는 DNA 주형을 제공하기 위한 레트로바이러스 벡터를 통한 RNA가 또한 가능하다. 당업자는 당 분야의 지식에 기여하는 본원의 교시내용으로부터 본 발명을 용이하게 실시할 수 있으며, 이와 관련하여 당 분야의 지식에 기여하는 본원의 교시내용으로부터 당업자가 상동성 암(arm) 길이에 관한 고려사항을 용이하게 인지하고 실행할 수 있는 것이 언급된다. 본원에 인용된 것을 포함하며, 본원의 발명자 Zhang을 포함하는 특허 출원 및 간행물이 언급된다. 수선 주형은 바람직하게는 CRISPR-Cas 시스템의 하나 이상의 구성요소와 공동 전달된다.

[0157] a) 표적 서열과 혼성화하는 CRISPR-Cas 시스템 가이드 RNA를 암호화하는 적어도 1개의 뉴클레오티드 서열에 작동 가능하게 연결된 진핵 세포에서 작동 가능한 제1 조절 구성요소, 및

[0158] b) 타입-II Cas9 단백질을 암호화하는 뉴클레오티드 서열에 작동 가능하게 연결된 진핵 세포에서 작동 가능한

제2 조절 구성요소를 포함하는,

[0159] 하나 이상의 벡터를 포함하는 유전자 조작된 비자연적으로 생기는 클러스터링되고 규칙적으로 산재된 짧은 팔린드로믹 반복부(CRISPR)-CRISPR 회합된(Cas)(CRISPR-Cas) 시스템을 표적 서열을 갖고 유전자 생성물을 암호화하는 DNA 분자를 함유하고 발현하는 진핵 간 세포, 예를 들어, 간세포로 도입시키는 것을 포함하는, 적어도 하나의 유사분열 후 세포 유전자 생성물의 발현을 변화시키는 방법이 또한 제공되며,

[0160] 상기 구성요소 (a) 및 (b)는 시스템의 동일하거나 상이한 벡터 상에 위치됨으로써, 가이드 RNA가 표적 서열을 표적화하고, Cas9 단백질이 DNA 분자를 절단하고, 이에 의해 적어도 하나의 유사분열 후 세포 유전자 생성물의 발현이 변화되고, Cas9 단백질 및 가이드 RNA는 함께 자연적으로 생기지 않는다.

[0161] 표적에 대한 하기의 언급은 달리 명백하지 않은 한 유사분열 후 세포 표적 또는 유사분열 후 세포에서 달리 발현된 유전자인 것으로 이해될 것이다.

[0162] 임의의 또는 모든 CRISPR 효소를 암호화하는 폴리뉴클레오티드 서열, 가이드 서열, tracr 메이트 서열 또는 tracr 서열은 RNA일 수 있다. CRISPR 효소를 암호화하는, 가이드 서열, tracr 메이트 서열 또는 tracr 서열을 암호화하는 폴리뉴클레오티드는 RNA일 수 있고, 리포좀, 나노입자, 엑소좀, 미세소포, 또는 유전자총을 통해 전달될 수 있다.

[0163] RNA인 폴리뉴클레오티드가 언급되고 tracr 메이트 서열과 같은 특징을 '포함하는' 것으로 언급되는 경우, RNA 서열은 해당 특징을 포함하는 것으로 인식될 것이다. 폴리뉴클레오티드가 DNA이고 tracr 메이트 서열과 같은 특징을 포함하는 것으로 언급되는 경우, DNA 서열은 당해 특징을 포함하는 RNA이거나 또는 당해 특징을 포함하는 RNA로 전사될 수 있다. 특징이 CRISPR 효소와 같은 단백질인 경우, 언급된 DNA 또는 RNA 서열은 번역되거나 또는 (처음 전사된 DNA의 경우에) 번역될 수 있다.

[0164] 따라서, 특정 구현예에서 본 발명은 조성물을 작동 가능하게 암호화하는 하나 이상의 바이러스 또는 플라스미드 벡터를 포함하는 바이러스 또는 플라스미드 벡터 시스템을 포함하는 비자연적으로 생기는 또는 유전자 조작된 조성물을 이의 발현을 위해 전달하는 단계를 포함하는 관심 대상의 게놈 유전자좌에서 표적 서열의 조작에 의해 유기체(예를 들어 유기체의 유사분열 후 세포를 변형시킴으로써), 예를 들어 인간 또는 비인간 포유류 또는 유기체를 포함하는 포유류를 변형시키는 방법을 제공하되, 상기 조성물은 (A) I. CRISPR-Cas 시스템 키메라 RNA(chiRNA) 폴리뉴클레오티드 서열에 대해 작동 가능하게 연결되는 제1 조절 구성요소로서, 폴리뉴클레오티드 서열은 (a) 진핵 세포 내 표적 서열에 혼성화할 수 있는 가이드 서열, (b) tracr 메이트 서열, 및 (c) tracr 서열을 포함하는, 제1 조절 구성요소, 및 II. 적어도 하나 이상의 핵 국소화 서열(또는 선택적으로 일부 구현예가 NLS를 포함할 수 없기 때문에 적어도 하나 이상의 핵 국소화 서열)을 포함하는 CRISPR 효소를 암호화하는 효소 암호화 서열에 작동 가능하게 연결된 제2 조절 구성요소를 포함하는 하나 이상의 벡터를 포함하는 벡터 시스템을 포함하는 비자연적으로 생기는 또는 유전자 조작된 조성물(여기서, (a), (b) 및 (c)는 5'에서 3' 방향으로 배열되고, 구성성분 I 및 II는 시스템의 동일 또는 상이한 벡터 상에 위치되며, 전사될 때, tracr 메이트 서열은 tracr 서열에 혼성화하고, 가이드 서열은 표적 서열에 대한 CRISPR 복합체의 서열-특이적 결합을 지시하며, CRISPR 복합체는 (1) 표적 서열에 혼성화된 가이드 서열, 및 (2) tracr 서열에 혼성화된 tracr 메이트 서열과 복합체화된 CRISPR 효소를 포함함, 또는 (B) I. (a) 진핵세포 내 표적 서열에 혼성화할 수 있는 가이드 서열, 및 (b) 적어도 하나 이상의 tracr 메이트 서열에 대해, 작동 가능하게 연결된 제1 조절 구성요소, II. CRISPR 효소를 암호화하는 효소 암호화 서열에 작동 가능하게 연결된 제2 조절 구성요소, 및 III. tracr 서열에 작동 가능하게 연결된 제3 조절 구성요소를 포함하는 하나 이상의 벡터를 포함하는 벡터 시스템을 포함하는 비자연적으로 생기는 또는 유전자 조작된 조성물(여기서, 구성성분 I, II 및 III는 시스템의 동일 또는 상이한 벡터 상에 위치되고, 전사될 때, tracr 메이트 서열은 tracr 서열에 혼성화하고, 가이드 서열은 표적 서열에 대한 CRISPR 복합체의 서열-특이적 결합을 지시하며, CRISPR 복합체는 (1) 표적 서열에 혼성화된 가이드 서열, 및 (2) tracr 서열에 혼성화된 tracr 메이트 서열과 복합체화된 CRISPR 효소를 포함함)을 포함한다. 일부 구현예에서, 구성성분 I, II 및 III은 동일한 벡터 상에 위치된다. 다른 구현예에서, 구성성분 I 및 II는 동일한 벡터 상에 위치되지만, 구성성분 III은 다른 벡터 상에 위치된다. 다른 구현예에서, 구성성분 I 및 III은 동일한 벡터 상에 위치되지만, 구성성분 II는 다른 벡터 상에 위치된다. 다른 구현예에서, 구성성분 II 및 III은 동일한 벡터 상에 위치되지만, 구성성분 I은 다른 벡터 상에 위치된다. 다른 구현예에서, 각각의 구성성분 I, II 및 III은 상이한 벡터 상에 위치된다. 본 발명은 또한 본 명세서에 기재된 바와 같은 바이러스 또는 플라스미드 벡터 시스템을 제공한다.

[0165] 바람직하게는, 벡터는 바이러스 벡터, 예컨대 렌티- 또는 바큘로- 또는 바람직하게는 아데노-바이러스/아데노-

연관 바이러스 벡터이지만, 다른 전달 수단(예컨대, 효모 시스템, 미세소포, 유전자총/금 나노입자에 벡터를 부착하는 수단)이 공지되며, 제공된다. 일부 구현예에서, 바이러스 또는 플라스미드 벡터 중 하나 이상은 리포좀, 나노입자, 엑소좀, 미세소포, 또는 유전자총을 통해 전달될 수 있다.

[0166] 표적 서열의 조작에 의해, 출원인은 또한 표적 서열의 후성적 조작을 의미한다. 이는 예컨대 표적 서열에 대한 접근성을 증가 또는 감소시키는, 표적 서열의 메틸화 상태의 변형(즉, 메틸화 또는 메틸화 패턴 또는 CpG 섬의 추가 또는 제거), 히스톤 변형에 의해, 또는 3D 폴딩을 촉진시키는 것에 의해 표적 서열의 염색질 상태를 가질 수 있다.

[0167] 관심대상의 게놈 유전자좌에서 표적 서열의 조작에 의해 인간 또는 비인간 포유류 또는 유기체를 포함하는 유기체 또는 포유류를 변형시키는 방법이 언급되는 경우, 이는 유기체로부터 전체 또는 단지 단일 세포 또는 세포의 집단으로서 유기체(또는 포유류)에 적용될 수 있다(유기체가 다세포라면)는 것이 인식될 것이다. 인간의 경우에, 예를 들어, 출원인은 특히 단일 세포 또는 세포의 집단을 생각하며, 이들은 바람직하게는 생체외에서 변형된 다음 재도입될 수 있다. 이 경우에, 생검 또는 다른 조직 또는 생물학적 유체 샘플이 필요할 수 있다. 이와 관련하여 줄기 세포가 또한 특히 바람직하다. 그러나, 물론 생체내 구현예가 또한 생각된다.

[0168] 특정 구현예에서 본 발명은 표적 서열의 조작에 의해 대상체 또는 비인간 대상체를 변형시키는 단계를 포함하는 치료가 필요한 대상체(예를 들어, 포유류 또는 인간) 또는 비인간 대상체(예를 들어, 포유류)에서 관심대상의 게놈 유전자좌 내 표적 서열에서 결함에 의해 야기되는 질환을 치료 또는 저해하는 방법을 제공하되, 질환은 조성물을 작동 가능하게 암호화하는 하나 이상의 AAV 또는 렌티바이러스 벡터를 포함하는 AAV 또는 렌티바이러스 벡터 시스템을 포함하는 비자연적으로 생기는 또는 유전자 조작된 조성물을 이의 발현을 위해 전달하는 단계를 포함하는 치료를 제공하는 것을 포함하는 표적 서열의 조작에 의해 치료 또는 저해에 영향을 받기 쉽되, 표적 서열은 발현될 때, 조성물에 의해 조작되고, 조성물은: (A) I. CRISPR-Cas 시스템 키메라 RNA(chiRNA) 폴리뉴클레오티드 서열에 대해 작동 가능하게 연결되는 제1 조절 구성요소로서, 폴리뉴클레오티드 서열은 (a) 진핵 세포 내 표적 서열에 혼성화할 수 있는 가이드 서열, (b) tracr 메이트 서열, 및 (c) tracr 서열을 포함하는, 제1 조절 구성요소, 및 II. 적어도 하나 이상의 핵 국소화 서열(또는 선택적으로 일부 구현예가 NLS를 포함할 수 없기 때문에 적어도 하나 이상의 핵 국소화 서열)을 포함하는 CRISPR 효소를 암호화하는 효소 암호화 서열에 작동 가능하게 연결된 제2 조절 구성요소를 포함하는 하나 이상의 벡터를 포함하는 벡터 시스템을 포함하는 비자연적으로 생기는 또는 유전자 조작된 조성물(여기서, (a), (b) 및 (c)는 5'에서 3' 방향으로 배열되고, 구성성분 I 및 II는 시스템의 동일 또는 상이한 벡터 상에 위치되며, 전사될 때, tracr 메이트 서열은 tracr 서열에 혼성화하고, 가이드 서열은 표적 서열에 대한 CRISPR 복합체의 서열-특이적 결합을 지시하며, CRISPR 복합체는 (1) 표적 서열에 혼성화된 가이드 서열, 및 (2) tracr 서열에 혼성화된 tracr 메이트 서열과 복합체화된 CRISPR 효소를 포함함), 또는 (B) I. (a) 진핵세포 내 표적 서열에 혼성화할 수 있는 가이드 서열, 및 (b) 적어도 하나 이상의 tracr 메이트 서열에 대해, 작동 가능하게 연결된 제1 조절 구성요소, II. CRISPR 효소를 암호화하는 효소 암호화 서열에 작동 가능하게 연결된 제2 조절 구성요소, 및 III. tracr 서열에 작동 가능하게 연결된 제3 조절 구성요소를 포함하는 하나 이상의 벡터를 포함하는 벡터 시스템을 포함하는 비자연적으로 생기는 또는 유전자 조작된 조성물(여기서, 구성성분 I, II 및 III는 시스템의 동일 또는 상이한 벡터 상에 위치되고, 전사될 때, tracr 메이트 서열은 tracr 서열에 혼성화하고, 가이드 서열은 표적 서열에 대한 CRISPR 복합체의 서열-특이적 결합을 지시하며, CRISPR 복합체는 (1) 표적 서열에 혼성화된 가이드 서열, 및 (2) tracr 서열에 혼성화된 tracr 메이트 서열과 복합체화된 CRISPR 효소를 포함함)을 포함한다. 일부 구현예에서, 구성성분 I, II 및 III은 동일한 벡터 상에 위치된다. 다른 구현예에서, 구성성분 I 및 II는 동일한 벡터 상에 위치되지만, 구성성분 III은 다른 벡터 상에 위치된다. 다른 구현예에서, 구성성분 I 및 III은 동일한 벡터 상에 위치되지만, 구성성분 II는 다른 벡터 상에 위치된다. 다른 구현예에서, 구성성분 II 및 III은 동일한 벡터 상에 위치되지만, 구성성분 I은 다른 벡터 상에 위치된다. 다른 구현예에서, 각각의 구성성분 I, II 및 III은 상이한 벡터 상에 위치된다. 본 발명은 또한 본 명세서에 기재된 바와 같은 바이러스(예를 들어, AAV 또는 렌티바이러스) 벡터 시스템을 제공하고, 본 명세서에 기재된 바와 같은 벡터 시스템의 부분일 수 있다.

[0169] 본 발명의 일부 방법은 발현을 유발하는 단계를 포함할 수 있다. 본 발명의 일부 방법에서, 유기체 또는 대상체는 진핵생물(인간을 포함하는 포유류를 포함) 또는 비인간 진핵생물 또는 비인간 동물 또는 비인간 포유류이다. 일부 구현예에서, 유기체 또는 대상체는 비인간 동물이고, 절지동물, 예를 들어, 곤충일 수 있거나, 또는 선충일 수 있다. 본 발명의 일부 방법에서 유기체 또는 대상체는 식물이다. 본 발명의 일부 방법에서 유기체 또는 대상체는 포유류 또는 비인간 포유류이다. 비인간 포유류는, 예를 들어 설치류(바람직하게는 마우스 또는 래트), 유제류, 또는 영장류일 수 있다. 본 발명의 일부 방법에서 유기체 또는 대상체는 미세조류를 포함한 조류

이거나, 균류이다. 본 발명의 일부 방법에서 바이러스 벡터는 AAV 또는 렌티바이러스이고, 본 명세서에 기재된 바와 같은 벡터 시스템의 부분일 수 있다. 본 발명의 일부 방법에서, CRISPR 효소는 Cas9이다. 본 발명의 일부 방법에서, 가이드 서열의 발현은 T7 프로모터의 제어 하에 있고, T7 중합효소의 발현에 의해 구동된다.

[0170] 본 발명은 일부 구현예에서 CRISPR 효소를 암호화하는 mRNA를 세포에 전달하는 단계를 포함하는 CRISPR 효소를 전달하는 방법을 포함한다. 일부의 이들 방법에서, CRISPR 효소는 Cas9이다.

[0171] 본 발명은 또한 본 발명의 벡터 세스템, 특히 본 명세서에 기재된 바와 같은 바이러스 벡터 시스템을 제조하는 방법을 제공한다. 본 발명은 일부 구현예에서 AAV에 대한 핵산 분자(들)의 암호를 함유하거나 또는 이로 본질적으로 이루어지는 플라스미드(들)를 AAV-감염 세포에 트렌스펙션시키는 단계, 및 AAV의 복제 및 패키징에 필수적인 AAV rep 및/또는 cap을 공급하는 단계를 포함하는 AAV 바이러스 벡터를 제조하는 방법을 포함한다. 일부 구현예에서 AAV의 복제 및 패키징에 필수적인 AAV rep 및/또는 cap은 헬퍼 플라스미드(들) 또는 헬퍼 바이러스(들)를 이용하여 세포를 트렌스펙션시킴으로써 공급된다. 일부 구현예에서 헬퍼 바이러스는 폭스바이러스, 아데노바이러스, 헤르페스바이러스 또는 바큘로바이러스이다. 일부 구현예에서 폭스바이러스는 백시니아 바이러스이다. 일부 구현예에서 세포는 포유류 세포이다. 일부 구현예에서, 세포는 곤충 세포이고, 헬퍼 바이러스는 바큘로바이러스이다. 다른 구현예에서, 바이러스는 렌티바이러스이다.

[0172] 식물에서, 병원균은 종종 숙주-특이적이다. 예를 들어, 푸사리움 옥시스포룸 f. 종 리코페르시시(*Fusarium oxysporum* f. sp. *lycopersici*)는 토마토 시들음을 야기하지만, 도마토만을 공격하며, F. 옥시스포룸 f. 디안티이 푸시니아 그라미니스 f. 종 트리티시(*F. oxysporum f. dianthii Puccinia graminis* f. sp. *tritici*)는 밀만을 공격한다. 식물은 대부분의 병원균에 저항하는 기존의 유도 방어를 가진다. 특히 병원균은 식물보다 더 빈번하게 번식하므로, 식물 세대를 통한 돌연변이 및 재조합 사건은 감수성을 발생시키는 유전적 변이성을 초래한다. 식물에서, 비숙주 저항성이 있을 수 있으며, 예를 들어 숙주 및 병원균은 양립할 수 없다. 또한, 통상적으로 다수의 유전자에 의해 제어되는 수평 저항성, 예를 들어 병원균의 모든 품종에 대한 부분적인 저항성, 및 통상적으로 소수의 유전자에 의해 제어되는 수직 저항성, 예를 들어 병원균의 일부 품종에 대해서는 완벽한 저항성을 갖지만, 다른 품종에 대해서는 그렇지 않은 저항성이 있을 수 있다. 유전자에 대한 유전자(Gene-for-Gene) 수준에서, 식물과 병원균은 함께 진화하고, 하나의 균형 상태에서의 유전자 변화는 다른 것을 변화시킨다. 따라서, 자연적 변동성을 사용하여, 재배자는 산출량, 품질, 균일성, 내한성, 저항성을 위하여 가장 유용한 유전자를 조합한다. 저항성 유전자의 공급원은 천연 또는 외래 품종, 세습 품종, 야생 식물 동족, 및 유도된 돌연변이체, 예를 들어 돌연변이 유발제로 식물 물질을 처리한 것을 포함한다. 본 발명을 사용하여, 식물 재배자에게 돌연변이를 유도하기 위한 새로운 도구가 제공된다. 따라서, 당업자는 저항성 유전자의 공급원의 게놈을 분석할 수 있으며, 원하는 특징 또는 특성을 가지는 품종에서 이전 돌연변이 유발제보다 더 정밀하게 저항성 유전자의 발생을 유도하기 위해 본 발명을 이용할 수 있고, 따라서 식물 육종 프로그램을 가속화하고 개선시킬 수 있다.

[0173] 본 발명은 추가로 의학에서 또는 요법에서 사용하기 위한 본 발명의 조성물 또는 이의 CRISPR 효소(CRISPR 효소를 암호화하는 mRNA를 포함하거나 또는 대안적으로 포함함)를 포함한다. 일부 구현예에서 본 발명은 본 발명에 따른 조성물 또는 이의 본 발명에 따른 방법에서 사용하기 위한 CRISPR 효소(CRISPR 효소를 암호화하는 mRNA를 포함하거나 또는 대안적으로 포함함)를 포함한다. 일부 구현예에서 본 발명은 생체외 유전자 또는 게놈 편집에서 본 발명의 조성물 또는 이의 CRISPR 효소(CRISPR 효소를 암호화하는 mRNA를 포함하거나 또는 대안적으로 포함함)의 용도를 제공한다. 특정 구현예에서 본 발명은 생체외 유전자 또는 게놈 편집을 위한 의약의 제조에서 본 발명에 따른 방법에서 사용하기 위한 본 발명의 조성물 또는 이의 CRISPR 효소(CRISPR 효소를 암호화하는 mRNA를 포함하거나 또는 대안적으로 포함함)의 용도를 포함한다. 일부 구현예에서 본 발명은 본 발명의 조성물 또는 이의 CRISPR 효소(CRISPR 효소를 암호화하는 mRNA를 포함하거나 또는 대안적으로 포함함)를 포함하되, 표적 서열은 그의 3'에서 Cas9가 스트렙토코커스 피오게네스 또는 스태필로코커스 아우레우스 Cas9인(또는 이들로부터 유래된) 경우, 5'-모티프를 포함하는 PAM(프로토스페이서 인접 모티프) 서열에 측접된다. 예를 들어, 적합한 PAM은 이하에 언급되는 바와 같이 SpCas9 또는 SaCas9 효소(또는 유래 효소)에 대해 각각 5'-NRG 또는 5'-NNGRR(N은 임의의 뉴클레오티드임)이다.

[0174] SpCas9 또는 SaCas9는 S. 피오게네스 또는 S. 아우레우스 Cas9로부터 생기거나 또는 유래된 것임이 인식될 것이다. 물론, 이는 본원에 기재된 바와 같이 의도된 용도를 적합화시키기 위해 야생형으로부터 돌연변이되거나 달리 변화될 수 있다. 동일 염색체의 상이한 가닥 상의 마주하는 부위로 지시된 2개의 중첩 가이드와 특별히 조합된 이중 닉카아제 D10A 돌연변이 또는 변이체가 바람직하다.

[0175] 본 발명의 양태들은 CRISPR 효소, 예컨대 Cas9 매개된 유전자 표적화의 특이성을 개선하는 것과, CRISPR 효소, 예컨대 Cas9에 의한 표적외 변형의 가능성을 감소시키는 것을 포함한다. 본 발명은 일부 구현예에서:

[0176] I. 제1 CRISPR-Cas 시스템 키메라 RNA(chiRNA) 폴리뉴클레오티드로서, 제1 폴리뉴클레오티드 서열은

[0177] (a) 제1 표적 서열에 혼성화할 수 있는 제1 가이드 서열,

[0178] (b) 제1 tracr 메이트 서열, 및

[0179] (c) 제1 tracr 서열을 포함하는, 폴리뉴클레오티드, 및

[0180] II. 제2 CRISPR-Cas 시스템 키메라 RNA(chiRNA) 폴리뉴클레오티드 서열, 제2 폴리뉴클레오티드 서열은

[0181] (a) 제2 표적 서열에 혼성화할 수 있는 제2 가이드 서열,

[0182] (b) 제2 tracr 메이트 서열, 및

[0183] (c) 제2 tracr 서열을 포함할 수 있는, 폴리뉴클레오티드 서열, 및

[0184] III. 적어도 하나 이상의 핵 국소화 서열을 포함하고 하나 이상의 돌연변이를 포함할 수 있하는 CRISPR 효소를 암호화하는 폴리뉴클레오티드 서열

[0185] 을 포함할 수 있는 비자연적으로 생기는 또는 유전자 조작된 조성물을 전달하는 단계를 포함하는, 세포에서 관심 대상의 게놈 유전자좌 내 DNA 듀플렉스의 마주하는 가닥에 대한 제1 및 제2 표적 서열의 조작에 의해 표적외 변형을 최소화함으로써 유기체 또는 비인간 유기체를 변형시키는 방법을 포함하되, (a), (b) 및 (c)는 5'에서 3' 방향으로 배열되고, 전사될 때, 제1 및 제2 tracr 메이트 서열은 각각 제1 및 제2 tracr 서열에 혼성화하며, 제1 및 제2 가이드 서열은 각각 제1 및 제2 표적 서열에 대한 제1 및 제2 CRISPR 복합체의 서열-특이적 결합을 지시하고, 제1 CRISPR 복합체는 (1) 제1 표적 서열에 혼성화되는 제1 가이드 서열 및 (2) 제1 tracr 서열에 혼성화된 제1 tracr 메이트 서열에 복합체화되는 CRISPR 효소를 포함하고, 제2 CRISPR 복합체는 (1) 제2 표적 서열에 혼성화되는 제2 가이드 서열 및 (2) 제2 tracr 서열에 혼성화된 제2 tracr 메이트 서열에 복합체화되는 CRISPR 효소를 포함하고, CRISPR 효소를 암호화하는 폴리뉴클레오티드 서열은 DNA 또는 RNA이며, 제1 가이드 서열은 제1 표적 서열 근처의 DNA 듀플렉스의 한 가닥의 절단을 지시하고, 제2 가이드 서열은 이중 가닥 파손을 유도하는 제2 표적 서열 근처의 다른 가닥의 절단을 지시함으로써 표적외 변형을 최소화하여 유기체 또는 비인간 유기체를 변형시킨다.

[0186] 본 발명의 일부 방법에서, CRISPR 효소를 암호화하는 임의의 또는 모든 폴리뉴클레오티드 서열, 제1 및 제2 가이드 서열, 제1 및 제2 tracr 메이트 서열 또는 제1 및 제2 tracr 서열은 RNA이다. 본 발명의 추가 구현예에서 CRISPR 효소를 암호화하는 서열을 암호화하는 폴리뉴클레오티드, 제1 및 제2 가이드 서열, 제1 및 제2 tracr 메이트 서열 또는 제1 및 제2 tracr 서열은 RNA이며, 리포좀, 나노입자, 엑소좀, 미세소포, 또는 유전자총을 통해 전달된다. 본 발명의 특정 구현예에서, 제1 및 제2 tracr 메이트 서열은 100% 동일성을 공유하고/하거나 제1 및 제2 tracr 서열은 100% 동일성을 공유한다. 일부 구현예에서, 폴리뉴클레오티드는 하나 이상의 벡터를 포함하는 벡터 시스템 내에 포함될 수 있다. 본 발명의 바람직한 구현예에서 CRISPR 효소는 Cas9 효소, 예를 들어 SpCas9이다. 본 발명의 양태에서, CRISPR 효소는 촉매적 도메인 내 하나 이상의 돌연변이를 포함하되, 하나 이상의 돌연변이는 D10A, E762A, H840A, N854A, N863A 및 D986A로 이루어진 군으로부터 선택된다. 고도로 바람직한 구현예에서, CRISPR 효소는 D10A 돌연변이를 가진다. 바람직한 구현예에서, 제1 CRISPR 효소는, 효소가 상보적 가닥 닉카아제이도록 하나 이상의 돌연변이를 가지고, 제2 CRISPR 효소는, 효소가 비상보적 가닥 닉카아제이도록 하나 이상의 돌연변이를 가진다. 대안적으로, 제1 효소는 비상보적 가닥 닉카아제일 수 있고, 제2 효소는 상보적 가닥 닉카아제일 수 있다.

[0187] 본 발명의 바람직한 방법에서, 제1 표적 서열 근처의 DNA 듀플렉스의 한 가닥의 절단을 지시하는 제1 가이드 서열 및 제2 표적 서열 근처의 다른 가닥의 절단을 지시하는 제2 가이드 서열은 5' 돌출부를 생성한다. 본 발명의 구현예에서, 5' 돌출부는 200 개 이하의 염기쌍, 바람직하게는 100 개 이하의 염기쌍, 또는 더 바람직하게는 50 개 이하의 염기쌍이다. 본 발명의 구현예에서, 5' 돌출부는 적어도 26 개 염기쌍, 바람직하게는 적어도 30 개 염기쌍 또는 더 바람직하게는 34 내지 50 개 염기쌍이다. 가장 바람직하게는, 중복은 5 개 내지 -1 개 염기쌍이다.

[0188] 본 발명은 일부 구현예에서

[0189] I. (a) 제1 표적 서열에 혼성화할 수 있는 제1 가이드 서열, 및

[0190] (b) 적어도 하나 이상의 tracr 메이트 서열에, 작동 가능하게 연결된 제1 조절 구성요소 ,

[0191] II. (a) 제2 표적 서열에 혼성화할 수 있는 제2 가이드 서열, 및

[0192] (b) 적어도 하나 이상의 tracr 메이트 서열에, 작동 가능하게 연결된 제2 조절 구성요소,

[0193] III. CRISPR 효소를 암호화하는 효소 암호화 서열에 작동 가능하게 연결된 제3 조절 구성요소, 및

[0194] IV. tracr 서열에 작동 가능하게 연결된 제4 조절 구성요소

[0195] 를 포함하는 하나 이상의 벡터를 포함하는 벡터 시스템을 포함하는 비자연적으로 생기는 또는 유전자 조작된 조성물을 전달하는 단계를 포함하는 세포에서 관심 대상의 게놈 유전자좌 내 DNA 듀플렉스의 마주하는 가닥에 대한 제1 및 제2 표적 서열의 조작에 의해 표적외 변형을 최소화함으로써 유기체 또는 비인간 유기체를 변형시키는 방법을 포함하되, 구성성분 I, II, III 및 IV는 시스템의 동일 또는 상이한 벡터 상에 위치되고, 전사될 때, tracr 메이트 서열은 tracr 서열에 혼성화하며, 제1 및 제2 가이드 서열은 제1 및 제2 표적 서열에 대해 각각 제1 및 제2 CRISPR 복합체의 서열 특이적 결합을 지시하고, 제1 CRISPR 복합체는 (1) 제1 표적 서열에 혼성화되는 제1 가이드 서열 및 (2) 제1 tracr 서열에 혼성화된 제1 tracr 메이트 서열에 복합체화되는 CRISPR 효소를 포함하고, 제2 CRISPR 복합체는 (1) 제2 표적 서열에 혼성화되는 제2 가이드 서열 및 (2) tracr 서열에 혼성화된 tracr 메이트 서열에 복합체화되는 CRISPR 효소를 포함하고, CRISPR 효소를 암호화하는 폴리뉴클레오티드 서열은 DNA 또는 RNA이며, 제1 가이드 서열은 제1 표적 서열 근처의 DNA 듀플렉스의 한 가닥의 절단을 지시하고, 제2 가이드 서열은 이중 가닥 파손을 유도하는 제2 표적 서열 근처의 다른 가닥의 절단을 지시함으로써 표적외 변형을 최소화하여 유기체 또는 비인간 유기체를 변형시킨다.

[0196] 본 발명은 또한 본 명세서에 기재된 바와 같은 벡터 시스템을 제공한다. 시스템은 1, 2, 3 또는 4 개의 상이한 벡터를 포함할 수 있다. 따라서 구성성분 I, II, III 및 IV는 1, 2, 3 또는 4 개의 상이한 벡터 상에 위치될 수 있고, 구성성분의 가능한 위치에 대한 모든 조합이 본 명세서에서 생각되고, 예를 들어: 구성성분 I, II, III 및 IV는 동일 벡터 상에 위치될 수 있으며; 구성성분 I, II, III 및 IV는 각각 상이한 벡터 상에 위치될 수 있고; 구성성분 I, II, II I 및 IV는 총 2 또는 3 개의 상이한 벡터 상에 위치될 수 있으며, 위치의 모든 조합 등이 생각된다.

[0197] 본 발명의 일부 방법에서, CRISPR 효소를 암호화하는 임의의 또는 모든 폴리뉴클레오티드 서열, 제1 및 제2 가이드 서열, 제1 및 제2 tracr 메이트 서열 또는 제1 및 제2 tracr 서열은 RNA이다. 본 발명의 추가 구현예는 제1 및 제2 tracr 메이트 서열이 100% 동일성을 공유하고/공유하거나 제1 및 제2 tracr 서열은 100% 동일성을 공유한다. 본 발명의 바람직한 구현예에서, CRISPR 효소는 Cas9 효소, 예를 들어 SpCas9이다. 본 발명의 양태에서, CRISPR 효소는 촉매적 도메인 내 하나 이상의 돌연변이를 포함하되, 하나 이상의 돌연변이는 D10A, E762A, H840A, N854A, N863A 및 D986A로 이루어진 군으로부터 선택된다. 고도로 바람직한 구현예에서, CRISPR 효소는 D10A 돌연변이를 가진다. 바람직한 구현예에서, 제1 CRISPR 효소는, 효소가 상보적 가닥 닉카아제이도록 하나 이상의 돌연변이를 가지며, 제2 CRISPR 효소는, 효소가 비상보적 가닥 닉카아제이도록 하나 이상의 돌연변이를 가진다. 대안적으로, 제1 효소는 비상보적 닉카아제일 수 있고, 제2 효소는 상보적 가닥 닉카아제일 수 있다. 본 발명의 추가 구현예에서, 바이러스 벡터 중 하나 이상은 리포좀, 나노입자, 엑소좀, 미세소포, 또는 유전자총을 통해 전달된다.

[0198] 본 발명의 바람직한 방법에서, 제1 표적 서열 근처의 DNA 듀플렉스의 한 가닥의 절단을 지시하는 제1 가이드 서열 및 제2 표적 서열 근처의 다른 가닥의 절단을 지시하는 제2 가이드 서열은 5' 돌출부를 생성한다. 본 발명의 구현예에서, 5' 돌출부는 200 개 이하의 염기쌍, 바람직하게는 100 개 이하의 염기쌍, 또는 더 바람직하게는 50 개 이하의 염기쌍이다. 본 발명의 구현예에서, 5' 돌출부는 적어도 26 개 염기쌍, 바람직하게는 적어도 30 개 염기쌍 또는 더 바람직하게는 34 내지 50 개 염기쌍이다.

[0199] 본 발명은 일부 구현예에서, 하나 이상의 돌연변이를 갖는 Cas 단백질 및 RNA 분자의 제1 및 제2 가닥을 각각 표적화하는 2 개의 가이드 RNA를 포함하는 유전자 조작된, 비자연적으로 생기는 CRISPR-Cas 시스템을 관심 대상의 유전자 산물을 암호화하는 이중 가닥 DNA 분자를 함유 및 발현시키는 세포 내로 도입함으로써 표적외 변형을 최소화하고, 이에 의해 가이드 RNA는 유전자 산물을 암호화하는 DNA 분자를 표적화하고, Cas 단백질은 유전자 산물을 암호화하는 DNA 분자의 각각의 제1 가닥 및 제2 가닥의 틈을 냄으로써, 유전자 산물의 발현이 변용되되; Cas 단백질 및 2 개의 가이드 RNA는 자연적으로 함께 생기지 않는, 관심 대상의 게놈 유전자좌를 변형시키는 방

법을 포함한다.

[0200] 본 발명의 바람직한 방법에서, 유전자 산물을 암호화하는 DNA 분자의 각각의 제1 가닥 및 제2 가닥의 틈을 내는 Cas 단백질은 5' 돌출부를 생성한다. 본 발명의 구현예에서, 5' 돌출부는 200 개 이하의 염기쌍, 바람직하게는 100 개 이하의 염기쌍, 또는 더 바람직하게는 50 개 이하의 염기쌍이다. 본 발명의 구현예에서, 5' 돌출부는 적어도 26 개 염기쌍, 바람직하게는 적어도 30 개 염기쌍 또는 더 바람직하게는 34 내지 50 개 염기쌍이다.

[0201] 본 발명의 구현예는 또한 tracr 메이트 서열 및 tracr 서열에 융합된 가이드 서열을 포함하는 가이드 RNA를 포함한다. 본 발명의 양태에서 Cas 단백질은 진핵 세포, 바람직하게는 포유류 세포 또는 인간 세포에서 발현을 위해 코돈 최적화된다. 본 발명의 추가 구현예에서 Cas 단백질은 II형 CRISPR-Cas 단백질, 예를 들어 Cas 9 단백질이다. 고도로 바람직한 구현예에서, Cas 단백질은 Cas9 단백질, 예를 들어 SpCas9이다. 본 발명의 양태에서, Cas 단백질은 D10A, E762A, H840A, N854A, N863A 및 D986A로 이루어진 군으로부터 선택된 하나 이상의 돌연변이를 가진다. 고도로 바람직한 구현예에서, Cas 단백질은 D10A 돌연변이를 가진다.

[0202] 본 발명의 양태는 감소될 유전자 산물의 발현 또는 유전자 산물을 암호화하는 DNA 분자 내로 추가로 도입될 주형 폴리뉴클레오티드 또는 2 개의 5' 돌출부를 다시 아닐링 및 결찰시킴으로써 정확하게 절단되는 개재 서열 또는 변용되는 유전자 산물의 활성 또는 기능 또는 증가될 유전자 산물의 발현에 관한 것이다. 본 발명의 구현예에서, 유전자 산물은 단백질이다.

[0203] 본 발명은 또한 세포 내 유전자 산물을 암호화하는 이중 가닥 DNA 분자의 제1 가닥 및 제2 가닥을 각각 표적화하는 2 개의 가이드 RNA 및 하나 이상의 돌연변이를 갖는 Cas 단백질을 포함하고, 이에 의해 가이드 RNA 는 유전자 산물을 암호화하는 DNA 분자를 표적화하고 Cas 단백질은 유전자 산물을 암호화하는 DNA 분자의 각각의 제1 가닥 및 제2 가닥의 틈을 만들어서 유전자 산물의 발현이 변용되되; Cas 단백질 및 2 개의 가이드 RNA는 함께 자연적으로 생기지 않는, 유전자 조작된, 비자연적으로 생기는 CRISPR-Cas 시스템을 포함한다.

[0204] 본 발명의 양태에서, 가이드 RNA는 tracr 메이트 서열 및 tracr 서열에 융합된 가이드 서열을 포함할 수 있다. 본 발명의 구현예에서, Cas 단백질은 II형 CRISPR-Cas 단백질이다. 본 발명의 양태에서, Cas 단백질은 진핵 세포, 바람직하게는 포유류 세포 또는 인간 세포에서 발현을 위해 코돈 최적화된다. 본 발명의 추가 구현예에서 Cas 단백질은 II형 CRISPR-Cas 단백질, 예를 들어 Cas 9 단백질이다. 고도로 바람직한 구현예에서, Cas 단백질은 Cas9 단백질, 예를 들어 SpCas9이다. 본 발명의 양태에서, Cas 단백질은 D10A, E762A, H840A, N854A, N863A 및 D986A로 이루어진 군으로부터 선택된 하나 이상의 돌연변이를 가진다. 고도로 바람직한 구현예에서, Cas 단백질은 D10A 돌연변이를 가진다.

[0205] 본 발명의 양태는 감소될 유전자 산물의 발현 또는 유전자 산물을 암호화하는 DNA 분자 내로 추가 도입될 주형 폴리뉴클레오티드 또는 2 개의 5' 돌출부를 다시 아닐링 및 결찰시킴으로써 정확하게 절단되는 개재 서열 또는 변용되는 유전자 산물의 활성 또는 기능 또는 증가될 유전자 산물의 발현에 관한 것이다. 본 발명의 구현예에서, 유전자 산물은 단백질이다.

[0206] 본 발명은 또한:

[0207] a) 유전자 산물을 암호화하는 이중 가닥 DNA 분자의 제1 가닥 및 제2 가닥을 각각 표적화하는 각각의 2 개 CRISPR-Cas 시스템 가이드 RNA에 작동 가능하게 연결된 제1 조절 구성요소,

[0208] b) Cas 단백질에 작동 가능하게 연결된 제2 조절 구성요소를 포함하는 유전자 조작된, 비자연적으로 생기는 하나 이상의 벡터를 포함하는 벡터 시스템을 포함하되,

[0209] 구성성분 (a) 및 (b)는 시스템의 동일 또는 상이한 벡터 상에 위치되며, 이에 의해 가이드 RNA는 유전자 산물을 암호화하는 DNA 분자를 표적화하고, Cas 단백질은 유전자 산물을 암호화하는 DNA 분자의 각각의 제1 가닥 및 제2 가닥의 틈을 내어서, 유전사 산물의 발현은 변형되되; Cas 단백질 및 2 개의 가이드 RNA는 자연적으로 함께 생기지 않는다.

[0210] 본 발명의 양태에서 가이드 RNA는 tracr 메이트 서열 및 tracr 서열에 융합된 가이드 서열을 포함할 수 있다. 본 발명의 구현예에서, Cas 단백질은 II형 CRISPR-Cas 단백질이다. 본 발명의 양태에서, Cas 단백질은 진핵 세포, 바람직하게는 포유류 세포 또는 인간 세포에서 발현을 위해 코돈 최적화된다. 본 발명의 추가 구현예에서, Cas 단백질은 II형 CRISPR-Cas 단백질, 예를 들어 Cas 9 단백질이다. 고도로 바람직한 구현예에서, Cas 단백질은 Cas9 단백질, 예를 들어 SpCas9이다. 본 발명의 양태에서, Cas 단백질은 D10A, E762A, H840A, N854A, N863A 및 D986A로 이루어진 군으로부터 선택되는 하나 이상의 돌연변이를 가진다. 고도로 바람직한 구현예에서, Cas

단백질은 D10A 돌연변이를 가진다.

[0211] 본 발명의 양태는 감소될 유전자 산물의 발현 또는 유전자 산물을 암호화하는 DNA 분자 내로 추가 도입될 주형 폴리뉴클레오티드 또는 2 개의 5' 돌출부를 다시 아닐링 및 결찰시킴으로써 정확하게 절단되는 개재 서열 또는 변용되는 유전자 산물의 활성 또는 기능 또는 증가될 유전자 산물의 발현에 관한 것이다. 본 발명의 구현예에서, 유전자 산물은 단백질이다. 본 발명의 바람직한 구현예에서, 시스템의 벡터는 바이러스 벡터이다. 추가 구현예에서, 시스템의 벡터는 리포좀, 나노입자, 엑소좀, 미세소포, 또는 유전자총을 통해 전달된다.

[0212] 일 양태에서, 본 발명은 진핵 세포에서의 표적 폴리뉴클레오티드의 변경 방법을 제공한다. 일부 구현예에서, 상기 방법은 CRISPR 복합체가 표적 폴리뉴클레오티드에 결합하게 하여, 상기 표적 폴리뉴클레오티드의 절단을 초래하여, 표적 폴리뉴클레오티드를 변경시키는 단계를 포함하며, 여기서, CRISPR 복합체는 상기 표적 폴리뉴클레오티드 내의 표적 서열에 혼성화되는 가이드 서열과 복합체화된 CRISPR 효소를 포함하며, 상기 가이드 서열은 tracr 메이트 서열에 연결되며, tracr 메이트 서열은 차례로 tracr 서열에 혼성화된다. 일부 구현예에서, 상기 절단은 상기 CRISPR 효소에 의한, 표적 서열의 위치에서의 1 개 또는 2 개의 가닥의 절단을 포함한다. 일부 구현예에서, 상기 절단은 감소된 표적 유전자의 전사를 야기한다. 일부 구현예에서, 상기 방법은 외인성 주형 폴리뉴클레오티드와의 상동성 재조합에 의해 상기 절단된 표적 폴리뉴클레오티드를 수복하는 단계를 더 포함하며, 상기 수복은 상기 표적 폴리뉴클레오티드의 하나 이상의 뉴클레오티드의 삽입, 결실 또는 치환을 포함하는 돌연변이를 야기한다. 일부 구현예에서, 상기 돌연변이는 표적 서열을 포함하는 유전자로부터 발현되는 단백질의 하나 이상의 아미노산 변화를 야기한다. 일부 구현예에서, 상기 방법은 하나 이상의 벡터를 상기 진핵 세포로 전달하는 단계를 더 포함하며, 하나 이상의 벡터는 CRISPR 효소, tracr 메이트 서열에 연결된 가이드 서열 및 tracr 서열 중 하나 이상의 발현을 유도한다. 일부 구현예에서, 상기 벡터는 대상체 내의 진핵 세포로 전달된다. 일부 구현예에서, 상기 변경은 세포 배양물 중의 상기 진핵 세포에서 발생한다. 일부 구현예에서, 상기 방법은 상기 변경 전에 상기 진핵 세포를 대상체로부터 분리하는 단계를 더 포함한다. 일부 구현예에서, 상기 방법은 상기 진핵 세포 및/또는 그로부터 유래된 세포를 상기 대상체로 복귀시키는 단계를 더 포함한다.

[0213] 일 양태에서, 본 발명은 진핵 세포에서의 폴리뉴클레오티드의 발현의 변경 방법을 제공한다. 일부 구현예에서, 상기 방법은 CRISPR 복합체가 폴리뉴클레오티드에 결합하게 하여, 상기 결합이 상기 폴리뉴클레오티드의 증가되거나 감소된 발현을 야기하도록 하는 단계를 포함하며; 여기서, CRISPR 복합체는 상기 폴리뉴클레오티드 내의 표적 서열에 혼성화되는 가이드 서열과 복합체화된 CRISPR 효소를 포함하고, 상기 가이드 서열은 tracr 메이트 서열에 연결되고, tracr 메이트 서열은 차례로 tracr 서열로 혼성화된다. 일부 구현예에서, 상기 방법은 하나 이상의 벡터를 상기 진핵 세포로 전달하는 단계를 더 포함하며, 여기서, 하나 이상의 벡터는 CRISPR 효소, tracr 메이트 서열에 연결된 가이드 서열 및 tracr 서열 중 하나 이상의 발현을 유도한다.

[0214] 일 양태에서, 본 발명은 돌연변이된 질병 유전자를 포함하는 모델 진핵 세포의 생성 방법을 제공한다. 일부 구현예에서, 질병 유전자는 질병을 갖거나 질병이 발생할 위험의 증가와 관련된 임의의 유전자이다. 일부 구현예에서, 상기 방법은 (a) 하나 이상의 벡터를 진핵 세포로 도입하는 단계로서, 하나 이상의 벡터는 CRISPR 효소, tracr 메이트 서열에 연결된 가이드 서열 및 tracr 서열 중 하나 이상의 발현을 유도하는 단계; 및 (b) CRISPR 복합체가 표적 폴리뉴클레오티드에 결합하게 하여, 상기 질병 유전자 내의 표적 폴리뉴클레오티드의 절단을 야기하여, 돌연변이된 질병 유전자를 포함하는 모델 진핵 세포를 생성하는 단계로서, CRISPR 복합체가 (1) 표적 폴리뉴클레오티드 내의 표적 서열에 혼성화되는 가이드 서열, 및 (2) tracr 서열에 혼성화되는 tracr 메이트 서열과 복합체화되는 CRISPR 효소를 포함하는 단계를 포함한다. 일부 구현예에서, 상기 절단은 상기 CRISPR 효소에 의한, 표적 서열의 위치에서의 1 개 또는 2 개의 가닥의 절단을 포함한다. 일부 구현예에서, 상기 절단은 표적 유전자의 감소된 전사를 야기한다. 일부 구현예에서, 상기 방법은 외인성 주형 폴리뉴클레오티드와의 상동성 재조합에 의해 상기 절단된 표적 폴리뉴클레오티드를 수복하는 단계를 더 포함하며, 상기 수복은 상기 표적 폴리뉴클레오티드의 하나 이상의 뉴클레오티드의 삽입, 결실 또는 치환을 포함하는 돌연변이를 야기한다. 일부 구현예에서, 상기 돌연변이는 표적 서열을 포함하는 유전자로부터의 단백질 발현의 하나 이상의 아미노산 변화를 야기한다.

[0215] 일 양태에서, 본 발명은 하나 이상의 돌연변이를 하나 이상의 원핵 세포(들) 내의 유전자에 도입함에 의한 하나 이상의 원핵 세포(들)의 선택 방법을 제공하며, 상기 방법은 하나 이상의 벡터를 원핵 세포(들)로 도입하는 단계로서, 하나 이상의 벡터가 CRISPR 효소, tracr 메이트 서열에 연결된 가이드 서열, tracr 서열 및 교정 주형 중 하나 이상의 발현을 유도하고; 교정 주형이 CRISPR 효소 절단을 없애는 하나 이상의 돌연변이를 포함하는 단계; 선택될 세포(들)에서 교정 주형과 표적 폴리뉴클레오티드의 상동성 재조합을 가능하게 하는 단계; CRISPR 복합체가 표적 폴리뉴클레오티드에 결합되게 하여, 상기 유전자 내의 표적 폴리뉴클레오티드의 절단을 초래하는

단계로서, CRISPR 복합체는 (1) 표적 폴리뉴클레오티드 내의 표적 서열에 혼성화되는 가이드 서열 및 (2) tracr 서열에 혼성화되는 tracr 메이트 서열과 복합체화되는 CRISPR 효소를 포함하고, 표적 폴리뉴클레오티드로의 CRISPR 복합체의 결합이 세포사를 유도하여, 하나 이상의 돌연변이가 도입된 하나 이상의 원핵 세포(들)가 선택되게 하는 단계를 포함한다. 바람직한 구현예에서, CRISPR 효소는 Cas9이다. 본 발명의 다른 양태에서, 선택될 세포는 진핵 세포, 예컨대 유사분열 후 진핵 세포일 수 있다. 본 발명의 양태는 선택 마커 또는 반대-선택 시스템을 포함할 수 있는 2-단계 과정을 필요로 하지 않고 특정 세포의 선택을 가능하게 한다.

[0216] 일 양태에서, 본 발명은 진핵 세포 내 표적 폴리뉴클레오티드를 변형시키는 방법을 제공한다. 일부 구현예에서, 상기 방법은 표적 폴리뉴클레오티드에 CRISPR 복합체를 결합시켜 상기 표적 폴리뉴클레오티드의 절단을 달성함으로써 표적 폴리뉴클레오티드를 변형시키는 단계를 포함하되, CRISPR 복합체는 상기 표적 폴리뉴클레오티드 내에서 표적 서열에 혼성화된 가이드 서열과 복합체화된 CRISPR 효소를 포함하고, 상기 가이드 서열은 tracr 메이트 서열에 연결되어, 이는 결국 tracr 서열과 혼성화한다.

[0217] 다른 구현예에서, 본 발명은 진핵 세포 내 폴리뉴클레오티드의 발현을 변형시키는 방법을 제공한다. 방법은 폴리뉴클레오티드에 결합하는 CRISPR 복합체를 사용함으로써 표적 폴리뉴클레오티드의 발현을 증가 또는 감소시키는 단계를 포함한다.

[0218] 원한다면, 세포 내 발현의 변형을 달성하기 위해, tracr 서열, tracr 메이트 서열에 연결된 가이드 서열, CRISPR 효소를 암호화하는 서열을 포함하는 하나 이상의 벡터는 세포에 전달된다. 일부 방법에서, 하나 이상의 벡터는 핵 국소화 서열을 포함하는 상기 CRISPR 효소를 암호화하는 효소 암호화 서열에 작동 가능하게 연결된 조절 구성요소; 및 tracr 메이트 서열에 작동 가능하게 연결된 조절 구성요소 및 tracr 메이트 서열 상류에 가이드 서열을 삽입하기 위한 하나 이상의 삽입 부위를 포함한다. 발현될 때, 가이드 서열은 세포 내 표적 서열에 대한 CRISPR 복합체의 서열-특이적 결합을 지시한다. 통상적으로, CRISPR 복합체는 (1) 표적 서열에 혼성화된 가이드 서열, 및 (2) tracr 서열에 혼성화된 tracr 메이트 서열과 복합체화된 CRISPR 효소를 포함한다.

[0219] 일부 방법에서, 표적 폴리뉴클레오티드는 불활성화되어 세포 내 발현의 변형을 달성할 수 있다. 예를 들어, 세포 내 표적 서열에 대한 CRISPR 복합체의 결합 시, 표적 폴리뉴클레오티드는 서열이 전사되지 않고, 암호 단백질이 생성되지 않도록, 또는 서열이 야생형 서열과 같이 작용하지 않도록 불활성화된다. 예를 들어, 단백질 또는 마이크로RNA 암호화 서열은 단백질이 생성되지 않도록 불활성화될 수 있다.

[0220] 특정 구현예에서, CRISPR 효소는 D10A, E762A, H840A, N854A, N863A 또는 D986A로 이루어진 군으로부터 선택되는 하나 이상의 돌연변이를 포함하고/포함하거나 하나 이상의 돌연변이는 CRISPR 효소의 RuvC1 또는 HNH 도메인에 있거나, 또는 다르게는 본 명세서에 논의되는 바와 같은 돌연변이이다. 일부 구현예에서, CRISPR 효소는 촉매적 도메인에서 하나 이상의 돌연변이를 가지되, 전사될 때, tracr 메이트 서열은 tracr 서열에 혼성화하고, 가이드 서열은 표적 서열의 CRISPR 복합체의 서열-특이적 결합을 지시하며, 효소는 기능성 도메인을 추가로 포함한다. 일부 구현예에서, 기능성 도메인은 전사 활성화 도메인, 바람직하게는 VP64이다. 일부 구현예에서, 기능성 도메인은 전사 억제 도메인, 바람직하게는 KRAB이다. 일부 구현예에서, 전사 억제 도메인은 SID, 또는 SID의 연쇄체(예를 들어, SID4X)이다. 일부 구현예에서, 후성적 변형 효소가 제공되도록 기능성 도메인은 후성적 변형 도메인이다. 일부 구현예에서, 기능성 도메인은 P65 활성화 도메인일 수 있는 활성화 도메인이다.

[0221] 일부 구현예에서, CRISPR 효소는 I형 또는 III형 CRISPR 효소이지만, 바람직하게는 II형 CRISPR 효소이다. 이II형 CRISPR 효소는 임의의 Cas 효소일 수 있다. Cas 효소는, II형 CRISPR 시스템으로부터의 다중 뉴클레아제 도메인을 지니는 가장 큰 뉴클레아제에 대해 상동성을 공유하는 일반적 효소 부류를 지칭할 수 있는 Cas9로서 동정될 수 있다. 가장 바람직하게는, Cas9 효소는 spCas9 또는 saCas9로부터 생기거나 또는 유래된다. 본 발명자들은 유래는, 유래된 효소가 야생형 효소와 높은 정도의 서열 동일성을 갖는 의미에서 야생형 효소에 크게 기반하지만, 이들은 본 명세서에 기재된 바와 같은 일부 방법에서 돌연변이(변형)되는 것을 의미한다.

[0222] 용어 Cas 및 CRISPR 효소는 달리 명확하지 않다면 본 명세서에서 일반적으로 상호호환적으로 사용될 것으로 인식될 것이다. 상기 언급한 바와 같이, 본 명세서에서 사용된 다수의 잔기 넘버링은 스트렙토코커스 피오게네스 내 II형 CRISPR 유전자좌로부터의 Cas9 효소를 지칭한다. 그러나, 본 발명은 SpCas9, SaCa9, St1Cas9 등과 같은 다른 미생물 종으로부터의 다수의 Cas9를 포함하는 것으로 인식될 것이다.

[0223] 이 예에서 인간에 대해 최적화된(즉, 인간에서의 발현을 위해 최적화됨), 코돈 최적화된 서열의 예가 본 명세서에 제공되며, SaCas9 인간 코돈 최적화된 서열을 참조한다. 이것이 바람직하지만, 다른 예가 가능하고 숙주종에 대한 코돈 최적화가 공지되어 있다는 것이 인식될 것이다.

[0224] 바람직하게는, 바이러스 벡터, 예컨대 렌티- 또는 바큘로- 또는 바람직하게는 아데노-바이러스/아데노-연관 바이러스 벡터일 수 있는 벡터의 형태로 전달되지만, 다른 전달 수단(예컨대, 효모 시스템, 미세소포, 유전자총/금 나노입자에 벡터를 부착하는 수단)이 공지되고, 제공된다. 벡터는 바이러스 또는 효모 시스템(예를 들어, 관심대상의 핵산이 프로모터에 작동 가능하게 연결되거나 또는 (궁극적으로 가공 RNA를 제공하는 것과 같은 발현에 대해) 프로모터의 제어하에 있을 수 있는 경우)뿐만 아니라 숙주 세포 내로의 핵산의 직접 전달을 의미한다. 본 명세서의 방법에서 벡터는 바이러스 벡터일 수 있고, 이는 유리하게는 AAV이지만, 본 명세서에서 논의된 바와 같은 다른 바이러스 벡터, 예컨대 렌티바이러스가 사용될 수 있다. 예를 들어, 바큘로바이러스는 곤충 세포 내 발현을 위해 사용될 수 있다. 이들 곤충 세포는 결국 본 발명의 전달에 적합한 AAV 또는 렌티바이러스 벡터와 같은 다량의 추가 벡터를 생성하는데 유용할 수 있다. 또한 CRISPR 효소를 암호화하는 mRNA를 세포에 전달하는 단계를 포함하는 본 CRISPR 효소를 전달하는 방법이 생각된다. 특정 구현예에서 CRISPR 효소는 절단되고/되거나 1000 개 미만의 아미노산 또는 4000 개 미만의 아미노산을 포함하고/포함하거나 뉴클레아제 또는 닉카아제이고/효소이거나 코돈 최적화되고/최적화되거나 하나 이상의 돌연변이를 포함하고/포함하거나 키메라 CRISPR 효소 및/또는 본 명세서에 논의되는 다른 선택사항을 포함하는 것이 인식될 것이다. AAV 및 렌티바이러스 벡터가 바람직하다.

[0225] 특정 구현예에서, 표적 서열은 그의 3' 말단에서 CRISPR 효소, 통상적으로 Cas 및 특히 Cas9에 적합한 PAM에 측접하거나 또는 뒤에 온다.

[0226] 예를 들어, 적합한 PAM은 SpCas9 또는 SaCas9 효소(또는 유래 효소) 각각에 대해 5'-NRG 또는 5'-NNGRR이다.

[0227] SpCas9 또는 SaCas9는 스트렙토코커스 피오게네스 또는 스태필로코커스 아우레우스 Cas9로부터 생기거나 또는 유래된 것임이 인식될 것이다.

[0228] 따라서, 임의의 이전에 공지된 산물, 산물의 제조방법 또는 산물의 사용 방법을 본 발명 내에 포함하지 않는 것이 본 발명의 목적이므로 출원인은 권리를 보존하며 이로써 임의의 이전에 공지된 산물, 공정 또는 방법에 대한 권리 포기를 드러낸다. 본 발명은 USPTO(35 U.S.C. §112, 제1항) 또는 EPO(EPC의 83항)의 기재된 설명 및 실시가능요건을 충족하지 않는 본 발명의 임의의 산물, 공정 또는 산물의 제조 또는 산물을 사용하는 방법의 범주 내에 포함하는 것으로 의도되지 않는다는 것이 추가로 언급되므로 출원인은 권리를 보존하고 이로써 임의의 이전에 기재된 산물, 산물의 제조방법 또는 산물을 사용하는 방법에 대한 권리 포기를 드러낸다.

[0229] 본 개시내용 및 특히 청구범위 및/또는 단락에서, "함유한다", "함유된", "함유하는" 등과 같은 용어가 미국 특허법에 귀속되는 의미를 가질 수 있고; 예를 들어, "포함한다", "포함된", "포함하는" 등을 의미할 수 있으며; "본질적으로 이루어지는" 및 "본질적으로 이루어진다"와 같은 용어가 미국 특허법에 귀속되는 의미를 갖고, 예를 들어, 명백하게 열거되지 않는 구성요소를 허용하지만, 선행 기술에서 발견되거나 본 발명의 기본적인 또는 새로운 특징에 영향을 미치는 구성요소를 배제함이 주목된다.

[0230] 상기 및 기타 구현예는 하기 상세한 설명으로부터 개시되거나, 그로부터 명백하고 그에 의해 포함된다.

도면의 간단한 설명

[0231] 본 발명의 신규의 특징은 특히 첨부된 청구범위에 개시되어 있다. 본 발명의 원리가 이용된 예시적인 구현예에 기재되어 있는 하기의 상세한 설명을 참조함으로써 본 발명의 특징 및 장점을 더욱 잘 이해할 것이며, 첨부된 도면은 다음과 같다:

도 1은 CRISPR 시스템의 개략적 모델을 보여준다. 스트렙토코커스 피오게네스 유래의 Cas9 뉴클레아제(황색)는 20-nt 가이드 서열(청색) 및 스캐폴드(적색)로 이루어진 합성 가이드 RNA(sgRNA)에 의해 게놈 DNA에 표적화된다. 가이드 서열은 필수 5'-NGG 프로토스페이서 인접 모티프(protospacer adjacent motif, PAM; 진홍색)의 인접 상류의 DNA 표적(청색)과 염기쌍을 형성하며, Cas9는 PAM의 약 3 bp 상류(적색 삼각형)에서 이중 가닥 파단(DSB)을 매개한다.

도 2a 내지 도 2f는 예시적인 CRISPR 시스템, 가능한 작용 메카니즘, 진핵 세포에서의 발현을 위한 예시적인 적합화, 및 핵 국소화 및 CRISPR 활성을 평가하는 시험의 결과를 보여준다.

도 3a 내지 도 3d는 예시적인 대상체에 대한 SpCas9 특이성의 평가의 결과를 보여준다.

도 4a 내지 도 4g는 예시적인 벡터 시스템 및 진핵 세포에서 상동성 재조합의 유도에서의 그의 사용에 대한 결과를 보여준다.

도 5는 프로토스페이서 서열의 표를 제공하며, 인간 및 마우스 게놈 내의 유전자좌(loci)에 대한 상응하는 PAM이 있는 예시적인 스트렙토코커스 피오게네스 및 스트렙토코커스 써모필러스 CRISPR 시스템에 기초하여 설계된 프로토스페이서 표적에 대한 변형 효율 결과를 요약한 것이다. 세포를 Cas9 및 pre-crRNA/tracrRNA 또는 키메라 RNA 중 어느 하나로 트랜스펙션시키고, 트랜스펙션 후 72 시간에 분석하였다. 삽입-결실(indel) 백분율은 표기된 세포주로부터의 서베이어(Surveyor) 검정 결과에 기초하여 계산된다(모든 프로토스페이서 표적에 대하여 N=3, 오차는 S.E.M.이고, N.D.는 서베이어 검정을 사용하여 검출가능하지 않음을 나타내며, N.T.는 이러한 연구에서 시험하지 않음을 나타낸다).

도 6a 내지 도 6c는 Cas9-매개의 유전자 표적화를 위한 상이한 tracrRNA 전사물의 비교를 보여준다.

도 7은 이중 가닥 파단-유도 마이크로-삽입 및 -결실의 검출을 위한 서베이어 뉴클레아제 검정의 개략도를 보여준다.

도 8a 및 도 8b는 진핵 세포에서 CRISPR 시스템 요소의 발현을 위한 예시적인 비시스트로닉(bicistronic) 발현 벡터를 보여준다.

도 9a 내지 도 9c는 인간 게놈에서 인접 스트렙토코커스 피오게네스 SF370 유전자좌 1 PAM(NGG)(도 9a) 간의 거리 및 스트렙토코커스 써모필러스 LMD9 유전자좌 2 PAM(NNAGAAW)(도 9b) 간의 거리; 및 염색체(Chr)에 의한 각 PAM에 대한 거리(도 9c)의 히스토그램을 보여준다.

도 10a 내지 도 10d는 예시적인 CRISPR 시스템, 진핵 세포에서의 발현을 위한 예시적인 적합화 및 CRISPR 활성을 평가하는 시험의 결과를 보여준다.

도 11a 내지 도 11c는 포유동물 세포에서 게놈 유전자좌의 표적화를 위한 CRISPR 시스템의 예시적인 조작을 보여준다.

도 12a 및 도 12b는 포유동물 세포에서 crRNA 가공의 노던 블롯(Northern blot) 분석의 결과를 보여준다.

도 13a 및 도 13b는 인간 PVALB 및 마우스 Th 유전자좌에서 프로토스페이서의 예시적인 선택을 보여준다.

도 14는 인간 EMX1 유전자좌에서 스트렙토코커스 써모필러스 CRISPR 시스템의 예시적인 프로토스페이서 및 상응하는 PAM 서열 표적을 보여준다.

도 15는 서베이어, RFLP, 게놈 시퀀싱 및 노던 블롯 검정을 위해 사용되는 프라이머 및 프로브에 대한 서열의 표를 제공한다.

도 16a 내지 도 16c는 키메라 RNA를 사용한 CRISPR 시스템의 예시적인 조작 및 진핵 세포에서 시스템 활성에 대한 서베이어 검정의 결과를 보여준다.

도 17a 및 도 17b는 진핵 세포에서 CRISPR 시스템 활성에 대한 서베이어 검정의 결과의 그래프 표현을 보여준다.

도 18은 UCSC 게놈 브라우저(browser)를 사용한 인간 게놈 내의 일부 스트렙토코커스 피오게네스 Cas9 표적 부위의 예시적인 가시화를 보여준다.

도 19a 내지 도 19d는 3 개 그룹의 큰 Cas9(약 1400 개 아미노산) 및 2 개 그룹의 작은 Cas9(약 1100 개 아미노산)를 포함하는 5 개 과의 Cas9를 보여주는 계통 분석의 원형 표기를 보여준다.

도 20a 내지 도 20f는 3 개 그룹의 큰 Cas9(약 1400 개 아미노산) 및 2 개 그룹의 작은 Cas9(약 1100 개 아미노산)를 포함하는 5 개 과의 Cas9를 보여주는 계통 분석의 선형 표기를 보여준다.

도 21a 내지 도 21d는 상동성 재조합을 통한 게놈 교정을 보여준다. (a) RuvC I 촉매 도메인 내에 D10A 돌연변이가 있는 SpCas9 닉카아제(nickase)의 개략도. (b) 수복 주형으로서 센스 또는 안티센스 단일 가닥 올리고뉴클레오티드를 사용하는 인간 EMX1 유전자좌에서의 상동성 재조합(HR)을 나타내는 개략도. 위의 적색 화살표는 sgRNA 절단 부위를 나타내며; 유전자형분석을 위한 PCR 프라이머(표 J 및 K)는 우측 패널에 화살표로 표시되어 있다. (c) HR에 의해 변형된 영역의 서열. d, 야생형(wt) 및 EMX1 표적 1 유전자좌에서의 닉카아제(D10A) SpCas9-매개의 삽입-결실에 대한 서베이어 검정(n=3). 화살표는 예상되는 단편 크기의 위치를 나타낸다.

도 22a 및 도 22b는 SpCas9에 대한 단일 벡터 설계를 보여준다.

도 23은 Cas9 오솔로그의 길이 분포를 나타내는 그래프를 나타낸다.

도 24a 내지 도 24m은 돌연변이 지점이 SpCas9 유전자 내에 위치된 서열을 나타낸다.

도 25a는 조건적 Cas9, Rosa26 표적화 벡터 맵을 나타낸다.

도 25b는 Cas9, Rosa26 표적화 벡터 맵을 나타낸다.

도 26은 구성적 및 조건적 Cas9 작제물 내 개략적인 중요한 구성요소를 나타낸다.

도 27은 전달 및 생체내 마우스 뇌 Cas9 발현 데이터를 나타낸다.

도 28은 세포 내로 Cas9 및 키메라 RNA의 RNA 전달 (A) Neuro-2A 세포 내로 DNA 또는 mRNA로서 GFP의 전달을 나타낸다. (B) RNA로서 Icam2 유전자에 대한 Cas9 및 키메라 RNA의 전달은 시험한 2 개의 스페이서 중 하나에 대한 절단을 초래한다. (C) RNA로서 F7 유전자에 대한 Cas9 및 키메라 RNA의 전달은 시험한 2 개의 스페이서 중 하나에 대한 절단을 초래한다.

도 29는 DNA 이중 가닥 파단(DSB) 수복이 유전자 교정을 촉진하는 방법을 보여준다. 오류-유발(error-prone) 비상동성 말단 연결(NHEJ) 경로에서, DSB의 말단을 내인성 DNA 수복 기구에 의해 가공하고, 함께 재연결하는데, 이는 연접 부위에서 무작위 삽입-결실(indel) 돌연변이를 야기할 수 있다. 유전자의 코딩 영역 내에서 발생한 삽입-결실 돌연변이는 해독틀 이동 및 조기 종결 코돈을 야기하여, 유전자 녹아웃을 유발할 수 있다. 대안적으로, 플라스미드 또는 단일-가닥 올리고데옥시뉴클레오티드(ssODN)의 형태의 수복 주형을 공급하여, 상동성-유도 수복(HDR) 경로를 활용할 수 있으며, 이는 높은 충실도와 정밀한 교정을 가능하게 한다.

도 30a 내지 도 30c는 HEK 및 HUES9 세포에서 HDR에 대해 예상한 결과를 나타낸다. (a) 표적화 플라스미드 또는 상동성 암(arm)을 지니는 ssODN(센스 또는 안티센스) 중 하나를 사용하여 Cas9(적색 삼각형)에 의해 절단된 표적 게놈 유전자좌에서 서열을 편집할 수 있다. HDR의 효율을 분석하기 위해, 출원인은 상동성 영역 밖에서 아닐링된 프라이머에 의해 PCR-증폭된 표적 유전자좌 내로 HindIII 부위(적색 막대)를 도입하였다. HindIII을 이용한 PCR 산물의 분해는 HDR 사건의 발생을 나타낸다. (b) 관심대상의 유전자좌에 대해 센스 또는 안티센스(s 또는 a) 방향 중 하나로 방향화된 ssODN은 Cas9와 조합으로 사용되어 표적 유전자좌에서 효율적인 HDR-매개 편집을 달성할 수 있다. 40 bp, 및 바람직하게는 90 bp의 최소 상동성 영역이 변형측 중 하나에서 권고된다(적색 막대). (c) EMX1 유전자좌 내 HDR에 대한 ssODN 효과의 예는 야생형 Cas9와 Cas9 닉카아제(D10A)를 사용하여 나타낸다. 각각의 ssODN은 2 개의 제한 부위의 12-bp 삽입에 측접하는 90 bp의 상동성 암을 함유한다.

도 31A 내지 도 31C는 낭성 섬유증 델타 F508 돌연변이에 대한 수선 전략을 나타낸다.

도 32a 내지 도 32b는 (a) FXN 인트론 1에서 개략적 GAA 반복부 확장을 나타내고, (b) CRISPR/Cas 시스템을 사용하여 GAA 확장 영역을 절단하도록 채택된 개략적 전략을 나타낸다.

도 33은 Tet1-3 및 Dnmt1, 3a 및 3b 유전자좌의 효율적 SpCas9 매개 표적화를 위한 선별을 나타낸다. 트렌스펙션된 N2A 세포로부터의 DNA에 대한 서베이어 분석은 상이한 gRNA를 사용함으로써 효율적인 DNA 절단을 입증한다.

도 34는 AAV1/2 전달 시스템에서 2-벡터 시스템을 사용하여 표적화하는 다중복합체 게놈의 전략을 나타낸다. U6 프로모터의 제어 하의 Tet1-3 및 Dnmt1, 3a 및 3b gRNA. 인간 시냅신 프로모터의 제어 하의 GFP-KASH. 제한측은 서브클로닝에 의한 단순한 gRNA 대체 전략을 나타낸다. 2 개의 핵 국소화 신호(NLS)에 측접한 HA-태그된 SpCas9를 나타낸다. 벡터는둘 다 1:1 비로 AAV1/2 바이러스에 의해 뇌에 전달된다.

도 35는 서베이어 분석을 사용하는 다중 DNMT 표적화 벡터 #1 작용성의 확인을 나타낸다. N2A 세포를 DNMT 유전자 패밀리 좌의 SpCas9 매개 절단을 시험하기 위해 DNMT 표적화 벡터 #1(+) 및 SpCas9 암호화 벡터에 의해 공동 트렌스펙션시켰다. gRNA는 단지(-) 음성 대조군이다. 트렌스펙션의 48 시간 후 DNA 정제 및 하류의 가공을 위해 세포를 채취하였다.

도 36은 서베이어 분석을 사용하는 다중 DNMT 표적화 벡터 #2 작용성의 확인을 나타낸다. N2A 세포를 DNMT 유전자 패밀리 좌의 SpCas9 매개 절단을 시험하기 위해 DNMT 표적화 벡터 #1 (+) 및 SpCas9 암호화 벡터로 공동 트렌스펙션시켰다. gRNA는 단지(-) 음성 대조군이다. 트렌스펙션의 48 시간 후 DNA 정제 및 하류의 가공을 위해 세포를 채취하였다.

도 37은 생체내 HA-SpCas9 발현을 위해 사용한 짧은 프로모터 및 짧은 폴리A 형태의 개략적 개요를 나타낸다. L-ITR로부터 R-ITR까지의 암호화 영역의 크기를 우측에 나타낸다.

도 38은 생체내 HA-SaCas9 발현에 대해 사용한 짧은 프로모터 및 짧은 폴리A 형태의 개략적 개요를 나타낸다. L-ITR로부터 R-ITR까지의 암호화 영역의 크기를 우측에 나타낸다.

도 39는 N2A 세포 내 SpCas9 및 SaCas9의 발현을 나타낸다. 상이한 짧은 프로모터의 제어 하의 그리고 짧은 폴리A(spA) 서열에 의한 HA-태그 SpCas9 및 SaCas9 형태의 대표적 웨스턴 블롯. 튜불린은 장입 대조군이다. mCherry(mCh)는 트렌스펙션 대조군이다. 세포를 채취하고 트렌스펙션의 48 시간 후에 웨스턴 블롯팅에 대해 추가로 처리하였다.

도 40은 Tet3 유전자좌의 효율적 SaCas9 매개 표적화에 대한 선별을 나타낸다. 트렌스펙션 N2A 세포로부터 DNA에 대한 서베이어 분석은 NNGGGT PUM 서열을 지니는 상이한 gRNA를 사용함으로써 효율적인 DNA 절단을 입증한다. GFP 트렌스펙션 세포 및 SaCas9만을 발현시키는 세포는 대조군이다.

도 41은 마우스 뇌에서 HA-SaCas9의 발현을 나타낸다. 동물을 인간 시냅신 프로모터의 제어 하에서 HA-SaCas9의 바이러스 구동 발현을 지니는 치아이랑(dentate gyri) 내로 주사하였다. 수술 2 주 후에 동물을 희생시켰다. 토끼 단클론성 항체 C29F4(세포 신호처리)를 사용하여 HA 태그를 검출하였다. 세포핵을 DAPI 염색에 의해 푸른색으로 염색하였다.

도 42는 트렌스펙션 후 7 일 배양에서 뇌 1 차 뉴런 내 SpCas9 및 SaCas9의 발현을 나타낸다. 상이한 프로모터의 제어 하에서 bgh 또는 짧은 폴리A(spA) 서열을 지니는 HA-태그 SpCas9 및 SaCas9 형태의 대표적 웨스턴 블롯. 튜블린은 장입 대조군이다.

도 43은 상이한 프로모터 및 다중 gRNA 작제물(DNMT에 대한 최종 패널 상에 나타낸 예)을 지니는 SpCas9를 운반하는 AAV1 입자를 이용하여 형질도입하고 7 일 후에 1 차 뇌 뉴런의 LIVE/DEAD 염색을 나타낸다. AAV 형질도입 후 뉴런을 대조군 비형질도입 뉴런과 비교하였다. 적색 핵은 침투되고 사멸된 세포(패널의 두 번째 줄)를 나타낸다. 살아있는 세포는 녹색으로 표시한다(패널의 세 번째 줄).

도 44는 상이한 프로모터를 지니는 SpCas9를 운반하는 AAV1 입자를 이용하여 형질도입하고 7 일 후에 1 차 뇌 뉴런의 LIVE/DEAD 염색을 나타낸다. 적색 핵은 침투되고 사멸된 세포(패널의 두 번째 줄)를 나타낸다. 살아있는 세포는 녹색으로 표시한다(패널의 세 번째 줄).

도 45는 TET 및 DNMT 유전자좌에 대해 SpCas9 및 gRNA 다중 복합체를 운반하는 AAV1 바이러스를 이용한 형질도입 후 뉴런 형상의 비교를 나타낸다. 형질도입이 없는 뉴런을 대조군으로서 나타낸다.

도 46은 1 차 뇌 뉴런에서의 서베이어 분석을 사용하는 다중복합 DNMT 표적화 벡터 #1 기능성의 확인을 나타낸다. DNMT 유전자 패밀리 좌의 SpCas9 매개 절단을 시험하기 위해 상이한 프로모터를 지니는 DNMT 표적화 벡터 #1 및 SpCas9로 세포에 공동형질도입하였다.

도 47은 뇌에서 SpCas9 절단의 생체내 효율을 나타낸다. 마우스에 2 개의 상이한 프로모터: 마우스 Mecp2 및 래트 Map1b의 제어 하에 SpCas9와 함께 DNMT 패밀리 유전자좌를 표적화하는 gRNA 다중복합체를 운반하는 AAV1/2 바이러스를 주사하였다. 주사하고 2 주 후, 뇌 조직을 추출하고 나서 핵을 준비하였고, gRNA 다중복합체 작제물로부터의 시냅신 프로모터에 의해 구동되는 GFP 발현에 기반하여 FACS를 사용하여 분류하였다. gDNA 추출 후에 서베이어 분석을 실행하였다. +는 GFP 양성핵을 나타내고, -는 대조군, 동일한 동물로부터의GFP-음성핵을 나타낸다. 겔 상의 번호는 평가한 SpCas9 효율을 나타낸다.

도 48은 해마 뉴런으로부터의 GFP-KASH 표지 세포핵의 정제를 나타낸다. 세포핵막의 외부 핵막(outer nuclear membrane: ONM)을 GFP와 KASH 단백질 막관통영역의 융합으로 태그한다. 주촉성 수술 및 AAV1/2 주사의 일주일 후 뇌에서의 강한 GFP 발현. 무결함 뇌로부터 세포핵을 정제하기 위한 밀도 구배 원심분리 단계. 정제 핵이 도시되어 있다. Vybrant® DyeCycle™ Ruby Stain에 의한 염색질 염색은 적색으로 나타나고, GFP 표지 핵은 녹색이다. GFP+ 및 GFP- 세포핵의 대표적인 FACS 프로파일(마젠타: Vybrant® DyeCycle™ Ruby Stain, 녹색: GFP).

도 49는 마우스 뇌에서 SpCas9 절단의 효율을 나타낸다. 마우스에 2 개의 상이한 프로모터: 마우스 Mecp2 및 래트 Map1b의 제어 하에 SpCas9 바이러스와 함께 TET 패밀리 유전자좌를 표적화하는 gRNA 다중복합체를 운반하는 AAV1/2 바이러스를 주사하였다. 주사하고 3 주 후, 뇌 조직을 추출하고 나서 핵을 준비하였고, gRNA 다중복합체 작제물로부터의 시냅신 프로모터에 의해 구동되는 GFP 발현에 기반하여 FACS를 사용하여 분류하였다. gDNA 추출 후에 서베이어 분석을 실행하였다. +는 GFP 양성핵을 나타내고, -는 대조군, 동일한 동물로부터의 GFP-음성핵을 나타낸다. 겔 상의 번호는 평가한 SpCas9 효율을 나타낸다.

도 50은 배양물 내 피질 뉴런에서의 GFP-KASH 발현을 나타낸다. TET 유전자좌를 표적화하는 AAV1 바이러스 운반

gRNA 다중복합체 작제물에 의해 뉴런을 형질도입하였다. 가장 강한 신호는 KASH 도메인 국소화에 기인하여 세포 핵 주위에서 국소화된다.

도 51은 (상부) 가이드 RNA의 쌍 사이의 간격의 목록(2 개의 PAM 서열에 대한 배열 패턴에 의해 표시한 바와 같음)을 나타낸다. 패턴 1, 2, 3, 4를 충족하는 가이드 RNA 쌍만이 SpCas9(D10A) 닉카아제와 함께 사용될 때 삽입결실을 나타내었다. (하부) 패턴 1, 2, 3, 4를 충족하는 가이드 RNA의 쌍과 SpCas9(D10A)의 조합을 나타내는 겔 이미지는 표적 부위 내 삽입결실의 형성을 야기한다.

도 52는 U6-가이드 RNA 발현 카세트를 생성하기 위해 사용되는 U6 역방향 프라이머 서열의 목록을 나타낸다. 각각의 프라이머는 U6 및 목적으로 하는 가이드 RNA를 함유하는 앰플리콘을 생성하기 위해 U6 정방향 프라이머 "gcactgagggcctatttcccatgattc"와 짝지어질 필요가 있다.

도 53은 도 33에서 열거한 24 개 패턴의 위치를 나타내는 인간 Emx1 좌로부터의 게놈 서열 맵을 나타낸다.

도 54는 (우측) 상이한 쌍의 가이드 RNA에 의해 표적화된 Cas9 닉카아제에 의한 절단 후 가변 5' 돌출부가 존재할 때 표적에서 삽입결실의 형성을 나타내는 겔 이미지를 나타낸다. (좌측) 표는 우측에서 겔의 레인 수 및 사용한 가이드 RNA 쌍 및 Cas9 닉카아제에 의한 절단 후 존재하는 5' 돌출부의 길이를 동정하는 것을 포함하는 다양한 파라미터를 나타낸다.

도 55는 도 54(우측)의 겔 패턴을 초래하고 실시예 35에서 추가로 기재되는 상이한 쌍의 가이드 RNA의 위치를 나타내는 인간 Emx1 유전자좌로부터의 게놈 서열 맵을 나타낸다.

도 56A 내지 K는 일차 피질 뉴런에서 *Mecp2*의 CRISPR-Cas9 표적화를 나타낸다. (A) AAV SpCas9 및 sgRNA 발현 벡터. sgRNA 벡터는 형질도입된 뉴런의 동정을 위한 GFP-KASH 융합 단백질의 암호화 서열을 함유한다. (B) HA-태깅된 Cas9(HA-Cas9) 및 GFP-KASH의 발현을 나타내는 Cas9 및 sgRNA 벡터로 공동-형질도입된 배양물 중 뉴런. DAPI로 표지화된 핵. 축척 막대, 20 μm. (c) 일차 피질 뉴런에서 GFP-KASH^{+}(n=635) 및 HA-Cas9(n=659)의 공동-감염 효율. (D) Cas9 표적 위치를 나타내는 마우스 *Mecp2* 유전자좌의 도표 표시; sgRNA는 청색으로 나타내었다. PAM 서열은 자주색으로 표시하였다. (E) 피질 뉴런에서 *Mecp2* 유전자좌의 서베이어(SURVEYOR)™ 분석 겔. (F) *Mecp2* 유전자좌의 CRISPR-Cas9 표적화 후 MeCP2 단백질 수준의 웨스턴 블롯 및 MeCP2 단백질 수준의 정량화(*t*-검정, ***p<0.0001, n=7). (G) *Mecp2* 유전자좌의 CRISPR-Cas9 표적화 후 뉴런의 가지돌기 나무에서의 감소된 복잡성. 축척 막대, 20 μm. (H) 수지상 말단의 수로 평가된 가지돌기 나무 형태 및 (I) Sholl 분석(t-검정, ***p<0.0001, n=40). (J) Cas9 및 *Mecp2* sgRNA로 표적화된 뉴런에서의 가지돌기 가시 형태의 변화. 축척 막대, 10 μm. (K) 가시 밀도 정량화(t-검정, ***p<0.0001, n=40).

도 57A 내지 I는 마우스 뇌에서 *Mecp2*의 표적화 및 CRISPR-Cas9 시스템 전달을 나타낸다. (A) 마우스 뇌로부터 CRISPR-Cas9의 표적화 세포의 세포 핵 정제의 전략. (B) 마우스 해마의 등쪽 치아이랑(DG)에서 HA-Cas9 및 GFP-KASH(sgRNA)의 발현. 축척 막대, 100 μm. (C) 2 개의 벡터 Cas9-CRISPR 시스템에 의해 효율적으로 표적화된 세포의 정량화. (D) DG 영역에서 AAV 전달 2 주 후에 *Mecp2* 유전자좌의 변형을 나타내는 서베이어™ 분석 겔. FACS 분류된 GFP-KASH 양성 세포는 더 높은 수준의 *Mecp2* 유전자좌 변형을 나타낸다. (E) 표적화된 뇌 영역에서 MeCP2 단백질 발현의 웨스턴 블롯 분석 및 등쪽 DG에서의 MeCP2 단백질 수준의 정량화(*t*-검정, **p<0.001, n=4). (F) *Mecp2* 유전자좌의 CRISPR-Cas9 표적화 2 주 후 등쪽 DG 영역의 이미지. 축척 막대, 150 μm. (G) 대조군 측면 부위와 비교하여 표적화된 뇌 영역에서 MeCP2 양성 세포의 집단의 정량화(*t*-검정, ****p<0.0001, 각각 n=290 개 및 249 개 세포,). (H) CRISPR-Cas9 전달 1 주 후 등쪽 DG 에서 과립 세포의 가지돌기 가시의 형태를 나타내는 골지-콕스(Golgi-Cox) 염색의 예. 축척 막대, 10 μm. (I) 등쪽 DG 영역에서 가지돌기 가시 밀도의 정량화(*t*-검정, ***p<0.0001, n=20).

도 58A 내지 F는 마우스 뇌에서 동시의 다중복합체 유전자 편집을 나타낸다. (A) 다중복합체 게놈 표적화를 위해 설계된 CRISPR-Cas9 시스템의 개략적 예시. (B) 표적화된 DNMT 마우스 유전자좌의 도표 표시. 가이드 RNA는 청색으로 표시된다. PAM 서열은 자주색으로 표시된다. (C) CRISPR-Cas9 전달 후에 치아이랑으로부터 FACS 분류된 핵에서의 DNMT 패밀리 유전자에 대한 표적내 변형률의 차세대 시퀀싱. MLE(maximum-likelihood estimator; 최대 가능도 추정량)가 나타내어져 있다. (D) DNMT 패밀리 유전자를 표적화하는 CRISPR-Cas9 시스템의 생체내 전달 후의 Dnmt3a 및 Dnmt1 단백질에 대한 웨스턴 블롯 분석(상부). 생체내 CRISPR-Cas9 표적화 후의 DG 에서의 Dnmt3a 및 Dnmt1 단백질 수준의 웨스턴 블롯 정량화(하부; *t*-검정, **p<0.001, *p<0.05, Dnmt3a: n=7; Dnmt1: n=5). (E, F) 트레이닝(E) 및 변경된 상황(F)에서 시험된, 해마의 DG 영역에서 SPR-Cas9를 사용한 DNMT 유전자의 표적화 8 주 후의 문맥상 학습 결핍(*t*-검정, ***p<0.0001, n=18).

도 59A 내지 E는 AAV 패키징에 대한 HA-태깅된 SpCas9(HA-Cas9)의 클로닝 및 발현을 나타낸다. (A) CRISPR/Cas9 시스템의 개략적 개관. Cas9의 단일 가이드 RNA(sgRNA) 매개 표적화는 표적화된 유전자 유전자좌의 이중가닥 브레이크(DSB)를 생성한다. 비-상동성 단부 연결(NHEJ) 메커니즘은 표적화된 게놈 유전자좌의 삽입결실 돌연변이를 생성한다. (B) 짧은 래트 *Map1b* 프로모터(*rMap1b*), 마우스 *Mecp2* 프로모터(*sMecp2*)의 트렁케이션된 형태 및 짧은 polyA 모티프(spA)를 사용한 Cas9 발현 카세트 크기를 최소화시키기 위한 다양한 클로닝 전략의 개략적 개관. (C) 다양한 Cas9 발현 카세트를 사용한 Cas9를 발현하는 일차 피질 뉴런 배양물의 웨스턴 블롯 분석. (D) *Mecp2* 프로모터는 뉴런(Map1b, NeuN; 화살표)에서 Cas9(적색) 발현을 구동하나, 성상교세포(GFAP, 화살촉)에서는 그렇지 않다. 핵은 DAPI(청색)로 표지화되었다. 축척 막대, 20 μm. (E) 세포는 바이러스 전달 7일 후에 LIFE/DEAD® 키트로 염색되었다. $DAPI^+$ 및 죽은($DEAD^+$) 세포의 정량화(ITR - 역위된 말단 반복; HA - 헤마글루티닌 태그; NLS - 핵 국소화 신호; spA - 합성 폴리아데닐화 신호; U6 - PolIII 프로모터; sgRNA - 단일 가이드 RNA; hSyn - 인간 시냅신 1 프로모터; GFP - 녹색 형광 단백질; KASH - Klarsicht, ANC1, Syne 상동성 핵 막횡단 도메인; bGH pA - 소 성장 호르몬 폴리아데닐화 신호; WPRE - 우드척 간염 바이러스 전사후 조절 요소).

도 60A 및 B는 Neuro-2a 세포에서 *Mecp2*의 표적화를 나타낸다. (A) *Mecp2* 표적화 서열 및 상응하는 프로토스페이서 인접 모티프(PAM). (B) Neuro-2a 세포로 Cas9와 공동-트랜스펙션된 6 개 Mecp2 sgRNA의 평가. 유전자좌 변형 효율이 서베이어™ 분석을 사용하여 트랜스펙션 48 시간 후에 분석되었다.

도 61은 일차 피질 뉴런에서의 CRISPR-Cas9 전달을 나타낸다. AAV-CRISPR 형질도입 7 일 후에 배양된 뉴런에서의 MeCP2(적색)의 면역형광 염색(녹색, GFP-KASH). *Mecp2*-표적화 AAV-CRISPR에 의해 형질도입된 세포에서 감소된 MeCP2 면역형광(중간 패널)이 나타내어져 있다. 핵은 DAPI(청색)로 표지화되었다. 축척 막대, 20 μm.

도 62A 내지 C는 표적화된 세포 핵의 GFP-표지화를 나타낸다. (A) GFP-표지화의 개략적 개관. 핵 막횡단 KASH 도메인에 융합된 향상된 녹색 형광 단백질(GFP) 및 핵 외막으로의 GFP-KASH의 통합이 예시된다. (B) 치아이랑으로의 바이러스 전달 4 주 후, GFP-KASH의 인간 시냅신 프로모터에 의해 구동된 발현. 헤마톡실린/에오신 염색(상부)은 어떠한 형태학적 이상도 나타내지 않았다. GFP-KASH 발현 헤마에서 정상적인 조직형태를 나타내는 면역형광 분석(NeuN, 적색으로 나타냄, 중간 패널) 및 성상교세포(GFAP, 적색으로 나타냄, 하부 패널)의 징후는 없다. 핵은 DAPI(청색)로 표지화되었다. 축척 막대, 200 μm. (C) 세포-유형 특이적 프로모터를 사용함으로써, GFP-KASH는 다양한 세포-유형에 표적화될 수 있다. 신경교 섬유질 산성 단백질(GFAP) 프로모터는 마우스 헤마의 성상교세포(적색)에서 GFP-KASH(녹색) 발현을 구동한다. 핵은 DAPI(청색)로 표지화되었다. 인세트는 더 큰 배율을 나타낸다. 축척 막대, 50 μm. (KASH - **K**larsicht, **A**NC1, **S**yne 상동성 핵 막횡단 도메인) (ONM - 외부 핵막; INM - 내부 핵막).

도 63A 및 B는 시험관내 DNMT 패밀리 일원의 다중복합체 게놈 표적화를 나타낸다. (A) *Dnmt3a*, *Dnmt1* 및 *Dnmt3b* 표적화 서열 및 상응하는 프로토스페이서 인접 모티프(PAM). (B) *Dnmt3a*, *Dnmt1* 및 *Dnmt3b* 유전자좌를 표적화하는 Cas9 및 DNMT 3xsgRNA 벡터를 이용한 트랜스펙션 48 시간 후 Neuro-2a 세포의 서베이어™ 뉴클레아제 검정 분석. 3 개 모두의 표적화된 유전자의 효과적인 게놈 편집이 나타내어져 있다.

도 64A 내지 C는 표적화된 *Dnmt3a*, *Dnmt1* 및 *Dnmt3b* 유전자좌의 차세대 시퀀싱을 나타낸다. 마우스 치아이랑으로의 Cas9 및 DNMT 3xsgRNA의 생체내 전달 후의 돌연변이된 *Dnmt3a*(A), *Dnmt1*(B) 및 *Dnmt3b*(C) 유전자좌의 시퀀싱 결과의 예. 녹색: 야생형 서열, 적색 점선: 결실된 염기, 적색 기선: 삽입 또는 돌연변이. 적색 화살촉은 CRISPR-Cas9 절단 부위를 나타낸다.

도 65a는 가이드(표적) 1이 ApoB에서 삽입결실의 가장 높은 백분율을 유도한 것을 나타낸다.

도 65b는 주사하고 4주 후, 삽입-결실 형성 효율에 대한 서베이어 뉴클레아제 겔 검정의 결과를 나타낸다.

도 66은 AAV-Cas9-sgRNA 전달 후 생체내에서 간 지질 축적 표현형을 검출하기 위한 오일 레드 염색을 나타낸다. 각각의 스퀘어 내의 축척 막대는 20 마이크로미터이다.

도 67은 회색 막대에 의해 표현된 21 ntds/염기쌍(bp)은 적어도 일정 범위의 표적 전체에 걸쳐 및 2개의 상이한 유전자(AAVS1 및 EMX1) 내에서 20 또는 22개의 염기쌍(각각 흑색 및 백색 막대에 의해 표현됨)에 비해 최적의 스페이서 길이인 것을 나타낸다.

도 68은 가이드 서열이 Cas9 인트론 서열로 삽입될 수 있는 지의 여부를 나타낸다.

도 69는 U6가 각각의 시험되는 표적에 대해 증가된 삽입-결실 백분율 형성을 나타냄에 따라 전장 H1 프로모터

(회색 막대)가 여전히 U6 프로모터(흑색 막대)보다 약한 것을 나타낸다.

도 70은 짧은 H1 프로모터가 전장 H1보다 약한 것을 나타낸다.

도 71은 D10A SaCAs9 이중 닉카아제의 절단 효율과 관련하여 측정된 작제물 내의 2개의 가이드 서열의 5' 말단 사이의 거리를 나타낸다.

도 72는 마우스 뇌에서의 CRISPR-Cas9 시스템 전달 및 *Mecp2* 유전자좌의 표적화를 나타낸다. (**a**) AAV-SpCas9 및 AAV-SpGuide(*Mecp2*) 발현 벡터. sgRNA 벡터는 형질도입된 뉴런의 확인을 위한 GFP-KASH 융합 단백질의 암호화 서열을 함유한다. (**b**) 마우스 해마의 등 치아이랑(DG) 내에서의 HA-Cas9 및 GFP-KASH의 발현. 축척 막대, 100 ㎛. (**c**) 이중-벡터 Cas9-CRISPR 시스템에 의해 효과적으로 표적화된 세포의 정량. (**d**) Cas9 표적 위치를 나타내는 마우스 *Mecp2* 유전자좌의 도표 표시; sgRNA는 청색으로 나타내었다. PAM 서열은 자주색으로 표시하였다. *Mecp2* 유전자좌의 시퀀싱에 의해 검출된 대표적 돌연변이 패턴이 하기에 제시되었다: 녹색 - 야생형 서열; 적색 점선 - 결실된 염기; 적색 기선: 삽입 또는 돌연변이; 적색 화살촉 모양은 CRISPR-Cas9 절단 부위를 나타낸다. (**e**) DG 영역 내에서의 AAV 전달 2주 후의 *Mecp2* 유전자좌의 변형을 나타내는 SURVEYOR™ 검정 겔. (**f**) 표적화된 뇌 영역에서의 MeCP2 단백질 발현의 웨스턴 블롯 분석 및 등 DG에서의 MeCP2 단백질 수준의 정량(*t*-검정, $^{**}p<0.001$, 3마리의 동물로부터의 n=4, 오차 막대: s.e.m.). (**g**) *Mecp2* 유전자좌의 CRISPR-Cas9 표적화 2주 후의 등 DG 영역의 이미지. 축척 막대, 150 ㎛. (**h**) 대조군 곁가지 부위에 비한 표적화된 뇌 영역에서의 모든 검출된 세포(DAPI 염색) 내의 MeCP2 양성 세포 집단의 정량(*t*-검정, $^{****}p<0.0001$, 각각 2마리의 동물로부터의 n=290 및 249 세포; 오차 막대: s.e.m). (ITR - 역위된 말단 반복; HA - 헤마글루티닌 태그; NLS - 핵 국소화 신호; spA - 합성 폴리아데닐화 신호; U6 - PolIII 프로모터; sgRNA - 단일 가이드 RNA; hSyn - 인간 시냅신 1 프로모터; GFP - 녹색 형광 단백질; KASH - Klarsicht, ANC1, Syne 상동성 핵 막횡단 도메인; bGH pA - 소 성장 호르몬 폴리아데닐화 신호; WPRE - 우드척(Woodchuck) 간염 바이러스 전사후 조절 구성요소).

도 73은 Cas9-매개 MeCP2 녹다운 뉴런에서의 유전자 발현의 분석을 나타낸다. (**a**) 마우스 뇌로부터의 CRISPR-Cas9 표적화 세포의 세포 핵 정제를 위한 전략. (**b**) RNAseq에 의해 검출된 차별적으로 발현된 유전자(*t*-검정, $p<0.01$, 8마리의 동물로부터의 분류된 핵의 n=19 집단)의 계통적 클러스터링. 유전자의 상대적 log2(TPM+1) 발현 수준이 각각의 열에 대해 표준화되고, 적색-청색 색 스케일로 표시되었다. 각각의 컬럼은 표시된 바와 같이 대조군 또는 *Mecp2* sgRNA 형질도입된 동물로부터의 분리된 세포의 치아이랑 집단으로부터 FACS 분류된 표적화된 100개의 뉴런 핵의 집단을 나타낸다.

도 74는 CRISPR-매개 MeCP2 녹다운 후의 뉴런의 세포 반응 특성에서의 세포-자율 결함을 나타낸다. (a) 마우스 시각 피질 및 시각 자극 파라미터로부터의 생체내 실험 형태를 나타내는 도면. GFP^+ 뉴런이 제시된다. 축척 막대, 20 ㎛. (b) 대측 및 동측 눈 특이적 투입을 받은 층 2/3 흥분 뉴런에서의 기록 형태를 나타내는 도면. 게놈 변형된 GFP^+ 세포는 녹색으로 존재하는 반면, 변형되지 않은 세포는 회색으로 존재한다. 표준화된 스파이크 형태는 정상 스파이킹 흥분 뉴런을 나타낸다. (c,d) 평균 OSI(c) 및 유발 FR(d)은 각각 *Mecp2* 및 대조군 sgRNA를 발현하는 GFP^+ 세포로부터 측정되었다(*t*-검정, $^{*}p<0.05$; 도면에서의 수는 기록된 세포의 수를 나타냄; n=2-3 동물; 오차 막대: s.e.m).

도 75는 마우스 뇌에서 동시의 다중복합체 유전자 편집을 나타낸다. (a) 다중복합체 게놈 표적화를 위해 설계된 CRISPR-Cas9 시스템의 개략적 예시. (b) 표적화된 DNMT 마우스 유전자좌의 도표 표시. 가이드 RNA는 청색으로 표시된다. PAM 서열은 자주색으로 표시된다. (c) DG 영역에서 AAV 전달 4주 후에 FACS 분류된 GFP-KASH 양성 세포에서의 DNMT 유전자좌의 변형을 나타내는 SURVEYOR™ 검정 겔. (d) 다수의 유전자좌에서의 변형의 공동-발생을 나타내는, 단일 세포에서의 DNMT 유전자좌 변형의 심층 시퀀싱-기반 분석. (e) DNMT 패밀리 유전자를 표적화하는 CRISPR-Cas9 시스템의 생체내 전달 후의 Dnmt3a 및 Dnmt1 단백질에 대한 웨스턴 블롯 분석(상부). 생체내 CRISPR-Cas9 표적화 후의 DG 에서의 Dnmt3a 및 Dnmt1 단백질 수준의 웨스턴 블롯 정량(하부; t-검정, $^{**}p<0.001$, $^{*}p<0.05$, Dnmt3a: n=7; Dnmt1: 5마리의 동물로부터의 n=5; 오차 막대: s.e.m). (f) 트레이닝 및 변경된 상황에서 시험된 해마의 DG 영역에서 SpCas9를 이용한 DNMT 유전자의 표적화 8주 후의 문맥상 학습 결핍(t-검정, $^{***}p<0.0001$, n=18 동물, 2개의 독립적 실험; 오차 막대: s.e.m).

도 76은 AAV 패키징에 대한 HA-태깅된 SpCas9(HA-SpCas9)의 클로닝 및 발현을 나타낸다. (a) 짧은 래트 *Map1b* 프로모터(pMap1b), 마우스 *Mecp2* 프로모터(pMecp2)의 트렁케이션된 형태 및 짧은 polyA 모티프(spA)를 이용한 SpCas9 발현 카세트 크기를 최소화시키기 위한 다양한 클로닝 전략의 개략적 개관. (**b**) 다양한 SpCas9 발현 카세트를 이용한 HA-SpCas9를 발현하는 일차 피질 뉴런 배양물의 웨스턴 블롯 분석. (**c**) *Mecp2* 프로모터는 뉴런

(Map1b, NeuN; 화살표)에서 HA-SpCas9(적색) 발현을 유도하나, 성상교세포(GFAP, 화살촉)에서는 그렇지 않다. HA-SpCas9와 GFP-KASH의 공동-발현이 제시된다(하부). 핵은 DAPI(청색)로 표지되었다. 축척 막대, 20 ㎛. (**d**) GFP-표지화의 개략적 개관. 핵 막횡단 KASH 도메인에 융합된 향상된 녹색 형광 단백질(GFP) 및 핵 외막으로의 GFP-KASH의 통합이 예시된다. (**e**) HA-SpCas9 및 GFP-KASH 둘 모두를 발현하는 세포의 집단을 나타내는 공동-감염 효율 계산(3개의 배양물로부터의 n=973 뉴런; 오차 막대: s.e.m). (**f**) 세포는 바이러스 전달 7일 후에 LIFE/DEAD® 키트로 염색되었다. $DAPI^+$ 및 죽은($DEAD^+$) 세포의 정량(대조군 n=518 $DAPI^+$ 핵; SpCas9/GFP-KASH 2개의 배양물로부터의 n=1003 $DAPI^+$ 핵; 오차 막대: s.e.m). (ITR - 역위된 말단 반복; HA - 헤마글루티닌 태그; NLS - 핵 국소화 신호; spA - 합성 폴리아데닐화 신호; U6 - PolIII 프로모터; sgRNA - 단일 가이드 RNA; hSyn - 인간 시냅신 1 프로모터; GFP - 녹색 형광 단백질; KASH - Klarsicht, ANC1, Syne 상동성 핵 막횡단 도메인; bGH pA - 소 성장 호르몬 폴리아데닐화 신호; WPRE - 우드척 간염 바이러스 전사후 조절 구성요소).

도 77은 Neuro-2a 세포에서 *Mecp2*의 표적화를 나타낸다. (a) *Mecp2* 표적화 서열 및 상응하는 프로토스페이서 인접 모티프(PAM). (b) Neuro-2a 세포로 SpCas9와 공동 트랜스펙션된 6 *Mecp2* sgRNA의 평가. 유전자좌 변형 효율이 SURVEYOR™ 검정을 이용한 트랜스펙션 48시간 후에 분석되었다.

도 78은 일차 피질 뉴런에서의 *Mecp2*의 CRISPR-SpCas9 표적화를 나타낸다. (a) AAV-CRISPR 형질도입 7일 후에 배양된 뉴런에서의 MeCP2(적색)의 면역형광 염색(녹색, GFP-KASH). 핵은 DAPI로 표지되었다(청색). 축척 막대, 20 ㎛. (**b**) SURVEYOR™ 검정 겔을 이용한 *Mecp2* sgRNA 또는 대조군(표적화 박테리아 *lacZ* 유전자) sgRNA와 함께 SpCas9 또는 dSpCas9를 이용한 *Mecp2* 유전자좌 표적화의 평가. (**c**) 뉴런(GFP^+)의 표적화된 집단에서의 MeCP2 양성 핵의 정량화. (**d**) *Mecp2* 유전자좌의 CRISPR-SpCas9 표적화 후의 MeCP2 단백질 수준의 웨스턴 블롯 및 MeCP2 단백질 수준의 정량화(*t*-검정, $^{**}p<0.001$, 3개의 배양물로부터의 n=5, 오차 막대: s.e.m).

도 79는 시험관 내에서 SpCas9-매개 MeCP2 녹다운 후의 뉴런의 가지돌기 나무에서의 형태학적 변화를 나타낸다. (a) *Mecp2* 유전자좌의 CRISPR-SpCas9 표적화 후의 뉴런에서의 가지돌기 나무의 감소된 복잡성. 축척 막대, 20 ㎛. (b) SpCas9 및 *Mecp2* sgRNA로 표적화된 뉴런에서의 가지돌기 가시 형태의 변화. 축척 막대, 10 ㎛. 세포의 형태가 mCherry 작제물과의 공동-트랜스펙션으로 시각화되었다. 형태 분석을 위한 세포는 Mecp2 염색의 결과를 기초로 하여 선택되었다. (c) 가지돌기 말단의 수로 평가된 가지돌기 나무 형태 및 (d) Sholl 분석(*t*-검정, $^{***}p<0.0001$, 2개의 배양물로부터의 n=40). (e) 가시 밀도 정량(*t*-검정, $^{***}p<0.0001$, 2개의 배양물로부터의 n=40, 오차 막대: s.e.m).

도 80은 대조군 동물로부터의 뉴런 핵의 RNAseq 및 SpCas9-매개 Mecp2 녹다운을 나타낸다. 발현 수준의 분위수당 RNA-seq 라이브러리(대조군 sgRNA 또는 *Mecp2* sgRNA 형질도입된 핵으로부터 수집된 각각 100개 핵의 19개의 라이브러리; n=4 동물/그룹) 전체에 걸쳐 검출된 유전자의 수를 나타내는 박스 플롯. 모든 유전자는 이들의 평균 log2(TPM+1) 발현 수준에 의해 10개의 분위수로 나누어진 후, 각각의 분위수에 대해, 검출되는 유전자의 수(log2(TPM+1)>2)가 각각의 샘플에서 계수되었다. 제시된 3개의 표적 서열은 Dnmt3a, Dnmt1 및 Dnmt3b에 대해 각각 서열 번호 ___, 서열 번호 ___ 및 서열 번호 ___이다.

도 81은 시험관 내에서 DNMT 패밀리 일원의 다중복합체 게놈 표적화를 나타낸다. (**a**) *Dnmt3a*, *Dnmt1* 및 *Dnmt3b* 표적화 서열 및 상응하는 프로토스페이서 인접 모티프(PAM). (**b**) *Dnmt3a*, *Dnmt1* 및 *Dnmt3b* 유전자좌를 표적으로 하는 SpCas9 및 DNMT 3xsgRNA 벡터를 이용한 트랜스펙션 48시간 후의 Neuro-2a 세포의 SURVEYOR™ 뉴클레아제 검정 분석. 3개 모두의 표적화된 유전자의 효과적인 게놈 편집이 제시된다.

도 82는 표적화된 *Dnmt3a*, *Dnmt1* 및 *Dnmt3b* 유전자좌의 차세대 시퀀싱을 나타낸다. 마우스 치아이랑으로의 SpCas9 및 DNMT 3xsgRNA의 생체내 전달 후의 돌연변이된 *Dnmt3a* (**a**), *Dnmt1* (**b**) 및 *Dnmt3b* (**c**) 유전자좌의 시퀀싱 결과의 예. 녹색: 야생형 서열, 적색 점선: 결실된 염기, 적색 기선: 삽입 또는 돌연변이. 적색 화살촉은 CRISPR-SpCas9 절단 부위를 나타낸다. 본 도면에서 사용된 전체 서열은 Dnmt3a, Dnmt1 및 Dnmt3b 유전자좌에 대해 각각 서열 번호 , 서열 번호 , 및 서열 번호 로 제공된다. 이들은 서열 번호 (Dnmt3a): CCT CCG TGT CAG CGA CCC ATG CCA A, 서열 번호 (Dnmt1): CCA GCG TCG AAC AGC TCC AGC CCG 및 서열 번호 (Dnmt3b) AGA GGG TGC CAG CGG GTA TAT GAG G이다.

본원의 도면은 단지 예시 목적을 위한 것으로, 반드시 축척에 의한 것은 아니다.

발명을 실시하기 위한 구체적인 내용

[0232] CRISPR-Cas 시스템에 대한 일반적 정보와 관련하여, 2013년 1월 30일, 2013년 3월 15일, 2013년 3월 28일,

2013년 4월 20일, 2013년 5월 6일 및 2013년 5월 28일에 각각 출원된 미국 가출원 61/758,468호, 61/802,174호, 61/806,375호, 61/814,263호, 61/819,803호 및 61/828,130호가 참조된다. 2013년 6월 17일에 출원된 미국 가출원 61/836,123호가 또한 참조된다. 2012년 12월 12일 및 2013년 1월 2일에 각각 출원된 미국 가출원 61/736,527호 및 61/748,427호가 또한 참조된다. 2013년 3월 15일에 출원된 미국 가출원 61/791,409호가 또한 참조된다. 2013년 3월 15일에 출원된 미국 가출원 61/799,800호가 또한 참조된다. 2013년 6월 17일에 각각 출원된 미국 가출원 61/835,931호, 61/835,936호, 61/836,127호, 61/836,101호, 61/836,080호 및 61/835,973호가 또한 참조된다. 2013년 8월 5일에 출원된 미국 가출원 61/862,468호 및 61/862,355호, 2013년 8월 28일에 출원된 61/871,301호, 2013년 9월 25일에 출원된 61/960,777호 및 2013년 10월 28일에 출원된 61/961,980호가 추가로 참조된다. 이들 출원 각각, 및 상기 출원에 또는 상기 출원의 절차 중에 인용된 모든 문헌("출원 인용 문헌") 및 상기 출원 인용 문헌에 인용되거나 참고된 모든 문헌과 함께 본원에 언급되거나 본원의 임의의 문헌 및 본원에 참조로서 포함된 임의의 문헌에서의 임의의 제품에 대한 임의의 지침서, 설명서, 제품 명세서, 및 제품 시이트(sheet)는 참조로서 본원에 포함되며, 본 발명의 실시에 이용될 수 있다. 모든 문헌(예를 들어, 이들 출원 및 출원 인용 문헌)은 각각의 개별적 문헌이 참조로서 포함하는 것이 구체적으로 그리고 개별적으로 표시된 것과 동일한 정도로 참조로서 본원에 포함된다.

[0233] 또한, CRISPR-Cas 시스템에 대한 일반적 정보와 관련하여, 하기가 언급된다:

- *Multiplex genome engineering using CRISPR/Cas systems*. Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., & Zhang, F. *Science* Feb 15;339(6121):819-23 (2013);
- *RNA-guided editing of bacterial genomes using CRISPR-Cas systems*. Jiang W., Bikard D., Cox D., Zhang F, Marraffini LA. Nat Biotechnol Mar;31(3):233-9 (2013);
- *One-Step Generation of Mice Carrying Mutations in Multiple Genes by CRISPR/Cas-Mediated Genome Engineering*. Wang H., Yang H., Shivalila CS., Dawlaty MM., Cheng AW., Zhang F., Jaenisch R. Cell May 9;153(4):910-8 (2013);
- *Optical control of mammalian endogenous transcription and epigenetic states*. Konermann S, Brigham MD, Trevino AE, Hsu PD, Heidenreich M, Cong L, Platt RJ, Scott DA, Church GM, Zhang F. Nature. 2013 Aug 22;500(7463):472-6. doi: 10.1038/Nature12466. Epub 2013 Aug 23;
- *Double Nicking by RNA-Guided CRISPR Cas9 for Enhanced Genome Editing Specificity*. Ran, FA., Hsu, PD., Lin, CY., Gootenberg, JS., Konermann, S., Trevino, AE., Scott, DA., Inoue, A., Matoba, S., Zhang, Y., & Zhang, F. *Cell* Aug 28. pii: S0092-8674(13)01015-5. (2013);
- *DNA targeting specificity of RNA-guided Cas9 nucleases*. Hsu, P., Scott, D., Weinstein, J., Ran, FA., Konermann, S., Agarwala, V., Li, Y., Fine, E., Wu, X., Shalem, O., Cradick, TJ., Marraffini, LA., Bao, G., & Zhang, F. *Nat Biotechnol* doi:10.1038/nbt.2647 (2013);
- *Genome engineering using the CRISPR-Cas9 system*. Ran, FA., Hsu, PD., Wright, J., Agarwala, V., Scott, DA., Zhang, F. Nature Protocols Nov;8(11):2281-308. (2013);
- *Genome-Scale CRISPR-Cas9 Knockout Screening in Human Cells*. Shalem, O., Sanjana, NE., Hartenian, E., Shi, X., Scott, DA., Mikkelson, T., Heckl, D., Ebert, BL., Root, DE., Doench, JG., Zhang, F. *Science* Dec 12. (2013). [Epub ahead of print];
- *Crystal structure of cas9 in complex with guide RNA and target DNA*. Nishimasu, H., Ran, FA., Hsu, PD., Konermann, S., Shehata, SI., Dohmae, N., Ishitani, R., Zhang, F., Nureki, O. *Cell* Feb 27. (2014). 156(5):935-49;
- *Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells*. Wu X., Scott DA., Kriz AJ., Chiu AC., Hsu PD., Dadon DB., Cheng AW., Trevino AE., Konermann S., Chen S., Jaenisch R., Zhang F., Sharp PA. *Nat Biotechnol*. (2014) Apr 20. doi: 10.1038/nbt.2889, and
- *Development and Applications of CRISPR-Cas9 for Genome Engineering*, Hsu et al, Cell 157, 1262-1278 (June 5, 2014) (Hsu 2014),

[0234]

[0235] 이들 각각은 참조로서 본원에 포함되고, 하기에 간략히 논의된다:

[0236] ■ 콩 등(Cong *et al.*)은 스트렙토코커스 써모필루스(*Streptococcus thermophilus*) Cas9 및 또한 스트렙토코커스 피오게네스(*Streptococcus pyogenes*) Cas9 둘 모두를 기초로 한 진핵 세포에서 사용하기 위한 타입 II CRISPR/Cas 시스템을 유전자 조작하였고, Cas9 뉴클레아제가 짧은 RNA에 의해 지시되어 인간 및 마우스 세포에서 DNA의 정확한 절단을 유도할 수 있는 것을 입증하였다. 이들의 연구는 닉킹 효소로 전환된 바와 같은 Cas9가 최소의 돌연변이유발 활성으로 진핵 세포에서 상동성-지시된 복구를 촉진하는데 이용될 수 있음을 추가로 나타내었다. 또한, 이들의 연구는 다수의 가이드 서열이 단일 CRISPR 어레이로 암호화되어 포유동물 게놈 내의 내인성 게놈 유전자좌 부위에서 여러 동시 편집을 가능케 할 수 있음을 입증하였으며, 이는 RNA-가이드된 뉴클레아제 기술의 용이한 프로그램성 및 폭넓은 응용가능성을 입증한다. 세포에서 서열 특이적 DNA 절단을 프로그램화하는 RNA를 이용하는 이러한 능력은 게놈 유전자 조작된 도구의 새로운 부류를 규정하였다. 이들 연구는 다른 CRISPR 유전자좌가 포유동물 세포로 이식될 수 있고, 또한 포유동물 게놈 절단을 매개할 수 있음을 추가로 나타내었다. 중요하게는, CRISPR/Cas 시스템의 여러 양태가 이의 효능 및 다능성을 증가시키기 위해 추가로 개선될 수 있음이 예견될 수 있다.

[0237] ■ 지앙 등(Jiang *et al.*)은 스트렙토코커스 뉴모니에 및 에스케리키아 콜라이의 게놈에서 정확한 돌연변이를 도입시키기 위해 이중-RNA와 복합체화된 클러스터링되고 규칙적으로 산재된 짧은 팔린드로믹 반복부(CRISPR)-회합된 Cas9 엔도뉴클레아제를 이용하였다. 상기 접근법은 돌연변이되지 않은 세포를 사멸시키기 위해 표적화된 게놈 부위에서의 이중-RNA:Cas9-지시된 절단에 의존하며, 선택 마커 또는 반대-선택 시스템의 필요를 회피한다. 상기 연구는 편집 주형에서 수행된 단일- 및 멀티뉴클레오티드 변화를 제조하기 위해 짧은 CRISPR RNA (crRNA)의 서열을 변화시킴으로써 재프로그램화 이중-RNA:Cas9 특이성을 보고하였다. 상기 연구는 2개의 crRNA의 동시 사용이 다중복합체 돌연변이유발을 가능케 한 것을 나타내었다. 또한, 상기 접근법이 리컴비니어링(recombineering)과 조합하여 이용된 경우, S. 뉴모니에에서, 기재된 접근법을 이용하여 회수된 세포의 거의 100%가 요망되는 돌연변이를 함유하였고, E. 콜라이에서, 회수된 65%가 돌연변이를 함유하였다.

[0238] ■ 코너만 등(Konermann *et al.*)은 DNA-결합 도메인 기반 CRISPR Cas9 효소 및 또한 전사 활성인자 유사 이펙터의 광학적 및 화학적 조절을 가능케 하는 다능성이고 강한 기술에 대한 당 분야에서의 필요성을 다루었다.

[0239] ■ 본 명세서에서 논의된 바와 같이, 미생물 CRISPR-Cas 시스템으로부터의 Cas9 뉴클레아제는 20 nt 가이드 서열에 의해 특정 게놈 유전자좌로 표적화되며, 이는 DNA 표적에 대한 특정 미스매치를 용인할 수 있고, 이에 의해 요망되지 않는 표적외 돌연변이유발을 촉진할 수 있다. 이를 다루기 위해, 란 등(Ran *et al.*)은 표적화된 이중-가닥 파손을 도입시키기 위해 쌍을 이룬 가이드 RNA와 Cas9 닉카아제 돌연변이를 조합한 접근법을 기재하였다. 게놈 내의 개별적 닉은 높은 정확도로 복구되므로, 적절한 오프셋 가이드 RNA를 통한 동시 닉킹이 이중-가닥 파손에 필요하고, 표적 절단을 위한 특별히 인지된 염기의 수를 연장시킨다. 상기 저자는 쌍을 이룬 닉킹을 이용하는 것이 세포주에서 50 내지 1,500배까지 표적외 활성을 감소시킬 수 있고, 표적에 대한 절단 효율을 희생시키지 않고 마우스 접합체에서 유전자 녹아웃을 촉진하는 것을 입증하였다. 이러한 다능성 전략은 높은 특이성을 필요로 하는 매우 다양한 게놈 편집 적용을 가능케 한다.

[0240] ■ 슈 등(Hsu *et al.*)은 표적 부위의 선택을 알리고, 표적외 효과를 피하기 위해 인간 세포에서 SpCas9 표적화 특이성을 특성규명하였다. 상기 연구는 293T 및 293FT 세포에서 100개 초과의 예측 게놈 표적외 유전자좌에서 700개 초과의 가이드 RNA 변이체 및 SpCas9-유도된 삽입-결실 돌연변이 수준을 평가하였다. 상기 저자는 SpCas9가 미스매치의 수, 위치 및 분포에 민감한 서열-의존성 방식으로 다양한 위치에서 가이드 RNA와 표적 DNA 사이의 미스매치를 용인하는 것을 나타내었다. 상기 저자는 SpCas9-매개 절단이 DNA 메틸화에 의해 영향을 받지 않고, SpCas9 및 sgRNA의 투여량이 적정되어 표적외 변형을 최소화시킬 수 있음을 추가로 나타내었다. 또한, 포유동물 게놈 유전자 조작 적용을 촉진하기 위해, 상기 저자는 표적 서열의 선택 및 확인뿐만 아니라 표적외 분석을 가이드하기 위해 웹-기반 소프트웨어 도구를 제공한 것을 보고하였다.

[0241] ■ 란 등(Ran *et al.*)은 포유동물 세포에서 비-상동성 말단 연결(NHEJ) 또는 상동성-지시 복구(HDR)뿐만 아니라 하류 기능성 연구를 위한 변형된 세포주의 생성을 통한 Cas9-매개 게놈 편집을 위한 한 세트의 도구를 기재하였다. 표적외 절단을 최소화시키기 위해, 상기 저자는 쌍을 이룬 가이드 RNA와 함께 Cas9 닉카아제 돌연변이를 이용한 이중-닉킹 전략을 추가로 기재하였다. 상기 저자에 의해 제공된 프로토콜은 표적 부위의 선택, 절단 효율의 평가 및 표적외 활성의 분석을 위한 가이드라인을 실험적으로 유래하였다. 상기 연구는 표적 설계로 시작하여, 유전자 변형은 1-2주만큼 적은 기간 내에 달성될 수 있고, 변형된 클로날 세포주가 2-3주 내에 유래될 수 있음을 나타내었다.

[0242] ■ 샬렘 등(Shalem *et al.*)은 게놈-범위에 걸친 규모의 유전자 기능을 질의하기 위한 새로운 방식을 기재한다. 이들의 연구는 64,751개의 독특한 가이드 서열을 이용한 게놈-규모의 CRISPR-Cas9 녹아웃(GeCKO) 라이브러리 표적화된 18,080개 유전자의 전달이 인간 세포에서 음성 및 양성 선택 스크리닝 둘 모두를 가능케 한 것을 나타내었다. 먼저, 상기 저자는 암 및 다능성 줄기 세포에서 세포 생활력에 필수적인 유전자를 확인하기 위해 GeCKO 라이브러리의 사용을 나타내었다. 다음으로, 흑색종 모델에서, 상기 저자는 상실이 돌연변이 단백질 키나제 BRAF를 억제하는 치료제인 베무라페니브(vemurafenib)에 대한 내성과 관련된 유전자를 스크리닝하였다. 이들의 연구는 가장 높은 순위의 후보자가 이전에 확인된 유전자 NF1 및 MED12뿐만 아니라 새로운 히트 NF2, CUL3, TADA2B, 및 TADA1을 포함한 것을 나타내었다. 상기 저자는 동일한 유전자를 표적으로 하는 독립적 가이드 RNA와 높은 비율의 히트 확인 사이에 높은 수준의 일치를 관찰하였고, 따라서 Cas9를 이용한 게놈-규모 스크리닝의 가망성을 입증하였다.

[0243] ■ 니시마스 등(Nishimasu *et al.*)은 2.5 A° 해상도로 sgRNA 및 이의 표적 DNA와의 복합체에서 스트렙토코커스 피오게네스 Cas9의 결정 구조를 보고하였다. 상기 구조는 계면에서 양성으로 하전된 홈에서 sgRNA:DNA 헤테로듀플렉스를 수용하는 표적 인지 및 뉴클레아제 엽으로 구성된 두엽 구조를 나타내었다. 인지 엽이 sgRNA 및 DNA

결합에 필수적인 반면, 뉴클레아제 엽이 표적 DNA의 상보적 및 비-상보적 가닥의 절단을 위해 적절하게 위치된 HNH 및 RuvC 뉴클레아제 도메인을 각각 함유한다. 뉴클레아제 엽은 또한 프로토스페이서 인접 모티프(PAM)와의 상호작용을 담당하는 카르복실-말단 도메인을 함유한다. 이러한 고-해상도 구조 및 수반하는 기능 분석은 Cas9에 의한 RNA-가이드된 DNA 표적화의 분자 메커니즘을 나타내었고, 이에 따라 새로운 다능성 게놈-편집 기술의 합리적 설계를 수월케 한다.

[0244] ■ 우 등(Wu *et al.*)은 마우스 배아 줄기 세포(mESCs)에서 단일 가이드 RNA(sgRNAs)가 적하된 스트렙토코커스 피오게네스로부터 촉매적으로 비활성인 Cas9(dCas9)의 게놈 범위에 걸친 결합 부위를 맵핑하였다. 저자는 4개의 sgRNA 각각이 sgRNA 및 NGG 프로토스페이서 인접 모티프(PAM) 내의 5-뉴클레오티드 시드 영역을 종종 특징으로 하는 수십 내지 수천개의 게놈 부위에 대한 표적 dCas9를 시험한 것을 나타내었다. 염색질 비접근성은 시드 서열과 일치하는 다른 부위로의 dCas9 결합을 감소시키고, 따라서 표적외 부위의 70%가 유전자와 관련된다. 상기 저자는 촉매적으로 활성인 Cas9로 트랜스펙션된 mESC 내의 295개의 dCas9 결합 부위의 표적화된 시퀀싱이 백그라운드 수준 초과로 돌연변이된 단지 하나의 부위를 확인한 것을 나타내었다. 상기 저자는 시드 매치가 결합을 촉발하나, 표적 DNA와의 광범위한 페어링이 절단에 필요한 Cas9 결합 및 절단을 위한 2-상태 모델을 제안하였다.

[0245] ■ 슈(Hsu 2014)는 세포의 유전적 스크리닝을 포함하는 게놈 편집을 위한 요거트 기원의 CRISPR-Cas9를 일반적으로 논의하는 리뷰 기사이며, 이는 2014년 6월 5일 이전에 출원된 본 명세서 계통의 출원의 정보, 데이터 및 결과에 속한다. 슈(Hsu 2014)의 전반적 교시내용은 본 출원의 특정 모델, 동물을 포함하지 않는다.

[0246] 본 발명은 CRISPR-Cas 시스템 및 이의 구성성분에 관한 서열 표적화를 수반하는 유전자 발현의 제어, 예컨대 게놈 교란(perturbation) 또는 유전자 편집을 위해 사용되는 시스템, 방법 및 조성물의 유전자 조작 및 최적화에 관한 것이다. 유리한 구현예에서, Cas 효소는 Cas9이다.

[0247] CRISPR-Cas 폴리뉴클레오티드 서열은 일반적으로 가이드 또는 심지어 가이드 RNA(sgRNA)로 본원에서 언급되나, 이러한 용어는 이전에는 평범한 것이 아니었던 것이 인지될 것이다. 또한, CRISPR-Cas9 시스템이 본원에서 언급되나, 이는 DSB, 닉 또는 이중 닉을 유도하는 뉴클레아제 기능을 갖는 것을 조건으로 임의의 Cas에 대한 광범위한 언급임이 인지될 것이나, Cas9가 바람직하고, SaCas9가 특히 바람직하다.

[0248] 본 발명의 간 데이터에서 주요 포인트 중 일부가 일반적으로 유사분열 후 세포를 통해 하기에 요약되어 있는데, 이는 간 세포가 통상적으로 유사분열 후 세포이기 때문이다:

[0249] **AAV2/8**

[0250] CRISPR-Cas 시스템에 대한 바람직한 전달은 바이러스 벡터를 통하는 것이다. 이러한 벡터는 본원에 어느 정도로 논의된 바와 같은 렌티바이러스 벡터 또는 AAV 벡터일 수 있다. 본 발명자가 특히 제시한 것은 AAV가 바이러스 벡터의 바람직한 예라는 점이다. AAV 내에서, 본 발명자는 AAV8 및 특히 AAV2/8(AAV2 패키징 신호 ITR과 패키징된 AAV8)이 간, 특히 생체내로의 전달에 유용한 것을 제시하였다.

[0251] **생체내에서 관찰된 표현형 변화**

[0252] 다른 곳에 논의된 바와 같이, 본 발명자는 표현형 변화가 검출될 수 있음을 생체내에서 제시할 수 있었다. 이는 RNAi에서 종종 평등화된 결핍이 존재하여 영구적인 효과가 관찰되지 않음에 따라 의미있는 진전이다. 본 발명을 이용하여, 간에서 표현형 변화가 처음으로 관찰될 수 있었다. 상기를 달성하기 위한 바람직한 배열은 실시예 36에 기재된 것을 이용하는 것이다. 이의 중요한 요소는 하기의 바람직한 단독 또는 조합이다:

[0253] · Sa Cas9;

[0254] · 가이드, tracr 서열 및 tracr 메이트를 포함하는 키메라 가이드 RNA의 이용;

[0255] · tracr 서열에 대해, Sa Cas9를 보충하기 위해 Sa tracr가 바람직하다;

[0256] · AAV8 또는 더욱 바람직하게는 AAV2/8;

[0257] · 실험 목적을 위해, Rosa26은 유용한 음성 대조군이다;

[0258] · AAV 벡터에서 CMV 프로모터의 사용이 도움이 되나, 간-특이적 프로모터(간 표적화를 위함), 예를 들어, TBG의 사용이 특히 효과적이다;

[0259] · CRISPR이 표적 전체에 걸쳐 광범위한 적용성을 갖는 것으로 밝혀짐에 따라 표적 또는 표적들은 광범위할 수

있고, CRISPR가 가이드한 후, 이들은 성공적으로 전달되고, Css9 효소가 적합하게 발현된다. 그러나, 그럼에도 불구하고 간(가이드가 이에 대해 설계될 수 있음)에서 바람직한 표적은 PCSK9; Hmgcr; SERPINA1; ApoB; 및/또는 LDL 중 하나 이상을 포함한다.

[0260] 따라서, 일부 구현예에서, Cas 효소가 Sa Cas9인 것이 특히 바람직하다. 바람직하게는, CRISPRS-Cas 폴리뉴클레오티드 서열은 키메라이고, 바람직하게는 Cas9이 Sa Cas9인 Sa tracr를 포함한다. 바람직하게는 AAV2/8인 바이러스 벡터가 이용될 수 있다. 또한, 간-특이적 프로모터가 이상적이고, 바람직한 예는 TBG이다. 이들 모두는 Sa tracr를 포함하는 키메라 CRISPRS-Cas 폴리뉴클레오티드 서열을 제공하기 위해 조합하여 이용될 수 있으며, 여기서 Cas9는 SaCas9이고, 벡터는 적어도 Cas9가 TBG와 같이 간-특이적 조절 하에 있는 AAV2/8이다. 상기 표적 중 임의의 표적, 특히 비만에서의 중요성으로 인해 ApoB가 본 시스템과 함께 이용될 수 있다.

[0261] Yin 및 Anderson의 이후의 문헌[Nature Biotech Paper (NBT 2884, 본원에서 참조됨)]은 본 발명자가 이미 제시한 생체내 표현형 변화에 대한 추가 뒷받침을 제공한다.

[0262] 본 발명자가 이후에 제공한 추가 데이터는 Cas9를 통한 체세포 간 조직의 효과적인 생체내 편집을 입증함으로써 추가 뒷받침을 제공한다. 또한, AAV2/8을 통한 전달 및 SaCas9의 사용은 다시 상기 생체내에서의 특정 접근법의 유용성을 나타낸다. 바람직한 ApoB가 다시 표적화되었다.

[0263] 이후의 실시예 36 및 37은 생체내에서 표현형 변화, 특히 지질 대사 유전자인 ApoB의 표현형 변화를 유도하는데 있어서의 효능에 대한 우수한 생체내 데이터를 제시한 한편, 실시예 38은 간이 중요한 예인 유사분열 후 세포에 대한 상기 기술의 적용성을 제시한다. 실시예 39는 다수의 에피토프 태그가 검출 목적을 위해 바람직한 것을 제시한다.

[0264] 바이러스 벡터가 바람직하나, 일부 구현예에서, 세포 투과 펩티드의 사용이 실행 가능한 대안이며, 따라서 이 또한 바람직하다.

[0265] 실시예 36은 유전자형 및 결정적으로 표현형 변화가 CRISPR-Cas 시스템을 이용하여 관찰되는 것을 제시하였다. 그 뿐만이 아니라, CRISPR-Cas9 시스템은 생체내에서 표현형 변화를 유도하는데 효과적이었다. 특히, 표적은 지질 대사 유전자인 ApoB였다. 고무적인 것은 ApoB가 간 전달에서 "최적 표준"인 것으로 언급될 수 있고, 비만의 마우스 모델에서 광범위하게 사용된다는 점이다. 간은 일부 구현예서 바람직한 유사분열 후 세포이나, 이는 또한 다른 구현예에서 배제될 수 있다. 어느 쪽이든, 상기 작업은 표현형 변화가 생체내에서도 관찰되며, 이는 다른 유사분열 후 세포에 동등하게 적용가능하다는 원리의 증거를 제공한다. 또한, 실시예 39는 유사분열 후 뉴런을 갖는 별개의 조직인 뇌에서의 상기의 추가 증거를 제공한다.

[0266] 실시예 37에서의 전달은 정맥내 주사를 통하는 것이었다. AAV 벡터가 사용되었을뿐만 아니라 Cas9에 대해 간-특이적 프로모터(TBG)가 사용되었다.

[0267] 본원에서 관찰된 바와 같은 바이러스 벡터로부터의 발현을 통한 전달은 전달 방법으로서 Anderson/Yin(NBT 2884)의 유체역학적 전달의 이용에 비해 개선인데, 이는 유체역학적 전달이 주사되기 위해 수 ml의 유체를 필요로 하며, 이는 뮤린 신체에 스트레스를 주고 치명적일 수 있기 때문이다. 유체역학적 전달은 플라스미드(네이키드) DNA의 전달에 가장 적합한 반면, 출원인은 바이러스 전달 벡터 내에 가이드 및 Cas9 서열을 패키징하는 것이 크게 증가된 효율과 관련하여 바람직한 것을 제시하였다. 또한, 단지 상대적으로 적은 부피가 도입되는 것이 필요하며, 이는 치료적으로 훨씬 더 용인될 수 있는 정맥내(i.v.)로 수행될 수 있다.

[0268] 특히 고무적인 것은 ApoB와 같은 간에 대한 "최적 표준" 유전자에서 유전자형 변화가 관찰되었을뿐만 아니라 표현형 변화도 또한 기록되었다는 점이다. PCSK9를 이용한 이전의 작업은 유전자형 변화를 제시하였으나, 표현형 변화는 제시하지 않았고, 따라서 ApoB로 관찰된 표현형 변화는 간으로의 CRISPR 전달의 개연성 및 간에서 표현형 변화를 초래하는 이의 능력을 확인한다. 이는 더욱 치료적으로 허용되는 전달 수단(유체역학적 전달에 비해 i.v.)과 조합된다. 이와 같이, CRISPR-Cas9 시스템(가이드 및 Cas9)의 바이러스 전달, 특히 정맥내가 바람직하다.

[0269] 잠재적 표적은 PCSK9, HMGCR, APOB, LDLR, ANGPTL3, F8, F9/FIX, AAT, FAH, HPD, TAT, ATP7B, UGT1A1, OTC, ARH를 포함하지만, 이로 제한되는 것은 아니다.

[0270] 따라서, 표적 세포, 예를 들어, 간으로 CRISPR-Cas9 시스템을 투여하는 것을 포함하는 생체내에서 표현형 변화를 유도하는 방법이 제공된다. 적합한 전달 경로가 본원에 기재되나, i.v. 주사가 일부 구현예에서 바람직하다. 바이러스 벡터, 특히 AAV, 특히 AAV 혈청형 2/8이 바람직하다.

[0271] 하나 이상의 가이드 표적화 지질 대사 유전자, 예를 들어, ApoB를 포함하는 CRISPR-Cas9 시스템이 또한 제공된다. 상기 CRISPR-Cas9 시스템을 투여하는 것을 포함하는 비만을 치료하는 방법이 또한 예견된다. 특히 지질 대사 유전자(들), 예를 들어, ApoB를 포함하는 하나 이상의 간 유전자 녹다운(들)을 포함하는 마우스 모델이 바람직하다.

[0272] Cas9에 대한 간 특이적 프로모터가 명백할 것이나, 본원에 언급된 것을 포함할 수 있다. 바람직한 예는 TBG이다.

[0273] 실시예 38에 제시된 바와 같이, 가이드는 18 내지 23개의 뉴클레오티드 길이일 수 있다. 이는 18 내지 22개, 또는 19 내지 22개, 또는 18 내지 21개, 20 내지 22개, 바람직하게는 22개, 가장 바람직하게는 21개의 뉴클레오티드 길이일 수 있다.

[0274] SaCas9 인트론으로의 가이드 서열의 성공적인 패키징의 원리의 증거가 또한 제공된다. 따라서, 하나 이상의 가이드 서열이 Cas9 인트론으로 패키징(위치결정 또는 삽입됨)된 CRISPR-Cas9 시스템이 바람직하다.

[0275] H1 프로모터가 사용될 수 있고, 일부 상황에서 바람직할 수 있다.

[0276] 란에 의한 작업(Cell, 154, 21 Aug 2013)으로부터 확장하여, D10A 이중-닉카아제를 이용한 이중 가이드 접근법에서의 중첩의 정도가 연구되었다. 최적 결과가 -5와 +1 bp 사이(5' 대 5')에 제시되었다. 따라서, 표적외 효과를 최소화시키는 이중 가이드 접근법을 이용하는 것이 바람직하다. 이들은 바람직하게는 게놈 표적에서 DNA의 상이한 가닥 상의 이들의 5' 말단에서 중첩되거나, 중첩에 가까워진다. 바람직하게는, 중첩은 -5 내지 +1 bp 범위이다. 이들 예에서, Cas9가 이중 닉카아제, 예를 들어, 바람직한 D10A 변이체인 것이 인지될 것이다.

[0277] 다수 또는 반복 에피토프 태그가 Cas9에 대해 바람직하다. 특히, 검출을 개선시키기 위해 실시예 39에서 삼중 에피토프가 제시되었다. 상기 태그는 바람직하게는 반복부이고, 더욱 바람직하게는 삼중 반복부이다. HA가 바람직한 Cas9 에피토프 태그이다. 따라서, 삼중 HA 에피토프 태그가 일부 구현예에서 바람직하다.

[0278] 실시예 39는 하기 특정한 포인트를 제공한다. 이는 하기를 제공한다:

[0279] - 생체내에서 성공적인 AAV-매개 Cas9 전달뿐만 아니라 유사분열 후 뉴런에서의 효과적인 게놈 변형의 첫 번째 증명;

[0280] - Cas9 및 sgRNA-발현 세포로부터의 뉴런 핵의 용이한 분리를 가능케 하는 핵 태깅 기술의 개발을 위함;

[0281] - 뉴런 전사체의 RNAseq 분석에 대한 적용성의 입증;

[0282] - 표현형 변화를 결정하기 위해 전기생리학적 연구 및 다른 기술이 Cas9-매개 게놈 교란과 결합될 수 있는 방법; 및

[0283] - Cas9-매개 게놈 편집을 이용한 설치류 행동에 대한 유전자 기능을 연구하는 능력 및 다중복합체 표적화의 증명.

[0284] 상기를 기초로 하여, 실시예 39가 하기의 2개의 주요 영역, 즉 모델의 생성 및 시험을 포함하는 유전자 기능의 이해 및 시험, 및 유전자 요법에서의 개념의 추가 증거를 제공하는 것이 관찰될 수 있다.

[0285] 하기에 추가로 논의되는 추가 양태는 핵 태깅을 위한 방법과 관련된다.

[0286] 본원에서 CRISPR-Cas9 시스템에 대한 언급은 하나 이상의 게놈 서열을 표적화하기 위해 이용된 가이드 또는 가이드들과 조합된 본원에 제공된 Cas9 효소 언급에 대한 속기임이 인지될 것이다. 가이드(들)에 대한 언급은 sgRNA뿐만 아니라 대상체의 게놈에서 표적 서열에 혼성화할 수 있는 가이드 서열, tracr 메이트 서열 및 tracr 서열을 포함하는 본원에 기재된 키메라 폴리뉴클레오티드 서열을 포함한다.

[0287] 데이터는 본질적으로 본 발명에 따른 2개의 독립된 CRISPR-Cas9 시스템(Cas9 효소와 조합된 가이드 RNA)을 이용한 유전자 녹다운으로부터 발생된 표현형 변화를 제시하며, 이러한 경우 유전자 기능을 성공적으로 교란시킨다. 선택된 조직은 뇌 조직이나, 결과는 광범위한 유사분열 후 조직에 대한 원리의 증거를 제공한다. 이는 중요한 차별성인데, 이는 이전의 연구가 분열하는 세포(즉, 유사분열 전)에 집중하였기 때문이다.

[0288] 즉, SpCas9가 분열하는 세포를 유전자 조작하기 위해 널리 사용된 반면, 출원인은 유사분열 후 뉴런의 게놈을 유전자 조작하기 위해 SpCas9가 또한 이용될 수 있음을 증명한다. 이는 녹다운을 발생시키기 위해 NHEJ-매개 삽입-결실 발생을 통해 고효율로 수행되나, HDR 메커니즘을 통한 교정(수선 주형의 제공시)을 포함하는 치료 용도

가 또한 예견된다. 둘 모두는 본원에 제시된 Cas9 및 RNA 가이드 또는 가이드들의 성공적인 전달 및 기능적 발현에 의존한다.

[0289] CRISPR-Cas9 시스템에 의해 유도된 유전자형 변화가 이후에 표현형 변화를 초래한다는 사실이 또한 상기 영역 둘 모두(유전자 기능 및 유전자 요법)에 중요하다.

[0290] 첫 번째 CRISPR-Cas9 시스템은 *Mecp2*에 지정된(표적화하는) 가이드 서열을 이용하였다. 하나의 벡터가 가이드를 포함하고, 하나가 Cas9를 포함하는 이중 벡터 CRISPR-Cas9 시스템이 성공적으로 이용되었고, 이는 상기 이중 벡터 시스템에 대한 원리의 추가 증거를 제시한다. 이중 벡터 CRISPR-Cas9 시스템은 정위 주사를 통해 뇌의 2개의 독립적 위치, 특히 해마 치아이랑 및 시각피질로 성공적으로 전달되었다. 둘 모두의 경우에서, 동일 유전자 *Mecp2*의 유전자 교란이 관찰되었고, 이는 Cas9 효소(이 경우, SpCas9)에서의 전사 및 기능적 활성, 및 가이드 서열에 의한 게놈 표적 서열로의 Cas9의 성공적인 보충과 함께 이중 벡터 시스템이 예상된 바와 같이 성공적으로 전달되고 작동된 것을 나타낸다.

[0291] SpCas9 및 sgRNA의 AAV-매개 생체내 전달은 무결한 신경 회로 내에서 정확한 게놈 교란을 달성하기 위한 신속하고 강력한 기술을 제공한다. 이와 같이, 사용된 벡터는 AAV 벡터였으며, 일반적으로 이들의 용도 및 특히 이중 벡터 CRISPR-Cas9 시스템, 특히 유사분열 후 세포 및 조직, 및 특히 뇌에서의 이들의 용도에 대한 추가 증거를 더한다.

[0292] 프로모터의 선택이 CRISPR-Cas9 시스템, 특히 Cas9 또는 가이드(들) 및 Cas9 둘 모두로부터의 발현을 달성하는 데 중요한 것이 물론 인지될 것이다. 세포 및 세포 생활주기 단계 특이성에 대한 적합한 예가 문헌으로부터 결정될 수 있다. 그럼에도 불구하고, 일부 비제한적인 예는 간-특이적 프로모터인 TBG를 포함하며, SaCas9, H1 프로모터; 트렁케이션된 H1 프로모터; U6 프로모터의 발현을 유도하기 위해 본원에서 사용된다. 또한, 가이드가 반드시 특이적 프로모터를 필요로 하지 않음에 따라, 하나 이상의 가이드가 Cas9 인트론으로 유사하게 패키징될 수 있다.

[0293] 이용된 두 번째 CRISPR-Cas9 시스템은 다중복합체 접근법을 포함하였다. SpCas9 시스템의 한 주요 장점은 다중복합체 게놈 편집을 돕는 이의 능력이다. 이러한 두 번째 시스템은 적합한 가이드를 포함시킴으로써 동일 패밀리로부터의 3개 이상의 유전자(이 경우, *Dmnt1, 3a* 및 *3b*)를 성공적으로 표적화하였고, 다수의 유전자의 안정적인 녹아웃을 발생시켰다. 이는 살아있는 동물의 조직에서 개별적 유전자뿐만 아니라 전체 유전자 패밀리의 기능을 탐지하기 위한 광범위한 관련성을 갖는다. 이는 이러한 것이 이전에 가능하지 않았거나, 긴 햇수의 고전적 유전학을 통해서만 달성될 수 있었던 뇌와 같은 조직에 특히 중요하다. 본 출원인은 단일 또는 다수의 유전자 교란(심지어 완전한 녹다운)이 정상 동물에서 유사분열 후 세포에서 발생할 수 있음을 제시하였다. 그러나, 이는 모델 유기체(예를 들어, 유전자 돌연변이 또는 교란을 이미 갖거나, 일부 종류의 변경된 발현을 포함하는 유기체) 또는 트랜스제닉 유기체에 동등하게 적용될 수 있어, 유전자 기능을 이해하기 위해 모델 유기체를 발생시키고 모델 유기체를 이용하는 현존하는 방법에 신속한 대안을 더할 수 있다. 추가 가이드(및/또는 전체 CRISPR-Cas9 시스템)가 동일 유기체 내에서 이후의 라운드의 유전자 교란 및/또는 복원(예를 들어, 수선 주형, 예를 들어, HDR에 적합한 ssDNA의 제공을 통한, 예를 들어, 교란된 유전자의 교정에 의한 유전자 기능 복구)을 이루기 위해 이용될 수 있다.

[0294] 사실, 일반적으로, 단일 또는 다수 유전자의 SpCas9-매개 표적화는 고전적인 더욱 시간 소모적인 유전학적 마우스 모델을 이용하여 관찰된 형태적, 전기생리학적 및 행동 표현형을 발생 반복시킬 수 있다.

[0295] 전체 유전자 패밀리 또는 관련 유전자를 녹다운시키는 것에 대해 대안적으로, 본원의 데이터는 또한 동시 녹다운 또는 3개 이상의 관련되지 않은 유전자가 동등하게 실행가능한 원리의 증거를 제공한다. 이는 모든 조직에 걸쳐 적용가능하나, 유사분열 후 조직, 특히 뇌와 관련하여 특히 강하게 제시된다.

[0296] 상기 작업의 또 다른 유용한 양태는 유전자형 변화를 수행하기 위해 CRISPR을 이용한 후, 존재시 CRISPR-Cas9 시스템에 의해 유도된 유전자형 변화로부터 표현형 변화가 발생한 것을 확립시키기 위해 전기생리학(특히, 뇌 및 CNS 조직과 관련됨), 생화학적, 시퀀싱, 전기생리학적 및/또는 행동 판독과 같은 고전적 도구를 이용한 유전자 기능을 연구하기 위해 조합되거나 통합된 접근법이 취해질 수 있음을 나타내었다. 예를 들어, 뇌에서, 이는 개별적 유전자뿐만 아니라 신경 과정의 유전자의 그룹의 기능 및 생체내에서 뇌 장애에서의 이들의 역할을 연구하는 것을 가능케 한다.

[0297] 이러한 연구에서 유전자의 성공적인 교란은 유전자 기능의 교정 또는 복원, 즉, 유전자 요법에서 CRISPR-Cas9 시스템의 사용에 동등하게 적용가능하다. 이는 유사분열 후 세포, 특히 뇌를 표적화하는 것과 특히 관련된다.

[0298] 일반적으로, CRISPR-Cas9 시스템의 사용은 설계하고 생성하는데 오랜 기간이 걸리고 다중복합체일 수 없는 Zn 핑거 및 너무 많은 표적외 효과를 갖는 shRNA와 같은 현존하는 기술에 비해 개선을 나타내는 반면, CRISPR의 표적외 효과는 이중-닉카아제 접근법을 이용함으로써 최소화될 수 있다.

[0299] **조직의 표적화**

[0300] 본원의 작업은 적절한 위치(즉, 관심대상의 기관 또는 조직 내의 세포)로의 CRISPR-Cas9 시스템의 전달을 통해 유사분열 후 세포 내의 표적 유전자에 대한 CRISPR-Cas9 시스템의 사용을 뒷받침한다. 바람직한 조직은 하기 기관 내에 존재한다:

[0301] · 신장;

[0302] · 위, 췌장, 십이지장, 회장 및/또는 결장을 포함하는 소화계통;

[0303] · 심장;

[0304] · 폐;

[0305] · 뇌, 특히 뉴런, 및/또는 일반적으로 CNS;

[0306] · 망막 조직을 포함하는 눈;

[0307] · 내이를 포함하는 귀;

[0308] · 피부;

[0309] · 근육;

[0310] · 뼈; 및/또는

[0311] · 일반적으로, 간.

[0312] 상기 중 많은 것이 유사분열 전 세포를 포함할 수 있으나, 본 발명의 상기 양태는 상기 기관 내의 유사분열 후 세포 또는 조직에 관한 것이 인지될 것이다.

[0313] 특히, 출원인은 상기 기관은 신장 또는 뇌인 것을 선호한다. 뇌 내에서, 데이터는 바람직한 조직인 해마 치아이랑 및 시각피질로의 전달을 특별히 제시하나, 일차 운동 피질, 일차 청각 피질, 일차 체성감각 피질, 소뇌, 주요 후각 망울, 전두엽 피질, 내조롱박핵(endopiriform nucleus), 편도, 흑색질, 선조체, 창백핵(pallidum), 시상, 시상 하부, 팔곁핵, 상올리브복합체, 달팽이핵, 유두핵 중 어느 하나 이상을 포함하는 다른 조직이 또한 일부 구현예에서 바람직하다.

[0314] 뇌로부터의 세포, 특히 뉴런이 특히 바람직하다.

[0315] CRISPR-Cas9 시스템, 특히 Cas9의 발현을 유도하는 프로모터의 선택이 상기 언급된 바와 같이 중요하다. 프로모터를 선택하는 경우에 고려되어야 하는 것은 세포 주기 단계(초기/후기) 및 세포 유형인데, 이는 프로모터가 세포 유형 및 세포-주기 단계 중 하나 이상에 특이적일 것이기 때문이다. 적합한 프로모터는 일부 구현예에서 하기 중 어느 하나 이상을 포함할 수 있다:

[0316]

세포 유형	프로모터
흥분성 뉴런	CamkII
신속 스파이킹 개재뉴런	파르브알부민
모든 개재뉴런	vGAT
도파민수용성 뉴런	DR1
도파민수용성 뉴런	DR2
성상교세포	GFAP
활성화된 뉴런	Arc

[0317] 뇌, 특히 해마 치아이랑을 표적화하는데 사용되는 이중 벡터 CRISPR-Cas9 시스템은 2개의 독립된 바이러스 벡터 상에 SpCas9 및 sgRNA 발현 카세트를 패키징하였다. 따라서, Cas9s, 특히 SpCAs9s는 바람직하게는 아데노바이러스 벡터, 특히 AAV(즉, AAV-SpCas9로서)에 의해 전달된다. 가이드는 바람직하게는 아데노바이러스 벡터, 특히

AAV에 의한 sgRNA 발현 카세트(즉, AAV-SpGuide)로서 전달된다. 상기 조직(해마 치아이랑) 및 뇌에 대한 바람직한 경로는 일반적으로 정위 주사이다.

[0318] **유전자 기능의 이해 및 시험, 및 이를 위한 모델의 생성 및 이용**

[0319] 고찰될 수 있는 질환은 헌팅던(Huntingdon) 질환을 포함하나, 본질적으로 유사분열 후 세포에서 발견되는 임의의 질환, 특히 생체내 연구로부터 이로울 수 있거나 유용한 모델이 결핍된 질환을 포함한다.

[0320] 상기에 언급된 바와 같이, 유사분열 후 세포에서 하나 이상의 유전자의 기능을 질의하기 위해 CRISPR-Cas9 시스템이 이용될 수 있다. 이는 유사분열 후 세포로의 CRISPR-Cas9 시스템의 전달 및 발현을 통해 달성될 수 있으며, 여기서 CRISPR-Cas9 시스템의 가이드(들)은 관심대상의 게놈 표적 또는 표적들로 Cas9를 보충하도록 설계된다. 동등하게, Cas9가 이미 유사분열 후 세포 내에 포함된 경우, 단백질이 (전사되어) 형성된 후, 유사분열 후 세포로의 가이드의 전달이 충분할 것이다. Cas9가 이미 뉴클레오티드 내(전사되지 않음)에 유사분열 후 세포에 포함된 경우, 유사분열 후 세포로의 가이드의 전달뿐만 아니라 Cas9 폴리뉴클레오티드의 전사의 유도가 필요할 것이다. Cas9가 유도성 또는 억제성 프로모터의 조절하에 있는 경우, tet(테트라사이클린) 온-오프 시스템이 본원에서 유리할 수 있다.

[0321] 특히 가망 있는 한 양태는 존재시 유전자 교란, 특히 녹다운으로부터 발생하는 표현형 변화를 결정하기 위한 표현형 검정과 CRISPR 기술의 통합이다. 예를 들어, 실시예 39는 특정 게놈 구성요소의 생물학적 기능에 대한 통찰을 제공하기 위해 정량적 판독과 커플링된 표적화된 게놈 교란으로 달성될 수 있는 것을 제시한다. 특히, 뇌에서의 Cas9-매개 생체내 게놈 편집은 또한 특정 세포 유형 또는 회로 구성요소에 대한 게놈 교란의 효과를 연구하기 위해 전기생리학적 기록과 커플링될 수 있다. 더 넓은 의미로, CRISPR-Cas9 시스템(Cas9-매개 게놈 교란)의 사용은 표적화된 게놈 구성요소의 기능을 연구하기 위해 생화학적, 시퀀싱, 전기생리학적, 및 행동 분석과 조합될 수 있다.

[0322] 따라서, 한 양태에서, 하기 기재되는 바와 같이 결함이 있는 유전자형 또는 유전자 녹다운을 유도하는 단계, 및 질환에서 하나 이상의 유전자의 발현에서의 변화를 결정함으로써 하나 이상의 유전자의 기능을 질의하는 단계를 포함하는, 유사분열 후 세포에서 하나 이상의 유전자의 기능을 질의하는 방법이 제공된다.

[0323] 선택적으로, 상기 방법은 또한 대상체로 세포의 제2 집단을 이식함으로써 결함이 있는 유전자형 또는 유전자 녹다운과 관련된 질환을 유도하는 단계를 포함할 수 있다. 이는 결정 단계 전일 수 있다.

[0324] 하기는 본 발명의 적절한 양태에 널리 적용된다. 세포는 유전자 기능이 생체내에서 질의되는 대상체, 예를 들어, 인간, 동물 또는 모델 유기체 내의 세포일 수 있다. 그러나, 세포는 생체외, 예를 들어, 세포 배양물 또는 모델 기관 또는 오가노이드에 존재할 수 있음이 또한 예견된다. 일부 구현예에서, 상기 방법은 대상체로부터의 세포를 선택적으로 배양하고, 이들을 하나 이상의 CRISPR-Cas9 시스템으로 형질도입시키는 단계를 포함하는, 대상체로부터의 세포의 제1 집단의 분리를 포함할 수 있다. 추가의 임의의 배양이 후속될 수 있다. 다시 대상체로의 형질도입된 세포의 이식이 이후에 발생할 수 있다.

[0325] 세포는 본원에 기재된 조직 또는 기관 중 임의의 것으로부터 유래될 수 있다. 뇌는 하나 이상의 유전자의 기능을 질의하는 상기 방법을 제공하는 한 바람직한 예이며, 여기서 유사분열 후 세포는 뇌 세포, 예를 들어, 뉴런이다. 특히 생체내에서, 이는 동물 행동에 대한 유전자 기능의 질의를 가능케 한다. 동물은 바람직하게는 포유동물, 예를 들어, 설치류이다. 이종성 세포 유형의 복잡한 네트워크로 구성된 신경계의 복잡성이 존재하는 경우, 생체내에서 뉴런의 게놈을 효과적으로 편집할 수 있는 것은 자연 상황에서 엠베딩된 관련 세포 유형에서의 유전자 기능의 직접적인 시험을 가능케 한다. 이는 녹아웃 마우스가 훈련된 상황 조건하에서 시험되는 경우 손상된 기억 경화를 나타낸 출원인의 데이터에 의해 뒷받침된다. 출원인의 결과는 치아이랑 뉴런에서의 DNMT 패밀리 일원의 CRIPSR-Cas9-매개 녹아웃이 행동 임무에서 유전자의 기능을 탐지하기에 충분한 것을 입증한다.

[0326] 이는 유전자 기능을 연구하기 위해, 특히 신경 회로의 해부를 위해 생체내에서 포유류 뇌에서의 표적화된 유전자 녹아웃을 촉진하는데 있어서의 Cas9s의 다능성을 나타낸다. 살아 있는 동물의 뇌에서 다수의 유전자의 안정적인 녹아웃을 유도하는 것은 생리학적 및 신경병리학적 질환에서 다인자 메커니즘의 원인적 질의와 같이 잠재적으로 광범위한 적용을 가질 것이다.

[0327] 상기 작업의 특성은 출원인이 배양된 일차 마우스 피질 뉴런에서 충분한 수준의 SpCas9 발현을 달성하는 능력을 기초로 하여 마우스 Mecp2 프로모터(235 bp, pMecp2)7 및 최소의 폴리아데닐화 신호(48 bp, spA)를 선택한다는 점이다. Mecp2 유전자는 자폐증 스펙트럼 장애의 한 유형인 레트 증후군에서 주요한 역할을 한다. Mecp2를 표적화하기 위해, 출원인은 먼저 마우스 Mecp2 유전자의 엑손 3를 표적으로 하는 여러 sgRNA를 설계하였고, Neuro-

2a 세포를 이용하여 이들의 효능을 평가하였다. 가장 효과적인 sgRNA가 SURVEYOR 뉴클레아제 검정을 이용하여 확인되었다. 전달은 고역가 AAV-SpCas9 및 AAV-SpGuide의 혼합물(1:1 비)의 정위 주사를 통하였다. 출원인은 또한 뇌에서의 다중복합체 게놈 편집의 가능성을 성공적으로 시험하였다. 본 출원인은 핵 표지화를 위한 GFP-KASH와 함께 연쇄적인 3개의 sgRNA로 구성된 다중복합체 sgRNA 발현 벡터를 설계하였다.

[0328] 따라서, 유사분열 후 세포에서 하나 이상의 유전자 넉다운을 특징으로 하는 질환을 유도하는 방법이 또한 제공된다. 상기 질환의 예는 많지만, 예로서 레트 증후군을 포함할 수 있다. 적합한 프로모터는 명백할 것이나, 레트 증후군에 대해 Mecp2 프로모터가 이상적이다. CRISPR-Cas9 시스템, 특히 Cas9의 발현을 유도하는 프로모터를 선택하는 한 방식은 관심대상 유전자에 대한 프로모터를 선택하는 것이다.

[0329] 따라서, 한 양태에서, 유사분열 후 세포에서 하나 이상의 결함이 있는 유전자(또는 유전자형) 또는 유전자 녹다운을 특징으로 하는 질환을 유도하는 방법이 제공되며, 상기 방법은

[0330] I. CRISPR-Cas 시스템 키메라 RNA(chiRNA) 폴리뉴클레오티드 서열에 대해 작동 가능하게 연결되는 제1 조절 구성요소로서, 폴리뉴클레오티드 서열은

[0331] (a) 대상체의 게놈 내의 3개 이상의 표적 서열에 혼성화할 수 있는 1개, 2개, 3개, 4개 이상의 가이드 서열,

[0332] (b) tracr 메이트 서열, 및

[0333] (c) tracr 서열을 포함하는, 제1 조절 구성요소, 및

[0334] II. 적어도 하나 이상의 핵 국소화 서열(NLSs)을 포함하는 CRISPR 효소를 암호화하는 효소 암호화 서열에 작동 가능하게 연결된 제2 조절 구성요소

[0335] ((a), (b) 및 (c)는 5'에서 3' 방향으로 배열되고,

[0336] 구성성분 I 및 II는 시스템의 동일 또는 상이한 벡터 상에 위치되며, 전사될 때, tracr 메이트 서열은 tracr 서열에 혼성화하며, 가이드 서열은 표적 서열에 대한 CRISPR 복합체의 서열-특이적 결합을 지시하며,

[0337] CRISPR 복합체는 (1) 표적 서열에 혼성화되는 가이드 서열 및 (2) tracr 서열에 혼성화된 tracr 메이트 서열과 복합체화되는 CRISPR 효소를 포함하고,

[0338] CRISPR 효소는 세포의 제1 집단의 게놈을 변화시켜 하나 이상의 결함이 있는 유전자 또는 녹다운된 유전자를 갖는 세포의 제2 집단이 수득됨)

[0339] 을 포함하는 하나 이상의 벡터를 포함하는 벡터 시스템을 포함하는 비자연적으로 생기는 또는 유전자 조작된 조성물로 세포의 제1 집단을 형질도입시키는 단계를 포함할 수 있다.

[0340] 선택적으로, 상기 방법은 또한 대상체로부터 세포의 제1 집단을 분리시키는 단계를 포함할 수 있다.

[0341] 선택적으로, 상기 방법은 또한 대상체로 세포의 제2 집단을 이식함으로써 증식 질환을 유도하는 단계를 포함할 수 있다.

[0342] 이는 표적 세포로 비-기능성(부분적 비-기능성을 포함함) 유전자형을 유도함으로써 연구를 위한 모델(기능성 유전자형의 이후의 복원을 포함함)을 제공하는 단계를 포함할 수 있다.

[0343] CRISPR-Cas9 시스템은 또한 유사분열 후 뉴런에서 표적화된 녹아웃을 가능케 함으로써 세포 검정에서 유전자 기능의 연구를 돕기 위해 이용될 수 있다.

[0344] 뉴런 세포로 뉴클레오티드를 전달하기 위한 방법은 널리 공지되어 있고, 문헌[The Journal of Neuroscience, by Karra and Dahm (5 May 2010, 30(18): 6171-6177; doi: 10.1523/JNEUROSCI.0183-10.2010)]에 개관되어 있다. 예로는 전기적 트랜스펙션 방법(예를 들어, 전기천공법, 뉴클레오펙션, 및 단일-세포 전기천공법); 화학적 트랜스펙션 방법(예를 들어, Ca2+ 포스페이트 공동/침전법 및 리포펙션); 바이러스 전달(예를 들어, 아데노바이러스, 아데노-연관 바이러스(AAV), 렌티바이러스 및 단순 포진 바이러스); 및 물리적 트랜스펙션 방법(예를 들어, 마이크로인젝션 및 비올리스틱스(biolistics)(DNA-코팅된 금 입자))을 포함한다. 이들 모두가 CRISPR-Cas9 시스템의 전달에 이용될 수 있으나, 리포펙션 또는 바이러스 방법, 특히 AAV 또는 렌티바이러스가 바람직하다.

[0345] **모델**

[0346] 단일 또는 다수의 유전자가 녹다운된 모델이 제공된다. 예로는 Mecp2 녹다운인 레트 증후군에 대한 설치류 모델

이 있다. 다른 예는 Dmnt 패밀리 녹다운, 특히 Dmnt1, 3a 및 3b 녹다운을 포함한다. 이와 같이, 신경학적 질환을 연구하는 모델이 제공된다. 수행될 필요가 있는 모두는 관심대상의 표적 유전자를 확인하고, 적합한 가이드(들)를 설계하고, 이들을 적합한 CRISPR-Cas9 시스템에 포함시키고, 이를 필요에 따라 생체내이거나 생체외이건 간에 유사분열 후 세포(들)로 전달하는 것이다. 예를 들어, 상기 모델은 변경된 가지돌기 나무 형태를 가질 수 있고/있거나, 가시 밀도가 제공된다.

[0347] 상기 언급된 바와 같이, 모델 조직, 예를 들어, 오가노이드 또는 "리버 온 어 칩" 또는 이의 비-간 동등물, 예를 들어, 귀, 신장 및 뇌 조직, 예를 들어, 칩 상에 존재하거나 또는 스캐폴드에서 지지된 귀, 신장 및 뇌 조직이 또한 제공된다. 동물 모델 및 모델 조직이 바람직하다. 이들은 이들이 상기 언급된 바와 같이 뉴클레오티드 또는 단백질 형태 내에 Cas9를 포함하므로 Cas9로 이미 형질전환될 수 있다. 이들은 Cas9가 가이드(들)와 접하여 전달될 필요가 없는 장점을 가지며, 이는 차례로 훨씬 더 많은 정도의 (가이드) 다중복합체화가 전달 벡터 내에 수용되는 것을 가능케 할 수 있다. 또한, 유도성 또는 억제성 시스템, 예를 들어, tet-온 또는 tet-오프의 사용이 본원에서 유리할 수 있다.

[0348] 이들 모델 모두는 상기 기재된 바와 같은 CRISPR-Cas9 시스템을 이용하여 수득 가능하다. CRISPR-Cas9 시스템의 다능성으로 인해, 인간이거나, 설치류이거나, 포유류이거나, 그 밖의 다른 것이건 간에 가능한 모델의 범위가 매우 다양하고, 이는 적절한 가이드(들)의 간단한 선택에 의해 확립될 수 있다. 하기를 포함하는 상기 모델을 생성시키는 방법이 또한 제공된다.

[0349] **유전자 요법**

[0350] 실시예 39의 데이터는 유전자 교란, 주로 녹다운에 집중한다. 유전자 녹다운은, 중요한 경우, 유전자 요법으로의 CRISPR-Cas9 시스템의 가능한 적용의 전체 쿼럼(total quorum)의 단지 적은 부분일 가능성이 있다. Yin 및 Anderson의 논문(Nature Biotech 2884 published online 30 March 2014)에 이미 제시된 바와 같이, 스플라이싱 동안 엑손 8의 스킵핑(skipping)을 야기하여, 트렁케이션된 불안정한 FAH 단백질의 형성을 초래하여, 독성 대사물의 축적을 야기하는 푸마릴아세토아세테이트 하이드롤라제(FAH)의 돌연변이(엑손 8의 마지막 뉴클레오티드에서 G에서 A로 돌연변이됨)에 의해 야기된 달리 치명적인 질환인 유전성 티로신혈증 타입 I(HTI)에서의 결함이 있는 돌연변이의 교정 후에 기능적 표현형이 복원될 수 있다. A 돌연변이의 다시 야생형 G 표현형으로의 교정은 복원된 표현형을 초래한다.

[0351] 이와 같이, 본원에서 취해진 접근법은 본 발명이 유전자 요법에 개연성 있게 적용될 수 있음을 증명한다. 특히, 이중 벡터 접근법, 핵 태깅 접근법, 뇌 전달의 특성(주사 형태, 사용되는 프로모터 및/또는 바이러스 벡터)뿐만 아니라 다중복합체화(동일한 유전자 내 또는 상이한 유전자 내에서의 다수의 표적에 대한 다수의 가이드의 사용)는 예시된 유전자 녹다운에 대해서와 같이 교정 유전자 요법(즉, 결함이 있는 유전자형이 교정됨)에 동등하게 적용될 수 있다. 교정 유전자 요법과 유전자 녹다운 사이의 주요 차이는 결함이 있는 유전자형, 예를 들어, 점돌연변이(예를 들어, 낭성 섬유증에서의 점돌연변이, 참고자료[Schwank *et al*, Cell Stem Cell 13, 653-658 5 Dec 2013] 참조)를 교정하기 위해, HDR 메커니즘을 자극하고, 적합한 Cas9 닉카아제를 또한 이상적으로 제공하기 위해 수선 주형을 제공하는 것이 유리하다는 점이다.

[0352] 따라서, 본 발명의 벡터는 바람직하게는 유사분열 후 세포를 표적으로 한다. 가이드 또는 가이드들이 결함이 있는 유전자형을 표적으로 하는 경우, 바람직하게는 수선 주형, 예를 들어, 교정된 서열에 상응하는 ssDNA(기능적 표현형을 제공하는 유전자형)가 또한 제공된다. 수선 주형은 본원에 기재된다. Cas9가 가이드 또는 가이드들로부터 동일하거나 상이한 벡터에 제공될 수 있다. 벡터는 바람직하게는 바이러스 벡터, 더욱 바람직하게는 아데노바이러스 벡터, 가장 바람직하게는 AAV 벡터이다. 세포로의 전달은 바람직하게는 적절한 경우 정맥내 주사 또는 정위 주사에 의한 것이다. 프로모터의 선택은 또한 중요할 수 있고, 유리한 예가 본원에 제공된다.

[0353] 적절한 세포로의 CRISPR-Cas9 시스템의 전달을 포함하는, 유사분열 후 세포에서 결함이 있는 유전자형에 의해 야기되거나 이와 관련된 유전 질병 또는 질환을 치료하는 방법이 제공된다. 결함이 있는 유전자형은 비-야생형 유전자형일 수 있다. 특히, 단일 점돌연변이 및/또는 단일유전자 장애가 CRISPR-Cas9 시스템을 이용한 치료에 특히 적합하다. 다수의 유전자가 편집 또는 교정을 필요로 하는 경우, 이들을 모두 동시에 표적화하기 위해 다중복합체 접근법이 이용될 수 있다. 대안적으로, 2개 이상의 라운드의 상이한 CRISPR-Cas9 시스템이 예견될 수 있다. 바람직하게는, 야생형 유전자형이 교정된다. 표현형에서 기능이 복원되거나 개선되는 한 반드시 가장 일반적인 유전자형일 필요는 없다.

[0354] 복원된 표현형의 예는 VGLUT3 기능을 복원시키기 위한 청력 및 이로부터 마우스의 내이에서의 청력의 복원이다

(Omar Akil, Rebecca P. Seal, Kevin Burke, Chuansong Wang, Aurash Alemi, Matthew During, Robert H. Edwards, Lawrence R. Lustig. Restoration of Hearing in the VGLUT3 Knockout Mouse Using Virally Mediated Gene Therapy.*Neuron*, 2012; 75 (2): 283 DOI: 10.1016/j.neuron.2012.05.019). 이는 VGLUT3 자체의 AAV-매개 전달을 이용하는 것이었으나, 내이의 세포를 표적화시키고, 비-기능성 VGLUT3 유전자형을 교정하여, 유사한 표현형 결과, 즉, 청력의 복원을 갖도록 하기 위해, 바람직하게는 또한 AAV 벡터의 이용과 함께 CRISPR-Cas9 시스템이 또한 이용될 수 있음이 전적으로 개연성이 있다. 이와 같이, 바람직하게는 AAV 벡터를 이용한 내이로의 CRISPR-Cas9 시스템의 전달이 바람직하며, 이에 따라 청력 손실을 치료한다. 또한, 감각 기관, 예를 들어, 망막을 포함한 눈, 코 및 귀(특히, 내이)에서의 유전자 기능의 복원이 바람직하다.

[0355] 유사분열 후 조직(눈, 귀 등)에서의 장애의 논의를 포함하는 비교적 최근의 개관은 문헌[Kaufmann *et al* (EMBO Mol Med (2013(, 5, p1642-1661)]이다. 이는 유사분열 후 조직에서의 단일유전자 장애의 교정에서의 AAV의 유용성을 입증한다. 여기에는 "낮은 염증 활성과 같은 다른 특징과 함께, 이들은 우수한 안전성 프로파일을 갖고, 따라서 생체내 유전자 요법을 위한 매우 매력적인 도구인 것으로 밝혀졌다. 또한, Glybera®이 직접 근육내 주사를 위한 재조합 AAV이다..."라고 언급되어 있다. 인용과 함께 상기 논문은 표적 조직으로서 망막, 중추신경계, 간, 골격근 및 심근에서의 유전자 요법이 개관되어 있다. 그리고, 인용과 함께, "최초 연구는 프로토타입 AAV 혈청형 2 벡터를 이용하였고, AAV 벡터의 포트폴리오는 추가 혈청형 및 심지어 유전자 조작된 캡시드를 포함하도록 최근에 확장되었다"라고 기재되어 있다. 카우프만(Kaufmann) 및 카우프만에 인용된 문헌은 참조로서 본원에 포함된다.

[0356] **전사체의 RNAseq 분석**

[0357] SpCas9-매개 게놈 교란 및 집단 수준 RNAseq 분석의 조합은 전사 조절을 특성규명하고, 고려 중인 세포에서의 특정 기능 또는 질병 과정에 중요할 수 있는 유전자를 암시하기 위한 방식을 제공한다. 특히, 세포는 뇌, 특히 뉴런으로부터의 세포이다. 신속-작용 기술, 예를 들어, CRISPR-Cas9 시스템이 본래 일시적인 전사체를 연구하기에 유리하다. 이와 같이, 전사체(RNAseq)의 분석에서 본 발명에 따른 CRISPR-Cas9 시스템의 이용이 제공된다.

[0358] **핵 태깅 방법**

[0359] SpCas9-발현 뉴런의 면역형광 확인을 돕기 위해, 출원인은 HA-에피토프 태그(인간 인플루엔자 헤마글루티닌으로부터 유래됨, 발현 벡터에서 널리 사용되는 일반 에피토프 태그)로 SpCas9를 태깅시켰다.

[0360] AAV-SpGuide 벡터에 대해, 출원인은 U6-sgRNA 발현 카세트뿐만 아니라 인간 시냅신 I 프로모터에 의해 유도된 KASH 핵 막횡단 도메인과 융합된 녹색 형광 단백질(GFP)를 패키징하였다. GFP-KASH 융합 단백질은 GFP를 핵 외막으로 지정하고, AAV-SpGuide에 의해 형질도입된 무결함 뉴런 핵의 형광-기반 확인 및 정제를 가능케 한다.

[0361] 따라서, 본 발명의 벡터는 바람직하게는 유사한 방식으로 적합화된다. 따라서, Cas9가 에피토프 태그, 예를 들어, HA-에피토프 태그로 태깅된 벡터가 제공된다. Cas9는 본원에 기재된 Cas9 중 임의의 것, 예를 들어, Sp 또는 SaCas9일 수 있고, 적절하게 태깅되거나 태깅될 수 있는 한 임의의 변이체(예를 들어, D10A 이중 닉카아제 등)일 수 있다.

[0362] 본 발명의 벡터는 또한 가이드 RNA가 발현 카세트 내에 패키징되도록 적합화될 수 있으며, 이는 하기를 포함한다:

[0363] · 리포터 단백질; 및

[0364] · 선택적으로, 가이드 RNA에 대한 적합한 프로모터, 예를 들어, U6;

[0365] 여기서, 리포터 단백질은 적합한 프로모터에 작동 가능하게 연결된 핵 막횡단 도메인과 융합된다.

[0366] 리포터 단백질은 바람직하게는 형광 단백질, 예를 들어, 녹색, 적색 또는 황색 형광 단백질(GFP, RFP, YFP) 중 하나 등이다.

[0367] 핵 막횡단 도메인의 예는 KASH-유사 도메인, Sun2 도메인, LEM 도메인을 포함한다. 일부 바람직한 구현예에서, 핵 막횡단 도메인은 KASH 핵 막횡단 도메인이다. 바람직하게는, 막횡단 도메인에 대한 프로모터는 인간 시냅신 I 프로모터이며;본원에 인용된 문헌을 또한 참조하라.

[0368] 이러한 태깅 접근법은 단일 또는 이중 벡터 시스템 내에서 이용될 수 있으나, 바람직하게는 이중 벡터 시스템 내에서 이용되는데, 이는 단일 벡터 시스템에서는 공간이 제한되고, 별개의 태그의 필요성이 또한 감소되기 때문이다.

[0369] 또한, 이러한 태깅 기술의 각각의 양태는 다른 것과 독립적으로 이용될 수 있어, Cas9의 에피토프 태깅이 단독으로 사용될 수 있거나, 리포터/형광 단백질 카세트 접근법이 단독으로 사용될 수 있거나, 더욱 바람직하게는 둘 모두가 함께 이용될 수 있다.

[0370] 2013년 3월 11일에 처음 제출되고, 2013년 10월 23일에 온라인에 공개된 카나스티 및 앤더슨의 문헌(Kanasty and Anderson)(Nature Materials, Vol 12 Nov 2013)은 RNAi 전달의 유용한 개관이다. RNAi와 CRISPR 가이드 서열 사이의 유사성으로 인해, RNAi와 관련하여 상기 문헌 및 기타 문헌의 교시내용은 출원인의 CRISPR-Cas9 시스템에서 가이드를 전달하는 메커니즘에 대한 정보를 제공한다. 기재된 기술 중 일부가 또한 Cas9의 전달에 또한 적합할 수 있다. 일부 예에서, Cas9와 독립적으로 출원인의 CRISPR-Cas9 시스템의 가이드를 전달하는 것이 유용할 수 있다. 이는 이중-벡터 전달 시스템의 일부일 수 있으며, 상기 벡터는 특별히 바이러스 벡터가 아니라 단순히 임의의 전달 수단으로서 벡터가 가장 광범위한 관점에서 고려된다. Cas9가 바이러스 벡터를 통해 전달될 수 있고, 게놈 표적에 특이적인 가이드가 별개로 전달되는 것이 예견된다. 본원에서 논의된 바와 같이, 가이드는 Cas9와 동일한 벡터 유형, 예를 들어, Cas9가 AAV 벡터 내에서 전달되고, 가이드(들)이 별개의 AAV 벡터 내에서 전달되는 이중-벡터 시스템을 통해 전달될 수 있다. 이는 실질적으로 동시에(즉, 공동-전달) 수행될 수 있으나, 이는 또한 심지어 몇 주 또는 몇 개월 떨어진 별개의 시점에서 수행될 수 있다. 예를 들어, 첫 번째 라운드의 CRISPR-Cas9 시스템이 전달 되었으나, 추가 가이드를 제공하는 것이 이후에 필요한 경우, 표적 세포에서 아마 여전히 기능성인 본래의 Cas9가 재이용될 수 있다. Cas9가 유도성 프로모터의 조절하에 있는 경우, 표적 세포 내에서의 새로운 CAs9의 전사의 유도가 바람직하다. 동등하게, 본원에 제공된 CAs9-발현 모델이 이용되는 경우, 단지 가이드(들)의 전달이 필요하다. 따라서, 가이드(들)의 전달이 Cas9와 별개로 필요한 경우, 이는 RNAi와 훨씬 동일한 방식으로 전달될 수 있다. 이와 같이, Kanasty에 의한 개관은 간에 특히 초점을 둔 적합한 다수의 공지된 접근법을 나타내는데 도움이 되나, 전달 수단은 일반적으로 광범위한 범위의 세포에 적절하다. 예로 하기를 포함한다:

[0371] ·"혈청 지질단백질과 상호작용하고, 이들 지질단백질을 흡수하는 간세포로의 이후의 진입을 달성하는 리포좀 전달 시스템뿐만 아니라 친지질성 분자에 컨쥬게이션된 siRNA";

[0372] ·페길화(PEGylation);

[0373] ·하기와 같은 컨쥬게이트:

[0374] a. RNAi를 전달하여 ApoB를 성공적으로 억제하는 것으로 밝혀진(이에 의해, CRISPR-Cas9 시스템을 통해 ApoB를 표적화하는 것에 대한 출원인의 작업과 교차됨) 동적 폴리컨쥬게이트(DPCs, 10nm 나노입자); 및

[0375] b. 트리안테나리(triantennary) GalNAc 컨쥬게이트

[0376] c. "둘 모두 매우 효과적인" 특히 GalNAc;

[0377] ·다른 나노입자는 하기를 포함한다:

[0378] d. 추가 제형 구성요소, 예를 들어, 아다만틴-PEG(AD-PEG) 및 아다만틴-PEG-트랜스페린(AD-PEG-Tf)을 포함하는 사이클로덱스트린 중합체 나노입자(CDP);

[0379] e. 양이온성 또는 이온화가능한 지질, 차폐 지질, 콜레스테롤 및 내인성 또는 외인성 표적화 리간드를 포함하는 지질 나노입자(LNP). 내인성 표적화 리간드의 예는 RBP 수용체를 발현하는 간 및 췌장 별세포를 표적화하는데 유용한 레티놀 결합 단백질(RBP)이다. 외인성 표적화 리간드의 예는 간세포 상의 아시알로당단백질 수용체를 통해 간을 또한 표적화 하는 GalNac이다. 조합된 접근법이 Anlylam's ALN-VSP에서 관찰된다;

[0380] ·"간 내피에서의 천공술은 직경이 100 내지 200 nm인 분자가 혈류로 확산되고, 간세포 및 다른 간 세포로의 접근을 획득하는 것을 가능케 한다";

[0381] ·GalNAc와 같은 리간드가 만노스 수용체를 발현하는 실질 간 세포, 및 GalNAc 리간드로의 적합한 siRNA의 컨쥬게이션이 PCSK9를 성공적으로 억제하는 것으로 밝혀진 간세포로의 전달에 적합하다;

[0382] ·미리-규정된 3D 구조로 혼성화하도록 설계된 상보성 DNA 단편으로 구성된 올리고뉴클레오티드 나노입자(ONPs). 적합한 3' 오버행 서열을 이용하여, 특정 위치에서도 6개의 siRNA 가닥이 각각의 입자에 부착될 수 있다. 유체역학적 직경은 약 29nm였다.

[0383] 이들 접근법은 CRISPR-Cas9 시스템에 대한 적어도 가이드의 전달을 위해 일부 구현예에서 바람직하다. 특히 바람직한 것은 동적 폴리컨쥬게이트 또는 내인성 표적화 리간드, 예를 들어, 레티놀 결합 단백질 또는 외인성 표

적화 리간드, 예를 들어, GalNac의 사용이다.

[0384] 또 다른 구현예에서, CRISPR-Cas9-매개 게놈 편집은 질환 돌연변이 및/또는 표현형을 교정하기 위해 사용될 수 있다. 상기 CRISPR-Cas9-매개 게놈 편집은 전체가 참조로서 본원에 포함되는, 웹사이트 nature.com/doifinder/10.1038/nbt.2884에서 이용가능한, 2014년 3월 30일에 Nature Biotechnology에 온라인 공개되고 2014년 3월 31일에 온라인 정정된 Hao Yin 등의 "Genome editing with Cas9 in adult mice corrects a disease mutation and phenotype"을 표제로 하는 매뉴스크립트에 예시된 바와 같이 간 및 또는 혜파토사이트에서 질환 돌연변이 및/또는 표현형을 교정하기 위해 사용될 수 있다. 상기 논문은 인간 질환 유전성 티로신혈증의 마우스 모델의 혜파토사이트에서 Fah 돌연변이의 CRISPR-Cas9-매개 교정에 관한 것이다. 유체역학적 주사에 의한 CRISPR-Cas9 시스템의 구성성분의 전달은 약 1/250 간 세포에서 야생형 Fah 단백질의 최초 발현을 초래한 것으로 밝혀졌다. Fah-양성 혜파토사이트의 확장이 체중 감소 표현형을 구제한 것이 추가로 밝혀졌다.

[0385] 본 방법의 이점은 CRISPR 시스템이 표적외 결합 및 그 결과의 부작용을 회피하는 것이다. 이는 표적 DNA에 대한 높은 정도의 서열 특이성을 갖도록 배열된 시스템을 사용하여 달성된다.

[0386] **Cas9**

[0387] Cas9 최적화는 기능을 향상시키거나 또는 새로운 기능을 개발하기 위해 사용될 수 있으며, 키메라 Cas9 단백질을 생성할 수 있다. 출원인이 생성한 예는 실시예 6에 제공된다. 키메라 Cas9 단백질은 상이한 Cas9 상동체로부터의 단편, 예를 들어, 본 명세서에 기재된 Cas9로부터의 2 개 예시적인 키메라 Cas9 단백질을 합함으로써 만들어질 수 있다. 예를 들어, 본 발명자들은 St1Cas9의 N-말단(이 단백질로부터의 단편은 볼드체임)을 SpCas9의 C-말단에 융합시켰다. 키메라 Cas9의 제조이점은, 감소된 독성; 진핵 세포 내 개선된 발현; 향상된 특이성; 단백질의 감소된 분자량, 예를 들어, 상이한 Cas9 상동체로부터의 가장 작은 도메인을 합함으로써; 및/또는 PAM 서열 필요를 변용시킴으로써 더 작은 단백질을 만드는 것 중 임의의 또는 모든 것을 포함한다.

[0388] Cas9는 일반적 DNA 결합 단백질로서 사용될 수 있다. 예를 들어, 실시예 7에 나타내는 바와 같이, DNA 표적의 가닥 둘 다의 절단을 초래하는 2 개의 촉매적 도메인(D10 및 H840)을 돌연변이시킴으로써 일반적 DNA 결합 단백질로서 Cas9를 사용하였다. 표적 유전자좌에서 유전자 전사를 상향조절하기 위해, 본 발명자들은 전사 활성화 도메인(VP64)을 Cas9에 융합시켰다. 다른 전사 활성화 도메인이 공지되어 있다. 실시예 17에서 나타낸 바와 같이, 전사 활성화가 가능하다. 또한 실시예 17에서 나타낸 바와 같이, 유전자 억제(베타-카테닌 유전자의 이 경우에)는 표적 유전자 서열에 결합하는 Cas9 리프레서(DNA-결합 도메인)을 사용하는 것이 가능하고, 따라서 그의 활성을 나타낸다.

[0389] Cas9 및 하나 이상의 가이드 RNA는 아데노 연관 바이러스(AAV), 렌티바이러스, 아데노바이러스 또는 다른 플라스미드 또는 바이러스 벡터 유형을 사용하여, 특히, 예를 들어, 미국 특허 제8,454,972호(아데노바이러스에 대한 제형, 용량), 제8,404,658호(AAV에 대한 제형, 용량) 및 제5,846,946호(DNA 플라스미드에 대한 제형, 용량)로부터의 제형 및 용량 및 렌티바이러스, AAV 및 아데노 바이러스를 수반하는 임상시험에 대한 임상시험 및 간행물로부터의 제형 및 용량을 사용하여 전달될 수 있다. 예를 들어, AAV에 대해, 투여 경로, 제형 및 용량은 미국 특허 제8,454,972호에서와 같고, AAV를 수반하는 임상시험에서와 같을 수 있다. 아데노바이러스에 대해, 투여 경로, 제형 및 용량은 미국 특허 제8,404,658호에서와 같을 수 있고 아데노바이러스를 수반하는 임상시험에서와 같을 수 있다. 플라스미드 전달에 대해, 투여 경로, 제형 및 용량은 미국 특허 제5,846,946호에서와 같고, 플라스미드를 수반하는 임상 연구에서와 같을 수 있다. 용량은 평균 70 kg 개체를 기준으로 하거나 또는 이를 기준으로 추론될 수 있고, 상이한 체중 및 종의 환자, 대상체, 포유류에 대해 조절될 수 있다. 투여 빈도는 환자 또는 대상체의 연령, 성별, 일반적 건강상태, 다른 질환 및 처리될 특정 질환 또는 증상을 포함하는 보통의 인자에 따라서 의학적 또는 수의학적 실행자(예를 들어, 의사, 수의사)의 영역 내에 있다.

[0390] 바이러스 벡터는 관심 대상의 조직 내에 주사될 수 있다. 세포 유형 특이적 게놈 변형에 대해, Cas9의 발현은 세포 유형 특이적 프로모터에 의해 구동될 수 있다. 예를 들어, 간 특이적 발현은 알부민 프로모터를 사용할 수 있고, 뉴런 특이적 발현은 시냅신 I 프로모터를 사용할 수 있다.

[0391] **트랜스제닉 동물 및 식물**

[0392] 트랜스제닉 동물(모델)이 또한 제공되며, 하기는 생체외 모델 조직 및 조직의 수집물, 예를 들어, 오가노이드, 리버 온 어 칩 등에 동등하게 적용된다. 바람직한 예는 폴리뉴클레오티드 암호화 Cas9 또는 이의 단백질에 대해 Cas9를 포함하는 동물을 포함한다. 마우스, 래트 및 토끼가 바람직하다. 작제물을 이용하여 유전자이식 마우스를 생성하기 위해, 본 명세서에 예시한 바와 같이 가성임신 암컷, 예를 들어 CB56 암컷으로부터의 접합자의 전

핵 내로 순수한 선형 DNA를 주사할 수 있다. 이어서 발명자들은 CB57 마우스에 대해 동정, 게노타이핑 및 역교배할 수 있다. 이어서 작제물은, 예를 들어 생어 시퀀싱에 의해 클로닝되고 선택적으로 확인될 수 있다. 녹아웃은, 예를 들어 하나 이상의 유전자가 모델에서 녹아웃되는 경우가 생각된다. 그러나, 녹인이 또한 생각된다(단독으로 또는 조합으로). 예시적 녹인 Cas9 마우스가 생성되고, 이것이 예시되지만, Cas9 녹인이 바람직하다. Cas9 녹인 마우스를 생성하기 위해, 본 명세서에 기재된 바와 같이(도 25a 및 도 25b 및 도 26) Rosa26 유전자좌에 대한 동일한 구성적 및 조건적 작제물을 표적화할 수 있다. Rosa 유전자좌를 표적화하도록 지시된 Sangamo BioSciences, Inc.가 출원한 미국 특허 공개 제20120017290호 및 제20110265198호의 방법은 본 발명의 CRISPR Cas 시스템을 이용하도록 변형될 수 있다. 다른 구현예에서, Rosa 유전자좌를 표적화하도록 지시된 Cellectis가 출원한 미국 특허 공개 제20130236946호의 방법은 또한 본 발명의 CRISPR Cas 시스템을 이용하도록 변형될 수 있다.

[0393] 조건적 Cas9 마우스의 효용: 본 발명자들은 Cas9 조건적 발현 작제물이 Cre와 공동발현에 의해 활성화될 수 있다는 것을 293 세포에서 나타내었다. 본 발명자들은 또한 Cre가 발현될 때 정확하게 표적화된 R1 mESC가 활성 Cas9를 가질 수 있다는 것을 나타낸다. Cas9 다음에 P2A 펩타이드 절단 서열 그 다음에 EGFP가 오기 때문에, 본 발명자들은 EGFP를 관찰함으로써 성공적 발현을 동정한다. 본 출원인은 mESC에서 Cas9 활성화를 나타내었다. 이 동일한 개념은 상당히 유용한 조건적 Cas9 마우스를 만드는 것이다. 본 발명자들은 그들의 조건적 Cas9 마우스를 Cre(ACTB-Cre 계통)를 편재하여 발현시키는 마우스와 교배시킬 수 있고, 세포마다 Cas9를 발현시키는 마우스에 도달할 수 있다. 배아 또는 성체 마우스에서 게놈 편집을 유발하기 위해서는 키메라 RNA만을 전달해야 한다. 흥미롭게도, 조건적 Cas9 마우스가 조직 특이적 프로모터 하에 Cre를 발현하는 마우스와 교배된다면, 또한 Cre를 발현시키는 조직에 Cas9만 있어야 한다. 이 접근은 동일 조직에 키메라 RNA를 전달함으로써 정확한 조직에서만 게놈을 편집하는데 사용될 수 있다.

[0394] 상기 언급한 바와 같이, 유전자이식 식물, 특히 작물 및 조류와 같이, 유전자이식 동물이 또한 제공된다. 유전자이식 식물은 질환 모델을 제공하는 이외의 적용에서 유용할 수 있다. 이들은, 예를 들어 야생형에서 정상적으로 관찰되는 것보다 더 높은 수준의 단백질, 탄수화물, 영양소 또는 비타민의 발현을 통해 식품 또는 사료 생성을 포함할 수 있다. 이와 관련하여, 유전자이식 식물, 특히 두류 및 괴경, 및 동물, 특히 포유류, 예컨대 가축(소, 양, 염소 및 돼지)뿐만 아니라, 가금류 및 식용 곤충이 바람직하다.

[0395] 유전자이식 조류 또는 다른 식물, 예컨대 유채는 예를 들어 식물성 오일 또는 바이오연료, 예컨대 알코올(특히 메탄올 및 에탄올)의 생성에 특히 유용할 수 있다. 이들은 오일 또는 바이오연료 산업에서의 사용을 위하여 높은 수준의 오일 또는 알코올을 발현 또는 과발현하도록 조작될 수 있다.

[0396] **아데노 연관 바이러스(AAV)**

[0397] 생체내 전달에 대해, AAV는 몇 가지 이유로 다른 바이러스 벡터 이상으로 유리하다:

[0398] 저독성(이는 면역 반응을 활성화할 수 있는 세포입자의 초원심분리를 필요로 하지 않는 정제 방법에 기인할 수 있다)

[0399] 숙주 게놈 내로 통합하지 않기 때문에 삽입 돌연변이유발을 야기할 낮은 확률.

[0400] AAV는 4.5 또는 4.75 Kb의 패키징 제한을 가진다. 이는 Cas9뿐만 아니라 프로모터 및 전사 종결자가 모두 동일한 바이러스 벡터에 적합하게 된다는 것을 의미한다. 4.5 또는 4.75 Kb보다 큰 작제물은 상당히 감소된 바이러스 생산을 야기할 것이다. SpCas9는 상당히 크며, 유전자 그 자체는 4.1 Kb 이상인데, 이는 그것이 AAV 내로 패키징되는 것을 어렵게 만든다. 따라서 본 발명의 구현예는 더 짧은 Cas9의 상동체를 이용하는 것을 포함한다. 예를 들어:

종	Cas9 크기
코리네박터 디프테리아	3252
유박테리움 벤트리오숨 (Eubacterium ventriosum)	3321
스트렙토코커스 파스테우리아누스 (Streptococcus pasteurianus)	3390
락토바실러스 파르시미니스 (Lactobacillus farciminis)	3378
스파에로카에타 글루부스 (Sphaerochaeta globus)	3537
아조스피릴룸 B510	3504
글루코나세토박터 디아조트로피쿠스 (Gluconacetobacter diazotrophicus)	3150
네이세리아 시네레아	3246
로세부리아 인테스티날리스 (Roseburia intestinalis)	3420
파르비바큘룸 라바멘티보란스 (Parvibaculum lavamentivorans)	3111
스타필로코커스 아우레우스	3159
니트라티프락토르 살수기니스 (Nitratifractor salsuginis) DSM 16511	3396
캄필로박터 라리 (Campylobacter lari) CF89-12	3009
스트렙토코커스 써모필루스 LMD-9	3396

[0401]

[0402] 따라서 이들 종은 일반적으로 Cas9 종이 바람직하다. 본 발명자들은 전달 및 생체내 마우스 뇌 Cas9 발현 데이터를 나타내었다.

[0403] 생체내 게놈 변형을 매개하기 위해 Cas9 암호화 핵산 분자, 예를 들어, DNA를 바이러스 벡터 내로 패키징하는 2가지 방법이 바람직하다:

[0404] NHEJ-매개 유전자 녹아웃을 달성하기 위해:

[0405] 단일 바이러스 벡터:

[0406] 2 이상의 발현 카세트를 함유하는 벡터:

[0407] 프로모터-Cas9 암호화 핵산 분자 -종결자

[0408] 프로모터-gRNA1-종결자

[0409] 프로모터-gRNA2-종결자

[0410] 프로모터-gRNA(N)-종결자 (벡터의 크기 제한까지)

[0411] 이중 바이러스 벡터:

[0412] Cas9의 발현을 구동하기 위한 하나의 발현 카세트를 함유하는 벡터 1

[0413] 프로모터-Cas9 암호화 핵산 분자-종결자

[0414] 하나 이상의 가이드RNA의 발현을 구동하기 위한 하나 이상의 발현 카세트를 함유하는 벡터 2

[0415] 프로모터-gRNA1-종결자

[0416] 프로모터-gRNA(N)-종결자(벡터의 크기 제한까지)

[0417] 상동체-관련 수선을 매개하는 것. 상기 기재한 단일 및 이중 바이러스 벡터 접근에 추가로, 추가적인 벡터를 사용하여 상동체-관련 수선 주형을 전달한다.

[0418] Cas9 암호화 핵산 분자 발현을 구동시키기 위해 사용되는 프로모터는 하기를 포함할 수 있다:

[0419] AAV ITR은 프로모터로서 작용할 수 있다: 이는 추가적인 프로모터 구성요소(벡터 내 공간을 취할 수 있음)에 대한 필요를 제거하는데 유리하다. 추가적인 공간 개방은 추가적인 구성요소(gRNA 등)의 발현을 구동하는데 사용될 수 있다. 또한, ITR 활성은 상대적으로 더 약하고, 따라서 Cas9의 과발현에 기인하는 독성을 감소시키기 위

해 사용될 수 있다.

[0420] 편재하는 발현에 대해, 다음의 프로모터를 사용할 수 있다: CMV, CAG, CBh, PGK, SV40, 페리틴 중쇄 또는 경쇄 등.

[0421] 뇌 발현에 대해, 다음의 프로모터를 사용할 수 있다: 모든 뉴런에 대해 시냅신I(SynapsinI), 흥분성 뉴런에 대해CaMKII알파, 가바성(GABAergic) 뉴런에 대해 GAD67 또는 GAD65 또는 VGAT 등.

[0422] 간 발현에 대해, 알부민 프로모터를 사용할 수 있다.

[0423] 폐 발현에 대해 SP-B를 사용할 수 있다.

[0424] 내피세포에 대해 ICAM을 사용할 수 있다.

[0425] 조혈모세포에 대해 IFN베타 또는 CD45를 사용할 수 있다.

[0426] 조골세포에 대해 OG-2를 사용할 수 있다.

[0427] 가이드 RNA를 구동하기 위해 사용되는 프로모터는:

[0428] **U6 또는 H1과 같은 Pol III 프로모터를 포함할 수 있다.**

[0429] **gRNA를 발현시키기 위한 Pol II 프로모터 및 인트론 카세트의 사용**

[0430] AAV에 관해, AAV는 AAV1, AAV2, AAV5 또는 이들의 임의의 조합일 수 있다. 표적화될 세포와 관련하여 AAV의 AAV를 선택할 수 있으며; 예를 들어, 뇌 또는 신경세포를 표적화하기 위해 AAV 혈청형 1, 2, 5 또는 혼성체 또는 캡시드 AAV1, AAV2, AAV5 또는 이들의 임의의 조합을 선택할 수 있으며; 심장 조직을 표적화하기 위한 AAV4를 선택할 수 있다. AAV8은 간에 대한 전달에 유용하다. 상기 프로모터 및 벡터는 개별적으로 바람직하다.

[0431] RNA 전달은 또한 유용한 생체내 전달방법이다. 도 27은 전달 및 생체내 마우스 뇌 Cas9 발현 데이터를 나타낸다. 리포좀 또는 나노입자를 사용하여 Cas9 및 gRNA(예를 들어, HR 수선 주형)을 세포 내로 전달할 수 있다. 따라서 Cas9와 같은 CRISPR 효소의 전달 및/또는 본 발명의 RNA의 전달은 RNA 형태일 수 있고 미세소포, 리포좀 또는 나노입자를 통할 수 있다. 예를 들어, Cas9 mRNA 및 gRNA는 생체내 전달을 위해 리포좀 입자 내로 패키징될 수 있다. Life Technologies사제의 리포펙타민과 같은 리포좀 트렌스펙션 시약 및 시판 중인 다른 시약은 RNA 분자를 간에 효과적으로 전달할 수 있다.

[0432] NHEJ 또는 HR 효율을 향상시키는 것은 또한 전달에 도움을 준다. NHEJ 효율은 Trex2와 같은 최종 가공 효소를 공동발현시킴으로써 향상되는 것이 바람직하다(Dumitrache et al. Genetics. 2011 August; 188(4): 787-797). HR 효율은 Ku70 및 Ku86과 같은 NHEJ 기작을 일시적으로 저해함으로써 증가되는 것이 바람직하다. HR 효율은 또한 RecBCD, RecA와 같은 원핵 또는 진핵 상동성 재조합 효소를 공동발현시킴으로써 증가될 수 있다.

[0433] 다양한 전달 수단이 본 명세서에 기재되고, 이 부문에서 추가로 논의된다.

[0434] 바이러스 전달: CRISPR 효소, 예를 들어 Cas9, 및/또는 임의의 본 발명의 RNA, 예를 들어 가이드 RNA는 아데노연관 바이러스(AAV), 렌티바이러스, 아데노바이러스 또는 다른 바이러스 벡터 유형, 또는 이들의 조합을 사용하여 전달될 수 있다. Cas9 및 하나 이상의 가이드 RNA는 하나 이상의 바이러스 벡터 내로 패키징될 수 있다. 일부 구현예에서, 바이러스 벡터는, 예를 들어 근육내 주사에 의해 관심대상의 조직에 전달되는 한편, 바이러스 전달은 정맥내, 경피, 비강내, 구강, 점막 또는 다른 전달방법을 통한다. 이러한 전달은 단일 용량 또는 다회 용량을 통한 것일 수 있다. 당업자는 본 명세서에 전달될 실제 투약량이 다양한 인자, 예컨대 선택된 벡터, 표적 세포, 유기체 또는 조직, 치료될 대상체의 일반적 질환, 추구되는 형질전환/변형 정도, 투여 경로, 투여 방식, 추구되는 형질전환/변형 유형 등에 따라서 크게 다를 수 있다는 것을 이해한다.

[0435] 이러한 투약량은, 예를 들어 담체(물, 식염수, 에탄올, 글리세롤, 락토스, 수크로스, 인산칼슘, 젤라틴, 덱스트란, 한천, 펙틴, 땅콩유, 참깨유 등), 희석제, 약제학적으로 허용가능한 담체(예를 들어, 인산완충 식염수), 약제학적으로 허용가능한 부형제, 및/또는 당업계에 공지된 다른 화합물을 추가로 함유할 수 있다. 이러한 투약량 제형은 당업자에 의해 용이하게 확인될 수 있다. 투약은 하나 이상의 약제학적으로 허용가능한 염, 예를 들어 무기염, 예컨대 염산, 브롬화수소산, 인산염, 황산염 등; 및 유기산의 염, 예컨대 아세트산염, 프로피온산염, 말론산염, 벤조산염 등을 추가로 함유할 수 있다. 추가적으로, 보조 물질, 예컨대 습윤제 또는 유화제, pH 완충 물질, 겔 또는 겔화 물질, 향미제, 착색제, 미소구체, 중합체, 현탁제 등은 또한 본 명세서에 존재할 수 있다. 추가로, 특히 제형이 재구성가능한 형태라면 하나 이상의 다른 통상적인 약제학적 성분, 예컨대 보존제,

습윤제, 현탁제, 계면활성제, 항산화제, 고화방지제, 충전제, 킬레이트제, 코팅제, 화학적 안정제 등이 또한 존재할 수 있다. 적합한 예시적 성분은 미정질 셀룰로스, 카복시메틸셀룰로스 나트륨, 폴리솔베이트 80, 페닐에틸 알코올, 클로로부탄올, 솔브산칼륨, 솔브산, 이산화황, 갈산프로필, 파라벤, 에틸 바닐린, 글리세린, 페놀, 파라클로로페놀, 젤라틴, 알부민 및 이들의 조합을 포함한다. 약제학적으로 허용가능한 부형제의 철저한 논의는 본 명세서에 참고로 포함되는 문헌[REMINGTON'S PHARMACEUTICAL SCIENCES (Mack Pub. Co., N.J. 1991)]에서 이용가능하다.

[0436] 본 명세서의 구현예에서, 전달은 아데노바이러스 벡터의 적어도 1 x 10^5 개 입자(또한 단위 입자로서 지칭됨, pu)를 함유하는 단일 부스터 용량에 있을 수 있는 아데노바이러스를 통한다. 본 명세서의 구현예에서, 용량은 바람직하게는 적어도 약 1 x 10^6 개 입자(예를 들어, 약 1 x 10^6-1 x 10^{12} 개 입자), 더 바람직하게는 적어도 약 1 x 10^7 개 입자, 더 바람직하게는 적어도 약 1 x 10^8 개 입자(예를 들어, 약 1 x 10^8-1 x 10^{11} 개 입자 또는 약 1 x 10^8-1 x 10^{12} 개 입자), 및 가장 바람직하게는 적어도 약 1 x 10^0 개 입자(예를 들어, 약 1* x 10^9-1 x 10^{10} 개 입자 또는 약 1 x 10^9-1 x 10^{12} 개 입자), 또는 심지어 적어도 약 1 x 10^{10} 개 입자(예를 들어, 약 1 x 10^{10}-1 x 10^{12} 개 입자)의 아데노바이러스 벡터이다. 대안적으로, 용량은 1 x 10^{14} 개 이하의 입자, 바람직하게는 약 1 x 10^{13} 개 이하의 입자, 훨씬 더 바람직하게는 약 1 x 10^{12} 개 이하의 입자, 훨씬 더 바람직하게는 약 1 x 10^{11} 개 이하의 입자, 및 가장 바람직하게는 약 1 x 10^{10} 개 이하의 입자(예를 들어, 약 1 x 10^9 개 이하의 입자)를 포함한다. 따라서, 용량은, 예를 들어 약 1 x 10^6 개 입자 단위(pu), 약 2 x 10^6 pu, 약 4 x 10^6 pu, 약 1 x 10^7 pu, 약 2 x 10^7 pu, 약 4 x 10^7 pu, 약 1 x 10^8 pu, 약 2 x 10^8 pu, 약 4 x 10^8 pu, 약 1 x 10^9 pu, 약 2 x 10^9 pu, 약 4 x 10^9 pu, 약 1 x 10^{10} pu, 약 2 x 10^{10} pu, 약 4 x 10^{10} pu, 약 1 x 10^{11} pu, 약 2 x 10^{11} pu, 약 4 x 10^{11} pu, 약 1 x 10^{12} pu, 약 2 x 10^{12} pu, 또는 약 4 x 10^{12} pu의 아데노바이러스 벡터에 의한 단일 용량의 아데노바이러스 벡터를 함유할 수 있다. 예를 들어, 본 명세서에 참고로 포함되는 2013년 6월 4일 등록된 Nabel 등의 미국 특허 제8,454,972 B2호에서의 아데노바이러스 벡터, 및 이의 29 칼럼, 36 내지 58줄에서의 투약량을 참조한다. 본 명세서의 구현예에서, 아데노바이러스는 다회용량을 통해 전달된다.

[0437] 본 명세서의 구현예에서, 전달은 AAV를 통한다. 인간에 대한 AAV의 생체내 전달을 위한 치료적 유효량은 약 1 x 10^{10} 내지 약 1 x 10^{10} 기능성 AAV/㎖ 용액을 함유하는 식염수 용액의 약 20 내지 약 50 ㎖의 범위에 있는 것으로 믿어진다. 투약량은 임의의 부작용에 대해 치료적 이점의 균형을 맞추도록 조절될 수 있다. 본 명세서의 구현예에서, AAV 용량은 일반적으로 약 1 x 10^5 내지 1 x 10^{50} 개 게놈 AAV, 약 1 x 10^8 개 내지 1 x 10^{20} 개 게놈 AAV, 약 1 x 10^{10} 내지 약 1 x 10^{16} 개 게놈, 또는 약 1 x 10^{11} 내지 약 1 x 10^{16} 개 게놈 AAV의 농도 범위에 있다. 인간 투약량은 약 1 x 10^{13} 개 게놈 AAV일 수 있다. 이러한 농도는 약 0.001 ㎖ 내지 약 100 ㎖, 약 0.05 내지 약 50 ㎖, 또는 약 10 내지 약 25 ㎖의 담체 용액으로 전달될 수 있다. 다른 유효한 투약량은 용량 반응 곡선을 확립하는 일상적 시도를 통해 당업자에 의해 용이하게 확립될 수 있다. 예를 들어, 2013년 3월 26일 등록된 Hajjar 등의 미국 특허 제8,404,658 B2호, 27 칼럼 45 내지 60줄.

[0438] 본 명세서의 구현예에서, 전달은 플라스미드를 통한다. 이러한 플라스미드 조성물에서, 투약량은 반응을 유발하는데 충분한 양의 플라스미드여야한다. 예를 들어, 플라스미드 조성물 중의 플라스미드 DNA의 적합한 양은 약 0.1 내지 약 2 ㎎, 또는 약 1 ㎍ 내지 약 10 ㎍일 수 있다.

[0439] 본 명세서의 용량은 평균 70 ㎏ 개체를 기준으로 한다. 투여 빈도는 의학적 또는 수의학적 실행자(예를 들어, 의사, 수의사)또는 당업자의 영역 내이다. 실험에 사용된 마우스는 약 20g이다. 20g 마우스에 투여되는 것으로부터, 70㎏ 개체가 외삽될 수 있다.

[0440] **렌티바이러스**

[0441] 렌티바이러스는 유사분열과 유사분열 후 세포 둘 다에서 그들의 유전자를 감염 및 발현시키는 능력을 갖는 복합 레트로바이러스이다. 가장 흔히 알려진 렌티바이러스는 광범위한 세포 유형을 표적화하기 위해 다른 바이러스의 외피 당단백질을 사용하는 인간 면역결핍바이러스(HIV)이다.

[0442] 렌티바이러스 다음과 같이 제조될 수 있다. (렌티 바이러스 전달 플라스미드 백본을 함유하는) pCasES10의 클로

닝 후에, 낮은 계대(p=5)에서 HEK293FT를 10% 소태아 혈청과 함께 그리고 항생제 없이 DMEM 중의 트렌스펙션 전날에 50% 컨플류언스로 T-75 플라스크에 파종하였다. 20 시간 후에, 배지를 OptiMEM(무혈청) 배지로 바꾸고, 4 시간 후에 트렌스펙션을 행하였다. 세포를 10 ㎍의 렌티바이러스 전달 플라스미드(pCasES10) 및 다음의 패키징 플라스미드: 5 ㎍의 pMD2.G(VSV-g 위형), 및 7.5 ㎍의 psPAX2 (gag/pol/rev/tat)로 트렌스펙션시켰다. 양이온성 지질 전달제(50㎕ 리포펙타민 2000 및 100㎕ 플러스 시약)을 이용하여 4 ㎖ OptiMEM 중에서 트렌스펙션을 행하였다. 6 시간 후, 배지를 10% 소태아 혈창을 지니는 무항생제 DMEM으로 바꾸었다.

[0443] 렌티바이러스는 다음과 같이 정제될 수 있다. 바이러스 상청액을 48 시간 후에 채취하였다. 상청액은 처음에 파편이 없으며, 0.45 um 저 단백질 결합(PVDF) 필터를 통해 여과시켰다. 이어서, 그들을 24,000 rpm에서 2 시간 동안 원심분리로 교반시켰다. 바이러스 펠렛을 4C에서 밤새 50 ㎕의 DMEM 중에서 재현탁시켰다. 이어서 그들을 알리쿼팅하고 즉시 -80C에서 냉동시켰다.

[0444] 다른 구현예에서, 말 전염성 빈혈 바이러스(equine infectious anemia virus: EIAV)에 기반한 최소 비영장류 렌티바이러스 벡터는 또한, 특히 눈 유전자 요법을 고려하였다(예를 들어, 문헌[Balagaan, J Gene Med 2006; 8: 275 - 285], Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/jgm.845에서 2005년 11월 21일 온라인 공개). 다른 구현예에서, RetinoStat®, 습윤 형태의 노인성 황반변형의 치료를 위한 망막하 주사를 통해 전달되는 혈관억제 단백질 엔도스타틴 및 앤지오스타틴을 발현시키는 말 전염성 빈혈 바이러스 기반 렌티 바이러스 유전자 요법 벡터가 또한 고려되며(예를 들어, 문헌[Binley et al., HUMAN GENE THERAPY 23:980-991 (September 2012)]) 본 발명의 CRISPR-Cas 시스템에 대해 변형될 수 있다.

[0445] 다른 구현예에서, HIV tat/rev에 의해 공유되는 통상적인 엑손을 표적화하는 siRNA, 핵소체-편재화 TAR 미끼, 및 항-CCR5-특이적 해머헤드 리보자임을 이용하는 자기-불활성화 렌티바이러스 벡터(예를 들어, DiGiusto et al. (2010) Sci Transl Med 2:36ra43)가 본 발명의 CRISPR-Cas 시스템에 대해 사용되고/되거나 적합하게 될 수 있다. 최소 2.5 × 10^6 CD34+ 세포/환자의 체중(킬로그램)은, 2 × 10^6 개 세포/㎖의 밀도에서 2 mM L-글루타민, 줄기 세포 인자(100 ng/㎖), Flt-3 리간드(Flt-3L)(100 ng/㎖), 및 트롬보포이에틴(10 ng/㎖)(CellGenix)을 함유하는 X-VIVO 15 배지(Lonza)에서 16 내지 20 시간 동안 수집하고 사전자극할 수 있다. 사전자극 세포는 피브로넥틴(25 ㎎/㎠)으로 코팅된 75-㎠ 조직 배양 플라스크에서 16 내지 24 시간 동안 감염다중도 5로 렌티바이러스에 의해 형질도입될 수 있다(RetroNectin, Takara Bio Inc.).

[0446] 렌티바이러스 벡터는 파킨슨병의 치료를 위한 치료의 경우와 같이 개시되었다, 예를 들어, 미국 특허 공개 제20120295960호 및 미국 특허 제7303910호 및 제 7351585호 참조. 렌티바이러스 벡터는 또한 안질환의 치료용으로 개시되었다, 예를 들어, 미국 특허 공개 제20060281180호, 제20090007284호, 미국 특허 제20110117189호; 미국 특허 제20090017543호; 미국 특허 제20070054961호, 미국 특허 제20100317109호 참조. 렌티바이러스 벡터는 또한 뇌에 대한 전달용으로 개시되었다, 예를 들어, 미국 특허 공개 제20110293571호; 미국 특허 제20110293571호, 미국 특허 제20040013648호, 미국 특허 제20070025970호, 미국 특허 제20090111106호 및 미국 특허 제7259015호 참조.

[0447] **RNA 전달**

[0448] RNA 전달: CRISPR 효소, 예를 들어 Cas9, 및/또는 밍의의 본 발명의 RNA, 예를 들어 가이드 RNA는 또한 RNA의 형태로 전달될 수 있다. Cas9 mRNA는 시험관내 전사를 사용하여 생성될 수 있다. 예를 들어, Cas9 mRNA는 다음의 구성요소를 함유하는 PCR 카세트를 사용하여 합성될 수 있다: 베타 글로빈-폴리A 꼬리(일련의 120개 이상의 아데닌)로부터의 T7_프로모터-코작 서열(GCCACC)-Cas9-3' UTR. 카세트는 T7 중합효소의 전사를 위해 사용될 수 있다. 가이드 RNA는 또한 T7_프로모터-GG-가이드 RNA 서열을 함유하는 카세트로부터의 시험관내 전사를 사용하여 전사될 수 있다.

[0449] 발현을 향상시키고 독성을 감소시키기 위해, CRISPR 효소 및/또는 가이드 RNA는 슈도-U 또는 5-메틸-C를 사용하여 변형될 수 있다.

[0450] mRNA 전달 방법은 현재 간 전달용으로 특히 촉망된다. 특히, AAV8에 대해 간에 대한 전달이 특히 바람직하다.

[0451] **나노입자**

[0452] CRISPR 효소 mRNA 및 가이드 RNA는 나노입자 또는 지질 외피를 사용하여 동시에 전달될 수 있다.

[0453] 예를 들어, Su X, Fricke J, Kavanagh DG, Irvine DJ의 문헌["In vitro and in vivo mRNA delivery using lipid-enveloped pH-responsive 중합체 nanoparticles" Mol Pharm. 2011 Jun 6;8(3):774-87. doi:

10.1021/mp100390w. Epub 2011 Apr 1]은 인지질 이중층 껍질에 의해 둘러싸인 폴리(β-아미노에스터)(PBAE) 코어를 지니는 생분해성 코어 껍질 구조의 나노입자를 기재한다. 이들은 생체내 mRNA 전달을 위해 개발되었다. pH-반응성 PBAE 구성성분은 엔도솜 파괴를 촉진하도록 선택되는 한편, 지질표면층은 폴리양이온 코어의 독성을 최소화하도록 선택되었다. 따라서 본 발명의 RNA를 전달하는 것이 바람직하다.

[0454] 일 구현예에서, 펩타이드의 구강 전달, 펩타이드의 정맥내 전달 및 펩타이드의 비강 전달에, 뇌에 대한 모두에 적용될 수 있는 생접착성 중합체의 자가 조립에 기반한 나노입자가 고려된다. 다른 구현예, 예컨대 구강 흡수 및 소수성 약물의 눈 전달이 또한 고려된다. 분자 외피는 기술적으로 질환의 부위에 대해 보호 및 전달되는 유전자 조작된 중합체 외피를 수반한다(예를 들어, 문헌[Mazza, M. et al. ACSNano, 2013. 7(2): 1016-1026; Siew, A., et al. Mol Pharm, 2012. 9(1):14-28; Lalatsa, A., et al. J Contr Rel, 2012. 161(2):523-36; Lalatsa, A., et al., Mol Pharm, 2012. 9(6):1665-80; Lalatsa, A., et al. Mol Pharm, 2012. 9(6):1764-74; Garrett, N.L., et al. J Biophotonics, 2012. 5(5-6):458-68; Garrett, N.L., et al. J Raman Spect, 2012. 43(5):681-688; Ahmad, S., et al. J Royal Soc Interface 2010. 7:S423-33; Uchegbu, I.F. Expert Opin Drug Deliv, 2006. 3(5):629-40; Qu, X.,et al. Biomacromolecules, 2006. 7(12):3452-9 및 Uchegbu, I.F., et al. Int J Pharm, 2001. 224:185-199] 참조). 표적 조직에 따라서 약 5 mg/kg의 단일 또는 다중 용량이 고려된다.

[0455] 일 구현예에서, MIT의 댄 앤더슨 연구소(Dan Anderson's lab)에 의해 개발된 종양 성장을 중단시키기 위해 암세포에 RNA를 전달할 수 있는 나노입자가 본 발명의 CRISPR Cas 시스템에 대해 사용되고/사용되거나 적합하게 될 수 있다. 특히, 앤더슨 연구소는 새로운 생체재료 및 나노제형의 합성, 정제, 특성규명 및 제형화를 위한 완전히 자동화된 조합적 시스템을 개발하였다. 예를 들어, 문헌[Alabi et al., Proc Natl Acad Sci U S A. 2013 Aug 6;110(32):12881-6; Zhang et al., Adv Mater. 2013 Sep 6;25(33):4641-5; Jiang et al., Nano Lett. 2013 Mar 13;13(3):1059-64; Karagiannis et al., ACS Nano. 2012 Oct 23;6(10):8484-7; Whitehead et al., ACS Nano. 2012 Aug 28;6(8):6922-9 및 Lee et al., Nat Nanotechnol. 2012 Jun 3;7(6):389-93] 참조.

[0456] 미국 특허 출원 제20110293703호는 또한 본 발명의 CRISPR Cas 시스템을 전달하기 위해 적용될 수 있는 폴리뉴클레오티드의 투여에서 특히 유용한 리피도이드 화합물에 관한 것이다. 일 양태에서, 아미노알코올 리피도이드 화합물은 작용제와 조합되어 세포 또는 대상체에게 전달되어 마이크로입자, 나노입자, 리포좀, 또는 미셀을 형성한다. 입자, 리포좀 또는 미셀에 의해 전달될 작용제는 기체, 액체 또는 고체의 형태일 수 있고, 작용제는 폴리뉴클레오티드, 단백질, 펩타이드 또는 수분자일 수 있다. 미노알코올 리피도이드 화합물은 다른 아미노알코올 리피도이드 화합물, 중합체(합성 또는 천연), 계면활성제, 콜레스테롤, 탄수화물, 단백질, 지질 등과 조합되어 입자를 형성할 수 있다. 이어서 이들 입자는 선택적으로 약제학적 부형제와 조합되어 약제학적 조성물을 형성할 수 있다.

[0457] 미국 특허 공개 제0110293703호는 또한 아미노알코올 리피도이드 화합물을 준비하는 방법을 제공한다. 1 이상의 당량의 아민은 본 발명의 아미노알코올 리피도이드 화합물을 형성하기에 적합한 조건 하에서 에폭시드-종결 화합물의 1 당량 이상과 반응된다. 특정 구현예에서, 아민의 모든 아미노기는 에폭시드-종결 화합물과 완전히 반응되어 3차 아민을 형성한다. 다른 구현예에서, 아민의 모든 아미노기는 3차 아민을 형성하는 에폭시드-종결 화합물과 완전히 반응되지 않으며, 이에 의해 아미노알코올 리피도이드 화합물에서 1차 또는 2차 아민을 초래한다. 이들 1차 또는 2차 아민은 상이한 에폭시드-종결 화합물과 같은 다른 친핵체로서 남아있거나 또는 이런 다른 친핵체와 반응될 수 있다. 당업자에 의해 인식될 바와 같이, 아민을 과량 미만의 에폭시드-종결 화합물과 반응시키는 것은 다수의 꼬리를 지니는 복수의 상이한 아미노알코올 리피도이드 화합물을 초래할 것이다. 특정 아민은 2개의 에폭시드 유래 화합물 꼬리에 의해 완전히 기능화될 수 있는 반면, 다른 분자는 에폭시드 유래 화합물 꼬리에 의해 완전히 기능화되지 않을 것이다. 예를 들어, 디아민 또는 폴리아민은 1차, 2차 및 3차 아민을 초래하는 분자의 다양한 아미노산 모이어티의를 감소시킨 1, 2, 3 또는 4개의 에폭시드 유래 화합물을 포함할 수 있다. 특정 구현예에서, 모든 아미노기는 완전히 기능화되지 않는다. 특정 구현예에서, 2 가지 동일 유형의 에폭시드 종결 화합물이 사용된다. 다른 구현예에서, 2 이상의 상이한 에폭시드 종결 화합물이 사용된다. 아미노알코올 리피도이드 화합물의 합성은 용매와 함께 또는 용매 없이 수행되고, 합성은 30 내지 100℃의 범위에 있는 더 고온에서, 바람직하게는 대략 50 내지 90℃에서 수행된다. 제조된 아미노알코올 리피도이드 화합물은 선택적으로 정제될 수 있다. 예를 들어, 아미노알코올 리피도이드 화합물의 혼합물은 정제되어 특정 수의 에폭시드 유래 화합물 꼬리를 지니는 아미노알코올 리피도이드 화합물을 수득할 수 있다. 또는 혼합물은 정제되어 특정 입체이성질체 또는 레지오이성질체를 수득할 수 있다. 아미노알코올 리피도이드 화합물은 또한 알킬 할로겐화물(예를 들어, 요오드화메틸) 또는 다른 알킬화제를 사용하여 알킬화될 수 있고/있거나 그들은 아실화될 수 있다.

[0458] 미국 특허 공개 제0110293703호는 본 발명의 방법에 의해 제조된 아미노알코올 리피도이드 화합물의 라이브러리를 제공한다. 이들 아미노알코올 리피도이드 화합물은 액체 조절기, 로봇, 마이크로역가 플레이트, 컴퓨터 등을 사용하여 제조 및/또는 선별될 수 있다. 특정 구현예에서, 아미노알코올 리피도이드 화합물은 폴리뉴클레오티드 또는 다른 작용제(예를 들어, 단백질, 펩타이드, 소분자)를 세포에 트렌스펙션하는 그들의 능력에 대해 선별된다.

[0459] 미국 특허 공개 제20130302401호는 조합적 중합화를 사용하여 제조된 폴리(베타-아미노 알코올)(PBAA)의 부류에 관한 것이다. 본 발명의 PBAA는 코팅(예컨대 의료장치 또는 이식물에 대한 필름 또는 다중층의 코팅), 첨가제, 재료, 부형제, 비생물오손제(non-biofouling agent), 미세패터닝제 및 세포 캡슐화제로서 생명공학 및 생체의학 적용분야에서 사용될 수 있다. 표면 코팅으로서 사용될 때, 이들 PBAA는 그들의 화학적 구조에 따라서 시험관내와 생체내 둘 다의 상이한 염증 수준을 유발하였다. 이 재료 부류의 큰 화학적 다양성은 본 발명자들이 시험관내 대식세포 활성화를 저해하는 중합체 코팅을 동정하게 하였다. 더 나아가, 이들 코팅은 카복실화된 폴리스티렌의 피하 이식 후에 염증 세포의 동원을 감소시키고, 섬유증을 감소시킨다. 이들 중합체를 사용하여 세포 캡슐화를 위한 고분자전해질 복합체 캡슐을 형성한다. 본 발명은 또한 미생물방지 코팅, DNA 또는 siRNA 전달, 및 줄기 세포 조직 유전자조작과 같은 다수의 다른 생물학적 적용분야를 가질 수 있다. 미국 특허 공개 제20130302401호의 교시는 본 발명의 CRISPR Cas 시스템에 적용될 수 있다.

[0460] 다른 구현예에서, 지질 나노입자(LNP)가 고려된다. 특히, 지질 나노입자에서 캡슐화된 항트렌스티레틴(antitransthyretin) 짧은 간섭 RNA(예를 들어, 문헌[Coelho et al., N Engl J Med 2013;369:819-29] 참조)는 본 발명의 CRISPR Cas 시스템에 적용될 수 있다. 정맥내로 투여된 체중의 kg 당 약 0.01 내지 약 1 mg의 용량이 고려된다. 주입 관련 반응의 위험을 감소시키기 위한 의약이 고려되며, 예컨대 덱사메타손, 아세탐피노펜, 디펜하이드라민 또는 세티리진 및 라니티딘이 고려된다. 5 용량에 대해 4 주마다 킬로그램 당 약 0.3 mg의 다회 용량이 또한 고려된다.

[0461] LNP는 siRNA를 간에 전달함에 있어 고도로 효과적인 것으로 나타났으며(예를 들어, 문헌[Tabernero et al., Cancer Discovery, April 2013, Vol. 3, No. 4, 363 내지 470 페이지] 참조) 따라서 간에 CRISPR Cas를 전달하는 것이 고려된다. 2 주마다 6 mg/kg의 LNP(또는 CRISPR-Cas 시스템의 RNA)의 약 4 회 용량의 투약량이 고려될 수 있다. Tabernero 등은 0.7 mg/kg에서 투약한 LNP의 처음 2 주기 후에 6 주기의 마지막까지 종양 억제를 관찰하였고, 환자는 림프절 전이의 퇴화 및 간 종양의 실질적 수축에 의한 부분적 반응을 달성하였음을 입증하였다. 26 개월에 걸쳐 용량을 받은 후에 관해가 남아있고 치료를 완료한 이 환자에게 40 용량 후에 완전한 반응을 얻었다. VEGF 경로 저해제를 이용한 사전 요법 후에 진행한 신장, 폐 및 림프절을 포함하는 RCC 및 간외영역 질환을 지니는 2 명의 환자는 대략 8 내지 12 개월 동안 모든 부위에서 안정한 질환을 가졌고, PNET 및 간 전이를 지니는 환자는 안정한 상태가 18개월(36회 용량) 동안의 연장 연구에 대해 계속되었다.

[0462] 그러나, LNP의 전하를 고려하여야 한다. 양이온성 지질은 세포내 전달을 용이하게 하는 비이중층 구조를 유발하기 위해 음으로 하전된 지질과 조합되기 때문이다. 하전된 LNP는 정맥내 주사 후 순환으로부터 빠르게 제거되기 때문에, 7 미만의 pKa 값을 지니는 이온화가능한 양이온성 지질을 개발하였다(예를 들어, 문헌[Rosin et al, Molecular Therapy, vol. 19, no. 12, pages 1286-2200, Dec. 2011] 참조). 음으로 하전된 중합체, 예컨대 siRNA 올리고뉴클레오티드는, 이온화가능한 지질이 양전하를 나타내는 경우 낮은 pH 값(예를 들어, pH 4)에서 LNP에 장입될 수 있다. 그러나, 생리적 pH 값에서, LNP는 더 긴 순환시간과 양립가능한 낮은 표면 전하를 나타낸다. 이온화가능한 양이온성 지질의 4 개 종, 즉, 1,2-디리네오일-3-디메틸암모늄-프로판(DLinDAP), 1,2-디리놀레일옥시-3-N,N-디메틸아미노프로판(DLinDMA), 1,2-디리놀레일옥시-케토-N,N-디메틸-3-아미노프로판(DLinKDMA), 및 1,2-디리놀레일-4-(2-디메틸아미노에틸)-[1,3]-디옥솔란(DLinKC2-DMA)에 중점을 두었다. 이들 지질을 함유하는 LNP siRNA 시스템은 생체내 간세포에서 현저하게 상이한 유전자 침묵 특성을 나타내는 것으로 나타났고, 효능은 인자 VII 유전자 침묵 모델을 사용하는 시리즈 DLinKC2-DMA>DLinKDMA>DLinDMA>>DLinDAP에 따라 다르다(예를 들어, 문헌[Rosin et al, Molecular Therapy,, vol. 19, no. 12, 1286내지 2200 페이지, Dec. 2011). 특히 DLinKC2-DMA을 함유하는 제형에 대해 1 ㎍/㎖ 수준의 투약량이 고려될 수 있다.

[0463] LNP 및 CRISPR Cas 캡슐화의 제조가 사용될 수 있고/있거나 문헌[Rosin et al, Molecular Therapy, vol. 19, no. 12, 1286 내지 2200 페이지, Dec. 2011]에 적합하게 될 수 있다. 양이온성 지질 1,2-디리네오일-3-디메틸암모늄-프로판(DLinDAP), 1,2-디리놀레일옥시-3-N,N-디메틸아미노프로판(DLinDMA), 1,2-디리놀레일옥시케토-N,N-디메틸-3-아미노프로판(DLinK-DMA), 1,2-디리놀레일-4-(2-디메틸아미노에틸)-[1,3]-디옥솔란(DLinKC2-DMA), (3-o-[2″-(메톡시폴리에틸렌글리콜 2000) 숙시노일]-1,2-디미리스토일-sn-글리콜(PEG-S-DMG) 및 R-3-[(ω-메톡시-폴리(에틸렌 글리콜)2000) 카바모일]-1,2-디미리스틸록시프로필-3-아민(PEG-C-DOMG)은 Tekmira

Pharmaceuticals(캐나다 밴쿠버에 소재)에 의해 제공될 수 있거나 합성될 수 있다. 콜레스테롤은 Sigma사(미조리주 세인트루이스에 소재)로부터 구입할 수 있다. 특이적 CRISPR Cas RNA는 DLinDAP, DLinDMA, DLinK-DMA 및 DLinKC2-DMA(40:10:40:10 몰비로 양이온성 지질:DSPC:CHOL: PEGS-DMG 또는 PEG-C-DOMG)를 함유하는 LNP 내에 캡슐화될 수 있다. 필요하다면, 0.2% SP-DiOC18(캐나다 벌링턴에 소재한 Invitrogen)은 혼입되어 세포내 흡수, 세포내 전달 및 생체분포를 평가할 수 있다. 캡슐화는 10 mmol/l의 최종 지질 농도로 에탄올 중에서 양이온성 지질:DSPC:콜레스테롤:PEG-c-DOMG (40:10:40:10 몰비)을포함한 지질 혼합물을 용해함으로써 수행될 수 있다. 지질의 이 에탄올 용액을 50 mmol/l 시트르산염, pH 4.0에 적가하여 다중 라멜라 소수포를 형성하여 30% 에탄올 vol/vol의 최종 농도를 생성할 수 있다. 압출기(캐나다 벤쿠버에 소재한 Northern Lipids)를 사용하여 2 개 적층 80 nm Nuclepore 폴리카보네이트 필터를 통해 다중 라멜라 소수포의 압출 후 거대 라멜라 소수포가 형성될 수 있다. 압출 프리폼의 거대 단일 라멜라 소수포에 적가하고 0.06/1 wt/wt의 최종 RNA/지질 중량비로 일정하게 혼합하면서 30 분 동안 31℃에서 인큐베이션한 30% 에탄올 vol/vol을 함유하는 50 mmol/l 시트르산염, pH 4.0 중의 2 mg/㎖에서 용해시킨 RNA를 첨가함으로써 캡슐화를 달성할 수 있다. Spectra/Por 2 재생 셀룰로스 투석막을 사용하여 16 시간 동안 인산염 완충 식염수(PBS), pH 7.4에 대해 투석에 의해 에탄올의 제거 및 제형 완충제의 중화를 수행하여다. NICOMP 370 입도 분석기, 소수포/강도 모드, 및 가우시안 피팅(캘리포니아주 산타 바바라에 소재한 Nicomp Particle Sizing)을 사용하여 역동적 광 분산에 의해 나노입자 크기 분포를 결정할 수 있다. 모두 3 개의 LNP 시스템에 대한 입자 크기는 직경이 약 70 nm일 수 있다. 투석 전 및 후에 수집한 샘플로부터 VivaPureD MiniH 칼럼(Sartorius Stedim Biotech)을 사용하여 유리 siRNA의 제거에 의해 siRNA 캡슐화 효율을 결정할 수 있다. 캡슐화된 RNA는 용리 나노입자로부터 추출될 수 있고, 260 nm에서 정량화된다. siRNA 대 지질비는 Wako Chemicals USA(버지니아주 리치몬드에 소재)제의 콜레스테롤 E 효소 분석을 사용하여 소수포 내 콜레스테롤 함량의 측정에 의해 결정될 수 있다. 페길화된 리포좀(또는 LNP)가 또한 전달에 사용될 수 있다.

[0464] 거대 LNP의 제조는 문헌[Rosin et al, Molecular Therapy, vol. 19, no. 12, 1286 내지 2200페이지, Dec. 2011]로부터 사용될 수 있고/있거나 이에 적합하게 될 수 있다. 지질 프레믹스 용액(20.4 mg/㎖ 총 지질 농도)는 50:10:38.5 몰비에서 DLinKC2-DMA, DSPC 및 콜레스테롤을 함유하는 에탄올 중에서 제조될 수 있다. 아세트산나트륨은 0.75:1의 몰비로 지질 프레믹스에 첨가될 수 있다(아세트산나트륨:DLinKC2-DMA). 지질은 격렬한 교반과 함께 1.85 용적의 시트르산염 완충제(10 mmol/l, pH 3.0)와 혼합물을 합함으로써 순차적으로 수화되어, 35% 에탄올을 함유하는 수성 완충제 중에서 자발적 리포좀 형성을 초래할 수 있다. 리포좀 용액은 37℃에서 인큐베이션되어 입자 크기에서 시간-의존적 증가를 허용한다. 알리퀴트는 역동적 광 산란에 의한 리포좀 크기에서의 변화를 조사하기 위해 인큐베이션 동안 다양한 시간에 제거될 수 있다(Zetasizer Nano ZS, 영국 우스터셔에 소재한 Malvern Instruments). 일단 목적으로 하는 입자 크기가 달성되면, 수성 PEG 지질 용액(저장액 = 35%(vol/vol) 에탄올 중의 10 mg/㎖ PEG-DMG)은 리포좀 혼합물에 첨가되어 3.5%의 총 지질의 최종 PEG 몰 농도를 수득할 수 있다. PEG-지질의 첨가 시, 리포좀을 추가 성장을 효과적으로 퀀칭시켜 그들의 크기를 해야 한다. 이어서, RNA는 대략 1:10(wt:wt)의 총 지질비로 siRNA에서 빈 리포좀에 첨가된 후에 37℃에서 30 분 동안 인큐베이션되어 LNP를 장입할 수 있다. 혼합물은 후속적으로 PBS 중에서 밤새 투석되고 나서, 0.45-μm 주사기 필터로 여과될 수 있다.

[0465] 구체 핵산(Spherical Nucleic Acid: SNA™) 작제물 및 다른 나노입자(특히 금 나노입자)는 또한 의도된 표적에 CRISPR/Cas 시스템을 전달하기 위한 수단으로서 고려된다. 유의한 데이터는 핵산 기능화된 금 나노입자에 기반한 AuraSense Therapeutics' Spherical Nucleic Acid(SNA™) 작제물이 하기와 같은 다수의 중요한 성공 인자에 기반한 대안의 플랫폼보다 우수하다는 것을 나타낸다:

[0466] 높은 생체내 안정성. 그들의 밀집한 장입 때문에, 다수의 화물(cargo)(DNA 또는 siRNA)은 세포 내부의 작제물에 결합되어 핵산 안정성 및 효소 분해에 대한 저항성을 부여한다.

[0467] 전달능력. 연구한 모든 세포 유형(예를 들어, 뉴런, 종양 세포주 등)에 대해, 작제물은 담체 또는 트렌스펙션제에 대한 필요 없이 99%의 트렌스펙션 효율을 입증한다.

[0468] 치료적 표적화. 작제물의 독특한 표적 결합 친화도 및 특이성은 매칭 표적 서열에 대한 강렬한 특이성(즉, 제한된 표적외 효과)을 허용한다.

[0469] 우수한 효능. 작제물은 선두적인 전통적 트렌스펙션 시약(리포펙타민 2000 및 사이토펙틴)을 상당히 능가한다.

[0470] 낮은 독성. 작제물은 다양한 배양 세포, 1차 세포 및 명확한 독성이 없는 조직에 유입될 수 있다.

[0471] 상당한 면역 반응 없음. 작제물은 전체-게놈 마이크로어레이 연구 및 사이토카인-특이적 단백질 분석에 의해 측

정되는 바와 같은 세계적 유전자 발현에서의 최소 변화를 유발한다.

[0472] 화학적 적용성. 다수의 단일 또는 조합제(예를 들어, 단백질, 펩타이드 소분자)는 작제물의 표면을 맞추는데 사용될 수 있다.

[0473] 핵산 기반 치료제에 대한 이 플랫폼은 염증 및 전염성 질환, 암, 피부 장애 및 심혈관 질환을 포함하는 수많은 질환 상태에 적용가능할 수 있다.

[0474] 인용가능한 문헌은 문헌[Cutler et al.,. J. Am. Chem. Soc. 2011 133:9254-9257, Hao et al., Small. 2011 7:3158-3162, Zhang et al., ACS Nano. 2011 5:6962-6970, Cutler et al., J. Am. Chem. Soc. 2012 134:1376-1391, Young et al.,. Nano Lett. 2012 12:3867-71, Zheng et al., Proc. Natl. Acad. Sci. USA. 2012 109:11975-80, Mirkin, Nanomedicine 2012 7:635-638 Zhang et al., J. Am. Chem. Soc. 2012 134:16488-1691, Weintraub, Nature 2013 495:S14-S16, Choi et al., Proc. Natl. Acad. Sci. USA. 2013 110(19):7625-7630, Jensen et al., Sci. Transl. Med. 5, 209ra152 (2013) 및 Mirkin, et al., Small, doi.org/10.1002/smll.201302143]을 포함한다.

[0475] siRNA를 지니는 자기 조립 나노입자는, 예를 들어 종양 혈관신생 발현 인테그린을 표적화하기 위한 수단으로서 폴리에틸렌 글리콜(PEG)의 원위 단부에 부착된 Arg-Gly-Asp(RGD) 펩타이드 리간드와 페길화된 폴리에틸렌이민(PEI)을 이용하여 구성될 수 있고, siRNA 저해 혈관표피성장인자 수용체-2(VEGF R2) 발현 및 이에 의한 종양 혈관형성을 전달하는데 사용될 수 있다(예를 들어, 문헌[Schiffelers et al., Nucleic Acids Research, 2004, Vol. 32, No. 19] 참조). 나노복합체는 동일한 용적의 양이온성 중합체 및 핵산의 수용액을 혼합함으로써 제조되어 2 내지 6의 범위에 걸쳐 인산염(핵산)에 대한 순 몰과량의 이온화가능한 질소(중합체)를 제공할 수 있다. 양이온성 중합체와 핵산 간의 정전기적 상호작용은 약 100 nm의 평균 입자 크기를 지니는 다중복합체 형성을 야기하며, 따라서 본 명세서에서 나노복합체로서 지칭된다. Schiffelers 등의 자기 조립 나노입자에서의 전달에 대해 약 100 내지 200 mg의 CRISPR Cas의 투약량이 생각된다.

[0476] Bartlett 등(PNAS, September 25, 2007,vol. 104, no. 39)의 나노복합체는 또한 본 발명에 적용될 수 있다. Bartlett 등의 나노복합체는 동일 용적의 양이온성 중합체 및 핵산의 수용액을 혼합함으로써 제조되어 2 내지 6의 범위에 걸쳐 인산염(핵산)에 대한 순 몰과량의 이온화가능한 질소(중합체)를 제공할 수 있다. 양이온성 중합체와 핵산 간의 정전기적 상호작용은 약 100 nm의 평균 입자 크기 분포를 지니는 다중복합체의 형성을 초래하였고, 따라서 본 명세서에서 나노복합체로서 지칭된다. Bartlett 등의 DOTA-siRNA는 다음과 같이 합성되었다: 1,4,7,10-테트라아자사이클로도데칸-1,4,7,10-테트라아세트산 모노(N-하이드록시숙신이미드 에스터)(DOTA-NHS에스터)는 Macrocyclics사제(텍사스주 달라스에 소재)로 주문하였다. 탄산염 완충제 중의 100 배 몰 과량의 DOTA-NHS-에스터에 의한 아민 변형 RNA 센스 가닥(pH 9)을 마이크로원심분리관에 첨가하였다. 내용물을 실온에서 4시간 동안 교반시킴으로써 반응시켰다. DOTA-RNA센스 컨쥬게이트를 에탄올 침전시키고 나서, 수 중에서 재현탁하고, 미변형 안티센스 가닥으로 아닐링하여 DOTA-siRNA를 수득하였다. 모든 액체를 Chelex-100(캘리포니아주 허큘러스에 소재한 Bio-Rad)으로 전처리하여 미량의 금속 오염물질을 제거하였다. Tf-표적화 및 비표적화 siRNA 나노입자는 사이클로덱스트린 함유 폴리양이온을 사용함으로써 형성될 수 있다. 통상적으로, 나노입자는 전하비 3 (+/-) 및 siRNA 농도 0.5 g/리터로 수 중에서 형성되었다. 표적화된 나노입자의 표면에 대한 아다만탄-PEG 분자의 1%는 Tf에 의해 변형되었다(아다만탄-PEG-Tf). 나노입자를 주사용 5%(중량/용적) 글루코스 담체 용액 중에서 현탁시켰다.

[0477] Davis 등 (Nature, Vol 464, 15 April 2010)은 표적화된 나노입자 전달 시스템(임상시험 등록 번호 NCT00689065)을 사용하는 siRNA 임상시험을 수행한다. 표준 치료 요법으로 난치성인 고형암을 지니는 환자에게 30 분 정맥내 주입에 의해 21 일 주기 중 제1 일, 제3 일, 제8 일 및 제10 일에 표적화된 나노입자의 용량을 투여한다. 나노입자는 하기를 함유하는 합성 전달 시스템으로 이루어진다: (1) 선형, 사이클로덱스트린 기반 중합체(CDP), (2) 암세포의 표면 상에서 TF 수용체(TFR)에 맞물리는 나노입자 외부에 나타난 인간 트랜스페린 단백질(TF) 표적화 리간드, (3) 친수성 중합체(생물학적 유체 중에서 나노입자 안정성을 촉진하기 위해 사용되는 폴리에틸렌 글리콜(PEG)), 및 (4) RRM2의 발현을 감소시키도록 설계된 siRNA(임상에서 사용한 서열은 앞서 정의한 siR2B+5였음). TFR는 악성 세포에서 상향조절되는 것으로 알려져 있고, RRM2는 확립된 항암 표적이다. 이들 나노입자(CALAA-01로서 정의된 임상형태)는 비인간 영장류에서의 다회 투약 연구에서 잘 용인되는 것으로 나타났다. 만성 골수성 백혈병을 지니는 단일 환자에게 리포좀 전달에 의해 siRNA가 투여되었음에도 불구하고, Davis 등의 임상시험은 표적화된 전달시스템을 이용하여 siRNA를 전신으로 전달하기 위한 그리고 고형암을 지니는 환자를 치료하기 위한 초기 인간 시험이다. 표적화된 전달 시스템이 인간 종양에 기능적 siRNA의 효과적인 전달을

제공할 수 있는지 여부를 확인하기 위해, Davis 등은 3 개의 상이한 투약 코호트로부터의 3 명 환자로부터의 생검을 조자하였고; 환자 A, B 및 C는 모두 전이성 흑색종을 가졌고, 각각 18, 24 및 30 mg m^{-2} siRNA의 CALAA-01 용량을 받았다. 유사한 용량은 또한 본 발명의 CRISPR Cas 시스템을 고려할 수 있다. 본 발명의 전달은 선형, 사이클로덱스트린 기반 중합체(CDP), 암세포의 표면 상에서 TF 수용체(TFR)에 맞물리는 나노입자 외부에 나타난 인간 트랜스페린 단백질(TF) 표적화 리간드 및/또는 친수성 중합체(예를 들어, 생물학적 유체 중에서 나노입자 안정성을 촉진하기 위해 사용되는 폴리에틸렌 글리콜(PEG))을 함유하는 나노입자에 의해 달성될 수 있다.

[0478] **엑소좀**

[0479] 엑소좀은 마우스의 뇌에 짧은 간섭 (si)RNA를 전달할 수 있는 단백질 및 RNA를 수송하는 내인성 나노소수포이다. 면역원성을 감소시키기 위해, Alvarez-Erviti 등(2011, Nat Biotechnol 29: 341)은 엑소좀 생성을 위한 자기 유래 수지상 세포를 사용하였다. 뉴런 특이적 RVG 펩티드3에 융합된 엑소좀 막 단백질인 Lamp2b를 발현시키기 위한 수지상 세포를 유전자 조작함으로써 표적화를 달성하였다. 전기천공법에 의한 내인성 siRNA에 의해 정제된 엑소좀을 장입하였다. 정맥내로 주사한 RVG-표적화 엑소좀을 뇌에서의 뉴런, 미세아교세포, 희소돌기신경교에 GAPDH siRNA를 특이적으로 전달하여, 특이적 유전자 녹다운을 초래하였다. RVG 엑소좀에 대한 사전 노출은 녹다운을 약화시키지 않았고, 다른 조직에서의 비특이적 흡수는 관찰되지 않았다. 엑소좀 매개 siRNA 전달의 치료적 가능성은 알츠하이머 질환에서 치료적 표적인 BACE1의 단백질(62%) 녹다운 및 강한 mRNA(60%)에 의해 입증되었다.

[0480] 면역적으로 비활성인 엑소좀의 풀을 얻기 위해, Alvarez-Erviti 등은 상동성 주조직 적합성 복합체(MHC) 단상형을 지니는 근친교배 C57BL/6 마우스로부터의 골수를 채취하였다. 미숙한 수지상 세포는 MHC-II 및 CD86과 같은 T-세포 활성인자가 없는 다량의 엑소좀을 생성하기 때문에, Alvarez-Erviti 등은 7일 동안 과립구/대식세포-집락자극인자(GM-CSF)를 지니는 수지상 세포에 대해 선택하였다. 잘 확립된 초원심분리 프로토콜을 사용하여 다음날 배양 상청액으로부터 엑소좀을 정제하였다. 생성된 엑소좀은 물리적으로 균질하였고, 크기 분포는 나노입자 트래킹 분석(NTA) 및 전자 현미경에 의해 결정하여 직경 80 nm에서 최고였다. Alvarez-Erviti 등은 10^6 개 세포당 6 내지 12 μg 엑소좀(단백질 농도를 기준으로 측정)을 얻었다.

[0481] 다음에, Alvarez-Erviti 등은 나노규모 적용에 적합한 전기천공법 프로토콜을 사용하여 외인성 화물을 지니는 장입 변형된 엑소좀의 가능성을 조사하였다. 나노미터 규모에서 막 입자에 대한 전기천공법은 제대로 특성규명되지 않기 때문에, 비특이적 Cy5-표지 siRNA는 전기천공법 프로토콜의 경험적 최적화에 대해 사용되었다. 캡슐화된 siRNA의 양을 엑소좀의 초원심분리 및 용해 후에 분석하였다. 400 V 및 125 μF에서의 전기천공법은 siRNA의 가장 큰 보유를 초래하였고, 모든 후속 실험에 대해 사용하였다.

[0482] Alvarez-Erviti 등은 정상 C57BL/6 마우스에 대해 150 μg의 RVG 엑소좀 중에서 캡슐화된 150 μg의 각각 BACE1 siRNA를 투여하였고, 4 개의 대조군(미처리 마우스, RVG 엑소좀만을 주사한 마우스, 생체내 양이온성 리포좀 시약에 복합체화된 BACE1 siRNA를 주사한 마우스 및 RVG-9R(siRNA에 정전기적으로 결합된 9 개의 D-알라닌에 컨쥬게이트된 RVG 펩타이드)에 복합체화된 BACE1 siRNA를 주사한 마우스)에 대한 녹다운 효율을 비교하였다. 투여 3일 후 피질 조직 샘플을 분석하였고, siRNA-RVG-9R-처리 마우스와 siRNARVG 엑소좀 처리 마우스 둘 다에서 BACE1 mRNA 수준 (66% [+ 또는 -] 15%, $P < 0.001$ 및 61% [+ 또는 -] 13% 각각, $P < 0.01$)에서 상당한 감소로부터 초래되는 상당한 단백질 녹다운(45%, $P < 0.05$, 대 62%, $P < 0.01$)을 관찰하였다. 게다가, 본 발명자들은 RVG-엑소좀 처리 동물에서 알츠하이머 병리에서의 아밀로이드 플라크의 주요 구성성분인 총 [베타]-아밀로이드 1 내지 42 수준에서 상당한 감소(55%, $P < 0.05$)를 입증하였다. 관찰한 감소는 BACE1 저해제의 정맥내 주사 후 정상 마우스에서 입증된 β-아밀로이드 1 내지 40 감소보다 더 컸다. Alvarez-Erviti 등은 BACE1 절단 산물에 대해 cDNA 단부의 5'-빠른 증폭(RACE)을 수행하였는데, 이는 siRNA에 의한 RNAi-매개 녹다운의 증거를 제공하였다.

[0483] 최종적으로, Alvarez-Erviti 등은 siRNA-RVG 엑소좀이 IL-6, IP-10, TNFα 및 IFN-α 혈청 농도를 평가함으로써 생체내 면역 반응을 유발하였는지 여부를 조사하였다. siRNA-RVG 엑소좀 처리 후에, 모든 사이토카인에서의 중요하지 않은 변화는 IL-6 분비를 잠재적으로 자극한 siRNA-RVG-9R과 대조적으로 siRNA-트렌스펙션 시약 처리와 유사하게 등록되었는데, 이는 엑소좀 처리의 면역적으로 비활성인 프로파일을 확인하였다. 엑소좀이 단지 20%의 siRNA를 캡슐화하는 것을 고려해 볼 때, RVG-엑소좀에 의한 전달은 비슷한 mRNA 녹다운으로서 RVG-9R 전달보다 더 효율적인 것으로 나타나며, 더 큰 단백질 녹다운은 대응하는 면역 자극 수준이 없는 5 배 더 적은 siRNA에 의해 달성되었다. 이 실험은 신경퇴행성 질환과 관련된 유전자의 장기간 침묵에 가능하게 적합한 RVG-엑소좀 기법의 치료 가능성을 입증하였다. Alvarez-Erviti 등의 엑소좀 전달 시스템은 치료 표적, 특히 신경퇴

행성 질환에 본 발명의 CRISPR-Cas 시스템을 전달하기 위해 적용될 수 있다. 약 100 내지 1000 mg의 RVG 엑소좀에서 캡슐화된약 100 내지 1000 mg의 CRISPR Cas의 투약량은 본 발명에 대해 고려될 수 있다.

[0484] El-Andaloussi 등(Nature Protocols 7,2112-2126(2012))은 배양 세포로부터 유래된 엑소좀이 시험관내 및 생체내 siRNA의 전달을 위해 이용될 수 있는 방법을 개시한다. 이 프로토콜은 펩티드 리간드와 융합된 엑소좀 단백질을 포함하는 발현 벡터의 트렌스펙션을 통한 표적화된 엑소좀의 생성을 처음으로 기재한다. 다음에, El-Andaloussi 등은 트렌스펙션 세포 상청액으로부터 엑소좀을 정제하고 특성규명하는 방법을 설명한다. 다음에, El-Andaloussi 등은 siRNA를 엑소좀에 장입하기 위한 중대한 단계를 상술한다. 최종적으로, El-Andaloussi 등은 마우스 뇌에서 시험관내 및 생체내 siRNA를 효율적으로 전달하기 위해 엑소좀을 사용하는 방법을 약술한다. 엑소좀-매개 siRNA 전달이 기능성 분석 및 영상에 의해 평가되는 예상된 결과의 예가 또한 제공된다. 전체 프로토콜은 약 3 주가 걸린다. 본 발명에 따른 전달 또는 투여는 자기 유래 수지상 세포로부터 생성된 엑소좀을 사용하여 수행될 수 있다.

[0485] 다른 구현예에서, Wahlgren 등의 혈장 엑소좀(Nucleic Acids Research, 2012, Vol. 40, No. 17 e130)이 고려된다. 엑소좀은 수지상 세포(DC), B 세포, T 세포, 비만 세포, 상피 세포 및 종양 세포를 포함하는 다수의 세포 유형에 의해 생성된 나노 크기 소수포(30 내지 90 nm 크기)이다. 이들 소수포는 후기 엔도솜의 안쪽으로 향하는 출아에 의해 형성되며, 이어서 혈장막과 융합시 세포밖 환경으로 방출된다. 엑소좀은 세포 간의 RNA를 자연적으로 운반하기 때문에, 이 특성은 유전자 요법에서 유용할 수 있다.

[0486] 혈장으로부터의 엑소좀은 20 분 동안 900 g에서 버피 코트의 원심분리로 혈장을 단리시킨 다음에 세포 상청액을 채취하고 10 분 동안 300 g에서 그리고 30 분 동안 16 500 g에서 원심분리하여 세포를 제거한 후 0.22 mm 필터를 통해 여과시켰다. 엑소좀은 120 000 g에서 70 분 동안 초원심분리에 의해 펠렛화한다. siRNA의 엑소좀 내로의 화학적 트렌스펙션을 RNAi 인간/마우스 스타터 키트(Mouse Starter Kit)(독일 힐덴에 소재한 Quiagen)에서 제조업자의 설명서에 따라 수행한다. siRNA를 2 mmol/㎖의 최종 농도로 100 ㎖ PBS에 첨가한다. HiPerFect 트렌스펙션 시약을 첨가한 후에, 혼합물을 실온에서 10분 동안 인큐베이션시킨다. 과량의 미셀을 제거하기 위해, 알데하이드/설페이트 라텍스 비드를 사용하여 엑소좀을 재단리시킨다. 엑소좀 내로 CRISPR Cas의 화학적 트렌스펙션은 siRNA와 유사하게 수행될 수 있다. 엑소좀은 단핵구와 공동배양될 수 있고, 림프구는 건강한 공여체의 말초혈액으로부터 단리되었다. 따라서, 엑소좀 함유 CRISPR Cas는 인간의 단핵구 및 림프구에 도입될 수도 있고 인간에게 자가 조직으로 재도입될 수도 있다. 따라서, 본 발명에 따른 전달 또는 투여는 혈장 엑소좀을 사용하여 수행될 수 있다.

[0487] **<u>리포좀</u>**

[0488] 본 발명에 따른 전달 또는 투여는 리포좀에 의해 수행될 수 있다. 리포좀은 내부 수성 격막을 둘러싸는 단일 또는 다중 라멜라 지질 이중층 및 상대적으로 불침투성인 외부 친유성 인지질 이중층으로 구성된 구체의 소수포 구조이다. 리포좀은 그들이 생물학적적합성, 비독성이기 때문에 약물 전달 담체로서 상당한 관심을 획득하였고, 친수성과 친유성 약물 분자를 둘 다 전달할 수 있으며, 혈장 효소에 의한 분해로부터 그들의 화물을 보호하고, 생물학적 막 및 혈액 뇌 장벽(BBB)을 가로질러 그들의 장입물을 수송한다(예를 들어, 검토를 위해 문헌[Spuch and Navarro, Journal of Drug Delivery, vol. 2011, Article ID 469679, 12 페이지], 2011. doi:10.1155/2011/469679 참조).

[0489] 리포좀은 몇몇 상이한 유형의 지질로부터 만들어질 수 있지만; 그러나, 인지질은 약물 담체로서 리포좀을 생성하는데 가장 통상적으로 사용된다. 지질 필름이 수용액과 혼합될 때 리포좀 형성은 자발적이지만, 또한 균질기, 초음파처리기 또는 압출 장치를 사용하는 것에 의한 진탕 형태로 힘을 적용함으로써 더 신속히 처리될 수 있다(예를 들어, 검토를 위해 문헌[Spuch and Navarro, Journal of Drug Delivery, vol. 2011, Article ID 469679, 12 페이지], 2011. doi:10.1155/2011/469679 참조).

[0490] 그들의 구조 및 특성을 변형시키기 위해 몇몇 다른 첨가제가 리포좀에 첨가될 수 있다. 예를 들어, 콜레스테롤 또는 스핑고미엘린이 리포좀 혼합물에 첨가되어 리포좀 구조를 안정화시키고 리포좀 내부 화물의 누출을 방지한다. 추가로, 리포좀은 수소화된 난황 포스파티딜콜린 또는 난황 포스파티딜콜린, 콜레스테롤 및 디세틸 포스페이트로부터 제조되고, 그들의 평균 소수포 크기는 약 50 및 100 nm로 조절되었다. (예를 들어, 검토를 위해 문헌[Spuch and Navarro, Journal of Drug Delivery, vol. 2011, Article ID 469679, 12 페이지], 2011. doi:10.1155/2011/469679 참조).

[0491] 통상적인 리포좀 제형은 주로 천연 인지질 및 지질, 예컨대 1,2-디스테아로릴-sn-글리세로-3-포스파티딜 콜린

(DSPC), 스핑고미엘린, 난황 포스파티딜콜린 및 모노시알로강글리오사이드를 주로 포함한다. 이 제형은 인지질만으로 구성되기 때문에, 리포좀 제형은 다수의 도전, 그 중에서도 혈장 내 불안정성과 접하게 되었다. 특이적으로 지질막의조작에서 이들 도전을 극복하기 위한 몇몇 시도가 있었다. 이들 시도 중 하나는 콜레스테롤의 조작에 집중되어 있다. 전통적인 제형에 대한 콜레스테롤의 첨가는 캡슐화된 생활성 화합물의 혈장 내로의 빠른 방출을 감소시키거나 또는 1,2-디올레오일-sn-글리세로-3-포스포에탄올아민(DOPE)은 안정성을 증가시킨다(예를 들어, 검토를 위해 문헌[Spuch and Navarro, Journal of Drug Delivery, vol. 2011, Article ID 469679, 12 페이지], 2011. doi:10.1155/2011/469679 참조).

[0492] 특히 유리한 구현예에서, 트로이목마 리포좀(또한 분자 트로이목마로서 알려짐)이 바람직하며, 프로토콜은 http://cshprotocols.cshlp.org/content/2010/4/pdb.prot5407.long에서 찾을 수 있다. 이들 입자는 혈관내 주사 후 전체 뇌에 이식유전자를 전달하게 한다. 제한에 의해 구속되는 일 없이, 표면에 컨쥬게이트된 특이적 항체를 지니는 중성 지질 입자는 내포작용을 통해 혈액 뇌 장벽을 가로지르는 것으로 믿어진다. 본 발명자들은 혈관내 주사를 통해 뇌에 뉴클레아제의 CRISPR 패밀리를 전달하기 위해 트로이목마 리포좀을 이용하는 것을 상정하는데, 이는 배아 조작에 대한 필요 없이 전체 뇌 유전자이식 동물을 허용한다. 약 1 내지 5 g의 핵산 분자, 예를 들어, DNA 또는 RNA가 리포좀 내 생체내 투여에 대해 고려될 수 있다.

[0493] 다른 구현예에서, CRISPR Cas 시스템은 안정한 핵산 지질 입자(SNALP)와 같이 리포좀에서 투여될 수 있다(예를 들어, 문헌[Morrissey et al., Nature Biotechnology, Vol. 23, No. 8, August 2005] 참조). SNALP에서 표적화된 특이적 CRISPR Cas의 약 1, 3 또는 5 mg/kg/일의 매일의 정맥내 주사가 고려된다. 매일의 치료는 약 3 일에 걸칠 수 있고, 이어서 약 5 주 동안 매주일 수 있다. 다른 구현예에서, abpit 1 또는 2.5 mg/kg의 용량에서 정맥내 주사에 의해 투여되는 구체적 CRISPR Cas 캡슐화된 SNALP가 또한 고려된다(예를 들어, 문헌[Zimmerman et al., Nature Letters, Vol. 441, 4 May 2006] 참조). SNALP 제형은 지질 3-N-[(w메톡시폴리(에틸렌 글리콜) 2000) 카바모일] -1,2-디미리스틸옥시-프로필아민 (PEG-C-DMA), 1,2-디리놀레일옥시-N,N-디메틸-3-아미노프로판(DLinDMA), 1,2-디스테아로일-sn-글리세로-3-포스포콜린(DSPC) 및 콜레스테롤을 2:40:10:48 몰 백분율 비로 함유할 수 있다(예를 들어, 문헌[Zimmerman et al., Nature Letters, Vol. 441, 4 May 2006] 참조).

[0494] 다른 구현예에서, 안정한 핵산 지질 입자(SNALP)는 고도로 혈관화된 HepG2-유래 간 종양에 대해 효과적인 전달 분자가 되는 것으로 증명되었지만, 불량하게 혈관화된 HCT-116 유래 간 종양에서는 그렇지 않다(예를 들어, Li, Gene Therapy (2012) 19, 775-780). SNALP 리포좀은 콜레스테롤/D-Lin-DMA/DSPC/PEG-C-DMA의 25:1 지질/siRNA 비 및 48/40/10/2 몰비를 사용하여 디스테아로일포스파티딜콜린(DSPC), 콜레스테롤 및 siRNA로 D-Lin-DMA 및 PEG-C-DMA를 제형화함으로써 제조될 수 있다. 얻어진 SNALP 리포좀은 크기가 약 80 내지 100 nm이다.

[0495] 또 다른 구현예에서, SNALP는 합성 콜레스테롤(미국 미조리주 세인트루이스에 소재한 Sigma-Aldrich), 디팔미토일포스파티딜콜린(미국 앨라배마주 앨러배스터에 소재한 Avanti Polar Lipids), 3-N-[(w-메톡시 폴리(에틸렌 글리콜)2000)카바모일]-1,2-디미레스틸옥시프로필아민, 및 양이온성 1,2-디리놀레일옥시-3-N,N디메틸아미노프로판(예를 들어, 문헌[Geisbert et al., Lancet 2010; 375: 1896-905] 참조)를 포함할 수 있다. 예를 들어, 볼루스 정맥재 주입으로서 투여되는 용량 당 약 2 mg/kg 총 CRISPR Cas의 투약량이 고려될 수 있다.

[0496] 또 다른 구현예에서, SNALP는 합성 콜레스테롤(Sigma-Aldrich), 1,2-디스테아로일-sn-글리세로-3-포스포콜린(DSPC; Avanti Polar Lipids Inc.), PEG-cDMA, 및 1,2-디리놀레일옥시-3-(N;N-디메틸)아미노프로판(DLinDMA)(예를 들어, 문헌[Judge, J. Clin. Invest. 119:661-673 (2009)] 참조)을 포함할 수 있다. 생체내 연구를 위해 사용되는 제형은 약 9:1의 최종 지질/RNA 질량비를 포함할 수 있다.

[0497] RNAi 나노의학의 안전성 프로파일은 Alnylam Pharmaceuticals의 Barros 및 Gollob에 의해 검토되었다(예를 들어, 문헌[Advanced Drug Delivery Reviews 64 (2012) 1730-1737] 참조). 안정항 핵산 지질 입자(SNALP)는 4 가지 상이한 지질 - 낮은 pH에서 양이온성인 이온화 가능한 지질(DLinDMA), 중성 헬퍼 지질, 콜레스테롤 및 확산성 폴리 에틸렌 글리콜(PEG)-지질을 포함한다. 입자는 직경이 대략 80 nm이고, 생리학적 pH에서 하전 중성이다. 제형 동안, 이온화가능한 지질은 입자 형성 동안 음이온성 siRNA와 지질을 축합하는 작용을 한다. 점점 더 산성인 엔도솜 조건 하에 양으로 하전될 때, 이온화가능한 지질은 또한 siRNA의 세포질 내로의 방출을 가능하게 하는 엔도솜 막과 SNALP의 융합을 매개한다. PEG-지질은 입자를 안정화하고, 제형 동안 응집을 감소시키며, 후속적으로 약동학적 특성을 개선시키는 외부의 중성 친수성을 제공한다.

[0498] 지금까지, SNALPsiRNA 제형을 사용하여 2 가지 임상 프로그램을 개시하였다. Tekmira Pharmaceuticals는 최근에 상승된 LDL 콜레스테롤을 지니는 성인 지원자에서 SNALP-ApoB의 I상 단일 용량 연구를 완료하였다. ApoB는 간 및 공장에서 대부분 발현되며, VLDL 및 LDL의 어셈블리 및 분비에 필수적이다. ApoB는 또한 출원인의

CRISPR-Cas 시스템에 의해 성공적으로 표적화되며, 실시예 37 내지 38을 참조하라. 17 명의 대상체에 단일 용량의 SNALP-ApoB를 투여하였다(7 용량 수준에 걸친 용량 상승). 간 독성의 증거는 없었다(전임상 연구에 기반한 잠재적 용량 제한 독성으로서 예상). 가장 높은 용량에서 (둘 중) 한 명의 대상체는 면역계 자극과 일치되는 감기-유사 증상을 경험하였고, 시험을 결론 짓는 결정을 하였다.

[0499] Alnylam Pharmaceuticals는 상기 기재한 SNALP 기법을 사용하는 유사하게 진보된 ALN-TTR01을 가지며, TTR 아밀로이드증(ATTR)을 치료하기 위한 돌연변이체와 야생형 TTR 둘 다의 간세포 생성을 표적화한다. 3 가지 ATTR 증후군이 기재되었다: 가족성 아밀로이드성 다발신경병증(FAP) 및 가족성 아밀로이드성 심근증(FAC) - 둘 다 TTR에서의 상염색체 우성 돌연변이에 의해 야기됨; 및 야생형 TTR에 의해 야기되는 노인 전신성 아밀로이드증(SSA). ALN-TTR01의 위약 제어, 단일 용량 상승 I상 시험을 ATTR을 지니는 환자에서 최근에 완료하였다. ALN-TTR01을 0.01 내지 1.0 mg/kg(siRNA 기준)의 용량 범위 내에서 31 명의 환자(연구 약물에 의한 23 명 및 위약에 의한 8 명)에 대한 15 분 IV 주입으로서 투여하였다. 간 기능 시험에서 상당한 증가 없이 치료는 잘 용인되었다. 주입 관련 반응은 0.4 mg/kg 이상에서 23 명의 환자 중 3 명에서 주목되었고; 모두 주입의 늦춤에 반응하였으며, 모두 연구일에 계속되었다. 혈청 사이토카인 IL-6, IP-10 및 IL-1ra의 최소 및 일시적 상승은 1 mg/kg의 가장 높은 용량에서 2 명의 환자에서 주목되었다(전임상 및 NHP 연구로부터 참여). 혈청 TTR의 저하, ALN-TTR01의 예상된 약동학적 효과를 1 mg/kg에 관찰하였다.

[0500] 또 다른 구현예에서, SNALP는 양이온성 지질, DSPC, 콜레스테롤 및 PEG-지질을 가용화함으로써 만들어질 수 있고, 각각 40:10:40:10의 몰비로 에탄올 중에서 가용화되었다(문헌[Semple et al., Nature Niotechnology, Volume 28 Number 2 February 2010, pp. 172-177] 참조). 지질 혼합물을 30% (vol/vol) 및 6.1 mg/㎖의 최종 에탄올 및 지질 농도로 각각 혼합하면서 수성 완충제(50 mM 시트르산염, pH 4)에 첨가하였고, 2 분 동안 22℃에서 평형상태로 만든 후 압출시켰다. 수화된 지질을 역동적 광 산란에 의해 결정된 바와 같은 소수포 직경 70 내지 90 nm까지 Lipex Extruder(Northern Lipids)를 사용하여 22℃에서 2층 80 nm 기공 크기 필터(Nuclepore)를 통해 압출하였고, 얻었다. 이는 일반적으로 1 내지 3 회 통과를 필요로 하였다. siRNA(50 mM 시트르산염 중에서 가용화됨, 30% 에탄올을 함유하는 pH 4 수용액)를 혼합하면서 약 5 ㎖/분의 속도로 사전 평형상태로 만든(35℃) 소수포에 첨가하였다. 0.06(wt/wt)의 최종 표적 siRNA/지질비가 도달될 후, 혼합물을 35℃에서 추가 30 분 동안 인큐베이션 시켜 siRNA를 소수포 재조직화 및 캡슐화시켰다. 이어서, 에탄올을 제거하고, 투석 또는 접선유동 정용여과에 의해 외부 완충제를 PBS로 대체하였다(155 mM NaCl, 3 mM Na2HPO4, 1 mM KH2PO4, pH 7.5). siRNA를 제어된 계단식 희석 방법 공정을 사용하여 SNALP에서 캡슐화하였다. KC2-SNALP의 지질 구성성분은 57.1:7.1:34.3:1.4의 몰비로 사용한 DLin-KC2-DMA(양이온성 지질), 디팔미토일포스파티딜콜린(DPPC; Avanti Polar Lipids), 합성 콜레스테롤(Sigma) 및 PEG-C-DMA이었다. 장입한 입자의 형성시, SNALP를 PBS에 대해 투석하였고, 사용 전 0.2 μm 필터를 통해 멸균여과시켰다. 평균 입자 크기는 75 내지 85 nm였고, siRNA의 90 내지 95%는 지질 입자 내에서 캡슐화되었다. 생체내 시험을 위해 사용한 제형 내 최종 siRNA/지질비는 약 0.15(wt/wt)였다. 인자 VII siRNA를 함유하는 LNP-siRNA 시스템을 사용 직전에 멸균 PBS 중에서 적절한 농도로 희석시켰고, 제형을 총 10 ㎖/kg 용적으로 옆 꼬리 정맥을 통해 정맥내로 투여하였다. 이 방법은 본 발명의 CRISPR Cas 시스템에 대해 추론될 수 있다.

[0501] **기타 지질**

[0502] 다른 양이온성 지질, 예컨대 아미노 지질 2,2-디리놀레일-4-디메틸아미노에틸-[1,3]-디옥솔란 (DLin-KC2-DMA)은 siRNA와 유사한 CRISPR Cas를 캡슐화하기 위해 이용될 수 있다(예를 들어, 문헌[Jayaraman, Angew. Chem. Int. Ed. 2012, 51, 8529 - 8533] 참조). 다음의 지질 조성물을 이용하여 사전 형성한 소수포가 고려될 수 있다: 40/10/40/10의 몰비 및 대략0.05(w/w) FVII siRNA/총 지질비로 각각 아미노지질, 디스테아로일포스파티딜콜린(DSPC), 콜레스테롤 및 (R)-2,3-비스(옥타데실옥시) 프로필-1-(메톡시 폴리(에틸렌 글리콜)2000)프로필카바메이트(PEG-지질). 70 내지 90 nm 및 낮은 다분산지수 0.11_0.04(n=56)의 범위에서 좁은 입자 크기 분포를 보장하기 위해, 입자는 CRISPR Cas RNA를 첨가하기 전에 80 nm 막을 통해 3 회까지 압출될 수 있다. 고도로 강한 아미노 지질 16(생체내 활성을 활성화하도록 추가로 최적화될 수 있는 4 가지 지질 구성성분 16, DSPC, 콜레스테롤 및 PEG-지질의 몰비(50/10/38.5/1.5))을 함유하는 입자가 사용될 수 있다.

[0503] Michael S D Kormann 등("Expression of therapeutic proteins after delivery of chemically modified mRNA in mice: Nature Biotechnology, Volume:29, Pages: 154-157 (2011) Published online 09 January 2011)은 RNA를 전달하기 위한 지질 외피의 사용을 기재한다. 지질 외피의 사용은 또한 본 발명에서 바람직하다.

[0504] 다른 구현예에서, 지질은 본 발명의 CRISPR Cas 시스템에 의해 제형화되어 지질 나노입자(LNP)를 형성할 수 있

다. 지질은 DLin-KC2-DMA4, C12-200 및 공지질 디스테로일포스파티딜 콜린, 콜레스테롤을 포함하지만, 이들로 제한되지 않으며, PEG-DMG은 자발적 소수포 형성 절차를 사용하여 siRNA 대신 CRISPR Cas를 이용하여 제형화될 수 있다(예를 들어, 문헌[Novobrantseva, Molecular Therapy-Nucleic Acids (2012) 1, e4; doi:10.1038/mtna.2011.3] 참조). 구성성분 몰비는 약 50/10/38.5/1.5(DLin-KC2-DMA 또는 C12-200/디스테로일포스파티딜 콜린/콜레스테롤/PEG-DMG)일 수 있다. 최종 지질:siRNA 중량비는 DLin-KC2-DMA 및 C12-200 지질 나노입자(LNPs)의 경우에 각각 약 12:1 및 9:1일 수 있다. 제형은 90% 초과의 포착 효율을 지니는 약 80 nm의 평균 입자 직경을 가질 수 있다. 3 mg/kg 용량이 고려될 수 있다.

[0505] Tekmira는 LNP 및 LNP 제형의 다양한 양태와 관련된 미국 및 해외에서의 대략 95개의 특허 패밀리의 포트폴리오를 가지며(예를 들어, 미국 특허 제7,982,027호; 제7,799,565호; 제8,058,069호; 제8,283,333호; 제7,901,708호; 제7,745,651호; 제7,803,397호; 제8,101,741호; 제8,188,263호; 제7,915,399호; 제8,236,943호 및 제7,838,658호 및 유럽 특허 제1766035호; 제1519714호; 제1781593호 및 제1664316호), 이들 모두는 본 발명에 대해 사용되고/사용되거나 적합하게 될 수 있다.

[0506] 단백질, 단백질 전구체 또는 단백질 또는 단백질 전구체의 부분적으로 또는 완전히 가공된 형태를 암호화할 수 있는 변형된 핵산 분자를 포함하는 조성물 제형의 양태에 관한 미국 출원 공개 제20130252281호 및 제20130245107호 및 제20130244279호(Moderna Therapeutics 출원)에서 추가로 기재된 것과 같은 PLGA 마이크로스피어에서 캡슐화된 CRISPR Cas 시스템이 전달될 수 있다. 제형은 몰비 50:10:38.5:1.5-3.0(양이온성 지질:융합생성 지질:콜레스테롤:PEG 지질)을 가질 수 있다. PEG 지질은 PEG-c-DOMG, PEG-DMG로부터 선택될 수 있지만, 이들로 제한되지 않는다. 융합생성 지질은 DSPC일 수 있다. 또한 Schrum 등의 미국 출원 공개 제20120251618호(Delivery and Formulation of Engineered Nucleic Acids)를 참조한다.

[0507] 나노머 기술은 저분자량 소수성 약물, 펩타이드 및 핵산 기반 치료제(플라스미드, siRNA, miRNA)를 포함하는 광범위 치료제에 대한 생체이용가능성 도전을 처리한다. 기술이 명확한 이점을 입증한 구체적 투여 경로는 경구 경로, 혈액뇌장벽에 걸친 수송, 고형 종양에 대한 전달뿐만 아니라 눈에 대한 전달을 포함한다. 예를 들어, 문헌[Mazza et al., 2013, ACS Nano. 2013 Feb 26;7(2):1016-26; Uchegbu and Siew, 2013, J Pharm Sci. 102(2):305-10 및 Lalatsa et al., 2012, J Control Release. 2012 Jul 20;161(2):523-36] 참조.

[0508] 미국 특허 공개 제20050019923호는 폴리뉴클레오티드 분자, 펩타이드 및 폴리펩타이드 및/또는 약제학적 작용제(agent)와 같은 생활성 분자를 포유류 신체에 전달하기 위한 양이온성 덴드리머를 기재한다. 덴드리머는, 예를 들어 간, 비장, 폐, 신장 또는 심장에 생활성 분자의 전달을 표적화하는데 적합하다. 덴드리머는 단순한 분지 단량체 단위로부터 단계적 방식으로 제조된 합성 3-차원 거대분자이며, 이의 특성 및 기능성은 용이하게 제어되고 변할 수 있다. 덴드리머는 다기능성 코어에 대해(합성에 대한 상이한 접근), 또는 다기능성 코어 쪽으로(합성에 대한 수렴적 접근) 빌딩 블록의 반복된 첨가로부터 합성되고, 빌딩 블록의 3차원 껍질의 각각의 첨가는 덴드리머의 더 고차 생성의 형성을 야기한다. 폴리프로필렌이민 덴드리머는 디아미노부탄 코어로부터 출발하여 이것에 아크릴로니트릴의 이중 마이클 첨가(Michael addition)에 의해 아미노기 수의 2 배로 1차 아민에 첨가된 후, 니트릴이 수소화된다. 이는 아미노기의 배증을 초래한다. 폴리프로필렌이민 덴드리머는 100% 양성자화 가능한 질소 및 64 개까지의 말단 아미노기를 함유한다(5세대, DAB 64). 양성자화 가능한기는 보통 중성 pH에서 양성자를 수용할 수 있는 아민기이다. 유전자 전달제로서 덴드리머의 용도는 컨쥬게이팅 단위로서 각각 아민/아미드 또는 N--P(O_2)S의 혼합물과 함께 폴리아미도아민 및 인 함유 화합물의 용도에 크게 집중되어 있었는데 유전자 전달을 위한 더 낮은 세대 폴리프로필렌이민 덴드리머의 용도에 대해 보고한 연구는 없었다. 폴리프로필렌이민 덴드리머는 또한 주위 아미노기에 의해 화학적으로 변형될 때 약물 전달을 위해 그리고 게스트 분자의 캡슐화를 위해 pH 민감 제어 방출 시스템으로서 연구되었다. 세포독성 및 DNA와 폴리프로필렌이민 덴드리머의 상호작용뿐만 아니라 DAB 64의 트렌스펙션 효능이 또한 연구되었다.

[0509] 미국 특허 공개 제20050019923호는 더 조기의 보고와 상반되게 양이온성 덴드리머, 예컨대 폴리프로필렌이민 덴드리머가 유전물질과 같은 생활성 분자의 표적화된 전달에서 사용하기 위한 적합한 특성, 예컨대 특이적 표적화 및 낮은 독성을 나타낸다는 관찰에 기반한다. 추가로, 양이온성 덴드리머의 유도체는 또한 생활성 분자의 표적화된 전달을 위한 적합한 특성을 나타낸다. 또한, 양이온성 폴리아민 중합체 및 덴드리머 중합체를 포함하는 "다양한 중합체"를 개시하는 미국 특허 공개 제20080267903호의 생활성 중합체는 항증식 활성을 갖는 것으로 나타나며, 따라서 신생물 및 종양과 같은 바람직하지 않은 세포 증식, 염증 장애(자가면역 장애를 포함), 건선 및 아테롬성 동맥 경화증을 특징으로 하는 장애의 치료에 유용할 수 있다. 중합체는 단독으로 활성제로서 또는 다른 치료제용 전달 비히클, 예컨대 유전자 요법을 위한 약물 분자 또는 핵산으로서 사용될 수 있다. 이러한 경우

에, 중합체 자체의 고유 항종양 활성은 전달될 작용제의 활성을 보완할 수 있다.

[0510] **과급 단백질**

[0511] 과급 단백질은 보통 높은 양 또는 음의 순 이론적 전하를 지니는 유전자 조작된 또는 자연적으로 생기는 단백질의 부류이다. 초 음으로 하전된 단백질과 초 양으로 하전된 단백질은 둘 다 열적으로 또는 화학적으로 유도된 응집을 견뎌내는 현저한 능력을 나타낸다. 초 양으로 하전된 단백질은 또한 포유류 세포를 침투할 수 있다. 플라스미드 DNA, siRNA 또는 다른 단백질과 같은 이들 단백질과 화물을 연관시키는 것은 시험관내와 생체내 둘 다에서 포유류 세포 내로 이들 거대분자의 기능성 전달을 가능하게 할 수 있다. David Liu 연구소는 2007년에 초 하전 단백질의 생성 및 특성규명을 보고하였다(Lawrence et al., 2007, Journal of the American Chemical Society 129, 10110-10112).

[0512] siRNA 및 플라스미드 DNA의 포유류 세포 내로의 비바이러스 전달은 연구와 치료적 적용 둘 다에 대해 가치있다(Akinc et al., 2010, Nat. Biotech. 26, 561-569). 정제된 +36 GFP 단백질(또는 다른 초 양으로 하전된 단백질)은 적절한 무혈청 배지에서 siRNA와 혼합되고, 세포에 첨가 전 복합체화되게 한다. 이 단계에서 혈청의 포함은 초 하전 단백질-siRNA 복합체의 형성을 저해하고, 치료 효능을 감소시킨다. 다음의 프로토콜은 다양한 세포주에 대해 효과적이 되는 것을 발견하였다(McNaughton et al., 2009, Proc. Natl. Acad. Sci. USA 106, 6111-6116). 그러나, 단백질 및 siRNA의 용량을 변화시키는 파일럿 실험은 특이적 세포주의 절차를 최적화하도록 수행되어야 한다.

[0513] (1) 치료 하루 전, 48-웰 플레이트에서 웰당 1 x 10^5 개를 플레이팅.

[0514] (2) 치료일에, 최종 농도 200 nM로 무혈청 배지에서 정제된 þ36 GFP 단백질을 희석시킴. siRNA를 50 nM의 최종 농도로 첨가함. 교반으로 혼합하고 실온에서 10 분 동안 인큐베이션시킴.

[0515] (3) 인큐베이션 동안, 세포로부터의 배지를 흡입하고, PBS로 1회 세척함.

[0516] (4) þ36 GFP 및 siRNA의 인큐베이션 후에, 단백질-siRNA 복합체를 세포에 첨가함.

[0517] (5) 37 C에서 4 시간 동안 복합체와 함께 세포를 인큐베이션시킴.

[0518] (6) 인큐베이션 후에, 배지를 흡입하고, 20 U/㎖ 헤파린 PBS로 3 회 세척. 녹다운에 대한 분석에 따라서 추가 48 시간 또는 그 이상 동안 혈청-함유 배지와 함께 세포를 인큐베이션시킴.

[0519] (7) 면역블롯, qPCR, 표현형 분석 또는 다른 적절한 방법에 의해 분석함.

[0520] 세포 범위에서 +36 GFP가 효과적인 플라스미드 전달 시약인 것으로 발견되었다. 플라스미드 DNA는 siRNA보다 더 큰 화물이기 때문에, 비례해서 더 많은 +36 GFP 단백질이 효과적으로 플라스미드와 복합체화하는데 필요로 된다. 효과적인 플라스미드 전달을 위해, 출원인은 인플루엔자 바이러스 혈구응집소 단백질로부터 유래된 공지된 엔도솜 파괴 펩타이드인 C-말단의 HA2 펩타이드 태그를 보유하는 +36 GFP의 변이체를 개발하였다. 다음의 프로토콜은 다양한 세포에서 효과적이었지만, 상기와 같이 플라스미드 DNA 및 초 하전 단백질은 특이적 세포주 및 전달 적용분야에 최적화된다는 것이 알려져 있다.

[0521] (1) 치료 하루 전, 48-웰 플레이트에서 1 x 10^5개 세포/웰을 플레이팅.

[0522] (2) 치료일에, 최종 농도 2 mM로 무혈청 배지에서 정제된 +36 GFP 단백질을 희석시킴. 1 ㎎의 플라스미드 DNA를 첨가. 교반으로 혼합하고 실온에서 10 분 동안 인큐베이션시킴.

[0523] (3) 인큐베이션 동안, 세포로부터의 배지를 흡입하고, PBS로 1 회 세척함.

[0524] (4) +36 GFP 및 플라스미드 DNA의 인큐베이션 후에, 단백질-DNA 복합체를 세포에 부드럽게 첨가함.

[0525] (5) 37 C에서 4 시간 동안 복합체와 함께 세포를 인큐베이션시킴.

[0526] (6) 인큐베이션 후에, 배지를 흡입하고, PBS로 세척. 혈청-함유 배지 내 세포를 인큐베이션시키고 추가 24 내지 48 시간 동안 인큐베이션시킴.

[0527] (7) 적절하다면 플라스미드 전달(예를 들어, 플라스미드-구동 유전자 발현에 의해)을 분석함.

[0528] 또한, 예를 들어, 문헌[McNaughton et al., Proc. Natl. Acad. Sci. USA 106, 6111-6116 (2009); Cronican et al., ACS Chemical Biology 5, 747-752 (2010); Cronican et al., Chemistry & Biology 18, 833-838 (2011);

Thompson et al., Methods in Enzymology 503, 293-319 (2012); Thompson, D.B., et al., Chemistry & Biology 19 (7), 831-843 (2012)] 참조. 본 발명의 CRISPR Cas 시스템의 전달을 위해 초 하전 단백질의 방법을 사용 및/또는 채택할 수 있다.

[0529] **세포 투과 펩티드**

[0530] 또 다른 구현예에서, 세포 침투 펩티드(CPP)가 CRISPR Cas 시스템의 전달에 고려된다. CPP는 다양한 분자 화물(나노크기 분자로부터 작은 화학적 분자 및 DNA의 큰 단편)의 세포 흡수를 돕는 짧은 펩티드이다. 본원에서 사용되는 용어 "화물"은 치료제, 진단 프로브, 펩티드, 핵산, 안티센스 올리고뉴클레오티드, 플라스미드, 단백질, 나노입자, 리포좀, 발색단, 소분자 및 방사성물질로 구성된 군을 포함하나, 이에 제한되지는 않는다. 본 발명의 양태에서, 화물은 또한 CRISPR Cas 시스템의 임의의 구성요소 또는 전체 기능적 CRISPR Cas 시스템을 포함할 수 있다. 본 발명의 양태는 (a) 본 발명의 세포 투과 펩티드 및 요망되는 화물을 포함하는 복합체를 제조하는 단계, 및 (b) 상기 복합체를 대상체에 경구, 관절내, 복막내, 수막강내, 동맥내, 비내, 실질내, 피하, 근내, 정맥내, 피부, 직장내, 또는 국소 투여하는 단계를 포함하는, 대상체로 요망되는 화물을 전달하기 위한 방법을 추가로 제공한다. 화물은 공유 결합을 통한 화학 결합 또는 비-공유 상호작용을 통해 펩티드와 회합된다.

[0531] CPP의 기능은 세포로 화물을 전달하는 것이며, 이는 세포내 이입을 통해 일반적으로 발생하는 과정으로, 화물은 살아 있는 포유류 세포의 엔도솜으로 전달된다. 세포-투과 펩티드는 다양한 크기, 아미노산 서열, 및 전하의 펩티드이나, 모든 CPP는 형질막에서 전위하고, 세포질 또는 소기관으로 다양한 분자 화물의 전달을 돕는 능력인 하나의 명백한 특징을 갖는다. CPP 전위는 막에서의 직접 투과, 세포내이입-매개 진입, 및 일시적 구조의 형성을 통한 전위의 3개의 주요 진입 메커니즘으로 분류될 수 있다. CPP는 암을 포함하는 다양한 질병의 치료에서의 약물 전달제 및 바이러스 억제제로서 의약뿐만 아니라 세포 표지를 위한 조영제에서 다수의 응용가능성이 발견되었다. 조영제의 예는 GFP에 대한 담체, MRI 조영제, 또는 양자점으로 작용하는 것을 포함한다. CPP는 연구 및 의약에서 사용하기 위한 시험관내 및 생체내 전달 벡터로서 큰 잠재성을 지닌다. CPP는 높은 상대 존재비의 양성으로 하전된 아미노산, 예를 들어, 리신 또는 아르기닌을 함유하거나, 극성/하전된 아미노산 및 비-극성, 소수성 아미노산의 교대 패턴을 함유하는 서열을 갖는 아미노산 조성을 갖는다. 이들 2개 유형의 구조는 각각 다가양이온 또는 양친매성으로 언급된다. 세 번째 부류의 CPP는 낮은 순 전하와 함께 무극성 잔기만을 함유하는 소수성 펩티드이거나, 세포 흡수에 중요한 소수성 아미노산 기를 갖는다. 발견된 최초 CPP 중 하나는 배양 중인 다수의 세포 유형에 의해 주위 배지로부터 효과적으로 흡수되는 것으로 밝혀진 인간 면역결핍 바이러스 1(HIV-1)로부터의 트랜스활성화 전사 활성인자(Tat)였다. 그 이후로, 공지된 CPP의 수가 매우 확장되었고, 더욱 효과적인 단백질 형질도입 특성을 갖는 소분자 합성 유사체가 생성되었다. CPP는 페네트라틴(Penetratin), Tat (48-60), 트랜스포르탄(Transportan), 및 (R-AhX-R4) (Ahx=아미노헥사노일)을 포함하나, 이에 제한되지는 않는다.

[0532] 미국 특허 제8,372,951호는 높은 세포-투과 효율 및 낮은 독성을 나타내는 호산구 양이온성 단백질(ECP)로부터 유래된 CPP를 제공한다. 척추동물 대상체로 화물과 함께 CPP를 전달하는 양태가 또한 제공된다. CPP 및 이의 전달의 추가 양태는 미국 특허 제8,575,305호; 제8,614,194호 및 제8,044,019호에 존재할 수 있다.

[0533] CRISPR-Cas 시스템을 전달하기 위해 이용될 수 있는 CPP가 또한 전체내용이 참조로서 포함되는 문헌[Suresh Ramakrishna, Abu-Bonsrah Kwaku Dad, Jagadish Beloor, et al. Genome Res. 2014 Apr 2. [Epub ahead of print]]의 매뉴스크립트 "Cas9 단백질 및 가이드 RNA의 세포-투과 펩티드-매개 전달에 의한 유전자 붕괴(Gene disruption by cell-penetrating peptide-mediated delivery of Cas9 protein and guide RNA)"에 또한 제공되며, 상기 문헌에서 CPP-컨쥬게이션된 재조합 Cas9 단백질 및 CPP-복합체화된 가이드 RNA를 이용한 처리가 인간 세포주에서 내인성 유전자 붕괴를 초래한 것이 입증되어 있다. 상기 논문에서, Cas9 단백질은 티오에테르 결합을 통해 CPP에 컨쥬게이션된 반면, 가이드 RNA는 CPP와 복합체화되어, 축합된 양성으로 하전된 나노입자를 형성한다. 배아 줄기 세포, 피부 섬유아세포, HEK293T 세포, HeLa 세포, 및 배아 암종 세포를 포함하는 인간 세포와 변형된 Cas9 및 가이드 RNA의 동시 및 순차적 처리가 플라스미드 트랜스펙션에 비해 감소된 표적외 돌연변이를 갖는 효과적인 유전자 붕괴를 초래한 것으로 밝혀졌다.

[0534] **이식가능한 장치**

[0535] 다른 구현예에서, CRISPR Cas 시스템의 전달을 위해 이식가능한 장치가 또한 고려된다. 예를 들어, 미국 특허 공개 제20110195123호는 국소적으로 그리고 연장된 기간에 약물을 용리하는 이식가능한 의료 장치를 개시하며, 이러한 장치의 몇 가지 유형, 실행의 처리 방식 및 이식 방법을 포함하여 제공된다. 장치는 중합체성 기질, 예를 들어 장치 몸체로서 사용되는 매트릭스, 및 약물, 및 일부 경우에 추가적인 스캐폴딩 재료, 예컨대 금속 또는 추가적인 중합체, 및 가시성 및 영상화를 향상시키는 재료를 포함한다. 약물의 선택은 국소적으로 그리고 연

장된 기간에 약물을 방출하는 것에 유리하게 기반하며, 여기서 약물은 종양, 염증, 변성과 같은 병에 걸린 영역의 세포외 기질(ECM)에 직접적으로, 또는 주사된 평활 세포에, 또는 예방을 위해 방출된다. 한 종류의 약물은 si RNA, sh RNA, 또는 안티센스 RNA/DNA, 리보자임 및 뉴클레오사이드 유사체를 포함하지만, 이들로 제한되지 않는 RNA 간섭(RNAi)에 기반한 유전자 침묵 약물이다. 따라서, 이 시스템은 본 발명의 CRISPR Cas 시스템에 사용될 수 있고/있거나 적합할 수 있다. 일부 구현예에서 이식 방식은 근접 요법 및 바늘 생검을 포함하는 다른 치료를 위해 오늘날 개발되고 사용되는 현재 사용되는 이식 절차이다. 이러한 경우에, 본 발명에서 기재된 새로운 이식물의 치수는 본래의 이식물과 유사하다. 통상적으로 소수의 장치는 동일한 치료 절차 동안 이식된다.

[0536] 미국 특허 공개 제20110195123호에 기재되어 있는 바와 같이, 복강과 같은 강에 적용가능한 시스템을 포함하는 약물 전달 이식형 또는 삽입형 시스템 및/또는 임의의 다른 유형의 투여가 제공되며, 이때 예를 들어 선택적으로 매트릭스일 수 있는 생체안정성 및/또는 분해가능한 및/또는 생체흡수성 중합체성 기질을 포함하는 약물 전달 시스템은 고정 또는 부착되지 않는다. 용어 "삽입"은 또한 이식을 포함하는 것으로 언급되어야 한다. 약물 전달 시스템은 바람직하게는 미국 특허 공개 제20110195123호에 기재되어 있는 바와 같은 "로더(Loder)"로서 실행된다.

[0537] 중합체 또는 복수의 중합체는 생체에 적합하며, 제어된 속도로 작용제를 방출시킬 수 있는 작용제 및/또는 복수의 작용제를 혼입하되, 중합체성 기질, 예컨대 매트릭스의 총 용적은, 일부 구현예에서 선택적으로 및 바람직하게는 작용제의 치료수준이 도달되게 하는 최대 용적 이하이다. 비제한적 예로서, 이러한 용적은 작용제 장입을 위한 용적으로 필요한 바와 같은 바람직하게는 0.1 m^3 내지 1000 mm^3 범위 내에 있다. 로더는, 예를 들어 크기가 기능성에 의해 결정되는 장치, 예를 들어 제한 없이, 무릎 관절, 자궁내 또는 자궁경부 링과 함께 혼입될 때, 선택적으로 더 클 수 있다.

[0538] (조성물을 전달하기 위한) 약물 전달 시스템은 바람직하게는 일부 구현예에서 분해성 중합체를 사용하도록 설계되거나(주요 방출 메커니즘은 벌크 침식(bulk erosion)임); 또는 일부 구현예에서, 분해되지 않는 또는 서서히 분해되는 중합체가 사용되며(주요 방출 메커니즘은 벌크 침식보다는 확산됨), 따라서 외부는 막으로서 기능하고, 내부는 약물 저장소로서 기능하는데, 이는 장기간(예를 들어 약 1 주 내지 약 몇 개월) 동안 주변에 의해 실행가능하게 영향받지 않는다. 상이한 방출 메커니즘을 지니는 상이한 중합체의 조합이 또한 선택적으로 사용될 수 있다. 표면에서 농도 구배는 바람직하게는 총 약물 방출 기간의 상당한 기간 동안 효과적으로 일정하게 유지되고, 따라서 확산 속도는 효과적으로 일정하다("제로 방식" 확신으로 지칭됨). 용어 "일정한"은 더 낮은 치료적 유효성의 역치 초과로 바람직하게 유지되지만, 여전히 선택적으로 초기 버스트를 특징으로 하고/하거나 변동하는, 예를 들어 특정 정도로 증가 및 감소할 수 있는 확산 속도를 의미한다. 확산 속도는 바람직하게는 연장된 기간 동안 이렇게 유지되며, 치료적으로 효과적인 기간, 예를 들어 효과적인 침묵 기간을 최적화하는 특정 수준으로 일정하게 고려될 수 있다.

[0539] 약물 전달 시스템은 선택적으로 및 바람직하게는 화학물질이 실재하든 아니든 또는 효소로부터의 공격 및 대상체의 신체내 다른 인자에 기인하여 뉴클레오티드 기반 치료제가 저하되는 것으로부터 보호하도록 설계된다.

[0540] 미국 특허 공개 제20110195123호에 기재된 바와 같은 약물 전달 시스템은 장치의 이식 시 및/또는 이식 후에, 활성화 및/또는 가속화/감속화의 비침습적 및/또는 최소 침습적 방법에 의해 작동되는 감지 및/또는 활성화 기기와 선택적으로 연관되며, 예를 들어 선택적으로 이하로 제한되는 것은 아니지만, 열적인 가열 및 냉각, 레이저 빔 및 집속 초음파를 포함하는 초음파 및/또는 RF(무선주파수) 방법 또는 장치를 포함한다.

[0541] 미국 특허 공개 제20110195123호의 일부 구현예에 따르면, 국소 전달을 위한 부위는 선택적으로 종양을 포함하는 세포의 고도 비정상 증식, 및 억제된 세포자멸사, 자가면역질환 상태를 포함하는 활성 및 또는 만성 염증 및 감염, 근육 및 신경 조직을 포함하는 퇴행성 조직, 만성 통증, 퇴행성 부위, 및 뼈 골절의 위치 및 조직 재생의 향상을 위한 다른 상처 위치, 및 손상된 심근, 평활근 및 횡문근을 특징으로 하는 표적 부위를 포함할 수 있다. 국소 전달을 위한 부위는 또한 선택적으로 임신, 감염의 예방 및 노화를 포함하는 예방적 활성을 수행할 수 있는 부위를 포함할 수 있다.

[0542] 조성물의 이식을 위한 부위, 또는 표적 부위는 바람직하게는 표적화된 국소 전달에 충분히 작은 반경, 면적 및/또는 용적을 특징으로 한다. 예를 들어, 표적 부위는 선택적으로 약 0.1 mm 내지 약 5 cm 범위의 직경을 가진다.

[0543] 표적 부위의 위치는 바람직하게는 최대 치료 효능에 대해 선택된다. 예를 들어, 약물 전달 시스템의 조성물(선택적으로 상기 기재한 바와 같은 이식을 위한 장치와 함께)은 선택적으로 및 바람직하게는 종양 환경 또는 이들

의 관련된 혈액 공급 내에서 또는 이에 근접하여 이식된다.

[0544] 예를 들어 조성물은 (선택적으로 장치와 함께) 선택적으로 혈관계 등 내에서 니플을 통해 췌장, 전립선, 유방, 간 내에서 또는 이에 근접하여 이식된다.

[0545] 표적 위치는 (선택적으로 신체 내 임의의 부위는 로더를 이식하는데 적합할 수 있기 때문에 단지 비제한적 예로서): 1. 기저핵, 백질 및 회백질에서 파킨슨병 또는 알츠하이머병에서와 같은 퇴행성 부위에서의 뇌; 2. 근위축성 측삭경화증(ALS)의 경우에서와 같은 척추; 3. HPV 감염을 예방하기 위한 자궁경; 4. 활성 및 만성 염증 관절; 5. 건선의 경우에서와 같은 진피; 6. 진통 효과를 위한 교감 및 감각 신경 부위; 7. 골내이식; 8. 급성 및 만성 감염 부위; 9. 질내; 10. 귀내--청각기관, 내이의 미로, 전정계; 11. 기관내; 12. 심장내; 관상동맥, 심외막; 13. 방광; 14. 담도계; 15. 신장, 간, 비장을 포함하지만, 이들로 제한되지 않는 실질세포; 16. 림프절; 17. 침샘; 18. 치과용 수지; 19. 관절내(관절 내로); 20. 안내; 21. 뇌조직; 22. 뇌실; 23. 복강을 포함하는 강(예를 들어, 제한되는 것은 아니지만, 난소암); 24. 내부식도 및 25. 직장내로 이루어진 군으로부터 선택적으로 선택된다.

[0546] 선택적으로 시스템의 삽입(예를 들어, 조성물을 함유하는 장치)는 표적 부위의 및 이러한 부위 부근의, 국소 pH 및/또는 온도 및/또는 ECM 내 약물의 확산 및/또는 약물 역학에 영향을 미치는 다른 생물학적 인자에 영향을 미치도록 표적 부위 및 해당 부위 부근에서 ECM에 대한 물질의 주사와 연관된다.

[0547] 선택적으로, 일부 구현예에 따르면, 상기 작용제의 방출은 삽입 전 및/또는 삽입 시 및/또는 삽입 후, 레이저빔, 조사, 열적 가열 및 냉각, 집속 초음파를 포함하는 초음파 및/또는 RF(무선주파수) 방법 또는 장치를 포함하는 활성화 및/또는 가속화/감속화의 비침습적 및/또는 최소 침습적 및/또는 그 밖의 방법에 의해 작동되는 감지 및/또는 활성화 기기와 연관될 수 있었다.

[0548] 미국 특허 공개 제20110195123호의 다른 구현예에 따르면, 약물은 바람직하게는, 예를 들어 이하에 기재되는 바와 같이 유방, 췌장, 뇌, 신장, 방광, 폐 및 전립선에서 국소화된 암 경우에 대해 유전자 침묵 생물학적 RNAi 약물을 포함한다. 게다가, 이러한 약물이 로더 기질, 예컨대 매트릭스와 함께 캡슐화될 수 있다면, siRNA 이외의 다수 약물은 로더에서 캡슐화되어 적용가능하며, 본 발명과 함께 사용될 수 있다. 이러한 약물은 진균 감염에 대해 암포테리신 B; 골수염에서와 같은 항생제; 진통제, 예컨대 마취제; 퇴행방지제, 예컨대 알츠하이머 또는 파킨슨병에서 등 통증의 경우에 척추 부근에 이식된 로더에서를 포함하는, 오늘날 본 발명 이외의 방법에 의해 전달되는 승인된 약물을 포함한다. 이러한 시스템은 본 발명의 CRISPR Cas 시스템을 전달하기 위해 사용되고/되거나 적합하게 될 수 있다.

[0549] 예를 들어, 평활근 세포(스텐팅 절차 동안 손상된 그리고 그 결과 증식하는 경향이 있는) 평활근 세포의 성장 또는 재성장의 방지와 같은 구체적 적용분야에 대해, 약물은 선택적으로 H19 침묵을 포함하는 평활근 세포를 침묵하게 하는 siRNA, 또는 탁솔, 라파마이신 및 라파마이신-유사체로 이루어진 군으로부터 선택된 약물일 수 있다. 이러한 경우에, 로더는 바람직하게는 일정한 속도로 연장 방출되는 약물 용리 스텐트(DES), 또는 스텐트와 함께, 별개로 이식되는 전용 장치이다. 이것 모두는 본 발명의 CRISPR Cas 시스템에 사용될 수 있고/있거나 적합할 수 있다.

[0550] 구체적 적용의 다른 예로서, 신경 및 근육 퇴행성 질환은 비정상 유전자 발현에 기인하여 발생한다. 침묵 RNA의 국소 전달은 이러한 비정상 유전자 발현을 방해하는 치료적 특성을 가질 수 있다. 소약물 및 거대분자를 포함하는 세포자멸사방지 및 염증방지 및 퇴행방지 약물의 국소 전달은 또한 선택적으로 치료적일 수 있다. 이러한 경우에, 로더는 일정한 속도로 및/또는 별도로 이식되는 전용 장치를 통해 연장된 방출을 위해 적용된다. 이것 모두는 본 발명의 CRISPR Cas 시스템에 사용될 수 있고/있거나 적합할 수 있다.

[0551] 구체적 적용의 또 다른 예로서, 정신과적 및 인지 장애가 유전자 변형자를 이용하여 치료된다. 침묵 RNA를 지니는 유전자 녹다운은 치료적 선택사항이다. 중추 신경계 부위에 뉴클레오티드 기반 작용제를 국소적으로 전달하는 로더는 정신이상, 조울증, 신경성 장애 및 행동 병폐를 포함하지만, 이들로 제한되지 않는 정신과적 및 인지 장애에 대한 치료적 선택사항이다. 로더는 또한 구체적 뇌 부위에 이식 시 소약물 및 거대분자를 포함하는 약물을 국소적으로 전달할 수 있었다. 이것 모두는 본 발명의 CRISPR Cas 시스템에 사용될 수 있고/있거나 적합할 수 있다.

[0552] 구체적 적용의 다른 예로서, 국소 부위에서 선천성 및/또는 적응성 면역의 침묵은 기관 이식 거부반응을 예방할 수 있다. 이식된 기관 및/또는 이식된 부위에 이식된 로더를 지니는 침묵 RNA 및 면역조절 시약의 국소 전달은 이식 기관에 대해 활성화된 CD8과 같은 면역 세포를 쫓아버림으로써 국소 면역 억제를 제공한다. 이것 모두는

본 발명의 CRISPR Cas 시스템에 사용될 수 있고/있거나 적합할 수 있다.

[0553] 구체적 적용의 다른 예로서, VEGF 및 앤지오앤지오제닌 및 기타를 포함하는 혈관성장인자는 신혈관형성에 필수적이다. 인자, 펩타이드, 펩티도모방체의 국소 전달 또는 그들의 리프레서의 억제는 중요한 치료적 양상이며; 리프레서의 침묵 및 로더에 의해 혈관형성을 자극하는 인자, 펩타이드, 거대분자 및 소약물은 말초, 전신 및 심혈관질환에 대한 치료제이다.

[0554] 이식과 같은 삽입 방법은 선택적으로 다른 유형의 조직 이식을 위해 및/또는 삽입을 위해 및/또는 조직 샘플링을 위해 선택적으로 변형 없이 또는 대안적으로 선택적으로 이러한 방법에서 주요하지 않은 변형만에 의해 이미 사용될 수 있다. 이러한 방법은 선택적으로 초음파와 함께 및/또는 초음파 없이 근접치료 방법, 생검, 내시경, 예컨대 ERCP, 뇌 조직 내로의 정위적 방법, 관절, 복부 기관, 방광벽 및 체강 내로 복강경의 이식을 포함하는 복강경검사를 포함하지만, 이들로 제한되지 않는다.

[0555] **CRISPR 효소 mRNA 및 가이드 RNA**

[0556] CRISPR 효소 mRNA 및 가이드 RNA는 또한 별개로 전달될 수 있다. CRISPR 효소 mRNA는 가이드 RNA 전에 전달되어 CRISPR 효소가 발현될 시간을 제공할 수 있다. CRISPR 효소 mRNA는 가이드 RNA의 투여 전 1 내지 12 시간(바람직하게는 약 2 내지 6 시간)에 투여될 수 있다.

[0557] 대안적으로, CRISPR 효소 mRNA 및 가이드 RNA는 함께 투여될 수 있다. 유리하게는, 가이드 RNA의 제2 부스터 용량은 CRISPR 효소 mRNA + 가이드 RNA의 초기 투여 후 1 내지 12 시간(바람직하게는 약 2 내지 6 시간)에 투여될 수 있다.

[0558] CRISPR 효소 mRNA 및/또는 가이드 RNA의 추가 투여는 가장 효율적인 게놈 변형 수준을 달성하는데 유용할 수 있다.

[0559] 독성 및 표적외 효과의 최소화를 위해, 전달되는 CRISPR 효소 mRNA 및 가이드 RNA의 농도를 제어하는 것은 중요할 것이다. 최적의 농도의 CRISPR 효소 mRNA 및 가이드 RNA는 세포 또는 동물 모델에서 상이한 농도를 시험함으로써, 그리고 잠재적인 표적외 게놈 유전자좌에서의 변형 정도를 분석하는 심층 시퀀싱(deep sequencing)을 사용함으로써 결정될 수 있다. 예를 들어, 인간 게놈의 EMX1 유전자에서 5'-GAGTCCGAGCAGAAGAAGAA-3'을 표적화하는 가이드 서열에 대해, 심층 시퀀싱은 다음의 2 개의 표적외 유전자좌, 즉 1: 5'-GAGTCCTAGCAGGAGAAGAA-3' 및 2: 5'-GAGTCTAAGCAGAAGAAGAA-3'에서 변형 수준을 평가하기 위해 사용될 수 있다. 표적 상의 변형의 가장 높은 수준을 제공하는 한편 표적외 변형의 수준을 최소화하는 농도가 생체내 전달을 위해 선택되어야 한다.

[0560] 대안적으로, 독성 수준 및 표적외 효과를 최소화하기 위해, CRISPR 효소 닉카아제 mRNA(예를 들어 D10A 돌연변이를 지니는 스트렙토코커스 피오게네스 Cas9)는 관심 대상의 부위를 표적화하는 가이드 RNA의 쌍에 의해 전달될 수 있다. 2 개의 가이드 RNA는 다음과 같이 이격될 필요가 있다. 각각 적색(단일 밑줄) 및 청색(이중 밑줄)의 가이드 서열(이들 예는 스트렙토코커스 피오게네스 Cas9에 대한 PAM 필요조건에 기반함).

[0561]

돌출부 길이 (bp)	가이드 RNA 설계 (가이드 서열 및 PAM 색표시한 암호)
14	5'- NNNNNNNNNNNNNNNNNNCCNNNNNNNNNNNNNNNNNNNGGNNNNNNNNNNNNN-3' 3'- NNNNNNNNNNNNNNNNNNGGNNNNNNNNNNNNNNNNNNNCCNNNNNNNNNNNNN-5'
13	
12	5'- NNNNNNNNNNNNNNNNNNCCNNNNNNNNNNNNNNNNNNGGNNNNNNNNNNNNN-3' 3'- NNNNNNNNNNNNNNNNNNGGNNNNNNNNNNNNNNNNNNCCNNNNNNNNNNNNN-5'
11	
10	5'- NNNNNNNNNNNNNNNNNNCCNNNNNNNNNNNNNNNNNNGGNNNNNNNNNNNNN-3' 3'-
9	

[0562]

8	NNNNNNNNNNNNNNNNNNGGNNNNNNNNNNNNNNNNNNNCCNNNNNNNNNNNNNNN-5’
7	5’-NNNNNNNNNNNNNNNNNNNCCNNNNNNNNNNNNNNNNNNNGGNNNNNNNNNNNNNNNN-3’
6	3’-NNNNNNNNNNNNNNNNNNNGGNNNNNNNNNNNNNNNNNNNCCNNNNNNNNNNNNNNNN-5’
5	5’-NNNNNNNNNNNNNNNNNNNCCNNNNNNNNNNNNNNNNNNGGNNNNNNNNNNNNNNNN-3’
4	3’-NNNNNNNNNNNNNNNNNNNGGNNNNNNNNNNNNNNNNNNCCNNNNNNNNNNNNNNNN-5’
3	5’-NNNNNNNNNNNNNNNNNNNCCNNNNNNNNNNNNNNNNNGGNNNNNNNNNNNNNNNNN-3’
2	3’-NNNNNNNNNNNNNNNNNNNGGNNNNNNNNNNNNNNNNNCCNNNNNNNNNNNNNNNNN-5’
1	5’-NNNNNNNNNNNNNNNNNNNCCNNNNNNNNNNNNNNNNNGGNNNNNNNNNNNNNNNNN-3’
블런트	3’-NNNNNNNNNNNNNNNNNNNGGNNNNNNNNNNNNNNNNNCCNNNNNNNNNNNNNNNNN-5’

[0563]

2	5'- NNNNNNNNNNNNNNNNNNNNNCCNNNNNNNNNNNNNNGGNN NNNNNNNNNNNNNNNNNNN-3'
3	3'- NNNNNNNNNNNNNNNNNNNNGGNNNNNNNNNNNNNNNCCNN NNNNNNNNNNNNNNNNNNN-5'
4	5'- NNNNNNNNNNNNNNNNNNNNCCNNNNNNNNNNNNNNGGNNN
5	NNNNNNNNNNNNNNNNNNN-3'
	3'- NNNNNNNNNNNNNNNNNNNNGGNNNNNNNNNNNNNNCCNNN
6	NNNNNNNNNNNNNNNNNNN-5'
	5'-
7	NNNNNNNNNNNNNNNNNNNCCNNNNNNNNNNNNNNGGNNNN NNNNNNNNNNNNNNNNNNN-3'
	3'-
8	NNNNNNNNNNNNNNNNNNNNGGNNNNNNNNNNNNNCCNNNN NNNNNNNNNNNNNNNNNNN-5'
12	5'- NNNNNNNNNNNNNNNNNNNCCNNNNNNNNNNNNNGGNNNNN NNNNNNNNNNNNNNNNNNN-3'
13	3'- NNNNNNNNNNNNNNNNNNNNGGNNNNNNNNNNNNCCNNNNN NNNNNNNNNNNNNNNNNNN-5'
14	5'-

[0564]

15	NNNNNNNNNNNNNNNNNNNCCNNNNNNNNNNGGNNNNNN NNNNNNNNNNNNNNNNNNN-3' 3'- NNNNNNNNNNNNNNNNNNNGGNNNNNNNNNNCCNNNNNN
16	NNNNNNNNNNNNNNNNNNN-5'
17	5'- NNNNNNNNNNNNNNNNNNNCCNNNNNNNNNNGGNNNNNNN NNNNNNNNNNNNNNNNNNN-3' 3'- NNNNNNNNNNNNNNNNNNNGGNNNNNNNNNNCCNNNNNNN NNNNNNNNNNNNNNNNNNN-5'
	5'- NNNNNNNNNNNNNNNNNNNCCNNNNNNNNNGGNNNNNNNN NNNNNNNNNNNNNNNNNNN-3' 3'- NNNNNNNNNNNNNNNNNNNGGNNNNNNNNNCCNNNNNNNN NNNNNNNNNNNNNNNNNNN-5'
	5'- NNNNNNNNNNNNNNNNNNNCCNNNNNNNNNGGNNNNNNNN NNNNNNNNNNNNNNNNNNN-3' 3'- NNNNNNNNNNNNNNNNNNNGGNNNNNNNNNCCNNNNNNNN NNNNNNNNNNNNNNNNNNN-5'
	5'- NNNNNNNNNNNNNNNNNNNCCNNNNNNNNNGGNNNNNNNN NNNNNNNNNNNNNNNNNNN-3'

[0565]

	3'- NNNNNNNNNNNNNNNNNNGGNNNNNNCCNNNNNNNNN NNNNNNNNNNNNNNNNNN-5' 5'- NNNNNNNNNNNNNNNNNNCCNNNNNGGNNNNNNNNNNN NNNNNNNNNNNNNNNNNN-3' 3'- NNNNNNNNNNNNNNNNNNGGNNNNNNCCNNNNNNNNNNN NNNNNNNNNNNNNNNNNN-5' 5'- NNNNNNNNNNNNNNNNNNCCNNNNNGGNNNNNNNNNNN NNNNNNNNNNNNNNNNNN-3' 3'- NNNNNNNNNNNNNNNNNNGGNNNNNCCNNNNNNNNNNN NNNNNNNNNNNNNNNNNN-5' 5'- NNNNNNNNNNNNNNNNNNCCNNNNGGNNNNNNNNNNNN NNNNNNNNNNNNNNNNNN-3' 3'- NNNNNNNNNNNNNNNNNNGGNNNNCCNNNNNNNNNNNN NNNNNNNNNNNNNNNNNN-5' 5'- NNNNNNNNNNNNNNNNNNCCNNNNGGNNNNNNNNNNNN NNNNNNNNNNNNNNNNNN-3' 3'- NNNNNNNNNNNNNNNNNNGGNNNCCNNNNNNNNNNNN

	NNNNNNNNNNNNNNNNNN-5' 5'- NNNNNNNNNNNNNNNNNNNCCNNGGNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNN-3' 3'- NNNNNNNNNNNNNNNNNNNGGNNCCNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNN-5' 5'- NNNNNNNNNNNNNNNNNNNCCNGGNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNN-3' 3'- NNNNNNNNNNNNNNNNNNNGGNCCNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNN-5' 5'- NNNNNNNNNNNNNNNNNNNNCCGGNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNN-3' 3'- NNNNNNNNNNNNNNNNNNNNGGCCNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNN-5' 5'- NNNNNNNNNNNNNNNNNNNNGGNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNN-3' 3'- NNNNNNNNNNNNNNNNNNNNCCGGNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNN-5'

[0566]

	5'- NNNNNNNNNNNNNNNNNNNNCGGNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNN-3' 3'- NNNNNNNNNNNNNNNNNNNNNNCCNGGNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNN-5' 5'- NNNNNNNNNNNNNNNNNNNNNCGGNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNN-3' 3'- NNNNNNNNNNNNNNNNNNNNNNCCNNGGNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNN-5' 5'- NNNNNNNNNNNNNNNNNNNNCGGNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNN-3' 3'- NNNNNNNNNNNNNNNNNNNNNNCCNNNGGNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNN-5' 5'- NNNNNNNNNNNNNNNNNNNNCGGNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNN-3' 3'- NNNNNNNNNNNNNNNNNNNNNNCCNNNNGGNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNN-5' 5'- NNNNNNNNNNNNNNNNNNNNCGGNNNNNNNNNNNNNNN

[0567]

	NNNNNNNNNNNNNNNNNNNN-3' 3'- NNNNNNNNNNNNNNNNNNNNNNCCNNNNNGGNNNNNNNNN NNNNNNNNNNNNNNNNNNNN-5'

[0568]

[0569] 시스템의 추가적 의문은 5' 돌출부의 출원인 증거로 주어졌다(예를 들어, 문헌[Ran et al., Cell. 2013 Sep 12;154(6):1380-9] 및 2013년 8월 28일자로 출원된 미국 가출원 특허 일련번호 제61/871,301호). 출원인은 2 개의 가이드 RNA와 조합될 때 Cas9 닉카아제 돌연변이체에 의한 효율적인 절단에 관한 파라미터를 추가로 동정하였고, 이들 파라미터는 5' 돌출부의 길이를 포함하지만, 이것으로 제한되지 않는다. 본 발명의 구현예에서, 5' 돌출부는 200 개 이하의 염기쌍, 바람직하게는 100 개 이하의 염기쌍, 또는 더 바람직하게는 50 개 이하의 염기쌍이다. 본 발명의 구현예에서, 5' 돌출부는 적어도 26 개의 염기쌍, 바람직하게는 적어도 30 개의 염기쌍 또는 더 바람직하게는 34 내지 50 개의 염기쌍 또는 1 내지 34 개의 염기쌍이다. 본 발명의 다른 바람직한 방법에서, 제1 표적 서열 근처의 DNA 듀플렉스의 한 개 가닥의 절단을 지시하는 제1 가이드 서열 및 제2 표적 서열 근처의 다른 가닥의 절단을 지시하는 제2 가이드 서열은 평활 절단 또는 3' 돌출부를 초래한다. 본 발명의 구현

예에서, 3' 돌출부는 150, 100 또는 25 개 이하의 염기쌍 또는 적어도 15, 10 또는 1 개의 염기쌍이다. 바람직한 구현예에서, 3' 돌출부는 1 내지 100 개의 염기쌍이다.

[0570] 본 발명의 양태는 감소될 유전자 산물의 발현 또는 유전자 산물을 암호화하는 DNA 분자 내로 추가로 도입될 주형 폴리뉴클레오티드 또는 2개의 5' 돌출부가 재아닐링 및 결찰되게 함으로써 정확하게 절단될 개재 서열 또는 변용될 유전자 산물의 활성 또는 작용 또는 증가될 유전자 산물의 발현에 관한 것이다. 본 발명의 구현예에서, 유전자 산물은 단백질이다.

[0571] 가이드 서열 사이에 8 bp 미만의 중복(-8 bp 초과의 상쇄)을 지니는 5' 돌출부를 생성하는 sgRNA 쌍만이 검출가능한 삽입결실 형성을 매개할 수 있었다. 중요하게는, 이들 분석에서 사용되는 각각의 가이드는 야생형 Cas9와 짝지어질 때 삽입결실을 효율적으로 유발할 수 있는데, 이는 가이드 쌍의 상대적 위치가 이중 닉카아제 활성을 예측함에 있어 가장 중요한 파라미터가 된다는 것을 나타낸다.

[0572] Cas9n 및 Cas9H840A는 DNA의 마주하는 가닥의 틈을 만들기 때문에, 주어진 sgRNA를 지니는 Cas9H840A로의 Cas9n의 치환은 돌출부 유형의 역위를 초래해야만 한다. 예를 들어, Cas9n을 지니는 5' 돌출부를 생성할 sgRNA의 쌍은 대신에 원칙적으로 대응하는 3' 돌출부를 생성하여야 한다. 따라서, Cas9n를 지니는 3' 돌출부의 생성을 야기하는 sgRNA 쌍은 Cas9H840A와 함께 사용되어 5' 돌출부를 생성할 수 있었다. 예상치 못하게, 본 발명자들은 5'와 3' 돌출부를 둘 다 생성하도록 설계된 sgRNA 쌍의 세트를 지니는(-278 내지 +58 bp 범위에 있는 상쇄) Cas9H840A를 시험하였지만, 삽입결실 형성을 관찰할 수 없었다. sgRNA 쌍이 Cas9H840A에 의한 이중 틈내기를 허용하는데 필요한 설계 규칙을 동정하기 위한 추가적인 작업이 필요할 수 있다.

[0573] **간, 프로단백질 전환효소 서브틸리신 켁신 9(PCSK9)**

[0574] 데이터는 표현형 전환을 나타낸다.

[0575] 프로단백질 전환효소 서브틸리신 켁신 9(PCSK9)는 서브틸리신 세린 프로테아제 패밀리의 구성원이다. PCSK9는 주로 간에 의해 발현되고, 간세포 LDL 수용체 발현의 하향 조절에 중요하다. 혈장 내 LDL-C 수준은 PCSK9에서 기능 돌연변이의 획득에 의해 인간에서 고도로 상승되는데, 그들을 중증의 고콜레스테롤혈증을 갖는 것으로 분류한다. 따라서, PCSK9는 CRISPR에 대해 매력적인 표적이다. PCS9K-표적화된 CRISPR은 지질 입자에서 제형화될 수 있으며, 예를 들어 정맥내로 약 15, 45, 90, 150, 250 및 400 ㎍/kg에서 투여된다(예를 들어, http://www.alnylam.com/capella/wp-content/uploads/2013/08/ALN-PCS02-001-Protocol-Lancet.pdf 참조).

[0576] Bailey 등은(J Mol Med (Berl). 1999 Jan;77(1):244-9) 생체외 체세포 유전자 요법에 의한 인슐린 전달이 당뇨병 환자로부터 비-B-세포 체세포(예를 들어, 섬유아세포)의 제거를 수반하고, 시험관내에서 그들을 유전적으로 변용시켜 인슐린을 생성하고 분비한다는 것을 개시한다. 세포는 배양물 중에서 성장될 수 있고, 인슐린 대체의 공급원으로서 공여체에 돌려보내진다. 이 방법으로 변형된 세포는 이식 전에 평가될 수 있고, 보존 저장액을 저온 보존될 수 있었다. 환자 자신의 세포를 사용함으로써, 절차는 면역억제에 대한 필요를 제거하여야 하며, 조직 공급의 문제점을 극복하는 한편, 세포 파괴의 재발을 회피한다. 생체외 체세포 유전자 요법은 다회 트렌스펙션으로 처리될 수 있고 증식이 제어되는 접근 가능하고 강한 세포 유형을 필요로 한다. 비-B-세포 체세포의 사용과 연관된 특별한 문제는 프로인슐린의 인슐린으로의 처리, 및 글루코스 자극 프로인슐린 생합성에 대한 민감도의 수여 및 조절된 인슐린 방출을 포함한다. 섬유아세포, 뇌하수체 세포, 신장(COS) 세포 및 난소(CHO) 세포를 사용하는 예비 연구는 이들 도전이 충족될 수 있고, 생체외 체세포 유전자 요법이 인슐린 대체 요법에 대한 실현가능한 접근을 제공한다는 것을 시사한다. Bailey 등의 시스템은 간에 전달에 대한 본 발명의 CRISPR Cas 시스템에 대해 사용되고/사용되거나 적합할 수 있다.

[0577] Sato 등의 방법(Nature Biotechnology Volume 26 Number 4 April 2008, 431 내지 442 페이지)은 간에 전달을 위해 본 발명의 CRISPR Cas 시스템에 적용될 수 있다. Sato 등은 siRNA-함유 비타민 A-결합 리포좀에 의한 치료가 간 섬유증을 거의 완전히 해결하였고, 용량 의존적 및 지속기간 의존적 방식으로 다른 치명적 디메틸니트로스아민 유발 간경변증을 지니는 래트에서 생존기간을 연장시킨 것을 발견하였다. 양이온성 지질로서 0,0'-디테트라데카노일-N-(a-트리메틸암모니오아세틸) 디에탄올아민 클로라이드(DC-6-14), 콜레스테롤 및 디올레오일포스파티딜 에탄올아민을 4:3:3의 몰비로 함유하는 양이온성 리포좀(리포트러스트(Lipotrust))(시험관내 및 생체내 유전자 전달을 위한 혈청 함유 조건 하에서 고 트렌스펙션 효율을 나타냄)을 Hokkaido System Science로부터 구입하였다. 냉동 건조 빈 리포좀 방법을 사용하여 리포좀을 제조하였고, 사용 전 교반 하에 동결건조 지질 혼합물에 대한 재증류수(DDW)의 첨가에 의해 1 mM(DC-16-4)의 농도에서 제조되었다. VA-결합 리포좀을 제조하기 위해, DMSO 중에 용해시킨 200 nmol의 비타민 A(레티놀, Sigma)를 1.5 ㎖ 관 내 25 1C에서 교반시킴으로써 리포좀

현탁액(DC-16-4로서 100 nmol)과 혼합하였다. VA-결합 리포좀 운반 siRNAgp46(VA-lip-siRNAgp46)을 제조하기 위해, siRNAgp46의 용액(DDW 중의 580 pmol/㎖)을 25C에서 교반시키면서 레티놀 결합 리포좀 용액에 첨가하였다. siRNA 대 DC-16-4 비는 1:11.5(mol/mol)이고, siRNA 대 리포좀 비(wt/wt)는 1:1이었다. 리포좀에 의해 취해지지 않은 임의의 유리 비타민 A 또는 siRNA는 마이크로파티션 시스템(VIVASPIN 2 농축기 30,000 MWCO PES, VIVASCIENCE)를 사용하여 리포좀 제제로부터 분리되었다. 리포좀 현탁액을 필터에 첨가하고 나서 25 1C에서 3 회 5 분 동안 1,500 g에서 원심분리시켰다. 분획을 수집하고 나서, 필터에 걸린 물질을 PBS로 재구성하여 시험관내 또는 생체내 용도를 위한 목적으로 하는 용량을 달성하였다. 0.75 mg/kg siRNA의 3 회 주사로 래트에 격일로 제공하였다. Sato 등의 시스템은 Sato 등에 의해 기재되는 리포좀에서 약 0.5 내지 1 mg/kg의 CRISPR Cas RNA를 인간에게 전달함으로써 간에 전달을 위한 본 발명의 CRISPR Cas 시스템에 사용되고/사용되거나 적합하게 될 수 있다.

[0578] Rozema 등이 siRNA Dynamic PolyConjugates로 명명한 시험관내와 생체내 둘 다에서 간세포에 siRNA의 전달을 위한 비히클에 대한 Rozema 등의 방법(PNAS, August 7, 2007, vol. 104, no. 32)이 또한 본 발명에 적용될 수 있다. 역동적 폴리-컨주게이트 기술의 중요한 특징은 막-활성 중합체, 엔도솜의 산성 환경에 도달될 때까지 이 중합체의 활성을 역으로 가리는 능력, 및 이 변형된 중합체 및 그의 siRNA 화물을 간단한 저압 i.v. 주사 후에 생체내 간세포에 대해 특이적으로 표적화하는 능력을 포함한다. SATA-변형 siRNA는 4℃에서 16 시간 동안 수 중에서 5' 아민 변형 siRNA와 1 중량 당량(wt eq)의 N숙신이미딜-S-아세틸티오아세테이트(SATA) 시약(Pierce) 및 0.36 중량 당량의 $NaHCO_3$의 반응에 의해 합성된다. 이어서, 변형된 siRNA는 9 vol의 에탄올의 첨가에 의해 침전되고, -80℃에서 2 시간 동안 인큐베이션시킨다. 침전물은 1X siRNA 완충제(Dharmacon) 중에서 재현탁되고, 260-nm 파장에서 흡광도를 측정함으로써 정량화한다. PBAVE(5mMTAPS 중의 30 mg/㎖, pH 9)는 1.5 중량% SMPT(Pierce)의 첨가에 의해 변형된다. 1-h 인큐베이션 후에, 0.8 mg의 SMPT-PBAVE를 5 mM TAPS 를 함유하는 등장성 글루코스 용액 400 ㎕에 첨가한다(pH 9). 이 용액에 50 ㎍의 SATA-변형 siRNA를 첨가하였다. 용량-반응 실험에 대해, [PBAVE]가 일정한 경우, 상이한 양의 siRNA를 첨가한다. 이어서, 혼합물을 16 시간 동안 인큐베이션시킨다. 이어서 용액에 5.6 mg의 Hepes 유리 염기를 첨가한 후에 3.7 mg의 CDM-NAG와 1.9 mg의 CDM-PEG의 혼합물을 첨가한다. 이어서, 주사 전에 용액을 적어도 1 시간 동안 실온에서 인큐베이션시킨다. CDM-PEG 및 CDM-NAG를 염화옥살릴을 사용하여 생성한 산염화물로부터 합성한다. 산 염화물에 1.1 몰 당량 폴리에틸렌 글리콜 모노메틸 에터(분자량 평균 450)를 첨가하여 CDM-PEG 또는 (아미노에톡시)에톡시-2-(아세틸아미노)-2-데옥시-β-D-글루코피라노시드를 생성하여 CDM-NAG를 생성한다. 0.1% TFA 물/아세토니트릴 구배를 이용하는 역상 HPLC에 의해 최종 생성물을 정제한다. 약 25 내지 50 ㎍의 siRNA를 마우스에 전달하였다. Rozema 등의 시스템은, 예를 들어 인간에 대한 전달을 위해 약 50 내지 약 200 mg의 투약량을 생각함으로써 간에 대한 전달을 위한 본 발명의 CRISPR Cas 시스템에 적용될 수 있다.

[0579] **뼈**

[0580] Oakes 및 Lieberman(Clin Orthop Relat Res. 2000 Oct;(379 Suppl):S101-12)은 뼈에 대한 유전자의 전달을 논의한다. 특이적 해부학적 부위에서 세포 내로 유전자를 전달함으로써, 성장 인자의 골유도 특성은 지속 기간 동안 생리학적 용량으로 사용되어 더 유의한 치유 반응을 가능하게 할 수 있다. 구체적 해부학적 부위, 뼈의 질 및 연조직 외피는 국부 유전자 요법을 위한 표적 세포의 선택에 영향을 미친다. 골전도 담체에서 치료 부위에 전달되는 유전자 요법 벡터는 유망한 결과를 수득하였다. 몇몇 조사자들은 동물 모델에서 생체외 및 생체내 국부 유전자 요법을 사용하여 흥미로운 결과를 나타내었다. 이러한 시스템은 뼈에 대한 절단을 위한 CRISPR Cas 시스템에 사용되고/되거나 적합할 수 있다.

[0581] **뇌**

[0582] 뇌에 대한 전달 선택사항은 리포좀 내로 DNA 또는 RNA 중 하나 형태로 CRISPR 효소 및 가이드 RNA의 캡슐화 및 트랜스-혈액뇌장벽(BBB) 전달을 위한 분자 트로이목마에 대한 컨쥬게이팅을 포함한다. 분자 트로이목마는 비인간 영장류의 뇌 내로 B-gal 발현 벡터의 전달에 효과적인 것으로 나타났다. 동일한 접근이 CRISPR 효소 및 가이드 RNA를 함유하는 벡터를 전달하는데 사용될 수 있다. 예를 들어, Xia CF and Boado RJ, Pardridge WM ("Antibody-mediated targeting of siRNA via the human insulin receptor using avidin-biotin technology." Mol Pharm. 2009 May-Jun;6(3):747-51. doi: 10.1021/mp800194)은 배양물에서 및 생체내에서 세포에 대한 짧은 간섭 RNA(siRNA)의 전달 방법이 수용체-특이적 단클론성 항체(mAb) 및 아비딘-바이오틴 기술의 조합된 사용에 의해 가능하다는 것을 기재한다. 저자는 또한 표적화 mAb와 siRNA 간의 결합이 아비딘-바이오틴 기술에 의해 안정하기 때문에, 뇌와 같은 먼 부위에서 RNAi 효과가 표적화된 siRNA의 정맥내 투여 후 생체내에서 관찰되었다는

것을 보고한다.

[0583] Zhang 등(Mol Ther. 2003 Jan;7(1):11-8.))은 루시퍼라제와 같은 리포터를 암호화하는 발현 플라스미드가 인간 인슐린 수용체(HIR)에 대해 단클론성 항체(MAb)를 지니는 생체내 레서스 원숭이 뇌에 대해 표적화된, 85 nm 페길화 면역리포좀을 포함하는 "인공 바이러스"의 내부에서 캡슐화되는 방법을 기재한다. HIRMAb는 리포좀이 혈액-뇌 장벽을 가로지르는 통과세포배출 및 정맥내 주사 후 신경 혈장막을 가로지르는 내포작용을 겪는 외인성 유전자를 운반하는 것을 가능하게 한다. 뇌에서 루시퍼라제 유전자 발현의 수준은 래트와 비교하여 레서스 원숭이에서 50 배 더 높았다. 영장류 뇌에서 베타-갈락토시다제 유전자의 널리 퍼진 신경 발현은 조직화학과 공초점 현미경 둘 다에 의해 입증되었다. 저자는 이 접근이 24 시간에 가역적 성체 유전자이식을 실현가능하게 만든다는 것을 나타낸다. 따라서, 면역리포좀의 사용이 바람직하다. 이들은 특이적 표적 또는 세포 표면 단백질을 표적화하는 항체와 함께 사용될 수 있다.

[0584] 나노입자(Cho, S., Goldberg, M., Son, S., Xu, Q., Yang, F., Mei, Y., Bogatyrev, S., Langer, R. and Anderson, D., Lipid-like nanoparticles for small interfering RNA delivery to endothelial cells, Advanced Functional Materials, 19: 3112-3118, 2010) 또는 엑소좀(Schroeder, A., Levins, C., Cortez, C., Langer, R., and Anderson, D., Lipid-based nanotherapeutics for siRNA delivery, Journal of Internal Medicine, 267: 9-21, 2010, PMID: 20059641)을 통하는 것과 같은 다른 전달 수단 또는 RNA가 또한 바람직하다.

[0585] 사실, 엑소좀은 CRISPR 시스템과 일부 병용되는 시스템인 전달 siRNA에서 특히 유용한 것으로 나타났다. 예를 들어, El-Andaloussi S, 등("exosomes-mediated delivery of siRNA in vitro and in vivo." Nat Protoc. 2012 Dec;7(12):2112-26. doi: 10.1038/nprot.2012.131. Epub 2012 Nov 15.)은 엑소좀이 상이한 생물학적 장벽을 가로질러 약물을 전달하기 위한 촉망받는 도구가 되고 시험관내 및 생체내 siRNA의 전달을 위해 이용될 수 있는 방법을 기재한다. 그들의 접근은 펩티드 리간드와 융합된 엑소좀 단백질을 포함하는 발현 벡터의 트렌스펙션을 통해 표적화된 엑소좀을 생성하는 것이다. 이어서, 엑소좀은 트렌스펙션된 세포 상청액으로부터 정제 및 특성규명된 다음, siRNA는 엑소좀에 장입된다. 본 발명에 따른 전달 또는 투여는 엑소좀에 의해 수행될 수 있으며, 특히 뇌로 제한되지 않는다.

[0586] 비타민 E(α-토코페롤)은 CRISPR Cas와 컨쥬게이션될 수 있고, 예를 들어 뇌에 짧은 간섭 RNA(siRNA)를 전달하기 위해 Uno 등(HUMAN GENE THERAPY 22:711-719 (June 2011))에 의해 행해지는 것과 유사한 방식으로 고밀도 지단백질(HDL)과 함께 뇌에 전달될 수 있다. 마우스에 인산염 완충 식염수(PBS) 또는 유리 TocsiBACE 또는 Toc-siBACE/HDL로 채우고 뇌 주입 키트 3(Brain Infusion Kit 3)(Alzet)과 연결한 삼투 미니펌프(Osmotic minipumps)(모델 1007D; Alzet, 캘리포니아주 쿠퍼티노)를 통해 주입하였다. 뇌 주입 캐뉼라를 등의 제3뇌실 내로 주입을 위해 중앙에서 브레그마(bregma)에 대해 약 0.5 mm 뒤에 위치시켰다. Uno 등은 HDL과 함께 3 nmol만큼 적은 Toc-siRNA가 동일한 ICV 주입 방법에 의해 유사한 정도로 표적 감소를 유발할 수 있었다는 것을 발견하였다. α-토코페롤에 컨주게이트되고 뇌에 표적화된 HDL과 공동투여된 유사한 투약량의 CRISPR Cas는 본 발명에서 인간에 대해 고려될 수 있으며, 예를 들어, 뇌에 표적화된 약 3 nmol 내지 약 3 μmol의 CRISPR Cas가 고려될 수 있다.

[0587] Zou 등((HUMAN GENE THERAPY 22:465-475 (April 2011))은 래트의 척수에서 생체내 유전자 침묵에 대한 짧은 헤어핀 RNA 표적화 PKCɣ의 렌티 바이러스-매개 전달 방법을 기재한다. Zou 등은 척추강내 카테터에 의해 1 x 10^9 형질도입 단위 (TU)/㎖의 역가를 갖는 재조합 렌티바이러스 약 10 ㎕를 투여하였다. 뇌에 표적화된 렌티바이러스 벡터에서 발현된 유사한 투약량의 CRISPR Cas는 본 발명에서 인간에 대해 고려될 수 있으며, 예를 들어, 1 x 10^9 형질도입 단위(TU)/㎖의 역가를 갖는 렌티바이러스에서 뇌에 표적화된 약 10 내지 50 ㎖의 CRISPR Cas가 고려될 수 있다.

[0588] **<u>표적화된 결실, 치료적 적용</u>**

[0589] 유전자의 표적화된 결실이 바람직하다. 예는 실시예 18에서 예시된다. 따라서, 콜레스테롤 생합성, 지방산 생합성 및 다른 대사 장애에 수반된 유전자, 아밀로이드 및 다른 질환에 수반된 미스-폴딩 단백질을 암호화하는 유전자, 세포 형질전환을 야기하는 발암유전자, 잠재 바이러스 유전자, 및 특히 우성-음성 장애를 야기하는 유전자가 바람직하다. 본 명세서에 예시하는 바와 같이, 본 발명자들은 바이러스 또는 나노입자 전달 시스템 중 하나를 사용하는, 대사 장애, 아밀로이드증 및 단백질 응집 관련 질환, 유전자 돌연변이 및 전좌로부터 생기는 세포 형질전환, 유전자 돌연변이의 우성 음성 효과, 잠재 바이러스 감염 및 다른 관련된 증상으로 고통받는 치료

가 필요한 대상체 또는 환자의 간, 뇌, 눈, 상피, 조혈, 또는 다른 조직에 CRISPR-Cas 시스템에 대한 유전자 전달을 선호한다.

[0590] CRISPR-Cas 시스템의 치료적 적용은 녹내장, 아밀로이드증, 및 헌팅턴병을 포함한다. 이들은 실시예 20에서 예시되며, 그에 기재된 특징은 단독으로 또는 조합으로 선호된다.

[0591] 본 발명에 의해 치료될 수 있는 폴리글루타민 확장 질환의 다른 예는 척수소뇌성 운동실조증 1형(spinocerebellar ataxia type 1: SCA1)을 포함한다. 소뇌내 주사시, 짧은 헤어핀 RNA를 발현시키는 재조합 아데노연관 바이러스(AAV) 벡터는 운동 협응을 완전히 개선시키고, 소뇌 형태를 회복시키며, SCA1 마우스의 푸르키네(Purkinje) 세포 내 특징적 아탁신(아탁신)-1 봉입을 분해하였다(예를 들어, 문헌[Xia et al., Nature Medicine, Vol. 10, No. 8, Aug. 2004] 참조). 특히, AAV1 및 AAV5 벡터가 바람직하며, 약 1 x 10^{12} 벡터게놈/㎖의 AAV 역가가 바람직하다.

[0592] 예로서, HIV-1에 의한 만성 감염은 치료 또는 예방될 수 있다. 이것을 달성하기 위해, 대다수의 HIV-1 게놈을 표적화하는 한편 최대 범위 및 유효성에 대해 HIV-1 균주 변이체를 고려하는 CRISPR-Cas 가이드 RNA를 생성할 수 있다. 숙주 면역 시스템의 전통적인 아데노 바이러스 또는 렌티 바이러스-매개 감염에 의한 CRISPR-Cas 시스템의 전달을 달성할 수 있다. 접근에 따라서, 숙주 면역 세포는 a) 단리, CRISPR-Cas로 형질도입, 선택 및 숙주에 재도입될 수 있거나, 또는 b) CRISPR-Cas 시스템의 전신 전달에 의해 생체내로 형질도입된다. 제1 접근은 내성 면역 집단을 생성시키는 반면, 제2 접근은 숙주 내에서 잠재 바이러스 저장소를 표적화할 가능성이 있다. 이는 실시예 부문에서 더 상세하게 논의된다.

[0593] 다른 예에서, Sangamo BioSciences, Inc.에 의해 양도 된 미국 특허 공개 제20130171732호는 게놈 내로 항-HIV 이식유전자의 삽입, 본 발명의 CRISPR Cas 시스템에 적용될 수 있는 방법에 관한 것이다. 다른 구현예에서, CXCR4 유전자는 표적화될 수 있고, Sangamo BioSciences, Inc.에 의해 양도 된 미국 특허 공개 제20100291048호의 TALE 시스템은 본 발명의 CRISPR Cas 시스템에 대해 변형될 수 있다. Sangamo BioSciences, Inc. 에 의해 양도 된 미국 특허 공개 제20130137104호 및 제20130122591호 및 Cellectis에 의해 양도 된 미국 특허 공개 제20100146651호의 방법은 그것이 유전자 변형의 빈도를 증가시키기 위해 하이포크산틴-구아닌 포스포리보실트랜스퍼라제(HPRT) 유전자좌의 변형을 수반하기 때문에 이식유전자 발현에 대해 더 일반적으로 적용될 수 있다.

[0594] 또한 본 발명은 유전자 녹아웃 세포 라이브러리를 생성하는 것으로 생각된다. 각각의 세포는 단일 유전자 녹아웃을 가질 수 있다. 이는 실시예 23에서 예시된다.

[0595] 각각의 세포가 단일 유전자 녹아웃을 갖는 ES 세포의 라이브러리를 만들 수 있으며, ES 세포의 전체 라이브러리는 모든 단일 유전자 녹아웃을 가질 것이다. 이 라이브러리는 세포 처리뿐만 아니라 질환에서 유전자 기능을 선별에 유용하다. 이 세포 라이브러리를 제조하기 위해, ES 세포 내로 유도성 프로모터(예를 들어, 독시사이클린 유도성 프로모터)에 의해 구동되는 Cas9를 통합할 수 있다. 추가로, 특이적 유전자를 표적화하는 단일 가이드 RNA를 ES 세포 내로 통합할 수 있다. ES 세포 라이브러리를 제조하기 위해, 인간 게놈 내 각각의 유전자를 표적화하는 가이드 RNA를 암호화하는 유전자의 라이브러리와 ES 세포를 단순히 혼합할 수 있다. 인간 ES 세포의 AAVS1 유전자좌 내로 단일 BxB1 attB 부위를 처음 도입할 수 있다. 이어서, AAVS1 유전자좌 내 BxB1 attB 부위 내로 개개 가이드 RNA 유전자의 통합을 용이하게 하기 위해 BxB1 인터그라제를 사용할 수 있다. 통합을 용이하게 하기 위해, 각각의 가이드 RNA 유전자는 단일 attP 부위를 운반하는 플라스미드 상에 함유될 수 있다. 이 방법의 BxB1은 게놈 내 attB 부위를 가이드 RNA 함유 플라스미드 상의 attP 부위와 재조합할 수 있다. 세포 라이브러리를 생성하기 위해, 통합된 단일 가이드 RNA를 가지고 Cas9 발현을 유도하는 세포의 라이브러리를 취할 수 있다. 유도 후, Cas9는 가이드 RNA에 의해 구체화된 부위에서 이중 가닥 파손을 매개한다.

[0596] 단백질 치료제의 만성 투여는 특이적 단백질에 대한 허용되지 않는 면역 반응을 유발할 수 있다. 단백질 약물의 면역원성은 소수의 면역우성 헬퍼 T 림프구(HTL) 에피토프 탓일 수 있다. 이들 단백질 내에 함유된 이들 HTL 에피토프의 MHC 결합 친화도를 감소시키는 것은 더 낮은 면역원성을 지니는 약물을 생성할 수 있다(Tangri S, 등의 문헌["Rationally engineered therapeutic proteins with reduced immunogenicity" J Immunol. 2005 Mar 15;174(6):3187-96.]). 본 발명에서, 특히 CRISPR 효소의 면역원성은 에리트로포이에틴에 관해 문헌[Tangri et al]에서 처음 제시된 접근에 따라 감소될 수 있고, 후속적으로 발행될 수 있다. 따라서, 방향 진화(directed evolution) 또는 합리적 설계는 숙주 종(인간 또는 다른 종)에서 CRISPR 효소(예를 들어, Cas9)의 면역원성을 감소시키기 위해 사용될 수 있다.

[0597] 실시예 28에서, 본 발명자들은 관심대상의 3개의 가이드 RNA를 사용하였고, 세포의 작은 서브세트에서만 생기는

생체내 효율적인 DNA 절단을 시각화할 수 있다. 본질적으로, 출원인은 본 명세서에서 나타낸 것은 표적화된 생체내 절단이다. 특히, 이는 포유류와 같은 더 고등의 유기체에서 특이적 표적화가 또한 달성될 수 있다는 개념의 증거를 제공한다. 또한 다중 가이드 서열(즉, 별개의 표적)이 동시에(공동 전달의 의미에서) 사용될 수 있다는 점에서 다중복합 양태를 강조한다. 다시 말해서, 본 발명자들은 동시에 그러나 독립적으로 표적화된 몇몇 상이한 서열에 의한 다중 접근을 사용하였다.

[0598] AAV에 대한 프로토콜의 적합한 예로서, 본 발명의 바람직한 벡터가 실시예 34에 제공된다.

[0599] 트리뉴클레오티드 반복 장애는 치료될 바람직한 질환이다. 이들은 또한 본 명세서에서 예시된다.

[0600] 예를 들어, 미국 특허 공개 제20110016540호는 트리뉴클레오티드 반복 확장 장애와 연관된 세포, 동물 및 단백질을 유전적으로 변형하기 위한 아연 핑거 뉴클레아제의 사용을 기재한다. 트리뉴클레오티드 반복 확장 장애는 발생 신경생물학과 관련되고 종종 인지뿐만 아니라 감각 운동 기능에 영향을 미치는 복합, 진행성 장애이다.

[0601] 트리뉴클레오티드 반복 확장 단백질은 트리뉴클레오티드 반복 확장 장애가 발생할 감수성, 트리뉴클레오티드 반복 확장 장애의 존재, 트리뉴클레오티드 반복 확장 장애의 중증도 또는 이들의 임의의 조합과 연관된 다양한 세트의 단백질이다. 트리뉴클레오티드 반복 확장 장애는 반복 유형에 의해 결정되는 2 가지 범주로 나누어진다. 가장 통상적인 반복은, 유전자의 암호 영역에서 존재할 때 아미노산 글루타민(Q)에 대해 암호화나는 트리플렛 CAG이다. 따라서, 이들 장애는 폴리글루타민(폴리Q) 장애로서 지칭되며, 다음의 질환을 포함한다: 헌팅턴병(HD); 척수수근 근위축증(SBMA); 척수소뇌실조증(SCA 유형 1, 2, 3, 6, 7 및 17); 및 치아적핵담창구루이체 위축증(DRPLA). 남아있는 트리뉴클레오티드 반복 확장 장애는 CAG 트리플렛을 수반하지 않거나 또는 CAG 트리플렛은 유전자의 암호 영역에 있지 않고, 따라서 비 폴리글루타민 장애로서 지칭된다. 비-폴리글루타민 장애는 취약 X 증후군(FRAXA); 취약 XE 정신지체(FRAXE); 프리드라이히 운동실조(FRDA); 근긴장성이영양증(DM); 및 척수소뇌실조증(SCA 8형 및 12형)을 포함한다.

[0602] 트리뉴클레오티드 반복 확장 장애와 관련된 단백질은 트리뉴클레오티드 반복 확장 장애에 대한 트리뉴클레오티드 반복 확장 장애와 연관된 단백질의 실험적 연관에 기반하여 통상적으로 선택된다. 예를 들어, 트리뉴클레오티드 반복 확장 장애와 연관된 단백질의 생성 속도 또는 순환 농도는 트리뉴클레오티드 반복 확장 장애가 없는 집단에 비해 트리뉴클레오티드 반복 확장 장애를 갖는 집단에서 상승 또는 억제될 수 있다. 단백질 수준에서의 차이는 웨스턴 블롯, 면역조직화학염색, 효소결합 면역흡착분석(ELISA) 및 질량분석법을 포함하지만, 이들로 제한되지 않는 프로테오믹(proteomic) 기법을 사용하여 평가될 수 있다. 대안적으로, 트리뉴클레오티드 반복 확장 장애와 연관된 단백질은 DNA 마이크로어레이 분석, 일련 분석 유전자 발현(SAGE), 및 정량적 실시간 중합효소 연쇄반응(Q-PCR)을 포함하지만, 이들로 제한되지 않는 게놈 기법을 사용하여 단백질을 암호화하는 유전자의 유전자 발현 프로파일을 얻음으로써 동정될 수 있다.

[0603] 트리뉴클레오티드 반복 확장 장애와 연관된 단백질의 비제한적 예는 AR(안드로겐 수용체), FMR1(취약 X 정신지체 1), HTT(헌팅틴), DMPK (근긴장성 이영양증-단백질 키나제), FXN(프라탁신), ATXN2(아탁신 2), ATN1(아트로핀 1), FEN1(플랩 구조-특이적 엔도뉴클레아제 1), TNRC6A (트리뉴클레오티드 반복부 함유 6A), PABPN1(폴리(A) 결합 단백질, 핵 1), JPH3(정크토필린 3), MED15(매개 복합체 서브유닛 15), ATXN1(아탁신 1), ATXN3(아탁신 3), TBP(TATA 박스 결합 단백질), CACNA1A(칼슘 채널, 전압-의존적, P/Q 유형, 알파 1A 서브유닛), ATXN80S (ATXN8 마주하는 가닥(비단백질 암호화)), PPP2R2B(단백질 포스파타제 2, 조절 서브유닛 B, 베타), ATXN7(아탁신 7), TNRC6B(트리뉴클레오티드 반복부 함유 6B), TNRC6C (트리뉴클레오티드 반복부 함유 6C), CELF3 (CUGBP, Elav-유사 패밀리 구성원 3), MAB21L1(mab-21-유사 1(예쁜꼬마선충(C. elegans))), MSH2(mutS 상동체 2, 결장암, 비용종증 1형(이콜라이(E. coli))), TMEM185A(막관통 단백질 185A), SIX5(SIX 호메오박스 5), CNPY3(캐노피 3 상동체(제브라피쉬)), FRAXE(취약 부위, 폴산 유형, 희쉬, fra(X)(q28) E), GNB2(구아닌 뉴클레오티드 결합 단백질(G 단백질), 베타 폴리펩티드 2), RPL14(리보솜 단백질 L14), ATXN8 (아탁신 8), INSR(인슐린 수용체), TTR(트랜스티레틴), EP400(E1A 결합 단백질 p400), GIGYF2(GRB10 상호작용 GYF 단백질 2), OGG1(8-옥소 구아닌 DNA 글리코실라제), STC1(스탄니오칼신 1), CNDP1(카르노신 디펩티다제 1(메탈로펩티다제 M20 패밀리)), C10orf2(염색체 10 오픈 리딩 프레임 2), MAML3 마스터마인드-유사 3(초파리), DKC1(선천성 이상각화증 1, 디스케린), PAXIP1(PAX 상호작용(전산-활성화 도메인과의) 단백질 1), CASK(칼슘/카모듈린-의존적 세린 단백질 키나제(MAGUK 패밀리)), MAPT(미세소관-결합 단백질 tau), SP1(Sp1 전사 인자), POLG(중합효소(DNA 관련), 감마), AFF2(AF4/FMR2 패밀리, 구성원 2), THBS1(트롬보스폰딘 1), TP53(종양 단백질 p53), ESR1(에스트로겐 수용체 1), CGGBP1(CGG 트리플렛 반복부 결합 단백질 1), ABT1(기본 전사 1의 활성인자), KLK3(칼리크레인-관련 펩티다제 3), PRNP(프리온 단백질), JUN(jun 발암유전자), KCNN3(칼륨 중간체/소 전도도 칼슘-활성화 채

널, 서브패밀리 N, 구성원 3), BAX(BCL2-연관 X 단백질), FRAXA(취약 부위, 폴산 유형, 희귀, fra(X)(q27.3) A(거대고환증, 정신지체)), KBTBD10(kelch 반복부 및 BTB(POZ) 도메인 함유 10), MBNL1(머슬블라인드-유사(초파리)), RAD51(RAD51 상동체(RecA 상동체, 이콜라이)(사카로마이세스 세레비시애(S. cerevisiae))), NCOA3(핵 수용체 공동활성인자 3), ERDA1(확장 반복부 도메인, CAG/CTG 1), TSC1(결절성 경화증 1), COMP(연골 올리고머 매트릭서ㅡ 단백질), GCLC(글루타메이트-시스테인 리가제, 촉매적 서브유닛), RRAD(당뇨병과 연관된 Ras-관련), MSH3(mutS 상동체 3(이콜라이)), DRD2(도파민 수용체 D2), CD44(CD44 분자(인디안 혈액 그룹)), CTCF(CCCTC-결합 인자(아연 핑거 단백질)), CCND1(사이클린 D1), CLSPN(클라스핀 상동체(남아프리카 발톱개구리)), MEF2A(근세포 인핸서 인자 2A), PTPRU(단백질 티로신 포스파타제, 수용체 유형, U), GAPDH(글리세르알데히드-3-포스페이트 탈수소효소), TRIM22(3부 모티프-함유 22), WT1(윌름 종양 1), AHR(아릴 탄화수소 수용체), GPX1(글루타티온 페록시다제 1), TPMT(티오퓨린 S-메틸트랜스퍼라제), NDP(노리병(가신경교종)), ARX(아리스탈레스 관련 호메오박스), MUS81(MUS81 엔도뉴클레아제 상동체(사카로마이세스 세레비시애)), TYR(티로시나제(안피부백피증IA)), EGR1(조기 성장 반응 1), UNG(유라실-DNA 글리코실라제), NUMBL(numb 상동체 (초파리)-유사), FABP2(지방산 결합 단백질 2, 장), EN2(인그레일드(engrailed) 호메오박스 2), CRYGC(크리스탈린, 감마 C), SRP14(신호 인식 입자 14 kDa(상동성 Alu RNA 결합 단백질)), CRYGB(크리스탈린, 감마 B), PDCD1(세포예정사 1), HOXA1(호메오박스 A1), ATXN2L(아탁신 2-유사), PMS2(PMS2 감수분열 후 분리 증가된 2(사카로마이세스 세레비시애)), GLA(갈락토시다제, 알파), CBL(Cas-Br-M(뮤린) 동종지향성 레트로바이러스 형질전환 서열), FTH1(페리틴, 중 폴리펩티드 1), IL12RB2(인터류킨 12 수용체, 베타 2), OTX2(교정돌기 호메오박스 2), HOXA5(호메오박스 A5), POLG2(중합효소(DNA 관련), 감마 2, 부속 서브유닛), DLX2(디스탈-리스(distal-less) 호메오박스 2), SIRPA(신호-조절 단백질 알파), OTX1(교정돌기 호메오박스 1), AHRR(아릴-탄화수소 수용체 리프레서), MANF(중뇌 성상교세포-유래 신경영양 인자), TMEM158(막관통 단백질 158(유전자/위유전자)), 및 ENSG00000078687을 포함한다.

[0604] 트리뉴클레오티드 반복 확장 장애와 연관된 바람직한 단백질은 HTT(헌팅틴), AR(안드로겐 수용체), FXN(프라탁신), Atxn3(아탁신), Atxn1(아탁신), Atxn2(아탁신), Atxn7(아탁신), Atxn10(아탁신), DMPK(근긴장성 이영양증-단백질 키나제), Atn1(아트로핀 1), CBP(creb 결합 단백질), VLDLR(매우 저밀도 지단백질 수용체), 및 이들의 임의의 조합을 포함한다.

[0605] 다른 양태에 따르면, CFTR 유전자에서 돌연변이를 갖는 대상체의 치료를 위한 유전자 요법의 방법이 제공되며, 선택적으로 생체에 적합한 약제학적 담체를 통해 대상체의 세포에 치료적 유효량의 CRISPR-Cas 유전자 요법 입자를 투여하는 단계를 포함한다. 바람직하게는, 표적 DNA는 돌연변이 델타F508를 포함한다. 일반적으로, 돌연변이는 야생형으로 수선되는 것이 바람직하다. 이 경우에, 돌연변이는 위치 508에서 페닐알라닌(F)의 코돈을 포함하는 3 개 뉴클레오티드의 결실이다. 따라서, 이 예에서 수선은 돌연변이체 내로 상실 코돈의 재도입을 필요로 한다.

[0606] 이 유전자 수선 전략을 실행하기 위해, 아데노바이러스/AAV 벡터 시스템이 숙주 세포, 세포들 또는 환자에게 도입되는 것이 바람직하다. 바람직하게는, 시스템은 F508 잔기를 함유하는 상동성 수선 주형을 포함하는 아데노바이러스/AAV 벡터 시스템과 함께 Cas9(또는 Cas9 닉카아제) 및 가이드 RNA를 포함한다. 이는 앞서 논의한 전달 방법 중 하나를 통해 대상체에게 도입될 수 있다. CRISPR-Cas 시스템은 CFTR델타 508 키메라 가이드 RNA에 의해 가이드될 수 있다. 이는 틈내기 또는 절단될 CFTR 게놈 유전자좌의 특이적 부위를 표적화한다. 절단 후에, 수선 주형은 낭포성 섬유증을 초래하거나 또는 낭포성 섬유증 관련 증상을 야기하는 결실을 보정하는 상동성 재조합을 통해 절단 부위에 삽입된다. 적절한 가이드 RNA를 지니는 CRISPR 시스템의 전신 도입을 제공하고 전달을 지시하기 위한 이 전략은 표 B에서와 같은 대사적, 간, 신장 및 단백질 질환 및 장애를 야기하는 유전자를 편집 또는 달리 조작하기 위한 유전자 돌연변이를 표적화하는데 사용될 수 있다.

[0607] **<u>게놈 편집</u>**

[0608] 본 발명의 CRISPR/Cas9 시스템은 TALEN 및 ZFN을 사용하여 제한된 성공을 지니는 이전에 시도한 유전자 돌연변이를 수정하기 위해 사용될 수 있다. 예를 들어, 듀크 유니버시티의 공개 출원 WO2013163628 A2(Genetic Correction of Mutated Genes)는, 예를 들어 디스트로핀 유전자 내 돌연변이에 기인하는 근육 퇴화를 초래하는 열성의, 치명적, X-연관 장애인 듀켄씨근이영양증(Duchenne Muscular Dystrophy, "DMD")을 초래하는 것과 같은 뉴클레아제 매개 비 상동성 말단 결합을 통해 수정될 수 있는 절단된 유전자 산물 및 조기 정지 코돈을 야기하는 틀 이동 돌연변이를 수정하기 위한 노력을 기재한다. DMD를 야기하는 대다수의 디스트로핀 돌연변이는 리딩 프레임을 파괴하고 디스트로핀 유전자에서의 조기 번역 종결을 야기하는 엑손의 결실이다. 디스트로핀은 근육 세포 완전성 및 기능의 조절을 초래하는 세포막의 디스트로글라이칸 복합체에 대해 구조적 안정성을 제공하는 세포질 단백질이다. 본 명세서에서 상호호환적으로 사용되는 바와 같은 디스트로핀 유전자 또는 "DMD 유전자"는

유전자좌 Xp21에서 2.2 메가베이스이다. 1차 전사는 약 2,400 kb로 측정되며 성숙 mRNA는 약 14 kb이다. 79 엑손은 3500 개 이상의 아미노산에 걸쳐 단백질을 암호화한다. 엑손 51은 DMD 환자에서 프레임-파괴 결실에 빈번하게 인접하며, 올리고뉴클레오티드 기반 엑손 스키핑(skipping)에 대한 임상 시험에서 표적화되었다. 엑손 51 스키핑 화합물 에테플러센(eteplirsen)에 대한 임상 시험은 최근에 48주에 걸쳐 상당한 기능적 이점을 보고하였고, 기준에 비해 평균 47% 디스트로핀 양성 섬유를 지녔다. 엑손 51에서의 돌연변이는 NHEJ-기반 게놈 편집에 의한 영구적 보정에 이상적으로 적합하다.

[0609] 인간 디스트로핀 유전자(DMD)로부터 표적 서열을 절단하기 위한 메가뉴클레아제 변이체에 관한, Cellectis에 의해 양도 된 미국 특허 공개 제20130145487호의 방법은 또한 본 발명의 CRISPR Cas 시스템에 대해 변형될 수 있다.

[0610] **혈액**

[0611] 본 발명은 또한 혈액에 CRISPR-Cas 시스템을 전달하는 것을 고려한다.

[0612] Wahlgren 등의 혈장 엑소좀(Nucleic Acids Research, 2012, Vol. 40, No. 17 e130)은 이전에 기재되었고, 혈액에 CRISPR Cas 시스템을 전달하기 위해 이용될 수 있다.

[0613] 본 발명의 CRISPR Cas 시스템은 또한 지중해 빈열 및 겸상적혈구증과 같은 혈색소병(hemoglobinopathies)을 치료하기 위해 생각된다. 예를 들어, 본 발명의 CRISPR Cas 시스템에 의해 표적화될 수 있는 가능한 표적에 대해 국제 특허 공개 제WO 2013/126794호 참조.

[0614] Cellectis에 의해 양도 된 미국 특허 공개 제20110225664호, 20110091441호, 제20100229252호, 제20090271881호 및 제20090222937호는 CREI 변이체에 관한 것이되, 2개의 I-CreI 단량체 중 적어도 하나는 적어도 2 개의 치환을 가지고 LAGLIDADG 코어 도메인의 각각 2 개의 기능성 서브도메인 중 하나는 각각 I-CreI의 위치 26 내지 40 및 44 내지 77로부터 위치되며, 상기 변이체는 통상적인 사이토카인 수용체 감마쇄 유전자 또는 감마 C 유전자로도 칭해지는 인간 인터류킨-2 수용체 감마쇄(IL2RG) 유전자로부터 DNA 표적 서열을 절단할 수 있다. 미국 특허 공개 제20110225664호, 제20110091441호, 제20100229252호, 제20090271881호 및 제20090222937호에서 동정되는 표적 서열은 본 발명의 CRISPR Cas 시스템에 대해 이용될 수 있다.

[0615] 중증 복합형 면역 부전증(Severe Combined Immune Deficiency: SCID)은 림프구 B에서의 기능성 결함과 항상 연관되는 림프구 T 성숙에서의 결함으로부터 초래된다(Cavazzana-Calvo et al., Annu. Rev. Med., 2005, 56, 585-602; Fischer et al., Immunol. Rev., 2005, 203, 98-109). 전반적인 발생률은 75 000 명의 출생 당 1명으로 추정된다. 치료되지 않은 SCID를 지니는 환자는 다수의 우발적 마이크로-유기체에 감염되고, 일반적으로 1년 이상 생존하지 못한다. SCID는 가족성 공여체로부터 동종이계 조혈 줄기세포 전달에 의해 처리될 수 있다. 공여체와의 조직적합성은 크게 다를 수 있다. SCID 형태 중 하나인 아데노신 데아미나제(Deaminase: ADA) 결핍증의 경우에, 환자는 재조합 아데노신 데아미나제 효소의 주사에 의해 치료될 수 있다.

[0616] ADA 유전자는 SCID 환자에서 돌연변이되는 것으로 나타났기 때문에(Giblett et al., Lancet, 1972, 2, 1067-1069), SCID에 수반된 몇몇 다른 유전자가 동정되었다(Cavazzana-Calvo et al., Annu. Rev. Med., 2005, 56, 585-602; Fischer et al., Immunol. Rev., 2005, 203, 98-109). SCID에 대한 4 가지 주된 원인이 있다: (i) SCID의 가장 흔한 형태인 SCID-X1(X-연결 SCID 또는 X-SCID)은 성숙 T 림프구 및 NK 세포의 부재를 초래하는 IL2RG 유전자 내 돌연변이에 의해 야기된다. IL2RG는 적어도 5 개의 인터류킨 수용체 복합체의 통상적인 구성성분인 감마 C 단백질을 암호화한다(Noguchi, et al., Cell, 1993, 73, 147-157). 이들 수용체는 JAK3 키나제를 통해 몇몇 표적을 활성화하는데(Macchi et al., Nature, 1995, 377, 65-68), 불활성화는 감마 C 불활성화와 동일한 증후군을 초래하고; (ii) ADA 유전자에서의 돌연변이는 림프구 전구체에 치명적인 퓨린 대사에서의 결함을 초래하는데, 이는 결국 B, T 및 NK 세포의 준 부재(quasi absence)를 초래하며; (iii) V(D)J 재조합은 면역글로불린 및 T 림프구 수용체(TCR)의 성숙에서 필수 단계이다. 이 과정에 수반된 3 가지 유전자인 재조합 활성화 유전자 1 및 2(RAG1 및 RAG2) 및 아르테미스(Artemis)에서의 돌연변이는 성숙 T 및 B 림프구의 부재를 초래하며; (iv) T 세포 특이적 신호처리에 수반된 CD45와 같은 다른 유전자에서의 돌연변이가 또한, 그들이 소수의 경우에 나타남에도 불구하고, 보고되었다(Cavazzana-Calvo et al., Annu. Rev. Med., 2005, 56, 585-602; Fischer et al., Immunol. Rev., 2005, 203, 98-109).

[0617] 그들의 유전자 염기가 동정되었지만, 상이한 SCID 형태가 2 가지 주된 이유로 유전자 요법 접근에 대한 패러다임이 되었다(Fischer et al., Immunol. Rev., 2005, 203, 98-109). 우선, 모든 혈액 질환에서와 같이, 생체외 치료가 생각될 수 있다. 조혈모세포(Hematopoietic Stem Cell: HSC)는 골수로부터 회수되고, 소수의 세포 분화

에 대해 그들의 다능 특성을 유지한다. 따라서, 그들은 시험관내에서 처리될 수 있고, 이어서 그들이 골수에서 다시 살 수 있는 경우 환자에게 재주사될 수 있다. 둘째로, 림프구의 성숙은 SCID 환자에서 손상되기 때문에, 보정된 세포는 선택적 이점을 가진다. 따라서, 보정된 세포는 기능성 면역계를 회복할 수 있다. 이 가설은 (i) SCID 환자에서 돌연변이의 복귀와 관련된 면역 기능의 부분적 복원(Hirschhorn et al., Nat. Genet., 1996, 13, 290-295; Stephan et al., N. Engl. J. Med., 1996, 335, 1563-1567; Bousso et al., Proc. Natl., Acad. Sci. USA, 2000, 97, 274-278; Wada et al., Proc. Natl. Acad. Sci. USA, 2001, 98, 8697-8702; Nishikomori et al., Blood, 2004, 103, 4565-4572), (ii) 조혈모세포에서 시험관내 SCID-X1 결핍증의 보정(Candotti et al., Blood, 1996, 87, 3097-3102; Cavazzana-Calvo et al., Blood, 1996, Blood, 88, 3901-3909; Taylor et al., Blood, 1996, 87, 3103-3107; Hacein-Bey et al., Blood, 1998, 92, 4090-4097), (iii) 동물 모델에서 생체내 SCID-X1(Soudais et al., Blood, 2000, 95, 3071-3077; Tsai et al., Blood, 2002, 100, 72-79), JAK-3(Bunting et al., Nat. Med., 1998, 4, 58-64; Bunting et al., Hum. Gene Ther., 2000, 11, 2353-2364) 및 RAG2(Yates et al., Blood, 2002, 100, 3942-3949)의 보정 및 (iv) 유전자 요법 임상 시험의 결과(Cavazzana-Calvo et al., Science, 2000, 288, 669-672; Aiuti et al., Nat. Med., 2002; 8, 423-425; Gaspar et al., Lancet, 2004, 364, 2181-2187)에 의해 몇 번 입증되었다.

[0618] Children's Medical Center Corporation 및 President and Fellows of Harvard College에 의해 양도된 미국 특허 공개 제20110182867호는 RNAi 및 항체와 같은 BCL11A 발현 또는 활성의 저해제를 통해 조혈 전구체 세포에서 태아 헤모글로빈 발현(HbF)을 조절하는 방법 및 용도에 관한 것이다. BCL11A와 같은 미국 특허 공개 제20110182867호에 개시된 표적은 태아 헤모글로빈 발현을 조절하기 위한 본 발명의 CRISPR Cas 시스템에 의해 표적화될 수 있다. 또한 추가적인 BCL11A 표적에 대해 Bauer 등의 문헌[Science 11 October 2013: Vol. 342 no. 6155 pp. 253-257) 및 Xu 등의 문헌[Science 18 November 2011: Vol. 334 no. 6058 pp. 993-996] 참조.

[0619] 귀

[0620] 본 발명은 또한 한쪽 또는 양쪽 귀에 CRISPR-Cas 시스템을 전달하는 것을 고려한다.

[0621] 연구자들은 유전자 요법이 현재의 난청 치료, 즉, 달팽이관 이식을 보조하기 위해 사용될 수 있는지 여부를 조사한다. 난청은 종종 청각신경세포에 신호를 전달할 수 없는 모세포(hair cell)의 손실 또는 손상에 의해 야기된다. 이러한 경우에, 달팽이관 이식은 소리에 반응하고 신경 세포에 전기 신호를 전달하기 위해 사용될 수 있다. 그러나 성장 인자가 손상된 모세포에 의해 거의 방출되지 않기 때문에 이들 뉴런은 종종 퇴화하고 달팽이관으로부터 뒤로 빠진다.

[0622] 미국 특허 출원 제20120328580호는, 예를 들어 주사기, 예를 들어 단일 용량 주사기를 사용하는 달팽이관의 귀 내로 루미내(luminae)(예를 들어, 중간계, 전정계 및 고실계) 내로 약제학적 조성물의 주사(예를 들어, 귀 내로 주사)를 기재한다. 예를 들어, 본 명세서에 기재된 하나 이상의 화합물은 고실내 주사(예를 들어, 중이 내로) 및/또는 외이, 중이 및/또는 내이 내로 주사에 의해 투여될 수 있다. 이러한 방법은 당업계에서, 예를 들어 인간 귀 내로 스테로이드 및 항생제의 투여를 위해 일상적으로 사용된다. 주사는, 예를 들어 귀의 원창을 통하거나 혹은 달팽이관 캡슐을 통할 수 있다. 다른 내이 투여 방법은 당업계에 공지되어 있다(예를 들어, 문헌[Salt and Plontke, Drug Discovery Today, 10:1299-1306, 2005] 참조).

[0623] 다른 투여 방식에서, 약제학적 조성물은 카테터 또는 펌프를 통해 인시추로 투여될 수 있다. 카테터 또는 펌프는, 예를 들어 귀의 달팽이관 또는 원창 및/또는 결장의 내강 내로 약제학적 조성물을 보낼 수 있다. 귀, 예를 들어 인간 귀 내로 본 명세서에 기재된 화합물 중 하나 이상을 투여하는데 적합한 예시적 약물 전달 장치 및 방법은 McKenna 등(미국 공개 제2006/0030837호) 및 Jacobsen 등(미국 특허 공개 제7,206,639호)에 기재되어 있다. 일부 구현예에서, 카테터 또는 펌프는 수술 절차 동안 환자의 귀(예를 들어, 외이, 중이 및/또는 내이)에 위치될 수 있다. 일부 구현예에서, 카테터 또는 펌프는 수술 절차에 대한 필요 없이, 예를 들어 환자의 귀(예를 들어, 외이, 중이 및/또는 내이)에 위치될 수 있다.

[0624] 대안적으로 또는 추가적으로, 본 명세서에 기재된 하나 이상의 화합물은 기계적 장치, 예컨대 달팽이관 이식 또는 외이에 착용되는 보청기와 병용하여 투여될 수 있다. 본 발명과 함께 사용하는데 적합한 예시적인 달팽이관 이식은 Edge et al., (미국 특허 공개 제2007/0093878호)에 의해 기재된다.

[0625] 일부 구현예에서, 상기 기재된 투여 방식은 임의의 순서로 조합될 수 있고, 동시 또는 산재성일 수 있다.

[0626] 대안적으로 또는 추가적으로, 본 발명은, 예를 들어, CDER 데이터 표준 메뉴얼, 버전 번호 004(fda.give/cder/dsm/DRG/drg00301.htm에서 이용가능)에 기재된 바와 같은 임의의 식품 및 약물 투여 승인

방법에 따라 투여될 수 있다.

[0627] 일반적으로, 미국 특허 출원 제20120328580호에 기재된 세포 요법 방법은 시험관내에서 내이(예를 들어, 모세포)의 성숙 세포 유형에 대해 또는 성숙 세포 유형 쪽으로 완전한 또는 부분적 분화를 촉진하기 위해 사용될 수 있다. 이어서, 이러한 방법으로부터 생긴 세포는 이러한 치료가 필요한 환자에게 이식 또는 삽입될 수 있다. 적합한 세포 유형을 동정 및 선택하기 위한 방법, 선택된 세포의 완전 또는 부분적 분화를 촉진하기 위한 방법, 완전 또는 부분적으로 분화된 세포 유형을 동정하기 위한 방법 및 완전 또는 부분적으로 분화된 세포를 이식하기 위한 방법을 포함하는, 이들 방법을 실행하는데 필요한 세포 배양 방법이 이하에 기재된다.

[0628] 본 발명에서 사용하기에 적합한 세포는, 예를 들어 본 명세서에 기재된 화합물 중 하나 이상과 시험관내에서 접촉될 때 내이의 성숙 세포, 예를 들어 모세포(예를 들어 내부 및/또는 외부 모세포)로 완전히 또는 부분적으로 분화할 수 있는 세포를 포함하지만, 이들로 제한되지 않는다. 모세포로 분화할 수 있는 예시적 세포는 줄기 세포(예를 들어, 내이 줄기 세포, 성체 줄기 세포, 골수 유래 줄기 세포, 배아 줄기 세포, 간엽 줄기 세포, 피부 줄기 세포, iPS 세포, 및 지방 유래 줄기 세포), 전구세포(예를 들어, 내이 전구세포), 지지 세포(예를 들어, 다이테르스세포(Deiters' cell), 주상세포, 내부 지골 세포, 시개(tectal) 세포 및 헨젠 세포), 및/또는 생식 세포를 포함하지만, 이들로 제한되지 않는다. 내이 감각 세포의 대체를 위한 줄기 세포의 사용은 Li 등(미국 특허 공개 제2005/0287127호) 및 Li 등(미국 특허 출원 제11/953,797호)에 기재되어 있다. 내이 감각 세포의 대체를 위한 골수 유래 줄기 세포의 용도는 Edge 등의 PCT/US2007/084654호에 기재되어 있다. iPS 세포는, 예를 들어, 문헌[Takahashi et al., Cell, Volume 131, Issue 5, 페이지 861-872 (2007); Takahashi and Yamanaka, Cell 126, 663-76 (2006); Okita et al., Nature 448, 260-262 (2007); Yu, J. et al., Science 318(5858):1917-1920 (2007); Nakagawa et al., Nat. Biotechnol. 26:101-106 (2008); 및 Zaehres and Scholer, Cell 131(5):834-835 (2007)]에 기재되어 있다.

[0629] 이러한 적합한 세포는 하나 이상의 조직 특이적 유전자의 존재를 분석함으로써(예를 들어, 정성적으로 또는 정량적으로) 동정될 수 있다. 예를 들어, 유전자 발현은 하나 이상의 조직-특이적 유전자의 단백질 생성물을 검출함으로써 검출될 수 있다. 단백질 검출 기법은 적절한 항원에 대해 항체를 사용하는(예를 들어, 세포 추출물 또는 전체 세포를 사용하는) 염색 단백질을 수반한다. 이 경우에, 적절한 항원은 조직-특이적 유전자 발현의 단백질 생성물이다. 원칙적으로, 제1 항체(즉, 항원에 결합한 항체)가 표지될 수 있지만, 제1(예를 들어, 항-IgG)에 대해 관련된 제2 항체를 사용하는 것이 더 통상적이다(그리고 시각화를 개선시킨다). 이 제2 항체는 형광색소 또는 비색법 반응에 대한 적절한 효소, 또는 금 비드(전자 현미경에 대해), 또는 바이오틴-아비딘 시스템과 컨주게이트되고, 따라서 1차 항체의 위치 및 그에 따른 항원이 인식될 수 있다.

[0630] 본 발명의 CRISPR Cas 분자는 미국 공개 출원 제20110142917호로부터 변형된 조성물을 이용하여, 외이에 대한 약제학적 조성물의 직접 적용에 의해 귀에 전달될 수 있다. 일부 구현예에서, 약제학적 조성물은 외이도에 적용된다. 귀에 대한 전달은 또한 청각 또는 귀 전달로서 지칭될 수 있다.

[0631] 일부 구현예에서, 본 발명의 RNA 분자는 리포좀 또는 리포펙틴 제형 등으로 전달되고, 당업자에게 잘 공지된 방법에 의해 제조될 수 있다. 이러한 방법은, 예를 들어, 본 명세서에 참고로 포함된 미국 특허 제5,593,972호, 제5,589,466호, 및 제5,580,859호에 기재되어 있다.

[0632] 포유류 세포 내로 siRNA의 향상 및 개선된 전달에서 구체적으로 목표로 하는 전달 시스템이 개발되었고, (예를 들어, 문헌[Shen et al FEBS Let. 2003, 539:111-114; Xia et al., Nat. Biotech. 2002, 20:1006-1010; Reich et al., Mol. Vision. 2003, 9: 210-216; Sorensen et al., J. Mol. Biol. 2003, 327: 761-766; Lewis et al., Nat. Gen. 2002, 32: 107-108 및 Simeoni et al., NAR 2003, 31, 11: 2717-2724] 참조) 본 발명에 적용될 수 있다. siRNA는 최근에 영장류에서 유전자 발현의 저해를 위해 성공적으로 사용되었다(예를 들어, 또한 본 명세서에 적용될 수 있는 문헌[Tolentino et al., Retina 24(4):660] 참조).

[0633] Qi 등은 본 발명의 CRISPR Cas 시스템에 적용될 수 있는 신규한 단백질 전달 기법에 의해 무결함 원창을 통해 내이에 효율적인 siRNA 트렌스펙션을 위한 방법을 개시한다(예를 들어, 문헌[Qi et al., Gene Therapy (2013), 1-9] 참조). 특히, 무결함 원창 침투를 통해 내부 및 외부 모세포, 팽대부릉(crista ampullaris), 난형낭반 및 구형낭반을 포함하는 내이의 세포 내로 Cy3-표지 siRNA를 트렌스펙션할 수 있는 TAT 이중 가닥 RNA-결합 도메인(TAT-DRBD)은 다양한 내이 질병을 치료하고 청각 기능의 보존을 위해 생체내 이중 가닥 siRNA를 전달하는데 성공적이었다. 약 40 ㎕의 10mM RNA가 귀에 투여를 위한 투약량으로서 생각될 수 있다.

[0634] Rejali 등에 따르면(Hear Res. 2007 Jun;228(1-2):180-7), 달팽이관 이식 기능은 이식물에 의한 전기적 자극의

표적인 나선청신경절 뉴런의 양호한 보정에 의해 개선될 수 있으며, 뇌 유래 신경성장 인자(brain derived neurotrophic factor: BDNF)는 실험적으로 청각을 잃은 귀에서 나선청신경절 생존을 향상시키는 것으로 이전에 나타났다. Rejali 등은 BDNF 유전자 삽입을 이용하여 바이러스 벡터에 의해 형질도입된 섬유아세포의 코팅을 포함하는 달팽이관 이식 전극의 변형된 설계를 시험하였다. 이 유형의 생체외 유전자 전달을 달성하기 위해, Rejali 등은 BDNF 유전자 카세트 삽입을 지니는 아데노바이러스를 이용하여 기니아 피그 섬유아세포를 형질도입하였고, 이들 세포가 BDNF를 분비한 다음 아가로스 겔을 통해 달팽이관 이식 전극에 BDNF-분비 세포를 부착한 것을 결정하고, 고실계 내 전극에 이식하였다. Rejali 등은 BDNF 발현 전극이 대조군 전극에 비해 이식 48일 후에 달팽이관의 기본 회전에서 상당히 더 많은 나선청신경절을 보존할 수 있다는 것을 결정하였고, 나선청신경절 뉴런 생존을 향상시키기 위한 생체외 유전자 전달과 달팽이관 이식 요법을 조합하는 것의 실행가능성을 입증하였다. 이러한 시스템은 본 발명의 CRISPR Cas 시스템을 귀에 전달하기 위해 적용될 수 있다.

[0635] Mukherjea 등(Antioxidants & Redox Signaling, Volume 13, Number 5, 2010)은 손상으로부터 OHC의 보호에 의해 입증되는 바와 같이 짧은 간섭 (si) RNA를 사용하는 NOX3의 녹다운이 시스플라틴 이독성(ototoxicity)을 없애며 청성뇌간반응(auditory brainstem response: ABR)에서 역치 이동을 감소시켰다는 것을 기록하였다. 상이한 용량의 siNOX3(0.3, 0.6 및 0.9 ㎍)가 래트에 투여되었고, NOX3 발현은 실시간 RT-PCR에 의해 평가되었다. 사용한 최저용량의 NOX3 siRNA(0.3 ㎍)는 스크램블 RNA 또는 미처리 달팽이관의 경고막 투여와 비교할 때 NOX3 mRNA의 임의의 저해를 나타내지 않았다. 그러나, 더 고용량의 NOX3 siRNA(0.6 및 0.9 ㎍)의 투여는 대조군 스크램블 siRNA에 비해 NOX3 발현이 감소되었다. 이러한 시스템은 인간에 대한 투여를 위해 약 2 ㎎ 내지 약 4 ㎎의 CRISPR Cas의 투약량으로 경고막 투여를 위한 본 발명의 CRISPR Cas 시스템에 적용될 수 있다.

[0636] Jung 등(Molecular Therapy, vol. 21 no. 4, 834-841 apr. 2013)은 Hes5 수준이 siRNA의 적용 후 감소되었고, 이들 포낭 내 모세포 수는 대조군 처리 수보다 상당히 더 컸다는 것을 입증한다. 데이터는 siRNA 기술이 내이 내 수선 및 재발생을 유발하는데 유용할 수 있고, 노치(Notch) 신호처리 경로는 특이적 유전자 발현 저해를 위한 잠재적으로 유용한 표적이라는 것을 시사한다. Jung 등은 2 ㎕ 용적 중의 8 ㎍의 Hes5 siRNA로 주사하였고, 귀의 전정 상피에 동결 siRNA에 대한 멸균 정상 식염수를 첨가함으로써 제조하였다. 이러한 시스템은 인간에 투여를 위해 약 1 내지 약 30 ㎎의 CRISPR Cas의 투약량으로 귀의 전정 상피에 투여를 위해 본 발명의 CRISPR Cas 시스템에 적용될 수 있다.

[0637] **눈**

[0638] 본 발명은 또한 한쪽 또는 양쪽 눈에 CRISPR-Cas 시스템을 전달하는 것을 고려한다.

[0639] 본 발명의 또 다른 양태에서, CRISPR-Cas 시스템은 문헌[Genetic Diseases of the Eye, Second Edition, edited by Elias I. Traboulsi, Oxford University Press, 2012]에서 추가로 기재되는 몇몇 유전자 돌연변이로부터 생긴 눈 결함을 보정하기 위해 사용될 수 있다.

[0640] 눈에 대한 투여를 위해, 렌티바이러스 벡터, 특히 말 전염성 빈혈 바이러스(EIAV)가 특히 바람직하다.

[0641] 다른 구현예에서, 말 전염성 빈혈 바이러스(EIAV)에 기반한 최소 비-영장류 렌티바이러스 벡터가 또한 특히 눈 유전자 요법에 대해 고려된다(예를 들어, Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/jgm.845에서 2005년 11월 21일 온라인 공개된 문헌[Balagaan, J Gene Med 2006; 8: 275 - 285] 참조). 벡터는 표적 유전자의 거대세포바이러스(CMV) 프로모터 구동 발현을 갖는 것으로 고려된다. 전방내, 망막하, 안내 및 유리체내 주사가 모두 고려된다(예를 들어, Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/jgm.845에서 2005년 11월 21일 온라인 공개된 문헌[Balagaan, J Gene Med 2006; 8: 275 - 285] 참조). 안내 주사는 수술용 현미경의 도움으로 수행될 수 있다. 망막하 및 유리체내 주사에 대해, 눈은 가벼운 지압에 의해 탈출될 수 있고, 전문가는 유리 현미경 슬라이드 커버슬립으로 덮은 각막 상에서 한 방울의 결합 매질 용액로 이루어진 콘택트 렌즈 시스템을 사용하여 시각화된다. 망막하 주사에 대해, 5-㎕ 해밀톤 주사기 상에 장착한 10-mm 34-게이지 니들의 팁은 바늘 구멍이 망막하 공간에서 시각화될 수 있을 때까지 후두극(posterior pole) 쪽으로 접하도록 적도부 공막을 통한 직접적 시각화 하에 전진할 수 있다. 이어서, 2 ㎕의 벡터 현탁액이 주사되어 우수한 수포성 망막 박리를 생성하며, 따라서, 망막하 벡터 투여를 확인할 수 있다. 이 접근은 RPE에 의해 흡수될 때까지 보통 절차의 48 시간 내에, 벡터 현탁액이 망막하 공간 내에 보유되게 하는 자기-밀봉 공막 절개를 생성한다. 이 절차는 하부 반구 내에서 반복되어 하부 망막 박리를 생성할 수 있다. 이 기법은 대략 70%의 감각신경 망막 및 RPE의 벡터 현탁액에 대한 노출을 야기한다. 유리체내 주사에 대해, 니들 팁은 각공막 경계 후측의 공막 1 ㎜를 통해 전진할 수 있고 2 ㎕의 벡터 현탁액이 유리체 내로 주사될 수 있다. 전방내 주사에 대해, 니들 팁은 각공막 윤부 천자를 통해 전진할 수 있고, 중심 각막쪽으로 향하며, 2 ㎕의 벡터 현탁액이 주

사될 수 있다. 전방내 주사에 대해, 니들 팁은 각공막 윤부 천자를 통해 전진할 수 있고, 중심 각막쪽으로 향하며, 2 ㎕의 벡터 현탁액이 주사될 수 있다. 이들 벡터는 1.0-1.4 × 10^{10} 또는 1.0-1.4 × 10^{9} 형질도입 단위(TU)/㎖ 중 하나의 역가로 주사될 수 있다.

[0642] 다른 구현예에서, 습윤 형태의 노인성 황반변성의 치료를 위한 망막하 주사를 통해 전달된 혈관형성억제 단백질 엔도스타틴 및 앤지오스타틴을 발현시키는 말 전염성 빈혈 바이러스-기반 렌티바이러스 유전자 요법 벡터인 RetinoStat®가 또한 고려된다(예를 들어, 문헌[Binley et al., 인간 유전자 요법 23:980-991 (September 2012)] 참조). 이러한 벡터는 본 발명의 CRISPR-Cas 시스템에 대해 변형될 수 있다. 각각의 눈은 총 용적 100 ㎕에서 1.1 x 10^{5} 형질도입 단위/눈(TU/눈)의 용량으로 RetinoStat® 중 하나로 처리될 수 있다.

[0643] 다른 구현예에서, E1-, 부분적 E3-, E4-결실 아데노바이러스 벡터는 눈에 전달을 위해 고려될 수 있다. 진행된 신생혈관 노인성 황반변성(AMD)을 지니는 28 명의 환자에게 E1-, 부분적 E3-, E4-결실 아데노바이러스 벡터 발현 인간 색소 상피-유래 인자(AdPEDF.11)의 단일 유리체강내 주사를 제공하였다(예를 들어, 문헌[Campochiaro et al., Human Gene Therapy 17:167-176 (February 2006)] 참조). 10^{6} 내지 $10^{9.5}$개 입자 단위(PU)의 범위에 있는 용량을 조사하였고, AdPEDF.11과 관련된 심각한 유해 사건 및 용량 의존적 독성은 없었다(예를 들어, 문헌[Campochiaro et al., 인간 유전자 요법 17:167-176 (February 2006)] 참조). 아데노바이러스 벡터-매개 눈 유전자 전달은 눈 장애의 치료를 위한 실행가능한 접근이 되는 것으로 나타나며, CRISPR Cas 시스템에 적용될 수 있었다.

[0644] 다른 구현예에서, RXi Pharmaceuticals의 sd-rxRNA® 시스템은 눈에 CRISPR Cas를 전달하기 위해 사용되고/사용되거나 적합하게 될 수 있다. 이 시스템에서, 3 ㎍의 sd-rxRNA의 단일 유리체내 투여는 14 일 동안 PPIB mRNA 수준의 서열-특이적 감소를 초래한다. sd-rxRNA® 시스템은 본 발명의 CRISPR Cas 시스템에 적용될 수 있으며, 인간에게 투여되는 약 3 내지 20 ㎎의 CRISPR 용량을 고려한다.

[0645] Millington-Ward 등(Molecular Therapy, vol. 19 no. 4, 642-649 apr. 2011)은 RNA 간섭(RNAi)-기반 로돕신 억제자 및 RNAi 표적 부위에 걸쳐 축퇴 위치에서 뉴클레오티드 변경에 기인하는 억제에 내성이 있는 코돈-변형 로돕신 대체 유전자를 전달하는 아데노-연관 바이러스(AAV) 벡터를 기재한다. 6.0 x 10^{8} vp 또는 1.8 x 10^{10} vp AAV의 주사는 Millington-Ward 등에 의해 눈에 망막하로 주사되었다. Millington-Ward 등의 AAV 벡터는 본 발명의 CRISPR Cas 시스템에 적용될 수 있으며, 인간을 통해 투여되는 약 2 x 10^{11} 내지 약 6 x 10^{13} vp 용량을 고려한다.

[0646] Dalkara 등(Sci Transl Med 5, 189ra76 (2013))은 또한 눈의 유리체액 내로 손상없는 주사 후에 망막을 통해 결함있는 유전자의 야생형 형태를 전달하는 AAV 벡터를 생성하는 생체내 관련 진화에 관한 것이다. Dalkara는 7량체 펩타이드 디스플레이 라이브러리 및 AAV1, 2, 4, 5, 6, 8 및 9로부터의 캡 유전자의 DNA 셔플링에 의해 구성된 AAV 라이브러리를 기재한다. CAG 또는 Rho 프로모터 하의 rcAAV 라이브러리 및 rAAV 벡터 발현 GFP를 패키징하고 나서, 데옥시리보뉴클레아제-내성 게놈 역가를 정량적 PCR을 통해 얻었다. 라이브러리를 풀링하고 나서, 2 회의 진화를 수행하였는데, 각각은 초기 라이브러리 다양화 다음에 3 개의 생체내 선택 단계로 이루어진다. 각각의 이러한 단계에서, P30 rho-GFP 마우스에 2 ㎖의 이오딕사놀-정제, 인산염 완충 식염수(PBS)로 유리체내 주사하였고-약 1 × 10^{12} vg/㎖의 게놈 역가로 라이브러리를 투석하였다. Dalkara 등의 AAV 벡터는 본 발명의 CRISPR Cas 시스템에 적용될 수 있으며, 인간에 대해 투여되는 약 1 x 10^{15} 내지 약 1 x 10^{16} vg/㎖의 용량을 고려한다.

[0647] 다른 구현예에서, 로돕신 유전자는 망막색소변성증(retinitis pigmentosa: RP)의 치료를 위해 표적화될 수 있되, Sangamo BioSciences, Inc.에 의해 양도 된 미국 특허 공개 제20120204282호의 시스템은 본 발명의 CRISPR Cas 시스템에 따라 변형될 수 있다.

[0648] 다른 구현예에서, 인간 로돕신 유전자로부터 표적 서열을 절단하는 방법에 관한 Cellectis에 의해 양도 된 미국 특허 공개 제20130183282호의 방법은 또한 본 발명의 CRISPR Cas 시스템에 대해 변형될 수 있다.

[0649] Academia Sinica에 의해 양도 된 미국 특허 공개 제20130202678호는 눈 내의 망막하 또는 유리체내 공간에 대해 Puf-A 유전자(이는 눈 조직의 망막 신경절 및 색소 세포에서 발현되며 독특한 항-세포자멸사 활성을 나타냄)를 전달하는 것에 관한 망막병증 및 약간 위협적인 안과적 장애를 치료하기 위한 방법에 관한 것이다. 특히, 바람직한 표적은 zgc:193933, prdm1a, spata2, tex10, rbb4, ddx3, zp2.2, Blimp-1 및 HtrA2이며, 이들 모두는 본

발명의 CRISPR Cas 시스템에 의해 표적화될 수 있다.

[0650] Wu(Cell Stem Cell,13:659-62, 2013)는 DNA 절단을 유발한 경우 마우스에서 백내장을 야기하는 단일 염기쌍 돌연변이에 Cas9를 안내하는 가이드 RNA를 설계하였다. 이어서 접합자 수선 메커니즘에 제공된 다른 야생형 대립유전자 또는 올리고는 파손 대립유전자의 서열을 보정하였고, 돌연변이체 마우스에서 백내장의 원인이 되는 유전자 결함을 보정하였다.

[0651] 미국 특허 공개 제20120159653호는 황반변성(MD)과 관련된 세포, 동물 및 단백질을 유전자 변형하기 위한 아연 핑거 뉴클레아제의 용도를 기재한다. 황반변성 (MD)은 노인에서 시각 손상의 주된 원인일 뿐만 아니라 아동 질환, 예컨대 스타르카르트병, 소르비 안저 및 치명적 아동 신경퇴행성 질환의 특징적 증상이고, 개시 연령이 유아로 어리다. 황반변성은 망막에 대한 손상 때문에 시야 중심(황반)에서 시력 상실을 야기한다. 현재 존재하는 동물 모델은 그것이 인간에서 관찰되기 때문에 질환의 주된 특징을 요약하지 않는다. MD와 관련된 단백질을 암호화하는 돌연변이체 유전자를 포함하는 이용가능한 동물 모델은 또한 고도로 가변성인 표현형을 생성하는데, 이는 인간 질환 및 요법 개발에 대한 번역에 문제를 만든다.

[0652] 미국 특허 공개 제20120159653호의 일 양태는 본 발명의 CRISPR Cas 시스템에 적용될 수 있는 MD와 관련된 단백질을 암호화하는 임의의 염색체 서열 편집에 관한 것이다. MD와 관련된 단백질은 통상적으로 MD와 관련된 단백질의 MD 장애에 대한 실험적 관련에 기반하여 선택된다. 예를 들어, MD와 관련된 단백질의 생성 속도 또는 순환 농도는 MD 장애가 없는 집단에 비해 MD 장애가 있는 집단에서 상승 또는 억제될 수 있다. 단백질 수준의 차이는 웨스턴 블롯, 면역조직화학적 염색, 효소결합 면역흡착법(ELISA) 및 질량 분석법을 포함하지만, 이들로 제한되지 않는 프로테오믹 기법을 사용하여 평가될 수 있다. 대안적으로, MD와 관련된 단백질은 DNA 마이크로어레이 분석, 일련 분석 유전자 발현(serial analysis of gene expression: SAGE), 및 정량적 실시간 중합효소 연쇄 반응(Q-PCR)을 포함하지만, 이들로 제한되지 않는 게놈 기법을 사용하여 단백질을 암호화하는 유전자의 유전자 발현 프로파일에 의해 동정될 수 있다.

[0653] 비제한적 예로서, MD와 관련된 단백질은 다음의 단백질을 포함하지만, 이들로 제한되지 않는다: (ABCA4) ATP-결합 카세트, 서브 패밀리 A(ABC1), 구성원 4 ACHM1 완전색맹(간상체 전색맹) 1 ApoE 아포지방단백질 E(ApoE) C1QTNF5(CTRP5) C1q 및 종양 괴사 인자 관련 단백질 5(C1QTNF5) C2 상보체 구성성분 2(C2) C3 상보체 구성성분(C3) CCL2 케모카인(C-C 모티프) 리간드 2(CCL2) CCR2 케모카인(C-C 모티프) 수용체 2(CCR2) 분화 36의 CD36 클러스터 CFB 상보체 인자 B CFH 상보체 인자 CFH H CFHR1 상보체 인자 H-관련 1 CFHR3 상보체 인자 H-관련 3 CNGB3 사이클릭 뉴클레오티드화된 채널 베타 3 CP 셀룰로플라스민(CP) CRP C 반응 단백질(CRP) CST3 시스타틴 C 또는 시스타틴 3(CST3) CTSD 카텝신 D(CTSD) CX3CR1 케모카인(C-X3-C 모티프) 수용체 1 ELOVL4 신장 of 매우 장쇄 지방산 4 ERCC6 절단 수선 교차 상보성 설치류 수선 결함, 상보체화 그룹 6 FBLN5 피불린-5 FBLN5 피불린 5 FBLN6 피불린 6 FSCN2 파신(FSCN2) HMCN1 헤미센트린 1 HMCN1 헤미센트린 1 HTRA1 HtrA 세린 펩티다제 1(HTRA1) HTRA1 HtrA 세린 펩티다제 1 IL-6 인터류킨 6 IL-8 인터류킨 8 LOC387715 가설 단백질 PLEKHA1 플렉스트린 상동성 도메인-함유 패밀리 A 구성원 1 (PLEKHA1) PROM1 프로미닌 1(PROM1 또는 CD133) PRPH2 페리페린-2 RPGR 망막색소변성증 GTPase 조절자 세르핀G1 세르핀 펩티다제 저해제, 클레이트 G, 구성원 1(C1- 저해제) TCOF1 트리클 TIMP3 메탈로프로테이나제 저해제 3 (TIMP3) TLR3 Toll-유사 수용체 3

[0654] 염색체 서열이 편집된 MD와 관련된 단백질의 동일성은 다를 수 있고, 다를 것이다. 바람직한 구현예에서, 염색체 서열이 편집된 MD와 관련된 단백질은 ATP-결합 카세트, ABCR 유전자에 의해 암호화된 서브 패밀리 A(ABC1) 구성원 4 단백질(ABCA4), APOE 유전자에 의해 암호화된 아포지방단백질 E 단백질(APOE), CCL2 유전자에 의해 암호화된 케모카인 (C-C 모티프) 리간드 2 단백질(CCL2), CCR2 유전자에 의해 암호화된 케모카인 (C-C 모티프) 수용체 2 단백질(CCR2), CP 유전자에 의해 암호화된 셀룰로플라스민 단백질(CP), CTSD 유전자에 의해 암호화된 카텝신 D 단백질(CTSD), 또는 TIMP3 유전자에 의해 암호화된 메탈로프로테이나제 저해제 3 단백질(TIMP3)일 수 있다. 예시적인 구현예에서, 유전자 변형된 동물은 래트이며, MD와 관련된 단백질을 암호화하는 편집된 염색체 서열은: (ABCA4) ATP-결합 카세트, NM_000350 서브 패밀리 A(ABC1), 구성원 4 APOE 아포지방단백질 E NM_138828(APOE) CCL2 케모카인(C-C NM_031530 모티프) 리간드 2 (CCL2) CCR2 케모카인(C-C NM_021866 모티프) 수용체 2(CCR2) CP 셀룰로플라스민(CP) NM_012532 CTSD 카텝신 D (CTSD) NM_134334 TIMP3 메탈로프로테이나제 NM_012886 저해제 3(TIMP3)일 수 있다. 동물 또는 세포는 MD와 관련된 단백질을 암호화하는 1, 2, 3, 4, 5, 6, 7 이상의 방해 염색체 서열 및 MD와 관련된 단백질을 암호화하는 0, 1, 2, 3, 4, 5, 6, 7 이상의 염색체 통합된 서열을 포함할 수 있다.

[0655] 편집 또는 통합된 염색체 서열은 MD와 관련된 변용된 단백질을 암호화하도록 변형될 수 있다. MD-관련 염색체

서열 내 몇몇 돌연변이는 MD와 관련되었다. MD와 관련된 염색체 서열 내 돌연변이의 비제한적 예는 ABCR 단백질에서, E471K(즉, 위치 471에서 글루타메이트가 리신으로 변함), R1129L(즉, 위치 1129에서 아르기닌이 류신으로 변함), T1428M(즉, 위치 1428에서 트레오닌이 메티오닌으로 변함), R1517S (즉, 위치 1517에서 아르기닌이 세린으로 변함), I1562T (즉, 위치 1562에서 이소류신이 트레오닌으로 변함), 및 G1578R(즉, 위치 1578에서 글리신이 아르기닌으로 변함); CCR2 단백질에서, V64I(즉, 위치 192에서 발린이 이소류신으로 변함); CP 단백질에서, G969B(즉, 위치 969에서 글리신이 아스파라긴 또는 아스파르테이트로 변함); TIMP3 단백질에서, S156C(즉, 위치 156에서 세린이 시스테인으로 변함), G166C(즉, 위치 166에서 글리신이 시스테인으로 변함), G167C(즉, 위치 167에서 글리신이 시스테인으로 변함), Y168C (즉, 위치 168에서 티로신이 시스테인으로 변함), S170C(즉, 위치 170에서 세린이 시스테인으로 변함), Y172C(즉, 위치 172에서 티로신이 시스테인으로 변함) 및 S181C(즉, 위치 181에서 세린이 시스테인으로 변함)를 비롯하여 MD를 야기할 수 있는 것을 포함한다. 다른 MD-관련 유전자 및 질환에서 유전적 변이체의 다른 연관은 당업계에 공지되어 있다.

[0656] **심장**

[0657] 본 발명은 또한 심장에 CRISPR-Cas 시스템을 전달하는 것을 고려한다. 심장에 대해, 심근 트로픽 아데나-연관 바이러스(AAVM), 특히 심장에서 우선적인 유전자 전달을 나타내는 AAVM41가 바람직하다(예를 들어, 문헌[Lin-Yanga et al., PNAS, March 10, 2009, vol. 106, no. 10] 참조). 투여는 전신 또는 국소일 수 있다. 약 1-10 x 10^{14} 벡터 게놈의 투약량이 전신 투여에 대해 고려된다. 또한 예를 들어, 문헌[Eulalio et al. (2012) Nature 492: 376 및 Somasuntharam et al. (2013) Biomaterials 34: 7790] 참조.

[0658] 예를 들어, 미국 특허 공개 제20110023139호는 심혈관 질환과 관련된 세포, 동물 및 단백질을 유전자 변형하기 위한 아연 핑거 뉴클레아제의 사용을 기재한다. 심혈관 질환은 일반적으로 고혈압, 심장 마비, 심부전 및 뇌졸중 및 TIA를 포함한다. 심혈관 질환에 수반된 임의의 염색체 서열 또는 심혈관 질환에 수반된 임의의 염색체 서열에 의해 암호화된 단백질은 본 개시내용에 기재된 방법에서 이용될 수 있다. 심혈관-관련 단백질은 통상적으로 심혈관 질환의 발생에 대한 심혈관 관련 단백질의 실험적 연관에 기반하여 선택된다. 예를 들어, 심혈관-관련 단백질의 생성 속도 또는 순환 농도는 심혈관 장애가 없는 집단에 비해 심혈관 장애를 갖는 집단에서 상승 또는 감소될 수 있다. 단백질 수준의 차이는 웨스턴 블롯, 면역조직화학염색, 효소결합 면역흡착분석(ELISA) 및 질량분석법을 포함하지만, 이들로 제한되지 않는 프로테오믹 기법을 사용하여 평가될 수 있다. 대안적으로, 심혈관-관련 단백질은 DNA 마이크로어레이 분석, 일련 분석 유전자 발현(SAGE), 및 정량적 실시간 중합효소 연쇄반응(Q-PCR)을 포함하지만, 이들로 제한되지 않는 게놈 기법을 사용하여 단백질을 암호화하는 유전자의 유전자 발현 프로파일을 얻음으로써 동정될 수 있다.

[0659] 예로서, 염색체 서열은 IL1B(인터류킨 1, 베타), XDH(잔틴 탈수소효소), TP53(종양 단백질 p53), PTGIS (프로스타글란딘 12(프로스타사이클린) 신타제), MB(미오글로빈), IL4(인터류킨 4), ANGPT1(앤지오포이에틴 1), ABCG8(ATP-결합 카세트, 서브 패밀리 G(WHITE), 구성원 8), CTSK(카텝신 K), PTGIR(프로스타글란딘 12(프로스타사이클린) 수용체(IP)), KCNJ11(칼륨 내부 교정 채널, 서브패밀리 J, 구성원 11), INS(인슐린), CRP(C-반응 단백질, 펜트락신-관련), PDGFRB(혈소판 유래 성장 인자 수용체, 베타 폴리펩티드), CCNA2(사이클린 A2), PDGFB(혈소판 유래 성장 인자 베타 폴리펩티드(시미안 육종 바이러스(v-sis) 발암유전자 상동체)), KCNJ5(칼륨 내부 교정 채널, 서브패밀리 J, 구성원 5), KCNN3(칼륨 중간체/소 전도도 칼슘-활성화 채널, 서브패밀리 N, 구성원 3), CAPN10(칼페인 10), PTGES (프로스타글란딘 E 신타제), ADRA2B(아드레날린, 알파-2B-, 수용체), ABCG5 (ATP-결합 카세트, 서브 패밀리 G(WHITE), 구성원 5), PRDX2(페록시레독신 2), CAPN5(칼페인 5), PARP14(폴리(ADP-리보스) 중합효소 패밀리, 구성원 14), MEX3C(mex-3 상동체 C(예쁜꼬마선충)), ACE 앤지오텐신 I 전환효소(펩티딜-디펩티다제 A) 1), TNF(종양 괴사 인자(TNF 수퍼패밀리, 구성원 2)), IL6(인터류킨 6(인터페론, 베타 2)), STN(스타틴), 세르핀E1(세르핀 펩티다제 저해제, 클레이드 E(넥신, 플라스미노겐 활성인자 저해제 1형), 구성원 1), ALB(알부민), ADIPOQ(아디포넥틴, C1Q 및 콜라겐 도메인 함유), APOB(아포지방단백질 B(Ag(x) 항원 포함)), APOE(아포지방단백질 E), LEP(렙틴), MTHFR(5,10-메틸렌테트라하이드로폴레이트 환원효소(NADPH)), APOA1(아포지방단백질 A-I), EDN1(엔도텔린 1), NPPB(나트륨이뇨 펩타이드 전구체 B), NOS3(일산화질소 신타제 3(내피세포)), PPARG(페록시좀 증식제-활성화 수용체 감마), PLAT(플라스미노겐 활성인자, 조직), PTGS2(프로스타글란딘-엔도페록시드 신타제 2 (프로스타글란딘 G/H 신타제 및 사이클로옥시게나제)), CETP (콜레스테릴 에스테르 전달 단백질, 혈장), AGTR1(앤지오텐신 II 수용체, 1형), HMGCR(3-하이드록시-3-메틸글루타릴-코엔자임 A 환원효소), IGF1(인슐린-유사 성장 인자 1(소마토메딘 C)), SELE(셀렉틴 E), REN(레닌), PPARA(페록시좀 증식제-활성화 수용체 알파), PON1(파라옥소나제 1), KNG1(키니노겐 1), CCL2(케모카인 (C-C 모티프) 리간드 2), LPL(지방단백질 리파제), VWF(폰 빌레브란트 인자), F2(응고인자 II(트롬빈)), ICAM1(세포내 부착

분자 1), TGFB1(형질전환 성장 인자, 베타 1), NPPA(나트륨이뇨 펩타이드 전구체 A), IL10 (인터류킨 10), EPO(에리트로포이에틴), SOD1 (수퍼옥시드 디스무타제 1, 가용성), VCAM1(혈관 세포 부착 분자 1), IFNG (인터페론, 감마), LPA (지방단백질, Lp(a)), MPO(미엘로퍼옥시다제), ESR1(에스트로겐 수용체 1), MAPK1(미토겐-활성화 단백질 키나제 1), HP(합토글로빈), F3 (응고인자 III(트롬보플라스틴, 조직인자)), CST3(시스타틴 C), COG2(올리고머 골지 복합체 2의 구성성분), MMP9(매트릭스 메탈로펩티다제 9(젤라티나제 B, 92 kDa 젤라티나제, 92 kDa IV형 콜라게나제)), 세르핀C1(세르핀 펩티다제 저해제, 클레이드 C(항트롬빈), 구성원 1), F8(응고인자 VIII, 응혈 촉진성 구성성분), HMOX1(헴 옥시게나제(디사이클링) 1), APOC3(아포지방단백질 C-III), IL8(인터류킨 8), PROK1(프로키네티신 1), CBS(시스타티오닌-베타-신타제), NOS2(일산화질소 신타제 2, 유도성), TLR4(톨-유사 수용체 4), SELP(셀렉틴 P(과립막 단백질 140 kDa, 항원 CD62)), ABCA1(ATP-결합 카세트, 서브 패밀리 A (ABC1), 구성원 1), AGT(앤지오텐시노겐(세르핀 펩티다제 저해제, 클레이드 A, 구성원 8)), LDLR(저밀도 지방단백질 수용체), GPT(글루타민-피루브산 트랜스아미나제(알라닌 아미노트랜스퍼라제)), VEGFA(혈관표피성장인자 A), NR3C2(핵 수용체 서브패밀리 3, 그룹 C, 구성원 2), IL18(인터류킨 18(인터페론-감마-유도 인자)), NOS1 (일산화질소 신타제 1(뉴런)), NR3C1(핵 수용체 서브패밀리 3, 그룹 C, 구성원 1(글루코코르티코이드 수용체)), FGB(피브리노겐 베타쇄), HGF(간세포 성장 인자(헤파포이에틴 A; 산란인자)), IL1A(인터류킨 1, 알파), RETN(레시스틴), AKT1(v-akt 뮤린 흉선종 바이러스 발암유전자 상동체 1), LIPC(리파제, 간), HSPD1(열 충격 60 kDa 단백질 1(샤페로닌)), MAPK14 (미토겐-활성화 단백질 키나제 14), SPP1(분비된 포스포단백질 1), ITGB3(인테그린, 베타 3(혈소판 글리코단백질 111a, 항원 CD61)), CAT(카탈라제), UTS2(유로텐신 2), THBD(트롬보모듈린), F10(응고인자 X), CP (셀룰로플라스민(페록시다제)), TNFRSF11B (종양 괴사 인자 수용체 수퍼 패밀리, 구성원 11b), EDNRA(엔도텔린 수용체 A형), EGFR(표피 성장 인자 수용체(적아세포 백혈병 바이러스(v-erb-b) 발암유전자 상동체, 조류)), MMP2(매트릭스 메탈로펩티다제 2(젤라티나제 A, 72 kDa 젤라티나제, 72 kDa IV형 콜라게나제)), PLG (플라스미노겐), NPY(뉴로펩타이드 Y), RHOD(ras 상동체 유전자 패밀리, 구성원 D), MAPK8(미토겐-활성화 단백질 키나제 8), MYC(v-myc 골수세포증 바이러스 발암유전자 상동체(조류)), FN1(피브로넥틴 1), CMA1 (키마제 1, 비만세포), PLAU(플라스미노겐 활성인자, 유로키나제), GNB3(구아닌 뉴클레오티드 결합 단백질(G 단백질), 베타 폴리펩티드 3), ADRB2(아드레날린, 베타-2-, 수용체, 표면), APOA5 (아포지방단백질 A-V), SOD2 (수퍼옥시드 디스무타제 2, 미토콘드리아), F5 (응고인자 V(프로악셀레린, 불안정인자)), VDR(비타민 D (1,25-디하이드록시 비타민 D3) 수용체), ALOX5 (아라키돈산 5-리폭시게나제), HLA-DRB1(주조직적합성 복합체, 클래스 II, DR 베타 1), PARP1(폴리(ADP-리보스) 중합효소 1), CD40LG(CD40 리간드), PON2 (파라옥소나제 2), AGER(진행된 글리코실화 최종 생성물-특이적 수용체), IRS1 (인슐린 수용체 기질 1), PTGS1(프로스타글란딘-엔도페록시드 신타제 1 (프로스타글란딘 G/H 신타제 및 사이클로옥시게나제)), ECE1 (엔도텔린 전환효소 1), F7(응고인자 VII(혈청 프로트롬빈 전환 촉진제)), URN (인터류킨 1 수용체 길항물질), EPHX2(에폭시드 가수분해효소 2, 세포질ic), IGFBP1(인슐린-유사 성장 인자 결합 단백질 1), MAPK10(미토겐-활성화 단백질 키나제 10), FAS(Fas (TNF 수용체 수퍼패밀리, 구성원 6)), ABCB1(ATP-결합 카세트, 서브 패밀리 B(MDR/TAP), 구성원 1), JUN(jun 발암유전자), IGFBP3(인슐린-유사 성장 인자 결합 단백질 3), CD14(CD14 분자), PDE5A(포스포디에스터라제 5A, cGMP-특이적), AGTR2(앤지오텐신 II 수용체, 2형), CD40(CD40 분자, TNF 수용체 수퍼패밀리 구성원 5), LCAT (레시틴-콜레스테롤 아실트랜스퍼라제), CCR5(케모카인(C-C 모티프) 수용체 5), MMP1(매트릭스 메탈로펩티다제 1(사이질 콜라게나제)), TIMP1(TIMP 메탈로펩티다제 저해제 1), ADM(아드레노메둘린), DYT10(근육긴장이상 10), STAT3(전사 3의 신호 변환 및 활성인자(급성기 반응 인자)), MMP3 (매트릭스 메탈로펩티다제 3(스트로멜리신 1, 프로젤라티나제)), ELN(엘라스틴), USF1(상류 전사 인자 1), CFH(상보체 인자 H), HSPA4(열 충격 70 kDa 단백질 4), MMP12(매트릭스 메탈로펩티다제 12(대식세포 엘라스타제)), MME(막 메탈로-엔도펩티다제), F2R(응고인자 II(트롬빈) 수용체), SELL(셀렉틴 L), CTSB(카텝신 B), ANXA5(아넥신 A5), ADRB1(아드레날린, 베타-1-, 수용체), CYBA(사이토크롬 b-245, 알파 폴리펩타이드), FGA(피브리노겐 알파 쇄), GGT1(감마-글루타밀트랜스퍼라제 1), LIPG(리파제, 내피), HIF1A(저산소증 유발 인자 1, 알파 서브유닛(기본나선-루프-나선 전사 인자)), CXCR4(케모카인(C-X-C 모티프) 수용체 4), PROC(단백질 C(응고인자 Va 및 VIIIa의 활성인자에서)), SCARB1(청소수용체 B형, 구성원 1), CD79A(CD79a 분자, 면역글로불린-연관 알파), PLTP(인지질 전달 단백질), ADD1(아듀신 1(알파)), FGG(피브리노겐 감마 쇄), SAA1(혈청 아밀로이드 A1), KCNH2(칼륨 전압 개폐 채널, 서브패밀리 H (eag-관련), 구성원 2), DPP4 (디펩티딜-펩티다제 4), G6PD(글루코스-6-인산염 탈수소효소), NPR1(나트륨이뇨 펩타이드 수용체 A/구아닐레이트 사이클라제 A(아트리오나트륨이뇨 펩타이드 수용체 A)), VTN (비트로넥틴), KIAA0101(KIAA0101), FOS(FBJ 뮤린 골육종 바이러스 발암유전자 상동체), TLR2(톨-유사 수용체 2), PPIG(펩티딜프롤린 이성화효소 G(사이클로필린 G)), IL1R1(인터류킨 1 수용체, I형), AR(안드로겐 수용체), CYP1A1(사이토크롬 P450, 패밀리 1, 서브패밀리 A, 폴리펩타이드 1), 세르핀A1(세르핀 펩티다제 저해제, 클레이드 A (알파-1 안티프로테이나제, 항트립신), 구성원 1), MTR(5-메틸테트라하이드로폴레이트-호모시스테인

메틸트랜스퍼라제), RBP4(레티놀 결합 단백질 4, 혈장), APOA4(아포지방단백질 A-IV), CDKN2A(사이클린-의존적 키나제 저해제 2A(흑색종, p16, CDK4를 저해)), FGF2(섬유아세포 성장 인자 2(염기성)), EDNRB(엔도텔린 수용체 B형), ITGA2(인테그린, 알파 2(CD49B, VLA-2 수용체의 알파 2 서브유닛)), CABIN1(칼시뉴린 결합 단백질 1), SHBG(성호르몬-결합 글로불린), HMGB1(고 이동성 그룹 박스 1), HSP90B2P(열 충격 단백질 90 kDa 베타(Grp94), 구성원 2(위유전자)), CYP3A4(사이토크롬 P450, 패밀리 3, 서브패밀리 A, 폴리펩타이드 4), GJA1(갭 연접 단백질, 알파 1, 43 kDa), CAV1(카베올린 1, 포낭 단백질, 22 kDa), ESR2(에스트로겐 수용체 2 (ER 베타)), LTA(림포톡신 알파(TNF 수퍼패밀리, 구성원 1)), GDF15(성장 분화 인자 15), BDNF(뇌 유래 신경영양인자), CYP2D6(사이토크롬 P450, 패밀리 2, 서브패밀리 D, 폴리펩타이드 6), NGF(신경 성장 인자(베타 폴리펩티드)), SP1(Sp1 전사 인자), TGIF1(TGFB-유발 인자 호메오박스 1), SRC(v-src 육종(슈미트-루핀 A-2) 바이러스 발암유전자 상동체(조류)), EGF(표피 성장 인자 (베타-유로가스트론)), PIK3CG(포스포이노시티드-3-키나제, 촉매적, 감마 폴리펩타이드), HLA-A(주조직적합성 복합체, 클래스 I, A), KCNQ1 (칼륨 전압 개폐 채널, KQT-유사 서브패밀리, 구성원 1), CNR1(칸나비노이드 수용체 1(뇌)), FBN1 (피브릴린 1), CHKA(콜린 키나제 알파), BEST1(베스트로핀 1), APP(아밀로이드 베타 (A4) 전구체 단백질), CTNNB1(카테닌(카데린-관련 단백질), 베타 1, 88 kDa), IL2(인터류킨 2), CD36(CD36 분자(트롬보스폰딘 수용체)), PRKAB1(단백질 키나제, AMP-활성화된, 베타 1 비-촉매적 서브유닛), TPO(갑상선 페록시다제), ALDH7A1(알데하이드 탈수소효소 7 패밀리, 구성원 A1), CX3CR1(케모카인(C-X3-C 모티프) 수용체 1), TH(티로신 하이드록실라제), F9(응고인자 IX), GH1(성장 호르몬 1), TF(트랜스페린), HFE(혈색소증), IL17A(인터류킨 17A), PTEN(포스파타제 및 텐신 상동체), GSTM1(글루타티온 S-트랜스퍼라제 뮤 1), DMD(디스트로핀), GATA4(GATA 결합 단백질 4), F13A1(응고인자 XIII, A1 폴리펩타이드), TTR(트랜스티레틴), FABP4(지방산 결합 단백질 4, 지방세포), PON3 (파라옥소나제 3), APOC1(아포지방단백질 C-I), INSR(인슐린 수용체), TNFRSF1B(종양 괴사 인자 수용체 수퍼패밀리, 구성원 1B), HTR2A(5-하이드록시트립타민(세로토닌) 수용체 2A), CSF3(집락자극인자 3(과립구)), CYP2C9(사이토크롬 P450, 패밀리 2, 서브패밀리 C, 폴리펩타이드 9), TXN(티오레독신), CYP11B2(사이토크롬 P450, 패밀리 11, 서브패밀리 B, 폴리펩타이드 2), PTH(부갑상선 호르몬), CSF2(집락자극인자 2(과립구-대식세포)), KDR(키나제 삽입 도메인 수용체(III형 수용체 티로신 키나제)), PLA2G2A(포스포리파제 A2, 그룹 IIA(혈소판, 활액)), B2M(베타-2-마이크로글로불린), THBS1(트롬보스폰딘 1), GCG(글루카곤), RHOA(ras 상동체 유전자 패밀리, 구성원 A), ALDH2(알데히드 탈수소효소 2 패밀리(미토콘드리아)), TCF7L2(전사 인자 7-유사 2 (T-세포 특이적, HMG-박스)), BDKRB2(브래디키닌 수용체 B2), NFE2L2(핵 인자(적혈구-유래 2)-유사 2), NOTCH1(Notch 상동체 1, 전좌-관련(초파리)), UGT1A1(UDP 글루쿠로노실 트랜스퍼라제 1 패밀리, 폴리펩타이드 A1), IFNA1(인터페론, 알파 1), PPARD(페록시좀 증식제-활성화 수용체 델타), SIRT1(시르루인(침묵 교배형 정보 조절 2 상동체) 1(사카로마이세스 세레비시에)), GNRH1(성선자극호르몬-방출 호르몬 1(황체형성-방출 호르몬)), PAPPA(임신-관련 혈장 단백질 A, 파팔리신1), ARR3(아레스틴 3, 레티날(X-아레스틴)), NPPC(나트륨이뇨 펩타이드 전구체 C), AHSP(알파 헤모글로빈 안정화 단백질), PTK2(PTK2 단백질 티로신 키나제 2), IL13(인터류킨 13), MTOR(라파마이신의 메커니즘 표적(세린/트레오닌 키나제)), ITGB2(인테그린, 베타 2(상보체 구성성분 3 수용체 3 및 4 서브유닛)), GSTT1(글루타티온 S-트랜스퍼라제 쎄타 1), IL6ST(인터류킨 6 신호 변환기(gp130, 온코스타틴 M 수용체)), CPB2(카복시펩티다제 B2(혈장)), CYP1A2(사이토크롬 P450, 패밀리 1, 서브패밀리 A, 폴리펩타이드 2), HNF4A(간세포 핵 인자 4, 알파), SLC6A4(용질 운반체 패밀리 6(신경전달물질 수송체, 세로토닌), 구성원 4), PLA2G6(포스포리파제 A2, VI 그룹(사이토졸, 칼슘-의존적)), TNFSF11(종양 괴사 인자(리간드) 수퍼패밀리, 구성원 11), SLC8A1(용질 운반체 패밀리 8(나트륨/칼슘 교환기), 구성원 1), F2RL1(응고인자 II(트롬빈) 수용체-유사 1), AKR1A1(알도-케토 환원효소 패밀리 1, 구성원 A1(알데히드 환원효소)), ALDH9A1(알데히드 탈수소효소 9 패밀리, 구성원 A1), BGLAP(뼈 감마-카복시글루타메이트(gla) 단백질), MTTP(마이크로솜 트리글리세리드 전달 단백질), MTRR (5-메틸테트라하이드로폴레이트-호모시스테인 메틸트랜스퍼라제 환원효소), SULT1A3(설포트랜스퍼라제 패밀리, 사이토졸, 1A, 페놀-선호, 구성원 3), RAGE(신장 종양 항원), C4B(상보체 구성성분 4B(치도(Chido) 혈액군), P2RY12(퓨린 수용체 P2Y, G-단백질 결합, 12), RNLS(레날라제, FAD-의존적 아민 옥시다제), CREB1(cAMP 반응성 구성요소 결합 단백질 1), POMC(프로오피오멜라노코르틴), RAC1(ras-관련 C3 보툴리눔 독소 기질 1(rho 패밀리, 소 GTP 결합 단백질 Rac1)), LMNA(라민 NC), CD59(CD59 분자, 상보체 조절 단백질), SCN5A(나트륨 채널, 전압 개폐, V형, 알파 서브유닛), CYP1B1(사이토크롬 P450, 패밀리 1, 서브패밀리 B, 폴리펩타이드 1), MIF(대식세포 이동 저해 인자(글리코실화-저해 인자)), MMP13(매트릭스 메탈로펩티다제 13(콜라게나제 3)), TIMP2(TIMP 메탈로펩티다제 저해제 2), CYP19A1(사이토크롬 P450, 패밀리 19, 서브패밀리 A, 폴리펩타이드 1), CYP21A2(사이토크롬 P450, 패밀리 21, 서브패밀리 A, 폴리펩타이드 2), PTPN22(단백질 티로신 포스파타제, 비 수용체 22형 (림포이드)), MYH14(마이오신, 중쇄 14, 비 근육), MBL2(만노스-결합 렉틴(단백질 C) 2, 가용성(옵소닌 결합)), SELPLG(셀렉틴 P 리간드), AOC3(아민 옥시다제, 구리 함유 3(혈관 부착 단백질 1)), CTSL1(카텝신 L1), PCNA(증식 세포 핵 항원),

IGF2(인슐린-유사 성장 인자 2(소마토메딘 A)), ITGB1(인테그린, 베타 1(피브로넥틴 수용체, 베타 폴리펩티드, 항원 CD29는 MDF2를 포함함, MSK12)), CAST(칼파스타틴), CXCL12(케모카인(C-X-C 모티프) 리간드 12(기질 세포-유래 인자 1)), IGHE(면역글로불린 중 불변 엡실론), KCNE1(칼륨 전압 개폐 채널, Isk-관련 패밀리, 구성원 1), TFRC(트랜스페린 수용체(p90, CD71)), COL1A1(콜라겐, I형, 알파 1), COL1A2(콜라겐, I형, 알파 2), IL2RB(인터류킨 2 수용체, 베타), PLA2G10(포스포리파제 A2, 그룹 X), ANGPT2(앤지오포이에틴 2), PROCR(단백질 C 수용체, 내피 (EPCR)), NOX4(NADPH 옥시다제 4), HAMP(헵시딘 항균성 펩타이드), PTPN11(단백질 티로신 포스파타제, 비-수용체 11형), SLC2A1(용질 운반체 패밀리 2(촉진 글루코스 수송체), 구성원 1), IL2RA(인터류킨 2 수용체, 알파), CCL5(케모카인(C-C 모티프) 리간드 5), IRF1(인터페론 조절 인자 1), CFLAR(CASP8 및 FADD-유사 세포자멸사 조절자), CALCA(칼시토닌-관련 폴리펩타이드 알파), EIF4E(진핵생물 번역 개시 인자 4E), GSTP1(글루타티온 S-트랜스퍼라제 pi 1), JAK2(야누스 키나제 2), CYP3A5(사이토크롬 P450, 패밀리 3, 서브패밀리 A, 폴리펩타이드 5), HSPG2(헤파린 설페이트 프로테오글라이칸 2), CCL3 (케모카인(C-C 모티프) 리간드 3), MYD88(골수성 분화 1차 반응 유전자(88)), VIP(혈관활성 장 펩타이드), SOAT1(스테롤 O-아실트랜스퍼라제 1), ADRBK1(아드레날린, 베타, 수용체 키나제 1), NR4A2(핵 수용체 서브패밀리 4, 그룹 A, 구성원 2), MMP8(매트릭스 메탈로펩티다제 8(호중구 콜라게나제)), NPR2(나트륨이뇨 펩타이드 수용체 B/구아닐산 사이클라제 B(아트리오나트륨이뇨 펩타이드 수용체 B)), GCH1(GTP 사이클로가수분해효소 1), EPRS(글루타밀-프롤릴-tRNA 합성효소), PPARGC1A (페록시좀 증식제-활성화 수용체 감마, 공동활성인자 1 알파), F12(응고인자 XII(하게만 인자)), PECAM1(혈소판/내피세포 부착 분자), CCL4(케모카인(C-C 모티프) 리간드 4), 세르핀A3(세르핀 펩티다제 저해제, 클레이드 A(알파-1 안티프로테이나제, 안티트립신), 구성원 3), CASR(칼슘-센싱 수용체), GJA5(갭 연접 단백질, 알파 5, 40 kDa), FABP2(지방산 결합 단백질 2, 장), TTF2(전사 종결 인자, RNA 중합효소 II), PROS1(단백질 S (알파)), CTF1(카디오트로핀 1), SGCB(사르코글리칸, 베타(43 kDa 디스트로핀-결합 글리코단백질)), YME1L1(YME1-유사 1(사카로마이세스 세레비시애)), CAMP(카텔리시딘 항균성 펩타이드), ZC3H12A(아연 핑거 CCCH-형 함유 12A), AKR1B1(알도-케토 환원효소 패밀리 1, 구성원 B1(알도스 환원효소)), DES(데스민), MMP7(매트릭스 메탈로펩티다제 7(마트리리신, 자궁)), AHR(아릴 탄화수소 수용체), CSF1(집락자극인자 1(대식세포)), HDAC9(히스톤 데아세틸라제 9), CTGF(결합조직 성장 인자), KCNMA1(칼륨 큰 전도도 칼슘-활성화 채널, 서브패밀리 M, 알파 구성원 1), UGT1A(UDP 글루쿠로노실트랜스퍼라제 1 패밀리, 폴리펩타이드 A 복합체 유전자좌), PRKCA(단백질 키나제 C, 알파), COMT(카테콜-.베타.-메틸트랜스퍼라제), S100B(S100 칼슘 결합 단백질 B), EGR1(조기 성장 반응 1), PRL(프롤락틴), IL15(인터류킨 15), DRD4(도파민 수용체 D4), CAMK2G(칼슘/카모둘린-의존 단백질 키나제 II 감마), SLC22A2(용질 운반체 패밀리 22(유기 양이온 수송체), 구성원 2), CCL11(케모카인(C-C 모티프) 리간드 11), PGF(B321 태반 성장 인자), THPO(트롬보포이에틴), GP6(글리코단백질 VI(혈소판)), TACR1(타키키닌 수용체 1), NTS(뉴로텐신), HNF1A (HNF1 호메오박스 A), SST(소마토스타틴), KCND1(칼륨 전압 개폐 채널, Shal-관련 서브패밀리, 구성원 1), LOC646627(포스포리파제 저해제), TBXAS1(트롬복산 A 신타제 1(혈소판)), CYP2J2(사이토크롬 P450, 패밀리 2, 서브패밀리 J, 폴리펩타이드 2), TBXA2R(트롬복산 A2 수용체), ADH1C(알코올 탈수소효소 1C(클래스 I), 감마 폴리펩타이드), ALOX12(아라키도네이트 12-리폭시게나제), AHSG(알파-2-HS-글리코단백질), BHMT(베타인-호모시스테인 메틸트랜스퍼라제), GJA4(갭 연접 단백질, 알파 4, 37 kDa), SLC25A4(용질 운반체 패밀리 25(미토콘드리아 운반체; 아데닌 뉴클레오티드 수송체), 구성원 4), ACLY(ATP 시트레이트 리가제), ALOX5AP(아라키도네이트 5-리폭시게나제-활성화 단백질), NUMA1(핵 유사분열 장치 단백질 1), CYP27B1(사이토크롬 P450, 패밀리 27, 서브패밀리 B, 폴리펩타이드 1), CYSLTR2(시스테이닐 류코트리엔 수용체 2), SOD3(수퍼옥시드 디스무타제 3, 세포밖), LTC4S(류코트리엔 C4 신타제), UCN(유로코틴), GHRL(그렐린/오베스타틴 프레프로펩타이드), APOC2(아포지방단백질 C-II), CLEC4A(C-형 렉틴 도메인 패밀리 4, 구성원 A), KBTBD10(켈크 반복부 및 BTB (POZ) 도메인 함유 10), TNC(테나스신 C), TYMS(티미딜레이트 합성효소), SHC1(SHC (Src 상동성 2 도메인 함유) 형질전환 단백질 1), LRP1(저밀도 지방단백질 수용체-관련 단백질 1), SOCS3(사이토카인 신호처리 3의 억제제), ADH1B(알코올 탈수소효소 1B(클래스 I), 베타 폴리펩티드), KLK3(칼리크레인-관련 펩티다제 3), HSD11B1(하이드록시스테로이드(11-베타) 탈수소효소 1), VKORC1(비타민 K 에폭시드 환원효소 복합체, 서브유닛 1), 세르핀B2(세르핀 펩티다제 저해제, 클레이드 B (오발부민), 구성원 2), TNS1(텐신 1), RNF19A(고리 핑거 단백질 19A), EPOR(에리트로포이에틴 수용체), ITGAM(인테그린, 알파 M(상보체 구성성분 3 수용체 3 서브유닛)), PITX2(쌍-유사 호메오도메인 2), MAPK7(미토겐-활성화 단백질 키나제 7), FCGR3A(IgG의 Fc 단편, 저 친화도 111a, 수용체(CD16a)), LEPR(렙틴 수용체), ENG(엔도글린), GPX1(글루타티온 페록시다제 1), GOT2(글루탐산-옥살로아세트산 트랜스아미나제 2, 미토콘드리아(아스파르테이트 아미노트랜스퍼라제 2)), HRH1(히스타민 수용체 H1), NR112(핵 수용체 서브패밀리 1, 그룹 I, 구성원 2), CRH(코르티코트로핀 방출 호르몬), HTR1A(5-하이드록시트립타민(세로토닌) 수용체 1A), VDAC1(전압-의존적 양이온 채널 1), HPSE(헤파라나제), SFTPD(계면활성제 단백질 D), TAP2(수송체 2, ATP-결합 카세트, 서브 패밀리 B (MDR/TAP)),

RNF123(고리 핑거 단백질 123), PTK2B (PTK2B 단백질 티로신 키나제 2 베타), NTRK2(신경영양 티로신 키나제, 수용체, 2형), IL6R(인터류킨 6 수용체), ACHE(아세틸콜린에스터라제(Yt 혈액 그룹)), GLP1R(글루카곤-유사 펩타이드 1 수용체), GHR(성장 호르몬 수용체), GSR(글루타티온 환원효소), NQO1(NAD(P)H 탈수소효소, 퀴논 1), NR5A1(핵 수용체 서브패밀리 5, A 그룹, 구성원 1), GJB2(갭 연접 단백질, 베타 2, 26 kDa), SLC9A1(용질 운반체 패밀리 9(나트륨/수소 교환기), 구성원 1), MAOA(모노아민 옥시다제 A), PCSK9(프로단백질 전환효소 서브틸리신/켁신 9형), FCGR2A(IgG의 Fc 단편, 저친화도 IIa, 수용체(CD32)), 세르핀F1(세르핀 펩티다제 저해제, 클레이드 F(알파-2 항플라스민, 색소 상피 유래 인자), 구성원 1), EDN3(엔도텔린 3), DHFR(디하이드로폴레이트 환원효소), GAS6(성장 저지-특이적 6), SMPD1(스핑고미엘린 포스포디에스터라제 1, 산 리파제), UCP2(미결합 단백질 2(미토콘드리아, 단백질 운반체)), TFAP2A(전사 인자 AP-2 알파(활성화 인핸서 결합 단백질 2 알파)), C4BPA(상보체 구성성분 4 결합 단백질, 알파), 세르핀F2(세르핀 펩티다제 저해제, 클레이드 F(알파-2 항플라스민, 색소 상피 유래 인자), 구성원 2), TYMP(티미딘 포스포릴라제), ALPP(알칼린 포스파타제, 태반(레간 이소자임)), CXCR2(케모카인(C-X-C 모티프) 수용체 2), SLC39A3(용질 운반체 패밀리 39(아연 수송체), 구성원 3), ABCG2(ATP-결합 카세트, 서브 패밀리 G(WHITE), 구성원 2), ADA(아데노신 데아미나제), JAK3(야누스 키나제 3), HSPA1A(열 충격 70 kDa 단백질 1A), FASN(지방산 신타제), FGF1(섬유아세포 성장 인자 1(산성)), F11(응고 인자 XI), ATP7A(ATP분해효소, Cu++ 수송, 알파 폴리펩타이드), CR1(상보체 구성성분(3b/4b) 수용체 1(놉스 혈액군)), GFAP(신경교 섬유질 산성 단백질), ROCK1(Rho-관련, 또 꼬인 나선 함유 단백질 키나제 1), MECP2(메틸 CpG 결합 단백질 2(Rett 증후군)), MYLK(미오신 경쇄 키나제), BCHE(부티릴콜린에스터라제), LIPE(리파제, 호르몬-민감성), PRDX5(페록시레독신 5), ADORA1(아데노신 A1 수용체), WRN(베르너 증후군, RecQ 헬리카제-유사), CXCR3(케모카인(C-X-C 모티프) 수용체 3), CD81 (CD81 분자), SMAD7(SMAD 패밀리 구성원 7), LAMC2(라미닌, 감마 2), MAP3K5(미토겐-활성화 단백질 키나제 키나제 키나제 5), CHGA(크로모그라닌 A(부갑상선 분비 단백질 1)), IAPP(섬 아밀로이드 폴리펩타이드), RHO(로돕신), ENPP1(엑토뉴클레오티드 피로포스파타제/포스포디에스터라제 1), PTHLH(부갑상선 호르몬-유사 호르몬), NRG1(뉴레귤린 1), VEGFC(혈관표피성장인자 C), ENPEP(글루타밀 아미노 펩티다제(아미노펩티다제 A)), CEBPB(CCAAT/인핸서 결합 단백질(C/EBP), 베타), NAGLU(N-아세틸글루코사미니다제, 알파-), F2RL3(응고인자 II(트롬빈) 수용체-유사 3), CX3CL1(케모카인(C-X3-C 모티프) 리간드 1), BDKRB1(브래디키닌 수용체 B1), ADAMTS13(트롬보스폰딘 1형 모티프를 지니는 ADAM 메탈로펩티다제, 13), ELANE(엘라스타제, 발현된 호중구), ENPP2(엑토뉴클레오티드 피로포스파타제/포스포디에스터라제 2), CISH(사이토카인 유도성 SH2-함유 단백질), GAST(가스트린), MYOC(섬유주세포, 섬유주대 그물 유도성 글루코코르티코이드 반응), ATP1A2(ATP 분해효소, Na+/K+ 수송, 알파 2 폴리펩타이드), NF1(뉴로피브로민 1), GJB1(갭 연접 단백질, 베타 1, 32 kDa), MEF2A(미오사이트 인핸서 인자 2A), VCL(빈쿨린), BMPR2(뼈 형성 단백질 수용체, II 형(세린/트레오닌 키나제)), TUBB(튜불린, 베타), CDC42(세포분화 주기 42(GTP 결합 단백질, 25 kDa)), KRT18(케라틴 18), HSF1(열 충격 전사 인자 1), MYB(v-myb 골수아페소 바이러스 발암유전자 상동체(조류)), PRKAA2(단백질 키나제, AMP-활성화된, 알파 2 촉매적 서브유닛), ROCK2(Rho-관련, 또 꼬인 나선 함유 단백질 키나제 2), TFPI(조직인자 경로 저해제(지방단백질-결합 응고 저해제)), PRKG1(단백질 키나제, cGMP-의존적, I형), BMP2(뼈 형성 단백질 2), CTNND1(카테닌(카데린-결합 단백질), 델타 1), CTH(시스타티오나제(시스타티오닌 감마-라이아제)), CTSS(카텝신 S), VAV2(vav 2 구아닌 뉴클레오티드 교환 인자), NPY2R(뉴로 펩타이드 Y 수용체 Y2), IGFBP2(인슐린-유사 성장 인자 결합 단백질 2, 36 kDa), CD28(CD28 분자), GSTA1(글루타티온 S-트랜스퍼라제 알파 1), PPIA(펩티딜프롤릴 이성질화효소 A(사이클로필린 A)), APOH(아포지방단백질 H(베타-2-당단백질 I)), S100A8(S100 칼슘 결합 단백질 A8), IL11(인터류킨 11), ALOX15(아라키도네이트 15-리폭시게나제), FBLN1(피불린 1), NR1H3(핵 수용체 서브패밀리 1, 그룹 H, 구성원 3), SCD(스테아로일-CoA 불포화효소(델타-9-불포화효소)), GIP(위 저해성 폴리펩타이드), CHGB(크로모그라닌 B(세크레토그라닌 1)), PRKCB(단백질 키나제 C, 베타), SRD5A1(스테로이드-5-알파-환원효소, 알파 폴리펩타이드 1(3-옥소-5 알파-스테로이드 델타 4-탈수소효소 알파 1)), HSD11B2(하이드록시스테로이드(11-베타) 탈수소효소 2), CALCRL(칼시토닌 수용체-유사), GALNT2(UDP-N-아세틸-알파-D-갈락토사민:폴리펩타이드 N-아세틸갈락토사미닐트랜스퍼라제 2(GalNAc-T2)), ANGPTL4(앤지오포이에틴-유사 4), KCNN4 (칼륨 중간체/소 전도도 칼슘-활성화 채널, 서브패밀리 N, 구성원 4), PIK3C2A(포스포이노시티드-3-키나제, 클래스 2, 알파 폴리펩타이드), HBEGF(헤파린-결합 EGF-유사 성장 인자), CYP7A1(사이토크롬 P450, 패밀리 7, 서브패밀리 A, 폴리펩타이드 1), HLA-DRB5(주조직적합성 복합체, 클래스 II, DR 베타 5), BNIP3 (BCL2/아데노바이러스 E1B 19 kDa 상호작용 단백질 3), GCKR (글루코키나제(헥소키나제 4) 조절제), S100A12(S100 칼슘 결합 단백질 A12), PADI4(펩티딜 아르기닌 데이미나제, IV형), HSPA14(열 충격 70 kDa 단백질 14), CXCR1(케모카인(C-X-C 모티프) 수용체 1), H19(H19, 모계 각인 발현 전사체(비 단백질 암호)), KRTAP19-3(케라틴 관련 단백질 19-3), IDDM2(인슐린-의존성 진성 당뇨병 2), RAC2(ras-관련 C3 보툴리눔 독소 기질 2(rho 패밀리, 소 GTP 결합 단백질 Rac2)), RYR1(리아노딘 수용체 1(골격)), CLOCK(클락 상동체(마

우스)), NGFR(신경 성장 인자 수용체(TNFR 수퍼패밀리, 구성원 16)), DBH(도파민 베타-하이드록실라제(도파민 베타-모노옥시게나제)), CHRNA4(콜린성 수용체, 니코틴, 알파 4), CACNA1C(칼슘 채널, 전압-의존성, L형, 알파 1C 서브유닛), PRKAG2(단백질 키나제, AMP-활성화된, 감마 2 비-촉매적 서브유닛), CHAT(콜린 아세틸트랜스퍼라제), PTGDS(프로스타글란딘 D2 신타제 21 kDa(뇌)), NR1H2(핵 수용체 서브패밀리 1, 그룹 H, 구성원 2), TEK(TEK 티로신 키나제, 내피), VEGFB(혈관표피성장인자 B), MEF2C(근세포 인핸서 인자 2C), MAPKAPK2(미토겐-활성화 단백질 키나제-활성화 단백질 키나제 2), TNFRSF11A(종양 괴사 인자 수용체 수퍼패밀리, 구성원 11a, NFKB 활성인자), HSPA9(열 충격 70 kDa 단백질 9 (모르탈린)), CYSLTR1(시스테이닐 류코트리엔 수용체 1), MAT1A(메티오닌 아데노실트랜스퍼라제 I, 알파), OPRL1(오피에이트 수용체-유사 1), IMPA1 (이노시톨(미오)-1(또는 4)-모노포스파타제 1), CLCN2(염화물 채널 2), DLD(디하이드로리포아미드 탈수소효소), PSMA6(프로테아좀(프로좀, 마크로파인) 서브유닛, 알파 유형, 6), PSMB8(프로테아좀(프로좀, 마크로파인) 서브유닛, 베타 유형, 8(거대 다기능성 펩티다제 7)), CHI3L1(키티나제 3-유사 1(연골 당단백질-39)), ALDH1B1(알데히드 탈수소효소 1 패밀리, 구성원 B1), PARP2(폴리 (ADP-리보스) 중합효소 2), STAR(스테로이도게닉 급성 조절 단백질), LBP(지질다당류 결합 단백질), ABCC6(ATP-결합 카세트, 서브 패밀리 C(CFTR/MRP), 구성원 6), RGS2(G 단백질 신호처리 2의 조절자, 24 kDa), EFNB2(에프린-B2), GJB6(갭 연접 단백질, 베타 6, 30 kDa), APOA2(아포지방단백질 A-II), AMPD1(아데노신 모노 인산염 데아미나제 1), DYSF(다이스페린, 대근이영양증 2B(상염색체 열성)), FDFT1(파르네실-이인산 파르네실트랜스퍼라제 1), EDN2(엔도텔린 2), CCR6(케모카인 (C-C 모티프) 수용체 6), GJB3(갭 연접 단백질, 베타 3, 31 kDa), IL1RL1(인터류킨 1 수용체-유사 1), ENTPD1(엑토뉴클레오사이드 삼인산염 디포스포 가수분해효소 1), BBS4(바르데트 비틀 증후군 4), CELSR2(카데린, EGF LAG 7-통과 G-형 수용체 2(플라밍고 상동체, 초파리)), F11R(F11 수용체), RAPGEF3(Rap 구아닌 뉴클레오티드 교환 인자(GEF) 3), HYAL1(히알루로노글루코사미니다제 1), ZNF259(아연 핑거 단백질 259), ATOX1(ATX1 항산화제 단백질 1 상동체(효모)), ATF6(활성화 전사 인자 6), KHK(케토헥소키나제(프럭토 키나제)), SAT1(스페르미딘/스페르민 N1-아세틸트랜스퍼라제 1), GGH(감마-글루타밀 가수분해효소(콘쥬가제, 폴릴폴리감마글루타밀 가수분해효소)), TIMP4(TIMP 메탈로펩티다제 저해제 4), SLC4A4(용질 운반체 패밀리 4, 중탄산나트륨 공동수송체, 구성원 4), PDE2A(포스포디에스터라제 2A, cGMP-자극), PDE3B (포스포디에스터라제 3B, cGMP-저해), FADS1(지방산 불포화효소 1), FADS2(지방산 불포화효소 2), TMSB4X(티모신 베타 4, X-연결), TXNIP(티오레독신 상호작용 단백질), LIMS1(LIM 및 노화 세포 항원-유사 도메인 1), RHOB(ras 상동체 유전자 패밀리, 구성원 B), LY96(림프구 항원 96), FOXO1(포크헤드 박스 O1), PNPLA2(파타틴-유사 포스포리파제 도메인 함유 2), TRH(티로트로핀-방출 호르몬), GJC1(갭 연접 단백질, 감마 1, 45 kDa), SLC17A5(용질 운반체 패밀리 17(음이온/당 수송체), 구성원 5), FTO(지방량 및 비만 관련), GJD2(갭 연접 단백질, 델타 2, 36 kDa), PSRC1(프롤린/세린-풍부 또 꼬인 나선 1), CASP12(카스파제 12(유전자/위유전자)), GPBAR1(G 단백질-결합 담즙산 수용체 1), PXK(PX 도메인 함유 세린/트레오닌 키나제), IL33(인터류킨 33), TRIB1(트리블 상동체 1(초파리)), PBX4(전-B-세포 백혈병 호메오박스 4), NUPR1(핵 단백질, 전사 조절자, 1), 15-Sep(15 kDa 셀레노단백질), CILP2(연골 중간체 층 단백질 2), TERC(텔로머라제 RNA 구성성분), GGT2(감마-글루타밀트랜스퍼라제 2), MT-CO1(미토콘드리아 암호화 사이토크롬 c 옥시다제 I), 및 UOX(요산염 옥시다제, 위유전자)를 포함할 수 있지만, 이들로 제한되지 않는다.

[0660] 추가 구현예에서, 염색체 서열은 Pon1(파라옥소나제 1), LDLR (LDL 수용체), ApoE(아포지방단백질 E), Apo B-100(아포지방단백질 B-100), ApoA (아포지방단백질(a)), ApoA1(아포지방단백질 A1), CBS(시스타티온 B-신타제), 글리코단백질 IIb/IIb, MTHRF(5,10-메틸렌테트라하이드로폴레이트 환원효소(NADPH), 및 이들의 조합으로부터 추가로 선택될 수 있다. 일 반복에서, 심혈관 질환에 연루된 염색체 서열, 및 염색체 서열에 의해 암호화된 단백질은 Cacna1C, Sod1, Pten, Ppar(알파), Apo E, 렙틴, 및 이들의 조합으로부터 선택될 수 있다.

[0661] **신장**

[0662] 본 발명은 또한 신장에 CRISPR-Cas 시스템을 전달하는 것을 고려한다. 치료적 핵산의 세포내 흡수를 유도하는 전달 전략은 물리적 힘 또는 벡터 시스템, 예컨대 바이러스-, 지질- 또는 복합체- 기반 전달, 또는 나노운반체를 포함한다. 덜 가능한 임상적 타당성을 지니는 초기 적용분야로부터, 전신에 유체역학적 고압 주사를 이용하여 핵산이 신세포로 보내질 때, 다양한 범위의 유전자 치료 바이러스 및 비-바이러스 운반체가 생체내 상이한 동물 신장 질환 모델에서 전사 후 사건을 표적화하는데 이미 적용되었다(Csaba Révész and Péter Hamar (2011). Delivery Methods to Target RNAs in the Kidney, Gene Therapy Applications, Prof. Chunsheng Kang (Ed.), ISBN: 978-953-307-541-9, InTech, Available from: http://www.intechopen.com/books/gene-therapy-applications/delivery-methods-to-target-rnas-in-the-kidney). 신장에 대한 전달 방법은 다음과 같다:

전달 방법	운반체	표적 RNA	질환	모델	기능성 분석	저자
유체역학/지질	TransIT 생체내 유전자 전달 시스템, DOTAP	p85α	급성 신손상	허혈-재관류	흡수, 생물학적 분류	Larson et al., Surgery, (Aug 2007), Vol. 142, No. 2, pp. (262-269)
유체역학/지질	리포펙타민 2000	Fas	급성 신손상	허혈-재관류	요소 질소, Fas 면역조직화학, 세포자멸사, 조직학적 스코어링	Hamar et al., Proc Natl Acad Sci, (Oct 2004), Vol. 101, No. 41, pp. (14883-14888)
유체역학	해당없음	세포자멸사 캐스케이드 구성요소	급성 신손상	허혈-재관류	해당없음	Zheng et al., Am J Pathol, (Oct 2008), Vol. 173, No. 4, pp. (973-980)
유체역학	해당없음	핵 인자 카파-b (NFkB)	급성 신손상	허혈-재관류	해당없음	Feng et al., Transplantation, (May 2009), Vol. 87, No. 9, pp. (1283-1289)
유체역학/바이러스	리포펙타민 2000	세포자멸사 길항 전사 인자 (AATF)	급성 신손상	허혈-재관류	세포자멸사, 산화적 스트레스, 카스파제 활성, 막지질 과산화반응	Xie & Guo, Am Soc Nephrol, (Dec 2006), Vol. 17, No. 12, pp. (3336-3346)

[0663]

전달 방법	운반체	표적 RNA	질환	모델	기능성 분석	저자
유체역학	pBAsi mU6 Neo/ TransIT-EE 유체역학 전달 시스템	그렘린	당뇨병성 신장 질환	스트렙토조토진 - 유발 당뇨병	단백뇨, 혈청 크레아티닌, 사구체 및 관 모양 직경, 콜라겐 IV형/BMP7 발현	Q. Zhang et al., PloS ONE, (Jul 2010), Vol. 5, No. 7, e11709, pp. (1-13)
바이러스/지질	pSUPER 벡터/리포펙타민	TGF-β II형 수용체	사이질 신장 섬유증	편측 요도폐쇄	α-SMA 발현, 콜라겐 함량,	Kushibikia et al., J Controlled Release, (Jul 2005), Vol. 105, No. 3, pp. (318-331)
바이러스	아데노-연관 바이러스 -2	무기염 코르티코이드 수용체	고혈압 야기 신장 손상	한랭 고혈압	혈압, 혈청 알부민, 혈청 요소질소, 혈청 크레아티닌, 신장 중량, 소변 나트륨	Wang et al., Gene Therapy, (Jul 2006), Vol. 13, No. 14, pp. (1097-1103)
유체역학/바이러스	pU6 벡터	루시퍼라제	해당없음	해당없음	흡수	Kobayashi et al., Journal of Pharmacology and Experimental Therapeutics, (Feb 2004), Vol. 308, No. 2, pp. (688-693)

[0664]

전달 방법	운반체	표적 RNA	질환	모델	기능성 분석	저자
지질	지방단백질 s, 알부민	apoB1, apoM	해당없음	해당없음	흡수, 지방단백질 및 알부민에 대한 결합 친화도	Wolfrum et al., Nature Biotechnology, (Sep 2007), Vol. 25, No. 10, pp. (1149-1157)
지질	리포펙타민 2000	p53	급성 신손상	허혈성 및 시스플라틴 - 유발 급성 손상	조직학적 스코어링, 세포자멸사	Molitoris et al., J Am Soc Nephrol, (Aug 2009), Vol. 20, No. 8, pp. (1754-1764)
지질	DOTAP/DOPE, DOTAP/DOPE/DOPE-PEG2000	COX-2	유방 아데노-암종	MDA-MB-231 유방암 이종이식 - 보유 마우스	세포생존, 흡수	Mikhaylova et al., Cancer Gene Therapy, (Mar 2011), Vol. 18, No. 3, pp. (217-226)
지질	콜레스테롤	12/15-리폭시게나제	당뇨병 신장병증	스트렙토조신 - 유발 당뇨병	알부민뇨, 소변 크레아티닌, 조직학, I형 및 IV형 콜라겐, TGF-β, 피브로넥틴, 플라스미노겐 활성인자 저해제 1	Yuan et al., Am J Physiol Renal Physiol, (Jun 2008), Vol. 295, pp. (F605-F617)

[0665]

전달 방법	운반체	표적 RNA	질환	모델	기능성 분석	저자
지질	리포펙타민 2000	미토콘드리아 막 44 (TIM44)	당뇨병 신장병증	스트렙토조신 - 유발 당뇨병	세포 증식 및 세포자멸사, 조직학, ROS, Mn-SOD의 미토콘드리아 유입 및 글루타티온 페록시다제, 세포막 분극	Y. Zhang et al., J Am Soc Nephrol, (Apr 2006), Vol. 17, No. 4, pp. (1090-1101)
유체역학/지질	프로테오리포솜	RLIP76	신장암종	Caki-2 신장암 이종이식 - 보유 마우스	흡수	Singhal et al., Cancer Res, (May 2009), Vol. 69, No. 10, pp. (4244-4251)
중합체	폐길화된 PEI	루시퍼라제 pGL3	해당없음	해당없음	흡수, 생물학적 분류, 적혈구 집괴	Malek et al., Toxicology and Applied Pharmacology, (Apr 2009), Vol. 236, No. 1, pp. (97-108)
중합체	폐길화된 폴리-L-리신	MAPK1	낭창 사구체신염	사구체 신염	단백뇨, 사구체경화증, TGF-β, 피브로넥틴, 플라스미노겐 활성인자 저해제 1	Shimizu et al., J Am Soc Nephrology, (Apr 2010), Vol. 21, No. 4, pp. (622-633)

[0666]

전달 방법	운반체	표적 RNA	질환	모델	기능성 분석	저자
중합체/Na 입자 없음	히알루론산/ 양자점/ PEI	VEGF	신장염/ 흑색종	B16F1 흑색종 종양-보유 마우스	생물학적 분류, 세포독성, 종양 용적, 내포작용	Jiang et al., Molecular Pharmaceutics, (May-Jun 2009), Vol. 6, No. 3, pp. (727-737)
중합체/Na 입자 없음	페길화된 폴리카프로락톤 나노섬유	GAPDH	해당없음	해당없음	세포 생존도, 흡수	Cao et al, J Controlled Release, (Jun 2010), Vol. 144, No. 2, pp. (203-212)
앱타머	스피에겔머 mNOX-E36	CC 케모카인 리간드 2	사구체 경화증	신장절제 마우스	소변 알부민, 소변 크레아티닌, 조직병리학, 사구체 여과율, 대식세포수, 혈청 Ccl2, Mac- 2+, Ki-67+	Ninichuk et al., Am J Pathol, (Mar 2008), Vol. 172, No. 3, pp. (628-637)
앱타머	어댑터 NOX-F37	바소프레신 (AVP)	울혈성 심부전	해당없음	D-AVP에 대한 결합 친화도, AVP 신호처리의 저해, 소변 삼투질 농도 및 나트륨 농도,	Purschke et al., Proc Natl Acad Sci, (Mar 2006), Vol. 103, No. 13, pp. (5173-5178)

[0667]

[0668] Yuan 등(Am J Physiol Renal Physiol 295: F605-F617, 2008)은 아라키돈산 대사의 12/15-리폭시게나제 (12/15-LO) 경로를 표적화하는 짧은 간섭 RNA(siRNA)의 생체내 전달이 1형 당뇨병의 스트렙토조신 주사 마우스 모델에서 신손상 및 당뇨병성 신장 질환(DN)을 개선시킬 수 있는지 여부를 조사하였다. 생체내 평가 및 신장에서의 siRNA 발현을 더 크게 달성하기 위해, Yuan 등은 콜레스테롤에 컨쥬게이트된 이중 가닥 12/15-LO siRNA 올리고뉴클레오티드를 사용하였다. 약 400 ㎍의 siRNA를 마우스에 피하로 주사하였다. Yuang 등의 방법은 신장에 전달을 위해 인간에게 콜레스테롤과 컨쥬게이트된 CRISPR Cas의 1 내지 2 g 피하 주사를 고려하여 본 발명의 CRISPR Cas 시스템에 적용될 수 있다.

[0669] Molitoris 등(J Am Soc Nephrol 20: 1754-1764, 2009)은 신장손상을 방지하기 위해 세포자멸사 경로로 중심 단백질인 p53에 표적화된 siRNA의 효능을 시험하도록 신장 내 올리고뉴클레오티드 재흡수 부위로서 기부세뇨관 세포(proximal tubule cell: PTC)를 이용하였다. 허혈성 손상이 PTC와 신장 기능을 둘 다 최대로 보호하고 4시간 후에 p53에 대한 네이키드 합성 siRNA를 정맥내로 주사하였다. Molitoris 등의 데이터는 기부 세뇨관 세포에 siRNA의 빠른 전달이 정맥내 투여를 따른다는 것을 나타낸다. 용량 반응 분석을 위해, 래트에 동일한 4 회 시점에 제공한 siP53, 0.33; 1, 3 또는 5mg/kg의 용량을 주사하여, 각각 1.32; 4, 12 및 20 mg/kg의 누적 용량을 야기하였다. 시험한 모든 siRNA 용량은 제1 일에 SCr 감소 효과를 생성하였으며, PBS-처리 허혈성 대조군 래트에 비해 대략 5 일에 걸쳐 용량이 커질수록 효과적으로 되었다. 12 및 20 mg/kg 누적 용량은 최고의 보호 효과를 제공하였다. Molitoris 등의 방법은 신장에 전달을 위해 인간에 대해 12 및 20 mg/kg 누적 용량을 고려하여 본 발명의 CRISPR Cas 시스템에 적용될 수 있다.

[0670] Thompson 등(Nucleic Acid Therapeutics, Volume 22, Number 4, 2012)은 설치류 및 비인간 영장류에서 정맥내 투여 후 합성의 짧은 간섭 RNA I5NP의 독성학적 및 약동학적 특성을 보호한다. I5NP는 세포자멸사 전 단백질 p53의 발현을 일시적으로 저해하기 위해 RNA 간섭(RNAi) 경로를 통해 작용하도록 설계되고, 주요 심장 수술동안 생길 수 있는 급성 신장 손상 및 신장 이식 후 생길 수 있는 지연된 이식편 기능과 같은 급성 허혈/재관류 손상

으로부터 세포를 보호하기 위해 개발 중에 있다. 설치류에서 800mg/kg I5NP, 및 비인간 영장류에서 1,000 mg/kg I5NP의 용량은 이상 반응을 유발하는데 필요로 되었는데, 원숭이에서 상보체의 무증상 활성화 및 약간 증가된 응고 시간을 포함하는 혈액에 대한 효과를 지시하도록 단리되었다. 래트에서, I5NP의 래트 유사체에 의해 추가적인 이상 반응은 관찰되지 않았는데, 이는 효과가 I5NP의 의도된 약학적 활성과 관련된 독성보다 합성 RNA 듀플렉스의 계열효과를 나타낼 가능성이 있다는 것을 나타낸다. 합쳐서 생각하면, 이들 데이터는 급성 허혈/재관류 손상 후 신기능의 보존을 위해 I5NP를 정맥내 투여하는 임상 시험을 뒷받침한다. 원숭이에서 이상 반응이 관찰되지 않은 수준(no observed adverse effect level: NOAEL)은 500 mg/kg이었다. 25 mg/kg까지 용량 수준에서 i.v. 투여 후 원숭이에서 심혈관, 호흡 및 신경학적 파라미터에 대한 효과는 관찰되지 않았다. 따라서, 인간의 신장에 CRISPR Cas의 정맥내 투여를 위해 유사한 투약량이 고려될 수 있다.

[0671] Shimizu 등(J Am Soc Nephrol 21: 622-633, 2010)은 폴리(에틸렌 글리콜)-폴리(L-리신)-기반 비히클을 통해 사구체에 siRNA의 전달을 표적화하는 시스템을 개발하였다. siRNA/나노담체 복합체는 메산지움에 접근하기 위해 그것이 창문내피(fenestrated endothelium)를 가로질러 이동하게 하는 크기인 대략 10 내지 20 nm의 직경이었다. 형광-표지된 siRNA/나노담체 복합체의 복강내 주사후에, Shimizu 등은 연장된 시간 동안 혈액 순환에서 siRNA를 검출하였다. 미토겐-활성화 단백질 키나제 1(MAPK1) siRNA/나노담체 복합체의 반복된 복강내 투여는 사구체신염의 마우스 모델에서 사구체 MAPK1 mRNA 및 단백질 발현을 억제하였다. siRNA 축적의 조사를 위해, PIC 나노담체(0.5 mℓ, 5 nmol의 siRNA 함량), 네이키드 Cy5-표지 siRNA(0.5 mℓ, 5 nmol), 또는 HVJ-E(0.5 mℓ, 5 nmol의 siRNA 함량) 중에서 캡슐화된 Cy5-표지 siRNA 와 복합체화된 Cy5-표지 siRNA가 BALB-c 마우스에 투여되었다. Shimizu 등의 방법은 복강내 투여 및 신장에 대한 전달을 위해 인간에 약 1 내지 2 리터 중의 나노담체와 복합체화된 약 10 내지 20 μmol CRISPR Cas의 용량을 고려하는 본 발명의 CRISPR Cas 시스템에 적용될 수 있다.

[0672] **폐**

[0673] 본 발명은 또한 한쪽 또는 양쪽 폐에 CRISPR-Cas 시스템을 전달하는 것을 고려한다.

[0674] AAV-2-기반 벡터는 본래 CF 기도에 CFTR 전달을 위해 제안되었지만, 다른 혈청형, 예컨대 AAV-1, AAV-5, AAV-6 및 AAV-9는 다양한 폐 상피 모델에서 개선된 유전자 전달 효율을 나타낸다(예를 들어, 문헌[Li et al., Molecular Therapy, vol. 17 no. 12, 2067-2077 Dec 2009] 참조). AAV-1이 AAV-5와 동일한 효율로 생체내에서 기관 기도 상피에 도입되었지만, AAV-1은 시험관내 인간 기도 상피 세포 도입에 있어서 AAV-2 및 AAV-5보다 약 100배 더 효율적인 것으로 입증되었다. 5 다른 연구는 AAV-5가 시험관내 인간 기도 상피에 대한 유전자 전달에서 AAV-2보다 50-배 더 효율적이며, 생체내 마우스 폐 기도 상피에서 상당히 더 효율적이라는 것을 나타내었다. AAV-6는 또한 시험관내 인간 기도 상피 세포 및 생체내 뮤린 기도에서 AAV-2보다 더 효율적인 것으로 나타났다.8 더 최근의 단리물인 AAV-9는 9 개월에 걸쳐 검출된 유전자 발현을 지니는 생체내 뮤린 비강 및 치조 상피에서 AAV-5보다 더 큰 유전자 전달 효율을 나타내는 것으로 나타났는데, 이는 AAV가 CFTR 유전자 전달 벡터에 대해 바람직한 특성인 생체내 장기간 유전자 발현을 가능하게 할 수 있다는 것을 시사한다. 더 나아가, AAV-9는 CFTR 발현 및 최소 면역 결과의 손실 없이 뮤린 폐에 재투여될 수 있다는 것이 입증되었다. CF 및 비-CF HAE 배양물은 몇 시간 동안 100 ㎕의 AAV 벡터를 이용하여 선단면 상에 접종될 수 있다(예를 들어, 문헌[Li et al., Molecular 요법, vol. 17 no. 12, 2067-2077 Dec 2009] 참조). MOI는 바이러스 농도 및 실험 목적에 따라서 1 × 10^3 내지 4 × 10^5 개 벡터 게놈/세포로 다를 수 있다. 상기 인용된 벡터는 본 발명의 전달 및/또는 투여를 위해 고려된다.

[0675] Zamora 등(Am J Respir Crit Care Med Vol 183. pp 531-538, 2011)은 인간 전염성 질환에 치료적인 RNA 간섭 적용의 예 및 또한 호흡기 융합 바이러스(RSV)-감염 폐 이식 수용자에서 항바이러스 약물의 무작위 시험을 보고하였다. Zamora 등은 RSV 호흡관 감염을 지니는 LTX 수용자에서 무작위, 이중맹검, 위약통제 시험을 수행하였다. 환자에게 RSV에 대한 표준 치료를 받게 하였다. 에어로졸화된 ALN-RSV01(0.6 mg/kg) 또는 위약은 3 일 동안 매일 투여되었다. 이 연구는 RSV를 치료적 표적으로 하는 RNAi가 RSV 감염을 지니는 LTX 수용자에게 안전하게 투여될 수 있다는 것을 입증한다. ALN-RSV01의 3 일의 매일 용량은 호흡관 증상의 임의의 악화 또는 폐 기능 장애를 초래하지 않았고, 임의의 전신 전염증 효과, 예컨대 사이토카인 또는 CRP의 유도를 나타내지 않았다. 약동학은 흡입 후 낮은 일시적 전신 노출만을 나타내었고, 정맥내 또는 흡입에 의해 투여된 ALN-RSV01이 엑소뉴클레아제 매개 절단 및 신장 배출을 통해 순환으로부터 빠르게 클리어런스된다는 것을 나타내는 전임상 동물 데이터와 일치된다. Zamora 등의 방법은 본 발명의 CRISPR Cas 시스템에 적용될 수 있고, 에어로졸화된 CRISPR Cas는, 예를 들어 0.6 mg/kg의 투약량으로 본 발명에 대해 고려될 수 있다.

[0676] CFTR델타508 키메라 가이드 RNA의 예에 대해, 아데노-연관 바이러스(AAV)를 사용하여, 낭포성 섬유증 또는 낭포성 섬유증(CF) 관련 증상으로 고통받는 치료가 필요한 대상체 또는 환자의 기도에서 CRISPR-Cas 시스템의 유전자 이동 또는 유전자 전달을 입증하는 실시예 22를 참조한다. 특히, 그들은 낭포성 섬유증 델타 F508 돌연변이에 대한 수선 전략을 예시한다. 이 유형의 전략은 모든 유기체에 걸쳐 적용되어야 한다. 특히 CF와 관련하여, 적합한 환자는 인간, 비영장류 인간, 개, 고양이, 소, 말 및 다른 가축 동물을 포함할 수 있다. 이 예에서, 본 발명자들은 델타F508 또는 다른 CFTR-유발 돌연변이를 표적화하기 위해 Cas9 효소를 포함하는 CRISPR-Cas 시스템을 이용하였다.

[0677] 이 예에서 치료되는 대상체는 폐에 대해 기관지로 전달되는 한편 자발적으로 호흡하여 약제학적 유효량의 에어로졸 AAV 벡터 시스템을 받는다. 이와 같이, 에어로졸화된 전달은 일반적으로 AAV 전달에 바람직하다. 아데노바이러스 또는 AAV 입자가 전달을 위해 사용될 수 있다. 하나 이상의 조절 서열에 각각 작동 가능하게 연결된 적합한 유전자 작제물은 전달 벡터 내로 클로닝될 수 있다. 이 예에서, 다음의 작제물은 예로서 제공된다: Cas9에 대해 Cbh 또는 EF1a 프로모터, 키메라 가이드 RNA에 대해 U6 또는 H1 프로모터),: 바람직한 배열은 선택적으로 하나 이상의 핵 국소화 신호 또는 서열(들)(NLS(들)), 예를 들어, 두 개의(2) NLS를 이용하는 델타 F508 돌연변이 및 코돈 최적화된 Cas9 효소(바람직한 Cas9는 뉴클레아제 또는 닉카아제 활성을 지니는 것임)에 대한 수선 주형인 CFTR델타508 표적화 키메라 가이드를 사용하는 것이다. NLS가 없는 작제물이 또한 생각된다.

[0678] Cas9 표적 부위를 동정하기 위해, 본 발명자는 인간 CFTR 게놈 유전자좌를 분석하였고, Cas9 표적 부위를 동정하였다. 바람직하게는, 일반적으로 및 이 CF 경우에, PAM은 NGG 또는 NNAGAAW 모티프를 함유할 수 있다.

[0679] 따라서, CF의 경우에, 본 방법은 하기 단계를 포함하는 관심 대상의 게놈 유전자좌에서의 표적 서열의 조작을 포함한다:

[0680] 조성물을 작동 가능하게 암호화하는 하나 이상의 바이러스 벡터를 포함하는 바이러스 벡터 시스템을 포함하는 비자연적으로 생기는 또는 유전자 조작된 조성물을 이의 발현을 위해 전달하는 단계로서, 상기 조성물은:

[0681] I. CRISPR-Cas 시스템 키메라 RNA(chiRNA) 폴리뉴클레오티드 서열에 작동 가능하게 연결된 제1 조절 구성요소로서, 폴리뉴클레오티드 서열은

[0682] (a) 적합한 포유류 세포에서 CF 표적 서열에 혼성화될 수 있는 가이드 서열,

[0683] (b) tracr 메이트 서열, 및

[0684] (c) tracr 서열을 포함하는, 제1 조절 구성요소, 및

[0685] II. 적어도 하나 이상의 핵 국소화 서열을 포함하는 CRISPR 효소를 암호화하는 효소 암호화 서열에 작동 가능하게 연결된 제2 조절 구성요소,

[0686] 를 포함하는 하나 이상의 벡터를 포함하는 벡터 시스템을 포함하는 비자연적으로 생기거나 또는 유전자 조작된 조성물을 포함하되,

[0687] (a), (b) 및 (c)는 5'에서 3' 방향으로 배열되며,

[0688] 구성성분 I 및 II는 시스템의 동일 또는 상이한 벡터 상에 위치되고,

[0689] 전사될 때, tracr 메이트 서열은 tracr 서열에 혼성화하며, 가이드 서열은 표적 서열에 대한 CRISPR 복합체의 서열-특이적 결합을 지시하고,

[0690] CRISPR 복합체는 (1) 표적 서열에 혼성화되는 가이드 서열 및 (2) tracr 서열에 혼성화된 tracr 메이트 서열에 복합체화되는 CRISPR 효소를 포함함)을 포함한다. CF에 대해, 바람직한 표적 DNA 서열은 CFTR델타508 돌연변이를 포함한다. 바람직한 PAM은 상기 기재되어 있다. 바람직한 CRISPR 효소는 임의의 Cas(본 명세서에 기재되어 있지만, 특히 실시예 22에 기재되어 있다.

[0691] CF에 대한 대안은 임의의 유전적 장애를 포함하며, 이들의 예는 잘 알려져 있다. 본 발명의 다른 바람직한 방법 또는 용도는 라포라병과 관련된 것으로 동정된 EMP2A 및 EMP2B 유전자에서의 결함을 보정하기 위한 것이다.

[0692] 일부 구현예에서, "가이드 서열"은 "가이드 RNA"과 별개일 수 있다. 가이드 서열은 표적 부위를 구체화하는 가이드 RNA 내의 대략 20 bp 서열을 지칭할 수 있다.

[0693] 일부 구현예에서, Cas9는 SpCas9이다(또는 이로부터 유래된다). 이러한 구현예에서, 바람직한 돌연변이는

SpCas9의 임의의 또는 모든 또는 위치 10, 762, 840, 854, 863 및/또는 986에 또는 다른 Cas9에서 대응하는 위치에 있다(예를 들어, 표준 서열 비교 도구에 의해 확인될 수 있다). 특히, SpCas9에서 임의의 또는 모든 다음의 돌연변이: D10A, E762A, H840A, N854A, N863A 및/또는 D986A가 바람직할 뿐만 아니라; 임의의 대체 아미노산에 대한 보존적 치환이 또한 생각된다. 다른 Cas9에서 대응하는 위치에서 동일한 돌연변이(또는 이들 돌연변이의 보존적 치환)이 또한 바람직하다. SpCas9에서 D10 및 H840이 특히 바람직하다. 그러나, 다른 Cas9에서, SpCas9 D10 및 H840에 대응하는 잔기가 또한 바람직하다. 이들은 그들이 닉카아제 활성을 제공하기 때문에 유리하다. 이러한 돌연변이는 CF의 치료분만 아니라 본 발명의 모든 양태에 적용될 수 있다.

[0694] Schwank 등(Cell Stem Cell, 13:653-58, 2013)은 인간 줄기 세포에서 낭포성 섬유증과 관련된 결함을 보정하기 위해 CRISPR/Cas9를 사용하였다. 상기 팀의 표적은 낭포성 섬유증 막통과 전도 수용체(CFTR)인 이온 채널에 대한 유전자였다. CFTR에서의 결함은 낭포성 섬유증 환자에서 단백질이 미스폴딩되게 한다. 낭포성 섬유증을 지니는 2명의 소아로부터의 세포 샘플로부터 발생시킨 배양한 장 줄기 세포를 사용하여, Schwank 등은 삽입되는 수선 서열을 함유하는 공여체와 함께 CRISPR을 사용하는 결함을 보정할 수 있었다. 이어서, 연구자들은 장 "유사기관(organoid)" 또는 미니어처 소화관 내로 세포를 성장시켰고, 그들이 정상적으로 기능한다는 것을 나타내었다. 이 경우에, 클론 유사기관의 약 절반은 적절한 유전자 보정을 겪었다.

[0695] **근육**

[0696] 본 발명은 또한 CRISPR-Cas 시스템을 근육(들)에 전달하는 것을 고찰한다.

[0697] Bortolanza 등(Molecular Therapy vol. 19 no. 11, 2055-2064 Nov. 2011)은 FRG1 마우스에서 RNA 간섭 발현 카세트의 전신 전달이, 안면견갑상완근이영양 (FSHD)의 개시 후, 독성 징후 없이 용량-의존적 장기적 FRG1 녹다운을 초래했음을 나타낸다. Bortolanza et al.은 5 × 10^{12} vg의 rAAV6-sh1FRG1의 단일 정맥내 주사가 FRG1 마우스의 근육 조직병리학 및 근육 기능을 회복시킨다는 것을 발견했다. 상세히 말하면, 생리학적 용액에 2 × 10^{12} 또는 5 × 10^{12} vg의 벡터를 함유하는 200㎕가 25-게이지 테루모 주사기를 사용하여 꼬리 정맥에 주사되었다. Bortolanza 등의 방법은 AAV 발현 CRISPR Cas에 적용될 수 있으며, 약 2 × 10^{15} 또는 2 × 10^{16} vg의 벡터의 용량으로 인간에게 주사될 수 있다.

[0698] Dumonceaux 등(Molecular Therapy vol. 18 no. 5, 881-887 May 2010)은 미오스타틴 수용체 AcvRIIb mRNA (sh-AcvRIIb)에 대해 지정된 RNA 간섭 기술을 사용하여 미오스타틴 경로를 억제한다. 유사-디스트로핀의 복구가 벡터화된 U7 엑손-스킵 기술 (U7-DYS)에 의해서 매개되었다. sh-AcvrIIb 작제물만, U7-DYS 작제물만, 또는 두 작제물의 조합을 지닌 아데노-관련 벡터가 이영양성 mdx 마우스의 전경골근 (TA) 근육에 주사되었다. 이 주사는 10^{11} AAV 바이러스 게놈을 사용하여 수행되었다. Dumonceaux 등의 방법은 AAV 발현 CRISPR Cas에 적용될 수 있으며, 약 10^{14} 또는 약 10^{15} vg의 벡터의 용량으로 인간에게 주사될 수 있다.

[0699] Kinouchi 등(Gene Therapy (2008) 15, 1126-1130)은 아텔로콜라겐 (ATCOL)과 함께 화학적으로 미변형된 siRNA의 나노입자 형성을 통한 정상 또는 질환 마우스의 골격근에 생체내 siRNA 전달의 유효성을 보고한다. 골격근 성장의 음성 조절제인 siRNA 표적화 미오스타틴의 마우스 골격근 또는 정맥내 ATCOL-매개 국소 적용은 적용 후 수주 이내에 근육량의 현저한 증가를 야기했다. 이들 결과는 siRNA의 ATCOL-매개 적용이 근위축증을 포함한 질환에 대한 향후의 치료적 사용을 위한 강력한 도구임을 내포한다. Mst-siRNAs (최종 농도, 10mM)가 제조자의 지시에 따라서 ATCOL(국소 투여를 위한 최종 농도, 0.5%)(AteloGene, Kohken, Tokyo, Japan)와 혼합되었다. 넴부탈(25mg/kg, i.p.)로 마우스(20-주령 수컷 C57BL/6)를 마취한 후, Mst-siRNA/ATCOL 복합체가 교근 및 대퇴이두근 근육에 주사되었다. Kinouchi et al의 방법은 CRISPR Cas에 적용될 수 있으며, 예를 들어 근육에 40 μM 용액 약 500 내지 1000 ㎖의 용량으로 인간에게 주사될 수 있다.

[0700] Hagstrom 등(Molecular Therapy Vol. 10, No. 2, August 2004)은 포유류의 팔다리 근육을 통해서 근유세포(근섬유)에 핵산의 효과적이며 반복가능한 전달을 가능하게 하는 혈관내 비-바이러스 방법을 설명한다. 이 과정은 지혈대 또는 혈압계 밴드에 의해서 일시적으로 고립된 팔의 말단 정맥에 네이키드 플라스미드 DNA 또는 siRNA를 주사하는 것을 포함한다. 근섬유에 핵산 전달은 근육 조직에서 핵산 용액의 분출을 가능하게 하기에 충분한 부피의 신속한 주사에 의해서 촉진된다. 골격근에서 높은 수준의 트랜스젠 발현이 최소 독성으로 소형 동물과 대형 동물에서 달성되었다. 팔다리 근육에 siRNA 전달의 증거도 또한 얻어졌다. 레서스 원숭이에 플라스미드 DNA 정맥내 주사를 위해서, 3방향 스톱콕이 2 개의 주사기 펌프(Model PHD 2000; Harvard Instruments)에 연결되었으며, 각각은 단일 주사기로 로딩되었다. 파파베린 주사 후 5 분 뒤에 pDNA (40-100㎖ 식염수 중 15.5 내지

25.7 mg)가 1.7 또는 2.0 ml/s의 속도로 주사되었다. 이것은 인간에 대해서 800 내지 2000 ml 식염수 중에 약 300 내지 500 mg의 주사로서 본 발명의 CRISPR Cas 발현 플라스미드 DNA에 대해 스케일업될 수 있다. 래트에 아데노바이러스 벡터 주사를 위해서, 2 x 10^9 감염성 입자가 3 ml의 정상 식염수 용액 (NSS) 중에서 주사되었다. 이것은 본 발명의 CRISPR Cas 발현 아데노바이러스 벡터에 대해 스케일업될 수 있으며, 인간에 대해서 10 리터의 NSS 중에 약 1 x 10^{13} 감염성 입자의 주사가 주사되었다. siRNA를 위해서, 래트에는 12.5 μg의 siRNA가 대복재정맥에 주사되었고, 영장류에는 750 μg의 siRNA가 대복재정맥에 주사되었다. 이것은 본 발명의 CRISPR Cas에 대해 스케일업될 수 있으며, 예를 들어, 인간의 대복재정맥에 약 15 내지 약 50 mg이 주사되었다.

[0701] **피부**

[0702] 본 발명은 또한 CRISPR-Cas 시스템을 피부에 전달하는 것을 고찰한다.

[0703] Hickerson 등(Molecular Therapy-Nucleic Acids (2013) 2, e129)은 인간 및 뮤린 피부에 자체-전달 (sd)-siRNA를 전달하기 위한 동력화된 마이크로니들 어레이 피부 전달 장치에 관한 것이다. siRNA-기반 피부 치료제를 병원용으로 전환하기 위한 주요 난제는 효과적인 전달 시스템의 개발이다. 다양한 피부 전달 기술에서 실질적인 노력이 있어왔고 제한적인 성공이 있었다. siRNA로 피부를 치료한 임상 연구에서, 피하주사 바늘 주사와 관련된 강렬한 통증이 이 시험에 추가의 환자의 등록을 막았으며, 이것은 개선된 더 "환자-친화적"(즉, 통증이 거의 없거나 없는) 전달 접근법에 대한 필요성을 강조한다. 마이크로니들은 1차 장벽인 각질층을 가로질러 siRNA를 포함하는 큰 하전된 물체를 전달하기 위한 유효한 방식이며, 일반적으로 종래의 피하주사 바늘보다 통증이 적다고 간주된다. Hickerson 등에 의해서 사용된 동력화된 마이크로니들 어레이 (MMNA)를 포함하는 동력화된 "스탬프 타입" 마이크로니들 장치는 무모 마우스 연구에서 안전한 것으로 나타났으며, (i) 화장품 산업에서의 광범한 사용 및 (ii) 거의 모든 지원자가 장치의 사용이 독감예방주사보다 훨씬 통증이 적다고 한 제한적 시험에 의해서 증명된 대로 통증을 거의 또는 전혀 야기하지 않는데, 이것은 이 장치를 사용한 siRNA 전달이 피하주사 바늘 주사를 사용한 이전의 임상 시험에서 경험된 것보다 훨씬 적은 통증을 가져올 것임을 시사한다. MMNA 장치 (Triple-M 또는 Tri-M으로 Bomtech Electronic Co, Seoul, South Korea에 의해서 시판됨)는 마우스 및 인간 피부에 siRNA의 전달을 위해서 개조되었다. sd-siRNA 용액 (0.1 mg/ml RNA 최대 300 μl)이 0.1 mm의 깊이로 설정된 일회용 Tri-M 니들 카트리지(Bomtech)의 챔버에 도입되었다. 인간 피부의 치료를 위해서, 치료 전에 식별되지 않은 피부 (수술 과정 직후 얻어진)를 손으로 잡아늘려 코르크 플랫폼으로 고정되었다. 모든 피내 주사는 28-게이지 0.5-인치 니들을 가진 인슐린 주사기를 사용하여 수행되었다. Hickerson 등의 MMNA 장치 및 방법이 사용될 수 있었고 및/또는 본 발명의 CRISPR Cas의 전달을 위해서, 예를 들어 피부에 0.1 mg/ml CRISPR Cas의 최대 300 μl의 용량의 전달을 위해서 개조될 수 있었다.

[0704] Leachman 등(Molecular Therapy, vol. 18 no. 2, 442-446 Feb. 2010)은 장애를 유발하는 발바닥 가피증을 포함한 상염색체 우성 증후군인 희귀한 피부 장애인 선천성 경조증 (PC)의 치료를 위한 Ib 상 임상 시험에 관한 것으로서, 피부를 위한 제1의 짧은 간섭 RNA(siRNA)-기반 치료제를 이용한다. 이 siRNA는 TD101로 불리는데 야생형 K6a mRNA에 영향 없이 케라틴 6a (K6a) N171K 돌연변이 mRNA를 특이적으로 강력하게 표적화한다. 용량-상승 스케줄을 이하에 제시한다:

[0705]

주	용량 번호	일 수	용적(ml)	TD101의 농도 (mg/ml)	총 용량 TD101 (mg)
1	1-2	1-7	0.1	1.0	0.10
2	3-4	8-14	0.25	1.0	0.25
3	5-6	15-21	0.50	1.0	0.50
4	7-8	22-28	1.0	1.0	1.0
5	9-10	29-35	1.5	1.0	1.5
6	11-12	36-42	2.0	1.0	2.0
7	13-14	43-49	2.0	1.5	3.0
8	15-16	50-56	2.0	2.0	4.0
9	17-18	57-63	2.0	2.5	5.0
10	19-20	64-70	2.0	3.0	6.0
11	21-22	71-77	2.0	3.5	7.0
12	23-24	78-84	2.0	4.0	8.0
13	25-26	85-91	2.0	4.5	9.0
14	27-28	92-98	2.0	5.0	10.0
15	29-30	99-105	2.0	6.0	12.0
16	31-32	106-112	2.0	7.0	14.0
17	33	113-119	2.0	8.5	17.0

[0706] 초기에, TD101 1.0 mg/㎖ 용액 0.1 ㎖ 또는 비히클(칼슘 또는 마그네슘 없는 둘베코 포스페이트-완충 식염수)만이 대칭적 굳은살에 투여되었다. 주사 당 TD101 용액을 0.1, 0.25, 0.5, 1.0, 1.5, 및 2.0 ㎖의 1.0 mg/㎖ 용액으로 증가시킬 때 부반응 없이 6 번의 씻어내는 투여가 완료되었다. 최고 설정된 부피 (2.0 ㎖)가 잘 용인되었으므로 이어서 TD101의 농도가 8.5 mg/㎖의 최종 농도까지 매주 1 mg/㎖씩 증가되었다. 케라틴 6a(K6a) N171K 돌연변이 mRNA를 특이적으로 강력하게 표적화하는 CRISPR Cas의 투여에 대해서도 유사한 투약량이 고찰된다.

[0707] Zheng 등(PNAS, July 24, 2012, vol. 109, no. 30, 11975-11980)은 금 코어가 잘 방향화된 공유 고정된 siRNA의 치밀한 외피로 둘러싸인 것인, 구형 핵산 나노입자 콘쥬게이트(SNA-NC)가 적용 후 수시간 내에 시험관내 마우스 피부, 및 인간 표피에서 각질세포에 거의 100% 자유롭게 침투한다는 것을 나타낸다. Zheng 등은 60 h 동안 25 nM 표피성장인자 수용체 (EGFR) SNA-NC의 단일 적용이 인간 피부에서 효과적인 유전자 녹아웃을 나타낸다는 것을 증명했다. 유사한 투약량이 피부에 투여하기 위해 SNA-NC에 고정된 CRISPR Cas에 대해서도 고찰될 수 있다.

[0708] **간염 바이러스**

[0709] 본 발명은 또한 B형 간염 바이러스(HBV)에도 적용될 수 있다. 그러나, CRISPR Cas 시스템은, 예를 들어 용량 및 서열을 최적화함으로써, 내인성 소 RNA 경로의 과포화의 위험과 같은, RNAi의 단점을 피할 수 있도록 개조되어야 한다(예를 들어, 문헌[Grimm et al., Nature vol. 441, 26 May 2006] 참조). 예를 들어, 인간에 대해서 약 1-10 x 10^{14} 입자와 같은 적은 용량이 고찰된다.

[0710] 다른 구현예에서, HBV에 대해 지정된 CRISPR Cas 시스템은 안정한 핵산-지질 입자(SNALP)와 같은 리포좀 중에서 투여될 수 있다(예를 들어, 문헌[Morrissey et al., Nature Biotechnology, Vol. 23, No. 8, August 2005] 참조). SNALP 중의 HBV RNA에 대해 표적화된 CRISPR Cas의 약 1, 3 또는 5 mg/kg/일의 일간 정맥내 주사가 고찰된다. 일간 치료는 약 3 일에 걸쳐 수행될 수 있고, 이어서 약 5 주 동안 주간 치료된다.

[0711] 다른 구현예에서, Chen 등(Gene Therapy (2007) 14, 11-19)의 시스템이 본 발명의 CRISPR Cas 시스템에 대해 사용될 수 있고/있거나 개조될 수 있다. Chen 등은 shRNA를 전달하기 위해서 이중-가닥 아데노연관 바이러스8-수도타입 벡터(dsAAV2/8)를 사용한다. HBV-특이적 shRNA를 지닌 dsAAV2/8 벡터의 단일 투여 (마우스 당 1 x 10^{12} 벡터 게놈)는 HBV 유전자이식 마우스의 간에서 HBV 단백질, mRNA 및 복제 DNA의 일정한 수준을 효과적으로 억제했으며, 순환계에서 HBV 로드의 최대 2-3 $\log_{10}$ 감소를 초래했다. 유의한 HBV 억제가 벡터 투여 후 적어도 120 일 동안 지속되었다. shRNA의 치료 효과는 표적 서열 의존적이었고, 인터페론의 활성화를 수반하지 않았다. 본 발명에 있어서, HBV에 대해 지정된 CRISPR Cas 시스템이 dsAAV2/8 벡터와 같은 AAV 벡터에 클로닝될 수 있으며, 예를 들어 인간 당 약 1 x 10^{15} 벡터 게놈 내지 약 1 x 10^{16} 벡터 게놈의 용량으로 인간에게 투여될 수 있다.

[0712] 다른 구현예에서, Wooddell 등(Molecular Therapy vol. 21 no. 5, 973-985 May 2013)의 방법이 본 발명의 CRISPR Cas 시스템에 대해 사용될 수 있고/있거나 개조될 수 있다. Woodell 등은 간세포-표적화된, N-아세틸갈락토사민-콘쥬게이트 멜리틴-유사 펩티드 (NAG-MLP)와 응고인자 VII (F7)을 표적화하는 간-향성 콜레스테롤-콘쥬게이트 siRNA (chol-siRNA) 의 간단한 동시주사가 임상 화학의 변화나 사이토카인의 유도 없이 마우스 및 비-인간 영장류에서 효과적인 F7 녹다운을 가져온다는 것을 나타낸다. HBV 감염의 일시적 유전자이식 마우스 모델을 사용하여 Wooddell 등은 NAG-MLP와 보존된 HBV 서열을 표적화하는 강력한 chol-siRNA의 단일 동시주사가 긴 지속기간으로 바이러스 RNA, 단백질 및 바이러스 DNA의 다중로그 억제를 가져왔음을 나타낸다. 예를 들어, 약 6 mg/kg의 NAG-MLP와 6 mg/kg의 HBV-특이적 CRISPR Cas의 정맥내 동시주사가 본 발명에 대해 고안될 수 있다. 대안으로서, 약 3 mg/kg의 NAG-MLP와 3 mg/kg의 HBV-특이적 CRISPR Cas가 어느날 한번 전달될 수 있고, 이어서 이후 2 주는 약 2-3 mg/kg의 NAG-MLP와 2-3 mg/kg의 HBV-특이적 CRISPR Cas가 투여된다.

[0713] 본 발명은 또한 C형 간염 바이러스(HCV)를 치료하기 위해서 적용될 수도 있다. Roelvinki 등(Molecular Therapy vol. 20 no. 9, 1737-1749 Sep 2012)의 방법이 CRISPR Cas 시스템에 적용될 수 있다. 예를 들어, AAV8과 같은 AAV 벡터가 고찰된 벡터일 수 있으며, 예를 들어 킬로그램 체중 당 약 1.25 × 10^{11} 내지 1.25 × 10^{13} 벡터 게놈 (vg/kg)의 용량이 고찰될 수 있다.

[0714] 또 다른 구현예에서, CRISPR-Cas9-매개 게놈 편집은 질병 돌연변이 및/또는 표현형을 교정하기 위해 이용될 수 있다. 상기 CRISPR-Cas9-매개 게놈 편집은 전체내용이 참조로서 본원에 포함되는, 웹사이트 nature.com/doifinder/10.1038/nbt.2884에서 이용가능한 2014년 3월 30일에 네이처 바이오테크놀로지(Nature

Biotechnology)에 온라인 공개되고 2014년 3월 31일에 온라인 정정된 하오 인 등(Hao Yin *et al.*)의 "Genome editing with Cas9 in adult mice corrects a disease mutation and phenotype"을 표제로 하는 매뉴스크립트에 예시된 바와 같이 간 및 간세포에서 질병 돌연변이 및/또는 표현형을 교정하기 위해 이용될 수 있다. 상기 논문은 인간 질병 유전성 티로신혈증의 마우스 모델의 간세포에서 Fah 돌연변이의 CRISPR-Cas9-매개 교정에 관한 것이다. 유체역학적 주사에 의한 CRISPR-Cas9 시스템의 구성요소의 전달은 약 1/250 간 세포에서 야생형 Fah 단백질의 최초 발현을 초래한 것으로 밝혀졌다. Fah-양성 간세포의 확장이 체중 감소 표현형을 구제한 것이 추가로 밝혀졌다.

[0715] 다른 질환의 숙주도 유사한 방식으로 치료될 수 있다는 것이 자명할 것이다. 돌연변이에 의해서 야기되는 유전 질환의 일부 예들이 여기 제공되며, 더 많은 것들이 알려져 있다. 상기 전략은 이들 질환에 적용될 수 있다.

[0716] **헌팅턴병 (HD)**

[0717] RNA 간섭 (RNAi)은 헌팅턴병의 질환-야기 유전자인 HTT의 발현을 감소시킴으로써 이 장애에 대한 치료적 잠재성을 제공하며(예를 들어, 문헌[McBride et al., Molecular Therapy vol. 19 no. 12 Dec. 2011, pp. 2152-2162]), 따라서 출원인은 그것이 CRISPR-Cas 시스템에 대해 사용될 수 있고/있거나 개조될 수 있다는 것을 가정한다. CRISPR-Cas 시스템은 안티센스 서열의 표적외 잠재성을 감소시키기 위한 알고리즘을 사용하여 생성될 수 있다. CRISPR-Cas 서열은 마우스, 레서스 또는 인간 헝틴틴의 엑손 52에 있는 서열을 표적화할 수 있으며, AAV와 같은 바이러스 벡터에서 발현될 수 있다. 인간을 포함한 동물에는 뇌반구 당 약 3 회의 미소주사 (총 6 회 주사)가 주사될 수 있는데, 첫 번째는 전방접합부에서 1 mm 문측(12 ㎕), 두 번의 나머지 주사(각각 12 ㎕ 및 10 ㎕)는 첫 번째 주사에서 3 및 6 mm 미측으로 이격하여 놓고, 1e12 vg/㎖의 AAV가 약 1 ㎕/분의 속도로 사용되며, 니들은 주사액이 니들 팁으로부터 확산할 수 있도록 추가로 5 분 동안 제자리에 유지되었다.

[0718] DiFiglia 등(PNAS, October 23, 2007, vol. 104, no. 43, 17204-17209)은 성체 선조체에 siRNA 표적화 Htt의 단일 투여가 돌연변이 Htt를 침묵화하고, 신경성 병증을 완화하며, HD의 신속-개시 바이러스 유전자이식 마우스 모델에서 관찰된 이상 행동 표현형을 지연시킬 수 있다는 것을 관찰했다. DiFiglia 주사된 마우스에 2 ㎕의 Cy3-표지된 cc-siRNA-Htt 또는 10μM의 비-콘쥬게이트 siRNA-Htt가 선조체내 주사되었다. Htt에 대해 표적화된 CRISPR Cas의 유사한 용량이 본 발명에서 인간에 대해 고찰될 수 있으며, 예를 들어 약 5-10 ㎖의 10 μM Htt 표적화된 CRISPR Cas가 선조체내 주사될 수 있다.

[0719] 다른 예에서, Boudreau 등(Molecular Therapy vol. 17 no. 6 june 2009)은 선조체에 htt-특이적 RNAi 바이러스를 발현하는 재조합 AAV 혈청형 2/1 벡터 5 ㎕를 주사한다(4 x 10^{12} 바이러스 게놈/㎖). Htt에 대해 표적화된 CRISPR Cas의 유사한 용량이 본 발명에서 인간에 대해 고찰될 수 있으며, 예를 들어 약 10-20 ㎖의 Htt에 대해 표적화된 CRISPR Cas (4 x 10^{12} 바이러스 게놈/㎖)가 선조체내 주사될 수 있다.

[0720] 다른 예에서, HTT에 대해 표적화된 CRISPR Cas는 연속 투여될 수 있다(예를 들어, 문헌[Yu et al., Cell 150, 895-908, August 31, 2012] 참조). Yu 등은 0.25㎖/hr을 전달하는 삼투 펌프(Model 2004)를 이용하여 28 일 동안 ss-siRNA 또는 포스페이트-완충 식염수(PBS) (Sigma Aldrich)를 300 ㎎/일로 전달하며, 0.5 ㎕/hr를 전달하도록 설계된 펌프(Model 2002)를 사용해서 14 일 동안 양성 대조군 MOE ASO를 75 ㎎/일로 전달했다. 펌프(Durect Corporation)는 멸균 PBS에 희석된 ss-siRNA 또는 MOE로 채워졌고, 이어서 이식 전에 24 시간 또는 48 시간 (Model 2004) 동안 37℃에서 인큐베이션되었다. 마우스가 2.5% 이소플루오란으로 마취되었고, 두개골 밑에 정중선 절개부가 만들어졌다. 입체정위 가이드를 사용하여 캐뉼라가 우측 뇌실에 이식되었고 록타이트 접착제로 고정되었다. 알젯 삼투 미니 펌프에 부착된 카테터가 캐뉼라에 부착되었고, 펌프는 정중견갑부 영역에 피하 위치되었다. 5.0 나일론 봉합실로 절개부가 봉합되었다. Htt에 대해 표적화된 CRISPR Cas의 유사한 용량이 본 발명에서 인간에 대해 고찰될 수 있으며, 예를 들어 약 500 내지 1000 g/일의 Htt 표적화된CRISPR Cas가 투여될 수 있다.

[0721] 연속 주입의 다른 예에서, Stiles 등(Experimental Neurology 233 (2012) 463-471)은 우측 피곡에 티타늄 니들 팁을 가진 뇌실질내 카테터를 이식했다. 카테터는 복부에 피하 이식된 SynchroMed® II 펌프(Medtronic Neurological, Minneapolis, MN)에 연결되었다. 6 ㎕/일로 포스페이트-완충 식염수를 7 일 주입한 후, 펌프가 시험 물질로 다시 채워졌고, 7 일 연속 전달을 위해 프로그램되었다. 약 2.3 내지 11.52 ㎎/d의 siRNA가 약 0.1 내지 0.5 ㎕/분의 다양한 주입 속도로 주입되었다. Htt에 대해 표적화된 CRISPR Cas의 유사한 용량이 본 발명에서 인간에 대해 고찰될 수 있으며, 예를 들어 약 20 내지 200 ㎎/일의 Htt 표적화된CRISPR Cas가 투여될 수 있다.

[0722] 다른 예에서, Sangamo에 양도된 미국 특허 공개 제20130253040의 방법이 또한 헌팅턴병의 치료를 위해서 TALES로부터 CRISPR Cas 시스템에 맞게 개조될 수 있다.

[0723] **핵산, 아미노산 및 단백질**

[0724] 본 발명은 표적 DNA 서열과 결합할 수 있는 핵산을 사용한다. 이것은 핵산이 단백질보다 훨씬더 용이하고 저렴하게 생산될 수 있고, 상동성이 추구되는 신장 길이에 따라 특이성이 변화될 수 있기 때문에 유익하다. 다중 핑거의 복합체 3-D 배치는 예를 들어 필요하지 않다.

[0725] 용어 "폴리뉴클레오티드", "뉴클레오티드", "뉴클레오티드 서열", "핵산" 및 "올리고뉴클레오티드"는 상호교환가능하게 사용된다. 그것들은 임의의 길이의 뉴클레오티드, 데옥시리보뉴클레오티드 또는 리보뉴클레오티드 중 어느 하나, 또는 그의 유사체의 중합체 형태를 말한다. 폴리뉴클레오티드는 임의의 3차원 구조를 가질 수 있으며, 기지의 또는 미지의 임의의 기능을 수행할 수 있다. 다음은 폴리뉴클레오티드의 비제한적인 예이다: 유전자 또는 유전자 단편의 코딩 또는 비-코딩 영역, 연관 분석으로부터 정의된 유전자좌들(유전자좌), 엑손, 인트론, 전령 RNA(mRNA), 운반 RNA, 리보솜 RNA, 짧은 간섭 RNA(siRNA), 짧은 헤어핀 RNA(shRNA), 마이크로-RNA(miRNA), 리보자임, cDNA, 재조합 폴리뉴클레오티드, 분지형 폴리뉴클레오티드, 플라스미드, 벡터, 임의의 서열의 분리된 DNA, 임의의 서열의 분리된 RNA, 핵산 프로브 및 프라이머. 이 용어는 또한 합성 백본을 가진 핵산-유사 구조를 포괄한다(WO 97/03211호; 및 WO 96/39154호). 폴리뉴클레오티드는 하나 이상의 변형된 뉴클레오티드, 예를 들어, 메틸화 뉴클레오티드 및 뉴클레오티드 유사체를 포함할 수 있다. 뉴클레오티드 구조에 대한 변형이 존재한다면, 중합체의 조립 전에 또는 후에 부여될 수 있다. 뉴클레오티드의 서열은 비-뉴클레오티드 성분에 의해 단속될 수 있다. 폴리뉴클레오티드는 중합화 후에, 예를 들어, 표지화 성분과의 컨쥬게이션에 의해 추가로 변형될 수 있다.

[0726] 본원에 사용되는 바와 같이, 용어 "야생형"은 당업자에 의해 이해되는 해당 분야의 용어이며, 그것이 돌연변이체 또는 변이체로부터 구별되는 정도로 천연에서 발생하는 것과 같은 통상적인 형태의 유기체, 균주, 유전자 또는 특징을 의미한다.

[0727] 본원에 사용되는 바와 같이, 용어 "변이체"는 천연에서 발생하는 것에서 벗어난 패턴을 갖는 특성의 표현을 의미하는 것으로 이해해야 한다.

[0728] 용어 "비-천연 발생" 또는 "조작된"은 상호교환가능하게 사용되며, 인간의 손의 개입을 나타낸다. 상기 용어는 핵산 분자 또는 폴리펩티드에 대하여 언급되는 경우, 핵산 분자 또는 폴리펩티드에 천연에서 천연적으로 관련되어 있고, 천연에서 관찰되는 적어도 하나의 다른 성분이 적어도 실질적으로 없음을 의미한다.

[0729] "상보성"은 통상의 왓슨-크릭(Watson-Crick) 염기 쌍형성 또는 기타 비-통상적 유형에 의해 다른 핵산 서열과 수소 결합(들)을 형성하는 핵산의 능력을 지칭한다. 상보성 백분율은 제2 핵산 서열과 수소 결합(예를 들어, 왓슨-크릭 염기 쌍형성)을 형성할 수 있는 핵산 분자 내의 잔기의 백분율을 나타낸다(예를 들어, 10 개 중 5, 6, 7, 8, 9, 10 개는 50%, 60%, 70%, 80%, 90% 및 100% 상보성임). "완전한 상보성"은 핵산 서열의 모든 연속 잔기가 동일한 수의 제2 핵산 서열 내의 연속 잔기와 수소 결합할 것임을 의미한다. 본원에 사용되는 바와 같이, "실질적인 상보성"은 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 35, 40, 45, 50 개 이상의 뉴클레오티드의 영역에 걸쳐 적어도 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99% 또는 100%인 상보성 정도를 지칭하거나, 엄격한 조건 하에서 혼성화하는 2개의 핵산을 지칭한다.

[0730] 본원에 사용되는 바와 같이, 혼성화를 위한 "엄격한 조건"은 표적 서열에 대하여 상보성을 갖는 핵산 서열이 대개 표적 서열과 혼성화하며, 비-표적 서열에는 실질적으로 혼성화하지 않는 조건을 지칭한다. 엄격한 조건은 일반적으로 서열-의존적이며, 다수의 요인에 따라 달라진다. 일반적으로, 서열이 길수록, 서열이 그의 표적 서열에 특이적으로 혼성화하는 온도가 더 높아진다. 엄격한 조건의 비제한적인 예는 문헌[Tijssen (1993), Laboratory Techniques In Biochemistry And Molecular Biology-Hybridization With Nucleic Acid Probes Part I, Second Chapter "Overview of principles of hybridization and the strategy of nucleic acid probe assay", Elsevier, N.Y.]에 상세히 기재되어 있다. 폴리뉴클레오티드 서열이 언급될 때, 상보성 또는 부분 상보성 서열도 생각된다. 이들은 바람직하게 고 긴축 조건에서 기준 서열과 혼성화할 수 있다. 일반적으로, 혼성화 비율을 최대화하기 위해서, 상대적으로 저-긴축 혼성화 조건이 선택된다: 열 용융점 (T_m)보다 약 20 내지 25℃ 낮은 온도. T_m은 특정 표적 서열의 50%가 규정된 이온 강도와 pH에서 용액 중에서 완전히 상보성인 프로브와 혼성화하는 온도이다. 일반적으로, 혼성화된 서열의 적어도 약 85% 뉴클레오티드 상보성을 요구하기 위해서, 고

긴축 세척 조건은 T_m보다 약 5 내지 15℃ 낮은 온도가 되도록 선택된다. 혼성화된 서열의 적어도 약 70% 뉴클레오티드 상보성을 요구하기 위해서, 중도-긴축 세척 조건은 T_m보다 약 15 내지 30℃ 낮은 온도가 되도록 선택된다. 아주 관대한 (초저 긴축) 세척 조건은 T_m보다 50℃ 정도 낮은 온도일 수 있으며, 이것은 혼성화된 서열들 간에 높은 수준의 미스매칭을 허용한다. 당업자는 표적 서열과 프로브 서열 간 상동성의 특정 수준으로부터 검출가능한 혼성화 신호의 결과에 영향을 미치기 위해서 혼성화 및 세척 단계에서 다른 물리화학적 변수들도 변경될 수 있다는 것을 인정할 것이다. 바람직한 고 긴축 조건은 42℃에서 50% 포름아미드, 5×SSC, 및 1% SDS 인큐베이션, 또는 65℃에서 5×SSC 및 1% SDS 인큐베이션과 65℃에서 0.2×SSC 및 0.1% SDS 세척을 포함한다.

[0731] "혼성화"는 하나 이상의 폴리뉴클레오티드가 반응하여, 복합체를 형성하고, 이 복합체는 뉴클레오티드 잔기의 염기 사이의 수소 결합을 통해 안정화되는 반응을 지칭한다. 수소 결합은 왓슨 크릭 염기 쌍형성, 후그스타인(Hoogstein) 결합 또는 임의의 다른 서열 특이적 방식으로 발생할 수 있다. 복합체는 듀플렉스 구조를 형성하는 2 개의 가닥, 다중 가닥 복합체를 형성하는 3 개 이상의 가닥, 단일의 자가 혼성화 가닥 또는 이들의 임의의 조합을 포함할 수 있다. 혼성화 반응은 PCR의 개시 또는 효소에 의한 폴리뉴클레오티드의 절단과 같은 보다 광범위한 과정에서 하나의 단계를 이룰 수 있다. 주어진 서열과 혼성화할 수 있는 서열은 주어진 서열의 "상보물"로 지칭된다.

[0732] 본원에서 사용된 용어 "게놈 유전자좌" 또는 "유전자좌"(복수 유전자좌)는 염색체에서 유전자 또는 DNA 서열의 특정 위치이다. "유전자"는 유기체에서 활동하는 기능적 역할을 가진 폴리펩티드 또는 RNA 사슬을 암호화하는 DNA 또는 RNA의 신축체로서, 살아있는 유기체에서 유전의 분자 단위를 말한다. 본 발명의 목적을 위해서, 유전자는 이러한 조절 서열이 코딩 서열 및/또는 전사된 서열에 인접하든 아니든 유전자 산물의 생성을 조절하는 영역을 포함하는 것으로 간주될 수 있다. 따라서, 반드시 제한되는 것은 아니지만, 유전자는 프로모터 서열, 터미네이터, 리보솜 결합 부위 및 내부 리보솜 진입 부위와 같은 번역 조절 서열, 인핸서, 사일런서, 인슐레이터, 경계 요소, 복제 기원, 바탕질 부착 부위 및 유전자좌 제어 영역을 포함한다.

[0733] 본원에서 사용된 "게놈 유전자좌의 발현" 또는 "유전자 발현"은 유전자로부터의 정보가 기능적 유전자 산물의 합성에서 사용되는 과정이다. 유전자 발현의 산물은 주로 단백질이지만, rRNA 유전자 또는 tRNA 유전자와 같은 비-단백질 코딩 유전자에서는 이 산물은 기능적 RNA이다. 유전자 발현 과정은 생존을 위한 기능적 산물들을 생성하기 위해서 모든 알려진 생명체, 즉 진핵생물 (다세포 유기체 포함), 원핵생물 (박테리아 및 고세균) 및 바이러스에 의해서 사용된다. 본원에서 사용된 유전자 또는 핵산의 "발현"은 세포 유전자 발현뿐만 아니라 클로닝 시스템 및 어떤 다른 맥락에서 핵산(들)의 전사 및 번역을 포괄한다. 본원에서 사용된 "발현"은 또한 폴리뉴클레오티드가 DNA 주형으로부터 전사되는 과정 (예컨대 mRNA 또는 다른 RNA 전사체로) 및/또는 전사된 mRNA가 이어서 펩티드, 폴리펩티드, 또는 단백질로 번역되는 과정을 말한다. 전사체 및 암호화된 폴리펩티드는 집합적으로 "유전자 산물"이라고 언급될 수 있다. 폴리뉴클레오티드가 게놈 DNA로부터 유래된다면, 발현은 진핵세포에서 mRNA의 스플라이싱을 포함할 수 있다.

[0734] 용어 "폴리펩티드", "펩티드" 및 "단백질"은 본원에서 임의의 길이의 아미노산의 중합체를 지칭하기 위해 상호교환가능하게 사용된다. 중합체는 선형 또는 분지형일 수 있으며, 그것은 변형된 아미노산을 포함할 수 있고, 그것은 비-아미노산에 의해 단속될 수 있다. 또한, 상기 용어는 변형된 아미노산 중합체, 예를 들어, 이황화 결합 형성, 글리코실화, 지질화(lipidation), 아세틸화, 인산화 또는 임의의 기타 조작, 예를 들어, 표지화 성분과의 컨쥬게이션을 포함한다. 본원에 사용되는 바와 같이, 용어 "아미노산"은 글리신 및 D 또는 L 광학 이성질체 및 아미노산 유사체 및 펩티도미메틱을 포함하는 천연 및/또는 비천연 또는 합성 아미노산을 포함한다.

[0735] 본원에서 사용된 용어 "도메인" 또는 "단백질 도메인"은 단백질 사슬의 나머지 부분과 독립적으로 존재하며 기능할 수 있는 단백질 서열의 일부분을 말한다.

[0736] 본 발명의 양태들에서 설명된 대로, 서열 동일성은 서열 상동성과 관련된다. 상동성 비교는 눈으로, 또는 더 일반적으로는 쉽게 이용가능한 서열 비교 프로그램의 도움하에 수행될 수 있다. 이들 상업적으로 이용가능한 컴퓨터 프로그램은 둘 이상의 서열 사이의 상동성 퍼센트(%)를 계산할 수 있고, 또한 둘 이상의 아미노산 또는 핵산 서열에 의해서 공유된 서열 동일성을 계산할 수 있다. 일부 바람직한 구현예에서, 여기 설명된 dTALE의 봉쇄 영역은 여기 제공된 봉쇄 영역 아미노산 서열과 적어도 95% 동일하거나 동일성을 공유하는 서열을 가진다.

[0737] 서열 상동성은 본 분야에 알려진 다수의 컴퓨터 프로그램, 예를 들어 BLAST 또는 FASTA 등의 어느 것에 의해서 생성될 수 있다. 이러한 정렬을 수행하기 위한 적합한 컴퓨터 프로그램은 GCG Wisconsin Bestfit 패키지이다 (University of Wisconsin, U.S.A; 문헌[Devereux et al., 1984, Nucleic Acids Research 12:387]). 서열 비

교를 수행할 수 있는 다른 소프트웨어의 예들은, 제한은 아니지만, BLAST 패키지(Ausubel et al., 1999 상기 참조 - Chapter 18 참조), FASTA(Atschul et al., 1990, J. Mol. Biol., 403-410) 및 GENEWORKS 비교 도구 일습을 포함한다. BLAST와 FASTA는 모두 오프라인과 온라인 검색으로 이용할 수 있다(Ausubel et al., 1999 ibid, pages 7-58 내지 7-60 참조). 그러나, GCG Bestfit 프로그램을 사용하는 것이 바람직하다.

[0738] 서열 상동성 퍼센트(%)는 연속 서열에 걸쳐서 계산될 수 있는데, 즉 하나의 서열이 다른 서열과 정렬되고, 하나의 서열에서 각 아미노산 또는 뉴클레오티드가 다른 서열에서 상응하는 아미노산 또는 뉴클레오티드와 한번에 하나의 잔기씩 직접 비교된다. 이것은 "비갭방식"(ungapped) 정렬이라고 불린다. 통상적으로, 이러한 비갭방식 정렬은 상대적으로 짧은 수의 잔기에 걸쳐서만 수행된다.

[0739] 이것은 매우 간단하고 일관된 방법이지만, 예를 들어 다른 동일한 쌍의 서열에서, 하나의 삽입 또는 결실이 이후의 아미노산 잔기를 정렬되지 않도록 할 수 있다는 것을 고려하지 않으므로, 전반적 정렬이 수행될 때는 상동성 %에 잠재적으로 상당한 감소가 있게 된다. 결론적으로, 대부분의 서열 비교 방법은 전체 상동성 또는 동일성 점수에 지나친 벌점부과 없이 가능한 삽입과 결실을 고려하여 최적 정렬을 생성하도록 설계된다. 이것은 서열 정렬에 "갭"을 삽입함으로써 달성되며, 이로써 국소 상동성 또는 동일성을 최대화하도록 시도할 수 있다.

[0740] 그러나, 이런 더 복잡한 방법은 정렬시 생긴 각 갭에 "갭 패널티"를 배정하며, 이로써 동일한 수의 동일한 아미노산에 대해, 가능한 적은 갭을 가진 서열 정렬 - 비교된 두 서열의 더 높은 관련성 반영 - 이 많은 갭을 가진 것보다 더 높은 점수를 달성할 수 있다. "친화성 갭 코스트"가 통상적으로 사용되는데, 이것은 갭의 존재에는 상대적으로 높은 코스트를 매기고, 갭의 각 연속 잔기에는 더 적은 패널티를 매긴다. 이것은 가장 흔히 사용되는 갭 점수배정 시스템이다. 높은 갭 패널티는 당연히 갭이 적을수록 최적화된 정렬을 생성할 수 있다. 대부분의 정렬 프로그램은 갭 패널티의 변형을 허용한다. 그러나, 서열 비교에 이러한 소프트웨어를 사용할 때는 디폴트 값을 사용하는 것이 바람직하다. 예를 들어, GCG Wisconsin Bestfit 패키지를 사용할 때, 아미노산 서열에 대한 디폴트 갭 패널티는 갭 당 -12, 각 연장 당 -4이다.

[0741] 따라서, 최대 상동성 %의 계산은 갭 패널티를 고려한 최적 정렬의 생성을 요구한다. 이러한 정렬을 수행하기 위한 적합한 컴퓨터 프로그램은 GCG Wisconsin Bestfit 패키지이다(Devereux et al., 1984 *Nuc. Acids Research* 12 p387). 서열 비교를 수행할 수 있는 다른 소프트웨어의 예들은, 제한은 아니지만, BLAST 패키지(Ausubel et al., 1999 *Short Protocols in Molecular Biology*, 4th Ed. Chapter 18 참조), FASTA (Altschul et al., 1990 *J. Mol. Biol.* 403-410) 및 GENEWORKS 비교 도구 일습을 포함한다. BLAST와 FASTA는 모두 오프라인과 온라인 검색으로 이용할 수 있다(Ausubel et al., 1999, *Short Protocols in Molecular Biology*, pages 7-58 내지 7-60 참조). 그러나, 일부 용도에서, GCG Bestfit 프로그램을 사용하는 것이 바람직하다. BLAST 2 서열이라고 불리는 새로운 도구도 단백질 및 뉴클레오티드를 비교하는데 이용할 수 있다(*FEMS Microbiol Lett.* 1999 174(2): 247-50; *FEMS Microbiol Lett.* 1999 177(1): 187-8 및 National Institutes for Health 웹사이트에서 National Center for Biotechnology 웹사이트의 정보 참조).

[0742] 최종 상동성%는 동일성 항목으로 측정될 수 있지만, 통상적으로 정렬 과정 자체는 모두-아니면-전부 쌍 비교에 기초하지 않는다. 대신, 축척된 유사성 점수 행렬이 일반적으로 사용되는데, 이것은 화학적 유사성 또는 진화적 거리에 기초하여 각 쌍-방식 비교에 점수를 배정한다. 흔히 사용되는 이러한 행렬의 예는 BLOSUM62 행렬로서, 이것은 BLAST 프로그램 일습의 디폴트 행렬이다. GCG Wisconsin 프로그램은 일반적으로 공개된 디폴트 값이나 제공되는 경우에는 사용자 지정 기호비교표 (추가의 상세한 설명을 위한 사용자 매뉴얼 참조)를 사용한다. 일부 용도에서, GCG 패키지에 대한 공개된 디폴트 값을 사용하거나, 또는 다른 소프트웨어의 경우에는 BLOSUM62와 같은 디폴트 행렬을 사용하는 것이 바람직하다.

[0743] 또는 달리, 상동성 퍼센트는 CLUSTAL와 유사한 알고리즘에 기초한, DNASIS™(Hitachi Software)의 다중 정렬 특징을 사용하여 계산될 수 있다(Higgins DG & Sharp PM (1988), *Gene* 73(1), 237-244). 일단 소프트웨어가 최적 정렬을 생성하면, 상동성 %, 바람직하게는 서열 동일성 %를 계산하는 것이 가능하다. 이 소프트웨어는 통상적으로 서열 비교의 일부로서 이것을 행하고, 수치 결과를 생성한다.

[0744] 서열은 또한 침묵성 변화를 야기하여 기능적으로 동등한 물질을 가져오는 아미노산 잔기의 결실, 삽입 또는 치환을 가질 수 있다. 아미노산 특성(예컨대 잔기의 극성, 전하, 용해성, 소수성, 친수성 및/또는 양친매성)의 유사성을 기초로 의도적인 아미노산 치환이 만들어질 수 있으며, 따라서 이것은 아미노산을 기능적 그룹으로 함께 그룹화하는데 유용하다. 아미노산은 그것의 측쇄만의 특성에 기초하여 그룹화될 수 있다. 그러나, 돌연변이 데이터를 포함하는 것도 역시 더 유용하다. 이렇게 유도된 아미노산 세트는 구조적 이유로 보존될 수 있다. 이들

세트는 벤다이어그램의 형태로 설명될 수 있다(Livingstone C.D. and Barton G.J. (1993) "Protein sequence alignments: a strategy for the hierarchical analysis of residue conservation" *Comput. Appl. Biosci.* 9: 745-756) (Taylor W.R. (1986) "The classification of amino acid conservation" *J. Theor. Biol.* 119; 205-218). 보존적 치환은, 예를 들어 일반적으로 용인된 벤다이어그램 아미노산 그룹화를 설명한 아래 표에 따라서 만들어질 수 있다.

[0745]

세트		하위세트	
소수성	F W Y H K M I L V A G C	방향족	F W Y H
		지방족	I L V
극성	W Y H K R E D C S T N Q	하전	H K R E D
		양으로 하전	H K R
		음으로 하전	E D
소형	V C A G S P T N D	최소형	A G S

[0746] 본 발명의 구현예는, 아미노산의 경우 유사한 것으로의 치환, 예컨대 염기성을 염기성으로, 산성을 산성으로, 극성을 극성으로 등등의 치환을 일으킬 수 있는 상동성 치환(치환 및 대체는 모두 본원에서 기존 아미노산 잔기 또는 뉴클레오티드의 다른 잔기 또는 뉴클레오티드로의 교환을 의미하기 위해서 사용된다)을 포함할 수 있는 서열(폴리뉴클레오티드 또는 폴리펩티드 모두)을 포함한다. 비-상동성 치환, 즉 한 부류의 잔기로부터 다른 부류의 잔기로의 치환이 일어날 수 있거나, 또는 달리 이것은 오르니틴(이후 Z로 언급), 디아미노부티르산 오르니틴(이후 B로 언급), 노르류신 오르니틴(이후 O로 언급), 피리일알라닌, 티엔일알라닌, 나프틸알라닌 및 페닐글리신과 같은 비천연 아미노산의 포함을 의 잔기로의 치환을 수반할 수 있다.

[0747] 변이체 아미노산 서열은 글리신 또는 β-알라닌 잔기와 같은 아미노산 스페이서에 더하여 메틸, 에틸 또는 프로필 기와 같은 알킬 기를 포함하는 서열의 임의의 두 아미노산 잔기 사이에 삽입될 수 있는 적합한 스페이서 기를 포함할 수 있다. 펩토이드 형태로 하나 이상의 아미노산 잔기의 존재를 포함하는 추가의 형태의 변형도 당업자에게 잘 이해될 수 있다. 의심을 피하기 위해서, "펩토이드 형태"는 α-탄소 치환체 기가 α-탄소가 아니라 잔기의 질소 원자에 존재하는 변이체 아미노산 잔기를 말하기 위해서 사용된다. 펩토이드 형태로 펩티드를 제조하는 과정은 본 분야에, 예를 들어 문헌[Simon RJ et al., *PNAS* (1992) 89(20), 9367-9371 및 Horwell DC, *Trends Biotechnol.* (1995) 13(4), 132-134]에 공지되어 있다.

[0748] 본 발명의 실시는 달리 나타내지 않는 한, 당업계의 기술 내에 있는 면역학, 생화학, 화학, 분자 생물학, 미생물학, 세포 생물학, 유전체학 및 재조합 DNA의 통상의 기술을 사용한다. 문헌[Sambrook, Fritsch and Maniatis, MOLECULAR CLONING: A LABORATORY MANUAL, 2nd edition (1989)]; 문헌[CURRENT PROTOCOLS IN MOLECULAR BIOLOGY (F. M. Ausubel, et al. eds., (1987))]; 시리즈 문헌[METHODS IN ENZYMOLOGY (Academic Press, Inc.): PCR 2: A PRACTICAL APPROACH (M.J. MacPherson, B.D. Hames and G.R. Taylor eds. (1995)), Harlow and Lane, eds. (1988) ANTIBODIES, A LABORATORY MANUAL, and ANIMAL CELL CULTURE (R.I. Freshney, ed. (1987))]을 참조한다.

[0749] **벡터**

[0750] 일 양태에서, 본 발명은 CRISPR-Cas 시스템의 조작 및 최적화에 사용되는 벡터를 제공한다.

[0751] 본원에서 사용된 "벡터"는 하나의 환경에서 다른 환경으로 어떤 독립체의 전달을 허용하거나 촉진하는 도구이다. 그것은 레플리콘, 예컨대 플라스미드, 파지, 또는 코스미드이며, 여기에 다른 DNA 세그먼트가 삽입될 수 있고, 이로써 삽입된 세그먼트의 복제가 이루어질 수 있다. 일반적으로, 벡터는 적절한 제어 요소와 관련되었을 때 복제가 가능하다. 일반적으로, 용어 "벡터"는 결합된 다른 핵산을 수송할 수 있는 핵산 분자를 말한다. 벡터는, 제한은 아니지만, 단일가닥, 이중가닥, 또는 부분적으로 이중가닥인 핵산 분자; 하나 이상의 자유 단부

를 포함하는 핵산 분자, 자유 단부가 없는 핵산 분자(예를 들어, 원형); DNA를 포함하거나, RNA를 포함하거나 또는 둘 다 포함하는 핵산 분자; 및 본 분야에 공지된 폴리뉴클레오티드의 다른 변종들을 포함한다. 벡터의 한 가지 종류는 "플라스미드"인데, 이것은 원형 이중가닥 DNA 루프를 말하며, 여기에 예컨대 표준 분자 클로닝 기술에 의해서 추가의 DNA 세그먼트가 삽입될 수 있다. 벡터의 다른 종류는 바이러스 벡터로서, 바이러스-유래된 DNA 또는 RNA 서열이 바이러스에 봉입되는 벡터에 존재한다(예를 들어, 레트로바이러스, 복제 결함 레트로바이러스, 아데노바이러스, 복제 결함 아데노바이러스, 및 아데노-관련 바이러스(AAV)). 바이러스 벡터는 또한 숙주 세포에 트랜스펙션되는 바이러스에 의해서 보유된 폴리뉴클레오티드를 포함한다. 특정 벡터는 그것이 도입된 숙주 세포에서 자율적 복제가 가능하다(예를 들어, 박테리아 복제 기원을 가진 박테리아 벡터 및 에피솜 포유류 벡터). 다른 벡터(예를 들어, 비-에피솜 포유류 벡터)는 숙주 세포에 도입시 숙주 세포의 게놈에 통합되며, 이로써 숙주 게놈과 함께 복제된다. 더욱이, 특정 벡터는 그것이 작동 가능하게 연결된 유전자의 발현을 지시할 수 있다. 이러한 벡터는 본원에서 "발현 벡터"라고 언급된다. 재조합 DNA 기술에서 활용되는 통상적인 발현 벡터는 주로 플라스미드의 형태이다.

[0752] 재조합 발현 벡터는 숙주 세포에서 핵산의 발현에 적합한 형태로 본 발명의 핵산을 포함할 수 있는데, 이것은 재조합 발현 벡터가 하나 이상의 조절 구성요소를 포함한다는 것을 의미하며, 하나 이상의 조절 구성요소는 발현에 사용될 숙주 세포에 기초하여 선택될 수 있고, 발현될 핵산 서열에 작동 가능하게 연결된다. 재조합 발현 벡터에서 "작동 가능하게 연결된"은 관심대상의 뉴클레오티드 서열이 그 뉴클레오티드 서열의 발현을 허용하는 방식으로 조절 구성요소(들)에 연결된 것을 의미한다(예를 들어, 시험관내 전사/번역 시스템에서 또는 벡터가 숙주 세포에 도입되었을 때는 숙주 세포에서). 재조합 및 클로닝 방법과 관련하여, US 2004-0171156 A1으로 2004년 9월 2일자 공개된 미국 특허출원 10/815,730을 참조하며, 이것의 내용은 그 전체가 여기 참고로 포함된다.

[0753] 본 발명의 양태는 키메라 RNA 및 Cas9에 대한 벡터에 관한 것이다. 키메라 RNA 및 Cas9에 대한 바이시스트로닉 발현 벡터가 바람직하다. 일반적으로 그리고 특히 이 구현예에서, Cas9는 바람직하게 CBh 프로모터에 의해서 유래된다. 키메라 RNA는 바람직하게 U6 프로모터에 의해서 유래된다. 이상적으로 이 두 개가 조합된다. 키메라 가이드 RNA는 통상적으로 20bp 가이드 서열(Ns)로 구성되며, 이것은 tracr 서열과 이어질 수 있다 (아래 가닥의 첫 번째 "U"에서 전사체의 끝까지 이어진다). tracr 서열은 나타낸 대로 다양한 위치에서 트렁케이트될 수 있다. 가이드 서열과 tracr 서열은 tracr-메이트 서열에 의해서 분리되는데, tracr-메이트 서열은 GUUUUAGAGCUA일 수 있다. 이것 다음에는 나타낸 대로 루프 서열 GAAA가 올 수 있다. 이들은 모두 바람직한 예들이다. 출원인은 서베이어 분석에 의해서 인간 *EMX1* 및 *PVALB* 유전자좌에서 Cas-9-매개 삽입결실부를 증명했다. ChiRNA는 "+n" 칭호에 의해서 나타내며, crRNA는 가이드 서열과 tracr 서열이 다른 전사체들로서 발현되는 혼성체 RNA를 말한다. 본 출원 전체에서, 키메라 RNA는 또한 단일 가이드, 또는 합성 가이드 RNA(sgRNA)라고 칭해질 수 있다. 이 루프는 바람직하게 GAAA이지만, 이 서열에만 제한되는 것은 아니며, 실제로 단지 4 bp 길이에만 제한되는 것도 아니다. 실제로, 헤어핀 구조에서 사용되는 바람직한 루프 형성 서열은 4 뉴클레오티드 길이이고, 가장 바람직하게 서열 GAAA를 가진다. 그러나, 더 길거나 짧은 루프 서열도 사용될 수 있으며, 대용 서열들도 사용될 수 있다. 서열은 바람직하게 뉴클레오티드 트리플(예를 들어, AAA), 및 추가의 뉴클레오티드(예를 들어, C 또는 G)를 포함한다. 루프 형성 서열의 예들은 CAAA 및 AAAG를 포함한다.

[0754] 용어 "조절 구성요소"는 프로모터, 인핸서, 내부 리보솜 진입 부위(IRES), 및 다른 발현 제어 요소들(예를 들어, 전사 종결 신호, 예컨대 폴리아데닐화 신호 및 폴리-U 서열)을 포함하도록 의도된다. 이러한 조절 구성요소는, 예를 들어 Goeddel, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990)에 설명된다. 조절 구성요소는 많은 종류의 숙주 세포에서 뉴클레오티드 서열의 구성적 발현을 지시하는 것들과 특정 숙주 세포에서만 뉴클레오티드 서열의 발현을 지시하는 것들을 포함한다(예를 들어, 조직-특이적 조절 서열). 조직-특이적 프로모터는 근육, 뉴런, 뼈, 피부, 혈액, 특정 장기(예를 들어, 간, 췌장)와 같은 관심 대상의 바람직한 조직, 또는 특정한 세포 타입(예를 들어, 림프구)에서 주로 발현을 지시할 수 있다. 또한, 조절 구성요소는 시간-의존적 방식으로, 예컨대 세포-주기 의존적 또는 발생 단계 의존적 방식으로 발현을 지시할 수 있으며, 이것은 조직 특이적이거나 세포-타입 특이적일 수도 있고 아닐 수도 있다. 일부 구현예에서, 벡터는 하나 이상의 pol III 프로모터(예를 들어, 1, 2, 3, 4, 5, 또는 그 이상의 pol III 프로모터), 하나 이상의 pol II 프로모터(예를 들어, 1, 2, 3, 4, 5, 또는 그 이상의 pol II 프로모터), 하나 이상의 pol I 프로모터(예를 들어, 1, 2, 3, 4, 5, 또는 그 이상의 pol I 프로모터), 또는 이들의 조합을 포함한다. pol III 프로모터의 예들은, 제한은 아니지만, U6 및 H1 프로모터를 포함한다. pol II 프로모터의 예들은, 제한은 아니지만, 레트로바이러스 Rous 육종 바이러스(RSV) LTR 프로모터(선택적으로 RSV 인핸서와 함께), 시토메갈로바이러스(CMV) 프로모터(선택적으로 CMV 인핸서와 함께)[예를 들어, Boshart et al, Cell, 41:521-530

(1985) 참조], SV40 프로모터, 디하이드로폴레이트 환원효소 프로모터, β-액틴 프로모터, 포스포글리세롤 키나제(PGK) 프로모터, 및 EF1α 프로모터를 포함한다. 또한, 용어 "조절 구성요소"는 인핸서 요소, 예컨대 WPRE; CMV 인핸서; HTLV-I의 LTR에서 R-U5' 세그먼트(Mol. Cell. Biol., Vol. 8(1), p. 466-472, 1988); SV40 인핸서; 토끼 β-글로빈의 엑손 2와 3 사이의 인트론 서열(Proc. Natl. Acad. Sci. USA., Vol. 78(3), p. 1527-31, 1981)도 포함된다. 발현 벡터의 디자인은 형질전환될 숙주 세포의 선택, 원하는 발현 수준 등의 요인들에 따를 수 있다는 것이 당업자에 의해서 인정될 것이다. 벡터는 숙주 세포에 도입될 수 있고, 이로써 본원에 설명된 대로 핵산에 의해서 암호화된, 융합 단백질 또는 펩티드를 포함하는, 전사체, 단백질, 또는 펩티드가 생성될 수 있다(예를 들어, 군락화된 규칙적으로 산재된 짧은 팔린드롬형 반복부 (CRISPR) 전사체, 단백질, 효소, 이들의 돌연변이 형태, 이들의 융합 단백질 등). 조절 서열과 관련하여, 미국 특허출원 10/491,026을 참조하며, 이것의 내용은 그 전체가 여기 참고로 포함된다. 프로모터와 관련하여, PCT 공보 WO 2011/028929 및 미국 출원 12/511,940을 참조하며, 이들의 내용은 그 전체가 여기 참고로 포함된다.

[0755] 벡터는 원핵 또는 진핵 세포에서 CRISPR 전사물(예를 들어, 핵산 전사물, 단백질 또는 효소)의 발현을 위해 설계될 수 있다. 예를 들어, CRISPR 전사물은 박테리아 세포, 예를 들어, 에스케리키아 콜라이, 곤충 세포(배큘로바이러스 발현 벡터 사용), 효모 세포 또는 포유동물 세포에서 발현될 수 있다. 적절한 숙주 세포는 문헌[Goeddel, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990)]에 추가로 논의되어 있다. 대안적으로, 재조합 발현 벡터는 예를 들어, T7 프로모터 조절 서열 및 T7 중합효소를 사용하여 시험관내에서 전사되고 번역될 수 있다.

[0756] 벡터는 원핵생물, 또는 원핵생물 세포에 도입되고, 그에서 증식될 수 있다. 일부 구현예에서, 원핵생물은 진핵 세포로 도입되거나 또는 진핵 세포로 도입되는 벡터의 생성에서 중간체 벡터(예를 들어, 바이러스 벡터 패키징 시스템의 일부로서 플라스미드 증폭)로서 벡터의 카피를 증폭시키기 위해서 사용될 수 있다. 일부 구현예에서, 원핵생물은 벡터의 카피를 증폭시키고, 하나 이상의 핵산을 발현하기 위해, 예를 들어, 숙주 세포 또는 숙주 유기체로의 전달을 위한 하나 이상의 단백질의 공급원을 제공하기 위해 사용된다. 원핵생물에서의 단백질의 발현은 자주 융합 또는 비-융합 단백질 중 어느 하나의 발현을 유도하는 구성성 또는 유도성 프로모터를 함유하는 벡터를 사용하여 에스케리키아 콜라이에서 수행된다. 융합 벡터는 거기에 인코딩된 단백질로, 예를 들어, 재조합 단백질의 아미노 말단으로 수많은 아미노산을 부가한다. 이러한 융합 벡터는 다음과 같은 하나 이상의 목적을 제공할 수 있다: (i) 재조합 단백질의 발현의 증가; (ii) 재조합 단백질의 용해도의 증가; 및 (iii) 친화성 정제에서 리간드로 작용함으로써 재조합 단백질의 정제의 보조. 종종, 융합 발현 벡터에서, 단백질분해 절단 부위는 융합 모이어티와 재조합 단백질의 연접부에 도입되어, 융합 단백질의 정제 이후에 융합 모이어티로부터 재조합 단백질의 분리를 가능하게 한다. 이러한 효소 및 그들의 동족 인식 서열은 인자 Xa, 트롬빈 및 엔테로키나아제를 포함한다. 예시적인 융합 발현 벡터는 pGEX(파마시아 바이오테크 인코포레이티드(Pharmacia Biotech Inc); 문헌[Smith and Johnson, 1988. Gene 67: 31-40]), pMAL(미국 매사추세츠주 비벌리 소재의 뉴 잉글랜드 바이오랩스(New England Biolabs)) 및 pRIT5(미국 뉴저지주 피스카타웨이 소재의 파마시아(Pharmacia))를 포함하며, 이는 각각 글루타티온 S-트랜스퍼라제(GST), 말토스 E 결합 단백질 또는 단백질 A를 표적 재조합 단백질에 융합시킨다.

[0757] 적절한 유도성 비-융합 에스케리키아 콜라이 발현 벡터의 예는 pTrc(문헌[Amrann et al., (1988) Gene 69:301-315]) 및 pET 11d(문헌[Studier et al., GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990) 60-89])를 포함한다.

[0758] 일부 구현예에서, 벡터는 효모 발현 벡터이다. 효모 사카로마이세스 세레비지애에서의 발현을 위한 벡터의 예에는 pYepSec1(문헌[Baldari, et al., 1987. EMBO J. 6: 229-234]), pMFa(문헌[Kuijan and Herskowitz, 1982. Cell 30: 933-943]), pJRY88(문헌[Schultz et al., 1987. Gene 54: 113-123]), pYES2(미국 캘리포니아주 샌 디에고 소재의 인비트로겐 코포레이션) 및 picZ(미국 캘리포니아주 샌 디에고 소재의 인비트로겐 코포레이션)가 포함된다.

[0759] 일부 구현예에서, 벡터는 배큘로바이러스 발현 벡터를 사용하여 곤충 세포에서 단백질 발현을 유도한다. 배양된 곤충 세포(예를 들어, SF9 세포)에서 단백질의 발현에 이용가능한 배큘로바이러스 벡터는 pAc 시리즈(문헌[Smith, et al., 1983. Mol. Cell. Biol. 3: 2156-2165]) 및 pVL 시리즈(문헌[Lucklow and Summers, 1989. Virology 170: 31-39])를 포함한다.

[0760] 일부 구현예에서, 벡터는 포유동물 발현 벡터를 사용하여 포유동물 세포에서 하나 이상의 서열의 발현을 유도할 수 있다. 포유동물 발현 벡터의 예는 pCDM8(문헌[Seed, 1987. Nature 329: 840]) 및 pMT2PC(문헌[Kaufman, et

al., 1987. EMBO J. 6: 187-195])를 포함한다. 포유동물 세포에서 사용되는 경우, 발현 벡터의 조절 기능은 통상적으로 하나 이상의 조절 요소에 의해 제공된다. 예를 들어, 통상적으로 사용되는 프로모터는 폴리오마, 아데노바이러스 2, 사이토메갈로바이러스, 유인원 바이러스 40 및 본원에 개시되고 당업계에 공지되어 있는 기타의 것으로부터 유래된다. 원핵 및 진핵 세포 둘 모두를 위한 다른 적절한 발현 시스템에 대하여, 예를 들어, 문헌[Sambrook, et al., MOLECULAR CLONING: A LABORATORY MANUAL. 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989]의 16 및 17장을 참조한다.

[0761] 일부 구현예에서, 재조합 포유동물 발현 벡터는 특정 세포 유형에서 우선적으로 핵산의 발현을 유도할 수 있다(예를 들어, 핵산을 발현하기 위하여 조직-특이적 조절 요소가 사용됨). 조직-특이적 조절 요소가 해당 분야에 공지되어 있다. 적절한 조직-특이적 프로모터의 비제한적인 예에는 알부민 프로모터(간-특이적; 문헌[Pinkert, et al., 1987. Genes Dev. 1: 268-277]), 림프-특이적 프로모터(문헌[Calame and Eaton, 1988. Adv. Immunol. 43: 235-275]), 특히, T 세포 수용체(문헌[Winoto and Baltimore, 1989. EMBO J. 8: 729-733]) 및 면역글로불린의 프로모터(문헌[Baneiji, et al., 1983. Cell 33: 729-740]; 문헌[Queen and Baltimore, 1983. Cell 33: 741-748]), 뉴런-특이적 프로모터(예를 들어, 신경섬유 프로모터; 문헌[Byrne and Ruddle, 1989. Proc. Natl. Acad. Sci. USA 86: 5473-5477]), 췌장-특이적 프로모터(문헌[Edlund, et al., 1985. Science 230: 912-916]) 및 유선-특이적 프로모터(예를 들어, 유장(milk whey) 프로모터; 미국 특허 제4,873,316호 및 유럽 출원 공개 제264,166호)가 포함된다. 발생-조절 프로모터, 예를 들어, 쥣과 hox 프로모터(문헌[Kessel and Gruss, 1990. Science 249: 374-379]) 및 α-태아단백질 프로모터(문헌[Campes and Tilghman, 1989. Genes Dev. 3: 537-546])도 또한 포함된다. 원핵세포 및 진핵세포 벡터와 관련하여, 미국특허 6,750,059를 참조하며, 이것의 내용은 그 전체가 여기 참고로 포함된다. 본 발명의 다른 구현예들은 바이러스 벡터의 사용에 관한 것일 수 있으며, 이것과 관련해서는 미국 특허출원 13/092,085를 참조하고, 이것의 내용은 그 전체가 여기 참고로 포함된다. 조직-특이적 조절 구성요소는 본 분야에 공지이며, 이것과 관련하여 미국특허 7,776,321을 참조하고, 이것의 내용은 그 전체가 여기 참고로 포함된다.

[0762] **조절 구성요소**

[0763] 일부 구현예에서, 조절 요소는 CRISPR 시스템의 하나 이상의 요소에 작동 가능하게 연결되어 CRISPR 시스템의 하나 이상의 요소의 발현을 유도한다. 일반적으로, SPIDR(스페이서 산재 직접 반복부)로도 공지되어 있는 CRISPR(클러스터링되고 규칙적으로 산재된 짧은 팔린드로믹 반복부)은 통상 특정 박테리아 종에 특이적인 DNA 유전자좌의 과를 구성한다. CRISPR 유전자좌는 에스케리키아 콜라이에서 인식되는 별개의 부류의 산재된 짧은 서열 반복부(SSR) 및 관련 유전자를 포함한다(문헌[Ishino et al., J. Bacteriol., 169:5429-5433 [1987]]; 및 문헌[Nakata et al., J. Bacteriol., 171:3553-3556 [1989]]). 유사한 산재된 SSR이 할로페락스 메디테라네이(Haloferax mediterranei), 스트렙토코커스 피오게네스, 아나바에나(Anabaena) 및 마이코박테리움 튜베르큘로시스(Mycobacterium tuberculosis)에서 확인되었다(문헌[Groenen et al., Mol. Microbiol., 10:1057-1065 [1993]]; 문헌[Hoe et al., Emerg. Infect. Dis., 5:254-263 [1999]]; 문헌[Masepohl et al., Biochim. Biophys. Acta 1307:26-30 [1996]]; 및 문헌[Mojica et al., Mol. Microbiol., 17:85-93 [1995]] 참조). CRISPR 유전자좌는 통상적으로 SRSR(규칙적으로 산재된 짧은 반복부(short regularly spaced repeats))로 명명된 반복부의 구조가 다른 SSR과 상이하다(문헌[Janssen et al., OMICS J. Integ. Biol., 6:23-33 [2002]]; 및 문헌[Mojica et al., Mol. Microbiol., 36:244-246 [2000]]). 일반적으로, 반복부는 실질적으로 고정된 길이를 갖는 독특한 개재 서열에 의해 규칙적으로 산재된 클러스터에 존재하는 짧은 요소이다(상기 문헌[Mojica et al., [2000]]). 반복 서열이 균주들 간에 고도로 보존되어 있지만, 산재된 반복부의 수와 스페이서 영역의 서열은 통상적으로 균주마다 상이하다(문헌[van Embden et al., J. Bacteriol., 182:2393-2401 [2000]]). CRISPR 유전자좌는 아에로피룸(Aeropyrum), 피로바쿨룸(Pyrobaculum), 술폴로부스(Sulfolobus), 아캐오글로부스(Archaeoglobus), 할로카르쿨라(Halocarcula), 메타노박테리움(Methanobacterium), 메타노코커스(Methanococcus), 메타노사르시나(Methanosarcina), 메타노피러스(Methanopyrus), 피로코커스(Pyrococcus), 피크로필러스(Picrophilus), 써모플라스마(Thermoplasma), 코리네박테리움(Corynebacterium), 마이코박테리움(Mycobacterium), 스트렙토마이세스(Streptomyces), 아퀴펙스(Aquifex), 포르피로모나스(Porphyromonas), 클로로비움(Chlorobium), 써머스(Thermus), 바실러스(Bacillus), 리스테리아(Listeria), 스태필로코커스(Staphylococcus), 클로스트리디움(Clostridium), 써모아나에로박터(Thermoanaerobacter), 마이코플라스마(Mycoplasma), 푸소박테리움(Fusobacterium), 아자쿠스(Azarcus), 크로모박테리움(Chromobacterium), 네이세리아(Neisseria), 니트로소모나스(Nitrosomonas), 데설포비브리오(Desulfovibrio), 게오박터(Geobacter), 믹소코커스(Myxococcus), 캄필로박터(Campylobacter), 볼리넬라(Wolinella), 아시네토박터(Acinetobacter), 에르위니아(Erwinia), 에스케리키아, 레지오넬라(Legionella), 메틸로코커스(Methylococcus), 파스퇴렐라

(Pasteurella), 포토박테리움(Photobacterium), 살모넬라(Salmonella), 잔토모나스(Xanthomonas), 예르시니아(Yersinia), 트레포네마(Treponema) 및 써모토가(Thermotoga)를 포함하나 이들에 한정되지 않는 40개 초과의 원핵생물에서 확인되었다(예를 들어, 문헌[Jansen et al., Mol. Microbiol., 43:1565-1575 [2002]]; 및 문헌[Mojica et al., [2005]] 참조).

[0764] 일반적으로, "CRISPR 시스템"은 집합적으로 Cas 유전자를 인코딩하는 서열, tracr(트랜스-활성화 CRISPR) 서열(예를 들어, tracrRNA 또는 활성 부분 tracrRNA), tracr-메이트 서열(내인성 CRISPR 시스템의 맥락에서 "직접 반복부" 및 tracrRNA-가공 부분 직접 반복부 포함), 가이드 서열(내인성 CRISPR 시스템의 맥락에서 "스페이서"로도 지칭) 또는 CRISPR 유전자좌로부터의 기타 서열 및 전사물을 포함하는 CRISPR-관련("Cas") 유전자의 발현에 수반되거나, 그의 활성을 유도하는 전사물 및 다른 요소를 지칭한다. 본 발명의 구현예에서, 용어 가이드 서열과 가이드 RNA는 상호 교환하여 사용된다. 일부 구현예에서, CRISPR 시스템의 하나 이상의 요소는 I형, II형 또는 III형 CRISPR 시스템으로부터 유래된다. 일부 구현예에서, CRISPR 시스템의 하나 이상의 요소는 내인성 CRISPR 시스템을 포함하는 특정 유기체, 예를 들어, 스트렙토코커스 피오게네스로부터 유래된다. 일반적으로, CRISPR 시스템은 표적 서열의 부위에서 CRISPR 복합체의 형성을 증진시키는 요소(내인성 CRISPR 시스템의 맥락에서 프로토스페이서로도 지칭)를 특징으로 한다. CRISPR 복합체의 형성의 맥락에서, "표적 서열"은 가이드 서열이 상보성을 갖도록 설계된 서열을 지칭하며, 여기서, 표적 서열과 가이드 서열 간의 혼성화는 CRISPR 복합체의 형성을 증진시킨다. 표적 서열은 임의의 폴리뉴클레오티드, 예를 들어, DNA 또는 RNA 폴리뉴클레오티드를 포함할 수 있다. 일부 구현예에서, 표적 서열은 세포의 핵 또는 세포질 내에 위치한다.

[0765] 일부 구현예에서, 직접 반복부는 다음의 기준을 일부 또는 전부 만족하는 반복 모티프를 검색함으로써 인실리코 방식으로 확인될 수 있다:

[0766] 1. 타입 II CRISPR 유전자좌가 측접한 게놈 서열의 2 Kb 창에서 발견됨;

[0767] 2. 20 내지 50 bp에 걸쳐 있음;

[0768] 3. 20 내지 50 bp에 산재되어 있음.

[0769] 일부 구현예에서, 이들 기준 중 두 가지가 사용될 수 있으며, 예를 들어 1과 2, 2와 3, 또는 1과 3이 사용될 수 있다. 일부 구현예에서, 세 가지 기준이 모두 사용될 수 있다.

[0770] 일부 구현예에서, 후보 tracrRNA가 다음의 기준을 일부 또는 전부 만족하는 서열에 의해서 연속하여 예측될 수 있다:

[0771] 1. 직접 반복부에 상동성인 서열(최대 18-bp 미스매치를 가진 Geneious에서 모티프 검색);

[0772] 2. 전사 방향으로 예측된 Rho-독립적 전사 터미네이터의 존재;

[0773] 3. tracr RNA와 직접 반복부 사이에 안정한 헤어핀 2차 구조.

[0774] 일부 구현예에서, 이들 기준 중 두 가지가 사용될 수 있으며, 예를 들어 1과 2, 2와 3, 또는 1과 3이 사용될 수 있다. 일부 구현예에서, 세 가지 기준이 모두 사용될 수 있다.

[0775] 일부 구현예에서, 키메라 합성 가이드 RNA (sgRNA) 디자인은 직접 반복부와 tracr RNA 사이에 적어도 12 bp의 듀플렉스 구조를 통합할 수 있다.

[0776] 본 발명의 바람직한 구현예에서, CRISPR 시스템은 타입 II CRISPR 시스템이고, Cas 효소는 Cas9이며, 이것은 DNA 절단을 촉매한다. *Streptococcus pyogenes*로부터 유래된 Cas9 또는 어떤 밀접하게 관련된 Cas9에 의한 효소 작용은 20 뉴클레오티드의 가이드 서열과 혼성화하는 표적 부위 서열에 이중가닥 브레이크를 생성하며, 표적 부위 서열은 20 뉴클레오티드의 표적 서열 뒤에 프로토스페이서-인접 모티프(PAM) 서열(예들은 여기 설명된 대로 결정될 수 있는 NGG/NRG 또는 PAM을 포함한다)을 가진다. 부위-특이적 DNA 인식 및 절단에 대한 Cas9를 통한 CRISPR 활성은 가이드 서열, 가이드 서열과 부분적으로 혼성화하는 tracr 서열 및 PAM 서열에 의해서 한정된다. CRISPR 시스템의 더 이상의 양태들이 Karginov and Hannon, The CRISPR System: small RNA-guided defence in bacteria and archaea, Mole Cell 2010, January 15; 37(1):7에 설명된다.

[0777] *Streptococcus pyogenes* SF370으로부터의 타입 II CRISPR 유전자좌는 4 개 유전자 Cas9, Cas1, Cas2 및 Csn1의 클러스터뿐만 아니라 2개의 비-코딩 RNA 요소, tracr RNA 및 비-반복 서열의 짧은 스트레 (스페이서, 각각 약 30bp) 가 개재된 반복 서열(직접 반복부)의 특징적인 어레이를 함유한다. 이 시스템에서, 표적화된 DNA 이중가닥 브레이크 (DSB)가 순차적 4 단계에서 생성된다(도 2a). 먼저, 2 개의 비-코딩 RNA, 예비-crRNA 어레이 및

tracr RNA가 CRISPR 유전자좌로부터 전사된다. 두 번째로, tracr RNA가 예비-crRNA의 직접 반복부와 혼성화하고, 이것은 이어서 개별적 스페이서 서열을 함유하는 성숙한 crRNA로 가공된다. 세 번째로, 성숙한 crRNA:tracr RNA 복합체가 Cas9를 crRNA의 스페이서 영역과 프로토스페이서 DNA 간의 헤테로듀플렉스 형성을 통해 프로토스페이서와 상응하는 PAM으로 구성된 DNA 표적으로 보낸다. 마지막으로, Cas9가 PAM의 상류에서 표적 DNA의 절단을 매개하여 프로토스페이서 내에 DSB를 생성한다(도 2a). 도 2b는 코돈 최적화된 Cas9의 핵 위치를 증명한다. 정확한 전사 개시를 촉진하기 위해서, RNA 폴리머라제 III-기반 U6 프로모터가 tracr RNA의 발현을 유도하기 위해 선택되었다(도 2c). 유사하게, U6 프로모터-기반 작제물이 2 개의 직접 반복부 (DRs, 용어 "tracr-메이트 서열"에 의해서도 포함된다; 도 2c)가 측접된 단일 스페이서로 구성된 예비-crRNA 어레이를 발현하도록 개발되었다. 초기 스페이서는 대뇌피질의 발생시 핵심 유전자인 인간 EMX1 유전자좌(도 2c)에서 33-염기쌍 (bp) 표적 부위(Cas9의 NGG 인식 모티프를 만족하는 30-bp 프로토스페이서 더하기 3-bp CRISPR 모티프(PAM) 서열)를 표적화하도록 설계되었다.

[0778] 통상적으로, 내인성 CRISPR 시스템의 맥락에서, CRISPR 복합체(표적 서열에 혼성화되고, 하나 이상의 Cas 단백질과 복합체화되는 가이드 서열을 포함)의 형성은 표적 서열 내의 또는 그 근처의(예를 들어, 그로부터 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 50 개 이상의 염기쌍 내의) 하나 또는 두 가닥 모두의 절단을 야기한다. 이론에 구속되지 않으면서, 야생형 tracr 서열의 전부 또는 그의 일부(예를 들어, 야생형 tracr 서열의 약 20, 26, 32, 45, 48, 54, 63, 67, 85 개 이상의 뉴클레오티드)를 포함하거나 그로 이루어질 수 있는 tracr 서열은 또한, 가이드 서열에 작동 가능하게 연결된 tracr 메이트 서열의 전부 또는 일부로의 tracr 서열의 적어도 일부분에 따른 혼성화에 의해서와 같이 CRISPR 복합체의 일부를 형성할 수 있다. 일부 구현예에서, CRISPR 시스템의 하나 이상의 요소의 발현을 유도하는 하나 이상의 벡터는 CRISPR 시스템의 요소의 발현이 하나 이상의 표적 부위에서 CRISPR 복합체의 형성을 유도하도록 숙주 세포 내로 도입된다. 예를 들어, Cas 효소, tracr-메이트 서열에 연결된 가이드 서열 및 tracr 서열은 각각 개별 벡터 상의 개별 조절 요소에 작동 가능하게 연결될 수 있다. 대안적으로, 동일하거나 상이한 조절 요소로부터 발현되는 요소 중 둘 이상은 단일의 벡터에서 조합될 수 있으며, 하나 이상의 추가의 벡터는 제1 벡터에 포함되지 않은 CRISPR 시스템의 임의의 성분을 제공한다. 단일의 벡터에서 조합되는 CRISPR 시스템 요소는 임의의 적절한 방향으로 배열될 수 있으며, 예를 들어, 하나의 요소는 제2 요소에 대하여 5'에(그의 "상류"에) 위치하거나 그에 대하여 3'에(그의 "하류"에) 위치한다. 하나의 요소의 코딩 서열은 제2 요소의 코딩 서열의 동일한 가닥 또는 마주하는 가닥에 위치할 수 있으며, 동일하거나 반대 방향으로 방향화될 수 있다. 일부 구현예에서, 단일의 프로모터는 CRISPR 효소를 인코딩하는 전사물 및 하나 이상의 인트론 서열 내에(예를 들어, 각각이 상이한 인트론 내에, 2 개 이상이 적어도 하나의 인트론 내에 또는 전부가 단일의 인트론 내에) 매립된 가이드 서열, tracr 메이트 서열(선택적으로 가이드 서열에 작동 가능하게 연결), 및 tracr 서열 중 하나 이상의 발현을 유도한다. 일부 구현예에서, CRISPR 효소, 가이드 서열, tracr 메이트 서열 및 tracr 서열이 동일한 프로모터에 작동 가능하게 연결되고, 그로부터 발현된다.

[0779] 일부 구현예에서, 벡터는 하나 이상의 삽입 부위, 예를 들어, 제한 엔도뉴클레아제 인식 서열("클로닝 부위"로도 지칭)을 포함한다. 일부 구현예에서, 하나 이상의 삽입 부위(예를 들어, 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 개 이상의 삽입 부위)는 하나 이상의 벡터의 하나 이상의 서열 요소의 상류 및/또는 하류에 위치한다. 일부 구현예에서, 벡터는 tracr 메이트 서열의 상류에 있고, 선택적으로 tracr 메이트 서열에 작동 가능하게 연결된 조절 요소의 하류에 있는 삽입 부위를 포함하여, 삽입 부위로의 가이드 서열의 삽입 후에, 그리고 발현 시에, 가이드 서열이 진핵 세포 내의 표적 서열로의 CRISPR 복합체의 서열-특이적 결합을 유도하게 한다. 일부 구현예에서, 벡터는 2 개 이상의 삽입 부위를 포함하며, 각각의 삽입 부위는 2 개의 tracr 메이트 서열 사이에 위치하여, 각 부위에서 가이드 서열의 삽입을 가능하게 한다. 이러한 배열에서, 2 개 이상의 가이드 서열은 단일의 가이드 서열의 2 개 이상의 카피, 2 개 이상의 상이한 가이드 서열 또는 이들의 조합을 포함할 수 있다. 다중의 상이한 가이드 서열이 사용되는 경우, 단일의 발현 작제물을 사용하여 세포 내의 다중의 상이한 상응하는 표적 서열에 CRISPR 활성을 표적화할 수 있다. 예를 들어, 단일의 벡터는 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20 개 이상의 가이드 서열을 포함할 수 있다. 일부 구현예에서, 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 개 이상의 이러한 가이드-서열-함유 벡터가 제공될 수 있으며, 선택적으로 세포로 전달될 수 있다.

[0780] 일부 구현예에서, 벡터는 CRISPR 효소, 예를 들어, Cas 단백질을 인코딩하는 효소-코딩 서열에 작동 가능하게 연결된 조절 요소를 포함한다. Cas 단백질의 비제한적인 예는 Cas1, Cas1B, Cas2, Cas3, Cas4, Cas5, Cas6, Cas7, Cas8, Cas9(Csn1 및 Csx12로도 알려짐), Cas10, Csy1, Csy2, Csy3, Cse1, Cse2, Csc1, Csc2, Csa5, Csn2, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Csb1, Csb2, Csb3, Csx17, Csx14, Csx10, Csx16, CsaX, Csx3, Csx1, Csx15, Csf1, Csf2, Csf3, Csf4, 그의 상동체 또는 그의 변형된 버전을 포함한다. 일부 구현예에서, 비변형 CRISPR 효소, 예를 들어, Cas9는 DNA 절단 활성을 갖는다. 일부 구현예에서,

CRISPR 효소는 표적 서열 내 및/또는 표적 서열의 상보물 내에서와 같은 표적 서열의 위치에서 1개 또는 2 개의 가닥의 절단을 유도한다. 일부 구현예에서, CRISPR 효소는 표적 서열의 처음 또는 마지막 뉴클레오티드로부터 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 50, 100, 200, 500 개 이상의 염기쌍에서 1 개 또는 2 개의 가닥의 절단을 유도한다. 일부 구현예에서, 벡터는 상응하는 야생형 효소에 대하여 돌연변이되어, 돌연변이된 CRISPR 효소에 표적 서열을 함유하는 표적 폴리뉴클레오티드의 1 개 또는 2 개 모두의 가닥의 절단 능력이 결여되게 한 CRISPR 효소를 인코딩한다. 예를 들어, 스트렙토코커스 피오게네스로부터의 Cas9의 RuvC I 촉매 도메인 내에서의 아스파르트산에서 알라닌으로의 치환(D10A)은 Cas9를 두 가닥 모두를 절단하는 뉴클레아제에서 닉카아제(단일 가닥 절단)로 전환시킨다. Cas9가 닉카아제가 되게 하는 돌연변이의 다른 예는 제한 없이, H840A, N854A 및 N863A를 포함한다. 추가의 예로서, Cas9의 2개 이상의 촉매 도메인(RuvC I, RuvC II 및 RuvC III 또는 HNH 도메인)을 돌연변이시켜, 모든 DNA 절단 활성이 실질적으로 결여된 돌연변이된 Cas9를 생성할 수 있다. 일부 구현예에서, D10A 돌연변이를 H840A, N854A 또는 N863A 돌연변이 중 하나 이상과 조합하여, 모든 DNA 절단 활성이 실질적으로 결여된 Cas9 효소를 생성한다. 일부 구현예에서, CRISPR 효소는 돌연변이된 효소의 DNA 절단 활성이 비-돌연변이 형태에 대하여 약 25%, 10%, 5%, 1%, 0.1%, 0.01% 이하인 경우 모든 DNA 절단 활성이 실질적으로 결여된 것으로 여겨진다. 효소가 SpCas9가 아닌 경우, 돌연변이는 SpCas9의 위치 10, 762, 840, 854, 863 및/또는 986에 상응하는 임의의 또는 모든 잔기에서 이루어질 수 있다(이것은 예를 들어 표준 서열 비교 도구에 의해서 확인될 수 있다). 특히, 다음의 돌연변이 중 일부 또는 전부가 SpCas9에서 바람직하다: D10A, E762A, H840A, N854A, N863A 및/또는 D986A; 뿐만 아니라, 대체 아미노산 중 임의의 것에 대한 보존적 치환도 생각된다. 다른 Cas9에서 상응하는 위치에서의 동일한 돌연변이 (또는 이들 돌연변이의 보존적 치환)도 바람직하다. D10 및 H840이 SpCas9에서 특히 바람직하다. 그러나, 다른 Cas9에서도 SpCas9 D10 및 H840에 상응하는 잔기가 또한 바람직하다.

[0781] SpCas9의 RuvC I 촉매 도메인에서 아스파르테이트-알라닌 치환 (D10A)은 이 뉴클레아제를 닉카아제(SpCas9n)로 전환하기 위하여 유전자 조작되었고(예를 들어, 문헌[Sapranauskas et al., 2011, Nucleic Acis Research, 39: 9275; Gasiunas et al., 2012, Proc. Natl. Acad. Sci. USA, 109:E2579] 참조), 이로써 흉터형성된 게놈 DNA가 고-충실도 상동성-지정 수선 (HDR)을 겪게 된다. 서베이어 분석이 SpCas9n이 EMX1 프로토스페이서 표적에 삽입결실부를 생성하지 않는다는 것을 확인했다. EMX1-표적화 키메라 crRNA (tracr RNA 구성성분도 가짐)와 SpCas9의 공-발현은 표적 부위에 삽입결실부를 생성했지만, SpCas9n과의 공-발현은 그렇지 않았다(n=3). 더욱이, 327 앰플리콘의 시퀀싱은 SpCas9n에 의해서 유도된 어떤 삽입결실부도 검출하지 않았다. HEK 293FT 세포와 키메라 RNA 표적화 EMX1, hSpCas9 또는 hSpCas9, 뿐만 아니라 프로토스페이서 근처에 한 쌍의 제한 부위(HindIII 및 NheI)를 도입하기 위한 HR 주형을 공-트랜스펙션함으로써 CRISPR-매개 HR을 시험하기 위해서 동일한 유전자좌가 선택되었다.

[0782] 바람직한 오솔로그가 여기 설명된다. Cas 효소는 Cas9로 확인될 수 있으며, 이것은 타입 II CRISPR 시스템으로부터의 다중 뉴클레아제 도메인을 가진 가장 큰 뉴클레아제에 대한 상동성을 공유하는 일반적인 효소 부류를 말한다. 가장 바람직하게, Cas9 효소는 spCas9 또는 saCas9로부터의 것이거나, 또는 그로부터 유래된다. 출원인에 의하면, 유래된다는 것은 그 유래된 효소가 높은 정도의 서열 상동성을 가진다는 의미에서 야생형 효소에 상당히 기초하지만, 그것이 여기 설명된 대로 어떤 방식으로 돌연변이 (변형)되었다는 것을 의미한다.

[0783] 명백히 다르지 않다면 용어 Cas 및 CRISPR 효소는 여기서 일반적으로 상호교환하여 사용된다는 것이 인정될 것이다. 상기 언급된 대로, 여기 사용된 잔기 넘버링은 대부분 *Streptococcus pyogenes*에서 타입 II CRISPR 유전자좌로부터의 Cas9 효소를 말한다. 그러나, 본 발명은 SpCas9, SaCa9, St1Cas9 등과 같은 다른 미생물 종들로부터의 더 많은 Cas9를 포함한다는 것이 인정될 것이다.

[0784] **코돈 최적화**

[0785] 코돈 최적화된 서열의 예가, 이 예에서는 인간에 대해 최적화된 예가 여기 제공되며 (즉, 인간에서의 발현을 위해 최적화됨), SaCas9 인간 코돈 최적화된 서열을 참조한다. 이것이 바람직하지만, 다른 예들도 가능하며, 숙주 종에 대한 코돈 최적화가 공지되어 있다는 것이 인정될 것이다.

[0786] 일부 구현예에서, CRISPR 효소를 암호화하는 효소 암호화 서열이 진핵 세포와 같은 특정 세포에서의 발현에 대해 코돈 최적화된다. 진핵 세포는, 제한은 아니지만, 인간, 마우스, 래트, 토끼, 개, 또는 비인간 포유류 또는 영장류와 같은 특정 유기체의 것들이거나, 또는 그로부터 유래될 수 있다. 일부 구현예에서, 사람이나 동물에 어떤 실질적인 의학적 유익 없이 그들을 고통스럽게 할 수 있는 인간의 생식계열 유전자 동일성을 변형하기 위한 과정 및/또는 동물의 유전자 동일성을 변형하기 위한 과정과 또한 이러한 과정으로부터 생긴 동물은 배제될

수 있다.

[0787] 일반적으로, 코돈 최적화는 고유 서열의 적어도 하나의 코돈(예를 들어, 약 1, 2, 3, 4, 5, 10, 15, 20, 25, 50 개 이상의 코돈)을 숙주 세포의 유전자에 더욱 빈번하게 또는 가장 빈번하게 사용되는 코돈으로 대체하면서, 고유 아미노산 서열을 유지함으로써 대상 숙주 세포에서의 발현의 증진을 위해 핵산 서열을 변형시키는 과정을 지칭한다. 다양한 종은 특정 아미노산의 특정 코돈에 대한 특정 편향을 나타낸다. 코돈 편향(유기체 간의 코돈 사용의 차이)은 종종 전령 RNA(mRNA)의 번역의 효율과 상호관련되며, 이는 차례로, 다른 것들 중에, 번역되는 코돈의 특성 및 특정 운반 RNA(tRNA) 분자의 이용가능성에 좌우되는 것으로 여겨진다. 세포에서의 선택된 tRNA의 우세는 일반적으로 펩티드 합성에 가장 빈번하게 사용되는 코돈을 반영하는 것이다. 따라서, 유전자는 코돈 최적화에 기초하여 주어진 유기체에서의 최적의 유전자 발현을 위해 맞춤화될 수 있다. 코돈 사용 표는 예를 들어, "코돈 사용 데이터베이스"(www.kazusa.orjp/codon/에서 이용가능함(2002년 7월 9일 방문))에서 용이하게 이용가능하며, 이들 표는 다수의 방식으로 적합하게 될 수 있다. 문헌[Nakamura, Y., et al. "Codon usage tabulated from the international DNA sequence databases: status for the year 2000" Nucl. Acids Res. 28:292 (2000)]을 참조한다. 특정 숙주 세포에서의 발현을 위해 특정 서열을 코돈 최적화시키는 컴퓨터 알고리즘도 또한 이용가능하며, 예를 들어, 진 포르지(Gene Forge)(압타젠(Aptagen); 미국 펜실베니아주 야코부스)도 또한 이용가능하다. 일부 구현예에서, CRISPR 효소를 인코딩하는 효소 내의 하나 이상의 코돈(예를 들어, 1, 2, 3, 4, 5, 10, 15, 20, 25, 50 개 이상 또는 모든 코돈)은 특정 아미노산에 대하여 가장 빈번하게 사용되는 코돈에 상응한다.

[0788] **핵 국소화 서열(NLS)**

[0789] 일부 구현예에서, 벡터는 하나 이상의 핵 국소화 서열(NLS), 예를 들어, 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 개 이상의 NLS를 포함하는 CRISPR 효소를 인코딩한다. 일부 구현예에서, CRISPR 효소는 아미노-말단에 또는 그 근처에 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 개 이상의 NLS, 카르복시-말단에 또는 그 근처에 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 개 이상의 NLS, 또는 이들의 조합(예를 들어, 아미노 말단에 하나 이상의 NLS 및 카르복시 말단에 하나 이상의 NLS)을 포함한다. 1 개 초과의 NLS가 존재하는 경우, 각각은 단일의 NLS가 1 개 초과의 카피로 존재하고/존재하거나 1 개 이상의 카피로 존재하는 하나 이상의 다른 NLS와 함께 존재할 수 있도록 다른 것들로부터 독립적으로 선택될 수 있다. 본 발명의 바람직한 구현예에서, CRISPR 효소는 최대 6개의 NLS를 포함한다. 일부 구현예에서, NLS는 NLS의 가장 가까운 아미노산이 N- 또는 C-말단으로부터 폴리펩티드 쇄를 따라 약 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 40, 50 개 이상의 아미노산 내에 존재하는 경우 N- 또는 C-말단 근처에 있는 것으로 여겨진다. NLS의 비제한적인 예는 하기로부터 유래된 NLS 서열을 포함한다: 아미노산 서열 PKKKRKV를 갖는 SV40 바이러스 대형 T-항원의 NLS; 뉴클레오플라스민(nucleoplasmin)으로부터의 NLS(예를 들어, 서열 KRPAATKKAGQAKKKK를 갖는 뉴클레오플라스민 이분(bipartite) NLS); 아미노산 서열 PAAKRVKLD 또는 RQRRNELKRSP를 갖는 c-myc NLS; 서열 NQSSNFGPMKGGNFGGRSSGPYGGGGQYFAKPRNQGGY를 갖는 hRNPA1 M9 NLS; 임포틴-알파로부터의 IBB 도메인의 서열 RMRIZFKNKGKDTAELRRRRVEVSVELRKAKKDEQILKRRNV; 마이오마(myoma) T 단백질의 서열 VSRKRPRP 및 PPKKARED; 인간 p53의 서열 POPKKKPL; 마우스 c-abl IV의 서열 SALIKKKKKMAP; 인플루엔자 바이러스 NS1의 서열 DRLRR 및 PKQKKRK; 간염 바이러스 델타 항원의 서열 RKLKKKIKKL; 마우스 Mx1 단백질의 서열 REKKKFLKRR; 인간 폴리(ADP-리보스) 중합효소의 서열 KRKGDEVDGVDEVAKKKSKK; 및 스테로이드 호르몬 수용체(인간) 글루코코르티코이드의 서열 RKCLQAGMNLEARKTKK.

[0790] 일반적으로, 하나 이상의 NLS는 진핵 세포의 핵에서 검출가능한 양의 CRISPR 효소의 축적을 유도하기에 충분한 세기의 것이다. 일반적으로, 핵 국소화 활성의 세기는 CRISPR 효소 내의 NLS의 수, 사용되는 특정 NLS(들) 또는 이들 인자의 조합으로부터 유래할 수 있다. 핵에서의 축적의 검출은 임의의 적절한 기술에 의해 수행될 수 있다. 예를 들어, 검출가능한 마커는 예를 들어, 핵의 위치를 검출하기 위한 수단(예를 들어, 핵에 특이적인 염색제, 예를 들어, DAPI)과 함께 세포 내의 위치가 가시화될 수 있도록 CRISPR 효소에 융합될 수 있다. 또한, 세포 핵을 세포로부터 분리할 수 있으며, 그 다음, 그의 내용물을 단백질을 검출하기 위한 임의의 적절한 과정, 예를 들어, 면역조직화학, 웨스턴 블롯 또는 효소 활성 검정에 의해 분석할 수 있다. 또한, 핵에서의 축적은 예를 들어, CRISPR 효소 또는 복합체에 노출되지 않거나, 하나 이상의 NLS가 결여된 CRISPR 효소에 노출된 대조군과 비교하여, 간접적으로, 예를 들어, CRISPR 복합체 형성의 영향에 대한 검정(예를 들어, 표적 서열에서의 DNA 절단 또는 돌연변이에 대한 검정, 또는 CRISPR 복합체 형성 및/또는 CRISPR 효소 활성에 의해 영향을 받는 변경된 유전자 발현 활성에 대한 검정)에 의해 결정될 수 있다.

[0791] **가이드 서열**

[0792] 특히 바람직한 가이드는 본원에 기재된 바와 같이 20 내지 22 ntd의 범위이고, 실시예 40을 참조하라. 일반적으로, 가이드 서열은 표적 서열과 혼성화하고, 표적 서열로의 CRISPR 복합체의 서열-특이적 결합을 유도하기에 충분한, 표적 폴리뉴클레오티드 서열과의 상보성을 갖는 임의의 폴리뉴클레오티드 서열이다. 일부 구현예에서, 가이드 서열과 그의 상응하는 표적 서열 간의 상보성의 정도는 적절한 정렬 알고리즘을 사용하여 최적으로 정렬되는 경우, 약 50%, 60%, 75%, 80%, 85%, 90%, 95%, 97.5%, 99% 이상이다. 최적의 정렬은 서열을 정렬하기에 적절한 임의의 알고리즘의 사용으로 결정될 수 있으며, 그의 비제한적인 예는 스미스-워터만(Smith-Waterman) 알고리즘, 니들만-분쉬(Needleman-Wunsch) 알고리즘, 버로우즈-휠러 트랜스폼(Burrows-Wheeler Transform)에 기초한 알고리즘(예를 들어, 버로우즈 휠러 얼라이너(Burrows Wheeler Aligner)), ClustalW, Clustal X, BLAT, 노보얼라인(Novoalign)(노보크라프트 테크놀로지즈(Novocraft Technologies)(www.novocraft.com에서 이용가능함), ELAND(일루미나(Illumina), 미국 캘리포니아주 샌 디에고), SOAP(soap.genomics.org.cn에서 이용가능) 및 Maq(maq.sourceforge.net에서 이용가능)를 포함한다. 일부 구현예에서, 가이드 서열은 약 5, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, 75 개 이상의 뉴클레오티드 길이이다. 일부 구현예에서, 가이드 서열은 약 75, 50, 45, 40, 35, 30, 25, 20, 15, 12 개 이하의 뉴클레오티드 길이이다. 표적 서열로의 CRISPR 복합체의 서열-특이적 결합을 유도하는 가이드 서열의 능력은 임의의 적절한 검정에 의해 평가될 수 있다. 예를 들어, 시험되는 가이드 서열을 포함하는 CRISPR 복합체를 형성하기에 충분한 CRISPR 시스템의 성분은 예를 들어, CRISPR 서열의 성분을 인코딩하는 벡터로의 트랜스펙션 후에, 예를 들어, 본원에 기술된 바와 같은 서베이어 검정에 의한 표적 서열 내의 우선적인 절단의 평가에 의해서와 같이, 상응하는 표적 서열을 갖는 숙주 세포로 제공될 수 있다. 유사하게, 표적 폴리뉴클레오티드 서열의 절단은 표적 서열, 시험되는 가이드 서열 및 시험 가이드 서열과 상이한 대조군 가이드 서열을 포함하는 CRISPR 복합체의 성분을 제공하고, 표적 서열에서 시험 및 대조군 가이드 서열 반응 간의 결합 또는 절단 비율을 비교함으로써 시험관에서 평가될 수 있다. 다른 검정이 가능하며, 당업자에게 떠오를 것이다.

[0793] 가이드 서열은 임의의 표적 서열을 표적화하도록 선택될 수 있다. 일부 구현예에서, 표적 서열은 세포의 게놈 내의 서열이다. 예시적인 표적 서열은 표적 게놈에서 독특한 것들을 포함한다. 예를 들어, 스트렙토코커스 피오게네스 Cas9에 대하여, 게놈 내의 독특한 표적 서열은 형태 MMMMMMMMNNNNNNNNNNNNXGG 의 Cas9 표적 부위를 포함할 수 있으며, 여기서, NNNNNNNNNNNNXGG(N은 A, G, T 또는 C이며; X는 임의의 것일 수 있음)는 게놈 내에 단일의 존재를 갖는다. 게놈 내의 독특한 표적 서열은 형태 MMMMMMMMMNNNNNNNNNNNNXGG의 스트렙토코커스 피오게네스 Cas9 표적 부위를 포함할 수 있으며, 여기서, NNNNNNNNNNXGG (N은 A, G, T 또는 C이며; X는 임의의 것일 수 있음)는 게놈 내에 단일의 존재를 갖는다. 스트렙토코커스 써모필러스 CRISPR1 Cas9에 대하여, 게놈 내의 독특한 표적 서열은 형태 MMMMMMMMNNNNNNNNNNNNXXAGAAW의 Cas9 표적 부위를 포함할 수 있으며, 여기서, NNNNNNNNNNNNXXAGAAW(N은 A, G, T 또는 C이고; X는 임의의 것일 수 있으며; W는 A 또는 T임)는 게놈 내에 단일의 존재를 갖는다. 게놈 내의 독특한 표적 서열은 형태 MMMMMMMMNNNNNNNNNNNNXXAGAAW의 스트렙토코커스 써모필러스 CRISPR1 Cas9 표적 부위를 포함할 수 있으며, 여기서, NNNNNNNNNNNXXAGAAW(N은 A, G, T 또는 C이고; X는 임의의 것일 수 있으며; W는 A 또는 T임)는 게놈 내에 단일의 존재를 갖는다. 스트렙토코커스 피오게네스 Cas9에 대하여, 게놈 내의 독특한 표적 서열은 형태 MMMMMMMMNNNNNNNNNNNNXGGXG의 Cas9 표적 부위를 포함할 수 있으며, 여기서, NNNNNNNNNNNNXGGXG(N은 A, G, T 또는 C이고; X는 임의의 것일 수 있음)는 게놈 내에 단일의 존재를 갖는다. 게놈 내의 독특한 표적 서열은 형태 MMMMMMMMMNNNNNNNNNNNXGGXG의 스트렙토코커스 피오게네스 Cas9 표적 부위를 포함할 수 있으며, 여기서, NNNNNNNNNNNXGGXG(N은 A, G, T 또는 C이고; X는 임의의 것일 수 있음)는 게놈 내에 단일의 존재를 갖는다. 이들 서열 각각에서, "M"은 A, G, T 또는 C일 수 있으며, 서열을 독특한 것으로 확인하는데 고려될 필요는 없다.

[0794] 일부 구현예에서, 가이드 서열은 가이드 서열 내의 2차 구조의 정도를 감소시키기 위해 선택된다. 일부 구현예에서, 가이드 서열의 뉴클레오티드의 약 75%, 50%, 40%, 30%, 25%, 20%, 15%, 10%, 5%, 1%, 또는 그 이하, 또는 미만이 최적으로 폴딩되었을 때 자기-상보성 염기쌍을 만드는데 참여한다. 최적 폴딩은 어떤 적합한 폴리뉴클레오티드 폴딩 알고리즘에 의해서 결정될 수 있다. 일부 프로그램은 최소 깁스 자유 에너지를 계산하는 것에 기초한다. 한 이러한 알고리즘의 예는 mFold로서, Zuker and Stiegler (Nucleic Acids Res. 9 (1981), 133-148)에서 설명되었다. 폴딩 알고리즘의 다른 예는 온라인 웹서버인 RNAfold인데, 이것은 도심 구조 예측 알고리즘을 사용하며, University of Vienna의 Institute for Theoretical Chemistry에서 개발되었다(예를 들어, 문헌[A.R. Gruber *et al.*, 2008, *Cell* 106(1): 23-24; 및 PA Carr and GM Church, 2009, *Nature Biotechnology* 27(12): 1151-62] 참조).

[0795] **tracr 메이트 서열**

[0796] 일반적으로, tracr 메이트 서열은 다음 중 하나 이상을 증진시키기에 충분한, tracr 서열과의 상보성을 갖는 임의의 서열을 포함한다: (1) 상응하는 tracr 서열을 함유하는 세포에서 tracr 메이트 서열이 측부 배치된 가이드 서열의 절제; 및 (2) 표적 서열에서의 CRISPR 복합체의 형성으로서, CRISPR 복합체가 tracr 서열에 혼성화되는 tracr 메이트 서열을 포함하는 표적 서열에서의 CRISPR 복합체의 형성. 일반적으로, 상보성의 정도는 2 개의 서열 중 더 짧은 서열의 길이에 따른 tracr 메이트 서열과 tracr 서열의 최적의 정렬을 참조한다. 최적의 정렬은 임의의 적절한 정렬 알고리즘에 의해 결정될 수 있으며, tracr 서열 또는 tracr 메이트 서열 중 어느 하나에서의 자가-상보성과 같이 2차 구조를 추가로 설명할 수 있다. 일부 구현예에서, 2개 중 보다 짧은 것의 길이를 따른 tracr 서열과 tracr 메이트 서열 간의 상보성의 정도는 최적으로 정렬되는 경우, 약 25%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 97.5%, 99% 이상이다. 일부 구현예에서, tracr 서열은 약 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 40, 50 개 이상의 뉴클레오티드 길이이다. 일부 구현예에서, tracr 서열 및 tracr 메이트 서열은 2 개 간의 혼성화가 헤어핀과 같은 2차 구조를 갖는 전사물을 생성하도록 단일의 전사물 내에 함유된다. 본 발명의 일 구현예에서, 전사물 또는 전사된 폴리뉴클레오티드 서열은 적어도 2 개 이상의 헤어핀을 갖는다. 바람직한 구현예에서, 전사물은 2, 3, 4 또는 5 개의 헤어핀을 갖는다. 본 발명의 추가의 구현예에서, 전사물은 최대 5 개의 헤어핀을 갖는다. 헤어핀 구조에서, 루프의 상류 및 마지막 "N"의 5' 서열의 부분은 tracr 메이트 서열에 상응하며, 루프의 3' 서열의 부분은 tracr 서열에 상응한다. 가이드 서열, tracr 메이트 서열 및 tracr 서열을 포함하는 단일의 폴리뉴클레오티드의 추가의 비제한적인 예는 하기와 같으며(5'에서 3'으로 표기), 여기서, "N"은 가이드 서열의 염기를 나타내고, 소문자의 제1 블록은 tracr 메이트 서열을 나타내며, 소문자의 제2 블록은 tracr 서열을 나타내고, 마지막 폴리-T 서열은 전사 종결자를 나타낸다: (1)
NNNNNNNNNNNNNNNNNNNNgttttgtactctcaagatttaGAAAtaaatcttgcagaagctacaaagataaggcttcatgccgaaatcaacaccctgtcattttatggcagggtgttttcgttatttaaTTTTTT; (2)
NNNNNNNNNNNNNNNNNNNNgttttgtactctcaGAAAtgcagaagctacaaagataaggcttcatgccgaaatcaacaccctgtcattttatggcagggtgttttcgttatttaaTTTTTT; (3)
NNNNNNNNNNNNNNNNNNNNgttttgtactctcaGAAAtgcagaagctacaaagataaggcttcatgccgaaatcaacaccctgtcattttatggcagggtgtTTTTTT; (4)
NNNNNNNNNNNNNNNNNNNNgttttagagctaGAAAtagcaagttaaaataaggctagtccgttatcaacttgaaaaagtggcaccgagtcggtgcTTTTTT; (5) NNNNNNNNNNNNNNNNNNNNgttttagagctaGAAATAGcaagttaaaataaggctagtccgttatcaacttgaaaaagtgTTTTTTT; 및 (6) NNNNNNNNNNNNNNNNNNNNgttttagagctagAAATAGcaagttaaaataaggctagtccgttatcaTTTTTTTT. 일부 구현예에서, 서열 (1) 내지 (3)은 스트렙토코커스 써모필러스 CRISPR1 유래의 Cas9와 함께 사용된다. 일부 구현예에서, 서열 (4) 내지 (6)은 스트렙토코커스 피오게네스 유래의 Cas9와 함께 사용된다. 일부 구현예에서, tracr 서열은 tracr 메이트 서열을 포함하는 전사물과 별개의 전사물이다.

[0797] **재조합 주형**

[0798] 일부 구현예에서, 재조합 주형도 또한 제공된다. 재조합 주형은 개별 벡터에 포함되거나, 개별 폴리뉴클레오티드로서 제공되는 본원에 기술된 바와 같은 다른 벡터의 성분일 수 있다. 일부 구현예에서, 재조합 주형은 상동성 재조합에서, 예를 들어, CRISPR 복합체의 일부로서 CRISPR 효소에 의해 닉이 생기거나 절단되는 표적 서열 내 또는 그 근처에서 주형으로 소용되도록 설계된다. 주형 폴리뉴클레오티드는 임의의 적절한 길이, 예를 들어, 약 10, 15, 20, 25, 50, 75, 100, 150, 200, 500, 1000 개 이상의 뉴클레오티드 길이의 것일 수 있다. 일부 구현예에서, 주형 폴리뉴클레오티드는 표적 서열을 포함하는 폴리뉴클레오티드의 부분에 상보적이다. 최적으로 정렬되는 경우, 주형 폴리뉴클레오티드는 표적 서열의 하나 이상의 뉴클레오티드(예를 들어, 약 1, 5, 10, 15, 20 개 이상의 뉴클레오티드)와 중첩할 수 있다. 일부 구현예에서, 주형 서열 및 표적 서열을 포함하는 폴리뉴클레오티드가 최적으로 정렬되는 경우, 주형 폴리뉴클레오티드의 가장 가까운 뉴클레오티드는 표적 서열로부터 약 1, 5, 10, 15, 20, 25, 50, 75, 100, 200, 300, 400, 500, 1000, 5000, 10000 개 이상의 뉴클레오티드 이내이다.

[0799] **융합 단백질**

[0800] 일부 구현예에서, CRISPR 효소는 하나 이상의 이종 단백질 도메인(예를 들어, CRISPR 효소에 더하여 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 개 이상의 도메인)을 포함하는 융합 단백질의 부분이다. CRISPR 효소 융합 단백질은 임의의 추가의 단백질 서열 및 선택적으로 임의의 2 개 도메인 사이의 링커 서열을 포함할 수 있다. CRISPR 효소에 융합될 수 있는 단백질 도메인의 예는 비제한적으로 에피토프 태그, 리포터 유전자 서열 및 하기의 활성 중 하나 이상을 갖는 단백질 도메인을 포함한다: 메틸라제 활성, 데메틸라제 활성, 전사 활성화 활성, 전사 억

제 활성, 전사 방출 인자(transcription release factor) 활성, 히스톤 변형 활성, RNA 절단 활성 및 핵산 결합 활성. 에피토프 태그의 비제한적인 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-S-트랜스퍼라제(GST), 호스라디시(horseradish) 과산화효소(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 포함하는 자가형광 단백질을 포함하나 이들에 한정되지 않는다. CRISPR 효소는 DNA 분자에 결합하거나, 말토스 결합 단백질(MBP), S-태그, Lex A DNA 결합 도메인(DBD) 융합체, GAL4 DNA 결합 도메인 융합체 및 단순 포진 바이러스(HSV) BP16 단백질 융합체를 포함하나 이들에 한정되지 않는 다른 세포 분자에 결합하는 단백질 또는 단백질의 단편을 인코딩하는 유전자 서열에 융합될 수 있다. CRISPR 효소를 포함하는 융합 단백질의 부분을 형성할 수 있는 추가의 도메인은 본원에 참조로 포함되는 US20110059502호에 기술되어 있다. 일부 구현예에서, 태그가 부착된 CRISPR 효소를 사용하여 표적 서열의 위치를 확인한다.

[0801] **유도성 시스템**

[0802] 일부 구현예에서, CRISPR 효소는 유도성 시스템의 구성성분을 형성할 수 있다. 이 시스템의 유도성 성질은 에너지 형태를 사용하여 유전자 편집이나 유전자 발현의 시공간적 제어를 허용할 것이다. 에너지 형태는, 제한은 아니지만, 전자기선, 소리 에너지, 화학 에너지 및 열 에너지를 포함할 수 있다. 유도성 시스템의 예들은 테트라사이클린 유도성 프로모터(Tet-On 또는 Tet-Off), 소 분자 2-하이브리드 전사 활성화 시스템(FKBP, ABA 등) 또는 광 유도성 시스템(피토크롬, LOV 도메인 또는 시토크롬)을 포함한다. 일 구현예에서, CRISPR 효소는 서열-특이적 방식으로 전사 활성에 변화를 지시할 수 있는 광 유도성 전사 이펙터(LITE)의 일부일 수 있다. 광의 구성성분은 CRISPR 효소, 광-반응성 시토크롬 헤테로다이머(예를 들어, *Arabidopsis thaliana*로부터), 및 전사 활성화/억제 도메인을 포함할 수 있다. 유도성 DNA 결합 단백질의 추가의 예들과 이들의 사용을 위한 방법은 US 61/736465 및 US 61/721,283에 제공되며, 이들은 그 전체가 여기 참고로 포함된다.

[0803] **전달**

[0804] 일부 양태에서, 본 발명은 하나 이상의 폴리뉴클레오티드, 예를 들어, 본원에 기술된 바와 같은 하나 이상의 벡터, 그의 하나 이상의 전사물 및/또는 그로부터 전사된 하나의 단백질 또는 단백질들을 숙주 세포에 전달하는 단계를 포함한다. 일부 양태에서, 본 발명은 이러한 방법에 의해 생성된 세포, 및 이러한 세포를 포함하거나 이로부터 생성된 동물을 추가로 제공한다. 일부 구현예에서, 가이드 서열과 조합된(선택적으로 복합체화된) CRISPR 효소는 세포로 전달된다. 통상의 바이러스 및 비-바이러스 기반의 유전자 운반 방법을 사용하여 핵산을 포유동물 세포 또는 표적 조직에 도입할 수 있다. 이러한 방법을 사용하여 CRISPR 시스템의 성분을 인코딩하는 핵산을 배양물 중의 또는 숙주 유기체 내의 세포로 투여할 수 있다. 비-바이러스 벡터 전달 시스템은 DNA 플라스미드, RNA(예를 들어, 본원에 기술된 벡터의 전사물), 네이키드(naked) 핵산 및 전달 비히클, 예를 들어, 리포좀과 복합체화된 핵산을 포함한다. 바이러스 벡터 전달 시스템은 DNA 및 RNA 바이러스를 포함하며, 이는 세포로의 전달 후에 에피솜 또는 통합된 게놈을 갖는다. 유전자 치료법 절차의 개요에 대해서는 문헌[Anderson, Science 256:808-813 (1992)]; 문헌[Nabel & Felgner, TIBTECH 11:211-217 (1993)]; 문헌[Mitani & Caskey, TIBTECH 11:162-166 (1993)]; 문헌[Dillon, TIBTECH 11:167-175 (1993)]; 문헌[Miller, Nature 357:455-460 (1992)]; 문헌[Van Brunt, Biotechnology 6(10):1149-1154 (1988)]; 문헌[Vigne, Restorative Neurology and Neuroscience 8:35-36 (1995)]; 문헌[Kremer & Perricaudet, British Medical Bulletin 51(1):31-44 (1995)]; 문헌[Haddada et al., in Current Topics in Microbiology and Immunology, Doerfler and Boehm (eds) (1995)]; 및 문헌[Yu et al., Gene Therapy 1:13-26 (1994)]을 참조한다.

[0805] 핵산의 비-바이러스 전달 방법은 리포펙션(lipofection), 미세주입, 비올리스틱스(biolistics), 비로좀(virosome), 리포좀, 면역리포좀, 다가양이온 또는 지질:핵산 컨쥬게이트, 네이키드 DNA, 인공 비리온 및 작용제-증진된 DNA의 흡수를 포함한다. 리포펙션은 예를 들어, 미국 특허 제5,049,386호, 제4,946,787호; 및 제4,897,355호에 기술되어 있으며, 리포펙션 시약은 상업적으로 시판된다(예를 들어, 트랜스펙탐(Transfectam)™ 및 리포펙틴(Lipofectin)™). 폴리뉴클레오티드의 효율적인 수용체-인식 리포펙션에 적절한 양이온 및 중성 지질은 펠그너(Felgner)의 WO 91/17424호; WO 91/16024호의 것들을 포함한다. 전달은 세포로(예를 들어, 시험관내 또는 생체외 투여) 또는 표적 조직으로(예를 들어, 생체내 투여) 이루어질 수 있다.

[0806] 표적화된 리포좀, 예를 들어, 면역지질 복합체를 포함하는 지질:핵산 복합체의 제제는 당업자에게 널리 공지되어 있다(예를 들어, 문헌[Crystal, Science 270:404-410 (1995)]; 문헌[Blaese et al., Cancer Gene Ther.

2:291-297 (1995)]; 문헌[Behr et al., Bioconjugate Chem. 5:382-389 (1994)]; 문헌[Remy et al., Bioconjugate Chem. 5:647-654 (1994)]; 문헌[Gao et al., Gene Therapy 2:710-722 (1995)]; 문헌[Ahmad et al., Cancer Res. 52:4817-4820 (1992)]; 미국 특허 제4,186,183호, 제4,217,344호, 제4,235,871호, 제4,261,975호, 제4,485,054호, 제4,501,728호, 제4,774,085호, 제4,837,028호 및 제4,946,787호 참조).

[0807] 핵산의 전달을 위한 RNA 또는 DNA 바이러스 기반의 시스템의 사용은 바이러스를 체내의 특정 세포에 표적화하고, 바이러스 페이로드(payload)를 핵에 수송하기 위한 고도로 발달된 과정을 이용한다. 바이러스 벡터를 환자에게 직접 투여하거나(생체내), 그들을 사용하여 시험관내에서 세포를 처리할 수 있으며, 변형된 세포가 선택적으로 환자에게 투여될 수 있다(생체외). 통상의 바이러스 기반의 시스템에는 유전자 운반을 위한 레트로바이러스, 렌티바이러스, 아데노바이러스, 아데노-관련 및 단순 포진 바이러스 벡터가 포함될 수 있다. 레트로바이러스, 렌티바이러스 및 아데노-관련 바이러스 유전자 운반 방법을 사용하여 숙주 게놈으로의 통합이 가능하며, 종종 삽입된 트랜스유전자의 장기간 발현을 야기한다. 또한, 높은 형질도입 효율이 많은 상이한 세포 유형 및 표적 조직에서 관찰되었다.

[0808] 레트로바이러스의 편향성은 외래 외피 단백질을 혼입시키고, 잠재적 표적 집단의 표적 세포를 증식시킴으로써 변경될 수 있다. 렌티바이러스 벡터는 비-분열 세포에 형질도입할 수 있거나, 그를 감염시킬 수 있으며, 통상적으로 높은 바이러스 역가를 생성하는 레트로바이러스 벡터이다. 따라서, 레트로바이러스 유전자 운반 시스템의 선택은 표적 조직에 따라 달라질 것이다. 레트로바이러스 벡터는 최대 6 내지 10 kb의 외래 서열에 대하여 패키징 능력을 갖는 시스-작용성 긴 말단 반복부로 이루어진다. 최소 시스-작용성 LTR은 벡터의 복제 및 패키징에 충분하며, 이는 이어서, 치료적 유전자를 표적 세포로 통합시켜, 영구적인 트랜스유전자 발현을 제공하는데 사용된다. 널리 사용되는 레트로바이러스 벡터는 쥣과 백혈병 바이러스(MuLV), 긴팔원숭이 유인원 백혈병 바이러스(GaLV), 원숭이 면역 결핍 바이러스(SIV), 인간 면역 결핍 바이러스(HIV) 및 그들의 조합에 기초한 것들을 포함한다(예를 들어, 문헌[Buchscher et al., J. Virol. 66:2731-2739 (1992)]; 문헌[Johann et al., J. Virol. 66:1635-1640 (1992)]; 문헌[Sommnerfelt et al., Virol. 176:58-59 (1990)]; 문헌[Wilson et al., J. Virol. 63:2374-2378 (1989)]; 문헌[Miller et al., J. Virol. 65:2220-2224 (1991)]; PCT/US94/05700호 참조).

[0809] 다른 구현예에서, Cocal 베지쿨로바이러스 외피보유 유사형 레트로바이러스 벡터 입자가 고려된다(예를 들어, 미국 특허 공개 제20120164118호 참조, Fred Hutchinson Cancer Research Center에 양도됨). Cocal 바이러스는 베지쿨로바이러스 속에 속하며, 포유류에서 수포성 구내염의 원인제이다. Cocal 바이러스는 원래 트리니나드에서 진드기로부터 분리되었으며 (문헌[Jonkers et al., Am. J. Vet. Res. 25:236-242(1964)]), 트리니나드, 브라질 및 아르헨티나에서 곤충, 소 및 말로부터 감염이 확인되었다. 포유류를 감염시키는 대부분의 베지쿨로바이러스는 자연적으로 감염된 절지동물로부터 분리되었는데, 이것은 이들이 매개체 감염성임을 시사한다. 베지쿨로바이러스에 대한 항체는 이 바이러스가 만연한 시골 지역에 살고 있는 사람들에게 흔하며, 실험실 획득되고, 인간 감염은 일반적으로 인플루엔자-유사 증상을 가져온다. Cocal 바이러스 외피보유 글리코단백질은 VSV-G Indiana와 아미노산 수준에서 71.5% 동일성을 공유하고, 베지쿨로바이러스의 이 외피보유 유전자의 계통발생적 비교는 Cocal 바이러스가 베지쿨로바이러스 중에서도 VSV-G Indiana 계통과 매우 밀접히 관련되지만 혈청학적으로는 다르다는 것을 나타낸다. 문헌[Jonkers et al., Am. J. Vet. Res. 25:236-242(1964) 및 Travassos da Rosa et al., Am. J. Tropical Med. & Hygiene 33:999-1006 (1984)]. Cocal 베지쿨로바이러스 외피보유 유사형 레트로바이러스 벡터 입자는, 예를 들어 렌티바이러스, 알파레트로바이러스, 베타 레트로바이러스, 감마 레트로바이러스, 델타 레트로바이러스, 및 엡실론 레트로바이러스 벡터 입자를 포함할 수 있으며, 이들은 레트로바이러스 Gag, Pol 및/또는 하나 이상의 부속 단백질(들)과 Cocal 베지쿨로바이러스 외피보유 단백질을 포함할 수 있다. 이들 구현예의 특정 양태 내에서, Gag, Pol, 및 부속 단백질들은 렌티바이러스 및/또는 감마 레트로바이러스이다.

[0810] 일시적 발현이 바람직한 출원에서, 아데노바이러스 기반의 시스템이 사용될 수 있다. 아데노바이러스 기반의 벡터는 많은 세포 유형에서 매우 높은 형질도입 효율을 가질 수 있으며, 세포 분열을 필요로 하지 않는다. 이러한 벡터를 사용하여, 높은 역가 및 발현 수준이 수득된다. 이러한 벡터는 비교적 간단한 시스템에서 대량 생성될 수 있다.

[0811] 예를 들어, 핵산 및 펩티드의 시험관내 생성에서, 그리고 생체내 및 생체외 유전자 치료법 절차를 위하여, 아데노-관련 바이러스("AAV") 벡터를 사용하여, 표적 핵산으로 세포를 형질도입시킬 수도 있다(예를 들어, 문헌[West et al., Virology 160:38-47 (1987)]; 미국 특허 제4,797,368호; WO 93/24641호; 문헌[Kotin, Human Gene Therapy 5:793-801 (1994)]; 문헌[Muzyczka, J. Clin. Invest. 94:1351 (1994)] 참조). 재조합 AAV 벡터의 작제는 미국 특허 제5,173,414호; 문헌[Tratschin et al., Mol. Cell. Biol. 5:3251-3260 (1985)]; 문헌

[Tratschin, et al., Mol. Cell. Biol. 4:2072-2081 (1984)]; 문헌[Hermonat & Muzyczka, PNAS 81:6466-6470 (1984)]; 및 문헌[Samulski et al., J. Virol. 63:03822-3828 (1989)]을 포함하는 수많은 간행물에 기술되어 있다.

[0812] 패키징 세포는 통상적으로 숙주 세포를 감염시킬 수 있는 바이러스 입자를 형성하기 위해 사용된다. 이러한 세포는 아데노바이러스를 패키징하는 293 세포, 레트로바이러스를 패키징하는 ψ2 세포 또는 PA317 세포를 포함한다. 유전자 치료법에 사용되는 바이러스 벡터는 통상적으로 생산자에 의해 핵산 벡터를 바이러스 입자로 패키징하는 세포주가 생성된다. 벡터는 통상적으로 패키징 및 이후의 숙주로의 통합에 필요한 최소 바이러스 서열을 함유하며, 다른 바이러스 서열은 발현될 폴리뉴클레오티드(들)에 대한 발현 카세트로 대체된다. 소실 바이러스 기능은 통상적으로 패키징 세포주에 의해 트랜스로 공급된다. 예를 들어, 유전자 치료법에 사용되는 AAV 벡터는 통상적으로 패키징 및 숙주 게놈으로의 통합에 필요한 AAV 게놈 유래의 ITR 서열만을 갖는다. 바이러스 DNA는 다른 AAV 유전자, 즉, rep 및 cap을 인코딩하나 ITR 서열이 결여된 헬퍼 플라스미드를 함유하는 세포주에서 패키징된다. 또한, 세포주는 헬퍼로서 아데노바이러스로 감염될 수 있다. 헬퍼 바이러스는 AAV 벡터의 복제 및 헬퍼 플라스미드로부터 AAV 유전자의 발현을 촉진시킨다. 헬퍼 플라스미드는 ITR 서열의 결여로 인해 충분한 양으로 패키징되지 않는다. 아데노바이러스로의 오염은 예를 들어, 아데노바이러스가 AAV보다 더 민감한 열 처리에 의해 감소될 수 있다.

[0813] 따라서, AAV가 형질도입 벡터로서 사용하기 위한 이상적인 후보로 고려된다. 이러한 AAV 형질도입 벡터는 트랜스로 제공된 아데노바이러스 또는 헤르페스바이러스 또는 폭스바이러스(예를 들어, 백시니아바이러스) 헬퍼 기능의 존재하에 복제할 수 있는 충분한 시스-작용 기능을 포함할 수 있다. 재조합 AAV (rAAV)는 다양한 계통의 세포에 외인성 유전자를 전달하기 위해서 사용될 수 있다. 이들 벡터에서, AAV cap 및/또는 rep 유전자가 바이러스 게놈으로부터 결실되고 선택된 DNA 세그먼트로 대체된다. 현재 AAV 벡터는 최대 4300 염기의 삽입된 DNA를 수용할 수 있다.

[0814] rAAV를 생성하기 위한 많은 방식이 있으며, 본 발명은 rAAV 및 rAAV의 제조 방법을 제공한다. 예를 들어, 원하는 바이러스 작제물을 함유하거나 필수적으로 구성된 플라스미드(들)이 AAV-감염된 세포에 트랜스펙션된다. 게다가, 제2 또는 추가의 헬퍼 플라스미드가 이들 세포에 공-트랜스펙션되며, 이로써 재조합 바이러스 작제물의 복제 및 패키징에 의무적인 AAV rep 및/또는 cap 유전자를 제공한다. 이들 조건에서, AAV의 rep 및/또는 cap 단백질은 rAAV 작제물의 복제 및 패키징을 자극하기 위하여 트랜스로 작용한다. 트랜스펙션 후 2-3일 뒤에 rAAV가 수거된다. 전통적으로, rAAV는 아데노바이러스와 함께 세포로부터 수거된다. 다음에, 오염된 아데노바이러스가 열처리에 의해서 비활성화된다. 본 발명에서, rAAV는 세포 자체로부터 수거되지 않고 세포 상청액으로부터 수거되는 것이 유익하다. 따라서, 초기 양태에서, 본 발명은 rAAV의 제조를 위해 제공되며, 전술한 것에 더하여, rAAV는 발현을 위한 DNA를 포함하는 외인성 DNA, 및 헬퍼 바이러스(예를 들어, 아데노바이러스, 헤르페스바이러스, 폭스바이러스, 예컨대 백시니아바이러스)를 함유하는 rAAV로 감수성 세포를 감염시키는 단계 (여기서 rAAV는 cap 및/또는 rep 기능을 결여하며, 헬퍼 바이러스(예를 들어, 아데노바이러스, 헤르페스바이러스, 폭스바이러스, 예컨대 백시니아바이러스)가 rAAV가 결여한 cap 및/또는 rev 기능을 제공한다); 또는 발현을 위한 DNA를 포함하는 외인성 DNA를 함유하는 rAAV로 감수성 세포를 감염시키고(여기서 재조합체는 cap 및/또는 rep 기능을 결여한다), 상기 세포를 rAAV가 결여한 cap 및/또는 rep 기능을 공급하는 플라스미드로 트랜스펙션하는 단계; 또는 발현을 위한 DNA를 포함하는 외인성 DNA를 함유하는 rAAV로 감수성 세포를 감염시키는 단계 (여기서 재조합체는 cap 및/또는 rep 기능을 결여하고, 상기 세포가 재조합체가 결여한 cap 및/또는 rep 기능을 공급한다); 또는 cap 및/또는 rep 기능을 결여한 AAV와 재조합체에 외인성 DNA를 삽입하기 위한 플라스미드로 감수성 세포를 트랜스펙션하는 단계를 포함하거나 필수적으로 구성되는 방법에 의해서 제조될 수 있으며, 이로써 외인성 DNA가 재조합에 의해서 발현되고, rep 및/또는 cap 기능을 공급함으로써 트랜스펙션의 결과 cap 및/또는 rep 기능을 결여한 발현을 위한 DNA를 포함하는 외인성 DNA를 함유하는 rAAV가 생성된다.

[0815] rAAV는 여기 설명된 AAV로부터의 것일 수 있으며, 유익하게는 AAV1, AAV2, AAV5 또는 이들의 어떤 조합을 포함할 수 있는 하이브리드 또는 캡시드를 가진 rAAV1, rAAV2, AAV5 또는 rAAV일 수 있다. rAAV에 의해서 표적화될 세포와 관련하여 rAAV들 중 AAV를 선택할 수 있는데, 예를 들어 뇌 또는 뉴런 세포를 표적화하기 위해 AAV 혈청형 1, 2, 5 또는 하이브리드 또는 캡시드 AAV1, AAV2, AAV5 또는 이들의 어떤 조합을 선택할 수 있고, 심장 조직을 표적화하기 위해 AAV4를 선택할 수 있다.

[0816] 293 세포에 더해서, 본 발명의 실시에서 사용될 수 있는 다른 세포 및 이들 세포에 대한 특정 AAV 혈청형의 시험관내 상대적 감염성은 다음과 같다(Grimm, D. et al, J. Virol. 82: 5887-5911 (2008) 참조):

세포주	AAV-1	AAV-2	AAV-3	AAV-4	AAV-5	AAV-6	AAV-8	AAV-9
Huh-7	13	100	2.5	0.0	0.1	10	0.7	0.0
HEK293	25	100	2.5	0.1	0.1	5	0.7	0.1
HeLa	3	100	2.0	0.1	6.7	1	0.2	0.1
HepG2	3	100	16.7	0.3	1.7	5	0.3	ND
Hep1A	20	100	0.2	1.0	0.1	1	0.2	0.0
911	17	100	11	0.2	0.1	17	0.1	ND
CHO	100	100	14	1.4	333	50	10	1.0
COS	33	100	33	3.3	5.0	14	2.0	0.5
MeWo	10	100	20	0.3	6.7	10	1.0	0.2
NIH3T3	10	100	2.9	2.9	0.3	10	0.3	ND
A549	14	100	20	ND	0.5	10	0.5	0.1
HT1180	20	100	10	0.1	0.3	33	0.5	0.1
단구세포	1111	100	ND	ND	125	1429	ND	ND
미성숙 DC	2500	100	ND	ND	222	2857	ND	ND
성숙 DC	2222	100	ND	ND	333	3333	ND	ND

[0817]

[0818] 본 발명은 CRISPR (군락화된 규칙적으로 산재된 짧은 팔린드롬형 반복부) 시스템을 암호화하는 외인성 핵산 분자를 함유하거나 필수적으로 구성되는 rAAV를 제공하며, 예를 들어 복수의 카세트는 프로모터, CRISPR-관련 (Cas) 단백질(뉴하수체 뉴클레아제 또는 헬리카아제 단백질), 예를 들어 Cas9를 암호화하는 핵산 분자 및 터미네이터를 포함하거나 필수적으로 구성되는 제1 카세트를 포함하고, 2 이상의, 유익하게는 벡터의 패키징 크기 제한 이하의, 예를 들어 총(제1 카세트를 포함하여) 5 개의 카세트가 프로모터, 가이드 RNA (gRNA) 및 터미네이터를 포함하거나 필수적으로 구성되고(예를 들어, 각 카세트는 프로모터-gRNA1-터미네이터, 프로모터-gRNA2-터미네이터 ... 프로모터-gRNA(N)-터미네이터로 도식적으로 표현된다(여기서 N은 삽입될 수 있는 수로서, 벡터의 패키징 크기 제한의 상한이다), 또는 2 이상의 개별 rAAV가 각각 CRISPR 시스템의 하나 이상의 카세트를 함유하는데, 예를 들어 제1 rAAV는 프로모터, Cas, 예를 들어 Cas9를 암호화하는 핵산 분자 및 터미네이터를 포함하거나 필수적으로 구성되는 제1 카세트를 함유하고, 제2 rAAV는 프로모터, 가이드 RNA (gRNA)를 암호화하는 핵산 분자 및 터미네이터를 포함하거나 필수적으로 구성되는 복수의, 4 개의 카세트를 함유한다(예를 들어, 각 카세트는 프로모터-gRNA1-터미네이터, 프로모터-gRNA2-터미네이터 ... 프로모터-gRNA(N)-터미네이터로 도식적으로 표현된다(여기서 N은 삽입될 수 있는 수로서, 벡터의 패키징 크기 제한의 상한이다). rAAV가 DNA 바이러스이기 때문에, AAV 또는 rAAV와 관련하여 여기서 논의된 핵산 분자는 DNA인 것이 유익하다. 프로모터는 일부 구현예에서 인간 시냅신 I 프로모터(hSyn)인 것이 유익하다.

[0819] 세포로의 핵산의 전달을 위한 추가의 방법은 당업자에게 공지되어 있다. 예를 들어, 본원에 참조로 포함되는 US20030087817호를 참조한다. 또한, 본원에 참조로서 포함되고 논의된 Kanasty를 참조한다.

[0820] 일부 구현예에서, 숙주 세포는 본원에 기술된 하나 이상의 벡터로 일시적으로 또는 비-일시적으로 트랜스펙션된다. 일부 구현예에서, 세포는 그것이 대상체에서 천연적으로 발생한 대로 트랜스펙션된다. 일부 구현예에서, 트랜스펙션되는 세포는 대상체로부터 취해진다. 일부 구현예에서, 세포는 세포주와 같이 대상체로부터 취해진 세포로부터 유래된다. 조직 배양을 위한 매우 다양한 세포주가 당업계에 공지되어 있다. 세포주의 예는 C8161, CCRF-CEM, MOLT, mIMCD-3, NHDF, HeLa-S3, Huh1, Huh4, Huh7, HUVEC, HASMC, HEKn, HEKa, MiaPaCell, Panc1, PC-3, TF1, CTLL-2, C1R, Rat6, CV1, RPTE, A10, T24, J82, A375, ARH-77, Calu1, SW480, SW620, SKOV3, SK-UT, CaCo2, P388D1, SEM-K2, WEHI-231, HB56, TIB55, Jurkat, J45.01, LRMB, Bcl-1, BC-3, IC21, DLD2, Raw264.7, NRK, NRK-52E, MRC5, MEF, Hep G2, HeLa B, HeLa T4, COS, COS-1, COS-6, COS-M6A, BS-C-1 원숭이 신장 상피, BALB/3T3 마우스 배아 섬유아세포, 3T3 Swiss, 3T3-L1, 132-d5 인간 태아 섬유아세포; 10.1 마우스 섬유아세포, 293-T, 3T3, 721, 9L, A2780, A2780ADR, A2780cis, A172, A20, A253, A431, A-549, ALC, B16, B35, BCP-1 세포, BEAS-2B, bEnd.3, BHK-21, BR 293, BxPC3, C3H-10T1/2, C6/36, Cal-27, CHO, CHO-7, CHO-IR, CHO-K1, CHO-K2, CHO-T, CHO Dhfr -/-, COR-L23, COR-L23/CPR, COR-L23/5010, COR-L23/R23, COS-7, COV-434, CML T1, CMT, CT26, D17, DH82, DU145, DuCaP, EL4, EM2, EM3, EMT6/AR1, EMT6/AR10.0, FM3, H1299, H69, HB54, HB55, HCA2, HEK-293, HeLa, Hepa1c1c7, HL-60, HMEC, HT-29, Jurkat, JY 세포, K562 세포, Ku812, KCL22, KG1, KYO1, LNCap, Ma-Mel 1-48, MC-38, MCF-7, MCF-10A, MDA-MB-231, MDA-MB-468, MDA-MB-

435, MDCK II, MDCK II, MOR/0.2R, MONO-MAC 6, MTD-1A, MyEnd, NCI-H69/CPR, NCI-H69/LX10, NCI-H69/LX20, NCI-H69/LX4, NIH-3T3, NALM-1, NW-145, OPCN / OPCT 세포주, Peer, PNT-1A / PNT 2, RenCa, RIN-5F, RMA/RMAS, Saos-2 세포, Sf-9, SkBr3, T2, T-47D, T84, THP1 세포주, U373, U87, U937, VCaP, Vero 세포, WM39, WT-49, X63, YAC-1, YAR 및 그의 트랜스제닉 변이형을 포함하나 이들에 한정되지 않는다. 세포주는 당업자에게 공지되어 있는 다양한 공급원으로부터 입수가능하다(예를 들어, 아메리칸 타입 컬쳐 콜렉션(American Type Culture Collection; ATCC)(미국 버지니아주 머내서스 소재) 참조). 일부 구현예에서, 본원에 기술된 하나 이상의 벡터로 트랜스펙션된 세포를 사용하여 하나 이상의 벡터-유래 서열을 포함하는 신규 세포주를 확립한다. 일부 구현예에서, 본원에 기술된 바와 같은 CRISPR 시스템의 성분이 일시적으로 트랜스펙션되고(예를 들어, 하나 이상의 벡터의 일시적 트랜스펙션 또는 RNA로의 트랜스펙션에 의해), CRISPR 복합체의 활성을 통해 변형된 세포를 사용하여, 변형을 포함하나 임의의 다른 외인성 서열이 결여된 세포를 포함하는 신규 세포주를 확립한다. 일부 구현예에서, 일시적으로 또는 비-일시적으로 본원에 기술된 하나 이상의 벡터로 트랜스펙션된 세포 또는 이러한 세포로부터 유래된 세포주가 하나 이상의 시험 화합물의 평가에서 사용된다.

[0821] 일부 구현예에서, 본원에 기술된 하나 이상의 벡터를 사용하여 비-인간 트랜스제닉 동물 또는 트랜스제닉 식물을 생성한다. 일부 구현예에서, 트랜스제닉 동물은 포유동물, 예를 들어, 마우스, 랫트 또는 토끼이다. 트랜스제닉 동물 및 식물의 생성 방법은 당업계에 공지되어 있으며, 일반적으로 본원에 기술된 바와 같은 세포 트랜스펙션 방법으로 시작한다.

[0822] 다른 구현예에서, 니들 어레이를 가진 유체 전달 장치(예를 들어, Fred Hutchinson Cancer Research Center에 양도된 미국 특허 공개 제20110230839 참조)가 CRISPR Cas의 고체 조직으로의 전달을 위해 고찰될 수 있다. 고체 조직에 유체의 전달을 위한 미국 특허 공개 제20110230839호의 장치는 어레이로 배열된 복수의 니들; 복수의 니들의 각자와 각각 유체 연통되는 복수의 저장소; 및 복수의 저장소의 각자와 작동 가능하게 연결되고 저장소 내의 유압을 제어하도록 구성된 복수의 작동장치를 포함할 수 있다. 특정 구현예에서, 복수의 작동장치의 각각은 복수의 플런저 중 하나를 포함할 수 있고, 복수의 플런저의 각각의 제1 단부가 복수의 저장소의 각자에 수용되어 있으며, 특정한 추가 구현예에서, 복수의 플런저의 플런저들은 동시에 압축될 수 있도록 각각의 제2 단부에서 함께 작동 가능하게 연결된다. 특정한 다른 추가 구현예는 선택적으로 가변적인 속도로 복수의 플런저를 전부 압축하도록 구성된 플런저 구동장치를 포함할 수 있다. 다른 구현예에서, 복수의 작동장치의 각각은 제1 및 제2 단부를 가진 복수의 유체 이송선 중 하나를 포함할 수 있고, 복수의 유체 이송선의 각각의 제1 단부가 복수의 저장소의 각자와 연결되어 있다. 다른 구현예에서, 장치는 유압원을 포함할 수 있고, 복수의 작동장치의 각각은 유압원과 복수의 저장소의 각자를 연결하는 유체를 포함한다. 추가 구현예에서, 유압원은 컴프레서, 진공 축적장치, 연동 펌프, 마스터 실린더, 미소유체 펌프, 및 밸브 중 적어도 하나를 포함할 수 있다. 다른 구현예에서, 복수의 니들의 각각은 길이를 따라 분포된 복수의 포트를 포함할 수 있다.

[0823] **표적 변형**

[0824] 일 양태에서, 본 발명은 진핵 세포에서 표적 폴리뉴클레오티드를 변형하는 방법을 제공하며, 이것은 생체내, 생체외 또는 시험관내 방식으로 이루어질 수 있다. 일부 구현예에서, 방법은 인간 또는 비인간 동물, 또는 식물로부터 세포 또는 세포 집단을 샘플링하는 단계, 및 세포 또는 세포들을 변형하는 단계를 포함한다. 배양은 생체외 방식으로 어떤 단계에서 일어날 수 있다. 세포 또는 세포들은 심지어 비인간 동물 또는 식물에 재도입될 수 있다. 재도입된 세포의 경우, 세포는 줄기세포인 것이 특히 바람직하다.

[0825] 일부 구현예에서, 방법은 상기 표적 폴리뉴클레오티드의 절단을 행하기 위하여 CRISPR 복합체를 표적 폴리뉴클레오티드와 결합시키는 단계를 포함하며, 이로써 표적 폴리뉴클레오티드가 변형되고, 여기서 CRISPR 복합체는 상기 표적 폴리뉴클레오티드 내의 표적 서열과 혼성화된 가이드 서열과 복합체화된 CRISPR 효소를 포함하며, 상기 가이드 서열은 tracr 메이트 서열과 결합되고, 이어서 tracr 서열과 혼성화된다.

[0826] 일 양태에서, 본 발명은 진핵 세포에서 폴리뉴클레오티드의 발현을 변형하는 방법을 제공한다. 일부 구현예에서, 방법은 결합이 상기 폴리뉴클레오티드의 증가된 또는 감소된 발현을 가져오도록 CRISPR 복합체를 폴리뉴클레오티드와 결합시키는 단계를 포함하며, 여기서 CRISPR 복합체는 상기 폴리뉴클레오티드 내의 표적 서열과 혼성화된 가이드 서열과 복합체화된 CRISPR 효소를 포함하고, 상기 가이드 서열은 tracr 메이트 서열과 결합되고, 이어서 tracr 서열과 혼성화된다. 유사한 고려사항 및 조건이 표적 폴리뉴클레오티드를 변형하는 방법에 대해 상기 설명된 것과 같이 적용된다. 실제로 이들 샘플링, 배양 및 재도입 옵션은 본 발명의 전 양태에 적용된다.

[0827] 실제로, 본 발명의 어떤 양태에서, CRISPR 복합체는 표적 서열과 혼성화된 가이드 서열과 복합체화된 CRISPR 효

소를 포함하며, 여기서 상기 가이드 서열은 tracr 메이트 서열과 결합되고, 이어서 tracr 서열과 혼성화될 수 있다. 유사한 고려사항 및 조건이 표적 폴리뉴클레오티드를 변형하는 방법에 대해 상기 설명된 것과 같이 적용된다.

[0828] **키트**

[0829] 일 양태에서, 본 발명은 상기 방법 및 조성물에서 개시된 요소들 중 어느 하나 이상을 함유하는 키트를 제공한다. 요소들은 개별적으로 또는 조합하여 제공될 수 있으며, 어떤 적합한 용기, 예컨대 바이알, 보틀 또는 튜브에 제공될 수 있다. 일부 구현예에서, 키트는 하나 이상의 언어로 된, 예를 들어 둘 이상의 언어로 된 설명서를 포함한다.

[0830] 일부 구현예에서, 키트는 본원에 기술된 요소 중 하나 이상을 사용하는 과정에 사용하기 위한 하나 이상의 시약을 포함한다. 시약은 임의의 적절한 용기에 제공될 수 있다. 예를 들어, 키트는 하나 이상의 반응 또는 저장 완충제를 제공할 수 있다. 시약은 특정 검정에 사용가능한 형태 또는 사용 전에 하나 이상의 다른 성분의 첨가를 필요로 하는 형태(예를 들어, 농축물 또는 동결건조 형태)로 제공될 수 있다. 완충제는 탄산나트륨 완충제, 중탄산나트륨 완충제, 붕산염 완충제, 트리스(Tris) 완충제, MOPS 완충제, HEPES 완충제 및 그들의 조합을 포함하나 이들에 한정되지 않는 임의의 완충제일 수 있다. 일부 구현예에서, 완충제는 알칼리성이다. 일부 구현예에서, 완충제는 약 7 내지 약 10의 pH를 갖는다. 일부 구현예에서, 키트는 가이드 서열과 조절 요소를 작동 가능하게 연결하도록, 벡터에 삽입하기 위한 가이드 서열에 상응하는 하나 이상의 올리고뉴클레오티드를 포함한다. 일부 구현예에서, 키트는 상동성 재조합 주형 폴리뉴클레오티드를 포함한다. 일부 구현예에서, 키트는 여기 설명된 벡터 중 하나 이상 및/또는 폴리뉴클레오티드 중 하나 이상을 포함한다. 키트는 본 발명의 시스템의 모든 요소를 제공하는 것이 유익할 수 있다.

[0831] **CRISPR 복합체**

[0832] 일 양태에서, 본 발명은 CRISPR 시스템의 하나 이상의 요소의 사용 방법을 제공한다. 본 발명의 CRISPR 복합체는 효율적인 표적 폴리뉴클레오티드의 변형 수단을 제공한다. 본 발명의 CRISPR 복합체는 다수의 세포 유형에서 표적 폴리뉴클레오티드를 변형시키는(예를 들어, 결실시키는, 삽입하는, 전위시키는, 불활성화시키는, 활성화시키는) 것을 포함하는 매우 다양한 유용성을 갖는다. 이와 같이, 본 발명의 CRISPR 복합체는 예를 들어, 유전자 치료법, 약물 스크리닝, 질병 진단 및 예후에서 넓은 스펙트럼의 응용을 갖는다. 예시적인 CRISPR 복합체는 표적 폴리뉴클레오티드 내의 표적 서열에 혼성화되는 가이드 서열과 복합체화되는 CRISPR 효소를 포함한다. 가이드 서열은 tracr 메이트 서열에 연결되며, tracr 메이트 서열은 차례로 tracr 서열에 혼성화된다.

[0833] 일 구현예에서, 본 발명은 표적 뉴클레오티드를 절단하는 방법을 제공한다. 방법은 표적 폴리뉴클레오티드와 결합하여 상기 표적 폴리뉴클레오티드의 절단을 행하는 CRISPR 복합체를 사용하여 표적 폴리뉴클레오티드를 변형하는 단계를 포함한다. 통상적으로, 본 발명은 CRISPR 복합체는 세포에 도입되었을 때 게놈 서열에 브레이크(예를 들어, 단일가닥 또는 이중가닥 브레이크)를 생성한다. 예를 들어, 방법은 세포에서 질환 유전자를 절단하기 위하여 사용될 수 있다.

[0834] CRISPR 복합체에 의해서 생성된 브레이크는 에러 경향 비-상동성 단부 연결 (NHEJ) 경로 또는 고 충실도 상동성-지정 수선 (HDR)과 같은 수선 과정에 의해서 수선될 수 있다(도 29). 이들 수선 과정 동안, 외인성 폴리뉴클레오티드 주형이 게놈 서열에 도입될 수 있다. 일부 방법에서, HDR 과정은 게놈 서열을 변형하기 위해서 사용된다. 예를 들어, 상류 서열과 하류 서열이 측접된 통합될 서열을 포함하는 외인성 폴리뉴클레오티드 주형이 세포에 도입된다. 상류 서열과 하류 서열은 염색체에의 통합 부위의 어느 사이드와 서열 유사성을 공유한다.

[0835] 바람직한 경우, 도너 폴리뉴클레오티드는 DNA, 예를 들어 DNA 플라스미드, 박테리아 인공 염색체 (BAC), 효모 인공 염색체 (YAC), 바이러스 벡터, DNA의 선형 조각, PCR 단편, 네이키드 핵산, 또는 리포솜 또는 폴록사머와 같은 전달 부형제와 복합체화된 핵산일 수 있다.

[0836] 외인성 폴리뉴클레오티드 주형은 통합될 서열을 포함한다(예를 들어, 돌연변이된 유전자). 통합 서열은 세포에 대해 내인성 또는 외인성인 서열일 수 있다. 통합될 서열의 예들은 단백질 또는 비-코딩 RNA(예를 들어, 마이크로RNA)를 암호화하는 폴리뉴클레오티드를 포함한다. 따라서, 통합 서열은 적절한 제어 서열 또는 서열들에 작동 가능하게 연결될 수 있다. 또는 달리, 통합될 서열은 조절 기능을 제공할 수 있다.

[0837] 외인성 폴리뉴클레오티드 주형에서 상류 및 하류 서열은 관심대상의 염색체 서열과 도너 폴리뉴클레오티드 사이의 재조합을 촉진하도록 선택된다. 상류 서열은 통합을 위해 표적화된 부위의 상류 염색체 서열과 서열 유사성을 공유하는 핵산 서열이다. 유사하게, 하류 서열은 통합 표적화된 부위의 하류 염색체 서열과 서열 유사성을

공유하는 핵산 서열이다. 외인성 폴리뉴클레오티드 주형의 상류 및 하류 서열은 표적화된 게놈 서열과 75%, 80%, 85%, 90%, 95%, 또는 100% 서열 동일성을 가질 수 있다. 바람직하게, 외인성 폴리뉴클레오티드 주형의 상류 및 하류 서열은 표적화된 게놈 서열과 약 95%, 96%, 97%, 98%, 99%, 또는 100% 서열 동일성을 가진다. 일부 방법에서, 외인성 폴리뉴클레오티드 주형의 상류 및 하류 서열은 표적화된 게놈 서열과 약 99% 또는 100% 서열 동일성을 가진다.

[0838] 상류 또는 하류 서열은 약 20bp 내지 약 2500bp, 예를 들어, 약 50, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000, 2100, 2200, 2300, 2400, 또는 2500 bp를 포함할 수 있다. 일부 방법에서, 예시적인 상류 또는 하류 서열은 약 200 bp 내지 약 2000 bp, 약 600 bp 내지 약 1000 bp, 또는 더 구체적으로 약 700 bp 내지 약 1000 bp를 가진다.

[0839] 일부 방법에서, 외인성 폴리뉴클레오티드 주형은 마커를 더 포함할 수 있다. 이러한 마커는 표적화된 통합의 스크리닝을 용이하게 할 수 있다. 적합한 마커의 예들은 제한 부위, 형광 단백질, 또는 선택성 마커를 포함한다. 본 발명의 외인성 폴리뉴클레오티드 주형은 재조합 기술을 사용하여 구성될 수 있다(예를 들어, 문헌[Sambrook et al., 2001 및 Ausubel et al., 1996] 참조).

[0840] 외인성 폴리뉴클레오티드 주형을 통합함으로써 표적 폴리뉴클레오티드를 변형하는 예시적인 방법에서, 이중가닥 브레이크가 CRISPR 복합체에 의해서 게놈 서열에 도입되며, 이 브레이크는 주형이 게놈에 통합되도록 외인성 폴리뉴클레오티드 주형의 상동성 재조합을 통해서 수선된다. 이중가닥 브레이크의 존재는 주형의 통합을 촉진한다.

[0841] 다른 구현예에서, 본 발명은 진핵 세포에서 폴리뉴클레오티드의 발현을 변형하는 방법을 제공한다. 방법은 폴리뉴클레오티드와 결합하는 CRISPR 복합체를 사용하여 표적 폴리뉴클레오티드의 발현을 증가시키거나 감소시키는 단계를 포함한다.

[0842] 일부 방법에서, 표적 폴리뉴클레오티드는 세포에서 발현의 변형을 행하는데 비활성화될 수 있다. 예를 들어, 세포에서 CRISPR 복합체와 표적 서열의 결합시, 표적 폴리뉴클레오티드가 비활성화되며, 이로써 서열이 번역되지 않거나, 코딩된 단백질이 생성되지 않거나, 또는 서열이 야생형 서열처럼 기능하지 않는다. 예를 들어, 단백질 또는 마이크로RNA 암호화 서열은 단백질이 생성되지 않도록 비활성화될 수 있다.

[0843] 일부 방법에서, 제어 서열이 더 이상 조절 서열로서 기능하지 않도록 비활성화될 수 있다. 본원에서 사용된 "제어 서열"은 전사, 번역 또는 핵산 서열의 접근성을 행하는 어떤 핵산 서열을 말한다. 제어 서열의 예들은 프로모터, 전사 터미네이터, 및 인핸서를 포함하며 이들이 제어 서열이다.

[0844] 비활성화된 표적 서열은 결실 돌연변이 (즉, 하나 이상의 뉴클레오티드의 결실), 삽입 돌연변이 (즉, 하나 이상의 뉴클레오티드의 삽입), 또는 논센스 돌연변이 (즉, 중단 코돈이 도입되도록 단일 뉴클레오티드의 다른 뉴클레오티드로의 치환)를 포함할 수 있다. 일부 방법에서, 표적 서열의 비활성화는 표적 서열의 "녹아웃"을 가져온다.

[0845] **질환 모델**

[0846] 본 발명의 방법은 질환 모델로서 사용될 수 있는 식물, 동물 또는 세포를 생성하기 위해서 사용될 수 있다. 본원에서 사용된 "질환"은 대상체에서의 질환, 장애 또는 징후를 말한다. 예를 들어, 본 발명의 방법은 질환과 관련된 하나 이상의 핵산 서열에 변형을 포함하는 동물이나 세포, 또는 질환과 관련된 하나 이상의 핵산 서열의 발현이 변경된 식물, 동물 또는 세포를 생성하기 위해서 사용될 수 있다. 이러한 핵산 서열은 질환 관련 단백질 서열을 암호화할 수 있거나, 또는 질환 관련 대조군 서열일 수 있다. 따라서, 본 발명의 구현예에서, 식물, 대상체, 환자, 유기체 또는 세포는 비인간 대상체, 환자, 유기체 또는 세포일 수 있다는 것이 이해된다. 따라서, 본 발명은 본 방법에 의해서 생성된 식물, 동물 또는 세포, 또는 그것의 자손을 제공한다. 자손은 생성된 동물의 클론일 수 있거나, 또는 바람직한 특성이 자손으로 유전자 이입되도록 동일한 종의 다른 개체와의 교배에 의한 유성생식의 결과일 수 있다. 세포는 다세포 유기체, 특히 동물 또는 식물의 경우 생체내 또는 생체외 방식일 수 있다. 세포가 배양중인 경우, 적절한 배양 조건이 충족된다면, 바람직하게 세포가 이 목적을 위해 적절히 개조된다면 세포주가 확립될 수 있다(예를 들어, 줄기세포). 본 발명에 의해 생성된 박테리아 세포주가 또한 고려된다. 따라서, 세포주가 또한 고려된다.

[0847] 일부 방법에서, 질환 모델은 동물 또는 세포에서 돌연변이의 효과 및 질환 연구에 통상 사용되는 척도를 사용하여 질환의 발생 및/또는 진행을 연구하기 위해서 사용될 수 있다. 또는 달리, 이러한 질환 모델은 질환에 대한

제약학적으로 활성인 화합물의 효과를 연구하는데 유용하다.

[0848] 일부 방법에서, 질환 모델은 잠재적 유전자 요법 전략의 효율을 평가하기 위해서 사용될 수 있다. 즉, 질환-관련 유전자 또는 폴리뉴클레오티드는 질환 발생 및/또는 진행이 억제되거나 감소되도록 변형될 수 있다. 특히, 방법은 변경된 단백질이 생성되고, 결과적으로 동물 또는 세포가 변경된 반응을 가지도록 질환-관련 유전자 또는 폴리뉴클레오티드를 변형하는 단계를 포함한다. 따라서, 일부 방법에서, 유전자 변형된 동물이 질환 발생 소인이 있는 동물과 비교될 수 있으며, 이로써 유전자 요법 사건의 효과가 평가될 수 있다.

[0849] 다른 구현예에서, 본 발명은 질환 유전자와 관련된 세포 신호화 사건을 조절하는 생물학적 활성제를 개발하는 방법을 제공한다. 방법은 시험 화합물을 CRISPR 효소, tracr 메이트 서열에 연결된 가이드 서열 및 tracr 서열 중 하나 이상의 발현을 유도하는 하나 이상의 벡터를 포함하는 세포와 접촉시키는 단계; 및 예를 들어 세포에 함유된 질환 유전자의 돌연변이와 관련된 세포 신호화 사건의 감소 또는 증강의 표시인 판독결과의 변화를 검출하는 단계를 포함한다.

[0850] 본원에 기재된 바와 같은 오가노이드 또는 세포 콜렉션을 포함하는 세포 모델, 또는 동물 모델은 세포 기능 변화를 스크리닝하기 위하여 본 발명의 방법과 조합하여 구성될 수 있다. 이러한 모델은 관심대상의 세포 기능에 대한 본 발명의 CRISPR 복합체에 의해서 변형된 게놈 서열의 효과를 연구하기 위해서 사용될 수 있다. 예를 들어, 세포 기능 모델은 세포내 신호화 또는 세포외 신호화에 대한 변형된 게놈 서열의 효과를 연구하기 위해서 사용될 수 있다. 또는 달리, 세포 기능 모델은 감각 인식에 대한 변형된 게놈 서열의 효과를 연구하기 위해서 사용될 수 있다. 일부 이러한 모델에서, 모델에서의 신호화 생화학적 경로와 관련된 하나 이상의 게놈 서열이 변형된다.

[0851] 몇 가지 질환 모델이 구체적으로 조사되었다. 이들은 새로운 자폐 위험 유전자 CHD8, KATNAL2, 및 SCN2A; 및 증후군성 자폐(앤젤만 증후군) 유전자 UBE3A를 포함한다. 이들 유전자와 결과의 자폐 모델이 물론 바람직하지만, 이들은 유전자 및 상응하는 모델 전체적으로 본 발명의 광범한 적용성을 나타내기 위해서 사용된다.

[0852] 신호화 생화학적 경로와 관련된 하나 이상의 게놈 서열의 변경된 발현은 후보 제제와 접촉되었을 때 시험 모델 세포와 대조군 세포 사이에 상응하는 유전자의 mRNA 수준의 차이를 분석함으로써 결정될 수 있다. 또는 달리, 신호화 생화학적 경로와 관련된 서열들의 차등적 발현이 암호화된 폴리펩티드 또는 유전자 생성물의 수준 차이를 검출함으로써 결정된다.

[0853] mRNA 전사체 또는 상응하는 폴리뉴클레오티드의 수준에서 제제-유도 변경을 분석하기 위하여, 샘플에 함유된 핵산이 먼저 본 분야의 표준 방법에 따라서 추출된다. 예를 들어, mRNA는 문헌[Sambrook et al. (1989)]에 제시된 과정에 따라서 다양한 세포용해 효소 또는 화학 용액을 사용하여 분리될 수 있거나, 또는 제조자에 의해서 제공된 첨부된 설명서에 따라서 핵산-결합 수지에 의해서 추출될 수 있다. 다음에, 추출된 핵산 샘플에 함유된 mRNA가 본 분야에 널리 공지된 방법에 따라서 또는 여기 예시된 방법에 기초하여 증폭 과정 또는 종래의 혼성화 분석(예를 들어, 노던 블롯 분석)에 의해서 검출된다.

[0854] 본 발명의 목적에 있어서, 증폭 수단은 합당한 충실도로 표적 서열을 복제할 수 있는 프라이머와 폴리머라제를 이용하는 어떤 방법을 의미한다. 증폭은 천연 또는 재조합 DNA 폴리머라제, 예컨대 TaqGold™, T7 DNA 폴리머라제, 이콜라이 DNA 폴리머라제의 Klenow 단편, 및 역전사효소에 의해서 수행될 수 있다. 바람직한 증폭 방법은 PCR이다. 특히, 분리된 RNA는 정량적 폴리머라제 연쇄 반응(RT-PCR)과 결합된 역전사 분석을 거칠 수 있고, 이로써 신호화 생화학적 경로와 관련된 서열의 발현 수준을 정량할 수 있다.

[0855] 유전자 발현 수준의 검출은 증폭 분석으로 실시한 수행될 수 있다. 일 양태에서, 증폭된 생성물은, 제한은 아니지만 DNA 인터칼레이터 및 DNA 그로브 바인더를 포함하는 형광 DNA-결합제로 직접 시각화될 수 있다. 이중가닥 DNA 분자에 통합된 인터칼레이터의 양은 통상적으로 증폭된 DNA 생성물의 양에 비례하기 때문에, 본 분야의 종래의 광학 시스템을 사용하여 인터칼레이트된 염료의 형광을 정량함으로써 증폭된 생성물의 양을 편리하게 결정할 수 있다. 본 출원에 적합한 DNA-결합 염료는 SYBR 그린, SYBR 블루, DAPI, 프로피듐 요오드, Hoeste, SYBR 골드, 에티듐 브로마이드, 아크리딘, 프로플라빈, 아크리딘 오렌지, 아크리플라빈, 플루오르쿠마닌, 엘립티신, 다우노마이신, 클로로퀸, 디스타마이신 D, 크로모마이신, 호미듐, 미스라마이신, 루테늄 폴리피리딜, 안트라마이신 등을 포함한다.

[0856] 다른 양태에서, 서열 특이적 프로브와 같은 다른 형광 표지가 증폭된 생성물의 검출 및 정량을 촉진하기 위해서 증폭 반응에 이용될 수 있다. 프로브-기반 정량 증폭은 원하는 증폭된 생성물의 서열-특이적 검출에 의존한다. 그것은 형광, 표적-특이적 프로브(예를 들어, TaqMan® 프로브)를 이용하며, 그 결과 특이성 및 감도가 증가된

다. 프로브-기반 정량 증폭을 수행하는 방법은 본 분야에 잘 확립되어 있으며, 미국특허 No. 5,210,015에 교시된다.

[0857] 또 다른 양태에서, 신호화 생화학적 경로와 관련된 서열과 서열 상동성을 공유하는 혼성화 프로브를 사용한 종래의 혼성화 분석이 수행될 수 있다. 통상적으로, 프로브는 혼성화 반응에서 시험 대상체로부터 유래된 생물학적 샘플에 함유된 신호화 생화학적 경로와 관련된 서열과 안정한 복합체를 형성하게 된다. 안티센스가 프로브 핵산으로 사용된 경우, 샘플에 제공되는 표적 폴리뉴클레오티드는 안티센스 핵산의 서열에 상보성이도록 선택된다는 것이 당업자에게 인정될 것이다. 반대로, 뉴클레오티드 프로브가 센스 핵산인 경우, 표적 폴리뉴클레오티드는 센스 핵산에 상보성이도록 선택된다.

[0858] 혼성화는 다양한 긴축도 조건에서 수행될 수 있다. 본 발명의 실시를 위한 적합한 혼성화 조건은 신호화 생화학적 경로와 관련된 서열과 프로브 사이의 인식 상호작용이 충분히 특이적이며 충분히 안정하게 되는 조건이다. 혼성화 반응의 긴축도를 증가시키는 조건은 본 분야에 널리 알려져 있으며 공개되어 있다(예를 들어, 문헌 [Sambrook, et al., (1989); Nonradioactive In Situ Hybridization Application Manual, Boehringer Mannheim, second edition] 참조). 혼성화 분석은, 제한은 아니지만, 니트로셀룰로스, 유리, 규소 및 다양한 유전자 어레이를 포함하는 어떤 고체 지지체 상에 고정된 프로브를 사용하여 형성될 수 있다. 바람직한 혼성화 분석은 미국특허 No. 5,445,934에 설명된 고밀도 유전자 칩에서 수행된다.

[0859] 혼성화 분석 동안 형성된 프로브-표적 복합체의 편리한 검출을 위하여, 뉴클레오티드 프로브는 검출가능한 표지와 콘쥬게이트된다. 본 발명에서 사용하기에 적합한 검출가능한 표지는 광화학, 생화학, 분광학, 면역화학, 전기, 광학 또는 화학적 수단에 의해서 검출될 수 있는 어떤 조성물을 포함한다. 광범한 다양한 적합한 검출가능한 표지가 본 분야에 알려져 있으며, 이들은 형광 또는 화학발광 표지, 방사성활성 동위원소 표지, 효소 또는 다른 리간드들을 포함한다. 바람직한 구현예에서, 당업자는 형광 표지 또는 효소 태그, 예컨대 디옥시게닌, β-갈락토시다제, 유레아제, 알칼리성 포스파타제 또는 퍼옥시다제, 아비딘/바이오틴 복합체를 이용하는 것을 주로 원할 것이다.

[0860] 혼성화 강도를 검출하거나 정량하기 위해서 사용되는 검출 방법은 통상적으로 상기 선택된 표지에 따를 것이다. 예를 들어, 방사성표지는 사진 필름 또는 포스포이미저를 사용하여 검출될 수 있다. 형광 마커는 방출된 광을 검출할 수 있는 포토디텍터를 사용하여 검출 및 정량될 수 있다. 효소 표지는 통상적으로 기질에 효소를 제공하고, 기질 에 대한 효소의 작용에 의해서 생성된 반응 생성물을 측정함으로써 검출되며; 마지막으로, 비색 표지는 착색된 표지를 간단히 시각화함으로써 검출된다.

[0861] 신호화 생화학적 경로와 관련된 서열의 발현에서 제제-유도 변화는 또한 상응하는 유전자 생성물을 검사함으로써 결정될 수 있다. 단백질 수준의 결정은 통상적으로 (a) 생물학적 샘플에 함유된 단백질을 신호화 생화학적 경로와 관련된 단백질과 특이적으로 결합하는 제제와 접촉시키는 단계; 및 (b) 이렇게 형성된 어떤 제제:단백질 복합체를 확인하는 단계를 수반한다. 이 구현예의 일 양태에서, 신호화 생화학적 경로와 관련된 단백질과 특이적으로 결합하는 제제는 항체, 바람직하게 단클론 항체이다.

[0862] 반응은 제제를, 신호화 생화학적 경로와 관련된 단백질과 제제 사이에 복합체가 형성될 수 있는 조건에서 시험 샘플로부터 유래된 신호화 생화학적 경로와 관련된 단백질의 샘플과 접촉시킴으로써 수행된다. 복합체의 형성은 본 분야의 표준 과정에 따라서 직접적으로 또는 간접적으로 검출될 수 있다. 직접 검출 방법에서, 제제는 검출가능한 표지와 함께 공급되고, 반응되지 않는 제제는 복합체로부터 제거될 수 있다. 이로써 남은 표지의 양이 형성된 복합체의 양을 나타낸다. 이러한 방법의 경우, 긴축 세척 조건 동안에도 제제에 부착된 상태를 유지하는 표지를 선택하는 것이 바람직하다. 표지는 결합 반응을 방해하지 않는 것이 바람직하다. 또는 대안적으로, 간접 검출 과정은 화학적으로 또는 효소적으로 도입된 표지를 함유하는 제제를 사용할 수 있다. 바람직한 표지는 일반적으로 결과의 제제:폴리펩티드 복합체의 결합 또는 안정성을 방해하지 않는다. 그러나, 표지는 통상적으로 유효한 결합을 위해 항체에 접근하여 검출가능한 신호를 생성하도록 설계된다.

[0863] 단백질 수준을 검출하는데 적합한 광범위한 다양한 표지가 본 분야에 알려져 있다. 비제한적 예들은 방사성동위원소, 효소, 콜로이드 금속, 형광 화합물, 생물발광 화합물, 및 화학발광 화합물을 포함한다.

[0864] 결합 반응 동안 형성된 제제:폴리펩티드 복합체의 양은 표준 정량 분석에 의해서 정량될 수 있다. 상기 예시된 대로, 제제:폴리펩티드 복합체의 형성은 결합 부위에 남은 표지의 양에 의해서 직접 측정될 수 있다. 또는 달리, 신호화 생화학적 경로와 관련된 단백질이 특정 제제에서 결합 부위에 대해 표지된 유사체와 경쟁하는 능력에 대해 시험된다. 이 경쟁 분석에서, 포착된 표지의 양은 시험 샘플에 존재하는 신호화 생화학적 경로와 관

련된 단백질 서열의 양에 역비례한다.

[0865] 상기 개략된 일반적인 원리에 기초한 단백질 분석을 위한 다수의 기술이 본 분야에서 이용될 수 있다. 이들은 제한은 아니지만 방사성면역분석, ELISA(효소 결합 면역방사성측정 분석), "샌드위치" 면역분석, 면역방사성측정 분석, 인시튜 면역분석(예를 들어, 콜로이드 금, 효소 또는 방사성동위원소 표지를 사용하는), 웨스턴 블롯 분석, 면역침전 분석, 면역형광 분석, 및 SDS-PAGE를 포함한다.

[0866] 신호화 생화학적 경로와 관련된 단백질을 특이적으로 인식하거나 그것과 결합하는 항체가 상기 언급된 단백질 분석을 수행하는데 바람직하다. 바람직한 경우, 특정 타입의 번역 후 변형(예를 들어, 신호화 생화학적 경로 유도성 변형)을 인식하는 항체가 사용될 수 있다. 번역 후 변형은, 제한은 아니지만 글리코실화, 지질화, 아세틸화 및 포스포릴화를 포함한다. 이들 항체는 상업적 판매자로부터 구입될 수 있다. 예를 들어, 티로신-포스포릴화된 단백질을 특이적으로 인식하는 항-포스포티로신 항체는 Invitrogen 및 Perkin Elmer를 포함하는 다수의 판매자로부터 입수할 수 있다. 항-포스포티로신 항체는 ER 스트레스에 반응하여 티로신 잔기에서 차등적으로 포스포릴화된 단백질들을 검출하는데 특히 유용하다. 이러한 단백질은, 제한은 아니지만 진핵생물 번역 개시 인자 2 알파(eIF-2α)를 포함한다. 또는 달리, 이들 항체는 종래의 다클론 또는 단클론 항체 기술을 사용하여 숙주 동물 또는 항체-생성 세포를 원하는 번역 후 변형을 나타내는 표적 단백질로 면역화함으로써 생성될 수 있다.

[0867] 본 방법을 실시하는데 있어서, 상이한 체조직, 상이한 세포 타입 및/또는 상이한 아세포 구조에서 신호화 생화학적 경로와 관련된 단백질의 발현 패턴을 조사하는 것이 바람직할 수 있다. 이들 연구는 특정 조직, 세포 타입 또는 아세포 구조에서 우선적으로 발현되는 단백질 마커와 결합할 수 있는 조직-특이적, 세포-특이적 또는 아세포 구조 특이적 항체의 사용에 의해서 수행될 수 있다.

[0868] 신호화 생화학적 경로와 관련된 유전자의 변경된 발현은 또한 대조군 세포에 비해서 유전자 생성물의 활성의 변화를 검사함으로써 결정될 수 있다. 신호화 생화학적 경로와 관련된 단백질의 활성에서 제제-유도 변화의 분석은 조사중인 신호전달 경로 및/또는 생물학적 활성에 따를 것이다. 예를 들어, 단백질이 키나제인 경우, 하류 기질(들)을 포스포릴화하는 능력의 변화가 본 분야에 공지된 다양한 분석에 의해서 결정될 수 있다. 대표적인 분석은, 제한은 아니지만 포스포릴화된 단백질을 인식하는 항-포스포티로신 항체와 같은 항체를 사용한 면역블롯팅 및 면역침전을 포함한다. 게다가, 키나제 활성이 AlphaScreen™(Perkin Elmer에서 입수가능) 및 eTag™ 분석(Chan-Hui, et al. (2003) Clinical Immunology 111: 162-174)과 같은 고-처리량 화학발광 분석에 의해서 검출될 수 있다.

[0869] 신호화 생화학적 경로와 관련된 단백질이 세포내 pH 조건의 변동을 초래하는 연쇄반응의 일부인 경우, 형광 pH 염료와 같은 pH 감응성 분자가 리포터 분자로서 사용될 수 있다. 신호화 생화학적 경로와 관련된 단백질이 이온 채널인 다른 예에서, 막 전위 및/또는 세포내 이온 농도의 변동이 모니터될 수 있다. 다수의 상업적 키트 및 고-처리량 장치가 이온 채널의 조절인자에 대한 신속하고 확실한 스크리닝에 특히 적합하다. 대표적인 기기는 FLIPRTM(Molecular Devices, Inc.)와 VIPR(Aurora Biosciences)을 포함한다. 이들 기기는 동시에 마이크로플레이트의 1000개 샘플 웰 이상에서 반응을 검출하고, 실시간 측정 및 1 초 안에 또는 미니세컨드 안에 기능적 데이터를 제공할 수 있다.

[0870] 여기 개시된 방법 중 어느 것을 실시하는데 있어서, 적합한 벡터가, 제한은 아니지만 마이크로인젝션, 전기천공, 소노포레이션, 바이오리스틱스, 칼슘 포스페이트-매개 트랜스펙션, 양이온 트랜스펙션, 리포솜 트랜스펙션, 덴드리머 트랜스펙션, 열충격 트랜스펙션, 뉴클레오펙션 트랜스펙션, 마그네토펙션, 리포펙션, 임페일펙션, 광학 트랜스펙션, 핵산의 전용 제제-증진 흡수, 및 리포솜, 면역리포솜, 비로솜 또는 인공 비리온을 통한 전달을 포함하는 본 분야에 공지된 하나 이상의 방법을 통해서 세포 또는 배아에 도입될 수 있다. 일부 방법에서, 벡터는 마이크로인젝션에 의해서 배아에 도입된다. 벡터 또는 벡터들은 배아의 핵 또는 세포질에 마이크로인젝션될 수 있다. 일부 방법에서, 벡터 또는 벡터들은 뉴클레오펙션에 의해서 세포에 도입될 수 있다.

[0871] CRISPR 복합체의 표적 폴리뉴클레오티드는 진핵 세포에 대해 내인성이거나 외인성인 임의의 폴리뉴클레오티드일 수 있다. 예를 들어, 표적 폴리뉴클레오티드는 진핵 세포의 핵에 존재하는 폴리뉴클레오티드일 수 있다. 표적 폴리뉴클레오티드는 유전자 산물(예를 들어, 단백질)을 코딩하는 서열 또는 비-코딩 서열(예를 들어, 조절 폴리뉴클레오티드 또는 정크(junk) DNA)일 수 있다.

[0872] 표적 폴리뉴클레오티드의 예는 신호전달 생화학 경로와 관련된 서열, 예를 들어, 신호전달 생화학적 경로-관련 유전자 또는 폴리뉴클레오티드를 포함한다. 표적 폴리뉴클레오티드의 예는 질병 관련 유전자 또는 폴리뉴클레오티드를 포함한다. "질병-관련" 유전자 또는 폴리뉴클레오티드는 질병이 없는 대조군의 조직 또는 세포와 비교하

여, 질병-발생 조직으로부터 유래된 세포에서 비정상적인 수준 또는 비정상적인 형태로 전사 또는 번역 산물을 생성하는 임의의 유전자 또는 폴리뉴클레오티드를 지칭한다. 그것은 비정상적으로 높은 수준으로 발현되는 유전자일 수 있으며; 그것은 비정상적으로 낮은 수준으로 발현되는 유전자일 수 있고, 여기서 변경된 발현은 질병의 발생 및/또는 진행과 관련이 있다. 또한, 질병-관련 유전자는 질병의 병인에 직접적인 원인이 있거나, 그에 원인이 있는 유전자(들)와 연관 불균형이 있는 돌연변이(들) 또는 유전적 변이를 갖는 유전자를 지칭한다. 전사 또는 번역된 산물은 공지된 것이거나 미공지된 것일 수 있으며, 정상 또는 비정상 수준으로 존재할 수 있다.

[0873] CRISPR 복합체의 표적 폴리뉴클레오티드는 진핵 세포에 대해 내인성이거나 외인성인 임의의 폴리뉴클레오티드일 수 있다. 예를 들어, 표적 폴리뉴클레오티드는 진핵 세포의 핵에 존재하는 폴리뉴클레오티드일 수 있다. 표적 폴리뉴클레오티드는 유전자 산물(예를 들어, 단백질)을 코딩하는 서열 또는 비-코딩 서열(예를 들어, 조절 폴리뉴클레오티드 또는 정크(junk) DNA)일 수 있다. 이론에 구속되지 않으면서, 표적 서열이 PAM(프로토스페이서 인접 모티프); 즉, CRISPR 복합체에 의해 인식되는 짧은 서열과 회합되어야 하는 것으로 여겨진다. PAM에 대한 정밀한 서열 및 길이 요건은 사용되는 CRISPR 효소에 따라 달라지지만, PAM은 통상적으로 프로토스페이서(즉, 표적 서열)에 인접한 2 내지 5 개 염기쌍 서열이다. PAM 서열의 예는 하기 실시예 섹션에 제공되어 있으며, 당업자는 주어진 CRISPR 효소와 함께 사용하기 위한 추가의 PAM 서열을 확인할 수 있을 것이다.

[0874] CRISPR 복합체의 표적 폴리뉴클레오티드는 모두의 내용이 본원에 참조로 포함되는, 각각 브로드 참조번호 BI-2011/008/WSGR 사건 번호 44063-701.101 및 BI-2011/008/WSGR 사건 번호 44063-701.102를 갖고, 둘 모두 명칭이 SYSTEMS METHODS AND COMPOSITIONS FOR SEQUENCE MANIPULATION이고, 각각 2012년 12월 12일 및 2013년 1월 2일에 출원된 미국 가출원 제61/736,527호 및 제61/748,427호에 열거된 바와 같은 다수의 질병-관련 유전자 및 폴리뉴클레오티드, 및 신호전달 생화학 경로-관련 유전자 및 폴리뉴클레오티드를 포함할 수 있다.

[0875] 표적 폴리뉴클레오티드의 예는 신호전달 생화학 경로와 관련된 서열, 예를 들어, 신호전달 생화학적 경로-관련 유전자 또는 폴리뉴클레오티드를 포함한다. 표적 폴리뉴클레오티드의 예는 질병 관련 유전자 또는 폴리뉴클레오티드를 포함한다. "질병-관련" 유전자 또는 폴리뉴클레오티드는 질병이 없는 대조군의 조직 또는 세포와 비교하여, 질병-발생 조직으로부터 유래된 세포에서 비정상적인 수준 또는 비정상적인 형태로 전사 또는 번역 산물을 생성하는 임의의 유전자 또는 폴리뉴클레오티드를 지칭한다. 그것은 비정상적으로 높은 수준으로 발현되는 유전자일 수 있으며; 그것은 비정상적으로 낮은 수준으로 발현되는 유전자일 수 있고, 여기서, 변경된 발현은 질병의 발생 및/또는 진행과 관련이 있다. 또한, 질병-관련 유전자는 질병의 병인에 직접적인 원인이 있거나, 그에 원인이 있는 유전자(들)와 연관 불균형이 있는 돌연변이(들) 또는 유전적 변이를 갖는 유전자를 지칭한다. 전사 또는 번역된 산물은 공지된 것이거나 미공지된 것일 수 있으며, 정상 또는 비정상 수준으로 존재할 수 있다.

[0876] 질병-관련 유전자 및 폴리뉴클레오티드의 예는 표 A 및 B에 열거되어 있다. 질병 특이적 정보는 맥쿠식-네이선스 유전의학연구소, 존스 홉킨스 대학(미국 메릴랜드주 볼티모어) 및 미국 국립생물공학정보센터, 국립 의학 도서관(미국 메릴랜드주 베데스다)로부터 이용가능하며, 월드 와이드 웹에서 이용가능하다. 신호전달 생화학 경로-관련 유전자 및 폴리뉴클레오티드의 예는 표 C에 열거되어 있다.

[0877] 이들 유전자 및 경로의 돌연변이는 기능에 영향을 미치는 부적절한 단백질 또는 부적절한 양의 단백질의 생성을 야기할 수 있다. 유전자, 질병 및 단백질의 추가의 예는 2012년 12월 12일에 출원된 미국 가출원 제61/736,527호로부터 본원에 참고로 포함된다. 이러한 유전자, 단백질 및 경로는 CRISPR 복합체의 표적 폴리뉴클레오티드일 수 있다.

[0878] 표 A

질병/장애	유전자(들)
신생물	PTEN; ATM; ATR; EGFR; ERBB2; ERBB3; ERBB4;
	Notch1; Notch2; Notch3; Notch4; AKT; AKT2; AKT3; HIF;
	HIF1a; HIF3a; Met; HRG; Bcl2; PPAR 알파; PPAR
	감마; WT1(윌름스 종양); FGF 수용체 과
	구성원(5 개 구성원: 1, 2, 3, 4, 5); CDKN2a; APC; RB
	(망막모세포종); MEN1; VHL; BRCA1; BRCA2; AR
	(안드로겐 수용체); TSG101; IGF; IGF 수용체; Igf1 (4 개
	변이체); Igf2 (3 개 변이체); Igf 1 수용체; Igf 2 수용체;
	Bax; Bcl2; 카스파제 과(9 개 구성원:
	1, 2, 3, 4, 6, 7, 8, 9, 12); Kras; Apc
연령-관련 황반	Abcr; Ccl2; Cc2; cp(세룰로플라스민); Timp3; 카텝신 D;
변성	Vldlr; Ccr2
정신분열병	뉴레귤린 1(Nrg1); Erb4(뉴레귤린에 대한 수용체);
	컴플렉신 1(Cplx1); Tph1 트립토판 하이드록실라제; Tph2
	트립토판 하이드록실라제 2; 뉴렉신 1; GSK3; GSK3a;
	GSK3b
장애	5-HTT(Slc6a4); COMT; DRD(Drd1a); SLC6A3; DAOA;
	DTNBP1; Dao(Dao1)
트리뉴클레오티드 반복	HTT(헌팅톤의 진단); SBMA/SMAX1/AR(케네디의
장애	진단); FXN/X25(프리드리히 실조증); ATX3(마카도-
	조셉의 진단); ATXN1 및 ATXN2(척수소뇌성
	실조증); DMPK(근긴장성 이영양증); 아트로핀(Atrophin)-1 및 Atn1
	(DRPLA Dx); CBP(Creb-BP - 전체 분안정성); VLDLR
	(알츠하이머); Atxn7; Atxn10
유약 X 증후군	FMR2; FXR1; FXR2; mGLUR5
세크레타제 관련	APH-1(알파 및 베타); 프레세닐린(Presenilin)(Psen1); 니카스트린(nicastrin)
장애	(Ncstn); PEN-2
기타	Nos1; Parp1; Nat1; Nat2
프리온-관련 장애	Prp
ALS	SOD1; ALS2; STEX; FUS; TARDBP; VEGF(VEGF-a;
	VEGF-b; VEGF-c)
약물 중독	Prkce(알콜); Drd2; Drd4; ABAT(알콜); GRIA2;
	Grm5; Grin1; Htr1b; Grin2a; Drd3; Pdyn; Grial(알콜)
자폐증	Mecp2; BZRAP1; MDGA2; Sema5A; 뉴렉신 1; 유약 X
	(FMR2 (AFF2); FXR1; FXR2; Mglur5)
알츠하이머병	E1; CHIP; UCH; UBB; Tau; LRP; PICALM; 클러스테린(Clusterin); PS1;
	SORL1; CR1; Vldlr; Uba1; Uba3; CHIP28 (Aqp1,

[0879]

	아쿠아포린 1); Uchl1; Uchl3; APP
염증	IL-10; IL-1 (IL-1a; IL-1b); IL-13; IL-17 (IL-17a (CTLA8); IL-
	17b; IL-17c; IL-17d; IL-17f); II-23; Cx3cr1; ptpn22; TNFa;
	IBD 에 대한 NOD2/CARD15; IL-6; IL-12(IL-12a; IL-12b);
	CTLA4; Cx3cl1
파킨슨병	x-시누클레인; DJ-1; LRRK2; Parkin; PINK1

[0880]

[0881] 표 B:

혈액 및 응고 질병 및 장애	빈혈증 (CDAN1, CDA1, RPS19, DBA, PKLR, PK1, NT5C3, UMPH1, PSN1, RHAG, RH50A, NRAMP2, SPTB, ALAS2, ANH1, ASB, ABCB7, ABC7, ASAT); 무포지 림프구 증후군 (TAPBP, TPSN, TAP2, ABCB3, PSF2, RING11, MHC2TA, C2TA, RFX5, RFXAP, RFX5), 출혈성 장애 (TBXA2R, P2RX1, P2X1); 인자 H 및 인자 H-유사 1 (HF1, CFH, HUS); 제 V 인자 및 제 VIII 인자 (MCFD2); 제 VII 인자 결핍 (F7); 제 X 인자 결핍 (F10); 제 XI 인자 결핍 (F11); 제 XII 인자 결핍 (F12, HAF); 제 XIIIA 인자 결핍 (F13A1, F13A); 제 XIIIB 인자 결핍 (F13B); 판코니 빈혈증 (FANCA, FACA, FA1, FA, FAA, FAAP95, FAAP90, FLJ34064, FANCB, FANCC, FACC, BRCA2, FANCD1, FANCD2, FANCD, FACD, FAD, FANCE, FACE, FANCF, XRCC9, FANCG, BRIP1, BACH1, FANCJ, PHF9, FANCL, FANCM, KIAA1596); 혈구탐식성 림프조직구증 장애 (PRF1, HPLH2, UNC13D, MUNC13-4, HPLH3, HLH3, FHL3); A 형 혈우병 (F8, F8C, HEMA); B 형 혈우병 (F9, HEMB), 출혈성 장애 (PI, ATT, F5); 백혈구 결핍 및 장애 (ITGB2, CD18, LCAMB, LAD, EIF2B1, EIF2BA, EIF2B2, EIF2B3, EIF2B5, LVWM, CACH, CLE, EIF2B4); 겸상적혈구 빈혈증 (HBB); 지중해빈혈증 (HBA2, HBB, HBD, LCRB, HBA1).
세포 조절이상 및 종양학 질병 및 장애	B-세포 비호지킨 림프종 (BCL7A, BCL7); 백혈병 (TAL1, TCL5, SCL, TAL2, FLT3, NBS1, NBS, ZNFN1A1, IK1, LYF1, HOXD4, HOX4B, BCR, CML, PHL, ALL, ARNT, KRAS2, RASK2, GMPS, AF10, ARHGEF12, LARG, KIAA0382, CALM, CLTH, CEBPA, CEBP, CHIC2, BTL, FLT3, KIT, PBT, LPP, NPM1, NUP214, D9S46E, CAN, CAIN, RUNX1, CBFA2, AML1, WHSC1L1, NSD3, FLT3, AF1Q, NPM1, NUMA1, ZNF145, PLZF, PML, MYL, STAT5B, AF10, CALM, CLTH, ARL11, ARLTS1, P2RX7, P2X7, BCR, CML, PHL, ALL, GRAF, NF1, VRNF, WSS, NFNS, PTPN11, PTP2C, SHP2, NS1, BCL2, CCND1, PRAD1, BCL1, TCRA, GATA1, GF1, ERYF1, NFE1, ABL1, NQO1, DIA4, NMOR1, NUP214, D9S46E, CAN, CAIN).
염증 및 면역 관련 질병 및 장애	AIDS (KIR3DL1, NKAT3, NKB1, AMB11, KIR3DS1, IFNG, CXCL12, SDF1); 자가면역 림프구증식 증후군 (TNFRSF6, APT1, FAS, CD95, ALPS1A); 복합 면역결핍, (IL2RG, SCIDX1, SCIDX, IMD4); HIV-1 (CCL5, SCYA5, D17S136E, TCP228), HIV 감수성 또는 감염 (IL10, CSIF, CMKBR2, CCR2, CMKBR5, CCCKR5 (CCR5)); 면역결핍 (CD3E, CD3G, AICDA, AID, HIGM2, TNFRSF5, CD40, UNG, DGU, HIGM4, TNFSF5, CD40LG, HIGM1, IGM, FOXP3, IPEX, AIID, XPID, PIDX, TNFRSF14B, TACI); 염증 (IL-10, IL-1 (IL-1a, IL-1b), IL-13, IL-17 (IL-17a (CTLA8), IL-17b, IL-17c, IL-17d, IL-

[0882]

	17f), Il-23, Cx3cr1, ptpn22, TNFa, IBD에 대한 NOD2/CARD15, IL-6, IL-12 (IL-12a, IL-12b), CTLA4, Cx3cl1); 중증 복합 면역결핍 (SCID) (JAK3, JAKL, DCLRE1C, ARTEMIS, SCIDA, RAG1, RAG2, ADA, PTPRC, CD45, LCA, IL7R, CD3D, T3D, IL2RG, SCIDX1, SCIDX, IMD4).
대사, 간, 신장 및 단백질 질병 및 장애	아밀로이드 신경병증 (TTR, PALB); 아밀로이드증 (APOA1, APP, AAA, CVAP, AD1, GSN, FGA, LYZ, TTR, PALB); 경화증 (KRT18, KRT8, CIRH1A, NAIC, TEX292, KIAA1988); 낭성 섬유증 (CFTR, ABCC7, CF, MRP7); 글리코겐 저장 질병 (SLC2A2, GLUT2, G6PC, G6PT, G6PT1, GAA, LAMP2, LAMPB, AGL, GDE, GBE1, GYS2, PYGL, PFKM); 간선종, 142330 (TCF1, HNF1A, MODY3), 간부전, 조기 발병 및 신경계 장애 (SCOD1, SCO1), 간 리파제 결핍 (LIPC), 간모세포종, 암 및 암종 (CTNNB1, PDGFRL, PDGRL, PRLTS, AXIN1, AXIN, CTNNB1, TP53, P53, LFS1, IGF2R, MPRI, MET, CASP8, MCH5; 수질 낭성 신장 질병 (UMOD, HNFJ, FJHN, MCKD2, ADMCKD2); 페닐케톤뇨증 (PAH, PKU1, QDPR, DHPR, PTS); 다낭성 신장 및 간 질병 (FCYT, PKHD1, ARPKD, PKD1, PKD2, PKD4, PKDTS, PRKCSH, G19P1, PCLD, SEC63).
근육 / 골격 질병 및 장애	베커 근이영양증 (DMD, BMD, MYF6), 뒤시엔느 근이영양증 (DMD, BMD); 에머리-드레이푸스 근이영양증 (LMNA, LMN1, EMD2, FPLD, CMD1A, HGPS, LGMD1B, LMNA, LMN1, EMD2, FPLD, CMD1A); 안면견갑상완근이영양증 (FSHMD1A, FSHD1A); 근이영양증 (FKRP, MDC1C, LGMD2I, LAMA2, LAMM, LARGE, KIAA0609, MDC1D, FCMD, TTID, MYOT, CAPN3, CANP3, DYSF, LGMD2B, SGCG, LGMD2C, DMDA1, SCG3, SGCA, ADL, DAG2, LGMD2D, DMDA2, SGCB, LGMD2E, SGCD, SGD, LGMD2F, CMD1L, TCAP, LGMD2G, CMD1N, TRIM32, HT2A, LGMD2H, FKRP, MDC1C, LGMD2I, TTN, CMD1G, TMD, LGMD2J, POMT1, CAV3, LGMD1C, SEPN1, SELN, RSMD1, PLEC1, PLTN, EBS1); 골화석증 (LRP5, BMND1, LRP7, LR3, OPPG, VBCH2, CLCN7, CLC7, OPTA2, OSTM1, GL, TCIRG1, TIRC7, OC116, OPTB1); 근이영양증 (VAPB, VAPC, ALS8, SMN1, SMA1, SMA2, SMA3, SMA4, BSCL2, SPG17, GARS, SMAD1, CMT2D, HEXB, IGHMBP2, SMUBP2, CATF1, SMARD1).
신경계 및 신경 질병 및 장애	ALS (SOD1, ALS2, STEX, FUS, TARDBP, VEGF (VEGF-a, VEGF-b, VEGF-c); 알츠하이머병 (APP, AAA, CVAP, AD1, APOE, AD2, PSEN2, AD4, STM2, APBB2, FE65L1, NOS3, PLAU, URK, ACE, DCP1, ACE1, MPO, PACIP1, PAXIP1L, PTIP, A2M, BLMH, BMH, PSEN1, AD3); 자폐증 (Mecp2, BZRAP1, MDGA2, Sema5A, 뉴렉신 1, GLO1, MECP2, RTT, PPMX, MRX16, MRX79, NLGN3, NLGN4, KIAA1260, AUTSX2); 유약 X 증후군 (FMR2, FXR1, FXR2,

[0883]

	mGLUR5); 헌팅톤 질병 및 질병 유사 장애 (HD, IT15, PRNP, PRIP, JPH3, JP3, HDL2, TBP, SCA17); 파킨슨병 (NR4A2, NURR1, NOT, TINUR, SNCAIP, TBP, SCA17, SNCA, NACP, PARK1, PARK4, DJ1, PARK7, LRRK2, PARK8, PINK1, PARK6, UCHL1, PARK5, SNCA, NACP, PARK1, PARK4, PRKN, PARK2, PDJ, DBH, NDUFV2); 레트 증후군 (MECP2, RTT, PPMX, MRX16, MRX79, CDKL5, STK9, MECP2, RTT, PPMX, MRX16, MRX79, x-시누클레인, DJ-1); 정신분열병 (뉴레귤린 1 (Nrg1), Erb4 (뉴레귤린에 대한 수용체), 컴플렉신 1 (Cplx1), Tph1 트립토판 하이드록실라제, Tph2, 트립토판 하이드록실라제 2, 뉴렉신 1, GSK3, GSK3a, GSK3b, 5-HTT (Slc6a4), COMT, DRD (Drd1a), SLC6A3, DAOA, DTNBP1, Dao (Dao1)); 세크레타제 관련 장애 (APH-1 (알파 및 베타), 프레세닐린 (Psen1), 니카스트린, (Ncstn), PEN-2, Nos1, Parp1, Nat1, Nat2); 트리뉴클레오티드 반복 장애 (HTT (헌팅톤의 진단), SBMA/SMAX1/AR (케네디의 진단), FXN/X25 (프리드리히 실조증), ATX3 (마카도-조셉의 진단), ATXN1 및 ATXN2 (척수소뇌성 실조증), DMPK (근긴장성 이영양증), 아트로핀-1 및 Atn1 (DRPLA Dx), CBP (Creb-BP - 전체 불안정성), VLDLR (알츠하이머), Atxn7, Atxn10).
안구 질병 및 장애	연령-관련 황반 변성 (Abcr, Ccl2, Cc2, cp (세룰로플라스민), Timp3, 카텝신 D, Vldlr, Ccr2); 백내장 (CRYAA, CRYA1, CRYBB2, CRYB2, PITX3, BFSP2, CP49, CP47, CRYAA, CRYA1, PAX6, AN2, MGDA, CRYBA1, CRYB1, CRYGC, CRYG3, CCL, LIM2, MP19, CRYGD, CRYG4, BFSP2, CP49, CP47, HSF4, CTM, HSF4, CTM, MIP, AQP0, CRYAB, CRYA2, CTPP2, CRYBB1, CRYGD, CRYG4, CRYBB2, CRYB2, CRYGC, CRYG3, CCL, CRYAA, CRYA1, GJA8, CX50, CAE1, GJA3, CX46, CZP3, CAE3, CCM1, CAM, KRIT1); 각막 혼탁 및 이영양증 (APOA1, TGFBI, CSD2, CDGG1, CSD, BIGH3, CDG2, TACSTD2, TROP2, M1S1, VSX1, RINX, PPCD, PPD, KTCN, COL8A2, FECD, PPCD2, PIP5K3, CFD); 선청성 낙작 각막(KERA, CNA2); 녹내장 (MYOC, TIGR, GLC1A, JOAG, GPOA, OPTN, GLC1E, FIP2, HYPL, NRP, CYP1B1, GLC3A, OPA1, NTG, NPG, CYP1B1, GLC3A); 레베르선천흑암시 (CRB1, RP12, CRX, CORD2, CRD, RPGRIP1, LCA6, CORD9, RPE65, RP20, AIPL1, LCA4, GUCY2D, GUC2D, LCA1, CORD6, RDH12, LCA3); 황반 이영양증 (ELOVL4, ADMD, STGD2, STGD3, RDS, RP7, PRPH2, PRPH, AVMD, AOFMD, VMD2).

[0884]

[0885] 표 C:

세포 기능	유전자
PI3K/AKT 신호전달	PRKCE; ITGAM; ITGA5; IRAK1; PRKAA2; EIF2AK2;
	PTEN; EIF4E; PRKCZ; GRK6; MAPK1; TSC1; PLK1;
	AKT2; IKBKB; PIK3CA; CDK8; CDKN1B; NFKB2; BCL2;
	PIK3CB; PPP2R1A; MAPK8; BCL2L1; MAPK3; TSC2;
	ITGA1; KRAS; EIF4EBP1; RELA; PRKCD; NOS3;
	PRKAA1; MAPK9; CDK2; PPP2CA; PIM1; ITGB7;
	YWHAZ; ILK; TP53; RAF1; IKBKG; RELB; DYRK1A;
	CDKN1A; ITGB1; MAP2K2; JAK1; AKT1; JAK2; PIK3R1;
	CHUK; PDPK1; PPP2R5C; CTNNB1; MAP2K1; NFKB1;
	PAK3; ITGB3; CCND1; GSK3A; FRAP1; SFN; ITGA2;
	TTK; CSNK1A1; BRAF; GSK3B; AKT3; FOXO1; SGK;
	HSP90AA1; RPS6KB1
ERK/MAPK 신호전달	PRKCE; ITGAM; ITGA5; HSPB1; IRAK1; PRKAA2;
	EIF2AK2; RAC1; RAP1A; TLN1; EIF4E; ELK1; GRK6;
	MAPK1; RAC2; PLK1; AKT2; PIK3CA; CDK8; CREB1;
	PRKC1; PTK2; FOS; RPS6KA4; PIK3CB; PPP2R1A;
	PIK3C3; MAPK8; MAPK3; ITGA1; ETS1; KRAS; MYCN;
	EIF4EBP1; PPARG; PRKCD; PRKAA1; MAPK9; SRC;
	CDK2; PPP2CA; PIM1; PIK3C2A; ITGB7; YWHAZ;
	PPP1CC; KSR1; PXN; RAF1; FYN; DYRK1A; ITGB1;
	MAP2K2; PAK4; PIK3R1; STAT3; PPP2R5C; MAP2K1;
	PAK3; ITGB3; ESR1; ITGA2; MYC; TTK; CSNK1A1;
	CRKL; BRAF; ATF4; PRKCA; SRF; STAT1; SGK
글루코코르티코이드 수용체 신호전달	RAC1; TAF4B; EP300; SMAD2; TRAF6; PCAF; ELK1;
	MAPK1; SMAD3; AKT2; IKBKB; NCOR2; UBE2I;
	PIK3CA; CREB1; FOS; HSPA5; NFKB2; BCL2;
	MAP3K14; STAT5B; PIK3CB; PIK3C3; MAPK8; BCL2L1;
	MAPK3; TSC22D3; MAPK10; NR1P1; KRAS; MAPK13;
	RELA; STAT5A; MAPK9; NOS2A; PBX1; NR3C1;
	PIK3C2A; CDKN1C; TRAF2; SERPINE1; NCOA3;
	MAPK14; TNF; RAF1; IKBKG; MAP3K7; CREBBP;
	CDKN1A; MAP2K2; JAK1; IL8; NCOA2; AKT1; JAK2;
	PIK3R1; CHUK; STAT3; MAP2K1; NFKB1; TGFBR1;
	ESR1; SMAD4; CEBPB; JUN; AR; AKT3; CCL2; MMP1;
	STAT1; IL6; HSP90AA1
축삭 유도 신호전달	PRKCE; ITGAM; ROCK1; ITGA5; CXCR4; ADAM12;
	IGF1; RAC1; RAP1A; EIF4E; PRKCZ; NRP1; NTRK2;
	ARHGEF7; SMO; ROCK2; MAPK1; PGF; RAC2;

[0886]

	PTPN11; GNAS; AKT2; PIK3CA; ERBB2; PRKCI; PTK2;
	CFL1; GNAQ; PIK3CB; CXCL12; PIK3C3; WNT11;
	PRKD1; GNB2L1; ABL1; MAPK3; ITGA1; KRAS; RHOA;
	PRKCD; PIK3C2A; ITGB7; GLI2; PXN; VASP; RAF1;
	FYN; ITGB1; MAP2K2; PAK4; ADAM17; AKT1; PIK3R1;
	GLI1; WNT5A; ADAM10; MAP2K1; PAK3; ITGB3;
	CDC42; VEGFA; ITGA2; EPHA8; CRKL; RND1; GSK3B;
	AKT3; PRKCA
에프린 수용체 신호전달	PRKCE; ITGAM; ROCK1; ITGA5; CXCR4; IRAK1;
	PRKAA2; EIF2AK2; RAC1; RAP1A; GRK6; ROCK2;
	MAPK1; PGF; RAC2; PTPN11; GNAS; PLK1; AKT2;
	DOK1; CDK8; CREB1; PTK2; CFL1; GNAQ; MAP3K14;
	CXCL12; MAPK8; GNB2L1; ABL1; MAPK3; ITGA1;
	KRAS; RHOA; PRKCD; PRKAA1; MAPK9; SRC; CDK2;
	PIM1; ITGB7; PXN; RAF1; FYN; DYRK1A; ITGB1;
	MAP2K2; PAK4; AKT1; JAK2; STAT3; ADAM10;
	MAP2K1; PAK3; ITGB3; CDC42; VEGFA; ITGA2;
	EPHA8; TTK; CSNK1A1; CRKL; BRAF; PTPN13; ATF4;
	AKT3; SGK
액틴 세포골격	ACTN4; PRKCE; ITGAM; ROCK1; ITGA5; IRAK1;
신호전달	PRKAA2; EIF2AK2; RAC1; INS; ARHGEF7; GRK6;
	ROCK2; MAPK1; RAC2; PLK1; AKT2; PIK3CA; CDK8;
	PTK2; CFL1; PIK3CB; MYH9; DIAPH1; PIK3C3; MAPK8;
	F2R; MAPK3; SLC9A1; ITGA1; KRAS; RHOA; PRKCD;
	PRKAA1; MAPK9; CDK2; PIM1; PIK3C2A; ITGB7;
	PPP1CC; PXN; VIL2; RAF1; GSN; DYRK1A; ITGB1;
	MAP2K2; PAK4; PIP5K1A; PIK3R1; MAP2K1; PAK3;
	ITGB3; CDC42; APC; ITGA2; TTK; CSNK1A1; CRKL;
	BRAF; VAV3; SGK
헌팅톤병	PRKCE; IGF1; EP300; RCOR1; PRKCZ; HDAC4; TGM2;
신호전달	MAPK1; CAPNS1; AKT2; EGFR; NCOR2; SP1; CAPN2;
	PIK3CA; HDAC5; CREB1; PRKCI; HSPA5; REST;
	GNAQ; PIK3CB; PIK3C3; MAPK8; IGF1R; PRKD1;
	GNB2L1; BCL2L1; CAPN1; MAPK3; CASP8; HDAC2;
	HDAC7A; PRKCD; HDAC11; MAPK9; HDAC9; PIK3C2A;
	HDAC3; TP53; CASP9; CREBBP; AKT1; PIK3R1;
	PDPK1; CASP1; APAF1; FRAP1; CASP2; JUN; BAX;
	ATF4; AKT3; PRKCA; CLTC; SGK; HDAC6; CASP3
아폽토시스 신호전달	PRKCE; ROCK1; BID; IRAK1; PRKAA2; EIF2AK2; BAK1;
	BIRC4; GRK6; MAPK1; CAPNS1; PLK1; AKT2; IKBKB;
	CAPN2; CDK8; FAS; NFKB2; BCL2; MAP3K14; MAPK8;

[0887]

	BCL2L1; CAPN1; MAPK3; CASP8; KRAS; RELA;
	PRKCD; PRKAA1; MAPK9; CDK2; PIM1; TP53; TNF;
	RAF1; IKBKG; RELB; CASP9; DYRK1A; MAP2K2;
	CHUK; APAF1; MAP2K1; NFKB1; PAK3; LMNA; CASP2;
	BIRC2; TTK; CSNK1A1; BRAF; BAX; PRKCA; SGK;
	CASP3; BIRC3; PARP1
B 세포 수용체 신호전달	RAC1; PTEN; LYN; ELK1; MAPK1; RAC2; PTPN11;
	AKT2; IKBKB; PIK3CA; CREB1; SYK; NFKB2; CAMK2A;
	MAP3K14; PIK3CB; PIK3C3; MAPK8; BCL2L1; ABL1;
	MAPK3; ETS1; KRAS; MAPK13; RELA; PTPN6; MAPK9;
	EGR1; PIK3C2A; BTK; MAPK14; RAF1; IKBKG; RELB;
	MAP3K7; MAP2K2; AKT1; PIK3R1; CHUK; MAP2K1;
	NFKB1; CDC42; GSK3A; FRAP1; BCL6; BCL10; JUN;
	GSK3B; ATF4; AKT3; VAV3; RPS6KB1
백혈구 혈관외 유출 신호전달	ACTN4; CD44; PRKCE; ITGAM; ROCK1; CXCR4; CYBA;
	RAC1; RAP1A; PRKCZ; ROCK2; RAC2; PTPN11;
	MMP14; PIK3CA; PRKCI; PTK2; PIK3CB; CXCL12;
	PIK3C3; MAPK8; PRKD1; ABL1; MAPK10; CYBB;
	MAPK13; RHOA; PRKCD; MAPK9; SRC; PIK3C2A; BTK;
	MAPK14; NOX1; PXN; VIL2; VASP; ITGB1; MAP2K2;
	CTNND1; PIK3R1; CTNNB1; CLDN1; CDC42; F11R; ITK;
	CRKL; VAV3; CTTN; PRKCA; MMP1; MMP9
인테그린 신호전달	ACTN4; ITGAM; ROCK1; ITGA5; RAC1; PTEN; RAP1A;
	TLN1; ARHGEF7; MAPK1; RAC2; CAPNS1; AKT2;
	CAPN2; PIK3CA; PTK2; PIK3CB; PIK3C3; MAPK8;
	CAV1; CAPN1; ABL1; MAPK3; ITGA1; KRAS; RHOA;
	SRC; PIK3C2A; ITGB7; PPP1CC; ILK; PXN; VASP;
	RAF1; FYN; ITGB1; MAP2K2; PAK4; AKT1; PIK3R1;
	TNK2; MAP2K1; PAK3; ITGB3; CDC42; RND3; ITGA2;
	CRKL; BRAF; GSK3B; AKT3
급성기 반응 신호전달	IRAK1; SOD2; MYD88; TRAF6; ELK1; MAPK1; PTPN11;
	AKT2; IKBKB; PIK3CA; FOS; NFKB2; MAP3K14;
	PIK3CB; MAPK8; RIPK1; MAPK3; IL6ST; KRAS;
	MAPK13; IL6R; RELA; SOCS1; MAPK9; FTL; NR3C1;
	TRAF2; SERPINE1; MAPK14; TNF; RAF1; PDK1;
	IKBKG; RELB; MAP3K7; MAP2K2; AKT1; JAK2; PIK3R1;
	CHUK; STAT3; MAP2K1; NFKB1; FRAP1; CEBPB; JUN;
	AKT3; IL1R1; IL6
PTEN 신호전달	ITGAM; ITGA5; RAC1; PTEN; PRKCZ; BCL2L11;
	MAPK1; RAC2; AKT2; EGFR; IKBKB; CBL; PIK3CA;
	CDKN1B; PTK2; NFKB2; BCL2; PIK3CB; BCL2L1;

[0888]

	MAPK3; ITGA1; KRAS; ITGB7; ILK; PDGFRB; INSR;
	RAF1; IKBKG; CASP9; CDKN1A; ITGB1; MAP2K2;
	AKT1; PIK3R1; CHUK; PDGFRA; PDPK1; MAP2K1;
	NFKB1; ITGB3; CDC42; CCND1; GSK3A; ITGA2;
	GSK3B; AKT3; FOXO1; CASP3; RPS6KB1
p53 신호전달	PTEN; EP300; BBC3; PCAF; FASN; BRCA1; GADD45A;
	BIRC5; AKT2; PIK3CA; CHEK1; TP53INP1; BCL2;
	PIK3CB; PIK3C3; MAPK8; THBS1; ATR; BCL2L1; E2F1;
	PMAIP1; CHEK2; TNFRSF10B; TP73; RB1; HDAC9;
	CDK2; PIK3C2A; MAPK14; TP53; LRDD; CDKN1A;
	HIPK2; AKT1; PIK3R1; RRM2B; APAF1; CTNNB1;
	SIRT1; CCND1; PRKDC; ATM; SFN; CDKN2A; JUN;
	SNAI2; GSK3B; BAX; AKT3
아릴 탄화수소 수용체	HSPB1; EP300; FASN; TGM2; RXRA; MAPK1; NQO1;
신호전달	NCOR2; SP1; ARNT; CDKN1B; FOS; CHEK1;
	SMARCA4; NFKB2; MAPK8; ALDH1A1; ATR; E2F1;
	MAPK3; NR1P1; CHEK2; RELA; TP73; GSTP1; RB1;
	SRC; CDK2; AHR; NFE2L2; NCOA3; TP53; TNF;
	CDKN1A; NCOA2; APAF1; NFKB1; CCND1; ATM; ESR1;
	CDKN2A; MYC; JUN; ESR2; BAX; IL6; CYP1B1;
	HSP90AA1
생체이물 대사	PRKCE; EP300; PRKCZ; RXRA; MAPK1; NQO1;
신호전달	NCOR2; PIK3CA; ARNT; PRKCI; NFKB2; CAMK2A;
	PIK3CB; PPP2R1A; PIK3C3; MAPK8; PRKD1;
	ALDH1A1; MAPK3; NR1P1; KRAS; MAPK13; PRKCD;
	GSTP1; MAPK9; NOS2A; ABCB1; AHR; PPP2CA; FTL;
	NFE2L2; PIK3C2A; PPARGC1A; MAPK14; TNF; RAF1;
	CREBBP; MAP2K2; PIK3R1; PPP2R5C; MAP2K1;
	NFKB1; KEAP1; PRKCA; EIF2AK3; IL6; CYP1B1;
	HSP90AA1
SAPK/JNK 신호전달	PRKCE; IRAK1; PRKAA2; EIF2AK2; RAC1; ELK1;
	GRK6; MAPK1; GADD45A; RAC2; PLK1; AKT2; PIK3CA;
	FADD; CDK8; PIK3CB; PIK3C3; MAPK8; RIPK1;
	GNB2L1; IRS1; MAPK3; MAPK10; DAXX; KRAS;
	PRKCD; PRKAA1; MAPK9; CDK2; PIM1; PIK3C2A;
	TRAF2; TP53; LCK; MAP3K7; DYRK1A; MAP2K2;
	PIK3R1; MAP2K1; PAK3; CDC42; JUN; TTK; CSNK1A1;
	CRKL; BRAF; SGK
PPAr/RXR 신호전달	PRKAA2; EP300; INS; SMAD2; TRAF6; PPARA; FASN;
	RXRA; MAPK1; SMAD3; GNAS; IKBKB; NCOR2;
	ABCA1; GNAQ; NFKB2; MAP3K14; STAT5B; MAPK8;
	IRS1; MAPK3; KRAS; RELA; PRKAA1; PPARGC1A;

[0889]

	NCOA3; MAPK14; INSR; RAF1; IKBKG; RELB; MAP3K7;
	CREBBP; MAP2K2; JAK2; CHUK; MAP2K1; NFKB1;
	TGFBR1; SMAD4; JUN; IL1R1; PRKCA; IL6; HSP90AA1;
	ADIPOQ
NF-KB 신호전달	IRAK1; EIF2AK2; EP300; INS; MYD88; PRKCZ; TRAF6;
	TBK1; AKT2; EGFR; IKBKB; PIK3CA; BTRC; NFKB2;
	MAP3K14; PIK3CB; PIK3C3; MAPK8; RIPK1; HDAC2;
	KRAS; RELA; PIK3C2A; TRAF2; TLR4; PDGFRB; TNF;
	INSR; LCK; IKBKG; RELB; MAP3K7; CREBBP; AKT1;
	PIK3R1; CHUK; PDGFRA; NFKB1; TLR2; BCL10;
	GSK3B; AKT3; TNFAIP3; IL1R1
뉴레귤린 신호전달	ERBB4; PRKCE; ITGAM; ITGA5; PTEN; PRKCZ; ELK1;
	MAPK1; PTPN11; AKT2; EGFR; ERBB2; PRKCI;
	CDKN1B; STAT5B; PRKD1; MAPK3; ITGA1; KRAS;
	PRKCD; STAT5A; SRC; ITGB7; RAF1; ITGB1; MAP2K2;
	ADAM17; AKT1; PIK3R1; PDPK1; MAP2K1; ITGB3;
	EREG; FRAP1; PSEN1; ITGA2; MYC; NRG1; CRKL;
	AKT3; PRKCA; HSP90AA1; RPS6KB1
Wnt & 베타 카테닌	CD44; EP300; LRP6; DVL3; CSNK1E; GJA1; SMO;
신호전달	AKT2; PIN1; CDH1; BTRC; GNAQ; MARK2; PPP2R1A;
	WNT11; SRC; DKK1; PPP2CA; SOX6; SFRP2; ILK;
	LEF1; SOX9; TP53; MAP3K7; CREBBP; TCF7L2; AKT1;
	PPP2R5C; WNT5A; LRP5; CTNNB1; TGFBR1; CCND1;
	GSK3A; DVL1; APC; CDKN2A; MYC; CSNK1A1; GSK3B;
	AKT3; SOX2
인슐린 수용체 신호전달	PTEN; INS; EIF4E; PTPN1; PRKCZ; MAPK1; TSC1;
	PTPN11; AKT2; CBL; PIK3CA; PRKCI; PIK3CB; PIK3C3;
	MAPK8; IRS1; MAPK3; TSC2; KRAS; EIF4EBP1;
	SLC2A4; PIK3C2A; PPP1CC; INSR; RAF1; FYN;
	MAP2K2; JAK1; AKT1; JAK2; PIK3R1; PDPK1; MAP2K1;
	GSK3A; FRAP1; CRKL; GSK3B; AKT3; FOXO1; SGK;
	RPS6KB1
IL-6 신호전달	HSPB1; TRAF6; MAPKAPK2; ELK1; MAPK1; PTPN11;
	IKBKB; FOS; NFKB2; MAP3K14; MAPK8; MAPK3;
	MAPK10; IL6ST; KRAS; MAPK13; IL6R; RELA; SOCS1;
	MAPK9; ABCB1; TRAF2; MAPK14; TNF; RAF1; IKBKG;
	RELB; MAP3K7; MAP2K2; IL8; JAK2; CHUK; STAT3;
	MAP2K1; NFKB1; CEBPB; JUN; IL1R1; SRF; IL6
간의 쓴개즙정체	PRKCE; IRAK1; INS; MYD88; PRKCZ; TRAF6; PPARA;
	RXRA; IKBKB; PRKCI; NFKB2; MAP3K14; MAPK8;
	PRKD1; MAPK10; RELA; PRKCD; MAPK9; ABCB1;

[0890]

	TRAF2; TLR4; TNF; INSR; IKBKG; RELB; MAP3K7; IL8;
	CHUK; NR1H2; TJP2; NFKB1; ESR1; SREBF1; FGFR4;
	JUN; IL1R1; PRKCA; IL6
IGF-1 신호전달	IGF1; PRKCZ; ELK1; MAPK1; PTPN11; NEDD4; AKT2;
	PIK3CA; PRKCI; PTK2; FOS; PIK3CB; PIK3C3; MAPK8;
	IGF1R; IRS1; MAPK3; IGFBP7; KRAS; PIK3C2A;
	YWHAZ; PXN; RAF1; CASP9; MAP2K2; AKT1; PIK3R1;
	PDPK1; MAP2K1; IGFBP2; SFN; JUN; CYR61; AKT3;
	FOXO1; SRF; CTGF; RPS6KB1
NRF2-매개의 산화	PRKCE; EP300; SOD2; PRKCZ; MAPK1; SQSTM1;
스트레스 반응	NQO1; PIK3CA; PRKCI; FOS; PIK3CB; PIK3C3; MAPK8;
	PRKD1; MAPK3; KRAS; PRKCD; GSTP1; MAPK9; FTL;
	NFE2L2; PIK3C2A; MAPK14; RAF1; MAP3K7; CREBBP;
	MAP2K2; AKT1; PIK3R1; MAP2K1; PPIB; JUN; KEAP1;
	GSK3B; ATF4; PRKCA; EIF2AK3; HSP90AA1
간 섬유증/간	EDN1; IGF1; KDR; FLT1; SMAD2; FGFR1; MET; PGF;
별세포 활성화	SMAD3; EGFR; FAS; CSF1; NFKB2; BCL2; MYH9;
	IGF1R; IL6R; RELA; TLR4; PDGFRB; TNF; RELB; IL8;
	PDGFRA; NFKB1; TGFBR1; SMAD4; VEGFA; BAX;
	IL1R1; CCL2; HGF; MMP1; STAT1; IL6; CTGF; MMP9
PPAR 신호전달	EP300; INS; TRAF6; PPARA; RXRA; MAPK1; IKBKB;
	NCOR2; FOS; NFKB2; MAP3K14; STAT5B; MAPK3;
	NRIP1; KRAS; PPARG; RELA; STAT5A; TRAF2;
	PPARGC1A; PDGFRB; TNF; INSR; RAF1; IKBKG;
	RELB; MAP3K7; CREBBP; MAP2K2; CHUK; PDGFRA;
	MAP2K1; NFKB1; JUN; IL1R1; HSP90AA1
Fc 엡실론 RI 신호전달	PRKCE; RAC1; PRKCZ; LYN; MAPK1; RAC2; PTPN11;
	AKT2; PIK3CA; SYK; PRKCI; PIK3CB; PIK3C3; MAPK8;
	PRKD1; MAPK3; MAPK10; KRAS; MAPK13; PRKCD;
	MAPK9; PIK3C2A; BTK; MAPK14; TNF; RAF1; FYN;
	MAP2K2; AKT1; PIK3R1; PDPK1; MAP2K1; AKT3;
	VAV3; PRKCA
G-단백질 결합	PRKCE; RAP1A; RGS16; MAPK1; GNAS; AKT2; IKBKB;
수용체 신호전달	PIK3CA; CREB1; GNAQ; NFKB2; CAMK2A; PIK3CB;
	PIK3C3; MAPK3; KRAS; RELA; SRC; PIK3C2A; RAF1;
	IKBKG; RELB; FYN; MAP2K2; AKT1; PIK3R1; CHUK;
	PDPK1; STAT3; MAP2K1; NFKB1; BRAF; ATF4; AKT3;
	PRKCA
이노시톨 포스페이트	PRKCE; IRAK1; PRKAA2; EIF2AK2; PTEN; GRK6;
대사	MAPK1; PLK1; AKT2; PIK3CA; CDK8; PIK3CB; PIK3C3;
	MAPK8; MAPK3; PRKCD; PRKAA1; MAPK9; CDK2;

[0891]

	PIM1; PIK3C2A; DYRK1A; MAP2K2; PIP5K1A; PIK3R1;
	MAP2K1; PAK3; ATM; TTK; CSNK1A1; BRAF; SGK
PDGF 신호전달	EIF2AK2; ELK1; ABL2; MAPK1; PIK3CA; FOS; PIK3CB;
	PIK3C3; MAPK8; CAV1; ABL1; MAPK3; KRAS; SRC;
	PIK3C2A; PDGFRB; RAF1; MAP2K2; JAK1; JAK2;
	PIK3R1; PDGFRA; STAT3; SPHK1; MAP2K1; MYC;
	JUN; CRKL; PRKCA; SRF; STAT1; SPHK2
VEGF 신호전달	ACTN4; ROCK1; KDR; FLT1; ROCK2; MAPK1; PGF;
	AKT2; PIK3CA; ARNT; PTK2; BCL2; PIK3CB; PIK3C3;
	BCL2L1; MAPK3; KRAS; HIF1A; NOS3; PIK3C2A; PXN;
	RAF1; MAP2K2; ELAVL1; AKT1; PIK3R1; MAP2K1; SFN;
	VEGFA; AKT3; FOXO1; PRKCA
자연 살해 세포 신호전달	PRKCE; RAC1; PRKCZ; MAPK1; RAC2; PTPN11;
	KIR2DL3; AKT2; PIK3CA; SYK; PRKCI; PIK3CB;
	PIK3C3; PRKD1; MAPK3; KRAS; PRKCD; PTPN6;
	PIK3C2A; LCK; RAF1; FYN; MAP2K2; PAK4; AKT1;
	PIK3R1; MAP2K1; PAK3; AKT3; VAV3; PRKCA
세포 사이클: G1/S	HDAC4; SMAD3; SUV39H1; HDAC5; CDKN1B; BTRC;
체크포인트 조절	ATR; ABL1; E2F1; HDAC2; HDAC7A; RB1; HDAC11;
	HDAC9; CDK2; E2F2; HDAC3; TP53; CDKN1A; CCND1;
	E2F4; ATM; RBL2; SMAD4; CDKN2A; MYC; NRG1;
	GSK3B; RBL1; HDAC6
T 세포 수용체 신호전달	RAC1; ELK1; MAPK1; IKBKB; CBL; PIK3CA; FOS;
	NFKB2; PIK3CB; PIK3C3; MAPK8; MAPK3; KRAS;
	RELA; PIK3C2A; BTK; LCK; RAF1; IKBKG; RELB; FYN;
	MAP2K2; PIK3R1; CHUK; MAP2K1; NFKB1; ITK; BCL10;
	JUN; VAV3
사멸 수용체 신호전달	CRADD; HSPB1; BID; BIRC4; TBK1; IKBKB; FADD;
	FAS; NFKB2; BCL2; MAP3K14; MAPK8; RIPK1; CASP8;
	DAXX; TNFRSF10B; RELA; TRAF2; TNF; IKBKG; RELB;
	CASP9; CHUK; APAF1; NFKB1; CASP2; BIRC2; CASP3;
	BIRC3
FGF 신호전달	RAC1; FGFR1; MET; MAPKAPK2; MAPK1; PTPN11;
	AKT2; PIK3CA; CREB1; PIK3CB; PIK3C3; MAPK8;
	MAPK3; MAPK13; PTPN6; PIK3C2A; MAPK14; RAF1;
	AKT1; PIK3R1; STAT3; MAP2K1; FGFR4; CRKL; ATF4;
	AKT3; PRKCA; HGF
GM-CSF 신호전달	LYN; ELK1; MAPK1; PTPN11; AKT2; PIK3CA; CAMK2A;
	STAT5B; PIK3CB; PIK3C3; GNB2L1; BCL2L1; MAPK3;
	ETS1; KRAS; RUNX1; PIM1; PIK3C2A; RAF1; MAP2K2;

[0892]

	AKT1; JAK2; PIK3R1; STAT3; MAP2K1; CCND1; AKT3;
	STAT1
근위축성 측삭	BID; IGF1; RAC1; BIRC4; PGF; CAPNS1; CAPN2;
경화증 신호전달	PIK3CA; BCL2; PIK3CB; PIK3C3; BCL2L1; CAPN1;
	PIK3C2A; TP53; CASP9; PIK3R1; RAB5A; CASP1;
	APAF1; VEGFA; BIRC2; BAX; AKT3; CASP3; BIRC3
JAK/Stat 신호전달	PTPN1; MAPK1; PTPN11; AKT2; PIK3CA; STAT5B;
	PIK3CB; PIK3C3; MAPK3; KRAS; SOCS1; STAT5A;
	PTPN6; PIK3C2A; RAF1; CDKN1A; MAP2K2; JAK1;
	AKT1; JAK2; PIK3R1; STAT3; MAP2K1; FRAP1; AKT3;
	STAT1
니코티네이트 및 니코틴아미드	PRKCE; IRAK1; PRKAA2; EIF2AK2; GRK6; MAPK1;
대사	PLK1; AKT2; CDK8; MAPK8; MAPK3; PRKCD; PRKAA1;
	PBEF1; MAPK9; CDK2; PIM1; DYRK1A; MAP2K2;
	MAP2K1; PAK3; NT5E; TTK; CSNK1A1; BRAF; SGK
케모카인 신호전달	CXCR4; ROCK2; MAPK1; PTK2; FOS; CFL1; GNAQ;
	CAMK2A; CXCL12; MAPK8; MAPK3; KRAS; MAPK13;
	RHOA; CCR3; SRC; PPP1CC; MAPK14; NOX1; RAF1;
	MAP2K2; MAP2K1; JUN; CCL2; PRKCA
IL-2 신호전달	ELK1; MAPK1; PTPN11; AKT2; PIK3CA; SYK; FOS;
	STAT5B; PIK3CB; PIK3C3; MAPK8; MAPK3; KRAS;
	SOCS1; STAT5A; PIK3C2A; LCK; RAF1; MAP2K2;
	JAK1; AKT1; PIK3R1; MAP2K1; JUN; AKT3
장기간 시냅스	PRKCE; IGF1; PRKCZ; PRDX6; LYN; MAPK1; GNAS;
저하	PRKCI; GNAQ; PPP2R1A; IGF1R; PRKD1; MAPK3;
	KRAS; GRN; PRKCD; NOS3; NOS2A; PPP2CA;
	YWHAZ; RAF1; MAP2K2; PPP2R5C; MAP2K1; PRKCA
에스트로겐 수용체	TAF4B; EP300; CARM1; PCAF; MAPK1; NCOR2;
신호전달	SMARCA4; MAPK3; NRIP1; KRAS; SRC; NR3C1;
	HDAC3; PPARGC1A; RBM9; NCOA3; RAF1; CREBBP;
	MAP2K2; NCOA2; MAP2K1; PRKDC; ESR1; ESR2
단백질 유비퀴틴화	TRAF6; SMURF1; BIRC4; BRCA1; UCHL1; NEDD4;
경로	CBL; UBE2I; BTRC; HSPA5; USP7; USP10; FBXW7;
	USP9X; STUB1; USP22; B2M; BIRC2; PARK2; USP8;
	USP1; VHL; HSP90AA1; BIRC3
IL-10 신호전달	TRAF6; CCR1; ELK1; IKBKB; SP1; FOS; NFKB2;
	MAP3K14; MAPK8; MAPK13; RELA; MAPK14; TNF;
	IKBKG; RELB; MAP3K7; JAK1; CHUK; STAT3; NFKB1;
	JUN; IL1R1; IL6
VDR/RXR 활성화	PRKCE; EP300; PRKCZ; RXRA; GADD45A; HES1;
	NCOR2; SP1; PRKCI; CDKN1B; PRKD1; PRKCD;

[0893]

	RUNX2; KLF4; YY1; NCOA3; CDKN1A; NCOA2; SPP1;
	LRP5; CEBPB; FOXO1; PRKCA
TGF-베타 신호전달	EP300; SMAD2; SMURF1; MAPK1; SMAD3; SMAD1;
	FOS; MAPK8; MAPK3; KRAS; MAPK9; RUNX2;
	SERPINE1; RAF1; MAP3K7; CREBBP; MAP2K2;
	MAP2K1; TGFBR1; SMAD4; JUN; SMAD5
톨(Toll)-유사 수용체 신호전달	IRAK1; EIF2AK2; MYD88; TRAF6; PPARA; ELK1;
	IKBKB; FOS; NFKB2; MAP3K14; MAPK8; MAPK13;
	RELA; TLR4; MAPK14; IKBKG; RELB; MAP3K7; CHUK;
	NFKB1; TLR2; JUN
p38 MAPK 신호전달	HSPB1; IRAK1; TRAF6; MAPKAPK2; ELK1; FADD; FAS;
	CREB1; DDIT3; RPS6KA4; DAXX; MAPK13; TRAF2;
	MAPK14; TNF; MAP3K7; TGFBR1; MYC; ATF4; IL1R1;
	SRF; STAT1
뉴로트로핀/TRK 신호전달	NTRK2; MAPK1; PTPN11; PIK3CA; CREB1; FOS;
	PIK3CB; PIK3C3; MAPK8; MAPK3; KRAS; PIK3C2A;
	RAF1; MAP2K2; AKT1; PIK3R1; PDPK1; MAP2K1;
	CDC42; JUN; ATF4
FXR/RXR 활성화	INS; PPARA; FASN; RXRA; AKT2; SDC1; MAPK8;
	APOB; MAPK10; PPARG; MTTP; MAPK9; PPARGC1A;
	TNF; CREBBP; AKT1; SREBF1; FGFR4; AKT3; FOXO1
장기간 시냅스	PRKCE; RAP1A; EP300; PRKCZ; MAPK1; CREB1;
상승작용	PRKCI; GNAQ; CAMK2A; PRKD1; MAPK3; KRAS;
	PRKCD; PPP1CC; RAF1; CREBBP; MAP2K2; MAP2K1;
	ATF4; PRKCA
갈슘 신호전달	RAP1A; EP300; HDAC4; MAPK1; HDAC5; CREB1;
	CAMK2A; MYH9; MAPK3; HDAC2; HDAC7A; HDAC11;
	HDAC9; HDAC3; CREBBP; CALR; CAMKK2; ATF4;
	HDAC6
EGF 신호전달	ELK1; MAPK1; EGFR; PIK3CA; FOS; PIK3CB; PIK3C3;
	MAPK8; MAPK3; PIK3C2A; RAF1; JAK1; PIK3R1;
	STAT3; MAP2K1; JUN; PRKCA; SRF; STAT1
심혈관계에서의 저산소증 신호전달	EDN1; PTEN; EP300; NQO1; UBE2I; CREB1; ARNT;
	HIF1A; SLC2A4; NOS3; TP53; LDHA; AKT1; ATM;
	VEGFA; JUN; ATF4; VHL; HSP90AA1
RXR 기능의 LPS/IL-1 매개의 억제	IRAK1; MYD88; TRAF6; PPARA; RXRA; ABCA1;
	MAPK8; ALDH1A1; GSTP1; MAPK9; ABCB1; TRAF2;
	TLR4; TNF; MAP3K7; NR1H2; SREBF1; JUN; IL1R1

[0894]

LXR/RXR 활성화	FASN; RXRA; NCOR2; ABCA1; NFKB2; IRF3; RELA;
	NOS2A; TLR4; TNF; RELB; LDLR; NR1H2; NFKB1;
	SREBF1; IL1R1; CCL2; IL6; MMP9
아밀로이드 가공	PRKCE; CSNK1E; MAPK1; CAPNS1; AKT2; CAPN2;
	CAPN1; MAPK3; MAPK13; MAPT; MAPK14; AKT1;
	PSEN1; CSNK1A1; GSK3B; AKT3; APP
IL-4 신호전달	AKT2; PIK3CA; PIK3CB; PIK3C3; IRS1; KRAS; SOCS1;
	PTPN6; NR3C1; PIK3C2A; JAK1; AKT1; JAK2; PIK3R1;
	FRAP1; AKT3; RPS6KB1
세포 사이클: G2/M DNA	EP300; PCAF; BRCA1; GADD45A; PLK1; BTRC;
손상 체크포인트	CHEK1; ATR; CHEK2; YWHAZ; TP53; CDKN1A;
조절	PRKDC; ATM; SFN; CDKN2A
심혈관계에서의 일산화질소 신호전달	KDR; FLT1; PGF; AKT2; PIK3CA; PIK3CB; PIK3C3;
	CAV1; PRKCD; NOS3; PIK3C2A; AKT1; PIK3R1;
	VEGFA; AKT3; HSP90AA1
퓨린 대사	NME2; SMARCA4; MYH9; RRM2; ADAR; EIF2AK4;
	PKM2; ENTPD1; RAD51; RRM2B; TJP2; RAD51C;
	NT5E; POLD1; NME1
cAMP-매개의 신호전달	RAP1A; MAPK1; GNAS; CREB1; CAMK2A; MAPK3;
	SRC; RAF1; MAP2K2; STAT3; MAP2K1; BRAF; ATF4
미토콘드리아 기능장애	SOD2; MAPK8; CASP8; MAPK10; MAPK9; CASP9;
	PARK7; PSEN1; PARK2; APP; CASP3
Notch 신호전달	HES1; JAG1; NUMB; NOTCH4; ADAM17; NOTCH2;
	PSEN1; NOTCH3; NOTCH1; DLL4
소포체	HSPA5; MAPK8; XBP1; TRAF2; ATF6; CASP9; ATF4;
스트레스 경로	EIF2AK3; CASP3
피리미딘 대사	NME2; AICDA; RRM2; EIF2AK4; ENTPD1; RRM2B;
	NT5E; POLD1; NME1
파킨슨의 신호전달	UCHL1; MAPK8; MAPK13; MAPK14; CASP9; PARK7;
	PARK2; CASP3
심장 및 베타 아드레날린	GNAS; GNAQ; PPP2R1A; GNB2L1; PPP2CA; PPP1CC;
신호전달	PPP2R5C
해당작용/글루코스신합성	HK2; GCK; GPI; ALDH1A1; PKM2; LDHA; HK1
인터페론 신호전달	IRF1; SOCS1; JAK1; JAK2; IFITM1; STAT1; IFIT3
소닉 헤지호그(Sonic Hedgehog) 신호전달	ARRB2; SMO; GLI2; DYRK1A; GLI1; GSK3B; DYRK1B
글리세로인지질	PLD1; GRN; GPAM; YWHAZ; SPHK1; SPHK2

[0895]

대사	
인지질 분해	PRDX6; PLD1; GRN; YWHAZ; SPHK1; SPHK2
트립토판 대사	SIAH2; PRMT5; NEDD4; ALDH1A1; CYP1B1; SIAH1
라이신 분해	SUV39H1; EHMT2; NSD1; SETD7; PPP2R5C
뉴클레오티드 절제 수복	ERCC5; ERCC4; XPA; XPC; ERCC1
경로	
전분 및 수크로스	UCHL1; HK2; GCK; GPI; HK1
대사	
아미노 당 대사	NQO1; HK2; GCK; HK1
아라키돈산	PRDX6; GRN; YWHAZ; CYP1B1
대사	
활동일주기 신호전달	CSNK1E; CREB1; ATF4; NR1D1
응고 시스템	BDKRB1; F2R; SERPINE1; F3
도파민 수용체	PPP2R1A; PPP2CA; PPP1CC; PPP2R5C
신호전달	
글루타티온 대사	IDH2; GSTP1; ANPEP; IDH1
글리세로지질 대사	ALDH1A1; GPAM; SPHK1; SPHK2
리놀레산 대사	PRDX6; GRN; YWHAZ; CYP1B1
메티오닌 대사	DNMT1; DNMT3B; AHCY; DNMT3A
피루베이트 대사	GLO1; ALDH1A1; PKM2; LDHA
아르기닌 및 프롤린	ALDH1A1; NOS3; NOS2A
대사	
에이코사노이드 신호전달	PRDX6; GRN; YWHAZ
프룩토스 및 만노스	HK2; GCK; HK1
대사	
갈락토스 대사	HK2; GCK; HK1
스틸벤, 쿠마린 및	PRDX6; PRDX1; TYR
리그닌 생합성	
항원 제시	CALR; B2M
경로	
스테로이드의 생합성	NQO1; DHCR7
부타노에이트 대사	ALDH1A1; NLGN1
시트레이트 사이클	IDH2; IDH1
지방산 대사	ALDH1A1; CYP1B1
글리세로인지질	PRDX6; CHKA
대사	
히스티딘 대사	PRMT5; ALDH1A1
이노시톨 대사	ERO1L; APEX1
사이토크롬 p450 에 의한 생체이물의 대사	GSTP1; CYP1B1

[0896]

메탄 대사	PRDX6; PRDX1
페닐알라닌 대사	PRDX6; PRDX1
프로파노에이트 대사	ALDH1A1; LDHA
셀레노아미노산	PRMT5; AHCY
대사	
스핑고지질 대사	SPHK1; SPHK2
아미노포스포네이트	PRMT5
대사	
안드로겐 및 에스트로겐	PRMT5
대사	
아스코르베이트 및 알다레이트	ALDH1A1
대사	
담즙산 생합성	ALDH1A1
시스테인 대사	LDHA
지방산 생합성	FASN
글루탐산염 수용체	GNB2L1
신호전달	
NRF2-매개의 산화	PRDX1
스트레스 반응	
벤토스 포스페이트	GPI
경로	
벤토스 및 글루쿠로네이트	UCHL1
상호전환	
레티놀 대사	ALDH1A1
리보플라빈 대사	TYR
티로신 대사	PRMT5, TYR
뉴비퀴논 생합성	PRMT5
발린, 류신 및	ALDH1A1
이소류신 분해	
글리신, 세린 및	CHKA
트레오닌 대사	
라이신 분해	ALDH1A1
통증/맛	TRPM5; TRPA1
몸종	TRPM7; TRPC5; TRPC6; TRPC1; Cnr1; cnr2; Grk2;
	Trpa1; Pomc; Cgrp; Crf; Pka; Era; Nr2b; TRPM5; Prkaca;
	Prkacb; Prkar1a; Prkar2a
미토콘드리아 기능	AIF; CytC; SMAC (Diablo); Aifm-1; Aifm-2

[0897]

발달 신경학	BMP-4; 코딘 (Chordin; Chrd); 노긴 (Noggin; Nog); WNT (Wnt2;
	Wnt2b; Wnt3a; Wnt4; Wnt5a; Wnt6; Wnt7b; Wnt8b;
	Wnt9a; Wnt9b; Wnt10a; Wnt10b; Wnt16); beta-catenin;
	Dkk-1; 프리즐드(Frizzled) 관련 단백질; Otx-2; Gbx2; FGF-8;
	Reelin; Dab1; unc-86 (Pou4f1 or Brn3a); Numb; Reln

[0898]

[0899] 또한, 본 발명의 구현예는 유전자의 녹아웃, 유전자의 증폭 및 DNA 반복부 불안정성 및 신경계 장애와 관련된 특정 돌연변이의 수복에 관한 방법 및 조성물에 관한 것이다(문헌[Robert D. Wells, Tetsuo Ashizawa, Genetic Instabilities and Neurological Diseases, Second Edition, Academic Press, Oct 13, 2011 - Medical]). 연쇄 반복(tandem repeat) 서열의 특정 양태는 20 개 초과의 인간 질병의 원인이 되는 것으로 관찰되었다(반복부 불안정성의 새로운 이해: RNA·DNA 하이브리드의 역할. 문헌[McIvor EI, Polak U, Napierala M. RNA Biol. 2010 Sep-Oct;7(5):551-8]). CRISPR-Cas 시스템은 이들 게놈 불안정성의 결함을 교정하도록 이용될 수 있다.

[0900] 본 발명의 추가의 양태는 라포라 질병(Lafora disease)과 관련이 있는 것으로 확인된 EMP2A 및 EMP2B 유전자의 결함을 교정하기 위하여 CRISPR-Cas 시스템을 사용하는 것에 관한 것이다. 라포라 질병은 청소년기에 간질성 발작으로 시작할 수 있는 진행성 간대성근경련 간질을 특징으로 하는 상염색체 열성 질환이다. 소수의 경우의 질병이 아직 확인되지 않은 유전자의 돌연변이에 의해 유발될 수 있다. 질병은 발작, 근육연축, 보행곤란, 치매 및 결국에는 사망을 야기한다. 질병 진행에 대하여 효율적인 것으로 입증된 치료법이 현재 존재하지 않는다. 또한, 간질과 관련된 다른 유전자 이상은 CRISPR-Cas 시스템에 의해 표적화될 수 있으며, 근본이 되는 유전학은

문헌[Genetics of Epilepsy and Genetic Epilepsies, edited by Giuliano Avanzini, Jeffrey L. Noebels, Mariani Foundation Paediatric Neurology:20; 2009]에 추가로 기술되어 있다.

[0901] T 세포 수용체 (TCR) 유전자를 비활성화하는데 수반되는 Sangamo BioSciences, Inc.에 양도된 미국 특허 공개 제20110158957호의 방법은 또한 본 발명의 CRISPR Cas 시스템에 대해 변형될 수 있다. 다른 예에서, 모두 글루타민 신세타제 유전자 발현 유전자를 비활성화하는데 수반되는, Sangamo BioSciences, Inc.에 양도된 미국 특허 공개 제20100311124호 및 Cellectis에 양도된 미국 특허 공개 제20110225664호의 방법도 또한 본 발명의 CRISPR Cas 시스템에 대해 변형될 수 있다.

[0902] 본 발명의 몇몇 추가의 양태는 토픽 세부항목 유전 장애하에 국립보건원의 웹사이트(health.nih.gov/topic/GeneticDisorders의 웹사이트)에 추가로 기술된 매우 다양한 유전 질병과 관련된 결함을 교정하는 것에 관한 것이다. 유전적 뇌 질병은 부신백질이영양증, 뇌들보 무발생, 에카르디 증후군, 알퍼스병, 알츠하이머병, 바르트 증후군, 배튼병, CADASIL, 소뇌변성, 파브리병, 게르스트만 슈투로이슬러 샤잉커 병, 헌팅톤병 및 기타 3중 반복 장애, 라이병, 레슈-니한 증후군, 멘케스 질병, 사립체성 근병증 및 NINDS 거대후두각을 포함할 수 있으나 이들에 한정되지 않는다. 이들 질병은 세부항목 유전 뇌 장애하에 국립보건원의 웹사이트에 추가로 기술되어 있다.

[0903] 일부 구현예에서, 질환은 신생물일 수 있다. 일부 구현예에서, 질환이 신생물인 경우, 표적화될 유전자는 표 A에 열거된 것들 중 임의의 것이다(이러한 경우에, PTEN 등). 일부 구현예에서, 질환은 연령-관련 황반 변성일 수 있다. 일부 구현예에서, 질환은 정신분열 장애일 수 있다. 일부 구현예에서, 질환은 트리뉴클레오티드 반복 장애일 수 있다. 일부 구현예에서, 질환은 유약 X 증후군일 수 있다. 일부 구현예에서, 질환은 세크레타제 관련 장애일 수 있다. 일부 구현예에서, 질환은 프리온-관련 장애일 수 있다. 일부 구현예에서, 질환은 ALS일 수 있다. 일부 구현예에서, 질환은 약물 중독일 수 있다. 일부 구현예에서, 질환은 자폐증일 수 있다. 일부 구현예에서, 질환은 알츠하이머병일 수 있다. 일부 구현예에서, 질환은 염증일 수 있다. 일부 구현예에서, 질환은 파킨슨병일 수 있다.

[0904] 예를 들어, 미국 특허 공개 제20110023145는 자폐 범주성 장애(ASD)와 관련된 세포, 동물 및 단백질을 유전자 변형시키기 위한 징크 핑거 뉴클레아제의 사용을 설명한다. 자폐 범주성 장애(ASD)는 사회적 상호작용 및 소통의 정량적 손상, 및 제한된 반복적이며 정형화된 행동, 관심 및 활동 패턴을 특징으로 하는 일군의 장애이다. 세 가지 장애, 즉 자폐, 아스퍼거 증후군(AS) 및 비전형 전반적 발달 장애(PDD-NOS)가 중증도의 정도, 관련된 지적 기능 및 의학적 상태가 다른 연속된 동일한 장애이다. ASD는 대개 유전적으로 결정되는 장애이며, 유전성은 약 90%이다.

[0905] US 특허 공개 제 20110023145호는 본 발명의 CRISPR Cas 시스템에 적용될 수 있는 ASD와 관련된 단백질을 암호화하는 어떤 염색체 서열의 편집을 포함한다. ASD와 관련된 단백질은 통상적으로 ASD와 관련관 단백질과 ASD의 유발 또는 징후의 실험적 연관성에 기초하여 선택된다. 예를 들어, ASD와 관련된 단백질의 혈중 농도나 생성 비율은 ASD를 결여한 집단에 비해 ASD를 가진 집단에서 상승되거나 억제될 수 있다. 단백질 수준의 차이는, 제한은 아니지만, 웨스턴 블롯, 면역조직화학적 염색, 효소 결합 면역흡착 분석(ELISA), 및 질량분광법을 포함하는 프로테오믹 기술을 사용하여 평가될 수 있다. 또는 달리, ASD와 관련된 단백질은, 제한은 아니지만, DNA 마이크로어레이 분석, 연속 유전자 발현 분석(SAGE), 및 정량적 실시간 폴리머라제 연쇄 반응(Q-PCR)을 포함하는 게놈 기술을 사용하여 단백질을 암호화하는 유전자의 유전자 발현 프로파일을 얻음으로써 확인될 수 있다.

[0906] ASD와 관련된 단백질과 관련될 수 있는 질환 상태 또는 장애의 비제한적 예들은 자폐, 아스퍼거 증후군(AS), 비전형 전반적 발달 장애(PDD-NOS), 레트 증후군, 결절성 경화증, 페닐케톤뇨증, 스미스-렘리-오피츠 증후군 및 취약 X 증후군을 포함한다. 비제한적 예로서, ASD와 관련된 단백질은 제한은 아니지만 다음의 단백질들을 포함한다: ATP10C 아미노인지질-MET MET 수용체 수송 ATP 분해효소 티로신 키나제(ATP10C) BZRAP1 MGLUR5(GRM5) 대사성 글루타메이트 수용체 5(MGLUR5) CDH10 캐드헤린-10 MGLUR6(GRM6) 대사성 글루타메이트 수용체 6(MGLUR6) CDH9 캐드헤린-9 NLGN1 뉴로리긴-1 CNTN4 컨택틴-4 NLGN2 뉴로리긴-2 CNTNAP2 컨택틴-관련 SEMA5A 뉴로리긴-3 단백질-유사 2(CNTNAP2) DHCR7 7-디하이드로콜레스테롤 NLGN4X 뉴로리긴-4 X- 환원효소(DHCR7) 결합 DOC2A 이중 C2-유사 도메인-NLGN4Y 뉴로리긴-4 Y-함유 단백질 알파 결합 DPP6 디펩티딜 NLGN5 뉴로리긴-5 아미노펩티다제-유사 단백질 6 EN2 분절형 2(EN2) NRCAM 뉴런 세포 부착 분자(NRCAM) MDGA2 취약 X 지적 장애 NRXN1 뉴렉신-11(MDGA2) FMR2(AFF2) AF4/FMR2 패밀리 멤버 2 OR4M2 올팩토리 수용체(AFF2) 4M2 FOXP2 포크헤드 박스 단백질 P2 OR4N4 올팩토리 수용체(FOXP2) 4N4 FXR1 취약 X 정신 OXTR 옥시토신 수용체 발달지연, 상염색체(OXTR) 상동체 1 (FXR1) FXR2 취약 X 정신 PAH 페닐알라닌 발달지연, 상염색체 하이드록실라제(PAH) 상동체 2(FXR2)

GABRA1 감마-아미노부티르산 PTEN 포스파타제 및 수용체 서브유닛 알파-1 텐신 상동성 (GABRA1)(PTEN) GABRA5 GABAA(.감마.-아미노부티르 PTPRZ1 수용체-타입 산) 수용체 알파 5 티로신-단백질 서브유닛 (GABRA5) 포스파타제 제타(PTPRZ1) GABRB1 감마-아미노부티르산 RELN 렐린 수용체 서브유닛 베타-1(GABRB1) GABRB3 GABAA(.감마.-아미노부티르 RPL10 60S 리보솜 산) 수용체 .베타.3 서브유닛 단백질 L10(GABRB3)GABRG1 감마-아미노부티르산 SEMA5A 세마포린-5A 수용체 서브유닛 감마-1(SEMA5A) (GABRG1) HIRIP3 HIRA-상호작용 단백질 3 SEZ6L2 발작 관련 6 상동체(마우스)-유사 2 HOXA1 호메오박스 단백질 Hox-A1 SHANK3 SH3 및 다중 (HOXA1) 안키린 반복부 도메인 3(SHANK3) IL6 인터류킨-6 SHBZRAP1 SH3 및 다중 안키린 반복부 도메인 3(SHBZRAP1) LAMB1 라미닌 서브유닛 베타-1 SLC6A4 세로토닌 (LAMB1) 수송인자(SERT) MAPK3 미토겐-활성화 단백질 TAS2R1 테이스트 수용체 키나제 3 타입 2 멤버 1 TAS2R1 MAZ Myc-관련 징크 핑거 TSC1 결절성 경화증 단백질 단백질 1 MDGA2 MAM 도메인 함유 TSC2 결절성 경화증 글리코실포스파티딜이노시톨 단백질 2 앵커 2(MDGA2) MECP2 메틸 CpG 결합 UBE3A 유비퀴틴 단백질 단백질 2(MECP2) 리가아제 E3A (UBE3A) MECP2 메틸 CpG 결합 WNT2 무익기형-타입 단백질 2(MECP2) MMTV 통합 부위 패밀리, 멤버 2(WNT2).

[0907] 염색체 서열이 편집된 ASD와 관련된 단백질의 동일성은 다양할 수 있으며 다양할 것이다. 바람직한 구현예에서, 염색체 서열이 편집된 ASD와 관련된 단백질은 BZRAP 1 유전자에 의해서 암호화된 벤조디아자핀 수용체(말초) 관련 단백질 1 (BZRAP1), AFF2 유전자에 의해서 암호화된 AF4/FMR2 패밀리 멤버 2 단백질(AFF2) (MFR2라고도 한다), FXR1 유전자에 의해서 암호화된 취약 X 지적 장애 상염색체 상동체 1 단백질(FXR1), FXR2 유전자에 의해서 암호화된 취약 X 지적 장애 상염색체 상동체 2 단백질(FXR2), MDGA2 유전자에 의해서 암호화된 MAM 도메인 함유 글리코실포스파티딜이노시톨 앵커 2 단백질(MDGA2), MECP2 유전자에 의해서 암호화된 메틸 CpG 결합 단백질 2(MECP2), MGLUR5-1 유전자에 의해서 암호화된 대사성 글루타메이트 수용체 5 (MGLUR5) (GRM5라고도 한다), NRXN1 유전자에 의해서 암호화된 뉴렉신 1 단백질, 또는 SEMA5A 유전자에 의해서 암호화된 세마포린-5A 단백질(SEMA5A)일 수 있다. 예시적인 구현예에서, 유전자 변형된 동물은 래트이며, ASD와 관련된 단백질을 암호화하는 편집된 염색체 서열은 아래 열거된 것과 같다: BZRAP1 벤조디아자핀 수용체 XM_002727789, (말초) 관련 XM_213427, 단백질 1 (BZRAP1) XM_002724533, XM_001081125 AFF2(FMR2) AF4/FMR2 패밀리 멤버 2 XM_219832, (AFF2) XM_001054673 FXR1 취약 X 정신 NM_001012179 발달지연, 상염색체 상동체 1 (FXR1) FXR2 취약 X 정신 NM_001100647 발달지연, 상염색체 상동체 2(FXR2) MDGA2 MAM 도메인 함유 NM_199269 글리코실포스파티딜이노시톨 앵커 2(MDGA2) MECP2 메틸 CpG 결합 NM_022673 단백질 2(MECP2) MGLUR5 대사성 글루타메이트 NM_017012(GRM5) 수용체 5 (MGLUR5) NRXN1 뉴렉신-1 NM_021767 SEMA5A 세마포린-5A (SEMA5A) NM_001107659

[0908] 예시적인 동물 또는 세포는 ASD와 관련된 단백질을 암호화하는 1, 2, 3, 4, 5, 6, 7, 8, 9 또는 그 이상의 비활성화된 염색체 서열 및 ASD와 관련된 단백질을 암호화하는 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 또는 그 이상의 염색체 통합 서열을 포함할 수 있다. 편집되거나 통합된 염색체 서열은 ASD와 관련된 변경된 단백질을 암호화하도록 변형될 수 있다. ASD와 관련된 단백질에서 돌연변이의 비제한적 예들은 위치 18의 류신이 글루타민으로 치환된 뉴렉신 1의 L18Q 돌연변이, 위치 451의 아르기닌이 시스테인으로 치환된 뉴로리긴 3의 R451C 돌연변이, 위치 87의 아르기닌이 트립토판으로 치환된 뉴로리긴 4의 R87W 돌연변이, 및 위치 425의 이소류신이 발린으로 치환된 세로토닌 수송인자의 I425V 돌연변이를 포함한다. ASD-관련 염색체 서열에서 다수의 다른 돌연변이와 염색체 재배열이 ASD와 관련되었으며 본 분야에 알려져 있다. 예를 들어, 문헌[Freitag et al. (2010) Eur. Child. Adolesc. Psychiatry 19:169-178, 및 Bucan et al. (2009) PLoS Genetics 5: e1000536]를 참조하며, 이것의 개시는 그 전체가 여기 참고로 포함된다.

[0909] 파킨슨병과 관련된 단백질의 예는 α-시누클레인, DJ-1, LRRK2, PINK1, Parkin, UCHL1, 신필린(Synphilin)-1 및 NURR1을 포함하나 이들에 한정되지 않는다.

[0910] 중독 관련 단백질의 예는 예를 들어, ABAT를 포함할 수 있다.

[0911] 염증-관련 단백질의 예는 예를 들어, Ccr2 유전자에 의해 인코딩된 단핵구 화학주성 단백질-1(MCP1), Ccr5 유전자에 의해 인코딩된 C-C 케모카인 수용체 5형(CCR5), Fcgr2b 유전자에 의해 인코딩된 IgG 수용체 IIB(FCGR2b, CD32로도 명명) 또는 Fcer1g 유전자에 의해 인코딩된 Fc 엡실론(epsilon) R1g(FCER1g) 단백질을 포함할 수 있다.

[0912] 심혈관 질병 관련 단백질의 예는 예를 들어, IL1B(인터류킨 1, 베타), XDH(잔틴 데하이드로게나제), TP53(종양 단백질 p53), PTGIS(프로스타글란딘 I2(프로스타사이클린(prostacyclin)) 신타제), MB(미오글로빈), IL4(인터류킨 4), ANGPT1(안지오포이에틴 1), ABCG8(ATP-결합 카세트, 하위-과 G(WHITE), 구성원 8) 또는 CTSK(카텝신 K)를 포함할 수 있다.

[0913] 예를 들어, 미국 특허 공개 제20110023153은 알츠하이머병과 관련된 세포, 동물 및 단백질을 유전자 변형시키기 위한 징크 핑거 뉴클레아제의 사용을 설명한다. 일단 변형된 세포 및 동물은 AD의 연구에 통상 사용되는 척도를 사용하여 AD의 발생 및/또는 진행에 대한 표적화된 돌연변이의 영향을 연구하기 위한 공지된 방법을 사용하여 더 시험될 수 있으며, 이들은 제한은 아니지만 예컨대 학습 및 기억, 불안, 우울감, 중독, 및 감각 운동 기능뿐만 아니라 행동, 기능, 병리, 대사 및 생화학적 기능을 측정하는 분석 등이다.

[0914] 본 개시는 AD 와 관련된 단백질을 암호화하는 어떤 염색체 서열의 편집을 포함한다. AD-관련 단백질은 통상적으로 AD-관련 단백질과 AD 장애의 실험적 연관성에 기초하여 선택된다. 예를 들어, AD-관련 단백질의 혈중 농도나 생성 비율은 AD 장애를 결여한 집단에 비해 AD 장애를 가진 집단에서 상승되거나 억제될 수 있다. 단백질 수준의 차이는, 제한은 아니지만, 웨스턴 블롯, 면역조직화학적 염색, 효소 결합 면역흡착 분석 (ELISA), 및 질량분광법을 포함하는 프로테오믹 기술을 사용하여 평가될 수 있다. 또는 달리, AD-관련 단백질은, 제한은 아니지만, DNA 마이크로어레이 분석, 연속 유전자 발현 분석 (SAGE), 및 정량적 실시간 폴리머라제 연쇄 반응 (Q-PCR)을 포함하는 게놈 기술을 사용하여 단백질을 암호화하는 유전자의 유전자 발현 프로파일을 얻음으로써 확인될 수 있다.

[0915] 알츠하이머병 관련 단백질의 예는 예를 들어, VLDLR 유전자에 의해 인코딩되는 극저밀도 리포단백질 수용체 단백질(VLDLR), UBA1 유전자에 의해 인코딩되는 유비퀴틴-유사 변형 활성화 효소(UBA1) 또는 UBA3 유전자에 의해 인코딩되는 NEDD8-활성화 효소 E1 촉매 서브유닛 단백질(UBE1C)을 포함할 수 있다.

[0916] 비제한적 예로서, AD와 관련된 단백질은 제한은 아니지만 다음에 열거된 단백질들을 포함한다: 염색체 서열 암호화된 단백질 ALAS2 델타-아미노레불리네이트 신타제 2(ALAS2) ABCA1 ATP-결합 카세트 수송인자(ABCA1) ACE 안지오텐신 I-전환 효소 (ACE) APOE 아포지방단백질 E 전구체 (APOE) APP 아밀로이드 전구체 단백질(APP) AQP1 아쿠아포린 1 단백질(AQP1) BIN1 Myc 박스-의존성-상호작용 단백질 1 또는 브릿지 통합인자 1 단백질(BIN1) BDNF 뇌-유래 신경영양 인자(BDNF) BTNL8 부티로필린-유사 단백질 8 (BTNL8) C1ORF49 염색체 1 오픈 리딩 프레임 49 CDH4 캐드헤린-4 CHRNB2 뉴런 아세틸콜린 수용체 서브유닛 베타-2 CKLFSF2 CKLF-유사 MARVEL 경막 도메인-함유 단백질 2(CKLFSF2) CLEC4E C-타입 렉틴 도메인 패밀리 4, 멤버 (CLEC4E) CLU 클러스테린 단백질(아포지방단백질 J로도 알려져 있다) CR1 적혈구 보체 수용체 1(CR1, CD35, C3b/C4b 수용체 및 면역 부착 수용체라고도 한다) CR1L 적혈구 보체 수용체 1(CR1L) CSF3R 과립구 콜로니-자극 인자 3 수용체 (CSF3R) CST3 시스타틴 C 또는 시스타틴 3 CYP2C 시토크롬 P450 2C DAPK1 사망-관련 단백질 키나제 1 (DAPK1) ESR1 에스트로겐 수용체 1 IgA의 FCAR Fc 단편 수용체 (FCAR, CD89라고도 한다) IgG의 FCGR3B Fc 단편, 저 친화성 IIIb, 수용체 (FCGR3B 또는 CD16b) FFA2 자유 지방산 수용체 2(FFA2) FGA 피브리노겐 (인자 I) GAB2 GRB2-관련-결합 단백질 2(GAB2) GAB2 GRB2-관련-결합 단백질 2(GAB2) GALP 갈라닌-유사 펩티드 GAPDHS 글리세르알데하이드-3-포스페이트 디하이드로게나제, 정자형성 (GAPDHS) GMPB GMBP HP 햅토글로빈 (HP) HTR7 5-하이드록시 트립타민 (세로토닌) 수용체 7 (아데닐레이트 시클라제-결합) IDE 인슐린 분해 효소 IF127 IF127 IFI6 인터페론, 알파-유도성 단백질 6 (IFI6) IFIT2 인터페론-유도 단백질 테트라트리코펩티드 반복부 2(IFIT2) IL1RN 인터류킨-1 수용체 길항제(IL-1RA) IL8RA 인터류킨 8 수용체, 알파 (IL8RA 또는 CD181) IL8RB 인터류킨 8 수용체, 베타 (IL8RB) JAG1 재그드 1 (JAG1) KCNJ15 칼륨 내부-정류 채널, 서브패밀리 J, 멤버 15 (KCNJ15) LRP6 저밀도 지방단백질 수용체-관련 단백질 6 (LRP6) MAPT 미세소관-관련 단백질 tau (MAPT) MARK4 MAP/미세소관 친화성-조절 키나제 4 (MARK4) MPHOSPH1 M-상 포스포단백질 1 MTHFR 5,10-메틸렌테트라하이드로폴레이트 환원효소 MX2 인터페론-유도 GTP-결합 단백질 Mx2 NBN 니브린, NBN NCSTN 니카스트린이라고도 함 NIACR2 니아신 수용체 2(NIACR2, GPR109B라고도 함) NMNAT3 니코틴아미드 뉴클레오티드 아데닐일트랜스페라제 3 NTM 뉴로트리민 (또는 HNT) ORM1 오로스뮤코이드 1 (ORM1) 또는 알파-1-산 글리코단백질 1 P2RY13 P2Y 퓨리노셉터 13 (P2RY13) PBEF1 니코틴아미드 포스포리보실트랜스페라제(NAmPRTase 또는 Nampt) 프레-B-세포 콜로니-증진 인자 1 (PBEF1) 또는 비스파틴 PCK1 포스포엔올피루베이트 카복시 키나제라고도 함 PICALM 포스파티딜이노시톨 결합 클라스린 조립체 단백질(PICALM) PLAU Uro 키나제-타입 플라스미노겐 활성인자(PLAU) PLXNC1 플렉신 C1 (PLXNC1) PRNP 프리온 단백질 PSEN1 프레세닐린 1 단백질(PSEN1) PSEN2 프레세닐린 2 단백질(PSEN2) PTPRA 단백질 티로신 포스파타제 수용체 타입 A 단백질 (PTPRA) RALGPS2 Ral GEF PH 도메인 및 SH3 결합 모티프 2(RALGPS2) RGSL2 조절인자 G-단백질 신호화 유사 2(RGSL2) SELENBP1 셀레늄 결합 단백질 1 (SELNBP1) SLC25A37 미토페린-1 SORL1 소틸린-관련 수용체 L(DLR 부류) A 반복부-함유 단백질(SORL1) TF 트랜스페린 TFAM 미토콘드리아 전사 인자 A TNF 종양 괴사 인자 TNFRSF10C 종양 괴사 인자 수용체 수퍼패밀리 멤버 10C (TNFRSF10C) TNFSF10 종양 괴사 인자 수용체 수퍼패밀리, (TRAIL) 멤버 10a (TNFSF10) UBA1 유비퀴틴-유사 변형인자 활성화 효소 1 (UBA1) UBA3 NEDD8-활성화 효소 E1 촉매 서브유닛 단백질(UBE1C) UBB 유비퀴틴 B 단백질(UBB) UBQLN1 유비퀼린-1 UCHL1 유비퀴틴 카복실-말단 에스테라제 L1

단백질(UCHL1) UCHL3 유비퀴틴 카복실-말단 하이드롤라제 이소엔자임 L3 단백질(UCHL3) VLDLR 초저밀도 지방단백질 수용체 단백질(VLDLR).

[0917] 예시적인 구현예에서, 염색체 서열이 편집된 AD와 관련된 단백질은 VLDLR 유전자에 의해서 암호화된 초저밀도 지방단백질 수용체 단백질(VLDLR), UBA1 유전자에 의해서 암호화된 유비퀴틴-유사 변형인자 활성화 효소 1(UBA1), UBA3 유전자에 의해서 암호화된 NEDD8-활성화 효소 E1 촉매 서브유닛 단백질(UBE1C), AQP1 유전자에 의해서 암호화된 아쿠아포린 1 단백질(AQP1), UCHL1 유전자에 의해서 암호화된 유비퀴틴 카복실-말단 에스테라제 L1 단백질(UCHL1), UCHL 3 유전자에 의해서 암호화된 유비퀴틴 카복실-말단 하이드롤라제 이소엔자임 L3 단백질(UCHL3), UBB 유전자에 의해서 암호화된 유비퀴틴 B 단백질(UBB), MAPT 유전자에 의해서 암호화된 미세소관-관련 단백질 tau (MAPT), PTPRA 유전자에 의해서 암호화된 단백질 티로신 포스파타제 수용체 타입 A 단백질(PTPRA), PICALM 유전자에 의해서 암호화된 포스파티딜이노시톨 결합 클라스린 조립체 단백질(PICALM), CLU 유전자에 의해서 암호화된 클러스테린 단백질(아포지방단백질 J로도 알려져 있다), PSEN1 유전자에 의해서 암호화된 프레세닐린 1 단백질, PSEN2 유전자에 의해서 암호화된 프레세닐린 2 단백질, SORL1 유전자에 의해서 암호화된 소틸린-관련 수용체 L (DLR 부류) A 반복부-함유 단백질(SORL1) 단백질, APP 유전자에 의해서 암호화된 아밀로이드 전구체 단백질(APP), APOE 유전자에 의해서 암호화된 아포지방단백질 E 전구체 (APOE), 또는 BDNF 유전자에 의해서 암호화된 뇌-유래 신경영양 인자(BDNF)일 수 있다. 예시적인 구현예에서, 유전자 변형된 동물은 래트이며, AD와 관련된 단백질을 암호화하는 편집된 염색체 서열은 다음과 같다: APP 아밀로이드 전구체 단백질(APP) NM_019288 AQP1 아쿠아포린 1 단백질(AQP1) NM_012778 BDNF 뇌-유래 신경영양 인자 NM_012513 CLU 클러스터린 단백질(NM_053021 아포지방단백질 J로도 알려져 있다) MAPT 미세소관-관련 단백질 NM_017212 tau (MAPT) PICALM 포스파티딜이노시톨 결합 NM_053554 클라스린 조립체 단백질(PICALM) PSEN1 프레세닐린 1 단백질(PSEN1) NM_019163 PSEN2 프레세닐린 2 단백질(PSEN2) NM_031087 PTPRA 단백질 티로신 포스파타제 NM_012763 수용체 타입 A 단백질(PTPRA) SORL1 소틸린-관련 수용체 L(DLR NM_053519, 부류) A 반복부-함유 XM_001065506, 단백질(SORL1) XM_217115 UBA1 유비퀴틴-유사 변형인자 활성화 NM_001014080 효소 1 (UBA1) UBA3 NEDD8-활성화 효소 E1 NM_057205 촉매 서브유닛단백질(UBE1C) UBB 유비퀴틴 B 단백질(UBB) NM_138895 UCHL1 유비퀴틴 카복실-말단 NM_017237 에스테라제 L1 단백질(UCHL1) UCHL3 유비퀴틴 카복실-말단 NM_001110165 하이드롤라제 이소엔자임 L3 단백질(UCHL3) VLDLR 초저밀도 지방단백질 NM_013155 수용체 단백질(VLDLR).

[0918] 동물 또는 세포는 AD와 관련된 단백질을 암호화하는 1, 2, 3, 4, 5, 6, 7, 8, 9,10, 11, 12, 13, 14, 15 또는 그 이상의 파괴된 염색체 서열 및 AD와 관련된 파괴된 단백질을 암호화하는 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 또는 그 이상의 염색체 통합 서열을 포함할 수 있다.

[0919] 편집된 또는 통합된 염색체 서열은 AD와 관련된 변경된 단백질을 암호화하도록 변형될 수 있다. AD-관련 염색체 서열에서 다수의 돌연변이가 AD와 관련되었다. 예를 들어, APP에서 V7171(즉, 위치 717에서 발린이 이소류신으로 변화됨) 미스센스 돌연변이는 가족성 AD를 야기한다. 프레세닐린-1 단백질에서의 다중 돌연변이, 예컨대 H163R (즉, 위치 163에서 히스티딘이 아르기닌으로 변화됨), A246E (즉, 위치 246에서 알라닌이 글루타메이트로 변화됨), L286V(즉, 위치 286에서 류신이 발린으로 변화됨) 및 C410Y (즉, 위치 410에서 시스테인이 티로신으로 변화됨)는 가족성 알츠하이머병 타입 3을 야기한다. 프레세닐린-2 단백질에서의 돌연변이, 예컨대 N141 I(즉, 위치 141에서 아스파라긴이 이소류신으로 변화됨), M239V (즉, 위치 239에서 메티오닌이 발린으로 변화됨), 및 D439A (즉, 위치 439에서 아스파르테이트가 알라닌으로 변화됨)는 가족성 알츠하이머 타입 4를 야기한다. AD-관련 유전자 및 질환에서 유전자 변이체의 다른 관련성도 본 분야에 잘 알려져 있다. 예를 들어, 문헌[Waring et al. (2008) Arch. Neurol. 65:329-334]를 참조하며, 이것의 개시는 그 전체가 여기 참고로 포함된다.

[0920] 자폐 스펙트럼 장애와 관련된 단백질의 예는 예를 들어, BZRAP1 유전자에 의해 인코딩되는 벤조디아자핀 수용체 (주변) 관련 단백질 1(BZRAP1), AFF2 유전자(MFR2로도 명명)에 의해 인코딩되는 AF4/FMR2 과 구성원 2 단백질 (AFF2), FXR1 유전자에 의해 인코딩되는 유약 X 정신 지체 상염색체 상동체 1 단백질(FXR1) 또는 FXR2 유전자에 의해 인코딩되는 유약 X 정신 지체 상염색체 상동체 2 단백질(FXR2)을 포함할 수 있다.

[0921] 황반 변성과 관련된 단백질의 예는 예를 들어, ABCR 유전자에 의해 인코딩되는 ATP-결합 카세트, 하위과 A(ABC1) 구성원 4 단백질(ABCA4), APOE 유전자에 의해 인코딩되는 아포리포단백질 E 단백질(APOE) 또는 CCL2 유전자에 의해 인코딩되는 케모카인(C-C 모티프) 리간드 2 단백질(CCL2)을 포함할 수 있다.

[0922] 정신분열증 관련 단백질의 예는 NRG1, ErbB4, CPLX1, TPH1, TPH2, NRXN1, GSK3A, BDNF, DISC1, GSK3B 및 그들의 조합을 포함할 수 있다.

[0923] 종양 저해에 수반되는 단백질의 예는 예를 들어, ATM(돌연변이된 혈관확장성 운동실조증), ATR(혈관확장성 운동

실조증 및 Rad3 관련), EGFR(상피 성장 인자 수용체), ERBB2(v-erb-b2 적혈모구성 백혈병 바이러스 암유전자 상동체 2), ERBB3(v-erb-b2 적혈모구성 백혈병 바이러스 암유전자 상동체 3), ERBB4(v-erb-b2 적혈모구성 백혈병 바이러스 암유전자 상동체 4), Notch 1, Notch2, Notch 3 또는 Notch 4를 포함할 수 있다.

[0924] 세크레타제 장애와 관련된 단백질의 예는 예를 들어, PSENEN(프레세닐린 인핸서 2 상동체(C. 엘레간스(C. elegans)), CTSB(카텝신 B), PSEN1(프레세닐린 1), APP(아밀로이드 베타(A4) 전구체 단백질), APH1B(앞인두 결함 1 상동체 B(C. 엘레간스)), PSEN2(페레세닐린 2(알츠하이머병 4)) 또는 BACE1(베타-부위 APP-절단 효소 1)을 포함할 수 있다.

[0925] 예를 들어, 미국 특허 공개 제20110023146은 세크레타제-관련 장애와 관련된 세포, 동물 및 단백질을 유전자 변형시키기 위한 징크 핑거 뉴클레아제의 사용을 설명한다. 세크레타제는 예비-단백질을 이들의 생물학적 활성 형태로 가공하는데 필수적이다. 세크레타제 경로의 다양한 구성성분에서 결함은 많은 장애들, 특히 특징적 아밀로이드형성 또는 아밀로이드반을 가진 것들, 예컨대 알츠하이머병 (AD)에 기여한다.

[0926] 세크레타제 장애 및 이들 장애와 관련된 단백질은 수많은 장애에 대한 감수성, 장애의 존재, 장애의 중증도, 또는 이들의 어떤 조합에 영향을 미치는 다양한 세트의 단백질이다. 본 개시는 세크레타제 장애와 관련된 단백질을 암호화하는 어떤 염색체 서열의 편집을 포함한다. 세크레타제 장애와 관련된 단백질은 통상적으로 세크레타제-관련 단백질과 세크레타제 장애의 발생의 실험적 연관성에 기초하여 선택된다. 예를 들어, 세크레타제 장애와 관련된 단백질의 혈중 농도나 생성 비율은 세크레타제 장애가 없는 집단에 비해 세크레타제 장애를 가진 집단에서 상승되거나 억제될 수 있다. 단백질 수준의 차이는, 제한은 아니지만, 웨스턴 블롯, 면역조직화학적 염색, 효소 결합 면역흡착 분석 (ELISA), 및 질량분광법을 포함하는 프로테오믹 기술을 사용하여 평가될 수 있다. 또는 달리, 세크레타제 장애와 관련된 단백질은, 제한은 아니지만, DNA 마이크로어레이 분석, 연속 유전자 발현 분석 (SAGE), 및 정량적 실시간 폴리머라제 연쇄 반응 (Q-PCR)을 포함하는 게놈 기술을 사용하여 단백질을 암호화하는 유전자의 유전자 발현 프로파일을 얻음으로써 확인될 수 있다.

[0927] 비제한적 예로서, 세크레타제 장애와 관련된 단백질은 PSENEN(프레세닐린 인핸서 2 상동체(예쁜꼬마선충)), CTSB (카텝신 B), PSEN1(프레세닐린 1), APP (아밀로이드 베타 (A4) 전구체 단백질), APH1B(앞인두 결함 1 상동체 B (예쁜꼬마선충)), PSEN2(프레세닐린 2(알츠하이머병 4)), BACE1 (베타-부위 APP-절단 효소 1), ITM2B (통합 막 단백질 2B), CTSD (카텝신 D), 노치1 (노치 상동체 1, 전위-관련 (초파리)), TNF (종양 괴사 인자 (TNF 수퍼패밀리, 멤버 2)), INS (인슐린), DYT10 (근육긴장이상 10), ADAM17 (ADAM 메탈로펩티다제 도메인 17), APOE (아포지방단백질 E), ACE(안지오텐신 I 전환 효소 (펩티딜-디펩티다제 A) 1), STN (스타틴), TP53 (종양 단백질 p53), IL6 (인터류킨 6 (인터페론, 베타 2)), NGFR (신경성장인자 수용체 (TNFR 수퍼패밀리, 멤버 16)), IL1B (인터류킨 1, 베타), ACHE (아세틸콜린에스테라제(Yt 혈액군)), CTNNB1 (카테닌 (캐드헤린-관련 단백질), 베타 1, 88kDa), IGF1 (인슐린-유사 성장인자 1 (소마토메딘 C)), IFNG (인터페론, 감마), NRG1 (뉴레굴린 1), CASP3 (카스파아제 3, 세포자멸사-관련 시스테인 펩티다제), MAPK1 (미토겐-활성화 단백질 키나제 1), CDH1 (캐드헤린 1, 타입 1, E-캐드헤린 (상피)), APBB1 (아밀로이드 베타 (A4) 전구체 단백질-결합, 패밀리 B, 멤버 1 (Fe65)), HMGCR (3-하이드록시-3-메틸글루타릴-코엔자임 A 환원효소), CREB1 (cAMP 반응성 요소 결합 단백질 1), PTGS2(프로스타글란딘-엔도퍼옥사이드 신타제 2(프로스타글란딘 G/H 신타제 및 시클로옥시게나제)), HES1(스플릿 1의 헤어리 및 인핸서, (초파리)), CAT (카탈라제), TGFB1 (형질전환 성장 인자, 베타 1), ENO2 (에놀라제 2(감마, 뉴런)), ERBB4(v-erb-a 적아세포 백혈병 바이러스 암유전자 상동체 4 (조류)), TRAPPC10 (트래픽킹 단백질 입자 복합체 10), MAOB(모노아민 옥시다제 B), NGF(신경성장인자(베타 폴리펩티드)), MMP12 (바탕질 메탈로펩티다제 12(대식세포 엘라스타제)), JAG1 (재그드 1 (알라질 증후군)), CD40LG (CD40 리간드), PPARG (퍼옥시솜 증식인자-활성화 수용체 감마), FGF2(섬유아세포 성장 인자 2(염기성)), IL3 (인터류킨 3 (콜로니-자극 인자, 다중)), LRP1 (저 밀도 지방단백질 수용체-관련 단백질 1), 노치4(노치 상동체 4 (초파리)), MAPK8 (미토겐-활성화 단백질 키나제 8), PREP (프롤릴 엔도펩티다제), 노치3 (노치 상동체 3 (초파리)), PRNP (프리온 단백질), CTSG (카텝신 G), EGF (상피성장인자(베타-유로가스트론)), REN (레닌), CD44 (CD44 분자(인디언 혈액군)), SELP (셀렉틴 P (과립 막 단백질 140 kDa, 항원 CD62)), GHR (성장 호르몬 수용체), ADCYAP1 (아데닐레이트 시클라제 활성화 폴리펩티드 1 (뇌하수체)), INSR (인슐린 수용체), GFAP(신경교 섬유질 산성 단백질), MMP3 (바탕질 메탈로펩티다제 3 (스트로멜리신 1, 프로젤라티나제)), MAPK10 (미토겐-활성화 단백질 키나제 10), SP1 (Sp1 전사 인자), MYC (v-myc 골수세포증 바이러스 암유전자 상동체 (조류)), CTSE (카텝신 E), PPARA (퍼옥시솜 증식인자-활성화 수용체 알파), JUN (jun 암유전자), TIMP1 (TIMP 메탈로펩티다제 저해제 1), IL5 (인터류킨 5 (콜로니-자극 인자, 호산구)), IL1A (인터류킨 1, 알파), MMP9 (바탕질 메탈로펩티다제 9 (젤라티나제 B, 92 kDa 젤라티나제, 92 kDa 타입 IV 콜라게나제)), HTR4(5-하이드록시트립타민 (세로토닌) 수용체

4), HSPG2(헤파란 설페이트 프로테오글리칸 2), KRAS (v-Ki-ras2 Kirsten 래트 육종 바이러스 암유전자 상동체), CYCS (시토크롬 c, 체세포), SMG1(SMG1 상동체, 포스파티딜이노시톨 3-키나제-관련 키나제(예쁜꼬마선충)), IL1R1 (인터류킨 1 수용체, 타입 I), PROK1 (프로키네티신 1), MAPK3 (미토겐-활성화 단백질 키나제 3), NTRK1(신경영양 티로신 키나제, 수용체, 타입 1), IL13(인터류킨 13), MME(막 메탈로-엔도펩티다제), TKT (트랜스케톨라제), CXCR2(케모카인 (C-X-C 모티프) 수용체 2), IGF1R(인슐린-유사 성장인자 1 수용체), RARA (레티노산 수용체, 알파), CREBBP (CREB 결합 단백질), PTGS1 (프로스타글란딘-엔도퍼옥사이드 신타제 1(프로스타글란딘 G/H 신타제 및 시클로옥시게나제)), GALT (갈락토오스-1-포스페이트 유리딜일트랜스페라제), CHRM1 (콜린성 수용체, 무스카리닉 1), ATXN1 (아탁신 1), PAWR (PRKC, 세포자멸사, WT1, 조절인자), 노치2(노치 상동체 2(초파리)), M6PR(만노오스-6-포스페이트 수용체 (양이온 의존성)), CYP46A1 (시토크롬 P450, 패밀리 46, 서브패밀리 A, 폴리펩티드 1), CSNK1 D(카제인 키나제 1, 델타), MAPK14 (미토겐-활성화 단백질 키나제 14), PRG2(프로테오글리칸 2, 골수(자연 사살 세포 활성인자, 호산구 과립 주 염기성 단백질)), PRKCA (단백질 키나제 C, 알파), L1 CAM(L1 세포 부착 분자), CD40 (CD40 분자, TNF 수용체 수퍼패밀리 멤버 5), NR1I2(핵 수용체 서브패밀리 1, 그룹 I, 멤버 2), JAG2(재그드 2), CTNND1 (카테닌 (캐드헤린-관련 단백질), 델타 1), CDH2(캐드헤린 2, 타입 1, N-캐드헤린 (뉴런)), CMA1(키마아제 1, 비만 세포), SORT1 (소틸린 1), DLK1 (델타-유사 1 상동체 (초파리)), THEM4(티오에스터라제 수퍼패밀리 멤버 4), JUP (접합부 플라코글로빈), CD46 (CD46 분자, 보체 조절 단백질), CCL11 (케모카인 (C-C 모티프) 리간드 11), CAV3(카베올린 3), RNASE3 (리보뉴클레아제, RNase A 패밀리, 3 (호산구 양이온 단백질)), HSPA8 (열충격 70kDa 단백질 8), CASP9 (카스파아제 9, 세포자멸사-관련 시스테인 펩티다제), CYP3A4 (시토크롬 P450, 패밀리 3, 서브패밀리 A, 폴리펩티드 4), CCR3 (케모카인 (C-C 모티프) 수용체 3), TFAP2A (전사 인자 AP-2 알파 (활성화 인핸서 결합 단백질 2 알파)), SCP2(스테롤 캐리어 단백질 2), CDK4 (사이클린-의존성 키나제 4), HIF1A (저산소증 유도성 인자 1, 알파 서브유닛 (염기성 헬릭스-루프-헬릭스 전사 인자)), TCF7L2(전사 인자 7-유사 2(T-세포 특이적, HMG-박스)), IL1R2(인터류킨 1 수용체, 타입 II), B3GALTL (베타 1,3-갈락토실트랜스페라제-유사), MDM2(Mdm2 p53 결합 단백질 상동체 (마우스)), RELA (v-rel 망내증 바이러스 암유전자 상동체 A (조류)), CASP7 (카스파아제 7, 세포자멸사-관련 시스테인 펩티다제), IDE (인슐린-분해 효소), FABP4 (지방산 결합 단백질 4, 지방세포), CASK (칼슘/칼모듈린-의존성 세린 단백질 키나제(MAGUK 패밀리)), ADCYAP1R1 (아데닐레이트 시클라제 활성화 폴리펩티드 1 (뇌하수체) 수용체 타입 I), ATF4 (활성화 전사 인자 4 (탁스-반응성 인핸서 요소 B67)), PDGFA(혈소판-유래 성장인자 알파 폴리펩티드), C21 또는 f33 (염색체 21 오픈 리딩 프레임 33), SCG5 (세크레토그라닌 V (7B2 단백질)), RNF123 (링 핑거 단백질 123), NFKB1 (B-세포의 카파 경쇄 폴리펩티드 유전자 인핸서의 핵 인자 1), ERBB2(v-erb-b2 적아세포 백혈병 바이러스 암유전자 상동체 2, 신경/교모세포종 유래 암유전자 상동체 (조류)), CAV1 (카베올린 1, 카베올라 단백질, 22kDa), MMP7 (바탕질 메탈로펩티다제 7 (매트릴리신, 유테린)), TGFA(형질전환 성장 인자, 알파), RXRA (레티노이드 X 수용체, 알파), STX1A (신탁신 1A (뇌)), PSMC4 (프로테아솜(프로솜, 매크로페인) 26S 서브유닛, ATP 분해효소, 4), P2RY2(퓨린 수용체 P2Y, G-단백질 결합, 2), TNFRSF21 (종양 괴사 인자 수용체 수퍼패밀리, 멤버 21), DLG1 (원판, 대형 상동체 1(초파리)), NUMBL (넘브 상동체 (초파리)-유사), SPN (시알로포린), PLSCR1(인지질 스크램블라제 1), UBQLN2(유비퀼린 2), UBQLN1(유비퀼린 1), PCSK7 (예비-단백질 컨버타제 섭틸리신/켁신 타입 7), SPON1 (스폰딘 1, 세포외 바탕질 단백질), SILV (은 상동체 (마우스)), QPCT (글루타미닐-펩티드 시클로트랜스페라제), HESS (스플릿 5의 헤어리 및 인핸서 (초파리)), GCC1 (GRIP 및 코일드-코일 도메인 함유 1), 및 이들의 어떤 조합을 포함한다.

[0928] 유전자 변형된 동물 또는 세포는 세크레타제 장애와 관련된 단백질을 암호화하는 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 또는 그 이상의 파괴된 염색체 서열 및 세크레타제 장애와 관련된 파괴된 단백질을 암호화하는 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 또는 그 이상의 염색체 통합 서열을 포함할 수 있다.

[0929] 근위축성 측삭경화증과 관련된 단백질의 예들은 SOD1(수퍼옥사이드 디스뮤타제 1), ALS2(근위축성 측삭경화증 2), FUS(육종 융합형), TARDBP(TAR DNA 결합 단백질), VAGFA(혈관 내피 성장인자 A), VAGFB(혈관 내피 성장인자 B), 및 VAGFC(혈관 내피 성장인자 C), 및 이들의 어떤 조합을 포함할 수 있다.

[0930] 예를 들어, 미국 특허 공개 제20110023144호는 근위축성 측삭경화증 (ALS) 질환과 관련된 세포, 동물 및 단백질을 유전자 변형시키기 위한 징크 핑거 뉴클레아제의 사용을 설명한다. ALS는 수의 운동에 수반되는 뇌 피질, 뇌줄기, 및 척수에서 특정 신경 세포의 점진적인 꾸준한 변성을 특징으로 한다.

[0931] 운동 뉴런 장애 및 이들 장애와 관련된 단백질은 운동 뉴런 장애 발생에 대한 감수성, 운동 뉴런 장애의 존재, 운동 뉴런 장애의 중증도, 또는 이들의 어떤 조합에 영향을 미치는 다양한 세트의 단백질이다. 본 개시는 특정한 운동 뉴런 장애인 ALS 질환과 관련된 단백질을 암호화하는 어떤 염색체 서열의 편집을 포함한다. ALS와 관련

된 단백질은 통상적으로 ALS-관련 단백질과 ALS의 실험적 연관성에 기초하여 선택된다. 예를 들어, ALS 와 관련된 단백질의 혈중 농도나 생성 비율이 ALS가 없는 집단에 비해 ALS를 가진 집단에서 상승되거나 억제될 수 있다. 단백질 수준의 차이는, 제한은 아니지만, 웨스턴 블롯, 면역조직화학적 염색, 효소 결합 면역흡착 분석(ELISA), 및 질량분광법을 포함하는 프로테오믹 기술을 사용하여 평가될 수 있다. 또는 달리, ALS 와 관련된 단백질은, 제한은 아니지만, DNA 마이크로어레이 분석, 연속 유전자 발현 분석 (SAGE), 및 정량적 실시간 폴리머라제 연쇄 반응 (Q-PCR)을 포함하는 게놈 기술을 사용하여 단백질을 암호화하는 유전자의 유전자 발현 프로파일을 얻음으로써 확인될 수 있다.

[0932] 비제한적 예로서, ALS와 관련된 단백질은, 제한은 아니지만 다음의 단백질들을 포함한다: SOD1 수퍼옥사이드 디스뮤타제 1, ALS3 근위축성 측삭 가용성 경화증 3 SETX 세나탁신 ALS5 근위축성 측삭경화증 5 FUS 융합 육종 ALS7 근위축성 측삭경화증 7 ALS2 근위축성 측삭 DPP6 디펩티딜-펩티다제 6 경화증 2 NEFH 신경세사, 중질 PTGS1 프로스타글란딘-폴리펩티드 엔도퍼옥사이드 신타제 1 SLC1A2 용질 캐리어 패밀리 1 TNFRSF10B 종양 괴사 인자(신경교 고 친화성 수용체 수퍼패밀리, 글루타메이트 수송인자), 멤버 10b 멤버 2 PRPH 페리페린 HSP90AA1 열충격 단백질 90kDa 알파 (시토졸), 부류 A 멤버 1 GRIA2 글루타메이트 수용체, IFNG 인터페론, 감마 이온성, AMPA 2 S100B S100 칼슘 결합 FGF2 섬유아세포 성장 인자 2 단백질 B AOX1 알데하이드 옥시다제 1 CS 시트레이트 신타제 TARDBP TAR DNA 결합 단백질 TXN 티오레독신 RAPH1 Ras 관련 MAP3K5 미토겐-활성화 단백질(RaIGDS/AF-6) 및 키나제 5 플렉스트린 상동체 도메인 1 NBEAL1 뉴로비친-유사 1 GPX1 글루타티온 퍼옥시다제 1 ICA1L 소섬 세포 자가항원 RAC1 ras-관련 C3 보툴리넘 1.69 kDa-유사 톡신 기질 1 MAPT 미세소관-관련 ITPR2 이노시톨 1,4,5- 단백질 tau 트리포스페이트 수용체, 타입 2 ALS2CR4 근위축성 측삭 GLS 글루타미나제 경화증 2(소아) 염색체 영역, 후보 4 ALS2CR8 근위축성 측삭 CNTFR 섬모성 신경영양 인자 경화증 2(소아) 수용체 염색체 영역, 후보 8 ALS2CR11 근위축성 측삭 FOLH1 폴레이트 하이드롤라제 1 경화증 2(소아) 염색체 영역, 후보 11 FAM117B 패밀리 서열 P4HB 프롤릴 4-하이드록실라제, 유사성 117, 멤버 B 베타 폴리펩티드 CNTF 섬모성 신경영양 인자 SQSTM1 시퀘스토솜 1 STRADB STE20-관련 키나제 NAIP NLR 패밀리, 세포자멸사 어댑터 베타 억제 단백질 YWHAQ 티로신 3- SLC33A1 용질 캐리어 패밀리 33 모노옥시게나제/트립토프(아세틸-CoA 수송인자), 5-모노옥시게나제 멤버 1 활성화 단백질, 세타 폴리펩티드 TRAK2 트래픽킹 단백질, 도 4 도 4 상동체, SAC1 키네신 결합 2 지질 포스파타제 도메인 함유 NIF3L1 NIF3 NGG1 상호작용 INA 인터넥신 뉴런 인자 3-유사 1 중간체 세사 단백질, 알파 PARD3B par-3 분배 COX8A 시토크롬 c 옥시다제 결함 3 상동체 B 서브유닛 VIIIA CDK15 사이클린-의존성 키나제 HECW1 HECT, C2 및 WW 15 도메인 함유 E3 유비퀴틴 단백질 리가아제 1 NOS1 산화질소 신타제 1 MET 메트 프로토-암유전자 SOD2 수퍼옥사이드 디스뮤타제 2, HSPB1 열충격 27 kDa 미토콘드리아 단백질 1 NEFL 신경세사, 경질 CTSB 카텝신 B 폴리펩티드 ANG 안지오게닌, HSPA8 열충격 70 kDa 리보뉴클레아제, RNase A 단백질 8 패밀리, 5 VAPB VAMP (소포- ESR1 에스트로겐 수용체 1 관련 막 단백질)-관련 단백질 B 및 C SNCA 시누클레인, 알파 HGF 간세포 성장 인자 CAT 카탈라제 ACTB 액틴, 베타 NEFM 신경세사, 중간 TH 티로신 하이드록실라제 폴리펩티드 BCL2 B-세포 CLL/림프종 2 FAS 파스(TNF 수용체 수퍼패밀리, 멤버 6) CASP3 카스파아제 3, 세포자멸사- CLU 클러스테린 관련 시스테인 펩티다제 SMN1 생존 운동 뉴런 G6PD 글루코오스-6-포스페이트 1, 텔로머디하이드로게나제 BAX BCL2-관련 X HSF1 열충격 전사 단백질 인자 1 RNF19A 링 핑거 단백질 19A JUN 준 암유전자 ALS2CR12 근위축성 측삭 HSPA5 열충격 70kDa 경화증 2(소아) 단백질 5 염색체 영역, 후보 12 MAPK14 미토겐-활성화 단백질 IL10 인터류킨 10 키나제 14 APEX1 APEX 뉴클레아제 TXNRD1 티오레독신 환원효소 1 (다기능 DNA 수선 효소) 1 NOS2 산화질소 신타제 2, TIMP1 TIMP 메탈로펩티다제 유도성 저해제 1 CASP9 카스파아제 9, 세포자멸사- XIAP X-결합 저해제 관련된 시스테인 세포자멸사 펩티다제 GLG1 골지 글리코단백질 1 EPO 에리트로포이에틴 VEGFA 혈관 내피 ELN 엘라스틴 성장 인자 A GDNF 신경교 세포 유래 NFE2L2 핵 인자(적혈구-신경영양 인자 유래 2)-유사 2 SLC6A3 용질 캐리어 패밀리 6 HSPA4 열충격 70 kDa (신경전달물질 단백질 4 수송인자, 도파민), 멤버 3 APOE 아포지방단백질 E PSMB8 프로테아솜 (프로솜, 매크로페인) 서브유닛, 베타 타입, 8 DCTN1 디낙틴 1 TIMP3 TIMP 메탈로펩티다제 저해제 3 KIFAP3 키네신-관련 SLC1A1 용질 캐리어 패밀리 1 단백질 3 (뉴런/상피 고친화성 글루타메이트 수송인자, 시스템 Xag), 멤버 1 SMN2 생존 운동 뉴런 CCNC 사이클린 C 2, 동심체 MPP4 막 단백질, STUB1 STIP1 상동체 및 U-팔미토일화 4 박스 함유 단백질 1 ALS2 아밀로이드 베타 (A4) PRDX6 퍼옥시레독신 6 전구체 단백질 SYP 시냅토피신 CABIN1 칼시뉴린 결합 단백질 1 CASP1 카스파아제 1, 세포자멸사- GART 포스포리보실글리신아미드 관련 시스테인 포르밀트랜스페라제, 펩티다제 포스포리보실글리신아미드 신세타제, 포스포리보실아미노이미다졸 신세타제 CDK5 사이클린-의존성 키나제 5 ATXN3 아탁신 3 RTN4 레티쿨론 4 C1QB 보체 구성성분 1, q 하위구성성분, B 사슬 VEGFC 신경성장인자 HTT 헌팅틴 수용체 PARK7 파킨슨병 7 XDH 잔틴 디하이드로게나제 GFAP 신경교 섬유질 산성 MAP2 미세소관-관련 단백질 단백질 2 CYCS 시토크롬 c, IgG의 체세포 FCGR3B Fc 단편, 저 친화성 IIIb, CCS 구리 샤페롱 UBL5 유비퀴틴-유사 5 수퍼옥사이드 디스뮤타제 MMP9 바탕질

메탈로펩티다제 SLC18A3 용질 캐리어 패밀리 18 9(혈관 아세틸콜린), 멤버 3 TRPM7 일시 수용체 HSPB2 열충격 27 kDa 잠재적 양이온 채널, 단백질 2 서브패밀리 M, 멤버 7 AKT1 v-akt 뮤린 흉선종 DERL1 Der1-유사 도메인 패밀리, 바이러스 암유전자 상동체 1 멤버 1 CCL2 케모카인 (C--C 모티프) NGRN 뉴그린, 뉴라이트 리간드 2 돌기 관련 GSR 글루타티온 환원효소 TPPP3 튜불린 중합-촉진 단백질 패밀리 멤버 3 APAF1 세포자멸사 펩티다제 BTBD10 BTB (POZ) 도메인 활성화 인자 1 함유 10 GLUD1 글루타메이트 CXCR4 케모카인 (C--X--C 모티프) 디하이드로게나제 1 수용체 4 SLC1A3 용질 캐리어 패밀리 1 FLT1 fms-관련 티로신 (신경교 고 친화성 글루타메이트 수송인자), 멤버 3 키나제 1 PON1 파라옥소나제 1 AR 안드로겐 수용체 LIF 백혈병 억제 인자 ERBB3 v-erb-b2 적아세포 백혈병 바이러스 암유전자 상동체 3 LGALS1 렉틴, 갈락토시드-CD44 CD44 분자 결합, 가용성, 1 TP53 종양 단백질 p53 TLR3 톨-유사 수용체 3 GRIA1 글루타메이트 수용체, GAPDH 글리세르알데하이드-3-이온성, AMPA 1 포스페이트 디하이드로게나제 GRIK1 글루타메이트 수용체, DES 데스민 이온성, 카이네이트 1 CHAT 콜린 아세틸트랜스페라제 FLT4 fms-관련 티로신 키나제 4 CHMP2B 크로마틴 변형 BAG1 BCL2-관련 단백질 2B 안타노젠 MT3 메탈로티오네인 3 CHRNA4 콜린성 수용체, 니코틴성, 알파 4 GSS 글루타티온 신세타제 BAK1 BCL2-길항제/킬러 1 KDR 키나제 삽입 도메인 GSTP1 글루타티온 S-트랜스페라제 수용체 (a 타입 III 파이 1 수용체 티로신 키나제) OGG1 8-옥소구아닌 DNA IL6 인터류킨 6 (인터페론, 글리코실라제 베타 2).

[0933] 동물 또는 세포는 ALS 와 관련된 단백질을 암호화하는 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 또는 그 이상의 파괴된 염색체 서열 및 ALS 와 관련된 파괴된 단백질을 암호화하는 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 또는 그 이상의 염색체 통합 서열을 포함할 수 있다. ALS와 관련된 바람직한 단백질은 SOD1 (수퍼옥사이드 디스뮤타제 1), ALS2(근위축성 측삭경화증 2), FUS (육종 융합형), TARDBP (TAR DNA 결합 단백질), VAGFA(혈관 내피 성장인자 A), VAGFB(혈관 내피 성장인자 B), 및 VAGFC(혈관 내피 성장인자 C), 및 이들의 어떤 조합을 포함한다.

[0934] 프리온 질병과 관련된 단백질의 예는 SOD1(슈퍼옥시드 디스뮤타제 1), ALS2(근위축성 측삭 경화증 2), FUS(육종에서 융합), TARDBP(TAR DNA 결합 단백질), VAGFA(혈관 내피 성장 인자 A), VAGFB(혈관 내피 성장 인자 B) 및 VAGFC(혈관 내피 성장 인자 C) 및 그들의 임의의 조합을 포함할 수 있다.

[0935] 프리온 장애에서의 신경변성 질환과 관련된 단백질의 예는 예를 들어, A2M(알파-2-마크로글로불린), AATF(아폽토시스 길항작용 전사 인자), ACPP(전립선 산 포스파타제), ACTA2(액틴 알파 2 평활근 대동맥), ADAM22(ADAM 메탈로펩티다제 도메인), ADORA3(아데노신 A3 수용체) 또는 ADRA1D(알파-1D 아드레노수용체에 대한 알파-1D 아드레날린 작용성 수용체)를 포함할 수 있다.

[0936] 면역결핍과 관련된 단백질의 예는 예를 들어, A2M[알파-2-마크로글로불린]; AANAT[아릴알킬아민 N-아세틸트랜스퍼라제]; ABCA1[ATP-결합 카세트, 하위과 A(ABC1), 구성원 1]; ABCA2[ATP-결합 카세트, 하위과 A(ABC1), 구성원 2]; 또는 ABCA3[ATP-결합 카세트, 하위과 A(ABC1), 구성원 3];을 포함할 수 있다.

[0937] 트리뉴클레오티드 반복 장애와 관련된 단백질의 예는 예를 들어, AR(안드로겐 수용체), FMR1(유약 X 정신 지체 1), HTT(헌팅틴) 또는 DMPK(근긴장성 이영양증-단백질 키나제), FXN(프라탁신(frataxin)), ATXN2(아탁신(ataxin) 2)를 포함한다.

[0938] 신경전달 장애와 관련된 단백질의 예는 예를 들어, SST(소마토스타틴), NOS1(산화질소 신타제 1(뉴런)), ADRA2A(아드레날린 작용성, 알파-2A-, 수용체), ADRA2C(아드레날린 작용성, 알파-2C-, 수용체), TACR1(타키키닌 수용체 1) 또는 HTR2c(5-하이드록시트립타민(세로토닌) 수용체 2C)를 포함한다.

[0939] 신경발달-관련 서열의 예는 예를 들어, A2BP1[아탁신 2-결합 단백질 1], AADAT[아미노아디페이트 아미노트랜스퍼라제], AANAT[아릴알킬아민 N-아세틸트랜스퍼라제], ABAT[4-아미노부티레이트 아미노트랜스퍼라제], ABCA1[ATP-결합 카세트, 하위과 A(ABC1), 구성원 1] 또는 ABCA13[ATP-결합 카세트, 하위과 A(ABC1), 구성원 13]을 포함한다.

[0940] 본 발명의 시스템으로 치료가능한 바람직한 질환의 추가의 예는 에카르디 고우티에레스(Aicardi-Goutieres) 증후군; 알렉산더병; 알란-헌든-두들리(Allan-Herndon-Dudley) 증후군; POLG-관련 장애; 알파-만노시도시스(Alpha-Mannosidosis)(II 및 III형); 알스트스트롬(Alstrom) 증후군; 안젤만(Angelman); 증후군; 혈관확장성 운동실조증; 신경 세로이드 리포푸신증; 베타-지중해빈혈; 양쪽성 시신경위축 및 (영아) 1형 시신경위축; 망막모세포종(양쪽성); 캐너번병; 뇌-눈-얼굴-골격 증후군 1[COFS1]; 뇌힘줄황색종증; 코넬리아디란지 증후군; MAPT-관련 장애; 유전적 프리온 질병; 드라벳 증후군; 조기-발병 가족성 알츠하이머병; 프리드리히 운동실조[FRDA]; 프린스 증후군; 푸코시드 축적증; 후쿠야마형 선천성 근이영양증; 갈락토시알리도시스; 고셰병; 유기산혈증; 혈구탐식성 림프조직구증; 허친슨-길포오드 조로증 증후군; II형 뮤코리피드증; 유아 유리 시알산 축적

병; PLA2G6-관련 신경변성; 제벨 랑쥐-닐슨 증후군; 연접부 수포성 표피박리증; 헌팅톤병; 크라베병(유아); 미토콘드리아 DNA-관련 레이 증후군 및 NARP; 레슈-니한 증후군; LIS1-관련 뇌회결손; 로우 증후군; 단풍시럽뇨병; MECP2 중복 증후군; ATP7A-관련 구리 수송 장애; LAMA2-관련 근이영양증; 아릴설파타제 A 결핍; I, II 또는 III형 점액다당류증; 퍼옥시좀 생물발생 장애; 젤웨거 증후군 스펙트럼; 뇌 철 축적 장애가 있는 신경변성; 산 스핑고미엘리나제 결핍; C형 니만 픽병; 글리신 뇌병증; ARX-관련 장애; 요소 사이클 장애; COL1A1/2-관련 불완전골형성; 미토콘드리아 DNA 결실 증후군; PLP1-관련 장애; 페리(Perry) 증후군; 펠란-맥더미드 증후군; II형 글리코겐 축적병(폼페병)(유아); MAPT-관련 장애; MECP2-관련 장애; 1형 어깨엉덩관절 점상연골형성이상; 로버츠 증후군; 샌드호프병; 쉰들러병 - 1형; 아데노신 탈아미노효소 결핍; 스미스-렘리-오피츠 증후군; 척수성 근위축; 유아-발병 척수소뇌성 실조증; 헥소사미니다제 A 결핍; 1형 치사성 이형성증; VI형 콜라겐-관련 장애; I형 어셔 증후군; 선천성 근이영양증; 윌프-허쉬호른 증후군; 리소좀산 리파제 결핍; 및 색소성 건피증으로부터 선택될 수 있다.

[0941] 명백한 바와 같이, 본 발명의 시스템이 임의의 대상 폴리뉴클레오티드 서열을 표적화하기 위해 사용될 수 있음이 예상된다. 본 발명의 시스템을 사용하여 유용하게 치료될 수 있는 질환 또는 질병의 일부 예는 상기 표에 포함되어 있으며, 그들 질환과 현재 관련되어 있는 유전자의 예도 또한 거기에 제공된다. 그러나, 예시된 유전자는 배타적인 것은 아니다.

[0942] 예를 들어, "야생형 StCas9"는 *스트렙토코커스 써모필루스*로부터의 야생형 Cas9를 말하며, 이것의 단백질 서열은 승인번호 G3ECR1로서 SwissProt 데이터베이스에 주어진다. 유사하게, *S. pyogenes* Cas9도 승인번호 Q99ZW2로서 SwissProt에 포함된다.

[0943] 효율적이고 비용 효과적인 유전자 편집 및 조작을 실행하기 위하여 CRISPR-Cas 시스템을 사용하는 능력은 단일 그리고 그리고 다중복합체 유전자 조작의 빠른 선택 및 비교가 개선된 생성 및 향상된 형질을 위하여 이와 같은 게놈을 형질전환시키는 것을 가능하게 할 것이다. 이와 관련하여, 미국 특허 및 공개, 즉 미국 특허 6,603,061호(아그로박테리움-매개 식물 형질전환 방법(Agrobacterium-Mediated Plant Transformation Method)); 미국 특허 7,868,149호(식물 게놈 서열 및 이의 용도(Plant Genome Sequences and Uses Thereof) 및 미국 2009/0100536호(향상된 농업 형질을 가지는 유전자이식 식물(Transgenic Plants with Enhanced Agronomic Traits))가 언급되며, 이들 각각의 내용 및 개시 내용 모두는 본원에 참조로 전체가 포함되어 있다. 본 발명의 실시에서, 문헌[Morrell et al “Crop genomics:advances and applications” Nat Rev Genet. 2011 Dec 29;13(2):85-96]의 내용 및 개시 내용 또한 본원에 참조로 전체가 포함되어 있다.

[0944] **실시예**

[0945] 하기의 실시예는 본 발명의 다양한 구현예를 예시할 목적으로 제공되며, 어떠한 방식으로든 본 발명을 제한하고자 하지 않는다. 본원에 기술된 방법과 함께 본 발명의 실시예는 본원에서 바람직한 구현예를 대표하는 것이며, 예시적이고, 본 발명의 범주에 대한 제한으로 의도되지 않는다. 거기에서의 변화 및 다른 용도는 청구범위의 범주에 의해 정의되는 바와 같은 본 발명의 목적에 포함되며, 당업자에게 수행될 것이다.

[0946] *실시예 1: 진핵 세포의 핵에서의 CRISPR 복합체 활성*

[0947] 예시적인 II형 CRISPR 시스템은 스트렙토코커스 피오게네스 SF370 유래의 II형 CRISPR 유전자좌이며, 이는 4 개 유전자, Cas9, Cas1, Cas2 및 Csn1의 클러스터 뿐 아니라 2 개의 비-코딩 RNA 요소, tracrRNA, 및 비-반복 서열의 짧은 스트레치(스페이서, 각각 30bp)에 의해 산재된 반복 서열의 특징적 어레이(직접 반복부)를 포함한다. 이러한 시스템에서, 표적화된 DNA 이중-가닥 파단(DSB)은 4 개의 순차적 단계로 생성된다(도 2a). 먼저, 2 개의 비-코딩 RNA, pre-crRNA 어레이 및 tracrRNA를 CRISPR 유전자좌로부터 전사시킨다. 두 번째로, tracrRNA를 pre-crRNA 의 직접 반복부에 혼성화시키고, 이어서 개별 스페이서 서열을 함유하는 성숙 crRNA로 가공한다. 세 번째로, 성숙 crRNA:tracrRNA 복합체는 crRNA의 스페이서 영역과 프로토스페이서 DNA 사이의 헤테로듀플렉스 형성을 통해 Cas9를 프로토스페이서 및 상응하는 PAM으로 구성된 DNA 표적으로 유도한다. 마지막으로, Cas9는 PAM의 상류의 표적 DNA의 절단을 매개하여 프로토스페이서 내에 DSB를 생성한다(도 2a). 이러한 실시예는 진핵 세포의 핵에서 CRISPR 복합체 활성을 유도하기 위해 이러한 RNA-프로그램화가능 뉴클레아제 시스템을 조정하는 예시적인 과정을 기술한다.

[0948] *세포 배양 및 트랜스펙션*

[0949] 인간 배아 신장(HEK) 세포주 HEK 293FT(라이프 테크놀로지즈(Life Technologies))를 10% 우태아혈청(하이클론(HyClone)), 2 mM GlutaMAX(라이프 테크놀로지즈), 100 U/㎖ 페니실린 및 100 ㎍/㎖ 스트렙토마이신이 보충된

둘베코 개질 이글스 배지(DMEM)에서 37℃에서 5% CO_2 인큐베이션과 함께 유지시켰다. 마우스 neuro2A(N2A) 세포주(ATCC)를 37℃에서 5% CO_2와 함께, 5% 우태아혈청(하이클론), 2 mM GlutaMAX(라이프 테크놀로지즈), 100 U/㎖ 페니실린 및 100 ㎍/㎖ 스트렙토마이신이 보충된 DMEM으로 유지시켰다.

[0950] 트랜스펙션 1 일 전, 웰당 200,000 개 세포의 밀도로 HEK 293FT 또는 N2A 세포를 24-웰 플레이트(코닝(Corning))에 씨딩하였다. 세포를 제조사의 권고 프로토콜에 따라 리포펙타민(Lipofectamine) 2000(라이프 테크놀로지즈)을 사용하여 트랜스펙션시켰다. 24-웰 플레이트의 각 웰에 대해 총 800 ng의 플라스미드를 사용하였다.

[0951] *게놈 변형에 대한 서베이어 검정 및 시퀀싱 분석*

[0952] HEK 293FT 또는 N2A 세포를 상기 기술한 바와 같이 플라스미드 DNA로 트랜스펙션시켰다. 트랜스펙션 후에, 게놈 DNA 추출 전 72 시간 동안 37℃에서 세포를 인큐베이션시켰다. 게놈 DNA를 제조사의 프로토콜에 따라 퀵익스트랙트(QuickExtract) DNA 추출 키트(에피센트레(Epicentre))를 사용하여 추출하였다. 간략하게, 세포를 퀵익스트랙트 용액에 재현탁시키고, 65℃에서 15 분 동안, 그리고 98℃에서 10 분 동안 인큐베이션시켰다. 추출된 게놈 DNA를 바로 처리하거나 20℃에서 보관하였다.

[0953] 각 유전자에 대한 CRISPR 표적 부위의 주변의 게놈 영역을 PCR 증폭시키고, 산물을 제조사의 프로토콜에 따라 퀴아퀵 스핀 컬럼(QiaQuick Spin Column)(퀴아젠(Qiagen))을 사용하여 정제하였다. 총 400ng의 정제된 PCR 산물을 2 ㎕ 10X Taq 중합효소 PCR 완충제(엔자이머틱스(Enzymatics)) 및 초순수와 총 20 ㎕ 부피로 혼합하고, 재-아닐링 과정을 거치게 하여 헤테로듀플렉스가 형성되게 하였다: 95℃에서 10 분, 95℃에서 85℃(-2℃/초로 램핑(ramping)), 85℃에서 25℃(-0.25℃/초), 및 25℃에서 1 분 유지. 재아닐링 후, 산물을 제조사의 권고 프로토콜에 따라 서베이어 뉴클레아제 및 서베이어 인핸서 S(트랜스게노믹스(Transgenomics))로 처리하고, 4-20% 노벡스(Novex) TBE 폴리-아크릴아미드 겔(라이프 테크놀로지즈)에서 분석하였다. 겔을 SYBR 골드(Gold) DNA 염색제(라이프 테크놀로지즈)로 30 분 동안 염색하고, Gel Doc 겔 영상화 시스템(바이오-라드(Bio-rad))으로 영상화하였다. 정량화는 절단된 DNA의 분율의 척도로서 상대적 밴드 세기를 기반으로 하였다. 도 7은 이러한 서베이어 검정의 개략적 예시를 제공한다.

[0954] 상동성 재조합의 검출을 위한 제한 단편 길이 다형성 검정

[0955] HEK 293FT 및 N2A 세포를 플라스미드 DNA로 트랜스펙션시키고, 37℃에서 72 시간 동안 인큐베이션시킨 후 상기 기재한 바와 같이 게놈 DNA를 추출하였다. 표적 게놈 영역을 상동성 재조합(HR) 주형의 상동성 암(arm) 외측의 프라이머를 사용하여 PCR 증폭시켰다. PCR 산물을 1% 아가로오스 겔 상에서 분리하고 MinElute GelExtraction 키트(퀴아젠)를 사용하여 추출하였다. 정제된 산물을 HindIII(퍼멘터스(Fermentas))로 분해시키고 6% 노벡스 TBE 폴리-아크릴아미드 겔(라이프 테크놀로지즈)에서 분석하였다.

[0956] RNA 2차 구조 예측 및 분석

[0957] 센트로이드 구조 예측 알고리즘을 사용하는 비엔나 대학의 이론 화학 기관에서 개발된 온라인 웹서버 RNAfold를 사용하여 RNA 2차 구조 예측을 수행하였다(예를 들어, 문헌[A.R. Gruber et al., 2008, Cell 106(1): 23-24]; 및 문헌[PA Carr and GM Church, 2009, Nature Biotechnology 27(12): 1151-62] 참조).

[0958] RNA 정제

[0959] HEK 293FT 세포를 상기 언급한 바와 같이 유지시키고 트랜스펙션시켰다. 세포를 트립신처리에 의해 수집한 후 인산 완충 염수(PBS)로 세척하였다. 전체 세포 RNA를 제조사의 프로토콜에 따라 TRI 시약(시그마(Sigma))으로 추출하였다. 추출한 전체 RNA를 나노드롭(Naonodrop)(써모 사이언티픽(Thermo Scientific))을 사용하여 정량화하고 동일 농도로 정규화하였다.

[0960] *포유류 세포 내의 crRNA 및 tracrRNA 발현의 노던 블롯 분석*

[0961] RNA를 동일 부피의 2X 로딩 완충제(앰비온(Ambion))와 혼합하고, 5 분 동안 95℃로 가열하고, 1 분 동안 얼음 상에서 냉각시킨 후, 적어도 30 분 동안 겔을 사전-전개시킨 후 8% 변성 폴리아크릴아미드 겔(SequaGel, 내셔널 디아그노스틱스(National Diagnostics)) 상에 로딩하였다. 샘플을 40 W 한도에서 1.5 시간 동안 전기영동하였다. 그 후, RNA를 실온에서 1.5 시간 동안 반-건조 운반 장치(바이오-라드) 내에서 300 mA에서 하이본드(Hybond) N+ 멤브레인(지이 헬쓰케어(GE Healthcare)) 에 옮겼다. RNA를 스트라타진(Stratagene) UV 가교제 스트라타링커(Stratalinker)(스트라타진) 상의 자동가교 버튼을 사용하여 멤브레인에 가교시켰다. 멤브레

인을 ULTRAhyb-Oligo 혼성화 완충제(앰비온) 중에서 30 분 동안, 42℃에서 회전시키면서 사전-혼성화한 후, 프로브를 첨가하고 밤새 혼성화하였다. 프로브를 IDT에서 주문하고 T4 폴리뉴클레오티드 키나제(뉴 잉글랜드 바이오랩스(New England Biolabs))와 함께 [감마-^{32}P] ATP(퍼킨 엘머(Perkin Elmer))로 표지하였다. 멤브레인을 사전가온된(42℃) 2xSSC, 0.5% SDS로 1분 동안 1 회 세척한 후 42℃ 에서 2 회 30 분 세척하였다. 멤브레인을 1 시간 동안 또는 밤새 실온에서 인광체 스크린에 노출시킨 후 포스포르이미저(phosphorimager)(타이푼(Typhoon))로 스캔하였다.

[0962] *박테리아 CRISPR 시스템 작제 및 평가*

[0963] tracrRNA, Cas9 및 리더를 포함하는 CRISPR 유전자좌 요소를 깁슨 조립을 위한 측부 배치 상동성 암과 함께, 스트렙토코커스 피오게네스 SF370 게놈 DNA로부터 PCR 증폭시켰다. 2 개의 BsaI IIS형 부위를 2 개의 직접 반복부 사이에 도입하여, 용이한 스페이서의 삽입을 가능하게 하였다(도 8). PCR 산물을 깁슨 조립 마스터 믹스(NEB)를 사용하여 tet 프로모터의 하류에 EcoRV-분해 pACYC184 내로 클로닝하였다. Csn2의 마지막 50 bp를 제외하고 다른 내인성 CRISPR 시스템 요소를 생략하였다. 상보성 오버행을 지니는 스페이서를 인코딩하는 올리고(인테그레이티드 디엔에이 테크놀로지(Integrated DNA Technology))를 BsaI-분해 벡터 pDC000(NEB) 내로 클로닝한 다음, T7 리가제(엔자이머틱스)를 사용하여 결찰시켜, pCRISPR 플라스미드를 생성하였다. 포유류 세포에서 PAM 발현과 함께 스페이서를 함유하는 챌린지 플라스미드 (도 6a에 도시된 발현 작제물, 기능성은 도 6b에 도시된 서베이어 분석의 결과에 의해서 결정된다). 전사 시작 부위는 +1로 표시되며, 전사 종결자 및 노던 블롯에 의해 프로빙되는 서열도 또한 표시되어 있다. 가공된 tracrRNA의 발현도 또한 노던 블롯에 의해 확인하였다. 도 6c는 긴 또는 짧은 tracrRNA, 및 SpCas9 및 DR-EMX1(1)-DR을 지니는 U6 발현 작제물이 트랜스펙션된 293FT 세포로부터 추출된 전체 RNA의 노던 블롯 분석의 결과를 보여준다. 좌측 및 우측 패널은 각각 SpRNase III의 부재 또는 존재 하에 트랜스펙션된 293FT 세포로부터의 것이다. U6은 인간 U6 snRNA를 표적화하는 프로브로 블롯팅된 로딩 대조군을 나타낸다. 짧은 tracrRNA 발현 작제물의 트랜스펙션은 풍부한 수준의 tracrRNA의 가공된 형태를 야기한다(약 75 bp). 매우 소량의 긴 tracrRNA가 노던 블롯에서 검출된다.

[0964] 정밀한 전사 개시를 촉진시키기 위하여 RNA 중합효소 III 기반의 U6 프로모터를 선택하여, tracrRNA의 발현을 유도하였다(도 2c). 유사하게, U6 프로모터 기반의 작제물을 2 개의 직접 반복부(DR, 또한 용어 "tracr-메이트 서열"에 포함; 도 2c)가 측부 배치된 단일의 스페이서로 구성된 pre-crRNA 어레이를 발현하도록 발생시켰다. 초기 스페이서를 대뇌 피질의 발생의 주요 유전자인 인간 EMX1 유전자좌 내의 33-염기-쌍(bp) 표적 부위(30-bp 프로토스페이서 + Cas9의 NGG 인식 모티프를 만족시키는 3-bp CRISPR 모티프(PAM) 서열)를 표적화하도록 설계하였다(도 2c).

[0965] 포유동물 세포에서 CRISPR 시스템(SpCas9, SpRNase III, tracrRNA 및 pre-crRNA)의 이종 발현이 포유동물 염색체의 표적화된 절단을 달성할 수 있는지를 시험하기 위하여, HEK 293FT 세포를 CRISPR 성분의 조합으로 트랜스펙션시켰다. 포유동물 핵에서 DSB가 삽입-결실의 형성을 야기하는 비-상동 말단 연결(NHEJ) 경로에 의해 부분적으로 수복되기 때문에, 서베이어 검정을 사용하여, 표적 EMX1 유전자좌에서 잠재적인 절단 활성을 검출하였다(도 7)(예를 들어, 문헌[Guschin et al., 2010, Methods Mol Biol 649: 247] 참조). 모든 4 개의 CRISPR 성분의 동시-트랜스펙션은 프로토스페이서 최대 5.0% 절단을 유도할 수 있었다(도 2d 참조). 또한, SpRNase III을 제한 모든 CRISPR 성분의 동시-트랜스펙션에 의해, 프로토스페이서에서 최대 4.7% 삽입-결실이 유도되었으며, 이는 crRNA 성숙을 보조할 수 있는 내인성 포유동물 RNase, 예를 들어, 관련 Dicer 및 Drosha 효소가 존재할 수 있음을 시사한다. 남아 있는 3 개 성분 중 임의의 것의 제거에 의해, CRISPR 시스템의 게놈 절단 활성이 없어졌다(도 2d). 표적 유전자좌를 함유하는 앰플리콘(amplicon)의 생거(Sanger) 시퀀싱에 의해, 절단 활성이 입증되었으며: 43 개의 시퀀싱된 클론 중에, 5 개의 돌연변이된 대립형질(11.6%)이 관찰되었다. 다양한 가이드 서열을 사용하는 유사한 실험에 의해, 29%만큼 높은 삽입-결실 백분율을 생성하였다(도 3 내지 6, 10 및 11 참조). 이들 결과는 포유동물 세포에서의 효율적인 CRISPR-매개의 게놈 변형을 위한 3-성분 시스템을 정의한다. 절단 효율을 최적화시키기 위하여, 본 발명자들은 또한 상이한 아이소형의 tracrRNA가 절단 효율에 영향을 미치는지를 시험하였으며, 이러한 예시적인 시스템에서, 오직 짧은(89-bp) 전사물 형태만이 인간 EMX1 게놈 유전자좌의 절단을 매개할 수 있는 것이 관찰되었다(도 6b).

[0966] 도 12는 포유동물 세포에서의 crRNA 가공의 추가의 노던 블롯 분석을 제공한다. 도 12a는 2 개의 직접 반복부가 측부 배치된 단일의 스페이서(DR-EMX1(1)-DR)에 대한 발현 벡터를 보여주는 개략도를 예시한다. 인간 EMX1 유전자좌 프로토스페이서 1을 표적화하는 30 bp 스페이서(도 6 참조) 및 직접 반복부 서열은 도 12a 아래의 서열에 나타나 있다. 선은 역-상보 서열을 사용하여 EMX1(1) crRNA 검출을 위한 노던 블롯 프로브를 생성한 영역을 나

타낸다. 도 12b는 DR-EMX1(1)-DR을 지니는 U6 발현 작제물로 트랜스펙션된 293FT 세포로부터 추출된 전체 RNA의 노던 블롯 분석을 보여준다. 좌측 및 우측 패널은 각각 SpRNase III의 부재 또는 존재 하에 트랜스펙션된 293FT 세포로부터의 것이다. DR-EMX1(1)-DR은 SpCas9 및 짧은 tracrRNA의 존재 하에서만 성숙 crRNA로 처리되었고, SpRNase III의 존재에 따라 달라지지 않았다. 트랜스펙션된 293FT 전체 RNA로부터 검출된 성숙 crRNA는 약 33 bp이며, 스트렙토코커스 피오게네스로부터의 39 내지 42 bp 성숙 crRNA보다 더 짧다. 이들 결과는 CRISPR 시스템이 진핵 세포로 이식될 수 있으며, 내인성 포유동물 표적 폴리뉴클레오티드의 절단을 용이하게 하도록 재프로그램화될 수 있음을 보여준다.₩

[0967] 도 2는 이러한 실시예에 기술된 박테리아 CRISPR 시스템을 예시한다. 도 2a는 스트렙토코커스 피오게네스 SF370으로부터의 CRISPR 유전자좌 1 및 이러한 시스템에 의한 제안된 CRISPR-매개의 DNA 절단의 메카니즘을 보여주는 개략도를 예시한다. 직접 반복부-스페이서 어레이로부터 가공된 성숙 crRNA는 Cas9를 상보성 프로토스페이서 및 프로토스페이서-인접 모티프(PAM)로 구성된 게놈 표적에 지향시킨다. 표적-스페이서 염기 쌍형성 시에, Cas9는 표적 DNA에서 이중 가닥 파단을 매개한다. 도 2b는 포유동물 핵으로의 유입을 가능하게 하는 핵 국소화 신호(NLS)가 있는 스트렙토코커스 피오게네스 Cas9(SpCas9) 및 RNase III(SpRNase III)의 조작을 예시한다. 도 2c는 정밀한 전사 개시 및 종결을 촉진하기 위한 구성성 EF1a 프로모터에 의해 유도되는 SpCas9 및 SpRNase III 및 RNA Pol3 프로모터 U6에 의해 유도되는 tracrRNA 및 pre-crRNA 어레이(DR-스페이서-DR)의 포유동물 발현을 예시한다. 만족스러운 PAM 서열이 있는 인간 EMX1 유전자좌로부터의 프로토스페이서는 pre-crRNA 검정에서 스페이서로 사용된다. 도 2d는 SpCas9-매개의 최소 삽입 및 결실을 위한 서베이어 뉴클레아제 검정을 예시한다. SpRNase III, tracrRNA 및 EMX1-표적 스페이서를 지니는 pre-crRNA 어레이의 존재 및 부재 하에 SpCas9를 발현시켰다. 도 2e는 표적 유전자좌와 EMX1-표적화 crRNA 사이의 염기 쌍형성의 개략적 표현, 및 SpCas9 절단 부위에 인접한 마이크로 결실을 보이는 예시적인 크로마토그램을 예시한다. 도 2f는 다양한 마이크로 삽입 및 결실을 보이는 43 개의 클론 앰플리콘의 시퀀싱 분석으로부터 확인된 돌연변이된 대립형질을 예시한다. 줄표는 결실된 염기를 나타내며, 비-정렬 또는 미스매치된 염기는 삽입 또는 돌연변이를 나타낸다. 축척 막대 = 10 ㎛.

[0968] 3-성분 시스템을 더욱 단순화시키기 위하여, 키메라 crRNA-tracrRNA 하이브리드 설계를 조정하였으며, 여기서, 성숙 crRNA(가이드 서열 포함)는 스템-루프를 통해 부분 tracrRNA에 융합되어, 천연 crRNA:tracrRNA 듀플렉스를 모방할 수 있다. 동시-전달 효율을 증가시키기 위하여, 비시스트로닉 발현 벡터를 생성하여, 트랜스펙션된 세포에서 키메라 RNA 및 SpCas9의 동시-발현을 유도하였다. 병행하여, 비시스트로닉 벡터를 사용하여 SpCas9와 함께 pre-crRNA(DR-가이드 서열-DR)를 발현하여, 따로 발현되는 tracrRNA와 함께 crRNA로 가공되도록 유도하였다(도 11b 상부 및 하부 비교). 도 8은 hSpCas9가 있는 pre-crRNA 어레이(도 8a) 또는 키메라 crRNA(도 8b에서 가이드 서열 삽입 부위의 하류 및 EF1α 프로모터의 상류에 짧은 선으로 표현)에 대한 비시스트로닉 발현 벡터의 개략적 표현을 제공하며, 다양한 요소의 위치 및 가이드 서열 삽입 점을 보여준다. 도 8b에서 가이드 서열 삽입 부위의 위치 주위의 확대된 서열은 또한 부분 DR 서열(GTTTAGAGCTA 서열 번호 ___) 및 부분 tracrRNA 서열(TAGCAAGTTAAAATAAGGCTAGTCCGTTTTT 서열 번호 ___)을 보여준다. 가이드 서열은 아닐링된 올리고뉴클레오티드를 사용하여 BbsI 부위 사이에 삽입될 수 있다. 올리고뉴클레오티드에 대한 서열 설계는 도 8에서 개략적 표현 밑에 나타나 있으며, 적절한 결찰 어댑터가 표기되어 있다. WPRE는 우드척(Woodchuck) 간염 바이러스 전사후 조절 요소를 나타낸다. 키메라 RNA-매개의 절단의 효율을 상기 기술된 동일한 EMX1 유전자좌를 표적화함으로써 시험하였다. 앰플리콘의 서베이어 검정 및 생거 시퀀싱 둘 모두를 사용하여, 본 발명자들은 키메라 RNA 설계가 대략 4.7% 변형률로, 인간 EMX1 유전자좌의 절단을 용이하게 하는 것을 확인하였다(도 3).

[0969] 인간 EMX1 및 PVALB, 및 마우스 Th 유전자좌 내의 다수의 영역을 표적화하는 키메라 RNA를 설계하여 인간 및 마우스 세포 둘 모두에서 추가의 게놈 유전자좌를 표적화함으로써 진핵 세포에서 CRISPR-매개의 절단의 일반화가능성을 시험하였다. 도 13은 인간 PVALB(도 13a) 및 마우스 Th(도 13b) 유전자좌에서의 일부 추가의 표적화된 프로토스페이서의 선택을 예시한 것이다. 유전자좌의 개략도 및 각각의 마지막 엑손 내의 3 개의 프로토스페이서의 위치가 제공된다. 밑줄이 있는 서열은 30 bp의 프로토스페이서 서열 및 3' 말단에서 PAM 서열에 상응하는 3 bp를 포함한다. 센스 및 안티-센스 가닥 상의 프로토스페이서는 각각 DNA 서열 상측 및 하측에 표기되어 있다. 인간 PVALB 및 마우스 Th 유전자좌 각각에 대하여 6.3% 및 0.75%의 변형률이 달성되었으며, 이는 다수의 유기체에 걸쳐 상이한 유전자좌의 변형에서의 CRISPR 시스템의 넓은 응용가능성을 입증한다(도 5). 키메라 작제물을 사용하여 각 유전자좌에 대하여 3 개의 스페이서 중 1 개에서만 절단이 검출되었지만, 동시-발현되는 pre-crRNA 배열을 사용하는 경우 27%에 미치는 삽입-결실 생성 효율로, 모든 표적 서열이 절단되었다(도 6 및 도 13).

[0970] 도 11은 SpCas9가 포유동물 세포에서 다수의 게놈 유전자좌를 표적화하도록 재프로그램화될 수 있다는 추가의

예시를 제공한다. 도 11a는 5 개의 프로토스페이서의 위치를 밑줄이 있는 서열로 나타내어 보여주는 인간 EMX1 유전자좌의 개략도를 제공한다. 도 11b는 pre-crRNA의 직접 반복부 영역과 tracrRNA 간의 혼성화를 보여주는 pre-crRNA/trcrRNA 복합체의 개략도(상측) 및 20 bp 가이드 서열, 및 헤어핀 구조로 혼성화되는 부분 직접 반복부 및 tracrRNA 서열로 구성된 tracr 메이트 서열 및 tracr 서열을 포함하는 키메라 RNA 설계의 개략도(하측)를 제공한다. 인간 EMX1 유전자좌 내의 5 개의 프로토스페이서에서의 Cas9-매개의 절단의 효능을 비교하는 서베이어 검정의 결과는 도 11c에 예시되어 있다. 각각의 프로토스페이서는 가공된 pre-crRNA/tracrRNA 복합체(crRNA) 또는 키메라 RNA(chiRNA)를 사용하여 표적화된다.

[0971] RNA의 2차 구조가 분자간 상호작용에 결정적일 수 있기 때문에, 최소 자유 에너지 및 볼쯔만(Boltzmann)-가중 구조 앙상블에 기초한 구조 예측 알고리즘을 사용하여 게놈 표적화 실험에 사용되는 모든 가이드 서열의 추정상 2차 구조를 비교하였다(예를 들어, 문헌[Gruber et al., 2008, Nucleic Acids Research, 36: W70] 참조). 분석에 의해, 대부분의 경우에, 키메라 crRNA 맥락에서 효율적인 가이드 서열에 2차 구조 모티프가 실질적으로 없지만, 비효율적인 가이드 서열은 표적 프로토스페이서 DNA와의 염기 쌍형성을 막을 수 있는 내부 2차 구조를 형성할 가능성이 더 큰 것으로 드러났다. 따라서, 키메라 crRNA를 사용하는 경우, 스페이서 2차 구조의 가변성이 CRISPR-매개 간섭의 효율에 영향을 미칠 수 있다.

[0972] SpCas9에 대한 추가의 벡터 설계는 도 22에 나타나 있으며, 이는 가이드 올리고에 대한 삽입 부위에 연결된 U6 프로모터 및 SpCas9 코딩 서열에 연결된 Cbh 프로모터를 혼입한 단일의 발현 벡터를 예시한다. 도 22b에 나타낸 벡터는 H1 프로모터에 연결된 tracrRNA 코딩 서열을 포함한다.

[0973] 박테리아 분석에서, 모든 스페이서는 효과적인 CRISPR 간섭을 촉진했다(도 3c). 이들 결과는 포유류 세포에서 CRISPR 활성의 효율에 영향을 미치는 추가적인 요인이 있을 수 있다는 것을 시사한다.

[0974] CRISPR-매개의 절단의 특이성을 조사하기 위하여, 포유동물 게놈에서 프로토스페이서 절단에 대한 가이드 서열 내의 단일-뉴클레오티드 돌연변이의 영향을, 단일의 점 돌연변이가 있는 일련의 EMX1-표적화 키메라 crRNA를 사용하여 분석하였다(도 3a). 도 3b는 상이한 돌연변이 키메라 RNA와 쌍을 형성하는 경우, Cas9의 절단 효율을 비교하는 서베이어 뉴클레아제 검정의 결과를 예시한다. PAM의 최대 12-bp 5'의 단일-염기 미스매치는 SpCas9에 의한 게놈 절단을 없애는 한편, 더 먼 상류 위치에 돌연변이가 있는 스페이서는 원래의 프로토스페이서 표적에 대한 활성을 보유하였다(도 3b). PAM에 더하여, SpCas9는 마지막 12-bp의 스페이서 내에 단일-염기 특이성을 갖는다. 또한, CRISPR은 동일한 EMX1 프로토스페이서를 표적화하는 TALE 뉴클레아제(TALEN)의 쌍만큼 효율적으로 게놈 절단을 매개할 수 있다. 도 3c는 EMX1을 표적화하는 TALEN의 설계를 보여주는 개략도를 제공하며, 도 3d는 TALEN 및 Cas9의 효율을 비교하는 서베이어 겔을 보여준다(n=3).

[0975] 오류-유발 NHEJ 메카니즘을 통해 포유동물 세포에서 CRISPR-매개의 유전자 교정을 달성하기 위한 성분의 세트를 확립하면, 상동성 재조합(HR), 게놈에서 정밀한 교정을 이루기 위한 고충실도 유전자 수복 경로를 자극하는 CRISPR의 능력을 시험하였다. 야생형 SpCas9는 부위-특이적 DSB를 매개할 수 있는데, 이는 NHEJ 및 HR 둘 모두를 통해 수복될 수 있다. 또한, SpCas9의 RuvC I 촉매 도메인에서의 아스파르트산염에서 알라닌으로의 치환(D10A)을 조작하여 뉴클레아제를 닉카아제로 전환시켜(SpCas9n; 도 4a에 예시)(예를 들어, 문헌[Sapranausaks et al., 2011, Nucleic Acids Research, 39: 9275]; 문헌[Gasiunas et al., 2012, Proc. Natl. Acad. Sci. USA, 109:E2579] 참조), 닉이 있는 게놈 DNA가 고-충실도 상동성-유도 수복(HDR)을 겪게 하였다. 서베이어 검정에 의해, SpCas9n이 EMX1 프로토스페이서 표적에서 삽입-결실을 생성하지 않음이 확인되었다. 도 4b에 예시된 바와 같이, EMX1-표적화 키메라 crRNA와 SpCas9의 동시-발현에 의해, 표적 부위에서 삽입-결실이 생성되는 한편, SpCas9n과의 동시-발현은 그렇지 않았다(n=3). 또한, 327 개 앰플리콘의 시퀀싱에 의해, SpCas9n에 의해 유도되는 임의의 삽입-결실이 검출되지 않았다. 동일한 유전자좌를 선택하여, HEK 293FT 세포를 EMX1, hSpCas9 또는 hSpCas9n을 표적화하는 키메라 RNA, 및 프로토스페이서 근처에 제한 부위(HindIII 및 NheI)의 쌍을 도입하기 위한 HR 주형으로 동시-트랜스펙션시킴으로써 CRISPR-매개의 HR을 시험하였다. 도 4c는 재조합 지점 및 프라이머 아닐링 서열(화살표)의 상대적 위치와 함께, HR 전략의 개략적 예시를 제공한다. SpCas9 및 SpCas9n은 실제로 HR 주형이 EMX1 유전자좌로 통합되는 것을 촉매작용시킨다. 표적 영역의 PCR 증폭 후에, HindIII를 사용한 제한 분해에 의해, 예상되는 단편 크기(도 4d에 나타낸 제한 단편 길이 다형성 겔 분석에서 화살표)에 상응하는 절단 산물이 나타났으며, SpCas9 및 SpCas9n은 유사한 수준의 HR 효율을 매개한다. 본 발명자들은 게놈 앰플리콘의 생거 시퀀싱을 사용하여 HR을 추가로 입증하였다(도 4e). 이들 결과는 포유동물 게놈 내로의 표적화된 유전자 삽입을 촉진시키기 위한 CRISPR의 유용성을 입증한다. 야생형 SpCas9의 14-bp(스페이서로부터 12-bp 및 PAM으로부터 2-bp) 표적 특이성을 고려하여, 단일 가닥 파단이 오류-유발 NHEJ 경로에 대한 기질이 아니기 때문

에, 닉카아제의 이용가능성은 표적외 변형 가능성을 상당히 감소시킬 수 있다.

[0976] 배열된 스페이서가 있는 CRISPR 유전자좌의 천연 구조를 모방하는 발현 작제물(도 2a)을 작제하여, 다중 서열 표적화의 가능성을 시험하였다. EMX1- 및 PVALB-표적화 스페이서의 쌍을 인코딩하는 단일의 CRISPR 어레이를 사용하여, 둘 모두의 유전자좌에서의 효율적인 절단이 검출되었다(도 4f, crRNA 어레이의 개략적 설계 및 효율적인 절단의 매개를 보여주는 서베이어 블롯을 보여줌). 119 bp 만큼 이격된 EMX1 내의 2 개의 표적에 대한 스페이서를 사용하여 동시 발생 DSB을 통한 보다 큰 게놈 영역의 표적화된 결실도 또한 시험하고, 1.6% 결실 효능(182 개 앰플리콘 중 3 개; 도 4g)을 검출하였다. 이는 CRISPR 시스템이 단일의 게놈 내에서 다중화 교정을 매개할 수 있음을 나타낸다.

[0977] *실시예 2: CRISPR 시스템 변형 및 대안*

[0978] 서열-특이적 DNA 절단을 프로그램화시키기 위하여 RNA를 사용하는 능력은 다양한 연구 및 산업 응용을 위한 신규한 부류의 게놈 조작 도구를 정한다. CRISPR 시스템의 몇몇 양태를 추가로 향상시켜, CRISPR 표적화의 효율 및 다능성을 증가시킬 수 있다. 최적의 Cas9 활성은 포유동물 핵에 존재하는 것보다 더 높은 수준의 유리 Mg^{2+}의 이용가능성에 따라 달라질 수 있으며(예를 들어, 문헌[Jinek et al., 2012, Science, 337:816] 참조), 프로토스페이서의 인접 하류 NGG 모티프에 대한 선호는 인간 게놈에서 평균하여 12-bp 마다를 표적화하는 능력을 제한한다(도 9, 인간 염색체 서열의 + 및 - 가닥 둘 모두를 평가). 이들 제약 중 일부는 미생물 메타게놈에 걸친 CRISPR 유전자좌의 다양성을 연구함으로써 극복될 수 있다(예를 들어, 문헌[Makarova et al., 2011, Nat Rev Microbiol, 9:467] 참조). 다른 CRISPR 유전자좌는 실시예 1에 기술된 것과 유사한 과정에 의해 포유동물 세포 환경으로 이식될 수 있다. 예를 들어, 도 10은 CRISPR-매개의 게놈 교정을 달성하기 위한, 포유동물 세포에서의 이종 발현을 위한 스트렙토코커스 써모필러스 LMD-9의 CRISPR 1으로부터의 II형 CRISPR 시스템의 조정을 예시한다. 도 10a는 스트렙토코커스 써모필러스 LMD-9 유래의 CRISPR 1의 개략적 표현을 제공한다. 도 10b는 스트렙토코커스 써모필러스 CRISPR 시스템에 대한 발현 시스템의 설계를 예시한다. 인간 코돈-최적화 hStCas9는 구성성 EF1α 프로모터를 사용하여 발현된다. 성숙 버전의 tracrRNA 및 crRNA는 U6 프로모터를 사용하여 발현되어, 정밀한 전사 개시를 조장한다. 성숙 crRNA 및 tracrRNA로부터의 서열이 예시되어 있다. crRNA 서열에서 소문자 "a"로 표기된 단일의 염기를 사용하여 RNA polIII 전사 종결자로 제공되는 polyU 서열을 제거한다. 도 10c는 인간 EMX1 유전자좌를 표적화하는 가이드 서열을 보여주는 개략도를 제공한다. 도 10d는 서베이어 검정을 사용한 표적 유전자좌에서의 hStCas9-매개의 절단의 결과를 보여준다. RNA 가이드 스페이서 1 및 2는 각각 14% 및 6.4%를 유도하였다. 이들 2 개의 프로토스페이서 부위에서의 생물학적 반복 검증에 걸친 절단 활성의 통계적 분석도 또한 도 5에 제공된다. 도 14는 인간 EMX1 유전자좌에서의 스트렙토코커스 써모필러스 CRISPR 시스템의 추가의 프로토스페이서 및 상응하는 PAM 서열 표적의 개략도를 제공한다. 2 개의 프로토스페이서 서열이 강조표시되어 있으며, NNAGAAW 모티프를 만족시키는 그들의 상응하는 PAM 서열은 상응하는 강조표시된 서열에 대하여 3'에 밑줄을 침으로써 표기된다. 둘 모두의 프로토스페이서는 안티-센스 가닥을 표적화한다.

[0979] *실시예 3: 샘플 표적 서열 선택 알고리즘*

[0980] 소프트웨어 프로그램을 설계하여, 특정 CRISPR 효소에 대하여 요망되는 가이드 서열 길이 및 CRISPR 모티프 서열(PAM)에 기초하여 투입 DNA 서열의 두 가닥 모두에서 후보 CRISPR 표적 서열을 확인한다. 예를 들어, PAM 서열 NGG가 있는 스트렙토코커스 피오게네스 유래의 Cas9에 대한 표적 부위는 투입 서열, 및 투입 서열의 역-상보물 둘 모두에서 5'-N_x-NGG-3'을 검색함으로써 확인될 수 있다. 마찬가지로, PAM 서열 NNAGAAW가 있는 스트렙토코커스 써모필러스 CRISPR1의 Cas9에 대한 표적 부위는 투입 서열, 및 투입 서열의 역-상보물 둘 모두에서 5'-N_x-NNAGAAW-3'을 검색함으로써 확인될 수 있다. 마찬가지로, PAM 서열 NGGNG가 있는 스트렙토코커스 써모필러스 CRISPR3의 Cas9에 대한 표적 부위는 투입 서열, 및 투입 서열의 역-상보물 둘 모두에서 5'-N_x-NGGNG-3'을 검색함으로써 확인될 수 있다. N_x에서 값 "x"는 프로그램에 의해 고정되거나, 사용자에 의해 지정될 수 있으며, 예를 들어, 20일 수 있다.

[0981] 게놈에서의 DNA 표적 부위의 다수의 존재가 비특이적인 게놈 교정을 야기할 수 있기 때문에, 모든 가능한 부위를 확인한 후에, 프로그램은 그들이 관련 참조 게놈에 나타나는 횟수에 기초하여 서열을 필터링한다. 서열 특이성이 '씨드(seed)' 서열, 예를 들어, PAM 서열 그 자체를 포함하여 PAM 서열로부터 5'의 11 내지 12 bp에 의해 결정되는 CRISPR 효소에 있어서, 필터링 단계는 씨드 서열에 기초하여 이루어질 수 있다. 따라서, 추가의 게놈 유전자좌에서 교정을 피하기 위하여, 결과는 관련 게놈에서의 씨드:PAM 서열의 발생 수에 기초하여 필터링된다. 사용자가 씨드 서열의 길이를 선택하게 할 수 있다. 또한, 필터를 통과시키기 위하여, 사용자가 게놈 내의

씨드:PAM 서열의 발생 수를 지정하게 할 수 있다. 디폴트는 독특한 서열에 대하여 스크리닝하는 것이다. 여과 수준은 씨드 서열의 길이와 게놈에서의 서열의 발생 수 둘 모두를 변화시킴으로써 변경된다. 프로그램은 추가로 또는 대안적으로, 확인된 표적 서열(들)의 역 상보물을 제공함으로써 보고된 표적 서열(들)에 상보적인 가이드 서열의 서열을 제공할 수 있다. 인간 게놈에서의 일부 표적 부위의 예시적 시각화가 도 18에 제공된다.

[0982] 서열 선택을 최적화시키기 위한 방법 및 알고리즘의 추가의 상세사항은 본원에 참조로 포함되는 미국 출원 제 61/064,798호(대리인 사건 번호 44790.11.2022; 브로드 참조번호 BI-2012/084)에서 찾을 수 있다.

[0983] *실시예 4: 다중 키메라 crRNA-tracrRNA 하이브리드의 평가*

[0984] 본 실시예는 상이한 길이의 야생형 tracrRNA 서열이 혼입된 tracr 서열을 갖는 키메라 RNA(chiRNA; 단일의 전사물에 가이드 서열, tracr 메이트 서열 및 tracr 서열을 포함)에 대해 수득된 결과를 기술한다. 도 16a는 키메라 RNA 및 Cas9에 대한 비시스트로닉 발현 벡터의 개략도를 예시한다. Cas9는 CBh 프로모터에 의해 유도되며, 키메라 RNA는 U6 프로모터에 의해 유도된다. 키메라 가이드 RNA는 표기된 바와 같은 다양한 위치에서 절단되는 tracr 서열(하부 가닥의 처음 "U"에서 전사물의 마지막까지 계속)에 연결된 20 bp 가이드 서열(N)로 구성된다. 가이드 및 tracr 서열은 tracr-메이트 서열 GUUUUAGAGCUA(서열 번호 ___)에 이어서 루프 서열 GAAA에 의해 분리된다. 인간 EMX1 및 PVALB 유전자좌에서의 Cas9-매개의 삽입-결실에 대한 서베이어 검정의 결과는 각각 도 16b 및 도 16c에 예시되어 있다. 화살표는 예상된 서베이어 단편을 나타낸다. chiRNA는 그들의 "+n" 표기로 표시되며, crRNA는 가이드 및 tracr 서열이 개별 전사물로서 발현되는 하이브리드 RNA를 지칭한다. 3 벌로 수행되는 이들 결과의 정량화는 각각 도 16b 및 도 16c에 상응하는 도 17a 및 도 17b의 히스토그램에 의해 예시되어 있다("N.D."는 삽입-결실이 검출되지 않음을 나타낸다). 프로토스페이서 ID 및 그들의 상응하는 게놈 표적, 프로토스페이서 서열, PAM 서열 및 가닥 위치는 표 D에 제공되어 있다. 하이브리드 시스템에서 개별 전사물의 경우에는 전체 프로토스페이서 서열에 상보적이거나, 키메라 RNA의 경우에는 밑줄이 있는 부분에만 상보성이도록 가이드 서열을 설계하였다.

[0985] **표 D:**

프로토스페이서 ID	게놈 표적	프로토스페이서 서열 (5' → 3')	PAM	가닥
1	*EMX1*	GGACATCGATGTCACCTCCAATGACTAGGG	TGG	+
2	*EMX1*	CATTGGAGGTGACATCGATGTCCTCCCCAT	TGG	-
3	*EMX1*	GGAAGGGCCTGAGTCCGAGCAGAAGAAGAA	GGG	+
4	*PVALB*	GGTGGCGAGAGGGGCCGAGATTGGGTGTTC	AGG	+
5	*PVALB*	ATGCAGGAGGGTGGCGAGAGGGGCCGAGAT	TGG	+

[0986] 이들은 각각 SEQ ID NOS:___ 내지 ___임

[0987] 가이드 서열을 최적화하기 위한 추가의 상세내용은 미국 출원 제61/836,127호 (Attorney docket 44790.08.2022; Broad Reference BI-2013/004G)에서 찾을 수 있으며, 이것은 여기 참고로 포함된다.

[0988] 먼저, 인간 HEK 293FT 세포에서 EMX1 유전자좌 내의 3 개의 부위를 표적화시켰다. 각 chiRNA의 게놈 변형 효율을 서베이어 뉴클레아제 검정을 사용하여 평가하였으며, 이 검정은 DNA 이중-가닥 파단(DSB) 및 비-상동성 말단 연결(NHEJ) DNA 손상 수복 경로에 의한 그들의 이후의 수복으로부터 야기되는 돌연변이를 검출한다. chiRNA(+n)로 표기된 작제물은 야생형 tracrRNA의 최대 +n 뉴클레오티드가 키메라 RNA 작제물에 포함되는 것을 나타내며, 48, 54, 67 및 85의 값이 n에 대해 사용된다. 보다 긴 야생형 tracrRNA의 단편을 함유하는 키메라 RNA(chiRNA(+67) 및 chiRNA(+85))는 모든 3 개의 EMX1 표적 부위에서 DNA 절단을 매개하며, chiRNA(+85)는 특히 개별 전사물에서 가이드 및 tracr 서열을 발현하는 상응하는 crRNA/tracrRNA 하이브리드보다 상당히 더 높은 수준의 DNA 절단을 나타낸다(도 16b 및 도 17a). 또한, 하이브리드 시스템(개별 전사물로 발현되는 가이드 서열 및 tracr 서열)을 사용하여 검출가능한 절단을 제공하지 않는 PVALB 유전자좌 내의 2 개의 부위를 chiRNA를 사용하여 표적화하였다. chiRNA(+67) 및 chiRNA(+85)는 2 개의 PVALB 프로토스페이서에서 상당한 절단을 매개할 수 있었다(도 16c 및 도 17b).

[0989] EMX1 및 PVALB 유전자좌 내의 모든 5 개의 표적에 있어서, tracr 서열 길이의 증가와 일치하는 게놈 변형 효율의 증가가 관찰되었다. 임의의 이론에 구속되지 않고, tracrRNA의 3' 말단에 의해 형성되는 2차 구조는 CRISPR 복합체 형성 비율을 향상시키는데 역할을 수행할 수 있다.

[0990] *실시예 5: Cas9 다양성*

[0991] CRISPR-Cas 시스템은 박테리아 및 조류에 걸쳐 다양한 종에 의해 사용되는 침투하는 외인성 DNA에 대한 적응 면역 메카니즘이다. II형 CRISPR-Cas9 시스템은 CRISPR 유전자좌로의 외래 DNA의 "획득"을 담당하는 단백질을 암호화하는 유전자의 세트 및 DNA 절단 메카니즘의 "실행"을 암호화하는 유전자의 세트로 구성되며; 이들은 DNA 뉴클레아제(Cas9), 비-코딩 트랜스활성화 cr-RNA(tracrRNA), 및 직접 반복부가 측부에 배치된 외래 DNA-유래 스페이서의 어레이(crRNAs)를 포함한다. Cas9에 의한 성숙 시에, tracRNA 및 crRNA 듀플렉스는 Cas9 뉴클레아제를 스페이서 가이드 서열에 의해 특정되는 표적 DNA 서열로 안내하며, 절단에 필요하며 각각의 CRISPR-Cas 시스템에 특이적인 표적 DNA 내의 짧은 서열 모티프 근처의 DNA에서 이중 가닥 파단을 매개한다. II형 CRISPR-Cas 시스템이 박테리아 계에서 관찰되며, Cas9 단백질 서열 및 크기, tracrRNA 및 crRNA 직접 반복부 서열, 이들 요소의 게놈 구성 및 표적 절단을 위한 모티프 요건이 매우 다양하다. 하나의 종이 다중의 별개의 CRISPR-Cas 시스템을 가질 수 있다.

[0992] 본 발명자들은 공지된 Cas9에 대한 서열 상동성 및 공지된 하위도메인에 이종상동성인 구조에 기초하여 확인된 박테리아 종으로부터, HNH 엔도뉴클레아제 도메인 및 RuvC 엔도뉴클레아제 도메인을 포함하는 207 개의 추정의 Cas9를 평가하였다[유진 쿠닌(Eugene Koonin) 및 키라 마카로바(Kira Makarova)로부터의 정보]. 이러한 세트의 단백질 서열 보존에 기초한 계통발생 분석에 의해, 3개 그룹의 큰 Cas9(약 1400 개 아미노산) 및 2 개의 작은 Cas9(약 1100 개 아미노산)를 포함하는 5 개 과의 Cas9가 드러났다(도 19 및 도 20a 내지 도 20f).

[0993] Cas9 및 닉카아제 또는 DNA 결합 단백질로의 전환을 위한 Cas9 효소의 돌연변이 및 변경된 기능성을 가진 그것의 사용에 대한 추가의 상세내용은 미국 출원 제61/836,101호 및 제61/835,936호(Attorney docket 44790.09.2022 및 4790.07.2022 및 Broad Reference BI-2013/004E 및 BI-2013/004F, 각각)에서 찾을 수 있으며, 이들은 여기 참고로 포함된다.

[0994] *실시예 6: Cas9 오솔로그*

[0995] 출원인은 관련 PAM 서열과 상응하는 키메라 가이드 RNA를 확인하기 위해서 Cas9 오솔로그를 분석했다. 확장된 PAM 세트는 게놈 전체에서 더 광범한 표적화를 제공하며, 또한 독특한 표적 부위의 수를 유의하게 증가시키고, 게놈에서 증가된 특이성 수준으로 신규 Cas9를 확인할 수 있는 가능성을 제공한다.

[0996] Cas9 오솔로그의 특이성은 가이드 RNA와 그것의 DNA 표적 사이의 미스매치를 용인할 수 있는 각 Cas9의 능력을 시험함으로써 평가할 수 있다. 예를 들어, SpCas9의 특이성이 절단 효율에 대한 가이드 RNA의 효과를 시험함으로써 특성화되었다. 가이드 RNA의 라이브러리가 가이드 서열과 표적 DNA 간에 단일 또는 다중 미스매치를 갖도록 작성하였다. 이들 발견에 기초하여, SpCas9에 대한 표적 부위가 다음의 가이드라인에 기초하여 선택할 수 있다:

[0997] 특정 유전자를 편집하는 것에 대한 SpCas9 특이성을 최대화하기 위해서, 잠재적 '표적외' 게놈 서열이 다음의 네 가지 제약을 따르도록 관심대상의 유전자좌 내에서 표적 부위를 선택해야 하며, 네 가지 제약은 다음과 같다: 가장 먼저 이들 뒤에 5' NGG 또는 NAG 서열을 가진 PAM 이 오지 않아야 한다. 둘째, 표적 서열에 대한 전체적 서열 유사성이 최소화되어야 한다. 셋째, 미스매치의 최대수는 표적외 부위의 PAM-근위 영역 내에 있어야 한다. 마지막으로, 미스매치의 최대수는 연속되거나 또는 4 개 염기보다 적게 떨어져 있어야 한다.

[0998] 다른 Cas9 오솔로그의 특이성을 평가하고 표적 종의 게놈 내에서 특정 표적 부위의 선택을 위한 기준을 확립하기 위해서 유사한 방법을 사용할 수 있다. 앞서 언급한 대로, 이 세트의 단백질 서열 보존에 기초한 계통발생적 분석은 세 그룹의 큰 Cas9(약 1400 아미노산)과 두 작은 Cas9(약 1100 아미노산)을 포함한 Cas9의 다섯 패밀리를 밝혀냈다(도 19 및 20a 내지 도 20f 참조). Cas 오솔로그에 대한 추가의 상세내용은 미국 출원 제61/836,101호 및 제61/835,936호(Attorney docket 44790.09.2022 및 4790.07.2022 및 Broad Reference BI-2013/004E 및 BI-2013/004F, 각각)에서 찾을 수 있으며, 이들은 여기 참고로 포함된다.

[0999] *실시예 7: 클로닝 및 전달을 단순화하기 위한 방법론적 개선*

[1000] 플라스미드 상에서 U6-프로모터와 가이드 RNA를 암호화하는 것보다 출원인은 가이드 RNA 상에 덧붙이기 위해 DNA 올리고와 함께 U6 프로모터를 증폭시켰다. 결과의 PCR 생성물은 가이드 RNA의 발현을 유도하기 위해 세포에 트랜스펙션될 수 있다.

[1001] U6-프로모터:: 인간 Emx1 유전자좌를 표적화하는 가이드 RNA 로 구성된 PCR 생성물의 생성을 허용하는 프라이머 쌍의 예:

[1002] 정방향 프라이머: AAACTCTAGAgagggcctatttcccatgattc (서열 번호 ___)

[1003] 역방향 프라이머 (가이드 RNA 보유, 밑줄로 표시된다):

[1004] acctctagAAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATGCTGTTTTGTTTCCAAAACAGCATAG CTCTAAAACCCCTAGTCATTGGAGGTGACGGTGTTTCGTCCTTTCCACaag (서열 번호 ___)

[1005] *실시예 8: 활성을 개선하기 위한 방법론적 개선:*

[1006] pol3 프로모터, 특히 RNA 폴리머라제 III(예를 들어, U6 또는 H1 프로모터)를 사용하는 것보다, 진핵 세포에서 가이드 RNA를 발현하기 위해서, 출원인은 T7 프로모터를 사용하여 가이드 RNA의 발현을 유도하기 위해 진핵 세포에서 T7 폴리머라제를 발현시킨다.

[1007] 이 시스템의 일례는 세 조각의 DNA의 도입을 수반할 수 있다:

[1008] 1. Cas9를 위한 발현 벡터

[1009] 2. T7 폴리머라제에 대한 발현 벡터

[1010] 3. T7 프로모터에 융합된 가이드 RNA를 함유하는 발현 벡터

[1011] *실시예 9: Cas9의 독성을 감소시키기 위한 방법론적 개선: mRNA 형태로 Cas9의 전달*

[1012] mRNA의 형태로 Cas9를 전달하는 것은 세포에서 Cas9의 일시적 발현을 가능하게 함으로써 독성을 감소시킨다. 예를 들어, 인간화된 SpCas9가 다음의 프라이머 쌍을 사용하여 증폭할 수 있다:

[1013] 정방향 프라이머 (시험관내 번역을 위해서 T7 프로모터에 부가):

[1014] TAATACGACTCACTATAGGAAGTGCGCCACCATGGCCCCAAAGAAGAAGCGG (서열 번호 ___)

[1015] 역방향 프라이머 (폴리A 테일에 부가):

[1016] GGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTttcttaCTTTTCTTTTTTGCCTGGCCG (서열 번호 ___)

[1017] 출원인은 진핵 세포에서 가이드 RNA 발현을 유도하기 위해 RNA 또는 DNA의 형태로 가이드 RNA로 세포에 Cas9 mRNA를 트랜스펙션한다.

[1018] *실시예 10: Cas9의 독성을 감소시키기 위한 방법론적 개선: 유도성 프로모터의 사용*

[1019] 출원인은 게놈 변형을 수행하기 위해 필요할 때만 Cas9를 일시적으로 변화시킨다. 유도성 시스템의 예는 테트라사이클린 유도성 프로모터(Tet-온 또는 Tet-오프), 소 분자 2-하이브리드 전사 활성화 시스템(FKBP, ABA 등), 또는 경질 유도성 시스템(피토크롬, LOV 도메인, 또는 크립토크롬)을 포함한다.

[1020] *실시예 11: 생체내 적용을 위한 Cas9 시스템의 개선*

[1021] 출원인은 작은 분자량을 가진 Cas9에 대한 메타 게놈 검색을 수행했다. 대부분의 Cas9 상동체는 꽤 크다. 예를 들어, SpCas9는 대략 1368aa 길이이며, 이것은 전달을 위해 바이러스 벡터에 쉽게 패키지되기에는 너무 크다. Cas9 상동체의 길이 분포를 나타내는 그래프가 GenBank에 기탁된 서열로부터 생성된다(도 23). 서열의 일부는 잘못 해석되었으며, 따라서 각 길이에 대한 정확한 빈도는 반드시 정확하지 않을 수 있다. 그렇지만, 그것은 Cas9 단백질의 분포에 일견을 제공하며, 이것은 더 짧은 Cas9 상동체가 있다는 것을 시사한다.

[1022] 컴퓨터 분석을 통해서 출원인은 박테리아 균주 *캄필로박터*에 1000 아미노산 미만의 2 개의 Cas9 단백질이 있다는 것을 발견했다. *캄필로박터 제주니*로부터의 하나의 Cas9의 서열이 아래 제시된다. 이 길이에서는 CjCas9가 1차 세포 및 동물 모델에 생체내 확실한 전달을 위해 AAV, 렌티바이러스, 아데노바이러스, 및 다른 바이러스 벡터에 쉽게 패키징될 수 있다. 본 발명의 바람직한 구현예에서, *스태필로코커스 아우레우스*로부터의 Cas9 단백질을 사용한다.

[1023] >캄필로박터 제주니 Cas9 (CjCas9)

MARILAFDIGISSIGWAFSENDELKDCGVRIFTKVENPKTGESLALPRRLAR
SARKRLARRKARLNHLKHLIANEFKLNYEDYQSFDESLAKAYKGSLISPYELRFRALN
ELLSKQDFARVILHIAKRRGYDDIKNSDDKEKGAILKAIKQNEEKLANYQSVGEYLYK
EYFQKFKENSKEFTNVRNKKESYERCIAQSFLKDELKLIFKKQREFGFSFSKKFEEEVL
SVAFYKRALKDFSHLVGNCSFFTDEKRAPKNSPLAFMFVALTRIINLLNNLKNTEGIL
YTKDDLNALLNEVLKNGTLTYKQTKKLLGLSDDYEFKGEKGTYFIEFKKYKEFIKAL
GEHNLSQDDLNEIAKDITLIKDEIKLKKALAKYDLNQNQIDSLSKLEFKDHLNISFKAL
KLVTPLMLEGKKYDEACNELNLKVAINEDKKDFLPAFNETYYKDEVTNPVVLRAIKE
YRKVLNALLKKYGKVHKINIELAREVGKNHSQRAKIEKEQNENYKAKKDAELECEK
LGLKINSKNILKLRLFKEQKEFCAYSGEKIKISDLQDEKMLEIDHIYPYSRSFDDSYMN
KVLVFTKQNQEKLNQTPFEAFGNDSAKWQKIEVLAKNLPTKKQKRILDKNYKDKEQ
KNFKDRNLNDTRYIARLVLNYTKDYLDFLPLSDDENTKLNDTQKGSKVHVEAKSGM
LTSALRHTWGFSAKDRNNHLHHAIDAVIIAYANNSIVKAFSDFKKEQESNSAELYAK
KISELDYKNKRKFFEPFSGFRQKVLDKIDEIFVSKPERKKPSGALHEETFRKEEEFYQS
YGGKEGVLKALELGKIRKVNGKIVKNGDMFRVDIFKHKKTNKFYAVPIYTMDFALK
VLPNKAVARSKKGEIKDWILMDENYEFCFSLYKDSLILIQTKDMQEPEFVYYNAFTSS
TVSLIVSKHDNKFETLSKNQKILFKNANEKEVIAKSIGIQNLKVFEKYIVSALGEVTKA
EFRQREDFKK. (SEQ ID NO:)

[1024]

[1025] 이 CjCas9에 대한 잠정적 tracrRNA 요소는 다음과 같다:

TATAATCTCATAAGAAATTTAAAAAGGGACTAAAATAAAGAGTTTGCG
[1026] GGACTCTGCGGGGTTACAATCCCCTAAAACCGCTTTTAAAATT (SEQ ID NO:)

[1027] 직접 반복부 서열은 다음과 같다:

[1028] ATTTTACCATAAAGAAATTTAAAAAGGGACTAAAAC (SEQ ID NO:)

[1029] CjCas9에 대한 키메라 가이드 RNA의 예는 다음과 같다:

NNNNNNNNNNNNNNNNNNNNGUUUUAGUCCCGAAAGGGACUAAAAU
[1030] AAAGAGUUUGCGGGACUCUGCGGGGUUACAAUCCCCUAAAACCGCUUUU (SEQ ID NO:)

[1031] *실시예 12: Cas9최적화*

[1032] 증진된 기능이나 새로운 기능을 개발하기 위해서, 출원인은 상이한 Cas9 상동체로부터의 단편들을 조합하여 키메라 Cas9 단백질을 생성한다. 예를 들어, 키메라 Cas9 단백질의 2 개의 예는 다음과 같다:

[1033] 예를 들어, 출원인은 St1Cas9의 N-말단 (이 단백질로부터의 단편은 굵은 글씨로 표시된다)과 SpCas9의 C 말단 (이 단백질로부터의 단편은 밑줄로 표시된다)을 융합했다.

[1034] >St1(N)Sp(C)Cas9

MSDLVLGLDIGIGSVGVGILNKVTGEIIHKNSRIFPAAQAENNLVRRTN
RQGRRLARRKKHRRVRLNRLFEESGLITDFTKISINLNPYQLRVKGLTDELSNEE
LFIALKNMVKHRGISYLDDASDDGNSSVGDYAQIVKENSKQLETKTPGQIQLERY
QTYGQLRGDFTVEKDGKKHRLINVFPTSAYRSEALRILQTQQEFNPQITDEFINR
YLEILTGKRKYYHGPGNEKSRTDYGRYRTSGETLDNIFGILIGKCTFYPDEFRAAK
ASYTAQEFNLLNDLNNLTVPTETKKLSKEQKNQIINYVKNEKAMGPAKLFKYIAK
LLSCDVADIKGYRIDKSGKAEIHTFEAYRKMKTLETLDIEQMDRETLDKLAYVLT
LNTEREGIQEALEHEFADGSFSQKQVDELVQFRKANSSIFGKGWHNFSVKLMME
LIPELYETSEEQMTILTRLGKQKTTSSSNKTKYIDEKLLTEEIYNPVVAKSVRQAIK
IVNAAIKEYGDFDNIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVE
NTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLT
RSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDK
AGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQ
FYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQ
EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKV
LSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSV
LVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYS
LFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLF
VEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNL
GAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD (SEQ ID NO:)

[1035]

[1036] >Sp(N)St1(C)Cas9

MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVE
EDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRG
HFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENL
IAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG
DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQ
QLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLL
RKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGN
SRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYE
YFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIE
CFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIE
ERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFAN
RNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELV
KVMGRHKPENIVIEMARETNEDDEKKAIQKIQKANKDEKDAAMLKAANQYNGKA
ELPHSVFHGHKQLATKIRLWHQQGERCLYTGKTISIHDLINNSNQFEVDHILPLSI
TFDDSLANKVLVYATANQEKGQRTPYQALDSMDDAWSFRELKAFVRESKTLSNK
KKEYLLTEEDISKFDVRKKFIERNLVDTRYASRVVLNALQEHFRAHKIDTKVSVVR
GQFTSQLRRHWGIEKTRDTYHHHAVDALIIAASSQLNLWKKQKNTLVSYSEDQL
LDIETGELISDDEYKESVFKAPYQHFVDTLKSKEFEDSILFSYQVDSKFNRKISDATI
YATRQAKVGKDKADETYVLGKIKDIYTQDGYDAFMKIYKKDKSKFLMYRHDPQT
FEKVIEPILENYPNKQINEKGKEVPCNPFLKYKEEHGYIRKYSKKGNGPEIKSLKYY
DSKLGNHIDITPKDSNNKVVLQSVSPWRADVYFNKTTGKYEILGLKYADLQFEKG
TGTYKISQEKYNDIKKKEGVDSDSEFKFTLYKNDLLLVKDTETKEQQLFRFLSRT
MPKQKHYVELKPYDKQKFEGGEALIKVLGNVANSGQCKKGLGKSNISIYKVRTD
VLGNQHIIKNEGDKPKLDF (SEQ ID NO:)

[1037]

[1038] 키메라 Cas9를 제조하는 것의 이점은 다음을 포함한다:

[1039] 독성 감소

[1040] 진핵 세포에서의 발현 개선

[1041] 특이성 증진

[1042] 단백질의 분자량 감소, 상이한 Cas9 상동체로부터 가장 작은 도메인들을 조합함으로써 더 작은 단백질 제조.

[1043] PAM 서열 요건 변경

[1044] *실시예 13: 일반적인 DNA 결합 단백질로서 Cas9의 활용*

[1045] 출원인은 DNA 표적의 양쪽 가닥의 절단을 초래하는 2개의 촉매 도메인 (D10 및 H840)을 돌연변이시킴으로써 일반적인 DNA 결합 단백질로서 Cas9를 사용했다. 표적 유전자좌에서 유전자 전사를 상향조절하기 위해서, 출원인은 전사 활성화 도메인 (VP64)과 Cas9를 융합했다. 출원인은 전사 인자 활성화 강도가 표적에서 소모되는 시간의 함수이기 때문에 Cas9-VP64 융합 단백질의 강한 핵 위치를 아는 것이 중요할 것이라는 가설을 세웠다. 따라서, 출원인은 일련의 Cas9-VP64-GFP 작제물을 클로닝하고, 이들을 293 세포에 트랜스펙션하고, 트랜스펙션 12시간 후에 형광 현미경 아래서 이들의 국소화를 평가했다.

[1046] 방해하기 위해 존재하는 벌키한 GFP 없이 작제물을 기능적으로 시험하기 위해서 동일한 작제물이 직접 융합이 아니라 2A-GFP로서 클로닝되었다. 출원인은 세포 리프로그램에 유용할 수 있기 때문에 Cas9 트랜스활성인자를 가진 Sox2 유전자좌를 표적화하도록 선택했으며, 이 유전자좌는 TALE-TF 매개 전사 활성화에 대한 표적으로서 이미 검증되었다. Sox2 유전자좌에 대해, 출원인은 전사 시작 부위(TSS) 근처에서 8 개의 표적을 선택했다. 각 표적은 20 bp 길이였고, 이웃한 NGG 프로토스페이서 인접 모티프(PAM)를 가진다. 각 Cas9-VP64 작제물이 293 세포에서 각 PCR 생성된 키메라 crispr RNA (chiRNA)로 공동 트랜스펙션시켰다. 트랜스펙션 72시간 후 전화 활성화를 RT-qPCR을 사용하여 평가하였다.

[1047] 전사 활성인자를 더 최적화하기 위하여, 출원인은 Cas9 (NLS-VP64-NLS-hSpCas9-NLS-VP64-NLS)에 대한 chiRNA (Sox2.1 및 Sox2.5)의 비율을 적정하고, 293 세포에 트랜스펙션하고, RT-qPCR을 사용하여 정량했다. 이들 결과는 Cas9가 표적 유전자좌에서 유전자 전사를 상향조절하기 위한 일반적인 DNA 결합 도메인으로서 사용될 수 있음을 나타낸다.

[1048] 출원인은 제2 세대의 작제물을 설계했다(아래 표).

[1049]

p 렌티-EF1a-GFP-2A-6xHis-NLS-VP64-NLS-hSpCsn1(D10A, H840A)-NLS
p 렌티-EF1a-GFP-2A-6xHis-NLS-VP64-NLS-hSpCsn1(D10A, H840A)
p 렌티-EF1a-GFP-2A-6xHis-NLS-VP64-NLS-NLS-hSpCsn1(D10A, H840A)
p 렌티-EF1a-GFP-2A-6xHis-NLS-hSpCsn1(D10A, H840A)-NLS
p 렌티-EF1a-GFP-2A-6xHis-NLS-hSpCsn1(D10A, H840A)
p 렌티-EF1a-GFP-2A-6xHis-NLS-NLS-hSpCsn1(D10A, H840A)

[1050] 출원인은 이들 작제물을 사용하여 RT-qPCR에 의해서 전사 활성화 (VP64 융합 작제물) 및 억제(Cas9 단독)를 평가한다. 출원인은 항-His 항체를 사용하여 각 작제물의 세포 국소화, 서베이어 뉴클레아제 분석을 사용하여 뉴클레아제 활성, 그리고 겔 이동 분석을 사용하여 DNA 결합 친화성을 평가한다.

[1051] *실시예 14: Cas9 유전자이식 및 녹인 마우스*

[1052] Cas9 뉴클레아제를 발현하는 마우스를 생성하기 위해서, 출원인은 두 가지 일반적인 전략을 제시하는데, 유전자이식과 녹인이다. 이들 전략은, 예를 들어 래트의 경우, 관심대상의 어떤 다른 모델 유기체를 생성하기 위해 적용될 수 있다. 일반적인 전략의 각각의 경우, 출원인은 구성적으로 활성인 Cas9와 조건적으로 발현되는 Cas9를 제작했다(Cre 재조합효소 의존적). 구성적으로 활성인 Cas9 뉴클레아제는 다음의 맥락에서 발현된다: pCAG-NLS-Cas9-NLS-P2A-EGFP-WPRE-bGH폴리A. pCAG는 프로모터, NLS는 핵 위치 신호, P2A는 펩티드 절단 서열, EGFP는 증진된 녹색 형광 단백질, WPRE는 우드처크 간염 바이러스 전사 후 조절 구성요소, bGHpolyA는 소 성장 호르몬 폴리-A 신호 서열이다(도 25a 내지 도 25b). 조건적 버전은 프로모터 뒤와 NLS-Cas9-NLS 앞에 하나의 추가적 중단 카세트 요소로서 loxP-SV40폴리A x3-loxP를 가진다(즉, pCAG-loxP-SV40폴리Ax3-loxP-NLS-Cas9-NLS-P2A-EGFP-

WPRE-bGH폴리A). 중요한 발현 요소는 도 26과 같이 시각화될 수 있다. 구성적 작제물은 발생을 통해서 모든 세포 타입에서 발현되지만, 조건적 작제물은 동일한 세포가 Cre 재조합효소를 발현할 때만 Cas9 발현을 허용할 것이다. 후자의 TBG 은 Cre가 조직 특이적 프로모터의 발현하에 있을 때 Cas9의 조직 특이적 발현을 허용할 것이다. 더욱이, Cas9 발현은 TET-온 또는 -오프 시스템과 같은 유도성 프로모터의 발현하에 Cre를 둠으로써 성체 마우스에서 유도될 수 있다.

[1053] Cas9 작제물의 확인: 각 플라스미드는 세 가지 방식으로 기능적으로 검증되었다: 1) 293 세포에서의 일시적 트랜스펙션 후 GFP 발현의 확인; 2) 293 세포에서의 일시적 트랜스펙션 후 P2A 서열을 인식하는 항체를 사용한 면역형광; 및 3) 일시적 트랜스펙션 후 서베이어 뉴클레아제 분석. 293 세포는 관심대상인 세포에 따라 293FT 또는 293 T 세포일 수 있다. 바람직한 구현예에서, 세포는 293FT 세포이다. 서베이어의 결과는 조건적 작제물과 구성적 작제물에 대해 각각 겔의 위와 아래에서 판독하였다. 각각은 hEMX1 유전자좌(키메라 RNA hEMX1.1)에 표적화된 키메라 RNA의 존재 또는 부재하에 시험되었다. 결과는 이 작제물이 키메라 RNA (및 조건적인 경우에는 Cre)가 존재할 때만 hEMX1 유전자좌를 성공적으로 표적화할 수 있음을 나타낸다. 겔이 정량화하며, 그 결과를 세 샘플의 평균 절단 효율 및 표준 편차로서 제시한다.

[1054] 유전자이식 Cas9 마우스: 작제물을 사용하여 유전자이식 마우스를 생성하기 위해서, 출원인은 거짓 임신 CB56 암컷으로부터의 접합체의 전핵에 순수한 선형 DNA를 주입한다. 시조가 확인되고, 제노타이핑되고, CB57 마우스에 다시 교배시킨다. 작제물은 성공적으로 클로닝하였고, 생어(Sanger) 시퀀싱에 의해서 검증하였다.

[1055] 녹인 Cas9 마우스: Cas9 녹인 마우스를 생성하기 위해서, 출원인은 Rosa26 유전자좌에 일부 구성적 및 조건적 작제물을 표적화한다. 출원인은 이것을 각 작제물을 다음의 요소를 가진 벡터를 표적화하는 Rosa26에 클로닝함으로써 행했다: Rosa26 짧은 상동체 암 - 구성적/조건적 Cas9 발현 카세트 - pPGK-Neo-Rosa26 긴 상동체 암 - pPGK-DTA. pPGK는 네오마이신에 대한 내성을 부여하는 양성 선택 마커 Neo, 1 kb 짧은 팔, 4.3 kb 긴 팔, 및 PGK에 의해서 유도되는 음성 선택 초파리 독소 (DTA)에 대한 프로모터이다.

[1056] 두 작제물을 R1 mESC에 전기천공하고, 네오마이신 선택이 적용되기 전에 2 일 동안 성장시켰다. 5-7 일까지 생존한 각 콜로니를 선별하고 개별 웰에서 성장시켰다. 5-7 일 후 콜로니를 수거하고, 절반을 냉동시키고, 나머지 절반은 제노타이핑에 사용했다. 제노타이핑은 게놈 PCR에 의해서 수행하였는데, 여기서는 하나의 프라이머는 도너 플라스미드 (AttpF) 내에서 아닐링되고, 나머지 프라이머는 짧은 상동체 암(Rosa26-R) 밖에서 아닐링되었다. 조건적 경우에 수거된 22 개의 콜로니 중 7 개가 양성이었다(왼쪽). 구성적 경우에 수거된 27 개의 콜로니 중 양성은 하나도 없었다(오른쪽). Cas9는 mESC에서 어떤 수준의 독성을 야기하고, 이 이유 때문에 양성 클론은 없었을 수 있다. 이것을 시험하기 위해서, 출원인은 정확히 표적화된 조건적 Cas9 세포에 Cre 발현 플라스미드를 도입했고, 수 일 배양 후에 매우 낮은 독성을 발견했다. 정확히 표적화된 조건적 Cas9 세포에서 Cas9의 감소된 복제수(세포당 1 내지 2 복제)는 안정한 발현을 허용하기에 충분하며, 세포독성은 비교적 없는 편이다. 더욱이, 이 데이터는 Cas9 복제수가 독성을 결정한다는 것을 나타낸다. 전기천공 후, 각 세포는 Cas9의 몇 개의 복제를 얻는데, 이것은 구성적 Cas9 작제물에서 양성 콜로니가 발견되지 않은 이유일 수 있다. 이것은 조건적, Cre-의존적 전략을 이용하는 것이 감소된 독성을 나타낸다는 것의 강력한 증거를 제공한다. 출원인은 정확히 표적화된 세포를 배반포에 주입해서 암컷 마우스에 이식한다. 키메라가 확인되고 역교배된다. 시조가 확인되고 제노타이핑된다.

[1057] 조건적 Cas9 마우스의 활용: 출원인은 293 세포에서 Cas9 조건적 발현 작제물이 Cre와의 공-발현에 의해서 활성화될 수 있다는 것을 나타냈다. 출원인은 또한 정확하게 표적화된 R1 mESC는 Cre가 발현될 때 활성 Cas9를 가질 수 있다는 것을 나타낸다. Cas9 뒤에는 P2A 펩티드 절단 서열이 오고 다음에 EGFP가 이어지기 때문에, 출원인은 EGFP를 관찰함으로써 성공적인 발현을 확인한다. 동일한 개념은 조건적 Cas9 마우스를 아주 유용하게 만드는 것이다. 출원인은 조건적 Cas9 마우스를 Cre를 편재적으로 발현하는 마우스(ACTB-Cre 계통)와 교배시킬 수 있고, 모든 세포에서 Cas9를 발현하는 마우스에 이를 수 있다. 배아 또는 성체 마우스에서 게놈 편집을 유도하기 위하여 키메라 RNA의 전달을 고려해야 한다. 흥미롭게도, 조건적 Cas9 마우스가 조직 특이적 프로모터 하에 Cre를 발현하는 마우스와 교배된다면, 또한 Cre를 발현하는 조직에 Cas9만 있어야 한다. 이 접근법은 동일한 조직에 키메라 RNA를 전달함으로써 정확한 조직에서만 게놈을 편집하기 위해서 사용될 수 있다.

[1058] *실시예 15: Cas9 전달 및 키메라 RNA*

[1059] CRISPR-Cas 시스템은 박테리아 및 고세균류를 통틀어서 다양한 종들에 의해서 이용되는 침입한 외인성 DNA에 대한 후천 면역 메커니즘이다. 타입 II CRISPR-Cas 시스템은 CRISPR 유전자좌에 외래 DNA의 "획득"을 책임지는 단백질을 암호화하는 일련의 유전자와 DNA 절단 메커니즘의 "실행"을 암호화하는 일련의 유전자로 구성되며; 이들

은 DNA 뉴클레아제(Cas9), 비-코딩 트랜스활성화 cr-RNA(tracrRNA), 및 직접 반복부(crRNA)가 측접된 외래 DNA-유래 스페이서 어레이를 포함한다. Cas9에 의한 성숙시, tracrRNA와 crRNA 듀플렉스는 스페이서 가이드 서열에 의해서 특정된 표적 DNA 서열로 Cas9 뉴클레아제를 가이드하고, 절단에 필요하며 각 CRISPR-Cas 시스템에 특이적인 표적 DNA의 짧은 서열 모티브 근처에서 DNA에 이중가닥 브레이크를 매개한다. 타입 II CRISPR-Cas 시스템은 박테리아 계 전체에서 발견되며, Cas9 단백질 서열 및 크기, tracr RNA 및 crRNA 직접 반복부 서열, 이들 요소들의 게놈 조직, 및 표적 절단을 위한 모티프 요건은 매우 다양하다. 하나의 종이 다수의 상이한 CRISPR-Cas 시스템을 가질 수 있다.

[1060] 출원인은 Cas9로 알려진 서열 상동체와 HNH 엔도뉴클레아제 도메인 및 RuvC 엔도뉴클레아제 도메인을 포함하는 서브 도메인에 오솔로그인 구조들에 기초하여 확인된, 박테리아 종들로부터의 207 개의 잠정적 Cas9를 평가했다[Eugene Koonin and Kira Makarova로부터의 정보]. 이 세트의 단백질 서열 보존에 기초한 계통발생적 분석은 세 그룹의 큰 Cas9 (약 1400 아미노산)와 두 그룹의 작은 Cas9 (약 1100 아미노산)을 포함하는 다섯 개의 Cas9 패밀리를 드러냈다(도 19a 내지 도 19d 및 도 20a 내지 도 20f).

[1061] 출원인은 또한 시험관내 방법을 사용하여 Cas9 가이드 RNA를 최적화했다.

[1062] *실시예 16: Cas9 돌연변이*

[1063] 이러한 실시예에서, 본 발명자들은 하기의 돌연변이가 SpCas9를 닉킹 효소로 전환시킬 수 있음을 보여준다: D10A, E762A, H840A, N854A, N863A, D986A.

[1064] 출원인은 돌연변이 점이 SpCas9 유전자 내에 위치하는 곳을 보여주는 서열을 제공한다(도 24a 내지 도 24m). 또한, 본 발명자들은 닉카아제가 여전히 상동성 재조합을 매개할 수 있음을 보여준다. 추가로, 본 발명자들은 이들 돌연변이가 있는 SpCas9(개별적으로)가 이중 가닥 파단을 유도하지 않음을 보여준다.

[1065] Cas9 오솔로그는 모두 3-4 RuvC 도메인과 HNH 도메인의 일반적인 조직을 공유한다. 5' 최단부 RuvC 도메인은 비-상보적 가닥을 절단하고, HNH 도메인은 상보적 가닥을 절단한다. 모든 표시법은 가이드 서열을 참조한다.

[1066] 5' RuvC 도메인의 촉매 잔기는 관심대상의 Cas9와 다른 Cas9 오솔로그 (*S. pyogenes* 타입 II CRISPR 유전자좌, *스트렙토코커스 써모필루스* CRISPR 유전자좌 1, *스트렙토코커스 써모필루스* CRISPR 유전자좌 3, 및 *Francisciella novicida* 타입 II CRISPR 유전자좌로부터의)의 상동성 비교를 통해서 확인되고, 보존된 Asp 잔기가 알라닌으로 돌연변이되며, 이로써 Cas9가 상보성-가닥 흡터형성 효소로 전환된다. 유사하게, HNH 도메인에서 보존된 His 및 Asn 잔기가 알라닌으로 돌연변이되어 Cas9가 비-상보성-가닥 흡터형성 효소로 전환된다.

[1067] *실시예 17: Cas9 전사 활성화 및 Cas9 억제인자*

[1068] Cas9 전사 활성화

[1069] 제2 세대 작제물이 설계되고 시험되었다(표 1). 이들 작제물을 사용하여 RT-qPCR에 의해서 전사 활성화 (VP64 융합 작제물) 및 억제(Cas9 단독)를 평가한다. 출원인은 항-His 항체를 사용하여 각 작제물의 세포 국소화, 서베이어 뉴클레아제 분석을 사용하여 뉴클레아제 활성, 그리고 겔 이동 분석을 사용하여 DNA 결합 친화성을 평가한다.

[1070] Cas 억제인자

[1071] dCas9가 유전자 발현을 억제하기 위한 일반적인 DNA 결합 도메인으로 사용될 수 있다는 것은 이미 밝혀졌다. 출원인은 개선된 dCas9 디자인뿐만 아니라 억제인자 도메인 KRAB 및 SID4x와 dCas9 융합을 보고한다. 표 1에서 Cas9를 사용하여 전사를 조절하기 위해 생성된 플라스미드 라이브러리로부터 다음의 억제인자 플라스미드들, pXRP27, pXRP28, pXRP29, pXRP48, pXRP49, pXRP50, pXRP51, pXRP52, pXRP53, pXRP56, pXRP58, pXRP59, pXRP61, 및 pXRP62가 qPCR에 의해서 기능적으로 특성화되었다.

[1072] 각 dCas9 억제인자 플라스미드는 베타-카테닌 유전자의 코딩 가닥에 표적화된 2 개의 가이드 RNA와 공-트랜스펙션되었다. 트랜스펙션 72 시간 후 RNA가 분리되었고, RT-qPCR에 의해서 유전자 발현이 정량되었다. 내인성 대조군 유전자는 GAPDH였다. 양성 대조군으로서 2 개의 검증된 shRNA가 사용되었다. 음성 대조군은 gRNA 없이 트랜스펙션된 특정 플라스미드였으며, 이들은 "pXRP## 대조군"으로 표시된다. 플라스미드 pXRP28, pXRP29, pXRP48, 및 pXRP49는 명시된 표적화 전략을 사용할 때 베타-카테닌 유전자를 억제할 수 있었다. 이들 플라스미드는 기능적 도메인이 없는 dCas9 (pXRP28 및 pXRP28) 및 SID4x와 융합된 dCas9 (pXRP48 및 pXRP49)에 상응한다.

[1073] 추가 연구 조사: 상기 실험을 반복하고, 상이한 유전자를 표적화하고, 다른 gRNA들을 이용하여 최적 표적화 위

치를 결정하고, 복합적인 억제를 달성한다.

표 1

pXRP024-p 렌티 2-EF1a-VP64-NLS-FLAG-링커-dCas9-NLS-gLuc-2A-GFP-WPRE
pXRP025-p 렌티 2-EF1a-VP64-NLS-$GGGGS_3$링커-dCas9-NLS-gLuc-2A-GFP-WPRE
pXRP026-p 렌티 2-EF1a-VP64-NLS-$EAAAK_3$링커-dCas9-NLS-gLuc-2A-GFP-WPRE
pXRP027-p 렌티 2-EF1a-NLS-FLAG-링커-dCas9-NLS-gLuc-2A-GFP-WPRE
pXRP028-p 렌티 2-EF1a-NLS-$GGGGS_3$링커-dCas9-NLS-gLuc-2A-GFP-WPRE
pXRP029-p 렌티 2-EF1a-NLS-$EAAAK_3$링커-dCas9-NLS-gLuc-2A-GFP-WPRE
pXRP030-p 렌티 2-pSV40-VP64-NLS-FLAG-링커-dCas9-NLS-gLuc-2A-GFP-WPRE
pXRP031-p 렌티 2-pPGK-VP64-NLS-FLAG-링커-dCas9-NLS-gLuc-2A-GFP-WPRE
pXRP032-p 렌티 2-LTR-VP64-NLS-FLAG-링커-dCas9-NLS-gLuc-2A-GFP-WPRE
pXRP033-p 렌티 2-pSV40-VP64-NLS-$GGGGS_3$링커-dCas9-NLS-gLuc-2A-GFP-WPRE
pXRP034-p 렌티 2-pPGK-VP64-NLS-$GGGGS_3$링커-dCas9-NLS-gLuc-2A-GFP-WPRE
pXRP035-p 렌티 2-LTR-VP64-NLS-$GGGGS_3$링커-dCas9-NLS-gLuc-2A-GFP-WPRE
pXRP036-p 렌티 2-pSV40-VP64-NLS-$EAAAK_3$링커-dCas9-NLS-gLuc-2A-GFP-WPRE
pXRP037-p 렌티 2-pPGK-VP64-NLS-$EAAAK_3$링커-dCas9-NLS-gLuc-2A-GFP-WPRE
pXRP038-p 렌티 2-LTR-VP64-NLS-$EAAAK_3$링커-dCas9-NLS-gLuc-2A-GFP-WPRE
pXRP048-p 렌티 2-EF1a-SID4x-NLS-FLAG-링커-dCas9-NLS-gLuc-2A-GFP-WPRE
pXRP049-p 렌티 2-EF1a-SID4X-NLS-$GGGGS_3$링커-dCas9-NLS-gLuc-2A-GFP-WPRE
pXRP050-p 렌티 2-EF1a-SID4X-NLS-$EAAAK_3$링커-dCas9-NLS-gLuc-2A-GFP-WPRE
pXRP051-p 렌티 2-EF1a-KRAB-NLS-FLAG-링커-dCas9-NLS-gLuc-2A-GFP-WPRE
pXRP052-p 렌티 2-EF1a-KRAB-NLS-$GGGGS_3$링커-dCas9-NLS-gLuc-2A-GFP-WPRE
pXRP053-p 렌티 2-EF1a-KRAB-NLS-$EAAAK_3$링커-dCas9-NLS-gLuc-2A-GFP-WPRE
pXRP054-p 렌티 2-EF1a-dCas9-링커-FLAG-NLS-VP64-gLuc-2A-GFP-WPRE
pXRP055-p 렌티 2-EF1a-dCas9-링커-FLAG-NLS-SID4X-gLuc-2A-GFP-WPRE
pXRP056-p 렌티 2-EF1a-dCas9-링커-FLAG-NLS-KRAB-gLuc-2A-GFP-WPRE
pXRP057-p 렌티 2-EF1a-dCas9-$GGGGGS_3$-NLS-VP64-gLuc-2A-GFP-WPRE
pXRP058-p 렌티 2-EF1a-dCas9-$GGGGGS_3$-NLS-SID4X-gLuc-2A-GFP-WPRE
pXRP059-p 렌티 2-EF1a-dCas9-$GGGGGS_3$-NLS-KRAB-gLuc-2A-GFP-WPRE

[1074]

pXRP060-p 렌티 2-EF1a-dCas9-EAAAK$_3$-NLS-VP64-gLuc-2A-GFP-WPRE
pXRP061-p 렌티 2-EF1a-dCas9-EAAAK$_3$-NLS-SID4X-gLuc-2A-GFP-WPRE
pXRP062-p 렌티 2-EF1a-dCas9-EAAAK$_3$-NLS-KRAB-gLuc-2A-GFP-WPRE
pXRP024-p 렌티 2-EF1a-VP64-NLS-FLAG-링커-Cas9-NLS-gLuc-2A-GFP-WPRE
pXRP025-p 렌티 2-EF1a-VP64-NLS-GGGGS$_3$링커-Cas9-NLS-gLuc-2A-GFP-WPRE
pXRP026-p 렌티 2-EF1a-VP64-NLS-EAAAK$_3$링커-Cas9-NLS-gLuc-2A-GFP-WPRE
pXRP027-p 렌티 2-EF1a-NLS-FLAG-링커-Cas9-NLS-gLuc-2A-GFP-WPRE
pXRP028-p 렌티 2-EF1a-NLS-GGGGS$_3$링커-Cas9-NLS-gLuc-2A-GFP-WPRE
pXRP029-p 렌티 2-EF1a-NLS-EAAAK$_3$링커-Cas9-NLS-gLuc-2A-GFP-WPRE
pXRP030-p 렌티 2-pSV40-VP64-NLS-FLAG-링커-Cas9-NLS-gLuc-2A-GFP-WPRE
pXRP031-p 렌티 2-pPGK-VP64-NLS-FLAG-링커-Cas9-NLS-gLuc-2A-GFP-WPRE
pXRP032-p 렌티 2-LTR-VP64-NLS-FLAG-링커-Cas9-NLS-gLuc-2A-GFP-WPRE
pXRP033-p 렌티 2-pSV40-VP64-NLS-GGGGS$_3$링커-Cas9-NLS-gLuc-2A-GFP-WPRE
pXRP034-p 렌티 2-pPGK-VP64-NLS-GGGGS$_3$링커-Cas9-NLS-gLuc-2A-GFP-WPRE
pXRP035-p 렌티 2-LTR-VP64-NLS-GGGGS$_3$링커-Cas9-NLS-gLuc-2A-GFP-WPRE
pXRP036-p 렌티 2-pSV40-VP64-NLS-EAAAK$_3$링커-Cas9-NLS-gLuc-2A-GFP-WPRE
pXRP037-p 렌티 2-pPGK-VP64-NLS-EAAAK$_3$링커-Cas9-NLS-gLuc-2A-GFP-WPRE
pXRP038-p 렌티 2-LTR-VP64-NLS-EAAAK$_3$링커-Cas9-NLS-gLuc-2A-GFP-WPRE
pXRP048-p 렌티 2-EF1a-SID4x-NLS-FLAG-링커-Cas9-NLS-gLuc-2A-GFP-WPRE
pXRP049-p 렌티 2-EF1a-SID4X-NLS-GGGGS$_3$링커-Cas9-NLS-gLuc-2A-GFP-WPRE
pXRP050-p 렌티 2-EF1a-SID4X-NLS-EAAAK$_3$링커-Cas9-NLS-gLuc-2A-GFP-WPRE
pXRP051-p 렌티 2-EF1a-KRAB-NLS-FLAG-링커-Cas9-NLS-gLuc-2A-GFP-WPRE
pXRP052-p 렌티 2-EF1a-KRAB-NLS-GGGGS$_3$링커-Cas9-NLS-gLuc-2A-GFP-WPRE
pXRP053-p 렌티 2-EF1a-KRAB-NLS-EAAAK$_3$링커-Cas9-NLS-gLuc-2A-GFP-WPRE
pXRP054-p 렌티 2-EF1a-Cas9-링커-FLAG-NLS-VP64-gLuc-2A-GFP-WPRE
pXRP055-p 렌티 2-EF1a-Cas9-링커-FLAG-NLS-SID4X-gLuc-2A-GFP-WPRE
pXRP056-p 렌티 2-EF1a-Cas9-링커-FLAG-NLS-KRAB-gLuc-2A-GFP-WPRE
pXRP057-p 렌티 2-EF1a-Cas9-GGGGGS$_3$-NLS-VP64-gLuc-2A-GFP-WPRE
pXRP058-p 렌티 2-EF1a-Cas9-GGGGGS$_3$-NLS-SID4X-gLuc-2A-GFP-WPRE
pXRP059-p 렌티 2-EF1a-Cas9-GGGGGS$_3$-NLS-KRAB-gLuc-2A-GFP-WPRE
pXRP060-p 렌티 2-EF1a-Cas9-EAAAK$_3$-NLS-VP64-gLuc-2A-GFP-WPRE
pXRP061-p 렌티 2-EF1a-Cas9-EAAAK$_3$-NLS-SID4X-gLuc-2A-GFP-WPRE
pXRP062-p 렌티 2-EF1a-Cas9-EAAAK$_3$-NLS-KRAB-gLuc-2A-GFP-WPRE

[1075]

[1076] *실시예 18: 콜레스테롤 생합성, 지방산 생합성 및 다른 대사 장애에 수반된 유전자, 아밀로이드 및 다른 질환에 수반된 미스-폴딩 단백질을 암호화하는 유전자, 세포 형질전환을 초래하는 암유전자, 잠복 바이러스 유전자, 및 특히 우성-음성 장애를 초래하는 유전자의 표적화된 결실*

[1077] 출원인은 바이러스 또는 나노입자 전달 시스템을 사용하여, 대사 장애, 아밀로이드증 및 단백질-응집 관련 질환, 유전자 돌연변이 및 전위로 인해 발생한 세포 형질전환, 유전자 돌연변이의 우성 음성 효과, 잠복 바이러스 감염, 및 다른 관련된 증상으로 고통받는, 필요한 대상체나 환자의 간, 뇌, 안구, 상피, 조혈, 또는 다른 조직에서 CRISPR-Cas 시스템의 유전자 전달을 증명한다.

[1078] **연구 설계:** 대사 장애, 아밀로이드증 및 단백질-응집 관련 질환으로 고통받는 치료가 필요한 대상체 또는 환자는, 제한은 아니지만, 인간, 비-영장류 인간, 개, 고양이, 소, 말, 다른 가축 및 관련 포유류를 포함한다. CRISPR-Cas 시스템은 키메라 가이드 RNA에 의해서 가이드되어 절단될 인간 게놈 유전자좌의 특정 부위를 표적화한다. 절단 및 비-상동성 단부-연결 매개 수선 후, 프레임-이동 돌연변이는 유전자의 녹아웃을 가져온다.

[1079] 출원인은 최소한의 표적외 활성으로 내인성 유전자좌에 특이적이 되는 상기 언급된 장애들에 수반되는 가이드-RNA 표적화 유전자를 선택한다. 둘 이상의 가이드 RNA가 단일 CRISPR 어레이에 암호화될 수 있으며, 이로써 DNA에 이중가닥 브레이크가 동시에 유도됨으로써 질환에 걸린 유전자 또는 염색체 영역의 미소-결실이 초래된다.

[1080] 유전자 표적의 확인 및 설계

[1081] 각 후보 질환 유전자에 대해, 출원인은 단백질-코딩 엑손, 기지의 우성 음성 돌연변이 부위를 포함하며 측접한 서열, 병리적 반복 서열을 포함하며 측접한 서열을 포함하는 관심대상의 DNA 서열을 선택한다. 유전자-녹아웃 접근법을 위해, 시작 코돈에 가장 가까운 조기 코딩 엑손이 완전한 녹아웃을 달성하고 일부 기능을 보유한 말단

절단된 단백질 생성물의 가능성을 최소화하기 위한 최상의 옵션을 제공한다.

[1082] 출원인은 NGG 모티프(SpCas9 시스템의 경우) 또는 NNAGAAW (St1Cas9 시스템의 경우)에 대해 바로 5'에 있는 모든 가능한 표적화가능한 20-bp 서열에 대해 관심대상의 서열을 분석한다. 출원인은 특이성을 결정하기 위한 컴퓨터 알고리즘에 기초하여 표적외 효과를 최소화하기 위해서 게놈에서 특유의 단일 RNA-가이드된 Cas9 인식을 위해 서열을 선택했다.

[1083] 전달 시스템에 가이드 서열의 클로닝

[1084] 이중가닥 20-24 bp 올리고뉴클레오티드로서 가이드 서열이 합성된다. 올리고의 5'포스포릴화 처리 및 듀플렉스를 형성하기 위한 아닐링 후, 전달 방법에 따라서 올리고가 적합한 벡터에 결찰된다:

[1085] 바이러스-기반 전달 방법

[1086] AAV-기반 벡터(PX260, 330, 334, 335)는 다른 곳에 설명되었다.

[1087] 렌티바이러스-기반 벡터는 가이드 서열을 U6 프로모터-유도된 키메라 RNA 스캐폴드와 EF1a 프로모터-유도된 Cas9 또는 Cas9 닉카아제를 지닌 단일 벡터에 직접 결찰하는 유사한 클로닝 전략을 사용한다.

[1088] 바이러스 생성은 다른 곳에 설명된다.

[1089] 나노입자-기반 RNA 전달 방법

[1090] 1. 가이드 서열은 T7 프로모터-가이드 서열-키메라 RNA를 암호화하는 올리고뉴클레오티드 듀플렉스로서 합성된다. T7 프로모터는 PCR 방법에 의해서 Cas9의 5'에 부가된다.

[1091] 2. T7-유도된 Cas9 및 가이드-키메라 RNA가 시험관내 전사되고, Cas9 mRNA가 더 봉쇄되며, 상업적 키트를 사용하여 A-테일 형식으로 된다. RNA 생성물을 키트 설명서에 따라서 정제한다.

[1092] 유체역학적 꼬리 정맥 전달 방법 (마우스의 경우)

[1093] 가이드 서열을 본 출원에서 다른 곳에 상기 설명된 대로 AAV 플라스미드에 클로닝한다.

[1094] *시험관내 세포주 검증*

[1095] 트랜스펙션

[1096] 1. DNA 플라스미드 트랜스펙션

[1097] 가이드 서열을 지닌 플라스미드가 지질-, 화학적- 또는 전기천공-기반 방법을 사용하여 인간 배아 신장 (HEK293T) 또는 인간 배아 줄기 (hES) 세포, 다른 관련 세포 타입에 트랜스펙션된다. HEK293T 세포의 24-웰 트랜스펙션을 위해 (약 260,000 세포), 총 500 ng의 DNA가 리포펙타민 2000을 사용하여 각 단일 웰에 트랜스펙션된다. hES 세포의 12-웰 트랜스펙션을 위해, 총 1 ㎍의 DNA는 Fugene HD를 사용하여 단일 웰에 트랜스펙션한다.

[1098] 2. RNA 트랜스펙션

[1099] 상기 설명된 정제된 RNA가 HEK293T 세포에의 트랜스펙션에 사용된다. 1-2 ㎍의 RNA가 제조자의 지시에 따라서 리포펙타민 2000을 사용하여 약 260,000에 트랜스펙션될 수 있다. Cas9 및 키메라 RNA의 RNA 전달이 도 28에 도시한다.

[1100] 시험관내 삽입결실부 형성 분석

[1101] 트랜스펙션 72 시간 후 세포가 수거되고, 이중가닥 브레이크의 표지로서 삽입결실부 형성에 대해 분석된다.

[1102] 간단히 말해서, 표적 서열 근처의 게놈 영역이 고 충실도 폴리머라제를 사용하여 PCR 증폭된다(약 400-600bp 앰플리콘 크기). 생성물이 정제되고, 동등한 농도로 표준화되고, 95℃에서 4℃까지 서서히 아닐링되며, 이로써 DNA 헤테로듀플렉스가 형성된다. 아닐링 후, Cel-I 효소를 사용하여 헤테로듀플렉스가 절단되고, 결과의 생성물이 폴리아크릴아미드 겔에서 분리되며, 삽입결실부 효율이 계산된다.

[1103] 동물에서 생체내 원리 입증

[1104] 전달 메커니즘

[1105] AAV 또는 렌티바이러스 제조는 다른 곳에 설명된다.

[1106] 나노입자 제제: RNA가 나노입자 제제에 혼합된다.

[1107] 마우스에 DNA 플라스미드의 유체역학적 꼬리 정맥 주사가 상업적 키트를 사용하여 수행된다.

[1108] Cas9 및 가이드 서열이 바이러스, 나노입자-코팅된 RNA 혼합물, 또는 DNA 플라스미드로서 전달되고, 대상체 동물에 주사된다. 병행하는 세트의 대조군 동물에는 멸균 식염수, Cas9 및 GFP, 또는 단독 가이드 서열 및 GFP가 주사된다.

[1109] 주사 3 주 후, 동물은 증상의 완화에 대해 시험되고 죽임을 당한다. 관련된 장기 시스템이 삽입결실부 형성에 대해 분석된다. 표현형 분석은 HDL, LDL, 지질의 혈중 수준을 포함한다.

[1110] 삽입결실부 형성 분석

[1111] DNA가 상업적 키트를 사용하여 조직으로부터 추출된다; 삽입결실부 분석은 시험관내 증명에 대해 설명된 대로 수행될 것이다.

[1112] CRISPR-Cas 시스템의 치료적 적용은 후보 질환 유전자의 조직-특이적이며 일시적으로 제어된 표적화된 결실을 달성하는 것을 잘 수행한다. 예들은 특히 콜레스테롤 및 지방산 대사, 아밀로이드 질환, 우성 음성 질환, 잠복 바이러스 감염에 수반되는 유전자를 포함한다.

[1113] 유전자 유전자좌에 표적화된 삽입결실부를 도입하기 위한 단일 가이드-RNA의 예들

[1114]

질환	유전자	스페이서	서열 번호 _)	PAM	메커니즘	참고자료
고콜레스테롤혈증	HMG-CR	GCCAAATTGGACGACCCTCG		CGG	녹아웃	Fluvastatin: a review of its pharmacology and use in the management of hypercholesterolaemia.(Plosker GL et al. Drugs 1996, 51(3):433-459)
고콜레스테롤혈증	SQLE	CGAGGAGACCCCCGTTTCGG		TGG	녹아웃	Potential role of nonstatin cholesterol lowering agents (Trapani et al. IUBMB Life, Volume 63, Issue 11, pages 964–971, November 2011)
고지혈증	DGAT1	CCCGCCGCCGCCGTGGCTCG		AGG	녹아웃	DGAT1 inhibitors as anti-obesity and anti-diabetic agents. (Birch AM et al. Current Opinion in Drug Discovery & Development [2010, 13(4):489-496)
백혈병	BCR-ABL	TGAGCTCTACGAGATCCACA		AGG	녹아웃	Killing of leukemic cells with a BCR/ABL fusion gene by RNA interference (RNAi).(Fuchs et al. Oncogene 2002, 21(37):5716-5724)

[1115] 유전자 유전자좌에 염색체 미소결실을 도입하기 위한 한 쌍의 가이드-RNA의 예들

질환	유전자	스페이서	서열 번호 _)	PAM	메커니즘	참고자료
고지혈증	PLIN2 가이드 1	CTCAAAATT CATACCGGT TG		TGG	미소결실	Perilipin-2 Null Mice are Protected Against Diet-Induced Obesity, Adipose Inflammation and Fatty Liver Disease (McManaman JL et al. The Journal of Lipid Research, jlr.M035063. First Published on February 12, 2013)
고지혈증	PLIN2 가이드 2	CGTTAAACA ACAACCGGA CT		TGG	미소결실	
고지혈증	SREBP 가이드 1	TTCACCCCG CGGCGCTGA AT		ggg	미소결실	Inhibition of SREBP by a Small Molecule, Betulin, Improves Hyperlipidemia and Insulin Resistance and Reduces Atherosclerotic Plaques (Tang J et al. Cell Metabolism, Volume 13, Issue 1, 44-56, 5 January 2011)
고지혈증	SREBP 가이드 2	ACCACTACC AGTCCGTCC AC		agg	미소결실	

[1116]

[1117] *실시예 19: 질환-원인 돌연변이를 지닌 유전자에 대한 수선의 표적화된 통합; 효소 결핍의 복구 및 다른 관련 질환*

[1118] 연구 설계

[1119] I. 유전자 표적의 확인 및 설계

[1120] · 실시예 22에 설명됨

[1121] II. 전달 시스템에 가이드 서열 및 수선 주형의 클로닝

[1122] · 실시예 22에 설명됨

[1123] · 출원인은 질환형 대립유전자를 가진 상동체 암과 야생형 수선 주형을 포함하도록 DNA 수선 주형을 클로닝한다.

[1124] *III.시험관내 세포주 검증*

[1125] a. 트랜스펙션은 상기 실시예 22에 설명된다; Cas9, 가이드 RNA, 및 수선 주형이 관련 세포 타입에 공-트랜스펙션된다.

[1126] b. 시험관내 수선 분석

[1127] i. 출원인은 트랜스펙션 72 시간 후 세포를 수거하고 수선에 대해 분석한다

[1128] ii. 간단히 말해서, 출원인은 고 충실도 폴리머라제 PCR을 사용하여 수선 주형 근처의 게놈 영역을 증폭한다. 서열 생성물은 돌연변이 대립유전자의 발생이 감소된다.

[1129] *IV.동물에서 생체내 원리 입증*

[1130] a. 전달 메커니즘은 상기 실시예 22 및 34에 설명된다.

[1131] b. 생체내 수선 분석

[1132] i. 출원인은 시험관내 증명세서 설명된 대로 수선 분석을 수행한다.

[1133] V. 치료적 적용

[1134] CRISPR-Cas 시스템은 후보 질환 유전자의 조직-특이적이며 일시적으로 제어된 표적화된 결실을 달성하는 것을 잘 수행한다. 예들은 특히 콜레스테롤 및 지방산 대사, 아밀로이드 질환, 우성 음성 질환, 잠복 바이러스 감염

에 수반되는 유전자를 포함한다.

[1135] 수선 주형을 가진 하나의 단일 미스센스 돌연변이의 예:

질환	유전자	스페이서	PAM
가족성 아밀로이드 다발신경장애	TTR	AGCCTTTCTGAACACATGCA	CGG

메커니즘 참고자료

V30M 수선 Transthyretin mutations in health and disease (Joao et al. Human Mutation, Volume 5, Issue 3, pages 191–196, 1995)

V30M CCTGCCATCAATGTGGCCATGCATGTGTTCAGAAAGGCT
대립유전자(서열 번호 ___)
WT CCTGCCATCAATGTGGCCGTGCATGTGTTCAGAAAGGCT
대립유전자(서열 번호 ___)

[1136]

[1137] *실시예 20: 녹내장, 아밀로이드증 및 헌팅턴병에서 CRISPR-Cas 시스템의 치료적 적용*

[1138] 녹내장: 출원인은 미실린 (MYOC) 유전자의 제1 엑손을 표적화하기 위한 가이드 RNA를 설계한다. 출원인은 Cas9와 MYOC 유전자를 표적화하는 가이드 RNA를 모두 패키지하기 위해 아데노바이러스 벡터(Ad5)를 사용한다. 출원인은 아데노바이러스 벡터를 녹내장의 병리생리학에서 세포가 이식된 섬유주대에 주사한다. 출원인은 이들이 시각의 선명함을 개선하고 안압을 감소시키는지의 여부를 알기 위해서 처음에 이것을 돌연변이된 MYOC 유전자를 지닌 마우스 모델에서 시험한다. 인간에서 치료적 적용도 유사한 전략을 이용한다.

[1139] 아밀로이드증: 출원인은 간에서 트랜스티레틴 (TTR) 유전자의 제1 엑손을 표적화하기 위한 가이드 RNA를 설계한다. 출원인은 Cas9와 TTR 유전자의 제1 엑손을 표적화하는 가이드 RNA를 패키지하기 위해 AAV8을 사용한다. AAV8는 간을 효과적으로 표적화하는 것으로 나타났으며, 정맥내 투여될 것이다. Cas9는 알부민 프로모터와 같은 간 특이적 프로모터를 사용하거나, 또는 구성적 프로모터를 사용하여 유도될 수 있다. pol3 프로모터는 가이드 RNA를 유도한다.

[1140] 또는 달리, 출원인은 TTR 유전자를 녹아웃시키기 위해 플라스미드 DNA의 유체역학적 전달을 이용한다. 출원인은 Cas9를 암호화하는 플라스미드와 TTR의 엑손1을 표적화하는 가이드 RNA를 전달한다.

[1141] 추가의 대안적 접근법으로서, 출원인은 RNA의 조합을 투여한다(Cas9의 경우 mRNA, 및 가이드 RNA). RNA는 Life Technologies로부터의 인비보펙타민을 사용하여 패키지되어 정맥내 전달될 수 있다. RNA-유도 면역원성을 감소시키고, Cas9 발현의 수준 및 가이드 RNA 안정성을 증가시키기 위해서, 출원인은 5' 봉쇄를 사용하여 Cas9 mRNA를 변형한다. 출원인은 또한 Cas9 mRNA 및 가이드 RNA에 변형된 RNA를 통합하며, 이로써 이들의 안정성이 증가하고 면역원성이 감소한다(예를 들어, TLR의 활성화). 효율을 증가시키기 위해서, 출원인은 바이러스, DNA 또는 RNA의 다수 용량을 투여한다.

[1142] 헌팅턴병: 출원인은 환자의 HTT 유전자에서 대립유전자 특이적 돌연변이에 기초하여 가이드 RNA를 설계한다. 예를 들어, 확장된 CAG 반복부를 가진 HTT에 대해 이형접합성인 환자에서, 출원인은 돌연변이 HTT 대립유전자에 특유한 뉴클레오티드 서열을 확인하며, 이것을 사용하여 가이드 RNA를 설계한다. 출원인은 돌연변이 염기가 가이드 RNA의 마지막 9 bp 안에 위치되는 것을 보장한다(출원인은 표적 크기와 가이드 RNA 사이에 단일 DNA 염기 미스매치를 구별할 수 있는 능력을 갖는 것을 확인했다).

[1143] 출원인은 돌연변이 HTT 대립유전자 특이적 가이드 RNA와 Cas9를 AAV9에 패키지해서 헌팅턴 환자의 선조체에 전달한다. 바이러스는 개두를 통해서 선조체에 주촉성 방식으로 주사된다. AAV9는 뉴런을 효과적으로 변환하는 것으로 알려져 있다. 출원인은 인간 시냅신 I과 같은 뉴런 특이적 프로모터를 사용하여 Cas9를 유도한다.

[1144] *실시예 21: HIV에서 CRISPR-Cas 시스템의 치료적 적용*

[1145] 만성 바이러스 감염은 유의한 사망률 및 유병율의 근원이다. 이들 바이러스 중 대부분에 대해 바이러스 복제의 다양한 양태를 효과적으로 표적화하는 종래의 항바이러스 요법이 존재하지만, 현재 치료 양식은 보통 "바이러스 잠복"으로 인해서 실제로는 비-치유적이다. 그 성질 때문에 바이러스 잠복은 활성 바이러스 생성 없는 바이러스 수명 주기의 휴면기로 특정된다. 이 기간 동안 바이러스는 면역 감시와 종래의 치료를 상당히 피할 수 있어서, 숙주 내에 장기적인 바이러스 저장소를 확립할 수 있고, 이로부터 후속 재활성화가 바이러스의 계속된 증식 및 전달을 허락할 수 있다. 바이러스 잠복에 대한 열쇠는 바이러스 게놈을 안정적으로 유지하는 능력인데, 이것은 세포질에 바이러스 게놈을 저장하거나 그것을 숙주 게놈에 통합하는 에피솜 또는 프로바이러스 잠복을 통해서

달성된다. 일차 감염을 방지하는 효과적인 백신접종이 없을 때는 잠복 저장소 및 세포용해 활성 사건에 의해서 특정된 만성 바이러스 감염이 유의한 결과를 가질 수 있다: 인간 유두종 바이러스(HPV)는 자궁경부암을 가져올 수 있고, C형 간염 바이러스(HCV)는 간세포 암종의 소인이며, 인간 면역결핍 바이러스는 궁극적으로 숙주 면역 시스템을 파괴하여 기회 감염에 대한 감수성을 가져온다. 이와 같이, 이들 감염은 현재 이용가능한 항바이러스 치료법의 일생동안의 사용을 필요로 한다. 더 나아가, 복잡한 문제로서 이들 바이러스 게놈의 대부분의 높은 변동성이 있는데, 이것은 효과적인 요법이 존재하지 않는 내성 균주의 발달을 초래한다.

[1146] CRISPR-Cas 시스템은 박테리아 후천 면역 시스템으로서, 복합적이며 서열-특이적 방식으로 이중가닥 DNA 브레이크 (DSB)를 유도할 수 있고, 최근에는 포유류 세포 시스템에서 재구성되었다. 하나 또는 다수의 가이드-RNA로 DNA를 표적화하는 것은 개재 서열의 삽입결실부와 결실을 각각 모두 가져올 수 있는 것으로 나타났다. 이와 같이, 이 신규 기술은 표적화된 복합적인 DNA 돌연변이유발이 높은 효율 및 특이성으로 단일 세포 내에서 달성될 수 있는 수단이다. 결과적으로, 바이러스 DNA 서열에 대해 지정된 CRISPR-Cas 시스템의 전달은 진행중인 바이러스 생성이 없을 때조차도 잠복 바이러스 게놈의 표적화된 파괴 및 결실을 허용할 수 있었다.

[1147] 예로서, HIV-1에 의한 만성적 감염은 전세계적 건강 문제로서, 3300만 명의 개체가 감염된 상태이며, 260만 건의 감염이 연간 발생한다. 바이러스 복제의 다수 양태를 동시에 표적화하는 다양식 고 활성 항레트로바이러스 요법 (HAART)의 사용은 만성적인 끝이 없는 질병으로서 HIV 감염이 상당히 관리될 수 있도록 했다. 치료하지 않으면 HIV의 AIDS로의 진행은 보통 9 내지 10 년 내에 일어나고, 결과적으로 숙주 면역 시스템이 고갈되며, 기회 감염이 발생하여 이후 곧 사망에 이르게 된다. 바이러스 잠복에 수반하여, HAART의 중단은 언제나 바이러스 복귀를 초래한다. 더욱이, 요법의 일시적 중단도 이용가능한 수단에 의해서 제어불가능한 HIV의 내성 균주에 대해 선택할 수 있다. 추가로, HAART 요법의 비용은 상당한데, 미국에서 년간 1 인당 $10,000 내지 15,0000가 소요된다. 이와 같이, 바이러스 복제 과정이 아니라 HIV 게놈을 직접 표적화하는 치료 접근법은 잠복 저장소의 박멸이 치유적 치료 옵션을 허용하는 수단이다.

[1148] HIV-1 표적화된 CRISPR-Cas 시스템의 개발 및 전달은 표적화된 DNA 돌연변이유발에 대한 기존 수단, 즉 ZFN 및 TALEN과는 차별화되는 독특한 접근법이며, 수많은 치료적 용도를 가진다. HAART와 함께 CRISPR-매개 DSB 및 삽입결실부에 의한 HIV-1 게놈의 표적화된 파괴 및 결실은 숙주에서 활성 바이러스 생성의 예방뿐만 아니라 잠복 바이러스 저장소의 고갈을 동시에 허용할 수 있었다.

[1149] 일단 숙주 면역 시스템에 통합되면, CRISPR-Cas 시스템은 완전한 바이러스 박멸의 없을 때조차도 HIV-1 내성 하위집단의 생성을 허용하고, 숙주 면역 활성의 유지와 재구성을 허용할 수 있다. 이것은 바이러스 게놈의 파괴에 의해 바이러스 생성 및 통합을 방지하여 1 차 감염을 방지하며, "백신접종"을 의미한다. CRISPR-Cas 시스템의 복합적 성질은 개별 세포에 동시에 게놈의 다수 양태를 표적화하는 것을 허용한다.

[1150] HAART에서와 같이, 돌연변이유발에 의한 바이러스 탈출은 다중 후천 돌연변이의 동시 획득을 요구함으로써 최소화된다. HIV-1의 다수의 균주가 동시에 표적화될 수 있으며, 이것은 수퍼감염의 기회를 최소화하고, 새로운 재조합 균주의 후속 생성을 방지한다. CRISPR-Cas 시스템의 단백질이 아니라 뉴클레오티드-매개된 서열 특이성은 전달 메커니즘의 유의한 변경의 필요 없이 치료제의 신속한 생성을 허용한다.

[1151] 이것을 달성하기 위해서, 출원인은 최대 커버리지 및 효율을 위해서 HIV-1 균주 변이체를 이용하여 거의 대부분의 HIV-1 게놈을 표적화하는 CRISPR-Cas 가이드 RNA를 생성한다. HIV-1 서브타입과 변이체 사이에 게놈 보존의 서열 분석은 개재된 바이러스 서열의 결실 또는 프레임-이동 돌연변이의 유도를 목표로 게놈의 측접한 보존된 영역의 표적화를 허용해야 하며, 이것은 바이러스 유전자 기능을 파괴할 것이다.

[1152] 출원인은 숙주 면역 시스템의 종래의 아데노바이러스 또는 렌티바이러스-매개 감염에 의해서 CRISPR-Cas 시스템의 전달을 달성한다. 접근법에 따라서, 숙주 면역 세포는 a) 분리되고, CRISPR-Cas로 형질도입되고, 선택되고, 숙주에 재도입되거나, 또는 b) CRISPR-Cas 시스템의 전신 전달에 의해서 생체내 형질도입될 수 있었다. 처음 접근법은 내성 면역 집단의 생성을 허용하고, 두 번째 접근법은 숙주 내에서 잠복 바이러스 저장소를 더 잘 표적화한다.

Mcintyre 등으로부터 개조된 잠재적 HIV-1 표적화된 스페이서의 예들 - 이들은 HIV-1 변이체의 최대 커버리지를 위해 최적화된 HIV-1에 대해 shRNA를 생성했다
CACTGCTTAAGCCTCGCTCGAGG (서열 번호 ___)
TCACCAGCAATATTCGCTCGAGG (서열 번호 ___)
CACCAGCAATATTCCGCTCGAGG (서열 번호 ___)
TAGCAACAGACATACGCTCGAGG (서열 번호 ___)
GGGCAGTAGTAATACGCTCGAGG (서열 번호 ___)
CCAATTCCCATACATTATTGTAC (서열 번호 ___)

[1153]

[1154] *실시예 22: 낭포성 섬유증에서 델타F508 또는 다른 돌연변이의 표적화된 보정*

[1155] 본 발명의 양태는 CRISPR-Cas 유전자 요법 입자와 생체적합성 제약학적 담체를 포함할 수 있는 제약 조성물을 제공한다. 다른 양태에 따라서, CFTR 유전자에 돌연변이를 가진 대상체의 치료를 위한 유전자 요법의 방법은 CRISPR-Cas 요법 입자의 치료적 유효량을 대상체의 세포에 투여하는 단계를 포함한다.

[1156] 이 실시예는 아데노-연관 바이러스(AAV) 입자를 사용한, 낭포성 섬유증 또는 낭포성 섬유증 관련 증상으로 고통받는, 필요한 대상체 또는 환자의 기도에 CRISPR-Cas 시스템의 유전자 전달 또는 유전자 전달을 증명한다.

[1157] 연구 설계: 필요한 대상체 또는 환자: 인간, 비-영장류 인간, 개, 고양이, 소, 말 및 다른 관련된 가축. 이 연구는 AAV 벡터에 의한 CRISPR-Cas 시스템의 유전자 전달의 효율을 시험한다. 출원인은 유전자 발현에 충분한 트랜스젠 수준을 결정하고, 델타508 또는 다른 CFTR-유도 돌연변이를 표적화하기 위해 Cas9 효소를 포함하는 CRISPR-Cas 시스템을 이용한다.

[1158] 관련된 대상체는 자발 호흡을 하는 동안 기관지내 전달된 폐에 대한 에어로졸화된 AAV 벡터 시스템의 제약학적 유효량을 수용한다. 대조군 대상체는 내부 대조군 유전자를 가진 거짓 타입 dAAV 벡터 시스템의 등량을 수용한다. 벡터 시스템은 제약학적으로 허용되는 또는 생체적합성 제약학적 담체와 함께 전달될 수 있다. 벡터 투여 후 3주 뒤에 또는 적절한 시간 간격으로, 치료된 대상체가 낭포성 섬유증 관련 증상의 완화에 대해 시험된다.

[1159] 출원인은 아데노바이러스 또는 AAV 입자를 사용한다.

[1160] 출원인은 하나 이상의 아데노바이러스 또는 AAV 벡터 또는 어떤 다른 양립성 벡터에, 하나 이상의 조절 서열(Cas9의 경우 Cbh 또는 EF1a 프로모터, 키메라 가이드 RNA의 경우 U6 또는 H1 프로모터)에 각각 작동 가능하게 연결된, 다음의 유전자 작제물을 클로닝한다: CFTR델타508 표적화 키메라 가이드 RNA (도 31B), 델타508 돌연변이를 위한 수선 주형 (도 31C) 및 선택적으로 하나 이상의 핵 위치 신호 또는 서열(들)을 가진 코돈 최적화된 Cas9 효소 (NLS(들)), 예를 들어 두(2) 개의 NLS).

[1161] Cas9 표적 부위의 확인

[1162] 출원인은 인간 CFTR 게놈 유전자좌를 분석했고, Cas9 표적 부위를 확인했다(도 31A). (PAM은 NGG 또는 NNAGAAW (서열 번호___) 모티프를 함유할 수 있다).

[1163] 유전자 수선 전략

[1164] 출원인은 앞서 논의된 전달 방법 중 하나를 통해서 대상체에 F508 잔기를 함유하는 상동체 수선 주형을 포함하는 아데노바이러스/AAV 벡터 시스템과 함께 Cas9 (또는 Cas9 닉카아제) 및 가이드 RNA를 포함하는 아데노바이러스/AAV 벡터 시스템을 도입한다. CRISPR-Cas 시스템은 CFTR델타508 키메라 가이드 RNA에 의해서 가이드되고, 흉터가 형성되거나 절단될 CFTR 게놈 유전자좌의 특정 부위를 표적화한다. 절단 후, 수선 주형이 낭포성 섬유증을 가져오거나 낭포성 섬유증 관련 증상을 야기하는 결실을 보정하는 상동성 재조합을 통해서 절단 부위에 삽입된다. 적절한 가이드 RNA에 의해 CRISPR 시스템의 전달을 지정하고 전신 도입을 제공하기 위한 이 전략은 표 B의 것들과 같은 대사성, 간, 신장 및 단백질 질환 및 장애를 야기하는 유전자를 편집하거나 조작하기 위해 유전자 돌연변이를 표적화하는데 이용한 수 있다.

[1165] *실시예 23: 유전자 녹아웃 세포 라이브러리 생성*

[1166] 이 실시예는 각 세포가 단일 유전자 녹아웃을 가진 세포들의 라이브러리를 생성하는 방법을 증명한다:

[1167] 출원인은 각 세포가 단일 유전자 녹아웃을 가진 ES 세포들의 라이브러리를 제작하며, 전체 ES 세포 라이브러리는 모든 단일 유전자 녹아웃을 가질 것이다. 이 라이브러리는 세포 과정뿐만 아니라 질환에서 유전자 기능의 스크리닝에 유용하다.

[1168] 이 세포 라이브러리를 제작하기 위해서, 출원인은 유도성 프로모터(예를 들어, 독시사이클린 유도성 프로모터)에 의해서 유도된 Cas9를 ES 세포에 통합한다. 게다가, 출원인은 ES 세포의 특정 유전자를 표적화하는 단일 가이드 RNA를 통합한다. ES 세포 라이브러리를 제작하기 위해서, 출원인은 ES 세포를 인간 게놈에서 각 유전자를 표적화하는 가이드 RNA를 암호화하는 유전자의 라이브러리와 간단히 혼합한다. 출원인은 먼저 단일 BxB1 attB 부위를 인간 ES 세포의 AAVS1 유전자좌에 도입한다. 다음에, 출원인은 BxB1 인테그라제를 사용하여 AAVS1 유전자좌에서 BxB1 attB 부위에 각 가이드 RNA 유전자의 통합을 촉진한다. 통합을 촉진하기 위해서, 각 가이드 RNA 유전자는 단일 attP 부위를 지닌 플라스미드 상에 함유된다. 이 방식에서, BxB1은 게놈의 attB 부위와 플라스미드를 함유하는 가이드 RNA 상의 attP 부위를 조합할 것이다.

[1169] 세포 라이브러리를 생성하기 위해서, 출원인은 통합된 단일 가이드 RNA를 가진 세포들의 라이브러리를 선택하여 Cas9 발현을 유도한다. 유도 후, Cas9는 가이드 RNA에 의해서 특정된 부위에 이중가닥 브레이크를 매개한다. 이 세포 라이브러리의 다양성을 검증하기 위해서, 출원인은 전체 엑솜 시퀀싱을 수행하며, 이로써 출원인이 모든 단일 표적화된 유전자에서 돌연변이를 관찰할 수 있도록 보장한다. 이 세포 라이브러리는 전체 라이브러리-기반 스크린을 포함하는 다양한 용도에 사용될 수 있거나, 또는 개별 인간 유전자 녹아웃을 가진 클론성 세포주의 신속한 생성을 촉진하기 위해서 개별 세포 클론으로 정렬할 수 있다.

[1170] 실시예 24: Cas9를 사용한 미세조류의 조작

[1171] Cas9의 전달 방법

[1172] 방법 1: 본 발명자들은 구성성 프로모터, 예를 들어, Hsp70A-Rbc S2 또는 베타2-투불린의 제어 하에 Cas9를 발현하는 벡터를 사용하여 Cas9 및 가이드 RNA를 전달한다.

[1173] 방법 2: 본 발명자들은 구성성 프로모터, 예를 들어, Hsp70A-Rbc S2 또는 베타2-투불린의 제어 하에 Cas9 및 T7 중합효소를 발현하는 벡터를 사용하여 Cas9 및 T7 중합효소를 전달한다. 가이드 RNA는 가이드 RNA를 유도하는 T7 프로모터를 함유하는 벡터를 사용하여 전달될 것이다.

[1174] 방법 3: 본 발명자들은 Cas9 mRNA 및 시험관내 전사된 가이드 RNA를 조류 세포로 전달한다. RNA는 시험관내 전사될 수 있다. Cas9 mRNA는 Cas9에 대한 코딩 영역, 및 Cop1으로부터의 3'UTR로 구성되어, Cas9 mRNA의 안정화를 보장할 것이다.

[1175] 상동성 재조합을 위하여, 본 발명자들은 추가의 상동성 유도 수복 주형을 제공한다.

[1176] 베타-2 투불린 프로모터에 이어서 Cop1의 3' UTR의 제어 하에 Cas9의 발현을 유도하는 카세트에 대한 서열

TCTTTCTTGCGCTATGACACTTCCAGCAAAAGGTAGGGCGGGCTGCGA
GACGGCTTCCCGGCGCTGCATGCAACACCGATGATGCTTCGACCCCCCGAAGCTC
CTTCGGGGCTGCATGGGCGCTCCGATGCCGCTCCAGGGCGAGCGCTGTTTAAATA
GCCAGGCCCCCGATTGCAAAGACATTATAGCGAGCTACCAAAGCCATATTCAAAC
ACCTAGATCACTACCACTTCTACACAGGCCACTCGAGCTTGTGATCGCACTCCGC
TAAGGGGGCGCCTCTTCCTCTTCGTTTCAGTCACAACCCGCAAACATGTACCCATA
CGATGTTCCAGATTACGCTTCGCCGAAGAAAAAGCGCAAGGTCGAAGCGTCCGA
CAAGAAGTACAGCATCGGCCTGGACATCGGCACCAACTCTGTGGGCTGGGCCGTG
ATCACCGACGAGTACAAGGTGCCCAGCAAGAAATTCAAGGTGCTGGGCAACACC
GACCGGCACAGCATCAAGAAGAACCTGATCGGAGCCCTGCTGTTCGACAGCGGC
GAAACAGCCGAGGCCACCCGGCTGAAGAGAACCGCCAGAAGAAGATACACCAG
ACGGAAGAACCGGATCTGCTATCTGCAAGAGATCTTCAGCAACGAGATGGCCAA
GGTGGACGACAGCTTCTTCCACAGACTGGAAGAGTCCTTCCTGGTGGAAGAGGAT
AAGAAGCACGAGCGGCACCCCATCTTCGGCAACATCGTGGACGAGGTGGCCTAC
CACGAGAAGTACCCCACCATCTACCACCTGAGAAAGAAACTGGTGGACAGCACC
GACAAGGCCGACCTGCGGCTGATCTATCTGGCCCTGGCCCACATGATCAAGTTCC
GGGGCCACTTCCTGATCGAGGGCGACCTGAACCCCGACAACAGCGACGTGGACA
AGCTGTTCATCCAGCTGGTGCAGACCTACAACCAGCTGTTCGAGGAAAACCCCAT
CAACGCCAGCGGCGTGGACGCCAAGGCCATCCTGTCTGCCAGACTGAGCAAGAG
CAGACGGCTGGAAAATCTGATCGCCCAGCTGCCCGGCGAGAAGAAGAATGGCCT
GTTCGGCAACCTGATTGCCCTGAGCCTGGGCCTGACCCCCAACTTCAAGAGCAAC
TTCGACCTGGCCGAGGATGCCAAACTGCAGCTGAGCAAGGACACCTACGACGAC
GACCTGGACAACCTGCTGGCCCAGATCGGCGACCAGTACGCCGACCTGTTTCTGG
CCGCCAAGAACCTGTCCGACGCCATCCTGCTGAGCGACATCCTGAGAGTGAACAC
CGAGATCACCAAGGCCCCCCTGAGCGCCTCTATGATCAAGAGATACGACGAGCA
CCACCAGGACCTGACCCTGCTGAAAGCTCTCGTGCGGCAGCAGCTGCCTGAGAAG
TACAAAGAGATTTTCTTCGACCAGAGCAAGAACGGCTACGCCGGCTACATTGACG
GCGGAGCCAGCCAGGAAGAGTTCTACAAGTTCATCAAGCCCATCCTGGAAAAGA
TGGACGGCACCGAGGAACTGCTCGTGAAGCTGAACAGAGAGGACCTGCTGCGGA
AGCAGCGGACCTTCGACAACGGCAGCATCCCCCACCAGATCCACCTGGGAGAGC
[1177] TGCACGCCATTCTGCGGCGGCAGGAAGATTTTTACCCATTCCTGAAGGACAACCG

[1178]

GGAAAAGATCGAGAAGATCCTGACCTTCCGCATCCCCTACTACGTGGGCCCTCTG
GCCAGGGGAAACAGCAGATTCGCCTGGATGACCAGAAAGAGCGAGGAAACCATC
ACCCCCTGGAACTTCGAGGAAGTGGTGGACAAGGGCGCTTCCGCCCAGAGCTTCA
TCGAGCGGATGACCAACTTCGATAAGAACCTGCCCAACGAGAAGGTGCTGCCCA
AGCACAGCCTGCTGTACGAGTACTTCACCGTGTATAACGAGCTGACCAAAGTGAA
ATACGTGACCGAGGGAATGAGAAAGCCCGCCTTCCTGAGCGGCGAGCAGAAAAA
GGCCATCGTGGACCTGCTGTTCAAGACCAACCGGAAAGTGACCGTGAAGCAGCT
GAAAGAGGACTACTTCAAGAAAATCGAGTGCTTCGACTCCGTGGAAATCTCCGGC
GTGGAAGATCGGTTCAACGCCTCCCTGGGCACATACCACGATCTGCTGAAAATTA
TCAAGGACAAGGACTTCCTGGACAATGAGGAAAACGAGGACATTCTGGAAGATA
TCGTGCTGACCCTGACACTGTTTGAGGACAGAGAGATGATCGAGGAACGGCTGA
AAACCTATGCCCACCTGTTCGACGACAAAGTGATGAAGCAGCTGAAGCGGCGGA
GATACACCGGCTGGGGCAGGCTGAGCCGGAAGCTGATCAACGGCATCCGGGACA
AGCAGTCCGGCAAGACAATCCTGGATTTCCTGAAGTCCGACGGCTTCGCCAACAG
AAACTTCATGCAGCTGATCCACGACGACAGCCTGACCTTTAAAGAGGACATCCAG
AAAGCCCAGGTGTCCGGCCAGGGCGATAGCCTGCACGAGCACATTGCCAATCTG
GCCGGCAGCCCCGCCATTAAGAAGGGCATCCTGCAGACAGTGAAGGTGGTGGAC
GAGCTCGTGAAAGTGATGGGCCGGCACAAGCCCGAGAACATCGTGATCGAAATG
GCCAGAGAGAACCAGACCACCCAGAAGGGACAGAAGAACAGCCGCGAGAGAAT
GAAGCGGATCGAAGAGGGCATCAAAGAGCTGGGCAGCCAGATCCTGAAAGAACA
CCCCGTGGAAAACACCCAGCTGCAGAACGAGAAGCTGTACCTGTACTACCTGCAG
AATGGGCGGGATATGTACGTGGACCAGGAACTGGACATCAACCGGCTGTCCGAC
TACGATGTGGACCATATCGTGCCTCAGAGCTTTCTGAAGGACGACTCCATCGACA
ACAAGGTGCTGACCAGAAGCGACAAGAACCGGGGCAAGAGCGACAACGTGCCCT
CCGAAGAGGTCGTGAAGAAGATGAAGAACTACTGGCGGCAGCTGCTGAACGCCA
AGCTGATTACCCAGAGAAAGTTCGACAATCTGACCAAGGCCGAGAGAGGCGGCC
TGAGCGAACTGGATAAGGCCGGCTTCATCAAGAGACAGCTGGTGGAAACCCGGC
AGATCACAAAGCACGTGGCACAGATCCTGGACTCCCGGATGAACACTAAGTACG
ACGAGAATGACAAGCTGATCCGGGAAGTGAAAGTGATCACCCTGAAGTCCAAGC
TGGTGTCCGATTTCCGGAAGGATTTCCAGTTTTACAAAGTGCGCGAGATCAACAA
CTACCACCACGCCCACGACGCCTACCTGAACGCCGTCGTGGGAACCGCCCTGATC

AAAAAGTACCCTAAGCTGGAAAGCGAGTTCGTGTACGGCGACTACAAGGTGTAC
GACGTGCGGAAGATGATCGCCAAGAGCGAGCAGGAAATCGGCAAGGCTACCGCC
AAGTACTTCTTCTACAGCAACATCATGAACTTTTTCAAGACCGAGATTACCCTGG
CCAACGGCGAGATCCGGAAGCGGCCTCTGATCGAGACAAACGGCGAAACCGGGG
AGATCGTGTGGGATAAGGGCCGGGATTTTGCCACCGTGCGGAAAGTGCTGAGCA
TGCCCCAAGTGAATATCGTGAAAAAGACCGAGGTGCAGACAGGCGGCTTCAGCA
AAGAGTCTATCCTGCCCAAGAGGAACAGCGATAAGCTGATCGCCAGAAAGAAGG
ACTGGGACCCTAAGAAGTACGGCGGCTTCGACAGCCCCACCGTGGCCTATTCTGT
GCTGGTGGTGGCCAAAGTGGAAAAGGGCAAGTCCAAGAAACTGAAGAGTGTGAA
AGAGCTGCTGGGGATCACCATCATGGAAAGAAGCAGCTTCGAGAAGAATCCCAT
CGACTTTCTGGAAGCCAAGGGCTACAAAGAAGTGAAAAAGGACCTGATCATCAA
GCTGCCTAAGTACTCCCTGTTCGAGCTGGAAAACGGCCGGAAGAGAATGCTGGC
CTCTGCCGGCGAACTGCAGAAGGGAAACGAACTGGCCCTGCCCTCCAAATATGT
GAACTTCCTGTACCTGGCCAGCCACTATGAGAAGCTGAAGGGCTCCCCCGAGGAT
AATGAGCAGAAACAGCTGTTTGTGGAACAGCACAAGCACTACCTGGACGAGATC
ATCGAGCAGATCAGCGAGTTCTCCAAGAGAGTGATCCTGGCCGACGCTAATCTGG
ACAAAGTGCTGTCCGCCTACAACAAGCACCGGGATAAGCCCATCAGAGAGCAGG
CCGAGAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCCCCTGCCGCCTT
CAAGTACTTTGACACCACCATCGACCGGAAGAGGTACACCAGCACCAAAGAGGT
GCTGGACGCCACCCTGATCCACCAGAGCATCACCGGCCTGTACGAGACACGGAT
CGACCTGTCTCAGCTGGGAGGCGACAGCCCCAAGAAGAAGAGAAAGGTGGAGGC
CAGCTAAGGATCCGGCAAGACTGGCCCCGCTTGGCAACGCAACAGTGAGCCCCTC
CCTAGTGTGTTTGGGGATGTGACTATGTATTCGTGTGTTGGCCAACGGGTCAACC
CGAACAGATTGATACCCGCCTTGGCATTTCCTGTCAGAATGTAACGTCAGTTGAT

[1179] GGTACT (서열 번호 __)

[1180] 베타-2 투불린 프로모터에 이어서 Cop1의 3' UTR의 제어 하에 T7 중합효소의 발현을 유도하는 카세트에 대한 서열:

TCTTTCTTGCGCTATGACACTTCCAGCAAAAGGTAGGGCGGGCTGCGA
GACGGCTTCCCGGCGCTGCATGCAACACCGATGATGCTTCGACCCCCCGAAGCTC
CTTCGGGGCTGCATGGGCGCTCCGATGCCGCTCCAGGGCGAGCGCTGTTTAAATA
GCCAGGCCCCCCGATTGCAAAGACATTATAGCGAGCTACCAAAGCCATATTCAAAC
ACCTAGATCACTACCACTTCTACACAGGCCACTCGAGCTTGTGATCGCACTCCGC
TAAGGGGGCGCCTCTTCCTCTTCGTTTCAGTCACAACCCGCAAACatgcctaagaagaaga
ggaaggttaacacgattaacatcgctaagaacgacttctctgacatcgaactggctgctatcccgttcaacactctggctgaccatt
acggtgagcgtttagctcgcgaacagttggcccttgagcatgagtcttacgagatgggtgaagcacgcttccgcaagatgtttga
gcgtcaacttaaagctggtgaggttgcggataacgctgccgccaagcctctcatcactaccctactcccctaagatgattgcacgc
atcaacgactggtttgaggaagtgaaagctaagcgcggcaagcgcccgacagccttccagttcctgcaagaaatcaagccgga
agccgtagcgtacatcaccattaagaccactctggcttgcctaaccagtgctgacaatacaaccgttcaggctgtagcaagcgca
atcggtcgggccattgaggacgaggctcgcttcggtcgtatccgtgaccttgaagctaagcacttcaagaaaaacgttgaggaa
caactcaacaagcgcgtagggcacgtctacaagaaagcatttatgcaagttgtcgaggctgacatgctctctaagggtctactcg
gtggcgaggcgtggtcttcgtggcataaggaagactctattcatgtaggagtacgctgcatcgagatgctcattgagtcaaccgg
aatggttagcttacaccgccaaaatgctggcgtagtaggtcaagactctgagactatcgaactcgcacctgaatacgctgaggct
atcgcaacccgtgcaggtgcgctggctggcatctctccgatgttccaaccttgcgtagttcctcctaagccgtggactggcattac
tggtggtggctattgggctaacggtcgtcgtcctctggcgctggtgcgtactcacagtaagaaagcactgatgcgctacgaagac
gtttacatgcctgaggtgtacaaagcgattaacattgcgcaaaacaccgcatggaaaatcaacaagaaagtcctagcggtcgcc
aacgtaatcaccaagtggaagcattgtccggtcgaggacatccctgcgattgagcgtgaagaactcccgatgaaaccggaaga
catcgacatgaatcctgaggctctcaccgcgtggaaacgtgctgccgctgctgtgtaccgcaaggacaaggctcgcaagtctcg
ccgtatcagccttgagttcatgcttgagcaagccaataagtttgctaaccataaggccatctggttcccttacaacatggactggcg
cggtcgtgtttacgctgtgtcaatgttcaacccgcaaggtaacgatatgaccaaaggactgcttacgctggcgaaaggtaaacca
atcggtaaggaaggttactactggctgaaaatccacggtgcaaactgtgcgggtgtcgacaaggttccgttccctgagcgcatca
agttcattgaggaaaaccacgagaacatcatggcttgcgctaagtctccactggagaacacttggtgggctgagcaagattctcc
gttctgcttccttgcgttctgctttgagtacgctggggtacagcaccacggcctgagctataactgctcccttccgctggcgtttgac
gggtcttgctctggcatccagcacttctccgcgatgctccgagatgaggtaggtggtcgcgcggttaacttgcttcctagtgaaac
cgttcaggacatctacgggattgttgctaagaaagtcaacgagattctacaagcagacgcaatcaatgggaccgataacgaagta
gttaccgtgaccgatgagaacactggtgaaatctctgagaaagtcaagctgggcactaaggcactggctggtcaatggctggct
tacggtgttactcgcagtgtgactaagcgttcagtcatgacgctggcttacgggtccaaagagttcggcttccgtcaacaagtgct
ggaagataccattcagccagctattgattccggcaagggtctgatgttcactcagccgaatcaggctgctggatacatggctaag
ctgatttgggaatctgtgagcgtgacggtggtagctgcggttgaagcaatgaactggcttaagtctgctgctaagctgctggctgc

[1181]

tgaggtcaaagataagaagactggagagattcttcgcaagcgttgcgctgtgcattgggtaactcctgatggtttccctgtgtggc
aggaatacaagaagcctattcagacgcgcttgaacctgatgttcctcggtcagttccgcttacagcctaccattaacaccaacaa
agatagcgagattgatgcacacaaacaggagtctggtatcgctcctaactttgtacacagccaagacggtagccaccttcgtaag
actgtagtgtgggcacacgagaagtacggaatcgaatctttgcactgattcacgactccttcggtacgattccggctgacgctgc
gaacctgttcaaagcagtgcgcgaaactatggttgacacatatgagtcttgtgatgtactggctgatttctacgaccagttcgctga
ccagttgcacgagtctcaattggacaaaatgccagcacttccggctaaaggtaacttgaacctccgtgacatcttagagtcggact
tcgcgttcgcgtaaGGATCCGGCAAGACTGGCCCCGCTTGGCAACGCAACAGTGAGCCC
CTCCCTAGTGTGTTTGGGGATGTGACTATGTATTCGTGTGTTGGCCAACGGGTCA
ACCCGAACAGATTGATACCCGCCTTGGCATTTCCTGTCAGAATGTAACGTCAGTT
GATGGTACT (서열 번호 ___)

[1182]

[1183] T7 프로모터(T7 프로모터, Ns는 표적 부위를 나타냄)에 의해 유도되는 가이드 RNA의 서열:

gaaat**TAATACGACTCACTATA**<u>NNNNNNNNNNNNNNNNNNNN</u>gttttagagc
taGAAAtagcaagttaaaataaggctagtccgttatcaacttgaaaaagtggcaccgagtcggtgcttttttt(서열 번호 __)

[1184]

[1185] 유전자 전달:

[1186] 클라미도모나스(Chlamydomonas) 자원 센터로부터의 클라미도모나스 라인하르티이(Chlamydomonas reinhardtii) 균주 CC-124 및 CC-125는 전기천공법을 위해 사용될 것이다. 전기천공법 프로토콜은 진아트(GeneArt) 클라미도

모나스 조작 키트로부터의 표준 권고 프로토콜을 따른다.

[1187] 또한, 본 발명자들은 Cas9를 구성적으로 발현하는 클로미도모나스 라인하르티이의 세포주를 생성한다. 이는 pChlamy1(PvuI을 사용하여 선형화)을 사용하고, 하이그로마이신 내성 콜로니를 선택하여 행해질 수 있다. Cas9를 함유하는 pChlamy1에 대한 서열은 하기에 나타나 있다. 이러한 방식으로 유전자 낙아웃을 달성하기 위하여, 간단하게 가이드 RNA를 위해 RNA를 전달할 필요가 있다. 상동성 재조합을 위하여, 본 발명자들은 가이드 RNA 및 선형화된 상동성 재조합 주형을 전달한다.

[1188] pChlamy1-Cas9:

TGCGGTATTTCACACCGCATCAGGTGGCACTTTTCGGGGAAATGTGC
GCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGA
GATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAA
ATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATC
AGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACT
CCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCT
GCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACC
AGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCAT
CCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGT
TTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGG
TATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCC
ATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTA
AGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACT
GTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATT
CTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGAT
AATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTT
CGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACC
CACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGT
GAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGG
AAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGG
TTATTGTCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAG
ACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATC
TGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATC
AAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACC
AAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTA
GCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGTTGCCAGTGG
CGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCG
CAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACG
ACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTC
CCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGA

[1189]

[1190]

GAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCG
GGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGG
AGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTG
GCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTA
TTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAG
CGAGTCAGTGAGCGAGGAAGCGGTCGCTGAGGCTTGACATGATTGGTGCGTATGT
TTGTATGAAGCTACAGGACTGATTTGGCGGGCTATGAGGGCGGGGGAAGCTCTGG
AAGGGCCGCGATGGGGCGCGCGGCGTCCAGAAGGCGCCATACGGCCCGCTGGCG
GCACCCATCCGGTATAAAAGCCCGCGACCCCGAACGGTGACCTCCACTTTCAGCG
ACAAACGAGCACTTATACATACGCGACTATTCTGCCGCTATACATAACCACTCAG
CTAGCTTAAGATCCCATCAAGCTTGCATGCCGGGCGCGCCAGAAGGAGCGCAGC
CAAACCAGGATGATGTTTGATGGGGTATTTGAGCACTTGCAACCCTTATCCGGAA
GCCCCCTGGCCCACAAAGGCTAGGCGCCAATGCAAGCAGTTCGCATGCAGCCCCT
GGAGCGGTGCCCTCCTGATAAACCGGCCAGGGGGCCTATGTTCTTTACTTTTTTAC
AAGAGAAGTCACTCAACATCTTAAAATGGCCAGGTGAGTCGACGAGCAAGCCCG
GCGGATCAGGCAGCGTGCTTGCAGATTTGACTTGCAACGCCCGCATTGTGTCGAC
GAAGGCTTTTGGCTCCTCTGTCGCTGTCTCAAGCAGCATCTAACCCTGCGTCGCCG
TTTCCATTTGCAGGAGATTCGAGGTACCATGTACCCATACGATGTTCCAGATTACG
CTTCGCCGAAGAAAAAGCGCAAGGTCGAAGCGTCCGACAAGAAGTACAGCATCG
GCCTGGACATCGGCACCAACTCTGTGGGCTGGGCCGTGATCACCGACGAGTACAA
GGTGCCCAGCAAGAAATTCAAGGTGCTGGGCAACACCGACCGGCACAGCATCAA
GAAGAACCTGATCGGAGCCCTGCTGTTCGACAGCGGCGAAACAGCCGAGGCCAC
CCGGCTGAAGAGAACCGCCAGAAGAAGATACACCAGACGGAAGAACCGGATCTG
CTATCTGCAAGAGATCTTCAGCAACGAGATGGCCAAGGTGGACGACAGCTTCTTC
CACAGACTGGAAGAGTCCTTCCTGGTGGAAGAGGATAAGAAGCACGAGCGGCAC
CCCATCTTCGGCAACATCGTGGACGAGGTGGCCTACCACGAGAAGTACCCCACCA
TCTACCACCTGAGAAAGAAACTGGTGGACAGCACCGACAAGGCCGACCTGCGGC
TGATCTATCTGGCCCTGGCCCACATGATCAAGTTCCGGGGCCACTTCCTGATCGA
GGGCGACCTGAACCCCGACAACAGCGACGTGGACAAGCTGTTCATCCAGCTGGT
GCAGACCTACAACCAGCTGTTCGAGGAAAACCCCATCAACGCCAGCGGCGTGGA
CGCCAAGGCCATCCTGTCTGCCAGACTGAGCAAGAGCAGACGGCTGGAAAATCT

[1191]

GATCGCCCAGCTGCCCGGCGAGAAGAAGAATGGCCTGTTCGGCAACCTGATTGC
CCTGAGCCTGGGCCTGACCCCCAACTTCAAGAGCAACTTCGACCTGGCCGAGGAT
GCCAAACTGCAGCTGAGCAAGGACACCTACGACGACGACCTGGACAACCTGCTG
GCCCAGATCGGCGACCAGTACGCCGACCTGTTTCTGGCCGCCAAGAACCTGTCCG
ACGCCATCCTGCTGAGCGACATCCTGAGAGTGAACACCGAGATCACCAAGGCCCC
CCTGAGCGCCTCTATGATCAAGAGATACGACGAGCACCACCAGGACCTGACCCTG
CTGAAAGCTCTCGTGCGGCAGCAGCTGCCTGAGAAGTACAAAGAGATTTTCTTCG
ACCAGAGCAAGAACGGCTACGCCGGCTACATTGACGGCGGAGCCAGCCAGGAAG
AGTTCTACAAGTTCATCAAGCCCATCCTGGAAAAGATGGACGGCACCGAGGAAC
TGCTCGTGAAGCTGAACAGAGAGGACCTGCTGCGGAAGCAGCGGACCTTCGACA
ACGGCAGCATCCCCCACCAGATCCACCTGGGAGAGCTGCACGCCATTCTGCGGCG
GCAGGAAGATTTTTACCCATTCCTGAAGGACAACCGGGAAAAGATCGAGAAGAT
CCTGACCTTCCGCATCCCCTACTACGTGGGCCCTCTGGCCAGGGGAAACAGCAGA
TTCGCCTGGATGACCAGAAAGAGCGAGGAAACCATCACCCCCTGGAACTTCGAG
GAAGTGGTGGACAAGGGCGCTTCCGCCCAGAGCTTCATCGAGCGGATGACCAAC
TTCGATAAGAACCTGCCCAACGAGAAGGTGCTGCCCAAGCACAGCCTGCTGTAC
GAGTACTTCACCGTGTATAACGAGCTGACCAAAGTGAAATACGTGACCGAGGGA
ATGAGAAAGCCCGCCTTCCTGAGCGGCGAGCAGAAAAAGGCCATCGTGGACCT
GCTGTTCAAGACCAACCGGAAAGTGACCGTGAAGCAGCTGAAAGAGGACTACTT
CAAGAAAATCGAGTGCTTCGACTCCGTGGAAATCTCCGGCGTGGAAGATCGGTTC
AACGCCTCCCTGGGCACATACCACGATCTGCTGAAAATTATCAAGGACAAGGACT
TCCTGGACAATGAGGAAAACGAGGACATTCTGGAAGATATCGTGCTGACCCTGA
CACTGTTTGAGGACAGAGAGATGATCGAGGAACGGCTGAAAACCTATGCCCACC
TGTTCGACGACAAAGTGATGAAGCAGCTGAAGCGGCGGAGATACACCGGCTGGG
GCAGGCTGAGCCGGAAGCTGATCAACGGCATCCGGGACAAGCAGTCCGGCAAGA
CAATCCTGGATTTCCTGAAGTCCGACGGCTTCGCCAACAGAAACTTCATGCAGCT
GATCCACGACGACAGCCTGACCTTTAAAGAGGACATCCAGAAAGCCCAGGTGTC
CGGCCAGGGCGATAGCCTGCACGAGCACATTGCCAATCTGGCCGGCAGCCCCGC
CATTAAGAAGGGCATCCTGCAGACAGTGAAGGTGGTGGACGAGCTCGTGAAAGT
GATGGGCCGGCACAAGCCCGAGAACATCGTGATCGAAATGGCCAGAGAGAACCA
GACCACCCAGAAGGGACAGAAGAACAGCCGCGAGAGAATGAAGCGGATCGAAG

[1192]

AGGGCATCAAAGAGCTGGGCAGCCAGATCCTGAAAGAACACCCCGTGGAAAACA
CCCAGCTGCAGAACGAGAAGCTGTACCTGTACTACCTGCAGAATGGGCGGGATA
TGTACGTGGACCAGGAACTGGACATCAACCGGCTGTCCGACTACGATGTGGACC
ATATCGTGCCTCAGAGCTTTCTGAAGGACGACTCCATCGACAACAAGGTGCTGAC
CAGAAGCGACAAGAACCGGGGCAAGAGCGACAACGTGCCCTCCGAAGAGGTCG
TGAAGAAGATGAAGAACTACTGGCGGCAGCTGCTGAACGCCAAGCTGATTACCC
AGAGAAAGTTCGACAATCTGACCAAGGCCGAGAGAGGCGGCCTGAGCGAACTGG
ATAAGGCCGGCTTCATCAAGAGACAGCTGGTGGAAACCCGGCAGATCACAAAGC
ACGTGGCACAGATCCTGGACTCCCGGATGAACACTAAGTACGACGAGAATGACA
AGCTGATCCGGGAAGTGAAAGTGATCACCCTGAAGTCCAAGCTGGTGTCCGATTT
CCGGAAGGATTTCCAGTTTTACAAAGTGCGCGAGATCAACAACTACCACCACGCC
CACGACGCCTACCTGAACGCCGTCGTGGGAACCGCCCTGATCAAAAAGTACCCTA
AGCTGGAAAGCGAGTTCGTGTACGGCGACTACAAGGTGTACGACGTGCGGAAGA
TGATCGCCAAGAGCGAGCAGGAAATCGGCAAGGCTACCGCCAAGTACTTCTTCTA
CAGCAACATCATGAACTTTTTCAAGACCGAGATTACCCTGGCCAACGGCGAGATC
CGGAAGCGGCCTCTGATCGAGACAAACGGCGAAACCGGGGAGATCGTGTGGGAT
AAGGGCCGGGATTTTGCCACCGTGCGGAAAGTGCTGAGCATGCCCCAAGTGAAT
ATCGTGAAAAAGACCGAGGTGCAGACAGGCGGCTTCAGCAAAGAGTCTATCCTG
CCCAAGAGGAACAGCGATAAGCTGATCGCCAGAAAGAAGGACTGGGACCCTAAG
AAGTACGGCGGCTTCGACAGCCCCACCGTGGCCTATTCTGTGCTGGTGGTGGCCA
AAGTGGAAAAGGGCAAGTCCAAGAAACTGAAGAGTGTGAAAGAGCTGCTGGGG
ATCACCATCATGGAAAGAAGCAGCTTCGAGAAGAATCCCATCGACTTTCTGGAAG
CCAAGGGCTACAAAGAAGTGAAAAAGGACCTGATCATCAAGCTGCCTAAGTACT
CCCTGTTCGAGCTGGAAAACGGCCGGAAGAGAATGCTGGCCTCTGCCGGCGAAC
TGCAGAAGGGAAACGAACTGGCCCTGCCCTCCAAATATGTGAACTTCCTGTACC
TGGCCAGCCACTATGAGAAGCTGAAGGGCTCCCCCGAGGATAATGAGCAGAAAC
AGCTGTTTGTGGAACAGCACAAGCACTACCTGGACGAGATCATCGAGCAGATCA
GCGAGTTCTCCAAGAGAGTGATCCTGGCCGACGCTAATCTGGACAAAGTGCTGTC
CGCCTACAACAAGCACCGGGATAAGCCCATCAGAGAGCAGGCCGAGAATATCAT
CCACCTGTTTACCCTGACCAATCTGGGAGCCCCTGCCGCCTTCAAGTACTTTGACA
CCACCATCGACCGGAAGAGGTACACCAGCACCAAAGAGGTGCTGGACGCCACCC

TGATCCACCAGAGCATCACCGGCCTGTACGAGACACGGATCGACCTGTCTCAGCT
GGGAGGCGACAGCCCCAAGAAGAAGAGAAAGGTGGAGGCCAGCTAACATATGATT
CGAATGTCTTTCTTGCGCTATGACACTTCCAGCAAAAGGTAGGGCG
GGCTGCGAGACGGCTTCCCGGCGCTGCATGCAACACCGATGATGCTTCGACCCCC
CGAAGCTCCTTCGGGGCTGCATGGGCGCTCCGATGCCGCTCCAGGGCGAGCGCTG
TTTAAATAGCCAGGCCCCCGATTGCAAAGACATTATAGCGAGCTACCAAAGCCAT
ATTCAAACACCTAGATCACTACCACTTCTACACAGGCCACTCGAGCTTGTGATCG
CACTCCGCTAAGGGGGCGCCTCTTCCTCTTCGTTTCAGTCACAACCCGCAAACAT
GACACAAGAATCCCTGTTACTTCTCGACCGTATTGATTCGGATGATTCCTACGCG
AGCCTGCGGAACGACCAGGAATTCTGGGAGGTGAGTCGACGAGCAAGCCCGGCG
GATCAGGCAGCGTGCTTGCAGATTTGACTTGCAACGCCCGCATTGTGTCGACGAA
GGCTTTTGGCTCCTCTGTCGCTGTCTCAAGCAGCATCTAACCCTGCGTCGCCGTTT
CCATTTGCAGCCGCTGGCCCGCCGAGCCCTGGAGGAGCTCGGGCTGCCGGTGCCG
CCGGTGCTGCGGGTGCCCGGCGAGAGCACCAACCCCGTACTGGTCGGCGAGCCC
GGCCCGGTGATCAAGCTGTTCGGCGAGCACTGGTGCGGTCCGGAGAGCCTCGCG
TCGGAGTCGGAGGCGTACGCGGTCCTGGCGGACGCCCCGGTGCCGGTGCCCCGC
CTCCTCGGCCGCGGCGAGCTGCGGCCCGGCACCGGAGCCTGGCCGTGGCCCTACC
TGGTGATGAGCCGGATGACCGGCACCACCTGGCGGTCCGCGATGGACGGCACGA
CCGACCGGAACGCGCTGCTCGCCCTGGCCCGCGAACTCGGCCGGGTGCTCGGCCG
GCTGCACAGGGTGCCGCTGACCGGGAACACCGTGCTCACCCCCCATTCCGAGGTC
TTCCCGGAACTGCTGCGGGAACGCCGCGCGGCGACCGTCGAGGACCACCGCGGG
TGGGGCTACCTCTCGCCCCGGCTGCTGGACCGCCTGGAGGACTGGCTGCCGGACG
TGGACACGCTGCTGGCCGGCCGCGAACCCCGGTTCGTCCACGGCGACCTGCACGG
GACCAACATCTTCGTGGACCTGGCCGCGACCGAGGTCACCGGGATCGTCGACTTC
ACCGACGTCTATGCGGGAGACTCCCGCTACAGCCTGGTGCAACTGCATCTCAACG
CCTTCCGGGGCGACCGCGAGATCCTGGCCGCGCTGCTCGACGGGGCGCAGTGGA
AGCGGACCGAGGACTTCGCCCGCGAACTGCTCGCCTTCACCTTCCTGCACGACTT
CGAGGTGTTCGAGGAGACCCCGCTGGATCTCTCCGGCTTCACCGATCCGGAGGAA
CTGGCGCAGTTCCTCTGGGGGCCGCCGGACACCGCCCCCGGCGCCTGATAAGGAT
CCGGCAAGACTGGCCCCGCTTGGCAACGCAACAGTGAGCCCCTCCCTAGTGTGTT
TGGGGATGTGACTATGTATTCGTGTGTTGGCCAACGGGTCAACCCGAACAGATTG

[1193]

ATACCCGCCTTGGCATTTCCTGTCAGAATGTAACGTCAGTTGATGGTACT

[1194] (서열 번호 ___)

[1195] 모든 변형된 클로미도모나스 라인하르티이 세포에 있어서, 본 발명자들은 PCR, 서베이어 뉴클레아제 검정 및 DNA 시퀀싱을 사용하여, 성공적인 변형을 확인하였다.

[1196] *실시예 25: 다양한 질환 타입을 표적화하기 위한 Cas9의 사용*

[1197] 단백질 암호화 서열에 돌연변이를 포함하는 질환:

[1198] 우성 질환은 우성 음성 대립유전자를 비활성화함으로써 표적화될 수 있다. 출원인은 우성 음성 대립유전자에서 특유의 서열을 표적화하고 NHEJ를 통해 돌연변이를 도입하기 위해서 Cas9를 사용한다. NHEJ-유도된 삽입결실부는 우성 음성 대립유전자에 프레임-이동 돌연변이를 도입하고, 우성 음성 단백질을 제거할 수 있다. 이것은 유전자가 하프로-충분인 경우 수행할 수 있다(예를 들어, MYOC 돌연변이 유도된 녹내장 및 헌팅턴병).

[1199] 열성 장애는 두 대립유전자에서 질환 돌연변이를 수선함으로써 표적화될 수 있다. 분열하는 세포의 경우, 출원인은 돌연변이 부위 근처에 이중가닥 브레이크를 도입하기 위해 Cas9를 사용하고, 외인성 재조합 주형을 사용하여 상동성 재조합의 비율을 증가시킨다. 분열하는 세포의 경우, 이것은 복합적인 닉카아제 활성을 사용하여 달성될 수 있으며, 이로써 상보성 돌출부를 지닌 외인성 DNA 단편의 NHEJ-매개 결찰을 통해서 두 대립유전자에서 돌연변이 서열의 치환을 촉매할 수 있다.

[1200] 또한, 출원인은 보호 돌연변이를 도입하기 위해 Cas9를 사용한다(예를 들어, HIV 감염을 방지하기 위한 CCR5의

비활성화, 콜레스테롤 감소를 위한 PCSK9의 비활성화, 또는 알츠하이머병의 가능성을 감소시키기 위한 APP에 A673T의 도입).

[1201] **비-암호화 서열을 수반하는 질환**

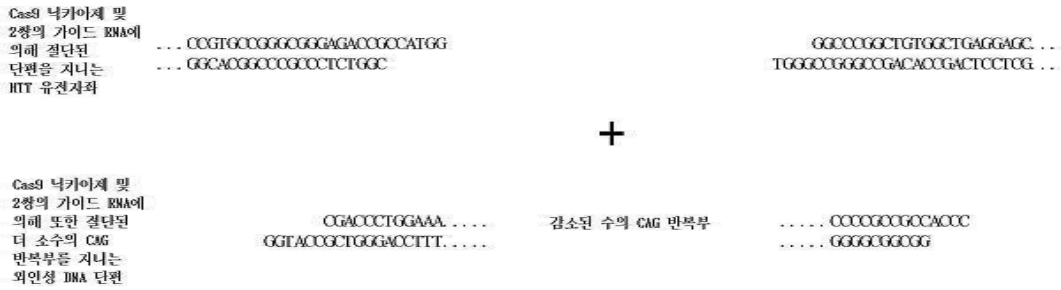
[1202] 출원인은 프로모터 영역에서 비-암호화 서열을 파괴하고, 전사 인자 결합 부위를 변경하고, 인핸서 또는 억제인자 요소를 변경하기 위해서 Cas9를 사용한다. 예를 들어, Cas9는 조혈 줄기세포에서 Klf1 인핸서 EHS1을 삭제하기 위해 사용될 수 있으며, 이로써 BCL11a 수준이 감소하고, 분화된 적혈구에서 글로빈 유전자 발현을 재활성화할 수 있다.

[1203] 출원인은 또한 5' 또는 3' 미번역 영역에서 기능적 모티프를 파괴하기 위해서 Cas9를 사용한다. 예를 들어, 근긴장 디스트로피의 치료를 위해, Cas9는 DMPK 유전자에서 CTG 반복부 확장을 제거하기 위해서 사용될 수 있다.

[1204] *실시예 26: 복합화된 닉카아제*

[1205] 본 출원에서 상세히 설명된 Cas9의 최적화 및 교시의 양태는 또한 Cas9 닉카아제를 생성하기 위해서 사용될 수 있다. 출원인은 규정된 돌출부를 가진 DNA 이중가닥 브레이크를 생성하기 위해서 가이드 RNA의 쌍과 조합하여 Cas9 닉카아제를 사용한다. 두 쌍의 가이드 RNA가 사용되는 경우, 개재 DNA 단편을 삭제하는 것이 가능하다. DNA의 외인성 조각이 두 쌍의 가이드 RNA에 의해서 절단되어 게놈 DNA를 가진 양립성 돌출부가 생성된다면, 삭제된 단편을 대체하기 위해서 외인성 DNA 단편이 게놈 DNA와 결찰될 수 있다. 예를 들어, 이것은 헌팅턴병을 치료하기 위해 헌팅틴 (HTT) 유전자에서 트리뉴클레오티드 반복부 확장을 제거하는데 사용될 수 있다.

[1206] 소수의 CAG 반복부를 지닌 외인성 DNA가 제공된다면, 그것은 동일한 돌출부를 지닌 DNA의 단편을 생성할 수 있고, HTT 게놈 유전자좌에 결찰되어 삭제된 단편을 대체할 수 있다.



[1207]

[1208] 이들은 각각 서열 번호 __ 내지 ____이다. 게놈으로 외인성 DNA 단편의 결찰은 상동성 재조합 기구를 필요로 하지 않으며, 따라서 이 방법은 뉴런과 같은 유사분열 후 세포에 사용될 수 있다.

[1209] *실시예 27: CRISPR 시스템의 전달*

[1210] Cas9 및 그것의 키메라 가이드 RNA, 또는 tracr RNA와 crRNA의 조합은 DNA 또는 RNA로서 전달될 수 있다. 둘 다 RNA (정상 또는 염기 또는 백본 변형을 함유하는) 분자로서 Cas9 및 가이드 RNA의 전달은 Cas9 단백질이 세포에 체류하는 시간량을 감소시키기 위해서 사용될 수 있다. 이것은 표적 세포에서 표적외 절단 활성의 수준을 감소시킬 수 있다. mRNA로서 Cas9의 전달은 단백질로 번역되는데 시간이 걸리므로, Cas9 mRNA의 전달 수시간 후에 가이드 RNA를 전달하는 것이 유익할 수 있으며, 이로써 Cas9 단백질과의 상호작용에 이용할 수 있는 가이드 RNA의 수준을 최대화할 수 있다.

[1211] 가이드 RNA 양이 제한된 상황에서, Cas9는 mRNA로서, 가이드 RNA는 가이드 RNA의 발현을 추진하는 프로모터를 가진 DNA 발현 카세트의 형태로서 도입하는 것이 바람직할 수 있다. 이 방식으로, 이용할 수 있는 가이드 RNA의 양이 전사를 통해 증폭될 것이다.

[1212] 다양한 전달 시스템이 숙주 세포에 Cas9 (DNA 또는 RNA) 및 가이드 RNA (DNA 또는 RNA)를 도입하기 위해서 소개될 수 있다. 이들은 리포좀, 바이러스 벡터, 전기천공, 나노입자, 나노와이어(Shalek et al., Nano Letters, 2012), 엑소좀의 사용을 포함한다. 분자 트로잔 호스 리포좀(Pardridge et al., Cold Spring Harb Protoc; 2010; doi:10.1101/pdb.prot5407)이 혈액 뇌 장벽을 가로질러 Cas9와 가이드 RNA를 전달하기 위해서 사용될 수 있다.

[1213] *실시예 28: 트리뉴클레오티드 반복부 장애를 위한 치료 전략*

[1214] 본 출원에서 앞서 언급된 대로, CRISPR 복합체의 표적 폴리뉴클레오티드는 다수의 질환-관련 유전자 및 폴리뉴클레오티드를 포함할 수 있고, 이들 질환 관련 유전자의 일부는 트리뉴클레오티드 반복부 장애라고 하는 일련의 유전적 장애에 속할 수 있다(트리뉴클레오티드 반복부 확장 장애, 삼중 반복부 확장 장애 또는 코돈 반복 장애라고도 한다).

[1215] 이들 질환은 특정 유전자의 트리뉴클레오티드 반복부가 정상적인 안정한 역치를 초과하는 돌연변이에 의해서 야기되며, 이것은 보통 유전자마다 상이할 수 있다. 더 많은 반복부 확장 장애의 발견은 기본적인 유사한 특징에 기초하여 다수의 범주로 이들 장애를 분류하는 것을 허용했다. 특정 유전자의 단백질-코딩 부분에서 CAG 반복부 확장에 의해서 야기되는 헌팅턴병 (HD)과 척수소뇌성 실조증은 범주 I에 포함된다. 표현형적으로 다양해지는 경향이 있으며 보통 규모가 작고 유전자의 엑손에서 발견되는 확장을 포함하는 확장을 가진 질환 또는 장애는 범주 III에 포함된다. 범주 III는 범주 I 또는 II보다 훨씬 더 큰 반복부 확장에 의해서 특정되며, 일반적으로 단백질 코딩 영역 밖에 위치된 질환 또는 장애를 포함한다. 범주 III 질환 또는 장애의 예들은, 제한은 아니지만 취약 X 증후군, 근긴장 디스트로피, 두 가지의 척수소뇌성 실조증, 소아 근간대성 간질, 및 프리드라이히 운동실조를 포함한다.

[1216] 아래 프리드라이히 운동실조에 대해서 언급한 것과 같은 유사한 치료 전략이 다른 트리뉴클레오티드 반복부 또는 확장 장애를 다루기 위해서 또한 채택될 수 있다. 예를 들어, 거의 동일한 전략을 사용하여 치료될 수 있는 다른 삼중 반복부 질환은 근긴장 디스트로피 1 (DM1)인데, 이 경우 3' UTR에 확장된 CTG 모티프가 있다. 프리드라이히 운동실조에서, 이 질환은 프라탁신 (FXN)의 제1 인트론에서 GAA 트리뉴클레오티드의 확장의 결과이다. CRISPR을 사용한 한 가지 치료 전략은 제1 인트론으로부터 GAA 반복부를 삭제하는 것이다. 확장된 GAA 반복부는 DNA 구조에 영향을 미치는 것으로 생각되며, 프라탁신 유전자로 바뀌는 헤테로크로마틴의 형성을 초래한다(도 32a).

[1217] 다른 치료 전략을 능가하는 경쟁적 이점이 아래 열거된다:

[1218] 질환이 프라탁신의 감소된 발현으로 인한 것이므로 siRNA 녹다운은 이 경우 적용될 수 없다. 바이러스 유전자 요법이 현재 조사중이다. HSV-1-기반 벡터가 동물 모델에 프라탁신 유전자를 전달하기 위해 사용되었고 치료 효과를 나타냈다. 그러나, 바이러스-기반 프라탁신 전달의 장기적 효율은 몇 가지 문제를 가지는데, 첫째 건강한 개체에서의 자연적 수준과 일치하도록 프라탁신의 발현을 조절하는 것이 어렵고, 둘째 프라탁신의 장기간 발현은 세포 사멸을 초래한다.

[1219] 뉴클레아제는 건강한 유전자형을 회복하기 위하여 GAA 반복부를 삭제하기 위해서 사용될 수 있지만, 징크 핑거 뉴클레아제 및 TALEN 전략은 전달과 뉴클레아제 조작에서 모두 상이한 두 쌍의 고 효율 뉴클레아제의 전달을 필요로 한다(ZFN 또는 TALEN에 의한 게놈 DNA의 효과적인 삭제는 달성하는 것이 어렵다).

[1220] 상기 전략과 달리, CRISPR-Cas 시스템은 명백한 이점을 가진다. Cas9 효소는 더 효과적이며 더 복합적인데, 이것은 하나 이상의 표적이 동시에 설정될 수 있다는 의미이다. 지금까지 인간 세포에서 Cas9에 의한 30%를 초과하는 게놈 DNA의 유효 정확도는 30%만큼 높을 수 있으며, 향후 개선될 수 있다. 또한, 헌팅턴병(HD)과 같은 특정 트리뉴클레오티드 반복부 장애와 관련해서, 코딩 영역의 트리뉴클레오티드 반복부는 두 대립유전자에 차이가 있는 경우 다뤄질 수 있다. 구체적으로, HD 환자가 돌연변이 HTT에 이형접합성이고, wt과 야생형 HTT 대립유전자 사이에 SNP와 같은 뉴클레오티드 차이가 있는 경우, 돌연변이 HTT 대립유전자를 특이적으로 표적화하기 위해서 Cas9가 사용될 수 있다. ZFN 또는 TALEN은 단일 염기 차이에 기초하여 두 대립유전자를 구별할 수 있는 능력을 갖지 않을 것이다.

[1221] 프리드라이히 운동실조를 다루기 위해 CRISPR-Cas9 효소를 사용한 전략을 채택하는데 있어서, 출원인은 GAA 확장이 측접한 부위를 표적화하는 다수의 가이드 RNA를 설계하고, 가장 효과적이며 특이적인 것이 선택된다(도 32b).

[1222] 출원인은 GAA 확장 영역의 정확성을 매개하기 위해서 Cas9와 함께 FXN의 인트론 1을 표적화하는 가이드 RNA의 조합을 전달한다. 척수에서 Cas9의 효과적인 전달을 매개하기 위해서 AAV9가 사용될 수 있다.

[1223] Gaa 확장에 인접한 Alu 요소가 중요하다고 고려된다면, 표적화될 수 있는 부위의 수에 제약이 있을 수 있지만, 출원인은 그것의 파괴를 피하기 위해서 전략을 수정할 수 있다.

[1224] 대안적 전략:

[1225] Cas9를 사용하여 게놈을 변형하는 것이 아니라, 출원인은 또한 FXN 유전자에 전사 활성화 도메인을 표적화하기

위해서 Cas9 (뉴클레아제 활성 결핍)-기반 DNA 결합 도메인을 사용하여 FXN 유전자를 직접 활성화할 수 있다. 출원인은 다른 방법에 비하여 이것이 충분히 강력하다는 것을 보장하기 위하여 Cas9-매개 인공 전사 활성의 강인성을 해결할 수 있어야 한다(Tremblay et al., Transcription Activator-Like Effector 단백질s Induce the Expression of the Frataxin Gene; Human Gene Therapy. August 2012, 23(8): 883-890.)

[1226] *실시예 29: Cas9 닉카아제를 사용한 표적외 절단을 최소화하기 위한 전략*

[1227] 본 출원에서 앞서 언급된 대로, Cas9는 다음의 돌연변이 중 하나 이상을 통해서 단일가닥 절단을 매개하기 위해 돌연변이될 수 있다: D10A, E762A, 및 H840A.

[1228] NHEJ를 통한 유전자 녹아웃을 매개하기 위해서, 출원인은 두 가이드 RNA와 함께 Cas9의 닉카아제 버전을 사용한다. 각 개별 가이드 RNA에 의한 표적외 흉터형성은 일차적으로 돌연변이 없이 수선될 수 있고, 이중가닥 브레이크 (이것은 NHEJ를 통해 돌연변이를 초래할 수 있다)는 표적 부위가 서로 인접해 있을 때만 발생한다. 이중 흉터형성에 의해서 도입된 이중가닥 브레이크는 블런트 형태가 아니므로, TREX1과 같은 단부-가공 효소의 공-발현이 NHEJ 활성의 수준을 증가시킬 것이다.

[1229] 표에서 다음의 표적 리스트는 다음의 질환에 수반되는 유전자에 대한 것이다:

[1230] 라포라병 - 뉴런에서 글리코겐을 감소시키기 위한 표적 GSY1 또는 PPP1R3C (PTG)

[1231] 고콜레스테롤혈증 - 표적 PCSK9

[1232] 표적 서열은 스페이서에 상이한 수의 뉴클레오티드 (0 내지 3 bp)를 가진 쌍들 (L 및 R)로 열거된다. 표적 유전자좌에 이중가닥 브레이크를 도입하기 위해 각 스페이스는 또한 단독으로 야생형 Cas9와 사용될 수 있다.

[1233]

		(SEQ ID NO:)
GYS1 (인간)	GGCC-L	ACCCTTGTTAGCCACCTCCC
	GGCC-R	GAACGCAGTGCTCTTCGAAG
	GGNCC-L	CTCACGCCCTGCTCCGTGTA
	GGNCC-R	GGCGACAACTACTTCCTGGT
	GGNNCC-L	CTCACGCCCTGCTCCGTGTA
	GGNNCC-R	GGGCGACAACTACTTCCTGG
	GGNNNCC-L	CCTCTTCAGGGCCGGGGTGG
	GGNNNCC-R	GAGGACCCAGGTGGAACTGC
PCSK9 (인간)	GGCC-L	TCAGCTCCAGGCGGTCCTGG
	GGCC-R	AGCAGCAGCAGCAGTGGCAG
	GGNCC-L	TGGGCACCGTCAGCTCCAGG
	GGNCC-R	CAGCAGTGGCAGCGGCCACC
	GGNNCC-L	ACCTCTCCCCTGGCCCTCAT
	GGNNCC-R	CCAGGACCGCCTGGAGCTGA
	GGNNNCC-L	CCGTCAGCTCCAGGCGGTCC
	GGNNNCC-R	AGCAGCAGCAGCAGTGGCAG
PPP1R3C (PTG) (인간)	GGCC-L	ATGTGCCAAGCAAAGCCTCA
	GGCC-R	TTCGGTCATGCCCGTGGATG
	GGNCC-L	GTCGTTGAAATTCATCGTAC
	GGNCC-R	ACCACCTGTGAAGAGTTTCC
	GGNNCC-L	CGTCGTTGAAATTCATCGTA
	GGNNCC-R	ACCACCTGTGAAGAGTTTCC
Gys1 (마우스)	GGCC-L	GAACGCAGTGCTTTTCGAGG
	GGCC-R	ACCCTTGTTGGCCACCTCCC
	GGNCC-L	GGTGACAACTACTATCTGGT
	GGNCC-R	CTCACACCCTGCTCCGTGTA
	GGNNCC-L	GGGTGACAACTACTATCTGG
	GGNNCC-R	CTCACACCCTGCTCCGTGTA
	GGNNNCC-L	CGAGAACGCAGTGCTTTTCG
	GGNNNCC-R	ACCCTTGTTGGCCACCTCCC
PPP1R3C (PTG) (마우스)	GGCC-L	ATGAGCCAAGCAAATCCTCA
	GGCC-R	TTCCGTCATGCCCGTGGACA
	GGNCC-L	CTTCGTTGAAAACCATTGTA
	GGNCC-R	CCACCTCTGAAGAGTTTCCT
	GGNNCC-L	CTTCGTTGAAAACCATTGTA
	GGNNCC-R	ACCACCTCTGAAGAGTTTCC
	GGNNNCC-L	CTTCCACTCACTCTGCGATT
	GGNNNCC-R	ACCATGTCTCAGTGTCAAGC
PCSK9 (마우스)	GGCC-L	GGCGGCAACAGCGGCAACAG

GGCC-R	ACTGCTCTGCGTGGCTGCGG
GGNNCC-L	CCGCAGCCACGCAGAGCAGT
GGNNCC-R	GCACCTCTCCTCGCCCCGAT

[1234]

[1235] 가이드 RNA의 안정성을 개선하고 특이성을 증가시키기 위한 다른 전략

[1236] 1. 가이드 RNA의 5'의 뉴클레오티드는 천연 RNA에서와 같이 포스포에스테르 결합이 아니라 티오에스테르 결합을 통해서 연결될 수 있다. 티오에스테르 결합은 가이드 RNA가 내인성 RNA 분해 기구에 의해서 소화되는 것을 방지할 수 있다.

[1237] 2. 가이드 RNA의 가이드 서열(5' 20 bp)의 뉴클레오티드는 결합 특이성을 개선하기 위해 염기로서 브릿지 핵산(BNA)을 사용할 수 있다.

[1238] *실시예 30: 신속하며 복합적인 게놈 편집을 위한 CRISPR-Cas*

[1239] 본 발명의 양태는 유전자 변형의 효율 및 특이성이 표적 설계 후 3-4일 내에 시험될 수 있고, 변형된 클론성 세포주가 2-3주 내에 유도될 수 있는 프로토콜 및 방법에 관한 것이다.

[1240] 프로그램가능한 뉴클레아제는 높은 정확도로 게놈 변경을 매개하기 위한 강력한 기술이다. 미생물 CRISPR 후천면역 시스템으로부터의 RNA-가이드된 Cas9 뉴클레아제는 그것의 가이드 RNA에서 20-nt 표적화 서열을 간단히 특정함으로써 진핵 세포에서의 효과적인 게놈 편짐을 촉진하는데 사용될 수 있다. 출원인은 포유류 세포에서 게놈 편집을 촉진하고, 하류 기능 연구를 위한 세포주를 생성하기 위해 Cas9를 적용하는 일련의 프로토콜을 설명한다. 표적 설계에서 시작하여 효과적이며 특이적인 유전자 변형이 3-4 일 내에 달성될 수 있고, 변형된 클론성 세포주가 2-3 주 내에 유도될 수 있다.

[1241] 생물학적 시스템과 유기체를 조작하는 능력은 기초과학, 의학 및 생물기술 분야에서의 용도에서 상당한 잠재력을 보유한다. 내인성 게놈 유전자좌의 정확한 편집을 촉진하는 프로그램가능한 서열-특이적 엔도 뉴클레아제가 이전에 유전적으로 다룰 수 없었던 것들을 포함해서 광범위한 종들에서 유전자 요소 및 결과적인 변형에 대한 현재 가능한 체계적 의제이다. 징크 핑거 뉴클레아제(ZFN), 전사 활성인자-유사 이펙터 뉴클레아제(TALEN), 및 RNA-가이드 CRISPR-Cas 뉴클레아제 시스템을 포함하여 많은 게놈 편집 기술이 최근 몇년간 출몰했다. 처음의 두 가지 기술은 특이적 게놈 유전자좌에 표적화된 DNA 이중가닥 브레이크 (DSB)를 유도하기 위해서 모듈형 DNA-결합 단백질에 엔도뉴클레아제 촉매 도메인을 매다는 공통된 전략을 사용한다. 반대로, Cas9는 표적 DNA와의 왓슨-크릭 염기쌍을 통해서 작은 RNA에 의해 가이드된 뉴클레아제로서, 설계가 쉽고 효과적이며 다양한 세포 타입과 유기체에 대한 고-처리량 및 복합적 유전자 편집에 아주 적합한 시스템을 제시한다. 여기서 출원인은 포유류 세포에서 효과적인 게놈 편집을 촉진하고, 하류 기능 연구를 위한 세포주를 생성하기 위해 최근 개발된 Cas9 뉴클레아제를 적용하기 위한 일련의 프로토콜을 설명한다.

[1242] ZFN 및 TALEN와 마찬가지로, Cas9는 표적 게놈 유전자좌에서 DSB를 자극함으로써 게놈 편집을 촉진한다. Cas9에 의한 절단시 표적 유전자좌는 DNA 손상 수선을 위한 두 가지 주요 경로, 즉 에러-경향 비-상동성 단부 연결 (NHEJ) 또는 고 충실도 상동체 지정 수선 (HDR) 경로 중 하나를 거친다. 두 경로 모두 바람직한 편집 결과를 달성하기 위해서 이용할 수 있다.

[1243] NHEJ: 수선 주형의 부재하에, NHEJ 과정은 DSB를 재결찰하며, 이것은 삽입결실부 돌연변이 형태의 흉터를 남길 수 있다. 이 과정은 유전자 녹아웃을 달성하기 위해서 활용할 수 있으며, 코딩 엑손 내에 발생한 삽입결실부는 프레임시프트 돌연변이와 조기성숙 중단 코돈을 초래할 수 있다. 다중 DBS가 또한 게놈에 더 큰 결실을 매개하기 위해서 이용될 수 있다.

[1244] HDR: 상동체 지정 수선은 NHEJ의 대안인 중요한 DNA 수선 경로이다. HDR은 전형적으로 NHEJ보다 낮은 빈도로 일어나지만, 그것은 외인성 도입된 수선 주형의 존재하에 표적 유전자좌에 정확하며 한정된 변형을 생성하기 위해서 이용될 수 있다. 수선 주형은 삽입 서열이 측접된 상동체 암을 가진 종래의 DNA 표적화 작제물과 유사하게 설계된, 이중가닥 DNA의 형태이거나, 또는 단일가닥 DNA 올리고뉴클레오티드 (ssODN)일 수 있다. 후자는 결과적인 유전자 변형을 검사하기 위한 단일 뉴클레오티드 돌연변이의 도입과 같은, 게놈에 작은 편집부를 만들기 위한 효과적이며 간단한 방법을 제공한다. NHEJ와는 달리, HDR는 일반적으로 분열하는 세포에서만 활성이고, 그것의 효율은 세포 타입 및 상태에 따라 다양하다.

[1245] CRISPR의 개요: 반대로, CRISPR-Cas 시스템은 Cas9 뉴클레아제 및 짧은 가이드 RNA로 구성된 최소 2-구성성분 시스템이다. 상이한 유전자좌에 Cas9의 재표적화 또는 다중 유전자의 동시 편집은 단순히 상이한 20-bp 올리고

뉴클레오티드의 클로닝을 필요로 한다. Cas9 뉴클레아제의 특이성은 더 충분히 해명되어야 하지만, CRISPR-Cas 시스템의 간단한 왓슨-크릭 염기쌍이 ZFN 또는 TALEN 도메인보다 더 잘 예측할 수 있을 것 같다.

[1246] 타입 II CRISPR-Cas (군락화된 규칙적으로 산재된 짧은 팔린드롬형 반복부)는 외래 유전자 요소를 절단하기 위해 Cas9를 사용하는 박테리아 후천 면역 시스템이다. Cas9는 한 쌍의 비-코딩 RNA, 가변 crRNA 및 필요한 보조 tracr RNA에 의해서 가이드된다. crRNA는 20-nt 가이드 서열을 함유하고, 왓슨-크릭 염기쌍을 통해 표적 DNA를 위치시킴으로써 특이성을 결정한다. 자생 박테리아 시스템에서, 다중 crRNA는 다양한 표적에 대해 Cas9를 지정하기 위해서 공-전사된다. *Streptococcus pyogenes* 로부터 유래된 CRISPR-Cas 시스템에서, 표적 DNA는 5'-NGG/NRG 프로토스페이서 인접 모티프(PAM) 바로 앞에 선행해야 하며, 이것은 다른 CRISPR 시스템의 경우 변할 수 있다.

[1247] CRISPR-Cas는 인간 코돈-최적화된 Cas9와 필수 RNA 구성성분의 이종성 발현을 통해서 포유류 세포에서 재구성된다. 또한, crRNA와 tracrRNA가 키메라 합성 가이드 RNA (sgRNA)를 생성하기 위해서 융합될 수 있다. 따라서, Cas9는 sgRNA 내에서 20-nt 가이드 서열을 변경함으로써 관심대상의 어떤 표적을 향해 재지정될 수 있다.

[1248] 실시 용이성과 복합적 능력이 주어진 경우, Cas9는 NHEJ와 HDR 모두를 통해 특이적 돌연변이를 지닌 유전자 조작된 진핵 세포를 생성하기 위해 사용되었다. 게다가, 배아에 sgRNA 및 Cas9를 암호화하는 mRNA의 직접 주입은 다수의 변형된 대립유전자를 가진 유전자이식 마우스의 신속한 발생을 가능하게 했으며, 이들 결과는 유전적으로 다루기 어려운 유기체를 편집하는 것에 대한 가능성을 가진다.

[1249] 촉매 도메인 중 하나가 파괴된 돌연변이 Cas9는 DNA를 절단하는 것이 아니라 흠집을 내도록 유전자 조작되었으며, 이것은 단일가닥 브레이크와 HDR을 통한 우선적 수선을 허용하고, 잠재적으로 표적외 DSB로부터의 원치않는 삽입결실부 돌연변이를 완화한다. 추가로, DNA-절단 촉매 잔기가 둘 다 돌연변이된 Cas9 돌연변이는 이콜라이에서의 전사 조절을 가능하게 하도록 개조되었으며, 이것은 다양한 용도에 대한 기능화된 Cas9의 잠재성을 증명한다. 본 발명의 특정 양태는 인간 세포의 복합적인 편집을 위한 Cas9의 구성 및 적용에 관한 것이다.

[1250] 출원인은 진핵생물 유전자 편집을 촉진할 수 있는 인간 코돈-최적화된, 핵 위치 서열-측접된 Cas9를 제공했다. 출원인은 20-nt 가이드 서열의 설계를 위한 고려사항, sgRNA의 신속한 구성 및 기능적 검증을 위한 프로토콜, 및 마지막으로 인간 배아 신장 (HEK-293FT) 및 인간 줄기세포 (HUES9) 계통에서 NHEJ- 및 HDR-기반 게놈 변형을 모두 매개하기 위한 Cas9 뉴클레아제의 사용을 설명한다. 이 프로토콜은 다른 세포 타입 및 유기체에도 마찬가지로 적용될 수 있다.

[1251] sgRNA를 위한 표적 선택: 유전자 표적화를 위한 20-nt 가이드 서열의 선택에는 두 가지 중요한 고려사항이 있다: 1) *S. pyogenes* Cas9의 경우 표적 서열이 5'-NGG PAM 앞에 와야 하고, 2) 가이드 서열은 표적외 활성을 최소화하도록 선택되어야 한다. 출원인은 관심대상의 투입 서열을 취해서 적합한 표적 부위를 확인하는 온라인 Cas9 표적화 디자인 툴을 제공했다. 각 sgRNA에 대한 표적외 변형을 실험적으로 평가하기 위해서, 출원인은 또한 각 의도된 표적에 대한 표적외 부위를 컴퓨터로 예측하고, 출원인의 정량적 특이성 분석에 따라서 염기쌍 미스매치 확인, 위치 및 분포의 영향에 순서를 매겼다.

[1252] 컴퓨터로 예측된 표적외 부위에 대한 상세한 정보는 다음과 같다:

[1253] 표적외 절단 활성을 위한 고려사항: 다른 뉴클레아제와 유사하게, Cas9는 감소된 빈도로 게놈에서 표적외 DNA 표적을 절단할 수 있다. 주어진 가이드 서열이 표적외 활성을 나타내는 정도는 효소 농도, 이용된 특정 가이드 서열의 열역학, 및 표적 게놈에서 유사한 서열의 풍부성을 포함하는 요인들의 조합에 따른다. Cas9의 통상적 용도에서, 표적외 절단 정도를 최소화하고, 또한 표적외 절단의 존재를 검출할 수 있는 방식을 고려하는 것이 중요하다.

[1254] 표적외 활성 최소화: 세포주에서의 용도를 위해서, 출원인은 표적외 게놈 변형의 정도를 감소시키기 위해 다음의 두 단계를 추천했다. 먼저, 본 발명자의 온라인 CRISPR 표적 선택 도구를 사용하여 주어진 가이드 서열이 표적외 부위를 가질 가능성을 컴퓨터로 평가하는 것이 가능하다. 이들 분석은 가이드 서열과 유사한 서열인 표적외 서열에 대해 게놈에서 철저한 검색을 통해 수행된다. sgRNA와 그것의 표적 DNA 사이에 미스매칭 염기의 효과에 대한 포괄적인 실험적 조사는 미스매치 용인성이 1) 위치 의존적 (가이드 서열의 3' 단부에서 8-14 bp는 5' 염기보다 미스매치에 덜 용인성이다), 2) 양 의존적(일반적으로 3 개를 초과하는 미스매치는 용인되지 않는다), 3) 가이드 서열 의존적(일부 가이드 서열은 다른 것보다 미스매치에 덜 용인성이다), 및 4) 농도 의존적 (표적외 절단은 트랜스펙션된 DNA의 양에 매우 민감하다)이라는 것을 드러냈다. 출원인의 표적 부위 분석 웹 도구(웹사이트 genome-engineering.org/tools에서 이용가능)는 표적 게놈에서 가능성 있는 표적외 부위에 대한 예측을

제공하기 위해 이들 기준을 통합한다. 두 번째, 출원인은 표적외 활성을 최소화하기 위해서 Cas9 및 sgRNA 발현 플라스미드의 양을 적정할 것을 추천했다.

[1255] 표적외 활성의 검출: 출원인의 CRISPR 표적화 웹 도구를 사용하여 가장 가능성 있는 표적외 부위와 이들 부위의 서베이어 또는 시퀀싱 분석을 수행하는 프라이머의 리스트를 생성하는 것이 가능하다. Cas9를 사용하여 생성된 동질유전자 클론의 경우, 출원인은 어떤 바람직하지 않은 돌연변이를 점검하기 위해 이들 후보 표적외 부위를 시퀀싱할 것을 강력하게 추천했다. 예측된 후보 리스트에 포함되지 않은 부위에 표적외 변형이 있을 수 있으며, 표적외 부위의 부재를 완벽히 검증하기 위해 완전한 게놈 서열에 대해 수행되어야 한다는 것은 그럴만하다. 또한, 몇 개의 DSB가 동일한 게놈 내에 유도되는 복합적 분석에서, 전위 사건은 낮은 비율로 있을 수 있으며, 심층 시퀀싱과 같은 다양한 기술을 사용하여 평가될 수 있다.

[1256] 온라인 툴은 1) sgRNA 작제물을 제조하고, 2) 표적 변형 효율을 분석하고, 3) 잠재적 표적외 부위에서 절단을 평가하는데 필요한 모든 올리고 및 프라이머의 서열을 제공한다. sgRNA를 발현하는데 사용된 U6 RNA 폴리머라제 III 프로모터는 그것의 전사체의 제1 염기로서 구아닌 (G) 뉴클레오티드를 선호하기 때문에, 20-nt 가이드 서열이 G로 시작하지 않는 경우 가외의 G가 sgRNA의 5'에 덧붙는 것이 좋다.

[1257] sgRNA 구성 및 전달을 위한 접근법: 원하는 용도에 따라서, sgRNA는 다음과 같이 전달될 수 있다: 1) 발현 카세트를 함유하는 PCR 앰플리콘 또는 2) sgRNA-발현 플라스미드. PCR-기반 sgRNA 전달은 U6 프로모터 주형을 증폭하는데 사용된 역 PCR 프라이머 상에 맞춤된 sgRNA 서열을 덧붙인다. 결과의 앰플리콘은 Cas9를 함유하는 플라스미드 (PX165)와 공-트랜스펙션될 수 있다. 이 방법은 다수의 후보 sgRNA의 신속한 스크리닝에 최적인데, 기능 시험을 위한 세포 트랜스펙션이 sgRNA-암호화 프라이머를 얻은 후 단지 수 시간 내에 수행될 수 있기 때문이다. 이 간단한 방법은 플라스미드-기반 클로닝 및 서열 검증에 대한 필요성을 배제하기 때문에, 대형 녹아웃 라이브러리를 생성하기 위해 상당수의 sgRNA를 시험하거나 공-트랜스펙션하는 경우나 다른 규모-민감성 용도에 아주 적합하다. sgRNA-암호화 프라이머는 플라스미드-기반 sgRNA 전달에 필요한 약 20-bp 올리고와 비교하여 100-bp를 넘는다는 것이 주지된다.

[1258] sgRNA를 위한 발현 플라스미드의 구성도 역시 간단하고 신속하며, 부분적으로 상보성인 올리고뉴클레오티드 쌍을 사용한 단일 클로닝 단계를 수반한다. 올리고 쌍의 아닐링 후, 결과의 가이드 서열이 Cas9를 지닌 플라스미드와 sgRNA 서열의 나머지를 지닌 비변이 스캐폴드 (PX330)에 삽입될 수 있다. 트랜스펙션 플라스미드도 생체내 전달을 위한 바이러스 생산을 가능하게 하도록 변형될 수 있다.

[1259] PCR 및 플라스미드-기반 전달 방법에 더하여, Cas9와 sgRNA는 둘 다 RNA로서 세포에 도입될 수 있다.

[1260] 수선 주형 설계: 전통적으로, 표적화된 DNA 변형은 변경 부위의 측접한 상동체 암을 함유하는 플라스미드-기반 도너 수선 주형의 사용이 필요했다. 각 사이드에서 상동체 암은 길이가 다양할 수 있지만, 전형적으로 500 bp를 초과한다. 이 방법은 형광 단백질 또는 항생물질 내성 마커와 같은 리포터 유전자의 삽입을 포함하여 큰 변형을 생성하기 위해서 사용될 수 있다. 표적화 플라스미드의 설계 및 구성은 다른 곳에서 설명되었다.

[1261] 더 최근에는, 단일가닥 DNA 올리고뉴클레오티드 (ssODN)가 클로닝 없이 한정된 유전자좌 내에서 짧은 변형을 위해 표적화 플라스미드 대신 사용되었다. 높은 HDR 효율을 달성하기 위해서, ssODN는 각 사이드에 적어도 40 bp의 측면 서열을 함유하며, 이들은 표적 영역에 상동성이고, 표적 유전자좌에 대해 센스 또는 안티센스 방향으로 방향화될 수 있다.

[1262] 기능 시험

[1263] 서베이어 뉴클레아제 분석: 출원인은 서베이어 뉴클레아제 분석 또는 PCR 앰플리콘 시퀀싱에 의해서 삽입결실부 돌연변이를 검출했다. 출원인의 온라인 CRISPR 표적 디자인 툴은 두 접근법을 위해 추천된 프라이머를 제공한다. 그러나, 서베이어 또는 시퀀싱 프라이머는 또한 게놈 DNA로부터 관심대상의 영역을 수동으로 증폭하고, NCBI 프라이머-BLAST를 사용하여 비-특이적 앰플리콘을 피할 수 있도록 설계될 수 있다. 서베이어 프라이머는 Cas9 표적의 어느 사이드에서 300-400 bp (총 600-800 bp 앰플리콘의 경우)를 증폭하도록 설계되어야 하며, 이로써 겔 전기영동에 의해서 절단 밴드의 분명한 시각화가 허용된다. 과도한 프라이머 다이머 형성을 방지하기 위해서, 서베이어 프라이머는 전형적으로 25-nt 길이 이하가 되고, 용융 온도는 약 60℃이도록 설계되어야 한다. 출원인은 특정 PCR 앰플리콘뿐만 아니라 서베이어 뉴클레아제 소화 과정 동안 비-특이적 절단의 존재에 대해 후보 프라이머의 각 쌍을 시험할 것을 추천했다.

[1264] 플라스미드- 또는 ssODN-매개 HDR: HDR은 변형된 영역의 PCR-증폭 및 시퀀싱을 통해서 검출될 수 있다. 이 목적을 위한 PCR 프라이머는 잔류 수선 주형의 거짓 검출을 피하기 위해서 상동체 암으로 걸쳐진 영역 밖에서 아닐

링되어야 한다(HDR Fwd 및 Rev, 도 30). ssODN-매개 HDR의 경우, 서베이어 PCR 프라이머가 사용될 수 있다.

[1265] 시퀀싱에 의한 삽입결실부 또는 HDR의 검출: 출원인은 Sanger 또는 차세대 심층 시퀀싱 (NGS)에 의해서 표적화된 게놈 변형을 검출했다. 전자의 경우, 변형된 영역으로부터의 게놈 DNA가 서베이어 또는 HDR 프라이머를 사용하여 증폭될 수 있다. 앰플리콘은 형질전환을 위해 pUC19와 같은 플라스미드에 서브클로닝되어야 하며, 클론 유전자형을 드러내기 위해서 개별 콜로니가 시퀀싱될 수 있다.

[1266] 출원인은 전형적으로 100-200 bp 크기 범위로 더 짧은 앰플리콘을 위해 차세대 시퀀싱 (NGS) 프라이머를 설계했다. NHEJ 돌연변이 검출을 위해서, 프라이밍 영역과 Cas9 표적 부위 사이에 적어도 10-20 bp가 있는 프라이머를 설계함으로써 더 긴 삽입결실부의 검출을 허용하는 것이 중요하다. 출원인은 복합적인 심층 시퀀싱을 위해 바코드 어댑터를 부착하기 위한 2-단계 PCR 방법을 위한 가이드라인을 제공한다. 출원인은 일반적으로 거짓 양성 삽입결실부 수준이 낮기 때문에 Illumina 플랫폼을 추천했다. 다음에, 표적외 분석 (앞서 설명된)이 ClustalW, Geneious, 또는 간단한 서열 분석 스크립트와 같은 판독 정렬 프로그램을 통해서 수행될 수 있다.

[1267] 재료 및 시약

[1268] <u>sgRNA 제조:</u>

[1269] UltraPure DNaseRNase-무함유 증류수(Life Technologies, 카탈로그 번호 10977-023)

[1270] 허큘라제 II 융합 폴리머라제(Agilent Technologies, 카탈로그 번호 600679)

[1271] CRITICAL. 표준 Taq 폴리머라제, 이것은 3'-5' 엑소뉴클레아제 교정 활성을 결여하고, 낮은 충실도를 가지며, 증폭 에러를 초래할 수 있다. 허큘라제 II는 최소한의 최적화로 PCR 생성물을 높은 수율로 생성하는 고 충실도 폴리머라제(Pfu와 동등한 충실도)이다. 다른 고 충실도 폴리머라제가 치환될 수 있다.

[1272] 허큘라제 II 반응 완충제(5x; Agilent Technologies, 폴리머라제와 함께 포함된다)

[1273] dNTP 용액 믹스(25 mM 각각; Enzymatics, 카탈로그 번호 N205L)

[1274] $MgCl_{2}$(25mM; ThermoScientific, 카탈로그 번호 R0971)

[1275] QIAquick 겔 추출 키트(Qiagen, 카탈로그 번호 28704)

[1276] QIAprep 스핀 미니프렙 키트(Qiagen, 카탈로그 번호 27106)

[1277] UltraPure TBE 완충제(10X; Life Technologies, 카탈로그 번호 15581-028)

[1278] SeaKem LE 아가로스(Lonza, 카탈로그 번호 50004)

[1279] SYBR Safe DNA 염색(10,000x; Life Technologies, 카탈로그 번호 S33102)

[1280] 1-kb Plus DNA 래더(Life Technologies, 카탈로그 번호 10787-018)

[1281] TrackIt CyanOrange 로딩 완충제(Life Technologies, 카탈로그 번호 10482-028)

[1282] FastDigest BbsI(BpiI)(Fermentas/ThermoScientific, 카탈로그 번호 FD1014)

[1283] Fermentas 탄고 완충제(Fermentas/ThermoScientific, 카탈로그 번호 BY5)

[1284] DL-디티오트레이톨(DTT; Fermentas/ThermoScientific, 카탈로그 번호 R0862)

[1285] T7 DNA 리가아제(Enzymatics, 카탈로그 번호 L602L)

[1286] 주: 더 흔히 사용되는 T4 리가아제로는 치환하지 않는다. T7 리가아제는 블런트 단부보다 점착성 단부에 대해 1,000-배 더 높은 활성을 가지고, 상업적으로 이용가능한 농축 T4 리가아제보다 높은 전체적 활성을 가진다.

[1287] T7 2X 신속 결찰 완충제(T7 DNA 리가아제와 함께 포함된다, Enzymatics, 카탈로그 번호 L602L)

[1288] T4 폴리뉴클레오티드 키나제(New England Biolabs, 카탈로그 번호 M0201S)

[1289] T4 DNA 리가아제 반응 완충제(10X; New England Biolabs, 카탈로그 번호 B0202S)

[1290] 아데노신 5'-트리포스페이트(10 mM; New England Biolabs, 카탈로그 번호 P0756S)

[1291] PlasmidSafe ATP-의존성 DNase (Epicentre, 카탈로그 번호 E3101K)

[1292] 원샷 Stbl3 화학적 컴페턴트 *Escherichia coli* (*E. coli*) (Life Technologies, 카탈로그 번호 C7373-03)

[1293] SOC 배지(New England Biolabs, 카탈로그 번호 B9020S)

[1294] LB 배지(Sigma, 카탈로그 번호 L3022)

[1295] LB 아가 배지(Sigma, 카탈로그 번호 L2897)

[1296] 암피시린, 멸균 여과(100 mg mℓ-1; Sigma, 카탈로그 번호 A5354)

[1297] **포유류 세포 배양물**:

[1298] HEK293FT 세포(Life Technologies, 카탈로그 번호 R700-07)

[1299] 둘베코 최소 이글 배지(DMEM, 1X, 고 글루코오스; Life Technologies, 카탈로그 번호 10313-039)

[1300] 둘베코 최소 이글 배지(DMEM, 1X, 고 글루코오스, 페놀 레드 없음; Life Technologies, 카탈로그 번호 31053-028)

[1301] 둘베코 포스페이트-완충 식염수(DPBS, 1X; Life Technologies, 카탈로그 번호 14190-250)

[1302] 태아 소 혈청, 정성적으로 열 불활성화됨 (Life Technologies, 카탈로그 번호 10438-034)

[1303] Opti-MEM I 감소된-혈청 배지(FBS; Life Technologies, 카탈로그 번호 11058-021)

[1304] 페니실린-스트렙토마이신 (100x; Life Technologies, 카탈로그 번호 15140-163)

[1305] TrypLE™ Express(1X, 페놀 레드 없음; Life Technologies, 카탈로그 번호 12604-013)

[1306] 리포펙타민 2000 트랜스펙션 시약(Life Technologies, 카탈로그 번호 11668027)

[1307] Amaxa SF 세포주 4D-Nucleofector® X 키트 S (32 RCT; Lonza, 카탈로그 번호 V4XC-2032)

[1308] HUES 9 세포주(HARVARD STEM CELL SCIENCE)

[1309] Geltrex LDEV-무함유 감소된 성장인자 기저막 바탕질(Life Technologies, 카탈로그 번호 A1413201)

[1310] mTeSR1 배지(Stemcell Technologies, 카탈로그 번호 05850)

[1311] 아큐타제 세포 탈착 용액(Stemcell Technologies, 카탈로그 번호 07920)

[1312] ROCK 저해제(Y-27632; Millipore, 카탈로그 번호 SCM075)

[1313] Amaxa P3 1차 세포 4D-Nucleofector® X 키트 S (32 RCT; Lonza 카탈로그 번호 V4XP-3032)

[1314] **제노타이핑 분석**:

[1315] QuickExtract DNA 추출 용액(Epicentre, 카탈로그 번호 QE09050)

[1316] 서베이어, RFLP 분석, 또는 시퀀싱용 PCR 프라이머(프라이머 표 참조)

[1317] 허큘라제 II 융합 폴리머라제(Agilent Technologies, 카탈로그 번호 600679)

[1318] 주: 서베이어 분석은 단일-염기 미스매치에 민감하므로, 고 충실도 폴리머라제를 사용하는 것이 특히 중요하다. 다른 고 충실도 폴리머라제가 치환될 수 있다.

[1319] 허큘라제 II 반응 완충제(5x; Agilent Technologies, 폴리머라제와 함께 포함된다)

[1320] dNTP 용액 믹스(25 mM 각각; Enzymatics, 카탈로그 번호 N205L)

[1321] QIAquick 겔 추출 키트(Qiagen, 카탈로그 번호 28704)

[1322] Taq 완충제(10x; Genscript, 카탈로그 번호 B0005)

[1323] 표준 겔 전기영동을 위한 서베이어 돌연변이 검출 키트(Transgenomic, 카탈로그 번호 706025)

[1324] UltraPure TBE 완충제(10x; Life Technologies, 카탈로그 번호 15581-028)

[1325] SeaKem LE 아가로스(Lonza, 카탈로그 번호 50004)

[1326] 4 내지 20% TBE 겔 1.0 mm, 15-웰(Life Technologies, 카탈로그 번호 EC62255BOX)

[1327] Novex® Hi-Density TBE 샘플 완충제(5X; Life Technologies, 카탈로그 번호 LC6678)

[1328] SYBR 금 핵산 겔 염색(SYBR Gold Nucleic Acid Gel Stain)(10,000X; Life Technologies, 카탈로그 번호 S-11494)

[1329] 1-kb Plus DNA 래더(Life Technologies, 카탈로그 번호 10787-018)

[1330] TrackIt CyanOrange 로딩 완충제(Life Technologies, 카탈로그 번호 10482-028)

[1331] FastDigest HindIII(Fermentas/ThermoScientific, 카탈로그 번호 FD0504)

[1332] **장비:**

[1333] 멸균 여과 피펫 팁(Corning)

[1334] 표준 1.5 ml 미소원심분리 튜브(Eppendorf, 카탈로그 번호 0030 125.150)

[1335] Axygen 96-웰 PCR 플레이트(VWR, 카탈로그 번호 PCR-96M2-HSC)

[1336] Axygen 8-스트립 PCR 튜브(Fischer Scientific, 카탈로그 번호 14-222-250)

[1337] 팔콘 튜브, 폴리프로필렌, 15 ml(BD Falcon, 카탈로그 번호 352097)

[1338] 팔콘 튜브, 폴리프로필렌, 50 ml(BD Falcon, 카탈로그 번호 352070)

[1339] 세포 스트레이너 캡을 가진 둥근-바닥 튜브, 5 ml(BD Falcon, 카탈로그 번호 352235)

[1340] 페트리 디쉬(60 mm × 15 mm; BD Biosciences, 카탈로그 번호 351007)

[1341] 조직 배양 플레이트(24-웰; BD Falcon, 카탈로그 번호 353047)

[1342] 조직 배양 플레이트(96-웰, 편평 바닥; BD Falcon, 카탈로그 번호 353075)

[1343] 조직 배양 디쉬(100 mm; BD Falcon, 353003)

[1344] 프로그램가능한 온도 단차 기능을 가진 96-웰 써모사이클러(Applied Biosystems Veriti, 카탈로그 번호 4375786).

[1345] 데스크탑 미소원심분리 5424, 5804 (Eppendorf)

[1346] 겔 전기영동 시스템(PowerPac 기본 전원, Bio-Rad, 카탈로그 번호 164-5050, 및 Sub-Cell GT 시스템 겔 트레이, Bio-Rad, 카탈로그 번호 170-4401)

[1347] Novex XCell SureLock 미니-셀(Life Technologies, 카탈로그 번호 EI0001)

[1348] 디지털 겔 영상화 시스템(GelDoc EZ, Bio-Rad, 카탈로그 번호 170-8270, 및 청색 샘플 트레이, Bio-Rad, 카탈로그 번호 170-8273)

[1349] 청색광 트랜스일루미네이터 및 오렌지 필터 고글(SafeImager 2.0; Invitrogen, 카탈로그 번호 G6600)

[1350] 겔 정량 소프트웨어(Bio-Rad, ImageLab, GelDoc EZ과 함께 포함됨, 또는 National Institutes of Health로부터의 공개 출처 ImageJ, 웹사이트 rsbweb.nih.gov/ij/에서 이용가능) UV 분광광도계(NanoDrop 2000c, Thermo Scientific)

[1351] 시약 셋업

[1352] 트리스-보레이트 EDTA(TBE) 전기영동 용액 - 아가로스 겔 캐스팅과 겔 전기영동의 완충제로 사용하기 위해 1X 작동 용액으로 증류수 중에 TBE 완충제를 희석한다. 완충제는 적어도 1 년 동안 실온에서 보관될 수 있다(18-22 ℃).

[1353] · ATP, 10 mM - 10 mM ATP를 50 μl 알리쿼트로 나누고, 최대 1 년까지 -20℃에 보관한다; 반복적인 냉동-해동 사이클은 피한다.

[1354] · DTT, 10 mM - 증류수 중에 10 mM DTT 용액을 제조하고, 20 μl 알리쿼트로 최대 2 년 동안 -70℃에 보관한다; DTT는 쉽게 산화되므로 각 반응마다 새로운 알리쿼트를 사용한다.

[1355] · D10 배양 배지 - HEK293FT 세포의 배양을 위해서, DMEM을 1X GlutaMAX 및 10% (vol/vol) 태아 소 혈청으로 보충하여 D10 배양 배지를 제조한다. 프로토콜에 나타낸 대로, 이 배지는 또한 1X 페니실린-스트렙토마이신으로 보충될 수 있다. D10 배지는 미리 제조될 수 있으며, 최대 1 개월 동안 4℃에 보관될 수 있다.

[1356] · mTeSR1 배양 배지 - 인간 배아 줄기세포의 배양을 위해서, 5X 보충액(mTeSR1 기본 배지와 함께 포함된다), 및 100 ㎍/㎖ 노모신을 보충함으로써 mTeSR1을 제조한다.

[1357] **과정**

[1358] 표적 구성성분 설계 및 온라인 툴의 사용 · 시기 1d

[1359] 1| 표적 게놈 DNA 서열 투입. 출원인은 관심대상의 투입 서열을 취하고 적합한 표적 부위를 확인하여 순위를 정하며, 각 의도된 표적에 대해 표적외 부위를 컴퓨터로 예측하는 온라인 Cas9 표적화 설계 도구를 제공한다. 또는 달리, 어떤 5'-NGG의 바로 상류에서 20-bp 서열을 확인함으로써 가이드 서열을 손수 선택할 수 있다.

[1360] 2| 온라인 툴에 의해서 특정된 필요한 올리고머 및 프라이머 순서. 부위가 손수 선택된 다면, 올리고머 및 프라이머가 설계되어야 한다.

[1361] sgRNA 발현 작제물의 제조

[1362] 3| sgRNA 발현 작제물을 생성하기 위해서 PCR- 또는 플라스미드-기반 프로토콜이 사용될 수 있다.

[1363] (A) PCR 증폭을 통한 방법 · 시기 2h

[1364] (i) 출원인은 희석된 U6 PCR 주형을 제조한다. 출원인은 PCR 주형으로서 PX330을 추천했지만, 어떤 U6-함유 플라스미드도 PCR 주형으로서 똑같이 사용될 수 있다. 출원인은 10 ng/㎕의 농도까지 ddH_2O를 사용하여 주형을 희석했다. U6-유도된 sgRNA를 이미 함유하는 플라스미드 또는 카세트가 주형으로 사용된다면 생성물이 의도된 sgRNA만을 함유하고 주형으로부터 운반된 미량의 sgRNA는 함유하지 않도록 보장하기 위해서 겔 추출을 수행할 필요가 있다는 것을 주의한다.

[1365] (ii) 출원인은 희석된 PCR 올리고를 제조했다. U6-Fwd 및 U6-sgRNA-Rev 프라이머가 ddH_2O 중에 10 uM의 최종 농도로 희석된다 (90 ㎕ ddH_2O에 100 uM 프라이머 10㎕ 첨가).

[1366] (iii) U6-sgRNA PCR 반응. 출원인은 각 U6-sgRNA-Rev 프라이머와 필요한 마스터믹스에 따라 다음의 반응을 설정했다:

구성성분:	양 (ul)	최종 농도
허큘라제 II PCR 완충제, 5X	10	1X
dNTP, 100mM (25mM 각각)	0.5	1 mM
U6 주형 (PX330)	1	0.2 ng/ul
U6-Fwd 프라이머	1	0.2 uM
U6-sgRNA-Rev 프라이머 (가변적)	1	0.2 uM
허큘라제 II 융합 폴리머라제	0.5	
증류수	36	

[1367] 합계 50

[1368] (iv) 출원인은 다음의 사이클링 조건을 사용하여 단계 (iii)으로부터의 반응물에 대해 PCR 반응을 수행했다:

사이클 수	변성	아닐	확장
1	95℃, 2m		
2-31	95℃, 20s	60℃, 20s	72℃, 20s
32			72℃, 3m

[1369]

[1370] (v) 반응이 완료된 후, 출원인은 성공적인 단일-밴드 증폭을 검증하기 위해 겔 상에 생성물을 전개시켰다. 1X SYBR 안전 염료를 가진 1X TBE 완충제 중 2% (wt/vol) 아가로스 겔을 캐스트한다. PCR 생성물 5 ㎕를 겔에 15V cm-1에서 20-30 분간 전개시킨다. 성공적인 앰플리콘은 하나의 단일 370-bp 생성물을 제공해야 하고, 주형은 보이지 않아야 한다. PCR 앰플리콘의 겔 추출이 반드시 필요한 것은 아니다.

[1371] (vi) 출원인은 PCR 생성물을 제조자의 지시에 따라서 QIAquick PCR 정제 키트를 사용하여 정제했다. 35㎕의 완충제 EB 또는 물에서 DNA를 용출시킨다. 정제된 PCR 생성물은 4℃ 또는 -20℃에 보관될 수 있다.

[1372] **(B) Cas9-함유 바이시스트로닉 발현 벡터에 sgRNA 클로닝 · 시기 3d**

[1373] (i) *sgRNA 올리고 삽입체 제조.* 출원인은 100 uM의 최종 농도까지 각 sgRNA 디자인에 대한 올리고의 상부 및 하부 가닥을 재현탁했다. 올리고의 포스포릴화 및 아닐링은 다음과 같다:

올리고 1 (100 uM)	1 ul
올리고 2 (100 uM)	1 ul
T4 결찰 완충제, 10X	1 ul
T4 PNK	1 ul
ddH_2O	6 ul
합계	10 ul

[1374]

[1375] (ii) 다음의 변수를 사용하여 서모사이클러에서 아닐링:

37℃, 30 m

95℃, 5 m

[1376] 25° C까지 온도 감소, 분당 5° C

[1377] (iii) 출원인은 포스포릴화되고 아닐링된 올리고를 199 ㎕ 실온 ddH_2O에 1 ㎕ 올리고를 참가하여 1:200까지 희석했다.

[1378] (iv) *PX330에 sgRNA 올리고 클로닝.* 출원인은 각 sgRNA에 대해 골든 게이트 반응을 설정했다. 출원인은 또한 음성 대조군인 비삽입체 PX330의 설정을 추천했다.

PX330 (100ng)	x ul
단계 (iii)으로부터의 희석된 올리고 듀플렉스	2 ul
탄고 완충제, 10X	2 ul
DTT, 10mM	1 ul
ATP, 10mM	1 ul
FastDigest BbsI	1 ul
T7 리가아제	0.5 ul
ddH_2O	x ul
합계	20 ul

[1379]

[1380] (v) 총 1h 동안 골든 게이트 반응을 인큐베이션:

사이클 수	조건
1-6	37℃, 5m 및 21℃, 5m

[1381]

[1382] (vi) 출원인은 어떤 잔류 선형 DNA를 소화시키기 위해서 PlasmidSafe 엑소뉴클레아제로 골든 게이트 반응물을 처리했다. 이 단계는 선택적이지만 강력히 추천된다.

단계 4로부터의 골든 게이트 반응물	11 ul
10X PlasmidSafe 완충제	1.5 ul
ATP, 10 mM	1.5 ul
PlasmidSafe 엑소뉴클레아제	1 ul
합계	15 ul

[1383]

[1384] (vii) 출원인은 30 분 동안 37℃에서 PlasmidSafe 반응물을 인큐베이션하고, 이어서 30 분 동안 70℃에서 비활성화했다. **유의점:** 완료 후, 반응물은 냉동되어 이후 계속될 수 있다. 원형 DNA는 적어도 1 주 동안 안정해야 한다.

[1385] (viii) *형질전환.* 출원인은 세포와 함께 공급된 프로토콜에 따라서 컴페턴트 *E. coli*에 Plasmidsafe-처리된 플라스미드를 형질전환했다. 출원인은 신속한 형질전환을 위해 Stbl3을 추천했다. 간단히 말해서, 출원인은 단계 (vii)로부터의 생성물 5 ㎕를 20 ㎕의 얼음-냉각된 화학적 컴페턴트 Stbl3 세포에 첨가했다. 다음에, 이것은 10 m 동안 얼음 위에서 인큐베이션되고, 30 s 동안 42℃에서 열충격을 가하고, 2 m 동안 얼음에 즉시 돌려보내고, SOC 배지 100 ㎕가 첨가되고, 100 ㎍/㎖ 암피실린을 함유하는 LB 평판 위에 평판되어 37℃에서 하룻밤 인큐베이션된다.

[1386] (ix) 제2 일: 출원인은 콜로니 성장에 대해 평판을 조사했다. 전형적으로 음성 대조군 평판에는 콜로니가 없고 (BbsI-digested PX330만 결찰, 아닐링된 sgRNA 올리고 없음), PX330-sgRNA 클로닝 평판에는 수십 내지 수백 개의 콜로니가 있다.

[1387] (x) 각 평판으로부터, 출원인은 sgRNA의 정확한 삽입을 점검하기 위해 2-3 개 콜로니를 선별했다. 출원인은 멸균 피펫 팁을 사용하여 단일 콜로니를 100 ㎍/㎖ 암피실린을 가진 LB 배지의 3 ㎖ 배양물에 접종했다. 인큐베이션하고 하룻밤 37℃에서 흔들었다.

[1388] (xi) 제3 일: 출원인은 제조자의 지시에 따라서 QiAprep 스핀 미니프렙 키트를 사용하여 하룻밤 배양물로부터

플라스미드 DNA를 분리했다.

[1389] (xii) *CRISPR플라스미드 서열 검증*. 출원인은 U6-Fwd 프라이머를 사용하여 U6 프로모터로부터 시퀀싱함으로써 각 콜로니의 서열을 검증했다. 선택사항: 다음의 프라이머 표에 열거된 프라이머를 사용하여 Cas9 유전자를 시퀀싱한다.

프라이머	서열 (5'에서 3')	(서열 번호 ___)	목적
U6-For	GAGGGCCTATTTCCCATGATTCC		U6-sgRNA 증폭
U6-Rev	AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACNNNNNNNNNNNNNNNNNNNCCGGTGTTTCGTCCTTTCCACAAG		U6-sgRNA 증폭; N 은 표적의 역 보체
sgRNA-상부	CACCGNNNNNNNNNNNNNNNNNNN		PX330 에 sgRNA 클로닝
sgRNA-하부	AAACNNNNNNNNNNNNNNNNNNNC		PX330 에 sgRNA 클로닝
U6-*EMX1*-Rev	AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACCCCTAGTCATTGGAGGTGACCGGTGTTTCGTCCTTTCCACAAG		U6-EMX1 sgRNA 증폭
EMX1-상부	CACCGTCACCTCCAATGACTAGGG		PX330 에 EMX1 sgRNA 클로닝
EMX1-하부	AAACCCCTAGTCATTGGAGGTGAC		PX330 에 EMX1 sgRNA 클로닝
ssODN-센스	CAGAAGAAGAAGGGCTCCCATCACATCAACCGGTGGCGCATTGCCACGAAGCAGGCCAATGGGGAGGACATCGATGTCACCTCCAATGACAAGCTTGCTAGCGGTGGGCAACCACAAACCCACGAGGGCAGAGTGCTGCTTGCTGCTGGCCAGGCCCCTGCGTGGGCCCAAGCTGGACTCTGGCCACTCCCT		EMX1 HDR (센스; 삽입부 밑줄 표시)
ssODN-안티센스	AGGGAGTGGCCAGAGTCCAGCTTGGGCCCACGCAGGGGCCTGGCCAGCAGCAAGCAGCACTCTGCCCTCGTGGGTTTGTGGTTGCCCACCGCTAGCAAGCTTGTCATTGGAGGTGACATCGATGTCCTCCCCCATTGGCCTGCTTCGTGGCAATGCGCCACCGGTTGATGTGATGGGAGCCCTTCTTCTTCTG		EMX1 HDR (안티센스; 삽입부 밑줄표시)

[1390]

[1391]

EMX1-SURV-F	CCATCCCCTTCTGTGAATGT		EMX1 서베이어 PCR 분석, 시퀀싱
EMX1-SURV-R	GGAGATTGGAGACACGGAGA		EMX1 서베이어 PCR 분석, 시퀀싱
EMX1-HDR-F	GGCTCCCTGGGTTCAAAGTA		EMX1 RFLP PCR 분석, 시퀀싱
EMX1-HDR-R	AGAGGGGTCTGGATGTCGTAA		EMX1 RFLP PCR 분석, 시퀀싱
pUC19-F	CGCCAGGGTTTTCCCAGTCACGAC		pUC19 다중 클로닝 부위 F 프라이머, 생어 시퀀싱용

[1392] 출원인은 20 bp 가이드 서열이 U6 프로모터와 sgRNA 스캐폴드의 나머지 부분 사이에 삽입되었는지 점검하기 위해서 PX330 클로닝 벡터 서열에 대한 시퀀싱 결과를 참조했다. PX330 맵의 상세한 내용 및 서열은 GenBank 백터 맵 형식 (*.gb 파일)으로 웹사이트 crispr.genome-engineering.org에서 찾을 수 있다.

[1393] **(선택사항) ssODN 주형의 설계 · 시기 3d 먼저 계획**

[1394] 3| *ssODN 설계 및 순서*. 센스 또는 안티센스 ssODN은 공급자로부터 구입될 수 있다. 출원인은 어느 한 사이드에 적어도 40 bp의 상동체 암과 최적 HDR 효율을 위 해 90 bp를 설계할 것을 추천했다. 출원인의 경험으로는 안티센스 올리고가 약간 더 높은 변형 효율을 가진다.

[1395] 4| 출원인은 10 uM의 최종 농도까지 ssODN 울트라머를 재현탁하고 희석했다. 센스와 안티센스 ssODN은 조합하거나 아닐링하지 않는다. -20℃에 보관한다.

[1396] 5| 주: HDR 적용의 경우, 출원인은 PX330 플라스미드에 sgRNA를 클로닝할 것을 추천했다.

[1397] **sgRNA의 기능적 검증: 세포 배양물 및 트랜스펙션 · 시기 3-4d**

[1398] CRISPR-Cas 시스템은 다수의 포유류 세포주에 사용되었다. 조건은 각 세포주마다 다양할 수 있다. 아래 프로토콜은 HEK239FT 세포에 대한 트랜스펙션 조건을 상세히 설명한다. ssODN-매개 HDR 트랜스펙션의 경우 Amaxa SF 세포주 뉴클레오펙터 키트가 ssODN의 최적 전달을 위해 사용된다는 것을 유의한다. 이것은 다음 섹션에서 설명된다.

[1399] 7| *HEK293FT 유지*. 세포는 제조자의 추천에 따라서 유지된다. 간단히 말해서, 출원인은 세포를 D10 배지(10% 태아 소 혈청으로 보충된 GlutaMax DMEM)에서 37℃ 및 5% CO_2에서 배양했다.

[1400] 8| 계대를 위해, 출원인은 배지를 제거하고 세포가 손실되지 않도록 용기 사이드에 DPBS를 부드럽게 가해서 한 번 헹궜다. 출원인은 TrypLE 2 ㎖를 T75 플라스크에 첨가하고, 5 m 동안 37℃에서 인큐베이션했다. 따뜻한 D10 배지 10 ㎖를 가하여 비활성화하고 50 ㎖ 팔콘 튜브로 옮겼다. 출원인은 부드럽게 분쇄해서 세포를 해리하고, 필요에 따라 새로운 플라스크에 재파종했다. 출원인은 전형적으로 세포를 2-3 d마다 1:4 또는 1:8의 분할비로 계대하며, 절대 세포가 70%를 초과하는 컨플루언시에 도달하지 않도록 한다. 계대수 15에 도달하면 세포주가 재시작된다.

[1401] 9| *트랜스펙션을 위한 세포를 제조한다*. 출원인은 웰 당 1.3 x 10^5 세포의 파종 밀도로 트랜스펙션 16-24 h 전에 항생물질 없이 D10 배지에서 24-웰 플레이트 위에 잘-해리된 세포를 평판했으며, 파종 부피는 500 ㎕이다. 필요에 따라 제조자의 매뉴얼에 따라서 규모를 증가시키거나 감소시킨다. 추천된 밀도보다 많은 세포를 평판하는 것은 제안되지 않으며, 그렇게 하면 트랜스펙션 효율이 감소할 수 있다.

[1402] 10| 트랜스펙션 일에 세포는 70-90% 컨플루언시에서 최적이다. 세포는 제조자의 프로토콜에 따라서 리포펙타민

2000 또는 Amaxa SF 세포주 뉴클레오펙터 키트로 트랜스펙션될 수 있다.

[1403] (A) PX330에 클로닝된 sgRNA의 경우, 출원인은 서열-검증된 CRISPR 플라스미드 500ng을 트랜스펙션했다; 하나보다 많은 플라스미드를 트랜스펙션하는 경우, 등몰 비율로 혼합하고 총 500ng을 넘지 않도록 한다.

[1404] (B) PCR에 의해서 증폭된 sgRNA의 경우, 출원인은 다음을 혼합했다:

PX165 (Cas9 단독)	200 ng
sgRNA 앰플리콘 (각각)	40 ng
pUC19	최대 총 500 ng 까지 DNA 충전

[1405]

[1406] 출원인은 신뢰성 있는 정량을 위한 3 번 중복 기술로 트랜스펙션하고, 트랜스펙션 효율의 모니터링을 위해 트랜스펙션 대조군 (예를 들어, GFP 플라스미드)을 포함시킬 것을 추천했다. 게다가, PX330 클로닝 플라스미드 및/또는 sgRNA 앰플리콘이 하류 기능 분석을 위한 음성 대조군으로서 단독으로 트랜스펙션될 수 있다.

[1407] 11| HEK293FT 세포가 평판으로부터 쉽게 탈착될 수 있기 때문에 출원인은 리포펙타민 복합체를 세포에 부드럽게 가했으며, 낮은 트랜스펙션 효율이 얻어진다.

[1408] 12| 출원인은 형광 현미경을 사용하여 대조군(예를 들어, GFP) 트랜스펙션에서 형광 세포의 비율을 추산함으로써 트랜스펙션 24 h 후 효율에 대해 세포를 점검했다. 전형적으로, 70%를 초과하는 세포가 트랜스펙션된다.

[1409] 13| 출원인은 배양 배지를 따뜻한 D10 배지 500 ㎕로 더 보충했다. 세포는 쉽게 탈착될 수 있으므로 웰의 사이드에 아주 서서히 D10를 가하며, 차가운 배지를 사용하지 않는다.

[1410] 14| 삽입결실부 분석을 위한 수거 전에 트랜스펙션 후 세포가 총 48-72 h 동안 인큐베이션된다. 삽입결실부 효율은 48 h 후에는 눈에 띄게 증가하지 않는다.

[1411] **(선택사항) HR을 위한 CRISPR 플라스미드와 ssODN 또는 표적화 플라스미드의 공-트랜스펙션 · 시기 3-4d**

[1412] 15| *표적화 플라스미드의 선형화.* 가능하다면 상동체 암 중 하나 근처의 벡터 백본의 제한 부위에서 또는 어느 한 상동체 암의 원단부에서 한번 절단함으로써 표적화 벡터가 선형화된다.

[1413] 16| 출원인은 0.8-1% 아가로스 겔 상에 절단되지 않은 플라스미드와 나란히 소량의 선형화된 플라스미드를 전개시켜 성공적인 선형화를 점검했다. 선형화된 플라스미드는 수퍼코일형 플라스미드 위에 전개되어야 한다.

[1414] 17| 출원인은 QIAQuick PCR 정제 키트로 선형화된 플라스미드를 정제했다.

[1415] 18| 트랜스펙션을 위한 세포를 제조한다. 출원인은 T75 또는 T225 플라스크에서 HEK293FT를 배양했다. 트랜스펙션 1 일 전에 충분한 세포 계수가 계획된다. Amaxa 스트립-큐벳 형식의 경우, 2 x 10^6 세포가 트랜스펙션 당 사용된다.

[1416] 19| 트랜스펙션을 위한 평판을 준비한다. 출원인은 12-웰 플레이트의 각 웰에 따뜻한 D10 배지 1 ㎖를 첨가했다. 평판은 배지를 따뜻하게 유지하기 위해 인큐베이터에 놓아 둔다.

[1417] 20| 뉴틀레오펙션. 출원인은Amaxa SF 세포주 뉴클레오펙터 4D 키트 제조자의 지시에 따라서 HEK293FT 세포를 트랜스펙션했으며, 이것은 아래 단계들처럼 개조된다.

[1418] a. ssODN와 CRISPR 공-트랜스펙션을 위하여, PCR 튜브에 다음의 DNA를 예비 혼합한다:

pCRISPR 플라스미드 (Cas9 + sgRNA)	500 ng
ssODN 주형 (10uM)	1 ul

[1419]

[1420] b. HDR 표적화 플라스미드와 CRISPR 공-트랜스펙션을 위하여, PCR 튜브에 다음의 DNA를 예비 혼합한다:

CRISPR 플라스미드 (Cas9 + sgRNA)	500 ng
선형화된 표적화 플라스미드	500 ng

[1421]

[1422] 트랜스펙션 대조군에 대해서는 이전 섹션을 참조한다. 게다가, 출원인은 음성 대조군으로서 ssODN을 트랜스펙션

하거나 플라스미드만을 표적화할 것을 추천했다.

[1423] 21| 단일 세포로 해리한다. 출원인은 배지를 제거하고, 세포가 손실되지 않도록 주의하면서 DPBS로 부드럽게 한 번 헹궜다. TrypLE 2 ㎖를 T75 플라스크에 첨가하고, 5 m 동안 37℃에서 인큐베이션했다. 비활성화하기 위해 따뜻한 D10 배지 10 ㎖가 첨가되고, 50 ㎖ 팔콘 튜브에서 부드럽게 분쇄된다. 세포가 부드럽게 분쇄되어 단일 세포로 해리되는 것이 추천된다. 큰 덩어리는 트랜스펙션 효율을 감소시킬 것이다. 출원인은 현탁액으로부터 10 ㎕ 알리쿼트를 취해서 계수를 위해 D10 배지 90 ㎕로 희석했다. 출원인은 세포를 계수하고, 트랜스펙션에 필요한 세포의 수와 현탁액의 부피를 계산했다. 출원인은 전형적으로 Amaxa 뉴클레오큐벳 스트립을 사용하여 조건당 2 x 10^5 세포를 트랜스펙션했으며, 후속 피펫팅 단계에서 부피 손실을 조정하기 위해서 필요한 것보다 20% 더 많은 세포에 대해 계산할 것을 추천했다. 필요한 부피가 새로운 팔콘 튜브로 옮겨진다.

[1424] 23| 출원인은 새로운 튜브를 5 m 동안 200 x g에서 하강 회전시킨다.

[1425] 출원인은 Amaxa에 의해서 추천된 대로 SF 용액과 S1 보충물을 혼합함으로써 트랜스펙션 용액을 제조했다. Amaxa 스트립-큐벳의 경우, 총 20 ㎕의 보충된 SF 용액이 트랜스펙션 당 요구된다. 마찬가지로, 출원인은 필요한 것의 20%를 넘는 부피에 대해 계산할 것을 추천한다.

[1426] 25| 출원인은 단계 23으로부터의 펠릿화된 세포로부터 배지를 완전히 제거하고, S1-보충된 SF 용액의 적절한 부피 (2 x 10^5 세포 당 20 ㎕)에 부드럽게 재현탁했다. 오랜 시간 기간 동안 SF 용액에 세포를 남겨두지 않는다.

[1427] 26| 재현탁된 세포 20 ㎕가 단계 20으로부터의 각 DNA 프레-믹스에 피펫팅된다. 피펫으로 부드럽게 혼합해서 뉴클레오큐벳 스트립 챔버로 옮긴다. 이것은 각 트랜스펙션 조건마다 반복된다.

[1428] Amaxa에 의해서 추천된 뉴클레오펙터 4D 프로그램 CM-130을 사용하여 세포를 전기천공한다.

[1429] 28| 출원인은 각 뉴클레오큐벳 스트립 챔버에 따뜻한 D10 배지 100 ㎕를 부드럽게 서서히 피펫팅하고, 모든 부피를 단계 19로부터의 미리 가온된 평판으로 옮겼다. 주: 이 단계에서 세포는 매우 취약하며, 가혹한 피펫팅은 세포 사멸을 야기할 수 있다. 24 h 동안 인큐베이션한다. 이 지점에서, 트랜스펙션 효율은 양성 트랜스펙션 대조군에서 형광 세포의 비율로부터 추산될 수 있다. 뉴클레오펙션은 전형적으로 70 내지 80%를 초과하는 트랜스펙션 효율을 가져온다. 출원인은 세포의 손실 없이 각 웰에 따뜻한 D10 배지 1 ㎖를 서서히 첨가했다. 총 72 h 동안 세포를 인큐베이션한다.

[1430] **인간 배아 줄기세포 (HUES 9) 배양 및 트랜스펙션 · 시기 3-4d**

[1431] hESC (HUES9) 계통 유지. 출원인은 mTesR1 배지를 사용하여 피더-프리 조건에서 HUES9 세포주를 통상적으로 유지한다. 출원인은 기본 배지에 포함된 5X 보충물과 100 ㎍/㎖ 노모신을 첨가함으로써 TeSR1 배지를 제조했다. 출원인은 10 uM Rock 저해제로 더 보충된 mTeSR1 배지의 10㎖ 알리쿼트를 제조했다. 조직 배양 평판을 코팅한다. 냉각된 DMEM 중에 냉각된 GelTrex을 1:100으로 희석하고, 100 mm 조직 배양 평판의 전체 표면을 코팅한다.

[1432] 37℃에서 적어도 30 m 동안 인큐베이터 안에 평판을 놓아 둔다. 15 ㎖ 팔콘 튜브에 37℃에서 세포 바이알을 해동하고, mTeSR1 배지 5 ㎖를 첨가하고,5 m 동안 200 x g에서 펠릿화한다. GelTrex 코팅을 흡인하여 없애고, Rock 저해제 를 함유하는 10 ㎖ mTeSR1 배지와 함께 약 1 x 106 세포를 파종한다. 트랜스펙션 24 h 후 정규 mTeSR1 배지로 교환하고 매일 재공급한다. 세포를 계대한다. 세포에 신선한 mTeSR1 배지를 매일 재공급하고, 70% 컨플루언시에 도달하기 전에 계대한다. mTeSR1 배지를 흡인하여 없애고, 세포를 DPBS로 한번 세척한다. 2 ㎖ 아큐타제에 의해서 세포를 해리하고, 3-5 m 동안 37℃에서 인큐베이션한다. 탈착된 세포에 10 ㎖ mTeSR1 배지를 첨가하고, 15 ㎖ 팔콘 튜브로 옮기고, 부드럽게 재현탁한다. 10 uM Rock 저해제 와 함께 mTeSR1 배지에서 GelTrex-코팅된 평판 위에 재평판한다. 평판 24 h 후 정규 mTeSR1 배지로 교환한다.

[1433] 트랜스펙션. 출원인은 Amaxa P3 1차 세포 4-D 뉴클레오펙터 키트(Lonza)를 사용하여 트랜스펙션 전 해동 후 적어도 1 주 동안 세포를 배양할 것을 추천했다. 트랜스펙션 2 h 전에 신선한 배지를 로그상 성장중인 세포에 재공급한다. 아큐타제로 부드럽게 재현탁해서 단일 세포 또는 10 세포 미만의 작은 클러스터로 해리한다. 뉴클레오펙션에 필요한 세포수를 계수하고, 5 m 동안 200 x g에서 하강 회전시킨다. 배지를 완전히 제거하고, S1-보충된 P3 뉴클레오펙션 용액의 추천된 부피에 재현탁한다. 전기천공된 세포를 1X Rock 저해제 의 존재하에 코팅된 평판에 부드럽게 평판한다.

[1434] 트랜스펙션 성공을 점검하고, 뉴클레오펙션 24 h 후 시작해서 mTeSR1 배지를 규칙적으로 매일 재공급한다. 전형

적으로, 출원인은 Amaxa 뉴클레오펙션으로 70%를 초과하는 트랜스펙션 효율을 관찰한다. DNA를 수거한다. 트랜스펙션 48-72 h 후, 세포를 아큐타제를 사용하여 해리하고, 5x 부피의 mTeSR1을 첨가함으로써 비활성화한다. 세포를 5 m 동안 200 x g에서 하강 회전시킨다. 펠릿화된 세포는 QuickExtract 용액으로 DNA 추출을 위해 직접 가공될 수 있다. 아큐타제 없이 기계적으로 세포를 해리하지 않는 것을 추천한다. 비활성화 아큐타제 없이 추천된 속도 이상으로 세포를 하강 회전시키지 않는 것을 추천하며, 그렇게 하면 세포 용해를 야기할 수 있다.

[1435] **FACS에 의한 클론성 세포주의 분리. 시기 · 2-3h 수동; 2-3주 확장**

[1436] FACS 또는 연속 희석에 의해서 트랜스펙션 24 h 후 클론 분리가 수행될 수 있다.

[1437] 54| *FACS 완충제 제조.* 착색된 형광을 사용하여 정렬될 필요가 없는 세포는 1X 페니실린/스트렙토마이신으로 보충된 정규 D10 배지에서 정렬될 수 있다. 착색된 형광 정렬이 또한 요구된다면, 페놀-무함유 DMEM 또는 DPBS가 정규 DMEM 대신 치환된다. 1X 페니실린/스트렙토마이신으로 보충하고, .22um Steriflip 필터를 통해 여과한다.

[1438] 55| *96웰 플레이트 준비.* 출원인은 웰당 1X 페니실린/스트렙토마이신으로 보충된 D10 배지 100 ㎕를 첨가하고, 원하는 수의 클론에 대해 필요에 따라 다수의 플레이트를 준비했다.

[1439] 56| *FACS를 위한 세포 제조.* 출원인은 배지를 완전히 흡인해서 세포를 해리하고, 24웰 플레이트의 웰당 100 ㎕ TrypLE를 첨가했다. 5 분 동안 인큐베이션하고, 따뜻한 D10 배지 400 ㎕를 첨가한다.

[1440] 57| 재현탁된 세포를 15 ㎖ 팔콘 튜브로 옮겨서 20 회 부드럽게 분쇄했다. 단일 세포의 해리를 보장하기 위해서 현미경 아래서 조사할 것이 추천된다.

[1441] 58| 5 분 동안 200 x g에서 세포를 하강 회전시켰다.

[1442] 59| 출원인은 배지를 흡인했고, FACS 배지 200 ㎕에 세포를 재현탁했다.

[1443] 60| 세포가 표지된 FACS 튜브에 35 um 메시 필터를 통해 여과된다. 출원인은 세포 스트레이터 캡을 가진 BD 팔콘 12 x 75 mm 튜브를 사용할 것을 추천했다. 정렬할 때까지 세포는 얼음 위에 놓아 둔다.

[1444] 61| 출원인은 단계 55로부터 준비된 96웰 플레이트에 단일 세포를 정렬시켰다. 출원인은 각 플레이트에서 하나의 지정된 웰에 양성 대조군으로서 100 세포를 정렬할 것을 추천했다.

[1445] 주: 나머지 세포는 유지될 수 있고, 전체 변형 효율을 가늠하기 위한 집단 수준의 제토타이핑에 사용될 수 있다.

[1446] 62| 출원인은 세포를 인큐베이터로 돌려보내고, 이들을 2-3 주 동안 확장시켰다. 따뜻한 D10 배지 100 ㎕가 정렬 5 d 후에 첨가된다. 필요에 따라 3-5 d마다 배지 100 ㎕를 교환한다.

[1447] 63| 정렬 1주 후 콜로니가 "클론" 출현에 대해 조사된다: 원형 콜로니가 중심점으로부터 방사상으로 형성되었다. 빈 웰과 이중 또는 다중 파종된 웰을 표시한다.

[1448] 64| 세포가 60% 초과의 컨플루언트에 도달했을 때, 출원인은 계대를 위하여 한 세트의 복제 평판을 준비했다. D10 배지 100 ㎕이 복제 평판의 각 웰에 첨가된다. 출원인은 20 회 격렬히 하방 피펫팅하여 세포를 직접 해리시켰다. 클론 계통을 유지하기 위하여 현탁된 부피의 20%가 준비된 복제 평판에 평판되었다. 이후 2-3 d마다 배지를 교환하고 그에 따라 계대한다.

[1449] 65| 세포의 나머지 80%를 DNA 분리 및 제노타이핑에 사용한다.

[1450] **선택사항: 희석에 의한 클론성 세포주의 분리. 시기 · 2-3h 수동; 2-3주 확장**

[1451] 66| 출원인은 상기 설명된 대로 24웰 플레이트로부터의 세포를 해리했다. 단일 세포로 해리되도록 확실히 했다. 세포가 뭉치는 것을 방지하기 위해서 세포 스트레이너가 사용될 수 있다.

[1452] 67| 각 조건에서 세포의 수가 계수된다. 100 ㎕ 당 0.5 세포의 최종 농도로 D10 배지에서 각 조건을 연속 희석한다. 각 96웰 플레이트에 대해, 출원인은 D10 12 ㎖에 60 세포의 최종 계수로 희석할 것을 추천했다. 적합한 클론 희석을 위해서 세포수의 정확한 계수가 권장된다. 세포는 정확성을 보장하기 위해서 중간 연속 희석 단계에서 다시 계수될 수 있다.

[1453] 68| 멀티채널 피펫을 사용하여 희석된 세포 100 ㎕가 96웰 플레이트의 각 웰에 피펫팅되었다.

[1454] 주: 나머지 세포는 유지될 수 있고, 전체 변형 효율을 가늠하기 위한 집단 수준의 제토타이핑에 사용될 수

있다.

[1455] 69| 출원인은 콜로니를 평판 약 1 주 후에 "클론성"에 대해 조사했다: 원형 콜로니가 중심점으로부터 방사상으로 형성되었다. 이중 또는 다중 파종된 웰을 표시한다.

[1456] 70| 출원인은 세포를 인큐베이터로 돌려보내고, 이들을 2-3주 동안 확장시켰다. 필요에 따라 앞의 섹션에서 상세히 설명한 대로 세포를 다시 공급했다.

[1457] **CRISPR 절단 효율에 대한 서베이어 분석. 시기 · 5-6h**

[1458] 트랜스펙션된 세포의 절단 효율을 분석하기 전에, 출원인은 아래 설명된 프로토콜을 사용하여 서베이어 뉴클레아제 소화 단계를 통해 음성 (트랜스펙션되지 않은) 대조군 샘플에 대해 각 새로운 서베이어 프라이머를 시험할 것을 추천했다. 때로, 단일 밴드 클린 서베이어 PCR 생성물조차도 비-특이적 서베이어 뉴클레아제 절단 밴드를 제공할 수 있으며, 잠재적으로 정확한 삽입결실부 분석을 방해할 수 있다.

[1459] 71| *DNA를 위한 세포 수거.* 세포를 해리하고 5 m 동안 200 x g에서 하강 회전시킨다. 주: 이 단계에서 필요에 따라 평판을 복제하여 트랜스펙션된 세포주를 유지한다.

[1460] 72| 상청액을 완전히 흡인한다.

[1461] 73| 출원인은 제조자의 지시에 따라서 QuickExtract DNA 추출 용액을 사용했다. 출원인은 전형적으로 24웰 플레이트의 각 웰에 50 ㎕의 용액을 사용했고, 96웰 플레이트에는 10 ㎕의 용액을 사용했다.

[1462] 74| 출원인은 추출된 DNA를 최종 농도 100-200 ng/㎕까지 ddH_2O로 표준화했다. **유의점:** 추출된 DNA는 수 개월 동안 -20℃에 보관될 수 있다.

[1463] 75| *서베이어 PCR 설정.* 출원인의 온라인/컴퓨터 알고리즘 툴에 의해서 제공된 서베이어 프라이머를 사용하여 다음의 마스터 믹스를 제조한다:

구성성분:	양 (ul)	최종 농도
허큘라제 II PCR 완충제, 5X	10	1X
dNTP, 100mM (25mM 각각)	1	1 mM
서베이어 Fwd 프라이머 (10uM)	1	0.2 uM
서베이어 Rev 프라이머 (10uM)	1	0.2 uM
허큘라제 II 융합 폴리머라제	1	
$MgCl_2$(25mM)	2	1 mM
증류수	33	
합계	49 (각 반응에 대해)	

[1464]

[1465] 76| 출원인은 각 반응에 단계 74로부터의 표준화된 게놈 DNA 주형 100-200 ng을 첨가했다.

[1466] 77| PCR 반응이 다음의 사이클링 조건을 사용하여 30 증폭 사이클을 초과하여 수행되었다:

사이클 수	변성	아닐	확장
1	95℃, 2m		
2-31	95℃, 20s	60℃, 20s	72℃, 30s
32			72℃, 3m

[1467]

[1468] 78| 출원인은 단일 밴드 생성물을 점검하기 위해 2-5 ㎕의 PCR 생성물을 1% 겔 상에 전개시켰다. 이들 PCR 조건

은 서베이어 프라이머의 대부분의 쌍에서 작동하도록 설계되지만, 일부 프라이머는 주형 농도, $MgCl_2$ 농도 및/또는 아닐링 온도를 조정함에 의한 추가의 최적화를 필요로 할 수 있다.

[1469] 79| 출원인은 PCR 반응물을 QIAQuick PCR 정제 키트를 사용하여 정제하고, 용출액을 20ng/㎕로 표준화했다. **유의점:** 정제된 PCR 생성물은 -20℃에 보관될 수 있다.

[1470] 80| *DNA 헤테로듀플렉스 형성.* 아닐링 반응이 다음과 같이 설정되었다:

Taq PCR 완충제, 10X	2 ul
표준화된 DNA (20 ng/ul)	18 ul

[1471] 합계 부피 20 ul

[1472] 81| 다음의 조건을 사용한 반응물 아닐링:

사이클 수	조건
1	95℃, 10 mn
2	95℃ 내지 85℃, -2℃/s
3	85℃, 1 분
4	85℃ 내지 75℃, -0.3℃/s
5	75℃, 1 분
6	75℃ 내지 65℃, -0.3℃/s
7	65℃, 1 분
8	65℃ 내지 55℃, -0.3℃/s
9	55℃, 1 분
10	55℃ 내지 45℃, -0.3℃/s
11	45℃, 1 분
12	45℃ 내지 35℃, -0.3℃/s
13	35℃, 1 분
14	35℃ 내지 25℃, -0.3℃/s
15	25℃, 1 분

[1473]

[1474] 82| *서베이어 뉴클레아제 S 소화.* 출원인은 마스터-믹스를 제조하고, 다음의 구성성분을 얼음 위에서 총 25 ㎕의 최종 부피가 되도록 단계 81로부터의 아닐링된 헤테로듀플렉스에 첨가했다:

구성성분	양 (ul)	최종 농도
$MgCl_2$ 용액, 0.15M	2.5	15mM
ddH_2O	0.5	
서베이어 뉴클레아제 S	1	1X
서베이어 인핸서 S	1	1X

[1475] 합계 5

[1476] 83| 웰을 와동시키고 하강 회전시킨다. 반응물을 42℃에서 1 h 동안 인큐베이션한다.

[1477] 84| 선택사항: 서베이어 키트로부터 2 ㎕의 중단 용액이 첨가될 수 있다. **유의점:** 소화된 PCR 생성물은 다음번

분석을 위해 -20℃에 보관될 수 있다.

[1478] 85| *서베이어반응물의 시각화*. 서베이어 뉴클레아제 소화 생성물은 2% 아가로스 겔에서 시각화될 수 있다. 보다 나은 분해능을 위해서, 생성물은 4-20% 구배 폴리아크릴아미드 TBE 겔 상에서 전개될 수 있다. 출원인은 10 ㎕ 생성물을 추천된 로딩 완충제와 함께 로딩하고, 제조자의 지시에 따라서 겔을 전개시켰다. 전형적으로, 출원인은 브로모페놀 청색 염료가 겔의 바닥으로 이동할 때까지 전개시켰다. 동일한 겔에 DNA 래더와 음성 대조군을 포함시킨다.

[1479] 86| 출원인은 겔을 TBE에 희석된 1X SYBR 골드 염료로 겔을 염색했다. 겔을 15 m 동안 부드럽게 흔들었다.

[1480] 87| 출원인은 밴드의 과다노출 없이 정량적 영상화 시스템을 사용하여 겔을 이미지화했다. 음성 대조군은 PCR 생성물의 크기에 상응하는 단지 하나의 밴드만을 가져야 하지만, 때로 다른 크기의 비-특이적 절단 밴드를 가질 수도 있다. 이들은 표적 절단 밴드와 크기가 상이한 경우 분석을 방해하지 않을 것이다. 출원인의 온라인/컴퓨터 알고리즘 툴에 의해서 제공되는 표적 절단 밴드 크기의 합계는 PCR 생성물의 크기와 동일해야 한다.

[1481] 88| *절단 강도 추산*. 출원인은 ImageJ 또는 다른 겔 정량 소프트웨어를 사용하여 각 밴드의 적분된 강도를 정량했다.

[1482] 89| 각 레인에 대해, 출원인은 다음의 식을 사용하여 PCR 생성물 절단 분획 (f_{cut})을 계산했다: $f_{cut} = (b + c) / (a + b + c)$, 여기서 a는 소화되지 않은 PCR 생성물의 적분된 강도이고, b 및 c는 각 절단 생성물의 적분된 강도이다. 90| 절단 효율은 듀플렉스 형성의 이항식 확률 분포에 기초한 다음 식을 사용하여 추산될 수 있다:

[1483] 91| $$indel\ (\%) = 100 \times (1 - \sqrt{(1 - f_{cut})})$$

[1484] <u>CRISPR 절단 효율을 평가하기 위한 Sanger 시퀀싱. 시기 · 3d</u>

[1485] 초기 단계는 서베이어 분석의 단계 71-79와 동일하다. 주: 서베이어 프라이머는 적절한 제한 부위가 정방향 및 역방향 프라이머에 덧붙은 경우 Sanger 시퀀싱에 사용될 수 있다. 추천된 pUC19 백본에 클로닝하는 경우, EcoRI이 Fwd 프라이머로, HindIII이 Rev 프라이머로 사용될 수 있다.

[1486] 92| *앰플리콘 소화*. 소화 반응을 다음과 같이 설정한다:

[1487]

구성성분	양 (ul)
FastDigest 완충제, 10X	3
FastDigest EcoRI	1
FastDigest HindIII	1
표준화된 DNA (20 ng/ul)	10
ddH_2O	15
합계 부피	30

[1488] 93| pUC19 백본 소화. 소화 반응을 다음과 같이 설정한다:

구성성분	양 (ul)
FastDigest 완충제, 10X	3
FastDigest EcoRI	1
FastDigest HindIII	1
FastAP 알칼리성 포스파타제	1
pUC19 벡터 (200 ng/ul)	5
ddH_2O	20
합계 부피	30

[1489]

[1490] 94| 출원인은 소화 반응물을 QIAQuick PCR 정제 키트를 사용하여 정제했다. **유의점:** 정제된 PCR 생성물은 -20℃에 보관될 수 있다.

[1491] 95| 출원인은 소화된 pUC19 백본과 Sanger 앰플리콘을 다음과 같이 1:3 벡터:삽입체의 비율로 결찰했다:

구성성분	양 (ul)
소화된 pUC19	x (50 ng)
소화된 삽입체	x (1:3 벡터:삽입체 몰비)
T7 리가아제	1
2X 신속 결찰 완충제	10
ddH_2O	x
합계 부피	20

[1492]

[1493] 96| *형질전환.* 출원인은 컴패턴트 이콜라이 균주에 PlasmidSafe-처리된 플라스미드를 세포와 함께 공급된 프로토콜에 따라서 형질전환했다. 출원인은 신속한 형질전환을 위해 Stbl3을 추천했다. 간단히 말해서, 단계 95로부터의 생성물 5 ㎕를 얼음 냉각된 화학적 컴패턴트 Stbl3 세포 20 ㎕에 첨가하고, 10 분 동안 얼음 위에서 인큐베이션하고, 30 s 동안 42℃에서 열충격을 가하고, 곧바로 2 m 동안 얼음에 돌려보내고, 100 ㎕ SOC 배지를 첨가하고, 100 ㎍/㎖ 암피실린을 함유하는 LB 평판 위에 평판했다. 이것은 37℃에서 하룻밤 인큐베이션된다.

[1494] 97| 제2 일: 출원인은 콜로니 성장에 대해 평판을 조사했다. 전형적으로, 음성 대조군 평판에는 콜로니가 없으며(EcoRI-HindIII 소화된 pUC19만의 결찰, Sanger 앰플리콘 삽입체 없음), pUC19-Sanger 앰플리콘 클로닝 평판에는 수백 개의 콜로니가 나타나는 경향이 있다.

[1495] 98| 제3 일: 출원인은 제조자의 지시에 따라서 QIAprep 스핀 미니프렙 키트를 사용하여 하룻밤 배양물로부터 플라스미드 DNA를 분리했다.

[1496] 99| *Sanger 시퀀싱.* 출원인은 pUC19-For 프라이머를 사용하여 pUC19 백본으로부터 시퀀싱에 의해서 각 콜로니의 서열을 검증했다. 출원인은 Cas9-유도된 NHEJ 돌연변이의 존재를 점검하기 위해서 예상된 게놈 DNA에 대한 시퀀싱 결과를 참조했다. 편집 효율 % = (# 변형된 클론)/(# 합계 클론). 정확한 변형 효율을 생성하기 위해서는 합당한 수의 클론 (>24)을 선별하는 것이 중요하다.

[1497] **미소결실에 대한 제노타이핑. 시기 · 2-3d 수동; 2-3 주 확장**

[1498] 100| 세포가 결실될 영역을 표적화하는 한 쌍의 sgRNA로 상기 설명된 대로 트랜스펙션되었다.

[1499] 101| 트랜스펙션 24 h 후, 클론 계통이 상기 설명된 대로 FACS 또는 연속 희석에 의해서 분리된다.

[1500] 102| 세포가 2-3 주 동안 확장된다.

[1501] 103| 출원인은 10㎕ QuickExtract 용액을 사용하여 상기 설명된 대로 클론 계통으로부터 DNA를 수거하고, 50-100 ng/㎕의 최종 농도까지 ddH_2O로 게놈 DNA를 표준화했다.

[1502] 104| *변형된 영역의 PCR 증폭*. PCR 반응이 다음과 같이 설정된다:

구성성분:	양 (ul)	최종 농도
허큘라제 II PCR 완충제, 5X 10	1X	
dNTP, 100 mM (25 mM 각각) 1	1 mM	
Out Fwd 프라이머 (10 uM) 1	0.2 uM	
Out Rev 프라이머 (10 uM) 1	0.2 uM	
허큘라제 II 융합 폴리머라제	1	
$MgCl_2$(25 mM)	2	1 mM
ddH_2O	32	
합계	48 (각 반응에 대해)	

[1503]

[1504] 주: 결실 크기가 1 kb를 초과하면, wt 대립유전자의 존재를 스크리닝하기 위해서 In-Fwd 및 In-Rev 프라이머로 병행 세트의 PCR 반응을 설정한다.

[1505] 105| 반전을 스크리닝하기 위해서 PCR 반응이 다음과 같이 설정된다:

구성성분:	양 (ul)	최종 농도
허큘라제 II PCR 완충제, 5X	10	1X
dNTP, 100 mM (25 mM 각각)	1	1 mM
Out Fwd 또는 Out-Rev 프라이머 (10 uM)	1	0.2 uM
In Fwd 또는 In-Rev 프라이머 (10 uM)	1	0.2 uM
허큘라제 II 융합 폴리머라제	1	
$MgCl_2$(25 mM)	2	1 mM
ddH_2O	32	
합계	48 (각 반응에 대해)	

[1506]

[1507] 주: 프라이머는 Out-Fwd + In Fwd, 또는 Out-Rev + In-Rev로서 쌍을 이룬다.

[1508] 106| 출원인은 각 반응에 단계 103으로부터의 표준화된 게놈 DNA 100-200 ng을 첨가했다.

[1509] 107| PCR 반응이 다음의 사이클링 조건을 사용하여 수행되었다:

사이클 수	변성	아닐	확장
1	95℃, 2m		
2-31	95℃, 20s	60℃, 20s	72℃, 30s
32			72℃, 3m

[1510]

[1511] 108| 출원인은 생성물의 점검을 위해 1-2% 겔 상에 PCR 생성물 2-5 $\mu\ell$를 전개시켰다. 이들 PCR 조건은 대부분의 프라이머에서 작동하도록 설계되지만, 일부 프라이머는 주형 농도, $MgCl_2$ 농도 및/또는 아닐링 온도를 조정함에 의한 추가의 최적화를 필요로 할 수 있다.

[1512] **HDR을 통한 표적화된 편형을 위한 제노타이핑. 시기 · 2-3d, 2-3h 수동**

[1513] 109| 출원인은 QuickExtract 용액을 사용하여 상기 설명된 대로 DNA를 수거하고, 100-200 ng/$\mu\ell$의 최종 농도까지 TE로 게놈 DNA를 표준화했다.

[1514] 110| *변형된 영역의 PCR 증폭*. PCR 반응이 다음과 같이 설정된다:

구성성분:	양 (ul)	최종 농도
허큘라제 II PCR 완충제, 5X	10	1X
dNTP, 100 mM (25 mM 각각)	1	1 mM
HDR Fwd 프라이머 (10 uM)	1	0.2 uM
HDR Rev 프라이머 (10 uM)	1	0.2 uM
허큘라제 II 융합 폴리머라제	1	
$MgCl_2$(25mM)	2	1 mM
ddH_2O	33	
합계	49 (각 반응에 대해)	

[1515]

[1516] 111| 출원인은 각 반응에 대해 단계 109로부터의 게놈 DNA 주형 100-200 ng을 첨가하고, 다음의 프로그램을 수행했다.

사이클 수	변성	아닐	확장
1	95℃, 2m		
2-31	95℃, 20s	60℃, 20 s	72℃, *kb 당 30-60s*
32			72℃, 3m

[1517]

[1518] 112| 출원인은 5 $\mu\ell$의 PCR 생성물을 0.8-1% 겔 상에서 전개시켜 단일 밴드 생성물을 점검했다. 프라이머는 주형 농도, $MgCl_2$ 농도 및/또는 아닐링 온도를 조정함에 의한 추가의 최적화를 필요로 할 수 있다.

[1519] 113| 출원인은 PCR 반응물을 QIAQuick PCR 정제 키트를 사용하여 정제했다.

[1520] 114| HDR 실시예에서, *Hind*III 제한 부위가 EMX1 유전자에 삽입된다.

[1521] 들은 PCR 앰플리콘의 제한 소화에 의해서 검출된다:

구성성분	양 (ul)
정제된 PCR 앰플리콘 (200-300ng)	x
F.D. 완충제, 그린	1
HindIII	0.5
ddH_2O	x
합계	10

[1522]

[1523] i. DNA는 37℃에서 10m 동안 소화된다:

[1524] ii. 출원인은 소화된 생성물 10 ㎕를 로딩 염료와 함께 4-20% 구배 폴리아크릴아미드 TBE 겔 상에서 자일렌 시아놀 밴드가 겔의 바닥으로 이동할 때까지 전개시켰다.

[1525] iii. 출원인은 겔을 1X SYBR 골드 염료로 15 m 동안 흔들면서 염색했다.

[1526] iv. 절단 생성물은 서베이어 분석 섹션에서 상기 설명된 대로 이미지화되고 정량된다. HDR 효율이 식 (b + c)/(a + b + c)에 의해서 추산되며, 여기서 a는 소화되지 않은 HDR PCR 생성물의 적분된 강도이고, b 및 c는 HindIII-절단 단편의 적분된 강도이다.

[1527] 115| 또는 달리, Sanger 시퀀싱 또는 NGS를 사용하여 단계 113으로부터의 정제된 PCR 앰플리콘이 클로닝되고 제노타이핑될 수 있다.

[1528] **심층 시퀀싱 및 표적외 분석 · 시기 제1일 내지 제2일**

[1529] 온라인 CRISPR 표적 설계 도구는 각 확인된 표적 부위에 대해 후보 게놈 표적외 부위를 생성한다. 이들 부위에서 표적외 분석은 서베이어 뉴클레아제 분석, Sanger 시퀀싱, 또는 차세대 심층 시퀀싱에 의해서 수행될 수 있다. 이들 부위의 대부분에서 낮거나 검출불가능한 변형 비율의 가능성이 주어진다면, 출원인은 높은 감도 및 정확도를 위해서 Illumina Miseq 플랫폼에 의한 심층 시퀀싱을 추천했다. 프로토콜은 시퀀싱 플랫폼에 따라 변할 것이며; 여기서 출원인은 간단히 시퀀싱 어댑터를 부착하기 위한 융합 PCR 방법을 설명한다.

[1530] 116| *프라이머 설계 심층 시퀀싱.* 차세대 시퀀싱 (NGS) 프라이머가 전형적으로 100-200 bp 크기 범위의 더 짧은 앰플리콘에 대해 설계된다. 프라이머는 NCBI 프라이머-블라스트를 사용하여 수동으로 설계되거나, 또는 온라인 CRISPR 표적 설계 도구로 생성될 수 있다(웹사이트 genome-engineering.org/tools).

[1531] 117| Cas9-표적화된 세포로부터 게놈 DNA 수거. QuickExtract 게놈 DNA를 ddH_2O로 100-200 ng/㎕까지 표준화한다.

[1532] 118| *초기 라이브러리 제조 PCR.* 단계 116으로부터의 NGS 프라이머를 사용하여 초기 라이브러리 제조 PCR을 준비한다.

구성성분:	양 (ul)	최종 농도
허큘라제 II PCR 완충제, 5X	10	1X
dNTP, 100 mM (25 mM 각각)	1	1 mM
NGS Fwd 프라이머 (10 uM)	1	0.2 uM
NGS Rev 프라이머 (10 uM)	1	0.2 uM
허큘라제 II 융합 폴리머라제	1	
$MgCl_2$(25 mM)	2	1 mM
ddH20	33	
합계	49 (각 반응에 대해)	

[1533]

[1534] 119| 각 반응에 표준화된 게놈 DNA 주형 100-200 ng을 첨가한다.

[1535] 120| 20 증폭 사이클을 넘지 않게 다음의 사이클링 조건을 사용하여 PCR 반응을 수행한다:

사이클 수	변성	아닐	확장
1	95℃, 2m		
2-21	95℃, 20s	60℃, 20s	72℃, 15s
22			72℃, 3m

[1536]

[1537] 121| 단일 밴드 생성물을 점검하기 위해 1% 겔 상에 PCR 생성물 2-5 ㎕를 전개시킨다. 모든 게놈 DNA PCR과 마찬가지로, NGS 프라이머는 주형 농도, $MgCl_2$ 농도 및/또는 아닐링 온도를 조정함에 의한 추가의 최적화를 필요로 할 수 있다.

[1538] 122| QIAQuick PCR 정제 키트를 사용하여 PCR 반응물을 정제하고, 용출액을 20 ng/㎕까지 표준화한다. **유의점:** 정제된 PCR 생성물은 -20℃에 보관될 수 있다.

[1539] 123| *Nextera XT DNA 샘플 제조 키트*. 제조자의 프로토콜에 따라서, 각 샘플에 대해 특유의 바코드를 가진 Miseq 시퀀싱-즉시이용 라이브러리를 생성한다.

[1540] 124| *시퀀싱 데이터 분석*. 표적외 분석이 ClustalW, Geneious, 또는 간단한 서열 분석 스트립트와 같은 판독 정렬 프로그램을 통해서 수행될 수 있다.

[1541] 시기

[1542] 단계 1 - 2 sgRNA 올리고 및 ssODN의 설계 및 합성: 1-5 d, 공급자에 따라 가변적

[1543] 단계 3 - 5 CRISPR 플라스미드 또는 PCR 발현 카세트의 구성: 2 h 내지 3 d

[1544] 단계 6 - 53 세포주에 트랜스펙션: 3 d (1 h 체험 시간)

[1545] 단계 54 - 70 클론 계통의 선택적 유도: 1-3 주, 세포 타입에 따라 가변적

[1546] 단계 71 - 91 서베이어를 통한 NHEJ의 기능적 검증: 5-6 h

[1547] 단계 92 - 124 Sanger 또는 차세대 심층 시퀀싱을 통한 제노타이핑: 2-3 d (3-4 h 체험 시간)

[1548] **실시예와 관련한 상황 취급**

상황	해결
sgRNA 미증폭	U6-주형 농도 적정
서베이어 또는 HDR PCR 오염 또는 미증폭	$MgCl_2$ 적정; 주형 농도 표준화 및 적정; 아닐링 온도 구배; 프라이머 재설계
미소결실 PCR에서 대립유전자의 불균등한 증폭	야생형 및 결실 대립유전자를 검출하기 위한 별도의 PCR 설정; 유사한 크기의 앰플리콘으로 프라이머 재설계
음성 대조군 평판 상의 콜로니	*Bbs*I 증가; 골든 게이트 반응 사이클 수 증가, 극성 포스페이트 처리와는 별도로 PX330 절단
sgRNA 서열 없음 또는 잘못된 서열	추가 콜로니 스크리닝
낮은 리포펙타민 트랜스펙션 효율	세포 건강 및 밀도 점검; DNA 적정; GFP 트랜스펙션 대조군 첨가
낮은 뉴클레오펙션 트랜스펙션 효율	세포 건강 및 밀도 점검; DNA 적정; 단일 세포로 현탁
FACS 후 덩어리 또는 세포 없음	FACS 전에 세포 여과; 단일 세포로 해리; 적절한 밀도로 재현탁
연속 희석후 덩어리 또는 세포 없음	세포 재계수; 단일 세포로 해리 및 스트레이너를 통한 여과; 연속 희석 점검
음성 샘플에서 높은 서베이어 바탕값	상이한 위치에서 프라이밍하도록 프라이머 재설계
겔 상에 오염된 서베이어 결과	PCR 생성물 정제; 투입 DNA 감소; 42° C 인큐베이션 30m 까지 감소
서베이어 절단 없음	PCR 생성물 정제 및 표준화; TaqB 완충제로 재아닐링; sgRNA 재설계; px330 백본에서 Cas9 서열 검증
TBE 아크릴아미드 겔에서 샘플이 가라앉지 않음	15mM의 최종 농도로 $MgCl_2$로 보충 또는 글리세롤 함유 로딩 완충제 첨가

[1549]

[1550] **고찰**

[1551] CRISPR-Cas는 몇 개 유전자의 동시 변형을 촉진하고 염색체 미소결실을 높은 효율로 매개하기 위해서 쉽게 복합화될 수 있다. 출원인은 HEK293FT 세포에서 최대 68%의 효율로 인간 *GRIN2B* 및 *DYRK1A* 유전자좌의 동시 표적화를 증명하기 위하여 2 개의 sgRNA를 사용했다. 마찬가지로, 한 쌍의 sgRNA가 엑손의 삭제와 같은 미소결실을 매개하기 위해 사용될 수 있으며, 이것은 클론 수준으로 PCR에 의해서 제노타이핑될 수 있다. 엑손 접합부의 정확한 위치는 변할 수 있다는 것을 유의한다. 출원인은 또한 HEK293 및 HUES9 세포에서 Cas9의 야생형과 닉카아제 돌연변이를 모두 가진 HDR을 매개하기 위한 ssODN 및 표적화 벡터의 사용을 증명했다(도 30). 출원인은 Cas9 닉카아제를 사용하여 HUES9 세포에서 HDR을 검출할 수 없었다는 것을 유의하며, 이것은 HUES9 세포에서 낮은 효율 또는 수선 활성의 잠재적 차이로 인한 것일 수 있다. 이들 값은 전형적이지만, 주어진 sgRNA의 절단 효율에 일부 가변성이 있으며, 특정 sgRNA가 아직 밝혀지지 않은 이유로 작동하지 않는 경우는 거의 없다. 출원인은 각 유전자좌에 대해 2개의 sgRNA를 설계하고, 의도된 세포 타입에서 이들의 효율을 시험할 것을 추천했다.

[1552] *실시예 31: NLS*

[1553] Cas9 전사 모듈레이터: 출원인은 Cas9/gRNA CRISPR 시스템을 DNA 절단을 넘어서는 기능이 실행될 수 있는 일반적인 DNA 결합 시스템으로 바꾸는 것을 제시했다. 예를 들어, 촉매 비활성 Cas9 위에 기능적 도메인(들)을 융합

함으로써, 출원인은 전사 활성화/억제, 메틸화/탈메틸화, 또는 크로마틴 변형과 같은 새로운 기능을 부여했다. 이 목표를 달성하기 위해서, 출원인은 뉴클레아제 활성에 필수적인 두 잔기, 즉 D10과 H840을 알라닌으로 바꿈으로써 촉매 비활성 Cas9 돌연변이를 제작했다. 이들 두 잔기를 돌연변이함으로써, 표적 DNA와 결합하는 능력을 유지한 채로 Cas9의 뉴클레아제 활성이 폐기된다. 출원인이 출원인의 가설을 시험하는데 중요하다고 결정한 기능적 도메인은 전사 활성인자 VP64와 전사 억제인자 SID 및 KRAB이다.

[1554] Cas9 핵 위치: 출원인은 가장 효과적인 Cas9 전사 모듈레이터는 핵에 강력히 국소화되며, 여기서 전사에 대해 최대 영향을 미칠 것이라는 가설을 세웠다. 더욱이, 세포질에서 어떤 잔류 Cas9는 원치않는 효과를 가질 수 있다. 출원인은 야생형 Cas9가 다중 핵 위치 신호 (NLS)를 포함하지 않고는 핵에 국소화되지 않는다는 것을 결정했다(CRISPR 시스템은 하나 이상의 NLS를 가질 필요가 없지만, 적어도 하나 이상의 NLS(들)를 갖는 것이 유익하다). 다중 NLS 서열이 요구되었기 때문에, 핵 위치를 파괴할 수 있는, Cas9와 융합되는 핵 및 어떤 추가의 도메인에 Cas9를 넣는 것이 어려운 이유가 되었다. 따라서, 출원인은 상이한 NLS 서열을 가진 4개의 Cas9-VP64-GFP 융합 작제물을 제작했다(pXRP02- p렌티2-EF1a-NLS-hSpCsn1(10A,840A)-NLS-VP64-EGFP, pXRP04- p렌티2-EF1a-NLS-hSpCsn1(10A,840A)-NLS-VP64-2A-EGFP-NLS, pXRP06- p렌티2-EF1a-NLS-EGFP-VP64-NLS-hSpCsn1(10A,840A)-NLS, pXRP08- p렌티2-EF1a-NLS-VP64-NLS-hSpCsn1(10A,840A)-NLS-VP64-EGFP-NLS). 이들 작제물은 인간 EF1a 프로모터의 발현하에 렌티 백본에 클로닝되었다. WPRE 요소가 또한 더욱 확실한 단백질 발현을 위해 첨가되었다. 각 작제물은 리포펙탐 2000을 사용하여 HEK 293FT 세포에 트랜스펙션되고 트랜스펙션 24 시간 후 이미지화되었다. 최상의 핵 위치는 융합 단백질이 융합 단백질의 N- 및 C-말단에 모두 NLS 서열을 가질 때 얻어진다. 최고로 관찰된 핵 위치는 4 개의 NLS 요소를 가진 작제물에서 발생된다.

[1555] Cas9에 대한 N1S 요소의 영향을 더 확실히 이해하기 위해서, 출원인은 0 내지 3 개 텐덤 반복부를 고려하여 N- 또는 C-말단에 동일한 알파 임포틴 NLS 서열을 첨가함으로써 16 개의 Cas9-GFP 융합체를 제작했다. 각 작제물은 리포펙탐 2000을 사용하여 HEK 293FT 세포에 트랜스펙션되고 트랜스펙션 24 시간 후 이미지화되었다. 주목할 점은 NLS 요소의 수가 핵 위치의 범위와 직접 상관되지 않는다는 것이다. C-말단에 NLS를 부가하는 것은 N-말단에 부가하는 것보다 핵 위치에 더 많은 영향을 미친다.

[1556] Cas9 전사 활성인자: 출원인은 Sox2 유전자좌를 표적화하고 RT-qPCR에 의해서 전사 활성화를 정량함으로써 Cas9-VP64 단백질을 기능적으로 시험했다. 8 개의 DNA 표적 부위가 Sox2의 프로모터에 걸쳐 놓이도록 선택되었다. 각 작제물은 리포펙탐 2000을 사용하여 HEK 293FT 세포에 트랜스펙션되었고, 트랜스펙션 72 시간 후 총 RNA가 세포로부터 추출되었다. 1 ㎍의 RNA가 40 ㎕ 반응물에서 cDNA로 역전사되었다(qScript 수퍼믹스). 2 ㎕의 반응 생성물이 TaqMan 분석 qPCR 반응물의 단일 20 ㎕에 첨가되었다. 각 실험은 생물학적 및 기술적으로 3 번 중복 수행되었다. RT 대조군 및 주형 대조군 반응물은 증폭을 나타내지 않았다. 강한 핵 위치를 나타내지 않은 작제물인 pXRP02 및 pXRP04는 활성화를 가져오지 않는다. 강한 핵 위치를 나타낸 작제물인 pXRP08의 경우, 중도의 활성화가 관찰되었다. 통계적으로 유의한 활성화가 가이드 RNAs Sox2.4 및 Sox2.5에서 관찰되었다.

[1557] *실시예 32: 생체내 마우스 데이터*

[1558] **재료 및 시약**

[1559] 허큘라제 II 융합 폴리머라제(Agilent Technologies, 카탈로그 번호 600679)

[1560] 10x NE완충제 4 (NEB, 카탈로그 번호 B7004S)

[1561] BsaI HF (NEB, 카탈로그 번호 R3535S)

[1562] T7 DNA 리가아제(Enzymatics, 카탈로그 번호 L602L)

[1563] FastDigest 완충제, 10X (ThermoScientific, 카탈로그 번호 B64)

[1564] FastDigest NotI(ThermoScientific, 카탈로그 번호 FD0594)

[1565] FastAP 알칼리성 포스파타제(ThermoScientific, 카탈로그 번호 EF0651)

[1566] 리포펙타민 2000 (Life Technologies, 카탈로그 번호 11668-019)

[1567] 트립신(Life Technologies, 카탈로그 번호 15400054)

[1568] 겸자 #4(Sigma, 카탈로그 번호 Z168777-1EA)

[1569] 겸자 #5(Sigma, 카탈로그 번호 F6521-1EA)

[1570] 10x 행크 균형 염 용액(Sigma, 카탈로그 번호 H4641-500ML)

[1571] 페니실린/스트렙토마이신 용액(Life Technologies, 카탈로그 번호 P4333)

[1572] 뉴로베이셜(Life Technologies, 카탈로그 번호 21103049)

[1573] B27 보충액(Life Technologies, 카탈로그 번호 17504044)

[1574] L-글루타민(Life Technologies, 카탈로그 번호 25030081)

[1575] 글루타메이트(Sigma, 카탈로그 번호 RES5063G-A7)

[1576] β-메르캅토에탄올(Sigma, 카탈로그 번호 M6250-100ML)

[1577] HA 토끼 항체(Cell Signaling, 카탈로그 번호 3724S)

[1578] LIVE/DEAD® 세포 이미징 키트(Life Technologies, 카탈로그 번호 R37601)

[1579] 30G World Precision Instrument 주사기(World Precision Instruments, 카탈로그 번호 NANOFIL)

[1580] 정위고정 장치 (Kopf Instruments)

[1581] UltraMicroPump3 (World Precision Instruments, 카탈로그 번호 UMP3-4)

[1582] 수크로스(Sigma, 카탈로그 번호 S7903)

[1583] 염화칼슘(Sigma, 카탈로그 번호 C1016)

[1584] 마그네슘 아세테이트(Sigma, 카탈로그 번호 M0631)

[1585] Tris-HCl(Sigma, 카탈로그 번호 T5941)

[1586] EDTA (Sigma, 카탈로그 번호 E6758)

[1587] NP-40 (Sigma, 카탈로그 번호 NP40)

[1588] 페닐메탄설포닐 플루오라이트(Sigma, 카탈로그 번호 78830)

[1589] 염화마그네슘(Sigma, 카탈로그 번호 M8266)

[1590] 염화칼륨(Sigma, 카탈로그 번호 P9333)

[1591] β-글리세로포스페이트(Sigma, 카탈로그 번호 G9422)

[1592] 글리세롤 (Sigma, 카탈로그 번호 G9012)

[1593] Vybrant® DyeCycle™ Ruby Stain (Life technologies, 카탈로그 번호 S4942)

[1594] FACS Aria Flu-act-세포 정렬기 (Koch Institute of MIT, Cambridge US)

[1595] DNAeasy 혈액 & 조직 키트(Qiagen, 카탈로그 번호 69504)

[1596] **과정**

[1597] ***뇌에서 생체내 사용하기 위한 gRNA 다중복합체의 구성***

[1598] 출원인은 마우스 TET 및 DNMT 패밀리 멤버를 표적화하는 단일 gRNA를 설계하고 PCR 증폭했다(본 명세서에 설명된 대로). 표적화 효율이 N2a 세포주에서 평가되었다(도 33). 몇 개 유전자의 생체내 동시 변형을 얻기 위하여, 효과적인 gRNA가 AAV-패키징 벡터에서 다중복합체화되었다(도 34). 시스템 효율의 추가 분석을 촉진하기 위해서, 출원인은 시스템에 인간 시냅신 I 프로모터의 제어하에 GFP-KASH 도메인 융합 단백질로 구성된 발현 카세트를 첨가했다(도 34). 이 변형은 뉴런 집단에서 시스템 효율의 추가 분석을 허용한다(더 상세한 과정은 **핵 정렬 및 생체내 결과** 섹션 참조).

[1599] 시스템의 모든 4 부분이 다음의 프라이머를 사용하여 허큘라제 II 융합 폴리머라제를 사용해서 PCR 증폭되었다:

1st U6 Fw: gagggtctcgtccttgcggccgcgctagcgagggcctatttcccatgattc (서열 번호 ___)
1st gRNA Rv: ctcggtctcggtAAAAAAgcaccgactcggtgccacttttcaagttgataacggactagc
cttattttaacttgctaTTTCtagctctaaaacNNNNNNNNNNNNNNNNNNNNGGTGTTTC
GTCCTTTCCAC (서열 번호 ___)
2nd U6 Fw:
gagggtctcTTTaccggtgagggcctatttcccatgattcc (서열 번호 ___)
2nd gRNA Rv: ctcggtctcctcAAAAAAgcaccgactcggtgccacttttcaagttgataacggactagc
cttattttaacttgctaTTTCtagctctaaaacNNNNNNNNNNNNNNNNNNNNGGTGTTTC
GTCCTTTCCAC (서열 번호 ___)
3rd U6 Fw:
gagggtctcTTTgagctcgagggcctatttcccatgattc (서열 번호 ___)
3rd gRNA Rv: ctcggtctcgcgtAAAAAAgcaccgactcggtgccacttttcaagttgataacggactag
ccttattttaacttgctaTTTCtagctctaaaacNNNNNNNNNNNNNNNNNNNNGGTGTTT
CGTCCTTTCCA (서열 번호 ___)
hSyn_GFP-kash Fw: gagggtctcTTacgcgtgtgtctagac (서열 번호 ___)
hSyn_GFP-kash Rv:
ctcggtctcAaggaCAGGGAAGGGAGCAGTGGTTCACGCCTGTAATCCCAGCAATTT
GGGA GGCCAAGGTGGGTAGATCACCTGAGATTAGGAGTTGC (서열 번호 ___)

[1600] (NNNNNNNNNNNNNNNNNNNN (서열 번호 ___)은 역 상보성 표적화 게놈 서열임)

[1601] 출원인은 단일 단계 반응에서 시스템의 모든 부분을 조립하기 위해 골든 게이트 전략을 사용했다(1:1 분자비):

[1602] 1st U6_gRNA 18 ng

[1603] 2nd U6_gRNA 18 ng

[1604] 3rd U6_gRNA 18 ng

[1605] Syn_GFP-kash 100 ng

[1606] 10x NE완충제 4 1.0 μℓ

[1607] 10x BSA 1.0 μℓ

[1608] 10 mM ATP 1.0 μℓ

[1609] BsaI HF 0.75 μℓ

[1610] T7 리가아제 0.25 μℓ

[1611] ddH_2O 10 μℓ

[1612] 사이클 수 조건

[1613] 1-50 37℃, 5m 및 21℃, 5m

[1614] 골든 게이트 반응 생성물은 허큘라제 II 융합 폴리머라제 및 다음의 프라이머를 사용하여 PCR 증폭되었다:

Fw 5' cctgtccttgcggccgcgctagcgagggcc (서열 번호 ___)

[1615] Rv 5' cacgcggccgcaaggacagggaagggagcag (서열 번호 ___)

[1616] PCR 생성물은 NotI 제한 부위를 사용하여 ITR 서열 사이에서 AAV 백본에 클로닝되었다:

[1617] PCR 생성물 소화:

FastDigest 완충제, 10X	3 μl
FastDigest NotI	1 μl
DNA	1 μg
ddH_2O	최대 30 μl

AAV 백본 소화:

FastDigest 완충제, 10X	3 μl
FastDigest NotI	1 μl
FastAP 알칼리성 포스파타제	1 μl
AAV 백본	1 μg
ddH_2O	최대 30 μl

[1618]

[1619] 37℃에서 20분 인큐베이션 후, 샘플이 QIAQuick PCR 정제 키트를 사용하여 정제되었다. 표준화된 샘플이 다음과 같이 1:3 벡터:삽입체 비로 결찰되었다:

[1620] 소화된 pUC19 50 ng

[1621] 소화된 삽입체 1:3 벡터:삽입체 몰비

[1622] T7 리가아제 1 $\mu\ell$

[1623] 2X 신속 결찰 완충제 5 $\mu\ell$

[1624] ddH_2O최대 10 $\mu\ell$

[1625] 결찰 반응 생성물로 박테리아의 형질전환 후, 출원인은 Sanger 시퀀싱에 의해서 얻어진 클론을 확인했다.

[1626] 양성 DNA 클론이 Cas9 작제물과 공-트랜스펙션 후 N2a 세포에서 시험되었다(도 35 및 36).

[1627] ***AAV 전달을 위한 신규 Cas9 작제물의 설계***

[1628] AAV 전달 시스템은 그것의 독특한 특징에도 불구하고 팩킹 제한을 가지는데, 발현 카세트를 생체내에 성공적으로 전달하기 위해서는 그것이 4.7 kb 미만의 크기를 가져야 한다. SpCas9 발현 카세트의 크기를 감소시키고 전달을 촉진하기 위해서, 출원인은 몇 가지 변경을 시험했다: 상이한 프로모터, 짧은 폴리A 신호 및 마지막으로 *Staphylococcus aureus* (SaCas9)로부터의 더 작은 버전의 Cas9 (도 37 및 38). 모든 시험된 프로모터는 마우스 *Mecp2*(Gray et al., 2011), 래트 *Map1b* 및 말단절단된 래트 *Map1b* (Liu and Fischer, 1996)를 포함하는 뉴런에서 활성인 것으로 이미 시험되고 공개되었다. 대안적인 합성 폴리A 서열도 역시 기능적인 것으로 이미 밝혀졌다(Levitt et al., 1989; Gray et al., 2011). 모든 클로닝된 작제물은 리포펙타민 2000에 의한 트랜스펙션 후 N2a 세포에서 발현되었고, 웨스턴 블롯팅 방법으로 시험되었다(도 39).

[1629] ***1차 뉴런에서 AAV다중복합체 시스템의 시험***

[1630] 뉴런에서 개발된 시스템의 기능성을 확인하기 위하여, 출원인은 시험관내 1차 뉴런 배양물을 사용한다. 마우스 피질 뉴런이 Banker and Goslin (Banker and Goslin, 1988)에 의해서 이미 공개된 프로토콜에 따라서 제조되었다.

[1631] 뉴런 세포는 16 일 배아로부터 얻어진다. 안락사시킨 임신한 암컷으로부터 배아를 추출하고 절두한 다음, 머리를 얼음 냉각된 HBSS에 넣어 두었다. 다음에, 겸자(4# 및 #5)를 사용하여 두개골로부터 뇌를 추출하고, 다른 교체된 얼음 냉각된 HBSS에 옮긴다. 얼음 냉각된 HBSS로 채워진 페트리 디쉬에서 #5 겸자를 사용하여 입체 현미경의 도움하에 추가의 단계를 수행한다. 각 다른 쪽 뇌줄기에서 뇌의 반구를 분리하며, 뇌척수막은 없다. 다음에,

해마가 주의깊게 절개되어 얼음 냉각된 HBSS로 채운 15㎖ 원뿔형 튜브에 넣는다. 해마 절개 후 남은 피질은 뇌줄기 잔류물과 후 신경구 제거 후 유사한 프로토콜을 사용하여 추가의 세포 분리에 사용될 수 있다. 분리된 해마는 10 ㎖ 얼음-냉각된 HBSS로 3번 세척하고, 37℃에서 HBSS (해마 당 10 ㎕ 2.5% 트립신 첨가된 4 ㎖ HBSS) 중에서 트립신과 함께 15분 인큐베이션하여 해리시킨다. 트립신화 후, 해마는 37℃로 예열된 HBSS로 어떤 미량의 트립신을 제거하기 위해서 매우 주의깊게 3 번 세척되고, 따뜻한 HBSS에서 해리된다. 출원인은 일반적으로 1 ㎖ 피펫 팁을 사용하여 1 ㎖ HBSS에서 10-12 개 배아로부터 얻어진 세포를 해리하고, 해리된 세포를 최대 4 ㎖로 희석한다. 세포는 250 세포/mm^2의 밀도로 평판되고, 최대 3 주 동안 37℃ 및 5% CO_2에서 배양된다.

[1632] **HBSS**

[1633] 435 ㎖ H_2O

[1634] 50 ㎖ 10x 행크 균형 염 용액

[1635] 16.5 ㎖ 0.3M HEPES pH 7.3

[1636] 5 ㎖ 페니실린-스트렙토마이신 용액

[1637] 여과(0.2 ㎛) 및 4℃ 보관

[1638] **뉴런 평판 배지**(100 ㎖)

[1639] 97 ㎖ 뉴로베이셜

[1640] 2 ㎖ B27 보충액

[1641] 1 ㎖ 페니실린-스트렙토마이신 용액

[1642] 250 ㎕ 글루타민

[1643] 125 ㎕ 글루타메이트

[1644] 뉴런은 배양물에서 4 내지 7 일 사이에 HEK293FT 세포의 여과된 배지로부터의 농축된 AAV1/2 바이러스 또는 AAV1 바이러스로 형질도입되고, 형질도입 후 적어도 1 주 동안 배양물에 유지되며, 이로써 전달된 유전자 발현이 허용된다.

[1645] *시스템의 AAV-유도된 발현*

[1646] 출원인은 웨스턴 블롯 방법을 사용하여 AAV 전달 후 뉴런 배양물에서 SpCas9 및 SaCas9의 발현을 확인했다(도 42). 형질도입 1 주 후에 뉴런이 β-메르캅토에탄올을 가진 NuPage SDS 로딩 완충제에 수집하였고, 5 분 동안 95℃에서 단백질 분해되었다. 샘플을 SDS PAGE 겔에서 분리하였고, WB 단백질 검출을 위해 PVDF 막으로 옮겼다. Cas9 단백질이 HA 항체로 검출되었다.

[1647] gRNA 다중복합체 AAV로부터 Syn-GFP-kash의 발현이 형광 현미경으로 확인되었다(도 50).

[1648] *독성*

[1649] CRISPR 시스템을 가진 AAV의 독성을 평가하기 위해서, 출원인은 바이러스 형질도입 1주 후에 뉴런의 전체 형태를 시험했다(도 45). 추가로, 출원인은 LIVE/DEAD® 세포 이미징 키트를 사용하여 설계된 시스템의 잠재적 독성을 시험했으며, 이것은 배양물에서 살아 있는 세포와 죽은 세포의 구별을 허용한다. 그것은 세포내 에스테라제 활성의 존재에 기초한다(비-형광 칼세인 AM의 짙은 녹색 형광 칼세인으로의 효소 전환에 의해서 결정된다). 한편, 키트의 적색, 세포-불침투 구성성분은 손상된 막을 가진 세포로 진입하여 DNA와 결합해서 죽은 세포에서 형광을 생성한다. 플루오로포어는 모두 형광 현미경을 사용하여 살아 있는 세포에서 쉽게 시각화될 수 있다. 일차 피질 뉴런에서 Cas9 단백질과 다중복합체 gRNA 작제물의 AAV-유도된 발현은 잘 용인되었으며 독성은 없었는데(도 43 및 44), 이것은 설계된 AAV 시스템이 생체내 시험에 적합하다는 것을 나타낸다.

[1650] ***바이러스 제조***

[1651] McClure et al., 2011에 설명된 방법에 따라서 농축 바이러스가 제조되었다. 상청액 바이러스 제조는 HEK293FT 세포에서 발생했다.

[1652] ***뇌 수술***

[1653] 바이러스 벡터 주사의 경우, 10-15 주령 수컷 C57BL/6N 마우스가 복강내 주사에 의해서 케타민/자일라진 칵테일로 마취시켰다(케타민 용량 100 mg/kg 및 자일라진 용량 10 mg/kg). Buprenex의 복강내 투여가 예비-선취 마취제로서 사용되었다(1mg/kg). 동물은 두개골을 움직이지 않도록 유지하기 위해서 이내 배치 스터드 및 투스 바를 사용하여 Kopf 정위고정 장치에 고정되었다. 해마의 CA1 영역에 주사하기 위해 핸드헬드 드릴을 사용하여 브레그마에서 앞쪽 -3.0 mm, 측면 3.5 mm에 구멍 (1-2 mm)을 만들었다. 2.5 mm 깊이의 30G World Precision Instrument 주사기를 사용하여 AAV 바이러스 입자 용액이 총 1 ㎕ 부피가 주사되었다. 주사는 조직 손상을 방지하기 위하여 0.5 ㎕/분의 유속의 'World Precision Instruments UltraMicroPump3' 주입 펌프에 의해서 모니터링되었다. 주입이 완료되었을 때, 0.5 mm/분의 속도로 주사 바늘이 서서히 제거되었다. 주사 후, 피부가 6-0 Ethilon 봉합실로 봉합되었다. 동물은 1 ㎖ 락테이트 링거액으로 수술 후 수화되었고(피하), 보행 회복을 달성할 때까지 온도 제어 (37℃) 환경에 수용되었다. 수술 3 주 후 동물은 깊은 마취에 의해서 안락사되었고, 이후 핵 정렬을 위해 조직이 제거되거나, 면역화학을 위해 4% 파라포름알데하이드로 관류되었다.

[1654] ***핵 정렬 및 생체내 결과***

[1655] 출원인은 표지된 세포 핵의 형광 활성화 세포 정렬 (FACS) 및 DNA, RNA 및 핵 단백질의 하류 가공을 위해 GFP로 gRNA 표적화된 뉴런 세포 핵을 특이적으로 유전적으로 태그화하는 방법을 설계했다. 이 목적을 위해서, 출원인의 다중복합체 표적화 벡터는 GFP와 마우스 핵 막 단백질 도메인 KASH (Starr DA, 2011, Current biology) 의 융합 단백질과 관심대상의 특정 유전자 유전자좌를 표적화하기 위한 3 gRNA를 모두 발현하도록 설계되었다(도 34). GFP-KASH는 인간 시냅신 프로모터의 제어하에 발현되었으며, 이로써 뉴런을 특이적으로 표적화할 수 있다. 융합 단백질 GFP-KASH의 아미노산은 다음과 같았다:

[1656] MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICT
TGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGN
YKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIK
VNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMV
LLEFVTAAGITLGMDELYKSGLRSREEEEETDSRMPHLDSPGSSQPRRSFLSRVIRAAL
PLQLLLLLLLLLACLLPASEDDYSCTQANNFARSFYPMLRYTNGPPPT (서열 번호 ___)

[1657] 뇌에 AAV1/2 매개 전달 후 1 주 뒤에 GFP-KASH의 확실한 발현이 관찰되었다. FACS 및 표지된 핵의 하류 가공을 위해서, 수술 3주 후 해마가 절개되고, 구배 원심분리 단계를 사용한 세포 핵 정제를 위해 가공되었다. 이 목적을 위해서, 조직은 320 mM 수크로스, 5 mM CaCl, 3 mM $Mg(Ac)_2$, 10 mM 트리스 pH 7.8, 0.1 mM EDTA, 0.1% NP40, 0.1 mM 페닐메탄설포닐 플루오라이드 (PMSF), 1 mM β-메르캅토에탄올에서 2 ㎖ Dounce 균질화장치 (Sigma)를 사용하여 균질화되었다. 균질화물은 4℃에서 3,500 rpm에서 30 분 동안 제조자의 프로토콜에 따라서 25% 내지 29% Optiprep® 구배에서 원심분리되었다. 핵 펠릿이 340 mM 수크로스, 2 mM $MgCl_2$, 25 mM KCl, 65 mM 글리세로포스페이트, 5% 글리세롤, 0.1 mM PMSF, 1 mM β-메르캅토에탄올에 재현탁되고, 세포 핵을 표지하기 위해 Vybrant® DyeCycle™ Ruby Stain(Life technologies)이 첨가되었다(DNA에 대해 근적외선 방출을 제공한다). 표지되고 정제된 핵은 FACS에 의해서 Aria Flu-act-세포 정렬장치 및 BDFACS Diva 소프트웨어를 사용하여 정렬되었다. 정렬된 GFP+ 및 GFP- 핵은 최종적으로 표적화된 게놈 영역의 서베이어 분석을 위하여 DNAeasy 혈액 & 조직 키트(Qiagen)를 사용하여 게놈 DNA를 정제하는데 사용되었다. 하류 가공을 위해 표적화된 세포로부터 핵 RNA 또는 단백질을 정제하는데도 동일한 접근법이 쉽게 사용될 수 있다. 2-벡터 시스템(도 34)으로 인하여, 출원인은 이 접근법에서 효과적인 Cas9 매개 DNA 절단이 뇌에서 세포의 작은 하위세트에서만 발생할 것이라고 예상했다(다중복합체 표적화 벡터와 Cas9 암호화 벡터로 공-감염된 세포). 여기 설명된 방법은 출원인이 DNA, RNA 및 핵 단백질을 관심대상의 3 gRNA를 발현하는 세포 집단으로부터 특이적으로 정제할 수 있도록 하며, 따라서 Cas9 매개 DNA 절단을 겪을 것으로 추정된다. 이 방법을 사용함으로써 출원인은 세포의 작은 하위세트에서만 발생한 효과적인 생체내 DNA 절단을 시각화할 수 있었다.

[1658] 본질적으로, 출원인이 여기 나타낸 것은 표적화된 생체내 절단이다. 또한, 출원인은 독립적이지만 동시에 몇 개의 상이한 서열이 표적화되는 다중 접근법을 사용했다. 제시된 시스템은 뇌 병리 상태 (유전자 녹아웃, 예를 들어 파킨슨 질환)를 연구하기 위해 적용될 수 있으며, 또한 뇌에서 게놈 편집 도구의 추가의 개발에 대한 분야도 열려 있다. 뉴클레아제 활성을 유전자 전사 조절인자 또는 후성적 조절인자로 대체함으로써 학습 및 기억 형성으로서 병리학적 상태뿐만 아니라 생리학적 과정에서도 유전자 조절 및 후성적 변호의 역할에 대한 전 범위의 과학적 질문에 답하는 것이 가능할 것이다. 마지막으로, 제시된 기술은 영장류와 같은 더 복잡한 포유류 시스템

에 적용될 수 있으며, 현재 기술의 제한을 극복할 수 있다.

[1659] *실시예 33: 모델 데이터*

[1660] 몇 가지 질환 모델이 구체적으로 조사되었다. 이들은 신규 자폐 위험 유전자들인 CHD8, KATNAL2, 및 SCN2A; 그리고 증후군성 자폐 (앤젤만 증후군) 유전자 UBE3A를 포함한다. 이들 유전자와 결과의 자폐 모델이 물론 바람직하지만, 본 발명은 어떤 유전자에도 적용될 수 있으며, 따라서 어떤 모델도 가능하다.

[1661] 출원인은 인간 배아 줄기세포 (hESC)에서 Cas9 뉴클레아제를 사용하여 이들 세포주를 제작했다. 세포주는 Cbh-Cas9-2A-EGFP 및 pU6-sgRNA에 의한 hESC의 일시적 트랜스펙션에 의해서 생성되었다. 환자 논센스(녹아웃) 돌연변이가 자폐 환자의 전체 엑솜 시퀀싱 연구로부터 최근 설명된 동일한 엑손을 표적화하는 각 유전자에 대해 2개의 sgRNA가 설계된다. Cas9-2A-EGFP 및 pU6 플라스미드가 이 프로젝트를 위해 구체적으로 생성되었다.

[1662] *실시예 34: AAV 생성 시스템 또는 프로토콜*

[1663] 고 처리량 스크리닝 사용을 위해 개발된, 특히 웰에서의 작업을 위한 AAV 생성 시스템 또는 프로토콜이 본원에서 제공되지만, 본 발명은 역시 더 넓은 적용성을 가진다. 내인성 유전자 발현의 조작은 다양한 난제를 나타내며, 발현의 비율은 조절 구성요소, mRNA 프로세싱, 및 전사체 안정성을 포함하는 많은 요인들에 따른다. 이 난제를 극복하기 위해서 출원인은 전달을 위한 아데노-연관 바이러스(AAV)-기반 벡터를 개발했다. AAV는 ssDNA-기반 게놈을 가지며, 따라서 재조합에 덜 민감하다.

[1664] AAV1/2(혈청형 AAV1/2, 즉 하이브리드 또는 모자이크 AAV1/AAV2 캡시드 AAV) 헤파린 정제된 농축된 바이러스 프로토콜

[1665] **<u>배지: D10 + HEPES</u>**

[1666] 500 mℓ 보틀 DMEM 고 글루코오스 + 글루타맥스(GIBCO)

[1667] 50 mℓ Hyclone FBS (열-비활성화) (Thermo Fischer)

[1668] 5.5 mℓ HEPES 용액 (1M, GIBCO)

[1669] 세포: 저 계대 HEK293FT (계대 <10 바이러스 제조시, 바이러스 제조를 위해 계대 2-4의 새로운 세포 해동, 3-5 계대 동안 성장)

[1670] **<u>트랜스펙션 시약: 폴리에틸렌이민 (PEI) "맥스"</u>**

[1671] 50 mg PEI "맥스"를 50 mℓ 멸균 Ultrapure H20에 용해

[1672] pH 7.1로 조정

[1673] 0.22 um 플립탑 필터로 여과

[1674] 튜브 밀봉 및 파라핀으로 래핑

[1675] -20℃에서 알리쿼트 냉동(보관의 경우, 즉시 사용될 수도 있다)

[1676] **<u>세포 배양</u>**

[1677] 저 계대 HEK293FT를 D10 + HEPES에서 배양

[1678] 1:2 내지 1:2.5 로 매일 계대

[1679] 세포는 85%를 초과하는 컨플루언시에는 도달하지 않도록 하는 것이 유익하다.

[1680] <u>T75의 경우</u>

[1681] - 따뜻한 10mℓ HBSS (-Mg2+, -Ca2+, GIBCO) + 1mℓ TrypLE Express (GIBCO), 플라스크 당, 37℃ (수조)

[1682] 배지를 완전히 흡인

[1683] - 10 mℓ 따뜻한 HBSS를 부드럽게 첨가 (배지를 완전히 세척하기 위해)

[1684] - 플라스크 당 1 mℓ TrypLE 첨가

[1685] - 1 분 동안 인큐베이터에 플라스크를 둔다(37℃)

[1686] - 세포가 탈착하도록 플라스크를 흔든다

[1687] - 9 ㎖ D10 + HEPES 배지 첨가 (37℃)

[1688] - 5번 위아래로 피펫팅해서 단일 세포 현탁액을 생성한다

[1689] - 1:2 - 1:2.5 (T75의 경우 12 ㎖ 배지) 비로 분할한다(세포가 더 서서히 성장중인 경우, 폐기하고 새로운 배치를 해동한다, 이들은 최적 성장 상태가 아니다)

[1690] - 충분한 세포가 존재하게 되면 곧 T225로 옮긴다(다량의 세포를 취급하는 경우)

[1691] **AAV 제조 (작제물 당 5*15cm 디쉬 규모):**

[1692] 21.5 ㎖ 배지 중의 1000만 세포를 15 cm 디쉬에 평판

[1693] 37℃에서 18-22 시간 동안 인큐베이션

[1694] 트랜스펙션은 평판 당 80% 컨플루언시에서 이상적이다.

[1695] **평판 당**

[1696] 예비 가온된 22 ㎖ 배지(D10 + HEPES)

[1697] DNA 혼합물로 튜브 준비(엔도프리 맥시프렙 DNA 사용):

[1698] 관심대상 플라스미드의 벡터 5.2 ㎍

[1699] 4.35 ㎍ AAV 1 혈청형 플라스미드

[1700] 4.35 ㎍ AAV 2 혈청형 플라스미드

[1701] 10.4 ㎍ pDF6 플라스미드 (아데노바이러스 헬퍼 유전자) 와동시켜 혼합

[1702] 434 ㎕ DMEM 첨가 (혈청 없음!)

[1703] 130 ㎕ PEI 용액 첨가

[1704] 5 내지 10 초 와동

[1705] 예비 가온된 배지에 DNA/DMEM/PEI 혼합물 첨가

[1706] 간단히 와동시켜 혼합

[1707] 15 cm 디쉬에서 배지를 DNA/DMEM/PEI 혼합물로 교체

[1708] 37℃ 인큐베이터로 복귀

[1709] 수거 전 48 h 인큐베이션 (배지가 너무 산성으로 변하지 않도록 보장한다)

[1710] **바이러스 수거:**

[1711] 1. 15 cm 디쉬로부터 배지를 주의깊게 흡인한다(세포가 손실되지 않는 것이 유익하다)

[1712] 2. 각 평판에 25 ㎖ RT DPBS (Invitrogen)를 첨가하고, 세포 스크래퍼로 세포를 부드럽게 제거한다. 50 ㎖ 튜브에 현탁액을 수집한다.

[1713] 3. 10 분 동안 800x g에서 세포를 펠릿화한다.

[1714] 4. 상청액을 버린다.

[1715] **유의점: 원한다면 -80C에서 세포 펠릿을 냉동시킨다**

[1716] 5. 펠릿을 150 mM NaCl, 20 mM Tris pH 8.0에 재현탁하고, 조직 배양 평판 당 10㎖를 사용한다.

[1717] 6. dH_2O 중에 10% 나트륨 데옥시콜레이트의 신선한 용액을 준비한다. 0.5%의 최종 농도를 위해서 조직 배양 평판 당 이것을 1.25 ㎖ 첨가한다. ㎖ 당 50 유닛의 최종 농도로 벤조나제 뉴클레아제를 첨가한다. 튜브를 완전히 혼합한다.

[1718] 7. 1시간 동안 37℃에서 인큐베이션(수조).

[1719] 8. 세포 파편을 15 분 동안 3000 x g에서 원심분리하여 제거한다. 새 50 ㎖ 튜브로 옮겨서 모든 세포 파편이 제거된 것을 보장하여 헤파린 칼럼의 차단을 방지한다.

[1720] **AAV1/2의 헤파린 칼럼 정제:**

[1721] 1. 분당 1㎖로 용액이 칼럼을 통해서 유동하도록 연동 펌프를 사용하여 HiTrap 헤파린 칼럼을 설정한다. 헤파린 칼럼에 기포가 도입되지 않도록 보정하는 것이 중요하다.

[1722] 2. 연동 펌프를 사용하여 칼럼을 10 ㎖ 150 mM NaCl, 20 mM Tris, pH 8.0과 평형을 이루도록 한다.

[1723] 3. 바이러스의 결합: 50 ㎖ 바이러스 용액을 칼럼에 적용하여 관통 유동시킨다.

[1724] 4. 세척 단계 1: 20 ㎖ 100 mM NaCl, 20 mM Tris, pH 8.0로 칼럼 세척(연동 펌프 사용)

[1725] 5. 세척 단계 2: 3 ㎖ 또는 5 ㎖ 주사기를 사용하여 1 ㎖ 200 mM NaCl, 20 mM Tris, pH 8.0로 칼럼 세척을 계속하고, 이어서 1 ㎖ 300 mM NaCl, 20 mM Tris, pH 8.0로 세척한다.

[1726] 관통 유동시켜 버린다.

[1727] (상기 바이러스 용액의 50 분 관통 유동 동안 상이한 완충제를 사용하여 주사기를 준비한다)

[1728] 6. 용출: 5 ㎖ 주사기를 사용하여 부드럽게 가압해서 칼럼으로부터 바이러스를 용출시킨다: 다음을 적용한다: (유속 < 1㎖/분)

[1729] 1.5 ㎖ 400 mM NaCl, 20 mM Tris, pH 8.0

[1730] 3.0 ㎖ 450 mM NaCl, 20 mM Tris, pH 8.0

[1731] 1.5 ㎖ 500 mM NaCl, 20 mM Tris, pH 8.0

[1732] 이들을 15 ㎖ 원심분리 튜브에 수집한다.

[1733] **AAV1/2의 농축:**

[1734] 1. 농축 단계 1: 100,000 분자량 컷오프를 가진 Amicon ultra 15 ㎖ 원심분리 필터를 사용하여 용출된 바이러스를 농축한다. 칼럼 용출액을 농축장치에 로딩하고 2 분 동안 2000x g에서 원심분리한다(실온에서). 농축된 부피 점검 - 대략 500 ㎕가 되어야 한다. 필요하다면 정확한 부피에 도달할 때까지 1 분 간격으로 원심분리한다.

[1735] 2. 완충제 교환: 여과 유닛에 1 ㎖ 멸균 DPBS를 첨가하고, 정확한 부피 (500 ㎕)에 도달할 때까지 1 분 간격으로 원심분리한다.

[1736] 3. 농축 단계 2: Amicon Ultra 0.5 ㎖ 100K 필터 유닛에 500 ㎕ 농축물을 가한다. 2 분 동안 6000 g에서 원심분리한다. 농축된 부피 점검 - 대략 100 ㎕가 되어야 한다. 필요하다면 정확한 부피에 도달할 때까지 1 분 간격으로 원심분리한다.

[1737] 4. 회수: 필터 삽입물을 뒤집어 반전시키고 새 수집관에 삽입한다. 2 분 동안 1000 g에서 원심분리한다.

[1738] 알리쿼트 분할 및 -80℃에서 냉동

[1739] 주사 부위 당 1 ㎕가 전형적으로 요구된다, 따라서 작은 알리쿼트(예를 들어, 5㎕)가 추천된다(바이러스의 냉동-해동은 피한다).

[1740] qPCR을 사용하여 DnaseI-내성 GC 입자 역가를 결정한다(별도 프로토콜 참조)

[1741] **재료**

[1742] Amicon Ultra, 0.5 ㎖, 100 K; MILLIPORE; UFC510024

[1743] Amicon Ultra, 15 ㎖, 100 K; MILLIPORE; UFC910024

[1744] 벤조나제 뉴클레아제; Sigma-Aldrich, E1014

[1745] HiTrap 헤파린 카트리지; Sigma-Aldrich; 54836

[1746] 나트륨 데옥시콜레이트; Sigma-Aldrich; D5670

[1747] AAV1 상청액 생성 프로토콜

[1748] 배지: D10 + HEPES

[1749] 500 mℓ 보틀 DMEM 고 글루코오스 + 글루타맥스(GIBCO)

[1750] 50 mℓ Hyclone FBS (열-비활성화) (Thermo Fischer)

[1751] 5.5 mℓ HEPES 용액 (1 M, GIBCO)

[1752] 세포: 저 계대 HEK293FT (계대 <10 바이러스 제조시)

[1753] 바이러스 제조를 위해 계대 2-4의 새로운 세포 해동, 2-5 계대 동안 성장

[1754] 트랜스펙션 시약: 폴리에틸렌이민 (PEI) "맥스"

[1755] 50 mg PEI "맥스"를 50 mℓ 멸균 Ultrapure H20에 용해

[1756] pH 7.1로 조정

[1757] 0.22 um 플립탑 필터로 여과

[1758] 튜브 밀봉 및 파라핀으로 래핑

[1759] -20℃에서 알리쿼트 냉동 (보관의 경우, 즉시 사용될 수도 있다)

[1760] 세포 배양

[1761] 저 계대 HEK293FT를 D10 + HEPES에서 배양 1:2 내지 1:2.5 로 매일 계대

[1762] 세포는 85%를 초과하는 컨플루언시에는 도달하지 않도록 하는 것이 유익하다

[1763] T75의 경우

[1764] - 따뜻한 10 mℓ HBSS (-Mg2+, -Ca2+, GIBCO) + 1 mℓ TrypLE Express (GIBCO), 플라스크 당, 37℃ (수조)

[1765] - 배지를 완전히 흡인

[1766] - 10 mℓ 따뜻한 HBSS를 부드럽게 첨가 (배지를 완전히 세척하기 위해)

[1767] - 플라스크 당 1 mℓ TrypLE 첨가

[1768] - 1 분 동안 인큐베이터에 플라스크를 둔다(37℃)

[1769] - 세포가 탈착하도록 플라스크를 흔든다

[1770] - 9 mℓ D10 + HEPES 배지 첨가(37℃)

[1771] - 5 번 위아래로 피펫팅해서 단일 세포 현탁액을 생성한다

[1772] - 1:2 - 1:2.5 (T75의 경우 12 mℓ 배지) 비로 분할한다(세포가 더 서서히 성장중인 경우, 폐기하고 새로운 배치를 해동한다, 이들은 최적 성장 상태가 아니다)

[1773] - 충분한 세포가 존재하게 되면 곧 T225로 옮긴다(다량의 세포를 취급하는 경우)

[1774] AAV 제조 (단일 15 cm 디쉬 규모)

[1775] 21.5 mℓ 배지 중의 1000만 세포를 15 cm 디쉬에 평판

[1776] 37℃에서 18-22 시간 동안 인큐베이션

[1777] 트랜스펙션은 평판 당 80% 컨플루언시에서 이상적이다

[1778] 예비 가온된 22 mℓ 배지(D10 + HEPES)

[1779] DNA 혼합물로 튜브 준비 (엔도프리 맥시프렙 DNA 사용):

[1780] 관심대상 플라스미드의 벡터 5.2 ㎍

[1781] 8.7 ㎍ AAV 1 혈청형 플라스미드

[1782] 10.4 ㎍ DF6 플라스미드(아데노바이러스 헬퍼 유전자)

[1783] 와동시켜 혼합

[1784] 434 ㎕ DMEM 첨가 (혈청 없음!) 130 ㎕ PEI 용액 첨가

[1785] 5-10 초 와동

[1786] 예비 가온된 배지에 DNA/DMEM/PEI 혼합물 첨가

[1787] 간단히 와동시켜 혼합

[1788] 15 cm 디쉬에서 배지를 DNA/DMEM/PEI 혼합물로 교체

[1789] 37℃ 인큐베이터로 복귀

[1790] 수거 전 48 h 인큐베이션 (배지가 너무 산성으로 변하지 않도록 보장하기 위해 모니터링하는 것이 유익하다)

[1791] 바이러스 수거:

[1792] 15 cm 디쉬로부터 상청액을 제거한다

[1793] 0.45 um 필터로 알리쿼트를 여과하고(낮은 단백질 결합), -80℃에서 냉동한다

[1794] 형질도입 (24-웰 형식으로 일차 뉴런 배양, 5DIV)

[1795] 신선한 뉴로베이셜로 전환되도록 뉴런의 각 웰에서 뉴로베이셜을 완전히 교체한다(일반적으로 웰 당 500 ㎕ 중 400 ㎕이 교체된다)

[1796] 37℃ 수조에서 AAV 상청액을 해동한다

[1797] 30 분 동안 인큐베이터에서 평형을 이루도록 둔다.

[1798] 각 웰에 250 ㎕ AAV 상청액을 첨가한다

[1799] 37℃에서 24h 인큐베이션한다

[1800] 배지/상청액을 제거하고, 신선한 뉴로베이셜로 완전히 교체한다

[1801] 발현은 48 시간 후 육안으로 볼 수 있게 되며, 감염 후 약 6-7 뒤에 포화된다

[1802] GOI를 가진 pAAV 플라스미드를 위한 작제물은 ITRS를 포함해서 4.8 kb를 초과하지 않아야 한다

[1803] 인간 코돈 최적화된 서열(즉, 인간에서의 발현에 대해 최적화된): SaCas9의 예가 아래 제공된다(서열 번호 ___):

[1804]

ACCGGTGCCACCATGTACCCATACGATGTTCCAGATTACGC
TTCGCCGAAGAAAAAGCGCAAGGTCGAAGCGTCCATGAAAAGGAACTACATTCT
GGGGCTGGACATCGGGATTACAAGCGTGGGGTATGGGATTATTGACTATGAAAC
AAGGGACGTGATCGACGCAGGCGTCAGACTGTTCAAGGAGGCCAACGTGGAAAA
CAATGAGGGACGGAGAAGCAAGAGGGGAGCCAGGCGCCTGAAACGACGGAGAA
GGCACAGAATCCAGAGGGTGAAGAAACTGCTGTTCGATTACAACCTGCTGACCG
ACCATTCTGAGCTGAGTGGAATTAATCCTTATGAAGCCAGGGTGAAAGGCCTGA
GTCAGAAGCTGTCAGAGGAAGAGTTTTCCGCAGCTCTGCTGCACCTGGCTAAGC
GCCGAGGAGTGCATAACGTCAATGAGGTGGAAGAGGACACCGGCAACGAGCTG
TCTACAAAGGAACAGATCTCACGCAATAGCAAAGCTCTGGAAGAGAAGTATGTC
GCAGAGCTGCAGCTGGAACGGCTGAAGAAAGATGGCGAGGTGAGAGGGTCAATT
AATAGGTTCAAGACAAGCGACTACGTCAAAGAAGCCAAGCAGCTGCTGAAAGTG
CAGAAGGCTTACCACCAGCTGGATCAGAGCTTCATCGATACTTATATCGACCTGC
TGGAGACTCGGAGAACCTACTATGAGGGACCAGGAGAAGGGAGCCCCTTCGGAT
GGAAAGACATCAAGGAATGGTACGAGATGCTGATGGGACATTGCACCTATTTTCC
AGAAGAGCTGAGAAGCGTCAAGTACGCTTATAACGCAGATCTGTACAACGCCCT
GAATGACCTGAACAACCTGGTCATCACCAGGGATGAAAACGAGAAACTGGAATA
CTATGAGAAGTTCCAGATCATCGAAAACGTGTTTAAGCAGAAGAAAAAGCCTAC
ACTGAAACAGATTGCTAAGGAGATCCTGGTCAACGAAGAGGACATCAAGGGCTA
CCGGGTGACAAGCACTGGAAAACCAGAGTTCACCAATCTGAAAGTGTATCACGA
TATTAAGGACATCACAGCACGGAAAGAAATCATTGAGAACGCCGAACTGCTGGA
TCAGATTGCTAAGATCCTGACTATCTACCAGAGCTCCGAGGACATCCAGGAAGAG
CTGACTAACCTGAACAGCGAGCTGACCCAGGAAGAGATCGAACAGATTAGTAAT
CTGAAGGGGTACACCGGAACACACAACCTGTCCCTGAAAGCTATCAATCTGATTC
TGGATGAGCTGTGGCATACAAACGACAATCAGATTGCAATCTTTAACCGGCTGAA
GCTGGTCCCAAAAAAGGTGGACCTGAGTCAGCAGAAAGAGATCCCAACCACACT
GGTGGACGATTTCATTCTGTCACCCGTGGTCAAGCGGAGCTTCATCCAGAGCATCAA
AGTGATCAACGCCATCATCAAGAAGTACGGCCTGCCCAATGATATCATTATCGAGCT
GGCTAGGGAGAAGAACAGCAAGGACGCACAGAAGATGATCAATGAGATGCAGAAA
CGAAACCGGCAGACCAATGAACGCATTGAAGAGATTATCCGAACTACCGGGAAAGA
GAACGCAAAGTACCTGATTGAAAAAATCAAGCTGCACGATATGCAGGAGGGAAAGT

GTCTGTATTCTCTGGAGGCCATCCCCCTGGAGGACCTGCTGAACAATCCATTCAACT
ACGAGGTCGATCATATTATCCCCAGAAGCGTGTCCTTCGACAATTCCTTTAACAACA
AGGTGCTGGTCAAGCAGGAAGAGAACTCTAAAAAGGGCAATAGGACTCCTTTCCAG
TACCTGTCTAGTTCAGATTCCAAGATCTCTTACGAAACCTTTAAAAAGCACATTCTG
AATCTGGCCAAAGGAAAGGGCCGCATCAGCAAGACCAAAAAGGAGTACCTGCTGG
AAGAGCGGGACATCAACAGATTCTCCGTCCAGAAGGATTTTATTAACCGGAATCTG
GTGGACACAAGATACGCTACTCGCGGCCTGATGAATCTGCTGCGATCCTATTTCCGG
GTGAACAATCTGGATGTGAAAGTCAAGTCCATCAACGGCGGGTTCACATCTTTTCTG
AGGCGCAAATGGAAGTTTAAAAAGGAGCGCAACAAAGGGTACAAGCACCATGCCG
AAGATGCTCTGATTATCGCAAATGCCGACTTCATCTTTAAGGAGTGGAAAAAGCTGG
ACAAAGCCAAGAAAGTGATGGAGAACCAGATGTTCGAAGAGAAGCAGGCCGAATC
TATGCCCGAAATCGAGACAGAACAGGAGTACAAGGAGATTTTCATCACTCCTCACC
AGATCAAGCATATCAAGGATTTCAAGGACTACAAGTACTCTCACCGGGTGGATAAA
AAGCCCAACAGAGAGCTGATCAATGACACCCTGTATAGTACAAGAAAAGACGATAA
GGGGAATACCCTGATTGTGAACAATCTGAACGGACTGTACGACAAAGATAATGACA
AGCTGAAAAAGCTGATCAACAAAAGTCCCGAGAAGCTGCTGATGTACCACCATGAT
CCTCAGACATATCAGAAACTGAAGCTGATTATGGAGCAGTACGGCGACGAGAAGAA
CCCACTGTATAAGTACTATGAAGAGACTGGGAACTACCTGACCAAGTATAGCAAAA
AGGATAATGGCCCCGTGATCAAGAAGATCAAGTACTATGGGAACAAGCTGAATGCC
CATCTGGACATCACAGACGATTACCCTAACAGTCGCAACAAGGTGGTCAAGCTGTC
ACTGAAGCCATACAGATTCGATGTCTATCTGGACAACGGCGTGTATAAATTTGTGAC
TGTCAAGAATCTGGATGTCATCAAAAAGGAGAACTACTATGAAGTGAATAGCAAGT
GCTACGAAGAGGCTAAAAAGCTGAAAAAGATTAGCAACCAGGCAGAGTTCATCGCC
TCCTTTTACAACAACGACCTGATTAAGATCAATGGCGAACTGTATAGGGTCATCGGG
GTGAACAATGATCTGCTGAACCGCATTGAAGTGAATATGATTGACATCACTTACCGA
GAGTATCTGGAAAACATGAATGATAAGCGCCCCCCTCGAATTATCAAAACAATTGC
CTCTAAGACTCAGAGTATCAAAAAGTACTCAACCGACATTCTGGGAAACCTGTATGA
GGTGAAGAGCAAAAAGCACCCTCAGATTATCAAAAAGGGCTAAGAATTC

[1805]

[1806] *실시예 35: Cas9 닉카아제 및 2 개의 가이드 RNA를 사용한 표적외 절단의 최소화*

[1807] Cas9는 RNA-안내된 DNA 뉴클레아제로서, 20bp RNA 가이드의 도움을 받아 게놈의 특정 위치에 표적화될 수 있다. 그러나, 가이드 서열은 가이드 서열과 DNA-표적 서열 사이에 일부 미스매치를 허용할 수 있다. 가이드 RNA가 가이드 서열과 상이한 몇 개의 염기를 가진 표적외 서열에 Cas9를 표적화하는 경우인 표적외 절단에 대한 가능성으로 인하여 이런 유연성은 바람직하지 않다. 모든 실험 용도 (유전자 표적화, 작물 조작, 치료 용도 등)에서, Cas9 매개 유전자 표적화의 특이성을 개선하고, Cas9에 의한 표적외 변형의 가능성을 감소시킬 수 있는 것이 중요하다.

[1808] 출원인은 표적외 변형 없이 게놈에 표적화된 이중가닥 브레이크를 촉진하기 위해서 2 개의 가이드 RNA와 조합하여 Cas9 닉카아제 돌연변이를 사용하는 방법을 개발했다. Cas9 닉카아제 돌연변이는 절단 활성을 불능화함으로써 Cas9 뉴클레아제로부터 생성될 수 있으며, 이로써 DNA 듀플렉스의 가닥이 둘 다 절단되는 대신 단지 하나의 가닥이 절단된다. Cas9 닉카아제는 Cas9 뉴클레아제의 하나 이상의 도메인에 돌연변이, 예를 들어 Ruvc1 또는 HNH를 유도함으로써 생성될 수 있다. 이들 돌연변이는, 제한은 아니지만, Cas9 촉매 도메인의 돌연변이를 포함할 수 있으며, 예를 들어 SpCas9에서 이들 돌연변이는 위치 D10 또는 H840에 있을 수 있다. 이들 돌연변이는, 제한은 아니지만, SpCas9에서는 D10A, E762A, H840A, N854A, N863A 또는 D986A를 포함할 수 있으며, 닉카아제는 다른 CRISPR 효소 또는 Cas9 오솔로그에서 상응하는 위치에 돌연변이를 유도함으로써 생성될 수 있다. 본 발명의 가장 바람직한 구현예에서, Cas9 닉카아제 돌연변이는 D10A 돌연변이를 가진 SpCas9 닉카아제다.

[1809] 이 방식에서, 이 작업은 Cas9 닉카아제와 조합된 각 가이드 RNA가 듀플렉스 DNA 표적의 표적화된 단일가닥 브레이크를 유도할 것이다. 각 가이드 RNA는 하나의 가닥에 흉터를 형성하므로, 순 결과는 이중가닥 브레이크가 된다. 이 방법이 표적외 돌연변이를 제거하는 이유는 가이드 서열과 높은 정도의 유사성을 가진 표적외 부위를 갖는 것이 매우 어렵기 때문이다(20 bp+2 bp(PAM) = 22 bp, 각 가이드, 및 2 개의 가이드에 대한 특이성은 어떤

표적외 부위가 44 bp의 상동성 서열에 근접해야 한다는 의미이다). 여전히 개별 가이드는 표적외를 가질 수 있지만, 표적외에는 단지 흉터가 형성될 뿐일 테고, 이것은 돌연변이유발 NHEJ 과정에 의해서 잘 수선되지 않을 것이다. 따라서, DNA 이중가닥 흉터형성의 복합화는 표적외 돌연변이유발 효과 없이 표적화된 DNA 이중가닥 브레이크를 도입하는 강력한 방식을 제공한다.

[1810] 출원인은 HEK293FT 세포와 Cas9(D10A) 닉카아제를 암호화하는 플라스미드 및 하나 이상의 가이드를 위한 DNA 발현 카세트의 공-트랜스펙션을 수반하는 실험을 수행했다. 출원인은 리포펙타민 2000을 사용하여 세포를 트랜스펙션했고, 트랜스펙션된 세포는 트랜스펙션 48 또는 72 시간 후에 수거되었다. 이중 흉터형성-유도 NHEJ가 앞서 설명된 서베이어 뉴클레아제 분석을 사용하여 검출되었다(도 51, 52 및 53).

[1811] 출원인은 두 가이드 RNA와 조합되었을 때 Cas9 닉카아제 돌연변이에 의한 효과적인 절단에 관한 변수들을 더 확인했으며, 이들 변수들은, 제한은 아니지만 5' 돌출부의 길이를 포함한다. 효과적인 절단은 적어도 26 염기쌍의 5' 돌출부에서 보고된다. 본 발명의 바람직한 구현예에서, 5' 돌출부는 적어도 30 염기쌍, 더 바람직하게 적어도 34 염기쌍이다. 절단을 위해 200 염기쌍 이하의 돌출부가 허용될 수 있지만, 100 염기쌍 미만의 5' 돌출부가 바람직하며, 50 염기쌍 미만의 5' 돌출부가 가장 바람직하다(도 54 및 55).

[1812] *실시예 36: Cas9를 사용한 포유류 뇌에서의 유전자 기능의 생체내 질의*

[1813] 생체내 뇌에서 뉴런의 게놈을 정확하게 조작하는 능력은 정상 및 질환-관련 뇌 공정에서 유전자의 기능의 빠른 해부를 가능하게 한다. 출원인은 성체 마우스 뇌에서 특이적인 세포-유형으로의 미생물 엔도뉴클레아제 Cas9 및 단일 가이드 RNA의 AAV-매개 전달이 단일 및 다중복합체 유전자 녹아웃을 용이하게 하였음을 나타낸다. 2 주 이내에, 메틸 CpG 결합 단백질(MeCP2)의 Cas9-매개 녹아웃은 표적화된 뉴런에서 형태학적 변화를 유도하였고, DNA 메틸트랜스페라제 단백질 패밀리(Dnmt1, Dnmt3a 및 Dnmt3b)의 다중복합체 녹아웃은 전후관계 공포-조건화(contextual fear conditioning)에서 마우스의 거동을 변경시켰다. 이와 함께, 출원인의 결과는 뇌에서의 유전자 기능의 빠른 역 유전자 연구를 용이하게 하는 Cas9의 잠재력을 입증하며, 시간이 걸리는 유전자이식 동물 모델의 사용에 대한 매력적인 대안을 제공한다.

[1814] 질환-연관 돌연변이를 가지는 유전자이식 동물 모델은 신경 정신 질환의 연구에 대단히 유용하며, 질환의 유전자 및 병리 생리학적 메커니즘을 해명하는 데 도움을 준다. 그러나, 일부 신경 정신 장애가 단일 유전자 결함과 연관된 것으로 알려져 있지만, 조현병, 자폐증 및 우울증을 포함하여 대부분의 인지 및 정신 장애는 다수의 유전자 변형과 연관되어 있다. 동시에 다수의 유전자 변형을 가지는 동물 모델의 생성은 특히 노동 집약적이고 다수의 세대를 거쳐서 시간이 소요되는 번식을 필요로 한다. S. 피오게네스(SpCas9)로부터의 CRISPR-연관 엔도뉴클레아제

[1815] Cas9는 진핵 세포를 복제하는 데 있어서 단일 그리고 다수의 유전자의 정확하고 효율적인 게놈 절단을 매개하는 것으로 나타났다. Cas9 뉴클레아제는 특이적인 게놈 유전자좌를 절단하는 것을 지시하여 표적화된 이중-가닥 절단을 유도할 수 있으며, 이는 프레임 이동 삽입/결실(삽입결실) 돌연변이를 초래한다(도 59A). 출원인은 뇌 과정에서 유전자 그룹 또는 개별 유전자의 기능을 연구할 뿐만 아니라 생체내 모노- 및 다중-유전자 장애를 모형화하기 위해 Cas9를 사용하는 것의 가능성을 탐색하고자 하였다.

[1816] 아데노-연관 바이러스(AAV)는 마우스 뇌로 재조합 유전자를 전달하는 데 일반적으로 사용된다. 출원인은 이중-벡터 AAV-기반 Cas9 및 sgRNA 전달 시스템을 설계하였다(도 56A). AAV 시스템의 주요 제한은 일정 크기의 이식 유전자 카세트(ITR 없이 약 4.5 kb)를 패키징하는 것이다. SpCas9의 크기는 약 4.2 kb이며, 이는 효율적이고 세포-유형 특이적 유전자 발현을 위해 필요한 다른 유전자 요소를 위하여 0.3 kb 미만을 남긴다. 출원인은 마우스 *Mecp2* 프로모터의 트렁케이션된 형태(235 bp, *sMecp2*) 및 짧은 폴리아데닐화 신호(48 bp, spA)가 배양된 마우스 피질 뉴런에서 Cas9의 강력한 발현을 구동하는 데 충분하다는 것을 확인하였다(도 56B; 도 59B 내지 D). 뉴런에서 sgRNA의 발현을 구동하기 위해, 출원인은 제2 sgRNA-암호화 벡터를 사용하였다. 추가적으로, 출원인은 sgRNA 전달을 표적화된 세포 핵의 동시 형광 표지화와 조합하였다. KASH 핵 막횡단 도메인과 녹색 형광 단백질(GFP)의 융합은 핵 외막으로 GFP를 유도하고, AAV 형질도입된 세포의 용이한 시각화를 가능하게 한다(도 56B).

[1817] 출원인은 먼저 일차 마우스 피질 뉴런을 사용하여 이러한 시험관내 이중-벡터 Cas9 전달 시스템의 전달 효율을 시험하였고, 80% 초과의 공동-형질도입 효율을 관찰하였다(도 56C). 중요한 것은, Cas9의 발현이 형질도입된 뉴런의 생존률 및 형태에 악영향을 미치지 않는다는 것이다(도 59D, E). 마우스 일차 뉴런에서 Cas9-매개 게놈 편집을 시험하기 위해, 출원인은 *Mecp2* 유전자를 표적화하였는데, 상기 *Mecp2* 유전자는 레트 증후군에서 주된 역할을 한다. MeCP2 결핍은 뉴런에서 가지돌기 나무 비정상 및 가시 형태 형성 결함과 연관되어 있는 것으로 나타

났으며, 형태학적 표현형 둘 다 레트 증후군이 있는 환자에서 발견되는 신경학적 증상에 기여하는 것으로 여겨진다.

[1818] *Mecp2*의 효율적인 표적화를 달성하기 위해, 출원인은 먼저 마우스 *Mecp2* 유전자좌의 엑손 3 이내 다수 부위를 표적화하는 여러 개의 sgRNA를 생성하였고(도 S2A,B), Neuro-2a 세포로부터의 서베이어™ 뉴클레아제 분석 결과를 기반으로 하여 가장 효율적인 sgRNA를 선택하였다(도 56D, 도 60). 시험관내 뉴런에서 출원인의 이중-벡터 시스템의 편집 효율을 평가하기 위해, 출원인은 시험관내에서 7일 째(7DIV)에 일차 마우스 피질 뉴런을 형질도입하였고, 형질 도입 7 일 후에 서베이어™ 뉴클레아제 분석을 사용하여 삽입결실률을 측정하였다(도 56E). 중요한 것은, Cas9 및 sgRNA가 공동-형질도입된 뉴런 배양물은 Cas9 단독이 형질도입된 뉴런과 비교하여 MeCP2 단백질 수준에서 대략 80% 감소를 나타내었다(도 56F; 도 61). 게다가, Cas9-단독 뉴런 집단과 비교하여, Cas9 및 sgRNA 공동-형질도입된 뉴런은 또한 변경된 가지돌기 나무 형태(도 56G 내지 I) 및 가시 밀도(도 69J, K)를 나타내었다.

[1819] 그 다음, 출원인은 Cas9가 살아 있는 마우스의 뇌에서 세포의 지정된 서브세트에서 안정적인 게놈 변형을 매개할 수 있는지 여부를 물어 보았다. AAV-형질도입된 세포를 구체적으로 분석하기 위해, 출원인은 형광 활성화 세포 정렬(FACS)을 사용하여 GFP-KASH 표지화 핵을 정제하는 방법을 개발하였다(도 57A, 도 62). 배양물에서 일차 뉴런과 유사하게, 출원인은 성체 마우스의 치아이랑으로 고역가의 Cas9 및 *Mecp2*-표적화 sgRNA를 가지는 AAV1/2의 혼합물(1:1 비율)을 공동-전달하였다(도 57B). 출원인은 바이러스 전달 4 주 후에 해마 과립 세포에서 벡터 둘 다의 높은 공동-형질도입 효율을 관찰하였다(도 57C). 생체내 *Mecp2* 유전자좌에서 게놈 변형의 효율을 시험하기 위해, 출원인은 GFP-KASH 표지화 핵을 정제하고, 정렬된 핵 집단에서 최대 34%까지의 삽입결실 빈도를 검출하였다(도 57D). *Mecp2* 유전자좌의 Cas9-매개 절단은, MeCP2 녹아웃 마우스에서 관찰된 효과와 유사하게, MeCP2 단백질 수준을 효율적으로 감소시켰고(도 57E 내지 G), 표적화된 뉴런 집단에서 가시 밀도를 감소시켰다(도 57H, I). 이러한 결과는 생체내 포유류 뇌에서 표적화된 유전자 녹아웃을 용이하게 하기 위한 Cas9의 다능성을 입증한다.

[1820] Cas9 시스템의 한 가지 주요 장점은 다중복합체 게놈 편집을 용이하게 하는 능력이다. 살아 있는 동물의 뇌에 다수 유전자의 안정적인 녹아웃을 도입하는 것은, 생리학적 및 신경병리학적 조건에서 다유전자 메커니즘의 인과적 질의와 같은, 잠재적으로 방대한 적용을 가질 것이다. 뇌에서의 다중복합체 게놈 편집의 가능성을 시험하기 위해, 출원인은 핵 표적화를 위한 GFP-KASH와 함께 연쇄적인 3 개의 sgRNA로 이루어지는 다중복합체 sgRNA 발현 벡터를 설계하였다(도 58A). 출원인은 *Dnmt1*, *Dnmt3a* 및 *Dnmt3b*로 이루어지는 DNA 메틸트랜스페라제 유전자 패밀리(DNMT)를 표적화하는 sgRNA를 선택하였다. *Dnmt1* 및 *3a*는 성체 뇌에서 발현되는 반면, *Dnmt3b*는 주로 신경발달 동안 발현된다. DNMT 활성이 뇌에서의 DNA 메틸화를 변경시키고, Dnmt3a 및 Dnmt1이 시냅스 가소성 및 학습 및 기억 형성에 필요한 것으로 이전에 밝혀졌다. 출원인은 3 가지 모두의 표적화된 유전자에 대한 고도의 동시 DNA 절단을 위해 높은 변형 효율을 갖는 개별적인 sgRNA 및 최적화된 가이드 조합을 선택하였다(도 58B; 도 63).

[1821] 생체내에서 다중복합체 게놈 편집의 효능을 시험하기 위해, 출원인은 성체 마우스의 등쪽 및 배쪽 치아이랑으로 고역가의 Cas9- 및 sgRNA-발현 AAV의 혼합물을 정위 전달하였다. 4 주 후, 해마를 해부하고, 표적화된 세포 핵을 FACS를 통해 정렬하였다. 딥 시퀀싱을 사용하여, 출원인은 3 가지 모든 표적화된 유전자좌에서 약 20% 내지 70% 범위의 높은 수준의 삽입결실 형성을 관찰하였다(도 58C, 도 64). 모두 성체 뇌에서 높게 발현되는 Dnmt3a 및 Dnmt1은 바이러스 전달 약 12 주 후에 해부된 치아이랑 조직에서 상당히 감소된 단백질 수준을 나타낸다(도 58D). 성체 뇌에서 *Dnmt3b*의 낮은 발현으로 인해, 출원인은 본 분석에서 Dnmt3b 단백질을 검출할 수 없었다.

[1822] Cas9를 이용한 이전 연구는 sgRNA와 부분적으로 매치되는 게놈 유전자좌가 표적외 삽입결실 형성을 초래할 수 있음을 나타내었다. 상기 예측된 표적외 유전자좌의 삽입결실 분석은 삽입결실 형성의 0% 내지 1.6% 비율을 나타내었고, 이는 Cas9 변형이 매우 특이적인 것을 입증한다(표 2). 관찰된 낮은 비율의 표적외의 것은 *sMecp2* 프로모터 및 짧은 polyA 신호로부터 생기는 상대적으로 약한 수준의 Cas9 발현으로 인한 것일 수 있다.

표 2

[1823] **DNMT 표적화를 위한 표적외 분석**

유전자	GI	잠재적인 표적외 서열	MLE(%)	SEM	(서열 번호 ___)
nt*Abca1*	NM_013454	GGAGCTGGAGCTGTTCACGTTGG	0.0000	0.00	
Mctp1	NM_030174	CGGGCAGCAGATGTTCGCGTAGG	0.0806	0.08	
Exd2	NM_133798	AGGGCTTGAGATGTTCGGGCTGG	0.0612	0.06	
Pik3r6	NM_001004435	CCGGCTGGGGCTGTCCTCGCTAG	0.0000	0.00	
Sobp	NM_175407	CGGGGTGCAGCTGCTCACGCCAG	0.0000	0.00	
Vac14	NM_146216	CTGGCGGGAGCTGGTCGCGTGAG	0.0083	0.00	
t*Efemp2*	NM_021474	TGAGCATGGGCCGCTGGCGGTGG	0.0050	0.01	
Bmpr1b	NM_001277217	ATGGCATAGGCCGCTGACAGAGG	0.0117	0.01	
Syce1	NM_001143765	TTGGCATGGTGAGCTGGCGGGGG	0.0067	0.00	
Atp8b3	NM_026094	TGGGCAGGGGTCTCTGAGGGCAG	0.0067	0.01	
Rdh11	NM_021557	TTGGCATGGGTCTCTTACCAAGG	0.0017	0.00	
t*Hecw2*	NM_001001883	ACATGGTTCCAGTGGGTATGTAG	0.0000	0.00	
Plekhg3	NM_153804	GGAGGTGGGCAGCGGGTATGTAG	0.0954	0.01	
Cdc25b	NM_001111075	AGAAGGTCCCCGCGGGCATGGAG	0.2421	0.12	
Top1mt	NM_028404	GGAGGGAACCAGCCGGTATGGGG	0.0167	0.01	
Sesn2	NM_144907	AGAGAGTGGCAGTGGGTAAGCAG	0.0000	0.00	
Ncan	NM_007789	AGAGGTGGCCAGCGGGCAGGAAG	0.0017	0.00	
Nacad	NM_001081652	TGAGGGGGCCAGCTGGGATGCAG	1.6254	0.76	

[1824] 이전 연구는 Dnmt3a 및 Dnmt1의 녹다운이 해마-의존성 기억 형성에 영향을 줄 수 있음을 시사하였다. 결과로서, 출원인은 기억 형성의 획득 및 경화(consolidation)에 대한 Cas9-매개 삼중 녹아웃(Dnmt3a, Dnmt1 및 Dnmt3b)의 효과를 연구하기 위한 전후관계 공포-조건화 행동 시험을 수행하였다. 출원인은 기억 획득 단계에서 대조군과 삼중 녹아웃 마우스 사이에 어떠한 차이도 관찰하지 않았으나, 녹아웃 마우스는 훈련 상황 조건 하에서 시험되는 경우 손상된 기억 경화를 나타내었다(도 58E, F). 이러한 효과는 마우스가 변경된 상황에서 시험된 경우에 폐지되었으며, 이는 전후관계 학습이 출원인의 마우스 모델에서 손상되었음을 입증한다. 출원인의 결과는 치아이랑 뉴런에서의 DNMT 패밀리 일원의 CRIPSR-Cas9-매개 녹아웃이 행동 임무에서 유전자의 기능을 탐지하기에 충분한 것을 입증한다.

[1825] 이와 함께, 출원인의 결과는 CRISPR-Cas9 시스템의 생체내 전달이 질환 모델을 생성하고 다중복합체 게놈 편집에 의해 마우스 행동을 조작하기 위한 정확하고 유연적이며 매우 효율적인 기술을 나타낸다는 것을 입증한다. 여기에서 나타낸 Cas9 시스템의 생체내 적용은 기초과학뿐만 아니라, 생물기술 및 의학에서의 광범위한 적용을 가질 것이다.

[1826] DNA 작제물: Cas9 표적 선택 및 단일 가이드 RNA(sgRNA)의 생성을 위해, 20-nt 표적 서열을 5'-NGG PAM 서열에 앞서도록 선택하였다. 표적외 효과를 최소화시키기 위해, CRIPSR 설계 도구를 사용하였다(웹사이트 tools.genome-engineering.org에서 이용가능함). sgRNA를 정방향 프라이머(5'-cgcacgcgtaattcgaacgctgacgtcatc-3') 및 20-nt DNA 표적 부위(***굵은 글씨***)와 함께 sgRNA를 함유하는 역방향 프라이머(5'-cacacgcgtAAAAAAgcaccgactcggtgccacttttcaagttgataacggactagccttattttaacttgctaTTTCtagctctaaaac***NNNNNNNNNNNNNNNNNNNC***GGTGTTTCGTCCTTTCCAC-3'; 서열 번호 ___)와 함께 주형으로서 U6 프로모터를 사용하여 PCR 증폭시켰다. EGFP-KASH 작제물은 인간 *시냅신* 프로모터(*hSyn*) 하에서 AAV 백본으로 암호화 카세트를 클로닝하기 위한 PCR 주형으로 사용하였다. 다음으로, U6-*Mecp2*_sgRNA 암호화 서열을 *Mlu*I 부위를 사용하여 도입시켰다. 다중복합체 유전자 표적화 전략을 위해, 개별적 sgRNA를 상기 기재된 바와 같이 PCR 증폭시켰다. 3 가지 모두의 sgRNA를 골든 게이트(Golden Gate) 클로닝 전략을 사용함으로써 PCR 증폭된 *hSyn*-GFP-KASH-bGHpA 카세트와 결찰시켰다(도 58A 참조). PCR 증폭 후, 3 sgRNA 및 *hSyn*-GFP-KASH-bGH pA를 함유하는 골든 게이트 결찰 산물을 AAV 백본으로 클로닝하였다. 모든 수득된 작제물을 시퀀싱 확인하였다. 뉴런에서 Cas9 발현을 구동시키는 최적 프로모터 서열을 찾기 위해, 출원인은 h*Syn1*, 마우스 트렁케이션된 *Mecp2* (*sMecp2*), 및 트렁케이션된 래트

Map1b(*rMap1b*) 프로모터 서열을 시험하였다(도 59B 참조). 하기 프라이머를 프로모터 영역을 증폭시키기 위해 사용하였다: hSyn_F: 5'-GTGTCTAGACTGCAGAGGGCCCTG-3'(서열 번호 ___); hSyn_R: 5'-GTGTCGTGCCTGAGAGCGCAGTCGAGAA-3'(서열 번호 ___); mMecp2_F 5'-gagaagcttAGCTGAATGGGGTCCGCCTC-3'(서열 번호 ___); mMecp2_R 5'-ctcaccggtGCGCGCAACCGATGCCGGGACC-3'(서열 번호 ___); rMap1b-283/-58_F 5'-gagaagcttGGCGAAATGATTTGCTGCAGATG-3'(서열 번호 ___); rMap1b-283/-58_R 5'-ctcaccggtGCGCGCGTCGCCTCCCCCTCCGC-3'(서열 번호 ___). 래트 *map1b* 프로모터의 또 다른 트렁케이션을 하기 올리고와 함께 어셈블리시켰다: 5'-agcttCGCGCCGGGAGGAGGGGGGACGCAGTGGGCGGAGCGGAGACAGCACCTTCGGAGATAATCCTTTCTCCTGCCGCAGAGCAGAGGAGCGGCGGGAGAGGAACACTTCTCCCAGGCTTTAGCAGAGCCGGa-3'(서열 번호 ___) 및 5'-ccggtCCGGCTCTGCTAAAGCCTGGGAGAAGTGTTCCTCTCCCGCCGCTCCTCTGCTCTGCGGCAGGAGAAAGGATTATCTCCGAAGGTGCTGTCTCCGCTCCGCCCACTGCGTCCCCCCTCCTCCCGGCGCGa-3'(서열 번호 ___). 짧은 합성 폴리아데닐화 신호(spA)(*3*)를 하기 올리고를 사용하여 어셈블리시켰다: 5'-aattcAATAAAAGATCTTTATTTTCATTAGATCTGTGTGTTGGTTTTTTG

[1827] TGTgc-3'(서열 번호 ___) 및 5'- ggccgcACACAA

[1828] AAAACCAACACACAGATCTAATGAAAATAAAGATCTTTTATTg-3'(서열 번호 ___). EF1α 프로모터의 제어 하의 플라스미드 암호화 적색 형광 단백질(mCherry)을 Lipofectamine®2000(Life Technologies)을 이용한 뉴런 트랜스펙션을 위해 이용하였다.

[1829] 세포주 배양 및 트랜스펙션: Neuro-2a(N2a) 세포를 5% 소 태아 혈청(BSA)을 함유하는 DMEM에서 성장시켰다. HEK293FT 세포에 대해, 10% 소 태아 혈청(FBS)을 함유하는 DMEM을 사용하였다. 세포를 5% CO_2 대기, 37℃에서 유지시켰다. 세포를 제조업체의 프로토콜에 따라 Lipofectamine®2000 또는 폴리에틸렌이민(PEI) "MAX" 시약(Polysciences)을 사용하여 트랜스펙션시켰다.

[1830] 농축된 AAV 벡터의 생성: 고역가 AAV1/2 입자를 동등한 비율의 AAV1 및 AAV2 혈청형 플라스미드 및 pDF6 헬퍼 플라스미드를 사용하여 생성시키고, 헤파린 친화성 컬럼에서 정제하였다. 바이러스 입자의 역가측정을 qPCR에 의해 수행하였다. 고역가 AAV1 입자를 UNC Vector Core Services(University of North Carolina at Chapel Hill)에 의해 생성시켰다. DMEM 내 저역가 AAV1 입자를 문헌[S. Konermann et al., Optical control of mammalian endogenous transcription and epigenetic states. Nature 500, 472 (Aug 22, 2013)]에 이전에 기재된 바와 같이 생성시켰다. 간단히, HEK293FT 세포를 PEI "MAX"를 사용하여 트랜스젠 플라스미드, pAAV1 혈청형 플라스미드 및 pDF6 헬퍼 플라스미드로 트랜스펙션시켰다. 배양 배지를 48 시간 후에 수집하고, 0.45 μm PVDF 필터(Millipore)를 통해 여과시켰다.

[1831] 일차 피질 뉴런 배양: 조직 배양을 위한 뉴런을 수득하기 위해 사용된 동물을 MIT 동물관리위원회(MIT CAC)에 의해 승인된 프로토콜에 따라 희생시켰다. 문헌[G. Banker, K. Goslin, Developments in neuronal cell culture. Nature 336, 185 (Nov 10, 1988)]에 기재된 바와 같이 16 일의 배아 마우스 뇌로부터 일차 배양물을 제조하였다. 어느 한 성별의 배아를 사용하였다. 세포를 폴리-D-리신(PDL) 코팅된 24-웰 플레이트(BD Biosciences) 또는 라미닌/PDL 코팅된 커버슬립(VWR) 상에 플레이팅하였다. 배양물을 B27, 글루타맥스(Life Technologies) 및 페니실린/스트렙토마이신 혼합물이 보충된 Neurobasal 배지에서 37℃ 및 5% CO_2에서 성장시켰다.

[1832] AAV 형질도입을 위해, 500 μl Neurobasal 배양 배지 중 피질 뉴런을 HEK293FT 세포로부터의 300 μl(1:1 비율의 이중 감염) AAV1-함유 적응용 배지를 이용하여 7 DIV로 인큐베이션하였다. 형질도입 1 주 후에, 뉴런을 하류 공정을 위해 수거하거나, 면역형광 염색 또는 형태학적 분석을 위해 4% 파라포름알데하이드 중에 고정시켰다.

[1833] 뉴런 형태의 시각화를 위해, DIV7의 세포를 문헌[L. Swiech et al., CLIP-170 and IQGAP1 cooperatively regulate dendrite morphology. The Journal of neuroscience: the official journal of the Society for Neuroscience 31, 4555 (Mar 23, 2011)]에 이전에 기재된 바와 같이 1 주 동안 Lipofectamine®2000(Life Technologies)을 사용하여 EF1α-mCherry 발현 벡터로 트랜스펙션시켰다. 전체 가지돌기 길이의 측정을 위해, 개별 뉴런의 모든 가지돌기를 ImageJ 소프트웨어를 사용하여 추적하였다. 일차 가지돌기의 수, 가지돌기 첨단 및 숄(Sholl) 분석의 정량화를 40x 대물렌즈에서 형광 현미경(Zeiss AxioCam Ax10 현미경, Axiocam MRm 카메라)으로 획득된 이미지에 대해 수행하였다. 가지돌기 수에 대해, 10 μm보다 긴 모든 비-축삭 돌출의 말단을 계수하였다. 숄 분석을 위해, 직경에 있어서 5 μm 단계를 갖는 동심원이 세포체 주위에 자동적으로 그려졌

고, 각각의 원을 가로지르는 가지돌기의 수를 Sholl 플러그-인을 갖는 ImageJ 소프트웨어를 이용하여 계수하였다.

[1834] 마우스 뇌로의 AAV1/2의 정위 주사: MIT CAC는 본원에 기재된 모든 동물 절차를 승인하였다. 성체(12 주령 내지 16 주령) 수컷 C57BL/6N 마우스를 100 mg/kg 케타민 및 10 mg/kg 자일라진의 복막내(i.p.) 주사에 의해 마취시켰다. 선행 마취를 제공하였다(Buprenex, 1 mg/kg, i.p.). 개두술을 승인된 절차에 따라 수행하였고, 1 μℓ의 1:1 AAV 혼합물($1x10^{13}$ Vg/mℓ의 *sMecp2*-SpCas9; $6x10^{12}$ Vg/mℓ의 DNMT 3xsgRNA; $3x10^{12}$ Vg/mℓ 내지 $5x10^{12}$ Vg/mℓ의 h*Syn*-GFP-KASH)을 등쪽 치아이랑(전방/후방: -1.7; 내외측: 0.6; 등쪽/배쪽: -2.15) 및/또는 배쪽 치아이랑(전방/후방: -3.52; 내외측: 2.65; 등쪽/배쪽: -3)에 주사하였다. 절개를 봉합시키고, 적절한 수술후 진통제(멜록시캄, 1 mg/kg 내지 2 mg/kg)를 수술 후 3 일 동안 투여하였다.

[1835] 뇌 조직으로부터 세포 핵의 정제: 전체 해마 또는 치아이랑을 얼음 저온 DPBS(Life Sciences)에서 신속하게 해부하고, 드라이아이스 상에서 쇼크 동결시켰다. 조직을 2 mℓ Dounce 균질화기(Sigma)를 사용하여 2 mℓ 얼음-저온 균질화 완충액(HB)(320 mM 수크로스, 5 mM CaCl, 3 mM $Mg(Ac)_2$, 10 mM 트리스 pH7.8, 0.1 mM EDTA, 0.1 % NP40, 0.1 mM PMSF, 1 mM 베타-머캅토에탄올)에서, 막자 A로 25 회 후, 막자 B로 25 회로 가볍게 균질화시켰다. 다음으로, 3 mℓ의 HB를 전체 5 mℓ까지 첨가하고, 5 분 동안 얼음 상에서 유지시켰다. 구배 원심분리를 위해, 5 mM CaCl, 3 mM $Mg(Ac)_2$, 10 mM 트리스 pH 7.8, 0.1 mM PMSF, 1 mM 베타-머컵토에탄올을 함유하는 5 mℓ의 50% OptiPrep™ 밀도 구배 배지(Sigma)를 첨가하고, 혼합하였다. 용해질을 코니컬 30 mℓ 원심분리 튜브(Beckman Coulter, SW28 로터) 내의 10 mℓ의 29% 등장성 OptiPrep™ 용액의 상부에 가볍게 로딩하였다. 샘플을 4℃에서 30 분 동안 10,100 xg(7,500 rpm)에서 원심분리시켰다. 상청액을 제거하고, 핵 펠렛을 65 mM 베타-글리세로포스페이트(pH 7.0), 2 mM $MgCl_2$, 25 mM KCl, 340 mM 수크로스 및 5% 글리세롤에 가볍게 재현탁시켰다. 정제된 핵의 수 및 품질을 명시야 현미경검사를 사용하여 조절하였다.

[1836] 세포 핵 분류: 정제된 GFP-양성(GFP^+) 및 음성(GFP^-) 무결함 핵을 Vybrant® DyeCycle™ Ruby 염색(1:500, Life Technologies)으로 공동-표지화시키고, BD FACSAria III(Koch Institute Flow Cytometry Core, MIT)을 사용하여 분류하였다. GFP^+ 및 GFP^- 핵을 1% BSA로 코팅되고, 400 μℓ의 재현탁 완충액(65 mM 베타-글리세로포스페이트 pH 7.0, 2 mM $MgCl_2$, 25 mM KCl, 340 mM 수크로스 및 5% 글리세롤)을 함유하는 1.5 mℓ 에펜도르프 튜브에서 수집하였다. 분류 후, 모든 샘플을 얼음 상에서 유지시키고, 4℃에서 20 분 동안 10,000 xg에서 원심분리시켰다. 핵 펠렛을 -80℃에서 저장하거나, 하류 공정을 위해 직접 사용하였다.

[1837] 게놈 DNA 추출 및 서베이어™ 검정: sgRNA의 기능 시험을 위해, 50% 내지 70% 컨플루언스(confluence) N2a 세포를 단일 PCR 증폭된 sgRNA 및 Cas9 벡터와 함께 공동-트랜스펙션시켰다. Cas9로만 트랜스펙션된 세포를 음성 대조군으로 제공하였다. 세포를 트랜스펙션 48 시간 후에 수거하고, DNA를 제조업체의 프로토콜에 따라 DNeasy Blood & Tissue Kit(Qiagen)를 사용하여 추출하였다. AAV1 형질도입된 일차 뉴런으로부터 게놈 DNA를 분리시키기 위해, DNeasy Blood & Tissue Kit를 제조업체의 설명서에 따라 AAV 형질도입 7 일 후에 사용하였다.

[1838] 분류된 핵 또는 해부된 조직을 30 분 동안 55℃에서 용해 완충액(10 mM 트리스, pH 8.0, 10 mM NaCl, 10 mM EDTA, 0.5 mM SDS, 프로테이나제 K(PK, 1 mg/mℓ) 및 RNAse A)에서 용해시켰다. 다음으로, 클로로포름-페놀 추출을 수행한 후, 표준 절차에 따라 에탄올을 이용하여 DNA 침전시켰다. DNA를 TE 완충액(10 mM 트리스 pH 8.0, 0.1 mM EDTA)에 최종적으로 재현탁시키고, 하류 분석에 이용하였다. 개별적 sgRNA의 기능 시험을 표 3에 나열된 PCR 프라이머를 사용하여 서베이어™ 뉴클레아제 검정(Transgenomics)에 의해 수행하였다. 밴드 강도 정량화를 문헌[F. A. Ran et al., Genome engineering using the CRISPR-Cas9 system. Nature protocols 8, 2281 (Nov, 2013)]에 이전에 기재된 바와 같이 수행하였다.

표 3

[1839] **서베이어™ 검정에서 사용된 PCR 프라이머**

표적	정방향 프라이머 서열(5'-3')	(서열 번호 ___)	역방향 프라이머 서열(5'-3')	(서열 번호 ___)
Mecp2	GGTCTCATGTGTGGCACTCA		TGTCCAACCTTCAGGCAAGG	
Dnmt3a	ATCCCTCCTCAGAGGGTCAGC		TACCTCATGCACAGCTAGCACC	

Dnmt1	TTCGGGCATAGCATGGTCTTCC		GTTCTATTTCAGAGGGCTGATCCC	
Dnmt3b	GTTCTGAGCCGCACAGTTTGG		GGATAAGAAGGGACAATACAGG	

[1840] 면역형광 염색

[1841] 세포 배양: 일차 뉴런의 면역형광 염색을 위해, 세포를 실온에서 20 분 동안 4% 파라포름알데하이드(PFA)를 이용하여 바이러스 전달 7 일 후에 고정시켰다. PBS를 이용하여 3 회 세척한 후, 세포를 실온에서 30 분 동안 PBS 중 5% 일반 염소 혈청(NGS)(Life Technologies), 5% 당나귀 혈청(DS)(Sigma) 및 0.1% Triton-X100(Sigma)로 블로킹시켰다. 세포를 실온에서 1 시간 또는 4℃에서 밤새 2.5% NGS, 2.5% DS 및 0.1% Triton-X100 중 일차 항체와 함께 인큐베이션하였다. PBST로 3 회 세척한 후, 세포를 실온에서 1 시간 동안 이차 항체와 함께 인큐베이션하였다. 최종적으로, 커버슬립을 DAPI를 갖는 VECTASHIELD HardSet 마운팅 배지(Vector Laboratories)를 사용하여 마운팅시키고, Zeiss AxioCam Ax10 현미경 및 Axiocam MRm 카메라를 사용하여 영상화시켰다. 이미지를 Zen 2012 소프트웨어(Zeiss)를 사용하여 처리하였다. 정량화를 ImageJ 소프트웨어 1.48 h 및 Neuron 검출기 플러그인을 사용하여 수행하였다.

[1842] 마우스를 치사 용량의 케타민/자일라진에 의해 바이러스 전달 4 주 내지 8 주 후에 희생시키고, PBS 및 이후 PFA로 경심 관류시켰다. 고정된 조직을 진동절편기(vibratome)(Leica, VT1000S)를 사용하여 절편화시켰다. 다음으로, 30 ㎛ 절편을 소듐 시트레이트 완충액(10 mM 트리-소듐 시트레이트 데하이드레이트, 0.05% Tween20, pH 6.0)에서 2 분 동안 비등시키고, 20 분 동안 실온에서 냉각시켰다. 절편을 1 시간 동안 TBST(137 mM NaCl, 20 mM Tris pH 7.6, 0.2% Tween-20) 중 4% 일반 염소 혈청(NGS)으로 블로킹시켰다. 파라핀 절편을 박절기(microtom)(Leica RM2125 RTS)를 사용하여 8 ㎛로 절단하고, 문헌[A. V. Tzingounis et al., The KCNQ5 potassium channel mediates a component of the afterhyperpolarization current in mouse hippocampus. Proceedings of the National Academy of Sciences of the United States of America 107, 10232 (Jun 1, 2010)]에 이전에 기재된 바와 같이 염색하였다[14].

[1843] 절편을 4℃에서 밤새 4% NGS와 함께 TBST에 희석된 일차 항체와 함께 인큐베이션하였다. TBST 중에서 3 회 세척 후, 샘플을 이차 항체와 함께 인큐베이션하였다. TBST로 3 회 세척 후, 절편을 DAPI를 갖는 VECTASHIELD HardSet 마운팅 배지를 사용하여 마운팅시키고, 동일초점 현미경(Zeiss LSM 710, Ax10 ImagerZ2, Zen 2012 소프트웨어)으로 시각화시켰다.

[1844] 다음과 같은 일차 항체를 사용하였다: 토끼 항-Dnmt3a(Santa Cruz, 1:100); 토끼 항-MeCP2(Millipore, 1:200); 마우스 항-NeuN(Millipore, 1:50 내지 1:400); 닭 항-GFAP(Abcam, 1:400); 마우스 항-Map2(Sigma, 1:500); 닭 항-GFP(Aves labs, 1:200 내지 1:400); 마우스 항-HA(Cell Signaling, 1:100). 이차 항체: AlexaFluor®488, 568 또는 633(Life Technologies, 1:500 내지 1:1,000).

[1845] 골지-콕스 염색: 골지-콕스 염색을 FD Rapid GolgiStain 키트(FD NeuroTechnologies)를 사용하여 수행하였다. *Mecp2* 유전자좌를 표적화하는 CRISPR-Cas9의 주사 1 주 내지 4 주 후 성체 수컷 마우스를 MIT CAC 프로토콜에 따른 경추 파열법에 의해 희생시켰다. 해부 직후, 뇌를 제조업체의 프로토콜에 따라서 처리하였다. 세포의 형태를 100x 대물렌즈 하 광학 현미경(전동 미세 초점 조절이 있는 Zeiss Axioplan 2, OcraER 디지털 카메라)을 사용하여 치아이랑의 100 μm 두께 극저온-절편(cryo-section)에 대하여 분석하였다. 가시 밀도는 MetaMorph4 소프트웨어를 사용하여 수동으로 계수하였다.

[1846] LIVE/DEAD® 검정의 정량화: 대조군 및 형질도입된 일차 뉴런을 제조업체의 설명서에 따라 LIVE/DEAD® 검정(Life technologies)을 사용하여 염색하였다. GFP-KASH 발현으로부터 GFP-신호 간섭을 피하기 위해, 세포를 DEAD(에티디움 동종이합체) 및 DAPI(모든 세포) 단독에 대해 염색하였다. 염색된 세포를 형광 현미경검사를 사용하여 영상화시키고, DEAD, GFP 및 DAPI 양성 세포를 ImageJ 1.48h 소프트웨어 및 Neuron 검출기 플러그인을 사용함으로써 계수하였다.

[1847] 웨스턴 블롯 분석: 형질도입된 일차 피질 뉴런(바이러스 전달 7 일 후) 및 형질도입된 조직 샘플(바이러스 전달 4 주 후)을 0.1% SDS 및 프로테아제 억제제(Roche, Sigma)를 함유하는 50 ㎕의 얼음-저온 RIPA 완충액(Cell Signaling) 중에서 용해시켰다. 세포 용해질을 Bioruptor 초음파파쇄기(Diagenode)에서 5 분 동안 초음파 파쇄시키고, 단백질 농도를 BCA 단백질 검정 키트(Pierce Biotechnology, Inc.)를 사용하여 결정하였다. 단백질 용해질을 SDS-PAGE 샘플 완충액에 용해시키고, 4% 내지 15% 트리스-HCl 겔(Bio-Rad) 상에서 환원 조건 하에서 분

리시키고, 다음과 같은 일차 항체를 사용한 웨스턴 블롯팅에 의해 분석하였다: 토끼 항-Dnmt3a(Santa Cruz, 1:500), 마우스 항-Dnmt1(Novus Biologicals, 1:800), 토끼 항-Mecp2(Millipore, 1:400), 토끼 항-튜불린(Cell Signaling, 1:10,000) 후, 이차 항-마우스 및 항-토끼 HRP 항체(Sigma-Aldrich, 1:10,000). GAPDH를 항-GAPDH 항체에 커플링된 토끼 HRP(Cell Signaling, 1:10,000)로 직접 시각화시켰다. 튜불린 또는 GAPDH를 로딩 대조군으로 제공하였다. 블롯을 ImageLab 4.1 소프트웨어(BioRad)와 함께 ChemiDoc™ MP 시스템으로 영상화시키고, ImageJ 소프트웨어 1.48h를 사용하여 정량하였다.

[1848] 지연 전후관계 공포 조건화(Delay Contextual Fear Conditioning; DCFC): 12주령의 C57BL/6N 수컷 마우스의 등쪽 및 배쪽 치아이랑으로의 양쪽 Cas9/DNMT 3xsgRNA 전달 8 주 후, 7 일 동안 동물을 실험자 및 행동 실험실(behavior room)에 익숙하게 하였다. Cas9/GFP-KASH 주사된 한배 새끼를 대조군으로 제공하였다. DCFC 1 일에서, 시험 전 및 후에 마우스 우리를 청각 신호로부터 보호하기 위해 분리된 앤터룸(anterroom)에 배치하였다. 개별적 마우스를 FC 챔버(Med Associates Inc.)에 배치하고, 12 분 익숙화(habituation) 기간을 수행하였다. 익숙화 후, 마우스를 이들의 홈케이지(homecage)로 다시 배치하였다. 다음날(훈련일) 개별적 마우스를 챔버에 배치하고, 4 분 동안 익숙하게 하였다. 또 다른 20 초(울림 전) 간격 후, 85 dB, 2.8 kHz 수준의 울림(청각 신호)을 20 초 동안 제시한 후, 18 초 지연 간격 후, 발-충격(0.5 mA, 2 초)을 제공하였다. 발-충격 후, 20 초 울림전 기간과 함께 다음의 동일한 시험 시작에 앞서 40 초 간격(울림/충격 후)을 선행시켰다. 훈련 시험을 6 회 반복한 후, 마우스를 이들의 홈케이지로 다시 배치하였다. 3 일(시험일)에서, 마우스를 먼저 3 분 동안 훈련 상황 챔버에 배치하였다. 다음으로, 마우스를 20 초 간격 후, 20 초 울림 및 60 초 울림 후 간격으로 시작하는 4x 100 초 시험 시도를 수행하였다. 최종적으로, 마우스를 변경 상황-검사 챔버(편평한 바닥 대 격자, 4합체 대 7합체 챔버, 바닐린 센트(vanillin scent))에 배치하고, 시험 시도를 반복하였다. 얼어붙는 행동을 기록하였고, 수작업으로 오프라인(off-line) 방식으로 분석하였고, Noldus EthoVision XT 소프트웨어(Noldus Information Technology)를 이용하여 확인하였다.

[1849] 심층 시퀀싱 분석 및 삽입결실 검출: 뇌에서 CRISPR-Cas9에 의해 표적화되는 DNMT 패밀리 유전자에 대한 잠재적인 표적외를 찾기 위해 CRISPR Design Tool(웹사이트 crispr.mit.edu/에서 이용가능함)을 사용하였다. 치아이랑으로부터의 표적화된 세포 핵을 바이러스 전달 12 주 후에 FACS 분류하고, 게놈 DNA를 상기 기재된 바와 같이 정제하였다. 각각의 관심대상 유전자에 대해, CRISPR 표적 부위에 측접한 게놈 영역을 융합 PCR 방법에 의해 증폭하여, Illumina P5 어댑터뿐만 아니라 독특한 샘플-특이적 바코드를 표적 앰플리콘에 부착시켰다(표적내 및 표적외 프라이머에 대해서는, 표 4 참조). 바코드화되고 정제된 DNA 샘플을 Qubit 2.0 Fluorometer(Life Technologies)에 의해 정량하고, 등몰비로 풀링시켰다. 이후, 시퀀싱 라이브러리를 300 bp 판독 길이를 갖는 Illumina MiSeq Personal Sequencer(Life Technologies)로 시퀀싱하였다.

표 4

[1850] **표적내 및 표적외 게놈 유전자좌 증폭을 위해 사용된 프라이머**

유전자좌	정방향 프라이머 서열(5'-3')	(서열 번호 ___)	역방향 프라이머 서열(5'-3')	(서열 번호 ___)
Dnmt1	GCCGGGGTCTCGTTCAGAGCT		CTACCGCCTGCGGACATGGT	
Dnmt3a	CCTGTCTCTCTGTCCTAGGGCTCC		CCGTTTGCTGATGTAGTAGGGGTCC	
Dnmt3b	CCCACAGGAAACAATGAAGGGAGAC		CATCCTTCGTGTCTGAGGACTGGTC	
Abca1	CCCTGACACCAGCTGTTCAGCAC		CTCTGGGTGACCACACACGATGC	
Mctp1	GAGCAGGCAGAGCCGAGCAAG		GGAGAGCGTCCGCCAGGAG	
Exd2	GGGTCTTGTTGTGAGTAGGGTGTG		GAAGCTCTCTTAACTACTGTTC	
Pik3r6	CCTGGAATACTATTTCCACGCCG		CAGGCCCTAGCAGCGAGCAG	
Sobp	GCAGCACACTCCACCCTCACAT		GGAAGGGGCTTTCCTCCGAGC	
Vac14	CGGCGTCACGTGACCTGAGTAAC		GCTCCGACCCTGCTCTCCCA	
Efemp2	GTGTCTGCCTCGCTCTGCTGC		CCTGTTCATCAGGCTCGTAGCCC	
Bmpr1b	CTATCTGAAATCCACCACCTTAGACGC		CGATTGCTGGCTTGCCTTGAG	
Syce1	GCCTGAGGGGGCCAGAGGT		GGTTCGCGTCCGCCCGCGTGAT	
Atp8b3	GGGACTCCCCGGGTGGTG		GAGAGGTGGTCCTGTCGCCTATG	
Rdh11	GACCCTGTGTTTCAAGTCTCTCTG		CCCAGCAGGTCACAGCTGACATC	
Hecw2	GGCCATCCAGTACATTCAATACG		AGCACAGTATGTATTCTATAAAATAATACGAC	

Plekhg3	GCAGAAGCCGTGACTCACAGCA		GTGGGAGGGGACAGAGACCATG	
Cdc25b	CTTGTGCTTGTGATTCTGTCCTTACTGC		CCTTACCTGTTCCTCTTCCTTATCCAGC	
Top1mt	CGAGAAGTCGATGCAGACACTTCAA		ATACCCAGTCCACATCCCTGCC	
Sesn2	GCTGAAGACTGGCGAGCACAGCT		CCTCTGCATCTCCCTCAGGAAGTATT	
Ncan	GACCTGAATGTTGTGGCTGAGAGTCC		GCCTCCTGTCCCCAGGTCCC	
Nacad	CCCTCACGTTCCTGTCCAGCAA		CACTAGGCTTGGGCTGCCCTCT	

[1851] MiSeq 판독을 이전에 기재된 바와 같이 분석하였다. 간단히, 판독을 프레드 품질(Phred quality)(Q 스코어)에 의해 필터링시키고, 스미스-워터맨 알고리즘(Smith-Waterman algorithm)을 사용하여 표적 부위의 상류 및 하류의 50 개 뉴클레오티드의 게놈 영역에 정렬시켰다. 표적 부위의 상류 5 개의 뉴클레오티드로부터 하류 5 개의 뉴클레오티드의 정렬된 영역(전체 30 bp)에서 삽입결실을 평가하였다. 추정 절단 사건으로서 삽입결실의 포함 또는 배제를 평가하기 위해 각각의 샘플에 대해 음성 대조군을 사용하였다. 출원인은 음성 대조군 샘플의 데이터로부터 판독당 표적 영역당 오류 비율(per-target-region-per-read error rate)을 사용하여 실제-삽입결실을 갖는 표적-영역을 갖는 판독의 분획에 대한 최대 가능도 추정량(MLE)을 계산하였다. 각각의 표적에 대한 MLE 스코어 및 절단 비율이 표 2에 나열되어 있다.

[1852] 통계 분석: 모든 실험을 최소 2 개의 독립적 생물학적 복제물을 이용하여 수행하였다. 스튜던츠 양측 *t*-검정을 사용하여 Prism6(GraphPad)로 통계를 수행하였다.

[1853] 뇌에서 CRISPR의 추가 특성규명을 위하여, 출원인은 사멸 세포 및 표준 조직에 대한 TUNEL 염색(예를 들어, 헤마톡실린 염색)을 사용하여 독성 효과를 시험한다. 더 상세한 방법으로 다수 유전자좌의 변형의 효율을 시험하기 위해, 이후 출원인은 GFP-양성 단일 세포를 분류하고 다음-유전자 시퀀싱을 사용하여 세포 당 변형 효율을 분석한다.

[1854] 출원인은 생체내 2-광자 현미경 하에서 패치 클램프를 사용하여 피질 내 표적화된 세포의 생리학적 표현형으로 뇌에서의 Mecp2 녹다운의 영향을 특성규명하였다. 성체 뇌에서 Mecp2 녹다운의 영향을 이해하기 위해, 출원인은 생체내 표적화된 세포의 전사체를 분석한다. 흥미로운 후보 유전자가 추가 연구를 위해 선택될 것이다.

[1855] 시스템의 2 가지 구성성분(Cas9 및 sgRNA)으로 표적화된 세포를 효율적으로 정제/시각화하기 위해, 출원인은 GFP의 분할 형태와 결합된 Cas9의 분할 형태를 구성하며, 이는 2 가지 AAV가 동일 세포내에 있을 경우에만 재구성된다. 마지막으로 출원인은 본원에 기재된 바와 같이 실험을 수행하여 시험관내 및 생체내 유사분열 후 뉴런에서의 상동성 재조합을 달성한다.

[1856] *실시예 37: AAV 벡터 및 간-특이적 Cas9 프로모터를 이용하여 간으로 정맥내 전달된 가이드 및 SaCas9로 생체내 관찰된 ApoB 유전자형 및 표현형 변화*

[1857] 본 실시예에서, 특히,

[1858] · AAV2/8은 간-표적화 아데노바이러스 벡터이다;

[1859] · TBG는 간-특이적 프로모터이고, SaCas9의 발현을 유도하기 위해 본 실시예에서 사용된다;

[1860] · U6는 sgRNA(가이드)의 발현을 유도하기 위해 본 실시예에서 사용된다;

[1861] · ApoB는 지질 대사 유전자이다. 이는 간 전달에서 "최적-표준"인 것으로 언급될 수 있으며, 비만의 마우스 모델에서 널리 사용된다;

[1862] · "표적1 내지 표적4"는 ApoB 내의 4개의 표적이 선택된 것을 의미하며, 이 중 표적 1 및 3(T1 및 T3)이 가장 유용하였다;

[1863] · 본 실시예에서 관찰되는 바와 같은 바이러스 벡터로부터의 발현을 통한 전달은 전달 방법으로서 앤더슨/인(Anderson/Yin)(NBT 2884)의 유체역학적 전달의 이용에 비해 개선인데, 이는 유체역학적 전달이 뮤린 신체에 스트레스를 줄 수 있고, 치명적일 수 있는 수 ml의 유체가 주입되는 것을 필요로 하기 때문이다. 유체역학적 전달은 플라스미드(네이키드) DNA의 전달에 충분히 적합한 반면, 본 출원인은 바이러스 전달 벡터 내에 가이드 및 Cas9 서열을 패키징시키는 것이 크게 증가된 효율과 관련하여 바람직한 것을 제시하였다. 또한, 단지 비교적 적은 부피가 도입되는 것을 필요로 하며, 이는 치료적으로 훨씬 더 용인될 수 있는 정맥내(i.v.)로 수행될 수 있다.

[1864] · 특히 고무적인 것은 간에 대한 "최적 표준" 유전자, 예를 들어, ApoB에서 유전자형 변화가 관찰될 뿐만 아니라, **표현형** 변화가 또한 기록되었다는 점이다. PCSK9를 이용한 이전의 작업은 유전자형 변화를 나타내었으나, 표현형 변화를 나타내지 않았고, 따라서 ApoB로 관찰된 표현형 변화는 간으로의 CRISPR 전달의 개연성 및 간에서 표현형 변화를 초래하는 이의 능력을 확인한다. 이는 더욱 치료적으로 허용되는 전달 수단(유체역학적 전달에 비해 i.v.)과 조합된다. 이와 같이, CRISPR(가이드 및 Cas9)의 바이러스 전달, 특히 정맥내 전달이 바람직하다.

[1865] · 표적은 *PCSK9, HMGCR, APOB, LDLR, ANGPTL3, F8, F9/FIX, AAT, FAH, HPD, TAT, ATP7B, UGT1A1, OTC, ARH*를 포함한다.

[1866] **재료 및 방법**

[1867] ***바이러스 및 주사 파라미터***

[1868] 사용된 작제물:

[1869] -AAV2/8 - TBG-SaCas9-U6-sgRNA (Apob-표적1 내지 표적4).

[1870] 시험관내 시험: 모두 Hepa 세포에서 10% 내지 15% 효율로 Apob 유전자좌의 절단을 유도하였다.

[1871] 생체내 결과: 마우스 - 8주, C57BL/6 (각각의 시점에 2마리의 동물 및 염수-주사된 야생형 대조군으로서 1마리의 동물)

[1872] ***꼬리 정맥 주사***:

[1873] *주사 부피*: 100㎕의 0.8E12 vp/㎖ (vp = 바이러스 입자)

[1874] *전달된 바이러스 입자*: 0.8E11 전체 vp/동물

[1875] **조직 처리 및 데이터 수집**

[1876] *조직 처리 및 데이터 수집을 하기와 같이 수행하였다*:

[1877] 첫 번째 시점 약 1주(8일). 두 번째 시점 ~ 4주.

[1878] 염수 관류 후 간 조직의 급성 해부.

[1879] (A) 절반의 간을 서베이어 & qPCR & 웨스턴 블롯 단백질 분석(X12 튜브/동물)을 위해 -80C 저장하였다.

[1880] (B) 절반의 간을 동해방지제에 두고, 냉동미세절단 과정을 위해 순간-동결시켰다. 동결절편을 H&E 및 오일 레드 (Oil Red) 염색시켰다.

[1881] 동물 당 2개 단편의 간으로부터 삽입-결실을 검출하고 정량하기 위해 QuickExtract 및 서베이어 검정을 이용하였다.

[1882] **결과**

[1883] *생체내 삽입-결실 평가*

[1884] 도면은 시간 경과에 걸친(주사 후 4주까지) ApoB 가이드(표적)에 대한 생체내 삽입-결실 평가를 나타낸다. 도 56a는 가이드(표적) 1이 ApoB에서 삽입-결실의 가장 높은 백분율을 유도하였다. 표적 2 및 4는 이들이 단지 삽입-결실 형성되지 않았거나, 매우 불충분한 삽입-결실을 초래한 한편, 표적 3은 일부 활성을 나타낸 의미에서 효과가 거의 없거나 효과가 없음을 나타내었다. 도 56b는 주사 4주 후의 삽입-결실 형성 효율에 대한 서베이어 뉴클레아제 겔 검정의 결과를 제시한다.

[1885] 표적 1은 거의 9%의 삽입-결실 형성을 갖는 것으로 관찰될 수 있으며, 이는 표적 유전자좌의 허용가능한 수준이다.

[1886] *표적에 대해 설계된 4개의 가이드 중 2개로 제시된 표현형 변화*

[1887] AAV-Cas9-sgRNA 전달 후 생체내에서 간 지질 축적 표현형을 검출하기 위한 오일 레드 염색을 제시하는 도 66에서 관찰되는 바와 같이, 사용된 3개의 가이드 중 2개(표적 1 및 3)로 표현형 변화가 관찰되었다. 좌측의 표적 1 및 3에서의 표적 66에서 축적되는 것으로 제시된 오일의 레드 패치는 ApoB가 하부 우측의 대조군에 비해 붕괴된 것을 나타낸다. Apob 유전자는 Cas9-유도된 표적화된 게놈 절단의 결과로서 붕괴되었고, 이는 상기 생리학적/표

현형 변화를 발생시켰다. 표적 2는 대조군에 비해 현저한 차이를 나타내지 않았고, 표적 4는 제시되지 않았다. 이러한 오일 레드 O 염색은 간 내의 지방이 조직학적 염색을 통해 시각화되는 검정이다. 이러한 염색은 간 내의 지방의 양을 평가하기 위해 연구에서 빈번하게 이용된다. 임상 실시에서, 간 이식 및 다른 절차 동안 간 내의 지방의 양을 평가하기 위해 오일 레드 O 염색은 주로 간 생검 표본의 동결 절편으로 지시된다. 본 실시예의 이러한 양태에 대한 프로토콜 및 정보를 위해, 문헌[Mehlem et al, "Imaging of neutral lipids by oil red O for analyzing the metabolic status in health and disease," Nature Protocols 8, 1149-1154(2013); Maczuga et al., "Therapeutic expression of hairpins targeting apolipoprotein B100 induces phenotypic and transcriptome changes in murine liver," *Gene Therapy* (2014) 21, 60-70; Koornneef et al, "Apolipoprotein B Knockdown by AAV-delivered shRNA Lowers Plasma Cholesterol in Mice," *Molecular Therapy* (2011) 19 4, 731-740; Tadin-Strapps et al., "siRNA-induced liver ApoB knockdown lowers serum LDL-cholesterol in a mouse model with human-like serum lipids," Journal of Lipid Research Volume 52, 1084-1097 (2011)]가 언급된다. 상기 도 66에서의 축척 막대는 20 마이크론이다.

[1888] *실시예 38: SaCas9 최적화 실험*

[1889] 하기를 연구하였다:

[1890] 가이드 길이 최적화

[1891] 인트론 시험

[1892] H1 프로모터

[1893] D10A 이중-닉카아제 시험

[1894] 추가 길이/DN 시험.

[1895] ***SaCas9 가이드 길이 시험:*** sgRNA 가이드 길이를 결정하기 위해, 20 대 21 대 22 bp뿐만 아니라 가이드의 시작(5' 말단)시에 'G'의 효과를 시험하였다. 본 실험에서,

[1896] *표적 부위*는 하기와 같다:

[1897] A1: AAVS1

[1898] E1: EMX1

[1899] T1, T2,...: 표적 부위의 넘버링

[1900] TGC, GTC,...: PAM의 5'-말단으로부터의 위치 23, 22, 21 nt에서의 염기 조성

[1901] 본 실시예의 실험은, 1. 2개의 관심 유전자 AAVS1 및 EMX1 내에서 PAM으로서 NNGRR을 이용한 표적 선택, 2. 표적에 상응하나, 20개 내지 21개, 또는 20개 내지 22개까지 sgRNA 내의 가이드 서열 부분의 길이가 다양한 올리고 합성, 3. sgRNA 발현 카세트를 생성시키고, SaCas9 단백질을 발현하는 플라스미드와 함께 HEK 293FT 세포주로 공동-트랜스펙션시키기 위해 올리고 이용, 4. 트랜스펙션 72시간 후, 세포의 수거 및 이후 삽입-결실을 검출하기 위한 서베이어 검정에 의한 분석, 5. 이후, Cas9에 의해 유도된 삽입-결실 형성 빈도의 계산 및 이와 함께 도면에의 요약에 의해 수행된다.

[1902] 도 67은 일정 범위의 표적에 걸쳐 2개의 상이한 유전자(AAVS1 및 EMX1) 내에서 회색 막대로 표시된 21 nts/염기쌍(bp)이 적어도 20 또는 22개 염기쌍(각각 흑색 및 백색 막대로 표시됨)에 비해 최적의 스페이서 길이임을 제시한다. 표적 및 유전자는 중요한 것으로 생각되지 않으며, 단지 대표이다. 이와 같이, SaCas9에서 또는 SaCas9에 관하여 21nt 또는 21개 염기쌍이 우수한 길이로서 최적임을 나타낸다. 도 67은 또한 가이드/표적 서열의 5' 말단에서의 G nt가, 예를 들어, U6 프로모터에 대해 유리할 수 있음을 제시한다. 최적 가이드 길이는 각각의 Cas9 단백질에 특이적일 수 있다. 예를 들어, SpCas9에 대해, 최적 가이드 길이는 19-20nt일 수 있다. 그러므로, 본 실시예는 최적 가이드 길이를 결정하고 이용할 수 있는 방법을 입증한다.

[1903] ***인트론 시험***

[1904] 본 실험은 가이드 서열이 Cas9 인트론 서열로 삽입될 수 있는지의 여부를 시험하기 위해 수행된다.

[1905] 하기 작제물을 이용하였다. 인트론 내부(Cas9 N'과 C' 말단 엑손 사이)의 가이드 RNA(sgRNA)의 존재를 인지하라.

[1906] CMV-SaCas9(N-term)-인트론(sgRNA)-SaCas9(C-term)

[1907] 작제물을 Hepa 세포에서 발현시켰다.

[1908] 각각의 인트론을 2개의 상이한 가이드 *Pcsk9* 및 *Hmgcr* sgRNA로 시험하였다.

[1909] 전체 9개의 작제물이 제시되었다: 3개의 EBV1, 3개의 EBV2 및 3개의 ADV:

[1910] - 레인 1-3은 EBV1-152(EBV 기반, EBV 게놈으로부터의 152bp 인트론 1)를 제시한다.

[1911] - 레인 4-6은 EBV2(EBV 기반, EBV 게놈의 W 반복부로부터의 인트론)를 제시한다.

[1912] - 레인 7-9는 ADV(아데노바이러스 기반 인트론, 문헌[Kiani et al., "CRISPR transcriptional repression devices and layered circuits in mammalian cells," Nature Methods doi:10.1038/nmeth.2969 Published online 5 May 2014 및 Nissim et al, "Multiplexed and Programmable Regulation of Gene Networks with an Integrated RNA and CRISPR/Cas Toolkit in Human Cells," Volume 54, Issue 4, p698-710, 22 May 2014; DOI: http://dx.doi.org/10.1016/j.molcel.2014.04.022]과 유사한 기원)를 제시한다.

[1913] 각각의 설계 그룹 내에서, 3개의 작제물은 인트론 내부의 sgRNA의 3개의 상이한 삽입 부위에 상응한다.

[1914] ADV-설계 3

[1915] 결과는 도 68에 제시된다. 이들 결과는 SaCas9 인트론으로의 가이드 서열의 성공적인 패키징의 원리가 확실히 가능하다는 증거를 제공한다. 가이드 서열을 가진 sgRNA는 아데노바이러스로부터 유래된 합성 인트론 내부에 삽입된 후, 이러한 전체 인트론-sgRNA 카세트가 SaCas9 유전자로 삽입된다. 인트론은 SaCas9 단백질의 정상 발현을 유의하게 붕괴시키지 않고 SaCas9 유전자 내부 어디에라도 삽입될 수 있다. sgRNA를 갖는 다수의 인트론은 SaCas9 유전자 내부의 다양한 위치에 삽입될 수 있다. 위치결정은 유연적이며, 이러한 폭넓은 접근법은 하기 두 방식을 포함하여 유리하다:

[1916] 크기 최소화는 작제물 내의 bp 또는 nts의 전체수가 감소되도록 한다.

[1917] 다중복합체화는 '공간'이 여기서도 또한 항상 문제임에 따라 더 큰 정도의 다중복합체화(다수의 가이드의 공동-전달)를 가능케 한다. 가이드는 특정 프로모터를 반드시 필요로 하지 않음에 따라, 하나 이상의 가이드는 Cas9 인트론으로 유사하게 패키징될 수 있다.

[1918] 상기 본문에서는 관사를 이용하는데, 이는 상기 논의된 바와 같이 다수의 합성 인트론이 Cas9에 도입될 수 있기 때문이다. 가깝지만 인트론의 5' 말단에 대해 적어도 5-15bp이고, 또한 인트론의 분지점 앞의 위치에 sgRNA를 삽입하는 것이 유리할 수 있다. 인트론의 중간의 분지점과 5' 스플라이스 공여 부위 사이의 인트론 스페이서 서열 일부는 당업자가 그렇게 하길 원하는 경우에 결실될 수 있다. 상기는 Cas9에서 달성되었으며, 특히 SaCas9가 놀라울 수 있는데, 이는 sgRNA 구조가 Sa와 Sp 사이에 상이하기 때문이다.

[1919] 지금으로서는, ADV가 바람직하나, 이러한 접근법은 일정 범위의 바이러스 및 Cas9s(Sa, Sp 등)에 걸쳐 폭넓은 적용성을 갖는다.

[1920] ***H1 프로모터 시험***

[1921] U6 프로모터에 대한 대안 프로모터를 연구하기 위해 본 실험을 설계하였다.

[1922] *A) 전장 H1*

[1923] 하기 작제물을 제조하였다:

[1924] 하나의 sgRNA(Pcsk9-표적201 또는 Hmgcr-신규표적5)를 유도하는 본래의 H1 프로모터를 갖는 CMV-SaCas9

[1925] 도 69에서 관찰될 수 있는 바와 같이, 전장 H1 프로모터(회색 막대)는 여전히 U6 프로모터(흑색 막대)보다 약한데, 이는 U6가 시험된 각각의 표적에 대해 증가된 삽입-결실 백분율 형성을 나타내기 때문이다.

[1926] *B) 이중 H1 프로모터 세트 (짧은 H1)*

[1927] 하기 작제물을 제조하였다:

[1928] 동일 방향 및 반대 방향으로 사용되는 이중의 짧은 H1 프로모터와 함께 동시에 2개의 sgRNA(Pcsk9-표적201 및 Hmgcr-신규표적5)를 유도하는 2개의 짧은 H1 프로모터를 갖는 TBG-SaCas9.

[1929] 도 70에서 관찰될 수 있는 바와 같이, 짧은 H1 프로모터는 전장 H1보다 약하다.

[1930] *SaCas9 닉카아제 시험(D10A 돌연변이를 이용함)*

[1931] 본 실험에서 작제물 내의 2개의 가이드 서열의 5' 말단 사이의 거리를 관찰하였고, 이후 D10A SaCAs9 이중 닉카아제의 절단 효율과 관련하여 이를 측정하였다. 표적은 인간 AAT1 유전자에 대한 것이었다. 이들 시험을 플라스미드로 클로닝된 20bp+G 가이드로 수행하였다.

[1932] 최적 결과가 -5 내지 +1 bp 사이(5'에서 5')에서 제시되었다. 도 71 참조.

[1933] *실시예 39: CRISPR-Cas9를 이용한 포유류 뇌에서의 유전자 기능의 생체내 질의*

[1934] 본 작업은 하기 주요 포인트를 제공한다:

[1935] · 생체내에서 성공적인 AAV-매개 Cas9 전달뿐만 아니라 유사분열 후 뉴런에서 효과적인 게놈 변형의 최초 입증

[1936] · Cas9 및 sgRNA-발현 세포로부터의 뉴런 핵의 용이한 분리를 가능케 하는 핵 태깅 기술의 개발

[1937] · 뉴런 전사체의 RNAseq 분석을 위한 적용의 입증

[1938] · 전기생리학적 연구와 Cas9-매개 게놈 교란의 통합, 및

[1939] · 다중복합체 표적화 및 Cas9-매개 게놈 편집을 이용한 설치류 행동에 대한 유전자 기능을 연구하는 능력의 입증.

[1940] 질병-관련 돌연변이를 갖는 트랜스제닉 동물 모델은 신경계 장애의 연구를 위해 매우 유용하며, 질병의 유전적 및 병태생리학적 메커니즘을 설명하는데 도움이 된다[1]. 그러나, 단일 또는 다수의 유전적 변형을 갖는 동물 모델의 발생은 특히 노동 집약적이며, 많은 세대에 걸친 시간-소모적 번식을 필요로 한다. 따라서, 정상 및 질병-관련 뇌 과정에서의 유전자 기능의 신속한 해부를 돕기 위해, 출원인은 생체내에서 뉴런의 게놈을 정확하고 효과적으로 조작하는 능력을 필요로 한다. 스트렙토코커스 피오게네스로부터의 CRISPR-관련 엔도뉴클레아제 Cas9(SpCas9)는 복제하는 진핵 세포에서 단일 및 다수의 유전자의 정확하고 효과적인 게놈 절단을 매개하여 프레임 이동 삽입/결실(삽입-결실) 돌연변이를 발생시키는 것으로 제시되었다[2, 3]. 여기서, 출원인은 신경 과정에서의 개별적 유전자뿐만 아니라 유전자의 그룹의 기능 및 생체내에서 뇌 장애에서의 이들의 역할을 연구하기 위해 Cas9-매개 게놈 교란과 생화학적, 시퀀싱, 전기생리학적, 및 행동 판독을 통합시켰다.

[1941] ***논의***

[1942] 마우스 뇌로 재조합 유전자를 전달하기 위해 아데노-관련 바이러스(AAV) 벡터가 흔히 사용된다[4]. AAV 시스템의 주요 한계는 ITR 없이 약 4.5 kb로 캡핑되는 이의 작은 패키징 크기이며[5], 이는 단일 벡터로 패키징될 수 있는 유전 물질의 양을 제한한다. SpCas9의 크기[6]가 이미 4.2 kb이어서, 단일 AAV 벡터 내의 다른 유전 성분에 대해서는 0.3 kb 미만이 남음에 따라, 출원인은 2개의 별개의 바이러스 벡터 상에 SpCas9(AAV-SpCas9) 및 sgRNA 발현 카세트(AAV-SpGuide)를 패키징시키는 이중-벡터 시스템을 설계하였다(도 72a). AAV-SpCas9 벡터를 설계하면서, 출원인은 SpCas9 발현을 최적화시키기 위해 다양한 짧은 뉴런-특이적 프로모터 뿐만 아니라 폴리아데닐화 신호를 비교하였다. 출원인의 최종 설계를 위해, 출원인은 배양된 일차 마우스 피질 뉴런에서 SpCas9 발현의 충분한 수준을 달성하는 이들의 능력을 기초로 하여 마우스 *Mecp2* 프로모터(235 bp, pMecp2)[7] 및 최소 폴리아데닐화 신호(48 bp, spA)[8]를 선택하였다(도 76-c). SpCas9-발현 뉴런의 면역형광 확인을 돕기 위해, 출원인은 HA-에피토프 태그로 SpCas9을 태깅하였다. AAV-SpGuide 벡터에 대해, 출원인은 U6-sgRNA 발현 카세트뿐만 아니라 인간 시냅신 I 프로모터에 의해 유도된 KASH 핵 막횡단 도메인과 융합된 녹색 형광 단백질(GFP)[9]을 패킹시켰다(도 72a). GFP-KASH 융합 단백질은 핵 외막으로 GFP를 유도하고(도 72c,d), AAV-Sp가이드에 의해 형질도입된 무결함 뉴런 핵의 형광-기반 확인 및 정제를 가능케 한다.

[1943] 출원인의 이중-벡터 전달 시스템의 전달 효능을 시험하기 위해, 출원인은 먼저 시험관 내에서 배양된 일차 마우스 피질 뉴런을 형질도입시키고, AAV-SpCas9 및 AAV-Sp가이드에 의한 강한 발현을 관찰하였으며(도 76c), 공동-형질도입 효율은 80%를 초과하였다(도 76e). 중요하게는, 형질도입되지 않은 뉴런과 비교하여, SpCas9의 발현은 형질도입된 뉴런의 형태 및 생존률에 불리하게 영향을 미치지 않았다(도 76c,f).

[1944] 효과적인 전달 시스템을 확립한 경우, 출원인은 다음으로 마우스 일차 뉴런에서 SpCas9-매개 게놈 편집을 시험하고자 하였다. 다양한 분열하는 세포 유형에서 효과적인 게놈 변형을 달성하기 위해 SpCas9가 사용된 반면, 유사분열 후 뉴런에서 게놈 편집을 효과적으로 달성하기 위해 SpCas9가 이용될 수 있는지의 여부는 불명확하다. 출원인의 최초 시험에 대해, 출원인은 자폐 범주성 장애의 한 유형인 레트 증후군[10]에서 주요 역할을 하는 *Mecp2* 유전자를 표적화하였다. MeCP2 단백질은 뇌 전체에 걸쳐 뉴런에서 편재성으로 발현되나, 아교세포[11, 12]에서는 거의 부재하며, 이의 결핍은 뉴런에서 심각한 형태적 및 전기생리학적 표현형과 관련된 것으로 밝혀졌고, 둘 모두는 레트 증후군을 가진 환자에서 관찰된 신경학적 증상에 기여하는 것으로 생각된다[13-16]. *Mecp2*를 표적화하기 위해, 출원인은 마우스 *Mecp2* 유전자의 엑손 3를 표적화하는 여러 sgRNA를 먼저 설계하였고(도 77a), Neuro-2a 세포를 이용하여 이들의 효능을 평가하였다. 가장 효과적인 sgRNA를 서베이어 뉴클레아제 검정을 이용하여 확인하였다(도 77b). 출원인은 이후의 시험관내 및 생체내 *Mecp2* 표적화 실험을 위해 가장 효과적인 sgRNA(*Mecp2* 표적 5)를 선택하였다.

[1945] 뉴런에서 출원인의 이중-벡터 시스템의 편집 효율을 평가하기 위해, 출원인은 시험관 내에서 7일에서 일차 마우스 피질 뉴런을 형질도입시켰고(7 DIV, 도 78a), 형질도입 7일 후에 서베이어 뉴클레아제 검정을 이용하여 삽입-결실률을 측정하였다(도 78b). 물론, *Mecp2*를 표적화하는 AAV-SpCas9 및 AAV-Sp가이드로 공동-형질도입된 뉴런 배양물은 대조군 뉴런에 비해 MeCP2 단백질 수준에서 80% 이하의 감소를 나타내었다(도 78c,d). 비교적 낮은 삽입-결실 빈도(약 14%)와 강한 단백질 고갈(약 80%) 사이에 관찰된 불일치에 대한 하나의 가능한 설명은 단지 표적 부위에서의 SpCas9에 의한 결합이 이콜라이[17, 18]에서 제시되었던 바와 같이 전사를 간섭할 수 있다는 것일 수 있다. 출원인은 RuvC 및 HNH 촉매 도메인 둘 모두가 비활성화된 SpCas9의 돌연변이[19,20](D10A 및 H840A, dSpCas9)를 이용하여 상기 가능성을 연구하였다. dSpCas9 및 *Mecp2*-표적화 sgRNA의 공동-발현은 MeCP2 단백질 수준을 감소시키지 않았으며(도 78a,d), 이는 활성 SpCas9의 존재하에서 MeCP2 수준의 관찰된 감소가 *Mecp2* 유전자좌에서의 변형의 발생으로 인한 것임을 암시한다. 검출된 삽입결실의 낮은 수준과 단백질 고갈의 높은 수준 사이의 불일치에 대한 또 다른 가능한 설명은 서베이어 뉴클레아제 검정에 의한 실제 삽입-결실률의 과소평가로 인한 것일 수 있는데, 서베이어의 검출 정확도는 삽입-결실 서열 조성에 민감한 것으로 이전에 밝혀졌었다[21].

[1946] MeCP2 기능 상실이 뉴런에서의 가지돌기 나무 이상 및 척추 형태형성 결함과 관련된 것으로 이전에 밝혀졌었다[14, 16]. MeCP2 박탈의 이들 표현형은 또한 MeCP-KO iPS 세포로부터 분화된 뉴런에서 재현되었다[15]. 따라서, 출원인은 뉴런에서의 SpCas9-매개 MeCP2-고갈이 레트 증후군의 형태학적 표현형을 유사하게 반복 발생할 수 있는지의 여부를 연구하였다. 또한, SpCas9 및 *Mecp2*-표적화 sgRNA를 공동-발현하는 뉴런은 대조군 뉴런과 비교하는 경우 변경된 가지돌기 나무 형태 및 척추 밀도를 나타내었다(도 79). 이들 결과는 유사분열 후 뉴런에서 표적화된 녹아웃을 가능케 함으로써 세포 검정에서 유전자 기능의 연구를 돕기 위해 SpCas9가 이용될 수 있음을 입증한다.

[1947] 이종성 세포 유형의 복잡한 네트워크로 구성된 신경계의 복잡성을 고려하면, 생체내에서 뉴런의 게놈을 효과적으로 편집할 수 있음은 자연 상황에서 엠베딩된 관련 세포 유형에서의 유전자 기능의 직접적인 시험을 가능케 할 것이다. 결과로서, 출원인은 성체 마우스의 해마 치아이랑으로 높은 역가의 AAV-SpCas9 및 AAV-Sp가이드의 혼합물(1:1 비)을 정위 주사하였다. 출원인은 바이러스 주사 4주 후(도 72b,c)에 해마 과립 세포에서 둘 모두의 벡터의 높은 공동-형질도입 효율(80% 초과)을 관찰하였고, 이는 *Mecp2* 유전자좌의 게놈 변형을 발생시켰다(도 72d). 서베이어 뉴클레아제 검정을 이용하여, 출원인은 주사된 뇌 영역으로부터 수득된 뇌 천공에서 약 13%의 삽입-결실 빈도를 검출하였다(도 72e). 배양된 일차 뉴런에서의 출원인의 발견과 유사하게, *Mecp2* 유전자좌의 SpCas9-매개 절단은 MeCP2 단백질 수준을 60% 초과까지 효과적으로 감소시켰다(도 72f). 또한, AAV-SpCas9 단독에 비해 AAV-SpCas9 및 AAV-Sp가이드로 주사된 경우 치아이랑에서의 MeCP2-양성 핵의 수가 75% 초과까지 감소하였다(도 72g-h). 이들 결과는 SpCas9가 무결함 생물학적 상황 내에서 특정 유전자를 직접 교란시키기 위해 이용될 수 있음을 암시한다.

[1948] 표적화된 게놈 교란은 특정 게놈 구성요소의 생물학적 기능에 대한 통찰을 제공하기 위해 정량적 판독과 커플링될 수 있다. AAV-SpCas9 및 AAV-Sp가이드 형질도입 세포의 분석을 돕기 위해, 출원인은 형광 활성화 세포 정렬(FACS)을 이용하여 GFP-KASH 표지된 핵을 정제하는 방법을 개발하였다(도 73a). 정렬된 핵은 하류 생화학 또는 시퀀싱 분석을 위해 핵 DNA 및 RNA를 정제하기 위해 직접 이용될 수 있다. 생어 시퀀싱을 이용하여, 출원인은 14개의 단일 GFP-양성 핵 중 13개가 sgRNA 표적 부위에서 삽입-결실 돌연변이를 함유한 것을 발견하였다.

[1949] 게놈 DNA 시퀀싱에 더하여, MeCP2 고갈의 전사 결과를 연구하기 위한 RNAseq 분석을 위해 정제된 GFP-양성 핵이

또한 이용될 수 있다(도 73b 및 도 80). 치아이랑으로부터 뉴런의 전사에 대한 *Mecp2* 녹아웃의 효과를 시험하기 위해, 출원인은 AAV-SpCas9뿐만 아니라 박테리아 *lacZ* 유전자를 표적화하나, 마우스 게놈은 그렇지 않도록 설계된 대조군 sgRNA 또는 *Mecp2*-표적화 sgRNA를 투여받는 동물로부터의 FACS 정제된 GFP^+ 핵을 이용하여 RNAseq 라이브러리를 제조하였다. 모든 sgRNA는 이들의 표적외 스코어를 최소화시키기 위해 최적화시켰다(CRISPR Design Tool: http://tools.genome-engineering.org)[2]. 출원인은 대조군과 *Mecp2* sgRNA 발현 핵 사이에서 차별적으로 발현된 유전자(도 70b)를 찾을 수 있었다($p<0.01$). 출원인은 *Mecp2* sgRNA를 발현하는 핵에서 하향 조절된 유전자 중에서 여러 흥미로운 후보를 확인하였다: 학습 행동에 있어서 중요한 역할을 하는 것으로 이전에 보고된 *Hpca*, *Olfm1*, 및 *Ncdn*[22-24]; 및 시냅스 소포 방출과 관련되고 뉴런 발화율과 관련된 것으로 밝혀진 *Cplx2*[25, 26]. 이들 결과는 SpCas9-매개 게놈 교란 및 집단 수준 RNAseq 분석의 조합이 뉴런에서 전사 조절을 특성규명하기 위한 방식을 제공하고, 특정 뉴런 기능 또는 질병 과정에 중요할 수 있는 유전자를 암시하는 것을 입증한다.

[1950] 특정 세포 유형 또는 회로 구성요소에 대한 게놈 교란의 효과를 연구하기 위해 뇌에서의 SpCas9-매개 생체내 게놈 편집은 또한 전기생리학적 기록과 커플링될 수 있다. 뉴런 생리학에 대한 MeCP2 고갈의 기능적 효과를 연구하기 위해, 출원인은 수컷 마우스의 일차 시각 피질(V1)의 표면층으로 *Mecp2*를 표적화하는 AAV-SpCas9 및 AAV-Sp가이드를 정위 공동-전달하였다. V1을 선택하였는데, 이는 표면층 피질 흥분 뉴런이 2-광자 영상화 및 2-광자 가이드된 표적화 기록에 더욱 접근가능하기 때문이다. SpCas9 전달 2주 후, 마우스 V1의 층 2/3에서 KASH-GFP^+ 뉴런 및 GFP^- 이웃 뉴런의 전기생리학적 반응(도 72a-c)을 비교하기 위해 마우스를 2-광자 가이드된 세포곁 기록에 적용시켰다(도 74). 출원인은 20-도 증분으로 18 드리프팅 격자(drifting grating)에 대한 뉴런 반응을 측정하였고, 벡터에 의한 세포의 발생된 발화율(FR) 및 방향 선택성 지수(OSI)를 계산하여, 반응에 대한 평균을 내었다. FR 및 OSI 둘 모두는 이웃 GFP^- 흥분 뉴런에 비해 흥분 GFP^+, MeCP2 녹아웃 뉴런에 대해 유의하게 감소되었다(도 73d-e). 비교시, SpCas9와 함께 대조군 sgRNA 발현은 이웃의 감염되지 않은 뉴런과 비교하는 경우 FR 및 OSI에 대해 어떠한 영향도 미치지 않았다(도 73d-e). 이들 결과는 성체 V1 피질 뉴런에서의 MeCP2의 SpCas9 매개 고갈이 2주 이내에 생체내에서 흥분 뉴런의 시각 반응 특성을 변경시키는 것을 제시하며, 유전자 기능 및 신경 회로의 해부를 연구하기 위해 생체내에서 포유류 뇌에서 표적화된 유전자 녹아웃을 돕는데 있어서의 SpCas9의 다능성을 추가로 입증한다.

[1951] SpCas9 시스템의 한 주요 장점은 다중복합체 게놈 편집을 돕는 이의 능력이다[2]. 살아 있는 동물의 뇌에 다수의 유전자의 안정적인 녹아웃을 도입시키는 것은 생리학적 및 신경병리학적 조건에서 다인자 메커니즘의 인과적 질의와 같은 잠재적으로 광범위한 적용성을 가질 것이다. 뇌에서의 다중복합체 게놈 편집의 가능성을 시험하기 위해, 출원인은 핵 표지화를 위한 GFP-KASH와 함께 연쇄적인 3개의 sgRNA로 구성된 다중복합체 sgRNA 발현 벡터를 설계하였다(도 74a). 출원인은 *Dnmt1*, *Dnmt3a* 및 *Dnmt3b*로 구성되는 DNA 메틸트랜스페라제 유전자 패밀리(DNMT)를 표적으로 하는 sgRNA를 선택하였다. *Dnmt1* 및 *3a*는 성체 뇌에서 고도로 발현되며, DNMT 활성이 DNA 메틸화를 변경시키고, Dnmt3a 및 Dnmt1 둘 모두가 시냅스 가소성 및 학습 및 기억 형성에 필요한 것으로 이전에 밝혀졌다[27]. 출원인은 높은 변형 효율을 갖는 *Dnmt3a* 및 *Dnmt1*에 대한 개별적 sgRNA를 설계하였다. *Dnmt3b*에 의한 임의의 잠재적 보상 효과를 피하기 위해, 출원인은 상기 유전자가 신경발달 동안 주로 발현되긴 하지만 상기 유전자를 추가로 표적화하는 것을 또한 결정하였다[27]. 출원인은 최종적으로 3개 모두의 표적화된 유전자에 대한 고도의 동시 DNA 절단을 위해 개별적 sgRNA를 최종적으로 선택하였다(도 75b 및 도 81).

[1952] 생체내에서 다중복합체 게놈 편집의 효능을 시험하기 위해, 출원인은 수컷 성체 마우스의 등쪽 및 배쪽 치아이랑으로 고역가 AAV-SpCas9 및 AAV-Sp가이드의 혼합물을 정위 전달하였다. 4주 후, 해마를 해부하고, 표적화된 세포 핵을 FACS에 의해 분류하였다. 출원인은 서베이어 핵 검정(도 75c) 및 시퀀싱(도 82)을 이용하여 분류된 핵 집단에서 약 19%(*Dnmt3a*), 18%(*Dnmt1*) 및 4%(*Dnmt3b*)의 삽입-결실 빈도를 검출하였다. 다수의 유전자좌를 표적화하는 것은 개별적 세포에서 다수의 녹아웃의 효과적인 비율에 관한 의문을 불러일으킨다. 표적화된 시퀀싱과 조합된 단일 핵 분류를 이용함으로써, 출원인은 개별적 뉴런 핵에서 다수의 DNMT 유전자좌의 동시 표적화를 정량하였다(도 75d). 적어도 하나의 *Dnmt* 유전자좌에서 변형을 갖는 뉴런 핵에서, 핵 중 70% 초과가 *Dnmt3a* 및 *Dnmt1* 둘 모두에서 삽입-결실을 함유하였다(약 40%는 3개 모두의 유전자좌에서 삽입-결실을 함유하였고, 약 30%는 *Dnmt3a* 및 *Dnmt1* 유전자좌 둘 모두에서 삽입-결실을 함유하였다). 이들 결과는 치아이랑에서의 Dnmt3a 및 Dnmt1 단백질 고갈 수준과 일치한다(도 75e). 성체 뇌에서의 *Dnmt3b*의 낮은 발현으로 인해, 출원인은 Dnmt3b 단백질을 검출할 수 없었다.

[1953] SpCas9를 이용한 최근의 연구는 20-nt sgRNA 서열 내의 각각의 염기가 종합적인 특이성에 기여하긴 하나, sgRNA와 부분적으로 매치되는 게놈 유전자좌가 표적외 이중 가닥 브레이크 및 삽입-결실 형성을 초래할 수 있는 것을 제시하였다[28, 29]. 표적외 변형의 비율을 평가하기 위해, 본 발명자는 매우 유사한 게놈 표적 부위의 목록을 전산적으로 확인하였고[2], 표적화된 심층 시퀀싱을 이용하여 변형의 비율을 정량하였다. 상기 예측된 표적외 유전자좌의 삽입-결실 분석은 삽입-결실 형성의 0 내지 1.6% 비율을 나타내었고, 이는 SpCas9 변형이 특이적인 것을 입증한다(부록 표 1). 생체내에서 SpCas9-매개 게놈 편집의 특이성을 증가시키기 위해, 추가 연구는 이중 닉킹[30, 31] 및 트렁케이션된 sgRNAs[28]와 같은 표적외 최소화 전략을 이용할 수 있다.

[1954] Dnmt3a 및 Dnmt1의 녹다운이 해마-의존성 기억 형성에 영향을 주는 것은 이전에 밝혀졌다[27]. 결과로서, 출원인은 기억 획득 및 경화(consolidation)에 대한 SpCas9-매개 삼중 녹다운(*Dnmt3a*, *Dnmt1* 및 *Dnmt3b*)의 효과를 연구하기 위한 전후관계 공포-조건화 행동 시험(contextual fear-conditioning behavior test)을 수행하였다. 출원인은 기억 획득 단계에서 대조군과 삼중 녹다운 마우스 사이에 어떠한 차이도 관찰하지 않았으나, 녹아웃 마우스는 훈련 상황 조건하에서 시험되는 경우 손상된 기억 경화를 나타내었다(도 75f). 이러한 효과는 마우스가 변경된 상황에서 시험된 경우에 폐지되었다. 출원인의 결과는 치아이랑 뉴런에서의 DNMT 패밀리 일원의 CRIPSR-Cas9-매개 녹아웃이 행동 임무에서 유전자의 기능을 탐지하기에 충분한 것을 입증한다.

[1955] 전체적으로, 출원인의 결과는 SpCas9 및 sgRNA의 AAV-매개 생체내 전달이 무결함 신경 회로 내에서 정확한 게놈 교란을 달성하기 위한 신속하고 강력한 기술을 제공하는 것을 입증한다. SpCas9가 분열하는 세포를 유전자 조작하기 위해 널리 사용된 반면, 출원인은 SpCas9가 또한 NHEJ-매개 삽입-결실 발생을 통해 높은 효율로 유사분열 후 뉴런의 게놈을 유전자 조작하기 위해 이용될 수 있음을 입증한다. SpCas9-매개 게놈 교란은 표적화된 게놈 구성요소의 기능을 연구하기 위해 생화학적, 시퀀싱, 전기생리학적, 및 행동 분석과 조합될 수 있다. 출원인은 단일 또는 다수 유전자의 SpCas9-매개 표적화가 고전적인 더욱 시간-소모적인 유전 마우스 모델을 이용하여 관찰된 형태적, 전기생리학적, 및 행동 표현형을 반복 발생시킬 수 있음을 입증한다. 2개의 AAV 벡터의 사용을 필요로 할뿐만 아니라 프로모터 구성요소의 크기를 제한하는 스트렙토코커스 피오게네스 Cas9를 이용한 현재의 연구가 세포 유형-특이적 표적화를 달성하기 위해 이용될 수 있다. 일부가 SpCas9보다 실질적으로 짧은 Cas9 오솔로그의 다양성이 주어지면[2, 32, 33], 본원에 기재된 바와 같이 Cas9 및 sgRNA 둘 모두를 발현하는 단일 AAV 벡터를 유전자 조작하는 것이 가능할 것이다.

[1956] 참고자료

1. Nestler, E.J. & Hyman, S.E. Animal models of neuropsychiatric disorders. Nat Neurosci 13, 1161-1169 (2010).

2. Cong, L. et al. Multiplex genome engineering using CRISPR/Cas systems. Science 339, 819-823 (2013).

[1957]

3. Mali, P. et al. RNA-guided human genome engineering via Cas9. Science 339, 823-826 (2013).

4. Burger, C., Nash, K. & Mandel, R.J. Recombinant adeno-associated viral vectors in the nervous system. Hum Gene Ther 16, 781-791 (2005).

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

6. Deltcheva, E. et al. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. Nature 471, 602-607 (2011).

7. Gray, S.J. et al. Optimizing promoters for recombinant adeno-associated virus-mediated gene expression in the peripheral and central nervous system using self-complementary vectors. Hum Gene Ther 22, 1143-1153 (2011).

8. Levitt, N., Briggs, D., Gil, A. & Proudfoot, N.J. Definition of an efficient synthetic poly(A) site. Genes Dev 3, 1019-1025 (1989).

9. Ostlund, C. et al. Dynamics and molecular interactions of linker of nucleoskeleton and cytoskeleton (LINC) complex proteins. J Cell Sci 122, 4099-4108 (2009).

10. Chahrour, M. & Zoghbi, H.Y. The story of Rett syndrome: from clinic to neurobiology. Neuron 56, 422-437 (2007).

11. Kishi, N. & Macklis, J.D. MECP2 is progressively expressed in post-migratory neurons and is involved in neuronal maturation rather than cell fate decisions. Molecular and cellular neurosciences 27, 306-321 (2004).

12. Skene, P.J. et al. Neuronal MeCP2 is expressed at near histone-octamer levels and globally alters the chromatin state. Molecular cell 37, 457-468 (2010).

13. Chen, R.Z., Akbarian, S., Tudor, M. & Jaenisch, R. Deficiency of methyl-CpG binding protein-2 in CNS neurons results in a Rett-like phenotype in mice. Nat Genet 27, 327-331 (2001).

14. Zhou, Z. et al. Brain-specific phosphorylation of MeCP2 regulates activity-dependent Bdnf transcription, dendritic growth, and spine maturation. Neuron 52, 255-269 (2006).

15. Li, Y. et al. Global transcriptional and translational repression in human-embryonic-stem-cell-derived Rett syndrome neurons. Cell Stem Cell 13, 446-458 (2013).

[1958]

16. Nguyen, M.V. et al. MeCP2 is critical for maintaining mature neuronal networks and global brain anatomy during late stages of postnatal brain development and in the mature adult brain. J Neurosci 32, 10021-10034 (2012).

17. Jiang, W., Bikard, D., Cox, D., Zhang, F. & Marraffini, L.A. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. Nature biotechnology 31, 233-239 (2013).

18. Qi, L.S. et al. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell 152, 1173-1183 (2013).

19. Sapranauskas, R. et al. The Streptococcus thermophilus CRISPR/Cas system provides immunity in Escherichia coli. Nucleic acids research 39, 9275-9282 (2011).

20. Jinek, M. et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science 337, 816-821 (2012).

21. Qiu, P. et al. Mutation detection using Surveyor nuclease. BioTechniques 36, 702-707 (2004).

22. Kobayashi, M. et al. Hippocalcin-deficient mice display a defect in cAMP response element-binding protein activation associated with impaired spatial and associative memory. Neuroscience 133, 471-484 (2005).

23. Dateki, M. et al. Neurochondrin negatively regulates CaMKII phosphorylation, and nervous system-specific gene disruption results in epileptic seizure. The Journal of biological chemistry 280, 20503-20508 (2005).

24. Nakaya, N. et al. Deletion in the N-terminal half of olfactomedin 1 modifies its interaction with synaptic proteins and causes brain dystrophy and abnormal behavior in mice. Experimental neurology 250, 205-218 (2013).

25. Reim, K. et al. Complexins regulate a late step in Ca2+-dependent neurotransmitter release. Cell 104, 71-81 (2001).

26. Edwardson, J.M. et al. Expression of mutant huntingtin blocks exocytosis in PC12 cells by depletion of complexin II. The Journal of biological chemistry 278, 30849-30853 (2003).

27. Feng, J. et al. Dnmt1 and Dnmt3a maintain DNA methylation and regulate synaptic function in adult forebrain neurons. Nat Neurosci 13, 423-430 (2010).

[1959]

28. Fu, Y. et al. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. Nature biotechnology 31, 822-826 (2013).

29. Hsu, P.D. et al. DNA targeting specificity of RNA-guided Cas9 nucleases. Nature biotechnology 31, 827-832 (2013).

30. Ran, F.A. et al. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. Cell 154, 1380-1389 (2013).

31. Mali, P. et al. CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. Nature biotechnology 31, 833-838 (2013).

32. Esvelt, K.M. & Wang, H.H. Genome-scale engineering for systems and synthetic biology. Molecular systems biology 9, 641 (2013).

33. Li, W., Teng, F., Li, T. & Zhou, Q. Simultaneous generation and germline transmission of multiple gene mutations in rat using CRISPR-Cas systems. Nat Biotechnol 31, 684-686 (2013).

[1960]

[1961] **방법**

[1962] **DNA 작제물**

[1963] SpCas9 표적 선택 및 단일 가이드 RNA(sgRNA)의 발생을 위해, 20-nt 표적 서열을 5'-NGG PAM 서열에 앞서도록 선택하였다. 표적외 효과를 최소화시키기 위해, CRIPSR 설계 도구를 이용하였다(http://tools.genome-engineering.org). sgRNA를 정방향 프라이머: 5'-cgcacgcgtaattcgaacgctgacgtcatc-3' 및 20-nt DNA 표적 부위(***굵은 글씨***)와 함께 sgRNA를 함유하는 역방향 프라이머와 함께 주형으로서 U6 프로모터를 이용하여 PCR 증폭시켰다:

[1964] 5'-CACACGCGTAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACG GACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAAC*NNNNNNNNNNNNNNNN NNNNC*GGTGTTTCGTCCTTTCCAC-3'. (SEQ ID NO:)

[1965] 대조군 sgRNA 서열을 이콜라이로부터 *lacZ* 유전자를 표적화하기 위해 설계하였다: 표적 서열: TGCGAATACGCCCACGCGATGGG (서열 번호). EGFP-KASH[1] 작제물은 워먼 교수(Prof. Worman)(Columbia University, NYC)에게서 제공받았고, 인간 *시냅신* 프로모터(*hSyn*) 하에서 AAV 백본으로 암호화 카세트를 클로닝하기 위한 PCR 주형으로 사용하였다. 다음으로, U6-*Mecp2*sgRNA 암호화 서열을 *Mlu*I 부위를 이용하여 도입시켰다. 다중복합체 유전자 표적화 전략을 위해, 개별적 sgRNA를 상기 기재된 바와 같이 PCR 증폭시켰다. 3개 모두의 sgRNA를 골든 게이트(Golden Gate) 클로닝 전략을 이용함으로써 PCR 증폭된 *hSyn*-GFP-KASH-bGHpA 카세트(도 4A 참조)와 결찰시켰다. PCR 증폭 후, 3 sgRNA 및 *hSyn*-GFP-KASH-bGH pA를 함유하는 골든 게이트 결찰 생성물을 AAV 백본으로 클로닝하였다. 모든 수득된 작제물을 시퀀싱 확인하였다. 뉴런에서 SpCas9 발현을 유도하는 최적 프로모터 서열을 찾기 위해, 출원인은 h*Syn1*, 마우스 트렁케이션된 *Mecp2*(pMecp2), 및 트렁케이션된 래트 *Map1b*(pMap1b) 프로모터 서열을 시험하였다[2](부록 도 1A를 참조). 하기 프라이머를 프로모터 영역을 증폭시키기 위해 사용하였다:

[1966] hSyn_F: 5'-GTGTCTAGACTGCAGAGGGCCCTG-3'; (SEQ ID NO:)
hSyn_R: 5'-GTGTCGTGCCTGAGAGCGCAGTCGAGAA-3'; (SEQ ID NO:)
Mecp2_F 5'-GAGAAGCTTAGCTGAATGGGGTCCGCCTC-3'; (SEQ ID NO:)
Mecp2_R 5'-CTCACCGGTGCGCGCAACCGATGCCGGGACC-3'; (SEQ ID NO:)
Map1b-283/-58_F 5'-GAGAAGCTTGGCGAAATGATTTGCTGCAGATG-3'; (SEQ ID NO:)
Map1b-283/-58_R 5'-CTCACCGGTGCGCGCGTCGCCTCCCCCTCCGC-3'. (SEQ ID NO:)

[1967] 래트 *map1b* 프로모터의 또 다른 트렁케이션을 하기 올리고와 함께 어셈블리시켰다:

5'-AGCTTCGCGCCGGGAGGAGGGGGGACGCAGTGGGCGGAGCGGAGACAGC
ACCTTCGGAGATAATCCTTTCTCCTGCCGCAGAGCAGAGGAGCGGCGGGAGAGG
AACACTTCTCCCAGGCTTTAGCAGAGCCGGA-3' 및

5'-CCGGTCCGGCTCTGCTAAAGCCTGGGAGAAGTGTTCCTCTCCCGCCGCTC
CTCTGCTCTGCGGCAGGAGAAAGGATTATCTCCGAAGGTGCTGTCTCCGCTCCGC
[1968] CCACTGCGTCCCCCCTCCTCCCGGCGCGA-3'. (SEQ ID NO:)

[1969] 짧은 합성 폴리아데닐화 신호(spA)[3]를 하기 올리고를 이용하여 어셈블리시켰다:

5'-AATTCAATAAAAGATCTTTATTTTCATTAGATCTGTGTGTTGGTTTTTTG
TGTGC-3' (SEQ ID NO:) 및
5'-GGCCGCACACAAAAAACCAACACACAGATCTAATGAAAATAAAGAT
[1970] CTTTTATTG-3'. (SEQ ID NO:)

[1971] SpCas9 및 이의 D10A 돌연변이 형태(dSpCas9)가 이전에 기재되었다[4, 5]. EF1α 프로모터의 조절하의 플라스미드 암호화 적색 형광 단백질(mCherry)을 Lipofectamine®2000(Life Technologies)을 이용한 뉴런 트랜스펙션을 위해 이용하였다.

[1972] **세포주 배양 및 트랜스펙션**

[1973] Neuro-2a(N2a) 세포를 5% 소 태아 혈청(BSA)을 함유하는 DMEM에서 성장시켰다. HEK293FT 세포에 대해, 10% 소 태아 혈청(FBS)을 함유하는 DMEM을 이용하였다. 세포를 5% CO_2 대기에서 37℃에서 유지시켰다. 세포를 제조업체의 프로토콜에 따라 Lipofectamine®2000 또는 폴리에틸렌이민(PEI) "MAX" 시약(Polysciences)을 이용하여 트랜스펙션시켰다.

[1974] **농축된 AAV 벡터의 생성**

[1975] 고역가 AAV1/2 입자를 동등한 비율의 AAV1 및 AAV2 혈청형 플라스미드 및 pDF6 헬퍼 플라스미드를 이용하여 생성시키고, 헤파린 친화성 컬럼에서 정제하였다[6]. 바이러스 입자의 역가측정을 qPCR에 의해 수행하였다. 고역가 AAV1 입자를 UNC Vector Core Services(University of North Carolina at Chapel Hill)에 의해 생성시켰다. DMEM 내 저역가 AAV1 입자를 이전에 기재된 바와 같이 생성시켰다[7]. 간단히, HEK293FT 세포를 PEI "MAX"를 이용하여 트랜스젠 플라스미드, pAAV1 혈청형 플라스미드 및 pDF6 헬퍼 플라스미드로 트랜스펙션시켰다. 배양 배지를 48시간 후에 수거하고, 0.45 ㎛ PVDF 필터(Millipore)를 통해 여과시켰다.

[1976] **일차 피질 뉴런 배양**

[1977] 조직 배양을 위한 뉴런을 수득하기 위해 사용된 동물을 MIT 동물관리위원회(MIT CAC)에 의해 승인된 프로토콜에 따라 희생시켰다. 16일의 배아 마우스 뇌로부터 일차 배양물을 제조하였다[8]. 어느 한 성별의 배아를 이용하였다. 세포를 폴리-D-리신(PDL) 코팅된 24-웰 플레이트(BD Biosciences) 또는 라미닌/PDL 코팅된 커버슬립(VWR) 상에 플레이팅하였다. 배양물을 B27, 글루타맥스(Life Technologies) 및 페니실린/스트렙토마이신 혼합물이 보충된 Neurobasal 배지에서 37℃ 및 5% CO_2에서 성장시켰다.

[1978] AAV 형질도입을 위해, 500 ㎕ Neurobasal 배양 배지 중 피질 뉴런을 HEK293FT 세포로부터의 300 ㎕ (1:1 비율의 이중 감염) AAV1-함유 적응용 배지를 이용하여 7 DIV로 인큐베이션하였다[7]. 형질도입 뉴런 1주 후에 하류 공정을 위해 수거하거나, 면역형광 염색 또는 형태학적 분석을 위해 4% 파라포름알데하이드에 고정시켰다.

[1979] 뉴런 형태의 시각화를 위해, DIV7의 세포를 이전에 기재된 바와 같이 1주 동안 Lipofectamine®2000(Life Technologies)을 이용하여 EF1α-mCherry 발현 벡터로 트랜스펙션시켰다[9]. 전체 가지돌기 길이의 측정을 위해, 개별 뉴런의 모든 가지돌기를 ImageJ 소프트웨어를 이용하여 추적하였다. 일차 가지돌기, 가지돌기 첨단 및 숄

(Sholl) 분석[10]의 정량화를 40x 대물렌즈에서 형광 현미경(Zeiss AxioCam Ax10 현미경, Axiocam MRm 카메라)으로 획득된 이미지에 대해 수행하였다. 가지돌기 수에 대해, 10 ㎛보다 긴 모든 비-축삭 돌출의 말단을 계수하였다. 숄 분석을 위해, 직경에 있어서 5 ㎛ 단계를 갖는 동심원이 세포체 주위에 자동적으로 그려졌고, 각각의 원을 가로지르는 가지돌기의 수를 Sholl 플러그-인을 갖는 ImageJ 소프트웨어를 이용하여 계수하였다.

[1980] **마우스 뇌로의 AAV1/2의 정위 주사**

[1981] MIT CAC는 본원에 기재된 모든 동물 절차를 승인하였다. 성체(12-16주령) 수컷 C57BL/6N 마우스를 100 mg/kg 케타민 및 10 mg/kg 자일라진의 복막내(i.p.) 주사에 의해 마취시켰다. 선행 마취를 제공하였다(Buprenex, 1 mg/kg, i.p.). 개두술을 승인된 절차에 따라 수행하였고, 1 ㎕의 1:1 AAV 혼합물(1x1013 Vg/㎖의 sMecp2-SpCas9; 6x1012 Vg/㎖의 DNMT 3xsgRNA; 3-5x1012 Vg/㎖의 hSyn-GFP-KASH)을 등쪽 치아이랑(전방/후방: -1.7; 내외측: 0.6; 등쪽/배쪽: -2.15) 및/또는 배쪽 치아이랑(전방/후방: -3.52; 내외측: 2.65; 등쪽/배쪽: -3)에 주사하였다. 생체내 전기생리학적 기록 실험을 위해(도 74), 바이러스 주사 좌표는 후방 봉합으로부터 3 mm 측면(정수리점으로부터) 및 1 mm 전방이었다. 두개골을 염수로 가끔 냉각시키면서 드레멜 드릴(dremel drill)을 이용하여 약화시키고, 남아있는 경막을 무기질유에 현탁된 바이러스로 충전된 유리 마이크로피펫을 이용하여 천공시켰다. 여러 주사(3-4)를 200-250 ㎛의 깊이로 이웃하는 부위에서 수행하였다. 150-200 nl의 바이러스 혼합물의 부피를 각각의 부위에 75 nl/분의 속도로 주사하였다. 각각의 주사 후, 누출을 방지하기 위해 수축 전에 3 내지 5분 동안 적소에 고정시켰다. 절개를 봉합시키고, 적절한 수술후 진통제(멜록시캄, 1 내지 2 mg/kg)를 수술 후 3일 동안 투여하였다.

[1982] **생체내 2-광자 가이드된 표적화 루즈 패치(loose patch) 기록**

[1983] 바이러스 주사 2주 후, 전기생리학적 실험을 위해 마우스를 이용하였다. 마우스를 2% 이소플루란으로 마취시키고, 0.8% 이소플루란을 이용하여 유지시켰다. 피부를 해부하고, 수기(sugi)로 세척하고, 금속 헤드 플레이트를 아교 및 치아 아크릴릭을 이용하여 두개골에 부착시키고, 일차 시각 피질(V1) 상에서 2 mm x 2 mm 개두술을 수행하였다. 이후, 노출된 영역을 인공 뇌척수액(aCSF; 140 mM NaCl, 5 mM KCl, 2 mM CaCl2, 1 mM MgCl2, 0.01 mM EDTA, 10 mM HEPES, 10 mM 글루코스; pH 7.4) 중 1.5% 아가로스의 얇은 층으로 덮었다. 동물 체온을 실험 동안 가열 재킷으로 37.5℃로 유지시켰다.

[1984] 보로실리케이트 피펫(WPI)을 Sutter P-2000 레이저 풀러(Sutter Instruments)를 이용하여 끌어당겼다. 첨단 직경은 약 1 ㎛였고, 저항은 3 내지 5 MΩ였다. Matlab(MathWorks)에 기록되고, MultiClamp 700B 증폭기(Axon)를 제어하는 맞춤 소프트웨어(Network Prism, Sur lab)를 이용하여 기록하였다. 유리 피펫 전극을 20-35° 의 각도로 뇌에 삽입하고, Ag/AgCl 접지 전극 펠렛(Warner Instruments)을 뇌 및 대물렌즈와 동일 용액에 위치시켰다. 형광 시각화를 위해, 피펫을 Alexa Fluor 594(Molecular Probes)로 충전하였다. 피펫을 먼저 10x 렌즈를 이용하여 주사 부위에 대해 표적화시킨 후, 770 nm에서 동시 2-광자 영상화를 통해 25x 렌즈를 이용하여 개별적 GFP+ 세포에 대해 표적화시켰다. 신속하게 시간-변화 5 mV 명령 전압 펄스 동안 전압 고정에서 관찰된 저항에서의 편향을 통해 세포 근접을 검출하였다. 저항이 5-10 MΩ까지 증가된 후, 증폭기를 전류 고정으로 전환시키고, 4 KHz의 베셀 필터(Bessel filter) 및 300 Hz의 AC 필터 하에서 제로 주입 전류로 기록하였다. 바이러스 주사된 뇌를 사후(post hoc) 관류시키고, 면역조직화학을 수행하였다.

[1985] **시각 자극 및 생체내 2-광자 가이드된 표적화 루즈 패치 기록으로부터의 데이터 분석**

[1986] 게놈-편집된 뉴런의 방향 선택성 및 동조(tuning)를 평가하기 위해, 출원인은 Matlab PsychToolbox-3에 기록된 맞춤 소프트웨어를 이용하여 방향화된 격자를 제시하였다. 격자를 세포 반응성에 대해 최적화시키고, 20 도의 단계로 0으로부터 360 도로 방향을 단계화시킴으로써 제시하였으며, 각각의 격자 제시는 144초의 전체 제시 기간 동안 "오프(off)"에 대해 4초, 그 후 "온(on)"에 대해 4초 선행하였다. 데이터를 Matlab로 직접 획득하였고, .mat 파일로 저장하였다. 스파이크 검출을 수작업으로 규정된 역치 및 이후 추가 확인을 위한 스파이크 형태 주형 매칭을 이용한 검정 과정을 통해 수행하였다. 각각의 스파이크를 그래픽 사용자 인터페이스로 스크린 상에서 태깅시키고 제시하였으며, 그 결과 실험자에 의해 거짓 양성 및 음성에 대해 수작업으로 재검토하였다. 이후, 모든 자극에 반응한 스파이크 시간을 시각 자극에 대한 이들의 시간을 기초로 하여 "온" 또는 "오프" 기간으로 그룹화시켰고, 각각의 자극에 대한 "온" 스파이크는 동등한 기간 동안 관찰된 "오프" 스파이크의 수에 의해 감소되었다. 방향 실험에 대해, 자극 당 스파이크의 #는 (스파이크 "온"의 #) - (스파이크 "오프"의 #)인데, 이는 "온" 및 "오프" 기간이 동일 기간이었기 때문이다.

[1987] 관심대상의 모든 세포에 대해, 각각의 방향화된 자극(0으로부터 360 도, 20 도의 단계)에 대한 반응을 수거하기

위해 상기 방법을 이용하였다. 이후, 이들 반응을 각각의 시도에 대한 반응에 대한 방향의 "동조 곡선"으로 전환시켰다. 하기와 같은 식에 따라 바람직한 방향에 대한 벡터 평균을 취함으로써 방향 선택성 지수(OSI)를 계산하였다:

[1988] $$OSI = \frac{\sqrt{(\sum_i R(\theta_i)\sin(2\theta_i))^2 + (\sum_i R(\theta_i)\cos(2\theta_i))^2}}{\sum_i R(\theta_i)}$$

[1989] **조직 제조 및 세포 핵의 정제**

[1990] 전체 해마 또는 치아이랑을 얼음 저온 DPBS(Life Sciences)에서 신속하게 해부하고, 드라이아이스 상에서 쇼크 동결시켰다. 세포 핵 정제를 위해, 조직을 2 ㎖ Dounce 균질화기(Sigma)를 이용하여 2 ㎖ 얼음-저온 균질화 완충액(HB)(320 mM 수크로스, 5 mM CaCl, 3 mM $Mg(Ac)_2$, 10 mM 트리스 pH7.8, 0.1 mM EDTA, 0.1 % NP40, 0.1 mM PMSF, 1 mM 베타-머캅토에탄올)에서 막자 A로 25회 후, 막자 B로 25회로 가볍게 균질화시켰다. 다음으로, 3 ㎖의 HB를 전체 5 ㎖까지 첨가하고, 5분 동안 얼음 상에서 유지시켰다. 구배 원심분리를 위해, 5 mM CaCl, 3 mM $Mg(Ac)_2$, 10 mM 트리스 pH 7.8, 0.1 mM PMSF, 1 mM 베타-머컵토에탄올을 함유하는 5 ㎖의 50% OptiPrep™ 밀도 구배 배지(Sigma)를 첨가하고, 혼합하였다. 용해질을 코니컬 30 ㎖ 원심분리 튜브(Beckman Coulter, SW28 rotor) 내의 10 ㎖의 29% 등장성 OptiPrep™ 용액의 상부에 가볍게 로딩하였다. 샘플을 4℃에서 30분 동안 10,100 xg(7,500 rpm)에서 원심분리시켰다. 상층액을 제거하고, 핵 펠렛을 65 mM 베타-글리세로포스페이트(pH 7.0), 2 mM $MgCl_2$, 25 mM KCl, 340 mM 수크로스 및 5% 글리세롤에 가볍게 재현탁시켰다. 정제된 핵의 수 및 품질을 명시야 현미경검사를 이용하여 조절하였다.

[1991] **세포 핵 분류**

[1992] 정제된 GFP-양성(GFP^+) 및 음성(GFP^-) 무결함 핵을 Vybrant® DyeCycle™ Ruby 염색(1:500, Life Technologies)으로 공동-표지시키고, BD FACSAria III(Koch Institute Flow Cytometry Core, MIT)를 이용하여 분류하였다. GFP^+ 및 GFP^- 핵을 1% BSA로 코팅되고, 400 ㎕의 재현탁 완충액(65 mM 베타-글리세로포스페이트 pH 7.0, 2 mM $MgCl_2$, 25 mM KCl, 340 mM 수크로스 및 5% 글리세롤)을 함유하는 1.5 ㎖ 에펜도르프 튜브에서 수거하였다. 분류 후, 모든 샘플을 얼음 상에서 유지시키고, 4℃에서 20분 동안 10,000 xg에서 원심분리시켰다. 핵 펠렛을 -80℃에서 저장하거나, 하류 공정을 위해 직접 이용하였다.

[1993] **게놈 DNA 추출 및 서베이어™ 검정**

[1994] sgRNA의 기능 시험을 위해, 50-70% 컨플루언스(confluence) N2a 세포를 단일 PCR 증폭된 sgRNA 및 SpCas9 벡터와 함께 공동-트랜스펙션시켰다. SpCas9로만 트랜스펙션된 세포를 음성 대조군으로 제공하였다. 세포를 트랜스펙션 48시간 후에 수거하고, DNA를 제조업체의 프로토콜에 따라 DNeasy Blood & Tissue Kit(Qiagen)를 이용하여 추출하였다. AAV1 형질도입된 일차 뉴런으로부터 게놈 DNA를 분리시키기 위해, DNeasy Blood & Tissue Kit를 제조업체의 설명서에 따라 AAV 형질도입 7일 후에 이용하였다.

[1995] 분류된 핵 또는 해부된 조직을 30분 동안 55℃에서 용해 완충액(10 mM Tris, pH 8.0, 10 mM NaCl, 10 mM EDTA, 0.5 mM SDS, 프로테이나제 K(PK, 1mg/㎖) 및 RNAse A)에서 용해시켰다. 다음으로, 클로로포름-페놀 추출을 수행한 후, 표준 절차에 따라 에탄올을 이용하여 DNA 침전시켰다. DNA를 TE 완충액(10 mM Tris pH 8.0, 0.1 mM EDTA)에 최종적으로 재현탁시키고, 하류 분석에 이용하였다. 개별적 sgRNA의 기능 시험을 부록 표 2에 나열된 PCR 프라이머를 이용하여 서베이어™ 뉴클레아제 검정(Transgenomics)에 의해 수행하였다. 밴드 강도 정량화를 이전에 기재된 바와 같이 수행하였다[11].

[1996] **RNA 라이브러리 제조 및 시퀀싱**

[1997] *Mecp2*를 표적화하는 가이드와 함께 SpCas9(4마리의 동물) 또는 *lacZ*를 표적으로 하는 gRNA와 함께 SpCas9의 양쪽 바이러스 전달 2주 후, 치아이랑을 얼음 저온 DPBS(Life Sciences)에서 신속히 해부하고, RNA-레이터(RNA-later) 용액(Ambion)으로 즉시 옮겼다. 4℃에서 24시간 후, 조직을 -80℃로 이동시켰다. 100개의 표적화된 뉴런 핵의 집단을 1% 2-머캅토에탄올(Qiagen)로 보충된 10 ㎕ TCL 완충액으로 FACS 분류하였다. 원심분리 후, 샘플을 -80℃에서 즉시 동결시켰다. RNA를 제조업체의 설명서에 따라 AMPure RNAcleanXP SPRI 비드(Beckman Coulter Genomics)에 의해 정제하고, 80% 에탄올로 3회 세척하였고, 최종 용리를 생략하였다. 포획된 RNA를 갖는 비드를

공기 건조시키고, cDNA 합성을 위해 즉시 처리하였다. 핵이 없는 샘플을 음성 대조군으로 이용하였다. 3개 집단 샘플을 역전사효소 효소를 0.1 ㎕의 Maxima H Minus 효소(200 U/㎕, Thermo Scientific)로만 대체하고, PCR 반응을 25 ㎕의 부피로 규모 감소시킨 SMART-seq2 프로토콜[12]에 따라 cDNA 라이브러리 제조물에서 각각의 동물에 대해 전체 24개 집단의 샘플에 대해 이용하였다. 태그멘테이션(tagmentation) 반응 및 최종 PCR 증폭을 하기 변형과 함께 Nextera XT DNA 샘플 제조 키트(Illumina)를 이용하여 수행하였다. 모든 반응 부피를 4배만큼 규모 축소시키고, 라이브러리를 2.5 ㎕의 각각의 샘플을 취함으로써 PCR 증폭 단계 후에 풀링시켰다. 풀링된 라이브러리를 세척하고, 2라운드의 0.7 부피의 AMPure XP SPRI 비드 정화(Beckman Coulter Genomics)를 이용하여 크기-선택하였다. 샘플을 라이브러리의 품질을 확인하기 위해 고-민감도 DNA 칩(Agilent) 상에 로딩한 한편, Qubit 고-민감성 DNA 키트(Invitrogen)로 정량화를 수행하였다. 풀링된 라이브러리를 4 nM 및 12 pmol의 최종 농도로 희석시키고, 75 bp의 쌍을 이룬 말단 판독과 함께 Illumina Miseq를 이용하여 시퀀싱하였다.

[1998] **RNA 라이브러리 데이터 분석**

[1999] Bowtie2 지수를 마우스 mm9 UCSC 게놈 및 공지된 유전자 전사체를 기초로 하여 생성시켰고[13], 쌍을 이룬 말단 판독을 명령 행 옵션 -q --phred33-quals -n 2 -e 99999999 -l 25 -I 1 -X 1000 -a -m 200 -p 4 --chunkmbs 512와 함께 Bowtie2를 이용하여 상기 지수에 직접 정렬시켰다. 다음으로, RSEM v1.27을 Bowtie2에 의해 생성된 정렬에 대한 디폴트 파라미터와 함께 수행하여 발현 수준을 평가하였다. RSEM의 유전자 수준 발현 평가(tau)에 1,000,000을 곱하여 각각의 유전자에 대한 백만개의 전사물 당 개수(TPM) 평가를 수득하고, TPM 평가를 log2(TPM+1)를 취하여 log-스페이스(log-space)로 전환시켰다. 전환된 발현 수준이 2(log2(TPM+1) 척도) 이상인 경우에 유전자가 검출된 것으로 간주하였다. 8000개 미만의 유전자가 검출된 경우 라이브러리를 필터링시켰다. 상기 기준을 기초로 하여, 하류 분석으로부터 4개의 라이브러리를 필터링시키고 배제시켰다. 대조군 동물과 Mecp2 sgRNA 발현 동물 사이의 차별적으로 발현된 유전자를 찾기 위해, 20개의 무작위 과변이 수행에서 스튜던츠 *t*-검정(Student's *t*-test)(Matlab V2013b) 및 교차 타당도를 이용하였으며, 여기서 각각의 수행에서 각각의 동물로부터 하나의 라이브러리를 배제하도록 무작위적으로 선택하였다(이는 *t*-검정 각 시기에서 이용된 전체 12개의 라이브러리를 초래하였다). 각각의 샘플에 대해 0.9 분위수(일반적으로 약 5 log2(TPM+1)) 초과의 평균 발현 수준을 갖는 모든 유전자에 대해 *t*-검정을 수행하였다. 이후, 과변이 수행의 ⅓ 초과에서 유의(p<0.01)한 유전자를 선택하였다. 샘플 전체에 걸친 상기 유전자의 log2(TPM+1) 발현 수준을 계층적 군집분석(Matlab V2013b)을 이용하여 군집화시켰다.

[2000] **면역형광 염색**

[2001] 세포 배양: 일차 뉴런의 면역형광 염색을 위해, 세포를 실온에서 20분 동안 4% 파라포름알데하이드(PFA)를 이용하여 바이러스 전달 7일 후에 고정시켰다. PBS를 이용하여 3회 세척한 후, 세포를 실온에서 30분 동안 PBS 중 5% 일반 염소 혈청(NGS)(Life Technologies), 5% 당나귀 혈청(DS)(Sigma) 및 0.1% Triton-X100(Sigma)로 블로킹시켰다. 세포를 실온에서 1시간 또는 4℃에서 밤새 2.5% NGS, 2.5% DS 및 0.1% Triton-X100 중 일차 항체와 함께 인큐베이션하였다. PBST로 3회 세척한 후, 세포를 실온에서 1시간 동안 이차 항체와 함께 인큐베이션하였다. 최종적으로, 커버슬립을 DAPI를 갖는 VECTASHIELD HardSet 마운팅 배지(Vector Laboratories)를 이용하여 마운팅시키고, Zeiss AxioCam Ax10 현미경 및 Axiocam MRm 카메라를 이용하여 영상화시켰다. 이미지를 Zen 2012 소프트웨어(Zeiss)를 이용하여 처리하였다. 정량화를 ImageJ 소프트웨어 1.48 h 및 Neuron 검출기 플러그인을 이용하여 수행하였다.

[2002] 마우스를 치사 용량의 케타민/자일라진에 의해 바이러스 전달 4주 후에 희생시키고, PBS 및 이후 PFA로 경심 관류시켰다. 고정된 조직을 진동절편기(vibratome)(Leica, VT1000S)를 이용하여 절편화시켰다. 다음으로, 30 ㎛ 절편을 소듐 시트레이트 완충액(10 mM 트리-소듐 시트레이트 데하이드레이트, 0.05% Tween20, pH 6.0)에서 2분 동안 비등시키고, 20분 동안 실온에서 냉각시켰다. 절편을 1시간 동안 TBST(137 mM NaCl, 20 mM Tris pH 7.6, 0.2% Tween-20) 중 4% 일반 염소 혈청(NGS)으로 블로킹시켰다. 파라핀 절편을 박절기(microtom)(Leica RM2125 RTS)를 이용하여 8 ㎛로 절단하고, 이전에 기재된 바와 같이 염색하였다[14].

[2003] 절편을 4℃에서 밤새 4% NGS와 함께 TBST에 희석된 일차 항체와 함께 인큐베이션하였다. TBST 중에서 3회 세척 후, 샘플을 이차 항체와 함께 인큐베이션하였다. TBST로 3회 세척 후, 절편을 DAPI를 갖는 VECTASHIELD HardSet 마운팅 배지를 이용하여 마운팅시키고, 동일초점 현미경(Zeiss LSM 710, Ax10 ImagerZ2, Zen 2012 Software)으로 시각화시켰다.

[2004] 다음과 같은 일차 항체를 사용하였다: 토끼 항-Dnmt3a(Santa Cruz, 1:100); 토끼 항-MeCP2(Millipore, 1:200);

마우스 항-NeuN(Millipore, 1:50 내지 1:400); 닭 항-GFAP(Abcam, 1:400); 마우스 항-Map2(Sigma, 1:500); 닭 항-GFP(Aves labs, 1:200 내지 1:400); 마우스 항-HA(Cell Signaling, 1:100). 이차 항체: AlexaFluor®488, 568 또는 633(Life Technologies, 1:500 내지 1:1,000).

[2005] **LIVE/DEAD® 검정의 정량화**

[2006] 대조군 및 형질도입된 일차 뉴런을 제조업체의 설명서에 따라 LIVE/DEAD® 검정(Life technologies)을 이용하여 염색하였다. GFP-KASH 발현으로부터 GFP-신호 간섭을 피하기 위해, 세포를 DEAD(에티디움 동종이합체) 및 DAPI(모든 세포) 단독에 대해 염색하였다. 염색된 세포를 형광 현미경검사를 이용하여 영상화시키고, DEAD, GFP 및 DAPI 양성 세포를 ImageJ 1.48h 소프트웨어 및 Neuron 검출기 플러그인을 이용하여 계수하였다.

[2007] **웨스턴 블롯 분석**

[2008] 형질도입된 일차 피질 뉴런(24웰, 바이러스 전달 7일 후) 및 형질도입된 조직 샘플(바이러스 전달 4주 후)을 0.1% SDS 및 프로테아제 억제제(Roche, Sigma)를 함유하는 50 ㎕의 얼음-저온 RIPA 완충액(Cell Signaling) 중에서 용해시켰다. 세포 용해질을 Bioruptor 초음파파쇄기(Diagenode)에서 5분 동안 초음파 파쇄시키고, 단백질 농도를 BCA 단백질 검정 키트(Pierce Biotechnology, Inc.)를 이용하여 결정하였다. 단백질 용해질을 SDS-PAGE 샘플 완충액에 용해시키고, 4-15% 트리스-HCl 겔(Bio-Rad) 상에서 환원 조건하에서 분리시키고, 다음과 같은 일차 항체를 이용한 웨스턴 블로팅에 의해 분석하였다: 토끼 항-Dnmt3a(Santa Cruz, 1:500), 마우스 항-Dnmt1(Novus Biologicals, 1:800), 토끼 항-Mecp2(Millipore, 1:400), 토끼 항-튜불린(Cell Signaling, 1:10,000) 후, 이차 항-마우스 및 항-토끼 HRP 항체(Sigma-Aldrich, 1:10,000). GAPDH를 항-GAPDH 항체에 커플링된 토끼 HRP(Cell Signaling, 1:10,000)로 직접 시각화시켰다. 튜불린 또는 GAPDH를 로딩 대조군으로 제공하였다. 블롯을 ImageLab 4.1 소프트웨어(BioRad)와 함께 ChemiDoc™ MP 시스템으로 영상화시키고, ImageJ 소프트웨어 1.48h를 이용하여 정량하였다.

[2009] **지연 전후관계 공포 조건화(Delay Contextual Fear Conditioning)(DCFC)**

[2010] 12주령의 C57BL/6N 수컷 마우스의 등쪽 및 배쪽 치아이랑으로의 양쪽 SpCas9/DNMT 3xsgRNA 전달 8주 후, 7일 동안 동물을 실험자 및 행동 실험실(behavior room)에 익숙하게 하였다. SpCas9/GFP-KASH 주사된 한배 새끼를 대조군으로 제공하였다. DCFC 1일에서, 시험 전 및 후에 마우스를 청각 신호로부터 보호하기 위해 분리된 앤터룸(anterroom)에 배치하였다. 개별적 마우스를 FC 챔버(Med Associates Inc.)에 배치하고, 12분 익숙화(habituation) 기간을 이행하였다. 익숙화 후, 마우스를 이들의 홈케이지(homecage)로 다시 배치하였다. 다음날(훈련일) 개별적 마우스를 챔버에 배치하고, 4분 동안 익숙하게 하였다. 또 다른 20초(울림 전) 간격 후, 85 dB, 2.8 kHz 수준의 울림(청각 신호)을 20초 동안 제시한 후, 18초 지연 간격 후, 발-충격(0.5 mA, 2초)을 제공하였다. 발-충격 후, 20초 울림전 기간과 함께 다음의 동일한 시험 시작에 앞서 40초 간격(울림/충격 후)을 선행시켰다. 훈련 시험을 6회 반복한 후, 마우스를 이들의 홈케이지로 다시 배치하였다. 3일(시험일)에서, 마우스를 먼저 3분 동안 훈련 상황 챔버에 배치하였다. 다음으로, 마우스를 20초 간격 후, 20초 울림 및 60초 울림 후 간격으로 시작하는 4x 100초 시험 시도를 수행하였다. 최종적으로, 마우스를 변경 상황-검사 챔버(편평한 바닥 대 격자, 4합체 대 7합체 챔버, 바닐린 센트(vanillin scent))에 배치하고, 시험 시도를 반복하였다. 얼어붙는 행동을 기록하였고, 분석을 수작업으로 맹검 오프라인(blind off-line) 방식으로 수행하였고, Noldus EthoVision XT 소프트웨어(Noldus Information Technology)를 이용하여 확인하였다.

[2011] **심층 시퀀싱 분석 및 삽입-결실 검출**

[2012] 뇌에서 CRISPR-SpCas9에 의해 표적화되는 DNMT 패밀리 일원에 대한 잠재적인 표적외를 찾기 위해 CRISPR Design Tool(http://crispr.mit.edu/)을 이용하였다. 치아이랑으로부터의 표적화된 세포 핵을 바이러스 전달 12주 후에 FACS 분류하고, 게놈 DNA를 상기 기재된 바와 같이 정제하였다. 각각의 관심대상 유전자에 대해, CRISPR 표적 부위의 측접한 게놈 영역을 융합 PCR 방법에 의해 증폭하여, Illumina P5 어댑터뿐만 아니라 독특한 샘플-특이적 바코드를 표적 앰플리콘에 부착시켰다(표적내 및 표적외 프라이머에 대해서는, 부록 표 3 참조)[15]. 바코드화되고 정제된 DNA 샘플을 Qubit 2.0 Fluorometer(Life Technologies)에 의해 정량하고, 등몰비로 풀링시켰다. 이후, 시퀀싱 라이브러리를 300 bp 판독 길이를 갖는 Illumina MiSeq Personal Sequencer(Life Technologies)로 시퀀싱하였다.

[2013] MiSeq 판독을 이전에 기재된 바와 같이 분석하였다[15]. 간단히, 판독을 프레드 품질(Phred quality)(Q 스코어)에 의해 필터링시키고, 스미스-워터맨 알고리즘(Smith-Waterman algorithm)을 이용하여 표적 부위의 상류 및 하류

의 50개 뉴클레오티드의 게놈 영역에 정렬시켰다. 표적 부위의 상류 5개의 뉴클레오티드로부터 하류 5개의 뉴클레오티드의 정렬된 영역(전체 30 bp)에서 삽입-결실을 평가하였다. 추정 절단 사건으로서 삽입-결실의 포함 또는 배제를 평가하기 위해 각각의 샘플에 대해 음성 대조군을 이용하였다. 출원인은 음성 대조군 샘플의 데이터로부터 판독당 표적 영역당 오류 비율(per-target-region-per-read error rate)을 이용하여 실제-삽입-결실을 갖는 표적-영역을 갖는 판독의 분획에 대한 최대 가능도 추정량(MLE)을 계산하였다. 각각의 표적에 대한 MLE 스코어 및 절단 비율이 부록 표 1에 나열되어 있다.

[2014] **통계 분석**

[2015] 모든 실험을 최소 2개의 독립적 생물학적 복제물을 이용하여 수행하였다. 스튜던츠 양측 *t*-검정을 이용하여 Prism6(GraphPad)로 통계를 수행하였다.

[2016] 부록 표 1. DNMT 표적화를 위한 표적외 분석

	유전자	GI	*잠재적인 표적외 서열*	MLE(%)	SEM	서열 번호
Dnmt1	*Abca1*	NM_013454	GGAGCTGGAGCTGTTCACGTTGG	0.0000	0.00	
	Mctp1	NM_030174	CGGGCAGCAGATGTTCGCGTAGG	0.0806	0.08	
	Exd2	NM_133798	AGGGCTTGAGATGTTCGGGCTGG	0.0612	0.06	
	Pik3r6	NM_001004435	CCGGCTGGGGCTGTCCTCGCTAG	0.0000	0.00	
	Sobp	NM_175407	CGGGGTGCAGCTGCTCACGCCAG	0.0000	0.00	
	Vac14	NM_146216	CTGGCGGGAGCTGGTCGCGTGAG	0.0083	0.00	
Dnmt3a	*Efemp2*	NM_021474	TGAGCATGGGCCGCTGGCGGTGG	0.0050	0.01	
	Bmpr1b	NM_001277217	ATGGCATAGGCCGCTGACAGAGG	0.0117	0.01	
	Syce1	NM_001143765	TTGGCATGGTGAGCTGGCGGGGG	0.0067	0.00	
	Atp8b3	NM_026094	TGGGCAGGGGTCTCTGAGGGCAG	0.0067	0.01	
	Rdh11	NM_021557	TTGGCATGGGTCTCTTACCAAGG	0.0017	0.00	
Dnmt3b	*Hecw2*	NM_001001883	ACATGGTTCCAGTGGGTATGTAG	0.0000	0.00	
	Plekhg3	NM_153804	GGAGGTGGGCAGCGGGTATGTAG	0.0954	0.01	
	Cdc25b	NM_001111075	AGAAGGTCCCCGCGGGCATGGAG	0.2421	0.12	
	Top1mt	NM_028404	GGAGGGAACCAGCCGGTATGGGG	0.0167	0.01	
	Sesn2	NM_144907	AGAGAGTGGCAGTGGGTAAGCAG	0.0000	0.00	
	Ncan	NM_007789	AGAGGTGGCCAGCGGGCAGGAAG	0.0017	0.00	
	Nacad	NM_001081652	TGAGGGGGCCAGCTGGGATGCAG	1.6254	0.76	

[2017]

[2018] 부록 표 2. 서베이어 검정에서 사용된 PCR 프라이머

유전자	정방향 프라이머 서열(5'-3')	서열 번호	역방향 프라이머 서열(5'-3')	서열 번호
Mecp2	GGTCTCATGTGTGGCACTCA		TGTCCAACCTTCAGGCAAGG	
Dnmt3a	ATCCCTCCTCAGAGGGTCAGC		TACCTCATGCACAGCTAGCACC	
Dnmt1	TTCGGGCATAGCATGGTCTTCC		GTTCTATTTCAGAGGGCTGATCCC	
Dnmt3b	GTTCTGAGCCGCACAGTTTGG		GGATAAGAAGGGACAATACAGG	

[2019]

[2020] 부록 표 3. 표적 내 및 표적외 게놈 유전자좌 증폭을 위해 사용된 프라이머

유전자	정방향 프라이머 서열(5'-3')	서열 번호	역방향 프라이머 서열(5'-3')	서열 번호
Dnmt1	GCCGGGGTCTC GTTCAGAGCT		CTACCGCCTGCGGA CATGGT	
Dnmt3a	CCTGTCTCTCTGT CCTAGGGCTCC		CCGTTTGCTGATGTAGTA GGGGTCC	
Dnmt3b	CCCACAGGAAA CAATGAAGGGAGAC		CATCCTTCGTGTCT GAGGACTGGTC	
Abca1	CCCTGACACCAGC TGTTCAGCAC		CTCTGGGTGAC CACACACGATGC	
Mctp1	GAGCAGGCAGA GCCGAGCAAG		GGAGAGCGTCC GCCAGGAG	
Exd2	GGGTCTTGTTGTG AGTAGGGTGTG		GAAGCTCTCTTAA CTACTGTTC	
Pik3r6	CCTGGAATACTAT TTCCACGCCG		CAGGCCCTAGCAGCG AGCAG	
Sobp	GCAGCACACTCCA CCCTCACAT		GGAAGGGGCTTTCC TCCGAGC	
Vac14	CGGCGTCACG TGACCTGAGTAAC		GCTCCGACCCTGCT CTCCCA	
Bfcmp2	GTGTCTGCCTC GCTCTGCTGC		CCTGTTCATCAGGCTC GTAGCCC	
Bmpr1b	CTATCTGAAATCC ACCACCTTAGACGC		CGATTGCTGGCTTGC CTTGAG	
Sycc1	GCCTGAGGGGG CCAGAGGT		GGTTCGCGTCCGCC CGCGTGAT	
Atp8b3	GGGACTCC CCGGGTGGTG		GAGAGGTGGTC CTGTCGCCTATG	
Rdh11	GACCCTGTGTTT CAAGTCTCTCTG		CCCAGCAGGTCACA GCTGACATC	
Hecw2	GGCCATCCAGTAC ATTCAATACG		AGCACAGTATGTATTC TATAAAATAATACGAC	
Plekhg3	GCAGAAGCCGT GACTCACAGCA		GTGGGAGGGGACAG AGACCATG	
Cdc25b	CTTGTGCTTG TGATTCTGTCCTTACTGC		CCTTACCTGTTCCTCT TCCTTATCCAGC	
Topbmt	CGAGAAGTC GATGCAGACACTTCAA		ATACCCAGTCCAC ATCCCTGCC	
Sesn2	GCTGAAGACTGGC GAGCACAGCT		CCTCTGCATCTCCCTCAGGAAGTATT	
Ncan	GACCTGAATGTTG TGGCTGAGAGTCC		GCCTCCTGTC CCCAGGTCCC	
Nacad	CCCTCACGTTCC TGTCCAGCAA		CACTAGGCTT GGGCTGCCCTCT	

[2021]

[2022] 참고자료

[2023] 1. Ostlund, C. et al. Dynamics and molecular interactions of linker of nucleoskeleton and cytoskeleton (LINC) complex proteins. *J Cell Sci* 122, 4099-4108 (2009).

2. Gray, S.J. et al. Optimizing promoters for recombinant adeno-associated virus-mediated gene expression in the peripheral and central nervous system using self-complementary vectors. *Hum Gene Ther* 22, 1143-1153 (2011).

3. Levitt, N., Briggs, D., Gil, A. & Proudfoot, N.J. Definition of an efficient synthetic poly(A) site. *Genes Dev* 3, 1019-1025 (1989).

4. Jinek, M. et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816-821 (2012).

5. Cong, L. et al. Multiplex genome engineering using CRISPR/Cas systems. *Science* 339, 819-823 (2013).

6. McClure, C., Cole, K.L., Wulff, P., Klugmann, M. & Murray, A.J. Production and titering of recombinant adeno-associated viral vectors. *J Vis Exp*, e3348 (2011).

7. Konermann, S. et al. Optical control of mammalian endogenous transcription and epigenetic states. *Nature* 500, 472-476 (2013).

8. Banker, G. & Goslin, K. Developments in neuronal cell culture. *Nature* 336, 185-186 (1988).

9. Swiech, L. et al. CLIP-170 and IQGAP1 cooperatively regulate dendrite morphology. *J Neurosci* 31, 4555-4568 (2011).

10. Sholl, D.A. Dendritic organization in the neurons of the visual and motor cortices of the cat. *J Anat* 87, 387-406 (1953).

11. Ran, F.A. et al. Genome engineering using the CRISPR-Cas9 system. *Nature protocols* 8, 2281-2308 (2013).

12. Picelli, S. et al. Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nature methods* 10, 1096-1098 (2013).

13. Fujita, P.A. et al. The UCSC Genome Browser database: update 2011. *Nucleic acids research* 39, D876-882 (2011).

14. Tzingounis, A.V. et al. The KCNQ5 potassium channel mediates a component of the afterhyperpolarization current in mouse hippocampus. *Proceedings of the National Academy of Sciences of the United States of America* 107, 10232-10237 (2010).

15. Hsu, P.D. et al. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nature biotechnology* 31, 827-832 (2013).

[2024]

16. Qiu, P. et al. Mutation detection using Surveyor nuclease. *BioTechniques* 36, 702-707 (2004).

[2025]

[2026] *실시예 40: 핵 태깅 기법로의 추가 연구*

[2027] 본 실시예 38은 Cas9의 에피토프 태깅에 관한 것이다. 요컨대, 출원인은 삼중 에피토프 태그(구체적으로, 3xHA)가 검출 신호를 개선시키는 것을 발견하였다.

[2028] **재료 및 방법**

[2029] *세포 배양 및 트랜스펙션*

[2030] 인간 배아 신장(HEK) 세포주 293FT(Life Technologies) 또는 마우스 Hepa1-6(Sigma-Aldrich) 세포주를 5% CO_2 인큐베이션과 함께 37℃에서 10% 소 태아 혈청(HyClone), 2 mM GlutaMAX(Life Technologies), 100 U/㎖ 페니실린, 및 100 ㎍/㎖ 스트렙토마이신이 보충된 둘베코 변형 이글 배지(Dulbecco's modified Eagle's

Medium)(DMEM)에서 유지시켰다.

[2031] 세포를 트랜스펙션 24 시간 전에 120,000 세포/웰의 밀도로 24-웰 플레이트(Corning) 상에 시딩하였다. 세포를 제조업체의 권장 프로토콜에 따라 80% 내지 90% 컨플루언스에서 Lipofectamine 2000(Life Technologies)을 사용하여 트랜스펙션시켰다. 전체 500 ng의 Cas9 플라스미드 및 100 ng의 U6-sgRNA PCR 산물을 트랜스펙션시켰다.

[2032] 게놈 변형에 대한 서베이어 뉴클레아제 검정

[2033] 293FT 및 HUES62 세포를 상기 기재된 바와 같이 DNA로 트랜스펙션시켰다. 세포를 게놈 DNA 추출 전에 트랜스펙션 후 72 시간 동안 37℃에서 인큐베이션하였다. 게놈 DNA를 제조업체의 프로토콜에 따라 QuickExtract DNA 추출 용액(Epicentre)을 사용하여 추출하였다. 간단히, 펠렛화된 세포를 QuickExtract 용액에 재현탁시키고, 15 분 동안 65℃, 15 분 동안 68℃, 및 10 분 동안 98℃에서 인큐베이션하였다.

[2034] 각각의 유전자에 대한 CRISPR 표적 부위에 측접한 게놈 영역을 PCR 증폭시키고, 산물을 제조업체의 프로토콜에 따라 QiaQuick Spin Column(Qiagen)을 사용하여 정제하였다. 전체 400 ng의 정제된 PCR 산물을 2 마이크로리터의 10X Taq DNA 중합효소 PCR 완충액(Enzymatics) 및 초순수와 20 마이크로리터의 최종 부피로 혼합하고, 헤테로듀플렉스 형성을 가능하게 하기 위해 재-아닐링시켰다: 10 분 동안 95℃, -2℃/s로 95℃에서 85℃로 램핑, -0.25℃/s로 85℃에서 25℃로 램핑, 및 1 분 동안 25℃에서 유지. 재-아닐링 후, 제조업체의 권장 프로토콜에 따라 산물에 서베이어 뉴클레아제 및 서베이어 인핸서 S(Transgenomics)를 처리하고, 4% 내지 20% Novex TBE 폴리아크릴아미드 겔(Life Technologies)에서 분석하였다. 겔을 30 분 동안 SYBR Gold DNA 염색제(Life Technologies)로 염색하고, Gel Doc 겔 영상화 시스템(Bio-rad)으로 영상화하였다. 정량화는 상대 밴드 강도를 기초로 하였다. 삽입결실 백분율을 식 $100 \times (1 - (1 - (b + c) / (a + b + c))^{1/2})$에 의해 결정하였고, 상기 식에서 a는 분해되지 않은 PCR 산물의 통합된 강도이고, b 및 c는 각각의 절단 산물의 통합된 강도이다.

[2035] *웨스턴 블롯*

[2036] HEK 293FT 세포를 트랜스펙션시키고, 프로테아제 억제제(Roche)가 보충된 1X RIPA 완충액(Sigma-Aldrich)에서 용해시켰다. 용해질을 Bolt 4% 내지 12% Bis-Tris Plus Gels(Invitrogen) 상에 로딩하고, 니트로셀룰로스 막으로 옮겼다. 막을 0.1% Tween-20 및 5% 블로킹 작용제(G-Biosciences)를 함유하는 트리스-완충 염수에서 블로킹시켰다. 막을 토끼 항-FLAG(1:5,000, Abcam), HRP-컨쥬게이션된 항-GAPDH(1:5,000 Cell Signaling Technology), 및 HRP-컨쥬게이션된 항-토끼(1:1,000) 항체로 탐지시키고, Gel Doc XR+ 시스템(Bio-Rad)으로 시각화시켰다.

[2037] **참고자료**

R. E. Amir *et al.*, Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nature genetics* **23**, 185 (Oct, 1999).

Banker G, Goslin K. Developments in neuronal cell culture. Nature. 1988 Nov 10;336(6195):185-6.

Bedell, V.M. et al. In vivo genome editing using a high-efficiency TALEN system. Nature 491, 114-U133 (2012).

Bhaya, D., Davison, M. & Barrangou, R. CRISPR-Cas systems in bacteria and archaea: versatile small RNAs for adaptive defense and regulation. Annu Rev Genet 45, 273-297 (2011).

Bobis-Wozowicz, S., Osiak, A., Rahman, S.H. & Cathomen, T. Targeted genome editing in pluripotent stem cells using zinc-finger nucleases. Methods 53, 339-346 (2011).

Boch, J. et al. Breaking the code of DNA binding specificity of TAL-type III effectors. Science 326, 1509-1512 (2009).

Bogenhagen, D.F. & Brown, D.D. Nucleotide sequences in Xenopus 5S DNA required for transcription termination. Cell 24, 261-270 (1981).

[2038]

Bultmann, S. et al. Targeted transcriptional activation of silent oct4 pluripotency gene by combining designer TALEs and inhibition of epigenetic modifiers. Nucleic Acids Res 40, 5368-5377 (2012).

C. Burger, K. Nash, R. J. Mandel, Recombinant adeno-associated viral vectors in the nervous system. *Human gene therapy* **16**, 781 (Jul, 2005).

Carlson, D.F. et al. Efficient TALEN-mediated gene knockout in livestock. Proc Natl Acad Sci U S A 109, 17382-17387 (2012).

M. Chahrour, H. Y. Zoghbi, The story of Rett syndrome: from clinic to neurobiology. *Neuron* **56**, 422 (Nov 8, 2007).

Chen, F.Q. et al. High-frequency genome editing using ssDNA oligonucleotides with zinc-finger nucleases. Nat Methods 8, 753-U796 (2011).

R. Z. Chen, S. Akbarian, M. Tudor, R. Jaenisch, Deficiency of methyl-CpG binding protein-2 in CNS neurons results in a Rett-like phenotype in mice. *Nature genetics* **27**, 327 (Mar, 2001).

Cho, S.W., Kim, S., Kim, J.M. & Kim, J.S. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. Nat Biotechnol 31, 230-232 (2013).

Christian, M. et al. Targeting DNA double-strand breaks with TAL effector nucleases. Genetics 186, 757-761 (2010).

Cong, L. et al. Multiplex genome engineering using CRISPR-Cas systems. Science 339, 819-823 (2013).

Deltcheva, E. et al. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. Nature 471, 602-607 (2011).

Deveau, H., Garneau, J.E. & Moineau, S. CRISPR-Cas system and its role in phage-bacteria interactions. Annu Rev Microbiol 64, 475-493 (2010).

Ding, Q. et al. A TALEN genome-editing system for generating human stem cell-based disease models. Cell Stem Cell 12, 238-251 (2013).

J. Feng *et al.*, Dnmt1 and Dnmt3a maintain DNA methylation and regulate synaptic function in adult forebrain neurons. *Nature neuroscience* **13**, 423 (Apr, 2010).

Y. Fu *et al.*, High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nature biotechnology* **31**, 822 (Sep, 2013).

[2039]

Garneau, J.E. et al. The CRISPR-Cas bacterial immune system cleaves bacteriophage and plasmid DNA. Nature 468, 67-71 (2010).

Gasiunas, G., Barrangou, R., Horvath, P. & Siksnys, V. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. Proc Natl Acad Sci U S A 109, E2579-2586 (2012).

Geurts, A.M. et al. Knockout Rats via Embryo Microinjection of Zinc-Finger Nucleases. Science 325, 433-433 (2009).

Gray SJ, Foti SB, Schwartz JW, Bachaboina L, Taylor-Blake B, Coleman J, Ehlers MD, Zylka MJ, McCown TJ, Samulski RJ. Optimizing promoters for recombinant adeno-associated virus-mediated gene expression in the peripheral and central nervous system using self-complementary vectors. Hum Gene Ther. 2011 Sep;22(9):1143-53. doi: 10.1089/hum.2010.245.

Guschin, D.Y. et al. A rapid and general assay for monitoring endogenous gene modification. Methods Mol Biol 649, 247-256 (2010).

Hasty, P., Rivera-Perez, J. & Bradley, A. The length of homology required for gene targeting in embryonic stem cells. Mol Cell Biol 11, 5586-5591 (1991).

Horvath, P. & Barrangou, R. CRISPR-Cas, the immune system of bacteria and archaea. Science 327, 167-170 (2010).

P. D. Hsu *et al.*, DNA targeting specificity of RNA-guided Cas9 nucleases. *Nature biotechnology* 31, 827 (Sep, 2013).

Hsu, P.D. & Zhang, F. Dissecting neural function using targeted genome engineering technologies. ACS Chem Neurosci 3, 603-610 (2012).

Hwang, W.Y. et al. Efficient genome editing in zebrafish using a CRISPR-Cas system. Nat Biotechnol 31, 227-229 (2013).

Jiang, W., Bikard, D., Cox, D., Zhang, F. & Marraffini, L.A. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. Nat Biotechnol 31, 233-239 (2013).

Jinek, M. et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science 337, 816-821 (2012).

Jinek, M. et al. RNA-programmed genome editing in human cells. eLife 2, e00471 (2013).

[2040]

Kaplitt, M.G., et al., Safety and tolerability of gene therapy with an adeno-associated virus (AAV) borne GAD gene for Parkinson's disease: an open label, phase I trial. Lancet. 2007 Jun 23;369(9579):2097-105.

S. Konermann *et al.*, Optical control of mammalian endogenous transcription and epigenetic states. *Nature* 500, 472 (Aug 22, 2013).

Levitt N. Briggs D. Gil A. Proudfoot N.J. Definition of an efficient synthetic poly(A) site. Genes Dev. 1989;3:1019–1025.

Y. Li *et al.*, Global transcriptional and translational repression in human-embryonic-stem-cell-derived Rett syndrome neurons. *Cell stem cell* **13**, 446 (Oct 3, 2013).

Liu D, Fischer I. Two alternative promoters direct neuron-specific expression of the rat microtubule-associated protein 1B gene. J Neurosci. 1996 Aug 15;16(16):5026-36.

Lopes, V.S., etc al., Retinal gene therapy with a large MYO7A cDNA using adeno-assocaited virus. Gene Ther, 2013 Jan 24. doi: 10.1038/gt 2013.3.[Epub ahead of print]

Mahfouz, M.M. et al. De novo-engineered transcription activator-like effector (TALE) hybrid nuclease with novel DNA binding specificity creates double-strand breaks. Proc Natl Acad Sci U S A 108, 2623-2628 (2011).

Makarova, K.S. et al. Evolution and classification of the CRISPR-Cas systems. Nat Rev Microbiol 9, 467-477 (2011).

Mali, P. et al. RNA-guided human genome engineering via Cas9. Science 339, 823-826 (2013).

S. A. McCarroll, S. E. Hyman, Progress in the genetics of polygenic brain disorders: significant new challenges for neurobiology. *Neuron* **80**, 578 (Oct 30, 2013).

McClure C, Cole KL, Wulff P, Klugmann M, Murray AJ. Production and titering of recombinant adeno-associated viral vectors. J Vis Exp. 2011 Nov 27;(57):e3348. doi: 10.3791/3348.

Michaelis, L.M., Maud "Die kinetik der invertinwirkung.". Biochem. z (1913).

Miller, J.C. et al. An improved zinc-finger nuclease architecture for highly specific genome editing. Nat Biotechnol 25, 778-785 (2007).

Miller, J.C. et al. A TALE nuclease architecture for efficient genome editing. Nat Biotechnol 29, 143-148 (2011).

[2041]

Moscou, M.J. & Bogdanove, A.J. A simple cipher governs DNA recognition by TAL effectors. Science 326, 1501 (2009).Porteus, M.H. & Baltimore, D. Chimeric nucleases stimulate gene targeting in human cells. Science 300, 763 (2003).

Mussolino, C. et al. A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. Nucleic acids research 39, 9283-9293 (2011).

Nathwani, A.C., et al., Adenovirus-associated virus vector-mediated gene transfer in hemophilia B. N Engl J Med. 2011 Dec 22;365(25):2357-65. doi: 10.1056/NEJMoa1108046. Epub 2011 Dec 10.

M. V. Nguyen *et al.*, MeCP2 is critical for maintaining mature neuronal networks and global brain anatomy during late stages of postnatal brain development and in the mature adult brain. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **32**, 10021 (Jul 18, 2012).

Oliveira, T.Y. et al. Translocation capture sequencing: a method for high throughput mapping of chromosomal rearrangements. J Immunol Methods 375, 176-181 (2012).

C. Ostlund et al., Dynamics and molecular interactions of linker of nucleoskeleton and cytoskeleton (LINC) complex proteins. Journal of cell science 122, 4099 (Nov 15, 2009).

Perez, E.E. et al. Establishment of HIV-1 resistance in CD4(+) T cells by genome editing using zinc-finger nucleases. Nat Biotechnol 26, 808-816 (2008).

Qi, L.S. et al. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell 152, 1173-1183 (2013).

F. A. Ran *et al.*, Genome engineering using the CRISPR-Cas9 system. *Nature protocols* 8, 2281 (Nov, 2013).

REMINGTON'S PHARMACEUTICAL SCIENCES (Mack Pub. Co., N.J. 1991)

Reyon, D. et al. FLASH assembly of TALENs for high-throughput genome editing. Nat Biotechnol 30, 460-465 (2012).

Saleh-Gohari, N. & Helleday, T. Conservative homologous recombination preferentially repairs DNA double-strand breaks in the S phase of the cell cycle in human cells. Nucleic Acids Res 32, 3683-3688 (2004).

Sander, J.D. et al. Selection-free zinc-finger-nuclease engineering by context-dependent assembly (CoDA). Nat Methods 8, 67-69 (2011).

[2042]

Sanjana, N.E. et al. A transcription activator-like effector toolbox for genome engineering. Nat Protoc 7, 171-192 (2012).

Sapranauskas, R. et al. The Streptococcus thermophilus CRISPR-Cas system provides immunity in Escherichia coli. Nucleic Acids Res 39, 9275-9282 (2011).

M. Shahbazian *et al.*, Mice with truncated MeCP2 recapitulate many Rett syndrome features and display hyperacetylation of histone H3. *Neuron* **35**, 243 (Jul 18, 2002).

Shen, B. et al. Generation of gene-modified mice via Cas9/RNA-mediated gene targeting. Cell Res 23, 720-723 (2013).

D. A. Sholl, Dendritic organization in the neurons of the visual and motor cortices of the cat. *Journal of anatomy* 87, 387 (Oct, 1953).

Smithies, O., Gregg, R.G., Boggs, S.S., Koralewski, M.A. & Kucherlapati, R.S. Insertion of DNA sequences into the human chromosomal beta-globin locus by homologous recombination. Nature 317, 230-234 (1985).

Soldner, F. et al. Generation of isogenic pluripotent stem cells differing exclusively at two early onset Parkinson point mutations. Cell 146, 318-331 (2011).

L. Swiech et al., CLIP-170 and IQGAP1 cooperatively regulate dendrite morphology. The Journal of neuroscience : the official journal of the Society for Neuroscience 31, 4555 (Mar 23, 2011).

Takasu, Y. et al. Targeted mutagenesis in the silkworm Bombyx mori using zinc finger nuclease mRNA injection. Insect Biochem Molec 40, 759-765 (2010).

Tangri S, et al., Rationally engineered therapeutic proteins with reduced immunogenicity, J Immunol. 2005 Mar 15;174(6):3187-96.

Thomas, K.R., Folger, K.R. & Capecchi, M.R. High frequency targeting of genes to specific sites in the mammalian genome. Cell 44, 419-428 (1986).

Tuschl, T. Expanding small RNA interference. Nat Biotechnol 20, 446-448 (2002).

A. V. Tzingounis et al., The KCNQ5 potassium channel mediates a component of the afterhyperpolarization current in mouse hippocampus. Proceedings of the National Academy of Sciences of the United States of America 107, 10232 (Jun 1, 2010).

Urnov, F.D., Rebar, E.J., Holmes, M.C., Zhang, H.S. & Gregory, P.D. Genome editing with engineered zinc finger nucleases. Nat Rev Genet 11, 636-646 (2010).

[2043]

Valton, J. et al. Overcoming transcription activator-like effector (TALE) DNA binding domain sensitivity to cytosine methylation. J Biol Chem 287, 38427-38432 (2012).

Wang, H. et al. One-Step Generation of Mice Carrying Mutations in Multiple Genes by CRISPR-Cas-Mediated Genome Engineering. Cell 153, 910-918 (2013).

Watanabe, T. et al. Non-transgenic genome modifications in a hemimetabolous insect using zinc-finger and TAL effector nucleases. Nat Commun 3 (2012).

Wilson, E.B. Probable inference, the law of succession, and statistical inference. J Am Stat Assoc 22, 209-212 (1927).

Wood, A.J. et al. Targeted genome editing across species using ZFNs and TALENs. Science 333, 307 (2011).

Wu, S., Ying, G.X., Wu, Q. & Capecchi, M.R. A protocol for constructing gene targeting vectors: generating knockout mice for the cadherin family and beyond. Nat Protoc 3, 1056-1076 (2008).

Z. Wu, H. Yang, P. Colosi, Effect of genome size on AAV vector packaging. *Molecular therapy : the journal of the American Society of Gene Therapy* **18**, 80 (Jan, 2010).

Zhang, F. et al. Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. Nat Biotechnol 29, 149-153 (2011).

Z. Zhou *et al.*, Brain-specific phosphorylation of MeCP2 regulates activity-dependent Bdnf transcription, dendritic growth, and spine maturation. *Neuron* **52**, 255 (Oct 19, 2006).

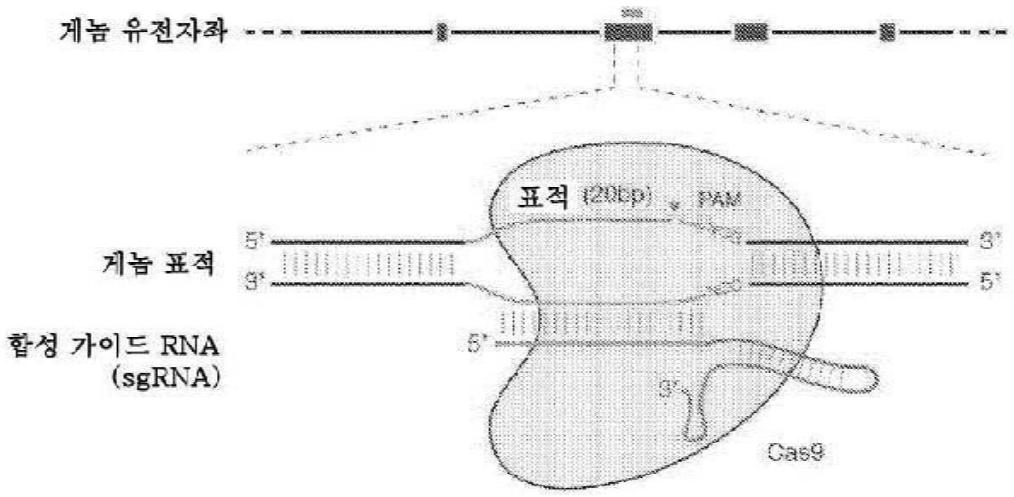
[2044]

[2045] * * *

[2046] 본 발명의 바람직한 구현예가 본원에 제시되고 기재되었으나, 상기 구체예는 단지 예로 제공된 것이 당업자에게 명백할 것이다. 본 발명을 벗어나지 않고 다수의 변동, 변화, 및 치환이 당업자에 의해 이제 발생될 것이다. 본원에 기재된 본 발명의 구현예에 대한 다양한 대안이 본 발명의 실시에서 이용될 수 있음이 이해되어야 한다.

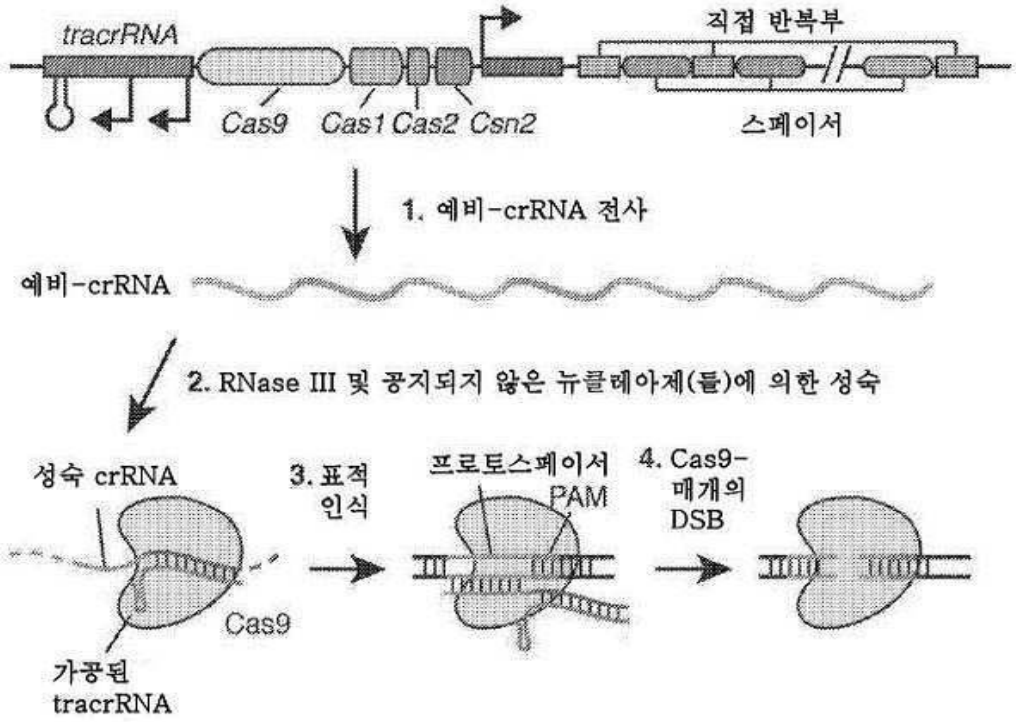
도면

도면1

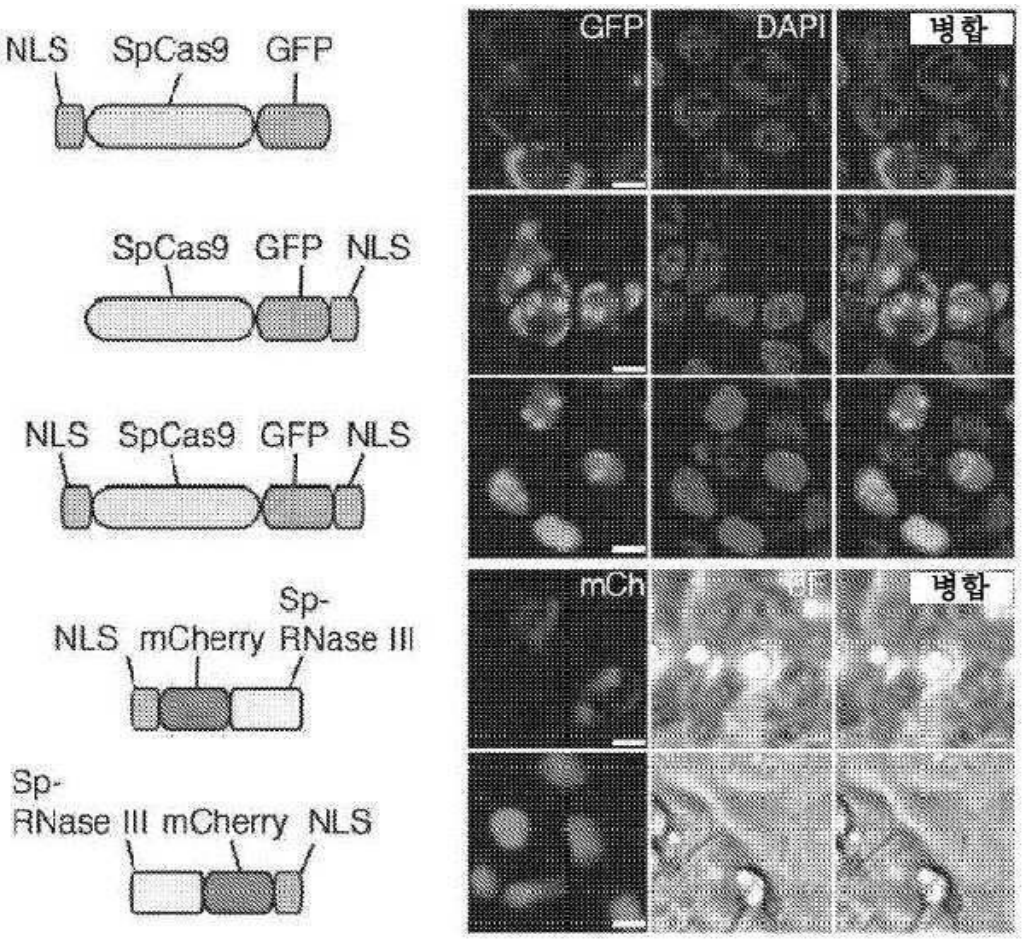


도면2a

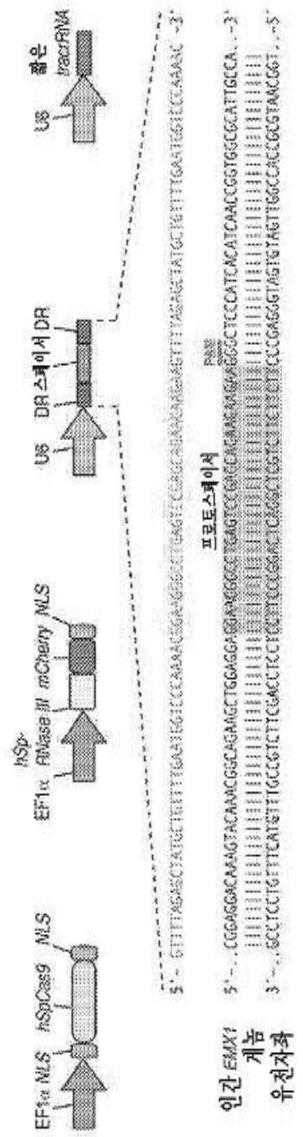
스트렙토코커스 피오게네스 SF370 CRISPR 유전자좌 1



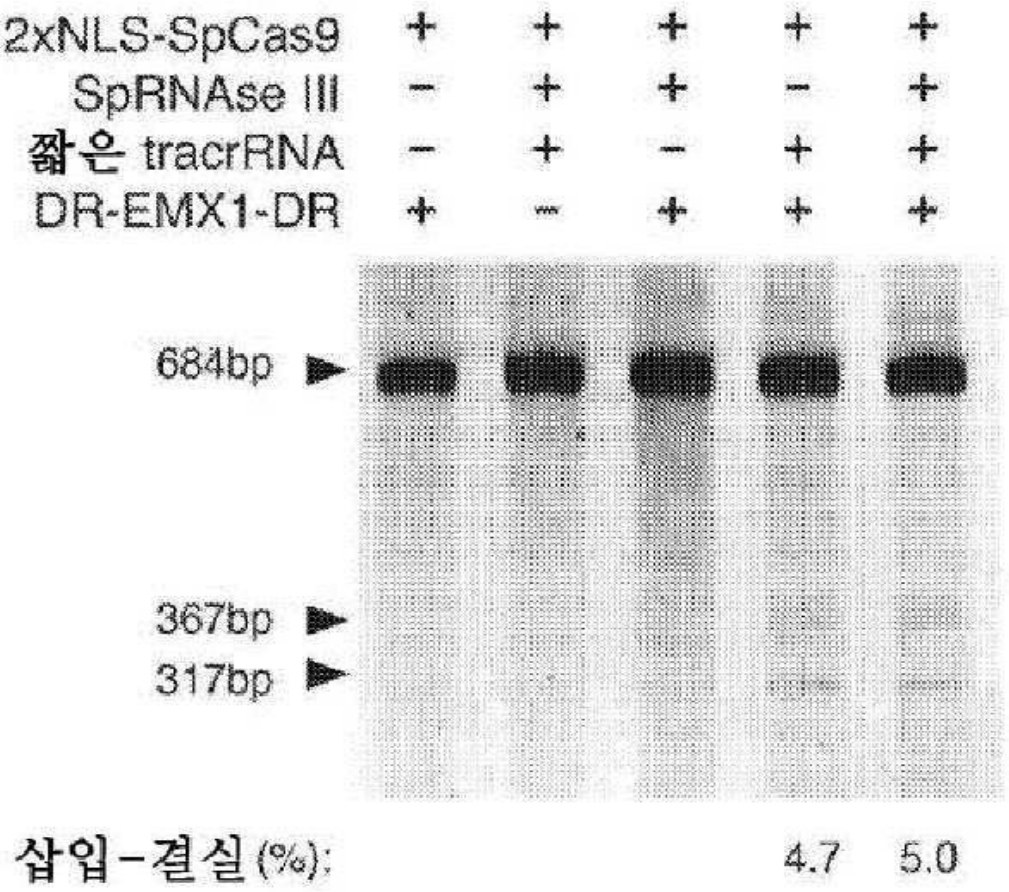
도면2b


NLS SpCas9 GFP
SpCas9 GFP NLS
NLS SpCas9 GFP NLS
Sp-
NLS mCherry RNase III
Sp-
RNase III mCherry NLS
GFP
DAPI
병합
mCh
병합

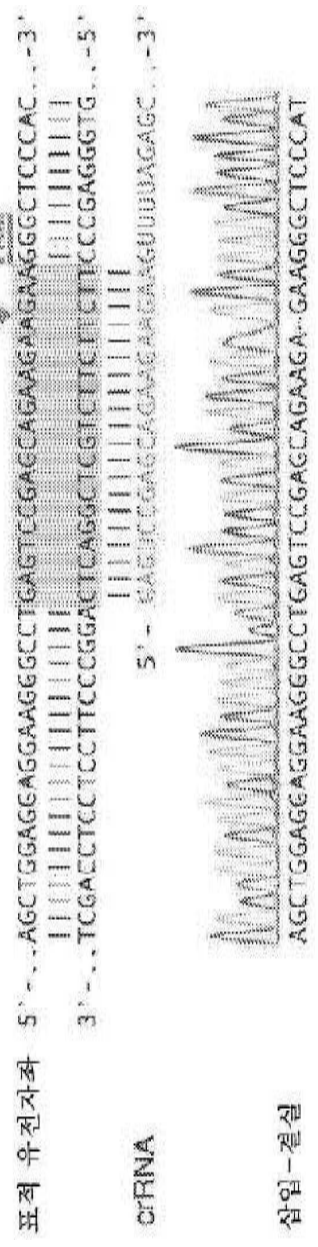
도면2c


EF1α NLS
hSpCas9
NLS
hSp-
EF1α RNase III
mCherry
NLS
U6
DR 스페이서 DR
U6
짧은
tracrRNA
프로토스페이서
PAM
인간 EMX1
게놈
유전자좌

도면2d


2xNLS-SpCas9 + + + + +
SpRNAse III − + + − +
짧은 tracrRNA − + − + +
DR-EMX1-DR + + + + +
684bp
367bp
317bp
삽입-결실(%): 4.7 5.0

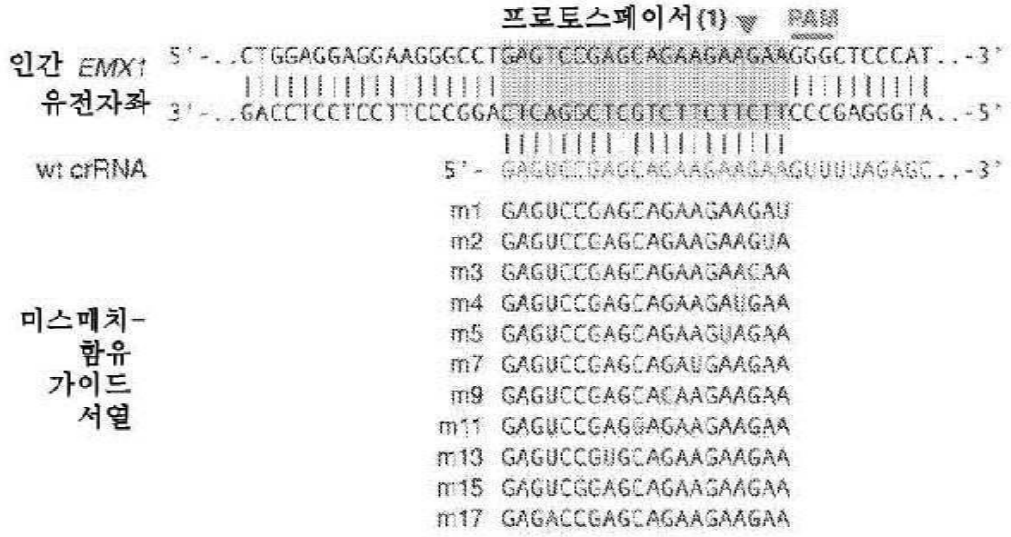
도면2e


PAM
표적 유전자좌
5'-..AGCTGGAGGAGGAAGGGCCTGAGTCCGAGCAGAAGAAGAAGGGCTCCCAC..-3'
3'-..TCGACCTCCTCCTTCCCGGACTCAGGCTCGTCTTCTTCTTCCCGAGGGTG..-5'
crRNA
5'-
삽입-결실
AGCTGGAGGAGGAAGGGCCTGAGTCCGAGCAGAAGA--GAAGGGCTCCCAT

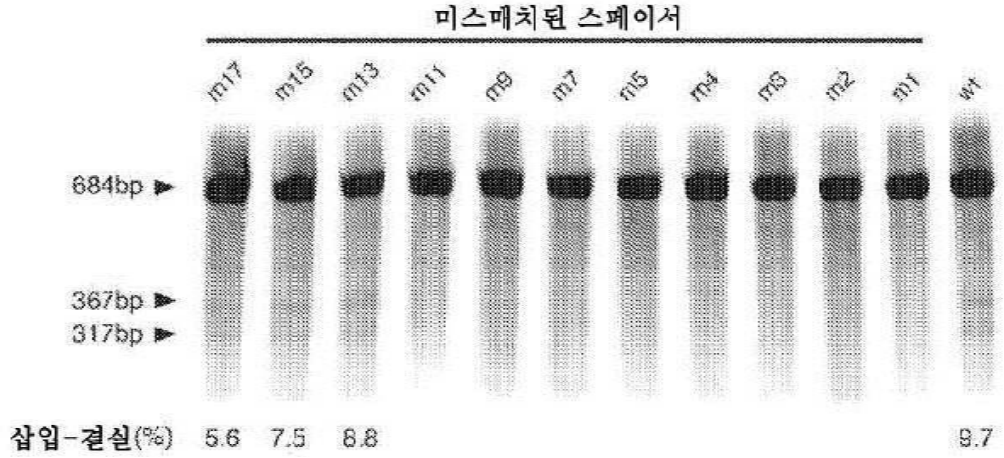
도면2f

인간 EMX1 프로토스페이서 표적 (43개의 시퀀싱된 클론 중 5개에서 돌연변이 = 11.6%)
WT 5'-..CTGGAGGAGAAGAGGCCTGAGTCCGAGCAGAAG AAGAAGGGCTCCCATCACAT..-3'
Δ1 CTGGAGGAGAAGAGGCCTGAGTCCGAGCAGAAG--AGAAGGGCTCCCATCACAT
+1 CTGGAGGAGAAGAGGCCTGAGTCCGAGCAGAAGAAGAAGGGCTCCCATCACAT
Δ3 CTGGAGGAGAAGAGGCCTGAGTCCGAGCAGAAG----AAGGGCTCCCATCACAT
m1, Δ6 CTGGAGGAGAAGAGGCCTGAGCCCGAGCAGAAG--------GGCTCCCATCACAT

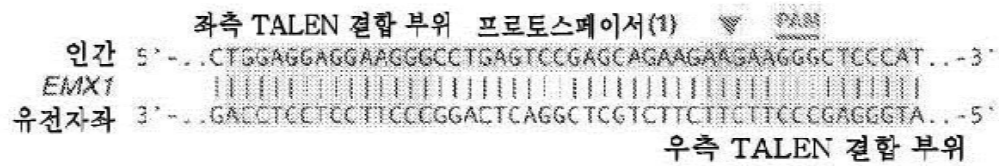
도면3a



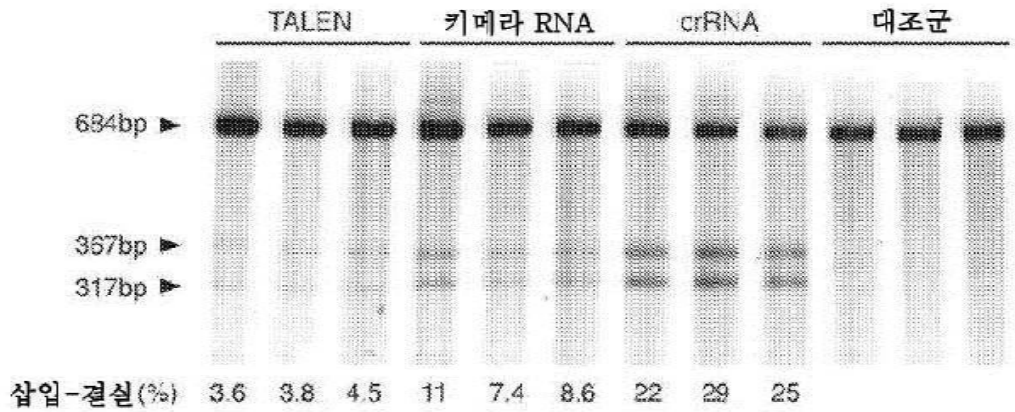
도면3b



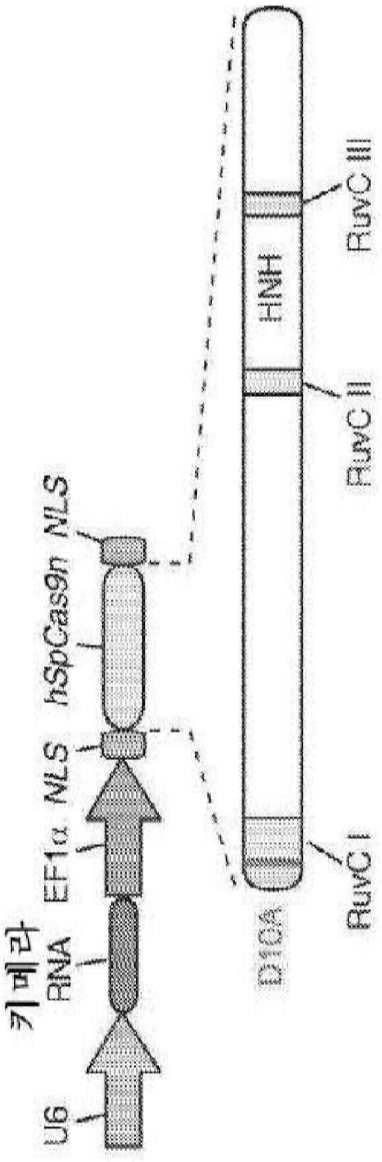
도면3c



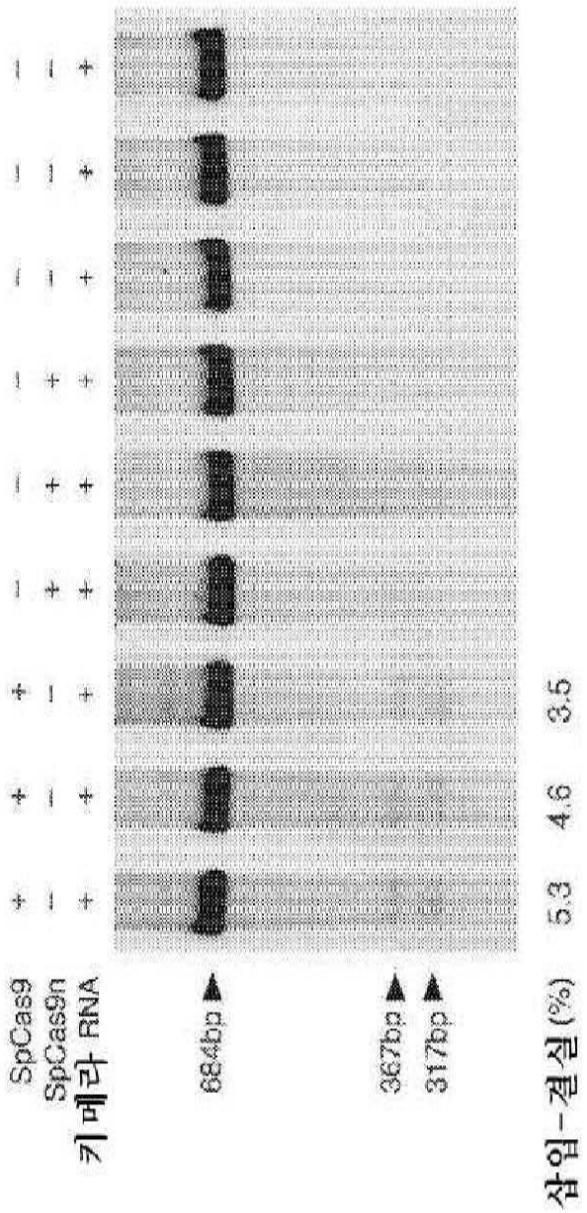
도면3d


TALEN
키메라 RNA
crRNA
대조군
684bp
367bp
317bp
삽입-결실(%)
3.6
3.8
4.5
11
7.4
8.6
22
29
25

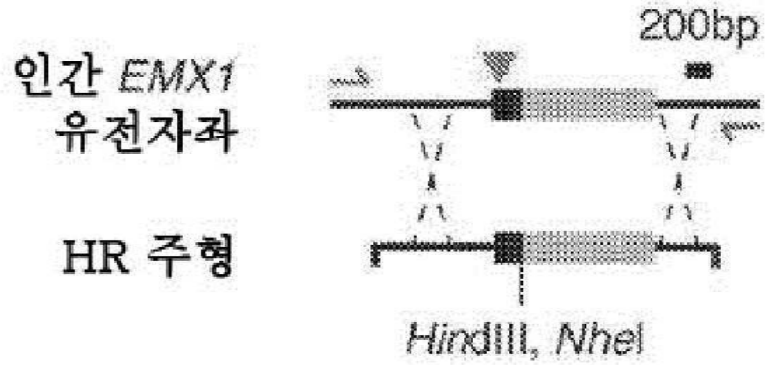
도면4a


키메라 RNA
U6
EF1α
NLS
hSpCas9n
NLS
RuvC I
RuvC II
HNH
RuvC III
D10A

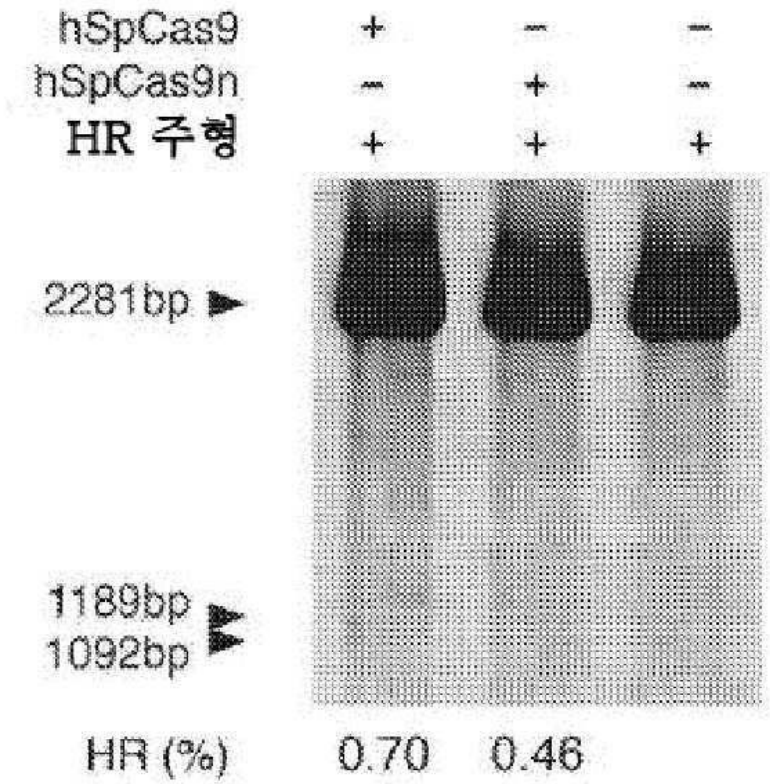
도면4b


SpCas9
SpCas9n
키메라 RNA
684bp
367bp
317bp
삽입-결실 (%)
5.3
4.6
3.5

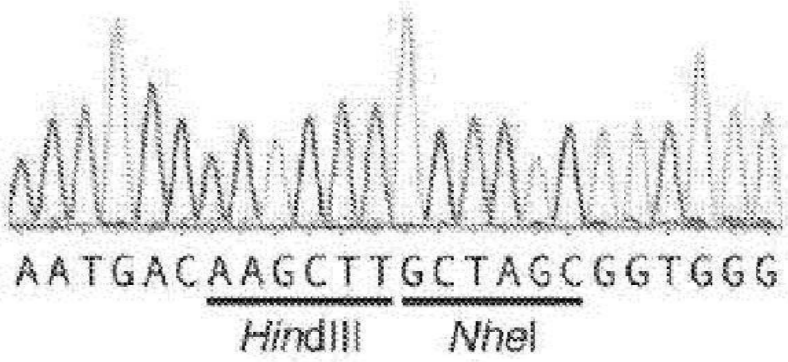
도면4c


인간 EMX1
유전자좌
HR 주형
200bp
HindIII, NheI

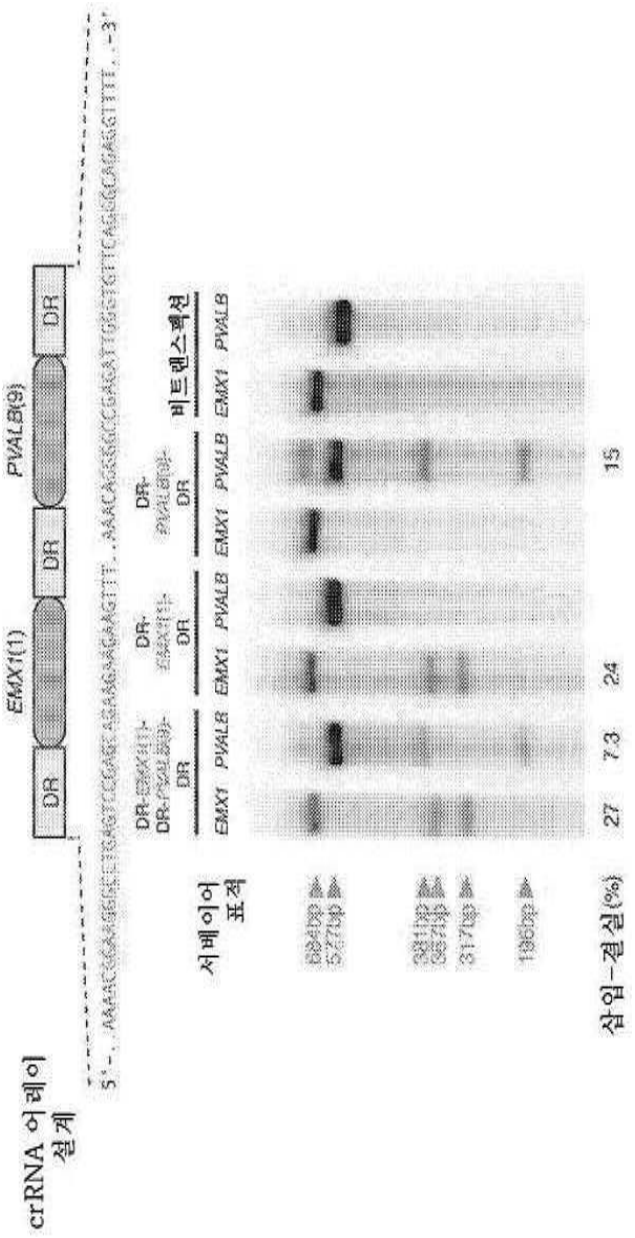
도면4d


hSpCas9 + - -
hSpCas9n - + -
HR 주형 + + +
2281bp
1189bp
1092bp
HR (%) 0.70 0.46

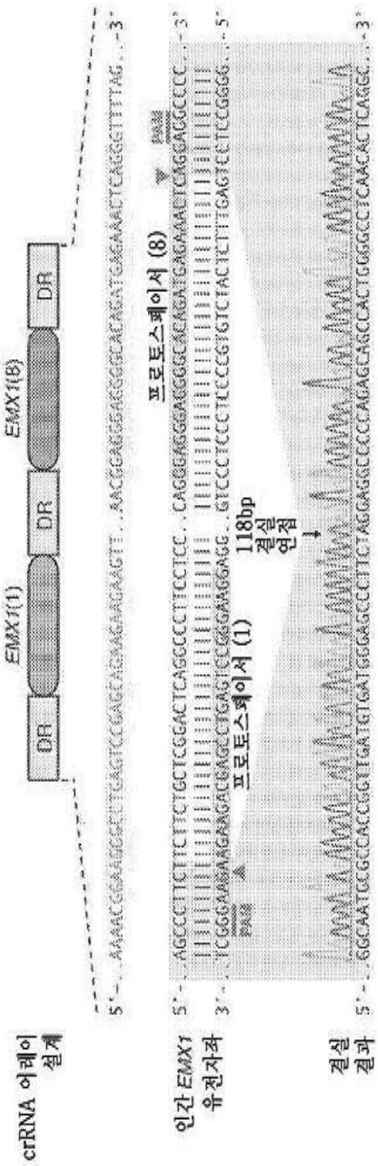
도면4e


AATGACAAGCTTGCTAGCGGTGGG
HindIII
NheI

도면4f


crRNA 어레이 설계
EMX1(1)
PVALB(9)
DR
DR
DR
DR
5'-, AAAACGAAGGGGTCTGAGTCGAGAAGAAGAAGTTT, AAACAGGGGCCGAGATTTGGGTGTTCAGGGCAGAGGTTTT, -3'
서베이어 표적
DR-EMX1(1)-DR-PVALB(9)-DR
DR-EMX1(1)-DR
DR-PVALB(9)-DR
비트렌스퀙션
EMX1
PVALB
EMX1
PVALB
EMX1
PVALB
EMX1
PVALB
884bp
577bp
381bp
357bp
317bp
186bp
삽입-결실(%)
27
7.3
24
15

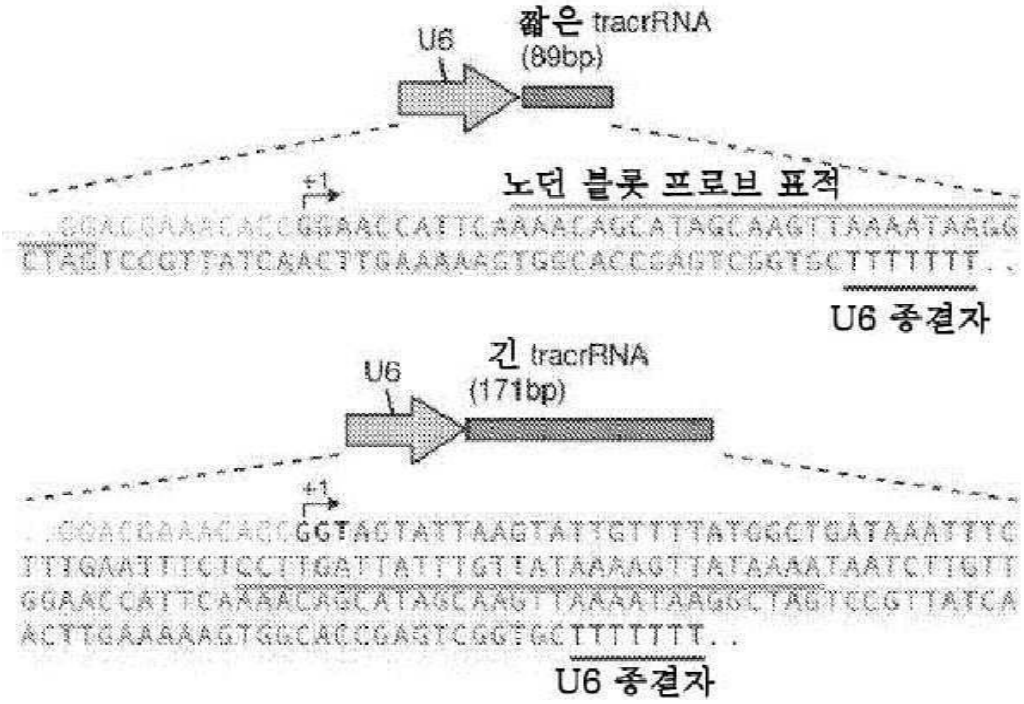
도면4g


crRNA 어레이 설계
EMX1(1)
EMX1(8)
DR
DR
DR
인간 EMX1 유전자좌
프로토스페이서 (1)
프로토스페이서 (8)
PAM
118bp
절실 연접
결실 결과

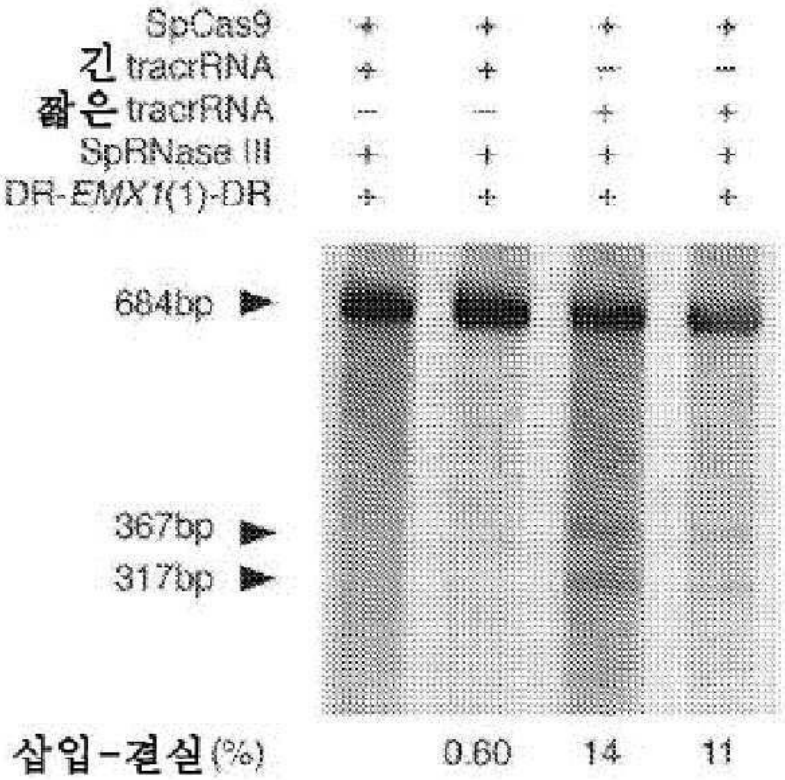
도면5

Cas9	표적 종	유전자	프로토스페이서 ID	프로토스페이서 서열 (5'으로부터 3')	PAM	가닥	시험된 세포주	삽입-결실% (혜비-crRNA + tracrRNA)(키메라 RNA)	삽입-결실%
스트렙토코커스 피오게네스 SF370 II형 CRISPR	호모 사피엔스	EMX1	1	GUAAGGCCUGAGUCGAGCAGAAGAAGAA	GGG	+	293FT	29±1.8	6.7±0.82
		EMX1	2	CATTGAAGUGAUAUCAUGUCUCCCAT	TGG	-	293FT	2.1±0.31	N.D.
		EMX1	3	GGACAUCGAUGUCACCUCAATGACUAGG	TGG	+	293FT	14±1.1	N.D.
		EMX1	4	CATCGATGTCCTCCCCATTGGCCTGCTTCG	TGG	-	293FT	11±1.7	N.D.
		EMX1	5	TTCGTGGCAATGCGCCACCGGTGATGTGA	TGG	-	293FT	4.3±0.40	2.1±0.51
		EMX1	6	TCGTGCCAATGCCCACCGGTGATGTGAT	GGG	-	293FT	4.0±0.86	0.41±0.25
		EMX1	7	TCCAGCTTCTGCCGTTTGTACTTTGTCCTC	CGG	-	293FT	1.5±0.12	N.D.
		EMX1	8	GCAGGCAGGGCACAGATGAGAACTCAGG	AGG	-	293FT	7.8±0.83	2.3±1.2
	호모 사피엔스	PVALB	9	AGGGGCCAGATTGGGTGTTCAGGCAGAG	AGG	+	293FT	21±2.6	6.5±0.32
		PVALB	10	ATGCAGGAGGGTGGCAGAGGGGCCGAGAT	TGG	+	293FT	N.D.	N.D.
		PVALB	11	GGTGGCAGAGGGCCAGATTGGGTGTTC	AGG	+	293FT	N.D.	N.D.
	무스 무스컬러스	Tb	12	CAAGCACTGAGTGCATTAGCTAAATGCAT	AGG	-	NeuroDA	27±4.3	41±22
		Tb	13	AATGCATAGGTACCACCCAGAGTGCCAG	GGG	-	NeuroDA	4.8±1.2	N.D.
		Tb	14	ACACACATGGGAAAGCCCTCTGGGCCAGGAA	AGG	+	NeuroDA	11.3±1.3	N.D.
스트렙토코커스 써모필러스 LMD-9 CRISPR1	호모 사피엔스	EMX1	15	GGAGGAGGTAGTATACAGAAACACAGCAA	GTAGAAT	-	293FT	14±0.88	N.T.
		EMX1	16	AGAATGTAGACGAGTCTACAGAAGTCAGCA	CTAGAAA	-	293FT	7.6±0.77	N.T.

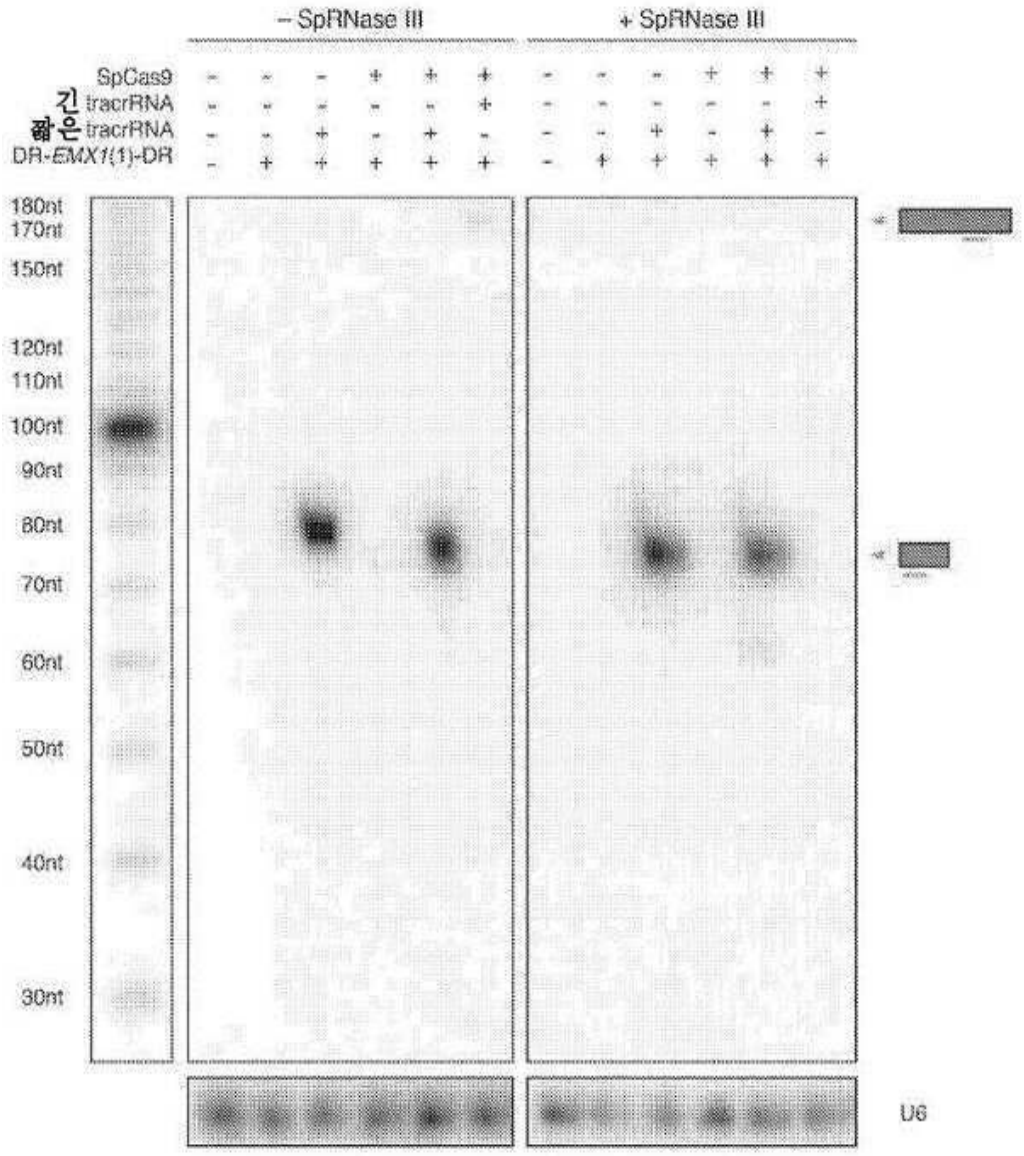
도면6a


U6
짧은 tracrRNA
(89bp)
+1
노던 블롯 프로브 표적
U6 종결자
U6
긴 tracrRNA
(171bp)
+1
U6 종결자

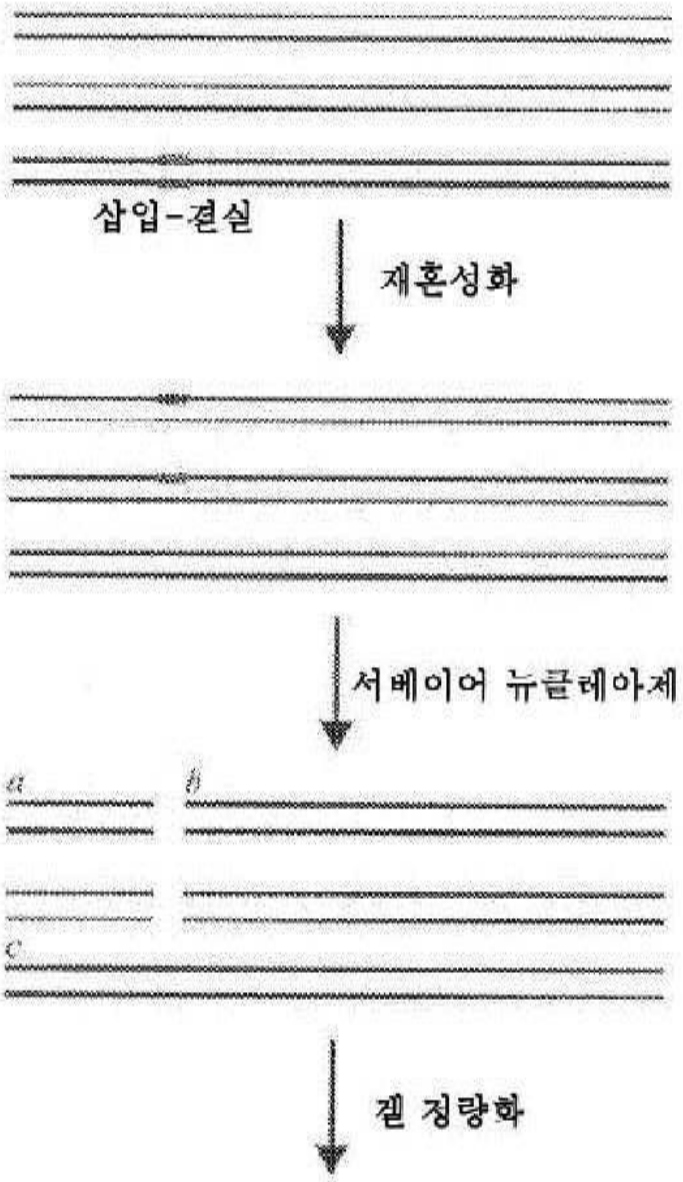
도면6b


SpCas9
긴 tracrRNA
짧은 tracrRNA
SpRNase III
DR-EMX1(1)-DR
684bp
367bp
317bp
삽입-결실(%)
0.60
14
11

도면6c

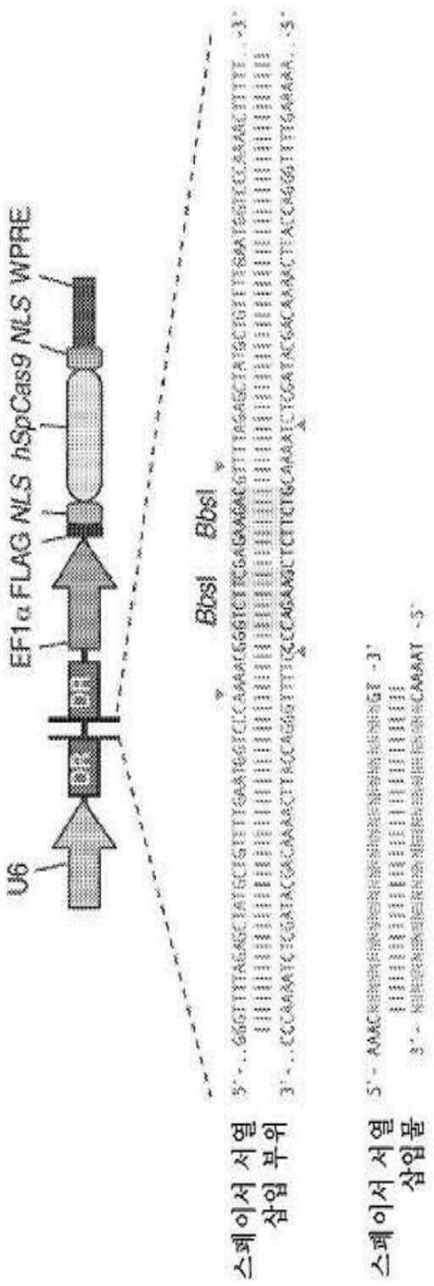

- SpRNase III
+ SpRNase III
SpCas9
긴 tracrRNA
짧은 tracrRNA
DR-EMX1(1)-DR
180nt
170nt
150nt
120nt
110nt
100nt
90nt
80nt
70nt
60nt
50nt
40nt
30nt
U6

도면7

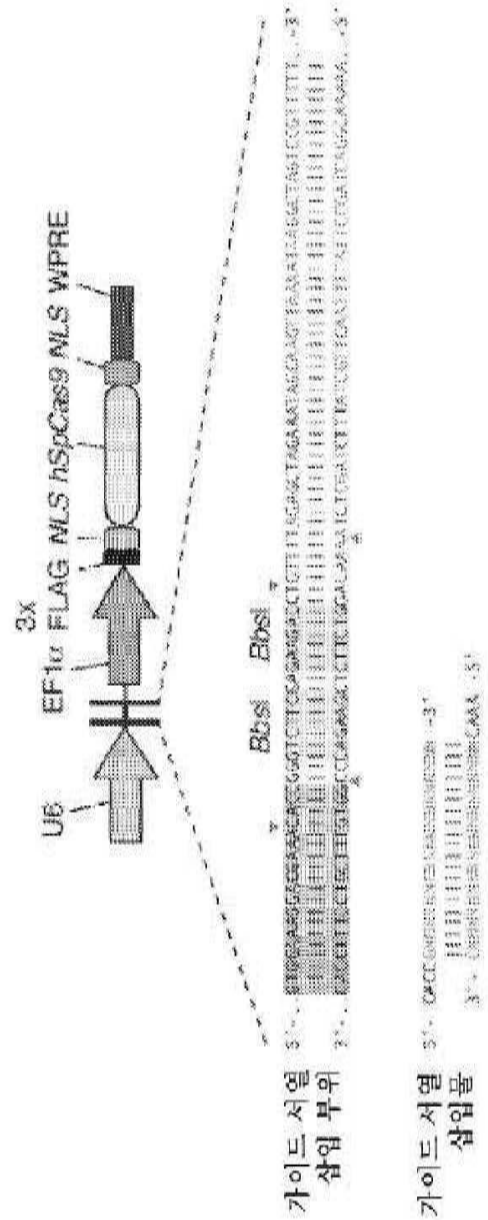


$$\text{삽입-결실 \%} = \left(1-\sqrt{1-\frac{(a+b)}{(a+b+c)}}\right)*100$$

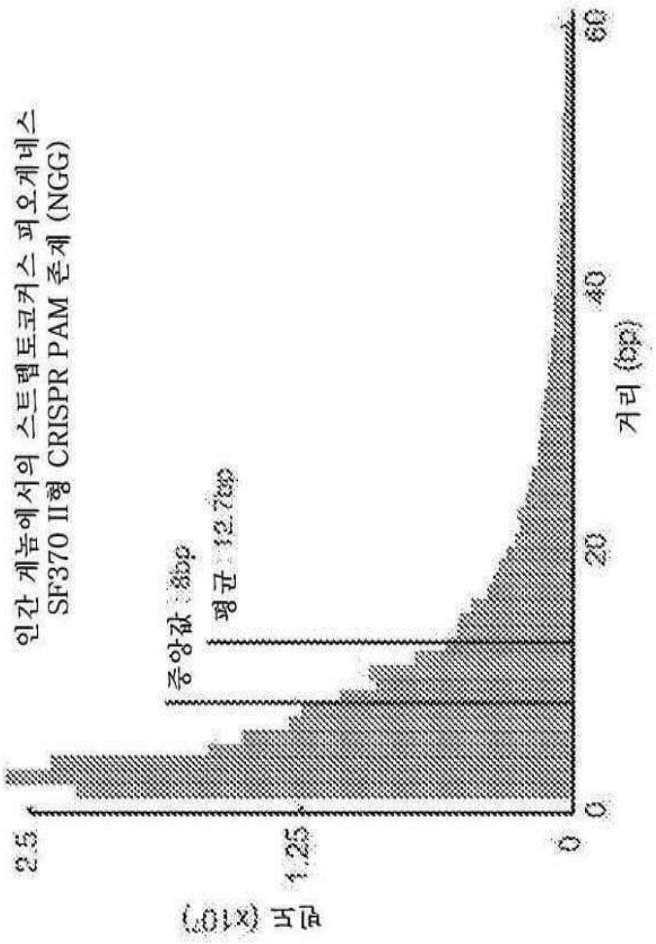
도면8a


U6
3x
EF1α FLAG NLS hSpCas9 NLS WPRE
BfuAI
BfuAI
BbsI
BbsI
스페이서 서열
삽입 부위
스페이서 서열
삽입물

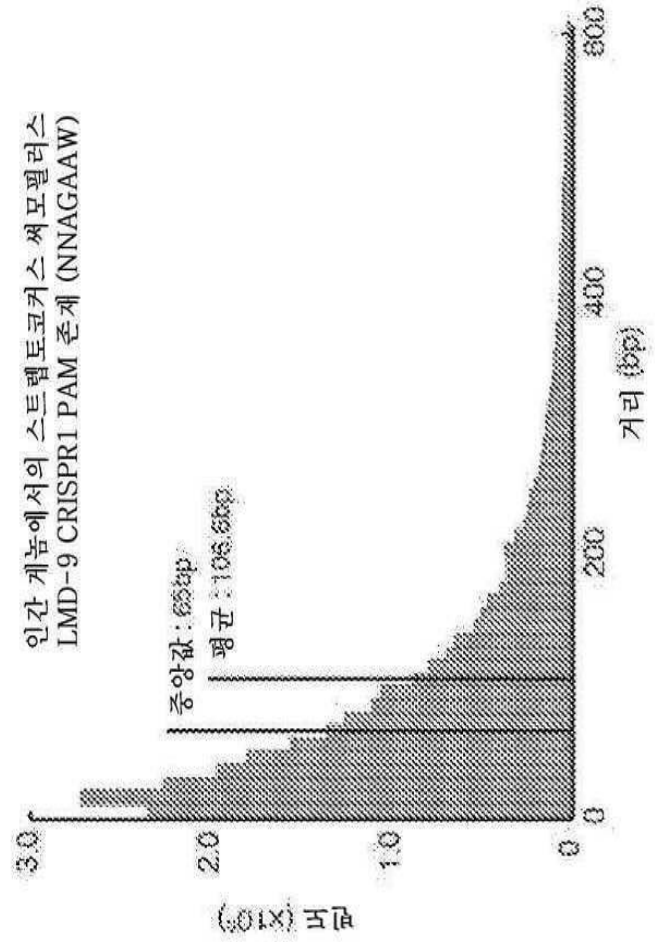
도면8b


U6
3x
EF1α FLAG MLS hSpCas9 MLS WPRE
BbsI
BbsI
가이드 서열
삽입 부위
가이드 서열
삽입물

도면9a


인간 게놈에서의 스트렙토코커스 피오게네스
SF370 II형 CRISPR PAM 존재 (NGG)
중앙값 : 8bp
평균 : 12.7bp
빈도 (x10³)
2.5
1.25
0
0
20
40
60
거리 (bp)

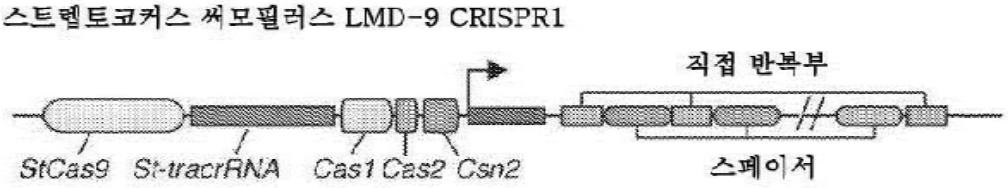
도면9b


인간 게놈에서의 스트렙토코커스 쎄모필러스 LMD-9 CRISPR1 PAM 존재 (NNAGAAW)
중앙값 : 65bp
평균 : 106.6bp
빈도 (x10⁵)
3.0
2.0
1.0
0
0
200
400
600
거리 (bp)

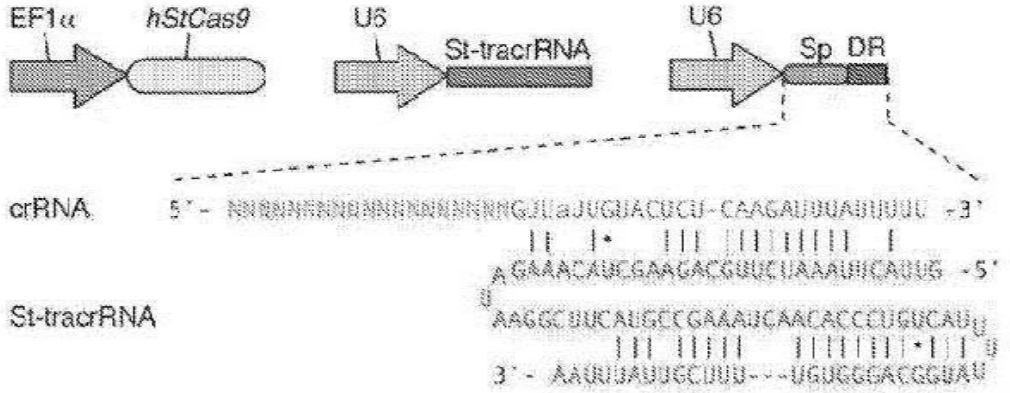
도면9c

	NGG		NNAGAAW	
염색체	중앙값	평균	중앙값	평균
1	7	12.8	67	115.8
2	8	12.7	64	100.8
3	8	13.0	63	98.5
4	9	14.0	61	94.5
5	8	13.1	63	97.9
6	8	13.1	63	98.5
7	8	12.4	64	102.9
8	8	12.8	64	100.9
9	7	13.9	65	120.5
10	7	12.1	66	107.0
11	7	12.0	65	105.6
12	8	12.4	65	103.5
13	8	13.6	62	94.6
14	8	12.0	65	101.5
15	7	11.5	68	107.7
16	7	11.7	74	136.8
17	6	10.3	76	127.9
18	8	13.4	63	101.8
19	6	9.4	82	145.4
20	7	11.1	72	121.6
21	7	13.4	64	111.4
22	6	9.2	85	140.3
X	8	13.2	63	99.0
Y	8	29.2	62	223.7

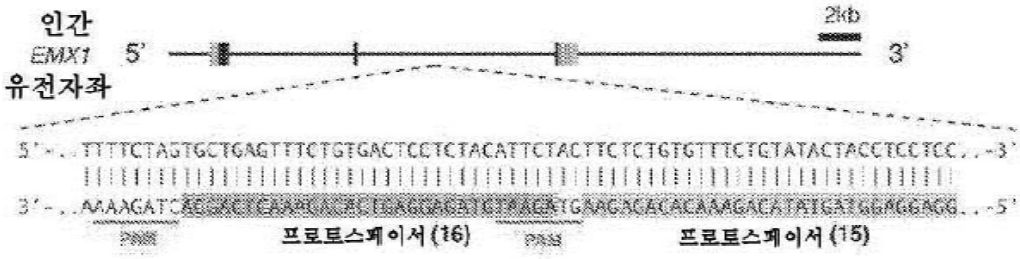
도면10a



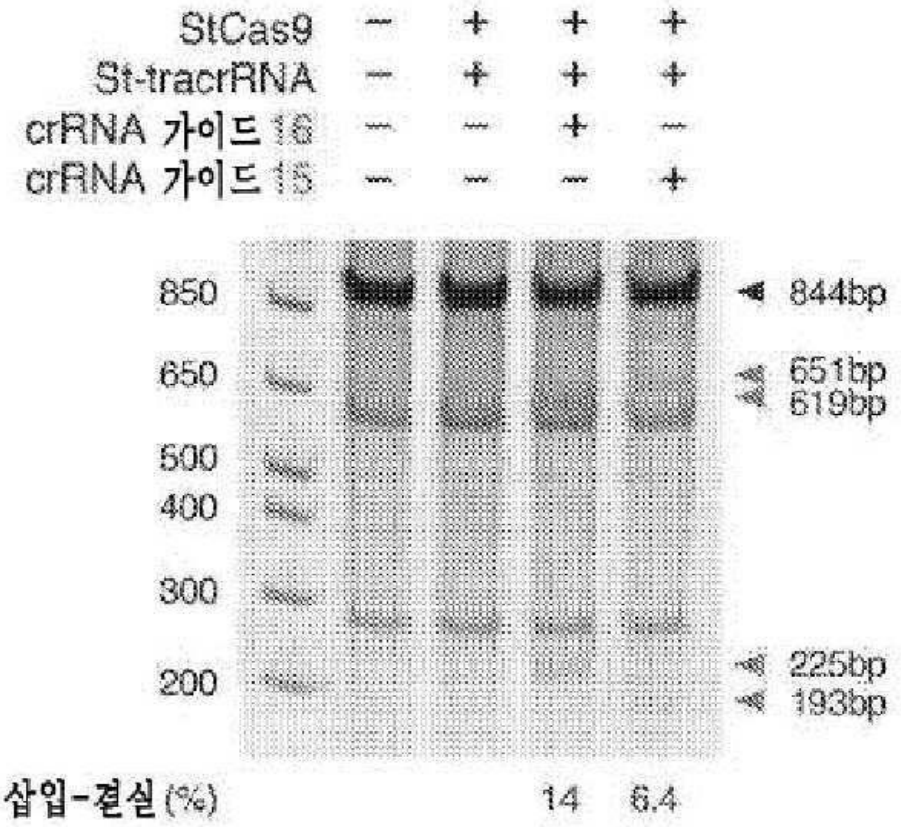
도면10b



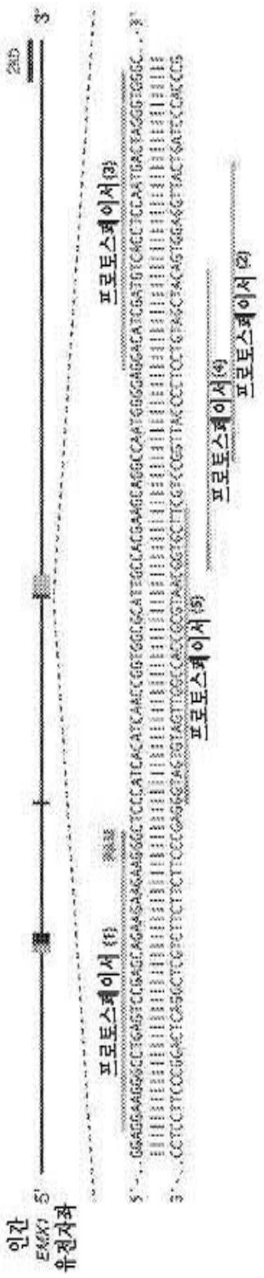
도면10c



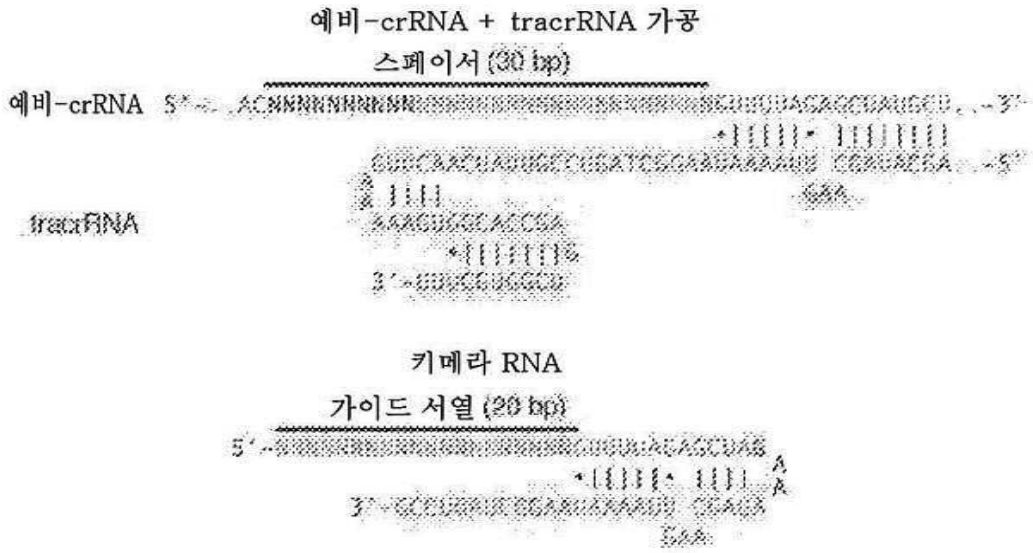
도면10d



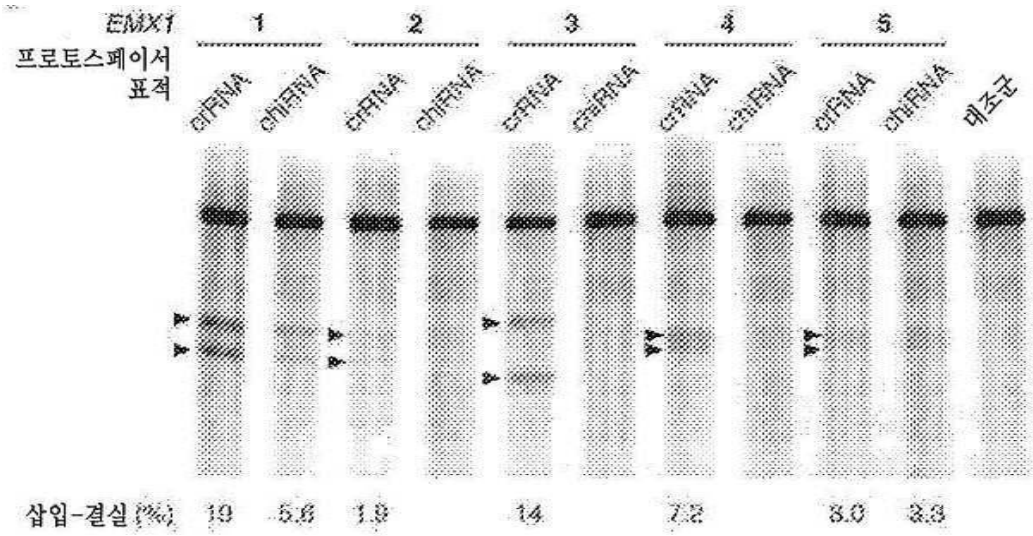
도면11a


인간
EMX1
유전자좌
5'
3'
프로토스페이서 (1)
프로토스페이서 (3)
프로토스페이서 (5)
프로토스페이서 (4)
프로토스페이서 (2)

도면11b



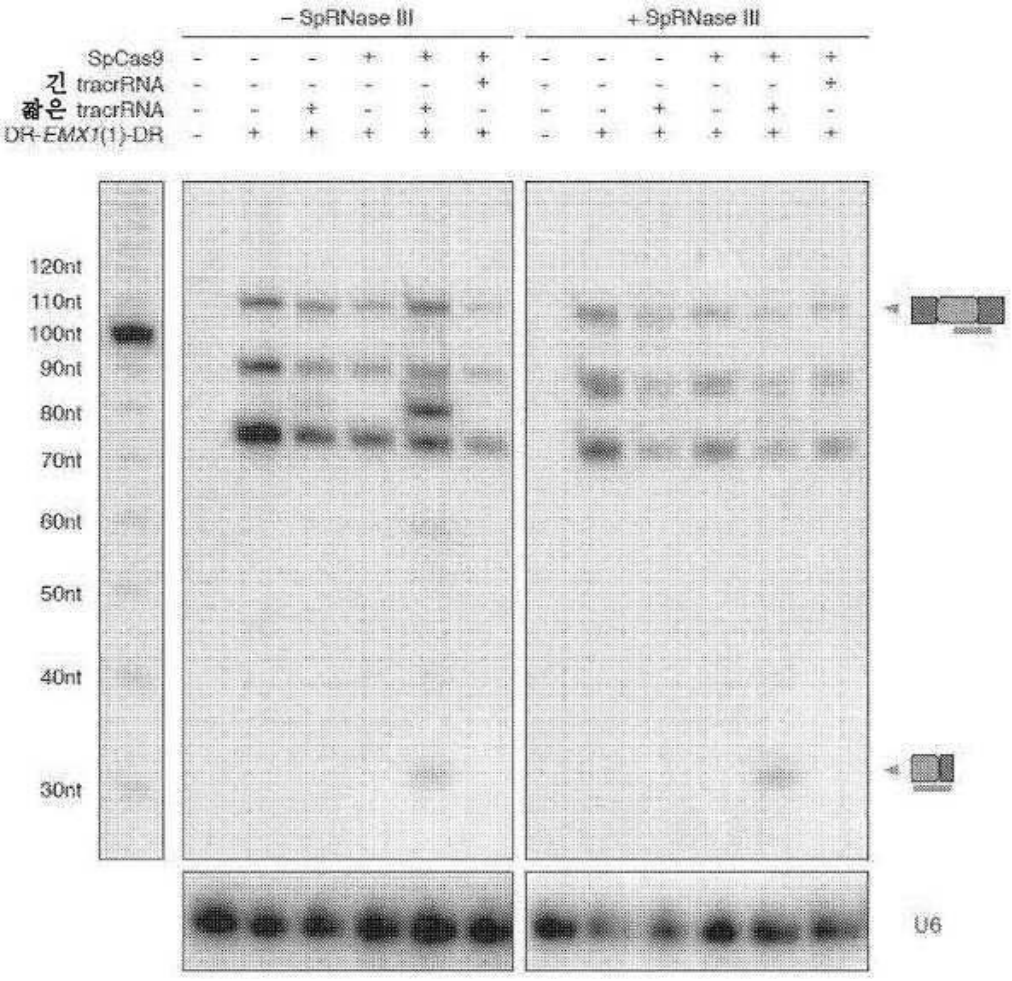
도면11c



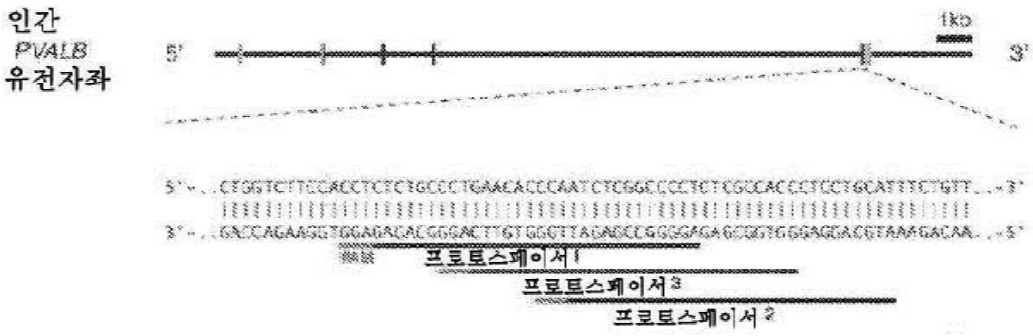
도면12a



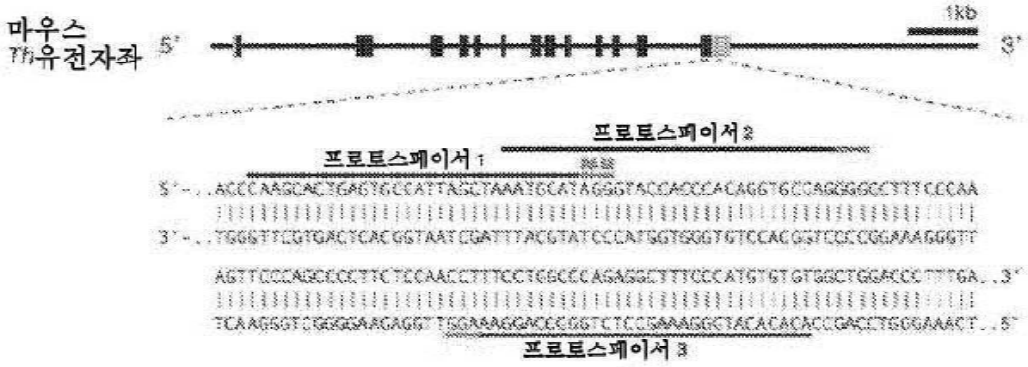
도면12b


– SpRNase III
+ SpRNase III
SpCas9
긴 tracrRNA
짧은 tracrRNA
DR-EMX1(1)-DR
120nt
110nt
100nt
90nt
80nt
70nt
60nt
50nt
40nt
30nt
U6

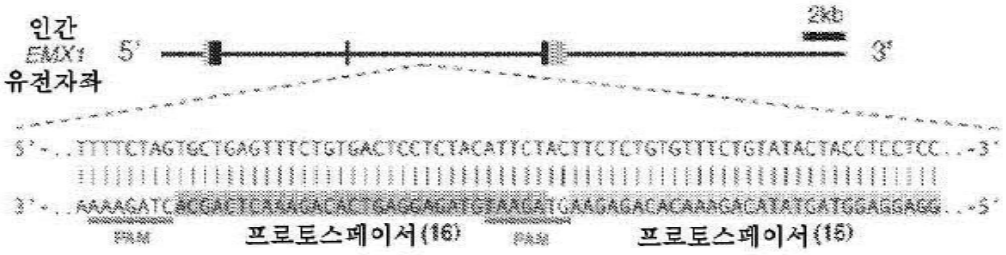
도면13a


인간
PVALB
유전자좌
5'
3'
1kb
프로토스페이서 1
프로토스페이서 3
프로토스페이서 2

도면13b


마우스
Tfh유전자좌
5'
1kb
3'
프로토스페이서 2
프로토스페이서 1
PAM
5'-..ACCCAAGCACTGAGTGCCATTAGCTAAATGCATAGGGTACCACCCACAGGTGCCAGGGGCCTTTCCCAA
3'-..TGGGTTCGTGACTCACGGTAATCGATTTACGTATCCCATGGTGGGTGTCCACGGTCCCCGGAAAGGGTT
AGTTCCCAGCCCCTTCTCCAACCTTTCCTCGCCCAGAGGCTTTCCCATGTGTGTGGCTGGACCCTTTGA..3'
TCAAGGGTCGGGGAAGAGGTTCGAAAGGACCCGGTCTCCGAAAGGCTACACACACCGACCTGGGAAACT..5'
프로토스페이서 3

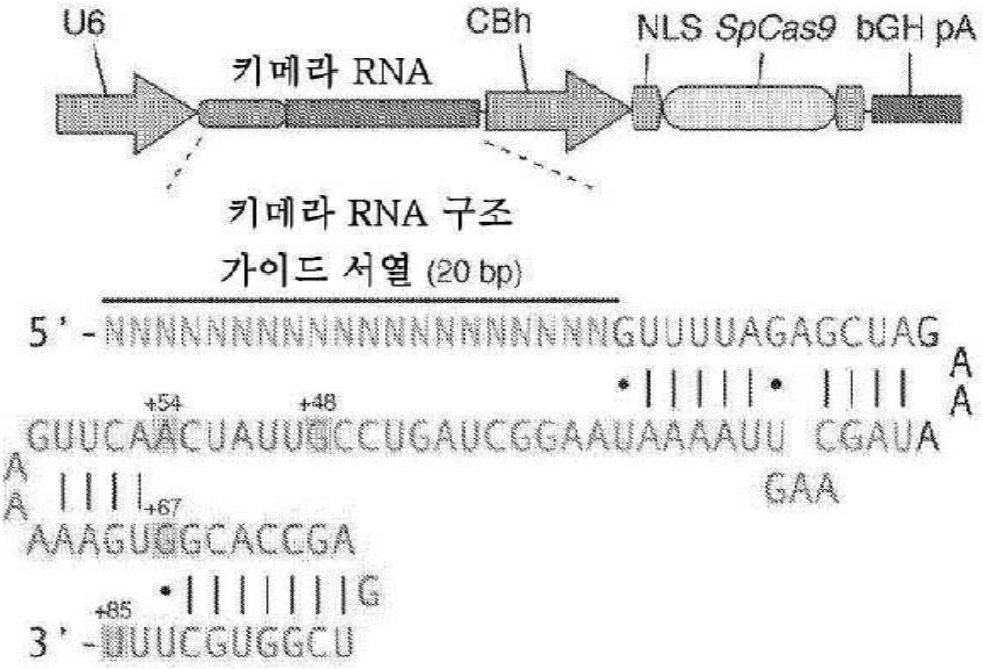
도면14


인간
EMX1
유전자좌
5'
2kb
3'
5'-..TTTTCTAGTGCTGAGTTTCTGTGACTCCTCTACATTCTACTTCTCTGTGTTTCTGTATACTACCTCCTCC..-3'
3'-..AAAAGATCACGACTCAAAGACACTGACGACATGTAACATGAAGAGACACACAAAGACATATGATGGAGGAGG..-5'
PAM
프로토스페이서(16)
PAM
프로토스페이서(15)

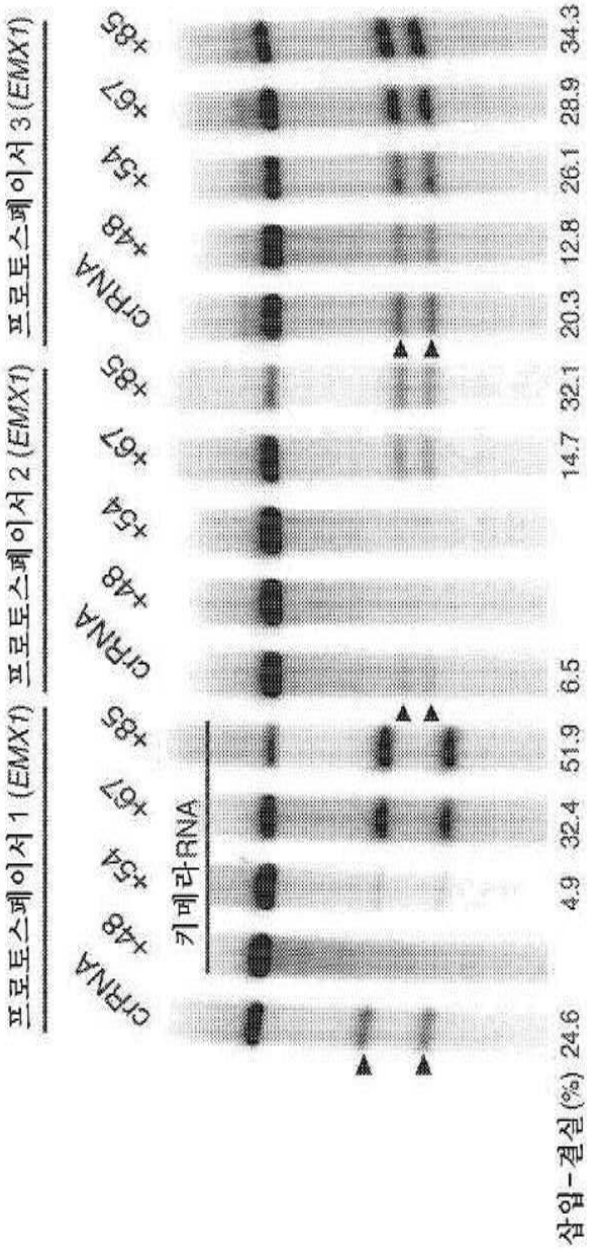
도면15

프라이머 명칭	검정	게놈 표적	프라이머 서열
Sp-EMX1-F	서베이어 검정, 시퀀싱	*EMX1*	AAAACCACCCTTCTCTCTGGC
Sp-EMX1-R	서베이어 검정, 시퀀싱	*EMX1*	GGAGATTGGAGACACGGAGAG
Sp-PVALB-F	서베이어 검정, 시퀀싱	*PVALB*	CTGGAAAGCCAATGCCTGAC
Sp-PVALB-R	서베이어 검정, 시퀀싱	*PVALB*	GGCAGCAAACTCCTTGTCCT
Sp-Th-F	서베이어 검정, 시퀀싱	*Th*	GTGCTTTGCAGAGGCCTACC
Sp-Th-R	서베이어 검정, 시퀀싱	*Th*	CCTGGAGCGCATGCAGTAGT
St-EMX1-F	서베이어 검정, 시퀀싱	*EMX1*	ACCTTCTGTGTTTCCACCATTC
St-EMX1-R	서베이어 검정, 시퀀싱	*EMX1*	TTGGGGAGTGCACAGACTTC
Sp-EMX1-RFLP-F	RFLP, 시퀀싱	*EMX1*	GGCTCCCTGGGTTCAAAGTA
Sp-EMX1-RFLP-R	RFLP, 시퀀싱	*EMX1*	AGAGGGGTCTGGATGTCGTAA
Pb_EMX1_sp1	노던 블롯 프로브	해당 없음	TAGCTCTAAAACTTCTTCTTCTGCTCGGAC
Pb_tracrRNA	노던 블롯 프로브	해당 없음	CTAGCCTTATTTTAACTTGCTATGCTGTTT

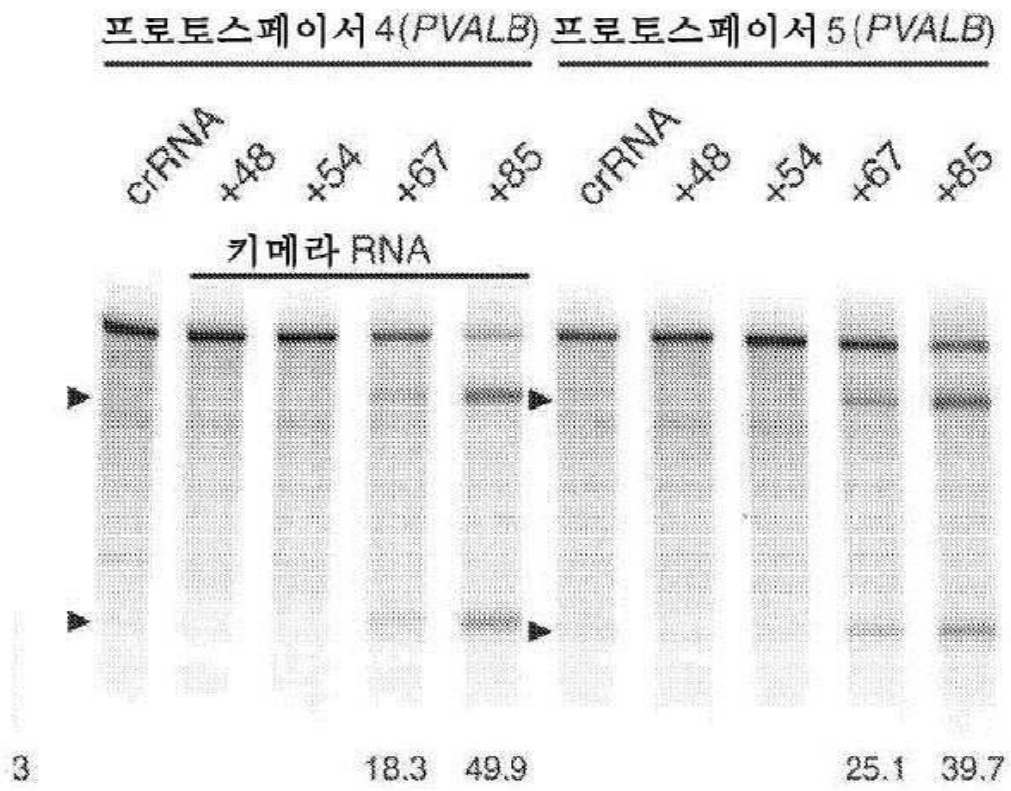
도면16a


U6
CBh
NLS SpCas9 bGH pA
키메라 RNA
키메라 RNA 구조
가이드 서열 (20 bp)
5' - NNNNNNNNNNNNNNNNNNNNGUUUUAGAGCUAG
A
A
+54
+48
GUUCAACUAUUGCCUGAUCGGAAUAAAAUU CGAUA
A
A
GAA
+67
AAAGUGGCACCGA
+85
3' - UUUCGUGGCU

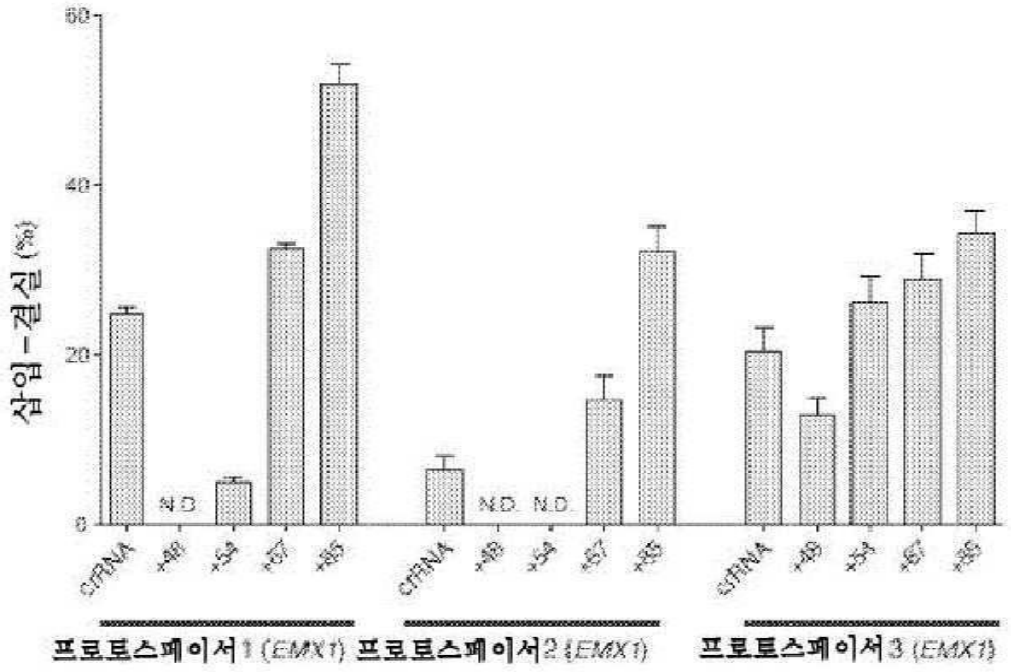
도면16b


프로토스페이서 1 (EMX1)
프로토스페이서 2 (EMX1)
프로토스페이서 3 (EMX1)
crRNA
+48
+54
+67
+85
키메라 RNA
삽입-결실 (%)
24.6
4.9
32.4
51.9
6.5
14.7
32.1
20.3
12.8
26.1
28.9
34.3

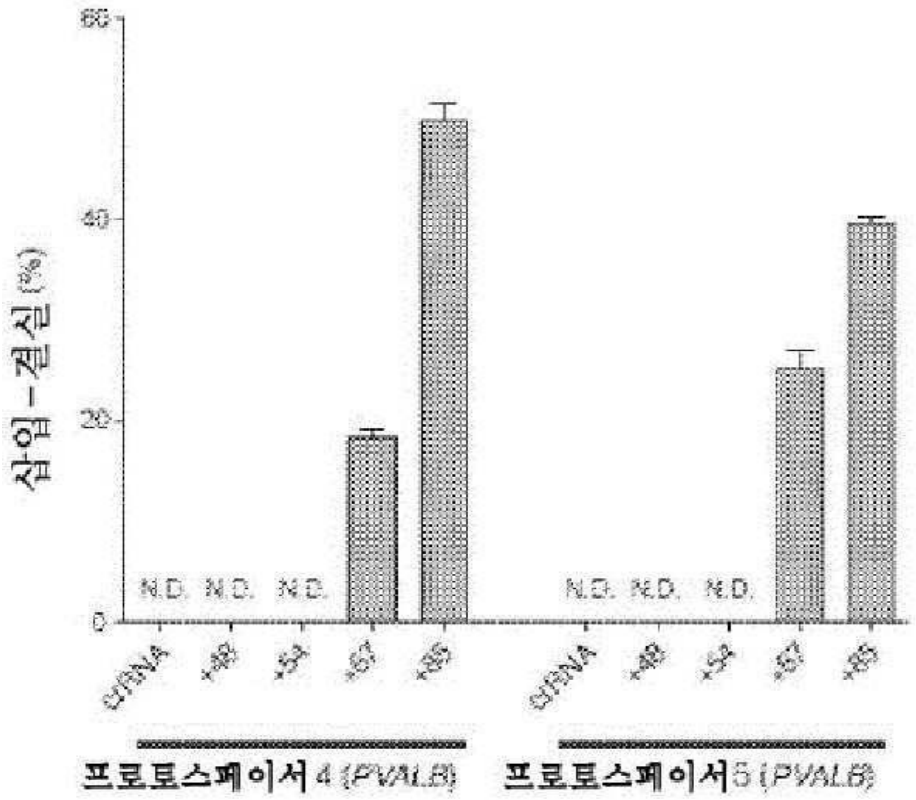
도면16c


프로토스페이서 4(PVALB)
프로토스페이서 5(PVALB)
crRNA
+48
+54
+67
+85
crRNA
+48
+54
+67
+85
키메라 RNA
3
18.3
49.9
25.1
39.7

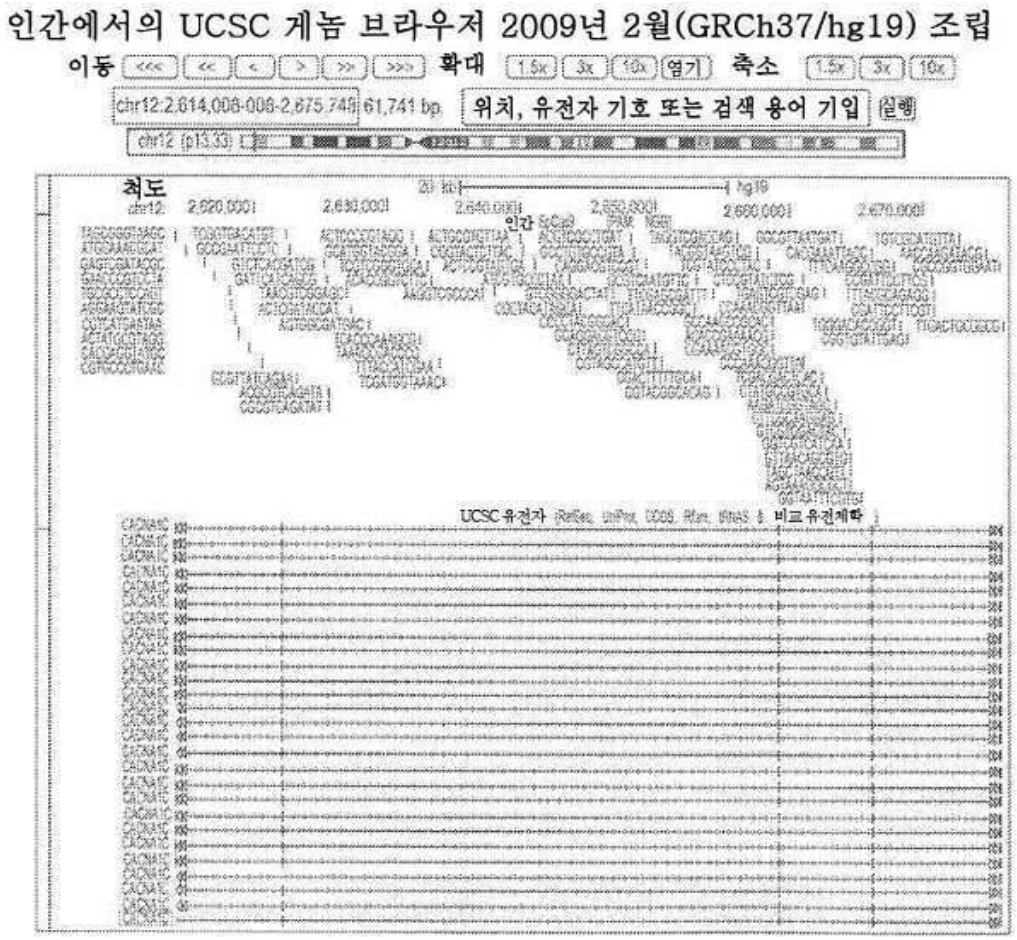
도면17a


60
40
20
0
삽입-결실 (%)
N.D.
N.D.
N.D.
crRNA
+48
+54
+67
+85
프로토스페이서 1 (EMX1)
프로토스페이서 2 (EMX1)
프로토스페이서 3 (EMX1)

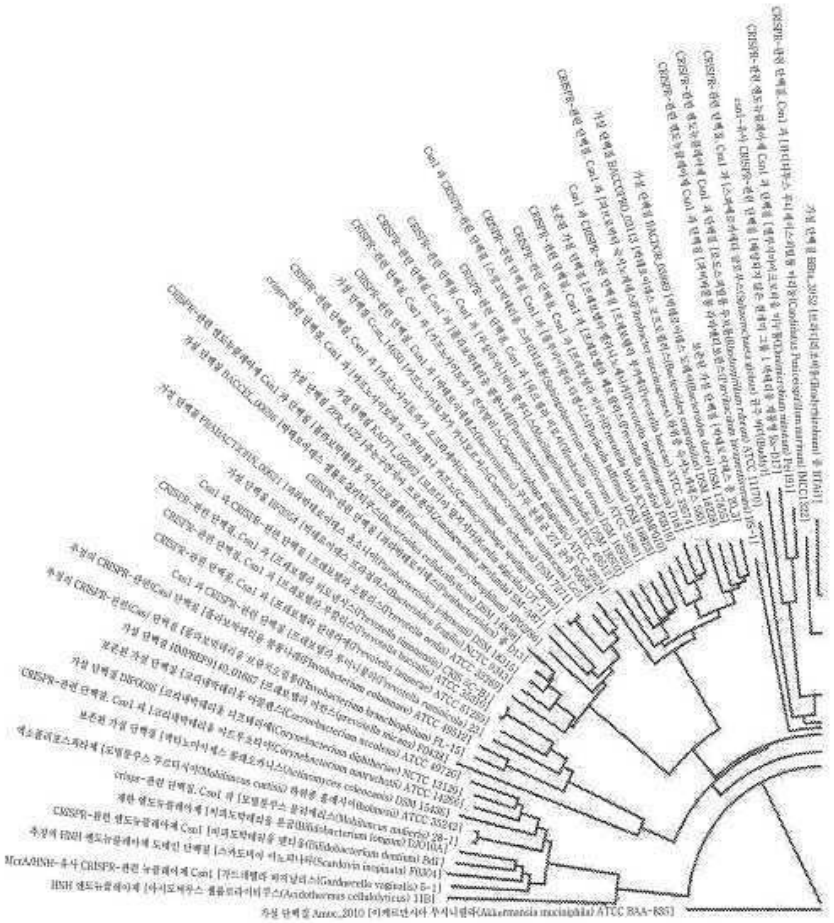
도면17b



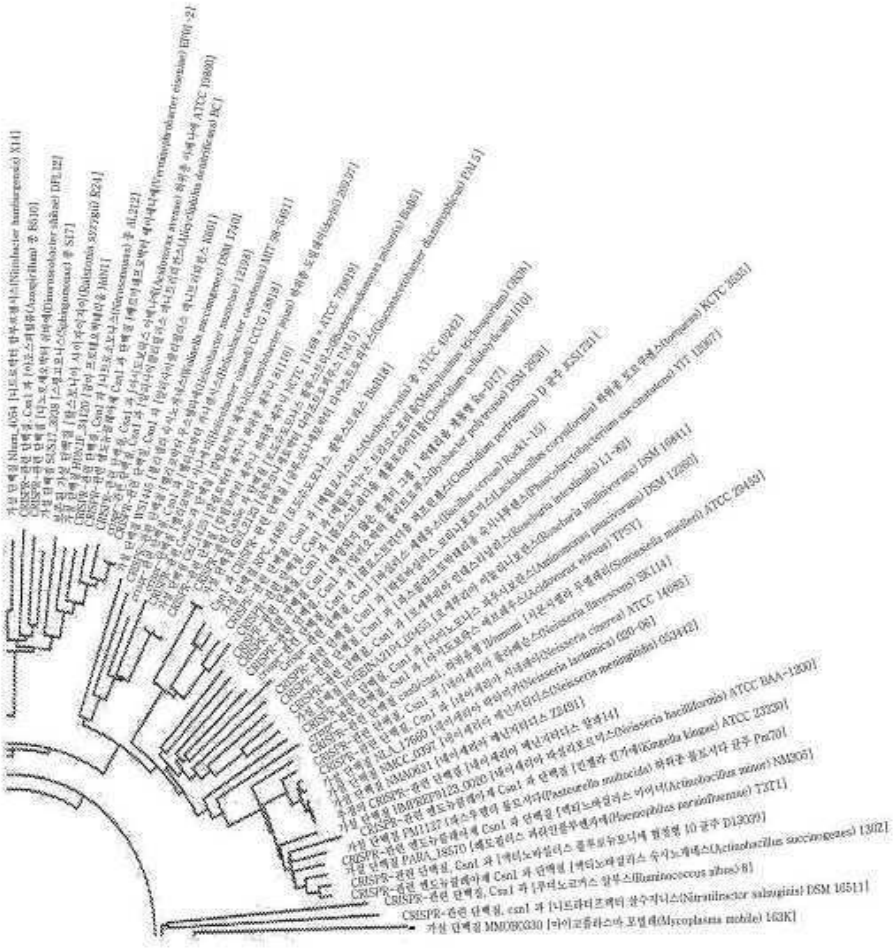
도면18


인간에서의 UCSC 게놈 브라우저 2009년 2월(GRCh37/hg19) 조립
이동
확대
축소
위치, 유전자 기호 또는 검색 용어 기입
실행
척도
인간
UCSC 유전자
비교유전체학

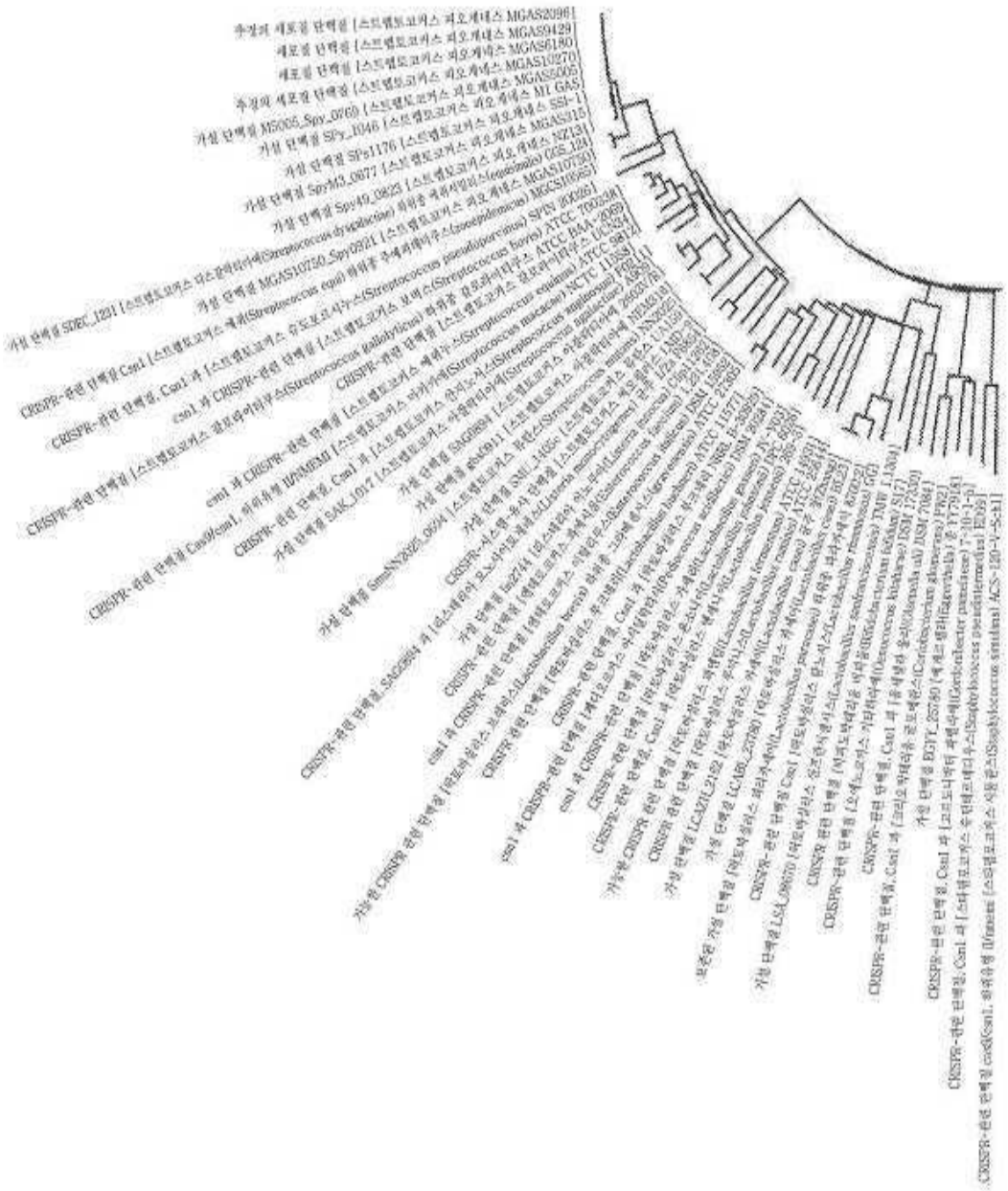
도면19a



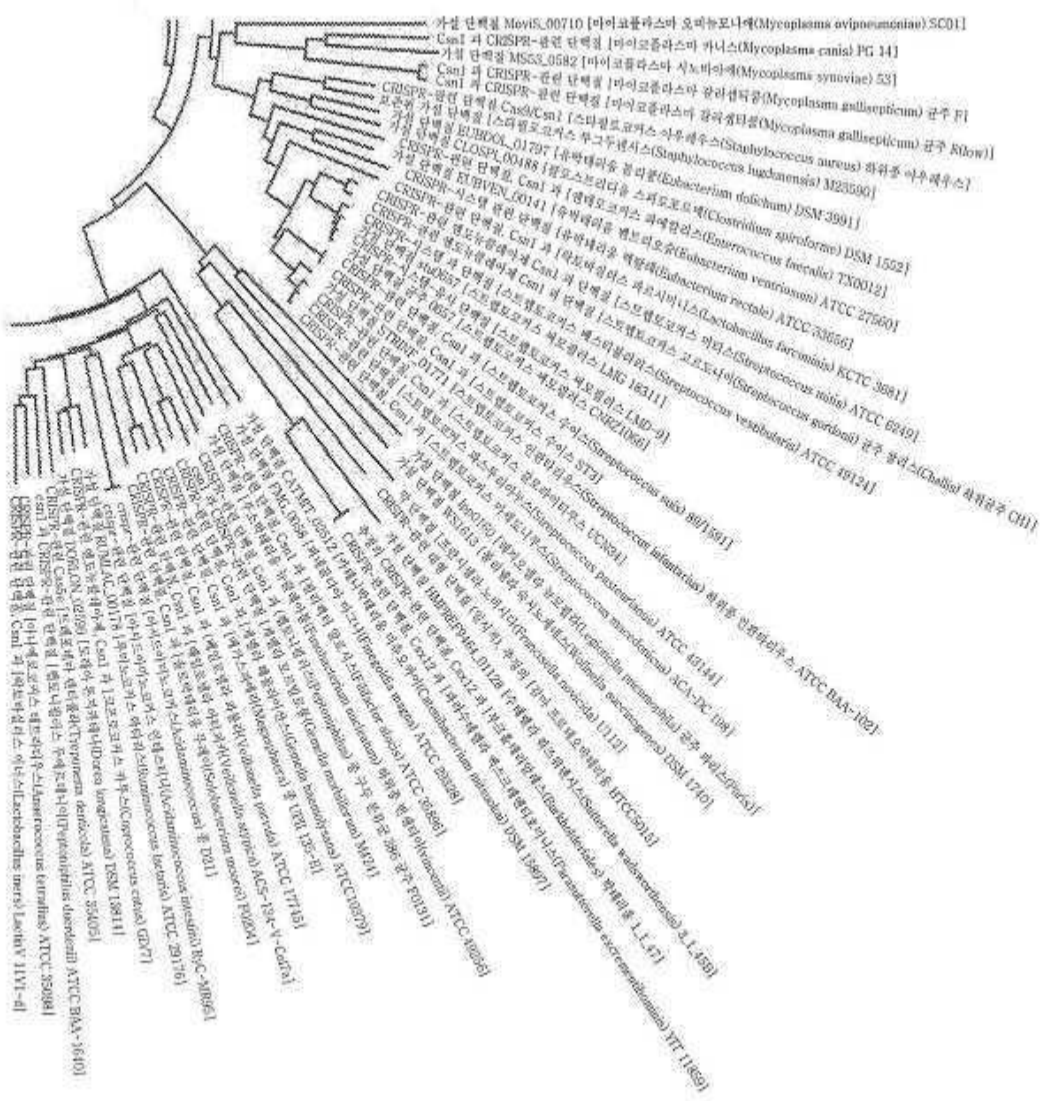
도면19b


CRISPR-관련 단백질, Csn1 과 [니트라티프랙터 살수기니스(Nitratifractor salsuginis) DSM 16511]
가설 단백질 MM0B0330 [마이코플라스마 포빌레(Mycoplasma mobile) 163K]

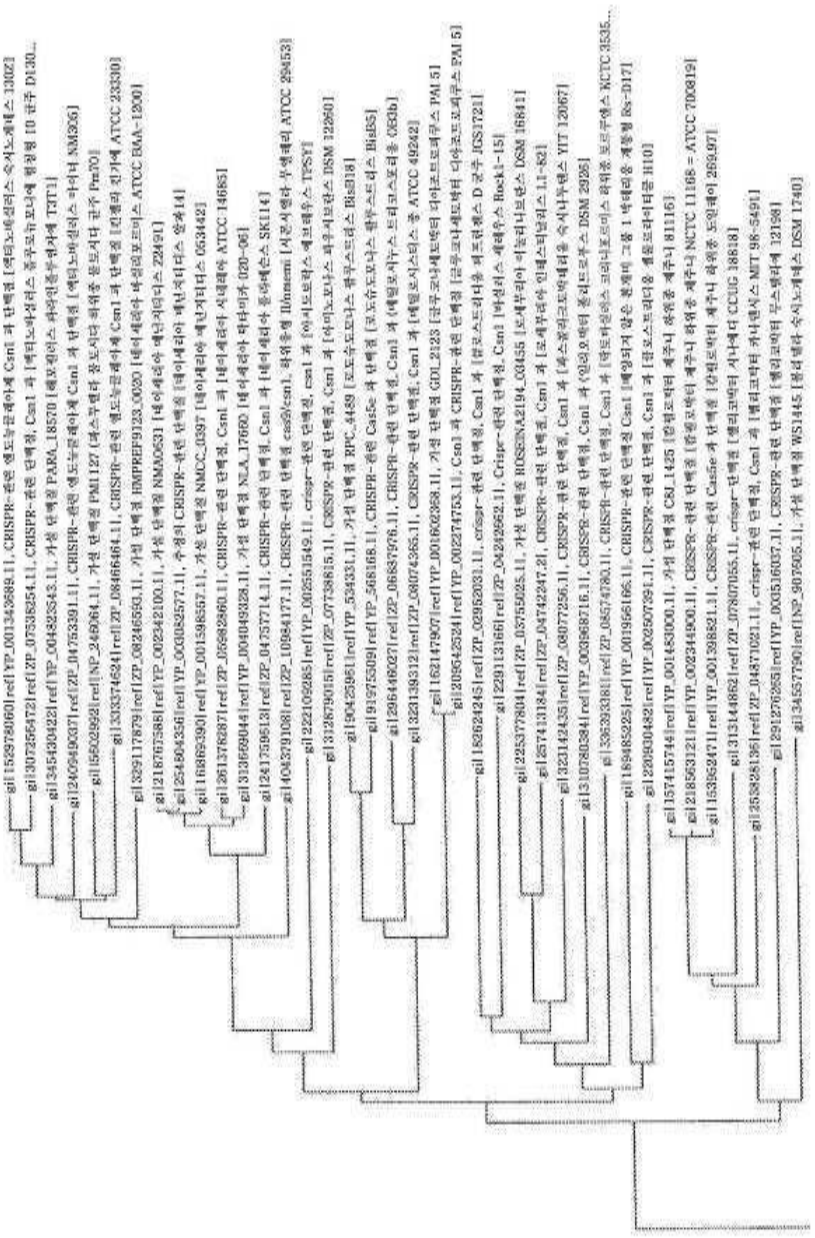
도면19c



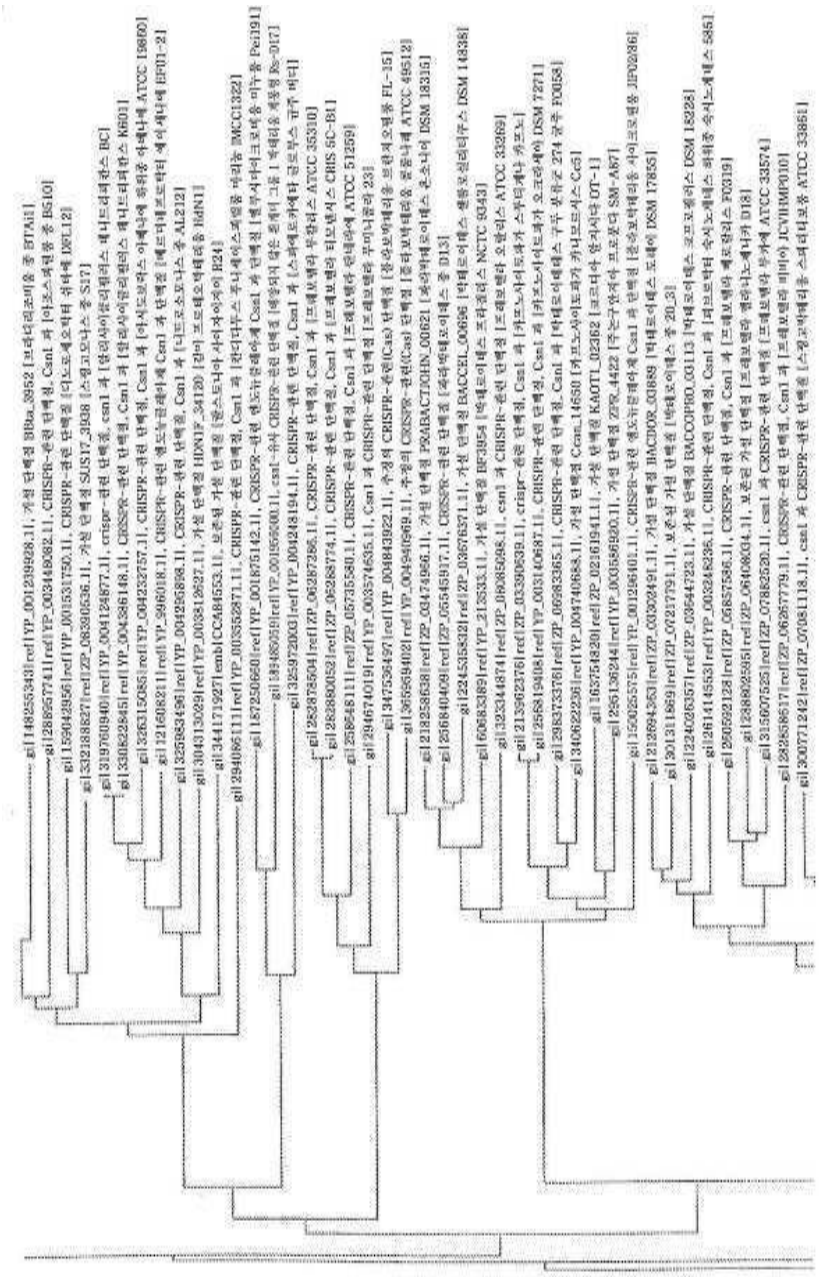
도면19d



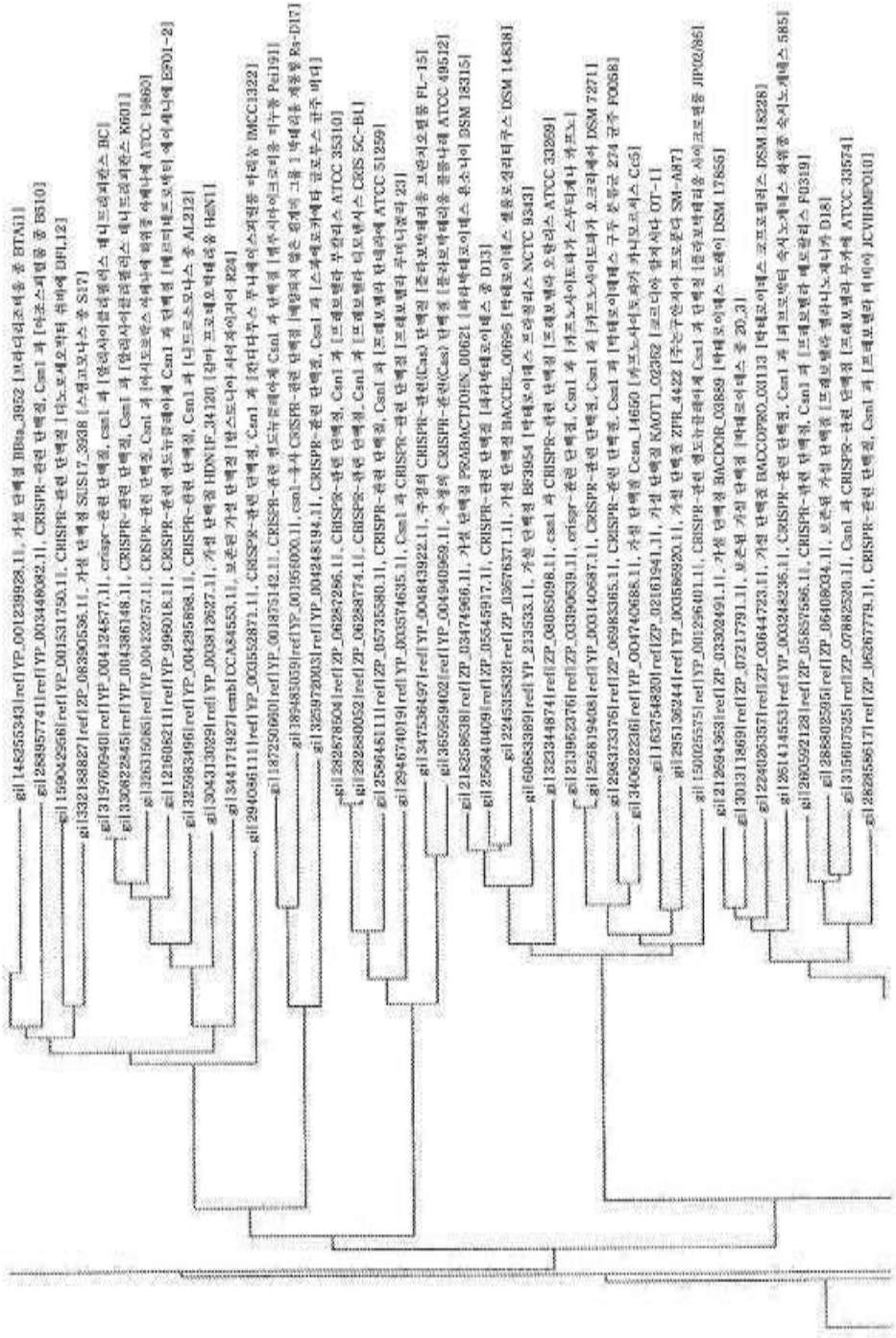
도면20a



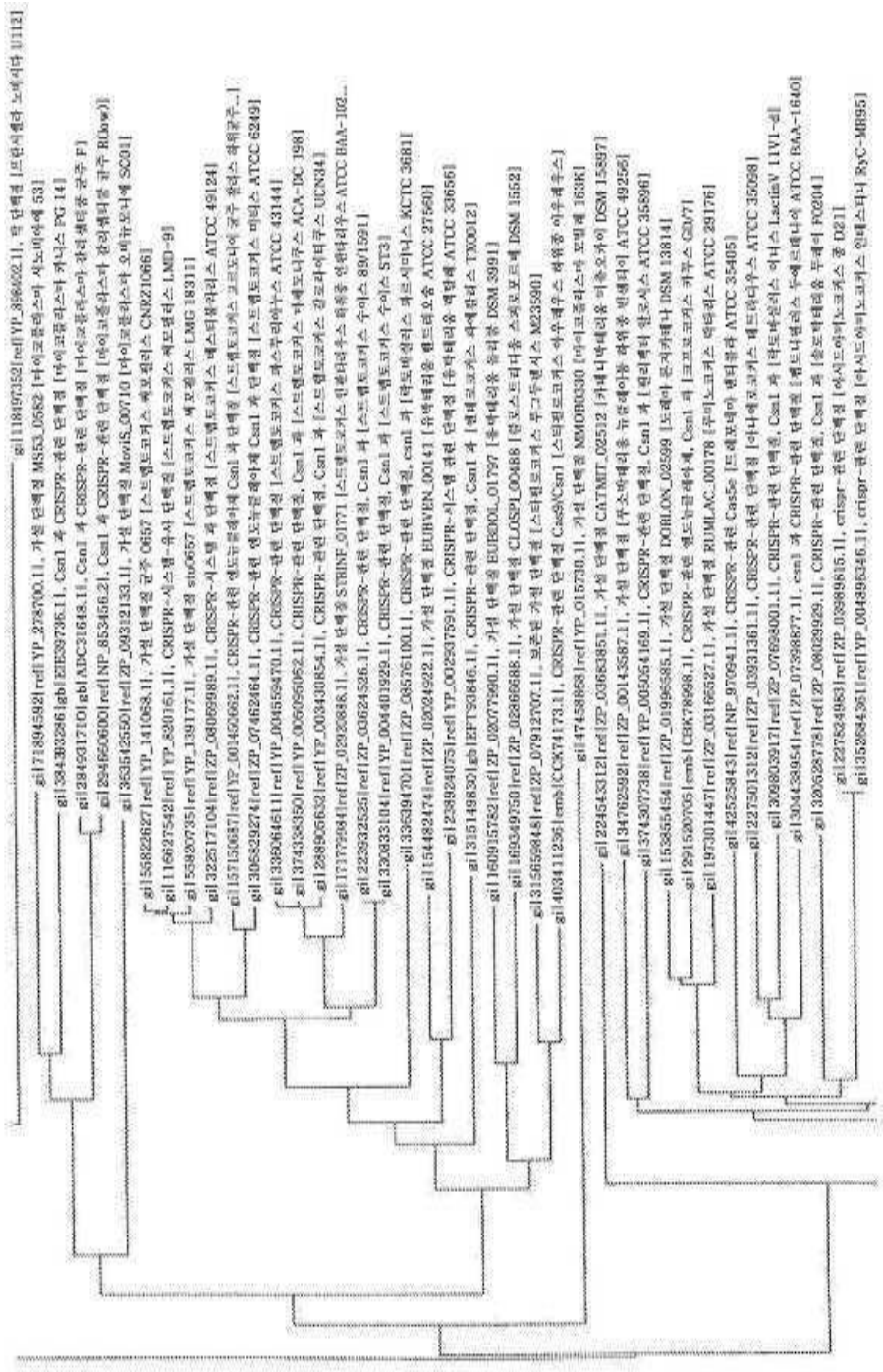
도면20b



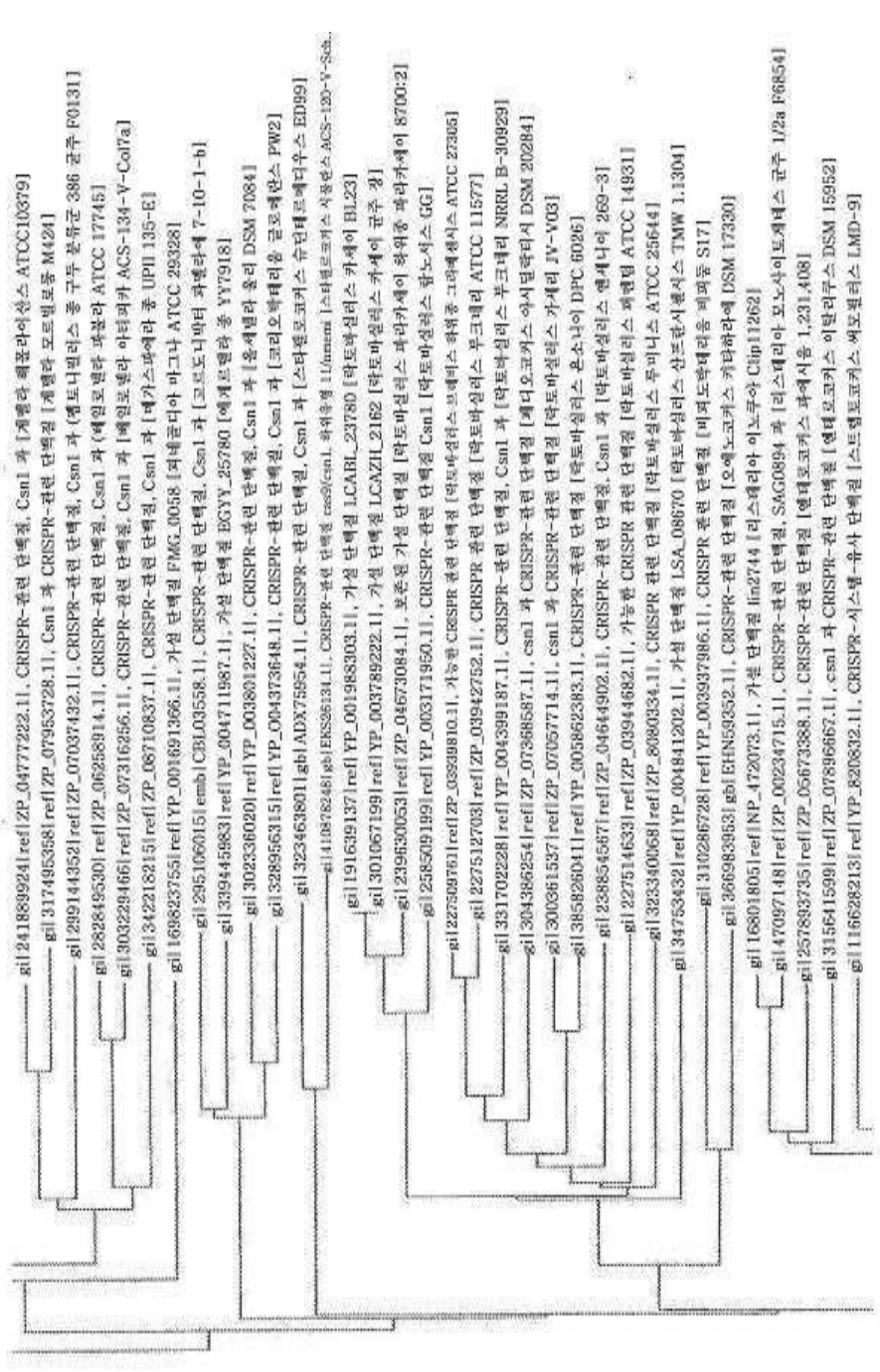
도면20c



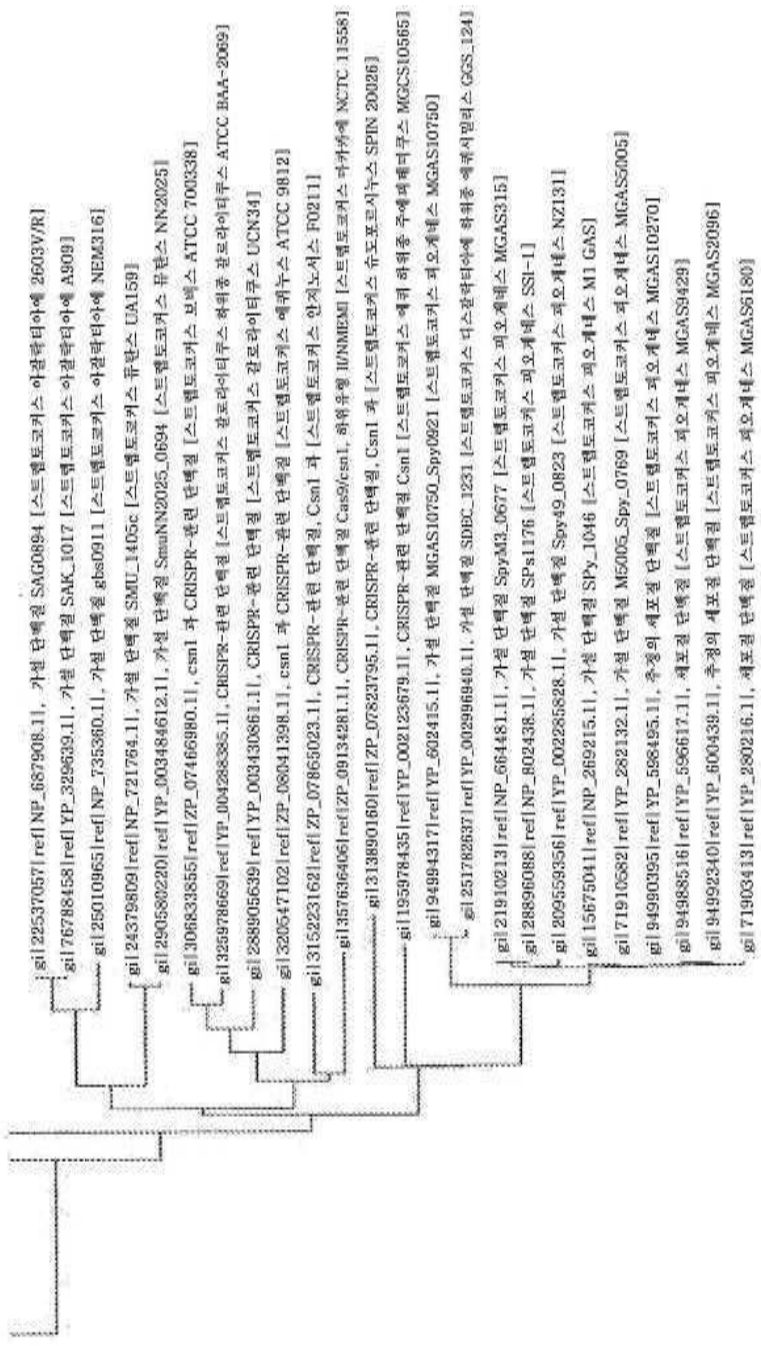
도면20d



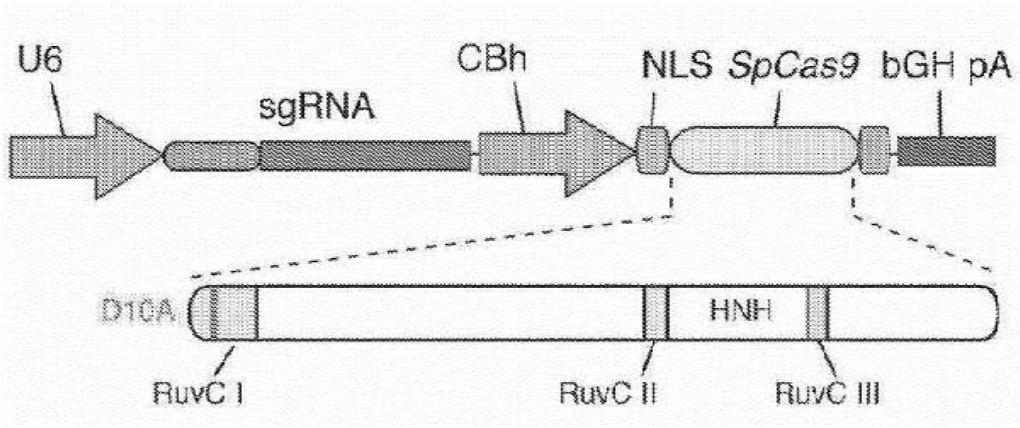
도면20e


gi|241889924|ref|ZP_04777222.1|, CRISPR-관련 단백질, Csn1 과 [계통군 헤몰라이신스 ATCC10379]
gi|317495358|ref|ZP_07953728.1|, Csn1 과 CRISPR-관련 단백질 [계통군 모르빌로종 M424]
gi|299144352|ref|ZP_07037432.1|, CRISPR-관련 단백질, Csn1 과 [캠토니필러스 종 구두 분류군 386 균주 F0131]
gi|282849530|ref|ZP_06258814.1|, CRISPR-관련 단백질, Csn1 과 [테일로발라 속할라 ATCC 17745]
gi|303228466|ref|ZP_07316256.1|, CRISPR-관련 단백질, Csn1 과 [베일로발라 아테피카 ACS-134-V-Col7a]
gi|342218215|ref|ZP_08710837.1|, CRISPR-관련 단백질, Csn1 과 [네가스파에라 종 UPII 135-E]
gi|169823755|ref|YP_001691366.1|, 가설 단백질 PMG_00518 [파에클디아 마그나 ATCC 29328]
gi|295106015|emb|CBL03558.1|, CRISPR-관련 단백질, Csn1 과 [코프도디바터 파톨라에 7-10-1-b]
gi|339445883|ref|YP_004711987.1|, 가설 단백질 EGYY_25780 [에게르넬라 종 YY7918]
gi|302336020|ref|YP_003801227.1|, CRISPR-관련 단백질, Csn1 과 [올세넬라 올리 DSM 7084]
gi|328956315|ref|YP_004373648.1|, CRISPR-관련 단백질, Csn1 과 [크리오뱅테리움 글로베란스 PW2]
gi|323463801|gb|ADX75954.1|, CRISPR-관련 단백질, Csn1 과 [스타필로코커스 슈던테르메리우스 ED99]
gi|410878248|gb|EKS26134.1|, CRISPR-관련 단백질 cas9/csn1 유사유형 II/Unemi [스타필로코커스 사프로피티스 ACS-120-V-Sch1]
gi|191639137|ref|YP_001988303.1|, 가설 단백질 LCABL_23780 [락토바실러스 카세이 BL23]
gi|301067199|ref|YP_003789222.1|, 가설 단백질 LCAZH_2162 [락토바실러스 파라카세이 파라카세이 하위종 파라카세이 8700:2]
gi|239630053|ref|ZP_04673084.1|, 보존된 가설 단백질 [락토바실러스 파라카세이 하위종 파라카세이 8700:2]
gi|258509199|ref|YP_003171950.1|, CRISPR-관련 단백질 Csn1 [락토바실러스 람노서스 GG]
gi|227509761|ref|ZP_03939810.1|, 가능한 CRISPR 관련 단백질 [락토바실러스 브레비스 하위종 그라에필러스 ATCC 27305]
gi|227512703|ref|ZP_03942752.1|, CRISPR 관련 단백질 [락토바실러스 루크레리 ATCC 11577]
gi|331702228|ref|YP_004399187.1|, CRISPR-관련 단백질, Csn1 과 [락토바실러스 루크레리 NRRL B-30929]
gi|304386254|ref|ZP_07368587.1|, csn1 과 CRISPR-관련 단백질 [페디오코커스 아시딜락티시 DSM 20284]
gi|300361537|ref|ZP_07057714.1|, csn1 과 CRISPR-관련 단백질 [락토바실러스 카세리 JV-V03]
gi|385826041|ref|YP_005862383.1|, CRISPR-관련 단백질 [락토바실러스 온스나이 DPC 6026]
gi|238854567|ref|ZP_04644902.1|, CRISPR-관련 단백질, Csn1 과 [락토바실러스 델브뤼이 269-3]
gi|227514633|ref|ZP_03944682.1|, 가능한 CRISPR 관련 단백질 [락토바실러스 파렌딘 ATCC 14931]
gi|323340068|ref|ZP_8080334.1|, CRISPR 관련 단백질 [락토바실러스 루미너스 ATCC 25644]
gi|347534321|ref|YP_004841202.1|, 가설 단백질 LSA_08670 [락토바실러스 살르타시넬시스 TMW 1.1304]
gi|310286728|ref|YP_003937986.1|, CRISPR 관련 단백질 [버크홀데리아 페크토룸 S17]
gi|366983953|gb|EHN59352.1|, CRISPR-관련 단백질 [오에노코커스 카타하레 DSM 17330]
gi|168018051|ref|NP_472073.1|, 가설 단백질 lin2744 [리스테리아 이노쿠아 Clip11262]
gi|47097148|ref|ZP_00234715.1|, CRISPR-관련 단백질, SAG0894 과 [리스테리아 모노사이토제네스 균주 1/2a F6854]
gi|257893735|ref|ZP_05673388.1|, CRISPR-관련 단백질 [엔테로코커스 패커시움 1,231,408]
gi|315641599|ref|ZP_07896667.1|, csn1 과 CRISPR-관련 단백질 [엔테로코커스 이탈리쿠스 DSM 15952]
gi|116628213|ref|YP_820832.1|, CRISPR-시스템-유사 단백질 [스트렙토코커스 써모필러스 LMD-9]

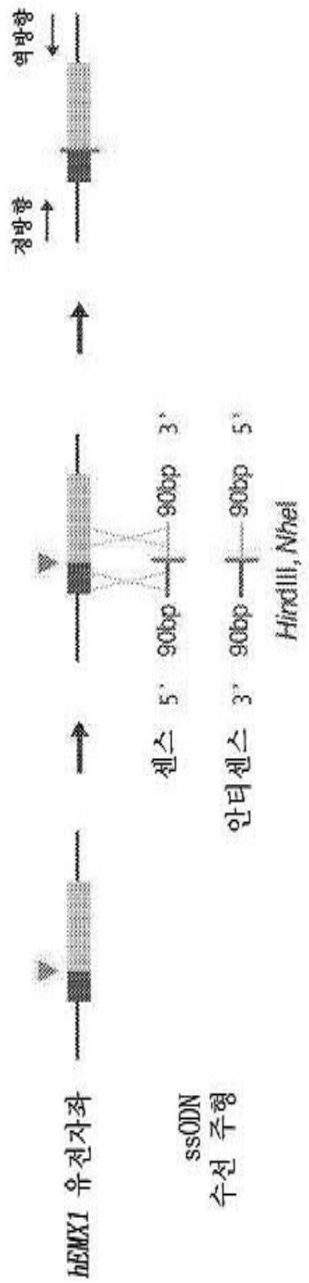
도면20f


gi|225370571|ref|NP_687908.1|, 가설 단백질 SAG0894 [스트렙토코커스 아갈락티아에 2603V/R]
gi|76788458|ref|YP_329639.1|, 가설 단백질 SAK_1017 [스트렙토코커스 아갈락티아에 A909]
gi|25010965|ref|NP_735360.1|, 가설 단백질 gbs0911 [스트렙토코커스 아갈락티아에 NEM316]
gi|243798809|ref|NP_721764.1|, 가설 단백질 SMU_140sc [스트렙토코커스 뮤탄스 UA159]
gi|290580220|ref|YP_003484612.1|, 가설 단백질 SmuNN2025_0694 [스트렙토코커스 뮤탄스 NN2025]
gi|306833855|ref|ZP_07466980.1|, csn1 과 CRISPR-관련 단백질 [스트렙토코커스 브베스 ATCC 700338]
gi|325978669|ref|YP_004288385.1|, CRISPR-관련 단백질 [스트렙토코커스 잘로라이터루스 잘로라이터루스 ATCC BAA-2069]
gi|288905639|ref|YP_003430861.1|, CRISPR-관련 단백질 [스트렙토코커스 잘로라이터루스 UCN34]
gi|320547102|ref|ZP_08041398.1|, csn1 과 CRISPR-관련 단백질 [스트렙토코커스 에퀴누스 ATCC 9812]
gi|315222162|ref|ZP_07865023.1|, CRISPR-관련 단백질, Csn1 과 [스트렙토코커스 인지노서스 F0211]
gi|357636406|ref|ZP_09134281.1|, CRISPR-관련 단백질 Cas9/csn1, 하위유형 II/NMEMI [스트렙토코커스 아카바에 NCTC 11558]
gi|313890160|ref|ZP_07823795.1|, CRISPR-관련 단백질, Csn1 과 [스트렙토코커스 슈도포르시누스 SPIN 20026]
gi|195978435|ref|YP_002123679.1|, CRISPR-관련 단백질 Csn1 [스트렙토코커스 에퀴 하위종 주에퍼테무스 MGCS10565]
gi|194994317|ref|YP_602415.1|, 가설 단백질 MGAS10750_Spy0921 [스트렙토코커스 피오게네스 MGAS10750]
gi|251782537|ref|YP_002996940.1|, 가설 단백질 SDEC_1231 [스트렙토코커스 디스갈락티아에 하위종 에퀴시밀리스 GGS_124]
gi|219910213|ref|NP_664481.1|, 가설 단백질 Spy_M3_0677 [스트렙토코커스 피오게네스 MGAS315]
gi|288906088|ref|NP_802438.1|, 가설 단백질 SPs1176 [스트렙토코커스 피오게네스 SSI-1]
gi|209559356|ref|YP_002286828.1|, 가설 단백질 Spy49_0823 [스트렙토코커스 피오게네스 NZ131]
gi|15675041|ref|NP_269215.1|, 가설 단백질 SPy_1046 [스트렙토코커스 피오게네스 M1 GAS]
gi|71910582|ref|YP_282132.1|, 가설 단백질 M5005_Spy_0769 [스트렙토코커스 피오게네스 MGAS5005]
gi|94990395|ref|YP_598495.1|, 추정의 세포질 단백질 [스트렙토코커스 피오게네스 MGAS10270]
gi|94988516|ref|YP_596617.1|, 세포질 단백질 [스트렙토코커스 피오게네스 MGAS9429]
gi|94992340|ref|YP_600439.1|, 추정의 세포질 단백질 [스트렙토코커스 피오게네스 MGAS2096]
gi|71903413|ref|YP_280216.1|, 세포질 단백질 [스트렙토코커스 피오게네스 MGAS6180]

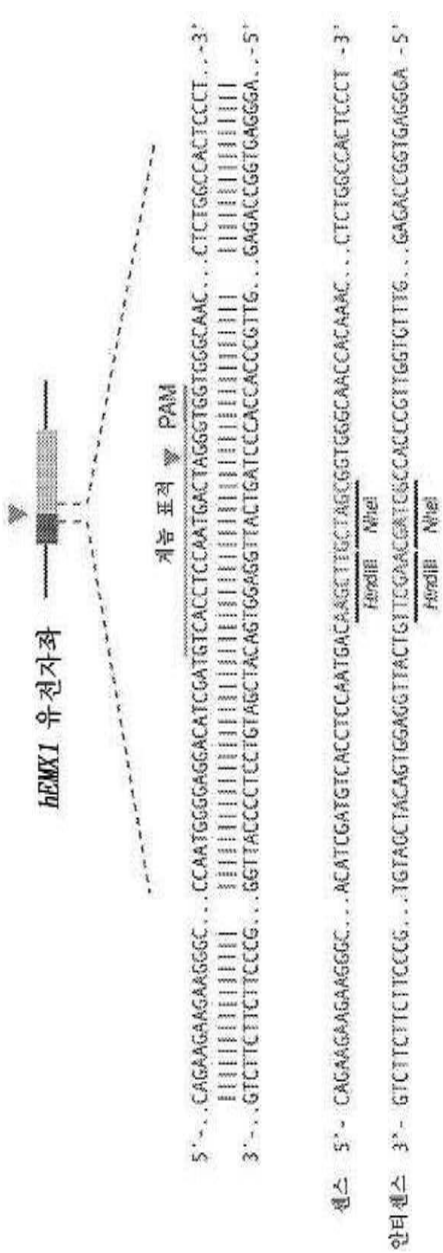
도면21a



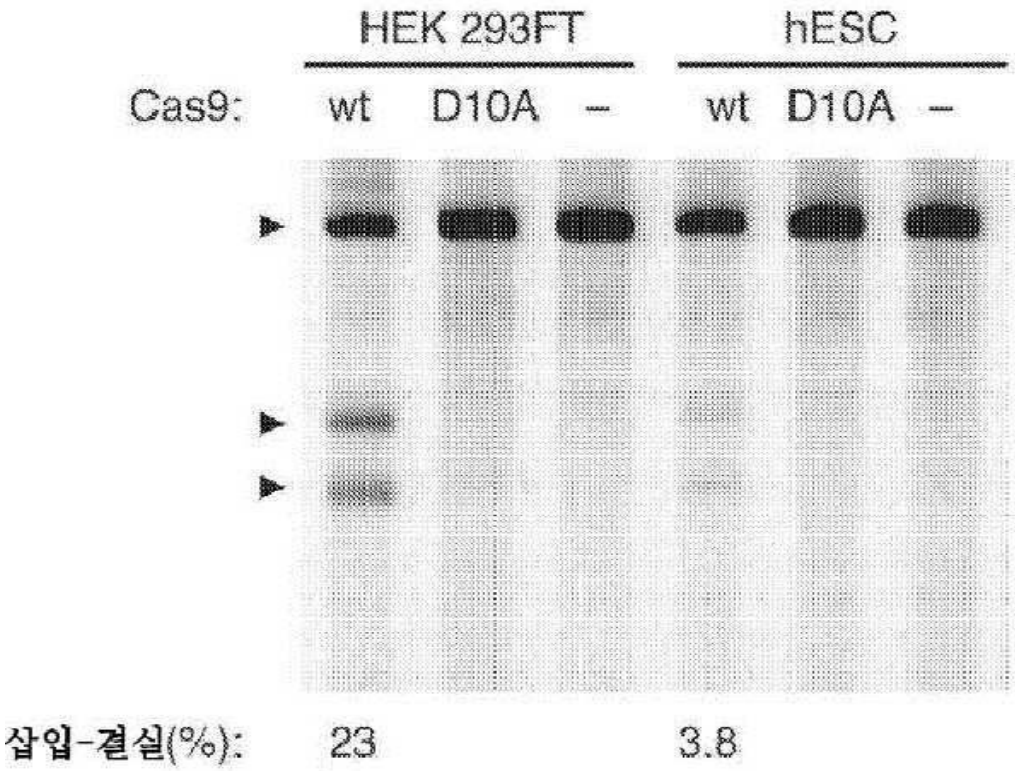
도면21b


정방향
역방향
hEMX1 유전자좌
ssODN
수선 주형
센스 5' 90bp 90bp 3'
안티센스 3' 90bp 90bp 5'
HindIII, NheI

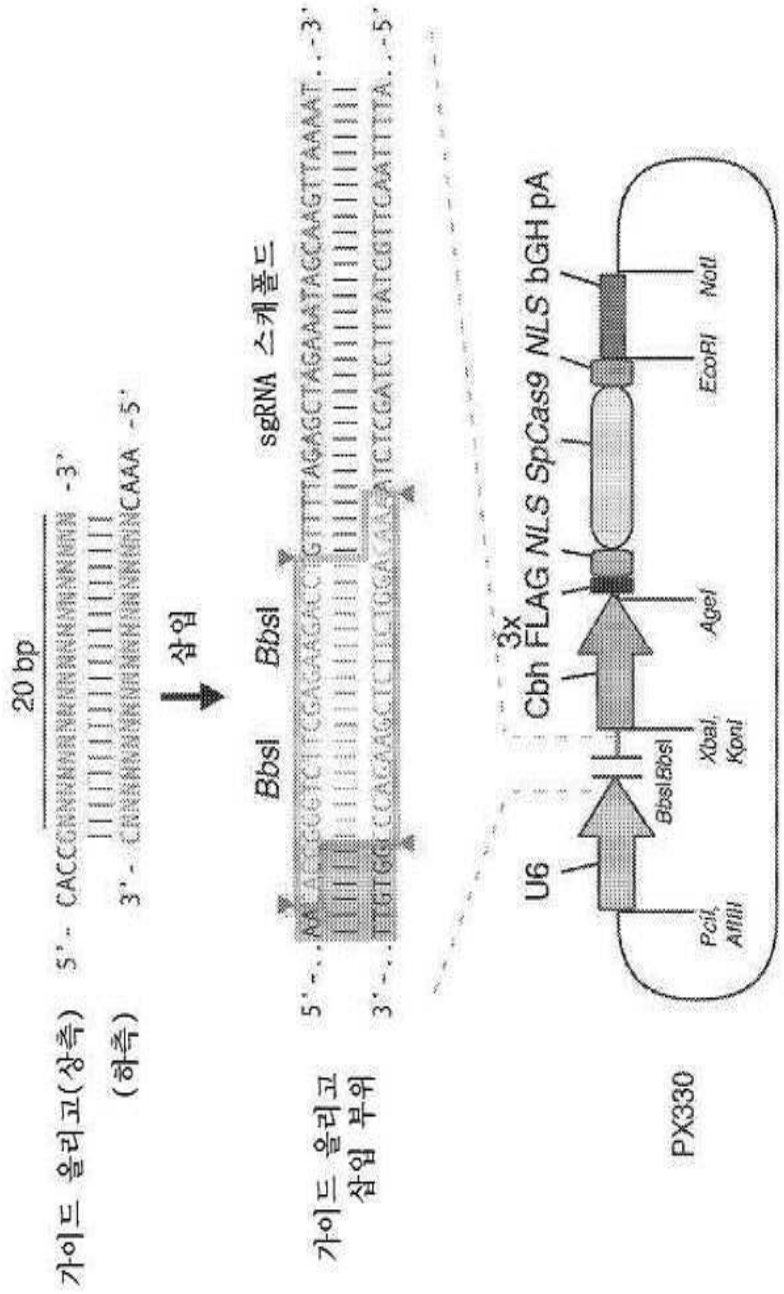
도면21c


hEMX1 유전자좌
게놈 표적
PAM
5'-..CAGAAGAAGAAGGGC...CCAATGGGGAGGACATCGATGTCACCTCCAATGACTAGGGTGGTGGGCAAC...CTCTGGCCACTCCCT..-3'
3'-..GTCTTCTTCTTCCCG...GGTTACCCCTCCTGTAGCTACAGTGGAAGGTTACTGATCCCACCACCCGTTG...GAGACCGGTGAGGGA..-5'
센스 5'- CAGAAGAAGAAGGGC...ACATCGATGTCACCTCCAATGACAAGCTTGCTAGCGGTGGGCAACCACAAAC...CTCTGGCCACTCCCT -3'
HindIII NheI
안티센스 3'- GTCTTCTTCTTCCCG...TGTAGCTACAGTGGAAGGTTACTGTTCGAACCATCGCCACCCGTTGGTGTTTG...GAGACCGGTGAGGGA -5'
HindIII NheI

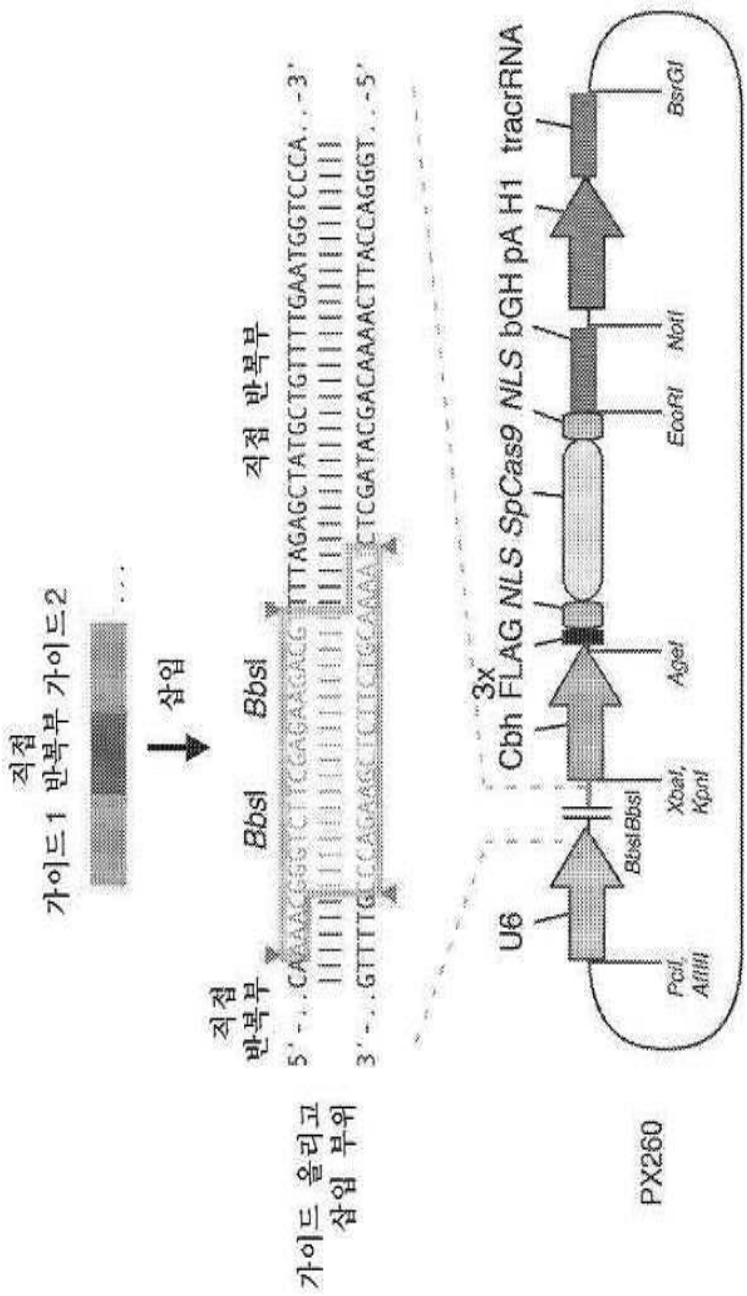
도면21d



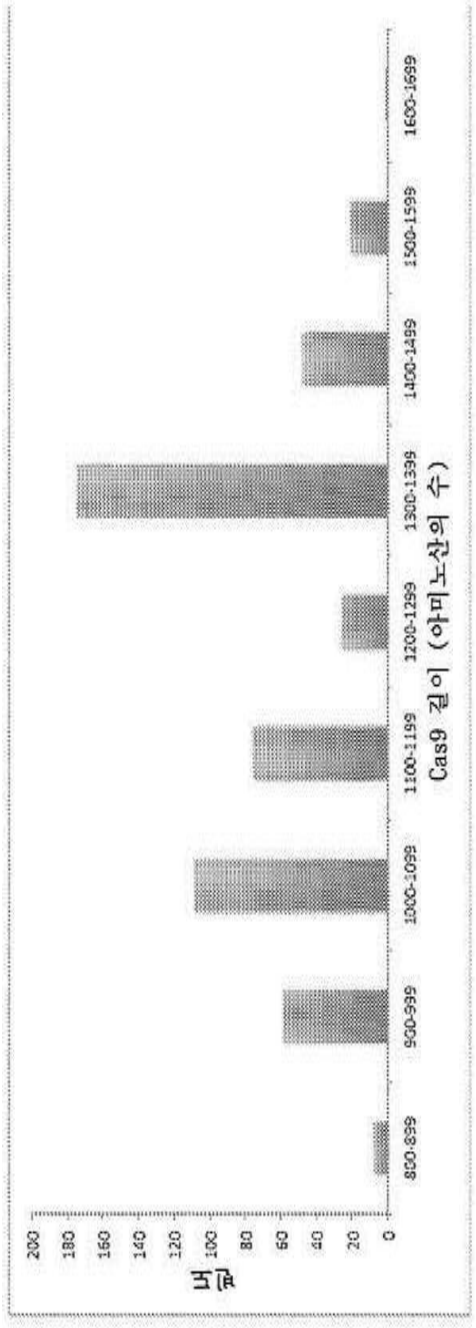
도면22a


20 bp
가이드 올리고(상측) 5'- CACC
(하측) 3'- CAAA -5'
삽입
BbsI
BbsI
sgRNA 스캐폴드
5'-...
3'-...
가이드 올리고 삽입 부위
U6
3x Cbh FLAG NLS SpCas9 NLS bGH pA
PX330
PciI, AflIII
BbsI BbsI
XbaI, KpnI
AgeI
EcoRI
NotI

도면22b

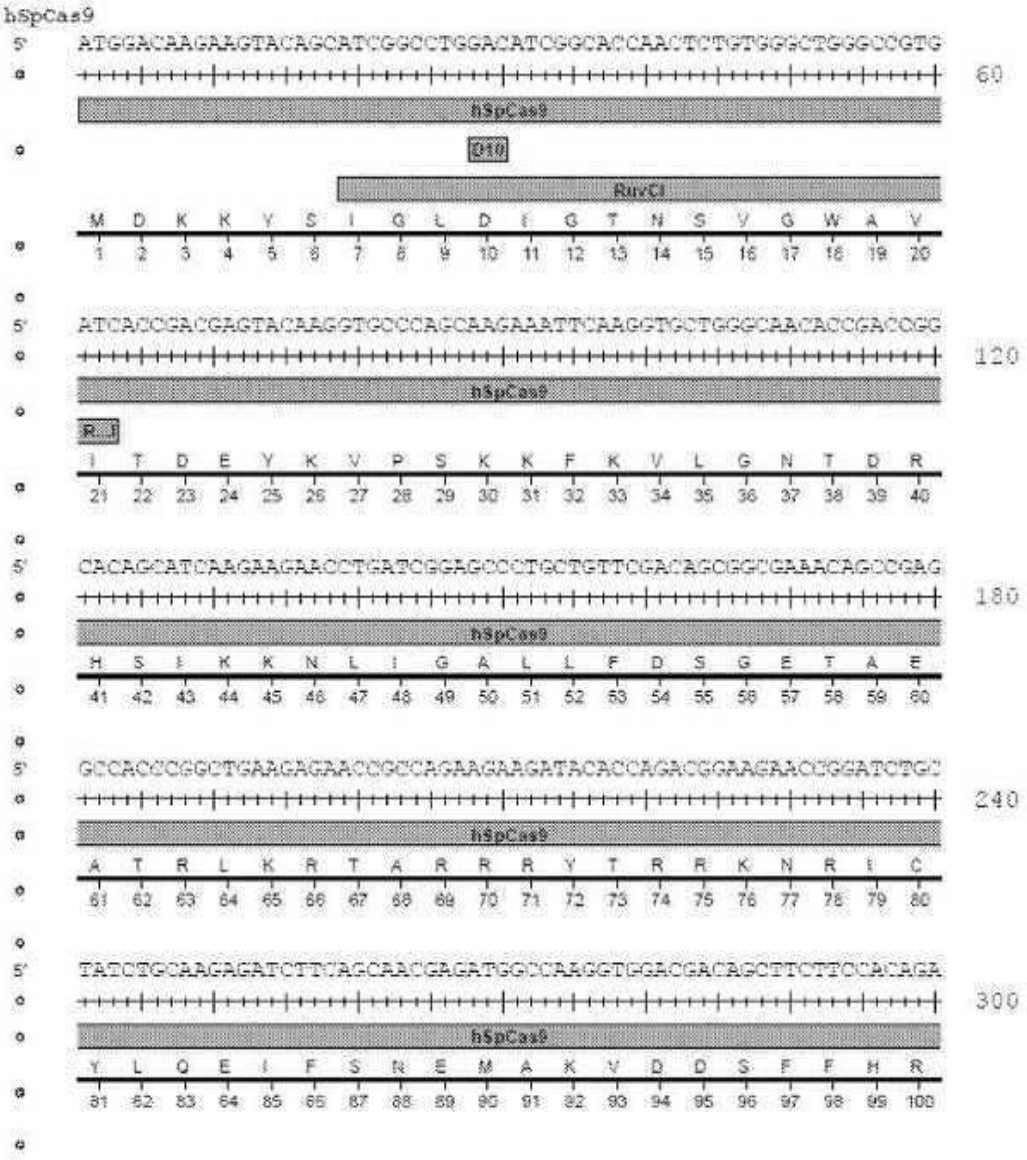

직접
가이드1 반복부 가이드2
삽입
BbsI
BbsI
직접 반복부
직접
반복부
가이드 올리고
삽입 부위
3x
U6
Cbh FLAG NLS SpCas9 NLS bGH pA H1 tracrRNA
PX260
PciI, AflIII
BbsI BbsI
XbaI, KpnI
AgeI
EcoRI
NotI
BsrGI

도면23

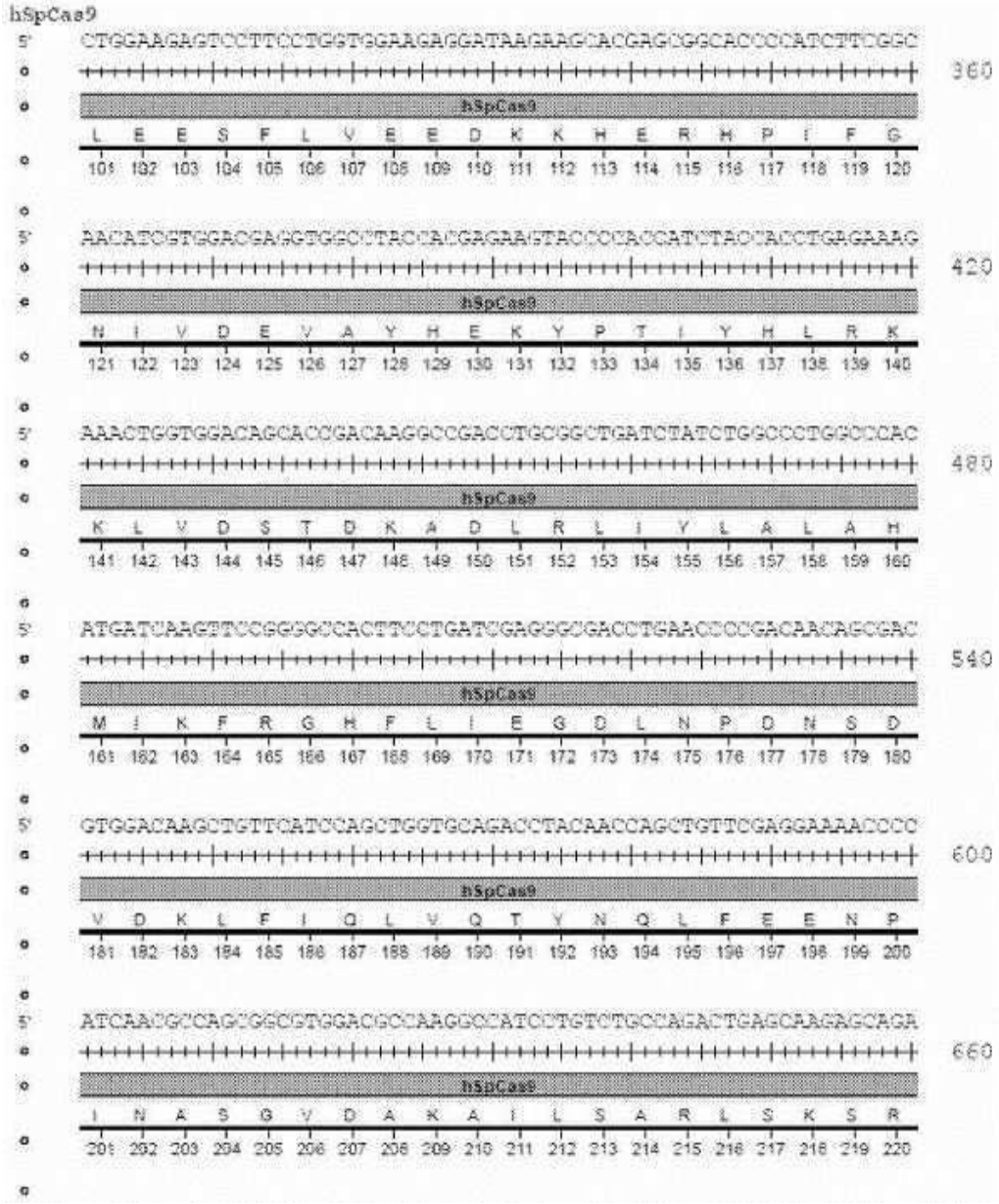

빈도
200
180
160
140
120
100
80
60
40
20
0
800-899
900-999
1000-1099
1100-1199
1200-1299
1300-1399
1400-1499
1500-1599
1600-1699
Cas9 길이 (아미노산의 수)

도면24a

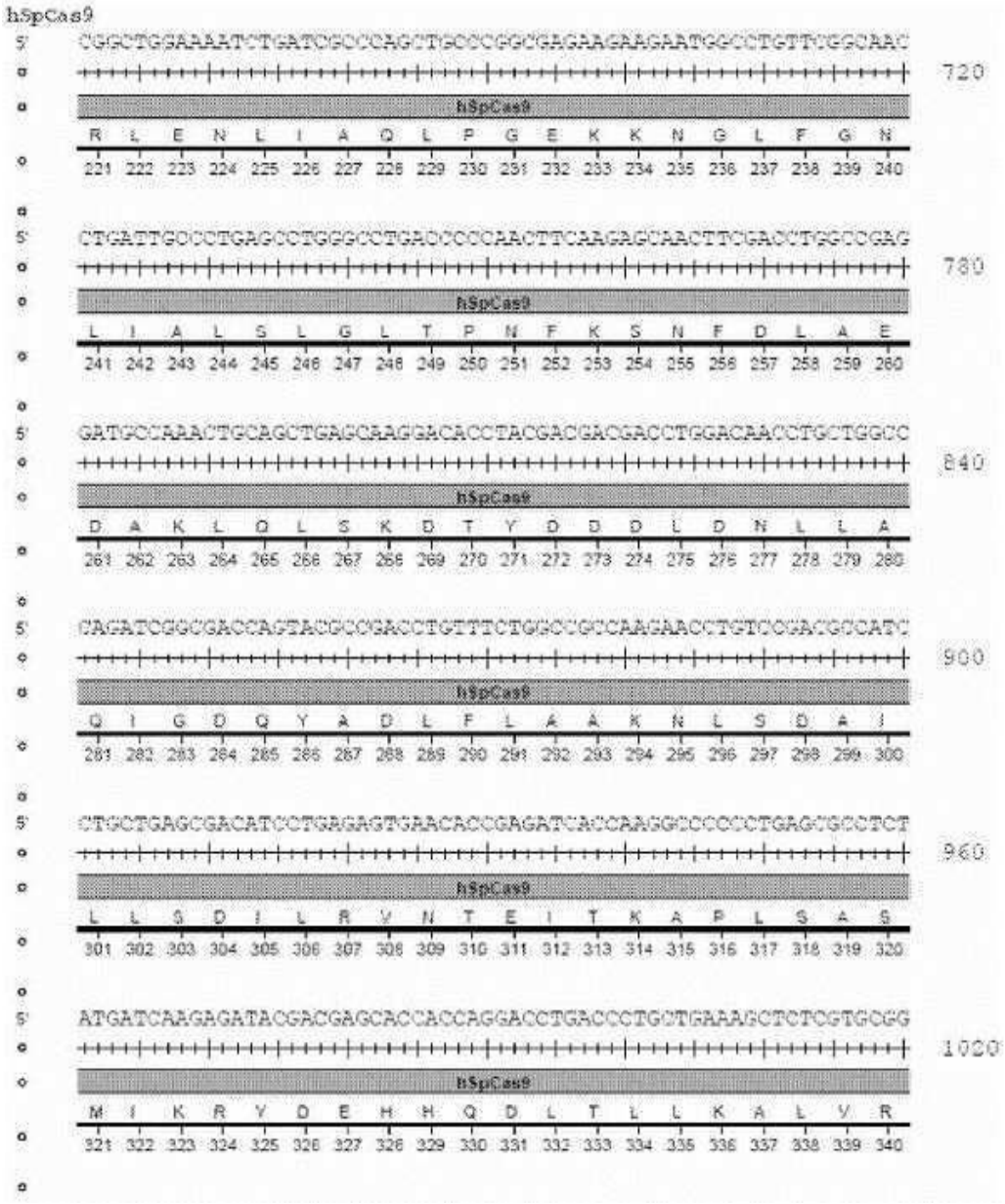
SpCas9 돌연변이 위치



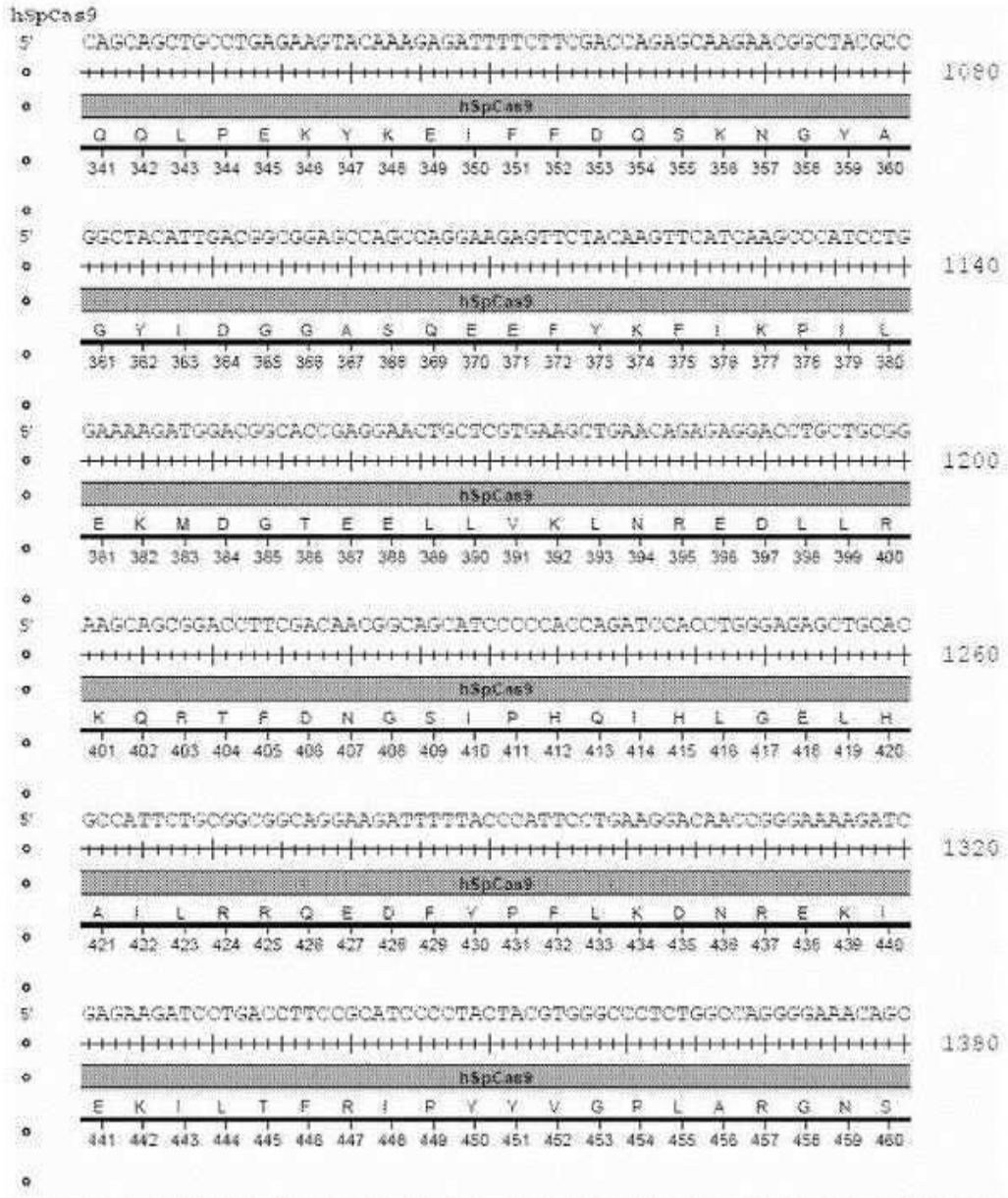
도면24b


hSpCas9
5'
CTGGAAGAGTCCTTCCTGGTGGAAGAGGATAAGAAGCACGAGCGGCACCCCATCTTCGGC
360
hSpCas9
L E E S F L V E E D K K H E R H P I F G
101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120
5'
AACATCGTGGACGAGGTGGCCTACCACGAGAAGTACCCCACCATCTACCACCTGAGAAAG
420
hSpCas9
N I V D E V A Y H E K Y P T I Y H L R K
121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140
5'
AAACTGGTGGACAGCACCGACAAGGCCGACCTGCGGCTGATCTATCTGGCCCTGGCCCAC
480
hSpCas9
K L V D S T D K A D L R L I Y L A L A H
141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160
5'
ATGATCAAGTTCCGGGGCCACTTCCTGATCGAGGGCGACCTGAACCCCGACAACAGCGAC
540
hSpCas9
M I K F R G H F L I E G D L N P D N S D
161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180
5'
GTGGACAAGCTGTTCATCCAGCTGGTGCAGACCTACAACCAGCTGTTCGAGGAAAACCCC
600
hSpCas9
V D K L F I Q L V Q T Y N Q L F E E N P
181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200
5'
ATCAACGCCAGCGGCGTGGACGCCAAGGCCATCCTGTCTGCCAGACTGAGCAAGAGCAGA
660
hSpCas9
I N A S G V D A K A I L S A R L S K S R
201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220

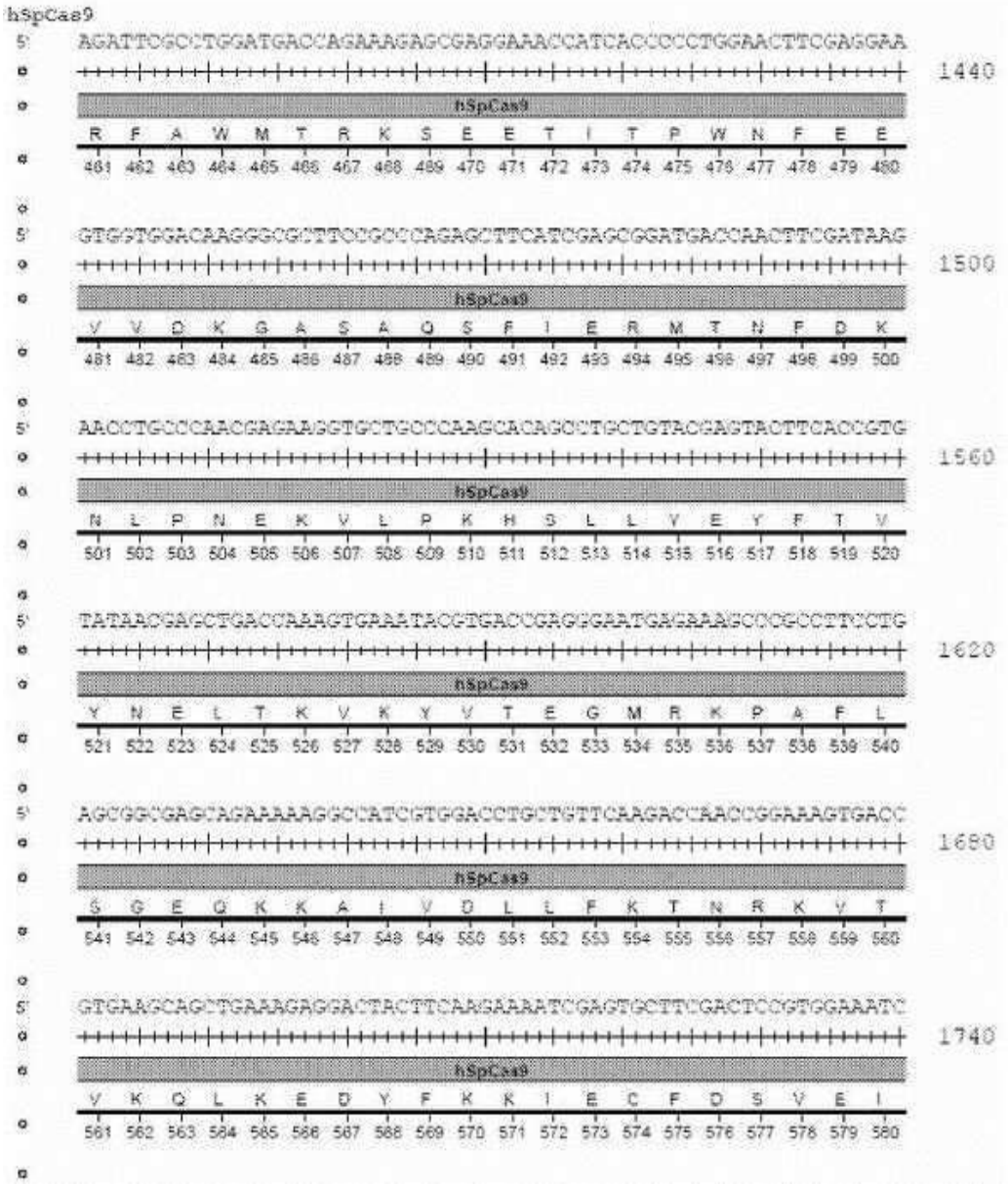
도면24c


hSpCas9
5'
CGGCTGGAAAATCTGATCGCCCAGCTGCCCGGCGAGAAGAAGAATGGCCTGTTCGGCAAC
720
hSpCas9
R L E N L I A Q L P G E K K N G L F G N
221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240
5'
CTGATTGCCCTGAGCCTGGGCCTGACCCCCAACTTCAAGAGCAACTTCGACCTGGCCGAG
780
hSpCas9
L I A L S L G L T P N F K S N F D L A E
241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260
5'
GATGCCAAACTGCAGCTGAGCAAGGACACCTACGACGACGACCTGGACAACCTGCTGGCC
840
hSpCas9
D A K L Q L S K D T Y D D D L D N L L A
261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280
5'
CAGATCGGCGACCAGTACGCCGACCTGTTTCTGGCCGCCAAGAACCTGTCCGACGCCATC
900
hSpCas9
Q I G D Q Y A D L F L A A K N L S D A I
281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300
5'
CTGCTGAGCGACATCCTGAGAGTGAACACCGAGATCACCAAGGCCCCCCTGAGCGCCTCT
960
hSpCas9
L L S D I L R V N T E I T K A P L S A S
301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320
5'
ATGATCAAGAGATACGACGAGCACCACCAGGACCTGACCCTGCTGAAAGCTCTCGTGCGG
1020
hSpCas9
M I K R Y D E H H Q D L T L L K A L V R
321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340

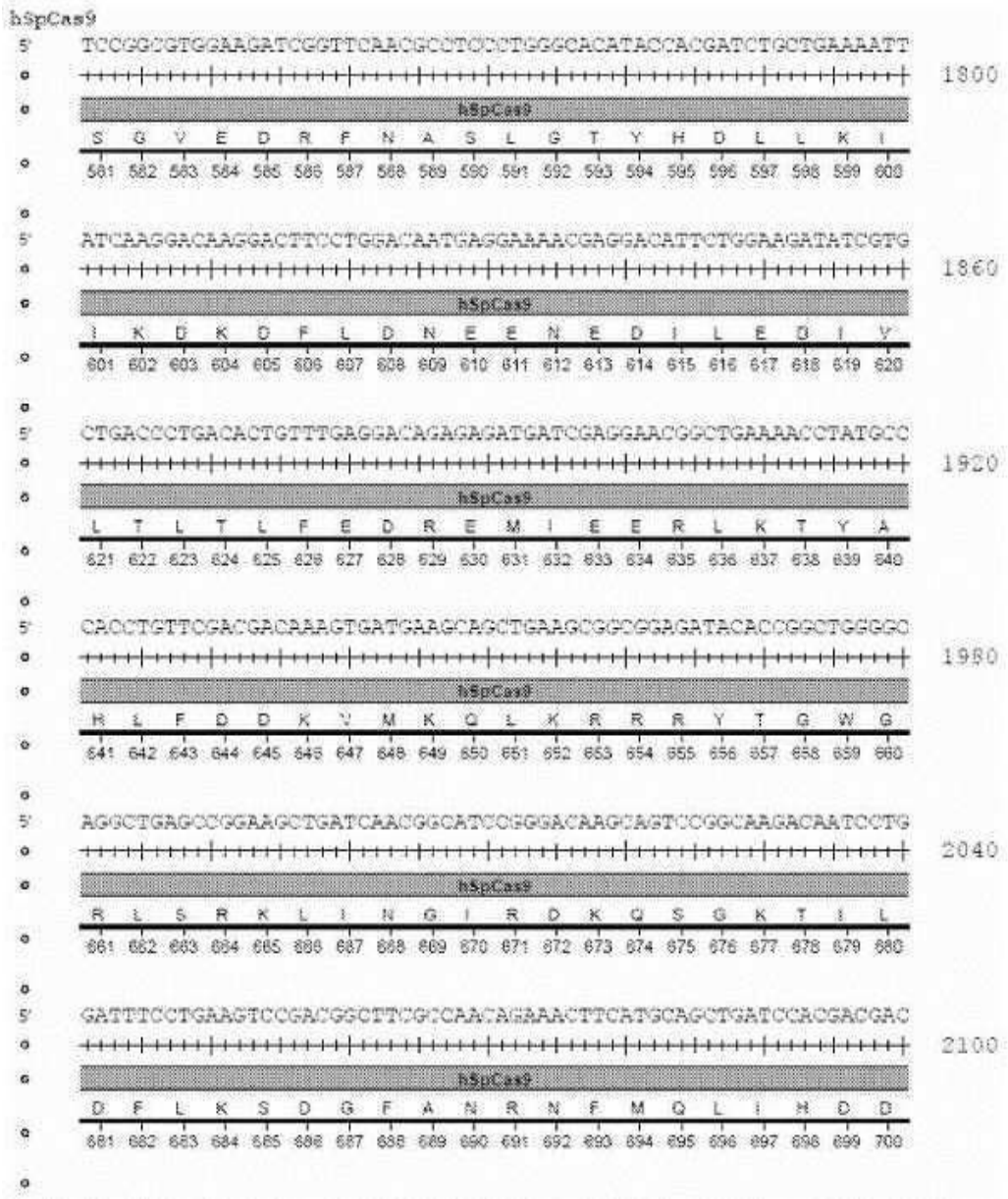
도면24d


hSpCas9
5' CAGCAGCTGCCTGAGAAGTACAAAGAGATTTTCTTCGACCAGAGCAAGAACGGCTACGCC
1080
hSpCas9
Q Q L P E K Y K E I F F D Q S K N G Y A
341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360
5' GGCTACATTGACGGCGGAGCCAGCCAGGAAGAGTTCTACAAGTTCATCAAGCCCATCCTG
1140
hSpCas9
G Y I D G G A S Q E E F Y K F I K P I L
361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380
5' GAAAAGATGGACGGCACCGAGGAACTGCTCGTGAAGCTGAACAGAGAGGACCTGCTGCGG
1200
hSpCas9
E K M D G T E E L L V K L N R E D L L R
381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400
5' AAGCAGCGGACCTTCGACAACGGCAGCATCCCCCACCAGATCCACCTGGGAGAGCTGCAC
1260
hSpCas9
K Q R T F D N G S I P H Q I H L G E L H
401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420
5' GCCATTCTGCGGCGGCAGGAAGATTTTTACCCATTCCTGAAGGACAACCGGGAAAAGATC
1320
hSpCas9
A I L R R Q E D F Y P F L K D N R E K I
421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440
5' GAGAAGATCCTGACCTTCCGCATCCCCTACTACGTGGGCCCTCTGGCCAGGGGAAACAGC
1380
hSpCas9
E K I L T F R I P Y Y V G P L A R G N S
441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460

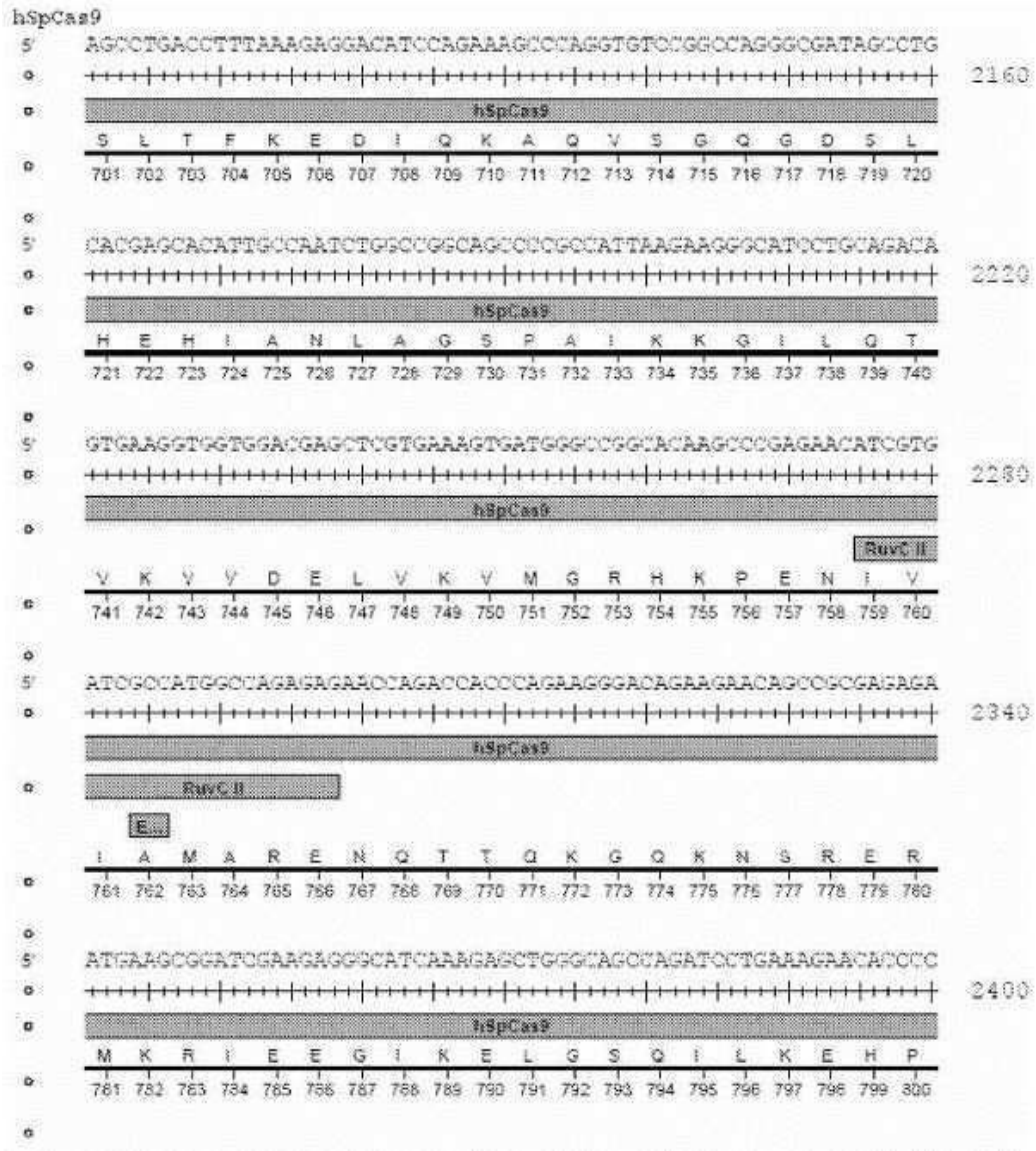
도면24e


hSpCas9
5'
AGATTCGCCTGGATGACCAGAAAGAGCGAGGAAACCATCACCCCCTGGAACTTCGAGGAA
1440
hSpCas9
R F A W M T R K S E E T I T P W N F E E
461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480
5'
GTGGTGGACAAGGGCGCTTCCGCCCAGAGCTTCATCGAGCGGATGACCAACTTCGATAAG
1500
hSpCas9
V V D K G A S A Q S F I E R M T N F D K
481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500
5'
AACCTGCCCAACGAGAAGGTGCTGCCCAAGCACAGCCTGCTGTACGAGTACTTCACCGTG
1560
hSpCas9
N L P N E K V L P K H S L L Y E Y F T V
501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520
5'
TATAACGAGCTGACCAAAGTGAAATACGTGACCGAGGGAATGAGAAAGCCCGCCTTCCTG
1620
hSpCas9
Y N E L T K V K Y V T E G M R K P A F L
521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540
5'
AGCGGCGAGCAGAAAAAGGCCATCGTGGACCTGCTGTTCAAGACCAACCGGAAAGTGACC
1680
hSpCas9
S G E Q K K A I V D L L F K T N R K V T
541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560
5'
GTGAAGCAGCTGAAAGAGGACTACTTCAAGAAAATCGAGTGCTTCGACTCCGTGGAAATC
1740
hSpCas9
V K Q L K E D Y F K K I E C F D S V E I
561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580

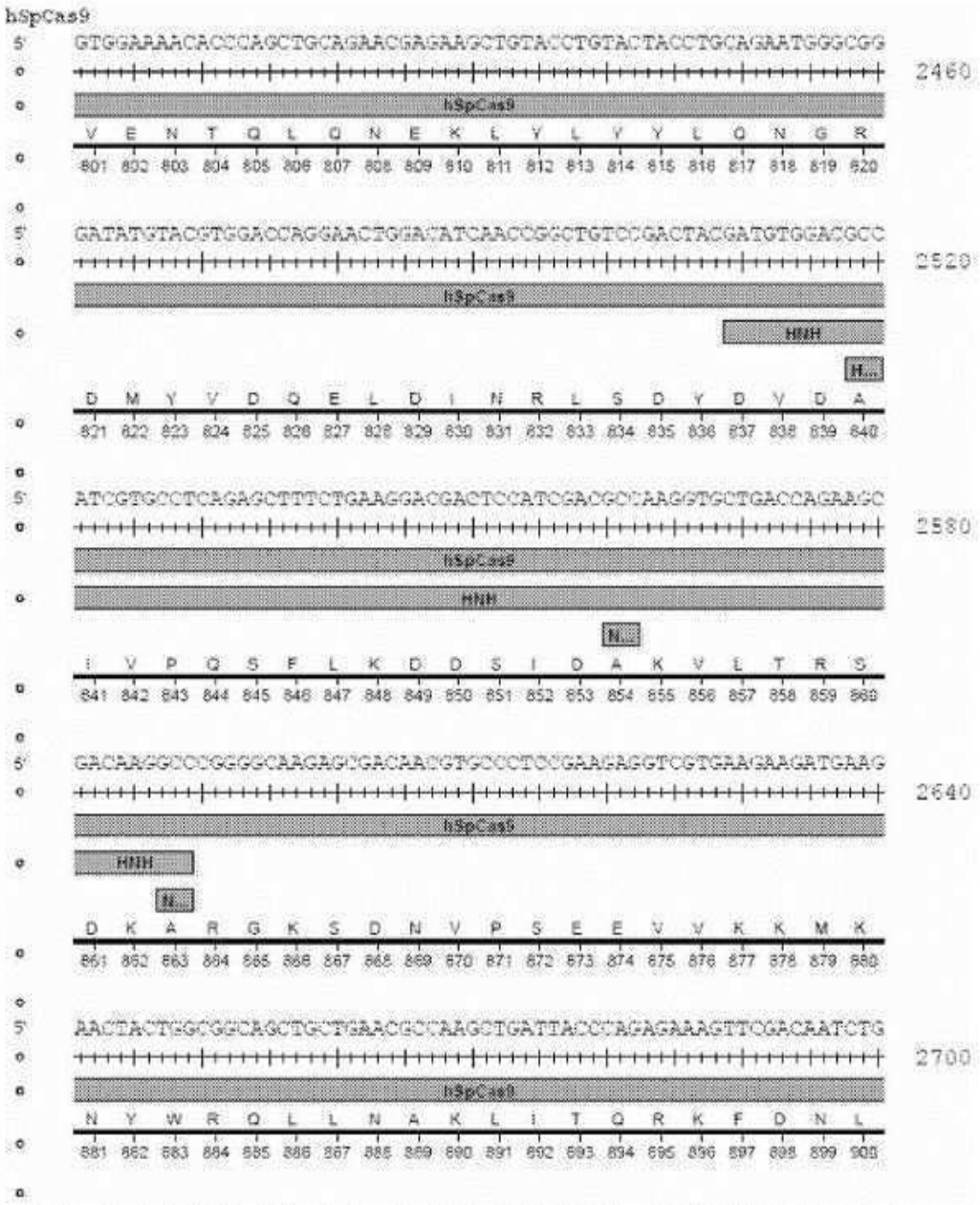
도면24f


hSpCas9
5'
TCCGGCGTGGAAGATCGGTTCAACGCCTCCCTGGGCACATACCACGATCTGCTGAAAATT
1800
hSpCas9
S G V E D R F N A S L G T Y H D L L K I
581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600
5'
ATCAAGGACAAGGACTTCCTGGACAATGAGGAAAACGAGGACATTCTGGAAGATATCGTG
1860
hSpCas9
I K D K D F L D N E E N E D I L E D I V
601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620
5'
CTGACCCTGACACTGTTTGAGGACAGAGAGATGATCGAGGAACGGCTGAAAACCTATGCC
1920
hSpCas9
L T L T L F E D R E M I E E R L K T Y A
621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640
5'
CACCTGTTCGACGACAAAGTGATGAAGCAGCTGAAGCGGCGGAGATACACCGGCTGGGGC
1980
hSpCas9
H L F D D K V M K Q L K R R R Y T G W G
641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660
5'
AGGCTGAGCCGGAAGCTGATCAACGGCATCCGGGACAAGCAGTCCGGCAAGACAATCCTG
2040
hSpCas9
R L S R K L I N G I R D K Q S G K T I L
661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680
5'
GATTTCCTGAAGTCCGACGGCTTCGCCAACAGAAACTTCATGCAGCTGATCCACGACGAC
2100
hSpCas9
D F L K S D G F A N R N F M Q L I H D D
681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700

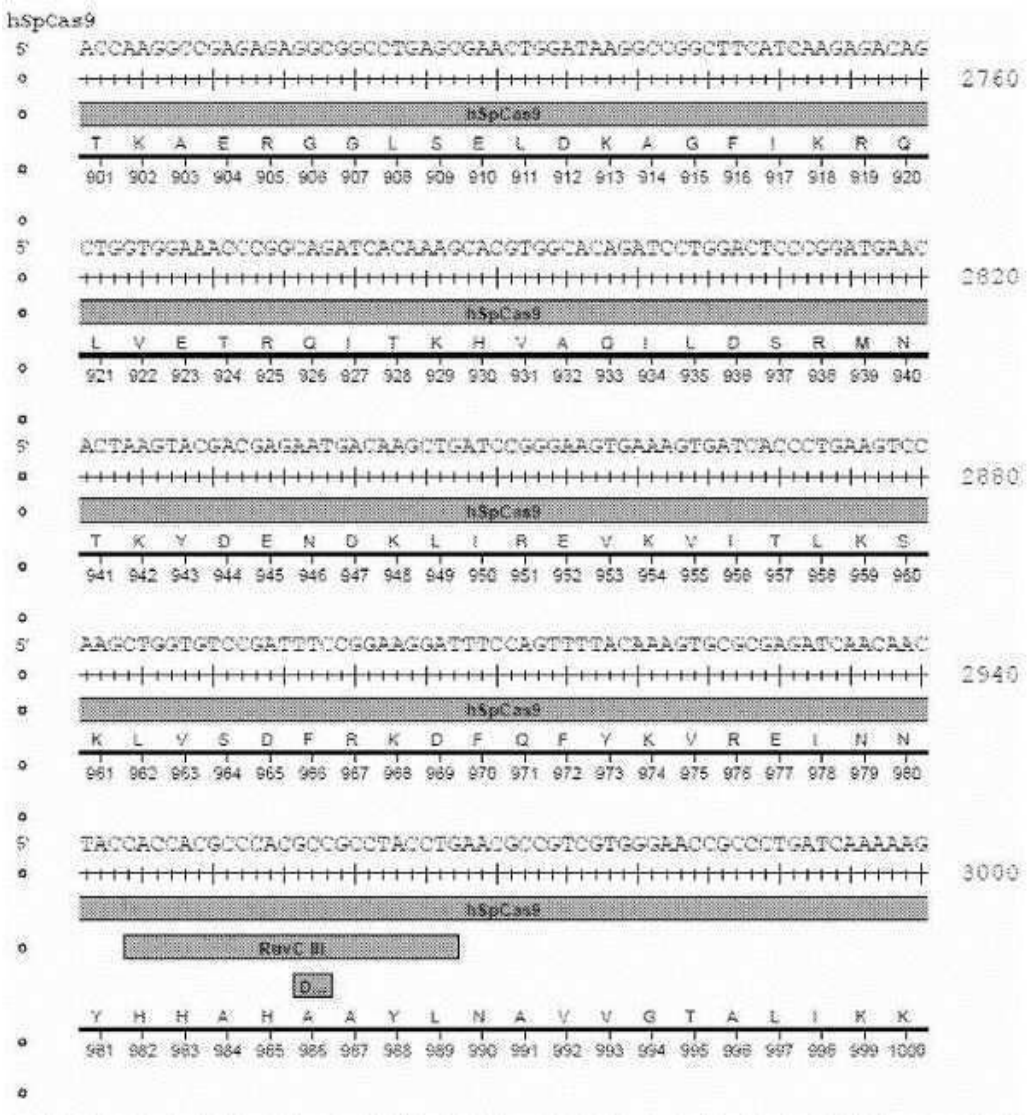
도면24g



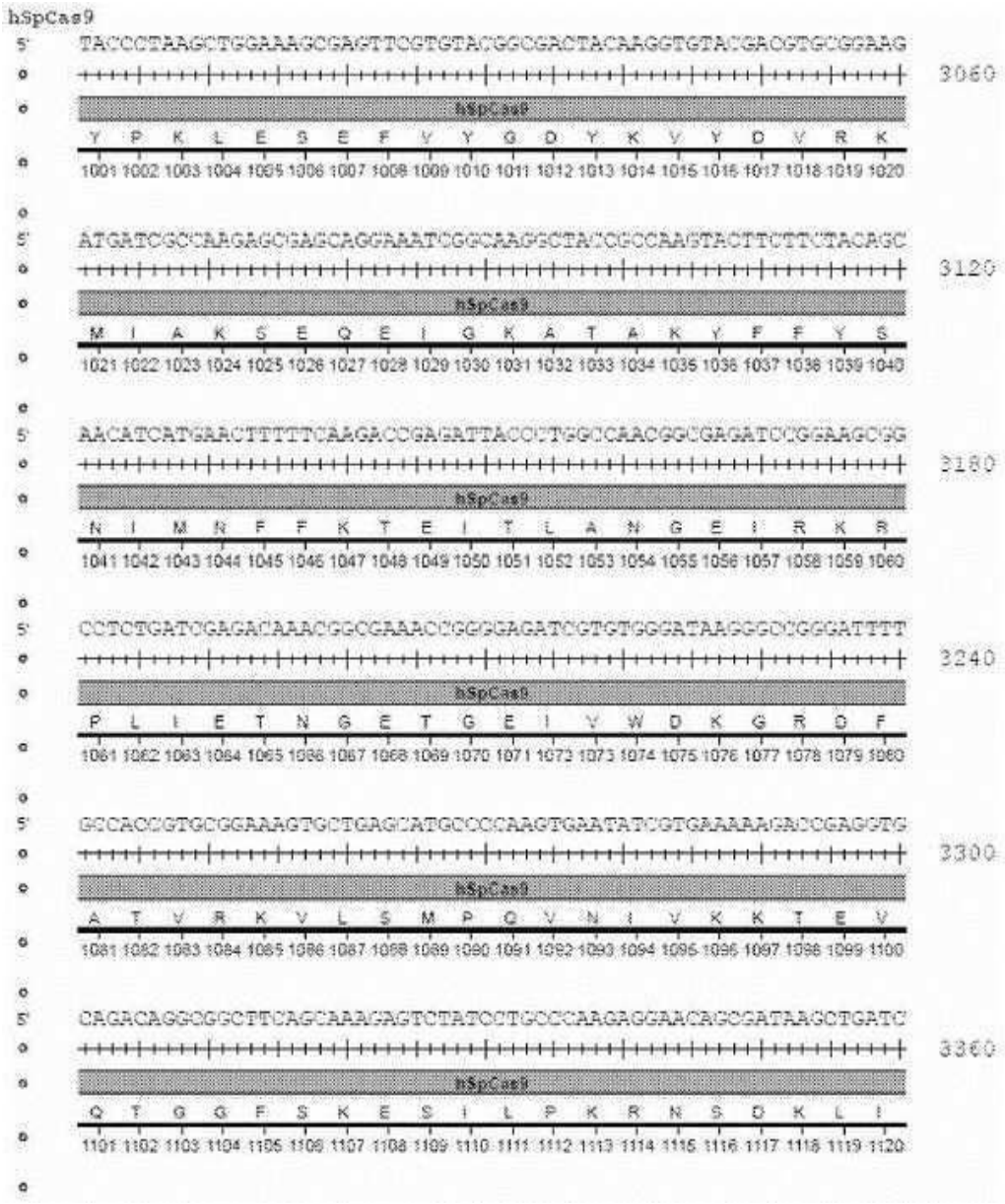
도면24h


hSpCas9
5'
GTGGAAAACACCCAGCTGCAGAACGAGAAGCTGTACCTGTACTACCTGCAGAATGGGCGG
2460
hSpCas9
V E N T Q L Q N E K L Y L Y Y L Q N G R
801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820
5'
GATATGTACGTGGACCAGGAACTGGACATCAACCGGCTGTCCGACTACGATGTGGACGCC
2520
hSpCas9
HNH
H...
D M Y V D Q E L D I N R L S D Y D V D A
821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840
5'
ATCGTGCCTCAGAGCTTTCTGAAGGACGACTCCATCGACGCCAAGGTGCTGACCAGAAGC
2580
hSpCas9
HNH
N...
I V P Q S F L K D D S I D A K V L T R S
841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860
5'
GACAAGGCCGGGGGCAAGAGCGACAACGTGCCCTCCGAAGAGGTCGTGAAGAAGATGAAG
2640
hSpCas9
HNH
N...
D K A R G K S D N V P S E E V V K K M K
861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880
5'
AACTACTGGCGGCAGCTGCTGAACGCCAAGCTGATTACCCAGAGAAAGTTCGACAATCTG
2700
hSpCas9
N Y W R Q L L N A K L I T Q R K F D N L
881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900

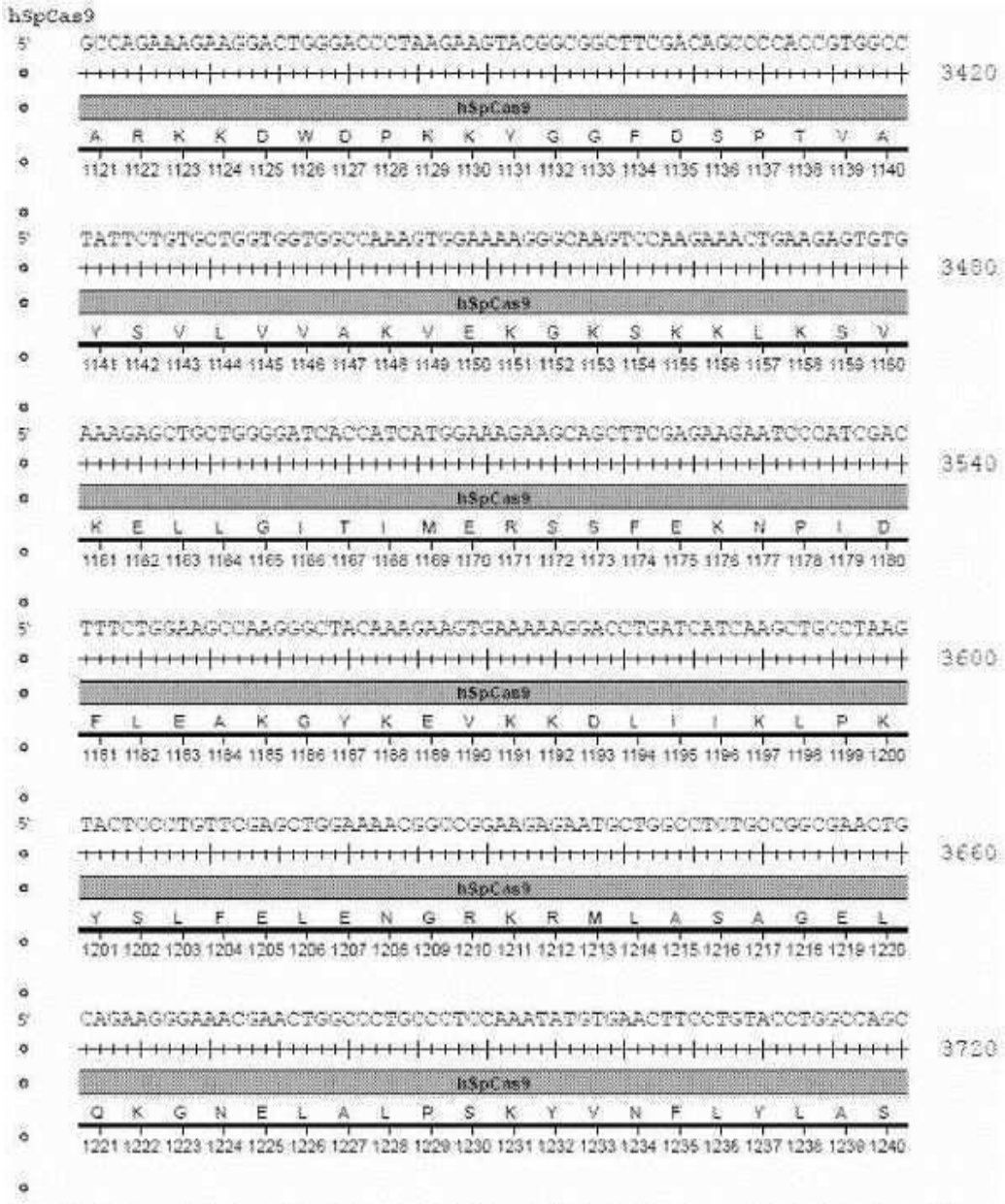
도면24i


hSpCas9
5'
ACCAAGGTCGAGAGAGGCGGCCTGAGCGAACTGGATAAGGCCGGCTTCATCAAGAGACAG
2760
hSpCas9
T K A E R G G L S E L D K A G F I K R G
901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920
5'
CTGGTGGAAACCCGGCAGATCACAAAGCACGTGGCACAGATCCTGGACTCCCGGATGAAC
2820
hSpCas9
L V E T R Q I T K H V A Q I L D S R M N
921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940
5'
ACTAAGTACGACGAGAATGACAAGCTGATCCGGGAAGTGAAAGTGATCACCCTGAAGTCC
2880
hSpCas9
T K Y D E N D K L I R E V K V I T L K S
941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960
5'
AAGCTGGTGTCCGATTTCCGGAAGGATTTCCAGTTTTACAAAGTGCGCGAGATCAACAAC
2940
hSpCas9
K L V S D F R K D F Q F Y K V R E I N N
961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980
5'
TACCACCACGCCCACGACGCGTACCTGAACGCCGTCGTGGGAACCGCCCTGATCAAAAAG
3000
hSpCas9
RuvC III
D
Y H H A H A A Y L N A V V G T A L I K K
981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000

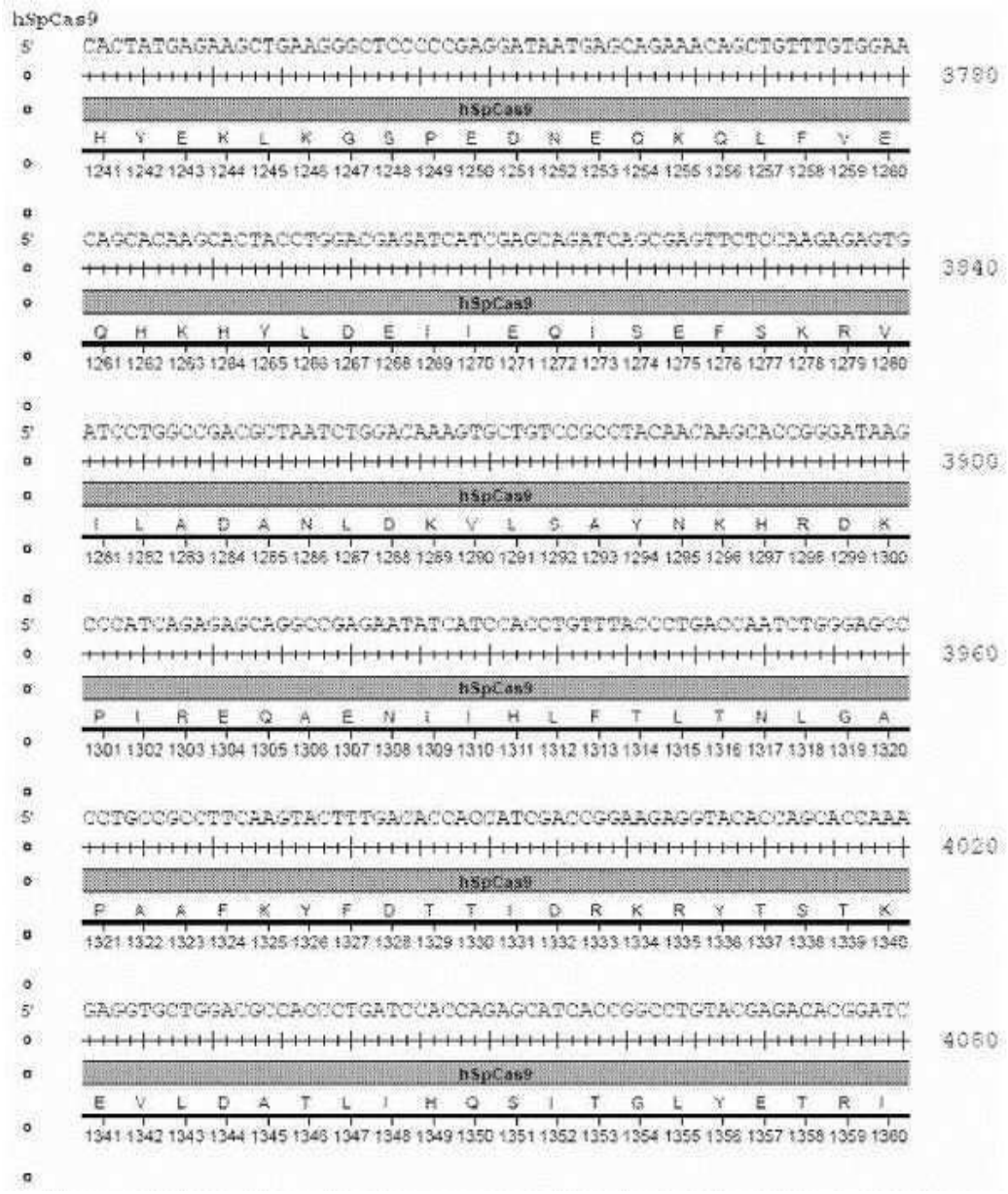
도면24j



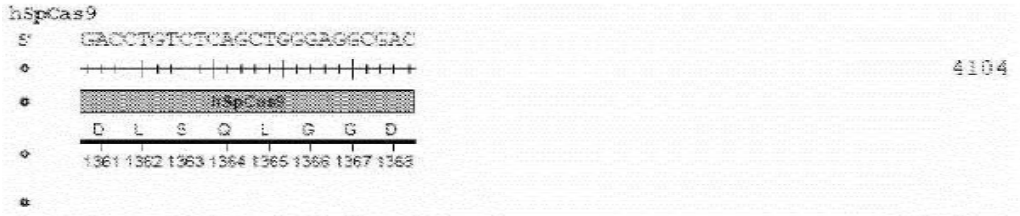
도면24k


hSpCas9
5'
GCCAGAAAGAAGGACTGGGACCCTAAGAAGTACGGCGGCTTCGACAGCCCCACCGTGGCC
3420
hSpCas9
A R K K D W D P K K Y G G F D S P T V A
1121 1122 1123 1124 1125 1126 1127 1128 1129 1130 1131 1132 1133 1134 1135 1136 1137 1138 1139 1140
5'
TATTCTGTGCTGGTGGTGGCCAAAGTGGAAAAGGGCAAGTCCAAGAAACTGAAGAGTGTG
3480
hSpCas9
Y S V L V V A K V E K G K S K K L K S V
1141 1142 1143 1144 1145 1146 1147 1148 1149 1150 1151 1152 1153 1154 1155 1156 1157 1158 1159 1160
5'
AAAGAGCTGCTGGGGATCACCATCATGGAAAGAAGCAGCTTCGAGAAGAATCCCATCGAC
3540
hSpCas9
K E L L G I T I M E R S S F E K N P I D
1161 1162 1163 1164 1165 1166 1167 1168 1169 1170 1171 1172 1173 1174 1175 1176 1177 1178 1179 1180
5'
TTTCTGGAAGCCAAGGGCTACAAAGAAGTGAAAAAGGACCTGATCATCAAGCTGCCTAAG
3600
hSpCas9
F L E A K G Y K E V K K D L I I K L P K
1181 1182 1183 1184 1185 1186 1187 1188 1189 1190 1191 1192 1193 1194 1195 1196 1197 1198 1199 1200
5'
TACTCCCTGTTCGAGCTGGAAAACGGCCGGAAGAGAATGCTGGCCTCTGCCGGCGAACTG
3660
hSpCas9
Y S L F E L E N G R K R M L A S A G E L
1201 1202 1203 1204 1205 1206 1207 1208 1209 1210 1211 1212 1213 1214 1215 1216 1217 1218 1219 1220
5'
CAGAAGGGAAACGAACTGGCCCTGCCCTCCAAATATGTGAACTTCCTGTACCTGGCCAGC
3720
hSpCas9
Q K G N E L A L P S K Y V N F L Y L A S
1221 1222 1223 1224 1225 1226 1227 1228 1229 1230 1231 1232 1233 1234 1235 1236 1237 1238 1239 1240

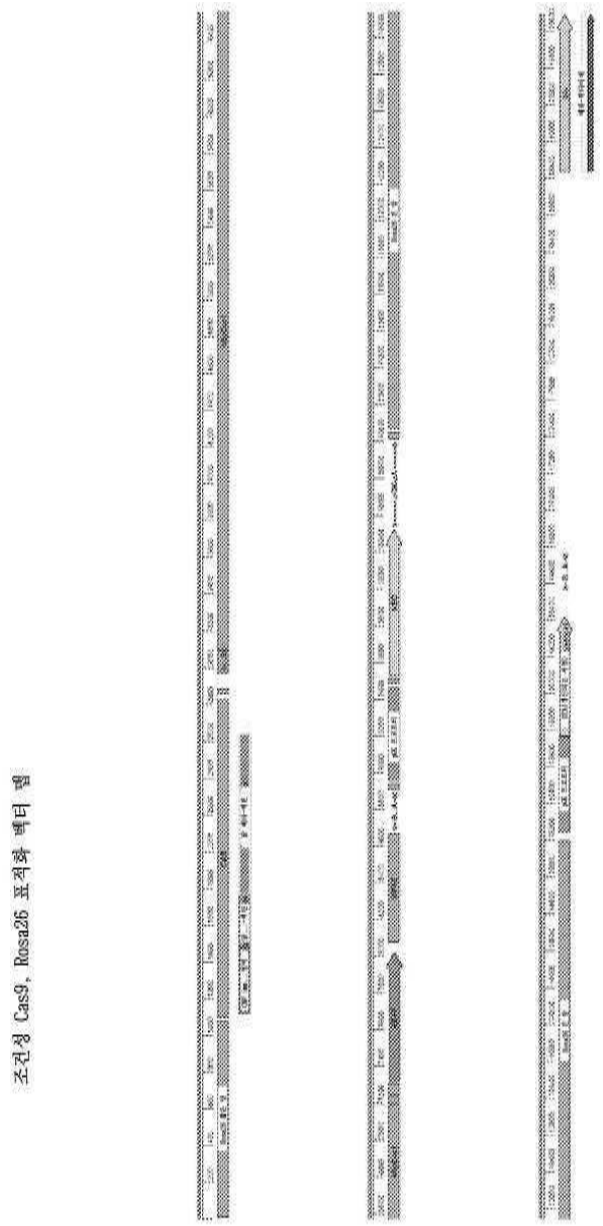
도면241


hSpCas9
5'
CACTATGAGAAGCTGAAGGGCTCCCCCGAGGATAATGAGCAGAAACAGCTGTTTGTGGAA
3780
hSpCas9
H Y E K L K G S P E D N E Q K Q L F V E
1241 1242 1243 1244 1245 1246 1247 1248 1249 1250 1251 1252 1253 1254 1255 1256 1257 1258 1259 1260
5'
CAGCACAAGCACTACCTGGACGAGATCATCGAGCAGATCAGCGAGTTCTCCAAGAGAGTG
3840
hSpCas9
Q H K H Y L D E I I E Q I S E F S K R V
1261 1262 1263 1264 1265 1266 1267 1268 1269 1270 1271 1272 1273 1274 1275 1276 1277 1278 1279 1280
5'
ATCCTGGCCGACGCTAATCTGGACAAAGTGCTGTCCGCCTACAACAAGCACCGGGATAAG
3900
hSpCas9
I L A D A N L D K V L S A Y N K H R D K
1281 1282 1283 1284 1285 1286 1287 1288 1289 1290 1291 1292 1293 1294 1295 1296 1297 1298 1299 1300
5'
CCCATCAGAGAGCAGGCCGAGAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCC
3960
hSpCas9
P I R E Q A E N I I H L F T L T N L G A
1301 1302 1303 1304 1305 1306 1307 1308 1309 1310 1311 1312 1313 1314 1315 1316 1317 1318 1319 1320
5'
CCTGCCGCCTTCAAGTACTTTGACACCACCATCGACCGGAAGAGGTACACCAGCACCAAA
4020
hSpCas9
P A A F K Y F D T T I D R K R Y T S T K
1321 1322 1323 1324 1325 1326 1327 1328 1329 1330 1331 1332 1333 1334 1335 1336 1337 1338 1339 1340
5'
GAGGTGCTGGACGCCACCCTGATCCACCAGAGCATCACCGGCCTGTACGAGACACGGATC
4080
hSpCas9
E V L D A T L I H Q S I T G L Y E T R I
1341 1342 1343 1344 1345 1346 1347 1348 1349 1350 1351 1352 1353 1354 1355 1356 1357 1358 1359 1360

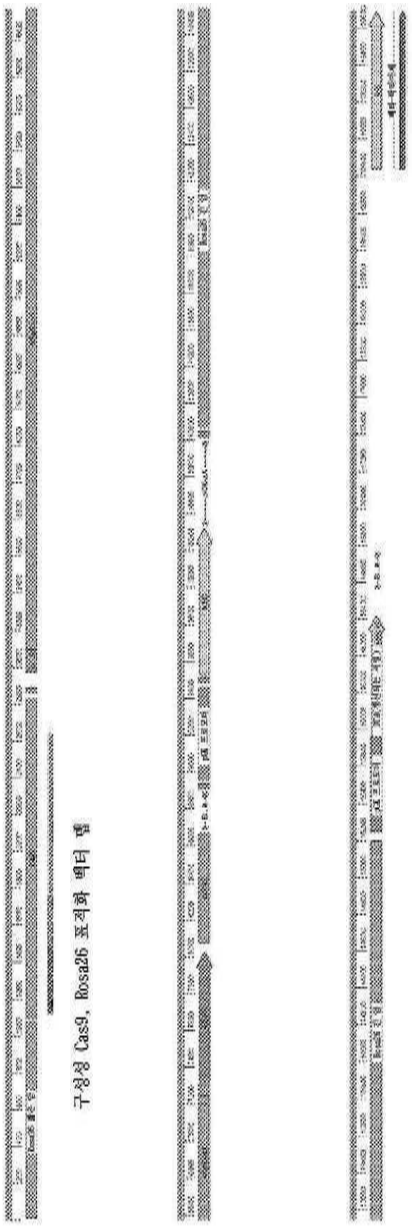
도면24m


hSpCas9
5'
GACCTGTCTCAGCTGGGAGGCGAC
4104
hSpCas9
D L S Q L G G D
1361 1362 1363 1364 1365 1366 1367 1368

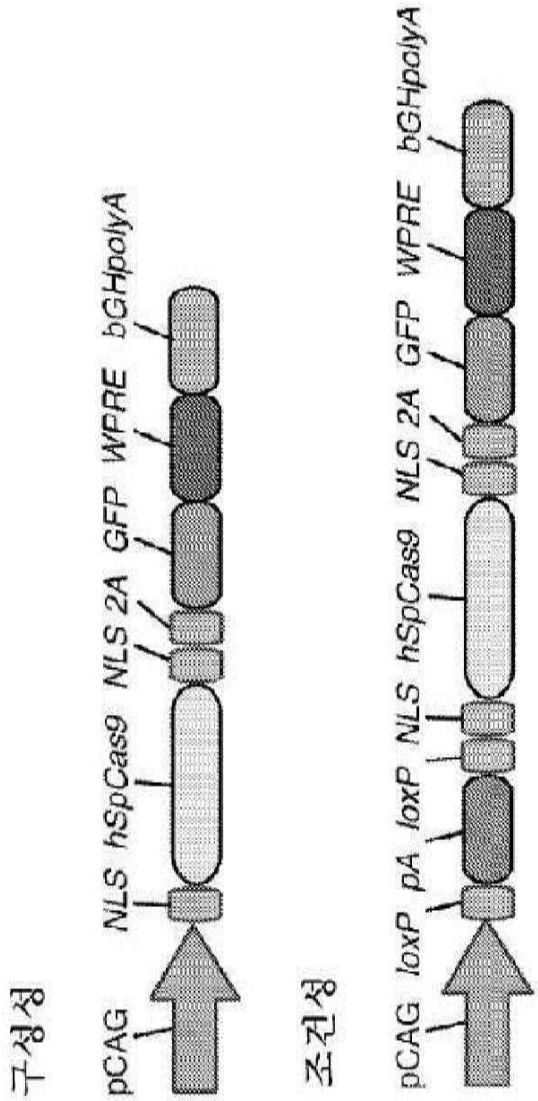
도면25a



도면25b

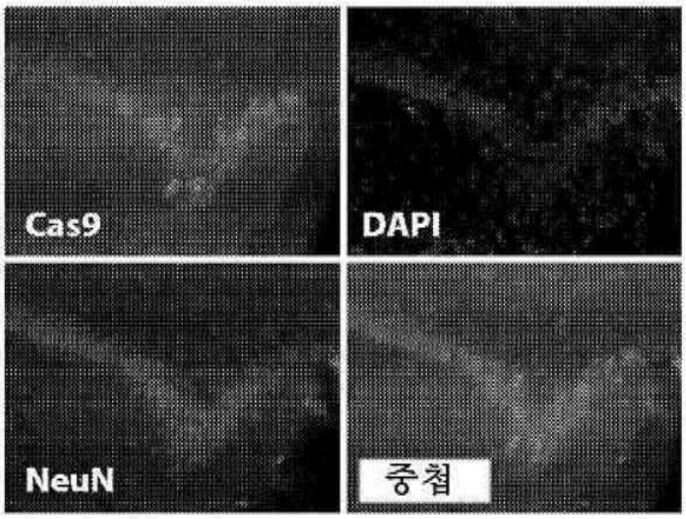


도면26

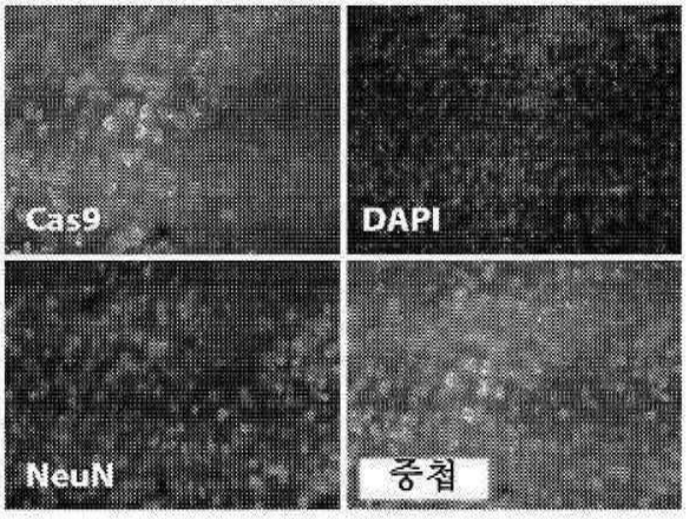

구성성
pCAG
NLS
hSpCas9
NLS
2A
GFP
WPRE
bGHpolyA
조건성
pCAG
loxP
pA
loxP
NLS
hSpCas9
NLS
2A
GFP
WPRE
bGHpolyA

도면27

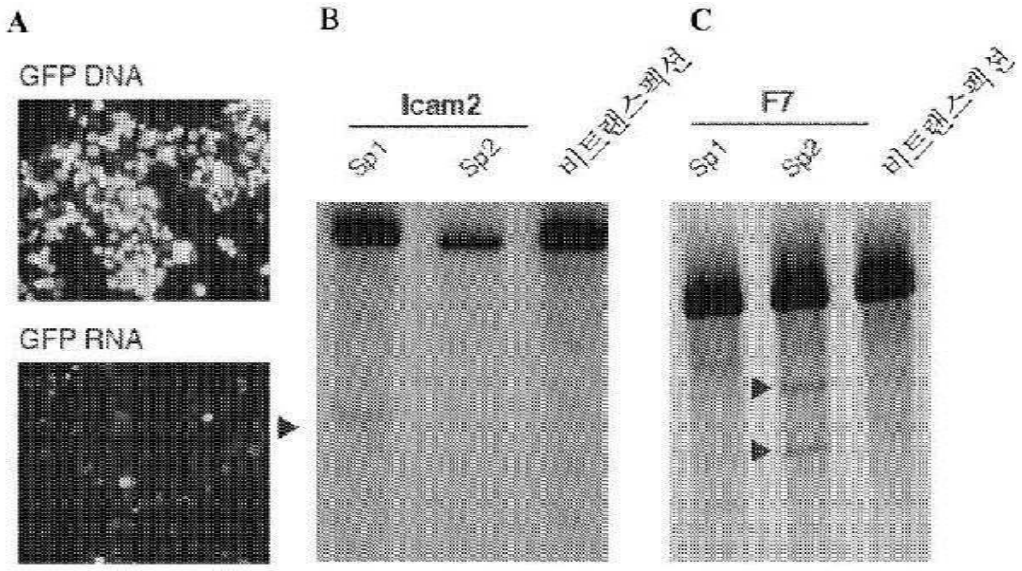
마우스 해마에서의 Cas9 발현 (AAV)



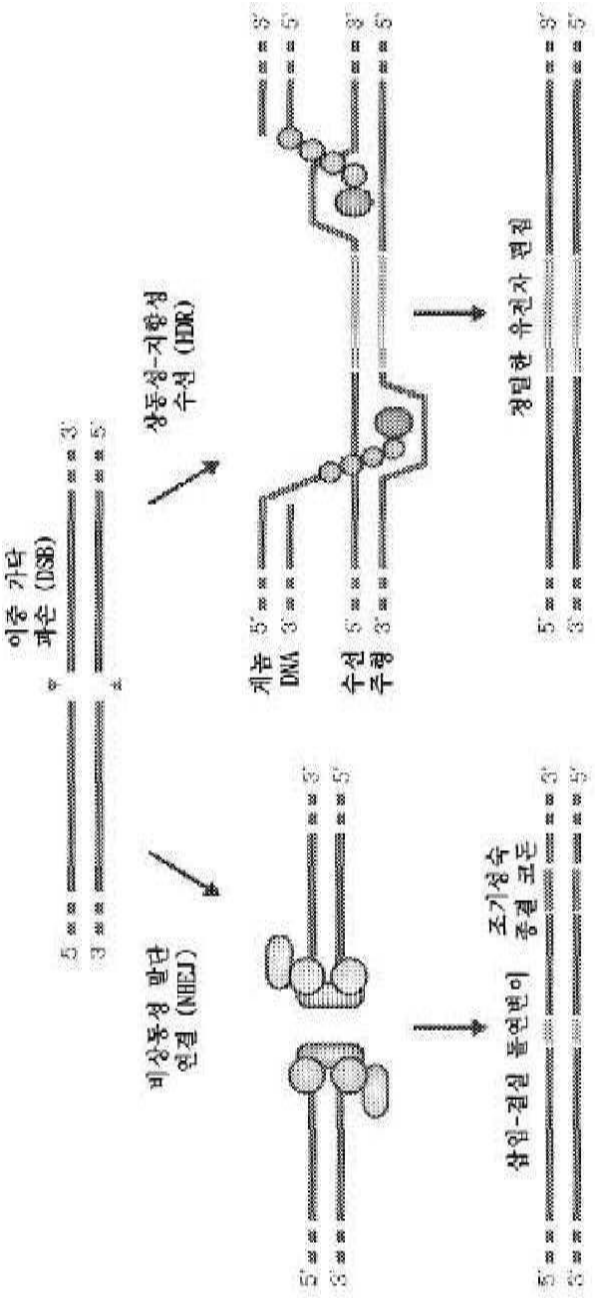
마우스 피질에서의 Cas9 발현 (AAV)



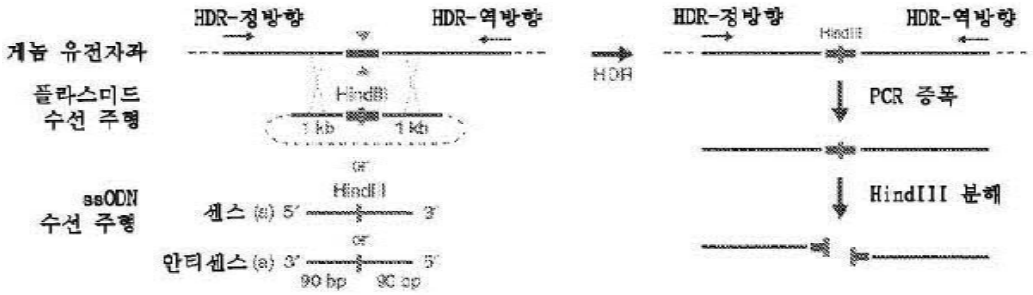
도면28



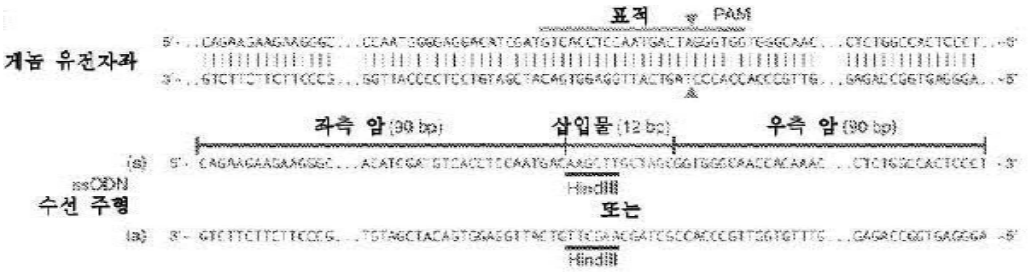
도면29


이중 가닥
파손 (DSB)
상동성-지향성
수선 (HDR)
게놈
DNA
수선
주형
정밀한 유전자 편집
비상동성 말단
연결 (NHEJ)
조기성숙
종결 코돈
삽입-결실 돌연변이

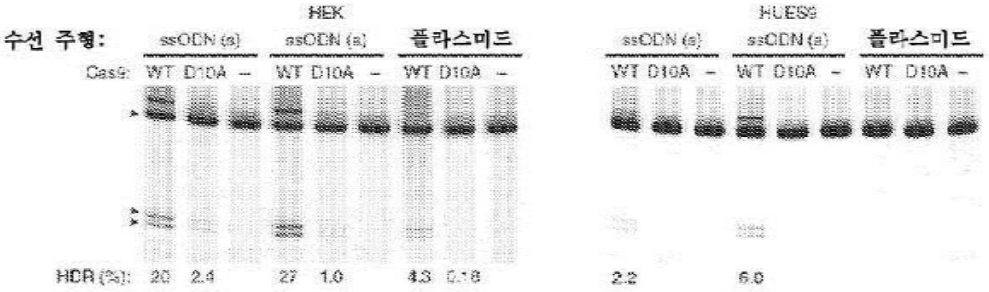
도면30a


게놈 유전자좌
플라스미드
수선 주형
ssODN
수선 주형
HDR-정방향
HDR-역방향
HindIII
1 kb
1 kb
or
HindI
센스 (s) 5′
3′
or
안티센스 (a) 3′
5′
90 bp
90 bp
HDR
HDR-정방향
HDR-역방향
HindIII
PCR 증폭
HindIII 분해

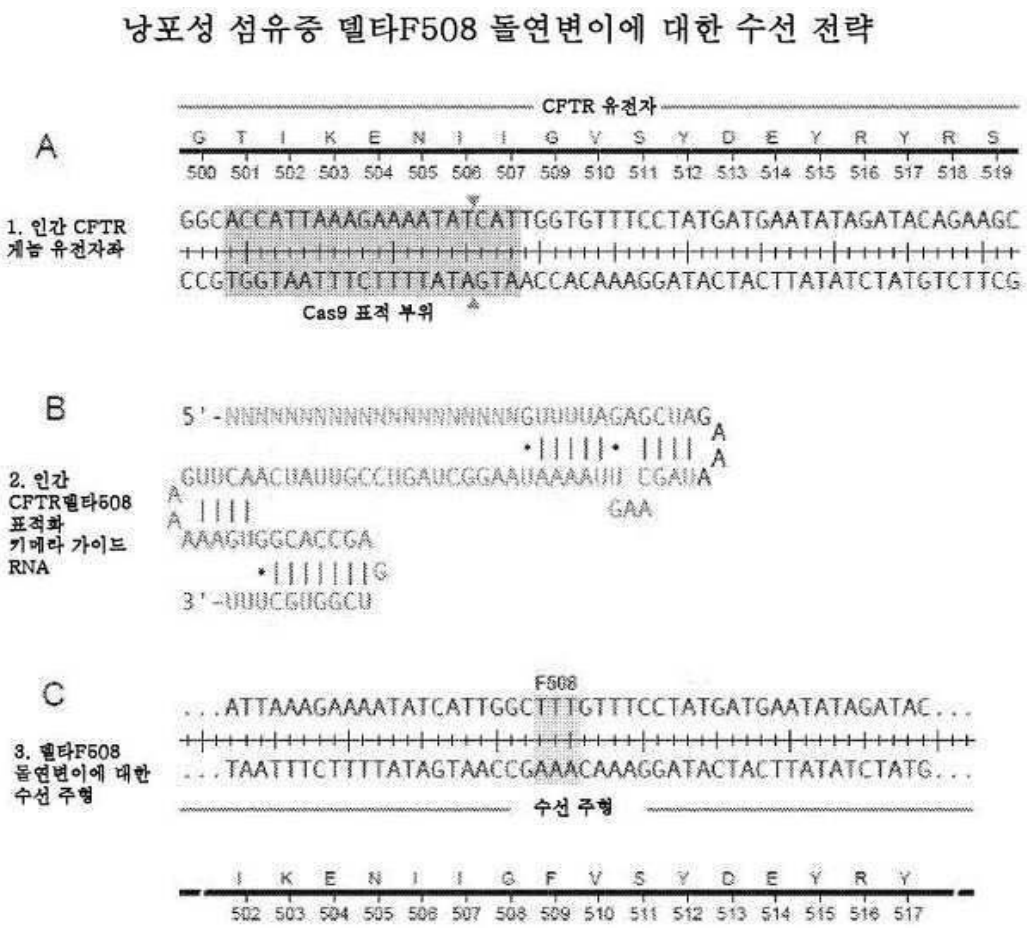
도면30b



도면30c



도면31



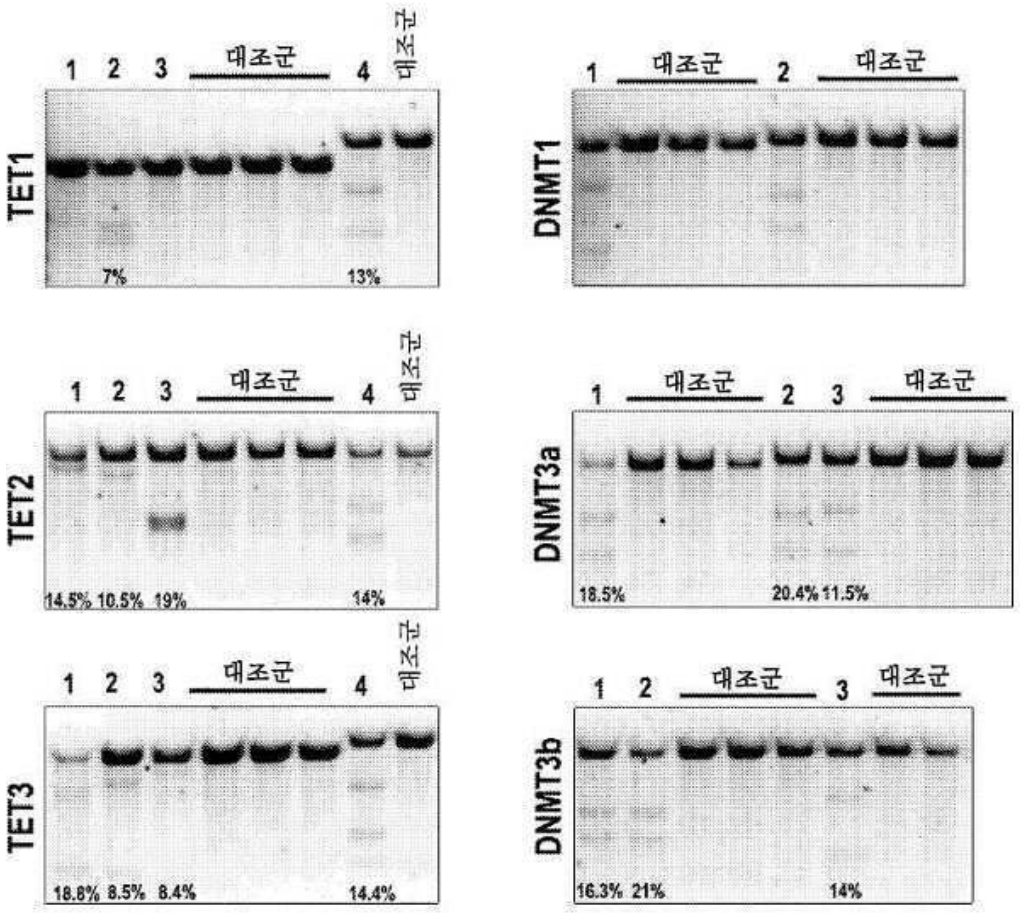
도면32a



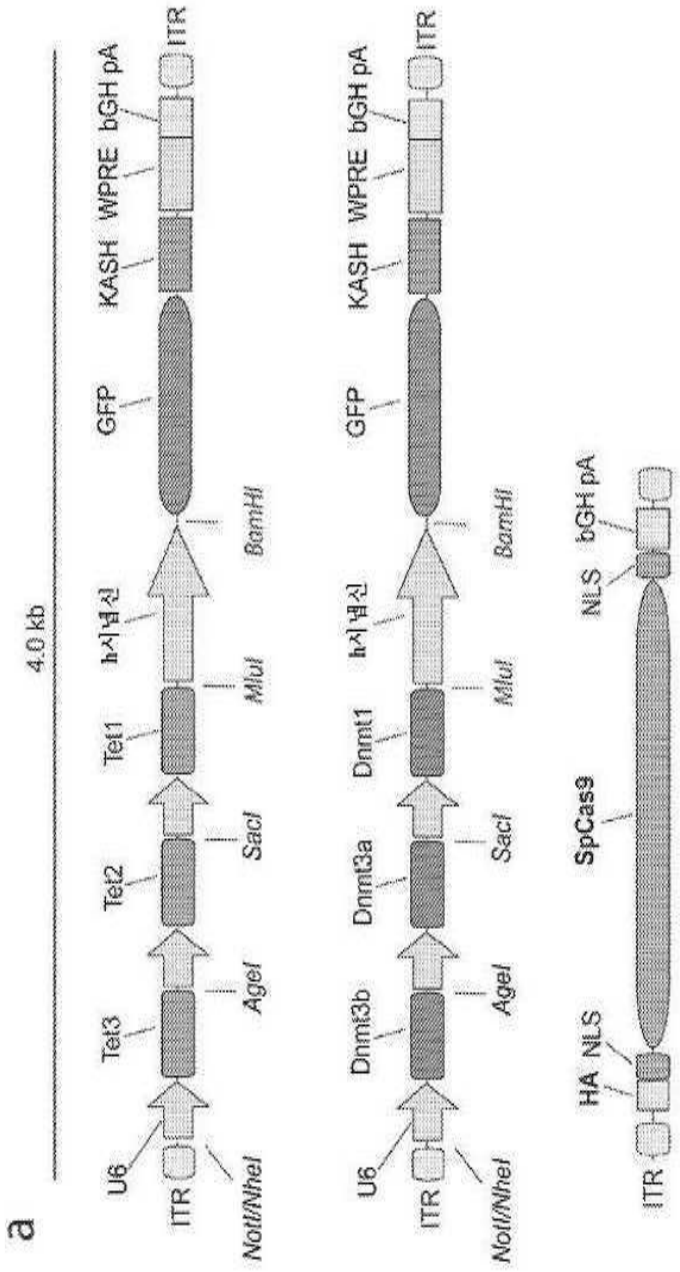
도면32b



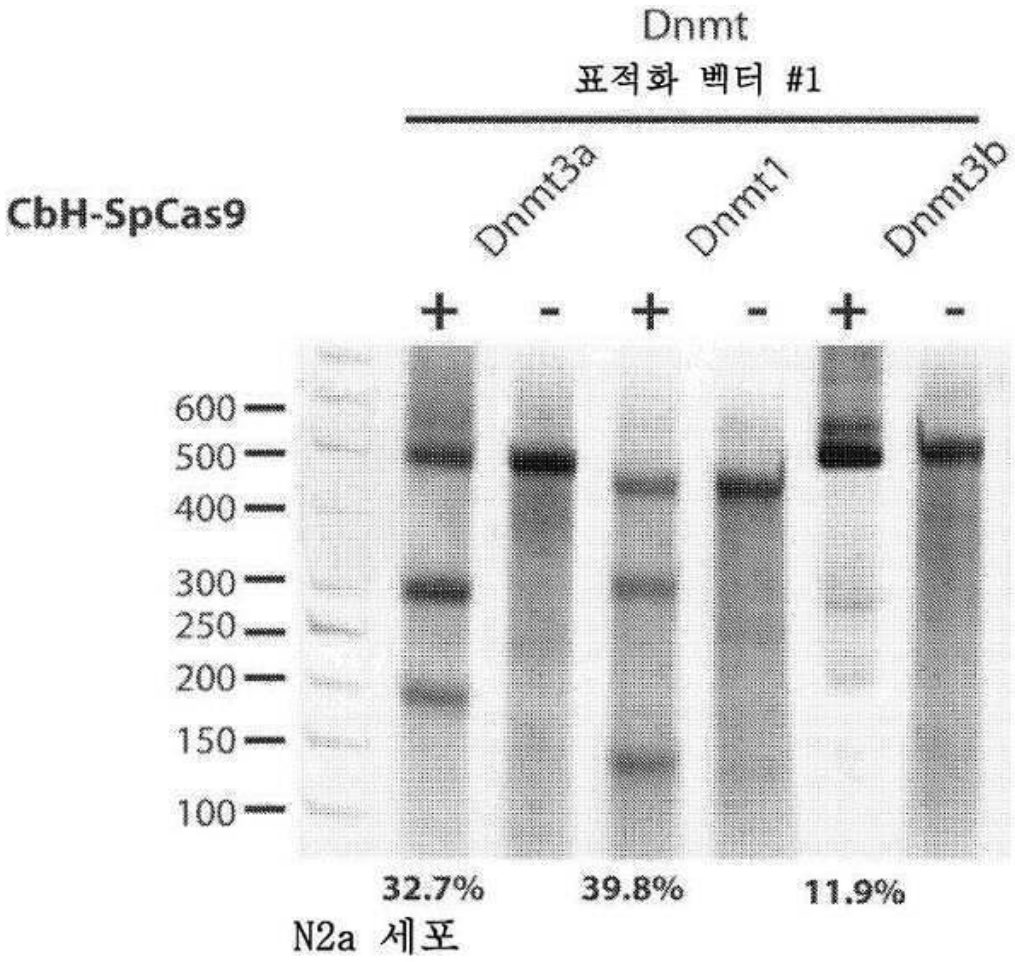
도면33



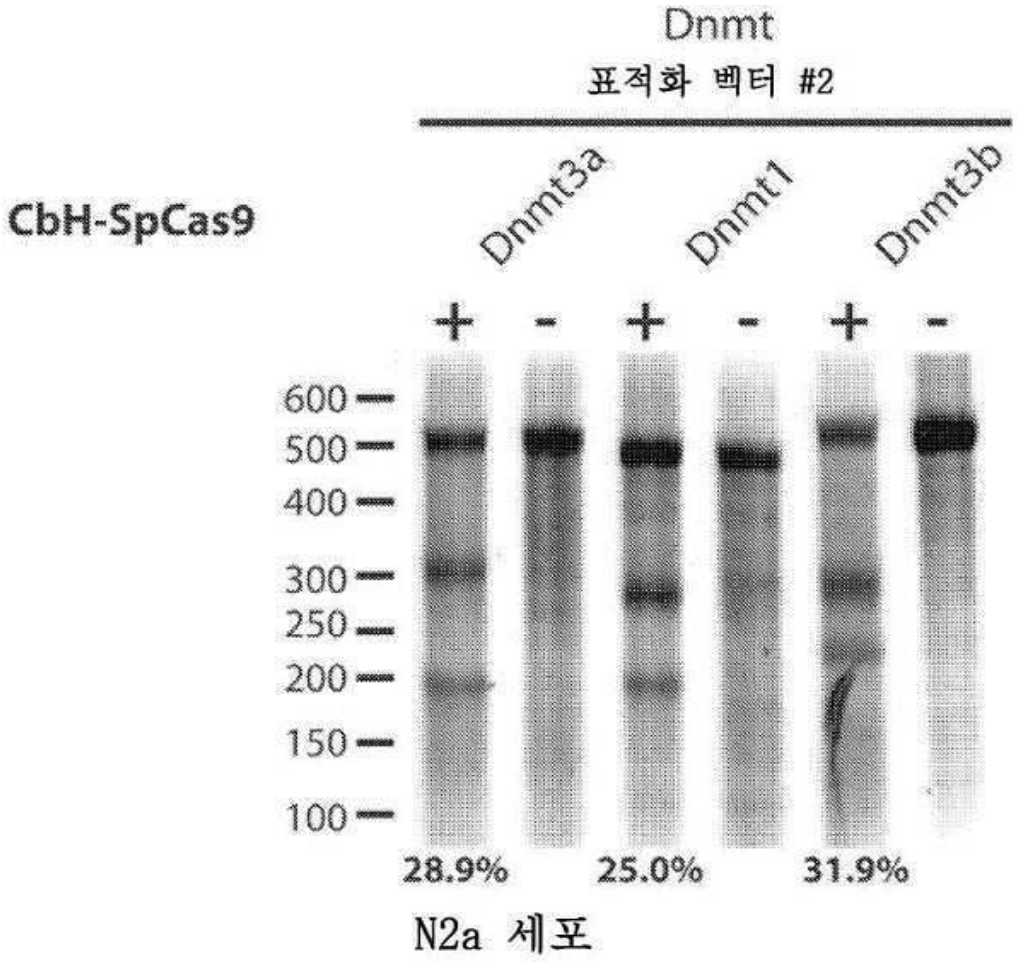
도면34


a
4.0 kb
U6
Tet3
Tet2
Tet1
hU6신
GFP
KASH
WPRE
bGH pA
ITR
ITR
NotI/NheI
AgeI
SacI
MluI
BamHI
U6
Dnmt3b
Dnmt3a
Dnmt1
hU6신
GFP
KASH
WPRE
bGH pA
ITR
ITR
NotI/NheI
AgeI
SacI
MluI
BamHI
HA
NLS
SpCas9
NLS
bGH pA
ITR

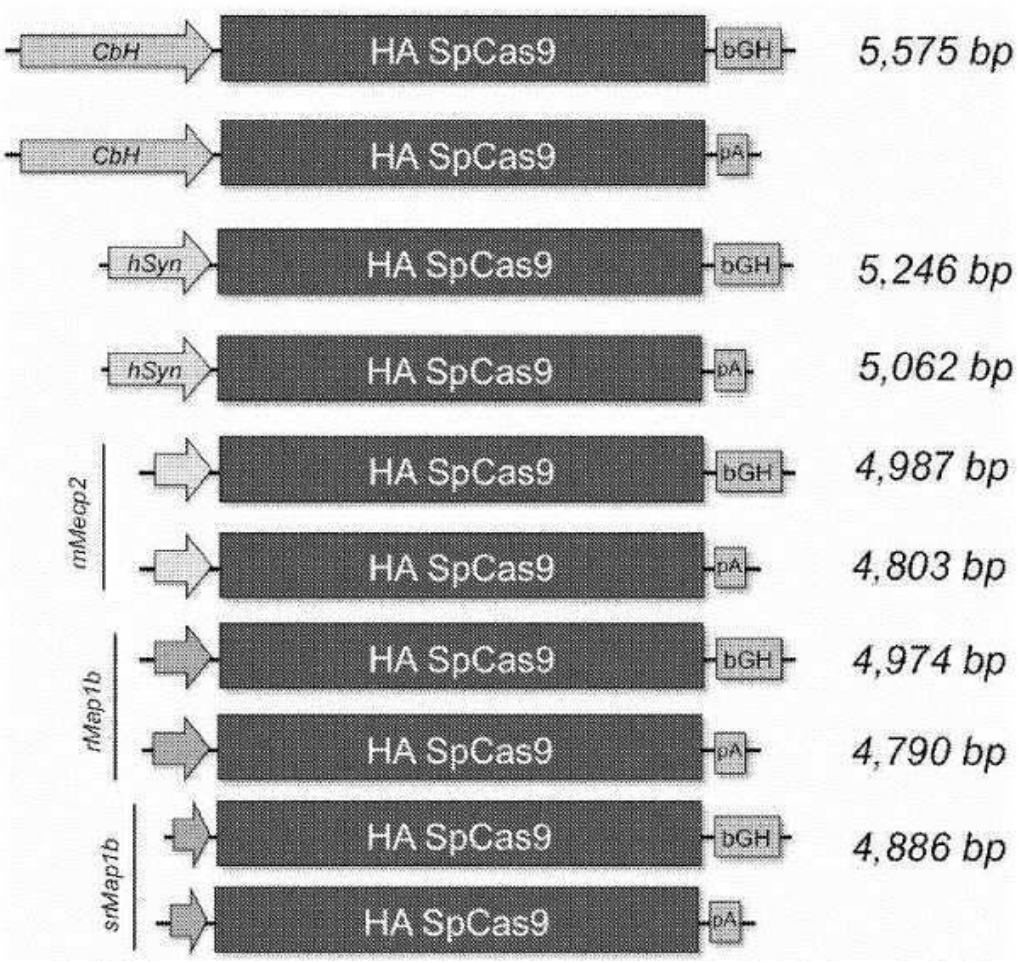
도면35


Dnmt
표적화 벡터 #1
CbH-SpCas9
Dnmt3a
Dnmt1
Dnmt3b
+
-
+
-
+
-
600
500
400
300
250
200
150
100
32.7%
39.8%
11.9%
N2a 세포

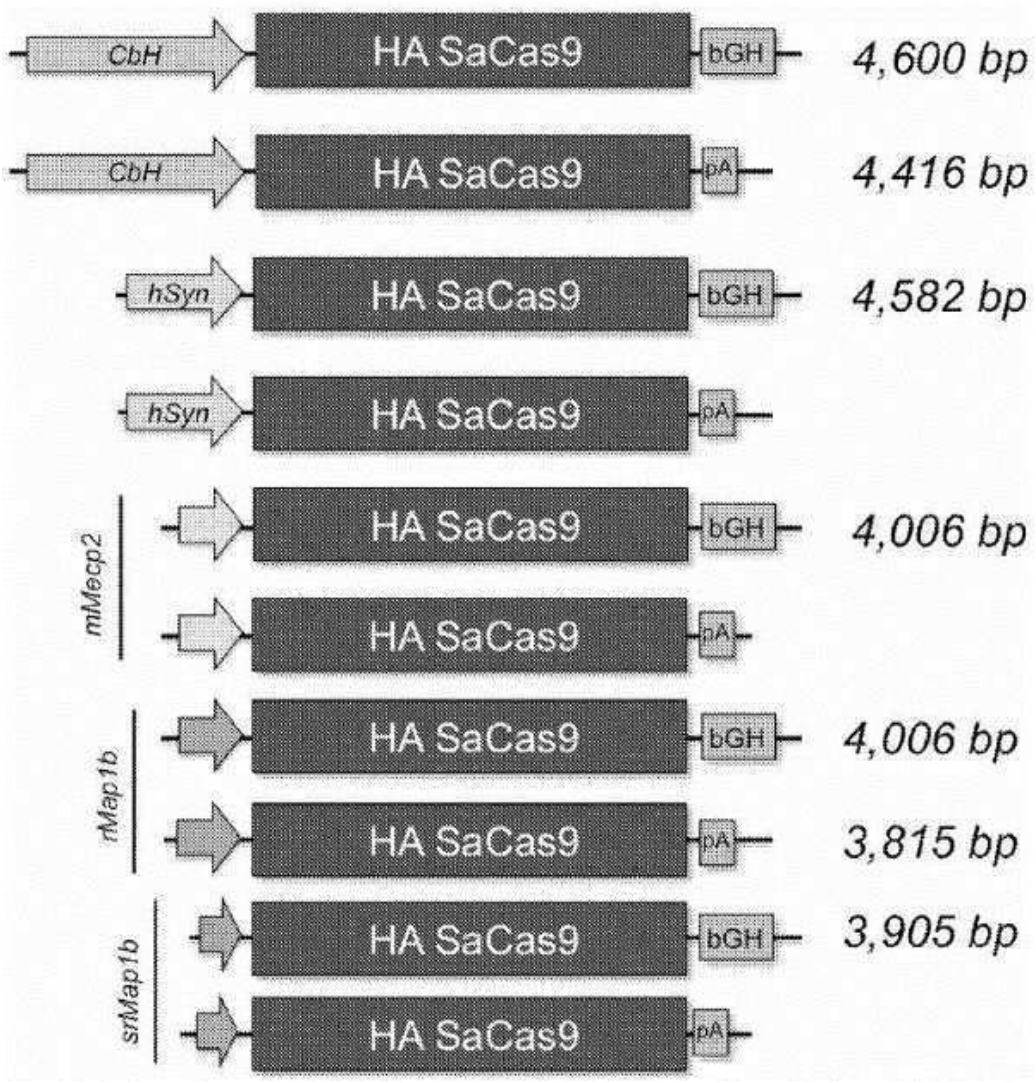
도면36


Dnmt
표적화 벡터 #2
CbH-SpCas9
Dnmt3a
Dnmt1
Dnmt3b
+ - + - + -
600
500
400
300
250
200
150
100
28.9%
25.0%
31.9%
N2a 세포

도면37

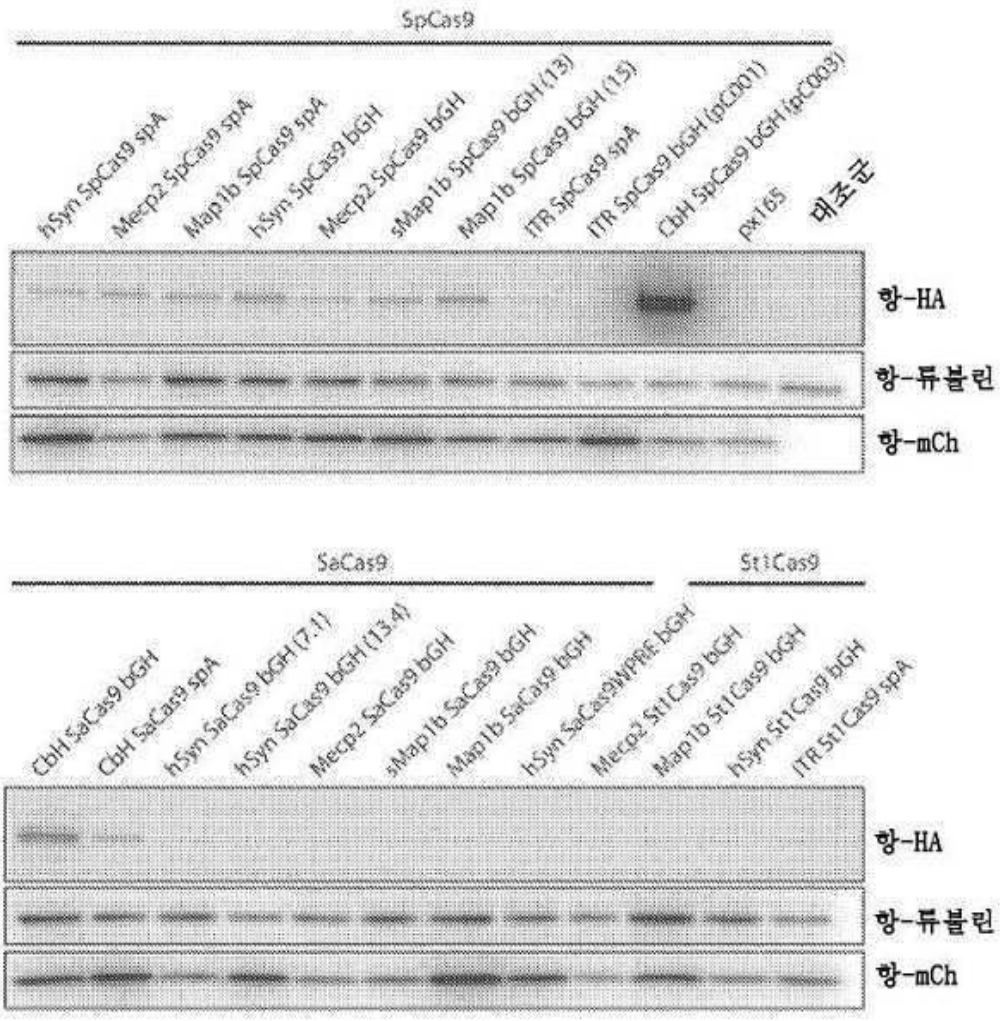

CbH
HA SpCas9
bGH
5,575 bp
CbH
HA SpCas9
pA
hSyn
HA SpCas9
bGH
5,246 bp
hSyn
HA SpCas9
pA
5,062 bp
rMecp2
HA SpCas9
bGH
4,987 bp
HA SpCas9
pA
4,803 bp
rMap1b
HA SpCas9
bGH
4,974 bp
HA SpCas9
pA
4,790 bp
srMap1b
HA SpCas9
bGH
4,886 bp
HA SpCas9
pA

도면38

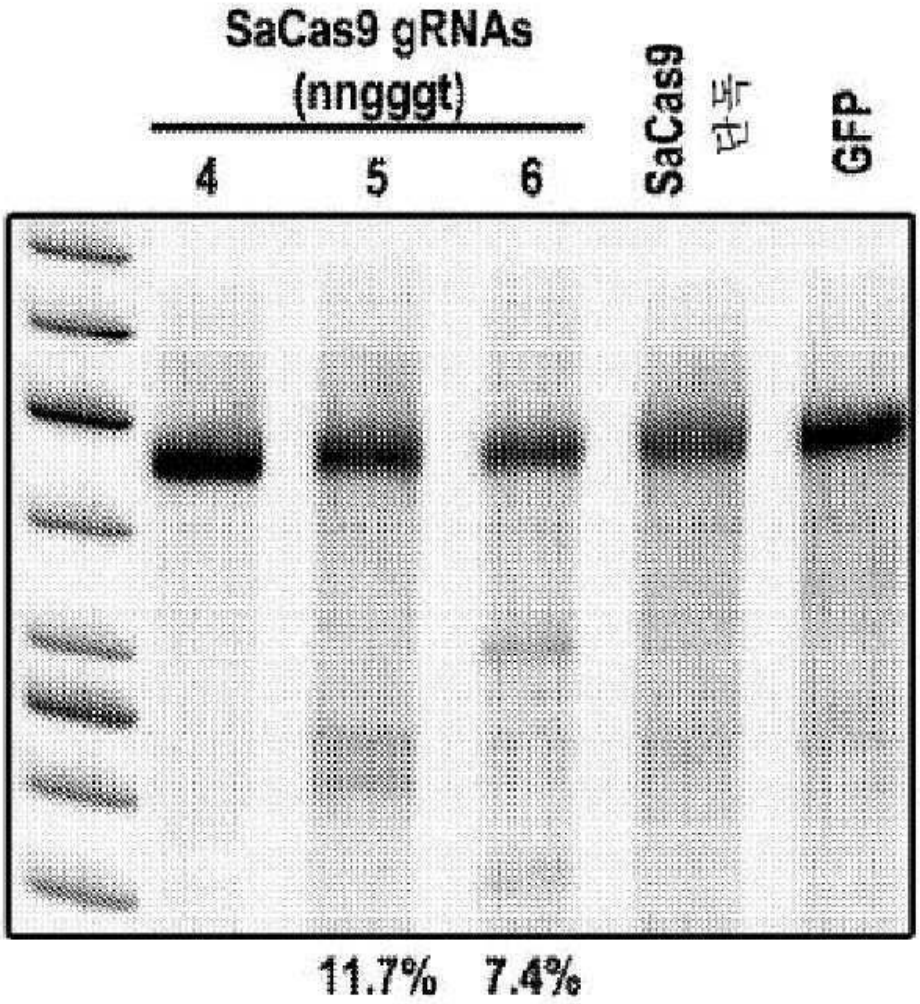

CbH
HA SaCas9
bGH
4,600 bp
CbH
HA SaCas9
pA
4,416 bp
hSyn
HA SaCas9
bGH
4,582 bp
hSyn
HA SaCas9
pA
mMecp2
HA SaCas9
bGH
4,006 bp
HA SaCas9
pA
rMap1b
HA SaCas9
bGH
4,006 bp
HA SaCas9
pA
3,815 bp
smMap1b
HA SaCas9
bGH
3,905 bp
HA SaCas9
pA

도면39

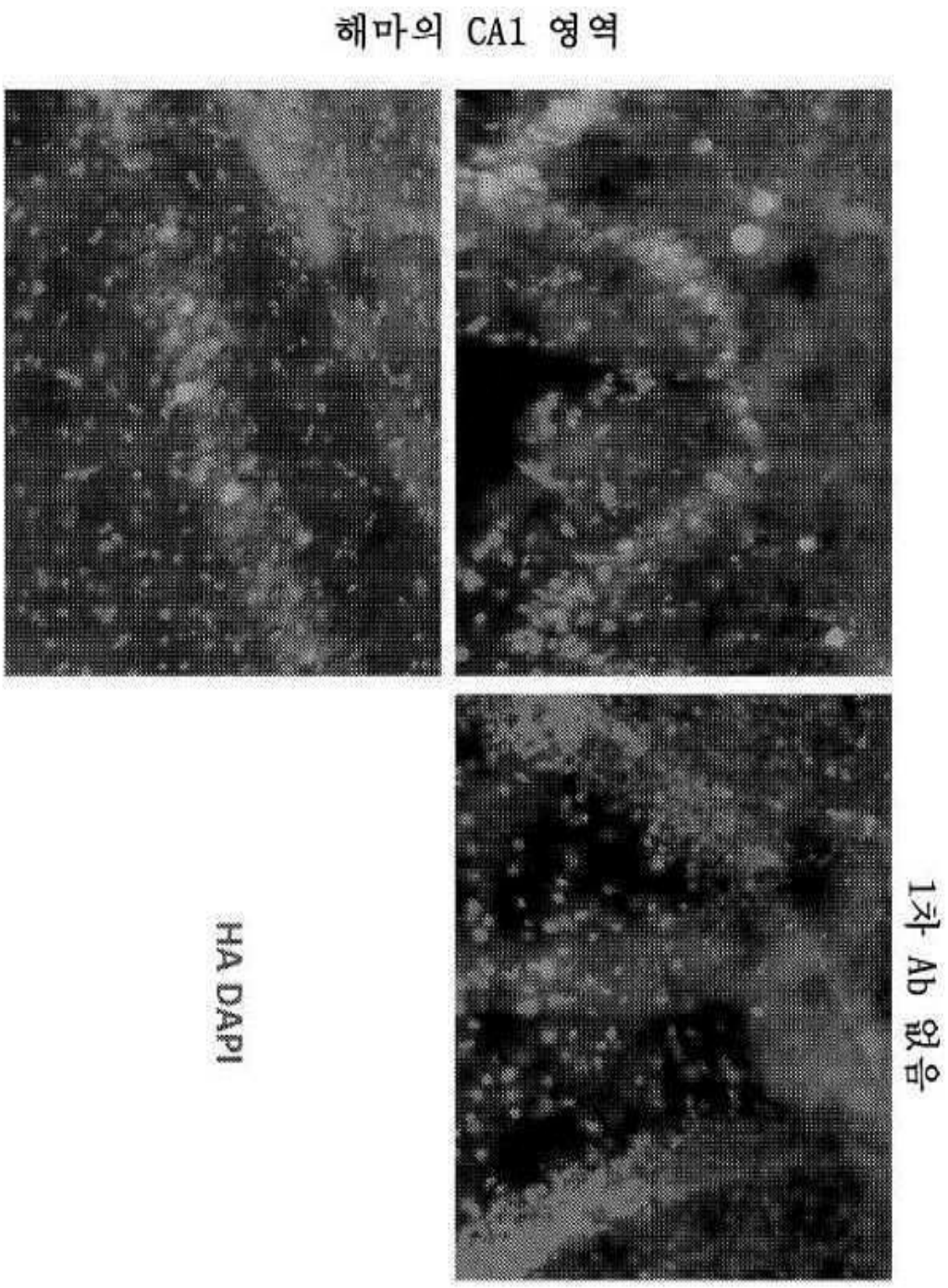
N2a 세포에서의 SpCas9 및 SaCas9의 발현



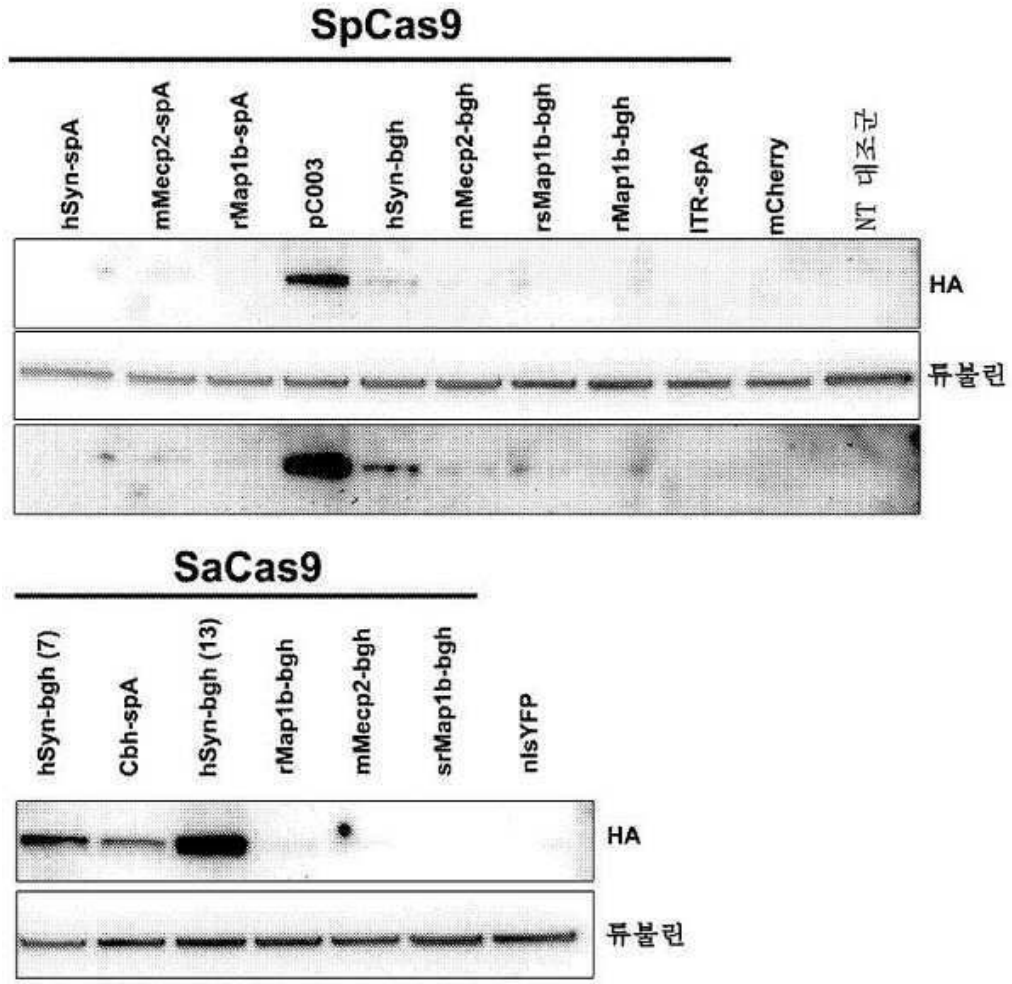
도면40



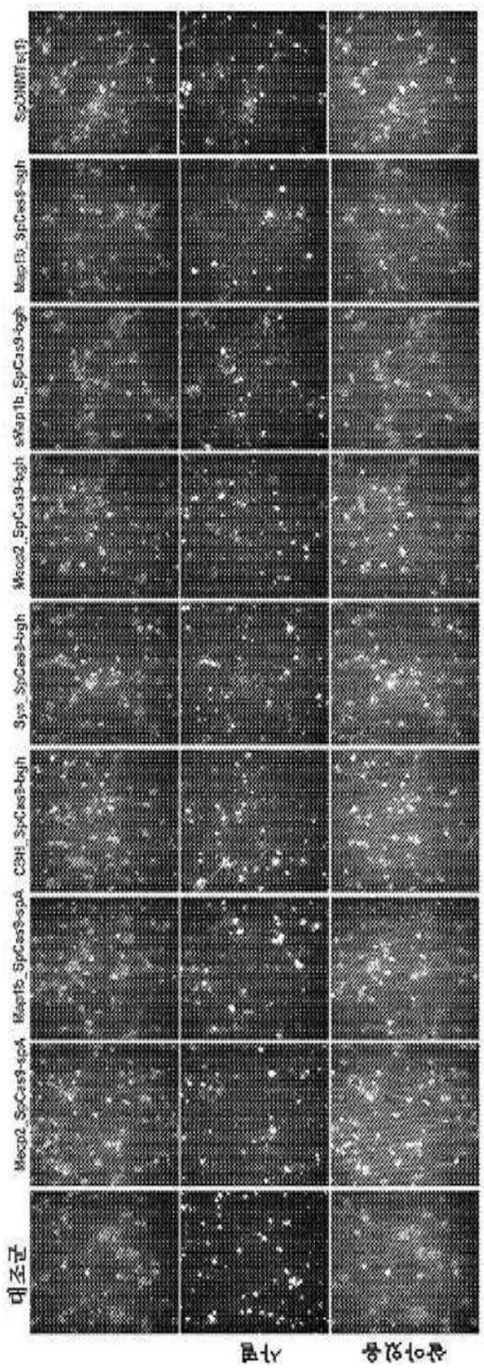
도면41



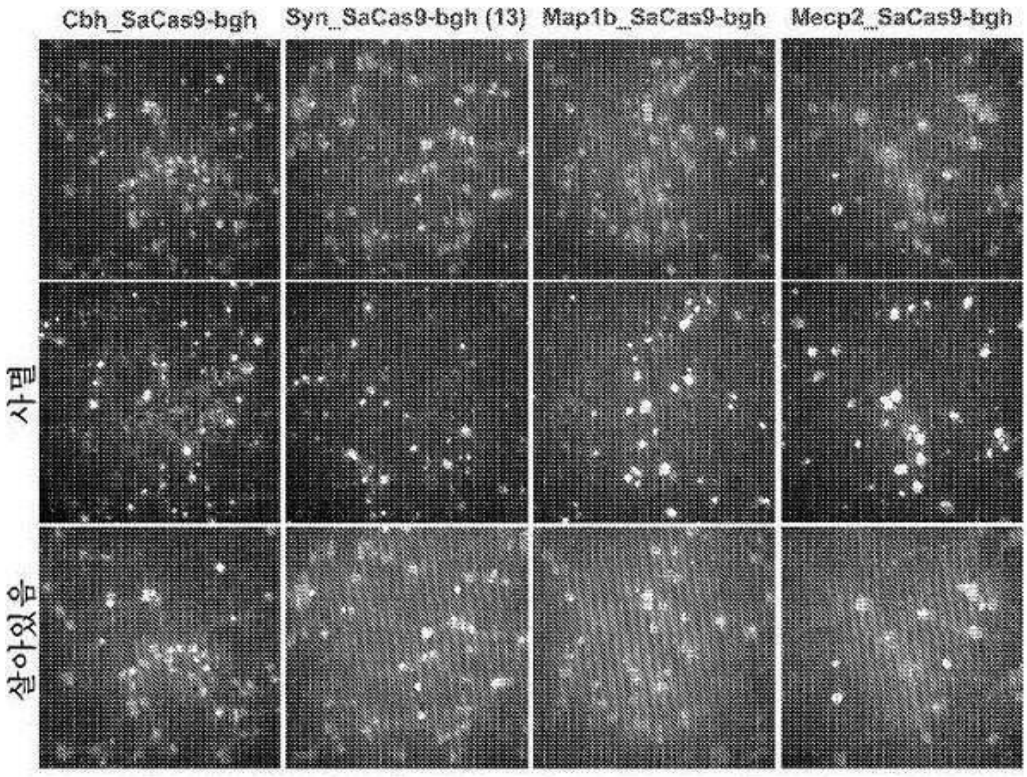
도면42


SpCas9
hSyn-spA
mMecp2-spA
rMap1b-spA
pC003
hSyn-bgh
mMecp2-bgh
rsMap1b-bgh
rMap1b-bgh
ITR-spA
mCherry
NT 대조군
HA
튜불린
SaCas9
hSyn-bgh (7)
Cbh-spA
hSyn-bgh (13)
rMap1b-bgh
mMecp2-bgh
srMap1b-bgh
nlsYFP
HA
튜불린

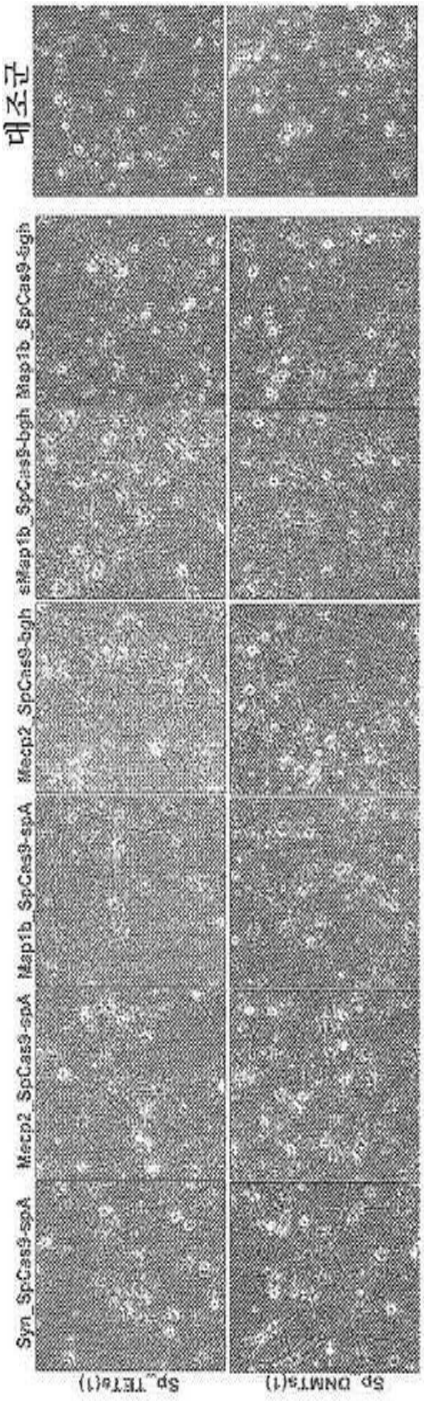
도면43


대조군
Mecp2_SpCas9-spA
Mapt5_SpCas9-spA
CBH_SpCas9-bgh
Syn_SpCas9-bgh
Mecp2_SpCas9-bgh
sMap1b_SpCas9-bgh
Mapt5_SpCas9-bgh
사멸
살아있음

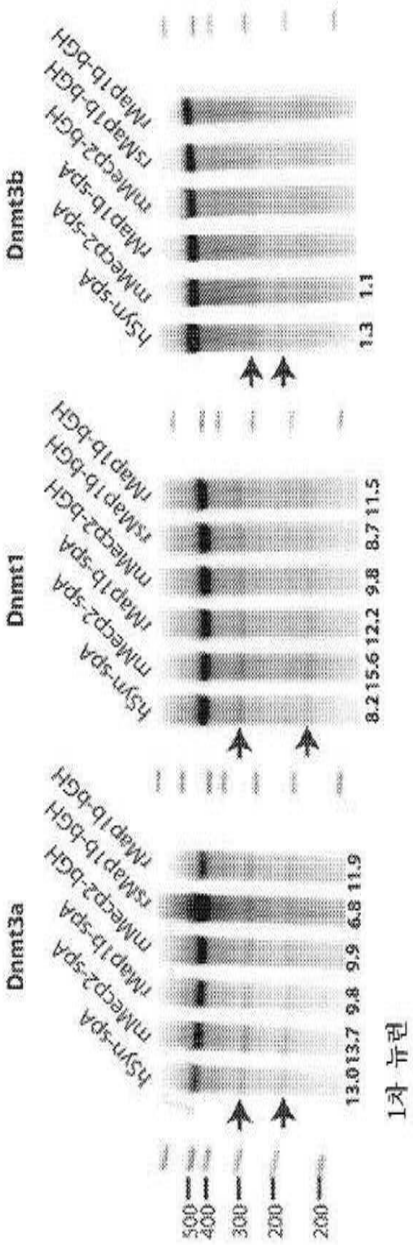
도면44


Cbh_SaCas9-bgh
Syn_SaCas9-bgh (13)
Map1b_SaCas9-bgh
Mecp2_SaCas9-bgh
사멸
살아있음

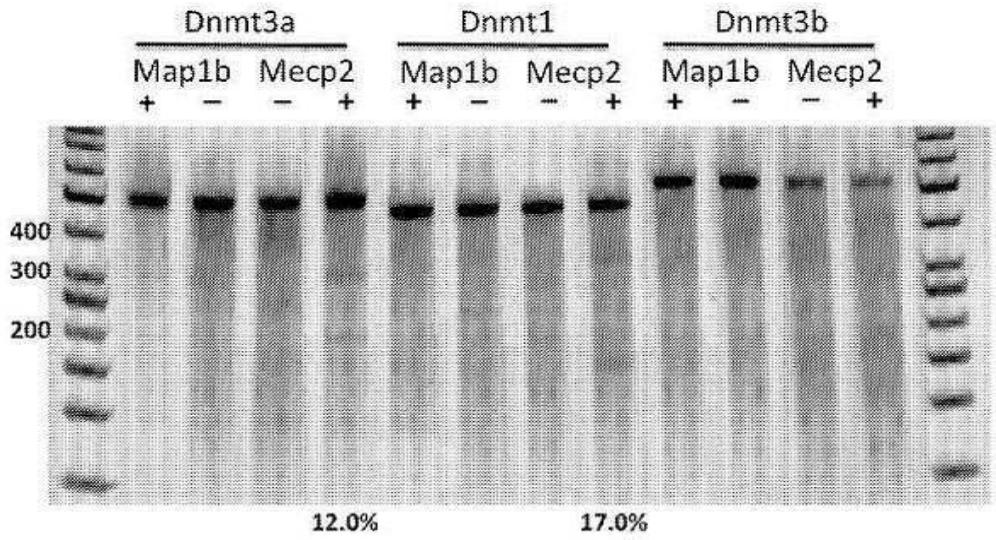
도면45


대조군
Syn_SpCas9-spA
Mecp2_SpCas9-spA
Map1b_SpCas9-spA
Mecp2_SpCas9-bgh
sMap1b_SpCas9-bgh
Map1b_SpCas9-bgh
Sp_TETs(1)
Sp_DNMTs(1)

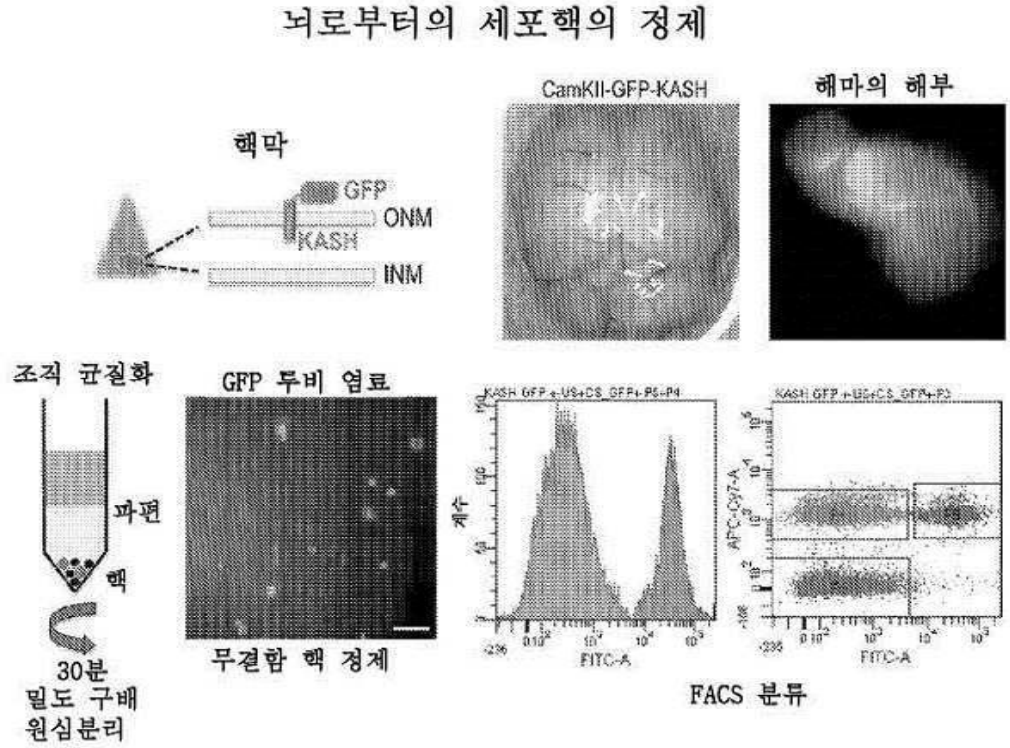
도면46


Dnmt3b
Dnmt1
Dnmt3a
hSyn-spA
mMecp2-spA
rMap1b-spA
rSMap1b-bGH
rMap1b-bGH
1.3
1.1
8.2
15.6
12.2
9.8
8.7
11.5
13.0
13.7
9.8
9.9
6.8
11.9
500
400
300
200
200
1차 뉴런

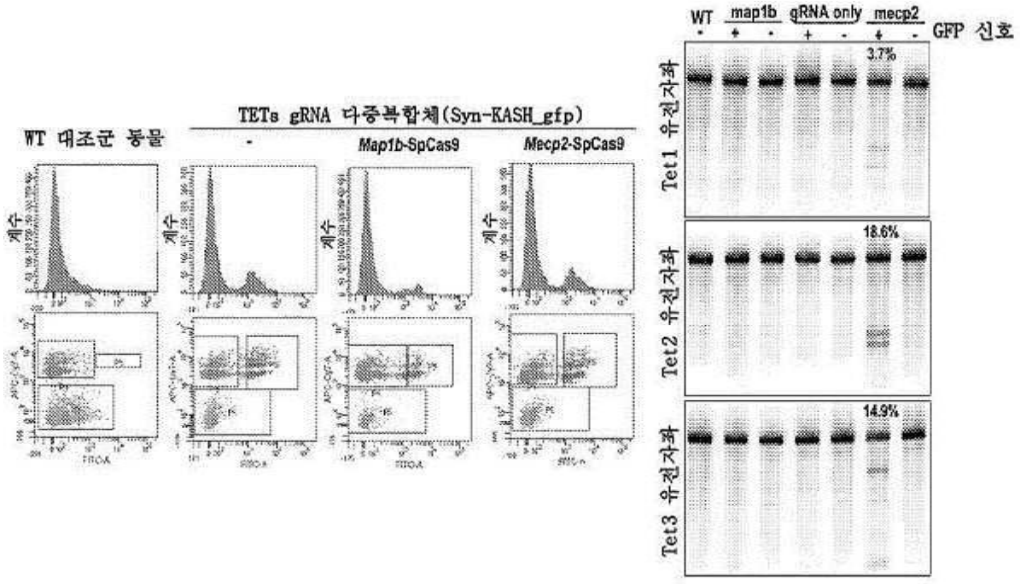
도면47


생체내 SpCas9: Mecp2
Dnmt3a
Dnmt1
Dnmt3b
Map1b
Mecp2
Map1b
Mecp2
Map1b
Mecp2
+ − − +
+ − − +
+ − − +
400
300
200
12.0%
17.0%

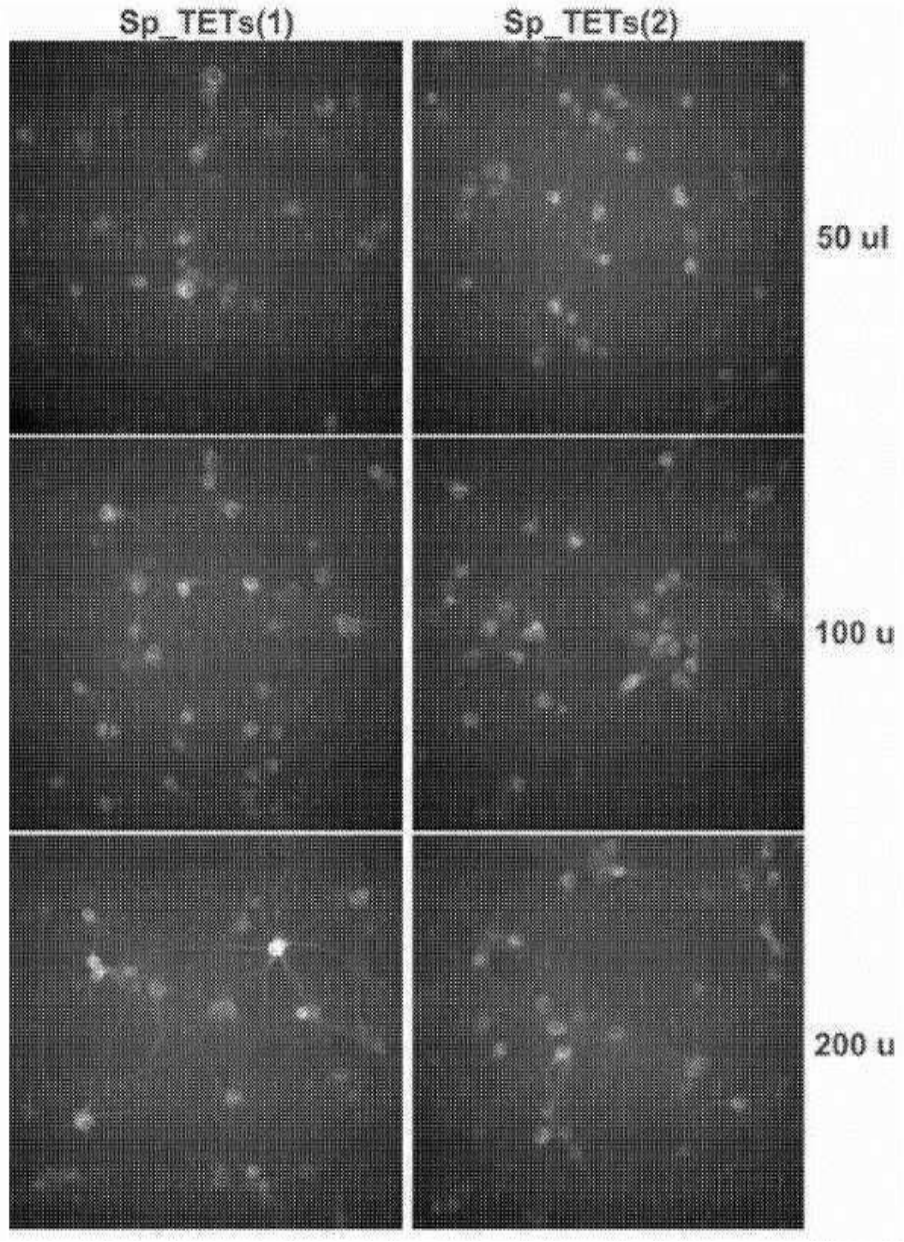
도면48


뇌로부터의 세포핵의 정제
핵막
GFP
ONM
KASH
INM
CamKII-GFP-KASH
해마의 해부
조직 균질화
파편
핵
30분
밀도 구배
원심분리
GFP 두비 염료
무결함 핵 정제
계수
FITC-A
FACS 분류

도면49

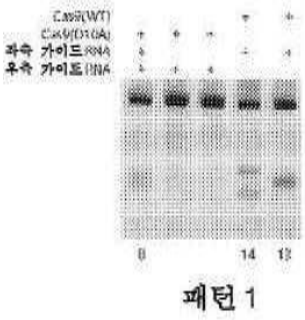

WT 대조군 동물
TETs gRNA 다중복합체(Syn-KASH_gfp)
-
Map1b-SpCas9
Mecp2-SpCas9
계수
WT
map1b
gRNA only
mecp2
GFP 신호
Tet1 유전자좌
Tet2 유전자좌
Tet3 유전자좌
3.7%
18.6%
14.9%

도면50

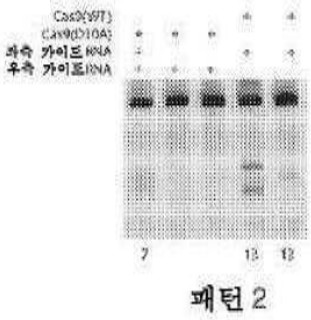


도면51

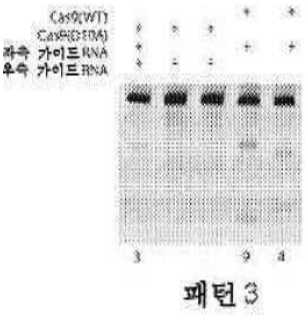
	패턴	돌출부
1	CCnNNNNNNNNNNNNNNNNnnnnnnnnnnnnNNNNNNNNNNNNNNNnGG	39
2	CCnNNNNNNNNNNNNNNNNnnnnnnnnnnNNNNNNNNNNNNNNNnGG	34
3	CCnNNNNNNNNNNNNNNNNnnnnnnnnNNNNNNNNNNNNNNNnGG	30
4	CCnNNNNNNNNNNNNNNNNnnnnnnnNNNNNNNNNNNNNNNnGG	26
5	CCnNNNNNNNNNNNNNNNNnnnnnnNNNNNNNNNNNNNNNnGG	22
6	CCnNNNNNNNNNNNNNNNNnnnnnNNNNNNNNNNNNNNNnGG	18
7	CCnNNNNNNNNNNNNNNNNnnnnNNNNNNNNNNNNNNNnGG	14
8	CCnNNNNNNNNNNNNNNNNnnnNNNNNNNNNNNNNNNnGG	10
9	CCnNNNNNNNNNNNNNNNNnnNNNNNNNNNNNNNNNnGG	6
10	CCnNNNNNNNNNNNNNNNNnNNNNNNNNNNNNNNNnGG	4
11	CCnNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNnGG	3
12	CCnNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNnGG	2
13	CCnNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNnGG	1
14	CCnNNNNNNNNNNNNNNNNNNNNNNNNNNNNNnGG	0
15	CCnNNNNNNNNNNNNNNNNNNNNNNNNNNNNnGG	1
16	CCnNNNNNNNNNNNNNNNNNNNNNNNNNNNnGG	2
17	CCnNNNNNNNNNNNNNNNNNNNNNNNNNNnGG	3
18	CCnNNNNNNNNNNNNNNNNNNNNNNNNNnGG	4
19	CCGG	8
20	GGCC	10
21	GCNNCC	12
22	GCNNNCC	14
23	GCNNNNCC	16
24	GCNNNNNNCC	20



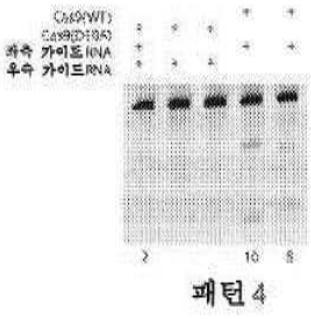
패턴 1



패턴 2



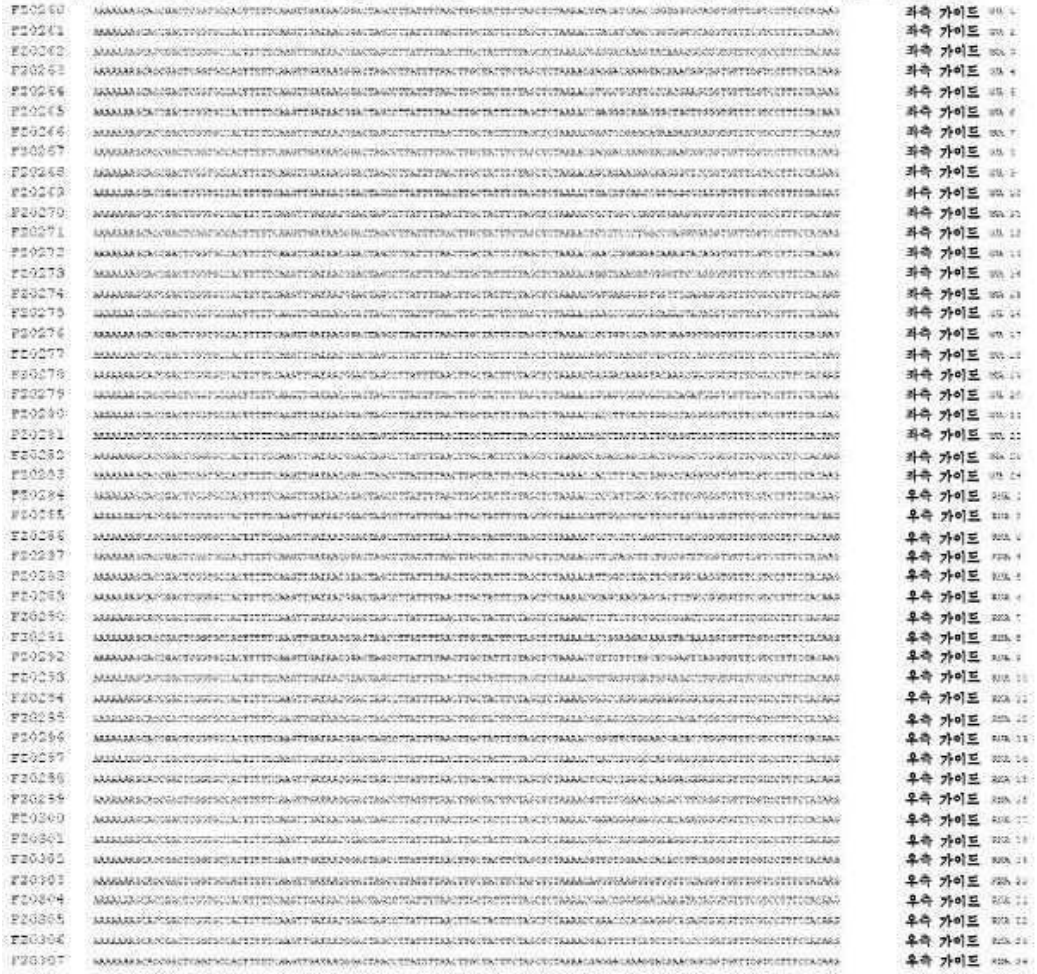
패턴 3



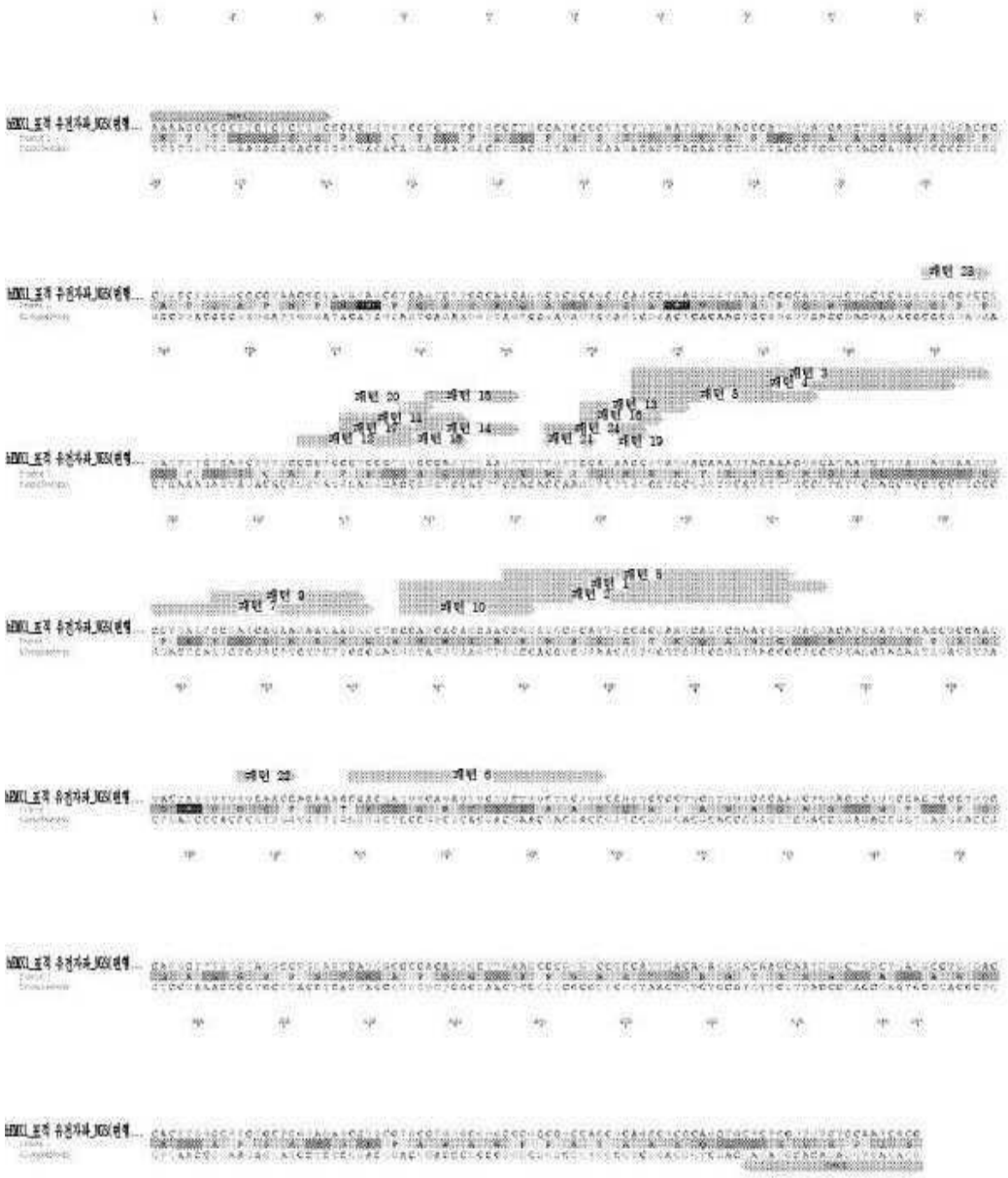
패턴 4

도면52

U6-가이드 RNA 발현 카세트를 생성하기 위한 역방향 PCR 프라이머 (U6 정방향 프라이머를 지니는 쌍)

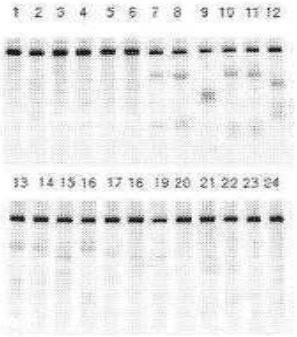


도면53

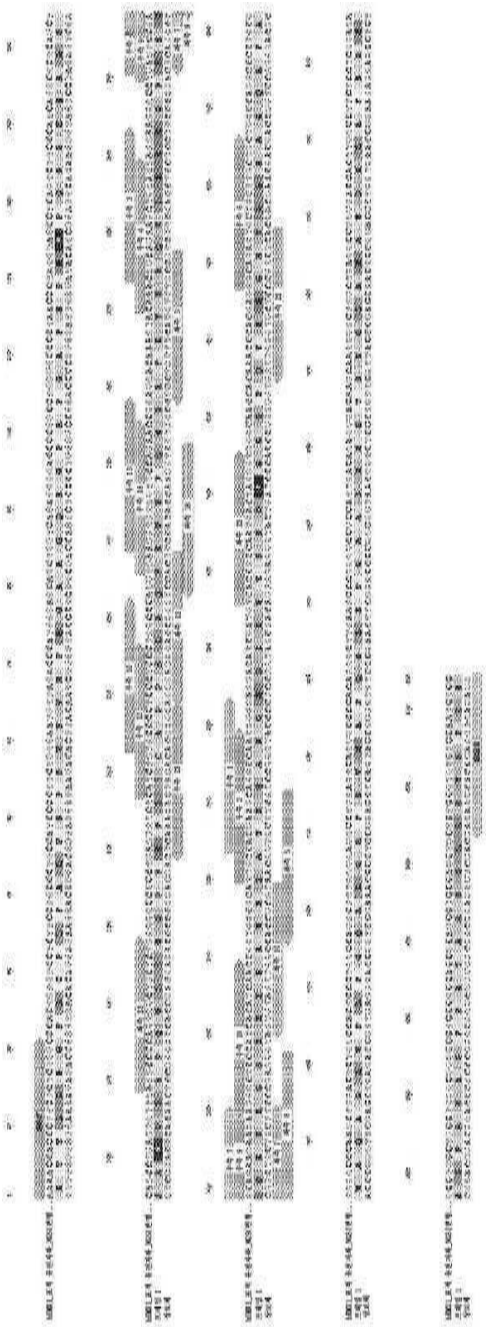


도면54

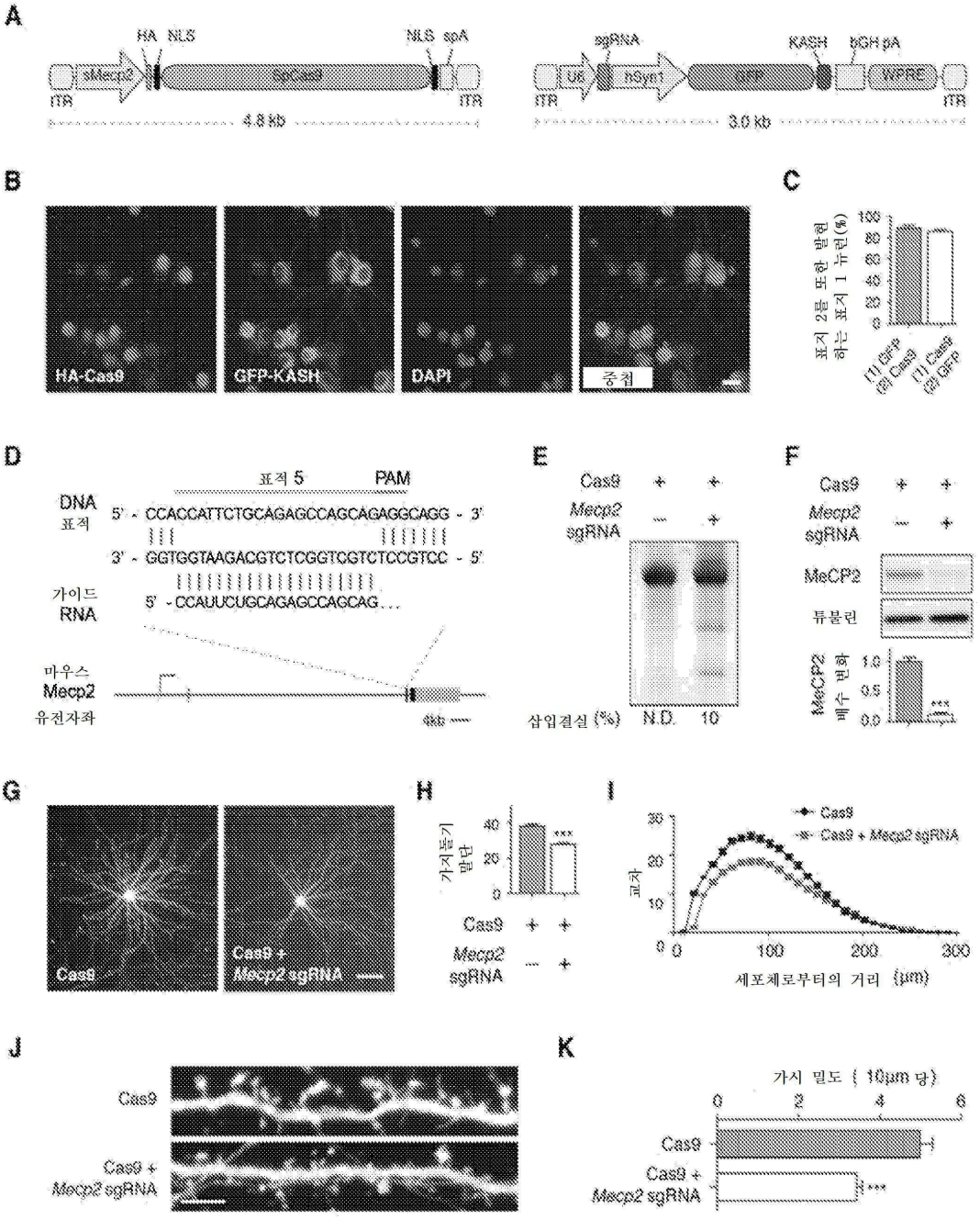
레인	사용을 위한 Cas9	좌측 gRNA	우측 gRNA	돌출부 (D10A)-5'으로부터 돌출된 염기의 수
1	D10A	좌측 23	좌측 12	-36
2	D10A	우측 4	좌측 9	-25
3	D10A	좌측 23	우측 23	-16
4	D10A	우측 7	좌측 10	-15
5	D10A	우측 16	좌측 3	-8
6	D10A	우측 22	우측 6	26
7	D10A	좌측 12	우측 16	31
8	D10A	좌측 12	우측 13	34
9	D10A	좌측 10	우측 1	38
10	D10A	우측 23	우측 16	51
11	D10A	우측 23	우측 13	54
12	D10A	좌측 3	우측 7	57
13	D10A	좌측 12	우측 4	65
14	D10A	좌측 12	우측 3	69
15	D10A	좌측 3	우측 10	76
16	D10A	우측 23	우측 4	85
17	D10A	좌측 12	우측 9	95
18	D10A	좌측 12	우측 10	115
19	D10A	우측 23	우측 10	135
20	D10A	좌측 12	우측 2	145
21	D10A	좌측 12	좌측 22	181
22	D10A	우측 23	좌측 22	201
23	D10A	좌측 12	우측 6	222
24	D10A	우측 23	우측 6	242



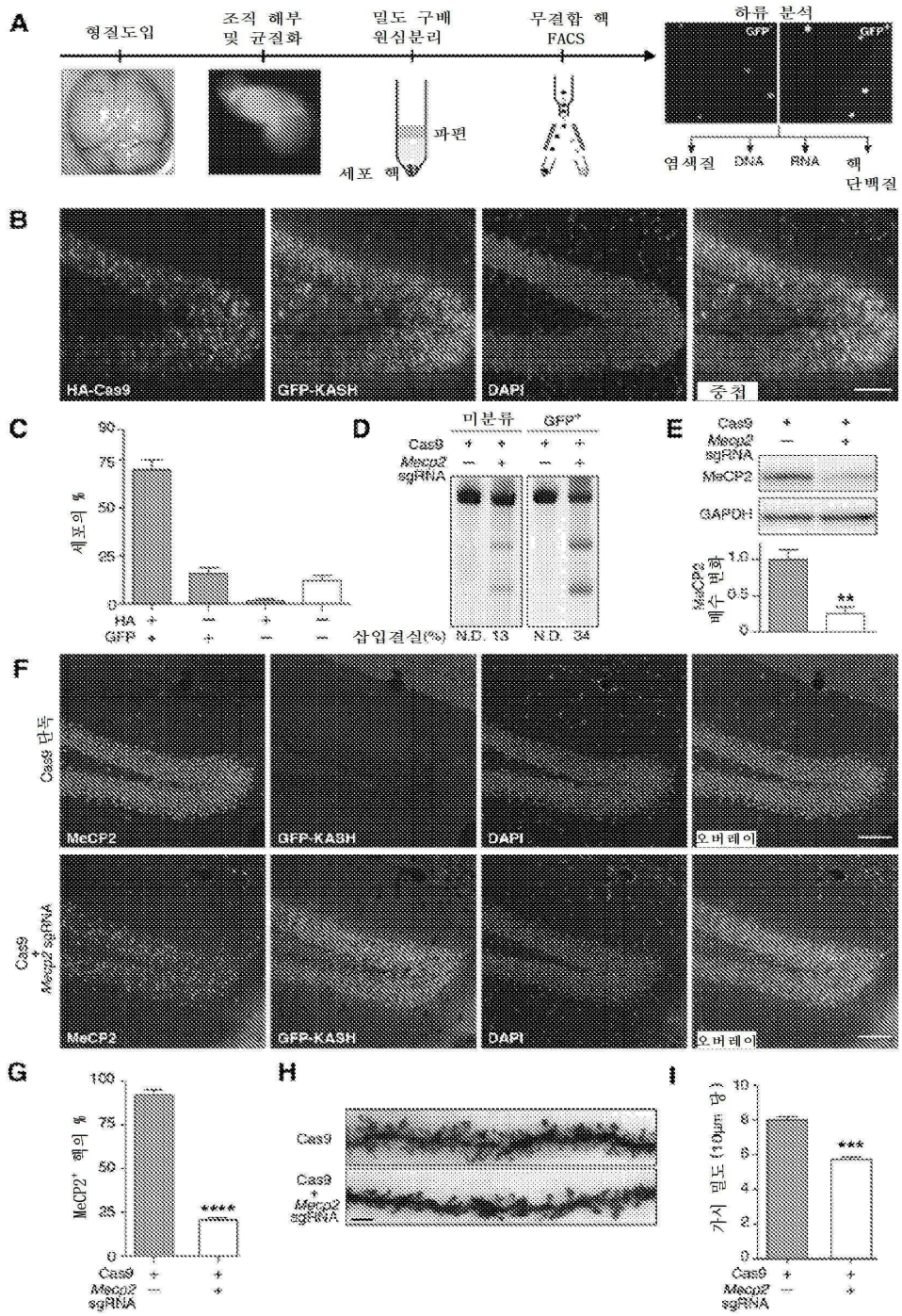
도면55



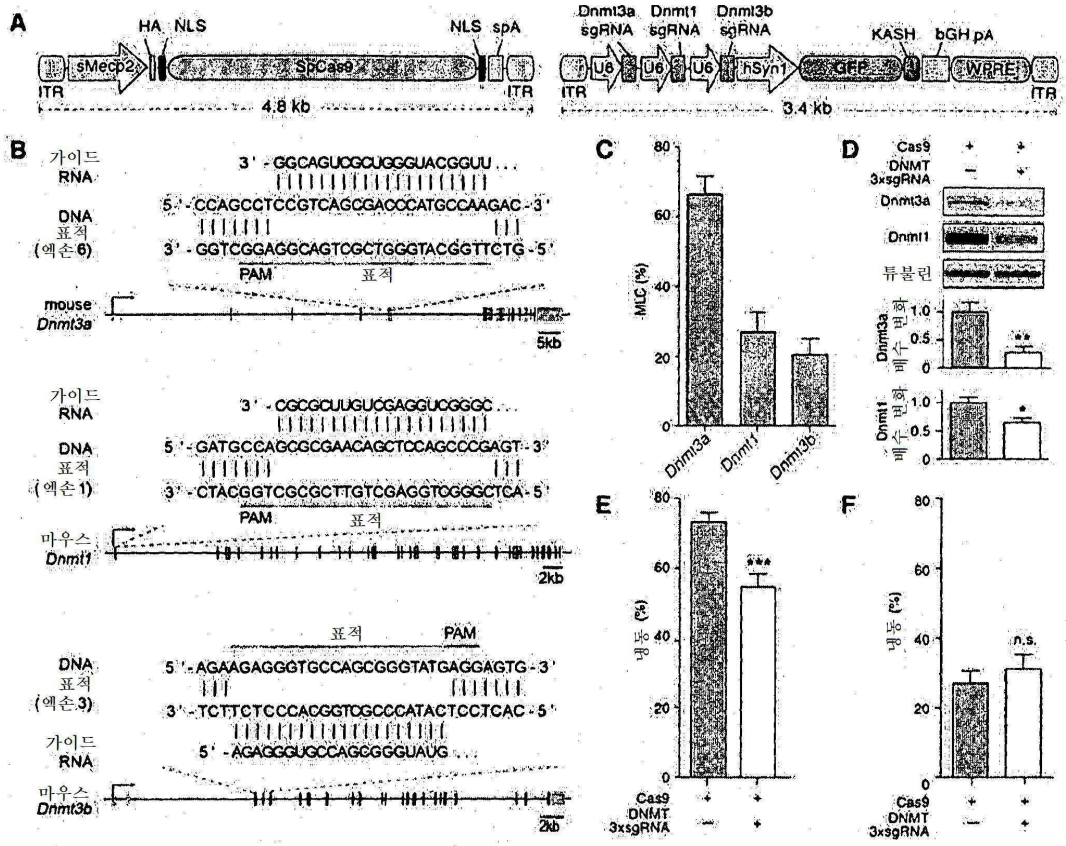
도면56


A
HA NLS
NLS spA
sMecp2
spCas9
ITR
ITR
4.8 kb
sgRNA
KASH
bGH pA
U6
hSyn1
GFP
WPRE
ITR
ITR
3.0 kb
B
HA-Cas9
GFP-KASH
DAPI
중첩
C
표지 2를 또한 발현하는 표지 1 뉴런(%)
[1] GFP [2] Cas9
[1] Cas9 [2] GFP
D
표적 5
PAM
DNA 표적
5' - CCACCATTCTGCAGAGCCAGCAGAGGCAGG - 3'
3' - GGTGGTAAGACGTCTCGGTCGTCTCCGTCC - 5'
가이드 RNA
5' - CCAUUCUGCAGAGCCAGCAG...
마우스 Mecp2 유전자좌
4kb
E
Cas9
Mecp2 sgRNA
삼입결실 (%)
N.D.
10
F
Cas9
Mecp2 sgRNA
MeCP2
튜불린
MeCP2 배수 변화
G
Cas9
Cas9 + Mecp2 sgRNA
H
가지돋기 밀단
Cas9
Mecp2 sgRNA
I
Cas9
Cas9 + Mecp2 sgRNA
교차
세포체로부터의 거리 (μm)
J
Cas9
Cas9 + Mecp2 sgRNA
K
가시 밀도 (10μm 당)
Cas9
Cas9 + Mecp2 sgRNA

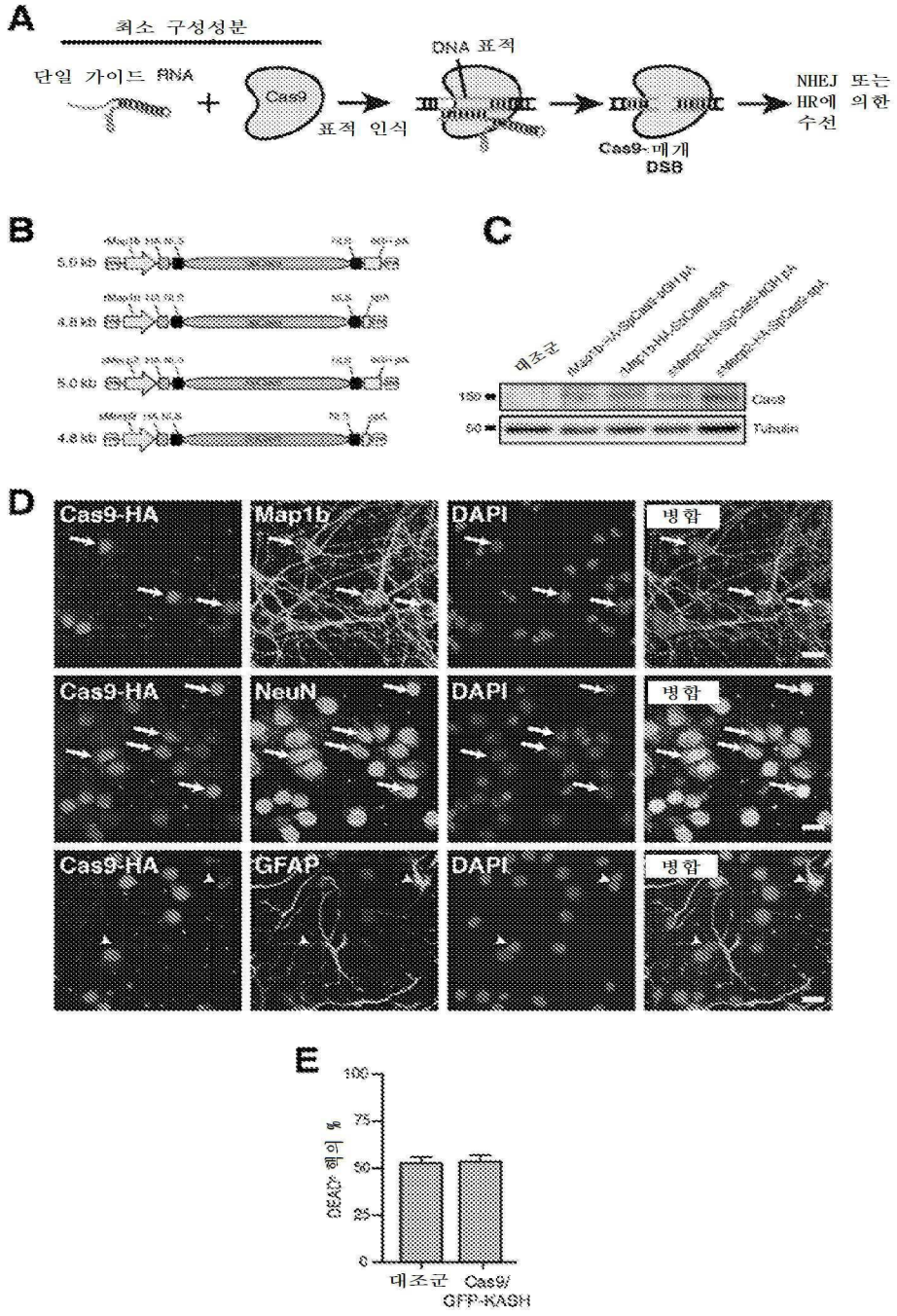
도면57


A
형질도입
조직 해부 및 균질화
밀도 구배 원심분리
무결합 핵 FACS
하류 분석
GFP-
GFP+
파편
세포 핵
염색질
DNA
RNA
핵 단백질
B
HA-Cas9
GFP-KASH
DAPI
중첩
C
세포의 %
HA
GFP
D
미분류
GFP+
Cas9
Mecp2 sgRNA
삽입결실(%)
N.D. 13
N.D. 34
E
MeCP2
GAPDH
MeCP2 베수 변화
F
Cas9 단독
Cas9 + Mecp2 sgRNA
MeCP2
GFP-KASH
DAPI
오버레이
G
MeCP2+ 핵의 %
H
Cas9
I
가시 밀도 (10μm 당)

도면58


A
HA NLS
NLS spA
sMecp2
SpCas9
ITR
ITR
4.8 kb
Dnmt3a sgRNA
Dnmt1 sgRNA
Dnmt3b sgRNA
U6
U6
U6
hSyn1
GFP
KASH
bGH pA
WPRE
ITR
ITR
3.4 kb
B
가이드 RNA
3' - GGCAGUCGCUGGGUACGGUU . . .
DNA 표적 (엑손6)
5' - CCAGCCTCCGTCAGCGACCCATGCCAAGAC - 3'
3' - GGTCGGAGGCAGTCGCTGGGTACGGTTCTG - 5'
PAM
표적
mouse Dnmt3a
5kb
가이드 RNA
3' - CGCGCUUGUCGAGGUCGGGC . . .
DNA 표적 (엑손1)
5' - GATGCCAGCGCGAACAGCTCCAGCCCGAGT - 3'
3' - CTACGGTCGCGCTTGTCGAGGTCGGGCTCA - 5'
PAM
표적
마우스 Dnmt1
2kb
표적
PAM
DNA 표적 (엑손3)
5' - AGAAGAGGGTGCCAGCGGGTATGAGGAGTG - 3'
3' - TCTTCTCCCACGGTCGCCCATACTCCTCAC - 5'
가이드 RNA
5' - AGAGGGUGCCAGCGGGUAUG . . .
마우스 Dnmt3b
2kb
C
MLC (%)
Dnmt3a
Dnmt1
Dnmt3b
D
Cas9
DNMT 3xsgRNA
Dnmt3a
Dnmt1
튜불린
Dnmt3a 발현 변수
Dnmt1 발현 변수
E
냉동 (%)
***
Cas9
DNMT 3xsgRNA
F
냉동 (%)
n.s.
Cas9
DNMT 3xsgRNA

도면59

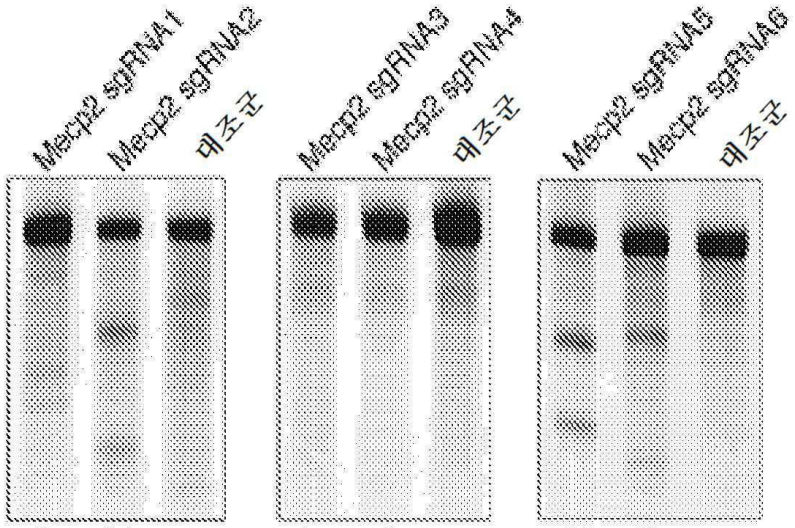

A
최소 구성성분
단일 가이드 RNA
Cas9
표적 인식
DNA 표적
Cas9-매개
DSB
NHEJ 또는
HR에 의한
수선
B
C
대조군
Cas9
Tubulin
D
Cas9-HA
Map1b
DAPI
병합
NeuN
GFAP
E
% 이중 가닥 절단
대조군
Cas9/
GFP-KASH

도면60

A

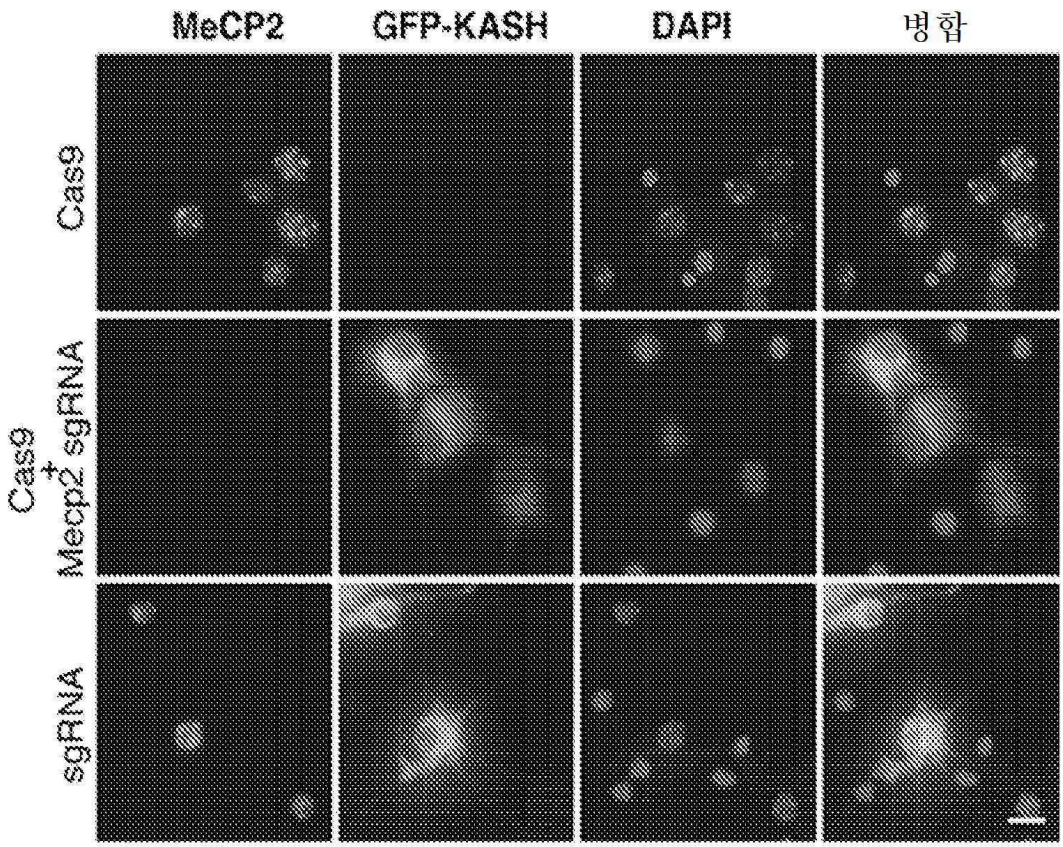
	Mecp2 표적 서열	PAM
표적 1	CTGGGAGAGGGAGCCCCTCC	AGG
표적 2	AAAGGTGGGAGACACCTCCT	TGG
표적 3	TCCAACCTTCAGGCAAGGTG	GGG
표적 4	AGGAAGTCTGGCCGATCTGC	TGG
표적 5	CCATTCTGCAGAGCCAGCAG	AGG
표적 6	CTCTGAGGCCCTGGAGATCC	TGG

B

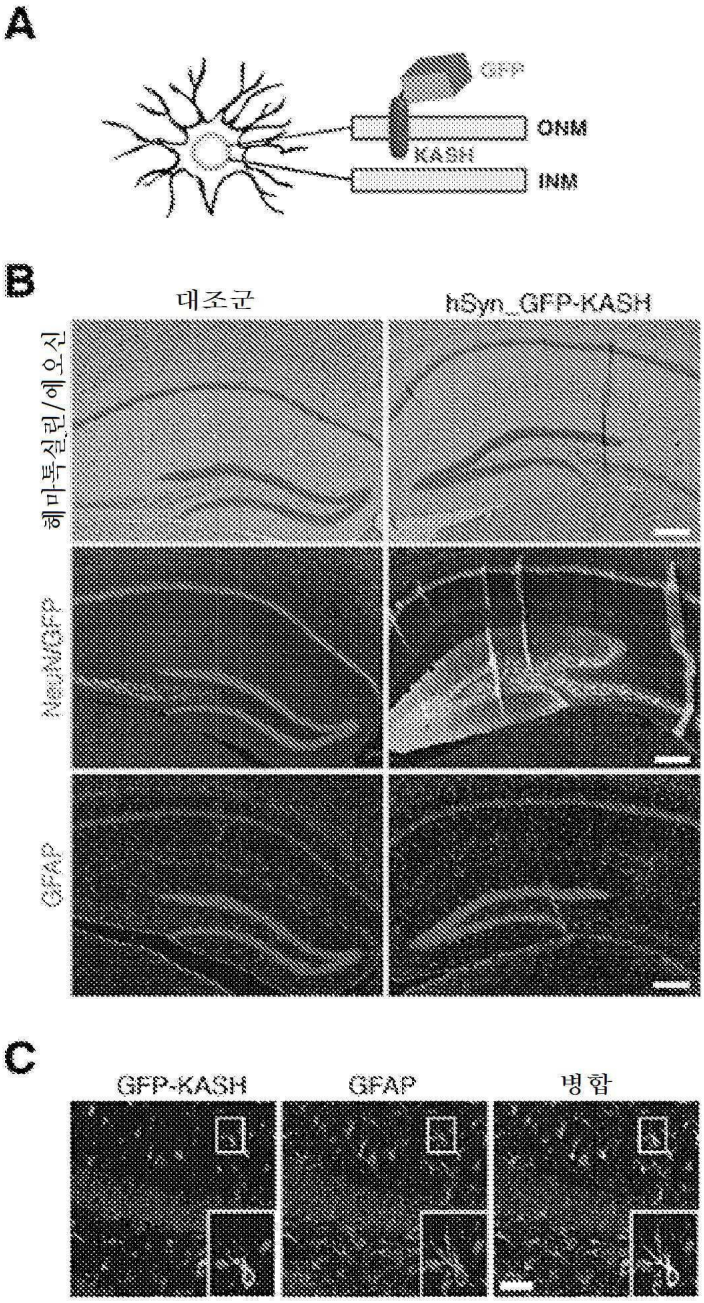


삽입결실 (%) 5 21 0 0 23 13

도면61



도면62

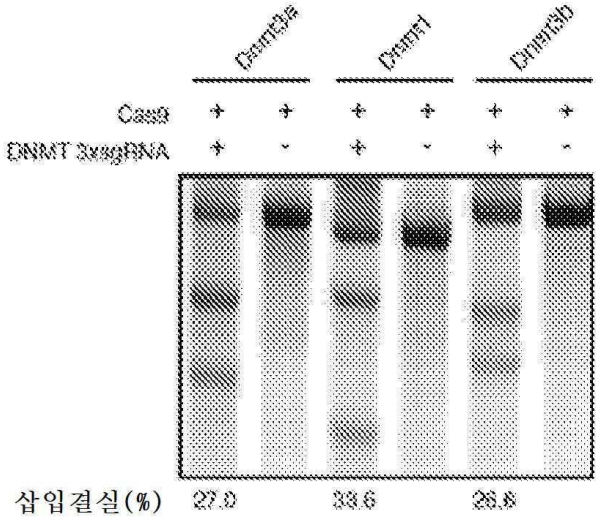

A
GFP
ONM
KASH
INM
B
대조군
hSyn_GFP-KASH
헤마톡실린/에오신
NeuN/GFP
GFAP
C
GFP-KASH
GFAP
병합

도면63

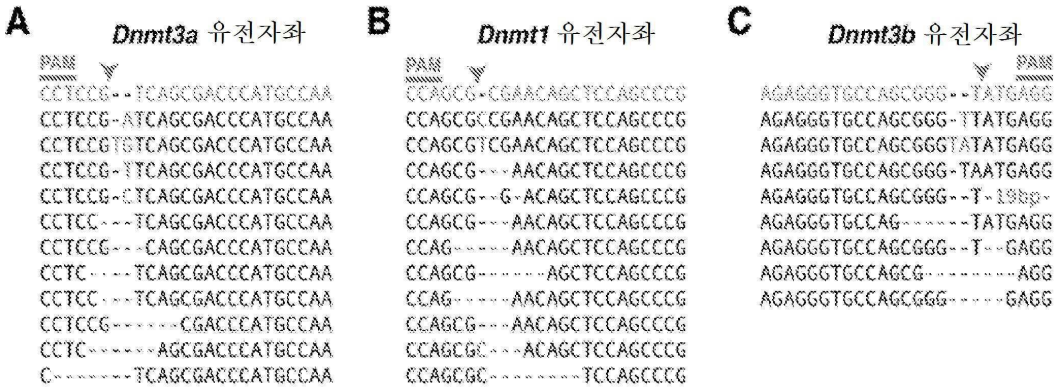
A

표적	표적 서열	PAM
Dnmt3a	TTGGCATGGGTCGCTGACGG	AGG
Dnmt1	CGGGCTGGAGCTGTTCGCGC	TGG
Dnmt3b	AGAGGGTGCCAGCGGGTATG	AGG

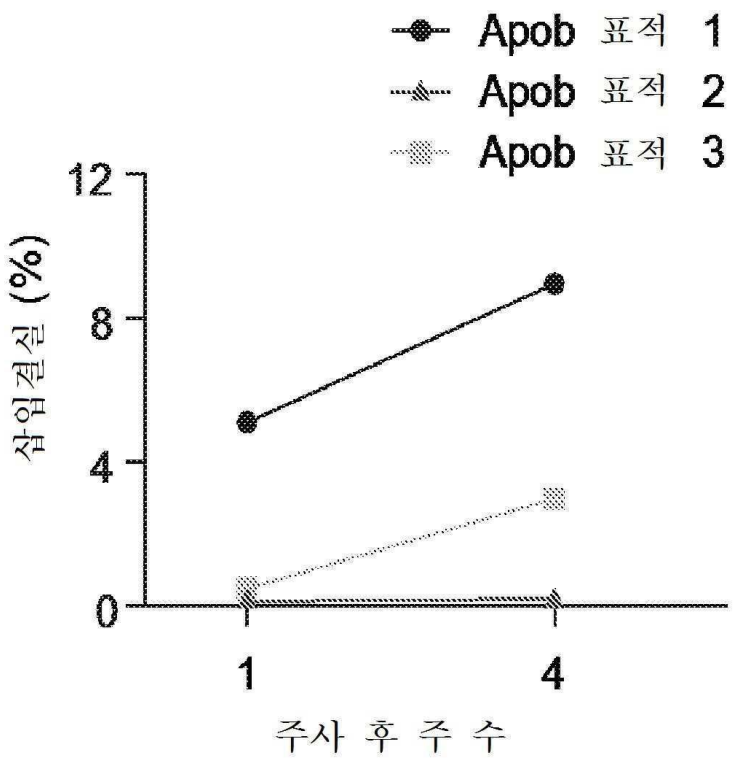
B



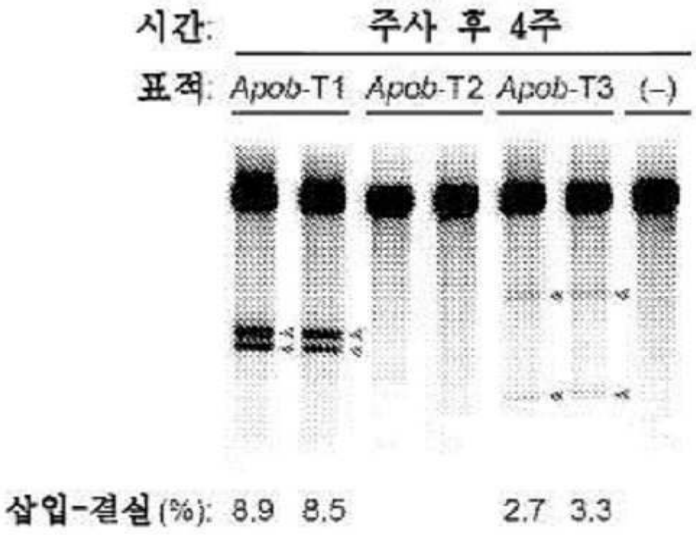
도면64



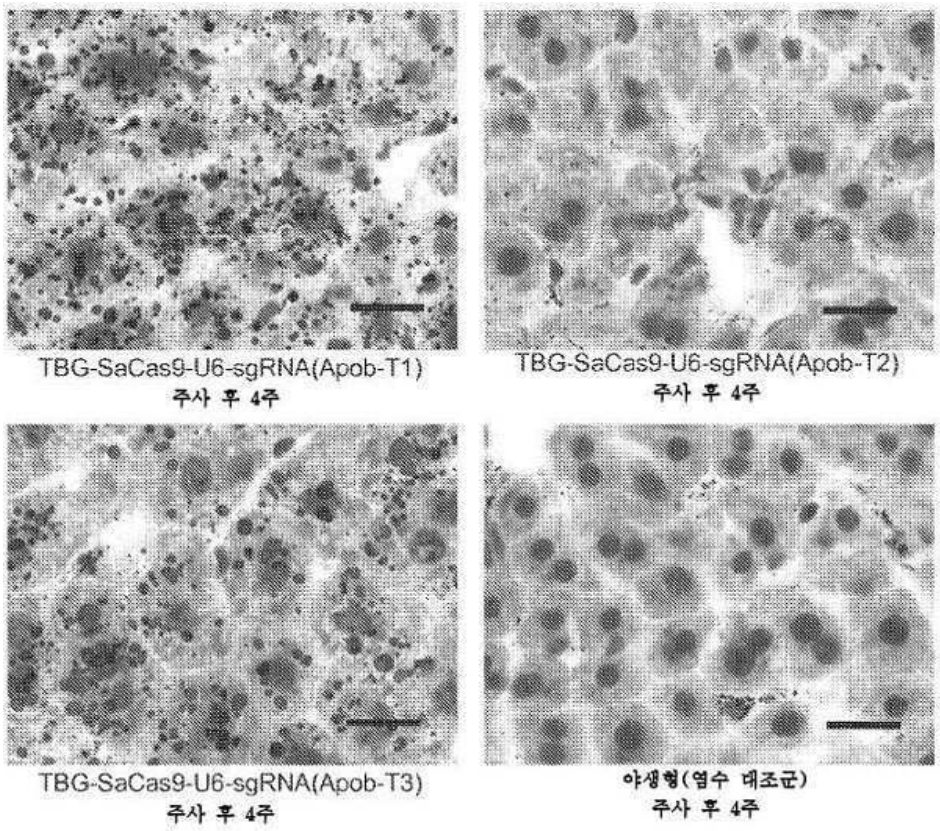
도면65a


Apob 표적 1
Apob 표적 2
Apob 표적 3
12
8
4
0
삽입결실 (%)
1
4
주사 후 주 수

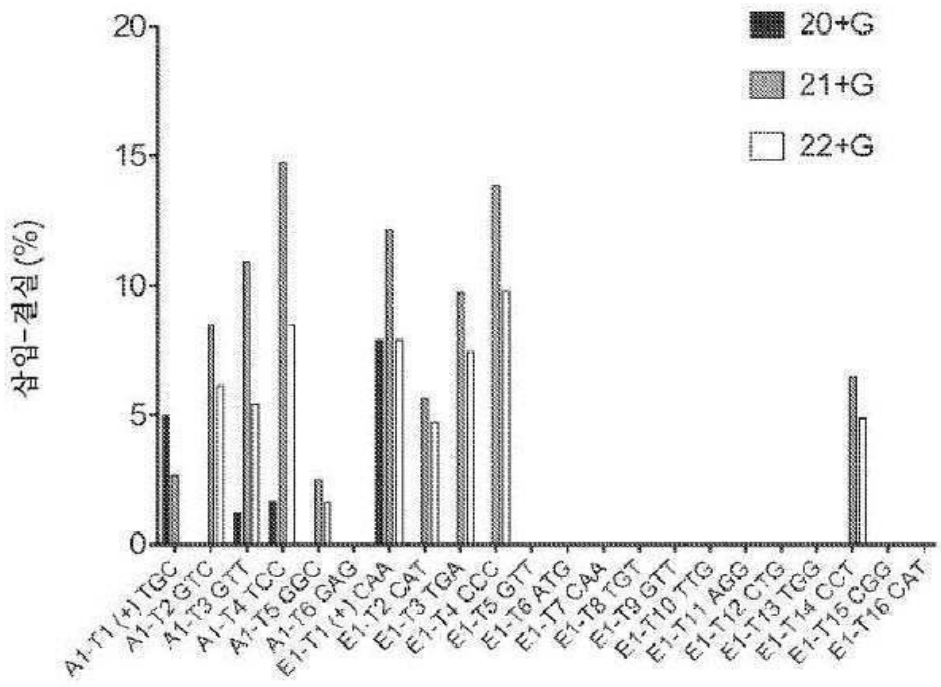
도면65b


시간:
주사 후 4주
표적:
Apob-T1
Apob-T2
Apob-T3
(-)
삽입-결실 (%):
8.9
8.5
2.7
3.3

도면66

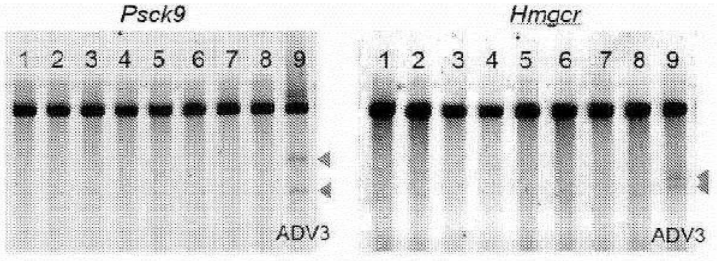


도면67

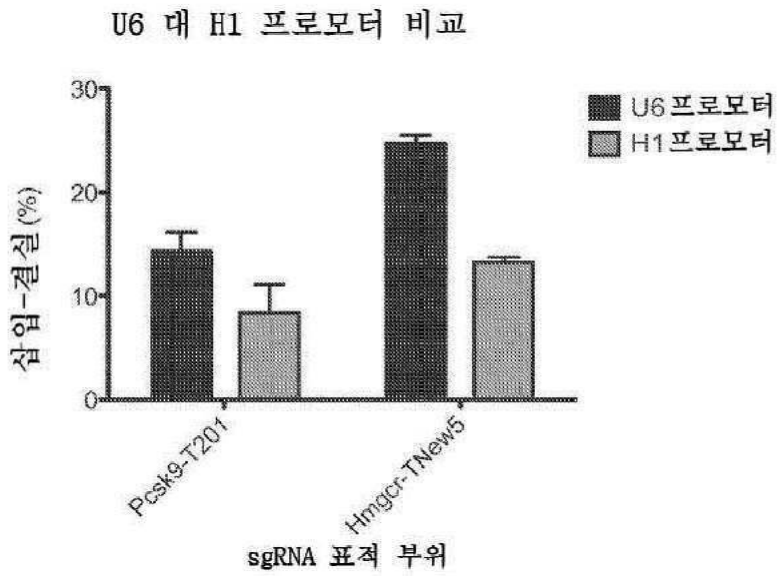


스페이서의 처음 3개의 염기(위치 22, 21, 20)를 갖는 AAVS1 및 EMX1 표적

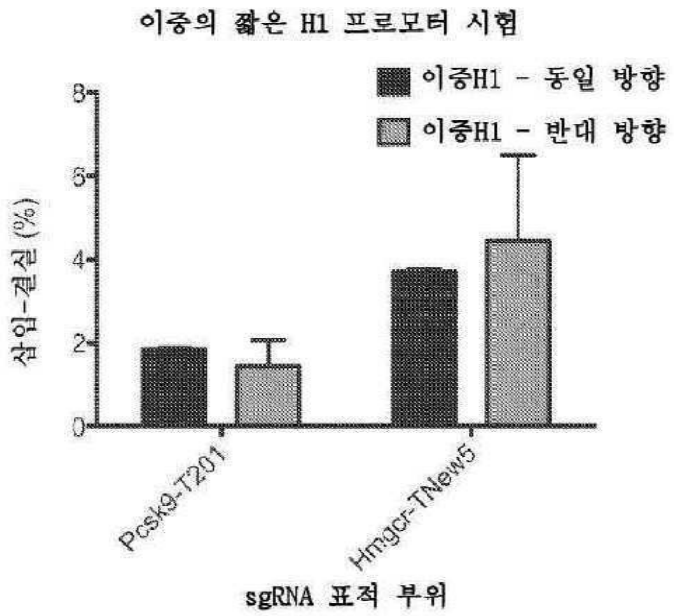
도면68



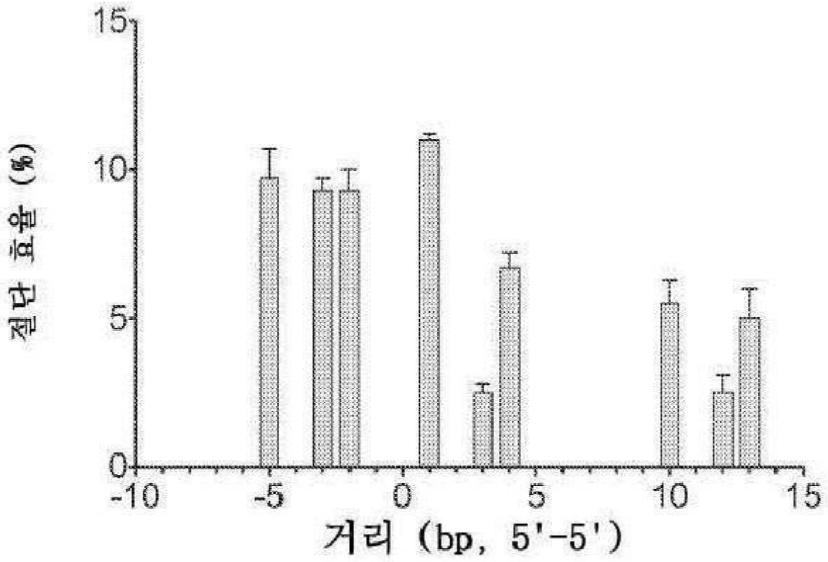
도면69



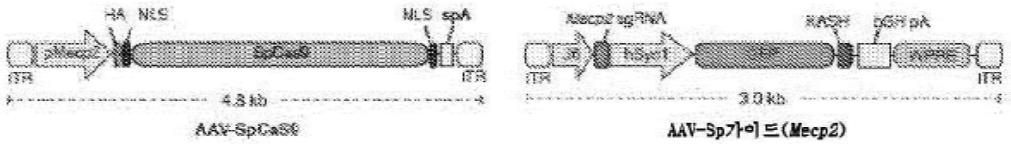
도면70



도면71



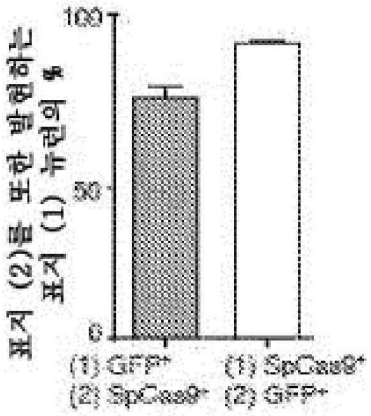
도면72a



도면72b



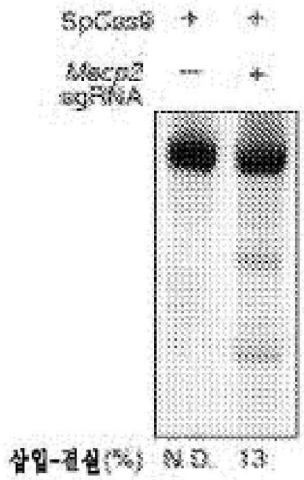
도면72c



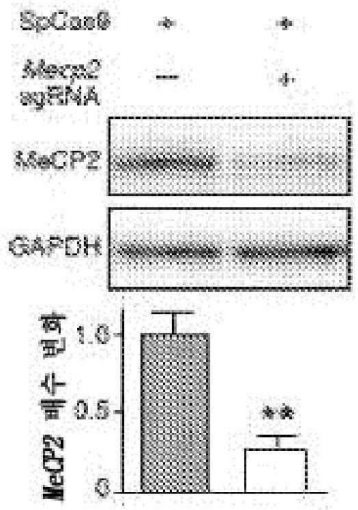
도면72d



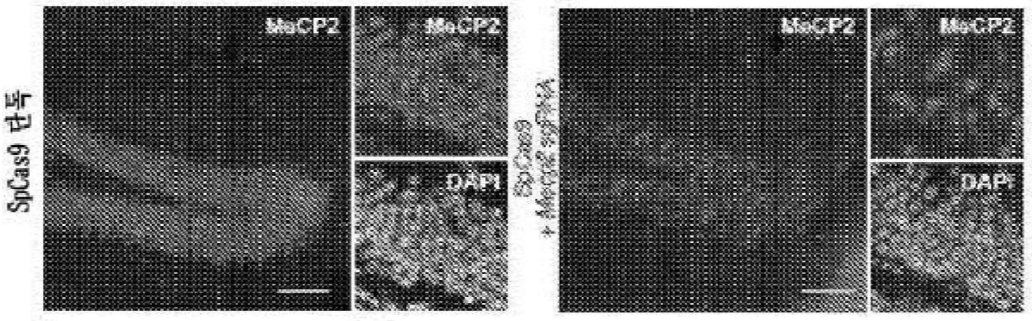
도면72e



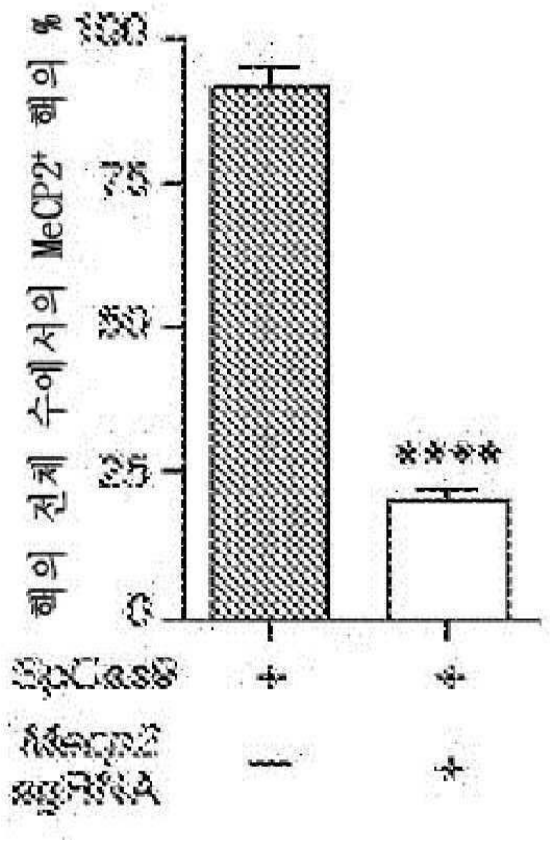
도면72f


SpCas9
Mecp2 sgRNA
MeCP2
GAPDH
MeCP2 발현 변화 배수
1.0
0.5
0
**

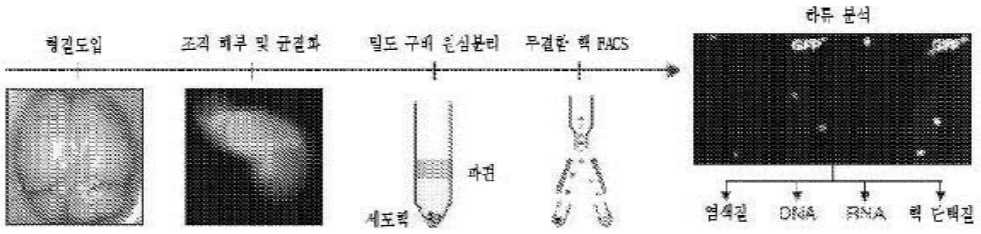
도면72g


SpCas9 단독
SpCas9 + Mecp2 sgRNA
MeCP2
MeCP2
DAPI
MeCP2
MeCP2
DAPI

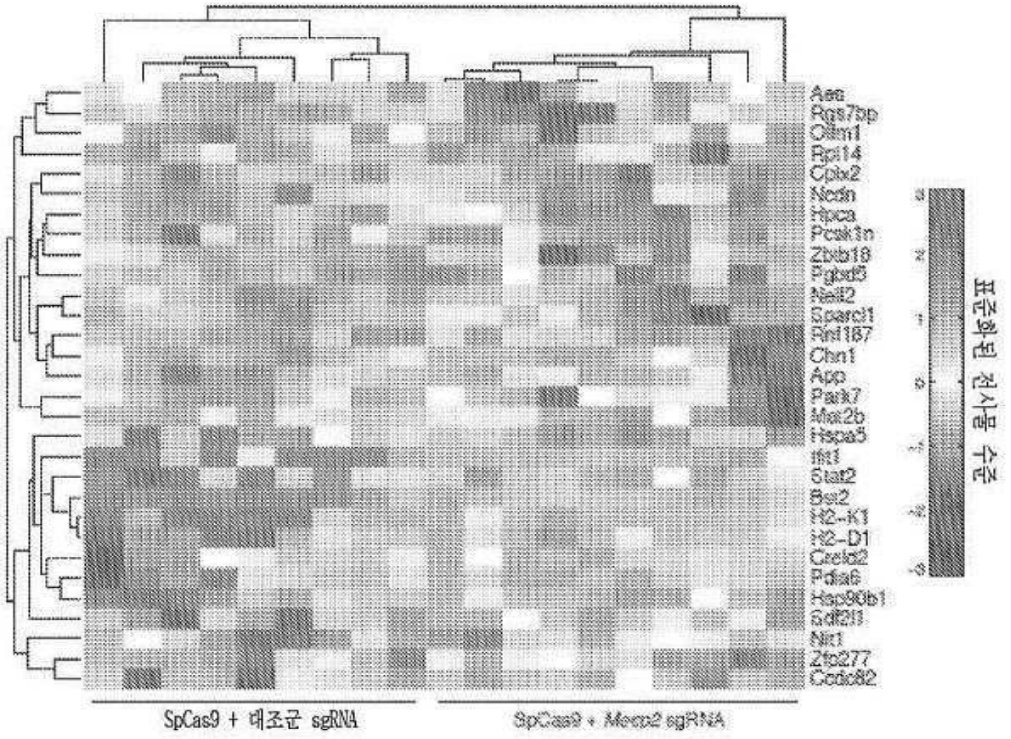
도면72h


해의 전체 수에서의 MeCP2+ 해의 %
****
dCas9
MeCP2
sgRNA

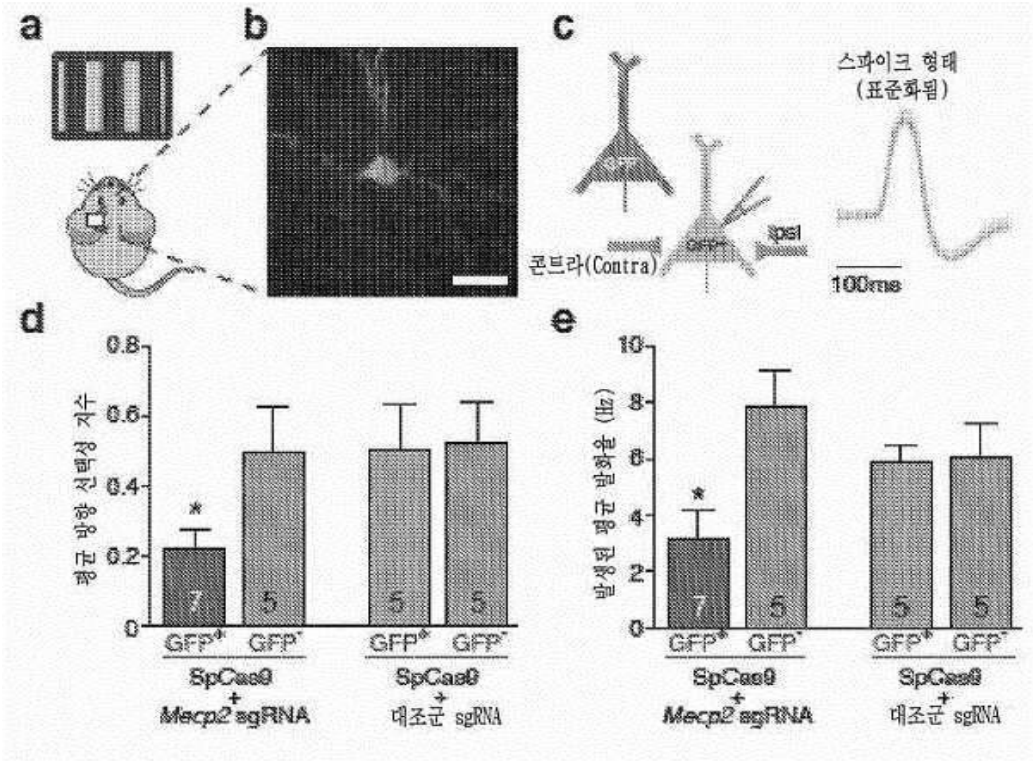
도면73a


형질도입
조직 해쿠 및 균질화
밀도 구배 원심분리
무결한 핵 FACS
라벨 분석
DAPI
세포핵
염색질
DNA
RNA
핵 단백질

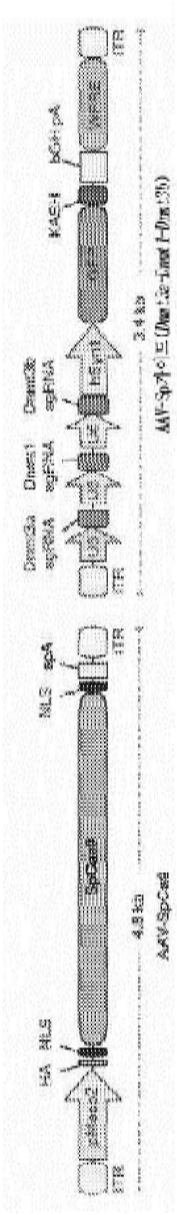
도면73b


Aes
Rgs7bp
Cltm1
Rpl14
Cplx2
Ncdn
Hpca
Pcsk1n
Zbtb18
Pgbd3
Nell2
Sparcl1
Rnf187
Chn1
App
Park7
Mx2b
Hspa5
Ifit1
Stat2
Bst2
H2-K1
H2-D1
Creb3l2
Pdia6
Hsp90b1
Sdf2l1
Nit1
Zfp277
Ccdc82
표준화된 전사물 수준
SpCas9 + 대조군 sgRNA
SpCas9 + Mecp2 sgRNA

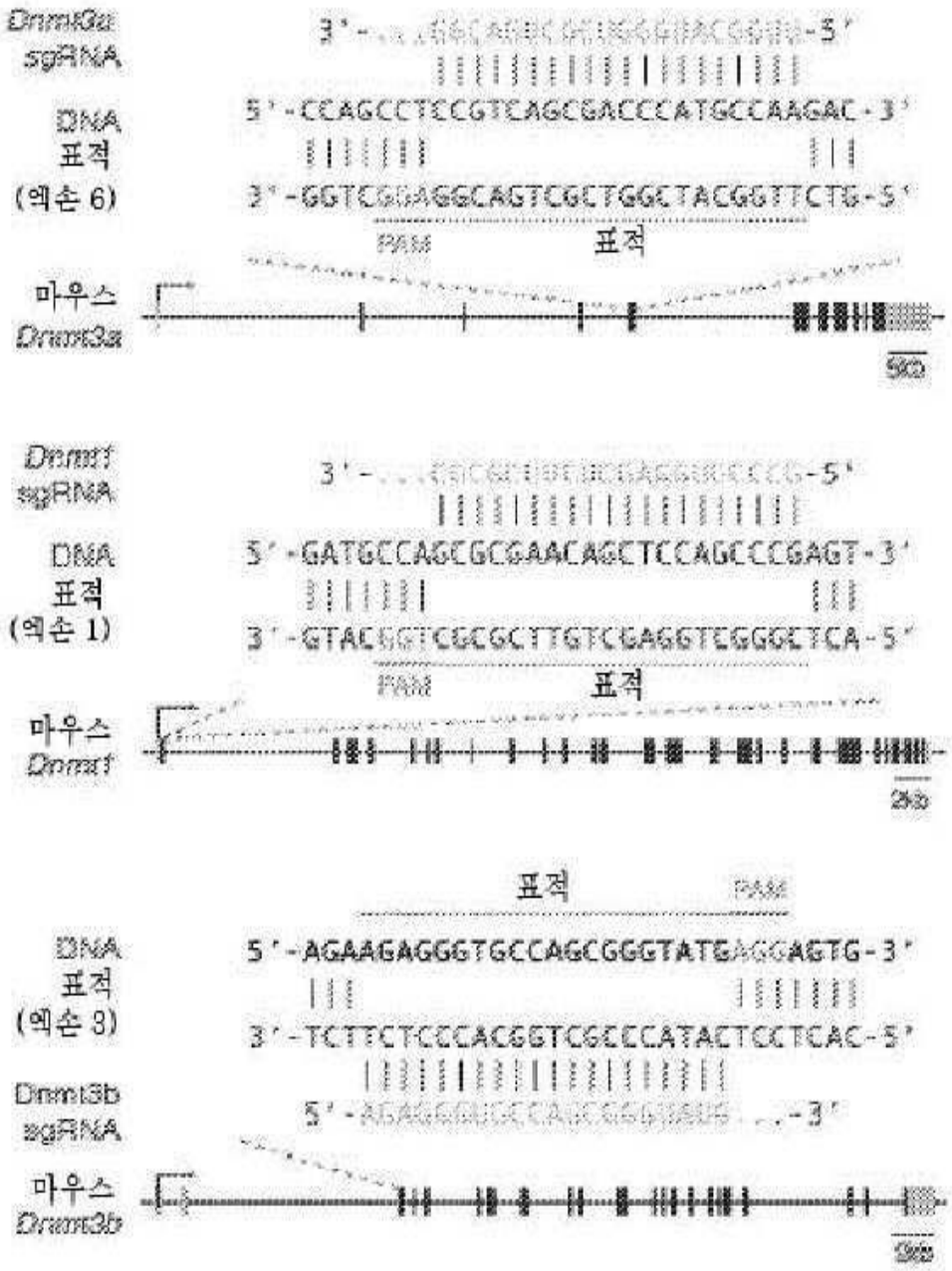
도면74


a
b
c
스파이크 형태
(표준화됨)
콘트라(Contra)
ipsi
100ms
d
평균 방향 선택성 지수
e
발화된 평균 발화율 (Hz)
GFP+
GFP-
SpCas9
+
Mecp2 sgRNA
SpCas9
+
대조군 sgRNA

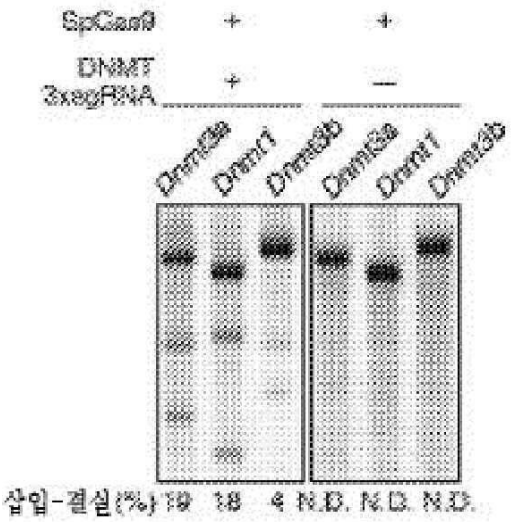
도면75a



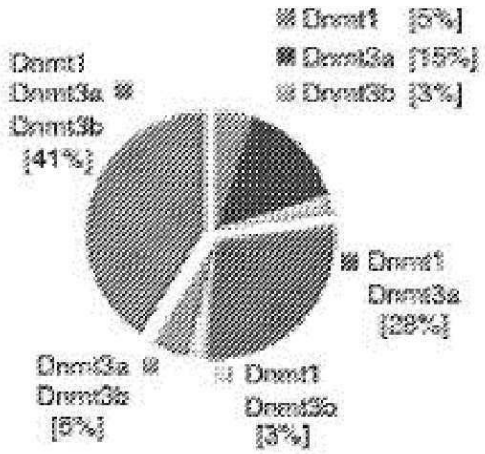
도면75b


Dnmt3a
sgRNA
3'-...-5'
DNA
표적
(엑손 6)
5'-CCAGCCTCCGTCAGCGACCCATGCCAAGAC-3'
3'-GGTCGGAGGCAGTCGCTGGCTACGGTTCTG-5'
PAM
표적
마우스
Dnmt3a
5kb
Dnmt1
sgRNA
3'-...-5'
DNA
표적
(엑손 1)
5'-GATGCCAGCGCGAACAGCTCCAGCCCGAGT-3'
3'-GTACGGTCGCGCTTGTCGAGGTCGGGCTCA-5'
PAM
표적
마우스
Dnmt1
2kb
표적
PAM
DNA
표적
(엑손 3)
5'-AGAAGAGGGTGCCAGCGGGTATGAGGAGTG-3'
3'-TCTTCTCCCACGGTCGCCCATACTCCTCAC-5'
Dnmt3b
sgRNA
5'-...-3'
마우스
Dnmt3b
9kb

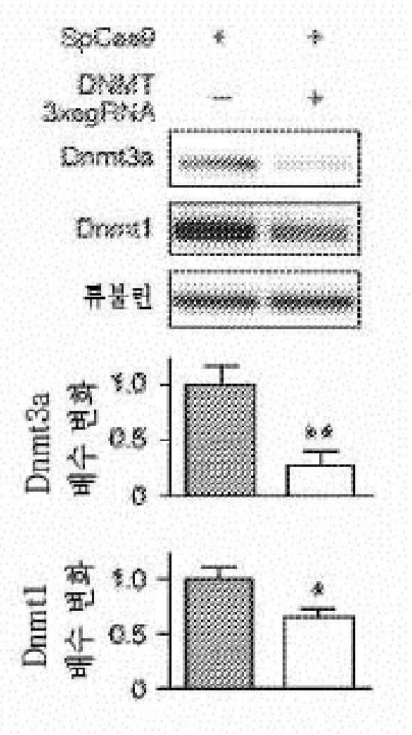
도면75c


SpCas9
DNMT 3xsgRNA
Dnmt3a
Dnmt1
Dnmt3b
Dnmt3a
Dnmt1
Dnmt3b
삽입-결실(%) 19 18 4 N.D. N.D. N.D.

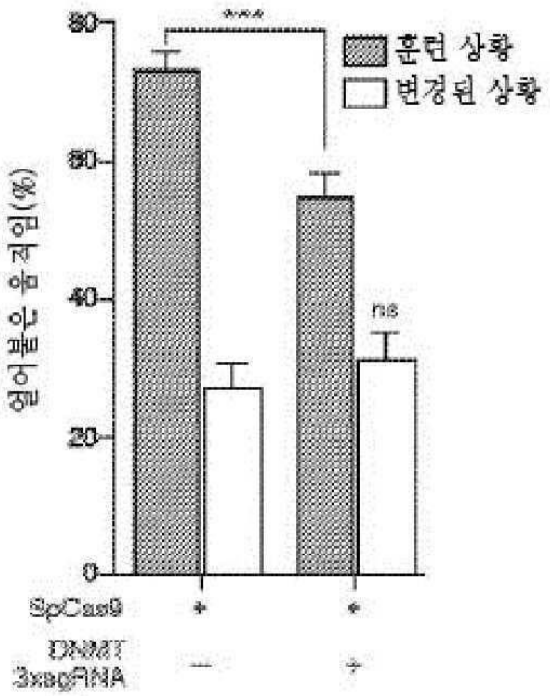
도면75d


Dnmt1 [5%]
Dnmt3a [15%]
Dnmt3b [3%]
Dnmt1
Dnmt3a
Dnmt3b
[41%]
Dnmt1
Dnmt3a
[29%]
Dnmt3a
Dnmt3b
[5%]
Dnmt1
Dnmt3b
[3%]

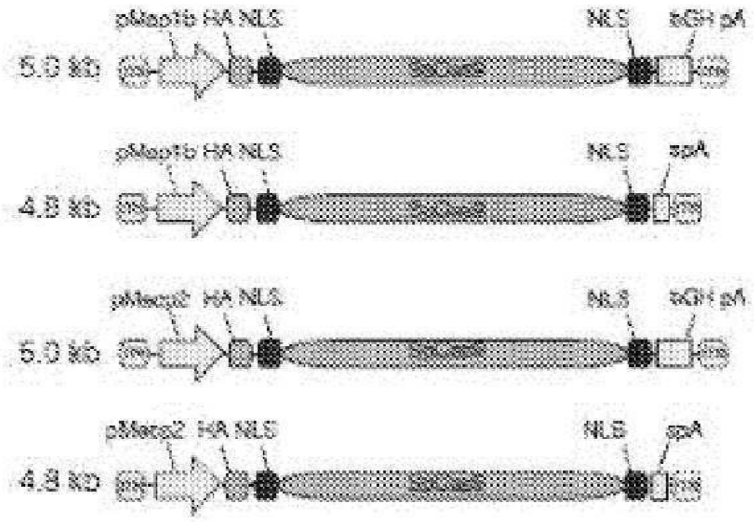
도면75e


SpCas9
DNMT 3xsgRNA
Dnmt3a
Dnmt1
튜불린
Dnmt3a 배수 변화
Dnmt1 배수 변화
1.0
0.5
0
**
*

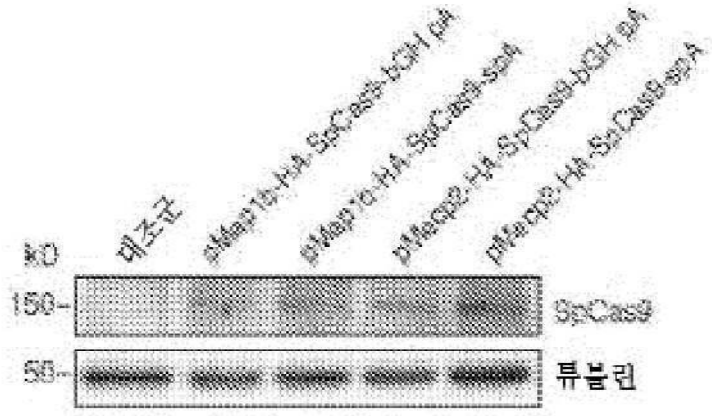
도면75f


훈련 상황
변경된 상황
얼어붙은 움직임(%)
80
60
40
20
0
****
ns
SpCas9
DNMT 3xsgRNA

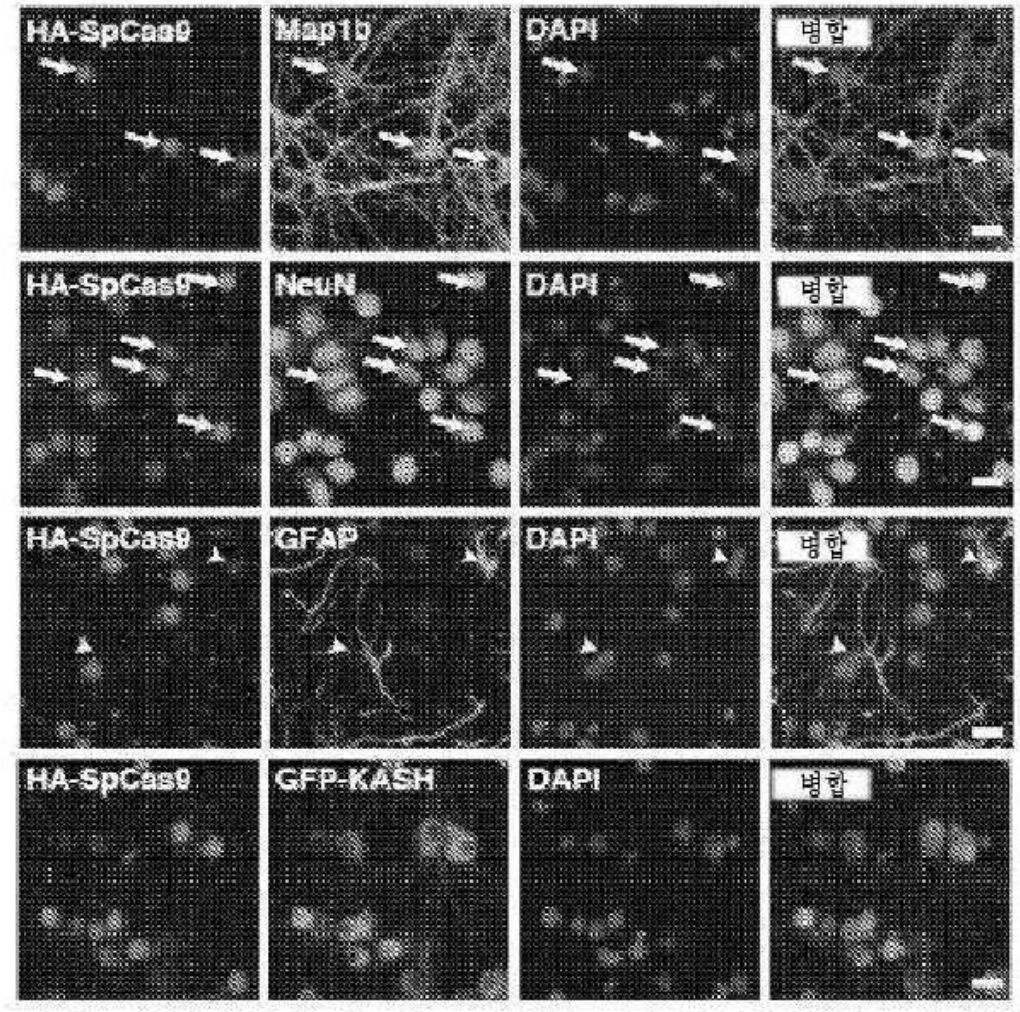
도면76a


pMecp1b HA NLS
NLS
bGH pA
5.0 kb
pMecp1b HA NLS
NLS
spA
4.8 kb
pMecp2 HA NLS
NLS
bGH pA
5.0 kb
pMecp2 HA NLS
NLS
spA
4.8 kb

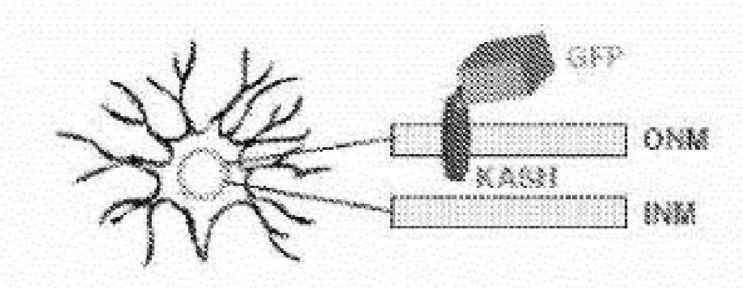
도면76b


kD
대조군
pMecp1b-HA-SpCas9-bGH pA
pMecp1b-HA-SpCas9-spA
pMecp2-HA-SpCas9-bGH pA
pMecp2-HA-SpCas9-spA
150-
SpCas9
55-
튜불린

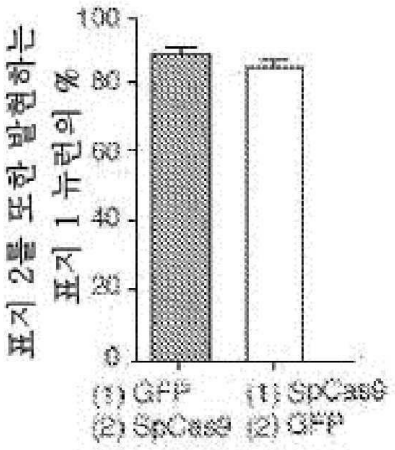
도면76c


HA-SpCas9
Map1b
DAPI
병합
HA-SpCas9
NeuN
DAPI
병합
HA-SpCas9
GFAP
DAPI
병합
HA-SpCas9
GFP-KASH
DAPI
병합

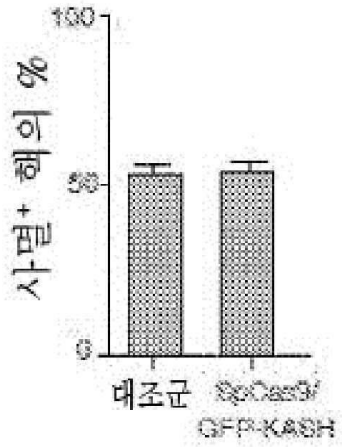
도면76d


GFP
ONM
KASH
INM

도면76e



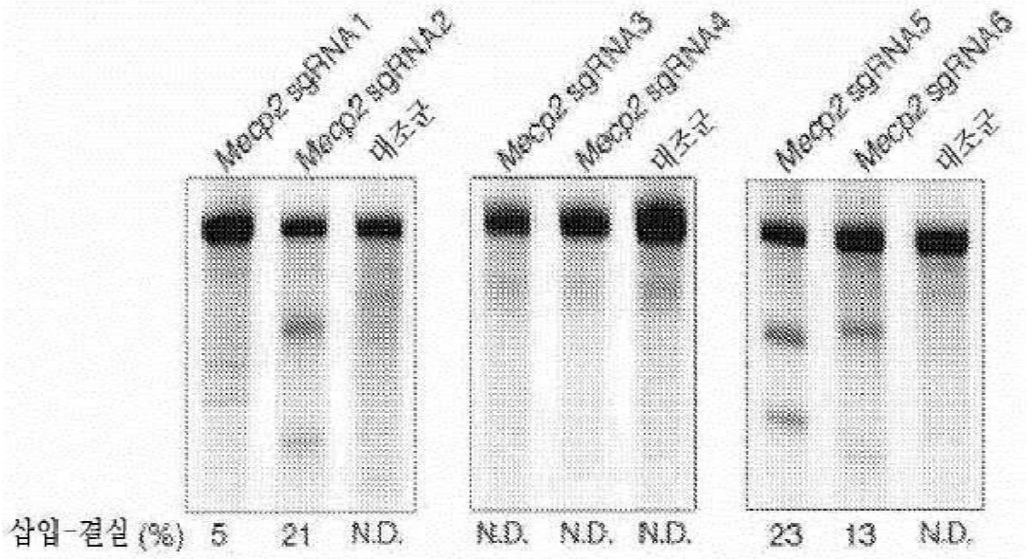
도면76f



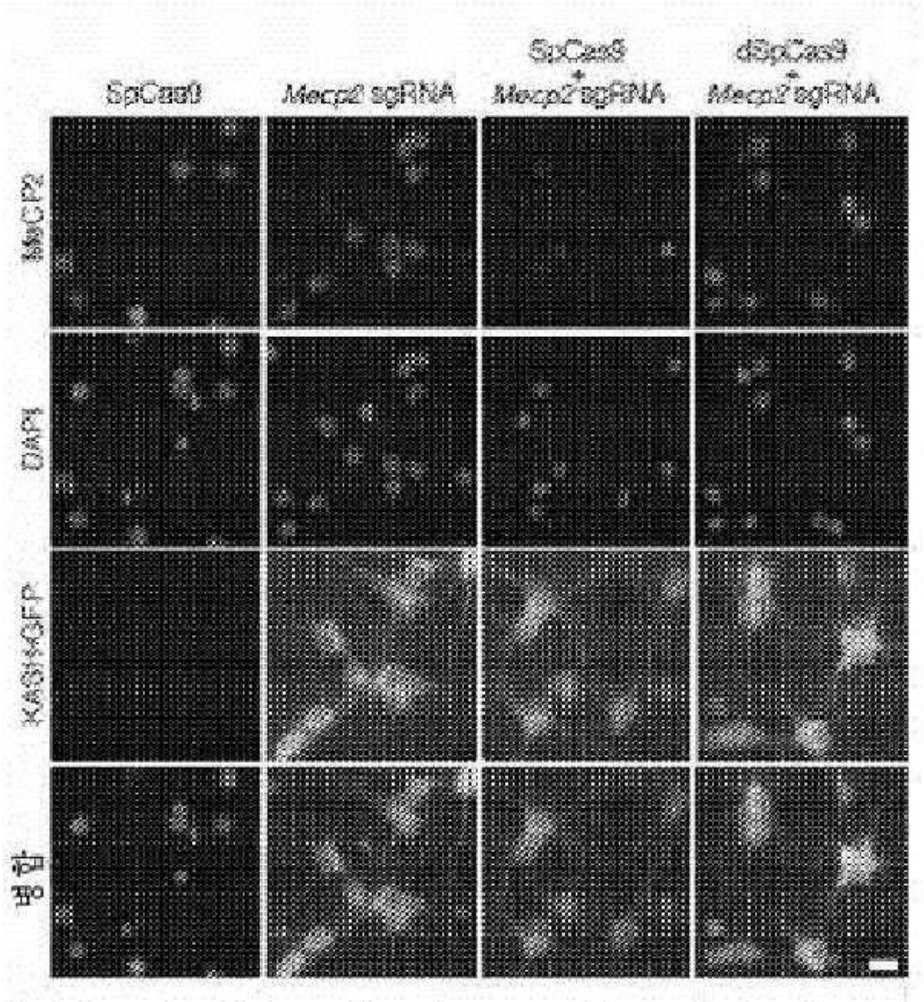
도면77a

	Mecp2 표적 서열	PAM
표적 1	CTGGGAGAGGGAGCCCCTCC	AGG
표적 2	AAAGGTGGGAGACACCTCCT	TGG
표적 3	TCCAACCTTCAGGCAAGGTG	GGG
표적 4	AGGAAGTCTGGCCGATCTGC	TGG
표적 5	CCATTCTGCAGAGCCAGCAG	AGG
표적 6	CTCTGAGGCCCTGGAGATCC	TGG

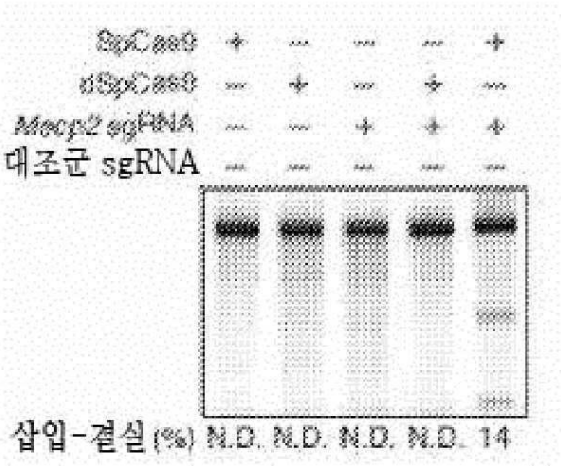
도면77b


Mecp2 sgRNA 1
Mecp2 sgRNA 2
대조군
Mecp2 sgRNA 3
Mecp2 sgRNA 4
대조군
Mecp2 sgRNA 5
Mecp2 sgRNA 6
대조군
삽입-결실 (%)
5
21
N.D.
N.D.
N.D.
N.D.
23
13
N.D.

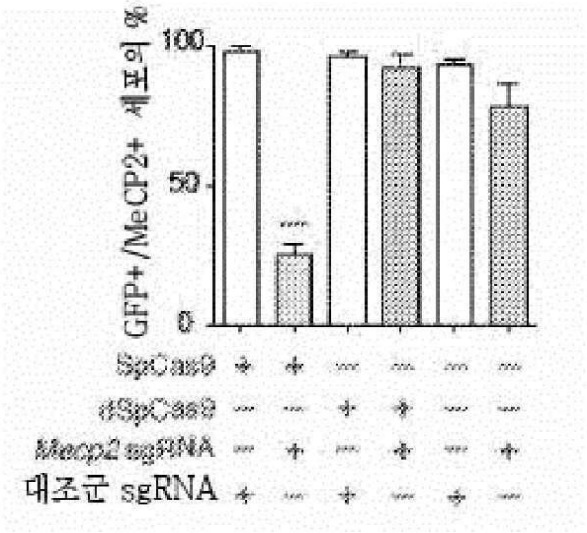
도면78a


SpCas9
Mecp2 sgRNA
SpCas9 + Mecp2 sgRNA
dSpCas9 + Mecp2 sgRNA
MeCP2
DAPI
KASH-GFP
병합

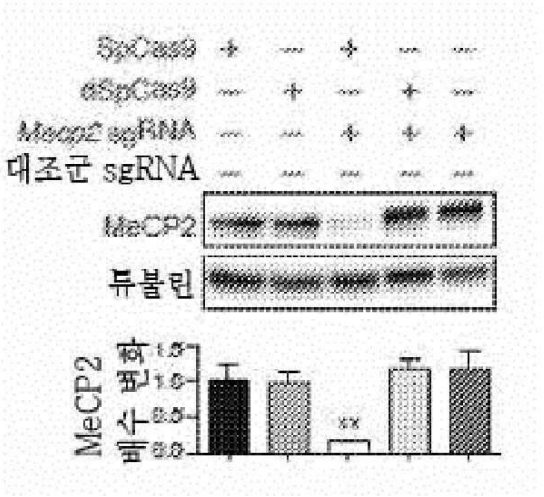
도면78b



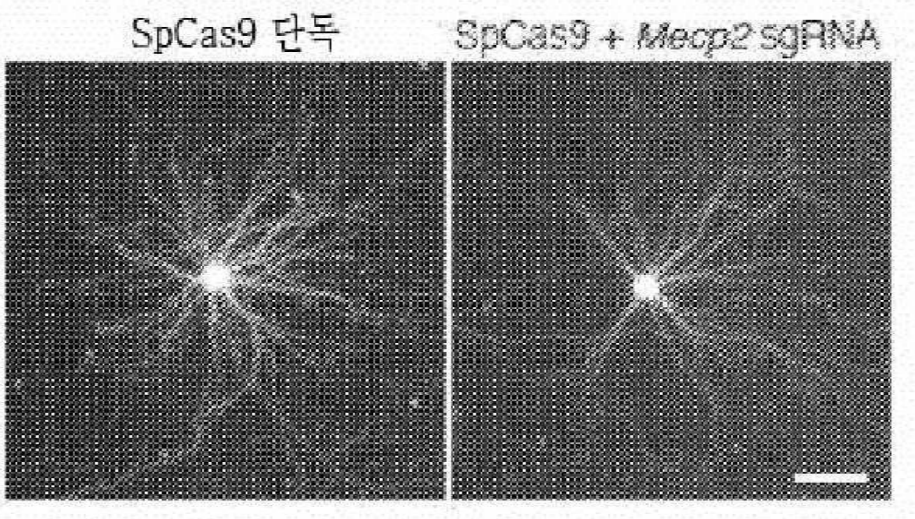
도면78c



도면78d



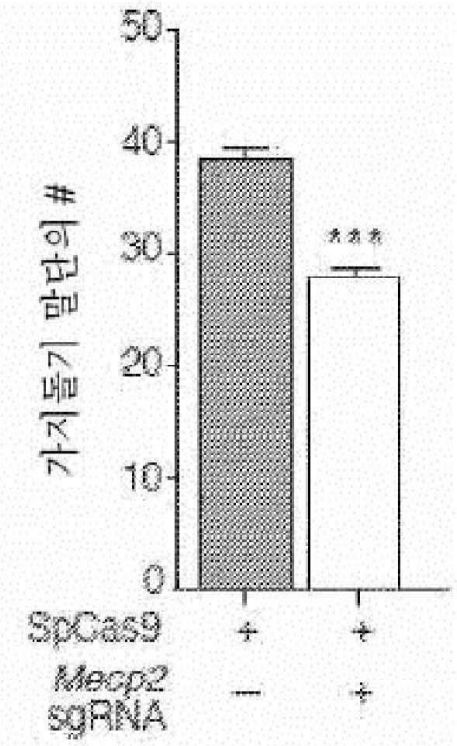
도면79a



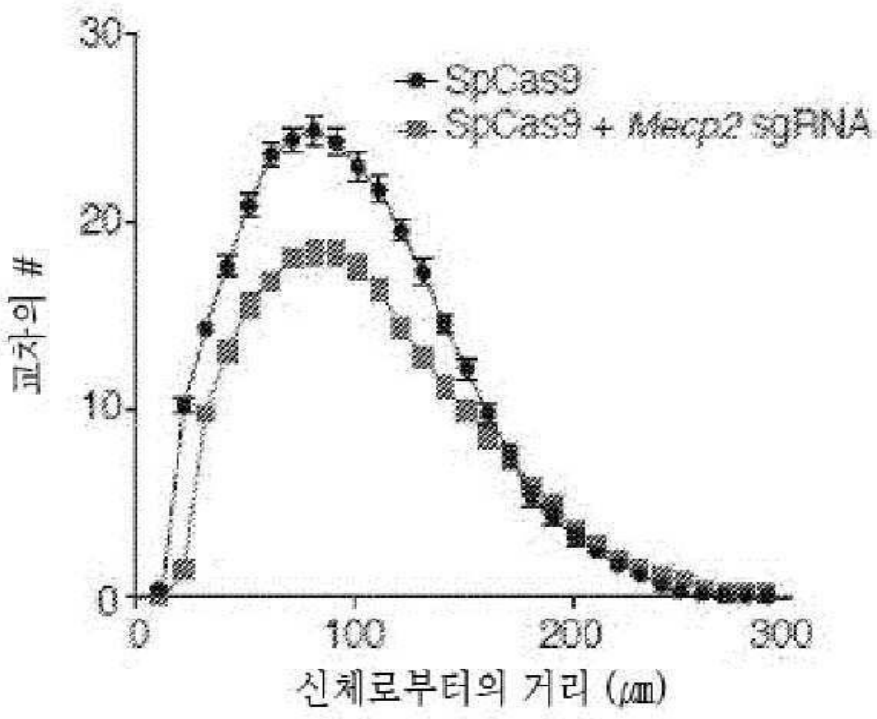
도면79b



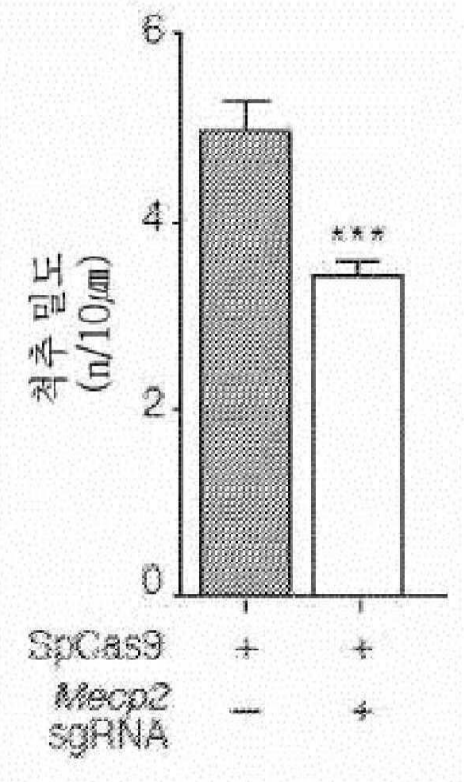
도면79c



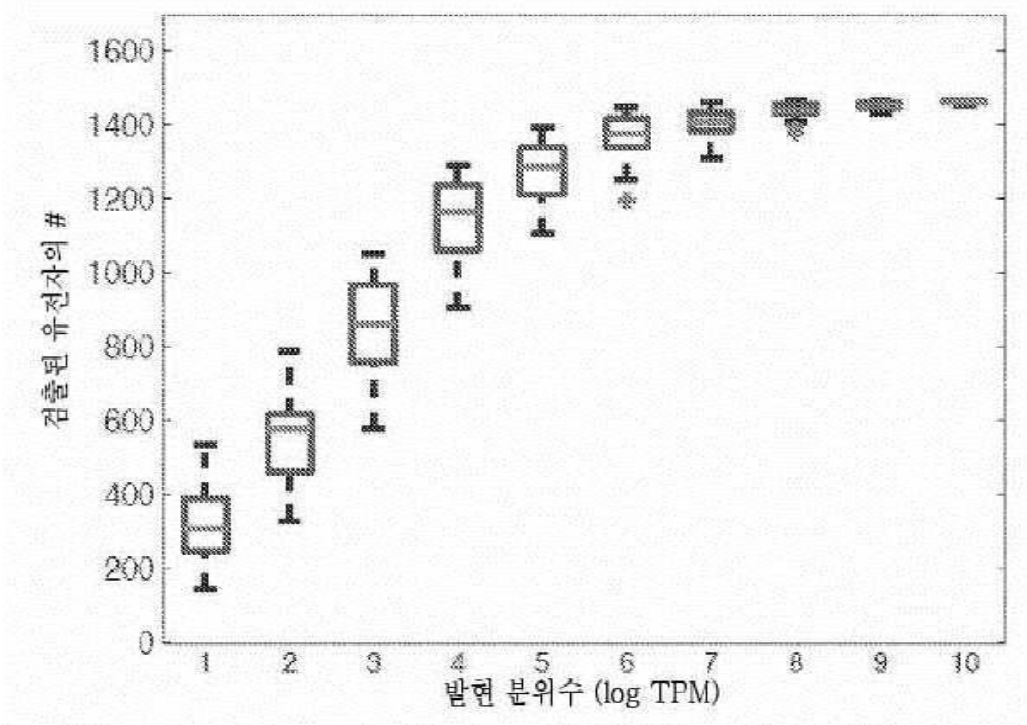
도면79d


30
20
10
0
교차의 #
SpCas9
SpCas9 + Mecp2 sgRNA
0
100
200
300
신체로부터의 거리 (㎛)

도면79e


6
4
2
0
척추 밀도 (n/10㎛)
***
SpCas9
+
+
Mecp2 sgRNA
−
+

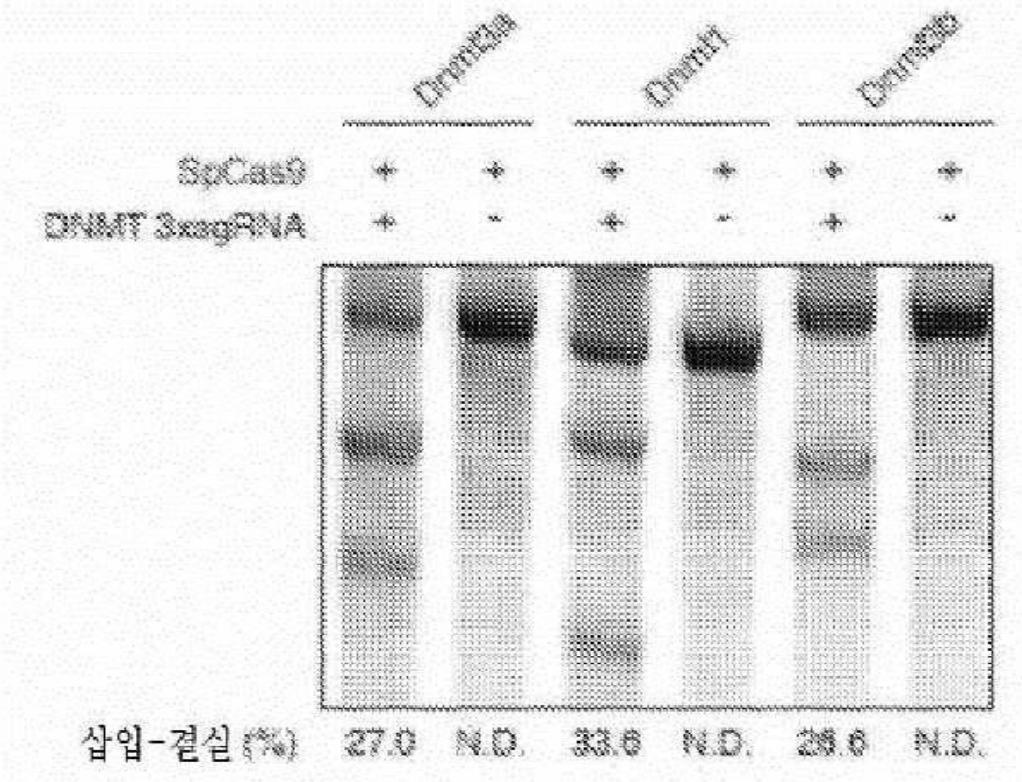
도면80



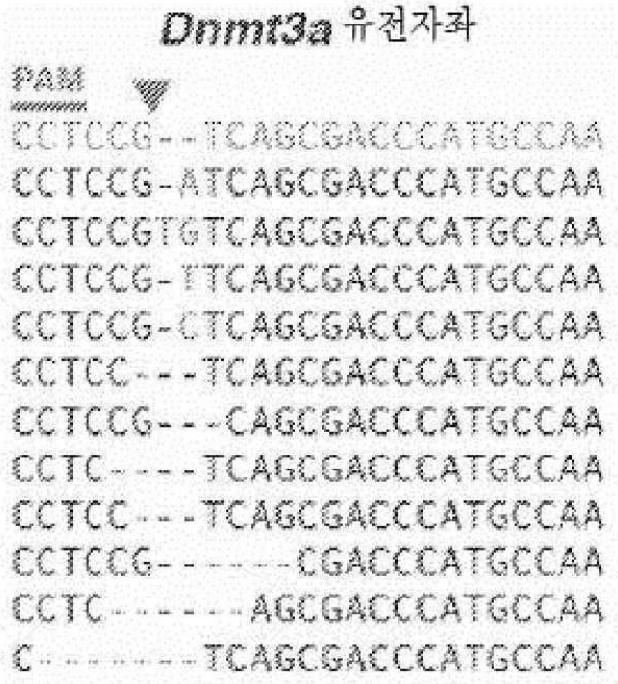
도면81a

표적	표적 서열	PAM
Dnmt3a	TTGGCATGGGTCGCTGACGG	AGG
Dnmt1	CGGGCTGGAGCTGTTCGCGC	TGG
Dnmt3b	AGAGGGTGCCAGCGGGTATG	AGG

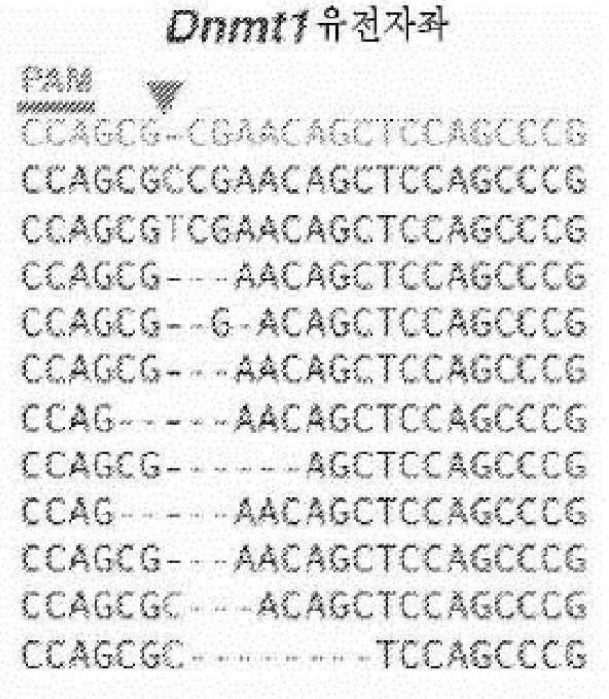
도면81b


Dnmt3a
Dnmt1
Dnmt3b
SpCas9
DNMT 3xsgRNA
삽입-결실 (%)
27.0
N.D.
33.6
N.D.
26.6
N.D.

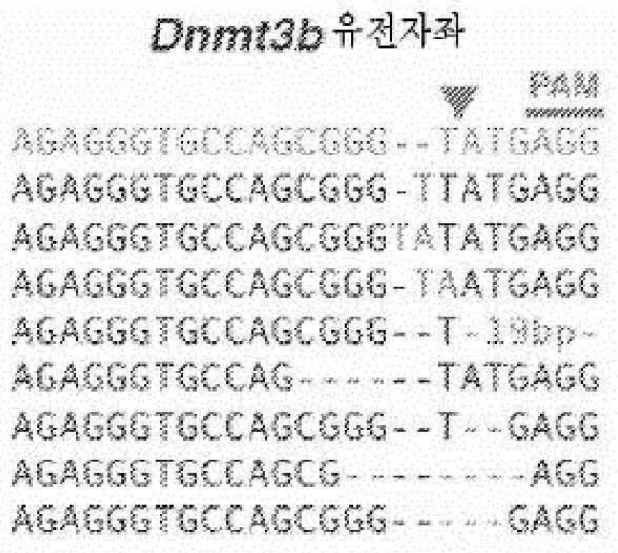
도면82a


Dnmt3a 유전자좌
PAM
CCTCCG--TCAGCGACCCATGCCAA
CCTCCG-ATCAGCGACCCATGCCAA
CCTCCGTGTCAGCGACCCATGCCAA
CCTCCG-TTCAGCGACCCATGCCAA
CCTCCG-CTCAGCGACCCATGCCAA
CCTCC---TCAGCGACCCATGCCAA
CCTCCG---CAGCGACCCATGCCAA
CCTC----TCAGCGACCCATGCCAA
CCTCC---TCAGCGACCCATGCCAA
CCTCCG------CGACCCATGCCAA
CCTC--------AGCGACCCATGCCAA
C---------TCAGCGACCCATGCCAA

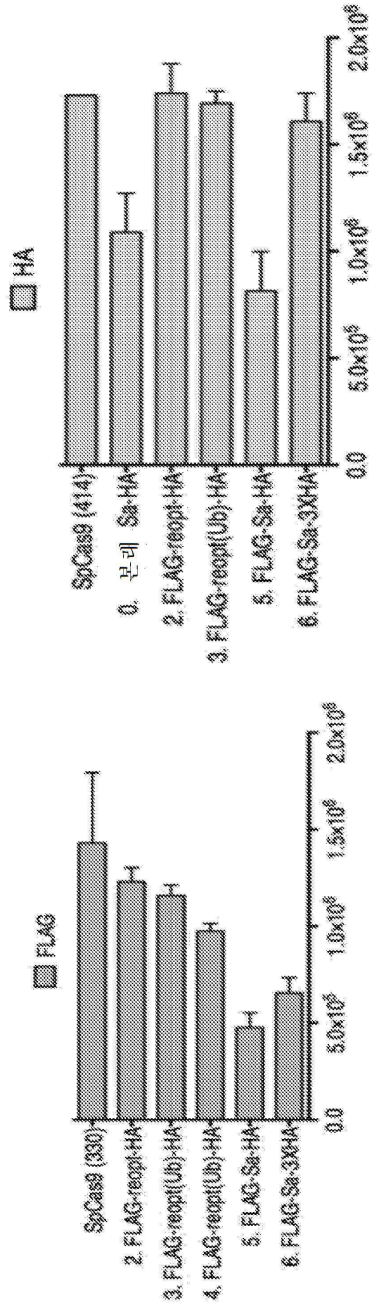
도면82b


Dnmt1 유전자좌
PAM
CCAGCG-CGAACAGCTCCAGCCCG
CCAGCGCCGAACAGCTCCAGCCCG
CCAGCGTCGAACAGCTCCAGCCCG
CCAGCG---AACAGCTCCAGCCCG
CCAGCG--G-ACAGCTCCAGCCCG
CCAGCG---AACAGCTCCAGCCCG
CCAG-----AACAGCTCCAGCCCG
CCAGCG------AGCTCCAGCCCG
CCAG-----AACAGCTCCAGCCCG
CCAGCG---AACAGCTCCAGCCCG
CCAGCGC----ACAGCTCCAGCCCG
CCAGCGC----------TCCAGCCCG

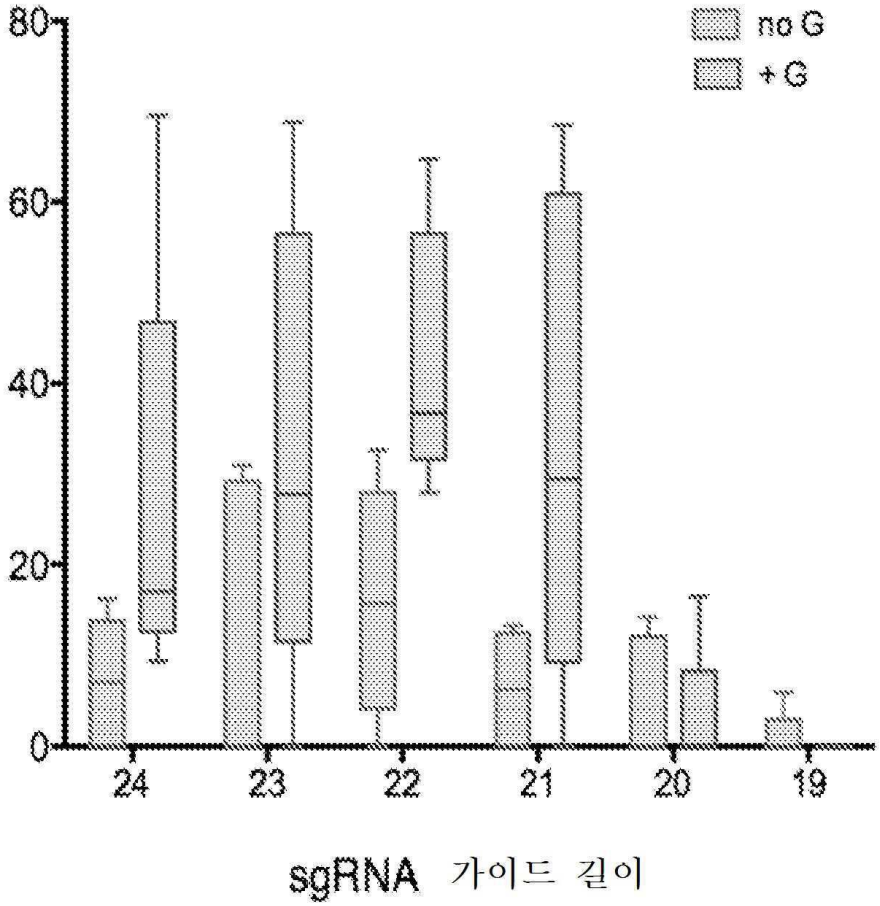
도면82c


Dnmt3b 유전자좌
PAM
AGAGGGTGCCAGCGGG--TATGAGG
AGAGGGTGCCAGCGGG-TTATGAGG
AGAGGGTGCCAGCGGGTATATGAGG
AGAGGGTGCCAGCGGG-TAATGAGG
AGAGGGTGCCAGCGGG--T-19bp-
AGAGGGTGCCAG------TATGAGG
AGAGGGTGCCAGCGGG--T--GAGG
AGAGGGTGCCAGCG--------AGG
AGAGGGTGCCAGCGGG-----GAGG

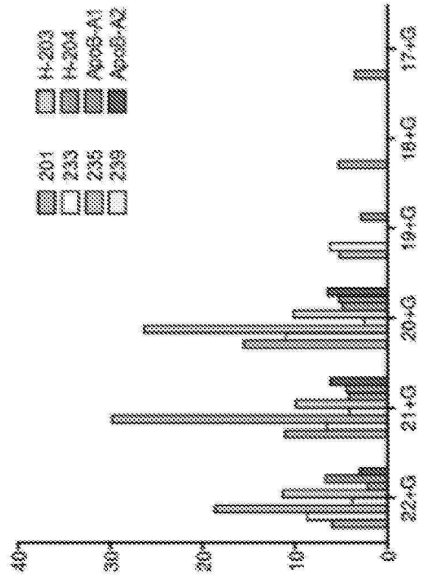
도면83

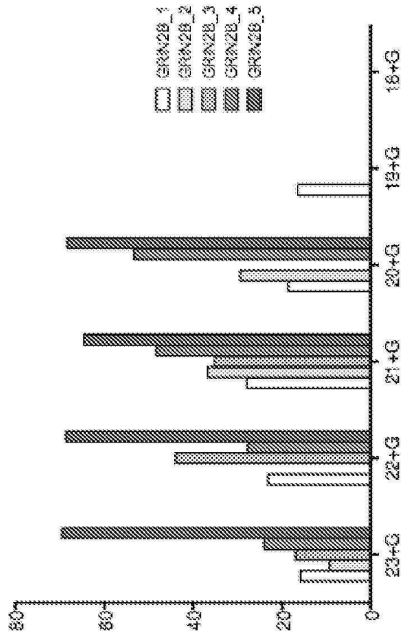

HA
SpCas9 (414)
0. 본래 Sa-HA
2. FLAG-reopt-HA
3. FLAG-reopt(Ub)-HA
5. FLAG-Sa-HA
6. FLAG-Sa-3XHA
0.0
5.0x10^5
1.0x10^6
1.5x10^6
2.0x10^6
FLAG
SpCas9 (930)
2. FLAG-reopt-HA
3. FLAG-reopt(Ub)-HA
4. FLAG-reopt(Ub)-HA
5. FLAG-Sa-HA
6. FLAG-Sa-3XHA
0.0
5.0x10^5
1.0x10^6
1.5x10^6
2.0x10^6

도면84



도면85





서 열 목 록

SEQUENCE LISTING

<110> THE BROAD INSTITUTE INC.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

<120> DELIVERY, USE AND THERAPEUTIC APPLICATIONS OF THE CRISPR-CAS SYSTEMS AND COMPOSITIONS FOR TARGETING DISORDERS AND DISEASES USING VIRAL COMPONENTS

<130> 44790.99.2058

<140> PCT/US2014/041809

<141> 2014-06-11

<150> 61/979,879

<151> 2014-04-15

<150> 61/915,225

<151> 2013-12-12

<150> PCT/US2013/074667

<151> 2013-12-12

<150> 61/871,301

<151> 2013-08-28

<150> 61/862,355

<151> 2013-08-05

<150> 61/847,537

<151> 2013-07-17

<150> 61/836,123

<151> 2013-06-17

<160> 575

<170> PatentIn version 3.5

<210> 1

<211> 27

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
primer"

<400> 1

gcactgaggg cctatttccc atgattc 27

<210> 2

<211> 25

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 2

cctccgtgtc agcgacccat gccaa 25

<210> 3

<211> 24

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 3

ccagcgtcga acagctccag cccg 24

<210> 4

<211> 25

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 4

agagggtgcc agcgggtata tgagg 25

<210> 5

<211> 20

<212> DNA

<213> Homo sapiens

<400> 5

gagtccgagc agaagaagaa 20

<210> 6

<211> 20

<212> DNA

<213> Homo sapiens

<400> 6

gagtcctagc aggagaagaa 20

<210> 7

<211> 20

<212> DNA

<213> Homo sapiens

<400> 7

gagtctaagc agaagaagaa 20

<210

> 8

<211> 60

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<220><221> modified_base

<222> (1)..(20)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (23)..(44)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (47)..(60)

<223> a, c, u, g, unknown or other

<400> 8

nnnnnnnnnn nnnnnnnnnn ccnnnnnnnn nnnnnnnnnn nnnnggnnnn nnnnnnnnnn 60

<210> 9

<211>

60

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<220><221> modified_base

<222> (1)..(14)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (17)..(38)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (41)..(60)

<223> a, c, u, g, unknown or other

<400> 9

nnnnnnnnnn nnnnccnnnn nnnnnnnnnn nnnnnnnngg nnnnnnnnnn nnnnnnnnnn 60

<210> 10

<211> 60

<212

> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<220><221> modified_base

<222> (1)..(20)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (23)..(43)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (46)..(60)

<223> a, c, u, g, unknown or other

<400> 10

nnnnnnnnnn nnnnnnnnnn ccnnnnnnnn nnnnnnnnnn nnnggnnnnn nnnnnnnnnn 60

<210> 11

<211> 60

<212> RNA

<

213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<220><221> modified_base

<222> (1)..(15)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (18)..(38)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (41)..(60)

<223> a, c, u, g, unknown or other

<400> 11

nnnnnnnnnn nnnnnccnnn nnnnnnnnnn nnnnnnnngg nnnnnnnnnn nnnnnnnnnn 60

<210> 12

<211> 60

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (1)..(20)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (23)..(42)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (45)..(60)

<223> a, c, u, g, unknown or other

<400> 12

nnnnnnnnnn nnnnnnnnnn ccnnnnnnnn nnnnnnnnnn nnggnnnnnn nnnnnnnnnn 60

<210> 13

<211> 60

<212> RNA

<213> Artificial Sequence

<220><221

> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (1)..(15)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (18)..(38)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (41)..(60)

<223> a, c, u, g, unknown or other

<400> 13

nnnnnnnnnn nnnnnccnnn nnnnnnnnnn nnnnnnnngg nnnnnnnnnn nnnnnnnnnn 60

<210> 14

<211> 60

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (1)..(20)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (23)..(41)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (44)..(60)

<223> a, c, u, g, unknown or other

<400> 14

nnnnnnnnnn nnnnnnnnnn ccnnnnnnnn nnnnnnnnnn nggnnnnnnn nnnnnnnnnn 60

<210> 15

<211> 60

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<220><221> modified_base

<222> (1)..(17)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (20)..(38)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (41)..(60)

<223> a, c, u, g, unknown or other

<400> 15

nnnnnnnnnn nnnnnnnccn nnnnnnnnnn nnnnnnnngg nnnnnnnnnn nnnnnnnnnn 60

<210> 16

<211> 60

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<220><221> modified_base

<222> (1)..(20)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (23)..(40)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (43)..(60)

<223> a, c, u, g, unknown or other

<400> 16

nnnnnnnnnn nnnnnnnnnn ccnnnnnnnn nnnnnnnnnn ggnnnnnnnn nnnnnnnnnn 60

<210> 17

<211> 60

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221
> modified_base

<222> (1)..(18)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (21)..(38)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (41)..(60)

<223> a, c, u, g, unknown or other

<400> 17

nnnnnnnnnn nnnnnnnncc nnnnnnnnnn nnnnnnnngg nnnnnnnnnn nnnnnnnnnn 60

<210> 18

<211> 60

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (1)..(20)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (23)..(39)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (42)..(60)

<223> a, c, u, g, unknown or other

<400> 18

nnnnnnnnnn nnnnnnnnnn ccnnnnnnnn nnnnnnnnng gnnnnnnnnn nnnnnnnnnn 60

<210> 19

<211> 60

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (1)..(19)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (22)..(38)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (41)..(60)

<223> a, c, u, g, unknown or other

<400> 19

nnnnnnnnnn nnnnnnnnnc cnnnnnnnnn nnnnnnnngg nnnnnnnnnn nnnnnnnnnn 60

<210> 20

<211> 60

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (1)..(20)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (23)..(38)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (41)..(60)

<223> a, c, u, g, unknown or other

<400> 20

nnnnnnnnnn nnnnnnnnnn ccnnnnnnnn nnnnnnnngg nnnnnnnnnn nnnnnnnnnn 60

<210> 21

<211> 60

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (1)..(20)

<223> a, c, u, g, unknown or other

<220><221
> modified_base

<222> (23)..(38)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (41)..(60)

<223> a, c, u, g, unknown or other

<400> 21

nnnnnnnnnn nnnnnnnnnn ccnnnnnnnn nnnnnnnngg nnnnnnnnnn nnnnnnnnnn 60

<210> 22

<211> 60

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (1)..(20)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (23)..(37)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (40)..(60)

<223> a, c, u, g, unknown or other

<400> 22

nnnnnnnnnn nnnnnnnnnn ccnnnnnnnn nnnnnnnggn nnnnnnnnnn nnnnnnnnnn 60

<210> 23

<211> 60

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (1)..(21)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (24)..(38)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (41)..(60)

<223> a, c, u, g, unknown or other

<400> 23

nnnnnnnnnn nnnnnnnnnn nccnnnnnnn nnnnnnnngg nnnnnnnnnn nnnnnnnnnn 60

<210> 24

<211> 60

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<220><221> modified_base

<222> (1)..(20)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (23)..(36)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (39)..(60)

<223> a, c, u, g, unknown or other

<400> 24

nnnnnnnnnn nnnnnnnnnn ccnnnnnnnn nnnnnnggnn nnnnnnnnnn nnnnnnnnnn 60

<210> 25

<211> 60

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<220><221> modified_base

<222> (1)..(22)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (25)..(38)

<223> a, c, u, g, unknown or other

<220><221

> modified_base

<222> (41)..(60)

<223> a, c, u, g, unknown or other

<400> 25

nnnnnnnnnn nnnnnnnnnn nnccnnnnnn nnnnnnnngg nnnnnnnnnn nnnnnnnnnn 60

<210> 26

<211> 60

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (1)..(20)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (23)..(35)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (38)..(60)

<223> a, c, u, g, unknown or other

<400> 26

nnnnnnnnnn nnnnnnnnnn ccnnnnnnnn nnnnnggnnn nnnnnnnnnn nnnnnnnnnn 60

<210> 27

<211> 60

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (1)..(23)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (26)..(38)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (41)..(60)

<223> a, c, u, g, unknown or other

<400> 27

nnnnnnnnnn nnnnnnnnnn nnnccnnnnn nnnnnnnngg nnnnnnnnnn nnnnnnnnnn 60

<210> 28

<211> 60

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (1)..(20)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (23)..(34)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (37)..(60)

<223> a, c, u, g, unknown or other

<400> 28

nnnnnnnnnn nnnnnnnnnn ccnnnnnnnn nnnnggnnnn nnnnnnnnnn nnnnnnnnnn 60

<210> 29

<211> 60

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (1)..(24)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (27)..(38)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (41)..(60)

<223> a, c, u, g, unknown or other

<400> 29

nnnnnnnnnn nnnnnnnnnn nnnnccnnnn nnnnnnnngg nnnnnnnnnn nnnnnnnnnn 60

<210> 30

<211> 60

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (1)..(20)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (23)..(33)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (36)..(60)

<223> a, c, u, g, unknown or other

<400> 30

nnnnnnnnnn nnnnnnnnnn ccnnnnnnnn nnnggnnnnn nnnnnnnnnn nnnnnnnnnn 60

<210> 31

<211> 60

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (1)..(25)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (28)..(38)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (41)..(60)

<223> a, c, u, g, unknown or other

<400> 31

nnnnnnnnnn nnnnnnnnnn nnnnnccnnn nnnnnnnngg nnnnnnnnnn nnnnnnnnnn 60

<210> 32

<211> 60

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (1)..(20)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (23)..(32)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (35)..(60)

<223> a, c, u, g, unknown or other

<400> 32

nnnnnnnnnn nnnnnnnnnn ccnnnnnnnn nnggnnnnnn nnnnnnnnnn nnnnnnnnnn 60

<210> 33

<211> 60

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (1)..(26)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (29)..(38)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (41)..(60)

<223> a, c, u, g, unknown or other

<400> 33

nnnnnnnnnn nnnnnnnnnn nnnnnnccnn nnnnnnnngg nnnnnnnnnn nnnnnnnnnn 60

<210> 34

<211> 60

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<220><221> modified_base

<222> (1)..(20)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (23)..(31)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (34)..(60)

<223> a, c, u, g, unknown or other

<400> 34

nnnnnnnnnn nnnnnnnnnn ccnnnnnnnn nggnnnnnnn nnnnnnnnnn nnnnnnnnnn 60

<210> 35

<211> 60

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<220><221> modified_base

<222> (1)..(27)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (30)..(38)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (41)..(60)

<223> a, c, u, g, unknown or other

<400> 35

nnnnnnnnnn nnnnnnnnnn nnnnnnnccn nnnnnnnngg nnnnnnnnnn nnnnnnnnnn 60

<210> 36

<211> 60

<212> RNA

<213> Artificial Sequence

<220><221

> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<220><221> modified_base

<222> (1)..(20)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (23)..(30)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (33)..(60)

<223> a, c, u, g, unknown or other

<400> 36

nnnnnnnnnn nnnnnnnnnn ccnnnnnnnn ggnnnnnnnn nnnnnnnnnn nnnnnnnnnn 60

<210> 37

<211> 60

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"
<220><221> modified_base
<222> (1)..(28)
<223> a, c, u, g, unknown or other
<220><221> modified_base
<222> (31)..(38)
<223> a, c, u, g, unknown or other
<220><221> modified_base
<222> (41)..(60)
<223> a, c, u, g, unknown or other
<400> 37
nnnnnnnnnn nnnnnnnnnn nnnnnnnncc nnnnnnnngg nnnnnnnnnn nnnnnnnnnn 60
<210> 38
<211> 60
<212> RNA
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"
<220><221> modified_base
<222> (1)..(20)
<223> a, c, u, g, unknown or other
<220><221> modified_base
<222> (23)..(29)
<223> a, c, u, g, unknown or other
<220><221> modified_base
<222> (32)..(60)
<223> a, c, u, g, unknown or other
<400> 38
nnnnnnnnnn nnnnnnnnnn ccnnnnnnng gnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 60
<210> 39

<211> 60

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (1)..(29)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (32)..(38)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (41)..(60)

<223> a, c, u, g, unknown or other

<400> 39

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnc cnnnnnnngg nnnnnnnnnn nnnnnnnnnn 60

<210> 40

<211> 60

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221
> modified_base

<222> (1)..(20)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (23)..(28)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (31)..(60)

<223> a, c, u, g, unknown or other

<400> 40

nnnnnnnnnn nnnnnnnnnn ccnnnnnngg nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 60

<210> 41

<211> 60

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (1)..(30)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (33)..(38)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (41)..(60)

<223> a, c, u, g, unknown or other

<400> 41

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn ccnnnnnngg nnnnnnnnnn nnnnnnnnnn 60

<210> 42

<211> 60

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (1)..(20)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (23)..(27)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (30)..(60)

<223> a, c, u, g, unknown or other

<400> 42

nnnnnnnnnn nnnnnnnnnn ccnnnnnggn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 60

<210> 43

<211> 60

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (1)..(31)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (34)..(38)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (41)..(60)

<223> a, c, u, g, unknown or other

<400> 43

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nccnnnnngg nnnnnnnnnn nnnnnnnnnn 60

<210> 44

<211> 60

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (1)..(20)

<223> a, c, u, g, unknown or other

<220><221

> modified_base

<222> (23)..(26)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (29)..(60)

<223> a, c, u, g, unknown or other

<400> 44

nnnnnnnnnn nnnnnnnnnn ccnnnnggnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 60

<210> 45

<211> 60

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (1)..(32)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (35)..(38)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (41)..(60)

<223> a, c, u, g, unknown or other

<400> 45

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnccnnnngg nnnnnnnnnn nnnnnnnnnn 60

<210> 46

<211> 60

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (1)..(20)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (23)..(25)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (28)..(60)

<223> a, c, u, g, unknown or other

<400> 46

nnnnnnnnnn nnnnnnnnnn ccnnnggnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 60

<210> 47

<211> 60

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<220><221> modified_base

<222> (1)..(33)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (36)..(38)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (41)..(60)

<223> a, c, u, g, unknown or other

<400> 47

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnccnnngg nnnnnnnnnn nnnnnnnnnn 60

<210> 48

<211> 60

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"
<220><221> modified_base
<222> (1)..(20)
<223> a, c, u, g, unknown or other
<220><221> modified_base
<222> (23)..(24)
<223> a, c, u, g, unknown or other
<220><221
> modified_base
<222> (27)..(60)
<223> a, c, u, g, unknown or other
<400> 48
nnnnnnnnnn nnnnnnnnnn ccnnggnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 60
<210> 49
<211> 60
<212> RNA
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"
<220><221> modified_base
<222> (1)..(34)
<223> a, c, u, g, unknown or other
<220><221> modified_base
<222> (37)..(38)
<223> a, c, u, g, unknown or other
<220><221> modified_base

<222> (41)..(60)
<223> a, c, u, g, unknown or other
<400> 49
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnccnngg nnnnnnnnnn nnnnnnnnnn 60
<210> 50
<211> 60

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (1)..(20)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (23)..(23)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (26)..(60)

<223> a, c, u, g, unknown or other

<400> 50

nnnnnnnnnn nnnnnnnnnn ccnggnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 60

<210> 51

<211> 60

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (1)..(35)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (38)..(38)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (41)..(60)

<223> a, c, u, g, unknown or other

<400> 51

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnccngg nnnnnnnnnn nnnnnnnnnn 60

<210> 52

<211> 60

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (1)..(21)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (26)..(60)

<223> a, c, u, g, unknown or other

<400> 52

nnnnnnnnnn nnnnnnnnnn nccggnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 60

<210

> 53

<211> 60

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (1)..(35)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (40)..(60)

<223> a, c, u, g, unknown or other

<400> 53

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnccggn nnnnnnnnnn nnnnnnnnnn 60

<210> 54

<211> 60

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<220><221> modified_base

<222> (1)..(23)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (26)..(60)

<223> a, c, u, g, unknown or other

<400> 54

nnnnnnnnnn nnnnnnnnnn nnnggnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 60

<210> 55

<211> 60

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<220><221> modified_base

<222> (1)..(33)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (38)..(60)

<223> a, c, u, g, unknown or other

<400> 55

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnggccnnn nnnnnnnnnn nnnnnnnnnn 60

<210> 56

<211> 60

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<220><221> modified_base

<222> (1)..(22)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (26)..(60)

<223> a, c, u, g, unknown or other

<400> 56

nnnnnnnnnn nnnnnnnnnn nncggnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 60

<210> 57

<211> 60

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<220><221> modified_base

<222> (1)..(32)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (35)..(35)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (38)..(60)

<223> a, c, u, g, unknown or other

<400> 57

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnggnccnnn nnnnnnnnnn nnnnnnnnnn 60

<210> 58

<211> 60

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<220><221> modified_base

<222> (1)..(31)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (34)..(35)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (38)..(60)

<223> a, c, u, g, unknown or other

<400> 58

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nggnnccnnn nnnnnnnnnn nnnnnnnnnn 60

<210> 59

<211> 60

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<220><221> modified_base

<222> (1)..(30)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (33)..(35)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (38)..(60)

<223> a, c, u, g, unknown or other

<400> 59

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn ggnnnccnnn nnnnnnnnnn nnnnnnnnnn 60

<210> 60

<211> 60

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (1)..(29)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (32)..(35)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (38)..(60)

<223> a, c, u, g, unknown or other

<400> 60

nnnnnnnnnn nnnnnnnnnn nnnnnnnnng gnnnnccnnn nnnnnnnnnn nnnnnnnnnn 60

<210> 61

<211> 60

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (1)..(28)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (31)..(35)

<223> a, c, u, g, unknown or other

<220><221> modified_base

<222> (38)..(60)

<223> a, c, u, g, unknown or other

<400> 61

nnnnnnnnnn nnnnnnnnnn nnnnnnnngg nnnnnccnnn nnnnnnnnnn nnnnnnnnnn 60

<210> 62

<211> 9

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:
'LAGLIDADG' family motif peptide"

<400> 62

Leu Ala Gly Leu Ile Asp Ala Asp Gly
1 5

<210> 63

<211> 12

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 63

guuuuagagc ua 12

<210> 64

<211> 7

<212> PRT

<213> Simian virus 40

<400> 64

Pro Lys Lys Lys Arg Lys Val
1 5

<210> 65

<211> 16

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:
Nucleoplasmin bipartite NLS sequence"

<400> 65

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

1 5 10 15

<210> 66

<211> 9

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

C-myc NLS sequence"

<400> 66

Pro Ala Ala Lys Arg Val Lys Leu Asp

1 5

<210> 67

<211> 11

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

C-myc NLS sequence"

<400> 67

Arg Gln Arg Arg Asn Glu Leu Lys Arg Ser Pro

1 5 10

<210> 68

<211> 38

<212> PRT

<213> Homo sapiens

<400> 68

Asn Gln Ser Ser Asn Phe Gly Pro Met Lys Gly Gly Asn Phe Gly Gly

1 5 10 15

Arg Ser Ser Gly Pro Tyr Gly Gly Gly Gly Gln Tyr Phe Ala Lys Pro

20 25 30

Arg Asn Gln Gly Gly Tyr

35

<210> 69

<211> 42

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

IBB domain from importin-alpha sequence"

<400> 69

Arg Met Arg Ile Glx Phe Lys Asn Lys Gly Lys Asp Thr Ala Glu Leu

1 5 10 15

Arg Arg Arg Arg Val Glu Val Ser Val Glu Leu Arg Lys Ala Lys Lys

20 25 30

Asp Glu Gln Ile Leu Lys Arg Arg Asn Val

35 40

<210> 70

<211> 8

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Myoma T protein sequence"

<400> 70

Val Ser Arg Lys Arg Pro Arg Pro

1 5

<210> 71

<211> 8

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Myoma T protein sequence"

<400> 71

Pro Pro Lys Lys Ala Arg Glu Asp

1 5

<210> 72

<211> 8

<212> PRT

<213> Homo sapiens

<400> 72

Pro Gln Pro Lys Lys Lys Pro Leu

1 5

<210> 73

<211> 12

<212> PRT

<213> Mus musculus

<400> 73

Ser Ala Leu Ile Lys Lys Lys Lys Lys Met Ala Pro

1 5 10

<210> 74

<211> 5

<212> PRT

<213> Influenza virus

<400> 74

Asp Arg Leu Arg Arg

1 5

<210> 75

<211> 7

<212> PRT

<213> Influenza virus

<400> 75

Pro Lys Gln Lys Lys Arg Lys

1 5

<210> 76

<211> 10

<212> PRT

<213> Hepatitis delta virus

<400> 76

Arg Lys Leu Lys Lys Lys Ile Lys Lys Leu

1 5 10

<210> 77

<211> 10

<212> PRT

<213> Mus musculus

<400> 77

Arg Glu Lys Lys Lys Phe Leu Lys Arg Arg

1 5 10

<210> 78

<211> 20

<212> PRT

<213> Homo sapiens

<400> 78

Lys Arg Lys Gly Asp Glu Val Asp Gly Val Asp Glu Val Ala Lys Lys

1 5 10 15

Lys Ser Lys Lys

20

<210> 79

<211> 17

<212> PRT

<213> Homo sapiens

<400> 79

Arg Lys Cys Leu Gln Ala Gly Met Asn Leu Glu Ala Arg Lys Thr Lys

1 5 10 15

Lys

<210> 80

<211> 27

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<220><221> modified_base

<222> (1)..(20)

<223> a, c, t or g

<220><221> modified_base

<222> (21)..(22)

<223

> a, c, t, g, unknown or other

<400> 80

nnnnnnnnnn nnnnnnnnnn nnagaaw 27

<210> 81

<211> 19

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<220><221> modified_base

<222> (1)..(12)

<223> a, c, t or g

<220><221> modified_base

<222> (13)..(14)

<223> a, c, t, g, unknown or other

<400> 81

nnnnnnnnnn nnnnagaaw 19

<210> 82

<211> 27

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<220><221> modified_base

<222> (1)..(20)

<223> a, c, t or g

<220><221> modified_base

<222> (21)..(22)

<223> a, c, t, g, unknown or other

<400> 82

nnnnnnnnnn nnnnnnnnnn nnagaaw 27

<210> 83

<211> 18

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<220><221> modified_base

<222> (1)..(11)

<223> a, c, t or g

<220><221> modified_base

<222> (12)..(13)

<223> a, c, t, g, unknown or other

<400> 83

nnnnnnnnnn nnnagaaw 18

<210> 84

<211> 137

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<220><221> modified_base

<222> (1)..(20)

<223> a, c, t, g, unknown or other

<

400> 84

nnnnnnnnnn nnnnnnnnnn gtttttgtac tctcaagatt tagaaataaa tcttgcagaa 60

gctacaaaga taaggcttca tgccgaaatc aacaccctgt cattttatgg cagggtgttt 120

tcgttattta atttttt 137

<210> 85

<211> 123

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<220><221> modified_base

<222> (1)..(20)

<223> a, c, t, g, unknown or other

<400> 85

nnnnnnnnnn nnnnnnnnnn gtttttgtac tctcagaaat gcagaagcta caaagataag 60

gcttcatgcc gaaatcaaca ccctgtcatt ttatggcagg gtgttttcgt tatttaattt 120

ttt 123

<210> 86

<211> 110

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<220><221> modified_base

<222> (1)..(20)

<223> a, c, t, g, unknown or other

<400> 86

nnnnnnnnnn nnnnnnnnnn gtttttgtac tctcagaaat gcagaagcta caaagataag 60

gcttcatgcc gaaatcaaca ccctgtcatt ttatggcagg gtgttttttt 110

<210> 87

<211> 102

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<220><221> modified_base

<222> (1)..(20)

<223> a, c, t, g, unknown or other

<400> 87

nnnnnnnnnn nnnnnnnnnn gttttagagc tagaaatagc aagttaaaat aaggctagtc 60

cgttatcaac ttgaaaaagt ggcaccgagt cggtgctttt tt 102

<210> 88

<211> 88

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<220><221> modified_base

<222> (1)..(20)

<223> a, c, t, g, unknown or other

<400> 88

nnnnnnnnnn nnnnnnnnnn gttttagagc tagaaatagc aagttaaaat aaggctagtc 60

cgttatcaac ttgaaaaagt gttttttt 88

<210> 89

<211> 76

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<220><221> modified_base

<222> (1)..(20)

<223> a, c, t, g, unknown or other

<400> 89

nnnnnnnnnn nnnnnnnnnn gttttagagc tagaaatagc aagttaaaat aaggctagtc 60

cgttatcatt tttttt 76

<210> 90

<211> 12

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 90

gttttagagc ta 12

<210> 91

<211> 31

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 91

tagcaagtta aaataaggct agtccgtttt t 31

<210> 92

<211> 27

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<220><221> modified_base

<222> (1)..(22)

<223> a, c, t, g, unknown or other

<400> 92

nnnnnnnnnn nnnnnnnnnn nnagaaw 27

<210> 93

<211> 33

<212> DNA

<213> Homo sapiens

<400> 93

ggacatcgat gtcacctcca atgactaggg tgg 33

<210> 94

<211> 33

<212> DNA

<213> Homo sapiens

<400> 94

cattggaggt gacatcgatg tcctccccat tgg 33

<210> 95

<211> 33

<212> DNA

<213> Homo sapiens

<400> 95

ggaagggcct gagtccgagc agaagaagaa ggg 33

<210> 96

<211> 33

<212> DNA

<213> Homo sapiens

<400> 96

ggtggcgaga ggggccgaga ttgggtgttc agg 33

<210> 97

<211> 33

<212> DNA

<213> Homo sapiens

<400> 97

atgcaggagg gtggcgagag gggccgagat tgg 33

<210> 98

<211> 32

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"
<400> 98
aaactctaga gagggcctat ttcccatgat tc 32
<210> 99
<211> 153
<212> DNA
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"
<400> 99
acctctagaa aaaaagcacc gactcggtgc cacttttca agttgataac ggactagcct 60
tattttaact tgctatgctg ttttgtttcc aaaacagcat agctctaaaa cccctagtca 120
ttggaggtga cggtgtttcg tcctttccac aag 153
<210> 100
<211> 52
<212> DNA
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"
<400> 100
taatacgact cactatagga agtgcgccac catggcccca aagaagaagc gg 52

<210> 101
<211> 60
<212> DNA
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"
<400> 101
ggtttttttt ttttttttt ttttttttt ttttcttact ttttcttttt tgcctggccg 60

<210> 102

<211> 984

<212> PRT

<213> Campylobacter jejuni

<400> 102

Met Ala Arg Ile Leu Ala Phe Asp Ile Gly Ile Ser Ser Ile Gly Trp
1 5 10 15
Ala Phe Ser Glu Asn Asp Glu Leu Lys Asp Cys Gly Val Arg Ile Phe

20 25 30
Thr Lys Val Glu Asn Pro Lys Thr Gly Glu Ser Leu Ala Leu Pro Arg
35 40 45
Arg Leu Ala Arg Ser Ala Arg Lys Arg Leu Ala Arg Arg Lys Ala Arg
50 55 60
Leu Asn His Leu Lys His Leu Ile Ala Asn Glu Phe Lys Leu Asn Tyr
65 70 75 80
Glu Asp Tyr Gln Ser Phe Asp Glu Ser Leu Ala Lys Ala Tyr Lys Gly

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

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

Lys Lys Gln Arg Glu Phe Gly Phe Ser Phe Ser Lys Lys Phe Glu Glu

210 215 220

Glu Val Leu Ser Val Ala Phe Tyr Lys Arg Ala Leu Lys Asp Phe Ser
225 230 235 240
His Leu Val Gly Asn Cys Ser Phe Phe Thr Asp Glu Lys Arg Ala Pro
245 250 255
Lys Asn Ser Pro Leu Ala Phe Met Phe Val Ala Leu Thr Arg Ile Ile
260 265 270
Asn Leu Leu Asn Asn Leu Lys Asn Thr Glu Gly Ile Leu Tyr Thr Lys

275 280 285

Asp Asp Leu Asn Ala Leu Leu Asn Glu Val Leu Lys Asn Gly Thr Leu
290 295 300
Thr Tyr Lys Gln Thr Lys Lys Leu Leu Gly Leu Ser Asp Asp Tyr Glu
305 310 315 320
Phe Lys Gly Glu Lys Gly Thr Tyr Phe Ile Glu Phe Lys Lys Tyr Lys
325 330 335
Glu Phe Ile Lys Ala Leu Gly Glu His Asn Leu Ser Gln Asp Asp Leu

340 345 350

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

405 410 415

Ala Cys Asn Glu Leu Asn Leu Lys Val Ala Ile Asn Glu Asp Lys Lys
420 425 430
Asp Phe Leu Pro Ala Phe Asn Glu Thr Tyr Tyr Lys Asp Glu Val Thr
435 440 445
Asn Pro Val Val Leu Arg Ala Ile Lys Glu Tyr Arg Lys Val Leu Asn

450 455 460

Ala Leu Leu Lys Lys Tyr Gly Lys Val His Lys Ile Asn Ile Glu Leu

465 470 475 480

Ala Arg Glu Val Gly Lys Asn His Ser Gln Arg Ala Lys Ile Glu Lys

485 490 495

Glu Gln Asn Glu Asn Tyr Lys Ala Lys Lys Asp Ala Glu Leu Glu Cys

500 505 510

Glu Lys Leu Gly Leu Lys Ile Asn Ser Lys Asn Ile Leu Lys Leu Arg

515 520 525

Leu Phe Lys Glu Gln Lys Glu Phe Cys Ala Tyr Ser Gly Glu Lys Ile

530 535 540

Lys Ile Ser Asp Leu Gln Asp Glu Lys Met Leu Glu Ile Asp His Ile

545 550 555 560

Tyr Pro Tyr Ser Arg Ser Phe Asp Asp Ser Tyr Met Asn Lys Val Leu

565 570 575

Val Phe Thr Lys Gln Asn Gln Glu Lys Leu Asn Gln Thr Pro Phe Glu

580 585 590

Ala Phe Gly Asn Asp Ser Ala Lys Trp Gln Lys Ile Glu Val Leu Ala

595 600 605

Lys Asn Leu Pro Thr Lys Lys Gln Lys Arg Ile Leu Asp Lys Asn Tyr

610 615 620

Lys Asp Lys Glu Gln Lys Asn Phe Lys Asp Arg Asn Leu Asn Asp Thr

625 630 635 640

Arg Tyr Ile Ala Arg Leu Val Leu Asn Tyr Thr Lys Asp Tyr Leu Asp

645 650 655

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

660 665 670

Lys Gly Ser Lys Val His Val Glu Ala Lys Ser Gly Met Leu Thr Ser

675 680 685

Ala Leu Arg His Thr Trp Gly Phe Ser Ala Lys Asp Arg Asn Asn His

690 695 700

Leu His His Ala Ile Asp Ala Val Ile Ile Ala Tyr Ala Asn Asn Ser
705 710 715 720
Ile Val Lys Ala Phe Ser Asp Phe Lys Lys Glu Gln Glu Ser Asn Ser

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

785 790 795 800
Ser Tyr Gly Gly Lys Glu Gly Val Leu Lys Ala Leu Glu Leu Gly Lys
805 810 815
Ile Arg Lys Val Asn Gly Lys Ile Val Lys Asn Gly Asp Met Phe Arg
820 825 830
Val Asp Ile Phe Lys His Lys Lys Thr Asn Lys Phe Tyr Ala Val Pro
835 840 845
Ile Tyr Thr Met Asp Phe Ala Leu Lys Val Leu Pro Asn Lys Ala Val

850 855 860
Ala Arg Ser Lys Lys Gly Glu Ile Lys Asp Trp Ile Leu Met Asp Glu
865 870 875 880
Asn Tyr Glu Phe Cys Phe Ser Leu Tyr Lys Asp Ser Leu Ile Leu Ile
885 890 895
Gln Thr Lys Asp Met Gln Glu Pro Glu Phe Val Tyr Tyr Asn Ala Phe
900 905 910
Thr Ser Ser Thr Val Ser Leu Ile Val Ser Lys His Asp Asn Lys Phe

915 920 925
Glu Thr Leu Ser Lys Asn Gln Lys Ile Leu Phe Lys Asn Ala Asn Glu
930 935 940
Lys Glu Val Ile Ala Lys Ser Ile Gly Ile Gln Asn Leu Lys Val Phe

945 950 955 960

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

965 970 975

Arg Gln Arg Glu Asp Phe Lys Lys

980

<210> 103

<211> 91

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 103

tataatctca taagaaattt aaaaagggac taaaataaag agtttgcggg actctgcggg 60

gttacaatcc cctaaaaccg cttttaaaat t 91

<210> 104

<211> 36

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 104

attttaccat aaagaaattt aaaaagggac taaaac 36

<210> 105

<211> 95

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<220><221> modified_base

<222> (1)..(20)

<223> a, c, u, g, unknown or other

<400> 105

nnnnnnnnnn nnnnnnnnnn guuuuagucc cgaaagggac uaaaauaaag aguuugcggg 60

acucugcggg guuacaaucc ccuaaaaccg cuuuu 95

<210> 106

<211> 1115

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 106

Met Ser Asp Leu Val Leu Gly Leu Asp Ile Gly Ile Gly Ser Val Gly
1 5 10 15

Val Gly Ile Leu Asn Lys Val Thr Gly Glu Ile Ile His Lys Asn Ser
20 25 30

Arg Ile Phe Pro Ala Ala Gln Ala Glu Asn Asn Leu Val Arg Arg Thr
35 40 45

Asn Arg Gln Gly Arg Arg Leu Ala Arg Arg Lys Lys His Arg Arg Val
50 55 60

Arg Leu Asn Arg Leu Phe Glu Glu Ser Gly Leu Ile Thr Asp Phe Thr
65 70 75 80

Lys Ile Ser Ile Asn Leu Asn Pro Tyr Gln Leu Arg Val Lys Gly Leu
85 90 95

Thr Asp Glu Leu Ser Asn Glu Glu Leu Phe Ile Ala Leu Lys Asn Met
100 105 110

Val Lys His Arg Gly Ile Ser Tyr Leu Asp Asp Ala Ser Asp Asp Gly
115 120 125

Asn Ser Ser Val Gly Asp Tyr Ala Gln Ile Val Lys Glu Asn Ser Lys
130 135 140

Gln Leu Glu Thr Lys Thr Pro Gly Gln Ile Gln Leu Glu Arg Tyr Gln
145 150 155 160

Thr Tyr Gly Gln Leu Arg Gly Asp Phe Thr Val Glu Lys Asp Gly Lys
165 170 175

Lys His Arg Leu Ile Asn Val Phe Pro Thr Ser Ala Tyr Arg Ser Glu
180 185 190
Ala Leu Arg Ile Leu Gln Thr Gln Gln Glu Phe Asn Pro Gln Ile Thr
195 200 205
Asp Glu Phe Ile Asn Arg Tyr Leu Glu Ile Leu Thr Gly Lys Arg Lys
210 215 220
Tyr Tyr His Gly Pro Gly Asn Glu Lys Ser Arg Thr Asp Tyr Gly Arg
225 230 235 240

Tyr Arg Thr Ser Gly Glu Thr Leu Asp Asn Ile Phe Gly Ile Leu Ile
245 250 255
Gly Lys Cys Thr Phe Tyr Pro Asp Glu Phe Arg Ala Ala Lys Ala Ser
260 265 270
Tyr Thr Ala Gln Glu Phe Asn Leu Leu Asn Asp Leu Asn Asn Leu Thr
275 280 285
Val Pro Thr Glu Thr Lys Lys Leu Ser Lys Glu Gln Lys Asn Gln Ile
290 295 300

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

Arg Glu Thr Leu Asp Lys Leu Ala Tyr Val Leu Thr Leu Asn Thr Glu
370 375 380
Arg Glu Gly Ile Gln Glu Ala Leu Glu His Glu Phe Ala Asp Gly Ser
385 390 395 400
Phe Ser Gln Lys Gln Val Asp Glu Leu Val Gln Phe Arg Lys Ala Asn

405 410 415

Ser Ser Ile Phe Gly Lys Gly Trp His Asn Phe Ser Val Lys Leu Met

420 425 430

Met Glu Leu Ile Pro Glu Leu Tyr Glu Thr Ser Glu Glu Gln Met Thr

435 440 445

Ile Leu Thr Arg Leu Gly Lys Gln Lys Thr Thr Ser Ser Ser Asn Lys

450 455 460

Thr Lys Tyr Ile Asp Glu Lys Leu Leu Thr Glu Glu Ile Tyr Asn Pro

465 470 475 480

Val Val Ala Lys Ser Val Arg Gln Ala Ile Lys Ile Val Asn Ala Ala

485 490 495

Ile Lys Glu Tyr Gly Asp Phe Asp Asn Ile Val Ile Glu Met Ala Arg

500 505 510

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

515 520 525

Lys Arg Ile Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys

530 535 540

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

545 550 555 560

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

565 570 575

Ile Asn Arg Leu Ser Asp Tyr Asp Val Asp His Ile Val Pro Gln Ser

580 585 590

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

595 600 605

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

610 615 620

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

625 630 635 640

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

645 650 655

Glu Leu Asp Lys Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr Arg
660 665 670
Gln Ile Thr Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr
675 680 685

Lys Tyr Asp Glu Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr
690 695 700
Leu Lys Ser Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr
705 710 715 720
Lys Val Arg Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu
725 730 735
Asn Ala Val Val Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu
740 745 750

Ser Glu Phe Val Tyr Gly Asp Tyr Lys Val Tyr Asp Val Arg Lys Met
755 760 765
Ile Ala Lys Ser Glu Gln Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe
770 775 780
Phe Tyr Ser Asn Ile Met Asn Phe Phe Lys Thr Glu Ile Thr Leu Ala
785 790 795 800
Asn Gly Glu Ile Arg Lys Arg Pro Leu Ile Glu Thr Asn Gly Glu Thr
805 810 815

Gly Glu Ile Val Trp Asp Lys Gly Arg Asp Phe Ala Thr Val Arg Lys
820 825 830
Val Leu Ser Met Pro Gln Val Asn Ile Val Lys Lys Thr Glu Val Gln
835 840 845
Thr Gly Gly Phe Ser Lys Glu Ser Ile Leu Pro Lys Arg Asn Ser Asp
850 855 860
Lys Leu Ile Ala Arg Lys Lys Asp Trp Asp Pro Lys Lys Tyr Gly Gly
865 870 875 880

Phe Asp Ser Pro Thr Val Ala Tyr Ser Val Leu Val Val Ala Lys Val
885 890 895
Glu Lys Gly Lys Ser Lys Lys Leu Lys Ser Val Lys Glu Leu Leu Gly

900 905 910

Ile Thr Ile Met Glu Arg Ser Ser Phe Glu Lys Asn Pro Ile Asp Phe

915 920 925

Leu Glu Ala Lys Gly Tyr Lys Glu Val Lys Lys Asp Leu Ile Ile Lys

930 935 940

Leu Pro Lys Tyr Ser Leu Phe Glu Leu Glu Asn Gly Arg Lys Arg Met

945 950 955 960

Leu Ala Ser Ala Gly Glu Leu Gln Lys Gly Asn Glu Leu Ala Leu Pro

965 970 975

Ser Lys Tyr Val Asn Phe Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu

980 985 990

Lys Gly Ser Pro Glu Asp Asn Glu Gln Lys Gln Leu Phe Val Glu Gln

995 1000 1005

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

1010 1015 1020

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

1025 1030 1035

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

1040 1045 1050

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

1055 1060 1065

Ala Ala Phe Lys Tyr Phe Asp Thr Thr Ile Asp Arg Lys Arg Tyr

1070 1075 1080

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

1085 1090 1095

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

1100 1105 1110

Gly Asp

1115

<210> 107

<211> 1374

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polypeptide"

<400> 107

Met Asp Lys Lys Tyr Ser Ile Gly Leu Asp Ile Gly Thr Asn Ser Val
1 5 10 15

Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe
20 25 30

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

Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu
50 55 60

Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys
65 70 75 80

Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser
85 90 95

Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys
100 105 110

His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr
115 120 125

His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp
130 135 140

Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His
145 150 155 160

Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro
165 170 175

Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr
180 185 190

Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala
195 200 205

Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn
210 215 220
Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn
225 230 235 240
Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe
245 250 255

Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp
260 265 270
Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp
275 280 285
Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp
290 295 300
Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser
305 310 315 320

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

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

Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp

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

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

Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly
            580                 585                 590
Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp
        595                 600                 605
Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr
    610                 615                 620
Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala
625                 630                 635                 640

His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr
                645                 650                 655
Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp
            660                 665                 670
Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe
        675                 680                 685
Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe
    690                 695                 700
```

Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu
705 710 715 720
His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly
725 730 735
Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly
740 745 750
Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Thr Asn
755 760 765

Glu Asp Asp Glu Lys Lys Ala Ile Gln Lys Ile Gln Lys Ala Asn Lys
770 775 780
Asp Glu Lys Asp Ala Ala Met Leu Lys Ala Ala Asn Gln Tyr Asn Gly
785 790 795 800
Lys Ala Glu Leu Pro His Ser Val Phe His Gly His Lys Gln Leu Ala
805 810 815
Thr Lys Ile Arg Leu Trp His Gln Gln Gly Glu Arg Cys Leu Tyr Thr
820 825 830

Gly Lys Thr Ile Ser Ile His Asp Leu Ile Asn Asn Ser Asn Gln Phe
835 840 845
Glu Val Asp His Ile Leu Pro Leu Ser Ile Thr Phe Asp Asp Ser Leu
850 855 860
Ala Asn Lys Val Leu Val Tyr Ala Thr Ala Asn Gln Glu Lys Gly Gln
865 870 875 880
Arg Thr Pro Tyr Gln Ala Leu Asp Ser Met Asp Asp Ala Trp Ser Phe
885 890 895

Arg Glu Leu Lys Ala Phe Val Arg Glu Ser Lys Thr Leu Ser Asn Lys
900 905 910
Lys Lys Glu Tyr Leu Leu Thr Glu Glu Asp Ile Ser Lys Phe Asp Val
915 920 925
Arg Lys Lys Phe Ile Glu Arg Asn Leu Val Asp Thr Arg Tyr Ala Ser
930 935 940
Arg Val Val Leu Asn Ala Leu Gln Glu His Phe Arg Ala His Lys Ile

945 950 955 960

Asp Thr Lys Val Ser Val Val Arg Gly Gln Phe Thr Ser Gln Leu Arg
965 970 975
Arg His Trp Gly Ile Glu Lys Thr Arg Asp Thr Tyr His His His Ala
980 985 990
Val Asp Ala Leu Ile Ile Ala Ala Ser Ser Gln Leu Asn Leu Trp Lys
995 1000 1005
Lys Gln Lys Asn Thr Leu Val Ser Tyr Ser Glu Asp Gln Leu Leu
1010 1015 1020

Asp Ile Glu Thr Gly Glu Leu Ile Ser Asp Asp Glu Tyr Lys Glu
1025 1030 1035
Ser Val Phe Lys Ala Pro Tyr Gln His Phe Val Asp Thr Leu Lys
1040 1045 1050
Ser Lys Glu Phe Glu Asp Ser Ile Leu Phe Ser Tyr Gln Val Asp
1055 1060 1065
Ser Lys Phe Asn Arg Lys Ile Ser Asp Ala Thr Ile Tyr Ala Thr
1070 1075 1080

Arg Gln Ala Lys Val Gly Lys Asp Lys Ala Asp Glu Thr Tyr Val
1085 1090 1095
Leu Gly Lys Ile Lys Asp Ile Tyr Thr Gln Asp Gly Tyr Asp Ala
1100 1105 1110
Phe Met Lys Ile Tyr Lys Lys Asp Lys Ser Lys Phe Leu Met Tyr
1115 1120 1125
Arg His Asp Pro Gln Thr Phe Glu Lys Val Ile Glu Pro Ile Leu
1130 1135 1140

Glu Asn Tyr Pro Asn Lys Gln Ile Asn Glu Lys Gly Lys Glu Val
1145 1150 1155
Pro Cys Asn Pro Phe Leu Lys Tyr Lys Glu Glu His Gly Tyr Ile
1160 1165 1170
Arg Lys Tyr Ser Lys Lys Gly Asn Gly Pro Glu Ile Lys Ser Leu
1175 1180 1185

Lys Tyr Tyr Asp Ser Lys Leu Gly Asn His Ile Asp Ile Thr Pro
1190 1195 1200

Lys Asp Ser Asn Asn Lys Val Val Leu Gln Ser Val Ser Pro Trp
1205 1210 1215
Arg Ala Asp Val Tyr Phe Asn Lys Thr Thr Gly Lys Tyr Glu Ile
1220 1225 1230
Leu Gly Leu Lys Tyr Ala Asp Leu Gln Phe Glu Lys Gly Thr Gly
1235 1240 1245
Thr Tyr Lys Ile Ser Gln Glu Lys Tyr Asn Asp Ile Lys Lys Lys
1250 1255 1260

Glu Gly Val Asp Ser Asp Ser Glu Phe Lys Phe Thr Leu Tyr Lys
1265 1270 1275
Asn Asp Leu Leu Leu Val Lys Asp Thr Glu Thr Lys Glu Gln Gln
1280 1285 1290
Leu Phe Arg Phe Leu Ser Arg Thr Met Pro Lys Gln Lys His Tyr
1295 1300 1305
Val Glu Leu Lys Pro Tyr Asp Lys Gln Lys Phe Glu Gly Gly Glu
1310 1315 1320

Ala Leu Ile Lys Val Leu Gly Asn Val Ala Asn Ser Gly Gln Cys
1325 1330 1335
Lys Lys Gly Leu Gly Lys Ser Asn Ile Ser Ile Tyr Lys Val Arg
1340 1345 1350
Thr Asp Val Leu Gly Asn Gln His Ile Ile Lys Asn Glu Gly Asp
1355 1360 1365
Lys Pro Lys Leu Asp Phe
1370
<210> 108
<211> 15
<212> PRT
<213> Artificial Sequence
<220
><221> source

<223> /note="Description of Artificial Sequence: Synthetic peptide"

<400> 108

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

1 5 10 15

<210> 109

<211> 15

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic peptide"

<400> 109

Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys

1 5 10 15

<210> 110

<211> 18

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic peptide"

<400> 110

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

1 5 10 15

Gly Ser

<210> 111

<211> 23

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 111

gccaaattgg acgaccctcg cgg 23

<210> 112

<211> 23

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 112

cgaggagacc cccgtttcgg tgg 23

<210> 113

<211> 23

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 113

cccgccgccg ccgtggctcg agg 23

<210> 114

<211> 23

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 114

tgagctctac gagatccaca agg 23

<210> 115

<211> 23

<212> DNA

<213> Artificial Sequence

<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"
<400> 115
ctcaaaattc ataccggttg tgg 23

<210> 116
<211> 23
<212> DNA
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"
<400> 116
cgttaaacaa caaccggact tgg 23
<210> 117
<211> 23
<212> DNA
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"
<400> 117
ttcaccccgc ggcgctgaat ggg 23

<210> 118
<211> 23
<212> DNA
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"
<400> 118
accactacca gtccgtccac agg 23
<210> 119

<211> 23

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 119

agcctttctg aacacatgca cgg 23

<210> 120

<211> 39

<212> DNA

<213> Homo sapiens

<400> 120

cctgccatca atgtggccat gcatgtgttc agaaaggct 39

<210> 121

<211> 39

<212> DNA

<213> Homo sapiens

<400> 121

cctgccatca atgtggccgt gcatgtgttc agaaaggct 39

<210> 122

<211> 23

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 122

cactgcttaa gcctcgctcg agg 23

<210> 123

<211> 23

<212> DNA

<213> Artificial Sequence

<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"
<400> 123
tcaccagcaa tattcgctcg agg 23
<210> 124
<211> 23
<212> DNA
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"
<400> 124
caccagcaat attccgctcg agg 23

<210> 125
<211> 23
<212> DNA
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"
<400> 125
tagcaacaga catacgctcg agg 23
<210> 126
<211> 23
<212> DNA
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"
<400> 126
gggcagtagt aatacgctcg agg 23

<210> 127

<211> 23

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 127

ccaattccca tacattattg tac 23

<210> 128

<211> 4677

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 128

tctttcttgc gctatgacac ttccagcaaa aggtagggcg ggctgcgaga cggcttcccg 60

gcgctgcatg caacaccgat gatgcttcga ccccccgaag ctccttcggg gctgcatggg 120

cgctccgatg ccgctccagg gcgagcgctg tttaaatagc caggcccccg attgcaaaga 180

cattatagcg agctaccaaa gccatattca aacacctaga tcactaccac ttctacacag 240

gccactcgag cttgtgatcg cactccgcta agggggcgcc tcttcctctt cgtttcagtc 300

acaacccgca aacatgtacc catacgatgt tccagattac gcttcgccga agaaaaagcg 360

caaggtcgaa gcgtccgaca agaagtacag catcggcctg gacatcggca ccaactctgt 420

gggctgggcc gtgatcaccg acgagtacaa ggtgcccagc aagaaattca aggtgctggg 480

caacaccgac cggcacagca tcaagaagaa cctgatcgga gccctgctgt tcgacagcgg 540

cgaaacagcc gaggccaccc ggctgaagag aaccgccaga agaagataca ccagacggaa 600

gaaccggatc tgctatctgc aagagatctt cagcaacgag atggccaagg tggacgacag 660

cttcttccac agactggaag agtccttcct ggtggaagag gataagaagc acgagcggca 720

ccccatcttc ggcaacatcg tggacgaggt ggcctaccac gagaagtacc ccaccatcta 780

ccacctgaga aagaaactgg tggacagcac cgacaaggcc gacctgcggc tgatctatct 840

ggccctggcc cacatgatca agttccgggg ccacttcctg atcgagggcg acctgaaccc 900

cgacaacagc gacgtggaca agctgttcat ccagctggtg cagacctaca accagctgtt 960

cgaggaaaac cccatcaacg ccagcggcgt ggacgccaag gccatcctgt ctgccagact 1020

gagcaagagc agacggctgg aaaatctgat cgcccagctg cccggcgaga agaagaatgg 1080

cctgttcggc aacctgattg ccctgagcct gggcctgacc cccaacttca agagcaactt 1140

cgacctggcc gaggatgcca aactgcagct gagcaaggac acctacgacg acgacctgga 1200

caacctgctg gcccagatcg gcgaccagta cgccgacctg tttctggccg ccaagaacct 1260

gtccgacgcc atcctgctga gcgacatcct gagagtgaac accgagatca ccaaggcccc 1320

cctgagcgcc tctatgatca agagatacga cgagcaccac caggacctga ccctgctgaa 1380

agctctcgtg cggcagcagc tgcctgagaa gtacaaagag attttcttcg accagagcaa 1440

gaacggctac gccggctaca ttgacggcgg agccagccag gaagagttct acaagttcat 1500

caagcccatc ctggaaaaga tggacggcac cgaggaactg ctcgtgaagc tgaacagaga 1560

ggacctgctg cggaagcagc ggaccttcga caacggcagc atcccccacc agatccacct 1620

gggagagctg cacgccattc tgcggcggca ggaagatttt tacccattcc tgaaggacaa 1680

ccgggaaaag atcgagaaga tcctgacctt ccgcatcccc tactacgtgg gccctctggc 1740

caggggaaac agcagattcg cctggatgac cagaaagagc gaggaaacca tcacccccctg 1800

gaacttcgag gaagtggtgg acaagggcgc ttccgcccag agcttcatcg agcggatgac 1860

caacttcgat aagaacctgc ccaacgagaa ggtgctgccc aagcacagcc tgctgtacga 1920

gtacttcacc gtgtataacg agctgaccaa agtgaaatac gtgaccgagg gaatgagaaa 1980

gcccgccttc ctgagcggcg agcagaaaaa ggccatcgtg gacctgctgt tcaagaccaa 2040

ccggaaagtg accgtgaagc agctgaaaga ggactacttc aagaaaatcg agtgcttcga 2100

ctccgtggaa atctccggcg tggaagatcg gttcaacgcc tccctgggca cataccacga 2160

tctgctgaaa attatcaagg acaaggactt cctggacaat gaggaaaacg aggacattct 2220

ggaagatatc gtgctgaccc tgacactgtt tgaggacaga gagatgatcg aggaacggct 2280

gaaaacctat gcccacctgt tcgacgacaa agtgatgaag cagctgaagc ggcggagata 2340

caccggctgg ggcaggctga gccggaagct gatcaacggc atccgggaca agcagtccgg 2400

caagacaatc ctggatttcc tgaagtccga cggcttcgcc aacagaaact tcatgcagct 2460

gatccacgac gacagcctga cctttaaaga ggacatccag aaagcccagg tgtccggcca 2520

gggcgatagc ctgcacgagc acattgccaa tctggccggc agccccgcca ttaagaaggg 2580

catcctgcag acagtgaagg tggtggacga gctcgtgaaa gtgatgggcc ggcacaagcc 2640

cgagaacatc gtgatcgaaa tggccagaga gaaccagacc acccagaagg gacagaagaa 2700

cagccgcgag agaatgaagc ggatcgaaga gggcatcaaa gagctgggca gccagatcct 2760

gaaagaacac cccgtggaaa acacccagct gcagaacgag aagctgtacc tgtactacct 2820

gcagaatggg cgggatatgt acgtggacca ggaactggac atcaaccggc tgtccgacta 2880

cgatgtggac catatcgtgc ctcagagctt tctgaaggac gactccatcg acaacaaggt 2940

gctgaccaga agcgacaaga accggggcaa gagcgacaac gtgccctccg aagaggtcgt 3000

gaagaagatg aagaactact ggcggcagct gctgaacgcc aagctgatta cccagagaaa 3060

gttcgacaat ctgaccaagg ccgagagagg cggcctgagc gaactggata aggccggctt 3120

catcaagaga cagctggtgg aaacccggca gatcacaaag cacgtggcac agatcctgga 3180

ctcccggatg aacactaagt acgacgagaa tgacaagctg atccgggaag tgaaagtgat 3240

caccctgaag tccaagctgg tgtccgattt ccggaaggat ttccagtttt acaaagtgcg 3300

cgagatcaac aactaccacc acgcccacga cgcctacctg aacgccgtcg tgggaaccgc 3360

cctgatcaaa aagtacccta agctggaaag cgagttcgtg tacggcgact acaaggtgta 3420

cgacgtgcgg aagatgatcg ccaagagcga gcaggaaatc ggcaaggcta ccgccaagta 3480

cttcttctac agcaacatca tgaacttttt caagaccgag attaccctgg ccaacggcga 3540

gatccggaag cggcctctga tcgagacaaa cggcgaaacc ggggagatcg tgtgggataa 3600

gggccgggat tttgccaccg tgcggaaagt gctgagcatg ccccaagtga atatcgtgaa 3660

aaagaccgag gtgcagacag gcggcttcag caaagagtct atcctgccca agaggaacag 3720

cgataagctg atcgccagaa agaaggactg ggaccctaag aagtacggcg gcttcgacag 3780

ccccaccgtg gcctattctg tgctggtggt ggccaaagtg gaaaagggca agtccaagaa 3840

actgaagagt gtgaaagagc tgctggggat caccatcatg gaaagaagca gcttcgagaa 3900

gaatcccatc gactttctgg aagccaaggg ctacaaagaa gtgaaaaagg acctgatcat 3960

caagctgcct aagtactccc tgttcgagct ggaaaacggc cggaagagaa tgctggcctc 4020

tgccggcgaa ctgcagaagg gaaacgaact ggccctgccc tccaaatatg tgaacttcct 4080

gtacctggcc agccactatg agaagctgaa gggctccccc gaggataatg agcagaaaca 4140

gctgtttgtg gaacagcaca agcactacct ggacgagatc atcgagcaga tcagcgagtt 4200

ctccaagaga gtgatcctgg ccgacgctaa tctggacaaa gtgctgtccg cctacaacaa 4260

gcaccgggat aagcccatca gagagcaggc cgagaatatc atccacctgt ttaccctgac 4320

caatctggga gcccctgccg ccttcaagta ctttgacacc accatcgacc ggaagaggta 4380

caccagcacc aaagaggtgc tggacgccac cctgatccac cagagcatca ccggcctgta 4440

cgagacacgg atcgacctgt ctcagctggg aggcgacagc cccaagaaga agagaaaggt 4500

ggaggccagc taaggatccg gcaagactgg ccccgcttgg caacgcaaca gtgagcccct 4560

ccctagtgtg tttggggatg tgactatgta ttcgtgtgtt ggccaacggg tcaacccgaa 4620

cagattgata cccgccttgg catttcctgt cagaatgtaa cgtcagttga tggtact 4677

<210> 129

<211> 3150

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 129

tctttcttgc gctatgacac ttccagcaaa aggtagggcg ggctgcgaga cggcttcccg 60

gcgctgcatg caacaccgat gatgcttcga ccccccgaag ctccttcggg gctgcatggg 120

cgctccgatg ccgctccagg gcgagcgctg tttaaatagc caggcccccg attgcaaaga 180

cattatagcg agctaccaaa gccatattca aacacctaga tcactaccac ttctacacag 240

gccactcgag cttgtgatcg cactccgcta agggggcgcc tcttcctctt cgtttcagtc 300

acaacccgca aacatgccta agaagaagag gaaggttaac acgattaaca tcgctaagaa 360

cgacttctct gacatcgaac tggctgctat cccgttcaac actctggctg accattacgg 420

tgagcgttta gctcgcgaac agttggccct tgagcatgag tcttacgaga tgggtgaagc 480

acgcttccgc aagatgtttg agcgtcaact taaagctggt gaggttgcgg ataacgctgc 540

cgccaagcct ctcatcacta ccctactccc taagatgatt gcacgcatca acgactggtt 600

tgaggaagtg aaagctaagc gcggcaagcg cccgacagcc ttccagttcc tgcaagaaat 660

caagccggaa gccgtagcgt acatcaccat taagaccact ctggcttgcc taaccagtgc 720

tgacaataca accgttcagg ctgtagcaag cgcaatcggt cgggccattg aggacgaggc 780

tcgcttcggt cgtatccgtg accttgaagc taagcacttc aagaaaaacg ttgaggaaca 840

actcaacaag cgcgtagggc acgtctacaa gaaagcattt atgcaagttg tcgaggctga 900

catgctctct aagggtctac tcggtggcga ggcgtggtct tcgtggcata aggaagactc 960

tattcatgta ggagtacgct gcatcgagat gctcattgag tcaaccggaa tggttagctt 1020

acaccgccaa aatgctggcg tagtaggtca agactctgag actatcgaac tcgcacctga 1080

atacgctgag gctatcgcaa cccgtgcagg tgcgctggct ggcatctctc cgatgttcca 1140

accttgcgta gttcctccta agccgtggac tggcattact ggtggtggct attgggctaa 1200

cggtcgtcgt cctctggcgc tggtgcgtac tcacagtaag aaagcactga tgcgctacga 1260

agacgtttac atgcctgagg tgtacaaagc gattaacatt gcgcaaaaca ccgcatggaa 1320

aatcaacaag aaagtcctag cggtcgccaa cgtaatcacc aagtggaagc attgtccggt 1380

cgaggacatc cctgcgattg agcgtgaaga actcccgatg aaaccggaag acatcgacat 1440

gaatcctgag gctctcaccg cgtggaaacg tgctgccgct gctgtgtacc gcaaggacaa 1500

ggctcgcaag tctcgccgta tcagccttga gttcatgctt gagcaagcca ataagtttgc 1560

taaccataag gccatctggt tcccttacaa catggactgg cgcggtcgtg tttacgctgt 1620

gtcaatgttc aacccgcaag gtaacgatat gaccaaagga ctgcttacgc tggcgaaagg 1680

taaaccaatc ggtaaggaag gttactactg gctgaaaatc cacggtgcaa actgtgcggg 1740

tgtcgacaag gttccgttcc ctgagcgcat caagttcatt gaggaaaacc acgagaacat 1800

catggcttgc gctaagtctc cactggagaa cacttggtgg gctgagcaag attctccgtt 1860

ctgcttcctt gcgttctgct ttgagtacgc tggggtacag caccacggcc tgagctataa 1920

ctgctccctt ccgctggcgt ttgacgggtc ttgctctggc atccagcact tctccgcgat 1980

gctccgagat gaggtaggtg gtcgcgcggt taacttgctt cctagtgaaa ccgttcagga 2040

catctacggg attgttgcta agaaagtcaa cgagattcta caagcagacg caatcaatgg 2100

gaccgataac gaagtagtta ccgtgaccga tgagaacact ggtgaaatct ctgagaaagt 2160

caagctgggc actaaggcac tggctggtca atggctggct tacggtgtta ctcgcagtgt 2220

gactaagcgt tcagtcatga cgctggctta cgggtccaaa gagttcggct tccgtcaaca 2280

agtgctggaa gataccattc agccagctat tgattccggc aagggtctga tgttcactca 2340

gccgaatcag gctgctggat acatggctaa gctgatttgg gaatctgtga gcgtgacggt 2400

ggtagctgcg gttgaagcaa tgaactggct taagtctgct gctaagctgc tggctgctga 2460

ggtcaaagat aagaagactg gagagattct tcgcaagcgt tgcgctgtgc attgggtaac 2520

tcctgatggt ttccctgtgt ggcaggaata caagaagcct attcagacgc gcttgaacct 2580

gatgttcctc ggtcagttcc gcttacagcc taccattaac accaacaaag atagcgagat 2640

tgatgcacac aaacaggagt ctggtatcgc tcctaacttt gtacacagcc aagacggtag 2700

ccaccttcgt aagactgtag tgtgggcaca cgagaagtac ggaatcgaat cttttgcact 2760

gattcacgac tccttcggta cgattccggc tgacgctgcg aacctgttca aagcagtgcg 2820

cgaaactatg gttgacacat atgagtcttg tgatgtactg gctgatttct acgaccagtt 2880

cgctgaccag ttgcacgagt ctcaattgga caaaatgcca gcacttccgg ctaaaggtaa 2940

cttgaacctc cgtgacatct tagagtcgga cttcgcgttc gcgtaaggat ccggcaagac 3000

tggccccgct tggcaacgca acagtgagcc cctccctagt gtgtttgggg atgtgactat 3060

gtattcgtgt gttggccaac gggtcaaccc gaacagattg atacccgcct tggcatttcc 3120

tgtcagaatg taacgtcagt tgatggtact 3150

<210> 130

<211> 125

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<220><221> modified_base

<222> (23)..(42)

<223> a, c, t, g, unknown or other

<400> 130

gaaattaata cgactcacta tannnnnnnn nnnnnnnnnn nngttttaga gctagaaata 60

gcaagttaaa ataaggctag tccgttatca acttgaaaaa gtggcaccga gtcggtgctt 120

ttttt 125

<210> 131

<211> 8452

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 131

tgcggtattt cacaccgcat caggtggcac ttttcgggga aatgtgcgcg gaaccctat 60

ttgtttattt ttctaaatac attcaaatat gtatccgctc atgagattat caaaaaggat 120

cttcacctag atccttttaa attaaaaatg aagttttaaa tcaatctaaa gtatatatga 180

gtaaacttgg tctgacagtt accaatgctt aatcagtgag gcacctatct cagcgatctg 240

tctatttcgt tcatccatag ttgcctgact ccccgtcgtg tagataacta cgatacggga 300

gggcttacca tctggcccca gtgctgcaat gataccgcga gacccacgct caccggctcc 360

agatttatca gcaataaacc agccagccgg aagggccgag cgcagaagtg gtcctgcaac 420

tttatccgcc tccatccagt ctattaattg ttgccgggaa gctagagtaa gtagttcgcc 480

agttaatagt ttgcgcaacg ttgttgccat tgctacaggc atcgtggtgt cacgctcgtc 540

gtttggtatg gcttcattca gctccggttc ccaacgatca aggcgagtta catgatcccc 600

catgttgtgc aaaaaagcgg ttagctcctt cggtcctccg atcgttgtca gaagtaagtt 660

ggccgcagtg ttatcactca tggttatggc agcactgcat aattctctta ctgtcatgcc 720

atccgtaaga tgcttttctg tgactggtga gtactcaacc aagtcattct gagaatagtg 780

tatgcggcga ccgagttgct cttgcccggc gtcaatacgg gataataccg cgccacatag 840

cagaacttta aaagtgctca tcattggaaa acgttcttcg gggcgaaaac tctcaaggat 900

cttaccgctg ttgagatcca gttcgatgta acccactcgt gcacccaact gatcttcagc 960

atcttttact ttcaccagcg tttctgggtg agcaaaaaca ggaaggcaaa atgccgcaaa 1020

aaagggaata agggcgacac ggaaatgttg aatactcata ctcttccttt ttcaatatta 1080

ttgaagcatt tatcagggtt attgtctcat gaccaaaatc ccttaacgtg agttttcgtt 1140

ccactgagcg tcagaccccg tagaaaagat caaaggatct tcttgagatc ctttttttct 1200

gcgcgtaatc tgctgcttgc aaacaaaaaa accaccgcta ccagcggtgg tttgtttgcc 1260

ggatcaagag ctaccaactc tttttccgaa ggtaactggc ttcagcagag cgcagatacc 1320

aaatactgtt cttctagtgt agccgtagtt aggccaccac ttcaagaact ctgtagcacc 1380

gcctacatac ctcgctctgc taatcctgtt accagtggct gttgccagtg gcgataagtc 1440

gtgtcttacc gggttggact caagacgata gttaccggat aaggcgcagc ggtcgggctg 1500

aacgggggggt tcgtgcacac agcccagctt ggagcgaacg acctacaccg aactgagata 1560

cctacagcgt gagctatgag aaagcgccac gcttcccgaa gggagaaagg cggacaggta 1620

tccggtaagc ggcagggtcg gaacaggaga gcgcacgagg gagcttccag ggggaaacgc 1680

ctggtatctt tatagtcctg tcgggtttcg ccacctctga cttgagcgtc gatttttgtg 1740

atgctcgtca ggggggcgga gcctatggaa aaacgccagc aacgcggcct ttttacggtt 1800

cctggccttt tgctggcctt ttgctcacat gttctttcct gcgttatccc ctgattctgt 1860

ggataaccgt attaccgcct ttgagtgagc tgataccgct cgccgcagcc gaacgaccga 1920

gcgcagcgag tcagtgagcg aggaagcggt cgctgaggct tgacatgatt ggtgcgtatg 1980

tttgtatgaa gctacaggac tgatttggcg ggctatgagg gcgggggaag ctctggaagg 2040

gccgcgatgg ggcgcgcggc gtccagaagg cgccatacgg cccgctggcg gcacccatcc 2100

ggtataaaag cccgcgaccc cgaacggtga cctccacttt cagcgacaaa cgagcactta 2160

tacatacgcg actattctgc cgctatacat aaccactcag ctagcttaag atcccatcaa 2220

gcttgcatgc cgggcgcgcc agaaggagcg cagccaaacc aggatgatgt ttgatggggt 2280

atttgagcac ttgcaaccct tatccggaag ccccctggcc cacaaaggct aggcgccaat 2340

gcaagcagtt cgcatgcagc ccctggagcg gtgccctcct gataaaccgg ccaggggccc 2400

tatgttcttt actttttac aagagaagtc actcaacatc ttaaaatggc caggtgagtc 2460

gacgagcaag cccggcggat caggcagcgt gcttgcagat ttgacttgca acgcccgcat 2520

tgtgtcgacg aaggcttttg gctcctctgt cgctgtctca agcagcatct aaccctgcgt 2580

cgccgtttcc atttgcagga gattcgaggt accatgtacc catacgatgt tccagattac 2640

gcttcgccga agaaaaagcg caaggtcgaa gcgtccgaca agaagtacag catcggcctg 2700

gacatcggca ccaactctgt gggctgggcc gtgatcaccg acgagtacaa ggtgcccagc 2760

aagaaattca aggtgctggg caacaccgac cggcacagca tcaagaagaa cctgatcgga 2820

gccctgctgt tcgacagcgg cgaaacagcc gaggccaccc ggctgaagag aaccgccaga 2880

agaagataca ccagacggaa gaaccggatc tgctatctgc aagagatctt cagcaacgag 2940

atggccaagg tggacgacag cttcttccac agactggaag agtccttcct ggtggaagag 3000

gataagaagc acgagcggca ccccatcttc ggcaacatcg tggacgaggt ggcctaccac 3060

gagaagtacc ccaccatcta ccacctgaga aagaaactgg tggacagcac cgacaaggcc 3120

gacctgcggc tgatctatct ggccctggcc cacatgatca agttccgggg ccacttcctg 3180

atcgagggcg acctgaaccc cgacaacagc gacgtggaca agctgttcat ccagctggtg 3240

cagacctaca accagctgtt cgaggaaaac cccatcaacg ccagcggcgt ggacgccaag 3300

gccatcctgt ctgccagact gagcaagagc agacggctgg aaaatctgat cgcccagctg 3360

cccggcgaga agaagaatgg cctgttcggc aacctgattg ccctgagcct gggcctgacc 3420

cccaacttca agagcaactt cgacctggcc gaggatgcca aactgcagct gagcaaggac 3480

acctacgacg acgacctgga caacctgctg gcccagatcg gcgaccagta cgccgacctg 3540

tttctggccg ccaagaacct gtccgacgcc atcctgctga gcgacatcct gagagtgaac 3600

accgagatca ccaaggcccc cctgagcgcc tctatgatca agagatacga cgagcaccac 3660

caggacctga ccctgctgaa agctctcgtg cggcagcagc tgcctgagaa gtacaaagag 3720

attttcttcg accagagcaa gaacggctac gccggctaca ttgacggcgg agccagccag 3780

gaagagttct acaagttcat caagcccatc ctggaaaaga tggacggcac cgaggaactg 3840

ctcgtgaagc tgaacagaga ggacctgctg cggaagcagc ggaccttcga caacggcagc 3900

atccccacc agatccacct gggagagctg cacgccattc tgcggcggca ggaagatttt 3960

tacccattcc tgaaggacaa ccgggaaaag atcgagaaga tcctgacctt ccgcatcccc 4020

tactacgtgg gccctctggc caggggaaac agcagattcg cctggatgac cagaaagagc 4080

gaggaaacca tcacccccctg gaacttcgag gaagtggtgg acaagggcgc ttccgcccag 4140

agcttcatcg agcggatgac caacttcgat aagaacctgc ccaacgagaa ggtgctgccc 4200

aagcacagcc tgctgtacga gtacttcacc gtgtataacg agctgaccaa agtgaaatac 4260

gtgaccgagg gaatgagaaa gcccgccttc ctgagcggcg agcagaaaaa ggccatcgtg 4320

gacctgctgt tcaagaccaa ccggaaagtg accgtgaagc agctgaaaga ggactacttc 4380

aagaaaatcg agtgcttcga ctccgtggaa atctccggcg tggaagatcg gttcaacgcc 4440

tccctgggca cataccacga tctgctgaaa attatcaagg acaaggactt cctggacaat 4500

gaggaaaacg aggacattct ggaagatatc gtgctgaccc tgacactgtt tgaggacaga 4560

gagatgatcg aggaacggct gaaaacctat gcccacctgt tcgacgacaa agtgatgaag 4620

cagctgaagc ggcggagata caccggctgg ggcaggctga gccggaagct gatcaacggc 4680

atccgggaca agcagtccgg caagacaatc ctggatttcc tgaagtccga cggcttcgcc 4740

aacagaaact tcatgcagct gatccacgac gacagcctga cctttaaaga ggacatccag 4800

aaagcccagg tgtccggcca gggcgatagc ctgcacgagc acattgccaa tctggccggc 4860

agccccgcca ttaagaaggg catcctgcag acagtgaagg tggtggacga gctcgtgaaa 4920

gtgatgggcc ggcacaagcc cgagaacatc gtgatcgaaa tggccagaga gaaccagacc 4980

acccagaagg gacagaagaa cagccgcgag agaatgaagc ggatcgaaga gggcatcaaa 5040

gagctgggca gccagatcct gaaagaacac cccgtggaaa acacccagct gcagaacgag 5100

aagctgtacc tgtactacct gcagaatggg cgggatatgt acgtggacca ggaactggac 5160

atcaaccggc tgtccgacta cgatgtggac catatcgtgc ctcagagctt tctgaaggac 5220

gactccatcg acaacaaggt gctgaccaga agcgacaaga accggggcaa gagcgacaac 5280

gtgccctccg aagaggtcgt gaagaagatg aagaactact ggcggcagct gctgaacgcc 5340

aagctgatta cccagagaaa gttcgacaat ctgaccaagg ccgagagagg cggcctgagc 5400

gaactggata aggccggctt catcaagaga cagctggtgg aaacccggca gatcacaaag 5460

cacgtggcac agatcctgga ctcccggatg aacactaagt acgacgagaa tgacaagctg 5520

atccgggaag tgaaagtgat caccctgaag tccaagctgg tgtccgattt ccggaaggat 5580

ttccagtttt acaaagtgcg cgagatcaac aactaccacc acgcccacga cgcctacctg 5640

aacgccgtcg tgggaaccgc cctgatcaaa aagtacccta agctggaaag cgagttcgtg 5700

tacggcgact acaaggtgta cgacgtgcgg aagatgatcg ccaagagcga gcaggaaatc 5760

ggcaaggcta ccgccaagta cttcttctac agcaacatca tgaacttttt caagaccgag 5820

attaccctgg ccaacggcga gatccggaag cggcctctga tcgagacaaa cggcgaaacc 5880

ggggagatcg tgtgggataa gggccgggat tttgccaccg tgcggaaagt gctgagcatg 5940

ccccaagtga atatcgtgaa aaagaccgag gtgcagacag gcggcttcag caaagagtct 6000

atcctgccca agaggaacag cgataagctg atcgccagaa agaaggactg ggaccctaag 6060

aagtacggcg gcttcgacag ccccaccgtg gcctattctg tgctggtggt ggccaaagtg 6120

gaaaagggca agtccaagaa actgaagagt gtgaaagagc tgctggggat caccatcatg 6180

gaaagaagca gcttcgagaa gaatcccatc gactttctgg aagccaaggg ctacaaagaa 6240

gtgaaaaagg acctgatcat caagctgcct aagtactccc tgttcgagct ggaaaacggc 6300

cggaagagaa tgctggcctc tgccggcgaa ctgcagaagg gaaacgaact ggccctgccc 6360

tccaaatatg tgaacttcct gtacctggcc agccactatg agaagctgaa gggctccccc 6420

gaggataatg agcagaaaca gctgtttgtg gaacagcaca agcactacct ggacgagatc 6480

atcgagcaga tcagcgagtt ctccaagaga gtgatcctgg ccgacgctaa tctggacaaa 6540

gtgctgtccg cctacaacaa gcaccgggat aagcccatca gagagcaggc cgagaatatc 6600

atccacctgt ttaccctgac caatctggga gcccctgccg ccttcaagta ctttgacacc 6660

accatcgacc ggaagaggta caccagcacc aaagaggtgc tggacgccac cctgatccac 6720

cagagcatca ccggcctgta cgagacacgg atcgacctgt ctcagctggg aggcgacagc 6780

cccaagaaga agagaaaggt ggaggccagc taacatatga ttcgaatgtc tttcttgcgc 6840

tatgacactt ccagcaaaag gtagggcggg ctgcgagacg gcttcccggc gctgcatgca 6900

acaccgatga tgcttcgacc ccccgaagct ccttcggggc tgcatgggcg ctccgatgcc 6960

gctccagggc gagcgctgtt taaatagcca ggcccccgat tgcaaagaca ttatagcgag 7020

ctaccaaagc catattcaaa cacctagatc actaccactt ctacacaggc cactcgagct 7080

tgtgatcgca ctccgctaag ggggcgcctc ttcctcttcg tttcagtcac aacccgcaaa 7140

catgacacaa gaatccctgt tacttctcga ccgtattgat tcggatgatt cctacgcgag 7200

cctgcggaac gaccaggaat tctgggaggt gagtcgacga gcaagcccgg cggatcaggc 7260

agcgtgcttg cagatttgac ttgcaacgcc cgcattgtgt cgacgaaggc ttttggctcc 7320

tctgtcgctg tctcaagcag catctaaccc tgcgtcgccg tttccatttg cagccgctgg 7380

cccgccgagc cctggaggag ctcgggctgc cggtgccgcc ggtgctgcgg gtgcccggcg 7440

agagcaccaa ccccgtactg gtcggcgagc ccggcccggt gatcaagctg ttcggcgagc 7500

actggtgcgg tccggagagc ctcgcgtcgg agtcggaggc gtacgcggtc ctggcggacg 7560

ccccggtgcc ggtgccccgc ctcctcggcc gcggcgagct gcggcccggc accggagcct 7620

ggccgtggcc ctacctggtg atgagccgga tgaccggcac cacctggcgg tccgcgatgg 7680

acggcacgac cgaccggaac gcgctgctcg ccctggcccg cgaactcggc cgggtgctcg 7740

gccggctgca cagggtgccg ctgaccggga acaccgtgct caccccccat tccgaggtct 7800

tcccggaact gctgcgggaa cgccgcgcgg cgaccgtcga ggaccaccgc gggtggggct 7860

acctctcgcc ccggctgctg gaccgcctgg aggactggct gccggacgtg gacacgctgc 7920

tggccggccg cgaaccccgg ttcgtccacg gcgacctgca cgggaccaac atcttcgtgg 7980
acctggccgc gaccgaggtc accgggatcg tcgacttcac cgacgtctat gcgggagact 8040
cccgctacag cctggtgcaa ctgcatctca acgccttccg gggcgaccgc gagatcctgg 8100
ccgcgctgct cgacggggcg cagtggaagc ggaccgagga cttcgcccgc gaactgctcg 8160
ccttcacctt cctgcacgac ttcgaggtgt tcgaggagac cccgctggat ctctccggct 8220
tcaccgatcc ggaggaactg gcgcagttcc tctgggggcc gccggacacc gccccggcg 8280
cctgataagg atccggcaag actggccccg cttggcaacg caacagtgag cccctcccta 8340

gtgtgtttgg ggatgtgact atgtattcgt gtgttggcca acgggtcaac ccgaacagat 8400
tgatacccgc cttggcattt cctgtcagaa tgtaacgtca gttgatggta ct 8452
<210> 132
<211> 27
<212> DNA
<213> Homo sapiens
<400> 132
ccgtgccggg cggggagacc gccatgg 27
<210> 133
<211> 22
<212> DNA
<213> Homo sapiens
<400> 133
ggcccggctg tggctgagga gc 22
<210> 134
<211> 20
<212> DNA
<213> Homo sapiens
<400> 134
cggtctcccg cccggcacgg 20

<210> 135
<211> 26
<212> DNA
<213> Homo sapiens

<400> 135

gctcctcagc cacagccggg ccgggt 26

<210> 136

<211> 12

<212> DNA

<213> Homo sapiens

<400> 136

cgaccctgga aa 12

<210> 137

<211> 14

<212> DNA

<213> Homo sapiens

<400> 137

ccccgccgcc accc 14

<210> 138

<211> 18

<212> DNA

<213> Homo sapiens

<400> 138

tttccagggt cgccatgg 18

<210> 139

<211> 10

<212> DNA

<213> Homo sapiens

<400> 139

ggcggcgggg 10

<210> 140

<211> 20

<212> DNA

<213> Homo sapiens

<400> 140

acccttgtta gccacctccc 20

<210> 141

<211> 20

<212> DNA

<213> Homo sapiens

<400> 141

gaacgcagtg ctcttcgaag 20

<210> 142

<211> 20

<212> DNA

<213> Homo sapiens

<400> 142

ctcacgccct gctccgtgta 20

<210> 143

<211> 20

<212> DNA

<213> Homo sapiens

<400> 143

ggcgacaact acttcctggt 20

<210> 144

<211> 20

<212> DNA

<213> Homo sapiens

<400> 144

ctcacgccct gctccgtgta 20

<210> 145

<211> 20

<212> DNA

<213> Homo sapiens

<400> 145

gggcgacaac tacttcctgg 20

<210> 146

<211> 20

<212> DNA

<213> Homo sapiens

<400> 146

cctcttcagg gccggggtgg 20

<210> 147

<211> 20

<212> DNA

<213> Homo sapiens

<400> 147

gaggacccag gtggaactgc 20

<210> 148

<211> 20

<212> DNA

<213> Homo sapiens

<400> 148

tcagctccag gcggtcctgg 20

<210> 149

<211> 20

<212> DNA

<213> Homo sapiens

<400> 149

agcagcagca gcagtggcag 20

<210> 150

<211> 20

<212> DNA

<213> Homo sapiens

<400> 150

tgggcaccgt cagctccagg 20

<210> 151

<211> 20

<212> DNA

<213> Homo sapiens

<400> 151

cagcagtggc agcggccacc 20

<210> 152

<211> 20

<212> DNA

<213> Homo sapiens

<400> 152

acctctcccc tggccctcat 20

<210> 153

<211> 20

<212> DNA

<213> Homo sapiens

<400> 153

ccaggaccgc ctggagctga 20

<210> 154

<211> 20

<212> DNA

<213> Homo sapiens

<400> 154

ccgtcagctc caggcggtcc 20

<210> 155

<211> 20

<212> DNA

<213> Homo sapiens

<400> 155

agcagcagca gcagtggcag 20

<210> 156

<211> 20

<212> DNA

<213> Homo sapiens

<400> 156

atgtgccaag caaagcctca 20

<210> 157

<211> 20

<212> DNA

<213> Homo sapiens

<400> 157

ttcggtcatg cccgtggatg 20

<210> 158

<211> 20

<212> DNA

<213> Homo sapiens

<400> 158

gtcgttgaaa ttcatcgtac 20

<210> 159

<211> 20

<212> DNA

<213> Homo sapiens

<400> 159

accacctgtg aagagtttcc 20

<210> 160

<211> 20

<212> DNA

<213> Homo sapiens

<400> 160

cgtcgttgaa attcatcgta 20

<210> 161

<211> 20

<212> DNA

<213> Homo sapiens

<400> 161

accacctgtg aagagtttcc 20

<210> 162

<211> 20

<212> DNA

<213> Mus musculus

<400> 162

gaacgcagtg cttttcgagg 20

<210> 163

<211> 20

<212> DNA

<213> Mus musculus

<400> 163

acccttgttg gccacctccc 20

<210> 164

<211> 20

<212> DNA

<213> Mus musculus

<400> 164

ggtgacaact actatctggt 20

<210> 165

<211> 20

<212> DNA

<213> Mus musculus

<400> 165

ctcacaccct gctccgtgta 20

<210> 166

<211> 20

<212> DNA

<213> Mus musculus

<400> 166

gggtgacaac tactatctgg 20

<210> 167

<211> 20

<212> DNA

<213> Mus musculus

<400> 167

ctcacaccct gctccgtgta 20

<210> 168

<211> 20

<212> DNA

<213> Mus musculus

<400> 168

cgagaacgca gtgcttttcg 20

<210> 169

<211> 20

<212> DNA

<213> Mus musculus

<400> 169

accettgttg gccacctccc 20

<210> 170

<211> 20

<212> DNA

<213> Mus musculus

<400> 170

atgagccaag caaatcctca 20

<210> 171

<211> 20

<212> DNA

<213> Mus musculus

<400> 171

ttccgtcatg cccgtggaca 20

<210> 172

<211> 20

<212> DNA

<213> Mus musculus

<400> 172

cttcgttgaa aaccattgta 20

<210> 173

<211> 20

<212> DNA

<213> Mus musculus

<400> 173

ccacctctga agagtttcct 20

<210> 174

<211> 20

<212> DNA

<213> Mus musculus

<400> 174

cttcgttgaa aaccattgta 20

<210> 175

<211> 20

<212> DNA

<213> Mus musculus

<400> 175

accacctctg aagagtttcc 20

<210> 176

<211> 20

<212> DNA

<213> Mus musculus

<400> 176

cttccactca ctctgcgatt 20

<210> 177

<211> 20

<212> DNA

<213> Mus musculus

<400> 177

accatgtctc agtgtcaagc 20

<210> 178

<211> 20

<212> DNA

<213> Mus musculus

<400> 178

ggcggcaaca gcggcaacag 20

<210> 179

<211> 20

<212> DNA

<213> Mus musculus

<400> 179

actgctctgc gtggctgcgg 20

<210> 180

<211> 20

<212> DNA

<213> Mus musculus

<400> 180

ccgcagccac gcagagcagt 20

<210> 181

<211> 20

<212> DNA

<213> Mus musculus

<400> 181

gcacctctcc tcgccccgat 20

<210> 182

<211> 23

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 182

gagggcctat ttcccatgat tcc 23

<210> 183

<211> 126

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<220><221> modified_base

<222> (84)..(102)

<223> a, c, t, g, unknown or other

<400> 183

aaaaaaagca ccgactcggt gccacttttt caagttgata acggactagc cttattttaa 60

cttgctattt ctagctctaa aacnnnnnnn nnnnnnnnnn nnccggtgtt tcgtcctttc 120

cacaag 126

<210> 184

<211> 24

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (6)..(24)

<223> a, c, t, g, unknown or other

<400> 184

caccgnnnnn nnnnnnnnnn nnnn 24

<210> 185

<211> 24

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (5)..(23)

<223> a, c, t, g, unknown or other

<400> 185

aaacnnnnnn nnnnnnnnnn nnnc 24

<210> 186

<211> 126

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 186

aaaaaaagca ccgactcggt gccacttttt caagttgata acggactagc cttattttaa 60

cttgctattt ctagctctaa aacccctagt cattggaggt gaccggtgtt tcgtcctttc 120

cacaag 126

<210> 187

<211> 24

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 187

caccgtcacc tccaatgact aggg 24

<210> 188

<211> 24

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 188

aaacccctag tcattggagg tgac 24

<210> 189

<211> 192

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 189

cagaagaaga agggctccca tcacatcaac cggtggcgca ttgccacgaa gcaggccaat 60

ggggaggaca tcgatgtcac ctccaatgac aagcttgcta gcggtgggca accacaaacc 120

cacgagggca gagtgctgct tgctgctggc caggcccctg cgtgggccca agctggactc 180

tggccactcc ct 192

<210> 190

<211> 192

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 190

agggagtggc cagagtccag cttgggccca cgcaggggcc tggccagcag caagcagcac 60

tctgccctcg tgggtttgtg gttgcccacc gctagcaagc ttgtcattgg aggtgacatc 120

gatgtcctcc ccattggcct gcttcgtggc aatgcgccac cggttgatgt gatgggagcc 180

cttcttcttc tg 192

<210> 191

<211> 20

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 191

ccatcccctt ctgtgaatgt 20

<210> 192

<211> 20

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 192

ggagattgga gacacggaga 20

<210> 193

<211> 20

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 193

ggctccctgg gttcaaagta 20

<210> 194

<211> 21

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 194

agaggggtct ggatgtcgta a 21

<210> 195

<211> 24

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 195

cgccagggtt ttcccagtca cgac 24

<210> 196

<211> 51

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 196

gagggtctcg tccttgcggc cgcgctagcg agggcctatt tcccatgatt c 51

<210> 197

<211> 133

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<220><221> modified_base

<222> (95)..(114)

<223> a, c, t, g, unknown or other

<400> 197

ctcggtctcg gtaaaaaagc accgactcgg tgccactttt tcaagttgat aacggactag 60

ccttatttta acttgctatt tctagctcta aaacnnnnnn nnnnnnnnnn nnnnggtgtt 120

tcgtcctttc cac 133

<210> 198

<211>

41

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 198

gagggtctct ttaccggtga gggcctattt cccatgattc c 41

<210> 199

<211> 133

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<220><221> modified_base

<222> (95)..(114)

<223> a, c, t, g, unknown or other

<400> 199

ctcggtctcc tcaaaaaagc accgactcgg tgccactttt tcaagttgat aacggactag 60

ccttatttta acttgctatt tctagctcta aaacnnnnnn nnnnnnnnnn nnnnggtgtt 120

tcgtcctttc cac 133

<210> 200

<211> 40

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 200

gagggtctct ttgagctcga gggcctattt cccatgattc 40

<210> 201

<211> 133

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<220><221> modified_base

<222> (96)..(115)

<223> a, c, t, g, unknown or other

<400> 201

ctcggtctcg cgtaaaaaag caccgactcg gtgccacttt ttcaagttga taacggacta 60

gccttatttt aacttgctat ttctagctct aaaacnnnnn nnnnnnnnnn nnnnnggtgt 120

ttcgtccttt cca 133

<210> 202

<211> 27

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 202

gagggtctct tacgcgtgtg tctagac 27

<210> 203

<211> 98

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 203

ctcggtctca aggacaggga agggagcagt ggttcacgcc tgtaatccca gcaatttggg 60

aggccaaggt gggtagatca cctgagatta ggagttgc 98

<210> 204

<211> 30

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 204

cctgtccttg cggccgcgct agcgagggcc 30

<210> 205

<211> 31

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 205

cacgcggccg caaggacagg gaagggagca g 31

<210> 206

<211> 327

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polypeptide"

<400> 206

Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu
1 5 10 15

Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly
20 25 30

Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile
35 40 45

Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr
50 55 60

Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys
65 70 75 80

Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu
85 90 95

Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu
100 105 110

Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly
115 120 125

Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr
130 135 140

Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn
145 150 155 160

Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser
165 170 175

Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly
180 185 190

Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu

195 200 205

Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe

210 215 220

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

225 230 235 240

Gly Leu Arg Ser Arg Glu Glu Glu Glu Glu Thr Asp Ser Arg Met Pro

245 250 255

His Leu Asp Ser Pro Gly Ser Ser Gln Pro Arg Arg Ser Phe Leu Ser

260 265 270

Arg Val Ile Arg Ala Ala Leu Pro Leu Gln Leu Leu Leu Leu Leu Leu

275 280 285

Leu Leu Leu Ala Cys Leu Leu Pro Ala Ser Glu Asp Asp Tyr Ser Cys

290 295 300

Thr Gln Ala Asn Asn Phe Ala Arg Ser Phe Tyr Pro Met Leu Arg Tyr

305 310 315 320

Thr Asn Gly Pro Pro Pro Thr

325

<210> 207

<211> 3243

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 207

accggtgcca ccatgtaccc atacgatgtt ccagattacg cttcgccgaa gaaaaagcgc 60

aaggtcgaag cgtccatgaa aaggaactac attctggggc tggacatcgg gattacaagc 120

gtggggtatg ggattattga ctatgaaaca agggacgtga tcgacgcagg cgtcagactg 180

ttcaaggagg ccaacgtgga aaacaatgag ggacggagaa gcaagagggg agccaggcgc 240

ctgaaacgac ggagaaggca cagaatccag agggtgaaga aactgctgtt cgattacaac 300

ctgctgaccg accattctga gctgagtgga attaatcctt atgaagccag ggtgaaaggc 360

ctgagtcaga agctgtcaga ggaagagttt tccgcagctc tgctgcacct ggctaagcgc 420

cgaggagtgc ataacgtcaa tgaggtggaa gaggacaccg gcaacgagct gtctacaaag 480

gaacagatct cacgcaatag caaagctctg gaagagaagt atgtcgcaga gctgcagctg 540

gaacggctga agaaagatgg cgaggtgaga gggtcaatta ataggttcaa gacaagcgac 600

tacgtcaaag aagccaagca gctgctgaaa gtgcagaagg cttaccacca gctggatcag 660

agcttcatcg atacttatat cgacctgctg gagactcgga gaacctacta tgagggacca 720

ggagaaggga gccccttcgg atggaaagac atcaaggaat ggtacgagat gctgatggga 780

cattgcacct attttccaga agagctgaga agcgtcaagt acgcttataa cgcagatctg 840

tacaacgccc tgaatgacct gaacaacctg gtcatcacca gggatgaaaa cgagaaactg 900

gaatactatg agaagttcca gatcatcgaa aacgtgttta agcagaagaa aaagcctaca 960

ctgaaacaga ttgctaagga gatcctggtc aacgaagagg acatcaaggg ctaccgggtg 1020

acaagcactg gaaaaccaga gttcaccaat ctgaaagtgt atcacgatat taaggacatc 1080

acagcacgga aagaaatcat tgagaacgcc gaactgctgg atcagattgc taagatcctg 1140

actatctacc agagctccga ggacatccag gaagagctga ctaacctgaa cagcgagctg 1200

acccaggaag agatcgaaca gattagtaat ctgaaggggt acaccggaac acacaacctg 1260

tccctgaaag ctatcaatct gattctggat gagctgtggc atacaaacga caatcagatt 1320

gcaatcttta accggctgaa gctggtccca aaaaaggtgg acctgagtca gcagaaagag 1380

atcccaacca cactggtgga cgatttcatt ctgtcacccg tggtcaagcg gagcttcatc 1440

cagagcatca aagtgatcaa cgccatcatc aagaagtacg gcctgcccaa tgatatcatt 1500

atcgagctgg ctagggagaa gaacagcaag gacgcacaga agatgatcaa tgagatgcag 1560

aaacgaaacc ggcagaccaa tgaacgcatt gaagagatta tccgaactac cgggaaagag 1620

aacgcaaagt acctgattga aaaaatcaag ctgcacgata tgcaggaggg aaagtgtctg 1680

tattctctgg aggccatccc cctggaggac ctgctgaaca atccattcaa ctacgaggtc 1740

gatcatatta tccccagaag cgtgtccttc gacaattcct ttaacaacaa ggtgctggtc 1800

aagcaggaag agaactctaa aaagggcaat aggactcctt tccagtacct gtctagttca 1860

gattccaaga tctcttacga aacctttaaa aagcacattc tgaatctggc caaaggaaag 1920

ggccgcatca gcaagaccaa aaaggagtac ctgctggaag agcgggacat caacagattc 1980

tccgtccaga aggattttat taaccggaat ctggtggaca caagatacgc tactcgcggc 2040

ctgatgaatc tgctgcgatc ctatttccgg gtgaacaatc tggatgtgaa agtcaagtcc 2100

atcaacggcg ggttcacatc ttttctgagg cgcaaatgga agtttaaaaa ggagcgcaac 2160

aaagggtaca agcaccatgc cgaagatgct ctgattatcg caaatgccga cttcatcttt 2220

aaggagtgga aaaagctgga caaagccaag aaagtgatgg agaaccagat gttcgaagag 2280

aagcaggccg aatctatgcc cgaaatcgag acagaacagg agtacaagga gattttcatc 2340

actcctcacc agatcaagca tatcaaggat ttcaaggact acaagtactc tcaccgggtg 2400
gataaaaagc ccaacagaga gctgatcaat gacaccctgt atagtacaag aaaagacgat 2460
aaggggaata ccctgattgt gaacaatctg aacggactgt acgacaaaga taatgacaag 2520
ctgaaaaagc tgatcaacaa aagtcccgag aagctgctga tgtaccacca tgatcctcag 2580
acatatcaga aactgaagct gattatggag cagtacggcg acgagaagaa cccactgtat 2640
aagtactatg aagagactgg gaactacctg accaagtata gcaaaaagga taatggcccc 2700
gtgatcaaga agatcaagta ctatgggaac aagctgaatg cccatctgga catcacagac 2760

gattacccta acagtcgcaa caaggtggtc aagctgtcac tgaagccata cagattcgat 2820
gtctatctgg acaacggcgt gtataaattt gtgactgtca agaatctgga tgtcatcaaa 2880
aaggagaact actatgaagt gaatagcaag tgctacgaag aggctaaaaa gctgaaaaag 2940
attagcaacc aggcagagtt catcgcctcc ttttacaaca acgacctgat taagatcaat 3000
ggcgaactgt atagggtcat cggggtgaac aatgatctgc tgaaccgcat tgaagtgaat 3060
atgattgaca tcacttaccg agagtatctg gaaaacatga atgataagcg ccccccctcga 3120
attatcaaaa caattgcctc taagactcag agtatcaaaa agtactcaac cgacattctg 3180

ggaaacctgt atgaggtgaa gagcaaaaag caccctcaga ttatcaaaaa gggctaagaa 3240
ttc 3243

<210> 208
<211> 23
<212> DNA
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"
<400> 208
ggagctggag ctgttcacgt tgg 23

<210> 209
<211> 23
<212> DNA
<213> Artificial Sequence
<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 209

cgggcagcag atgttcgcgt agg 23

<210> 210

<211> 23

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 210

agggcttgag atgttcgggc tgg 23

<210> 211

<211> 23

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 211

ccggctgggg ctgtcctcgc tag 23

<210> 212

<211> 23

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 212

cggggtgcag ctgctcacgc cag 23

<210> 213

<211> 23

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 213

ctggcgggag ctggtcgcgt gag 23

<210> 214

<211> 23

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 214

tgagcatggg ccgctggcgg tgg 23

<210> 215

<211> 23

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 215

atggcatagg ccgctgacag agg 23

<210> 216

<211> 23

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 216

ttggcatggt gagctggcgg ggg 23

<210> 217

<211> 23

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 217

tgggcagggg tctctgaggg cag 23

<210> 218

<211> 23

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 218

ttggcatggg tctcttacca agg 23

<210> 219

<211> 23

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 219

acatggttcc agtgggtatg tag 23

<210> 220

<211> 23

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"
<400> 220
ggaggtgggc agcgggtatg tag 23
<210> 221
<211> 23
<212> DNA
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"
<400> 221
agaaggtccc cgcgggcatg gag 23
<210> 222
<211> 23
<212> DNA
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"
<400> 222
ggagggaacc agccggtatg ggg 23
<210> 223
<211> 23
<212> DNA
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"
<400> 223
agagagtggc agtgggtaag cag 23
<210> 224
<211> 23

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 224

agaggtggcc agcgggcagg aag 23

<210> 225

<211> 23

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 225

tgaggggcc agctgggatg cag 23

<210> 226

<211> 120

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 226

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 60

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 120

<210> 227

<211> 30

<212> DNA

<213> Artificial Sequence

<

220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 227

cgcacgcgta attcgaacgc tgacgtcatc 30

<210> 228

<211> 130

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<220><221> modified_base

<222> (92)..(110)

<223> a, c, t, g, unknown or other

<400> 228

cacacgcgta aaaaagcacc gactcggtgc cacttttca agttgataac ggactagcct 60

tattttaact tgctatttct agctctaaaa cnnnnnnnnn nnnnnnnnnn cggtgtttcg 120

tcctttccac 130

<210> 229

<211> 24

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 229

gtgtctagac tgcagagggc cctg 24

<210> 230

<211> 28

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 230

gtgtcgtgcc tgagagcgca gtcgagaa 28

<210> 231

<211> 29

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 231

gagaagctta gctgaatggg gtccgcctc 29

<210> 232

<211> 31

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 232

ctcaccggtg cgcgcaaccg atgccgggac c 31

<210> 233

<211> 32

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 233

gagaagcttg gcgaaatgat ttgctgcaga tg 32

<210> 234

<211> 32

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 234

ctcaccggtg cgcgcgtcgc ctccccctcc gc 32

<210> 235

<211> 134

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 235

agcttcgcgc cgggaggagg ggggacgcag tgggcggagc ggagacagca ccttcggaga 60

taatcctttc tcctgccgca gagcagagga gcggcgggag aggaacactt ctcccaggct 120

ttagcagagc cgga 134

<210> 236

<211> 134

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 236

ccggtccggc tctgctaaag cctgggagaa gtgttcctct cccgccgctc ctctgctctg 60

cggcaggaga aaggattatc tccgaaggtg ctgtctccgc tccgcccact gcgtcccccc 120

tcctcccggc gcga 134

<210> 237

<211> 55

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 237

aattcaataa aagatcttta ttttcattag atctgtgtgt tggttttttg tgtgc 55

<210> 238

<211> 55

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 238

ggccgcacac aaaaaaccaa cacacagatc taatgaaaat aaagatcttt tattg 55

<210> 239

<211> 20

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 239

ggtctcatgt gtggcactca 20

<210> 240

<211> 21

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 240

atccctcctc agagggtcag c 21

<210> 241

<211> 22

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 241

ttcgggcata gcatggtctt cc 22

<210> 242

<211> 21

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 242

gttctgagcc gcacagtttg g 21

<210> 243

<211> 20

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 243

tgtccaacct tcaggcaagg 20

<210> 244

<211> 22

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 244

tacctcatgc acagctagca cc 22

<210> 245

<211> 24

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 245

gttctatttc agagggctga tccc 24

<210> 246

<211> 22

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 246

ggataagaag ggacaataca gg 22

<210> 247

<211> 21

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 247

gccggggtct cgttcagagc t 21

<210> 248

<211> 24

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 248

cctgtctctc tgtcctaggg ctcc 24

<210> 249

<211> 25

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 249

cccacaggaa acaatgaagg gagac 25

<210> 250

<211> 23

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 250

ccctgacacc agctgttcag cac 23

<210> 251

<211> 21

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 251

gagcaggcag agccgagcaa g 21

<210> 252

<211> 24

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 252

gggtcttgtt gtgagtaggg tgtg 24

<210> 253

<211> 23

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 253

cctggaatac tatttccacg ccg 23

<210> 254

<211> 22

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 254

gcagcacact ccaccctcac at 22

<210> 255

<211

> 23

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 255

cggcgtcacg tgacctgagt aac 23

<210> 256

<211> 21

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
primer"

<400> 256

gtgtctgcct cgctctgctg c 21

<210> 257

<211> 27

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
primer"

<400> 257

ctatctgaaa tccaccacct tagacgc 27

<210> 258

<211> 19

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
primer"

<400> 258

gcctgagggg gccagaggt 19

<210> 259

<211> 18

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
primer"

<400> 259

gggactcccc gggtggtg 18

<210> 260

<211> 24

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 260

gaccctgtgt ttcaagtctc tctg 24

<210> 261

<211> 23

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 261

ggccatccag tacattcaat acg 23

<210> 262

<211> 22

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 262

gcagaagccg tgactcacag ca 22

<210> 263

<211> 28

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 263

cttgtgcttg tgattctgtc cttactgc 28

<210> 264

<211> 25

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 264

cgagaagtcg atgcagacac ttcaa 25

<210> 265

<211> 23

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 265

gctgaagact ggcgagcaca gct 23

<210> 266

<211> 26

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 266

gacctgaatg ttgtggctga gagtcc 26

<210> 267

<211> 22

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 267

ccctcacgtt cctgtccagc aa 22

<210> 268

<211> 20

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 268

ctaccgcctg cggacatggt 20

<210> 269

<211> 25

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 269

ccgtttgctg atgtagtagg ggtcc 25

<210> 270

<211

> 25

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 270

catccttcgt gtctgaggac tggtc 25

<210> 271

<211> 23

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
primer"

<400> 271

ctctgggtga ccacacacga tgc 23

<210> 272

<211> 19

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
primer"

<400> 272

ggagagcgtc cgccaggag 19

<210> 273

<211> 22

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
primer"

<400> 273

gaagctctct taactactgt tc 22

<210> 274

<211> 20

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
primer"

<400> 274

caggccctag cagcgagcag 20

<210> 275

<211> 21

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 275

ggaaggggct ttcctccgag c 21

<210> 276

<211> 20

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 276

gctccgaccc tgctctccca 20

<210> 277

<211> 23

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 277

cctgttcatc aggctcgtag ccc 23

<210> 278

<211> 21

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 278

cgattgctgg cttgccttga g 21

<210> 279

<211> 22

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 279

ggttcgcgtc cgcccgcgtg at 22

<210> 280

<211> 23

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 280

gagaggtggt cctgtcgcct atg 23

<210> 281

<211> 23

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 281

cccagcaggt cacagctgac atc 23

<210> 282

<211> 32

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 282

agcacagtat gtattctata aaataatacg ac 32

<210> 283

<211> 22

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 283

gtgggagggg acagagacca tg 22

<210> 284

<211> 28

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 284

ccttacctgt tcctcttcct tatccagc 28

<210> 285

<211

> 22

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 285

atacccagtc cacatccctg cc 22

<210> 286

<211> 26

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 286

cctctgcatc tccctcagga agtatt 26

<210> 287

<211> 20

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 287

gcctcctgtc cccaggtccc 20

<210> 288

<211> 22

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 288

cactaggctt gggctgccct ct 22

<210> 289

<211> 23

<212> DNA

<213> Escherichia coli

<400> 289

tgcgaatacg cccacgcgat ggg 23

<210> 290

<211> 102

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 290

gttttagagc tatgctgttt tgaatggtcc caaaacggaa gggcctgagt ccgagcagaa 60

gaagaagttt tagagctatg ctgttttgaa tggtcccaaa ac 102

<210> 291

<211> 100

<212> DNA

<213> Homo sapiens

<400> 291

cggaggacaa agtacaaacg gcagaagctg gaggaggaag ggcctgagtc cgagcagaag 60

aagaagggct cccatcacat caaccggtgg cgcattgcca 100

<210> 292

<211> 50

<212> DNA

<213> Homo sapiens

<400> 292

agctggagga ggaagggcct gagtccgagc agaagaagaa gggctcccac 50

<210> 293

<211> 30

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 293

gaguccgagc agaagaagaa guuuuagagc 30

<210> 294

<211> 49

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 294

agctggagga ggaagggcct gagtccgagc agaagagaag ggctcccat 49

<210> 295

<211> 53

<212> DNA

<213> Homo sapiens

<400> 295

ctggaggagg aagggcctga gtccgagcag aagaagaagg gctcccatca cat 53

<210> 296

<211> 52

<212> DNA

<213> Homo sapiens

<400> 296

ctggaggagg aagggcctga gtccgagcag aagagaaggg ctcccatcac at 52

<210> 297

<211> 54

<212> DNA

<213> Homo sapiens

<400> 297

ctggaggagg aagggcctga gtccgagcag aagaaagaag ggctcccatc acat 54

<210> 298

<211> 50

<212> DNA

<213> Homo sapiens

<400> 298

ctggaggagg aagggcctga gtccgagcag aagaagggct cccatcacat 50

<210> 299

<211> 47

<212> DNA

<213> Homo sapiens

<400> 299

ctggaggagg aagggcctga gcccgagcag aagggctccc atcacat 47

<210> 300

<211> 48

<212> DNA

<213> Homo sapiens

<400> 300

ctggaggagg aagggcctga gtccgagcag aagaagaagg gctcccat 48

<210> 301

<211> 20

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 301

gaguccgagc agaagaagau 20

<210> 302

<211> 20

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 302

gaguccgagc agaagaagua 20

<210> 303

<211> 20

<212> RNA

<213> Artificial Sequence

<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"
<400> 303
gaguccgagc agaagaacaa 20
<210> 304
<211> 20
<212> RNA
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"
<400> 304
gaguccgagc agaagaugaa 20

<210> 305
<211> 20
<212> RNA
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"
<400> 305
gaguccgagc agaaguagaa 20
<210> 306
<211> 20
<212> RNA
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"
<400> 306
gaguccgagc agaugaagaa 20

<210> 307

<211> 20

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 307

gaguccgagc acaagaagaa 20

<210> 308

<211> 20

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 308

gaguccgagg agaagaagaa 20

<210> 309

<211> 20

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 309

gaguccgugc agaagaagaa 20

<210> 310

<211> 20

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 310

gagucggagc agaagaagaa 20

<210> 311

<211> 20

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 311

gagaccgagc agaagaagaa 20

<210> 312

<211> 24

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 312

aatgacaagc ttgctagcgg tggg 24

<210> 313

<211> 39

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 313

aaaacggaag ggcctgagtc cgagcagaag aagaagttt 39

<210> 314

<211> 39

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 314

aaacaggggc cgagattggg tgttcagggc agaggtttt 39

<210> 315

<211> 38

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 315

aaaacggaag ggcctgagtc cgagcagaag aagaagtt 38

<210> 316

<211> 40

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 316

aacggaggga ggggcacaga tgagaaactc agggttttag 40

<210> 317

<211> 38

<212> DNA

<213> Homo sapiens

<400> 317

agcccttctt cttctgctcg gactcaggcc cttcctcc 38

<210> 318

<211> 40

<212> DNA

<213> Homo sapiens

<400> 318

cagggaggga ggggcacaga tgagaaactc aggaggcccc 40

<210> 319

<211> 80

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 319

ggcaatgcgc caccggttga tgtgatggga gcccttctag gaggccccca gagcagccac 60

tggggcctca acactcaggc 80

<210> 320

<211> 33

<212> DNA

<213> Homo sapiens

<400> 320

catcgatgtc ctccccattg gcctgcttcg tgg 33

<210> 321

<211> 33

<212> DNA

<213> Homo sapiens

<400> 321

ttcgtggcaa tgcgccaccg gttgatgtga tgg 33

<210> 322

<211> 33

<212> DNA

<213> Homo sapiens

<400> 322

tcgtggcaat gcgccaccgg ttgatgtgat ggg 33

<

210> 323

<211> 33

<212> DNA

<213> Homo sapiens

<400> 323

tccagcttct gccgtttgta ctttgtcctc cgg 33

<210> 324

<211> 33

<212> DNA

<213> Homo sapiens

<400> 324

ggagggaggg gcacagatga gaaactcagg agg 33

<210> 325

<211> 33

<212> DNA

<213> Homo sapiens

<400> 325

aggggccgag attgggtgtt cagggcagag agg 33

<210> 326

<211> 33

<212> DNA

<213> Mus musculus

<400> 326

caagcactga gtgccattag ctaaatgcat agg 33

<210> 327

<211> 33

<212> DNA

<213> Mus musculus

<400> 327

aatgcatagg gtaccaccca caggtgccag ggg 33

<210> 328

<211> 33

<212> DNA

<213> Mus musculus

<400> 328

acacacatgg gaaagcctct gggccaggaa agg 33

<210> 329

<211> 37

<212> DNA

<213> Homo sapiens

<400> 329

ggaggaggta gtatacagaa acacagagaa gtagaat 37

<210> 330

<211> 37

<212> DNA

<213> Homo sapiens

<400> 330

agaatgtaga ggagtcacag aaactcagca ctagaaa 37

<210> 331

<211> 98

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 331

ggacgaaaca ccggaaccat tcaaaacagc atagcaagtt aaaataaggc tagtccgtta 60

tcaacttgaa aaagtggcac cgagtcggtg ctttttt 98

<210> 332

<211> 186

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 332

ggacgaaaca ccggtagtat taagtattgt tttatggctg ataaatttct ttgaatttct 60

ccttgattat ttgttataaa agttataaaa taatcttgtt ggaaccattc aaaacagcat 120

agcaagttaa aataaggcta gtccgttatc aacttgaaaa agtggcaccg agtcggtgct 180

tttttt 186

<210> 333

<211> 95

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 333

gggttttaga gctatgctgt tttgaatggt cccaaaacgg gtcttcgaga agacgtttta 60

gagctatgct gttttgaatg gtcccaaaac ttttt 95

<210> 334

<211> 36

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<220><221> modified_base

<222> (5)..(34)

<223> a, c, t, g, unknown or other

<400> 334

aaacnnnnnn nnnnnnnnnn nnnnnnnnnn nnnngt 36

<210> 335

<211> 36

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<220><221> modified_base

<222> (7)..(36)

<223> a, c, t, g, unknown or other

<400> 335

taaaacnnnn nnnnnnnnnn nnnnnnnnnn nnnnnn 36

<210> 336

<211> 84

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 336

gtggaaagga cgaaacaccg ggtcttcgag aagacctgtt ttagagctag aaatagcaag 60

ttaaaataag gctagtccgt tttt 84

<210> 337

<211> 46

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (1)..(19)

<223> a, c, u, g, unknown or other

<400> 337

nnnnnnnnnn nnnnnnnnng uuauuguacu cucaagauuu auuuuu 46

<210> 338

<211> 91

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400

> 338

guuacuuaaa ucuugcagaa gcuacaaaga uaaggcuuca ugccgaaauc aacacccugu 60

cauuuuaugg caggguguuu ucguuauuua a 91

<210> 339

<211> 70

<212> DNA

<213> Homo sapiens

<400> 339

ttttctagtg ctgagtttct gtgactcctc tacattctac ttctctgtgt ttctgtatac 60

tacctcctcc 70

<210> 340

<211> 122

<212> DNA

<213> Homo sapiens

<400> 340

ggaggaaggg cctgagtccg agcagaagaa gaagggctcc catcacatca accggtggcg 60

cattgccacg aagcaggcca atggggagga catcgatgtc acctccaatg actagggtgg 120

gc 122

<210> 341

<211> 48

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (3)..(32)

<223> a, c, u, g, unknown or other

<400> 341

acnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnguuuuaga gcuaugcu 48

<210> 342

<211> 67

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> source

<223> /note="Description of Combined DNA/RNA Molecule: Synthetic

oligonucleotide"

<400> 342

agcauagcaa guuaaaauaa ggctaguccg uuaucaacuu gaaaaagugg caccgagucg 60

gugcuuu 67

<210> 343

<211> 62

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<220><221> modified_base

<222> (1)..(20)

<223> a, c, u, g, unknown or other

<400> 343

nnnnnnnnnn nnnnnnnnnn guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cg 62

<210> 344

<211> 73

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 344

tgaatggtcc caaaacggaa gggcctgagt ccgagcagaa gaagaagttt tagagctatg 60

ctgttttgaa tgg 73

<210> 345

<211> 69

<212> DNA

<213> Homo sapiens

<400> 345

ctggtcttcc acctctctgc cctgaacacc caatctcggc ccctctcgcc accctcctgc 60

atttctgtt 69

<210> 346

<211> 138

<212> DNA

<213> Mus musculus

<400> 346

acccaagcac tgagtgccat tagctaaatg catagggtac cacccacagg tgccaggggc 60

ctttcccaaa gttcccagcc ccttctccaa cctttcctgg cccagaggct ttcccatgtg 120

tgtggctgga ccctttga 138

<210> 347

<211> 21

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 347

aaaaccaccc ttctctctgg c 21

<210> 348

<211> 21

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 348

ggagattgga gacacggaga g 21

<210> 349

<211> 20

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 349

ctggaaagcc aatgcctgac 20

<210> 350

<211> 20

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 350

ggcagcaaac tccttgtcct 20

<210> 351

<211> 20

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 351

gtgctttgca gaggcctacc 20

<210> 352

<211> 20

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 352

cctggagcgc atgcagtagt 20

<210> 353

<211> 22

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 353

accttctgtg tttccaccat tc 22

<210> 354

<211> 20

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 354

ttggggagtg cacagacttc 20

<210> 355

<211> 30

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 355

tagctctaaa acttcttctt ctgctcggac 30

<210> 356

<211> 30

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 356

ctagccttat tttaacttgc tatgctgttt 30

<210> 357

<211> 99

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<220><221> modified_base

<222> (1)..(20)

<223> a, c, u, g, unknown or other

<400> 357

nnnnnnnnnn nnnnnnnnnn guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuu 99

<210> 358

<211> 12

<212> DNA

<213> Homo sapiens

<400> 358

tagcgggtaa gc 12

<210> 359

<211> 12

<212> DNA

<213> Homo sapiens

<400> 359

tcggtgacat gt 12

<210> 360

<211> 12

<212> DNA

<213> Homo sapiens

<400> 360

actccccgta gg 12

<210> 361

<211> 12

<212> DNA

<213> Homo sapiens

<400> 361

actgcgtgtt aa 12

<210> 362

<211> 12

<212> DNA

<213> Homo sapiens

<400> 362

acgtcgcctg at 12

<210> 363

<211> 12

<212> DNA

<213> Homo sapiens

<400> 363

taggtcgacc ag 12

<210> 364

<211> 12

<212> DNA

<213> Homo sapiens

<400> 364

ggcgttaatg at 12

<210> 365

<211> 12

<212> DNA

<213> Homo sapiens

<400> 365

tgtcgcatgt ta 12

<210> 366

<211> 12

<212> DNA

<213> Homo sapiens

<400> 366

atggaaacgc at 12

<210> 367

<211> 12

<212> DNA

<213> Homo sapiens

<400> 367

gccgaattcc tc 12

<210> 368

<211> 12

<212> DNA

<213> Homo sapiens

<400> 368

gcatggtacg ga 12

<210> 369

<211> 12

<212> DNA

<213> Homo sapiens

<400> 369

cggtactctt ac 12

<210> 370

<211> 12

<212> DNA

<213> Homo sapiens

<400> 370

gcctgtgccg ta 12

<210> 371

<211> 12

<212> DNA

<213> Homo sapiens

<400> 371

tacggtaagt cg 12

<210> 372

<211> 12

<212> DNA

<213> Homo sapiens

<400> 372

cacgaaatta cc 12

<210> 373

<211> 12

<212> DNA

<213> Homo sapiens

<400> 373

aaccaagata cg 12

<210> 374

<211> 12

<212> DNA

<213> Homo sapiens

<400> 374

gagtcgatac gc 12

<210> 375

<211> 12

<212> DNA

<213> Homo sapiens

<400> 375

gtctcacgat cg 12

<210> 376

<211> 12

<212> DNA

<213> Homo sapiens

<400> 376

tcgtcgggtg ca 12

<210> 377

<211> 12

<212> DNA

<213> Homo sapiens

<400> 377

actccgtagt ga 12

<210> 378

<211> 12

<212> DNA

<213> Homo sapiens

<400> 378

caggacgtcc gt 12

<210> 379

<211> 12

<212> DNA

<213> Homo sapiens

<400> 379

tcgtatccct ac 12

<210> 380

<211> 12

<212> DNA

<213> Homo sapiens

<400> 380

tttcaaggcc gg 12

<210> 381

<211> 12

<212> DNA

<213> Homo sapiens

<400> 381

cgccggtgga at 12

<210> 382

<211> 12

<212> DNA

<213> Homo sapiens

<400> 382

gaacccgtcc ta 12

<210> 383

<211> 12

<212> DNA

<213> Homo sapiens

<400> 383

gattcatcag cg 12

<210> 384

<211> 12

<212> DNA

<213> Homo sapiens

<400> 384

acaccggtct tc 12

<210> 385

<211> 12

<212> DNA

<213> Homo sapiens

<400> 385

atcgtgccct aa 12

<210> 386

<211> 12

<212> DNA

<213> Homo sapiens

<400> 386

gcgtcaatgt tc 12

<210> 387

<211> 12

<212> DNA

<213> Homo sapiens

<400> 387

ctccgtatct cg 12

<210> 388

<211> 12

<212> DNA

<213> Homo sapiens

<400> 388

ccgattcctt cg 12

<210> 389

<211> 12

<212> DNA

<213> Homo sapiens

<400> 389

tgcgcctcca gt 12

<210> 390

<211> 12

<212> DNA

<213> Homo sapiens

<400> 390

taacgtcgga gc 12

<210> 391

<211> 12

<212> DNA

<213> Homo sapiens

<400> 391

aaggtcgccc at 12

<210> 392

<211> 12

<212> DNA

<213> Homo sapiens

<400> 392

gtcggggact at 12

<210> 393

<211> 12

<212> DNA

<213> Homo sapiens

<400> 393

ttcgagcgat tt 12

<210> 394

<211> 12

<212> DNA

<213> Homo sapiens

<400> 394

tgagtcgtcg ag 12

<210> 395

<211> 12

<212> DNA

<213> Homo sapiens

<400> 395

tttacgcaga gg 12

<210> 396

<211> 12

<212> DNA

<213> Homo sapiens

<400> 396

aggaagtatc gc 12

<210> 397

<211> 12

<212> DNA

<213> Homo sapiens

<400> 397

actcgatacc at 12

<210> 398

<211> 12

<212> DNA

<213> Homo sapiens

<400> 398

cgctacatag ca 12

<210> 399

<211> 12

<212> DNA

<213> Homo sapiens

<400> 399

ttcataaccg gc 12

<210> 400

<211> 12

<212> DNA

<213> Homo sapiens

<400> 400

ccaaacggtt aa 12

<210> 401

<211> 12

<212> DNA

<213> Homo sapiens

<400> 401

cgattccttc gt 12

<210> 402

<211> 12

<212> DNA

<213> Homo sapiens

<400> 402

cgtcatgaat aa 12

<210> 403

<211> 12

<212> DNA

<213> Homo sapiens

<400> 403

agtggcgatg ac 12

<210> 404

<211> 12

<212> DNA

<213> Homo sapiens

<400> 404

cccctacggc ac 12

<210> 405

<211> 12

<212> DNA

<213> Homo sapiens

<400> 405

gccaacccgc ac 12

<210> 406

<211> 12

<212> DNA

<213> Homo sapiens

<400> 406

tgggacaccg gt 12

<210> 407

<211> 12

<212> DNA

<213> Homo sapiens

<400> 407

ttgactgcgg cg 12

<210> 408

<211> 12

<212> DNA

<213> Homo sapiens

<400> 408

actatgcgta gg 12

<210> 409

<211> 12

<212> DNA

<213> Homo sapiens

<400> 409

tcacccaaag cg 12

<210> 410

<211> 12

<212> DNA

<213> Homo sapiens

<400> 410

gcaggacgtc cg 12

<210> 411

<211> 12

<212> DNA

<213> Homo sapiens

<400> 411

acaccgaaaa cg 12

<210> 412

<211> 12

<212> DNA

<213> Homo sapiens

<400> 412

cggtgtattg ag 12

<210> 413

<211> 12

<212> DNA

<213> Homo sapiens

<400> 413

cacgaggtat gc 12

<210> 414

<211> 12

<212> DNA

<213> Homo sapiens

<400> 414

taaagcgacc cg 12

<210> 415

<211> 12

<212> DNA

<213> Homo sapiens

<400> 415

cttagtcggc ca 12

<210> 416

<211> 12

<212> DNA

<213> Homo sapiens

<400> 416

cgaaaacgtg gc 12

<210> 417

<211> 12

<212> DNA

<213> Homo sapiens

<400> 417

cgtgccctga ac 12

<210> 418

<211> 12

<212> DNA

<213> Homo sapiens

<400> 418

tttaccatcg aa 12

<210> 419

<211> 12

<212> DNA

<213> Homo sapiens

<400> 419

cgtagccatg tt 12

<210> 420

<211> 12

<212> DNA

<213> Homo sapiens

<400> 420

cccaaacggt ta 12

<210> 421

<211> 12

<212> DNA

<213> Homo sapiens

<400> 421

gcgttatcag aa 12

<210> 422

<211> 12

<212> DNA

<213> Homo sapiens

<400> 422

tcgatggtaa ac 12

<210> 423

<211> 12

<212> DNA

<213> Homo sapiens

<400> 423

cgacttttg ca 12

<210> 424

<211> 12

<212> DNA

<213> Homo sapiens

<400> 424

tcgacgactc ac 12

<210> 425

<211> 12

<212> DNA

<213> Homo sapiens

<400> 425

acgcgtcaga ta 12

<210> 426

<211> 12

<212> DNA

<213> Homo sapiens

<400> 426

cgtacggcac ag 12

<210> 427

<211> 12

<212> DNA

<213> Homo sapiens

<400> 427

ctatgccgtg ca 12

<210> 428

<211> 12

<212> DNA

<213> Homo sapiens

<400> 428

cgcgtcagat at 12

<210> 429

<211> 12

<212> DNA

<213> Homo sapiens

<400> 429

aagatcggta gc 12

<210> 430

<211> 12

<212> DNA

<213> Homo sapiens

<400> 430

cttcgcaagg ag 12

<210> 431

<211> 12

<212> DNA

<213> Homo sapiens

<400> 431

gtcgtggact ac 12

<210> 432

<211> 12

<212> DNA

<213> Homo sapiens

<400> 432

ggtcgtcatc aa 12

<210> 433

<211> 12

<212> DNA

<213> Homo sapiens

<400> 433

gttaacagcg tg 12

<210> 434

<211> 12

<212> DNA

<213> Homo sapiens

<400> 434

tagctaaccg tt 12

<210> 435

<211> 12

<212> DNA

<213> Homo sapiens

<400> 435

agtaaaggcg ct 12

<210> 436

<211> 12

<212> DNA

<213> Homo sapiens

<400> 436

ggtaatttcg tg 12

<210> 437

<211> 15

<212> DNA

<213> Homo sapiens

<400> 437

cagaagaaga agggc 15

<210> 438

<211> 51

<212> DNA

<213> Homo sapiens

<400> 438

ccaatgggga ggacatcgat gtcacctcca atgactaggg tggtgggcaa c 51

<210> 439

<211> 15

<212> DNA

<213> Homo sapiens

<400> 439

ctctggccac tccct 15

<210> 440

<211> 52

<212> DNA

<213> Homo sapiens

<400> 440

acatcgatgt cacctccaat gacaagcttg ctagcggtgg gcaaccacaa ac 52

<210> 441

<211> 25

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (6)..(25)

<223> a, c, t, g, unknown or other

<400> 441

caccgnnnnn nnnnnnnnnn nnnnn 25

<210> 442

<211> 25

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (5)..(24)

<223> a, c, t, g, unknown or other

<400>

442

aaacnnnnnn nnnnnnnnnn nnnnc 25

<210> 443

<211> 54

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 443

aacaccgggt cttcgagaag acctgtttta gagctagaaa tagcaagtta aaat 54

<210> 444

<211> 54

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400

> 444

caaaacgggt cttcgagaag acgttttaga gctatgctgt tttgaatggt ccca 54

<210> 445

<211> 4104

<212> DNA

<213> Homo sapiens

<220><221> CDS

<222> (1)..(4104)

<400> 445

atg gac aag aag tac agc atc ggc ctg gac atc ggc acc aac tct gtg 48
Met Asp Lys Lys Tyr Ser Ile Gly Leu Asp Ile Gly Thr Asn Ser Val
1 5 10 15

ggc tgg gcc gtg atc acc gac gag tac aag gtg ccc agc aag aaa ttc 96
Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe

20 25 30

aag gtg ctg ggc aac acc gac cgg cac agc atc aag aag aac ctg atc 144

```
Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile
        35                  40                  45
gga gcc ctg ctg ttc gac agc ggc gaa aca gcc gag gcc acc cgg ctg        192
Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu
    50                  55                  60

aag aga acc gcc aga aga aga tac acc aga cgg aag aac cgg atc tgc        240
Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys
65                  70                  75                  80
tat ctg caa gag atc ttc agc aac gag atg gcc aag gtg gac gac agc        288
Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser
                85                  90                  95
ttc ttc cac aga ctg gaa gag tcc ttc ctg gtg gaa gag gat aag aag        336

Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys
            100                 105                 110
cac gag cgg cac ccc atc ttc ggc aac atc gtg gac gag gtg gcc tac        384
His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr
        115                 120                 125
cac gag aag tac ccc acc atc tac cac ctg aga aag aaa ctg gtg gac        432
His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp

    130                 135                 140
agc acc gac aag gcc gac ctg cgg ctg atc tat ctg gcc ctg gcc cac        480
Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His
145                 150                 155                 160
atg atc aag ttc cgg ggc cac ttc ctg atc gag ggc gac ctg aac ccc        528
Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro
                165                 170                 175

gac aac agc gac gtg gac aag ctg ttc atc cag ctg gtg cag acc tac        576
Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr
            180                 185                 190
aac cag ctg ttc gag gaa aac ccc atc aac gcc agc ggc gtg gac gcc        624
Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala
```

```
        195                 200                 205
aag gcc atc ctg tct gcc aga ctg agc aag agc aga cgg ctg gaa aat        672

Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn
    210                 215                 220
ctg atc gcc cag ctg ccc ggc gag aag aag aat ggc ctg ttc ggc aac        720
Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn
225                 230                 235                 240
ctg att gcc ctg agc ctg ggc ctg acc ccc aac ttc aag agc aac ttc        768
Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe

                245                 250                 255
gac ctg gcc gag gat gcc aaa ctg cag ctg agc aag gac acc tac gac        816
Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp
            260                 265                 270
gac gac ctg gac aac ctg ctg gcc cag atc ggc gac cag tac gcc gac        864
Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp
        275                 280                 285

ctg ttt ctg gcc gcc aag aac ctg tcc gac gcc atc ctg ctg agc gac        912
Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp
    290                 295                 300
atc ctg aga gtg aac acc gag atc acc aag gcc ccc ctg agc gcc tct        960
Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser
305                 310                 315                 320
atg atc aag aga tac gac gag cac cac cag gac ctg acc ctg ctg aaa       1008

Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys
                325                 330                 335
gct ctc gtg cgg cag cag ctg cct gag aag tac aaa gag att ttc ttc       1056
Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe
            340                 345                 350
gac cag agc aag aac ggc tac gcc ggc tac att gac ggc gga gcc agc       1104
```

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

355 360 365

cag gaa gag ttc tac aag ttc atc aag ccc atc ctg gaa aag atg gac 1152

Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp

370 375 380

ggc acc gag gaa ctg ctc gtg aag ctg aac aga gag gac ctg ctg cgg 1200

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

385 390 395 400

aag cag cgg acc ttc gac aac ggc agc atc ccc cac cag atc cac ctg 1248

Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu

405 410 415

gga gag ctg cac gcc att ctg cgg cgg cag gaa gat ttt tac cca ttc 1296

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

420 425 430

ctg aag gac aac cgg gaa aag atc gag aag atc ctg acc ttc cgc atc 1344

Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile

435 440 445

ccc tac tac gtg ggc cct ctg gcc agg gga aac agc aga ttc gcc tgg 1392

Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp

450 455 460

atg acc aga aag agc gag gaa acc atc acc ccc tgg aac ttc gag gaa 1440

Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu

465 470 475 480

gtg gtg gac aag ggc gct tcc gcc cag agc ttc atc gag cgg atg acc 1488

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

485 490 495

aac ttc gat aag aac ctg ccc aac gag aag gtg ctg ccc aag cac agc 1536

Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser

500 505 510

ctg ctg tac gag tac ttc acc gtg tat aac gag ctg acc aaa gtg aaa 1584
Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys
515 520 525
tac gtg acc gag gga atg aga aag ccc gcc ttc ctg agc ggc gag cag 1632
Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln
530 535 540
aaa aag gcc atc gtg gac ctg ctg ttc aag acc aac cgg aaa gtg acc 1680

Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr
545 550 555 560
gtg aag cag ctg aaa gag gac tac ttc aag aaa atc gag tgc ttc gac 1728
Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp
565 570 575
tcc gtg gaa atc tcc ggc gtg gaa gat cgg ttc aac gcc tcc ctg ggc 1776
Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly

580 585 590
aca tac cac gat ctg ctg aaa att atc aag gac aag gac ttc ctg gac 1824
Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp
595 600 605
aat gag gaa aac gag gac att ctg gaa gat atc gtg ctg acc ctg aca 1872
Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr
610 615 620

ctg ttt gag gac aga gag atg atc gag gaa cgg ctg aaa acc tat gcc 1920
Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala
625 630 635 640
cac ctg ttc gac gac aaa gtg atg aag cag ctg aag cgg cgg aga tac 1968
His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr
645 650 655
acc ggc tgg ggc agg ctg agc cgg aag ctg atc aac ggc atc cgg gac 2016

Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp
660 665 670
aag cag tcc ggc aag aca atc ctg gat ttc ctg aag tcc gac ggc ttc 2064

Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe

675 680 685

gcc aac aga aac ttc atg cag ctg atc cac gac gac agc ctg acc ttt 2112

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

690 695 700

aaa gag gac atc cag aaa gcc cag gtg tcc ggc cag ggc gat agc ctg 2160

Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu

705 710 715 720

cac gag cac att gcc aat ctg gcc ggc agc ccc gcc att aag aag ggc 2208

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

725 730 735

atc ctg cag aca gtg aag gtg gtg gac gag ctc gtg aaa gtg atg ggc 2256

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

740 745 750

cgg cac aag ccc gag aac atc gtg atc gcc atg gcc aga gag aac cag 2304

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

755 760 765

acc acc cag aag gga cag aag aac agc cgc gag aga atg aag cgg atc 2352

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

770 775 780

gaa gag ggc atc aaa gag ctg ggc agc cag atc ctg aaa gaa cac ccc 2400

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

785 790 795 800

gtg gaa aac acc cag ctg cag aac gag aag ctg tac ctg tac tac ctg 2448

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

805 810 815

cag aat ggg cgg gat atg tac gtg gac cag gaa ctg gac atc aac cgg 2496

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

820 825 830

ctg tcc gac tac gat gtg gac gcc atc gtg cct cag agc ttt ctg aag 2544

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

```
        835                 840                 845

gac gac tcc atc gac gcc aag gtg ctg acc aga agc gac aag gcc cgg       2592
Asp Asp Ser Ile Asp Ala Lys Val Leu Thr Arg Ser Asp Lys Ala Arg
    850                 855                 860
ggc aag agc gac aac gtg ccc tcc gaa gag gtc gtg aag aag atg aag       2640
Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys
865                 870                 875                 880
aac tac tgg cgg cag ctg ctg aac gcc aag ctg att acc cag aga aag       2688

Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys
                885                 890                 895
ttc gac aat ctg acc aag gcc gag aga ggc ggc ctg agc gaa ctg gat       2736
Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp
            900                 905                 910
aag gcc ggc ttc atc aag aga cag ctg gtg gaa acc cgg cag atc aca       2784
Lys Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr

        915                 920                 925
aag cac gtg gca cag atc ctg gac tcc cgg atg aac act aag tac gac       2832
Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp
    930                 935                 940
gag aat gac aag ctg atc cgg gaa gtg aaa gtg atc acc ctg aag tcc       2880
Glu Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser
945                 950                 955                 960

aag ctg gtg tcc gat ttc cgg aag gat ttc cag ttt tac aaa gtg cgc       2928
Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg
                965                 970                 975
gag atc aac aac tac cac cac gcc cac gcc gcc tac ctg aac gcc gtc       2976
Glu Ile Asn Asn Tyr His His Ala His Ala Ala Tyr Leu Asn Ala Val
            980                 985                 990
gtg gga acc gcc ctg atc aaa aag  tac cct aag ctg gaa  agc gag ttc     3024
```

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

995 1000 1005

gtg tac ggc gac tac aag gtg tac gac gtg cgg aag atg atc gcc 3069

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

1010 1015 1020

aag agc gag cag gaa atc ggc aag gct acc gcc aag tac ttc ttc 3114

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

1025 1030 1035

tac agc aac atc atg aac ttt ttc aag acc gag att acc ctg gcc 3159

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

1040 1045 1050

aac ggc gag atc cgg aag cgg cct ctg atc gag aca aac ggc gaa 3204

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

1055 1060 1065

acc ggg gag atc gtg tgg gat aag ggc cgg gat ttt gcc acc gtg 3249

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

1070 1075 1080

cgg aaa gtg ctg agc atg ccc caa gtg aat atc gtg aaa aag acc 3294

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

1085 1090 1095

gag gtg cag aca ggc ggc ttc agc aaa gag tct atc ctg ccc aag 3339

Glu Val Gln Thr Gly Gly Phe Ser Lys Glu Ser Ile Leu Pro Lys

1100 1105 1110

agg aac agc gat aag ctg atc gcc aga aag aag gac tgg gac cct 3384

Arg Asn Ser Asp Lys Leu Ile Ala Arg Lys Lys Asp Trp Asp Pro

1115 1120 1125

aag aag tac ggc ggc ttc gac agc ccc acc gtg gcc tat tct gtg 3429

Lys Lys Tyr Gly Gly Phe Asp Ser Pro Thr Val Ala Tyr Ser Val

1130 1135 1140

ctg gtg gtg gcc aaa gtg gaa aag ggc aag tcc aag aaa ctg aag 3474

Leu Val Val Ala Lys Val Glu Lys Gly Lys Ser Lys Lys Leu Lys

```
    1145                1150                1155
agt gtg  aaa gag ctg ctg ggg  atc acc atc atg gaa  aga agc agc        3519
Ser Val  Lys Glu Leu Leu Gly  Ile Thr Ile Met Glu  Arg Ser Ser
    1160                1165                1170

ttc gag  aag aat ccc atc gac  ttt ctg gaa gcc aag  ggc tac aaa        3564
Phe Glu  Lys Asn Pro Ile Asp  Phe Leu Glu Ala Lys  Gly Tyr Lys
    1175                1180                1185
gaa gtg  aaa aag gac ctg atc  atc aag ctg cct aag  tac tcc ctg        3609
Glu Val  Lys Lys Asp Leu Ile  Ile Lys Leu Pro Lys  Tyr Ser Leu
    1190                1195                1200
ttc gag  ctg gaa aac ggc cgg  aag aga atg ctg gcc  tct gcc ggc        3654

Phe Glu  Leu Glu Asn Gly Arg  Lys Arg Met Leu Ala  Ser Ala Gly
    1205                1210                1215
gaa ctg  cag aag gga aac gaa  ctg gcc ctg ccc tcc  aaa tat gtg        3699
Glu Leu  Gln Lys Gly Asn Glu  Leu Ala Leu Pro Ser  Lys Tyr Val
    1220                1225                1230
aac ttc  ctg tac ctg gcc agc  cac tat gag aag ctg  aag ggc tcc        3744
Asn Phe  Leu Tyr Leu Ala Ser  His Tyr Glu Lys Leu  Lys Gly Ser

    1235                1240                1245
ccc gag  gat aat gag cag aaa  cag ctg ttt gtg gaa  cag cac aag        3789
Pro Glu  Asp Asn Glu Gln Lys  Gln Leu Phe Val Glu  Gln His Lys
    1250                1255                1260
cac tac  ctg gac gag atc atc  gag cag atc agc gag  ttc tcc aag        3834
His Tyr  Leu Asp Glu Ile Ile  Glu Gln Ile Ser Glu  Phe Ser Lys
    1265                1270                1275

aga gtg  atc ctg gcc gac gct  aat ctg gac aaa gtg  ctg tcc gcc        3879
Arg Val  Ile Leu Ala Asp Ala  Asn Leu Asp Lys Val  Leu Ser Ala
    1280                1285                1290
tac aac  aag cac cgg gat aag  ccc atc aga gag cag  gcc gag aat        3924
Tyr Asn  Lys His Arg Asp Lys  Pro Ile Arg Glu Gln  Ala Glu Asn
    1295                1300                1305
```

atc atc cac ctg ttt acc ctg acc aat ctg gga gcc cct gcc gcc 3969

Ile Ile His Leu Phe Thr Leu Thr Asn Leu Gly Ala Pro Ala Ala
1310 1315 1320

ttc aag tac ttt gac acc acc atc gac cgg aag agg tac acc agc 4014

Phe Lys Tyr Phe Asp Thr Thr Ile Asp Arg Lys Arg Tyr Thr Ser
1325 1330 1335

acc aaa gag gtg ctg gac gcc acc ctg atc cac cag agc atc acc 4059

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

1340 1345 1350

ggc ctg tac gag aca cgg atc gac ctg tct cag ctg gga ggc gac 4104

Gly Leu Tyr Glu Thr Arg Ile Asp Leu Ser Gln Leu Gly Gly Asp
1355 1360 1365

<210> 446

<211> 1368

<212> PRT

<213> Homo sapiens

<400> 446

Met Asp Lys Lys Tyr Ser Ile Gly Leu Asp Ile Gly Thr Asn Ser Val
1 5 10 15

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

20 25 30

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

Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu
50 55 60

Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys
65 70 75 80

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

85 90 95

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

100 105 110

His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr
115 120 125
His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp
130 135 140
Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His

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

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

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

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

355 360 365

Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp

370 375 380

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

385 390 395 400

Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu

405 410 415

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

420 425 430

Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile

435 440 445

Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp

450 455 460

Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu

465 470 475 480

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

485 490 495

Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser

500 505 510

Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys

515 520 525

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

530 535 540

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

545 550 555 560

Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp

565 570 575

Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly

580 585 590

Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp

595 600 605

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

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

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

785 790 795 800
Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu
805 810 815
Gln Asn Gly Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg
820 825 830
Leu Ser Asp Tyr Asp Val Asp Ala Ile Val Pro Gln Ser Phe Leu Lys
835 840 845
Asp Asp Ser Ile Asp Ala Lys Val Leu Thr Arg Ser Asp Lys Ala Arg

850 855 860

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

865 870 875 880

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

885 890 895

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

900 905 910

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

915 920 925

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

930 935 940

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

945 950 955 960

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

965 970 975

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

980 985 990

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

995 1000 1005

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

1010 1015 1020

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

1025 1030 1035

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

1040 1045 1050

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

1055 1060 1065

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

1070 1075 1080

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

1085 1090 1095

Glu Val Gln Thr Gly Gly Phe Ser Lys Glu Ser Ile Leu Pro Lys

1100 1105 1110

Arg Asn Ser Asp Lys Leu Ile Ala Arg Lys Lys Asp Trp Asp Pro

1115 1120 1125

Lys Lys Tyr Gly Gly Phe Asp Ser Pro Thr Val Ala Tyr Ser Val

1130 1135 1140

Leu Val Val Ala Lys Val Glu Lys Gly Lys Ser Lys Lys Leu Lys

1145 1150 1155

Ser Val Lys Glu Leu Leu Gly Ile Thr Ile Met Glu Arg Ser Ser

1160 1165 1170

Phe Glu Lys Asn Pro Ile Asp Phe Leu Glu Ala Lys Gly Tyr Lys

1175 1180 1185

Glu Val Lys Lys Asp Leu Ile Ile Lys Leu Pro Lys Tyr Ser Leu

1190 1195 1200

Phe Glu Leu Glu Asn Gly Arg Lys Arg Met Leu Ala Ser Ala Gly

1205 1210 1215

Glu Leu Gln Lys Gly Asn Glu Leu Ala Leu Pro Ser Lys Tyr Val

1220 1225 1230

Asn Phe Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu Lys Gly Ser

1235 1240 1245

Pro Glu Asp Asn Glu Gln Lys Gln Leu Phe Val Glu Gln His Lys

1250 1255 1260

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

1265 1270 1275

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

1280 1285 1290

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

1295 1300 1305

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

1310 1315 1320

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

1325 1330 1335

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

1340 1345 1350

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

1355 1360 1365

<210> 447

<211> 57

<212> DNA

<213> Homo sapiens

<400> 447

ggcaccatta aagaaaatat cattggtgtt tcctatgatg aatatagata cagaagc 57

<210> 448

<211> 19

<212> PRT

<213> Homo sapiens

<400> 448

Gly Thr Ile Lys Glu Asn Ile Ile Gly Val Ser Tyr Asp Glu Tyr Arg

1 5 10 15

Tyr Arg Ser

<210> 449

<211> 48

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 449

attaaagaaa atatcattgg ctttgtttcc tatgatgaat atagatac 48

<210> 450

<211> 16

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

peptide"

<400> 450

Ile Lys Glu Asn Ile Ile Gly Phe Val Ser Tyr Asp Glu Tyr Arg Tyr

1 5 10 15

<210> 451

<211> 50

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<220><221> modified_base

<222> (3)..(48)

<223> a, c, t, g, unknown or other

<400> 451

ccnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnngg 50

<210> 452

<211> 46

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<220><221> modified_base

<222> (3)..(44)

<223> a, c, t, g, unknown or other

<400> 452

ccnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnngg 46

<210> 453

<211> 42

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (3)..(40)

<223> a, c, t, g, unknown or other

<400> 453

ccnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn gg 42

<210> 454

<211> 38

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (3)..(36)

<223> a, c, t, g, unknown or other

<400> 454

ccnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnngg 38

<210> 455

<211> 34

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220

><221> modified_base

<222> (3)..(32)

<223> a, c, t, g, unknown or other

<400> 455

ccnnnnnnnn nnnnnnnnnn nnnnnnnnnn nngg 34

<210> 456

<211> 30

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (3)..(28)

<223> a, c, t, g, unknown or other

<400> 456

ccnnnnnnnn nnnnnnnnnn nnnnnnnngg 30

<210> 457

<211> 26

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (3)..(24)

<223> a, c, t, g, unknown or other

<400> 457

ccnnnnnnnn nnnnnnnnnn nnnngg 26

<210> 458

<211> 22

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (3)..(20)

<223> a, c, t, g, unknown or other

<400> 458

ccnnnnnnnn nnnnnnnnnn gg 22

<210> 459

<211> 18

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (3)..(16)

<223> a, c, t, g, unknown or other

<400> 459

ccnnnnnnnn nnnnnngg 18

<210> 460

<211> 16

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (3)..(14)

<223> a, c, t, g, unknown or other

<400> 460

ccnnnnnnnn nnnngg 16

<210> 461

<211> 15

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (3)..(13)

<223> a, c, t, g, unknown or other

<400> 461

ccnnnnnnnn nnngg 15

<210> 462

<211> 14

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<220><221> modified_base

<222> (3)..(12)

<223> a, c, t, g, unknown or other

<400> 462

ccnnnnnnnn nngg 14

<210> 463

<211> 13

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<220><221> modified_base

<222> (3)..(11)

<223> a, c, t, g, unknown or other

<400> 463

ccnnnnnnnn ngg 13

<210> 464

<211> 12

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (3)..(10)

<223> a, c, t, g, unknown or other

<400> 464

ccnnnnnnnn gg 12

<210> 465

<211> 11

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (3)..(9)

<223> a, c, t, g, unknown or other

<400> 465

ccnnnnnnng g 11

<210> 466

<211> 10

<212> DNA

<213> Artificial Sequence

<220><221> source

<223

> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (3)..(8)

<223> a, c, t, g, unknown or other

<400> 466

ccnnnnnngg 10

<210> 467

<211> 12

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220><221> modified_base

<222> (3)..(10)

<223> a, c, t, g, unknown or other

<400> 467

ggnnnnnnnn cc 12

<210> 468

<211> 125

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
primer"

<400> 468

aaaaaaagca ccgactcggt gccacttttt caagttgata acggactagc cttattttaa 60

cttgctattt ctagctctaa aactcacatc aaccggtggc gcaggtgttt cgtccttcc 120

acaag 125

<210> 469

<211> 125

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
primer"

<400> 469

aaaaaaagca ccgactcggt gccacttttt caagttgata acggactagc cttattttaa 60

cttgctattt ctagctctaa aactcacatc aaccggtggc gcaggtgttt cgtccttcc 120

acaag 125

<210> 470

<211> 125

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 470

aaaaaaagca ccgactcggt gccacttttt caagttgata acggactagc cttattttaa 60

cttgctattt ctagctctaa aacgaggaca aagtacaaac ggcggtgttt cgtcctttcc 120

acaag 125

<210> 471

<211> 125

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 471

aaaaaaagca ccgactcggt gccacttttt caagttgata acggactagc cttattttaa 60

cttgctattt ctagctctaa aacgaggaca aagtacaaac ggcggtgttt cgtcctttcc 120

acaag 125

<210> 472

<211> 125

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 472

aaaaaaagca ccgactcggt gccacttttt caagttgata acggactagc cttattttaa 60

cttgctattt ctagctctaa aacgtggcgc attgccacga agcggtgttt cgtcctttcc 120

acaag 125

<210> 473

<211> 125

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 473

aaaaaaagca ccgactcggt gccacttttt caagttgata acggactagc cttattttaa 60

cttgctattt ctagctctaa aaccgagggc agagtgctgc ttgggtgttt cgtcctttcc 120

acaag 125

<210> 474

<211> 125

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 474

aaaaaaagca ccgactcggt gccacttttt caagttgata acggactagc cttattttaa 60

cttgctattt ctagctctaa aacgagtccg agcagaagaa gaaggtgttt cgtcctttcc 120

acaag 125

<210> 475

<211> 125

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 475

aaaaaaagca ccgactcggt gccacttttt caagttgata acggactagc cttattttaa 60

cttgctattt ctagctctaa aacgaggaca aagtacaaac ggcggtgttt cgtcctttcc 120

acaag 125

<210> 476

<211> 125

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 476

aaaaaaagca ccgactcggt gccacttttt caagttgata acggactagc cttattttaa 60

cttgctattt ctagctctaa aacagcagaa gaagaagggc tccggtgttt cgtcctttcc 120

acaag 125

<210> 477

<211> 125

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 477

aaaaaaagca ccgactcggt gccacttttt caagttgata acggactagc cttattttaa 60

cttgctattt ctagctctaa aactcacatc aaccggtggc gcaggtgttt cgtcctttcc 120

acaag 125

<210> 478

<211> 125

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 478

aaaaaaagca ccgactcggt gccacttttt caagttgata acggactagc cttattttaa 60

cttgctattt ctagctctaa aacccctggc ccaggtgaag gtgggtgttt cgtcctttcc 120

acaag 125

<210> 479

<211> 125

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 479

aaaaaaagca ccgactcggt gccacttttt caagttgata acggactagc cttattttaa 60

cttgctattt ctagctctaa aactccctcc ctggcccagg tgaggtgttt cgtcctttcc 120

acaag 125

<210> 480

<211> 125

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 480

aaaaaaagca ccgactcggt gccacttttt caagttgata acggactagc cttattttaa 60

cttgctattt ctagctctaa aacgaaccgg aggacaaagt acaggtgttt cgtcctttcc 120

acaag 125

<210> 481

<211> 125

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 481

aaaaaaagca ccgactcggt gccacttttt caagttgata acggactagc cttattttaa 60

cttgctattt ctagctctaa aacaggtgaa ggtgtggttc cagggtgttt cgtcctttcc 120

acaag 125

<210> 482

<211> 125

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 482

aaaaaaagca ccgactcggt gccacttttt caagttgata acggactagc cttattttaa 60

cttgctattt ctagctctaa aacggtgaag gtgtggttcc agaggtgttt cgtcctttcc 120

acaag 125

<210> 483

<211> 125

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 483

aaaaaaagca ccgactcggt gccacttttt caagttgata acggactagc cttattttaa 60

cttgctattt ctagctctaa aacgaaccgg aggacaaagt acaggtgttt cgtcctttcc 120

acaag 125

<210> 484

<211> 125

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 484

aaaaaaagca ccgactcggt gccacttttt caagttgata acggactagc cttattttaa 60

cttgctattt ctagctctaa aacccctggc ccaggtgaag gtgggtgttt cgtcctttcc 120

acaag 125

<210> 485

<211> 125

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 485

aaaaaaagca ccgactcggt gccacttttt caagttgata acggactagc cttattttaa 60

cttgctattt ctagctctaa aacaggtgaa ggtgtggttc cagggtgttt cgtcctttcc 120

acaag 125

<210> 486

<211> 125

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 486

aaaaaaagca ccgactcggt gccacttttt caagttgata acggactagc cttattttaa 60

cttgctattt ctagctctaa aacgaggaca aagtacaaac ggcggtgttt cgtcctttcc 120

acaag 125

<210> 487

<211> 125

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 487

aaaaaaagca ccgactcggt gccacttttt caagttgata acggactagc cttattttaa 60

cttgctattt ctagctctaa aacgggaggg aggggcacag atgggtgttt cgtcctttcc 120

acaag 125

<210> 488

<211> 125

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 488

aaaaaaagca ccgactcggt gccacttttt caagttgata acggactagc cttattttaa 60

cttgctattt ctagctctaa aaccaccttc acctgggcca gggggtgttt cgtcctttcc 120

acaag 125

<210> 489

<211> 125

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 489

aaaaaaagca ccgactcggt gccacttttt caagttgata acggactagc cttattttaa 60

cttgctattt ctagctctaa aacaccctag tcattggagg tgaggtgttt cgtcctttcc 120

acaag 125

<210> 490

<211> 125

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 490

aaaaaaagca ccgactcggt gccacttttt caagttgata acggactagc cttattttaa 60

cttgctattt ctagctctaa aaccagagca gccactgggg cctggtgttt cgtccttcc 120

acaag 125

<210> 491

<211> 125

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 491

aaaaaaagca ccgactcggt gccacttttt caagttgata acggactagc cttattttaa 60

cttgctattt ctagctctaa aaccaccttc acctgggcca gggggtgttt cgtcctttcc 120

acaag 125

<210> 492

<211> 125

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 492

aaaaaaagca ccgactcggt gccacttttt caagttgata acggactagc cttattttaa 60

cttgctattt ctagctctaa aacccccatt ggcctgcttc gtgggtgttt cgtcctttcc 120

acaag 125

<210> 493

<211> 125

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 493

aaaaaaagca ccgactcggt gccacttttt caagttgata acggactagc cttattttaa 60

cttgctattt ctagctctaa aacattggcc tgcttcgtgg caaggtgttt cgtcctttcc 120

acaag 125

<210> 494

<211> 125

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 494

aaaaaaagca ccgactcggt gccacttttt caagttgata acggactagc cttattttaa 60

cttgctattt ctagctctaa aactcctcct ccagcttctg ccgggtgttt cgtcctttcc 120

acaag 125

<210> 495

<211> 125

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 495

aaaaaaagca ccgactcggt gccacttttt caagttgata acggactagc cttattttaa 60

cttgctattt ctagctctaa aaccctccag cttctgccgt ttgggtgttt cgtcctttcc 120

acaag 125

<210> 496

<211> 125

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 496

aaaaaaagca ccgactcggt gccacttttt caagttgata acggactagc cttattttaa 60

cttgctattt ctagctctaa aacattggcc tgcttcgtgg caaggtgttt cgtcctttcc 120

acaag 125

<210> 497

<211> 125

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 497

aaaaaaagca ccgactcggt gccacttttt caagttgata acggactagc cttattttaa 60

cttgctattt ctagctctaa aacgcagcaa gcagcactct gccggtgttt cgtcctttcc 120

acaag 125

<210> 498

<211> 125

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 498

aaaaaaagca ccgactcggt gccacttttt caagttgata acggactagc cttattttaa 60

cttgctattt ctagctctaa aacttcttct tctgctcgga ctcggtgttt cgtcctttcc 120

acaag 125

<210> 499

<211> 125

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 499

aaaaaaagca ccgactcggt gccacttttt caagttgata acggactagc cttattttaa 60

cttgctattt ctagctctaa aacaccggag gacaaagtac aaaggtgttt cgtccttcc 120

acaag 125

<210> 500

<211> 125

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 500

aaaaaaagca ccgactcggt gccacttttt caagttgata acggactagc cttattttaa 60

cttgctattt ctagctctaa aactcttctt ctgctcggac tcaggtgttt cgtccttcc 120

acaag 125

<210> 501

<211> 125

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 501

aaaaaaagca ccgactcggt gccacttttt caagttgata acggactagc cttattttaa 60

cttgctattt ctagctctaa aacgttgatg tgatgggagc cctggtgttt cgtccttcc 120

acaag 125

<210> 502

<211> 125

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 502

aaaaaaagca ccgactcggt gccacttttt caagttgata acggactagc cttattttaa 60

cttgctattt ctagctctaa aacgggccag ggagggaggg gcaggtgttt cgtcctttcc 120

acaag 125

<210> 503

<211> 125

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 503

aaaaaaagca ccgactcggt gccacttttt caagttgata acggactagc cttattttaa 60

cttgctattt ctagctctaa aacgggaggg aggggcacag atgggtgttt cgtcctttcc 120

acaag 125

<210> 504

<211> 125

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 504

aaaaaaagca ccgactcggt gccacttttt caagttgata acggactagc cttattttaa 60

cttgctattt ctagctctaa aacccggttc tggaaccaca cctggtgttt cgtcctttcc 120

acaag 125

<210> 505

<211> 125

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 505

aaaaaaagca ccgactcggt gccacttttt caagttgata acggactagc cttattttaa 60

cttgctattt ctagctctaa aactcacctg ggccagggag ggaggtgttt cgtcctttcc 120

acaag 125

<210> 506

<211> 125

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 506

aaaaaaagca ccgactcggt gccacttttt caagttgata acggactagc cttattttaa 60

cttgctattt ctagctctaa aactcacctg ggccagggag ggaggtgttt cgtcctttcc 120

acaag 125

<210> 507

<211> 125

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 507

aaaaaaagca ccgactcggt gccacttttt caagttgata acggactagc cttattttaa 60

cttgctattt ctagctctaa aacgttctgg aaccacacct tcaggtgttt cgtcctttcc 120

acaag 125

<210> 508

<211> 125

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 508

aaaaaaagca ccgactcggt gccacttttt caagttgata acggactagc cttattttaa 60

cttgctattt ctagctctaa aacgggaggg aggggcacag atgggtgttt cgtcctttcc 120

acaag 125

<210> 509

<211> 125

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 509

aaaaaaagca ccgactcggt gccacttttt caagttgata acggactagc cttattttaa 60

cttgctattt ctagctctaa aacgggccag ggagggaggg gcaggtgttt cgtcctttcc 120

acaag 125

<210> 510

<211> 125

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 510

aaaaaaagca ccgactcggt gccacttttt caagttgata acggactagc cttattttaa 60

cttgctattt ctagctctaa aacgttctgg aaccacacct tcaggtgttt cgtcctttcc 120

acaag 125

<210> 511

<211> 125

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 511

aaaaaaagca ccgactcggt gccacttttt caagttgata acggactagc cttattttaa 60

cttgctattt ctagctctaa aacaggtgaa ggtgtggttc cagggtgttt cgtcctttcc 120

acaag 125

<210> 512

<211> 125

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 512

aaaaaaagca ccgactcggt gccacttttt caagttgata acggactagc cttattttaa 60

cttgctattt ctagctctaa aacgaaccgg aggacaaagt acaggtgttt cgtcctttcc 120

acaag 125

<210> 513

<211> 125

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic primer"

<400> 513

aaaaaaagca ccgactcggt gccacttttt caagttgata acggactagc cttattttaa 60

cttgctattt ctagctctaa aaccaaaccc acgagggcag agtggtgttt cgtcctttcc 120

acaag 125

<210> 514

<211> 125

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

primer"

<400> 514

aaaaaaagca ccgactcggt gccacttttt caagttgata acggactagc cttattttaa 60

cttgctattt ctagctctaa aacgagtttc tcatctgtgc cccggtgttt cgtcctttcc 120

acaag 125

<210> 515

<211> 684

<212> DNA

<213> Homo sapiens

<220><221> CDS

<222> (1)..(123)

<220><221> CDS

<222> (127)..(159)

<220><221> CDS

<222> (163)..(399)

<220><221> CDS

<222> (403)..(684)

<400> 515

aaa acc acc ctt ctc tct ggc cca ctg tgt cct ctt cct gcc ctg cca 48

Lys Thr Thr Leu Leu Ser Gly Pro Leu Cys Pro Leu Pro Ala Leu Pro

1 5 10 15

tcc cct tct gtg aat gtt aga ccc atg gga gca gct ggt cag agg gga 96

Ser Pro Ser Val Asn Val Arg Pro Met Gly Ala Ala Gly Gln Arg Gly

20 25 30

ccc cgg cct ggg gcc cct aac cct atg tag cct cag tct tcc cat cag 144

Pro Arg Pro Gly Ala Pro Asn Pro Met Pro Gln Ser Ser His Gln

35 40 45

gct ctc agc tca gcc tga gtg ttg agg ccc cag tgg ctg ctc tgg ggg 192

Ala Leu Ser Ser Ala Val Leu Arg Pro Gln Trp Leu Leu Trp Gly

50 55 60

cct cct gag ttt ctc atc tgt gcc cct ccc tcc ctg gcc cag gtg aag 240

Pro Pro Glu Phe Leu Ile Cys Ala Pro Pro Ser Leu Ala Gln Val Lys

65 70 75

gtg tgg ttc cag aac cgg agg aca aag tac aaa cgg cag aag ctg gag 288
Val Trp Phe Gln Asn Arg Arg Thr Lys Tyr Lys Arg Gln Lys Leu Glu
80 85 90
gag gaa ggg cct gag tcc gag cag aag aag aag ggc tcc cat cac atc 336
Glu Glu Gly Pro Glu Ser Glu Gln Lys Lys Lys Gly Ser His His Ile
95 100 105 110
aac cgg tgg cgc att gcc acg aag cag gcc aat ggg gag gac atc gat 384

Asn Arg Trp Arg Ile Ala Thr Lys Gln Ala Asn Gly Glu Asp Ile Asp
115 120 125
gtc acc tcc aat gac tag ggt ggg caa cca caa acc cac gag ggc aga 432
Val Thr Ser Asn Asp Gly Gly Gln Pro Gln Thr His Glu Gly Arg
130 135 140
gtg ctg ctt gct gct ggc cag gcc cct gcg tgg gcc caa gct gga ctc 480
Val Leu Leu Ala Ala Gly Gln Ala Pro Ala Trp Ala Gln Ala Gly Leu

145 150 155
tgg cca ctc cct ggc cag gct ttg ggg agg cct gga gtc atg gcc cca 528
Trp Pro Leu Pro Gly Gln Ala Leu Gly Arg Pro Gly Val Met Ala Pro
160 165 170
cag ggc ttg aag ccc ggg gcc gcc att gac aga ggg aca agc aat ggg 576
Gln Gly Leu Lys Pro Gly Ala Ala Ile Asp Arg Gly Thr Ser Asn Gly
175 180 185

ctg gct gag gcc tgg gac cac ttg gcc ttc tcc tcg gag agc ctg cct 624
Leu Ala Glu Ala Trp Asp His Leu Ala Phe Ser Ser Glu Ser Leu Pro
190 195 200 205
gcc tgg gcg ggc ccg ccc gcc acc gca gcc tcc cag ctg ctc tcc gtg 672
Ala Trp Ala Gly Pro Pro Ala Thr Ala Ala Ser Gln Leu Leu Ser Val
210 215 220
tct cca atc tcc 684

Ser Pro Ile Ser

225

<210> 516

<211> 41

<212> PRT

<213> Homo sapiens

<400> 516

Lys Thr Thr Leu Leu Ser Gly Pro Leu Cys Pro Leu Pro Ala Leu Pro

1 5 10 15

Ser Pro Ser Val Asn Val Arg Pro Met Gly Ala Ala Gly Gln Arg Gly

20 25 30

Pro Arg Pro Gly Ala Pro Asn Pro Met

35 40

<210> 517

<211> 11

<212> PRT

<213> Homo sapiens

<400> 517

Pro Gln Ser Ser His Gln Ala Leu Ser Ser Ala

1 5 10

<210> 518

<211> 79

<212> PRT

<213> Homo sapiens

<400> 518

Val Leu Arg Pro Gln Trp Leu Leu Trp Gly Pro Pro Glu Phe Leu Ile

1 5 10 15

Cys Ala Pro Pro Ser Leu Ala Gln Val Lys Val Trp Phe Gln Asn Arg

20 25 30

Arg Thr Lys Tyr Lys Arg Gln Lys Leu Glu Glu Glu Gly Pro Glu Ser

35 40 45

Glu Gln Lys Lys Lys Gly Ser His His Ile Asn Arg Trp Arg Ile Ala

50 55 60

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

65 70 75

<210> 519

<211> 94

<212> PRT

<213> Homo sapiens

<400> 519

Gly Gly Gln Pro Gln Thr His Glu Gly Arg Val Leu Leu Ala Ala Gly

1 5 10 15

Gln Ala Pro Ala Trp Ala Gln Ala Gly Leu Trp Pro Leu Pro Gly Gln

20 25 30

Ala Leu Gly Arg Pro Gly Val Met Ala Pro Gln Gly Leu Lys Pro Gly

35 40 45

Ala Ala Ile Asp Arg Gly Thr Ser Asn Gly Leu Ala Glu Ala Trp Asp

50 55 60

His Leu Ala Phe Ser Ser Glu Ser Leu Pro Ala Trp Ala Gly Pro Pro

65 70 75 80

Ala Thr Ala Ala Ser Gln Leu Leu Ser Val Ser Pro Ile Ser

85 90

<210> 520

<211> 30

<212> DNA

<213> Mus sp.

<400> 520

ccaccattct gcagagccag cagaggcagg 30

<210> 521

<211> 20

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 521

ccauucugca gagccagcag 20

<210> 522

<211> 20

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 522

uuggcauggg ucgcugacgg 20

<210> 523

<211> 30

<212> DNA

<213> Mus sp.

<400> 523

ccagcctccg tcagcgaccc atgccaagac 30

<210> 524

<211> 20

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 524

cgggcuggag cuguucgcgc 20

<210> 525

<211> 30

<212> DNA

<213> Mus sp.

<400> 525

gatgccagcg cgaacagctc cagcccgagt 30

<210> 526

<211> 30

<212> DNA

<213> Mus sp.

<400> 526

agaagagggt gccagcgggt atgaggagtg 30

<210> 527

<211> 20

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 527

agagggugcc agcggguaug 20

<210> 528

<211> 23

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 528

ctgggagagg gagcccctcc agg 23

<210> 529

<211> 23

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 529

aaaggtggga gacacctcct tgg 23

<210> 530

<211> 23

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 530

tccaaccttc aggcaaggtg ggg 23

<210> 531

<211> 23

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 531

aggaagtctg gccgatctgc tgg 23

<210> 532

<211> 23

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 532

ccattctgca gagccagcag agg 23

<210> 533

<211> 23

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 533

ctctgaggcc ctggagatcc tgg 23

<210> 534

<211> 23

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 534

ttggcatggg tcgctgacgg agg 23

<210> 535

<211> 23

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 535

cgggctggag ctgttcgcgc tgg 23

<210> 536

<211> 23

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 536

agagggtgcc agcgggtatg agg 23

<210> 537

<211> 23

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 537

cctccgtcag cgacccatgc caa 23

<210> 538

<211> 24

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 538

cctccgatca gcgacccatg ccaa 24

<210> 539

<211> 24

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 539

cctccgttca gcgacccatg ccaa 24

<210> 540

<211> 24

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 540

cctccgctca gcgacccatg ccaa 24

<210> 541

<211> 22

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 541

cctcctcagc gacccatgcc aa 22

<210> 542

<211> 22

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 542

cctccgcagc gacccatgcc aa 22

<210> 543

<211> 21

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 543

cctctcagcg acccatgcca a 21

<210> 544

<211> 19

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 544

cctccgcgac ccatgccaa 19

<210> 545

<211> 19

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 545

cctcagcgac ccatgccaa 19

<210> 546

<211> 18

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 546

ctcagcgacc catgccaa 18

<210> 547

<211> 23

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 547

ccagcgcgaa cagctccagc ccg 23

<210> 548

<211> 24

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 548

ccagcgccga acagctccag cccg 24

<210> 549

<211> 21

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 549

ccagcgaaca gctccagccc g 21

<210> 550

<211> 21

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 550

ccagcggaca gctccagccc g 21

<210> 551

<211> 19

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 551

ccagaacagc tccagcccg 19

<210> 552

<211> 18

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 552

ccagcgagct ccagcccg 18

<210> 553

<211> 21

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 553

ccagcgcaca gctccagccc g 21

<210> 554

<211> 16

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 554

ccagcgctcc agcccg 16

<210> 555

<211> 24

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 555

agagggtgcc agcgggttat gagg 24

<210> 556

<211> 24

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 556

agagggtgcc agcgggtaat gagg 24

<210> 557

<211> 36

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<220><221> modified_base

<222> (18)..(36)

<223> a, c, t, g, unknown or other

<400> 557

agagggtgcc agcgggtnnn nnnnnnnnnn nnnnnn 36

<210> 558

<211> 19

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 558

agagggtgcc agtatgagg 19

<210> 559

<211> 21

<212> DNA

<213> Artificial Sequence

<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"
<400> 559
agagggtgcc agcgggtgag g 21
<210> 560
<211> 17
<212> DNA
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"
<400> 560
agagggtgcc agcgagg 17
<210> 561
<211> 20
<212> DNA
<213> Artificial Sequence

<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"
<400> 561
agagggtgcc agcggggagg 20
<210> 562
<211> 24
<212> DNA
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"
<400> 562
ccattctgca gagccagagg cagg 24

<210> 563

<211> 21

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 563

ccattctgca gagcccagag g 21

<210> 564

<211> 22

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 564

ccattctgca gagccagaga gg 22

<210> 565

<211> 19

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 565

ccattctgca gagcagagg 19

<210> 566

<211> 22

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 566

ccattctgca gagccccaga gg 22

<210> 567

<211> 25

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 567

ccattctgca gagccaggag gcagg 25

<210> 568

<211> 67

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<220><221> modified_base

<222> (18)..(61)

<223> a, c, t, g, unknown or other

<400> 568

ccattctgca gagccagnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 60

ncagagg 67

<210> 569

<211> 55

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<220><221> modified_base

<222> (18)..(52)

<223> a, c, t, g, unknown or other

<400> 569

ccattctgca gagccagnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnagg 55

<210> 570

<211> 23

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 570

ccattctgca gagccagaag agg 23

<210> 571

<211> 25

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 571

ccattctgca gagccagcac agagg 25

<210> 572

<211> 22

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 572

ccattctgca gagccacaga gg 22

<210> 573

<211> 24

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 573

ccattctgca gagccagtca gagg 24

<210> 574

<211> 6

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

peptide"

<220><221> MOD_RES

<222> (2)..(2)

<223> Ahx

<400> 574

Arg Xaa Arg Arg Arg Arg

1 5

<210> 575

<211> 6

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

6xHis tag"

<400> 575

His His His His His His

1 5