



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2022-0040983
(43) 공개일자 2022년03월31일

(51) 국제특허분류(Int. Cl.)
C12N 15/10 (2017.01) C12N 15/113 (2010.01)
C12N 15/63 (2006.01) C12N 9/12 (2006.01)
C12N 9/22 (2006.01)
(52) CPC특허분류
C12N 15/102 (2013.01)
C12N 15/113 (2013.01)
(21) 출원번호 10-2021-0106637
(22) 출원일자 2021년08월12일
심사청구일자 2021년08월12일
(30) 우선권주장
1020200123432 2020년09월24일 대한민국(KR)

(71) 출원인
고려대학교 산학협력단
서울특별시 성북구 안암로 145, 고려대학교 (안암동5가)
서울대학교산학협력단
서울특별시 관악구 관악로 1 (신림동)
재단법인 국가마우스표현형분석사업단
서울특별시 관악구 관악로 1 (신림동, 서울대학교수의과대학생명공학연구동5층)
(72) 발명자
김경미
서울특별시 종로구 통일로18길 9, 109동 701호 (무악동, 인왕산아이파크)
박수지
경기도 평택시 동삭로 397, 106동 1704호 (동삭동, 삼익사이버아파트)
(74) 대리인
특허법인 무한

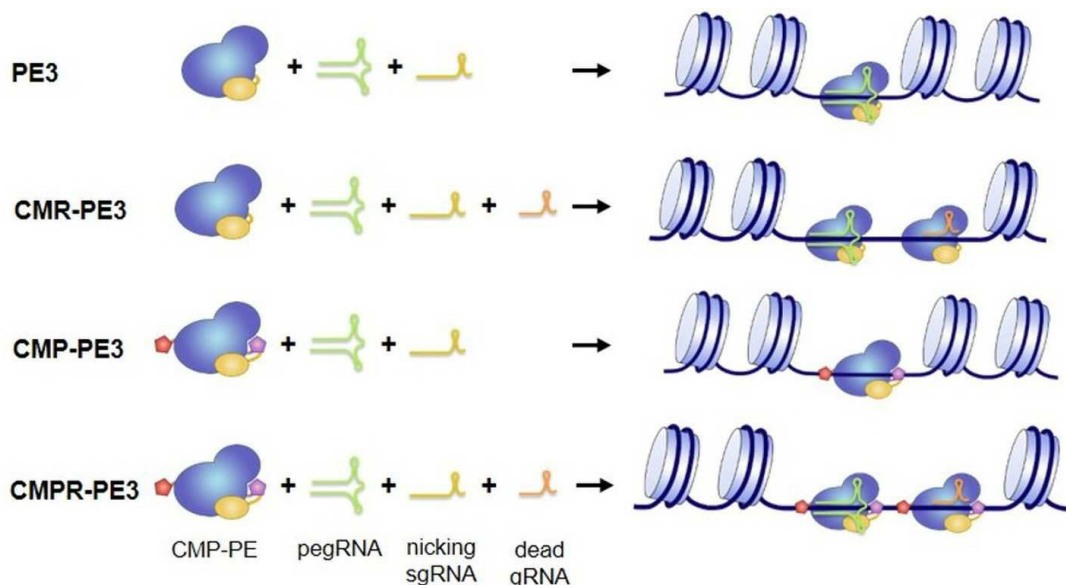
전체 청구항 수 : 총 20 항

(54) 발명의 명칭 편집 효율이 향상된 프라임 편집 기반 유전자 교정용 조성물 및 이의 용도

(57) 요약

본 발명은 유전자 편집 효율이 향상된 프라임 편집 기반 유전자 교정용 조성물, 상기 조성물을 이용한 유전자 교정 방법, 유전자 교정용 키트 및 유전자 변형 포유동물의 제조방법에 관한 것이다. 본 발명에서 개발한 프라임 편집기는 현저히 증진된 게놈 편집 효율 및 표적 특이성을 나타내며 이를 이용해 제작한 돌연변이 동물모델에서 다음 세대로의 변이 전달 및 표현형 변화가 나타나는 것을 확인하였는바, 상기 개선된 프라임 편집기 또는 이를 포함하는 유전자 교정용 조성물은 인간화 동물모델의 제작 및 연구, 유전공학 기술 분야 및 유전질환의 치료 수단 등의 다양한 용도로 유용하게 활용될 수 있을 것이다.

대표도 - 도11



(52) CPC특허분류

C12N 15/63 (2013.01)
C12N 9/1276 (2013.01)
C12N 9/22 (2013.01)
C12Y 207/07049 (2013.01)
C12N 2310/20 (2017.05)

성제경

서울특별시 양천구 목동동로 430, 602-702 (목동,
 목동신시가지아파트6단지)

(72) 발명자

정태영

경기도 하남시 미사강변대로 135, 1909동 1101호
 (망월동, 미사강변 하우스디 더 레이크)

신승균

서울특별시 성북구 보문로26길 83-11, 302호 (안암
 동1가)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711117313
과제번호	2020M3A9D5A01082439
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	(원천)국가마우스표현형분석사업
연구과제명	상동제조합 효율을 높인 새로운 유전자편집 기술 개발 및 이를 활용한 인간화 마우스 제작
기 여 율	40/100
과제수행기관명	고려대학교
연구기간	2020.04.13 ~ 2021.12.31

이 발명을 지원한 국가연구개발사업

과제고유번호	1711108999
과제번호	2019R1A2C2087198
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	(이공)(우형1-2)중견연구(연평균연구비 1억원~2억원 이내)
연구과제명	크리스퍼시스템 기반 전사체 활성인자를 이용한 세포전환기술 개발
기 여 율	30/100
과제수행기관명	고려대학교
연구기간	2020.03.01 ~ 2021.02.28

이 발명을 지원한 국가연구개발사업

과제고유번호	1711116112
과제번호	2013M3A9D5072550
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	바이오. 의료기술개발(R&D)
연구과제명	마우스 표현형서비스 구축 및 대사 운동 표현형 분석
기 여 율	30/100
과제수행기관명	재단법인 국가마우스표현형분석사업단
연구기간	2020.03.20 ~ 2020.12.31

명세서

청구범위

청구항 1

- (a) i) CRISPR/Cas9 단백질 또는 이의 변이체, 및
- ii) 역전사효소 (Reverse Transcriptase) 또는 이의 변이체를 포함하는, 융합단백질 또는 이를 암호화하는 핵산; 및
- (b) 가이드 RNA 또는 이를 암호화하는 핵산을 포함하고,
- 이때, 상기 가이드 RNA는 pegRNA (prime editing guide RNA) 및 데드 단일 가닥 RNA (dead single guide RNA, dsgrNA)를 포함하며,
- 상기 dsgrNA는 10~20 뉴클레오티드 (nt) 길이인 것을 특징으로 하는, 유전자 교정용 조성물.

청구항 2

- 제1항에 있어서,
- 상기 dsgrNA는 pegRNA 결합 부위로부터 5 내지 70nt 떨어진 위치에 결합하여 상기 융합단백질의 염색질 접근성을 증가시키는 것을 특징으로 하는, 유전자 교정용 조성물.

청구항 3

- 제1항에 있어서,
- 상기 유전자 교정용 조성물은 비표적 DNA 가닥에 상보적으로 결합하여 표적 DNA 가닥의 절단을 유도하는 단일 가이드 RNA (single guide RNA, sgRNA)를 추가로 포함하는 것을 특징으로 하는, 유전자 교정용 조성물.

청구항 4

- (a) i) CRISPR/Cas9 단백질 또는 이의 변이체,
- ii) 역전사효소 (Reverse Transcriptase) 또는 이의 변이체, 및
- iii) 염색질 조절 펩타이드 (Chromatin-modulating peptides)를 포함하는, 융합단백질 또는 이를 암호화하는 핵산; 및
- (b) 가이드 RNA 또는 이를 암호화하는 핵산을 포함하고,
- 상기 가이드 RNA는 pegRNA (prime editing guide RNA) 및 데드 단일 가닥 RNA (dead single guide RNA, dsgrNA)인 것을 특징으로 하는, 유전자 교정용 조성물.

청구항 5

- 제4항에 있어서,
- 상기 염색질 조절 펩타이드는 고-이동성 그룹 뉴클레오솜 결합 도메인 1 (high-mobility group nucleosome binding domain 1, HN1), 히스톤 H1 중심 구형 도메인 (histone H1 central globular domain, H1G), 또는 이의 조합인 것을 특징으로 하는, 유전자 교정용 조성물.

청구항 6

제4항에 있어서,

상기 염색질 조절 펩타이드는 화학적 결합에 의해 직접적으로, 링커에 의해 간접적으로, 또는 이의 조합으로 상기 CRISPR/Cas9 단백질 또는 역전사효소에 연결된 것을 특징으로 하는, 유전자 교정용 조성물.

청구항 7

제4항 또는 제5항에 있어서,

상기 융합단백질은

N말단-[HN1]-[Cas9]-[H1G]-[역전사효소]-C말단; 또는

N말단-[HN1]-[Cas9]-[역전사효소]-[H1G]-C말단의 구성으로 이루어진 것을 특징으로 하는, 유전자 교정용 조성물.

청구항 8

제4항 또는 제5항에 있어서,

상기 융합단백질은 N-말단 및 C-말단에 각각 핵 위치화 신호 (nuclear localization signal, NLS) 서열을 추가로 포함하는 것을 특징으로 하는, 유전자 교정용 조성물.

청구항 9

제4항에 있어서,

상기 CRISPR/Cas9 단백질 변이체는 니카아제 (nickase)인 것을 특징으로 하는, 유전자 교정용 조성물.

청구항 10

제9항에 있어서,

상기 CRISPR/Cas9 단백질 변이체는 RuvC 도메인 또는 HNH 도메인 중 어느 하나가 불활성화된 것을 특징으로 하는, 유전자 교정용 조성물.

청구항 11

제4항에 있어서,

상기 역전사효소 (Reverse Transcriptase) 또는 이의 변이체는 몰로니 마우스 백혈병 바이러스 (Moloney murine leukemia virus, M-MLV) 유래인 것을 특징으로 하는, 유전자 교정용 조성물.

청구항 12

제4항에 있어서,

상기 융합단백질은 서열번호 1 또는 서열번호 2로 표시되는 아미노산 서열로 이루어진 것을 특징으로 하는, 유전자 교정용 조성물.

청구항 13

제4항에 있어서,

상기 dsgrRNA는 10~20 뉴클레오티드 (nt) 길이로 구성된 것을 특징으로 하는, 유전자 교정용 조성물.

청구항 14

제4항에 있어서,

상기 dsgrRNA는 pegRNA 결합 부위로부터 5 내지 70 nt 떨어진 위치에 결합하여 상기 융합단백질의 염색질 접근성을 향상시키는 것을 특징으로 하는, 유전자 교정용 조성물.

청구항 15

제4항에 있어서,

상기 유전자 교정용 조성물은 비표적 DNA 가닥에 상보적으로 결합하여 표적 DNA 가닥의 절단을 유도하는 단일 가이드 RNA (single guide RNA, sgRNA)를 추가로 포함하는 것을 특징으로 하는, 유전자 교정용 조성물.

청구항 16

제1항 또는 제4항에 있어서,

상기 조성물은 유전자 교정 효율 및 표적 특이성이 증진된 것을 특징으로 하는, 유전자 교정용 조성물.

청구항 17

제1항 또는 제4항의 유전자 교정용 조성물을 *in vitro* 또는 *ex vivo* 상에서 표적 핵산 서열을 포함하는 표적 영역 (region)과 접촉시키는 단계를 포함하는, 유전자 교정 방법.

청구항 18

제1항 또는 제4항의 유전자 교정용 조성물을 포함하는, 유전자 교정용 키트.

청구항 19

제1항 또는 제4항의 유전자 교정용 조성물을 인간을 제외한 포유동물 세포에 도입하여 유전자 변형 포유동물 세포를 얻는 단계; 및

상기 얻어진 유전자 변형 포유동물 세포를 인간을 제외한 포유동물 위탁모 난관에 이식하는 단계를 포함하는, 인간을 제외한 유전자 변형 포유동물의 제조방법.

청구항 20

제19항에 있어서,

상기 포유동물 세포는 포유동물의 배아 세포인 것을 특징으로 하는, 제조방법.

발명의 설명

기술 분야

[0001] 본 발명은 유전자 편집 효율이 향상된 프라임 편집 기반 유전자 교정용 조성물, 상기 조성물을 이용한 유전자 교정 방법, 유전자 교정용 키트 및 유전자 변형 포유동물의 제조방법에 관한 것이다.

배경 기술

[0003] CRISPR-Cas 시스템은 원하는 표적 돌연변이를 효율적으로 유도할 수 있는 뉴클레아제 (nucleases), 염기 편집기 (base editors), 유전자전위효소 (transposases)와 같은 다양한 진보된 게놈 편집 도구로 진화되어왔다. 특히 크리스퍼 (CRISPR) 시스템을 기반으로 개발된 사이토신 염기 교정 유전자가위 (cytosine base editor, CBE)와 아데닌 염기 교정 유전자가위 (adenine base editor, ABE)는 마우스를 포함하는 다양한 유기체에서 C·G를 T·A로, A·T를 G·C로 효율적으로 치환할 수 있다. 또한 최근 연구를 통해 인간 세포에서 C에서 G로 염기 교정이 가능한 C-to-G 염기 편집기 (C-to-G base editors)로써 CGBE1이 보고되었다. 그러나 하나 이상 염기의 삽입, 치환 또는 절단과 같은 정확한 표적 돌연변이의 생성은 세포 내 상동직접수선 (homology-directed repair, HDR)의 낮은 효율성으로 인한 유전자 편집의 제한으로 인해 여전히 어려운 실정이다.

[0004] 이러한 요구에 따라 개발된 새로운 개념의 게놈 편집 도구인 프라임 편집기 (Prime editor, PE)는 DNA 이중가닥 중 한 가닥만 절단하도록 변형된 Cas9 니카아제 (Cas9 nickase-H840A) 및 역전사효소 (Reverse transcriptase, RT)로 구성된 융합단백질이다. 프라임 편집기는 원하는 편집 서열을 암호화하는 새로운 형태의 가이드 RNA인 pegRNA (prime editing guide RNA)를 필요로 한다. 상기 pegRNA는 표적 유전자의 비표적 가닥에 상보적인 염기 서열 (primer binding site, PBS) 및 교정하고자 하는 염기서열을 포함하는 역전사효소의 주형 가닥 부위 (RT template)를 가진다. 프라임 편집기의 Cas9 니카아제는 표적 유전자의 비표적 가닥을 절단하고 역전사효소가 pegRNA의 RT 주형 가닥을 기반으로 교정된 염기서열이 포함된 새로운 DNA 가닥을 합성하고, 이후 세포 내 복구 과정에 의해 기존의 DNA 염기서열이 제거되고 새로 합성된 교정된 형태의 DNA 염기서열이 이를 대신한다. 이러한 정교한 게놈 편집 시스템을 통해 이중 가닥 DNA의 절단이나 공여 DNA 없이 염기에서 염기로의 치환, 작은 크기의 삽입 및 결실을 포함한 표적 돌연변이 생성이 가능하다. 그러나 프라임 편집기 역시 염기 편집기보다 그 빈도는 낮으나 오프-타겟 (off-target)의 문제가 발생하고, 특히 역전사효소에 의해 추가적인 염기서열의 삽입 및 결실이 발생하여 정확한 편집이 어려운 문제가 있다. 더욱이, 상기 문제점과 함께 현재 개발된 염기 편집기 및 프라임 편집기는 매우 낮은 효율의 한계점이 있다. 예컨대, 프라임 편집기 중에서 가장 효율이 좋은 PE3는 유전자 편집 효율이 20~50%에 불과하다.

[0005] 인간을 포함한 생체체의 세포 내 유전 물질은 항상 돌연변이에 노출되어 있으며, 돌연변이의 축적은 다양한 변이 표현형을 유도하고 그 일부는 질환과 관련이 있다. 따라서 프라임 편집을 이용한 정교한 유전자 교정 기술은 유전질환의 치료를 위한 도구로 이용될 수 있으며, 병원성 변이 인자를 확인하고 질병의 기전 연구 등에 폭넓게 이용될 수 있다. 따라서 공지된 프라임 편집기보다 더욱 정교하고 편집 효율이 개선된 프라임 편집기 및 이를 이용한 유전자 교정 기술의 개발이 필요하다.

선행기술문헌

비특허문헌

[0007] (비특허문헌 0001) Int J Mol Sci. 2020 Aug 28;21(17):6240.

발명의 내용

해결하려는 과제

- [0008] 프라임 편집 기반 유전자 교정의 상기 문제점들을 극복하기 위하여, 본 발명자들은 proximal dead sgRNA (dsgRNA) 및/또는 염색질-조절 펩타이드 (chromatin-modulating peptides, CMPs)를 사용하여 개선된 프라임 편집기를 개발하고 이의 현저히 증진된 게놈 편집 효율 및 표적 특이성 등을 구체적인 실험을 통해 확인함으로써 본 발명을 완성하게 되었다.
- [0009] 이에, 본 발명은 유전자 편집 효율이 개선된 프라임 편집 기반 유전자 교정용 조성물을 제공하는 것을 목적으로 한다.
- [0010] 또한, 본 발명은 상기 유전자 교정용 조성물을 이용한 유전자 교정 방법을 제공하는 것을 다른 목적으로 한다.
- [0011] 또한, 본 발명은 상기 유전자 교정용 조성물을 포함하는 유전자 교정용 키트를 제공하는 것을 또 다른 목적으로 한다.
- [0012] 또한, 본 발명은 상기 유전자 교정용 조성물을 이용하여 인간을 제외한 유전자 변형 포유동물을 제조하는 방법을 제공하는 것을 또 다른 목적으로 한다.
- [0014] 그러나 본 발명이 이루고자 하는 기술적 과제는 이상에서 언급한 과제에 제한되지 않으며, 언급되지 않은 또 다른 과제들은 아래의 기재로부터 당업자에게 명확하게 이해될 수 있을 것이다.

과제의 해결 수단

- [0016] 상기와 같은 본 발명의 목적을 달성하기 위하여, 본 발명은 (a) i) CRISPR/Cas9 단백질 또는 이의 변이체, 및 ii) 역전사효소 (Reverse Transcriptase) 또는 이의 변이체를 포함하는, 융합단백질 또는 이를 암호화하는 핵산; 및
- [0017] (b) 가이드 RNA 또는 이를 암호화하는 핵산을 포함하고,
- [0018] 이때, 상기 가이드 RNA는 pegRNA (prime editing guide RNA) 및 데드 단일 가닥 RNA (dead single guide RNA, dsgRNA)를 포함하며, 상기 dsgRNA는 10~20bp 길이인 것을 특징으로 하는, 유전자 교정용 조성물을 제공한다.
- [0019] 본 발명의 일구현예로, 상기 dsgRNA는 pegRNA 결합 부위로부터 5 내지 70개 뉴클레오티드가 떨어진 위치에 결합하여 상기 융합단백질의 염색질 접근성을 증가시킬 수 있다.
- [0020] 본 발명의 다른 구현예로, 상기 유전자 교정용 조성물은 비표적 DNA 가닥에 상보적으로 결합하여 표적 DNA 가닥의 절단을 유도하는 단일 가이드 RNA (single guide RNA, sgRNA)를 추가로 포함할 수 있다.
- [0021] 또한, 본 발명은 (a) i) CRISPR/Cas9 단백질 또는 이의 변이체, ii) 역전사효소 (Reverse Transcriptase) 또는 이의 변이체, 및 iii) 염색질 조절 펩타이드 (Chromatin-modulating peptides)를 포함하는, 융합단백질 또는 이를 암호화하는 핵산; 및
- [0022] (b) 가이드 RNA 또는 이를 암호화하는 핵산을 포함하고,
- [0023] 상기 가이드 RNA는 pegRNA (prime editing guide RNA) 및 데드 단일 가닥 RNA (dead single guide RNA, dsgRNA)인 것을 특징으로 하는, 유전자 교정용 조성물을 제공한다.
- [0024] 본 발명의 일 구현예로, 상기 염색질 조절 펩타이드는 고-이동성 그룹 뉴클레오솜 결합 도메인 1 (high-mobility group nucleosome binding domain 1, HN1), 히스톤 H1 중심 구형 도메인 (histone H1 central globular domain, H1G), 또는 이의 조합일 수 있다.
- [0025] 본 발명의 다른 구현예로, 상기 염색질 조절 펩타이드는 화학적 결합에 의해 직접적으로, 링커에 의해 간접적으로, 또는 이의 조합으로 상기 CRISPR/Cas9 단백질 또는 역전사효소에 연결될 수 있다.
- [0026] 본 발명의 또 다른 구현예로, 상기 융합단백질은
- [0027] N말단-[HN1]-[Cas9]-[H1G]-[역전사효소]-C말단; 또는 N말단-[HN1]-[Cas9]-[역전사효소]-[H1G]-C말단의 구성으로 이루어진 것일 수 있다.

- [0028] 본 발명의 또 다른 구현예로, 상기 융합단백질은 N-말단 및 C-말단에 각각 핵 위치화 신호 (nuclear localization signal, NLS) 서열을 추가로 포함하는 것일 수 있다.
- [0029] 본 발명의 또 다른 구현예로, 상기 CRISPR/Cas9 단백질 변이체는 니카아제 (nickase)일 수 있다.
- [0030] 본 발명의 또 다른 구현예로, 상기 CRISPR/Cas9 단백질 변이체는 RuvC 도메인 또는 HNH 도메인 중 어느 하나가 불활성화된 것일 수 있다.
- [0031] 본 발명의 또 다른 구현예로, 상기 역전사효소 (Reverse Transcriptase) 또는 이의 변이체는 몰로니 마우스 백혈병 바이러스 (Moloney murine leukemia virus, M-MLV) 유래인 것일 수 있다.
- [0032] 본 발명의 또 다른 구현예로, 상기 융합단백질은 서열번호 1 또는 서열번호 2로 표시되는 아미노산 서열로 이루어진 것일 수 있다.
- [0033] 본 발명의 또 다른 구현예로, 상기 dsgrNA는 10 내지 20개 뉴클레오티드 길이로 이루어진 것일 수 있다.
- [0034] 본 발명의 또 다른 구현예로, 상기 dsgrNA는 pegRNA 결합 부위로부터 5 내지 70개 뉴클레오티드가 떨어진 위치에 결합하여 상기 융합단백질의 염색질 접근성을 향상시킬 수 있다.
- [0035] 본 발명의 또 다른 구현예로, 상기 유전자 교정용 조성물은 유전자 교정 효율 및 표적 특이성이 증진된 것을 특징으로 한다.
- [0036] 또한, 본 발명은 (a) i) CRISPR/Cas9 단백질 또는 이의 변이체; ii) 역전사효소 (Reverse Transcriptase) 또는 이의 변이체; 및 iii) 염색질 조절 펩타이드 (Chromatin-modulating peptides)를 포함하는 단백질 또는 이를 암호화하는 핵산; 및
- [0037] (b) 가이드 RNA 또는 이를 암호화하는 핵산을 포함하고,
- [0038] 상기 가이드 RNA는 pegRNA (prime editing guide RNA) 및 데드 단일 가닥 RNA (dead single guide RNA, dsgrNA)인 것을 특징으로 하는, 유전자 교정용 조성물을 제공한다.
- [0039] 본 발명의 일 구현예로 상기 i) CRISPR/Cas9 단백질 또는 이의 변이체; 및 상기 ii) 역전사효소 (Reverse Transcriptase) 또는 이의 변이체는 융합되어 세포 내로 전달되거나, i) 및 ii)를 각각 별도로 발현시켜서 플라스미드 DNA, RNA, 또는 단백질 형태로 세포 내로 전달할 수 있다.
- [0041] 또한, 본 발명은 상기 유전자 교정용 조성물을 *in vitro* 또는 *ex vivo* 상에서 표적 핵산 서열을 포함하는 표적 영역 (region)과 접촉시키는 단계를 포함하는, 유전자 교정 방법을 제공한다.
- [0042] 또한, 본 발명은 상기 유전자 교정용 조성물을 포함하는, 유전자 교정용 키트를 제공한다.
- [0043] 또한, 본 발명은 상기 유전자 교정용 조성물을 인간을 제외한 포유동물 세포에 도입하여 유전자 변형 포유동물 세포를 얻는 단계; 및
- [0044] 상기 얻어진 유전자 변형 포유동물 세포를 인간을 제외한 포유동물 위탁모 난관에 이식하는 단계를 포함하는, 인간을 제외한 유전자 변형 포유동물의 제조방법을 제공한다.
- [0045] 본 발명의 일 구현예로, 상기 포유동물 세포는 포유동물의 배아 세포일 수 있다.

발명의 효과

- [0047] 본 발명에서는 dead sgRNA (dsgrNA) 및/또는 염색질-조절 펩타이드 (chromatin-modulating peptides, CMPs)를 사용하여 성능이 개선된 프라임 편집기를 개발하였고, 이의 현저히 증진된 게놈 편집 효율 및 표적 특이성을 확인하였으며 표적 유전자의 돌연변이 동물모델을 제작하여 다음 세대로의 돌연변이 전달 및 표현형 변화 등을 확인하였다. 따라서 본 발명에 따른 개선된 프라임 편집기를 포함하는 유전자 교정용 조성물은 인간화 동물모델의 제작 및 연구, 유전공학 기술 분야 및 유전질환의 치료 수단 등의 다양한 용도로 유용하게 활용될 수 있을 것이다.

도면의 간단한 설명

[0049]

도 1은 포유동물의 게놈에서 프라임 편집 시스템을 검증한 결과로서, 도 1a는 tdTomato를 발현하는 리포터 HEK293T 세포의 AAVS1 유전자에서 tdTomato 유전자에 대한 표적 돌연변이 설계를 도시한 개략도, 및 상기 세포에서 PE3의 편집 효율을 나타낸 현미경 이미지 및 생어 염기서열분석 크로마토그램 결과를 나타낸 것이고, 도 1b는 PE3를 형질감염시키거나 시키지 않은 세포를 11일간 배양한 후 FACS를 통해 tdTomato-음성 및 tdTomato-양성 세포를 계수한 결과이다.

도 2는 마우스 유전자 *Igf2* 및 *Adamts20*의 표적 부위에서 다양한 길이의 pegRNAs 및 dsgrNA를 이용하여 프라임 편집 효율을 최적화한 결과로서, 도 2a는 상기 2종의 유전자 표적에서 각각 표적 돌연변이 설계를 도시한 개략도이고, 도 2b는 마우스 유래 NIH/3T3 세포에서 프라이머 결합 부위 (PBS) 및 역전사효소 주형 길이의 다양한 조합에 따른 PE3 및 CMR-PE3의 프라임 편집 효율을 비교하여 나타낸 결과이다.

도 3은 마우스 유래 NIH/3T3 또는 C2C12 세포에서 순서대로 *Igf2*, *Adamts20*, *Casp1*, *Hoxd13*, *Angpt1* 및 *Ksr2* 각각의 유전자 내 표적에 대하여 dsgrNA 사용 여부 (PE3, PE3 + dsgrNA)에 따른 PE3의 프라임 편집 효율을 비교 분석한 결과이다.

도 4는 염색질 조절 펩타이드 (CMP) 결합에 의한 프라임 편집 효율 증진 효과를 확인한 결과로서, 도 4a는 CMP가 결합된 2종류의 융합단백질 (CMP-PE-V1, CMP-PE-V2) 구조를 그림으로 도시한 것이고, 도 4b 및 도 4c는 각각 CMP-PE-V1 및 CMP-PE-V2의 아미노산 서열을 구성별로 구분하여 나타낸 것이며, 도 4d는 NIH/3T3 또는 C2C12 세포에서 *Igf2*, *Adamts20*, *Casp1*, *Hoxd13*, *Angpt1* 및 *Ksr2* 각각의 유전자 내 표적에 대하여 CMP (HN1, H1G) 결합 여부에 따른 프라임 편집 효율을 비교 분석한 결과이다. 도 4e는 인간세포인 HEK293T 세포에서 HEK3 서열을 표적으로 프라임 편집 효율을 비교 분석한 결과이다.

도 5는 NIH/3T3 및 C2C12 세포에서 *Igf2*, *Adamts20*, *Casp1*, *Hoxd13*, *Angpt1* 및 *Ksr2* 각각의 유전자 내 표적에 대하여 PE3, dsgrNA를 이용한 PE3 + dsgrNA, CMP가 결합된 CMP-PE3-V1, 및 CMP와 dsgrNA를 함께 이용한 CMP-PE3-V1 + dsgrNA의 프라임 편집 효율을 비교 분석한 결과이다.

도 6은 PE3, PE3 + dsgrNA, CMP-PE3-V1 및 CMP-PE3-V1 + dsgrNA의 각 프라임 편집 시스템이 주입된 마우스 배아에서 표적 돌연변이의 생성 빈도를 분석한 결과로서, 도 6a는 *Igf2*에서의 결과이고, 도 6b는 *Adamts20*, *Hoxd13*, *Angpt1*, *Ksr2* 및 *Ar*에서의 결과를 나타낸 것이다.

도 7은 NIH/3T3 및 C2C12 세포에서 DNase I digestion 분석 후 실시간 qPCR을 통해 각각 온전한 게놈 DNA의 분획을 측정된 결과로서, 도 7a는 상기 각 세포에서 *Igf2*, *Adamts20*, *Casp1*, *Hoxd13*, *Angpt1* 및 *Ksr2* 표적 유전자좌에서 온전한 게놈 DNA의 분획을 측정된 결과이고, 도 7b는 상기 두 세포주에 DNase I을 2 내지 16U으로 상이하게 처리하고 온전한 게놈 DNA의 분획을 측정된 결과이며, 도 7c는 C2C12 세포에 PE3, PE3 + dsgrNA, CMP-PE3-V1, 또는 CMP-PE3-V1 + dsgrNA를 암호화하는 플라스미드를 형질감염시킨 후 *Igf2* 표적 유전자좌에서 온전한 게놈 DNA의 분획을 측정된 결과이다.

도 8은 dsgrNA를 이용한 PE3를 매개로 마우스에서 표적 돌연변이 생성을 유도한 결과로서, 도 8a는 *Igf2* 유전자의 exon4에서 표적 돌연변이 설계를 나타낸 개략도이고, 도 8b는 *Igf2*에서 proximal dsgrNA +7를 이용한 PE3에 의해 유도된 표적 돌연변이 (G에서 C로의 치환 및 TA 삽입)를 갖는 두 마우스의 유전형 및 생어 염기서열분석 크로마토그램 결과를 나타낸 것이며, 도 8c는 제작된 *Igf2* 돌연변이 마우스의 F1 세대 마우스 7마리에 대하여 생어 염기서열분석 및 유전형 분석을 통해 표적 돌연변이 여부를 확인한 결과이다.

도 9는 본 발명에 따른 프라임 편집기의 오프-타겟 효과를 분석한 결과로서, 도 9a는 야생형 및 상기 도 8에서 제작된 돌연변이 *Igf2* #1 및 *Igf2* #2 마우스에 대하여 *Igf2* 표적 돌연변이 생성에 이용된 pegRNA 및 nsgRNA의 잠재적 오프-타겟 부위에서 삽입/결실 (indel) 빈도를 측정하여 나타낸 결과이고, 도 9b는 야생형 (*Igf2* WT) 및 *Igf2* #1 마우스에 대하여 전장 유전체 염기서열분석을 실시하고 비교한 결과이며, 도 9c는 잠재적 오프-타겟 (OT) 위치에 대한 염기서열을 비교하고 생어 염기서열분석 크로마토그램을 나타낸 결과이다.

도 10은 *Igf2* 돌연변이 마우스의 표현형을 확인한 결과로서, 도 10a는 *Igf2*^{p+/m+} 수컷 (F1)을 야생형 암컷 마우스와 교배하여 얻어진 *Igf2* 돌연변이 마우스 (*MUT(Igf2*^{p+/m+}*)*)의 왜소증 표현형을 보여주는 이미지이고, 도 10b는 상기 *Igf2* 돌연변이 마우스와 *Igf2* 야생형 마우스의 체중을 측정하여 비교한 결과이다.

도 11은 본 발명에 따른 3종류의 편집 효율이 개선된 프라임 편집 시스템의 구성 및 작용 기전을 종래 공지된

PE3와 비교하여 그림으로 도시한 것이다.

발명을 실시하기 위한 구체적인 내용

- [0050] 본 발명자들은 proxymal dsgrRNA (dead sgRNA) 및/또는 염색질-조절 펩타이드 (chromatin-modulating peptides, CMPs)를 사용하여 종래의 문제점이 개선된 프라임 편집기를 개발하고 이의 우수한 편집 효율 및 표적 특이성을 확인함으로써 본 발명을 완성하였다.
- [0051] 이에, 본 발명은 (a) i) CRISPR/Cas9 단백질 또는 이의 변이체, 및 ii) 역전사효소 (Reverse Transcriptase) 또는 이의 변이체를 포함하는, 융합단백질 또는 이를 암호화하는 핵산; 및
- [0052] (b) 가이드 RNA 또는 이를 암호화하는 핵산을 포함하고,
- [0053] 이때, 상기 가이드 RNA는 pegRNA (prime editing guide RNA) 및 데드 단일 가닥 RNA (dead single guide RNA, dsgrRNA)를 포함하며, 상기 dsgrRNA는 10~20nt 길이인 것을 특징으로 하는, 유전자 교정용 조성물을 제공한다.
- [0054] 상기 유전자 교정용 조성물은 비표적 DNA 가닥에 상보적으로 결합하여 표적 DNA 가닥의 절단을 유도하는 단일 가이드 RNA (single guide RNA, sgRNA)를 추가로 포함할 수 있으며, 바람직하게 상기 가이드 RNA는 본 발명에서 nicking sgRNA를 의미하는 것일 수 있다.
- [0055] 본 발명에서 사용되는 용어, "유전자 교정 (gene editing)"은 유전자 편집, 게놈 편집 등과 동일한 의미로 사용될 수 있다. 유전자 교정은 표적 유전자 내의 표적 부위에 하나 이상의 염기에 대하여 돌연변이를 유발하는 변이 (치환, 삽입 또는 결실)를 의미하는 것이다. 바람직하게 상기 유전자 교정은 표적 유전자의 이중가닥 절단 (double-stranded DNA cleavage)을 수반하지 않는 것일 수 있으며, 더욱 바람직하게는 프라임 편집 (prime editing)을 매개로 이루어지는 것일 수 있다.
- [0056] 본 발명의 일 구현예에서, 상기 하나 이상의 염기에 대하여 돌연변이를 유발하는 변이 또는 유전자 교정은 표적 부위에 종결코돈을 생성시키거나, 야생형과 다른 아미노산을 코딩하는 코돈을 생성시킴으로써 표적 유전자를 불활성화 (knock-out) 시킨다. 또는 개시 코돈을 다른 아미노산으로 바꾸어서 유전자를 불활성화 시키거나 또는 유전자 변이를 교정하거나, 삽입 또는 결실에 의한 프레임쉬프트로(frameshift)로 유전자를 불활성화 시키거나 또는 유전자 변이를 교정하거나, 단백질을 생성하지 않는 비코딩 DNA 서열에 변이를 도입하거나, 질병을 유발하는 야생형과 다른 서열의 DNA를 야생형과 동일한 서열로 변화시키는 등의 다양한 형태일 수 있으나, 이에 제한되는 것은 아니다.
- [0057] 본 발명에서, 용어 "염기서열"은 해당 염기를 포함하는 뉴클레오타이드의 서열을 의미하는 것으로, 뉴클레오타이드 서열, 핵산 서열 또는 DNA 서열과 동일한 의미로 사용될 수 있다.
- [0058] 본 발명에서, 상기 '표적 유전자 (target gene)'는 유전자 교정의 대상이 되는 유전자를 의미하고, '표적 부위 (target site 또는 target region)'는 표적 유전자 내의 표적 특이적 뉴클레아제에 의한 유전자 편집 또는 교정이 일어나는 부위를 의미하는 것으로, 일 예에서 상기 표적 특이적 뉴클레아제가 RNA-가이드 뉴클레아제 (RNAguided engineered nuclease, RGEN)를 포함하는 것인 경우, 표적 유전자 내의 RNA-가이드 뉴클레아제가 인식하는 서열 (PAM 서열)의 5' 말단 및/또는 3' 말단에 인접하여 위치하는 것일 수 있다.
- [0059] 또한, 본 발명은 (a) i) CRISPR/Cas9 단백질 또는 이의 변이체, ii) 역전사효소 (Reverse Transcriptase) 또는 이의 변이체, 및 iii) 염색질 조절 펩타이드 (Chromatin-modulating peptides)를 포함하는, 융합단백질 또는 이를 암호화하는 핵산; 및
- [0060] (b) 가이드 RNA 또는 이를 암호화하는 핵산을 포함하고,
- [0061] 상기 가이드 RNA는 pegRNA (prime editing guide RNA) 및 데드 단일 가닥 RNA (dead single guide RNA, dsgrRNA)인 것을 특징으로 하는, 유전자 교정용 조성물을 제공한다.
- [0062] 상기 유전자 교정용 조성물은 비표적 DNA 가닥에 상보적으로 결합하여 표적 DNA 가닥의 절단을 유도하는 단일 가이드 RNA (single guide RNA, sgRNA)를 추가로 포함할 수 있으며, 바람직하게 상기 가이드 RNA는 본 발명에서 nicking sgRNA를 의미하는 것일 수 있다.
- [0063] 본 발명에서, 상기 염색질 조절 펩타이드는 뉴클레오솜 재배열 및/또는 염색질 재형성을 용이하게 하기 위하여 뉴클레오솜 및/또는 염색체 단백질들과 상호작용하는 염색체 단백질 또는 이의 단편들을 의미한다. 보다 구체적으로, 상기 염색질 조절 펩타이드는 고-이동성 그룹 뉴클레오솜 결합 도메인 1 (high-mobility group

nucleosome binding domain 1, HN1) 또는 이의 단편, 히스톤 H1 중심 구형 도메인 (histone H1 central globular domain, H1G) 또는 이의 단편, 또는 이의 조합일 수 있으나, 이에 제한되는 것은 아니다.

[0064] 상기 고-이동성 그룹 뉴클레오솜 결합 도메인 (HMGN)은 염색질의 구조 및 기능을 조절하는 염색체 단백질이고, 상기 히스톤 H1 중심 구형 도메인 (histone H1 central globular domain, H1G)은 '링커 히스톤'이라고도 알려진 히스톤 H1을 구성하는 도메인이다. 히스톤 H1은 뉴클레오솜 어레이의 압축 상태를 조절하고 형태에 영향을 주며, 상기 중심 구형 도메인은 뉴클레오솜 상의 링커 DNA의 entry/exit 부위 근처에 결합한다고 알려져 있다.

[0065] 상기 염색질 조절 펩타이드는 화학적 결합에 의해 직접적으로, 링커에 의해 간접적으로, 또는 이의 조합으로 상기 CRISPR/Cas9 단백질 또는 역전사효소에 연결될 수 있다. 구체적으로, 최소한 하나의 염색질 조절 펩타이드는 CRISPR/Cas9 단백질의 N-말단, C-말단, 및/또는 내부 위치에 연결될 수 있다. 바람직하게, 본 발명의 융합단백질은 CRISPR/Cas9 단백질 또는 역전사효소에 연결된 2개의 염색질 조절 펩타이드를 포함한다. 보다 바람직하게, 상기 융합단백질에서 HMGN1 (HN1)은 Cas9 단백질 또는 이의 변이체의 N-말단에 연결될 수 있고, 상기 히스톤 H1 중심 구형 도메인 (H1G)은 Cas9 단백질 또는 이의 변이체의 C-말단 또는 역전사효소의 C-말단에 연결될 수 있다. 가장 바람직하게 상기 융합단백질은 i) N말단-[HN1]-[Cas9]-[H1G]-[역전사효소]-C말단; 또는 ii) N말단-[HN1]-[Cas9]-[역전사효소]-[H1G]-C말단의 구성으로 이루어진 것일 수 있다.

[0066] 본 발명에 있어서, 상기 융합단백질은 최소한 하나의 핵 위치화 신호, 최소한 하나의 세포-침투 도메인, 최소한 하나의 마커 도메인, 또는 이의 조합을 추가로 포함할 수 있으며, 바람직하게는 N-말단 및 C-말단에 각각 핵 위치화 신호 (nuclear localization signal, NLS) 서열을 추가로 포함할 수 있으나, 이에 제한되는 것은 아니다.

[0067] 본 발명에 따른 상기 융합단백질은 서열번호 1 또는 서열번호 2로 표시되는 아미노산 서열로 이루어진 것일 수 있다. 이때, 상기 융합단백질은 서열번호 1 또는 서열번호 2로 표시되는 아미노산 서열과 70% 이상, 바람직하게는 80% 이상, 더욱 바람직하게는 90% 이상, 가장 바람직하게는 95%, 96%, 97%, 98%, 99% 이상의 서열 상동성을 가지는 아미노산 서열을 포함할 수 있다.

[0068] 본 발명에서, 'Cas9 (CRISPR associated protein 9) 단백질'은 DNA 바이러스에 대한 특정 박테리아의 면역학적 방어에서 중요한 역할을 하는 단백질로 유전공학 응용에 많이 사용되는데, 상기 단백질의 주요 기능이 DNA를 절단하는 것이기 때문에 세포의 게놈을 변형하는 데에 적용할 수 있다. 구체적으로, CRISPR/Cas9은 3세대 유전자 가위로써 이용하고자 하는 특정 염기서열을 인식하여 절단하고 편집하며, 게놈의 목적 장소에 특정 유전자를 삽입하거나 특정 유전자의 활동을 정지시키는 조작을 간단하고 신속하고 효율성 있게 실시하는데 유용하다. Cas9 단백질 또는 유전자 정보는 NCBI (National Center for Biotechnology Information)의 GenBank와 같은 공지의 데이터베이스에서 얻을 수 있으나, 이에 제한되는 것은 아니다. 또한, Cas9 단백질은 그 목적에 따라 당업자가 추가적인 도메인을 적절하게 연결할 수 있다. 본 발명에 있어서, Cas9 단백질은 야생형 Cas9 뿐만 아니라, 유전자 편집을 위한 핵산분해효소의 기능을 갖는 것이라면 Cas9의 변이체를 모두 포함할 수 있다.

[0069] 본 발명에 있어서, 상기 Cas9 변이체는 DNA 이중 가닥을 절단하는 엔도뉴클레아제 활성을 상실하도록 변이된 것을 의미할 수 있다. 예컨대, 상기 Cas9 변이체는 엔도뉴클레아제 활성을 상실하고 니카아제 활성을 갖도록 변이된 Cas9 단백질 및 엔도뉴클레아제 활성과 니카아제 활성을 모두 상실하도록 변이된 Cas9 단백질 중에서 선택된 1종 이상일 수 있으며, 바람직하게는 Cas9 니카아제 (Cas9 nickase)일 수 있다.

[0070] 상기 Cas9 니카아제는 뉴클레아제의 촉매 활성 도메인 (예컨대, Cas9의 RuvC 또는 HNH 도메인)에서 변이가 일어나 불활성화된 것일 수 있다. 구체적으로, 10번째 위치의 아스파르트산 (D10), 762번째 위치의 글루탐산 (E762), 840번째 위치의 히스티딘 (H840), 854번째 위치의 아스파라긴 (N854), 863번째 위치의 아스파라긴 (N863) 및 986번째 위치의 아스파르트산 (D986) 등으로 이루어진 군에서 선택된 하나 이상이 임의의 다른 아미노산으로 치환된 돌연변이를 포함할 수 있으며, 바람직하게 본 발명의 Cas9 니카아제는 840번째 위치의 히스티딘이 알라닌으로 치환 (H840A)된 변이를 포함하는 것일 수 있으나, 이에 제한되지 않는다.

[0071] 상기 Cas9 단백질 또는 이의 변이체는 그 유래가 제한되지 않으며, 비제한적인 예시로서 스트렙토코커스 피요제네스 (*Streptococcus pyogenes*), 프란시셀라 노비시다 (*Francisella novicida*), 스트렙토코커스 써모필러스 (*Streptococcus thermophilus*), 레지오넬라 뉴모필라 (*Legionella pneumophila*), 리스테리아 이노큐아 (*Listeria innocua*), 또는 스트렙토코커스 뮤탄스 (*Streptococcus mutans*) 유래일 수 있다.

[0072] 상기 Cas9 단백질 또는 이의 변이체는 미생물에서 분리된 것 또는 재조합적 방법 또는 합성적 방법 등과 같이 인위적 또는 비자연적 생산된 것 (non-naturally occurring)일 수 있다. 상기 Cas9은 *in vitro*에서 미리 전사된 mRNA 또는 미리 생산된 단백질 형태, 또는 표적 세포 또는 생체 내에서 발현하기 위하여 재조합 벡터에 포함

된 형태로 사용될 수 있다. 일 실시예에서, 상기 Cas9은 재조합 DNA(Recombinant DNA, rDNA)에 의하여 만들어진 재조합 단백질일 수 있다. 재조합 DNA는 다양한 유기체로부터 얻어진 이중 또는 동종 유전 물질을 포함하기 위하여 분자 클로닝과 같은 유전자 재조합 방법에 의해 인공적으로 만들어진 DNA 분자를 의미한다.

[0073] 본 발명에서, '역전사효소 (reverse transcriptase)'는 RNA를 주형으로 삼아 DNA를 합성해내는 능력을 갖고 있는 효소를 말한다. 본 발명에 있어서, 역전사효소는 야생형 역전사효소뿐만 아니라 상기와 같이 RNA를 주형으로 삼아 DNA를 합성해내는 기능을 갖는 것이라면 역전사효소의 변이체를 모두 포함할 수 있으며, 역전사효소 또는 이의 변이체는 바람직하게 몰로니 마우스 백혈병 바이러스 (Moloney murine leukemia virus, M-MLV)에서 유래된 것일 수 있으나, 이에 제한되는 것은 아니다.

[0074] 본 발명에서, 사용되는 용어 "가이드 RNA (guide RNA)"는 표적 유전자 내 표적 부위 내의 특이적인 염기 서열 (표적서열)에 혼성화 가능한 표적화 서열을 포함하는 RNA를 의미하며, 생체 외 (*in vitro*) 또는 생체 (또는 세포) 내에서 Cas와 같은 뉴클레아제 단백질과 결합하여 이를 표적 유전자 (또는 표적 부위)로 인도하는 역할을 한다.

[0075] 상기 가이드 RNA는 복합체를 형성할 뉴클레아제의 종류 및/또는 그 유래 미생물에 따라서 적절히 선택될 수 있다. 예컨대, 본 발명의 가이드 RNA는 pegRNA, dead sgRNA 또는 nicking sgRNA일 수 있다.

[0076] 상기 가이드 RNA는 표적 유전자 (표적 부위) 내의 표적 서열과 상보적인 서열 (표적화 서열)을 가지는 부분인 스페이서 영역 (Spacer region, Target DNA recognition sequence, base pairing region 등으로도 명명함) 및 Cas9 단백질 결합을 위한 hairpin 구조를 포함할 수 있다. 보다 구체적으로는 표적 유전자 내의 표적 서열과 상보적인 서열을 포함하는 부분, Cas 단백질 결합을 위한 hairpin 구조, 및 Terminator 서열을 포함할 수 있다. 추가적으로 상기 pegRNA는 표적 유전자의 비표적 가닥에 상보적인 염기서열 (primer binding site, PBS) 및 고정하고자 하는 염기서열을 포함하는 역전사효소의 주형가닥 부위 (RT template)를 갖는다.

[0077] 상기 가이드 RNA의 표적 서열과 혼성화 가능한 가이드 RNA의 표적화 서열은 상기 표적 서열이 위치하는 DNA 가닥 (즉, PAM 서열(5'-NGG-3' (N은 A, T, G, 또는 C임))이 위치하는 DNA 가닥) 또는 이의 상보적인 가닥의 뉴클레오타이드 서열과 50% 이상, 60% 이상, 70% 이상, 80% 이상, 90% 이상, 95% 이상, 99% 이상, 또는 100%의 서열 상보성을 갖는 뉴클레오타이드 서열을 의미하는 것으로, 상기 상보적 가닥의 뉴클레오타이드 서열과 상보적 결합이 가능하다.

[0078] 상기 가이드 RNA는 RNA 형태로 사용 (또는 상기 조성물에 포함)되거나, 이를 코딩하는 DNA를 포함하는 플라스미드 형태로 사용 (또는 상기 조성물에 포함)될 수 있다.

[0079] 본 발명에서, 상기 dsgrRNA는 10~20nt (뉴클레오타이드) 길이로 구성된 것일 수 있고, 바람직하게는 11 내지 19nt, 12 내지 18nt, 13 내지 17nt, 13 내지 16nt, 가장 바람직하게는 14 내지 15nt 길이일 수 있으나, 이것으로 한정되는 것은 아니다.

[0080] 또한 상기 dsgrRNA는 pegRNA 결합 부위, 바람직하게는 pegRNA의 스페이서 (spacer) 부위로부터 5 내지 70nt (뉴클레오타이드), 바람직하게는 6 내지 65nt, 가장 바람직하게는 7 내지 62nt 떨어진 위치에 결합하여 상기 융합단백질의 염색질 접근성을 증가시킬 수 있다.

[0081] 본 발명에서 사용되는 용어, "염색질 접근성 (Chromatin accessibility)"이란 주로 히스톤 (histone), 전사인자 (transcription factor, TF), 염색질 변형 효소 (chromatin-modifying enzymes) 및 염색질 리모델링 복합체 (chromatin-remodelling complexes)로 구성된 DNA 및 관련 단백질에 의해 형성된 복합체인 염색질의 물리적 압축 수준을 의미한다. 진핵생물 게놈은 일반적으로 히스톤 8량체 (octamer)를 감싸고 있는 ~147bp의 DNA를 포함하는 뉴클레오솜으로 압축되지만, 뉴클레오솜의 점유율은 게놈에서 균일하지 않으며 조직 및 세포 유형에 따라 다르게 나타난다. 뉴클레오솜은 일반적으로 전사 조절제 (ex. 전사인자)와 상호작용하는 시스 조절 요소 (인핸서 및 프로모터)가 존재하는 게놈 위치에서 고갈되어 접근 가능한 염색질을 생성하게 된다. 유전자 교정 (유전자 편집)과 관련해서는 닫힌 게놈 영역 보다 열린 게놈 영역을 표적하는 gRNA의 활성화에 유의한 차이가 있어 Cas9의 효율이 국소적 염색질 접근성에 의해 영향을 받으므로 염색질 접근성과 CRISPR-Cas9 매개 유전자 편집 효율성 사이의 양의 상관관계가 있음이 공지되어 있다.

[0082] 본 발명의 다른 양태로서, 본 발명은 상기 유전자 교정용 조성물을 *in vitro* 또는 *ex vivo* 상에서 표적 핵산 서열을 포함하는 표적 영역 (region)과 접촉시키는 단계를 포함하는, 유전자 교정 방법을 제공한다.

[0083] 상기 유전자 교정용 조성물은 바람직하게 진핵세포에 적용될 수 있으며, 상기 진핵세포는 바람직하게 인간 등의

영장류, 마우스 등의 설치류를 포함하는 포유동물에서 유래된 것일 수 있으나 이에 한정되는 것은 아니다.

- [0084] 본 발명의 또 다른 양태로서, 본 발명은 상기 유전자 교정용 조성물을 포함하는, 유전자 교정용 키트를 제공한다.
- [0085] 본 발명에 있어서, 상기 키트는 상기 유전자 교정용 조성물과 함께 버퍼 (buffer) 및 데옥시리보뉴클레오티드-5-트리포스페이트 (dNTP)와 같은 유전자 교정을 실시하는데 필요한 물질 (시약)을 모두 포함할 수 있다. 또한, 상기 키트의 특정 반응에서 사용되는 시약의 최적량은, 본 명세서에 개시 사항을 습득한 당업자에 의해서 용이하게 결정될 수 있다.
- [0086] 본 발명의 또 다른 양태로서, 상기 유전자 교정용 조성물을 인간을 제외한 포유동물 세포에 주입하여 유전자 변형 포유동물 세포를 얻는 단계; 및 상기 얻어진 유전자 변형 포유동물 세포를 인간을 제외한 포유동물 위탁모 난관에 이식하는 단계를 포함하는, 인간을 제외한 유전자 변형 포유동물의 제조방법을 제공한다.
- [0087] 본 발명의 상기 유전자 교정 조성물의 도입에 의해 포유동물 세포에서 40% 이상, 45% 이상, 50% 이상, 55% 이상, 60% 이상, 65% 이상, 70% 이상, 75% 이상, 80% 이상, 85% 이상, 90% 이상, 95% 이상, 97% 이상, 99% 이상 또는 100%의 유전자 교정효율 (예컨대, 염기 치환, 삽입 또는 결실)을 달성할 수 있다.
- [0088] 본 발명에서, 상기 포유동물 세포에 유전자 교정용 조성물을 도입하는 단계는, i) 상기 세포를 본 발명에 따른 프라임 편집용 융합단백질, pegRNA, nsgRNA 및 dsgrNA를 암호화하는 플라스미드 벡터 또는 바이러스 벡터로 형질감염시키거나,
- [0089] ii) 상기 세포에 융합단백질을 암호화하는 mRNA, pegRNA, nsgRNA 및 dsgrNA의 혼합물 또는 이들 각각을 직접 주입하거나,
- [0090] iii) 상기 세포에 융합단백질, pegRNA, nsgRNA 및 dsgrNA의 혼합물 또는 복합체 형태의 리보핵산단백질을 직접 주입하여 수행될 수 있다.
- [0091] 상기 직접 주입은 상기 ii) 또는 iii)의 각 mRNA 및 가이드 RNA 또는 리보핵산단백질이 재조합 벡터를 사용하지 않고, 세포막 및/또는 핵막을 통과하여 유전체(genome)에 전달되는 것을 의미할 수 있으며, 예컨대, 나노입자, 전기천공법, 리포펙션 (lipofection), 미세주입 (microinjection) 등에 의하여 수행될 수 있다.
- [0092] 상기 유전자 교정 조성물이 도입되는 포유동물 세포는 인간 등의 영장류, 마우스 등의 설치류를 포함하는 포유동물의 배아일 수 있으며, 바람직하게는 인간을 제외한 포유동물의 배아일 수 있다. 예컨대, 상기 배아는 과배란 유도된 암컷 포유동물과 수컷 포유동물을 교배하여 얻어진 수정된 배아를 상기 암컷 포유동물의 난관에서 채취한 것일 수 있다. 상기 염기 교정용 조성물이 적용 (주입)되는 배아는 수정된 1-세포기의 배아 (zygote)일 수 있다.
- [0093] 상기 얻어진 유전자 변형 포유동물 세포는 상기 유전자 교정 조성물의 도입에 의하여 표적 유전자에 염기 치환, 삽입 또는 결실 돌연변이가 발생한 세포일 수 있다.
- [0094] 상기 유전자 변형 포유동물 세포, 바람직하게 유전자 변형 배아 세포를 난관에 이식받는 포유동물은 상기 배아 세포가 유래하는 포유동물과 동종의 포유류 (위탁모)일 수 있다.
- [0095] 또한, 본 발명은 상기 방법에 의해 제조된, 유전자 변형 포유동물을 제공한다.
- [0097] 이하, 본 발명의 이해를 돕기 위하여 바람직한 실시예를 제시한다. 그러나 하기의 실시예는 본 발명을 보다 쉽게 이해하기 위하여 제공되는 것일 뿐, 하기 실시예에 의해 본 발명의 내용이 한정되는 것은 아니다.

[실시예]

실시예 1. 실험방법

1-1. 플라스미드 DNA 구축

- [0102] NEBuilder® HiFi DNA Assembly Master Mix (E2621L, NEB)를 사용하여 HN1 (High-mobility group nucleosome binding domain 1) 및 H1G (Histone H1 central globular domain) 올리고를 pCMV-PE2 (#132775, Addgene)에서 nCas9의 양면에 융합시켜 제조사의 프로토콜에 따라 염색질 조절 펩타이드 프라임 편집기 (chromatin-

modulating peptide prime editor)를 구축하였다.

[0103] 또한 *Igf2*, *Adamts20*, *Casp1*, 및 *Hoxd13* 유전자에 특정 돌연변이를 유발시키기 위한 pegRNA 발현 벡터를 제작하기 위해, pU6-pegRNA-GG-수용체 벡터 (#132777, Addgene)에 스페이서 (spacer), 프라이م 결합 부위 (prime binding site) 및 역전사효소 주형 (reverse transcriptase template oligos) 서열들을 BsaI 효소 절단 부위에 삽입하였다. nsgRNA 및 dsgrNA 발현 벡터는 pRG2-GG 벡터 (#104174, Addgene)로 삽입하였다. 추가적으로 본 실시예에서 이용된 각 표적 유전자에 특이적인 pegRNAs, nsgRNA 및 dsgrNA의 서열을 하기 표 1 내지 3에 정리하여 나타내었다.

[0104] [표 1]

pegRNA	서열 (5' to 3')		PBS 길이 (nt)	RT 주형 길이 (nt)	서열번호
tdTomato_G to C / T ins_8-17	Spacer	CGCATGGAGGGCTCCATGAA	8	17	3
	Primer Binding Site	ATGGAGCC			4
	RT template	GAACTCAGTGGCGGTTC			5
<i>Igf2</i> _G to C / T A ins_9-14	Spacer	TATTGGAAGAAGTGGCCAC	9	14	6
	Primer Binding Site	GGCAAGTTC			7
	RT template	AGATACCGCGTGTA			8
<i>Adamts20</i> _CG to AA_11-13	Spacer	AGTGAATAAGAAGACGTACT	11	13	9
	Primer Binding Site	ACGTCTTCTTA			10
	RT template	GACCGGCCTTAGT			11
<i>Casp1</i> _TAGG del_12-12	Spacer	GTCTTGCTCTTATAGGAGA	12	12	12
	Primer Binding Site	CCTATAAGAGAC			13
	RT template	ACCTCTTCACT			14
<i>Hoxd13</i> _G to T_10-15	Spacer	GAGGCATACATCTCCATGGA	10	15	15
	Primer Binding Site	ATGGAGATGT			16
	RT template	GACTGGTAGACCTCC			17
<i>Angpt1</i> _CGG to TGA_11-13	Spacer	CACATTGCCCCATGTTGAATC	11	13	18
	Primer Binding Site	TCAACATGGGC			19
	RT template	GATATAACTGAAT			20
<i>Ksr2</i> _TGAT ins_8-14	Spacer	TCCTGCCCTGGCTCCGTGGT	8	14	21
	Primer Binding Site	ACGGAGCC			22
	RT template	GTGGCCACCTGAT			23
<i>Ar</i> _G to T_13-13	Spacer	TCTCACTGTGGCAGCTGCA	13	13	24
	Primer Binding Site	AGCTGCCACAAGT			25
	RT template	AAGAAGACCTTGA			26

[0105]

[0106] [표 2]

Nicking sgRNA	서열 (5' to 3')	서열번호
tdTomato-nsgRNA	GCCCTCGATCTCGAACTCAG	27
<i>Igf2</i> -nsgRNA	CTTCCCCAGATACCGCGTGT	28
<i>Adamt20</i> -nsgRNA	AGCTCGTGTTCAAGGACATG	29
<i>Casp1</i> -nsgRNA	CTGTGAGAAGTCTTGTGCTC	30
<i>Hoxd13</i> -nsgRNA	GATCCTTGGCACAGTACACC	31
<i>Angpt1</i> -nsgRNA	GATATAACTGAATTCAACAT	32
<i>Ksr2</i> -nsgRNA	ACCTGATACGGAGCCAGGGC	33
<i>Ar</i> -nsgRNA	AGGGGAAAAATATCAGGAAGT	34

[0107]

[0108] [표 3]

dsgRNA	서열 (5' to 3')	서열번호
<i>Igf2</i> _dsgRNA_-34	AGTCAGCAGGTTTC	35
<i>Igf2</i> _dsgRNA_-27	AGGTGCTAGTCAGC	36
<i>Igf2</i> _dsgRNA_+7	TGGAGACAGTCCGC	37
<i>Igf2</i> _dsgRNA_+22	GGACGCCTGCGCAG	38
<i>Adamts20</i> _dsgRNA_-13	TCAAGAGCACAGCC	39
<i>Adamts20</i> _dsgRNA_+14	TTTCTGGTGCTTGG	40
<i>Adamts20</i> _dsgRNA_+24	CGATTTTGACTTTC	41
<i>Adamts20</i> _dsgRNA_+41	CTTCCGCTCACTTC	42
<i>Casp1</i> _dsgRNA_-13	CTTTGACTTCTCTA	43
<i>Casp1</i> _dsgRNA_-12	ACTTTGACTTCTCTA	44
<i>Casp1</i> _dsgRNA_+7	CAGCAAATTCTTTC	45
<i>Casp1</i> _dsgRNA_+62	GTATTCATGTCTCA	46
<i>Hoxd13</i> _dsgRNA_-12	CCCGGATCCAAAAG	47
<i>Hoxd13</i> _dsgRNA_-11	CACTTTTGGATCCG	48
<i>Hoxd13</i> _dsgRNA_+20	GCTGTTCCACCCGT	49
<i>Hoxd13</i> _dsgRNA_+24	CGGGTGGAAACAGCC	50
<i>Angpt1</i> _dsgRNA_-15	CGAAATCCAGAAAA	51
<i>Angpt1</i> _dsgRNA_-11	ATCCAGAAAACGGA	52
<i>Angpt1</i> _dsgRNA_+14	CTTCCAGAACACGA	53
<i>Angpt1</i> _dsgRNA_+17	TTCCCGTCGTGTTC	54
<i>Ksr2</i> _dsgRNA_-12	TCTCCAAACAAGAT	55
<i>Ksr2</i> _dsgRNA_+7	TCCGGGGGGGCACAC	56
<i>Ar</i> _dsgRNA_-15	GGAGATGAAGCTTC	57

[0109]

[0111] 1-2. 리포펙션 (lipofection) 및 전기천공 (electroporation)

[0112] AAVS1 유전자좌에서 tdTomato를 발현하는 리포터 시스템이 도입된 HEK293T 세포, NIH/3T3 세포 (ATCC, CRL-1658) 및 C2C12 세포 (ATCC, CRL-1772)는 10% FBS (S 001-01, Welgene) 또는 BCS (26170-043, Gibco)가 보충된 Dulbecco's Modified Eagle's Medium (DMEM; LM001-05, Welgene)에서 5% CO₂ 및 37°C 조건하에 배양되었다. 본 실시예에서는 제조사의 프로토콜에 따라 0.5μg pegRNA, 2.15μg PE2, 0.22μg nsgrRNA의 3가지 플라스미드와 1 μl Lipofectamine 2000 시약 (11668019, Thermo Fisher Scientific)이 포함된 50μl Opti-MEM (31985070, Gibco)을 2 x 10⁴ 세포에 처리하여 형질감염시킨 후 11일 동안 세포를 배양하였다.

[0113] 전기천공은 각각 1 x 10⁵개의 NIH/3T3 세포 및 C2C12 세포에 대해 실시하였으며, 상기 각 세포주를 3μg PE2 또는 CMP-PE, 0.7μg pegRNA, 0.3μg nsgrRNA 및 0.25μg dsgRNA의 플라스미드와 혼합하고 제조사의 프로토콜에 따라 Neon™ Transfection System (MPK1096, Thermo Fisher Scientific)을 이용하여 형질감염시켰다. 이후 72시간 동안 세포를 배양한 다음 회수하고 표적 심층 염기서열분석 (targeted deep sequencing)을 실시하였다.

[0115] 1-3. DNA 앰플리콘의 제조

[0116] 형질감염 후 tdTomato 발현 리포터 HEK293T, NIH/3T3, C2C12 세포 및 마우스 배아로부터 DNeasy Blood & Tissue Kits (69506, Qiagen)를 사용하여 게놈 DNA를 추출하였다. 이후 Phusion™ High-Fidelity DNA 중합효소 (F-530XL, Thermo Fisher Scientific) 및 Sungen (SG-PT02, Sun genetics)을 사용하여 편집된 표적 서열을 증폭시켰다.

[0118] 1-4. In vitro 전사

[0119] mMESSAGE mMACHINE T7 Ultra Kit(AM1345, Invitrogen)를 사용하여 PE2 및 CMP-PE의 전사체를 제조하고

MEGAclean™ Transcription Clean-Up Kit(AM1908, Invitrogen)를 이용하여 정제하였다. 제조사의 프로토콜에 따라 T7 RNA 중합효소(M0251, NEB)를 사용하여 pegRNA, nsgRNA, 및 dsRNA의 전사를 유도하였고, Expin™ CleanUp SV(113-150, GeneA11)를 사용하여 전사된 RNAs를 정제하였다. 이후 NanoDrop One UV-Vis(Thermo Fisher Scientific)를 이용해 상기 정제된 RNAs를 정량화하였다.

[0121] 1-5. 실험동물

[0122] 본 실시예의 *in vivo* 실험은 서울대학교 동물실험윤리위원회 (Institutional Animal Care and Use Committee, IACUC)의 승인을 받고 진행하였다. 배아 기증을 위해 C57BL/6N 및 ICR 마우스를 이용하였고, 대리모는 특정 병원체가 없는 실험실에서 사육되었다.

[0124] 1-6. 미세주입 (Microinjection)

[0125] C57BL/6N 암컷 마우스에 HyperOva (KYD-010-EX-x5, CARD) 및 hCG (CG10-1v1, Sigma)를 48시간 간격으로 주입하여 난소 과자극을 유도하였고, hCG 호르몬을 주입한 후 암컷 마우스를 야생형 C57BL/6N 수컷 마우스와 교배시켰다. 수정된 1세포 단계의 배아를 팽대부 (ampulla)에서 얻었고 두 개의 전핵 (pronuclei)이 나타날 때까지 M2 배지 (M7167, Sigma)에서 배양하였다. 미세주입을 위한 프라임 편집용 혼합물은 100μl의 Tris-EDTA (pH 7.4)하에 100ng의 PE2 또는 CMP-PE mRNA, 65ng pegRNA, 32.5ng nsgRNA, 22ng dsRNA의 농도로 준비하였다. 미세주입 후에는 배반포 (blastocyst)로의 발달을 위해 배아를 KSOM 배지 (MR-121-D, Millipore)하에 37℃ 배양기에서 4일 동안 배양하였다. 이후 2세포 단계 배아의 일부를 가임신 양모의 난관으로 이식하였다.

[0127] 1-7. 유전형 분석 (Genotyping) 및 표적 심층 염기서열분석 (targeted deep sequencing)

[0128] 각 표적은 특이적 프라이머를 사용하여 nested PCR을 통해 증폭시켰다. PCR 앰플리콘으로 구성된 라이브러리에 대하여 iSeq™ 100 시퀀싱 시스템 (Illumina, Inc.)을 이용하여 염기서열분석을 실시하였다. 이후 CRISPR REGN Tools 프로그램 (<http://www.rgenome.net/>)과 EUN 프로그램 (<https://daeunyo.com/>)을 통해 서열분석 데이터를 분석하였다.

[0130] 1-8. 전장 유전체 염기서열분석 (Whole genome sequencing) 및 변이 검출 (variant calling)

[0131] DNeasy Blood & Tissue kit (69506, Qiagen)를 사용하여 마우스 (C57BL/6N) 귀에서 게놈 DNA를 분리한 다음, 제조사의 지침에 따라 Covaris S2 초음파 장치 시스템을 사용하여 게놈 DNA를 전단하고 Truseq Nano DNA sample prep kit를 사용하여 paired-end DNA 라이브러리를 제조하였다. 딥 커버리지 (30x) 전장 유전체 염기서열분석은 Illumina Novaseq 6000 플랫폼 (Illumina)에서 101 base paired-end 시퀀싱을 통해 수행되었고, 서열 판독은 BWA-MEM을 사용하여 마우스 참조 게놈 GRCm38/mm10에 정렬하여 수행되었다. 단일 뉴클레오타이드 변이와 작은 삽입/결실 (indel)은 GATK4 HaplotypeCaller를 이용하여 검출하였고, 마우스 v142 (dbSNP142)에 대하여 dbSNP에서 존재하는 공지된 변이들은 ANNOVAR로 주석을 달았다. dbSNP142에 존재하지 않는 새로운 변이체들은 추가로 분석하여 오프-타겟 부위에서 위치를 확인하였다. 추정되는 오프-타겟 부위는 최대 7-bp 또는 2-bp 벌지 + 5bp 불일치를 고려하여 Cas-OFFinder의 후보와 비교하였다. 모든 변이는 판독 플롯을 시각화하여 수동으로 확인하였다.

[0133] 1-9. DNase I digestion assay 및 qPCR

[0134] 본 발명자들은 종래 공지된 방법에 따라 DNase I digestion assay를 실시하였다. 구체적으로, 세포를 떼어내고 차가운 1X PBS로 반복해서 씻어낸 다음 900rpm에서 5분 동안 2회 스핀 다운하였다. 이어서 차가운 RSB 완충액 (10mM Tris-HCl, 10mM NaCl 및 3mM MgCl₂) + 0.1% IGEPAL CA-630 (I8896, Sigma)을 사용하여 세포를 용해시키고 4deg에서 10분 동안 500g로 스핀 다운하여 핵을 펠렛화하였다. 다음으로 상층액을 제거하고 DNase I (2-16U)을 처리하거나 처리하지 않은 조건에서 핵을 37℃에서 20-30분 동안 배양하였다. 이후 50mM EDTA를 첨가하여 반응을 정지시키고, DNeasy Blood & Tissue kit (69506, Qiagen)를 이용하여 DNase I으로 분해된 DNA를 정

제하였다. 실시간 qPCR (real-time qPCR)은 KAPA SYBR FAST qPCR Master Mix Kit (KR0389, Kapa Biosystems)를 이용해 수행하였으며, 온전한 게놈 DNA의 분획은 비교 C_T 방법을 이용하여 측정되었다. 상기 실시간 qPCR에 사용된 프라이머 서열은 하기 표 4에 나열하였다.

[표 4]

타겟	서열 (5' to 3')		서열번호
Closed chromatin	Forward	GCAGCAGATGGCAAGTAATACTAAGAT	58
	Reverse	CCCTTATTCTCTGAGCATTAGACAGTTATA	59
Open chromatin	Forward	TGAGTCACAGCATCAGCATCGGGTCTCT	60
	Reverse	GGAGAGGGCCTTTTCTCTTCAAGGTTC	61
<i>Igf2</i>	Forward	TCTCCAGGACGACTTCCCCAGA	62
	Reverse	CTCCAGGTGTCATATTGGAAGAAC	63
<i>Adamts20</i>	Forward	GTGGAATCAAGAGCACAGC	64
	Reverse	CAGTGAATAAGAAGACGTAC	65
<i>Casp1</i>	Forward	AAATCACTGGTCTTGTCTCTTATAG	66
	Reverse	CAGCAAATTCTTTCACCTCTTTC	67
<i>Hoxd13</i>	Forward	TCGGCACGAGGCATACATCTCCATG	68
	Reverse	CGTTGGCTAGCGTCCAGGACTG	69
<i>Angpt1</i>	Forward	CCAGAAAACGAGGGGAGAAGA	70
	Reverse	TAGGCACATTGCCCATGTTGA	71
<i>Ksr2</i>	Forward	TCTCTCCAAACAAGATTGGATCAT	72
	Reverse	TCTCCTGCCCTGGCTCC	73

1-10. 통계분석

데이터는 평균과 표준 편차로 제시하였고, 3 또는 5회 독립적 반복 실험을 수행하였다. P값은 unpaired 및 two-sided Student's t-검정을 사용하여 계산되었다.

실시예 2. 포유류 게놈에서 프라임 편집 시스템의 검증 및 편집 효율 확인

본 발명자들은 먼저 포유류 게놈에서 프라임 편집 시스템의 적용 가능성을 검증하고자 하였으며, 이를 위해 HEK293T 세포의 AAVS1 유전자좌에서 tdTomato를 발현하는 리포터 시스템을 이용하였다. 구체적으로, 도 1a에 그림으로 도시한 바와 같이 프라이머 결합 부위 (PBS) 길이가 8 nt이고 역전사효소 (RT) 주형의 길이가 17 nt (PBS8-RT17)인 프라임 편집기 3 (prime editor 3, PE3)를 이용하여 상기 tdTomato 서열 내에 티민 (thymine, T) 염기를 삽입함으로써 정지 코돈을 생성시켰다. 또한, pegRNA (prime-editing guide RNA)는 편집되는 가닥에서 편집을 억제하기 위해 비표적 가닥 상의 PAM 서열을 제거하도록 설계되었다. 이후 본 발명자들은 상기 PE3에 의한 tdTomato 유전자의 편집 효율을 확인하기 위해 유세포 분석 (flow cytometry)을 통해 tdTomato-음성 세포를 게이팅하여 편집 효율을 평가하였다.

그 결과, 도 1b에서 볼 수 있는 바와 같이 tdTomato를 발현하는 양성대조군(tdTomato expressing HEK293T)과 비교할 때 PE3를 이용해 tdTomato 유전자 편집을 유도한 경우 (tdTomato expressing HEK293T + PE3) 32%의 세포가 tdTomato 음성인 것을 확인하였다. 이러한 결과는 프라임 편집기가 본 발명의 시스템에서 원하는 표적 돌연변이의 생성을 가능하게 함을 시사하는 것이다.

다음으로, 프라임 편집기를 이용하여 마우스 모델을 제작하기 위해, 8개의 전이 돌연변이 또는 하나 이상의 뉴클레오티드 삽입 또는 결실을 포함하는 표적을 설계하고자 하였다. 먼저, 도 2a에 도시한 바와 같이 2종의 마우스 유전자 즉, *Igf2* (insulin-like growth factor 2) 및 *Adamts20* (a disintegrin and metalloproteinase domain with thrombospondin type-1 motifs 20)에서 종결 코돈의 생성을 유도하였다. *Igf2* 유전자는 부계로부터 유전된 *Igf2* 대립 유전자의 돌연변이에 의해 왜소증 표현형을 유도할 수 있다. 구체적으로, 본 발명자들은 *Igf2* 유전자의 exon4에 TA 염기를 삽입함으로써 종결 코돈을 생성시켜 유전자의 기능 상실을 유도하였다. 또한, 편집되는 가닥의 지속적 편집을 억제하기 위해 PAM 서열의 뉴클레오티드를 NGG에서 NCG로 치환시켰다. 한편, *Adamts20*은 멜라닌 세포 (melanocytes)의 발달에 관여하는 유전자로서, *Adamts20* 유전자좌의 E584 부위에서 조

기 정지 코돈의 생성은 전형적인 백색 벨트 표현형과 관련이 있다고 알려져 있다. 본 발명자들은 *Adamts20* 유전자좌의 exon12에서 CG를 TT로 치환을 유도하여 조기 종결 코돈 (E584*) 및 PAM 서열의 변형 (NGG에서 NAG로)을 유도하였다. 또한 상기와 같은 마우스 유전자 표적에서 상기 돌연변이를 유도하기 위해 PE (조작된 M-MLV RT와 융합된 nCas9), pegRNA 및 nicking sgRNA (nsgRNA)로 구성된 PE3 시스템을 사용했다. 이때, 상기 nsgRNA는 표적 DNA 가닥 즉, 비-편집 가닥의 절단을 통해 DNA 복구 활성을 촉진함으로써 편집 효율성을 향상시키기 위한 목적으로 이용된 것이다.

[0145] 다음으로, 본 발명자들은 *Igf2* 및 *Adamts20* 표적 부위에서 프라임 편집을 최적화하기 위해 다양한 PBS 길이 (8-14 nt) 및 RT 주형 길이 (10-18 nt)로 구성된 pegRNA에 따른 편집 효율성을 평가하였다. 이때, 전사 종결 신호의 일부가 될 수 있는 3'-말단에 티민이 있는 PBS의 길이와 pegRNA 구조를 방해할 수 있는 5'-말단에 시토신이 있는 RT의 길이는 제외하였다. 이후 전기천공법을 통해 PE, pegRNA 및 nsgRNA를 암호화하는 3가지 플라스미드를 NIH/3T3 세포로 형질감염시키고 72시간 동안 세포를 배양한 후 회수하여 표적 심층 염기서열분석을 실시하였다.

[0146] 실험 결과, 도 2b에 나타난 바와 같이 NIH/3T3 세포에서 *Igf2* 및 *Adamts20* 표적에 대한 PE3 매개 편집 효율은 3% 미만으로 나타난 것을 확인하였다. 이러한 결과로부터 프라임 편집기를 마우스 모델 제작에 이용하기 위해서는 편집 효율의 개선이 필요할 것으로 판단하였다.

[0148] 실시예 3. dsgrRNA에 의한 프라임 편집기의 편집 효율 증진 확인

[0149] 본 발명자들은 프라임 편집기의 편집 효율성을 개선시키기 위한 시도로써 proxy-CRISPR 아이디어에 기반하여 촉매적으로 불활성인 엔도뉴클레아제 (catalytically dead endonuclease) 대신 표적 부위의 염색질 구조를 풀기 위한 dsgrNAs를 이용하였다. dsgrRNA는 불활성화된 촉매 작용을 나타내면서 Cas 엔도뉴클레아제를 안내하고 표적 부위에 결합하는 14~15 nt 길이의 가이드 RNA이다. 이에, 본 발명자들은 프라임 편집기가 다음과 같은 두 가지 역할을 할 것이라고 가정하였다. 한가지는 표적 부위에서 pegRNA와 함께 프라임 편집 기능을 하는 것이고, 다른 하나는 dsgrRNA와 표적 부위에 인접한 염색질을 조절하는 것이다. 본 발명자들은 pegRNA의 스페이서 (spacer)로부터 7-62 뉴클레오타이드 위치 범위에서 *Igf2* 및 *Adamts20* 표적 부위에 인접한 proximal dsgrNAs를 설계하였다. 이후 편집 효율을 확인하기 위해 *Igf2* 및 *Adamts20* 부위에서 다양한 pegRNA 길이에 proximal dsgrRNA를 적용하였으며, 흥미롭게도 proximal dsgrRNA를 이용한 PE3의 편집 효율은 대부분의 그룹에서 향상된 것으로 나타났다. 따라서 본 발명자들은 *Igf2* 및 *Adamts20* 표적에 대하여 각각 가장 높은 효율을 나타낸 PBS9-RT14 pegRNA 및 PBS11-RT13 pegRNA를 선택하여 후속 실험을 진행하였다.

[0150] 다음으로, 편집 효율이 높은 dsgrRNA를 선별하기 위해 *Igf2*, *Adamts20*, *Casp1* (4bp 결실), *Hoxd13* (G에서 T로 치환), *Angpt1* (CGG에서 TGA로 치환) 및 *Ksr2* (TGAT 삽입) 유전자에 대한 추가적인 proximal dsgrRNA를 설계하고 테스트하고자 하였다. 이를 위해, PE, pegRNA, nsgRNA 및 각각의 proximal dsgrRNA를 암호화하는 플라스미드를 전기천공법을 통해 NIH/3T3 및 C2C12 세포로 형질감염시켰으며, 이후 표적 심층 염기서열분석을 실시하였다.

[0151] 그 결과, 도 3에서 볼 수 있는 바와 같이 proximal dsgrRNA가 PE3에 비해 대부분의 표적에서 편집 효율성을 선택적으로 개선시켰음을 확인하였다. 이러한 결과로부터, dsgrRNA 적용에 따른 유전자 편집 효율은 proximal dsgrRNA의 위치에 따라 다르며, 효과적인 표적 돌연변이를 유도하기 위해서는 각 표적 및 세포 유형별로 최적의 dsgrRNA를 위한 스크리닝 프로세스가 필요하다는 것을 알 수 있었다.

[0153] 실시예 4. CMP가 결합된 프라임 편집기의 편집 효율 증진 확인

[0154] 본 발명자들은 프라임 편집기의 편집 효율성을 개선시키기 위한 다른 시도로서, 염색질 조절 펩타이드 (chromatin-modulating peptides, CMP)인 고-이동성 그룹 뉴클레오솜 결합 도메인 1 (high-mobility group nucleosome binding domain 1, HN1) 및 히스톤 H1 중앙 구상 도메인 (histone H1 central globular domain, H1G)을 사용해 프라임 편집기를 조작하였다. 구체적으로, 프라임 편집기의 Cas9 니카아제 (Cas9 nickase, nCas9)와 역전사효소가 융합된 단백질에 상기 HN1 및 H1G를 추가하여 2가지 버전으로 제조하였다. 도 4a에 도시된 바와 같이 nCas9의 N-말단 쪽에 HN1 및 C-말단 쪽에 H1G를 결합시킨 구성의 융합단백질을 CMP-PE-V1으로 명명하였고, nCas9의 N-말단 쪽에 HN1, nCas9의 C-말단 쪽에 조작된 M-MLV RT 및 상기 RT의 C-말단에 H1G가 결합되도록 조작한 융합단백질을 CMP-PE-V2로 명명하였다. 또한 상기 CMP-PE-V1 및 CMP-PE-V2의 아미노산 서열을 도 4b 및 도 4c에 각각 나타내었다.

[0155] 본 발명자들은 상기 CMP-PE3-V1 (pegRNA/nsgrRNA와 CMP-PE-V1) 또는 CMP-PE3-V2 (pegRNA/nsgrRNA와 CMP-PE-V2)를 두 가지 마우스 세포주에 각각 전달하고 CMP를 결합시키지 않은 PE3를 세포에 도입한 경우와 편집 효율을 비교하였다. 그 결과, 도 4b에 나타난 바와 같이 대부분의 표적 부위에서 CMP-PE3-V1의 경우 PE3에 비해 훨씬 더 편집 효율이 높게 나타난 것을 확인하였다. 특히, CMP-PE3-V1에 의한 편집 효율은 NIH/3T3 세포에서 *Igf2*의 경우 2.55배, *Adamts20*의 경우 3.92배 더 높게 나타났다. 또한 인간세포인 HEK293T 세포에서 HEK3 서열을 표적으로 향상된 프라임 편집기와 dead sgRNA를 적용하여 편집 효율을 확인하였다. 그 결과를 도 4e에 나타내었다. PE3에 비해 향상된 프라임 편집 방법인 CMP-PE-V1또는 CMP-PE-V2에서 편집 효율이 향상되는 것을 확인할 수 있었다. 또한, dead sgRNA를 향상된 프라임 편집기에 함께 적용하였을 때 PE3나 향상된 편집기만을 처리했을 때 보다 향상된 효율을 보이는 것을 확인할 수 있었다(CMP-PE3-V2 + dsgrRNA(-11)). 이러한 결과는 염색질 조절 펩타이드 HN1 및 H1G를 이용하여 조작된 프라임 편집기가 편집 효율성을 유의하게 향상시킬 수 있음을 보여주는 것이다.

[0157] 실시예 5. dsgrRNA 및 CMP가 적용된 프라임 편집기의 편집 효율 상승효과 확인

[0158] 나아가 상기 실시예 3 및 4의 결과를 바탕으로 본 발명자들은 CMP-PE3-V1 및 dsgrRNA (CMP-PE3-V1 + dsgrRNA)를 마우스 세포에 함께 전달하여 편집 효율에 대한 시너지 효과가 나타나는지 여부를 분석하였다.

[0159] 그 결과, 도 5에서 볼 수 있는 바와 같이 CMP-PE3-V1 및 dsgrRNA를 마우스 세포에 함께 전달한 경우 (CMP-PE3-V1 + dsgrRNA), PE3 대조군에 비해 *Igf2*, *Adamts20* 및 *Hoxd13* 표적 부위에서 각각 편집 효율이 최대 4.20배, 5.11배 및 3.56배 향상된 것으로 나타났으며, 이러한 시너지 효과는 세포주와 표적에 따라 상이한 수준으로 나타났다. 또한, 상기 결과로부터 CMP-PE3만을 세포로 도입시킨 경우 (CMP-PE3-V1) 실험한 모든 표적 및 두 가지 세포 유형에서 프라임 편집기의 효율이 유의하게 향상된 반면, dsgrRNA만을 세포로 도입한 경우 (PE3 + dsgrRNA) 또는 CMP-PE3-V1 및 dsgrRNA를 함께 도입한 경우 (CMP-PE3 + dsgrRNA)에는 편집 효율이 표적 부위 및 세포 유형에 따라 달라질 수 있음을 알 수 있었다.

[0160] 나아가, 본 발명자들은 진보된 프라임 편집기 시스템을 미세주입을 통해 마우스 배아에 주입하여 표적화 돌연변이 생성을 유도하고 그 효율을 분석하고자 하였다. 구체적으로, 설계된 마우스 표적 중에서 원치 않는 돌연변이가 상대적으로 낮은 *Igf2* 표적 부위를 선택하였다. 그 결과, 도 6a에서 볼 수 있는 바와 같이 CMP-PE3-V1가 주입된 배아 및 CMP-PE3-V1 + dsgrRNA가 주입된 배아에서 *Igf2* 표적에 대하여 상당히 높은 수준의 편집 효율을 보였다. 특히 CMP-PE3-V1 + dsgrRNA의 경우, *Igf2* 유전자의 표적 부위에서 원하는 돌연변이가 22개 배아 중 21개 (95%)에서 관찰되어 CMP-PE3-V1 + dsgrRNA가 PE3 및 PE3+dsgrRNA에 비해 매우 높은 편집 효율을 나타내는 것을 확인하였다.

[0161] 이에 더하여, 마우스 배아에서 *Adamts20*, *Hoxd13*, *Angpt1*, *Ksr2*, 및 *Ar* 유전자에서 편집 효율성을 검증하기 위해 CMP-PE-V1 및 proximal dsgrRNA를 사용하여 테스트하였다. 표적화 심층 시퀀싱 수행 결과, 도 6b에 나타난 바와 같이 PE3를 주입한 경우 비해 CMP-PE-V1 및 proximal dsgrRNA를 주입한 배아에서 *Ksr2* 및 *Ar* 표적을 제외하고 *Adamts20*, *Hoxd13*, 및 *Angpt1* 표적에서 편집 효율이 개선되었음을 확인하였다. 종합적으로, 이러한 결과는 마우스 배아에서 proximal dsgrRNA 및 염색질 조절 펩타이드를 이용하여 편집 효율이 증진된 프라임 편집이 가능함을 시사하는 것이다.

[0163] 실시예 6. dsgrRNA 및 CMP에 의한 표적 부위 염색질 접근성 향상 확인

[0164] 본 발명자들은 CMP-PE-V1과 dsgrRNA가 표적 부위의 염색질 구조를 풀어 염색질 접근성을 향상시킬 수 있는지 여부를 확인하기 위해, DNaseI digestion 분석 및 qPCR을 수행하여 각 표적 부위의 염색질 상태를 확인하였다. 이때, 음성 대조군 (닫힌 염색질)으로 Chr 3 : 71,026,628-71,026,685 (마우스 게놈 build mm9)에 위치한 유전자를 사용하였고, 양성 대조군 (열린 염색질)으로는 *Col6a1* 유전자를 사용하였다. 실험 결과, 도 7a 및 도 7b에 나타난 바와 같이 NIH/3T3 세포주에서 *Igf2*, *Adamts20* 및 *Hoxd13*은 상대적으로 닫힌 상태의 염색질 (closed chromatin)로 나타났으며, C2C12에서는 *Igf2*를 제외한 모든 표적이 열린 상태의 염색질인 것으로 분석되었다. 이러한 결과는 두 세포주에서 표적 서열이 동일하더라도 염색질 구조의 상태가 다르다는 것을 시사한다.

[0165] 나아가 닫힌 염색질 구조를 갖는 것으로 확인된 대표적 유전자인 *Igf2*의 표적 부위에서 CMP-PE-V1 또는 dsgrRNA가 염색질 상태를 변화시킬 수 있는지 분석하였다. 그 결과, 도 7c에서 볼 수 있는 바와 같이 PE3와 비교할 때 CMP-PE-V1, dsgrRNA 또는 CMP-PE-V1 + dsgrRNA에 의해 닫힌 상태의 염색질 구조가 열린 상태로 점진적으로 바뀌

는 것을 확인하였다. 이러한 결과는 CMP-PE-V1 또는 dsgrRNA를 이용하면 닫힌 상태의 염색질 구조를 풀고 염색질 접근성을 개선하여 프라임 편집의 효율성을 향상시킬 수 있음을 입증하는 직접적인 증거이다.

[0167] **실시예 7. 표적 돌연변이가 유발된 마우스 모델 제작 및 오프-타겟 효과 분석**

[0168] 다음으로, 본 발명자들은 원치 않는 돌연변이의 발생 빈도가 상대적으로 낮은 PBS9-RT14 및 dsgrRNA +7을 사용하여 미세주입을 통해 마우스 배아에서 도 8a에 도시된 바와 같은 *Igf2*의 표적 돌연변이가 생성을 유도하였으며, 상기 마우스 배아를 대리모에게 이식하였다. 이후 상기 대리모로부터 태어난 새끼를 관찰 및 분석한 결과, 도 8b에서 볼 수 있는 바와 같이 최대 47%의 편집 빈도 (10마리 중 2마리)로 *Igf2* 유전자좌에서 G에서 C로의 치환 및 TA 삽입이 일어난 것을 확인하였다. 또한, 상기와 같은 *Igf2* 돌연변이가 다음 세대로 전달되는지 여부를 분석한 결과, 도 8c와 같이 *Igf2* 돌연변이 마우스로부터 태어난 F1 한배 새끼 9마리 중에서 7마리가 동일한 돌연변이를 갖는 것을 통해 표적 돌연변이의 생식세포를 통한 다음 세대로의 전달이 가능함을 알 수 있었다.

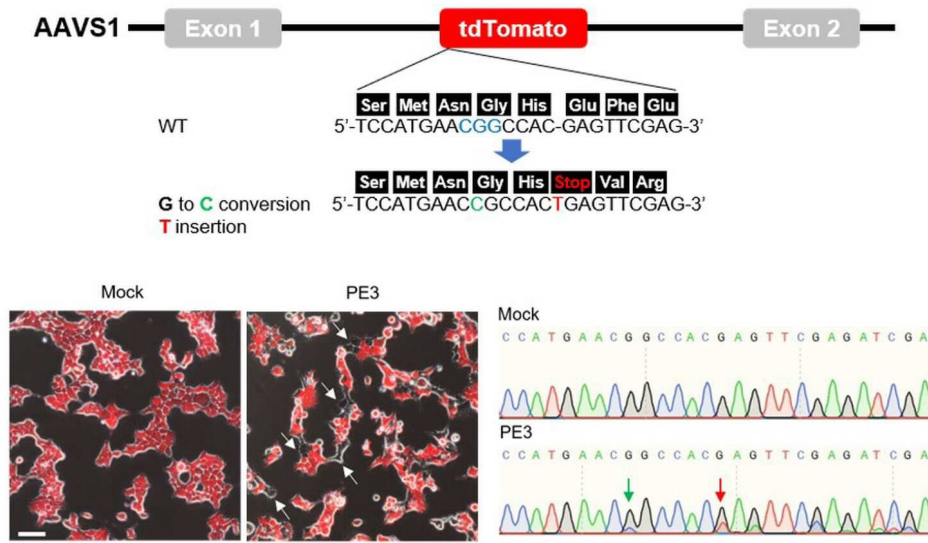
[0169] 본 발명자들은 프라임 편집기의 오프-타겟 효과 (off-target effect)를 평가하기 위해, Cas-OFFinder를 사용하여 마우스 게놈에서 각각 최대 3개의 뉴클레오티드 불일치가 있는 *Igf2* 타겟의 pegRNA 및 nsgRNA에 의한 잠재적인 오프 타겟 부위를 확인하였다. 그 결과, 도 9a에 나타난 바와 같이 야생형(Wild type)과 비교하였을 때 잠재적인 오프-타겟 돌연변이는 검출되지 않았다. 또한, 제작된 *Igf2* 돌연변이 마우스에서 오프-타겟 효과를 확인하기 위해 전장 유전체 염기서열분석(WGS)을 실시하였다. 그 결과, 도 9b 및 도 9c에 나타난 바와 같이 nsgRNA의 단일 오프-타겟 부위를 발견되었으나, 게놈 DNA를 이용한 Sanger 시퀀싱을 통해 상기 부위가 위양성임을 확인하였다.

[0170] 마지막으로, *Igf2* 돌연변이 마우스의 표현형을 확인하기 위해, *Igf2*^{p+/m-} 수컷 (F1)을 야생형 암컷 마우스와 교배시켰다. 그 결과, 도 10a 및 도 10b에서 볼 수 있는 바와 같이 부계 대립 유전자로부터 유전된 *Igf2* 유전자의 돌연변이를 보유한 *Igf2*^{p-/m+} 마우스는 원하는 돌연변이 유전자형과 일치하는 왜소증 표현형을 나타냈다. 이러한 결과는 개선된 본 발명에 따른 프라임 편집 시스템이 마우스 모델 제작에 효과적으로 적용될 수 있음을 시사하는 것이다.

[0172] 상기 진술한 본 발명의 설명은 예시를 위한 것이며, 본 발명이 속하는 기술분야의 통상의 지식을 가진 자는 본 발명의 기술적 사상이나 필수적인 특징을 변경하지 않고서 다른 구체적인 형태로 쉽게 변형이 가능하다는 것을 이해할 수 있을 것이다. 그러므로 이상에서 기술한 실시예들은 모든 면에서 예시적인 것이며 한정적이 아닌 것으로 이해해야만 한다.

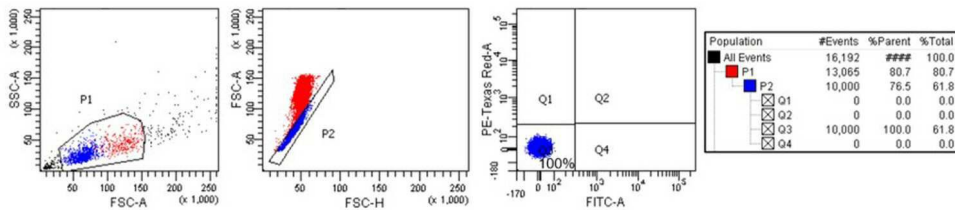
도면

도면1a

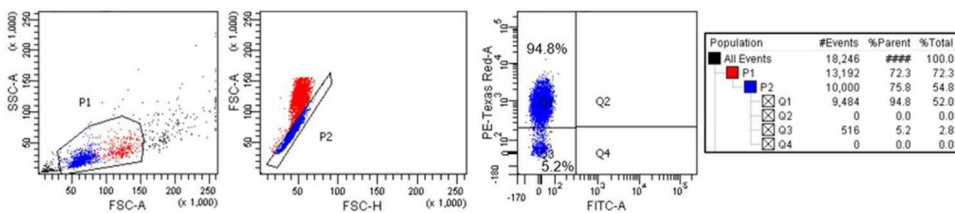


도면1b

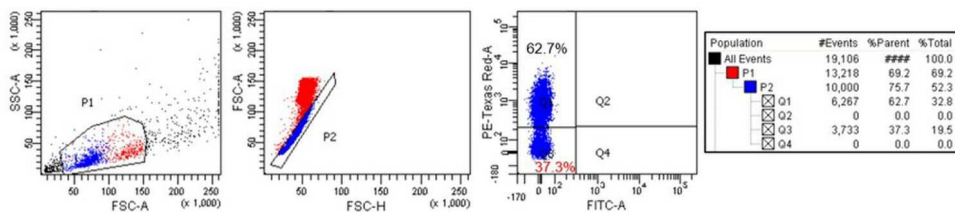
HEK293T (Negative control)



tdTomato expressing HEK293T (Positive control)

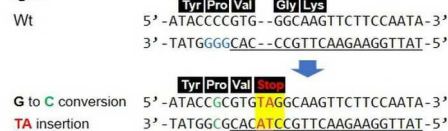


tdTomato expressing HEK293T + PE3

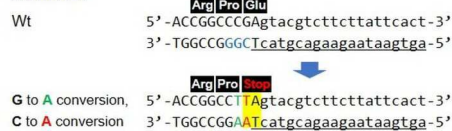


도면2a

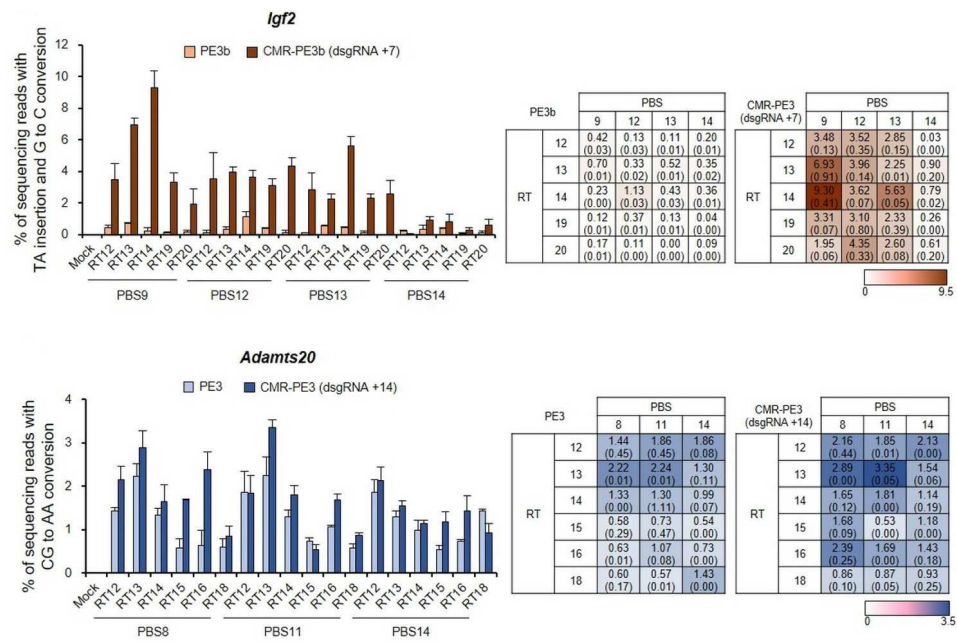
Igf2



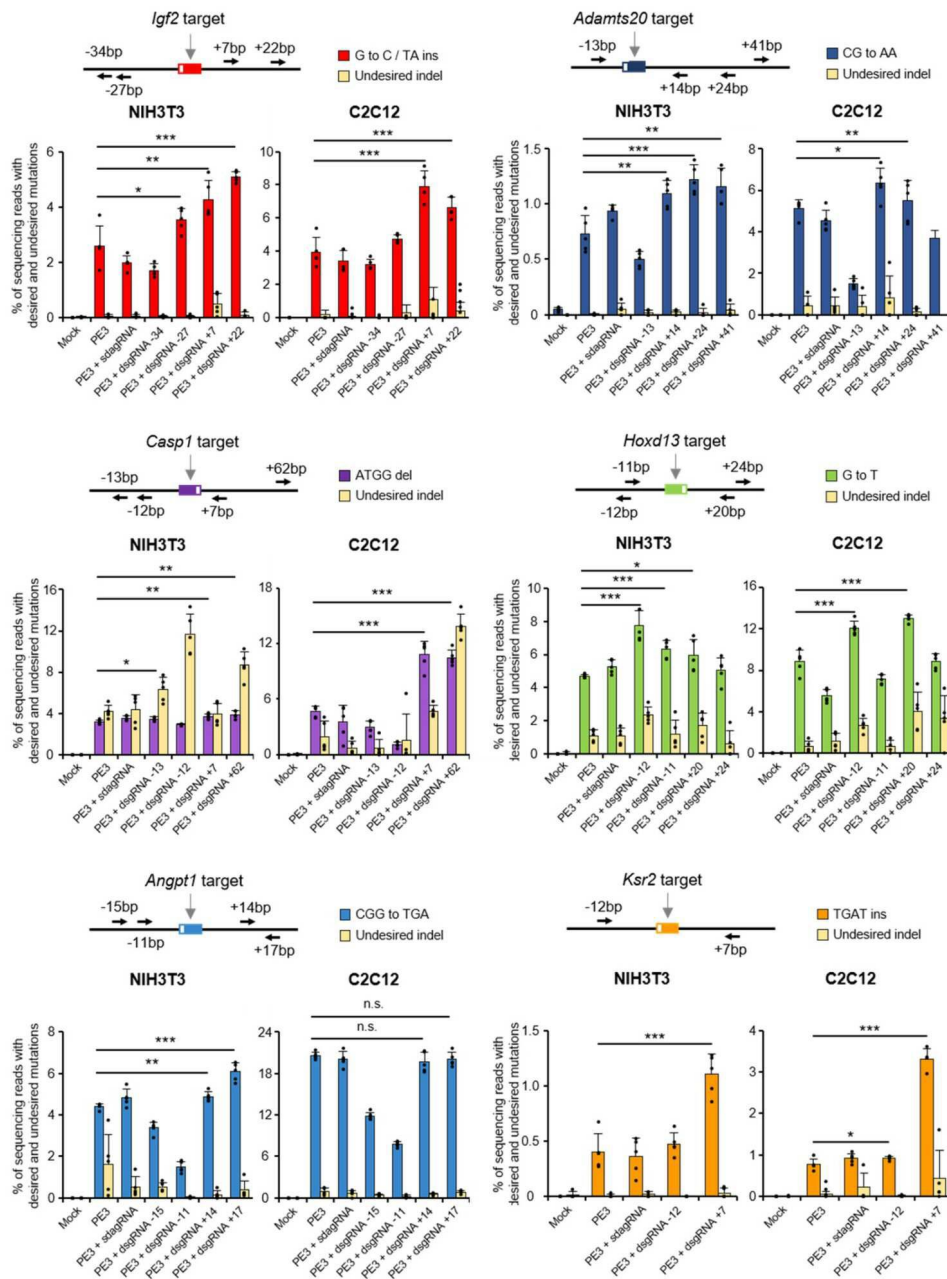
Adams20



도면2b



도면3



도면4a



도면4b

Purple: Bipartite NLS

Green: HN1

Yellow: Linker

Grey : nCas9

Light green: H1G

Blue: M-MLV RT

CMP-PE-V1

MKRTADGSEFESPKKKRKVM^{Purple}PKRKVSSAEGAAKEPKRRSARLSAKPPAKVEAKPKKA
AAKDKSSDKKVQTKGKRGAKGKQAEVANQETKEDLPAENGETKTEESPASDEAGEKEA
KSD^{Yellow}TSGSDKKYSIGLDIGTNSVGVAVITDEYKVPSSKFKVLGNTDRHSIKKNLIGALLFD
SGETAETRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLSEESFLVEEDKKHE
RHPIFGNIVDEVAYHEKYPTIYHLRKKLV DSTDKADLRLIYLAALAHMIKFRGHFLIEGDL
NPDNSDVKLFIQLVQTYNQLFEEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKK
NGLFGNLIALSLGLTPNFKNFDLAEDAKLQLSKD TYDDLDNLLAQIGDQYADFLAA
KNLSAAILSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQ
SKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQI
HLGELHAILRRQEDFYPLKDNREKIEKILTRIPYYVGPLARGNSRFAWMTRKSEETITP
WNFEVVDKGASAQSFIERMTNFDKNLPNEKVLPHSLLYEYFTVYNELTKVKYVTEG
MRKPAFLSGEQKKAIVDLLFKTRNKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLG
TYHDLKIKDKDFLDNEENEDILEDIVLTLLFEDREMIEERLKYAHLFDDKVMKQLK
RRRYTGWGRLSRKLINGIRDQSGKTILDFLKSDFANRNFMQLIHDDSLTFKEDIQKAQ
VSGQGDSLHEHIANLAGSPAICKGILQTVKVVDLVKVMGRHKPENIVIEAMARENQTTQ
KGQKNSRERMKRIEIEGKELGSQILKEHPVENTQLQNEKLYLYLQNGRDMYVDQELDI
NRLSDYDVDAIVPQSFLKDDSIDNKVLTSDKNRKGSDNVPSEEVVKMKMKNYWRQLLN
AKLITQRKFDNLTKAERGGSELDAKAGFIKQVLVETRQITKHVAQILDSRMNTKYDEND
KLIREVKVITLKSILVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTAIKKYPKLESE
FVYGDYKVVYDVRKMIKSEQIEGKATAKYFFYSNIMNFFKTEITLANGEIRKRPIETNG
ETGEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDW
DPKKYGGFDSPTVAYSVLVAKVEKGKSKKLKSVKELLGITIMERSSEKPNPIDFLEAKG
YKEVKKDLIILPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKL
KGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKHHRDKPIREQ
AENIIHLFTLTNLGAPAAFYFDTTIDRKRYTSTKEVLDATLIHQSIITGLYETRIDLSQLGG
D^{Light green}LEGGGSSSTDHPKYSMDMVAIAIQA^{Green}EKNRAGSSRSIQKYIKSHYKVENADSQIKLSIK
RLVTTGVLKQTKGVGASGSFRLAKSDEPSGGSSGGSSGSETPGTSESATPESGGSSGGSS
TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIPLKATSTPVSI
KQYPMSEQEARLGKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRVPVDLREVNK
RVEDIHPTVPNPYNLLSGLPPSHQWYTVL DLKD AFFCLRLHPTSQPLFAFEWRDPPEMGI
GQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDL LLAATSELD CQQG
TRALLQTLGNLGYRASAKKAQICQKQVKYLYLLKEGQRWLTEARKETVMGQPTPKT
PRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQKAYQEIKQALLTAPA
LGLPDLTKPFELFVDEKQGYAKGVLTKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVA
AIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQF
GPVVALNPATLLPLPEEGLQHNC LDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQE
GQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLVYTDSTRYAFAT
AHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIHCPGHQKGHSAEARGNR
MADQAARKAAITETPDTSTLLIENSSPSGGSK^{Purple}RTADGSEFEPKKRKV

도면4c

Purple: Bipartite NLS

Green: HN1

Yellow: Linker

Grey : nCas9

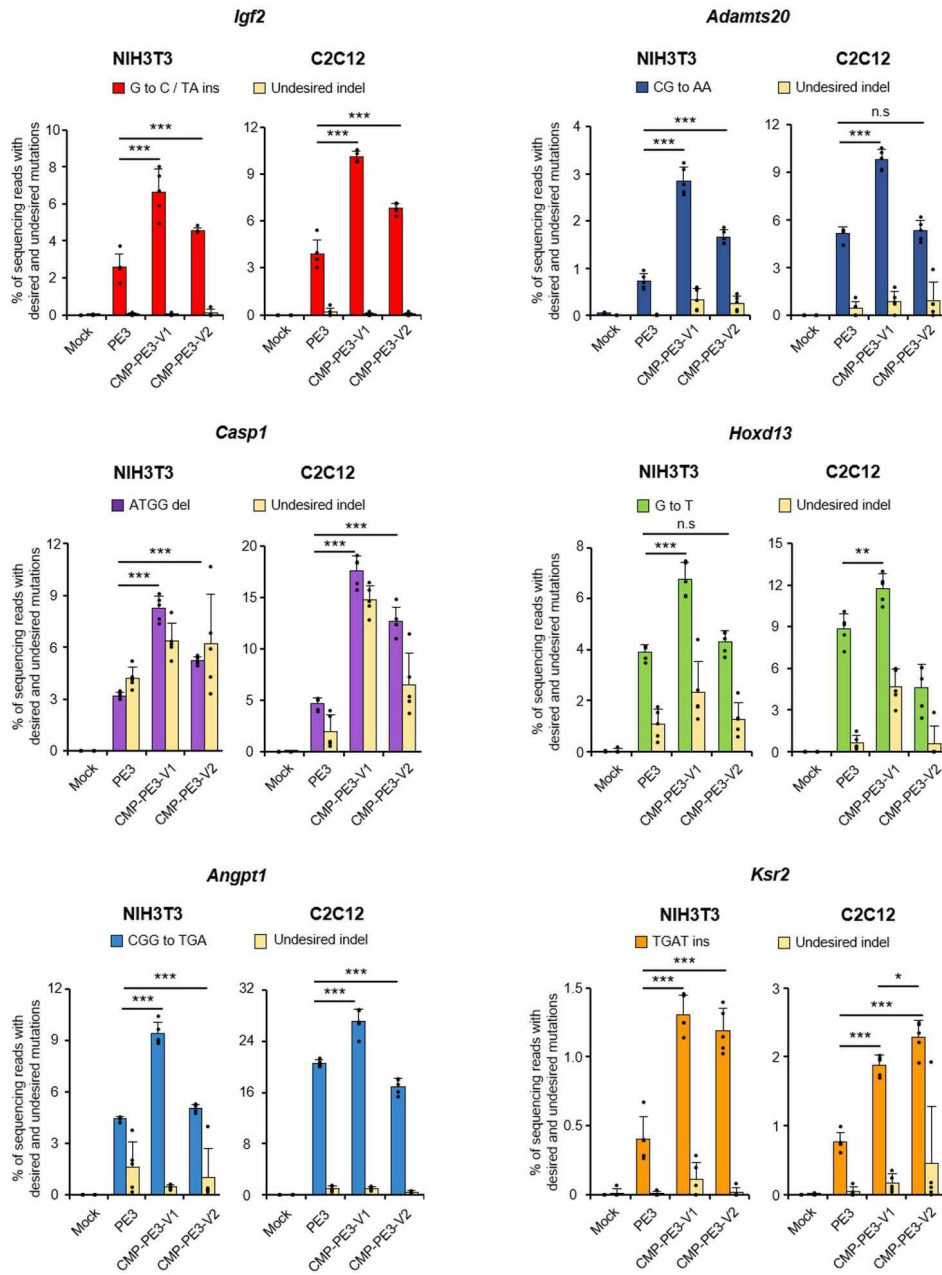
Light green: H1G

Blue: M-MLV RT

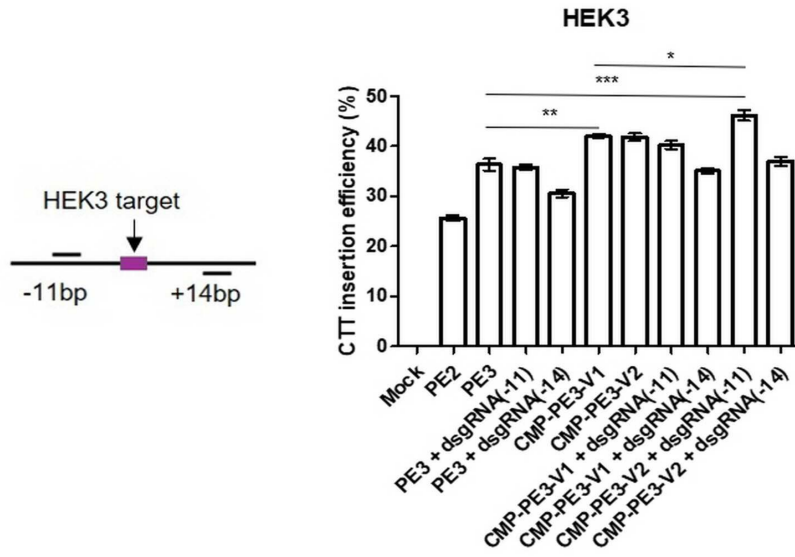
CMP-PE-V2

MKRTADGSEFESPKKKRKVMPKRVSSAEGAAKEPKRRSARLSAKPPAKVEAKPKKA
AAKDKSSDKKVQTKGKRGAKGKQAEVANQETKEDLPAENGETKTEESPASDEAGEKEA
KSDTSGGDKKYSIGLDIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFD
SGETAETRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLSEESFLVEEDKKHE
RHPIFGNIVDEVAYHEKYPTTYHLRKKLVDSDDKADLRILIYLAALAHMIKFRGHFLIEGDL
NPDNSDVKLFQIVQTYNQLEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKK
NGLFGNLIASLGLTPNFKSNFDLAEDAKLQLSKDITYDDDLNLLAQIGDQYADFLAA
KNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQ
SKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQI
HLGELHAILRRQEDFYFPLKDNREKIEKILTRIPYYVGPLARGNSRFAWMTRKSEETITP
WNFEVVVDKGASQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEG
MRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLG
TYHDLKKIHKDKDFLDNEENEDILEDIVLTLTLFEDREMIERLKYAHLFDDKVMKQLK
RRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDFANRNFQMQLIHDDSLTFKEDIKKAQ
VSGQGDSLHEHIANLAGSPAIIKGLQTVKVVDELVKVMGRHKPENIVIAMARENQTTQ
KGQKNSRERMKRIEKGELGSQLKEHPVENTQLQNEKLYLYLQNGRDMYVDQELDI
NRLSDYDVIDAIVPQSLKDDSIDNKVLTSDKNRGKSDNVPSEEVVKKMKNYWRQLLN
AKLITQRKFDNLTKAERGGLSELDKAGFIKRLVETRQITKHVAQILDSRMNTKYDEND
KLIREVKVITLKSCLVSDFRKDFQFYKREINNYHHAHDAYLNAVVGTAIIKKYKPLESE
FVYGDYKVYDVRKMIKSEQEI GKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNG
ETGEIVWDKGRDFATVRKVL SMPQVNIVKKEVQTGGFSKESILPKRNSDKLIARKKDW
DPKKYGGFDSPTVAYSVLVAKVEKGKSKKLKSVKELLGITIMERSSEFKNPIDFLEAKG
YKEVKKDLIIKLPKYSLEFENGRKRLASAGELQKGNELALPSKYVNFLYLASHYEKL
KGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKHDKPIREQ
AENIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSTGLYETRIDLSQLGG
DSGGSSGGSSGSETPGTSESATPSSGGSSGGSSTLNIEDEYRLHETSKEPDVSLGSTWLS
DFPQAWAETGGMGLAVRQAPLIPLKATSTPVSIIKQYPMSEARLGIIKPHIQRLLDQGILV
PCQSPWNTPLLPVKKPGTNDYRPVQDLREVNRVEDIHPTVNPYNLLSGLPPSHQWYT
VLDLKDAAFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPGGFKNSPTLFNEALHRD
LADFRIQHPDLILLQYVDDLLAATSELDCCQGTALLQTLGNLGYRASAKKAQICQKQ
VKYLYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIPGFAEMAAPL
YPLTKPGTLFNWGPQQKAYQEI KQALLTAPALGLPDLTKPFELFVDEKQGYAKGVL TQ
KLGPWRRPVAYLSKKLPVAAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEA
LVKQPPDRWLSNARMTHYQALLDTRVQFGPVVALNPATLLPLPEEGLQHNCLDILAE
AHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETETVWAKALPAGTSAQ
RAELIALTQALKMAEGKKLVNVTDSRYAFATAHIIHGEIYRRRGWLTSEGKEIKNKDEILA
LLKALFLPKRLSIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSPSG
GSLEGGSSSTDHPKYSDMIIVAAIAEKNRAGSSRSIQKYIKSHYKVGENADSQIKLSI
KRLVTTGVLLKQTKGVGASGSFRLAKSDEPKRTADGSEFEPKKKKRKV

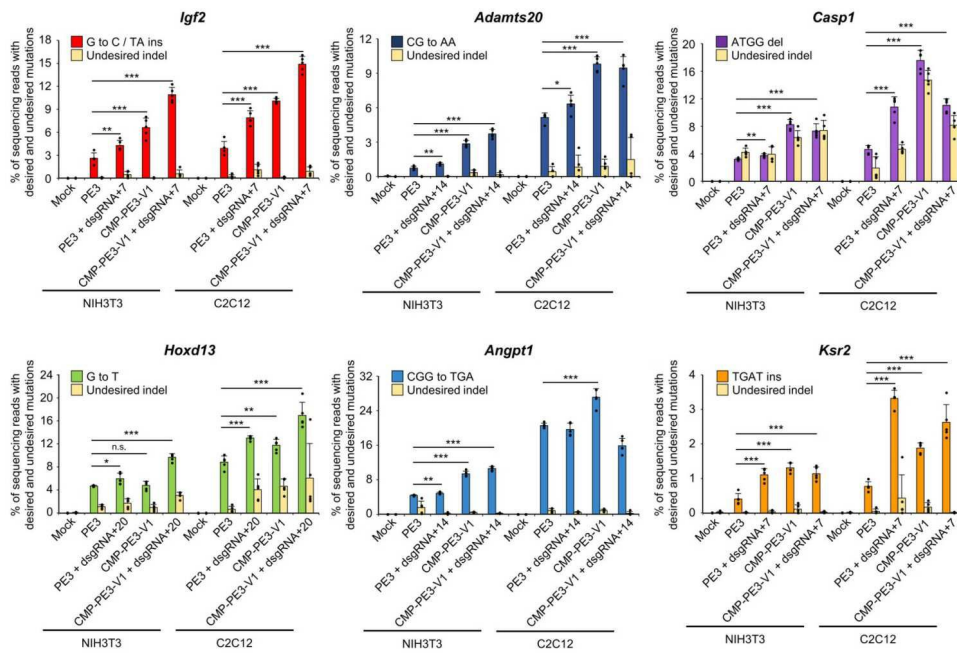
도면4d



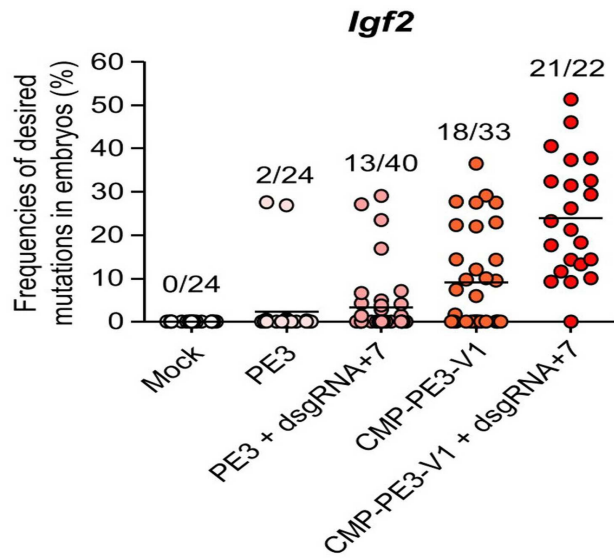
도면4e



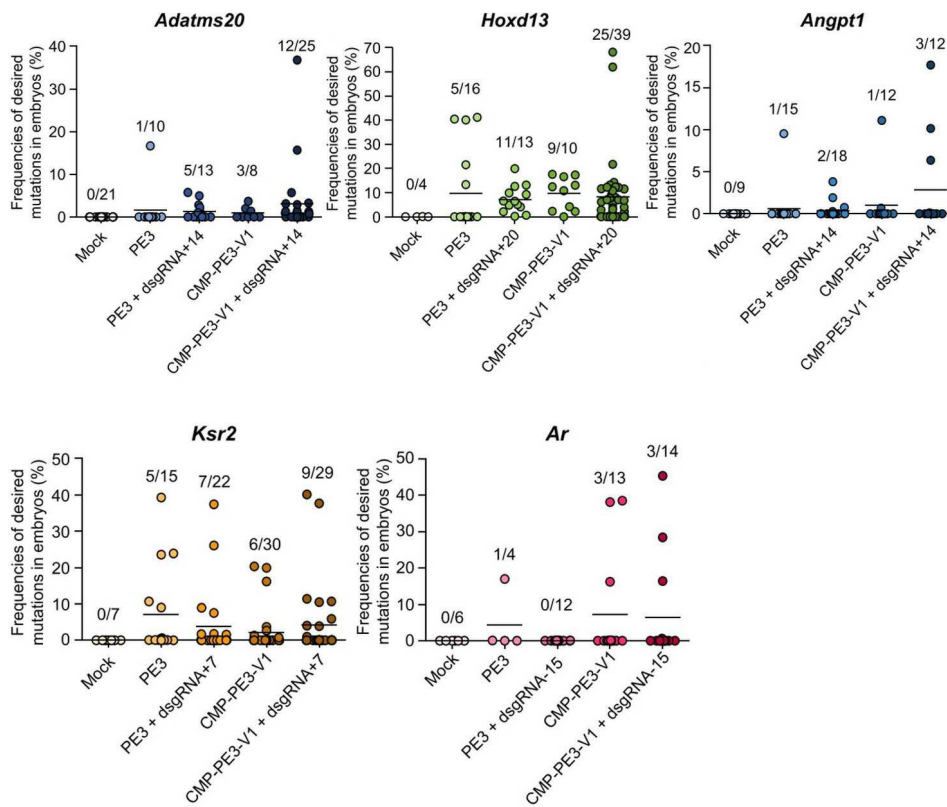
도면5



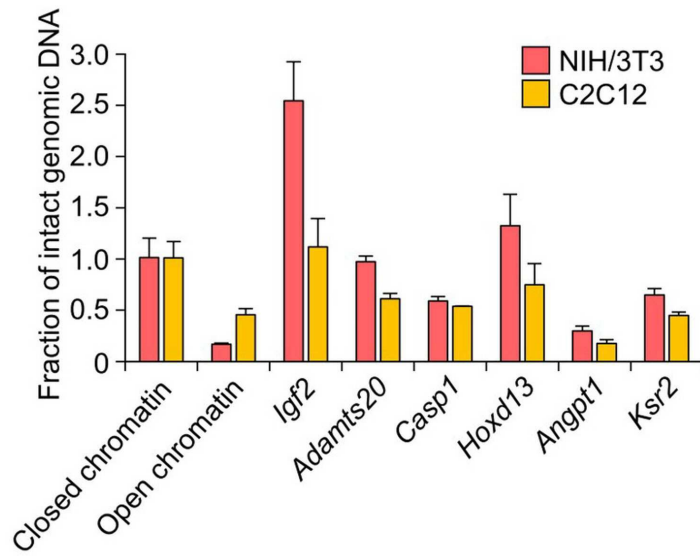
도면6a



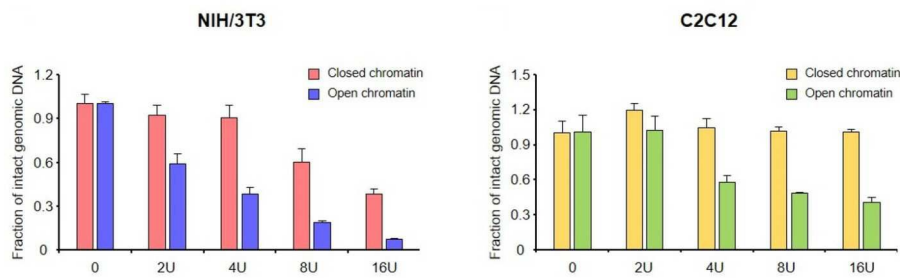
도면6b



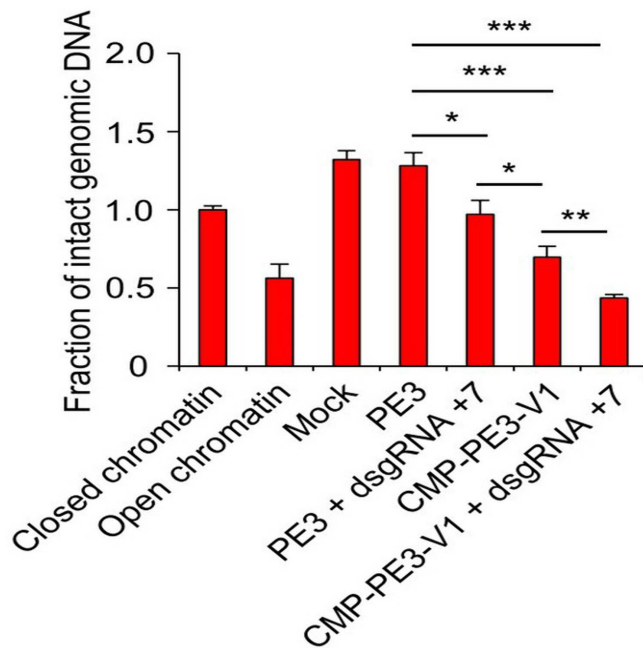
도면7a



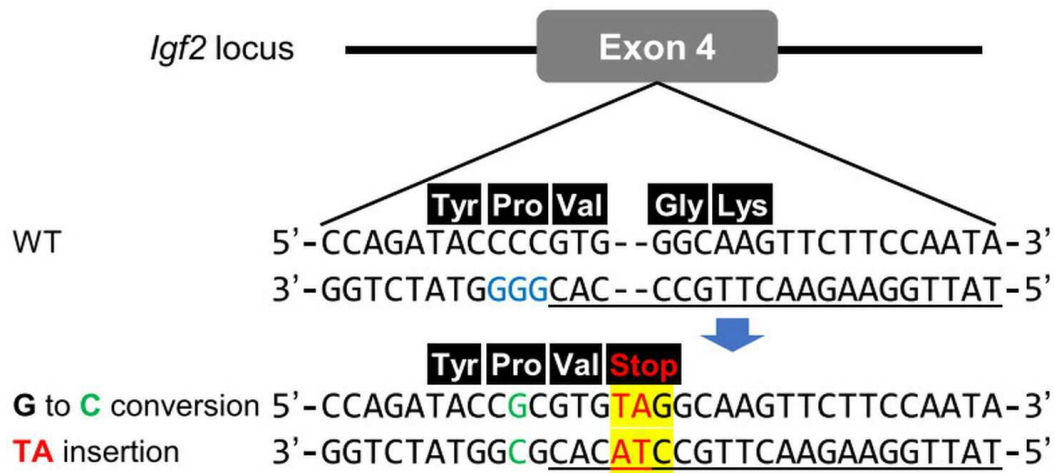
도면7b



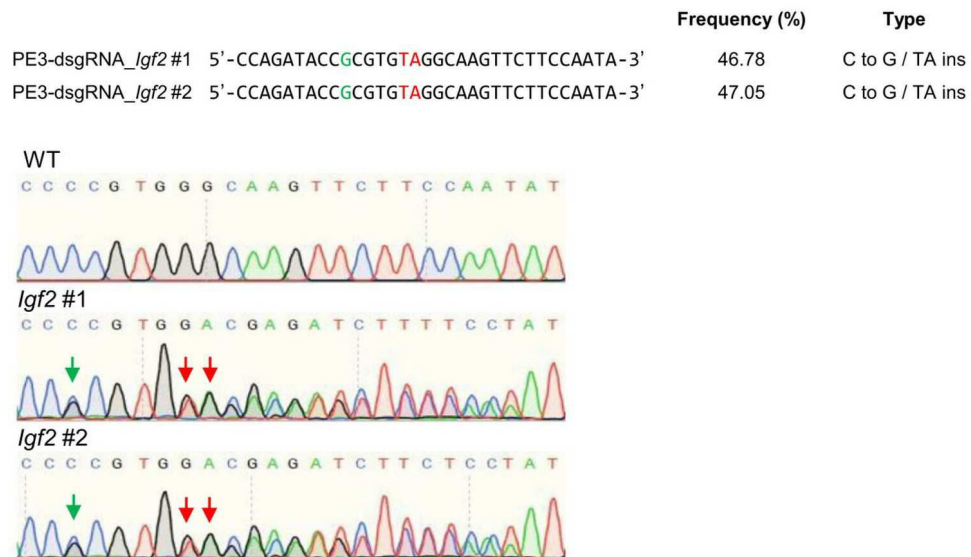
도면7c



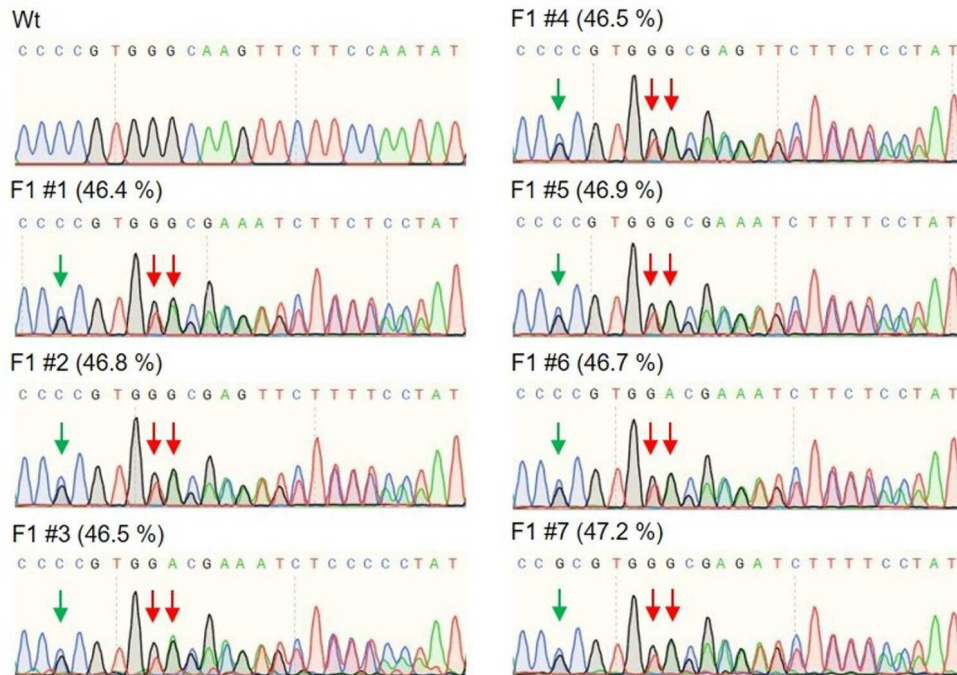
도면8a



도면8b



도면8c



Igf2

Wt	5' - CCAGATACCCCGTG - - GGCAAGTCTTCCAATA - 3'		
	3' - GGTCTATGGGCAC - - CCGTTCAAGAAGTTAT - 5'		
		Frequency (%)	Type
	5' - CCAGATACCCCGTG - - GGCAAGTCTTCCAATA - 3'	48.13	WT
F1_#1	5' - CCAGATACCCCGTGTAAGGCAAGTCTTCCAATA - 3'	46.39	C to G / TA ins
	5' - CCAGATACCCCGTG - - GGCAAGTCTTCCAATA - 3'	47.69	WT
F1_#2	5' - CCAGATACCCCGTGTAAGGCAAGTCTTCCAATA - 3'	46.79	C to G / TA ins
	5' - CCAGATACCCCGTG - - GGCAAGTCTTCCAATA - 3'	48.04	WT
F1_#3	5' - CCAGATACCCCGTGTAAGGCAAGTCTTCCAATA - 3'	46.47	C to G / TA ins
	5' - CCAGATACCCCGTG - - GGCAAGTCTTCCAATA - 3'	48.06	WT
F1_#4	5' - CCAGATACCCCGTGTAAGGCAAGTCTTCCAATA - 3'	46.46	C to G / TA ins
	5' - CCAGATACCCCGTG - - GGCAAGTCTTCCAATA - 3'	48.31	WT
F1_#5	5' - CCAGATACCCCGTGTAAGGCAAGTCTTCCAATA - 3'	46.87	C to G / TA ins
	5' - CCAGATACCCCGTG - - GGCAAGTCTTCCAATA - 3'	48.45	WT
F1_#6	5' - CCAGATACCCCGTGTAAGGCAAGTCTTCCAATA - 3'	46.65	C to G / TA ins
	5' - CCAGATACCCCGTG - - GGCAAGTCTTCCAATA - 3'	47.26	WT
F1_#7	5' - CCAGATACCCCGTGTAAGGCAAGTCTTCCAATA - 3'	47.17	C to G / TA ins

도면9a

Whole genome sequencing analysis

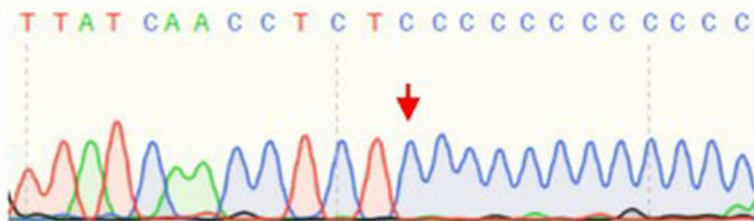
	<i>Igf2</i> WT	<i>Igf2</i> MUT
Total variants (SNPs + Indels)	75030 + 99307	79772 + 107911
SNPs	75030	79772
SNPs after excluding dbSNP 142	43767	47876
unique SNPs	6515	10624
On-target (pegRNA)		
Up to 2bp bulge + 5bp mismatches with NRG PAM (5723)	0	0
Up to 7bp mismatches with NRG PAM (3202)	0	0
On-target (nicking sgRNA)		
Up to 2bp bulge + 5bp mismatches with NRG PAM (5815)	0	1
Up to 7bp mismatches with NRG PAM (3212)	0	0
On-target (dead sgRNA)		
Up to 3bp mismatches with NRG PAM (1224)	0	0

도면9b

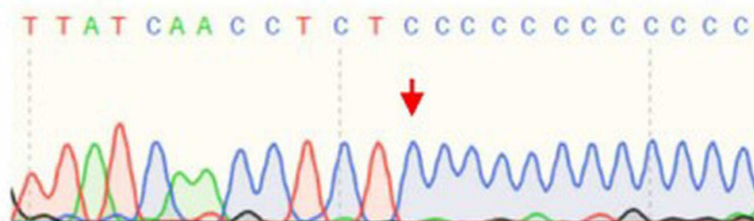
WT 5' -TTATCAACCTCTCCCCCCCCCCC-3'

OT 5' -TTATCAACCTCT-CCCCCCCCCCC-3' (-1 bp)

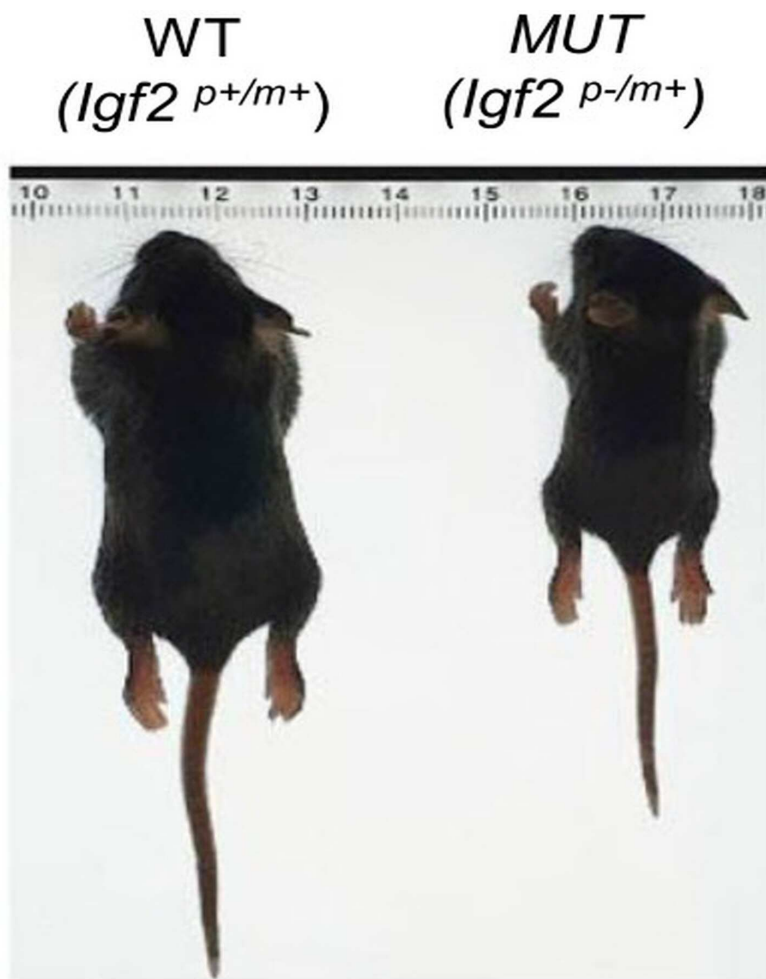
WT



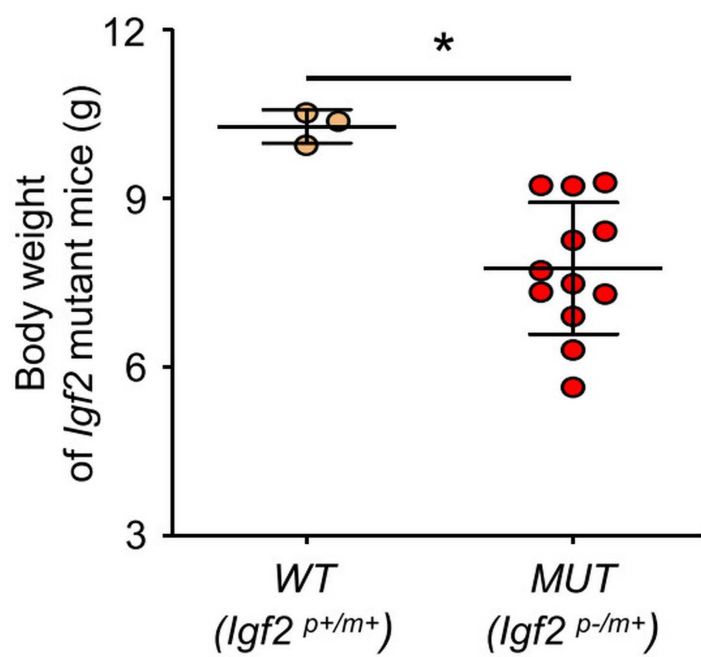
MUT



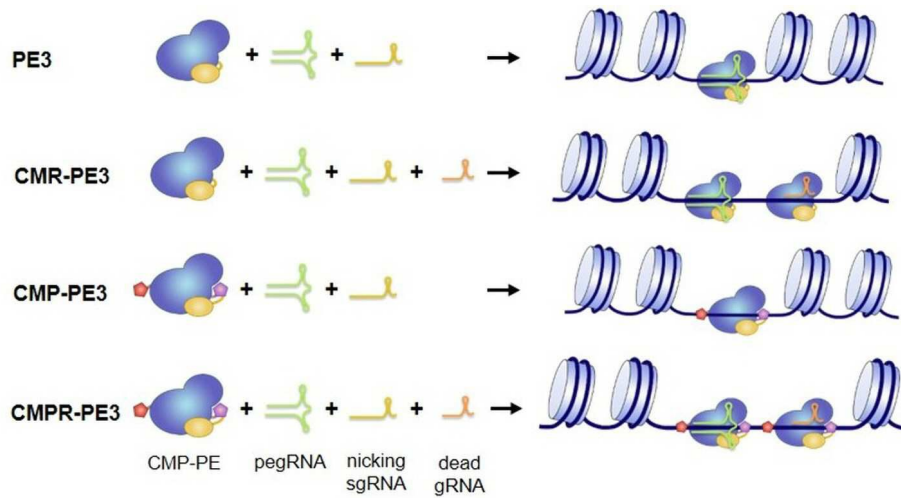
도면10a



도면10b



도면11



서열 목록

- <110> Korea University Research and Business Foundation
Seoul National University R&DB Foundation
Korea Mouse Phenotyping Center
- <120> Prime editing-based gene editing composition with improved editing efficiency and use thereof
- <130> DHP20-561-P1
- <150> KR 10-2020-0123432
- <151> 2020-09-24
- <160> 73
- <170> KoPatent In 3.0
- <210> 1
- <211> 2308
- <212> PRT
- <213> Artificial Sequence
- <220><223> CMP-PE-V1
- <400> 1

Met Lys Arg Thr Ala Asp Gly Ser Glu Phe Glu Ser Pro Lys Lys Lys

1 5 10 15
Arg Lys Val Met Pro Lys Arg Lys Val Ser Ser Ala Glu Gly Ala Ala
20 25 30
Lys Glu Glu Pro Lys Arg Arg Ser Ala Arg Leu Ser Ala Lys Pro Pro

35 40 45
 Ala Lys Val Glu Ala Lys Pro Lys Lys Ala Ala Ala Lys Asp Lys Ser
 50 55 60
 Ser Asp Lys Lys Val Gln Thr Lys Gly Lys Arg Gly Ala Lys Gly Lys
 65 70 75 80

 Gln Ala Glu Val Ala Asn Gln Glu Thr Lys Glu Asp Leu Pro Ala Glu
 85 90 95
 Asn Gly Glu Thr Lys Thr Glu Glu Ser Pro Ala Ser Asp Glu Ala Gly
 100 105 110
 Glu Lys Glu Ala Lys Ser Asp Thr Gly Ser Gly Asp Lys Lys Tyr Ser
 115 120 125
 Ile Gly Leu Asp Ile Gly Thr Asn Ser Val Gly Trp Ala Val Ile Thr
 130 135 140
 Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe Lys Val Leu Gly Asn Thr

 145 150 155 160
 Asp Arg His Ser Ile Lys Lys Asn Leu Ile Gly Ala Leu Leu Phe Asp
 165 170 175
 Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu Lys Arg Thr Ala Arg Arg
 180 185 190
 Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys Tyr Leu Gln Glu Ile Phe
 195 200 205
 Ser Asn Glu Met Ala Lys Val Asp Asp Ser Phe Phe His Arg Leu Glu
 210 215 220

 Glu Ser Phe Leu Val Glu Glu Asp Lys Lys His Glu Arg His Pro Ile
 225 230 235 240
 Phe Gly Asn Ile Val Asp Glu Val Ala Tyr His Glu Lys Tyr Pro Thr
 245 250 255
 Ile Tyr His Leu Arg Lys Lys Leu Val Asp Ser Thr Asp Lys Ala Asp
 260 265 270
 Leu Arg Leu Ile Tyr Leu Ala Leu Ala His Met Ile Lys Phe Arg Gly
 275 280 285

His Phe Leu Ile Glu Gly Asp Leu Asn Pro Asp Asn Ser Asp Val Asp

290

295

300

Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr Asn Gln Leu Phe Glu Glu

305

310

315

320

Asn Pro Ile Asn Ala Ser Gly Val Asp Ala Lys Ala Ile Leu Ser Ala

325

330

335

Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn Leu Ile Ala Gln Leu Pro

340

345

350

Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn Leu Ile Ala Leu Ser Leu

355

360

365

Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe Asp Leu Ala Glu Asp Ala

370

375

380

Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp Asp Asp Leu Asp Asn Leu

385

390

395

400

Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp Leu Phe Leu Ala Ala Lys

405

410

415

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

420

425

430

Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser Met Ile Lys Arg Tyr Asp

435

440

445

Glu His His Gln Asp Leu Thr Leu Leu Lys Ala Leu Val Arg Gln Gln

450

455

460

Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe Asp Gln Ser Lys Asn Gly

465

470

475

480

Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser Gln Glu Glu Phe Tyr Lys

485

490

495

Phe Ile Lys Pro Ile Leu Glu Lys Met Asp Gly Thr Glu Glu Leu Leu

500

505

510

Val Lys Leu Asn Arg Glu Asp Leu Leu Arg Lys Gln Arg Thr Phe Asp

515

520

525

Asn Gly Ser Ile Pro His Gln Ile His Leu Gly Glu Leu His Ala Ile

530 535 540
 Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe Leu Lys Asp Asn Arg Glu
 545 550 555 560
 Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile Pro Tyr Tyr Val Gly Pro
 565 570 575
 Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp Met Thr Arg Lys Ser Glu

 580 585 590
 Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu Val Val Asp Lys Gly Ala
 595 600 605
 Ser Ala Gln Ser Phe Ile Glu Arg Met Thr Asn Phe Asp Lys Asn Leu
 610 615 620
 Pro Asn Glu Lys Val Leu Pro Lys His Ser Leu Leu Tyr Glu Tyr Phe
 625 630 635 640
 Thr Val Tyr Asn Glu Leu Thr Lys Val Lys Tyr Val Thr Glu Gly Met
 645 650 655

 Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln Lys Lys Ala Ile Val Asp
 660 665 670
 Leu Leu Phe Lys Thr Asn Arg Lys Val Thr Val Lys Gln Leu Lys Glu
 675 680 685
 Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp Ser Val Glu Ile Ser Gly
 690 695 700
 Val Glu Asp Arg Phe Asn Ala Ser Leu Gly Thr Tyr His Asp Leu Leu
 705 710 715 720
 Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp Asn Glu Glu Asn Glu Asp

 725 730 735
 Ile Leu Glu Asp Ile Val Leu Thr Leu Thr Leu Phe Glu Asp Arg Glu
 740 745 750
 Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala His Leu Phe Asp Asp Lys
 755 760 765
 Val Met Lys Gln Leu Lys Arg Arg Arg Tyr Thr Gly Trp Gly Arg Leu
 770 775 780
 Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp Lys Gln Ser Gly Lys Thr

785					790								795								800						
Ile	Leu	Asp	Phe	Leu	Lys	Ser	Asp	Gly	Phe	Ala	Asn	Arg	Asn	Phe	Met												
				805								810								815							
Gln	Leu	Ile	His	Asp	Asp	Ser	Leu	Thr	Phe	Lys	Glu	Asp	Ile	Gln	Lys												
				820								825								830							
Ala	Gln	Val	Ser	Gly	Gln	Gly	Asp	Ser	Leu	His	Glu	His	Ile	Ala	Asn												
				835								840								845							
Leu	Ala	Gly	Ser	Pro	Ala	Ile	Lys	Lys	Gly	Ile	Leu	Gln	Thr	Val	Lys												
				850								855								860							
Val	Val	Asp	Glu	Leu	Val	Lys	Val	Met	Gly	Arg	His	Lys	Pro	Glu	Asn												
865								870								875								880			
Ile	Val	Ile	Glu	Met	Ala	Arg	Glu	Asn	Gln	Thr	Thr	Gln	Lys	Gly	Gln												
				885								890								895							
Lys	Asn	Ser	Arg	Glu	Arg	Met	Lys	Arg	Ile	Glu	Glu	Gly	Ile	Lys	Glu												
				900								905								910							
Leu	Gly	Ser	Gln	Ile	Leu	Lys	Glu	His	Pro	Val	Glu	Asn	Thr	Gln	Leu												
				915								920								925							
Gln	Asn	Glu	Lys	Leu	Tyr	Leu	Tyr	Tyr	Leu	Gln	Asn	Gly	Arg	Asp	Met												
				930								935								940							
Tyr	Val	Asp	Gln	Glu	Leu	Asp	Ile	Asn	Arg	Leu	Ser	Asp	Tyr	Asp	Val												
945								950								955								960			
Asp	Ala	Ile	Val	Pro	Gln	Ser	Phe	Leu	Lys	Asp	Asp	Ser	Ile	Asp	Asn												
				965								970								975							
Lys	Val	Leu	Thr	Arg	Ser	Asp	Lys	Asn	Arg	Gly	Lys	Ser	Asp	Asn	Val												
				980								985								990							
Pro	Ser	Glu	Glu	Val	Val	Lys	Lys	Met	Lys	Asn	Tyr	Trp	Arg	Gln	Leu												
				995								1000								1005							
Leu	Asn	Ala	Lys	Leu	Ile	Thr	Gln	Arg	Lys	Phe	Asp	Asn	Leu	Thr	Lys												
1010								1015								1020											
Ala	Glu	Arg	Gly	Gly	Leu	Ser	Glu	Leu	Asp	Lys	Ala	Gly	Phe	Ile	Lys												
1025								1030								1035								1040			

Arg Gln Leu Val Glu Thr Arg Gln Ile Thr Lys His Val Ala Gln Ile
1045 1050 1055

Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp Glu Asn Asp Lys Leu Ile
1060 1065 1070

Arg Glu Val Lys Val Ile Thr Leu Lys Ser Lys Leu Val Ser Asp Phe
1075 1080 1085

Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg Glu Ile Asn Asn Tyr His
1090 1095 1100

His Ala His Asp Ala Tyr Leu Asn Ala Val Val Gly Thr Ala Leu Ile
1105 1110 1115 1120

Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe Val Tyr Gly Asp Tyr Lys
1125 1130 1135

Val Tyr Asp Val Arg Lys Met Ile Ala Lys Ser Glu Gln Glu Ile Gly
1140 1145 1150

Lys Ala Thr Ala Lys Tyr Phe Phe Tyr Ser Asn Ile Met Asn Phe Phe
1155 1160 1165

Lys Thr Glu Ile Thr Leu Ala Asn Gly Glu Ile Arg Lys Arg Pro Leu
1170 1175 1180

Ile Glu Thr Asn Gly Glu Thr Gly Glu Ile Val Trp Asp Lys Gly Arg
1185 1190 1195 1200

Asp Phe Ala Thr Val Arg Lys Val Leu Ser Met Pro Gln Val Asn Ile
1205 1210 1215

Val Lys Lys Thr Glu Val Gln Thr Gly Gly Phe Ser Lys Glu Ser Ile
1220 1225 1230

Leu Pro Lys Arg Asn Ser Asp Lys Leu Ile Ala Arg Lys Lys Asp Trp
1235 1240 1245

Asp Pro Lys Lys Tyr Gly Gly Phe Asp Ser Pro Thr Val Ala Tyr Ser
1250 1255 1260

Val Leu Val Val Ala Lys Val Glu Lys Gly Lys Ser Lys Lys Leu Lys
1265 1270 1275 1280

Ser Val Lys Glu Leu Leu Gly Ile Thr Ile Met Glu Arg Ser Ser Phe
1285 1290 1295

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

1300 1305 1310

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

1315 1320 1325

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

1330 1335 1340

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

1345 1350 1355 1360

Ala Ser His Tyr Glu Lys Leu Lys Gly Ser Pro Glu Asp Asn Glu Gln

1365 1370 1375

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

1380 1385 1390

Glu Gln Ile Ser Glu Phe Ser Lys Arg Val Ile Leu Ala Asp Ala Asn

1395 1400 1405

Leu Asp Lys Val Leu Ser Ala Tyr Asn Lys His Arg Asp Lys Pro Ile

1410 1415 1420

Arg Glu Gln Ala Glu Asn Ile Ile His Leu Phe Thr Leu Thr Asn Leu

1425 1430 1435 1440

Gly Ala Pro Ala Ala Phe Lys Tyr Phe Asp Thr Thr Ile Asp Arg Lys

1445 1450 1455

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

1460 1465 1470

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

1475 1480 1485

Gly Asp Leu Glu Gly Gly Gly Gly Ser Ser Thr Asp His Pro Lys Tyr

1490 1495 1500

Ser Asp Met Ile Val Ala Ala Ile Gln Ala Glu Lys Asn Arg Ala Gly

1505 1510 1515 1520

Ser Ser Arg Gln Ser Ile Gln Lys Tyr Ile Lys Ser His Tyr Lys Val

1525 1530 1535

Gly Glu Asn Ala Asp Ser Gln Ile Lys Leu Ser Ile Lys Arg Leu Val

1540	1545	1550	
Thr Thr Gly Val Leu Lys Gln Thr Lys Gly Val Gly Ala Ser Gly Ser			
1555	1560	1565	
Phe Arg Leu Ala Lys Ser Asp Glu Pro Ser Gly Gly Ser Ser Gly Gly			
1570	1575	1580	
Ser Ser Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu			
1585	1590	1595	1600
Ser Ser Gly Gly Ser Ser Gly Gly Ser Ser Thr Leu Asn Ile Glu Asp			
1605	1610	1615	
Glu Tyr Arg Leu His Glu Thr Ser Lys Glu Pro Asp Val Ser Leu Gly			
1620	1625	1630	
Ser Thr Trp Leu Ser Asp Phe Pro Gln Ala Trp Ala Glu Thr Gly Gly			
1635	1640	1645	
Met Gly Leu Ala Val Arg Gln Ala Pro Leu Ile Ile Pro Leu Lys Ala			
1650	1655	1660	
Thr Ser Thr Pro Val Ser Ile Lys Gln Tyr Pro Met Ser Gln Glu Ala			
1665	1670	1675	1680
Arg Leu Gly Ile Lys Pro His Ile Gln Arg Leu Leu Asp Gln Gly Ile			
1685	1690	1695	
Leu Val Pro Cys Gln Ser Pro Trp Asn Thr Pro Leu Leu Pro Val Lys			
1700	1705	1710	
Lys Pro Gly Thr Asn Asp Tyr Arg Pro Val Gln Asp Leu Arg Glu Val			
1715	1720	1725	
Asn Lys Arg Val Glu Asp Ile His Pro Thr Val Pro Asn Pro Tyr Asn			
1730	1735	1740	
Leu Leu Ser Gly Leu Pro Pro Ser His Gln Trp Tyr Thr Val Leu Asp			
1745	1750	1755	1760
Leu Lys Asp Ala Phe Phe Cys Leu Arg Leu His Pro Thr Ser Gln Pro			
1765	1770	1775	
Leu Phe Ala Phe Glu Trp Arg Asp Pro Glu Met Gly Ile Ser Gly Gln			
1780	1785	1790	
Leu Thr Trp Thr Arg Leu Pro Gln Gly Phe Lys Asn Ser Pro Thr Leu			

1795 1800 1805
 Phe Asn Glu Ala Leu His Arg Asp Leu Ala Asp Phe Arg Ile Gln His
 1810 1815 1820
 Pro Asp Leu Ile Leu Leu Gln Tyr Val Asp Asp Leu Leu Leu Ala Ala
 1825 1830 1835 1840
 Thr Ser Glu Leu Asp Cys Gln Gln Gly Thr Arg Ala Leu Leu Gln Thr
 1845 1850 1855
 Leu Gly Asn Leu Gly Tyr Arg Ala Ser Ala Lys Lys Ala Gln Ile Cys
 1860 1865 1870
 Gln Lys Gln Val Lys Tyr Leu Gly Tyr Leu Leu Lys Glu Gly Gln Arg

 1875 1880 1885
 Trp Leu Thr Glu Ala Arg Lys Glu Thr Val Met Gly Gln Pro Thr Pro
 1890 1895 1900
 Lys Thr Pro Arg Gln Leu Arg Glu Phe Leu Gly Lys Ala Gly Phe Cys
 1905 1910 1915 1920
 Arg Leu Phe Ile Pro Gly Phe Ala Glu Met Ala Ala Pro Leu Tyr Pro
 1925 1930 1935
 Leu Thr Lys Pro Gly Thr Leu Phe Asn Trp Gly Pro Asp Gln Gln Lys
 1940 1945 1950

 Ala Tyr Gln Glu Ile Lys Gln Ala Leu Leu Thr Ala Pro Ala Leu Gly
 1955 1960 1965
 Leu Pro Asp Leu Thr Lys Pro Phe Glu Leu Phe Val Asp Glu Lys Gln
 1970 1975 1980
 Gly Tyr Ala Lys Gly Val Leu Thr Gln Lys Leu Gly Pro Trp Arg Arg
 1985 1990 1995 2000
 Pro Val Ala Tyr Leu Ser Lys Lys Leu Asp Pro Val Ala Ala Gly Trp
 2005 2010 2015
 Pro Pro Cys Leu Arg Met Val Ala Ala Ile Ala Val Leu Thr Lys Asp

 2020 2025 2030
 Ala Gly Lys Leu Thr Met Gly Gln Pro Leu Val Ile Leu Ala Pro His
 2035 2040 2045

Ala Val Glu Ala Leu Val Lys Gln Pro Pro Asp Arg Trp Leu Ser Asn
2050 2055 2060

Ala Arg Met Thr His Tyr Gln Ala Leu Leu Leu Asp Thr Asp Arg Val
2065 2070 2075 2080

Gln Phe Gly Pro Val Val Ala Leu Asn Pro Ala Thr Leu Leu Pro Leu
2085 2090 2095

Pro Glu Glu Gly Leu Gln His Asn Cys Leu Asp Ile Leu Ala Glu Ala
2100 2105 2110

His Gly Thr Arg Pro Asp Leu Thr Asp Gln Pro Leu Pro Asp Ala Asp
2115 2120 2125

His Thr Trp Tyr Thr Asp Gly Ser Ser Leu Leu Gln Glu Gly Gln Arg
2130 2135 2140

Lys Ala Gly Ala Ala Val Thr Thr Glu Thr Glu Val Ile Trp Ala Lys
2145 2150 2155 2160

Ala Leu Pro Ala Gly Thr Ser Ala Gln Arg Ala Glu Leu Ile Ala Leu
2165 2170 2175

Thr Gln Ala Leu Lys Met Ala Glu Gly Lys Lys Leu Asn Val Tyr Thr
2180 2185 2190

Asp Ser Arg Tyr Ala Phe Ala Thr Ala His Ile His Gly Glu Ile Tyr
2195 2200 2205

Arg Arg Arg Gly Trp Leu Thr Ser Glu Gly Lys Glu Ile Lys Asn Lys
2210 2215 2220

Asp Glu Ile Leu Ala Leu Leu Lys Ala Leu Phe Leu Pro Lys Arg Leu
2225 2230 2235 2240

Ser Ile Ile His Cys Pro Gly His Gln Lys Gly His Ser Ala Glu Ala
2245 2250 2255

Arg Gly Asn Arg Met Ala Asp Gln Ala Ala Arg Lys Ala Ala Ile Thr
2260 2265 2270

Glu Thr Pro Asp Thr Ser Thr Leu Leu Ile Glu Asn Ser Ser Pro Ser
2275 2280 2285

Gly Gly Ser Lys Arg Thr Ala Asp Gly Ser Glu Phe Glu Pro Lys Lys
2290 2295 2300

Lys Arg Lys Val

2305

<210> 2

<211> 2308

<212> PRT

<213> Artificial Sequence

<220><223> CMP-PE-V2

<400> 2

Met Lys Arg Thr Ala Asp Gly Ser Glu Phe Glu Ser Pro Lys Lys Lys

1 5 10 15

Arg Lys Val Met Pro Lys Arg Lys Val Ser Ser Ala Glu Gly Ala Ala

20 25 30

Lys Glu Glu Pro Lys Arg Arg Ser Ala Arg Leu Ser Ala Lys Pro Pro

35 40 45

Ala Lys Val Glu Ala Lys Pro Lys Lys Ala Ala Ala Lys Asp Lys Ser

50 55 60

Ser Asp Lys Lys Val Gln Thr Lys Gly Lys Arg Gly Ala Lys Gly Lys

65 70 75 80

Gln Ala Glu Val Ala Asn Gln Glu Thr Lys Glu Asp Leu Pro Ala Glu

85 90 95

Asn Gly Glu Thr Lys Thr Glu Glu Ser Pro Ala Ser Asp Glu Ala Gly

100 105 110

Glu Lys Glu Ala Lys Ser Asp Thr Gly Ser Gly Asp Lys Lys Tyr Ser

115 120 125

Ile Gly Leu Asp Ile Gly Thr Asn Ser Val Gly Trp Ala Val Ile Thr

130 135 140

Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe Lys Val Leu Gly Asn Thr

145 150 155 160

Asp Arg His Ser Ile Lys Lys Asn Leu Ile Gly Ala Leu Leu Phe Asp

165 170 175

Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu Lys Arg Thr Ala Arg Arg

180 185 190

Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys Tyr Leu Gln Glu Ile Phe

195 200 205

Ser Asn Glu Met Ala Lys Val Asp Asp Ser Phe Phe His Arg Leu Glu

210 215 220

Glu Ser Phe Leu Val Glu Glu Asp Lys Lys His Glu Arg His Pro Ile

225 230 235 240

Phe Gly Asn Ile Val Asp Glu Val Ala Tyr His Glu Lys Tyr Pro Thr

245 250 255

Ile Tyr His Leu Arg Lys Lys Leu Val Asp Ser Thr Asp Lys Ala Asp

260 265 270

Leu Arg Leu Ile Tyr Leu Ala Leu Ala His Met Ile Lys Phe Arg Gly

275 280 285

His Phe Leu Ile Glu Gly Asp Leu Asn Pro Asp Asn Ser Asp Val Asp

290 295 300

Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr Asn Gln Leu Phe Glu Glu

305 310 315 320

Asn Pro Ile Asn Ala Ser Gly Val Asp Ala Lys Ala Ile Leu Ser Ala

325 330 335

Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn Leu Ile Ala Gln Leu Pro

340 345 350

Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn Leu Ile Ala Leu Ser Leu

355 360 365

Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe Asp Leu Ala Glu Asp Ala

370 375 380

Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp Asp Asp Leu Asp Asn Leu

385 390 395 400

Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp Leu Phe Leu Ala Ala Lys

405 410 415

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

420 425 430

Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser Met Ile Lys Arg Tyr Asp

435 440 445
 Glu His His Gln Asp Leu Thr Leu Leu Lys Ala Leu Val Arg Gln Gln
 450 455 460
 Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe Asp Gln Ser Lys Asn Gly
 465 470 475 480
 Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser Gln Glu Glu Phe Tyr Lys

 485 490 495
 Phe Ile Lys Pro Ile Leu Glu Lys Met Asp Gly Thr Glu Glu Leu Leu
 500 505 510
 Val Lys Leu Asn Arg Glu Asp Leu Leu Arg Lys Gln Arg Thr Phe Asp
 515 520 525
 Asn Gly Ser Ile Pro His Gln Ile His Leu Gly Glu Leu His Ala Ile
 530 535 540
 Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe Leu Lys Asp Asn Arg Glu
 545 550 555 560

 Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile Pro Tyr Tyr Val Gly Pro
 565 570 575
 Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp Met Thr Arg Lys Ser Glu
 580 585 590
 Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu Val Val Asp Lys Gly Ala
 595 600 605
 Ser Ala Gln Ser Phe Ile Glu Arg Met Thr Asn Phe Asp Lys Asn Leu
 610 615 620
 Pro Asn Glu Lys Val Leu Pro Lys His Ser Leu Leu Tyr Glu Tyr Phe

 625 630 635 640
 Thr Val Tyr Asn Glu Leu Thr Lys Val Lys Tyr Val Thr Glu Gly Met
 645 650 655
 Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln Lys Lys Ala Ile Val Asp
 660 665 670
 Leu Leu Phe Lys Thr Asn Arg Lys Val Thr Val Lys Gln Leu Lys Glu
 675 680 685
 Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp Ser Val Glu Ile Ser Gly

690					695					700									
Val	Glu	Asp	Arg	Phe	Asn	Ala	Ser	Leu	Gly	Thr	Tyr	His	Asp	Leu	Leu				
705					710					715					720				
Lys	Ile	Ile	Lys	Asp	Lys	Asp	Phe	Leu	Asp	Asn	Glu	Glu	Asn	Glu	Asp				
725					730					735									
Ile	Leu	Glu	Asp	Ile	Val	Leu	Thr	Leu	Thr	Leu	Phe	Glu	Asp	Arg	Glu				
740					745					750									
Met	Ile	Glu	Glu	Arg	Leu	Lys	Thr	Tyr	Ala	His	Leu	Phe	Asp	Asp	Lys				
755					760					765									
Val	Met	Lys	Gln	Leu	Lys	Arg	Arg	Arg	Tyr	Thr	Gly	Trp	Gly	Arg	Leu				
770					775					780									
Ser	Arg	Lys	Leu	Ile	Asn	Gly	Ile	Arg	Asp	Lys	Gln	Ser	Gly	Lys	Thr				
785					790					795					800				
Ile	Leu	Asp	Phe	Leu	Lys	Ser	Asp	Gly	Phe	Ala	Asn	Arg	Asn	Phe	Met				
805					810					815									
Gln	Leu	Ile	His	Asp	Asp	Ser	Leu	Thr	Phe	Lys	Glu	Asp	Ile	Gln	Lys				
820					825					830									
Ala	Gln	Val	Ser	Gly	Gln	Gly	Asp	Ser	Leu	His	Glu	His	Ile	Ala	Asn				
835					840					845									
Leu	Ala	Gly	Ser	Pro	Ala	Ile	Lys	Lys	Gly	Ile	Leu	Gln	Thr	Val	Lys				
850					855					860									
Val	Val	Asp	Glu	Leu	Val	Lys	Val	Met	Gly	Arg	His	Lys	Pro	Glu	Asn				
865					870					875					880				
Ile	Val	Ile	Glu	Met	Ala	Arg	Glu	Asn	Gln	Thr	Thr	Gln	Lys	Gly	Gln				
885					890					895									
Lys	Asn	Ser	Arg	Glu	Arg	Met	Lys	Arg	Ile	Glu	Glu	Gly	Ile	Lys	Glu				
900					905					910									
Leu	Gly	Ser	Gln	Ile	Leu	Lys	Glu	His	Pro	Val	Glu	Asn	Thr	Gln	Leu				
915					920					925									
Gln	Asn	Glu	Lys	Leu	Tyr	Leu	Tyr	Tyr	Leu	Gln	Asn	Gly	Arg	Asp	Met				
930					935					940									

Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg Leu Ser Asp Tyr Asp Val
 945 950 955 960
 Asp Ala Ile Val Pro Gln Ser Phe Leu Lys Asp Asp Ser Ile Asp Asn
 965 970 975
 Lys Val Leu Thr Arg Ser Asp Lys Asn Arg Gly Lys Ser Asp Asn Val
 980 985 990

 Pro Ser Glu Glu Val Val Lys Lys Met Lys Asn Tyr Trp Arg Gln Leu
 995 1000 1005
 Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys Phe Asp Asn Leu Thr Lys
 1010 1015 1020
 Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp Lys Ala Gly Phe Ile Lys
 1025 1030 1035 1040
 Arg Gln Leu Val Glu Thr Arg Gln Ile Thr Lys His Val Ala Gln Ile
 1045 1050 1055
 Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp Glu Asn Asp Lys Leu Ile

 1060 1065 1070
 Arg Glu Val Lys Val Ile Thr Leu Lys Ser Lys Leu Val Ser Asp Phe
 1075 1080 1085
 Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg Glu Ile Asn Asn Tyr His
 1090 1095 1100
 His Ala His Asp Ala Tyr Leu Asn Ala Val Val Gly Thr Ala Leu Ile
 1105 1110 1115 1120
 Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe Val Tyr Gly Asp Tyr Lys
 1125 1130 1135

 Val Tyr Asp Val Arg Lys Met Ile Ala Lys Ser Glu Gln Glu Ile Gly
 1140 1145 1150
 Lys Ala Thr Ala Lys Tyr Phe Phe Tyr Ser Asn Ile Met Asn Phe Phe
 1155 1160 1165
 Lys Thr Glu Ile Thr Leu Ala Asn Gly Glu Ile Arg Lys Arg Pro Leu
 1170 1175 1180
 Ile Glu Thr Asn Gly Glu Thr Gly Glu Ile Val Trp Asp Lys Gly Arg
 1185 1190 1195 1200

Asp Phe Ala Thr Val Arg Lys Val Leu Ser Met Pro Gln Val Asn Ile

1205

1210

1215

Val Lys Lys Thr Glu Val Gln Thr Gly Gly Phe Ser Lys Glu Ser Ile

1220

1225

1230

Leu Pro Lys Arg Asn Ser Asp Lys Leu Ile Ala Arg Lys Lys Asp Trp

1235

1240

1245

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

1250

1255

1260

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

1265

1270

1275

1280

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

1285

1290

1295

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

1300

1305

1310

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

1315

1320

1325

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

1330

1335

1340

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

1345

1350

1355

1360

Ala Ser His Tyr Glu Lys Leu Lys Gly Ser Pro Glu Asp Asn Glu Gln

1365

1370

1375

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

1380

1385

1390

Glu Gln Ile Ser Glu Phe Ser Lys Arg Val Ile Leu Ala Asp Ala Asn

1395

1400

1405

Leu Asp Lys Val Leu Ser Ala Tyr Asn Lys His Arg Asp Lys Pro Ile

1410

1415

1420

Arg Glu Gln Ala Glu Asn Ile Ile His Leu Phe Thr Leu Thr Asn Leu

1425

1430

1435

1440

Gly Ala Pro Ala Ala Phe Lys Tyr Phe Asp Thr Thr Ile Asp Arg Lys

1445 1450 1455
 Arg Tyr Thr Ser Thr Lys Glu Val Leu Asp Ala Thr Leu Ile His Gln
 1460 1465 1470
 Ser Ile Thr Gly Leu Tyr Glu Thr Arg Ile Asp Leu Ser Gln Leu Gly
 1475 1480 1485
 Gly Asp Ser Gly Gly Ser Ser Gly Gly Ser Ser Gly Ser Glu Thr Pro

 1490 1495 1500
 Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Ser Gly Gly Ser Ser Gly
 1505 1510 1515 1520
 Gly Ser Ser Thr Leu Asn Ile Glu Asp Glu Tyr Arg Leu His Glu Thr
 1525 1530 1535
 Ser Lys Glu Pro Asp Val Ser Leu Gly Ser Thr Trp Leu Ser Asp Phe
 1540 1545 1550
 Pro Gln Ala Trp Ala Glu Thr Gly Gly Met Gly Leu Ala Val Arg Gln
 1555 1560 1565

 Ala Pro Leu Ile Ile Pro Leu Lys Ala Thr Ser Thr Pro Val Ser Ile
 1570 1575 1580
 Lys Gln Tyr Pro Met Ser Gln Glu Ala Arg Leu Gly Ile Lys Pro His
 1585 1590 1595 1600
 Ile Gln Arg Leu Leu Asp Gln Gly Ile Leu Val Pro Cys Gln Ser Pro
 1605 1610 1615
 Trp Asn Thr Pro Leu Leu Pro Val Lys Lys Pro Gly Thr Asn Asp Tyr
 1620 1625 1630
 Arg Pro Val Gln Asp Leu Arg Glu Val Asn Lys Arg Val Glu Asp Ile

 1635 1640 1645
 His Pro Thr Val Pro Asn Pro Tyr Asn Leu Leu Ser Gly Leu Pro Pro
 1650 1655 1660
 Ser His Gln Trp Tyr Thr Val Leu Asp Leu Lys Asp Ala Phe Phe Cys
 1665 1670 1675 1680
 Leu Arg Leu His Pro Thr Ser Gln Pro Leu Phe Ala Phe Glu Trp Arg
 1685 1690 1695
 Asp Pro Glu Met Gly Ile Ser Gly Gln Leu Thr Trp Thr Arg Leu Pro

1700 1705 1710
 Gln Gly Phe Lys Asn Ser Pro Thr Leu Phe Asn Glu Ala Leu His Arg
 1715 1720 1725
 Asp Leu Ala Asp Phe Arg Ile Gln His Pro Asp Leu Ile Leu Leu Gln
 1730 1735 1740
 Tyr Val Asp Asp Leu Leu Leu Ala Ala Thr Ser Glu Leu Asp Cys Gln
 1745 1750 1755 1760
 Gln Gly Thr Arg Ala Leu Leu Gln Thr Leu Gly Asn Leu Gly Tyr Arg
 1765 1770 1775
 Ala Ser Ala Lys Lys Ala Gln Ile Cys Gln Lys Gln Val Lys Tyr Leu

 1780 1785 1790
 Gly Tyr Leu Leu Lys Glu Gly Gln Arg Trp Leu Thr Glu Ala Arg Lys
 1795 1800 1805
 Glu Thr Val Met Gly Gln Pro Thr Pro Lys Thr Pro Arg Gln Leu Arg
 1810 1815 1820
 Glu Phe Leu Gly Lys Ala Gly Phe Cys Arg Leu Phe Ile Pro Gly Phe
 1825 1830 1835 1840
 Ala Glu Met Ala Ala Pro Leu Tyr Pro Leu Thr Lys Pro Gly Thr Leu
 1845 1850 1855

 Phe Asn Trp Gly Pro Asp Gln Gln Lys Ala Tyr Gln Glu Ile Lys Gln
 1860 1865 1870
 Ala Leu Leu Thr Ala Pro Ala Leu Gly Leu Pro Asp Leu Thr Lys Pro
 1875 1880 1885
 Phe Glu Leu Phe Val Asp Glu Lys Gln Gly Tyr Ala Lys Gly Val Leu
 1890 1895 1900
 Thr Gln Lys Leu Gly Pro Trp Arg Arg Pro Val Ala Tyr Leu Ser Lys
 1905 1910 1915 1920
 Lys Leu Asp Pro Val Ala Ala Gly Trp Pro Pro Cys Leu Arg Met Val

 1925 1930 1935
 Ala Ala Ile Ala Val Leu Thr Lys Asp Ala Gly Lys Leu Thr Met Gly
 1940 1945 1950

Gln Pro Leu Val Ile Leu Ala Pro His Ala Val Glu Ala Leu Val Lys
1955 1960 1965

Gln Pro Pro Asp Arg Trp Leu Ser Asn Ala Arg Met Thr His Tyr Gln
1970 1975 1980

Ala Leu Leu Leu Asp Thr Asp Arg Val Gln Phe Gly Pro Val Val Ala
1985 1990 1995 2000

Leu Asn Pro Ala Thr Leu Leu Pro Leu Pro Glu Glu Gly Leu Gln His
2005 2010 2015

Asn Cys Leu Asp Ile Leu Ala Glu Ala His Gly Thr Arg Pro Asp Leu
2020 2025 2030

Thr Asp Gln Pro Leu Pro Asp Ala Asp His Thr Trp Tyr Thr Asp Gly
2035 2040 2045

Ser Ser Leu Leu Gln Glu Gly Gln Arg Lys Ala Gly Ala Ala Val Thr
2050 2055 2060

Thr Glu Thr Glu Val Ile Trp Ala Lys Ala Leu Pro Ala Gly Thr Ser
2065 2070 2075 2080

Ala Gln Arg Ala Glu Leu Ile Ala Leu Thr Gln Ala Leu Lys Met Ala
2085 2090 2095

Glu Gly Lys Lys Leu Asn Val Tyr Thr Asp Ser Arg Tyr Ala Phe Ala
2100 2105 2110

Thr Ala His Ile His Gly Glu Ile Tyr Arg Arg Arg Gly Trp Leu Thr
2115 2120 2125

Ser Glu Gly Lys Glu Ile Lys Asn Lys Asp Glu Ile Leu Ala Leu Leu
2130 2135 2140

Lys Ala Leu Phe Leu Pro Lys Arg Leu Ser Ile Ile His Cys Pro Gly
2145 2150 2155 2160

His Gln Lys Gly His Ser Ala Glu Ala Arg Gly Asn Arg Met Ala Asp
2165 2170 2175

Gln Ala Ala Arg Lys Ala Ala Ile Thr Glu Thr Pro Asp Thr Ser Thr
2180 2185 2190

Leu Leu Ile Glu Asn Ser Ser Pro Ser Gly Gly Ser Leu Glu Gly Gly
2195 2200 2205

Gly Gly Ser Ser Thr Asp His Pro Lys Tyr Ser Asp Met Ile Val Ala

2210 2215 2220
Ala Ile Gln Ala Glu Lys Asn Arg Ala Gly Ser Ser Arg Gln Ser Ile
2225 2230 2235 2240
Gln Lys Tyr Ile Lys Ser His Tyr Lys Val Gly Glu Asn Ala Asp Ser
2245 2250 2255
Gln Ile Lys Leu Ser Ile Lys Arg Leu Val Thr Thr Gly Val Leu Lys
2260 2265 2270
Gln Thr Lys Gly Val Gly Ala Ser Gly Ser Phe Arg Leu Ala Lys Ser
2275 2280 2285

Asp Glu Pro Lys Arg Thr Ala Asp Gly Ser Glu Phe Glu Pro Lys Lys
2290 2295 2300
Lys Arg Lys Val

2305

<210> 3

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> tdTomato_G to C / T ins_8-17_spacer

<400> 3

cgcacggagg gctccatgaa

20

<210> 4

<211> 8

<212> DNA

<213> Artificial Sequence

<220><223> tdTomato_G to C / T ins_8-17_PBS

<400> 4

atggagcc

8

<210> 5

<211> 17

<212> DNA

<213> Artificial Sequence

<220><223> tdTomato_G to C / T ins_8-17_RT template
 <400> 5
 gaactcagtg gcggttc 17
 <210> 6
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Igf2_G to C / TA ins_9-14_Spacer
 <400> 6
 tattggaaga acttgccac 20
 <210> 7
 <211> 9
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Igf2_G to C / TA ins_9-14_PBS
 <400> 7
 ggcaagttc 9
 <210> 8
 <211> 14
 <212> DNA
 <213> Artificial Sequence
 <220><223> Igf2_G to C / TA ins_9-14_RT template
 <400> 8
 agataccgcg tgta 14
 <210> 9
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Adamts20_CG to AA_11-13_Spacer
 <400> 9
 agtgaataag aagacgtact 20
 <210> 10
 <211> 11

<212> DNA

<213> Artificial Sequence

<220><223> Adamts20_CG to AA_11-13_PBS

<400> 10

acgtcttctt a 11

<210> 11

<211> 13

<212> DNA

<213> Artificial Sequence

<220><223> Adamts20_CG to AA_11-13_RT template

<400> 11

gaccggcctt agt 13

<210> 12

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223>

> Casp1_TAGG del_12-12_Spacer

<400> 12

gtcttgtctc ttataggaga 20

<210> 13

<211> 12

<212> DNA

<213> Artificial Sequence

<220><223> Casp1_TAGG del_12-12_PBS

<400> 13

cctataagag ac 12

<210> 14

<211> 12

<212> DNA

<213> Artificial Sequence

<220><223> Casp1_TAGG del_12-12_RT template

<400> 14

acctctttca ct 12

<210>	15	
<211>	20	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	Hoxd13_G to T_10-15_Spacer	
<400>	15	
	gaggcataca tctccatgga	20
<210>	16	
<211>	10	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	Hoxd13_G to T_10-15_PBS	
<400>	16	
	atggagatgt	10
<210>	17	
<211>	15	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	Hoxd13_G to T_10-15_RT template	
<400>	17	
	gactggtaga cctcc	15
<210>	18	
<211>	20	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	Angpt1_CGG to TGA_11-13_Spacer	
<400>	18	
	cacattgccc atgttgaatc	20
<210>	19	
<211>	11	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	Angpt1_CGG to TGA_11-13_PBS	
<400>	19	

tcaacatggg c	11
<210> 20	
<211> 13	
<212> DNA	
<213> Artificial Sequence	
<220><223> Angpt1_CGG to TGA_11-13_RT template	
<400> 20	
gatataactg aat	13
<210> 21	
<211> 20	
<212> DNA	
<213> Artificial Sequence	
<220><223> Ksr2_TGAT ins_8-14_Spacer	
<400> 21	
tcctgccctg gctccgtggt	20
<210> 22	
<211> 8	
<212> DNA	
<213> Artificial Sequence	
<220><223> Ksr2_TGAT ins_8-14_PBS	
<400> 22	
acggagcc	8
<210> 23	
<211> 14	
<212> DNA	
<213> Artificial Sequence	
<220><223> Ksr2_TGAT ins_8-14_RT template	
<400> 23	
gtggcccacc tgat	14
<210> 24	
<211> 20	
<212> DNA	
<213> Artificial Sequence	

<220><223> Ar_G to T_13-13_Spacer

<400> 24

tctcacttgt ggcagctgca

20

<210> 25

<211> 13

<212> DNA

<213> Artificial Sequence

<220><223> Ar_G to T_13-13_PBS

<400> 25

agctgccaca agt

13

<210> 26

<211> 13

<212> DNA

<213> Artificial Sequence

<220><223> Ar_G to T_13-13_RT template

<400> 26

aagaagacct tga

13

<210> 27

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> tdTomato-nsgRNA

<400> 27

gccctcgatc tcgaactcag

20

<210> 28

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Igf2-nsgRNA

<400> 28

cttccccaga taccgcgtgt

20

<210> 29

<211> 20

<212>	DNA	
<213>	Artificial Sequence	
<220><223>	Adamt20-nsgRNA	
<400>	29	
	agctcgtgtt caaggacatg	20
<210>	30	
<211>	20	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	Casp1-nsgRNA	
<400>	30	
	ctgtcagaag tcttgtgctc	20
<210>	31	
<211>	20	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	Hoxd13-nsgRNA	
<400>	31	
	gataccttggc acagtacacc	20
<210>	32	
<211>	20	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	Angpt1-nsgRNA	
<400>	32	
	gatataactg aattcaacat	20
<210>	33	
<211>	20	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	Ksr2-nsgRNA	
<400>	33	
	acctgatacg gagccagggc	20

<210> 34
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Ar-nsgRNA
 <400> 34
 aggggaaaat atcaggaagt 20
 <210> 35
 <211> 14
 <212> DNA
 <213> Artificial Sequence
 <220><223> Igf2_dsgRNA_-34
 <400> 35
 agtcagcagg tttc 14

 <210> 36
 <211> 14
 <212> DNA
 <213> Artificial Sequence
 <220><223> Igf2_dsgRNA_-27
 <400> 36
 aggtgctagt cagc 14
 <210> 37
 <211> 14
 <212> DNA
 <213> Artificial Sequence
 <220><223> Igf2_dsgRNA_+7
 <400> 37
 tggagacagt ccgc 14
 <210> 38
 <211> 14
 <212> DNA
 <213> Artificial Sequence
 <220><223> Igf2_dsgRNA_+22

<400> 38

ggacgcctgc gcag

14

<210> 39

<211> 14

<212> DNA

<213> Artificial Sequence

<220><223> Adamts20_dsgRNA_-13

<400> 39

tcaagagcac agcc

14

<210> 40

<211> 14

<212> DNA

<213> Artificial Sequence

<220><223> Adamts20_dsgRNA_+14

<400> 40

tttctggtgc ttgg

14

<210> 41

<211> 14

<212> DNA

<213> Artificial Sequence

<220><223> Adamts20_dsgRNA_+24

<400> 41

cgattttgac tttc

14

<210> 42

<211> 14

<212> DNA

<213> Artificial Sequence

<220><223> Adamts20_dsgRNA_+41

<400> 42

cttccgetca cttc

14

<210> 43

<211> 14

<212> DNA

<213> Artificial Sequence

<220><223> Casp1_dsgRNA_-13

<400> 43

ctttgacttc tcta

14

<210> 44

<211> 15

<212> DNA

<213> Artificial Sequence

<220><223> Casp1_dsgRNA_-12

<400> 44

actttgactt ctcta

15

<210> 45

<211> 14

<212> DNA

<213> Artificial Sequence

<220><223> Casp1_dsgRNA_+7

<400> 45

cagcaaattc tttc

14

<210> 46

<211> 14

<212> DNA

<213> Artificial Sequence

<220><223> Casp1_dsgRNA_+62

<400> 46

gtattcatgt ctca

14

<210> 47

<211> 14

<212> DNA

<213> Artificial Sequence

<220><223> Hoxd13_dsgRNA_-12

<400> 47

cccggatcca aaag

14

<210> 48

<211> 14
 <212> DNA
 <213> Artificial Sequence
 <220><223> Hoxd13_dsgRNA_-11
 <400> 48
 cacttttgga tccg 14
 <210> 49
 <211> 14
 <212>
 > DNA
 <213> Artificial Sequence
 <220><223> Hoxd13_dsgRNA_+20
 <400> 49
 gctgttcac ccgt 14
 <210> 50
 <211> 14
 <212> DNA
 <213> Artificial Sequence
 <220><223> Hoxd13_dsgRNA_+24
 <400> 50
 cgggtggaac agcc 14
 <210> 51
 <211> 14
 <212> DNA
 <213> Artificial Sequence
 <220><223> Angpt1_dsgRNA_-15
 <400> 51
 cgaaatccag aaaa 14

 <210> 52
 <211> 14
 <212> DNA
 <213> Artificial Sequence
 <220><223> Angpt1_dsgRNA_-11
 <400> 52

atccagaaaa cgga	14
<210> 53	
<211> 14	
<212> DNA	
<213> Artificial Sequence	
<220><223> Angpt1_dsgRNA_+14	
<400> 53	
cttcagaac acga	14
<210> 54	
<211> 14	
<212> DNA	
<213> Artificial Sequence	
<220><223> Angpt1_dsgRNA_+17	
<400>	
> 54	
ttcccgctcgt gttc	14
<210> 55	
<211> 14	
<212> DNA	
<213> Artificial Sequence	
<220><223> Ksr2_dsgRNA_-12	
<400> 55	
tctccaaaca agat	14
<210> 56	
<211> 14	
<212> DNA	
<213> Artificial Sequence	
<220><223> Ksr2_dsgRNA_+7	
<400> 56	
tccggggggc acac	14
<210> 57	
<211> 14	
<212> DNA	
<213> Artificial Sequence	

<220><223> Ar_dsgRNA_-15

<400> 57

ggagatgaag cttc 14

<210> 58

<211> 27

<212> DNA

<213> Artificial Sequence

<220><223> Closed chromatin_Forward

<400> 58

gcagcagatg gcaagtaata ctaagat 27

<210> 59

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Closed chromatin_Reverse

<400> 59

cccttattct ctgagcatta gacagttata 30

<210> 60

<211> 28

<212> DNA

<213> Artificial Sequence

<220><223> Open chromatin_Forward

<400> 60

tgagtcacag catcagcatc gggctctct 28

<210> 61

<211> 28

<212> DNA

<213> Artificial Sequence

<220><223> Open chromatin_Reverse

<400> 61

ggagagggcc tttttctctt caaggttc 28

<210> 62

<211> 22

<212> DNA

<213> Artificial Sequence

<220><223> Igf2_Forward

<400> 62

tctccaggac gacttcccca ga

22

<210> 63

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Igf2_Reverse

<400> 63

ctccaggtgt catattggaa gaac

24

<210> 64

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Adamts20_Forward

<400> 64

gtggaatcaa gagcacagc

19

<210> 65

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Adamts20_Reverse

<400> 65

cagtgaataa gaagacgtac

20

<210> 66

<211> 25

<212> DNA

<213> Artificial Sequence

<220><223> Casp1_Forward

<400> 66

aaatcactgg tcttgtctct tatag

25

<210> 67

<211> 23
 <212> DNA
 <213> Artificial Sequence
 <220><223> Casp1_Reverse
 <400> 67
 cagcaaattc tttcacctct ttc 23

<210> 68
 <211> 25
 <212> DNA
 <213> Artificial Sequence
 <220><223> Hoxd13_Forward
 <400> 68
 tcggcacgag gcatacatct ccatg 25

<210> 69
 <211> 22
 <212> DNA
 <213> Artificial Sequence
 <220><223> Hoxd13_Reverse
 <400> 69
 cgttggctag cgtccaggac tg 22

<210> 70
 <211> 21
 <212> DNA
 <213> Artificial Sequence
 <220><223> Angpt1_Forward
 <400> 70

ccagaaaacg gagggagaag a 21
 <210> 71
 <211> 21
 <212> DNA
 <213> Artificial Sequence
 <220><223> Angpt1_Reverse
 <400> 71

taggcacatt gcccatgttg a	21
<210> 72	
<211> 24	
<212> DNA	
<213> Artificial Sequence	
<220><223> Ksr2_Forward	
<400> 72	
tctctccaaa caagattgga tcat	24
<210> 73	
<211> 17	
<212> DNA	
<213> Artificial Sequence	
<220><223> Ksr2_Reverse	
<400> 73	
tctcctgccc tggctcc	17