



(19) 대한민국특허청(KR)  
(12) 등록특허공보(B1)

(45) 공고일자 2021년10월01일  
(11) 등록번호 10-2307280  
(24) 등록일자 2021년09월24일

(51) 국제특허분류(Int. Cl.)  
C12N 15/85 (2006.01) A61K 48/00 (2006.01)  
C07K 14/005 (2006.01) C12N 15/113 (2010.01)  
C12N 15/62 (2006.01) C12N 5/074 (2010.01)  
C12N 5/077 (2010.01) C12N 9/22 (2006.01)  
(52) CPC특허분류  
C12N 15/85 (2013.01)  
A61K 48/00 (2013.01)  
(21) 출원번호 10-2016-7000166  
(22) 출원일자(국제) 2014년06월05일  
심사청구일자 2019년06월05일  
(85) 번역문제출일자 2016년01월05일  
(65) 공개번호 10-2016-0023765  
(43) 공개일자 2016년03월03일  
(86) 국제출원번호 PCT/US2014/041190  
(87) 국제공개번호 WO 2014/197748  
국제공개일자 2014년12월11일  
(30) 우선권주장  
61/831,481 2013년06월05일 미국(US)  
(뒷면에 계속)  
(56) 선행기술조사문헌  
Mol Ther., 21(9):1718-1726(2013.6.4.)\*  
Nat Biotechnol., 31(3):230-232(2013.1.29.)\*  
Science., 339(6121):823-826(2013.2.15.)\*  
\*는 심사관에 의하여 인용된 문헌

(73) 특허권자  
듀크 유니버시티  
미국, 노스캐롤라이나 27705, 더럼, 스위트 406,  
2812 에르윈 로드  
(72) 발명자  
거스배치, 찰스, 에이.  
미국 27712 노스캐롤라이나주 더럼 월더니스 로드  
3220  
힐튼, 아이작, 비.  
미국 27713 노스캐롤라이나주 더럼 그레이 엘름  
트레일 118  
(뒷면에 계속)  
(74) 대리인  
양영준, 김영

전체 청구항 수 : 총 19 항

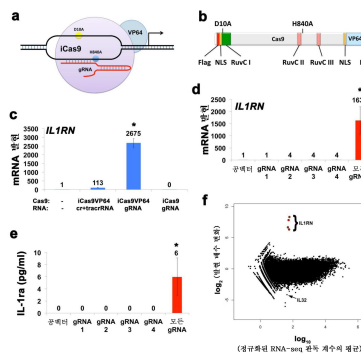
심사관 : 최성호

(54) 발명의 명칭 RNA-가이드 유전자 편집 및 유전자 조절

(57) 요약

클러스터링된 규칙적 간격의 짧은 회문식 반복부 (CRISPR)/CRISPR-연관 (Cas) 9-기반 시스템 관련 조성물, 및 유전자 발현 변경 및 게놈 조작을 위한 상기 CRISPR/Cas9-기반 시스템 관련 조성물의 사용 방법이 본원에 개시된다. 또한, 근육, 예컨대 골격근 및 심장 근육에서 유전자 발현 변경 및 게놈 조작을 위한 조성물 및 상기 조성물의 사용 방법이 개시된다.

대표도 - 도1



(52) CPC특허분류

**C07K 14/005** (2013.01)  
**C12N 15/113** (2013.01)  
**C12N 15/62** (2013.01)  
**C12N 5/0656** (2013.01)  
**C12N 5/0696** (2013.01)  
**C12N 9/22** (2013.01)  
**C07K 2319/00** (2013.01)

(72) 발명자

**페레즈-피네라, 파블로**

미국 01902 매사추세츠주 린 세인트 클레어 스트리트 #3 9

**카바디, 아미, 엠.**

미국 27705 노스캐롤라이나주 더럼 포코노 드라이브 16

**타코어, 프래딕샤, 아이.**

미국 27705 노스캐롤라이나주 더럼 레밍턴 서클 1004

**아우스터라웃, 데이빗, 쥐.**

미국 27703 노스캐롤라이나주 더럼 유닛 105 티더블유 알렉산더 드라이브 1725

**블랙, 조슈아, 비.**

미국 27708 노스캐롤라이나주 더럼 로즈 가든 레인 13305

(30) 우선권주장

61/839,127	2013년06월25일	미국(US)
61/904,911	2013년11월15일	미국(US)
61/967,466	2014년03월19일	미국(US)
61/981,575	2014년04월18일	미국(US)

## 명세서

### 청구범위

#### 청구항 1

Cas9 및 적어도 1개의 가이드 RNA (gRNA)를 포함하는, 디스트로핀 유전자에 결합하는 DNA 표적화 시스템으로서,

a) 적어도 1개의 gRNA가 서열 ggggctccaccctcacgagt, gcacaaaagtcaaatcggaa, gatttcaatataagattcgg, gtgagggtccaccctcacga, gaaggattgagggtccaccc, ggctccaccctcacgagtggg 및 gtgagggtccaccctcacga로부터 선택된 폴리뉴클레오티드 서열을 표적화하거나;

b) 적어도 1개의 gRNA가 서열 gtacacttttcaaatgctt를 포함하는 폴리뉴클레오티드 서열을 표적화하거나;

c) 적어도 1개의 gRNA가 서열 gtcgctcactcaccctgcaa를 포함하는 폴리뉴클레오티드 서열을 표적화하거나;

d) 적어도 1개의 gRNA가 서열 ggatctgtcaaatcgctgc를 포함하는 폴리뉴클레오티드 서열을 표적화하거나;

e) 적어도 1개의 gRNA가 서열 gattggctttgatttccta, gtgtagagtaagtcagccta 및 gcagttgcctaagaactggt로부터 선택된 폴리뉴클레오티드 서열을 표적화하거나;

f) 적어도 1개의 gRNA가 서열 gattggctttgatttccta, gcagttgcctaagaactggt, gcctactcagactgttactc, ggggctccaccctcacgagt, gccttctttatccctatcg 및 gttggacagaacttaccgac로부터 선택된 폴리뉴클레오티드 서열을 표적화하거나;

g) 적어도 1개의 gRNA가 gattggctttgatttccta, gcctactcagactgttactc, gcagttgcctaagaactggt, ggggctccaccctcacgagt, gccttctttatccctatcg, gtgtagagtaagtcagccta, gttggacagaacttaccgac, gtttgcctcgtataaaacg, gtctgaggatggggccgcaa, ggatctgtcaaatcgctgc, gccaggatggcatgtggcag, gctgaatctgcgggtggcagg, gttcttttcttctttagcc, ggaaaagcttgagcaagtca, ggaagagtggccctgcgcc, gacaaatctccagtggataa, gtgtttctcaggtaaagctc, ggaaggaccatttgacgtta, gaactgctatttcagttcc, gccagccactcagccagtga, ggtatgcttttctgttaaag, gctcctggactgaccactat, ggaacagaggcgtccccagt, ggaggctagaacaatcatta, gacaagaacaccttcagaac, ggggtttctgtgattttctt, gggccaaagacctccgcccag, gttggagaagcattcataaa, gtcgctcactcaccctgcaa, gaaaagagctgatgaaacaa, gtacacttttcaaatgctt, ggagatgatcatcaagcaga, gctttgaaagagcaataaaa, gcacaaaagtcaaatcggaa, gatttcaatataagattcgg, gcttaagcaatcccgaactc, gaggccaaacctcggcttac, gttcgaaaatttcaggttaag, ggcagaacaggagataacag, ggcgccctcgcccttctct, gtatgtatcgtggatacag, gtacagccctcggtgtatat, gggaaggaattaagcccgaa, gataaagtcagtgatc, gaaaaccagagcttcggtca, ggagtcttctgggcaggcttaaggctaac, gtcgggtgagcatgtcttaatctacctga, ggtgtcaccagagtaacagt, gtgatcatcaagcagaaggt, gaacttcgaaaatttcaggta, ggaaactcatcaaatatgcgt, gtcatttacactaacacgcat, ggaatgaaactcatcaaatat, gtcacatcaatatttgaagga, gtgtttcataggaaaaatag, gaattggaaaatgtgatggga, gatgatcatcaagcagaaggt, gagatgatcatcaagcagaag, gcatttttctcataccttct, gtcctactcagactgttactc, gacagggtgtgtcaccagagt, gttatcatttttctcatacc, gtgcctaagaactggtggga, gaaacagttgcctaagaactg, gtttccaccagttcttaggc, gtggctttgatttccttaggg, tagggaaatcaaagccaatg, ggaccctagggaaatcaaagc, gtgagggtccaccctcacga, gaaggattgagggtccaccc, ggctccaccctcacgagtggg, gtatccctatcgaggaaacc, ggataaagaaggcctatttca, gaggccttctttatccctat, gtgagggtccaccctcacga 및 ggataaagaaggcctatttca로부터 선택된 폴리뉴클레오티드 서열을 표적화하는 것인

DNA 표적화 시스템.

#### 청구항 2

제1항에 있어서, 적어도 1개의 gRNA가 디스트로핀 유전자의 인트론 또는 디스트로핀 유전자의 엑손을 표적화하는 것인 DNA 표적화 시스템.

#### 청구항 3

제1항의 DNA 표적화 시스템을 코딩하는 단리된 폴리뉴클레오티드.

#### 청구항 4

제3항의 단리된 폴리뉴클레오티드를 포함하는 벡터.

#### 청구항 5

제3항의 단리된 폴리뉴클레오티드를 포함하는 세포.

#### 청구항 6

돌연변이체 디스트로핀 유전자를 가지며, 아래 질환의 치료를 필요로 하는 대상체를 치료하기 위해 사용되는, 제1항의 DNA 표적화 시스템을 포함하는 조성물로서,

대상체는 변성 또는 약화를 유발하는 골격근 또는 심장 근육 병태를 앓고 있거나, 뒤시엔느 근육 이영양증 (DMD), 혈우병, 낭성 섬유증, 헌팅톤 무도병, 가족성 고콜레스테롤혈증 (LDL 수용체 결함), 간모세포종, 윌슨병, 선천성 간 포르피린증, 간 대사의 유전성 장애, 레쉬 니한 증후군, 겸상 적혈구성 빈혈, 지중해 빈혈, 색소성 건피증, 판코니 빈혈, 색소성 망막염, 모세혈관확장성 운동실조, 블루움 증후군, 망막모세포종 또는 테이-삭스병으로부터 선택된 유전 질환을 앓고 있는 것인 조성물.

#### 청구항 7

제6항에 있어서, 대상체가 뒤시엔느 근육 이영양증을 앓고 있는 것인 조성물.

#### 청구항 8

돌연변이체 디스트로핀 유전자를 함유하는 세포에서 돌연변이체 디스트로핀 유전자를 교정하기 위해 사용되는, 제1항의 DNA 표적화 시스템을 포함하는 조성물.

#### 청구항 9

제8항에 있어서, 돌연변이체 디스트로핀 유전자가 조기 정지 코돈, 유전자 결실을 통해 파괴된 리딩 프레임, 이상 스플라이스 수용자 부위 또는 이상 스플라이스 공여자 부위를 포함하고, 표적 영역이 조기 정지 코돈, 파괴된 리딩 프레임, 이상 스플라이스 수용자 부위 또는 이상 스플라이스 공여자 부위의 상류 또는 하류인 조성물.

#### 청구항 10

제8항에 있어서, 돌연변이체 디스트로핀 유전자의 교정이 상동성-지정 복구를 포함하는 것인 조성물.

#### 청구항 11

제10항에 있어서, 공여자 DNA를 추가로 포함하는 조성물.

#### 청구항 12

제8항에 있어서, 돌연변이체 디스트로핀 유전자가 조기 정지 코돈 및 말단절단된 유전자 산물을 유발하는 프레임시프트 돌연변이를 포함하는 것인 조성물.

#### 청구항 13

제12항에 있어서, 돌연변이체 디스트로핀 유전자의 교정이 뉴클레아제 매개된 비-상동 말단 연결을 포함하는 것인 조성물.

#### 청구항 14

제8항에 있어서, 돌연변이체 디스트로핀 유전자의 교정이 조기 정지 코돈의 결실, 파괴된 리딩 프레임의 교정, 또는 스플라이스 수용자 부위의 파괴 또는 스플라이스 공여자 서열의 파괴에 의한 스플라이싱의 조절을 포함하는 것인 조성물.

#### 청구항 15

제14항에 있어서, 돌연변이체 디스트로핀 유전자의 교정이 엑손 45-55 또는 엑손 51의 결실을 포함하는 것인 조



성물.

#### 청구항 16

변형된 아데노-연관 바이러스 (AAV) 벡터 및 제1항의 DNA 표적화 시스템을 코딩하는 뉴클레오티드 서열을 포함하는, 골격근 또는 심장 근육인 대상체의 근육에서의 게놈 편집을 위한 조성물.

#### 청구항 17

제16항에 있어서, 변형된 AAV 벡터가 증진된 심장 근육 및 골격근 조직 형성을 갖는 것인 조성물.

#### 청구항 18

제16항에 있어서, 대상체가 골격근 병태 또는 유전 질환을 앓고 있는 것인 조성물.

#### 청구항 19

제16항에 있어서, 대상체가 뒤시엔느 근육 이영양증을 앓고 있는 것인 조성물.

#### 청구항 20

삭제

#### 청구항 21

삭제

#### 청구항 22

삭제

#### 청구항 23

삭제

#### 청구항 24

삭제

#### 청구항 25

삭제

#### 청구항 26

삭제

#### 청구항 27

삭제

#### 청구항 28

삭제

#### 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

## 발명의 설명

### 기술 분야

- [0001] 관련 출원에 대한 상호 참조
- [0002] 본원은 2013년 6월 5일 출원된 미국 가출원 번호 61/831,481, 2013년 6월 25일 출원된 미국 가출원 번호 61/839,127, 2013년 11월 15일 출원된 미국 가출원 번호 61/904,911, 2014년 3월 19일 출원된 미국 가출원 번호 61/967,466, 및 2014년 4월 18일 출원된 미국 가출원 번호 61/981,575를 우선권 주장하며, 상기 모든 가출원은 그 전문이 본원에 참조로 포함된다.
- [0003] 정부 이권의 진술
- [0004] 본 발명은 NIH에 의해 수여된 연방 승인 번호 DP2-OD008586 및 R01DA036865 및 국립 과학 재단에 의해 수여된 CBET-1151035 하에 정부 지원으로 만들어졌다. 미국 정부는 본 발명에 대한 특정 권리를 갖는다.
- [0005] 기술 분야
- [0006] 본 개시내용은 클러스터링된 규칙적 간격의 짧은 회문식 반복부 (CRISPR)/CRISPR-연관 (Cas) 9-기반 시스템 및 바이러스 전달 시스템을 사용하는 유전자 발현 변경, 게놈 조작 및 유전자의 게놈 변경 분야에 관한 것이다. 본 개시내용은 또한 근육, 예컨대 골격근 및 심장 근육에서의 게놈 조작 및 유전자의 게놈 변경에 관한 것이다.

### 배경 기술

- [0007] 합성 전사 인자는 포유동물 시스템에서 조직 재생의 자극, 약물 스크리닝, 유전적 결합의 보상, 침묵화된 종양 억제자의 활성화, 줄기 세포 분화의 제어, 유전적 스크린의 수행 및 합성 유전자 회로의 생성을 비롯한 많은 상이한 의학적 및 과학적 적용을 위해 유전자 발현을 제어하도록 조작되었다. 이들 전사 인자는 내인성 유전자의 프로모터 또는 인핸서를 표적화할 수 있거나, 또는 트랜스진 조절을 위해 포유동물 게놈에 적교인 서열을 인식하도록 의도적으로 설계될 수 있다. 사용자-정의 서열을 표적화하는 신규 전사 인자를 조작하기 위한 가장 통상적인 전략은 아연 핑거 단백질 및 전사-활성화제 유사 이펙터 (TALE)의 프로그램가능한 DNA-결합 도메인을 기반으로 하였다. 이들 접근법 둘 다는 이들 도메인의 단백질-DNA 상호작용의 원리를 공유한 DNA-결합 특이성을 갖는 새로운 단백질을 조작하는데 적용하는 것을 수반한다. 이들 방법이 많은 적용에서 폭넓게 성공하였지만, 단백질-DNA 상호작용을 조작하는데 필요한 단백질 조작은 힘들고 전문화된 지식을 요구할 수 있다.
- [0008] 추가로, 이들 새로운 단백질이 항상 효과적인 것은 아니다. 그 이유는 아직 알려지지 않지만, 후성적 변형 및 염색질 상태가 게놈 표적 부위에서의 단백질 결합에 미치는 효과와 관련이 있을 수 있다. 또한, 이들 새로운 단백질 뿐만 아니라 다른 성분이 각 세포로 전달되도록 보장하는데 과제가 존재한다. 이들 새로운 단백질 및 그의 다중 성분을 전달하기 위한 기존 방법은 개별 플라스미드 또는 벡터 상에서 세포로의 전달을 포함하며, 이는 카피수의 차이로 인해 각 세포에서 고도로 가변적인 발현 수준을 유도한다. 추가로, 형질감염 후의 유전자 활성화는 플라스미드 DNA의 회색으로 인해 일시적이고, 일시적 유전자 발현은 치료 효과를 유도하기에 충분하지 않을 수 있다. 또한, 이러한 접근법은 용이하게 형질감염되지 않는 세포 유형에는 맞지 않는다. 따라서 이들 새로운 단백질의 또 다른 제한은 전사 활성화의 효력이다.
- [0009] 부위-특이적 뉴클레아제는 표적화된 게놈 유전자좌에 부위-특이적 이중 가닥 파단을 도입하는데 사용될 수 있다. 이러한 DNA 절단은 천연 DNA-복구 기구를 자극하여, 2개의 가능한 복구 경로 중 하나를 유도한다. 공여



자 주형의 부재 하에, 파단은 DNA의 작은 삽입 또는 결실을 유도하는 오류-유발 복구 경로인 비-상동 말단 연결 (NHEJ)에 의해 복구될 것이다. 이러한 방법은 표적화된 유전자 서열의 리딩 프레임을 의도적으로 파괴하거나, 결실시키거나 또는 변경하는데 사용될 수 있다. 그러나, 공여자 주형이 뉴클레아제와 함께 제공된다면, 세포 기구는 상동 재조합에 의해 파단을 복구할 것이며, 이는 DNA 절단의 존재 하에 여러 자릿수 증진된다. 이러한 방법은 표적 부위에서의 DNA 서열에 특정한 변화를 도입하는데 사용될 수 있다. 조작된 뉴클레아제는 다양한 인간 줄기 세포 및 세포주에서의 유전자 편집 및 마우스 간에서의 유전자 편집에 사용되었다. 그러나, 이들 기술의 실시를 위한 주요 한계는 효과적이고 효율적이며 성공적인 게놈 변형을 촉진하는 방식으로 생체내 특정한 조직으로 전달하는 것이다.

[0010] 유전성 유전 질환은 미국에서 소아에 대해 엄청난 피해를 주었다. 이들 질환은 현재 치유법이 없으며 단지 증상을 완화시키기 위한 시도로 관리될 수 있다. 수십년간, 유전자 요법 분야는 이들 질환에 대한 치유법을 약속해왔다. 그러나, 세포 및 환자로의 치료 유전자의 안전하고 효율적인 전달에 관한 기술적 한계는 이러한 접근법을 제한해왔다. 뒤시엔느 근육 이영양증 (DMD)은 가장 통상적인 유전성 단일유전자 질환이며 3500명의 남성 중 1명에서 발생한다. DMD는 디스트로핀 유전자에서의 유전성 또는 자발적 돌연변이의 결과이다. 디스트로핀은 근육 세포 완전성 및 기능의 조절을 담당하는 단백질 복합체의 핵심 성분이다. DMD 환자는 전형적으로 소아기 동안 그들 자신을 신체적으로 지탱할 능력을 상실하고, 10대 동안 진행하여 점점 더 약해지며, 20대에 사망한다. DMD를 위한 현재 실험 유전자 요법 전략은 일시적 유전자 전달 비효율의 반복 투여를 요구하거나 또는 게놈 DNA 내로의 외래 유전 물질의 영구적 통합에 의존한다. 이들 방법 둘 다는 심각한 안전성 문제를 갖는다. 또한, 이들 전략은 크고 복잡한 디스트로핀 유전자 서열을 전달할 수 없다는 것에 의해 제한되어 왔다.

### 발명의 내용

[0011] 본 발명은 2개의 이중 폴리펩티드 도메인을 포함하는 융합 단백질에 관한 것이다. 제1 폴리펩티드 도메인은 클러스터링된 규칙적 간격의 짧은 회문식 반복부 연관 (Cas) 단백질을 포함하고, 제2 폴리펩티드 도메인은 전사 활성화, 전사 억제, 전사 방출 인자 활성화, 히스톤 변형 활성화, 뉴클레아제 활성화, 핵산 회합 활성화, 메틸라제 활성화 및 데메틸라제 활성화로 이루어진 군으로부터 선택된 활성을 갖는다. Cas 단백질은 Cas9를 포함할 수 있다. Cas9는 Cas9의 뉴클레아제 활성을 녹아웃시키는 적어도 1개의 아미노산 돌연변이를 포함할 수 있다. 적어도 1개의 아미노산 돌연변이는 D10A 및 H840A 중 적어도 하나일 수 있다. Cas 단백질은 iCas9 (서열 1의 아미노산 36-1403)를 포함할 수 있다. 제2 폴리펩티드 도메인은 전사 활성화 활성을 가질 수 있다. 제2 폴리펩티드 도메인은 적어도 1개의 VP16 전사 활성화 도메인 반복부를 포함할 수 있다. 제2 폴리펩티드 도메인은 VP16 사랑체 ("VP64") 또는 p65 활성화 도메인을 포함할 수 있다. 융합 단백질은 제2 폴리펩티드 도메인에 제1 폴리펩티드 도메인을 연결하는 링커를 추가로 포함할 수 있다. 융합 단백질은 iCas9-VP64를 포함할 수 있다.

[0012] 본 발명은 상기 융합 단백질 및 적어도 1개의 가이드 RNA (gRNA)를 포함하는 DNA 표적화 시스템에 관한 것이다. 적어도 1개의 gRNA는 표적 DNA 서열의 12-22개 염기 쌍의 상보적 폴리뉴클레오티드 서열 및 그에 이어서 프로토스페이서-인접 모티프를 포함할 수 있다. 적어도 1개의 gRNA는 유전자의 프로모터 영역, 유전자의 인핸서 영역 또는 유전자의 전사된 영역을 표적화할 수 있다. 적어도 1개의 gRNA는 유전자의 인트론을 표적화할 수 있다. 적어도 1개의 gRNA는 유전자의 엑손을 표적화할 수 있다. 적어도 1개의 gRNA는 ASCL1, BRN2, MYT1L, NANOG, VEGFA, TERT, IL1B, IL1R2, IL1RN, HBG1, HBG2 및 MYOD1로 이루어진 군으로부터 선택된 유전자의 프로모터 영역을 표적화할 수 있다. 적어도 1개의 gRNA는 서열 5-40, 65-144, 492-515, 540-563 및 585-625 중 적어도 하나를 포함할 수 있다.

[0013] 본 발명은 Cas9 및 적어도 1개의 가이드 RNA (gRNA)를 포함하는, 디스트로핀 유전자에 결합하는 DNA 표적화 시스템에 관한 것이다. 적어도 1개의 gRNA는 디스트로핀 유전자의 인트론을 표적화할 수 있다. 적어도 1개의 gRNA는 디스트로핀 유전자의 엑손을 표적화할 수 있다. 적어도 1개의 가이드 RNA는 서열 5-40, 65-144, 492-515, 540-563 및 585-625 중 적어도 하나를 포함할 수 있다. DNA 표적화 시스템은 1 내지 10개의 상이한 gRNA를 포함할 수 있다.

[0014] 본 발명은 상기 융합 단백질 또는 상기 DNA 표적화 시스템을 코딩하는 단리된 폴리뉴클레오티드에 관한 것이다.

[0015] 본 발명은 상기 단리된 폴리뉴클레오티드를 포함하는 벡터에 관한 것이다.

[0016] 본 발명은 상기 단리된 폴리뉴클레오티드 또는 상기 벡터를 포함하는 세포에 관한 것이다.

[0017] 본 발명은 세포에서 포유동물 유전자 발현을 조절하는 방법에 관한 것이다. 방법은 세포를 상기 융합 단백질, 상기 DNA 표적화 시스템, 상기 단리된 폴리뉴클레오티드 또는 상기 벡터와 접촉시키는 것을 포함한다. 유전자

발현이 유도될 수 있다.

[0018] 본 발명은 세포를 전환분화시키거나 또는 세포의 분화를 유도하는 방법에 관한 것이다. 방법은 세포를 상기 융합 단백질, 상기 DNA 표적화 시스템, 상기 단리된 폴리뉴클레오티드 또는 상기 벡터와 접촉시키는 것을 포함한다. 세포는 섬유모세포 또는 유도된 만능 줄기 세포일 수 있다. 섬유모세포는 뉴런 세포 또는 근원성 세포로 전환분화될 수 있다. DNA 표적화 시스템은 세포와 접촉될 수 있고, 적어도 1개의 gRNA는 ASCL1, BRN2, MYOD1 및 MYT1L로 이루어진 군으로부터 선택된 적어도 1개의 유전자의 프로모터 영역을 표적화한다. DNA 표적화 시스템은 ASCL1 유전자의 프로모터 영역을 표적화하는 적어도 1개의 gRNA 및 BRN2 유전자의 프로모터 영역을 표적화하는 적어도 1개의 gRNA를 포함할 수 있다. DNA 표적화 시스템은 1 내지 20개의 상이한 gRNA를 포함할 수 있다. DNA 표적화 시스템은 8 또는 16개의 상이한 gRNA를 포함할 수 있다. DNA 표적화 시스템은 dCas9-VP64를 포함할 수 있다. DNA 표적화 시스템은 바이러스로 또는 비-바이러스로 세포에 전달될 수 있다.

[0019] 본 발명은 세포에서 돌연변이체 유전자를 교정하는 방법에 관한 것이다. 방법은 상기 DNA 표적화 시스템을 함유하는 세포에 상기 단리된 폴리뉴클레오티드 또는 상기 벡터를 투여하는 것을 포함한다. 돌연변이체 유전자의 교정은 상동성-지정 복구를 포함할 수 있다. 방법은 세포에 공여자 DNA를 투여하는 것을 추가로 포함할 수 있다. 돌연변이체 유전자는 조기 정지 코돈 및 말단절단된 유전자 산물을 유발하는 프레임시프트 돌연변이를 포함할 수 있다. 돌연변이체 유전자의 교정은 뉴클레아제 매개된 비-상동 말단 연결을 포함할 수 있다. 돌연변이체 유전자의 교정은 조기 정지 코돈의 결실, 스플라이스 수용자 부위의 파괴, 하나 이상의 엑손의 결실 또는 스플라이스 공여자 서열의 파괴를 포함할 수 있다. 하나 이상의 엑손의 결실은 리딩 프레임의 교정을 유발할 수 있다.

[0020] 본 발명은 돌연변이체 디스트로핀 유전자를 갖는 치료를 필요로 하는 대상체를 치료하는 방법에 관한 것이다. 방법은 대상체에게 상기 DNA 표적화 시스템, 상기 단리된 폴리뉴클레오티드 또는 상기 벡터를 투여하는 것을 포함한다. 대상체는 뒤시엔스 근육 이영양증을 앓고 있을 수 있다.

[0021] 본 발명은 세포에서 돌연변이체 디스트로핀 유전자를 교정하는 방법에 관한 것이다. 방법은 돌연변이체 디스트로핀 유전자를 함유하는 세포에 상기 DNA 표적화 시스템, 상기 단리된 폴리뉴클레오티드, 상기 벡터 또는 상기 세포를 투여하는 것을 포함한다. 돌연변이체 디스트로핀 유전자는 조기 정지 코돈, 유전자 결실을 통해 파괴된 리딩 프레임, 이상 스플라이스 수용자 부위 또는 이상 스플라이스 공여자 부위를 포함할 수 있고, 여기서 표적 영역은 조기 정지 코돈, 파괴된 리딩 프레임, 이상 스플라이스 수용자 부위 또는 이상 스플라이스 공여자 부위의 상류 또는 하류이다. 돌연변이체 디스트로핀 유전자의 교정은 상동성-지정 복구를 포함할 수 있다. 방법은 세포에 공여자 DNA를 투여하는 것을 추가로 포함할 수 있다. 돌연변이체 디스트로핀 유전자는 조기 정지 코돈 및 말단절단된 유전자 산물을 유발하는 프레임시프트 돌연변이를 포함할 수 있다. 돌연변이체 디스트로핀 유전자의 교정은 뉴클레아제 매개된 비-상동 말단 연결을 포함할 수 있다. 돌연변이체 디스트로핀 유전자의 교정은 조기 정지 코돈의 결실, 파괴된 리딩 프레임의 교정, 또는 스플라이스 수용자 부위의 파괴 또는 스플라이스 공여자 서열의 파괴에 의한 스플라이싱의 조절을 포함할 수 있다. 돌연변이체 디스트로핀 유전자의 교정은 엑손 45-55 또는 엑손 51의 결실을 포함할 수 있다.

[0022] 본 발명은 상기 융합 단백질, 상기 DNA 표적화 시스템, 상기 단리된 폴리뉴클레오티드, 상기 벡터 또는 상기 세포를 포함하는 키트에 관한 것이다.

[0023] 본 발명은 세포에서 포유동물 유전자 발현을 조절하는 방법에 관한 것이다. 방법은 세포를 DNA 표적화 시스템을 코딩하는 폴리뉴클레오티드와 접촉시키는 것을 포함한다. DNA 표적화 시스템은 상기 융합 단백질 및 적어도 1개의 가이드 RNA (gRNA)를 포함한다. DNA 표적화 시스템은 1 내지 10개의 상이한 gRNA를 포함할 수 있다. 상이한 gRNA는 표적 유전자 내의 상이한 표적 영역에 결합할 수 있다. 표적 영역은 적어도 1개의 뉴클레오티드에 의해 분리될 수 있다. 표적 영역은 약 15 내지 약 700개 염기 쌍에 의해 분리될 수 있다. 각각의 상이한 gRNA는 적어도 1개의 상이한 표적 유전자에 결합할 수 있다. 상이한 표적 유전자는 동일한 염색체 상에 위치할 수 있다. 상이한 표적 유전자는 상이한 염색체 상에 위치할 수 있다. 적어도 1개의 표적 영역은 비-개방 염색질 영역, 개방 염색질 영역, 표적 유전자의 프로모터 영역, 표적 유전자의 인핸서 영역, 표적 유전자의 전사된 영역 또는 표적 유전자의 전사 개시 부위의 상류 영역 내에 있을 수 있다. 적어도 1개의 표적 영역은 표적 유전자의 전사 개시 부위의 상류 약 1 내지 약 1000개 염기 쌍 사이에 위치할 수 있다. 적어도 1개의 표적 영역은 표적 유전자의 전사 개시 부위의 상류 약 1 내지 약 600개 염기 쌍 사이에 위치할 수 있다. 유전자 발현이 유도될 수 있다. DNA 표적화 시스템은 2개의 상이한 gRNA, 3개의 상이한 gRNA, 4개의 상이한 gRNA, 5개의 상이한 gRNA, 6개의 상이한 gRNA, 7개의 상이한 gRNA, 8개의 상이한 gRNA, 9개의 상이한 gRNA 또는 10개의 상이한

gRNA를 포함할 수 있다. 적어도 1개의 가이드 RNA는 ASCL1, BRN2, MYT1L, NANOG, VEGFA, TERT, IL1B, IL1R2, IL1RN, HBG1, HBG2 및 MYOD1로 이루어진 군으로부터 선택된 유전자의 프로모터 영역을 표적화할 수 있다. 적어도 1개의 가이드 RNA는 서열 5-40, 65-144, 492-515, 540-563 및 585-625 중 적어도 하나를 포함할 수 있다. 적어도 1개의 표적 영역은 표적 유전자의 인트론 또는 엑손 내에 있을 수 있다.

[0024] 본 발명은 세포에서 포유동물 유전자 발현을 유도하기 위한 조성물에 관한 것이다. 조성물은 상기 융합 단백질 및 적어도 1개의 가이드 RNA (gRNA)를 포함한다.

[0025] 본 발명은 세포에서 포유동물 유전자 발현을 유도하기 위한 조성물에 관한 것이다. 조성물은 상기 융합 단백질을 코딩하는 단리된 폴리뉴클레오티드 서열 및 적어도 1개의 가이드 RNA (gRNA)를 포함한다. 적어도 1개의 가이드 RNA는 ASCL1, BRN2, MYT1L, NANOG, VEGFA, TERT, IL1B, IL1R2, IL1RN, HBG1, HBG2 및 MYOD1로 이루어진 군으로부터 선택된 유전자의 프로모터 영역을 표적화할 수 있다. 적어도 1개의 가이드 RNA는 서열 5-40, 65-144, 492-515, 540-563 및 585-625 중 적어도 하나를 포함할 수 있다.

[0026] 본 발명은 세포에서 포유동물 유전자 발현을 유도하기 위한 상기 조성물을 포함하는 세포에 관한 것이다.

[0027] 본 발명은 세포에서 포유동물 유전자 발현을 유도하기 위한 상기 조성물 또는 세포에서 포유동물 유전자 발현을 유도하기 위한 상기 조성물을 포함하는 상기 세포를 포함하는 키트에 관한 것이다.

[0028] 본 발명은 세포에서 포유동물 유전자 발현을 유도하기 위한 키트에 관한 것이다. 키트는 세포에서 포유동물 유전자 발현을 유도하기 위한 상기 조성물 또는 세포에서 포유동물 유전자 발현을 유도하기 위한 상기 조성물을 포함하는 상기 세포를 포함한다.

[0029] 본 발명은 대상체의 근육에서의 게놈 편집을 위한 조성물에 관한 것이다. 조성물은 변형된 아데노-연관 바이러스 (AAV) 벡터 및 부위-특이적 뉴클레아제를 코딩하는 뉴클레오티드 서열을 포함한다. 근육은 골격근 또는 심장 근육이다. 변형된 AAV 벡터는 증진된 심장 근육 및 골격근 조직 향성을 가질 수 있다. 부위-특이적 뉴클레아제는 아연 핑거 뉴클레아제, TAL 이펙터 뉴클레아제 또는 CRISPR/Cas9 시스템을 포함할 수 있다. 부위-특이적 뉴클레아제는 근육 세포에서의 유전자 또는 유전자좌에 결합할 수 있다. 유전자 또는 유전자좌는 디스트로핀 유전자일 수 있다. 조성물은 공여자 DNA 또는 트랜스진을 추가로 포함할 수 있다.

[0030] 본 발명은 대상체의 근육에서의 게놈 편집을 위한 상기 조성물을 포함하는 키트에 관한 것이다.

[0031] 본 발명은 대상체의 근육에서 게놈 편집하는 방법에 관한 것이다. 방법은 대상체의 근육에서의 게놈 편집을 위한 상기 조성물을 근육에 투여하는 것을 포함하며, 여기서 근육은 골격근 또는 심장 근육이다. 게놈 편집은 돌연변이체 유전자를 교정하거나 또는 트랜스진을 삽입하는 것을 포함할 수 있다. 돌연변이체 유전자를 교정하는 것은 돌연변이체 유전자를 결실시키거나, 재배열하거나 또는 대체하는 것을 포함할 수 있다. 돌연변이체 유전자를 교정하는 것은 뉴클레아제-매개 비-상동 말단 연결 또는 상동성-지정 복구를 포함할 수 있다.

[0032] 본 발명은 대상체를 치료하는 방법에 관한 것이다. 방법은 대상체의 근육에서의 게놈 편집을 위한 상기 조성물을 대상체의 근육에 투여하는 것을 포함하며, 여기서 근육은 골격근 또는 심장 근육이다. 대상체는 골격근 병태 또는 유전 질환을 앓고 있을 수 있다. 대상체는 뒤시엔느 근육 이영양증을 앓고 있을 수 있다.

[0033] 본 발명은 대상체의 근육에서의 게놈 편집을 위한 상기 조성물을 투여하는 것을 포함하는, 대상체에서 돌연변이체 유전자를 교정하는 방법에 관한 것이다. 근육은 골격근 또는 심장 근육이다. 조성물은 대상체의 골격근 내로 주사될 수 있다. 조성물은 대상체에 전신 주사될 수 있다. 골격근은 전경골근일 수 있다.

[0034] 본 발명은 상기 융합 단백질을 코딩하는 제1 폴리뉴클레오티드 서열 및 적어도 1개의 sgRNA를 코딩하는 제2 폴리뉴클레오티드 서열을 포함하는, 대상체에서의 게놈 편집을 위한 변형된 렌티바이러스 벡터에 관한 것이다. 제1 폴리뉴클레오티드 서열은 제1 프로모터에 작동가능하게 연결될 수 있다. 제1 프로모터는 구성적 프로모터, 유도성 프로모터, 억제성 프로모터 또는 조절성 프로모터일 수 있다. 제2 폴리뉴클레오티드 서열은 1 내지 10개의 상이한 sgRNA를 코딩할 수 있다. 제2 폴리뉴클레오티드 서열은 2개의 상이한 sgRNA, 3개의 상이한 sgRNA, 4개의 상이한 sgRNA, 5개의 상이한 sgRNA, 6개의 상이한 sgRNA, 7개의 상이한 sgRNA, 8개의 상이한 sgRNA, 9개의 상이한 sgRNA 또는 10개의 상이한 sgRNA를 코딩할 수 있다. 상이한 sgRNA를 코딩하는 각각의 폴리뉴클레오티드 서열은 프로모터에 작동가능하게 연결될 수 있다. 상이한 sgRNA에 작동가능하게 연결된 각각의 프로모터는 동일한 프로모터일 수 있다. 상이한 sgRNA에 작동가능하게 연결된 각각의 프로모터는 상이한 프로모터일 수 있다. 프로모터는 구성적 프로모터, 유도성 프로모터, 억제성 프로모터 또는 조절성 프로모터일 수 있다. sgRNA는 표적 유전자에 결합할 수 있다. 각각의 sgRNA는 1개의 표적 유전자좌 내의 상이한 표적 영역에 결합할

수 있다. 각각의 sgRNA는 상이한 유전자와 내의 상이한 표적 영역에 결합할 수 있다. 융합 단백질은 Cas9 단백질 또는 iCas9-VP64 단백질을 포함할 수 있다. 융합 단백질은 VP64 도메인, p300 도메인 또는 KRAB 도메인을 포함할 수 있다. 2개 이상의 내인성 유전자가 전사 활성화될 수 있다. 2개 이상의 내인성 유전자가 억제될 수 있다.

[0035] 본 발명은 세포에서 내인성 유전자를 활성화시키는 방법에 관한 것이다. 방법은 세포를 상기 변형된 렌티바이러스 벡터와 접촉시키는 것을 포함한다. 내인성 유전자는 일시적으로 활성화될 수 있다. 내인성 유전자는 안정하게 활성화될 수 있다. 내인성 유전자는 일시적으로 억제될 수 있다. 내인성 유전자는 안정하게 억제될 수 있다. 융합 단백질은 sgRNA와 유사한 수준으로 발현될 수 있다. 융합 단백질은 sgRNA와 상이한 수준으로 발현될 수 있다. 세포는 1차 인간 세포일 수 있다.

[0036] 본 발명은 세포에서의 멀티플렉스 유전자 편집 방법에 관한 것이다. 방법은 세포를 상기 변형된 렌티바이러스 벡터와 접촉시키는 것을 포함한다. 멀티플렉스 유전자 편집은 적어도 1개의 돌연변이체 유전자를 교정하거나 또는 트랜스진을 삽입하는 것을 포함할 수 있다. 돌연변이체 유전자를 교정하는 것은 적어도 1개의 돌연변이체 유전자를 결실시키거나, 재배열하거나 또는 대체하는 것을 포함할 수 있다. 적어도 1개의 돌연변이체 유전자를 교정하는 것은 뉴클레아제-매개 비-상동 말단 연결 또는 상동성-지정 복구를 포함할 수 있다. 멀티플렉스 유전자 편집은 적어도 1개의 유전자를 결실시키는 것을 포함하며, 여기서 유전자는 내인성 정상 유전자 또는 돌연변이체 유전자이다. 멀티플렉스 유전자 편집은 적어도 2개의 유전자를 결실시키는 것을 포함할 수 있다. 멀티플렉스 유전자 편집은 2 내지 10개의 유전자를 결실시키는 것을 포함할 수 있다.

[0037] 본 발명은 세포에서 적어도 1개의 표적 유전자의 유전자 발현을 조절하는 방법에 관한 것이다. 방법은 세포를 상기 변형된 렌티바이러스 벡터와 접촉시키는 것을 포함한다. 적어도 2개의 유전자의 유전자 발현이 조절될 수 있다. 2개의 유전자 내지 10개의 유전자의 유전자 발현이 조절될 수 있다. 적어도 1개의 표적 유전자의 유전자 발현 수준이 적어도 1개의 표적 유전자에 대한 정상 유전자 발현 수준과 비교하여 증가되거나 또는 감소되는 경우에 적어도 1개의 표적 유전자의 유전자 발현이 조절될 수 있다.

### 도면의 간단한 설명

[0038] 도 1은 iCas9-VP64에 의한 인간 IL1RN 유전자의 RNA-가이드 활성화를 보여준다. (a,b) RNA-가이드 전사 활성화제는 불활성화 Cas9 (iCas9, D10A/H840A)를 VP64 전사활성화 도메인에 융합시켜 생성하였다. iCas9-VP64는 20 bp 표적 서열에 대한 가이드 RNA (gRNA)의 혼성화를 통해 게놈 표적 부위를 인식한다. (c) IL1RN 프로모터에서의 서열을 표적화하는 4개의 gRNA 또는 crRNA/tracrRNA에 대한 발현 플라스미드를 iCas9-VP64 발현 플라스미드와 함께 HEK293T 세포 내로 공동-형질감염시켰다. IL1RN 발현의 활성화를 qRT-PCR에 의해 평가하였다. (d) 4개의 gRNA 발현 플라스미드를 iCas9-VP64와 함께 개별적으로 또는 조합으로 공동-형질감염시켰다. 강한 유전자 활성화는 단지 gRNA의 조합에 반응하여서만 qRT-PCR에 의해 관찰되었다. (e) IL1RN 발현의 활성화는 IL-1ra 유전자 산물의 배지 내로의 분리를 ELISA에 의해 평가함으로써 확인하였다. IL-1ra는 단지 gRNA의 조합으로 처리된 6개 샘플 중 3개에서 검출되었다. (c-e)의 경우에, 데이터는 평균  $\pm$  s.e.m. (n = 3 독립적 실험)으로 제시되어 있다. gRNA의 조합을 사용한 처리는 터키 검정에 의해 다른 모든 처리와 통계적으로 상이하였다 (\*P  $\leq$  0.02). (f) RNA-seq는 공 발현 벡터로 처리한 샘플 (n = 2)에 대해 수행하거나 또는 iCas9-VP64 및 IL1RN을 표적화하는 4개의 gRNA에 대한 발현 플라스미드로 공동-형질감염된 샘플 (n = 2)에 대해 수행하였다. 이들 처리 사이에서 단지 통계적으로 유의한 유전자 발현 변화는 4개의 IL1RN 이소형의 증가 (오류 발견율  $\leq$  3 x 10<sup>-4</sup>) 및 IL32의 감소 (오류 발견율 = 0.03)였다.

도 2는 세포 및 유전자 요법, 유전적 재프로그래밍 및 재생 의학과 관련된 인간 유전자의 RNA-가이드 활성화를 보여준다. HEK293T 세포를 iCas9-VP64 발현 플라스미드 및 4개의 gRNA로 개별적으로 또는 조합으로 형질감염시켰다. 표적 유전자 발현을 qRT-PCR에 의해 측정하고, GAPDH mRNA 수준으로 정규화하였다. 데이터는 평균  $\pm$  s.e.m. (n = 3 독립적 실험)으로 제시되어 있다. gRNA의 조합을 사용한 처리는 터키 검정에 의해 다른 모든 처리와 통계적으로 상이하였다 (\*P < 0.05).

도 3은 iCas9-VP64의 발현을 보여준다. 형질감염된 HEK293 세포에서 iCas9-VP64의 발현은 N-말단 Flag 에피토프 태그에 대한 웨스턴 블롯에 의해 확인하였다. wt Cas9 발현 플라스미드는 에피토프 태그를 함유하지 않는다.

도 4는 gRNA 표적 부위의 위치 및 인간 표적 유전자의 DNase 과민증을 보여준다. 각 유전자좌에 대한 4개의 gRNA의 표적 부위는 각 유전자 위에 맞춤 트랙으로 지정되고 있고, DNase-과민성 개방 염색질 영역을 나타내는



DNase-seq 데이터는 각 유전자 아래에 제시되어 있다. DNase-seq를 이전에 기재된 바와 같이 HEK293T 세포에서 수행하여 DNase 과민성 영역을 확인하였다 (Song et al., Cold Spring Harbor protocols 2010, pdb prot5384 (2010); Song et al. Genome Res 21, 1757-1767 (2011)). 결과는 개방 염색질이 gRNA의 iCas9-VP64와의 조합에 의한 유전자 활성화에 대한 요건이 아니었음을 제시한다.

도 5는 iCas9-VP64에 의한 뉴클레아제 활성의 부재를 보여준다. 야생형 Cas9 또는 불활성화된 (D10A, H840A) iCas9-VP64 발현 플라스미드를, IL1RN 프로모터를 표적화하는 4개의 상이한 가이드 RNA에 대한 발현 플라스미드와 함께 공동-형질감염시켰다. 뉴클레아제 활성을 서베이어(Surveyor) 검정에 의해 결정하였다 (Guschin et al., Methods Mol Biol 649, 247-256 (2010)). 뉴클레아제 활성 및 비-상동 말단 연결에 의한 DNA 복구를 나타내는 저분자량 밴드는 단지 야생형 Cas9를 사용한 처리 후에만 존재하였으며, 이는 iCas9-VP64에 의해 뉴클레아제 활성이 사라진다는 것을 지지한다.

도 6은 HBG1 및 HBG2를 표적화하는 gRNA로 처리된 샘플에 대한 RNA-seq를 보여준다. RNA-seq는 대조군 공 발현 벡터로 처리된 샘플 (n = 3)에 대해 수행하거나 또는 iCas9-VP64 및 HBG1을 표적화하는 4개의 gRNA에 대한 발현 플라스미드로 공동-형질감염된 샘플 (n = 2)에 대해 수행하였다. 이들 gRNA 중 3개는 또한 HBG2를 표적화한다. 대조군에 비해 HBG1 및 HBG2 둘 다에서 증가가 관찰되었지만, 낮은 발현 수준으로 인해 통계학적으로 유의하지는 않았다. 이들 처리 사이에서 단지 통계적으로 유의한 유전자 발현 변화는 IL32의 감소 (오류 발견율 = 0.0007) 및 TNFRS9의 감소 (오류 발견율 = 0.002)였다.

도 7은 iCas9-VP64에 의한 Ascl1 및  $\gamma$ -글로빈의 상향조절을 보여준다. HEK293T 세포를 iCas9-VP64 및 ASCL1 또는 HBG1 프로모터를 표적화하는 4개의 gRNA로 형질감염시켰다. 상응하는 Ascl1 및  $\gamma$ -글로빈 단백질 생산의 수준을 웨스턴 블롯에 의해 평가하였다. 이들 단백질의 낮은 수준은 HEK293T 세포에서 검출가능하였고, 발현의 증가는 2개의 독립적 실험에서 iCas9-VP64 처리 후에 검출가능하였다.

도 8은 iCas9-VP64-처리 무린 배아 섬유모세포에서 Ascl1의 하류 표적의 활성화를 보여준다. 마우스 배아 섬유모세포 (MEF)를 대조군 GFP 발현 플라스미드 또는 50:50 또는 75:25 비의 iCas9-VP64 발현 플라스미드 및 ASCL1을 표적화하는 4개의 gRNA 발현 플라스미드의 조합으로 형질감염시켰다. (a) 인간 ASCL1 프로모터 (서열 3) 내의 gRNA 표적 부위는 마우스 ASCL1 프로모터 (서열 4)에서 보존된다. 표적 부위는 실선으로 나타나 있고, 전사된 영역은 파선으로 나타나 있다. (b) MEF에서의 ASCL1 발현은 qRT-PCR에 의해 결정된 바와 같이 iCas9-VP64/gRNA 처리 2일 후에 증가하였다. (c-h) 신경 유도 배지에서의 10일 후에, 세포를 Ascl1 및 뉴런 분화의 초기 마커인 Tuj1에 대해 (c-d) 또는 Tuj1 및 보다 성숙한 뉴런 분화의 마커인 MAP2에 대해 (d-f) 염색하였다. 일부 Tuj1-양성 세포는 뉴런 형태를 채택하였고 (f-g), 단세포는 Tuj1 및 MAP2에 대해 양성인 것으로 밝혀졌다 (g). (h) Tuj1-양성 세포는 iCas9-VP64/gRNA-처리 배양물 (~0.05%)에서 용이하게 확인되었지만, 대조군에서는 부재하였다. n = 3 독립적 샘플 및 데이터는 평균  $\pm$  평균의 표준 오차로 나타내어진다. gRNA 75/25는 gRNA 50/50 및 대조군과 유의하게 상이하다 (\*P < 0.01, 터키 검정).

도 9는 (a) iCas9-VP64 단백질 서열 (서열 1) 및 (b) U6 프로모터를 갖는 gRNA 발현 카세트의 서열 (서열 2)을 보여준다.

도 10은 qRT-PCR에 대한 표준 곡선을 보여준다. 각 유전자에 대해, 최고 발현 수준을 갖는 실험 샘플을 희석하여 qRT-PCR에 의해 검정된 표준 곡선을 생성함으로써 적절한 동적 범위에 걸친 효율적인 증폭을 확실히 하였다. 모든 증폭 반응의 효율은 90-115% 내에 있었다.

도 11(a)-11(b)는 RNA-가이드 복구의 검증을 보여준다. 도 11(a)는 Cas9를 공벡터 (음성 대조군) 또는 gRNA와 함께 세포 내로 공동-형질감염시킨지 2일 후에 HEK 293T 세포로부터 수거한 게놈 DNA의 서베이어 검정 결과를 보여준다. 도 11(b)는 gRNA 표적의 위치를 보여준다. 도 11(c)는 각 gRNA에 대한 예상된 절단 크기를 보여준다.

도 12는 서베이어 검정에 의해 제시된 바와 같은 DMD 8036 (del48-50) 세포에서의 RNA-가이드 복구를 보여준다.

도 13은 전체 유전자좌에 걸쳐 PCR에 의해 제시된 바와 같은 DMD 8036 (del48-50) 세포에서의 RNA-가이드 복구를 보여준다. 야생형 디스트로핀 유전자의 PCR은 1447 bp 크기의 단편을 생성하고, 반면에 DMD 8036 세포주에서의 돌연변이체 유전자의 PCR은 대략 817 bp의 결실을 보여준다. CRISPR/Cas9-기반 시스템의 도입 후의 결실 밴드는 대략 630 bp였다.

도 14는 MANDYS8 (항-디스트로핀 항체) 및 GAPDH 항체 (양성 대조군)를 사용한 웨스턴 블롯에 의해 제시된 바와

같은 DMD 8036 (de148-50) 세포에서의 RNA-가이드 복구를 보여준다.

도 15는 IL1RN 프로모터를 표적화하는 iCas9-VP64의 특이적 결합을 예시하는 ChIP 서열분석 데이터를 보여준다. HEK 293T 세포는 IL1RN 프로모터를 표적화하는 iCas9-VP64로 형질감염시켰다.

도 16은 디스트로핀 유전자에 대한 CRISPR/Cas9 표적화를 보여준다. (A) sgRNA 서열을 디스트로핀 유전자의 엑손 45-55 돌연변이 핫스팟 영역의 서열에 결합하도록 설계하여, 유전자 편집이 폭넓게 다양한 환자-특이적 돌연변이로부터 디스트로핀 발현을 복원할 수 있도록 하였다. 인트론 내의 화살표는 게놈으로부터 전체 엑손을 결실시키도록 설계된 sgRNA 표적을 나타낸다. 엑손 내의 화살표는 디스트로핀 유전자에서 표적화된 프레임시프트를 생성하도록 설계된 sgRNA 표적을 나타낸다. (B) 작은 삽입 또는 결실의 도입 후에 CR3 sgRNA를 사용한 엑손 51에서의 NHEJ DNA 복구에 의한 프레임 교정의 예. (C) 엑손 51을 결실시키고 엑손 48-50의 결실을 갖는 환자 돌연변이에서 디스트로핀 리딩 프레임을 복원하도록 설계된 멀티플렉스 sgRNA 표적의 개략도. (D) 다양한 DMD 환자 돌연변이를 다루기 위해 전체 엑손 45-55 영역을 결실시키도록 설계된 멀티플렉스 sgRNA 표적의 개략도.

도 17은 표 7에서의 제3일 유전자 변형을 측정하기 위한 서베이어 검정 결과를 정량화하는데 사용된 TBE-PAGE 겔의 영상을 보여준다. 별표는 뉴클레아제 활성을 나타내는 밴드의 예상된 크기를 표시한다.

도 18은 표 7에서의 제10일 유전자 변형을 측정하기 위한 서베이어 검정 결과를 정량화하는데 사용된 TBE-PAGE 겔의 영상을 보여준다. 별표는 뉴클레아제 활성을 나타내는 밴드의 예상된 크기를 표시한다.

도 19는 유전자 변형된 DMD 근모세포를 보강하기 위한 형광-활성화 유동 분류를 보여준다. (A) T2A 리보솜 스키프 펩티드 서열을 사용하여 GFP 마커에 연결된 인간-코돈 최적화된 SpCas9 단백질을 발현하는 플라스미드를, sgRNA 발현 카세트를 보유하는 1 또는 2개의 플라스미드와 함께 인간 DMD 근모세포 내로 공동-전기천공하였다. (B) 나타낸 sgRNA 발현 카세트를, T2A 리보솜 스키프 펩티드 서열에 의해 SpCas9에 연결된 GFP 마커와 함께 (하단) 또는 그 없이 (상단) SpCas9를 발현하는 별개의 플라스미드와 함께 HEK293T 내로 독립적으로 공동-형질감염시켰다. 유전자 변형 빈도를 형질감염 3일 후에 서베이어 검정에 의해 평가하였다. (C) 디스트로핀 유전자에서 엑손 48-50의 결실을 갖는 DMD 근모세포를, 이들 환자 세포에서 디스트로핀 리딩 프레임을 교정하는 sgRNA로 처리하였다. 유전자 변형을 미분류 (별크) 또는 GFP+ 분류 세포에서 전기천공 20일 후에 평가하였다. (D) 나타낸 발현 플라스미드로 전기천공한지 3일 후의 DMD 근모세포의 GFP 발현. 형질감염 효율 및 분류된 세포 집단은 게이팅 영역에 의해 나타나 있다.

도 20은 CRISPR/Cas9를 사용하여 디스트로핀 리딩 프레임을 복원하기 위한 표적화된 프레임시프트를 보여준다. (A) 엑손 51의 5' 영역은 제1 아웃-오브-프레임 정지 코돈의 바로 상류에 결합하는 sgRNA인 CR3를 사용하여 표적화하였다. PAM: 프로토스페이서-인접 모티프. (B) 엑손 51 유전자좌는 SpCas9 및 CR3 발현 카세트로 처리된 HEK293T 세포로부터 PCR 증폭시켰다. 개별 클론의 서열은 생어(Sanger) 서열분석에 의해 결정하였다. 상단 서열 (볼드체, 적색의 엑손)은 천연 비변형 서열이다. 각 서열에 대한 클론의 개수는 괄호에 나타나 있다. (C) (B)에 제시된 유전자 변형으로부터 생성된 총 유전자 편집 효율 및 리딩 프레임 전환율의 요약. (D) 디스트로핀 리딩 프레임을 복원하기 위한 표적화된 프레임시프트를 생성하는, SpCas9 및 CR3 sgRNA 발현 카세트로 처리된 인간 DMD 근모세포에서 (도 19C)의 디스트로핀 발현에 대한 웨스턴 블롯. 디스트로핀 발현은 분화 6일 후에 디스트로핀 단백질의 막대형-도메인에 대한 항체를 사용하여 프로빙하였다.

도 21은 멀티플렉스 CRISPR/Cas9 유전자 편집을 사용하는 인간 게놈으로부터의 엑손 51의 결실을 보여준다. (A) 엑손 48-50의 결실을 갖는 인간 DMD 근모세포에서의 엑손 51 유전자좌에 걸친 엔드-포인트 게놈 PCR. 상단 화살표는 전장 PCR 앰플리콘의 예상된 위치를 나타내고, 2개의 하단 화살표는 나타낸 sgRNA 조합에 의해 유발된 결실을 갖는 PCR 앰플리콘의 예상된 위치를 나타낸다. (B) (A)로부터의 PCR 산물을 클로닝하고, 개별 클론을 서열분석하여 표적화된 유전자좌에 존재하는 삽입 및 결실을 결정하였다. 상단 열은 야생형 미변형 서열을 보여주고, 삼각형은 SpCas9 절단 부위를 나타낸다. 예상된 결실 접합부의 서열을 보여주는 대표적인 크로마토그램이 우측에 있다. (C) 나타낸 sgRNA로 처리된 CRISPR/Cas9-변형 인간 Δ48-50 DMD 근모세포에서의 디스트로핀 mRNA 전사체의 엔드-포인트 RT-PCR 분석. 예상된 결실 PCR 산물의 대표적인 크로마토그램이 우측에 제시되어 있다. 별표: 미변형 가닥에 대한 결실 산물 가닥의 혼성화로부터 생성된 밴드. (D) CRISPR/Cas9 게놈 편집에 의한 디스트로핀 단백질 발현의 구조는 로딩 대조군으로서 GAPDH를 갖는 디스트로핀 단백질에 대한 웨스턴 블롯에 의해 평가하였다. 화살표는 예상된 복원된 디스트로핀 단백질 밴드를 나타낸다.

도 22는 멀티플렉스 CRISPR/Cas9 유전자 편집에 의한 인간 DMD 근모세포에서의 전체 엑손 45-55 영역의 결실을 보여준다. (A) HEK293T 또는 DMD 근모세포를 나타낸 sgRNA로 처리한 후 인트론 44와 인트론 55 사이의 영역의

결실을 검출하기 위한 게놈 DNA의 엔드-포인트 게놈 PCR. (B) (A)에서의 DMD 근모세포로부터의 결실에 대한 예상된 크기의 PCR 산물의 개별 클론을 생어 서열분석에 의해 분석하여 표적화된 유전자좌에 존재하는 게놈 결실의 서열을 결정하였다. 예상된 결실 접합부의 서열을 보여주는 대표적인 크로마토그램이 하단에 있다. (C) 나타낸 sgRNA로 처리된 CRISPR/Cas9-변형 인간 Δ48-50 DMD 근모세포에서의 디스트로핀 mRNA 전사체의 엔드-포인트 RT-PCR 분석. 예상된 결실 PCR 산물의 대표적인 크로마토그램이 우측에 제시되어 있다. (D) 인트론 44 및/또는 인트론 55를 표적화하는 sgRNA로 DMD 근모세포를 전기천공한 후 복원된 디스트로핀 단백질 발현의 웨스턴 블롯에 의한 분석.

도 23은 생체내 세포 이식 실험에 사용된 유전자-변형 DMD 근모세포의 유동 세포측정법-기반 보강의 확인을 보여준다. DMD 근모세포를 CR1 및 CR5에 대한 sgRNA 발현 벡터와 함께 또는 그 없이 Cas9로 처리하고, 유동 세포측정법에 의해 GFP+ 세포를 분류하였다. 엑손 51 유전자좌에서의 결실은 유전자좌에 플랭킹된 프라이머를 사용하여 엔드-포인트 PCR에 의해 검출하였다. 음성 대조군: 단지 Cas9로 처리되고 GFP+ 세포에 대해 분류된 DMD 근모세포.

도 24는 면역결핍 마우스 내로 CRISPR/Cas9-처리 인간 DMD 근모세포의 이식 후에 복원된 인간 디스트로핀의 생체내 발현을 보여준다. 인간 Δ48-50 DMD 근모세포를 SpCas9, CR1 및 CR5로 처리하여 엑손 51을 결실시키고, 도 19에 제시된 바와 같이 GFP 발현에 대해 분류하였다. 이들 분류 세포 및 미처리 대조군 세포를 면역결핍 마우스의 뒷다리 내로 주사하고, 이식 후 4주 후에 근섬유에서 인간-특이적 단백질 발현을 평가하였다. 동결절편은 나타낸 바와 같이, 마우스 근섬유 내로 융합된 비교정 및 교정 근모세포 둘 다에 의해 발현되는 항-인간 스펙트린으로 염색하거나, 또는 항-인간 디스트로핀 항체로 염색되었다. 백색 화살표는 인간 디스트로핀에 대해 양성인 근섬유를 나타낸다.

도 25는 인간 디스트로핀 발현을 프로빙하는 추가의 면역형광 영상을 보여준다. 항-인간 스펙트린으로 염색된 영역으로부터의 연속 절편은 상단 좌측에 있는 삽도에 제시되어 있다. (A-C) 미처리 인간 DMD 근모세포가 주사된 근육으로부터의 절편. (D-F) 유동 세포측정법에 의해 보강된, CR1/5 처리된 인간 DMD 근모세포가 주사된 근육으로부터의 절편. 백색 화살표는 디스트로핀 양성 섬유를 나타낸다.

도 26은 인간 세포에서 엑손 51의 CR1/CR5-매개 결실에 대한 CRISPR/Cas9 독성 및 오프-타겟 효과의 평가를 보여준다. (A) 인간-최적화된 SpCas9 및 나타낸 sgRNA 구축물로 처리된 HEK293T 세포에서의 세포독성 검정의 결과. 세포독성은 나타낸 뉴클레아제로 공동-형질감염된 GFP-양성 세포의 생존을 기반으로 한다. I-SceI는 잘-특성화된 비-독성 메가뉴클레아제이고, GZF3은 공지된 독성 아연 핑거 뉴클레아제이다. (B) Cas9를 코딩하는 발현 카세트 및 나타낸 sgRNA로 처리된 분류된 hDMD 세포 내 오프-타겟 부위에서의 서베이어 분석. hDMD 세포에서 시험된 이들 3개의 오프-타겟 부위는 HEK293T 세포에서 시험된 50개의 예측된 부위의 패널로부터 확인하였다 (도 27 및 표 4). TGT: 나타낸 sgRNA에 대한 온-타겟 유전자좌. OT: 오프-타겟 유전자좌. (C, D) Cas9 및 CR1로 처리된 HEK293T 세포 (C) 또는 Cas9, CR1 및 CR5로 처리된 분류된 hDMD 세포 (D)에서 염색체 전위를 검출하기 위한 엔드-포인트 네스티드 PCR. 개략도는 각 전위 사건에 대해 맞춤형 네스티드 프라이머 쌍의 상대 위치를 도시한다. 각 밴드의 예상된 크기는 프라이머 크기 및 각 유전자좌에서 예측된 sgRNA 컷 부위의 위치를 기반으로 하여 추정하였다. 별표는 예상된 크기에서 검출된 밴드를 나타낸다. (C)에서의 밴드의 정체는 각 말단으로부터 생어 서열분석에 의해 확인하였다 (도 30). HEK293T 세포에서의 P2/P5 전위에 대한 대표적인 크로마토그램이 제시되어 있다.

도 27은 표 4에서의 온-타겟 및 오프-타겟 유전자 변형을 측정하기 위한 서베이어 검정 결과를 정량화하는데 사용된 TBE-PAGE 겔의 영상을 보여준다. 별표는 뉴클레아제 활성을 나타내는 밴드의 예상된 크기를 표시한다.

도 28은 인간 세포에서 CR3 및 CR6/CR36에 대한 CRISPR/Cas9 오프-타겟 활성화에 의해 유발된 염색체 전위를 검출하기 위한 엔드-포인트 네스티드 PCR을 보여준다. 네스티드 엔드-포인트 PCR 분석을 사용하여, (A) 나타낸 바와 같이 Cas9 및 CR3으로 처리된 HEK293T 또는 분류된 hDMD 세포, (B) Cas9 및 CR36 단독으로 처리된 HEK293T 세포, 또는 (C) Cas9, CR6 및 CR36 발현 카세트로 처리된 분류된 hDMD 세포에서 전위를 검출하였다. 전위에 대한 제2 네스티드 PCR 반응을 특이성을 최대화하기 위한 각 예측된 전위 유전자좌에 대한 맞춤 프라이머를 사용하여 증폭시켰다 (표 4 참조). 개략도는 전위의 존재를 프로빙하는데 사용된 네스티드 프라이머 쌍의 상대 위치를 도시한다. 먼저 각각의 가능한 전위 사건은 나타낸 sgRNA(들)와 함께 또는 그 없이 처리된 세포로부터 단리된 게놈 DNA로부터 증폭시켰다. 제2 네스티드 PCR 반응은 전위로부터 생성될 예측된 PCR 앰플리콘 내의 프라이머를 사용하여 수행하였다. 예상된 크기는 나타낸 프라이머 결합 부위 및 각 유전자좌에서의 예측된 sgRNA 컷 부위를 기반으로 하여 추정하였다. \*는 예상된 크기에서 검출되며 각 말단으로부터의 생어 서열분석에 의해

확인된 밴드를 나타낸다. #는 생어 서열분석에서 가능하게는 네스티드 PCR 동안 미스프라이밍의 결과인 예측된 전위 이외의 서열이 제시된 앰플리콘을 나타낸다.

도 29는 Cas9 및 CR3 유전자 카세트로 처리된 HEK293T 세포에서, 각각 염색체 X 및 1 상에, CR3과 CR3-OT1 사이의 전위로부터 생성된 도 28에서 검출된 밴드에 대한 생어 서열분석 크로마토그램을 보여준다. 화살표는 적절한 sgRNA에 의해 유발된 예상 파단점 근처에서 나타난 염색체에 대한 상동성의 영역을 보여준다. 서열분석 판독은 비-상동 말단 연결에 의한 DNA 복구의 오류-유발 성질로 인해 파단점 근처에서 상이 달라진다는 것을 주목한다.

도 30은 Cas9 및 CR1 유전자 카세트로 처리된 HEK293T 세포에서, 각각 염색체 X 및 16 상에, CR1과 CR1-OT1 사이의 전위로부터 생성된 도 26에서 검출된 밴드에 대한 생어 서열분석 크로마토그램을 보여준다. 화살표는 적절한 sgRNA에 의해 유발된 예상 파단점 근처에서 나타난 염색체에 대한 상동성의 영역을 보여준다. 서열분석 판독은 비-상동 말단 연결에 의한 DNA 복구의 오류-유발 성질로 인해 파단점 근처에서 상이 달라진다는 것을 주목한다.

도 31은 생체내 AAV 주사 및 조직 수거의 개관을 보여준다.

도 32는 시험관내 및 생체내에서 AAV-SASTG-ROSA의 전달 후 골격근에서의 Rosa26 ZFN 활성의 서베이어 분석을 보여준다. 화살표는 서베이어 절단으로부터 생성된 예상 밴드를 나타낸다. n.d.: 검출되지 않음. (a) 증식 C2C12에 나타난 양의 바이러스를 형질도입하고, 감염 4일 후에 수거하였다. 화살표는 서베이어 절단으로부터 생성된 예상 밴드 크기를 나타낸다. (b) C2C12를 5일 동안 분화 배지에서 인큐베이션한 다음, 24 웰 플레이트에서 나타난 양의 AAV-SASTG-ROSA 바이러스를 형질도입하였다. 샘플을 형질도입 10일 후에 수집하였다. (c) 나타난 양의 AAV-SASTG-ROSA를 C57BL/6J 마우스의 전경골근 내로 직접 주사하고, 근육을 감염 4주 후에 수거하였다. 수거된 TA 근육은 게놈 DNA 분석을 위해 8개의 별개의 조각으로 분할하였으며, 각각 별개의 레인에 제시되어 있다.

도 33은 Rosa T2A opt DNA 서열 (서열 434) 및 Rosa T2A opt 단백질 서열 (서열 435)을 보여준다.

도 34는 SASTG 캡시드 DNA 서열 (서열 436) 및 SASTG 캡시드 펩티드 서열 (서열 437)을 보여준다.

도 35는 DZF16 ZFN 표적 부위 서열 (서열 442), DZF16-L6 좌측 전체 아미노산 서열 (서열 443) 및 DZF16-R6 우측 전체 아미노산 서열 (서열 444)을 보여준다.

도 36은 E51C3 ZFN 표적 부위 서열 (서열 445), E51C3-3L 좌측 전체 아미노산 서열 (서열 446) 및 E51C3-3R 우측 전체 아미노산 서열 (서열 447)을 보여준다.

도 37은 DZF15 ZFN 표적 부위 서열 (서열 448), DZF15-L6 좌측 전체 아미노산 서열 (서열 449), DZF15-R6 우측 전체 아미노산 서열 (서열 450), DZF15-L5 우측 전체 아미노산 서열 (서열 451), DZF15-R5 우측 전체 아미노산 서열 (서열 452)을 보여준다.

도 38은 E51C4 ZFN 표적 부위 서열 (서열 453), E51C4-4L 좌측 전체 아미노산 서열 (서열 454) 및 E51C4-4R 우측 전체 아미노산 서열 (서열 455)을 보여준다.

도 39는 "단일 벡터, 멀티플렉스 CRISPR 시스템", "이중 벡터, 멀티플렉스 CRISPR 시스템" 및 "단일 벡터, 단일 gRNA 시스템"의 개략적 다이어그램을 보여준다.

도 40은 SaCas9-NLS (NLS는 밑줄표시됨) (서열 64) 및 SaCas9 gRNA (서열 116)의 뉴클레오티드 서열을 보여준다.

도 41은 NmCas9 (NLS 1은 밑줄표시되고, NLS 2는 밑줄표시되며 볼드체이고, HA 태그는 볼드체임)의 뉴클레오티드 서열, 톰슨(Thomson) PNAS 2013으로부터의 NmCas9 짧은 헤어핀 (서열 118) 및 처치 네이처 바이오테크 (Church Nature Biotech) 2013으로부터의 NmCas9 긴 헤어핀 (서열 119)을 보여준다.

도 42는 sgRNA 및 렌티바이러스 Cas9 발현 구축물의 검증을 보여준다. (A) AAVS1 유전자좌를 표적화하는 sgRNA를 발현하는 고유한 Pol III 프로모터를 코딩하는 구축물, 또는 hU6 프로모터에 바로 이어서 발현을 종결시키는 폴리-티미딘 ("PolyT")을 함유하는 구축물을 HEK293T 세포 내로 형질감염시켰다. 엔드-포인트 RT-PCR을 사용하여 형질감염 2일 후에 각각 나타난 프로모터/sgRNA 구축물의 발현을 프로빙하였다. -RT: 무 역전사효소 대조군. (B) HEK293T를 AAVS1 아연-핑거 뉴클레아제 또는 Cas9-T2A-GFP를 코딩하는 발현 벡터 및 나타난 프로모터/sgRNA 발현 카세트에 형질감염시키고, 서베이어 검정을 사용하여 형질감염 3일 후에 유전자 변형 수준에



대해 평가하였다. (C) HEK293T 세포에 sgRNA 없이 나타낸 Cas9-T2A-GFP 구축물을 코딩하는 렌티바이러스 구축물을 형질도입하고, 형질도입 7일 후에 웨스턴 블롯에 의해 Cas9 단백질의 N-말단 상의 FLAG 에피토프 태그를 프로빙함으로써 Cas9 발현을 평가하였다.

도 43은 단일 렌티바이러스 CRISPR/Cas9 발현 카세트의 골든 게이트(Golden Gate) 조립을 보여준다.

도 44는 멀티플렉스 CRISPR/Cas9 시스템의 단일 렌티바이러스 전달을 보여준다. (A) 별개의 게놈 유전자좌를 표적화하는 4개의 sgRNA를, 활성 Cas9 뉴클레아제를 발현하는 렌티바이러스 벡터 내로 클로닝하였다. (B) HEK293T 및 1차 인간 피부 섬유모세포에 나타낸 sgRNA를 발현하는 렌티바이러스를 형질도입하고, 서베이어 검정을 사용하여 절단 사건에 대해 검정하였다. HEK293T를 형질도입 7일 후에 검정하였다. 인간 섬유모세포를 형질도입 10일 후에 검정하였다.

도 45는 dCas9-VP64를 안정하게 발현하는 HEK293T에서의 일시적 유전자 활성화를 보여준다. HEK293T에 dCas9-VP64를 안정하게 발현하는 렌티바이러스를 형질도입하고, 후속하여 나타낸 sgRNA 조합을 발현하는 플라스미드로 형질감염시켰다. 전달된 sgRNA의 개수를 변화시킴으로써, 내인성 IL1RN (A) 및 HBG1 (B) 유전자좌의 조정가능한 내인성 유전자 활성화를 형질감염 3일 후에 달성하였다. 내인성 IL1RN (C) 및 HBG1 (D)의 피크 수준은 형질감염 3-6일 후에 관찰되었고, 활성화의 수준은 15-20일 사이에 배경 수준으로 돌아갔다. 중요하게, 세포주는 제20일에 제2 형질감염 후에, 비록 이전에 관찰된 것보다 낮은 수준일지라도 반응성일 수 있었다.

도 46은 단일 렌티바이러스 멀티플렉스 dCas9-VP64 벡터를 사용하는 HEK293T에서의 안정한 유전자 활성화를 보여준다. HEK293T를 dCas9-VP64 및 나타낸 gRNA 조합을 안정하게 발현하는 렌티바이러스로 형질도입하였다. 전달된 sgRNA의 개수를 변화시킴으로써, 내인성 IL1RN (A) 및 HBG1 (B) 유전자좌의 조정가능한 내인성 유전자 활성화를 형질감염 7일 후에 달성하였다. 내인성 IL1RN (C) 및 HBG1 (D)의 피크 수준은 형질감염 6일 후에 관찰되었고, 활성화의 수준은 제21일까지 지속되었다.

도 47은 IL1RN mRNA 발현 수준을 보여준다.

도 48은 BAM 뉴런 전사 인자의 이소성 발현을 통한 섬유모세포에서 뉴런으로의 직접적 전환을 나타내는 개략도를 보여준다.

도 49는 (A) dCas9-VP64 구축물의 개략도를 보여준다. dCas9-VP64는 VP16 전사 활성화 도메인의 사랑체에 융합된 Cas9 단백질의 촉매적 불활성 형태이다. (B) 게놈 표적에 대한 dCas9-VP64의 RNA-가이드 동원의 메카니즘을 보여주는 개략도. (C) CRISPR/Cas9 전사 인자를 사용하여 iN을 생성하는 실험 프로토콜의 개략도.

도 50은 (A) qRT-PCR에 의해 제3일에 결정된 내인성 ASCL1 발현 또는 (B) dCas9-VP64가 형질도입되고 ASCL1 프로모터, ASCL1 cDNA 또는 루시페라제를 표적화하는 gRNA로 형질감염된 MEF에서의 면역형광에 의해 검출된 총 ASCL1 단백질을 보여준다. 별표 (\*)는 4개의 gRNA와 비교하여 8개의 gRNA의 공동-전달에 의해 ASCL1 발현의 유의한 ( $p < 0.05$ ) 증가를 나타낸다. ASCL1의 이소성 발현은 dCas9-VP64 및 Ascl1 프로모터를 표적화하는 8개의 gRNA에 의해 유도되는 것보다 더 많은 단백질을 생산했지만, 배양에서 제3일까지 내인성 유전자좌를 활성화시키지 않았다.

도 51은 (A) 이소성 BAM 인자에 의해 또는 dCas9-VP64 및 BRN2 및 ASCL1 프로모터를 표적화하는 gRNA에 의해 생성된 TUJ1 및 MAP2-양성 세포, (B) N3 배지에서 제11일에 hSyn-RFP 리포터를 발현하는 뉴런 형태를 갖는 세포를 보여준다.

도 52는 (a) 배양 배지에서 KC1의 존재 하에 (하단) 또는 부재 하에 (상단) GCaMP5 칼슘 지시자에 대해 양성인 뉴런 형태를 갖는 세포, (b) KC1 첨가에 반응한 세포의 탈분극을 보여주는 시간에 따른 정규화된 형광 강도의 자취를 보여준다.

도 52는 섬유모세포를 뉴런으로 전환시키기 위한 dCas9-VP64 전사 인자를 사용한 iCas9-VP64-처리 무린 배아 섬유모세포에서의 Ascl1 및 Brn2, 즉 마스터 조절 유전자의 하루 표적의 활성화를 보여준다. 마우스 배아 섬유모세포 (MEF)를 대조군 GFP 발현 플라스미드 또는 iCas9-VP64 발현 플라스미드 및 ASCL1 및 BRN2를 표적화하는 8개의 gRNA 발현 플라스미드의 조합으로 형질감염시켰다. dCas9 전사 인자는 바이러스에 의해 전달하였다. 신경 유도 배지에서 10일 후에, 세포를 뉴런 분화의 초기 마커인 Tuj1 및 보다 성숙한 뉴런 분화의 마커인 MAP2에 대해 염색하였다. 뉴런으로의 전환은 효율적이었다.

도 53은 포유동물 유전자 조절의 제어를 위한 CRISPR/Cas9 플랫폼을 보여준다. A. Cas9-기반 이펙터는 표적 부위 특이성을 부여하는 교환가능한 20bp 프로토스페이서가 선행하는 Cas9와 복합체를 이루는 불변 영역으로 이루

어진 키메라 gRNA 분자의 존재 하에 게놈 서열에 결합한다. B. Cas9-기반 합성 전사 인자는 RNA 폴리머라제 활성을 방해함으로써 또는 프로모터 내에 결합하여 내인성 전사 인자의 결합 부위를 차단함으로써 표적 유전자의 전사를 억제한다. C. 조절 요소, 예컨대 인핸서를 표적화하는 것은 또한 다중 원위 유전자의 발현을 잠재적으로 차단할 수 있다.

도 54는 CRISPR/dCas9-KRAB를 사용하여 HS2 인핸서를 표적화하는 것을 보여준다. HS2 영역은 하류 >10kb의 글로빈 유전자의 발현을 원위 조절하는 강력한 인핸서이다. 단일 gRNA의 패넬은 인핸서 영역을 따른 부위를 표적화하도록 설계되었다.

도 55는 HS2 인핸서를 표적화하는 단일 gRNA가 글로빈 유전자의 강력한 전사 억제를 가져오는 것을 보여준다. a. dCas9 및 dCas9-KRAB 리프레서는 렌티바이러스 벡터 상에서 전달되었다. 단일 gRNA는 스크리닝을 위해 일시적으로 형질감염되었다. 형질감염 3일 후에 정량적 RT-PCR에 의해 검정하는 경우에, dCas9-KRAB를 발현하는 K562는 gRNA 처리를 받지 않은 대조군 세포와 비교하여 b.  $\gamma$ -글로빈, c.  $\epsilon$ -글로빈 및 d.  $\beta$ -글로빈 유전자의 최대 80% 억제를 달성한다. d. dCas9 또는 dCas9-KRAB를 발현하며 Cr4 또는 Cr8로 처리된 세포에서의 단백질 발현은  $\beta$ -액틴과 비교하여 제3일에  $\gamma$ -글로빈 발현의 정도 발현을 보여준다.

도 56은 a. 렌티바이러스로 처리되지 않은 세포, b. dCas9 렌티바이러스로 처리된 세포, 또는 c. dCas9-KRAB 렌티바이러스로 처리된 세포로 전달된 가변 용량의 gRNA 플라스미드를 갖는 글로빈 유전자와 유전자의 발현을 보여준다. 전달된 Cr4 gRNA 플라스미드의 용량을 증가시키는 것은 dCas9-KRAB 처리된 세포에서의 억제를 증진시켰으며, 이는 dCas9-KRAB 이펙터 및 표적화된 gRNA 둘 다가 억제를 달성하는데 역할을 하는 것을 나타낸다.

도 57은 단일 gRNA를 dCas9-KRAB와 함께 안정하게 전달하는 것이 글로빈 유전자의 발현을 침묵화시킨다는 것을 보여준다. a. dCas9 및 dCas9-KRAB 리프레서는 단일 gRNA를 갖는 렌티바이러스 벡터 상에서 공발현되었다. 형질도입 7일 후에 정량적 RT-PCR에 의해 검정하는 경우에, dCas9-KRAB를 발현하는 K562는 렌티바이러스 처리를 받지 않은 대조군 세포와 비교하여 b.  $\gamma$ -글로빈, c.  $\epsilon$ -글로빈 및 d.  $\beta$ -글로빈 유전자의 최대 95% 억제를 달성한다.

도 58은 단지 dCas9 융합을 통한 히스톤의 표적화된 후성적 변형을 위해 p300 HAT "코어"를 단리하는 것을 보여준다.

도 59는 에스. 피오게네스(*S. pyogenes*) dCas9-VP64 융합 (상단) 및 dCas9-p300 코어 융합 (하단)의 단순화된 개략도를 보여준다. 프로토스페이서 인접 모티프 (PAM)는 표적 유전자 유전자좌에 화살표로 제시되고 합성 가이드 RNA (gRNA)는 평행무늬 화살표로 제시된다.

도 60a-60c는 인간 293T 세포주 배양에서 임의의 융합된 이펙터 도메인 없이 dCas9-VP64 및 dCas9와 관련된 dCas9-p300을 사용하여 활성화의 효능을 입증하는 3개의 인간 유전자좌에서의 대표적인 데이터를 보여준다.

도 61a-61c는 dCas9 구축물의 아미노산 서열을 보여준다. 모든 도 61a-61c에 대한 범례는 도 61a에 제시된다.

도 62는 HAT-dCase9-p300 융합 단백질이 유전자 발현을 활성화시키지 못하였다는 것을 보여준다.

도 63은 gRNA가 또한 dCas9-p300 코어와 상승작용적으로 작용한다는 것을 보여준다.

도 64는 디스트로핀 유전자에 상이한 결실을 보유하는 인간 DMD 환자로부터 유래된 골격 근모세포주에서 디스트로핀의 Dp427m 골격근 이소형의 5'UTR에서의 미니디스트로핀의 TALEN 매개된 통합을 보여준다. DMD 환자 세포를 5'UTR 유전자좌에서 활성인 TALEN 쌍 및 미니디스트로핀 유전자를 보유하는 공여자 주형을 코딩하는 구축물로 전기천공하였다. (a) 미니디스트로핀을 5'UTR 내로 통합시키는 방법을 보여주는 개략도. (b) 히그로마이신-내성 클론 세포주를 단리하고, (a)에 제시된 프라이머를 사용한 5'UTR에서의 성공적인 부위-특이적 통합에 대해 PCR로 스크리닝하였다. 별표는 (c)에서의 추가의 분석에 대해 선택된 클론을 나타낸다. (c) 검출된 통합 사건을 갖는 클론 단리된 DMD 근모세포를 6일 동안 분화시키고, 미니디스트로핀의 C 말단에 융합된 HA 태그의 발현에 대해 평가하였다.

### 발명을 실시하기 위한 구체적인 내용

[0039]

본원에 기재된 바와 같은 특정 방법 및 조작된 CRISPR/CRISPR-연관 (Cas) 9-기반 시스템 조성물은 유전자 발현의 변경, 게놈 조작, 및 유전 질환에 수반되는 유전자에서의 돌연변이의 효과를 교정하거나 감소시키기에 유용한 것으로 발견되었다. CRISPR/Cas9-기반 시스템은 Cas9 단백질 및 적어도 1개의 가이드 RNA를 수반하며, 이는 그 시스템에 DNA 표적화 특이성을 제공한다. 특히, 본 개시내용은 CRISPR/Cas9-기반 시스템의 DNA 서열 표적화

기능을 추가의 활성화와 조합하여 유전자 발현 및/또는 후성적 상태의 변화를 가능하게 하는 Cas9 융합 단백질을 기재한다. 시스템은 또한 게놈 조작 및 유전자 돌연변이의 효과를 교정하거나 감소시키는데 사용될 수 있다.

[0040] 본 개시내용은 또한 CRISPR/CRISPR-연관 (Cas) 9-기반 시스템 및 다중 gRNA를 전달하여 하나 이상의 내인성 유전자를 표적화하기 위한 특정 조성물 및 방법을 제공한다. 단일 프로모터를 표적화하는 다중 sgRNA의 공동-형질감염은 상승작용적 활성화를 가능하게 하지만, 다중 플라스미드의 공동-형질감염은 카피수의 차이로 인해 각 세포에서 가변적인 발현 수준을 유도한다. 추가로, 형질감염 후의 유전자 활성화는 시간에 따른 플라스미드 DNA의 희석으로 인해 일시적이다. 더욱이, 많은 세포 유형은 용이하게 형질감염되지 않고, 일시적 유전자 발현은 치료 효과를 유도하기에 충분하지 않을 수 있다. 이들 제한을 다루기 위해, Cas9 및 독립적 프로모터로부터 최대 4개의 sgRNA를 발현하는 단일 렌티바이러스 시스템이 개발되었다. 단일 렌티바이러스 벡터로부터 Cas9 또는 dCas9 융합 단백질 및 최대 4개의 gRNA를 발현하는 플랫폼이 개시된다. 렌티바이러스 벡터는 독립적 프로모터로부터 발현된 1, 2, 3 또는 4개의 gRNA 이외에도 구성적 또는 유도성 Cas9 또는 dCas9-VP64를 발현한다. 이 시스템은 CRISPR/Cas9-기반 유전자 조절의 크기 및 시기 둘 다의 제어를 가능하게 한다. 또한, 렌티바이러스 플랫폼은 1차 세포에서 CRISPR/Cas9 시스템의 치료 적용을 촉진할 유전자 발현의 강력하고 지속적인 수준을 제공한다. 최종적으로, 이 시스템은 다중 유전자를 동시에 편집하는데, 예컨대 여러 종양유전자의 병행 녹아웃에 사용될 수 있다.

[0041] 본 개시내용은 또한 변형된 아데노-연관 바이러스 (AAV) 벡터를 사용하여 부위-특이적 뉴클레아제를 골격근 및 심장 근육으로 전달하기 위한 특정 조성물 및 방법을 제공한다. 조작될 수 있는 부위-특이적 뉴클레아제는 유전자 발현의 변경, 게놈 조작, 유전 질환에 수반되는 유전자에서의 돌연변이의 효과를 교정하거나 감소시키는 것, 또는 골격근 또는 심장 근육에 영향을 미치는 다른 상태 또는 근육 재생에 수반되는 유전자를 조작하는 것에 유용하다. 조작된 부위-특이적 뉴클레아제는 게놈 편집을 위한 아연 핑거 뉴클레아제 (ZFN), TAL 이펙터 뉴클레아제 (TALEN) 및/또는 CRISPR/Cas9 시스템을 포함할 수 있다. 본원에 기재된 바와 같은, 골격근 조직에서의 유전자는 이러한 고유한 전달 시스템을 사용하여 생체내에서 성공적으로 편집되었다. 개시된 발명은 치료 적용을 위한 인간 게놈 및 기초 과학 적용을 위한 표적 모델 종을 재기록하기 위한 수단을 제공한다.

[0042] 유전자 편집은 조직마다 달라지는 세포 주기 및 복잡한 DNA 복구 경로에 매우 의존한다. 골격근은 매우 복잡한 환경이며, 세포당 100개 초과의 핵을 갖는 큰 근섬유로 이루어져 있다. 유전자 요법 및 생물체제는 일반적으로 생체내 전달 한계에 의해 수십년간 제한되어 왔다. 이들 과제는 생체내에서의 담체의 안정성, 올바른 조직의 표적화, 충분한 유전자 발현 및 활성 유전자 산물의 수득, 및 유전자 편집 도구에서 흔한 것인 활성을 이기는 독성의 회피를 포함한다. 다른 전달 비히클, 예컨대 플라스미드 DNA의 직접 주사는 다른 맥락에서는 골격근 및 심장 근육에서 유전자를 발현하는 작용을 하지만, 게놈 편집의 검출가능한 수준을 달성하기 위한 이들 부위-특이적 뉴클레아제와는 잘 작용하지 않는다.

[0043] 많은 유전자 서열이 AAV 벡터에서 불안정하며 따라서 전달불가능하지만, 이들 부위-특이적 뉴클레아제는 놀랍게도 AAV 벡터에서 안정하다. 이들 부위-특이적 뉴클레아제가 전달되어 발현되는 경우에, 그들은 골격근 조직에서 활성인 채로 남아있다. 부위-특이적 뉴클레아제의 단백질 안정성 및 활성은 고도로 조직 유형-및 세포 유형-의존적이다. 이들 활성 및 안정한 뉴클레아제는 골격근의 복잡한 환경에서 유전자 서열을 변형시킬 수 있다. 본 발명은 효과적이고 효율적이며 성공적인 게놈 변형을 촉진하는, 골격근 또는 심장 근육으로 상기 부류의 치료제의 활성 형태를 전달하는 방식을 기재한다.

[0044] 본 개시내용은 또한 dCas9-VP64 융합체와 비교하여, 합성 전사 조절을 위한 강하고도 잠재적으로 보다 폭넓게 적용가능한 도구를 제공하는 특정 융합 후성적 이펙터 분자인 dCas9-p300 융합 단백질을 제공한다. 그의 활성화물은 시험된 모든 유전자좌에서 dCas9-VP64 융합 단백질보다 실질적으로 더 큰 정도로 유전자를 표적화한다. 또한, p300은 인간 게놈 내 인헨서에 고유 내인성 활성을 갖는다. dCas9-p300 융합 단백질은 내인성 표적 유전자 프로모터 및 인헨서 영역을 활성화/활성화시킬 수 있다.

[0045] dCas9-p300 융합 단백질은 인간 조직 배양 세포주에서 사용되어 유전자 발현을 활성화할 수 있다. 이러한 융합 단백질은 분화를 제어하고, 세포 조절을 조절하고, 혁신적인 잠재적 요법을 적용하기 위해 정확성 및 예측가능성을 가지고 인간 세포 내에서 표적 유전자좌의 후성적 상태를 지시하는데 사용될 수 있다. 현재 기술은 활성화의 강도 및 후성적 조절의 정도와 지속성에 있어서 제한되며; 장애물은 이러한 새로운 융합 단백질의 이용을 통해 제거될 수 있다.

[0046] 본 섹션에 사용된 섹션 표제 및 본원의 전체 개시내용은 단지 유기적 구조의 목적을 위한 것이며 제한하려는 것으로 의도되지 않는다.

- [0047] 1. 정의
- [0048] 달리 정의되지 않는 한, 본원에 사용된 모든 기술 과학 용어는 통상의 기술자에 의해 통상적으로 이해되는 것과 동일한 의미를 갖는다. 모순되는 경우에는, 정의를 비롯하여 본원이 우선할 것이다. 본원에 기재된 것과 유사하거나 또는 동등한 방법 및 물질이 본 발명의 실시 또는 시험에 사용될 수 있지만, 바람직한 방법 및 물질이 하기 기재된다. 본원에 언급된 모든 공개문헌, 특허 출원, 특허 및 다른 참고문헌은 그 전문이 참조로 포함된다. 본원에 개시된 물질, 방법 및 예는 단지 설명적이고, 제한하려는 것으로 의도되지 않는다.
- [0049] 본원에 사용된 용어 "포함한다", "비롯한다", "갖는", "갖는다", "일 수 있다", "함유한다" 및 그의 변형은 추가의 작용 또는 구조의 가능성을 배제하지 않는, 제약을 두지 않은 번역 어구, 용어 또는 단어인 것으로 의도된다. 단수 형태는 문맥에서 달리 명확히 기재되지 않는 한 복수 대상을 포함한다. 본 개시내용은 또한 분명히 기재되어 있든지 아니든지 본원에 제공된 실시양태 또는 요소를 "포함하는", 그로 "이루어진" 및 그로 "본질적으로 이루어진"의 다른 실시양태를 고려한다.
- [0050] 본원에서 수치 범위의 언급의 경우에, 그 사이에 개재된 숫자 각각이 동일한 정밀도로 분명히 고려된다. 예를 들어, 6-9의 범위의 경우에, 6 및 9 이외에도 7 및 8이 고려되고, 범위 6.0-7.0의 경우에, 숫자 6.0, 6.1, 6.2, 6.3, 6.4, 6.5, 6.6, 6.7, 6.8, 6.9 및 7.0이 분명히 고려된다.
- [0051] 본원에서 상호교환적으로 사용된 "아데노-연관 바이러스" 또는 "AAV"는 인간 및 일부 다른 영장류 종을 감염시키는 파르보비리다에(Parvoviridae) 과의 데펜도바이러스(Dependovirus) 속에 속하는 작은 바이러스를 지칭한다. AAV는 현재 질환을 유발하는 것으로 알려져 있지 않으며 따라서 그 바이러스는 매우 경도의 면역 반응을 유발한다.
- [0052] 본원에 사용된 "결합 영역"은 뉴클레아제에 의해 인식되고 결합되는 뉴클레아제 표적 영역 내의 영역을 지칭한다.
- [0053] 본원에서 상호교환적으로 사용된 "심장 근육" 또는 "심근육"은 심장의 벽 및 조직학적 토대에서 발견된 불수의 가로무늬근의 유형인 심근을 의미한다. 심장 근육은 심근세포 또는 심장근육세포로 구성된다. 심장근육세포는 골격근 세포 상의 가로무늬와 유사한 가로무늬를 보이지만, 다핵 골격 세포와는 달리 단지 하나의 독특한 핵을 함유한다.
- [0054] 본원에 사용된 "심장 근육 병태"는 심장 근육과 관련된 병태, 예컨대 심근병증, 심부전, 부정맥 및 염증성 심장 질환을 지칭한다.
- [0055] 본원에 사용된 "코딩 서열" 또는 "코딩하는 핵산"은 단백질을 코딩하는 뉴클레오티드 서열을 포함하는 핵산(RNA 또는 DNA 분자)을 의미한다. 코딩 서열은 핵산이 투여되는 개체 또는 포유동물의 세포에서 발현을 지시할 수 있는 프로모터 및 폴리아데닐화 신호를 포함하는 조절 요소에 작동가능하게 연결된 개시 및 종결 신호를 추가로 포함할 수 있다. 코딩 서열은 코돈 최적화될 수 있다.
- [0056] 본원에 사용된 "보체" 또는 "상보적"은 핵산이 핵산 분자의 뉴클레오티드 또는 뉴클레오티드 유사체 사이에 왓슨-크릭(Watson-Crick) (예를 들어, A-T/U 및 C-G) 또는 후그스틴(Hoogsteen) 염기 쌍형성을 이룰 수 있는 것을 의미한다. "상보성"은 2개의 핵산 서열 사이에 공유되는 특성으로, 그들이 역평행으로 정렬되는 경우에 각 위치에서의 뉴클레오티드 염기가 상보적이도록 하는 특성을 지칭한다.
- [0057] 본원에 사용된 "교정", "계놈 편집" 및 "복원"은 말단절단된 단백질을 코딩하거나 또는 단백질을 전혀 코딩하지 않는 돌연변이체 유전자를 변화시켜 전장 기능적 또는 부분적 전장 기능적 단백질 발현이 수득되도록 하는 것을 지칭한다. 돌연변이체 유전자를 교정하거나 또는 복원하는 것은 돌연변이를 갖는 유전자의 영역 또는 전체 돌연변이체 유전자를, 복구 메카니즘, 예컨대 상동성-지정 복구 (HDR)에 의해 돌연변이를 갖지 않는 유전자의 카피로 대체하는 것을 포함할 수 있다. 돌연변이체 유전자를 교정하거나 또는 복원하는 것은 또한 비-상동 말단 연결 (NHEJ)을 사용하여 나중에 복구되는 유전자에의 이중 가닥 파단 생성에 의해 조기 정지 코돈, 이상 스플라이스 수용자 부위 또는 이상 스플라이스 공여자 부위를 유발하는 프레임시프트 돌연변이를 복구하는 것을 포함할 수 있다. NHEJ는 적절한 리딩 프레임을 복원하고 조기 정지 코돈을 제거할 수 있는 복구 동안 적어도 1개의 염기 쌍을 부가하거나 또는 결실시킬 수 있다. 돌연변이체 유전자를 교정하거나 또는 복원하는 것은 또한 이상 스플라이스 수용자 부위 또는 스플라이스 공여자 서열을 파괴하는 것을 포함할 수 있다. 돌연변이체 유전자를 교정하거나 또는 복원하는 것은 또한 2개 뉴클레아제 표적 부위 사이의 DNA를 제거하고 NHEJ에 의해 DNA 파단을 복구함으로써 적절한 리딩 프레임을 복원하기 위해 동일한 DNA 가닥에 대한 2개 뉴클레아제의 동시 작용에 의해



비-필수 유전자 절편을 결실시키는 것을 포함할 수 있다.

- [0058] 본원에서 상호교환적으로 사용된 "공여자 DNA", "공여자 주형" 및 "복구 주형"은 관심 유전자의 적어도 일부분을 포함하는 이중-가닥 DNA 단편 또는 분자를 지칭한다. 공여자 DNA는 완전-기능적 단백질 또는 부분-기능적 단백질을 코딩할 수 있다.
- [0059] 본원에서 상호교환적으로 사용된 "뒤시엔느 근육 이영양증" 또는 "DMD"는 근육 변성 및 최후 사망을 초래하는 열성 치사 X-연관 장애를 지칭한다. DMD는 통상적인 유전성 단일유전자 질환이며 3500명의 남성 중 1명에서 발생한다. DMD는 디스트로핀 유전자에서 넨센스 또는 프레임 시프트 돌연변이를 유발하는 유전성 또는 자발적 돌연변이의 결과이다. DMD를 유발하는 디스트로핀 돌연변이의 대부분은 디스트로핀 유전자에서 리딩 프레임을 파괴하고 조기 번역 종결을 유발하는 엑손의 결실이다. DMD 환자는 전형적으로 소아기 동안 그들 자신을 신체적으로 지탱할 능력을 상실하고, 10대 동안 진행하여 점점 더 약해지며, 20대에 사망한다.
- [0060] 본원에 사용된 "디스트로핀"은 세포 막을 통해 주위 세포외 매트릭스에 근섬유의 세포골격을 연결하는 단백질 복합체의 일부인 막대-형상 세포질 단백질을 지칭한다. 디스트로핀은 근육 세포 완전성 및 기능의 조절을 담당하는 세포 막의 디스트로글리칸 복합체에 구조적 안정성을 제공한다. 본원에서 상호교환적으로 사용된 디스트로핀 유전자 또는 "DMD 유전자"는 유전자좌 Xp21에서 2.2 메가염기이다. 1차 전사는 약 2,400 kb로 측정되며, 이때 성숙 mRNA는 약 14 kb이다. 79개 엑손은 3500개 아미노산을 초과하는 단백질을 코딩한다.
- [0061] 본원에 사용된 "엑손 51"은 디스트로핀 유전자의 51번째 엑손을 지칭한다. 엑손 51은 빈번하게는 DMD 환자에서 프레임-파괴 결실에 인접해 있으며 올리고뉴클레오타이드-기반 엑손 스킵핑에 대한 임상 시험에서 표적화되었다. 엑손 51 스킵핑 화합물인 에테플러센에 대한 임상 시험은 최근에 48주에 걸쳐 유의한 기능적 이익을 보고하였으며, 기준선과 비교하여 평균 47% 디스트로핀 양성 섬유를 나타냈다. 엑손 51에서의 돌연변이는 NHEJ-기반 게놈 편집에 의한 영구적 교정에 이상적으로 적합하다.
- [0062] 본원에서 상호교환적으로 사용된 "프레임시프트" 또는 "프레임시프트 돌연변이"는 1개 이상의 뉴클레오타이드의 부가 또는 결실이 mRNA에서 코돈의 리딩 프레임의 변화를 유발하는 유전자 돌연변이의 유형을 지칭한다. 리딩 프레임의 변화는 단백질 번역 시 아미노산 서열의 변경, 예컨대 미스센스 돌연변이 또는 조기 정지 코돈을 유도할 수 있다.
- [0063] 본원에 사용된 "기능적" 및 "완전-기능적"은 생물학적 활성을 갖는 단백질을 기재한다. "기능적 유전자"는 기능적 단백질로 번역되는 mRNA로 전사되는 유전자를 지칭한다.
- [0064] 본원에 사용된 "융합 단백질"은 본래 별개의 단백질을 코딩하는 2개 이상의 유전자의 연결을 통해 생성된 키메라 단백질을 지칭한다. 융합 유전자의 번역은 각각의 본래 단백질로부터 유래된 기능적 특성을 갖는 단일 폴리펩티드를 생성한다.
- [0065] 본원에 사용된 "유전적 구축물"은 단백질을 코딩하는 뉴클레오타이드 서열을 포함하는 DNA 또는 RNA 분자를 지칭한다. 코딩 서열은 핵산이 투여되는 개체의 세포에서 발현을 지시할 수 있는 프로모터 및 폴리아데닐화 신호를 포함하는 조절 요소에 작동가능하게 연결된 개시 및 종결 신호를 포함한다. 본원에 사용된 용어 "발현가능한 형태"는 단백질을 코딩하는 코딩 서열에 작동가능하게 연결된 필요한 조절 요소를 함유하여 상기 코딩 서열이 개체의 세포에 존재하는 경우에 발현되도록 하는 유전자 구축물을 지칭한다.
- [0066] 본원에 사용된 "유전 질환"은 게놈에서 하나 이상의 이상에 의해 부분적으로 또는 완전히, 직접적으로 또는 간접적으로 유발되는 질환, 특히 출생시부터 존재하는 병태를 지칭한다. 이상은 돌연변이, 삽입 또는 결실일 수 있다. 이상은 유전자의 코딩 서열 또는 그의 조절 서열에 영향을 미칠 수 있다. 유전 질환은 DMD, 혈우병, 낭성 섬유증, 헌팅톤 무도병, 가족성 고콜레스테롤혈증 (LDL 수용체 결함), 간모세포종, 윌슨병, 선천성 간 포르피린증, 간 대사의 유전성 장애, 레쉬 니한 증후군, 겸상 적혈구성 빈혈, 지중해 빈혈, 색소성 건피증, 판코니 빈혈, 색소성 망막염, 모세혈관확장성 운동실조, 블루움 증후군, 망막모세포종 및 테이-삭스병일 수 있으나, 이에 제한되지는 않는다.
- [0067] 본원에서 상호교환적으로 사용된 "상동성-지정 복구" 또는 "HDR"은 DNA의 상동 조각이 대부분 세포 주기의 G2 및 S 기에서 핵에 존재하는 경우에 이중 가닥 DNA 병변을 복구하기 위한 세포에서의 메커니즘을 지칭한다. HDR은 복구를 가이드하기 위한 공여자 DNA 주형을 사용하며, 전체 유전자의 표적화된 부가를 비롯한 특정한 서열 변화를 게놈에 생성하는데 사용될 수 있다. 공여자 주형이 부위 특이적 뉴클레아제와 함께, 예컨대 CRISPR/Cas9-기반 시스템과 함께 제공된다면, 세포 기구는 상동 재조합에 의해 파단을 복구할 것이며, 이는 DNA 절단의 존재 하에 여러 자릿수 증진된다. 상동 DNA 조각이 부재하는 경우에, 비-상동 말단 연결이 그 대신에

일어날 수 있다.

- [0068] 본원에 사용된 "게놈 편집"은 유전자를 변화시키는 것을 지칭한다. 게놈 편집은 돌연변이체 유전자를 교정하거나 또는 복원하는 것을 포함할 수 있다. 게놈 편집은 유전자, 예컨대 돌연변이체 유전자 또는 정상 유전자를 녹아웃시키는 것을 포함할 수 있다. 게놈 편집은 관심 유전자를 변화시킴으로써 질환을 치료하거나 또는 근육 복구를 증진시키는데 사용될 수 있다.
- [0069] 본원에 사용된 "동일한" 또는 "동일성"은 2개 이상의 핵산 또는 폴리펩티드 서열과 관련하여 서열이 특정된 영역에 걸쳐 동일한 특정된 핵분열의 잔기를 갖는 것을 의미한다. 핵분열은, 2개의 서열을 최적으로 정렬시키고, 특정된 영역에 걸쳐 2개의 서열을 비교하고, 두 서열에서 동일한 잔기가 발생하는 위치의 개수를 결정하여 매칭되는 위치의 개수를 산출하고, 매칭되는 위치의 개수를 특정된 영역 내의 위치의 전체 개수로 나누고, 결과에 100을 곱하여 서열 동일성의 핵분열을 산출함으로써 계산될 수 있다. 2개의 서열이 상이한 길이를 가지거나 또는 정렬이 하나 이상의 스테거형 말단을 생산하고, 특정된 비교 영역이 단일 서열만을 포함하는 경우에, 단일 서열의 잔기는 계산의 분모에는 포함되나 분자에는 포함되지 않는다. DNA 및 RNA를 비교하는 경우에, 티민 (T) 및 우라실 (U)은 동등한 것으로 간주될 수 있다. 동일성은 수동으로 또는 컴퓨터 서열 알고리즘, 예컨대 블라스트(BLAST) 또는 블라스트 2.0을 사용하여 수행될 수 있다.
- [0070] 본원에서 상호교환적으로 사용된 "돌연변이체 유전자" 또는 "돌연변이된 유전자"는 검출가능한 돌연변이를 겪은 유전자를 지칭한다. 돌연변이체 유전자는 유전 물질의 변화, 예컨대 상실, 획득 또는 교환을 겪으며, 이는 유전자의 정상 전달 및 발현에 영향을 미친다. 본원에 사용된 "파괴된 유전자"는 조기 정지 코돈을 유발하는 돌연변이를 갖는 돌연변이체 유전자를 지칭한다. 파괴된 유전자 산물은 파괴되지 않은 전장 유전자 산물에 비해 말단절단되어 있다.
- [0071] 본원에 사용된 "비-상동 말단 연결 (NHEJ) 경로"는 상동 주형에 대한 필요 없이 파단 말단을 직접적으로 라이게이션함으로써 DNA에서의 이중-가닥 파단을 복구하는 경로를 지칭한다. NHEJ에 의한 DNA 말단의 주형-독립적 재-라이게이션은 DNA 파단 지점에 무작위 미세-삽입 및 미세-결실 (indel)을 도입하는 확률적 오류-유발 복구 과정이다. 이 방법을 사용하여 표적화된 유전자 서열의 리딩 프레임의 의도적으로 파괴하거나, 결실시키거나 또는 변경할 수 있다. NHEJ는 전형적으로 복구를 가이드하기 위한 미세상동성으로 칭해지는 짧은 상동 DNA 서열을 사용한다. 이들 미세상동성은 종종 이중-가닥 파단의 말단 상에 단일-가닥 오버행으로 존재한다. 오버행이 완벽하게 적합성인 경우에, NHEJ는 통상 파단을 정확하게 복구하지만, 뉴클레오타이드의 상실을 유도하는 부정확한 복구도 또한 발생할 수 있으며, 이는 오버행이 적합성이 아닌 경우에 훨씬 더 일반적이다.
- [0072] 본원에 사용된 "정상 유전자"는 유전 물질의 변화, 예컨대 상실, 획득 또는 교환을 겪지 않은 유전자를 지칭한다. 정상 유전자는 정상 유전자 전달 및 유전자 발현을 겪는다.
- [0073] 본원에 사용된 "뉴클레아제 매개된 NHEJ"은 뉴클레아제, 예컨대 cas9가 이중 가닥 DNA를 컷팅한 후에 개시되는 NHEJ를 지칭한다.
- [0074] 본원에 사용된 "핵산" 또는 "올리고뉴클레오타이드" 또는 "폴리뉴클레오타이드"는 함께 공유적으로 연결된 적어도 2개의 뉴클레오타이드를 의미한다. 단일 가닥의 설명은 또한 상보적 가닥의 서열을 정의한다. 따라서, 핵산은 또한 설명된 단일 가닥의 상보적 가닥을 포괄한다. 핵산의 많은 변이체는 주어진 핵산과 동일한 목적에 사용될 수 있다. 따라서, 핵산은 또한 실질적으로 동일한 핵산 및 그의 보체를 포괄한다. 단일 가닥은 엄격한 혼성화 조건 하에서 표적 서열에 혼성화할 수 있는 프로브를 제공한다. 따라서, 핵산은 또한 엄격한 혼성화 조건 하에서 혼성화하는 프로브를 포괄한다.
- [0075] 핵산은 단일 가닥 또는 이중 가닥일 수 있거나, 또는 이중 가닥 및 단일 가닥 서열 둘 다의 부분을 함유할 수 있다. 핵산은 DNA, 게놈 및 cDNA 둘 다, RNA 또는 하이브리드일 수 있으며, 여기서 핵산은 데옥시리보- 및 리보-뉴클레오타이드의 조합, 및 우라실, 아데닌, 티민, 시토신, 구아닌, 이노신, 크산틴 하이포크산틴, 이소시토신 및 이소구아닌을 비롯한 염기의 조합을 포함할 수 있다. 핵산은 화학적 합성 방법에 의해 또는 재조합 방법에 의해 수득될 수 있다.
- [0076] 본원에 사용된 "작동가능하게 연결된" 유전자의 발현이 공간적으로 연결된 프로모터의 제어 하에 있다는 것을 의미한다. 프로모터는 그의 제어 하의 유전자의 5' (상류) 또는 3' (하류)에 위치할 수 있다. 프로모터와 유전자 사이의 거리는, 프로모터가 유래된 유전자에서 그 프로모터와 그가 제어하는 유전자 사이의 거리와 대략 동일할 수 있다. 관련 기술분야에 공지된 바와 같이, 이러한 거리의 변화는 프로모터 기능의 상실 없이 수용될 수 있다.

- [0077] 본원에 사용된 "부분-기능적"은 돌연변이체 유전자에 의해 코딩되며 기능적 단백질보다 더 적은 생물학적 활성을 갖지만 비-기능적 단백질보다는 더 많은 생물학적 활성을 갖는 단백질을 기재한다.
- [0078] 본원에서 상호교환적으로 사용된 "조기 정지 코돈" 또는 "아웃-오브-프레임 정지 코돈"은 야생형 유전자에서 통상적으로 발견되지 않는 위치에 정지 코돈을 생성하는 DNA 서열에서의 넌센스 돌연변이를 지칭한다. 조기 정지 코돈은 단백질이 단백질의 전장 버전과 비교하여 말단절단되거나 또는 더 짧아지도록 할 수 있다.
- [0079] 본원에 사용된 "프로모터"는 세포에서 핵산의 발현을 부여하거나, 활성화시키거나 또는 증진시킬 수 있는 합성 또는 자연-유래 분자를 의미한다. 프로모터는 추가로 발현을 증진시키고/거나 그의 공간적 발현 및/또는 시간적 발현을 변경하는 하나 이상의 특정한 전사 조절 서열을 포함할 수 있다. 프로모터는 또한 원위 인핸서 또는 리프레서 요소를 포함할 수 있으며, 이는 전사의 개시 부위로부터 수천개 염기 쌍만큼 떨어진 곳에 위치할 수 있다. 프로모터는 바이러스, 박테리아, 진균, 식물, 곤충 및 동물을 비롯한 공급원으로부터 유래될 수 있다. 프로모터는 발현이 일어나는 세포, 조직 또는 기관에 대하여, 또는 발현이 일어나는 발달 단계에 대하여, 또는 생리학적 스트레스, 병원체, 금속 이온 또는 유도 작용제와 같은 외부 자극에 반응하여, 유전자 성분의 발현을 구성적으로 또는 차등적으로 조절할 수 있다. 프로모터의 대표적인 예는 박테리오파지 T7 프로모터, 박테리오파지 T3 프로모터, SP6 프로모터, lac 오퍼레이터-프로모터, tac 프로모터, SV40 후기 프로모터, SV40 초기 프로모터, RSV-LTR 프로모터, CMV IE 프로모터, SV40 초기 프로모터 또는 SV40 후기 프로모터 및 CMV IE 프로모터를 포함한다.
- [0080] 본원에서 상호교환적으로 사용된 "반복 가변 이잔기" 또는 "RVD"는 TALE DNA-결합 도메인의 33-35개 아미노산을 포함하는 DNA 인식 모티프 (또한 "RVD 모듈"로 알려짐) 내 한 쌍의 인접한 아미노산 잔기를 지칭한다. RVD는 RVD 모듈의 뉴클레오티드 특이성을 결정한다. RVD 모듈은 조합되어 RVD 어레이를 생성할 수 있다. 본원에 사용된 "RVD 어레이 길이"는 TALEN에 의해 인식되는 TALEN 표적 영역, 즉 결합 영역 내 뉴클레오티드 서열의 길이에 상응하는 RVD 모듈의 개수를 지칭한다.
- [0081] 본원에 사용된 "부위-특이적 뉴클레아제"는 DNA 서열을 특이적으로 인식하여 절단할 수 있는 효소를 지칭한다. 부위-특이적 뉴클레아제는 조작될 수 있다. 조작된 부위-특이적 뉴클레아제의 예는 아연 핑거 뉴클레아제 (ZFN), TAL 이펙터 뉴클레아제 (TALEN) 및 CRISPR/Cas9-기반 시스템을 포함한다.
- [0082] 본원에 사용된 "골격근"은 체성 신경계의 제어 하에 있으며 건으로 알려진 콜라겐 섬유 다발에 의해 골에 부착되어 있는 가로무늬근의 유형을 지칭한다. 골격근은 근세포 또는 "근육 세포"로 알려져 있으며 때때로 구어체로 "근섬유"로 칭해지는 개별 성분으로 구성된다. 근세포는 근발생으로 알려진 과정에서 발달 근모세포 (근육 세포를 야기하는 배아 전구 세포의 유형)의 융합으로부터 형성된다. 이들 긴, 원통형, 다핵 세포는 또한 근섬유로 칭해진다.
- [0083] 본원에 사용된 "골격근 병태" 골격근과 관련된 병태, 예컨대 근육 이영양증, 노화, 근육 변성, 상처 치유 및 근육 약화 또는 위축을 지칭한다.
- [0084] 본원에서 상호교환적으로 사용된 "스페이스" 및 "스페이스 영역"은 2개의 TALEN 또는 ZFN에 대한 결합 영역 사이에 있는, 그러나 그의 일부는 아닌, TALEN 또는 ZFN 표적 영역 내의 영역을 지칭한다.
- [0085] 본원에서 상호교환적으로 사용된 "대상체" 및 "환자"는 포유동물 (예를 들어, 소, 돼지, 낙타, 라마, 말, 염소, 토끼, 양, 햄스터, 기니 피그, 고양이, 개, 래트 및 마우스, 비-인간 영장류 (예를 들어, 원숭이, 예컨대 시노몰구스 또는 레서스 원숭이, 침팬지 등) 및 인간)을 포함하나 이에 제한되지는 않는 임의의 척추동물을 지칭한다. 일부 실시양태에서, 대상체는 인간 또는 비-인간일 수 있다. 대상체 또는 환자는 다른 형태의 처리를 겪을 수 있다.
- [0086] 본원에 사용된 "표적 유전자"는 공지 또는 추정 유전자 산물을 코딩하는 임의의 뉴클레오티드 서열을 지칭한다. 표적 유전자는 유전 질환에 수반되는 돌연변이된 유전자일 수 있다.
- [0087] 본원에 사용된 "표적 영역"은 부위-특이적 뉴클레아제가 결합하여 절단하도록 설계된 표적 유전자의 영역을 지칭한다.
- [0088] 본원에 사용된 "전사 활성화제-유사 이펙터" 또는 "TALE"는 특정한 DNA 서열을 인식하여 그에 결합하는 단백질 구조를 지칭한다. "TALE DNA-결합 도메인"은, 각각이 DNA의 단일 염기 쌍을 특이적으로 인식하는 RVD 모듈로도 알려진 직렬 33-35개 아미노산 반복부의 어레이를 포함하는 DNA-결합 도메인을 지칭한다. RVD 모듈은 정의된 서열을 인식하는 어레이로 조립되기 위해 임의의 순서로 배열될 수 있다.

- [0089] TALE DNA-결합 도메인의 결합 특이성은 RVD 어레이에 이어서 20개 아미노산의 단일 말단절단 반복부에 의해 결정된다. TALE DNA-결합 도메인은, 각각이 RVD를 함유하며 DNA의 단일 염기 쌍을 인식하는 12 내지 27개 RVD 모듈을 가질 수 있다. 각각의 4개의 가능한 DNA 뉴클레오티드 (A, T, C 및 G)를 인식하는 특정한 RVD가 확인되었다. TALE DNA-결합 도메인은 모듈식이기 때문에, 4개의 상이한 DNA 뉴클레오티드를 인식하는 반복부는 함께 연결되어 임의의 특정한 DNA 서열을 인식할 수 있다. 이어서, 이들 표적화된 DNA-결합 도메인은 촉매 도메인과 조합되어 인공 전사 인자, 메틸트랜스퍼라제, 인테그라제, 뉴클레아제 및 레코미니아제를 비롯한 기능적 효소를 생성할 수 있다.
- [0090] 본원에서 상호교환적으로 사용된 "전사 활성화제-유사 이펙터 뉴클레아제" 또는 "TALEN"은 뉴클레아제, 예컨대 엔도뉴클레아제 FokI의 촉매 도메인, 및 맞춤 DNA 서열에 표적화될 수 있는 설계된 TALE DNA-결합 도메인의 조합된 융합 단백질을 지칭한다. "TALEN 단량체"는 촉매 뉴클레아제 도메인 및 설계된 TALE DNA-결합 도메인을 갖는 조합된 융합 단백질을 지칭한다. 2개의 TALEN 단량체가 TALEN 표적 영역을 표적화하여 절단하도록 설계될 수 있다.
- [0091] 본원에 사용된 "트랜스진"은 하나의 유기체로부터 단리되었으며 상이한 유기체에 도입되는 유전자 서열을 함유하는 유전자 또는 유전 물질을 지칭한다. 이러한 DNA의 비-천연 절편은, 트랜스제닉 유기체에 RNA 또는 단백질을 생산하는 능력을 유지할 수 있거나, 또는 트랜스제닉 유기체의 유전자 코드의 정상 기능을 변경할 수 있다. 트랜스진의 도입은 유기체의 표현형을 변화시키는 잠재력을 갖는다.
- [0092] 핵산에 대하여 본원에 사용된 "변이체"는 (i) 언급된 뉴클레오티드 서열의 부분 또는 단편; (ii) 언급된 뉴클레오티드 서열 또는 그의 부분의 보체; (iii) 언급된 핵산 또는 그의 보체와 실질적으로 동일한 핵산; 또는 (iv) 엄격한 조건 하에서 언급된 핵산, 그의 보체 또는 그와 실질적으로 동일한 서열에 혼성화하는 핵산을 의미한다.
- [0093] 펩티드 또는 폴리펩티드에 대하여 "변이체"는 아미노산의 삽입, 결실 또는 보존적 치환에 의해 아미노산 서열에 있어 상이하지만 적어도 하나의 생물학적 활성을 유지한다. 변이체는 또한 언급된 단백질과 실질적으로 동일한 아미노산 서열을 가지며 적어도 하나의 생물학적 활성을 유지하는 아미노산 서열을 갖는 단백질을 의미할 수 있다. 아미노산의 보존적 치환, 즉 아미노산을 유사한 특성 (예를 들어, 하전된 영역의 친수성, 정도 및 분포)의 상이한 아미노산으로 대체하는 것은 관련 기술분야에서 전형적으로 작은 변화를 수반하는 것으로 인식된다. 이들 작은 변화는 부분적으로는, 관련 기술분야에서 이해된 바와 같이 아미노산의 소수친수성 지수를 고려함으로써 확인될 수 있다. 문헌 [Kyte et al., J. Mol. Biol. 157:105-132 (1982)]. 아미노산의 소수친수성 지수는 그의 소수친수성 및 전하의 고찰을 기반으로 한다. 유사한 소수친수성 지수의 아미노산이 치환될 수 있으며 여전히 단백질 기능을 유지할 수 있다는 것은 관련 기술분야에 공지되어 있다. 한 측면에서,  $\pm 2$ 의 소수친수성 지수를 갖는 아미노산이 치환된다. 아미노산의 친수성은 또한 생물학적 기능을 유지하는 단백질을 생성할 치환을 밝히는데 사용될 수 있다. 펩티드와 관련하여 아미노산의 친수성의 고찰은 그 펩티드의 최대 국부 평균 친수성의 계산을 가능하게 한다. 치환은 서로  $\pm 2$  내의 친수성 값을 갖는 아미노산으로 수행될 수 있다. 아미노산의 소수성 지수 및 친수성 값 둘 다는 그 아미노산의 특정한 측쇄에 의해 영향을 받는다. 그 관찰과 일치하게, 생물학적 기능과 적합성인 아미노산 치환은, 소수성, 친수성, 전하, 크기 및 다른 특성에 의해 밝혀진 바와 같이, 아미노산의 상대 유사성, 및 특히 상기 아미노산의 측쇄에 좌우되는 것으로 이해된다.
- [0094] 본원에 사용된 "벡터"는 복제 기점을 함유하는 핵산 서열을 의미한다. 벡터는 바이러스 벡터, 박테리오파지, 박테리아 인공 염색체 또는 효모 인공 염색체일 수 있다. 벡터는 DNA 또는 RNA 벡터일 수 있다. 벡터는 자기-복제 염색체외 벡터일 수 있고, 바람직하게는 DNA 플라스미드이다. 예를 들어, 벡터는 서열 1의 아미노산 서열을 포함하는 iCas9-VP64 융합 단백질을 또는 서열 5-40, 65-144, 492-515, 540-563 및 585-625 중 어느 하나의 적어도 1개의 gRNA 뉴클레오티드 서열을 코딩할 수 있다. 대안적으로, 벡터는 Cas9 및 서열 5-40, 65-144, 492-515, 540-563 및 585-625 중 어느 하나의 적어도 1개의 gRNA 뉴클레오티드 서열을 코딩할 수 있다.
- [0095] 본원에 사용된 "아연 핑거"는 DNA 서열을 인식하고 그에 결합하는 단백질 구조를 지칭한다. 아연 핑거 도메인은 인간 프로테오미에서 가장 통상적인 DNA-결합 모티프이다. 단일 아연 핑거는 대략 30개 아미노산을 함유하고, 도메인은 전형적으로 염기 쌍당 단일 아미노산 측쇄의 상호작용을 통해 DNA의 3개 연속적 염기 쌍에 결합함으로써 기능한다.
- [0096] 본원에서 상호교환적으로 사용된 "아연 핑거 뉴클레아제" 또는 "ZFN"은 적어도 1개의 뉴클레아제 또는 완전히 조합되었을 때 DNA를 절단할 수 있는 뉴클레아제의 일부에 효과적으로 연결된 적어도 1개의 아연 핑거 DNA 결합 도메인을 포함하는 키메라 단백질을 지칭한다.



- [0097] 본원에서 달리 정의되지 않는 한, 본 개시내용과 관련하여 사용된 과학 기술 용어는 통상의 기술자가 통상적으로 이해하는 의미를 가질 것이다. 예를 들어, 본원에 기재된 세포 및 조직 배양, 분자 생물학, 면역학, 미생물학, 유전학 및 단백질 및 핵산 화학 및 혼성화와 관련하여 사용된 임의의 명명법 및 이의 기술은 관련 기술분야에 널리 공지되어 있으며 통상적으로 사용되는 것들이다. 용어의 의미 및 범주는 명확하여야 하나; 임의의 잠재적 모호성의 사례에서, 본원에 제공된 정의가 임의의 사전 또는 외부 정의보다 선행한다. 또한, 문맥에서 달리 요구되지 않는 한, 단수 용어는 복수대상을 포함할 것이고, 복수 용어는 단수대상을 포함할 것이다.
- [0098] 2. 게놈 편집을 위한 조성물
- [0099] 본 발명은 게놈 편집, 게놈 변경 또는 표적 유전자의 유전자 발현의 변경을 위한 조성물에 관한 것이다. 조성물은 바이러스 벡터 및 융합 단백질, 예컨대 부위-특이적 뉴클레아제 또는 CRISPR/Cas9-시스템과 적어도 1개의 gRNA를 포함할 수 있다.
- [0100] a. 근육에서의 게놈 편집을 위한 조성물
- [0101] 본 발명은 대상체의 골격근 또는 심장 근육에서의 표적 유전자를 게놈 편집하기 위한 조성물에 관한 것이다. 조성물은 부위-특이적 뉴클레아제를 코딩하는 변형된 AAV 벡터 및 뉴클레오티드 서열을 포함한다. 조성물은 부위-특이적 뉴클레아제의 활성 형태를 골격근 또는 심장 근육으로 전달한다. 조성물은 공여자 DNA 또는 트랜스진을 추가로 포함할 수 있다. 이들 조성물은 게놈 편집, 게놈 조작, 및 유전 질환 및/또는 다른 골격 또는 심장 근육 병태에 수반되는 유전자에서의 돌연변이의 효과를 교정하거나 또는 감소시키는데 사용될 수 있다.
- [0102] 표적 유전자는 유전자의 활성화, 억제 또는 파괴가 요구될 수 있는 세포의 분화 또는 임의의 다른 과정에 수반될 수 있거나, 또는 돌연변이, 예컨대 결실, 프레임시프트 돌연변이 또는 넌센스 돌연변이를 가질 수 있다. 표적 유전자가 조기 정지 코돈, 이상 스플라이스 수용자 부위 또는 이상 스플라이스 공여자 부위를 유발하는 돌연변이를 갖는 경우에, 부위-특이적 뉴클레아제는 조기 정지 코돈, 이상 스플라이스 수용자 부위 또는 이상 스플라이스 공여자 부위로부터의 상류 또는 하류 뉴클레오티드 서열을 인식하고 그에 결합하도록 설계될 수 있다. 부위-특이적 뉴클레아제는 또한 스플라이스 수용자 및 공여자를 표적화함으로써 정상 유전자 스플라이싱을 파괴하는데 사용되어 조기 정지 코돈의 스킵핑을 유도하거나 또는 파괴된 리딩 프레임을 복원할 수 있다. 부위-특이적 뉴클레아제는 게놈의 단백질-코딩 영역에 오프-타겟 변화를 매개할 수 있거나 또는 매개하지 않을 수 있다.
- [0103] 3. CRISPR 시스템
- [0104] 본원에서 상호교환적으로 사용된 "클러스터링된 규칙적 간격의 짧은 회문식 반복부" 및 "CRISPR"은 대략 서열분석된 박테리아의 40% 및 서열분석된 고세균의 90%의 게놈에서 발견되는 다중의 짧은 직접 반복부를 함유하는 유전자좌를 지칭한다. CRISPR 시스템은 후천성 면역의 형태를 제공하는, 침입 파지 및 플라스미드에 대한 방어에 수반되는 미생물 뉴클레아제 시스템이다. 미생물 숙주에서의 CRISPR 유전자좌는 CRISPR-연관 (Cas) 유전자의 조합 뿐만 아니라 CRISPR-매개 핵산 절단의 특이성을 프로그램화할 수 있는 비-코딩 RNA 요소를 함유한다. 스페이서로 칭해지는 외래 DNA의 짧은 절편이 CRISPR 반복부 사이의 게놈 내로 통합되며 지난 노출의 '메모리'로서의 역할을 한다. Cas9는 sgRNA의 3' 말단과 복합체를 형성하고, 단백질-RNA 쌍은 sgRNA 서열의 5' 말단과 프로토스페이서로 알려져 있는 미리 정의된 20 bp DNA 서열 사이의 상보적 염기 쌍형성에 의해 그의 게놈 표적을 인식한다. 이러한 복합체는 crRNA 내의 코딩된 영역, 즉 프로토스페이서, 및 병원체 게놈 내의 프로토스페이서-인접 모티프 (PAM)를 통해 병원체 DNA의 상동 유전자좌에 지시된다. 비-코딩 CRISPR 어레이는 직접 반복부 내에서 전사 및 절단되어 개별 스페이서 서열을 함유하는 짧은 crRNA가 되고, 이는 Cas 뉴클레아제를 표적 부위 (프로토스페이서)로 지시한다. 발현된 sgRNA의 20 bp 인식 서열을 간단하게 교환함으로써, Cas9 뉴클레아제는 새로운 게놈 표적에 지시될 수 있다. CRISPR 스페이서는 진핵 유기체에서 RNAi와 유사한 방식으로 외인성 유전 요소를 인식하고 침묵화하는데 사용된다.
- [0105] CRISPR 시스템의 3개 유형 (유형 I, II 및 III 이펙터 시스템)이 공지되어 있다. 유형 II 이펙터 시스템은 dsDNA를 절단하기 위한 단일 이펙터 효소인 Cas9를 사용하여, 4개의 순차적 단계로 표적화된 DNA 이중 가닥 파단을 수행한다. 복합체로서 작용하는 다중의 별개의 이펙터를 요구하는 유형 I 및 유형 III 이펙터 시스템과 비교하여, 유형 II 이펙터 시스템은 대안적인 맥락에서, 예컨대 진핵 세포에서 기능할 수 있다. 유형 II 이펙터 시스템은 스페이서-함유 CRISPR 유전자좌로부터 전사되는 긴 프리-crRNA, Cas9 단백질, 및 프리-crRNA 프로세싱에 수반되는 tracrRNA로 이루어진다. tracrRNA는 프리-crRNA의 스페이서를 분리하는 반복 영역에 혼성화하여, 내인성 RNase III에 의한 dsRNA 절단을 개시한다. 이러한 절단에 이어서, 각 스페이서 내에서 Cas9에 의한

제2 절단이 수행되어, tracrRNA 및 Cas9와 회합된 채로 남아있는 성숙 crRNA를 생성함으로써, Cas9:crRNA-tracrRNA 복합체가 형성된다.

[0106] Cas9:crRNA-tracrRNA 복합체는 DNA 듀플렉스를 풀고, crRNA와 매칭되는 서열을 찾아 절단한다. 표적 인식은 표적 DNA 내의 "프로토스페이서" 서열과 crRNA 내의 나머지 스페이서 서열 사이의 상보성의 검출 시에 발생한다. Cas9는 정확한 프로토스페이서-인접 모티프 (PAM)가 또한 프로토스페이서의 3' 말단에 존재하는 경우에 표적 DNA의 절단을 매개한다. 프로토스페이서 표적화를 위해, 서열에 바로 이어서, DNA 절단에 요구되는 Cas9 뉴클레아제에 의해 인식되는 짧은 서열인 프로토스페이서-인접 모티프 (PAM)가 존재하여야 한다. 상이한 유형 II 시스템은 상이한 PAM 요건을 가진다. 예스. 피오게네스 CRISPR 시스템은 5'-NRG-3' (여기서 R은 A 또는 G임)으로서 이 Cas9 (SpCas9)에 대한 PAM 서열을 가질 수 있으며, 인간 세포에서 이 시스템의 특이성을 특성화하였다. CRISPR/Cas9 시스템의 고유한 능력은 단일 Cas9 단백질을 2개 이상의 sgRNA와 공-발현시킴으로써 다중의 별개의 게놈 유전자좌를 동시에 표적화하는 간단한 능력이다. 예를 들어, 스트렙토코쿠스 피오게네스(*Streptococcus pyogenes*) 유형 II 시스템은 자연에서 "NGG" 서열 (여기서 "N"은 임의의 뉴클레오티드일 수 있음)을 사용하는 것을 선호하지만, 또한 조작된 시스템에서 다른 PAM 서열, 예컨대 "NAG"를 수용한다 (Hsu et al., Nature Biotechnology (2013) doi:10.1038/nbt.2647). 유사하게, 정상적으로 네이세리아 메닝기티디스(*Neisseria meningitidis*)로부터 유래된 Cas9 (NmCas9)는 통상적으로 NNNNGATT의 천연 PAM을 갖지만, 고도로 축중성인 NNNNGNNN PAM을 비롯한 다양한 PAM에 걸쳐 활성을 갖는다 (Esvelt et al. Nature Methods (2013) doi:10.1038/nmeth.2681).

#### [0107] 4. CRISPR/Cas9-기반 시스템

[0108] 스트렙토코쿠스 피오게네스의 유형 II 이펙터 시스템의 조작된 형태는 게놈 조작을 위한 인간 세포에서 기능을 하는 것으로 제시되었다. 이 시스템에서, Cas9 단백질은, 일반적으로 RNase III 및 crRNA 프로세싱에 대한 필요를 제거하는 crRNA-tracrRNA 융합체인 합성적으로 재구성된 "가이드 RNA" ("gRNA", 또한 키메라 단일 가이드 RNA ("sgRNA")로서 본원에서 상호교환적으로 사용됨)에 의해 게놈 표적 부위에 지시되었다 (도 53A 참조). 게놈 편집 및 유전 질환의 치료에 사용하기 위한 CRISPR/Cas9-기반 조작 시스템이 본원에 제공된다. CRISPR/Cas9-기반 조작 시스템은 유전 질환, 노화, 조직 재생 또는 상처 치유에 수반되는 유전자를 비롯한 임의의 유전자를 표적화하도록 설계될 수 있다. CRISPR/Cas9-기반 시스템은 Cas9 단백질 또는 Cas9 융합 단백질 및 적어도 1개의 gRNA를 포함할 수 있다. Cas9 융합 단백질은, 예를 들어 Cas9에 대해 내인성인 것과 상이한 활성을 갖는 도메인, 예컨대 전사활성화 도메인을 포함할 수 있다.

[0109] 표적 유전자는 유전자의 활성화가 요구될 수 있는 세포의 분화 또는 임의의 다른 과정에 수반될 수 있거나, 또는 돌연변이, 예컨대 프레임시프트 돌연변이 또는 넨센스 돌연변이를 가질 수 있다. 표적 유전자가 조기 정지 코돈, 이상 스플라이스 수용자 부위 또는 이상 스플라이스 공여자 부위를 유발하는 돌연변이를 갖는 경우에, CRISPR/Cas9-기반 시스템은 조기 정지 코돈, 이상 스플라이스 수용자 부위 또는 이상 스플라이스 공여자 부위로 부터의 상류 또는 하류 뉴클레오티드 서열을 인식하고 그에 결합하도록 설계될 수 있다. CRISPR/Cas9-기반 시스템은 또한 조기 정지 코돈의 스킵을 유도하거나 또는 파괴된 리딩 프레임을 복원하기 위해 스플라이스 수용자 및 공여자를 표적화함으로써 정상 유전자 스플라이싱을 파괴하는데 사용될 수 있다. CRISPR/Cas9-기반 시스템은 게놈의 단백질-코딩 영역에 오프-타겟 변화를 매개할 수 있거나 또는 매개하지 않을 수 있다.

#### [0110] a. Cas9

[0111] CRISPR/Cas9-기반 시스템은 Cas9 단백질 또는 Cas9 융합 단백질을 포함할 수 있다. Cas9 단백질은 핵산을 절단하는 엔도뉴클레아제이고, CRISPR 유전자좌에 의해 코딩되며, 유형 II CRISPR 시스템에 수반된다. Cas9 단백질은 임의의 박테리아 또는 고세균 종, 예컨대 스트렙토코쿠스 피오게네스로부터의 것일 수 있다. Cas9 단백질은 돌연변이되어 뉴클레아제 활성을 불활성화시킬 수 있다. 어떤 엔도뉴클레아제 활성도 갖지 않는 스트렙토코쿠스 피오게네스로부터의 불활성화 Cas9 단백질 (iCas9, 또한 "dCas9"로 지칭됨)은 최근에 박테리아, 효모 및 인간 세포에서 gRNA에 의해 유전자에 표적화되어 입체 장애를 통해 유전자 발현을 침묵화하였다. 본원에 사용된 "iCas9" 및 "dCas9" 둘 다는 아미노산 치환 D10A 및 H840A를 갖는 Cas9 단백질을 지칭하며, 그의 뉴클레아제 활성이 불활성화되어 있다. 예를 들어, CRISPR/Cas9-기반 시스템은 서열 459 또는 461의 Cas9를 포함할 수 있다.

#### [0112] b. Cas9 융합 단백질

[0113] CRISPR/Cas9-기반 시스템은 융합 단백질을 포함할 수 있다. 융합 단백질은 2개의 이중 폴리펩티드 도메인을 포함할 수 있으며, 여기서 제1 폴리펩티드 도메인은 Cas 단백질을 포함하고, 제2 폴리펩티드 도메인은 활성, 예컨

대 전사 활성화 활성, 전사 억제 활성, 전사 방출 인자 활성, 히스톤 변형 활성, 뉴클레아제 활성, 핵산 회합 활성, 메틸라제 활성 또는 데메틸라제 활성을 갖는다. 융합 단백질은 활성, 예컨대 전사 활성화 활성, 전사 억제 활성, 전사 방출 인자 활성, 히스톤 변형 활성, 뉴클레아제 활성, 핵산 회합 활성, 메틸라제 활성 또는 데메틸라제 활성을 갖는 제2 폴리펩티드 도메인에 융합된 상기 기재된 바와 같은 Cas9 단백질 또는 돌연변이된 Cas9 단백질을 포함할 수 있다.

[0114] (1) 전사 활성화 활성

[0115] 제2 폴리펩티드 도메인은 전사 활성화 활성, 즉 전사활성화 도메인을 가질 수 있다. 예를 들어, 내인성 포유동물 유전자, 예컨대 인간 유전자의 유전자 발현은 gRNA의 조합을 통해 포유동물 프로모터에 iCas9의 융합 단백질 및 전사활성화 도메인을 표적화함으로써 달성될 수 있다. 전사활성화 도메인은 VP16 단백질, 다중 VP16 단백질, 예컨대 VP48 도메인 또는 VP64 도메인, 또는 NF 카파 B 전사 활성화제 활성의 p65 도메인을 포함할 수 있다. 예를 들어, 융합 단백질은 iCas9-VP64일 수 있다.

[0116] (2) 전사 억제 활성

[0117] 제2 폴리펩티드 도메인은 전사 억제 활성을 가질 수 있다. 제2 폴리펩티드 도메인은 크루펠(Kruppel) 연관 박스 활성, 예컨대 KRAB 도메인, ERF 리프레서 도메인 활성, Mxi1 리프레서 도메인 활성, SID4X 리프레서 도메인 활성, Mad-SID 리프레서 도메인 활성 또는 TATA 박스 결합 단백질 활성을 가질 수 있다. 예를 들어, 융합 단백질은 dCas9-KRAB일 수 있다.

[0118] (3) 전사 방출 인자 활성

[0119] 제2 폴리펩티드 도메인은 전사 방출 인자 활성을 가질 수 있다. 제2 폴리펩티드 도메인은 진행 방출 인자 1 (ERF1) 활성 또는 진행 방출 인자 3 (ERF3) 활성을 가질 수 있다.

[0120] (4) 히스톤 변형 활성

[0121] 제2 폴리펩티드 도메인은 히스톤 변형 활성을 가질 수 있다. 제2 폴리펩티드 도메인은 히스톤 데아세틸라제, 히스톤 아세틸트랜스퍼라제, 히스톤 메틸라제 또는 히스톤 메틸트랜스퍼라제 활성을 가질 수 있다. 히스톤 아세틸트랜스퍼라제는 p300 또는 CREB-결합 단백질 (CBP) 단백질 또는 그의 단편일 수 있다. 예를 들어, 융합 단백질은 dCas9-p300일 수 있다.

[0122] (5) 뉴클레아제 활성

[0123] 제2 폴리펩티드 도메인은 Cas9 단백질의 뉴클레아제 활성과 상이한 뉴클레아제 활성을 가질 수 있다. 뉴클레아제 또는 뉴클레아제 활성을 갖는 단백질은 핵산의 뉴클레오타이드 서브유닛 사이의 포스포디에스테르 결합을 절단할 수 있는 효소이다. 뉴클레아제는 통상 엔도뉴클레아제 및 엑소뉴클레아제로 추가로 분류되며, 효소 중 일부는 두 카테고리에 속할 수 있다. 잘 알려져 있는 뉴클레아제는 데옥시리보뉴클레아제 및 리보뉴클레아제이다.

[0124] (6) 핵산 회합 활성

[0125] 제2 폴리펩티드 도메인은 핵산 회합 활성 또는 핵산 결합 단백질을 가질 수 있으며 - DNA-결합 도메인 (DBD)은 이중 또는 단일 가닥 DNA를 인식하는 적어도 1개의 모티프를 함유하는 독립적으로 폴딩된 단백질 도메인이다. DBD는 특정한 DNA 서열 (인식 서열)을 인식할 수 있거나 또는 DNA에 대해 일반적 친화도를 갖는다. 헬릭스-턴-헬릭스 영역, 류신 지퍼 영역, 윈드 헬릭스 영역, 윈드 헬릭스-턴-헬릭스 영역, 헬릭스-루프-헬릭스 영역, 이류노글로불린 폴드, B3 도메인, 아연 핑거, HMG-박스, Wor3 도메인, TAL 이펙터 DNA-결합 도메인으로 이루어진 군으로부터 선택된 핵산 회합 영역.

[0126] (7) 메틸라제 활성

[0127] 제2 폴리펩티드 도메인은 메틸 기를 DNA, RNA, 단백질, 소분자, 시토신 또는 아데닌으로 전달하는 것을 수반하는 메틸라제 활성을 가질 수 있다. 제2 폴리펩티드 도메인은 DNA 메틸트랜스퍼라제를 포함할 수 있다.

[0128] (8) 데메틸라제 활성

[0129] 제2 폴리펩티드 도메인은 데메틸라제 활성을 가질 수 있다. 제2 폴리펩티드 도메인은 메틸 (CH<sub>3</sub>-) 기를 핵산, 단백질 (특히 히스톤) 및 다른 분자로부터 제거하는 효소를 포함할 수 있다. 대안적으로, 제2 폴리펩티드는 DNA를 탈메틸화하기 위한 메커니즘에서 메틸 기를 히드록시메틸시토신으로 전환시킬 수 있다. 제2 폴리펩티드는 이러한 반응을 촉매할 수 있다. 예를 들어, 이러한 반응을 촉매하는 제2 폴리펩티드는 Tet1일 수 있다.

[0130] c. gRNA

[0131] gRNA는 CRISPR/Cas9-기반 시스템의 표적화를 제공한다. gRNA는 2개의 비코딩 RNA: crRNA 및 tracrRNA의 융합체이다. sgRNA는 목적하는 DNA 표적과의 상보적 염기 쌍형성을 통해 표적화 특이성을 부여하는 20 bp 프로토스페이서를 코딩하는 서열을 교환함으로써 임의의 목적하는 DNA 서열을 표적화할 수 있다. gRNA는 유형 II 이펙터 시스템에 수반된 자연 발생 crRNA:tracrRNA 듀플렉스를 모방한다. 예를 들어 42개-뉴클레오타이드 crRNA 및 75개-뉴클레오타이드 tracrRNA를 포함할 수 있는 이러한 듀플렉스는 표적 핵산을 절단하는 Cas9에 대한 가이드로서의 역할을 한다. 본원에서 상호교환적으로 사용된 "표적 영역", "표적 서열" 또는 "프로토스페이서"는 CRISPR/Cas9-기반 시스템이 표적화하는 표적 유전자의 영역을 지칭한다. CRISPR/Cas9-기반 시스템은 적어도 1개의 gRNA를 포함할 수 있으며, 여기서 gRNA는 상이한 DNA 서열을 표적화한다. 표적 DNA 서열은 중복될 수 있다. 표적 서열 또는 프로토스페이서에 이어서, 프로토스페이서의 3' 말단에 PAM 서열이 존재한다. 상이한 유형 II 시스템은 상이한 PAM 요건을 갖는다. 예를 들어, 스트렙토코쿠스 피오게네스 유형 II 시스템은 "NGG" 서열을 사용하며, 여기서 "N"은 임의의 뉴클레오타이드일 수 있다.

[0132] 세포에 투여되는 gRNA의 개수는 적어도 1개의 gRNA, 적어도 2개의 상이한 gRNA, 적어도 3개의 상이한 gRNA 적어도 4개의 상이한 gRNA, 적어도 5개의 상이한 gRNA, 적어도 6개의 상이한 gRNA, 적어도 7개의 상이한 gRNA, 적어도 8개의 상이한 gRNA, 적어도 9개의 상이한 gRNA, 적어도 10개의 상이한 gRNA, 적어도 11개의 상이한 gRNA, 적어도 12개의 상이한 gRNA, 적어도 13개의 상이한 gRNA, 적어도 14개의 상이한 gRNA, 적어도 15개의 상이한 gRNA, 적어도 16개의 상이한 gRNA, 적어도 17개의 상이한 gRNA, 적어도 18개의 상이한 gRNA, 적어도 18개의 상이한 gRNA, 적어도 20개의 상이한 gRNA, 적어도 25개의 상이한 gRNA, 적어도 30개의 상이한 gRNA, 적어도 35개의 상이한 gRNA, 적어도 40개의 상이한 gRNA, 적어도 45개의 상이한 gRNA, 또는 적어도 50개의 상이한 gRNA일 수 있다. 세포에 투여되는 gRNA의 개수는 적어도 1개의 gRNA 내지 적어도 50개의 상이한 gRNA, 적어도 1개의 gRNA 내지 적어도 45개의 상이한 gRNA, 적어도 1개의 gRNA 내지 적어도 40개의 상이한 gRNA, 적어도 1개의 gRNA 내지 적어도 35개의 상이한 gRNA, 적어도 1개의 gRNA 내지 적어도 30개의 상이한 gRNA, 적어도 1개의 gRNA 내지 적어도 25개의 상이한 gRNA, 적어도 1개의 gRNA 내지 적어도 20개의 상이한 gRNA, 적어도 1개의 gRNA 내지 적어도 16개의 상이한 gRNA, 적어도 1개의 gRNA 내지 적어도 12개의 상이한 gRNA, 적어도 1개의 gRNA 내지 적어도 8개의 상이한 gRNA, 적어도 1개의 gRNA 내지 적어도 4개의 상이한 gRNA, 적어도 4개의 gRNA 내지 적어도 50개의 상이한 gRNA, 적어도 4개의 상이한 gRNA 내지 적어도 40개의 상이한 gRNA, 적어도 4개의 상이한 gRNA 내지 적어도 35개의 상이한 gRNA, 적어도 4개의 상이한 gRNA 내지 적어도 30개의 상이한 gRNA, 적어도 4개의 상이한 gRNA 내지 적어도 25개의 상이한 gRNA, 적어도 4개의 상이한 gRNA 내지 적어도 20개의 상이한 gRNA, 적어도 4개의 상이한 gRNA 내지 적어도 16개의 상이한 gRNA, 적어도 4개의 상이한 gRNA 내지 적어도 12개의 상이한 gRNA, 적어도 4개의 상이한 gRNA 내지 적어도 8개의 상이한 gRNA, 적어도 8개의 상이한 gRNA 내지 적어도 50개의 상이한 gRNA, 적어도 8개의 상이한 gRNA 내지 적어도 45개의 상이한 gRNA, 적어도 8개의 상이한 gRNA 내지 적어도 40개의 상이한 gRNA, 적어도 8개의 상이한 gRNA 내지 적어도 35개의 상이한 gRNA, 8개의 상이한 gRNA 내지 적어도 30개의 상이한 gRNA, 적어도 8개의 상이한 gRNA 내지 적어도 25개의 상이한 gRNA, 8개의 상이한 gRNA 내지 적어도 20개의 상이한 gRNA, 적어도 8개의 상이한 gRNA 내지 적어도 16개의 상이한 gRNA, 또는 8개의 상이한 gRNA 내지 적어도 12개의 상이한 gRNA일 수 있다.

[0133] gRNA는 표적 DNA 서열의 상보적 폴리뉴클레오타이드 서열 및 그에 이어서 PAM 서열을 포함할 수 있다. gRNA는 상보적 폴리뉴클레오타이드 서열의 5' 말단에 "G"를 포함할 수 있다. gRNA는 표적 DNA 서열의 적어도 10개 염기 쌍, 적어도 11개 염기 쌍, 적어도 12개 염기 쌍, 적어도 13개 염기 쌍, 적어도 14개 염기 쌍, 적어도 15개 염기 쌍, 적어도 16개 염기 쌍, 적어도 17개 염기 쌍, 적어도 18개 염기 쌍, 적어도 19개 염기 쌍, 적어도 20개 염기 쌍, 적어도 21개 염기 쌍, 적어도 22개 염기 쌍, 적어도 23개 염기 쌍, 적어도 24개 염기 쌍, 적어도 25개 염기 쌍, 적어도 30개 염기 쌍 또는 적어도 35개 염기 쌍의 상보적 폴리뉴클레오타이드 서열 및 그에 이어서 PAM 서열을 포함할 수 있다. PAM 서열은 "NGG"일 수 있으며, 여기서 "N"은 임의의 뉴클레오타이드일 수 있다. gRNA는 표적 유전자의 프로모터 영역, 인핸서 영역 또는 전사된 영역 중 적어도 하나를 표적화할 수 있다. gRNA는 서열 5-40, 65-144, 492-515, 540-563, 585-625, 462 (도 40), 464 (도 41) 및 465 (도 41) 중 적어도 하나의 핵산 서열을 포함할 수 있다.

[0134] gRNA는 임의의 핵산 서열을 표적화할 수 있다. 핵산 서열 표적은 DNA일 수 있다. DNA는 임의의 유전자일 수 있다. 예를 들어, gRNA는 유전자, 예컨대 BRN2, MYT1L, ASCL1, NANOG, VEGFA, TERT, IL1B, IL1R2, IL1RN, HBG1, HBG2, MYOD1, OCT4 및 DMD를 표적화할 수 있다.



[0135] (1) 디스트로핀

[0136] "디스트로핀"은 막대-형상 세포질 단백질이며, 이는 근섬유의 세포골격을 세포막을 통해 주위 세포의 매트릭스에 연결하는 단백질 복합체의 일부이다. 디스트로핀은 세포 막의 디스트로글리칸 복합체에 구조적 안정성을 제공한다. 디스트로핀 유전자는 유전자좌 Xp21에서 2.2 메가염기이다. 1차 전사는 약 2,400 kb로 측정되며, 이때 성숙 mRNA는 약 14 kb이다. 79개 엑손은 3500개 아미노산을 초과하는 단백질을 코딩한다. 정상 골격 근육 조직은 단지 소량의 디스트로핀을 함유하지만 그의 비정상적 발현의 부재는 중증의 치유불가능한 증상의 발달을 유도한다. 디스트로핀 유전자에서의 일부 돌연변이는 이환된 환자에서 결합성 디스트로핀 및 중증 이영양 표현형의 생성을 유도한다. 디스트로핀 유전자에서의 일부 돌연변이는 이환된 환자에서 부분-기능적 디스트로핀 단백질 및 훨씬 더 경도인 이영양 표현형을 유도한다.

[0137] DMD는 디스트로핀 유전자에서 넌센스 또는 프레임 시프트 돌연변이를 유발하는 유전성 또는 자발적 돌연변이의 결과이다. 자연 발생 돌연변이 및 그의 결과는 DMD에 대해 비교적 잘 이해된다. 막대형 도메인 내에 함유된 엑손 45-55 영역에서 발생하는 인-프레임 결실은 고도 기능적 디스트로핀 단백질을 생산할 수 있고, 다수의 보균자는 무증상이거나 또는 경도 증상을 나타낼 수 있는 것으로 공지되어 있다. 또한, 60% 초과환자는 이론적으로 이러한 디스트로핀 유전자 영역에서 엑손을 표적화함으로써 치료될 수 있다. DMD 환자에서 mRNA 스플라이싱 동안 비-필수 엑손을 스킵하는 것에 의해 파괴된 디스트로핀 리딩 프레임을 복원하여 내부적으로는 결실되었지만 기능적인 디스트로핀 단백질을 생산하려는 노력이 이루어졌다. 내부 디스트로핀 엑손의 결실은 적절한 리딩 프레임을 유지하지만, 보다 덜 중증인 베커 근육 이영양증을 유발한다.

[0138] (2) 디스트로핀을 표적화하기 위한 CRISPR/Cas9-기반 시스템

[0139] 디스트로핀 유전자에 특이적인 CRISPR/Cas9-기반 시스템이 본원에 개시된다. CRISPR/Cas9-기반 시스템은 디스트로핀 유전자를 표적화하기 위해 Cas9 및 적어도 1개의 gRNA를 포함할 수 있다. CRISPR/Cas9-기반 시스템은 표적 영역에 결합하고 그를 인식할 수 있다. 표적 영역은 가능한 아웃-오브-프레임 정지 코돈의 바로 상류에서 선택되어, 복구 과정 동안의 삽입 또는 결실이 프레임 전환에 의해 디스트로핀 리딩 프레임을 복원하도록 할 수 있다. 표적 영역은 또한 스플라이스 수용자 부위 또는 스플라이스 공여자 부위여서, 복구 과정 동안의 삽입 또는 결실이 스플라이싱을 파괴하고 스플라이스 부위 파괴 및 엑손 제거에 의해 디스트로핀 리딩 프레임을 복원하도록 할 수 있다. 표적 영역은 또한 이상 정지 코돈이어서, 복구 과정 동안의 삽입 또는 결실이 정지 코돈을 제거하거나 또는 파괴함으로써 디스트로핀 리딩 프레임을 복원하도록 할 수 있다.

[0140] 단일 또는 멀티플렉스화 sgRNA는 엑손 45-55에서의 돌연변이 핫스팟을 표적화하고 엑손내 작은 삽입 및 결실 또는 하나 이상의 엑손의 큰 결실을 도입함으로써 디스트로핀 리딩 프레임을 복원하도록 설계될 수 있다. Cas9 및 1개 이상의 sgRNA를 사용한 처리 후에, 디스트로핀 발현은 시험관내 뒤시엔스 환자 근육 세포에서 복원될 수 있다. 인간 디스트로핀은 유전자 교정된 환자 세포의 면역결핍 마우스 내로의 이식 후에 생체내에서 검출되었다. 유의하게, CRISPR/Cas9 시스템의 고유한 멀티플렉스 유전자 편집 능력은 보편적 또는 환자-특이적 유전자 편집 접근법에 의해 환자 돌연변이의 최대 62%를 교정할 수 있는 이러한 돌연변이 핫스팟 영역의 큰 결실을 효율적으로 생성할 수 있게 한다.

[0141] CRISPR/Cas9-기반 시스템은 다양한 서열 및 길이의 gRNA를 사용할 수 있다. gRNA의 예는 표 6에서 찾아볼 수 있다. CRISPR/Cas9-기반 시스템은 서열 65-144의 핵산 서열 또는 그의 보체를 표적화할 수 있다. gRNA는 서열 65-144로 이루어진 군으로부터 선택된 뉴클레오타이드 서열 또는 그의 보체를 포함할 수 있다. 예를 들어, 개시된 CRISPR/Cas9-기반 시스템은 디스트로핀 유전자의 엑손 51에 고도로 효율적인 유전자 편집을 매개하도록 조작되었다. 이들 CRISPR/Cas9-기반 시스템은 DMD 환자로부터의 세포에서 디스트로핀 단백질 발현을 복원하였다.

[0142] (a) 엑손 51 및 45-55

[0143] 엑손 51은 DMD에서 프레임-파괴 결실에 빈번하게 인접해있다. 엑손 스킵에 의한 디스트로핀 전사체로부터의 엑손 51의 제거는 모든 DMD 환자의 대략 15%를 치료하는데 사용될 수 있다. DMD 돌연변이의 이러한 부류는 NHEJ-기반 게놈 편집 및 HDR에 의한 영구적 교정에 이상적으로 적합하다. 본원에 기재된 CRISPR/Cas9-기반 시스템은 인간 디스트로핀 유전자에서 엑손 51의 표적화된 변형을 위해 개발되었다. 이들 CRISPR/Cas9-기반 시스템은 인간 DMD 세포 내로 형질감염되고, 효율적인 유전자 변형 및 정확한 리딩 프레임으로의 전환을 매개하였다. 단백질 복원은 프레임 복원에 동반되며 CRISPR/Cas9-기반 시스템-처리 세포의 벌크 집단에서 검출되었다. 유사하게, 디스트로핀 전사체의 엑손 45-55의 제거는 모든 DMD 환자의 대략 62%를 치료하는데 사용될 수 있다.

- [0144] (3) AAV/CRISPR 구축물
- [0145] AAV는 다양한 구축물 구성을 사용하여 CRISPR을 전달하는데 사용될 수 있다 (도 39 참조). 예를 들어, AAV는 별개의 벡터 상에서 Cas9 및 gRNA 발현 카세트를 전달할 수 있다. 대안적으로, 스타필로코쿠스 아우레우스 또는 네이세리아 메니기티디스와 같은 종으로부터 유래된 작은 Cas9 단백질이 사용된다면, Cas9 및 2개 이하의 gRNA 발현 카세트 둘 다는 4.7 kb 패키징 제한 내에서 단일 AAV 벡터에 조합될 수 있다 (도 39 참조).
- [0146] 5. 멀티플렉스 CRISPR/Cas9-기반 시스템
- [0147] 본 개시내용은 하나 이상의 내인성 유전자를 표적화하기 위한 CRISPR/CRISPR-연관 (Cas) 9-기반 시스템, 예컨대 Cas9 또는 dCas9 및 다중 gRNA를 포함하는 멀티플렉스 CRISPR/Cas9-기반 시스템에 관한 것이다. 이러한 플랫폼은 편리한 골든 게이트 클로닝 방법을 이용하여 최대 4개의 독립적 sgRNA 발현 카세트를 단일 렌티바이러스 벡터 내로 신속하게 통합시킨다. 각 sgRNA는 효율적으로 발현되었고, 불멸화 1차 인간 세포주에서의 다양한 유전자좌에 멀티플렉스 유전자 편집을 매개할 수 있다. dCas9-VP64를 안정하게 발현하는 세포주에서의 일시적 전사 활성화는 1 내지 4개의 sgRNA를 사용한 상승작용적 활성화에 의해 조정가능한 것으로 입증되었다. 또한, 단일 렌티바이러스 벡터는 불멸화 1차 인간 세포에서 지속적이고 장기적인 내인성 유전자 발현을 유도할 수 있다. 이러한 시스템은 모델 및 1차 세포주에서 효율적 멀티플렉스 유전자 편집 또는 활성화를 가능하게 하는 단일 렌티바이러스 벡터의 신속한 조립을 가능하게 한다.
- [0148] 멀티플렉스 CRISPR/Cas9-기반 시스템은 전사 활성화의 효력 및 전사 활성화의 조정가능한 유도를 제공한다. 골든 게이트 조립에 의해 용이하게 생성되는 최종 벡터는 독립적 프로모터로부터 발현된 1, 2, 3 또는 4개의 sgRNA 이외에도 구성적 Cas9 또는 dCas9-VP64를 발현한다. 각 프로모터는 Cas9 뉴클레아제 활성의 유사한 수준을 지시하는 sgRNA를 효율적으로 발현할 수 있다. 또한, Cas9 및 독립적 유전자좌를 표적화하는 4개의 sgRNA를 발현하는 단일 벡터의 렌티바이러스 전달은 모든 4개의 유전자좌의 동시 멀티플렉스 유전자 편집을 생성하였다. 일시적 및 안정적 둘 다의 맥락에서 2개의 내인성 유전자에서의 조정가능한 전사 활성화는 sgRNA와 함께 또는 그 없이 Cas9의 렌티바이러스 전달을 사용하여 달성되었다. 1차 인간 세포에서 고도로 효율적이고 장기적인 유전자 활성화가 달성된다. 따라서, 이러한 시스템은 인간 세포에서 멀티플렉스 유전자 편집 및 장기적 전사 활성화를 생성하기 위한 매력적이고 효율적인 방법이다.
- [0149] 멀티플렉스 CRISPR/Cas9-기반 시스템은 다중 유전자를 동시에 불활성화시키기 위한 효율적 멀티플렉스 유전자 편집을 가능하게 한다. CRISPR/Cas9 시스템은 단일 Cas9 단백질을 2개 이상의 sgRNA와 공발현시킴으로써 다중의 별개의 게놈 유전자좌를 동시에 표적화하여, 이러한 시스템이 멀티플렉스 유전자 편집 또는 상승작용적 활성화 적용에 고유하게 적합하도록 만들 수 있다. CRISPR/Cas9 시스템은 발현된 sgRNA 분자를 간단하게 변형시킴으로써 새로운 부위로의 분자 표적화의 과정을 매우 촉진한다. 단일 렌티바이러스 벡터는 화학 또는 광유전학 조절에 의해 이들 성분의 유도성 제어를 달성하는 방법과 조합되어, 시간상 및 공간상 둘 다에서의 유전자 조절의 역학의 조사를 촉진할 수 있다.
- [0150] 멀티플렉스 CRISPR/Cas9-기반 시스템은 2개 이상의 내인성 유전자를 전사 활성화시킬 수 있다. 멀티플렉스 CRISPR/Cas9-기반 시스템은 2개 이상의 내인성 유전자를 전사 억제할 수 있다. 예를 들어, 적어도 2개의 내인성 유전자, 적어도 3개의 내인성 유전자, 적어도 4개의 내인성 유전자, 적어도 5개의 내인성 유전자 또는 적어도 10개의 내인성 유전자는 멀티플렉스 CRISPR/Cas9-기반 시스템에 의해 활성화되거나 또는 억제될 수 있다. 2 내지 15개의 유전자, 2 내지 10개의 유전자, 2 내지 5개의 유전자, 5 내지 15개의 유전자, 또는 5 내지 10개의 유전자가 멀티플렉스 CRISPR/Cas9-기반 시스템에 의해 활성화되거나 또는 억제될 수 있다.
- [0151] (1) 변형된 렌티바이러스 벡터
- [0152] 멀티플렉스 CRISPR/Cas9-기반 시스템은 변형된 렌티바이러스 벡터를 포함한다. 변형된 렌티바이러스 벡터는 융합 단백질을 코딩하는 제1 폴리뉴클레오티드 서열 및 적어도 1개의 sgRNA를 코딩하는 제2 폴리뉴클레오티드 서열을 포함한다. 융합 단백질은 상기 기재된 바와 같은 CRISPR/Cas9-기반 시스템의 융합 단백질일 수 있다. 제1 폴리뉴클레오티드 서열은 프로모터에 작동가능하게 연결될 수 있다. 프로모터는 구성적 프로모터, 유도성 프로모터, 억제성 프로모터 또는 조절성 프로모터일 수 있다.
- [0153] 제2 폴리뉴클레오티드 서열은 적어도 1개의 sgRNA를 코딩한다. 예를 들어, 제2 폴리뉴클레오티드 서열은 적어도 1개의 sgRNA, 적어도 2개의 sgRNA, 적어도 3개의 sgRNA, 적어도 4개의 sgRNA, 적어도 5개의 sgRNA, 적어도 6개의 sgRNA, 적어도 7개의 sgRNA, 적어도 8개의 sgRNA, 적어도 9개의 sgRNA, 적어도 10개의 sgRNA, 적어도 11개의 sgRNA, 적어도 12개의 sgRNA, 적어도 13개의 sgRNA, 적어도 14개의 sgRNA, 적어도 15개의 sgRNA, 적어도

16개의 sgRNA, 적어도 17개의 sgRNA, 적어도 18개의 sgRNA, 적어도 19개의 sgRNA, 적어도 20개의 sgRNA, 적어도 25개의 sgRNA, 적어도 30개의 sgRNA, 적어도 35개의 sgRNA, 적어도 40개의 sgRNA, 적어도 45개의 sgRNA, 또는 적어도 50개의 sgRNA를 코딩할 수 있다. 제2 폴리뉴클레오티드 서열은 1개의 sgRNA 내지 50개의 sgRNA, 1개의 sgRNA 내지 45개의 sgRNA, 1개의 sgRNA 내지 40개의 sgRNA, 1개의 sgRNA 내지 35개의 sgRNA, 1개의 sgRNA 내지 30개의 sgRNA, 1개의 sgRNA 내지 25개의 상이한 sgRNA, 1개의 sgRNA 내지 20개의 sgRNA, 1개의 sgRNA 내지 16개의 sgRNA, 1개의 sgRNA 내지 8개의 상이한 sgRNA, 4개의 상이한 sgRNA 내지 50개의 상이한 sgRNA, 4개의 상이한 sgRNA 내지 45개의 상이한 sgRNA, 4개의 상이한 sgRNA 내지 40개의 상이한 sgRNA, 4개의 상이한 sgRNA 내지 35개의 상이한 sgRNA, 4개의 상이한 sgRNA 내지 30개의 상이한 sgRNA, 4개의 상이한 sgRNA 내지 25개의 상이한 sgRNA, 4개의 상이한 sgRNA 내지 20개의 상이한 sgRNA, 4개의 상이한 sgRNA 내지 16개의 상이한 sgRNA, 4개의 상이한 sgRNA 내지 8개의 상이한 sgRNA, 8개의 상이한 sgRNA 내지 50개의 상이한 sgRNA, 8개의 상이한 sgRNA 내지 45개의 상이한 sgRNA, 8개의 상이한 sgRNA 내지 40개의 상이한 sgRNA, 8개의 상이한 sgRNA 내지 35개의 상이한 sgRNA, 8개의 상이한 sgRNA 내지 30개의 상이한 sgRNA, 8개의 상이한 sgRNA 내지 25개의 상이한 sgRNA, 8개의 상이한 sgRNA 내지 20개의 상이한 sgRNA, 8개의 상이한 sgRNA 내지 16개의 상이한 sgRNA, 16개의 상이한 sgRNA 내지 50개의 상이한 sgRNA, 16개의 상이한 sgRNA 내지 45개의 상이한 sgRNA, 16개의 상이한 sgRNA 내지 40개의 상이한 sgRNA, 16개의 상이한 sgRNA 내지 35개의 상이한 sgRNA, 16개의 상이한 sgRNA 내지 30개의 상이한 sgRNA, 16개의 상이한 sgRNA 내지 25개의 상이한 sgRNA, 또는 16개의 상이한 sgRNA 내지 20개의 상이한 sgRNA를 코딩할 수 있다. 상이한 sgRNA를 코딩하는 각각의 폴리뉴클레오티드 서열은 프로모터에 작동가능하게 연결될 수 있다. 상이한 sgRNA에 작동가능하게 연결된 프로모터는 동일한 프로모터일 수 있다. 상이한 sgRNA에 작동가능하게 연결된 프로모터는 상이한 프로모터일 수 있다. 프로모터는 구성적 프로모터, 유도성 프로모터, 억제성 프로모터 또는 조절성 프로모터일 수 있다.

[0154] 적어도 1개의 sgRNA는 표적 유전자 또는 유전자좌에 결합할 수 있다. 1개 초과와 sgRNA가 포함되는 경우에, 각각의 sgRNA는 1개의 표적 유전자좌 내의 상이한 표적 영역에 결합하거나 또는 각각의 sgRNA는 상이한 유전자좌 내의 상이한 표적 영역에 결합한다. 융합 단백질은 Cas9 단백질 또는 iCas9-VP64 단백질을 포함할 수 있다. 융합 단백질은 VP64 도메인, p300 도메인 또는 KRAB 도메인을 포함할 수 있다.

[0155] 6. 부위-특이적 뉴클레아제

[0156] 상기 기재된 바와 같은 조성물은 표적 영역에 결합하여 그를 절단하는 부위-특이적 뉴클레아제를 코딩하는 뉴클레오티드 서열을 포함한다. 부위-특이적 뉴클레아제는 조작될 수 있다. 예를 들어, 조작된 부위-특이적 뉴클레아제는 CRISPR/Cas9-기반 시스템, ZFN 또는 TALEN일 수 있다. 부위-특이적 뉴클레아제는 골격근 또는 심장 근육에서의 세포의 게놈의 유전자 또는 유전자좌에 결합하여 그를 절단할 수 있다. 예를 들어, 유전자 또는 유전자좌는 Rosa26 유전자좌 또는 디스트로핀 유전자일 수 있다.

[0157] a. CRISPR/Cas9-기반 시스템

[0158] 상기 기재된 바와 같은 CRISPR/Cas9-기반 시스템은 표적화된 게놈 유전자좌에 부위-특이적 이중 가닥 파단을 도입하는데 사용될 수 있다.

[0159] b. 아연 핑거 뉴클레아제 (ZFN)

[0160] 부위-특이적 뉴클레아제는 ZFN일 수 있다. 단일 아연 핑거는 염기 쌍당 단일 아미노산 측쇄의 상호작용을 통해 DNA의 3개 연속적 염기 쌍에 결합함으로써 대략 30개 아미노산 및 도메인 기능을 함유한다. 아연 핑거 모티프의 모듈식 구조는 연속적으로 여러 도메인의 결합을 허용하여, 3개 뉴클레오티드의 배수로 연장된 서열의 인식 및 표적화를 가능하게 한다. 이들 표적화된 DNA-결합 도메인은 뉴클레아제 도메인, 예컨대 FokI와 조합되어, 부위-특이적 뉴클레아제를 생성할 수 있으며, 이는 표적화된 게놈 유전자좌에 부위-특이적 이중 가닥 파단을 도입하는데 사용될 수 있는 "아연 핑거 뉴클레아제" (ZFN)로 칭해진다. 이러한 DNA 절단은 천연 DNA-복구 기구를 자극하여, 2개의 가능한 복구 경로인 NHEJ 및 HDR 중 하나를 유도할 수 있다. 예를 들어, ZFN은 표적 Rosa26 유전자좌 (Perez-Pinera et al. Nucleic Acids Research (2012) 40:3741-3752) 또는 디스트로핀 유전자를 표적화할 수 있다. ZFN의 예는 표 1 및 도 35-38에 제시되어 있다. 표 1에서, DNA 인식 나선은 밑줄표시되어 있고, "Fok ELD-S" 및 "Fok KKR-S"는 아연 핑거 단백질 DNA-결합 도메인에 융합된 FokI 뉴클레아제 도메인을 지칭한다. 도 35-38에서, 표적 부위에서 (즉, 서열 442, 445, 448 및 453에서) 표적 DNA 서열 및 ZFN 아미노산 서열에서 (즉, 서열 443, 444, 446, 447, 449-452, 454 및 455에서) DNA 인식 나선은 각각 밑줄표시되어 있다.

[0161] <표 1>

[0162] 확인된 ZFN의 전체 아미노산 서열

<p><b>ZFN B 좌측 Fok ELD-S 전체 아미노산 서열 (서열 438)</b></p> <p>MDYKDHDGDYKDHDIDYKDDDDKMAPKKRKVGRLEPGEKPYKCPECGKSFSR<u>KDALRGH</u>          QRTHTGEKPYKCPECGKSFS<u>SHRTTLTNH</u>QRTHTGEKPYKCPECGKSFS<u>QRNALAGH</u>QRTHT          TGEKPYKCPECGKSFS<u>SHKNALQN</u>HQRTHTGEKPYKCPECGKSFS<u>DPGHLVRH</u>QRTHTGEK          PYKCPECGKSFS<u>TSGNLVRH</u>QRTHTGAAARALVKSELEEKSELRHKLKYPHEYIELIE          IARNPTQDRILEMKVMEFFMKVYGYRGEHLGGSRKPDGAIYTVGSPIDYGVIVDTKAYSG          GYNLPIGQADEMERYVEENQTRDKHLNPNEWWKVYPSSVTEFKFLFVSGHFGKGYKAQLT          RLNHITNCNGAVLSVEELLIGGEMIKAGTTLTLEEVRRKFNNGEINF*</p>
<p><b>ZFN B 우측 Fok KKR-S 전체 아미노산 서열 (서열 439)</b></p> <p>MDYKDHDGDYKDHDIDYKDDDDKMAPKKRKVGRLEPGEKPYKCPECGKSFS<u>QQRSLVGH</u>          QRTHTGEKPYKCPECGKSFS<u>DKKDLTRH</u>QRTHTGEKPYKCPECGKSFS<u>TSGHLVRH</u>QRTHT          TGEKPYKCPECGKSFS<u>QRAHLERH</u>QRTHTGEKPYKCPECGKSFS<u>TSGSLVRH</u>QRTHTGEK          PYKCPECGKSFS<u>TSGNLVRH</u>QRTHTGAAARALVKSELEEKSELRHKLKYPHEYIELIE          IARNPTQDRILEMKVMEFFMKVYGYRGEHLGGSRKPDGAIYTVGSPIDYGVIVDTKAYSG          GYNLPIGQADEMQRYVKENQTRNKHINPNEWWKVYPSSVTEFKFLFVSGHFGKGYKAQLT          RLNRKTNCNGAVLSVEELLIGGEMIKAGTTLTLEEVRRKFNNGEINF*</p>
<p><b>ZFN J 좌측 Fok KKR-S 전체 아미노산 서열 (서열 440)</b></p> <p>MDYKDHDGDYKDHDIDYKDDDDKMAPKKRKVGRLEPGERPFQCRICMRNFS<u>SKQALAVH</u>          TRTHTGEKPFQCRICMRNFS<u>QSTTLKRHLR</u>THRTHTGEKPFQCRICMRNFS<u>RSDDLHL</u>LHLKTH          LRGSQVLVKSELEEKSELRHKLKYPHEYIELIEIARNPTQDRILEMKVMEFFMKVYGYR          GEHLGGSRKPDGAIYTVGSPIDYGVIVDTKAYSGGYNLPIGQADEMQRYVKENQTRNKH          NPNEWWKVYPSSVTEFKFLFVSGHFGKGYKAQLTRLNRKTNCNGAVLSVEELLIGGEMIK          AGTTLTLEEVRRKFNNGEINF*</p>
<p><b>ZFN J 우측 Fok ELD-S 전체 아미노산 서열 (서열 441)</b></p> <p>MDYKDHDGDYKDHDIDYKDDDDKMAPKKRKVGRLEPGERPFQCRICMRNFS<u>RRRAHLQNH</u>          TRTHTGEKPFQCRICMRNFS<u>QSTTLKRHLR</u>THRTHTGEKPFQCRICMRNFS<u>DGGHLTRHL</u>LHLKTH          LRGSQVLVKSELEEKSELRHKLKYPHEYIELIEIARNPTQDRILEMKVMEFFMKVYGYR          GEHLGGSRKPDGAIYTVGSPIDYGVIVDTKAYSGGYNLPIGQADEMERYVEENQTRDKHL          NPNEWWKVYPSSVTEFKFLFVSGHFGKGYKAQLTRLNHITNCNGAVLSVEELLIGGEMIK          AGTTLTLEEVRRKFNNGEINF*</p>

[0163]

[0164] c. TAL 이펙터 뉴클레아제 (TALEN)

[0165] TALEN은 표적화된 게놈 유전자좌에 부위-특이적 이중 가닥 파단을 도입하는데 사용될 수 있다. 부위-특이적 이중-가닥 파단은 2개의 독립적 TALEN이 DNA 서열 근처에 결합하여 그로 인해 FokI의 이량체화 및 표적 DNA의 절단을 허용하는 경우에 생성된다. TALEN의 절단은 성공적이고 효율적인 유전자 변형의 높은 비율로 인해 게놈 편집을 진보시켰다. 이러한 DNA 절단은 천연 DNA-복구 기구를 자극하여, 2개의 가능한 복구 경로: 상동성-지정 복구 (HDR) 또는 비-상동 말단 연결 (NHEJ) 경로 중 하나를 유도할 수 있다. TALEN은 유전 질환에 수반되는 임의의 유전자를 표적화하도록 설계될 수 있다.

[0166] TALEN은 뉴클레아제 및 TALEN 표적 영역에서의 표적 유전자에 결합하는 TALE DNA-결합 도메인을 포함할 수 있다. 표적 유전자는 돌연변이, 예컨대 프레임시프트 돌연변이 또는 넌센스 돌연변이를 가질 수 있다. 표적 유전자가 조기 정지 코돈을 유발하는 돌연변이를 갖는 경우에, TALEN은 조기 정지 코돈으로부터의 상류 또는 하류 뉴클레오타이드 서열을 인식하고 그에 결합하도록 설계될 수 있다. "TALEN 표적 영역"은 2개의 TALEN에 대한 결합 영역 및 그 결합 영역 사이에 발생하는 스페이서 영역을 포함한다. 2개의 TALEN은 TALEN 표적 영역 내의 상이한 결합 영역에 결합하고, 그 후에 TALEN 표적 영역은 절단된다. TALEN의 예는 그의 전문이 참조로 포함된 국제 특허 출원 번호 PCT/US2013/038536에 기재되어 있다.

[0167] 7. 전사 활성화제



- [0168] 상기 기재된 바와 같은 조성물은 표적 유전자를 활성화시키는 전사 활성화제를 코딩하는 뉴클레오타이드 서열을 포함한다. 전사 활성화제는 조작될 수 있다. 예를 들어, 조작된 전사 활성화제는 CRISPR/Cas9-기반 시스템, 아연 핑거 융합 단백질 또는 TALE 융합 단백질일 수 있다.
- [0169] a. CRISPR/Cas9-기반 시스템
- [0170] 상기 기재된 바와 같은 CRISPR/Cas9-기반 시스템은 RNA로 표적 유전자의 전사를 활성화시키는데 사용될 수 있다. CRISPR/Cas9-기반 시스템은 상기 기재된 바와 같은 융합 단백질을 포함할 수 있으며, 여기서 제2 폴리펩티드 도메인은 전사 활성화 활성 또는 히스톤 변형 활성을 갖는다. 예를 들어, 제2 폴리펩티드 도메인은 VP64 또는 p300을 포함할 수 있다.
- [0171] b. 아연 핑거 융합 단백질
- [0172] 전사 활성화제는 아연 핑거 융합 단백질일 수 있다. 상기 기재된 바와 같은 아연 핑거 표적화된 DNA-결합 도메인은 전사 활성화 활성 또는 히스톤 변형 활성을 갖는 도메인과 조합될 수 있다. 예를 들어, 도메인은 VP64 또는 p300을 포함할 수 있다.
- [0173] c. TALE 융합 단백질
- [0174] TALE 융합 단백질은 표적 유전자의 전사를 활성화시키는데 사용될 수 있다. TALE 융합 단백질은 TALE DNA-결합 도메인 및 전사 활성화 활성 또는 히스톤 변형 활성을 갖는 도메인을 포함할 수 있다. 예를 들어, 도메인은 VP64 또는 p300을 포함할 수 있다.
- [0175] 8. 조성물
- [0176] 본 발명은 세포 또는 대상체에서 유전자 발현을 변경하고 게놈 DNA를 조작 또는 변경하기 위한 조성물에 관한 것이다. 조성물은 또한 바이러스 전달 시스템을 포함할 수 있다.
- [0177] a. 근육에서의 게놈 편집을 위한 조성물
- [0178] 본 발명은 대상체의 골격근 또는 심장 근육에서의 표적 유전자를 게놈 편집하기 위한 조성물에 관한 것이다. 조성물은 부위-특이적 뉴클레아제를 코딩하는 변형된 AAV 벡터 및 뉴클레오타이드 서열을 포함한다. 조성물은 부위-특이적 뉴클레아제의 활성 형태를 골격근 또는 심장 근육으로 전달한다. 조성물은 공여자 DNA 또는 트랜스진을 추가로 포함할 수 있다. 이들 조성물은 게놈 편집, 게놈 조작, 및 유전 질환 및/또는 다른 골격 또는 심장 근육 병태에 수반되는 유전자에서의 돌연변이의 효과를 교정하거나 또는 감소시키는데 사용될 수 있다.
- [0179] 표적 유전자는 유전자의 활성화, 억제 또는 파괴가 요구될 수 있는 세포의 분화 또는 임의의 다른 과정에 수반될 수 있거나, 또는 돌연변이, 예컨대 결실, 프레임시프트 돌연변이 또는 넌센스 돌연변이를 가질 수 있다. 표적 유전자가 조기 정지 코돈, 이상 스플라이스 수용자 부위 또는 이상 스플라이스 공여자 부위를 유발하는 돌연변이를 갖는 경우에, 부위-특이적 뉴클레아제는 조기 정지 코돈, 이상 스플라이스 수용자 부위 또는 이상 스플라이스 공여자 부위로부터의 상류 또는 하류 뉴클레오타이드 서열을 인식하고 그에 결합하도록 설계될 수 있다. 부위-특이적 뉴클레아제는 또한 조기 정지 코돈의 스킵핑을 유도하거나 또는 파괴된 리딩 프레임을 복원하기 위해 스플라이스 수용자 및 공여자를 표적화함으로써 정상 유전자 스플라이싱을 파괴하는데 사용될 수 있다. 부위-특이적 뉴클레아제는 게놈의 단백질-코딩 영역에 오프-타겟 변화를 매개할 수 있거나 또는 매개하지 않을 수 있다.
- [0180] b. 아데노-연관 바이러스 벡터
- [0181] 상기 기재된 바와 같은 조성물은 변형된 아데노-연관 바이러스 (AAV) 벡터를 포함한다. 변형된 AAV 벡터는 심장 근육 및 골격근 조직 향성을 증진시켰을 수 있다. 변형된 AAV 벡터는 포유동물의 세포에 부위-특이적 뉴클레아제를 전달하고 발현시킬 수 있다. 예를 들어, 변형된 AAV 벡터는 AAV-SASTG 벡터일 수 있다 (Piacentino et al. (2012) Human Gene Therapy 23:635-646). 변형된 AAV 벡터는 뉴클레아제를 골격근 및 심장 근육으로 생체내 전달할 수 있다. 변형된 AAV 벡터는 AAV1, AAV2, AAV5, AAV6, AAV8 및 AAV9를 비롯한 여러 캡시드 유형 중 하나 이상을 기반으로 할 수 있다. 변형된 AAV 벡터는 전신 및 국부 전달에 의해 골격근 또는 심장 근육으로 효율적으로 형질도입될 수 있는 AAV2 유사형과 대안적인 근육-향성 AAV 캡시드, 예컨대 AAV2/1, AAV2/6, AAV2/7, AAV2/8, AAV2/9, AAV2.5 및 AAV/SASTG 벡터를 기반으로 할 수 있다 (Seto et al. Current Gene Therapy (2012) 12:139-151).

[0182] c. CRISPR/Cas9-기반 시스템

[0183] 본 개시내용은 또한 상기 기재된 바와 같은 적어도 1개의 CRISPR/Cas9-기반 시스템의 DNA 표적화 시스템 또는 조성물을 제공한다. 이들 조성물은 게놈 편집, 게놈 조작, 및 유전 질환에 수반되는 유전자에서의 돌연변이의 효과를 교정하거나 또는 감소시키는데 사용될 수 있다. 조성물은 상기 기재된 바와 같은 Cas9 단백질 또는 Cas9 융합 단백질을 포함하는 CRISPR/Cas9-기반 시스템을 포함한다. CRISPR/Cas9-기반 시스템은 또한 상기 기재된 바와 같은 적어도 1개의 gRNA를 포함할 수 있다.

[0184] d. 멀티플렉스 CRISPR/Cas9-기반 시스템

[0185] 본 개시내용은 또한 상기 기재된 바와 같은 멀티플렉스 CRISPR/Cas9-기반 시스템을 제공한다. 이들 조성물은 게놈 편집, 게놈 조작, 및 유전 질환에 수반되는 유전자에서의 돌연변이의 효과를 교정하거나 또는 감소시키는데 사용될 수 있다. 이들 조성물은 1개 초과 유전자를 표적화하는데 사용될 수 있다. 조성물은 상기 기재된 바와 같은 Cas9 단백질 또는 Cas9 융합 단백질 및 상기 기재된 바와 같은 1개 초과 gRNA를 포함하는 CRISPR/Cas9-기반 시스템을 포함하는 변형된 렌티바이러스 벡터를 포함한다.

[0186] 9. 사용 방법

[0187] 조성물의 잠재적 적용은 과학 및 생명공학의 많은 영역에 걸쳐 다양하다. 개시된 조성물은 질환을 유발하는 유전자 돌연변이를 복구하는데 사용될 수 있다. 개시된 조성물은 유전자를 파괴하는데 사용되어, 유전자 파단이 근육 재생 또는 근육 강도의 증가 또는 근육 노화의 감소를 유도하도록 할 수 있다. 개시된 조성물은 골격근 또는 심장 근육으로부터 전신으로 발현되는 치료 유전자, 예컨대 응고 인자 또는 모노클로날 항체를 도입하는데 사용될 수 있다. 개시된 조성물은 포유동물 유전자 발현을 조절하는데 사용될 수 있다. 개시된 조성물은 세포를 전환분화시키거나 또는 세포의 분화를 유도하거나 또는 세포에서의 돌연변이체 유전자를 교정하는데 사용될 수 있다. 세포 및 유전자 요법, 유전적 재프로그래밍 및 재생 의학과 관련된 유전자의 활성화의 예가 제공된다. RNA-가이드 전사 활성화제는 세포 계통 사양을 재프로그래밍하는데 사용될 수 있다. 이들 인자의 강제 과다발현 보다는 세포 운명의 주요 조절제를 코딩하는 내인성 유전자의 활성화가 유전적 재프로그래밍, 전환분화 및/또는 분화 유도를 위한 보다 신속한, 효율적인, 안정한 또는 구체적인 방법을 잠재적으로 유도할 수 있다.

[0188] 10. 근육에서 게놈 편집하는 방법

[0189] 본 개시내용은 대상체의 골격근 또는 심장 근육에서 게놈 편집하는 방법에 관한 것이다. 방법은 대상체의 골격근 또는 심장 근육에 상기 기재된 바와 같은 골격근 또는 심장 근육에서의 게놈 편집을 위한 조성물을 투여하는 것을 포함한다. 게놈 편집은 돌연변이체 유전자를 교정하거나 또는 트랜스진을 삽입하는 것을 포함할 수 있다. 돌연변이체 유전자를 교정하는 것은 돌연변이체 유전자를 결실시키거나, 재배열하거나 또는 대체하는 것을 포함할 수 있다. 돌연변이체 유전자를 교정하는 것은 뉴클레아제-매개 NHEJ 또는 HDR을 포함할 수 있다.

[0190] 11. CRISPR/Cas9-기반 시스템의 사용 방법

[0191] CRISPR/Cas9-기반 시스템의 잠재적 적용은 과학 및 생명공학의 많은 영역에 걸쳐 다양하다. 개시된 CRISPR/Cas9-기반 시스템은 포유동물 유전자 발현을 조절하는데 사용될 수 있다. 개시된 CRISPR/Cas9-기반 시스템은 세포를 전환분화시키거나 또는 세포의 분화를 유도하거나 또는 세포에서의 돌연변이체 유전자를 교정하는데 사용될 수 있다. 세포 및 유전자 요법, 유전적 재프로그래밍 및 재생 의학과 관련된 유전자의 활성화의 예가 제공된다. RNA-가이드 전사 활성화제는 세포 계통 사양을 재프로그래밍하는데 사용될 수 있다. 재프로그래밍이 이들 실험에서 불완전하고 비효율적이었지만, iCas9-VP64/gRNA 조합의 반복된 형질감염, 이들 인자의 안정한 발현, 및 다중 유전자, 예컨대 뉴런 표현형으로의 전환분화를 위한 Ascl1 이외에도 Brn2 및 Myt1l의 표적화를 비롯한, 상기 방법을 개선할 수 있는 다수의 방식이 존재한다. 이들 인자의 강제 과다발현 보다는 세포 운명의 주요 조절제를 코딩하는 내인성 유전자의 활성화가 유전적 재프로그래밍 및 세포의 전환분화 또는 분화 유도를 위한 보다 신속한, 효율적인, 안정한 또는 구체적인 방법을 잠재적으로 유도할 수 있다. 최종적으로, 억제성 및 후성적-변형 도메인을 비롯한 다른 도메인에의 Cas9 융합은 RNA-가이드 전사 조절제의 보다 큰 다양성을 제공하여 포유동물 세포 조작을 위한 다른 RNA-기반 도구를 보완할 수 있다.

[0192] a. 유전자 발현의 활성화 방법

[0193] 본 개시내용은 상기 기재된 바와 같은 CRISPR/Cas9 기반 시스템을 사용하여 전사 활성화제를 RNA를 통해 프로모터에 표적화하는 것을 기반으로, 내인성 유전자, 예컨대 포유동물 유전자의 발현을 활성화시키기 위한 메커니즘

을 제공한다. 이것은 서열-특이적 DNA-결합 단백질의 조작을 기반으로 하는 이전에 기재된 방법과 근본적으로 상이하며, 표적화된 유전자 조절에 대한 기회를 제공할 수 있다. gRNA 발현 플라스미드의 생성은 간단하게 2개의 짧은 맞춤형 올리고뉴클레오타이드의 합성 단계 및 1개의 클로닝 단계를 수반하기 때문에, 많은 새로운 유전자 활성화제를 신속하고 경제적으로 생성하는 것이 가능하다. gRNA는 또한 시험관내 전사 후에 세포로 직접 감염시킬 수 있다. 단일 프로모터를 표적화하는 다중 gRNA가 제시되었지만, 다중 프로모터의 동시 표적화가 또한 가능할 수 있다. 단백질보다는 RNA로 게놈 표적 부위를 인식하는 것이 또한 후성적으로 변형된 부위, 예컨대 메틸화 DNA를 표적화하는 제한을 면할 수 있다.

[0194] DNA-결합 단백질의 조작을 기반으로 하는 현재 방법과 달리, 전사 활성화 도메인에 융합된 Cas9는 가이드 RNA 분자의 조합에 표적화되어 내인성 인간 유전자의 발현을 유도할 수 있다. 표적화된 유전자 활성화에 대한 이러한 간단하고 다양한 접근법으로 새로운 단백질의 조작에 대한 필요를 면하고 폭넓은 합성 유전자 조절이 가능해진다.

[0195] 방법은 세포 또는 대상체에게 상기 기재된 바와 같은 CRISPR/Cas9-기반 시스템, 상기 CRISPR/Cas9-기반 시스템을 코딩하는 폴리뉴클레오타이드 또는 벡터, 시스템 또는 적어도 1개의 CRISPR/Cas9-기반 시스템의 DNA 표적화 시스템 또는 조성물을 투여하는 것을 포함할 수 있다. 방법은 CRISPR/Cas9-기반 시스템을 투여하는 것, 예컨대 전사 활성화 도메인을 함유하는 Cas9 융합 단백질 또는 상기 Cas9 융합 단백질을 코딩하는 뉴클레오타이드 서열을 투여하는 것을 포함할 수 있다. Cas9 융합 단백질은 전사 활성화 도메인, 예컨대 aVP16 단백질 또는 전사 보조 활성화제, 예컨대 p300 단백질을 포함할 수 있다.

[0196] (1) dCas9-VP16

[0197] Cas9 융합 단백질은 전사 활성화 도메인, 예컨대 VP16 단백질을 포함할 수 있다. 전사 활성화 도메인은 적어도 1개의 VP16 단백질, 적어도 2개의 VP16 단백질, 적어도 3개의 VP16 단백질, 적어도 4개의 VP16 단백질 (즉, VP64 활성화제 도메인), 적어도 5개의 VP16 단백질, 적어도 6개의 VP16 단백질, 적어도 6개의 VP16 단백질 또는 적어도 10개의 VP16 단백질을 함유할 수 있다. Cas9 단백질은 뉴클레아제 활성이 불활성화된 Cas9 단백질일 수 있다. 예를 들어, 융합 단백질에서의 Cas9 단백질은 iCas9 (서열 1의 아미노산 36-1403)일 수 있으며, 이는 D10A 및 H840A의 아미노산 치환을 포함한다. Cas9 융합 단백질은 iCas9-VP64일 수 있다.

[0198] (2) dCas9-p300

[0199] Cas9 융합 단백질은 전사 공동-활성화 도메인, 예컨대 p300 단백질을 포함할 수 있다. p300 단백질 (또한 EP300 또는 E1A 결합 단백질 p300으로 공지됨)은 EP300 유전자에 의해 코딩되며, 신체 전체에 걸쳐 조직에서의 많은 유전자의 활성을 조절한다. p300 단백질은 세포 성장 및 분열을 조절하고, 세포가 성숙하여 특화된 기능을 맡도록 (분화시키도록) 촉진하고, 암성 종양의 성장을 방지하는데 역할을 한다. p300 단백질은 세포의 핵에서 전사를 수행하는 단백질의 복합체와 전사 인자를 연결함으로써 전사를 활성화시킨다. 전사 인자와의 p300 상호작용은 p300 도메인: 핵 수용체 상호작용 도메인 (RID), CREB 및 MYB 상호작용 도메인 (KIX), 시스테인/히스티딘 영역 (TAZ1/CH1 및 TAZ2/CH3) 및 인터페론 반응 결합 도메인 (IBiD) 중 하나 이상에 의해 다루어진다. p300의 마지막 4개 도메인 KIX, TAZ1, TAZ2 및 IBiD는 각각 전사 인자 p53의 전사활성화 도메인 9aaTAD 둘 다에 걸쳐있는 서열에 단단히 결합한다. 단백질은 염색질 재형성을 통해 전사를 조절하며 세포 증식 및 분화의 과정에서 중요한 히스톤 아세틸트랜스퍼라제로서 기능한다. 그것은 인산화 CREB 단백질에 특이적으로 결합함으로써 cAMP-유전자 조절을 매개한다.

[0200] p300 단백질은 테카펜타플렉스 상동체 7에 대한 모체, MAF, TSG101, 퍼옥시솜 증식자-활성화 수용체 알파, NPAS2, PAX6, DDX5, MYBL2, 테카펜타플렉스 상동체 1에 대한 모체, 테카펜타플렉스 상동체 2에 대한 모체, 림프성 인핸서-결합 인자 1, SNIP1, TRERF1, STAT3, EID1, RAR-관련 고아 수용체 알파, ELK1, HIF1A, ING5, 퍼옥시솜 증식자-활성화 수용체 감마, SS18, TCF3, Zif268, 에스트로젠 수용체 알파, GPS2, MyoD, YY1, ING4, PROX1, CITED1, HNF1A, MEF2C, MEF2D, MAML1, 트위스트 전사 인자, PTMA, IRF2, DTX1, 플랩 구조-특이적 엔도뉴클레아제 1, 근세포-특이적 인핸서 인자 2A, CDX2, BRCA1, HNRPU, STAT6, CITED2, RELA, TGS1, CEBPB, Mdm2, NCOA6, NFATC2, 갑상선 호르몬 수용체 알파, BCL3, TFAP2A, PCNA, P53 및 TAL1을 활성화시킬 수 있다.

[0201] 전사 공동-활성화 도메인은 인간 p300 단백질 또는 그의 단편을 포함할 수 있다. 전사 공동-활성화 도메인은 야생형 인간 p300 단백질 또는 돌연변이체 인간 p300 단백질 또는 그의 단편을 포함할 수 있다. 전사 공동-활성화 도메인은 인간 p300 단백질의 코어 리신-아세틸트랜스퍼라제 도메인, 즉 p300 HAT 코어 (또한 "p300 WT 코어"로 공지됨; 도 58 참조)를 포함할 수 있다. Cas9 단백질은 뉴클레아제 활성이 불활성화된 Cas9 단백질일 수

있다. 예를 들어, 융합 단백질에서의 Cas9 단백질은 iCas9 (서열 1의 아미노산 36-1403)일 수 있으며, 이는 D10A 및 H840A의 아미노산 치환을 포함한다. Cas9 융합 단백질은 iCas9-p300 WT 코어일 수 있다.

[0202] (3) gRNA

[0203] 방법은 또한 세포 또는 대상체에게 적어도 1개의 gRNA를 포함하는 CRISPR/Cas9-기반 시스템을 투여하는 것을 포함할 수 있으며, 여기서 gRNA는 상이한 DNA 서열을 표적화할 수 있다. 표적 DNA 서열은 중복될 수 있다. 세포에 투여되는 gRNA의 개수는 적어도 1개의 gRNA, 적어도 2개의 상이한 gRNA, 적어도 3개의 상이한 gRNA, 적어도 4개의 상이한 gRNA, 적어도 5개의 상이한 gRNA, 적어도 6개의 상이한 gRNA, 적어도 7개의 상이한 gRNA, 적어도 8개의 상이한 gRNA, 적어도 9개의 상이한 gRNA, 적어도 10개의 상이한 gRNA, 적어도 11개의 상이한 gRNA, 적어도 12개의 상이한 gRNA, 적어도 13개의 상이한 gRNA, 적어도 14개의 상이한 gRNA, 적어도 15개의 상이한 gRNA, 적어도 16개의 상이한 gRNA, 적어도 17개의 상이한 gRNA, 적어도 18개의 상이한 gRNA, 적어도 18개의 상이한 gRNA, 적어도 20개의 상이한 gRNA, 적어도 25개의 상이한 gRNA, 적어도 30개의 상이한 gRNA, 적어도 35개의 상이한 gRNA, 적어도 40개의 상이한 gRNA, 적어도 45개의 상이한 gRNA, 또는 적어도 50개의 상이한 gRNA일 수 있다. 세포에 투여되는 gRNA의 개수는 적어도 1개의 gRNA 내지 적어도 50개의 상이한 gRNA, 적어도 1개의 gRNA 내지 적어도 45개의 상이한 gRNA, 적어도 1개의 gRNA 내지 적어도 40개의 상이한 gRNA, 적어도 1개의 gRNA 내지 적어도 35개의 상이한 gRNA, 적어도 1개의 gRNA 내지 적어도 30개의 상이한 gRNA, 적어도 1개의 gRNA 내지 적어도 25개의 상이한 gRNA, 적어도 1개의 gRNA 내지 적어도 20개의 상이한 gRNA, 적어도 1개의 gRNA 내지 적어도 16개의 상이한 gRNA, 적어도 1개의 gRNA 내지 적어도 12개의 상이한 gRNA, 적어도 1개의 gRNA 내지 적어도 8개의 상이한 gRNA, 적어도 1개의 gRNA 내지 적어도 4개의 상이한 gRNA, 적어도 4개의 gRNA 내지 적어도 50개의 상이한 gRNA, 적어도 4개의 상이한 gRNA 내지 적어도 45개의 상이한 gRNA, 적어도 4개의 상이한 gRNA 내지 적어도 40개의 상이한 gRNA, 적어도 4개의 상이한 gRNA 내지 적어도 35개의 상이한 gRNA, 적어도 4개의 상이한 gRNA 내지 적어도 30개의 상이한 gRNA, 적어도 4개의 상이한 gRNA 내지 적어도 25개의 상이한 gRNA, 적어도 4개의 상이한 gRNA 내지 적어도 20개의 상이한 gRNA, 적어도 4개의 상이한 gRNA 내지 적어도 16개의 상이한 gRNA, 적어도 4개의 상이한 gRNA 내지 적어도 12개의 상이한 gRNA, 적어도 4개의 상이한 gRNA 내지 적어도 8개의 상이한 gRNA, 적어도 8개의 상이한 gRNA 내지 적어도 50개의 상이한 gRNA, 적어도 8개의 상이한 gRNA 내지 적어도 45개의 상이한 gRNA, 적어도 8개의 상이한 gRNA 내지 적어도 40개의 상이한 gRNA, 적어도 8개의 상이한 gRNA 내지 적어도 35개의 상이한 gRNA, 8개의 상이한 gRNA 내지 적어도 30개의 상이한 gRNA, 적어도 8개의 상이한 gRNA 내지 적어도 25개의 상이한 gRNA, 8개의 상이한 gRNA 내지 적어도 20개의 상이한 gRNA, 적어도 8개의 상이한 gRNA 내지 적어도 16개의 상이한 gRNA, 8개의 상이한 gRNA 내지 적어도 12개의 상이한 gRNA, 적어도 8개의 상이한 gRNA 내지 적어도 8개의 상이한 gRNA일 수 있다.

[0204] gRNA는 표적 DNA 서열의 상보적 폴리뉴클레오타이드 서열 및 그에 이어서 NGG를 포함할 수 있다. gRNA는 상보적 폴리뉴클레오타이드 서열의 5' 말단에 "G"를 포함할 수 있다. gRNA는 표적 DNA 서열의 적어도 10개 염기 쌍, 적어도 11개 염기 쌍, 적어도 12개 염기 쌍, 적어도 13개 염기 쌍, 적어도 14개 염기 쌍, 적어도 15개 염기 쌍, 적어도 16개 염기 쌍, 적어도 17개 염기 쌍, 적어도 18개 염기 쌍, 적어도 19개 염기 쌍, 적어도 20개 염기 쌍, 적어도 21개 염기 쌍, 적어도 22개 염기 쌍, 적어도 23개 염기 쌍, 적어도 24개 염기 쌍, 적어도 25개 염기 쌍, 적어도 30개 염기 쌍 또는 적어도 35개 염기 쌍의 상보적 폴리뉴클레오타이드 서열 및 그에 이어서 NGG를 포함할 수 있다. gRNA는 표적 유전자의 프로모터 영역, 인핸서 영역 또는 전사된 영역 중 적어도 하나를 표적화할 수 있다. gRNA는 서열 5-40, 65-144, 492-515, 540-563 및 585-625 중 적어도 하나의 핵산 서열을 포함할 수 있다.

[0205] b. 유전자 발현의 억제 방법

[0206] 본 개시내용은 내인성 유전자, 예컨대 포유동물 유전자의 발현을, 상기 기재된 바와 같은 CRISPR/Cas9 기반 시스템을 사용하여 RNA를 통해 게놈 조절 요소, 예컨대 원위 인핸서를 표적화하는 것을 기반으로 하여 억제하는 메커니즘을 제공한다. Cas9 융합 단백질은 전사 리프레서, 예컨대 KRAB 리프레서를 포함할 수 있다. Cas9 융합 단백질은 dCas9-KRAB일 수 있다. dCas9-KRAB는 표적화된 유전자좌에 이질염색질-형성 인자를 동원함으로써 후성적 유전자 조절에 추가로 영향을 미칠 수 있다. CRISPR/dCas9-KRAB 시스템은 유전자의 전사를 억제하는데 사용될 수 있지만, 또한 전형적인 억제 방법, 예컨대 RNA 간섭에 의해 이전에 접근하기 어려웠던 게놈 조절 요소를 표적화하는데 사용될 수 있다 (도 53B). 원위 인핸서를 표적화하는 gRNA를 갖는 dCas9-KRAB를 전달하는 것은 표적화된 인핸서에 의해 조절되는 다중 유전자의 발현을 파괴할 수 있다 (도 53C 참조). 표적화된 인핸서는 유전자에 대한 임의의 인핸서, 예컨대 HS2 인핸서일 수 있다.



- [0207] a. 전환분화 또는 분화 유도 방법
- [0208] 본 개시내용은 상기 기재된 바와 같은 CRISPR/Cas9-기반 시스템을 사용하여 RNA를 통해 내인성 유전자를 활성화 시킴으로써 세포를 전환분화시키거나 또는 세포의 분화를 유도하기 위한 메커니즘을 제공한다.
- [0209] (1) 전환분화
- [0210] CRISPR/Cas9-기반 시스템은 세포를 전환분화시키는데 사용될 수 있다. 계통 재프로그래밍화 또는 직접적 전환으로 또한 공지된 전환분화는 세포가 중간 만능 상태 또는 전구 세포 유형을 겪는 것 없이 한 분화 세포 유형으로부터 또 다른 것으로 전환되는 과정이다. 그것은 줄기 세포의 상호전환을 비롯한 모든 세포 운명 스위치를 포함하는 화생의 유형이다. 세포의 전환분화는 질량 모델링, 약물 발견, 유전자 요법 및 재생 의학을 위한 잠재적인 용도를 갖는다. 상기 기재된 CRISPR/Cas9 기반 시스템을 사용한 내인성 유전자, BRN2, MYT1L, ASCL1, NANOG 및/또는 MYOD1의 활성화는 여러 세포 유형, 예컨대 섬유모세포, 심근세포, 간세포, 연골세포, 중간엽 전구 세포, 조혈성 줄기 세포 또는 평활근 세포를 각각 뉴런 및 근원성 표현형으로 전환분화시킬 수 있다.
- [0211] (2) 분화 유도
- [0212] CRISPR/Cas9-기반 시스템은 세포, 예컨대 줄기 세포, 심근세포, 간세포, 연골세포, 중간엽 전구 세포, 조혈성 줄기 세포 또는 평활근 세포의 분화를 유도하는데 사용될 수 있다. 예를 들어, 줄기 세포, 예컨대 배아 줄기 세포 또는 만능 줄기 세포는 근육 세포 또는 혈관 내피 세포로 분화시키도록 유도될 수 있으며, 즉 뉴런 또는 근원성 분화를 유도할 수 있다.
- [0213] 12. 멀티플렉스 CRISPR/Cas9-기반 시스템의 사용
- [0214] 멀티플렉스 CRISPR/Cas9-기반 시스템은 sgRNA 디자인의 단순성 및 저비용의 이점을 이용하며, CRISPR/Cas9 기술을 사용하는 고처리량 게놈 연구의 진전을 조사하는데 도움을 줄 수 있다. 예를 들어, 본원에 기재된 단일 렌티바이러스 벡터는 곤란한 세포주, 예컨대 본원에 기재된 1차 섬유모세포에서 Cas9 및 다수의 sgRNA를 발현하는데 유용하다 (도 47). 멀티플렉스 CRISPR/Cas9-기반 시스템은 상기 기재된 바와 같은 CRISPR/Cas9-기반 시스템과 동일한 방식으로 사용될 수 있다.
- [0215] 기재된 전사 활성화 및 뉴클레아제 기능성 이외에도, 이러한 시스템은 게놈 아키텍처의 질의 및 내인성 유전자 조절의 경로를 비롯한 다양한 목적을 위한 후성적 변형을 제어하는 다른 신규 Cas9-기반 이펙터를 발현하기에 유용할 것이다. 내인성 유전자 조절이 다중 효소 사이에서 정교하게 균형을 이루기 때문에, 상이한 기능성을 갖는 Cas9 시스템을 멀티플렉스화하는 것은 상이한 조절 신호 사이의 복잡한 상호작용을 검사하는 것을 가능하게 할 것이다. 본원에 기재된 벡터는 sgRNA의 단일 세트를 사용하여 독립적 유전자 조작을 가능하게 하기 위해 압타머-변형 sgRNA 및 직교 Cas9와 상용성이어야 한다.
- [0216] 멀티플렉스 CRISPR/Cas9-기반 시스템은 세포에서 적어도 1개의 내인성 유전자를 활성화시키는데 사용될 수 있다. 방법은 세포를 변형된 렌티바이러스 벡터와 접촉시키는 것을 포함한다. 내인성 유전자는 일시적으로 활성화되거나 또는 안정하게 활성화될 수 있다. 내인성 유전자는 일시적으로 억제되거나 또는 안정하게 억제될 수 있다. 융합 단백질은 sgRNA와 유사한 수준으로 발현될 수 있다. 융합 단백질은 sgRNA와 비교하여 상이한 수준으로 발현될 수 있다. 세포는 1차 인간 세포일 수 있다.
- [0217] 멀티플렉스 CRISPR/Cas9-기반 시스템은 세포에서 멀티플렉스 유전자 편집 방법에 사용될 수 있다. 방법은 세포를 변형된 렌티바이러스 벡터와 접촉시키는 것을 포함한다. 멀티플렉스 유전자 편집은 돌연변이체 유전자를 교정하거나 또는 트랜스젠을 삽입하는 것을 포함할 수 있다. 돌연변이체 유전자를 교정하는 것은 돌연변이체 유전자를 결실시키거나, 재배열하거나 또는 대체하는 것을 포함할 수 있다. 돌연변이체 유전자를 교정하는 것은 뉴클레아제-매개 비-상동 말단 연결 또는 상동성-지정 복구를 포함할 수 있다. 멀티플렉스 유전자 편집은 적어도 1개의 유전자를 결실시키거나 또는 편집하는 것을 포함할 수 있으며, 여기서 유전자는 내인성 정상 유전자 또는 돌연변이체 유전자이다. 멀티플렉스 유전자 편집은 적어도 2개의 유전자를 결실시키거나 또는 교정할 수 있다. 예를 들어, 적어도 2개의 유전자, 적어도 3개의 유전자, 적어도 4개의 유전자, 적어도 5개의 유전자, 적어도 6개의 유전자, 적어도 7개의 유전자, 적어도 8개의 유전자, 적어도 9개의 유전자 또는 적어도 10개의 유전자가 결실되거나 또는 교정될 수 있다.
- [0218] 멀티플렉스 CRISPR/Cas9-기반 시스템은 세포에서 유전자 발현의 멀티플렉스 조절 방법에 사용될 수 있다. 방법은 세포를 변형된 렌티바이러스 벡터와 접촉시키는 것을 포함한다. 방법은 적어도 1개의 유전자의 유전자 발현 수준을 조절하는 것을 포함할 수 있다. 적어도 1개의 표적 유전자의 유전자 발현은 적어도 1개의 표적 유전자

의 유전자 발현 수준이 적어도 1개의 표적 유전자에 대한 정상 유전자 발현 수준과 비교하여 증가되거나 또는 감소되는 경우에 조절된다. 유전자 발현 수준은 RNA 또는 단백질 수준일 수 있다.

[0219] 13. 돌연변이체 유전자를 교정하고, 대상체를 치료하는 방법

[0220] 본 개시내용은 또한 대상체에서 돌연변이체 유전자를 교정하는 방법에 관한 것이다. 방법은 대상체의 골격근 또는 심장 근육에 상기 기재된 바와 같은 골격근 또는 심장 근육에서의 게놈 편집을 위한 조성물을 투여하는 것을 포함한다. 부위-특이적 뉴클레아제를 골격근 또는 심장 근육으로 전달하기 위한 조성물의 사용은 전체 유전자 또는 돌연변이를 함유하는 영역을 대체할 수 있는 복구 주형 또는 공여자 DNA로 완전-기능적 또는 부분-기능적 단백질의 발현을 복원할 수 있다. 부위-특이적 뉴클레아제는 표적화된 게놈 유전자좌에 부위-특이적 이중 가닥 파단을 도입하는데 사용될 수 있다. 부위-특이적 이중-가닥 파단은, 부위-특이적 뉴클레아제가 표적 DNA 서열에 결합하여, 그로 인해 표적 DNA의 절단을 허용하는 경우에 생성된다. 이러한 DNA 절단은 천연 DNA 복구 기구를 자극시켜, 2개의 가능한 복구 경로: 상동성-지정 복구 (HDR) 또는 비-상동 말단 연결 (NHEJ) 경로 중 하나를 유도할 수 있다.

[0221] 본 개시내용은 효율적으로 리딩 프레임 교정하고 유전 질환에 수반되는 기능적 단백질의 발현을 복원할 수 있는, 복구 주형 없이 부위-특이적 뉴클레아제를 사용하는 게놈 편집에 관한 것이다. 개시된 부위-특이적 뉴클레아제는 상동성-지정 복구 또는 뉴클레아제-매개 비-상동 말단 연결 (NHEJ)-기반 교정 접근법을 사용하는 것을 수반할 수 있으며, 이는 상동 재조합 또는 선택-기반 유전자 교정에 맞지 않을 수 있는 증식-제한 1차 세포주에서의 효율적인 교정을 가능하게 한다. 이러한 전략은 프레임시프트, 조기 정지 코돈, 이상 스플라이스 공여자 부위 또는 이상 스플라이스 수용자 부위를 유발하는 비필수 코딩 영역에서의 돌연변이에 의해 유발된 유전 질환의 치료를 위한 효율적인 유전자 편집 방법에 활성 부위-특이적 뉴클레아제의 신속하고 강한 조립을 통합시킨다.

[0222] a. 뉴클레아제 매개된 비-상동 말단 연결

[0223] 내인성 돌연변이된 유전자로부터의 단백질 발현의 복구는 주형-무함유 NHEJ-매개 DNA 복구를 통해서 이루어질 수 있다. 표적 유전자 RNA를 표적화하는 일시적 방법과 달리, 일시적으로 발현된 부위-특이적 뉴클레아제에 의한 게놈에서의 표적 유전자 리딩 프레임의 교정은 각각의 변형된 세포 및 모든 그의 자손에 의해 영구적으로 복원된 표적 유전자 발현을 유도할 수 있다.

[0224] 뉴클레아제 매개된 NHEJ 유전자 교정은 돌연변이된 표적 유전자를 교정할 수 있으며 HDR 경로에 대한 여러 잠재적인 이점을 제공한다. 예를 들어, NHEJ는 비특이적인 삽입 돌연변이유발을 야기할 수 있는 공여자 주형을 요구하지 않는다. HDR과 달리, NHEJ는 세포 주기의 모든 단계에서 효율적으로 작동하고, 따라서 주기 및 유사분열 후 세포 둘 다에서, 예컨대 근섬유에서 효과적으로 이용될 수 있다. 이것은 올리고뉴클레오타이드-기반 엑손 스킵핑 또는 정지 코돈의 약리학적 강제 번역-초과에 대한 강한 영구적 유전자 회복을 제공하며, 이론적으로 겨우 하나의 약물 치료를 요구할 수 있다. CRISPR/Cas9-기반 시스템 뿐만 아니라 메가뉴클레아제 및 아연 핑거 뉴클레아제를 포함하는 다른 조작된 뉴클레아제를 사용하는 NHEJ-기반 유전자 교정은 본원에 기재된 플라스미드 전기천공 접근법 이외에도 세포 및 유전자-기반 요법에 대한 다른 기존 생체외 및 생체내 플랫폼과 조합될 수 있다. 예를 들어, mRNA-기반 유전자 전달에 의한 또는 정제된 세포 투과가능한 단백질로서의 CRISPR/Cas9-기반 시스템의 전달은 삽입 돌연변이유발의 임의의 가능성을 면할 DNA-무함유 게놈 편집 접근법을 가능하게 할 수 있었다.

[0225] b. 상동성-지정 복구

[0226] 내인성 돌연변이된 유전자로부터의 단백질 발현의 복구는 상동성-지정 복구를 수반할 수 있다. 상기 기재된 바와 같은 방법은 세포에 공여자 주형을 투여하는 것을 추가로 포함한다. 공여자 주형은 완전-기능적 단백질 또는 부분-기능적 단백질을 코딩하는 뉴클레오타이드 서열을 포함할 수 있다. 예를 들어, 공여자 주형은 미니디스트로핀 ("minidy")으로 불리는 소형화된 디스트로핀 구축물, 돌연변이체 디스트로핀 유전자를 복원하기 위한 완전-기능적 디스트로핀 구축물, 또는 상동성-지정 복구 후에 돌연변이체 디스트로핀 유전자의 복원을 유도하는 디스트로핀 유전자의 단편을 포함할 수 있다.

[0227] c. CRISPR/Cas9를 사용하여, 돌연변이체 유전자를 교정하고, 대상체를 치료하는 방법

[0228] 본 개시내용은 또한 전체 유전자 또는 돌연변이를 함유하는 영역을 대체할 수 있는 복구 주형 또는 공여자 DNA로 완전-기능적 또는 부분-기능적 단백질의 발현을 복원하기 위한 CRISPR/Cas9-기반 시스템을 사용하는 게놈 편집에 관한 것이다. CRISPR/Cas9-기반 시스템은 표적화된 게놈 유전자좌에 부위-특이적 이중 가닥 파단을 도입

하는데 사용될 수 있다. 부위-특이적 이중-가닥 파단은, CRISPR/Cas9-기반 시스템이 gRNA를 사용하여 표적 DNA 서열에 결합하여, 그로 인해 표적 DNA의 절단을 허용하는 경우에 생성된다. CRISPR/Cas9-기반 시스템은 그의 성공적 및 효율적인 유전자 변형의 높은 비율로 인한 진보된 게놈 편집의 이점을 갖는다. 이러한 DNA 절단은 천연 DNA 복구 기구를 자극시켜, 2개의 가능한 복구 경로: 상동성-지정 복구 (HDR) 또는 비-상동 말단 연결 (NHEJ) 경로 중 하나를 유도할 수 있다. 예를 들어, 디스트로핀 유전자를 향해 지시된 CRISPR/Cas9-기반 시스템은 서열 65-115 중 어느 하나의 핵산 서열을 갖는 gRNA를 포함할 수 있다.

[0229] 본 개시내용은 효율적으로 리딩 프레임을 교정하고 유전 질환에 수반되는 기능적 단백질의 발현을 복원할 수 있는, 복구 주형 없이 CRISPR/Cas9-기반 시스템을 사용하는 게놈 편집에 관한 것이다. 개시된 CRISPR/Cas9-기반 시스템 및 방법은 상동성-지정 복구 또는 뉴클레아제-매개 비-상동 말단 연결 (NHEJ)-기반 교정 접근법을 사용하는 것을 수반할 수 있으며, 이는 상동 재조합 또는 선택-기반 유전자 교정에 맞지 않을 수 있는 증식-제한 1차 세포주에서의 효율적인 교정을 가능하게 한다. 이러한 전략은 프레임시프트, 조기 정지 코돈, 이상 스플라이스 공여자 부위 또는 이상 스플라이스 수용자 부위를 유발하는 비필수 코딩 영역에서의 돌연변이에 의해 유발된 유전 질환의 치료를 위한 효율적인 유전자 편집 방법에 활성 CRISPR/Cas9-기반 시스템의 신속하고 강한 조립을 통합시킨다.

[0230] 본 개시내용은 세포에서의 돌연변이체 유전자를 교정하고, 유전 질환, 예컨대 DMD를 앓고 있는 대상체를 치료하는 방법을 제공한다. 방법은 세포 또는 대상체에 CRISPR/Cas9-기반 시스템, 상기 CRISPR/Cas9-기반 시스템을 코딩하는 폴리뉴클레오티드 또는 벡터, 또는 상기 기재된 바와 같은 상기 CRISPR/Cas9-기반 시스템의 조성물을 투여하는 것을 포함할 수 있다. 방법은 CRISPR/Cas9-기반 시스템을 투여하는 것, 예컨대 뉴클레아제 활성을 갖는 제2 도메인을 함유하는 Cas9 단백질 또는 Cas9 융합 단백질을, 상기 Cas9 단백질 또는 Cas9 융합 단백질을 코딩하는 뉴클레오티드 서열, 및/또는 적어도 1개의 gRNA를 투여하는 것을 포함할 수 있으며, 여기서 gRNA는 상이한 DNA 서열을 표적화한다. 표적 DNA 서열은 중복될 수 있다. 세포에 투여되는 gRNA의 개수는 상기 기재된 바와 같이, 적어도 1개의 gRNA, 적어도 2개의 상이한 gRNA, 적어도 3개의 상이한 gRNA, 적어도 4개의 상이한 gRNA, 적어도 5개의 상이한 gRNA, 적어도 6개의 상이한 gRNA, 적어도 7개의 상이한 gRNA, 적어도 8개의 상이한 gRNA, 적어도 9개의 상이한 gRNA, 적어도 10개의 상이한 gRNA, 적어도 15개의 상이한 gRNA, 적어도 20개의 상이한 gRNA, 적어도 30개의 상이한 gRNA, 또는 적어도 50개의 상이한 gRNA일 수 있다. gRNA는 서열 65-115 중 적어도 하나의 핵산 서열을 포함할 수 있다. 방법은 상동성-지정 복구 또는 비-상동 말단 연결을 수반할 수 있다.

[0231] 14. 질환의 치료 방법

[0232] 본 개시내용은 치료를 필요로 하는 대상체를 치료하는 방법에 관한 것이다. 방법은 대상체의 조직에 상기 기재된 바와 같은 세포 또는 대상체 게놈 편집에서의 유전자 발현의 변경 또는 게놈 DNA의 조작 또는 변경을 위한 조성물을 투여하는 것을 포함한다. 방법은 대상체의 골격근 또는 심장 근육에 상기 기재된 바와 같은 골격근 또는 심장 근육에서의 게놈 편집을 위한 조성물을 투여하는 것을 포함할 수 있다. 대상체는 변성 또는 약화 또는 유전 질환을 유발하는 골격근 또는 심장 근육 병태를 앓고 있을 수 있다. 예를 들어, 대상체는 상기 기재된 바와 같은 뒤시엔느 근육 이영양증을 앓고 있을 수 있다.

[0233] a. 뒤시엔느 근육 이영양증

[0234] 상기 기재된 바와 같은 방법은 디스트로핀 유전자를 교정하고 상기 돌연변이된 디스트로핀 유전자의 완전-기능적 또는 부분-기능적 단백질 발현을 회복하는데 사용될 수 있다. 일부 측면 및 실시양태에서, 본 개시내용은 환자에서 DMD의 효과 (예를 들어, 임상 증상/징후)를 감소시키기 위한 방법을 제공한다. 일부 측면 및 실시양태에서, 본 개시내용은 환자에서 DMD를 치료하는 방법을 제공한다. 일부 측면 및 실시양태에서, 본 개시내용은 환자에서 DMD를 예방하는 방법을 제공한다. 일부 측면 및 실시양태에서, 본 개시내용은 환자에서 DMD의 추가의 진행을 예방하는 방법을 제공한다.

[0235] 15. 구축물 및 플라스미드

[0236] 상기 기재된 바와 같은 조성물은 본원에 개시된 바와 같은 CRISPR/Cas9-기반 시스템을 코딩하는 유전적 구축물을 포함할 수 있다. 유전적 구축물, 예컨대 플라스미드는 CRISPR/Cas9-기반 시스템, 예컨대 Cas9 단백질 및 Cas9 융합 단백질 및/또는 적어도 1개의 gRNA를 코딩하는 핵산을 포함할 수 있다. 상기 기재된 바와 같은 조성물은 본원에 개시된 바와 같은 변형된 AAV 벡터를 코딩하는 유전적 구축물 및 부위-특이적 뉴클레아제를 코딩하는 핵산 서열을 포함할 수 있다. 유전적 구축물, 예컨대 플라스미드는 부위-특이적 뉴클레아제를 코딩하는 핵산을 포함할 수 있다. 상기 기재된 바와 같은 조성물은 본원에 개시된 바와 같은 변형된 렌티바이러스 벡터를



코딩하는 유전적 구축물을 포함할 수 있다. 유전적 구축물, 예컨대 플라스미드는 Cas9-융합 단백질 및 적어도 1개의 sgRNA를 코딩하는 핵산을 포함할 수 있다. 유전적 구축물은 기능 염색체의 분자로서 세포에 존재할 수도 있다. 유전적 구축물은 동원체, 텔로미어 또는 플라스미드 또는 코스미드를 비롯한 선형 미니염색체일 수 있다.

[0237] 유전적 구축물은 또한 재조합 렌티바이러스, 재조합 아데노바이러스 및 재조합 아데노바이러스 연관 바이러스를 비롯한 재조합 바이러스 벡터의 계통의 일부일 수 있다. 유전적 구축물은 세포에서 살아있는 약독화 생 미생물 또는 재조합 미생물 벡터의 유전 물질의 일부일 수 있다. 유전적 구축물은 핵산의 코딩 서열의 유전자 발현을 위한 조절 요소를 포함할 수 있다. 조절 요소는 프로모터, 인핸서, 개시 코돈, 정지 코돈 또는 폴리아데닐화 신호일 수 있다.

[0238] 핵산 서열은 벡터일 수 있는 유전적 구축물을 구성할 수 있다. 벡터는 포유동물의 세포에서 융합 단백질, 예컨대 Cas9-융합 단백질 또는 부위-특이적 뉴클레아제를 발현할 수 있다. 벡터는 재조합일 수 있다. 벡터는 융합 단백질, 예컨대 Cas9-융합 단백질 또는 부위-특이적 뉴클레아제를 코딩하는 이중 핵산을 포함할 수 있다. 벡터는 플라스미드일 수 있다. 벡터는 Cas9-융합 단백질 또는 부위-특이적 뉴클레아제를 코딩하는 핵산으로 세포를 형질감염시키는데 유용할 수 있으며, 형질감염된 숙주 세포는 Cas9-융합 단백질 또는 부위-특이적 뉴클레아제 시스템의 발현이 일어나는 조건 하에서 배양 및 유지된다.

[0239] 코딩 서열은 안정성 및 높은 발현 수준에 대해 최적화될 수 있다. 일부 경우에, 코돈은 분자내 결합으로 인해 형성된 것과 같은 RNA의 2차 구조 형성을 감소시키도록 선택된다.

[0240] 벡터는 CRISPR/Cas9-기반 시스템 또는 부위-특이적 뉴클레아제를 코딩하는 이중 핵산을 포함할 수 있으며, CRISPR/Cas9-기반 시스템 또는 부위-특이적 뉴클레아제 코딩 서열의 상류에 있을 수 있는 개시 코돈, 및 CRISPR/Cas9-기반 시스템 또는 부위-특이적 뉴클레아제 코딩 서열의 하류에 있을 수 있는 정지 코돈을 추가로 포함할 수 있다. 개시 및 종결 코돈은 CRISPR/Cas9-기반 시스템 또는 부위-특이적 뉴클레아제 코딩 서열과 인 프레임으로 있을 수 있다. 벡터는 또한 CRISPR/Cas9-기반 시스템 또는 부위-특이적 뉴클레아제 코딩 서열에 작동가능하게 연결된 프로모터를 포함할 수 있다. CRISPR/Cas9-기반 시스템 또는 부위-특이적 뉴클레아제 코딩 서열에 작동가능하게 연결된 프로모터는 원숭이 바이러스 40 (SV40), 마우스 유방 종양 바이러스 (MMTV) 프로모터, 인간 면역결핍 바이러스 (HIV) 프로모터, 예컨대 소 면역결핍 바이러스 (BIV) 긴 말단 반복부 (LTR) 프로모터, 폴로니 바이러스 프로모터, 조류 백혈증 바이러스 (ALV) 프로모터, 시토메갈로바이러스 (CMV) 프로모터, 예컨대 CMV 극초기 프로모터, 엡스타인 바르 바이러스 (EBV) 프로모터 또는 라우스 육종 바이러스 (RSV) 프로모터로부터의 프로모터일 수 있다. 프로모터는 또한 인간 유전자, 예컨대 인간 유비퀴틴 C (hUbC), 인간 액틴, 인간 미오신, 인간 헤모글로빈, 인간 근육 크레아틴 또는 인간 메탈로티오네인으로부터의 프로모터일 수 있다. 프로모터는 또한 천연 또는 합성인 조직 특이적 프로모터, 예컨대 근육 또는 피부 특이적 프로모터일 수 있다. 이러한 프로모터의 예는 그의 내용 전문이 본원에 참조로 포함된 미국 특허 출원 공개 번호 US20040175727에 기재되어 있다.

[0241] 벡터는 또한 CRISPR/Cas9-기반 시스템 또는 부위-특이적 뉴클레아제의 하류에 있을 수 있는 폴리아데닐화 신호를 포함할 수 있다. 폴리아데닐화 신호는 SV40 폴리아데닐화 신호, LTR 폴리아데닐화 신호, 소 성장 호르몬 (bGH) 폴리아데닐화 신호, 인간 성장 호르몬 (hGH) 폴리아데닐화 신호 또는 인간  $\beta$ -글로빈 폴리아데닐화 신호일 수 있다. SV40 폴리아데닐화 신호는 pCEP4 벡터 (인비트로젠(Invitrogen), 캘리포니아주 샌디에고)로부터의 폴리아데닐화 신호일 수 있다.

[0242] 벡터는 또한 CRISPR/Cas9-기반 시스템, 즉, Cas9 단백질 또는 Cas9 융합 단백질 코딩 서열 또는 sgRNA 또는 부위-특이적 뉴클레아제 상류의 인핸서를 포함할 수 있다. 인핸서는 DNA 발현을 위해 필요할 수 있다. 인핸서는 인간 액틴, 인간 미오신, 인간 헤모글로빈, 인간 근육 크레아틴 또는 바이러스 인핸서, 예컨대 CMV, HA, RSV 또는 EBV로부터의 바이러스 인핸서일 수 있다. 폴리뉴클레오타이드 기능 인핸서는 각각의 내용이 완전히 참조로 포함된 미국 특허 번호 5,593,972, 5,962,428 및 W094/016737에 기재되어 있다. 벡터는 또한 벡터를 염색체외에서 유지하고 세포에서 벡터의 다중 카피를 생산하기 위해 포유동물 복제 기점을 포함할 수 있다. 벡터는 또한 벡터가 투여되는 포유동물 또는 인간 세포에서 유전자 발현에 매우 적합할 수 있는 조절 서열을 포함할 수 있다. 벡터는 또한 리포터 유전자, 예컨대 녹색 형광 단백질 ("GFP") 및/또는 선택 마커, 예컨대 히그로마이신 ("Hygro")을 포함할 수 있다.

[0243] 벡터는 완전히 참조로 포함된 문헌 [Sambrook et al., Molecular Cloning and Laboratory Manual, Second Ed., Cold Spring Harbor (1989)]을 비롯한 상용 기술 및 용이하게 입수가능한 출발 물질에 의해 단백질을 생산하기

위한 발현 벡터 또는 시스템일 수 있다. 일부 실시양태에서 벡터는 Cas9 단백질 또는 Cas9 융합 단백질을 코딩하는 핵산 서열을 포함하는, CRISPR/Cas9-기반 시스템을 코딩하는 핵산 서열, 및 서열 5-40, 65-144, 492-515, 540-563 및 585-625 중 적어도 하나의 핵산 서열을 포함하는 적어도 1개의 gRNA를 코딩하는 핵산 서열을 포함할 수 있다.

[0244] 16. 제약 조성물

[0245] 조성물은 제약 조성물로 존재할 수 있다. 제약 조성물은 약 1 ng 내지 약 10 mg의, CRISPR/Cas9-기반 시스템을 코딩하는 DNA 또는 CRISPR/Cas9-기반 시스템 단백질 성분, 즉 Cas9 단백질 또는 Cas9 융합 단백질을 포함할 수 있다. 제약 조성물은 약 1 ng 내지 약 10 mg의, 변형된 AAV 벡터의 DNA 및 부위-특이적 뉴클레아제를 코딩하는 뉴클레오티드 서열을 포함할 수 있다. 제약 조성물은 약 1 ng 내지 약 10 mg의 변형된 렌티바이러스 벡터의 DNA를 포함할 수 있다. 본 발명에 따른 제약 조성물은 사용될 투여 방식에 따라 제제화된다. 제약 조성물이 주사가능한 제약 조성물인 경우에, 그것들은 멸균, 발열원 무함유 및 미립자 무함유이다. 등장성 제제가 바람직하게 사용된다. 일반적으로, 등장성을 위한 첨가제는 염화나트륨, 텍스트로스, 만니톨, 소르비톨 및 락토스를 포함할 수 있다. 일부 경우에서, 등장성 용액, 예컨대 포스페이트 완충 염수가 바람직하다. 안정화제는 젤라틴 및 알부민을 포함한다. 일부 실시양태에서, 혈관수축 작용제가 제제에 첨가된다.

[0246] 조성물은 제약상 허용되는 부형제를 추가로 포함할 수 있다. 제약상 허용되는 부형제는 비히클, 아주반트, 담체 또는 희석제로서 기능적 분자일 수 있다. 제약상 허용되는 부형제는 표면 활성제, 예컨대 면역-자극 복합체 (ISCOMS), 프로인트 불완전 아주반트, 모노포스포릴 지질 A를 비롯한 LPS 유사체, 류라밀 펩티드, 퀴논 유사체, 소포, 예컨대 스쿠알렌 및 스쿠알렌, 히알루론산, 지질, 리포솜, 칼슘 이온, 바이러스 단백질, 다가음이온, 다가양이온 또는 나노입자를 포함할 수 있는 형질감염 촉진 작용제, 또는 다른 공지된 형질감염 촉진 작용제일 수 있다.

[0247] 형질감염 촉진 작용제는 다가음이온, 폴리-L-글루타메이트 (LGS)를 비롯한 다가양이온, 또는 지질이다. 형질감염 촉진 작용제는 폴리-L-글루타메이트이고, 보다 바람직하게는, 폴리-L-글루타메이트는 골격근 또는 심장 근육에서의 게놈 편집을 위한 조성물 중에 6 mg/ml 미만의 농도로 존재한다. 형질감염 촉진 작용제는 또한 표면 활성제, 예컨대 면역-자극 복합체 (ISCOMS), 프로인트 불완전 아주반트, 모노포스포릴 지질 A를 비롯한 LPS 유사체, 류라밀 펩티드, 퀴논 유사체 및 소포, 예컨대 스쿠알렌 및 스쿠알렌을 포함할 수 있고, 히알루론산은 또한 유전적 구축물과 함께 투여되는데 사용될 수 있다. 일부 실시양태에서, 조성물을 코딩하는 DNA 벡터는 또한 형질감염 촉진 작용제, 예컨대 지질, 레시틴 리포솜 또는 DNA-리포솜 혼합물 (예를 들어 W09324640 참조)로서 관련 기술분야에 공지된 다른 리포솜을 비롯한 리포솜, 칼슘 이온, 바이러스 단백질, 다가음이온, 다가양이온, 또는 나노입자, 또는 다른 공지된 형질감염 촉진 작용제를 포함할 수 있다. 바람직하게는, 형질감염 촉진 작용제는 다가음이온, 폴리-L-글루타메이트 (LGS)을 비롯한 다가양이온, 또는 지질이다.

[0248] 17. 전달 방법

[0249] 유전적 구축물을 제공하기 위한 제약 제제, 바람직하게는 상기 기재된 조성물을 전달하는 방법이 본원에 제공된다. 조성물의 전달은 세포에서 발현되어 세포의 표면으로 전달되는 핵산 분자로서 조성물의 형질감염 또는 전기천공일 수 있다. 핵산 분자는 바이오라드 진 펄서 엑스셀(BioRad Gene Pulser Xcell) 또는 아막사 뉴클레오펙터(Amaxa Nucleofector) IIb 장치를 사용하여 전기천공될 수 있다. 바이오라드 전기천공 용액, 시그마 (Sigma) 포스페이트-완충 염수 제품 #D8537 (PBS), 인비트로젠 OptiMEM I (OM) 또는 아막사 뉴클레오펙터 용액 V (N.V.)를 비롯한 여러 상이한 완충제가 사용될 수 있다. 형질감염은 형질감염 시약, 예컨대 리포펙타민 (Lipofectamine) 2000을 포함할 수 있다.

[0250] 조직으로 조성물의 전달 시 및 포유동물의 세포 내로의 벡터의 전달 시에, 형질감염된 세포는 융합 단백질을, 예컨대 CRISPR/Cas9-기반 시스템 및/또는 부위-특이적 뉴클레아제를 발현할 것이다. 조성물은 유전자 발현을 변경하거나 또는 게놈을 재-조작 또는 변경하기 위해 투여될 수 있다. 예를 들어, 조성물은 포유동물에서 디스트로핀 유전자를 교정하기 위해 포유동물에게 투여될 수 있다. 포유동물은 인간, 비-인간 영장류, 소, 돼지, 양, 염소, 영양, 들소, 물소, 소과, 사슴, 헛지호그, 코끼리, 라마, 알파카, 마우스, 래트 또는 닭, 및 바람직하게는 인간, 소, 돼지 또는 닭일 수 있다.

[0251] a. CRISPR/Cas9-기반 시스템

[0252] CRISPR/Cas9-기반 시스템 단백질 성분, 즉 Cas9 단백질 또는 Cas9 융합 단백질을 코딩하는 벡터는 생체내 전기천공, 리포솜 매개, 나노입자 촉진 및/또는 재조합 벡터와 함께 또는 그 없이 DNA 주사 (또한 DNA 백신접종으로

지침됨)에 의해 포유동물로 전달될 수 있다. 재조합 벡터는 임의의 바이러스 모드에 의해 전달될 수 있다. 바이러스 모드는 재조합 렌티바이러스, 재조합 아데노바이러스 및/또는 재조합 아데노-연관 바이러스일 수 있다.

[0253] CRISPR/Cas9-기반 시스템 단백질 성분, 즉 Cas9 단백질 또는 Cas9 융합 단백질을 코딩하는 뉴클레오티드는 표적 유전자를 유전자 교정하거나 또는 유전자의 유전자 발현을 변경하기 위해, 예컨대 내인성 유전자를 활성화시키거나 또는 억제하기 위해 세포 내로 도입될 수 있다. 예를 들어, gRNA에 의해 돌연변이체 디스트로핀 유전자를 향해 지시된 CRISPR/Cas9-기반 시스템 단백질 성분, 즉 Cas9 단백질 또는 Cas9 융합 단백질을 코딩하는 뉴클레오티드는 DMD 환자로부터의 근모세포 내로 도입될 수 있다. 대안적으로, 그것들은 DMD 환자로부터의 섬유모세포 내로 도입될 수 있고, 유전자 교정된 섬유모세포는 MyoD로 처리되어 근모세포로의 분화를 유도할 수 있으며, 이것은 대상체 내에, 예컨대 대상체의 손상된 근육 내로 이식되어 교정된 디스트로핀 단백질이 기능적이었다는 것을 검증하고/거나 대상체를 치료할 수 있다. 변형된 세포는 또한 줄기 세포, 예컨대 만능 줄기 세포, 골수-유래 전구세포, 골격근 전구세포, DMD 환자로부터의 인간 골격 근모세포, CD133+ 세포, 중간혈관모세포, 및 MyoD- 또는 Pax7-형질도입 세포 또는 다른 근원성 전구 세포일 수 있다. 예를 들어, CRISPR/Cas9-기반 시스템은 유도된 만능 줄기 세포의 뉴런 또는 근원성 분화를 유발할 수 있다.

[0254] 18. 투여 경로

[0255] 조성물은 경구로, 비경구로, 설하로, 경피로, 직장으로, 경점막으로, 국소로, 흡입을 통해, 협측 투여를 통해, 흉막내로, 정맥내, 동맥내, 복강내, 피하, 근육내, 비강내, 척수강내 및 관절내 또는 그의 조합을 비롯한 여러 경로에 의해 대상체에게 투여될 수 있다. 수의학적 용도를 위해, 조성물은 정상적인 수의학 실시에 따라 적합하게 허용가능한 제제로서 투여될 수 있다. 수의사는 특정한 동물에 가장 적절한 투여 요법 및 투여 경로를 용이하게 결정할 수 있다. 조성물은 전형적인 시린지, 무바늘 주사 장치, "미세입자 투사 총", 또는 다른 물리적 방법, 예컨대 전기천공 ("EP"), "유체역학 방법" 또는 초음파에 의해 투여될 수 있다.

[0256] 조성물은 생체내 전기천공, 리포솜 매개, 나노입자 촉진 및/또는 재조합 벡터, 예컨대 재조합 렌티바이러스, 재조합 아데노바이러스 및 재조합 아데노바이러스 연관된 바이러스와 함께 또는 그 없이 DNA 주사 (또한 DNA 백신 접종으로 지칭됨)를 비롯한 여러 기술에 의해 포유동물로 전달될 수 있다. 조성물은 골격근 또는 심장 근육 내로 주사될 수 있다. 예를 들어, 조성물은 전경골근 내로 주사될 수 있다.

[0257] 19. 세포 유형

[0258] 임의의 이들 전달 방법 및/또는 투여 경로는 무수한 세포 유형, 예를 들어 세포-기반 요법을 위해 현재 연구 하에 있는 세포 유형으로 이용될 수 있다. 세포 유형은 섬유모세포, 만능 줄기 세포, 심근세포, 간세포, 연골세포, 중간엽 전구 세포, 조혈성 줄기 세포, 평활근 세포 또는 K562 인간 적백혈병 세포주일 수 있다.

[0259] a. DMD

[0260] DMD의 세포-기반 요법을 위해 현재 연구 하에 있는 세포 유형은 불멸화 근모세포, 예컨대 야생형 및 DMD 환자 유래 세포주, 예를 들어 Δ48-50 DMD, DMD 8036 (del48-50), C25C14 및 DMD-7796 세포주, 1차 DMD 피부 섬유모세포 유도된 만능 줄기 세포, 골수-유래 전구세포, 골격근 전구세포, DMD 환자로부터의 인간 골격 근모세포, CD133+ 세포, 중간혈관모세포, 심근세포, 간세포, 연골세포, 중간엽 전구 세포, 조혈성 줄기 세포, 평활근 세포, 및 MyoD 또는 Pax7-형질도입 세포, 또는 다른 근원성 전구 세포를 포함한다. 인간 근원성 세포의 불멸화는 유전자 교정된 근원성 세포의 클론 도출에 사용될 수 있다. 세포는 유전자 교정된 디스트로핀 유전자를 함유하며 게놈의 단백질 코딩 영역에 다른 뉴클레아제-도입 돌연변이가 없는 불멸화 DMD 근모세포의 클론 집단을 단리하여 확장시키기 위해 생체외에서 변형될 수 있다. 대안적으로, 비-바이러스 또는 비-통합 바이러스 유전자 전달에 의한 또는 정제된 단백질 및 gRNA 함유 세포-침투 모티프의 직접적 전달에 의한 뉴클레아제의 일시적 생체내 전달은 최소의 외인성 DNA 통합 하에 또는 그의 위험 없이 계내에서 고도로 특이적인 교정을 가능하게 할 수 있다.

[0261] 20. 키트

[0262] 골격근 또는 심장 근육의 게놈을 편집하는데, 예컨대 돌연변이체 유전자를 교정하는데 사용될 수 있는 키트가 본원에 제공된다. 키트는 상기 기재된 바와 같은 골격근 또는 심장 근육에서의 게놈 편집을 위한 조성물 및 상기 조성물의 사용 지침서를 포함한다. 키트에 포함된 지침서는 포장 재료에 고정되어 있을 수 있거나 또는 포장 삽입물로서 포함될 수 있다. 지침서는 전형적으로 기록된 또는 인쇄된 재료이지만, 이러한 것으로 제한되지는 않는다. 이러한 지침을 저장하여 이를 최종 사용자에게 전달할 수 있는 임의의 매체가 본 개시내용에서 고려된다. 이러한 매체는 전자 저장 매체 (예를 들어, 자기 디스크, 테이프, 카트리지, 칩), 광학 매체 (예를 들어,

어, CD ROM) 등을 포함하나, 이에 제한되지는 않는다. 본원에 사용된 용어 "지침서"는 지침서를 제공하는 인터넷 사이트의 주소를 포함할 수 있다.

[0263] 골격근 또는 심장 근육의 게놈 편집을 위한 조성물은 상기 기재된 바와 같은 변형된 AAV 벡터 및 부위-특이적 뉴클레아제를 코딩하는 뉴클레오티드 서열을 포함할 수 있다. 부위-특이적 뉴클레아제는 돌연변이된 유전자에 특이적으로 결합하여 그를 절단하는, 상기 기재된 바와 같은 ZFN, TALEN 또는 CRISPR/Cas9-기반 시스템을 포함할 수 있다. 상기 기재된 바와 같은 부위-특이적 뉴클레아제는 돌연변이된 유전자의 특정한 영역에 특이적으로 결합하여 그를 표적화하기 위해 키트 내에 포함될 수 있다. 부위-특이적 뉴클레아제는 상기 기재된 바와 같은 돌연변이된 디스트로핀 유전자에 대해 특이적일 수 있다. 키트는 상기 기재된 바와 같은 공여자 DNA, gRNA 또는 트랜스진을 추가로 포함할 수 있다.

[0264] a. CRISPR/Cas9-기반 시스템

[0265] 돌연변이된 유전자를 교정하는데 사용될 수 있는 키트가 본원에 제공된다. 키트는 돌연변이된 유전자를 교정하기 위한 적어도 1개의 성분 및 CRISPR/Cas9-기반 시스템의 사용 지침서를 포함한다. 키트에 포함된 지침서는 포장 재료에 고정되어 있을 수 있거나 또는 포장 삽입물로서 포함될 수 있다. 지침서는 전형적으로 기록된 또는 인쇄된 재료이지만, 이러한 것으로 제한되지는 않는다. 이러한 지침을 저장하여 이를 최종 사용자에게 전달할 수 있는 임의의 매체가 본 개시내용에서 고려된다. 이러한 매체는 전자 저장 매체 (예를 들어, 자기 디스크, 테이프, 카트리지, 칩), 광학 매체 (예를 들어, CD ROM) 등을 포함하나, 이에 제한되지는 않는다. 본원에 사용된 용어 "지침서"는 지침서를 제공하는 인터넷 사이트의 주소를 포함할 수 있다.

[0266] 적어도 1개의 성분은 유전자를 특이적으로 표적화하는, 상기 기재된 바와 같은 적어도 1개의 CRISPR/Cas9-기반 시스템을 포함할 수 있다. 키트는 Cas9 단백질 또는 Cas9 융합 단백질을, 상기 Cas9 단백질 또는 Cas9 융합 단백질을 코딩하는 뉴클레오티드 서열, 및/또는 적어도 1개의 gRNA를 포함할 수 있다. 상기 기재된 바와 같은 CRISPR/Cas9-기반 시스템은 특정한 표적 영역 상류, 표적 유전자의 코딩 영역 내 또는 하류에 특이적으로 결합하여 그를 표적화하기 위해 키트 내에 포함될 수 있다. 예를 들어, CRISPR/Cas9-기반 시스템은 표적 유전자의 프로모터 영역에 대해 특이적일 수 있거나 또는 CRISPR/Cas9-기반 시스템은 돌연변이된 유전자, 예를 들어 상기 기재된 바와 같은 돌연변이된 디스트로핀 유전자에 대해 특이적일 수 있다. 키트는 상기 기재된 바와 같은 공여자 DNA를 포함할 수 있다.

[0267] 21. 실시예

[0268] 본원에 기재된 본 개시내용의 방법의 다른 적합한 변형 및 적합화는 용이하게 적용가능하고 이해가능하며, 본 개시내용 또는 본원에 개시된 측면 및 실시양태의 범주에서 벗어남 없이 적합한 등가물을 사용하여 이루어질 수 있다는 것이 통상의 기술자에게 자명할 것이다. 지금 본 개시내용에 상세하게 기재되었지만, 이는 하기 실시예를 참조로 하여 보다 명확하게 이해될 것이고, 실시예는 본 개시내용의 일부 측면 및 실시양태를 예시하기 위해 서만 단지 의도된 것이지, 본 발명의 범주를 제한하는 것으로 간주되어서는 안된다. 본원에 언급된 모든 저널 참고문헌, 미국 특허 및 공개문헌의 개시내용은 그 전문이 본원에 참조로 포함된다.

[0269] 본 발명은 하기 비제한적 실시예에 의해 예시되는 다중 측면을 갖는다.

[0270] 실시예 1

[0271] 물질 및 방법

[0272] 세포 배양 및 형질감염. HEK293T 세포를 듀크 대학교 암 센터 기관(Duke University Cancer Center Facilities)을 통해 아메리칸 티슈 콜렉션 센터(American Tissue Collection Center) (ATCC)로부터 획득하고, 이를 5% CO<sub>2</sub> 하에 37°C에서 10% 태아 소 혈청 및 1% 페니실린/스트렙토마이신으로 보충된 DMEM에서 유지하였다. HEK293T 세포를 리포펙타민 2000 (인비트로젠)으로 제조업체의 지침서에 따라 형질감염시켰다. 형질감염 효율은 일상적으로 대조군 eGFP 발현 플라스미드의 전달 후 형광 현미경검사에 의해 결정 시에 80% 초과였다. Cas9 발현 플라스미드를 개별 gRNA 발현 플라스미드에 또는 동등량의 4개의 gRNA의 혼합물로 이루어진 동일한 양의 gRNA 발현 플라스미드에 3:1의 질량 비로 형질감염시켰다.

[0273] 1차 마우스 배아 섬유모세포 (PMEF-HL, 밀리포어(Millipore), 매사추세츠주 빌리리카)를 24-웰 TCPS 플레이트(BD, 뉴저지주 프랭클린 레이크스)에 시딩하고 (웰당 75,000개), 10% 프리미엄 셀렉트(Premium Select) FBS (아트란타 바이올로지칼스(Atlanta Biologicals), 조지아주 로렌스빌), 25 µg mL<sup>-1</sup> 겐타미신 (인비트로젠), 1X 글루타맥스(GlutaMAX), 비-필수 아미노산, 피루브산나트륨 및 β-메르캅토에탄올 (인비트로젠)로 보충된 높은



글루코스 DMEM으로 이루어진 완전 MEF 배지에서 37℃ 및 5% CO<sub>2</sub>에서 유지하였다. MEF 형질감염은 이전에 기재된 바와 같이 (Adler et al. Molecular therapy. Nucleic acids 1, e32 (2012)) 혈청- 및 항생제-무함유 OptiMEM에서 폴리(CBA-ABOL)로 정전기적 축합 후에 양이온성 나노복합체로서 전달되는, 총 플라스미드 DNA의 단일 1 µg cm<sup>-2</sup> 용량으로 수행하였다. OptiMEM을 형질감염 4시간 후에 완전 MEF 배지로 대체하였다. 형질감염 48시간 후에, MEF를 qRT-PCR을 위해 처리하거나, 또는 완전 MEF 배지를 다음을 함유하는 N3 신경 유도 배지로 대체하였다: DMEM/F-12 (인비트로젠), 1X N-2 보충제 (인비트로젠), 10 ng mL<sup>-1</sup> 인간 bFGF2 (스텨젠티 (Stemgent), 매사추세츠주 캠브리지) 및 25 µg mL<sup>-1</sup> 겐타미신 (인비트로젠). GFP 리포터 벡터 (pmax-GFP, 3486 bp, 아막사, 독일 쾰른)를 형질감염 조건을 최적화하는데 사용하였다. Cas9 발현 플라스미드를 동등한 4개의 gRNA 발현 플라스미드의 혼합물에 3:1 또는 1:1의 질량 비로 형질감염시켰다.

[0274] 플라스미드. 야생형 및 H840A Cas9를 코딩하는 플라스미드는 애드진(Addgene)으로부터 획득하였다 (플라스미드 #39312 및 플라스미드 #39316; 문헌 [Jinek, et al. *Science* 337, 816-821 (2012)]). H840A Cas9는 D10A 돌연변이를 도입한 프라이머 쌍을 갖는 N-말단에 FLAG 에피토프 태그 및 핵 국재화 서열 (NLS)과 인 프레임으로 벡터 pcDNA3.1 내로 클로닝하였다. VP64 도메인, NLS 및 HA 에피토프 태그는 C-말단에 Cas9 ORF와 인 프레임으로 클로닝하였다 (도 1a, 도 9a). tracrRNA 및 crRNA 발현 카세트 (Cong et al. *Science* 339, 819-823 (2013))는 gBlocks (인티그레이티드 DNA 테크놀로지스(Integrated DNA Technologies (IDT)))로 주문하여 KpnI 및 SacII 부위를 갖는 pZDonor 플라스미드 (시그마) 내로 클로닝하였다. 키메라 가이드 RNA 발현 카세트 (Mali et al. *Science* 339, 823-826 (2013))는 또한 새로운 가이드 RNA 스페이서 서열의 신속한 클로닝을 촉진하기 위한 BbsI 제한 부위를 포함하는 변형을 갖는 gBlocks로 주문하였다 (도 9b). 표적 서열을 함유하는 올리고뉴클레오타이드는 IDT로부터 획득하고, 혼성화하고, 인산화하고, BbsI 부위를 사용하여 적절한 플라스미드에 클로닝하였다. 표적 서열은 표 2에 제공된다.

[0275] <표 2>

[0276] gRNA의 표적 서열

표적	명칭	서열	서열 번호
ASCL1	CR1	GCTGGGTGTCCCATTGAAA	5
	CR2	CAGCCGCTCGCTGCAGCAG	6
	CR3	TGGAGAGTTTGCAAGGAGC	7
	CR4	GTTTATTCAGCCGGGAGTC	8
NANOG	CR1	CGCCAGGAGGGGTGGGTCTA	9
	CR2	CCTTGGTGAGACTGGTAGA	10
	CR3	GTCTTCAGGTCTGTGTGCT	11
	CR4	ATATTCCTGATTTAAAAGT	12
VEGFA	CR1	TTAAAAGTCGGCTGGTAGC	13
	CR2	CGGGCCGGGGGCGGGGTCC	14
	CR3	GCCCGAGCCGCGTGTGGAA	15
	CR4	CCTTCATTGCGGCGGGCTG	16
TERT	CR1	CCGACCCCTCCCGGGTCCC	17
	CR2	CAGGACCGCGCTTCCCACG	18
	CR3	TGCACCCTGGGAGCGCGAG	19
	CR4	CCGCACGCACCTGTTCCCA	20
IL1B	CR1	AAAACAGCGAGGGAGAAAC	21
	CR2	TTAACTTGATTGTGAAATC	22
	CR3	AAAACAATGCATATTGCA	23
	CR4	AAAATCCAGTATTTTAATG	24
IL1R2	CR1	ACCCAGCACTGCAGCCTGG	25
	CR2	AACTTATGCGGCGTTTCCT	26
	CR3	TCACTTTAAAACACCTCT	27
	CR4	GCATCTTTTCTCTTTAAT	28
IL1RN	CR1	TGTACTCTCTGAGGTGCTC	29
	CR2	ACGCAGATAAGAACCAGTT	30
	CR3	CATCAAGTCAGCCATCAGC	31
	CR4	GAGTCACCCTCCTGGAAAC	32
HBG1/2	CR1	GCTAGGGATGAAGAATAAA	33
	CR2	TTGACCAATAGCCTTGACA	34
	CR3	TGCAAATATCTGTCTGAAA	35
	CR4	AAATTAGCAGTATCCTCTT	36
MYOD1	CR1	CCTGGGCTCCGGGGCGTTT	37
	CR2	GGCCCCCTGCGGCCACCCCG	38
	CR3	CTCCCTCCCTGCCCCGGTAG	39
	CR4	AGGTTTGGAAAGGGCGTGC	40

[0277]

[0278]

웨스턴 블롯. 세포를 50 mM 트리스-Cl (pH 7.4), 150 mM NaCl, 0.5% 트리톤 X-100 및 0.1% SDS 중에 용해시켰다. 용해물을 로딩 완충제와 혼합하고, 5분 동안 비등시키고, 동등 부피의 단백질을 NuPAGE® 노백스(Novex) 4-12% 또는 10% 비스-트리스 겔 폴리아크릴아미드 겔에서 전개시키고, 니트로셀룰로스 막으로 전달하였다. 비-특이적 항체 결합을 30분 동안 5% 탈지유를 함유하는 50 mM 트리스/150 mM NaCl/0.1% 트윈-20 (TBS-T)으로 차단하였다. 막을 밤새 1:1000으로 희석한 TBS-T 중 5% BSA 중 1차 항체 (HRP-접합 항-Flag (셀 시그널링(Cell Signaling), Cat#2044)); 30분 동안 1:5000으로 희석한 TBS-T 중 5% 밀크 중 항-GAPDH (셀 시그널링, 클론 14C10); 1:500으로 희석한 5% BSA 중 항-ASCL1 (산타 크루즈(Santa Cruz), 클론 sc-48449); 또는 1:500으로 희석한 5% 밀크 중 항-g-글로빈 (산타 크루즈, 클론 51-7)과 함께 인큐베이션하였고, 막을 30분 동안 TBS-T로 세척하였다. 1차 항체로 표지된 막을 30분 동안 1:5000으로 희석한 항-토끼 HRP-접합 항체 (시그마-알드리치(Sigma-Aldrich)), 항-염소 (1:3000) 또는 항-마우스 (1:5000)와 함께 인큐베이션하고, 30분 동안 TBS-T로 세척하였다. 막을 이문-스타 웨스턴씨(Immun-Star WesternC)™ 화학발광 키트 (바이오-라드(Bio-Rad))를 사용하여 가시화하였고, 영상을 케미독(ChemiDoc)™ XRS+ 시스템을 사용하여 포획하고, 이미지랩(ImageLab) 소프트웨어 (바이오-라드)를 사용하여 처리하였다.

[0279]

ELISA. 혈청-무함유 배양 배지 (OPTI-MEM)를 수집하고, -80℃에서 동결시켰다. 배양 배지 내로의 인간 IL-1ra 분비는 제조업체의 프로토콜 (알덴디 시스템즈(R&D Systems), Cat. 번호 DY280)에 따라 효소-연결된 면역흡착 검정 (ELISA)을 통해 정량화하였다. 표준 곡선은 OPTI-MEM 중에 재조합 인간 IL-1ra를 희석함으로써 제조하였고, 배양 배지 중 IL-1ra는 희석하지 않은 채로 측정하였다. 샘플을 20분 동안 3 kDa MWCO 필터 (아미콘 울트라(Amicon Ultra), Cat # UFC500396)를 통한 원심분리를 통해 약 8배로 농축시켰다. 보고된 값은 각 샘플에 대한 농도 인자에 의해 보정하였다.

[0280] 광학 밀도는 450 nm에서 측정하였으며, 540 nm에서 파장 보정하였다. 각 표준물 및 샘플을 이중으로 검정하였다. 이중 판독물을 평균내고, 평균 0 표준 광학 밀도를 차감하여 정규화하였다. 표준 곡선은 데이터를 log-변환하고 광학 밀도 대비 IL-1ra 농도의 선형 회귀를 수행함으로써 작성하였다. 보고된 값은 상이한 날에 수행한 3개의 독립적 실험 (n = 3)으로부터의 평균 및 평균의 표준 오차이며, 여기서 기술적 반복이 각 실험에 대해 평균이 내려졌다.

[0281] qRT-PCR. 총 RNA는 RNeasy 플러스 RNA 단리 키트 (퀴아젠(Qiagen))를 사용하여 단리하였다. cDNA 합성은 슈퍼스크립트(SuperScript)® VILO™ cDNA 합성 키트 (인비트로젠)를 사용하여 수행하였다. 퍼펙타(PerfeCTa)® SYBR® 그린 패스트믹스(Green FastMix)를 사용한 실시간 PCR은 프라이머3플러스(Primer3Plus) 소프트웨어를 사용하여 설계하고 IDT로부터 구입한 표 3에 보고된 올리고뉴클레오타이드 프라이머를 갖는 CFX96 실시간 PCR 검출 시스템 (바이오-라드)으로 수행하였다.

[0282] <표 3>

[0283] qRT-PCR에 사용된 프라이머의 서열

표적	정방향 프라이머	서열 번호	역방향 프라이머	서열 번호
<i>hASCL1</i>	GGAGCTTCTCGACTTCACCA	41	AACGCCACTGACAAGAAAGC	53
<i>NANOG</i>	GATTGTGGGCTGAAGAAA	42	CAGATCCATGGAGGAAGGAA	54
<i>VEGFA</i>	AAGGAGGAGGGCAGAATCAT	43	GGGTACTCTGGAAGATGTCC	55
<i>TERT</i>	AAACCTTCCTCAGCTATGCC	44	GTTTGCACGCATGTTCTC	56
<i>IL1B</i>	AGCTGATGGCCCTAAACAGA	45	AAGCCCTTGCTGTAGTGGTG	57
<i>IL1R2</i>	CAGGAGGACTCTGGCACCTA	46	CGGCAGGAAAGCATCTGTAT	58
<i>IL1RN</i>	GGAATCCATGGAGGGAAGAT	47	TGTTCTCGCTCAGGTCAGTG	59
<i>HBG1/2</i>	GCTGAGTGAATGCACTGTGA	48	GAATTCTTGGCGAAATGGA	60
<i>MYOD1</i>	CTCTCTGCTCCTTTGCCACA	49	GTGCTCTTCGGGTTTCAGGA	61
<i>GAPDH</i>	CAATGACCCCTTCATTGACC	50	TTGATTTTGGAGGGATCTCG	62
<i>mASCL1</i>	GGAACAAGAGCTGCTGGACT	51	GTTTTTCTGCCTCCCCATTT	63
<i>mGAPDH</i>	AACTTTGGCATTGTGAAGG	52	GGATGCAGGGATGATGTCT	64

[0284]

[0285] 프라이머 특이성을 아가로스 겔 전기영동 및 용융 곡선 분석에서 확인하였다. 적절한 동적 범위에 걸친 반응 효율을 계산하여 표준 곡선의 선형성을 확실히 하였다 (도 10). 결과는  $\Delta \Delta C_T$  방법에 의해 GAPDH 발현에 대해 정규화된 관심 유전자의 mRNA 발현 배수 증가로서 표현된다. 보고된 값은 상이한 날에 수행한 3개의 독립적 실험 (n = 3)으로부터의 평균 및 평균의 표준 오차이며, 여기서 기술적 반복이 각 실험에 대해 평균이 내려졌다.

[0286] RNA-seq. RNA seq 라이브러리를 구축하였다. 간략하게, 제1 가닥 cDNA는 슈퍼스크립트® VILO™ cDNA 합성 키트 (인비트로젠)를 사용하여 올리고 dT 디나비드(Dynabead)® (인비트로젠) 포획된 mRNA로부터 합성하였다. 제2 가닥 cDNA는 DNA 폴리머라제 I (뉴 잉글랜드 바이오랩스(New England Biolabs))를 사용하여 합성하였다. cDNA는 아젠코트(Agencourt) AMPure XP 비드 (베크만 쿨터(Beckman Coulter))를 사용하여 정제하고, 넥스테라(Nextera) 트랜스포사제 (일루미나(Illumina); 55°C에서 5분)를 사용하여 서열분석 프라이머를 단편화함과 동시에 이중-가닥 cDNA 내로 삽입하였다. 전위 반응은 QG 완충제 (퀴아젠)를 사용하여 정지시키고, 단편화된 cDNA는 AMPure XP 비드 상에서 정제하였다. 색인을 단 서열분석 라이브러리는 6 사이클의 PCR에 의해 생성하였다.

[0287] 라이브러리를 일루미나 HiSeq 2000 기기의 2개 레인 상에서 50-bp 단일 말단 판독물을 사용하여 서열분석하여, 라이브러리당 29백만 및 74백만 판독물을 생성하였다. 판독물을 보타이(Bowtie)를 사용하여 인간 RefSeq 전사체에 정렬시켰다 (Langmead et al. Genome biology 10, R25 (2009)). 다중 가설 검정을 위한 보정을 포함하는, 차등 발현의 통계적 유의성은 DESeq를 사용하여 계산하였다 (Anders et al. Genome biology 11, R106 (2010)). 미가공 RNA-seq 판독물 및 각 RefSeq 전사체에 정렬된 판독물의 개수는 진 익스프레션 옴니버스(Gene Expression Omnibus (GEO))에 공중의 접근을 위해 기탁되었으며, 수탁 번호는 현재 계류중이다.

[0288] 면역형광 염색. Tuj1 및 MAP2 발현의 검출을 위해, 형질감염된 MEF를 20분 동안 실온 (RT)에서 4% PFA를 갖는 N3 배지 (EMS, 펜실베이니아주 하트필드) 중 배양 제10일에 고정시켰다. 이어서, 세포를 0.2% 트리톤 X-100, 3% w/v BSA 및 10% 염소 혈청을 함유하는 차단 완충제 (시그마-알드리치, 미주리주 세인트 루이스)에서, 토끼 항-Tuj1 (코바스(Covance), 뉴저지주 프린스턴, 클론 Tuj1 1-15-79, 1:500) 및 마우스 항-MAP2 (BD, 클론 Ap20,



1:500)와 함께 실온에서 2시간 동안 또는 마우스 항-Ascl1 (BD 클론 24B72D11.1, 1:100)과 함께 4℃에서 추가의 24시간 동안 인큐베이션하였다. 이어서, 세포를 PBS로 3회 세척하고, 알렉사 플루오르(Alexa Fluor) 488 염소 항-마우스 IgG 및 알렉사 플루오르 594 염소 항-토끼 IgG를 갖는 차단 완충제 (인비트로젠, 1:200) 중 실온에서 1시간 동안 인큐베이션하고, PBS로 3회 세척하였다. 이어서, 염색된 MEF를 프로스캔II(ProScanII) 자동화 단계 (프라이어 사이언티픽(Prior Scientific), 매사추세츠주 록랜드)를 갖는 니콘 이클립스(Nikon Eclipse) TE2000-U 도립 형광 현미경으로 스캐닝하여 각각의 완전 배양 영역의 거대 모자이크 영상을 생성하였다. 이들 모자이크를 FIJI 매크로로 처리하여 국부 콘트라스트에 따라 각 영상을 자동적으로 및 균일하게 임계화하고, 소형 파편을 배제하고, 각 웰에서의  $Tuj1^{+}$  세포의 개수를 계수하였다.

[0289] 통계. 통계적 분석을 JMP 10 프로에서 알파가 0.05인 터키 검정에 의해 수행하였다.

[0290] 실시예 2

[0291] 결과

[0292] CRISPR/Cas9-기반 전사 활성화 시스템을 생성하기 위해, Cas9의 촉매 잔기 (D10A, H840A)를 돌연변이시켜 iCas9를 생성하고, C-말단 VP64 산성 전사활성화 도메인과 유전자 융합하였다 (도 1a,b). iCas9-VP64의 강한 발현이 N-말단 Flag 에피토프 태그의 웨스턴 블롯에 의해 인간 배아 신장 (HEK) 293T 세포에서 형질감염된 플라스미드로부터 관찰되었다 (도 3). CRISPR 시스템은 상보적 DNA 표적, 이어서 NGG 프로토스페이서-인접 모티프 (PAM) 서열 (여기서 N은 임의의 염기 쌍임)에 대하여 gRNA에서 20 bp 서열의 염기 쌍형성을 통해 그의 표적을 인식한다. 내인성 인간 프로모터를 표적화하는 합성 전사 인자의 조합은 유전자 발현의 상승작용적 및 강한 활성화를 유발한다. 따라서 4개의 gRNA 표적 부위 및 그에 이어서 NGG PAM 서열이 전사 개시 부위의 500 bp 내의 IL1RN 유전자의 프로모터에서 확인되었다 (도 4, 표 2). crRNA- 및 gRNA-기반 표적화 전략을 비교하기 위해, 4개의 표적 부위 서열을 crRNA 및 gRNA 발현 플라스미드<sup>17</sup> 내로 도입하고, iCas9-VP64 발현 플라스미드와 함께 HEK293T 세포 내로 공동-형질감염시켰다. IL1RN 발현의 실질적 유도가 crRNA의 조합으로 처리된 샘플에서 qRT-PCR에 의해 관찰되었지만, 보다 높은 수준은 gRNA의 조합에 의해 달성되었다 (도 1c). 유전자 발현에 대한 변화는 VP64 없이 gRNA 및 iCas9에 대한 발현 플라스미드로 처리된 세포에서 관찰되었으며, 이는 유전자 발현 조절에서 활성화 도메인의 중요한 역할을 입증한다 (도 1c). 이들 표적 부위에서의 뉴클레아제 활성은 iCas9-VP64 및 야생형 Cas9로 처리된 샘플에서의 DNA 복구 사건을 검출하기 위한 서베이어 검정을 수행함으로써 iCas9-VP64 시스템에서 사라진 것으로 확인되었다 (도 5). 4개의 gRNA 각각을 개별적으로 또는 조합으로 형질감염시킴으로써, gRNA의 조합에 의한 프로모터에서의 다중 부위 표적화가 유전자 발현의 강한 증가를 제시하였다 (도 1d). 조작된 전사 인자의 다른 부류에서 보여진 것처럼, IL1RN 발현의 높은 수준은 gRNA 조합이 iCas9-VP64와 공동-형질감염되었을 때만 관찰되었다 (도 1d). 유사하게, IL1RN 유전자에 의해 코딩된 IL-1 수용체 길항제 (IL-1ra) 단백질의 생산은 단지 3개의 상이한 실험에 걸쳐 gRNA의 조합으로 처리된 6개의 샘플 중 3개에서만 관찰되었고, 반면에 단일 gRNA 또는 대조군 플라스미드로 처리된 샘플에서는 결코 검출되지 않았다 (도 1e). iCas9-VP64에 의한 유전자 활성화의 특이성을 검사하기 위해, 4개의 gRNA의 조합으로 처리된 HEK293T 세포의 전반적 유전자 발현을 RNA-seq에 의해 평가하였다 (도 1f). 주목할 만하게, 대조군에 비해 유의하게 증가된 발현을 갖는 유일한 유전자 (오류 발견을  $\leq 3 \times 10^{-4}$ )는 IL1RN 유전자로부터 발현된 4개의 이소형이었으며 (도 4), 이는 유전자 활성화의 특이성의 높은 수준을 나타낸다.

[0293] 이러한 시스템의 일반적 적용가능성을 입증하기 위해, 4개의 gRNA는 ASCL1, NANOG, HBG1/2, MYOD1, VEGFA, TERT, IL1B, 및 IL1R2를 포함하는, 의약 및 생명공학과 관련된 8개의 다른 유전자의 각각의 프로모터를 표적화하도록 설계하였다 (도 4, 표 2). ASCL1 및 MYOD1의 강제 발현은 여러 세포 유형을 각각 뉴런 및 근원성 표현형으로 전환분화시키도록 유도한다. NANOG는 다능성의 마커이며, 또한 유전적 재프로그래밍 전략에 사용된다. 태아 발생 동안  $\gamma$ -글로빈을 코딩하는 상동체 HBG1 및 HBG2의 활성화는 겸상 적혈구 질환에서  $\beta$ -글로빈 돌연변이를 보충하기 위한 치료 전략으로서 사용될 수 있다. 합성 전사 인자에 의한 VEGFA의 상향조절은 조직 재생 및 상처 치유를 증진시키기 위한 전략으로서 조사되었다. TERT 유전자에 의해 코딩된 텔로머라제의 강제 발현이 세포주를 불멸화시키는데 사용될 수 있다. IL1B는 염증 및 자가면역을 매개하는 IL-1 $\beta$  시토카인을 코딩한다. IL-1 $\beta$  신호전달은 IL-1ra 또는 IL1R2에 의해 코딩된 디코이 수용체의 발현에 의해 차단될 수 있다. 각각의 이들 유전자의 발현은 qRT-PCR에 의해 결정된 바와 같이 iCas9-VP64 및 4개의 gRNA에 대한 발현 플라스미드를 HEK293T 세포 내로 공동-형질감염시켜 증진시켰다 (도 2). 일부 경우에서, 단일 gRNA의 발현은 유전자 발현을 유도하기에 충분하였지만, 모든 경우에서 4개의 gRNA의 공동-형질감염은 상승작용 효과를 유도하였다 (도 2a-d). 주목할 만하게, DNase-seq에 의해 결정된 바와 같은 염색질 접근성은 성공적인 유전자 활성화의 예측인

자가 아니었다 (도 4). RNA-seq는 iCas9-VP64 및 HBG1을 표적화하는 4개의 gRNA (이들 중 3개는 또한 HBG2를 표적화함)로 형질감염된 세포에서 수행하였다. 이것은, 비록 통계적 유의성이 낮은 전체 발현 수준으로 인해 달성되지는 않았지만, RNA-seq에 의해 구별될 수 없는 HBG1 및 HBG2 둘 다의 발현의 특이적 및 재현가능한 증가를 밝혀냈다 (도 6). iCas9-VP64 및 4개의 gRNA를 사용한 처리 후 Ascl1 및  $\gamma$ -글로빈의 단백질 발현의 증가를 웨스턴 블롯에 의해 검출하였으며 (도 7), 이는 qRT-PCR에 의해 관찰된 보다 높은 mRNA 수준을 확증해주었다 (도 2). Ascl1 및  $\gamma$ -글로빈 단백질 발현의 낮은 기준선 수준은 공백터 대조군에서 검출가능하였다. iCas9-VP64에 의한 유전자 표적의 활성화가 유전자 네트워크 및 세포 표현형에서의 2차 변화를 유도할 수 있다는 예비 증거로서, iCas9-VP64 및 ASCL1을 표적화하는 4개의 gRNA에 대한 발현 플라스미드를 뮤린 배아 섬유모세포 (MEF) 내로 공동-형질감염시켰다 (도 8). MEF에서의 Ascl1의 강제 발현은 하류 표적 Tuj1을 포함하는 뉴런 유전자 네트워크를 부분적으로 활성화시키는 것으로 나타났다. gRNA 표적 부위가 인간 및 마우스 ASCL1 프로모터 내에 보존되기 때문에 (도 8a), ASCL1 발현의 활성화는 또한 iCas9-VP64 및 4개의 gRNA를 코딩하는 플라스미드로 처리된 MEF에서 관찰되었다 (도 8b). 또한, Ascl1을 발현하는 세포 및 뉴런 마커 Tuj1은 iCas9-VP64/gRNA-처리 샘플에서의 형질감염 12일 후에 면역형광 염색에 의해 용이하게 검출되었다 (도 8c-h). Tuj1-양성 세포는 대조군 플라스미드로 처리된 세포에서 관찰되지 않았다.

[0294] 따라서 지금까지 포유동물 세포에서 Cas9/CRISPR 활성의 특이성의 포괄적 조사는 전혀 없었다. RNA-seq를 사용하여, 표적화된 유전자 활성화는 정교하게 특이적인 것으로 나타났으며, 이때 검출가능한 오프-타겟 유전자 활성화는 없었다 (도 1f, 도 6). 유전자 산물 IL1ra 및  $\gamma$ -글로빈은 HEK293T 세포에서 유전자 발현에 대한 2차 효과를 생성하지 않을 수 있기 때문에, IL1RN 및 HBG1/2를 상기 특이성 분석을 위해 선택하였다. 단일 강 활성화제를 사용하는 경우와 달리, 다중 약 전사 활성화제의 상승작용적 활성을 이용하는 것은, 다중 인접 오프-타겟 부위가 또 다른 유전자좌에 존재할 가능성이 없기 때문에, 특이적 유전자 조절을 증가시킬 수 있다. 흥미롭게도, IL32 유전자는 단지 공 발현 플라스미드로 처리된 대조군 샘플과 비교하여 iCas9-VP64 및 IL1RN- 또는 HBG1/2-표적화 gRNA로 처리된 샘플 둘 다에서 중등도로 하향조절되었다 (오류 발견율 < 0.03) (도 1f, 도 6). IL1RN 및 HBG1/2-표적화 샘플 둘 다가 유사하게 영향을 받았기 때문에, 이것이 표적 서열의 동일성과 관련된 오프-타겟 iCas9-VP64 활성의 결과일 가능성은 없다.

[0295] iCas9-VP64가 게놈에 결합하는 특이성을 평가하기 위해, ChIP 서열분석을 iCas9-VP64 및 IL1RN 프로모터를 표적화하는 4개의 gRNA로 처리된 세포 상에서 항-HA 항체를 사용하여 수행하였다. 실험은 iCas9가 IL1RN 프로모터를 표적화한다는 것을 밝혀냈다 (도 15). 더욱이, 실험은 매우 높은 수준의 특이성을 밝혀냈다. iCas9는 단지 10개의 잠재적 오프-타겟 결합 부위를 가졌다 (FDR < 5%). 특이성을 추가로 문의하기 위해, RNA 서열분석 실험을 iCas9 EGEM으로 수행하고, 단지 IL1RN 유전자 이소형만이 대조군에 비해 발현을 증가시켰다는 것이 밝혀졌다 (FDR  $\leq 3 \times 10^{-4}$ ).

[0296] 실시예 3

[0297] 디스트로핀 유전자를 표적화하는 CRISPR - 방법 및 물질

[0298] 플라스미드 구축물. 에스. 피오게네스 sgRNA 및 인간 코돈 최적화된 Cas9 (hCas9) 뉴클레아제에 대한 발현 카세트를 이전에 기재된 바와 같이 사용하였다 (Perez-Pinera et al., Nat Methods 10:973-976 (2013)). CRISPR/Cas9-변형 세포를 보강하기 위한 형광 리포터 시스템을 생성하기 위해, 다중 클로닝 부위 바로 상류의 T2A 스킵핑 펩티드에 융합된 Cas9 코딩 서열의 3' 말단 부분을 함유하는 진블록(GeneBlock) (IDT)을 합성하고, 이후 hCas9 발현 벡터 내로 클로닝하였다. 이어서, eGFP 리포터 유전자를 T2A 벡터 내로 클로닝하여, 동일한 발현 벡터 (hCas9-T2A-GFP, 서열 116)로부터 Cas9 및 eGFP 단백질의 공동-번역을 가능하게 하였다.

[0299] 세포 배양 및 형질감염. HEK293T 세포를 듀크 세포 배양 기관(Duke Cell Culture Facility)을 통해 미국 티슈 컬렉션 센터 (ATCC)로부터 획득하고, 이를 10% 소 송아지 혈청 및 1% 페니실린/스트렙토마이신으로 보충된 DMEM에서 유지하였다. 불멸화 근모세포 (Mamchaoui, K. et al. Skelet Muscle 1, 1-11 (2011)) (야생형 공여자로부터 1개 및 2개의  $\Delta 48-50$  DMD 환자 유래 세포주)를 20% 소 송아지 혈청 (시그마), 50  $\mu$ g/ml 페투인, 10 ng/ml 인간 표피 성장 인자 (시그마), 1 ng/ml 인간 기초 섬유모세포 성장 인자 (시그마), 10  $\mu$ g/ml 인간 인슐린 (시그마), 1% 글루타맥스 (인비트로젠) 및 1% 페니실린/스트렙토마이신 (인비트로젠)으로 보충된 골격근 배지 (프로모셀(PromoCell))에서 유지하였다. 1차 DMD 피부 섬유모세포를 코리엘 셀(Coriell Cell) 저장소로부터 획득하고 (GM05162A,  $\Delta 46-50$ ), 이를 10% 태아 소 혈청, 1 ng/ml 인간 기초 섬유모세포 성장 인자 및 1% 페니실린/스트렙토마이신으로 보충된 DMEM에서 유지하였다. 모든 세포주를 37°C 및 5% CO<sub>2</sub>에서 유지하였다.

[0300] HEK293T 세포를 24 웰 플레이트에서 400 ng의 각 발현 벡터와 함께 리포펙타민 2000 (인비트로젠)으로 제조업체

의 프로토콜에 따라 형질감염시켰다. 불멸화 근모세포 및 1차 섬유모세포는 각 세포주에 대한 최적화된 조건을 사용하여, 전기천공 완충제로서의 PBS와 함께, 진 펄서 엑스셀 (바이오라드)을 사용하는 전기천공에 의해 5  $\mu$ g의 각 발현 벡터로 형질감염시켰다 (도 1) (Ousterout et al. Mol Ther 21:1718-1726 (2013)). 형질감염 효율은 eGFP 발현 플라스미드 (pmaxGFP, 클론테크)를 전달하고 유동 세포측정법을 사용함으로써 측정하였다. 이들 효율은 일상적으로 HEK293T에 대해  $\geq 95\%$ 이고, 1차 섬유모세포 및 불멸화 근모세포에 대해  $\geq 70\%$ 였다. 모든 실험에 대해, 전기천공된 플라스미드의 나타난 질량은 각 CRISPR/Cas9-기반 시스템에 사용된 양에 상응한다.

[0301]

내인성 유전자 변형의 Cel-I 정량화 (서베이어 검정). 내인성 표적 부위에서의 CRISPR/Cas9-기반 시스템 유도 병변은 뉴클레아제-매개 NHEJ의 돌연변이 특성을 검출할 수 있는 서베이어 뉴클레아제 검정을 사용하여 정량화하였다 (Guschin, D.Y. et al. Meth Mol Biol 649, 247-256 (2010)). 형질감염 후에, 세포를 37°C에서 3 또는 10일 동안 인큐베이션하고, 게놈 DNA를 DNeasy 혈액 및 조직 키트 (키아젠)를 사용하여 추출하였다. 표적 유전자좌를 각 유전자좌 (표 4 참조), 예컨대 5'-GAGTTGGCTCAAATTGTTACTCTT-3' (서열 60) 및 5'-GGGAAATGGTCTAGGAGAGTAAAGT-3' (서열 61)에 특이적인 프라이머를 사용하여 아큐프라임(AccuPrime) 고 충실도 PCR 키트 (인비트로젠)로 35 사이클의 PCR에 의해 증폭시켰다.

[0302]

<표 4>

[0303]

인 실리코로 예측된 상위 10개의 오프 타겟 부위 및 Cas9와 나타난 sgRNA 발현 카세트로 형질감염된 HEK293T 세포에서 서베이어 검정에 의해 검출된 각 부위에서의 활성에 관한 요약. n.d.: 검출되지 않음.

서열 번호	가이드	표적	서열	PAM	점수	Chr	유전자	인트론/ 엑손	# MM	% indel
67	CR3	가이드	GCCTACTCAGACTGTTACTC	-	-	-	-	-	-	-
150		표적	tCCTACTCAGACTGTTACTC	TGG	-	X	DMD	엑손	1	13.0
151		OT1	tCCTACTCAcACTGTTACTC	AGG	7.4	1	STRIP1	인트론	2	9.3
152		OT2	aCCTgCTCAcACTGTTACTC	CAG	2.5	2	ARHGAP25	인트론	3	n.d.
153		OT3	GCaTtCTCAaACTGTTACTC	AGG	2.4	13	없음	없음	3	n.d.
154		OT4	GgaTtCTCAcACTGTTACTC	GGG	1.3	14	PGPEP1	엑손	4	n.d.
155		OT5	aCaTACTtAtACTGTTACTC	TAG	1.3	19	MDGA2	인트론	4	n.d.
156		OT6	taTtCTaAGACTGTTACTC	AAG	0.9	8	LPPR1	인트론	5	n.d.

[0304]

157		OT7	aaggACTaAGACTGTTACTC	GGG	0.9	9	RNF122	인트론	5	n.d.
158		OT8	GagctCTCAaCTGTTACTC	TAG	0.8	3	DNMBP	엑손	5	n.d.
159		OT9	GCaaAaTgAGACTGTTACTC	CAG	0.8	5	SLC12A2	인트론	4	n.d.
160		OT10	cCtAcATCAGACTGTTACTC	AAG	0.8	4	KCNIP4	인트론	4	n.d.
65		가이드	GATTGGCTTTGATTTCCTTA	-	-	-	-	-	-	-
161		표적	cATTGGCTTTGATTTCCTTA	GGG	-	X	DMD	인트론	1	8.3
162		OT1	aATTGGCATTGATTTCCTTA	GAG	7.1	16	없음	없음	2	0.8
163		OT2	cATTGGCTTTaATTTCCTTA	TAG	4.8	4	없음	없음	2	n.d.
164		OT3	GATaGGCTgTGATTTCCTTA	GAG	3.9	9	없음	없음	2	n.d.
165		OT4	GAaTaGCcTTGATTTCCTTA	AAG	2.4	1	없음	없음	3	n.d.
166		OT5	aATTGCTTTGATTTCCTTg	AGG	1.5	1	TIMM17A	인트론	3	n.d.
167		OT6	GATgtGCTTTGATTTCCTTt	GGG	1.4	17	MYO1D	인트론	3	n.d.
168		OT7	aATTGGGTTTaATTTCCTTA	AAG	1.1	8	PIK1A	인트론	3	n.d.
169		OT8	aATTGGGTTTGATTTCCTTt	TGG	1.1	11	MS4A1	인트론	3	n.d.
170		OT9	GATgGGGTTtATTTCCTTA	GAG	1.0	11	없음	없음	3	n.d.
171		OT10	GAaTGGGTTTGATTTCCTTg	GAG	1.0	11	없음	없음	3	n.d.
69		가이드	GCAGTTGCCTAAGAAGTGGT	-	-	-	-	-	-	-
172		표적	aCAGTTGCCTAAGAAGTGGT	GGG	-	X	DMD	인트론	1	14.0
173		OT1	cCAGTTGCTAAGAAGTGGg	GAG	1.5	5	NRG1	인트론	3	n.d.
174		OT2	GCAGTTGCCTgGAAGTGGT	AGG	1.4	X	없음	없음	2	n.d.
175		OT3	GCAGaTGcAgAAGAAGTGGT	GAG	1.4	19	SMIM7	인트론	3	n.d.
176		OT4	GCAGTTcCagAAGAAGTGGT	GAG	0.9	11	GLB1L2	인트론	3	n.d.
177		OT5	caAcTTGCCTAtGAAGTGGT	AGG	0.7	8	ASAP1	인트론	4	n.d.
178		OT6	aCAccTGCCTAAGAAGTGGa	GGG	0.7	11	없음	없음	4	n.d.
179		OT7	tCAGgTGgCTAAGAAGTGGg	TGG	0.7	14	NIN	인트론	4	n.d.
180		OT8	GaAGTTTGcCaAGAAGTGGa	GAG	0.6	7	없음	없음	4	n.d.
181		OT9	GcTGeTGCCcAAGAAGTGGc	AGG	0.6	11	AMOTL1	인트론	4	n.d.
182		OT10	tCAGtTGcCTAAGAAGcGGT	AAG	0.6	7	ACTR3C	인트론	4	n.d.
70		가이드	GGGGCTCCACCCTCACGAGT	-	-	-	-	-	-	-
183		표적	aGGGCTCCACCCTCACGAGT	GGG	-	X	DMD	인트론	1	19.9
184		OT1	GcaGCTCagCCCTCACGAGT	CAG	0.8	3	없음	없음	4	n.d.
185		OT2	GGGGCTtCAgCaTCACGAGT	GAG	0.8	8	없음	없음	3	n.d.
186		OT3	GGGGCTcCCCTCACtAGT	GAG	0.6	8	없음	없음	3	n.d.
187		OT4	GGGGaTCCACCCTCACaAGT	CAG	0.6	2	없음	없음	3	n.d.
188		OT5	aGGGCTggACCCTCACaAGT	AAG	0.4	16	AXIN1	인트론	4	n.d.
189		OT6	tGGtCTCCtCCcCACGAGT	GGG	0.4	2	없음	없음	4	n.d.
190		OT7	aGGGCTCCcaCCcCACGAGT	GAG	0.3	5	없음	없음	4	n.d.
191		OT8	GaGGCTCCAtaCTCACcAGT	GAG	0.3	11	없음	없음	4	n.d.
192		OT9	GGaGCTgCcCCtTCACGAGT	GGG	0.3	3	없음	없음	4	n.d.
193		OT10	atGaCTCCACCCTCAaGAGT	AAG	0.3	8	AGPAT5	없음	4	n.d.
100	CR36	가이드	GCCTTCTTTATCCCCTATCG	-	-	-	-	-	-	-

[0305]

194		표적	GCCTTCTTTATCCCCTATCG	AGG	-	X	DMD	인트론	0	20.6
195		OT1	GtCTgCTgTgTCCCCTATCG	GGG	1.3	21	없음	없음	4	n.d.
196		OT2	cCCTTCTcTATCCCCTgTCG	TGG	1.3	8	없음	없음	3	n.d.
197		OT3	GCCTTCTTTATCCCCTcTct	TGG	0.9	10	없음	없음	2	0.5
198		OT4	GCgcTCTTHCCCCATATct	TAG	0.6	16	없음	없음	4	n.d.
199		OT5	GCCcTCTgTcTCCCCTgTCG	CAG	0.5	1	NFASC	없음	4	n.d.
200		OT6	tCCAATCTHtTCCCCTATg	AGG	0.5	10	없음	없음	4	n.d.
201		OT7	aCCHTCTCTcTCCCCTATaG	AGG	0.5	5	LOC100996485	인트론	4	n.d.
202		OT8	GttTtCTTTTCCCCTATgG	GAG	0.5	3	없음	없음	4	n.d.
203		OT9	tgCTTCTTaATCCCCTATCa	AAG	0.4	7	없음	없음	4	n.d.
204		OT10	aCCTTCTTAcTCCCCTATCc	GGG	0.4	10	ADARB2	없음	4	n.d.

[0306]

[0307]

생성된 PCR 산물을 무작위로 용융시키고, 하기 프로그램으로 쉼 사이클러에서 재어닐링하였다: 단계 사이에 -0.3℃/초의 비율로 240초 동안 95℃, 이어서 60초 동안 85℃, 60초 동안 75℃, 60초 동안 65℃, 60초 동안 55℃, 60초 동안 45℃, 60초 동안 35℃, 및 60초 동안 25℃. 재어닐링 후에, 8 μL의 PCR 산물을 1 μL의 서베이어 뉴클레아제 S 및 1 μL의 인헨서 S (트랜스게노믹(Transgenomic))와 혼합하고, 42℃에서 1시간 동안 인큐베이션하였다. 인큐베이션 후에, 6 μL의 소화 산물을 10% TBE 폴리악릴아미드 겔 상에 로딩하고, 30분 동안 200V에서 전개시켰다. 겔을 브로민화메틸로 염색하고, 이미징랩 (바이오-라드)을 사용하여 이전에 기재된 바와 같은 밀도측정법에 의해 정량화하였다 (Guschin, et al. Meth Mol Biol 649, 247-256 (2010)).

[0308] 근모세포의 형광-활성화 세포 분류. DMD 근모세포를 5 마이크로그램의 각 hCas9-T2A-GFP 및 sgRNA 발현 벡터로 전기천공하고, 37℃ 및 5% CO<sub>2</sub>에서 인큐베이션하였다. 전기천공 3일 후에, 세포를 트립신처리하고, FACS벤티지 (FACSVantage) II 분류 기계를 사용하는 FACS 분류를 위해 수집하였다. GFP-양성 세포를 분석을 위해 수집하고 성장시켰다.

[0309] 게놈 결실을 검출하기 위한 PCR 기반 검정. 엑손 51 또는 엑손 45-55 유전자좌를 각 유전자좌에 플랭킹하는 프라이머를 사용하여 PCR (인비트로젠 아큐프라이 고 충실도 PCR 키트)에 의해 게놈 DNA로부터 증폭시켰다. 플랭킹 프라이머는 엑손 51에 대해 CeII-CR1/2-F 및 CeII-CR5-R이거나 또는 엑손 45-55 분석에 대해 CeII-CR6-F 및 CeII-CR36-R이었다 (표 4). PCR 산물을 TAE-아가로스 겔 상에서 분리하고, 분석을 위해 브로민화에티듬으로 염색하였다.

[0310] 전위의 PCR 기반 검출. 예측된 가능한 전위를 갖는 유전자좌는 Cas9 단독으로 (대조군) 또는 Cas9와 sgRNA로 형질감염된 세포로부터의 게놈 DNA의 2-단계 네스티드 PCR (각 단계에 대해 인비트로젠 아큐프라이 고 충실도 PCR 키트)에 의해 증폭시켰다. 제1 단계에서, 온-타겟 및 오프-타겟 sgRNA 표적 부위 각각에서 일어날 수 있는 전위는 클로닝 및 서열 분석을 촉진하기 위한 제한 부위를 포함하도록 변형된 각 유전자좌에 대해 서베이어 프라이머의 조합을 사용하여 35 사이클의 PCR에 의해 증폭시켰다 (표 4). 1 마이크로리터의 각 PCR 반응물을 각 개별 예측 전위에 대해 맞춤 설계된 네스티드 프라이머 세트를 사용하여 35 사이클의 PCR에 의해 제2 라운드의 증폭에 적용하였다 (표 4). 각 제2 네스티드 PCR 프라이머는 1차 앰플리콘 내의 동일한 대략적 영역 내에 결합하지만; 그러나, 각 쌍은 각 전위의 특이적 검출을 보장하도록 프라이머3 온라인 생물정보학 소프트웨어를 사용하여 최적화되었다. 예측된 전위의 예상된 길이에 상응하며 sgRNA로 처리된 세포에만 단지 존재하는 PCR 앰플리콘을 정제하고 (퀴아젠 겔 추출 키트), 생어(Sanger) 서열분석에 의해 분석하였다.

[0311] mRNA 분석. 불멸화 근모세포는 5일 동안 성장 배지를 1% 인슐린-트랜스페린-셀레늄 (인비트로젠 #51500056) 및 1% 페니실린/스트렙토마이신 (인비트로젠 #15140)으로 보충된 DMEM으로 대체함으로써 근섬유로 분화시킨 후에 세포를 트립신처리하고 수집하였다. 총 RNA는 이들 세포로부터 RNeasy 플러스 미니 키트 (퀴아젠)를 제조업체의 지침서에 따라 사용하여 분리하였다. RNA는 VILO cDNA 합성 키트 (라이프 테크놀로지스(Life Technologies) #11754) 및 1.5 마이크로그램의 RNA를 2시간 동안 42℃에서 제조업체의 지침서에 따라 사용하여 cDNA로 역전사시켰다. 표적 유전자좌는 엑손 44 및 52에 어닐링하여 CR1/5 또는 CR2/5에 의한 엑손 51 결실을 검출하는 프라이머, 또는 엑손 44 및 60에 어닐링하여 CR6/36에 의한 엑손 45-55 결실을 검출하는 프라이머를 사용하여 아큐프라이 고 충실도 PCR 키트 (인비트로젠)로 35 사이클의 PCR에 의해 증폭시켰다 (표 4). PCR 산물을 TAE-아가로스 겔 상에 전개시키고, 분석을 위해 브로민화에티듬으로 염색하였다. 해상된 PCR 밴드를 클로닝하고, 생어 서열분석에 의해 분석하여 예상된 엑손 접합부를 검증하였다. 표 5는 실시예 4에 사용된 프라이머의 서열을 열거한다.

[0312] <표 5>

서열 번호	프라이머 명칭	프라이머 서열	설명
205	Cell-CR1/2-F	GAGAGGTTATGTGGCTTTACCA	CR1/2에 대한 정방향 서베이어 프라이머
206	Cell-CR1-R	AAAAATGCTTCCCACTTTGC	CR1에 대한 역방향 서베이어 프라이머
207	Cell-CR2-R	CTCATTCTCATGCCTGGACA	CR2에 대한 역방향 서베이어 프라이머
208	Cell-CR3-F	GAGTTTGGCTCAAATTGTTACTCTT	CR3에 대한 정방향 서베이어 프라이머
209	Cell-CR3-R	GGGAAATGGTCTAGGAGAGTAAAG T	CR3에 대한 역방향 서베이어 프라이머
210	Cell-CR4/31-F	GTTTGGCTCAAATTGTTACTCTTCA	CR4 또는 CR31에 대한 정방향 서베이어 프라이머
211	Cell-CR4/31-R	GTGAGAGTAATGTGTTTGTGAGAG	CR4 또는 CR31에 대한 역방향 서베이어 프라이머
212	Cell-CR5-F	CGGGCTTGGACAGAACTTAC	CR5에 대한 정방향 서베이어 프라이머

[0313]



213	Cell-CR5-R	CTGCGTAGTGCCAAAACAAA	CR5 에 대한 역방향 서베이어 프라이머
214	Cell-CR6-F	TAATTTTCATTGAAGAGTGGCTGAA	CR6 에 대한 정방향 서베이어 프라이머
215	Cell-CR6-R	AAGCCCTGTGTGGTAGTAGTCAGT	CR6 에 대한 역방향 서베이어 프라이머
216	Cell-CR7-F	TGAGTCATGTTGGATAACCAGTCT	CR7 에 대한 정방향 서베이어 프라이머
217	Cell-CR7-R	GAAGGTCAGGAACATACAATTCAA	CR7 에 대한 역방향 서베이어 프라이머
218	Cell-CR10/11-F	GATATGGGCATGTCAGTTTCATAG	CR10 또는 CR11 에 대한 정방향 서베이어 프라이머
219	Cell-CR10/11-R	TGCTGTTGATTAATGGTTGATAGG	CR10 또는 CR11 에 대한 역방향 서베이어 프라이머
220	Cell-CR12/13-F	TTTAAATGCCATGTTTGTGTC	CR12 또는 CR13 에 대한 정방향 서베이어 프라이머
221	Cell-CR12/13-R	ATGAATAACCTAATGGGCAGAAAA	CR12 또는 CR13 에 대한 역방향 서베이어 프라이머
222	Cell-CR14/15-F	TCAAGTCGCTTCATTTGATAGAC	CR14 또는 CR15 에 대한 정방향 서베이어 프라이머
223	Cell-CR14/15-R	CACAACAAAACATATAGCCAAAGC	CR14 또는 CR15 에 대한 역방향 서베이어 프라이머
224	Cell-CR16/17-F	TGCTGCTAAAATAACACAAATCAGT	CR16 또는 CR17 에 대한 정방향 서베이어 프라이머
225	Cell-CR16/17-R	CTGTGCCTATTGTGGTTATCCTG	CR16 또는 CR17 에 대한 역방향 서베이어 프라이머
226	Cell-CR18/19-F	ATTGATCTGCAATACATGTGGAGT	CR18 또는 CR19 에 대한 정방향 서베이어 프라이머
227	Cell-CR18/19-R	TTTGCCTCTGCTATTACAGTATGG	CR18 또는 CR19 에 대한 역방향 서베이어 프라이머
228	Cell-CR20/21-F	TGTAGGGTGGTTGGCTAAAATAAT	CR20 또는 CR21 에 대한 정방향 서베이어 프라이머
229	Cell-CR20/21-R	TTTTTGCACAGTCAATAACACAAA	CR20 또는 CR21 에 대한 역방향 서베이어 프라이머
230	Cell-CR22/23-F	GGCTGGTCTCACAATTGTACTTTA	CR22 또는 CR23 에 대한 정방향 서베이어 프라이머
231	Cell-CR22/23-R	CATTATGGACTGAAAAATCTCAGCA	CR22 또는 CR23 에 대한 역방향 서베이어 프라이머
232	Cell-CR24/25-F	ATCATCCTAGCCATAACACAATGA	CR24 또는 CR25 에 대한 정방향 서베이어 프라이머
233	Cell-CR24/25-R	TTCAGCTTTAACGTGATTTTCTGT	CR24 또는 CR25 에 대한 역방향 서베이어 프라이머
234	Cell-CR26/27-F	GGATTTCAGAAAGCTGTTTACGAAGT	CR26 또는 CR27 에 대한 정방향 서베이어 프라이머
235	Cell-CR26/27-R	TTTAGCTGGATTGGAAAAACAAAT	CR26 또는 CR27 에 대한 역방향 서베이어 프라이머
236	Cell-CR28/29-F	AACTCACCCCATTTGTTGGTATATT	CR28 또는 CR29 에 대한 정방향 서베이어 프라이머

[0314]



237	Cell-CR28/29-R	CCTTGTCCAAATACCGAAATACAT	CR28 또는 CR29 에 대한 역방향 서베이어 프라이머
238	Cell-CR33-F	CACATAATTCATGAACTTGGCTTC	CR33 에 대한 정방향 서베이어 프라이머
239	Cell-CR33-R	TAGTAGCTGGGGAGGAAGATACAG	CR33 에 대한 역방향 서베이어 프라이머
240	Cell-CR34-F	TTTTGTITTAATGCGACTGTGT	CR34 에 대한 정방향 서베이어 프라이머
241	Cell-CR34-R	AGAAAAGGGGTTTTCTTTGACTT	CR34 에 대한 역방향 서베이어 프라이머
242	Cell-CR35-F	CATTGTGACTGGATGAGAAGAAAC	CR35 에 대한 정방향 서베이어 프라이머
243	Cell-CR35-R	AACGGCTGTTATTAAGTCCTCAG	CR35 에 대한 역방향 서베이어 프라이머
244	Cell-CR36-F	CAAGTCAGAAGTCACTTGCTTTGT	CR36 에 대한 정방향 서베이어 프라이머
245	Cell-CR36-R	TTTATGTGCGAGGAATCAGTCTGT	CR36 에 대한 역방향 서베이어 프라이머
246	Dys-E44-F	TGGCGGCGTTTTTCATTAT	엑손 44 에서의 정방향 RT-PCR 프라이머 결합
247	Dys-E52-R	TTCGATCCGTAATGATTGTTCTAGCC	엑손 52 에서의 역방향 RT-PCR 프라이머 결합
248	Dys-E60-R	GGTCTTCCAGAGTGCTGAGG	엑손 60 에서의 역방향 RT-PCR 프라이머 결합
249	CR3-Cell-OT1-F	TGTGTGCTTCTGTACACATCATCT	CR3 오프-타겟 1 에 대한 정방향 서베이어 프라이머
250	CR3-Cell-OT1-R	AGATTCAACCCTCAAAACTGAG	CR3 오프-타겟 1 에 대한 역방향 서베이어 프라이머
251	CR3-Cell-OT2-F	TAAACTCTTTCTTTCCGCAATTC	CR3 오프-타겟 2 에 대한 정방향 서베이어 프라이머
252	CR3-Cell-OT2-R	CAAGGTGACCTGCTACCTAAAAAT	CR3 오프-타겟 2 에 대한 역방향 서베이어 프라이머
253	CR3-Cell-OT3-F	TATGACCAAGGCTATGTGTTCACT	CR3 오프-타겟 3 에 대한 정방향 서베이어 프라이머
254	CR3-Cell-OT3-R	ACAGCCTCTCTCCAGTAACATTCT	CR3 오프-타겟 3 에 대한 역방향 서베이어 프라이머
255	CR3-Cell-OT4-F	TATTCTTGCACTGGTTTCACATTT	CR3 오프-타겟 4 에 대한 정방향 서베이어 프라이머
256	CR3-Cell-OT4-R	ATATTTTAAGCCAAGACCCAACAA	CR3 오프-타겟 4 에 대한 역방향 서베이어 프라이머
257	CR3-Cell-OT5-F	CTTCAACTGTCTGTCTGATTGCT	CR3 오프-타겟 5 에 대한 정방향 서베이어 프라이머
258	CR3-Cell-OT5-R	AACAGCCTCTCTTCATTGTCTCT	CR3 오프-타겟 5 에 대한 역방향 서베이어 프라이머
259	CR3-Cell-OT6-F	CTCTGGAACCTGTCTCTGTCTTGA	CR3 오프-타겟 6 에 대한 정방향 서베이어 프라이머
260	CR3-Cell-OT6-R	CTTTCCTGCGTTCTCATGTTACTA	CR3 오프-타겟 6 에 대한 역방향 서베이어 프라이머

[0315]

261	CR3-Cell-OT7-F	CCTTATATCCGTATCGCTCACTCT	CR3 오프-타겟 7 에 대한 정방향 서베이어 프라이머
262	CR3-Cell-OT7-R	CATATCTGTCTAACTTCCGCACAC	CR3 오프-타겟 7 에 대한 역방향 서베이어 프라이머
263	CR3-Cell-OT8-F	ACAGGTGTTATGTTGTCTGCATCT	CR3 오프-타겟 8 에 대한 정방향 서베이어 프라이머
264	CR3-Cell-OT8-R	ACTCCATTCCCAGATTAGTTATGC	CR3 오프-타겟 8 에 대한 역방향 서베이어 프라이머
265	CR3-Cell-OT9-F	CTGTTTTCTTTGTGAGAGTGGAGA	CR3 오프-타겟 9 에 대한 정방향 서베이어 프라이머
266	CR3-Cell-OT9-R	TGTAAGGTGGTCAAACCTTGCTCTA	CR3 오프-타겟 9 에 대한 역방향 서베이어 프라이머
267	CR3-Cell-OT10-F	TTTTTCCTAGTACCCACAGATTTT	CR3 오프-타겟 10 에 대한 정방향 서베이어 프라이머
268	CR3-Cell-OT10-R	TCCCTGATTCTCTCATTTGTGTA	CR3 오프-타겟 10 에 대한 역방향 서베이어 프라이머
269	CR1-Cell-OT1-F	TTGGGAACATCAGAGAAAGTATGA	CR1 오프-타겟 1 에 대한 정방향 서베이어 프라이머
270	CR1-Cell-OT1-R	ACAAATTACAGTCTCCTGGGAAAG	CR1 오프-타겟 1 에 대한 역방향 서베이어 프라이머
271	CR1-Cell-OT2-F	AGTAGCTTACCTTGGCAGAGAAAA	CR1 오프-타겟 2 에 대한 정방향 서베이어 프라이머
272	CR1-Cell-OT2-R	TGACATACTGTTACCCTTTGCACT	CR1 오프-타겟 2 에 대한 역방향 서베이어 프라이머
273	CR1-Cell-OT3-F	GAAAGGCTCAGTGAATGTTGTT	CR1 오프-타겟 3 에 대한 정방향 서베이어 프라이머
274	CR1-Cell-OT3-R	CACTGCATCATCTCATTAATCAA	CR1 오프-타겟 3 에 대한 역방향 서베이어 프라이머
275	CR1-Cell-OT4-F	CCCATATATTCATGATTACCCACA	CR1 오프-타겟 4 에 대한 정방향 서베이어 프라이머
276	CR1-Cell-OT4-R	TATCAGAACGAGCACTAAAAGCAC	CR1 오프-타겟 4 에 대한 역방향 서베이어 프라이머
277	CR1-Cell-OT5-F	TTGGGAGGCTGAGGTACAAG	CR1 오프-타겟 5 에 대한 정방향 서베이어 프라이머
278	CR1-Cell-OT5-R	GAATGAAAAACAAACAGAAGGTGA	CR1 오프-타겟 5 에 대한 역방향 서베이어 프라이머
279	CR1-Cell-OT6-F	CTCCTCATCTGTACCCTTCAATCT	CR1 오프-타겟 6 에 대한 정방향 서베이어 프라이머
280	CR1-Cell-OT6-R	AGAGTGGCATCTAGTGTCACTGAG	CR1 오프-타겟 6 에 대한 역방향 서베이어 프라이머
281	CR1-Cell-OT7-F	TACCAAAAGCTTCTCCTGTTACC	CR1 오프-타겟 7 에 대한 정방향 서베이어 프라이머
282	CR1-Cell-OT7-R	GTAAGTTGGATGGCCTATTCTTTG	CR1 오프-타겟 7 에 대한 역방향 서베이어 프라이머
283	CR1-Cell-OT8-F	GAAGGAAATGCAAGGATACAAGAT	CR1 오프-타겟 8 에 대한 정방향 서베이어 프라이머
284	CR1-Cell-OT8-R	TGATTGAAAGAATCATTCCAGAAA	CR1 오프-타겟 8 에 대한 역방향 서베이어 프라이머

[0316]

285	CR1-Cell-OT9-F	TCAGAAGGAAAATTGAAATTGGTT	CR1 오프-타겟 9 에 대한 정방향 서베이어 프라이머
286	CR1-Cell-OT9-R	CAGATGTGTTCTTCATCATTCCTC	CR1 오프-타겟 9 에 대한 역방향 서베이어 프라이머
287	CR1-Cell-OT10-F	TTCTCTTTAGGGAAAAGCTCTCAAA	CR1 오프-타겟 10 에 대한 정방향 서베이어 프라이머
288	CR1-Cell-OT10-R	GGGTATAGATCATATGGAGGGAAG	CR1 오프-타겟 10 에 대한 역방향 서베이어 프라이머
289	CR5-Cell-OT1-F	AGATGATCTGCCACCTCAG	CR5 오프-타겟 1 에 대한 정방향 서베이어 프라이머
290	CR5-Cell-OT1-R	CTTTCTTCCTCATTTAGTGGAAT	CR5 오프-타겟 1 에 대한 역방향 서베이어 프라이머
291	CR5-Cell-OT2-F	ATGAATTGCAGATTGATGGTACTG	CR5 오프-타겟 2 에 대한 정방향 서베이어 프라이머
292	CR5-Cell-OT2-R	TCTCACCAAGAACCAAATTGTCTA	CR5 오프-타겟 2 에 대한 역방향 서베이어 프라이머
293	CR5-Cell-OT3-F	GTAGGATACCTTGGAACAGTCTT	CR5 오프-타겟 3 에 대한 정방향 서베이어 프라이머
294	CR5-Cell-OT3-R	TTAAGCAATTGTGAGATTTGCTGT	CR5 오프-타겟 3 에 대한 역방향 서베이어 프라이머
295	CR5-Cell-OT4-F	TCAGAAAGTCAAGTAGCACACACA	CR5 오프-타겟 4 에 대한 정방향 서베이어 프라이머
296	CR5-Cell-OT4-R	AGAAGCACACACTCAGGTAAAGC	CR5 오프-타겟 4 에 대한 역방향 서베이어 프라이머
297	CR5-Cell-OT5-F	TCTTTGGGGGAATAATGACTAAAA	CR5 오프-타겟 5 에 대한 정방향 서베이어 프라이머
298	CR5-Cell-OT5-R	TTTGGCATTATGGGAATAAAACT	CR5 오프-타겟 5 에 대한 역방향 서베이어 프라이머
299	CR5-Cell-OT6-F	ACTAATTCTGGTCAAGCCCATCA	CR5 오프-타겟 6 에 대한 정방향 서베이어 프라이머
300	CR5-Cell-OT6-R	TTAAGACATCGGATGAACAGAAAG	CR5 오프-타겟 6 에 대한 역방향 서베이어 프라이머
301	CR5-Cell-OT7-F	AGAAGCTTTCTGACATGATCTGC	CR5 오프-타겟 7 에 대한 정방향 서베이어 프라이머
302	CR5-Cell-OT7-R	TCAATTGCATTAGGACTTAGACCA	CR5 오프-타겟 7 에 대한 역방향 서베이어 프라이머
303	CR5-Cell-OT8-F	GTAAATTACCTGTGAAGCCCTTG	CR5 오프-타겟 8 에 대한 정방향 서베이어 프라이머
304	CR5-Cell-OT8-R	CGGAAAACAGATCCACTTTATGAT	CR5 오프-타겟 8 에 대한 역방향 서베이어 프라이머
305	CR5-Cell-OT9-F	AAATCCACTGGAAACATCTTGAGT	CR5 오프-타겟 9 에 대한 정방향 서베이어 프라이머
306	CR5-Cell-OT9-R	AGTCTCTTCAGAATCATGCCCTAT	CR5 오프-타겟 9 에 대한 역방향 서베이어 프라이머
307	CR5-Cell-OT10-F	GCTTGGTGGCACATACCTGTAG	CR5 오프-타겟 10 에 대한 정방향 서베이어 프라이머
308	CR5-Cell-OT10-R	GGTAGGTAGATTTGCTTGCTTGT	CR5 오프-타겟 10 에 대한 역방향 서베이어 프라이머

[0317]

309	CR6-Cell-OT1-F	AGCTCTCAGCAGAGTAGGGATTTA	CR6 오프-타겟 1 에 대한 정방향 서베이어 프라이머
310	CR6-Cell-OT1-R	GTGAGTCTACTGCACCCCATC	CR6 오프-타겟 1 에 대한 역방향 서베이어 프라이머
311	CR6-Cell-OT2-F	TGACACTGTGAAGTCAATTCTGTC	CR6 오프-타겟 2 에 대한 정방향 서베이어 프라이머
312	CR6-Cell-OT2-R	TCAAGAACTTGACAATGAGCAAAT	CR6 오프-타겟 2 에 대한 역방향 서베이어 프라이머
313	CR6-Cell-OT3-F	TATCCGATCCACTGTTGTGTGT	CR6 오프-타겟 3 에 대한 정방향 서베이어 프라이머
314	CR6-Cell-OT3-R	CAGGAGACCCAAAACCACTCTAC	CR6 오프-타겟 3 에 대한 역방향 서베이어 프라이머
315	CR6-Cell-OT4-F	TTGTCTACAAATAGGGCTTCCTT	CR6 오프-타겟 4 에 대한 정방향 서베이어 프라이머
316	CR6-Cell-OT4-R	TGTTAAGTTGGGCTTATGTTCTT	CR6 오프-타겟 4 에 대한 역방향 서베이어 프라이머
317	CR6-Cell-OT5-F	CACAAGTCTCACTGCACAAACAT	CR6 오프-타겟 5 에 대한 정방향 서베이어 프라이머
318	CR6-Cell-OT5-R	TGACCCATGATTATCTCTCTTTGA	CR6 오프-타겟 5 에 대한 역방향 서베이어 프라이머
319	CR6-Cell-OT6-F	TTCAGCTTCTGATTGGTTTAAATG	CR6 오프-타겟 6 에 대한 정방향 서베이어 프라이머
320	CR6-Cell-OT6-R	CCAATTCCTTAATTTCCCTACAG	CR6 오프-타겟 6 에 대한 역방향 서베이어 프라이머
321	CR6-Cell-OT7-F	ATCTCAGACCAGGAGGGAGAC	CR6 오프-타겟 7 에 대한 정방향 서베이어 프라이머
322	CR6-Cell-OT7-R	CCTCAGGGTCAGTACATTTTCAG	CR6 오프-타겟 7 에 대한 역방향 서베이어 프라이머
323	CR6-Cell-OT8-F	TTCTTAGGACATTGCTCCACATAC	CR6 오프-타겟 8 에 대한 정방향 서베이어 프라이머
324	CR6-Cell-OT8-R	GCAAACATAATGCAACTCGTAATC	CR6 오프-타겟 8 에 대한 역방향 서베이어 프라이머
325	CR6-Cell-OT9-F	GCAAGGGAGTCTGTGTCTTTG	CR6 오프-타겟 9 에 대한 정방향 서베이어 프라이머
326	CR6-Cell-OT9-R	TCATTTAAGTGGCTGTTCTGTGTT	CR6 오프-타겟 9 에 대한 역방향 서베이어 프라이머
327	CR6-Cell-OT10-F	ACAAAACAGAGAGAAAAGGCAGAG	CR6 오프-타겟 10 에 대한 정방향 서베이어 프라이머
328	CR6-Cell-OT10-R	GTTTTGATTTCTGGTGCCTACAG	CR6 오프-타겟 10 에 대한 역방향 서베이어 프라이머
329	CR36-Cell-OT1-F	ACTGAAGCTGAAGCCAGTC	CR36 오프-타겟 1 에 대한 정방향 서베이어 프라이머
330	CR36-Cell-OT1-R	ACATGAGCTCTCAGGTTTCTGAC	CR36 오프-타겟 1 에 대한 역방향 서베이어 프라이머
331	CR36-Cell-OT2-F	TCAAACCTAGATGGTTCCTATGTT	CR36 오프-타겟 2 에 대한 정방향 서베이어 프라이머
332	CR36-Cell-OT2-R	GTACCCTGAAAATGTAGGGTGACT	CR36 오프-타겟 2 에 대한 역방향 서베이어 프라이머

[0318]

333	CR36-Cell-OT3-F	CACTTCCCAAGTGAGGCAAT	CR36 오프-타겟 3에 대한 정방향 서베이어 프라이머
334	CR36-Cell-OT3-R	CTATACTTGGGGCTGACTTGCTAC	CR36 오프-타겟 3에 대한 역방향 서베이어 프라이머
335	CR36-Cell-OT4-F	TCGTATAGTTACTTTGGCTCACA	CR36 오프-타겟 4에 대한 정방향 서베이어 프라이머
336	CR36-Cell-OT4-R	AGGGATCTTTACTCCTCAGTGTGT	CR36 오프-타겟 4에 대한 역방향 서베이어 프라이머
337	CR36-Cell-OT5-F	TGTAGAAGTTGGAATATCCTGCTG	CR36 오프-타겟 5에 대한 정방향 서베이어 프라이머
338	CR36-Cell-OT5-R	GTCAACAATTTGATCTCAGGCTTC	CR36 오프-타겟 5에 대한 역방향 서베이어 프라이머
339	CR36-Cell-OT6-F	CTCAGTACTAAAGATGGACGCTTG	CR36 오프-타겟 6에 대한 정방향 서베이어 프라이머
340	CR36-Cell-OT6-R	AATCATTTTCAGTCTTCCCAACAAT	CR36 오프-타겟 6에 대한 역방향 서베이어 프라이머
341	CR36-Cell-OT7-F	GGAATCACAGTAGATGTTTGTC	CR36 오프-타겟 7에 대한 정방향 서베이어 프라이머
342	CR36-Cell-OT7-R	AGACCAGGAGGTAAGAACATTTTG	CR36 오프-타겟 7에 대한 역방향 서베이어 프라이머
343	CR36-Cell-OT8-F	CCACATAGAAAAGAGACTTGCAGAA	CR36 오프-타겟 8에 대한 정방향 서베이어 프라이머
344	CR36-Cell-OT8-R	AGAGATGCCAAAAGAACAGTCAAT	CR36 오프-타겟 8에 대한 역방향 서베이어 프라이머
345	CR36-Cell-OT9-F	TGTGCCTTAGGCTATGTAACTGT	CR36 오프-타겟 9에 대한 정방향 서베이어 프라이머
346	CR36-Cell-OT9-R	AAACCCCTTGTAAACAAAATTACCA	CR36 오프-타겟 9에 대한 역방향 서베이어 프라이머
347	CR36-Cell-OT10-F	TAACTGCATCAGAAGTCCTTGCTA	CR36 오프-타겟 10에 대한 정방향 서베이어 프라이머
348	CR36-Cell-OT10-R	GGAGACCAAGCTGCTAAAGTCA	CR36 오프-타겟 10에 대한 역방향 서베이어 프라이머
349	Cell-CR3-F-네스티드	GTGGTGccgcggGAGTTTGGCTCAAAT TGTTACTCTT	네스티드 PCR 제 1 라운드 프라이머
350	Cell-CR3-R-네스티드	GTGGTGccgcggGGGAAATGGTCTAG GAGAGTAAAGT	네스티드 PCR 제 1 라운드 프라이머
351	Cell-CR1-F-네스티드	GTGGTGccgcggGAGAGGTTATGTGGC TTACCA	네스티드 PCR 제 1 라운드 프라이머
352	Cell-CR1-R-네스티드	GTGGTGccgcggCTCATTCATGCCT GGACA	네스티드 PCR 제 1 라운드 프라이머
353	Cell-CR5-F-네스티드	GTGGTGccgcggCGGGCTTGACAG ACTTAC	네스티드 PCR 제 1 라운드 프라이머
354	Cell-CR5-R-네스티드	GTGGTGccgcggCTGCGTAGTGCCAAA ACAAA	네스티드 PCR 제 1 라운드 프라이머
355	Cell-CR6-F-네스티드	GTGGTGccgcggTAATTTTCATTGAAGA GTGGCTGAA	네스티드 PCR 제 1 라운드 프라이머
356	Cell-CR6-R-네스티드	GTGGTGccgcggAAGCCCTGTGTGGTA GTAGTCAGT	네스티드 PCR 제 1 라운드 프라이머

[0319]

357	Cell-CR36-F-네스티드	GTGGTGccgcggCAAGTCAGAAGTCAC TTGCTTTGT	네스티드 PCR 제 1 라운드 프라이머
358	Cell-CR36-R-네스티드	GTGGTGccgcggTTTTATGTGCAGGAA TCAGTCTGT	네스티드 PCR 제 1 라운드 프라이머
359	CR3-Cell-OT1-F-네스티드	GTGGTGccgcggTGTGTGCTTCTGTAC ACATCATCT	네스티드 PCR 제 1 라운드 프라이머
360	CR3-Cell-OT1-R-네스티드	GTGGTGccgcggAGATTTCAACCCTCA AAAACTGAG	네스티드 PCR 제 1 라운드 프라이머
361	CR1-Cell-OT1-F-네스티드	GTGGTGccgcggTTGGGAACATCAGAG AAAGTATGA	네스티드 PCR 제 1 라운드 프라이머
362	CR1-Cell-OT1-R-네스티드	GTGGTGccgcggACAAATTACAGTCTC CTGGGAAAG	네스티드 PCR 제 1 라운드 프라이머
363	CR36-Cell-OT3-F-네스티드	GTGGTGccgcggCACTTCCCAAGTGAG GCAAT	네스티드 PCR 제 1 라운드 프라이머
364	CR36-Cell-OT3-R-네스티드	GTGGTGccgcggCTATACTTGGGGCTG ACTTGCTAC	네스티드 PCR 제 1 라운드 프라이머
365	CR3-P1/P3-F	GTGGTGccgcggTTGGCTCTTTAGCTT GTGTTTC	네스티드 PCR 제 2 라운드 프라이머
366	CR3-P1/P3-R	GTGGTGccgcggTGAGACTCCCAAAGG CAATC	네스티드 PCR 제 2 라운드 프라이머
367	CR3-P1/P4-F	GTGGTGccgcggTTGGCTCTTTAGCTT GTGTTTC	네스티드 PCR 제 2 라운드 프라이머
368	CR3-P1/P4-R	GTGGTGccgcggACTGAGGGGTGATCT TGGTG	네스티드 PCR 제 2 라운드 프라이머
369	CR3-P2/P3-F	GTGGTGccgcggGCAGAGAAAGCCAG TCGGTA	네스티드 PCR 제 2 라운드 프라이머
370	CR3-P2/P3-R	GTGGTGccgcggTGAGACTCCCAAAGG CAATC	네스티드 PCR 제 2 라운드 프라이머
371	CR3-P2/P4-F	GTGGTGccgcggGCAGAGAAAGCCAG TCGGTA	네스티드 PCR 제 2 라운드 프라이머
372	CR3-P2/P4-R	GTGGTGccgcggACTGAGGGGTGATCT TGGTG	네스티드 PCR 제 2 라운드 프라이머
373	CR1-P1/P5-F	GTGGTGccgcggCCAGAGTTCCTAGGG CAGAG	네스티드 PCR 제 2 라운드 프라이머
374	CR1-P1/P5-R	GTGGTGccgcggAGCTAGTCCCCACAT TCCAC	네스티드 PCR 제 2 라운드 프라이머
375	CR1-P1/P6-F	GTGGTGccgcggCCAGAGTTCCTAGGG CAGAG	네스티드 PCR 제 2 라운드 프라이머
376	CR1-P1/P6-R	GTGGTGccgcggGGTGGAGGGAAACT TTAGGC	네스티드 PCR 제 2 라운드 프라이머
377	CR1-P2/P5-F	GTGGTGccgcggCTCATTCATGCCT GGACA	네스티드 PCR 제 2 라운드 프라이머
378	CR1-P2/P5-R	GTGGTGccgcggAGCTAGTCCCCACAT TCCAC	네스티드 PCR 제 2 라운드 프라이머
379	CR1-P2/P6-F	GTGGTGccgcggTCTCATGCCTGGACA AGTAACT	네스티드 PCR 제 2 라운드 프라이머
380	CR1-P2/P6-R	GTGGTGccgcggGGTGGAGGGAAACT TTAGGC	네스티드 PCR 제 2 라운드 프라이머

[0320]



381	CR5-P3/P5-F	GTGGTGccgcggGGCTTGACAGAACT TACCG	네스티드 PCR 제 2 라운드 프라이머
382	CR5-P3/P5-R	GTGGTGccgcggCACCACTGTCTGCCT AAGGA	네스티드 PCR 제 2 라운드 프라이머
383	CR5-P4/P6-F	GTGGTGccgcggGGCTTGACAGAACT TACCG	네스티드 PCR 제 2 라운드 프라이머
384	CR5-P4/P6-R	GTGGTGccgcggGGTGGAGGAACT TTAGGC	네스티드 PCR 제 2 라운드 프라이머
385	CR5-P3/P5-F	GTGGTGccgcggCGTAGTGCCAAAACA AACAGT	네스티드 PCR 제 2 라운드 프라이머
386	CR5-P3/P5-R	GTGGTGccgcggCACCACTGTCTGCCT AAGGA	네스티드 PCR 제 2 라운드 프라이머
387	CR5-P4/P6-F	GTGGTGccgcggCGTAGTGCCAAAACA AACAGT	네스티드 PCR 제 2 라운드 프라이머
388	CR5-P4/P6-R	GTGGTGccgcggGGTGGAGGAACT TTAGGC	네스티드 PCR 제 2 라운드 프라이머
389	CR6-P1/P5-F	GTGGTGccgcggGCGAGGCCTACTTG ATATG	네스티드 PCR 제 2 라운드 프라이머
390	CR6-P1/P5-R	GTGGTGccgcggCTTCCAAGTGAGGC AATGC	네스티드 PCR 제 2 라운드 프라이머
391	CR6-P1/P6-F	GTGGTGccgcggACGTTTTGTGCTGCT GTAACA	네스티드 PCR 제 2 라운드 프라이머
392	CR6-P1/P6-R	GTGGTGccgcggCTGCAGGCACATTCT CTTCC	네스티드 PCR 제 2 라운드 프라이머
393	CR6-P2/P5-F	GTGGTGccgcggGCCCTGTGTGTAGT AGTCA	네스티드 PCR 제 2 라운드 프라이머
394	CR6-P2/P5-R	GTGGTGccgcggCTTCCAAGTGAGGC AATGC	네스티드 PCR 제 2 라운드 프라이머
395	CR6-P2/P6-F	GTGGTGccgcggCAGTATTAAGGGGTG GGAGCT	네스티드 PCR 제 2 라운드 프라이머
396	CR6-P2/P6-R	GTGGTGccgcggTCTCTTCTCACACA GCTGA	네스티드 PCR 제 2 라운드 프라이머
397	CR36-P3/P5-F	GTGGTGccgcggGGAGCTTGAGGGA AGAGAA	네스티드 PCR 제 2 라운드 프라이머
398	CR36-P3/P5-R	GTGGTGccgcggCTTCCAAGTGAGGC AATGC	네스티드 PCR 제 2 라운드 프라이머
399	CR36-P4/P6-F	GTGGTGccgcggATGGATGGGGAAGA CACTGG	네스티드 PCR 제 2 라운드 프라이머
400	CR36-P4/P6-R	GTGGTGccgcggCTGCAGGCACATTCT CTTCC	네스티드 PCR 제 2 라운드 프라이머
401	CR36-P3/P5-F	GTGGTGccgcggGGATGAAACAGGGC AGGAAC	네스티드 PCR 제 2 라운드 프라이머
402	CR36-P3/P5-R	GTGGTGccgcggTTCCAAGTGAGGCA ATGC	네스티드 PCR 제 2 라운드 프라이머
403	CR36-P4/P6-F	GTGGTGccgcggTTTGCAGAGCCATGA TGAGG	네스티드 PCR 제 2 라운드 프라이머
404	CR36-P4/P6-R	GTGGTGccgcggCGACAGCCAAAACA GCCG	네스티드 PCR 제 2 라운드 프라이머

[0321]

[0322]

웨스턴 블롯 분석. 디스트로핀 단백질 발현을 평가하기 위해, 불멸화 근모세포는 4-7일, 예컨대 6 또는 7일 동안 성장 배지를 1% 인슐린-트랜스페린-셀레늄 (인비트로젠) 및 1% 항생제/항진균제 (인비트로젠)로 보충된 DMEM으로 대체함으로써 근섬유로 분화시켰다. 섬유모세포는 MyoD 과다발현을 유도하고 15일 동안 1% 인슐린-트랜스페린-셀레늄 (인비트로젠), 1% 항생제/항진균제 (인비트로젠) 및 3  $\mu$ g/mL 독시시클린으로 보충된 DMEM에서 세포를 인큐베이션함으로써 근모세포로 전환분화시켰다. 디스트로핀 발현은 HEK293T 세포를 형질감염시킨지 3일 후에 평가하였다. 세포를 트립신처리하고, 수집하고, 프로테아제 억제제 콕테일 (시그마)로 보충된 RIPA 완충제 (시그마) 중에 용해시켰으며, 총 단백질 양을 제조업체의 명령 (피어스(Pierce))에 따라 비신코닌산 검정을 사용하여 정량화하였다. 이어서, 샘플을 NuPAGE 로딩 완충제 (인비트로젠) 및 5%  $\beta$ -메르캅토에탄올과 혼합하고, 10분 동안 85°C로 가열하였다. 25 마이크로그램 단백질을 MES 완충제 (인비트로젠)로 4-12% NuPAGE 비스-트리스 겔 (인비트로젠) 상에서 분리하였다. 단백질을 10-20% 메탄올, 예컨대 10% 메탄올 및 0.01% SDS를 함유하는 전달 완충제 중에서 1-2시간 동안 니트로셀룰로스 막으로 전달하였다. 이어서, 블롯을 실온에서 5% 밀크-TBST에 의해 1시간 동안 차단하였다. 블롯을 하기 1차 항체에 의해 프로빙하였다: 디스트로핀을 검출하기 위한 MANDYS8 (1:100, 시그마 D8168) 및 토끼 항-GAPDH (1:5000, 셀 시그널링 2118S). 이어서, 블롯을 마우스 또는 토끼 양고추냉이 퍼옥시다제-접합 2차 항체와 함께 인큐베이션하고, 케미독 화학발광 시스템 (바이오라드) 및 웨스턴-C ECL 기질 (바이오라드)을 사용하여 가시화하였다.

[0323]

면역결핍 마우스 내로의 이식. 모든 동물 실험은 듀크 동물 실험 윤리 위원회(Duke Institutional Animal Care

& Use Committee)에 의해 승인된 프로토콜 하에 수행하였다. 세포를 트립신처리하고, 수집하고, 1X 헹크 평형 염 용액 (HBSS, 시그마)으로 세척하였다. 2백만개의 세포를 펠릿화하고, 주사 직전에 심장독소 (시그마 #C9759)로 보충된 5  $\mu$ L 1X HBSS (시그마) 중에 재현탁시켰다. 이들 세포를 근육내 주사에 의해 NOD.SCID.감마 (NSG) 마우스 (듀크 CCIF 브리딩 코어(Duke CCIF Breeding Core))의 뒷다리 전경골근 (TA) 근육 내로 이식하였다. 주사 4주 후에, 마우스를 안락사시키고, TA 근육을 수거하였다.

[0324] 면역형광 염색. 수거된 TA 근육을 4℃에서 밤새 30% 글리세롤 중에서 인큐베이션한 후에 최적 커팅 온도 (Optimal Cutting Temperature) 화합물 중에 탑재하여 동결시켰다. 일련의 10 마이크로미터 절편을 -20℃에서 포매된 근육 조직의 동결절편화에 의해 획득하였다. 이어서, 동결절편을 PBS로 세척하여 OCT 화합물을 제거하고, 이후 스펙트린 검출을 위한 10% 열-불활성화 태아 소 혈청 또는 디스트로핀 검출을 위한 5% 열-불활성화 태아 소 혈청을 함유하는 PBS 중 실온에서 30-60분 동안 차단하였다. 동결절편을 단지 인간 에피토프에 특이적인 하기 1차 항체와 함께 4℃에서 밤새 인큐베이션하였다: 항-스펙트린 (1:20, 라이카(Leica) NCL-SPEC1) 또는 항-디스트로핀 (1:2, 라이카 NCL-DYS3). 1차 염색 후에, 스펙트린 또는 디스트로핀 발현은 티라미드 -기반 면역형광 신호 증폭 검출 키트 (라이프 테크놀로지스, TSA 키트 #22, 카탈로그 #T-20932)를 사용하여 검출하였다. 간략하게, 동결절편을 실온에서 1시간 동안 차단 완충제에서 2차적인 1:200 염소 항-마우스 비오틴-XX (라이프 테크놀로지스 #B2763)와 함께 인큐베이션하였다. 이어서, 신호를 실온에서 1시간 동안 차단 완충제 중 스트렙타 비딘-HRP 접합체 (1:100, TSA 키트로부터의 것)를 사용하여 증폭시켰다. 최종적으로, 동결절편을 실온에서 10분 동안 제조업체-제공 증폭 완충제에서 티라미드-알렉사플루오르488 접합체 (1:100, TSA 키트)와 함께 인큐베이션하였다. 이어서, 염색된 동결절편을 프로롱 안티페이드(ProLong AntiFade) (라이프 테크놀로지스 #P36934)에 탑재하고, 통상의 형광 현미경검사로 가시화하였다.

[0325] 세포독성 검정. 잠재적 sgRNA 또는 SpCas9 뉴클레아제-연관 세포독성을 정량적으로 평가하기 위해, HEK293T 세포를 제조업체의 지침서 (인비트로젠)에 따라 리포펙타민 2000을 사용하여 10 ng의 GFP 리포터 및 100 ng SpCas9 발현 벡터 및 100 ng sgRNA 발현 벡터로 형질감염시켰다. GFP 양성 세포의 백분율은 유동 세포측정법에 의해 2 및 5일에 평가하였다. 생존율은 제2일 내지 제5일에서의 GFP 양성 세포의 감소로서 계산하고, 문헌 [Cornu et al., Meth Mol Biol 649:237-245 (2010)]에 기재된 바와 같은 공 뉴클레아제 발현 벡터로 형질감염된 세포에 대해 정규화하였다.

[0326] 실시예 4

[0327] 디스트로핀 유전자를 표적화하는 CRISPR - 결과

[0328] CRISPR/Cas9-기반 시스템은 디스트로핀 유전자를 표적화하도록 설계하였다. NNNNN NNNNN NNNNN NNNNN NGG 및 GNNNN NNNNN NNNNN NNNNN NGG를 기반으로 하는, 인간 및 마우스 디스트로핀 유전자의 상이한 영역을 표적화하는 다양한 gRNA를 선택하였다 (표 6, 7 및 8 참조).

[0329] <표 6>

명칭 (서열 번호)	종	유전자	표적	가닥	키메라에 대한 19bp (5' 말단 상에 G 부가)	PAM
DCR1 (65)	인간	DMD	인트론 50	+	attgcttggattcccta	GGG
DCR2 (66)	인간	DMD	인트론 50	-	tgtagagtaagtcagccta	TGG
DCR3 (67)	인간	DMD	엑손 51-55'	+	cctactcagactgttactc	TGG
DCR4 (68)	인간	DMD	엑손 51-53'	+	ttggacagaactaccgac	TGG
DCR5 (69)	인간	DMD	인트론 51	-	cagttgcctaagaactggt	GGG
DCR6 (70)	인간	DMD	인트론 44	-	GGGCTCCACCCACAGAGT	GGG

[0330]

명칭 (서열 번호)	종	유전자	표적	가닥	키메라에 대한 19bp (5' 말단 상에 G 부가)	PAM
DCR7 (71)	인간	DMD	인트론 55	+	TTTGCTTCGCTATAAAACG	AGG
DCR8 (72)	인간	DMD	엑손 41	+	TCTGAGGATGGGGCCGCAA	TGG
DCR9 (73)	인간	DMD	엑손 44	-	GATCTGTCAAATCGCCTGC	AGG
DCR10 (74)	인간	DMD	엑손 45	+	CCAGGATGGCATTGGGCAG	CGG
DCR11 (75)	인간	DMD	엑손 45	+	CTGAATCTGCGGTGGCAGG	AGG
DCR12 (76)	인간	DMD	엑손 46	-	TTCTTTTGTTCCTCTAGC <sub>e</sub>	TGG
DCR13 (77)	인간	DMD	엑손 46	+	GAAAAGCTTGAGCAAGTCA	AGG
DCR14 (78)	인간	DMD	엑손 47	+	GAAGAGTTGCCCTGCGCC	AGG
DCR15 (79)	인간	DMD	엑손 47	+	ACAAATCTCCAGTGGATAA	AGG
DCR16 (80)	인간	DMD	엑손 48	-	TGTTTCTCAGGTAAAGCTC	TGG
DCR17 (81)	인간	DMD	엑손 48	+	GAAGGACCATTTGACGTT <sub>a</sub>	AGG
DCR18 (82)	인간	DMD	엑손 49	-	AACTGCTATTTCAGTTTC <sub>c</sub>	TGG
DCR19 (83)	인간	DMD	엑손 49	+	CCAGCCACTCAGCCAGTGA	AGG
DCR20 (84)	인간	DMD	엑손 50	+	gtatgctttctgttaaag	AGG
DCR21 (85)	인간	DMD	엑손 50	+	CTCCTGGACTGACCACTAT	TGG
DCR22 (86)	인간	DMD	엑손 52	+	GAACAGAGGCGTCCCAGT	TGG
DCR23 (87)	인간	DMD	엑손 52	+	GAGGCTAGAACAATCATTA	CGG
DCR24 (88)	인간	DMD	엑손 53	+	ACAAGAACACCTTCAGAAC	CGG
DCR25 (89)	인간	DMD	엑손 53	-	GGTTTCTGTGATTTTCTTT	TGG
DCR26 (90)	인간	DMD	엑손 54	+	GGCCAAAGACCTCCGCCAG	TGG
DCR27 (91)	인간	DMD	엑손 54	+	TTGGAGAAGCATTCATAAA	AGG
DCR28 (92)	인간	DMD	엑손 55	-	TCGCTCACTCACCctgcaa	AGG
DCR29 (93)	인간	DMD	엑손 55	+	AAAAGAGCTGATGAAACAA	TGG
DCR30 (94)	인간	DMD	5'UTR/엑손 1	+	TAcACTTTTCaAAATGCTT	TGG
DCR31 (95)	인간	DMD	엑손 51	+	gagatgatcatcaagcaga	AGG
DCR32 (96)	마우스	DMD	mdx 돌연변이	+	cttgaaagagcaaTaaaa	TGG
DCR33 (97)	인간	DMD	인트론 44	-	CACAAAAGTCAAATCGGAA	TGG
DCR34 (98)	인간	DMD	인트론 44	-	ATTTCATATAAGATTCGG	AGG

[0331]

명칭 (서열 번호)	종	유전자	표적	가닥	키메라에 대한 19bp (5' 말단 상에 G 부가)	PAM
DCR35 (99)	인간	DMD	인트론 55	-	CTTAAGCAATCCCGAACTC	TGG
DCR36 (100)	인간	DMD	인트론 55	-	CCTTCTTTATCCCCTATCG	AGG
DCR40 (104)	마우스	DMD	엑손 23	-	aggccaaacctcggttac	NNGRR
DCR41 (105)	마우스	DMD	엑손 23	+	TTCGAAAATTTTCAGgtaag	NNGRR
DCR42 (106)	마우스	DMD	엑손 23	+	gcagaacaggagataacag	NNGRRT
DCR43 (107)	마우스	ACVR2B	엑손 1	+	gcggccctcgcccttctct	ggggat
DCR48 (108)	인간	DMD	인트론 45	-	TAGTGATCGTGGATACGAG	AGG
DCR49 (109)	인간	DMD	인트론 45	-	TACAGCCCTCGTGTATAT	TGG
DCR50 (110)	인간	DMD	인트론 52	-	GGAAGGAATTAAGCCCGAA	TGG
DCR51 (111)	인간	DMD	인트론 53	-	GGAACAGCTTTCGTAGTTG	AGG
DCR52 (112)	인간	DMD	인트론 54	+	ATAAAGTCCAGTGTGATC	AGG
DCR53 (113)			인트론 54	+	AAAACCAGAGCTTCGGTCA	AGG
DCR54 (114)	마우스	Rosa26	ZFN 영역	+	GAGTCTTCTGGGCAGGCTTAA AGGCTAACC	TGG
DCR55 (115)	마우스	Rosa26	mRNA	-	TCGGGTGAGCATGTCTTTAAT CTACCTCGA	TGG
DCR49 (116)	인간	DMD	엑손 51	-	gtgtcaccagagtaacagt	ctgagt
DCR50 (117)	인간	DMD	엑손 51	+	tgatcatcaagcagaaggt	atgag
DCR60 (118)	마우스	DMD	엑손 23	+	AACTTCGAAAAATTCAGgta	agccgagg
DCR61 (119)	마우스	DMD	인트론 22	+	gaaactcatcaaatatgcgt	gttagtgt
DCR62 (120)	마우스	DMD	인트론 22	-	tcatttacactaacacgcat	atttgatg
DCR63 (121)	마우스	DMD	인트론 22	+	gaatgaaactcatcaaatat	gcgtgtta
DCR64 (122)	마우스	DMD	인트론 23	-	tcacaaatatttgaagga	ctctgggt
DCR65 (123)	마우스	DMD	인트론 23	-	tgtttcatagaaaaatag	gcaagtgt
DCR66 (124)	마우스	DMD	인트론 23	+	aattggaaaatgtgatggga	aacagata
DCR67 (125)	인간	DMD	엑손 51	+	atgatcatcaagcagaaggt	atgagaaa
DCR68 (126)	인간	DMD	엑손 51	+	agatgatcatcaagcagaag	gtatgaga
DCR69 (127)	인간	DMD	엑손 51	-	cattttttctacacttct	gcttgatg
DCR70 (128)	인간	DMD	엑손 51	+	tcctactcagactgttactc	tggtgaca
DCR71 (129)	인간	DMD	엑손 51	-	acaggttgtgtcaccagagt	aacagtct

[0332]

명칭 (서열 번호)	종	유전자	표적	가닥	키메라에 대한 19bp (5' 말단 상에 G 부가)	PAM
DCR72 (130)	인간	DMD	엑손 51	-	ttatcattttttctcatacc	ttctgctt
DCR73 (131)	인간	DMD	인트론 51	-	ttgcctaagaactggtggga	aatggtct
DCR74 (132)	인간	DMD	인트론 51	-	aaacagttgcctaagaactg	gtgggaaa
DCR75 (133)	인간	DMD	인트론 51	+	ttcccaccagtcttaggc	aactgttt
DCR76 (134)	인간	DMD	인트론 50	+	tggtttgatttcctaggg	tccagctt
DCR77 (135)	인간	DMD	인트론 50	-	tagggaaatcaaagccaatg	aaacgttc
DCR78 (136)	인간	DMD	인트론 50	-	gacctagggaatcaaagc	caatgaaa
DCR79 (137)	인간	DMD	인트론 44	-	TGAGGGCTCCACCCTCACGA	GTGGGTTT
DCR80 (138)	인간	DMD	인트론 44	-	AAGGATTGAGGGCTCCACCC	TCACGAGT
DCR81 (139)	인간	DMD	인트론 44	-	GCTCCACCCTCACGAGTGGG	TTTGGTTC
DCR82 (140)	인간	DMD	인트론 55	-	TATCCCCTATCGAGGAAACC	ACGAGTTT
DCR83 (141)	인간	DMD	인트론 55	+	GATAAAGAAGGCCTATTTC	TAGAGTTG
DCR84 (142)	인간	DMD	인트론 55	-	AGGCCTTCTTTATCCCCTAT	CGAGGAAA
DCR85 (143)	인간	DMD	인트론 44	-	TGAGGGCTCCACCCTCACGA	GTGGGT
DCR86 (144)	인간	DMD	인트론 55	+	GATAAAGAAGGCCTATTTC	TAGAGT

[0333]

[0334]

<표 7>

명칭	설명	% 변형
DCR1	엑손 51 을 결실시킴	6.6
DCR2	엑손 51 을 결실시킴	10.3
DCR3	프레임시프트	13
DCR4	엑손 51 을 결실시킴	11.9
DCR5	엑손 51 을 결실시킴	12.4
DCR6	인트론 44 에서 가능한 한 엑손 44 에 근접하게 (환자 결실의 경우에서)	16.1
DCR7	인트론 55 에서 가능한 한 엑손 56 에 근접하게 (환자 결실의 경우에서)	6.8
DCR8	단지 엑손 42-43 결실 (-1/+2)을 교정할 수 있음, (-2/+1)은 이것에 의해 교정가능하지 않음	17.3
DCR9	엑손 44 를 스킵핑함 (5')	14.4
DCR10	프레임시프트	14.9
DCR11	엑손 45 의 하류를 교정함	<1
DCR12	5' 스플라이스 수용자/프레임시프트	<1
DCR13	엑손 46 의 하류를 교정함	16.9
DCR14	프레임시프트	17.2

[0335]



명칭	설명	% 변형
DCR15	엑손 47 의 하류를 교정함	15.4
DCR16	프레임시프트	11.5
DCR17	엑손 48 의 하류를 교정함	<1
DCR18	5' 스플라이스 수용자/프레임시프트	1.8
DCR19	엑손 49 의 하류를 교정함	33.7
DCR20	5' 스플라이스 수용자	14.9
DCR21	엑손 50 의 하류를 교정함	24.1
DCR22	프레임시프트	25.9
DCR23	엑손 52 의 하류를 교정함	25.2
DCR24	프레임시프트 (단지 +1 프레임을 교정할 수 있음)	24.8
DCR25	엑손 53 의 하류를 교정함	2.6
DCR26	프레임시프트	24.5
DCR27	엑손 54 의 하류를 교정함	13.4
DCR28	5' 스플라이스 수용자	21.6
DCR29	엑손 55 의 하류를 교정함	19.2
DCR30	엑손 1 에 미니디스트로폰을 통합시킴	시험되지 않음
DCR31	엑손 51 의 하류를 교정함	18.9
DCR32	정지 코돈을 결실시킴	시험되지 않음
DCR33	CR6 에 대한 대체	1.3
DCR34	CR6 에 대한 대체	13.2
DCR35	CR7 에 대한 대체	22.5
DCR36	CR7 에 대한 대체	26.4
DCR40	엑손 23 5' 스플라이스 공여자를 파괴함 (mdx 돌연변이를 교정함)	
DCR41	엑손 23 5' 스플라이스 공여자를 파괴함 (mdx 돌연변이를 교정함)	
DCR42	엑손 53 mdx4cv 돌연변이를 결실시킴	
DCR43	미오스타틴 수용체를 파괴함	

[0336]

[0337] <표 8>

명칭	Cas9		설명	사용된 Cas9
DCR49	에스. 아우레우스		엑손 51 에서의 프레임시프트	SaCas9 (장(Zhang) pX441 로부터) (NNGRRT PAM)
DCR50	에스. 아우레우스		엑손 51 의 5' 말단을 과괴함	SaCas9 (장 pX441 로부터) (NNGRR PAM)
DCR60	엔. 메닌기티디스	NNNNGANN	엑손 23 의 3' 스플라이스 공여자를 우회 mdx 돌연변이에 표적화함	NmCas9 (NNNNGANN PAM)
DCR61	엔. 메닌기티디스	NNNNGNNT	엑손 23 및 우회 mdx 돌연변이를 결실시킴	NmCas9 (NNNNGNNT PAM)
DCR62	엔. 메닌기티디스	NNNNGANN	엑손 23 및 우회 mdx 돌연변이를 결실시킴	NmCas9 (NNNNGANN PAM)
DCR63	엔. 메닌기티디스	NNNNGTTN	엑손 23 및 우회 mdx 돌연변이를 결실시킴	NmCas9 (NNNNGTTN PAM)
DCR64	엔. 메닌기티디스	NNNNGNNT	엑손 23 및 우회 mdx 돌연변이를 결실시킴	NmCas9 (NNNNGNNT PAM)
DCR65	엔. 메닌기티디스	NNNNGTTN	엑손 23 및 우회 mdx 돌연변이를 결실시킴	NmCas9 (NNNNGTTN PAM)
DCR66	엔. 메닌기티디스	NNNNGANN	엑손 23 및 우회 mdx 돌연변이를 결실시킴	NmCas9 (NNNNGANN PAM)
DCR67	엔. 메닌기티디스	NNNNGANN	엑손 51 의 3' 스플라이스 공여자를 스킵 엑손에 표적화함	NmCas9 (NNNNGANN PAM)
DCR68	엔. 메닌기티디스	NNNNGANN	엑손 51 의 3' 스플라이스 공여자를 스킵 엑손에 표적화함	NmCas9 (NNNNGANN PAM)
DCR69	엔. 메닌기티디스	NNNNGANN	엑손 51 의 3' 스플라이스 공여자를 스킵 엑손에 표적화함	NmCas9 (NNNNGANN PAM)
DCR70	엔. 메닌기티디스	NNNNGANN	엑손 51 에서의 프레임시프트	NmCas9 (NNNNGANN PAM)

[0338]

DCR71	엔. 메닌기타디스	NNNNGNNT	엑손 51에서의 프레임시프트	NmCas9 (NNNNGNNT PAM)
DCR72	엔. 메닌기타디스	NNNNGNNT	엑손 51의 3' 스플라이스 공여자를 스킵 엑손에 표적화함	NmCas9 (NNNNGNNT PAM)
DCR73	엔. 메닌기타디스	NNNNGNNT	엑손 51을 결실시킴 (가능한 한 DCR5에 근접하게 결함함)	NmCas9 (NNNNGNNT PAM)
DCR74	엔. 메닌기타디스	NNNNGANN	엑손 51을 결실시킴 (가능한 한 DCR5에 근접하게 결함함)	NmCas9 (NNNNGANN PAM)
DCR75	엔. 메닌기타디스	NNNNGTTN	엑손 51을 결실시킴 (가능한 한 DCR5에 근접하게 결함함)	NmCas9 (NNNNGTTN PAM)
DCR76	엔. 메닌기타디스	NNNNGNNT	엑손 51을 결실시킴 (가능한 한 DCR1/2에 근접하게 결함함)	NmCas9 (NNNNGNNT PAM)
DCR77	엔. 메닌기타디스	NNNNGTTN	엑손 51을 결실시킴 (가능한 한 DCR1/2에 근접하게 결함함)	NmCas9 (NNNNGTTN PAM)
DCR78	엔. 메닌기타디스	NNNNGANN	엑손 51을 결실시킴 (가능한 한 DCR1/2에 근접하게 결함함)	NmCas9 (NNNNGANN PAM)
DCR79	엔. 메닌기타디스	NNNNGNNT	엑손 45-55를 결실시킴 - NNNNGTTN PAM과 중복됨, 가능한 한 DCR6에 근접하게 결함함	NmCas9 (NNNNGNNT PAM)
DCR80	엔. 메닌기타디스	NNNNGANN	엑손 45-55를 결실시킴, 가능한 한 DCR6에 근접하게 결함함	NmCas9 (NNNNGANN PAM)
DCR81	엔. 메닌기타디스	NNNNGTTN	엑손 45-55를 결실시킴, 가능한 한 DCR6에 근접하게 결함함	NmCas9 (NNNNGTTN PAM)

[0339]

DCR82	엔. 메닌기타디스	NNNNGNNT	엑손 45-55를 결실시킴, 가능한 한 DCR36에 근접하게 결함함 - NNNNGTTN PAM과 중복됨	NmCas9 (NNNNGNNT PAM)
DCR83	엔. 메닌기타디스	NNNNGTTN	엑손 45-55를 결실시킴, 가능한 한 DCR36에 근접하게 결함함	NmCas9 (NNNNGTTN PAM)
DCR84	엔. 메닌기타디스	NNNNGANN	엑손 45-55를 결실시킴, 가능한 한 DCR36에 근접하게 결함함	NmCas9 (NNNNGANN PAM)
DCR85	에스. 아우레우스	NNGRRT	엑손 45-55를 결실시킴, 가능한 한 DCR6에 근접하게 결함함	SaCas9 (장 pX441로부터) (NNGRRT PAM)
DCR86	에스. 아우레우스	NNGRRT	엑손 45-55를 결실시킴, 가능한 한 DCR36에 근접하게 결함함	SaCas9 (장 pX441로부터) (NNGRRT PAM)

[0340]

[0341] 특히, Cas9 400 ng은 엑손 51을 포괄하는 영역, 즉 CR1, CR2, CR3, CR4 및 CR5를 표적화하는 공백터 또는 gRNA

400 ng과 함께 HEK 293T 세포 내로 공동-형질감염시켰다 (도 11(b) 참조). 게놈 DNA는 형질감염 2일 후에 수거하고, 서베이어 검정을 사용하여 분석하였다 (도 11(a) 및 11(c) 참조).

[0342] CRISPR/Cas9-기반 시스템을 시스템이 돌연변이체 디스트로핀 유전자를 복구할 수 있는지의 여부를 결정하기 위해 DMD 8036 (del148-50) 세포에서 사용하였다. 5  $\mu$ g의 Cas9를 7.5  $\mu$ g의 공벡터 또는 gRNA와 함께 DMD 8036 (del148-50) 세포 내로 공동-형질감염시켰다. 특히, 7.5  $\mu$ g의 CR1 ("DCR1"), 7.5  $\mu$ g의 CR5 ("DCR5"), 15  $\mu$ g의 CR3 ("DCR3") 또는 7.5  $\mu$ g의 CR1과 CR5의 조합 (DCR1+DCR5)을 사용하였다. 게놈 DNA를 형질감염 3일 후에 수거하고, 서베이어 검정 (도 12) 또는 전체 유전자좌에 걸친 PCR 분석 (도 13)을 사용하여 분석하였다. 이러한 유전자좌를 CR1 및 CR5에 대한 게놈 표적을 함유하는 영역에 플랭킹하는 프라이머를 사용하여 PCR에 의해 증폭시켜 (정방향 프라이머: 5'-gagaggttatgtggctttacca (서열 457), 역방향 프라이머: 5'-ctgcgtagtgccaaaacaaa (서열 458)), 야생형 유전자좌에 대해 1447 bp 밴드 또는 결실된 유전자좌에 대해 대략 630 bp의 예상된 크기를 생성하였다. 분화 7일 후에, 처리된 세포의 웨스턴 블롯은 디스트로핀 단백질의 발현을 제시하였다 (도 14 참조).

[0343] 실시예 5

[0344] CRISPR/Cas9의 인간 디스트로핀 유전자 내 핫스팟 돌연변이에의 표적화

[0345] 폭넓은 범위의 디스트로핀 돌연변이를 교정하기 위한 CRISPR/Cas9 유전자 편집 플랫폼을 이용하기 위해, 엑손 45-55 사이의 핫스팟 돌연변이 영역을 표적화하는 다수의 sgRNA를 생성하였다 (도 16). 효율적 부위-특이적 유전자 편집을 가이드하기 위해 인간-코돈 최적화된 SpCas9 뉴클레아제 및 키메라 단일-가이드 RNA (sgRNA) 발현 벡터를 이용하는 에스. 피오게네스 시스템을 사용하였다. TALEN을 사용하여 엑손 51을 표적화하는 실시예 4와 유사하게, SpCas9의 5'-NRG-3' PAM 요건을 충족하는 엑손 45 내지 55의 5' 및 3' 말단을 표적화하는 프로토스페이서를 선택하였다. 이들 엑손 내에서 NHEJ-기반 DNA 복구에 의해 생성된 작은 삽입 또는 결실은 각 엑손을 둘러싸는 다양한 디스트로핀 돌연변이를 다루는 표적화된 프레임시프트 돌연변이를 생성할 수 있다 (도 16A-16B). 예를 들어, CR3은 작은 삽입 또는 결실을 엑손 51의 5' 말단에 도입하여 하류 디스트로핀 리딩 프레임을 복원함으로써 엑손 51을 둘러싸는 디스트로핀 돌연변이 또는 결실을 교정하도록 설계하였다 (도 16B). 추가로, sgRNA는 CRISPR/Cas9 시스템의 멀티플렉스 능력을 사용하고, 올리고뉴클레오타이드-기반 엑손 스킵핑의 방법과 유사하게 개별 엑손 또는 일련의 엑손을 특이적으로 결실시켜 디스트로핀 리딩 프레임을 복원하도록 설계하였다. 이러한 목적을 위해, sgRNA는 엑손 51 (도 16C) 또는 엑손 45-55 (도 16D)를 둘러싸는 인트론 영역을 표적화하였다. 이들 sgRNA는 생성된 전사체에 포함되도록 의도된 하류 또는 상류 엑손에 가장 가까운 부위를 의도적으로 표적화하여, 배경 환자 결실이 인트론 sgRNA 표적 부위를 포함할 가능성을 최소화하였다.

[0346] 실시예 6

[0347] 인간 세포에서의 디스트로핀 유전자를 표적화하는 sgRNA의 스크리닝

[0348] 인간 HEK293T 세포주에서의 유전자 편집 빈도를 평가하여 상이한 sgRNA 표적화 효율을 신속하게 결정하였다. HEK293T를 인간 코돈-최적화된 SpCas9 및 나타낸 sgRNA를 코딩하는 구축물로 형질감염시켰다. 각 sgRNA는 나타낸 바와 같은 디스트로핀 유전자를 변형시키도록 설계하였다. 형질감염 후 제3일 또는 제10일에서의 유전자 변형의 빈도를 서베이어 검정에 의해 결정하였다. 제3일 및 제10일에 측정된 서베이어 신호의 비를 계산하여 인간 세포에서의 각 sgRNA에 대한 유전자 편집 빈도의 안정성을 정량화하였다. 형질감염 3일 후에 서베이어 검정에 의해 정량화된 바와 같이, 시험된 sgRNA의 29/32 (~90%)는 의도된 유전자좌에서 고도로 효율적인 유전자 변형을 매개할 수 있었다 (표 9, 도 17). 유전자 편집 빈도는 거의 모든 sgRNA에 대해 안정하였으며 (제3일 내지 제10일에 <25% 신호 변화, 표 9, 도 18), 이는 각 개별 sgRNA에 의해 매개된 유전자 편집이 잘-용인되었음을 나타낸다. 주목할 만한 예외는 CR33이며, 이는 활성이 서베이어 검정의 감수성 미만일 수 있지만 제10일에 검출 가능한 활성을 갖지 않았다 (추정치 ~1%).

[0349] <표 9>

[0350] 인간 세포에서의 sgRNA의 측정된 활성

표적	sgRNA #	% 제 3 일에서의 변형된 대립유전자	% 제 10 일에서의 변형된 대립유전자	% 제 10 일/제 3 일 변화
엑손 51의 멀티플렉스 결실				
인트론 50	CR1	6.6	9.3	41.8
인트론 50	CR2	10.3	14.0	36.2
엑손 51	CR4	11.9	14.4	21.3
인트론 51	CR5	12.4	13.3	7.8
엑손 45-55의 멀티플렉스 결실				
인트론 44	CR6	16.1	16.9	4.3
인트론 44	CR33	1.3	<1	n.d.
인트론 44	CR34	13.2	11.0	-16.6
인트론 55	CR7	6.8	7.1	5.3
인트론 55	CR35	22.5	20.9	-7.1
인트론 55	CR36	26.4	24.7	-6.4
표적화된 프레임시프트				
엑손 45	CR10	14.9	16.3	9.3
엑손 45	CR11	<1	<1	n.d.
엑손 46	CR12	<1	<1	n.d.
엑손 46	CR13	16.9	18.4	9.2
엑손 47	CR14	17.2	17.6	2.9
엑손 47	CR15	15.4	15.3	-0.9
엑손 48	CR16	11.5	10.9	-5.0
엑손 48	CR17	<1	<1	n.d.
엑손 49	CR18	1.8	2.2	20.1
엑손 49	CR19	33.7	38.4	13.9
엑손 50	CR20	14.9	13.7	-7.6
엑손 50	CR21	24.1	20.8	-13.5
엑손 51	CR3	13.0	16.7	28.0
엑손 51	CR31	18.9	16.9	-10.2
엑손 52	CR22	25.9	20.3	-21.6
엑손 52	CR23	25.2	24.0	-4.8
엑손 53	CR24	24.8	23.6	-4.6
엑손 53	CR25	2.6	2.9	9.5
엑손 54	CR26	24.5	22.0	-10.1
엑손 54	CR27	13.4	12.6	-5.9
엑손 55	CR28	21.6	19.8	-8.4
엑손 55	CR29	19.2	19.6	2.2

[0351]

[0352] 실시예 7

[0353] 형광-기반 리포터 시스템을 사용하는 유전자-편집 세포의 보강

[0354] sgRNA를 DMD 환자 근모세포주에서 특정한 돌연변이를 교정하기 위해 선택하였다. DMD 근모세포 내로의 형질감염 후에, 예상외로 낮은 또는 검출불가능한 유전자 변형 활성이 서베이어 검정에 의해 측정된 바와 같이 관찰되었다 (도 19C, 벌크 집단). 유동 세포측정법을 사용하여, SpCas9 단백질에 연결된 2A 리보솜 스키피нг 펩티드를 통해 GFP를 공-발현하는 형질감염된 세포를 선택하였다 (도 19A). SpCas9 발현 벡터에 이러한 형광 리포터의 부가는 HEK293T 세포에서 유전자 편집 활성에 유의한 영향을 미치는 것처럼 보이지 않았다 (도 19B). 대조군 GFP 발현 플라스미드의 높은 형질감염 효율 (전형적으로 >70%, 도 19D, pmaxGFP)에도 불구하고, 형질감염된 근모세포의 낮은 백분율 (~0.5-2%)이 전기천공 3일 후에 형광 리포터를 발현하였다. 용이하게 형질감염된 HEK293T 세포주에서 CRISPR/Cas9 활성의 높은 수준을 고려할 때, SpCas9-T2A-GFP 및 sgRNA 구축물의 DMD 세포 내로의 전기천공 후 비효율적인 트랜스진 발현은 비분류 세포에서 관찰된 낮은 유전자 편집 효율을 설명할 수 있다. GFP-양성 DMD 근모세포를 분류한 후, 실질적인 증가가 대부분의 sgRNA 표적 유전자좌에서 검출가능한 활성으로 관찰되었다 (도 19C). 따라서, 모든 후속 실험은 이러한 형광 리포터의 발현에 의한 SpCas9 발현에 대해 분류된 세포를 사용하였다.

[0355] 실시예 8



- [0356] 표적화된 프레임시프트에 의한 디스트로핀 발현의 복원
- [0357] NHEJ DNA 복구에 의해 생성된 작은 삽입 및 결실을 사용하여 표적화된 프레임시프트를 생성함으로써 이상 리딩 프레임을 교정할 수 있다. sgRNA인 CR3은 엑손 51 내에 작은 삽입 및 결실을 도입함으로써 디스트로핀 리딩 프레임을 복원하도록 설계하였다 (도 16B, 20A). 이러한 유전자좌에서 CRISPR/Cas9에 의해 생성된 삽입 및 결실의 유형은 SpCas9 및 CR3 sgRNA에 대한 발현 플라스미드로 공동-형질감염시킨 HEK293T 세포의 게놈 DNA로부터의 대립유전자를 생어 서열분석하여 평가하였다 (도 20B). 주목할 만하게, 삽입 및 결실은 모든 3개의 리딩 프레임으로의 전환을 유발하였다 (도 20B, 20C). 관련 환자 세포주에서 유전자 교정을 입증하기 위해, SpCas9 및 CR3 sgRNA에 대한 발현 플라스미드를, 엑손 51에서 프레임시프트를 생성함으로써 교정가능한 엑손 48-50의 결실을 갖는 DMD 근모세포주 내로 전기천공하였다. 처리된 세포를 분류하고, 서베이어 검정에 의해 유전자 변형 활성을 가지고 있다는 것을 검증하고 (CR3, 도 19C 분류된 집단), 복원된 디스트로핀 발현을 시험하기 위해 근관으로 분화시켰다. 디스트로핀 단백질의 발현은 검출가능한 뉴클레아제 활성을 동반하는 것으로 관찰되었다 (도 20D). 에스. 피오케네스 CRISPR/Cas9 시스템은 다양한 환자 돌연변이를 다루고 인간 디스트로핀 유전자의 발현을 복원하기 위해 표적화된 프레임시프트를 신속하게 생성하는 강력한 방법을 제시한다.
- [0358] 실시예 9
- [0359] 멀티플렉스 CRISPR/Cas9 유전자 편집은 엑손 51의 게놈 결실을 매개하고 디스트로핀 단백질 발현을 구제함
- [0360] CRISPR/Cas9 시스템의 멀티플렉스화 능력은 표적화된 유전자 교정을 위한 특정한 엑손의 게놈 결실을 효율적으로 생성하는 신규 방법을 제시한다. 엑손 51 스킵핑에 의해 교정가능한 배경 결실을 갖는 DMD 환자 근모세포를 엑손 51에 플랭킹하는 sgRNA의 2개 조합 (CR1/CR5 또는 CR2/CR5)으로 처리하고, 도 19에서와 같이 유전자-편집 세포를 보강하도록 분류하였다. 이들 처리된 세포로부터의 게놈 DNA의 엔드-포인트 PCR에 의해 검출된 바와 같이, 예상된 게놈 결실은 단지 sgRNA 둘 다가 SpCas9와 함께 세포 내로 전기천공된 경우에만 제시되었다 (도 21A). 생어 서열분석은 둘 다의 결실에 대해 원위 염색체 절편의 예상된 접합부를 확인하였다 (도 21B). 분류된 근모세포를 구별한 후에, mRNA 전사체로부터의 엑손 51의 결실은 단지 둘 다의 sgRNA로 처리된 세포에서만 검출되었다 (도 21C). 최종적으로, 엑손 51의 게놈- 및 mRNA-수준 결실이 관찰된 것에 동반하여, 복원된 디스트로핀 단백질 발현이 처리된 세포에서 검출되었다 (도 21D).
- [0361] 실시예 10
- [0362] 다중-엑손 거대 게놈 결실에 의한 디스트로핀 구제
- [0363] 환자-특이적 돌연변이를 다루는 것이 CRISPR/Cas9 시스템을 강력하게 사용하는 것이지만, 무수한 통상의 환자 결실을 다룰 수 있는 단일 방법을 개발하는 것이 유리할 것이다. 예를 들어, 유망한 전략은 공지된 환자 결실의 최대 62%를 교정하는 방법으로서 전체 엑손 45-55 영역을 배제시키는 것이다. 멀티플렉스 CRISPR/Cas9-기반 유전자 편집을 그것이 인간 세포에서 엑손 45-55 유전자좌의 효율적 결실을 생성할 수 있는지의 여부를 결정하기 위해 시험하였다. HEK293T 세포 내로의 형질감염 후에, ~336,000 bp의 예상된 결실이 게놈 DNA의 PCR에 의해 검출되었다 (도 22A). 유사하게, 이러한 결실은 미지 길이의 엑손 48-50의 배경 결실을 보유하는 SpCas9/sgRNA-처리 DMD 환자 세포로부터의 게놈 DNA의 PCR에 의해 검출되었다 (도 22A). 처리된 DMD 세포의 게놈 DNA로부터의 이러한 결실 밴드의 생어 서열분석은 sgRNA 표적 부위에 바로 인접한 인트론 44 및 인트론 55의 예상된 접합부를 밝혀냈다 (도 22B). 처리된 DMD 세포의 분화 후에, 엑손 45-55의 예상된 결실이 디스트로핀 mRNA 전사체에서 검출되었고, 생어 서열분석에 의해 엑손 44 및 56의 융합이 있었음을 검증하였다 (도 22C). 복원된 단백질 발현은 게놈 및 생성된 mRNA 전사체로부터 엑손 45-55의 CRISPR/Cas9-유도 결실을 함유하는 분류된 세포 집단에서 웨스턴 블롯에 의해 관찰되었다 (도 22D). 이들 데이터는 멀티플렉스 CRISPR/Cas9 편집이 DMD 환자 돌연변이의 60% 초과에서 디스트로핀 리딩 프레임을 복원하는 단일의 보편적 방법을 제시한다는 것을 입증한다.
- [0364] 실시예 11
- [0365] 면역결핍 마우스 내로의 교정된 근모세포의 이식
- [0366] DMD 요법을 위한 유망한 방법은 디스트로핀 발현을 구제하기 위해 환자의 골격근 조직 내로 생착될 수 있는 자가 환자 근육 전구 세포의 집단을 교정하는 것이다. 생체내에서 인간 디스트로핀을 발현하는 교정된 세포의 능력을 입증하기 위해, 엑손 51에 플랭킹하는 sgRNA CR1 및 CR5로 처리된 DMD 근모세포 집단을 이식하고, 상기와 같이 GFP의 발현에 대해 분류하였다 (도 19, 도 23). 4주 후에, 교정 세포 및 비교정 세포 둘 다에 의해 발현

되는 인간 스펙트린에 대해 양성인 근섬유가 주사된 근육 조직의 동결절편에서 검출되었다 (도 24). 수많은 이들 섬유는 또한, 근초에 국재화된 발현을 갖는 인간 디스트로핀에 대해 양성이며, 이는 이들 세포에서의 기능적 유전자 교정을 입증한다 (도 24, 도 25). 인간 디스트로핀에 대해 양성인 섬유는 미처리 DMD 근모세포가 주사된 마우스로부터의 절편에서 관찰되지 않았으며 (도 24, 도 25), 이는 CRISPR/Cas9-변형 세포가 인간 디스트로핀 발현의 공급원이라는 것을 나타낸다.

[0367] 실시예 12

[0368] 오프-타겟 및 세포독성 분석

[0369] CRISPR/Cas9 시스템의 상대 세포독성은 인간 세포에서 선택 sgRNA에 대해, 이전에 기재된 바와 같은 유동 세포 측정법-기반 GFP 체류 검정을 적합화시킴으로써 평가하였다 (Ousterout et al., Mol Ther 21:1718-1726 (2013)). 최소 세포독성은 인간 세포 내로의 형질감염 후에 sgRNA와 함께 또는 그 없이 공-발현된 SpCas9에 대해 관찰되었다 (도 26A). 공중이 이용가능한 도구는 sgRNA 프로토스페이서 서열에서의 소정의 미스매치의 예측된 위치 편향 및 의도된 표적 부위에 대한 미스매치의 총 개수를 기준으로 하여 오프-타겟 유전자좌에서의 잠재적 CRISPR/Cas9 활성을 평가하고 우선순위를 정하는데 이용가능하다 (Hsu et al., Nat Biotechnol 31:827-826 (2013)). 이러한 공중 웹서버는 이러한 연구에서 디스트로핀 유전자를 교정하는데 사용된 sgRNA에 대해 가장 가능성있는 오프-타겟 부위를 예상하는데 사용되었다 (표 4). 상위 10개의 잠재적 오프-타겟 부위는 SpCas9 및 CR1, CR3, CR5, CR6 또는 CR36에 대한 개별 sgRNA 발현 카세트에 처리된 HEK293T 세포에서 서베이어 검정에 의해 평가하였다. CR1, CR3 및 CR36은 각각 이들 10개의 예측된 오프-타겟 유전자좌 중 하나를 가졌으며 유의한 수준의 유전자 변형을 입증하였다 (표 4 및 도 27). 흥미롭게도, CR3 오프-타겟 서열은 의도된 온-타겟에 대한 상당한 상동성 및 유사한 변형 빈도를 가졌다 (OT-1에서 9.3% vs. 의도된 부위에서 13.3% (표 4 및 도 27)). 주목할 만하게, CR3-OT1은 분류된 hDMD 세포에서 서베이어 검정에 의해 유의한 수준의 활성을 보이는, 이들 3개의 오프-타겟 부위 중 유일한 하나였다 (도 26B).

[0370] 오프-타겟 부위에서의 뉴클레아제 활성은 별개의 염색체 상에서 절단된 표적과 오프-타겟 유전자좌 사이의 원위 재-라이게이션에 의해 의도치 않은 염색체 재배열을 유발할 수 있다. 이것은 멀티플렉스 CRISPR/Cas9 유전자 편집에서와 같이 2개 이상의 뉴클레아제를 사용함으로써 오프-타겟 활성에 대한 잠재력 증가로 인해 결실-기반 유전자 교정 전략에 대한 상당한 관심을 제시한다. 잠재적 전위는 단일 및 멀티플렉스 CRISPR/Cas9 편집 전략 둘 다 동안에 검증된 오프-타겟 유전자좌 (표 4)에서 전위를 검출하기 위해 고도로 감수성인 네스티드 게놈 PCR 검정을 사용하는 것에 대해 프로빙되었다. 이러한 검정을 사용하여, 전위는 높은 수준의 오프-타겟 활성을 또한 보이는 모델 HEK293T 세포주에서 온-타겟 부위와 오프-타겟 부위 사이에서 용이하게 검출되었다 (도 26C 및 도 28A, 28B). PCR 앰플리콘의 생어 서열분석은 각 프라이머 쌍에 대한 예측된 전위 사건의 동일성을 확인시켜 주었다 (도 29-30). 신호가 상당히 약했고 서열 동일성이 산물의 낮은 수율로 인해 확인되지 않았지만, HEK293T 세포에서 검출된 전위의 하위세트는 분류된 hDMD 근모세포에서 네스티드 PCR에 의해 또한 검출가능하였다 (도 26D 및 도 28A, 28C). 전위는 각각 CR6 또는 CR6/CR36으로 처리된 HEK293T 또는 분류된 hDMD 세포에서 이러한 검정을 사용하여 검출되지 않았으며 (도 28), HEK293T 세포에서의 단지 CR6-OT3에서만 오프-타겟 활성의 낮은 수준을 가졌다 (표 4). 이들 결과는 특히 멀티플렉스 편집 적용에 대해 고도로 특이적인 sgRNA를 선택하는 것이 중요하다는 것을 강조하고, 이러한 접근법이 CRISPR/Cas9 시스템의 특이성을 개선하기 위한 진행중인 노력으로부터 이익을 얻을 수 있다는 것을 제시한다. 이들 데이터는 선택된 sgRNA가 유의한 독성 없이 및 단지 검출가능한 수준의 활성을 갖는 단일의 강하게 예측된 오프-타겟 부위를 이용하여 디스트로핀 유전자를 교정할 수 있다는 것을 시사한다.

[0371] 실시예 13

[0372] 논의

[0373] 게놈 편집은 유전 질환을 교정하기 위한 강력한 도구이고, CRISPR/Cas9 시스템의 최근 개발은 이러한 분야에서 극적으로 가속화되는 진전이다. 또한 현재 어떠한 승인된 치료 옵션도 없는 가장 통상적인 유전 질환인 DMD의 교정이 입증되었다. DMD에 대한 많은 유전자- 및 세포-기반 요법이 전임상 개발 및 임상 시험 중에 있고, 게놈 편집 방법은 많은 이들 접근법과 상용성이다. 예를 들어, 게놈 편집은 DMD에 대한 환자-특이적 세포-기반 요법과 조합될 수 있다. CRISPR/Cas9 시스템은 제시된 바와 같이 인간 만능 줄기 세포 및 다른 인간 세포주, 뿐만 아니라 인간 골격 근모세포에서 기능할 수 있다. 중요하게, CRISPR/Cas9를 사용한 유전자 편집은, 면역결핍 마우스 내로의 이식 후에 시험관내 및 생체내 효율적인 디스트로핀 발현에 의해 입증된 바와 같이, 이들 세포의 근원성 능력을 제거하지 않았다. 따라서, 이러한 전략은 DMD에 대한 세포-기반 요법과 상용성이어야 한다.

- [0374] 추가로, 유전자-교정 세포의 보강된 풀은 면역결핍 마우스 내로의 생체 후 생체내 인간 디스트로핀의 발현을 입증하였다. CRISPR/Cas9 유전자 편집은 안정한 유전자 편집 빈도 및 여러 sgRNA의 최소 세포독성에 의해 관찰된 바와 같이 인간 근모세포에서 유의한 독성 효과를 갖지 않았다. 그러나, 유전자 편집 활성은 5개의 sgRNA에 걸쳐 50개의 예측된 오프-타겟 부위 중 3개에서 확인되었고, 온-타겟 부위와 오프-타겟 부위 사이의 CRISPR/Cas9-유도 염색체 전위는 검출가능하였다. CRISPR/Cas9 기술은 디스트로핀 돌연변이의 유의한 분획을 교정하기 위한 효율적이고 다양한 방법이며, 유전 질환을 치료하기 위한 일반적 플랫폼의 역할을 할 수 있다.
- [0375] 추가로, 본원에 사용된 플라스미드-기반 전달 방법과 달리, sgRNA 및 Cas9 mRNA의 직접적 형질감염은 Cas9 발현의 지속시간을 감소시키고 무작위 플라스미드 통합의 가능성을 제거함으로써 특이성 및 안전성을 증가시키는데 사용될 수 있다. 대안적으로, 바이러스, 플라스미드 또는 RNA 전달 벡터에 의한 골격근 및/또는 심장 근육으로의 CRISPR/Cas9 시스템의 직접 전달은 생체내 게놈 편집 및 이러한 접근법의 번역에 사용될 수 있다. 에스. 피오게네스 Cas9 유전자의 거대 크기 (~4.2 킬로베이스)는 크기-제한 아데노-연관 바이러스 벡터에서의 그의 사용에 대한 과제를 제시한다. 그러나, 다른 종, 예컨대 엔. 메닌기티디스 및 에스. 썬모필루스(*S. thermophilus*)로부터의 Cas9 유전자는 생체내 유전자 편집 적용을 위해 Cas9 및 sgRNA 발현 카세트 둘 다를 단일 AAV 벡터 내로 효율적으로 패키징하기에 충분히 짧다.
- [0376] CRISPR/Cas9 시스템은 다양한 유전자좌에서 이러한 시스템의 강한 활성의 다른 보고와 일치하는, 시험된 표적의 거의 90%의 효율적 변형을 가능하게 하였다. 이러한 기술의 강인성 및 다양성은 환자-특이적 유전자 편집의 임의적 생성을 향한 유의한 진보이다. 야생형 발현의 4%만큼을 포함하는 낮은 수준의 디스트로핀은 마우스 모델에서 생존, 운동 기능 및 심장 기능을 개선하는데 충분할 수 있다. CRISPR/Cas9 활성의 수준은 치료 이익에 충분할 수 있다.
- [0377] 엑손을 결실시키기 위한 CRISPR/Cas9와의 멀티플렉스화의 사용은 또한 기회 및 과제의 고유 세트를 제시한다. 단일 뉴클레아제의 작용 후 NHEJ-기반 DNA 복구에 의해 생성된 작은 indel로 디스트로핀 유전자의 리딩 프레임을 복원하는 것과 달리, 디스트로핀 발현을 복원하기 위해 게놈으로부터 완전한 엑손의 결실을 수행하였다. 각 DNA 복구 사건으로부터 신규 에피토프의 생성을 유도하게 될 단일 뉴클레아제의 내부엑손 작용에 의해 생성된 무작위 indel과 달리, 편집된 유전자의 단백질 산물은 예측가능하고 자연 발생 결실을 갖는 베키 근육 이영양증 환자인 것으로 이미 특성화되었다. 또한, 엑손 결실로부터 생성된 산물은 모든 성공적인 유전자 편집 사건에 대해 복원된 디스트로핀을 유도할 것이고, 반면에 엑손 내의 무작위 indel로 유전자를 변형시키는 것은 단지 정확한 리딩 프레임을 유도하는 편집 사건 중 3분의 1에서 리딩 프레임을 복원할 것이다.
- [0378] 시험된 모든 sgRNA는 인간 세포에서 유의한 세포독성 효과와 연관되지 않았다. 사용된 5개 sgRNA에 대해 시험된 총 50개 부위 중 3개의 잠재적 오프-타겟 부위가 디스트로핀 발현을 복원하는 것으로 확인되었다. 추가로, 의도된 온-타겟 부위와 이들 오프-타겟 부위 사이의 염색체 전위는 높은 수준의 Cas9 및 sgRNA를 발현하는 HEK293T 세포에서 고도로 감수성인 네스티드 PCR 검정에 의해 검출가능하였다. 주목할 만하게, 매우 높은 수준의 Cas9 및 sgRNA를 발현하는 불멸화 및 이수성 세포주인 HEK293T 세포에서 확인된 오프-타겟 활성 및 전위는 그 만큼 높은 수준으로 발생하지 않았고, 일부 경우에는 hMDM 근모세포에서 검출가능하지 않았다. 중요하게, 이러한 특이성 수준은 인간 세포에서 분명한 세포독성 효과의 결핍인 DMD의 중증도를 고려할 때 용인가능할 수 있다.
- [0379] 실시예 14
- [0380] SASTG (도 34; 서열 436 및 437)로 불리는 조작된 AAV 캡시드가 증진된 심장 근육 및 골격근 조직 형성을 위해 개발되었다 (Piacentino et al. (2012) Human Gene Therapy 23:635-646). Rosa26 유전자좌를 표적화하는 ZFN ("Rosa26 ZFN"; 도 33; 서열 434 및 435)은 마우스 세포에서 고도로 활성인 것으로 제시되었다 (Perez-Pinera et al. Nucleic Acids Research (2012) 40:3741-3752). Rosa26 ZFN 단백질을 코딩하는 AAV-SASTG 벡터를 설계하고, 이후 UNC 바이럴 벡터 코어(UNC Viral Vector Core)에 의해 생성하고 정제하였다. 서베이어 검정 (Guschin et al., Methods Mol Biol 649, 247-256 (2010))을 사용하여, 혈청 제거에 의해 분화로 활발히 사이클링 또는 강제된 배양 C2C12 근모세포에 AAV-SASTG Rosa26 ZFN의 전달 후 Rosa26 유전자좌에서의 NHEJ 돌연변이 유발을 입증하였다 (제시되지 않음).
- [0381] 성체의 유사분열 후 골격근이 AAV 전달 후 Rosa26 ZFN에 의해 효율적으로 표적화되었다는 것을 검증하기 위해, Rosa26 ZFN을 코딩하는 AAV-SASTG 벡터를 6주령 C57BL6/J 마우스의 전경골근 (TA) 근육 내로, 근육당 1e10 벡터 게놈 (vg) 또는 2.5e10 vg로 직접 주사하였다. 마우스를 주사 4주 후에 희생시키고, TA 근육을 수거하여 게놈 DNA 추출 및 분석을 위한 여러 단편으로 분할하였다 (도 31). 게놈 DNA를 PCR 증폭시키고, Rosa26 표적 부

위에서의 ZFN 돌연변이유발의 NHEJ 돌연변이 특성을 검출하기 위해 서베이어 검정에 적용하였다 (도 32). 도 32는 AAV-SASTG-ROSA의 전달 후 시험관내 및 생체내 골격근에서의 Rosa26 ZFN 활성의 서베이어 분석을 보여준다. 증식 C2C12를 나타낸 양의 바이러스로 형질도입하고, 형질감염 4일 후에 수거하였다 (도 32(a)). C2C12를 5일 동안 분화 배지에서 인큐베이션하고, 이어서 24웰 플레이트에서 나타낸 양의 AAV-SASTG-ROSA 바이러스로 형질도입하였다 (도 32(b)). 샘플을 형질도입 10일 후에 수집하였다. 나타낸 양의 AAV-SASTG-ROSA를 C57BL/6J 마우스의 전경골근 내로 직접 주사하고, 근육을 감염 4주 후에 수거하였다. 수거된 TA 근육을 게놈 DNA 분석을 위한 8개의 별개의 조각으로 분할하였으며, 이들 각각은 별개의 레인에 제시되어 있다 (도 32(c)). 주목할 만하게, 높은 수준의 유전자 변형은 가장 높은 용량 (2.5e10 vg)의 모든 단편에서 검출되었다.

[0382] 실시예 15

[0383] 돌연변이체 디스트로핀 유전자를 표적화하는 AAV-CRISPR 구축물

[0384] AAV 구축물은 뒤시엔느 근육 이영양증 및 골격근 및 심장 근육 변성을 유발하는 디스트로핀 유전자의 돌연변이를 치료상 교정하도록 설계한다. CRISPR/Cas9 시스템은 AAV를 사용하여 전달되어, 엑손 51을 결실시키거나, 엑손 45-55를 결실시키거나, 스플라이스 공여자 또는 수용자 부위를 파괴하거나, 또는 엑손 51 내에 프레임시프트를 생성함으로써 디스트로핀 리딩 프레임을 복원하여 (Ousterout et al., Molecular Therapy 2013) 디스트로핀 리딩 프레임 및 단백질 발현을 복원할 수 있다. CRISPR/Cas9 시스템은 서열 64 또는 114를 갖는 Cas9를 포함할 것이다 (도 40 및 41 참조). 각 PAM 서열을 표적화하는, 이들 Cas9와 조합될 수 있는 gRNA가 제공된다 (도 40 및 41 참조; 표 2 및 3 또한 참조).

[0385] 실시예 16

[0386] 유도된 뉴런 (iN)의 생성

[0387] 다른 세포 계통으로부터 유도된 뉴런 (iN)의 생성은 재생 의학 및 신경계 질환의 연구에서 잠재적 적용을 갖는다. 마우스 배아 섬유모세포 (MEF)의 기능적 뉴런 세포로의 직접적 전환은 3개의 뉴런 전사 인자의 카테일 - BRN2, ASCL1 및 MYT1L의 전달을 통해 발생할 수 있다 (BAM 인자, 도 48). 다른 방법은 다양한 뉴런 하위유형을 유도하기 위한 추가의 인자를 포함할 수 있다. 이들 실험은 전사 인자가 이소성으로 전달되고 상응하는 내인성 유전자와의 활성화가 뉴런 표현형을 지속시키도록 요구한다. CRISPR/Cas9 시스템은 RNA-가이드 메커니즘을 통해 계놈에서 임의의 프로모터를 표적화하는 능력을 갖는 포유동물 세포에서 내인성 유전자를 활성화시키기 위한 다양한 전사 인자로서 조작하였다 (도 49A, 49B).

[0388] 물질 & 방법. CRISPR/Cas9-전사 인자를 사용하여 ASCL1 및 BRN2를 코딩하는 내인성 유전자를 활성화시켜 MEF를 기능적 유도된 뉴런으로 직접적으로 재프로그래밍하였다.

[0389] 세포 배양: MEF를 24-웰 TCPS 플레이트에 또는 폴리-D-리신/라미닌-코딩 커버슬립 상에 시딩하였다. dCas9-VP64의 형질도입 및 gRNA의 형질감염 후에 (gRNA의 서열에 대해 표 10 및 11 참조), 세포를 24 시간 동안 MEF 배지 (Adler et al. Mol Ther Nucleic Acids 1:e32 (2012))에서 배양한 다음, 실험의 지속기간 동안 N3 신경 유도 배지 (Vierbuchen et al. Nature 463:1035-1041 (2010))로 전달하였다 (도 49B).

[0390] <표 10>



[0391] 마우스 ASCL1에 대한 gRNA (CR13)

올리고 (5'에서 3'로)	5' 오버행		ASCL1 표적 서열	서열 번호
CR13-1_S:	cacc	G	CAGCCGCTCGCTGCAGCAG (서열 468)	492
CR13-1_AS:	AAAC		CTGCTGCAGCGAGCGGCTG (서열 469)	C 493
CR13-2_S:	cacc	G	GCTGGGTGTCCCATTGAAA (서열 470)	494
CR13-2_AS:	AAAC		TTTCAATGGGACACCCAGC (서열 471)	C 495
CR13-3_S:	cacc	G	GTTTATTTCAGCCGGGAGTC (서열 472)	496
CR13-3_AS:	AAAC		GACTCCCGGCTGAATAAAC (서열 473)	C 497
CR13-4_S:	cacc	G	TGGAGAGTTTGCAAGGAGC (서열 474)	498
CR13-4_AS:	AAAC		GCTCCTTGCAAACCTCTCCA (서열 475)	C 499
CR13-5_S:	cacc	G	CCCTCCAGACTTTCCACCT (서열 476)	500
CR13-5_AS:	AAAC		AGGTGGAAAGTCTGGAGGG (서열 477)	C 501
CR13-6_S:	cacc	G	AATTTTCTTCCAAGTTCTC (서열 478)	502
CR13-6_AS:	AAAC		GAGAACTTGAAGAAAATT (서열 479)	C 503
CR13-7_S:	cacc	G	CTGCGGAGAGAAGAAAGGG (서열 480)	504
CR13-7_AS:	AAAC		CCCTTTCTTCTCTCCGCAG (서열 481)	C 505
CR13-8_S:	cacc	G	AGAGCCACCCCTGGCTCC (서열 482)	506
CR13-8_AS:	AAAC		GGAGCCAGGGGGTGGCTCT (서열 483)	C 507
CR13-9_S:	cacc	G	cgaagccaaccgcggcg (서열 484)	508
CR13-9_AS:	AAAC		cccgcgcggttggttcg (서열 485)	C 509
CR13-10_S:	cacc	G	agagggaagacgatcgccc (서열 486)	510
CR13-10_AS:	AAAC		ggcgcatcgtcttccctct (서열 487)	C 511
CR13-11_S:	cacc	G	ccccttaactttctcgcg (서열 488)	512
CR13-11_AS:	AAAC		cggaggaaagttaagggg (서열 489)	C 513
CR13-12_S:	cacc	G	gcagccccgttcttcaa (서열 490)	514
CR13-12_AS:	AAAC		tgaaggaagcggggctgc (서열 491)	C 515

[0392]

[0393] <표 11>

[0394] 마우스 BRN2에 대한 gRNA (CR16)

올리고 (5'에서 3'로)	5' 오버행		BRN2 표적 서열	서열 번호
CR16-1_S:	cacc	G	CGAGAGCGAGAGGAGGGAG (서열 516)	540
CR16-1_AS:	AAAC		CTCCCTCTCTCGCTCTCG (서열 517)	C 541
CR16-2_S:	cacc	G	GAGAGAGCTTGAGAGCGCG (서열 518)	542
CR16-2_AS:	AAAC		CGCGCTCTCAAGCTCTCTC (서열 519)	C 543
CR16-3_S:	cacc	G	GGTGGAGGGGGCGGGGCC (서열 520)	544
CR16-3_AS:	AAAC		GGGCCCCGCCCTCCACC (서열 521)	C 545
CR16-4_S:	cacc	G	GGTATCCACGTAAATCAA (서열 522)	546
CR16-4_AS:	AAAC		TTTGATTACGTGGATACC (서열 523)	C 547
CR16-5_S:	cacc	G	CCAATCACTGGCTCCGTC (서열 524)	548
CR16-5_AS:	AAAC		GACCGGAGCCAGTGATTGG (서열 525)	C 549
CR16-6_S:	cacc	G	GGCGCCGAGGGAAGAAGA (서열 526)	550
CR16-6_AS:	AAAC		TCTTCTTCCCTCGGGCGCC (서열 527)	C 551
CR16-7_S:	cacc	G	GGGTGGGGTACCAGAGGA (서열 528)	552
CR16-7_AS:	AAAC		TCCTCTGGTACCCCCACCC (서열 529)	C 553
CR16-8_S:	cacc	G	CCGGGGACAGAAGAGAGGG (서열 530)	554
CR16-8_AS:	AAAC		CCCTCTTCTGTCCCCGG (서열 531)	C 555
CR16-9_S:	cacc	G	gagagagagtgggagaagc (서열 532)	556
CR16-9_AS:	AAAC		gcttctccactctctctc (서열 533)	C 557
CR16-10_S:	cacc	G	aaagtaactgtcaaatgcg (서열 534)	558
CR16-10_AS:	AAAC		cgcatttgacagttacttt (서열 535)	C 559
CR16-11_S:	cacc	G	ttaaccagagcgccagtc (서열 536)	560
CR16-11_AS:	AAAC		gactggcgctctgtgtaa (서열 537)	C 561
CR16-12_S:	cacc	G	cgtcggagctgcccgtag (서열 538)	562
CR16-12_AS:	AAAC		ctagcgggcagctccgacg (서열 539)	C 563

[0395]

[0396] qRT-PCR & IF: 내인성 ASCL1의 활성화는 dCas9-VP64 및 gRNA, ASCL1 cDNA, 또는 루시페라제를 코딩하는 음성 대조군 벡터의 전달 후 제3일에 MEF에서 qRT-PCR 및 면역형광에 의해 평가하였다. iN의 생성은 TUJ1 및 MAP2 공동-염색 및 뉴런 형태와 확장된 과정을 사용한 세포의 확인에 의해 평가하였다.



- [0397] 살아있는 세포 리포터: N3 배지에서 7-8일 후에, 폴리-D-리신/라미닌-코팅 커버슬립 상에서 배양된 MEF를 hSyn-RFP 및 MAP2-GCaMP5 리포터를 보유하는 바이러스로 형질도입하여 칼슘 영상화 및 전기생리학을 통한 기능적 특성화를 위해 가장 성숙한 iN를 확인하였다 (도 49B).
- [0398] 결과. dCas9-VP64 및 ASCL1 프로모터를 표적화하는 gRNA는 MEF에서 내인성 유전자를 활성화시켰다. 8개의 gRNA의 공동-전달은 내인성 유전자를, 4개의 gRNA의 공동-전달에 의해 유도된 100-배 활성화에 비해 유의한 증가 ( $p < 0.05$ )인 400배 활성화시켰다 (도 50A). 핵-국제화 Asc11 단백질을 MEF에서 면역형광에 의해 검출하였다. 이소성 Asc11 발현은 임의의 gRNA 카테일을 갖는 dCas9-VP64보다 더 많은 Asc11 단백질을 생산하였지만, 제3일까지 내인성 유전자좌를 활성화시키지 않았다 (도 50A, 50B). 확장된 과정을 갖는 TUJ1 및 MAP2 공동-양성 세포는 dCas9-VP64 및 ASCL1 및 BRN2 프로모터를 표적화하는 gRNA의 전달 13일 후에 신경성 배지에서 면역형광에 의해 확인되었다 (도 4a 제1 열). 유사한 개수의 TUJ1 및 MAP2 공동-양성 세포는 BAM 인자의 이소성 발현으로 확인되었다 (도 51A 제2 열). hSyn-RFP 리포터를 발현하는 뉴런 형태를 갖는 세포는 신경성 배지에서 제11일 정도로 일찍 배양 중에 보였다 (도 51B). MAP2-GCaMP5 칼슘 지시제를 발현하는 세포는 형광 현미경으로 검출된 KC1-유도 탈분극을 나타냈다 (도 52a, 52b).
- [0399] 마우스 배아 섬유모세포의, 뉴런 형태를 갖는 TUJ1 및 MAP2 공동-양성 세포로의 직접적 전환은 CRISPR/Cas9-기반 전사 인자에 의한 내인성 BRN2 및 ASCL1의 활성화를 통해 수행하였다. dCas9-VP64가 ASCL1의 이소성 발현보다 단백질을 덜 생산하였지만 (도 50B), 뉴런-유사 세포의 생성은 유사하다. 내인성 유전자좌의 활성화는 이소성 발현으로 생성된 경우와 기계적으로 동일하지 않은 사건의 재프로그래밍 캐스케이드를 유도할 수 있다.
- [0400] dCas9-VP64는 이질염색질을 침투할 수 있었고, 단지 "선구자" 전사 인자의 하위세트의 특성을 갖는 안정하게 침묵화된 내인성 유전자를 활성화시켰다. 그 결과, CRISPR/Cas9-전사 인자를 사용하여 세포 계통을 전환시키는 것은 특히 재프로그래밍하기 어려운 세포 유형, 예컨대 성체 인간 세포에서 전사 인자의 이소성 발현을 재프로그래밍하는 것에 대한 후성적 장벽을 더 잘 극복할 수 있다. 이것은, 세포 대체 요법에서 자가 공급원을 사용하는 것이 종종 바람직한 것처럼, 재생 의학 분야에서 임상 중요성을 가질 수 있다.
- [0401] 실시예 17
- [0402] 멀티플렉스 CRISPR/Cas9-기반 게놈 조작 - 물질 및 방법
- [0403] 플라스미드 구축물. 상기 기재된 바와 같은 에스. 피오게네스 sgRNA 및 인간 코돈 최적화된 Cas9 (hCas9) 뉴클레아제에 대한 발현 카세트를 사용하였다. 추가의 프로모터 mU6 (Ohshima et al., Nucleic Acids Res 9:5145-5158 (1981)), H1 (Myslinski et al., Nucleic Acids Res 29:2502-2509 (2001)), 및 7SK (Murphy et al., Cell 51:81-87 (1987)) POL-III 프로모터는 진블록 (IDT)을 사용하여 합성하고, hU6 sgRNA 발현 카세트 대신에 클로닝하였다. 진블록 (IDT)은 Cas9 코딩 서열의 3' 말단 상에 클로닝하여 T2A 스킵핑 펩티드 및 eGFP 유전자를 Cas9 바로 뒤에 융합시켜 백터 발현을 모니터링하였다. 이어서, hCas9-T2A-GFP (서열 145)에 대한 코딩 영역은, hCas9-T2A-GFP의 발현을 구동하기 위한 인간 유비퀴틴 C (hUbC) 프로모터, 뿐만 아니라 hUbC 프로모터 바로 상류의 sgRNA 발현 카세트의 골든 게이트 클로닝을 촉진하기 위한 제한 부위를 함유하는 렌티바이러스 발현 백터 내로 전달하였다 (도 42A).
- [0404] 맞춤 렌티바이러스 백터의 조립을 위한 프로토콜. 최대 4개의 선택 sgRNA 및 활성 Cas9, dCas9 또는 dCas9-VP64를 발현하는 맞춤 렌티바이러스 백터의 조립을 5일 미만으로 수행하였다. 클로닝 방법은 골든 게이트 클로닝, 및 인식 서열 밖을 절단하여 고유한 오버행을 생성하는 유형 IIS 제한 효소를 사용하였다. 골든 게이트 조립은 클로닝을 촉진하였으며, 모든 4개의 발현 카세트가 한 단계에서 최종 렌티바이러스 백터 내로 라이게이션되었다. 렌티바이러스 백터는 독립적 프로모터로부터 발현된 1, 2, 3 또는 4개의 sgRNA 이외에도 활성 Cas9, cCas9 또는 dCas9-VP64를 발현하였다.
- [0405] 단계 1: 각 20 bp 프로토스페이서를 함유하는 단일 가닥 올리고를 점착성 말단을 생성하는 방식으로 어닐링하고, 목적하는 pZDonor-프로모터 백터 내로 라이게이션하였다. 각 목적하는 게놈 표적을 위해 2개의 단일 가닥 올리고를 주문하였다. 상보적 올리고를 어닐링하기 위해, 8  $\mu$ L 센스 올리고 + 8  $\mu$ L 안티센스 올리고 (둘 다 10mM) + 2  $\mu$ L 10X 리가제 완충제를 혼합하였다. 올리고를 용융시키고, 하기 프로그램으로 PCR 기계에서 제어닐링하였다: 단계 사이에 -0.3°C/초의 비율로 300초 동안 96°C, 이어서 20초 동안 85°C, 20초 동안 75°C, 20초 동안 65°C, 20초 동안 55°C, 20초 동안 45°C, 20초 동안 35°C, 및 20초 동안 25°C. 점착성 말단을 인산화하기 위해, 1  $\mu$ L 25mM ATP + 1 마이크로리터 T4 폴리뉴클레오티드 키나제 (NEB)를 첨가하고, 37°C에서 60분 동안, 이어서 65°C에서 20분 동안 인큐베이션하여 효소를 열 불활성화시켰다. 각 프로토스페이서는, 제조

업체의 지침서에 따라 10  $\mu$ L 반응 부피 중 50ng의 벡터 및 1  $\mu$ L의 어닐링된 올리고뉴클레오타이드를 사용하여 16°C에서 60분 동안 인큐베이션된 T4 DNA 리가제 (NEB)를 사용하여 목적하는 발현 벡터 내로 라이게이션하였다. 5 마이크로리터의 각 라이게이션물을 제조업체의 지침서에 따라 XL1 블루 화학적 적격 박테리아 (애질런트 (Agilent)) 내로 형질전환시켰다. 50  $\mu$ g/mL 카나마이신 (시그마)을 함유하는 LB 한천 플레이트 상으로 형질전환체를 플레이트팅하고, 37°C에서 밤새 인큐베이션하였다. 본 발명자들의 경험상, 콜로니의 >90%는 목적하는 라이게이션 산물을 함유할 것이다. M13 역 표준 서열분석 프라이머를 사용하는 서열분석을 수행하여 단계 2 상으로 이동하기 전에 각 최종 sgRNA 구축물을 검증하였다.

[0406] 단계 2: 골든 게이트 조립을 사용하여 렌티바이러스 목적 벡터 내로 4개의 프로모터-gRNA 카세트를 구축하였다. 단계 1의 완료 후에, 상이한 프로모터로부터 상이한 sgRNA를 각각 발현하는 4개의 독립적 플라스미드가 존재하였다. 바람직한 목적 벡터 내로 4개의 상이한 프로모터-sgRNA 구축물을 조립하기 위해, 200 ng의 각 sgRNA 발현 플라스미드 및 20  $\mu$ L 반응 부피 중 1  $\mu$ L T4 DNA 리가제 (NEB), 1  $\mu$ L BsmBI FastDigest (피셔 사이언티픽 (Fisher Scientific)) 및 2  $\mu$ L 10X T4 리가제 완충제 (NEB)를 갖는 바람직한 렌티바이러스 목적 벡터를 조합하였다. 반응물을 하기와 같이 인큐베이션하였다: 10분 동안 37°C, 15분 동안 16°C, 30분 동안 37°C, 5분 동안 80°C. 5  $\mu$ L의 라이게이션 반응물을 SURE 2 화학적 적격 세포 (애질런트) 내로 제조업체의 지침서에 따라 형질전환시켰다. 100  $\mu$ g/mL 암피실린을 함유하는 LB 한천 상에 형질전환체를 플레이트팅하고, 37°C에서 밤새 인큐베이션하였다. 임의로, 콜로니를 IPTG 및 X-gal을 사용하는 lacZ-기반 블루/화이트 스크리닝에 의해 스크리닝할 수 있었지만; 그러나 본 발명자들의 경험상, 형질전환체의 >90%는 적절한 라이게이션 산물을 함유한다. 대항하는 sgRNA 발현 카세트에 의해 형성된 역전된 반복부호 인해, 최종 구축물은 불안정할 수 있고, 따라서 본 발명자들은 이들 플라스미드를 SURE 2 세포주에서 유지하고 최종 플라스미드를 센스 프라이머 5'-TCGGGTTTATTACAGGGACAGCAG-3' (서열 464) 및 안티센스 프라이머 5'-TCTAAGGCCGAGTCTTATGAGCAG-3' (서열 465)를 사용하는 시험 PCR로 스크리닝할 것을 권고한다. 이들 프라이머는 4개 프로모터-gRNA 영역에 걸쳐 증폭시킨다. 반복적 성질로 인해, 뚜렷한 밴딩 패턴은 크기가 대략 1800 bp인 가장 큰 산물로 관찰될 것이다.

[0407] 세포 배양 및 형질감염. HEK293T 세포를 듀크 대학교 암 센터 기관을 통해 아메리칸 티슈 콜렉션 센터 (ATCC, 버지니아주 마나사스)로부터 획득하였고, 10% FBS 및 1% 페니실린 스트렙토마이신으로 보충된 DMEM에서 유지하였다. 1차 인간 피부 섬유모세포 (카탈로그 ID: GM03348)를 코리엘 연구소 (뉴저지주 캄덴)로부터 획득하였고, 10% FBS 및 1% 페니실린 스트렙토마이신으로 보충된 DMEM에서 유지하였다. 모든 세포를 5% CO<sub>2</sub> 하에 37°C에서 배양하였다. HEK293T 세포를 24 웰 플레이트에서 200 ng의 각 sgRNA 발현 벡터 (800 ng 총 pDNA)와 함께 리포펙타민 2000 (라이프 테크놀로지스)으로 제조업체의 프로토콜에 따라 형질감염시켰다.

[0408] 바이러스 생산 및 형질도입. 사용된 모든 렌티바이러스 벡터는 제2 세대였고, 표준 바이러스 생산 방법을 사용하여 생산하였다. 간략하게, 3.5백만개의 HEK293T를 10 cm 접시마다 플레이트팅하였다. 다음 날, 세포를 20  $\mu$ g의 전달 벡터, 6  $\mu$ g의 pMD2G 및 10  $\mu$ g psPAX2를 사용하여 인산칼슘 형질감염 방법에 의해 형질감염시켰다. 배지를 형질감염 12-14시간 후에 교체하였다. 바이러스 상청액을 이러한 배지 교체 24 및 48시간 후에 수집하고, 0.45 마이크로미터 필터를 통과시키고, 폴링하였다. 형질도입을 위해, 세포 배지를 4  $\mu$ g/mL 폴리브렌으로 보충된 바이러스 상청액으로 대체하였다. 바이러스 상청액을 12-24시간 후에 신선한 배지로 교환하였다.

[0409] 역전사 PCR. RNA는 miRNeasy 미니 RNA 단리 키트 (키아젠)를 사용하여 단리하였다. DNase 소화는 DNA-무함유 키트 (어플라이드 바이오시스템즈(Applied Biosystems))를 사용하여 수행하였다. cDNA 합성은 슈퍼스크립트 VILO cDNA 합성 키트 (인비트로젠)를 사용하여 수행하였다. cDNA는 Taq DNA 폴리머라제 (NEB)를 사용하여 PCR 증폭시켰고, 생성된 산물을 TAE 아가로스 겔 상에 전개시켰다. 영상은 케미독 XRS+ 시스템을 사용하여 포획하고, 이미지랩 소프트웨어 (바이오-라드)를 사용하여 처리하였다.

[0410] 정량적 실시간 PCR. RNA는 RNeasy 플러스 RNA 단리 키트 (키아젠)를 사용하여 단리하였다. cDNA 합성은 슈퍼스크립트 VILO cDNA 합성 키트 (인비트로젠)를 사용하여 수행하였다. 퍼펙타 SYBR 그린 패스트믹스 (퀀타 바이오사이언시스(Quanta Biosciences))를 사용하는 실시간 PCR은 CFX96 실시간 PCR 검출 시스템 (바이오-라드)으로 수행하였다. 프라이머 특이성은 아가로스 겔 전기영동 및 용융 곡선 분석에 의해 확인하였다. 적절한 동적 범위에 걸친 반응 효율을 계산하여 표준 곡선의 선형성을 확실히 하였다. 결과는  $\Delta\Delta C_t$  방법을 사용하여  $\beta$ -액틴 발현에 대해 정규화된 관심 유전자의 mRNA 발현 배수 증가로서 표현된다. 보고된 값은 2개의 독립적 실험 (n = 2)으로부터의 평균 및 S.E.M.이며, 여기서 기술적 반복이 각 실험에 대해 평균이 내려졌다.

[0411] 웨스턴 블롯. 세포는 프로테아제 억제제 콕테일 (시그마)로 보충된 RIPA 완충제 (시그마) 중에 용해시켰다. 단백질 농도는 BCA 단백질 검정 시약 (써모 사이언티픽(Thermo Scientific)) 및 바이오텍 시너지 2 다중-모드

마이크로플레이트 판독기를 사용하여 측정하였다. 용해물을 로딩 완충제와 혼합하고 5분 동안 비등시켰다; 25  $\mu$ g의 단백질을 NuPage 10% 비스-트리스 겔 폴리아크릴아미드 겔 (바이오-라드)에서 전개시키고, 니트로셀룰로스 막으로 전달하였다. 비특이적 항체 결합은 실온에서 1시간 동안 5% 탈지유를 갖는 TBST (50mM 트리스, 150mM NaCl 및 0.1% 트윈-20)로 차단하였다. 막을 하기 1차 항체와 함께 인큐베이션하였다: 4°C에서 밤새 TBST 중 5% BSA 중 항-MyoD (1:250, 산타 크루즈 Sc-32758); 4°C에서 밤새 5% BSA 중 항-미오제닌 (1:250, 산타 크루즈 Sc-12732); 실온에서 60분 동안 TBST 중 5% 밀크 중 항-FLAG-HRP (1:1000, 셀 시그널링 2044); 실온에서 30분 동안 TBST 중 5% 밀크 중 항-GAPDH (1:5000, 셀 시그널링, 클론 14C10). 이어서, 막을 총 15분 동안 TBST로 3회 세척하였다. 막을 30분 동안 1:5,000로 희석되고 각각 15분 동안 TBST로 3회 세척된 항-토끼 HRP-접합 항체 (시그마, A 6154) 또는 항-마우스 HRP-접합 항체 (산타 크루즈, SC-2005)와 함께 인큐베이션하였다. 막을 이문스타 웨스턴씨 화학발광 키트 (바이오-라드)를 사용하여 가시화하였고, 영상을 케미독 XRS+ 시스템을 사용하여 포획하고, 이미지랩 소프트웨어 (바이오-라드)를 사용하여 처리하였다.

[0412] 내인성 유전자 변형의 Cel-I 정량화. 내인성 표적 부위에서의 CRISPR/Cas9 뉴클레아제 병변은 뉴클레아제-매개 NHEJ의 돌연변이 특성을 검출할 수 있는 서베이어 뉴클레아제 검정을 사용하여 정량화하였다. 형질감염 또는 형질도입 후에, 세포를 37°C에서 3-10일 동안 인큐베이션하고, 게놈 DNA를 DNeasy 혈액 및 조직 키트 (키아젠)를 사용하여 추출하였다. 표적 유전자좌를 아큐프라임 고 충실도 PCR 키트 (인비트로젠)로 35 사이클의 PCR에 의해 증폭시켰다. 생성된 PCR 산물은 무작위로 용융시키고, 하기 프로그램으로 PCR 기계에서 재어닐링하였다: 단계 사이에 -0.3°C/초의 비율로 240초 동안 95°C, 이어서 60초 동안 85°C, 60초 동안 75°C, 60초 동안 65°C, 60초 동안 55°C, 60초 동안 45°C, 60초 동안 35°C, 및 60초 동안 25°C. 재-어닐링 후에, 8  $\mu$ L의 PCR 산물을 1  $\mu$ L의 서베이어 뉴클레아제 S 및 1  $\mu$ L의 인헨서 S (트랜스게노믹, 네브라스카주 오마하)와 혼합하고, 42°C에서 1시간 동안 인큐베이션하였다. 인큐베이션 후에, 6  $\mu$ L의 소화 산물을 10% TBE 폴리아크릴아미드 겔 상에 로딩하고, 30분 동안 200V에서 전개시켰다. 겔을 브로민화메틸황으로 염색하고, 이미지랩 (바이오-라드)을 사용하여 밀도측정법에 의해 정량화하였다 (Perez-Pinera et al., *Nucleic Acids Res* 40:3741-3751 (2012)).

[0413] 통계적 분석. 적어도 2개의 독립적 실험을 평균 및 평균의 표준 오차로서 컴파일링하였다. 효과는 JMP 10 프로그램을 사용하여 다변량 ANOVA 및 던넛 사후 검정으로 평가하였다.

[0414] 실시예 18

[0415] 멀티플렉스 CRISPR/Cas9 적용을 위한 단일 렌티바이러스 벡터의 개발

[0416] 현재 CRISPR/Cas9 유전자 편집 시스템, 특히 전사활성인자 시스템의 제한은, 특히 형질감염시키기에 곤란한 세포 유형에서, 멀티플렉스 유전자 편집 및 상승작용적 유전자 활성화 적용을 위해 다중 sgRNA 및 Cas9 단백질을 동시에 효율적으로 전달하는 것이다. 이러한 제한을 극복하기 위해, 본 발명자들은 Cas9 및 최대 4개의 sgRNA를 효율적으로 발현하는 단일 렌티바이러스 벡터를 개발하였다. 각 sgRNA의 발현 효율을 최대화하기 위해, 이러한 벡터는 4개의 독립적 pol III 프로모터 (인간 U6 프로모터, 마우스 U6 프로모터, 7SK 및 H1)로부터 4개의 sgRNA를 발현한다. 본 발명자들은 AAVS1 유전자좌를 표적화하는 sgRNA를 검출하기 위한 엔드-포인트 RT-PCR을 사용하여 각 프로모터로부터의 sgRNA 발현을 검증하였다 (도 42A). 각 sgRNA 발현 구축물의 활성을 시험하기 위해, 본 발명자들 활성 Cas9 발현 구축물과 독립적으로 AAVS1을 표적화하는 sgRNA를 발현하는 각 프로모터 구축물을 인간 HEK293T 세포 내로 공동-형질감염시켰다. 주목할 만하게, 본 발명자들은 AAVS1 유전자좌에서 높은 활성을 갖는 잘-특성화된 아연-핑거 뉴클레아제와 필적할 만한, 각 sgRNA에 대한 표적 유전자좌에서 일관되고 높은 수준의 유전자 변형을 검출하였다 (도 42B). 또한, 활성 Cas9 뉴클레아제, 사멸 Cas9, 및 VP64 전사활성인자 도메인에 융합된 사멸 Cas9를 포함하는 상이한 Cas9-기반 구축물의 렌티바이러스 전달은 웨스턴 블롯에 의해 결정된 바와 같이 HEK293T 세포에서 전장 Cas9 단백질의 발현을 유발하였다 (도 42C).

[0417] 이들 성분을 사용하여, 본 발명자들은 목적하는 Cas9 이펙터를 발현하는 단일 렌티바이러스 벡터 내로 다중 sgRNA 발현 카세트의 신속하고 효율적인 클로닝을 촉진하기 위한 골든 게이트 클로닝 방법을 개발하였다 (도 43). 제1 단계에서, sgRNA 프로토스페이서 서열을 코딩하는 올리고뉴클레오타이드를, 각각 sgRNA 발현을 구동하는 별개의 프로모터를 갖는 상이한 발현 벡터 내로 독립적으로 클로닝하였다. 제2 단계에서, 각 sgRNA 발현 구축물을 골든 게이트 조립에 의해 선택 렌티바이러스 Cas9 발현 벡터 내로 서브클로닝하였다. 이러한 전략은 유전자 편집 또는 활성화 적용을 위해 최대 4개의 sgRNA를 단일 렌티바이러스 벡터 내로 강하고 신속하게 클로닝하도록 하였다. 4개 미만의 sgRNA를 발현하기 위해, 폴리T 종결인자 서열을 미사용 프로모터 하류에 클로닝하여 미사용 프로모터로부터의 전사를 방지하였다. 각 벡터는 2A 스킵핑 펩티드를 통해 선택 Cas9과 eGFP를 공-발현하여 형광-활성화 분류 및 높은 감염 다중도로의 세포의 보강을 가능하게 하였다. 최종적으로, sgRNA 및



Cas9 발현 카세트를 함유하는 전체 영역은 loxP 부위에 플랜킹되어 Cre-lox 절제에 의한 제거를 매개하였다.

[0418]

실시예 19

[0419]

멀티플렉스 게놈 조작을 위한 단일 렌티바이러스 sgRNA/Cas9 발현 벡터의 검증

[0420]

각 sgRNA의 독립적 활성을 검증하기 위해, 본 발명자들은 활성 Cas9 및 각각 독립적 유전자좌를 표적화하는 4개의 sgRNA를 발현하는 단일 렌티바이러스 벡터를 조립하였다 (도 44A). 대조군 벡터로서, 본 발명자들은 다른 3개 위치에서 폴리T 프로토스페이서와 함께 단지 1개의 sgRNA를 발현하는 구축물을 조립하였다. 본 발명자들은 HEK293T 및 1차 섬유모세포를, 나타낸 sgRNA를 발현하는 렌티바이러스 벡터로 형질도입하고, 형질도입 7 또는 10일 후에 각각 유전자 변형 빈도를 모니터링하였다 (도 44B). 두 세포 유형에서, 단일 렌티바이러스 벡터는 모든 4개의 유전자좌에서 고도로 효율적인 멀티플렉스 편집을 매개하였다 (도 44B). 흥미롭게도, 모든 4개의 sgRNA 함께 발현은 섬유모세포에서의 4개의 유전자좌 중 3개에서 단일 sgRNA 단독보다 더 높은 변형 빈도를 유발하였다 (도 44B). 본 발명자들은 통상적으로 형질감염시키기에 곤란한 세포 유형인 섬유모세포에서 효율적인 멀티플렉스 유전자 편집을 관찰하였다. 이들 데이터는 단일 렌티바이러스가 4개의 활성 sgRNA를 효율적으로 발현할 수 있고, 이러한 렌티바이러스 플랫폼이 멀티플렉스 CRISPR/Cas9 유전자 편집을 위한 4개의 별개의 유전자좌를 표적화하는데 사용될 수 있다는 것을 입증한다.

[0421]

실시예 20

[0422]

렌티바이러스 Cas9-기반 전사활성인자를 안정하게 발현하는 세포주에서의 일시적 RNA-가이드 유전자 활성화

[0423]

다음에, 본 발명자들은 Cas9를 안정하게 발현하는 모델 세포주 내로 sgRNA를 형질감염시킴으로써 일시적 유전자 활성화를 가능하게 하는 시스템을 개발하는 것에 관심이 있었다. HEK293T를 상이한 Cas9-T2A-GFP로 형질도입하고, GFP 발현을 유동 세포측정법을 사용하여 모니터링하였다. 2-3일마다의 정상 계대배양 후에, 각 세포주는 형질도입 후 최대 35일 동안 안정한 GFP 발현을 나타냈다. 이어서, 형질도입된 HEK293T를 IL1RN 또는 HBG1 프로모터를 표적화하는 1 내지 4개의 별개 sgRNA 발현 구축물로 형질감염시켰다. 안정한 dCas9-VP64 발현 세포주에서의 이들 sgRNA 구축물의 일시적 형질감염은 조정가능한 내인성 유전자 활성화를 유발하였다 (도 45A, 45B). dCas9-VP64를 발현하는 세포에 sgRNA 구축물의 일시적 형질감염 후 유전자 활성화는 형질감염 3-6일 후 최대 수준의 활성화에 도달하였고, 형질감염 20일 후까지 검출불가능한 수준으로 하락하였다 (도 45C, 45D). 또한, 본 발명자들은, 활성화 수준이 제1 형질감염으로부터 관찰된 것보다 유의하게 낮았지만, 각 프로모터를 표적화하는 모든 4개의 sgRNA 구축물의 제2 형질감염에 의해 각 유전자를 재활성화시킬 수 있었다 (도 45C, 45D). 제2 형질감염 후 이러한 활성화 감소는 감소된 벡터 발현 또는 미형질도입 세포의 경쟁적 성장으로 인한 것일 수 있다. 이것에도 불구하고, 이들 데이터는 일시적 sgRNA 전달과 조합된 렌티바이러스 Cas9가 다양한 시스템으로서 사용되어 Cas9 안정적 형질도입 세포주에서 표적 유전자를 조정가능하게 및 일시적으로 활성화 및 재-활성화시킬 수 있다는 것을 입증한다.

[0424]

실시예 21

[0425]

단일 렌티바이러스 sgRNA/Cas9 전사활성인자 발현 벡터를 사용하는 HEK293T 세포에서의 안정한 유전자 활성화

[0426]

렌티바이러스 전달은 CRISPR/Cas9 전사활성화에 의한 안정한 장기적 유전자 활성화를 가능하게 할 수 있다. 이것을 시험하기 위해, HEK293T를 dCas9-VP64 및 1 내지 4개의 sgRNA 발현 카세트를 코딩하는 단일 렌티바이러스 벡터를 사용하여 형질도입하였다. 본 발명자들의 일시적 형질감염 결과 (도 45)와 유사하게, 본 발명자들은 조정가능하게 및 강하게 내인성 IL1RN 및 HBG1 유전자의 발현을 활성화시킬 수 있었다 (도 46A, 46B). dCas9-VP64 및 IL1RN 및 HBG1 프로모터를 표적화하는 4개의 sgRNA로 HEK293T 세포를 공동-형질감염시켜 유도된 유전자 활성화는 형질감염 3일 후에 피크였고, 유전자 발현은 형질감염 15-20일 후에 배경 수준으로 돌아갔다 (도 4c). 대조적으로, dCas9-VP64 및 동일한 4개의 IL1RN 또는 HBG1-표적화 sgRNA의 렌티바이러스 전달은 형질도입 20일 초과 후 동안 지속적인 유전자 활성화를 유도하였다 (도 46C, 46D). 따라서, 멀티플렉스 dCas9-VP64 전사활성인자의 단일 렌티바이러스 전달은 표적 내인성 유전자를 효율적으로 및 안정하게 상향조절하기 위한 유용한 플랫폼이다.

[0427]

실시예 22

[0428]

dCas9-KRAB - HS2 인핸서를 표적화함

[0429]

HS2 인핸서는 글로빈 유전자좌의 활성화를 위해 필요한 잘-특성화된 원위 조절 요소이다. HS2 인핸서를 표적화하는 gRNA를 갖는 dCas9-KRAB는 이러한 시스템이 K562 인간 적백혈병 세포주에서  $\gamma$ -,  $\epsilon$ - 및  $\beta$ -글로빈 발현을

억제하는지의 여부를 결정하기 위해 전달되었다 (도 54). HS2 인헨서 (서열 467)의 코어 영역을 따라 상이한 부위를 표적화하는 gRNA의 패널을 생성하였다. 표 12 참조.

<표 12>

HS2 gRNA 표적 서열

Cr#	프로토스페이서	완전한 gRNA 서열
1	gagacacacagaaatgtaac (서열 564)	gagacacacagaaatgtaacggttTAGAGCTAGAAATAGCAAGTTAAAATA AGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGG TGCTTTTTTTCC (서열 585)
2	ggfggggcactgaccccgac (서열 565)	ggfggggcactgaccccgacggttTAGAGCTAGAAATAGCAAGTTAAAATA AGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGG TGCTTTTTTTCC (서열 586)
3	ctagagtgatgactcctatc (서열 566)	ctagagtgatgactcctatcggttTAGAGCTAGAAATAGCAAGTTAAAATA GGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGT GCTTTTTTTCC (서열 587)
4	gactaaactccacccaac (서열 567)	gactaaactccacccaacggttTAGAGCTAGAAATAGCAAGTTAAAATA GGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGT GCTTTTTTTCC (서열 588)
5	aatatgtcacattctgtctc (서열 568)	aatatgtcacattctgtctcggttTAGAGCTAGAAATAGCAAGTTAAAATA GGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGT CTTTTTTTCC (서열 589)
6	ggactatgggaggtcactaa (서열 569)	ggactatgggaggtcactaaggttTAGAGCTAGAAATAGCAAGTTAAAATA AGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGG TGCTTTTTTTCC (서열 590)
7	gctcatgcttgactatggg (서열 570)	gctcatgcttgactatgggggttTAGAGCTAGAAATAGCAAGTTAAAATA GGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGT GCTTTTTTTCC (서열 591)
8	gttctggccaggccctgtc (서열 571)	gttctggccaggccctgtcggttTAGAGCTAGAAATAGCAAGTTAAAATA GGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGT GCTTTTTTTCC (서열 592)
9	agtgcacaccccccgccttc (서열 572)	agtgcacaccccccgccttcggttTAGAGCTAGAAATAGCAAGTTAAAATA GGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGT GCTTTTTTTCC (서열 593)
10	gtggggcactgaccccgaca (서열 573)	gtggggcactgaccccgacggttTAGAGCTAGAAATAGCAAGTTAAAATA AGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGG TGCTTTTTTTCC (서열 594)
11	aaccttctaagcaaaccttc (서열 574)	aaccttctaagcaaaccttcggttTAGAGCTAGAAATAGCAAGTTAAAATA GGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGT GCTTTTTTTCC (서열 595)
12	gttacacagaaccagaaggc (서열 575)	gttacacagaaccagaaggcggttTAGAGCTAGAAATAGCAAGTTAAAATA AGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGG TGCTTTTTTTCC (서열 596)
13	gaaggttacacagaaccaga (서열 576)	gaaggttacacagaaccagaggttTAGAGCTAGAAATAGCAAGTTAAAATA AGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGG TGCTTTTTTTCC (서열 597)
14	agtcatgatgatcatgctg (서열 577)	agtcatgatgatcatgctgggttTAGAGCTAGAAATAGCAAGTTAAAATA GGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGT GCTTTTTTTCC (서열 598)



15	gatgagtcagctgaggctt (서열 578)	gatgagtcagctgaggcttTTAGAGCTAGAAATAGCAAGTTAAAATAA GGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGT GCTTTTTTTC (서열 599)
16	actctaggctgagaacatct (서열 579)	actctaggctgagaacatctTTAGAGCTAGAAATAGCAAGTTAAAATAA GGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGT GCTTTTTTTC (서열 600)
17	gtccccagcaggatgcttac (서열 580)	gtccccagcaggatgcttacTTAGAGCTAGAAATAGCAAGTTAAAATAA GGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGT GCTTTTTTTC (서열 601)
18	gcctgtgaagcatcctgctg (서열 581)	gcctgtgaagcatcctgctgTTAGAGCTAGAAATAGCAAGTTAAAATAA GGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGT GCTTTTTTTC (서열 602)
19	cagggcagatggcaaaaaa (서열 582)	cagggcagatggcaaaaaaTTAGAGCTAGAAATAGCAAGTTAAAATAA AGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGG TGCTTTTTTTC (서열 603)
20	gaggtggagtttagtcagg (서열 583)	gaggtggagtttagtcaggTTAGAGCTAGAAATAGCAAGTTAAAATAA GGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGT GCTTTTTTTC (서열 604)
21	aaacggcatcataaagaaaa (서열 584)	aaacggcatcataaagaaaaTTAGAGCTAGAAATAGCAAGTTAAAATAA AGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGG TGCTTTTTTTC (서열 605)

[0433]

[0434]

CRISPR/dCas9에 의한 글로빈 유전자좌에서의 단일 gRNA 스크리닝. dCas9 및 dCas9-KRAB 이펙터를 U6-sgRNA 발현을 코딩하는 플라스미드 5  $\mu$ g으로 전기천공하기 5-8일 전에 렌티바이러스로 전달하였다 (도 55a). gRNA로 전기천공되지 않은 세포 (gRNA 없음) 및 상이한 유전자좌를 표적화하는 gRNA로 처리된 세포 (IL1RN)는 대조군으로서 포함되었다. 다중 gRNA는 형질감염 3일 후에 검정하였을 때  $\epsilon$ -,  $\gamma$ - 및  $\beta$ -글로빈 유전자의 강력한 억제력을 가져왔으며, 최대 80% 녹다운을 달성하였다 (도 55b, 55c, 55d). dCas9 또는 dCas9-KRAB와 함께 gRNA를 발현시키는 것은 글로빈 유전자좌에서 유전자 발현을 억제하였다. 일반적으로, dCas9-KRAB를 사용한 처리는 dCas9 단독과 비교하여 소정의 gRNA의 보다 강한 억제를 유발하였으며, 이는 억제를 증진시키는 이질염색질 인자를 동원함에 있어서 KRAB 도메인의 중요한 역할을 시사한다. 달성된 억제 수준은 단지 dCas9-KRAB 처리 세포에서만 형질감염에 의해 전달된 gRNA 플라스미드의 양에 의존하였다 (도 56a, 56b, 56c). Cr4 gRNA 플라스미드의 용량을 최대 10  $\mu$ g으로 증가시키는 것은 dCas9-KRAB-처리 세포에서 글로빈 유전자의 침묵화 수준을 증가시켰다.

[0435]

dCas9-KRAB에 의한 글로빈 유전자의 안정한 침묵화. dCas9/dCas9-KRAB를 K562에서 렌티바이러스로 단일 gRNA와 함께 공-발현시켰다 (도 57a). 렌티바이러스로 처리되지 않은 세포 (NT), gRNA 없이 dCas9/dCas9-KRAB로 처리된 세포 (gRNA 없음), 및 dCas9/dCas9-KRAB와 상이한 유전자좌를 표적화하는 gRNA로 처리된 세포 (IL1RN)는 대조군으로서 포함되었다. 렌티바이러스로 처리된 세포를 제4일 내지 제7일에 선택하였다. 다중 gRNA는 형질도입 7일 후에 검정하였을 때  $\epsilon$ -,  $\gamma$ - 및  $\beta$ -글로빈 유전자의 강력한 전사 억제를 가져왔으며, 최대 95% 녹다운을 달성하였다 (도 57b, 57c, 57d).  $\epsilon$ -글로빈의 발현은 HS2 인핸서를 표적화하는 반응 gRNA에서 대부분 침묵화되었다. dCas9-KRAB와 gRNA로의 처리는 dCas9와 gRNA로의 처리보다 훨씬 더 많은 억제를 유발하였다.

[0436]

이들 발견은 gRNA에 의해 HS2 인핸서를 표적화하는 dCas9-KRAB가 원위 글로빈 유전자의 강력한 억제를 가져온다는 것을 입증한다. 이것은 CRISPR/Cas9 시스템에 의한 포유동물 세포에서의 원위 조절 요소의 표적화된 후성적 제어의 제1 예이다. 인핸서는 발달 및 질환을 조절하고, 본 개시내용은 인핸서 기능을 프로빙 및 제어하는 방법을 제공하며, 국부 염색질 접근성 및 계놈-전반 발현에 대한 dCas9-KRAB의 효과를 결정하는데 사용될 수 있다.

[0437]

실시예 23

[0438]

dCas9-p300

[0439]

dCas9-p300 융합 단백질을 설계하고, dCas9-VP64 융합 단백질과 비교하였다 (도 59 참조). dCas9 구축의 아미노산 구축물이 도 61a-61c에 제시되어 있다. 인간 배아 신장 조직 배양 세포주 HEK293T (ATCC; CRL-11268)로부터의 세포를 리포펙타민 2000 형질감염 시약 (라이프 테크놀로지스)으로의 형질감염 1일 전에 24-웰 조직 배양 접시에서 웰당 1.5e5개 세포의 밀도로 시딩하였다. 24시간 후에 세포를 1  $\mu$ L 리포펙타민 2000, 375 ng dCas9 발현 구축물 (각각 dCas9, dCas9VP64 또는 dCas9p300), 및 125 ng의 풀링된 gRNA 발현 플라스미드 (4개 각각 등몰비)로 형질감염시켰다. 표 13은 gRNA 정보를 제시한다.

[0440] <표 13>

[0441] gRNA 정보

표적 위치	프로토스페이서 서열 (5'-3')	개념 위치 (GRCh38 주요 조립)	참고문헌
IL1RN 프로모터	TGTACTCTCTGAGGTGCTC (서열 606)	Chr2: 113117865- 113117883	Perez-Pinera et al., Nat Methods, 2013
IL1RN 프로모터	ACGCAGATAAGAACCAGTT (서열 607)	Chr2: 113117714- 113117732	Perez-Pinera et al., Nat Methods, 2013
IL1RN 프로모터	CATCAAGTCAGCCATCAGC (서열 608)	Chr2: 113117781- 113117799	Perez-Pinera et al., Nat Methods, 2013
IL1RN 프로모터	GAGTCACCCCTCTGGAAAC (서열 609)	Chr2: 113117749- 113117767	Perez-Pinera et al., Nat Methods, 2013
MyoD 프로모터	CCTGGGCTCCGGGGCGTTT (서열 610)	Chr11: 17719509- 17719527	Perez-Pinera et al., Nat Methods, 2013
MyoD 프로모터	GGCCCCTGCGGCCACCCCG (서열 611)	Chr11: 17719422- 17719440	Perez-Pinera et al., Nat Methods, 2013
MyoD 프로모터	CTCCCTCCCTGCCCCGGTAG (서열 612)	Chr11: 17719350- 17719368	Perez-Pinera et al., Nat Methods, 2013
MyoD 프로모터	AGGTTTGGAAGGGCGTGC (서열 613)	Chr11: 17719290- 17719308	Perez-Pinera et al., Nat Methods, 2013
Oct4 프로모터	ACTCCACTGCACTCCAGTCT (서열 614)	Chr6: 31170953- 31170934	Hu et al., Nucleic Acids Res, 2014
Oct4 프로모터	TCTGTGGGGGACCTGCACTG (서열 615)	Chr6: 31170885- 31170866	Hu et al., Nucleic Acids Res, 2014
Oct4 프로모터	GGGGCGCCAGTTGTGTCTCC (서열 616)	Chr6: 31170855- 31170836	Hu et al., Nucleic Acids Res, 2014
Oct4 프로모터	ACACCATTGCCACCACCAT (서열 617)	Chr6: 31170816- 31170797	Hu et al., Nucleic Acids Res, 2014

[0442]

[0443] 형질감염 3일 후 세포를 수거하고, mRNA 발현에 대해 RT-QPCR에 의해 검정하였다. RT-QPCR 프라이머 서열은 표 14에 열거되어 있다.

[0444] <표 14>

[0445] RT-QPCR 프라이머

<u>프라이머</u> <u>표적</u>	<u>프라이머 서열</u> <u>(5'-3')</u>	<u>참고문헌</u>
GAPDH 정방향	CAATGACCCCTTCATTGACC (서열 618)	Perez-Pinera et al., Nat Methods, 2013
GAPDH 역방향	TTGATTTTGGAGGGATCTCG (서열 619)	Perez-Pinera et al., Nat Methods, 2013
IL1RN 정방향	GGAATCCATGGAGGGAAGAT (서열 620)	Perez-Pinera et al., Nat Methods, 2013
IL1RN 역방향	TGTTCTCGCTCAGGTCAGTG (서열 621)	Perez-Pinera et al., Nat Methods, 2013
MyoD 정방향	CTCTCTGCTCCTTTGCCACA (서열 622)	Perez-Pinera et al., Nat Methods, 2013
MyoD 역방향	GTGCTCTTCGGGTTTCAGGA (서열 623)	Perez-Pinera et al., Nat Methods, 2013
Oct4 정방향	CGAAAGAGAAAGCGAACCAGTATCGAGAAC (서열 624)	Hu et al., Nucleic Acids Res, 2014
Oct4 역방향	CGTTGTGCATAGTCGCTGCTTGATCGC (서열 625)	Hu et al., Nucleic Acids Res, 2014

[0446]

[0447] RT-QPCR을  $\Delta\Delta C_t$  방법을 사용하여 GAPDH 발현에 대해 정규화하였다. 결과는 DNA 형질감염 없이 단지 리포펙타민으로 처리된 세포에 대한 관심 유전자 발현의 배수-증가로서 표시하였다 ("DNA 없음") (도 60a-60c).

[0448] 도 62는 p300 HAT 도메인에서의 잔기를 돌연변이시키는 것이 그의 유전자 발현 활성화 능력의 상실을 유발한다는 것을 보여준다. 도 63은 dCas-9-VP64에 의해 제시된 바와 같이, 다중 gRNA가 dCas-9-p300와 상승작용적으로 작용한다는 것일 보여준다.

[0449] 실시예 24

[0450] 도 66은 디스트로핀 유전자에 상이한 결실을 보유하는 인간 DMD 환자로부터 유래된 골격 근모세포주에서 디스트로핀의 Dp427m 골격근 이소형의 5'UTR에서의 미니디스트로핀의 TALEN 매개 통합을 보여준다. DMD 환자 세포를 5'UTR 유전자좌에서의 TALEN 쌍 활성 및 미니디스트로핀 유전자를 보유하는 공여자 주형을 코딩하는 구축물로 전기천공하였다. 도 66(a)는 미니디스트로핀을 5'UTR 내로 통합시키는 방법의 개략도를 보여준다. 도 66(b)는 히그로마이신-내성 클론 세포주를 단리하고, 도 66(a)에 제시된 프라이머를 사용하여 5'UTR에서의 성공적인 부위-특이적 통합에 대해 PCR에 의해 스크리닝한 것을 보여준다. 별표는 도 66(c)에서 추가의 분석을 위해 선택된 클론을 나타낸다. 도 66(c)는 통합 사건이 검출된 클론 단리된 DMD 근모세포를 6일 동안 분화시키고, 미니디스트로핀의 C 말단에 융합된 HA 태그의 발현에 대해 평가한 것을 보여준다.

[0451] 상기한 상세한 설명 및 첨부된 실시예는 단지 예시적이며, 첨부된 청구범위 및 그의 등가물에 의해 단독으로 정의되는 본 발명의 범주에 대한 제한으로서 간주되지 않는 것으로 이해된다.

[0452] 개시된 실시양태에 대한 다양한 변화 및 변형은 통상의 기술자에게 명백할 것이다. 비제한적으로 본 발명의 화학 구조, 치환기, 유도체, 중간체, 합성, 조성물, 제제 또는 사용 방법과 관련된 것을 포함하는 이러한 변화 및 변형은 그의 취지 및 범주에서 벗어남 없이 이루어질 수 있다.

[0453] 부록

[0454] hCas9-T2A-GFP DNA 서열: 서열 145 (SpCas9 인간 최적화된 서열, HA 태그, T2A 펩티드, eGFP 서열)

gccaccATGGACAAGAGTACTCCATTGGGCTCGATATCGGCACAAACAGCGTCGGCTGGGCCGTCATTACGGACGAGTA  
CAAGGTGCCGAGCAAAAAATTCAAAGTCTTGGGCAATACCGATCGCCACAGCATAAAGAAAGAACCTCATTGGCGCCCTCC  
TGTTTCGACTCCGGGGAGACGGCCGAAGCCACGCGGCTCAAAAGAACAGCAGCGCGCAGATATACCCGCAGAAAGAAATCGG  
ATCTGCTACCTGCAGGAGATCTTTAGTAATGAGATGGCTAAGGTGGATGACTCTTTCTTCCATAGGCTGGAGGAGTCCTT  
TTTGGTGGAGGAGGATAAAAGCAGCAGCGCCACCCAACTCTTGGCAATATCGTGGACGAGGTGGCGTACCATGAAAGT  
ACCAACCATATATCATCTGAGGAAGAGCTTGTAGACAGTACTGATAAGGCTGACTTGCAGTTGATCTATCTCGCGCTG  
GCGCATATGATCAAAATTCGGGGACACTTCTCATCGAGGGGGACCTGAACCCAGACAAACAGCGATGTCGACAACTCTT  
TATCCAACTGGTTCAGACTTACAATCAGCTTTTCGAAAGAAACCCGATCAACGCATCCGGAGTTGACGCCAAAGCAATCC  
TGAGCGCTAGGCTGTCCAAATCCCGCGGCTCGAAAACCTCATCGCACAGCTCCCTGGGGAGAAAGAACGGCCTGTTT  
GGTAATCTTATCGCCCTGTCACTCGGGCTGACCCCACTTTAAATCTAACTTCGACCTGGCCGAAGATGCCAAGCTTCA  
ACTGAGCAAAAGACACCTACGATGATGATCTCGACAATCTGCTGGCCAGATCGGCGACCACTACCGAGACCTTTTTTGG  
CGGCAAAAGAACTGTACAGCGCCATCTGCTGAGTGATATTCTGCGAGTGAACACGGAGATCACCAGAGCTCCGCTGAGC  
GCTAGTATGATCAAGCGCTATGATGAGCACCACCAAGACTTGACTTGTCTGAAGGCCCTTGTGACACGCAACTGCCTGA  
GAAGTACAAGGAAATTTCTTCGATCAGCTCAAAAAATGGCTACGCCGATACATTGACGGCGGAGCAAGCCAGGAGGAAT  
TTTACAAATTTATTAAAGCCCATCTTGGAAAAATGGACGGCACCGAGGAGCTGCTGGTAAAGCTTAACAGAGAAAGATCTG  
TTGCCCAACAGCGCACTTTGCAAAATGGAAGCATCCCCCACCAGATTCACCTGGCGCAACTGCACGCTATCTCAGGCG  
GCAAGAGGATTTCTACCCCTTTTGAAGATAACAGGGAAAAGATTGAGAAAACTCTCACATTTCCGATACCTTACTATG  
TAGGCCCTCGCCCGGGGAAATTCAGATTCTGCTGGATGACTCGCAATCAGAAGAGACCATCTCCCTGGAACTTC  
GAGGAAGTCGTGGATTAAGGGGCTCTGCCAGTCTTCATCGAAAGGATGACTAACTTTGATAAAATCTGCCTAACGA  
AAAGTGTCTTCTAACAACCTCTCTGCTGTACGAGTACTTCAAGTTTATAACGAGCTCACCAGGTCAAATACGTCACAG  
AAGGGATGAGAAAGCCAGCATCTCTGCTGGAGAGCAGAAGAAAGCTATCGTGGACCTCCTCTTCAAGACGAACCGGAAA  
GTTACCGGTGAACAGCTCAAGAGAGCATTTTCAAAAAGATTGAATGTTTCGACTCTGTTGAAATCAGCGGAGTGGAGGA  
TCGCTTCAACGCATCCCTGGGAACGTATCACGATCTCTGAAATCATTAAAGACAAGGACTTCTGGACAATGAGGAGA  
ACGAGGACATCTTGAGGACATTGTCTCACCTTACGTTGTTTGAAGATAGGGAGATGATTGAAGAACGCTTGAAACT  
TACGCTCATCTCTCGACGACAAAGTCATGAAACAGCTCAAGAGGCGCGGATATACAGGATGGGGCGGCTGTCAAGAAA  
ACTGATCAATGGGATCCGAGACAAGCAGAGTGGAAAGACAACTCTGGATTTTCTTAAGTCCGATGGATTGGCCAAACCGGA  
ACTTCATGCACTGATCCATGATGACTCTCTCACCTTTAAGGAGGACATCCAGAAAGCAAGTTTCTGGCCAGGGGGAC  
AGTCTTACAGGACATCGCTAACTTTCAGGATAGCCAGCTATCAAAAGGGAACTAGTCAGACCGTTAAGGTCTGTGA  
TGACTCGTCAAGTAAATGGGAAGGCATAAGCCCGAGAAATATCGTTATCGAGATGGCCCGAGAGAACCAAACTACCCAGA  
AGGGACAGAAAGACAGTAGGGAAGGATGAAGAGGATTAAGAGGGTATAAAAGAACTGGGGTCCCAATCTTAAAGGAA  
CACCCAGTTGAAAACACCCAGCTTCAGAAATGAGAAGCTCTACCTGTACTACCTGCAGAACCGCAGGGACATGTACGTGGA  
TCAGGAACCTGGACATCAATCGGCTCTCCGACTACGACGTGGATCATATCGTGCCCGAGTCTTTTCAAAAGATGATTCTA  
TTGATAATAAGTGTGACAAAGATCCGATAAAAAATAGAGGGAAGAGTGATAACGTCCCTCAGAAAGAGTTGTCAAGAAA  
ATGAAAAATTTATGGCGGAGCTGCTGAACGCCAACTGATCACACAACCGGAAGTTTCGATAATCTGACTAAGGCTGAACG  
AGGTGGCCTGTCTGAGTTGGATAAAGCCGGCTTCATCAAAAGGCAGCTTGTGAGACACGCCAGATCACCAGCAGCTGG  
CCCAATTTCTCGATTACGCGATGAACACCAAGTACGATGAAAATGACAACTGATTTCGAGAGGTGAAAGTTATTACTCTG  
AAGTCTAAGCTGGTCTCAGATTTTCAAAAGGACTTTTCAAGTTTATAAGGTGAGAGAGATCAACAATTACCACCATGCGCA  
TGATGCCCTACCTGAATGCAGTGGTAGGACCTGCACTTATCAAAAAATATCCCAAGCTTGAATCTGAATTTGTTTACGGAG  
ACTATAAAGTGTACGATGTTAGGAAAATGATCGCAAGTCTGAGCAGGAAATAGGCAAGGCCACCGCTAAGTACTTCTTT  
TACAGCAATATTATGAATTTTTTCAAGACCGAGATTACACTGGCCAAATGGAGAGATTGGAAGCGACCACTTATCGAAAC  
AAACGGAGAAACAGGAGAAATCGTGTGGGACAAGGGTAGGGATTTCGCGACAGTCCGGAAGGTCTGTCCATGCCGAGG  
TGAACATCGTTAAAAAGACCGAAGTACAGACCGGAGGCTTCTCAAGGAAAGTATCCTCCCGAAAGGAACAGCGACAAAG  
CTGATCGCAGCAGCAAAAAGATTGGGACCCCAAGAAATACGGCGGATTCGATTCTCTACAGTCTGTTACAGTGTACTGGT  
TGTGGCCAAAGTGGGAAAGGGAAGTCTAAAAAATCAAAAGCGTCAAGGAAGTCTGGGCATCACAATCATGGAGCGAT  
CAAGCTTCGAAAAAACCCCATCGACTTTCTCGAGGCGAAAGGATATAAGAGGTCAAAAAGACCTCATCTAAAGCTT  
CCCAAGTACTCTCTTTGAGCTTGAACGCGCCGAAACGAATGCTCGCTAGTGGCGGCGAGCTGCAGAAAGGTAACGA  
GCTGGCACTGCCCTCTAAATACGTTAATTTCTTGTATCTGGCCAGCCACTATGAAAAGCTCAAAGGGTCTCCCGAAGATA  
ATGAGCAGAAAGCAGCTGTTCTGGAAACACACAACTACCTTGTATGAGATCATCGAGCAAAATAGCGAATTTCTCCAAA  
AGAGTGATCCTCGCGCAGCTAACTCGATAAGGTGCTTTCTGCTTACAATAAGCACAGGGATAAGCCCATCAGGGAGCA  
GGCAGAAAACATTTATCCACTTGTATTCTGACCACTTGGGCGCGCTGCAGCCTTCAAGTACTTCGACACCACCATAG  
ACAGAAAGCGGTACACCTCTACAAAGGAGGTCTGGACGCCACACTGATTTCATCAGTCAATTACGGGGCTCTATGAAACA  
AGAATCGACCTCTCTCAGCTCGGTGGAGACAGCAGGGCTGACCTCAAGAGAGAGAGAGTCTGCTAGCGAGGGCAGAGG  
AAGTCTTCTAACATCGGTGACGTGGAGGAGATCCCGGCCCTGgtacGgtgagcaaggcgaggagctgttcacccggg  
tggtgcccatcctggtcgagctggaagcgagctaaacggccacaagttcagcggtgcccggcgaggcgaggcgatgcc  
acctacggcaagctgacctgaagttcactgacccaccggcaagctgcccgtgcccgtgcccaccctcgtagccacct  
gacctacggcggtgacgtgcttcagccgctaccccgaccacatagaagcagcagcacttcttcaagtccgccatgcccgaag  
gctacgtccaggagcgaccactcttcttcaaggacgagcgaactacaagaccccgccgaggtgaagttcagggcgac  
acctggtgaaccgcatcgagctgaagggaactgaacttcaaggaggacggcaactcctggggcacaagctggagtacaa  
ctacaacagccacaacgctctatactatgcccgaagcagaagaacggcatcaaggtgaacttcaagatccgccacaaca  
tcgaggacggcagctgacgtcgccgaccactaccagcagaacacccccatcgggcagcgcccgctgctgctgcccagac  
aacactacctgagcaccagctccgcccctgagcaagaccccaacgagaagcgcatcactggtcctgctggaggttcgt  
gaccgcccggggtacactctcgcatggacgagctgtacaagaccggtTAG

[0455]

[0456] AAV/Rosa26 구축물 (서열 456)

gggggggggggggggggttgccactccctctctgcgcgctcgctcgctcactgagggcgggcgaccaaaggtcgcccg  
acgcccgggcttggccggggcgccctcagtgagcgagcgagcgcgagagggagtgggcaactccatcactaggggtt  
cctagatctgaattcggtaccoggttacataaacttacggtaaatggcccgctggctgaccgccaacgacccccgcca  
tgacgtcaaatagcgtatgttcccatagtaacgccaatagggaacttccattgacgtcaatgggtggagattttacg  
taaaactgccacttggcagtaacatcaagtgtatcatatgccaagtacgccccctattgacgtcaatgacggtaaatggcc  
cgctggcattatgcccagtagcatgaccttatgggactttccacttggcagtagcatctacgtattagtcacgtattat  
ccatggtgatgcggttttggcagtagcatcaatggcggtggatagcggtttgactcaggggatttccaagtctccacccc  
attgacgtcaatgggagtttggtttggccacaaaatcaacgggacttccaagtctgtaacaactccgccccattgac  
gcaaatggcggttagcggtgtacgggtgggaggtctatataagcagagctcggttagtgaaacgtcagatcgctggagac  
gcatccacgctgttttgacctccatagaagacacgggagccgacccgactcagaggatccggtactcagag  
gaactgaaaaacagaaagtttaactggttaagtttagtcttttcttttatttcaaggtcccggtacgggtgggtgca  
aatcaagaactgctcctcagtggtatgttgctttacttctaggcctgtacggaagtgttacttctgctctaaaagctgc  
ggaattgtacccgcgccgggataccacgggtTGGCTAGCgtctataggccccccccTTGGTGGAAATTCGCCATGAGGTC  
TGACTACAAAGACCATGACGGTGTATTATAAAGATCATGACATCGATTACAGGATGACGATGACAAGATGGCCCCAAGA  
AGAAgaggaaggtggcgctcgaGCCCGGAGAAAAACCTACAAAGTGCCTGAGTGCGGGAAATCATTTCTCCGACCTGGG  
GCGCTCGTCCGGCACCAAGGACGCATACAGGGGAAAAGCCGTATAAGTGCCCCGAGTGTGGAAAGAGCTTCTCGCAGAG  
AGCCCACTTGAACGACACCAAGACACACACACTGGTGAGAAAACCTATAAGTGTCCAGAGTGCGGCAATCGTTTGA  
GATCCGATGACTTGGTGCGCCACCGAGCGACACACACGGGTGAAAAGCCCTACAAATGCCCGGAGTGTGGGAAGTCGTTT  
TCAAGGTCGGATCATCTGACTACCATCAGCGCACCCATACGGGAGCggcgcgccgcgccctGGTGAAGAGCGAGCTGGA  
GGAGGAAGAGTCCGAGCTCGCGCACCAAGCTGAAGTACGTGCCCCACGAGTACATCGAGCTGATCGAGATCGCCAGGAAC  
CCACCCAGGACCGCATCTCTGGAGATGAAGGTGATGGAGTTCTTATGAAGGTGTACGGCTACAGGGGAGAGCACCTGGGCG  
GGAGGCAAAAGCTGACGGCCACGCTGACACGGCTGAACCGCAAAACCACTGCAATGGCGCGGTGCTGAGCGTGGAGGAGCT  
GCTGATCGGCGCGAGATGATCAAAGCCGGCACCTTGACACTGGAAGGAGGTGCGCGCGCAAGTTCAACAAACGGCGAGATCA  
ACTTCGAGGGCAGAGGAAGTCTTCTAATCATGCGGTGACGTGGAGGAGAATCCCGGCCCTAgatcTGACTACAAAGACCAT  
GAGGCTGATTATAGGATCATGACATCTGATTACAAAGGATGACGATGACAAGATGGCCCCAAGAAAGAAgaggaaggtggg  
cctcgagccGGGAGAGAAGCCCTACAAAGTGTCCGAATGTGGAAGAGCTTCTCACAGTGGGGGACCTTTCGGCGCCACC  
AGCGCACACATACTGGTGAAGAGCCGTATAAGTGTCCAGAATGTGCAAAATCATTTCTCCACATCAGGAGCCCTGGTCAGG  
CACCGCAAGTCCGAGCTGAGAGCCCTATAAGTGTCCCGCAATGCGGGAAGTCCTTTTCGACAGAGAGCCCACTTGA  
GAGGCAACAGAGAGCCATACGGGGGAGAAACCTTACAAAGTGCCTGAATGCGGAAAGTCTGTTCTCGACCATCTGGATC  
TCATCAGACATCAGAGAAGCGCACACTGGAGAGAAACCTTCAAAATGTCCGAGTGTGGGAAGTCGTTTAGCCGAAAGGAC  
AATCTCAAAACCATCAACGGACACACACGGGTGAAAAACCATACAAATGCCCGGAGTGGCGCAAAATCGTTTTCCTCAACT  
TGGCGACTTTCGGGGCACCAACGACGACGATCTGAGCGGCGCCCGcgcgccctGGTGAAGAGCGAGCTGGAGGAGAAGA  
AGTCCGAGTGGCGGCACAGCTGAAGTACGTGCCCCACGAGTACATCGAGTGTGATCGAGATCGCCAGGAACCCACCCAG  
GACCGCATCTCGGAGATGAAGGTGATGGAGTTCTTATGAAGGTGTACGGCTACAGGGGAGAGCACCTGGCGGAAGCAG  
AAAGCCTGACGGCGCCATCTATACAGTGGGCGAGCCCATCGATTACGGCGTGTGCTGGACACAAAGGCTACAGCGGCG  
GCTCAACTTGGCTATCGCCACGAGCCGACGAGATGGAGAGATACGTGGAGGAGAACAGACACGGGATAGCACCTCAAC  
CCCAACGAGTGGTGGAAAGGTGATACCTAGCAGCGTGACCGAGTTCAAGTTCCTGTTGCTGAGCGGCGCACTTCAAGGGCAA  
CTACAAGGCCAGCTGACCAAGGTGAACCATACCAACTGCAATGGCGCCGTGCTGAGCGTGGAGGAGCTGCTGATCG  
GCGGCGAGATGATCAAGAGCCGACCTGACACTGGAGGAGGTGCGGCGCAAGTTCACAAACGGCGAGATCAACTTCTGA  
TTAATTAACTAATCTAGAGTcgaactagagctcgctgatcagcctcgactgtgccttctagttgccagccatctgttgttt  
gccccctcccccgctgcttcttgacctggaaggtgccactcccactgtcctttctctaataaaatgaggaatgtcatcg  
cattgtctgagtagtgctcattctattctgggggtgggtgggacagcagaagggggaggttgggaagacaatag  
cagggcatgctggggagagatctaggaacccctagtgtgaggttggccactccctctctgcgcgctcgctcgctcactga  
ggcgcccgggcaaaagcccgccgtcgggcgaccttgggtcgcccgccctcagtgagcgagcgagcgcgagagagggag  
tgcccaacccccccccccccctcgagcccgactgcattaatgaatcgcccaacgcgcggggagaggggttgggt  
attggcgctcttccgcttctcgtcactgactcgctcgctcggtcggtcggtcggtcggtcggtcggtcggtcggtcggt  
aaaggcggttaatacgggttatccacagatcaggggataaacgcaggaagaacaatgtgagcaaaagccagcaaaaggcca  
ggaaccgtaaaaaggcgcggtgtgtggcggtttttccataggctccgccccctgacgagcatcacaataatcgacgctca  
agtcagaggttggcgaaaccccgacagagctataaagataccaggggttcccccgtggaagctccctcgctcgctctcctgt  
tcgcgacccctcgcgcttaccggatacctgtccgcttctcctctcgggaaagcggtggcgctttctcaatgctcagcggtta  
ggtatctcaggttcggtgtaggttcgttcgctccaagctgggtgtgtgcacgaaccccccggttcagcccgaccgctgccc  
ttatccggtaactatcgcttctgagtcacacccggtaagacacgacttatcgccactggcagcagccactggtaacaggat  
tagcagagcgaggtatgtgagcggtgtgtacagagttcttgaaagtggtggcgtaactacggtacactagaaggacagtat  
ttggtatctgctcgtctgctgaagcagttaccttcggaaaaagagttggtagctcttgatccggcaacaaacacccgct  
ggtagcggtggttttttggttgcaagcagcagattacgcgcagaaaaaaaggatctcaagaagatcctttgatctttc  
tacggggtcagcagctcagtcagcagcaaaactcaggttaagggattttggtcactgagattatcaaaaaggatcttccact  
agatccttttaaatataaatgaagtttttaaatcaatctaaagtataatgagtaaaacttggtctgacaggttaccatgc  
ttaatcagtagggcacctatctcagcgatctgtctatttctgctcatccatagttgctgactccccgctcggtgtagataac

[0457]

tacgatagggagggcttaccatctggccccagtgctgcaatgataccgagagaccacgctcaccggctccagatttat  
cagcaataaacacagccagcggaaggccgagcgagagagtggtcctgcaactttatccgctccatccaggtctattaat  
tgttgcggggaagctagagtagtgctgcaagtttgctgcaactgttggtgcaattgctacagggatcggtgt  
gtcagcgctcgctggtttggtatggtctcatcagctccggttcccaacgatcaaggcgagttacatgatcccccatggtgt  
gcaaaaaagcggttagctccttcggtcctccgatcggttgtcagaagtaagttggcgcgaggtttatcactcatggttatg  
gcagcactgcataattctcttactgtcatgccatccgttaagatgcttttctgtgactggtgagtagtactcaaccaagtcat  
ctgagaatagtgatgctggcgagcaggttgctcttgcggcgctcaatacgggataataccgcgccacatagcagaactt  
taaaagtgctcatcatttgaaaaacgtttcttcggggcgaaaaactctcaaggatcttaccgctgttgagatccaggttcgatg  
taaccactcgtgcacccaactgatcttcagcatctttactttccaccgctttctgggtgagcaaaaaacaggaaggca  
aaatgcccgaaaaaagggaataaggggcgacaggaatagttgaaatactcactcttctcttcaatattatgaaagca  
ttatcaggggtattgtctcactgagcgagatacatattgaaatgtatttagaaaaataaaacaaataggggttcgcgcgaca  
tttccccgaaaagtgccacctgacgtcttaagaaacattattatcatgacattaacctataaaaataggcggtatcacgag  
gccccttctgctcgcgcgtttcggtgatgacgggtgaaaaactctgacacatgcagctcccgagagcggtcacagcttgtc  
tgtaacgggatgcggggagcagacaagcccgctcagggcggtgtggcggtgtggcggtgtggcggtgtggcggtgtggcggt  
gcggcatcagagcagattgtactgagagtgacacatgctggtgtgaaataccgcacagatcggtgaaggagaaaaatccg  
catcaggaatgtgaaagcttaattttgttaaaattcggttaaaattttgttaaaatcagctcatttttaaccaata  
ggcggaactcgcaaaatcccttataaatcaaaagaaatagacagagtaggggttgaggtgtgttccagtttggaacaaga  
gtccatatttaagaacgttgactccaacgtcaaaaggcgcaaaaaacgctctacggggcgatggccccatcgtgaaacca  
tcacctaatacgaatttttgggtcgaggtgcgtaaaagcactaaatcggaaacccataaaggagagcccccgatttagagc  
ttgacggggaaagccggcgcaacgtggcgagaaagggaaggaaggaagcgaagggcggttagggcgctggcaagtg  
tagcggtcagcgtgcggttaaacccacaccccgcgcgcttaatgcccgtacagggcggtcgcgccattcgccattc  
aggctacgcaactgttgggaaggcgatcggtgcgggctcttctgctattacgcaagctgggtgca

[0458]



[0459] HS2 인헨서 표적 서열 (서열 467)

taagcttcagtttttctttagttctgtttacattttctgtgtgtctccatttagtgacctcccatagtcacagcatgagcag  
ttctgtggccaggccctgtcggggtcagtgccccaccccgccctctgtgtgttaaccttcttaagcaaaccttctgtgc  
tcaagcacagcaatgctgagtcatagtgagtcatagtctgaggcttaggggtgtgtgtgcccagatgttctcagcctagagtgat  
gactcctatctgggttccccagcaggatgctttacagggcagatggcaaaaaaaggagaagctgaccacctgactaaaact  
ccacctcaaacggcatcataaagaaaatggatgcctgagacagaatgtgacatatcttagaatatatt

[0460]

[0461] dSpCas9-KRAB 서열 (서열 466)

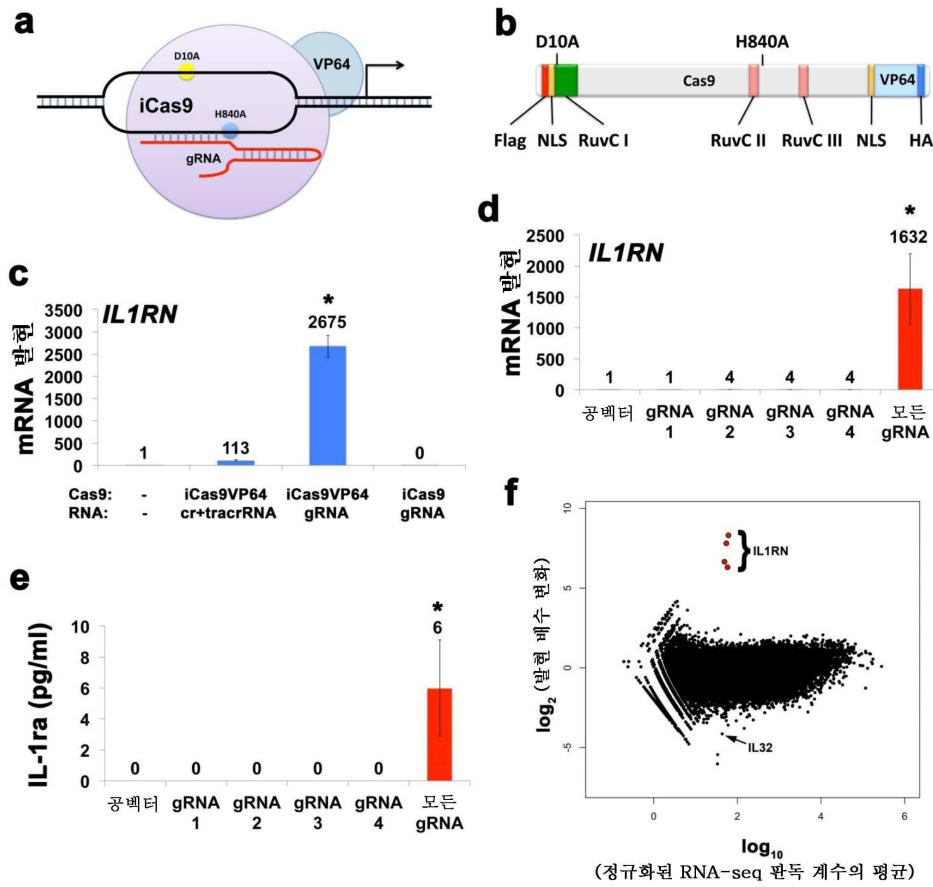
ATGGACTACAAAGACCATGACGGTGATTATAAAGATCATGACATCGATTACAAGGATGACGATGACAAAGatggccccc  
gaagaagagggaaggtgggcccgggaATGGACAGAAGTACTCCATTGGGCTCGCCATCGGCACAAACAGCGTCGGCTGGG  
CCGTCAATTACGGACGAGTACAAGGTGCCGAGCAAAAATTCAAAGTTCTGGGCAATACCGATCGCCACAGCATAAAGAAG  
AACCTCATTGGCGCCCTCTGTTCGACTCCGGGGAAACCGCCGAAGCCACGCGGCTCAAAAGAACAGCACGGCGCAGATA  
TACCCGCAGAAAGAAATCGGATCTGCTACCTgcaGGAGATCTTTAGTAATGAGATGGCTAAGGTGGATGACTCTTTCTTCC  
ATAGGCTGGAGGAGTCTTTTGGTGGAGGAGGATAAAAAGCACGAGCGCCACCCAATCTTTGGCAATATCTGGACGAG  
GTGGCGTACCATGAAAGTACCCAACCATATATCATCTGAGGAAGAAGCTTGTAGACAGTACTGATAAGGCTGACTTGGC  
GTTGATCTATCTCGCGCTGGCGCATATGATCAAAATTCGGGGGACACTTCCTCATCGAGGGGGACCTGAACCCAGACAA  
CGCATGTCGACAAACTCTTTATCCAATGTTTCAGACTTACAATCAGCTTTTCGAAGAGAACCCGATCAACGCATCCGGA  
GTTGACGCCAAAGCAATCTGAGCGCTAGGCTGTCCAATCCCGGCGGCTCGAAACCTCATCGCACAGCTCCCTGGGGA  
GAAGAAGAACGGCCTGTTTGGTAATCTTATCGCCCTGTCACTCGGGCTGACCCCAACTTTAAATCTAATCTCGACCTGG  
CCGAAGATGCCAAGCTTCAACTGAGCAAGAGACACCTACGATGATGATCTCGACAATCTGCTGGCCAGATCGCGCAGCAG  
TACGCAGACCTTTTGGCGGCAAGAACCTGTCAAGCGCATTTCTGCTGAGTGATATTCTGCGAGTGAAACCGGAGAT  
CACCAAGCTCCGCTGAGCGCTAGTATGATCAAGCGCTATGATGAGCACCACCAAGACTTGACTTTGCTGAAGGCCCTTG  
TCAGACAGCAACTGCCTGAGAAGTACAAGGAATTTTCTCGATCAGTCTAAAAATGGCTACGCCGGATACATTGACGGC  
GGAGCAAGCCAGGAGGAATTTTCAAAATTTAAGCCCATCTTGGAAAAATGGACGGCACCGGAGAGCTGCTGGTAA  
GCTTAACAGAGAAGATCTGTTGGCGCAACAGCGCACTTCGACAAATGGAAGCATCCCCACCAGATTACCTGGGCGAAC  
TGCACGCTATCCTCAGGCGGCAAGAGGATTTTACCCCTTTTGAAGATAACAGGGAAGAGTTGAGAAAACTCCTCACA  
TTTCGGATACCTTACTATGTAGGCCCTTCGCCCGGGGAAATTCAGATTTCGCGTGGATGACTCGCAAAATCAGAAAGAG  
CATCACTCCCTGGAATCTCGAGGAAGTCTGTGATAAGGGGGCCTCTGCCAGTCTTATCGAAAGGATGACTAACTTTG  
ATAAAAAATCTGCCTAACGAAAGGTGCTTCTTAAACACTCTCTGCTGTACGAGTACTTCACAGTTTATAACGAGCTCACC  
AAGGTCAAATACGTACAGAAAGGATGAGAAAGCCAGCATTCCTGTCTGGAGAGCAGAAGAAAGCTATCGTGGACCTCCT  
CTTCAAGACGAACCGGAAAGTTACCTGTAAACAGCTCAAAGAAAGACTATTTCAAAAAGATTGAATGTTTCGACTCTGTTG  
AAATCAGCGGAGTGGAGGATCGCTTCAACGCATCCCTGGGAACGTATCAGCATCTCCTGAAAAATCATTAAGACAAGGAC  
TTCTGGACAATGAGGAGAACGAGGACATTTCTGAGGACATTTGCTCCTCACCTTACGTTGTTTGAAGATAGGGAGATGAT  
TGAAGAACCGCTTGAAGAACTTACGCTCATCTCTTCGACGACAAAGTCATGAAACAGCTCAAGAGGCGCCGATATACAGGAT  
GGGCGCGGCTGTCAAGAAAGTATCAATGGGatcCGAGACAAGCAGAGTGGAAAGACAATCTGGATTCTTAAGTCC  
GATGGATTTGCCAACCGGAACCTTATGCAAGTTGATCCATGATGACTCTCTCACCTTTAAGGAGGACATCCAGAAAGCACA  
AGTTTCTGGCCAGGGGGACAGTCTTCAAGAGCACATCGCTAATCTTGCAAGTAGCCAGCTATCAAAAAGGGAATACTGC  
AGACCGTTAAGGTCTGGATGAACTGTCAAAGTAATGGGAAGGCATAAGCCCGAGAATATCGTTATCGAGATGGCCCGA  
GAGAACCAAACTACCCAGAGGAGACAGAAGACAGTAGGGAAGGATGAAGAGGATTGAAGAGGGTATAAAAGAACTGGG  
GTCCCAATCCTTAAGGAACCCAGTTGAAAAACCCAGCTTCAGAAATGAGAAGCTCTACCTGTACTACTTGCAGAACG  
GCAGGACATGTACGTGGATCAGGAACCTGGACATCAATCGGCTCTCCGACTACGACGTGGATGCCATCGTCCCGAGTCT  
TTTCTCAAGATGATTCTATTGATAATAAAGTGTGACAAGATCCGATAAAAAATAGAGGGGAAGAGTGATAACGTCCTC  
AGAAGAAAGTTGTCAAGAAATGAAAAATATTGGCGGCAGCTGCTGAACGCCAACTGATCACACACCGGAAGTTTCGATA  
ATCTGACTAAGGCTGAACGAGGTGGCTGTCTGAGTTGGATAAAGCCGGCTTCATCAAAAGGCAGCTTGTGTGAGACACGC  
CAGATCACCAAGcacGTGGCCCAATTTCTCGATTACGCATGAACACCAAGTACGATGAAAAATGACAAACTGATTCGAGA  
GGTGAAAGTTATTACTCTGAAGTCTAAGCTGGTCTCAGATTTCAGAAAGGACTTTCAGTTTATAAGGTGAGAGAGATCA  
ACAATTACCACTATGCGCATGATGCTACCTGAATGCAAGTGGTAGGCAGTGCATTTATCAAAAAATATCCCAAGCTTGAA  
TCTGAATTTGTTTACGGAGACTATAAAGTGTACGATGTTAGGAAAAATGATCGCAAGGTCTGAGCAGGAAATAGGCAAGGC  
CACCGCTAAGTACTTCTTTACAGCAATATTATGAATTTTTCAGAACCGAGATTACACTGGCCAATGGAGAGATTTCGGA  
AGCGACCACTTATCGAAACAAACGGAGAAACAGGAGAAATCGTGTGGGACAAGGATAGGATTTTCGCAGACATCCGGAAG  
GTCCTGTCCATGCCGAGGTGAACATCGTTAAAAAGACCGAAGTACAGACCGGAGGCTTCTCCAAGGAAGTATCCTCCC  
GAAAAGGAACAGCGACAAGCTGATCGCACGCAAAAAAGATTGGGACCCCAAGAAATACGGCGGATTCGATTCTCCTACAG  
TCGCTTACAGTGTACTGTTGTGGCCAAAGTGGAGAAAGGGAAGTCTAAAAAATCAAAAGCGTCAAGGAACCTGCTGGGC  
ATCACAACTATGGAGCGATCAAGCTTCGAAAAAACCCCATCGACTTCTCGAGGCGAAAGGATATAAGAGAGTCAAAAA  
AGACCTCATCATTAAAGCTTCCCAAGTACTCTCTCTTTGAGCTTGAACCGGCCGGAACGAATGCTCGCTAGTGCAGGCG  
AGCTGCAGAAAGGTAAACGAGCTGGCACTGCCCTCTAAATACGTAAATTTCTGTATCTGGCCAGCCACTATGAAAAGCTC  
AAAGGCTCTCCCGAAGATAATGAGCAGAAGCAGCTGTTCTGTGGAACAACAACAACACTACCTTGATGAGATCATCGAGCA  
AATAAGCGAATTTCCAAAAGAGTGATCTCGCCGACGCTAACCTCGATAAGGTGCTTTCTGCTTACAATAAGCACAGGG  
ATAAGCCCATCAGGGAGCAGGCAGAAAAATATCAACTGTTTACTCTGACCAACTTGGGCGCGCTGCAGCCTTCAAG  
TACTTCGACACCAACCATAGACAGAAAGCGGTACACCTCTACAAAGGAGGTCTGGACGCCACACTGATTATCAGTCAAT  
TACGGGCTCTATGAACAAGAACTGACCTCTCTCAGCTCGGTGGAGACAGCAGGGCTGACCCCAAGAGAAAGAGGAAG  
TGGctagCGATGCTAAGTCACTGACTGCCTGGTCCCGGACACTGGTGACCTTCAAGGATGTGTTTGTGGACTTCACAGG  
GAGGAGTGGAAAGCTGCTGGACACTGCTCAGCAGATCCTGTACAGAAATGTGATGCTGGAGAACTATAAGAACCTGGTTTC  
TTTGGTTATCAGCTTACTAAGCCAGATGTATCCTCCGTTGGAGAGGAGAGAGCCCTGGCTGGTGGAGAGAGAAA  
TTCACCAAGAGACCATCTGATTGAGAGATGCATTTGAATCAAATCATCAGTTCGAAAAAGAACGCAAGTtGct  
agCG

[0462]

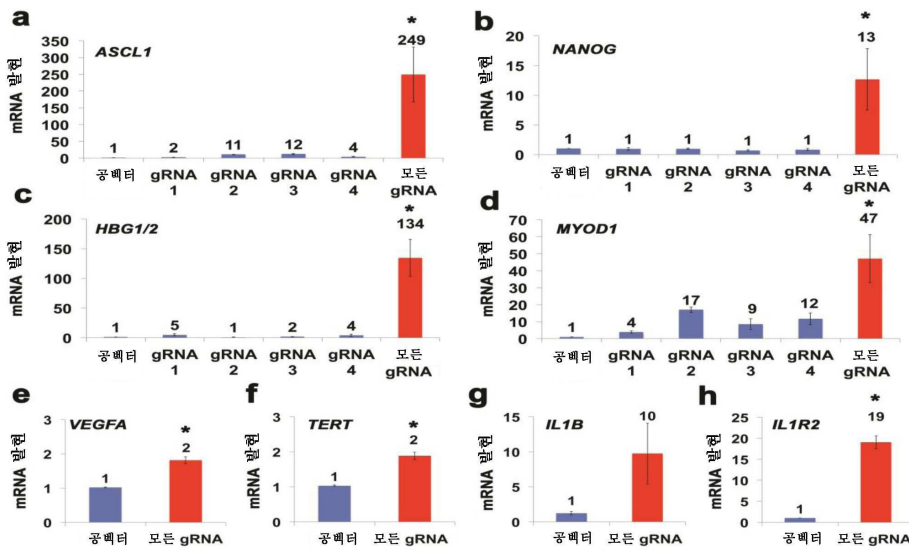


도면

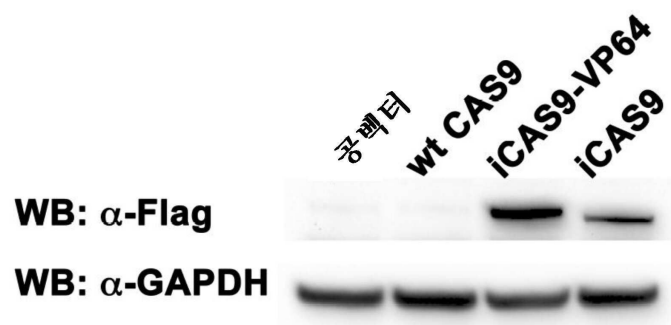
도면1



도면2



도면3



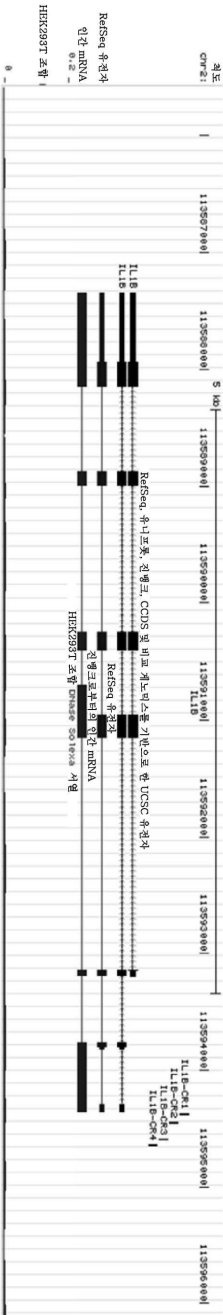




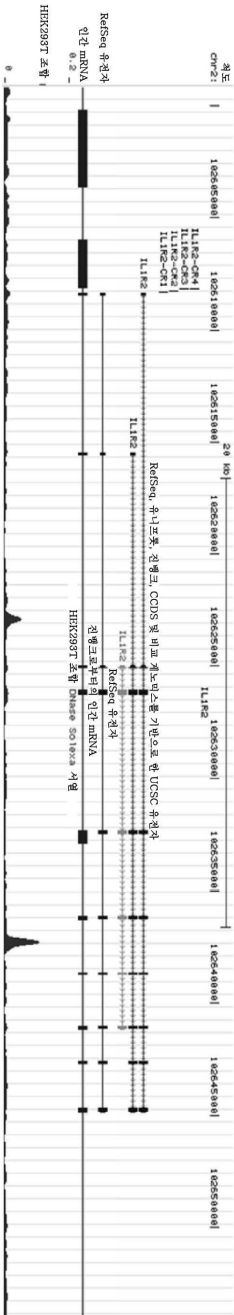


도면4d

IL1B.

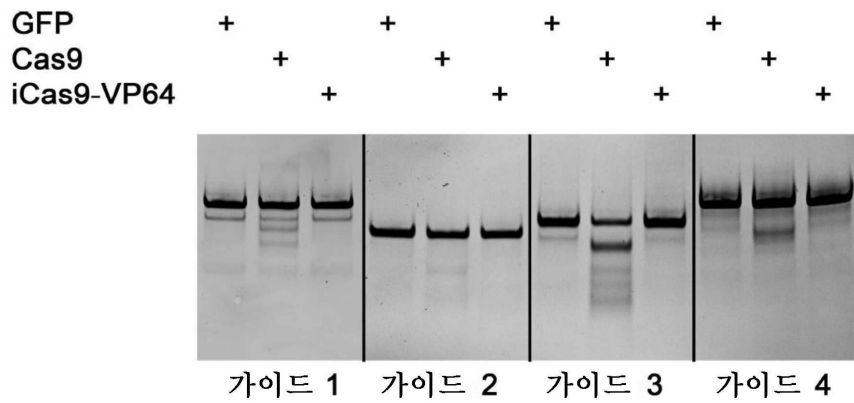


IL1R2.

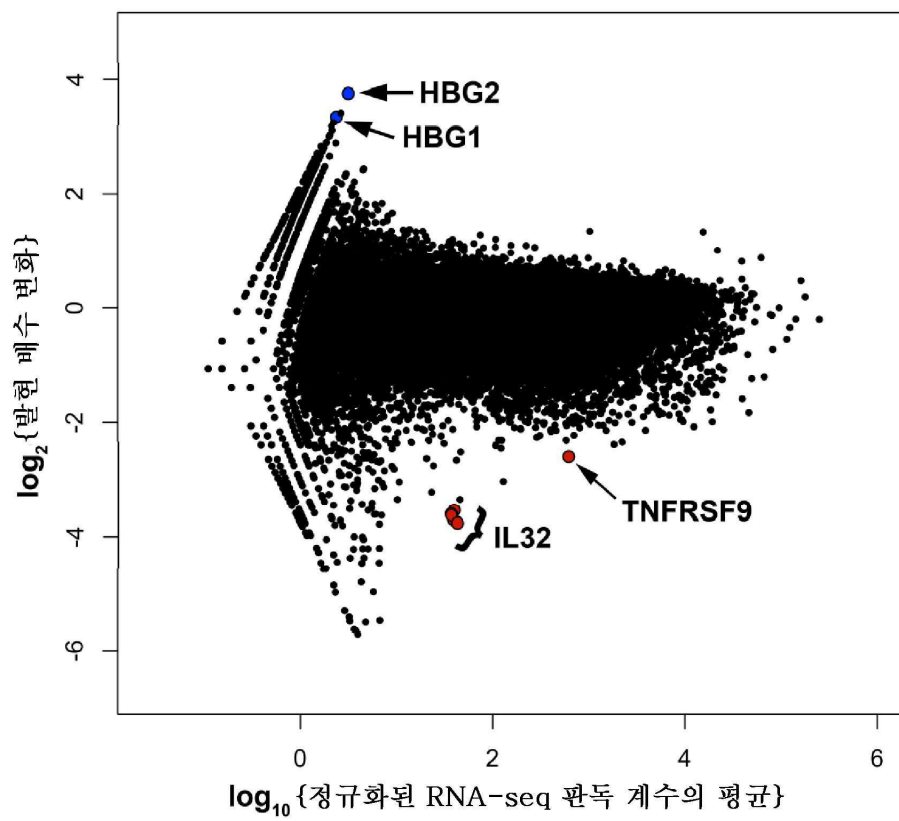




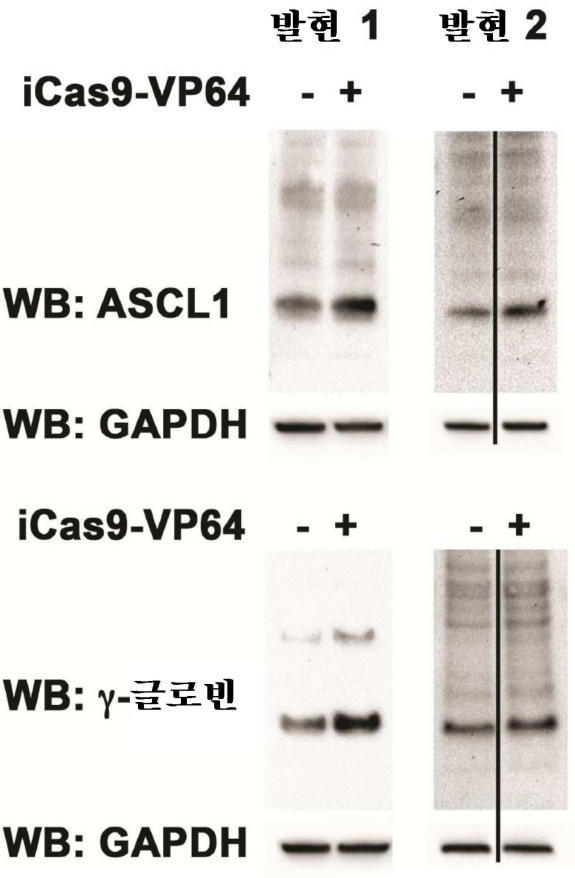
도면5



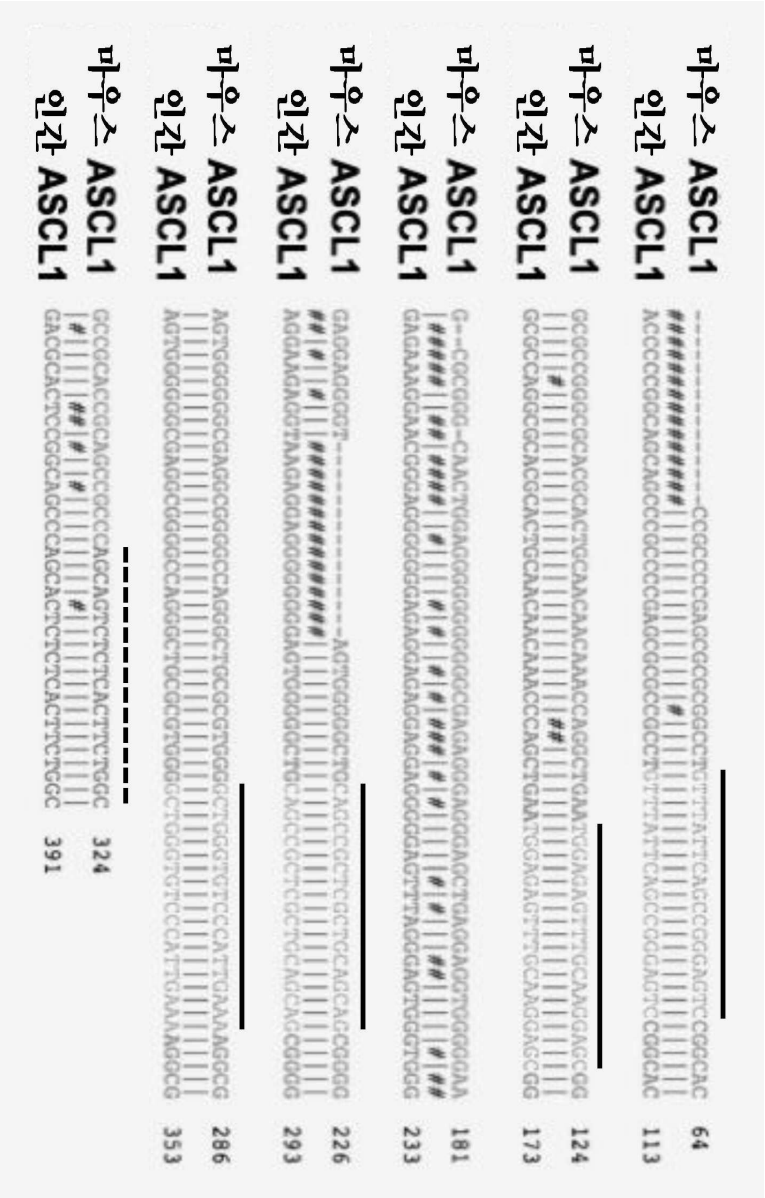
도면6



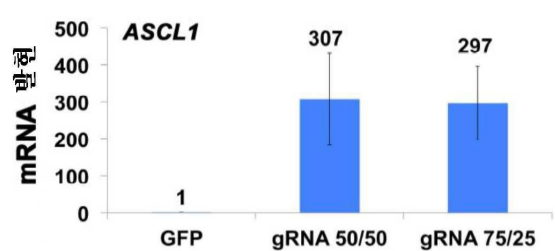
도면7



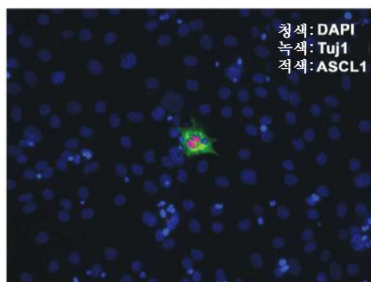
도면8a



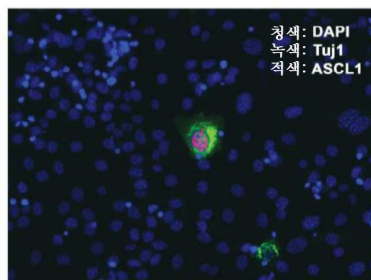
도면8b



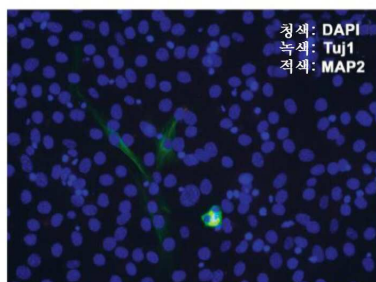
도면8c



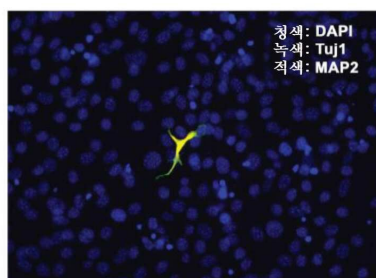
도면8d



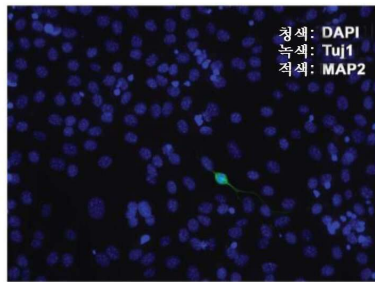
도면8e



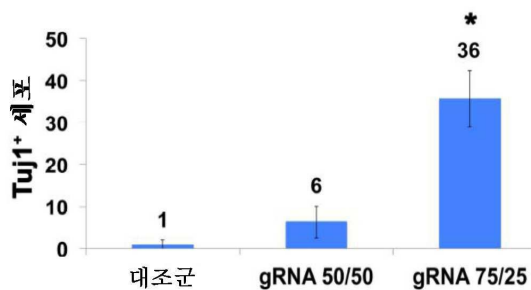
도면8f



도면8g



도면8h



도면9a

MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGRGMDKKYSIGLAIGTNSVGWAVITD  
EYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQ  
EIFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPFGNIVDEVAYHEKYPTIYHLRKKLVD  
STDKADLRLLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFQLVQTYNQLFEENPINASG  
VDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQL  
SKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEH  
HQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEE  
LLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTRIPY  
YVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASQSFIERMTNFDKNLPNEKVL  
PKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDY  
FKKIECFDSVEISGVEDRFNASLGTYHDLLKIKDKDFLDNEENEDILEDIVLTTLTFEDRE  
MIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFAN  
RNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIAKKGILQTVKVVDELVKV  
MGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNE  
KLYLYYLQNGRDMYVDQELDINRLSDYDVAIVPQSFLKDDSIDNKVLTRSDKNRGKSD  
NVPSEEVVKMKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQIT  
KHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAY  
LNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEGEIGKATKAYFFYSNIMNFFKTE  
ITLANGEIRKRPLIETNGETGEIWWDKGRDFATVRKVLSPQVNIKKTEVQTGGFSKESI  
LPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVAKVEKGKSKKLKSVKELLGITIME  
RSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLEFENGRKRMLASAGELQKGNELALPSK  
YVNFYLYASHYEKLGSPEDNEQKQLFVEQHKHYLDEIEQISEFSKRVLADANLNDKVL  
AYNKHDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSTIG  
LYETRIDLSQLGGDPAGSKASPKKKRKVGRADALDDFDLMDLGSALDDFDLMDLGS  
DALDDFDLMDLGSALDDFDLMDLINYPYDVDPDYAS (서열 1)

FLAG 에피토프 태그 = 이탤릭체

핵 국재화 서열 = 볼드체

스트렙토코쿠스 피오게네스 Cas9 = 밑줄표시

VP64 (4x 최소 VP16 도메인) = 이탤릭체 및 볼드체

HA 에피토프 태그 = 이탤릭체 및 밑줄표시

도면9b

GAGGGCCTATTTCCCATGATTCCTTCATATTTGCATATACGATACAAGGCTG  
 TTAGAGAGATAATTGGAATTAATTTGACTGTAAACACAAAGATATTAGTACA  
 AAATACGTGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAAA  
 TTATGTTTTTAAAAATGGACTATCATATGCTTATCGTAACTTGAAAGTATTTTGA  
 TTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACCGGTCTTCGAGA  
AGACTGTTTAGAGCTAGAATAGCAAGTTAAATAAGGCTAGTCCGTTATC  
 AACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTT (서열 2)

U6 프로모터 = 볼드체

+1 전사 개시 부위 = 밑줄표시

가이드 RNA에서 클로닝하기 위한 BbsI 제한 부위 = 이탤릭체 및 밑줄표시

키메라 가이드 RNA 서열 = 이탤릭체

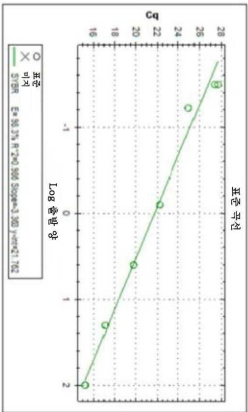
폴리-T 종결인자 서열 = 볼드체 및 밑줄표시



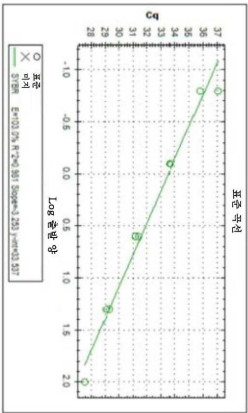


도면10b

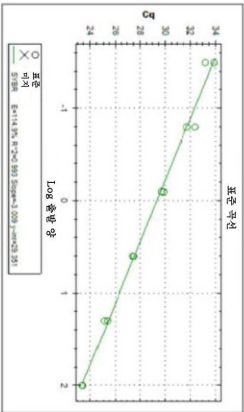
GAPDH



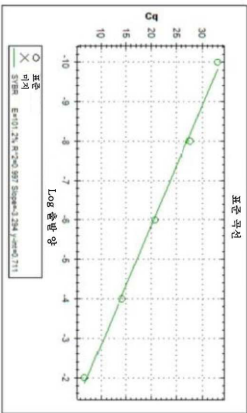
MASCL1



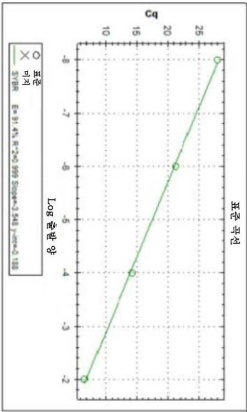
IL1RN



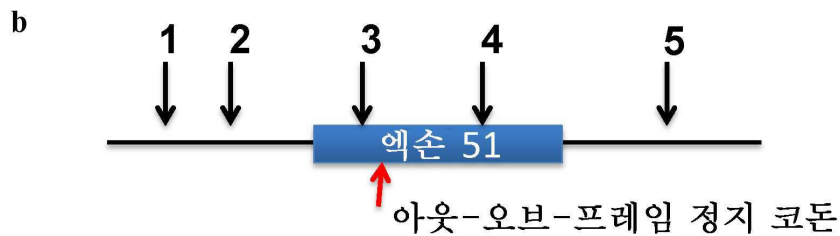
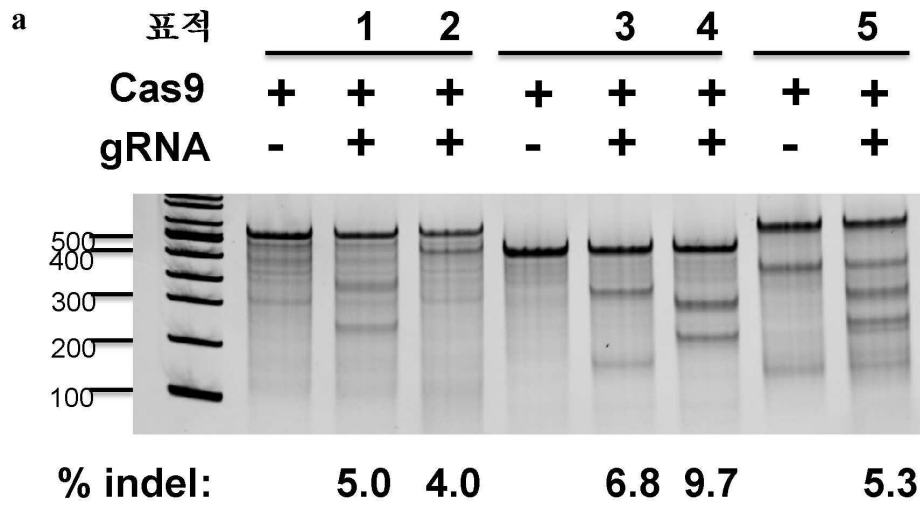
HGB1



MYOD1



도면11



**c**

예상된 절단 크기

CR1=352/237

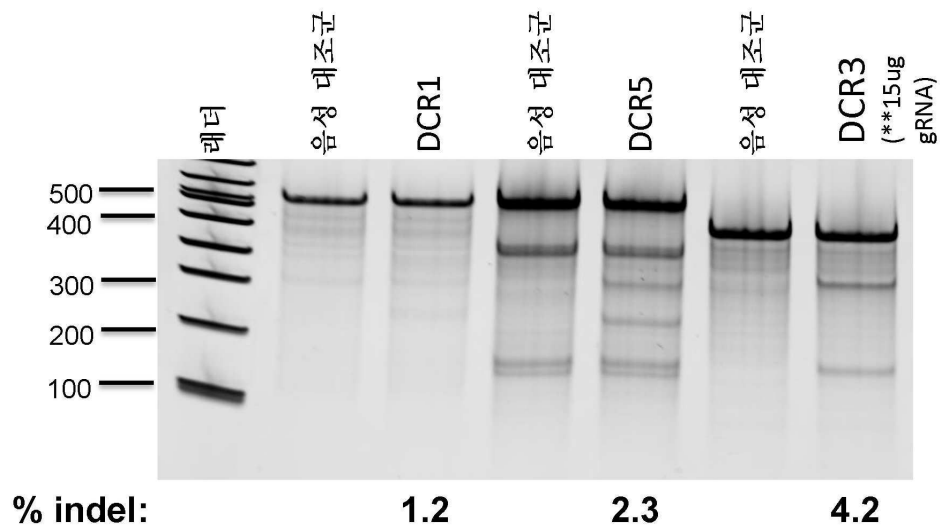
CR2=461/128

CR3=301/150

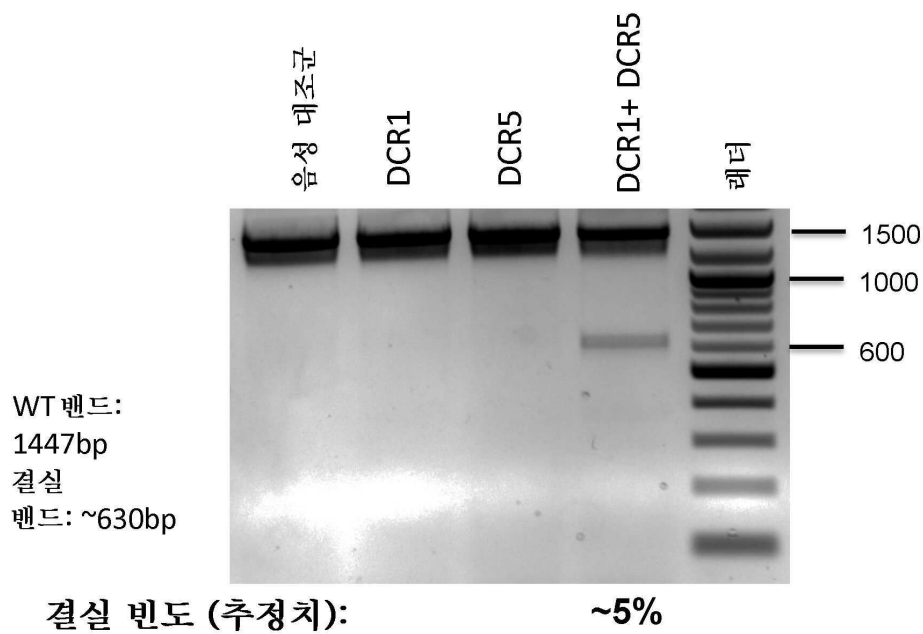
CR4=272/179

CR5=280/203

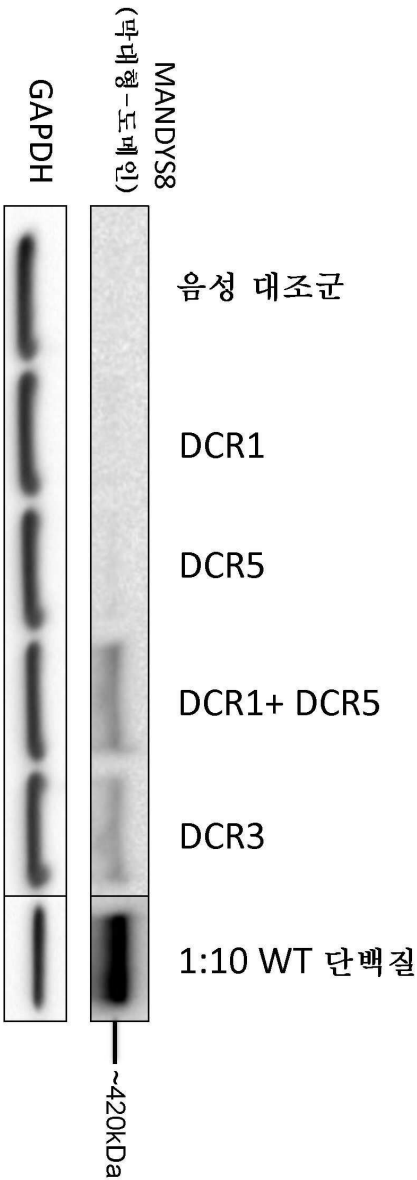
도면12



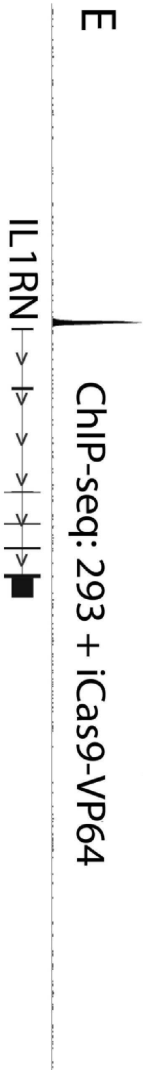
도면13



도면14

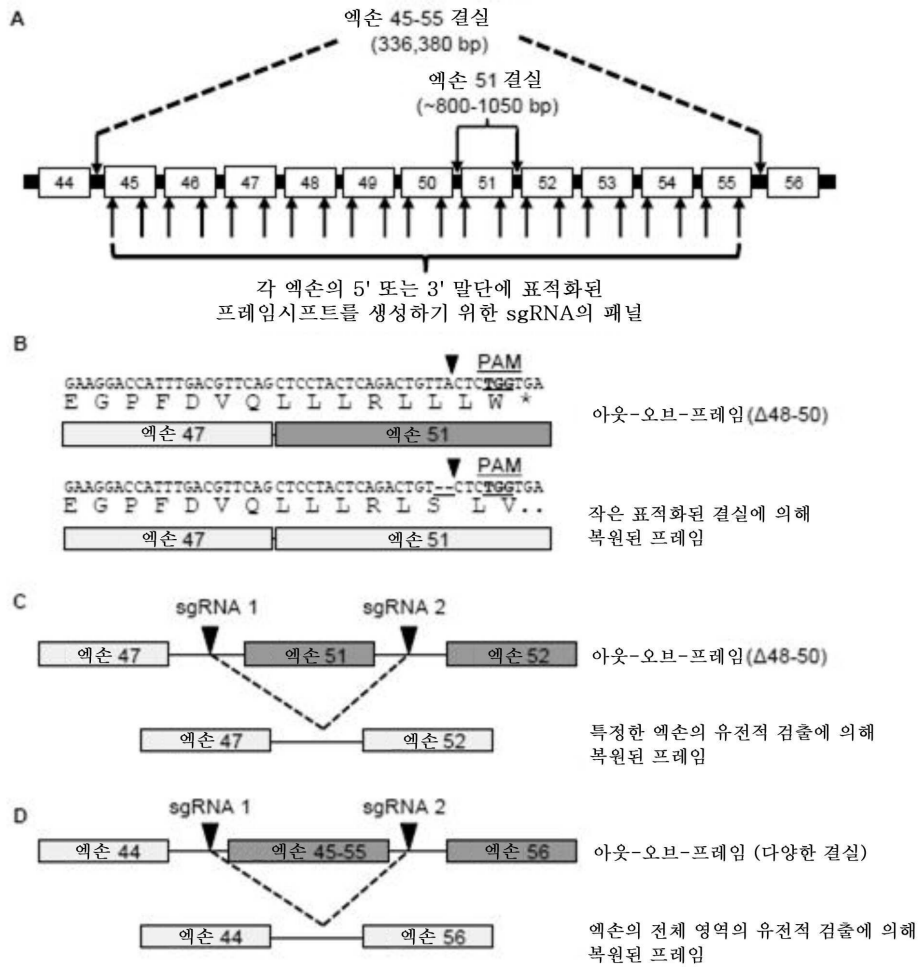


도면15

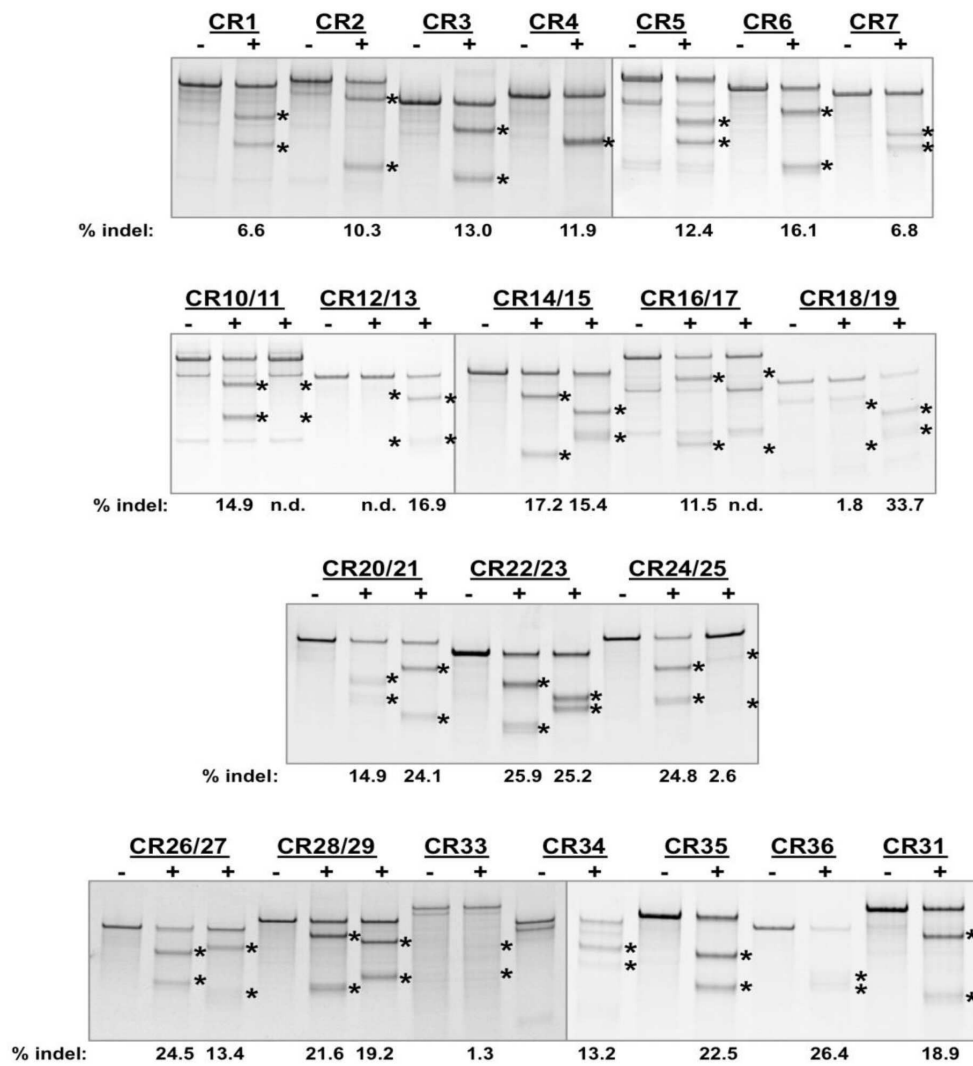




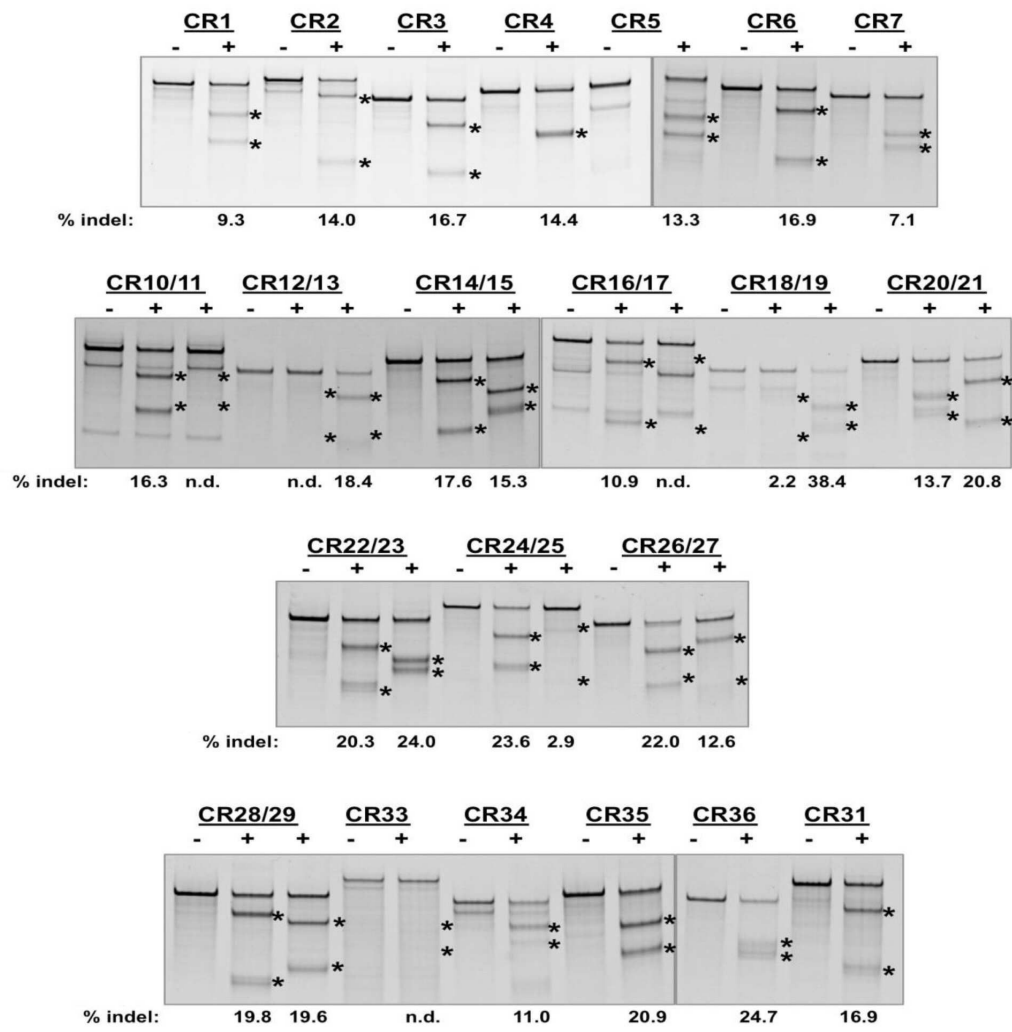
도면16



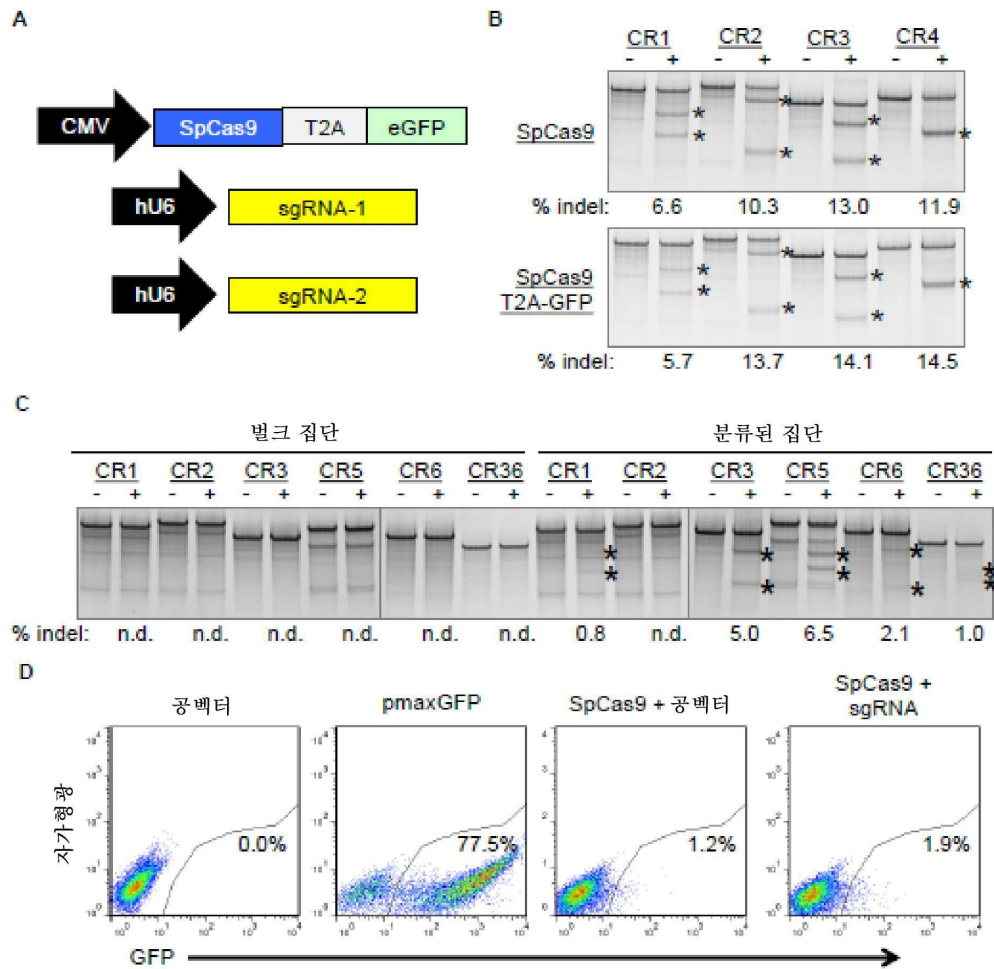
도면17



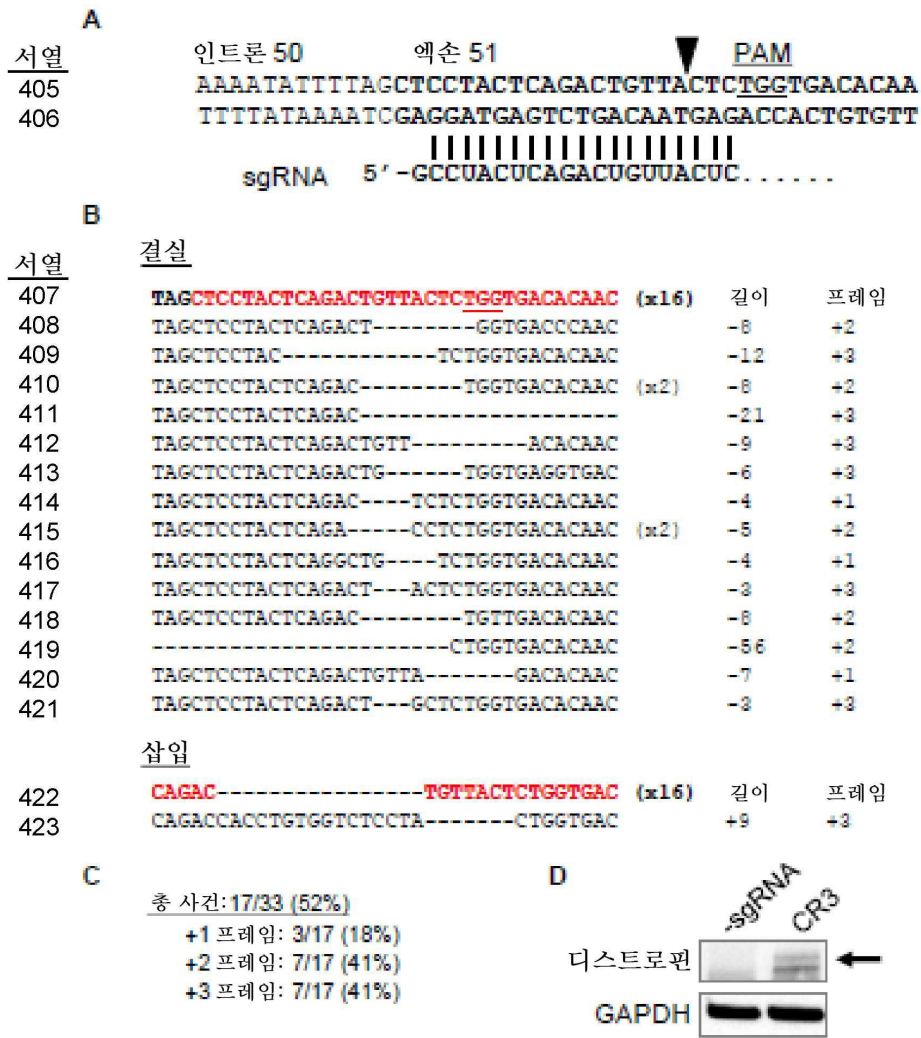
도면18



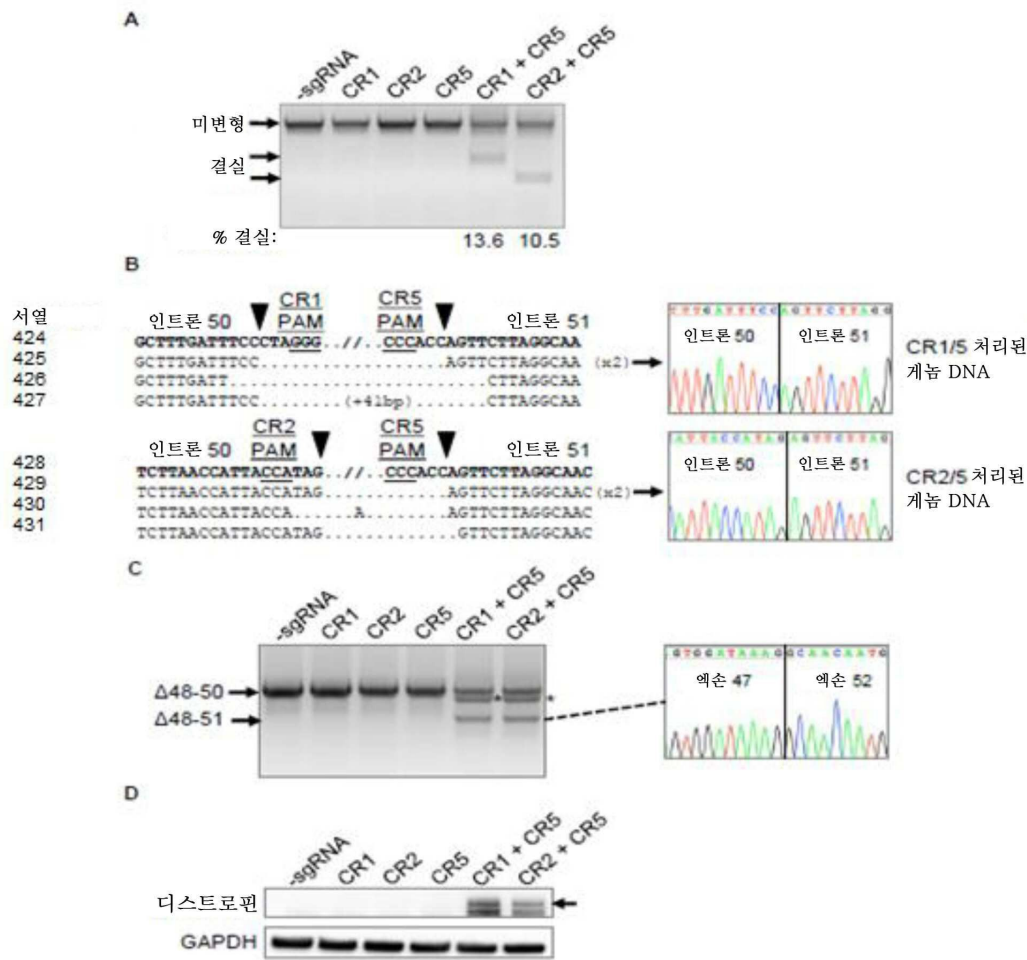
도면19



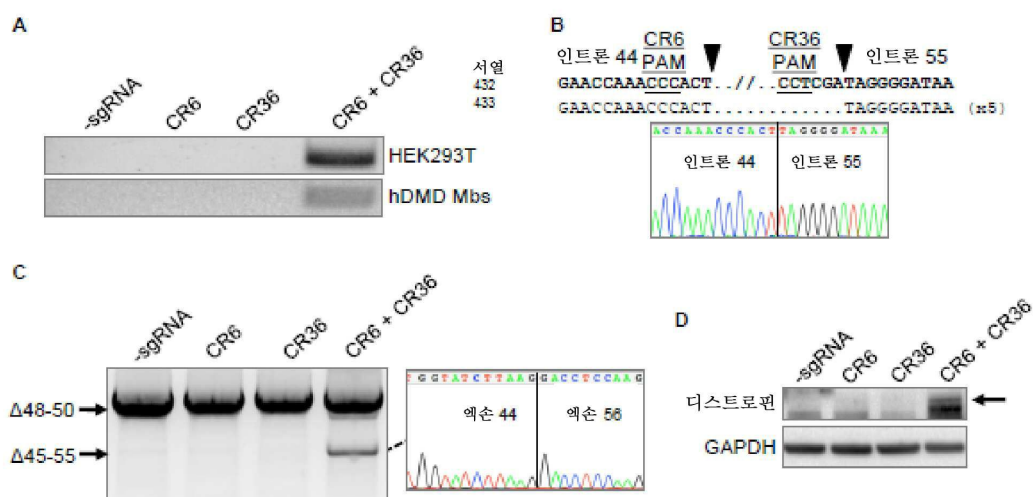
도면20



도면21

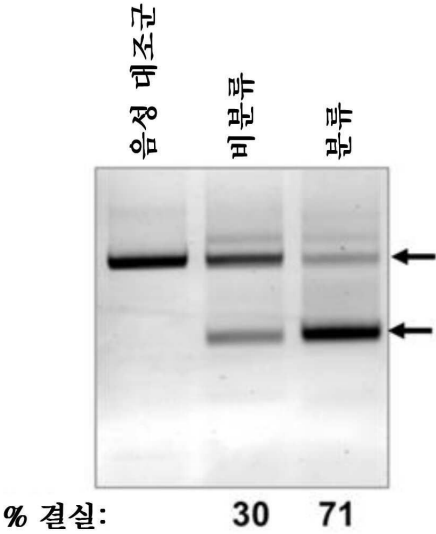


도면22

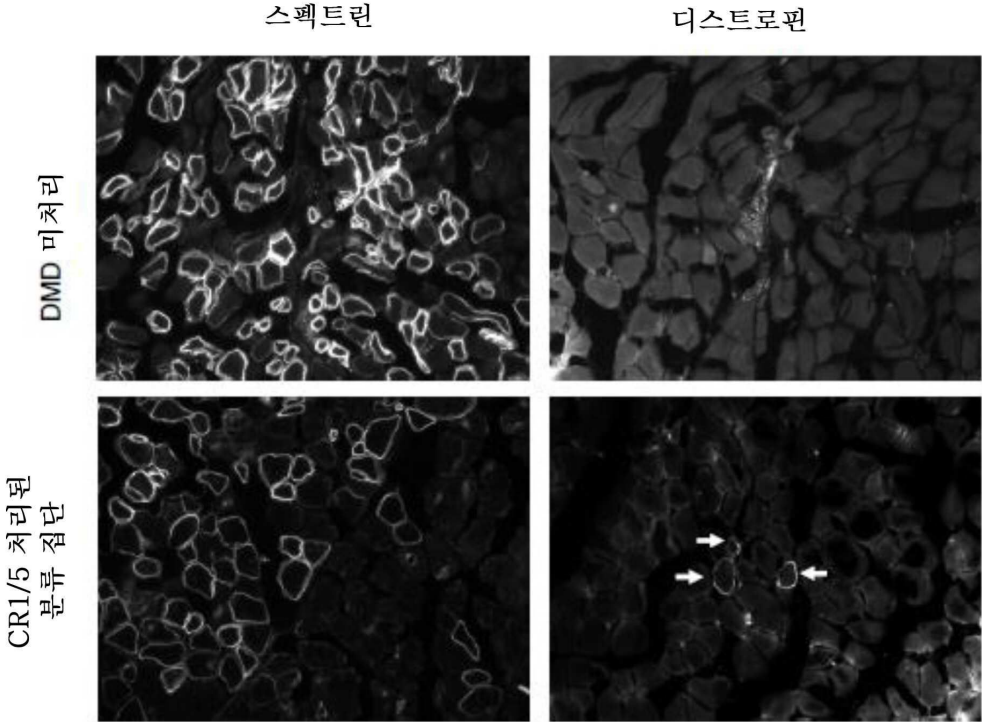




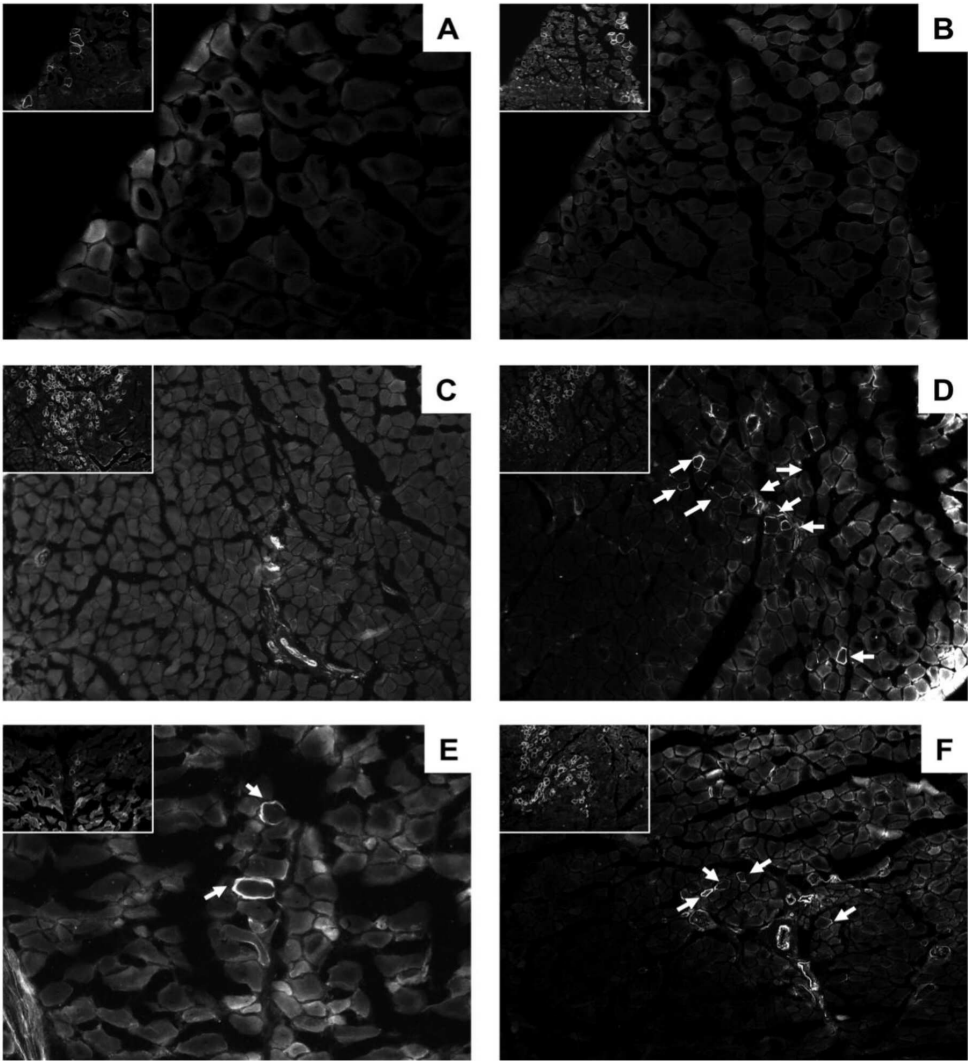
도면23



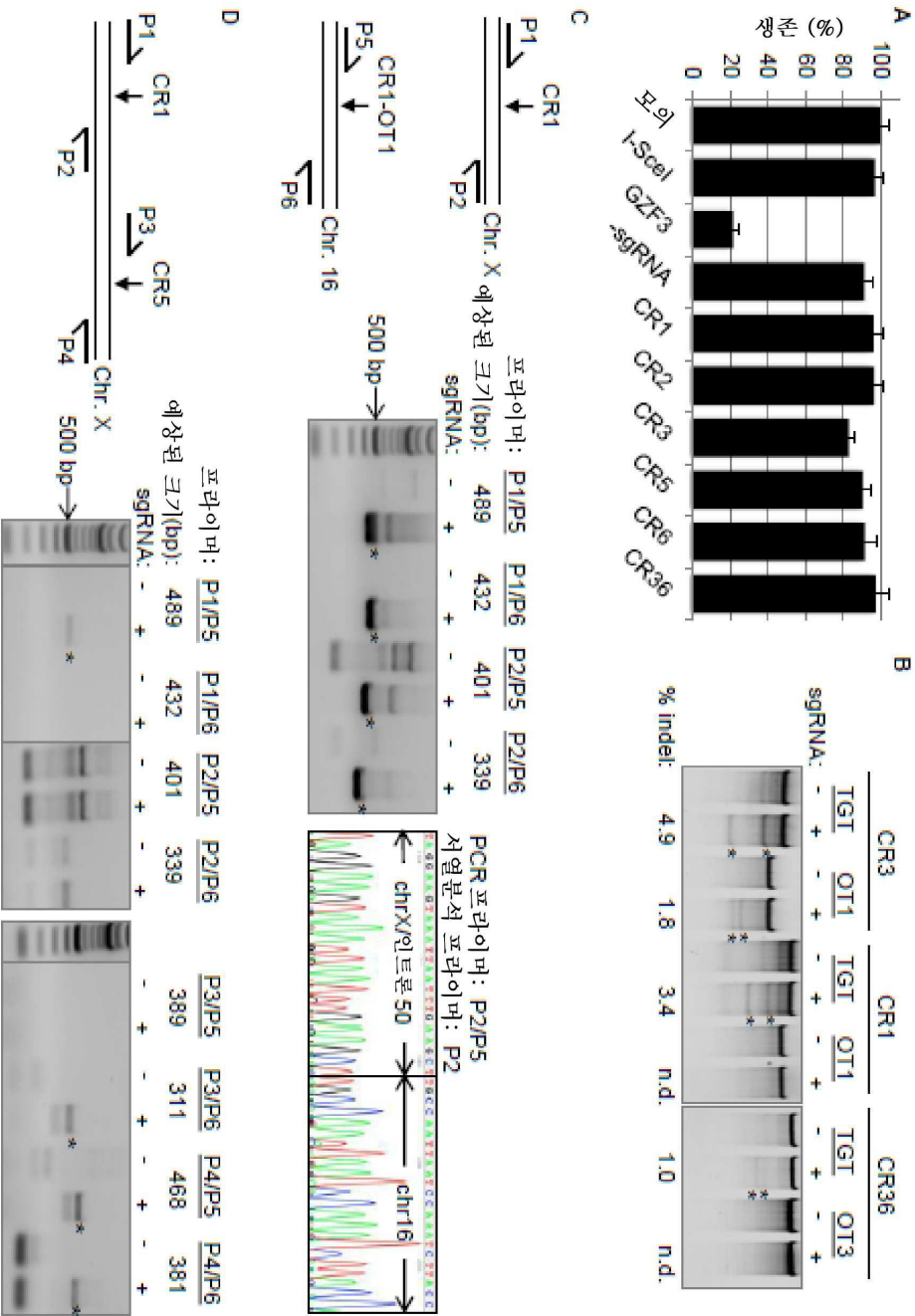
도면24



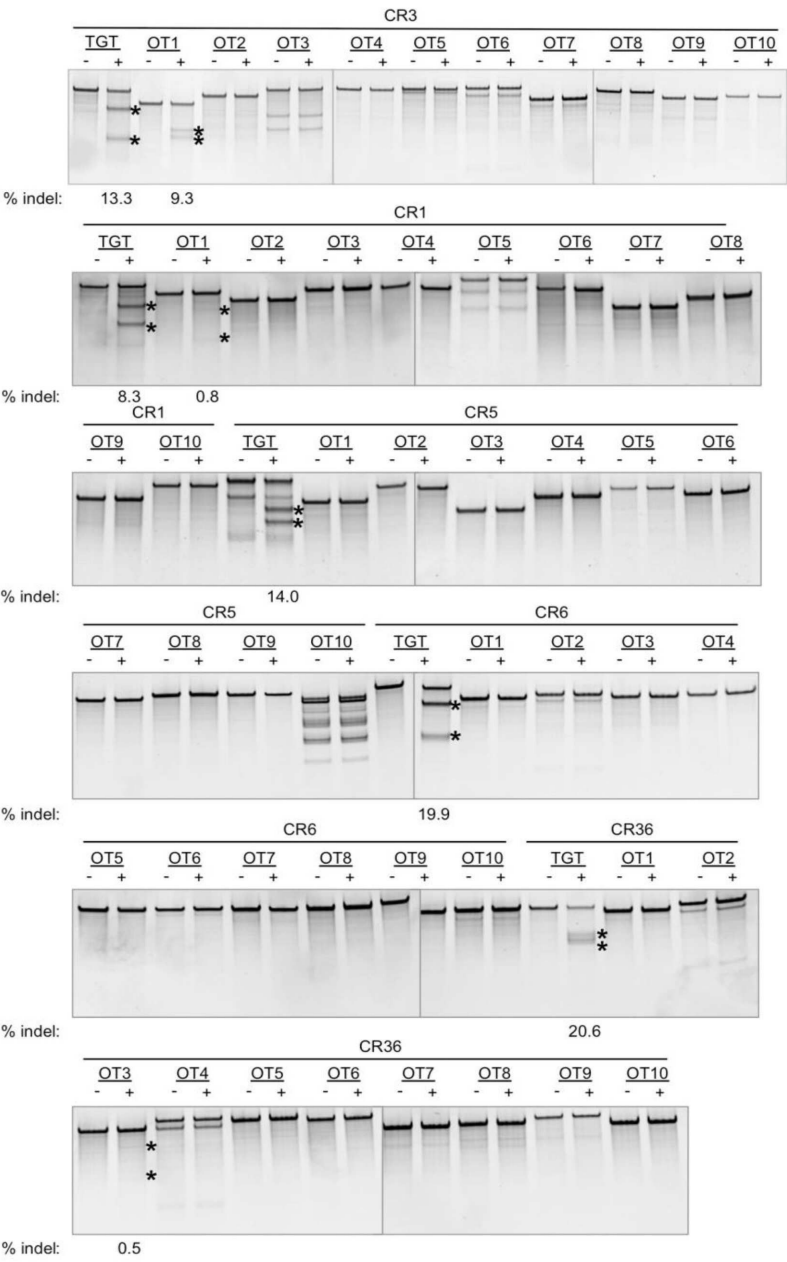
도면25



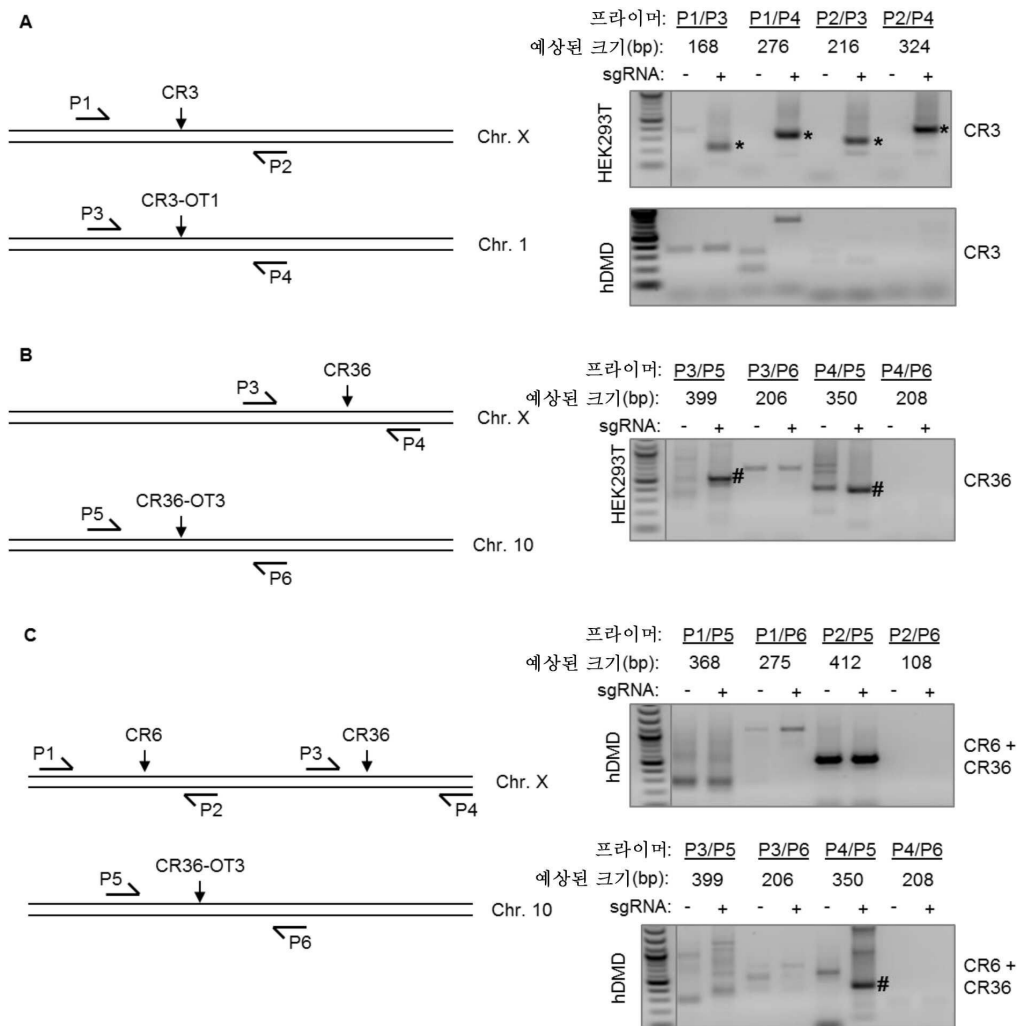
도면26



도면27



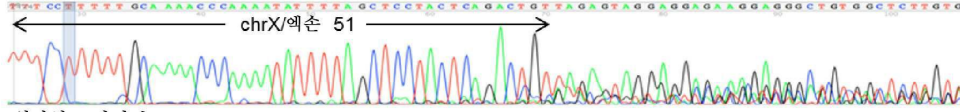
도면28



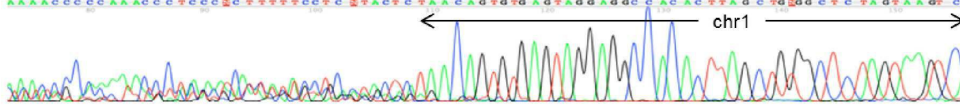
도면29

**P1/P3 밴드**

정방향 프라이머

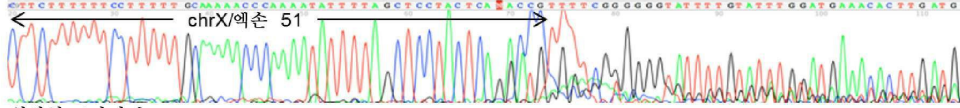


역방향 프라이머

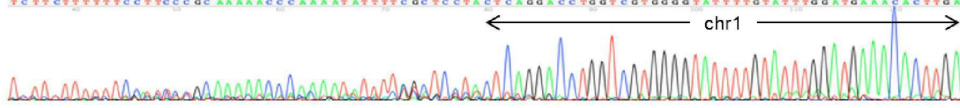


**P1/P4 밴드**

정방향 프라이머

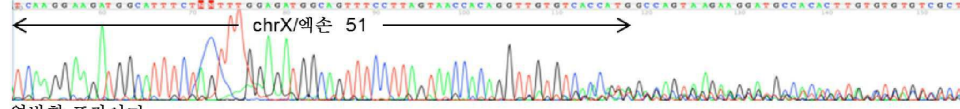


역방향 프라이머

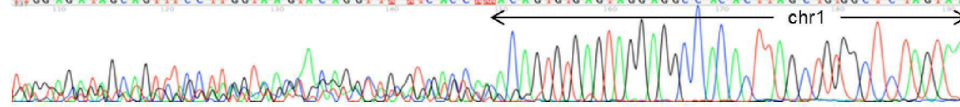


**P2/P3 밴드**

정방향 프라이머

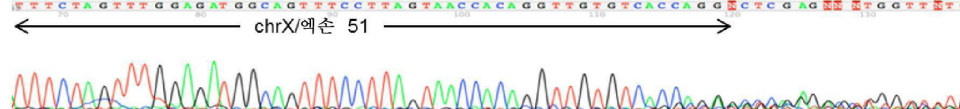


역방향 프라이머

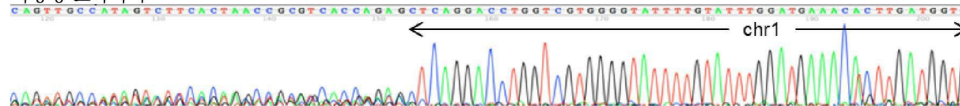


**P2/P4 밴드**

정방향 프라이머

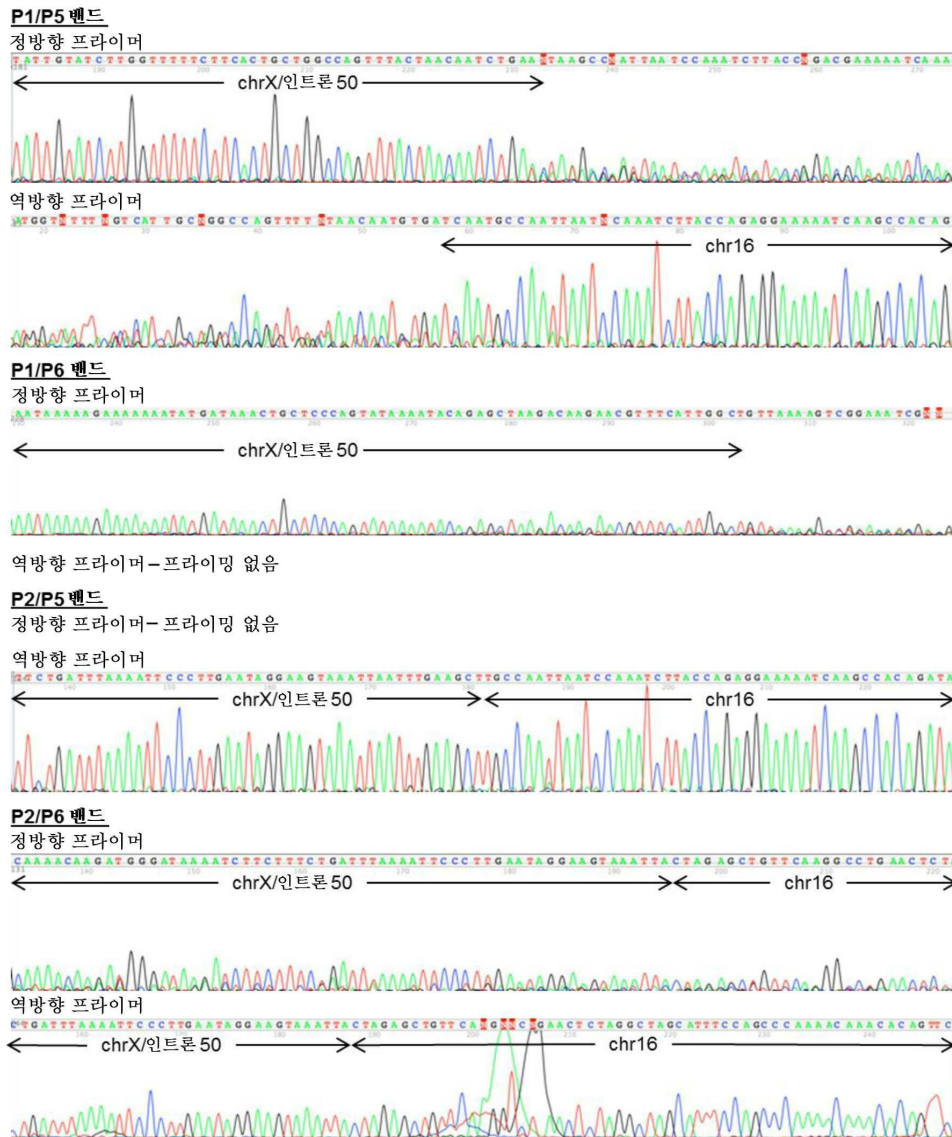


역방향 프라이머

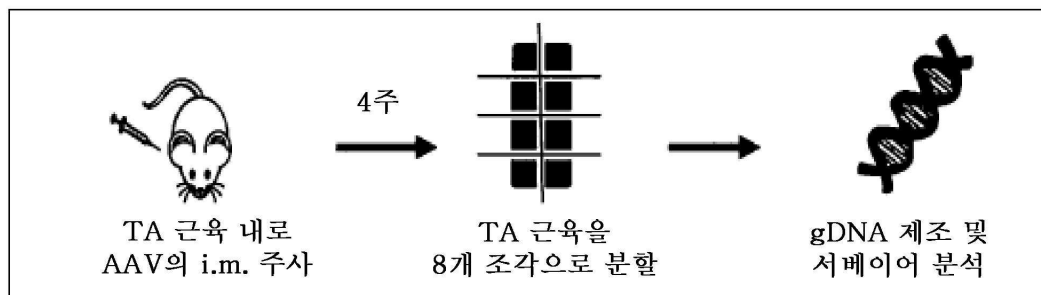




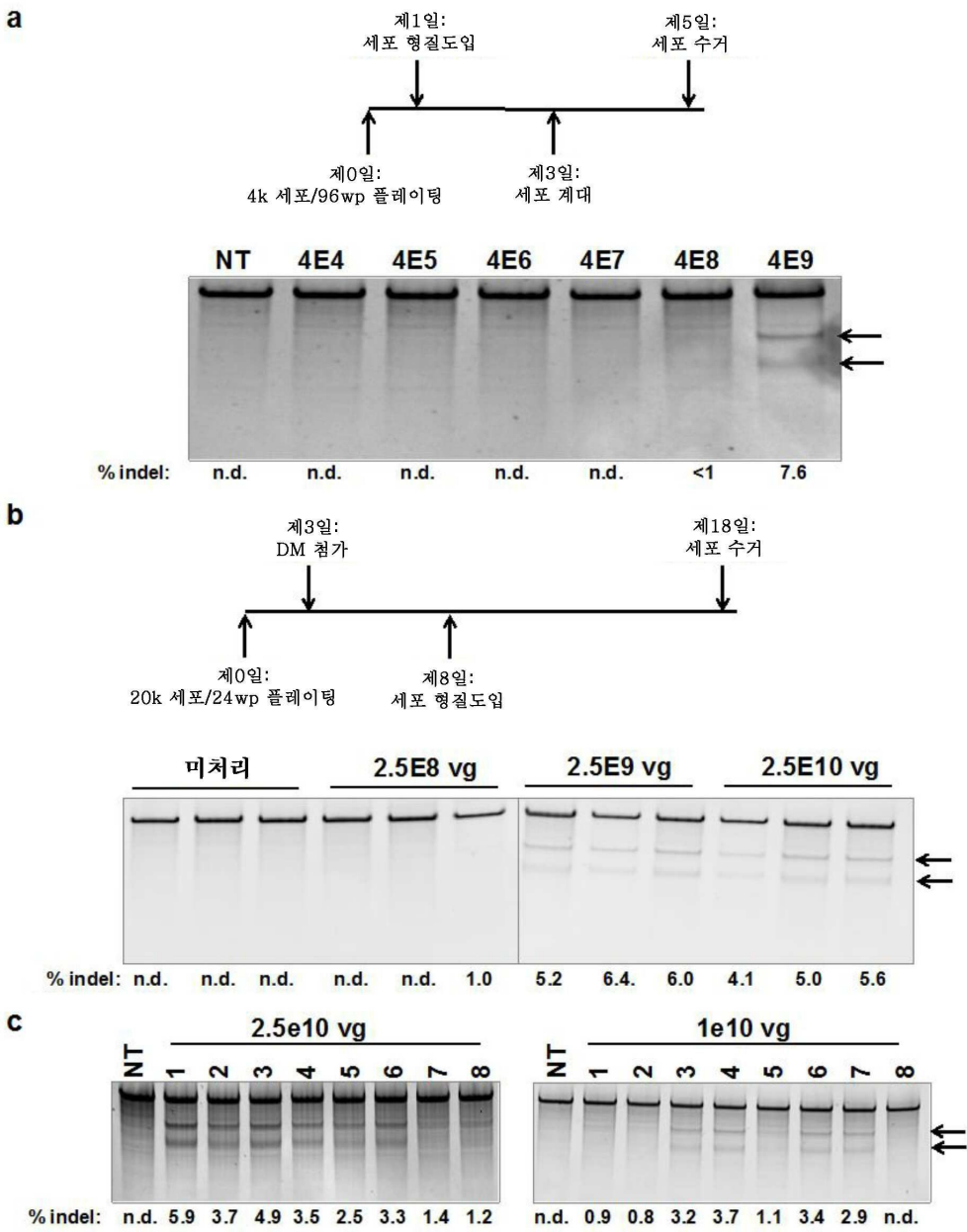
도면30



도면31



도면32



도면33

**Rosa T2A opt DNA 서열 (서열 434)**

ATGAGGTCTGACTACAAAGACCATGACGGTGATTATAAAGATCATGACATCGATTACAAGGATGACGATGACAAGAT  
GGCCCCCAAGAAAGAGgaaggtgggcctcgaGCCCCGAGAAAAACCGTACAAGTGCCCTGAGTGCGGGAAATCAT  
TCTCCGACCCCTGGGGCGCTCGTCCGGCACCACAAAGGACGCATACAGGGGAAAAGCCGTATAAGTGCCCCGAGTGTGGA  
AAGAGCTTCTCGCAGAGAGCCACCTTGAACGACACCAAAGAACACACACTGGTGAGAAACCTATAAGTGTCAGAG  
GTGCGGCAAAATCGTTTAGCAGATCCGATGACTTGGTGCGCCACCAGCGGACACACACGGGTGAAAAGCCCTACAAAT  
GCCCCGAGTGTGGGAAGTCTTTTTCAAGGTGCGATCATCTGACTACCCATCAGCGCACCCATACGGGAGCGgcgcgc  
cgcgccctGGTGAAGAGCGAGCTGGAGGAGAAGAAGTCCGAGCTGCGGCACAAGCTGAAGTACGTGCCCCACGAGTA  
CATCGAGCTGATCGAGATCGCCAGGAACCCACCCAGGACCGCATCTGGAGATGAAGGTGATGGAGTTCTTCATGA  
AGGTGTACGGTACAGGGGAGAGCCTGGGCGGAAGCAGAAAGCCTGACGGCGCCATCTATACAGTGGGCAGCCCC  
ATCGATTACGGCGTGATCGTGGACACAAAGGCCCTACAGCGCGCGGTACAATCTGCCTATCGGCCAGGCCGACGAGAT  
GCAGAGATACGTGAAGGAGAACCAGACCCGGAATAAGCACATCAACCCCAACGAGTGGTGGAAGGTGTACCCTAGCA  
GCGTGACCGAGTTCAAGTTCCTGTTCGTGAGCGGCCACTTCAAGGGCAACTACAAGGCCAGCTGACAGGCTGAAC  
CGCAAAACCAACTGCAATGGCGCCGTGCTGAGCGTGGAGGAGTCTGATCGGCGCGGAGATGATCAAAGCCGGCAC  
CCTGACACTGGAGGAGTGCAGCGCAAGTTCAACAACGGCGAGATCAACTTCGAGGGCAGAGGAAGTCTTCTAACAT  
GCGGTGACGTGGAGGAGAATCCCGGCCCTAgatcTGACTACAAAGACCATGACGGTGATTATAAAGATCATGACATC  
GATTACAAGGATGACGATGACAAGATGGCCCCCAAGAAAGAGgaaggtgggcctcgaGCCCCGAGAGAAGCCGTA  
CAAGTGTCCCGAATGTGAAAGAGCTTCTCACAGTGGGGGACCTTCGGCGCCACCAGCGCACACATACTGGTGAAA  
AGCCGTATAAGTGTCCAGAATGTGGCAAATCATTCTCCACATCAGGGAGCCTGGTCAGGCACCAGCGAACCCACACG  
GGTGAGAAGCCCTATAAGTGCCCCGAATGCGGGAAGTCTTTTTCGCAGAGAGCCCACTTGGAGAGGCACAGAGGAC  
CCATACGGGGGAGAAACCTTACAAGTGCCCTGAATGCGGAAAGTCTTCTCGACCCATCTGGATCTCATCAGACATC  
AGAGAACGCACACTGGAGAGAAAACCTACAAATGTCCGAGTGTGGGAAGTCTGTTAGCCGAAAGGACAATCTCAAA  
AACCATCAACGGACACACACGGGTGAAAAACCATACAAATGCCCGGAGTGCGGCAAAATCGTTTTCCCACTTGCGCA  
CTTGCGGGCACACCAACGCACGCATACTGGAGCGGCCGCGcggcgcctGGTGAAGAGCGAGCTGGAGGAGAAGAAGT  
CCGAGCTGCGGCACAAGCTGAAGTACGTGCCCCACGAGTACATCGAGCTGATCGAGATCGCCAGGAACCCACCCAG  
GACCGCATCTGGAGATGAAGGTGATGGAGTTCTTCATGAAGGTGTACGGCTACAGGGGAGAGCACCTGGGCGGAAG  
CAGAAAGCCTGACGGCGCCATCTATACAGTGGGCGAGCCCATCGATTACGGCGTGATCGTGAGACACAAAGGCCATCA  
GCGGCGGCTACAATCTGCCTATCGGCCAGGCCGACGAGATGGAGAGATACGTGGAGGAGAACCAGACACGGGATAAG  
CACCTCAACCCCAACGAGTGGTGGAAGGTGTACCCTAGCAGCGTGACCGAGTTCAAGTTCCTGTTCTGTGAGCGGCCA  
CTTCAAGGGCAACTACAAGGCCAGCTGACCAGGCTGAACCACATCACCAACTGCAATGGCGCCGTGCTGAGCGTGG  
AGGAGCTGCTGATCGGCGGCGAGATGATCAAAGCCGGCACCCCTGACACTGGAGGAGGTGCGGCGCAAGTTCACAAC  
GGCGAGATCAACTTCTGA

**Rosa T2A opt 단백질 서열 (서열 435)**

MRSDYKDHDGDYKDHDIDYKDDDDKMAPKKRKRVGLEPGEKPYKCECGKSFSDPGALVRHQRTHTGEKPYKCECG  
KSFSQRAHLERHQRTHTGEKPYKCECGKSFSDDLVRHQRTHTGEKPYKCECGKSFSDHLLTHQRTHTGAAA  
RALVKSELEEKSELRLKLYVPHEYIELIEIARNPTQDRILEMKVMEFFMKVGYRGEHLGGSRKPDGAIYTVGSP  
IDYGVIVDTKAYSGGYNLPIGQADEMQRYVKENQTRNKHINPNEWWKVYPSSVTEFKFLFVSGHFKGNYKAQLTRLN  
RKTNCNGAVLSVEELLIGGEMIKAGTLTLEEVRKFNNGEINFEGRGSLLTCGDVEENPGPRSDYKDHDGDYKDHD  
DYKDDDDKMAPKKRKRVGLEPGEKPYKCECGKSFSGDLRRHQRTHTGEKPYKCECGKSFSTSGSLVRHQRTHT  
GEKPYKCECGKSFQRAHLERHQRTHTGEKPYKCECGKSFSTHLDLRHQRTHTGEKPYKCECGKSFSRKDNLK  
NHQRTHTGEKPYKCECGKSFSQLAHLRAHQRTHTGAAARALVKSELEEKSELRLKLYVPHEYIELIEIARNPTQ  
DRILEMKVMEFFMKVGYRGEHLGGSRKPDGAIYTVGSPIDYGVIVDTKAYSGGYNLPIGQADEMERYVEENQTRDK  
HLNPNNEWWKVYPSSVTEFKFLFVSGHFKGNYKAQLTRLNHITNCNGAVLSVEELLIGGEMIKAGTLTLEEVRKFN  
GEINF

도면34a

**SASTG 캡시드 서열 (서열 436)**

tggccactccctctatgcgcaactcgctcgctcggtggggcctggcgaccaaagtgccagacggacgtgctttgca  
cgccggccccaccgagcgagcgagtgcgcatagagggagtgccaaactccatcactagaggtatggcagtgacgta  
acgcgaagcgcggaagcgagaccacgcctaccagctgctgcagcagtgacgtgacccttttgcgacagtttgcgac  
accacgtggcgcgtgaggtatataattctcgagtgagcggaaccaggagctccattttgaccgcgaaatttgaacgag  
cagcagccatgccccgggttctacgagattgtcctgaaggtcccgagtgacctggacgagcacctgccccgatttct  
aactcgtttgttaactgggtggcgagaaaggaatgggagtgccgcccgtattctgacatggatccgaatctgattga  
gcaggcaccctgacgtggccgaaagcttcagcgcgagttcctggtggagtgggcgccgctgagtaagggcccg  
aggccctctttttgtccagttcgaaaaaggggagacctaactccacctgcacgtgctgattgagaccatcggggtc  
aaatccatggtggtcgccgctacgtgagccagattaaagagaagctggtgacctgcacatctaccgccccggtcgacc  
gcagcttcgcgaactggttcggtgacgcaaaacgcgaatggcgccggggcggaacaaggtggtggacgactgct  
acatcccccaactactgctccccagaccagcccgagctccagtggtgggactaactggaccagttttaagc  
gctgtttgaatctcgcgagcgtaaacggctggtggcgagcatctgacgcagctgtcgagacgcaggagcagaa  
caaagagaatcagaacccaattctgacgcgcggtcatcaggtcaaaaacctcagccaggtacatggagctggtcg  
ggtggtggtggagcggttcgagcagcaaaaagcaatggattcaggaggaccagggcctcgatctctctcaac  
ggcgctccaactcggtggtccagatcaaggccgctggacaatgctccaagatcatgagcctgacaaagacggc  
tcgggactacgtggtggcgagcaaccgcccggagagacattacaaaaatcggtatctacaaaatcctggagctgaacg  
ggtacgatccgcagtagcgccctcctcttctgggctggcgcaaaagaagttcggaagaggaaacaccatctgg  
ctctttggcgcccgacgagcggttaaaacacacatcgcggaagcctcgccacgcctgcccctctacggctgct  
aaactggaccaatgagaactttcccttcaacgattgctgcgacaagatggtgatctggtgggaggagggcaagatga  
cggccaaggtcggtggagagcgcaaggccattctggcggaagcaaggtgcgcgtggaccaaagtgcaagtcacg  
gcccagatcgaaacccactcccgtagctgcaccccaacacacacatgctgcgcgtgattgacgggaacagcaccac  
cttgagcatcgaaacgcggtctcaggaaccggtggttaatttgaacttaccgcccgttggaccatgacttctcaac  
aggtcaccaaacaggaagtaaggacttttccggtgggcttccgatcacgtgactgacgtggctcatgagttctac  
gtcagaaaggtggtgagctaaagaaacgccccgcctccaatgacgcggatgtaagcgagccaaaacggcagtgacgctc  
acttgcgagcgacaaacgtcagacgcggaagcaccggcggtacgcggacaggtaccaaaacaaatgttctcgtc  
acgtggcgatgaatctgattcttccctgtaaaacacatgagagagaatgaatcaaatttccaatgtctgttttacg  
catggtcaagagactgtgggaatgcttccctggaatgtcagaatctcaaccggtttctgctgcaaaaagaagac  
ttatcagaaactgtgtccaattcatcatatcctgggaagggcaccgcagattgctgttcggcctgcgatttggcca  
atgtggacttggatgactgtgttcttgagcaataaattgacttaaacaggtatggctgctgacggttatcttccaga  
atcggtcgaggacaaccttctcagggcattctgagtggtggctctgaaacctggagtcctcaaccacaaagcga  
accaacaacaccaggacaacgctcggtgttctgcttccgggttacaataacctcgaccggtaacggactcgac  
aaaggagagccggtcaacgagggcgacgcggcaccctcgaaacagcaaaagcttacgaccagcagctcaaggccg  
tgacaaccctgacctcaagtacaaccacgcgcgagcttccaggagcgtcttcaagaagatagctcttttgggg  
gcaaccttgccagagcaggtcttcagggccaaaagaggaatccttgagcctcttggctggttgaggaagcagctaaa  
acggctcctggaaagaagaggcctgtagatcagctcctcaggaaacggactcatcatctggtgttggcaaatcggg  
caaacagcctgccagaaaaagactaaatttcggtcagactggcgactcagagtcagtcaccagaccctcaacctctcg  
gagaaccaccagcagccccacaaagtttgggatctaatacaatggcttcaggcggtggcgaccaatggcagacaat  
aacgaggtgcccagtgagtggtggttaattcctcaggaaattggcattgagcttcccaatggctggcgacagagtcac  
caccaccagcaccagaacctgggcccctgcccacttacaacaacctctctacaagcaaatctccagcGCTtcaACGg  
gagcttcaaacgcaaccactactttggtacagcacccttgggggtattttgactttaacagattccactgccac  
ttctcaccacgtgactggcagcactcattaacaacaactgggattccggcccaagaaactcagcttcaagctctt  
caacatccaagttaaagaggtcacgcagacgagtgccagcagactattgccaataaccttaccagcagcgttcaag  
tggttacggactcgagtatcagctcccgtagctgctcggtcgccgaccaaaggtgtctcccgcggtttccagcg  
gacgtcttcatggtccctcagtatggataacctcaccctgaacaacggaagtcaagcggtgggacgctcatccttta  
ctgcttgagtagtcttcccttcgagatgctaaggactggaaataaacttccaattcagctataccttcgaggtatgac  
cttttcacagcagctacgctcacagccagagtttggatcgcttgatgaatcctcttattgatcagtatctgtactac  
ctgaacagaacgcaaggaaacacctctggaacaaccaaccaatcacggctgctttttagccaggctgggcccctcagtc  
tatgtctttgcaggccagaaattggctacctgggcccgtctacccggcaacagagactttcaaagactgctaacgaca  
acaacaacagtaactttccttgacagcgccagcaaatatcatctcaatggccgcgactcgctggtgaatccagga  
ccagctatggccagtcacaaggacgatgaagaaaaattttccctatgcacggcaatctaattttggcaagaaagg  
gacaacggcaagtaacgcagaattagataatgtaattgattacggatgaagaagagattcgtaaccaccaatcctgtgg  
caacagagcagtatggaactgtggcaataaacttcagagctcaaatacagctcccacgactagaactgtcaatgat  
cagggggccttaccttgcatggtgtggcaagatcgtagctgtaccttcaaggacctatctgggcaagattctca  
cacggatggacactttcatccttctcctctgatgggaggttttggactgaaacatccgctcctcaaatcatgtca  
aaaaactccggtaccggcaaatcctccgacgactttcagccccggcaagtttcttcatcttactcagtaactcc

### 도면34b

actggacaggtcagcgtggaaattgagtgaggactacagaaagaaaacagcaaacgttggaatccagagattcagta  
cacttccaactacaacaagtctgttaattgtggactttactgttagacactaatggtgtttatagtgaaacctcgcccta  
ttggaacccggtatctcacacgaaactgttaactctgttaatacaataaacgtttaattcgtttcagttgaacttt  
ggctcttgtgcacttcttattcttattcttggttccatggctactgcttagataagcagcgccctgcgcgcttgcgct  
tcgcggtttacaactgctgggttaataatttaactctcgccatacctctagtgatggagttggccactccctctatgcg  
cactcgctcgctcggtggggccggacgtgcaaagcacgtccgtctggcgacctttggtcgccaggccccaccgagcg  
agcgagtgcgcatagaggagtgggcaa

#### SASTG 캡시드 캡티드 서열 (서열 437)

MAADGYLPDWLEDNLSEGIREWALKPGVFPQPKANQQHQDNRRGLVLPGYKYLPGNGLDKGEFVNEADAAALEHDK  
AYDQQLKAGDNPYLKYNHADAERLQEDTSFGGNLGRAVFQAKKRILEPLGLVEEAAKTAPGKKRPVDQS PQEPD  
SSSGVGKSGKQPKARKLNFQGTGDSVDPDQPLGEPAPAPTSLSGNTMASGGGAPMADNNEGADGVNSSGNWHCD  
SQWLGDVRVITTTSTRTWALPTYNHLYKQISSASTGASNDNHYFGYSTPWGYFDENRFHCHFS PRDWQRLINNNWGF  
PKKLSFKLFNIQKVEVTQNDGTTIANNLSTVQVFTDSEYQLPYVLGSAHQCLPPFPADVFMVPQYGYLTNNGS  
QAVGRSSFYCLEYFPSQMLRTGNFQFSYTFEDVFPFHSSYAHSSQLDRLMNPIDQYLYYLNRTQGTSTGTTNQSR  
LFSQAGPQSMQLQARNWLPGPCYRQQLSKTANDNNNSNFPWTAASKYHLNGRDSLVPNGPAMASHKDDEEKFFPMH  
GNLIFGKEGTTASNAELDNVMTDEEEIRTTNPVATEQYGTVANLQSSNTAPTTRTVNDQGALPGMVWQDRDVYLQ  
GPIWAKI PHTDGHFHPSPLMGGFGLKHPPQIMIKNTVPANPPTTSPAKFASFITQYSTGQVSVIEWELQKENS  
KRWNPEIQYTSNYNKS NVDFVTNGVYSEPRPIGTRYLTRNL

### 도면35

#### DZF16 ZFN 표적 부위 (서열 442) :

5' - CAAACTAGAAATGCCATCTTCCTTGATGTTGGAGGTACCTGC - 3'  
3' - GTTTGATCTTTACGGTAGAAGGAAC TACAACCTCCATGGACG - 5'

#### DZF16-L6 좌측 전체 아미노산 서열 (서열 443)

MDYKDHDGDYKDHDIDYKDDDDKMAPKKRKVGRLEPGEKPYKCPECGKSF SRKDALRGHQRTHTGEKPY  
KCPECGKSF SHRTTLTNHQRTHTGEKPYKCPECGKSF SQRNALAGHQRTHTGEKPYKCPECGKSF SHKNA  
LQNHQRTHTGEKPYKCPECGKSF SDPGHLVRHQRTHTGEKPYKCPECGKSF STSGNLVRHQRTHTGAAAR  
ALVKSELEEKSELRHKLKYVPHEYIELIEIARNPTQDRILEMKVMEFFMKVYGYRGEHLGGSRKPDGAI  
YTVGSPIDYGVIVDTKAYSGGYNLPIGQADEMERYVEENQTRDKHLNPNEWVKVYSSVTEFKFLFVSGH  
FKGNYKAQLTRLNHNITNCNGAVLSVEELLIGGEMIKAGTLTLEEVRKFNNGEINF\*

#### DZF16-R6 우측 전체 아미노산 서열 (서열 444)

MDYKDHDGDYKDHDIDYKDDDDKMAPKKRKVGRLEPGEKPYKCPECGKSF SQRSVLVGHQRTHTGEKPY  
KCPECGKSF SDKKDLTRHQRTHTGEKPYKCPECGKSF STSGHLVRHQRTHTGEKPYKCPECGKSF QRAH  
LERHQRTHTGEKPYKCPECGKSF STSGSLVRHQRTHTGEKPYKCPECGKSF STSGNLVRHQRTHTGAAAR  
ALVKSELEEKSELRHKLKYVPHEYIELIEIARNPTQDRILEMKVMEFFMKVYGYRGEHLGGSRKPDGAI  
YTVGSPIDYGVIVDTKAYSGGYNLPIGQADEMQRYVKENQTRNKHINPNEWVKVYSSVTEFKFLFVSGH  
FKGNYKAQLTRLNRKTNCGAVLSVEELLIGGEMIKAGTLTLEEVRKFNNGEINF\*

### 도면36

#### E51C3 표적 부위 (서열 445) :

5' - (t) ATCTGCCCATGACTGGCGCAGGG (a) -3'  
3' - (a) TAGACGGGTACTGACCGCTCCC (t) -5'

#### E51C-3L 좌측 전체 아미노산 서열 (서열 446)

MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGRLPEGERPFQCRICMRNFSKQALAVHTRTHTGEKPF  
QCRICMRNFSQSTTLKRHLRTHTEKPFQCRICMRNFSRSDHLSLHLKTHLRGSQLVKSELEKKSELRH  
KLKYPVPEYIELIEIARNPTQDRILEMKVMEFFMKVGYRGEHLGGSRKPDGAIYTVGSPIDYGVIVDTK  
AYSGGYNLPIGQADEMQRYVKENQTRNKHINPNEWWKVYPSSVTEFKFLFVSGHFKGNYKAQLTRLNRKT  
NCNGAVLSVEELLIGGEMIKAGTLTLEEVRRKFNNGEINF\*

#### E51C-3R 우측 전체 아미노산 서열 (서열 447)

MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGRLPEGERPFQCRICMRNFSRAHLQNHTRTHTGEKPF  
QCRICMRNFSQSTTLKRHLRTHTEKPFQCRICMRNFSDDGHLTRHLKTHLRGSQLVKSELEKKSELRH  
KLKYPVPEYIELIEIARNPTQDRILEMKVMEFFMKVGYRGEHLGGSRKPDGAIYTVGSPIDYGVIVDTK  
AYSGGYNLPIGQADEMERYVEENQTRDKHLNPNNEWWKVYPSSVTEFKFLFVSGHFKGNYKAQLTRLNHIT  
NCNGAVLSVEELLIGGEMIKAGTLTLEEVRRKFNNGEINF\*

### 도면37

#### DZF15 표적 부위 (서열 448) :

5' -ACTAGAAATGCCATCTTCCTTGATGTTGGAGGTACCTGCTCT-3'  
3' -TGATCTTTACGGTAGAAGGAACCTACACCTCCATGGACGAGA-5'

#### DZF15-L6 좌측 전체 아미노산 서열 (서열 449)

MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGRLPEGEKPYKCPECGKSFSHRTTLTNHQRTHTGEKPY  
KCPECGKSFSQRNALAGHQRTHTGEKPYKCPECGKSFSHKNALQNHQRTHTGEKPYKCPECGKSFSDDPGH  
LVRHQRTHTGEKPYKCPECGKSFSSTSGNLVRHQRTHTGEKPYKCPECGKSFSQSSNLVRHQRTHTGAAAR  
ALVKSELEKKSELRHKLKYPVPEYIELIEIARNPTQDRILEMKVMEFFMKVGYRGEHLGGSRKPDGAI  
YTVGSPIDYGVIVDTKAYSGGYNLPIGQADEMERYVEENQTRDKHLNPNNEWWKVYPSSVTEFKFLFVSGH  
FKGNYKAQLTRLNHITNCNGAVLSVEELLIGGEMIKAGTLTLEEVRRKFNNGEINF\*

#### DZF15-R6 우측 전체 아미노산 서열 (서열 450)

MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGRLPEGEKPYKCPECGKSFSQRNALAGHQRTHTGEKPY  
KCPECGKSFSQQRSLVGHQRTHTGEKPYKCPECGKSFSDDKDLTRHQRTHTGEKPYKCPECGKSFSSTSGH  
LVRHQRTHTGEKPYKCPECGKSFSQRAHLERHQRTHTGEKPYKCPECGKSFSSTSGSLVRHQRTHTGAAAR  
ALVKSELEKKSELRHKLKYPVPEYIELIEIARNPTQDRILEMKVMEFFMKVGYRGEHLGGSRKPDGAI  
YTVGSPIDYGVIVDTKAYSGGYNLPIGQADEMQRYVKENQTRNKHINPNEWWKVYPSSVTEFKFLFVSGH  
FKGNYKAQLTRLNRKTNCNGAVLSVEELLIGGEMIKAGTLTLEEVRRKFNNGEINF\*

#### DZF15-L5 좌측 전체 아미노산 서열 (서열 451)

MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGRLPEGEKPYKCPECGKSFSQRNALAGHQRTHTGEKPY  
KCPECGKSFSHKNALQNHQRTHTGEKPYKCPECGKSFSDDPGHLVRHQRTHTGEKPYKCPECGKSFSSTSGN  
LVRHQRTHTGEKPYKCPECGKSFSQSSNLVRHQRTHTGAAARALVKSELEKKSELRHKLKYPVPEYIEL  
IEIARNPTQDRILEMKVMEFFMKVGYRGEHLGGSRKPDGAIYTVGSPIDYGVIVDTKAYSGGYNLPIGQ  
ADEMERYVEENQTRDKHLNPNNEWWKVYPSSVTEFKFLFVSGHFKGNYKAQLTRLNHITNCNGAVLSVEEL  
LIGGEMIKAGTLTLEEVRRKFNNGEINF\*

#### DZF15-R5 우측 전체 아미노산 서열 (서열 452)

MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGRLPEGEKPYKCPECGKSFSQQRSLVGHQRTHTGEKPY  
KCPECGKSFSDDKDLTRHQRTHTGEKPYKCPECGKSFSSTSGHLVRHQRTHTGEKPYKCPECGKSFSQRAH  
LERHQRTHTGEKPYKCPECGKSFSSTSGSLVRHQRTHTGAAARALVKSELEKKSELRHKLKYPVPEYIEL  
IEIARNPTQDRILEMKVMEFFMKVGYRGEHLGGSRKPDGAIYTVGSPIDYGVIVDTKAYSGGYNLPIGQ  
ADEMQRYVKENQTRNKHINPNEWWKVYPSSVTEFKFLFVSGHFKGNYKAQLTRLNRKTNCNGAVLSVEEL  
LIGGEMIKAGTLTLEEVRRKFNNGEINF\*



도면38

**E51C4 표적 부위 (서열 453) :**

5' - (t) GCCATCTTCCTTGATGTTGGAGGT (a) -3'  
3' - (a) CGGTAGAAGGAACCTACAACCTCCA (t) -5'

**E51C-4L 좌측 전체 아미노산 서열 (서열 454)**

MDYKDHDGDYKDHDIDYKDDDDKMAPKKRKVGRLEPGERPFQCRICMRNFSSPSKLARHTRHTGKPF  
QCRICMRNFSVRHNLTRHLRTHHTGKPFQCRICMRNFSQRNNLGRHLKTHTGAAARALVKSELEEKSEL  
RHKLKYPHEYIELIEIARNPTQDRILEMKVMEFFMKVYGYRGEHLGGSRKPDGAIYTVGSPIDYGVVD  
TKAYSGGYNLPIGQADEMQRYVKENQTRNKHINPNEWVKVPSSVTEFKFLFVSGHFKGNYKAQLTRLNR  
ITNCNGAVLSVEELLIGGEMIKAGTLTLEEVRKFNNGEINF\*

**E51C-4R 우측 전체 아미노산 서열 (서열 455)**

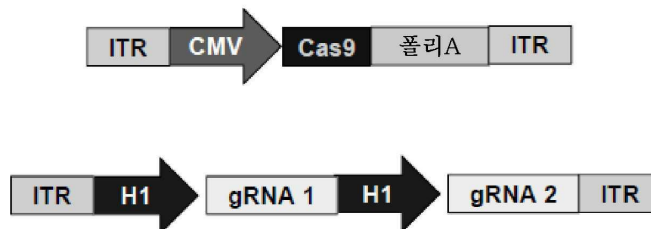
MDYKDHDGDYKDHDIDYKDDDDKMAPKKRKVGRLEPGERPFQCRICMRNFSIPNHLARHTRHTGKPF  
QCRICMRNFSQSAHLKRHLRTHHTGKPFQCRICMRNFSHHNSLTRHLKTHTGAAARALVKSELEEKSEL  
RHKLKYPHEYIELIEIARNPTQDRILEMKVMEFFMKVYGYRGEHLGGSRKPDGAIYTVGSPIDYGVVD  
TKAYSGGYNLPIGQADEMERYVEENQTRDKHLNPNEWVKVPSSVTEFKFLFVSGHFKGNYKAQLTRLNH  
ITNCNGAVLSVEELLIGGEMIKAGTLTLEEVRKFNNGEINF\*

도면39

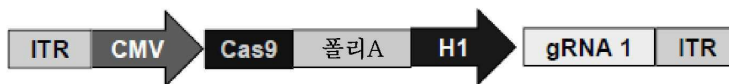
"단일 벡터, 멀티플렉스 CRISPR 시스템"



"이중 벡터, 멀티플렉스 CRISPR 시스템"



"단일 벡터, 단일 gRNA 시스템"



도면40

**SaCas9-NLS (NLS 밀줄표시) (서열 459)**

atgaaaaggaaactacattctggggctggacatcgggattacaagcgtggggatgggattattgactatgaaacaag  
ggacgtgatcgacgcagcagcgtcagactgttcaaggaggcccaacgtggaacaacatgaggacggagaagcaagagg  
gagccagcgcctgaaacgacggagaaggcacagaatccagagggtgaagaaactgctgttcgattacaacctgctg  
accgaccattctgagctgagtggaattaatccttatgaagccagggtgaaaggcctgagtcagaagctgtcagagga  
agagttttccgcagctctgtgacacctggctaagcgcaggaggtgcataacgtcaatgaggtggaagaggacaccg  
gcaacgagctgtctacaagggaacagatctcacgcaatagcaaaagctctggaagagaagtagtcgcagagctgcag  
ctggaacggctgaagaaagatggcgaggtgagagggtcaattaatagggttcaagacaagcgactacgtcaaagaagc  
caagcagctgctgaaagtgcagaaggcttaccaccagctggatcagagcttcatcgatacttatatcgacctgctgg  
agactcggagaacctactatgagggaccaggagaaggggagccccttcggatggaagacatcaaggaatggtacgag  
atgctgattgggacattgcacctattttccagaagagctgagaagcgtcaagtacgcttataacgcagatctgtacaa  
cgccctgaatgacctgaacaacctggtcatcaccagggtgaaaacgagaacctggaatactatgagaagtccaga  
tcacgaaaacgtgtttaagcagaagaaaaagcctacactgaaacagattgctaaggagatcctggtaacgaagag  
gacatcaagggtaccggtgacaagcactggaaaaccagagttcccaatctgaaagtgtatcacgataattaagga  
catcacagcagcgaagaaatcattgagaacgccgaactgctggatcagattgctaagatcctgactatctaccaga  
gctccgagcagatccaggaagagctgactaacctgaacagcgagctgaccaggaagagatcgaacagattagta  
ctgaaggggtacaccggaacacacaacctgtccctgaaagctatcaatctgattctggatgagctgtggcatacaa  
cgacaatcagattgcaactctttaaccggtgaaagctggtcccaaaaagggtggacctgagtcagcagaagagatcc  
caaccacactggtggacgatttctattctgtcaccctggttcaagcggagcttcatccagagcatcaaagtgatcaac  
gccatcatcaagaagtacggcctgcccgaatgatattatcgagctggctagggaagaagaacgacaagcagcaca  
gaagatgatcaatgagatgcagaacgaaaccggcagaccaatgaacgcattgaagagattatccgaactaccggga  
aagagaacgcgaagctgattgaaaaatcaagctgcagcagatgacagagggaagtgctgtattctctgtgag  
gccatccccctggaggacctgctgaacaatccattcaactacgaggtcgatcatattatccccagaagcgtgtcctt  
cgacaattcctttaacaacaaggtgctggtcaagcaggaagagaactctaaaaagggcaataggactcctttccagt  
acctgtctagttcagattccaagatctcttacgaaacctttaaaaagcacattctgaatctggccaaaggaaaggcc  
cgatcagcagaagaccacaaaaggagtagctgtggaagagcgggacatcaacagattctccgtccagaaggattttat  
taaccggaatctggtggacacaagatcgcgtactcgcggcctgatgaatctgtgcgactcctatttccgggtgaaca  
atctggatgtgaaagtcaagtcacatcaacggcgggttcacatcttttctgaggcgcaaatggaagttaaaaaggag  
cgcaacaaagggtacaagcaccatgccgaagatgctctgattatcgcaaatgccgacttcatctttaaggagtggaa  
aaagctggacaagaagccaagaagtgatggagaaccagatgttcgaagagaagcagggccgaatctatgccgaaatcg  
agacagaacaggagtacaaggagattttcatcactcctcaccagatcaagcatatcaaggatttcaaggactacaag  
tactctcaccgggtggataaaaaagcccaacagagagctgatcaatgacacctgtatagtaacagaaaagacgataa  
ggggaataccctgattgtgaacaatctgaacggactgtacgacaaaagataatgacaagctgaaaaagctgatcaaca  
aaagtccccgagaagctgctgatgtaccacatgatcctcagacatacagaaactgaagctgattatggagcagtac  
ggcgacgagaagaacccactgtataagtactatgaagagactgggaactacctgaccaagtatagcaaaaaggataa  
tgccccctgtatcaagaagatcaagtactatgggaacaagctgaatgcccatctggacatcacagacgattacccta  
acagtcgcaacaagggtggtcaagctgtcactgaagccatacagattcgatgtctatctggacaacggcgtgtataaa  
tttgtgactgtcaagaatctggatgtcatcaaaaaggagaactactatgaagtgaatagcaagtgtacgaagaggc  
taaaaagctgaaaaagatttagcaaccaggcagagttcatcgctccttttacaacaacgacctgattaaagtcaatg  
gcgaactgtataggtcatcggggtgaacaatgatctgctgaaccgcattgaagtgaatatgattgacatcacttac  
cgagagtatctggaaaacatgaatgataagcgccccctcgaattatcaaaacaattgcctctaagactcagagtat  
caaaaagtactcaaccgacattctgggaaacctgtatgaggtgaagagcaaaaagcaccctcagattatcaaaaagg  
gcagcggaggcaagcgtcctgctgctactaagaaagctggtcaagctaaagaaaaagaaaggaatcctaccatacagat  
gttccagattacgcttaa

**SaCas9 gRNA (서열 460)**

[프로토스페이서]gttttagtactctggaacagaatctactaaaaacaaggcaaaatgccgtgtttatctcgtcaac  
ttgttgccgagattttttt

도면41

**NmCas9 (NLS 1 밀줄표시 NLS 2 밀줄표시/볼드체, HA 태그 볼드체) (서열 461)**

atgggtgcctaagaagaagagaaaggtggctgccttcaaacctaattcaatcaactacatcctcgccctcgatatcg  
catcgcatccgctcggtggcgatggttagaaattgacgaagaagaaaccccatccgctgattgatttggcgctgc  
gcgtattttgagcgtgccgaagtaccgaaaacagcgagctcccttgccatggcaaggcggttggcgcgagtggtcgc  
cgctgaccgcgctgcgcgccaccgcctgcttcggaccgcgcctattgaaacgcgaaggcgattacaagccgc  
caattttgacgaaaaacggcttgattaaattcctaccgaatacaccatggcaacttcgcgcgagcgcgattagaccgca  
aactgacgcctttagatggctgcgcagctctgttgcatttaatacaaacatcgcggtatttatcgcaacggaaaaac  
gagggcgaaactgcccagataaggagcttgccgctttgcttaaaggcgtagccggcaatgcccatgcccattacagacag  
cgatttccgcacacccggcgaattggctttaaataaatttgagaaagaaagcgccatattccgcaatcagcgagcg  
attattccgcatacgttcagccgcaagatttacaggcgagctgattttgctgtttgaaaaacaaaaagaatttggc  
aatccgcgtatttcaggcgcccttaaagaaggtattgaaacccctactgatgacgcaacgcctgcccgtgcccggga  
tgccgttcaaaaaatgttggggcattgacaccttcgaacgcgcagagcggaagccgctaaaaacacctacacagccg  
aacgttttcatctggctgaccaagctgaacaacctgctgatttttagagcaaggcagcgagcgccattgaccgatacc  
gaagccgcccagctttagagcgctgttgaaacacatcaaaactgacttacgcacaaagccgtaagctgctgggttt  
agaagataccgcctttttcaaaggcttgcgctatggttaaagacaatgccgaagcctcaacattgatggaaatgaagg  
cctaccatgccatcagccgtgactggaaaaaagaggttgaaagacaaaaatccccattaaacctttctcccgaa  
ttacaagacgaaatcggcacggcatttctccctgttcaaaacccgatgaagacattacaggcgcgtcgaagaccgtat  
acagcccgcaatcgttagagcgctgttgaaacacatcagccttcgataaagtctgcctcaaatcttcttgaagcatgtc  
gccgaattgtgctctaatggaacaaggcaaacgttacgatgaagcctgcccgaatctacggagaccattacggc  
aagaagaatcaggaagaaagatttatctgcgcgcgattcccgccgacgaaatccgcaaccccgctcgtcttgccgcg  
cttatctcaagcacgtaaggtcattaacggcggtggtacgctttagcgtccccagctcgtatccatattgaaactg  
acagccggaagttagttaaattcgtttaaagaccgcaagaaattgagaaacgcaagaagaaacgcaagaccgggaa  
aaagccgcccgaatctccgagagtatttccccaatttgtcggagaacccaaatccaagatatttctgaaactgcg  
cctgtacgagcaacacacggcaaatgctgtatttcgggcaagaaatcaacttagccgctcgaacgaaaaaggct  
atgtcgaatcgcacatgcccgttctgcgcacatgggacgacagtttcaacaataaagtactggtattgggc  
agcgaaccccaaaacaaaggcaatcaaaccccttacgaatacttcaacggcaagacaacagccgcaatggcagga  
atttaaagcgcgtgctgaacacagccgttccccgcgcagtaaaaaacaacggattctgctgcaaaaattcgtatgaag  
acggctttaaagaacgcaatctgaacgacacgcgctacgtcaacccgttctcgtgtcaatttgttgccgaccgtatg  
cggctgacaggttaaaggcaagaaacgtgtctttgcatccaacggacaatatacaatctggtgcgcggcttttggg  
attgcgcaaaagtgcgtgcggaaaaacgacccgcatcagccttgagcgcgctcgtcgttgctgctcgcacgttgcca  
tgacgcagaaaattacccgttttgcacgtataaagagatgaacgcgtttgacggtaaaacccatagacaaagaaaca  
ggagaagtgtgtcatcaaaaaacacacttcccacaaccttgggaatttttcgcacaagaagtcatttcgcgtctt  
ggcgaacccggacggcaaacccgaattcgaagaagcgcgataccctagaaaaactgacgacgttgcttgccgaaaaat  
tatcatctcgcgccgaagccgtacacgaatacgttacgcccactgtttgtttcacgcgcgcccacatcggaagatgagc  
gggcaagggcatatggagaccgtcaaatccgccaacgactggacgaaggcgctcagcgtgttgccgctaccgctgac  
acagttaaaactgaaagacttggaaaaaatggtaaatcgggagcgcgaaacctaaagctatacgaagcactgaaagcac  
ggctggaagcacataaagacgatcctgccaagcctttgcccagccgttttacaataacgataaagcagggcaacccg  
accacaacaggtaaaagccgtacgcgtagagcaagtacagaaaacccggcgtatgggtgcgcaaccataacggatttgc  
cgacaacgcaacccatggctgcgcgtagatgtgtttgagaagggcacaagatttatctggtaccgatttacagttggc  
aggtagcgaagggatttttgccgatagggtgtgtgtacaaggaagaaagatgaagaagattggcaacttattgatgat  
agtttcaacttttaattctcattacacccctaattgatatttagtcgaggtataacaaaaaaagctagaatgttttggtta  
ctttgccagctgccatcgaggcacaggttaatatcaatatcgcattcatgatcttgatcataaaattggcaaaaatg  
gaatactggaaggtatcgccgtcaaaacccgctttcattccaaaaataccaaatgacgaactgggcaagaaatc  
agaccatgcccgtctgaaaaaacgcccgcctgtccgt **tacccatacgatgttccagatttacgctgcagctccagcagc**  
**gaagaaaaagaagctgga**ttta

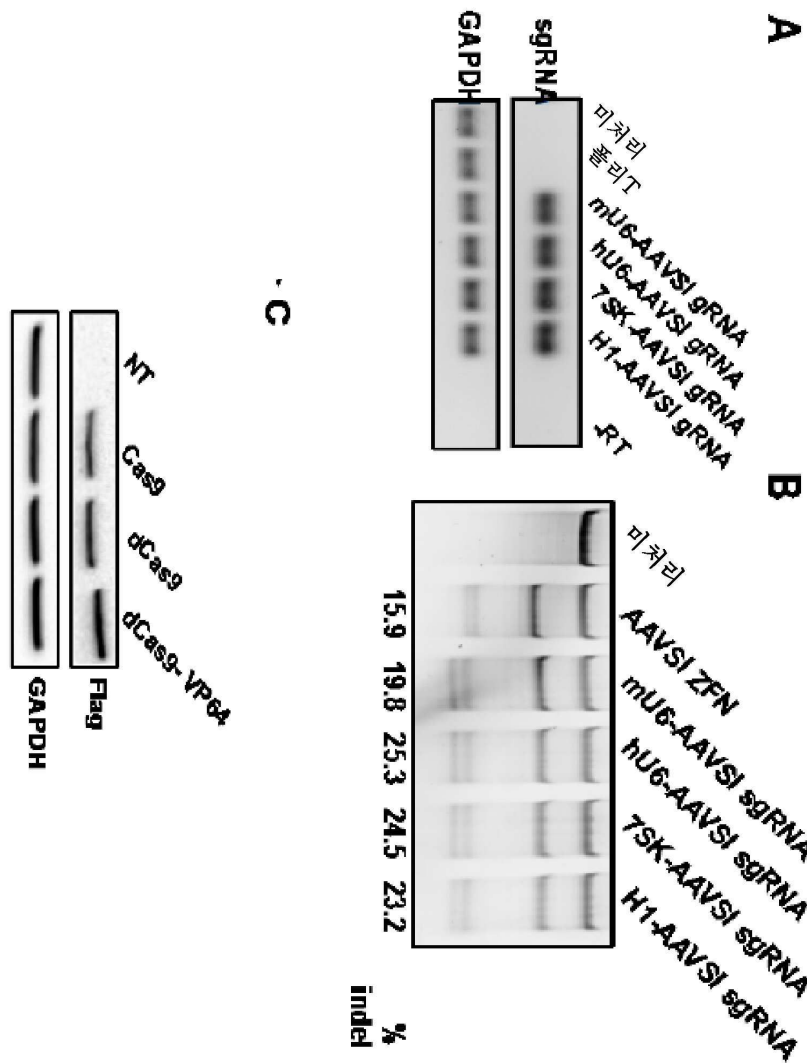
**롬슨 PNAS 2013으로부터의 NmCas9 짧은 헤어핀 (서열 462)**

[프로토스페이스] GTTGTAGCTCCCTTTCTCATTTTCGGAAACGAAATGAGAACCGTTGCTACAATAAGGCCGTCTGA  
AAAGATGTGCGCGCAACGCTCTGCCCTTAAAGCTTCTGCTTTAAGGGGCTTTTTTT

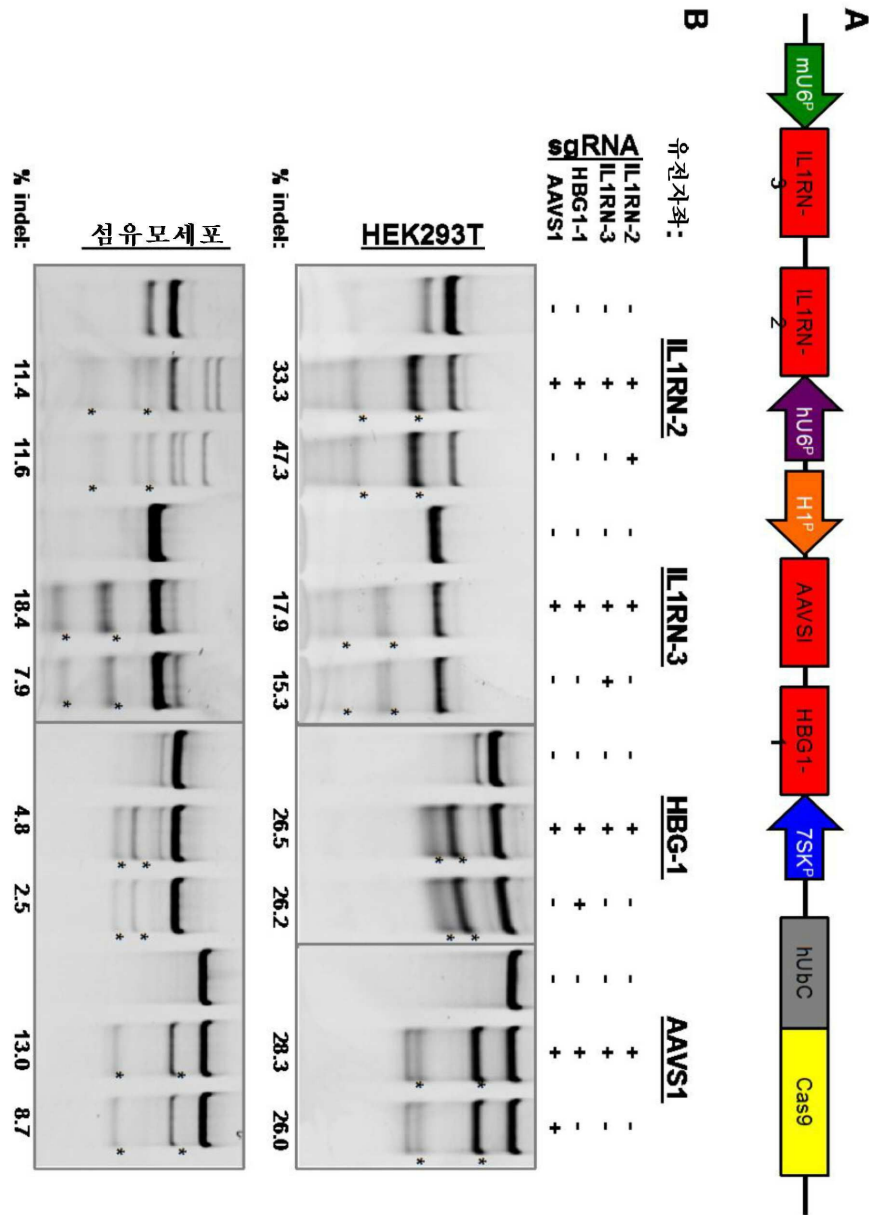
**처치 네이처 바이오테크 2013으로부터의 NmCas9 긴 헤어핀 (서열 463)**

[프로토스페이스] GTTGTAGCTCCCTTTCTCATTTTCGCAGTGCTACAATGAAAATTGTGCGCACTGCGAAATGAGAACCG  
GTTGCTACAATAAGGCCGTCTGAAAAGATGTGCCGCAACGCTCTGCCCTTAAAGCTTCTGCTTTAAGGGGCTTTTTT  
TT

도면42

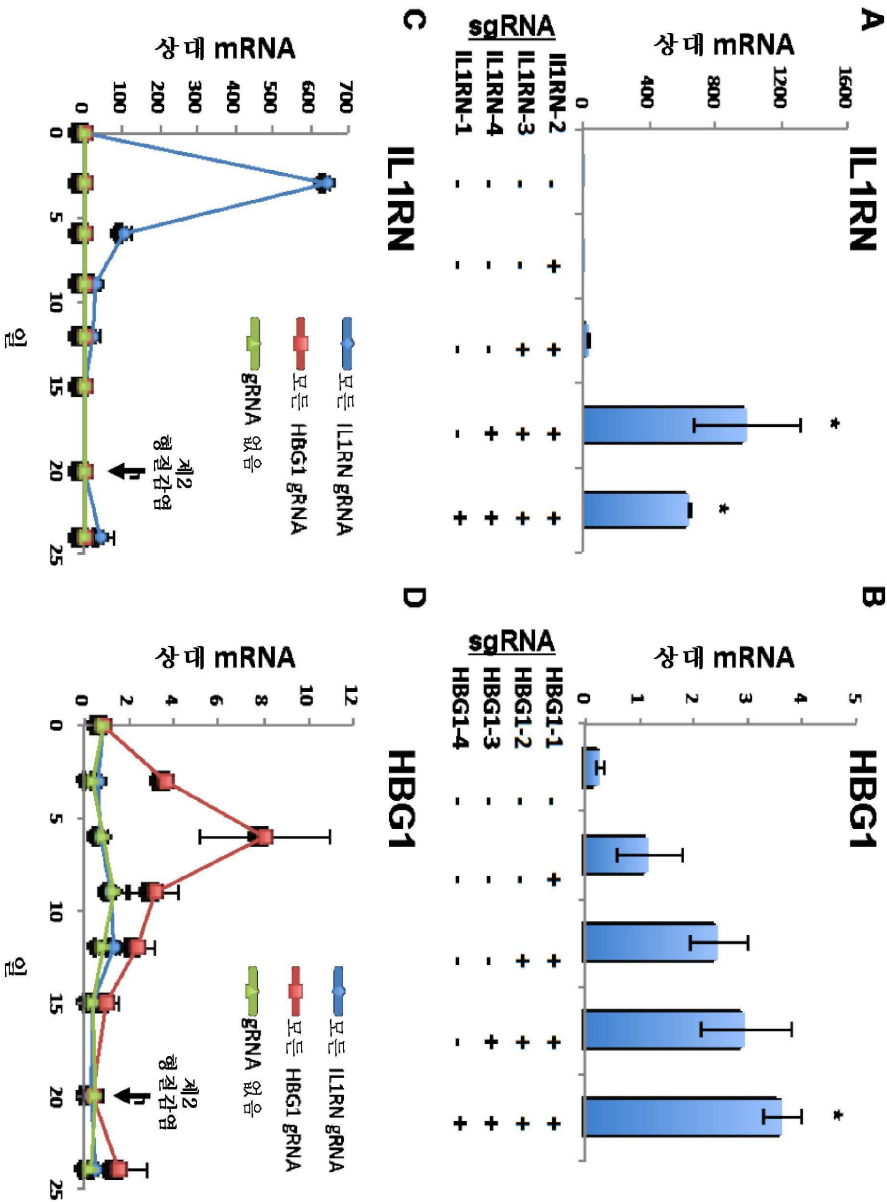




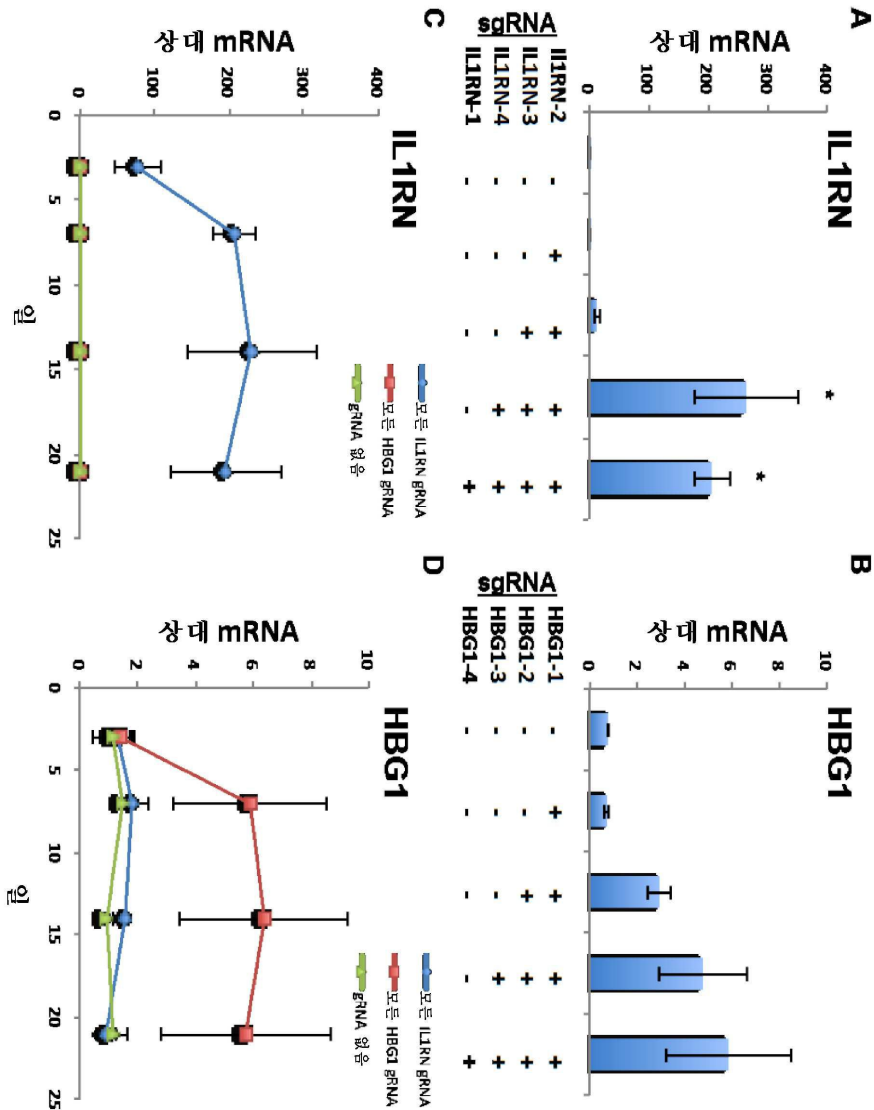




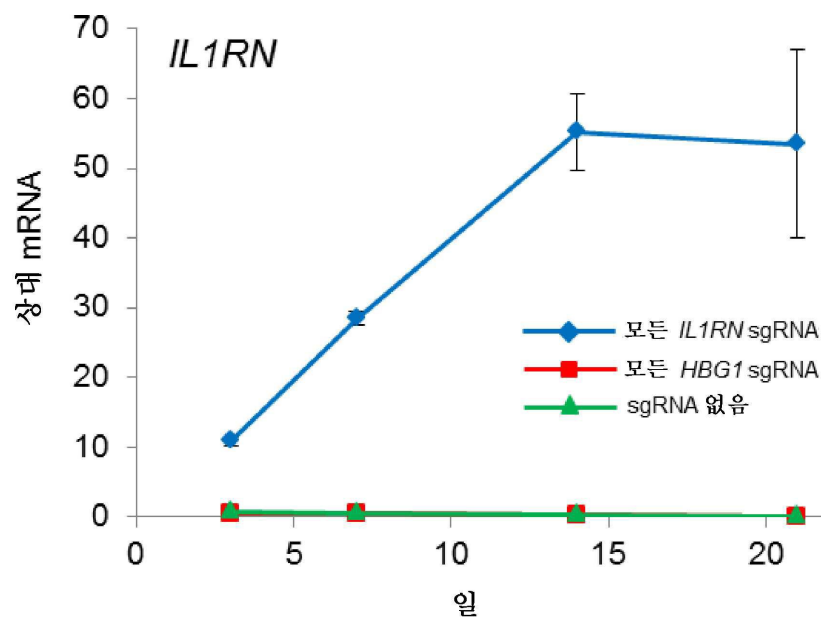
도면45



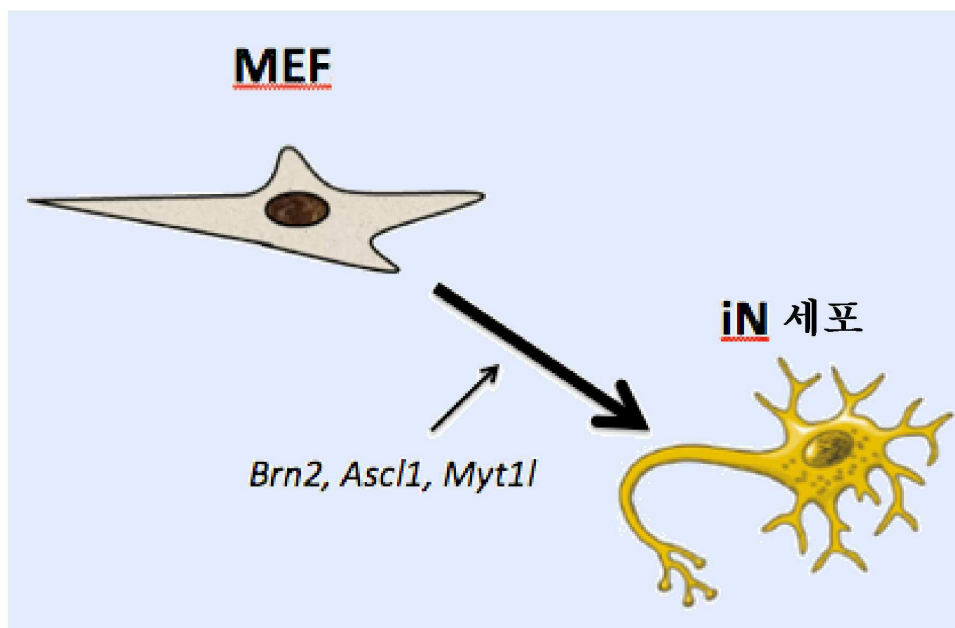
도면46



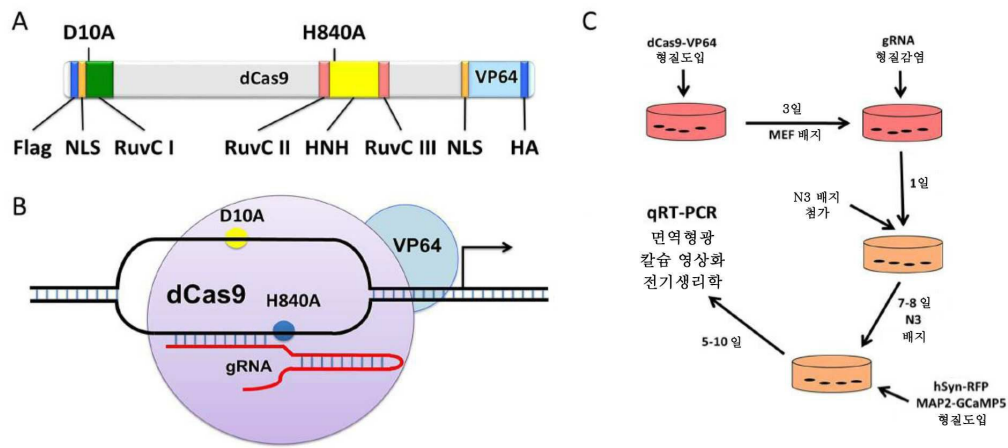
도면47



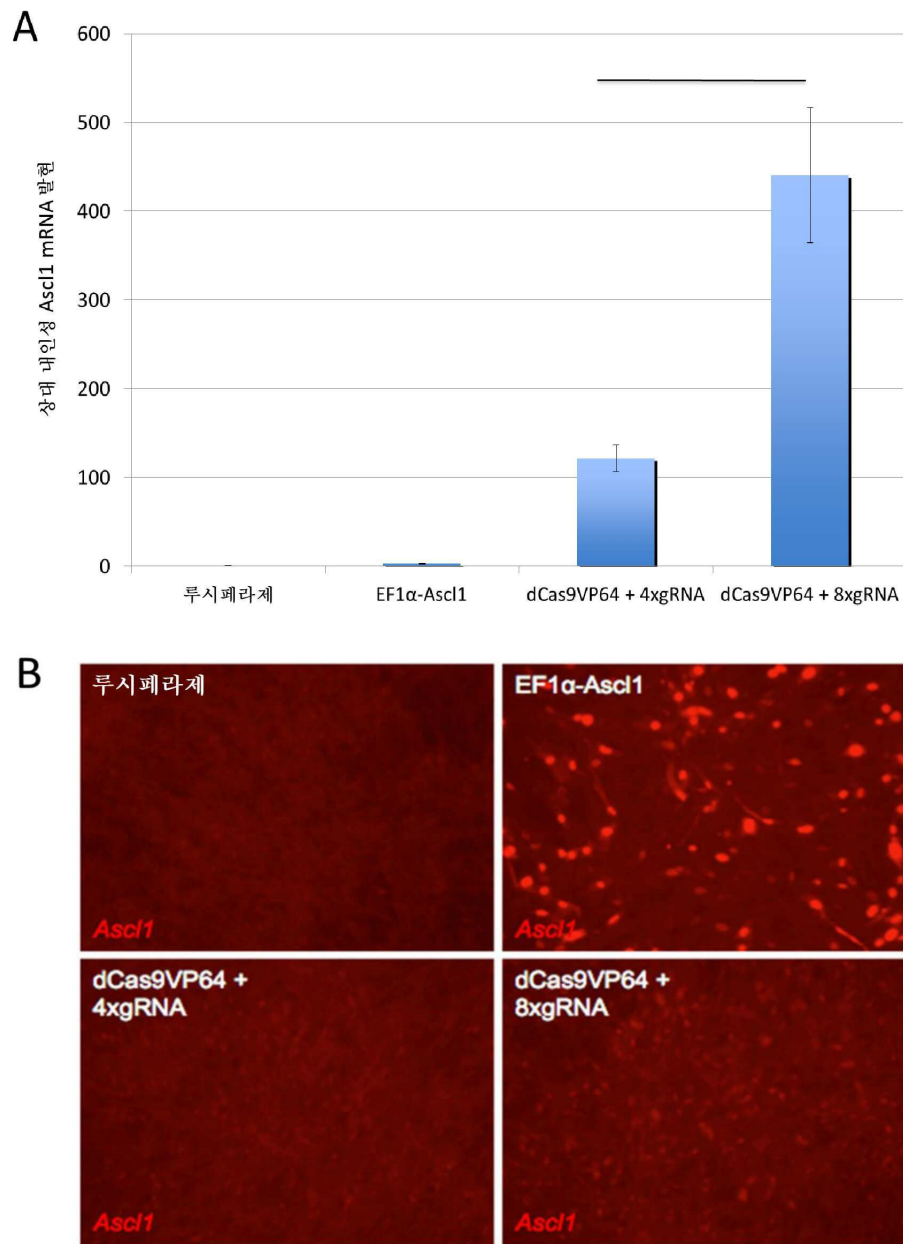
도면48



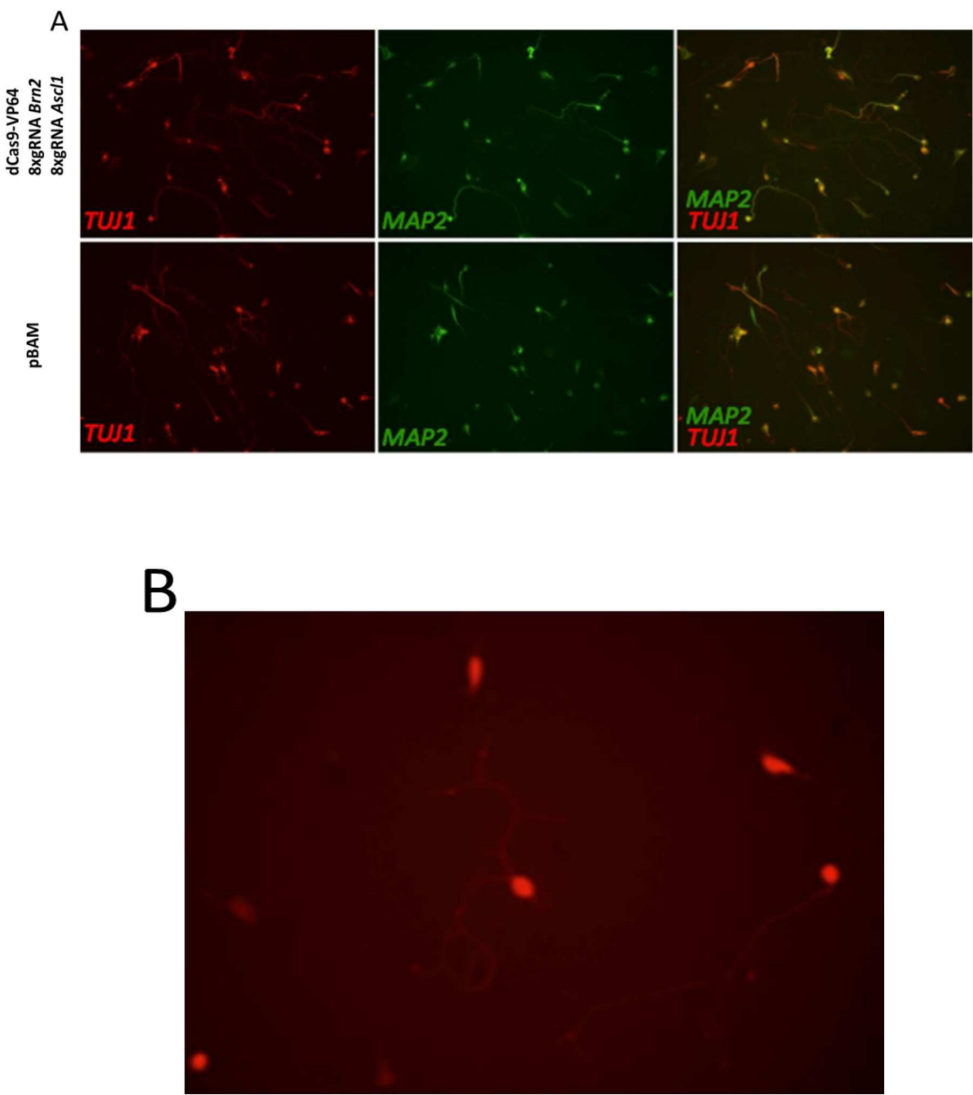
도면49



도면50

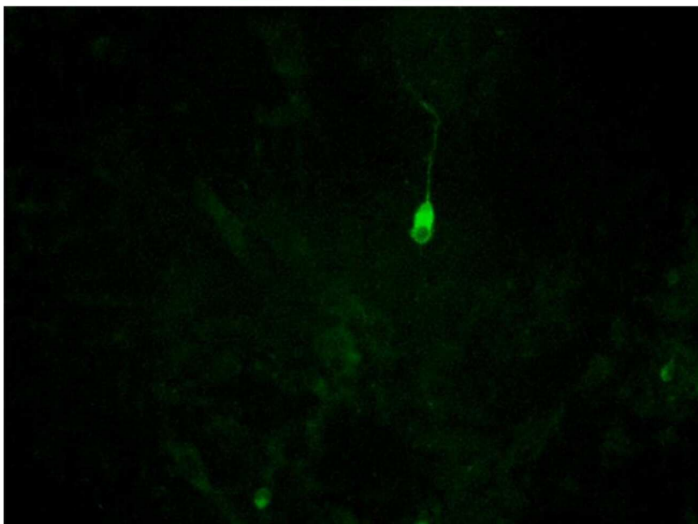
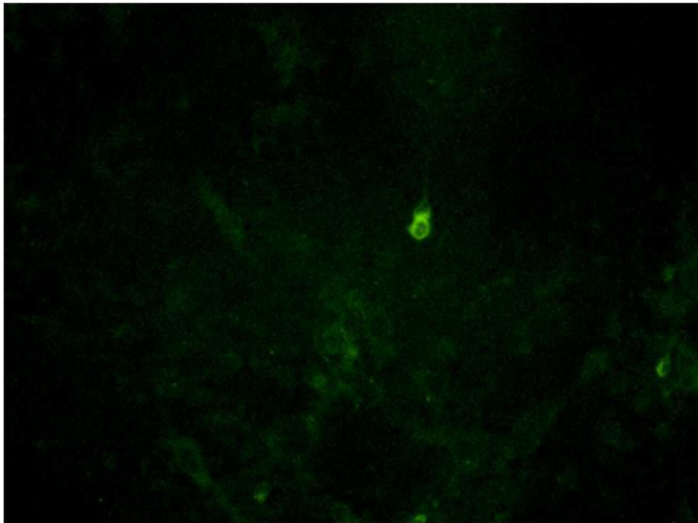


도면51

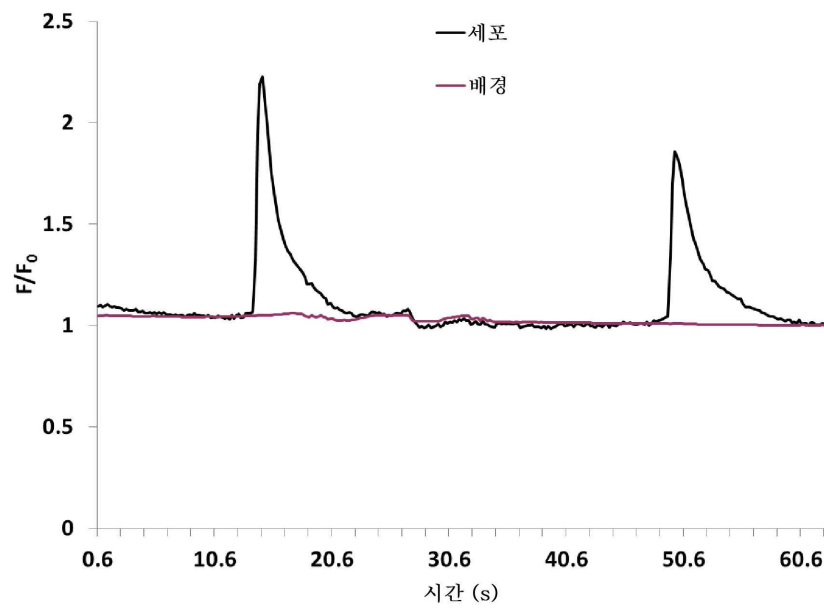




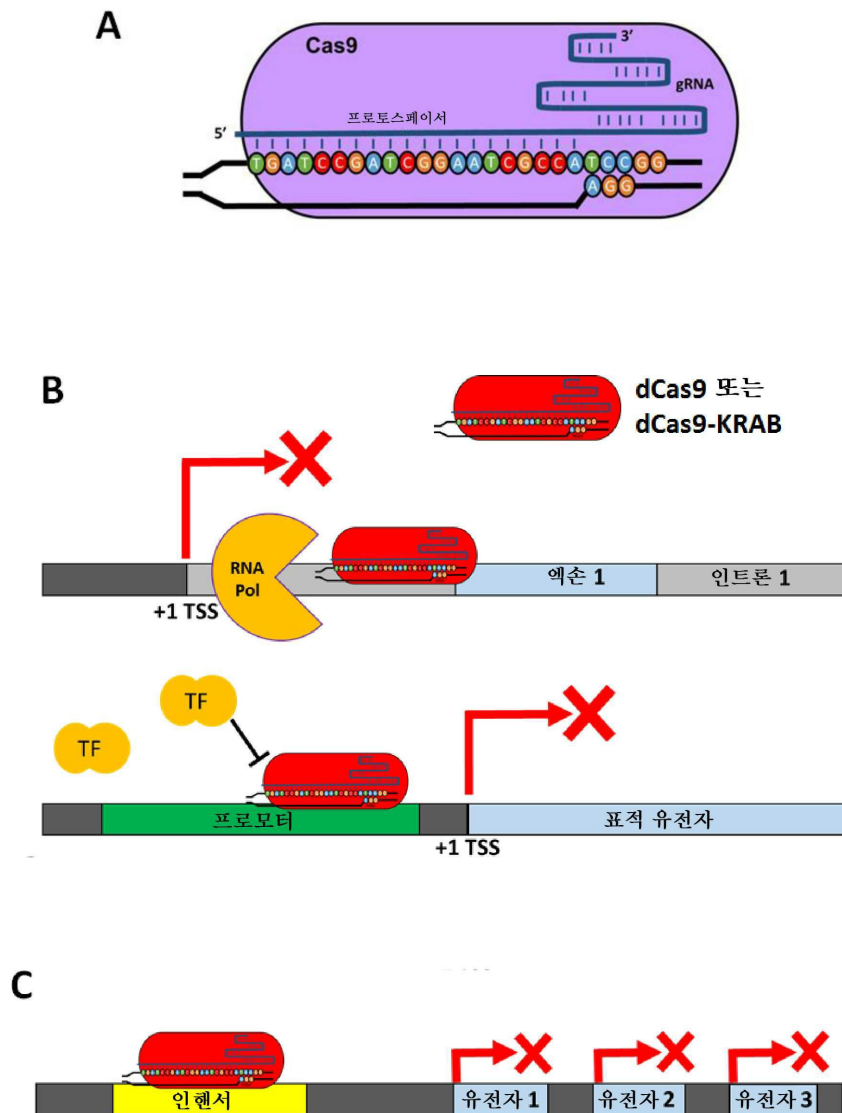
도면52a



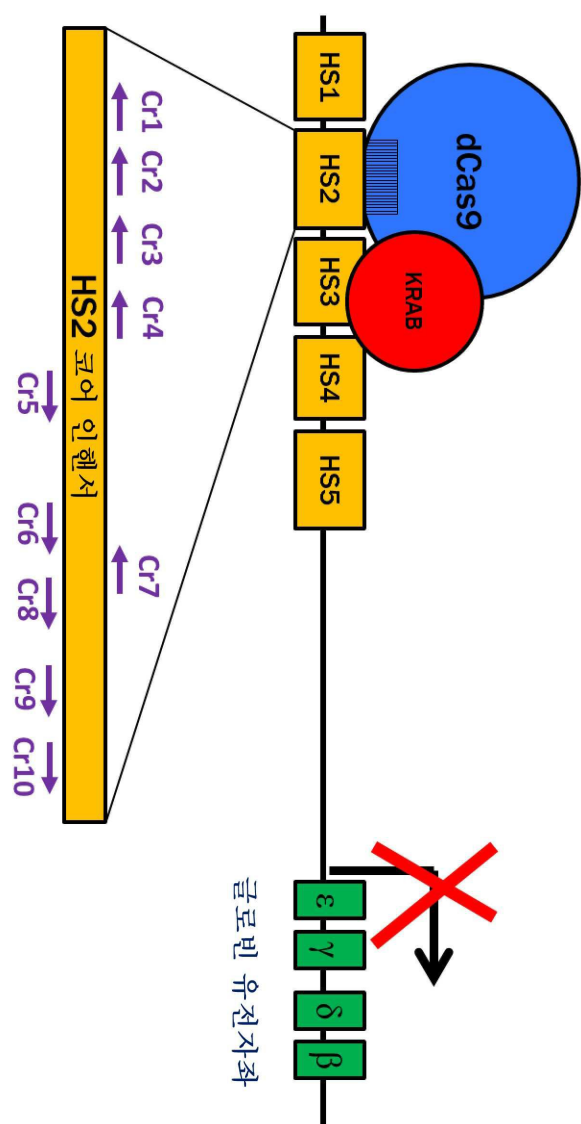
도면52b



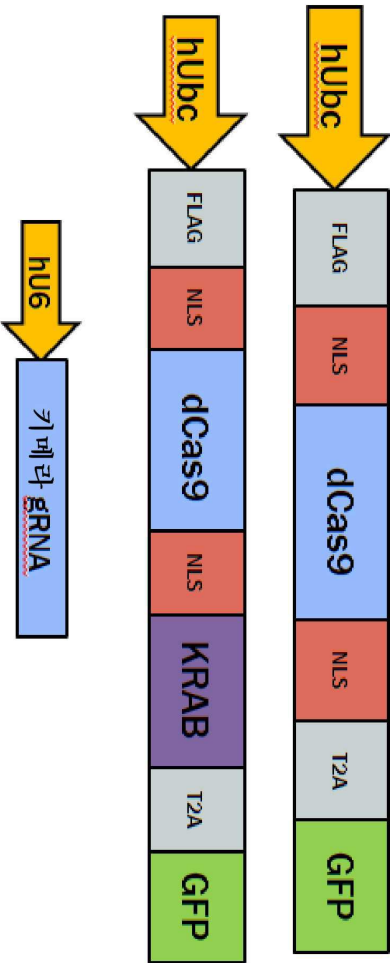
도면53



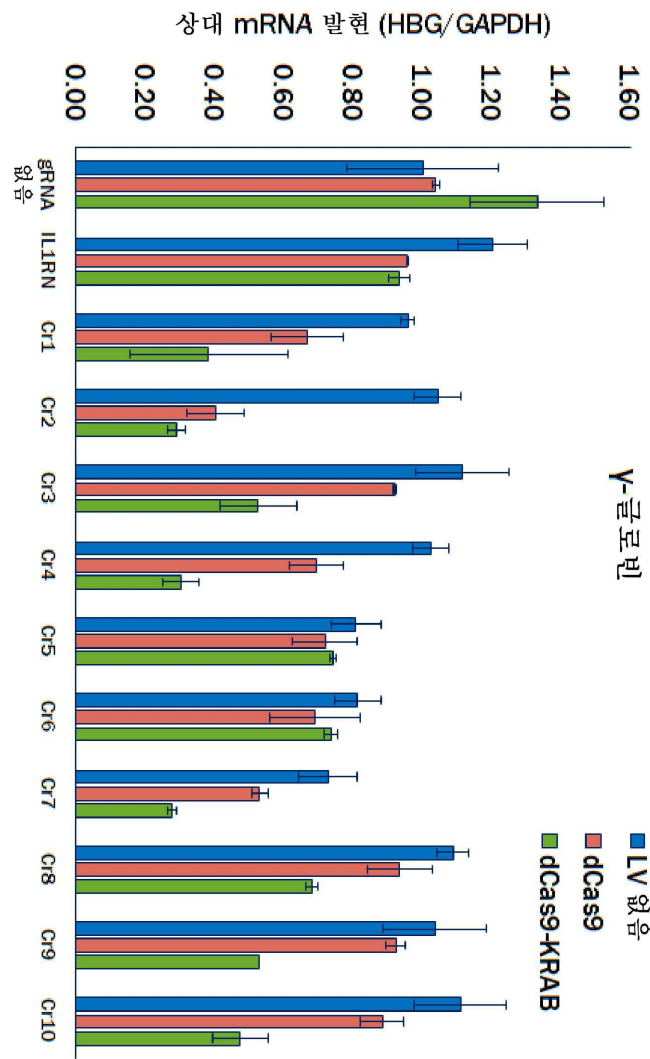
도면54



도면55a

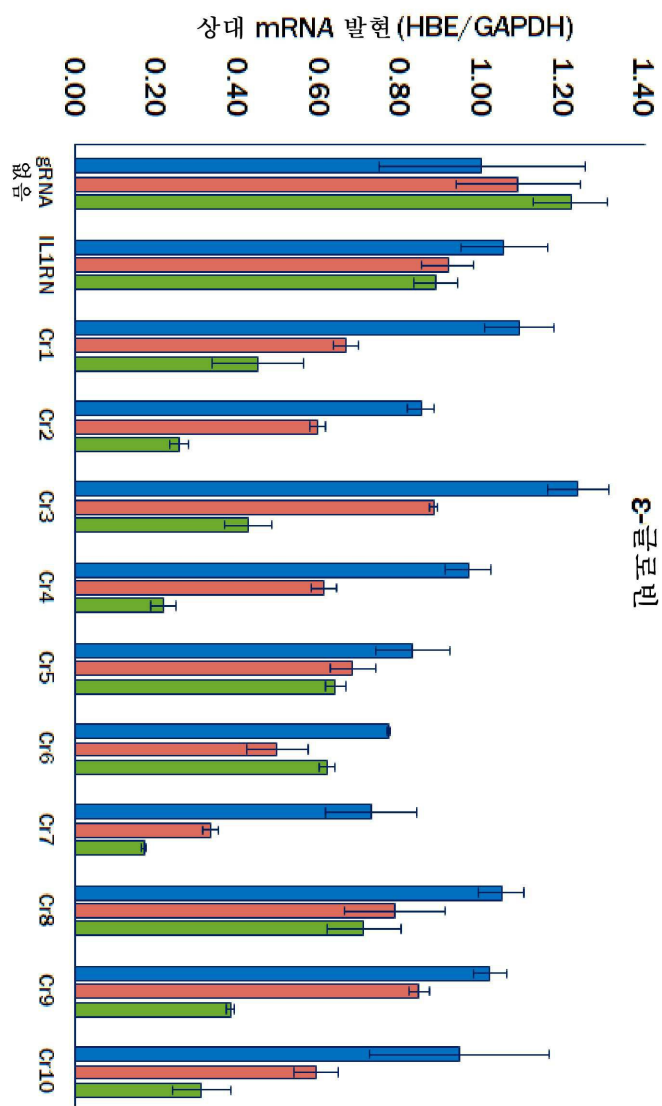


도면55b

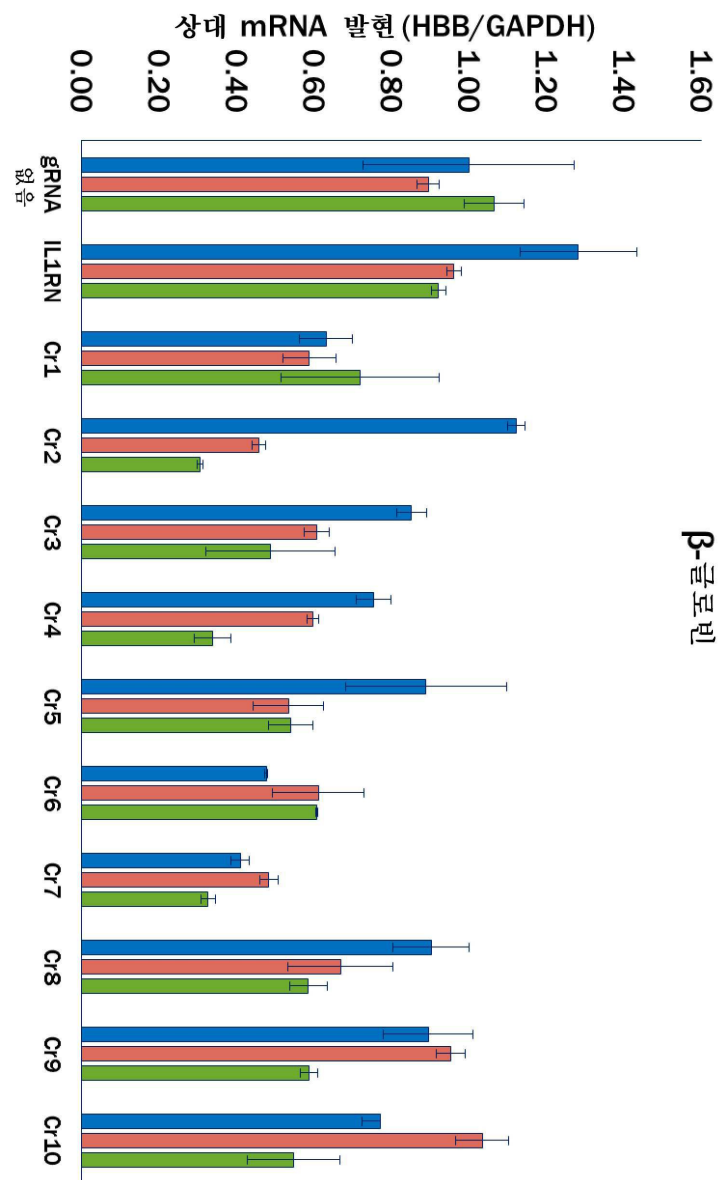




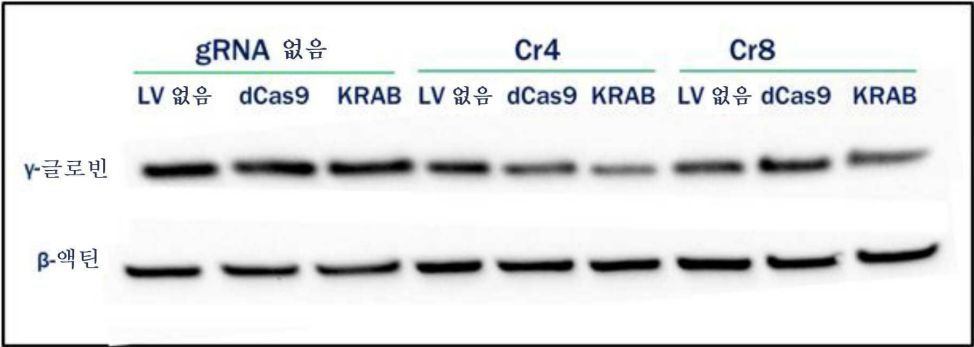
도면55c



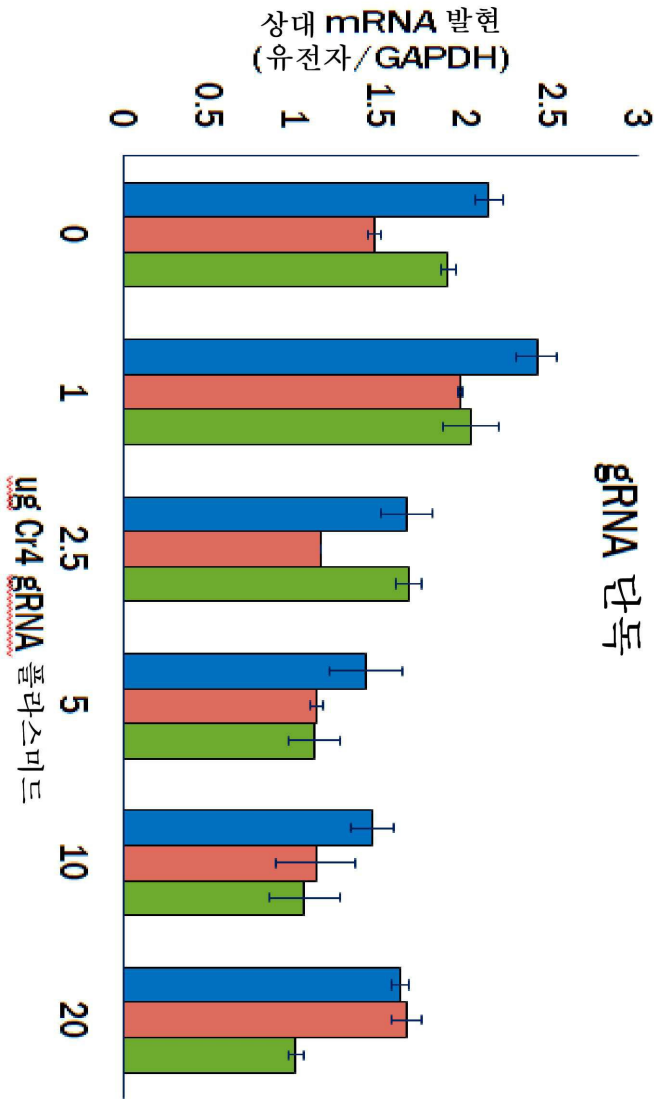
도면55d



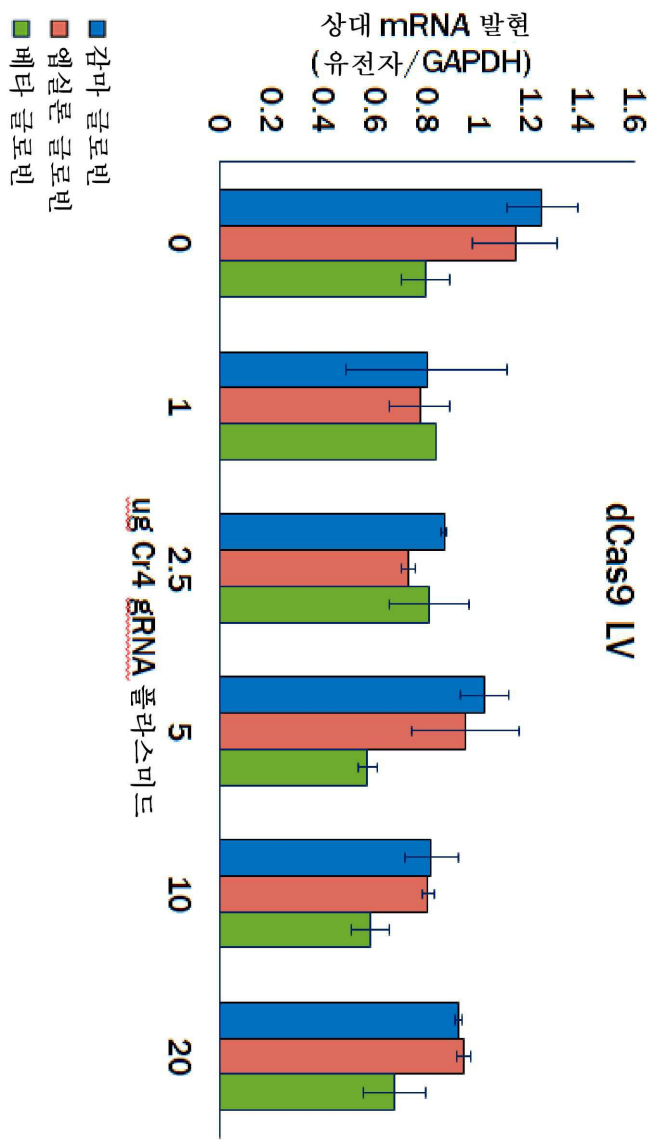
도면55e



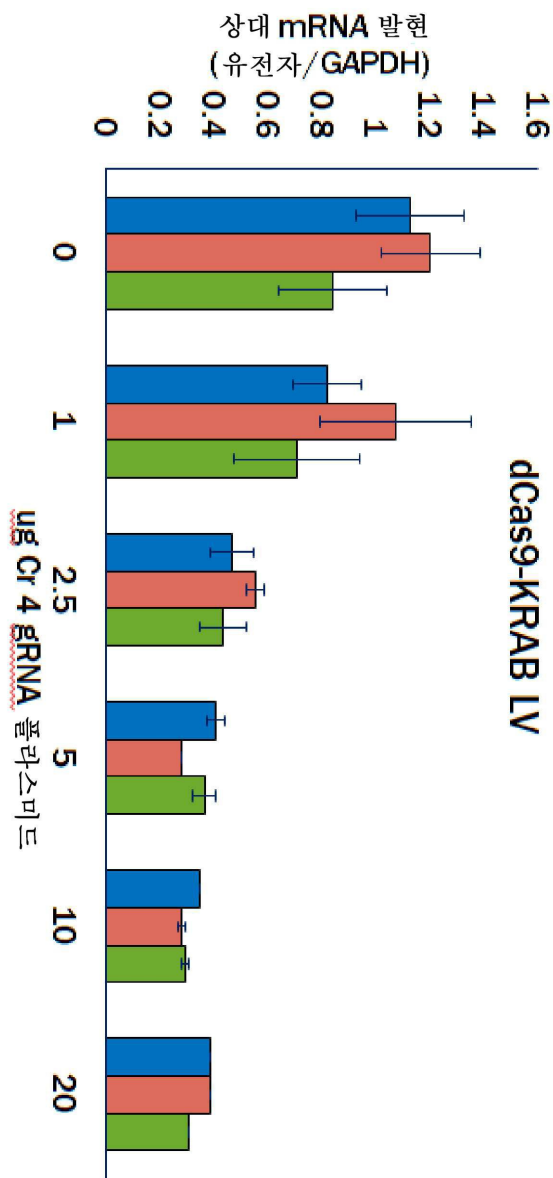
도면56a



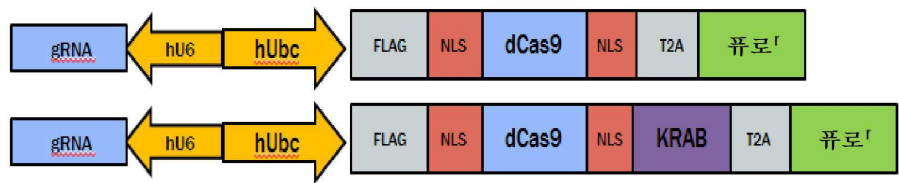
도면56b



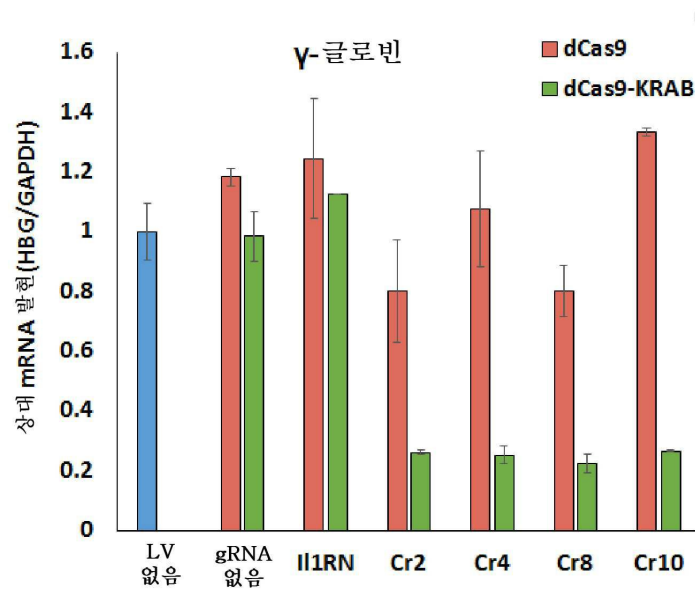
도면56c



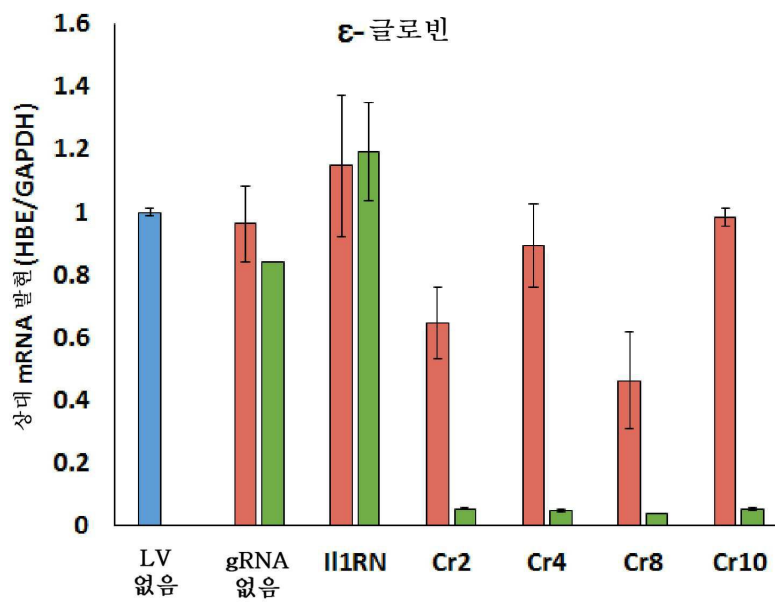
도면57a



도면57b

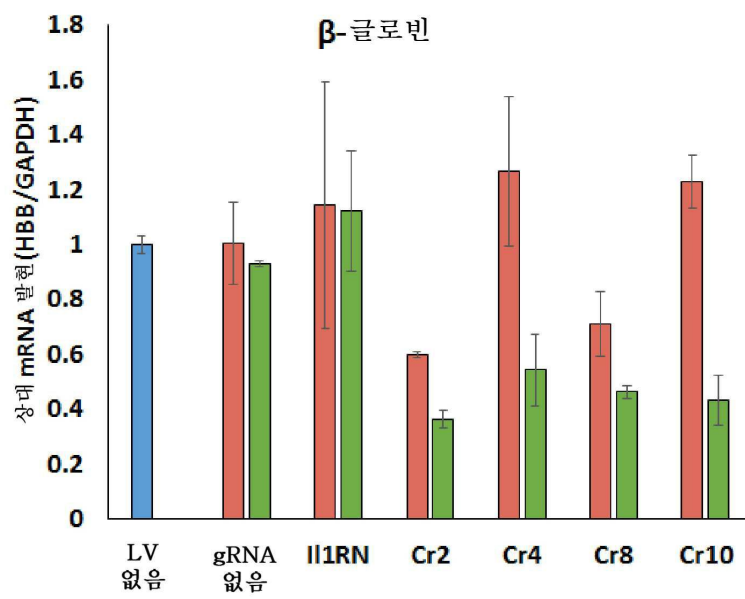


도면57c

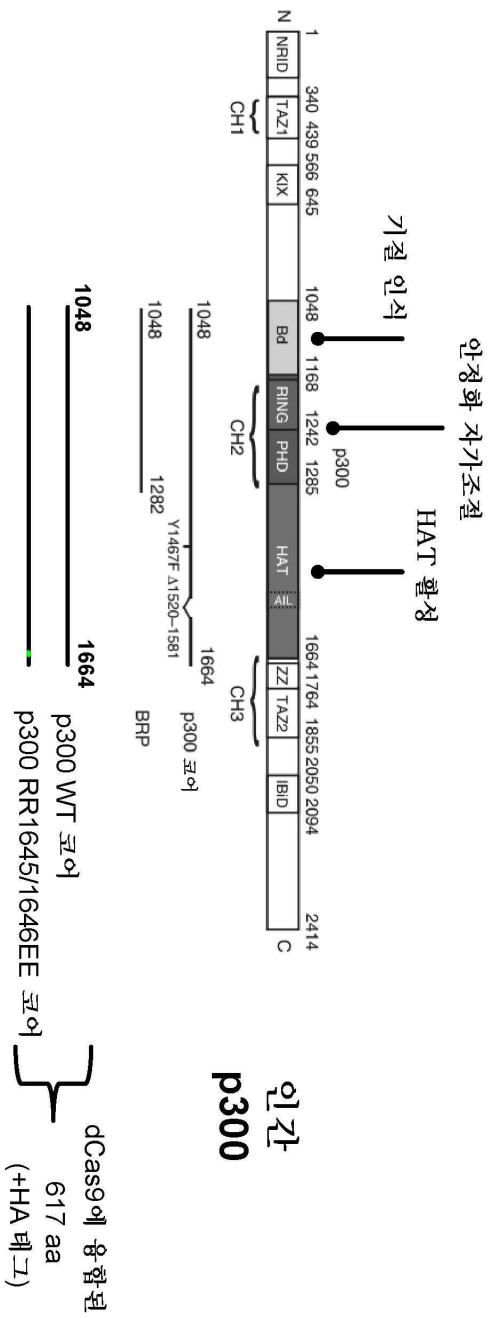




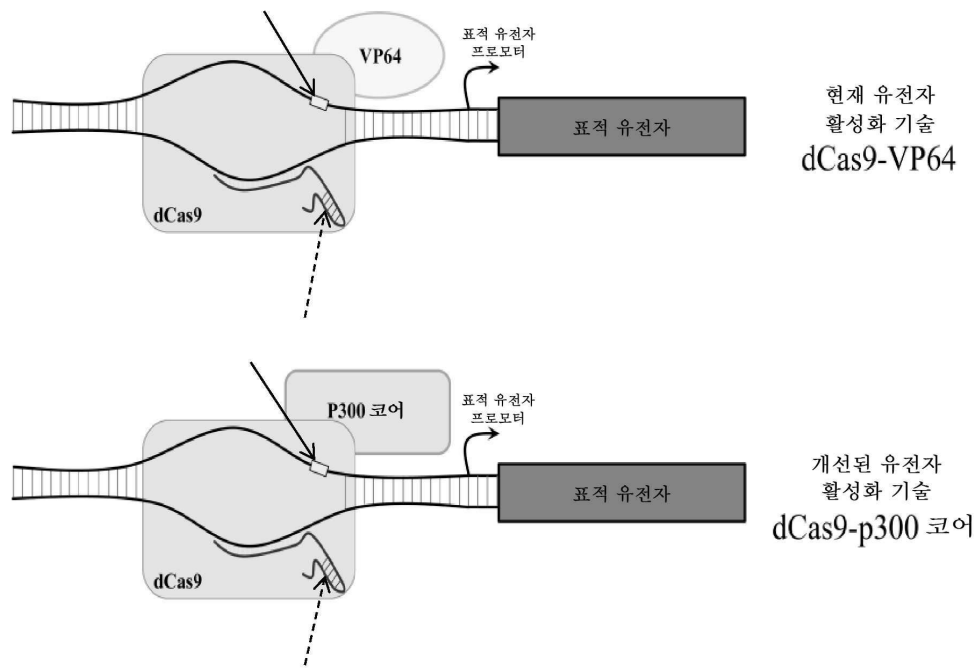
도면57d



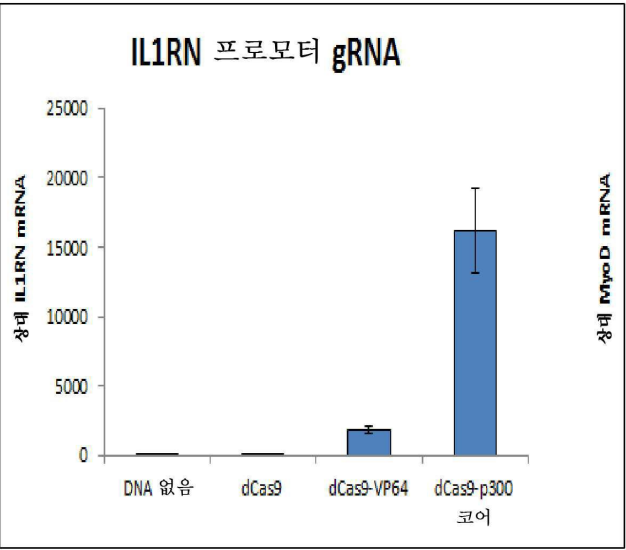
도면58



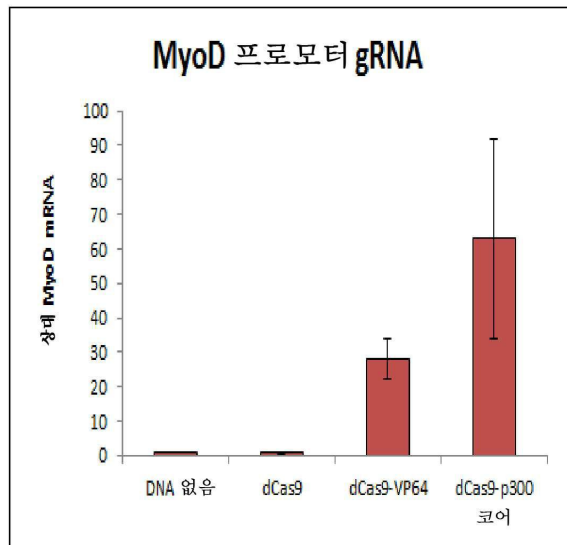
도면59



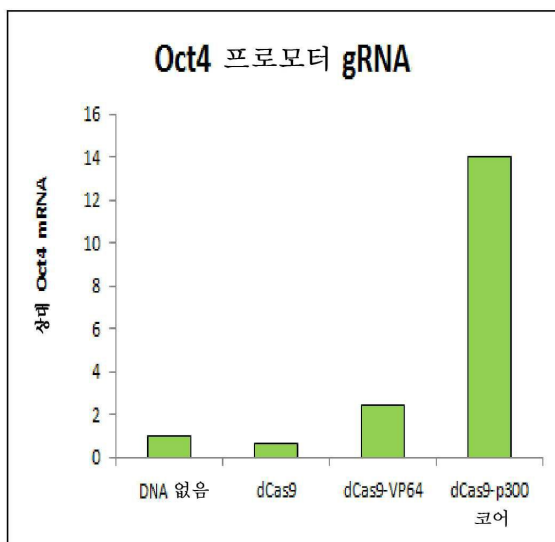
도면60a



도면60b



도면60c



도면61a

i. dCas9

MDYKDHDGDYKDHDIDYKDDDDMAPKKKKVGRGMDKKYSIGLAIGTNSVGWAVITDEY  
 KVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLOEI  
 FSNEMAKVDDSFHRLVESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDS  
 TDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASG  
 VDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLO  
 LSKDTYDDDLNLLAQIGDQYADFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRY  
 DEHHQDLTLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDG  
 TEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILT  
 FRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEVVDKGASAQSFIERMTNFDKNLPNE  
 KVLPKHSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLK  
 EDYFKKIECFDSVEISGVEDRFNASLGTYHDLKIIKDKDFLDNEENEDILEDIVLTTL  
 FEDREMIEERLKYAHLFDDKVMKQLKRRRTGWGRLSRKLINGIRDKQSGKTILDFLKS  
 DGFANRNFMLIHDDSLTFKEDIQKAQVSGQDSLHEHIANLAGSPAIIKKGILQTVKVVD  
 ELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIELGSQILKEHPVENTQ  
 LQNEKLYLYLQNGRDMYVDQELDINRLSDYDVIDHIVPQSFLKDDSIDNKVLRSDKNRG  
 KSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETR  
 QITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAH  
 DAYLNAVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNF  
 FKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNIKKTEVQTGGF  
 SKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKEGKSKKLKSVKELLG  
 ITIMERSSEKNPIDFLEAGYKEVKKDLIIKLPKYSLELENGRKRMLASAGELQKGNE  
 LALPSKYVNFLYLASHYEKLGSPEDNEQKQLFVEQHKHYLDEIEQISEFSKRVILADA  
 NLDKVL SAYNKHRRDKPIREQAENIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDA  
 TLIHQSIITGLYETRIDLSQLGGDPKKKKRVG

아미노산 범례:

"Flag" 에피토프

핵 국제화 서열

스트렙토코쿠스 피오게네스 Cas9 (D10A, H840A)

VP64 이펙터

p300 코어 이펙터

"HA" 에피토프

도면61b

ii. dCas9<sup>VP64</sup>

MDYKDHDGDYKDHDIDYKDDDDKMAPKKKKRVGRGMDKKYSIGLAIGTNSVGWAVITDEY  
 KVPSKKFKVLGNTDRHSIKKNLIGALLPDSGETAEATRLKRTARRRYTRRKNRICYLQEI  
 FSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDS  
 TDKADLRILIYLAHAMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASG  
 VDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLO  
 LSKDTYDDDLDNLLAQIGDQYADFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRY  
 DEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDG  
 TEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILT  
 FRIPYYVGPLARGNSRFWMTRKSEETITPWNFEVVDKGASAQSFIERMTNFDKNLPNE  
 KVLPKHSLLEYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLK  
 EDYFKKIECFDSVEISGVEDRFNASLGTYHDLKIIKDKDFLDNEENEDILEDIVLTTLTL  
 FEDREMIERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTLIDFLKS  
 DGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQDSLHEHIANLAGSPAIIKKGILQTVKVV  
 ELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQ  
 LQNEKLYLYLQNGRDMYVDQELDINRLSDYDVAIVPQSFLKDDSIDNKVLTNRSDKNRG  
 KSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETR  
 QITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAH  
 DAYLNAVVGTAIIKKYKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNF  
 FKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNIKKTEVQTGGF  
 SKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVAKVEKGSKKLKSVKELLG  
 ITIMERSSEFKNPIDFLEAKGYKEVKDLIIKLPKYSLFELENGRKRMLASAGELQKGNE  
 LALPSKYVNFYLYLASHYEKLKGSPEDEQQLFVEQHKHYLDEIIIEQISEFSKRVILADA  
 NLDKVL SAYNKHDKPIREQAENI IHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDA  
 TLIHQSITGLYETRIDLSQLGGDPIAGSKASPKKKRKVGRADALDDFDLMLGSDALDDF  
 DLDMLGSDALDDFDLMLGSDALDDFDLMLINYPYDVPDYAS

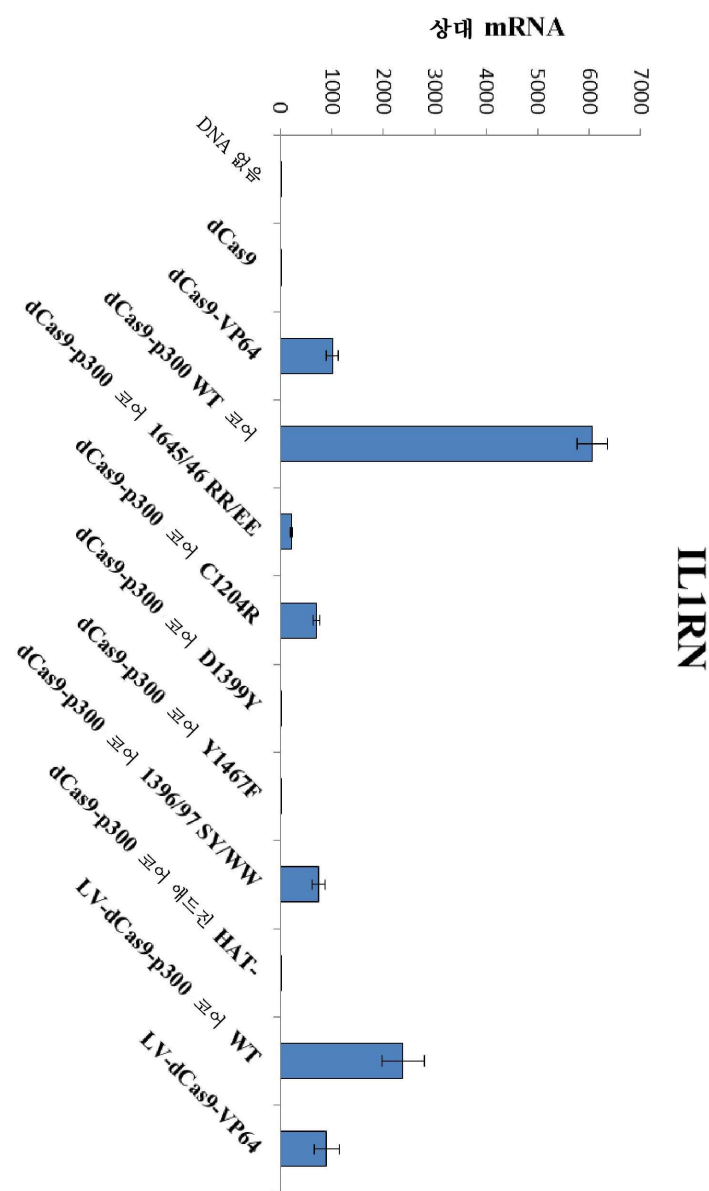


도면61c

iii. dCas9<sup>p300</sup> 코어

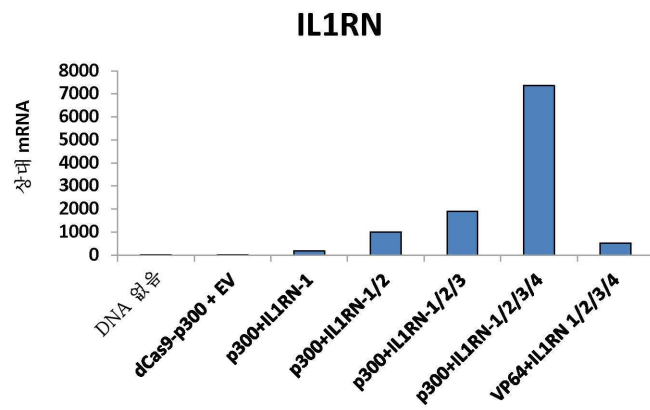
MDYKDHDGDYKDHDIDYKDDDDKMAPKKKKRVGRGMDKKYSIGLAIGTNSVGWAVITDEY  
 KVPSKKFKVLGNTDRHS IKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEI  
FSNEMAKVDDSFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDS  
TDKADLRILIYLAHAMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASG  
VDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQ  
LSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRY  
DEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDG  
TEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILT  
FRIPYVYVGPLARGNSRFAMTRKSEETITPWNFEVVDKGASQSFIERMTNFDKNLPNE  
KVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLK  
EDYFKKIECFDSVEISGVEDRFNASLGTYHDLKIIKDKDFLDNEENEDILEDIVLTTLTL  
FEDREMIERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTIIDFLKS  
DGFANRNFMLIHDDSLTFKEDIQKAQVSGQDSLHEHIANLAGSPAIIKKGILQTVKVVVD  
ELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQ  
LQNEKLYLYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLRSDKNRG  
KSDNVPSSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELKAGFIKQQLVETR  
QITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAH  
DAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEQIEGKATAKYFFYSNIMNF  
FKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNIKKTEVQTGGF  
SKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVAKVEKGKSKKLKSVKELLG  
ITIMERSSEFEKNPIDFLEAGKYKEVKKDLIIKLPKYSLEFLENGRKRMLASAGELQKGNE  
LALPSKYVNFYLYLASHYEKLKGSPEDEQKQLFVEQHKHYLDEIEQISEFSKRVILADA  
NLDKVL SAYNKHDKPIREQAENIHLFTLTNLGAPAAFKYFDTTIDRKRYSTKEVLDA  
TLIHQSIGLYETRIDLSQLGGDPIAGSKASPKKKRKVGRAIFKPEELRQALMPTLEALY  
RQDPESLPFRQPVDPQLLGPDIYFDIVKSPMDLSTIKRKLDTGOYQEPWOYVDDIWL MFN  
NAWLYNRKTSRVYKYCSKLESEVFEQEI DPVMQSLGYCCGRKLEFSQTLCCYKQLCTIP  
RDATYYSYONRYHFCEKCFNEIQGESVSLGDDPSQOPOTTINKEOFSKRKNDTLDPELFVE  
CTECGRKMHOICVLHHEIIWPAGFVCDGCLKKSARTRKENKFSKRLPSTRLGTFLENRV  
NDFLRRONHPESGEVTVRVVHASDKTVEVKPGMKARFVDSGEMAESFPYRTKALFAFEEI  
DGVDLCCFFGMHVQEQYGSDCPPPNQRRVYISYLDVHFFRPKCLRTAVYHEILIGYLEYVK  
KLGYTTGHIWACPPSEGDDYIFHCHPPDQIKPKPKRLQEWYKMLDKAVSERIVHDYKDI  
FKQATEDRLTSAKELPYFEGDFWPNVLEESIKELEQEEEEERKREENTSNESTDVTKGDSK  
NAKKKNKKTSKNKSSLSRGNKKKPGMPNVSNDSLQKLYATMEKHKEVFFVIRLIAGPAA  
NSLPPIVDPDPLIPCDLMDGRDAFLT LARDKHLEFSLLRRAQWSTMCMLELHTQSODYP  
YDVFDYAS

도면62

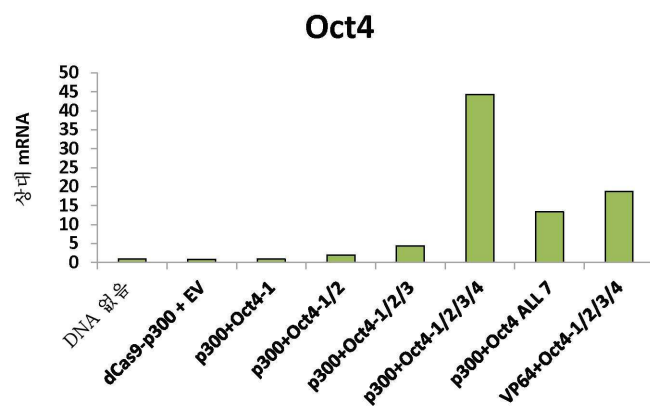


도면63

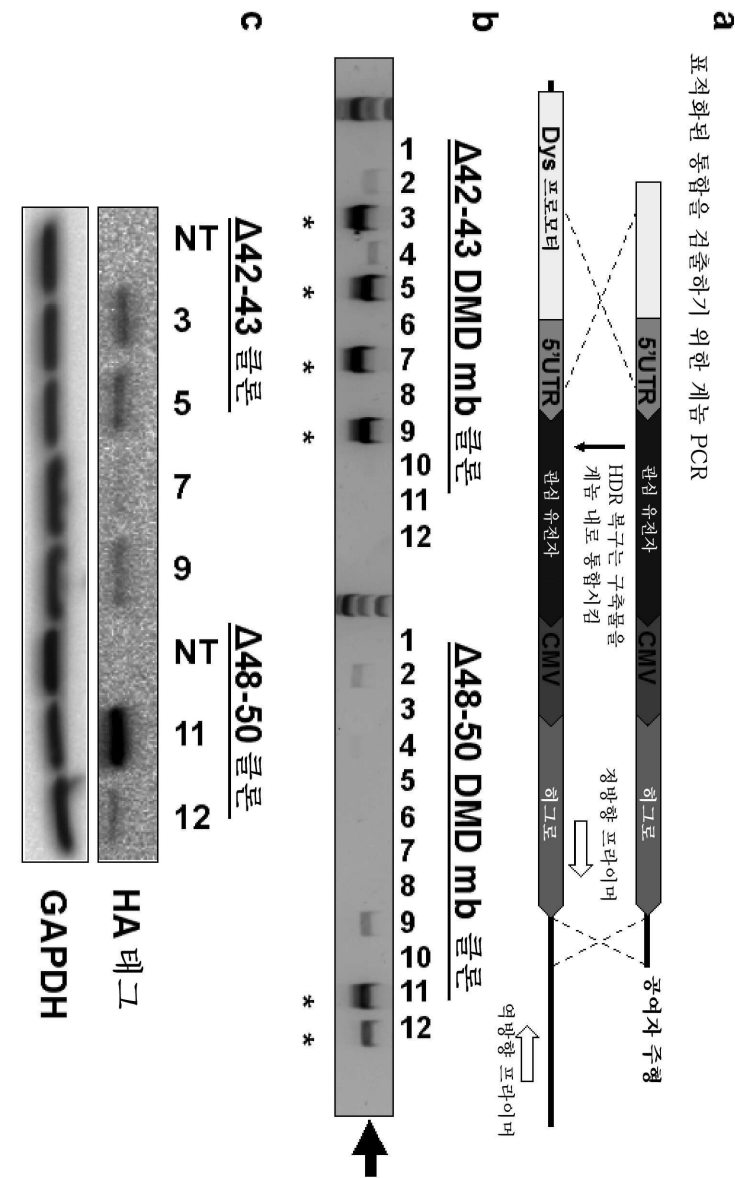
A



B



도면64



서열목록

SEQUENCE LISTING

<110> DUKE UNIVERSITY

<120> RNA-GUIDED GENE EDITING AND GENE REGULATION

<130> 028193-9164-US03

<140> 14/895,316

<141> 2015-12-02

<150> PCT/US2014/041190

<151> 2014-06-05

<150> 61/831,481

<151> 2013-06-05

<150> 61/839,127

<151> 2013-06-25

<150> 61/904,911

<151> 2013-11-15

<150> 61/967,466

<151> 2014-03-19

<150> 61/981,575

<151> 2014-04-18

<160> 625

<170> PatentIn version 3.5

<210> 1

<211> 1483

<212> PRT

<213> Artificial Sequence

<220><223>

Description of Artificial Sequence: Synthetic  
polypeptide

<400> 1

Met Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp His Asp Ile Asp

1 5 10 15

Tyr Lys Asp Asp Asp Asp Lys Met Ala Pro Lys Lys Lys Arg Lys Val

20 25 30

Gly Arg Gly Met Asp Lys Lys Tyr Ser Ile Gly Leu Ala Ile Gly Thr

35 40 45

Asn Ser Val Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser

50 55 60

Lys Lys Phe Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys

65 70 75 80

Asn Leu Ile Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala

85 90 95

Thr Arg Leu Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn

100 105 110

Arg Ile Cys Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val

115 120 125  
Asp Asp Ser Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu

130 135 140  
Asp Lys Lys His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu  
145 150 155 160

Val Ala Tyr His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys  
165 170 175  
Leu Val Asp Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala

180 185 190  
Leu Ala His Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp  
195 200 205

Leu Asn Pro Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val  
210 215 220

Gln Thr Tyr Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly  
225 230 235 240  
Val Asp Ala Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg

245 250 255  
Leu Glu Asn Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu  
260 265 270

Phe Gly Asn Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys  
275 280 285

Ser Asn Phe Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp  
290 295 300

Thr Tyr Asp Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln

305 310 315 320  
Tyr Ala Asp Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu  
325 330 335

Leu Ser Asp Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu  
340 345 350

Ser Ala Ser Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr

355 360 365  
Leu Leu Lys Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu

370 375 380  
Ile Phe Phe Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly  
385 390 395 400

Gly Ala Ser Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu  
405 410 415

Lys Met Asp Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp  
420 425 430

Leu Leu Arg Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln

435 440 445  
Ile His Leu Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe  
450 455 460

Tyr Pro Phe Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr  
465 470 475 480

Phe Arg Ile Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg  
485 490 495

Phe Ala Trp Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn

500 505 510  
Phe Glu Glu Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu  
515 520 525

Arg Met Thr Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro  
530 535 540

Lys His Ser Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr  
545 550 555 560

Lys Val Lys Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser

565 570 575  
Gly Glu Gln Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg  
580 585 590

Lys Val Thr Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu  
595 600 605



Cys Phe Asp Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala  
 610 615 620  
 Ser Leu Gly Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp  
  
 625 630 635 640  
 Phe Leu Asp Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu  
 645 650 655  
 Thr Leu Thr Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys  
 660 665 670  
 Thr Tyr Ala His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg  
 675 680 685  
 Arg Arg Tyr Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly  
  
 690 695 700  
 Ile Arg Asp Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser  
 705 710 715 720  
 Asp Gly Phe Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser  
 725 730 735  
 Leu Thr Phe Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly  
 740 745 750  
 Asp Ser Leu His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile  
  
 755 760 765  
 Lys Lys Gly Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys  
 770 775 780  
 Val Met Gly Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg  
 785 790 795 800  
 Glu Asn Gln Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met  
 805 810 815  
 Lys Arg Ile Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys  
  
 820 825 830  
 Glu His Pro Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu  
 835 840 845  
 Tyr Tyr Leu Gln Asn Gly Arg Asp Met Tyr Val Asp Gln Glu Leu Asp

850                      855                      860  
 Ile Asn Arg Leu Ser Asp Tyr Asp Val Asp Ala Ile Val Pro Gln Ser  
 865                      870                      875                      880  
 Phe Leu Lys Asp Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser Asp  
  
                     885                      890                      895  
 Lys Asn Arg Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys  
                     900                      905                      910  
 Lys Met Lys Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr  
                     915                      920                      925  
 Gln Arg Lys Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser  
                     930                      935                      940  
 Glu Leu Asp Lys Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr Arg  
  
 945                      950                      955                      960  
 Gln Ile Thr Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr  
                     965                      970                      975  
 Lys Tyr Asp Glu Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr  
                     980                      985                      990  
 Leu Lys Ser Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr  
                     995                      1000                      1005  
 Lys Val Arg Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr  
  
 1010                      1015                      1020  
 Leu Asn Ala Val Val Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys  
                     1025                      1030                      1035  
 Leu Glu Ser Glu Phe Val Tyr Gly Asp Tyr Lys Val Tyr Asp Val  
                     1040                      1045                      1050  
 Arg Lys Met Ile Ala Lys Ser Glu Gln Glu Ile Gly Lys Ala Thr  
                     1055                      1060                      1065  
 Ala Lys Tyr Phe Phe Tyr Ser Asn Ile Met Asn Phe Phe Lys Thr  
  
 1070                      1075                      1080  
 Glu Ile Thr Leu Ala Asn Gly Glu Ile Arg Lys Arg Pro Leu Ile  
                     1085                      1090                      1095

Glu Thr	Asn Gly	Glu Thr	Gly	Glu Ile	Val Trp	Asp	Lys Gly	Arg
1100			1105			1110		
Asp Phe	Ala Thr	Val Arg	Lys	Val Leu	Ser Met	Pro	Gln Val	Asn
1115			1120			1125		
Ile Val	Lys Lys	Thr Glu	Val	Gln Thr	Gly Gly	Phe	Ser Lys	Glu
1130			1135			1140		
Ser Ile	Leu Pro	Lys Arg	Asn	Ser Asp	Lys Leu	Ile	Ala Arg	Lys
1145			1150			1155		
Lys Asp	Trp Asp	Pro Lys	Lys	Tyr Gly	Gly Phe	Asp	Ser Pro	Thr
1160			1165			1170		
Val Ala	Tyr Ser	Val Leu	Val	Val Ala	Lys Val	Glu	Lys Gly	Lys
1175			1180			1185		
Ser Lys	Lys Leu	Lys Ser	Val	Lys Glu	Leu Leu	Gly	Ile Thr	Ile
1190			1195			1200		
Met Glu	Arg Ser	Ser Phe	Glu	Lys Asn	Pro Ile	Asp	Phe Leu	Glu
1205			1210			1215		
Ala Lys	Gly Tyr	Lys Glu	Val	Lys Lys	Asp Leu	Ile	Ile Lys	Leu
1220			1225			1230		
Pro Lys	Tyr Ser	Leu Phe	Glu	Leu Glu	Asn Gly	Arg	Lys Arg	Met
1235			1240			1245		
Leu Ala	Ser Ala	Gly Glu	Leu	Gln Lys	Gly Asn	Glu	Leu Ala	Leu
1250			1255			1260		
Pro Ser	Lys Tyr	Val Asn	Phe	Leu Tyr	Leu Ala	Ser	His Tyr	Glu
1265			1270			1275		
Lys Leu	Lys Gly	Ser Pro	Glu	Asp Asn	Glu Gln	Lys	Gln Leu	Phe
1280			1285			1290		
Val Glu	Gln His	Lys His	Tyr	Leu Asp	Glu Ile	Ile	Glu Gln	Ile
1295			1300			1305		
Ser Glu	Phe Ser	Lys Arg	Val	Ile Leu	Ala Asp	Ala	Asn Leu	Asp
1310			1315			1320		
Lys Val	Leu Ser	Ala Tyr	Asn	Lys His	Arg Asp	Lys	Pro Ile	Arg

1325                      1330                      1335  
 Glu Gln Ala Glu Asn Ile Ile His Leu Phe Thr Leu Thr Asn Leu  
 1340                      1345                      1350  
 Gly Ala Pro Ala Ala Phe Lys Tyr Phe Asp Thr Thr Ile Asp Arg  
 1355                      1360                      1365  
 Lys Arg Tyr Thr Ser Thr Lys Glu Val Leu Asp Ala Thr Leu Ile

1370                      1375                      1380  
 His Gln Ser Ile Thr Gly Leu Tyr Glu Thr Arg Ile Asp Leu Ser  
 1385                      1390                      1395  
 Gln Leu Gly Gly Asp Pro Ile Ala Gly Ser Lys Ala Ser Pro Lys  
 1400                      1405                      1410  
 Lys Lys Arg Lys Val Gly Arg Ala Asp Ala Leu Asp Asp Phe Asp  
 1415                      1420                      1425  
 Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp

1430                      1435                      1440  
 Met Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu  
 1445                      1450                      1455  
 Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Ile Asn  
 1460                      1465                      1470  
 Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ser  
 1475                      1480

<210> 2

<211> 350

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 2

gagggcctat ttcccatgat tccttcataat ttgcatatac gatacaaggc tgtagagag	60
ataattggaa ttaatttgac tgtaaacaca aagatattag tacaaaatac gtgacgtaga	120
aagtaataat ttcttgggta gtttgcagtt ttaaaattat gttttaaat ggactatcat	180
atgcttatcg taacttgaaa gtatttcgat ttcttggcctt tatatatctt gtggaaagga	240

cgaaacaccg ggctttcgag aagacctgtt ttagagctag aaatagcaag ttaaaataag 300  
gctagtccgt tatcaacttg aaaaagtggc accgagtcgg tgcttttttt 350

<210> 3

<211> 338

<212> DNA

<213> Homo sapiens

<400> 3

acccccggca gcagcccgcc ccgagcgcgc cgctgttta ttcagccggg agtccggcac 60  
gcgccaggcg cagcactgc aacaacaaac ccagctgaat ggagagtttg caaggagcgg 120  
gagaaaggaa cgggaggggg ggagaggaga ggaggagggg gagtttaggg agtgggtggg 180  
aggaagaggt aagaggaggg gggggagtgg gggctgcagc cgctcgctgc agcagcgggg 240  
agtggggggc gaggcggggc cagggtcgcg cgtggggctg ggtgtcccat tgaaaaggcg 300  
gacgcactcc ggcagcccag cactctctca cttctggc 338

<210> 4

<211> 305

<212> DNA

<213> Mus musculus

<400> 4

ccgccccgag cgcgcggcct gtttattcag ccgggagtcg ggcacgcgcc gggcgcacgc 60  
actgcaacaa caaaccaggc tgaatggaga gtttgcaagg agcgggcgcg ggcaactgga 120  
gggggggggg gcgagaggga gggagctgag gaggtggggg aagaggaggg gtagtggggg 180  
ctgcagccgc tcgctgcagc agcggggagt ggggggcgag gcggggccag ggctgcgcgt 240  
ggggctgggt gtccattga aaaggcggcc gcaccgcagc cgcccagcag tctctcactt 300  
ctggc 305

<210> 5

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 5

gctgggtgtc ccattgaaa 19

<210> 6

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 6

cagccgctcg ctgcagcag 19

<210> 7

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 7

tggagagttt gcaaggagc 19

<210> 8

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 8

gtttattcag ccgggagtc 19

<210> 9

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 9

cgccaggagg ggtgggtcta 20

<210> 10

<211> 19  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
     oligonucleotide  
 <400> 10  
 ccttggtgag actggtaga 19  
 <210> 11  
 <211> 19  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
     oligonucleotide  
 <400> 11  
 gtcttcaggt tctgttgct 19  
  
 <210> 12  
 <211> 19  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
     oligonucleotide  
 <400> 12  
 atattcctga tttaaaagt 19  
 <210> 13  
 <211> 19  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
     oligonucleotide  
 <400> 13  
 ttaaaagtcg gctggtagc 19  
 <210> 14  
 <211> 19  
 <212> DNA



<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 14

cgggccgggg gcgggtcc 19

<210> 15

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 15

gcccagccg cgtgtgaa 19

<210> 16

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 16

ccttcattgc ggcgggctg 19

<210> 17

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 17

ccgacccctc ccgggtccc 19

<210> 18

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 18

caggaccgcg cttcccacg

19

<210> 19

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 19

tgcaccctgg gagcgcgag

19

<210> 20

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 20

ccgcacgcac ctgttccca

19

<210> 21

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 21

aaaacagcga gggagaaac

19

<210> 22

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 22

ttaacttgat tgtgaaatc 19

<210> 23

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 23

aaaacaatgc atatttgca 19

<210> 24

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 24

aaaatccagt attttaatg 19

<210> 25

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 25

accgcact gcgcctgg 19

<210> 26

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 26  
aacttatgcg gcgtttcct 19  
<210> 27  
<211> 19  
<212> DNA  
<213> Artificial Sequence  
<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide  
<400> 27  
tcactttaaa accacctct 19  
<210> 28  
<211> 19  
<212> DNA  
<213> Artificial Sequence  
<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide  
<400> 28  
gcatcttttt ctctttaat 19  
<210> 29  
<211> 19  
<212> DNA  
<213> Artificial Sequence  
<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide  
<400> 29  
tgtactctct gaggtgctc 19  
<210> 30  
<211> 19  
<212> DNA  
<213> Artificial Sequence  
<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide  
<400> 30

acgcagataa gaaccagtt 19

<210> 31

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 31

catcaagtca gccatcagc 19

<210> 32

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 32

gagtcaccct cctggaaac 19

<210> 33

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 33

gctagggatg aagaataaa 19

<210> 34

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 34

ttgaccaata gccttgaca 19

<210> 35

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 35

tgcaaatatc tgtctgaaa

19

<210> 36

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 36

aaattagcag tatcctctt

19

<210> 37

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 37

cctgggctcc ggggcgttt

19

<210> 38

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 38

ggcccctgcg gccaccccg

19

<210> 39

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 39

ctccctccct gcccgtag 19

<210> 40

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 40

aggtttggaa agggcgtgc 19

<210> 41

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 41

ggagcttctc gacttcacca 20

<210> 42

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 42

gatttgtggg cctgaagaaa 20

<210> 43

<211> 20

<212> DNA



<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 43

aaggaggagg gcagaatcat

20

<210> 44

<211> 21

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 44

aaaccttctt cagctatgcc c

21

<210> 45

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 45

agctgatggc cctaaacaga

20

<210> 46

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 46

caggaggact ctggcaccta

20

<210> 47

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
 primer  
 <400> 47  
 ggaatccatg gaggaagat 20  
 <210> 48  
 <211> 21  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
 primer  
 <400> 48  
 gctgagtga ctgcactgtg a 21  
 <210> 49  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 ><223> Description of Artificial Sequence: Synthetic  
 primer  
 <400> 49  
 ctctctgctc ctttgccaca 20  
 <210> 50  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
 primer  
 <400> 50  
 caatgacccc ttcattgacc 20  
 <210> 51  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic

primer	
<400> 51	
ggaacaagag ctgctggact	20
<210> 52	
<211> 20	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic	
primer	
<400> 52	
aactttggca ttgtggaagg	20
<210> 53	
<211> 20	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic	
primer	
<400> 53	
aacgccactg acaagaaagc	20
<210> 54	
<211> 20	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic	
primer	
<400> 54	
cagatccatg gaggaaggaa	20
<210> 55	
<211> 21	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic	
primer	

<400> 55  
 gggtactcct ggaagatgtc c 21  
 <210> 56  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220  
 ><223> Description of Artificial Sequence: Synthetic  
 primer  
 <400> 56  
 gtttgcgacg catgttcctc 20  
 <210> 57  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
 primer  
 <400> 57  
 aagcccttgc ttagtggtg 20  
 <210> 58  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
 primer  
 <400> 58  
 cggcaggaaa gcatctgtat 20  
 <210> 59  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
 primer  
 <400> 59

tgttctcgct caggtcagtg	20
<210> 60	
<211> 20	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic	
primer	
<400> 60	
gaattctttg ccgaaatgga	20
<210> 61	
<211> 20	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic	
primer	
<400> 61	
gtgctcttcg ggtttcagga	20
<210> 62	
<211> 20	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic	
primer	
<400> 62	
ttgattttgg agggatctcg	20
<210> 63	
<211> 20	
<212> DNA	
<213> Artificial Sequence	
<220>	
><223> Description of Artificial Sequence: Synthetic	
primer	
<400> 63	
gtttttctgc ctccccattt	20

<210> 64  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
           primer  
 <400> 64  
 ggatgcaggg atgatgttct 20  
 <210> 65  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
           oligonucleotide  
 <400> 65  
 gattggcttt gatttccta 20  
 <210> 66  
 <211> 23  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
           oligonucleotide  
 <400> 66  
 gtgtagagta agtcagccta tgg 23  
 <210> 67  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
           oligonucleotide  
 <400> 67  
 gcctactcag actgttactc 20  
 <210> 68

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 68

gttgacaga acttaccgac tgg 23

<210> 69

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 69

gcagttgcct aagaactggt 20

<210> 70

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 70

ggggctccac cctcacgagt 20

<210> 71

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 71

gtttgcttcg ctataaaacg agg 23

<210> 72

<211> 23

<212> DNA



<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 72

gtctgaggat ggggccgcaa tgg 23

<210> 73

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 73

ggatctgtca aatgcctgc agg 23

<210> 74

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 74

gccaggatgg cattgggcag cgg 23

<210> 75

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 75

gctgaatctg cgggtggcagg agg 23

<210> 76

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 76

gttcttttgt tcttctagcc tgg

23

<210> 77

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 77

ggaaaagctt gagcaagtca agg

23

<210> 78

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 78

ggaagagttg cccctgcgcc agg

23

<210> 79

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 79

gacaaatctc cagtggataa agg

23

<210> 80

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 80

gtgtttctca ggtaaagctc tgg 23

<210> 81

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 81

ggaaggacca ttgacgtta agg 23

<210> 82

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 82

gaactgctat ttcagtttcc tgg 23

<210> 83

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 83

gccagccact cagccagtga agg 23

<210> 84

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 84  
 ggtatgcttt tctgttaaag agg 23  
 <210> 85  
 <211> 23  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
 oligonucleotide  
 <400> 85  
 gctcctggac tgaccactat tgg 23  
 <210> 86  
 <211> 23  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
 oligonucleotide  
 <400> 86  
 ggaacagagg cgtccccagt tgg 23  
 <210> 87  
 <211> 23  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
 oligonucleotide  
 <400> 87  
 ggaggctaga acaatcatta cgg 23  
 <210> 88  
 <211> 23  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
 oligonucleotide  
 <400> 88

gacaagaaca ccttcagaac cgg	23
<210> 89	
<211> 23	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic oligonucleotide	
<400> 89	
gggtttctgt gattttcttt tgg	23
<210> 90	
<211> 23	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic oligonucleotide	
<400> 90	
gggccaaaga cctccgccag tgg	23
<210> 91	
<211> 23	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic oligonucleotide	
<400> 91	
gttggaag cattcataaa agg	23
<210> 92	
<211> 23	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic oligonucleotide	
<400> 92	
gtcgctcact caccctgcaa agg	23

<210> 93

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 93

gaaaagagct gatgaaacaa tgg 23

<210> 94

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 94

gtacactttt caaatgctt tgg 23

<210> 95

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 95

ggagatgatc atcaagcaga agg 23

<210> 96

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 96

gctttgaaag agcaataaaa tgg 23

<210> 97

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 97

gcacaaaagt caaatcggaa tgg 23

<210> 98

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 98

gatttcaata taagattcgg agg 23

<210> 99

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 99

gttaagcaa tccgaactc tgg 23

<210> 100

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 100

gcctttcttta tcccctatcg 20

<210> 101

<400> 101



000

<210> 102

<400> 102

000

<210> 103

<400> 103

000

<210> 104

<211> 25

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<220><221> modified\_base

<222> (21)..(22)

<223> a, c, t, g, unknown or other

<400> 104

gaggccaaac ctcggttac nngrr 25

<210> 105

<211> 25

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<220><221> modified\_base

<222> (21)..(22)

<223> a, c, t, g, unknown or other

<400> 105

gttcgaaaat ttcaggttaag nngrr 25

<210> 106

<211> 26

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<220><221> modified\_base

<222> (21)..(22)

<223

> a, c, t, g, unknown or other

<400> 106

ggcagaacag gagataacag nngrrt 26

<210> 107

<211> 26

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 107

ggcggccctc gcccttctct ggggat 26

<210> 108

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 108

gtagtgatcg tggatacgag agg 23

<210> 109

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 109

gtacagccct cgggtgtatat tgg 23

<210> 110

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 110

gggaaggaat taagcccgaa tgg 23

<210> 111

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 111

gggaacagct ttcgtagttg agg 23

<210> 112

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 112

gataaagtcc agtgcgatc agg 23

<210> 113

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 113

gaaaaccaga gcttcggtca agg 23

<210> 114

<211> 34

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 114

ggagtcttct gggcaggctt aaaggctaac ctgg

34

<210> 115

<211> 34

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 115

gtcgggtgag catgtcttta atctacctcg atgg

34

<210> 116

<211> 26

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 116

ggtgtcacca gagtaacagt ctgagt

26

<210> 117

<211> 25

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 117

gtgatcatca agcagaaggt atgag

25

<210> 118

<211> 29

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 118

gaacttcgaa aatttcaggt aagccgagg 29

<210> 119

<211> 29

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 119

ggaaactcat caaatatgcg tgtagtgt 29

<210> 120

<211> 29

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 120

gtcatttaca ctaacacgca tatttgatg 29

<210> 121

<211> 29

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 121

ggaatgaaac tcatcaaata tgcgtgtta 29

<210> 122

<211> 29

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 122  
 gtcacatcaata tctttgaagg actctgggt 29

<210> 123  
 <211> 29  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> 123  
 gtgttttcat aggaaaaata ggcaagttg 29

<210> 124  
 <211> 29  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> 124  
 gaattggaaa atgtgatggg aaacagata 29

<210> 125  
 <211> 29  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> 125  
 gatgatcatc aagcagaagg tatgagaaa 29

<210> 126  
 <211> 29  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> 126  
 gagatgatca tcaagcagaa ggtatgaga 29

<210> 127

<211> 29

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 127

gcattttttc tcataccttc tgcttgatg 29

<210> 128

<211> 29

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 128

gtcctactca gactgttact ctggtgaca 29

<210> 129

<211> 29

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 129

gacaggttgt gtcaccagag taacagtct 29

<210> 130

<211> 29

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 130

gttatcattt ttctcatcac cttctgctt 29

<210> 131

<211> 29

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 131

gttgcctaag aactggtggg aaatggtct 29

<210> 132

<211> 29

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 132

gaaacagttg cctaagaact ggtgggaaa 29

<210> 133

<211> 29

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 133

gtttccacc agttcttagg caactgttt 29

<210> 134

<211> 29

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 134

gtggctttga tttccctagg gtccagctt 29

<210> 135

<211> 29



<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 135

gtagggaaat caaagccaat gaaacgttc 29

<210> 136

<211> 29

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 136

ggaccctagg gaaatcaaag ccaatgaaa 29

<210> 137

<211> 29

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 137

gtgagggtc caccctcacg agtgggttt 29

<210> 138

<211> 29

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 138

gaaggattga gggctccacc ctcacgagt 29

<210> 139

<211> 29

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 139

ggctccaccc tcacgagtgg gtttggttc 29

<210> 140

<211> 29

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 140

gtatccccta tcgaggaaac cacgagttt 29

<210> 141

<211> 29

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 141

ggataaagaa ggcctatttc atagagttg 29

<210> 142

<211> 29

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 142

gaggccttct ttatccccta tcgaggaaa 29

<210> 143

<211> 27

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 143

gtgagggtc caccctcacg agtgggt 27

<210> 144

<211> 27

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 144

ggataaagaa ggcctatttc atagagt 27

<210> 145

<211> 4932

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 145

gccaccatgg acaagaagta ctccattggg ctcgatatcg gcacaaacag cgtcggctgg 60

gccgtcatta cggacgagta caagtgccg agcaaaaaat tcaaagttct gggcaatacc 120

gatcgccaca gcataaagaa gaacctcatt ggcgcctcc tgttcgactc cggggagacg 180

gccgaagcca cgcggctcaa aagaacagca cggcgcagat ataccgcag aaagaatcgg 240

atctgtacc tgcaggagat ctttagtaat gagatggcta aggtggatga ctctttcttc 300

cataggtcgg aggagtcctt tttggtggag gaggataaaa agcacgagcg ccaccaatc 360

tttggcaata tcgtggacga ggtggcgtac catgaaaagt acccaacat atatcatctg 420

aggaagaagc ttgtagacag tactgataag gctgacttgc ggttgatcta tctcgcgtg 480

gcgcatatga tcaaatttcg gggacacttc ctcacgagg gggacctgaa ccagacaac 540

agcgatgtcg acaaactctt tatccaactg gttcagactt acaatcagct tttcgaagag 600

aaccgatca acgcatccgg agttgacgcc aaagcaatcc tgagcgctag gctgtccaaa 660

tcccggcggc tcgaaaacct catcgcacag ctccctgggg agaagaagaa cggcctgttt 720

ggtaatctta tcgccctgtc actcgggctg accccaact ttaaatctaa cttcgacctg 780

gccgaagatg ccaagcttca actgagcaaa gacacctacg atgatgatct cgacaatctg 840  
ctggcccaga tggcgacca gtacgcagac ctttttttgg cggcaaagaa cctgtcagac 900

gccattctgc tgagtgatat tctgcgagtg aacacggaga tcaccaaagc tccgctgagc 960  
gctagtatga tcaagcgcta tgatgagcac caccaagact tgactttgct gaaggccctt 1020  
gtcagacagc aactgcctga gaagtacaag gaaattttct tcgatcagtc taaaaatggc 1080  
tacgccggat acattgacgg cggagcaagc caggaggaat ttacaaaatt tattaagccc 1140  
atcttgaaaa aaatggacgg caccgaggag ctgctggtaa agcttaacag agaagatctg 1200  
ttgcgcaaac agcgacttt cgacaatgga agcatcccc accagattca cctgggcgaa 1260  
ctgcacgcta tcctcaggcg gcaagaggat ttctaccct tttgaaaga taacagggaa 1320

aagattgaga aaatctcac atttcggata ccctactatg taggccccct cgcccgggga 1380  
aattccagat tcgctggat gactcgcaaa tcagaagaga ccatcactcc ctggaacttc 1440  
gaggaagtgc tggataagg ggcctctgcc cagtccttca tcgaaaggat gactaacttt 1500  
gataaaaatc tgcctaacga aaagtgctt cctaaacact ctctgctgta cgagtacttc 1560  
acagtttata acgagctcac caaggtcaaa tacgtcacag aagggatgag aaagccagca 1620  
ttcctgtctg gagagcagaa gaaagctatc gtggacctcc tcttcaagac gaaccggaaa 1680  
gttaccgtga aacagctcaa agaagactat ttcaaaaaga ttgaatgttt cgactctgtt 1740

gaaatcagcg gagtggagga tcgcttcaac gcatccctgg gaacgtatca cgatctcctg 1800  
aaaatcatta aagacaagga cttcctggac aatgaggaga acgaggacat tcttgaggac 1860  
attgtcctca cccttacgtt gtttgaagat agggagatga ttgaagaacg cttgaaaact 1920  
tacgtcatc tcttcgacga caaagtcatg aaacagctca agaggcgccg atatacagga 1980  
tggggcgccg tgtcaagaaa actgatcaat gggatccgag acaagcagag tggaaagaca 2040  
atcctggatt ttcttaagtc cgatggattt gccaaccgga acttcatgca gttgatccat 2100  
gatgactctc tcacctttaa ggaggacatc cagaaagcac aagtttctgg ccagggggac 2160

agtcttcacg agcacatcg taatcttgca ggtagcccag ctatcaaaaa gggaatactg 2220  
cagaccgtta aggtcgtgga tgaactcgtc aaagtaatgg gaaggcataa gcccgagaat 2280  
atcgttatcg agatggcccg agagaaccaa actaccaga agggacagaa gaacagtagg 2340  
gaaaggatga agaggattga agagggtata aaagaactgg ggtcccaaat ccttaaggaa 2400  
caccagttg aaaacacca gcttcagaat gagaagctct acctgtacta cctgcagaac 2460  
ggcagggaca tgtacgtgga tcaggaaactg gacatcaatc ggctctccga ctacgacgtg 2520

gatcatatcg tgccccagtc ttttctcaaa gatgattcta ttgataataa agtgttgaca	2580
agatccgata aaaatagagg gaagagtgat aacgtcccct cagaagaagt tgtcaagaaa	2640
atgaaaaatt attggcggca gctgctgaac gccaaactga tcacacaacg gaagttcgat	2700
aatctgacta aggctgaacg aggtggccctg tctgagttgg ataaagccgg cttcatcaaa	2760
aggcagcttg ttgagacacg ccagatcacc aagcacgtgg cccaaattct cgattcacgc	2820
atgaacacca agtacgatga aaatgacaaa ctgattcgag aggtgaaagt tattactctg	2880
aagtctaagc tggctcaga tttcagaaag gactttcagt tttataaggt gagagagatc	2940
aacaattacc accatgcgca tgatgcctac ctgaatgcag tggtaggcac tgcacttacc	3000
aaaaaatatc ccaagcttga atctgaattt gtttacggag actataaagt gtacgatgtt	3060
aggaaaatga tcgcaaagtc tgagcaggaa ataggcaagg ccaccgctaa gtacttcttt	3120
tacagcaata ttatgaattt tttcaagacc gagattacac tggccaatgg agagattcgg	3180
aagcgaccac ttatcgaac aaacggagaa acaggagaaa tcgtgtggga caagggtagg	3240
gatttcgcga cagtcggaa ggtcctgtcc atgccgcagg tgaacatcgt taaaaagacc	3300
gaagtacaga ccggaggtct ctccaaggaa agtatcctcc cgaaaaggaa cagcgacaag	3360
ctgatcgac gcaaaaaaga ttgggacccc aagaaatcgc gcgattcga ttctcctaca	3420
gtcgcttaca gtgtactggt tgtggccaaa gtggagaaaag ggaagtctaa aaaactcaaa	3480
agcgtcaagg aactgctggg catcacaatc atggagcgat caagcttcga aaaaaacccc	3540
atcgactttc tcgaggcgaa aggatataaa gaggtcaaaa aagacctcat cattaagctt	3600
cccaagtact ctctctttga gcttgaaaac ggccggaaac gaatgctcgc tagtgcgggc	3660
gagctgcaga aaggtaacga gctggcacctg cctctaaat acgttaattt cttgtatctg	3720
gccagccact atgaaaagct caaagggctct cccgaagata atgagcagaa gcagctgttc	3780
gtggaacaac acaaacacta ccttgatgag atcatcgagc aaataagcga attctccaaa	3840
agagtgatcc tcgccgacgc taacctcgat aaggtgcttt ctgcttaca taagcacagg	3900
gataagccca tcaggagca ggcagaaaaac attatccact tgtttactct gaccaacttg	3960
ggcgcgcctg cagccttcaa gtacttcgac accaccatag acagaaagcg gtacacctct	4020
acaaaggagg tcctggacgc cacttgatt catcagtcaa ttacgggct ctatgaaca	4080
agaatcgacc tctctcagct cggtaggagc agcagggtg accccaagaa gaagaggaag	4140
gtggctagcg agggcagagg aagtcttcta acatgcggtg acgtggagga gaatcccggc	4200
cctggtaccg tgagcaaggg cgaggagctg ttcaccgggg tggtagccat cctggtcgag	4260

ctggacggcg acgtaaacgg ccacaagttc agcgtgtccg gcgagggcga gggcgatgcc 4320  
acctacggca agctgaccct gaagttcac tgcaccaccg gcaagctgcc cgtgccctgg 4380  
ccccccctcg tgaccaccct gacctacggc gtgcagtgtc tcagccgcta ccccgaccac 4440  
atgaagcagc acgacttctt caagtccgcc atcccgaag gctacgtcca ggagcgcacc 4500  
atcttcttca aggacgacgg caactacaag acccgcgccg aggtgaagtt cgagggcgac 4560  
accttggatga accgcatcga gctgaagggc atcgacttca aggaggacgg caacatcctg 4620  
gggcacaagc tggagtacaa ctacaacagc cacaacgtct atatcatggc cgacaagcag 4680

aagaacggca tcaagtgaa cttcaagatc cgccacaaca tcgaggacgg cagcgtgcag 4740  
ctcgccgacc actaccagca gaacaccccc atcggcgacg gccccgtgtc gctgcccgac 4800  
aaccactacc tgagcaccca gtccgccctg agcaaagacc ccaacgagaa gcgcgatcac 4860  
atggtcctgc tggagttcgt gaccgccgcc gggatcactc tcggcatgga cgagctgtac 4920  
aagaccggtt ag 4932

<210> 146

<400> 146

000

<210> 147

<400> 147

000

<210> 148

<400> 148

000

<210> 149

<400> 149

000

<210> 150

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 150

tcctactcag actgttactc tgg

23

<210> 151

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 151

tcctactcac actgttactc agg 23

<210> 152

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 152

acctgctcac actgttactc cag 23

<210> 153

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 153

gcattctcaa actgttactc agg 23

<210> 154

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 154

ggattctcac actgttactc ggg 23

<210> 155

<211> 23

<212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
     oligonucleotide  
 <400> 155  
 acatacttat actgttactc tag 23  
 <210> 156  
 <211> 23  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
     oligonucleotide  
 <400> 156  
 tattcctaag actgttactc aag 23  
 <210> 157  
 <211> 23  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
     oligonucleotide  
 <400> 157  
 aaggactaag actgttactc ggg 23  
 <210> 158  
 <211> 23  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
     oligonucleotide  
 <400> 158  
 gagctctcat actgttactc tag 23  
 <210> 159  
 <211> 23  
 <212> DNA



<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 159

gcaaaatgag actgttactc cag 23

<210> 160

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 160

cctcattcag actgttactc aag 23

<210> 161

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 161

cattggcttt gatttccta ggg 23

<210> 162

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 162

aattggcatt gatttccta gag 23

<210> 163

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 163

cattggcttt aatttccta tag 23

<210> 164

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 164

gataggctgt gatttccta gag 23

<210> 165

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 165

gaatagcctt gatttccta aag 23

<210> 166

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 166

aatttgcttt gatttcctg agg 23

<210> 167

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 167

gatgtgcttt gatttcctt ggg 23

<210> 168

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 168

aattggtttt aatttccta aag 23

<210> 169

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 169

aattgggttt gatttcctt tgg 23

<210> 170

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 170

gatgggtttt tatttccta gag 23

<210> 171

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 171  
gaatggtttt gatttcctg gag 23  
<210> 172  
<211> 23  
<212> DNA  
<213> Artificial Sequence  
<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide  
<400> 172  
acagttgcct aagaactggt ggg 23  
<210> 173  
<211> 23  
<212> DNA  
<213> Artificial Sequence  
<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide  
<400> 173  
ccagttgtct aagaactggg gag 23  
<210> 174  
<211> 23  
<212> DNA  
<213> Artificial Sequence  
<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide  
<400> 174  
gcagttgcct gtgaactggt agg 23  
<210> 175  
<211> 23  
<212> DNA  
<213> Artificial Sequence  
<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide  
<400> 175

gcagatgcag aagaactggt gag 23

<210> 176

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 176

gcagttccag aagaactggt gag 23

<210> 177

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 177

caacttcct atgaactggt agg 23

<210> 178

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 178

acacctgcct aagaactgga ggg 23

<210> 179

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 179

tcaggtggct aagaactggg tgg 23

<210> 180

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 180

gaagttggcc aagaactgga gag 23

<210> 181

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 181

gctgctgccc aagaactggc agg 23

<210> 182

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 182

tcagctggct aagaacgggt aag 23

<210> 183

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 183

agggtccac cctcacgagt ggg 23

<210> 184

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 184

gcagctcagc cctcacgagt cag 23

<210> 185

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 185

ggggcttcag catcacgagt gag 23

<210> 186

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 186

ggggctctcc cctcactagt gag 23

<210> 187

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 187

ggggatccac cttcaccagt cag 23

<210> 188

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 188

agggtggac cctcacaagt aag 23

<210> 189

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 189

tggctctctc ccccacgagt ggg 23

<210> 190

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 190

agggtccca ccccacgagt gag 23

<210> 191

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 191

gaggctccat actcaccagt gag 23

<210> 192

<211> 23

<212> DNA

<213> Artificial Sequence



<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 192

ggagctgccc cttcacgagt ggg

23

<210> 193

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 193

atgactccac cctcaagagt aag

23

<210> 194

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 194

gcctttcttta tcccctatcg agg

23

<210> 195

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 195

gtctgctgtg tcccctatcg ggg

23

<210> 196

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 196

cccttctcta tcccctgtcg tgg 23

<210> 197

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 197

gccttcttta tcccctctct tgg 23

<210> 198

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 198

gcgctctttt tcccctatct tag 23

<210> 199

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 199

gccctctgtc tcccctgtcg cag 23

<210> 200

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 200

tccatctttg tcccctattg agg	23
<210> 201	
<211> 23	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic oligonucleotide	
<400> 201	
accttctctc tcccctatag agg	23
<210> 202	
<211> 23	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic oligonucleotide	
<400> 202	
gttttctttt tcccctatgg gag	23
<210> 203	
<211> 23	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic oligonucleotide	
<400> 203	
tgcttcttaa tcccctatca aag	23
<210> 204	
<211> 23	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic oligonucleotide	
<400> 204	
accttcttac tcccctatcc ggg	23

<210> 205  
 <211> 22  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
     primer  
 <400> 205  
 gagaggttat gtggctttac ca 22

<210> 206  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
     primer  
 <400> 206  
 aaaaatgctt cccactttgc 20

<210> 207  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
     primer  
 <400> 207  
 ctcatctca tgcttgaca 20

<210> 208  
 <211> 25  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
     primer  
 <400> 208  
 gagtttgct caaattgtta ctctt 25

<210> 209

<211> 25  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
     primer  
 <400> 209  
 gggaaatggt ctaggagagt aaagt 25  
 <210> 210  
 <211> 25  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
     primer  
 <400> 210  
 gtttggtca aattgttact ctca 25  
 <210> 211  
 <211> 25  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
     primer  
 <400> 211  
 gtgagagtaa tgtgtttgct gagag 25  
 <210> 212  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
     primer  
 <400> 212  
 cgggcttgga cagaacttac 20  
 <210> 213  
 <211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 213

ctgcgtagtg ccaaaacaaa 20

<210> 214

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 214

taatttcatt gaagagtggc tgaa 24

<210> 215

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 215

aagccctgtg tggtagtagt cagt 24

<210> 216

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 216

tgagtcatgt tggataacca gtct 24

<210> 217

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 217

gaaggtcagg aacatacaat tcaa

24

<210> 218

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 218

gatatgggca tgcagtttc atag

24

<210> 219

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 219

tgctgttgat taatggttga tagg

24

<210> 220

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 220

ttttaaattg ccatgtttgt gtc

23

<210> 221

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 221

atgaataacc taatgggcag aaaa 24

<210> 222

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 222

tcaagtcgct tcattttgat agac 24

<210> 223

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 223

cacaacaaaa catatagcca aagc 24

<210> 224

<211> 25

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 224

tgctgctaaa ataacacaaa tcagt 25

<210> 225

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer



<400> 225  
ctgtgcctat tgtggttatc ctg 23

<210> 226  
<211> 24  
<212> DNA  
<213> Artificial Sequence  
<220><223> Description of Artificial Sequence: Synthetic  
primer  
<400> 226  
attgatctgc aatacatgtg gagt 24

<210> 227  
<211> 24  
<212> DNA  
<213> Artificial Sequence  
<220><223> Description of Artificial Sequence: Synthetic  
primer  
<400> 227  
tttgctctg ctattacagt atgg 24

<210> 228  
<211> 24  
<212> DNA  
<213> Artificial Sequence  
<220><223> Description of Artificial Sequence: Synthetic  
primer  
<400> 228  
tgtaggtgg ttggctaaaa taat 24

<210> 229  
<211> 24  
<212> DNA  
<213> Artificial Sequence  
<220><223> Description of Artificial Sequence: Synthetic  
primer  
<400> 229

tttttgcaca gtcaataaca caaa	24
<210> 230	
<211> 24	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic	
primer	
<400> 230	
ggctggtctc acaattgtac tttt	24
<210> 231	
<211> 24	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic	
primer	
<400> 231	
cattatggac tgaaaatctc agca	24
<210> 232	
<211> 24	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic	
primer	
<400> 232	
atcatcctag ccataacaca atga	24
<210> 233	
<211> 24	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic	
primer	
<400> 233	
ttcagcttta acgtgatttt ctgt	24

<210> 234  
 <211> 24  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
     primer  
 <400> 234  
 ggattcagaa gctgtttacg aagt 24  
 <210> 235  
 <211> 24  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
     primer  
 <400> 235  
 tttagctgga ttggaaaaac aaat 24  
 <210> 236  
 <211> 24  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
     primer  
 <400> 236  
 aactcacccc attgttggtat tatt 24  
 <210> 237  
 <211> 24  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
     primer  
 <400> 237  
 ccttgtccaa ataccgaaat acat 24  
 <210> 238  
 <211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 238

cacataattc atgaacttgg cttc 24

<210> 239

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 239

tagtagctgg ggaggaagat acag 24

<210> 240

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 240

tttttgtttt aattgcgact gtgt 24

<210> 241

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 241

agaaaagggg ttttcttttg actt 24

<210> 242

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 242

cattgtgact ggatgagaag aaac 24

<210> 243

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 243

aacggctgtt attaaagtcc tcag 24

<210> 244

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 244

caagtcagaa gtcacttgct ttgt 24

<210> 245

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 245

ttttatgtgc aggaatcagt ctgt 24

<210> 246

<211> 18

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
 primer  
 <400> 246  
 tggcggcgtt ttcattat 18  
 <210> 247  
 <211> 26  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
 primer  
 <400> 247  
 ttcgatccgt aatgattgtt ctagcc 26  
 <210> 248  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
 primer  
 <400> 248  
 ggtcttccag agtgctgagg 20  
 <210> 249  
 <211> 24  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
 primer  
 <400> 249  
 tgtgtgcttc tgtacacatc atct 24  
 <210> 250  
 <211> 24  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic

```

    primer
<400> 250
agatttcaac cctcaaaaac tgag
24
<210> 251
<211> 24
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
    primer
<400> 251
taaactcttt cttttccgca attc
24
<210> 252
<211> 24
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
    primer
<400> 252
caaggtgacc tgctacctaa aaat
24
<210> 253
<211> 24
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
    primer
<400> 253
tatgaccaag gctatgtgtt cact
24
<210> 254
<211> 24
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
    primer
<400> 254

```

acagcctctc tccagtaaca ttct	24
<210> 255	
<211> 24	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic primer	
<400> 255	
tattcttgca gtggtttcac attt	24
<210> 256	
<211> 24	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic primer	
<400> 256	
atattttaag ccaagaccca acaa	24
<210> 257	
<211> 24	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic primer	
<400> 257	
ctttcaactg tctgtctgat tgct	24
<210> 258	
<211> 24	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic primer	
<400> 258	
aacagcctct cttcattggt ctct	24



<210> 259

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 259

ctctggaact tgtctctgtc ttga 24

<210> 260

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 260

ctttcctgcg ttctcatgtt acta 24

<210> 261

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 261

ccttatatcc gtatcgctca ctct 24

<210> 262

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 262

catatctgtc taacttcgc acac 24

<210> 263

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 263

acaggtgtta tgtgtctgc atct 24

<210> 264

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 264

actccattcc cagattagtt atgc 24

<210> 265

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 265

ctgttttctt tgtgagagtg gaga 24

<210> 266

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 266

tgtaaggtagg tcaaacttgc tcta 24

<210> 267

<211> 25

<212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
     primer  
 <400> 267  
 tttttcctag tacccacaga ttttt 25  
 <210> 268  
 <211> 24  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
     primer  
 <400> 268  
 tccctgattc tctcatttgt gtta 24  
 <210> 269  
 <211> 24  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
     primer  
 <400> 269  
 ttgggaacat cagagaaagt atga 24  
 <210> 270  
 <211> 24  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
     primer  
 <400> 270  
 acaaattaca gtctcctggg aaag 24  
 <210> 271  
 <211> 24  
 <212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 271

agtagcttac cttggcagag aaaa 24

<210> 272

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 272

tgacatactg ttaccctttg cagt 24

<210> 273

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 273

gaaaggctca gtgaatgttt gtt 23

<210> 274

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 274

cactgcatca ttcattaaa tcaa 24

<210> 275

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 275

cccatatatt catgattacc caca

24

<210> 276

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 276

tatcagaacg agcactaaaa gcac

24

<210> 277

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 277

ttgggaggct gaggtacaag

20

<210> 278

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 278

gaatgaaaaa caaacagaag gtga

24

<210> 279

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 279

ctcctcatct gtacccttca atct 24

<210> 280

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 280

agagtggcat ctagtgtcag tgag 24

<210> 281

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 281

taccaaaagc ttctcctgtt tacc 24

<210> 282

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 282

gtaagttgga tggcctattc ttg 24

<210> 283

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 283  
gaaggaaatg caaggataca agat 24  
<210> 284  
<211> 24  
<212> DNA  
<213> Artificial Sequence  
<220><223> Description of Artificial Sequence: Synthetic  
primer  
<400> 284  
tgattgaaag aatcattcca gaaa 24  
<210> 285  
<211> 24  
<212> DNA  
<213> Artificial Sequence  
  
<220><223> Description of Artificial Sequence: Synthetic  
primer  
<400> 285  
tcagaaggaa aattgaaatt gggt 24  
<210> 286  
<211> 24  
<212> DNA  
<213> Artificial Sequence  
<220><223> Description of Artificial Sequence: Synthetic  
primer  
<400> 286  
cagatgtgtt cttcatcatt cctc 24  
<210> 287  
<211> 24  
<212> DNA  
<213> Artificial Sequence  
<220><223> Description of Artificial Sequence: Synthetic  
primer  
<400> 287

ttctcttttag ggaaagctct caaa 24

<210> 288

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 288

gggtatagat catatggagg gaag 24

<210> 289

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 289

agatgatctg cccacctcag 20

<210> 290

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 290

ctttcttctt catttagtgg caat 24

<210> 291

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 291

atgaattgca gattgatggt actg 24

<210> 292



<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 292

tctcaccaag aaccaaattg tcta

24

<210> 293

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 293

gtaggatacc ttggcaacag tctt

24

<210> 294

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 294

ttaacgaatt gtgagatttg ctgt

24

<210> 295

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 295

tcagaaagtc aagtagcaca caca

24

<210> 296

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 296

agaagcacac actcaggtaa agc 23

<210> 297

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 297

tctttggggg aataatgact aaaa 24

<210> 298

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 298

tttggcattt atgggaataa aact 24

<210> 299

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 299

actaattctg gtcaagccca tca 23

<210> 300

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 300

ttaagacatc ggatgaacag aaag 24

<210> 301

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 301

agaagctttc tgacatgac tgc 23

<210> 302

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 302

tcaattgcat taggacttag acca 24

<210> 303

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 303

gttaaattac ctgtgaagcc cttg 24

<210> 304

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 304

cggaaaacag atccacttta tgat 24

<210> 305

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 305

aaatccactg gaaacatctt gagt 24

<210> 306

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 306

agtctcttca gaatcatgcc ctat 24

<210> 307

<211> 22

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 307

gcttggtggc acatacctgt ag 22

<210> 308

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 308

ggtaggtaga ttgcttgct tggt 24

<210> 309

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 309

agctctcagc agagtaggga tttta 24

<210> 310

<211> 21

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 310

gtgagtctac tgcaccccat c 21

<210> 311

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 311

tgacactgtg aagtcaattc tgtc 24

<210> 312

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 312  
tcaagaactt gacaatgagc aaat 24

<210> 313  
<211> 22  
<212> DNA  
<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 313  
tatccgatcc actgttgtgt gt 22

<210> 314  
<211> 23  
<212> DNA  
<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 314  
caggagaccc aaaaccactc tac 23

<210> 315  
<211> 24  
<212> DNA  
<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 315  
ttgttctaca aatagggttt cctt 24

<210> 316  
<211> 24  
<212> DNA  
<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 316

tgtaaagttt gggcttatgt tcct	24
<210> 317	
<211> 23	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic primer	
<400> 317	
cacaagtctc actgcacaaa cat	23
<210> 318	
<211> 24	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic primer	
<400> 318	
tgacctatga ttatctctct ttga	24
<210> 319	
<211> 24	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic primer	
<400> 319	
ttcagcttct gattggtttt aatg	24
<210> 320	
<211> 24	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic primer	
<400> 320	
ccaattcctt aattttccct acag	24

<210> 321  
 <211> 21  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
     primer  
 <400> 321  
 atctcagacc aggagggaga c 21  
 <210> 322  
 <211> 24  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
     primer  
 <400> 322  
 cctcagggtc agtacatttt tcag 24  
 <210> 323  
 <211> 24  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
     primer  
 <400> 323  
 ttcttaggac attgctccac atac 24  
 <210> 324  
 <211> 24  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
     primer  
 <400> 324  
 gcaaacataa tgcaactcgt aatc 24  
 <210> 325



<211> 21

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 325

gcaagggagt ctgtgtcttt g 21

<210> 326

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 326

tcatttaagt ggctgttctg tggt 24

<210> 327

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 327

acaaaacaga gagaaaaggc agag 24

<210> 328

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 328

gttttgattt ctggtgccta cag 23

<210> 329

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 329

actgaagctg aagcccagtc 20

<210> 330

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 330

acatgagctc tcaggtttct gac 23

<210> 331

<211> 25

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 331

tcaaaccttag atggttcctt atgtt 25

<210> 332

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 332

gtaccctgaa aatgtagggt gact 24

<210> 333

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 333

cacttcccaa gtgaggcaat

20

<210> 334

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 334

ctatacttgg ggctgacttg ctac

24

<210> 335

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 335

tcgtataggt tactttggct caca

24

<210> 336

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 336

agggatcttt actcctcagt gtgt

24

<210> 337

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer  
<400> 337  
tgtagaagtt ggaatatacct gctg 24  
<210> 338  
<211> 24  
<212> DNA  
<213> Artificial Sequence  
<220><223> Description of Artificial Sequence: Synthetic  
primer  
<400> 338  
gtcaacaatt tgatctcagg cttc 24  
<210> 339  
<211> 24  
<212> DNA  
<213> Artificial Sequence  
<220><223> Description of Artificial Sequence: Synthetic  
primer  
<400> 339  
ctcagtacta aagatggacg cttg 24  
<210> 340  
<211> 24  
<212> DNA  
<213> Artificial Sequence  
<220><223> Description of Artificial Sequence: Synthetic  
primer  
<400> 340  
aatcatttca gtcttcctaa caat 24  
<210> 341  
<211> 24  
<212> DNA  
<213> Artificial Sequence  
<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 341  
gggaatcaca gtagatgttt gtca 24  
<210> 342  
<211> 24  
<212> DNA  
<213> Artificial Sequence  
<220><223> Description of Artificial Sequence: Synthetic  
primer  
<400> 342  
agaccaggag gtaagaacat ttg 24  
<210> 343  
<211> 24  
<212> DNA  
<213> Artificial Sequence  
<220><223> Description of Artificial Sequence: Synthetic  
primer  
<400> 343  
ccacatagaa agagacttgc agaa 24  
<210> 344  
<211> 24  
<212> DNA  
<213> Artificial Sequence  
<220><223> Description of Artificial Sequence: Synthetic  
primer  
<400> 344  
agagatgcca aaagaacagt caat 24  
<210> 345  
<211> 24  
<212> DNA  
<213> Artificial Sequence  
<220><223> Description of Artificial Sequence: Synthetic  
primer  
<400> 345

tgtagccttag gctatgtaaa ctgt	24
<210> 346	
<211> 24	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic primer	
<400> 346	
aaacccttgt aaccaaaatt acca	24
<210> 347	
<211> 24	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic primer	
<400> 347	
taactgcatc agaagtcctt gcta	24
<210> 348	
<211> 22	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic primer	
<400> 348	
ggagaccaag ctgctaaagt ca	22
<210> 349	
<211> 37	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic primer	
<400> 349	
gtggtgccgc gggagtttgg ctcaaattgt tactctt	37

<210> 350

<211> 37

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 350

gtggtgccgc ggggaaatg gtctaggaga gtaaagt 37

<210> 351

<211> 34

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 351

gtggtgccgc gggagaggtt atgtggcttt acca 34

<210> 352

<211> 32

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 352

gtggtgccgc ggctcattct catgcctgga ca 32

<210> 353

<211> 32

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 353

gtggtgccgc ggcgggcttg gacagaactt ac 32

<210> 354

<211> 32

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 354

gtggtgccgc ggctgcgtag tgccaaaaca aa 32

<210> 355

<211> 36

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 355

gtggtgccgc ggtaatttca ttgaagagt gctgaa 36

<210> 356

<211> 36

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 356

gtggtgccgc ggaagccctg tgtgtagta gtcagt 36

<210> 357

<211> 36

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 357

gtggtgccgc ggcaagtcag aagtcacttg ctttgt 36

<210> 358

<211> 36



<212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
     primer  
 <400> 358  
 gtggtgccgc ggttttatgt gcaggaatca gtctgt 36  
 <210> 359  
 <211> 36  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
     primer  
 <400> 359  
 gtggtgccgc ggtgtgtgct tctgtacaca tcattc 36  
 <210> 360  
 <211> 36  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
     primer  
 <400> 360  
 gtggtgccgc ggagatttca accctcaaaa actgag 36  
 <210> 361  
 <211> 36  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
     primer  
 <400> 361  
 gtggtgccgc ggttggaac atcagagaaa gtatga 36  
 <210> 362  
 <211> 36  
 <212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 362

gtggtgccgc ggacaaatta cagtctcctg ggaaag 36

<210> 363

<211> 32

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 363

gtggtgccgc ggcacttccc aagtgaggca at 32

<210> 364

<211> 36

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 364

gtggtgccgc ggctatactt ggggctgact tgctac 36

<210> 365

<211> 34

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 365

gtggtgccgc ggttggtctt ttagcttggt tttc 34

<210> 366

<211> 32

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 366

gtggtgccgc ggtgagactc ccaaaggcaa tc

32

<210> 367

<211> 34

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 367

gtggtgccgc ggttggctct ttagcttggtg tttc

34

<210> 368

<211> 32

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 368

gtggtgccgc ggactgaggg gtgatcttgg tg

32

<210> 369

<211> 32

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 369

gtggtgccgc gggcagagaa agccagtcgg ta

32

<210> 370

<211> 32

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 370

gtggtgccgc ggtgagactc ccaaaggcaa tc 32

<210> 371

<211> 32

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 371

gtggtgccgc gggcagagaa agccagtcgg ta 32

<210> 372

<211> 32

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 372

gtggtgccgc ggactgaggg gtgatcttgg tg 32

<210> 373

<211> 32

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 373

gtggtgccgc ggccagagtt cctagggcag ag 32

<210> 374

<211> 32

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 374  
gtggtgccgc ggagctagtc cccacattcc ac 32  
<210> 375  
<211> 32  
<212> DNA  
<213> Artificial Sequence  
<220><223> Description of Artificial Sequence: Synthetic  
primer  
<400> 375  
gtggtgccgc ggccagagtt cctagggcag ag 32  
<210> 376  
<211> 32  
<212> DNA  
<213> Artificial Sequence  
<220><223> Description of Artificial Sequence: Synthetic  
primer  
<400> 376  
gtggtgccgc ggggtggagg gaaactttag gc 32  
<210> 377  
<211> 32  
<212> DNA  
<213> Artificial Sequence  
<220><223> Description of Artificial Sequence: Synthetic  
primer  
<400> 377  
gtggtgccgc ggctcattct catgcctgga ca 32  
<210> 378  
<211> 32  
<212> DNA  
<213> Artificial Sequence  
<220><223> Description of Artificial Sequence: Synthetic  
primer  
<400> 378

gtggtgccgc ggagctagtc cccacattcc ac 32

<210> 379

<211> 34

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 379

gtggtgccgc ggtctcatgc ctggacaagt aact 34

<210> 380

<211> 32

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 380

gtggtgccgc ggggtggagg gaaactttag gc 32

<210> 381

<211> 32

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 381

gtggtgccgc ggggcttgga cagaacttac cg 32

<210> 382

<211> 32

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 382

gtggtgccgc ggcaccactg tctgcctaag ga 32

<210> 383

<211> 32

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 383

gtggtgccgc ggggcttgga cagaacttac cg 32

<210> 384

<211> 32

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 384

gtggtgccgc ggggtggagg gaaactttag gc 32

<210> 385

<211> 33

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 385

gtggtgccgc ggcgtagtgc caaaacaaac agt 33

<210> 386

<211> 32

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 386

gtggtgccgc ggcaccactg tctgcctaag ga 32

<210> 387

<211> 33

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 387

gtggtgccgc ggcgtagtgc caaaacaaac agt 33

<210> 388

<211> 32

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 388

gtggtgccgc ggggtggagg gaaacttttag gc 32

<210> 389

<211> 32

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 389

gtggtgccgc gggcgagggc ctacttgata tg 32

<210> 390

<211> 32

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 390

gtggtgccgc ggcttcccaa gtgaggcaat gc 32

<210> 391

<211> 33

<212> DNA



<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 391

gtggtgccgc ggacgttttg tgctgctgta aca 33

<210> 392

<211> 32

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 392

gtggtgccgc ggctgcaggc acattctctt cc 32

<210> 393

<211> 32

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 393

gtggtgccgc gggccctgtg tggtagtagt ca 32

<210> 394

<211> 32

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 394

gtggtgccgc ggcttcctaa gtgaggcaat gc 32

<210> 395

<211> 33

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 395

gtggtgccgc ggcagtatta aggggtggga gct

33

<210> 396

<211> 32

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 396

gtggtgccgc ggtctcttcc tcacacagct ga

32

<210> 397

<211> 32

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 397

gtggtgccgc ggggagcttg gaggaagag aa

32

<210> 398

<211> 32

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 398

gtggtgccgc ggcttcccaa gtgaggcaat gc

32

<210> 399

<211> 32

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 399

gtggtgccgc ggatggatgg ggaagacact gg 32

<210> 400

<211> 32

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 400

gtggtgccgc ggctgcaggc acattctctt cc 32

<210> 401

<211> 32

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 401

gtggtgccgc ggggatgaaa cagggcagga ac 32

<210> 402

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 402

gtggtgccgc ggttccaag tgaggcaatg c 31

<210> 403

<211> 32

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 403  
gtggtgccgc ggtttgcaga gccatgatga gg 32

<210> 404  
<211> 30  
<212> DNA  
<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 404  
gtggtgccgc ggcgacagcc aaaacagccg 30

<210> 405  
<211> 44  
<212> DNA  
<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 405  
aaaatatttt agctcctact cagactgtta ctctggtgac acaa 44

<210> 406  
<211> 44  
<212> DNA  
<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 406  
ttgtgtcacc agagtaacag tctgagtagg agctaaaata tttt 44

<210> 407  
<211> 36  
<212> DNA  
<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 407

tagctcctac tcagactggt actctggtga cacaac	36
<210> 408	
<211> 28	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic oligonucleotide	
<400> 408	
tagctcctac tcagactggt gaccaaac	28
<210> 409	
<211> 24	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic oligonucleotide	
<400> 409	
tagctcctac tctggtgaca caac	24
<210> 410	
<211> 28	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic oligonucleotide	
<400> 410	
tagctcctac tcagactggt gacacaac	28
<210> 411	
<211> 16	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic oligonucleotide	
<400> 411	
tagctcctac tcagac	16

<210> 412

<211> 27

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 412

tagctcctac tcagactgtt acacaac 27

<210> 413

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 413

tagctcctac tcagactgtg gtgaggtgac 30

<210> 414

<211> 32

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 414

tagctcctac tcagactctc tggtagacaca ac 32

<210> 415

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 415

tagctcctac tcagacctct ggtgacacaa c 31

<210> 416

<211> 32

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 416

tagtcctac tcaggctgtc tggtagacac ac 32

<210> 417

<211> 33

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 417

tagtcctac tcagactact ctggtgacac aac 33

<210> 418

<211> 28

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 418

tagtcctac tcagactgtt gacacaac 28

<210> 419

<211> 13

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 419

ctggtgacac aac 13

<210> 420

<211> 29

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 420

tagctcctac tcagactgtt agacacaac 29

<210> 421

<211> 33

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 421

tagctcctac tcagactgct ctggtagac aac 33

<210> 422

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 422

cagactgtta ctctggtgac 20

<210> 423

<211> 29

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 423

cagaccacct gtggtctcct actggtgac 29

<210> 424

<211> 18

<212> DNA

<213> Artificial Sequence



<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide  
<400> 424  
gctttgattt ccctaggg 18  
<210> 425  
<211> 25  
<212> DNA  
<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide  
<400> 425  
gctttgattt ccagttctta ggcaa 25  
<210> 426  
<211> 18  
<212> DNA  
<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide  
<400> 426  
gctttgattc ttaggcaa 18  
<210> 427  
<211> 62  
<212> DNA  
<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide  
<220><221> modified\_base  
<222> (13)..(53)  
<223> a, c, t, g, unknown or other  
<400> 427  
gctttgattt ccnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnncttaggc 60  
aa 62  
<210> 428

<211> 18

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 428

tcttaaccat taccatag 18

<210>

429

<211> 32

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 429

tcttaaccat taccatagag ttcttaggca ac 32

<210> 430

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 430

tcttaaccat taccaaagtt cttaggcaac 30

<210> 431

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 431

tcttaaccat taccataggt tcttaggcaa c 31

<210> 432

<211> 14

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 432

gaaccaaacc cact

14

<210> 433

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 433

gaaccaaacc cacttagggg ataa

24

<210> 434

<211> 2328

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 434

atgaggtctg actacaaaga ccatgacggt gattataaag atcatgacat cgattacaag 60

gatgacgatg acaagatggc cccaagaag aagaggaagg tgggcctcga gcccggagaa 120

aaaccgtaca agtgcctga gtgcgggaaa tcattctccg accctggggc gctcgtccgg 180

caccaaagga cgcatacagg ggaaaagccg tataagtgcc ccgagtgtgg aaagagcttc 240

tcgcagagag cccaccttga acgacaccaa agaacacaca ctggtgagaa accctataag 300

tgtccagagt gcggcaaate gtttagcaga tccgatgact tggcgcgcca ccagcggaca 360

cacacgggtg aaaagcccta caaatgcccg gagtgtggga agtcgttttc aaggtcggat 420

catctgacta cccatcagcg caccatacgg ggagcggccg cccgcgcctt ggtgaagagc 480

gagctggagg agaagaagtc cgagctgcgg cacaagctga agtacgtgcc ccacgagtac 540

atcgagctga tcgagatcgc caggaacccc acccaggacc gcatcctgga gatgaaggtg 600

atggagtctt tcatgaaggt gtacggctac aggggagagc acctgggcgg aagcagaaag 660

cctgacggcg ccatctatac agtgggcagc cccatcgatt acggcgtgat cgtggacaca 720  
aaggcctaca gggcggtcta caatctgcct atcgccagg cgcagagat gcagagatac 780  
gtgaaggaga accagaccgc gaataagcac atcaaccca acgagtgggtg gaaggtgtac 840  
cctagcagcg tgaccgagtt caagtccctg ttcgtgagcg gccacttcaa gggcaactac 900  
aagggccagc tgaccaggct gaaccgcaaa accaactgca atggcgccgt gctgagcgtg 960  
gaggagctgc tgatcggcgg cgagatgac aaagccggca ccctgacact ggaggaggtg 1020  
cggcgcaagt tcaacaacgg cgagatcaac ttcgagggca gaggaagtct tctaactgc 1080

ggtgacgtgg aggagaatcc cgccctaga tctgactaca aagaccatga cggtgattat 1140  
aaagatcatg acatcgatta caaggatgac gatgacaaga tggcccccga gaagaagagg 1200  
aaggtgggccc tcgagccggg agagaagccg tacaagtgtc ccgaatgtgg aaagagcttc 1260  
tcacagtcgg gggaccttcg gcgccaccag cgcacacata ctggtgaaaa gccgtataag 1320  
tgtccagaat gtggcaaatc attctccaca tcaggagacc tggtcaggca ccagcgaacc 1380  
cacacgggtg agaagcccta taagtgcgcc gaatgcggga agtccttttc gcagagagcc 1440  
cacttggaga ggcaccagag gaccatacg ggggagaaac cttacaagtg ccctgaatgc 1500

ggaaagtctg tctcgacca tctggatctc atcagacatc agagaacgca cactggagag 1560  
aaacctaca aatgtccga gtgtgggaag tcgttttagcc gaaaggacaa tctcaaaaac 1620  
catcaacgga cacacacggg tgaaaaacca tacaatgcc cggagtgcgg caaatcgttt 1680  
tcccaacttg cgcaattcgg ggcacaccaa cgcacgcata ctggagcggc cgcccgcc 1740  
ctggtgaaga gcgagctgga ggagaagaag tccgagctgc ggcacaagct gaagtacgtg 1800  
ccccacgagt acatcgagct gatcgagatc gccagggaacc ccaccagga ccgcatcctg 1860  
gagatgaagg tgatggagtt cttcatgaag gtgtacggct acaggggaga gcacctgggc 1920

ggaagcagaa agcctgacgg cgccatctat acagtgggca gcccacatga ttacggcgtg 1980  
atcgtggaca caaaggccta cagcggcggc tacaatctgc ctatcgcca ggccgacgag 2040  
atggagagat acgtggagga gaaccagaca cgggataagc acctaaccc caacgagtgg 2100  
tggaagggtg accctagcag cgtgaccgag ttcaagtcc tgttcgtgag cggccacttc 2160  
aagggaact acaaggccca gctgaccagg ctgaaccaca tcaccaactg caatggcgcc 2220  
gtgctgagcg tggaggagct gctgatcggc ggcgagatga tcaaagccgg caccctgaca 2280  
ctggaggagg tgcggcgcaa gttcaacaac ggcgagatca acttctga 2328

<210> 435

<211> 775

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
polypeptide

<400> 435

Met Arg Ser Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp His Asp

1 5 10 15

Ile Asp Tyr Lys Asp Asp Asp Asp Lys Met Ala Pro Lys Lys Lys Arg

20 25 30

Lys Val Gly Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys

35 40 45

Gly Lys Ser Phe Ser Asp Pro Gly Ala Leu Val Arg His Gln Arg Thr

50 55 60

His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe

65 70 75 80

Ser Gln Arg Ala His Leu Glu Arg His Gln Arg Thr His Thr Gly Glu

85 90 95

Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Ser Asp

100 105 110

Asp Leu Val Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys

115 120 125

Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Ser Asp His Leu Thr Thr

130 135 140

His Gln Arg Thr His Thr Gly Ala Ala Ala Arg Ala Leu Val Lys Ser

145 150 155 160

Glu Leu Glu Glu Lys Lys Ser Glu Leu Arg His Lys Leu Lys Tyr Val

165 170 175

Pro His Glu Tyr Ile Glu Leu Ile Glu Ile Ala Arg Asn Pro Thr Gln

180 185 190

Asp Arg Ile Leu Glu Met Lys Val Met Glu Phe Phe Met Lys Val Tyr

195 200 205

Gly Tyr Arg Gly Glu His Leu Gly Gly Ser Arg Lys Pro Asp Gly Ala

210 215 220

Ile Tyr Thr Val Gly Ser Pro Ile Asp Tyr Gly Val Ile Val Asp Thr  
225 230 235 240

Lys Ala Tyr Ser Gly Gly Tyr Asn Leu Pro Ile Gly Gln Ala Asp Glu  
245 250 255

Met Gln Arg Tyr Val Lys Glu Asn Gln Thr Arg Asn Lys His Ile Asn  
260 265 270

Pro Asn Glu Trp Trp Lys Val Tyr Pro Ser Ser Val Thr Glu Phe Lys  
275 280 285

Phe Leu Phe Val Ser Gly His Phe Lys Gly Asn Tyr Lys Ala Gln Leu  
290 295 300

Thr Arg Leu Asn Arg Lys Thr Asn Cys Asn Gly Ala Val Leu Ser Val  
305 310 315 320

Glu Glu Leu Leu Ile Gly Gly Glu Met Ile Lys Ala Gly Thr Leu Thr  
325 330 335

Leu Glu Glu Val Arg Arg Lys Phe Asn Asn Gly Glu Ile Asn Phe Glu  
340 345 350

Gly Arg Gly Ser Leu Leu Thr Cys Gly Asp Val Glu Glu Asn Pro Gly  
355 360 365

Pro Arg Ser Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp His Asp  
370 375 380

Ile Asp Tyr Lys Asp Asp Asp Asp Lys Met Ala Pro Lys Lys Lys Arg  
385 390 395 400

Lys Val Gly Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys  
405 410 415

Gly Lys Ser Phe Ser Gln Ser Gly Asp Leu Arg Arg His Gln Arg Thr  
420 425 430

His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe  
435 440 445

Ser Thr Ser Gly Ser Leu Val Arg His Gln Arg Thr His Thr Gly Glu  
450 455 460

Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Gln Arg Ala

465                      470                      475                      480  
 His Leu Glu Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys  
                                  485                      490                      495  
  
 Cys Pro Glu Cys Gly Lys Ser Phe Ser Thr His Leu Asp Leu Ile Arg  
                                  500                      505                      510  
 His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys  
                                  515                      520                      525  
 Gly Lys Ser Phe Ser Arg Lys Asp Asn Leu Lys Asn His Gln Arg Thr  
                                  530                      535                      540  
 His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe  
 545                      550                      555                      560  
  
 Ser Gln Leu Ala His Leu Arg Ala His Gln Arg Thr His Thr Gly Ala  
                                  565                      570                      575  
 Ala Ala Arg Ala Leu Val Lys Ser Glu Leu Glu Glu Lys Lys Ser Glu  
                                  580                      585                      590  
 Leu Arg His Lys Leu Lys Tyr Val Pro His Glu Tyr Ile Glu Leu Ile  
                                  595                      600                      605  
 Glu Ile Ala Arg Asn Pro Thr Gln Asp Arg Ile Leu Glu Met Lys Val  
                                  610                      615                      620  
  
 Met Glu Phe Phe Met Lys Val Tyr Gly Tyr Arg Gly Glu His Leu Gly  
 625                      630                      635                      640  
 Gly Ser Arg Lys Pro Asp Gly Ala Ile Tyr Thr Val Gly Ser Pro Ile  
                                  645                      650                      655  
 Asp Tyr Gly Val Ile Val Asp Thr Lys Ala Tyr Ser Gly Gly Tyr Asn  
                                  660                      665                      670  
 Leu Pro Ile Gly Gln Ala Asp Glu Met Glu Arg Tyr Val Glu Glu Asn  
                                  675                      680                      685  
  
 Gln Thr Arg Asp Lys His Leu Asn Pro Asn Glu Trp Trp Lys Val Tyr  
                                  690                      695                      700  
 Pro Ser Ser Val Thr Glu Phe Lys Phe Leu Phe Val Ser Gly His Phe  
 705                      710                      715                      720

Lys Gly Asn Tyr Lys Ala Gln Leu Thr Arg Leu Asn His Ile Thr Asn

725

730

735

Cys Asn Gly Ala Val Leu Ser Val Glu Glu Leu Leu Ile Gly Gly Glu

740

745

750

Met Ile Lys Ala Gly Thr Leu Thr Leu Glu Glu Val Arg Arg Lys Phe

755

760

765

Asn Asn Gly Glu Ile Asn Phe

770

775

<210> 436

<211> 4725

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 436

tggccactcc ctctatgcgc actcgtctgc tcggtggggc ctggcgacca aaggtcgcca 60

gacggacgtg ctttgcacgt cgggccccac cgagcgagcg agtgcgcata gagggagtgg 120

ccaactccat cactagaggt atggcagtga cgtaacgcga agcgcgcgaa gcgagaccac 180

gcctaccagc tgcgtcagca gtcaggtgac ctttttgcga cagtttgcga caccacgtgg 240

ccgctgaggg tatatatctt cgagtgcgcg aaccaggagc tccattttga ccgcgaaatt 300

tgaacagca gcagccatgc cggggttcta cgagattgtc ctgaaggtcc cgagtgcct 360

ggacgagcac ctgccgggca tttctaactc gtttgttaac tgggtggccg agaaggaatg 420

ggagctgccg ccggattctg acatggatcc gaattctgatt gagcaggcac ccctgaccgt 480

ggccgaaaag cttcagcgcg agttctctgt ggagtggcgc cgcgtgagta aggccccgga 540

ggccctcttt tttgtccagt tcgaaaaggg ggagacctac ttccacctgc acgtgctgat 600

tgagaccatc ggggtcaaatt ccatggtggt cggccgctac gtgagccaga ttaaagagaa 660

gctgtgtgacc cgcattctacc gcggggtcga gccgcagctt ccgaactggt tcgcggtgac 720

caaaacgcga aatggcgccg ggggcgggaa caagtggtg gacgactgct acatcccaa 780

ctacctgtc cccaagacc agcccagct ccagtggcg tggactaaca tggaccagta 840

tttaagcgc ttttgaatc tcgcggagcg taaacgctg gtggcgcagc atctgacga 900

cgtgtcgcag acgcaggagc agaacaaaga gaatcagaac cccaattctg acgcgccgt 960



catcaggta aaaacctcag ccaggtaacat ggagctggc gggtggctgg tggaccgcgg	1020
gatcacgtca gaaaagcaat ggattcagga ggaccaggcc tcgtacatct ctttcaacgc	1080
cgcctccaac tcgcggcccc agatcaaggc cgcgctggac aatgcctcca agatcatgag	1140
cctgacaaag acggctccgg actacctggi gggcagcaac ccgccggagg acattaccaa	1200
aaatcggatc taccaaatcc tggagctgaa cgggtacgat ccgcagtacg cggcctccgt	1260
cttctctgggc tgggcgcaaa agaagtccgg gaagaggaac accatctggc tctttgggcc	1320
ggccacgacg ggtaaaacca acatcgcgga agccatcgcc cagcccgatgc ctttctacgg	1380
ctgcgtaaac tggaccaatg agaactttcc cttcaacgat tgcgtcgaca agatggtgat	1440
ctggtgggag gagggcaaga tgacggccaa ggtcgtggag agcgccaagg ccattctggg	1500
cggaagcaag gtgcgcgtgg accaaaagtg caagtcatcg gccagatcg aaccactcc	1560
cgtgatcgc accccaaca ccaacatgtg cgcgctgatt gacgggaaca gcaccacctt	1620
cgagcatcag cagccgctgc aggaccggat gtttaattt gaacttacc gccgtttgga	1680
ccatgacttt ggggaaggta ccaaacagga agtaaaaggac ttttccggt gggcttccga	1740
tcacgtgact gacgtggctc atgagttcta cgtcagaaag ggtggagcta agaaacgccc	1800
cgcctccaat gacgcggatg taagcgagcc aaaacggcag tgcacgtcac ttgcgcagcc	1860
gacaacgtca gacgcggaag caccggcgga ctacgcggac aggtacaaa acaaatgttc	1920
tcgtcacgtg ggcatgaatc tgatgctttt tcctgtaaa acatgcgaga gaatgaatca	1980
aatttccaat gtctgtttta cgcatggta aagagactgt ggggaatgct tccttggat	2040
gtcagaatct caaccgctt ctgtcgtcaa aaagaagact tatcagaaac tgtgtccaat	2100
tcatcatatc ctgggaaggg cacccgagat tgcctgttcg gcctgcgatt tggccaatgt	2160
ggacttggat gactgtgttt ctgagcaata aatgacttaa accaggtatg gctgctgacg	2220
gttatcttcc agattggctc gaggacaacc tttctgaagg cattcgtgag tgggtggctc	2280
tgaaacctgg agtccctcaa cccaaagcga accaacaaca ccaggacaac cgtcggggtc	2340
ttgtgtctcc gggttacaaa tacctcgga cgggtaacgg actcgacaaa ggagagccgg	2400
tcaacgaggc ggacgcggca gccctcgaac acgacaaagc ttacgaccag cagctcaagg	2460
ccggtgacaa cccgtacctc aagtacaacc acgccgacgc cgagtttcag gagcgtcttc	2520
aagaagatac gtcttttggg ggcaaccttg gcagagcagt cttccaggcc aaaaagagga	2580
tccttgagcc tcttggctcg gttgaggaag cagctaaaac ggctcctgga aagaagaggc	2640
ctgtagatca gtctctcag gaaccggact catcatctgg tgttggcaaa tcgggcaaac	2700

agcctgccag aaaaagacta aatttcggtc agactggcga ctacagagtc gtcccagacc 2760  
 ctcaacctct cggagaacca ccagcagccc ccacaagttt gggatctaat acaatggctt 2820  
 caggcgggtgg cgcaccaatg gcagacaata acgagggtgc cgatggagtg ggtaattcct 2880  
 caggaaattg gcattgcgat tcccaatggc tgggcgacag agtcatcacc accagcacca 2940  
 gaacctgggc cctgcccact tacaacaacc atctctacaa gcaaatctcc agcgccttcaa 3000  
 cgggagcttc aaacgacaac cactactttg gctacagcac cccttggggg tatittgact 3060  
 ttaacagatt ccaactgccac ttctcaccac gtgactggca gcgactcatt aacaacaact 3120  
  
 ggggattccg gccaagaaa ctacagcttca agctcttcaa catccaagtt aaagaggtea 3180  
 cgcagaacga tggcacgacg actattgcc aataacctac cagcacggtt caagtgttta 3240  
 cggactcgga gtatcagctc ccgtacgtgc tcgggtcggc gcaccaaggc tgtctcccgc 3300  
 cgtttccagc ggacgtcttc atggtccttc agtatggata cctcacctg aacaacggaa 3360  
 gtcaagcggg gggacgctca tcttttact gcctggagta cttcccttcg cagatgctaa 3420  
 ggactggaaa taacttcaa ttacgtata ccttcgagga tgtacctttt cacagcagct 3480  
 acgtcacag ccagagtttg gatcgcttga tgaatctctt tattgatcag tatctgtact 3540  
  
 acctgaacag aacgcaagga acaacctctg gaacaaccaa ccaatcacgg ctgcttttta 3600  
 gccaggttgg gcctcagctt atgtctttgc aggccagaaa ttggctacct gggccctgct 3660  
 accggcaaca gagactttca aagactgcta acgacaacaa caacagtaac tttccttggg 3720  
 cagcggccag caaatatcat ctcaatggc gcgactcgct ggtgaatcca ggaccagcta 3780  
 tggccagtca caaggacgat gaagaaaaat ttttcctat gcacggcaat ctaatatattg 3840  
 gcaaagaagg gacaacggca agtaacgcag aattagataa tgtaatgatt acggatgaag 3900  
 aagagattcg taccaccaat cctgtggcaa cagagcagta tggaactgtg gcaaataact 3960  
  
 tgcagagctc aaatacagct cccacgacta gaactgtcaa tgatcagggg gccttacctg 4020  
 gcatggtgtg gcaagatcgt gacgtgtacc ttcaaggacc tatctgggca aagattcctc 4080  
 acacggatgg acactttcat ccttctctc tgatgggagg ctttggactg aaacatccgc 4140  
 ctctcaaat catgatcaaa aatactccgg taccggcaaa tctccgacg actttcagcc 4200  
 cggccaagtt tgcttcattt atcactcagt actccactgg acaggtcagc gtggaaattg 4260  
 agtgggagct acagaaagaa aacagcaaac gttggaatcc agagattcag tacacttcca 4320  
 actacaacaa gtctgttaat gtggacttta ctgtagacac taatgggtgtt tatagtgaac 4380  
  
 ctgcacctat tggaaccggg tatctcacac gaaacttgta atccttggtt atcaataaac 4440  
 cgtttaattc gtttcagttg aactttggct cttgtgcact tcttatctta tctgtttcc 4500  
 atggctactg cgtagataag cagcggcctg cggcgcttgc gcttcgcggt ttacaactgc 4560

tggttaatat ttaactctcg ccatacctct agtgatggag ttggccactc cctctatgcg 4620  
 cactcgctcg ctcggtgggg ccggacgtgc aaagcacgtc cgtctggcga cctttggtcg 4680  
 ccaggcccca ccgagcgagc gactgcgcat agagggagtg gccaa 4725

<210> 437

<211> 737

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polypeptide

<400> 437

Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser

1 5 10 15

Glu Gly Ile Arg Glu Trp Trp Ala Leu Lys Pro Gly Val Pro Gln Pro

20 25 30

Lys Ala Asn Gln Gln His Gln Asp Asn Arg Arg Gly Leu Val Leu Pro

35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Gly Asn Gly Leu Asp Lys Gly Glu Pro

50 55 60

Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp

65 70 75 80

Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala

85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly

100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Ile Leu Glu Pro

115 120 125

Leu Gly Leu Val Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg

130 135 140

Pro Val Asp Gln Ser Pro Gln Glu Pro Asp Ser Ser Ser Gly Val Gly

145 150 155 160

Lys Ser Gly Lys Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln Thr

165 170 175

Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro Pro

180 185 190

Ala Ala Pro Thr Ser Leu Gly Ser Asn Thr Met Ala Ser Gly Gly Gly

195 200 205

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

210 215 220

Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile

225 230 235 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu

245 250 255

Tyr Lys Gln Ile Ser Ser Ala Ser Thr Gly Ala Ser Asn Asp Asn His

260 265 270

Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe

275 280 285

His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn

290 295 300

Trp Gly Phe Arg Pro Lys Lys Leu Ser Phe Lys Leu Phe Asn Ile Gln

305 310 315 320

Val Lys Glu Val Thr Gln Asn Asp Gly Thr Thr Thr Ile Ala Asn Asn

325 330 335

Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro

340 345 350

Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala

355 360 365

Asp Val Phe Met Val Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly

370 375 380

Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro

385 390 395 400

Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Thr Phe

405 410 415

Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp

420 425 430  
Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Asn Arg

435 440 445  
Thr Gln Gly Thr Thr Ser Gly Thr Thr Asn Gln Ser Arg Leu Leu Phe

450 455 460  
Ser Gln Ala Gly Pro Gln Ser Met Ser Leu Gln Ala Arg Asn Trp Leu

465 470 475 480  
Pro Gly Pro Cys Tyr Arg Gln Gln Arg Leu Ser Lys Thr Ala Asn Asp

485 490 495  
Asn Asn Asn Ser Asn Phe Pro Trp Thr Ala Ala Ser Lys Tyr His Leu

500 505 510  
Asn Gly Arg Asp Ser Leu Val Asn Pro Gly Pro Ala Met Ala Ser His

515 520 525  
Lys Asp Asp Glu Glu Lys Phe Phe Pro Met His Gly Asn Leu Ile Phe

530 535 540  
Gly Lys Glu Gly Thr Thr Ala Ser Asn Ala Glu Leu Asp Asn Val Met

545 550 555 560  
Ile Thr Asp Glu Glu Glu Ile Arg Thr Thr Asn Pro Val Ala Thr Glu

565 570 575  
Gln Tyr Gly Thr Val Ala Asn Asn Leu Gln Ser Ser Asn Thr Ala Pro

580 585 590  
Thr Thr Arg Thr Val Asn Asp Gln Gly Ala Leu Pro Gly Met Val Trp

595 600 605  
Gln Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro

610 615 620  
His Thr Asp Gly His Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly

625 630 635 640  
Leu Lys His Pro Pro Pro Gln Ile Met Ile Lys Asn Thr Pro Val Pro

645 650 655  
Ala Asn Pro Pro Thr Thr Phe Ser Pro Ala Lys Phe Ala Ser Phe Ile

660 665 670

Thr Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu  
675 680 685  
Gln Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser  
690 695 700  
Asn Tyr Asn Lys Ser Val Asn Val Asp Phe Thr Val Asp Thr Asn Gly  
705 710 715 720  
Val Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn  
725 730 735  
Leu

<210> 438

<211> 406

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
polypeptide

<400> 438

Met Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp His Asp Ile Asp

1 5 10 15  
Tyr Lys Asp Asp Asp Lys Met Ala Pro Lys Lys Lys Arg Lys Val  
20 25 30  
Gly Arg Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly  
35 40 45  
Lys Ser Phe Ser Arg Lys Asp Ala Leu Arg Gly His Gln Arg Thr His  
50 55 60  
Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser  
65 70 75 80  
His Arg Thr Thr Leu Thr Asn His Gln Arg Thr His Thr Gly Glu Lys  
85 90 95  
Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Gln Arg Asn Ala  
100 105 110  
Leu Ala Gly His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys

115 120 125  
Pro Glu Cys Gly Lys Ser Phe Ser His Lys Asn Ala Leu Gln Asn His

130 135 140  
Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly  
145 150 155 160  
Lys Ser Phe Ser Asp Pro Gly His Leu Val Arg His Gln Arg Thr His  
165 170 175  
Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser  
180 185 190  
Thr Ser Gly Asn Leu Val Arg His Gln Arg Thr His Thr Gly Ala Ala

195 200 205  
Ala Arg Ala Leu Val Lys Ser Glu Leu Glu Glu Lys Lys Ser Glu Leu  
210 215 220  
Arg His Lys Leu Lys Tyr Val Pro His Glu Tyr Ile Glu Leu Ile Glu  
225 230 235 240  
Ile Ala Arg Asn Pro Thr Gln Asp Arg Ile Leu Glu Met Lys Val Met  
245 250 255  
Glu Phe Phe Met Lys Val Tyr Gly Tyr Arg Gly Glu His Leu Gly Gly

260 265 270  
Ser Arg Lys Pro Asp Gly Ala Ile Tyr Thr Val Gly Ser Pro Ile Asp  
275 280 285  
Tyr Gly Val Ile Val Asp Thr Lys Ala Tyr Ser Gly Gly Tyr Asn Leu  
290 295 300  
Pro Ile Gly Gln Ala Asp Glu Met Glu Arg Tyr Val Glu Glu Asn Gln  
305 310 315 320  
Thr Arg Asp Lys His Leu Asn Pro Asn Glu Trp Trp Lys Val Tyr Pro

325 330 335  
Ser Ser Val Thr Glu Phe Lys Phe Leu Phe Val Ser Gly His Phe Lys  
340 345 350  
Gly Asn Tyr Lys Ala Gln Leu Thr Arg Leu Asn His Ile Thr Asn Cys  
355 360 365

Asn Gly Ala Val Leu Ser Val Glu Glu Leu Leu Ile Gly Gly Glu Met

370

375

380

Ile Lys Ala Gly Thr Leu Thr Leu Glu Glu Val Arg Arg Lys Phe Asn

385

390

395

400

Asn Gly Glu Ile Asn Phe

405

<210> 439

<211> 406

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polypeptide

<400> 439

Met Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp His Asp Ile Asp

1

5

10

15

Tyr Lys Asp Asp Asp Asp Lys Met Ala Pro Lys Lys Lys Arg Lys Val

20

25

30

Gly Arg Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly

35

40

45

Lys Ser Phe Ser Gln Gln Arg Ser Leu Val Gly His Gln Arg Thr His

50

55

60

Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser

65

70

75

80

Asp Lys Lys Asp Leu Thr Arg His Gln Arg Thr His Thr Gly Glu Lys

85

90

95

Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Thr Ser Gly His

100

105

110

Leu Val Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys

115

120

125

Pro Glu Cys Gly Lys Ser Phe Ser Gln Arg Ala His Leu Glu Arg His

130

135

140

Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly



145                      150                      155                      160  
 Lys Ser Phe Ser Thr Ser Gly Ser Leu Val Arg His Gln Arg Thr His  
                                  165                      170                      175  
 Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser  
                                  180                      185                      190  
 Thr Ser Gly Asn Leu Val Arg His Gln Arg Thr His Thr Gly Ala Ala  
                                  195                      200                      205  
 Ala Arg Ala Leu Val Lys Ser Glu Leu Glu Glu Lys Lys Ser Glu Leu  
                                  210                      215                      220  
  
 Arg His Lys Leu Lys Tyr Val Pro His Glu Tyr Ile Glu Leu Ile Glu  
 225                      230                      235                      240  
 Ile Ala Arg Asn Pro Thr Gln Asp Arg Ile Leu Glu Met Lys Val Met  
                                  245                      250                      255  
 Glu Phe Phe Met Lys Val Tyr Gly Tyr Arg Gly Glu His Leu Gly Gly  
                                  260                      265                      270  
 Ser Arg Lys Pro Asp Gly Ala Ile Tyr Thr Val Gly Ser Pro Ile Asp  
                                  275                      280                      285  
  
 Tyr Gly Val Ile Val Asp Thr Lys Ala Tyr Ser Gly Gly Tyr Asn Leu  
                                  290                      295                      300  
 Pro Ile Gly Gln Ala Asp Glu Met Gln Arg Tyr Val Lys Glu Asn Gln  
 305                      310                      315                      320  
 Thr Arg Asn Lys His Ile Asn Pro Asn Glu Trp Trp Lys Val Tyr Pro  
                                  325                      330                      335  
 Ser Ser Val Thr Glu Phe Lys Phe Leu Phe Val Ser Gly His Phe Lys  
                                  340                      345                      350  
  
 Gly Asn Tyr Lys Ala Gln Leu Thr Arg Leu Asn Arg Lys Thr Asn Cys  
                                  355                      360                      365  
 Asn Gly Ala Val Leu Ser Val Glu Glu Leu Leu Ile Gly Gly Glu Met  
                                  370                      375                      380  
 Ile Lys Ala Gly Thr Leu Thr Leu Glu Glu Val Arg Arg Lys Phe Asn  
 385                      390                      395                      400

Asn Gly Glu Ile Asn Phe

405

<210> 440

<211> 320

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polypeptide

<400> 440

Met Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp His Asp Ile Asp

1 5 10 15

Tyr Lys Asp Asp Asp Asp Lys Met Ala Pro Lys Lys Lys Arg Lys Val

20 25 30

Gly Arg Leu Glu Pro Gly Glu Arg Pro Phe Gln Cys Arg Ile Cys Met

35 40 45

Arg Asn Phe Ser Ser Lys Gln Ala Leu Ala Val His Thr Arg Thr His

50 55 60

Thr Gly Glu Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser

65 70 75 80

Gln Ser Thr Thr Leu Lys Arg His Leu Arg Thr His Thr Gly Glu Lys

85 90 95

Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Arg Ser Asp His

100 105 110

Leu Ser Leu His Leu Lys Thr His Leu Arg Gly Ser Gln Leu Val Lys

115 120 125

Ser Glu Leu Glu Glu Lys Lys Ser Glu Leu Arg His Lys Leu Lys Tyr

130 135 140

Val Pro His Glu Tyr Ile Glu Leu Ile Glu Ile Ala Arg Asn Pro Thr

145 150 155 160

Gln Asp Arg Ile Leu Glu Met Lys Val Met Glu Phe Phe Met Lys Val

165 170 175

Tyr Gly Tyr Arg Gly Glu His Leu Gly Gly Ser Arg Lys Pro Asp Gly

180 185 190  
Ala Ile Tyr Thr Val Gly Ser Pro Ile Asp Tyr Gly Val Ile Val Asp  
195 200 205  
Thr Lys Ala Tyr Ser Gly Gly Tyr Asn Leu Pro Ile Gly Gln Ala Asp  
210 215 220  
Glu Met Gln Arg Tyr Val Lys Glu Asn Gln Thr Arg Asn Lys His Ile  
225 230 235 240  
Asn Pro Asn Glu Trp Trp Lys Val Tyr Pro Ser Ser Val Thr Glu Phe

245 250 255  
Lys Phe Leu Phe Val Ser Gly His Phe Lys Gly Asn Tyr Lys Ala Gln  
260 265 270  
Leu Thr Arg Leu Asn Arg Lys Thr Asn Cys Asn Gly Ala Val Leu Ser  
275 280 285  
Val Glu Glu Leu Leu Ile Gly Gly Glu Met Ile Lys Ala Gly Thr Leu  
290 295 300  
Thr Leu Glu Glu Val Arg Arg Lys Phe Asn Asn Gly Glu Ile Asn Phe

305 310 315 320

<210> 441

<211> 320

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
polypeptide

<400> 441

Met Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp His Asp Ile Asp  
1 5 10 15  
Tyr Lys Asp Asp Asp Asp Lys Met Ala Pro Lys Lys Lys Arg Lys Val  
20 25 30  
Gly Arg Leu Glu Pro Gly Glu Arg Pro Phe Gln Cys Arg Ile Cys Met

35 40 45

Arg Asn Phe Ser Arg Arg Ala His Leu Gln Asn His Thr Arg Thr His  
 50 55 60  
 Thr Gly Glu Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser  
 65 70 75 80  
 Gln Ser Thr Thr Leu Lys Arg His Leu Arg Thr His Thr Gly Glu Lys  
 85 90 95  
 Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Asp Gly Gly His  
 100 105 110  
 Leu Thr Arg His Leu Lys Thr His Leu Arg Gly Ser Gln Leu Val Lys  
 115 120 125  
 Ser Glu Leu Glu Glu Lys Lys Ser Glu Leu Arg His Lys Leu Lys Tyr  
 130 135 140  
 Val Pro His Glu Tyr Ile Glu Leu Ile Glu Ile Ala Arg Asn Pro Thr  
 145 150 155 160  
 Gln Asp Arg Ile Leu Glu Met Lys Val Met Glu Phe Phe Met Lys Val  
 165 170 175  
 Tyr Gly Tyr Arg Gly Glu His Leu Gly Gly Ser Arg Lys Pro Asp Gly  
 180 185 190  
 Ala Ile Tyr Thr Val Gly Ser Pro Ile Asp Tyr Gly Val Ile Val Asp  
 195 200 205  
 Thr Lys Ala Tyr Ser Gly Gly Tyr Asn Leu Pro Ile Gly Gln Ala Asp  
 210 215 220  
 Glu Met Glu Arg Tyr Val Glu Glu Asn Gln Thr Arg Asp Lys His Leu  
 225 230 235 240  
 Asn Pro Asn Glu Trp Trp Lys Val Tyr Pro Ser Ser Val Thr Glu Phe  
 245 250 255  
 Lys Phe Leu Phe Val Ser Gly His Phe Lys Gly Asn Tyr Lys Ala Gln  
 260 265 270  
 Leu Thr Arg Leu Asn His Ile Thr Asn Cys Asn Gly Ala Val Leu Ser  
 275 280 285  
 Val Glu Glu Leu Leu Ile Gly Gly Glu Met Ile Lys Ala Gly Thr Leu

290 295 300  
 Thr Leu Glu Glu Val Arg Arg Lys Phe Asn Asn Gly Glu Ile Asn Phe  
 305 310 315 320

<210> 442

<211> 42

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 442

caaactagaa atgcatctt ccttgatgtt ggaggtacct gc

42

<210> 443

<211> 406

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polypeptide

<400> 443

Met Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp His Asp Ile Asp

1 5 10 15

Tyr Lys Asp Asp Asp Asp Lys Met Ala Pro Lys Lys Lys Arg Lys Val

20 25 30

Gly Arg Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly

35 40 45

Lys Ser Phe Ser Arg Lys Asp Ala Leu Arg Gly His Gln Arg Thr His

50 55 60

Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser

65 70 75 80

His Arg Thr Thr Leu Thr Asn His Gln Arg Thr His Thr Gly Glu Lys

85 90 95

Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Gln Arg Asn Ala

100 105 110

Leu Ala Gly His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys

115	120	125	
Pro Glu Cys Gly Lys Ser Phe Ser His Lys Asn Ala Leu Gln Asn His			
130	135	140	
Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly			
145	150	155	160
Lys Ser Phe Ser Asp Pro Gly His Leu Val Arg His Gln Arg Thr His			
	165	170	175
Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser			
	180	185	190
Thr Ser Gly Asn Leu Val Arg His Gln Arg Thr His Thr Gly Ala Ala			
	195	200	205
Ala Arg Ala Leu Val Lys Ser Glu Leu Glu Glu Lys Lys Ser Glu Leu			
	210	215	220
Arg His Lys Leu Lys Tyr Val Pro His Glu Tyr Ile Glu Leu Ile Glu			
225	230	235	240
Ile Ala Arg Asn Pro Thr Gln Asp Arg Ile Leu Glu Met Lys Val Met			
	245	250	255
Glu Phe Phe Met Lys Val Tyr Gly Tyr Arg Gly Glu His Leu Gly Gly			
	260	265	270
Ser Arg Lys Pro Asp Gly Ala Ile Tyr Thr Val Gly Ser Pro Ile Asp			
	275	280	285
Tyr Gly Val Ile Val Asp Thr Lys Ala Tyr Ser Gly Gly Tyr Asn Leu			
	290	295	300
Pro Ile Gly Gln Ala Asp Glu Met Glu Arg Tyr Val Glu Glu Asn Gln			
305	310	315	320
Thr Arg Asp Lys His Leu Asn Pro Asn Glu Trp Trp Lys Val Tyr Pro			
	325	330	335
Ser Ser Val Thr Glu Phe Lys Phe Leu Phe Val Ser Gly His Phe Lys			
	340	345	350
Gly Asn Tyr Lys Ala Gln Leu Thr Arg Leu Asn His Ile Thr Asn Cys			
	355	360	365

Asn Gly Ala Val Leu Ser Val Glu Glu Leu Leu Ile Gly Gly Glu Met  
370 375 380

Ile Lys Ala Gly Thr Leu Thr Leu Glu Glu Val Arg Arg Lys Phe Asn  
385 390 395 400

Asn Gly Glu Ile Asn Phe  
405

<210> 444  
<211> 406  
<212> PRT  
<213> Artificial Sequence  
<220><223> Description of Artificial Sequence: Synthetic  
polypeptide  
<400> 444

Met Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp His Asp Ile Asp  
1 5 10 15

Tyr Lys Asp Asp Asp Asp Lys Met Ala Pro Lys Lys Lys Arg Lys Val  
20 25 30

Gly Arg Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly  
35 40 45

Lys Ser Phe Ser Gln Gln Arg Ser Leu Val Gly His Gln Arg Thr His  
50 55 60

Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser  
65 70 75 80

Asp Lys Lys Asp Leu Thr Arg His Gln Arg Thr His Thr Gly Glu Lys  
85 90 95

Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Thr Ser Gly His  
100 105 110

Leu Val Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys  
115 120 125

Pro Glu Cys Gly Lys Ser Phe Ser Gln Arg Ala His Leu Glu Arg His  
130 135 140

Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly

145 150 155 160

Lys Ser Phe Ser Thr Ser Gly Ser Leu Val Arg His Gln Arg Thr His

165 170 175

Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser

180 185 190

Thr Ser Gly Asn Leu Val Arg His Gln Arg Thr His Thr Gly Ala Ala

195 200 205

Ala Arg Ala Leu Val Lys Ser Glu Leu Glu Glu Lys Lys Ser Glu Leu

210 215 220

Arg His Lys Leu Lys Tyr Val Pro His Glu Tyr Ile Glu Leu Ile Glu

225 230 235 240

Ile Ala Arg Asn Pro Thr Gln Asp Arg Ile Leu Glu Met Lys Val Met

245 250 255

Glu Phe Phe Met Lys Val Tyr Gly Tyr Arg Gly Glu His Leu Gly Gly

260 265 270

Ser Arg Lys Pro Asp Gly Ala Ile Tyr Thr Val Gly Ser Pro Ile Asp

275 280 285

Tyr Gly Val Ile Val Asp Thr Lys Ala Tyr Ser Gly Gly Tyr Asn Leu

290 295 300

Pro Ile Gly Gln Ala Asp Glu Met Gln Arg Tyr Val Lys Glu Asn Gln

305 310 315 320

Thr Arg Asn Lys His Ile Asn Pro Asn Glu Trp Trp Lys Val Tyr Pro

325 330 335

Ser Ser Val Thr Glu Phe Lys Phe Leu Phe Val Ser Gly His Phe Lys

340 345 350

Gly Asn Tyr Lys Ala Gln Leu Thr Arg Leu Asn Arg Lys Thr Asn Cys

355 360 365

Asn Gly Ala Val Leu Ser Val Glu Glu Leu Leu Ile Gly Gly Glu Met

370 375 380

Ile Lys Ala Gly Thr Leu Thr Leu Glu Glu Val Arg Arg Lys Phe Asn



385 390 395 400

Asn Gly Glu Ile Asn Phe

405

<210> 445

<211> 25

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<220><221> misc\_feature

<222> (1)..(1)

<223> May or may not be present

<220><221> misc\_feature

<222> (25)..(25)

<223> May or may not be present

<400> 445

tatctgccca tgactggcgc aggga

25

<210> 446

<211> 320

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polypeptide

<400> 446

Met Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp His Asp Ile Asp

1 5 10 15

Tyr Lys Asp Asp Asp Lys Met Ala Pro Lys Lys Lys Arg Lys Val

20 25 30

Gly Arg Leu Glu Pro Gly Glu Arg Pro Phe Gln Cys Arg Ile Cys Met

35 40 45

Arg Asn Phe Ser Ser Lys Gln Ala Leu Ala Val His Thr Arg Thr His

50 55 60

Thr Gly Glu Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser  
 65                                70                                75                                80  
 Gln Ser Thr Thr Leu Lys Arg His Leu Arg Thr His Thr Gly Glu Lys  
    85                                90                                95  
 Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Arg Ser Asp His  
    100                                105                                110  
 Leu Ser Leu His Leu Lys Thr His Leu Arg Gly Ser Gln Leu Val Lys  
    115                                120                                125  
  
 Ser Glu Leu Glu Glu Lys Lys Ser Glu Leu Arg His Lys Leu Lys Tyr  
    130                                135                                140  
 Val Pro His Glu Tyr Ile Glu Leu Ile Glu Ile Ala Arg Asn Pro Thr  
 145                                150                                155                                160  
 Gln Asp Arg Ile Leu Glu Met Lys Val Met Glu Phe Phe Met Lys Val  
    165                                170                                175  
 Tyr Gly Tyr Arg Gly Glu His Leu Gly Gly Ser Arg Lys Pro Asp Gly  
    180                                185                                190  
  
 Ala Ile Tyr Thr Val Gly Ser Pro Ile Asp Tyr Gly Val Ile Val Asp  
    195                                200                                205  
 Thr Lys Ala Tyr Ser Gly Gly Tyr Asn Leu Pro Ile Gly Gln Ala Asp  
    210                                215                                220  
 Glu Met Gln Arg Tyr Val Lys Glu Asn Gln Thr Arg Asn Lys His Ile  
 225                                230                                235                                240  
 Asn Pro Asn Glu Trp Trp Lys Val Tyr Pro Ser Ser Val Thr Glu Phe  
    245                                250                                255  
  
 Lys Phe Leu Phe Val Ser Gly His Phe Lys Gly Asn Tyr Lys Ala Gln  
    260                                265                                270  
 Leu Thr Arg Leu Asn Arg Lys Thr Asn Cys Asn Gly Ala Val Leu Ser  
    275                                280                                285  
 Val Glu Glu Leu Leu Ile Gly Gly Glu Met Ile Lys Ala Gly Thr Leu  
    290                                295                                300  
 Thr Leu Glu Glu Val Arg Arg Lys Phe Asn Asn Gly Glu Ile Asn Phe

305 310 315 320

<210> 447

<211> 320

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polypeptide

<400> 447

Met Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp His Asp Ile Asp

1 5 10 15

Tyr Lys Asp Asp Asp Lys Met Ala Pro Lys Lys Lys Arg Lys Val

20 25 30

Gly Arg Leu Glu Pro Gly Glu Arg Pro Phe Gln Cys Arg Ile Cys Met

35 40 45

Arg Asn Phe Ser Arg Arg Ala His Leu Gln Asn His Thr Arg Thr His

50 55 60

Thr Gly Glu Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser

65 70 75 80

Gln Ser Thr Thr Leu Lys Arg His Leu Arg Thr His Thr Gly Glu Lys

85 90 95

Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Asp Gly Gly His

100 105 110

Leu Thr Arg His Leu Lys Thr His Leu Arg Gly Ser Gln Leu Val Lys

115 120 125

Ser Glu Leu Glu Glu Lys Lys Ser Glu Leu Arg His Lys Leu Lys Tyr

130 135 140

Val Pro His Glu Tyr Ile Glu Leu Ile Glu Ile Ala Arg Asn Pro Thr

145 150 155 160

Gln Asp Arg Ile Leu Glu Met Lys Val Met Glu Phe Phe Met Lys Val

165 170 175

Tyr Gly Tyr Arg Gly Glu His Leu Gly Gly Ser Arg Lys Pro Asp Gly

180 185 190  
Ala Ile Tyr Thr Val Gly Ser Pro Ile Asp Tyr Gly Val Ile Val Asp  
195 200 205  
Thr Lys Ala Tyr Ser Gly Gly Tyr Asn Leu Pro Ile Gly Gln Ala Asp  
210 215 220  
Glu Met Glu Arg Tyr Val Glu Glu Asn Gln Thr Arg Asp Lys His Leu  
225 230 235 240

Asn Pro Asn Glu Trp Trp Lys Val Tyr Pro Ser Ser Val Thr Glu Phe  
245 250 255  
Lys Phe Leu Phe Val Ser Gly His Phe Lys Gly Asn Tyr Lys Ala Gln  
260 265 270  
Leu Thr Arg Leu Asn His Ile Thr Asn Cys Asn Gly Ala Val Leu Ser  
275 280 285  
Val Glu Glu Leu Leu Ile Gly Gly Glu Met Ile Lys Ala Gly Thr Leu  
290 295 300

Thr Leu Glu Glu Val Arg Arg Lys Phe Asn Asn Gly Glu Ile Asn Phe  
305 310 315 320

<210> 448

<211> 42

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 448

actagaaatg ccattcttct tgatgttgga ggtacctgct ct

42

<210> 449

<211> 406

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
polypeptide

<400>

> 449

Met Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp His Asp Ile Asp  
1 5 10 15  
Tyr Lys Asp Asp Asp Asp Lys Met Ala Pro Lys Lys Lys Arg Lys Val  
20 25 30  
Gly Arg Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly  
35 40 45  
Lys Ser Phe Ser His Arg Thr Thr Leu Thr Asn His Gln Arg Thr His  
50 55 60  
  
Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser  
65 70 75 80  
Gln Arg Asn Ala Leu Ala Gly His Gln Arg Thr His Thr Gly Glu Lys  
85 90 95  
Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser His Lys Asn Ala  
100 105 110  
Leu Gln Asn His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys  
115 120 125  
  
Pro Glu Cys Gly Lys Ser Phe Ser Asp Pro Gly His Leu Val Arg His  
130 135 140  
Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly  
145 150 155 160  
Lys Ser Phe Ser Thr Ser Gly Asn Leu Val Arg His Gln Arg Thr His  
165 170 175  
Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser  
180 185 190  
  
Gln Ser Ser Asn Leu Val Arg His Gln Arg Thr His Thr Gly Ala Ala  
195 200 205  
Ala Arg Ala Leu Val Lys Ser Glu Leu Glu Glu Lys Lys Ser Glu Leu  
210 215 220  
Arg His Lys Leu Lys Tyr Val Pro His Glu Tyr Ile Glu Leu Ile Glu  
225 230 235 240  
Ile Ala Arg Asn Pro Thr Gln Asp Arg Ile Leu Glu Met Lys Val Met

245 250 255

Glu Phe Phe Met Lys Val Tyr Gly Tyr Arg Gly Glu His Leu Gly Gly

260 265 270

Ser Arg Lys Pro Asp Gly Ala Ile Tyr Thr Val Gly Ser Pro Ile Asp

275 280 285

Tyr Gly Val Ile Val Asp Thr Lys Ala Tyr Ser Gly Gly Tyr Asn Leu

290 295 300

Pro Ile Gly Gln Ala Asp Glu Met Glu Arg Tyr Val Glu Glu Asn Gln

305 310 315 320

Thr Arg Asp Lys His Leu Asn Pro Asn Glu Trp Trp Lys Val Tyr Pro

325 330 335

Ser Ser Val Thr Glu Phe Lys Phe Leu Phe Val Ser Gly His Phe Lys

340 345 350

Gly Asn Tyr Lys Ala Gln Leu Thr Arg Leu Asn His Ile Thr Asn Cys

355 360 365

Asn Gly Ala Val Leu Ser Val Glu Glu Leu Leu Ile Gly Gly Glu Met

370 375 380

Ile Lys Ala Gly Thr Leu Thr Leu Glu Glu Val Arg Arg Lys Phe Asn

385 390 395 400

Asn Gly Glu Ile Asn Phe

405

<210> 450

<211> 406

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polypeptide

<400> 450

Met Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp His Asp Ile Asp

1 5 10 15

Tyr Lys Asp Asp Asp Asp Lys Met Ala Pro Lys Lys Lys Arg Lys Val

20	25	30	
Gly Arg Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly			
35	40	45	
Lys Ser Phe Ser Gln Arg Asn Ala Leu Ala Gly His Gln Arg Thr His			
50	55	60	
Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser			
65	70	75	80
Gln Gln Arg Ser Leu Val Gly His Gln Arg Thr His Thr Gly Glu Lys			
85	90	95	
Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Asp Lys Lys Asp			
100	105	110	
Leu Thr Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys			
115	120	125	
Pro Glu Cys Gly Lys Ser Phe Ser Thr Ser Gly His Leu Val Arg His			
130	135	140	
Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly			
145	150	155	160
Lys Ser Phe Ser Gln Arg Ala His Leu Glu Arg His Gln Arg Thr His			
165	170	175	
Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser			
180	185	190	
Thr Ser Gly Ser Leu Val Arg His Gln Arg Thr His Thr Gly Ala Ala			
195	200	205	
Ala Arg Ala Leu Val Lys Ser Glu Leu Glu Glu Lys Lys Ser Glu Leu			
210	215	220	
Arg His Lys Leu Lys Tyr Val Pro His Glu Tyr Ile Glu Leu Ile Glu			
225	230	235	240
Ile Ala Arg Asn Pro Thr Gln Asp Arg Ile Leu Glu Met Lys Val Met			
245	250	255	
Glu Phe Phe Met Lys Val Tyr Gly Tyr Arg Gly Glu His Leu Gly Gly			
260	265	270	

Ser Arg Lys Pro Asp Gly Ala Ile Tyr Thr Val Gly Ser Pro Ile Asp

275 280 285  
Tyr Gly Val Ile Val Asp Thr Lys Ala Tyr Ser Gly Gly Tyr Asn Leu  
290 295 300  
Pro Ile Gly Gln Ala Asp Glu Met Gln Arg Tyr Val Lys Glu Asn Gln  
305 310 315 320  
Thr Arg Asn Lys His Ile Asn Pro Asn Glu Trp Trp Lys Val Tyr Pro  
325 330 335  
Ser Ser Val Thr Glu Phe Lys Phe Leu Phe Val Ser Gly His Phe Lys

340 345 350  
Gly Asn Tyr Lys Ala Gln Leu Thr Arg Leu Asn Arg Lys Thr Asn Cys  
355 360 365  
Asn Gly Ala Val Leu Ser Val Glu Glu Leu Leu Ile Gly Gly Glu Met  
370 375 380  
Ile Lys Ala Gly Thr Leu Thr Leu Glu Glu Val Arg Arg Lys Phe Asn  
385 390 395 400  
Asn Gly Glu Ile Asn Phe  
405

<210> 451

<211> 378

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
polypeptide

<400> 451

Met Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp His Asp Ile Asp  
1 5 10 15  
Tyr Lys Asp Asp Asp Asp Lys Met Ala Pro Lys Lys Lys Arg Lys Val  
20 25 30  
Gly Arg Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly  
35 40 45



Lys Ser Phe Ser Gln Arg Asn Ala Leu Ala Gly His Gln Arg Thr His  
 50 55 60  
 Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser  
 65 70 75 80  
 His Lys Asn Ala Leu Gln Asn His Gln Arg Thr His Thr Gly Glu Lys  
 85 90 95  
 Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Asp Pro Gly His  
 100 105 110  
  
 Leu Val Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys  
 115 120 125  
 Pro Glu Cys Gly Lys Ser Phe Ser Thr Ser Gly Asn Leu Val Arg His  
 130 135 140  
 Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly  
 145 150 155 160  
 Lys Ser Phe Ser Gln Ser Ser Asn Leu Val Arg His Gln Arg Thr His  
 165 170 175  
  
 Thr Gly Ala Ala Ala Arg Ala Leu Val Lys Ser Glu Leu Glu Glu Lys  
 180 185 190  
 Lys Ser Glu Leu Arg His Lys Leu Lys Tyr Val Pro His Glu Tyr Ile  
 195 200 205  
 Glu Leu Ile Glu Ile Ala Arg Asn Pro Thr Gln Asp Arg Ile Leu Glu  
 210 215 220  
 Met Lys Val Met Glu Phe Phe Met Lys Val Tyr Gly Tyr Arg Gly Glu  
 225 230 235 240  
  
 His Leu Gly Gly Ser Arg Lys Pro Asp Gly Ala Ile Tyr Thr Val Gly  
 245 250 255  
 Ser Pro Ile Asp Tyr Gly Val Ile Val Asp Thr Lys Ala Tyr Ser Gly  
 260 265 270  
 Gly Tyr Asn Leu Pro Ile Gly Gln Ala Asp Glu Met Glu Arg Tyr Val  
 275 280 285  
 Glu Glu Asn Gln Thr Arg Asp Lys His Leu Asn Pro Asn Glu Trp Trp

290

295

300

Lys Val Tyr Pro Ser Ser Val Thr Glu Phe Lys Phe Leu Phe Val Ser

305 310 315 320

Gly His Phe Lys Gly Asn Tyr Lys Ala Gln Leu Thr Arg Leu Asn His

325 330 335

Ile Thr Asn Cys Asn Gly Ala Val Leu Ser Val Glu Glu Leu Leu Ile

340 345 350

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

355 360 365

Arg Lys Phe Asn Asn Gly Glu Ile Asn Phe

370 375

<210> 452

<211> 378

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polypeptide

<400> 452

Met Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp His Asp Ile Asp

1 5 10 15

Tyr Lys Asp Asp Asp Asp Lys Met Ala Pro Lys Lys Lys Arg Lys Val

20 25 30

Gly Arg Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly

35 40 45

Lys Ser Phe Ser Gln Gln Arg Ser Leu Val Gly His Gln Arg Thr His

50 55 60

Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser

65 70 75 80

Asp Lys Lys Asp Leu Thr Arg His Gln Arg Thr His Thr Gly Glu Lys

85 90 95

Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Thr Ser Gly His

100	105	110	
Leu Val Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys			
115	120	125	
Pro Glu Cys Gly Lys Ser Phe Ser Gln Arg Ala His Leu Glu Arg His			
130	135	140	
Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly			
145	150	155	160
Lys Ser Phe Ser Thr Ser Gly Ser Leu Val Arg His Gln Arg Thr His			
165	170	175	
Thr Gly Ala Ala Ala Arg Ala Leu Val Lys Ser Glu Leu Glu Glu Lys			
180	185	190	
Lys Ser Glu Leu Arg His Lys Leu Lys Tyr Val Pro His Glu Tyr Ile			
195	200	205	
Glu Leu Ile Glu Ile Ala Arg Asn Pro Thr Gln Asp Arg Ile Leu Glu			
210	215	220	
Met Lys Val Met Glu Phe Phe Met Lys Val Tyr Gly Tyr Arg Gly Glu			
225	230	235	240
His Leu Gly Gly Ser Arg Lys Pro Asp Gly Ala Ile Tyr Thr Val Gly			
245	250	255	
Ser Pro Ile Asp Tyr Gly Val Ile Val Asp Thr Lys Ala Tyr Ser Gly			
260	265	270	
Gly Tyr Asn Leu Pro Ile Gly Gln Ala Asp Glu Met Gln Arg Tyr Val			
275	280	285	
Lys Glu Asn Gln Thr Arg Asn Lys His Ile Asn Pro Asn Glu Trp Trp			
290	295	300	
Lys Val Tyr Pro Ser Ser Val Thr Glu Phe Lys Phe Leu Phe Val Ser			
305	310	315	320
Gly His Phe Lys Gly Asn Tyr Lys Ala Gln Leu Thr Arg Leu Asn Arg			
325	330	335	
Lys Thr Asn Cys Asn Gly Ala Val Leu Ser Val Glu Glu Leu Leu Ile			
340	345	350	

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

355 360 365

Arg Lys Phe Asn Asn Gly Glu Ile Asn Phe

370 375

<210> 453

<211> 26

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<220><221> misc\_feature

<222> (1)..(1)

<223> May or may not be present

<220><221> misc\_feature

<222> (26)..(26)

<223> May or may not be present

<400> 453

tgccatcttc ctgatgttg gaggtg

26

<210> 454

<211> 322

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polypeptide

<400> 454

Met Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp His Asp Ile Asp

1 5 10 15

Tyr Lys Asp Asp Asp Asp Lys Met Ala Pro Lys Lys Lys Arg Lys Val

20 25 30

Gly Arg Leu Glu Pro Gly Glu Arg Pro Phe Gln Cys Arg Ile Cys Met

35 40 45

Arg Asn Phe Ser Ser Pro Ser Lys Leu Ala Arg His Thr Arg Thr His

50	55	60	
Thr Gly Glu Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser			
65	70	75	80
Val Arg His Asn Leu Thr Arg His Leu Arg Thr His Thr Gly Glu Lys			
	85	90	95
Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Gln Arg Asn Asn			
	100	105	110
Leu Gly Arg His Leu Lys Thr His Thr Gly Ala Ala Ala Arg Ala Leu			
	115	120	125
Val Lys Ser Glu Leu Glu Glu Lys Lys Ser Glu Leu Arg His Lys Leu			
	130	135	140
Lys Tyr Val Pro His Glu Tyr Ile Glu Leu Ile Glu Ile Ala Arg Asn			
145	150	155	160
Pro Thr Gln Asp Arg Ile Leu Glu Met Lys Val Met Glu Phe Phe Met			
	165	170	175
Lys Val Tyr Gly Tyr Arg Gly Glu His Leu Gly Gly Ser Arg Lys Pro			
	180	185	190
Asp Gly Ala Ile Tyr Thr Val Gly Ser Pro Ile Asp Tyr Gly Val Ile			
	195	200	205
Val Asp Thr Lys Ala Tyr Ser Gly Gly Tyr Asn Leu Pro Ile Gly Gln			
	210	215	220
Ala Asp Glu Met Gln Arg Tyr Val Lys Glu Asn Gln Thr Arg Asn Lys			
225	230	235	240
His Ile Asn Pro Asn Glu Trp Trp Lys Val Tyr Pro Ser Ser Val Thr			
	245	250	255
Glu Phe Lys Phe Leu Phe Val Ser Gly His Phe Lys Gly Asn Tyr Lys			
	260	265	270
Ala Gln Leu Thr Arg Leu Asn Arg Lys Thr Asn Cys Asn Gly Ala Val			
	275	280	285
Leu Ser Val Glu Glu Leu Leu Ile Gly Gly Glu Met Ile Lys Ala Gly			
	290	295	300

Thr Leu Thr Leu Glu Glu Val Arg Arg Lys Phe Asn Asn Gly Glu Ile  
 305 310 315 320  
 Asn Phe

<210> 455

<211> 322

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polypeptide

<400> 455

Met Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp His Asp Ile Asp  
 1 5 10 15  
 Tyr Lys Asp Asp Asp Lys Met Ala Pro Lys Lys Lys Arg Lys Val  
 20 25 30

Gly Arg Leu Glu Pro Gly Glu Arg Pro Phe Gln Cys Arg Ile Cys Met  
 35 40 45  
 Arg Asn Phe Ser Ile Pro Asn His Leu Ala Arg His Thr Arg Thr His  
 50 55 60  
 Thr Gly Glu Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser  
 65 70 75 80  
 Gln Ser Ala His Leu Lys Arg His Leu Arg Thr His Thr Gly Glu Lys  
 85 90 95

Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser His His Asn Ser  
 100 105 110  
 Leu Thr Arg His Leu Lys Thr His Thr Gly Ala Ala Ala Arg Ala Leu  
 115 120 125  
 Val Lys Ser Glu Leu Glu Glu Lys Lys Ser Glu Leu Arg His Lys Leu  
 130 135 140  
 Lys Tyr Val Pro His Glu Tyr Ile Glu Leu Ile Glu Ile Ala Arg Asn  
 145 150 155 160

Pro Thr Gln Asp Arg Ile Leu Glu Met Lys Val Met Glu Phe Phe Met

165 170 175  
 Lys Val Tyr Gly Tyr Arg Gly Glu His Leu Gly Gly Ser Arg Lys Pro  
 180 185 190  
 Asp Gly Ala Ile Tyr Thr Val Gly Ser Pro Ile Asp Tyr Gly Val Ile  
 195 200 205  
 Val Asp Thr Lys Ala Tyr Ser Gly Gly Tyr Asn Leu Pro Ile Gly Gln  
 210 215 220

Ala Asp Glu Met Glu Arg Tyr Val Glu Glu Asn Gln Thr Arg Asp Lys  
 225 230 235 240  
 His Leu Asn Pro Asn Glu Trp Trp Lys Val Tyr Pro Ser Ser Val Thr  
 245 250 255  
 Glu Phe Lys Phe Leu Phe Val Ser Gly His Phe Lys Gly Asn Tyr Lys  
 260 265 270  
 Ala Gln Leu Thr Arg Leu Asn His Ile Thr Asn Cys Asn Gly Ala Val  
 275 280 285

Leu Ser Val Glu Glu Leu Leu Ile Gly Gly Glu Met Ile Lys Ala Gly  
 290 295 300  
 Thr Leu Thr Leu Glu Glu Val Arg Arg Lys Phe Asn Asn Gly Glu Ile  
 305 310 315 320  
 Asn Phe

<210> 456

<211> 6626

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
 polynucleotide

<400> 456

gggggggggg ggggggggtt ggccactccc tctctgcgcg ctcgctcgct cactgaggcc 60

gggcgaccaa aggtcgcccg acgcccgggc ttgcccggg cggcctcagt gagcgagcga 120

gcgcgcagag agggagtggc caactccatc actaggggtt cctagatctg aattcggtac 180

ccgttacata acttacggta aatggccgcg ctggctgacc gccaacgac ccccgcccat 240

tgacgtcaat aatgacgtat gttcccatag taacgccaat agggactttc cattgacgtc	300
aatgggtgga gtatttacgg taaactgccc acttggcagt acatcaagtg tatcatatgc	360
caagtacgcc ccctattgac gtcaatgacg gtaaatggcc cgcctggcat tatgcccagt	420
acatgacctt atgggacttt cctacttggc agtacatcta cgtattagtc atcgctatta	480
ccatggtgat gcggttttgg cagtacatca atgggcgtgg atagcggttt gactcacggg	540
gatttccaag tctccacccc attgacgtca atgggagttt gttttggcac caaatcaac	600
gggactttcc aaaatgtcgt aacaactccg ccccatgac gcaaatgggc ggtaggcgtg	660
tacggtggga ggtctatata agcagagctc gtttagtgaa ccgtcagatc gcctggagac	720
gccatccacg ctgttttgac ctccatagaa gacaccggga ccgatccagc ctccggactc	780
tagaggatcc ggtactcgag gaactgaaaa accagaaagt taactggtaa gtttagtctt	840
tttgtctttt atttcaggtc ccggatccgg tgggtgtgca aatcaaagaa ctgctcctca	900
gtggatgttg cctttacttc taggcctgta cggaagtgtt acttctgctc taaaagctgc	960
ggaattgtac ccgcggcccg ggatccaccg gtggctagcg tctataggcc cacccttg	1020
gtggaattcg ccatgaggtc tgactacaaa gaccatgacg gtgattataa agatcatgac	1080
atcgattaca aggatgacga tgacaagatg gcccccaaga agaagaggaa ggtgggcctc	1140
gagcccgag aaaaaccgta caagtgcctt gagtgcggga aatcattctc cgacctggg	1200
gcgctcgtec ggcacaaaag gacgcataca ggggaaaagc cgtataagtg ccccagtg	1260
ggaagagctt tctcgagag agcccacctt gaacgacacc aaagaacaca cactggtgag	1320
aaacctata agtgtccaga gtgcggcaaa tcgttttagca gatccgatga cttggtgcgc	1380
caccagcga cacacacggg tgaaaagccc tacaatgcc cggagtgtgg gaagtcgttt	1440
tcaagtcgg atcatctgac taccatcag cgcaccata cgggagcggc cgcccgcgc	1500
ctggtgaaga gcgagctgga ggagaagaag tccgagctgc ggcacaagct gaagtacgtg	1560
ccccacgagt acatcgagct gatcgagatc gccaggaacc ccaccagga ccgcatcctg	1620
gagatgaagg tgatggagtt cttcatgaag gtgtacggct acaggggaga gcacctgggc	1680
ggaagcagaa agcctgacgg cgccatctat acagtgggca gcccacatga ttacggcgtg	1740
atcgtggaca caaaggccta cagcggcggc tacaatctgc ctatcgcca gcccacgag	1800
atgcagagat acgtgaagga gaaccagacc cggaataagc acatcaacc caacgagtgg	1860
tggaagggtg accctagcag cgtgaccgag ttcaagttcc tgttcgtgag cggccacttc	1920
aagggaact acaaggccca gctgaccagg ctgaaccgca aaaccaactg caatggcgcc	1980
gtgctgagcg tggaggagct gctgatcggc ggcgagatga tcaaagccgg caccctgaca	2040
ctggaggagg tgcggcgcaa gttcaacaac ggcgagatca acttcgaggg cagaggaagt	2100



cttctaacat gcggtgacgt ggaggagaat cccggcccta gatctgacta caaagaccat 2160

gacggtgatt ataaagatca tgacatcgat tacaaggatg acgatgacaa gatggccccc 2220

aagaagaaga ggaaggtggg cctcgagccg ggagagaagc cgtacaagtg tcccgaatgt 2280

ggaaagagct tctcacagtc gggggacctt cggcgccacc agcgcacaca tactggtgaa 2340

aagccgtata agtgtccaga atgtggcaaa tcattctcca catcaggag cctggtcagg 2400

caccagcgaa cccacacggg tgagaagccc tataagtgcc ccgaatgcgg gaagtccttt 2460

tcgcagagag ccacttggg gaggcaccag aggaccata cgggggagaa accttacaag 2520

tgccctgaat gcggaaagtc gttctcgacc catctggatc tcacagaca tcagagaacg 2580

cacactggag agaaacccta caaatgtccc gagtgtggga agtcgttttag ccgaaaggac 2640

aatctcaaaa accatcaacg gacacacacg ggtgaaaaac catacaaatg cccggagtgc 2700

ggcaaatcgt tttcccaact tgcgcacttg cgggcacacc aacgcacgca tactggagcg 2760

gccgcccgcg ccctggtgaa gagcgagctg gaggagaaga agtccgagct gcggcacaag 2820

ctgaagtacg tgccccacga gtacatcgag ctgatcgaga tcgccaggaa ccccccag 2880

gaccgcatcc tggagatgaa ggtgatggag ttcttcatga aggtgtacgg ctacagggga 2940

gagcacctgg gcggaagcag aaagcctgac ggcgccatct atacagtggg cagcccatc 3000

gattacggcg tgatcgtgga cacaaaggcc tacagcggcg gctacaatct gcctatcggc 3060

caggccgacg agatggagag atacgtggag gagaaccaga cacgggataa gcacctcaac 3120

cccaacgagt ggtggaaggt gtaccctagc agcgtgaccg agttcaagtt cctgttcgtg 3180

agcggccact tcaagggcaa ctacaaggcc cagctgacca ggctgaacca catcaccaac 3240

tgcaatggcg ccgtgctgag cgtggaggag ctgctgatcg gcggcgagat gatcaaagcc 3300

ggcacctga cactggagga ggtgcggcgc aagtccaaca acggcgagat caacttctga 3360

ttaattaact aatctagagt cgactagagc tcgctgatca gcctcgactg tgccttctag 3420

ttgccagcca tctgttgttt gccctcccc cgtgccttcc ttgacctgg aaggtgccac 3480

tcccactgtc ctttctaata aaaatgagga aattgcatcg cattgtctga gtaggtgtca 3540

ttctattctg gggggtgggg tggggcagga cagcaagggg gaggattggg aagacaatag 3600

caggcatgct ggggagagat ctaggaacct ctagtgatgg agttggccac tccctctctg 3660

cgcgctcgct cgctcactga ggccgcccgg gcaaagcccg ggctcgggc gacctttggt 3720

cgcccgccct cagtgagcga gcgagcgcgc agagaggag tggccaacct ccccccccc 3780

ccccctgcag ccagctgca ttaatgaatc ggccaacgcg cggggagagg cggtttgcgt 3840

attgggcgct cticcgcctc ctgcctcact gactcgtgc gctcggctcgt tcggctgcgg 3900  
cgagcggat cagctcactc aaaggcggta atacggttat ccacagaatc aggggataac 3960  
gcaggaaaga acatgtgagc aaaaggccag caaaaggcca ggaaccgtaa aaaggccgcg 4020  
ttgctggcgt ttttccatag gctccgcccc cctgacgagc atcacaaaaa tcgacgtca 4080  
agtcagaggt ggcgaaaccc gacaggacta taaagatacc aggcgtttcc ccctggaagc 4140  
tccctcgtgc gctctcctgt tccgacctg ccgcttaccg gatacctgtc cgcctttctc 4200  
ccttcgggaa gcgtggcgct ttctcaatgc tcacgctgta ggtatctcag ttcgggtgag 4260

gtcgttcgct ccaagctggg ctgtgtgcac gaaccccccg ttcagcccga ccgtgcgcc 4320  
ttatccggt aatctcgtct tgagtcacac ccggttaagac acgacttatc gccactggca 4380  
gcagccactg gtaacaggat tagcagagcg aggtatgtag gcggtgtac agagtcttg 4440  
aagtgtggc ctaactacgg ctacactaga aggacagtat ttggtatctg cgtctgtctg 4500  
aagccagtta ccttcgaaa aagagttagt agctcttgat ccggcaaaca aaccaccgt 4560  
ggtagcgggt gttttttgt ttgcaagcag cagattacgc gcagaaaaaa aggatctcaa 4620  
gaagatcctt tgatctttc tacggggtct gacgctcagt ggaacgaaaa ctacgttaa 4680

gggattttgg tcatgagatt atcaaaaagg atcttcacct agatcctttt aaattaaaaa 4740  
tgaagttta aatcaatcta aagtatatat gagttaactt ggtctgacag ttaccaatgc 4800  
ttaatcagt aggcacctat ctacgcgac tgtctatttc gttcatccat agttgcctga 4860  
ctccccgtc ttagataac tacgatacgg gagggcttac catctggccc cagtgtgca 4920  
atgataccgc gagaccacg ctacccggt ccagatttat cagcaataaa ccagccagcc 4980  
ggaagggccg agcgcagaag tggctctgca actttatccg cctccatcca gtctattaat 5040  
tgttgccggg aagctagagt aagtagttcg ccagttaata gtttgcgcaa cgttgttgcc 5100

attgctacag gcacgtggt gtcacgctcg tcgtttggta tggttcatt cagctccggt 5160  
tccaacgat caaggcgagt tacatgatcc ccatgttgt gcaaaaaagc ggtagctcc 5220  
ttcggtctc cgatcgttgt cagaagtaag ttggccgcag tgttatcact catggttatg 5280  
gcagcactgc ataattctct tactgtcatg ccatccgtaa gatgcttttc tgtactggt 5340  
gagtactcaa ccaagtcatt ctgagaatag tgtatgcggc gaccgagttg ctcttgcccc 5400  
gcgtcaatac gggataatac gcgcacacat agcagaactt taaaagtgt catcattgga 5460  
aaacgttctt cggggcgaaa actctcaagg atcttaccgc tgttgagatc cagttcgatg 5520

taaccactc gtgcaccaa ctgatcttca gcactttta ctttaccag cgtttctggg 5580  
tgagcaaaaa caggaaggca aaatgccgca aaaaaggga taaggcgac acggaatgt 5640  
tgaatactca tactcttct tttcaatat tattgaagca tttatcaggg ttattgtctc 5700

atgagcggat acatatttga atgtatttag aaaaataaac aaataggggt tccgcgcaca 5760  
 tttccccgaa aagtgccacc tgacgtctaa gaaaccatta ttatcatgac attaacctat 5820  
 aaaaataggc gtatcacgag gccctttcgt ctgcgcggtt tcggtgatga cggtgaaaac 5880  
 ctctgacaca tgcagctccc ggagacggtc acagcttgtc tgtaagcgga tgccgggagc 5940

agacaagccc gtcaggcgcc gtcagcgggt gttggcgggt gtcggggctg gcttaactat 6000  
 gcggcatcag agcagattgt actgagagtg caccatatgc ggtgtgaaat accgcacaga 6060  
 tgcgtaagga gaaaataacc catcaggaaa ttgtaaacgt taatattttg ttaaaattcg 6120  
 cgttaaattt ttgttaaatc agctcatttt ttaaccaata ggccgaaatc ggcaaaatcc 6180  
 cttataaatc aaaagaatag accgagatag ggttgagtgt tgttcagtt tggaacaaga 6240  
 gtccactatt aaagaacgtg gactccaacg tcaaaggcg aaaaaccgtc tatcaggcg 6300  
 atggcccact acgtgaacca tcacccta atcaagttttt ggggtcgagg tgccgtaaag 6360

cactaaatcg gaaccctaaa gggagcccc gatttagagc ttgacgggga aagccggcga 6420  
 acgtggcgag aaaggaaggg aagaaagcga aaggagcggg cgctaggcg ctggcaagtg 6480  
 tagcgtcac gctgcgcgta accaccacac ccgccgcgt taatgcgcg ctacaggcg 6540  
 cgtcgcgcca ttcgccattc aggctacga actgttggga agggcgatcg gtgcgggcct 6600  
 cttegtatt acgccagctg gctgca 6626

<210> 457

<211> 22

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 457

gagaggttat gtggctttac ca 22

<210> 458

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 458

ctgcgtagt ccaaaacaaa 20

<210> 459

<211> 3252

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 459

```

atgaaaagga actacattct ggggctggac atcgggatta caagcgtggg gtatgggatt      60

attgactatg aaacaaggga cgtgatcgac gcaggcgtca gactgttcaa ggaggccaac      120
gtggaaaaca atgaggagcg gagaagcaag aggggagcca ggcgcctgaa acgacggaga      180
aggcacagaa tccagagggt gaagaaactg ctgttcgatt acaacctgct gaccgaccat      240
tctgagctga gtggaattaa tccttatgaa gccagggtga aaggcctgag tcagaagctg      300
tcagaggaag agttttccgc agctctgctg cacctggcta agcggcgagg agtgcataac      360
gtcaatgagg tggaagagga caccggcaac gagctgtcta caaaggaaca gatctcacgc      420
aatagcaaag ctctggaaga gaagtatgtc gcagagctgc agctggaacg gctgaagaaa      480

gatggcgagg tgagagggtc aattaatagg ttcaagacaa gcgactacgt caaagaagcc      540
aagcagctgc tgaaagtga gaaggcttac caccagctgg atcagagctt catcgatact      600
tatatcgacc tgctggagac tcggagaacc tactatgagg gaccaggaga agggagcccc      660
ttcgatgga aagacatcaa ggaatggtac gagatgctga tgggacattg cacctathtt      720
ccagaagagc tgagaagcgt caagtacgt tataacgcag atctgtacaa cgccctgaat      780
gacctgaaca acctggtcat caccagggat gaaaacgaga aactggaata ctatgagaag      840
ttccagatca tcgaaaacgt gtttaagcag aaaaaaagc ctacactgaa acagattgct      900

aaggagatcc tggtaacga agaggacatc aagggtacc gggtgacaag cactggaaaa      960
ccagagtcca ccaatctgaa agtgtatcac gatattaagg acatcacagc acgaaaagaa     1020
atcattgaga acgccgaact gctggatcag attgctaaga tctgactat ctaccagagc     1080
tccgaggaca tccaggaaga gctgactaac ctgaacagcg agctgacca ggaagagatc     1140
gaacagatta gtaatctgaa ggggtacacc ggaacacaca acctgtccct gaaagctatc     1200
aatctgattc tggatgagct gtggcataca aacgacaatc agattgcaat cttaaccgg      1260
ctgaagctgg tcccaaaaaa ggtggacctg agtcagcaga aagagatccc aaccacactg     1320

gtggacgatt tcattctgtc acccgtggtc aagcgagct tcattccagag catcaaagtg     1380
atcaacgcca tcatcaagaa gtacggcctg cccaatgata tcattatcga gctggctagg     1440

```

gagaagaaca gcaaggacgc acagaagatg atcaatgaga tgcagaaacg aaaccggcag	1500
accaatgaac gcattgaaga gattatccga actaccggga aagagaacgc aaagtacctg	1560
attgaaaaaa tcaagctgca cgatatgcag gagggaaagt gtctgtattc tctggaggcc	1620
atccccctgg aggacctgct gaacaatcca ttcaactacg aggtcgatca tattatcccc	1680
agaagcgtgt ccttcgacaa ttcctttaac aacaagggtc tggtaagca ggaagagaac	1740
tctaaaaagg gcaataggac tcctttccag tacctgtcta gttcagattc caagatctct	1800
tacgaaacct ttaaaaagca cattctgaat ctggccaaag gaaagggccg catcagcaag	1860
acaaaaaagg agtacctgct ggaagagcgg gacatcaaca gattctccgt ccagaaggat	1920
tttattaacc ggaatctggt ggacacaaga tacgtactc gcggcctgat gaatctgctg	1980
cgatcctatt tccgggtgaa caatctggat gtgaaagtca agtccatcaa cggcgggttc	2040
acatcttttc tgaggcgcaa atggaagttt aaaaaggagc gcaacaaagg gtacaagcac	2100
catgccgaag atgctctgat tatcgcaaat gccgacttca tctttaagga gtggaaaaag	2160
ctggacaaaag ccaagaaagt gatggagaac cagatgttcg aagagaagca ggccgaatct	2220
atgcccgaaa tcgagacaga acaggagtac aaggagattt tcatcactcc tcaccagatc	2280
aagcatatca aggatttcaa ggactacaag tactctcacc gggaggataa aaagcccaac	2340
agagagctga tcaatgacac cctgtatagt acaagaaaag acgataaggg gaataacctg	2400
attgtgaaca atctgaacgg actgtacgac aaagataatg acaagctgaa aaagctgac	2460
aacaaaagtc ccgagaagct gctgatgtac caccatgac ctcagacata tcagaaactg	2520
aaagctgatta tggagcagta cggcgacgag aagaaccac tgtataagta ctatgaagag	2580
actgggaact acctgaccaa gtatagcaaa aaggataatg gcccctgat caagaagatc	2640
aagtactatg ggaacaagct gaatgcccat ctggacatca cagacgatta ccctaacagt	2700
cgcaacaagg tggtaagct gtactgaag ccatacagat tcgatgtcta tctggacaac	2760
ggcgtgtata aatttgtgac tgtcaagaat ctggatgtca tcaaaaagga gaactactat	2820
gaagtgaata gcaagtgtca cgaagggtc aaaaagctga aaaagattag caaccaggca	2880
gagttcatcg cctcctttta caacaacgac ctgattaaga tcaatggcga actgtatagg	2940
gtcatcgggg tgaacaatga tctgtgaac cgcatgaag tgaatatgat tgacatcact	3000
taccgagagt atctggaaaa catgaatgat aagcgcccc ctcgaattat caaaacaatt	3060
gcctctaaga ctcagagtat caaaaagtac tcaaccgaca ttctgggaaa cctgtatgag	3120
gtgaagagca aaaagcacc tcagattatc aaaaaggga gcggaggcaa gcgtcctgct	3180
gtactaaga aagctggta agctaagaaa aagaaaggat cctaccata cgatgttcca	3240
gattacgctt aa	3252

<210> 460

<211> 83

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 460

gttttagtac tctggaagca gaatctacta aaacaaggca aaatgccgtg tttatctcgt	60
caacttggtg gcgagatttt ttt	83

<210> 461

<211> 3333

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
polynucleotide

<400> 461

atggtgccta agaagaagag aaaggtggct gccttcaaac ctaattcaat caactacatc	60
ctcggcctcg atatcgcat cgcacccgtc ggctgggcga tggtagaaat tgacgaagaa	120
gaaaacccca tccgcctgat tgatttgggc gtgcgcgtat ttgagcgtgc cgaagtaccg	180
aaaacaggcg actcccttgc catggcaagg cgtttggcgc gcagtgttcg cgcctgacc	240
cgccgtcgcg cccaccgctt gcttcggacc cgccgcctat tgaaacgcga aggcgtatta	300
caagccgcca attttgacga aaacggcttg attaaatcct taccgaatac accatggcaa	360
cttcgcgcag ccgcattaga ccgcaaactg acgccttttag agtggtcggc agtcttgttg	420
catttaatac aacatcgagg ctatttatcg caacggaaaa acgagggcga aactgccgat	480
aaggagcttg gcgctttgct taaaggcgta gccggcaatg cccatgcctt acagacaggc	540
gatttccgca caccggccga attggcttta aataaatttg agaaagaaag cggccatata	600
cgcaatcagc gcagcgatta ttgcatacgc ttcagccgca aagattttaca ggcggagctg	660
atattgctgt ttgaaaaaca aaaagaattt ggcaatccgc atgtttcagg cggccttaaa	720
gaaggtattg aaaccctact gatgacgcaa cgccctgccc tgtccggcga tgccgttcaa	780
aaaatgttgg ggcattgcac cttcgaaccg gcagagccga aagccgctaa aaacacctac	840
acagccgaac gtttcatctg gctgaccaag ctgaacaacc tgcgtatttt agagcaaggc	900
agcgagcggc cattgaccga taccgaacgc gccacgttta tggacgagcc atacagaaaa	960

tccaaactga cttacgcaca agcccgtgaag ctgctgggtt tagaagatac cgcctttttc	1020
aaaggcttgc gctatggtaa agacaatgcc gaagcctcaa cattgatgga aatgaaggcc	1080
taccatgcca tcagccgtgc actggaaaaa gaaggattga aagacaaaaa atccccatta	1140
aacctttctc ccgaattaca agacgaaatc ggcacggcat tctccctgtt caaaaccgat	1200
gaagacatta caggccgtct gaaagaccgt atacagcccg aaatcttaga agcgtgtttg	1260
aaacacatca gtttcgataa gttcgtccaa atttccttga aagcattgcg ccgaattgtg	1320
cctctaattgg aacaaggcaa acgttacgat gaagcctgcg ccgaaatcta cggagaccat	1380
tacggcaaga agaatacggg agaaaagatt tatctgccgc cgattcccgc cgacgaaatc	1440
cgcaaccccg tcgtcttgcg cgccttatct caagcacgta aggtcattaa cggcgtggta	1500
cgccgttacg gctccccagc tcgtatccat attgaaactg caagggaagt aggtaaatcg	1560
tttaaagacc gcaaagaaat tgagaaacgc caagaagaaa accgcaaaga ccgggaaaaa	1620
gccgcgcgca aattccgaga gtatttcccc aattttgtcg gagaacccaa atccaaagat	1680
attctgaaac tgcgcctgta cgagcaacaa cacggcaaat gcctgtattc gggcaaagaa	1740
atcaacttag gccgtctgaa gaaaaaaggc tatgtcgaaa tcgaccatgc cctgccgttc	1800
tcgcgcacat gggacgacag tttcaacaat aaagtactgg tattgggcag cgaaaaccaa	1860
aacaaaggca atcaaacccc ttacgaatac ttcaacggca aagacaacag ccgcgaatgg	1920
caggaattta aagcgcgtgt cgaaaccagc cgtttccgcg gcagtaaaaa acaacggatt	1980
ctgctgcaaa aattcgatga agacggcttt aaagaacgca atctgaacga cacgcgtac	2040
gtcaaccgtt tctgtgtca atttgttgc gaccgtatgc ggctgacagg taaaggcaag	2100
aaacgtgtct ttgcatccaa cggacaaatt accaatctgt tgcgcggctt ttggggattg	2160
cgaaagtgc gtgcggaaaa cgaccgcat cagccttgg acgccgtcgt cgttgccctgc	2220
tcgaccgttg ccatgcagca gaaaattacc cgttttgtac gctataaaga gatgaacgcg	2280
tttgacggta aaacataga caaagaaaca ggagaagtgc tgcatcaaaa aacacacttc	2340
ccacaacctt gggaattttt cgcacaagaa gtcatgattc gcgtcttcgg caaacggac	2400
ggcaaaccgg aattcgaaga agccgatacc ctagaaaaac tgcgcacgtt gcttgccgaa	2460
aaattatcat ctgccccga agccgtacac gaatacgtta cgccactgtt tgtttcacgc	2520
gcgcccaatc ggaagatgag cgggcaaggg catatggaga ccgtcaaatc cgccaaacga	2580
ctggacgaag gcgtcagcgt gttgcgcgta ccgctgacac agttaaact gaaagacttg	2640
gaaaaaatgg tcaatcgga ggcggaacct aagctatacg aagcactgaa agcacggctg	2700

gaagcacata aagacgatcc tgccaaagcc ttgcccagc cgttttataa atacgataaa 2760  
gcaggcaacc gcaccaaca ggtaaaagcc gtacgcgtag agcaagtaca gaaaaccggc 2820  
glatgggtgc gcaaccataa cggattgcc gacaacgcaa ccatggtgcg cgtagatgtg 2880  
tttgagaaag gcgacaagta ttatctggta ccgatttaca gttggcaggt agcgaaaggg 2940  
attttgccgg atagggctgt tgtacaagga aaagatgaag aagattggca acttattgat 3000  
gatagtttca actttaaatt ctcatcacac cctaattgatt tagtcgaggt tataacaaaa 3060  
aaagctagaa tgtttggta ctttgcagc tgccatcgag gcacaggtaa tatcaatata 3120

cgcattcatg atcttgatca taaaattggc aaaaatggaa tactggaagg tatcggcgtc 3180  
aaaaccgccc tticattcca aaaataccaa attgacgaac tgggcaaaga aatcagacca 3240  
tgccgtctga aaaaacgccc gcctgtccgt tacccatagc atgttccaga ttacgtgca 3300  
gctccagcag cgaagaaaaa gaagctggat taa 3333

<210> 462

<211> 120

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 462

gttgtagctc cttttctcat ttcgaaacg aaatgagaac cgttgctaca ataaggccgt 60

ctgaaaagat gtgccgaac gctctgccc ttaaagcttc tgctttaagg ggcttttttt 120

<210> 463

<211> 144

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 463

gttgtagctc cttttctcat ttcgagtcg tacaatgaaa attgtgcac tgcgaaatga 60

gaaccgttgc tacaataagg ccgtctgaaa agatgtgccg caacgctctg ccccttaaag 120

cttctgcttt aaggggcttt tttt 144

<210> 464

<211> 24

<212> DNA



<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 464

tcgggtttat tacagggaca gcag 24

<210> 465

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 465

tctaaggccg agtcttatga gcag 24

<210> 466

<211> 4564

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 466

atggactaca aagaccatga cgggtattat aaagatcatg acatcgatta caaggatgac 60

gatgacaaga tggcccccaa gaagaagagg aaggtgggcc gcggaatgga caagaagtac 120

tccattgggc tcgccatcgg cacaacagc gtcggctggg ccgtcattac ggacgagtac 180

aaggtgccga gcaaaaaatt caaagttctg ggcaataccg atcgccacag cataaagaag 240

aacctcattg gcgccctcct gttcgactcc ggggaaaccg ccgaagccac gcggctcaaa 300

agaacagcac ggcgagata taccgcaga aagaatcgga tctgctacct gcaggagatc 360

tttagtaatg agatggctaa ggtggatgac tctttcttcc ataggctgga ggagtccttt 420

ttggtggagg aggataaaaa gcacgagcgc cacccaatct ttggcaatat cgtggacgag 480

gtggcgtacc atgaaaagta cccaaccata tatcatctga ggaagaagct ttagacagt 540

actgataagg ctgacttgcg gttgatctat ctgcgctgg cgcataatgat caaatctcg 600

ggacacttcc tcatcgaggg ggacctgaac ccagacaaca gcgatgtcga caaactcttt 660

atccaactgg ttcagactta caatcagctt ttcgaagaga acccgatcaa cgcacccgga 720

gttgacgcca aagcaatcct gagcgctagg ctgtccaaat cccggcggct cgaaaacctc	780
atcgcacagc tccctgggga gaagaagaac ggctgtttg gtaatcttat cgccctgtca	840
ctcgggctga cccccaactt taaatctaac ttcgacctgg ccgaagatgc caagcttcaa	900
ctgagcaaag acacctacga tgatgatctc gacaatctgc tggcccagat cggcgaccag	960
tacgcagacc tttttttggc ggcaaagaac ctgtcagacg ccattctgct gagtgatatt	1020
ctgcgagtga acacggagat caccaaagct ccgctgagcg ctagtatgat caagcgctat	1080
gatgagcacc accaagactt gactttgctg aaggcccttg tcagacagca actgcctgag	1140
aagtacaagg aaattttctt cgatcagtct aaaaatggct acgccggata cattgacggc	1200
ggagcaagcc aggaggaatt ttacaaattt attaagccca tcttgaaaaa aatggacggc	1260
accgaggagc tgctggtaaa gcttaacaga gaagatctgt tgcgcaaaca gcgcactttc	1320
gacaatggaa gcatccccc ccagattcac ctgggcgaac tgcacgtat cctcaggcgg	1380
caagaggatt tctaccctt tttgaaagat aacagggaag agattgagaa aatcctcaca	1440
tttcggatac cctactatgt agggccctc gcccggggaa attccagatt cgcgtggatg	1500
actcgcaaat cagaagagac catcactccc tggaacttcg aggaagtcgt ggataagggg	1560
gcctctgccc agtccttcac cgaaaggatg actaaccttg ataaaaatct gcctaacgaa	1620
aaggtgcttc ctaaacactc tctgtgtac gagtacttca cagtttataa cgagctcacc	1680
aaggtcaaat acgtcacaga agggatgaga aagccagcat tcctgtctgg agagcagaag	1740
aaagctatcg tggacctcct cttcaagacg aaccggaaag ttaccgtgaa acagctcaaa	1800
gaagactatt tcaaaaagat tgaatgtttc gactctgttg aaatcagcgg agtggaggat	1860
cgcttcaacg catccctggg aacgtatcac gatctctga aaatcattaa agacaaggac	1920
ttcctggaca atgaggagaa cgaggacatt cttgaggaca ttgtcctcac ccttacgttg	1980
tttgaagata gggagatgat tgaagaacgc ttgaaaactt acgtcatct cttcgacgac	2040
aaagtcatga aacagctcaa gaggcgccga tatacaggat gggggcggct gtcaagaaaa	2100
ctgatcaatg ggatccgaga caagcagagt ggaaagacaa tectggattt tcttaagtcc	2160
gatggatttg ccaaccggaa cttcatgcag ttgatccatg atgactctct cacctttaag	2220
gaggacatcc agaaagcaca agtttctggc cagggggaca gtcttcacga gcacatcgt	2280
aatcttgcat gtagccagc tatcaaaaag ggaatactgc agaccgttaa ggtcgtggat	2340
gaactcgtca aagtaatggg aaggcataag cccgagaata tcgttatcga gatggcccga	2400
gagaacaaaa ctaccagaa gggacagaag aacagtaggg aaaggatgaa gaggattgaa	2460

gagggtataa aagaactggg gtcccaaattc ctttaaggaac acccagttga aaacacccag 2520

cttcagaatg agaagctcta cctgtactac ctgcagaacg gcaggacat gtacgtggat 2580

caggaactgg acatcaatcg gctctccgac tacgacgtgg atgccatcgt gccccagtct 2640

tttctcaaag atgattctat tgataataaa gtgttgacaa gatccgataa aaatagaggg 2700

aagagtata acgtccctc agaagaagt gtcaagaaaa tgaataatta ttggcgccag 2760

ctgctgaacg ccaaactgat cacacaacgg aagttcgata atctgactaa ggctgaacga 2820

ggtggcctgt ctgagttgga taaagccggc ttcatacaaa ggagccttgt tgagacacgc 2880

cagatcacca agcagctggc ccaaattctc gattcacgca tgaacaccaa gtacgatgaa 2940

aatgacaaac tgattcgaga ggtgaaagt attactctga agtctaagct ggtctcagat 3000

ttcagaaagg actttcagtt ttataagggt agagagatca acaattacca ccatgcgcat 3060

gatgcctacc tgaatgcagt ggtaggcact gcacttatca aaaaatatcc caagcttgaa 3120

tctgaatttg ttacggaga ctataaagt tacgatgtta ggaaatgat cgcaaagtct 3180

gagcaggaag taggcaaggc caccgctaag tacttctttt acagcaatat tatgaatttt 3240

ttcaagaccg agattacact ggccaatgga gagattcgga agcgaccact tatcgaaaca 3300

aacggagaaa caggagaaat cgtgtgggac aagggtaggg atttcgcgac agtccggaag 3360

gtcctgtcca tgccgcaggt gaacatcgtt aaaaagaccg aagtacagac cggaggttc 3420

tccaaggaaa gtatcctccc gaaaaggaac agcgacaagc tgatcgacg caaaaaagat 3480

tgggacccca agaaatacgg cggattcgat tctctacag tcgcttacag tgtactggtt 3540

gtggccaaag tggagaaagg gaagtctaaa aaactcaaaa gcgtcaagga actgctgggc 3600

atcacaatca tggagcgatc aagcttcgaa aaaaacccca tcgactttct cgaggcgaaa 3660

ggataataag aggtcaaaaa agacctcatc attaaagctt ccaagtactc tctctttgag 3720

cttgaacacg gccggaaacg aatgctcgt agtgcgggag agctgcagaa aggtaacgag 3780

ctggcactgc cctctaaata cgttaatttc ttgtatctgg ccagccacta tgaaaagctc 3840

aaagggtctc ccgaagataa tgagcagaag cagctgttcg tggaacaaca caaacactac 3900

cttgatgaga tcatcgagca aataagcgaa ttctcaaaa gagtgatcct cgccgacgct 3960

aacctgata aggtgctttc tgcttacaat aagcacaggg ataagcccat caggagcag 4020

gcagaaaaca ttatccactt gtttactctg accaacttgg gcgcgcctgc agccttcaag 4080

tacttcgaca ccacataga cagaaagcgg tacacctcta caaaggaggt cctggacgcc 4140

acactgattc atcagtcaat tacggggctc tatgaaacaa gaatcgacct ctctcagctc 4200

ggtaggagaca gcagggctga cccaagaag aagaggaagg tggctagcga tgctaagtca 4260  
ctgactgcct ggtcccgac actggtgacc ttcaaggatg tgtttgtgga cttcaccagg 4320  
gaggagtgga agctgctgga cactgctcag cagatcctgt acagaaatgt gatgctggag 4380  
aactataaga acctggtttc cttgggttat cagcttacta agccagatgt gatcctccgg 4440  
ttggagaagg gagaagagcc ctggctggtg gagagagaaa ttcaccaaga gacccatcct 4500  
gattcagaga ctgcatttga aatcaaatca tcagttccga aaaagaaacg caaagttgct 4560  
agcg 4564

<210> 467

<211> 388

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 467

taagcttcag tttttcctta gttcctgtta catttctgtg tgtctccatt agtgacctcc 60  
catagtccaa gcatgagcag ttctggccag gccctgtcg gggtcagtgc cccacccccg 120  
ccttctggtt ctgtgtaacc ttctaagcaa accttctggc tcaagcacag caatgctgag 180  
tcatgatgag tcatgctgag gcttaggggtg tgtgccaga tgttctcagc ctagagtgat 240  
gactcctatc tgggtcccca gcaggatgct tacagggcag atggcaaaaa aaaggagaag 300

ctgaccacct gactaaaact ccacctcaaa cgcatcata aagaaaatgg atgcctgaga 360  
cagaatgtga catattctag aatatatt 388

<210> 468

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 468

cagccgctcg ctgcagcag 19

<210> 469

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 469

ctgctgcagc gagcggctg

19

<210> 470

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 470

gctgggtgtc ccattgaaa

19

<210> 471

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 471

tttcaatggg acaccagc

19

<210> 472

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 472

gtttattcag ccgggagtc

19

<210> 473

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide  
 <400> 473  
 gactcccggc tgaataaac 19  
 <210> 474  
 <211> 19  
 <212> DNA  
 <213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
 oligonucleotide  
 <400> 474  
 tggagagttt gcaaggagc 19  
 <210> 475  
 <211> 19  
 <212> DNA  
 <213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
 oligonucleotide  
 <400> 475  
 gctccttgca aactctcca 19  
 <210> 476  
 <211> 19  
 <212> DNA  
 <213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
 oligonucleotide  
 <400> 476  
 ccctccagac ttccacct 19  
 <210> 477  
 <211> 19  
 <212> DNA  
 <213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
 oligonucleotide

<400> 477  
aggtggaaag tctggaggg 19

<210> 478  
<211> 19  
<212> DNA  
<213> Artificial Sequence  
<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 478  
aattttcttc caagttctc 19

<210> 479  
<211> 19  
<212> DNA  
<213> Artificial Sequence  
<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 479  
gagaacttgg aagaaaatt 19

<210> 480  
<211> 19  
<212> DNA  
<213> Artificial Sequence  
<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 480  
ctgcggagag aagaaaggg 19

<210> 481  
<211> 19  
<212> DNA  
<213> Artificial Sequence  
<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 481

ccctttcttc tctccgcag	19
<210> 482	
<211> 19	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic	
oligonucleotide	
<400> 482	
agagccaccc cctggctcc	19
<210> 483	
<211> 19	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic	
oligonucleotide	
<400> 483	
ggagccaggg ggtggctct	19
<210> 484	
<211> 19	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic	
oligonucleotide	
<400> 484	
cgaagccaac cgcggcggg	19
<210> 485	
<211> 19	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic	
oligonucleotide	
<400> 485	
cccgccgcgg ttggcttcg	19



<210> 486  
 <211> 19  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
           oligonucleotide  
 <400> 486  
 agagggaaga cgatcgccc 19  
 <210> 487  
 <211> 19  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
           oligonucleotide  
 <400> 487  
 gggcgatcgt cttccctct 19  
 <210> 488  
 <211> 19  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
           oligonucleotide  
 <400> 488  
 cccctttaac tttcctcgg 19  
 <210> 489  
 <211> 19  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
           oligonucleotide  
 <400> 489  
 cggaggaaag ttaaagggg 19  
 <210> 490  
 <211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 490

gcagccccgc ttccttcaa

19

<210> 491

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 491

ttgaaggaag cggggctgc

19

<210> 492

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 492

caccgcagcc gctcgctgca gcag

24

<210> 493

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 493

aaacctgctg cagcgagcgg ctgc

24

<210> 494

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 494

caccggctgg gtgtccatt gaaa 24

<210> 495

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 495

aaactttcaa tgggacaccc agcc 24

<210> 496

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 496

caccggttta ttcagccggg agtc 24

<210> 497

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 497

aaacgactcc cggctgaata aacc 24

<210> 498

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 498

caccgtggag agtttgcaag gagc 24

<210> 499

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 499

aaacgctcct tgcaaaactct ccac 24

<210> 500

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 500

caccgccttc cagactttcc acct 24

<210> 501

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 501

aaacaggtgg aaagtctgga gggc 24

<210> 502

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 502

caccgaattt tcttccaagt tctc 24

<210> 503

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 503

aaacgagaac ttggaagaaa attc 24

<210> 504

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 504

caccgctgcg gagagaagaa aggg 24

<210> 505

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 505

aaaccctttt cttctctccg cagc 24

<210> 506

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 506

caccgagagc cacccttggtg ctcc	24
<210> 507	
<211> 24	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic oligonucleotide	
<400> 507	
aaacggagcc aggggtggc tctc	24
<210> 508	
<211> 24	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic oligonucleotide	
<400> 508	
caccgcgaag ccaaccgcgg cggg	24
<210> 509	
<211> 24	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic oligonucleotide	
<400> 509	
aaaccccgcc gcggttggt tcgc	24
<210> 510	
<211> 24	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic oligonucleotide	
<400> 510	
caccgagagg gaagacgatc gccc	24

<210> 511

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 511

aaacgggcga tcgtcttccc tctc 24

<210> 512

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 512

caccgcccct ttaactttcc tccg 24

<210> 513

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 513

aaaccggagg aaagttaaag gggc 24

<210> 514

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 514

caccggcagc cccgcttctc tcaa 24

<210> 515

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 515

aaacttgaag gaagcggggc tgcc 24

<210> 516

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 516

cgagagcgag aggaggag 19

<210> 517

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 517

ctccctctc tcgtctcg 19

<210> 518

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 518

gagagagctt gagagcgcg 19

<210> 519

<211> 19



<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 519

cgcgctctca agctctctc 19

<210> 520

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 520

ggtggagggg gcggggccc 19

<210> 521

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 521

ggggcccgcc ccctccacc 19

<210> 522

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 522

ggtatccacg taaatcaaa 19

<210> 523

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 523

tttgatttac gtggatacc

19

<210> 524

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 524

ccaatcactg gctccggtc

19

<210> 525

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 525

gaccggagcc agtgattgg

19

<210> 526

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 526

ggcgcccag ggaagaaga

19

<210> 527

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 527

tctttcttccc tcgggcgcc

19

<210> 528

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 528

gggtgggggt accagagga

19

<210> 529

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 529

tcctctggta cccccaccc

19

<210> 530

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 530

ccggggacag aagagaggg

19

<210> 531

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 531

ccctctcttc tgtccccg 19

<210> 532

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 532

gagagagagt gggagaagc 19

<210> 533

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 533

gcttctccca ctctctctc 19

<210> 534

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 534

aaagtaactg tcaaatgcg 19

<210> 535

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 535  
cgcatTTgac agttacttt 19  
<210> 536  
<211> 19  
<212> DNA  
<213> Artificial Sequence  
<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide  
<400> 536  
ttaaccagag cgcccagtc 19  
<210> 537  
<211> 19  
<212> DNA  
<213> Artificial Sequence  
<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide  
<400> 537  
gactggggcgc tctggttaa 19  
<210> 538  
<211> 19  
<212> DNA  
<213> Artificial Sequence  
<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide  
<400> 538  
cgtcggagct gcccgctag 19  
<210> 539  
<211> 19  
<212> DNA  
<213> Artificial Sequence  
<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide  
<400> 539

ctagcgggca gctccgacg	19
<210> 540	
<211> 24	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic	
oligonucleotide	
<400> 540	
caccgcgaga gcgagaggag ggag	24
<210> 541	
<211> 24	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic	
oligonucleotide	
<400> 541	
aaacctccct cctctcgctc tcgc	24
<210> 542	
<211> 24	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic	
oligonucleotide	
<400> 542	
caccggagag agcttgagag cgcg	24
<210> 543	
<211> 24	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic	
oligonucleotide	
<400> 543	
aaaccgcgct ctcaagctct ctcc	24
<210> 544	

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 544

caccgggtgg agggggcggg gcc

24

<210> 545

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 545

aaacgggccc cgccccctcc accc

24

<210> 546

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 546

caccgggtat ccacgtaaat caaa

24

<210> 547

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 547

aaactttgat ttacgtggat accc

24

<210> 548

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 548

caccgccaat cactggtcc ggtc 24

<210> 549

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 549

aaacgaccgg agccagtgat tggc 24

<210> 550

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 550

caccgggcgc ccgagggaag aaga 24

<210> 551

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 551

aaactcttct tccctcgggc gccc 24

<210> 552

<211> 24

<212> DNA



<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 552

caccggggtg ggggtaccag agga 24

<210> 553

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 553

aaactcctct ggtaccccca cccc 24

<210> 554

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 554

caccgccggg gacagaagag aggg 24

<210> 555

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 555

aaaccctct cttctgtccc cggc 24

<210> 556

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 556

caccggagag agagtgggag aagc

24

<210> 557

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 557

aaacgttct cccactctct ctcc

24

<210> 558

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 558

caccgaaagt aactgtcaaa tgcg

24

<210> 559

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 559

aaaccgcatt tgacagttac ttcc

24

<210> 560

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 560

caccgttaac cagagcgccc agtc 24

<210> 561

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 561

aaacgactgg gcgctctggt taac 24

<210> 562

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 562

caccgcgtcg gagctgcccg ctag 24

<210> 563

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 563

aaacctagcg ggcagctccg acgc 24

<210> 564

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 564

gagacacaca gaaatgtaac	20
<210> 565	
<211> 20	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic oligonucleotide	
<400> 565	
ggtggggcac tgaccccgac	20
<210> 566	
<211> 20	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic oligonucleotide	
<400> 566	
ctagagtgat gactcctatc	20
<210> 567	
<211> 20	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic oligonucleotide	
<400> 567	
gactaaaact ccacctcaaa	20
<210> 568	
<211> 20	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic oligonucleotide	
<400> 568	
aatatgtcac attctgtctc	20

<210> 569  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
     oligonucleotide  
 <400> 569  
 ggactatggg aggtcactaa 20

<210> 570  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
     oligonucleotide  
 <400> 570  
 gctcatgctt ggactatggg 20

<210> 571  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
     oligonucleotide  
 <400> 571  
 gttctggcca ggcccctgtc 20

<210> 572  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
     oligonucleotide  
 <400> 572  
 agtgccccac cccgccttc 20

<210> 573

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 573

gtggggcact gaccccgaca 20

<210> 574

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 574

aaccttctaa gcaaaccttc 20

<210> 575

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 575

gttacacaga accagaaggc 20

<210> 576

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 576

gaaggttaca cagaaccaga 20

<210> 577

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 577

agtcgatgatg agtcgatgctg 20

<210> 578

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 578

gatgagtcac gctgaggctt 20

<210> 579

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 579

actctaggct gagaacatct 20

<210> 580

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 580

gtccccagca ggatgcttac 20

<210> 581

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 581

gccctgtaag catcctgctg

20

<210> 582

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 582

cagggcagat ggcaaaaaaa

20

<210> 583

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 583

gaggtggagt tttagtcagg

20

<210> 584

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 584

aaacggcatc ataaagaaaa

20

<210> 585

<211> 105

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic



polynucleotide

<400> 585

gagacacaca gaaatgtaac gttttagagc tagaaatagc aagttaaaat aaggctagtc 60

cgttatcaac ttgaaaaagt ggcaccgagt cggtgctttt tttcc 105

<210> 586

<211> 105

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 586

ggtggggcac tgaccccgac gttttagagc tagaaatagc aagttaaaat aaggctagtc 60

cgttatcaac ttgaaaaagt ggcaccgagt cggtgctttt tttcc 105

<210> 587

<211> 105

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 587

ctagagtgat gactcctatc gttttagagc tagaaatagc aagttaaaat aaggctagtc 60

cgttatcaac ttgaaaaagt ggcaccgagt cggtgctttt tttcc 105

<210> 588

<211> 105

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 588

gactaaaact ccacctcaaa gttttagagc tagaaatagc aagttaaaat aaggctagtc 60

cgttatcaac ttgaaaaagt ggcaccgagt cggtgctttt tttcc 105

<210> 589

<211> 105

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 589

aatatgtcac attctgtctc gttttagagc tagaaatagc aagttaaaat aaggctagtc 60

cgttatcaac ttgaaaaagt ggcaccgagt cggtgctttt tttcc 105

<210> 590

<211> 105

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 590

ggactatggg aggtcactaa gttttagagc tagaaatagc aagttaaaat aaggctagtc 60

cgttatcaac ttgaaaaagt ggcaccgagt cggtgctttt tttcc 105

<210> 591

<211> 105

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 591

gctcatgctt ggactatggg gttttagagc tagaaatagc aagttaaaat aaggctagtc 60

cgttatcaac ttgaaaaagt ggcaccgagt cggtgctttt tttcc 105

<210> 592

<211> 105

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 592

gttctggcca ggcccctgtc gttttagagc tagaaatagc aagttaaaat aaggctagtc 60

cgttatcaac ttgaaaaagt ggcaccgagt cggcgctttt tttcc 105

<210> 593

<211> 105

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
polynucleotide

<400> 593

agtgccccac cccgccttc gtttagagc tagaaatagc aagttaaat aaggctagtc 60

cgttatcaac ttgaaaaagt ggcaccgagt cggcgctttt tttcc 105

<210> 594

<211> 105

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
polynucleotide

<400> 594

gtggggcact gaccccgaca gtttagagc tagaaatagc aagttaaat aaggctagtc 60

cgttatcaac ttgaaaaagt ggcaccgagt cggcgctttt tttcc 105

<210> 595

<211> 105

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
polynucleotide

<400> 595

aaccttctaa gcaaaccttc gtttagagc tagaaatagc aagttaaat aaggctagtc 60

cgttatcaac ttgaaaaagt ggcaccgagt cggcgctttt tttcc 105

<210> 596

<211> 105

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 596

gttacacaga accagaaggc gttttagagc tagaaatagc aagttaaaat aaggctagtc 60

cgttatcaac ttgaaaaagt ggcaccgagt cggtgctttt tttcc 105

<210> 597

<211> 105

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 597

gaaggttaca cagaaccaga gttttagagc tagaaatagc aagttaaaat aaggctagtc 60

cgttatcaac ttgaaaaagt ggcaccgagt cggtgctttt tttcc 105

<210> 598

<211> 105

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 598

agtcgatg agtcgatgctg gttttagagc tagaaatagc aagttaaaat aaggctagtc 60

cgttatcaac ttgaaaaagt ggcaccgagt cggtgctttt tttcc 105

<210> 599

<211> 105

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 599

gatgagtc gctgaggctt gttttagagc tagaaatagc aagttaaaat aaggctagtc 60

cgttatcaac ttgaaaaagt ggcaccgagt cggtgctttt tttcc 105

<210> 600

<211> 105  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
 polynucleotide  
 <400> 600  
 actctaggct gagaacatct gttttagagc tagaaatagc aagttaaaat aaggctagtc 60  
 cgttatcaac ttgaaaaagt ggcaccgagt cggtgctttt tttcc 105  
 <210> 601  
 <211> 105  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
 polynucleotide  
 <400> 601  
 gtccccagca ggatgcttac gttttagagc tagaaatagc aagttaaaat aaggctagtc 60  
 cgttatcaac ttgaaaaagt ggcaccgagt cggtgctttt tttcc 105  
 <210> 602  
 <211> 105  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
 polynucleotide  
 <400> 602  
 gccctgtaag catcctgctg gttttagagc tagaaatagc aagttaaaat aaggctagtc 60  
 cgttatcaac ttgaaaaagt ggcaccgagt cggtgctttt tttcc 105  
 <210> 603  
 <211> 105  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Description of Artificial Sequence: Synthetic  
 polynucleotide  
 <400> 603

cagggcagat ggcaaaaaa gttttagagc tagaaatagc aagttaaaat aaggctagtc 60  
cgttatcaac ttgaaaaagt ggcaccgagt cggtgctttt tttcc 105  
<210> 604  
<211> 105  
<212> DNA  
<213> Artificial Sequence  
<220><223> Description of Artificial Sequence: Synthetic  
polynucleotide  
<400> 604  
gaggtggagt ttagtcagg gttttagagc tagaaatagc aagttaaaat aaggctagtc 60  
cgttatcaac ttgaaaaagt ggcaccgagt cggtgctttt tttcc 105  
<210> 605  
<211> 105  
<212> DNA  
<213> Artificial Sequence  
<220><223> Description of Artificial Sequence: Synthetic  
polynucleotide  
<400> 605  
aaacggcatc ataaagaaaa gttttagagc tagaaatagc aagttaaaat aaggctagtc 60  
cgttatcaac ttgaaaaagt ggcaccgagt cggtgctttt tttcc 105  
<210> 606  
<211> 19  
<212> DNA  
<213> Artificial Sequence  
<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide  
<400> 606  
tgtactctct gaggtgctc 19  
<210>  
> 607  
<211> 19  
<212> DNA  
<213> Artificial Sequence  
<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 607

acgcagataa gaaccagtt 19

<210> 608

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 608

catcaagtca gccatcagc 19

<210> 609

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 609

gagtcaccct cctggaaac 19

<210> 610

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 610

cctgggctcc ggggcgttt 19

<210> 611

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 611	
ggcccctgcg gccaccccg	19
<210> 612	
<211> 19	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic oligonucleotide	
<400> 612	
ctccctccct gcccggtag	19
<210> 613	
<211> 19	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic oligonucleotide	
<400> 613	
aggtttggaa agggcgtgc	19
<210> 614	
<211> 20	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic oligonucleotide	
<400> 614	
actccactgc actccagtct	20
<210> 615	
<211> 20	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic oligonucleotide	
<400> 615	
tctgtggggg acctgcactg	20



<210> 616

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 616

ggggcgccag ttgtgtctcc 20

<210> 617

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
oligonucleotide

<400> 617

acaccattgc caccaccatt 20

<210> 618

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 618

caatgacccc ttcattgacc 20

<210> 619

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 619

ttgattttgg agggatctcg 20

<210> 620

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 620

ggaatccatg gaggaagat 20

<210> 621

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 621

tggttctcgct caggtcagtg 20

<210> 622

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 622

ctctctgctc ctttgcaca 20

<210> 623

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic  
primer

<400> 623

gtgctcttcg gggttcagga 20

<210> 624

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 624

cgaaagagaa agcgaaccag tatcgagaac

30

<210> 625

<211> 27

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 625

cgttgtgcat agtcgctgct tgatcgc

27