



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2025년02월26일
(11) 등록번호 10-2773555
(24) 등록일자 2025년02월21일

(51) 국제특허분류(Int. Cl.)
C12N 15/85 (2006.01) C12N 15/113 (2010.01)
C12N 15/52 (2006.01) C12N 15/90 (2006.01)
C12N 9/16 (2006.01) C12N 9/22 (2006.01)
(52) CPC특허분류
C12N 15/85 (2013.01)
C12N 15/113 (2013.01)
(21) 출원번호 10-2016-7004663
(22) 출원일자(국제) 2014년07월25일
심사청구일자 2019년03월15일
(85) 번역문제출일자 2016년02월23일
(65) 공개번호 10-2016-0035017
(43) 공개일자 2016년03월30일
(86) 국제출원번호 PCT/US2014/048140
(87) 국제공개번호 WO 2015/013583
국제공개일자 2015년01월29일
(30) 우선권주장
61/858,866 2013년07월26일 미국(US)
(56) 선행기술조사문헌
KR1020150105634 A*
(뒷면에 계속)
전체 청구항 수 : 총 28 항

(73) 특허권자
프레지던트 앤드 펠로우즈 오브 하바드 칼리지
미합중국, 메사추세츠 02138, 캠브리지, 퀸시스트리트17
(72) 발명자
처치, 조지 엠.
미국 02446 매사추세츠주 브루클린 켄트 스트리트 218
양, 루한
미국 02115 매사추세츠주 보스턴 애비뉴 데 루이 파스퇴르 107
(뒷면에 계속)
(74) 대리인
양영준, 김영

심사관 : 김승범

(54) 발명의 명칭 **게놈 조작**

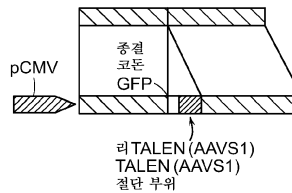
(57) 요약

반복부 서열이 결핍된 TALEN 또는 Cas9를 이용하여 세포에서 게놈을 조작하는 방법을 제공한다.

대표도 - 도1a

IGFP 도너

염색체 리포터



(52) CPC특허분류

C12N 15/52 (2013.01)
C12N 15/907 (2013.01)
C12N 9/16 (2013.01)
C12N 9/22 (2013.01)

(72) 발명자

카폴, 마크 구엘

미국 02114 매사추세츠주 보스턴 아파트먼트 7 사
우스 러셀 스트리트 58

양, 조이스 리치

미국 02139 매사추세츠주 캠브리지 퍼시픽 스트리
트 70

(56) 선행기술조사문헌

KR1020150107739 A
Science, 2013.01., Vol. 339, pp. 823-826.*
Nature Methods volume 6, pages363-369 (2009)*
JP2003522536 A*
Current Gene Therapy (2013.04) 13(2):73-92
Nature (2011) 487:391-394
Disease Models& Mechanisms (2009) 2:446-453
Cell (2011) 146:318-331
BMC Medicine (2010) 8:42
PNAS (2000) 97(4):1495-1500
The Journal of Biological Chemistry (2000)
275(43):33850-33860
Cell Stem Cell (2013.04) 12:393-394
Science (2013.01) 339:819-823
Nature Biotechnology (2013.01) 31:230-232
*는 심사관에 의하여 인용된 문헌

명세서

청구범위

청구항 1

세포에서 표적 DNA를 변경시키는 시험관내 또는 생체의 방법이며, 상기 세포는 표적 DNA에 상보적인 가이드 RNA와 공동 국제화 복합체를 형성하고 부위 특이적인 방식으로 표적 DNA를 절단하는 Cas9 효소를 코딩하는 핵산을 포함하도록 유전적으로 변형되고 상기 핵산이 세포의 게놈 내로 통합된 것이며, 상기 방법은

(a) 상기 세포 내로, 표적 DNA에 상보적이고 Cas9 효소를 표적 DNA로 가이드하는 가이드 RNA를 코딩하는 제1 외래 핵산을 도입하는 단계로서, 여기서 가이드 RNA 및 Cas9 효소는 표적 DNA에 대한 공동 국제화 복합체의 구성 원인 단계, 및

(b) 세포 내로 도너 핵산 서열을 도입하는 단계를 포함하고,

여기서 가이드 RNA를 코딩하는 제1 외래 핵산이 발현되고,

여기서 가이드 RNA 및 Cas9 효소가 표적 DNA로 공동 국제화되고, Cas9 효소가 표적 DNA를 절단하고, 도너 핵산이 표적 DNA 내로 삽입되어 세포에서 변경된 DNA를 생성하고,

단계 (a) 및 (b)를 반복하여 복수 개의 도너 핵산 서열 삽입을 야기하는 것인, 시험관내 또는 생체의 방법.

청구항 2

제1항에 있어서, 가이드 RNA가 10 내지 500개의 뉴클레오티드인 시험관내 또는 생체의 방법.

청구항 3

제1항에 있어서, 가이드 RNA가 20 내지 100개의 뉴클레오티드인 시험관내 또는 생체의 방법.

청구항 4

제1항에 있어서, 가이드 RNA가 tracrRNA-crRNA 융합체인 시험관내 또는 생체의 방법.

청구항 5

제1항에 있어서, DNA가 게놈 DNA, 미토콘드리아 DNA, 바이러스 DNA, 또는 외인성 DNA인 시험관내 또는 생체의 방법.

청구항 6

제1항에 있어서, 도너 핵산 서열이 재조합에 의해 삽입되는 것인 시험관내 또는 생체의 방법.

청구항 7

제1항에 있어서, 도너 핵산 서열이 상동성 재조합에 의해 삽입되는 것인 시험관내 또는 생체의 방법.

청구항 8

제1항에 있어서, 도너 핵산 서열이 비상동성 단부 연결에 의해 삽입되는 것인 시험관내 또는 생체의 방법.

청구항 9

제1항에 있어서, 가이드 RNA 및 도너 핵산 서열이 하나 이상의 플라스미드 상에 존재하는 것인 시험관내 또는 생체의 방법.

청구항 10

제1항에 있어서, 세포에서 변경된 DNA를 제조한 후, Cas9 효소를 코딩하는 핵산이 세포 게놈으로부터 제거되는

것인 시험관내 또는 생체의 방법.

청구항 11

제1항에 있어서, 가이드 RNA 및 도너 핵산 서열이 함께 결합된 것인 시험관내 또는 생체의 방법.

청구항 12

표적 DNA에 상보적인 가이드 RNA와 공동 국제화 복합체를 형성하고 부위 특이적인 방식으로 표적 DNA를 절단하는 Cas9 효소를 코딩하는 제1 외래 핵산을 포함하고 상기 핵산이 세포의 게놈으로 통합된 단리된 세포이며, 여기서 단리된 세포는 인간 배아 또는 인간 배아 줄기 세포가 아니고, Cas9 효소를 코딩하는 통합된 제1 외래 핵산을 포함하는 효소 제거가능한 벡터 또는 카세트를 세포의 게놈 DNA 내로 삽입함으로써 유전적으로 변형되고, 여기서 세포는 복수 개의 가이드 RNA 및 복수 개의 도너 핵산 서열을 포함하는 것인 단리된 세포.

청구항 13

제12항에 있어서, 표적 DNA에 상보적이고 Cas9 효소를 표적 DNA로 가이드하는 가이드 RNA를 코딩하는 제2 외래 핵산을 추가로 포함하고, 여기서 가이드 RNA 및 Cas9 효소는 표적 DNA에 대한 공동 국제화 복합체의 구성원인 단리된 세포.

청구항 14

제13항에 있어서, 도너 핵산 서열을 추가로 포함하는 단리된 세포.

청구항 15

제12항에 있어서, Cas9 효소의 발현을 촉진시키기 위한 유도성 프로모터를 추가로 포함하는 단리된 세포.

청구항 16

제12항에 있어서, 제1 외래 핵산이 트랜스포사제를 사용하여 세포의 게놈 DNA로부터 제거가능한 것인 단리된 세포.

청구항 17

제12항에 있어서, 가이드 RNA가 10 내지 500개의 뉴클레오티드인 단리된 세포.

청구항 18

제12항에 있어서, 가이드 RNA가 20 내지 100개의 뉴클레오티드인 단리된 세포.

청구항 19

제12항에 있어서, 가이드 RNA가 tracrRNA-crRNA 융합체인 단리된 세포.

청구항 20

제12항에 있어서, 표적 DNA가 게놈 DNA, 미토콘드리아 DNA, 바이러스 DNA, 또는 외인성 DNA인 단리된 세포.

청구항 21

제12항 내지 제20항 중 어느 한 항에 있어서, 줄기 세포인 단리된 세포.

청구항 22

표적 DNA에 상보적인 가이드 RNA와 공동 국제화 복합체를 형성하고 부위 특이적인 방식으로 표적 DNA를 절단하는 Cas9 효소를 코딩하는 제1 외래 핵산을 포함하고 상기 핵산이 세포의 게놈으로 통합되었으며, Cas9 효소의 발현을 촉진시키기 위한 유도성 프로모터를 포함하는 단리된 세포이며, 여기서 단리된 세포는 인간 배아 또는 인간 배아 줄기 세포가 아니고,

여기서 세포는 복수 개의 가이드 RNA 및 복수 개의 도너 핵산 서열을 포함하는 것인 단리된 세포.

청구항 23

제22항에 있어서, 줄기 세포인 단리된 세포.

청구항 24

표적 DNA에 상보적인 가이드 RNA와 공동 국제화 복합체를 형성하고 부위 특이적인 방식으로 표적 DNA를 절단하는 Cas9 효소를 코딩하는 제1 외래 핵산을 포함하고 상기 핵산이 세포의 게놈 내로 통합된 단리된 세포이며, 여기서 제1 외래 핵산은 트랜스포사제를 사용하여 세포의 게놈 DNA로부터 제거가능한 것이고, 단리된 세포는 인간 배아 또는 인간 배아 줄기 세포가 아니고,

여기서 세포는 복수 개의 가이드 RNA 및 복수 개의 도너 핵산 서열을 포함하는 것인 단리된 세포.

청구항 25

제24항에 있어서, 줄기 세포인 단리된 세포.

청구항 26

표적 DNA에 상보적인 가이드 RNA와 공동 국제화 복합체를 형성하고 부위 특이적인 방식으로 표적 DNA를 절단하는 Cas9 효소를 코딩하는 제1 외래 핵산을 포함하고 상기 핵산이 세포의 게놈 내로 통합된 단리된 세포이며, 여기서 제1 외래 핵산은 세포의 게놈 DNA 내로 가역적으로 삽입되는 것이고, 단리된 세포는 인간 배아 또는 인간 배아 줄기 세포가 아니고,

여기서 세포는 복수 개의 가이드 RNA 및 복수 개의 도너 핵산 서열을 포함하는 것인 단리된 세포.

청구항 27

제26항에 있어서, 줄기 세포인 단리된 세포.

청구항 28

제1항 내지 제11항 중 어느 한 항에 있어서, 세포가 줄기 세포인 방법.

청구항 29

삭제

청구항 30

삭제

청구항 31

삭제

청구항 32

삭제

청구항 33

삭제

청구항 34

삭제

청구항 35

삭제

청구항 36

삭제

청구항 37

삭제

청구항 38

삭제

청구항 39

삭제

청구항 40

삭제

청구항 41

삭제

청구항 42

삭제

청구항 43

삭제

청구항 44

삭제

청구항 45

삭제

청구항 46

삭제

청구항 47

삭제

청구항 48

삭제

청구항 49

삭제

청구항 50

삭제

청구항 51

삭제

청구항 52

삭제

청구항 53

삭제

청구항 54

삭제

청구항 55

삭제

청구항 56

삭제

청구항 57

삭제

청구항 58

삭제

청구항 59

삭제

청구항 60

삭제

청구항 61

삭제

청구항 62

삭제

청구항 63

삭제

청구항 64

삭제

청구항 65

삭제

청구항 66

삭제

청구항 67

삭제

청구항 68

삭제

청구항 69

삭제

청구항 70

삭제

청구항 71

삭제

청구항 72

삭제

청구항 73

삭제

청구항 74

삭제

청구항 75

삭제

청구항 76

삭제

청구항 77

삭제

청구항 78

삭제

청구항 79

삭제

청구항 80

삭제

청구항 81

삭제

청구항 82

삭제

청구항 83

삭제

청구항 84

삭제

청구항 85

삭제

청구항 86

삭제

청구항 87

삭제

청구항 88

삭제

청구항 89

삭제

청구항 90

삭제

청구항 91

삭제

청구항 92

삭제

청구항 93

삭제

청구항 94

삭제

청구항 95

삭제

청구항 96

삭제

청구항 97

삭제

청구항 98

삭제

청구항 99

삭제

청구항 100

삭제

청구항 101

삭제

청구항 102

삭제

청구항 103

삭제

청구항 104

삭제

청구항 105

삭제

청구항 106

삭제

청구항 107

삭제

청구항 108

삭제

청구항 109

삭제

청구항 110

삭제

청구항 111

삭제

청구항 112

삭제

청구항 113

삭제

청구항 114

삭제

청구항 115

삭제

청구항 116

삭제

청구항 117

삭제

청구항 118

삭제

청구항 119

삭제

청구항 120

삭제

청구항 121

삭제

청구항 122

삭제

청구항 123

삭제

청구항 124

삭제

청구항 125

삭제

청구항 126

삭제

청구항 127

삭제

청구항 128

삭제

청구항 129

삭제

청구항 130

삭제

청구항 131

삭제

청구항 132

삭제

청구항 133

삭제

청구항 134

삭제

청구항 135

삭제

청구항 136

삭제

청구항 137

삭제

청구항 138

삭제

청구항 139

삭제

청구항 140

삭제

청구항 141

삭제

청구항 142

삭제

청구항 143

삭제

청구항 144

삭제

청구항 145

삭제

청구항 146

삭제

청구항 147

삭제

청구항 148

삭제

청구항 149

삭제

청구항 150

삭제

청구항 151

삭제

청구항 152

삭제

청구항 153

삭제

청구항 154

삭제

청구항 155

삭제

청구항 156

삭제

청구항 157

삭제

청구항 158

삭제

발명의 설명

기술 분야

[0001] 관련 출원에 관한 데이터

[0002] 본 출원은 2013년 7월 26일 출원된 미국 가특허 출원 번호 61/858,866을 우선권 주장하며, 이 출원은 그 전문이 모든 목적을 위해 본원에서 참조로 포함된다.

[0003] 정부 이익에 관한 언급

[0004] 본 발명은 국립 인간 게놈 연구소 게놈 과학 우수 센터(National Human Genome Research Center for Excellence in Genomics Science)로부터의 P50 HG003170 하의 정부 지원으로 만들어졌다. 정부가 본 발명에 대한 특정 권리를 가진다.

배경 기술

[0005] 서열-특이 뉴클레아제를 통한 게놈 편집은 공지되어 있다. 그의 전문이 본원에 참조로 포함된 참고 문헌 1, 2, 및 3을 참조할 수 있다. 게놈 중 뉴클레아제-매개 이중 가닥 DNA (dsDNA) 절단은, 빈번하게는 비-특이 삽입 및 결실 (indels)의 도입을 결과로서 초래하는 비-상동성 단부 연결 (NHEJ), 또는 수복 주형으로서 상동성 가닥을 도입하는 상동성 지정 수복 (HDR)이라는 2가지 주요 기전에 의해 수복될 수 있다. 그 전문이 본원에 참조로 포

함된 참고 문헌 4를 참조할 수 있다. 서열-특이 뉴클레아제가 원하는 돌연변이를 함유하는 상동성 도너 DNA를 따라 전달될 때, 유전자 표적화 효율은 단지 도너 구축물 단독인 것과 비교하였을 때 1,000배 증가되어 있다. 그 전문이 본원에 참조로 포함된 참고 문헌 5를 참조할 수 있다. DNA 도너로서 단일 가닥 올리고데옥시리보뉴클레오티드 ("ssODN")를 사용하는 것이 보고된 바 있다. 그 전문이 본원에 참조로 포함된 참고 문헌 21 및 22를 참조할 수 있다.

[0006] 유전자 편집 도구가 크게 진보했음에도 불구하고, 인간 유도 다능성 줄기 세포 ("hiPSC") 조작에서 맞춤형 조작된 뉴클레아제 사용과 관련해서는 많은 도전 과제와 의문이 여전히 남아있다. 첫째, 그의 디자인이 단순함에도 불구하고, 전사 활성인자-유사 효과기 뉴클레아제 (TALEN)는 반복 가변 이잔기 (RVD) 도메인의 일렬 카피를 포함하는 특정 DNA 서열을 표적화한다. 그 전문이 본원에 참조로 포함된 참고 문헌 6을 참조할 수 있다. RVD의 모듈식 성질을 통해 TALEN 디자인이 간소화되었지만, 그의 반복 서열은 그의 DNA 구축물을 합성하는 방법을 복잡하게 만들고 (그 전문이 본원에 참조로 포함된 참고 문헌 2, 9, 및 15-19를 참조할 수 있다), 이는 또한 렌티바이러스 유전자 전달 비히클을 이용하는 그의 사용을 손상시킨다. 그 전문이 본원에 참조로 포함된 참고 문헌 13을 참조할 수 있다.

[0007] 현 관행상, NHEJ 및 HDR은 빈번하게는 별개의 검정법을 사용하여 평가된다. NHEJ를 검정하는 데에는 대개 미스매치-고감도 엔도뉴클레아제 검정법 (그 전문이 본원에 참조로 포함된 참고 문헌 14를 참조할 수 있다)이 사용되지만, 상기 방법의 정량적 정확성은 가변적이고, 감도는 ~3% 초과로 NHEJ 빈도에 한정된다. 그 전문이 본원에 참조로 포함된 참고 문헌 15를 참조할 수 있다. HDR은 빈번하게는, 완전히 다르고, 대개는 번잡한 방법인 클로닝 및 시퀀싱에 의해 평가된다. 일부 세포 유형, 예컨대, U2OS 및 K562의 경우에는 대략 50%로 편집 빈도가 높은 것으로 빈번하게 보고되어 있지만 (그 전문이 본원에 참조로 포함된 참고 문헌 12 및 14를 참조할 수 있다), hiPSC에서 빈도는 일반적으로 더 낮기 때문에, 감도가 여전히 문제가 된다. 그 전문이 본원에 참조로 포함된 참고 문헌 10을 참조할 수 있다. 최근, hiPSC 및 hESC에서 TALEN을 사용하는 경우, 편집 빈도가 높은 것으로 보고되었고 (그 전문이 본원에 참조로 포함된 참고 문헌 9를 참조할 수 있다), CRISPR Cas9-gRNA 시스템을 사용하는 경우에는 빈도가 훨씬 더 높은 것으로 보고되었다 (그 전문이 본원에 참조로 포함된 참고 문헌 16-19를 참조할 수 있다). 그러나, 다른 부위에서의 편집률은 매우 다른 것으로 보이며 (그 전문이 본원에 참조로 포함된 참고 문헌 17을 참조할 수 있다), 편집은 종종 일부 부위에서는 전혀 검출되지 않는다 (그 전문이 본원에 참조로 포함된 참고 문헌 20을 참조할 수 있다).

[0008] 박테리아 및 고세균 CRISPR-Cas 시스템은 침입 외래 핵산 내에 존재하는 상보적인 서열의 분해를 지시하는, Cas 단백질과 함께 복합된 짧은 가이드 RNA에 의존한다. 문헌 [Deltcheva, E. et al. CRISPR RNA maturation by trans-encoded small RNA 및 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. *Science* 337, 816-821 (2012)]; [Sapranauskas, R. et al. The *Streptococcus thermophilus* CRISPR/Cas systems 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 참조할 수 있다. 최근 *S. pyogenes* (S. 피오게네스) II형 CRISPR의 시험관내 재구성을 통해 보통 트랜스 코딩된 tracrRNA ("트랜스-활성 CRISPR RNA")에 융합된 crRNA ("CRISPR RNA")가 Cas9 단백질을 지시하여 상기 crRNA에 매칭되는 표적 DNA를 서열 특이적으로 절단하는 데에는 충분하다는 것이 입증되었다. 표적 부위에 상동성인 gRNA가 발현되면, Cas9가 동원되고, 표적 DNA가 분해된다. 문헌 [H. Deveau et al., Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. *Journal of Bacteriology* 190, 1390 (Feb, 2008)]을 참조할 수 있다.

발명의 내용

[0009] 본 개시내용의 측면은 세포, 예컨대, 체세포 또는 줄기 세포를 유전적으로 변형시키기 위한, 변형된 전사 활성인자-유사 효과기 뉴클레아제 (TALEN)의 용도에 관한 것이다. TALEN은 반복부 서열을 포함하는 것으로 알려져 있다. 본 개시내용의 측면은 반복부 서열 100 bp 이상이 결핍된 TALEN을 세포 내로 도입하는 단계를 포함하고, 여기서, 세포에서 TALEN은 표적 DNA를 절단하고, 세포는 비상동성 단부 연결을 거치고, 이로써 변경된 DNA가 생산되는 것인, 세포에서 표적 DNA를 변경시키는 방법에 관한 것이다. 특정 측면에 따라, 원하는 길이의 반복부

서열이 TALEN으로부터 제거될 수 있다. 특정 측면에 따라, TALEN에는 특정의 원하는 길이의 반복부 서열이 없다. 특정 측면에 따라, 원하는 길이의 반복부 서열이 제거된 TALEN을 제공한다. 특정 측면에 따라, TALEN을 변형시켜 원하는 길이의 반복부 서열을 제거한다. 특정 측면에 따라, TALEN을 조작하여 원하는 길이의 반복부 서열을 제거한다.

[0010] 본 개시내용의 측면은 세포에서 반복부 서열 100 bp 이상이 결핍된 TALEN 및 도너 핵산 서열을 조합하는 단계를 포함하고, 여기서, 세포에서 TALEN은 표적 DNA를 절단하고, 도너 핵산 서열은 DNA 내로 삽입되는 것인, 세포에서 표적 DNA를 변경시키는 방법에 관한 것이다. 본 개시내용의 측면은 반복부 서열 100 bp 이상이 결핍된 TALEN을 코딩하는 핵산 서열을 포함하는 바이러스에 관한 것이다. 본 개시내용의 측면은 반복부 서열 100 bp 이상이 결핍된 TALEN을 코딩하는 핵산 서열을 포함하는 세포에 관한 것이다. 본원에 기술된 특정 측면에 따라, TALEN에는 반복부 서열 100 bp 이상, 90 bp 이상, 80 bp 이상, 70 bp 이상, 60 bp 이상, 50 bp 이상, 40 bp 이상, 30 bp 이상, 20 bp 이상, 19 bp 이상, 18 bp 이상, 17 bp 이상, 16 bp 이상, 15 bp 이상, 14 bp 이상, 13 bp 이상, 12 bp 이상, 11 bp 이상, 또는 10 bp 이상이 결핍되어 있다.

[0011] 본 개시내용의 측면은 엔도뉴클레아제, DNA 폴리머라제, DNA 리가제, 엑소뉴클레아제, 반복 가변 이장기 도메인을 코딩하는 복수 개의 핵산 이량체 블록 및 엔도뉴클레아제 절단 부위를 포함하는 TALE-N/TF 골격 벡터를 조합하는 단계, 엔도뉴클레아제를 활성화시켜 엔도뉴클레아제 절단 부위에서 TALE-N/TF 골격 벡터를 절단하여 제1 단부 및 제2 단부를 제조하는 단계, 엑소뉴클레아제를 활성화시켜 TALE-N/TF 골격 벡터 및 복수 개의 핵산 이량체 블록 상에 3' 및 5' 오버행을 형성하고, TALE-N/TF 골격 벡터 및 복수 개의 핵산 이량체 블록을 원하는 순서로 어닐링하는 단계, 및 DNA 폴리머라제 및 DNA 리가제를 활성화시켜 TALE-N/TF 골격 벡터 및 복수 개의 핵산 이량체 블록을 연결시키는 단계를 포함하는, TALE 제조에 관한 것이다. 통상의 기술자는 본 개시내용에 기초하여 적합한 엔도뉴클레아제, DNA 폴리머라제, DNA 리가제, 엑소뉴클레아제, 반복 가변 이장기 도메인을 코딩하는 핵산 이량체 블록 및 TALE-N/TF 골격 벡터를 쉽게 확인할 수 있을 것이다.

[0012] 본 개시내용의 측면은 (a) 줄기 세포 내로 표적 DNA에 상보적이고, 효소를 표적 DNA로 가이드하는 RNA를 코딩하는 제1 외래 핵산을 도입하는 단계로서, 여기서, RNA 및 효소가 표적 DNA에 대한 공동 국제화 복합체의 구성원인 것인 단계, 줄기 세포 내로 도너 핵산 서열을 코딩하는 제2 외래 핵산을 도입하여 줄기 세포에서 변경된 DNA를 제조하는 단계로서, 여기서, RNA 및 도너 핵산 서열이 발현되고, 여기서, RNA 및 효소는 표적 DNA로 공동 국제화되고, 효소는 표적 DNA를 절단하고, 도너 핵산은 표적 DNA 내로 삽입되는 것인 단계를 포함하는, 표적 DNA에 상보적인 RNA와 공동 국제화 복합체를 형성하고, 부위 특이적인 방식으로 표적 DNA를 절단하는 효소를 발현하는 줄기 세포에서 표적 DNA를 변경시키는 방법에 관한 것이다.

[0013] 본 개시내용의 측면은 표적 DNA에 상보적인 RNA와 공동 국제화 복합체를 형성하고, 부위 특이적인 방식으로 표적 DNA를 절단하는 효소를 코딩하는 제1 외래 핵산을 포함하는 줄기 세포에 관한 것이다.

[0014] 본 개시내용의 측면은 표적 DNA에 상보적인 RNA와 공동 국제화 복합체를 형성하고, 부위 특이적인 방식으로 표적 DNA를 절단하는 효소를 코딩하는 제1 외래 핵산을 포함하고, 효소의 발현을 촉진시키는 유도성 프로모터를 포함하는 것인 세포에 관한 것이다. 이러한 방식으로, 발현은 조절될 수 있고, 예를 들어, 개시될 수 있고, 종결될 수 있다.

[0015] 본 개시내용의 측면은 표적 DNA에 상보적인 RNA와 공동 국제화 복합체를 형성하고, 부위 특이적인 방식으로 표적 DNA를 절단하는 효소를 코딩하는 제1 외래 핵산을 포함하는 세포로서, 여기서, 제1 외래 핵산은 제거 효소, 예컨대, 트랜스포사제를 사용하여 세포의 게놈 DNA로부터 제거가능한 것인, 세포에 관한 것이다.

[0016] 본 개시내용의 측면은 (a) 세포 내로 도너 핵산 서열을 코딩하는 제1 외래 핵산을 도입하는 단계, 세포 주변 배지로부터 세포 내로 표적 DNA에 상보적이고, 효소를 표적 DNA로 가이드하는 RNA를 도입하여 세포에서 변경된 DNA를 제조하는 단계로서, 여기서, RNA 및 효소가 표적 DNA에 대한 공동 국제화 복합체의 구성원이고, 여기서, 도너 핵산 서열이 발현되고, 여기서, RNA 및 효소는 표적 DNA로 공동 국제화되고, 효소는 표적 DNA를 절단하고, 도너 핵산은 표적 DNA 내로 삽입되는 것인 단계를 포함하는, 표적 DNA에 상보적인 RNA와 공동 국제화 복합체를 형성하고, 부위 특이적인 방식으로 표적 DNA를 절단하는 효소를 발현하는 세포에서 표적 DNA를 변경시키는 방법에 관한 것이다.

[0017] 본 개시내용의 측면은 줄기 세포를 유전적으로 변형시키기 위한 RNA에 의해 가이드되는 DNA 결합 단백질의 용도에 관한 것이다. 한 측면에서, 줄기 세포는 RNA에 의해 가이드되는 DNA 결합 단백질을 코딩하는 핵산을 포함하도록 유전적으로 변형되고, 줄기 세포는 RNA에 의해 가이드되는 DNA 결합 단백질을 발현한다. 특정 측면에 따

라, 특히 돌연변이를 도입하기 위한 도너 핵산은 변형된 TALEN 또는 RNA에 의해 가이드되는 DNA 결합 단백질을 사용하여 게놈 편집을 위해 최적화된다.

- [0018] 본 개시내용의 측면은 하나 이상의 가이드 RNA (리보핵산)를 사용하여 줄기 세포에 의해 발현된 뉴클레아제 활성을 가지는 효소, 예컨대, 뉴클레아제 활성을 가지는 DNA 결합 단백질을 DNA (데옥시리보핵산) 상의 표적 위치로 유도하는 것으로서, 여기서, 효소는 DNA를 절단하고, 외인성 도너 핵산이 예컨대, 상동성 재조합에 의해 DNA 내로 삽입되는 것인, 줄기 세포에서의 DNA 변형, 예컨대, DNA의 다중 변형에 관한 것이다. 본 개시내용의 측면은 줄기 세포 상에서 DNA 변형 단계를 사이클링 또는 반복하여 세포 내에서 DNA의 다중 변형을 가지는 줄기 세포를 제조하는 것을 포함한다. 변형은 외인성 도너 핵산의 삽입을 포함할 수 있다.
- [0019] 다중 외인성 핵산 삽입은 복수 개의 RNA를 코딩하는 핵산 및 복수 개의 외인성 도너 핵산을 예컨대, 공동 형질 전환에 의해 효소를 발현하는 줄기 세포 내로 도입하는 단일 단계로서, 여기서, RNA는 발현되고, 여기서, 복수 개의 각 RNA는 효소를 DNA의 특정 부위로 가이드하고, 효소는 DNA를 절단하고, 복수 개의 외인성 핵산 중 하나는 절단 부위에서 DNA 내로 삽입되는 것인 단계에 의해 달성될 수 있다. 따라서, 본 측면에 따라, 세포에서의 DNA의 다수의 변경 또는 변형은 단일 사이클로 형성된다.
- [0020] 다중 외인성 핵산 삽입은 하나 이상의 RNA 또는 복수 개의 RNA를 코딩하는 하나 이상의 핵산 및 하나 이상의 외인성 핵산 또는 복수 개의 외인성 핵산을 효소를 발현하는 줄기 세포 내로 도입하는 반복 단계 또는 사이클로서, 여기서, RNA는 발현되고, 효소를 DNA의 특정 부위로 가이드하고, 효소는 DNA를 절단하고, 외인성 핵산은 절단 부위에서 DNA 내로 삽입되어 줄기 세포 내에서 DNA 내로의 다수의 변경 또는 외인성 DNA의 삽입을 가지는 세포를 생성하는 것인 단계 또는 사이클에 의해 세포에서 달성될 수 있다. 따라서, 한 측면에 따라, 효소를 발현하는 줄기 세포는 예컨대, 세포 내로 효소를 코딩하고, 줄기 세포에 의해 발현될 수 있는 핵산을 도입함으로써 효소를 발현하도록 유전적으로 변경되었다. 이러한 방식으로, 본 개시내용의 측면은 효소를 발현하는 줄기 세포 내로 RNA를 도입하는 단계, 줄기 세포 내로 외인성 도너 핵산을 도입하는 단계, RNA를 발현하는 단계, RNA, 효소, 및 DNA의 공동 국제화 복합체를 형성하는 단계, 효소에 의해 DNA를 효소적으로 절단하는 단계, 및 도너 핵산을 DNA 내로 삽입하는 단계를 사이클링하는 것을 포함한다. 상기 단계를 사이클링 또는 반복함으로써 줄기 세포를 다중 유전자좌에서 다중으로 유전적으로 변형시킬 수 있고, 즉, 다중의 유전적 변형을 가지는 줄기 세포를 생성할 수 있다.
- [0021] 특정 측면에 따라, 본 개시내용의 범주 내의 DNA 결합 단백질 또는 효소는 가이드 RNA와 복합체를 형성하는 단백질을 포함하고, 가이드 RNA는 상기 복합체를 이중 가닥 DNA 서열로 가이드하며, 여기서, 복합체는 DNA 서열에 결합한다. 한 측면에 따라, 효소는 RNA에 의해 가이드되는 DNA 결합 단백질, 예컨대, DNA에 결합하고, RNA에 의해 가이드되는, II형 CRISPR 시스템의 RNA에 의해 가이드되는 DNA 결합 단백질일 수 있다. 한 측면에 따라, RNA에 의해 가이드되는 DNA 결합 단백질은 Cas9 단백질이다.
- [0022] 본 개시내용의 상기 측면은 이중 가닥 DNA에의, 또는 그와의 RNA 및 DNA 결합 단백질의 공동 국제화로 지칭될 수 있다. 이러한 방식으로, DNA 결합 단백질-가이드 RNA 복합체는 다중으로 유전적 변형된 줄기 세포, 예컨대, 외인성 도너 DNA가 다중으로 삽입된 줄기 세포를 생성하기 위해 이중 가닥 DNA의 다중 부위를 절단하는 데 사용될 수 있다.
- [0023] 특정 측면에 따라, (a) 줄기 세포 내로 표적 DNA에 상보적이고, 효소를 표적 DNA로 가이드하는 하나 이상의 RNA를 코딩하는 제1 외래 핵산을 도입하는 단계로서, 여기서, 하나 이상의 RNA 및 효소가 표적 DNA에 대한 공동 국제화 복합체의 구성원인 것인 단계, 줄기 세포 내로 하나 이상의 도너 핵산 서열을 코딩하는 제2 외래 핵산을 도입하여 줄기 세포에서 변경된 DNA를 제조하는 단계로서, 여기서, 하나 이상의 RNA 및 하나 이상의 도너 핵산 서열이 발현되고, 여기서, 하나 이상의 RNA 및 효소는 표적 DNA로 공동 국제화되고, 효소는 표적 DNA를 절단하고, 도너 핵산은 표적 DNA 내로 삽입되는 것인 단계, 및 단계 (a)를 다회에 걸쳐 반복하여 줄기 세포에서 DNA를 다중으로 변경시키는 단계를 포함하는, 표적 DNA에 상보적인 RNA와 공동 국제화 복합체를 형성하고, 부위 특이적인 방식으로 표적 DNA를 절단하는 효소를 발현하는 줄기 세포에서 표적 DNA를 다중으로 변경시키는 방법을 제공한다.
- [0024] 한 측면에 따라, RNA는 약 10 내지 약 500개의 뉴클레오티드이다. 한 측면에 따라, RNA는 약 20 내지 약 100개의 뉴클레오티드이다.
- [0025] 한 측면에 따라, 하나 이상의 RNA는 가이드 RNA이다. 한 측면에 따라, 하나 이상의 RNA는 tracrRNA-crRNA 융합체이다.

[0026] 한 측면에 따라, DNA는 게놈 DNA, 미토콘드리아 DNA, 바이러스 DNA, 또는 외인성 DNA이다.

[0027] 한 측면에 따라, 세포는 효소를 사용하여 쉽게 제거될 수 있는 벡터를 이용하여 DNA 결합 효소를 코딩하는 핵산을 가역적으로 포함하도록 유전적으로 변형될 수 있다. 유용한 벡터 방법은 통상의 기술자에게 공지되어 있고, 렌티바이러스, 다에노 관련 바이러스, 뉴클레아제 및 인테그라제 매개 표적 삽입 방법 및 트랜스포존 매개 삽입 방법을 포함한다. 한 측면에 따라, 예컨대, 카세트 또는 벡터에 의해 부가된 DNA 결합 효소를 코딩하는 핵산은 예를 들어, 그 전체로 카세트 및 벡터와 함께, 및 게놈 DNA 중 상기 핵산, 카세트 또는 벡터의 일부를 남기지 않고 제거될 수 있다. 상기와 같은 제거는, 게놈이 핵산, 카세트 또는 벡터의 부가 전과 동일하기 때문에 관련 기술분야에서 "흔적 없는" 제거로 지칭된다. 삽입 및 흔적 없는 제거에 대한 한 예시적인 실시양태는 시스템 바이오사이언시즈(System Biosciences)로부터 상업적으로 이용가능한 피기백(PiggyBac) 벡터이다.

[0028] 본 발명의 특정 실시양태의 추가의 특징 및 이점은 하기 실시양태에 대한 상세한 설명 및 그의 도면에서, 및 청구범위로부터 더욱 충분하게 자명해질 것이다.

도면의 간단한 설명

[0029] 본 실시양태의 상기 및 다른 특징 및 이점은 첨부된 도면과 함께 하기의 예시적인 실시양태에 관한 상세한 설명으로부터 더욱 충분히 이해될 것이다:

도 1은 인간 체세포 및 줄기 세포에서의 리TALEN의 기능적 시험에 관한 것이다.

(a) 게놈 표적화 효율을 시험하기 위한 실험 디자인의 도식적 대표도. 종결 코돈 및 AAVS1 유전자좌로부터 유래된 68 bp 게놈 단편의 삽입에 의해 게놈적으로 통합된 GFP 코딩 서열을 파괴한다 (하단). tGFP 도너와의 뉴클레아제-매개 상동성 재조합에 의한 GFP 서열의 복원 (상단)을 통해 FACS에 의해 정량화될 수 있는 GFP+ 세포를 수득한다. 리TALEN 및 TALEN은 AAVS1 단편 내의 동일한 서열을 표적화한다.

(b) FACS에 의해 측정된, tGFP 도너만이 단독으로, tGFP 도너와 함께 TALEN, 및 tGFP 도너와 함께 리TALEN이 표적 유전자좌에 도입된 GFP+ 세포의 비율(%)을 도시한 막대 그래프. (N=3, 오차 막대=SD). 대표 FACS 플롯이 하기에 제시되어 있다.

(c) 천연 AAVS1 유전자좌에 대한 표적화 전략을 도시한 도식적 개요. 스플라이싱 엑셉터 (SA)- 2A (자가-절단 펩티드), 퓨로마이신 저항성 유전자 (PURO) 및 GFP를 함유하는 도너 플라스미드가 기술되었다 (그 전문이 본원에 참조로 포함된 참고 문헌 10을 참조할 수 있다). 성공적인 편집 이벤트를 검출하는 데 사용된 PCR 프라이머의 위치는 청색 화살표로 표시되어 있다.

(d) 2주 동안 퓨로마이신 (0.5ug/mL)을 이용하여 PGP1 hiPSC의 성공적으로 표적화된 클론을 선별하였다. 대표적인 GFP+ 클론 3개에 대한 현미경 영상이 제시되어 있다. 세포를 또한 다능성 마커 TRA-1-60에 대하여 염색하였다. 스케일 바: 200 μm.

(e) 상기의 단일클론 GFP+ hiPSC 클론에서 수행된 PCR 검정법을 통해 도너 카세트가 AAVS1 부위에 성공적으로 삽입된 반면 (레인 1,2,3), 단순 hiPSC는 성공적인 삽입에 관한 어떤 증거도 보이지 않았다는 것 (레인 C)이 입증되었다.

도 2는 iPSC 중 CCR5 상에서의 리TALEN 및 Cas9-gRNA 게놈 표적화 효율 비교에 관한 것이다.

(a) 게놈 조작 실험 디자인의 도식적 대표도. 리TALEN 쌍 또는 Cas9-gRNA 표적화 부위에서 게놈 DNA에 대한 2 bp 미스매치를 보유하는 90mer ssODN을 리TALEN 또는 Cas9-gRNA 구축물과 함께 PGP1 hiPSC 내로 전달하였다. 도면에서 뉴클레아제의 절단 부위는 적색 화살표로 표시되어 있다.

(b) 리TALEN 쌍 (CCR5 #3) 및 ssODN, 또는 Cas9-gRNA 및 ssODN에 대한 HDR 및 NHEJ 효율에 관한 심층 서열분석 분석. hiPSC의 게놈 중의 변형을 GEAS에 의한 고처리량 서열 데이터로부터 분석하였다. 상단: HDR은 ssODN의 중앙부 내로 구축된 2 bp 점 돌연변이를 함유하는 리드의 분율로부터 정량화하고 (적색), NHEJ 활성은 게놈 내의 각 특이 위치에서의 결실 (회색)/삽입 (적색)의 분율로부터 정량화하였다. 리TALEN 및 ssODN 그래프의 경우, 녹색 점선을 플로팅하여 두 리TALEN 결합 부위의 중앙부에 대해 상대적인 위치로 -26 bp 내지 +26 bp 위치인 리TALEN 쌍의 결합 부위의 외곽 경계를 마킹하였다. Cas9-gRNA 및 ssODN 그래프의 경우, 녹색 점선은 PAM 서열에 대해 상대적인 위치로 -20 내지 -1 bp 위치인 gRNA 표적화 부위의 외곽 경계를 마킹한다. 하단: 상기 명시된 바와 같이 처리된 전체 NHEJ 집단으로부터 분석된 hiPSC 중의 결실/삽입 크기 분포.

(c) PGP1 hiPSC에서의 리TALEN 및 Cas9-gRNA 표적화 CCR5의 게놈 편집 효율.

상단: CCR5에서 표적화된 게놈 편집 부위에 관한 도식적 대표도. 15개의 표적화 부위가 하기에 청색 화살표로 도시되어 있다. 각 부위에 대하여 세포를 리TALEN, 및 게놈 DNA에 대한 2 bp 미스매치를 보유하는 그의 상응하는 ssODN 도너로 이루어진 쌍으로 공동 형질감염시켰다. 형질감염 후 6일 경과하였을 때 게놈 편집 효율을 검정하였다. 유사하게, 15개의 Cas9-gRNA를 그의 상응하는 ssODN을 이용하여 개별적으로 PGP1-hiPSC 내로 형질감염시켜 같은 15개의 부위를 표적화하고, 형질감염 후 6일 경과하였을 때 효율을 분석하였다. 하단: PGP1 hiPSC에서의 리TALEN 및 Cas9-gRNA 표적화 CCR5의 게놈 편집 효율. 패널 1 및 2는 리TALEN에 의해 매개된 NHEJ 및 HDR 효율을 나타낸다. 패널 3 및 4는 Cas9-gRNA에 의해 매개된 NHEJ 및 HDR 효율을 나타낸다. NHEJ 비율은 표적화 부위에 결실 또는 삽입을 보유하는 게놈 대립 유전자의 빈도에 의해 계산하였고; HDR 비율은 2 bp 미스매치를 보유하는 게놈 대립 유전자의 빈도에 의해 계산하였다. 패널 5는 ENCODE 데이터베이스로부터의 hiPSC 세포주의 DNaseI HS 프로파일이다 (듀크(Duke) DNase HS, iPS NIHi7 DS). 특히, 다른 패널의 스케일은 상이하다.

(f) 3개의 표적화된 hiPSC 콜로니로부터의 PCR 앰플리콘의 생어(Sanger) 서열분석을 통해 게놈-삽입 경계에 예상 DNA 염기가 존재한다는 것을 확인하였다. M: DNA 래더. C: 단순 hiPSC 게놈 DNA를 포함하는 대조군.

도 3은 PGP1 hiPSC에서의 리TALEN 또는 Cas9-gRNA를 이용하는 ssODN-매개 HDR을 지배하는 기능적 파라미터에 관한 연구에 관한 것이다.

(a) PGP1 hiPSC를 리TALEN 쌍 (#3) 및 상이한 길이 (50, 70, 90, 110, 130, 150, 170 nts)의 ssODN으로 공동 형질감염시켰다. 모든 ssODN은 그의 서열 중간에 게놈 DNA에 대하여 동일한 2 bp 미스매치를 포함하였다. 90mer ssODN은 표적화된 게놈에서 최적의 HDR을 달성하였다. HDR, NHEJ 발생된 결실 및 삽입 효율 평가는 본원에 기술되어 있다.

(b) 각각이 중앙에 2 bp 미스매치 (A) 및 A로부터의 상이한 위치 상쇄부 (여기서, 상쇄부는 -30 bp→30 bp로 달랐다)에 추가의 2 bp 미스매치 (B)를 함유하는 리TALEN 쌍 #3에 상응하는 90mer ssODN을 사용하여 ssODN에 따른 상동성으로부터의 이탈의 효과를 시험하였다. 각 ssODN의 게놈 편집 효율은 PGP1 hiPSC에서 평가하였다. 하단의 막대 그래프는 표적화된 게놈에서의 오직 A만, 오직 B만, 및 A + B의 도입 빈도를 보여주는 것이다. HDR 비율은 중앙으로부터의 상동성 이탈 거리가 증가할수록 감소한다.

(c) 리TALEN 쌍 #3의 표적 부위로부터 다양한 거리 (-620 bp ~ 480 bp)를 두고 떨어져 있는 부위로 표적화되는 ssODN을 시험함으로써 돌연변이를 도입하는 데 ssODN이 배치될 수 있는 최대 거리를 평가하였다. 모든 ssODN은 그의 서열 한가운데에 2 bp 미스매치를 보유하였다. ssODN 미스매치가 리TALEN 쌍의 결합 부위의 한가운데로부터 40 bp만큼 떨어져 위치할 때 최소 HDR 효율 (<=0.06%)이 관찰되었다.

(d) PGP1 hiPSC를 Cas9-gRNA (AAVS1) 및 상이한 배향 (Oc: gRNA에 대해 상보적; On: gRNA에 대해 비상보적) 및 상이한 길이 (30, 50, 70, 90, 110 nt)의 ssODN으로 공동 형질감염시켰다. 모든 ssODN은 그의 서열 한가운데에 게놈 DNA에 대하여 동일한 2 bp 미스매치를 포함하였다. 70 mer Oc가 표적화된 게놈에서 최적의 HDR을 달성하였다.

도 4는 선별 없이 단일클론 게놈 편집된 hiPSC를 획득하기 위해 리TALEN 및 ssODN을 사용하는 것에 관한 것이다.

(a) 실험 타임라인.

(b) 도 2b에 기술된 NGS 플랫폼에 의해 평가된 리TALEN 쌍 및 ssODN (#3)의 게놈 조작 효율.

(c) 게놈 편집 이후의 단일클론 hiPSC 콜로니의 생어 서열분석 결과. 2 bp 이형접합성 유전자형 (CT/CT→TA/CT)이 PGP1-iPS-3-11, PGP1-iPS-3-13 콜로니의 게놈 내로 성공적으로 도입되었다.

(d) 표적화된 PGP1-iPS-3-11의 면역형광 염색. 세포를 다능성 마커 Tra-1-60 및 SSEA4에 대해 염색하였다.

(e) 단일클론 PGP1-iPS-3-11 세포로부터 생성된 테라토마 섹션의 헤마톡실린 및 에오신 염색.

도 5. 리TALE 디자인. (a) 리TALE-16.5에서의 단량체와 함께 원래의 TALE RVD 단량체로 이루어진 서열 정렬 (리TALE-M1→리TALE-M17). 원래의 서열로부터의 뉴클레오티드 변경은 회색으로 강조 표시되어 있다. (b) PCR에 의한 리TALE의 반복성 시험. 상단의 패널은 PCR 반응에서의 프라이머의 위치 및 리TALE/TALE의 구조를 도시한 것이다. 하단의 패널은 PCR 밴드를 도시한 것이고, 조건은 하기에 명시되어 있다. 원래의 TALE 주형 (우측 레

인)과 함께 PCR 래더링이 제시되어 있다.

도 6. TALE 단일-인큐베이션 조립 (TASA) 조립 디자인 및 실시.

(a) TASA 조립을 위한 리TALE 이량체 블록의 라이브러리의 도식적 대표도. 2개의 RVD를 코딩하는 10개의 리TALE 이량체 블록의 라이브러리가 존재한다. 각 블록 내에서 16개의 이량체는 모두 RVD를 코딩하는 서열을 제외하면 동일한 DNA 서열을 공유한다; 다른 블록 중의 이량체는 독특한 서열을 가지지만, 그가 인접 블록과 32 bp의 중첩부를 공유하도록 디자인된다. 한 이량체 (블록6_AC)의 DNA 및 아미노산 서열은 우측에 열거되어 있다.

(b) TASA 조립의 도식적 대표도. 좌측 패널은 TASA 조립 방법을 도시한 것이다: 효소 혼합물/리TALE 블록/리TALE-N/TF 골격 벡터를 이용하여 원-포트 인큐베이션 반응을 수행한다. 반응 생성물은 박테리아 형질전환을 위해 직접 사용될 수 있다. 우측 패널은 TASA의 기전을 도시한 것이다. 37°C에서 엔도뉴클레아제에 의해 목표 벡터를 선형화하여 ccdB 역 선택 카세트를 절단하고; 블록의 단부 및 선형화된 벡터를 프로세싱하는 엑소뉴클레아제는 단편의 단부의 ssDNA 오버행을 노출시켜 블록 및 벡터 골격이 지정된 순서로 어닐링될 수 있도록 한다. 온도가 최대 50°C까지 상승하였을 때, 폴리머라제 및 리가제는 함께 작용하여 겹을 실링하고, 이로써, 형질전환을 위해 준비된 최종 구축물이 생성된다.

(c) 상이한 단량체 길이를 가지는 리TALE에 대한 TASA 조립 효율. 조립에 사용된 블록은 좌측에 도시되어 있고, 조립 효율은 우측에 제시되어 있다.

도 7. 렌터-리TALE의 기능성 및 서열 완전성.

도 8. GEAS의 감도 및 재현가능성.

(a) HDR 검출 한계의 정보 기반 분석. 리TALEN (#10)/ssODN인 데이터세트 전제하에, 예상되는 편집 (HDR)을 함유하는 리드를 확인하고, 이들 HDR 리드를 체계적으로 제거하여 편집 신호가 "희석된" 상이한 인공 데이터세트를 작성하였다. 제거율이 100, 99.8, 99.9, 98.9, 97.8, 89.2, 78.4, 64.9, 21.6, 10.8, 2.2, 1.1, 0.2, 0.1, 0.02, 및 0%인 HDR 리드를 포함하는 데이터세트는 HR 효율 범위가 0~0.67%인 인공 데이터세트로 작성되었다. 각각의 개별 데이터세트에 대하여, 배경 신호 (보라색), 및 표적화 부위에서 수득된 신호 (녹색)의 상호 정보 (MI)를 추정하였다. HDR 효율이 0.0014%를 초과할 때, 표적화 부위에서의 MI가 배경보다 현저히 더 높다. HDR 검출 한계는 0.0014% 내지 0.0071%인 것으로 추정되었다. MI 계산은 본원에 기술된 바와 같이 수행된다.

(b) 게놈 편집 평가 시스템의 재현가능성 시험. 플롯 쌍 (상단 및 하단)은 상기 명시된 리TALEN 쌍 및 세포 유형을 이용한 두 복제물의 HDR 및 NHEJ 평가 결과를 보여주는 것이다. 각 실험을 위해, 뉴클레오펙션, 표적화된 게놈 증폭, 심층 서열분석 및 데이터 분석을 독립적으로 수행하였다. 복제물의 게놈 편집 평가 편차는 $\sqrt{2(|HDR1-HDR2|)/((HDR1+HDR2)/2)} = \Delta HDR/HDR$ 및 $\sqrt{2(|NHEJ1-NHEJ2|)/((NHEJ1+NHEJ2)/2)} = \Delta NHEJ/NHEJ$ 로 계산하였고, 편차 결과는 하기 플롯에 열거되어 있다. 시스템의 평균 편차는 $(19\%+11\%+4\%+9\%+10\%+35\%)/6=15\%$ 였다. 편차에 기여할 수 있는 인자로는 뉴클레오펙션하에 있는 세포의 상태, 뉴클레오펙션 효율, 및 서열분석 커버리지 및 품질을 포함한다.

도 9. CCR5 상에서의 리TALEN 및 Cas9-gRNA에 의한 NHEJ 및 HDR 효율의 통계학적 분석

(a) iPSC 중 동일 부위에서의 리TALEN에 의해 매개된 HR 및 NHEJ 효율의 상관관계 ($r=0.91, P < 1 \times 10^{-5}$).

(b) iPSC 중 동일 부위에서의 Cas9-gRNA에 의해 매개된 HR 및 NHEJ 효율의 상관관계 ($r=0.74, P=0.002$).

(c) iPSC에서의 Cas9-gRNA에 의해 매개된 NHEJ 효율 및 gRNA 표적화 부위의 Tm 온도의 상관관계 ($r=0.52, P=0.04$).

도 10. 게놈 편집 효율 및 후성적 상태의 상관관계 분석.

피어슨 상관관계를 사용하여 DNase I 감도와 게놈 조작 효율 (HR, NHEJ) 사이의 가능한 연관성을 연구하였다. 관찰된 상관관계를 임의 추출된 세트 (N=100000)와 비교하였다. 관찰된 상관관계가 모의 분포의 95번째 백분위수보다 더 높거나, 또는 5번째 백분위수보다 더 낮은 것은 잠재적으로 연관성이 있는 것으로 간주되었다. DNase I 감도와 NHEJ/HR 효율 사이에 어떤 뚜렷한 상관관계도 관찰되지 않았다.

도 11. ssODN-매개 게놈 편집에서의 상동성 쌍 형성의 영향.

(a) 도 3b에 기술된 실험에서, 중간 2 b 미스매치 (A)가 도입된 속도로 측정된 바, 전체 HDR은 제2 미스매치 B

의 A로부터의 그의 거리가 증가함에 따라 감소하였다 (B에서 A까지의 상대적인 위치는 -30→30 bp로 다르다). B가 A로부터 단지 10 bp만큼 떨어져 있을 때 (-10 bp 내지 +10b)의 더 빠른 도입 속도는 dsDNA 절단에 인접한 게놈 DNA에 대한 ssODN의 쌍 형성이 덜 요구된다는 것을 반영하는 것일 수 있다.

(b) ssODN에 따른 유전자 전환 길이의 분포. A로부터의 B의 각 거리에서, HDR 이벤트 중 일부는 오직 A만을 도입하는 반면, 또 다른 일부는 A 및 B, 둘 모두를 도입한다. 상기 두 이벤트는 유전자 전환 트랙으로 해석될 수 있는데 (Elliott et al., 1998), 여기서, A+B 이벤트는 B 너머까지 확장된 긴 전환 트랙을 나타내고, 오직 A만의 이벤트는 B에 도달하지 못한 더 짧은 것을 나타낸다. 이러한 해석시, 올리고를 따른 양 방향의 유전자 전환 길이의 분포는 추정될 수 있다 (ssODN의 중간을 0으로, ssODN의 5' 단부쪽으로서의 전환 트랙은 (-) 방향으로, 및 3' 단부 쪽으로서의 것은 (+) 방향으로 정의된다). 유전자 전환 트랙은 그의 길이가 증가함에 따라 점진적으로 발생률이 감소하는데, 이 결과는 dsDNA 도너를 이용하였을 때 관찰된 유전자 전환 트랙 분포와 매우 유사한 결과로서, 단, dsDNA 도너의 경우에는 수백 개의 염기인데 반해, ssDNA 올리고의 경우에는 수십 bp 만큼의 고도로 압축된 거리 규모일 때이다.

(c) 일련의 돌연변이를 함유하는 단일 ssODN을 이용하고, 인접한 일련의 도입을 측정하는 유전자 전환 트랙에 대한 검정법. 중앙의 2 bp 미스매치 양측에 10 nt 만큼의 간격을 두고 이격되어 있는 3쌍의 2 bp 미스매치 (오렌지색)를 포함하는 ssODN 도너 (상단)를 사용하였다. 상기 영역을 서열분석하였을 때 >300,000개의 리드 중 ssODN에 의해 정의되는 미스매치를 >=1 보유하는 게놈 서열분석 리드는 거의 검출되지 않았다 (참고 문헌 62 (상기 문헌은 그 전문이 본원에서 참조로 포함된다) 참조). 상기 리드 모두 플롯팅하였고 (하단), 리드 서열은 색깔로 코딩되어 있다. 오렌지색: 정의된 미스매치; 녹색: 야생형 서열. 상기 ssODN을 사용하여 게놈 편집을 수행하였을 때, 오직 한가운데 돌연변이만이 도입된 패턴을 상기 경우의 85% (53/62)가 되게 하였고, 다른 경우에는 다중 B 미스매치가 도입되었다. 비록 B 도입 이벤트 횟수가 너무 낮아 길이가 > 10 bp인 트랙 분포를 추정하기 어렵기는 했지만, -10-10 bp의 짧은 트랙 영역이 지배적이라는 것은 명백하다.

도 12. Cas9-gRNA 뉴클레아제 및 니카제 게놈 편집 효율.

뉴클레아제 (C2) (Cas9-gRNA) 또는 니카제 (Cc) (Cas9D10A-gRNA) 및 상이한 배향 (Oc 및 On)의 ssODN의 조합으로 PGP1 iPSC를 공동 형질감염시켰다. 모든 ssODN은 그의 서열 한가운데에 게놈 DNA에 대하여 동일한 2 bp 미스매치를 포함하였다. HDR 평가는 본원에 기술되어 있다.

도 13. 리TALE 서열 디자인 및 최적화

반복부를 제거하기 위해 수회의 디자인 사이클을 거쳐 리TALE 서열을 진화시켰다. 매회 사이클에서, 각 반복부로부터의 동의적 서열을 평가한다. 진화 DNA까지의 해밍 거리가 가장 큰 것을 선별한다. 최종 서열의 cai = 0.59이고, ΔG= -9.8 kcal/mol이다. 합성 단백질 디자인을 위해 상기의 일반 골격을 수행하기 위하여 R 패키지를 제공하였다.

도 14는 PGP1 세포에서의 Cas 9의 게놈 삽입의 PCR 확인을 보여주는 겔 영상이다. 라인 3, 6, 9, 12는 단순 PGP1 세포주의 PCR 생성물이다.

도 15는 유도하의 Cas9 mRNA의 mRNA 발현 수준을 나타낸 그래프이다.

도 16은 상이한 RNA 디자인에 의한 게놈 표적화 효율을 보여주는 그래프이다.

도 17은 가이드 RNA - 도너 DNA 융합체에 의해 달성된 44%의 상동성 재조합의 게놈 표적화 효율을 보여주는 그래프이다.

도 18은 본원에 기술된 시스템에 의해 생성된 동질유전자 PGP1 세포주의 유전자형을 보여주는 다이어그램이다. PGP1-iPS-BTHH는 BTHH 환자와 같은 단일 뉴클레오티드 결실 표현형을 가진다. PGP1-NHEJ는 다른 방식으로 프레임 쉬프트 돌연변이를 생성한 4 bp 결실을 가진다.

도 19는 환자 특이 세포에서 입증된 바, 동질유전자 PGP1 iPSC로부터 유래된 심장 근육 세포가 결손 ATP 생산 및 F1F0 ATPase 특이 활성을 반복하였음을 보여주는 그래프이다.

발명을 실시하기 위한 구체적인 내용

[0030] 본 발명의 측면은 예를 들어, 이중 가닥 핵산 절단에 의해 이루어지는 핵산 조작을 위한, 특정 반복부 서열이 결핍된 TALEN의 용도에 관한 것이다. 이중 가닥 핵산을 절단하는 데 TALEN을 사용하면, 비상동성 단부 연결 (NHEJ) 또는 상동성 재조합 (HR)이 이루어질 수 있다. 본 개시내용의 측면은 또한 예를 들어, 도너 핵산 존재

하의 이중 가닥 핵산 절단에 의해 이루어지는 핵산 조작을 위한, 및 예컨대, 비상동성 단부 연결 (NHEJ) 또는 상동성 재조합 (HR)에 의한 도너 핵산의 이중 가닥 핵산 내로의 삽입을 위한, 특정 반복부 서열이 결핍된 TALEN의 용도도 고려한다.

- [0031] 전사 활성인자-유사 효과기 뉴클레아제 (TALEN)는 관련 기술분야에 공지되어 있고, TAL 효과기 DNA 결합 도메인을 DNA 절단 도메인에 융합시킴으로써 생성된 인공 제한 효소를 포함한다. 제한 효소는 특이 서열에서 DNA 가닥을 절단하는 효소이다. 전사 활성인자-유사 효과기 (TALE) 뉴클레아제는 원하는 DNA 서열에 결합하도록 조작될 수 있다. 문헌 [Boch, Jens (February 2011). "TALEs of genome targeting". *Nature Biotechnology* 29 (2): 135-6] (상기 문헌은 그 전문이 본원에서 참조로 포함된다)을 참조할 수 있다. 상기 조작된 TALE를 (DNA 가닥을 절단하는) DNA 절단 도메인과 조합함으로써 임의의 원하는 DNA 서열에 대해 특이적인 제한 효소인 TALEN이 제조된다. 특정 측면에 따라, TALEN을 계내 표적 핵산 편집을 위해, 예컨대, 계내 게놈 편집을 위해 세포 내로 도입한다.
- [0032] 한 측면에 따라, FokI 엔도뉴클레아제의 단부로부터의 비-특이 DNA 절단 도메인을 이용하여 효모 세포, 식물 세포 및 동물 세포에서 활성인 하이브리드 뉴클레아제를 구축할 수 있다. FokI 도메인은 적절한 배향 및 간격을 두고 표적 게놈 중의 부위에 대해 독특한 DNA 결합 도메인을 가진 두 구축물을 필요로 하면서, 이량체로서의 기능을 한다. TALE DNA 결합 도메인과 FokI 절단 도메인 사이의 아미노산 잔기의 개수, 및 두 개별 TALEN 결합 부위 사이의 염기의 개수, 둘 모두 활성에 영향을 미친다.
- [0033] 아미노산 서열과 TALE 결합 도메인의 DNA 인식 사이의 관계를 통해 디자인가능한 단백질을 얻을 수 있다. 소프트웨어 프로그램, 예컨대, DNA웍스(DNAWorks)를 이용하여 TALE 구축물을 디자인할 수 있다. TALE 구축물을 디자인하는 다른 방법은 통상의 기술자에게 공지되어 있다. 문헌 [Cermak, T.; Doyle, E. L.; Christian, M.; Wang, L.; Zhang, Y.; Schmidt, C.; Baller, J. A.; Somia, N. V. et al. (2011). "Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting". *Nucleic Acids Research*. doi:10.1093/nar/gkr218]; [Zhang, Feng; et.al. (February 2011). "Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription", *Nature Biotechnology* 29 (2): 149-53]; [Morbiter, R.; Elsaesser, J.; Hausner, J.; Lahaye, T. (2011). "Assembly of custom TALE-type DNA binding domains by modular cloning". *Nucleic Acids Research*. doi:10.1093/nar/gkr151]; [Li, T.; Huang, S.; Zhao, X.; Wright, D. A.; Carpenter, S.; Spalding, M. H.; Weeks, D. P.; Yang, B. (2011). "Modularly assembled designer TAL effector nucleases for targeted gene knockout and gene replacement in eukaryotes". *Nucleic Acids Research*. doi:10.1093/nar/gkr188]; [Geißler, R.; Scholze, H.; Hahn, S.; Streubel, J.; Bonas, U.; Behrens, S. E.; Boch, J. (2011). "Transcriptional Activators of Human Genes with Programmable DNA-Specificity". In Shiu, Shin-Han. *PLoS ONE* 6 (5): e19509]; [Weber, E.; Gruetzner, R.; Werner, S.; Engler, C.; Marillonnet, S. (2011). "Assembly of Designer TAL Effectors by Golden Gate Cloning". In Bendahmane, Mohammed. *PLoS ONE* 6 (5): e19722] (상기 문헌들은 그 전문이 본원에서 참조로 포함된다)을 참조할 수 있다.
- [0034] 예시적인 측면에 따라, 일단 TALEN 유전자를 조립하고 나면, 이를 특정 실시양태에 따라 플라스미드로 내로 삽입할 수 있고; 이어서, 플라스미드를 이용하여 표적 세포를 형질감염시키고, 상기 표적 세포에서 유전자 생성물이 발생되고, 이는 핵으로 진입하여 게놈에 접근하게 된다. 예시적인 측면에 따라, 본원에 기술된 바와 같은 TALEN을 이용하여, 세포는 수복 기전으로 대응하는 이중 가닥 절단 (DSB)을 유도함으로써 표적 핵산, 예컨대, 게놈을 편집할 수 있다. 예시적인 수복 기전으로는 어닐링을 위한 서열 중복부가 거의 없거나, 전혀 없는 이중 가닥 절단의 양측으로부터의 DNA를 재연결하는 비-상동성 단부 연결 (NHEJ)을 포함한다. 이러한 수복 기전은 삽입 또는 결실 (indel), 또는 염색체 재배열을 통하여 게놈 중 오류를 유도하고; 상기와 같은 임의의 오류로 인하여 상기 위치에서 코딩된 유전자 생성물은 비-기능성이 될 수 있다. 문헌 [Miller, Jeffrey; et.al. (February 2011). "A TALE nuclease architecture for efficient genome editing". *Nature Biotechnology* 29 (2): 143-8] (상기 문헌은 그 전문이 본원에서 참조로 포함된다)을 참조할 수 있다. 이러한 활성은 사용되는 중, 세포 유형, 표적 유전자, 및 뉴클레아제에 따라 달라질 수 있기 때문에, PCR에 의해 증폭되는 두 대립유전자 사이의 임의의 차이를 검출하는 이중이중체 절단 검정법을 이용함으로써 활성을 모니터링할 수 있다. 절단 생성물은 단순 아가로스 겔 또는 슬랩 겔 시스템 상에서 시각화될 수 있다.
- [0035] 별법으로, DNA는 외인성 이중 가닥 DNA 단편의 존재하에서 NHEJ를 통해 게놈 내로 도입될 수 있다. 형질감염된 이중 가닥 서열이 수복 효소에 대한 주형으로서 사용되는 바, 상동성 지정 수복 또한 DSB에 의해 DNA를 도입할 수 있다. 특정 측면에 따라, 본원에 기술된 TALEN을 이용하여 안정적으로 변형된 인간 배아 줄기 세포 및 유도

다능성 줄기 세포 (iPSC) 클론을 생성할 수 있다. 특정 측면에 따라, 본원에 기술된 TALEN을 이용하여 녹아웃 중, 예컨대, *C. elegans*, 녹아웃 래트, 녹아웃 마우스 또는 녹아웃 제브라피시를 생성할 수 있다.

[0036] 본 개시내용의 한 측면에 따라, 실시양태는 줄기 세포 내에서 DNA로 공동 국제화하고, DNA를 분해 또는 절단시키면서, 외인성 DNA를 삽입하기 위한, 외인성 DNA, 뉴클레아제 효소, 예컨대, DNA 결합 단백질 및 가이드 RNA의 용도에 관한 것이다. 상기 DNA 결합 단백질은 다양한 목적을 위해 DNA에 결합하는 것으로 통상의 기술자에게 쉽게 공지되어 있다. 상기 DNA 결합 단백질은 천연적으로 발생된 것일 수 있다. 본 개시내용의 범주 내에 포함되는 DNA 결합 단백질로는, 본원에서 가이드 RNA로 지칭되는 RNA에 의해 가이드될 수 있는 것을 포함한다. 상기 측면에 따라, 가이드 RNA 및 RNA에 의해 가이드되는 DNA 결합 단백질은 DNA에서 공동 국제화 복합체를 형성한다. 뉴클레아제 활성을 가진 상기 DNA 결합 단백질은 통상의 기술자에게 공지되어 있고, 이는 뉴클레아제 활성을 가지는 천연적으로 발생된 DNA 결합 단백질, 예를 들어, II형 CRISPR 시스템에 존재하는 Cas9 단백질과 같은 것을 포함한다. 상기 Cas9 단백질 및 II형 CRISPR 시스템은 관련 기술분야에 문서로 잘 기록되어 있다. 문헌 [Makarova et al., Nature Reviews, Microbiology, Vol. 9, June 2011, pp. 467-477] (모든 보충 정보 포함) (상기 문헌은 그 전문이 본원에서 참조로 포함된다)을 참조할 수 있다.

[0037] 뉴클레아제 활성을 가진 예시적인 DNA 결합 단백질은 이중 가닥 DNA를 니킹하거나, 또는 절단하는 기능을 한다. 상기 뉴클레아제 활성은 뉴클레아제 활성을 보이는 하나 이상의 폴리펩티드 서열을 가진 DNA 결합 단백질로부터 생성될 수 있다. 상기 예시적인 DNA 결합 단백질은, 각 도메인이 이중 가닥 DNA의 특정 가닥을 절단 또는 니킹하는 것을 담당하는 것인, 별개의 두 뉴클레아제 도메인을 가질 수 있다. 통상의 기술자에게 공지된 뉴클레아제 활성을 가진 예시적인 폴리펩티드 서열은 McrA-HNH 뉴클레아제 관련 도메인 및 RuvC-유사 뉴클레아제 도메인을 포함한다. 따라서, 예시적인 DNA 결합 단백질은 자연상에서 McrA-HNH 뉴클레아제 관련 도메인 및 RuvC-유사 뉴클레아제 도메인 중 하나 이상의 것을 함유하는 것이다.

[0038] 예시적인 DNA 결합 단백질은 II형 CRISPR 시스템의 RNA에 의해 가이드되는 DNA 결합 단백질이다. 예시적인 DNA 결합 단백질은 Cas9 단백질이다.

[0039] S. 피오게네스에서, Cas9는 단백질 중 두 촉매성 도메인: DNA의 상보적인 가닥을 절단하는 HNH 도메인 및 비-상보적인 가닥을 절단하는 RuvC-유사 도메인에 의해 매개되는 프로세스를 통해 프로토스페이서-인접 모티프 (PAM)로부터 3 bp만큼 상류쪽에 평활 말단 이중 가닥 절단을 생성한다. 문헌 [Jinke et al., Science 337, 816-821 (2012)] (상기 문헌은 그 전문이 본원에서 참조로 포함된다)를 참조할 수 있다. Cas9 단백질은 문헌 [Makarova et al., Nature Reviews, Microbiology, Vol. 9, June 2011, pp. 467-477]에 대한 보충 정보에서 확인되는 것과 같은 하기의 것을 비롯한 다수의 II형 CRISPR 시스템: 메타노코쿠스 마리팔루디스(*Methanococcus maripaludis*) C7; 코리네박테리움 디프테리아(*Corynebacterium diphtheriae*); 코리네박테리움 에피시엔스(*Corynebacterium efficiens*) YS-314; 코리네박테리움 글루타미쿰(*Corynebacterium glutamicum*) ATCC 13032 키타사토(Kitasato); 코리네박테리움 글루타미쿰 ATCC 13032 빌레펠트(Bielefeld); 코리네박테리움 글루타미쿰 R; 코리네박테리움 크롭펜스테드티(*Corynebacterium kroppenstedtii*) DSM 44385; 마이코박테리움 앵세스수스(*Mycobacterium abscessus*) ATCC 19977; 노카르디아 파르시니카(*Nocardia farcinica*) IFM10152; 로도코쿠스 에리트로폴리스(*Rhodococcus erythropolis*) PR4; 로도코쿠스 조스티(*Rhodococcus jostii*) RHA1; 로도코쿠스 오파쿠스(*Rhodococcus opacus*) B4 uid36573; 아시도써무스 셀룰로리티쿠스(*Acidothermus cellulolyticus*) 11B; 아르트로박터 클로로페놀리쿠스(*Arthrobacter chlorophenolicus*) A6; 크립벨라 플라비다(*Kribbella flavida*) DSM 17836 uid43465; 썬모모노스포라 쿠르바타(*Thermomonospora curvata*) DSM 43183; 비피도박테리움 덴티움(*Bifidobacterium dentium*) Bd1; 비피도박테리움 롱검(*Bifidobacterium longum*) DJ010A; 슬랙키아 헬리오트리니레두센스(*Slackia heliotrinireducens*) DSM 20476; 퍼셉포넬라 마리나(*Persephonella marina*) EX H1; 박테로이데스 프라질리스(*Bacteroides fragilis*) NCTC 9434; 카프노시토파가 오크라세아(*Capnocytophaga ochracea*) DSM 7271; 플라보박테리움 프시크로필룸(*Flavobacterium psychrophilum*) JIP02 86; 아커만시아 뮤시니필라(*Akkermansia muciniphila*) ATCC BAA 835; 로세이플렉수스 카스텐홀지(*Roseiflexus castenholzii*) DSM 13941; 로세이플렉수스 RS1; 시네코시스티스(*Synechocystis*) PCC6803; 엘루시마이크로비움 미누툼(*Elusimicrobium minutum*) Pei191; 미배양된 흰개미(Termite) 군 1 박테리움 계통형 Rs D17; 피프로박터 속시노게네스(*Fibrobacter succinogenes*) S85; 바실렛 세레우스(*Bacillus cereus*) ATCC 10987; 리스테리아 인노쿠아(*Listeria innocua*); 락토바실러스 카제이(*Lactobacillus casei*); 락토바실러스 람노수스(*Lactobacillus rhamnosus*) GG; 락토바실러스 살리바리우스(*Lactobacillus salivarius*) UCC118; 스트렙토코쿠스 아갈락티아에(*Streptococcus agalactiae*) A909; 스트렙토코쿠스 아갈락티아에 NEM316; 스트렙토코쿠스 아갈락티아에 2603; 스트렙토코쿠스 디스갈락티아에 이퀴시밀리스(*Streptococcus dysgalactiae equisimilis*) GGS 124; 스트

렙토코쿠스 이퀴 주에피데미쿠스(*Streptococcus equi zooepidemicus*) MGCS10565; 스트렙토코쿠스 갈로리티쿠스(*Streptococcus gallolyticus*) UCN34 uid46061; 스트렙토코쿠스 고르도니이 칼리스 섭스트(*Streptococcus gordonii Challis subst*) CH1; 스트렙토코쿠스 뮤탄스(*Streptococcus mutans*) NN2025 uid46353; 스트렙토코쿠스 뮤탄스; 스트렙토코쿠스 피오게네스 M1 GAS; 스트렙토코쿠스 피오게네스 MGAS5005; 스트렙토코쿠스 피오게네스 MGAS2096; 스트렙토코쿠스 피오게네스 MGAS9429; 스트렙토코쿠스 피오게네스 MGAS10270; 스트렙토코쿠스 피오게네스 MGAS6180; 스트렙토코쿠스 피오게네스 MGAS315; 스트렙토코쿠스 피오게네스 SSI-1; 스트렙토코쿠스 피오게네스 MGAS10750; 스트렙토코쿠스 피오게네스 NZ131; 스트렙토코쿠스 썬모필레스(*Streptococcus thermophiles*) CNRZ1066; 스트렙토코쿠스 썬모필레스 LMD-9; 스트렙토코쿠스 썬모필레스 LMG 18311; 클로스트리디움 보툴리눔(*Clostridium botulinum*) A3 로크 마리(Loch Maree); 클로스트리디움 보툴리눔 B 에클룬드(Eklund) 17B; 클로스트리디움 보툴리눔 Ba4 657; 클로스트리디움 보툴리눔 F 랑엘란드(Langeland); 클로스트리디움 셀룰로리티쿰(*Clostridium cellulolyticum*) H10; 피네골디아 마그나(*Fingoldia magna*) ATCC 29328; 유박테리움 렉탈레(*Eubacterium rectale*) ATCC 33656; 미코플라스마 갈리셉티쿰(*Mycoplasma gallisepticum*); 미코플라스마 모바일레(*Mycoplasma mobile*) 163K; 미코플라스마 페넬트란스(*Mycoplasma penetrans*); 미코플라스마 시노비아에(*Mycoplasma synoviae*) 53; 스트렙토바실러스 모닐리포르미스(*Streptobacillus moniliformis*) DSM 12112; 브라디리조비움(*Bradyrhizobium*) BTAi1; 니트로박터 함부르겐시스(*Nitrobacter hamburgensis*) X14; 로도슈도모나스 팔루스트리스(*Rhodopseudomonas palustris*) BisB18; 로도슈도모나스 팔루스트리스 BisB5; 파르비바쿨럼 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; 네이세리아 메닝기티데스(*Neisseria meningitides*) 053442; 네이세리아 메닝기티데스 알파14; 네이세리아 메닝기티데스 Z2491; 데술포비브리오 살렉시겐스(*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*); 올리넬라 숙시노게네스(*Wolinella succinogenes*); 톨루모나스 아우엔시스(*Tolomonas auensis*) DSM 9187; 슈도알테로모나스 아틀란티카(*Pseudoalteromonas atlantica*) T6c; 쉬와넬라 페알레아나(*Shewanella pealeana*) ATCC 700345; 레지오넬라 뉴모필라(*Legionella pneumophila*) 파리시(Paris); 악티노바실러스 숙시노게네스(*Actinobacillus succinogenes*) 130Z; 파스퇴렐라 물토시다(*Pasteurella multocida*); 프란시셀라 투라렌시스 노비시다(*Francisella tularensis novicida*) U112; 프란시셀라 투라렌시스 홀라르티카(*Francisella tularensis holarctica*); 프란시셀라 투라렌시스 FSC 198; 프란시셀라 투라렌시스 투라렌시스(*Francisella tularensis tularensis*); 프란시셀라 투라렌시스 WY96-3418; 및 트레포네마 덴티콜라(*Treponema denticola*) ATCC 35405에 존재하는 것으로 공지되어 있다. 따라서, 본 개시내용의 측면은 II형 CRISPR 시스템에 존재하는 Cas9 단백질에 관한 것이다.

[0040] Cas9 단백질은 통상의 기술자에 의해 문헌에서 Csn1로 지칭될 수 있다. S. 피오게네스 Cas9 단백질은 하기에 제시되어 있다. 문헌 [Deltcheva et al., Nature 471, 602-607 (2011)] (상기 문헌은 그 전문이 본원에서 참조로 포함된다)을 참조할 수 있다.

MDKKYSIGLDIGTNSVGVAVITDEYKVPSSKFKVLGNTDRHSIKKNLIGALLFDSGETAE
 ATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHERHPIFG
 NIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSD
 VDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLLENLIAQLPGEKKNGLFGN
 LIALSLGLTPNFKSNFDLAEDAQLQSKDQYDDDLNLLAQIGDQYADLFLAAKNLSDAI
 LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYA
 GYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELH
 AILRRQEDFYFPLKDNREKIEKILTRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEE
 VVDKGASAQSFIERMTNFDKNLPNEKVLPHKSHLLYEYFTVYNELTKVKYVTEGMRKPAFL
 SGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI
 IKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWG
 RLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMLIHDDSLTFKEDIQKAQVSGQGDSL
 HEHIANLAGSPAIKKILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRER
 MKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVDH
 IVPQSFLKDDSIDNKVLRSDKNRGSNDNPSEEVVKKMKNYWRQLLNAKLITQRKFDNL
 TKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS
 KLVSDFRKDFQFYKREINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVDVDR
 K
 MIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDF
 ATVRKVLSPQVNIIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKGFFDSPTVA
 YSVLVVAKVEKGSKLLKSVKELLGITIMERSSEKPNIDFLEAKGYKEVKKDLIILPK
 YSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLGSPEDNEQQLFVE
 QHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKHRDKPIREQAENIIHLFTLTNLGA
 PAAFKYFDTTIDRKRYTSTKEVLDATLIHQSIITGLYETRIDLSQLGGD-

[0041]

[0042]

한 측면에 따라, RNA에 의해 가이드되는 DNA 결합 단백질은 DNA에 결합할 수 있고, RNA에 의해 가이드될 수 있고, DNA를 절단할 수 있는 단백질의 능력을 유지하는 Cas9의 호모로그 및 오르토로그를 포함한다. 한 측면에 따라, Cas9 단백질은 S. 피오케네스로부터의 천연적으로 발생된 Cas9에 대해 기재된 것과 같은 서열, 및 그와 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 98% 또는 99% 이상 상동성이고, DNA 결합 단백질, 예컨대, RNA에 의해 가이드되는 DNA 결합 단백질인 것인 단백질 서열을 포함한다.

[0043]

한 측면에 따라, 줄기 세포에서 부위 특이적인 방식으로 RNA에 의해 가이드되는 게놈 절단을 수행할 수 있고, 원하는 경우, 외인성 도너 핵산의 삽입에 의해 줄기 세포 게놈을 변형시킬 수 있는 조작된 Cas9-gRNA 시스템을 제공한다. 가이드 RNA는 DNA 상의 표적 부위 또는 표적 유전자좌에 상보적이다. 가이드 RNA는 crRNA-tracrRNA 키메라일 수 있다. 가이드 RNA는 세포 주변의 배지로부터 도입될 수 있다. 이러한 방식으로 각종 가이드 RNA를 주변 배지에 제공하고, 세포에 의해 가이드 RNA가 흡수되고, 배지를 추가의 가이드 RNA로 보충하는 정도로 세포를 연속적으로 변형시키는 방법을 제공한다. 보충은 연속적인 방식으로 이루어질 수 있다. Cas9는 표적 게놈 DNA에 또는 그 주변에 결합한다. 하나 이상의 가이드 RNA는 표적 게놈 DNA에 또는 그 주변에 결합한다. Cas9는 표적 게놈 DNA를 절단하고, 외인성 도너 DNA가 DNA 내로 그 절단 부위에 삽입된다.

[0044]

따라서, 본 방법은 RNA를 코딩하는 (또는 주변 배지로부터 RNA를 제공하는) 핵산 및 외인성 도너 핵산을 삽입하고, RNA를 발현하고 (또는 RNA를 흡수하고), DNA를 절단하는 방식으로 RNA, Cas9 및 DNA를 공동 국재화시키고, 외인성 도너 핵산을 삽입하는 것을 사이클링함으로써 Cas9를 발현하는 줄기 세포 내에서 외인성 도너 핵산을 DNA 내로 다중으로 삽입하기 위한 Cas9 단백질 및 외인성 도너 핵산과 함께 가이드 RNA의 용도에 관한 것이다. 본 방법 단계는 임의의 원하는 개수만큼 DNA를 변형시킬 수 있도록 임의의 원하는 횟수로 사이클링될 수 있다. 따라서, 본 개시내용의 방법은 본원에 개시된 Cas9 단백질 및 가이드 RNA를 사용하여 표적 유전자를 편집하여 줄기 세포를 다중으로 유전적 및 후성적으로 조작하는 것에 관한 것이다.

[0045]

추가로, 본 개시내용의 측면은 일반적으로 줄기 세포, 예컨대, 인간 줄기 세포의 DNA, 예컨대, 게놈 DNA 내로의

외인성 도너 핵산의 다중 삽입을 위한 DNA 결합 단백질 또는 시스템 (예컨대, 본원에 기술된 변형된 TALENS 또는 Cas9)의 용도에 관한 것이다. 통상의 기술자는 본 개시내용에 기초하여 예시적인 DNA 결합 시스템을 쉽게 확인할 수 있을 것이다.

[0046] 달리 명시되지 않는 한, 본 개시내용에 따른 세포로는 본원에 기술된 바와 같이 외래 핵산이 그 내부로 도입될 수 있고, 거기서 발현될 수 있는 임의의 세포를 포함한다. 본원에 기술된 본 개시내용의 기본 개념은 세포 유형에 의해 한정되지 않는다는 것을 이해하여야 한다. 본 개시내용에 따른 세포로는 체세포, 줄기 세포, 진핵 세포, 원핵 세포, 동물 세포, 식물 세포, 진균 세포, 고세균 세포, 진정세균 세포 등을 포함한다. 세포로는 진핵 세포 예컨대, 효모 세포, 식물 세포, 및 동물 세포를 포함한다. 특정 세포로는 포유동물 세포, 예컨대, 인간 세포를 포함한다. 추가로, 세포로는 DNA를 변형시키는 것이 유익하거나, 바람직한 임의의 것을 포함한다.

[0047] 표적 핵산으로는, 본원에 기술된 바와 같은 뉴클레아제 활성을 가지는 TALEN 또는 RNA에 의해 가이드되는 DNA 결합 단백질이 그를 니킹 또는 절단하는 데 유용할 수 있는 것인 임의의 핵산 서열을 포함한다. 표적 핵산으로는 본원에 기술된 바와 같은 공동 국제화 복합체가 그를 니킹 또는 절단하는 데 유용할 수 있는 것인 임의의 핵산 서열을 포함한다. 표적 핵산으로는 유전자를 포함한다. 본 개시내용의 목적을 위해, DNA, 예컨대, 이중 가닥 DNA는 표적 핵산을 포함할 수 있고, 공동 국제화 복합체는 DNA에 결합할 수 있거나, 또는 다르게는 그와 공동 국제화할 수 있거나, 또는 TALEN는 다르게는 공동 국제화 복합체 또는 TALEN이 표적 핵산에 대하여 원하는 효과를 발휘할 수 있는 방식으로 표적 핵산에서 또는 그에 인접하게 또는 그 부근에서 DNA에 결합할 수 있다. 상기 표적 핵산은 내인성 (또는 천연적으로 발생된) 핵산 및 외인성 (또는 외래) 핵산을 포함할 수 있다. 본 개시내용에 기초하여 통상의 기술자는 DNA로 공동 국제화하는 가이드 RNA 및 Cas9 단백질, 또는 DNA에 결합하는 TALEN을 쉽게 확인하거나, 또는 디자인할 수 있을 것이다. 통상의 기술자는 추가로, 유사하게, 표적 핵산을 비롯한 DNA로 공동 국제화하는 전사 조절인자 단백질 또는 도메인, 예컨대, 전사 활성인자 또는 전사 억제인자를 확인할 수 있을 것이다. DNA로는 게놈 DNA, 미토콘드리아 DNA, 바이러스 DNA 또는 외인성 DNA를 포함한다. 한 측면에 따라, 본 개시내용의 실시에서 유용한 물질 및 방법으로는 예시적인 균주 및 배지, 플라스미드 구축, 플라스미드 형질전환, 일과성 gRNA 카세트 및 도너 핵산의 전기천공, 도너 DNA를 포함하는 gRNA 플라스미드의 Cas9-발현 세포로의 형질전환, Cas9의 갈락토스 유도, 효모 게놈에서의 CRISPR-Cas 표적 확인 등을 비롯한, 문헌 [Di Carlo, et al., *Nucleic Acids Research*, 2013, vol. 41, No. 7 4336-4343] (상기 문헌은 그 전문이 모든 목적을 위해 본원에서 참조로 포함된다)에 기술되어 있는 것을 포함한다. 본 발명을 수행하는 데 있어 통상의 기술자에게 유용한 정보, 물질 및 방법을 비롯한 추가의 참조 사항은 문헌 [Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E. and Church, G.M. (2013) RNA-Guided human genome engineering via Cas9. *Science*, 10.1126/science.1232033]; [Storici, F., Durham, C.L., Gordenin, D.A. and Resnick, M.A. (2003) Chromosomal site-specific double-strand breaks are efficiently targeted for repair by oligonucleotides in yeast. *PNAS*, 100, 14994-14999] 및 [Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A. and Charpentier, E. (2012) A programmable dual-RNA-Guided DNA endonuclease in adaptive bacterial immunity. *Science*, 337, 816-821] (상기 문헌들은 각각 그 전문이 모든 목적을 위해 본원에서 참조로 포함된다)에 제공되어 있다.

[0048] 외래 핵산 (즉, 세포의 천연 핵산 조성의 일부가 아닌 것)은 상기 도입을 위해 통상의 기술자에게 공지된 임의의 방법을 사용하여 세포 내로 도입될 수 있다. 상기 방법으로는 형질감염, 형질도입, 바이러스 형질도입, 미세주입, 리포펙션, 뉴클레오펙션, 나노입자 충격법, 형질전환, 접합 등을 포함한다. 통상의 기술자는 쉽게 확인 가능한 문헌 자료를 이용하여 상기 방법을 쉽게 이해하고, 적합화시킬 것이다.

[0049] 도너 핵산은 본원에 기술된 바와 같이 핵산 서열 내로 삽입되는 임의의 핵산을 포함한다.

[0050] 하기 실시예는 본 개시내용을 대표하는 것으로서 기술되는 것이다. 상기 실시양태 및 다른 등가의 실시양태는 본 개시내용, 도면 및 첨부된 청구범위를 고려하면 자명한 바, 본 실시예는 본 개시내용의 범주를 한정하는 것으로 해석되지 않아야 한다.

[0051] 실시예 I

[0052] 가이드 RNA 조립

[0053] 선택된 표적 서열의 19 bp (즉, 5'-N19-NGG-3'의 5'-N19)를 2개의 상보적인 100 mer 올리고뉴클레오티드 내로 도입하였다 (TTCTTGGCTTTATATATCTGTGGAAAGGACGAAACACCGN19GTTTGTAGAGCTAGAAATAGCAAGTTAAATAAGGCTAGTCC). 각 100 mer 올리고뉴클레오티드를 물 중에 100 mM으로 현탁시키고, 동량으로 혼합하고, 씨모사이클 장치 (95°C, 5

min; 4℃까지, 0.1℃/sec의 경사로)에서 어닐링시켰다. 목표 벡터를 제조하기 위해, AfIII를 이용하여 gRNA 클로닝 벡터 (애드진(Addgene) 플라스미드 ID 41824)를 선형화시키고, 벡터를 정제하였다. 10 ng 어닐링된 100 bp 단편, 100 ng 목표 골격, 1X 깁슨(Gibson) 조립 반응 믹스 (뉴 잉글랜드 바이오랩스(New England Biolabs))를 이용하여 50℃에서 30 min 동안 (10 ul) gRNA 조립 반응을 수행하였다. 박테리아 형질전환을 위해서 직접 본 반응을 프로세싱하여 개별 조립체를 전이증식(colonize)시킬 수 있다.

[0054] 실시예 II

[0055] 재코딩된 TALE 디자인 및 조립

[0056] 조립을 촉진시키고, 발현을 개선시키기 위해 리TALE를 다른 수준으로 최적화시켰다. 먼저 리TALE DNA 서열을 인간 코돈 선호도, 및 5' 단부의 낮은 mRNA 폴딩 에너지에 맞게 함께 최적화시켰다 (진GA(GeneGA), 바이오컨덕터(Bioconductor)). 수회의 사이클을 거쳐 수득된 서열을 진화시켜 11 bp 초과 반복부 (동방향 또는 역방향)를 제거하였다 (도 12 참조). 각 사이클에서 각 반복부에 대한 동의적 서열을 평가하였다. 진화 DNA까지의 해밍 거리가 가장 큰 것을 선별하였다. 16.5개의 단량체를 가지는 리TALE 중 하나의 서열은 하기와 같다:

```
CTAACCCCTGAACAGGTAGTCGCTATAGCTTCAAATATCGGGGGCAAGCAAGCACTTG
AGACCGTTCAACGACTCCTGCCAGTGCTCTGCCAAGCCATGGATTGACTCCGGAGCA
AGTCGTCGCGATCGCGAGCAACGGCGGGGGGAAGCAGGCGCTGGAAACTGTTTCAGAG
ACTGCTGCCTGTACTTTGTCAGGCGCATGGTCTCACCCCCGAACAGTTGTCGCAATA
GCAAGTAATATAGGCGGTAAGCAAGCCCTAGAGACTGTGCAACGCCTGCTCCCCGTGC
TGTGTCAGGCTCACGGTCTGACACCTGAACAAGTTGTCGCGATAGCCAGTCACGACGG
GGGAAAACAAGCTCTAGAAACGGTTCAAAGTTGTTGCCCGTTCTGTGCCAAGCACAT
GGGTTAACACCCGAACAAGTAGTAGCGATAGCGTCAAATAACGGGGGTAAACAGGCT
TTGGAGACGGTACAGCGGTTATTGCCGGTCTCTGCCAGGCCACGACTTACGCCAG
AACAGGTGGTTGCAATTGCCTCCAACATCGGCGGGAAACAAGCGTTGGAAACTGTGC
AGAGACTCCTTCTGTTTTGTGTCAGCCACGGCTTGACGCCTGAGCAGGTTGTGGC
CATCGCTAGCCACGACGGAGGGAAGCAGGCTCTTGAAACCGTACAGCGACTTCTCCCA
GTTTTGTGCCAAGCTCACGGGCTAACCCCCGAGCAAGTAGTTGCCATAGCAAGCAACG
GAGGAGGAAAACAGGCATTAGAAAACAGTTCAGCGCTTGCTCCCGGTAAGTCTGTGTCAGG
CACACGGTCTAACTCCGGAACAGGTCGTAGCCATTGCTTCCCATGATGGCGGCAAACA
GGCGCTAGAGACAGTCCAGAGGCTCTTGCCGTGTGTTATGCCAGGCACATGGCCTCACC
CCGGAGCAGGTCGTTGCCATCGCCAGTAATATCGGCGGAAAGCAAGCTCTCGAAACA
GTACAACGGCTGTTGCCAGTCTATGTCAAGCTCATGGACTGACGCCCAGCAGGTTAG
TGGCAATCGCATCTCACGATGGAGGTAACAAGCACTCGAGACTGTCCAAAGATTGTT
ACCCGTAAGTATGCCAAGCGCATGGTTTAAACCCAGAGCAAGTTGTGGCTATTGCATCT
AACGGCGGTGGCAAACAAGCCTTGGAGACAGTGCAACGATTACTGCCTGTCTTATGTC
AGGCCCATGGCCTTACTCCTGAGCAAGTCGTAGCTATCGCCAGCAACATAGGTGGGAA
ACAGGCCCTGGAAACCGTACAACGTCTCCTCCAGTACTTTGTCAAGCACACGGGTTG
ACACCGGAACAAGTGGTGGCGATTGCGTCCAACGGCGGAGGCAAGCAGGCACTGGAG
ACCGTCCAACGGCTTCTTCCGGTCTTTGCCAGGCTCATGGGCTCACGCCAGAGCAGG
TGGTAGCAATAGCGTCGAACATCGGTGGTAAGCAAGCGCTTCAAACGGTCCAGCGTCT
TCTGCCGGTGTGTGCCAGGCGCACGGACTCACACCAGAACAAGTGGTTGCTATTGCT
AGTAACAACGGTGGAAAGCAGGCCCTCGAGACGGTGCAGAGGTTACTTCCCGTCTCTCT
GTCAAGCGCACGGCCTCACTCCAGAGCAAGTGGTTGCGATCGCTTCAAACAATGGTGG
AAGACCTGCCCTGGAA
```

[0057]

[0058] 특정 측면에 따라, 상기 서열과 80% 이상의 서열 동일성, 85% 이상의 서열 동일성, 90% 이상의 서열 동일성, 95% 이상의 서열 동일성, 98% 이상의 서열 동일성, 또는 99% 이상의 서열 동일성을 가지는 TALE가 사용될 수 있

다. TALE의 DNA 결합 활성은 그대로 유지하면서, 상기 서열은 달라질 수 있는 경우를 통상의 기술자는 쉽게 이해할 것이다.

[0059] 표준 카파 HIFI (KPAP) PCR 조건하에, 1회차 PCR에서는 RVD 코딩 서열을 도입하였고, 2회차 PCR에서는 인접 블록과의 36 bp의 중복부를 포함하는 전체 이량체 블록을 생성한, 2회에 걸친 PCR에 의해 2개의 RVD를 코딩하는 리TALE 이량체 블록 (도 6a 참조)을 생성하였다. QIA퀵(QIAquick) 96 PCR 정제 키트 (퀴아젠(QIAGEN))를 이용하여 PCR 생성물을 정제하고, 나노-드롭(Nano-drop)에 의해 농도를 측정하였다. 프라이머 및 주형 서열은 하기 표 1 및 표 2에 열거되어 있다.

[0060] <표 1>

리TALE 블록 서열

블록 0	CGCAATGCGCTCACGGGAGCACCCCTCAACCTAACCCCTGAACAGGTA GTCGCTATAGCTTCANNNNNNGGGGGCAAGCAAGCACTTGAGACCGT TCAACGACTCCTGCCAGTGCTCTGCCAAGCCCATGGATTGACTCCGGA GCAAGTCGTCGCGATCGCGAGCNNNNNNNGGGGGGAAGCAGGCGCTGG
------	---

[0061]

	AAACTGTTGAGAGACTGCTGCCTGTACTTTGTGTCAGGCGCATGGTCTC
블록 1	AGACTGCTGCCTGTACTTTGTGTCAGGCGCATGGTCTACCCCCGAACAG GTTGTGCGCAATAGCAAGTNNNNNNGGCGGTAAGCAAGCCCTAGAGAC TGTGCAACGCCTGCTCCCCGTGCTGTGTCAGGCTCACGGTCTGACACC TGAACAAGTTGTGCGGATAGCCAGTNNNNNNGGGGAAAAACAAGCTC TAGAAAACGGTTCAAAGGTTGTTGCCCGTTCTGTGCCAAGCACATGGGT TA
블록 1'	TGCGCTCACGGGAGCACCCCTCAACCTCACCCCCGAACAGGTTGTGCG AATAGCAAGTNNNNNNGGCGGTAAGCAAGCCCTAGAGACTGTGCAAC GCCTGCTCCCCGTGCTGTGTCAGGCTCACGGTCTGACACCTGAACAAG TTGTGCGGATAGCCAGTNNNNNNGGGGAAAAACAAGCTCTAGAAAACG GTTCAAAGGTTGTTGCCCGTTCTGTGCCAAGCACATGGGTTA
블록 2	AGGTTGTTGCCCGTTCTGTGCCAAGCACATGGGTAAACACCCGAACAA GTAGTAGCGATAGCGTCANNNNNNGGGGTAACAGGCTTTGGAGAC GGTACAGCGGTTATTGCCGGTCTCTGCCAGGCCACGGACTTACGCC AGAACAGGTGGTTGCAATTGCCTCCNNNNNNGGCGGAAAAACAAGCGT TGAAACTGTGCAGAGACTCCTTCTGTTTTGTGTCAGGCCACGGCTT GACGCCT
블록 3	AGACTCCTTCCTGTTTTGTGTCAGGCCACGGCTTGACGCCTGAGCAG GTTGTGGCCATCGCTAGCNNNNNNGGAGGGAAGCAGGCTCTTGAAAC CGTACAGCGACTTCTCCAGTTTTGTGCCAAGCTCACGGGCTAACCCC CGAGCAAGTAGTTGCCATAGCAAGCNNNNNNGGAGGAAAAACAGGCAT TAGAAAACAGTTCAGCGCTTGTCCCCGTTACTCTGTGTCAGGCACACGGTC TA
블록 4	CGCTTGCTCCCGTACTCTGTGTCAGGCACACGGTCTAACTCCGGAACAG GTCGTAGCCATTGCTTCCNNNNNNGGCGGCAAACAGGCGCTAGAGAC CGTCCAGAGGCTCTTGCTGTGTTATGCCAGGCACATGGCCTCACCCC GGAGCAGGTCGTTGCCATCGCCAGTNNNNNNGGCGGAAAAGCAAGCTC TCGAAAACAGTACAACGGCTGTTGCCAGTCCTATGTCAAGCTCATGGAC TG
블록 5	CGGCTGTTGCCAGTCCTATGTCAAGCTCATGGACTGACGCCCCGAGCAG GTAGTGGCAATCGCATCTNNNNNNGGAGGTAACAAGCACTCGAGAC TGTCAAAAGATTGTTACCCGTTACTATGCCAAGCGCATGGTTTAAACCC

[0062]

	AGAGCAAGTTGTGGCTATTGCATCTNNNNNNGGTGGCAAACAAGCCTT GGAGACCGTGCAACGATTACTGCCTGTCTTATGTCAGGCCCATGGCCT T
블록 6	CGATTACTGCCTGTCTTATGTCAGGCCCATGGCCTTACTCCTGAGCAGG TGGTCGCTATCGCCAGCNNNNNNGGGGCAAGCAAGCACTGGAAACA GTCCAGCGTTTGTCTCCAGTACTTTGTCAGGCGCATGGATTGACACCG GAACAAGTGGTGGCTATAGCCTCANNNNNNGGAGGAAAGCAGGCGCT GGAAACCGTCCAACGTCTTTTACCGGTGCTTTGCCAGGCGCACGGGCT C
블록 6'	CGATTACTGCCTGTCTTATGTCAGGCCCATGGCCTTACTCCTGAGCAAG TCGTAGCTATCGCCAGCNNNNNNGGTGGGAAACAGGCCCTGGAAACC GTACAACGTCTCCTCCCAGTACTTTGTCAAGCACACGGGTTGACACCG GAACAAGTGGTGGCGATTGCGTCCNNNNNNGGAGGCAAGCAGGCACT GGAGACCGTCCAACGGCTTCTTCCGGTTCTTTGCCAGGCTCATGGGCT C
블록 7	CGGCTTCTTCCGGTTCTTTGCCAGGCTCATGGGCTCACGCCAGAGCAG GTGGTAGCAATAGCGTCGNNNNNNGGTGGTAAGCAAGCGCTTGAAC GGTCCAGCGTCTTCTGCCGGTGTGTGCCAGGCGCACGGACTCACACC AGAACAAGTGGTTGCTATTGCTAGTNNNNNNGGTGGAAAGCAGGCC TCGAGACGGTGCAGAGGTTACTTCCCGTCCTCTGTCAAGCGCACGGCC TC

[0063]

[0064] <표 2>

리TALE 블록 프라이머 서열

블록 0-F	CGCAATGCGCTCACGGGAGCACCCCTCAACctAACCCCTGAACAGGT* A*G
블록 0-R	GAGACCATGCGCCTGACAAAGTACAGGCAGCAGTCTCTGAACAG*T* T
블록 1'-F	TGGCGCAATGCGCTCACGGGAGCACCCCTCA*A*C
블록 1-F	AGACTGCTGCCTGTACTTTGTGTCAGGCGCATGGTCTCACCCCGAACA* G*G
블록 1-R/ 블록 1'-R	TAACCCATGTGCTTGGCACAGAACGGGCAACAACCTTTGAACCG*T*T
블록 2-F	AGGTTGTTGCCCGTTCTGTGCCAAGCACATGGGTTAACACCCgaac*a*a
블록 2-R	AGGCGTCAAGCCGTGGGCTTGACACAAAACAGGAAGGAGTCTCTGCA CAG*T*t
블록 3-F	AGACTCCTTCCTGTTTTGTGTCAAGCCACGGCTTGACGCCTG*A*G
블록 3-R	TAGACCGTGTGCCTGACAGAGTACCGGGAGCAAGCGCT*G*A
블록 4-F	CGCTTGCTCCCGGTACTCTGTCAGGCACACGGTCTAA*C*T
블록 4-R	CAGTCCATGAGCTTGACATAGGACTGGCAACAGCCGTT*G*T
블록 5-F	CGGCTGTTGCCAGTCCTATGTCAAGCTCATGGACTGA*C*G
블록 5-R	AAGGCCATGGGCCTGACATAAGACAGGCAGTAATCGTT*G*C
블록 6-F	CGATTACTGCCTGTCTTATGTCAGGCCCATGGCCTTA*C*T
블록 6-R	GAGCCCGTGCCTGGCAAAGCACCGGTAAAAGACGTTGGA*C*G
블록 6'-F	CGATTACTGCCTGTCTTATGTCAGGCCCATGGCCTTACTCCTGAGCAA *G*T
블록 6'-R	GAGCCCATGAGCCTGGCAAAGAACCGGAAGAAGCCGTT*G*G
블록 7-F	CGGCTTCTCCCGTTCTTTGCCAGGCTCATGGGCTCACGCCAGAGCAG G*T*G
블록 7-R	GAGGCCGTGCGCTTGACAGAGGACGGGAAGTAACCTCT*G*C

[0065]

[0066]

TALE-TF 및 TALEN 클로닝 골격을 변형시킴으로써 리TALEN 및 리TALE-TF 목표 벡터를 구축하였다 (참고 문헌 24 (상기 문헌은 그 전문이 본원에서 참조로 포함된다) 참조). 벡터 상의 0.5 RVD 영역을 채코딩하고, SapI 절단 부위를 지정된 리TALE 클로닝 부위에 도입하였다. 리TALEN 및 리TALE-TF 골격의 서열은 도 20에 제공되어 있다. 플라스미드를 제조사가 권고한 조건에 따라 SapI (뉴 잉글랜드 바이오랩스)로 미리 처리하고, QIA퀵 PCR 정제용 키트 (키아젠)로 정제할 수 있다.

[0067]

200 ng의 각 블록, 500 ng 목표 골격, 1X TASA 효소 혼합물 (2 U SapI, 100 U 앰프리가제(Ampligase) (에피상트르(Epicentre)), 10 mU T5 엑소뉴클레아제 (에피상트르), 2.5 U 퓨전(Phusion) DNA 폴리머라제 (뉴 잉글랜드 바이오랩스)) 및 상기 기술된 바와 같은 1X 등은 조립 반응 완충제 (참고 문헌 25 (상기 문헌은 그 전문이 본원에서 참조로 포함된다) 참조) (5% PEG-8000, 100 mM 트리스-HCl (pH 7.5), 10 mM MgCl₂, 10 mM DTT, 0.2 mM의 4가지의 dNTP 각각 및 1 mM NAD)를 이용하여 (10 ul) 원-포트 TASA 조립 반응을 수행하였다. 37°C에서 5 min 및 50°C에서 30 min 동안 인큐베이션을 수행하였다. 박테리아 형질전환을 위해 직접 TASA 조립 반응을 프로세싱하여 개별 조립체를 전이증식시킬 수 있다. 본 접근법을 이용한 경우, 전장의 구축물의 수득률은 ~20%이다. 별법으로, >90% 효율은 3단계 조립에 의해 달성될 수 있다. 먼저, 50°C에서 30 min 동안 200 ng의 각 블록, 1X 리TALE 효소 혼합물 (100 U 앰프리가제, 12.5 mU T5 엑소뉴클레아제, 2.5 U 퓨전 DNA 폴리머라제) 및 1X 등은 조립 완충제를 이용하여 10 ul 리TALE 조립 반응을 수행한 후, 표준화된 카파 HIFI PCR 반응, 아가로스 겔 전기

영동, 및 QIA퀵 겔 추출 (키아젠(Qiagen))을 수행하여 전장의 리TALE를 농축시킨다. 이어서, 200 ng 리TALE 엠플리콘을 500 ng의, SapI로 미리 처리된 목표 골격, 1X 리TALE 조립 혼합물 및 1X 등은 조립 반응 완충제와 함께 혼합하고, 50 °C에서 30 min 동안 인큐베이션시킬 수 있다. 박테리아 형질전환을 위해 직접 리TALE 최종 조립 반응을 프로세싱하여 개별 조립체를 전이증식시킬 수 있다. 통상의 기술자는 공지된 것으로부터 엔도뉴클레아제, 엑소뉴클레아제, 폴리머라제 및 리가제를 쉽게 선택함으로써 본원에 기술된 방법을 실시할 수 있다. 예를 들어, II형 엔도뉴클레아제, 예컨대, FokI, BtsI, EarI, SapI가 사용될 수 있다. 적정가능한 엑소뉴클레아제, 예컨대, 람다 엑소뉴클레아제, T5 엑소뉴클레아제 및 엑소뉴클레아제 III이 사용될 수 있다. 비-핫스타트 (hotstart) 폴리머라제, 예컨대, 퓨전 DNA 폴리머라제, Taq DNA 폴리머라제 및 VentR DNA 폴리머라제가 사용될 수 있다. 열안정 리가제, 예컨대, 앰프리가제, pfu DNA 리가제, Taq DNA 리가제가 본 반응에서 사용될 수 있다. 추가로, 사용되는 특정 종에 따라 상기 엔도뉴클레아제, 엑소뉴클레아제, 폴리머라제 및 리가제를 활성화시키는 데 상이한 반응 조건이 사용될 수 있다.

- [0068] 실시예 III
- [0069] 세포주 및 세포 배양
- [0070] PGP1 iPS 세포를 mTeSR1 (스텝셀 테크놀로지스 Technologies(Stemcell Technologies)) 중 매트릭셀 (BD 바이오 사이언시즈(BD Biosciences))-코팅 플레이트 상에 유지시켰다. 배양물을 매 5-7일마다 TrypLE 익스프레스 (TrypLE Express) (인비트로젠(Invitrogen))로 계대접종하였다. 293T 및 293FT 세포를 10% 우태아 혈청 (FBS, 인비트로젠), 페니실린/스트렙토마이신 (pen/strep, 인비트로젠), 및 비필수 아미노산 (NEAA, 인비트로젠)으로 보충된 둘베코스 변형된 이글즈 배지 (DMEM, 인비트로젠) 고 글루코스 중에서 성장시키고 유지시켰다. K562 세포는 10% 우태아 혈청 (FBS, 인비트로젠 15%) 및 페니실린/스트렙토마이신 (pen/strep, 인비트로젠)으로 보충된 RPMI (인비트로젠) 중에서 성장시키고 유지시켰다. 모든 세포를 습윤화된 인큐베이터 중 37°C 및 5% CO2에서 유지시켰다.
- [0071] HDR 효율을 검출하기 위해 안정한 293T 세포주를 참고 문헌 26 (상기 문헌은 그 전문이 본원에서 참조로 포함된다)에 기술되어 있는 바와 같이 확립시켰다. 구체적으로, 리포터 세포주는 종결 코돈의 삽입에 의해 파괴된 게놈적으로 통합된 GFP 코딩 서열 및 AAVS1 유전자좌로부터 유래된 68 bp 게놈 단편을 포함한다.
- [0072] 실시예 IV
- [0073] 리TALEN 활성 시험
- [0074] 293T 리포터 세포를 24 웰 플레이트 중 2×10^5 개의 세포/웰의 밀도로 시딩하고, 리포펙타민(Lipofectamine) 2000을 사용하여 제조사의 프로토콜에 따라 1 μ g의 각각의 리TALEN 플라스미드 및 2 μ g DNA 도너 플라스미드로 형질감염시켰다. 형질감염 후 ~18 h 경과하였을 때, TrypLE 익스프레스 (인비트로젠)를 사용하여 세포를 수거하고, LSRFortessa 세포 분석기 (BD 바이오사이언시즈)를 사용하여 유세포 분석법 분석을 위해 200 μ l 배지 중에 재현탁시켰다. 유세포 분석법 데이터를 플로우조(FlowJo) (플로우조)를 이용하여 분석하였다. 각 형질감염 샘플당 25,000개 이상의 이벤트를 분석하였다. 293T에서의 내인성 AAVS1 유전자좌 표적화 실험을 위해, 형질감염 방법은 상기 기술된 바와 동일하였고, 형질감염 후 1주 경과하였을 때, 3 μ g/ml의 약물 농도를 이용하여 플로마이신 선별을 수행하였다.
- [0075] 실시예 V
- [0076] 기능성 렌티바이러스 생성 평가
- [0077] 표준 PCR 및 클로닝 기법에 의해 렌티바이러스 벡터를 생성하였다. 렌티바이러스 플라스미드를 리포펙타민 2000에 의해 렌티바이러스 패키징 믹스 (인비트로젠)를 이용하여 293FT 세포 (인비트로젠) 내로 형질감염시켜 렌티바이러스를 생성하였다. 형질감염 후 48 및 72 h 경과하였을 때, 상청액을 수집하고, 멸균 여과시키고, 100 μ l 여과된 상청액을 폴리브렌과 함께 5×10^5 개의 신선한 293T 세포에 첨가하였다. 하기 공식: 바이러스 적정량 = (GFP+ 293T 세포의 비율(%)) * 형질도입하에 있는 초기 세포 개수)/(형질도입 실험에 사용된, 원래의 바이러스 수집 상청액 부피)에 기초하여 렌티바이러스 적정량을 계산하였다. 렌티바이러스의 기능성을 시험하기 위하여, 형질감염 후 3일 경과하였을 때, 렌티바이러스가 형질도입된 293T 세포를 리포펙타민 2000 (인비트로젠)을 이용하여 mCherry 리포터를 포함하는 30 ng 플라스미드 및 500 ng pUC19 플라스미드로 형질감염시켰다. 형질감염 후 18시간 경과하였을 때, 악시오 옵저버(Axio Observer) Z.1 (자이스(Zeiss))을 이용하여 세포 영상을 분석하고, TrypLE 익스프레스 (인비트로젠)를 이용하여 수거하고, LSR포르테사(LSRFortessa) 세포 분석

기 (BD 바이오사이언시즈)를 이용하여 유세포 분석법 분석을 수행하기 위해 200 μ l 배지 중에 재현탁시켰다. BD FACS디바 (BD 바이오사이언시즈)를 이용하여 유세포 분석법 데이터를 분석하였다.

[0078] 실시예 VI

[0079] 리TALEN 및 Cas9-gRNA 게놈 편집 효율 시험

[0080] 뉴클레오펙션 수행하기 2 h 전에 PGP1 iPSC를 Rho 키나제 (ROCK) 억제제 Y-27632 (Calbiochem) 중에서 배양하였다. P3 프라이머리 셀 4D-뉴클레오펙터 X 키트 (론자(Lonza))를 이용하여 형질감염을 수행하였다. 구체적으로, TrypLE 익스프레스 (인비트로젠)를 이용하여 세포를 수거하고, 16.4 μ l P3 뉴클레오펙터 용액, 3.6 μ l 보충물, 1 μ g의 각각의 리TALEN 플라스미드 또는 1 μ g Cas9 및 1 μ g gRNA 구축물, 2 μ l의 100 μ M ssODN을 함유하는 20 μ l 뉴클레오펙션 혼합물 중에 2×10^6 개의 세포를 재현탁시켰다. 이어서, 혼합물을 20 μ l 뉴클레오펙스트립으로 옮기고, CB150 프로그램을 이용하여 뉴클레오펙션을 수행하였다. 세포를 처음 24 hr 동안 ROCK 억제제로 보충된 mTeSR1 배지 중 매트릭셀 코팅된 플레이트 상에 플레이팅시켰다. dsDNA 도너를 이용하여 내인성 AAVS1 유전자좌 표적화 실험을 수행하는 경우, 동일한 방법을 수행하되, 단, 예외적으로, 2 μ g dsDNA 도너를 사용하였고, mTeSR1 배지를 형질감염 후 1주 경과하였을 때, 0.5 μ g/mL 농도의 푸로마이신으로 보충하였다.

[0081] 본 실시예에서 사용된 리TALEN, gRNA 및 ssODN에 관한 정보는 하기 표 3 및 표 4에 열거되어 있다.

[0082] <표 3> 리TALEN 쌍/Cas9-gRNA 표적화 CCR5에 관한 정보

표적화 부위 번호	리 TAL EN	리 TAL EN	리TALEN-L 표적화 서열	리TALEN-R 표적화 서열	gRNA 표적화 서열	gRNA 표적화 서열 출발 위치
	쌍 표적화 부위 (출발)	쌍				
	/chr3:	표적화 부위 (단부)				
		/chr3:				
1	4640 9942	4640 9993	TCCCCACTTTCT TGTGAA	TAACCACTCAG GACAGGG	CACTTCTTGTG AATCCTT	4640 9946
2	4641 0227	4641 0278	TCACACAGCAA GTCAGCA	TAGCGGAGCAG GCTCGGA	TGGGCTAGCGG AGCAGGCT	4641 0264
3	4641 1260	4641 1311	TACCCAGACGA GAAAGCT	TCAGACTGCCA AGCTTGA	ACCCAGACGAG AAAGCTGA	4641 1261
4	4641 1464	4641 1515	TCTTGTGGCTC GGGAGTA	TATTGTCAGCA GAGCTGA	AGAGGGCATCTT GTGGCTC	4641 1456
5	4641 1517	4641 1568	TTGAGATTTTC AGATGTC	TATACAGTCAT ATCAAGC	ATCAAGCTCTCT TGGCGGT	4641 1538
6	4641 1634	4641 1685	TTCAGATAGAT TATATCT	TGCCAGATACA TAGGTGG	GCTTCAGATAGA TTATATC	4641 1632
7	4641 2396	4641 2447	TTATACTGTCT ATATGAT	TCAGCTCTTCT GGCCAGA	ACGGATGTCTCA GCTCTTC	4641 2437
8	4641 2432	4641 2483	TGGCCAGAAGA GCTGAGA	TTACCGGGGAG AGTTTCT	CCGGGGAGAGT TTCTTGTA	4641 2461

[0083]

9	4641 2750	4641 2801	TTTGCAGAGAG ATGAGTC	TTAGCAGAAGA TAAGATT	GAAATCTTATCT TCTGCTA	4641 2782
10	4641 3152	4641 3203	TATAAGACTAA ACTACCC	TCGTCTGCCAC CACAGAT	AATGCATGACAT TCATCTG	4641 3172
11	4641 4305	4641 4356	TAAAACAGTTT GCATTCA	TATAAAGTCCT AGAATGT	AACAGTTTGCAT TCATGGA	4641 4308
12	4641 4608	4641 4659	TGGCCATCTCT GACCTGT	TAGTGAGCCCA GAAGGGG	CCAGAAGGGGA CAGTAAGA	4641 4632
13	4641 4768	4641 4820	TAGGTACCTGG CTGTCTG	TGACCGTCCTG GCTTTTA	CTGACAATCGAT AGGTACC	4641 4757
14	4641 5017	4641 5068	TGTCATGGTCA TCTGCTA	TCGACACCGAA GCAGAGT	ACACCGAAGCA GAGTTTTT	4641 5046
15	4642 0034	4642 0084	TGCCCCGCGA GGCCACA	TCTGGAAGTTG AACACCC	GGAAGTTGAAC ACCCTTGC	4642 0062

[0084]

[0085]

<표 4> ssODN-매개 게놈 편집 연구를 위한 ssODN 디자인

도 3b	2차 돌연 변이 와 DSB 사이 의 거리	90- *1	CTACTGTCATTTCAGGGCAATACCCAGACGAGAAAGCTGAGGGTAT AACAGGTTTCAAGCTTGGCAGTCTGACTACAGAGGCCACTGGCTT
		90- *2	CTACTGTCATTTCAGCCCAATACCCTAACGAGAAAGCTGAGGGTATA ACAGGTTTCAAGCTTGGCAGTCTGACTACAGAGGCCACTGGCTT
		90- *3	CTACTGTCATTTCAGCCCAATACCCAGACGAGAAAAGTGAGGGTATA ACAGGTTTCAAGCTTGGCAGTCTGACTACAGAGGCCACTGGCTT
		90M- 0	CTACTGTCATTTCAGCCCAATACCCAGACGAGAAAGCTGAGGGTATA ACAGGTTTCAAGCTTGGCAGTCTGACTACAGAGGCCACTGGCTT
		90- *4	CTACTGTCATTTCAGCCCAATACCAGACGAGAAAGCTGAGGGTATA ACAGGTTTGTAGCTTGGCAGTCTGACTACAGAGGCCACTGGCTT
		90- *5	CTACTGTCATTTCAGCCCAATACCCAGACGAGAAAGCTGAGGGTATA ACAGGTTTCAAGCTTGGCTCTCTGACTACAGAGGCCACTGGCTT

[0086]

		90- *6	CTACTGTCATTTCAGCCCAATACCCAGACGAGAAAGCTGAGGGTATA ACAGGTTTCAAGCTTGGCAGTCTGACTAGTGAGGCCACTGGCTT
도 3c	ss ODN 과 DSB 사이의 거리	L670 bp_9 0M	CACTTTATATTTCCCTGCTTAAACAGTCCCCGAGGGTGGGTGCGG AAAAGGCTCTACACTTGTTATCATTCCCTCTCCACCACAGGCAT
		L570 bp_9 0M	TTTGTATTTGGGTTTTTTTTAAACCTCCACTCTACAGTTAAGAATTC TAAGGCACAGAGCTTCAATAATTTGGTCAGAGCCAAGTAGCAG
		L480 bp_9 0M	GGAGGTAAACCCAGCAGCATGACTGCAGTTCTTAATCAATGCCCC TTGAATTGCACATATGGGATGAACTAGAACATTTTCTCGATGAT
		L394 bp_9 0M	CTCGATGATTCGCTGCCTTGTTATGATTATGTTACTGAGCTCTACT GTAGCACAGACATATGTCCCTATATGGGGCGGGGGTGGGGGTG
		L290 bp_9 0M	GGTGTCTTGATCGCTGGGCTATTTCTATACTGTTCTGGCTTTTCGGA AGCAGTCATTTCTTCTATTCTCCAAGCACCAGCAATTAGCTT
		L200 bp_9 0M	GCTTCTAGTTTGTCTGAAACTAATCTGCTATAGACAGAGACTCCGAC GAACCAATTTTATTAGGATTTGATCAAATAAACTCTCTCTGACA
		L114 bp_9 0M	GAAAGAGTAACTAAGAGTTTGTATGTTACTGAGTGCATAGTATGCA CTAGATGCTGGCCGTGGATGCCTCATAGAATCCTCCCAACAACCT
		L45b p_90 M	GCTAGATGCTGGCCGTGGATGCCTCATAGAATCCTCCCAACAACCG ATGAAATGACTACTGTCATTCAGCCCAATACCCAGACGAGAAAG
		R40b p_90 M	ACAGGTTTCAAGCTTGGCAGTCTGACTACAGAGGCCACTGGCTTTA CCCCTGGGTTAGTCTGCCTCTGTAGGATTGGGGGCACGTAATTT
		R100 bp_9 0M	TTAGTCTGCCTCTGTAGGATTGGGGGCACGTAATTTGCTGTTTAAG GTCTCATTGCTTCTTAGAGATCACAAAGCCAAAGCTTTTTAT
		R200 bp_9 0M	GGAAGCCCAGAGGGCATCTTGTGGCTCGGGAGTAGCTCTCTGCTAC CTTCTCAGCTCTGCTGACAATACTTGAGATTTTCAGATGTCACC

[0087]

	R261 bp_9 0M	TCAGCTCTGCTGACAATACTTGAGATTTTCAGATGTCACCAACCAG CAAGAGAGCTTGATATGACTGTATATAGTATAGTCATAAAGAAC	
	R322 bp_9 0M	CATAAAGAACCTGAACTTGACCATATACTTATGTCATGTGGAAATC TTCTCATAGCTTCAGATAGATTATATCTGGAGTGAAGAATCCTG	
	R375 M_9 0M	GTGGAAAATTTCTCATAGCTTCAGATAGATTATATCTGGAGTGAGC AATCCTGCCACCTATGTATCTGGCATAGTGTGAGTCCTCATAAA	
	R448 bp_9 0M	GGTTTGAAGGGCAACAAAATAGTGAACAGAGTGAAAATCCCCACC TAGATCCTGGGTCCAGAAAAAGATGGGAAACCTGTTTAGCTCACC	
	상보체 - 30me r	GGCCACTAGGGACAAAATTGGTGAcagaaa	
도 3d	Cas9- gRNA 표적화 를 위한 ssODN 길이 및 배향	상보체 - 50me r	CCCACAGTGGGGCCACTAGGGACAAAATTGGTGAcagaaaagcccatcc
		상보체 - 70me r	TCCCCTCCACCCACAGTGGGGCCACTAGGGACAAAATTGGTGAcag aaaagcccatccttaggcctcc
		상보체 - 90me r	ctTTATCTGTCCCCTCCACCCACAGTGGGGCCACTAGGGACAAAAT TGGTGAcagaaaagcccatccttaggcctcctcctctag
		상보체 -	gttctgggtacttTTATCTGTCCCCTCCACCCACAGTGGGGCCACTAGGGA CAAAATTGGTGAcagaaaagcccatccttaggcctcctcctctagtcctgata

[0088]

	110mer	
	비상보체 - 30mer	TTTCTGTCACCAATGGTGTCCCTAGTGGCC
	비상보체 - 50mer	GGATGGGGCTTTTCTGTCACCAATGGTGTCCCTAGTGGCCCCACTGTGGG
	비상보체 - 70mer	GGAGGCCTAAGGATGGGGCTTTTCTGTCACCAATGGTGTCCCTAGTGGCCCCACTGTGGGGTGGAGGGGA
	비상보체 - 90mer	CTAGGAAGGAGGAGGCCTAAGGATGGGGCTTTTCTGTCACCAATGGTGTCCCTAGTGGCCCCACTGTGGGGTGGAGGGGACAGATAAAAAG
	비상보체 - 110mer	TATCAGGAGACTAGGAAGGAGGAGGCCTAAGGATGGGGCTTTTCTGTCACCAATGGTGTCCCTAGTGGCCCCACTGTGGGGTGGAGGGGACAGATAAAAAGTACCCAGAAC
도	CCR5를 표적화하는 Cas9 gRN	TTCTAGTAACCACTCAGGACAGGGGGTTCAGCCCCAAAATTCACAAGAAAAGTGGGGACCCATGGGAAAT

[0089]

2c	Cas9-gRNA를 위한 ssODN 도너	A- CCR 5-1	
		Cas9 - gRN A- CCR 5-2	CAGCAAGTCAGCAGCACAGCGTGTGTGACTCCGAGGGTGCTCCGCT AGCCACATTGCCCTCTGGGGGTG
		Cas9 - gRN A- CCR 5-3	GTCAGACTGCCAAGCTTCAAACCTGTCTTACCCTCTACTTTCTCGTC TGGGTATTGGGCTGAATGACAGT
		Cas9 - gRN A- CCR 5-4	CAGAGCTGAGAAGACAGCAGAGAGCTACTCCGAAGCACAAGATG CCCTCTGGGCTTCCGTGACCTTGGC
		Cas9 - gRN A- CCR 5-5	CTGACAATACTTGAGATTTTCAGATGTCACCAACGACCAAGAGAGC TTGATATGACTGTATATAGTATAG
		Cas9 - gRN A- CCR 5-6	CAGATACATAGGTGGCAGGATTCCTCACTCCAGACTTAATCTATCT GAAGCTATGAGAAATTTCCACAT
		Cas9 -	TATATGATTGATTTGCACAGCTCATCTGGCCAGATAAGCTGAGACA TCCGTTCCCCTACAAGAAACTCTC

[0090]

gRN A- CCR 5-7	
Cas9 - gRN A- CCR 5-8	ATCTGGCCAGAAGAGCTGAGACATCCGTTCCCCTTGAAGAACTCT CCCCGGTAAGTAACCTCTCAGCTG
Cas9 - gRN A- CCR 5-9	AGGCATCTCACTGGAGAGGGTTTAGTTCTCCTTAAGAGAAGATAAG ATTTCAAGAGGGAAGCTAAGACTC
Cas9 - gRN A- CCR 5-10	ATAATATAATAAAAAATGTTTCGTCTGCCACCACTAATGAATGTCA TGCATTCTGGGTAGTTTAGTCTTA
Cas9 - gRN A- CCR 5-11	TTTATAAAGTCCTAGAATGTATTTAGTTGCCCTCGTTGAATGCAAAC TGTTTTATACATCAATAGGTTTT
Cas9 - gRN A- CCR 5-12	GCTCAACCTGGCCATCTCTGACCTGTTTTTCCTCCCCTGTCCCCT TCTGGGCTCACTATGCTGCCGCC
Cas9	TTTTAAAGCAAACACAGCATGGACGACAGCCAGGCTCCTATCGATT

[0091]

	- gRN A- CCR 5-13	GTCAGGAGGATGATGAAGAAGATT
	Cas9 - gRN A- CCR 5-14	GCTTGTGCATGGTCATCTGCTACTCGGGAATCCTAATTACTCTGCTTC GGTGTGCGAAATGAGAAGAAGAGG
	Cas9 - gRN A- CCR 5-15	ATACTGCCCCCGGAGGCCACATTGGCAAACCAGCTTGGGTGTTCA ACTTCCAGACTTGGCCATGGAGAA
CCR5 를 표적화 하는 리 TALEN 을 위한 ssODN 도너	리TA LEN- CCR 5-1	CTGAAGAATTTCCCATGGGTCCCCTTTCTTGTGAATCCTTGGAGT GAACCCCTGTCTGAGTGGTTACTAGAACACACCTCTGGAC
	리TA LEN- CCR 5-2	TGGAAGTATCTTGCCGAGGTCACACAGCAAGTCAGCAGCACAGCC AGTGTGACTCCGAGCCTGCTCCGCTAGCCCACATTGCCCTCTGGG
	리TA LEN- CCR 5-3	CTACTGTCATTCAGCCAATACCCAGACGAGAAAGCTGAGGGTATA ACAGGTTTCAAGCTTGGCAGTCTGACTACAGAGGCCACTGGCTT
	리TA LEN- CCR 5-4	GGAAGCCCAGAGGGCATCTTGTGGCTCGGGAGTAGCTCTCTGCTAC CTTCTCAGCTCTGCTGACAATACTTGAGATTTTCAGATGTCACC
	리TA LEN-	TCAGCTCTGCTGACAATACTTGAGATTTTCAGATGTCACCAACGCC CAAGAGAGCTTGATATGACTGTATATAGTATAGTCATAAAGAAC

[0092]

CCR 5-5	
리TA LEN- CCR 5-6	GTGGAAAATTTCTCATAGCTTCAGATAGATTATATCTGGAGTGAGC AATCCTGCCACCTATGTATCTGGCAGTAGTGTGAGTCCTCATAAA
리TA LEN- CCR 5-7	GAAACAGCATTTCCTACTTTTATACTGTCTATATGATTGATTTGGTC AGCTCATCTGGCCAGAAGAGCTGAGACATCCGTTCCCCTACAA
리TA LEN- CCR 5-8	TTGATTTGCACAGCTCATCTGGCCAGAAGAGCTGAGACATCCGTAT CCCTACAAGAACTCTCCCCGGTAAGTAACCTCTCAGCTGCTTG
리TA LEN- CCR 5-9	GGAGAGGGTTTAGTTCTCCTTAGCAGAAGATAAGATTCAAGATGA GAGCTAAGACTCATCTCTCTGCAAATCTTCTTTTGAGAGGTAA
리TA LEN- CCR 5-10	TAATATAATAAAAAATGTTTCGTCTGCCACCACAGATGAATGTCGA GCATTCTGGGTAGTTTAGTCTTATAACCAGCTGTCTTGCCTAGT
리TA LEN- CCR 5-11	TTAAAAACCTATTGATGTATAAAACAGTTTGCATTCATGGAGGGTG ACTAAATACATTCTAGGACTTTATAAAAGATCACTTTTTATTTA
리TA LEN- CCR 5-12	GACATCTACCTGCTCAACCTGGCCATCTCTGACCTGTTTTTCCTATT TACTGTCCCCTTCTGGGCTCACTATGCTGCCGCCAGTGGGAC
리TA LEN- CCR 5-13	TCATCCTCCTGACAATCGATAGGTACCTGGCTGTCGTCATGCTAC GTTTGCTTTAAAAGCCAGGACGGTCACCTTTGGGGTGGTGACAA

[0093]

리TA LEN- CCR 5-14	GGCTGGTCCTGCCGCTGCTTGTCATGGTCATCTGCTACTCGGGAGA CCTAAAAACTCTGCTTCGGTGTGCGAAATGAGAAGAAGAGGCACA
리TA LEN- CCR 5-15	GGCAAGCCTTGGGTCACTGCCCCGCGAGGCCACATTGGCAAGT CAGCAAGGGTGTTCAACTTCCAGACTTGCCATGGAGAAGACAT

[0094]

[0095]

실시예 VII

[0096]

표적화 영역의 앰플리콘 라이브러리 제조

[0097] 뉴클레오펙션 후 6일 경과하였을 때, 세포를 수거하고, 0.1 μ l 프렙GEM(preGEM) 조직 프로테아제 효소 (자이GEM(ZyGEM)) 및 1 μ l 프렙GEM 골드 완충제 (자이GEM)를 8.9 μ l의 배지 중 2-5 X 10⁵개의 세포에 첨가하였다. 이어서, 1 μ l의 반응물을, 5 μ l 2X 카파 히피 핫스타트 레디믹스(KAPA Hifi Hotstart Readymix) (카파 바이오시스템즈(KAPA Biosystems)) 및 100 nM 상응하는 증폭 프라이머 쌍을 함유하는 9 μ l PCR 믹스에 첨가하였다. 반응물을 95°C에서 5 min 동안, 이어서, 98°C, 20 s; 65°C, 20 s 및 72°C, 20 s의 15 사이클로 인큐베이션시켰다. 이어서, 사용되는 일루미나(Illumina) 서열 어댑터를 부가하기 위해, 5 μ l 반응 생성물을, 12.5 μ l 2X 카파 히피 핫스타트 레디믹스 (카파 바이오시스템즈) 및 200 nM의 일루미나 서열 어댑터를 포함하는 프라이머를 함유하는 20 μ l의 PCR 믹스에 첨가하였다. 반응물을 95°C에서 5 min 동안, 이어서, 98°C, 20 s; 65°C, 20 s 및 72°C, 20 s의 25 사이클로 인큐베이션시켰다. QIA퀵 PCR 정제용 키트에 의해 PCR 생성물을 정제하고, 거의 동일한 농도로 혼합하고, MiSeq 퍼스널 시퀀서(MiSeq Personal Sequencer)를 이용하여 서열분석하였다. PCR 프라이머는 하기 표 5에 열거되어 있다.

[0098] <표 5> CCR5 표적화 부위 PCR 프라이머 서열

CCR5 중 표적화 번호	명칭	프라이머 서열
1	부위1-F1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGTGATTTT GCAGTGTGCGTTACTCC
	부위1-F2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTACATCGTTT GCAGTGTGCGTTACTCC
	부위1-F3	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGCCTAATTT GCAGTGTGCGTTACTCC
	부위1-F4	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTGGTCATTT GCAGTGTGCGTTACTCC
	부위1-R	CTCGGCATTCTGCTGAACCGCTCTTCCGATCTCCAAGCAA

[0099]

		CTAAGTCACAGCA
2	부위2-F1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGTGATATG AGGAAATGGAAGCTTG
	부위2-F2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTACATCGAT GAGGAAATGGAAGCTTG
	부위2-F3	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGCCTAAAT GAGGAAATGGAAGCTTG
	부위2-F4	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTGGTCAATG AGGAAATGGAAGCTTG
	부위2-R	CTCGGCATTCTGTGAACCGCTCTTCCGATCTCATTAGGG TATTGGAGGA
3	부위3-F1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGTGATAAT CCTCCCAACAACATCAT
	부위3-F2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTACATCGAA TCCTCCCAACAACATCAT
	부위3-F3	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGCCTAAAA TCCTCCCAACAACATCAT
	부위3-F4	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTGGTCAAAT CCTCCCAACAACATCAT
	부위3_R	CTCGGCATTCTGTGAACCGCTCTTCCGATCTCCAATCCT ACAGAGGCAG
4	부위4-F1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGTGATAA GCCAAAGCTTTTTATTCT
	부위4-F2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTACATCGAA GCCAAAGCTTTTTATTCT
	부위4-F3	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGCCTAAAA GCCAAAGCTTTTTATTCT
	부위4-F4	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTGGTCAAA GCCAAAGCTTTTTATTCT
	부위4_R	ACACTCTTTCCCTACACGACGCTCTTCCGATCTAAGCCAAA GCTTTTTATTCT
5	부위5-F1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGTGATATC TTGTGGCTCGGGAGTAG
	부위5-F2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTACATCGATC TTGTGGCTCGGGAGTAG

[0100]

	부위5-R	CTCGGCATTCTGCTGAACCGCTCTTCCGATCTTGGCAGGA TTCTTCACTCCA
6	부위6-F1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGTGATCTA TTTTGTTGCCCTTCAA
	부위6-F2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTACATCGCTA TTTTGTTGCCCTTCAA
	부위6-R	CTCGGCATTCTGCTGAACCGCTCTTCCGATCTAACCTGAA CTTGACCATATACT
7	부위7-F1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGTGATCA GCTGAGAGGTTACTTACC
	부위7-F2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTACATCGCA GCTGAGAGGTTACTTACC
	부위7-R	CTCGGCATTCTGCTGAACCGCTCTTCCGATCTAATGATTTA ACTCCACCCTC
8	부위8-F1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGTGATACT CCACCCTCTTCAAAGA
	부위8-F2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTACATCGACT CCACCCTCTTCAAAGA
	부위8-R	CTCGGCATTCTGCTGAACCGCTCTTCCGATCTTGGTGTG CCAAATGTCT
9	부위9_F1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGTGATGG GCACATATTCAGAAGGCA
	부위9_F2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTACATCGGG GCACATATTCAGAAGGCA
	부위9_R	CTCGGCATTCTGCTGAACCGCTCTTCCGATCTAGTGAAAG ACTTTAAAGGGAGCA
10	부위10-F1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGTGATCAC AATTAAGAGTTGTCATA
	부위10-F2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTACATCGCA CAATTAAGAGTTGTCATA
	부위10-R	CTCGGCATTCTGCTGAACCGCTCTTCCGATCTCTCAGCTA GAGCAGCTGAAC
11	부위11-F1	CTCGGCATTCTGCTGAACCGCTCTTCCGATCTGACACTTG ATAATCCATC
	부위11-F2	ACACTCTTTCCCTACACGACGCTCTTCCGATCTACATCGTCA

[0101]

		ATGTAGACATCTATGTAG
	부위11-R	ACACTCTTTCCTACACGACGCTCTTCCGATCTCGTGATTCA ATGTAGACATCTATGTAG
12	부위12-F1	ACACTCTTTCCTACACGACGCTCTTCCGATCTCGTGATACT GCAAAAGGCTGAAGAGC
	부위12-F2	ACACTCTTTCCTACACGACGCTCTTCCGATCTACATCGACT GCAAAAGGCTGAAGAGC
	부위12-F3	ACACTCTTTCCTACACGACGCTCTTCCGATCTGCCTAAACT GCAAAAGGCTGAAGAGC
	부위12-F4	ACACTCTTTCCTACACGACGCTCTTCCGATCTTGGTCAACT GCAAAAGGCTGAAGAGC
	부위12-R	CTCGGCATTCTGCTGAACCGCTCTTCCGATCTGCCTATAA AATAGAGCCCTGTCAA
13	부위13-F1	ACACTCTTTCCTACACGACGCTCTTCCGATCTCGTGATCTC TATTTTATAGGCTTCTTC
	부위13-F2	ACACTCTTTCCTACACGACGCTCTTCCGATCTACATCGCTC TATTTTATAGGCTTCTTC
	부위13-R	CTCGGCATTCTGCTGAACCGCTCTTCCGATCTAGCCACCA CCCAAGTGATC
14	부위14-F1	ACACTCTTTCCTACACGACGCTCTTCCGATCTACATCGTTC CAGACATTAAGATAGTC
	부위14-F2	ACACTCTTTCCTACACGACGCTCTTCCGATCTCGTGATTTC CAGACATTAAGATAGTC
	부위14-R	CTCGGCATTCTGCTGAACCGCTCTTCCGATCTAATCATGA TGGTGAAGATAAG
15	부위15-F1	ACACTCTTTCCTACACGACGCTCTTCCGATCTCGTGATCCG GCAGAGACAAACATTAAG
	부위15-F2	ACACTCTTTCCTACACGACGCTCTTCCGATCTCCGGCAGA GACAAACATTAAG
	부위15-R	CTCGGCATTCTGCTGAACCGCTCTTCCGATCTAGCTAGGA AGCCATGGCAAG
일루미나 어댑터	PE-PCR-F	AATGATACGGCGACCACCGAGATCTACACTCTTTCCTACA cgac*g*c
	PE-PCR-R	CAAGCAGAAGACGGCATAACGAGATCGGTCTCGGCATTCT

[0102]

		GCTGAACc*g*c
다중 서열분석 PCR 프라이머		
3	부위3-M-F	ACACTCTTTCCTACACGACGCTCTTCCGATCTAGTGCATA GTATGTGCTAGATGCTG
	부위3-M-R	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTGATCTC TAAGAAGGCAAATGAGAC
일루미나 어댑터	인덱스-PCR	CAAGCAGAAGACGGCATAACGAGATN1N2N3N4N5N6GTGAC TGGAGTTCAGACGTGTGCTCTTCCGATCT
	범용-PCR	AATGATACGGCGACCACCGAGATCTACACTCTTTCCTACA CGACGCTCTTCCGATCT

[0103]

- [0104] *인덱스-PCR 프라이머는 에피상트르 (스크립트Seq™ 인덱스 PCR 프라이머(ScriptSeq™ Index PCR Primers))로부터 구입
- [0105] 실시예 VIII
- [0106] 게놈 편집 평가 시스템 (GEAS)
- [0107] 회귀 게놈 변경을 검출하기 위해 차세대 서열분석을 사용하여 왔다. 참고 문헌 27-30 (상기 문헌들은 그 전문이 본원에서 참조로 포함된다)을 참조할 수 있다. 상기 접근법을 광범위하게 사용함으로써 hiPSCs 중 HDR 및 NHEJ 효율을 신속하게 평가할 수 있도록 하기 위해, "파이프라인"으로 지칭되는 소프트웨어를 작성하여 게놈 조각 데이터를 분석하였다. 상기 파이프라인은 상이한 도구, 예컨대, R, BLAT, 및 FASTX 툴킷(FASTX Toolkit)를 사용하는 하나의 단일 유닉스(Unix) 모듈로 통합된다.
- [0108] 바코드 분할: 샘플 균을 함께 풀링하고, MiSeq 150 bp 쌍 형성 단부 (PE150) (일루미나 넥스트 젠 시퀀싱 (Illumina Next Gen Sequencing))를 이용하여 서열분석한 후, 이어서, FASTX 툴킷을 이용하여 DNA 바코드에 기초하여 분류하였다.
- [0109] 품질 여과: 서열 품질이 더 낮은 (프레드 점수<20) 뉴클레오티드를 트리밍하였다. 트리밍한 후, 길이가 80개의 뉴클레오티드보다 더 짧은 리드는 폐기하였다.
- [0110] 지도화: BLAT를 이용하여 쌍 형성 리드를 독립적으로 참조 게놈으로 지도화하고, 출력값으로서 .psl 파일을 작성하였다.
- [0111] Indel 호출: Indel은 정렬 중 2 블록의 매치를 함유하는 전장의 리드로서 정의되었다. 2개의 쌍 형성 단부 리드 모두에서 하기 패턴 다음의 리드만이 고려되었다. 품질 관리자로서 indel 리드는 참조 게놈과 매칭되는 최소 70 nt의 블록을 가져야 하고, 두 블록은 모두 길이가 20 nt 이상이어야 한다. 참조 게놈에 대해 상대적인 각 블록의 위치에 의해 indel의 크기 및 위치를 계산하였다. indel을 함유하는 리드의 비율(%) (하기 방정식 1 참조)로서 비-상동성 단부 연결 (NHEJ)을 추정하였다. NHEJ 이벤트의 대다수는 표적화 부위 인접부에서 검출되었다.
- [0112] 상동성 지정 재조합 (HDR) 효율: DSB 상의 중심부의 12 bp 윈도우 내의 패턴 매칭 (grep)을 이용하여 참조 서열을 함유하는 리드, 참조 서열의 변형 (2 bp 의도된 미스매치)을 함유하는 리드, 및 2 bp 의도된 미스매치 내에 오직 1 bp 돌연변이만을 함유하는 리드에 상응하는 특이 시그니처를 계수하였다 (하기 방정식 1 참조).
- [0113] 방정식 1. NHEJ 및 HDR 추정
- [0114] A= 참조와 동일한 리드: XXXXXABXXXXX
- [0115] B= ssODN에 의해 프로그램화된 2 bp 미스매치를 함유하는 리드: XXXXXabXXXXX
- [0116] C= 표적 부위 중 오직 1 bp 돌연변이만을 함유하는 리드: 예컨대, XXXXXaBXXXXX 또는 XXXXXAbXXXXX
- [0117] D = 상기 기술된 바와 같은 indel을 함유하는 리드
- $$\text{NHEJ 효율} = \left(100 \times \frac{D}{A+B+C+D}\right)\%$$
- $$\text{HDR 효율} = \left(100 \times \frac{B}{A+B+C+D}\right)\%$$
- [0118]
- [0119] 실시예 IX
- [0120] 전이증식된 hiPSC의 유전자형 스크리닝
- [0121] FACS 분류 전 적어도 2시간 동안 SMC4 (5 μ M 티아조비빈, 1 μ M CHIR99021, 0.4 μ M PD0325901, 2 μ M SB431542)로 보충된 mTesr-1 배지로 피더 세포 무함유 배양물 상의 인간 iPS 세포를 미리 처리하였다 (참고 문헌 23 (상기 문헌은 그 전문이 본원에서 참조로 포함된다) 참조). 아큐타제(Accutase) (밀리포어(Millipore))를 이용하여 배양물을 해리시키고, SMC4 및 1~2 X 10⁷ /mL 농도의 생존력 염료 ToPro-3 (인비트로젠)으로 보충된 mTesr-1 배지 중에 재현탁시켰다. 멸균 조건하에서 100 μ m 노즐이 장착된 BD FACSAria II SORP UV (BD 바이오사이언시즈)를 이용하여 생 hiPS 세포를 방사선 조사된 CF-1 마우스 배아 섬유아세포 (글로벌 스템(Global Stem))으로 코팅된 96 웰 플레이트로 단일 세포 분류하였다. 각 웰은 SMC4 및 5 μ g/ml 피브로넥틴 (시그마(Sigma))으로 보

충된, 100 ng/ml 재조합 인간 염기성 섬유아세포 성장 인자 (bFGF) (밀리포어)를 포함하는 hES 세포 배지를 함유하였다 (참고 문헌 31 (상기 문헌은 그 전문이 본원에서 참조로 포함된다) 참조). 분류 후, 플레이트를 3 min 동안 70 x g로 원심분리하였다. 분류 후 4일째 콜로니 형성을 관찰할 수 있었으며, 배양 배지를 SMC4가 보충된 hES 세포 배지로 교체하였다. 분류 후 8일 경과하였을 때, SMC4를 hES 세포 배지로부터 제거할 수 있다.

[0122] 형광 활성화된 세포 분류 (FACS) 후 8일 경과하였을 때, 수천 개의 세포를 수거하고, 0.1 ul 프렙GEM 조직 프로테아제 효소 (자이GEM) 및 1 ul 프렙GEM 골드 완충제 (자이GEM)를 8.9 μ l의 배지 중 세포에 첨가하였다. 이어서, 반응물을, 35.5 ml 백금 1.1X 슈퍼믹스(Supermix) (인비트로젠), 250 nM의 각 dNTP 및 400 nM 프라이머를 함유하는 40 μ l PCR 믹스에 첨가하였다. 반응물을 95°C에서 3 min 동안, 이어서, 95°C, 20 s; 65°C, 30 s 및 72°C, 20 s의 30 사이클로 인큐베이션시켰다. 생성물을 표 5의 PCR 프라이머 중 하나를 이용하여 생어 서열분석하고, DNASTAR (DNASTAR)을 이용하여 서열을 분석하였다.

[0123] 실시예 X

[0124] hiPSC의 면역염색 및 테라토마 검정

[0125] 세포를 하기 항체: 항-SSEA-4 PE (밀리포어) (1:500으로 희석); Tra-1-60 (BD 파미젠(BD Pharmingen)) (1:100으로 희석)를 이용하여 37°C에서 60분 동안 닉아웃 DMEM/F-12 배지 중에서 인큐베이션시켰다. 인큐베이션 후, 세포를 닉아웃 DMEM/F-12로 3회에 걸쳐 세척하고, 악시오 옵저버 Z.1 (자이스) 상에서 영상화하였다.

[0126] 테라토마 형성 분석을 수행하기 위해, 콜라게나제 IV형 (인비트로젠)을 이용하여 인간 iPSC를 수거하고, 세포를 200 μ l 매트릭셀에 재현탁시키고, Rag2감마 녹아웃 마우스의 뒷다리 내로 근육내로 주사하였다. 주사 후 4-8주 경과하였을 때, 테라토마를 단리시키고, 포르말린 중에 고정시켰다. 이어서, 테라토마를 헤마톡실린 및 에오신 염색에 의해 분석하였다.

[0127] 실시예 XI

[0128] 리TALEN을 이용한 인간 체세포 및 인간 줄기 세포 중 게놈 유전자좌 표적화

[0129] 특정 측면에 따라, 통상의 기술자에게 공지된 TALE를 변형 또는 재코딩하여 반복부 서열을 제거한다. 본원에 기술된 바이러스 전달 비히클에서 및 각종 세포주 및 유기체에서의 게놈 편집 방법에서 변형 및 사용에 적합한 상기 TALE는 참고 문헌 2, 7-12 (상기 문헌들은 그 전문이 본원에서 참조로 포함된다)에 개시되어 있다. 반복적 TALE RVD 어레이 서열을 조립하는 수개의 전략법이 개발되었다 (참고 문헌 14 및 32-34 (상기 문헌들은 그 전문이 본원에서 참조로 포함된다)를 참조). 그러나, 일단 조립되고 나면, TALE 서열 반복부는 불안정한 상태 그대로 유지되며, 이로 인해 상기 도구의 광범위한 유용성은 제한되고, 특히 바이러스 유전자 전달 비히클인 경우에 그러하다 (참고 문헌 13 및 35 (상기 문헌들은 그 전문이 본원에서 참조로 포함된다)를 참조). 따라서, 본 개시내용의 한 측면은 반복부가 결핍된, 예컨대, 반복부가 완전히 TALEN에 관한 것이다. 상기 재코딩된 TALE는 그를 통해서 확장된 TALE RVD 어레이를 더욱 신속하고, 간단하게 합성할 수 있기 때문에 이롭다.

[0130] 반복부를 제거하기 위하여, TALE RVD 어레이의 뉴클레오티드 서열을 연산적으로 진화시켜 아미노산 조성은 유지시키면서, 서열 반복부의 개수를 최소화시켰다. 16개의 일렬 RVD DNA 인식 단량체와, 마지막 절반 RVD 반복부를 코딩하는 재코딩된 TALE (리TALEN)에는 임의의 12 bp 반복부가 없다 (도 5a 참조). 특히, 이러한 재코딩 수준은 전장의 리TALEN 구축물로부터 임의의 특이 단량체 또는 서브섹션을 PCR 증폭시키는 데 충분하다 (도 5b 참조). 반복부 중쇄 서열과 관련된 추가 비용 또는 방법을 초래하지 않으면서, 표준 DNA 합성 기술을 사용하여 개선된 디자인의 리TALEN을 합성할 수 있다 (참고 문헌 36 (상기 문헌은 그 전문이 본원에서 참조로 포함된다) 참조). 추가로, 재코딩된 서열 디자인을 통해서 본원에 기술된 방법에 기술된 바와 같이, 및 도 6을 참조하여 변형된 등은 조립 반응을 이용함으로써 리TALEN 구축물을 효율적으로 조립할 수 있다.

[0131] 게놈 편집 NGS 데이터를 하기와 같이 통계학적으로 분석하였다. HDR 특이성 분석을 위해, 정확한 이항 시험을 사용하여 2 bp 미스매치를 함유하는 다양한 개수의 서열 리드를 관찰할 수 있는 확률을 계산하였다. 부위 표적화 전후의 10 bp 윈도우에 관한 서열분석 결과에 기초하여, 두 윈도우의 최대 염기 변화율 (P1 및 P2)을 추정하였다. 두 표적 bp 각각의 변화는 독립적이라는 널 가설을 사용하여, 우연히 표적화 부위에서 2 bp 미스매치를 관찰할 수 있는 예상 확률은 상기 두 확률의 곱 (P1*P2)으로서 계산하였다. 총 리드 개수 N 및 HDR 리드 개수 n을 함유하는 데이터셋 전체하여, 본 발명자들은 관찰된 HDR 효율의 p 값을 계산하였다. HDR 감도 분석을 위해, ssODN DNA 도너는 표적화 게놈에 대하여 2 bp 미스매치를 함유하였으며, 그 결과, ssODN이 표적화 게놈 내로 도입될 경우, 가능하게는 2개의 표적 bp 중 염기 변이가 공동으로 존재하게 되었다. 다른 의도되지 않은 관찰된 서열 변이는 가능하게는 동시에 바뀌지 않을 것이다. 따라서, 의도되지 않은 변이는 상호의존성이 훨씬

더 적었다. 이러한 가정하에, 상호 정보 (MI)를 이용하여 모든 다른 위치 쌍에서의 동시의 두 염기쌍 변이의 상호 의존성을 측정하고, 2 bp 부위를 표적화하는 것의 MI가 모든 다른 위치 쌍의 MI보다 높은 최소 HDR로서 HDR 검출 한계를 추정하였다. 주어진 실험의 경우, 원 fastq 파일로부터의 의도된 2 bp 미스매치를 포함하는 HDR 리드를 확인하였고, 원래의 데이터 세트로부터 상이한 개수의 HDR 리드를 체계적으로 제거함으로써 HDR 효율이 희석된 fastq 파일 세트를 모의하였다. 표적화 부위를 중심으로 한 20 bp의 윈도우 내의 모든 위치 쌍 사이의 상호 정보 (MI)를 계산하였다. 상기 계산에서는 임의의 두 위치 사이의 염기 조성에 관한 상호 정보를 계산한다. 상기 기술된 HDR 특이성 척도와 달리, 상기 척도는 임의의 특정 표적 염기 쌍으로 변이될 수 있는 위치 쌍의 경향을 평가하는 것이 아니라, 오직 동시에 변이될 수 있는 그의 경향만을 평가할 수 있다 (도 8a). 하기 표 6은 CCR5를 표적화하는 리TALEN/ssODN의 HDR 및 NHEJ 효율 및 Cas9-gRNA의 NHEJ 효율을 보여주는 것이다. 본 발명자들은 패키지 이포테오(infotheo)를 이용하여 R 및 MI에 관한 본 발명자들의 분석을 계산하였다.

[0132] <표 6>

표적화 부위 번호	세포 유형	HDR (리TALEN) (%)	NHEJ (리TALEN) (%)	정보 분석에 기반한 HDR 검출 한계	NHEJ (Cas9-gRNA)	HDR (Cas9-gRNA)
1	PGP1-iPS	0.06%	0.80%	0.04%	0.58%	0.38%
2	PGP1-iPS	0.48%	0.26%	0.01%	16.02%	3.71%
3	PGP1-iPS	1.71%	0.07%	0.03%	3.44%	3.20%
4	PGP1-iPS	0.02%	1.20%	0.02%*	1.50%	0.14%
5	PGP1-iPS	0.80%	0.04%	0.00%	3.70%	0.39%
6	PGP1-iPS	0.20%	0.73%	0.00%	1.12%	0.49%
7	PGP1-iPS	0.01%	0.15%	0.01%*	1.98%	1.78%
8	PGP1-iPS	0.03%	0.00%	0.00%	1.85%	0.03%
9	PGP1-iPS	1.60%	0.06%	0.00%	0.50%	0.13%
10	PGP1-iPS	0.68%	1.25%	0.01%	8.77%	1.32%
11	PGP1-iPS	0.06%	0.27%	0.00%	0.62%	0.44%
12	PGP1-iPS	1.60%	0.03%	0.04%	0.18%	0.99%
13	PGP1-iPS	0.00%	1.47%	0.00%	0.65%	0.02%
14	PGP1-iPS	0.47%	0.13%	0.02%	2.50%	0.31%
15	PGP1-iPS	0.8	0.14	0.08%	1.50	1.10%

[0133]

[0134] * HDR 검출 한계가 검출된 실제 HDR을 초과하는 군

[0135]

편집 효율과 후성적 상태 사이의 상관관계는 하기와 같이 처리하였다. 피어슨 상관관계 계수를 계산하여 후성적 파라미터 (DNase I HS 또는 뉴클레오솜 점유) 및 계놈 조작 효율 (HDR, NHEJ) 사이의 가능한 연관성을 연구하였다. UCSC 계놈 브라우저 (hiPSC DNase I HS: /gdb/hg19/bbi/wgEncodeOpenChromDnaseIpsnihi7Sig.bigWig)로부터 DNAaseI 하이퍼센시티비티(DNAaseI Hypersensitivity) 데이터세트를 다운로드 받았다.

[0136]

P 값을 계산하기 위해, 후성적 파라미터의 위치를 임의 추출하여 (N=100,000) 구축된 모의 분포와 관찰된 상관관계를 비교하였다. 관찰된 상관관계가 모의 분포의 95번째 백분위수보다 더 높거나, 또는 5번째 백분위수보다 더 낮은 것은 잠재적으로 연관성이 있는 것으로 간주되었다.

[0137]

인간 세포에서 상용하는 재코딩되지 않는 TALEN과 비교하여 리TALEN의 기능을 측정하였다. 참고 문헌 37 (상기 문헌은 그 전문이 본원에서 참조로 포함된다)에 기술된 바와 같이 프레임 쉬프팅 삽입을 보유하는 GFP 리포터 카세트를 함유하는 HEK 293 세포주를 사용하였다. 도 1a 또한 참조할 수 있다. 프로모터가 없는 GFP 도너 구축물과 함께 삽입 서열을 표적화하는 TALEN 또는 리TALEN을 전달하면, GFP 카세트의 DSB 유도성 HDR 수복이 이루어지고, 이로써, GFP 수복 효율은 뉴클레아제 절단 효율을 평가하는 데 사용될 수 있다. 참고 문헌 38 (상기 문헌은 그 전문이 본원에서 참조로 포함된다)을 참조할 수 있다. 리TALEN은 TALEN에 의해 달성된 것과 유사한 정도로 (1.2%), 형질감염된 세포의 1.4%로 GFP 수복을 유도하였다 (도 1b 참조). PGP1 hiPSC 중 AAVS1 유전자 좌에서의 리TALEN의 활성을 시험하고 (도 1c 참조), 특히 삽입을 함유하는 세포 클론을 성공적으로 회수함으로써

써 (도 1d, e 참조), 리TALEN이 체세포 및 다능성 인간 세포, 둘 모두에서 활성임을 확인할 수 있었다.

[0138] 반복부를 제거함으로써 리TALE 카고를 포함하는 기능성 렌티바이러스를 생성할 수 있었다. 구체적으로, 리TALE-2A-GFP를 코딩하는 렌티바이러스 입자를 패키징하고, 렌티-리TALE-2A-GFP 감염된 293T 세포 폴로 mCherry 리포터를 형질감염시켜 바이러스 입자에 의해 코딩된 리TALE-TF의 활성을 시험하였다. 렌티-리TALE-TF가 형질 도입된 293T 세포는 오직 리포터만이 있는 음성의 것과 비교하였을 때 36X 리포터 발현 활성화를 보였다 (도 7a,b,c 참조). 렌티바이러스 감염된 세포에서의 리TALE-TF의 서열 완전성을 체크하였고, 시험된 10개의 클론 모두에서 전장의 리TALE가 검출되었다 (도 7d 참조).

[0139] 실시예 XII

[0140] 게놈 편집 평가 시스템 (GEAS)을 이용한 hiPSC에서의 리TALE 및 Cas9-gRNA 효율 비교

[0141] hiPSC에서의 리TALEN 대 Cas9-gRNA의 편집 효율을 비교하기 위하여, NHEJ 및 HDR 유전자 편집 이벤트, 둘 모두를 확인하고, 정량화하는 차세대 서열분석 플랫폼 (게놈 편집 평가 시스템)을 개발하였다. 2 bp 미스매치를 제외하면 표적 부위와 동일한 90nt ssODN 도너와 함께, CCR5의 상류 영역 (표 3에서 리TALEN, Cas9-gRNA 쌍 #3)을 표적화하는 리TALEN 쌍 및 Cas9-gRNA를 디자인하고, 구축하였다 (도 2a 참조). 뉴클레아제 구축물 및 도너 ssODN을 hiPSC 내로 형질감염시켰다. 유전자 편집 효율을 정량화하기 위해, 형질감염 후 3일 경과하였을 때, 표적 게놈 영역에 대한 쌍 형성 단부에 관한 심층 서열분석을 수행하였다. 정확한 2 bp 미스매치를 함유하는 리드의 비율(%)에 의해 HDR 효율을 측정하였다. NHEJ 효율은 indel을 보유하는 리드의 비율(%)에 의해 측정하였다.

[0142] hiPSC 내로 ssODN만을 전달한 결과, 최소 비율의 HDR 및 NHEJ를 얻은 반면, 리TALEN 및 ssODN을 전달하였을 때 HDR 효율은 1.7%이고, NHEJ 효율은 1.2%였다 (도 2b 참조). ssODN과 함께 Cas9-gRNA를 도입하였을 때 HDR 효율은 1.2%이고, NHEJ 효율은 3.4%였다. 특히, 게놈 결실 및 삽입의 비율은 두 리TALEN 결합 부위 사이의 스페이스 영역 중간에서 가장 컸지만, Cas9-gRNA 표적화 부위의 프로토스페이스 결합 모티프 (PAM) 서열 상류의 3-4 bp에서 가장 컸는데 (도 2b 참조), 이는 이중 가닥 절단이 상기 영역에서 발생하는 바, 그에 따라 예상되는 바와 같은 것이었다. 리TALEN에 의해 생성된 게놈 결실 크기의 중간값은 6 bp이고, 삽입 크기의 중간값은 3 bp인 것으로 관찰되었고, Cas9-gRNA에 의한 게놈 결실 크기의 중간값은 7 bp이고, 삽입 크기의 중간값은 1 bp인 것으로 관찰되었으며 (도 2b 참조), 이는 일반적으로 NHEJ에 의해 생성된 DNA 손상 패턴과 일치하는 것이었다 (참고 문헌 4 (상기 문헌은 그 전문이 본원에서 참조로 포함된다) 참조). 차세대 서열분석 플랫폼의 수회의 분석을 통해 GEAS는 0.007%만큼 낮은 HDR 검출율을 검출할 수 있는 것으로 나타났는데, 이는 고도로 재현가능할 뿐만 아니라 (복제물 사이의 변동 계수 = $\pm 15\%$ * 측정된 효율), 가장 일반적으로 사용되는 미스매치 감도성 엔도뉴클레아제 검정법보다 400X 더 큰 고감도성이다 (도 8 참조).

[0143] 편집 효율을 측정하기 위하여 CCR5 게놈 유전자좌에서 15개의 부위로 표적화되는 리TALEN 쌍 및 Cas9-gRNA를 구축하였다 (도 2c 참조, 표 3 참조). 광범위한 DNaseI 감도를 나타내기 위하여 상기 부위를 선택하였다 (참고 문헌 39 (상기 문헌은 그 전문이 본원에서 참조로 포함된다) 참조). 상응하는 ssODN 도너를 이용하여 (표 3 참조) 뉴클레아제 구축물을 PGP1 hiPSC 내로 형질감염시켰다. 형질감염 후 6일 경과하였을 때, 상기 부위에서의 게놈 편집 효율을 프로파일링하였다 (하기 표 6). ssODN 도너와 함께 15개의 리TALEN 쌍 중 13개의 경우, NHEJ 및 HDR은 통계학적 검출 역치보다 큰 수준으로 검출되었고, 평균 NHEJ 효율은 0.4%이고, 평균 HDR 효율은 0.6%였다 (도 2c 참조). 추가로, 같은 표적화 유전자좌에서의 HR과 NHEJ 효율 사이에 통계학적 유의적인 양의 상관관계 ($r^2 = 0.81$)가 있는 것으로 밝혀졌고 ($P < 1 \times 10^{-4}$) (도 9a 참조), 이는 HDR과 NHEJ, 둘 모두의 공통된 상류 단계인 DSB 생성이 리TALEN-매개 게놈 편집을 위한 속도 제한 단계임을 제안하는 것이다.

[0144] 대조적으로, 15개의 Cas9-gRNA 쌍은 모두 평균 NHEJ 효율 3% 및 평균 HDR 효율 1.0%로, 유의적인 수준의 NHEJ 및 HR을 보였다 (도 2c 참조). 추가로, Cas9-gRNA에 의해 도입된 NHEJ 및 HDR 효율 사이에 양의 상관관계 또한 검출되었으며 (도 9b 참조) ($r^2 = 0.52$, $p = 0.003$), 이는 리TALEN에 대한 관찰결과와 일치하는 것이었다. Cas9-gRNA에 의해 달성된 NHEJ 효율은 리TALEN에 의해 달성된 것보다 유의적으로 더 높았다 (t 검정, 쌍 형성 단부, $P = 0.02$). gRNA 표적화 서열의 용융 온도와 NHEJ 효율 사이의, 중간 정도이지만, 통계학적 유의적인 상관관계가 관찰되었는데 (도 9c 참조) ($r^2 = 0.28$, $p = 0.04$), 이는 gRNA 및 그의 게놈 표적 사이의 염기쌍 형성의 강도가 Cas9-gRNA-매개 DSB 생성 효율에서 편차가 28% 정도라는 것을 설명할 수 있다는 것을 제안한다. 비록 Cas9-gRNA가 상응하는 리TALEN보다 평균 7배 더 높은 NHEJ 수준을 생성하기는 하지만, Cas9-gRNA는 단지 상응하는 리TALEN의 것 (평균 = 0.6%)과 유사한 HDR 수준을 달성하였다 (평균=1.0%). 과학적 이론으로 한정하고자 하지 않으면서, 본 결과는 DSB에서의 ssODN 농도가 HDR에 대한 제한 인자이거나, 또는 Cas9-gRNA에 의해 생성된 게놈

절단 구조가 효과적인 HDR에 대하여 바람직하지 않다는 것을 제안하는 것일 수 있다. 어느 방법의 경우에도 DNaseI HS와 게놈 표적화 효율 사이에는 어떤 상관관계도 관찰되지 않았다 (도 10 참조).

[0145] 실시예 XIII

[0146] HDR을 위한 ssODN 도너 디자인의 최적화

[0147] hiPSC에서의 고성능 ssODN을 하기와 같이 디자인하였다. 모두가 CCR5 리TALEN 쌍 #3 표적 부위의 스페이서 영역의 한가운데 동일한 2 bp 미스매치를 보유하는 것인, 상이한 길이의 (50-170 nt) ssODN 도너 세트를 디자인하였다. HDR 효율은 ssODN 길이에 따라 달라지는 것으로 관찰되었고, 90 nt ssODN의 경우, 최적의 HDR 효율은 ~1.8%인 것을 관찰된 반면, 그보다 더 긴 ssODN은 HDR 효율을 감소시켰다 (도 3a 참조). dsDNA 도너가 뉴클레아제와 함께 사용될 때, 더욱 긴 상동성 영역은 HDR 비율을 개선시키기 때문에 (참고 문헌 40 (상기 문헌은 그 전문가 본원에서 참조로 포함된다) 참조), 본 결과에 대하여 가능한 이유는 ssODN이 선택적 게놈 수복 프로세스에서 사용될 수 있고; 더욱 긴 ssODN이 게놈 수복 장치에 덜 이용가능하거나; 또는 더욱 긴 ssODN은 dsDNA 도너에 비하여, 더욱 긴 상동성에 의해 획득되는 임의의 개선을 상쇄시키는 부정적 효과를 초래할 수 있다는 점이다 (참고 문헌 41 (상기 문헌은 그 전문가 본원에서 참조로 포함된다) 참조). 그러나, 상기의 처음 두 이유 중 어느 것이 되는 경우, 이때 NHEJ 비율은, NHEJ 수복이 ssODN 도너를 포함하지 않기 때문에, 영향을 받지 않거나, 또는 ssODN이 더 길어짐에 따라 증가하여야 한다. 그러나, NHEJ 비율은 HDR과 함께 감소하는 것으로 관찰되었으며 (도 3a 참조), 이는 더욱 긴 ssODN이 상쇄 효과를 나타낸다는 것을 제안한다. 더욱 긴 ssODN이 세포에 독성이거나 (참고 문헌 42 (상기 문헌은 그 전문가 본원에서 참조로 포함된다) 참조), 또는 더욱 긴 ssODN의 형질감염이 DNA 프로세싱 기구를 포화시킴으로써 DNA 물 흡수량을 감소시키고, 리TALEN 플라스미드를 흡수하거나, 그를 발현시키는 세포의 능력을 감소시킨다는 것이 가능한 가설이 될 것이다.

[0148] ssODN 도너가 보유하는 미스매치의 도입 비율이 그의 이중 가닥 절단 ("DSB")까지의 거리에 따라 어떻게 달라지는지를 조사하였다. 모두가 리TALEN 쌍 #3의 스페이서 영역의 중앙에 동일한 2 bp 미스매치 (A)를 포함하는 일련의 90 nt ssODN을 디자인하였다. 각 ssODN은 또한 중앙부로부터 다른 거리에 제2의 2 bp 미스매치 (B)를 함유하였다 (도 3b 참조). 오직 중앙부의 2 bp 미스매치만을 포함하는 ssODN을 대조군으로 사용하였다. 상기 ssODN을 각각 리TALEN 쌍 #3과 함께 개별적으로 도입하였고, 그 결과는 GEAS를 이용하여 분석하였다. A 미스매치가 도입된 (A만 또는 A+B) 비율에 의해 측정된 바, 전체 HDR은 B 미스매치가 중앙부로부터 더 멀어짐에 따라 감소하였다는 것을 (도 3b 참조, 도 11a 참조) 본 발명자들은 알 수 있었다. B가 A로부터 단지 10 bp만큼 떨어져 있을 때 관찰된 바 전체 HDR 비율이 더 높다는 것은 dsDNA 절단 바로 옆에 인접해 있는 게놈 DNA에 대해 ssODN이 어닐링될 필요가 더 적다는 것을 반영할 수 있다.

[0149] A로부터의 B까지의 각 거리의 경우, HDR 이벤트 중 일부는 오직 A 미스매치만을 도입한 반면, 또 다른 일부는 A 및 B 미스매치, 둘 모두를 도입하였다 (도 3b 참조 (A만 및 A+B)). 상기의 두 결과는 ssDNA 올리고 길이에 따른 유전자 전환 트랙에 기인하는 것일 수 있으며 (참고 문헌 43 (상기 문헌은 그 전문가 본원에서 참조로 포함된다) 참조), 이로서, A+B 미스매치의 도입은 B 미스매치 너머까지 확장된 긴 전환 트랙으로부터 초래되었고, 오직 A 미스매치만의 도입은 B에 도달하지 못한 더 짧은 트랙으로부터 초래되었다. 이러한 해석시, ssODN을 따른 양 방향의 유전자 전환 길이의 분포를 추정하였다 (도 11b 참조). 추정된 분포는 유전자 전환 트랙이 그의 길이가 증가함에 따라 점진적으로 덜 빈번하게 발생하게 된다는 것을 암시하는 것이며, 이러한 결과는 dsDNA 도너를 이용하였을 때 관찰된 유전자 전환 트랙 분포와 매우 유사한 결과로서, 단, dsDNA 도너의 경우에는 수백 개의 염기인데 반해, ssDNA 도너의 경우에는 수십개 염기 만큼의 고도로 압축된 거리 규모일 때이다. 상기 결과와 일치하는 것으로서, 중앙의 2 bp 미스매치 "A" 양측에 10 nt 만큼의 간격을 두고 이격되어 있는 3쌍의 2 bp 미스매치를 포함하는 ssODN을 사용하여 수행한 실험에서는 A만이 단독으로 도입된 경우의 패턴을 다른 경우에는 다중의 B 미스매치가 도입된 경우에는 그때의 것의 86%가 되도록 하였다 (도 11c). 비록 B만의 도입 이벤트 횟수가 너무 낮아 길이가 10 np 미만인 트랙 분포를 추정하기 어렵기는 했지만, 뉴클레아제 부위의 10 bp 이내의 짧은 트랙 영역이 지배적이라는 것은 명백하다 (도 11b 참조). 최종적으로, 단일 B 미스매치를 이용하여 수행한 모든 실험에서, 작은 분율의 B만의 도입 이벤트가 관찰되었으며 (0.04%~0.12%), 이는 대략적으로는 A로부터 B까지의 모든 길이 간에 걸쳐 일정하였다.

[0150] 추가로, 여전히 도입은 관찰되면서, ssODN 도너가 리TALEN 유도 dsDNA 절단으로부터 얼마나 멀리 떨어져 배치될 수 있는지에 관하여 분석하였다. 리TALEN 유도 dsDNA 절단 부위로부터 더 큰 거리 범위 (-600 bp 내지 +400 bp) 만큼 떨어져 있는 부위를 표적화하는, 중앙 2 bp 미스매치를 포함하는 90 nt ssODN 세트를 시험하였다. ssODN이 ≥ 40 bp 만큼 떨어져 매칭될 때, 본 발명자들은 절단 영역에 걸쳐 중앙부에 위치한 대조군 ssODN과 비

교하였을 때 >30x 더 낮은 HDR 효율을 관찰하였다 (도 3c 참조). 게놈이 ssDNA 도너 단독에 의해 변경된 실험에서 관찰된 바와 같이, 관찰된 낮은 수준의 도입은 dsDNA 절단과 관련이 없는 프로세스에 기인하는 것일 수 있다 (참고 문헌 42 (상기 문헌은 그 전문이 본원에서 참조로 포함된다) 참조). 한편, ssODN이 ~40 bp 만큼 떨어져 있을 때, 낮은 수준의 HDR이 존재하는 것은 dsDNA 절단의 미스매치를 함유하는 쪽에 위치하는 약화된 상동성과, dsDNA 절단의 나머지 다른 한쪽 상의 불충분한 ssODN 올리고 길이와의 조합에 기인하는 것일 수 있다.

[0151] Cas9-gRNA 매개 표적화를 위한 ssODN DNA 도너 디자인을 시험하였다. AAVS1 유전자좌를 표적화하는 Cas9-gRNA (C2)를 구축하고, 다양한 배향 (Oc: gRNA에 상보적 및 On: gRNA에 비상보적) 및 길이 (30, 50, 70, 90, 110 nt)의 ssODN 도너를 디자인하였다. Oc는 On보다 더 우수한 효율을 달성하였고, 70 mer Oc는 1.5%의 최적의 HDR 비율을 달성하였다 (도 3d 참조). ssODN과 함께 Cas9-유래 니카제 (Cc: Cas9_D10A)에 의해 매개된 HDR 효율이 C2보다 유의적으로 더 낮았다는 사실에도 불구하고 (t 검정, 쌍 형성 단부, P=0.02), Cc를 사용하였을 때, 같은 ssODN 가닥 편향성이 검출되었다 (도 12 참조).

[0152] 실시예 XIV

[0153] 교정된 세포의 hiPSC 클론 단리

[0154] GEAS를 통해 리TALEN 쌍 #3이 hiPSC에서, 정확하게 편집된 세포가 일반적으로 클론 스크리닝에 의해 단리될 수 있는 수준인 ~1%의 효율로 정확한 게놈 편집을 달성하였다는 것이 밝혀졌다. HiPSC는 단일 세포만큼 열등한 생존 능력을 가진다. 단일 세포 FACS 분류 방법과 함께 참고 문헌 23 (상기 문헌은 그 전문이 본원에서 참조로 포함된다)에 기술된 최적화된 프로토콜을 사용하여 단일 hiPSC 분류 및 유지를 위한 강력한 플랫폼을 확립하였고, 여기서, hiPSC 클론은 >25%의 생존율로 회수될 수 있다. 상기 방법을 신속하고 효율적인 유전형 분석 시스템과 조합하여 1시간 이내에 단일 튜브 반응에서 염색체 DNA 추출 및 표적화된 게놈 증폭을 수행하였으며, 이로써 편집된 hiPSC의 대규모 유전형 분석을 수행할 수 있었다. 동시에, 상기 방법은 선별 없이 게놈 편집된 hiPSC를 강건하게 수득하기 위한 파이프라인을 포함한다.

[0155] 본 시스템을 입증하기 위하여 (도 4a 참조), PGP1 hiPSC를 부위 #3에서 ssODN 표적화 CCR5 및 리TALEN 쌍으로 형질감염시켰다 (표 3 참조). 형질감염된 세포 중 일부를 이용하여 GEAS를 수행하였고, 이 결과, HDR 빈도는 1.7%인 것으로 나타났다 (도 4b 참조). 분류된 단일 세포 클론의 회수율이 25%라는 것과 함께, 상기의 정보를 통해 5개의 96 웰 플레이트로부터 1개 이상의 정확하게 편집된 클론을 98%의 포아송 확률(Poisson probability) ($\mu=0.017*0.25*96*5*2$ 로 가정할 때)로 수득할 수 있다고 추정할 수 있다. 형질감염 후 6일 경과하였을 때, hiPSC를 FACS 분류하고, 분류 후 8일 경과하였을 때, 100개의 hiPSC 클론을 스크리닝하였다. 생어 서열분석 결과, 상기의 비선별된 hiPSC 콜로니 100개 중 2개가 ssODN 도너에 의해 도입된 2 bp 돌연변이를 포함하는 이형접합성 유전자형을 함유하는 것으로 밝혀졌다 (도 4c 참조). 1% ($1\%=2/2*100$, 스크리닝된 100개의 세포 중 2개의 단일 대립유전자 교정 클론)의 표적화 효율은 차세대 서열분석 분석 (1.7%)과 일치하는 것이었다 (도 4b 참조). 생성된 hiPSC의 다능성은 SSEA4 및 TRA-1-60에 대한 면역염색법으로 확인하였다 (도 4d 참조). 성공적으로 표적화된 hiPSC 클론은 3개의 배엽 모두를 특징으로 하는 성숙한 테라토마를 생성할 수 있었다 (도 4e 참조).

[0156] 실시예 XV

[0157] 연속적인 세포 게놈 편집 방법

[0158] 특정 측면에 따라, 세포가 표적 DNA에 상보적인 RNA와 공동 국제화 복합체를 형성하고, 부위 특이적인 방식으로 표적 DNA를 절단하는 효소를 코딩하는 핵산을 포함하도록 유전적으로 변형된 것인, 인간 세포, 예를 들어, 인간 줄기 세포를 비롯한 세포에서의 게놈 편집 방법을 제공한다. 상기 효소는 RNA에 의해 가이드되는 DNA 결합 단백질, 예컨대, II형 CRISPR 시스템의 RNA에 의해 가이드되는 DNA 결합 단백질을 포함한다. 예시적인 효소는 Cas9이다. 본 측면에 따라, 세포는 효소를 발현하고, 세포 주변 배지로부터 세포에 가이드 RNA를 제공한다. 가이드 RNA 및 효소는 효소가 DNA를 절단하는 곳인 표적 DNA에서 공동 국제화 복합체를 형성한다. 임의적으로, 도너 핵산은 예를 들어, 비상동성 단부 연결 또는 상동성 재조합에 의해 절단 부위에서 DNA 내로의 삽입을 위해 존재할 수 있다. 한 측면에 따라, 표적 DNA에 상보적인 RNA와 공동 국제화 복합체를 형성하고, 부위 특이적인 방식으로 표적 DNA를 절단하는 효소, 예컨대, Cas9를 코딩하는 핵산은 프로모터의 영향하에 있으며, 예컨대, 핵산은 활성화되고, 침묵화될 수 있다. 상기 프로모터는 통상의 기술자에게 주지되어 있다. 예시적인 한 프로모터는 dox 유도성 프로모터이다. 한 측면에 따라, 세포는 표적 DNA에 상보적인 RNA와 공동 국제화 복합체를 형성하고, 부위 특이적인 방식으로 표적 DNA를 절단하는 효소를 코딩하는 핵산을 그의 게놈 내로 가역적으로 삽입함으로써 유전적으로 변형된다. 일단 삽입되고 나면, 핵산은 시약, 예컨대, 트랜스포사제의 사용에 의해 제거

될 수 있다. 상기 방식으로, 핵산은 사용 이후에 쉽게 제거될 수 있다.

- [0159] 한 측면에 따라, CRISPR 시스템을 이용한 인간 유도 다능성 줄기 세포 (hiPSC)에서의 연속적인 게놈 편집 시스템을 제공한다. 예시적인 측면에 따라, 본 방법은 Cas9가 게놈 내에 가역적으로 삽입되어 있는 hiPSC 세포주 (Cas9-hiPSC); 및 그의 천연 형태로부터 Cas9와 함께 사용될 수 있도록 세포 주변 배지로부터 세포로 그가 통과할 수 있도록 변형된 gRNA를 이용하는 것을 포함한다. 포스페이트 기를 제거하는 방식으로 포스파타제를 이용하여 상기 gRNA를 처리하였다. 세포에서의 게놈 편집은 Cas9와 함께 조직 배양 배지 중 포스파타제로 처리된 gRNA를 보충함으로써 수행된다. 본 접근법을 통해서 현재까지 보고된 최고 효율보다 2-10X 배 더 효율적인 방식으로, 단일 처리일인 경우에는 최대 50%까지의 효율로 HiPSC에서 흔적 없는 게놈 편집을 수행할 수 있다. 추가로, 본 방법은 사용이 용이하고, 세포 독성이 유의적으로 더 낮다. 본 개시내용의 실시양태는 생물학적 연구 및 진단 적용을 위한 hiPSC의 단일 편집, 생물학적 연구 및 진단 적용을 위한 hiPSC의 다중 편집, hiPSC의 지향적 진화 및 hiPSC 및 그의 유도체 세포의 표현형 스크리닝을 포함한다.
- [0160] 특정 측면에 따라, 줄기 세포 이외에도 본원에 기술된 다른 세포주 및 유기체가 사용될 수 있다. 예를 들어, 안정한 Cas9 통합 마우스 세포 및 래트 세포를 생성할 수 있고, 포스파타제 처리된 gRNA를 세포 주변 배지로부터 국소적으로 도입함으로써 조직 특이 게놈 편집을 수행할 수 있도록 동물 세포, 예컨대, 마우스 또는 래트 세포에 대해 본원에 기술된 방법이 사용될 수 있다. 또한, 다른 Cas9 유도체가 다수의 세포주 및 유기체 내로 삽입될 수 있고, 표적화된 게놈 조작, 예컨대, 서열 특이 니킹, 유전자 활성화, 억제 및 후성적 변형이 수행될 수 있다.
- [0161] 본 개시내용의 측면은 Cas9가 게놈 내로 삽입된 안정한 hiPSC를 제조하는 것에 관한 것이다. 본 개시내용의 측면은 세포벽을 통해 세포 내로 진입할 수 있고, 세포의 면역 반응을 피하면서 Cas9와 공동 국재화할 수 있도록 RNA를 변형시키는 것에 관한 것이다. 상기 변형된 가이드 RNA를 최소의 독성으로 최적의 형질감염 효율을 달성할 수 있다. 본 개시내용의 측면은 포스파타제 처리된 gRNA를 사용한 Cas9-hiPSC에서의 최적화된 게놈 편집에 관한 것이다. 본 개시내용의 측면은 hiPSC로부터 Cas9를 제거하여 흔적 없는 게놈 편집을 달성하고, 여기서, Cas9를 코딩하는 핵산은 게놈 내로 가역적으로 배치된 것인 것을 포함한다. 본 개시내용의 측면은 원하는 유전자 돌연변이를 생성하기 위해 Cas9가 게놈 내로 삽입된 hiPSC를 사용하는 생체 의학적 조작을 포함한다. 상기 조작된 hiPSC는 다능성을 유지하고, 환자 세포주의 표현형을 완전하게 재현하는, 심장 근육 세포를 비롯한 각종 세포 유형으로 성공적으로 분화될 수 있다.
- [0162] 본 개시내용의 측면은 다중 게놈 편집을 위한 포스파타제 처리된 gRNA 라이브러리를 포함한다. 본 개시내용의 측면은 약물 스크리닝을 위한 자원으로서의 역할을 할 수 있는 것인, 각각의 것이 1개를 게놈 중 수개의 지정된 돌연변이를 운반하는 PGP 세포주의 라이브러리를 생성하는 것을 포함한다. 본 개시내용의 측면은 레트로트랜스 요소 모두가 상기 요소의 위치 및 활성을 추적하기 위해 상이한 서열로 바코딩된 것인 PGP1 세포주를 생성하는 것을 포함한다.
- [0163] 실시예 XVI
- [0164] Cas9가 게놈 내로 삽입된 안정한 hiPSC 생성
- [0165] dox 유도성 프로모터 하에 Cas9를 코딩하고, 피기백 트랜스포사제의 도움으로 게놈 내로 삽입되고, 그에서 제거될 수 있는 피기백 벡터 내로 구축물을 배치하였다. PCR 반응을 통해 벡터의 안정한 삽입을 확인하였다 (도 14 참조). RT-QPCR을 통해 유도성 Cas9 발현을 측정하였다. 배양 배지 중 1 ug/mL DOX 보충 후 8시간 경과하였을 때, Cas9의 mRNA 수준은 1,000X 증가하였고, DOX 제거 후 ~20시간 경과시 Cas9 mRNA 수준은 정상 수준으로 하락하였다 (도 15 참조).
- [0166] 한 측면에 따라, Cas9-hiPSC 시스템 기반 게놈 편집은 hiPSC에서의 형질감염 효율이 보통 <1%인 큰 구축물인 Cas9 플라스미드/RNA의 형질감염 방법을 우회한다. 본 Cas9-hiPSC 시스템은 인간 줄기 세포에서 고도로 효율적인 게놈 조작을 수행하기 위한 플랫폼으로서의 역할을 할 수 있다. 추가로, 피기백 시스템을 사용하여 hiPSC 내로 도입된 Cas9 카세트는 트랜스포사제 도입시 게놈으로부터 쉽게 제거될 수 있다.
- [0167] 실시예 XVII
- [0168] 포스파타제 처리된 가이드 RNA
- [0169] Cas9-hiPSC에서의 연속적인 게놈 편집을 수행할 수 있도록 하기 위해, gRNA를 코딩하는 일련의 변형된 RNA를 생성하고, 리포솜과 함께 복합체로 Cas9-iPS 배양 배지 내로 보충되었다. 캡핑되지 않은, 포스파타제 처리된 천

연 RNA은 앞서 보고된 5' 캡-Mod RNA보다 30X 더 큰 13%의 최적의 HDR 효율을 달성하였다 (도 16 참조).

- [0170] 한 측면에 따라, 가이드 RNA는 물리적으로 도너 DNA에 부착된다. 이러한 방식으로, 본 방법은 Cas9 매개 게놈 절단 및 ssODN-매개 HDR을 커플링시켜 서열 특이 게놈 편집을 자극하는 것을 제공한다. 최적의 농도로 DNA ssODN 도너와 연결된 gRNA는 44% HDR 및 비특이 NHEJ 2%를 달성하였다 (도 17 참조). 특히, 본 방법은 뉴클레오핵션 또는 전기천공시에 관찰되는 것과 같은 뚜렷한 독성을 초래하지 않는다.
- [0171] 한 측면에 따라, 본 개시내용은 게놈적으로 삽입된 Cas9와 함께 높은 형질감염 효율, 게놈 편집 효율을 달성한, gRNA를 코딩하는 시험관내 조작된 RNA 구조를 제공한다. 추가로, 본 개시내용은 상동성 지정 재조합 반응과 함께 게놈 절단 이벤트를 커플링하는 gRNA-DNA 키메라 구축물을 제공한다.
- [0172] 실시예 XVIII
- [0173] 가역적으로 조작된 Cas9를 hiPSC로부터 제거함으로써 흔적 없는 게놈 편집 달성
- [0174] 특정 측면에 따라, 가역성 벡터를 이용하여 Cas9 카세트를 hiPSC 세포의 게놈 내로 삽입한다. 따라서, 피기백 벡터를 이용함으로써 hiPSC 세포의 게놈 내로 Cas9 카세트를 가역적으로 삽입하였다. 세포를 트랜스포사제 코딩 플라스미드로 형질감염시킴으로써 게놈 편집된 hiPSC로부터 Cas9 카세트를 제거하였다. 따라서, 본 개시내용의 측면은 통상의 기술자에게 주지되어 있는, 가역성 벡터를 사용하는 것을 포함한다. 가역성 벡터는 게놈 내로 삽입될 수 있고, 이어서, 예를 들어, 상응하는 벡터 제거 효소로 제거될 수 있는 벡터이다. 상기 벡터 및 상응하는 벡터 제거 효소는 통상의 기술자에게 공지되어 있다. 전이증식된 iPS 세포에서 스크린을 수행하였고, PCR 반응에 의해 확인하여 Cas9-카세트가 없는 콜로니를 회수하였다. 따라서, 본 개시내용은 세포에 영구적인 Cas9 카세트가 존재하도록 함으로써 나머지 게놈에는 영향을 미치지 않으면서 게놈 편집을 수행할 수 있는 방법을 제공한다.
- [0175] 실시예 XIX
- [0176] iPGP1 세포에서의 게놈 편집
- [0177] 심근증의 병인론에 관한 연구는 역사적으로 적합한 모델 시스템의 부족으로 인해 제대로 이루어지지 못했다. 환자 유래의 유도 다능성 줄기 세포 (iPSC)의 심장 근육 세포 분화는 이러한 장벽을 극복할 수 있는 한 유망한 방안을 제공하며, 심근증의 iPSC 모델링에 관한 보고가 출현하기 시작하였다. 그러나, 이러한 유망성을 실현하기 위해서는 환자 유래의 iPSC 세포주의 유전적 이형접합성을 극복할 수 있는 접근법을 필요로 할 것이다.
- [0178] Cas9-iPGP1 세포주 및 DNA에 결합한 포스파타제 처리된 가이드 RNA를 사용하여, 바르트 증후군 환자에서는 단일 뉴클레오티드 결실을 보유하는 것으로 확인된 TAZ 엑손 6에서의 서열을 제외하면 동질유전자인 것인 3개의 iPSC 세포주를 생성하였다. 1회차의 RNA 형질감염으로 ~30%의 HDR 효율을 달성하였다. 원하는 돌연변이를 포함하는 변형된 Cas9-iPGP1 세포를 전이증식시키고 (도 18 참조), 세포주를 심장 근육 세포로 분화시켰다. 조작된 Cas9-iPGP1로부터 유래된 심장 근육 세포는 환자 유래의 iPSC에서, 및 신생 래트의 TAZ 너다운 모델에서 관찰되는 카르디올리핀, 미토콘드리아, 및 ATP 결핍을 완전하게 재현하였다 (도 19 참조). 따라서, 다능성 세포에서 질환을 유발하는 돌연변이를 교정한 후, 세포를 원하는 세포 유형으로 분화시키는 방법을 제공한다.
- [0179] 실시예 XX
- [0180] 물질 및 방법
- [0181] 1. 피기백 Cas9 dox 유도성 안정한 인간 iPS/ES 세포주 확립.
- [0182] 1. 세포가 70% 컨플루언스에 도달한 후, 배양물을 최종 농도 10 uM의 ROCK 억제제 Y27632로 밤새도록 전처리 한다.
- [0183] 2. 다음날, 멸균 1.5 ml 에펜도르프 튜브 중에서 82 μ l의 인간 줄기 세포 뉴클레오핵터 용액 및 18 uL의 보충제 1을 조합하여 뉴클레오핵션 용액을 제조한다. 잘 혼합한다. 용액을 37°C에서 5 min 동안 인큐베이션시킨다.
- [0184] 3. mTeSR1을 흡인하고; 6 웰 플레이트 중 2 mL/웰로 DPBS를 이용하여 세포를 온화하게 세정한다.
- [0185] 4. DPBS를 흡인하고, 2 mL/웰의 베르센(Versene)을 첨가하고, 배양물이 한데 모이고, 느슨한 부착성을 띠되, 단, 탈착되지는 않을 때까지 배양물을 37°C 인큐베이터에 다시 돌려 놓는다. 3-7 min이 소요된다.
- [0186] 5. 온화하게 베르센을 흡인하고, mTeSR1을 첨가한다. 1 ml mTeSR1을 첨가하고, 1,000 uL 마이크로피펫을 이

용하여 세포 상으로 mTeSR1을 온화하게 유도시킴으로써 세포를 떼어낸다.

[0187] 6. 떼어낸 세포를 수집하고, 그를 단일 세포 현탁액으로 온화하게 분쇄시키고, 혈구계수기에 의해 정량화하고, 세포 밀도를 1 ml당 세포 100만개인 밀도로 조정한다.

[0188] 7. 1 ml 세포 현탁액을 1.5 ml 에펜도르프 튜브로 첨가하고, 벤치 탑 원심분리기에서 5 min 동안 1,100 RPM으로 원심분리시킨다.

[0189] 8. 단계 2로부터의 100 μ l 인간 줄기 세포 뉴클레오펙터 용액 중에 세포를 재현탁시킨다.

[0190] 9. 1 ml 피펫 팁을 이용하여 세포를 뉴클레오펙터 큐벳으로 옮겨 놓는다. 큐벳 중 세포 현탁액으로 1 μ g의 플라스미드 트랜스포사제 및 5ug PB Cas9 플라스미드를 첨가한다. 온화하게 와동시켜 세포 및 DNA를 혼합한다.

[0191] 10. 큐벳을 뉴클레오펙터 안으로 넣는다. 프로그램 B-016을 선택하고, X 버튼을 눌러 세포를 뉴클레오펙션시킨다.

[0192] 11. 뉴클레오펙션시킨 후, 큐벳에 ROCK 억제제와 함께 500 ul mTeSR1 배지를 첨가한다.

[0193] 12. 제공된 파스퇴르(Pasteur) 플라스틱 피펫을 이용하여 큐벳으로부터 뉴클레오펙션된 세포를 흡인한다. 이어서, ROCK 억제제와 함께 mTeSR1 배지가 있는 6 웰 플레이트의 매트릭셀 코팅 웰 내로 세포를 적가 방식으로 옮겨 놓는다. 세포를 37°C에서 밤새도록 인큐베이션시킨다.

[0194] 13. 다음날, 배지를 mTesr1로 교체하고, 72시간 동안의 형질감염 후; 1 ug/ml의 최종 농도로 푸로마이신을 첨가한다. 이어서, 7일 이내에 세포주가 확립될 것이다.

[0195] 2. RNA 제조

[0196] 1. gRNA 코딩 서열 상류에 T7 프로모터를 포함하는 DNA 주형을 제조한다.

[0197] 2. 메가 클리어 퓨리피케이션(Mega Clear Purification)을 이용하여 DNA를 정제하고, 농도를 정규화시킨다.

[0198] 3. 상이한 gRNA 제조를 위해 시판용 NTPS 혼합물을 제조한다.

#1 천연 RNA 믹스	[최종] (mM)
GTP	7.5
ATP	7.5
CTP	7.5
UTP	7.5
전체 부피	

#2 캡핑된 천연 RNA 믹스	[최종] (mM)
3'-O-Me-m7G 캡 구조 유사체 (NEB)	6
GTP	1.5
ATP	7.5
CTP	7.5
UTP	7.5
전체 부피	

#3 변형된 RNA 믹스	[최종] (mM)
GTP	7.5
ATP	7.5
5-Me-CTP (트리-링크)	7.5
슈도-UTP (트리-링크)	7.5
전체 부피	

[0201]

#4 캡핑된/ 변형된 RNA 믹스	[최종] (mM)
3'-O-Me-m7G 캡 구조 유사체 (NEB)	6
GTP	1.5
ATP	7.5
5-Me-CTP (트리-링크)	7.5
슈도-UTP (트리-링크)	7.5
전체 부피	

[0202]

[0203]

4. 실온에서 시험관내 전사 믹스를 제조한다.

Amt (ul)	
얼음상의 시판용 NTPS (*상기 명시된 바와 같이 vol/IVT rxn으로 첨가)	NA
PCR 생성물 (100ng/ul) = 총 1,600 ng ([최종]=40 ng/ul)	16
RT에서 완충제 X10 (엠비온(Ambion)으로 부터의 메가스크립트 키트 (MEGAscript kit))	4
T7 효소 (엠비온으로부터의 메가스크립트 키트)	4

[0204]

[0205]

5. 37°C (써모사이클러)에서 4시간 (3-6 hr도 허용) 동안 인큐베이션시킨다.

[0206]

6. 각 샘플에 2 μl 터보(Turbo) DNase (엠비온으로부터의 메가스크립트 키트)를 첨가한다. 온화하게 혼합하고, 37°C에서 15' 동안 인큐베이션시킨다.

[0207]

7. 제조사의 설명서에 따라 엠비온으로부터의 메가클리어를 이용하여 DNase 처리된 반응물을 정제한다.

[0208]

8. 메가클리어를 이용하여 RNA를 정제한다 (정제된 RNA를 수개월 동안 -80에서 보관할 수 있다).

[0209]

9. 숙주 세포로부터의 톨(To11) 2 면역 반응을 피하기 위해 포스페이트 기를 제거한다.

RNA 포스파타제 처리	1X	12
각 RNA 샘플에 대하여	~100ul	NA
10X 엔탁틱(Antarctic) 포스파타제 완충제	11ul	132
엔탁틱 포스파타제	2ul	24

[0210]

[0211]

샘플을 온화하게 혼합하고, 37°C에서 30' (30'-1 hr도 허용) 동안 인큐베이션시킨다.

- [0212] 3. RNA 형질감염
- [0213] 1. 항생제 없이 48웰당 10K-20K 세포를 플레이팅한다. 형질감염을 위해 세포는 30-50% 컨플루언트여야 한다.
- [0214] 2. 형질감염 이전 적어도 2시간 전에 세포 배지를 B18R (200 ng/ml), DOX (1 ug/ml), 푸로마이신 (2 ug/ml)을 포함하는 것으로 교체한다.
- [0215] 3. gRNA (0.5 ug~2 ug), 도너 DNA (0.5 ug~2 ug) 및 RNAiMax를 함유하는 형질감염 시약을 제조하고, 실온에서 15분 동안 혼합물을 인큐베이션시키고, 세포로 옮겨 놓는다.
- [0216] 4. 단일 인간 iPS 세포 시딩 및 단일클론 수집
- [0217] 1. 4일간의 dox 유도 및 1일간의 dox 제거 후, 배지를 흡인하고, DPBS를 이용하여 온화하게 세정한다. 2 ml/웰의 베르센을 첨가하고, 배양물을 그라 라운드 업되고, 분리되는 것이 아니라, 느슨하게 부착성을 떨 때까지 37°C 인큐베이터에 다시 놓는다. 이를 통해 3-7 min이 소요된다.
- [0218] 2. 베르센을 온화하게 흡인하고, mTeSR1을 첨가한다. 1,000 uL 마이크로피펫을 이용하여 그 위로 mTeSR1을 온화하게 유동시킴으로써 1 ml mTeSR1을 첨가하고, 세포를 떼어낸다.
- [0219] 3. 떼어낸 세포를 수집하고, 그를 단일 세포 현탁액으로 온화하게 연마하고, 혈구계수기에 의해 정량화하고, 세포 밀도를 100K 세포/ml로 조정한다.
- [0220] 4. 세포를 10 cm 디쉬당 50K, 100K 및 400K의 세포 밀도로 mTeSR1 + ROCK 억제제를 포함하는 매트릭셀 코팅된 10 cm 디쉬 내로 시딩한다.
- [0221] 5. 단일 세포 형성 클론 스크리닝
- [0222] 1. 10 cm 디쉬에서 12일 동안 배양한 후, 클론은 육안으로 확인할 수 있고, 콜론 마커에 의해 표지화될 정도로 충분히 크다. 클론은 너무 크지 않게, 및 서로 부착되지 않게 한다.
- [0223] 2. 필터 팁이 있는 (10 uL로 설정된) P20 피펫을 사용하여 10 cm 디쉬를 배양 후드에 놓는다. 24 웰 플레이트의 한 웰에 대하여 10 uL 배지를 흡인한다. 클론을 작은 조각으로 스크래치하여 클론을 수집하고, 24 웰 플레이트 중 한 웰로 옮겨 놓는다. 각 클론에 대해 각각의 필터 팁을 이용한다.
- [0224] 3. 4-5일 경과 후, 24 웰 플레이트 중 한 웰 내의 클론의 크기는 분할할 수 있을 정도로 충분히 커진다.
- [0225] 4. 배지를 흡인하고, 2 mL/웰 DPBS로 세정한다.
- [0226] 5. DPBS를 흡인하고, 250 uL/웰 디스파제 (0.1 U/mL)로 교체하고, 37°C에서 7 min 동안 디스파제 중에서 세포를 인큐베이션시킨다.
- [0227] 6. 디스파제를 2 ml DPBS로 교체한다.
- [0228] 7. 250 uL mTeSR1을 첨가한다. 세포 스크래퍼를 이용하여 세포를 떼어내고, 세포를 수집한다.
- [0229] 8. 125 uL 세포 현탁액을 매트릭셀 코팅된 24 웰 플레이트의 웰로 옮겨 놓는다.
- [0230] 9. 게놈 DNA 추출을 위해 125 uL 세포 현탁액을 1.5 ml 에펜도르프 튜브로 옮겨 놓는다.
- [0231] 6. 클론 스크리닝
- [0232] 1. 단계 7.7로부터의 튜브를 원심분리한다.
- [0233] 2. 배지를 흡입하고, 웰당 250 uL 용해 완충제 (10 mM+트리스pH7.5+ (또는 +8.0), 10 mM EDTA, 10 mM을 첨가한다.
- [0234] 3. NaCl, +10% SDS, 40 ug/mL +프로테이나제 K(완충제 사용 전 새로 첨가)를 첨가한다.
- [0235] 4. 55에서 밤새도록 인큐베이션시킨다.
- [0236] 5. 250 uL 이소프로판올을 첨가하여 DNA를 침전시킨다.
- [0237] 6. 고속으로 30분 동안 회전시킨다. 70% 에탄올로 세척한다.
- [0238] 7. 에탄올을 온화하게 제거한다. 5 min 동안 대기 건조시킨다.

- [0239] 8. 100-200 ul dH2O를 이용하여 gDNA를 재현탁시킨다.
- [0240] 9. 특이 프라이머를 이용하여 표적화된 게놈을 PCR 증폭시킨다.
- [0241] 10. 각 프라이머를 이용하여 PCR 생성물을 생어 서열분석한다.
- [0242] 11. 생어 서열 데이터를 분석하고, 표적화된 클론을 확장시킨다.
- [0243] 7. 피기백 벡터 제거
- [0244] 1. 단계 2.1-2.9를 반복한다.
- [0245] 2. 1 ml 피펫 팁을 이용하여 세포를 뉴클레오펙터 큐벳으로 옮겨 놓는다. 2 μ g의 트랜스포사제 플라스미드를 큐벳 중의 세포 현탁액으로 첨가한다. 온화하게 와동시켜 세포 및 DNA를 혼합한다.
- [0246] 3. 단계 2.10-2.11을 반복한다.
- [0247] 4. 제공된 파스퇴르 플라스틱 피펫을 이용하여 큐벳으로부터 뉴클레오펙션된 세포를 흡입한다. 이어서, ROCK 억제제와 함께 mTeSR1 배지를 포함하는 10 cm 디쉬의 매트릭셀 코팅된 웰 내로 세포를 적가 방식으로 옮겨 놓는다. 세포를 37°C에서 밤새도록 인큐베이션시켰다.
- [0248] 5. 다음날, 배지를 mTesr1로 교체하고, 다음 4일 동안 매일 배지를 교체한다.
- [0249] 6. 클론 크기가 충분히 크게 된 후, 20-50개의 클론을 선택하고, 24 웰에 시딩한다.
- [0250] 7. PB Cas9 피기백 벡터 프라이머를 이용하여 클론의 유전자형을 분석하고, 음성 클론을 확장시킨다.
- [0251] 참고 문헌
- [0252] 참고 문헌은 명세서 전역에 걸쳐 하기의 그의 번호로 명시되어 있으며, 마치 본원에 전체가 기술되어 있는 것과 같이 본 명세서에 내에 포함된다. 하기 참고 문헌들은 각각 그 전문이 본원에서 참조로 포함된다.

1. Carroll,D. (2011) Genome engineering with zinc-finger nucleases. *Genetics*, 188, 773-82.
2. Wood,A.J., Lo,T.-W., Zeitler,B., Pickle,C.S., Ralston,E.J., Lee,A.H., Amora,R., Miller,J.C., Leung,E., Meng,X., et al. (2011) Targeted genome editing across species using ZFNs and TALENs. *Science (New York, N.Y.)*, 333, 307.
3. Perez-Pinera,P., Ousterout,D.G. and Gersbach,C.A. (2012) Advances in targeted genome editing. *Current opinion in chemical biology*, 16, 268-77.
4. Symington,L.S. and Gautier,J. (2011) Double-strand break end resection and repair pathway choice. *Annual review of genetics*, 45, 247-71.

[0253]

5. Urnov,F.D., Miller,J.C., Lee,Y.-L., Beausejour,C.M., Rock,J.M., Augustus,S., Jamieson,A.C., Porteus,M.H., Gregory,P.D. and Holmes,M.C. (2005) Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature*, 435, 646–51.
6. Boch,J., Scholze,H., Schornack,S., Landgraf,A., Hahn,S., Kay,S., Lahaye,T., Nickstadt,A. and Bonas,U. (2009) Breaking the code of DNA binding specificity of TAL-type III effectors. *Science (New York, N.Y.)*, 326, 1509–12.
7. Cell,P., Replacement,K.S., Talens,A., Type,A., Collection,C., Ccl-,A. and Quickextract,E. Genetic engineering of human pluripotent cells using TALE nucleases.
8. Mussolino,C., Morbitzer,R., Lütge,F., Dannemann,N., Lahaye,T. and Cathomen,T. (2011) A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. *Nucleic acids research*, 39, 9283–93.
9. Ding,Q., Lee,Y., Schaefer,E.A.K., Peters,D.T., Veres,A., Kim,K., Kuperwasser,N., Motola,D.L., Meissner,T.B., Hendriks,W.T., et al. (2013) Resource A TALEN Genome-Editing System for Generating Human Stem Cell-Based Disease Models.
10. Hockemeyer,D., Wang,H., Kiani,S., Lai,C.S., Gao,Q., Cassidy,J.P., Cost,G.J., Zhang,L., Santiago,Y., Miller,J.C., et al. (2011) Genetic engineering of human pluripotent cells using TALE nucleases. *Nature biotechnology*, 29, 731–4.
11. Bedell,V.M., Wang,Y., Campbell,J.M., Poshusta,T.L., Starker,C.G., Krug li,R.G., Tan,W., Penheiter,S.G., Ma,A.C., Leung,A.Y.H., et al. (2012) In vivo genome editing using a high-efficiency TALEN system. *Nature*, 490, 114–118.
12. Miller,J.C., Tan,S., Qiao,G., Barlow,K. a, Wang,J., Xia,D.F., Meng,X., Paschon,D.E., Leung,E., Hinkley,S.J., et al. (2011) A TALE nuclease architecture for efficient genome editing. *Nature biotechnology*, 29, 143–8.
13. Holkers,M., Maggio,I., Liu,J., Janssen,J.M., Miselli,F., Mussolino,C., Recchia,A., Cathomen,T. and Gonçalves,M. a F. V (2012) Differential integrity of TALE nuclease genes following adenoviral and lentiviral vector gene transfer into human cells. *Nucleic acids research*, 10.1093/nar/gks1446.

[0254]

14. Reyon,D., Tsai,S.Q., Khayter,C., Foden,J. a, Sander,J.D. and Joung,J.K. (2012) FLASH assembly of TALENs for high-throughput genome editing. *Nature Biotechnology*, 30, 460–465.
15. Qiu,P., Shandilya,H., D’Alessio,J.M., O’Connor,K., Durocher,J. and Gerard,G.F. (2004) Mutation detection using Surveyor nuclease. *BioTechniques*, 36, 702–7.
16. Mali,P., Yang,L., Esvelt,K.M., Aach,J., Guell,M., DiCarlo,J.E., Norville,J.E. and Church,G.M. (2013) RNA-guided human genome engineering via Cas9. *Science (New York, N.Y.)*, 339, 823–6.
17. Ding,Q., Regan,S.N., Xia,Y., Oostrom,L.A., Cowan,C.A. and Musunuru,K. (2013) Enhanced Efficiency of Human Pluripotent Stem Cell Genome Editing through Replacing TALENs with CRISPRs. *Cell Stem Cell*, 12, 393–394.
18. Cong,L., Ran,F.A., Cox,D., Lin,S., Barretto,R., Habib,N., Hsu,P.D., Wu,X., Jiang,W., Marraffini,L. a, et al. (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science (New York, N.Y.)*, 339, 819–23.
19. Cho,S.W., Kim,S., Kim,J.M. and Kim,J.-S. (2013) Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nature biotechnology*, 31, 230–232.
20. Hwang,W.Y., Fu,Y., Reyon,D., Maeder,M.L., Tsai,S.Q., Sander,J.D., Peterson,R.T., Yeh,J.-R.J. and Joung,J.K. (2013) Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nature biotechnology*, 31, 227–229.
21. Chen,F., Pruett-Miller,S.M., Huang,Y., Gjoka,M., Duda,K., Taunton,J., Collingwood,T.N., Frodin,M. and Davis,G.D. (2011) High-frequency genome editing using ssDNA oligonucleotides with zinc-finger nucleases. *Nature methods*, 8, 753–5.
22. Soldner,F., Laganière,J., Cheng,A.W., Hockemeyer,D., Gao,Q., Alagappan,R., Khurana,V., Golbe,L.I., Myers,R.H., Lindquist,S., et al. (2011) Generation of isogenic pluripotent stem cells differing exclusively at two early onset Parkinson point mutations. *Cell*, 146, 318–31.
23. Valamehr,B., Abujarour,R., Robinson,M., Le,T., Robbins,D., Shoemaker,D. and Flynn,P. (2012) A novel platform to enable the high-throughput derivation and characterization of feeder-free human iPSCs. *Scientific reports*, 2, 213.

[0255]

24. Sanjana, N.E., Cong, L., Zhou, Y., Cunniff, M.M., Feng, G. and Zhang, F. (2012) A transcription activator-like effector toolbox for genome engineering. *Nature protocols*, 7, 171–92.
25. Gibson, D.G., Young, L., Chuang, R., Venter, J.C., Iii, C.A.H., Smith, H.O. and America, N. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *6*, 12–16.
26. Zou, J., Maeder, M.L., Mali, P., Pruett-Miller, S.M., Thibodeau-Beganny, S., Chou, B.-K., Chen, G., Ye, Z., Park, I.-H., Daley, G.Q., et al. (2009) Gene targeting of a disease-related gene in human induced pluripotent stem and embryonic stem cells. *Cell stem cell*, 5, 97–110.
27. Perez, E.E., Wang, J., Miller, J.C., Jouvenot, Y., Kim, K. a, Liu, O., Wang, N., Lee, G., Bartsevich, V. V, Lee, Y.-L., et al. (2008) Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. *Nature biotechnology*, 26, 808–16.
28. Bhakta, M.S., Henry, I.M., Ousterout, D.G., Das, K.T., Lockwood, S.H., Meckler, J.F., Wallen, M.C., Zykovich, A., Yu, Y., Leo, H., et al. (2013) Highly active zinc-finger nucleases by extended modular assembly. *Genome research*, 10.1101/gr.143693.112.
29. Kim, E., Kim, S., Kim, D.H., Choi, B.-S., Choi, I.-Y. and Kim, J.-S. (2012) Precision genome engineering with programmable DNA-nicking enzymes. *Genome research*, 22, 1327–33.
30. Gupta, A., Meng, X., Zhu, L.J., Lawson, N.D. and Wolfe, S. a (2011) Zinc finger protein-dependent and -independent contributions to the in vivo off-target activity of zinc finger nucleases. *Nucleic acids research*, 39, 381–92.
31. Park, I.-H., Lerou, P.H., Zhao, R., Huo, H. and Daley, G.Q. (2008) Generation of human-induced pluripotent stem cells. *Nature protocols*, 3, 1180–6.
32. Cermak, T., Doyle, E.L., Christian, M., Wang, L., Zhang, Y., Schmidt, C., Baller, J.A., Somia, N. V, Bogdanove, A.J. and Voytas, D.F. (2011) Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic acids research*, 39, e82.
33. Briggs, A.W., Rios, X., Chari, R., Yang, L., Zhang, F., Mali, P. and Church, G.M. (2012) Iterative capped assembly: rapid and scalable synthesis of repeat-module DNA such as TAL effectors from individual monomers. *Nucleic acids research*, 10.1093/nar/gks624.
34. Zhang, F., Cong, L., Lodato, S., Kosuri, S., Church, G.M. and Arlotta, P. (2011) LETTERs Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. *29*, 149–154.

[0256]

35. Pathak,V.K. and Temin,H.M. (1990) Broad spectrum of in vivo forward mutations, hypermutations, and mutational hotspots in a retroviral shuttle vector after a single replication cycle: substitutions, frameshifts, and hypermutations. *Proceedings of the National Academy of Sciences of the United States of America*, 87, 6019–23.
36. Tian,J., Ma,K. and Saaem,I. (2009) Advancing high-throughput gene synthesis technology. *Molecular bioSystems*, 5, 714–22.
37. Zou,J., Mali,P., Huang,X., Dowey,S.N. and Cheng,L. (2011) Site-specific gene correction of a point mutation in human iPS cells derived from an adult patient with sickle cell disease. *Blood*, 118, 4599–608.
38. Mali,P., Yang,L., Esvelt,K.M., Aach,J., Guell,M., Dicarlo,J.E., Norville,J.E. and Church,G.M. (2013) RNA-Guided Human Genome.
39. Boyle,A.P., Davis,S., Shulha,H.P., Meltzer,P., Margulies,E.H., Weng,Z., Furey,T.S. and Crawford,G.E. (2008) High-resolution mapping and characterization of open chromatin across the genome. *Cell*, 132, 311–22.
40. Orlando,S.J., Santiago,Y., DeKolver,R.C., Freyvert,Y., Boydston,E. a, Moehle,E. a, Choi,V.M., Gopalan,S.M., Lou,J.F., Li,J., et al. (2010) Zinc-finger nuclease-driven targeted integration into mammalian genomes using donors with limited chromosomal homology. *Nucleic acids research*, 38, e152.
41. Wang,Z., Zhou,Z.-J., Liu,D.-P. and Huang,J.-D. (2008) Double-stranded break can be repaired by single-stranded oligonucleotides via the ATM/ATR pathway in mammalian cells. *Oligonucleotides*, 18, 21–32.
42. Rios,X., Briggs,A.W., Christodoulou,D., Gorham,J.M., Seidman,J.G. and Church,G.M. (2012) Stable gene targeting in human cells using single-strand oligonucleotides with modified bases. *PloS one*, 7, e36697.
43. Elliott,B., Richardson,C., Winderbaum,J., Jac,A., Jasin,M. and Nickoloff,J.A.C.A. (1998) Gene Conversion Tracts from Double-Strand Break Repair in Mammalian Cells Gene Conversion Tracts from Double-Strand Break Repair in Mammalian Cells. 18.
44. Lombardo,A., Genovese,P., Beausejour,C.M., Colleoni,S., Lee,Y.-L., Kim,K. a, Ando,D., Urnov,F.D., Galli,C., Gregory,P.D., et al. (2007) Gene editing in human stem cells using zinc

[0257]

finger nucleases and integrase-defective lentiviral vector delivery. Nature biotechnology, 25, 1298–306.

45. Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. a and Charpentier, E. (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science (New York, N.Y.), 337, 816–21.
46. Shrivastav, M., De Haro, L.P. and Nickoloff, J. a (2008) Regulation of DNA double-strand break repair pathway choice. Cell research, 18, 134–47.
47. Kim, Y.G., Cha, J. and Chandrasegaran, S. (1996) Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. Proceedings of the National Academy of Sciences of the United States of America, 93, 1156–60.
48. Mimitou, E.P. and Symington, L.S. (2008) Sae2, Exo1 and Sgs1 collaborate in DNA double-strand break processing. Nature, 455, 770–4.
49. Doyon, Y., Choi, V.M., Xia, D.F., Vo, T.D., Gregory, P.D. and Holmes, M.C. (2010) Transient cold shock enhances zinc-finger nuclease-mediated gene disruption. Nature methods, 7, 459–60.

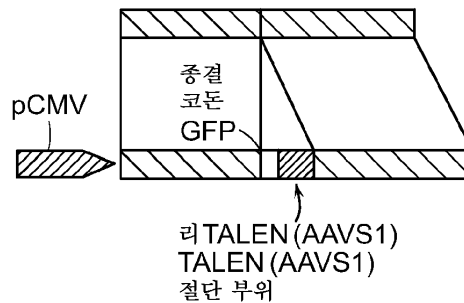
[0258]

도면

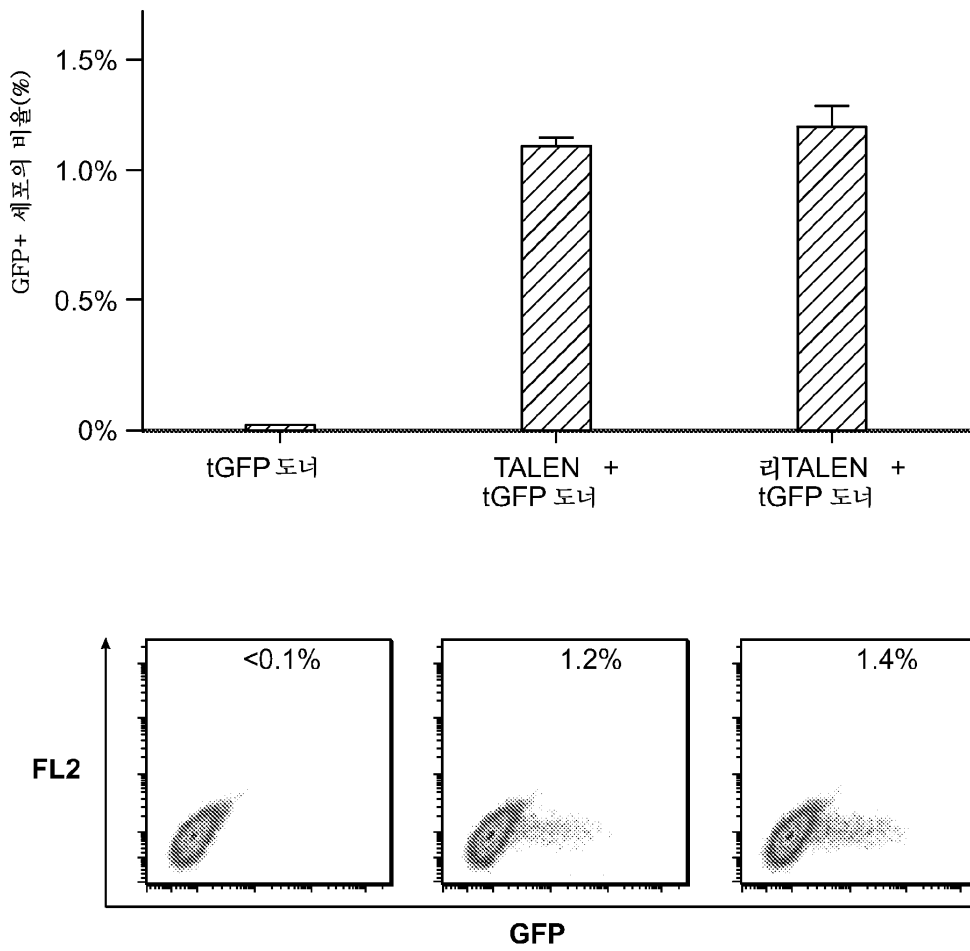
도면 1a

tGFP 도너

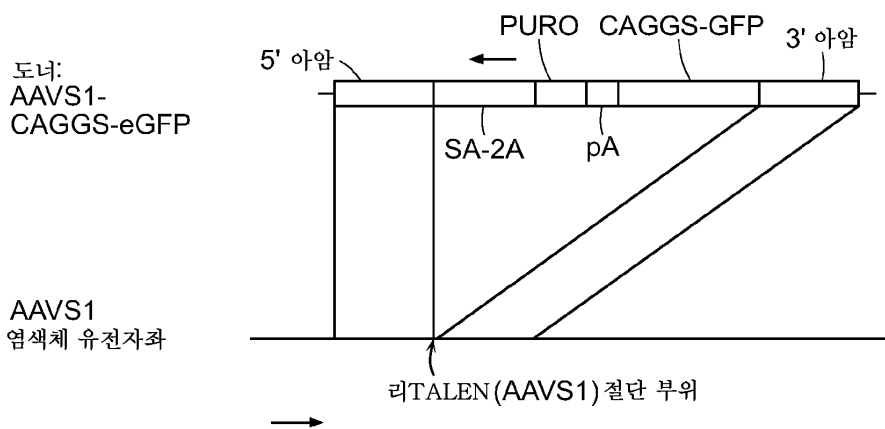
염색체 리포터



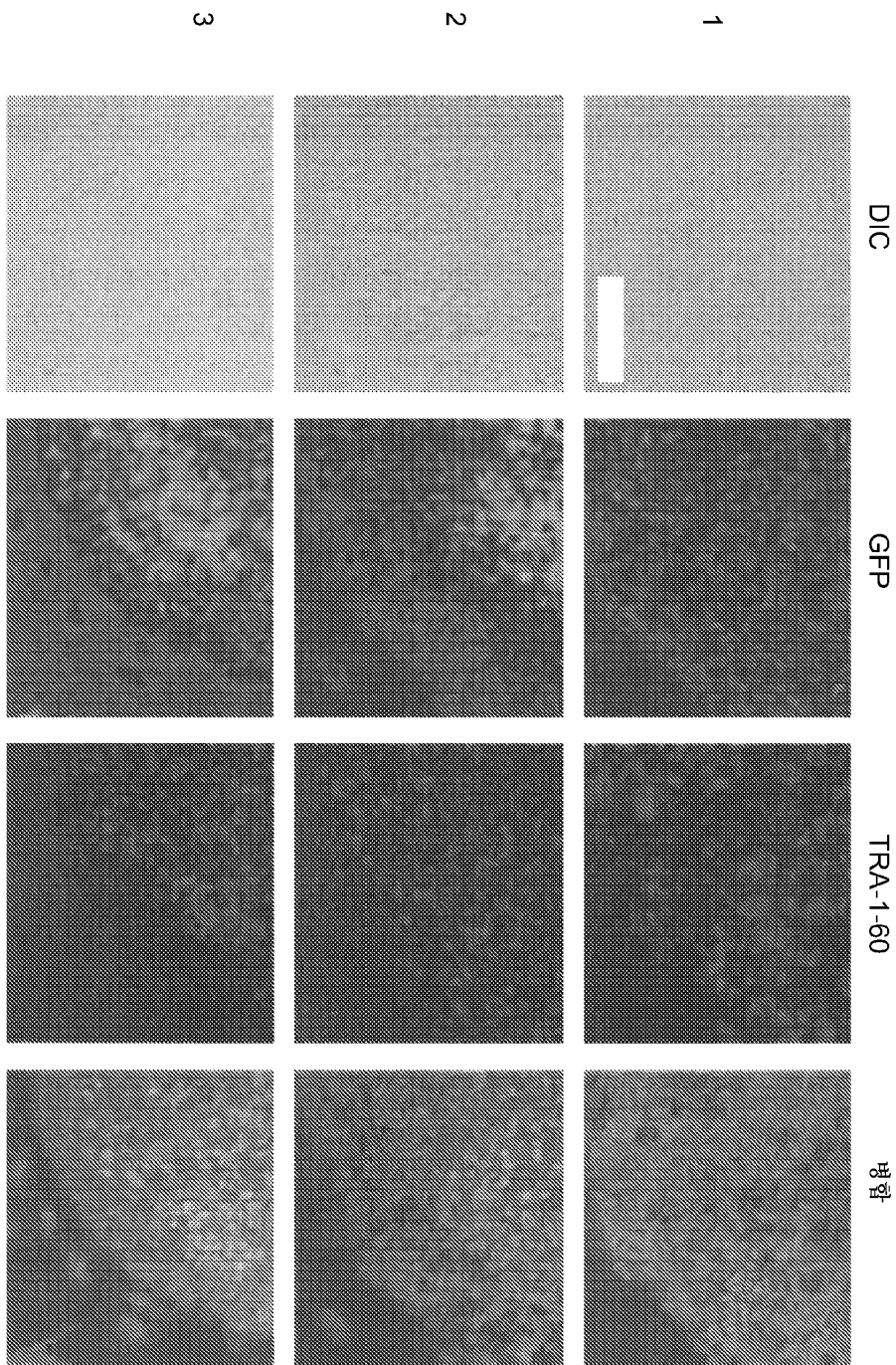
도면1b



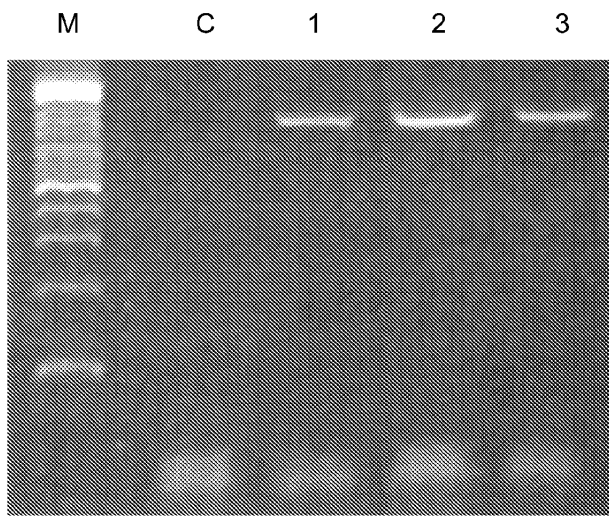
도면1c



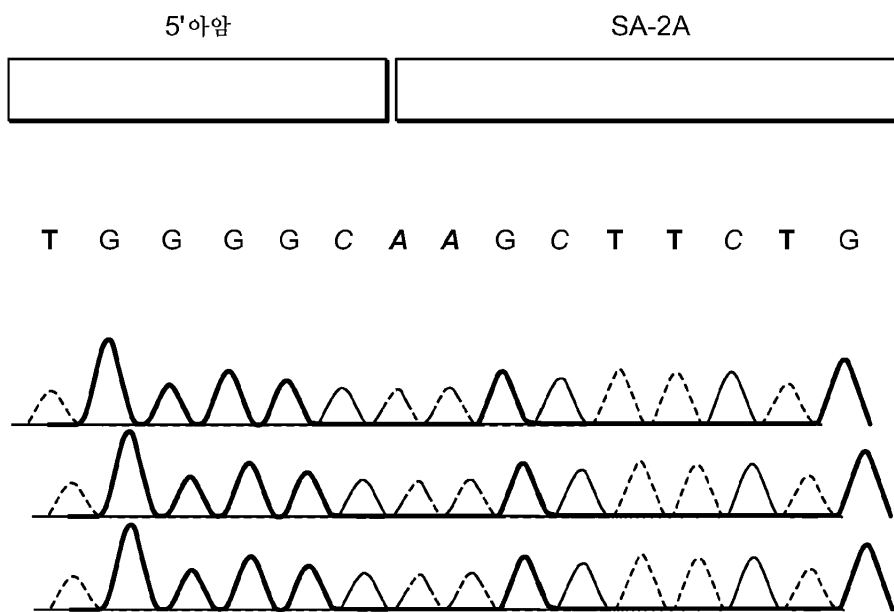
도면1d



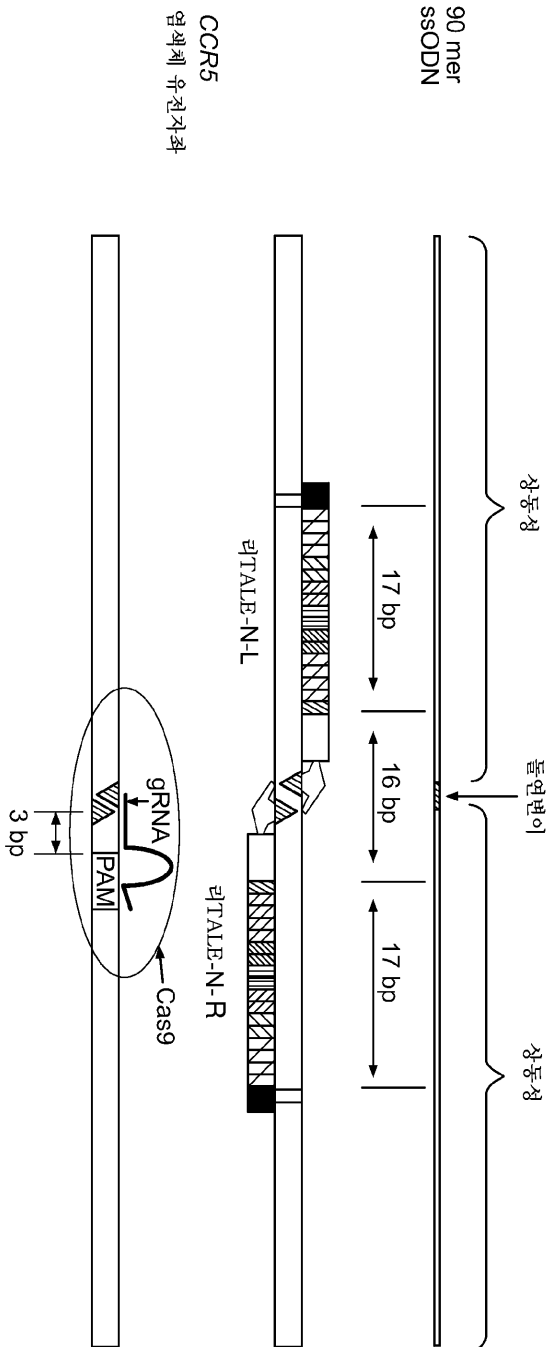
도면1e



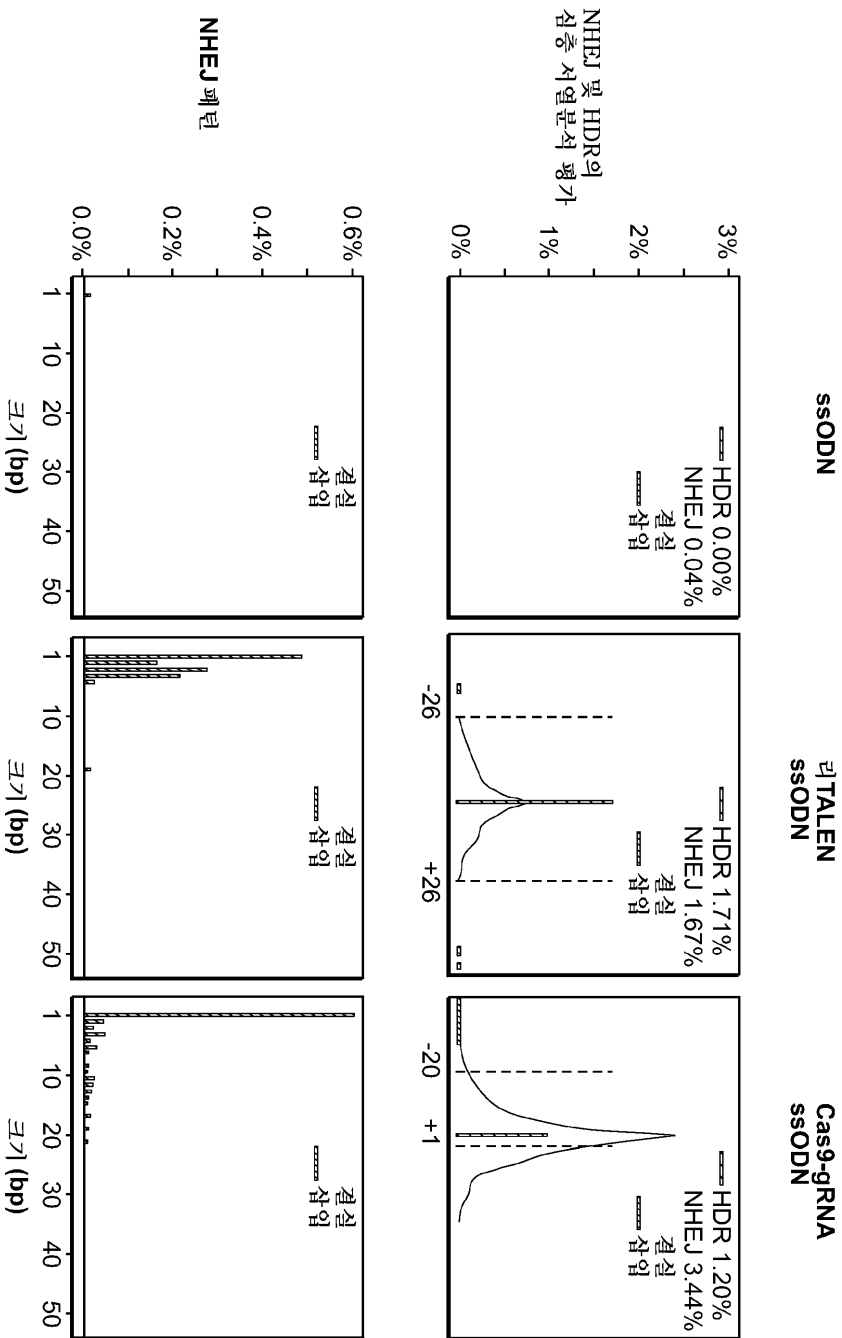
도면1f



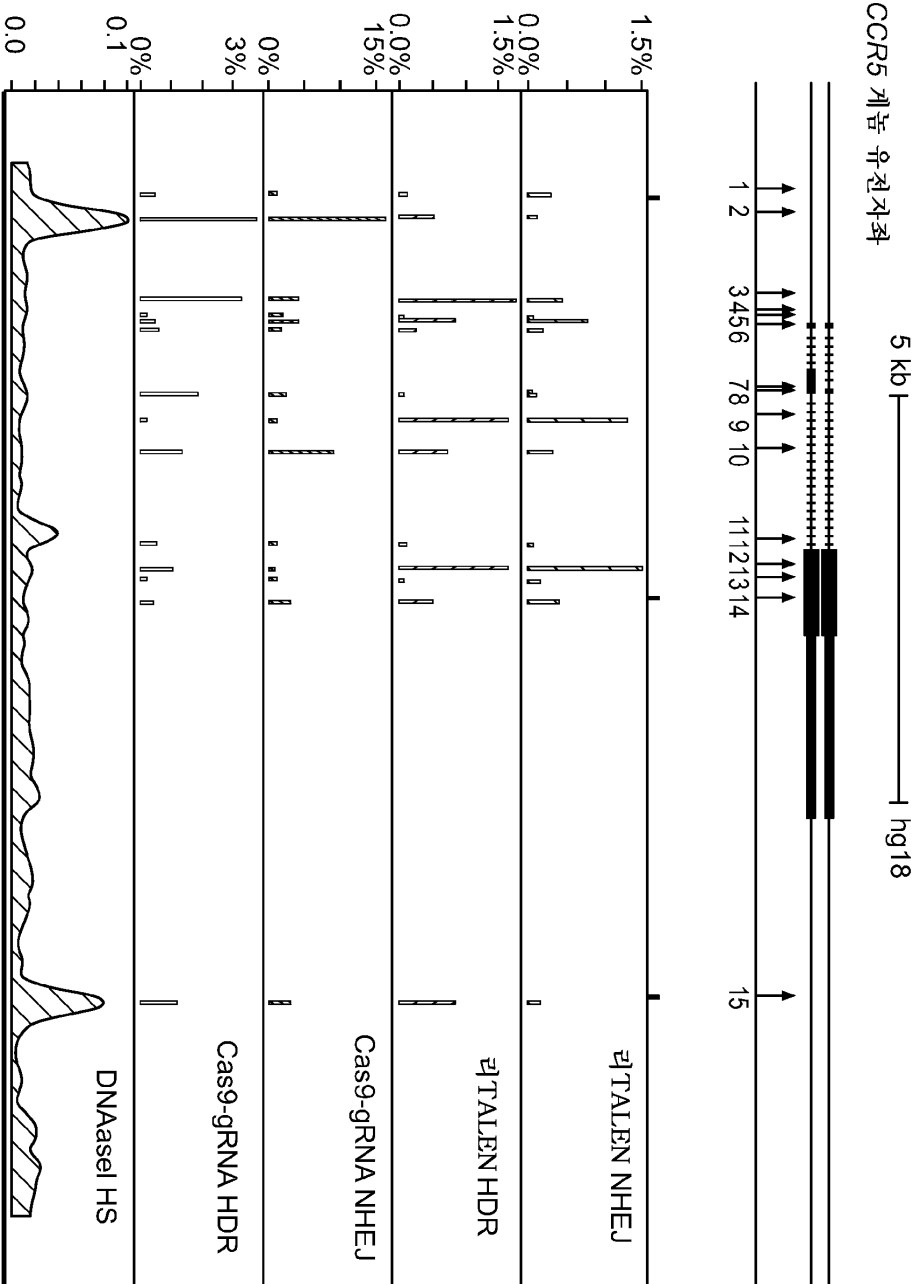
도면2a



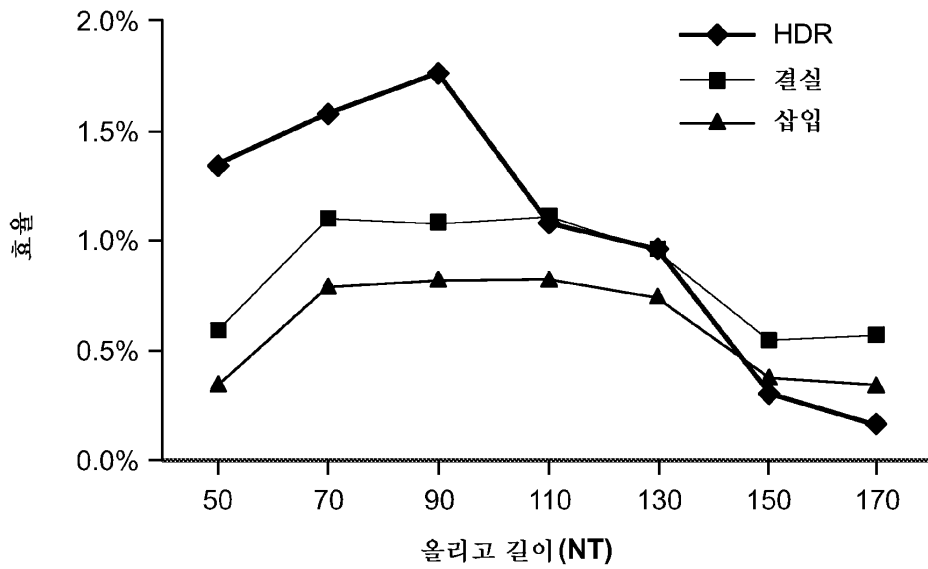
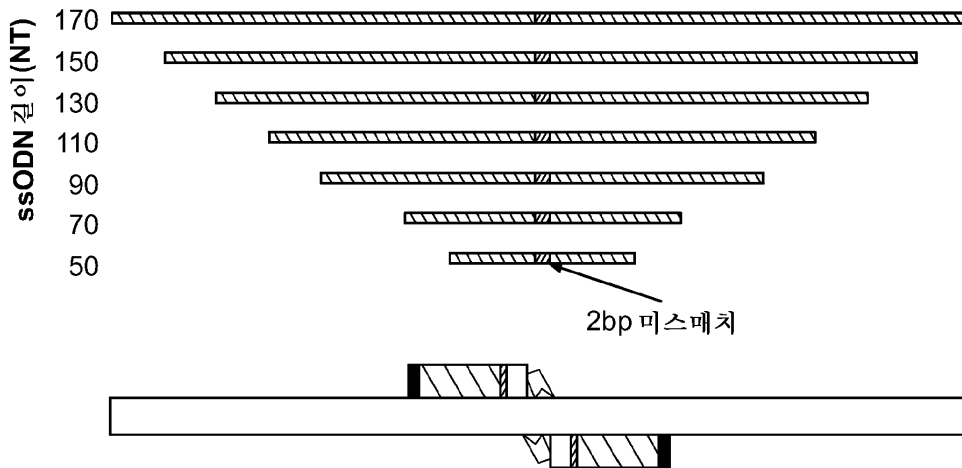
도면 2b



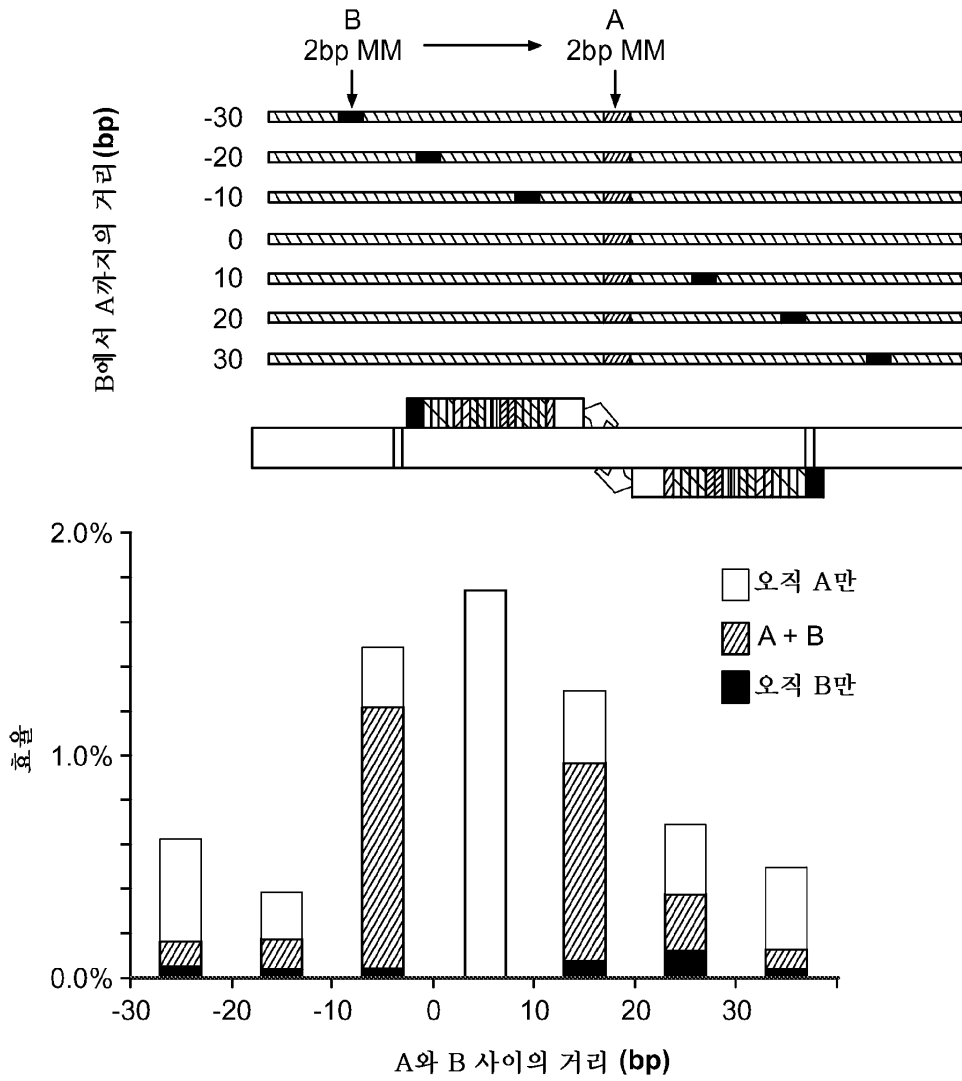
도면2c



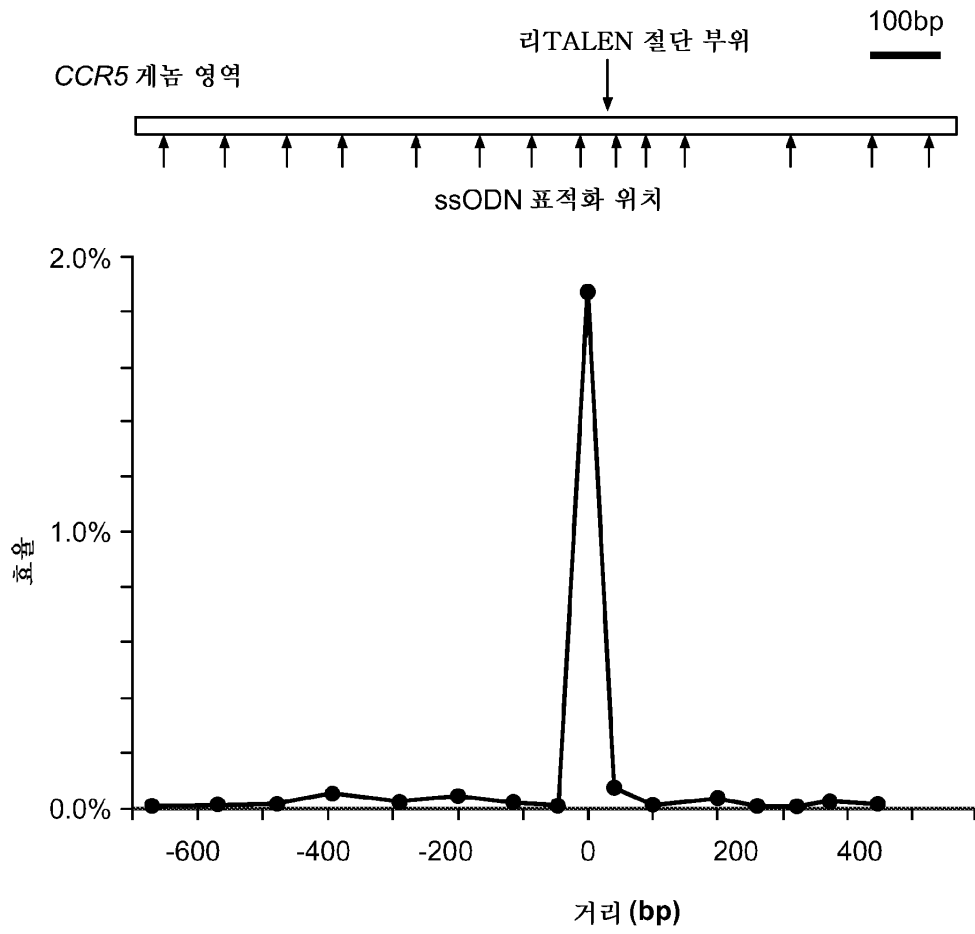
도면3a



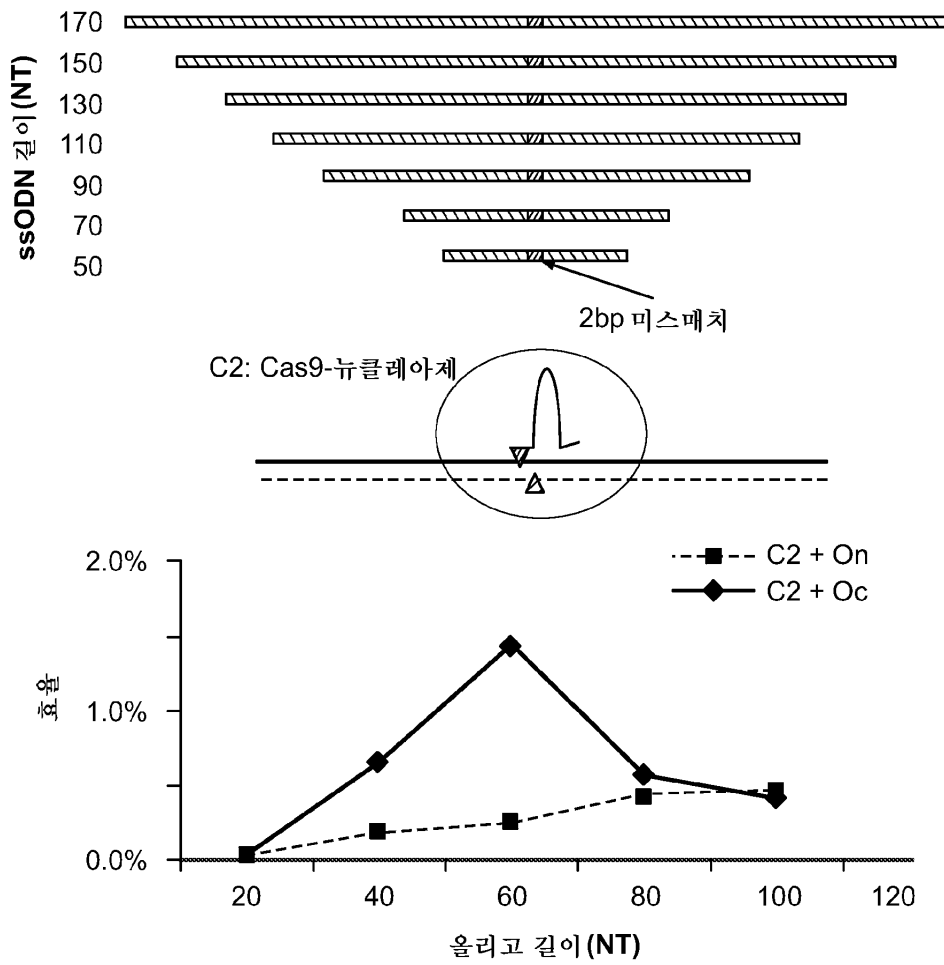
도면3b



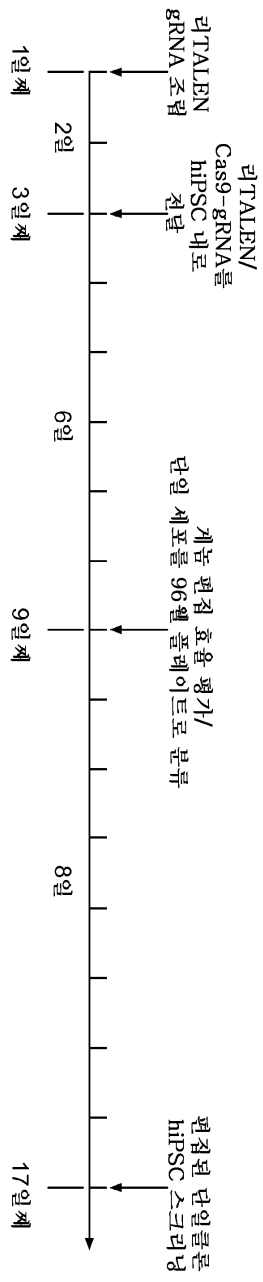
도면3c



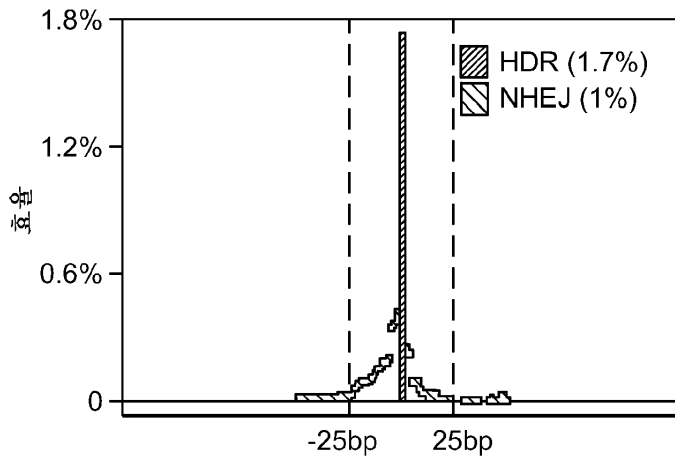
도면3d



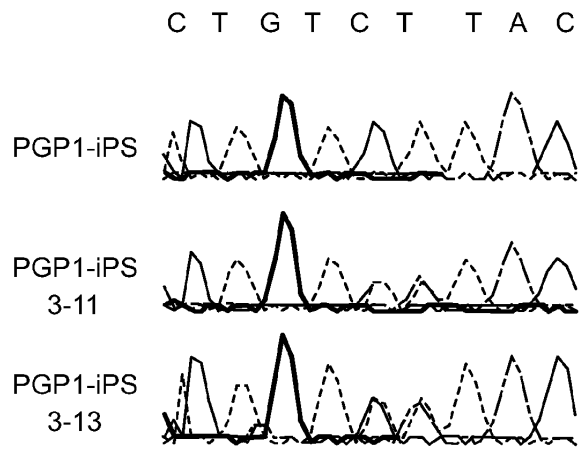
도면4a



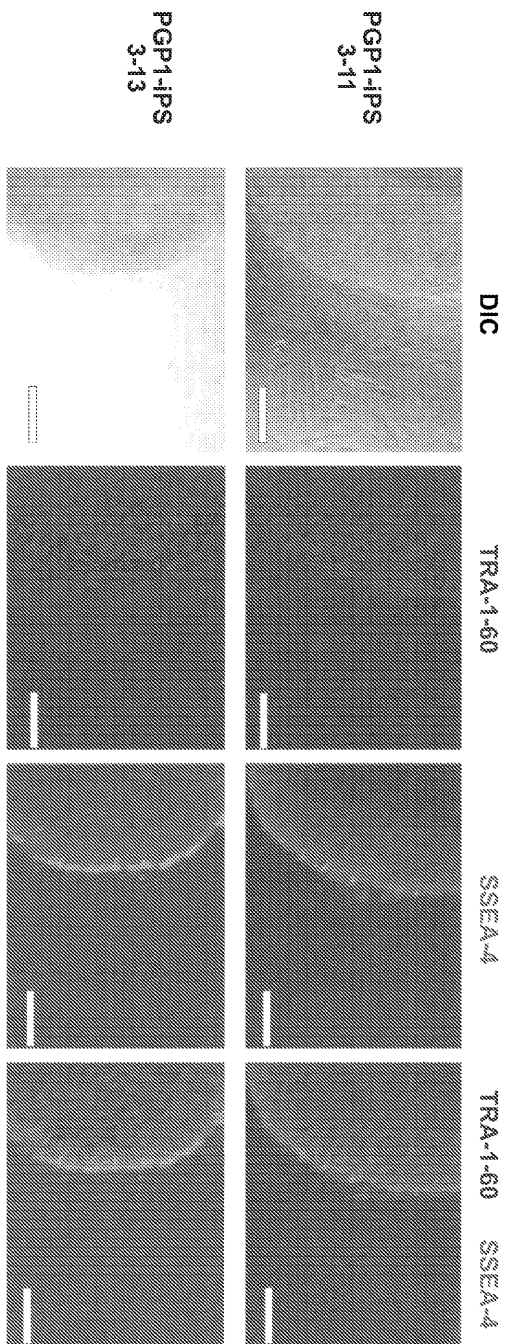
도면4b



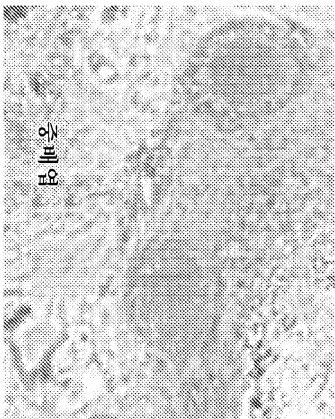
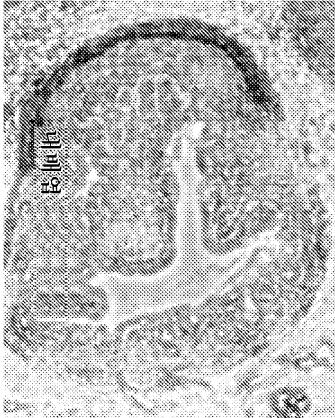
도면4c



도면4d



도면4e



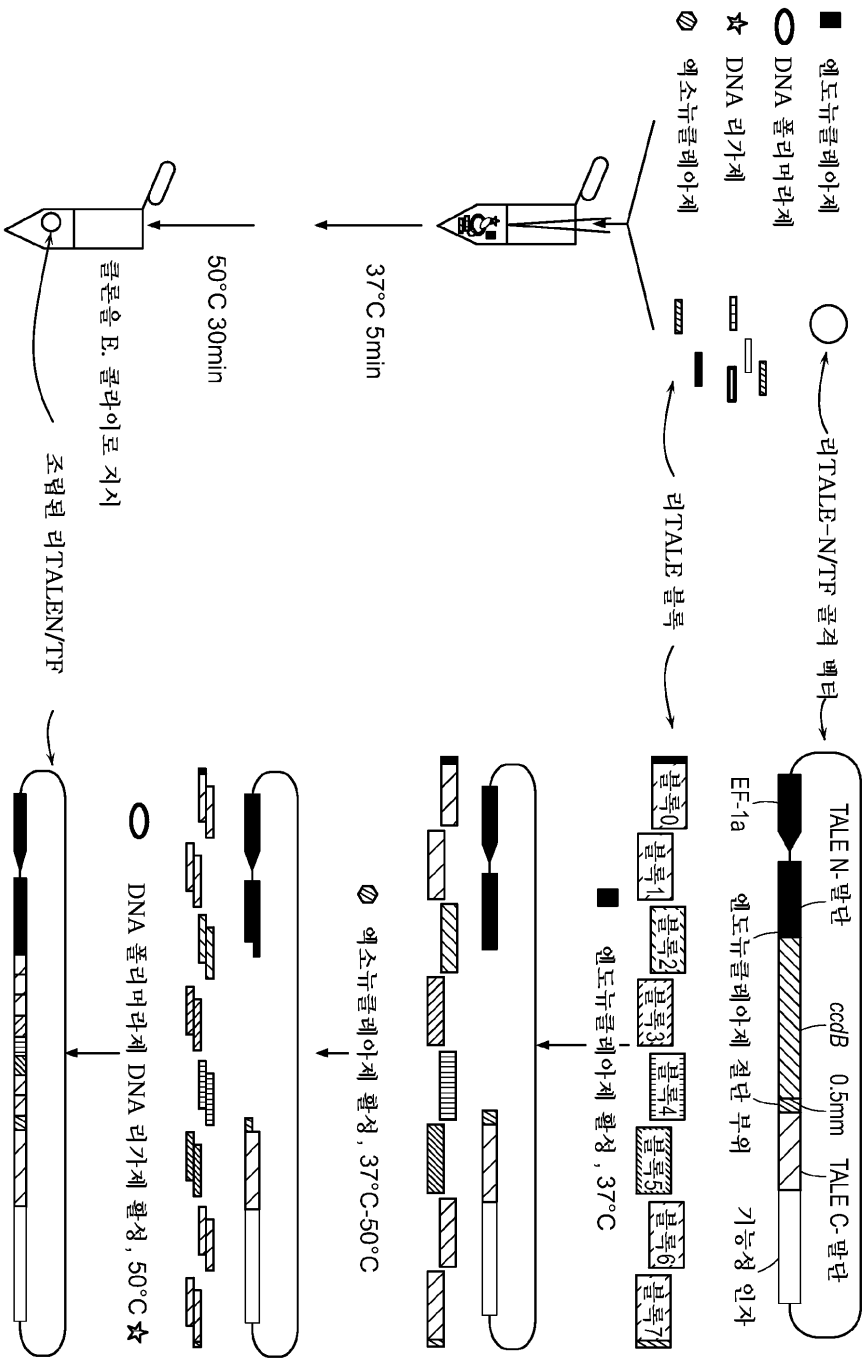
L T P E Q V V A I A S X X G G K Q A L E T V Q R L L P V L C Q A H G
 10 20 30 40 50 60 70 80 90 100
 TALE-MM CTACCCGAGAGGTTGTCGTAATCGCCTCCNNNNNNGGGGAAACAGGCAATTGGAAACAATAAGCGGCTGTCGGGGTCTGTGCCAAGCGCAGCGA
 리TALE-M1 CTAACCCCTGAACAGGTAGTGGCTATAGCTTCANNNNNNGGGGCAAGCAAGCACTTGAACCGCTTCAACGACTCCTGCCAAGTCTGTGCCAAGCCCATGGG
 리TALE-M2 TTGACTCCGGAGCAAGTCGTGCGGATCGCGAGCANNNNNGGGGGAAAGCAGGCGCTGGAACCTGTCAGAGACTGCTGCTGACTTTTGTCAAGGCGCATGGT
 리TALE-M3 CTCACCCCGAACAAGTTGTCCGAAATAGCAAGTNNNNNNNGGGGGTAAACAAGCTCTAAGAACGGTTCAAAGGTTGTGCCCGTTCTGTGCCAAGCATGGG
 리TALE-M4 CTGACACCTGAACAAGTTGTCCGGAATAGCCAGTNNNNNNNGGGGGTAAACAAGCTCTAAGAACGGTTCAAAGGTTGTGCCCGTTCTGTGCCAAGCATGGG
 리TALE-M5 TTAACACCCGAAACAAGTAGTAGGATAGCGTGANNNNNNNGGGGTAACAAGGCTTTGGAAGCGGTACAGCGGTTATTGCCGGTCTGTGCCAAGCGCAGGAA
 리TALE-M6 CTTAGCGCAACAAGGTGGTTGCAATTCCTCCNNNNNGGGGGAAACAAGCGCTTGAACCTGTGAGAGACTCCTTCTGTTTGTGTCAAGCCCAAGCGG
 리TALE-M7 TTGACGCTGAGCAAGTTGTGCCATCGCTAAGCANNNNNGGAGGAAAGCAGGCTTGAACCGTACAGCACTTCCCAATTTGTGCCAAGCTCAAGGG
 리TALE-M8 CTAACCCCGAAGTAGTTGCCATAGCAAGCANNNNNNNGGAGGAAACAAGGCATTAAGAACGTTCAAGCGCTTGTCCCGTACTGTCAAGGCAAGCGG
 리TALE-M9 CTAACCTCCGGAACAAGTCGTAGCCATTCCTCCNNNNNGGGGCAAAACAAGGCGCTAAGACAGTCCAGAGGCTCTTGCCCTGTTAAGCCAGCATGGC
 리TALE-M10 CTCACCCCGGAGCAAGTCGTTGCCATCGCCAGTNNNNNGGGGAAAGCAAGCTCTGAACAATAACAGGCTGTGCCAAGCGGCTGTAAGCTCATGGA
 리TALE-M11 CTGACGCCGAGGAGTAGTGGCAATCGCACTNNNNNNNGGAGGTAAACAAGCACTCGAGACTGTCCAAAATTTTACCCTACTAATGCCAAGCGCATGGT
 리TALE-M12 TTAACCCCGAAGTAGTTGGCTATTGGACTNNNNNNNGGTGGCAACAAGCCCTTGGAGACAGTGAACGATTTACTGCCCTGTCTTAATGTCAAGGCCATGGC
 리TALE-M13 CTTACTCTGAGCAAGTCGTACTATCGCCAGCANNNNNGGTGGGAAACAAGGCCCTGGAACCCGTACAACGCTCTCCTCCAGTACTTTGTCAAGCAAGCGG
 리TALE-M14 TTGACACCCGGAACAAGTGGTGGCATTCGCTCCNNNNNGAGGCAAGCAAGCACTGGAACCGCTCCAGCGCTTCTTCCGGTTCTTTGCCAGGCTCATGGG
 리TALE-M15 CTCAGGCCAAGAGGTTGTTAGCAATAGCGCTCGNNNNNGGTGGTAAAGCAAGCGCTTGAACCGGTTCAAGCGCTTCTGCGGGTGTGCCAGGCGCACGGA
 리TALE-M16 CTCACACCAACAAGTGGTTGCTATTGCTAGTNNNNNGGTGGAAAGCAGGCGCTCGAAGCGGTGCAAGAGTTACTTCCCGTCTCTGTCAAGCGCACGCG
 리TALE-M17 CTCACCTCAGAGCAAGTGGTGGCATCGCTTCANNNNNNGGTGGAAAGCCTCGCCCTGGAA

도면5a

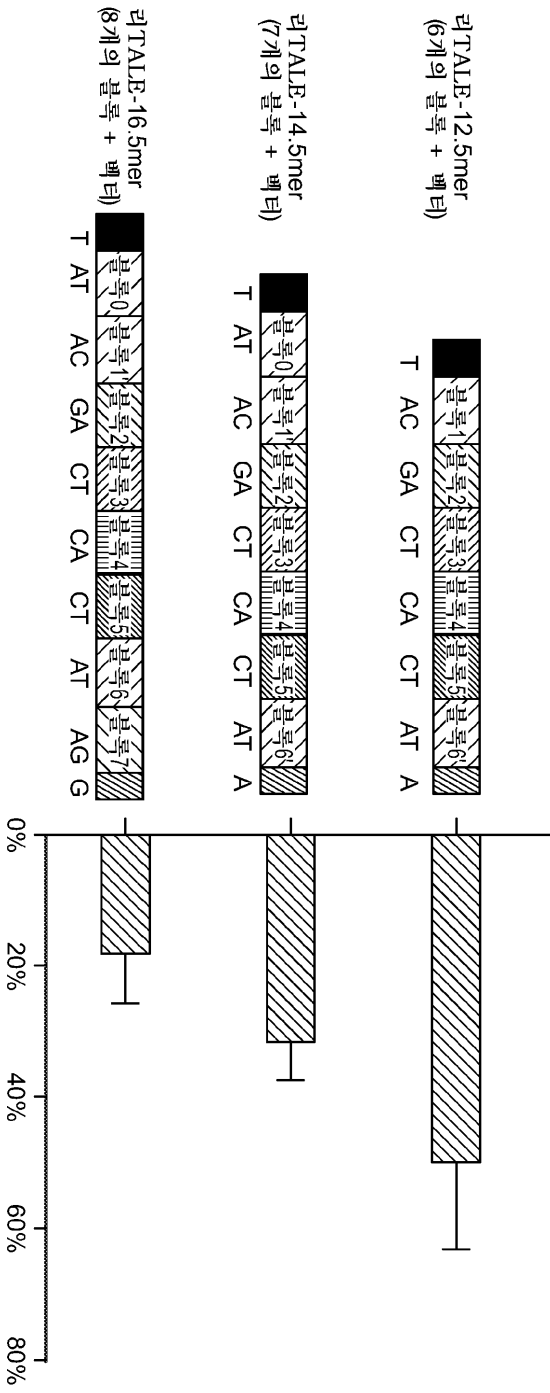
도면6a



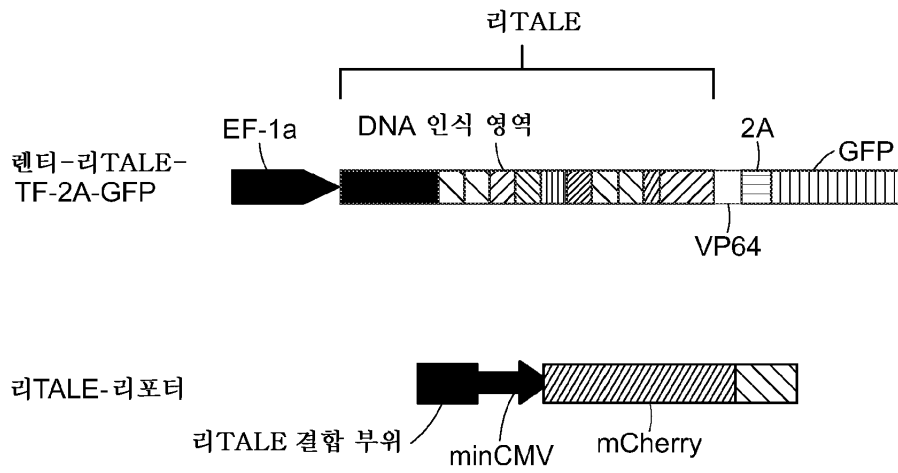
도면6b



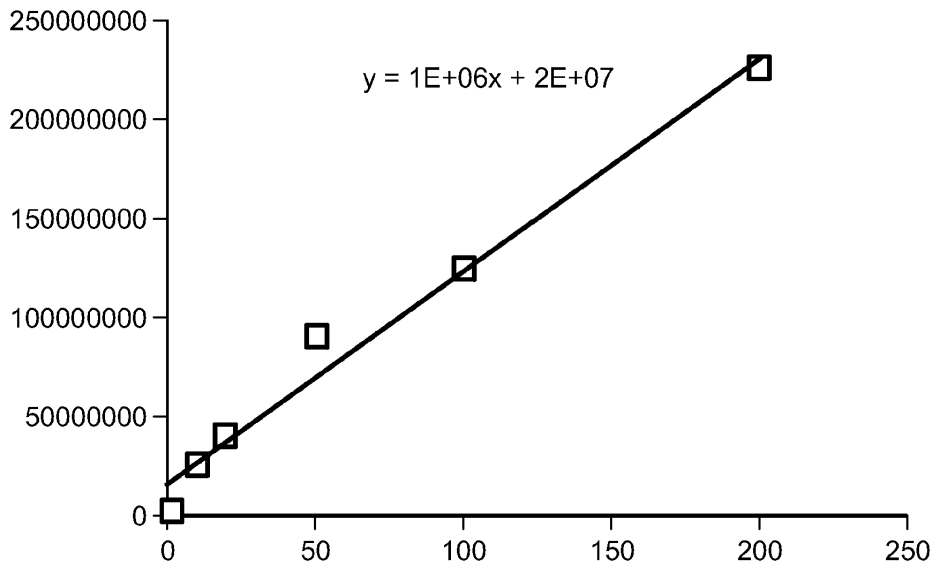
도면6c



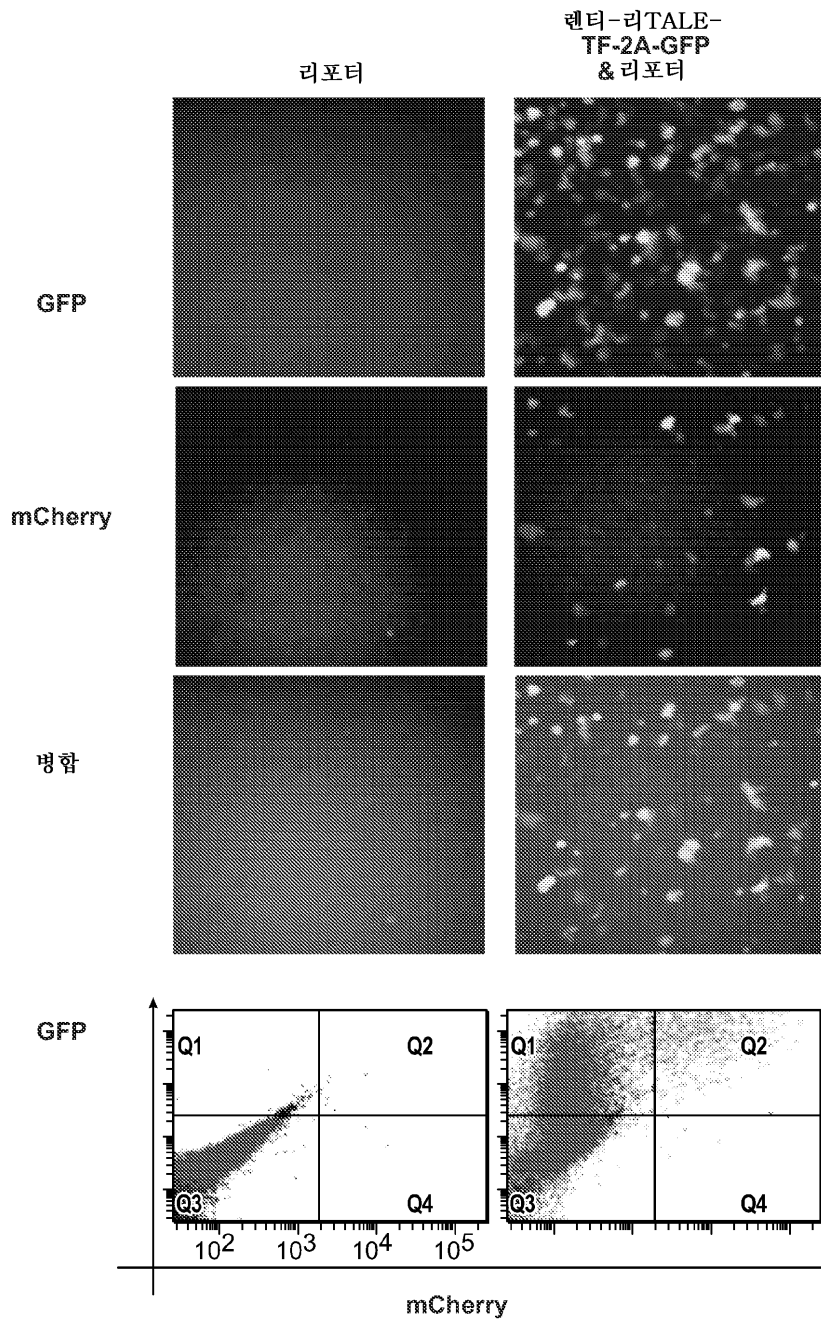
도면7a



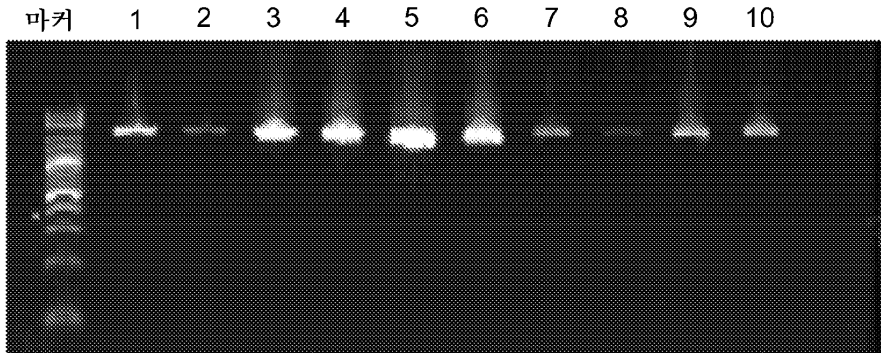
도면7b



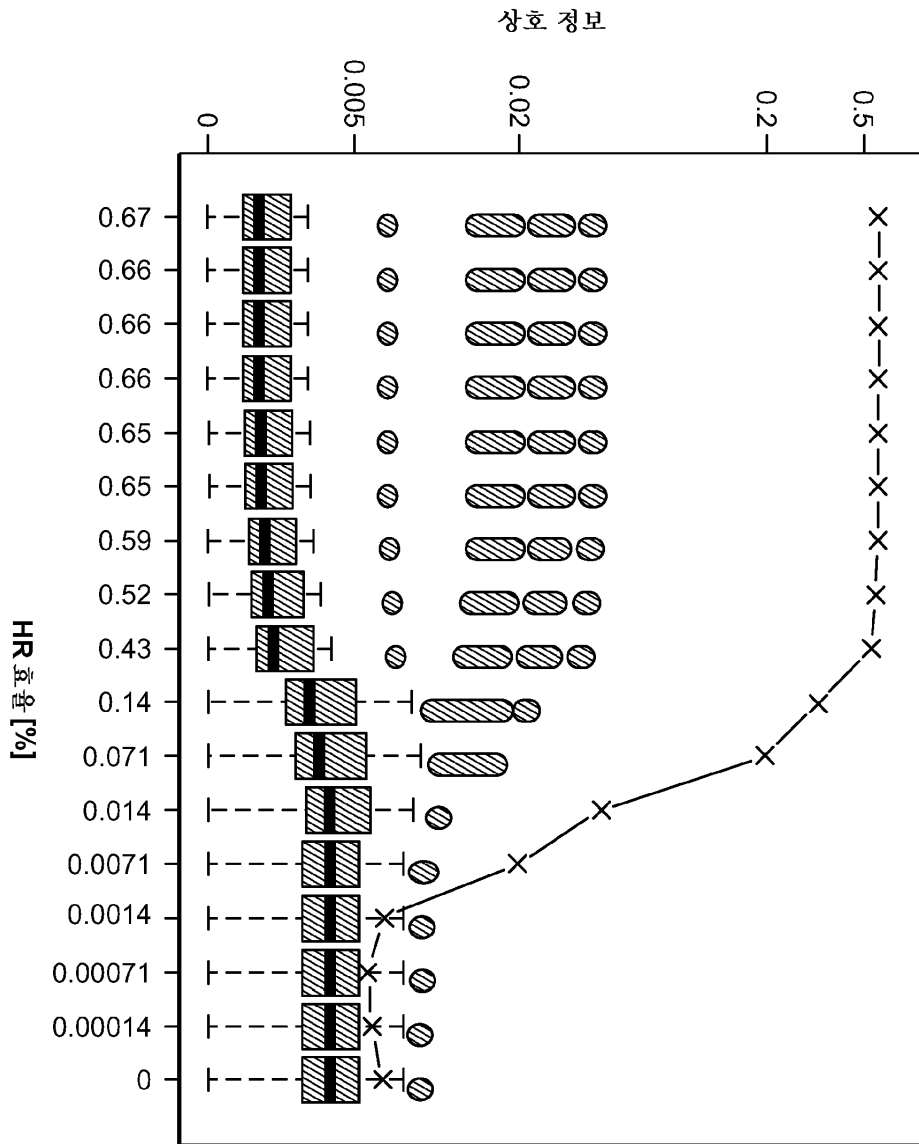
도면7c



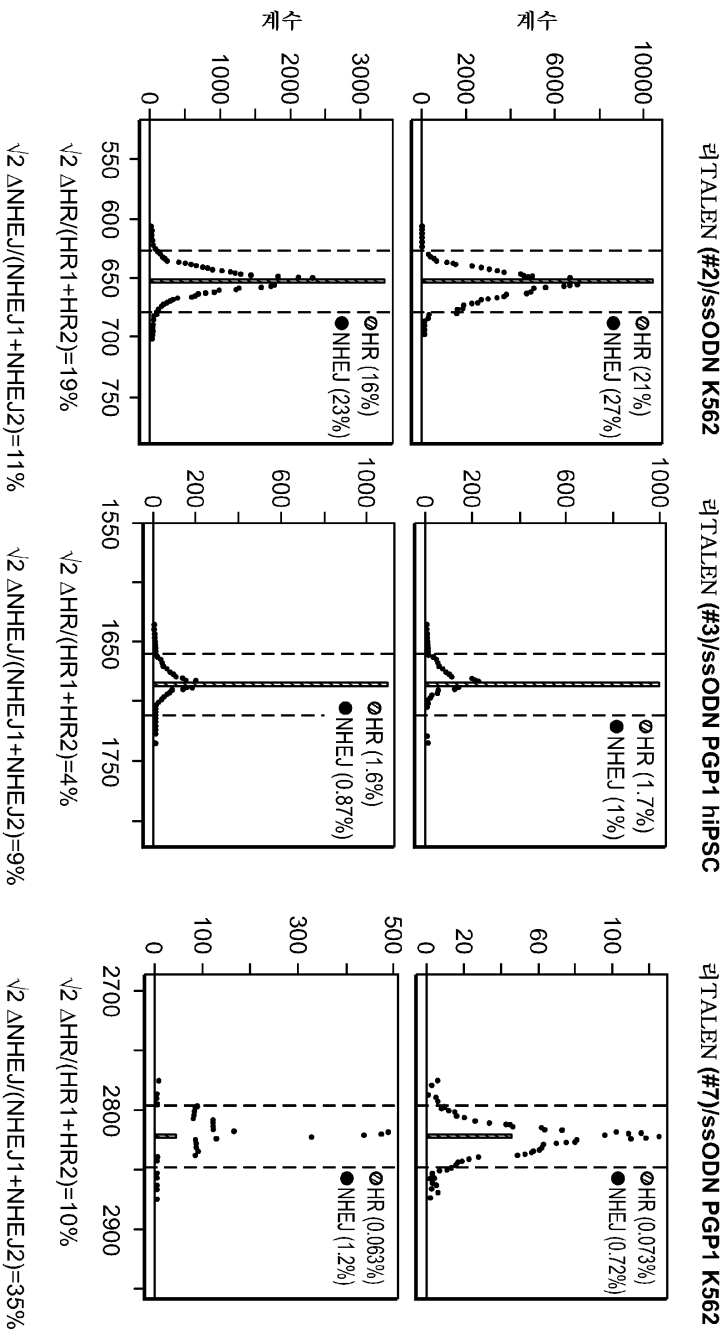
도면7d



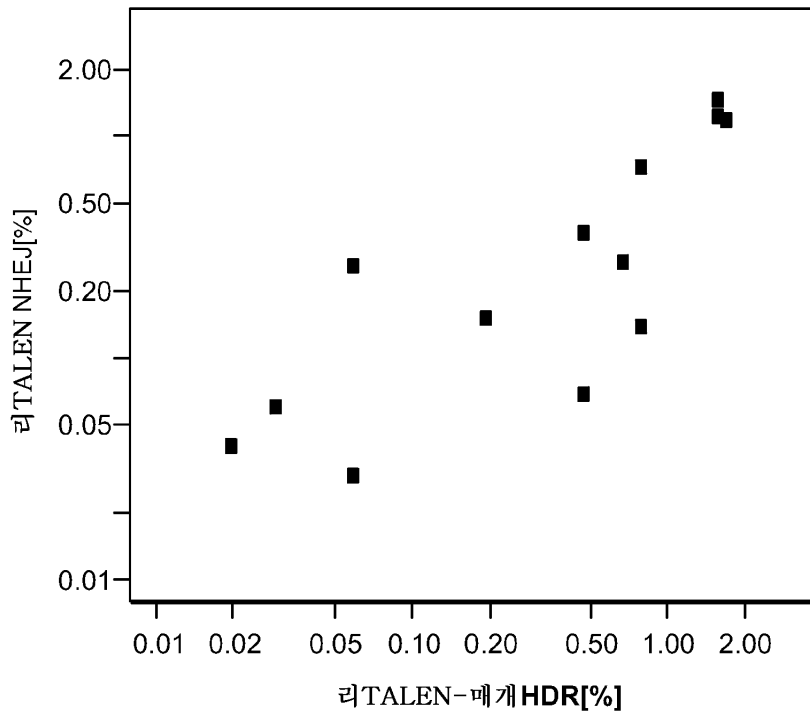
도면8a



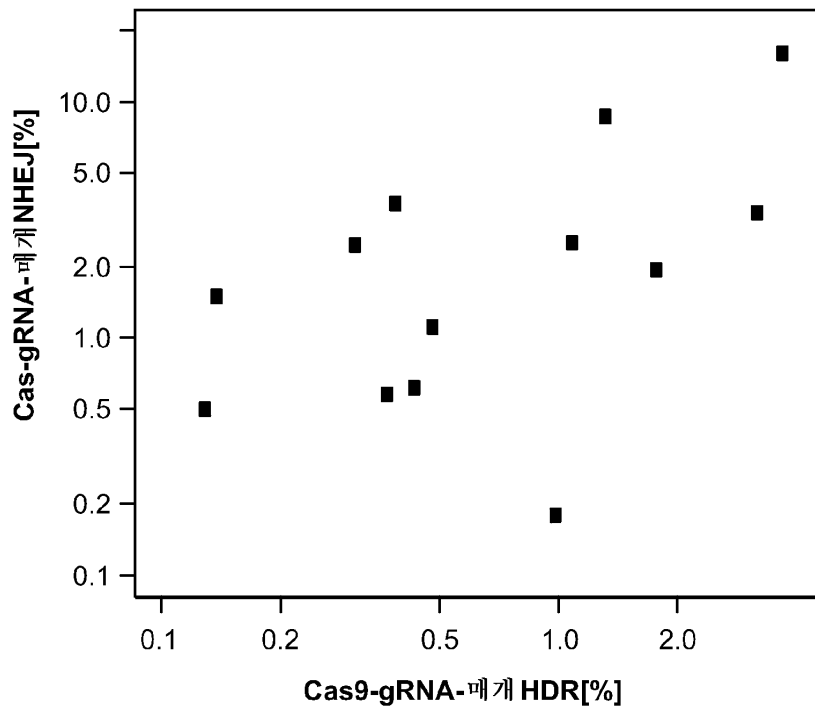
도면 8b



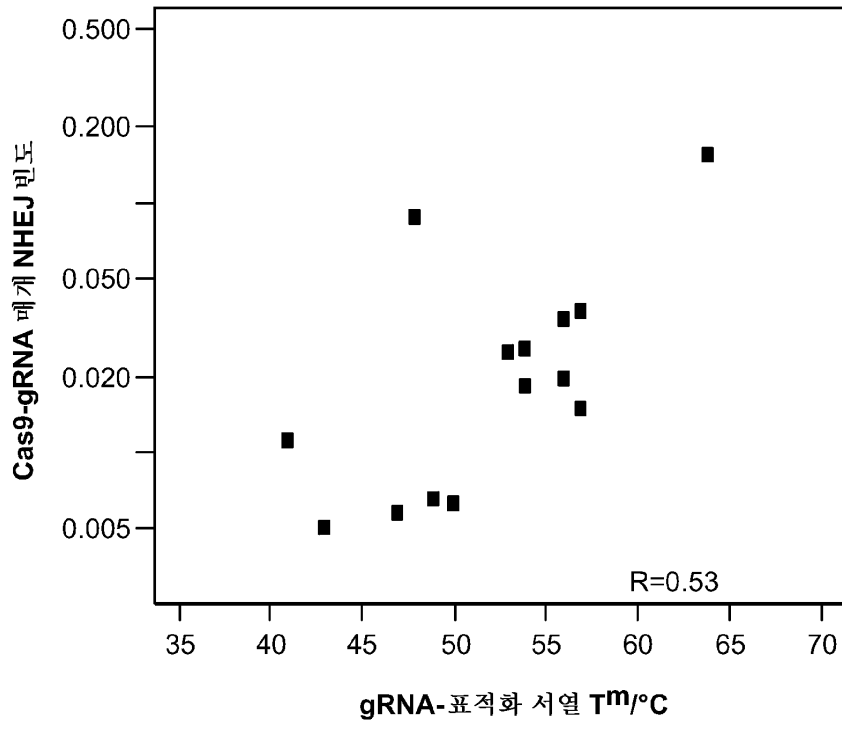
도면9a



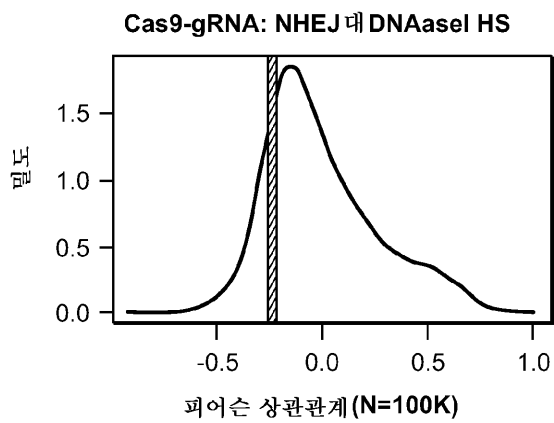
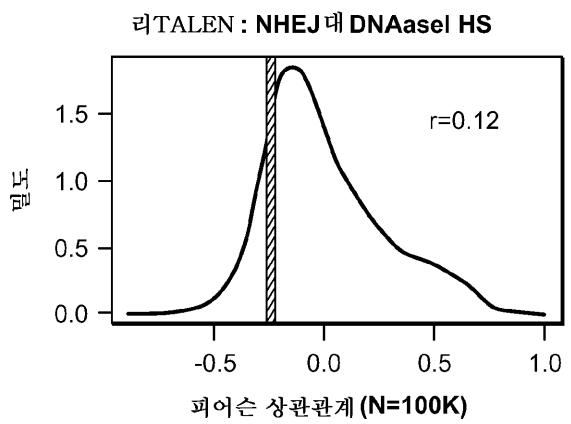
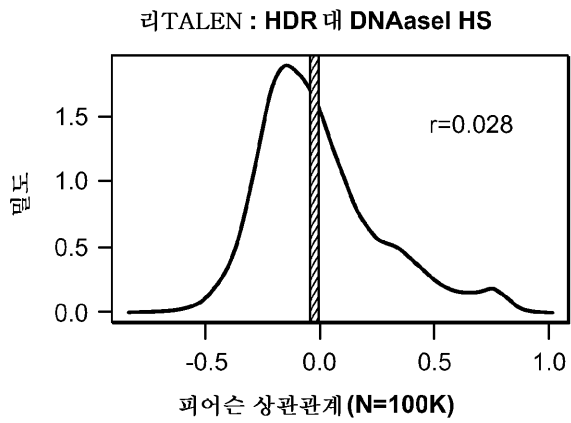
도면9b



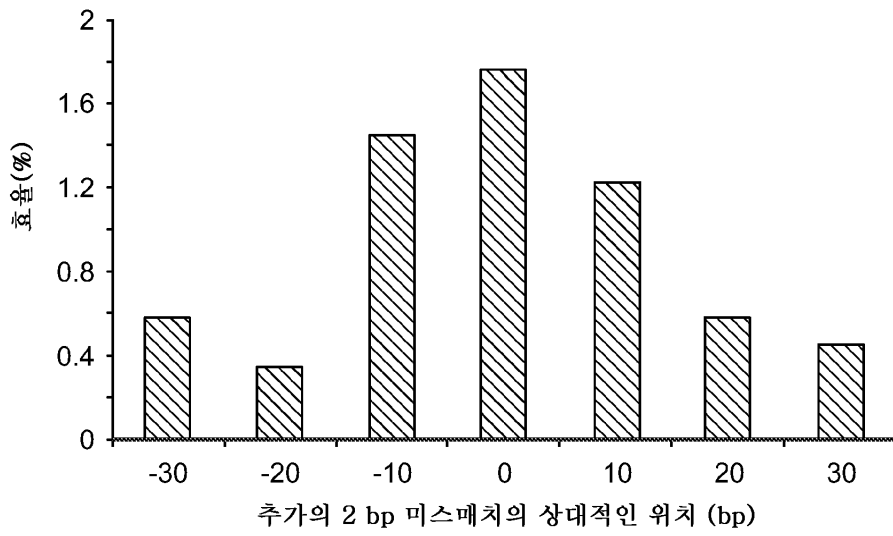
도면9c



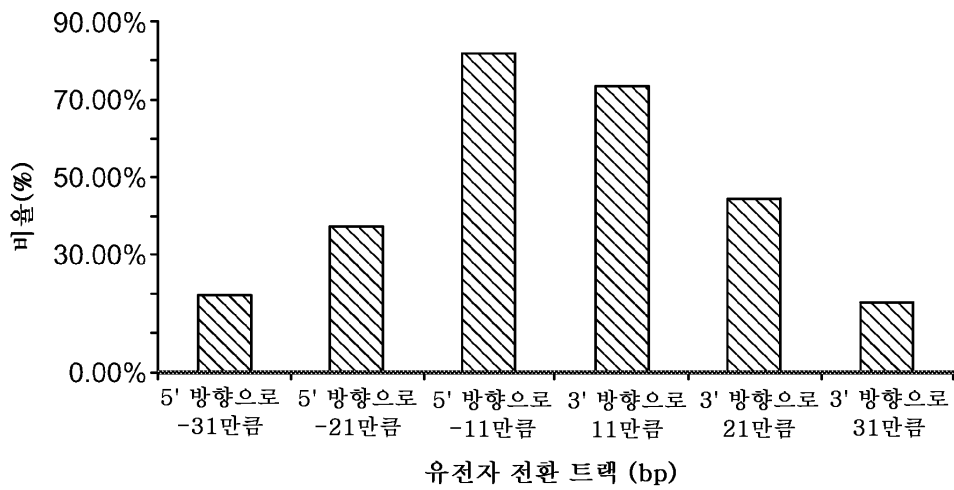
도면10



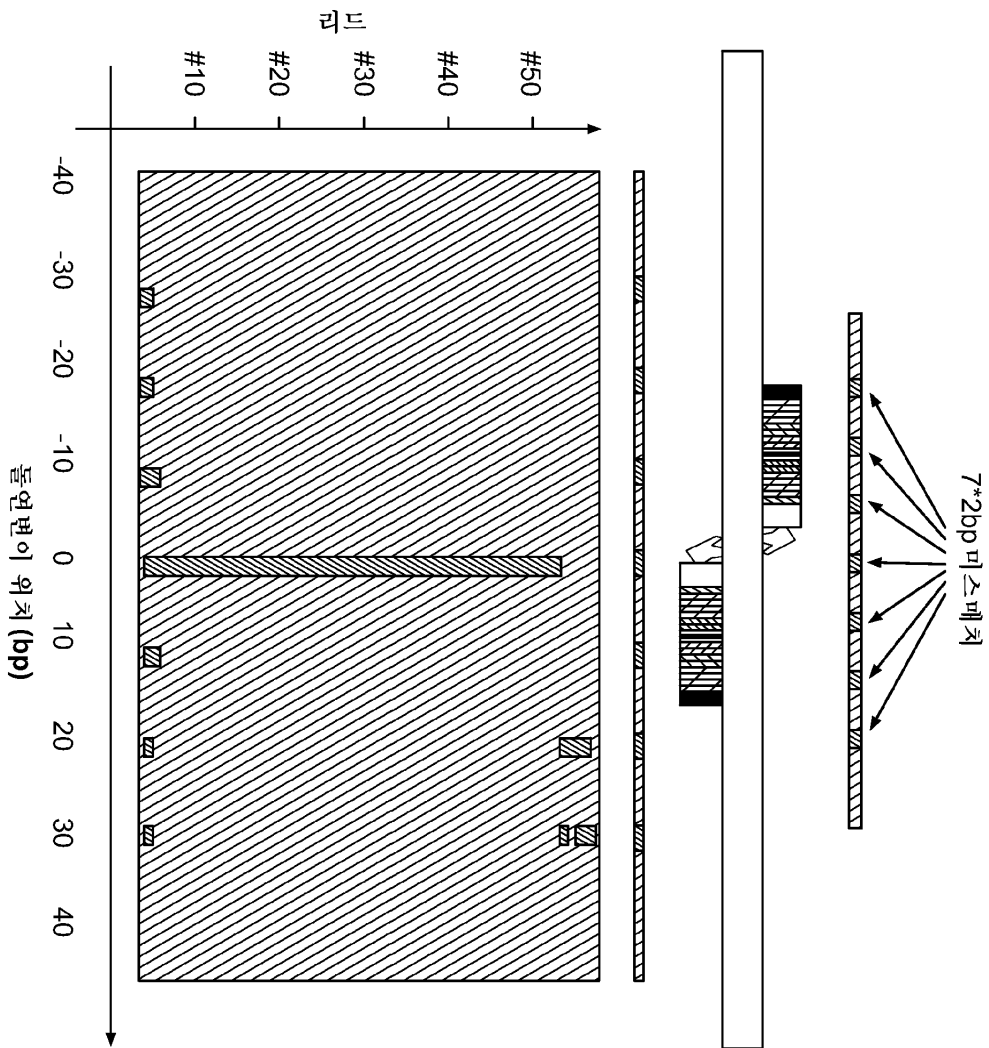
도면11a



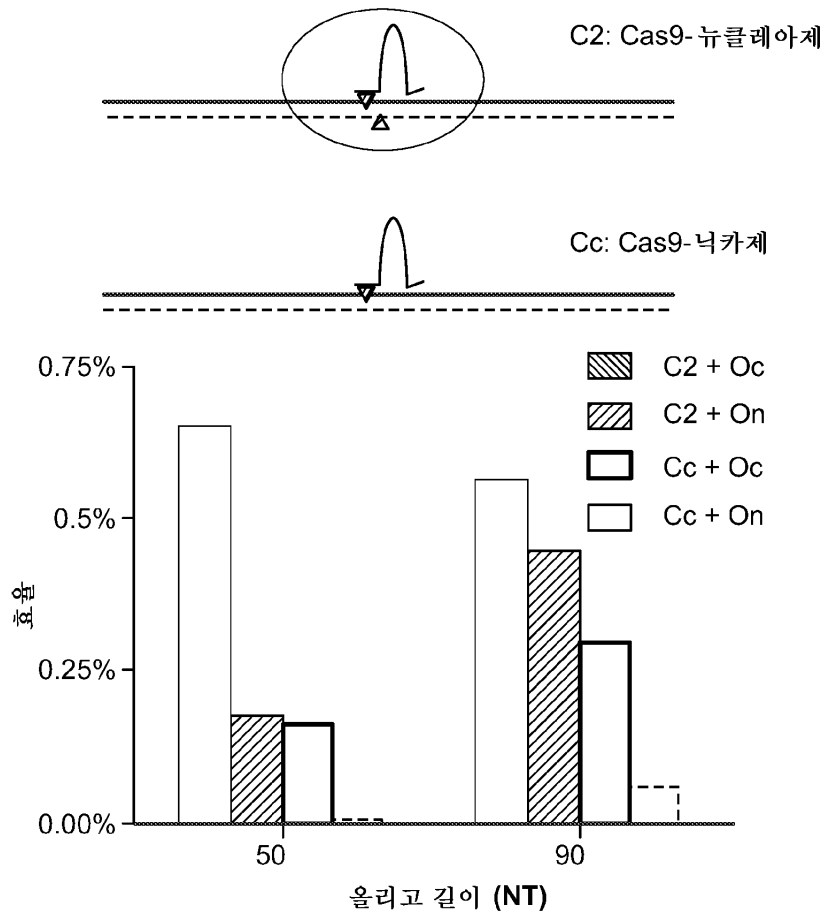
도면11b



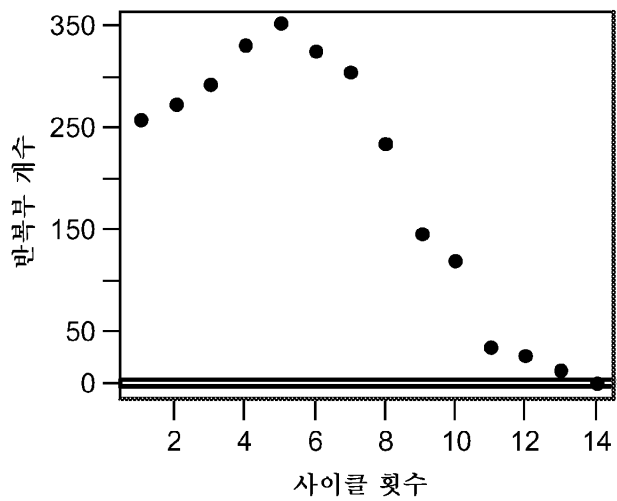
도면11c



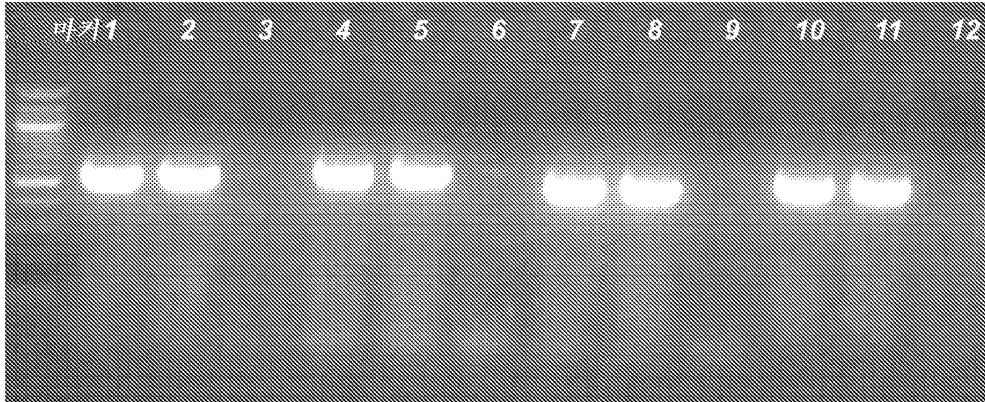
도면12



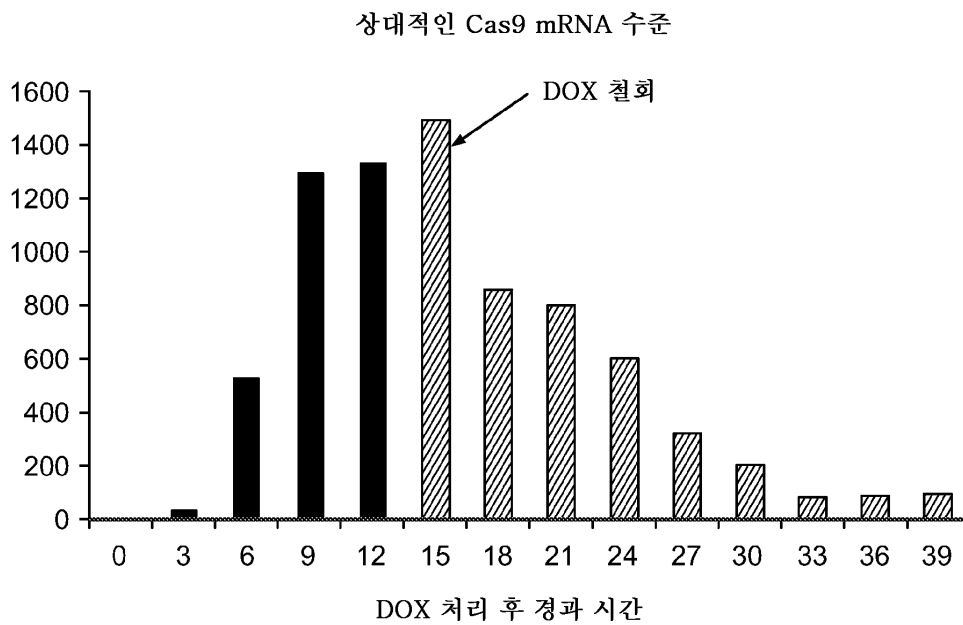
도면13



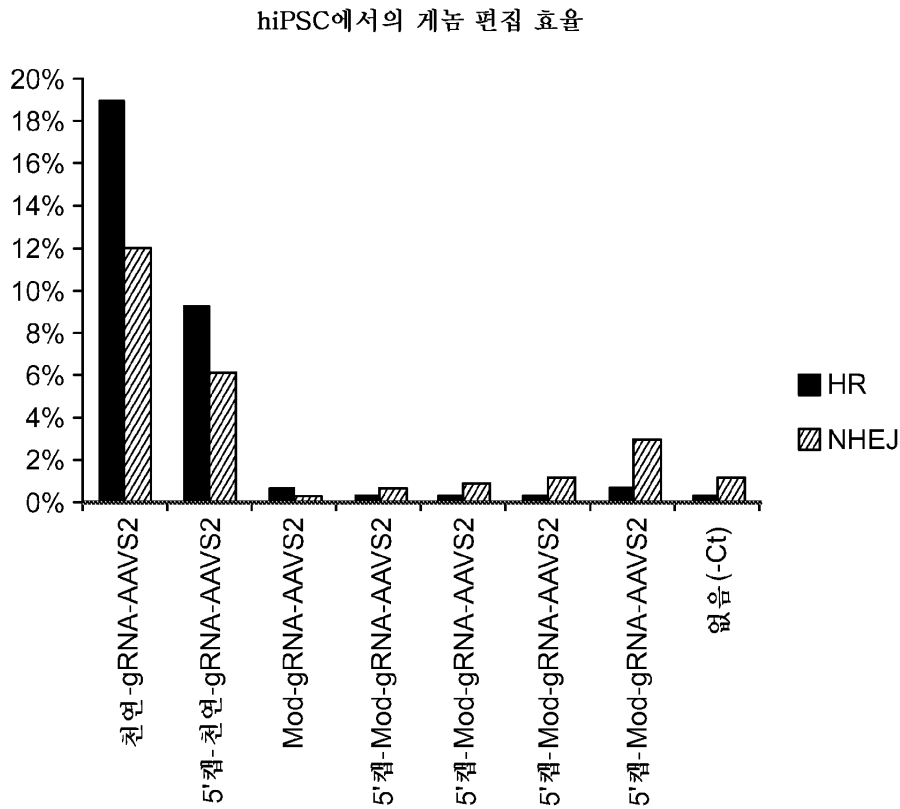
도면14



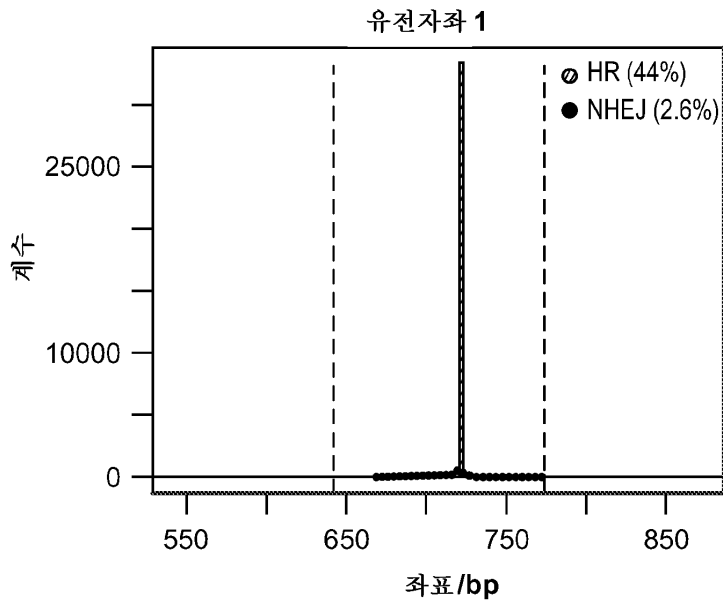
도면15



도면16

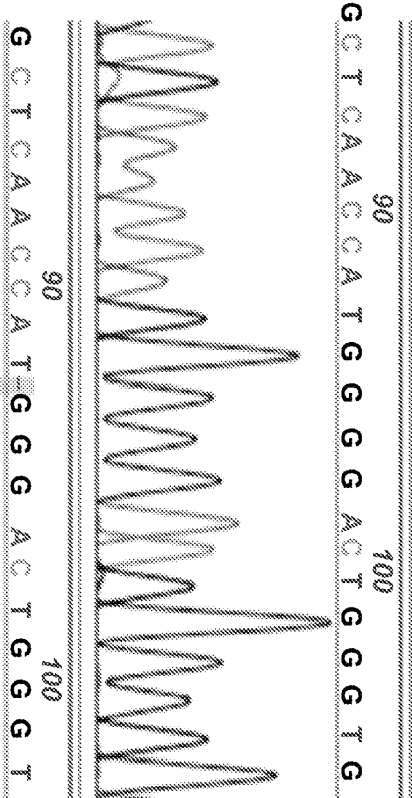


도면17

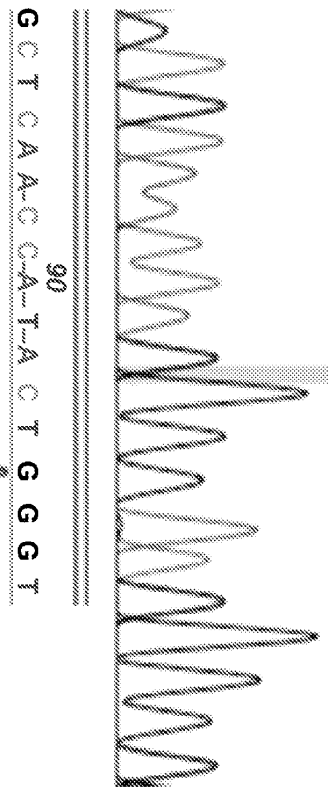


도면18

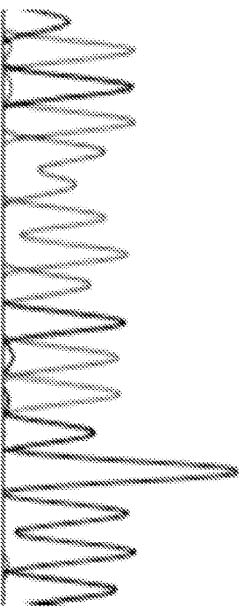
PGP1-IPS-NT



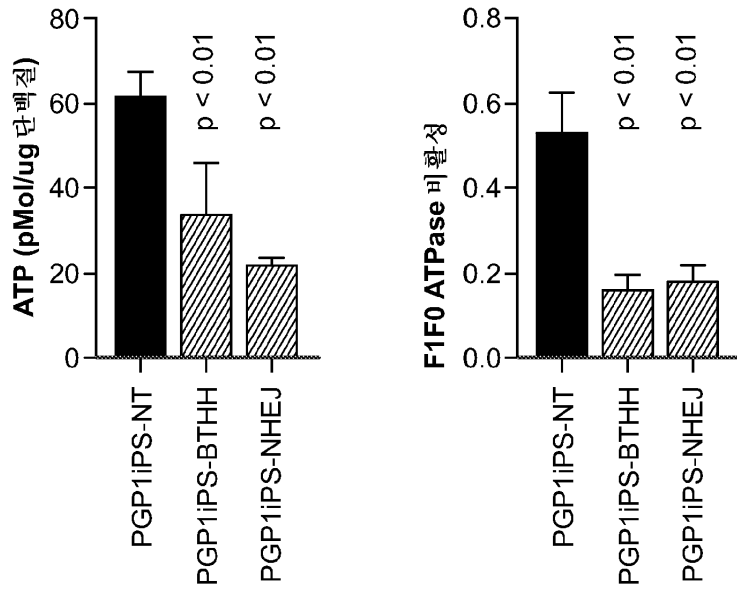
PGP1-IPS-BTHH



PGP1-IPS-NHEJ



도면19



리TALLEN-클릭 서열

(보라색: 리TALBE-N; 적색: SapI 부위; 녹색: 0.5 단량체, 청색: 리TALBE-C; 옐로우색: Fok I)

ATGTCGGGACCCGGCTCCCTTCCCAACCCGGCAACCAGCCGCTTTTCGGCCGAC
 TCGTTCACAGACCTGCTTAGGCAGTTCGACCCCTCAGTGTTAACACATCGTTGTTCCG
 ACTCCCTTCCCTCCGTTTGGGGCCACCATACGGAGCGGCCACCGGGAGTGGAT
 GAGTGCAGTCCGGATTGAGAGCTGCCGATGCACCAACCCCAACCATGCCGGTGGC
 CGTCAACCGCTGCCCGACCCCGAGGGCCGAAGCCCGCACCAAGCGGAGGGCAGCCG
 AACCGTCCGACCGCAAGCCCGCAGCGCAAGTAGATTGAGAACTTTGGGATATTCA
 CAGCAGCAGCAGGAAGAATCAAGCCCAAGTAGGTCGACAGTCCGCCAGCATC
 ACGAAGCGCTGGTGGGTCAATGGGTTTACACATGCCACATCGTAGCCCTTGTCCGACG
 ACCCTGCAGCCCTTGGCAGCGGTCGCCGTCAGTAGCAGGACATGATGCCGGCTTGC
 CGGAAGCCACACATGAGGGCGATCGTCCGTTGGGAAACAGTGGAGCGGACCCCG
 AGCGCTTGAGGCCCTGTGACGGTCCCGGGAGAGCTGAGAGGGCTCCCTTCAGC
 TGGACACGGGCCAGTTGCTGAAGATCCGGAAGCGGGAGGAGTCAACGGCTCGAG
 GCGGTGCACCGCTGGCCCAATGCCCTCACGGGAGCACCCCTCAACAGTTTCAAGTTG
 ACAGAGACCCGGCCGCATTAAGGCACCCCAAGCCTTTACACTTATGCTTCCGGCTCG
 TATAATGTGTGATTTTGAAGTTAGGATCCGTCGAGATTTTCAGGAGCTAAGGAAGCT
 AAAATGGAGAAAAAAATCACTGATATACCACCGCTTGATATATCCCAATGGCATCG
 TAAAAGAACATTTTGAAGCATTTAGGCAATTCAGTTGCTCAATGTACCTATAACAGACCGT
 TCAGCTGGATATTACGGCCTTTTAAAGACCGTAAAGAAAAATAAGCACAAAGTTTAA
 TCCGGCCTTTATTACACATTTCTTGCCCGCCCTGATGAATGCTCATCCGGAATTCGGTATG
 GCAATGAAAAGACGGTGAAGCTGGTGAATATGGGATAGTGTTCACCCCTTGTTACACCCGTT

도면 20a

TTCATGAGCAAACCTGA AACGTTTTCATCGCTCTGGAGTGAATACCACGCGATTTCC
 CGGCAGTTTCTACACATATAATTCGCAAGATGTGGCGTGTACGGTGAA AACCTGGCC
 TATTTCCCTAAGGGTTTATTTGAGAAATATGTTTTTTCGTCACGCCAATCCCTGGGTGA
 GTTTCACCAGTTTTGATTTAAACGTGGCCAATATGGACAACCTTCCGCCCCCGTTTTT
 CACCATGGCAAATATTTATACGCAAGGCGACAAGGTGCTGATGCCGCTGGCGATTC
 AGGTTCAATGCCCCGTTGTGATGGCTTCCATGTCCGCAGAAATGCTTAATGAATTTAC
 AACAGTACTGCCATGAGTGGCAGGGCGGCGTAAAGATCTGGATCCGGCTTACTTA
 AAAGCCAGATTAACAGTATGCCGTTATTTGCCGCTGATTTTTGCGGTATAAGAAATATAT
 ACTGATATGTAATACCCGAAAGTATGTCAAAAAAGAGGTATGCTATGAAAGCAGCGTATTT
 ACAGTGACAGTTTGACAGCGACAGCTATCAGTTGCTCAAGGCATATATGATGTCAATA
 TCTCCGGTCTGTTAAGCACAAACCATGCAGAAATGAAAGCCCGTCTGCTGCCGTGCCGAAC
 GCTGGAAGCGGAAATCAGGAAGGATGGCTGAGGTCGCCCGTTTATTGAAATG
 AACGGCTCTTTTGGCTGACGAGAACAGGGGCTGGTGAATGCAGTTTAAGGTTTACAC
 CTATAAAGAGAGAGCCGTTATTCGCTGTGTTTGTGGAATGTACAGAGTATATTAATTGA
 CACGCCCGGGCGACGGATGGTATCCCGTGGCCAGTGCACGTCCTGCTGTCAGATA
 AAGTCTCCCGTGAACCTTTACCCGGTGGTGCATATCGGGATGAAAGCTGGCCGATGA
 TGACCAACCGATATGGCCAGTGTGCCGCTCTCCGTTATCGGGGAAGAGTGGCTGATC
 TCAGCCACCCGGAATAATGACATCAAAAAACGCCATTAAACCTGATGTTCTGGGGAATA
 TAAATGTCAGGCTCCCTTATACACAGCCAGTCTGCAAGGTCGACGGTCTCC**GCTCTTTCG**
 AAGTTACTTCCCGTCTCTGTCAAAGCCAGCGCCCTCACCTCCAGAGCAAAGTGGTTGC
 GATCGCTTCAAACAACGGGTGGAAGACCTGCCCTGGAATCAATCGTGGCCCA**GCTTTC**

도면20b

GAGGCCGGACCCCCGCCTGGCCGCACTCACTAATGATCATCTTGTAGCCCTGGCCTG
CCTCGCGGACGACCCCGCCTTGGATGCGGTGAAGAGGGGCTCCCGACCGCCCTG
CATTGATTTAAGCCGGACCAACAGAAAGATTCCCGAGAGGACATCATCGAGTGGCA
GGTCCCAACTCGTGAAGAGTGAACCTTGAAGAAATAAGTCGAGCTGCCGCAACA
ATTGAAATACGTACCGCATGAATACATCGAACTTATCGAAATTGCTAGGAACTCGAC
TCAAGACAGAAATCCTTGAGATGAAGGTAATGAGTTCTTTATGAAGTTTATGGATA
CCGAGGGAAGCATCTCGGTGGATCACGAAACCCGACGGAGCAATCTATACGGTGG
GAGCCCGATTGATTACGGAGTGTCTCGACACGAAAGCCTACAGCGGTGGGTAC
AATCTTCCCATCGGCGAGCAGATGAGATGCAACGTTATGTGGAAGAAATCAGAC
CAGGAACAACAACATCAATCCAATGAGTGGTGGAAAGTGTATCTTCATCAGTGA
CCGAGTTTAAAGTTTGTCTCTCTGGCATTTCAAAAGCACTATAAGGCCCAAGCT
CACACGGTTGAATCACATTAACGAACTGCAATGGTGGGTTTGTCCGTAGAGGAACCT
GCTCATTTGGTGAGAAATGATCAAAAGCCGGAACCTCTGACACTGGAAGAAGTCAGAC
GCAAGTTTAAACAATGGCGAGATCAATTTCCGC

리TALB-TF 클럭 서열

(보라색: 리TALB-N; 적색: SapI 부위; 녹색: 0.5 단량체, 청색: 리TALB-C; 오렌지색: NLS-VP64; 2x-GFP는 녹색으로 강조 표시되어 있다)

ATGTCCGCGACCCCGGCTCCCTTCCACCACCCGCAACCAGCCAGCGTTTTCGGCCGAC
TCGTTCTCAGACCCTGCTTAGGCAGTTTCGACCCCTCACTGTTTAAACAACATCGTTGTTCCG

ACTCCCTTCCCTCGTTTGGGGCCCACTACGGAGCGGGCCACCAGGAGTGGGAT
 GAGGTGCAGTCGGGATTTGAGAGCTGCCGATGCACCACCACCACCATGCCGGTGGC
 CGTCAACCGCTGCCCGACCCTGGAGGGCGAAGCCCGCACCAAGCGGAGGGCAGCCG
 AACCGTCCGACGCAAGCCCCCGCAGCGCAAGTAGATTTGAGAAGCTTTGGGATATTCA
 CAGCAGCAGCAGGAAAAGATCAAGCCCAAAGTGAGGTCGACAGTCGCCGACGATC
 ACGAAGCGCTGGTGGGTCTGGGTTCACACATGCCACATCGTAGCCTTGTCCGAGC
 ACCCTGCAGCCCTTTGGCACCGTCCGCCGTCAAGTACCAGACATGATTTGCCGGTTGC
 CGGAAGCCACACATGAGGCCGATCGTCGGTGTGGGAAACAGTGGAGCGGAGCCCG
 AGCGCTTGAAGCCCTGTGACCGGTCGCGGGAGAGCTGAGAGGCCCTCCCTTCAGC
 TGGACACGGGGCCAGTTGCTGAAGATCCCGAAGCGGGGAGGAGTCAAGCGGTCGAG
 GCGGTGCACGCGTGGCGCAATGCCCTCAGCGGAGCACCCCTCAACAGTTCAAGCTG
 ACAGAGACCGCGCCGATTAAGCAACCCAGGCTTTACACTTTATGCTTCCGGCTCG
 TATAATGTGTGATTTTGAAGTTAGGATCCGTCGAGATTTTCAGGAGCTTAAGGAAGCT
 AAAAAATGGAAAAAAAATCACTGGATATACCACCCGTTGATATATCCCAATGGCATCG
 TAAAGAACAATTTTGAAGCATTTCAAGTCAAGTTGCTCAATGTACCTATAACCAAGACCGT
 TCAGCTGGATATTACGGCCCTTTTAAAGACCGTAAAGAAAAAATAAGCACAAAGTTTAA
 TCCGGCCTTTATTTCACATTTCTTGCCCGCCCTGATGAATGCTCATCCGGAATTCGGTATG
 GCAATGAAAAGACCGGTGAGCTGGTGATATGGGATAGTGTTCACCCCTTGTACACCCGTT
 TTCCATGAGCAAACTGAAAACGTTTTCATCGCTCTGGAGTGAATACCAACGACGATTTTC
 CGGCAGTTTCTACACATATATTCGCAAGATGTGGCGTGTACCGGTGAAAAACCTGGCC
 PATTTCCCTAAAGGGTTTATTGAGAAATATGTTTTTCTCAGCCAAATCCCTGGGTGA
 GTTTCACCAAGTTTTGATTTAAACGTGGCCAAATATGGACAACCTTCTTCCGCCCCGTTTTT

CACCATGGGCAAAATATTAACGCCAAGGCGACAAAGGTGCTGATGCCGCTGGCGATTTC
 AGGTTCAATCATGCCGTTTGTGALGGCTTCCATGTCCGCAGAAATGCTTAATGAATTTAC
 AACAGTACTGCCGATGAGTGGCAGGGCGGGCGTAAAGATCTGGATCCGGCTTACTTA
 AAAGCCAGATAACAGTATGCCGTATTTGCCGCTGATTTTGGGTLATAAGAATATAT
 ACTGATATGTAACCAGAGTATGTCAAAAGAAGGTATGCTATGAAGCAGCCGTAAT
 ACAGTGACAGTTGACAGCGACAGCCTATCAGTTGCTCAAGGCATATATGATGTCAATA
 TCTCCGGTCTGGTAAACAACCATGCAGATGAAGCCCGTCTGCCGTGCCGAAC
 GCTGGAAGCGGAAATCAGGAAGGATGCTGAGGTCGCCCGTTTATTGAATG
 AACGGCTCTTTGCTGACGAGAACAGGGGCTGGTGAATGCAAGTTTAAAGTTTACAC
 CTATAAAAAGAGAGAGCCGTTATCGTCTGTTTGTGGATGTACAGAGTATATTATTGA
 CACGCCCGGGCGACCGGATGGTATCCCCCTGGCCAGTGCACGCTCTGCTGCAGATA
 AAGTCTCCCGTGAACCTTTACCCGGTGGTGCATATCCGGGATGAAGCTGGCCGATGA
 TGACCACCGATATGGCCAGTGTGCCGCTCTCCGTTATCCGGGAAGAGTGGCTGATC
 TCAGCCACCCGGAATGACATCAAAAACGCCATTAACCTGATGTTCTGGGGAATTA
 TAAATGTCAGGCTCCCTTATACACAGCCAGTCTGCAGGTCGACCGTCTC**GCTCTTCG**
 AAGGTTACTTCCGCTCTCTGTCAAGCCGACGGCCCTCACTCCAGAGCAAAGTGGTTGC
 GATCGCTTCAACAACAGCGGTGGAAGAACCTGCCCTGGAAATCAATCGTGGCCCAAGCTTTC
 GAGGCCGGACCCCGCCCTGGCCGCACTCACTAATGATCACTCTTGTAGCCCTGGCCCTG
CCTCGCGGACGACCCCGCCTTGATGCCGTTGAAGAAGGGCTCCCGCACCGCCCTG
CAITGATTTAAGCCGACCAACAAGGATTTCCCGAAGAGGACATAGCCCAAGAGAA
GAGAAAGGTGAGGCCAAGCGGTTCCCGACGGGCTGACCGCATTTGACGATTTTGATC

도면 20e

<141> 2014-06-30
 <150> US 61/858,866
 <151> 2013-07-26
 <160> 224
 <170> PatentIn version 3.5
 <210> 1
 <211> 15
 <212> DNA
 <213> Artificial
 <220><223> Oligonucleotide sequence
 <400> 1
 tggggcaagc ttctg 15
 <210> 2
 <211> 102
 <212> DNA
 <213> Artificial
 <220><223> TALE template
 <220><221> misc_feature
 <222> (34)..(39)
 <223> n is a, c, g, or t
 <400> 2
 ctaaccagc agcaggtcgt ggcaatgcc tcnnnnnng gcggaaaaca ggcattggaa 60
 acagtacagc ggtgctgcc ggtgctgtgc caagcgacg ga 102
 <210> 3
 <211> 102
 <212> DNA
 <213> Artificial
 <220><223> PCR primer
 <220><221> misc_feature
 <222> (34)..(39)
 <223> n is a, c, g, or t
 <400> 3
 ctaaccctg aacaggtagt cgctatagct tcannnnng ggggcaagca agcacttgag 60

accgttcaac gactcctgcc agtgcctctgc caagcccatg ga 102

<210> 4

<211> 102

<212> DNA

<213> Artificial

<220><223> PCR primer

<220><221> misc_feature

<222> (34)..(39)

<223> n is a, c, g, or t

<400> 4

ttgactccgg agcaagtcgt cgcgatcgcg agcnnnnnng gggggaagca ggcgctggaa 60

actgttcaga gactgctgcc tgtactttgt caggcgcatg gt 102

<210> 5

<211> 102

<212> DNA

<213> Artificial

<220><223> PCR primer

<220><221> misc_feature

<222>

(34)..(39)

<223> n is a, c, g, or t

<400> 5

ctcacccecg aacaggttgt cgcaatagca agtnnnnnng gcggtaaagca agccctagag 60

actgtgcaac gectgctccc cgtgctgtgt caggctcacg gt 102

<210> 6

<211> 102

<212> DNA

<213> Artificial

<220><223> PCR primer

<220><221> misc_feature

<222> (34)..(39)

<223> n is a, c, g, or t

<400> 6

ctgacacctg aacaagttgt cgcgatagcc agtnnnnnng ggggaaaaca agctctagaa 60

acggttcaaa ggttggtgcc cgttctgtgc caagcacatg gg 102

<210> 7
 <211> 102
 <212> DNA
 <213> Artificial
 <220><223> PCR primer
 <220><221> misc_feature
 <222> (34)..(39)
 <223> n is a, c, g, or t
 <400> 7

ttaacacccg aacaagtagt agcgatagcg tcannnnnng ggggtaaaca ggctttggag 60
 acggtacagc ggttattgcc ggtcctctgc caggcccacg ga 102

<210> 8
 <211> 102
 <212> DNA
 <213> Artificial
 <220><223> PCR primer
 <220><221> misc_feature
 <222> (34)..(39)
 <223> n is a, c, g, or t
 <400> 8

cttacgccag aacaggtggt tgcaattgcc tcnnnnnng gcggaaca agcgttgaa 60

actgtcaga gactccttcc tgttttgtgt caagcccacg gc 102

<210> 9
 <211> 102
 <212> DNA
 <213> Artificial
 <220><223> PCR primer
 <220><221> misc_feature
 <222> (34)..(39)
 <223> n is a, c, g, or t
 <400> 9

ttgacgctg agcaggttgt ggccatcgct agcnnnnng gaggaagca ggctcttgaa 60

accgtacagc gacttctccc agttttgtgc caagctcacg gg 102

<210> 10

<211> 102

<212> DNA

<213> Artificial

<220><223> PCR primer

<220><221> misc_feature

<222>

> (34)..(39)

<223> n is a, c, g, or t

<400> 10

ctaaccctccg agcaagtagt tgccatagca agcnnnnnng gaggaaaaca ggcattagaa 60

acagttcagc gcttgctccc ggtactctgt caggcacacg gt 102

<210> 11

<211> 102

<212> DNA

<213> Artificial

<220><223> PCR primer

<220><221> misc_feature

<222> (34)..(39)

<223> n is a, c, g, or t

<400> 11

ctaactccgg aacaggtcgt agccattgct tcnnnnnng gcggcaaaca ggcgctagag 60

acagtccaga ggctcttgcc tgtgttatgc caggcacatg gc 102

<210> 12

<211> 102

<212> DNA

<213> Artificial

<220><223> PCR primer

<220><221> misc_feature

<222> (34)..(39)

<223> n is a, c, g, or t

<400> 12

ctcaccctccg agcaggtcgt tgccatcgcc agtnnnnnng gcggaaagca agctctcgaa 60

acagtacaac ggctgttgcc agtcctatgt caagctcatg ga 102
 <210> 13
 <211> 102
 <212> DNA
 <213> Artificial
 <220><223> PCR primer
 <220><221> misc_feature
 <222> (34)..(39)
 <223> n is a, c, g, or t
 <400> 13

ctgacgccg agcaggtagt ggcaatcgca tctnnnnng gaggtaaaca agcactcgag 60

actgtccaaa gattgttacc cgtactatgc caagcgcattg gt 102
 <210> 14
 <211> 102
 <212> DNA
 <213> Artificial
 <220><223> PCR primer
 <220><221> misc_feature
 <222> (34)..(39)
 <223> n is a, c, g, or t
 <400> 14

ttaacccag agcaagttgt ggctattgca tctnnnnng gtggcaaaca agccttggag 60

acagtgaac gattactgcc tgtcttatgt caggcccatg gc 102
 <210> 15
 <211> 102
 <212> DNA
 <213> Artificial
 <220><223> PCR primer
 <220><221> misc_feature
 <222>
 > (34)..(39)
 <223> n is a, c, g, or t
 <400> 15

cttactcctg agcaagtcgt agctatgcc agcnnnnng gtgggaaaca ggcctggaa 60

accgtacaac gtctcctccc agtactttgt caagcacacg gg 102

<210> 16

<211> 102

<212> DNA

<213> Artificial

<220><223> PCR primer

<220><221> misc_feature

<222> (34)..(39)

<223> n is a, c, g, or t

<400> 16

ttgacaccgg aacaagtggg ggcgattgcg tcnnnnnng gaggcaagca ggcactggag 60

accgtccaac ggettcttcc ggttctttgc caggctcatg gg 102

<210> 17

<211> 102

<212> DNA

<213> Artificial

<220><223> PCR primer

<220><221> misc_feature

<222> (34)..(39)

<223> n is a, c, g, or t

<400> 17

ctcacgccag agcaggtggg agcaatagcg tcgnnnnng gtggtaaagca agcgcttgaa 60

acggtccagc gtcttctgcc ggtgttgtgc caggcgcacg ga 102

<210> 18

<211> 102

<212> DNA

<213> Artificial

<220><223> PCR primer

<220><221> misc_feature

<222> (34)..(39)

<223> n is a, c, g, or t

<400> 18

ctcacaccag aacaagtggg tgctattgct agtnnnnng gtggaaagca ggcctcgag 60

acgggtgcaga ggttacttcc cgtcctctgt caagcgcacg gc 102

<210> 19

<211> 60

<212> DNA

<213> Artificial

<220><223> PCR primer

<220><221> misc_feature

<222> (34)..(39)

<223> n is a, c, g, or t

<400> 19

ctcactccag agcaagtgg tgcgatcgct tcannnnng gtggaagacc tgcctggaa 60

<210> 20

<211> 5714

<212> DNA

<213> Artificial

<220><223> Oligonucleotide sequence encoding TALEN backbone

<400> 20

atgtcgcgga cccggtccc tccccacc gcaccagcc cagcgtttc ggccgactcg 60

ttctcagacc tgcttaggca gttcgacccc tactgttta acacatcgtt gttcgactcc 120

cttctccgtt ttggggcgca ccatacggag gcggccaccg gggagtggga tgaggtgcag 180

tcgggatgga gagctgcgga tgcaccacc ccaaccatgc gggaggccgt caccgctgcc 240

cgaccgccga gggcgaagcc cgcaccaagg cggagggcag cgcaaccgtc cgacgcaagc 300

cccgcagcgc aagtagattt gagaactttg ggatattcac agcagcagca ggaaaagatc 360

aagcccaaag tgaggtcgac agtcgcgcag catcacgaag cgctggtggg tcatgggttt 420

acacatgccc acatcgtagc cttgtcgcag caccctgcag cccttggcac ggtcgcctgc 480

aagtaccagg acatgattgc ggcgttgccg gaagccacac atgaggcgat cgtcgggtgtg 540

gggaaacagt ggagcggagc ccgagcgtt gaggccctgt tgacggtcgc gggagagctg 600

agagggcctc cccttcagct ggacacgggc cagttgctga agatcgcgaa gcggggagga 660

gtcacggcgg tcgaggcggg gcacgcgtgg cgcaatgcgc tcacgggagc acccctcaac 720

agttcacgct gacagagacc gcggccgcat taggcacccc aggttttaca ctttatgctt 780

ccggtcgtga taatgtgtgg attttgagtt aggatccgtc gagattttca ggagctaagg 840

aagctaaaat ggagaaaaaa atcactggat ataccaccgt tgatatatcc caatggcatc 900

gtaaagaaca ttttgaggca tttcagtcag ttgctcaatg tacctataac cagaccgttc 960
 agctggatat tacggccttt ttaaagaccg taaagaaaaa taagcacaag ttttatccgg 1020
 cctttattca cattcttgcc cgctgatga atgctcatcc ggaattccgt atggcaatga 1080
 aagacgggta gciggtgata tgggatagtg ttcacccttg ttacaccgtt ttccatgagc 1140
 aaactgaaac gttttcatcg ctctggagtg aataccacga cgatttccgg cagtttctac 1200
 acatatattc gcaagatgtg gcgtgttacg gtgaaaacct ggctatttc cctaaagggt 1260
 ttattgagaa taigtttttc gtctcagcca atccctgggt gagtttcacc agttttgatt 1320

 taaacgtggc caatatggac aacttcttcg cccccgttt caccatgggc aaatattata 1380
 cgcaaggcga caaggtgctg atgccgctgg cgattcaggt tcatcatgcc gtttggatg 1440
 gcttccatgt cggcagaatg cttaatgaat tacaacagta ctgcgatgag tggcagggcg 1500
 gggcgtaaaag atctggatcc ggcttactaa aagccagata acagtatgcg tatttgcgcg 1560
 ctgatttttg cgttataaga atatatactg atatgtatac ccgaagtatg tcaaaaagag 1620
 gtatgctatg aagcagcgtg ttacagtgac agttgacagc gacagctatc agttgctcaa 1680
 ggcatatatg atgtcaatat ctccggtctg gtaagcacia ccatgcagaa tgaagcccgt 1740

 cgtctcgtg ccgaacgctg gaaagcggaa aatcaggaag ggatggctga ggtcgcccgg 1800
 tttattgaaa tgaacggctc ttttctgac gagaacaggg gctggtgaaa tgcagtttaa 1860
 ggtttacacc tataaaagag agagccgtta tcgtctgttt gtggatgtac agagtgatat 1920
 tattgacacg cccggcgac ggatggatgat cccctggcc agtgcacgctc tgctgtcaga 1980
 taaagtctcc cgtgaacttt acccgggtgt gcatatcggg gatgaaagct ggcgcatgat 2040
 gaccaccgat atggccagtg tgccggtctc cgttatcggg gaagaagtgg ctgatctcag 2100
 ccaccgcgaa aatgacatca aaaacccat taacctgatg ttctggggaa tataaatgtc 2160

 aggtccctt atacacagcc agtctgcagg tcgacggtct cgctcttcga aggttacttc 2220
 ccgtcctctg tcaagcgcac ggcctcactc cagagcaagt ggttgcgac gcttcaaaca 2280
 acggtggaag acctgcctg gaatcaatcg tgcccagct ttcgagggcg gaccccgcc 2340
 tggccgact cactaatgat catctttag cgctggcctg cctcggcgga cgaccgcct 2400
 tggatcgggt gaagaagggg ctccgcacg cgctgcatt gattaagcgg accaacagaa 2460
 ggattccga gaggacatca catcgagtgg caggttccca actcgtgaag agtgaacttg 2520
 aggagaaaaa gtcggagctg cggcaciaat tgaatacgt accgcatgaa tacatcgaac 2580

 ttatcgaat tgctaggaac tcgactcaag acagaatcct tgagatgaag gtaatggagt 2640
 tctttatgaa ggtttatgga taccgaggga agcatctcgg tggatcacga aaaccgcagc 2700
 gagcaatcta tacggtgggg agcccgattg attacggagt gatcgtcgac acgaaagcct 2760

acagcggtag gtacaatctt cccatcgggc aggcagatga gatgcaacgt tatgtcgaag 2820
 aaaatcagac caggaacaaa cacatcaatc caaatgagtg gtggaaagtg tatccttcat 2880
 cagtgaccga gtttaagttt ttgtttgtct ctgggcattt caaaggcaac tataaggccc 2940
 agctcacacg gttgaatcac attacgaact gcaatggtgc ggttttgtcc gtagaggaac 3000

 tgctcattgg tggagaaatg atcaaagcgg gaactctgac actggaagaa gtcagacgca 3060
 agtttaacaa tggcgagatc aatttccgca agctaaaatg gagaaaaaaaa tcaactggata 3120
 taccaccgtt gatatatccc aatggcatcg taaagaacat tttgaggcat ttcagtcagt 3180
 tgctcaatgt acctataacc agaccgttca gctggatatt acggcctttt taaagaccgt 3240
 aaagaaaaat aagcacaagt tttatccggc ctttattcac attcttgccc gcctgatgaa 3300
 tgctcatccg gaattccgta tggcaatgaa agacggtgag ctggtgatat gggatagtgt 3360
 tcacccttgt tacaccgttt tccatgagca aactgaaacg ttttcatcgc tctggagtga 3420

 ataccacgac gatttccggc agtttctaca catatatcgc caagatgtgg cgtgttacgg 3480
 tgaaaacctg gcctatttcc ctaaagggtt tattgagaat atgtttttcg tctcagccaa 3540
 tccctgggtg agtttaccac gttttgattt aaacgtggcc aatatggaca acttcttcgc 3600
 ccccttttc accatgggca aatattatac gcaaggcgac aaggctgctga tgccgctggc 3660
 gattcaggtt catcatgccg tttgtgatgg ctccatgtc ggcagaatgc ttaatgaatt 3720
 acaacagtac tgcgatgagt ggcagggcgg ggcgtaaaga tctggatccg gcttactaaa 3780
 agccagataa caglatgcgt atttgcgcgc tgatttttgc ggtataagaa tatatactga 3840

 tatgtatacc cgaagtatgt caaaaagagg tatgctatga agcagcgtat tacagtgaca 3900
 gttgacagcg acagctatca gttgctcaag gcataatga tgtcaatc tccggctctgg 3960
 taagcacaac catgcagaat gaagccctgc gctcgcgtgc cgaacgctgg aaagcggaaa 4020
 atcaggaagg gatggctgag gtcgcccgtt ttattgaaat gaacggctct tttgctgacg 4080
 agaacagggg ctggtgaaat gcagtttaag gtttacacct ataaaagaga gagccgttat 4140
 cgtctgtttg tggatgtaca gactgatatt attgacacgc cggggcgacg gatggtgatc 4200
 cccctggcca gtgcacgtct gctgicagat aaagtctccc gtgaacttta cccggtggtg 4260

 catatcgggg atgaaagctg gcgcatgatg accaccgata tggccagtgt gccggtctcc 4320
 gttatcgggg aagaagtggc tgatctcagc caccgcgaaa atgacatcaa aaacgccatt 4380
 aacctgatgt tctggggaat ataaatgtca ggctccctta tacacagcca gctctcaggt 4440
 cgacggtctc gctcttcgaa ggttacttcc cgtcctctgt caagcgcacg gcctcactcc 4500
 agagcaagtg gttcgatcg ctccaacaaa cgggtggaaga cctgccttgg aatcaatcgt 4560
 ggcccagctt tcgaggccgg accccgcgct ggccgcactc actaatgatc atctttagc 4620

gctggcctgc ctcggcggac gacccgcctt ggatgcggtg aagaaggggc tcccgcacgc 4680

gcctgcattg attaagcggc ccaacagaag gattcccag aggacatagc cccaagaaga 4740

agagaaaggt ggaggccagc ggttcggac gggctgacgc attggacgat tttgatctgg 4800

atatgctggg aagtgacgcc ctcgatgatt ttgacctga catgcttggt tcggatgccc 4860

ttgatgactt tgacctcgac atgctcggca gtgacgcct tgatgattc gacctggaca 4920

tgctgattaa ctctagaggc agtggagagg gcagaggaag tctgctaaca tgcggtgacg 4980

tcgaggagaa tcctggccca gtgagcaagg gcgaggagct gttcaccggg gtggtgccca 5040

tcctggtcga gctggacggc gacgtaaacg gccacaagtt cagcgtgtcc ggcgagggcg 5100

agggcgatgc cacctacggc aagctgaccc tgaagtcat ctgcaccacc ggcaagctgc 5160

ccgtgcctg gccaccctc gtgaccacc tgacctacgg cgtgcagtgc ttcagccgt 5220

accccgacca catgaagcag cagcacttct tcaagtccgc catgcccga ggctacgtcc 5280

aggagcgac catcttctt aaggacgac gcaactaca gaccgcgcc gaggtgaagt 5340

tcgagggcga caccctggtg aaccgatcg agctgaagg catcgactt aaggaggacg 5400

gcaacatcct ggggcacaag ctggagtaca actacaacag ccacaacgtc tatatcatgg 5460

ccgacaagca gaagaacggc atcaaggtga acttcaagat ccgccacaac atcgaggacg 5520

gcagcgtgca gctcggcag cactaccagc agaacacccc catcggcgac ggccccgtgc 5580

tgctgcccga caaccactac ctgagcacc agtccgcct gagcaaagac cccaacgaga 5640

agcgcgatca catggtcctg ctggagtctg tgaccgccgc cgggatcact ctcgcatgg 5700

acgagctgta caag 5714

<210> 21

<211> 3465

<212> DNA

<213> Artificial

<220><223> Oligonucleotide sequence encoding TALE backbone sequence

<400> 21

atgtcggga cccggctccc ttcccacc gcaccagcc cagcgttttc ggccgactcg 60

ttctcagacc tgettaggca gttcgacccc tcaactgitta acacatcgtt gttcgactcc 120

cttctccgt ttggggcga ccatacggag gcggccaccg gggagtggga tgaggigcag 180

tcgggattga gagctcgga tgcaccacc ccaaccatgc gggtggccgt caccgtgcc 240

cgaccccga gggcgaagcc cgcaccaagg cggagggcag cgcaaccgtc cgacgcaagc 300

cccgcagcgc aagtagattt gagaactttg ggatattcac agcagcagca ggaaaagatc 360

aagcccaaag tgaggctgac agtcgcgag catcacgaag cgctggggg tcatgggttt 420
acacatgccc acatcgtagc cttgtcgag caccctgcag cccttggcac ggtcgccgtc 480

aagtaccagg acatgattgc ggcgttgccg gaagccacac atgaggcgat cgtcgggtgtg 540
gggaaacagt ggagcggagc ccgagcgctt gaggccctgt tgacggctgc gggagagctg 600
agagggcctc cccttcagct ggacacgggc cagttgtga agatcgcgaa gcggggagga 660
gtcacggcgg tegaggcggg gcacgcgtgg cgcaatgcgc tcacgggagc acccctcaac 720
agttcacgct gacagagacc gcggccgcat taggcacccc aggctttaca ctttatgctt 780
ccggctcgta taatgtgtgg attttgagtt aggatccgtc gagattttca ggagctaagg 840
aagctaaaat ggagaaaaa atcactggat ataccaccgt tgatatatcc caatggcatc 900

gtaaagaaca ttttgaggca tttcagtcag ttgctcaatg tacctataac cagaccgttc 960
agctggatat tacggccttt ttaaagaccg taaagaaaa taagcacaag ttttatccgg 1020
cctttattca cattcttgcc cgcctgatga atgctcatcc ggaattccgt atggcaatga 1080
aagacggatga gctggatgata tgggatagtg ttacccttg ttacaccgtt ttccatgagc 1140
aaactgaaac gttttcatcg ctctggagtg aataccacga cgatttccgg cagttttctac 1200
acatatattc gcaagatgtg gcgtgttacg gtgaaaacct ggcctatttc cctaaagggt 1260
ttattgagaa tatgtttttc gtctcagcca atccctgggt gagtttcacc agttttgatt 1320

taaactgtgc caatatggac aacttcttcg cccccgttt caccatgggc aaatattata 1380
cgcaaggcga caaggtgtg atgccgctgg cgattcaggt tcatcatgcc gtttgtgatg 1440
gcttccatgt cggcagaatg cttaatgaat tacaacagta ctgcgatgag tggcagggcg 1500
gggcgtaaag atctggatcc ggcttactaa aagccagata acagtatgcg tatttgcgcg 1560
ctgatttttg cggataaaga atataactg atatgtatac ccgaagtatg tcaaaaagag 1620
gtatgctatg aagcagcgtg ttacagtgc agttgacagc gacagctatc agttgctcaa 1680
ggcatatatg atgtcaatat ctccggtctg gtaagcaca ccatgcagaa tgaagcccgt 1740

cgtctgcgtg ccgaacgctg gaaagcggaa aatcaggaag ggatggctga ggtcgcccgg 1800
tttattgaaa tgaacggctc ttttctgac gagaacaggg gctggtgaaa tgcagtttaa 1860
ggtttacacc tataaaagag agagccgtta tcgtctgttt gtggatgtac agagtgatat 1920
tattgacag cccggcgagc ggatggatgat cccctggcc agtgcacgtc tgctgtcaga 1980
taaagtctcc cgigaacttt acccgggtgt gcatatcggg gatgaaagct ggcgcatgat 2040
gaccaccgat atggccagtg tgccggtctc cgttatcggg gaagaagtgg ctgatctcag 2100

ccaccgcgaa aatgacatca aaaacgccat taacctgatg ttctggggaa tataaatgtc 2160

aggctccctt atacacagcc agtctgcagg tcgacggtct cgctcttcga aggttacttc 2220

ccgtcctctg tcaagcgcac ggcctcactc cagagcaagt ggttgcgac gettcaaaaca 2280

acggtggaag acctgccctg gaatcaatcg tggcccagct ttcgaggccg gaccccgcgc 2340

tggccgcact cactaatgat catctttagt cgctggcctg cctcggcgga cgacccgctt 2400

tggatgcggt gaagaagggg ctcccgcacg gcctgcatt gattaagcgg accaacagaa 2460

ggattcccga gaggacatag cccaagaag aagagaaagg tggaggccag cggttccgga 2520

cgggctgacg cattggacga ttttgatctg gatatgctgg gaagtgacgc cctcgatgat 2580

tttgacctg acatgcttgg ttcggatgcc cttgatgact ttgacctga catgctcggc 2640

agtgacgcc ttgatgattt cgacctggac atgctgatta actctagagg cagtggagag 2700

ggcagaggaa gtctgctaac atgcggtgac gtcgaggaga atcctggccc agtgagcaag 2760

ggcgaggagc tgttcaccgg ggtggtgcc atcctggtcg agctggacgg cgactaac 2820

ggccacaagt tcagcgtgtc cggcgagggc gagggcgatg ccacctacgg caagctgacc 2880

ctgaagtta tctgaccac cggaagctg cccgtgcctt ggeccacct cgtgaccacc 2940

ctgacctacg gcgtgcagtg cttcagccgc taccgccacc acatgaagca gcacacttc 3000

ttcaagtccg ccatgcccga aggctacgtc caggagcgca ccatcttctt caaggacgac 3060

ggcaactaca agaccccgcg cgaggtgaag ttcgaggcg acacctggt gaaccgcatc 3120

gagctgaagg gcatcgactt caaggaggac ggcaacatcc tggggcacia gctggagtac 3180

aactacaaca gccacaact ctatatcatg gccgacaagc agaagaacgg catcaaggtg 3240

aacttaaga tccgccaaa catcgaggac ggacgcgtgc agctcgccga cactaccag 3300

cagaacacc ccatcgcgca cggccccgtg ctgctgccc acaaccacta cctgagcacc 3360

cagtcgccc tgagcaaaga cccaacgag aagcgcgac acatggtcct getggagttc 3420

gtgaccgccg ccgggatcac tctcgcatg gacgagctgt acaag 3465

<210> 22

<211> 1368

<212> PRT

<213> Streptococcus pyogenes

<400> 22

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

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

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> 23

<211> 82

<212> DNA

<213> Artificial

<220><223> Oligonucleotide sequence

<220><221> misc_feature

<222> (41)..(41)

<223> n is a, c, g, or t

<400> 23

ttcttggctt tatatatctt gtggaagga cgaacaccg ngtttagag ctagaaatag 60

caagttaaaa taaggctagt cc 82

<210> 24

<211> 1692

<212> DNA

<213> Artificial

<220><223> Oligonucleotide sequence encoding re-TALE

<400> 24

ctaaccctg aacaggtagt cgctatagct tcaaatatcg ggggcaagca agcacttgag 60

accgttcaac gactcctgcc agtgcctctgc caagcccatg gattgactcc ggagcaagtc 120

gtcgcgatcg cgagcaacgg cggggggaag caggcctgg aaactgttca gagactgctg 180

cctgtacttt gtcaggcgca tggctctacc cccgaacagg ttgtcgcaat agcaagtaat 240

ataggcggtg agcaagccct agagactgtg caagcctgc tccccgtgct gtgtcaggct 300

cacggtctga cacctgaaca agttgtcgcg atagccagtc acgacggggg aaaacaagct 360

ctagaaacgg ttcaaaggtt gttgcccgtt ctgtgccaag cacatgggtt aacaccgaa 420

caagtagtag cgatagcgtc aaataacggg ggtaaacagg ctttggagac ggtacagcgg 480

ttattgccgg tcctctgcca ggcccacgga cttacgccag aacaggtggt tgcaattgcc 540

tccaacatcg gcgggaaaca agcgttggaa actgtgcaga gactccttcc tgttttgtgt 600

caagcccacg gcttgaccgc tgagcagggt gtggccatcg ctagccacga cggaggggaag 660

caggctcttg aaaccgtaca gcgacttctc ccagttttgt gccaaagtca cgggctaacc 720

cccagcaag tagttgcat agcaagcaac ggaggaggaa aacaggcatt agaaacagtt 780

cagcgttgc tcccgtact ctgtcaggca cacggtetaa ctccggaaca ggtcgtagcc 840

attgcttccc atgatggcgg caaacaggcg ctagagacag tccagaggct cttgcctgtg 900

ttatgccagg cacatggcct caccocggag caggctgttg ccatcgccag taatatcggc 960

ggaaagcaag ctctcgaac agtacaacgg ctgttgccag tcctatgtca agctcatgga 1020

ctgaccccc agcaggtagt ggcaatcgca tctcacgatg gaggtaaaca agcactcgag 1080

actgtccaaa gattgttacc cgtaactatgc caagcgcgatg gtttaacccc agagcaagtt 1140

gtggctattg catctaacgg cgggtggcaaa caagccttgg agacagtgca acgattactg 1200

cctgtcttat gtcaggccca tggccttact cctgagcaag tcgtagctat cgccagcaac 1260

ataggtggga aacaggccct ggaaaccgta caactctcc tcccagtact ttgtcaagca 1320

cacgggttga caccggaaca agtgggtggcg attgcgtcca acggcggagg caagcaggca 1380

ctggagaccg tccaacggct tcttccggtt ctttgccagg ctcatgggct cacgccagag 1440

caggtggtag caatagcgtc gaacatcggg ggtaaagcaag cgcttgaaac ggtccagcgt 1500

cttctgccgg tgttgtgcca ggccgacgga ctcacaccag aacaagtggg tgctattgct 1560

agtaacaacg gtggaaagca ggccctcgag acgggtgcaga ggttacttcc cgtcctctgt 1620

caagcgcacg gcctcactcc agagcaagtg gttgcgatcg cttcaaaca tgggtgaaga 1680

cctgccctgg aa 1692

<210> 25
 <211> 237
 <212> DNA
 <213> Artificial
 <220><223> Template sequence
 <220><221> misc_feature
 <222> (64)..(69)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (166)..(171)
 <223> n is a, c, g, or t
 <400> 25
 cgcaatgcgc tcacgggagc acccctcaac ctaaccctg aacaggtagt cgctatagct 60
 tcannnnnng ggggcaagca agcacttgag accgttcaac gactcctgcc agtgctctgc 120
 caagcccatg gattgactcc ggagcaagtc gtcgcatcg cgagcnnnnn nggggggaag 180
 caggcgctgg aaactgttca gagactgctg cctgtacttt gtcaggcgca tggctctc 237

<210> 26
 <211> 240
 <212> DNA
 <213> Artificial
 <220><223> Template sequence
 <220><221> misc_feature
 <222> (67)..(72)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (169)..(174)
 <223> n is a, c, g, or t
 <400> 26
 agactgctgc ctgtactttg tcaggcgcat ggtctcacc cgaacaggt tgcgcaata 60
 gcaagtnnnn nngcggttaa gcaagcccta gagactgtgc aacgcctgct ccccgctctg 120
 tgtcaggctc acggtctgac acctgaacaa gttgtcgca tagccagttn nnnnggggga 180
 aaacaagctc tagaaacggt tcaaaggttg ttgccgttc tgtgccaagc acatgggtta 240

<210> 27

<211> 232
 <212> DNA
 <213> Artificial
 <220><223> Template sequence
 <220><221> misc_feature
 <222> (59)..(64)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (161)..(166)
 <223> n is a, c, g, or t
 <400> 27
 tgcgctcacg ggagcacccc tcaacctcac ccccgaaacag gttgtcgcaa tagcaagtnn 60
 nnnnggcggt aagcaagccc tagagactgt gcaacgcctg ctccccgtgc tgtgtcaggc 120
 tcacggtctg acacctgaac aagttgtcgc gatagccagt nnnnnngggg gaaaacaagc 180
 tctagaaacg gttcaaaggt tgttgcccgt tctgtgccaa gcacatgggt ta 232

<210> 28
 <211> 246
 <212> DNA
 <213> Artificial
 <220><223> Template sequence
 <220><221> misc_feature
 <222> (67)..(72)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (169)..(174)
 <223> n is a, c, g, or t
 <400> 28
 aggttgttc cgttctgtg ccaagcacat gggttaacac ccgaacaagt agtagcgata 60
 gcgtcannnn nngggggtaa acaggctttg gagacggtac agcggttatt gccggtcctc 120
 tgccaggccc acggacttac gccagaacag gtggttgcaa ttgcctccnn nnnnggcggg 180
 aaacaagcgt tggaaactgt gcagagactc cttcctgttt tgtgtcaagc ccacggcttg 240
 acgcct 246
 <210> 29

<211> 240
 <212> DNA
 <213> Artificial
 <220><223> Template sequence
 <220><221> misc_feature
 <222> (67)..(72)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (169)..(174)
 <223> n is a, c, g, or t
 <400> 29
 agactccttc ctgttttgtg tcaagcccac ggcttgacgc ctgagcaggt tgtggccatc 60
 gctagcnnnn nnggagggaa gcaggtcttt gaaaccttac agcgacttct cccagttttg 120
 tgccaagctc acgggctaac ccccgagcaa gtagttgcca tagcaagcnn nnnnggagga 180

 aaacaggcat tagaaacagt tcagcgcttg ctcccgttac tctgtcaggc acacggtcta 240
 <210> 30
 <211> 240
 <212> DNA
 <213> Artificial
 <220><223> Template sequence
 <220><221> misc_feature
 <222> (67)..(72)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (169)..(174)
 <223> n is a, c, g, or t
 <400> 30
 cgcttgctcc cggctactctg tcaggcacac ggtctaactc cggaacaggt cgtagccatt 60
 gcttcnnnn nnggcggcaa acaggcgcta gagacctcc agaggctctt gcctgtgtta 120
 tgccaggcac atggcctcac cccggagcag gtcggtgcca tcgccagtnn nnnnggcgga 180

 aagcaagctc tcgaaacagt acaacggctg ttgccagtcc tatgtcaagc teatggactg 240
 <210> 31
 <211> 240

<212> DNA
 <213> Artificial
 <220><223> Template sequence
 <220><221> misc_feature
 <222> (67)..(72)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (169)..(174)
 <223> n is a, c, g, or t
 <400> 31
 cggctgttgc cagtcctatg tcaagctcat ggactgacgc ccgagcaggt agtggcaatc 60
 gcactcnnnn nnggaggtaa acaagcactc gagactgtcc aaagattgtt acccgtacta 120
 tgccaagcgc atggtttaac ccagagcaa gttgtggcta ttgcatctnn nnnnggtggc 180

 aaacaagcct tggagaccgt gcaacgatta ctgctgtct tatgtcaggc ccattggcctt 240
 <210> 32
 <211> 240
 <212> DNA
 <213> Artificial
 <220><223> Template sequence
 <220><221> misc_feature
 <222> (67)..(72)
 <223> n is a, c, g, or t
 <220><221> misc_feature
 <222> (169)..(174)
 <223> n is a, c, g, or t
 <400> 32
 cgattactgc ctgtcttatg tcaggcccat ggccttactc ctgagcaggt ggtcgcctatc 60
 gccagcnnnn nngggggcaa gcaagcactg gaaacagtcc agcgtttgct tccagtactt 120
 tgtcaggcgc atggattgac accggaacaa gtggtggcta tagcctcann nnnnggagga 180

 aagcaggcgc tggaaaccgt ccaacgtctt ttaccgtgac tttgccagc gcacgggctc 240
 <210> 33
 <211> 240
 <212> DNA

<213> Artificial

<220><223> Template sequence

<220><221> misc_feature

<222> (67)..(72)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (169)..(174)

<223> n is a, c, g, or t

<400> 33

cgattactgc ctgtcttatg tcaggcccat ggccttactc ctgagcaagt cgtagctatc 60

gccagcnmnn nnggtgggaa acaggccctg gaaaccgtac aacgtctcct cccagtactt 120

tgtaagcac acgggttgac accggaacaa gtggtggcga ttgcgtccnn nnnnggaggc 180

aagcaggcac tggagaccgt ccaacggctt cttccggttc tttgccaggc tcatgggctc 240

<210> 34

<211> 240

<212> DNA

<213> Artificial

<220><223> Template sequence

<220><221> misc_feature

<222> (67)..(72)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (169)..(174)

<223> n is a, c, g, or t

<400> 34

cggtctcttc cggttctttg ccaggctcat gggctcagc cagagcaggt ggtagcaata 60

gcgtcgnmnn nnggtggtaa gcaagcgtt gaaacgtcc agcgtcttct gccggtgttg 120

tgccaggegc acggactcac accagaacaa gtggttgcta ttgctagttn nnnnggtgga 180

aagcaggccc tcgagacggt gcagaggtaa cttcccgtcc tctgtcaagc gcacggcctc 240

<210> 35

<211> 49

<212> DNA

<213> Artificial

<220><223> Primer sequence
 <400> 35
 cgcaatgcgc tcacgggagc acccctcaac ctaaccctg aacaggtag 49
 <210> 36
 <211> 46
 <212> DNA
 <213> Artificial
 <220><223> Primer sequence
 <400> 36
 gagacatgc gcctgacaaa gtacaggcag cagtctctga acagtt 46
 <210> 37
 <211> 33
 <212> DNA
 <213> Artificial
 <220><223> Primer sequence
 <400> 37
 tggcgcaatg cgctcacggg agcaccctc aac 33
 <210> 38
 <211> 49
 <212> DNA
 <213> Artificial
 <220><223> Primer sequence
 <400> 38
 agactgctgc ctgtactttg tcaggcgcat ggtctcacc ccgaacagg 49
 <210> 39
 <211> 46
 <212> DNA
 <213> Artificial
 <220><223> Primer sequence
 <400> 39
 taacctatgt gcttggcaca gaacgggcaa caacctttga accgtt 46
 <210> 40
 <211> 48
 <212> DNA

<213> Artificial
 <220><223> Primer sequence

<400> 40
 aggttggtgc ccgttctgtg ccaagcacat gggttaacac ccgaacaa 48
 <210> 41
 <211> 52
 <212> DNA
 <213> Artificial
 <220><223> Primer sequence

<400> 41
 aggcgtcaag ccgtgggctt gacacaaaac aggaaggagt ctctgcacag tt 52
 <210> 42
 <211> 45
 <212> DNA
 <213> Artificial
 <220><223> Primer sequence

<400> 42
 agactccttc ctgttttgtg tcaagccac ggcttgacgc ctgag 45
 <210> 43
 <211> 40
 <212> DNA
 <213> Artificial
 <220><223> Primer sequence

<400> 43
 tagaccgtgt gcctgacaga gtaccgggag caagcgtga 40
 <210> 44
 <211> 39
 <212> DNA
 <213> Artificial
 <220><223> Primer sequence

<400> 44
 cgcttgctcc cggtactctg tcaggcacac ggtctaact 39
 <210> 45

<211> 40
 <212> DNA
 <213> Artificial
 <220><223> Primer sequence
 <400> 45
 cagtccatga gcttgacata ggactggcaa cagccgttgt 40
 <210> 46
 <211> 39
 <212> DNA
 <213> Artificial
 <220><223> Primer sequence

 <400> 46
 cggctgttgc cagtcctatg tcaagctcat ggactgacg 39
 <210> 47
 <211> 40
 <212> DNA
 <213> Artificial
 <220><223> Primer sequence
 <400> 47
 aaggccatgg gcctgacata agacaggcag taatcgttgc 40
 <210> 48
 <211> 39
 <212> DNA
 <213> Artificial
 <220><223> Primer sequence
 <400> 48
 cgattactgc ctgtcttatg tcaggcccat ggccttact 39
 <210> 49
 <211> 43
 <212> DNA
 <213> Artificial
 <220><223> Primer sequence

 <400> 49

gagcccgtgc gcctggcaaa gcaccggtaa aagacgttgg acg 43
 <210> 50
 <211> 50
 <212> DNA
 <213> Artificial
 <220><223> Primer sequence
 <400> 50
 cgattactgc ctgtcttatg tcaggcccat ggccttactc ctgagcaagt 50
 <210> 51
 <211> 40
 <212> DNA
 <213> Artificial
 <220><223> Primer sequence
 <400> 51
 gagcccatga gcctggcaaa gaaccggaag aagccgttgg 40
 <210> 52
 <211> 51
 <212> DNA
 <213> Artificial
 <220><223> Primer sequence
 <400> 52
 cggcttcttc cggttctttg ccaggctcat ggcctcagc cagagcaggt g 51
 <210> 53
 <211> 40
 <212> DNA
 <213> Artificial
 <220><223> Primer sequence
 <400> 53
 gaggccgtgc gcttgacaga ggacgggaag taacctctgc 40
 <210> 54
 <211> 18
 <212> DNA
 <213> Artificial
 <220><223> Targeting sequence

<400> 54
 tccccacttt cttgtgaa 18
 <210> 55
 <211> 18
 <212> DNA
 <213> Artificial
 <220><223> Targeting sequence

<400> 55
 taaccactca ggacaggg 18
 <210> 56
 <211> 19
 <212> DNA
 <213> Artificial
 <220><223> Targeting sequence

<400> 56
 cactttcttg tgaatcctt 19
 <210> 57
 <211> 18
 <212> DNA
 <213> Artificial
 <220><223> Targeting sequence

<400> 57
 tcacacagca agtcagca 18
 <210> 58
 <211> 18
 <212> DNA
 <213> Artificial
 <220><223>
 > Targeting sequence

<400> 58
 tagcggagca ggctcgga 18
 <210> 59
 <211> 19
 <212> DNA

<213> Artificial
 <220><223> Targeting sequence
 <400> 59
 tgggctagcg gagcaggct 19
 <210> 60
 <211> 18
 <212> DNA
 <213> Artificial
 <220><223> Targeting sequence
 <400> 60
 taccagacg agaaagct 18
 <210> 61
 <211> 18
 <212> DNA
 <213>
 Artificial
 <220><223> Targeting sequence
 <400> 61
 tcagactgcc aagcttga 18
 <210> 62
 <211> 19
 <212> DNA
 <213> Artificial
 <220><223> Targeting sequence
 <400> 62
 acccagacga gaaagctga 19
 <210> 63
 <211> 18
 <212> DNA
 <213> Artificial
 <220><223> Targeting sequence
 <400> 63
 tcttgtggct cgggagta 18
 <210> 64

<211> 18

<212> DNA

<213> Artificial

<220><223> Targeting sequence

<400> 64

tattgtcagc agagctga 18

<210> 65

<211> 19

<212> DNA

<213> Artificial

<220><223> Targeting sequence

<400> 65

agaggcatc ttgtggctc 19

<210> 66

<211> 18

<212> DNA

<213> Artificial

<220><223> Targeting sequence

<400> 66

ttgagatddd cagatgtc 18

<210> 67

<211> 18

<212> DNA

<213> Artificial

<220><223> Targeting sequence

<400> 67

tatacagtca tatcaagc 18

<210> 68

<211> 19

<212> DNA

<213> Artificial

<220><223> Targeting sequence

<400> 68

atcaagctct cttggcggg 19

<210> 69

<211> 18

<212> DNA

<213> Artificial

<220><223> Targeting sequence

<400> 69

ttcagataga ttatatct 18

<210> 70

<211> 18

<212> DNA

<213> Artificial

<220><223> Targeting sequence

<400> 70

tgccagatac ataggtgg 18

<210> 71

<211> 19

<212> DNA

<213> Artificial

<220><223> Targeting sequence

<400> 71

gcttcagata gattatattc 19

<210> 72

<211> 18

<212> DNA

<213> Artificial

<220><223> Targeting sequence

<400> 72

ttatactgtc tatatgat 18

<210> 73

<211> 18

<212> DNA

<213> Artificial

<220><223> Targeting sequence
 <400> 73
 tcagctcttc tggccaga 18
 <210> 74
 <211> 19
 <212> DNA
 <213> Artificial
 <220><223> Targeting sequence
 <400> 74
 acggatgtct cagctcttc 19
 <210> 75
 <211> 18
 <212> DNA
 <213> Artificial
 <220><223> Targeting sequence
 <400> 75
 tggccagaag agctgaga 18
 <210> 76
 <211> 18
 <212> DNA
 <213> Artificial
 <220><223> Targeting sequence
 <400> 76
 ttaccgggga gagtttct 18
 <210> 77
 <211> 19
 <212> DNA
 <213> Artificial
 <220><223> Targeting sequence
 <400> 77
 ccggggagag tttcttga 19
 <210> 78
 <211> 18
 <212> DNA

<213> Artificial
 <220><223> Targeting sequence
 <400> 78
 ttgacagaga gatgagtc 18

<210> 79
 <211> 18
 <212> DNA
 <213> Artificial
 <220><223> Targeting sequence
 <400> 79
 ttagcagaag ataagatt 18

<210> 80
 <211> 19
 <212> DNA
 <213> Artificial
 <220><223> Targeting sequence
 <400> 80
 gaaatcttat cttctgcta 19

<210> 81
 <211> 18
 <212> DNA
 <213> Artificial
 <220><223> Targeting sequence
 <400> 81
 tataagacta aactacc 18

<210> 82
 <211> 18
 <212> DNA
 <213> Artificial
 <220><223> Targeting sequence
 <400> 82
 tcgtctgcca ccacagat 18

<210> 83

<211> 19
 <212> DNA
 <213> Artificial
 <220><223> Targeting sequence
 <400> 83
 aatgcatgac attcatctg 19
 <210> 84
 <211> 18
 <212> DNA
 <213> Artificial
 <220><223> Targeting sequence
 <400> 84
 taaaacagtt tgcattca 18

 <210> 85
 <211> 18
 <212> DNA
 <213> Artificial
 <220><223> Targeting sequence
 <400> 85
 tataaagtcc tagaatgt 18
 <210> 86
 <211> 19
 <212> DNA
 <213> Artificial
 <220><223> Targeting sequence
 <400> 86
 aacagtttgc attcatgga 19
 <210> 87
 <211> 18
 <212> DNA
 <213> Artificial
 <220><223> Targeting sequence
 <400> 87

tggccatctc tgacctgt 18

<210> 88
 <211> 18
 <212> DNA
 <213> Artificial
 <220><223> Targeting sequence
 <400> 88

tagtgagccc agaagggg 18

<210> 89
 <211> 19
 <212> DNA
 <213> Artificial
 <220><223> Targeting sequence
 <400> 89

ccagaagggg acagtaaga 19

<210> 90
 <211> 18
 <212> DNA
 <213> Artificial
 <220><223> Targeting sequence
 <400> 90

taggtacctg gctgtcgt 18

<210> 91
 <211> 18
 <212> DNA
 <213> Artificial
 <220><223> Targeting sequence
 <400> 91

tgaccgtcct ggctttta 18

<210> 92
 <211> 19
 <212> DNA
 <213> Artificial

<220><223> Targeting sequence
 <400> 92
 ctgacaatcg ataggtacc 19
 <210> 93
 <211> 18
 <212> DNA
 <213> Artificial
 <220><223> Targeting sequence
 <400> 93
 tgtcatggtc atctgcta 18

 <210> 94
 <211> 18
 <212> DNA
 <213> Artificial
 <220><223> Targeting sequence
 <400> 94
 tcgacaccga agcagagt 18
 <210> 95
 <211> 19
 <212> DNA
 <213> Artificial
 <220><223> Targeting sequence
 <400> 95
 acaccgaagc agagttttt 19
 <210> 96
 <211> 18
 <212> DNA
 <213> Artificial
 <220><223> Targeting sequence
 <400> 96
 tgccccgcg aggccaca 18

 <210> 97
 <211> 18

<212> DNA
 <213> Artificial
 <220><223> Targeting sequence
 <400> 97
 tctggaagtt gaacaccc 18
 <210> 98
 <211> 19
 <212> DNA
 <213> Artificial
 <220><223> Targeting sequence
 <400> 98
 ggaagttgaa cacccttgc 19
 <210> 99
 <211> 90
 <212> DNA
 <213> Artificial
 <220><223> Target sequence
 <400> 99
 ctactgtcat tcaggccaat acccagacga gaaagctgag ggtataacag gtttcaagct 60

 tggcagtctg actacagagg ccaactggctt 90
 <210> 100
 <211> 90
 <212> DNA
 <213> Artificial
 <220><223> Target sequence
 <400> 100
 ctactgtcat tcagcccaat accctaacga gaaagctgag ggtataacag gtttcaagct 60
 tggcagtctg actacagagg ccaactggctt 90
 <210> 101
 <211> 90
 <212> DNA
 <213> Artificial
 <220><223> Target sequence
 <400> 101

ctactgtcat tcagcccaat acccagacga gaaaagtgag ggtataacag gtttcaagct 60
 tggcagtctg actacagagg ccaactggctt 90

<210> 102
 <211> 90
 <212> DNA
 <213> Artificial
 <220><223> Target sequence
 <400> 102

ctactgtcat tcagcccaat acccagacga gaaagctgag ggtataacag gtttcaagct 60
 tggcagtctg actacagagg ccaactggctt 90

<210> 103
 <211> 90
 <212> DNA
 <213> Artificial
 <220><223> Target sequence
 <400> 103

ctactgtcat tcagcccaat acccagacga gaaagctgag ggtataacag gttttagct 60
 tggcagtctg actacagagg ccaactggctt 90

<210>
 > 104

<211> 90
 <212> DNA
 <213> Artificial
 <220><223> Target sequence
 <400> 104

ctactgtcat tcagcccaat acccagacga gaaagctgag ggtataacag gtttcaagct 60
 tggctctctg actacagagg ccaactggctt 90

<210> 105
 <211> 90
 <212> DNA
 <213> Artificial
 <220><223> Target sequence
 <400> 105

ctactgtcat tcagcccaat acccagacga gaaagctgag ggtataacag gtttcaagct 60

tggcagtctg actagtgagg ccactggcctt 90
 <210> 106

<211> 90
 <212> DNA
 <213> Artificial
 <220><223> Target sequence
 <400> 106

cactttatat ttcctgctt aaacagtccc cggagggtgg gtgcggaaaa ggctctacac 60
 ttgttatcat tcctctcca ccacaggcat 90

<210> 107
 <211> 90
 <212> DNA
 <213> Artificial
 <220><223> Target sequence
 <400> 107

tttgtatttg ggttttttta aaacctccac tctacagtta agaattctaa ggcacagagc 60
 ttcaataatt tggtcagagc caagtagcat 90

<210> 108
 <211>
 > 90
 <212> DNA
 <213> Artificial
 <220><223> Target sequence
 <400> 108

ggaggttaaa cccagcagca tgactgcagt tcttaatcaa tgccccttga attgcacata 60
 tgggatgaac tagaacattt tctcgatgat 90

<210> 109
 <211> 90
 <212> DNA
 <213> Artificial
 <220><223> Target sequence
 <400> 109

ctcgatgatt cgctgtcctt gttatgatta tgttactgag ctctactgta gcacagacat 60
 atgtccctat atggggcggg ggtgggggtg 90

<210> 110

<211> 90

<212> DNA

<213> Artificial

<220><223> Target sequence

<400> 110

gggtgtcttga tcgctgggct atttctatac tgttctggct tttcgggaagc agtcatttct 60

ttctattctc caagcaccag caattagctt 90

<210> 111

<211> 90

<212> DNA

<213> Artificial

<220><223> Target sequence

<400> 111

gcttctagtt tgctgaaact aatctgctat agacagagac tccgacgaac caattttatt 60

aggatttgat caaataaact ctctctgaca 90

<210> 112

<211> 90

<212

> DNA

<213> Artificial

<220><223> Target sequence

<400> 112

gaaagagtaa ctaagagttt gatgtttact gagtgcatag tatgcactag atgctggccg 60

tggatgcctc atagaatcct cccaacaact 90

<210> 113

<211> 90

<212> DNA

<213> Artificial

<220><223> Target sequence

<400> 113

gctagatgct ggccgtggat gcctcataga atcctcccaa caaccgatga aatgactact 60

gtcattcagc ccaataccca gacgagaaag 90

<210> 114

<211> 90
 <212> DNA

<213> Artificial

<220><223> Target sequence

<400> 114

acaggtttca agcttggcag tctgactaca gaggccactg gctttacccc tgggtagtc 60
 tgcctctgta ggattggggg cacgtaattt 90

<210> 115

<211> 90
 <212> DNA

<213> Artificial

<220><223> Target sequence

<400> 115

ttagtctgcc tctgtaggat tgggggcacg taattttgct gtttaaggtc tcatttgct 60
 tcttagagat cacaagccaa agctttttat 90

<210> 116

<211> 90
 <212> DNA

<213>

> Artificial

<220><223> Target sequence

<400> 116

ggaagcccag agggcatctt gtggetcggg agtagctctc tgctacctc tcagctctgc 60
 tgacaatact tgagattttc agatgtcacc 90

<210> 117

<211> 90
 <212> DNA

<213> Artificial

<220><223> Target sequence

<400> 117

tcagctctgc tgacaatact tgagattttc agatgtcacc aaccagcaag agagcttgat 60
 atgactgtat atagtatagt cataaagaac 90

<210> 118

<211> 90

<212> DNA

<213> Artificial

<220><223> Target sequence

<400> 118

cataaagaac ctgaacttga ccatatactt atgtcatgtg gaaatcttct catagcttca 60

gatagattat atctggagtg aagaatcctg 90

<210> 119

<211> 90

<212> DNA

<213> Artificial

<220><223> Target sequence

<400> 119

gtggaaaatt tctcatagct tcagatagat tatatctgga gtgagcaatc ctgccaccta 60

tgtatctggc atagtgtgag tcttcataaa 90

<210> 120

<211> 90

<212> DNA

<213> Artificial

<220

><223> Target sequence

<400> 120

ggtttgaagg gcaacaaaat agtgaacaga gtgaaaatcc ccacctagat cctgggtcca 60

gaaaaagatg gaaaacctgt ttagctcacc 90

<210> 121

<211> 30

<212> DNA

<213> Artificial

<220><223> Target sequence

<400> 121

ggccactagg gacaaaattg gtgacagaaa 30

<210> 122

<211> 50

<212> DNA

<213> Artificial

<220><223> Target sequence
 <400> 122
 cccacagtgg ggccactagg gacaaaattg gtgacagaaa agcccatcc 50

<210> 123
 <211> 70
 <212> DNA
 <213> Artificial
 <220><223> Target sequence
 <400> 123
 tcccctccac cccacagtgg ggccactagg gacaaaattg gtgacagaaa agcccatcc 60
 ttaggcctcc 70

<210> 124
 <211> 90
 <212> DNA
 <213> Artificial
 <220><223> Target sequence
 <400> 124
 cttttatctg tcccctccac cccacagtgg ggccactagg gacaaaattg gtgacagaaa 60
 agcccatcc ttaggcctcc tccttctag 90

<210>
 > 125
 <211> 110
 <212> DNA
 <213> Artificial
 <220><223> Target sequence
 <400> 125
 gttctgggta cttttatctg tcccctccac cccacagtgg ggccactagg gacaaaattg 60
 gtgacagaaa agcccatcc ttaggcctcc tccttctag tctctgata 110

<210> 126
 <211> 30
 <212> DNA
 <213> Artificial
 <220><223> Target sequence
 <400> 126

tttctgtcac caatggtgtc cctagtggcc 30

<210> 127

<211> 50

<212> DNA

<213> Artificial

<220><223> Target sequence

<400> 127

ggatggggct tttctgtcac caatggtgtc cctagtggcc cactgtggg 50

<210> 128

<211> 70

<212> DNA

<213> Artificial

<220><223> Target sequence

<400> 128

ggaggcctaa ggatggggct tttctgtcac caatggtgtc cctagtggcc cactgtggg 60

gtggagggga 70

<210> 129

<211> 90

<212> DNA

<213> Artificial

<220><223> Target sequence

<400> 129

ctaggaagga ggaggcctaa ggatggggct tttctgtcac caatggtgtc cctagtggcc 60

ccactgtggg gtggagggga cagataaaag 90

<210> 130

<211> 110

<212> DNA

<213> Artificial

<220><223> Target sequence

<400> 130

tatcaggaga ctaggaagga ggaggcctaa ggatggggct tttctgtcac caatggtgtc 60

cctagtggcc ccactgtggg gtggagggga cagataaaag taccagaac 110

<210> 131

<211> 70
 <212> DNA
 <213> Artificial
 <220><223> Target sequence
 <400> 131
 ttctagtaac cactcaggac aggggggttc agcccaaaaa ttcacaagaa agtggggacc 60
 catgggaaat 70

<210> 132
 <211> 70
 <212> DNA
 <213> Artificial
 <220><223> Target sequence
 <400> 132
 cagcaagtca gcagcacagc gtgtgtgact ccgagggtgc tccgctagcc cacattgccc 60
 tctgggggtg 70

<210> 133
 <211> 70
 <212> DNA
 <213> Artificial
 <220><223> Target sequence
 <400> 133
 gtcagactgc caagcttgaa acctgtctta ccctctactt tctcgtctgg gtattgggct 60
 gaatgacagt 70

<210>
 > 134
 <211> 70
 <212> DNA
 <213> Artificial
 <220><223> Target sequence
 <400> 134
 cagagctgag aagacagcag agagctactc ccgaagcaca agatgccctc tgggcttccg 60
 tgaccttgge 70

<210> 135
 <211> 70

<212> DNA
 <213> Artificial
 <220><223> Target sequence
 <400> 135
 ctgacaatac ttgagatfff cagatgtcac caacgaccaa gagagcttga tatgactgta 60
 tatagtatag 70
 <210> 136

 <211> 70
 <212> DNA
 <213> Artificial
 <220><223> Target sequence
 <400> 136
 cagatacata ggtggcagga ttcttcactc cagacttaat ctatctgaag ctatgagaaa 60
 ttttccacat 70
 <210> 137
 <211> 70
 <212> DNA
 <213> Artificial
 <220><223> Target sequence
 <400> 137
 tatatgattg atttgcacag ctcatctggc cagataagct gagacatccg ttcccctaca 60
 agaaactctc 70
 <210> 138
 <211>
 > 70
 <212> DNA
 <213> Artificial
 <220><223> Target sequence
 <400> 138
 atctggccag aagagctgag acatccgttc cccttgaaga aactctcccc ggtaagtaac 60
 ctctcagctg 70
 <210> 139
 <211> 70
 <212> DNA

<213> Artificial
 <220><223> Target sequence
 <400> 139
 aggcattca ctggagaggg tttagtctc cttaaagaa gataagattt caagaggaa 60
 gctaagactc 70
 <210> 140
 <211> 70

 <212> DNA
 <213> Artificial
 <220><223> Target sequence
 <400> 140
 ataataaat aaaaaatgtt tcgtctgcca ccactaatga atgcatgca ttctgggtag 60
 tttagtctta 70
 <210> 141
 <211> 70
 <212> DNA
 <213> Artificial
 <220><223> Target sequence
 <400> 141
 tttataaagt cctagaatgt atttagttgc cctcgttgaa tgcaaactgt ttataacatc 60
 aataggtttt 70
 <210> 142
 <211> 70
 <212>
 > DNA
 <213> Artificial
 <220><223> Target sequence
 <400> 142
 gctcaactg gccattctg acctgttttt ccttcccact gtcaccttct gggctcacta 60
 tgctgccgcc 70
 <210> 143
 <211> 70
 <212> DNA
 <213> Artificial

<220><223> Target sequence
 <400> 143
 ttttaaagca aacacagcat ggacgacagc caggctccta tcgattgtca ggaggatgat 60
 gaagaagatt 70
 <210> 144
 <211> 70
 <212> DNA

 <213> Artificial
 <220><223> Target sequence
 <400> 144
 gcttgtcatg gtcactgtct actcgggaat cctaattact ctgcttcggt gtcgaaatga 60
 gaagaagagg 70
 <210> 145
 <211> 70
 <212> DNA
 <213> Artificial
 <220><223> Target sequence
 <400> 145
 atactgcccc cgcgaggcca cattggcaaa ccagcttggg tgttcaactt ccagacttgg 60
 ccatggagaa 70
 <210> 146
 <211> 90
 <212> DNA
 <213>
 > Artificial
 <220><223> Target sequence
 <400> 146
 ctgaagaatt tcccatgggt cccactttc ttgtgaatcc ttggagtga cccccctgtc 60
 ctgagtggtt actagaacac acctctggac 90
 <210> 147
 <211> 90
 <212> DNA
 <213> Artificial
 <220><223> Target sequence

<400> 147
 tggaagtatc ttgccgaggt cacacagcaa gtcagcagca cagccagtgt gactccgagc 60
 ctgctccgct agccacatt gccctctggg 90

<210> 148
 <211> 90
 <212> DNA
 <213> Artificial

<220><223> Target sequence

<400> 148
 ctactgtcat tcagcccaat acccagacga gaaagctgag ggtataacag gtttcaagct 60
 tggcagtctg actacagagg ccaactggctt 90

<210> 149
 <211> 90
 <212> DNA
 <213> Artificial

<220><223> Target sequence

<400> 149
 ggaagcccag agggcatctt gtggctcggg agtagctctc tgctaccttc tcagctctgc 60
 tgacaatact tgagattttc agatgtcacc 90

<210> 150
 <211> 90
 <212> DNA
 <213> Artificial

<220>

><223> Target sequence

<400> 150
 tcagctctgc tgacaatact tgagattttc agatgtcacc aacgcccaag agagcttgat 60
 atgactgtat atagtatagt cataaagaac 90

<210> 151
 <211> 90
 <212> DNA
 <213> Artificial

<220><223> Target sequence

<400> 151

gtggaaaatt tcicatagct tcagatagat tataatctgga gtgagcaatc ctgccaccta 60
 tgtatctggc atagtgtgag tcctcataaa 90
 <210> 152
 <211> 90
 <212> DNA
 <213> Artificial
 <220><223>
 > Target sequence
 <400> 152
 gaaacagcat ttctacttt tatactgtct atatgattga tttggtcagc tcatctggcc 60
 agaagagctg agacatccgt tcccctaaa 90
 <210> 153
 <211> 90
 <212> DNA
 <213> Artificial
 <220><223> Target sequence
 <400> 153
 ttgatttga cagctcatct ggccagaaga gctgagacat ccgtatccct acaagaaact 60
 ctccccggtg agtaacctct cagctgcttg 90
 <210> 154
 <211> 90
 <212> DNA
 <213> Artificial
 <220><223> Target sequence
 <400> 154
 ggagagggtt tagttctcct tagcagaaga taagatttca agatgagagc taagactcat 60
 ctctctgcaa atctttcttt tgagaggtaa 90
 <210> 155
 <211> 90
 <212> DNA
 <213> Artificial
 <220><223> Target sequence
 <400> 155
 taatataata aaaaatgttt cgtctgccac cacagatgaa tgtcgagcat tctgggtagt 60

ttagtcttat aaccagctgt ctgectagt 90

<210> 156

<211> 90

<212> DNA

<213> Artificial

<220><223> Target sequence

<400>

> 156

ttaaaaacct attgatgtat aaaacagttt gcattcatgg agggtgacta aatacattct 60

aggactttat aaaagatcac tttttattta 90

<210> 157

<211> 90

<212> DNA

<213> Artificial

<220><223> Target sequence

<400> 157

gacatctacc tgcTcaacct ggccatctct gacctgtttt tcctatttac tgtccccttc 60

tgggctcact atgctgccgc ccagtgggac 90

<210> 158

<211> 90

<212> DNA

<213> Artificial

<220><223> Target sequence

<400> 158

tcacTctct gacaatcgat aggtacctgg ctgctgtcca tgctacgttt gctttaaaag 60

ccaggacggT cacctttggg gtggTgacaa 90

<210> 159

<211> 90

<212> DNA

<213> Artificial

<220><223> Target sequence

<400> 159

ggctggTctt gccgctgctt gTcatggTca tctgctactc gggagaccta aaaactctgc 60

ttcggtgtcg aaatgagaag aagaggcaca 90

<210> 160
 <211> 90
 <212> DNA
 <213> Artificial
 <220><223> Target sequence
 <400> 160
 ggcaagcctt gggtcatact gccccgcga ggccacattg gcaagtcagc aagggtgttc 60

 aacttccaga cttggccatg gagaagacat 90
 <210> 161
 <211> 59
 <212> DNA
 <213> Artificial
 <220><223> PCR primer
 <400> 161
 acactctttc cctacacgac gctcttccga tctcgtgatt ttgcagtgtg cgttactcc 59
 <210> 162
 <211> 59
 <212> DNA
 <213> Artificial
 <220><223> PCR primer
 <400> 162
 acactctttc cctacacgac gctcttccga tctacatcgt ttgcagtgtg cgttactcc 59
 <210> 163
 <211> 59
 <212> DNA
 <213> Artificial
 <220><223> PCR primer
 <400> 163

 acactctttc cctacacgac gctcttccga tctgcctaatt ttgcagtgtg cgttactcc 59
 <210> 164
 <211> 59
 <212> DNA
 <213> Artificial

<220><223> PCR primer
 <400> 164
 acactctttc cctacacgac gctcttccga tcttggatcat ttgcagtgtg cgttactcc 59
 <210> 165
 <211> 54
 <212> DNA
 <213> Artificial
 <220><223> PCR primer
 <400> 165
 ctcggcattc ctgctgaacc gctcttccga tctccaagca actaagtcac agca 54
 <210> 166
 <211> 58
 <212> DNA
 <213> Artificial
 <220><223> PCR primer
 <400> 166

 acactctttc cctacacgac gctcttccga tctcgtgata tgaggaaatg gaagcttg 58
 <210> 167
 <211> 58
 <212> DNA
 <213> Artificial
 <220><223> PCR primer
 <400> 167
 acactctttc cctacacgac gctcttccga tctacatcga tgaggaaatg gaagcttg 58
 <210> 168
 <211> 58
 <212> DNA
 <213> Artificial
 <220><223> PCR primer
 <400> 168
 acactctttc cctacacgac gctcttccga tctgcctaaa tgaggaaatg gaagcttg 58
 <210> 169
 <211> 58
 <212> DNA

<213> Artificial
 <220><223> PCR primer
 <400> 169

acactctttc cctacacgac gctcttccga tcttggtaa tgaggaaatg gaagcttg 58
 <210> 170
 <211> 51
 <212> DNA
 <213> Artificial
 <220><223> PCR primer
 <400> 170

ctcggcattc ctgctgaacc gctcttccga tctcattagg gtattggagg a 51
 <210> 171
 <211> 58
 <212> DNA
 <213> Artificial
 <220><223> PCR primer
 <400> 171

acactctttc cctacacgac gctcttccga tctcgtgata atcctcccaa caactcat 58
 <210> 172
 <211> 58
 <212> DNA
 <213> Artificial
 <220><223> PCR primer
 <400> 172

acactctttc cctacacgac gctcttccga tctacatcga atcctcccaa caactcat 58
 <210> 173
 <211> 58
 <212> DNA
 <213> Artificial
 <220><223> PCR primer
 <400> 173

acactctttc cctacacgac gctcttccga tctgcctaaa atcctcccaa caactcat 58
 <210> 174

<211> 58
 <212> DNA
 <213> Artificial
 <220><223> PCR primer
 <400> 174
 acactctttc cctacacgac gctcttccga tcttggtcaa atcctcccaa caactcat 58
 <210> 175
 <211> 52
 <212> DNA
 <213> Artificial
 <220><223> PCR primer
 <400> 175

 ctcggcattc ctgctgaacc gctcttccga tctccaatc ctacagaggc ag 52
 <210> 176
 <211> 58
 <212> DNA
 <213> Artificial
 <220><223> PCR primer
 <400> 176
 acactctttc cctacacgac gctcttccga tctcgtgata agccaaagct ttttattc 58
 <210> 177
 <211> 58
 <212> DNA
 <213> Artificial
 <220><223> PCR primer
 <400> 177
 acactctttc cctacacgac gctcttccga tctacatcga agccaaagct ttttattc 58
 <210> 178
 <211> 58
 <212> DNA
 <213> Artificial
 <220><223> PCR primer
 <400> 178

acactctttc cctacacgac gctcttccga tctgcctaaa agccaaagct ttttattc 58
 <210> 179
 <211> 58
 <212> DNA
 <213> Artificial
 <220><223> PCR primer
 <400> 179

acactctttc cctacacgac gctcttccga tcttggtcaa agccaaagct ttttattc 58
 <210> 180
 <211> 53
 <212> DNA
 <213> Artificial
 <220><223> PCR primer
 <400> 180

acactctttc cctacacgac gctcttccga tctaagccaa agctttttat tct 53
 <210> 181
 <211> 59
 <212> DNA
 <213> Artificial
 <220><223> PCR primer
 <400> 181

acactctttc cctacacgac gctcttccga tctcgtgata tcttgtggct cgggagtag 59
 <210> 182
 <211> 59
 <212> DNA
 <213> Artificial
 <220><223> PCR primer
 <400> 182

acactctttc cctacacgac gctcttccga tctacatcga tcttgtggct cgggagtag 59
 <210> 183
 <211> 53
 <212> DNA
 <213> Artificial
 <220><223> PCR primer

<400> 183
ctcggcattc ctgctgaacc gctcttcgga tcttggcagg attcttact cca 53

<210> 184
<211> 59
<212> DNA
<213> Artificial
<220><223> PCR primer
<400> 184

acactcttc cctacacgac gctcttcgga tctcgtgatc tattttgtg cccttcaaa 59

<210> 185
<211> 59
<212> DNA
<213> Artificial
<220><223> PCR primer
<400> 185

acactcttc cctacacgac gctcttcgga tctacatgc tattttgtg cccttcaaa 59

<210> 186
<211> 55
<212> DNA
<213> Artificial
<220><223> PCR primer
<400> 186

ctcggcattc ctgctgaacc gctcttcgga tctaacctga acttgacat atact 55

<210> 187
<211> 59
<212> DNA
<213> Artificial
<220><223> PCR primer
<400> 187

acactcttc cctacacgac gctcttcgga tctcgtgatc agctgagagg ttacttacc 59

<210> 188
<211> 59
<212> DNA

<213> Artificial
 <220><223> PCR primer
 <400> 188
 acactctttc cctacacgac gctcttccga tctacatcgc agctgagagg ttacttacc 59
 <210> 189
 <211> 53
 <212> DNA
 <213> Artificial
 <220><223> PCR primer
 <400> 189
 ctcggcattc ctgctgaacc gctcttccga tctaattgatt taactccacc etc 53
 <210> 190
 <211> 60
 <212> DNA
 <213> Artificial
 <220><223> PCR primer
 <400> 190

 acactctttc cctacacgac gctcttccga tctcgtgata ctccaccctc cttcaaaaga 60
 <210> 191
 <211> 60
 <212> DNA
 <213> Artificial
 <220><223> PCR primer
 <400> 191
 acactctttc cctacacgac gctcttccga tctacatcga ctccaccctc cttcaaaaga 60
 <210> 192
 <211> 52
 <212> DNA
 <213> Artificial
 <220><223> PCR primer
 <400> 192
 ctcggcattc ctgctgaacc gctcttccga tcttgggtgtt tgccaaatgt ct 52
 <210> 193
 <211> 59

<212> DNA
 <213> Artificial
 <220><223> PCR primer
 <400> 193

 acactctttc cctacacgac gctcttccga tctcgtgatg ggcacatatt cagaaggca 59
 <210> 194
 <211> 59
 <212> DNA
 <213> Artificial
 <220><223> PCR primer
 <400> 194

 acactctttc cctacacgac gctcttccga tctacatcgg ggcacatatt cagaaggca 59
 <210> 195
 <211> 56
 <212> DNA
 <213> Artificial
 <220><223> PCR primer
 <400> 195

 ctcggcattc ctgctgaacc gctcttccga tctagtgaaa gactttaag ggagca 56
 <210> 196
 <211> 59
 <212> DNA
 <213> Artificial
 <220><223> PCR primer
 <400> 196

 acactctttc cctacacgac gctcttccga tctcgtgatc acaattaaga gttgtcata 59
 <210> 197
 <211> 59
 <212> DNA
 <213> Artificial
 <220><223> PCR primer
 <400> 197

 acactctttc cctacacgac gctcttccga tctacatcgc acaattaaga gttgtcata 59

<210> 198
 <211> 53
 <212> DNA
 <213> Artificial
 <220><223> PCR primer
 <400> 198
 ctcggcattc ctgctgaacc gctcttccga tctctcagct agagcagctg aac 53
 <210> 199
 <211> 51
 <212> DNA
 <213> Artificial
 <220><223> PCR primer
 <400> 199

 ctcggcattc ctgctgaacc gctcttccga tctgacactt gataatccat c 51
 <210> 200
 <211> 60
 <212> DNA
 <213> Artificial
 <220><223> PCR primer
 <400> 200
 acactctttc cctacacgac gctcttccga tctacatcgt caatgtagac atctatgtag 60
 <210> 201
 <211> 60
 <212> DNA
 <213> Artificial
 <220><223> PCR primer
 <400> 201
 acactctttc cctacacgac gctcttccga tctcgtgatt caatgtagac atctatgtag 60
 <210> 202
 <211> 59
 <212> DNA
 <213> Artificial
 <220><223> PCR primer

<400> 202

acactctttc cctacacgac gctcttccga tctcgtgata ctgcaaaaagg ctgaagagc 59

<210> 203

<211> 59

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 203

acactctttc cctacacgac gctcttccga tctacatcga ctgcaaaaagg ctgaagagc 59

<210> 204

<211> 59

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 204

acactctttc cctacacgac gctcttccga tctgcctaaa ctgcaaaaagg ctgaagagc 59

<210> 205

<211> 59

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 205

acactctttc cctacacgac gctcttccga tcttgggtcaa ctgcaaaaagg ctgaagagc 59

<210> 206

<211> 57

<212> DNA

<213> Artificial

<220><223> PCR primer

<400> 206

ctcggcattc ctgctgaacc gctcttccga tctgcctata aaatagagcc ctgtcaa 57

<210> 207

<211> 60

<212> DNA

<213> Artificial
 <220><223> PCR primer
 <400> 207
 acactctttc cctacacgac gctcttccga tctcgtgac tctatttat aggtctcttc 60
 <210> 208
 <211> 60
 <212> DNA
 <213> Artificial
 <220><223> PCR primer
 <400> 208

 acactctttc cctacacgac gctcttccga tctacatgc tctatttat aggtctcttc 60
 <210> 209
 <211> 52
 <212> DNA
 <213> Artificial
 <220><223> PCR primer
 <400> 209
 ctcgcatc ctgctgaacc gctcttccga tctagccacc acccaagtga tc 52
 <210> 210
 <211> 60
 <212> DNA
 <213> Artificial
 <220><223> PCR primer
 <400> 210
 acactctttc cctacacgac gctcttccga tctacatgt tccagacatt aaagatagtc 60
 <210> 211
 <211> 60
 <212> DNA
 <213> Artificial
 <220><223> PCR primer
 <400> 211

 acactctttc cctacacgac gctcttccga tctcgtgatt tccagacatt aaagatagtc 60
 <210> 212

<211> 54
 <212> DNA
 <213> Artificial
 <220><223> PCR primer
 <400> 212
 ctcggcattc ctgctgaacc gctcttccga tctaatcatg atggtgaaga taag 54
 <210> 213
 <211> 60
 <212> DNA
 <213> Artificial
 <220><223> PCR primer
 <400> 213
 acactctttc cctacacgac gctcttccga tctcgtgatc cggcagagac aacattaaa 60
 <210> 214
 <211> 54
 <212> DNA
 <213> Artificial
 <220><223> PCR primer
 <400> 214

 acactctttc cctacacgac gctcttccga tctccggcag agacaaacat taaa 54
 <210> 215
 <211> 53
 <212> DNA
 <213> Artificial
 <220><223> PCR primer
 <400> 215
 ctcggcattc ctgctgaacc gctcttccga tctagctagg aagccatggc aag 53
 <210> 216
 <211> 47
 <212> DNA
 <213> Artificial
 <220><223> PCR primer
 <400> 216
 aatgataccg cgaccaccga gatctacaact ctttcctac acgacgc 47

<210> 217
 <211> 50
 <212> DNA
 <213> Artificial
 <220><223> PCR primer
 <400> 217

 caagcagaag acggcatacg agatcggctct cggcattcct gctgaaccgc 50
 <210> 218
 <211> 58
 <212> DNA
 <213> Artificial
 <220><223> PCR primer
 <400> 218
 acactctttc cctacacgac gctcttccga tctagtgcac agtatgtgct agatgctg 58
 <210> 219
 <211> 59
 <212> DNA
 <213> Artificial
 <220><223> PCR primer
 <400> 219
 gtgactggag ttcagacgtg tgctcttccg atcttgatct ctaagaagc aatgagac 59
 <210> 220
 <211> 64
 <212> DNA
 <213> Artificial
 <220><223> PCR primer
 <220><221
 > misc_feature
 <222> (25)..(30)
 <223> n is a, c, g, or t
 <400> 220
 caagcagaag acggcatacg agatnnnnn gtgactggag ttcagacgtg tgctcttccg 60
 atct 64
 <210> 221

<211> 58
 <212> DNA
 <213> Artificial
 <220><223> PCR primer
 <400> 221
 aatgatacgg cgaccaccga gatctacaact ctttccctac acgacgtctt tccgatct 58
 <210> 222
 <211> 34
 <212> PRT
 <213> Artificial
 <220><223> Deduced TALE template amino acid sequence

 <220><221> misc_feature
 <222> (12)..(13)
 <223> Xaa can be any naturally occurring amino acid
 <400> 222
 Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Xaa Xaa Gly Gly Lys
 1 5 10 15
 Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala
 20 25 30
 His Gly

 <210> 223
 <211> 246
 <212> DNA
 <213> Artificial
 <220><223> TALE dimer nucleotide sequence
 <400> 223

 cgattactgc ctgtcttatg tcaggcccat ggccttactc ctgagcaagt cgtagctatc 60
 gccagcaaca taggtgggaa acaggccctg gaaaccgtac aacgtctcct cccagtactt 120
 tgtcaagcac acgggttgac accggaacaa gtggtggcga ttgctgccca cgatggaggc 180
 aagcaggcac tggagaccgt ccaacggctt cttccggttc tttgccaggc tcatgggctc 240
 acgcca 246
 <210> 224

<211> 82

<212> PRT

<213> Artificial

<220><223> Deduced TALE dimer amino acid sequence

<400> 224

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln

1 5 10 15
 Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr

 20 25 30
 Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro

 35 40 45
 Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu

 50 55 60
 Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu

65 70 75 80
 Thr Pro