(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2024-0104191
(43) 공개일자 2024년07월04일

(51) 국제특허분류(Int. Cl.)
C12N 15/63 (2006.01) *C12N 15/10* (2017.01)
C12N 15/113 (2010.01) *C12N 15/90* (2006.01)
C12N 9/22 (2006.01)
(52) CPC특허분류
C12N 15/63 (2013.01)
C12N 15/102 (2013.01)
(21) 출원번호 10-2024-7020768(분할)
(22) 출원일자(국제) 2014년11월19일
심사청구일자 없음
(62) 원출원 특허 10-2022-7022535
원출원일자(국제) 2014년11월19일
심사청구일자 2022년06월30일
(85) 번역문제출일자 2024년06월21일
(86) 국제출원번호 PCT/US2014/066324
(87) 국제공개번호 WO 2015/077290
국제공개일자 2015년05월28일
(30) 우선권주장
61/906,188 2013년11월19일 미국(US)

(71) 출원인
프레지던트 앤드 펠로우즈 오브 하바드 칼리지
미합중국, 메사추세츠 02138, 캠브리지, 퀸시스트리트17
(72) 발명자
바이른, 수산 엠.
미국 02446 매사추세츠주 브루클라인 세월 애비뉴 9
처치, 조지 엠.
미국 02446 매사추세츠주 브루클라인 켄트 스트리트 218
(74) 대리인
양영준, 김영

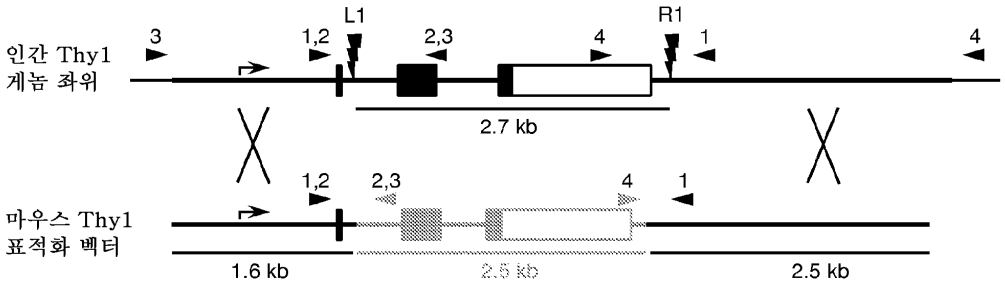
전체 청구항 수 : 총 26 항

(54) 발명의 명칭 **대형 유전자 절제 및 삽입**

(57) 요 약

DNA 결합 단백질 뉴클레아제를 사용하여 동시에, 표적 핵산으로부터 대형 핵산 서열을 절제하고, 표적 핵산 서열로 대형 외래 핵산 서열을 삽입하는 방법이 기술된다.

대 표 도 - 도1a



(52) CPC특허분류
C12N 15/113 (2013.01)
C12N 15/907 (2013.01)
C12N 9/22 (2013.01)
C12N 2310/20 (2017.05)

명 세 서

청구범위

청구항 1

유전자의 일 측에서의 표적 핵산 서열에 대하여 상보적인 가이드 RNA 서열 또는 가이드 RNA 서열을 코딩하는 제1 외래 핵산을 진핵 세포 내로 도입하는 단계,

유전자의 일 측에서의 절단 부위를 생성하기 위한 Cas9 단백질 또는 Cas9 단백질을 코딩하는 제2 외래 핵산을 진핵 세포 내로 도입하는 단계, 및

유전자의 어느 한 측에 대하여 상보적인 상동성 아암 서열에 의해 측접되는, 유전자를 대체하기 위한 외인성 핵산 서열을 포함하는 표적화 벡터를 진핵 세포 내로 도입하는 단계이며, 여기서 절단 부위와 먼 측에서의 상동성 아암 서열은 절단 부위와 가까운 측에서의 상동성 아암 서열보다 적어도 24 염기 더 길고, 각 상동성 아암 서열의 길이가 100 내지 2000 염기인 것인, 단계를 포함하고,

상기 제1 외래 핵산 및 제2 외래 핵산이 진핵 세포 내로 도입되었을 때 가이드 RNA 및 Cas9 단백질이 발현되고,

상기 가이드 RNA 및 Cas9 단백질이 상기 표적 핵산 서열에 공동-국재화되며, 상기 Cas9 단백질이 절단 부위를 생성하고, 상기 외인성 핵산 서열이 절단 부위에서 삽입되는 것인,

단리된 진핵 세포에서 유전자를 대체하는 방법이며, 상기 단리된 진핵 세포는 인간 배아 또는 인간 배아 줄기 세포가 아닌 방법.

청구항 2

제1항에 있어서, 삽입되는 외인성 핵산 서열이 유전자인 방법.

청구항 3

제1항에 있어서, 외인성 핵산 서열의 길이가 1,000개 염기쌍 내지 100,000개 염기쌍 사이인 방법.

청구항 4

제1항에 있어서, 단리된 진핵 세포가 인간 세포인 방법.

청구항 5

제1항에 있어서, 단리된 진핵 세포가 줄기 세포인 방법.

청구항 6

제1항에 있어서, 단리된 진핵 세포가 인간 줄기 세포인 방법.

청구항 7

제1항에 있어서, 단리된 진핵 세포가 유도 만능성 줄기 세포인 방법.

청구항 8

제1항에 있어서, 단리된 진핵 세포가 인간 유도 만능성 줄기 세포인 방법.

청구항 9

제1항에 있어서, 가이드 RNA가 tracrRNA-crRNA 융합체인 방법.

청구항 10

제1항에 있어서, 유전자를 가지는 핵산이 게놈 DNA, 미토콘드리아 DNA, 바이러스 DNA 또는 외인성 DNA인 방법.

청구항 11

제1항에 있어서, 외인성 핵산 서열이 상동성 재조합에 의해 표적 핵산 서열에 삽입되는 것인 방법.

청구항 12

제1항에 있어서, 외인성 핵산 서열의 길이가 10,000개 염기쌍 내지 100,000개 염기쌍 사이인 방법.

청구항 13

제1항에 있어서, 외인성 핵산 서열의 길이가 1,000개 염기쌍 내지 10,000개 염기쌍 사이인 방법.

청구항 14

제1항에 있어서, 총 상동성 아암 서열의 길이가 1.5kb 내지 2kb인 방법.

청구항 15

유전자의 일 측에서의 표적 핵산 서열에 대하여 상보적인 가이드 RNA 서열 또는 가이드 RNA 서열을 코딩하는 제1 외래 핵산,

유전자의 일 측에서의 절단 부위를 생성하기 위한 Cas9 단백질 또는 Cas9 단백질을 코딩하는 제2 외래 핵산, 및

유전자의 어느 한 측에 대하여 상보적인 상동성 아암 서열에 의해 측접되는, 유전자를 대체하기 위한 외인성 핵산 서열을 포함하는 표적화 벡터를 포함하고, 여기서 절단 부위와 먼 측에서의 상동성 아암 서열은 절단 부위와 가까운 측에서의 상동성 아암 서열보다 적어도 24 염기 더 길고, 각 상동성 아암 서열의 길이가 100 내지 2000 염기인 것인,

단리된 진핵 세포에서 유전자를 대체하기 위한 시스템이며, 상기 단리된 진핵 세포는 인간 배아 또는 인간 배아 줄기 세포가 아닌 시스템.

청구항 16

제15항에 있어서, 외인성 핵산 서열이 유전자인 시스템.

청구항 17

제15항에 있어서, 외인성 핵산 서열의 길이가 1,000개 염기쌍 내지 100,000개 염기쌍 사이인 시스템.

청구항 18

제15항에 있어서, 가이드 RNA가 tracrRNA-crRNA 융합체인 시스템.

청구항 19

제15항에 있어서, 외인성 핵산 서열의 길이가 10,000개 염기쌍 내지 100,000개 염기쌍 사이인 시스템.

청구항 20

제15항에 있어서, 외인성 핵산 서열의 길이가 1,000개 염기쌍 내지 10,000개 염기쌍 사이인 시스템.

청구항 21

제15항에 있어서, 총 상동성 아암 서열의 길이가 1.5kb 내지 2kb인 시스템.

청구항 22

제15항의 시스템을 포함하는, 인간 배아 또는 인간 배아 줄기 세포가 아닌 단리된 세포.

청구항 23

제22항에 있어서, 진핵 세포인 단리된 세포.

청구항 24

제22항에 있어서, 효모 세포, 식물 세포 또는 동물 세포인 단리된 세포.

청구항 25

제22항에 있어서, 인간 세포인 단리된 세포.

청구항 26

제22항에 있어서, 인간 유도 만능성 줄기 세포인 단리된 세포.

발명의 설명

기 술 분 야

[0001] **[관련 출원 데이터]**

[0002] 본 출원은 2013년 11월 19일자 U.S. 특허 가출원 제61/906,188호의 우선권을 주장하며, 이는 모든 목적에 있어서 그 전체가 본원에 참조로 포함된다.

[0003] **[정부 권리에 대한 진술]**

[0004] 본 발명은 국립 보건원(National Institutes of Health)이 수여하는 교부금 제P50 HG005550호 하의 정부 후원으로 이루어졌다. 정부는 본 발명에 있어서 소정의 권리를 가진다.

배 경 기 술

[0005] 세균 및 고세균 CRISPR-Cas 시스템은 Cas 단백질과의 복합체로 짧은 가이드(short guide) RNA에 의존하여 침입 외래 핵산 내에 존재하는 상보성 서열의 분해를 유도한다. 문헌 [Deltcheva, E. et al. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* 471, 602-607 (2011)]; [Gasiunas, G., Barrangou, R., Horvath, P. & Siksnys, V. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proceedings of the National Academy of Sciences of the United States of America* 109, E2579-2586 (2012)]; [Jinek, M. et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Jinek, M. *et al. Science* 337, 816-821 (2012)]; [Sapranauskas, R. et al. The Streptococcus thermophilus CRISPR/Cas system provides immunity in Escherichia coli. *Nucleic acids research* 39, 9275-9282 (2011)]; 및 [Bhaya, D., Davison, M. & Barrangou, R. CRISPR-Cas systems in bacteria and archaea: versatile small RNAs for adaptive defense and regulation. *Annual review of genetics* 45, 273-297 (2011)]을 참조하라. *에스. 피오게네스(S. pyogenes)* 유형 II CRISPR 시스템의 최근의 시험관내 재구축은 정상적으로 트랜스-코딩된 tracrRNA ("트랜스-활성화 CRISPR RNA")에 융합된 crRNA ("CRISPR RNA")이면 crRNA와 매치하는 표적 DNA 서열을 서열-특이적으로 절단하도록 Cas9 단백질에게 지시하기에 충분하다는 것을 입증하였다. 표적 부위에 대하여 상동성인 gRNA를 발현시키는 것은 Cas9 동원 및 표적 DNA의 분해를 초래한다. 문헌 [H. Deveau *et al.*, Phage response to CRISPR-encoded resistance in Streptococcus thermophilus. *Journal of Bacteriology* 190, 1390 (Feb, 2008)]을 참조하라.

발명의 내용

과제의 해결 수단

[0006] **[발명의 개요]**

[0007] 본 개시내용의 측면은 표적 핵산에의 대형 유전자 삽입의 효율을 향상시키는 것에 관한 것이다. 본 개시내용의 측면은 세포 내 표적 핵산으로부터 대형 핵산 서열, 예컨대 대형 유전자의 절제에 관한 것이다. 본 개시내용의 측면은 세포 내 표적 핵산으로의 대형 핵산 서열, 예컨대 대형 유전자의 삽입에 관한 것이다. 본 개시내용의 측면은 세포 내 표적 핵산으로부터의 대형 핵산 서열, 예컨대 대형 유전자의 절제, 및 세포 내 표적 핵산으로의 대형 핵산 서열, 예컨대 대형 유전자의 삽입에 관한 것이다.

[0008] 본 개시내용의 측면은 표적 핵산 내 대형 핵산 대체물을 위한 표적화 벡터의 설계에 관한 것이다.

[0009] 소정 측면에서는, 표적 핵산 서열에 이중 가닥 절단을 생성시키는 데에 서열 특이적 뉴클레아제와 같은 DNA 결합 단백질이 사용된다. 하나 이상 또는 다수의 이중 가닥 절단들이 표적 핵산 서열에서 이루어질 수 있다. 일 측면에 따르면, 표적 핵산 서열로부터 제1 핵산 서열이 제거된 후, 외인성 핵산 서열이 표적 핵산 서열의 절단된 부위들 또는 절단된 말단들 사이에서 표적 핵산에 삽입된다. 소정 측면에 따르면, 각 상동성 아암(homology arm)에서의 이중 가닥 절단은 예컨대 상동성 재조합에 의해 핵산 서열 삽입 또는 대체의 효율을 증가시키거나 향상시킨다. 소정 측면에 있어서, 다수의 이중 가닥 절단 또는 절단 부위(cut site)들은 표적화 벡터로부터의 핵산 서열의 도입 효율을 향상시킨다.

[0010] 본 개시내용의 소정 측면은 인간 유도 만능성 줄기 세포 (iPSC)를 포함한 포유동물 세포와 같은 세포에서의 다수-킬로베이스인 내인성 유전자의 동형접합 녹-인(knock-in) 표적화 대체 및 절제 방법에 관한 것이다. 소정 측면에 따르면, 상기 방법은 선택 마커 없이 실행된다.

[0011] 소정 측면에서는, 뉴클레아제 활성을 가지는 DNA 결합 단백질, 예컨대 뉴클레아제 활성을 가지는 RNA 가이드된 DNA 결합 단백질을 사용한 표적 핵산 서열에의 대형 유전자의 삽입 방법이 제공된다. 소정 측면에 따르면, DNA (데옥시리보핵산)에 상보적인 1종 이상의 RNA (리보핵산)를 코딩하는 제1 외래 핵산이 세포에 도입되는데, 여기서 상기 DNA는 표적 핵산을 포함한다. DNA에 결합하며 1종 이상의 RNA에 의해 가이드되는 뉴클레아제 활성을 가지는 RNA 가이드된 DNA 결합 단백질을 코딩하는 제2의 외래 핵산이 세포에 도입된다. 상기 1종 이상의 RNA 및 RNA 가이드된 DNA 결합 단백질은 발현되는데, 이 경우 1종 이상의 RNA 및 RNA 가이드된 DNA 결합 단백질은 DNA에 공동-국재화되고, DNA 결합 단백질이 표적 핵산을 절단함으로써 관심 제1 핵산 서열을 제거한다. 관심 외인성 핵산 서열은 표적 핵산 서열의 절단 부위들 사이에 삽입됨으로써 관심 제1 핵산 서열의 제거를 초래한다. 소정 측면에 따르면, 다수의 가이드 RNA가 사용될 수도 있다.

[0012] 본 개시내용의 범주에 속하는 대형 핵산 서열 (제거될 관심 제1 핵산 서열 또는 삽입될 외인성 핵산 서열일 수 있음)은 100개 염기쌍 초과 내지 약 100,000개 염기쌍 사이, 100개 염기쌍 초과 내지 약 10,000개 염기쌍 사이, 약 200개 염기쌍 내지 약 100,000개 염기쌍 사이, 약 300개 염기쌍 내지 약 100,000개 염기쌍 사이, 약 400개 염기쌍 내지 약 100,000개 염기쌍 사이, 약 500개 염기쌍 내지 약 100,000개 염기쌍 사이, 약 600개 염기쌍 내지 약 100,000개 염기쌍 사이, 약 700개 염기쌍 내지 약 100,000개 염기쌍 사이, 약 800개 염기쌍 내지 약 100,000개 염기쌍 사이, 약 900개 염기쌍 내지 약 100,000개 염기쌍 사이, 약 1000개 염기쌍 내지 약 100,000개 염기쌍 사이, 약 2000개 염기쌍 내지 약 100,000개 염기쌍 사이, 약 3000개 염기쌍 내지 약 100,000개 염기쌍 사이, 약 4000개 염기쌍 내지 약 100,000개 염기쌍 사이, 약 5000개 염기쌍 내지 약 100,000개 염기쌍 사이, 약 6000개 염기쌍 내지 약 100,000개 염기쌍 사이, 약 7000개 염기쌍 내지 약 100,000개 염기쌍 사이, 약 8000개 염기쌍 내지 약 100,000개 염기쌍 사이, 약 9000개 염기쌍 내지 약 100,000개 염기쌍 사이, 약 10,000개 염기쌍 내지 약 100,000개 염기쌍 사이, 약 20,000개 염기쌍 내지 약 100,000개 염기쌍 사이, 약 30,000개 염기쌍 내지 약 100,000개 염기쌍 사이, 약 40,000개 염기쌍 내지 약 100,000개 염기쌍 사이, 약 50,000개 염기쌍 내지 약 100,000개 염기쌍 사이, 약 60,000개 염기쌍 내지 약 100,000개 염기쌍 사이, 약 70,000개 염기쌍 내지 약 100,000개 염기쌍 사이, 약 80,000개 염기쌍 내지 약 100,000개 염기쌍 사이, 약 90,000개 염기쌍 내지 약 100,000개 염기쌍 사이, 약 500개 염기쌍 내지 약 10,000개 염기쌍 사이, 약 1000개 염기쌍 내지 약 10,000개 염기쌍 사이, 약 2000개 염기쌍 내지 약 10,000개 염기쌍 사이 또는 약 1000개 염기쌍 내지 약 5,000개 염기쌍 사이의 길이를 가지는 핵산 서열이다. 대형 핵산 서열은 본원에서 "킬로베이스" 또는 "다수-킬로베이스"의 핵산 서열로 지칭될 수 있는데, 다시 말하자면 1000개 염기쌍을 초과하는 것이다. 소정 측면에 따르면, 대형 핵산은 1000개 염기쌍 초과이다. 소정 측면에 따르면, 삽입될 핵산 서열은 그것이 삽입될 세포의 게놈에 대하여 이종유래(heterologous)이다. 소정 측면에 따르면, 표적 핵산 서열에 삽입되는 외인성 핵산은 그것이 대체하는 관심 제1 핵산 서열과 상이할 수 있거나, 또는 세포의 표적 핵산 또는 게놈 중의 어떠한 핵산 서열과도 상이할 수 있는 외래 핵산 서열이다. 일 측면에 따르면, 관심 제1 핵산 서열의 제거 및 외인성 핵산의 삽입은 동시에, 또는 실질적으로 동시에 이루어진다. 이는 결실이 이루어진 후 별도의 삽입이 이어지는 방법과 구별되어야 한다.

[0013] 소정 측면에서는, 대체물의 일 측에 위치하는 단일 절단 부위를 사용한 유전자 대체를 조작하는 데에 사용되는 외래 서열을 포함하는 표적화 벡터가 제공된다. 상기 외래 서열에는 대체물 부근의 게놈과 동일한 서열이 측접(flank)한다. 이후, 일 측에서의 단지 1회의 절단으로 유전자 대체가 수행될 수 있다. 타 말단은 게놈과 표적화 벡터의 외래 서열 부근에 위치하는 그의 상보적 서열 (상동성 아암) 사이의 자연적인 재조합에 의해 제거되

게 된다. 그러나, 본 개시내용의 측면은 대체되는 영역의 양 측을 절단하는 것도 포함한다.

[0014] 일 측면에 있어서, 세포는 진핵 세포이다. 일 측면에 있어서, 세포는 효모 세포, 식물 세포 또는 동물 세포이다. 일 측면에 있어서, 세포는 포유동물 세포이다.

[0015] 일 측면에 있어서, RNA는 약 10 내지 약 500개 뉴클레오티드 사이이다. 일 측면에 있어서, RNA는 약 20 내지 약 100개 뉴클레오티드 사이이다.

[0016] 일 측면에 있어서, 1종 이상의 RNA는 가이드 RNA이다. 일 측면에 있어서, 1종 이상의 RNA는 crRNA이다. 일 측면에 있어서, 1종 이상의 RNA는 tracrRNA이다. 일 측면에 있어서, 1종 이상의 RNA는 tracrRNA-crRNA 융합체이다.

[0017] 일 측면에 있어서, DNA는 게놈 DNA, 미토콘드리아 DNA, 바이러스 DNA 또는 외인성 DNA이다.

[0018] 일 측면에 있어서, RNA 가이드된 DNA 결합 단백질은 DNA에 결합하며 1종 이상의 RNA에 의해 가이드되는, 유형 II CRISPR 시스템의 것이다. 일 측면에 있어서, RNA 가이드된 DNA 결합 단백질은 DNA에 결합하며 1종 이상의 RNA에 의해 가이드되는 Cas9 단백질이다.

[0019] 본 발명 소정 실시양태의 다른 특징 및 장점들은 실시양태에 대한 하기의 설명 및 그의 도면, 그리고 청구범위로부터 더 완전히 드러나게 될 것이다.

도면의 간단한 설명

[0020] 본 특허 또는 출원 파일은 컬러로 작성된 하나 이상의 도면을 포함하고 있다. 컬러 도면(들)을 포함하는 이와 같은 특허 또는 특허 출원 공개의 사본은 요청 및 필요 비용의 납부시 특허청에 의해 제공될 것이다. 본 발명의 전기 및 기타 특징 및 장점들은 첨부 도면과 함께 선택된 예시적인 실시양태들의 하기 상세한 설명으로부터 더 완전하게 이해될 것인 바, 도면 중:

도 1(a)-(d). 1개 또는 2개의 CRISPR sgRNA를 사용한 동형접합 표적화 유전자 대체. 도 1(a). 인트론 1 내 (L1) 또는 폴리아데닐화 부위 후 (R1)의 2개의 Crispr sgRNA 표적 hThy1. mThy1 표적화 벡터 플라스미드는 sgRNA 부위들 외부에 hThy1 상동성 아암들이 측접되어 있는 mThy1 엑손 2 및 3 (오렌지색)을 포함한다 - hThy1 프로모터 및 엑손 1 (리더 서열을 코딩함)은 유지되나, sgRNA 부위가 붕괴됨. 소형 삼각형은 4종의 제노타이핑 PCR 반응을 위한 프라이머 부위들을 표시한다. 도 1(b). PGP1 iPSC를 mThy1 표적화 벡터, Cas9 뉴클레아제 및 L1, R1, 양자를 코딩하거나 sgRNA를 코딩하지 않는 플라스미드로 뉴클레오펙션시켰다. 5일 후, 유동 세포측정법에 의해 세포를 분석하였다. 수득된 mThy1 발현 및/또는 상실된 hThy1 발현을 가지는 세포의 백분율을 표시하였다. 도 1(c). 각 사분면으로부터 단일 iPSC를 FACS 분류하고, 개별 웰에서 배양한 후, 4종의 PCR 반응을 사용하여 제노타이핑하였다. PCR 생성물의 크기 및 생거 서열분석을 바탕으로 대립유전자들을 확인하였다: 자연 인간 (+); 재조합 마우스 (m); 2개 sgRNA 부위들 사이에서 절제 (Δ); 및 2개 sgRNA 부위들 사이에서 역전 (i). +/+ 야생형, m/+ 이형접합, m/m 동형접합, m/Δ 이형접합, Δ/Δ 동형접합 및 i/Δ 이형접합 집락(colony)으로부터의 대표적인 겔을 나타내었다. 도 1(d). FACS-분류된 iPS 집락간 유전형의 빈도. 결과는 3회 개별 실험의 대표이다.

도 2의 (A)-(E). CRISPR-생성 동형접합 및 이형접합 결실의 빈도. 도 2(A). 인간 Thy1 유전자를 표적으로 하는 Crispr sgRNA를 생성시켰다: 2개는 인트론 1 내 (좌측: L1 및 L2), 10개는 hThy1 후의 다양한 거리 (우측: R1 내지 R10). B+C) sgRNA의 좌측 하나와 우측 하나의 쌍을 도 2(B) PGP1 iPSC 또는 도 2(C) $Thy1^{m/+}$ PGP1 iPSC 클론 중 어느 하나에 뉴클레오펙션시켰다. 음성 대조군으로는, 좌측 sgRNA만을 뉴클레오펙션시켰다 (우측 컬럼). 5일 후, 도 2(B)의 인간 Thy1 대립유전자들 양자의 동형접합 결실 또는 도 2(C)의 나머지 인간 Thy1 대립유전자의 이형접합 결실 중 어느 하나, 그리고 마우스 Thy1 대립유전자의 보유에 대하여 유동 세포측정법에 의해 세포를 분석하였다. sgRNA 부위들 사이의 거리 (Thy1 Δ) 및 $hThy1^-$ 세포의 빈도를 표시하였다. 도 2(D)-(E). 각 sgRNA 쌍으로부터의 $hThy1^-$ 세포의 퍼센트에서 좌측 sgRNA-단독 대조군으로부터의 퍼센트를 뺀 것을 Thy1 결실의 크기에 대하여 플로팅하였다. L1 또는 L2를 포함하는 sgRNA 쌍은 각각 흑색 또는 적색으로 나타내었다. 오차 막대는 2회 개별 실험의 평균 ± s.e.m.을 나타낸다.

도 3의 (A)-(C). 절제 및 역전 접합의 서열. FACS-분류된 hThy1-mThy1- iPSC 클론으로부터 정제된 게놈 DNA에서 제노타이핑 PCR 반응 #1을 수행하고 (도 1에서 기술된 바와 같음), 동일한 PCR 프라이머들을 사용하여 생거 서열분석으로 분석하였다. 생성되는 생거 서열분석 추적에서의 이중 피크는 각 클론의 이중대립유전자 서열을

밝히기 위하여 디컨볼루션하였다. 도 3(A). 인간 Thy1 좌위의 자연 서열 (서열(SEQ ID NO) 108). Crispr sgRNA 표적화 부위 L1 (보라색) 및 R1 (적색)은 인간 Thy1 좌위에서 2.7 kb 이격되어 있다. 예측 뉴클레아제 절단 부위는 PAM의 6 bp 상류에 표시되어 있다 (밑줄침). 도 3(B)-(C). 도 3(B)의 4종의 Δ/Δ 이중-절제 (서열 109-112) 및 도 3(C)의 도 1d에 기술되어 있는 5종의 i/Δ 역전-및-절제 (서열 113-120) hThy1-mThy1- 집락들의 이중대립유전자 서열. 2종의 별도 클론에서 발견된 서열들은 ×2로 표시하였다. 결과는 2회 개별 실험의 대표이다.

도 4의 (A)-(B). 마우스 CD147 유전자의 인간 CD147 게놈 좌위에의 표적화된 대체. 도 4(A). 인트론 1 내 (L147, 좌측) 및 폴리아데닐화 부위 후 (R147, 우측)의 인간 CD147 유전자를 표적으로 하는 2개의 Crispr sgRNA를 9.8 kb 이격시켜 설계하였다. 마우스 CD147 표적화 벡터 플라스미드는 sgRNA 부위 외부에 인간 CD147 서열과 매치하는 상동성 아암이 측접되는 마우스 CD147 엑손 2-7 (갈색)을 포괄하는 5.8 kb 서열로 구성되었다. 표적화 구축물에서, 인간 CD147 프로모터 및 엑손 1 (리더 서열을 코딩함)은 유지되었으나, sgRNA 부위들은 붕괴되었다. 도 4(B). PGP1 iPSC를 마우스 CD147 표적화 벡터, Cas9 뉴클레아제 및 어느 하나, 양자를 코딩하거나 sgRNA를 코딩하지 않는 플라스미드로 뉴클레오펙션시켰다. 9일 후, 유동 세포측정법에 의해 세포를 분석하였다. 수득된 마우스 CD147의 발현 및/또는 상실된 인간 CD147의 발현을 가지는 세포의 백분율을 표시하였다. 결과는 3회 개별 실험의 대표이다.

도 5의 (A)-(B). Crispr 또는 TALEN을 가지는 유전자 배향 중 어느 하나의 표적화된 유전자 대체. 도 5(A). 엑손 2 내 (L3, 좌측) 및 폴리아데닐화 부위 후 (R1, 우측)의 인간 Thy1 유전자를 표적으로 하는 2개의 Crispr sgRNA를 2.2 kb 이격시켜 설계하였다. 동일한 L3 및 R1 부위를 표적으로 하는 2개의 TALEN 쌍도 설계하였다. "녹-인" 표적화 벡터용으로, 구성적 pGK 프로모터 (적색) 하에 형광 mCherry 유전자를 포괄하는 1.9 kb 서열에, sgRNA 부위 외부에서 인간 Thy1 서열과 매치하는 상동성 아암을 측접시켰다. pGK-mCherry 삽입물을 Thy1 유전자 대비 순방향 또는 역방향 배향 중 어느 하나로 클로닝하였다. 양 mCherry 표적화 구축물에서 sgRNA 및 TALEN 부위들을 붕괴시켰다. 도 5(B). 하기를 코딩하는 플라스미드와 함께 센스 또는 안티센스 mCherry 표적화 벡터를 사용하여 PGP1 iPSC를 뉴클레오펙션시켰다: sgRNA 어느 하나 또는 양자를 포함하는 Cas9 뉴클레아제; TALEN 쌍 어느 하나 또는 양자; 또는 빈 벡터 플라스미드. 10일 후, 유동 세포측정법에 의해 세포를 분석하였다. 수득된 mCherry의 발현 및/또는 상실된 인간 Thy1의 발현을 가지는 세포의 백분율을 표시하였다. 결과는 2회 개별 실험의 대표이다.

도 6의 (A)-(B). 원형 또는 선형 표적화 벡터를 사용한 표적화된 유전자 대체. 도 6(A). 상동성 아암의 말단 (소형 삼각형)에서 PCR 프라이머들을 사용하여 도 1의 원형 플라스미드 마우스 Thy1 표적화 벡터를 증폭함으로써, 마우스 Thy1 표적화 벡터의 선형 형태를 생성시켰다. 도 6(B). Cas9 뉴클레아제, 및 L1, R1, 양자를 코딩하거나 sgRNA를 코딩하지 않는 플라스미드와 함께 원형 플라스미드 또는 선형 PCR 생성물 마우스 Thy1 표적화 벡터를 사용하여 PGP1 iPSC를 뉴클레오펙션시켰다. 6일 후, 유동 세포측정법에 의해 세포를 분석하였다. 수득된 마우스 Thy1의 발현 및/또는 상실된 인간 Thy1의 발현을 가지는 세포의 빈도를 표시하였다. 결과는 2회 개별 실험의 대표이다.

도 7(a)-(b). 재조합 효율에 대한 상동성 아암 길이의 효과. 다양한 길이의 상동성 아암을 가지나 여전히 마우스 Thy1 엑손 2 및 3을 포괄하는 2.5 kb 서열을 포함하는 마우스 Thy1 표적화 벡터 플라스미드의 버전들 (도 1a)을 구축하였다. 각 벡터의 상류 및 하류 상동성 아암의 길이를 표시하였다. Cas9 뉴클레아제, 및 L1, R1, 양자를 코딩하거나 sgRNA를 코딩하지 않는 플라스미드와 함께 도 7(a)의 PGP1 또는 도 7(b)의 PGP4 iPSC에 각 마우스 Thy1 표적화 벡터를 뉴클레오펙션시켰다. 10일 후, 마우스 및 인간 Thy1 유전자의 발현에 대하여 유동 세포측정법에 의해 세포를 분석하였다. 각 형광 사분면에서의 세포의 빈도를 표시하였다. 결과는 2회 개별 실험의 대표이다.

도 8(a)-(c). 각 절단 부위 어느 한 측에 상동성을 가지는 표적화된 유전자 대체. 도 8(a). 도 1로부터의 mThy1 표적화 벡터 (외부)를 인간 Thy1 상동성 아암이 L1 및 R1 sgRNA 부위 내부로 연장되도록 변형시켰다. 마우스 Thy1 엑손 2 및 3 (오렌지색)이 이와 같은 표적화 벡터에 완전히 보유되어 있기는 하지만, 마우스 Thy1 인트론 1 350 bp 및 폴리아데닐화 부위 후 마우스 Thy1 서열 150 bp를 상응하는 인간 서열로 대체하였다. 생성되는 표적화 벡터는 무손상 L1 및 R1 sgRNA 부위를 포함한다 (무손상). 다음에, 표적 벡터 내 각 sgRNA 부위로부터 단일 염기쌍을 결실시킴으로써, 유사한 상동성 아암을 가지나 붕괴된 sgRNA 부위를 가지는 대안적인 버전 (붕괴)을 개발하였다. 도 8(b)-(c). Cas9 뉴클레아제, 및 L1, R1, 양자를 코딩하거나 sgRNA를 코딩하지 않는 플라스미드와 함께 마우스 Thy1 표적화 벡터들 (외부, 무손상 또는 붕괴) 중 하나를 사용하여 PGP1 또는 PGP4 iPSC를 뉴클레오펙션시켰다. 도 8(b). 뉴클레오펙션 2일 후, 생존력 염료 토프로3(ToPro3)을 사용하여 각 조건

의 세포 샘플을 염색하고, 일정한 유량 및 수집 시간을 사용하여 유동 세포측정법에 의해 분석하였다. 생존가능 세포 계수를 sgRNA를 가지지 않는 외부 마우스 Thy1 표적화 벡터의 것 (100)에 정규화하였다. 도 8(c). 뉴클레오펙션 5일 후, 유동 세포측정법에 의해 세포를 분석하였다. 수득된 마우스 Thy1의 발현 및/또는 상실된 인간 Thy1의 발현을 가지는 세포의 백분율을 표시하였다. 결과는 3회 (PGP1) 및 2회 (PGP4) 개별 실험의 대표이다.

발명을 실시하기 위한 구체적인 내용

[0021] 본 개시내용의 실시양태는 표적 핵산 서열로부터 제1 핵산 서열을 제거함으로써 그 안에의 외인성 핵산 서열의 삽입을 가능케 하는 뉴클레아제 활성을 가지는 DNA 결합 단백질의 사용을 바탕으로 한다. 이와 같은 DNA 결합 단백질은 다양한 목적으로 DNA에 결합하는 것으로 통상의 기술자에게 잘 알려져 있다. 이와 같은 DNA 결합 단백질은 자연 발생일 수 있다. 본 개시내용의 영역 내에 포함되는 DNA 결합 단백질에는 본원에서 가이드 RNA로 지칭되는 RNA에 의해 가이드될 수 있는 것들이 포함된다. 이와 같은 측면에 따르면, 가이드 RNA와 RNA 가이드된 DNA 결합 단백질은 DNA에서 공동-국재화 복합체를 형성한다. 뉴클레아제 활성을 가지는 이와 같은 DNA 결합 단백질에 대해서는 통상의 기술자에게 알려져 있는데, 예를 들면 유형 II CRISPR 시스템에 존재하는 Cas9 단백질과 같은 뉴클레아제 활성을 가지는 자연 발생 DNA 결합 단백질이 포함된다. 이와 같은 Cas9 단백질 및 유형 II CRISPR 시스템에 대해서는 관련 기술분야에 충분히 입증되어 있다. 그 전체가 참조로 포함되는 모든 보충 정보를 포함한 문헌 [Makarova et al., *Nature Reviews, Microbiology*, Vol. 9, June 2011, pp. 467-477]을 참조하라.

[0022] 뉴클레아제 활성을 가지는 대표적인 DNA 결합 단백질은 이중 가닥 DNA에 틈을 형성하거나(nick) 그것을 절단하는 기능을 한다. 이와 같은 뉴클레아제 활성은 뉴클레아제 활성을 나타내는 하나 이상의 폴리펩티드 서열을 가지는 DNA 결합 단백질로부터 기인할 수 있다. 이와 같은 대표적인 DNA 결합 단백질은 각 도메인이 이중 가닥 DNA를 절단하거나 그의 특정 가닥에 틈새를 형성하는(nicking) 작용을 하는 2개의 별도 뉴클레아제 도메인을 가질 수 있다. 통상의 기술자에게 알려져 있는 뉴클레아제 활성을 가지는 대표적인 폴리펩티드 서열에는 McrA-HNH 뉴클레아제 관련 도메인 및 RuvC-유사 뉴클레아제 도메인이 포함된다. 이에 따라, 대표적인 DNA 결합 단백질은 자연상에서 McrA-HNH 뉴클레아제 관련 도메인 및 RuvC-유사 뉴클레아제 도메인 중 하나 이상을 포함하는 것들이다.

[0023] 일 측면에 있어서, 2개 이상의 뉴클레아제 도메인을 가지는 DNA 결합 단백질은 뉴클레아제 도메인들 중 하나를 제외한 모두를 불활성화하도록 변형 또는 변경될 수 있다. DNA 결합 단백질이 이중 가닥 DNA의 하나의 가닥만을 절단하거나 거기에 틈새를 형성한다는 점에서, 그와 같은 변형되거나 변경된 DNA 결합 단백질은 DNA 결합 단백질 니카제로 지칭된다. RNA에 의해 DNA로 가이드되는 경우, 상기 DNA 결합 단백질 니카제는 RNA 가이드된 DNA 결합 단백질 니카제로 지칭된다.

[0024] 대표적인 DNA 결합 단백질은 유형 II CRISPR 시스템의 RNA 가이드된 DNA 결합 단백질이다. 대표적인 DNA 결합 단백질은 Cas9 단백질이다.

[0025] *에스. 피오게네스*에서, Cas9는 프로토스페이서(protospacer)-인접 모티프 (PAM)의 3 bp 상류에서 단백질의 하기 2개 촉매촉진 도메인에 의해 매개되는 과정을 통하여 둔단 이중-가닥 절단을 생성시킨다: DNA의 상보성 가닥을 절단하는 HNH 도메인 및 비-상보성 가닥을 절단하는 RuvC-유사 도메인. 그 전체가 참조로 포함되는 문헌 [Jinke et al., *Science* 337, 816-821 (2012)]을 참조하라. 문헌 [Makarova et al., Nature Reviews, Microbiology, Vol. 9, June 2011, pp. 467-477]의 보충 정보에 확인되어 있는 바와 같이, Cas9 단백질은 하기를 포함한 많은 유형 II CRISPR 시스템에 존재하는 것으로 알려져 있다: 메타노코쿠스 마리팔루디스(Methanococcus maripaludis) C7; 코리네박테리움 디프테리아에(Corynebacterium diphtheriae); 코리네박테리움 에피시엔스(Corynebacterium efficiens) YS-314; 코리네박테리움 글루타미쿰(Corynebacterium glutamicum) ATCC 13032 Kitasato; 코리네박테리움 글루타미쿰 ATCC 13032 Bielefeld; 코리네박테리움 글루타미쿰 R; 코리네박테리움 크로펜스테드티이(Corynebacterium kroppenstedtii) DSM 44385; 마이코박테리움 아브세수스(Mycobacterium abscessus) ATCC 19977; 노카르디아 파르시니카(Nocardia farcinica) IFM10152; 로도코쿠스 에리트로폴리스(Rhodococcus erythropolis) PR4; 로도코쿠스 조스티이(Rhodococcus jostii) RHA1; 로도코쿠스 오파쿠스(Rhodococcus opacus) B4 uid36573; 아시도테르무스 셀룰롤리티쿠스(Acidothermus cellulolyticus) 11B; 아르트로박터 클로로페놀리쿠스(Arthrobacter chlorophenolicus) A6; 크리벨라 플라비다(Kribbella flavida) DSM 17836 uid43465; 서모모노스포라 쿠르바타(Thermomonospora curvata) DSM 43183; 비피도박테리움 덴티움(Bifidobacterium dentium) Bd1; 비피도박테리움 롱굼(Bifidobacterium longum) DJ010A; 슬라키아 헬리오트리

니레두센스(Slackia heliotrinireducens) DSM 20476; 페르세포넬라 마리나(Persephonella marina) EX H1; 박테로이데스 프라길리스(Bacteroides fragilis) NCTC 9434; 카프노사이토파가 오크라세아(Capnocytophaga ochracea) DSM 7271; 플라보박테리움 사이크로필룸(Flavobacterium psychrophilum) JIP0286; 알케르만시아 무시니필라(Akkermansia muciniphila) ATCC BAA 835; 로세이플렉수스 카스텐홀지이(Roseiflexus castenholzii) DSM 13941; 로세이플렉수스 RS1; 시네코시스티스(Synechocystis) PCC6803; 엘루시미크로븀 미누툼(Elusimicrobium minutum) Pei191; 비배양 흰개미 군 1 세균 계통 Rs D17; 피브로박터 숙시노제네스(Fibrobacter succinogenes) S85; 바실루스 세레우스(Bacillus cereus) ATCC 10987; 리스테리아 이노쿠아(Listeria innocua); 락토바실루스 카세이(Lactobacillus casei); 락토바실루스 람노수스(Lactobacillus rhamnosus) GG; 락토바실루스 살리바리우스(Lactobacillus salivarius) UCC118; 스트렙토코쿠스 아갈락티아에(Streptococcus agalactiae) A909; 스트렙토코쿠스 아갈락티아에 NEM316; 스트렙토코쿠스 아갈락티아에 2603; 스트렙토코쿠스 디스갈락티아에 에퀴시밀리스(Streptococcus dysgalactiae equisimilis) GGS 124; 스트렙토코쿠스 에퀴주에피데미쿠스(Streptococcus equi zooepidemicus) MGCS10565; 스트렙토코쿠스 갈롤리티쿠스(Streptococcus gallolyticus) UCN34 uid46061; 스트렙토코쿠스 고르도니이 칼리스(Streptococcus gordonii Challis) subst CH1; 스트렙토코쿠스 뮤탄스(Streptococcus mutans) NN2025 uid46353; 스트렙토코쿠스 뮤탄스; 스트렙토코쿠스 피오게네스(Streptococcus pyogenes) M1 GAS; 스트렙토코쿠스 피오게네스 MGAS5005; 스트렙토코쿠스 피오게네스 MGAS2096; 스트렙토코쿠스 피오게네스 MGAS9429; 스트렙토코쿠스 피오게네스 MGAS 10270; 스트렙토코쿠스 피오게네스 MGAS6180; 스트렙토코쿠스 피오게네스 MGAS315; 스트렙토코쿠스 피오게네스 SSI-1; 스트렙토코쿠스 피오게네스 MGAS10750; 스트렙토코쿠스 피오게네스 NZ131; 스트렙토코쿠스 서모필레스(Streptococcus thermophiles) CNRZ1066; 스트렙토코쿠스 서모필레스 LMD-9; 스트렙토코쿠스 서모필레스 LMG18311; 클로스트리듐 보툴리눔(Clostridium botulinum) A3 Loch Maree; 클로스트리듐 보툴리눔 B Eklund 17B; 클로스트리듐 보툴리눔 Ba4 657; 클로스트리듐 보툴리눔 F Langeland; 클로스트리듐 셀룰로라이티쿰(Clostridium cellulolyticum) H10; 피네골리다 마그나(Finegoldia magna) ATCC 29328; 유박테리움 렉탈레(Eubacterium rectale) ATCC 33656; 마이코플라스마 갈리셉티쿰(Mycoplasma gallisepticum); 마이코플라스마 모빌(Mycoplasma mobile) 163K; 마이코플라스마 페네트란스(Mycoplasma penetrans); 마이코플라스마 시노비아에(Mycoplasma synoviae) 53; 스트렙토바실루스 모닐리포르미스(Streptobacillus moniliformis) DSM12112; 브라디리조븀(Bradyrhizobium) BTAi1; 니트로박터 함부르젠시스(Nitrobacter hamburgensis) X14; 로도슈도모나스 팔루스트리스(Rhodopseudomonas palustris) BisB18; 로도슈도모나스 팔루스트리스 BisB5; 파르비바큘룸 라바멘티보란스(Parvibaculum lavamentivorans) DS-1; 디노로세오박터 쉬바에(Dinoroseobacter shibae) DFL 12; 글루코나세토박터 디아조트로피쿠스(Gluconacetobacter diazotrophicus) Pal 5 FAPERJ; 글루코나세토박터 디아조트로피쿠스 Pal 5 JGI; 아조스피릴룸(Azospirillum) B510 uid46085; 로도스피릴룸 루브룸(Rhodospirillum rubrum ATCC 11170; 디아포로박터(Diaphorobacter) TPSY uid29975; 베르미네프로박터 에이세니아에(Verminephrobacter eiseniae) EF01-2; 네이쎄리아 메닝기티데스(Neisseria meningitides) 053442; 네이쎄리아 메닝기티데스 alpha14; 네이쎄리아 메닝기티데스 Z2491; 데술포비브리오 살렉시젠스(Desulfovibrio salexigens) DSM 2638; 캄필로박터 제주니 도일레이(Campylobacter jejuni doylei) 269 97; 캄필로박터 제주니 81116; 캄필로박터 제주니; 캄필로박터 라리(Campylobacter lari) RM2100; 헬리코박터 헤파티쿠스(Helicobacter hepaticus); 월리넬라 숙시노제네스(Wolinella succinogenes); 톨루모나스 아우엔시스(Tolumonas auensis) DSM 9187; 슈도알테로모나스 아틀란티카(Pseudoalteromonas atlantica) T6c; 세와넬라 페알레아나(Shewanella pealeana) ATCC 700345; 레지오넬라 뉴모필라 파리스(Legionella pneumophila Paris); 악티노바실루스 숙시노제네스(Actinobacillus succinogenes) 130Z; 파스튜렐라 물토시다(Pasteurella multocida); 프란시셀라 툴라렌시스 노비시다(Francisella tularensis novicida) U112; 프란시셀라 툴라렌시스 홀락티카(Francisella tularensis holarctica); 프란시셀라 툴라렌시스 FSC 198; 프란시셀라 툴라렌시스 툴라렌시스(Francisella tularensis tularensis); 프란시셀라 툴라렌시스 WY96-3418; 및 트레포네마 덴티콜라(Treponema denticola) ATCC 35405. 이에 따라, 본 개시내용의 측면은 유형 II CRISPR 시스템에 존재하는 Cas9 단백질에 관한 것이다.

[0026] Cas9 단백질은 통상의 기술자에 의해 문헌에서 Csn1로 지칭될 수도 있다. 본원에서 기술되는 실험의 대상인 *에스. 피오게네스* Cas9 단백질 서열을 하기에 나타내었다. 그 전체가 참조로 포함되는 문헌 [Deltcheva et al., *Nature* 471, 602-607 (2011)]을 참조하라.

MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAE
ATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFG
NIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSD
VDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGN
LIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAI
LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYA
GYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELH
AILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEE
VVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFL
SGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI
IKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWG
RLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSL
HEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRER
MKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDH
IVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNL
TKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS
KLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVR
K
MIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDF
ATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVA
YSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPK
YSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVE
QHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGA
PAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD- (SEQ ID NO:1)

[0027]

[0028] 일 측면에 있어서, DNA 결합 단백질 뉴클레아제에는 그의 동족체 및 오르토로그, 그리고 이들에 대하여 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, 90 %, 95 %, 98 % 또는 99 % 이상의 상동성을 가지며 뉴클레아제 활성을 가지는 DNA 결합 단백질인 단백질 서열들이 포함된다.

[0029] 본 개시내용의 다른 측면은 일반적으로 예컨대 표적 핵산에의 더 큰 유전자의 삽입에 의한 표적 핵산, 예컨대 표적 유전자의 유전적 변형 또는 편집을 위한, DNA 결합 단백질 또는 시스템의 사용에 관한 것이다. 통상의 기술자라면, 본 개시내용을 바탕으로 대표적인 DNA 결합 시스템들을 용이하게 식별하게 될 것이다. 이와 같은 DNA 결합 시스템에는 ZFN, TALE, TALENS 또는 CRISPR/Cas9 뉴클레아제가 포함된다.

[0030] 소정 측면에서는, 1종 이상, 2종 이상 또는 다수의 외래 핵산을 세포에 도입하는 것을 포함하는, 세포에서의 핵산 편집 방법이 본원에서 기술된다. 세포에 도입되는 외래 핵산은 가이드 RNA 또는 가이드 RNA들, Cas9 단백질 또는 단백질들, 및 삽입될 대형 핵산 서열을 코딩하고 있다. 가이드 RNA와 Cas9 단백질은 합쳐서 공동-국재화 복합체로 지칭되는데, 통상의 기술자에 의하면 가이드 RNA와 Cas9 단백질이 DNA에 결합한 후 Cas9 단백질이 DNA를 절단함으로써 관심 제1 핵산 서열을 제거한다는 부분으로 이와 같은 용어가 이해되기 때문이다. 상기 대형 핵산 서열은 DNA에 삽입된다. 소정의 추가적인 측면에 따르면, 세포에 도입되는 외래 핵산은 가이드 RNA 또는 가이드 RNA들 및 Cas9 단백질을 코딩한다.

[0031] 본 개시내용에 따른 세포에는 본원에서 기술되는 바와 같이 외래 핵산이 도입 및 발현될 수 있는 임의의 세포가 포함된다. 본원에서 기술되는 본 개시내용의 기본적인 개념이 세포 유형에 의해 제한되는 것은 아니라는 것이 이해되어야 한다. 본 개시내용에 따른 세포에는 진핵 세포, 원핵 세포, 동물 세포, 식물 세포, 진균 세포, 고세균 세포, 진정세균 세포 등이 포함된다. 세포에는 효모 세포, 식물 세포 및 동물 세포와 같은 진핵 세포가 포함된다. 구체적인 세포에는 포유동물 세포가 포함된다. 구체적인 세포에는 줄기 세포, 예컨대 만능성 줄기 세포, 예컨대 인간 유도 만능성 줄기 세포가 포함된다.

[0032] 표적 핵산에는 DNA 결합 단백질 뉴클레아제가 틈새를 형성하거나 절단하는 데에 유용할 수 있는 임의의 핵산 서

열, 예컨대 본원에서 기술되는 바와 같은 공동-국재화 복합체를 형성하는 RNA 가이드된 DNA 결합 단백질이 포함된다. 표적 핵산에는 유전자가 포함된다. 본 개시내용의 목적상, 이중 가닥 DNA와 같은 DNA에는 표적 핵산이 포함될 수 있으므로, 공동-국재화 복합체는 표적 핵산에서, 또는 그와 인접하여, 또는 그 부근에서 공동-국재화 복합체가 표적 핵산에 대하여 원하는 효과를 가질 수 있는 방식으로 DNA에 결합하거나 달리 그와 공동-국재화될 수 있다. 이와 같은 표적 핵산에는 내인성 (또는 자연 발생) 핵산 및 외인성 (또는 외래) 핵산이 포함될 수 있다. 통상의 기술자라면, 본 개시내용을 바탕으로, 표적 핵산을 포함하는 DNA에 공동-국재화되는 가이드 RNA 및 Cas9 단백질을 용이하게 확인 또는 설계할 수 있을 것이다. DNA에는 게놈 DNA, 미토콘드리아 DNA, 바이러스 DNA 또는 외인성 DNA가 포함된다.

[0033] 외래 핵산 (즉 세포의 자연 핵산 조성의 일부가 아닌 것들)은 해당 도입용으로 통상의 기술자에게 알려져 있는 임의의 방법을 사용하여 세포에 도입될 수 있다. 그와 같은 방법에는 형질감염, 형질도입, 바이러스 형질도입, 미세주사, 리포펙션, 뉴클레오펙션, 나노입자 폭격, 형질전환, 접합 등이 포함된다. 통상의 기술자라면, 용이하게 확인가능한 문헌 자료들을 사용하여 그와 같은 방법들을 용이하게 이해하고 적합화하게 될 것이다.

[0034] 본 개시내용의 대표적인 것으로써 하기의 실시예가 제시된다. 이러한 실시예가 본 개시내용의 영역을 제한하는 것으로 간주되어서는 아니되는데, 본 개시내용, 도면 및 첨부된 청구범위를 보게 되면, 이들 및 기타 등가의 실시양태들이 드러나게 될 것이기 때문이다.

[0035] **[실시예]**

[0036] 실시예 1

[0037] sgRNA 표적 서열

[0038] 컴퓨터 알고리즘을 사용하여, 인간 Thy1 및 CD147 유전자 부근의 다양한 위치에서 인간 게놈에서의 그의 독특성을 바탕으로 sgRNA (단일 가이드 RNA) 서열들 (문헌 [Mali, P. *et al. Science* 339, 823-826 (2013)])을 확인하였다.

[0039] 인간 *Thy1*:

L1 CACAG TCTCA GAAAA GCGC AGG (SEQ ID NO:2)
L2 AAATA TCAGC GCGGT GGAT TGG (SEQ ID NO:3)
L3 GGTCA GGCTG AACTC GTAC TGG (SEQ ID NO:4)
R1 TTAGT AGCAA CGCTA CCCC AGG (SEQ ID NO:5)
R2 GTGTG CAGTC ATTAG CCCC TGG (SEQ ID NO:6)
R3 GGGCA AATGT GTCTC GTTA GGG (SEQ ID NO:7)
R4 TTCTC CTTTC CGAAG TCCG TGG (SEQ ID NO:8)
R5 GCCGC TGTCG CCTGG CAAA AGG (SEQ ID NO:9)
R6 GATGG TAGAC ATCGA CCAT GGG (SEQ ID NO:10)
R7 TTCAA TTTCG GGCCC GATC TGG (SEQ ID NO:11)
R8 TGAGT CGCGT CACGG CTAT TGG (SEQ ID NO:12)
R9 CATTT GCGGT GGTAA TCGC AGG (SEQ ID NO:13)
R10 GATCG GATCG GGTCG CGTC GGG (SEQ ID NO:14)

[0040]

[0041] 인간 *CD147*:

L147 TTTCC TGCGC TGAAT CGGG TGG (SEQ ID NO:15)
R147 GGCTC CTGTC TGTGC CTGA CGG (SEQ ID NO:16)

[0042]

[0043] 실시예 II

[0044] Cas9 및 sgRNA 플라스미드 구축

[0045] 문헌 [Mali, P. *et al. Science* 339, 823-826 (2013)] (IDT 사)에 기술되어 있는 바와 같이 U6 프로모터 및 sgRNA 백본 서열을 합성하고, 프라이머 5' CTTTCTTGTACAAAGTTGGCATTA_ttagacgtcaggtggcacttttc 3' (서열 17)

및 5' CCTTTAAAGCCTGCTTTTTTGTACA GTTTGCGTATTGGGCGCTCTTC 3' (서열 18)을 사용하여 pUC19로부터 PCR 증폭된 암피실린 내성 유전자 pBR322 ori를 포함하는 최소 플라스미드 백본에 등온 조립으로 클로닝하였다. 등온 조립을 사용하여, 다양한 sgRNA 서열들을 이와 같은 벡터에 클로닝하였다. 등온 조립을 위한 중첩 분절에는 밑줄을 쳤다. 모든 프라이머는 IDT 사의 것이며; 모든 PCR 반응은 카파 하이파이 핫스타트(KAPA HiFi HotStart) PCR 키트를 사용하여 수행하였다. 플라스미드는 TOP10 또는 Stbl3 세균 (인비트로겐(Invitrogen) 사) 중 어느 하나에서 유지하였다.

[0046] 프라이머 5' GCCACCATGGACAAGAAGTACTCC 3' (서열 19) 및 5' TCACACCTTCCTCTTCTTCTTGGG 3' (서열 20)을 사용하여 인간 코돈-최적화된 Cas9 뉴클레아제 유전자 (문헌 [Mali, P. *et al. Science* 339, 823-826 (2013)])를 PCR 증폭하고, 등온 조립을 사용하여 pCDNA3 플라스미드 백본상의 EF1α 프로모터와 bGH 폴리아데닐화 서열 사이에 클로닝하였다. 상기 EF1α 프로모터는 프라이머 5' CCGAAAAGTGCCACCTGACGTCGACGGA tgaaaggagtgGGAATTggc 3' (서열 21) 및 5' GGAGTACTTCTTGTCCATGGTGGC GGCC AACTAGCCAGCTTGGGTCTCCC 3' (서열 22)를 사용하여 pEGIP (Addgene #26777)로부터 PCR 증폭하였다. 상기 bGH 폴리아데닐화 서열은 프라이머 5' GCTGACCCCAAGAAGAAGAGGAAGGTGTGA CATCACCATTGAGTTTAAACCCGC 3' (서열 23) 및 5' CCAAGCTCTAGCTAGAGGTCGACG GTAT C GAGCCCCAGCTGGTTC 3' (서열 24)를 사용하여 pST1374 (Addgene #13426)로부터 PCR 증폭하였다. 플라스미드 백본은 프라이머 5' ATACCGTCGACCTCTAGCTAG 3' (서열 25) 및 5' TCCGTCGACGTCAGGTGG 3' (서열 26)을 사용하여 pCDNA3 (인비트로겐 사)로부터 PCR 증폭하였다.

[0047] 실시예 III

[0048] 표적화 플라스미드 구축

[0049] 등온 조립을 사용하여, sgRNA 부위 부근에서 상동성을 가지는 마우스 Thy1 표적화 벡터 (외부)를 클로닝하였다. 프라이머 5' TGGTGGTGGTTGTGGTACACACC 3' (서열 27) 및 5' AATAGGGAGGGCCGGGGTACC 3' (서열 28)을 사용하여 C57BL/6J 게놈 DNA (잭슨 래보래토리스(Jackson Laboratories) 사)로부터 마우스 Thy1의 엑손 2 및 3을 PCR 증폭하였다. 프라이머 5' AC CCTTCCCCTCCTAGATCCCAAGCC 3' (서열 29) 및 5' GATTAAAGGTGTGTACCACAACCACCACCA CTTTTCTGAGACTGTGAGGGAG 3' (서열 30)을 사용하여 PGP1 iPSC 게놈 DNA로부터 상류 인간 Thy1 상동성 아암을 PCR 증폭하였다. 프라이머 5' AGACTCTGGGGTACCCCGGCCCTCCCTATT CCCAGGGGCTAATGACTGC 3' (서열 31) 및 5' GCACCTCCAGCCATCACAGC 3' (서열 32)를 사용하여 PGP1 인간 iPSC 게놈 DNA로부터 하류 인간 Thy1 상동성 아암을 PCR 증폭하였다. 플라스미드 백본은 프라이머 5' CCAGGAAGGGGCTGTGATGGCTGGAGGTGC ttagacgtcaggtggcacttttc 3' (서열 33) 및 5' GGGCTTGGGATCTAGGAGGGGAAGG GTTTGCGTATTGGGCGCTCTTC 3' (서열 34)를 사용하여 pUC19로부터 PCR 증폭하였다.

[0050] 프라이머 5' AC CCTTCCCCTCCTAGATCCCAAGCC 3' (서열 35) 및 5' GCACCTCCAGCCATCACAGC 3' (서열 36)을 사용하여 원래의 외부 표적화 벡터로부터 마우스 Thy1 표적화 벡터의 선형 버전을 PCR 증폭하였다. 키아퀵(Qiaquick) PCR 정제 키트 (키아젠(Qiagen) 사)를 사용하여 PCR 생성물들을 세척하였다.

[0051] 하기 3종의 순방향 및 3종의 역방향 PCR 프라이머들의 조합을 사용하여 등온 조립으로, 원래의 외부 Thy1 표적화 벡터 플라스미드로부터 더 짧은 상동성 아암을 가지는 마우스 Thy1 표적화 벡터의 버전을 클로닝하였다. 상류 상동성 아암의 길이를 결정하는 순방향 프라이머:

[0052] 1550 bp, L: 5' ACCCTTCCCCTCCTAGATCCCAAGCC 3' (SEQ ID NO:37);
821 bp, M: 5' AAGATTCAGAGAGATTCATTCATTCATTCACAA 3' (SEQ ID NO:38);
100 bp, S: 5' CCTGCTAACAGGTACCCGGCATG 3' (SEQ ID NO:39).

[0053] 하류 상동성 아암의 길이를 결정하는 역방향 프라이머:

[0054] 2466 bp, L: 5' GCACCTCCAGCCATCACAGC 3' (SEQ ID NO:40);
797 bp, M: 5' CAGCATCTTGCTAAGGGGTTGTCAG 3' (SEQ ID NO:41);
94 bp, S: 5' GTCAGCAGACATGGGATGTTCGTTT 3' (SEQ ID NO:42).

[0055] 상류 및 하류 Thy1 상동성 아암에 대한 상보적 오버행(overhang)을 가지는 하기 3종의 순방향 및 역방향 프라이머들을 사용하여 pUC19로부터 플라스미드 백본을 PCR 증폭하였다.

[0056] 상류 오버행:

1550 bp: 5' GGGCTTGGGATCTAGGAGGGGAAGG GTTTGCGTATTGGGCGCTCTTC 3' (SEQ ID NO:43);

821 bp: 5' TGAATGAATGAATGAATCTCTCTGAATCTT GTTTGCGTATTGGGCGCTCTT C 3' (SEQ ID NO:44);

100 bp: 5' TCCTGCCCCATGCCGGGTACCTGTTAGCAG GTTTGCGTATTGGGCGCTCTTC 3' (SEQ ID NO:45).

[0057]

[0058] 하류 오버행:

2466 bp: 5' CCAGGAAGGGGCTGTGATGGCTGGAGGTGC ttagacgtcaggtggcacttttc 3' (SEQ ID NO:46);

797 bp: 5' GGAGGCTGACAACCCCTTAGCAAGATGCTG ttagacgtcaggtggcacttttc 3' (SEQ ID NO:47);

94 bp: 5' CAAATAAACGAACATCCCATGTCTGCTGAC ttagacgtcaggtggcacttttc 3' (SEQ ID NO:48).

[0059]

[0060] 등온 조립을 사용하여, 더 긴 상동성 아암을 가지는 마우스 Thy1 표적화 벡터의 버전 (XX)을 클로닝하였다. 프라이머 5' CCTTCCCCTCCTAGATCCCAAGCC 3' (서열 49) 및 5' GCACCTCCAGCCATCACAGC 3' (서열 50)을 사용하여 원래의 외부 Thy1 표적화 벡터 플라스미드를 PCR 증폭하였다. 프라이머 5' TCTTGTTTGAGATGTTGTGCGGG 3' (서열 51) 및 5' CTGGTTTCAGCACTCCGATCCTATC 3' (서열 52)를 사용하여 PGP1 게놈 DNA로부터 Thy1 상류 상동성 아암의 추가 3 kb를 PCR 증폭하였다. 프라이머 5' TGTGGCTCTGCACCAGGAAG 3' (서열 53) 및 5' CCTCTCCCTTTTCCCTGGTTTTG 3' (서열 54)를 사용하여 PGP1 게놈 DNA로부터 Thy1 하류 상동성 아암의 추가 2.4 kb를 PCR 증폭하였다. 상보적 오버행을 가지는 플라스미드 백본은 프라이머 5' TACTCTGCAAAACCAGGGAAAAGGGAGAGG ttagacgtcaggtggcacttttc 3' (서열 55) 및 5' CTGTGGATAGGATCGGAGTGCTGAAACCAG GTTTGCGTATTGGGCGCTCTTC 3' (서열 56)을 사용하여 pUC19로부터 PCR 증폭하였다.

[0061] 등온 조립을 사용하여, 각 sgRNA 부위 부근에서 상동성을 가지는 마우스 Thy1 표적화 벡터 (무손상)를 클로닝하였다. 프라이머 5' ATCTCTCCACTTCAGGTGGGTGGGAGGCCCCTGT GGTCTGTGTCTCCCCAAATT 3' (서열 57) 및 5' CAGGTGGACAGGAGGACAGATTCCAGAGGC TTGGTTTTATTGTGCAGTTTTCTTTC 3' (서열 58)을 사용하여 원래의 외부 Thy1 표적화 벡터 플라스미드로부터 마우스 Thy1 엑손 2 및 3의 더 짧은 단편을 PCR 증폭하였다. 프라이머 5' GGCTTCCTTCCCTCCAGAG 3' (서열 59) 및 5' ACAGGGGCCTCCCACCC 3' (서열 60)을 사용하여 PGP1 게놈 DNA로부터 sgRNA 부위 내 상류 상동성 아암의 연장된 단편을 PCR 증폭하였다. 프라이머 5' CAAGCCTCTGGAATCTGTCCTC 3' (서열 61) 및 5' GCCCAGTGTGCAGTCATTAGC 3' (서열 62)를 사용하여 PGP1 게놈 DNA로부터 sgRNA 부위 내 하류 상동성 아암의 연장된 단편을 PCR 증폭하였다. 나머지 상류 및 하류 상동성 아암 단편을 가지는 플라스미드 백본은 프라이머 5' CTTTTCTGAGACTGTGAGGGAG 3' (서열 63) 및 5' TACCCCAGGGGCTAATGACTGCAC 3' (서열 64)를 사용하여 원래의 외부 Thy1 표적화 벡터로부터 PCR 증폭하였다.

[0062] 등온 조립을 사용하여, 각 붕괴된 sgRNA 부위 부근에서 상동성을 가지는 마우스 Thy1 표적화 벡터 (붕괴)를 클로닝하였다. 각 sgRNA 부위로부터 하나의 뉴클레오티드를 결실시키기 위하여, 프라이머 5' ACAGTCTCAGAAAACGCAGGTGACAAAG 3' (서열 65) 및 5' CATTAGCCCCTGGGTAGCGTTGCTACTAAG 3' (서열 66), 그리고 이어서 5' TTGTCACCTGCGTTTTCTGAGACTGTGAG 3' (서열 67) 및 5' CTTAGTAGCAACGCTACCCAGGGGCTAATG 3' (서열 68)을 사용하여 무손상 Thy1 표적화 벡터 플라스미드로부터 2개 부문을 PCR 증폭하였다.

[0063] 등온 조립을 사용하여 mCherry Thy1 표적화 벡터를 클로닝하였다. 프라이머 5' GAGAATACCAGCAGTTCACCCATCCAGTAC GAAATTCTACCGGGTAGGGGAG 3' (서열 69) 및 5' CCCAGTGTGCAGTCATTAGCCCCTGGGGTA CGACGGCCAGTGAATTGTAATACG 3' (서열 70)을 사용하여 bGH 폴리아데닐화 서열을 가지는 pGK 프로모터의 조절하에 mCherry를 포함하는 플라스미드 구축물로부터 mCherry 트랜스진을 PCR 증폭하였다. 프라이머 5' AC CCTTCCCCTCCTAGATCCCAAGCC 3' (서열 71) 및 5' GTACTGGATGGGTGAACTGCTGGTATTC 3' (서열 72)를 사용하여 PGP1 게놈 DNA로부터 상류 상동성 아암을 PCR 증폭하였다. 프라이머 5' TACCCCAGGGGCTAATGACTGCAC 3' (서열 73) 및 5' GCACCTCCAGCCATCACAGC 3' (서열 74)를 사용하여 PGP1 게놈 DNA

로부터 하류 상동성 아암을 PCR 증폭하였다. 플라스미드 백본은 프라이머 5' CCAGGAAGGGGCTGTGATGGCTGGAGGTGC ttagacgtcaggtggcacttttc 3' (서열 75) 및 5' GGGCTTGGGATCTAGGAGGGGAAGG GTTTGCGTATTGGGCGCTCTTC 3 ' (서열 76)을 사용하여 pUC19로부터 PCR 증폭하였다.

[0064] 등온 조립을 사용하여 마우스 CD147 표적화 벡터를 클로닝하였다. 프라이머 5' GAAGTCGAGGTTCCAAGGTCACAGTGAG GGGGCCCTGGCCACCC CTTGCAGGTTCTCCATAGTCCACAG 3' (서열 77) 및 5' CAACAACCCCTCCTGTATATGACCT 3' (서열 78)을 사용하여 C57BL/6J 게놈 DNA (잭슨 래보래토리스 사)로부터 마우스 CD147의 엑손 2-7을 PCR 증폭하였다. 프라이머 5' ACACACTTTCAACCTCCAAGAGACG 3' (서열 79) 및 5' CTCACTGTGACCTTGGAACCTCG 3' (서열 80)을 사용하여 PGP1 게놈 DNA로부터 상류 상동성 아암을 PCR 증폭하였다. 프라이머 5' TGTTGAGGTCATATACAGGAGGGGTTGTTG CCTGACGGGGTTGGGTTTTCC 3' (서열 81) 및 5' AA GGGAGCCCTGAGGCCTTTTCC 3' (서열 82)를 사용하여 PGP1 게놈 DNA로부터 하류 상동성 아암을 PCR 증폭하였다.

[0065] 상보적 오버행을 가지는 플라스미드 백본은 프라이머 5' TCAGGAAAAGGCCTCAGGGCTCCC ttagacgtcaggtggcacttttc 3' (서열 83) 및 5' CGTCTCTTGGAGGTTGAAAGTGTGT GTTTGCGTATTGGGCGCTCTTC 3' (서열 84)를 사용하여 pUC19로부터 PCR 증폭하였다.

[0066] 실시예 IV

[0067] TALEN 조립체

[0068] 반복 캡핑된 조립체를 사용하여 인간 Thy1 유전자를 표적으로 하는 TALE 쌍 (16.5량체)를 조립하였다. L3 sgRNA 부위를 표적으로 하는 TALE 쌍: 좌측: 5' T ACCAGCAGTTCACCCAT 3' (서열 85); 우측: 5' T CTTTGTCTCACGGGTCA 3' (서열 86). R1 sgRNA 부위를 표적으로 하는 TALEN 쌍: 좌측: 5' T CTCCCCAACCACTTAGT 3' (서열 87); 우측: 5' T GTGCAGTCATTAGCCCC 3' (서열 88). 등온 조립을 사용하여 FokI 이종이량체 뉴클레아제 도메인상에 TALE를 클로닝하였다. 프라이머 5' GGCCGCCACCATGGATTATAAGGAC 3' (서열 89) 및 5' GGAACCTGCCACTCGATGTG 3' (서열 90)을 사용하여 조립된 TALE를 PCR 증폭하였다. 샤키(Sharkey) 돌연변이를 가지는 상기 FokI 이종이량체 뉴클레아제 도메인 KKR 및 ELD는 퀵체인지 라이트닝(QuikChange Lightning) 부위-지정 돌연변이유발 키트 (스트라타젠(Stratagene) 사)를 사용하여 pMG10 (Addgene #31238)로부터 유도하였다. 프라이머 5' GACATCACATCGAGTGGCAGGTTCC CAGCTGGTGAAGTCCGAGC 3' (서열 91) 및 5' CAACTAGAAGGCACAGTCGAGGC TGATCAGCG GGGTTA GAAATTGATTTCACCATTGTTGAAC 3' (서열 92)를 사용하여 KKR-샤키 FokI 도메인을 PCR 증폭하였다. 프라이머 5' GACATCACATCGAGTGGCAGGTTCC CAACTAGTCAAAAGTGAACTGGAGG 3' (서열 93) 및 5' CAACTAGAAGGCACAGTCGAGGC TGATCAGCG CCCTTAAAAGTTTATCTCGCCG 3' (서열 94)를 사용하여 ELD-샤키 FokI 도메인을 PCR 증폭하였다. 각 TALE 및 FokI 이종이량체 도메인을 EF1α 프로모터 및 bGH 폴리아데닐화 서열을 포함하는 플라스미드 백본에 클로닝하고; 프라이머 5' GCCTCGACTGTGCCTTCTAGTTG 3' (서열 95) 및 5' CTTATAA TCCATGGTGGCGGCC 3' (서열 96)을 사용하여 상기에서 기술된 Cas9 발현 벡터로부터 이를 증폭하였다.

[0069] 실시예 V

[0070] iPSC 배양 및 형질감염

[0071] 코리엘(Coriell) 사를 통하여, 개인 게놈 프로젝트(Personal Genome Project) 공여자 PGP1 및 PGP4로부터의 입증된 인간 iPSC를 입수하였다. 세포주를 마트리겔(Matrigel)-코팅된 플레이트 (BD 사)상에서 유지하면서, mTesr1 (스템 셀 테크놀로지스(Stem Cell Technologies) 사) 중에서 제조자의 지침에 따라 성장시켰다. 아큐타제 (밀리포어(Millipore) 사)를 사용한 배양 전, 동안 및 후에, 10 μM의 록(Rock) 억제제 Y-27632 (밀리포어 사)를 배양물에 첨가하였다. TRA-1/60 FACS 염색 (BD 사)에 의해 iPSC 배양물의 만능성을 입증하였다.

[0072] 모든 플라스미드는 키아젠 엔도-프리 플라스미드 맥시프레프(Qiagen Endo-free Plasmid Maxiprep) 키트를 사용하여 정제하였다. 론자(Lonza) 4D-뉴클레오펙터(Nucleofector) X 장치 (버퍼 P3, 프로그램 CB-150)를 사용하여 제조자의 지침에 따라, 플라스미드들을 iPSC 세포에 뉴클레오펙션시켰다(Nucleofected). 각 20 μl의 뉴클레오펙션 반응 당, 0.2-0.5×10^6 iPSC를 4 ㎍ 이하의 플라스미드 DNA로 형질감염시켰다. 뉴클레오펙션-후, mTesr1 배지 더하기 10 μM Y-27632를 함유하는 24- 및 96-웰 마트리겔-코팅 플레이트상에 iPSC를 플레이팅하였다.

[0073] 표적화 벡터를 사용한 CRISPR-기반 뉴클레오펙션 (도 1(a)-(d) 및 도 3의 (A)-(C) 내지 도 8(a)-(c) 참조)을 위하여, 2 ㎍의 표적화 벡터 플라스미드, 0.5 ㎍의 Cas9 플라스미드 및 1.5 ㎍의 총 sgRNA 플라스미드를 사용하였다. 2종의 sgRNA가 사용된 경우에는, 각 플라스미드 0.75 ㎍을 조합하였다. sgRNA가 사용되지 않은 경우,

대신 1.5 ㎍의 pUC19를 사용하였다.

[0074] 표적화 벡터가 없는 CRISPR-기반 뉴클레오펙션 (도 2의 (A)-(E) 참조)을 위해서는, 2 ㎍의 하기 총 플라스미드를 사용하였다: 0.75 ㎍의 각 sgRNA 플라스미드 또는 0.75 ㎍의 pUC19와 함께인 0.5 ㎍의 Cas9 플라스미드.

[0075] 표적화 벡터를 사용한 TALEN-기반 뉴클레오펙션 (도 5의 (A)-(B) 참조)의 경우, 2 ㎍의 표적화 벡터 플라스미드 더하기 총 2 ㎍의 TALEN 플라스미드를 사용하였다. 하나의 TALEN 쌍을 사용한 1회의 dsDNA 절단 당, 1 ㎍의 각 TALEN-발현 이종이량체 플라스미드를 사용하였다. 2개 TALEN 쌍을 사용한 2회의 dsDNA 절단 당으로는, 0.5 ㎍의 각 TALEN-발현 이종이량체 플라스미드를 사용하였다.

[0076] 실시예 VI

[0077] FACS 염색

[0078] 트라이플 익스프레스(TrypLE Express) (인비트로겐 사)를 사용하여 iPSC를 해리시키고, 하기 FACS 완충제에서 세척하였다: PBS (인비트로겐 사) + 0.2 % 소 혈청 알부민 (시그마(Sigma) 사). FACS 완충제 더하기 10 % 소 태아 혈청 중에서 4 ℃로 30분 동안 세포를 염색하였다. 하기의 항체들을 이바이오사이언스(eBioscience) 사로부터 구매하였다: 항-인간 Thy1 APC (eBio5E10), 항-마우스 Thy1.2 PE (30-H12), 항-인간 CD147 APC (8D12), 항-마우스 CD147 PE (RL73), 동형 대조군 마우스 IgG1 κ APC (P3.6.2.8.1), 동형 대조군 마우스 IgG2b PE (eBMG2b). FACS 완충제 중에서 세포를 2회 세척한 다음, 생존력 염료 사이톡스 블루(SYTOX Blue) (인비트로겐 사)와 함께 FACS 완충제 중에 재현탁하였다. 고처리량 샘플러 (HTS)가 구비된 BD LSRFortessa 유동 세포측정장치에서 샘플을 수집하고, 플루오조(FlowJo) 소프트웨어 (트리 스타(Tree Star) 사)를 사용하여 분석하였다.

[0079] 도 8(a)-(c)에 나타낸 일정한 생존가능 세포 계수용으로, 96-웰 플레이트의 하나의 플레이트 중에서 각 샘플을 성장시켰다. 50 μl의 트라이플을 사용하여 각 웰을 해리시켰다. 다음에, 생존력 염료 토프로(ToPro)3 (인비트로겐 사)을 함유하는 150 μl의 FACS 완충제를 첨가하였다. 혼합하면서, HTS를 사용하여 1 μl/s의 일정한 속도로 각 웰로부터의 100 μl를 분석하였다.

[0080] 실시예 VII

[0081] 단일-세포 iPSC FACS 분류

[0082] FACS 분류를 위하여, 공급자 세포가 담긴 96-웰 플레이트에 1개 세포/웰로 iPSC를 분류하였다. 전날 밤에, 96-웰 편평-바닥 조직 배양 플레이트를 젤라틴 (밀리포어 사)으로 코팅한 후, 조사된 CF-1 마우스 배아 섬유모세포 (플레이트 당 10^6 MEF; 글로벌 스템(Global Stem) 사)를 사용하여 배양하였다. 분류 전에, 플레이트 중 배지를 100 ng/ml의 bFGF (밀리포어 사), SMC4 억제제 (바이오 비전(Bio Vision) 사) 및 5 ㎍/ml의 피브로넥틴 (시그마 사)을 포함하는 hES 세포 유지 배지로 교체하였다.

[0083] FACS 분류 전 2시간 이상 동안, SMC4 억제제를 함유하는 mTesr1을 사용하여 iPSC를 예비-처리하였다. 아큐타제를 사용하여 세포를 해리하고, 상기한 바와 같이 염색하였다. BD FACS 아리아(Aria)를 사용하여 MEF-코팅된 96-웰 플레이트로 iPSC를 분류하였다. 다음에, 확립된 iPSC 집락을 새로운 MEF-코팅 웰상에 기계적 계대배양하였다.

[0084] 실시예 VIII

[0085] PCR 제노타이핑

[0086] 분류된 iPSC 클론들로부터의 게놈 DNA를 96-웰 플레이트로부터 정제하였다. 4 세트의 PCR 프라이머들 (도 1(a)-(d) 참조)을 사용하고 카파 하이파이 핫스타(KAPA HiFi HotStart) 폴리머라제를 사용하여 각 클론을 제노타이핑하고(genotyped), 0.8 % 아가로스 겔상에서 전개하였다.

[0087] 반응 1: 5' AGGGACTTAGATGACTGCCATAGCC 3' (서열 97) 및 5' ATGTTGGCAGTAAGCATGTTGTTCC 3' (서열 98). 야생형 Thy1 (+) 또는 역전 대립유전자 (i): 3129 bp; 표적화된 마우스 Thy1 대립유전자 (m): 2904 bp; 절제된 대립유전자 (Δ): 387 bp.

[0088] 반응 2: 5' AGGGACTTAGATGACTGCCATAGCC 3' (서열 99), 5' CTCACCTCTGAGCACTGTGACGTTC 3' (서열 100) 및 5' ACTGAAGTTCTGGGTCCCAACAATG 3' (서열 101). 야생형 대립유전자 (+): 993 bp; 표적화된 마우스 대립유전자 (m): 490 bp; 절제되거나 (Δ) 역전된 (i) 대립유전자: PCR 생성물 없음.

[0089] 반응 3: 5' ATGAATACAGACTGCACCTCCCCAG 3' (서열 102),

[0090] 5' CTCACCTCTGAGCACTGTGACGTTC 3' (서열 103) 및 5' CCATCAATCTACTGAAGTTCTGGGTCCCAACAATG 3' (서열 104). 야생형 대립유전자 (+): 2393 bp; 표적화된 마우스 대립유전자 (m): 1891 bp; 절제되거나 (Δ) 역전된 (i) 대립유전자: PCR 생성물 없음.

[0091] 반응 4: 5' TGAAGTGAAACCCTAAAGGGGGAAG 3' (서열 105), 5' AAACCACACACTTCAACCTGGATGG 3' (서열 106) 및 5' GTTTGGCCCAAGTTTCTAAGGGAGG 3' (서열 107). 야생형 대립유전자 (+): 3064 bp; 표적화된 마우스 대립유전자 (m): 2707; 절제되거나 (Δ) 역전된 (i) 대립유전자: PCR 생성물 없음.

[0092] sgRNA 뉴클레아제 부위 외부로 연장되는 제노타이핑 반응 1의 프라이머들을 사용하여 분류된 iPSC 클론들로부터의 게놈 DNA를 PCR 증폭하였다. 아가로스 겔 추출 (키아젠 사)을 사용하여 서로 다른 크기의 PCR 생성물들을 분류하였다. 동일한 2종의 프라이머들을 사용하여 PCR 생성물들을 생거(Sanger) 서열분석 (제네위즈(Genewiz) 사)하였다. 생거 추적 파일로부터 단일 및 이중 피크를 분석하고, 각 클론의 이중대립유전자 서열을 확인하기 위하여 디컨볼루션하였다.

[0093] 실시예 IX

[0094] 본 개시내용의 측면은 항체 선택 마커를 사용하는 경우에도 표적화된 유전자 대체를 제한할 수 있는 상동성 재조합 (HR)의 낮은 빈도 (10^{-3}-10^{-7})를 향상시키는 것에 관한 것이다. 문헌 [Deng, C. & Capecchi, M. R. *Mol. Cell. Biol.* 12, 3365-3371 (1992)]을 참조하라. 본 개시내용의 측면은 긴 핵산 서열, 예컨대 긴 유전자 서열의 절제 및 대체와 같은 효율적인 게놈 변형에 있어서의 징크 핑거 뉴클레아제(zinc finger nuclease) (ZFN) (문헌 [Urnov, F. D., Rebar, E. J., Holmes, M. C., Zhang, H. S. & Gregory, P. D. *Nat. Rev. Genet.* 11, 636-646 (2010)]), 전사 활성화인자-유사 이펙터 뉴클레아제 (TALEN) (문헌 [Joung, J. K. & Sander, J. D. *Nature Reviews Molecular Cell Biology* 14, 49-55 (2012)]) 또는 CRISPR/Cas9 뉴클레아제 (문헌 [Mali, P. *et al. Science* 339, 823-826 (2013)])와 같은 맞춤-조작 뉴클레아제 시스템의 사용을 고려한다. 본 개시내용의 측면은 통상적인 방법을 사용한 유전자 편집에 내성일 수 있는 세포 유형의 사용을 고려한다.

[0095] 본 개시내용의 측면은 이후 NHEJ 복구 경로 (문헌 [Chapman, J. R., Taylor, M. R. G. & Boulton, S. J. *Molecular Cell* 47, 497-510 (2012)])가 유전자를 돌연변이시키고 붕괴시킬 수 있는 특정 표적 부위에서의 하나 이상의 dsDNA 절단의 사용을 고려한다 (문헌 [Urnov, F. D., Rebar, E. J., Holmes, M. C, Zhang, H. S. & Gregory, P. D. *Nat. Rev. Genet.* 11, 636-646 (2010)], [Mali, P. *et al. Science* 339, 823-826 (2013)]). 본 개시내용의 측면은 게놈의 개재 부분을 절제하거나 (문헌 [Lee, H. J., Kim, E. & Kim, J.-S. *Genome Research* 20, 81-89 (2010)]), 전위를 생성시킬 수 있는 (문헌 [Piganeau, M. et al. *Genome Research* 23, 1182-1193 (2013)]) 2개의 dsDNA 절단을 고려한다. 본 개시내용의 측면은 돌연변이 또는 트랜스진을 도입하기 위한 (문헌 [Moehle, E. A. *et al. Proc. Natl. Acad. Sci. U.S.A.* 104, 3055-3060 (2007)], [Hockemeyer, D. *et al. Nat Biotechnol* 27, 851-857 (2009)]) ssODN을 사용하는 HR (문헌 [Chen, F. *et al. Nature Methods* 8, 753-755 (2011)], [Yang, L. et al. Nucleic Acids Research (2013). doi:10.1093/nar/gkt555]) 또는 플라스미드 표적화 벡터 (문헌 [Yang, L. et al. Nucleic Acids Research (2013). doi:10.1093/nar/gkt555], [Mali, P. *et al. Science* 339, 823-826 (2013)])의 사용을 고려한다. 본 개시내용의 측면은 효율적일 수 있는 더 큰 삽입물의 유전자 삽입 효율을 향상시키기 위한 본원에서 기술되는 방법을 고려한다 (문헌 [Moehle, E. A. *et al. Proc. Natl. Acad. Sci. U.S.A.* 104, 3055-3060 (2007)], [Urnov, F. D., Rebar, E. J., Holmes, M. C, Zhang, H. S. & Gregory, P. D. *Nat. Rev. Genet.* 11, 636-646 (2010)]).

[0096] 본 개시내용의 측면은 ZFN 절단으로부터 기인할 수 있는 짧은 단일-가닥 오버행들 사이의 본원에서 기술되는 방법 및 미세상동성-매개 말단 연결 (MMEJ)을 사용한 다수-킬로베이스의 표적화된 유전자 대체물을 생성시키는 것을 고려한다 (문헌 [Orlando, S. J. *et al. Nucleic Acids Research* 38, e152 (2010)], [Cristea, S. *et al. Biotechnol. Bioeng.* 110, 871-880 (2013)], [Maresca, M., Lin, V. G., Guo, N. & Yang, Y. *Genome Research* 23, 539-546 (2013)]). 본 개시내용의 측면은 각 상동성 아암에 dsDNA 절단을 생성시키는 것에 의한 측접 상동성 아암들 사이의 HR 교차의 효율을 향상시키는 것을 고려한다. 추가적으로 고려되는 것은 대형 핵산 삽입의 효율을 향상시키기 위한 각 상동성 아암에서의 다중 절단 부위의 사용이다.

[0097] 실시예 X

[0098] 표적 벡터 설계의 최적화

[0099] 본 개시내용의 측면은 대형 핵산 대체용 표적화 벡터 설계를 최적화하는 것에 관한 것이다. 대표적인 측면에 따르면, 개인 게놈 프로젝트 (PGP) 공여자로부터 유래하는 인간 iPSC에서 2.7 kb 인간 *Thy1* 유전자 (hThy1)를 그의 마우스 동족체 (mThy1)로 대체하는 것이다. 인간 *Thy1* (CD90)은 본원에서 기술되는 방법을 사용한 대형 유전자 대체의 예를 입증하는 데에 유리한데, 그것이 인간 iPSC의 표면에서 발현되며, 시험관 내에서의 세포 생존에 그것이 필수적이 아니고, 종-특이적 염색 항체가 이용가능하기 때문이다. 본 실시예는 단지 예시로써, 인간 Thy1의 절제 및 마우스 Thy1을 사용한 대체로 본 개시내용의 영역을 제한하고자 하는 것은 아님이 이해되어야 한다.

[0100] 인간 *Thy1*의 인트론 1 내 또는 폴리아데닐화 서열 후를 표적으로 하는 2종의 단일 가이드 RNA (sgRNA)를 설계하였다. 상기 mThy1 표적화 벡터 플라스미드는 절단 부위 외부에 인간 *Thy1* 상동성 아암이 측접하는 마우스 *Thy1*의 엑손 2 및 3을 포함하였다 (도 1a). 양 dsDNA 절단이 이루어진 경우, 6.7 %의 iPSC가 선택 없이 $mThy1^{+}hThy1^{-}$가 되었다 (도 1b). 단일 세포 FACS-분류 $mThy1^{+}hThy1^{-}$ iPSC 클론의 PCR 제노타이핑은 동형접합 표적화 대체 (m/m; 4.7 %) 및 한 인간 대립유전자 대체 와 다른 것의 절제 (m/Δ; 2 %)의 혼재를 나타냈다 (도 1(c)-(d)). 또한, 1.6 %의 세포가 $mThy1^{+}hThy1^{+}$ 이중 양성이었다 (m/+; 이형접합 표적화 대체).

[0101] 마지막으로, 16.2 %의 세포가 $mThy1^{-}hThy1^{-}$ 이중 음성이었는데: PCR 및 생거 서열분석은 sgRNA 부위들 사이의 동형접합 절제 (Δ/Δ) 및 이형접합 역전 및 절제 (i/Δ)의 혼재를 나타냈다 (도 1(a)-(d)). 절제 및 역전 부위에서 소수의 indel 및 삽입된 염기가 관찰되기는 하였지만, 최대 indel이 겨우 15 bp이었으며, 대부분의 대립유전자들이 절단 부위들 사이에서 바로 재-연결되었다 (도 3(A)-(C)). ZFN 또는 TALEN을 사용하여 2개의 dsDNA 절단을 생성시키는 이전의 기록은 대부분의 절제 및 역전 대립유전자에서 200 bp까지 indel을 관찰하였었다 (문헌 [Lee, H. J., Kim, E. & Kim, J.-S. *Genome Research* 20, 81-89 (2010)], [Lee, H. J., Kweon, J., Kim, E., Kim, S. & Kim, J.S. *Genome Research* 22, 539-548 (2012)], [Piganeau, M. *et al. Genome Research* 23, 1182-1193 (2013)]). ZFN 및 TALEN에 의해 생성되는 5' 오버행과 달리, Cas9 뉴클레아제는 둔단 dsDNA 절단을 생성시키는데 (문헌 [Jinek, M. *et al. Science* 337, 816-821 (2012)]), 이는 더 짧은 19 bp (문헌 [Mali, P. *et al. Science* 339, 823-826 (2013)]) 및 118 bp (문헌 [Cong, L. *et al. Science* 339, 819-823 (2013)]) Cas9-매개 절제에서도 나타났던 재-연결의 증가된 정확성에 기여할 수 있다.

[0102] 단일 sgRNA만이 사용된 경우, mThy1 동형접합 대체는 > 2 %의 세포에서 발생하였으며 ($mThy1^{+}hThy1^{-}$; m/m), 이형접합 대체는 4-6 %의 세포에서 발생하였다 ($mThy1^{+}hThy1^{+}$; m/+). 매우 적은 $mThy1^{-}hThy1^{-}$ 이중 음성 세포가 관찰되었으며, 절제된 hThy1 대립유전자 (Δ)는 관찰되지 않았다 (도 1(a)-(d)). 9.8 kb 인간 CD147 유전자를 그의 마우스 동족체로 대체하거나 hThy1을 형광 리포터로 대체한 경우, 유사한 패턴의 결과가 나타났다 (도 4(A)-(B) 및 도 5(A)-(B)).

[0103] 실시예 XI

[0104] 인간 iPSC에서의 표적화된 유전자 대체에 대한 상동성 길이의 효과 측정

[0105] 통상적인 유전자 표적화 벡터는 일반적으로 선형화된 플라스미드로 형질감염된다. ZFN의 경우, MMEJ-매개 유전자 삽입에는 선형의 구축물이 더 효과적이었지만, 원형의 플라스미드 표적화 구축물이 선형화된 플라스미드에 비해 더 높은 HR-매개 유전자 삽입 비율을 산출하였다 (문헌 [Orlando, S. J. *et al. Nucleic Acids Research* 38, e152 (2010)], [Cristea, S. *et al. Biotechnol. Bioeng.* 110, 871-880 (2013)]). 선형화된 mThy1 표적화 벡터는 원형 플라스미드에 비해 훨씬 더 적은 유전자 표적화를 산출하였는데 (도 (a)-(b)), 선형화된 플라스미드의 감소된 뉴클레오펙션 효율 또는 증가된 분해로 인한 것일 수 있다 (문헌 [Cristea, S. *et al. Biotechnol. Bioeng.* 110, 871-880 (2013)]).

[0106] 세균 인공 염색체를 사용한 추가적인 상동성 아암 길이 (~70 kb까지)가 약하거나 비-동종인 표적화 벡터를 향상시킬 수 있기는 하지만 (문헌 [Valenzuela, D.M. *et al. Nat Biotechnol* 21, 652-659 (2003)]), 통상적인 HR-매개 유전자 표적화에 있어서, 표적화 빈도는 ~14 kb까지 상동성 아암 길이에 따라 증가된다 (문헌 [Deng, C. & Capecchi, M. R. *Mol. Cell. Biol.* 12, 3365-3371 (1992)]). 그러나, dsDNA 절단을 도입하는 것은 단일 절단 부위에의 트랜스진 삽입을 위한 필요 상동성 아암 길이를 ~0.2-0.8 kb까지 감소시킨다 (문헌 [Elliott, B., Richardson, C., Winderbaum, J., Nickoloff, J.A. & Jasin, M. *Mol. Cell. Biol.* 18, 93-101 (1998)], [Moehle, E. A. *et al. Proc. Natl. Acad. Sci. U.S.A.* 104, 3055-3060 (2007)], [Hockemeyer, D. et al. Nat Biotechnol 27, 851-857 (2009)], [Orlando, S. J. et al. Nucleic Acids Research 38, e152 (2010)],

[Beumer, K.J., Trautman, J.K., Mukherjee, K. & Carroll, D. G3 Genes|Genomes|Genetics 3, 657-664 (2013)]). 인간 iPSC에서의 표적화된 유전자 대체에 대한 상동성 길이의 효과를 조사하기 위하여, 다양한 길이의 상동성 아암을 가지는 mThy1 표적화 벡터의 버전들을 구축하였다 (도 7(a)). 원래 표적화 벡터로부터의 ~2 kb 상동성 아암 (장, L) 이외에, ~800 bp (중간, M) 또는 ~100 bp (단, S)의 더 짧은 길이들을 선택하였는데, 이러한 길이들이 HR-매개 유전자 삽입 (문헌 [Moehle, E. A. *et al. Proc. Natl. Acad. Sci. U.S.A.* 104, 3055-3060 (2007)], [Hockemeyer, D. *et al. Nat Biotechnol* 27, 851-857 (2009)]) 또는 ssODN 보정 (문헌 [Chen, F. *et al. Nature Methods* 8, 753-755 (2011)], [Yang, L. et al. Nucleic Acids Research (2013). doi:10.1093/nar/gkt555])에 종종 사용되기 때문이다. 더 긴 ~5 kb (초장, X)의 상동성 아암도 시험하였다. 2개 dsDNA 절단의 경우, 표적화 빈도가 ~2 kb 상동성 아암 (LL 및 ML)에서 가장 높았으며, 대체적으로 상동성이 더 적을수록 감소하였지만, ~1.5 kb 총 상동성 (MM, LS 및 SL)까지 내려가도 > 1 %의 빈도가 달성되었다. 초장 상동성 아암은 유전자 표적화 효율을 향상시키지 않았다 (XX 대 LL).

[0107] 하나의 dsDNA 절단만을 사용한 경우, 절단 부위 반대편 아암 상에서의 상동성 길이가 가장 중요하였다. LM 및 LS 벡터는 좌측에 비해 우측 sgRNA에서 더 높은 유전자 표적화를 나타내었으며; ML 및 SL 벡터는 우측에 비해 좌측 sgRNA에서 더 높은 유전자 표적화를 나타내었다 (도 7(a)). 표적화 효율이 더 낮기는 하였다 할지라도, 동일한 패턴의 결과가 PGP4 iPSC에서 관찰되었다 (도 7(b)). 이러한 결과는 dsDNA 절단 외부에서 절단된 염색체가 표적화 벡터 플라스미드의 상응하는 서열에 어닐링되는 합성-의존 가닥 어닐링(Synthesis-Dependent Strand Annealing)의 모델 (문헌 [Moehle, E. A. et al. Proc. Natl. Acad. Sci. U.S.A. 104, 3055-3060 (2007)], [Urnov, F. D., Rebar, E. J., Holmes, M. C., Zhang, H. S. & Gregory, P. D. *Nat. Rev. Genet.* 11, 636-646 (2010)], [Chapman, J.R., Taylor, M. R. G. & Boulton, S. J. Molecular Cell 47, 497-510 (2012)])과 일치한다. 표적화 벡터의 이종유래 마우스 *Thy1* 서열은 도입되어, 반대편 상동성 아암상의 인간 서열에 도달할 때까지 D-루프를 형성한다. 이와 같은 상동성 아암의 충분한 길이는 D-루프의 해상(resolution)을 위한 염색체의 상응 부분에 대한 이후의 그의 상동성 탐색 및 재-어닐링을 가능케 한다.

[0108] 이와 같은 모델에서, 상동성 아암 서열은 dsDNA 절단 부위 외부로 연장되어야 하며, 이종유래 대체 서열은 내부에 존재해야 한다. 각 절단 부위의 일 측을 포괄하는 상동성 아암은 PGP1 또는 PGP4 iPSC에서 표적화 효율을 향상시키지 않았다 (도 8(a)-(c)). 표적화 벡터가 이제는 무손상 sgRNA 표적 부위를 포함하기 때문에, 2배 더 많은 iPSC사멸이 관찰되었는데 (도 8(b)), 세포 내 dsDNA 절단의 압도적인 수에 기인하는 것일 가능성이 있다. 이러한 sgRNA 부위들이 단일 bp 결실에 의해 붕괴된 경우, 증가된 세포 사멸은 관찰되지 않았으나, 원래의 mThy1 표적화 벡터에 비해 표적화 효율이 여전히 감소되었다. 과학적 이론에 얽매이고자 하는 것은 아니나, 상기 결과가 각 dsDNA 절단의 양 말단이 표적화 벡터의 상응하는 상동성 서열에 어닐링되어 별도의 교차를 형성하는 이중 홀리데이(Holliday) 접합의 탁월한 HR 기작을 암시하는 것은 아니다 (문헌 [Chapman, J. R., Taylor, M. R. G. & Boulton, S. J. Molecular Cell 47, 497-510 (2012)]).

[0109] 실시예 XII

[0110] 인간 iPSC에서의 Cas9-매개 결실 빈도와 크기 사이의 관계 측정

[0111] 대체적으로 결실 크기가 더 클수록 결실 빈도가 감소하기는 하였지만, 종양 세포주에서 ZFN을 사용하여 다수-킬로베이스의 결실을 달성한 바 있다 (문헌 [Lee, H. J., Kim, E. & Kim, J.-S. *Genome Research* 20, 81-89 (2010)], [Chen, F. *et al. Nature Methods* 8, 753-755 (2011)]). 인간 iPSC에서의 빈도와 Cas9-매개 결실 크기 사이의 관계를 설명하기 위하여, 다양한 거리로 이격되어 hThy1을 표적으로 하는 2개의 좌측 및 10개의 우측 sgRNA를 설계하였다 (도 2A). 표적화 벡터가 사용되지 않았기 때문에, 붕괴된 양 hThy1 대립유전자를 가지는 세포는 $hThy1^-$가 되었다 (도 2B). 좌측 sgRNA로만 뉴클레오펙션된 세포를 사용하여 $hThy1^-$ 세포의 바탕 수준을 측정하였다 (도 2B, 우측 컬럼). 바탕 수준을 초과하는 동형접합 결실의 빈도가 더 짧은 거리에서 더 높은 경향 - 2.7 kb 결실에서 24 % 및 86 kb 결실에서 8 %까지 -이 있었지만, 다른 sgRNA 부위들이 훨씬 더 낮은 결실 빈도를 산출함으로써, 이는 항상 크기에 상응하지는 않았다 (도 2D).

[0112] 도 1(a)-(d)에 기술되어 있는 바와 같이 생성된 iPSC의 $mThy1^+hThy1^+$ 클론 주를 사용하여, 단일대립유전자 결실의 빈도를 측정하였다. mThy1 대립유전자는 2개의 좌측 sgRNA 부위를 포함하지 않기 때문에, 하나의 나머지 hThy1 대립유전자만이 결실에 적용되었다. 이형접합 결실의 빈도가 때로는 동일한 sgRNA 쌍으로부터의 동형접합 결실에서의 것에 비해 더 높기는 하였지만, 그것은 항상 수 % 이내였다 (도 2(C)-(E)). 과학 이론에 얽매이고자 하는 것은 아니나, 동일한 우측 sgRNA와 쌍을 이루었을 때 L1 또는 L2 sgRNA 중 어느 것도 부합하여 더 많을 결실을 산출하지 않았기 때문에, 표적 부위에서의 융점, 유전자 발현 또는 염색질 환경의 차이로 인한 sgRNA

부위들 간에 가변적인 뉴클레아제 활성 (문헌 [Yang, L. *et al. Nucleic Acids Research* (2013). doi:10.1093/nar/gkt555])은 유전자 결실 빈도에서 관찰되는 편차의 원인이 아닐 가능성이 있다. 쌍-특이적 변수, 예컨대 2개 dsDNA 절단 부위들 사이의 미세상동성은 결실 빈도에 영향을 줄 수 있다.

[0113] 실시예 XIII

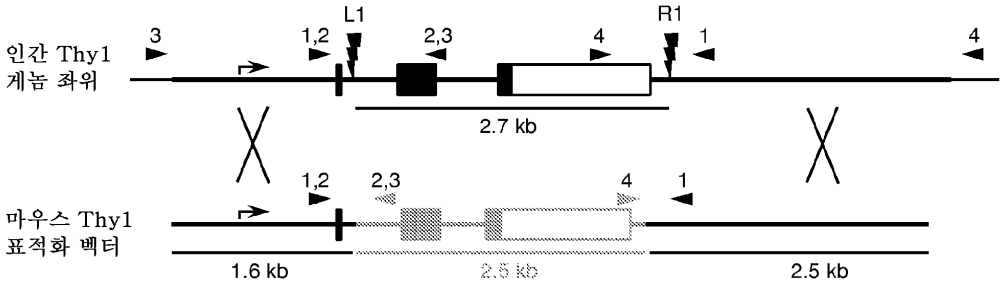
[0114] 유전자 대체 벡터를 위한 상이한 뉴클레아제 및 최적 설계를 사용한 유전자 대체

[0115] 소정 측면에서는, 본원에서 기술되는 유전자 대체 벡터용 설계가 상이한 뉴클레아제와 함께 사용될 수 있는데, 그것이 DNA 절단을 생성시키는 시스템에 있어서 특정되는 것은 아니기 때문이다. 효율적인 다수-킬로베이스 유전자 대체는 ZFN, TALEN 및 CRISPR 뉴클레아제를 사용하여 달성되었다 (도 5(A)-(B), 데이터 미도시). 하나의 절단 부위를 가지는 표적화된 유전자 대체가 덜 효율적이기는 하였지만, 단일 절단 부위의 사용은 잠재적인 유전형 및 표적외 돌연변이 형성을 감소시킨다. 현재의 기술은 dsDNA 절단이 돌연변이 또는 삽입 부위 100 bp 이내에서 이루어져야 할 것을 필요로 하는데, 이는 잠재적으로 가용한 sgRNA 부위를 제한한다 (문헌 [Elliott, B., Richardson, C., Winderbaum, J., Nickoloff, J.A. & Jasin, M. *Mol. Cell. Biol.* 18, 93-101 (1998)], [Yang, L. *et al. Nucleic Acids Research* (2013). doi:10.1093/nar/gkt555]). 소정 측면에서는, 유전자 부근 또는 인트론 내에 위치하는 더 넓은 범위의 뉴클레아제 부위들을 사용하여 대형 유전자 대체가 이루어질 수 있다. 추가적인 고유 sgRNA 부위들은 유전자 계열 내의 보존되는 코딩 서열을 회피함으로써, 본 발명의 방법 내에서 유용하다. 이는 특정 영역에 대한 다수 sgRNA의 시험을 용이하게 한다 (도 2(A)-(E)). 본원에서 기술되는 방법에 따르면, 2 kb까지의 측접 상동성 아암이 유전자 표적화 효율을 향상시킨다 (도 7(a)-(b)).

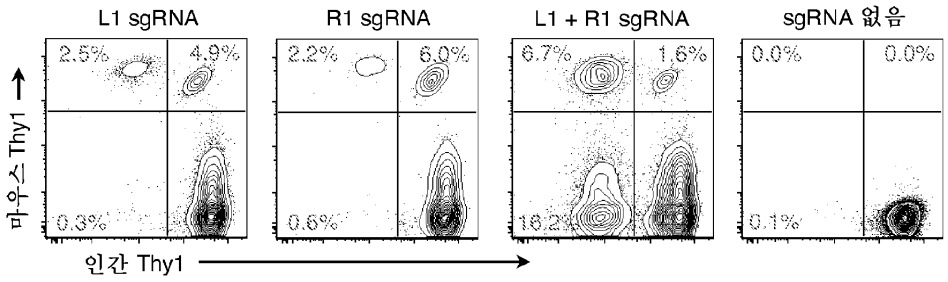
[0116] 소정 측면에서는, 게놈 편집에 ZFN, TALEN 및 CRISPR/Cas 뉴클레아제와 같이 뉴클레아제 활성을 가지는 DNA 결합 단백질을 사용하는 표적화된 유전자 대체물의 제조 방법 (유전자 붕괴 또는 삽입과 대조적임)이 제공된다. 소정 측면에서는, 본원에서 기술되는 유전자 대체의 높은 표적화 효율로 인하여 선택 없이 스크리닝하는 것에 의한, 표적화된 클론의 단리 방법이 제공된다. 소정 측면에 따르면, 유전자는 형광 단백질로도 대체될 수 있으며, 그에 따라 성공적으로 표적화된 세포는 FACS에 의해 선택 및 클로닝될 수 있다. 이종유래 서열을 사용한 유전자 대체는 인간 세포주에서 "녹-인" 동물 또는 질환 모델을 생성시키는 데에 특히 유익하게 된다. 구체적인 적용분야에는 하기가 포함된다: 내인성 프로모터하에 리포터 구축물을 위치시키는 것; 내인성 유전자를 재코딩된 트랜스진으로 대체하는 것; 또는 상이한 종들에 걸친 비교 유전체학.

도면

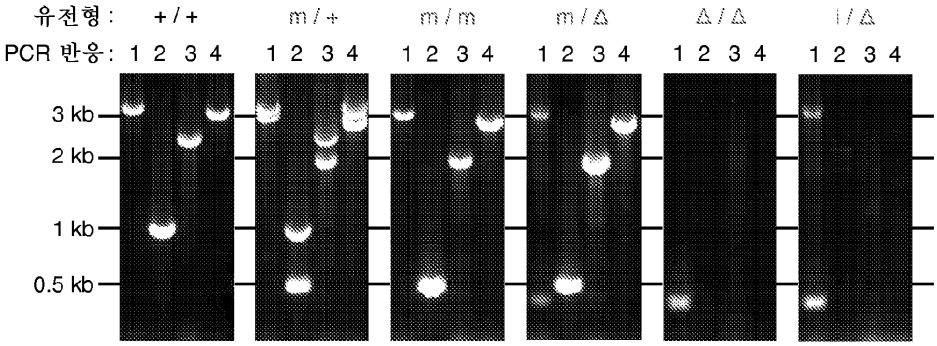
도면1a



도면1b



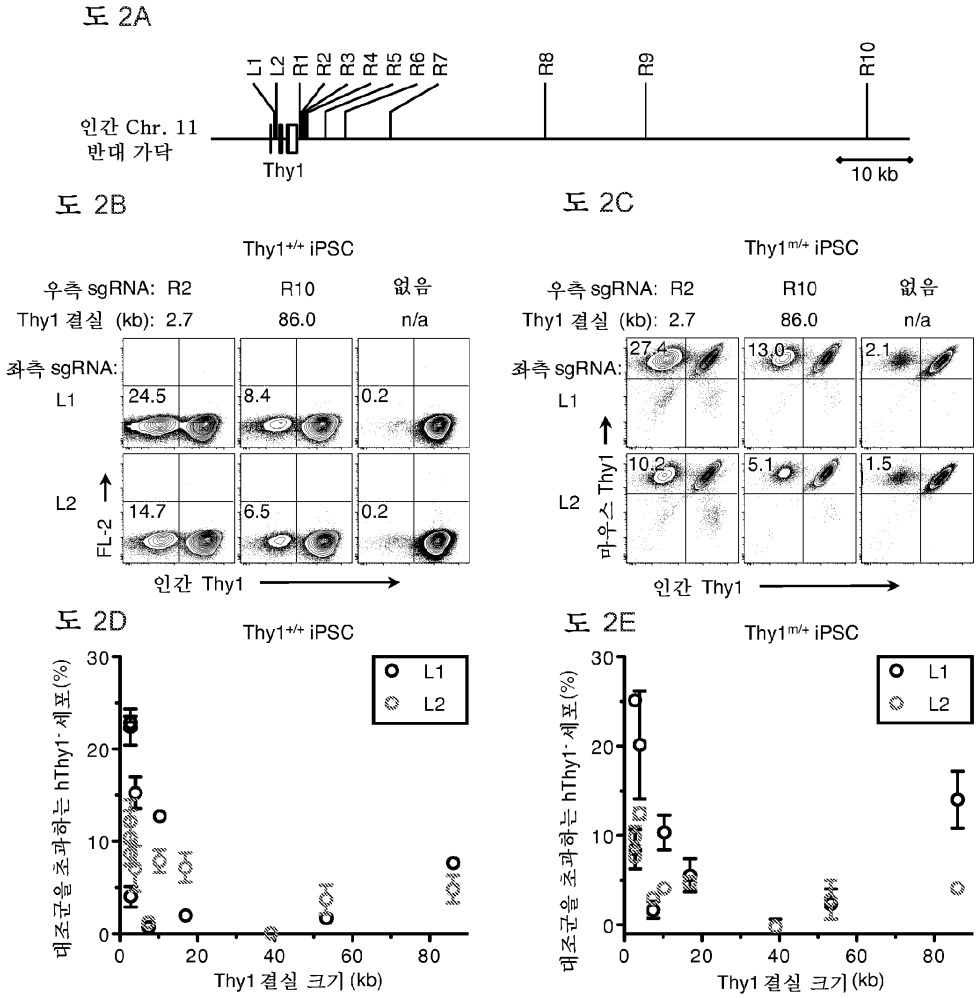
도면1c



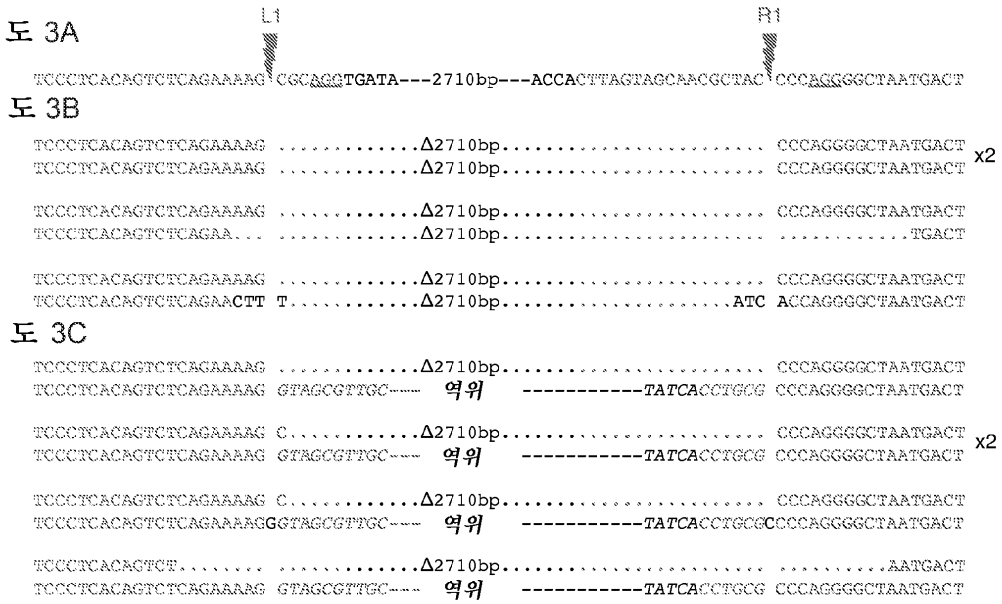
도면1d

샘플	하위세트	샘플의 빈도	분류된 집락의 유전형
L1 sgRNA	mThy1⁺ hThy1⁻ mThy1⁺ hThy1⁺ mThy1⁻ hThy1⁺	2.5 % 4.9 % 92.4 %	8 / 8 = m/m 9 / 11 = m/+ 5 / 6 = +/+
R1 sgRNA	mThy1⁺ hThy1⁻ mThy1⁺ hThy1⁺ mThy1⁻ hThy1⁺	2.2 % 6.0 % 91.2 %	16 / 16 = m/m 12 / 13 = m/+ 6 / 6 = +/+
L1 + R1 sgRNA	mThy1⁺ hThy1⁻ mThy1⁺ hThy1⁺ mThy1⁻ hThy1⁻ mThy1⁻ hThy1⁺	6.7 % 1.6 % 16.2 % 75.5 %	17 / 24 = m/m; 7 / 24 = m/Δ 10 / 11 = m/+ 4 / 9 = Δ/Δ; 5 / 9 = I/Δ 7 / 7 = +/+

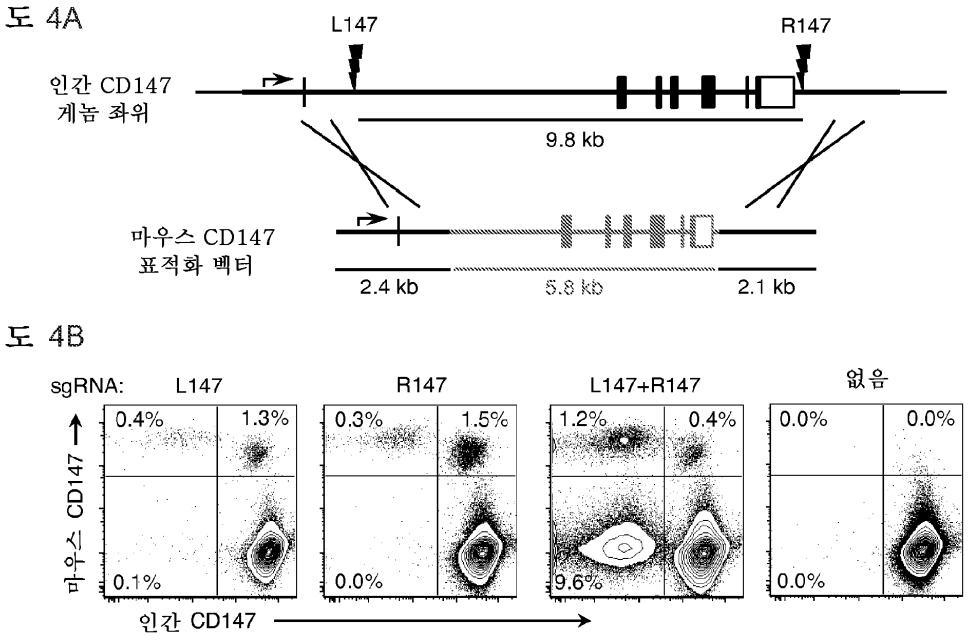
도면2


도 2A
L1
L2
R1
R2
R3
R4
R5
R6
R7
R8
R9
R10
인간 Chr. 11
반대 가닥
Thy1
10 kb
도 2B
Thy1+/+ iPSC
우측 sgRNA: R2
R10
없음
Thy1 결실 (kb): 2.7
86.0
n/a
좌측 sgRNA:
L1
L2
24.5
8.4
0.2
14.7
6.5
0.2
FL-2
인간 Thy1
도 2C
Thy1m/+ iPSC
우측 sgRNA: R2
R10
없음
Thy1 결실 (kb): 2.7
86.0
n/a
좌측 sgRNA:
L1
L2
27.4
13.0
2.1
10.2
5.1
1.5
마우스 Thy1
인간 Thy1
도 2D
Thy1+/+ iPSC
L1
L2
대조군을 초과하는 hThy1- 세포(%)
Thy1 결실 크기 (kb)
도 2E
Thy1m/+ iPSC
L1
L2
대조군을 초과하는 hThy1- 세포(%)
Thy1 결실 크기 (kb)

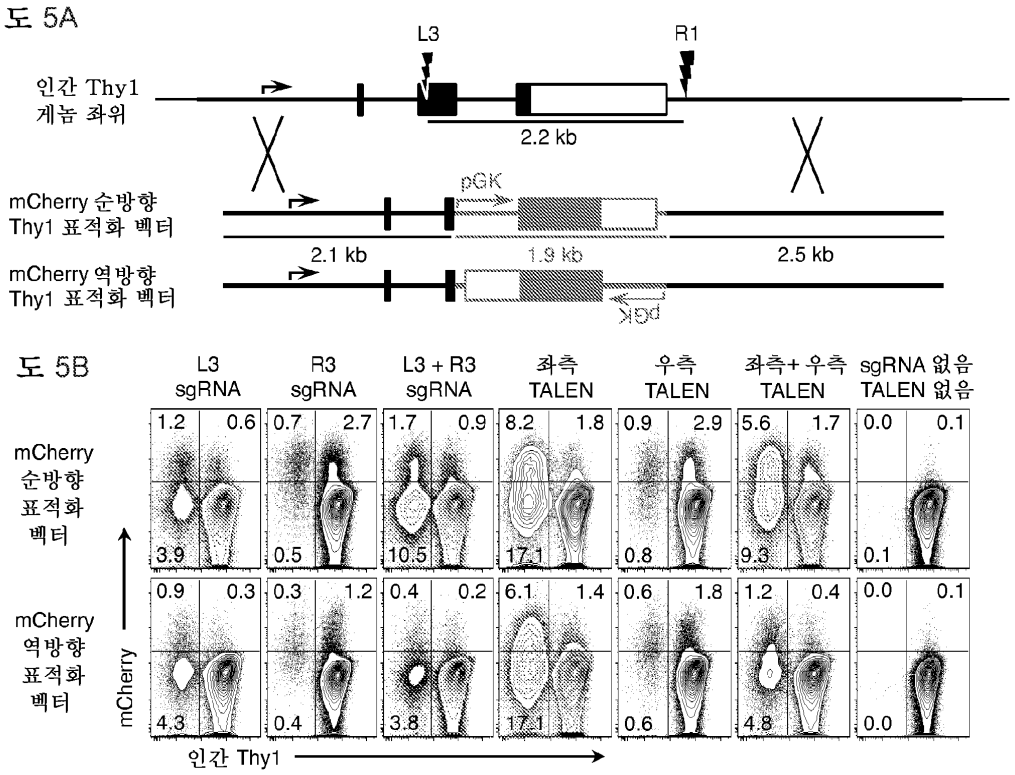
도면3


도 3A
L1
R1
TCCCTCACAGTCTCAGAAAAG CGCACGTGATA---2710bp---ACCACTTAGTAGCAACGCTAC CCCACGGGGCTAATGACT
도 3B
TCCCTCACAGTCTCAGAAAAG .....Δ2710bp..... CCCAGGGGCTAATGACT
TCCCTCACAGTCTCAGAAAAG .....Δ2710bp..... CCCAGGGGCTAATGACT
x2
TCCCTCACAGTCTCAGAAAAG .....Δ2710bp..... CCCAGGGGCTAATGACT
TCCCTCACAGTCTCAGAA... .....Δ2710bp..... .....TGACT
TCCCTCACAGTCTCAGAAAAG .....Δ2710bp..... CCCAGGGGCTAATGACT
TCCCTCACAGTCTCAGAACTT T.....Δ2710bp.....ATC ACCAGGGGCTAATGACT
도 3C
TCCCTCACAGTCTCAGAAAAG .....Δ2710bp..... CCCAGGGGCTAATGACT
TCCCTCACAGTCTCAGAAAAG GTAGCGTTGC---- 역위 ----------TATCACCTGCG CCCAGGGGCTAATGACT
TCCCTCACAGTCTCAGAAAAG C.....Δ2710bp..... CCCAGGGGCTAATGACT
x2
TCCCTCACAGTCTCAGAAAAG GTAGCGTTGC---- 역위 ----------TATCACCTGCG CCCAGGGGCTAATGACT
TCCCTCACAGTCTCAGAAAAG C.....Δ2710bp..... CCCAGGGGCTAATGACT
TCCCTCACAGTCTCAGAAAAGGGTAGCGTTGC---- 역위 ----------TATCACCTGCGCCCCAGGGGCTAATGACT
TCCCTCACAGTCT..... .....Δ2710bp..... .....AATGACT
TCCCTCACAGTCTCAGAAAAG GTAGCGTTGC---- 역위 ----------TATCACCTGCG CCCAGGGGCTAATGACT

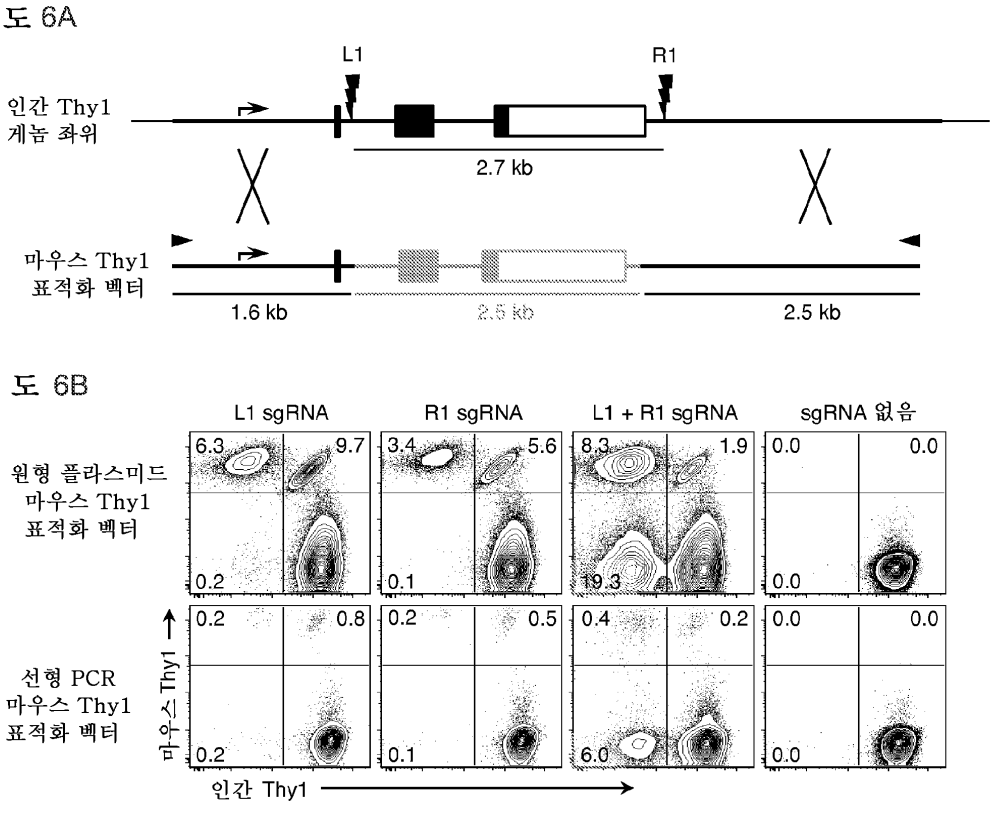
도면4


도 4A
L147
R147
인간 CD147
게놈 좌위
9.8 kb
마우스 CD147
표적화 벡터
2.4 kb
5.8 kb
2.1 kb
도 4B
sgRNA:
L147
R147
L147+R147
없음
0.4%
1.3%
0.1%
0.3%
1.5%
0.0%
1.2%
0.4%
9.6%
0.0%
0.0%
0.0%
마우스 CD147
인간 CD147

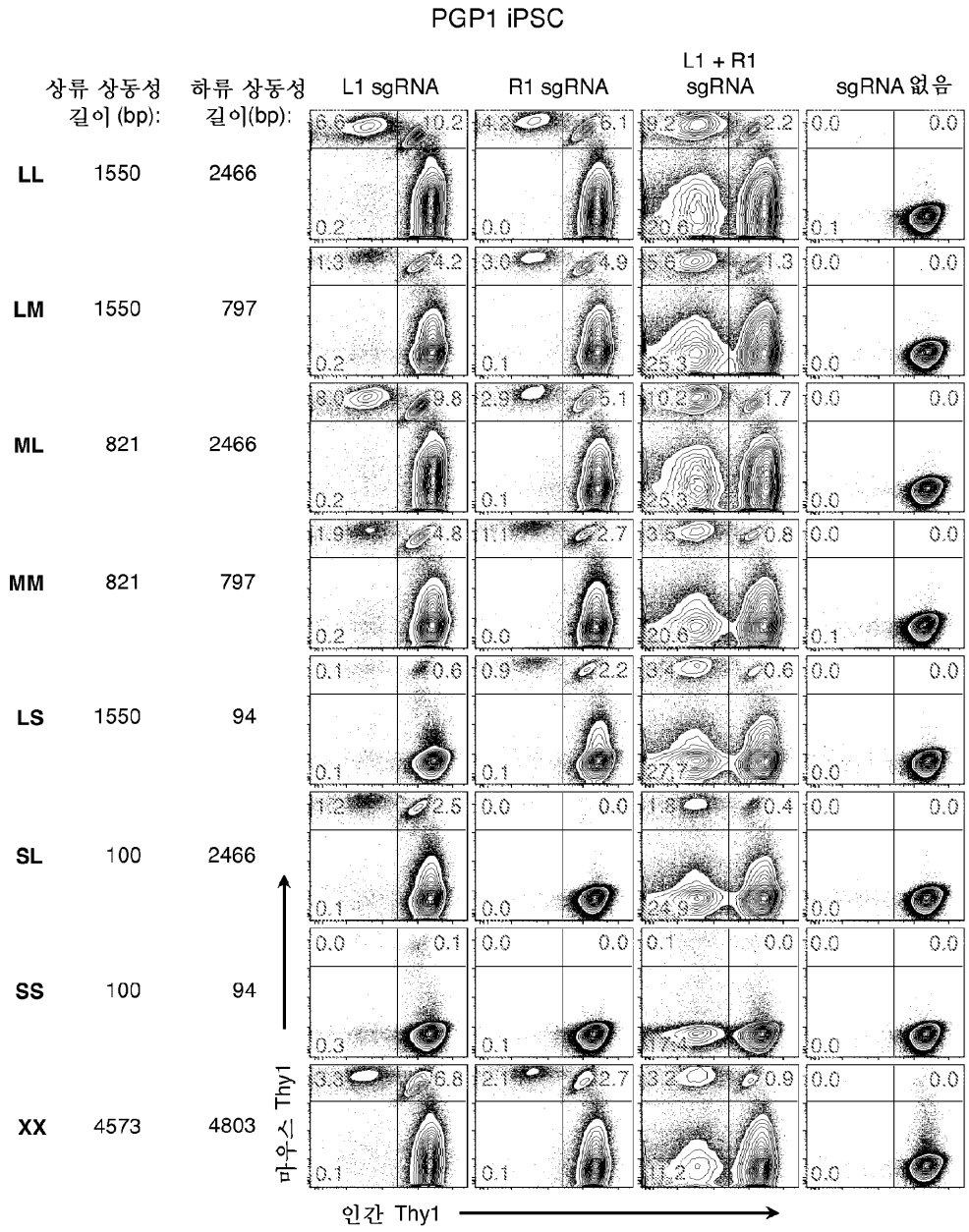
도면5


도 5A
인간 Thy1
게놈 좌위
L3
R1
2.2 kb
mCherry 순방향
Thy1 표적화 벡터
pGK
2.1 kb
1.9 kb
2.5 kb
mCherry 역방향
Thy1 표적화 벡터
pGK
도 5B
L3 sgRNA
R3 sgRNA
L3 + R3 sgRNA
좌측 TALEN
우측 TALEN
좌측+ 우측 TALEN
sgRNA 없음 TALEN 없음
mCherry 순방향 표적화 벡터
mCherry 역방향 표적화 벡터
mCherry
인간 Thy1

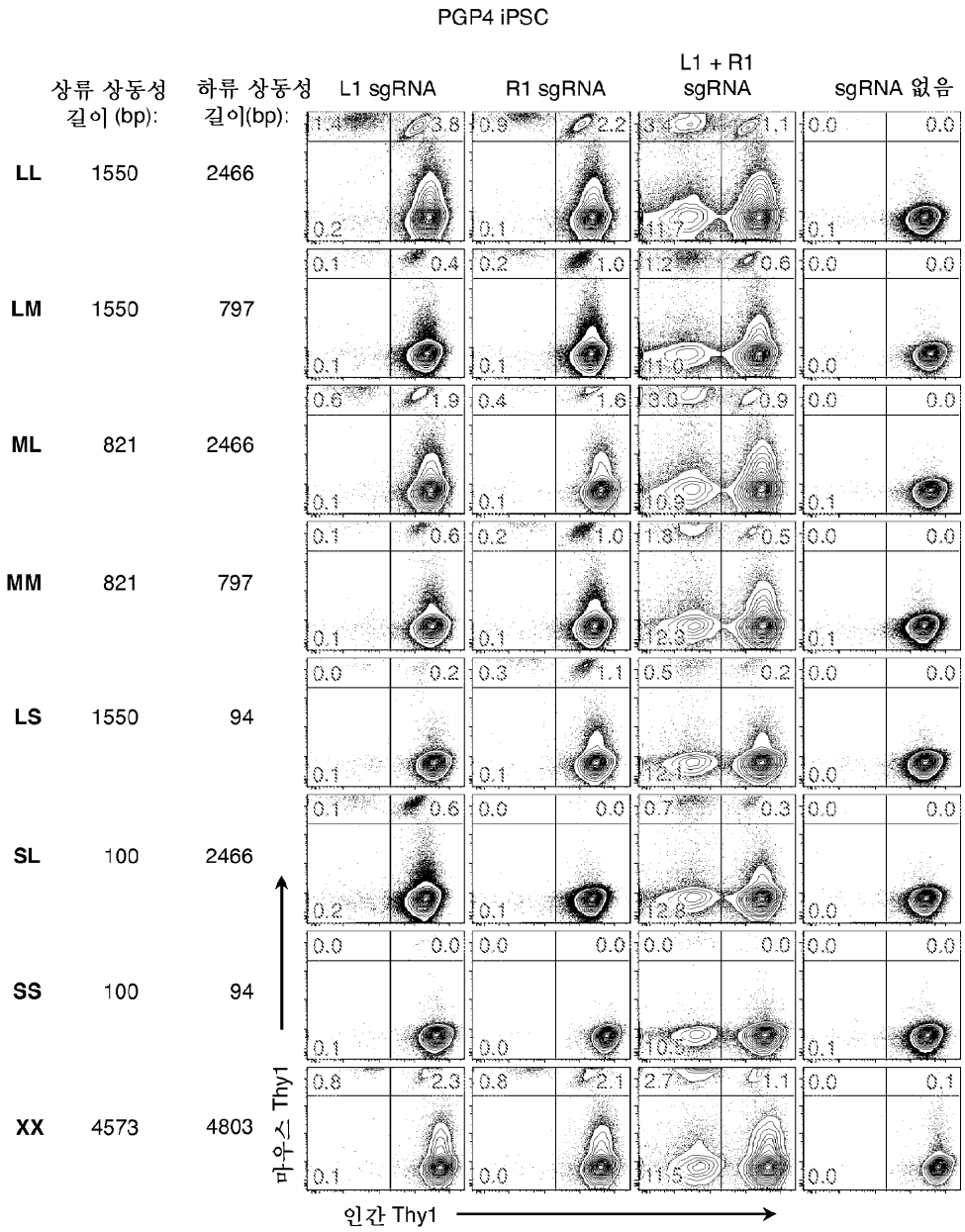
도면6


도 6A
인간 Thy1
게놈 좌위
L1
R1
2.7 kb
마우스 Thy1
표적화 벡터
1.6 kb
2.5 kb
2.5 kb
도 6B
L1 sgRNA
R1 sgRNA
L1 + R1 sgRNA
sgRNA 없음
원형 플라스미드
마우스 Thy1
표적화 벡터
선형 PCR
마우스 Thy1
표적화 벡터
마우스 Thy1
인간 Thy1

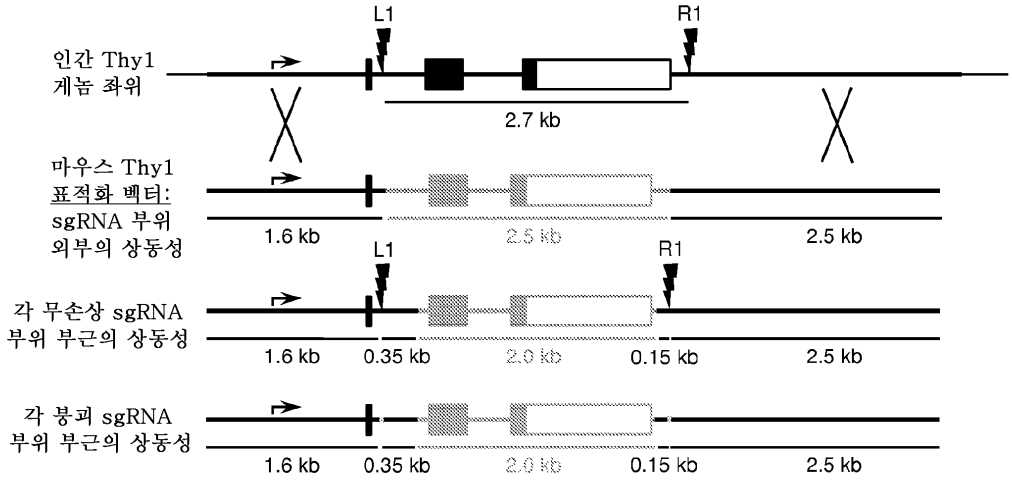
도면7a


PGP1 iPSC
상류 상동성 길이 (bp):
하류 상동성 길이(bp):
L1 sgRNA
R1 sgRNA
L1 + R1 sgRNA
sgRNA 없음
LL 1550 2466
LM 1550 797
ML 821 2466
MM 821 797
LS 1550 94
SL 100 2466
SS 100 94
XX 4573 4803
마우스 Thy1
인간 Thy1

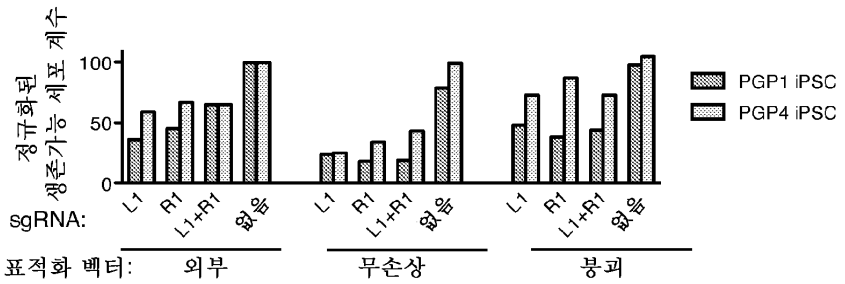
도면7b


PGP4 iPSC
상류 상동성 길이 (bp):
하류 상동성 길이(bp):
L1 sgRNA
R1 sgRNA
L1 + R1 sgRNA
sgRNA 없음
LL 1550 2466
LM 1550 797
ML 821 2466
MM 821 797
LS 1550 94
SL 100 2466
SS 100 94
XX 4573 4803
마우스 Thy1
인간 Thy1

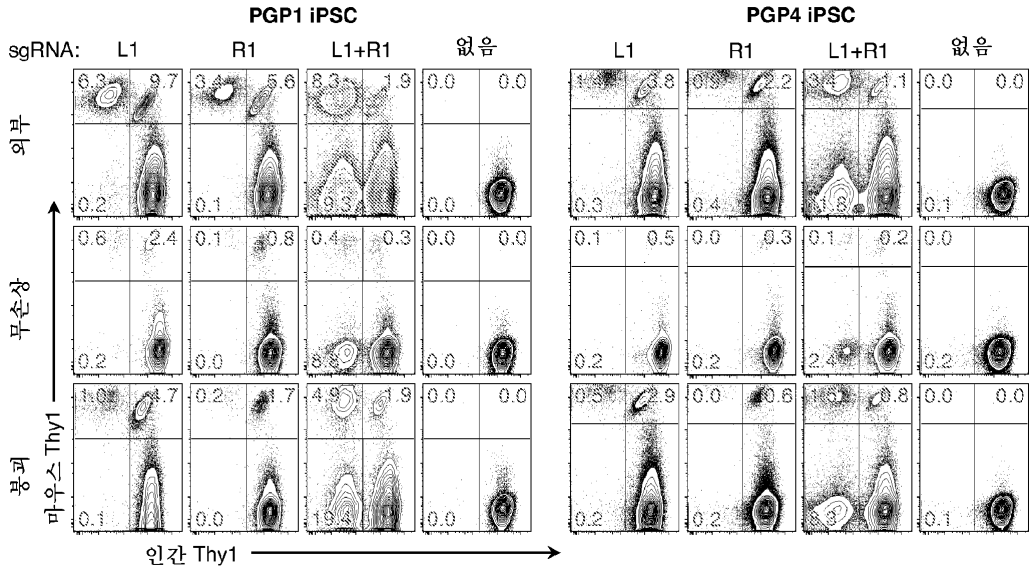
도면8a



도면8b



도면8c



서 열 목 록

SEQUENCE LISTING

<110> Byrne, Susan

Church, George

<120> Large Gene Excision and Insertion

<130> 010498.00552

<140> PCT/US2014/066324

<141> 2014-06-30

<150> US 61/906,188

<151> 2013-11-19

<160> 120

<170> PatentIn version 3.5

<210> 1

<211> 1368

<212> PRT

<213> Streptococcus pyogenes

<400> 1

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 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 His Ile Val Pro Gln Ser Phe Leu Lys

835 840 845

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

850 855 860

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

865 870 875 880

Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys
885 890 895
Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp
900 905 910
Lys Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr

915 920 925
Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp
930 935 940
Glu Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser
945 950 955 960
Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg
965 970 975
Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val

980 985 990
Val Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe
995 1000 1005
Val Tyr Gly Asp Tyr Lys Val Tyr Asp Val Arg Lys Met Ile Ala
1010 1015 1020
Lys Ser Glu Gln Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe
1025 1030 1035
Tyr Ser Asn Ile Met Asn Phe Phe Lys Thr Glu Ile Thr Leu Ala

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

1100 1105 1110
Arg Asn Ser Asp Lys Leu Ile Ala Arg Lys Lys Asp Trp Asp Pro

1115 1120 1125

Lys Lys Tyr Gly Gly Phe Asp Ser Pro Thr Val Ala Tyr Ser Val

1130 1135 1140

Leu Val Val Ala Lys Val Glu Lys Gly Lys Ser Lys Lys Leu Lys

1145 1150 1155

Ser Val Lys Glu Leu Leu Gly Ile Thr Ile Met Glu Arg Ser Ser

1160 1165 1170

Phe Glu Lys Asn Pro Ile Asp Phe Leu Glu Ala Lys Gly Tyr Lys

1175 1180 1185

Glu Val Lys Lys Asp Leu Ile Ile Lys Leu Pro Lys Tyr Ser Leu

1190 1195 1200

Phe Glu Leu Glu Asn Gly Arg Lys Arg Met Leu Ala Ser Ala Gly

1205 1210 1215

Glu Leu Gln Lys Gly Asn Glu Leu Ala Leu Pro Ser Lys Tyr Val

1220 1225 1230

Asn Phe Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu Lys Gly Ser

1235 1240 1245

Pro Glu Asp Asn Glu Gln Lys Gln Leu Phe Val Glu Gln His Lys

1250 1255 1260

His Tyr Leu Asp Glu Ile Ile Glu Gln Ile Ser Glu Phe Ser Lys

1265 1270 1275

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

1280 1285 1290

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

1295 1300 1305

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

1310 1315 1320

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

1325 1330 1335

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

1340 1345 1350

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

1355 1360 1365

<210> 2

<211> 22

<212> DNA

<213> Artificial

<220><223> Target oligonucleotide

<400> 2

cacagtctca gaaaagcgca gg 22

<210> 3

<211> 22

<212> DNA

<213> Artificial

<220><223> Target oligonucleotide

<400> 3

aaatatcagc gcggtggatt gg 22

<210> 4

<211> 22

<212> DNA

<213> Artificial

<220><223> Target oligonucleotide

<400> 4

ggtcaggctg aactcgtact gg 22

<210> 5

<211> 22

<212> DNA

<213> Artificial

<220><223> Target oligonucleotide

<400> 5

ttagtagcaa cgctacccca gg 22

<210> 6

<211> 22

<212> DNA

<213> Artificial

<220><223> Target oligonucleotide

<400> 6

gtgtgcagtc attagcccct gg 22

<210> 7

<211> 22

<212> DNA

<213> Artificial

<220><223> Target oligonucleotide

<400> 7

gggcaaatgt gtctcgttag gg 22

<210> 8

<211> 22

<212> DNA

<213> Artificial

<220><223> Target oligonucleotide

<400> 8

ttctcctttc cgaagtccgt gg 22

<210> 9

<211> 22

<212> DNA

<213> Artificial

<220><223> Target oligonucleotide

<400> 9

gccgctgtcg cctggcaaaa gg 22

<210> 10

<211> 22

<212> DNA

<213> Artificial

<220><223> Target oligonucleotide

<400> 10

gatggtagac atcgaccatg gg 22

<210> 11

<211> 22

<212> DNA

<213> Artificial

<220><223> Target oligonucleotide

<400> 11

ttcaatttcg ggcccgatct gg 22

<210> 12

<211> 22

<212> DNA

<213> Artificial

<220><223> Target oligonucleotide

<400> 12

tgagtcgcgt cacggctatt gg 22

<210> 13

<211> 22

<212> DNA

<213> Artificial

<220><223> Target oligonucleotide

<400> 13

catttgcggt ggtaatcgca gg 22

<210> 14

<211> 22

<212> DNA

<213> Artificial

<220><223> Target oligonucleotide

<400> 14

gatcggatcg ggtcgcgtcg gg 22

<210> 15

<211> 22

<212> DNA

<213> Artificial

<220><223> Target oligonucleotide

<400> 15

tttcctgcgc tgaatcgggt gg 22

<210> 16

<211> 22

<212> DNA

<213> Artificial

<220><223> Target oligonucleotide

<400> 16

ggctcctgtc tgtgcctgac gg 22

<210> 17

<211> 47

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 17

ctttcttgta caaagttggc attattagac gtcaggtggc acttttc 47

<210> 18

<211> 47

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 18

cctttaaagc ctgctttttt gtacagtttg cgtattgggc gctcttc 47

<210> 19

<211> 24

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 19

gccaccatgg acaagaagta ctcc 24

<210> 20

<211> 24

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 20

tcacaccttc ctcttcttct tggg 24

<210> 21

<211> 48

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 21

ccgaaaagtg ccacctgacg tcgacggatg aaaggagtgg gaattggc 48

<210> 22

<211> 50

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 22

ggagtacttc ttgtccatgg tggcggccaa ctagccagct tgggtctccc 50

<210> 23

<211> 54

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 23

gctgacccca agaagaagag gaaggtgtga catcaccatt gagtttaaac ccgc 54

<210> 24

<211> 45

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 24

ccaagctcta gctagaggtc gacggtatcg agccccagct ggttc 45

<210> 25

<211> 21

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 25

ataccgtcga cctctagcta g 21

<210> 26

<211> 18

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 26

tccgtcgacg tcaggtgg 18

<210> 27

<211> 23

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 27

tggtggtggt tgtggtacac acc 23

<210> 28

<211> 21

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 28

aatagggagg gccggggtac c 21

<210> 29

<211> 26

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 29

acccttcccc tcctagatcc caagcc 26

<210> 30

<211> 52

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 30

gattaaaggt gtgtaccaca accaccacca cttttctgag actgtgaggg ag 52

<210> 31

<211> 49

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 31

agactctggg gtaccccggc cctccctatt cccaggggct aatgactgc 49

<210> 32

<211> 20

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 32

gcacctccag ccatcacagc 20

<210> 33

<211> 53

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 33

ccaggaaggg gctgtgatgg ctggaggtgc ttagacgtca ggtggcactt ttc 53

<210> 34

<211> 47

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 34

gggcttggga tctaggaggg gaagggtttg cgtattgggc gctcttc 47

<210> 35

<211> 26

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 35

acccttcccc tcctagatcc caagcc 26

<210> 36

<211> 20

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 36

gcacctccag ccatcacagc 20

<210> 37

<211> 26

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 37

acccttcccc tcctagatcc caagcc 26

<210> 38

<211> 33

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 38

aagattcaga gagattcatt cattcattca caa 33

<210> 39

<211> 23

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 39

cctgctaaca ggtacccggc atg 23

<210> 40

<211> 20

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 40

gcacctccag ccatcacagc 20

<210> 41

<211> 25

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 41

cagcatcttg ctaaggggtt gtcag 25

<210> 42

<211> 25

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 42

gtcagcagac atgggatgtt cgttt 25

<210> 43

<211> 47

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 43

gggcttggga tctaggaggg gaagggtttg cgtattgggc gctcttc 47

<210> 44

<211> 52

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 44

tgaatgaatg aatgaatctc tctgaatctt gtttgcgtat tgggcgctct tc 52

<210> 45

<211> 52

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 45

tcctgcccca tgccgggtac ctgttagcag gtttgcgtat tgggcgctct tc 52

<210> 46

<211> 53

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 46

ccaggaaggg gctgtgatgg ctggaggtgc ttagacgtca ggtggcactt ttc 53

<210> 47

<211> 53

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 47

ggaggctgac aacccctttag caagatgctg ttagacgtca ggtggcactt ttc 53

<210> 48

<211> 53

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 48

caaataaacg aacatcccat gtctgctgac ttagacgtca ggtggcactt ttc 53

<210> 49

<211> 24

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 49

ccttcccctc ctagatccca agcc 24

<210> 50

<211> 20

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 50

gcacctccag ccatcacagc 20

<210> 51

<211> 23

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 51

tcttgtttga gatgttgtgc ggg 23

<210> 52

<211> 25

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 52

ctggtttcag cactccgatc ctatc 25

<210> 53

<211> 20

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 53

tgtggctctg caccaggaag 20

<210> 54

<211> 23

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 54

cctctccctt ttccctggtt ttg 23

<210> 55

<211> 53

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 55

tactctgcaa aaccagggaa aagggagagg ttagacgtca ggtggcactt ttc 53

<210> 56

<211> 52

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 56

ctgtggatag gatcggagtg ctgaaaccag gtttgcgtat tgggcgctct tc 52

<210> 57

<211> 54

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 57

atctctccac ttcaggtggg tgggaggccc ctgtggtctg tgtctcccca aatt 54

<210> 58

<211> 56

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 58

caggtggaca ggaggacaga ttccagaggc ttggttttat tgtgcagttt tctttc 56

<210> 59

<211> 19

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 59

ggcttccttc cctccagag 19

<210> 60

<211> 17

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 60

acaggggcct cccaccc 17

<210> 61

<211> 22

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 61

caagcctctg gaatctgtcc tc 22

<210> 62

<211> 21

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 62

gcccagtgtg cagtcattag c 21

<210> 63

<211> 22

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 63

cttttctgag actgtgaggg ag 22

<210> 64

<211> 24

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 64

taccccaggg gctaatgact gcac 24

<210> 65

<211> 28

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 65

acagtctcag aaaacgcagg tgacaaag 28

<210> 66

<211> 30

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 66

cattagcccc tgggtagcgt tgctactaag 30

<210> 67

<211> 29

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 67

ttgtcacctg cgttttctga gactgtgag 29

<210> 68

<211> 30

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 68

cttagtagca acgctaccca ggggctaatg 30

<210> 69

<211> 52

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 69

gagaatacca gcagttcacc catccagtac gaaattctac cgggtagggg ag 52

<210> 70

<211> 54

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 70

cccagtgtgc agtcattagc ccctggggta cgacggccag tgaattgtaa tacg 54

<210> 71

<211> 26

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 71

acccttcccc tcctagatcc caagcc 26

<210> 72

<211> 28

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 72

gtactggatg ggtgaactgc tggtattc 28

<210> 73

<211> 24

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 73

taccccaggg gctaatgact gcac 24

<210> 74

<211> 20

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 74

gcacctccag ccatcacagc 20

<210> 75

<211> 53

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 75

ccaggaaggg gctgtgatgg ctggaggtgc ttagacgtca ggtggcactt ttc 53

<210> 76

<211> 47

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 76

gggcttggga tctaggaggg gaagggtttg cgtattgggc gctcttc 47

<210> 77

<211> 69

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 77

gaagtcgagg ttccaaggtc acagtgaggg ggccctggcc acccccttgca ggttctccat 60

agtccacag 69

<210> 78

<211> 25

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 78

caacaacccc tcctgtatat gacct 25

<210> 79

<211> 25

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 79

acacactttc aacctccaag agacg 25

<210> 80

<211> 23

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 80

ctcactgtga ccttggaacc tcg 23

<210> 81

<211> 51

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 81

tgttgaggtc atatacagga ggggttgttg cctgacgggg ttgggttttc c 51

<210> 82

<211> 23

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 82

aagggagccc tgaggccttt tcc 23

<210> 83

<211> 47

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 83

tcaggaaaag gcctcagggc tcccttagac gtcaggtggc acttttc 47

<210> 84

<211> 47

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 84

cgtctcttgg aggttgaaag tgtgtgtttg cgtattgggc gctcttc 47

<210> 85

<211> 18

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 85

taccagcagt tcacccat 18

<210> 86

<211> 18

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 86

tctttgtctc acgggtca 18

<210> 87

<211> 18

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 87

tctccccaac cacttagt 18

<210> 88

<211> 18

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 88

tgtgcagtca ttagcccc 18

<210> 89

<211> 25

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 89

ggccgccacc atggattata aggac 25

<210> 90

<211> 20

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 90

ggaacctgcc actcgatgtg 20

<210> 91

<211> 44

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 91

gacatcacat cgagtggcag gttcccagct ggtgaagtcc gagc 44

<210> 92

<211> 63

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 92

caactagaag gcacagtcga ggctgatcag cggggttaga aattgatttc accattgttg 60

aac 63

<210> 93

<211>

50

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 93

gacatcacat cgagtggcag gttcccaact agtcaaaagt gaactggagg 50

<210> 94

<211> 54

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 94

caactagaag gcacagtcga ggctgatcag cgcccttaaa agtttatctc gccg 54

<210> 95

<211> 23

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 95

gcctcgactg tgccttctag ttg 23

<210> 96

<211>

22

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 96

cttataatcc atggtggcgg cc 22

<210> 97

<211> 25

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 97

agggacttag atgactgcca tagcc 25

<210> 98

<211> 25

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 98

atgttggcag taagcatgtt gttcc 25

<210> 99

<211>

25

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 99

agggacttag atgactgcca tagcc 25

<210> 100

<211> 25

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 100

ctcacctctg agcactgtga cgttc 25

<210> 101

<211> 25

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 101

actgaagttc tgggtcccaa caatg 25

<210> 102

<211> 25

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 102

atgaatacag actgcacctc cccag 25

<210> 103

<211> 25

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 103

ctcacctctg agcactgtga cgttc 25

<210> 104

<211> 35

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 104

ccatcaatct actgaagttc tgggtcccaa caatg 35

<210> 105

<211> 25

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 105

tgaagtgaaa ccctaaaggg ggaag 25

<210> 106

<211> 25

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 106

aaaccacaca cttcaacctg gatgg 25

<210> 107

<211> 25

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 107

gtttggccca agtttctaag ggagg 25

<210> 108

<211> 70

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 108

tccctcacag tctcagaaaa gcgcaggtga taaccactta gtagcaacgc taccccaggg 60

gctaatgact 70

<210> 109

<211> 38

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 109

tccctcacag tctcagaaaa gcccaggggc taatgact 38

<210> 110

<211> 23

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 110

tccctcacag tctcagaatg act 23

<210> 111

<211> 38

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 111

tccctcacag tctcagaaaa gcccaggggc taatgact 38

<210> 112

<211> 42

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 112

tccctcacag tctcagaact ttatcaccag gggctaatga ct 42

<210> 113

<211> 38

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 113

tccctcacag tctcagaaaa gcccaggggc taatgact 38

<210> 114

<211> 59

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 114

tccctcacag tctcagaaaa ggtagcgttg ctatcacctg cgcccagggg ctaatgact 59

<210> 115

<211> 39

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 115

tccctcacag tctcagaaaa gccccagggg ctaatgact 39

<210> 116

<211> 59

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 116

tccctcacag tctcagaaaa ggtagcgttg ctatcacctg cgcccagggg ctaatgact 59

<210> 117

<211> 39

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 117

tccctcacag tctcagaaaa gccccagggg ctaatgact 39

<210> 118

<211> 61

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 118

tccctcacag tctcagaaaa gggtagcgtt gctatcacct gcgccccagg ggctaatgac 60

t 61

<210> 119

<211> 20

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 119

tccctcacag tctaatgact 20

<210> 120

<211> 59

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 120

tccctcacag tctcagaaaa ggtagcgttg ctatcacctg cgcccagggg ctaatgact 59