



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2023년03월30일

(11) 등록번호 10-2515727

(24) 등록일자 2023년03월27일

(51) 국제특허분류(Int. Cl.)
C12N 15/10 (2017.01) **C12N 15/113** (2010.01)
C12N 9/22 (2006.01)
(52) CPC특허분류
C12N 15/102 (2013.01)
C12N 15/113 (2013.01)
(21) 출원번호 10-2018-0003981
(22) 출원일자 2018년01월11일
심사청구일자 2021년01월11일
(65) 공개번호 10-2018-0082981
(43) 공개일자 2018년07월19일
(30) 우선권주장
62/445,130 2017년01월11일 미국(US)
(56) 선행기술조사문헌
KR1020160050069 A
US20160153006 A1
US20160237455 A1
KR1020150101446 A

(73) 특허권자
주식회사 톨젠
서울특별시 강서구 마곡중앙로 172, 8층(마곡동, 바이오이노베이션파크)
서울대학교산학협력단
서울특별시 관악구 관악로 1 (신림동)
연세대학교 산학협력단
서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)
(72) 발명자
이한웅
서울특별시 중구 소월로2길 30, 남산트라펠리스 704호
이재훈
서울특별시 마포구 공덕동(마포대로7길 22) 래미안공덕3차 아파트 305동 204호
(뒷면에 계속)
(74) 대리인
특허법인아이플레이

전체 청구항 수 : 총 11 항

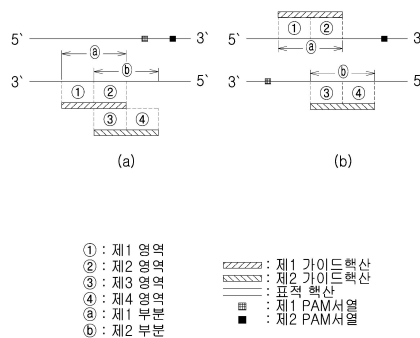
심사관 : 한지혜

(54) 발명의 명칭 **중첩된 가이드핵산을 이용한 표적 핵산에 특정 핵산 서열을 삽입하기 위한 조성물 및 방법**

(57) 요약

본 출원에 의해 개시되는 기술은 타겟 시퀀스를 조작하기 위한 조성물 및 표적 핵산에 특정 핵산 서열을 삽입하는 방법에 관한 것으로, 본 출원에 의해 개시되는 조성물 및 방법은 중첩된 가이드핵산을 사용하면, 특정 핵산 서열이 삽입되는 효율이 증가된다.

대 표 도 - 도1



(52) CPC특허분류

C12N 9/22 (2013.01)

C12N 2310/20 (2017.05)

(72) 발명자

염수청

강원도 평창군 대화면 평창대로 1447-1, 103동 20
9호

이재영

서울시 서초구 효령로 68길 81, 101동 602호

구옥재

경기도 수원시 장안구 이목로 24, 103동 1204호

이 발명을 지원한 국가연구개발사업

과제고유번호	1403-20170074
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	기초연구사업/신진연구자지원사업
연구과제명	cas9 녹인마우스 기반 생체 유전자적중 체장암 마우스모델 개발
기 여 율	1/2
과제수행기관명	서울대학교
연구기간	2017.11.01 ~ 2018.10.31

이 발명을 지원한 국가연구개발사업

과제고유번호	1403-20170004
부처명	식품의약품안전처
과제관리(전문)기관명	식품의약품안전평가원
연구사업명	R&D분야용역연구개발사업
연구과제명	기타 주요 질환 관련 모델마우스 제작 및 확인연구
기 여 율	1/2
과제수행기관명	서울대학교
연구기간	2017.01.01 ~ 2017.11.30

명세서

청구범위

청구항 1

표적 핵산을 조작하기 위한 조성물로서,

상기 조성물은,

Streptococcus pyogenes 유래의 Cas9 단백질 또는 이를 암호화하는 핵산;

제 1 가이드 RNA(first guide RNA) 또는 이를 암호화하는 핵산; 및

제 2 가이드 RNA(second guide RNA) 또는 이를 암호화하는 핵산;

을 포함하고,

상기 제 1 가이드 RNA는 제 1 가이드 도메인을 포함하는 제 1 crRNA(CRISPR RNA), 제 1 tracrRNA(trans-activating crRNA)를 포함하고,

이 때, 상기 제 1 가이드 도메인은 상기 표적 핵산과 전부 상보적인 결합을 할 수 있고,

상기 제 1 crRNA, 제 1 tracrRNA는 상기 Cas9 단백질과 상호작용하고,

상기 제 2 가이드 RNA는 제 2 가이드 도메인을 포함하는 제 2 crRNA(CRISPR RNA), 제 2 tracrRNA(trans-activating crRNA)를 포함하고,

이 때, 상기 제 2 가이드 도메인은 상기 표적 핵산과 전부 상보적인 결합을 할 수 있고,

상기 제 2 crRNA, 제 2 tracrRNA는 상기 Cas9 단백질과 상호작용하고,

상기 제 1 가이드 도메인의 제 1지점에서 3' 말단까지의 서열과 상기 제 2 가이드 도메인의 5' 말단에서 제 2지점까지의 서열은 서로 90% 내지 100% 동일하고,

이 때, 상기 제 2가이드 RNA의 3'말단과 상기 제 2지점 사이에 상기 SpCas9 단백질에 대응되는 PAM(Proto-spacer-Adjacent Motif)서열을 포함하고 있는 것을 특징으로 하는 조성물.

청구항 2

표적 핵산을 조작하기 위한 조성물로서,

상기 조성물은,

Streptococcus pyogenes 유래의 Cas9 단백질 또는 이를 암호화하는 핵산;

제 1 가이드 RNA(first guide RNA) 또는 이를 암호화하는 핵산; 및

제 2 가이드 RNA(second guide RNA) 또는 이를 암호화하는 핵산;

을 포함하고,

상기 제 1 가이드 RNA는 제 1 가이드 도메인을 포함하는 제 1 crRNA(CRISPR RNA), 제 1 tracrRNA(trans-activating crRNA)를 포함하고,

이 때, 상기 제 1 가이드 도메인은 상기 표적 핵산과 전부 상보적인 결합을 할 수 있고,

상기 제 1 crRNA, 제 1 tracrRNA는 상기 Cas9 단백질과 상호작용하고,

상기 제 2 가이드 RNA는 제 2 가이드 도메인을 포함하는 제 2 crRNA(CRISPR RNA), 제 2 tracrRNA(trans-activating crRNA)를 포함하고,

이 때, 상기 제 2 가이드 도메인은 상기 표적 핵산과 전부 상보적인 결합을 할 수 있고,

상기 제 2 crRNA, 제 2 tracrRNA는 상기 Cas9 단백질과 상호작용하고,

상기 제 1 가이드 도메인의 5'말단에서 제 1지점까지의 서열과 상기 제 2 가이드 도메인의 5'말단에서 제 2지점까지의 서열은 서로 90% 내지 100% 상보성을 가지는 것을 특징으로 하는 조성물.

청구항 3

제 1항 또는 제 2항에 있어서,

상기 제 2 가이드 도메인의 5'말단에서 제 2지점까지의 길이는 2 base 내지 16base인 것을 특징으로 하는 조성물.

청구항 4

제 1항 또는 제 2항에 있어서,

제 3 가이드 RNA(third guide RNA) 또는 이를 암호화하는 핵산;을 더 포함할 수 있고,

상기 제 3 가이드 RNA는 제 3 가이드 도메인을 포함하는 제 3 crRNA(CRISPR RNA), 제 3 tracrRNA(trans-activating crRNA)를 포함하고,

이 때, 상기 제 3 가이드 도메인은 상기 표적 핵산과 전부 상보적인 결합을 할 수 있고,

상기 제 3 crRNA, 제 3 tracrRNA는 상기 Cas9 단백질과 상호작용하는 것을 특징으로 하는 조성물.

청구항 5

제 4항에 있어서,

상기 제 3 가이드 도메인의 5'말단에서 제 3지점까지의 서열과 상기 제 2 가이드 도메인의 제 2지점에서 3'말단까지의 서열은 서로 90% 내지 100% 동일한 것을 특징으로 하는 조성물.

청구항 6

제 1항 또는 제 2항에 있어서,

상기 제 1 가이드 RNA 및 상기 제 2 가이드 RNA는,

서로 동일한 엑손 또는 인트론을 표적하는 것을 특징으로 하는 조성물.

청구항 7

제 1 항 또는 제 2항에 있어서,

표적 핵산을 조작하기 위해 도너(donor)를 더 포함할 수 있고, 이 때, 상기 도너는 표적 핵산에 삽입하고자 하는 특정 핵산 서열의 양 말단에 상보적으로 결합할 수 있는 호몰로지 암(homology arm)인 것을 특징으로 하는 조성물.

청구항 8

제 1항 또는 제 2항에 있어서,

상기 표적 핵산의 조작은, 넉인(Knock-in)인 것을 특징으로 하는 조성물.

청구항 9

제 1항 또는 제 2항에 있어서,

상기 표적 핵산은 진핵생물의 게놈 시퀀스인 것을 특징으로 하는 조성물.

청구항 10

제 1항에 있어서,

상기 PAM 서열은,

5'-NGG-3'(N은 A, T, C 또는 G임)인 것을 특징으로 하는 조성물.

청구항 11

표적 핵산을 조작하는 방법으로,

상기 제 1항 또는 2항 중 선택되는 어느 하나의 조성물을 세포에 도입시키는 단계;를 포함하고,

이 때, 상기 세포는 원핵세포 또는 분리된 진핵세포인 것을 특징으로 하는 방법.

청구항 12

삭제

청구항 13

삭제

청구항 14

삭제

청구항 15

삭제

청구항 16

삭제

청구항 17

삭제

청구항 18

삭제

발명의 설명

기술 분야

본 출원에 의해 개시되는 기술은 타겟 시퀀스를 조작하기 위한 조성물 및 표적 핵산에 특정 핵산 서열을 삽입하는 방법에 관한 것이다. 보다 구체적으로, 중첩된 가이드핵산을 이용하여 타겟 시퀀스를 조작하기 위한 조성물 및 표적 핵산에 특정 핵산 서열을 삽입하는 방법에 관한 것이다.

[0001]

배경 기술

- [0003] 유전자 변이를 인위적으로 초래하기 위해서, 이전에는 i) 방사선 또는 UV를 사용한 물리적 방법이나, ii) DNA 구조를 불안정하게 만드는 화학적 방법 등을 이용해 왔다. 하지만 이러한 방법들은 유전자 변이가 발생하는 위치를 특정한 위치로 제어하는 데에 어렵다는 단점이 있다.
- [0004] 이와 달리, 최근 개발된 유전자 가위 기술(예를 들어, Clustered Regularly Interspaced Short Palindromic Repeat, CRISPR)은 종래의 기술보다는 유전자 변이의 위치를 특정한 위치로 제어하는 것에 용이해졌다. 그럼에도 불구하고, 원하는 유전자를 인위적으로 도입하는데 있어서는 여전히 낮은 효율성을 지니고 있다.
- [0005] 따라서, 최근의 유전자 가위 기술과 분자생물학적 방법을 잘 융합시킴으로써 원하는 유전자의 도입 효율을 극대화한 유전자 편집 기술이 요구되고 있는 실정이다.

선행기술문헌

특허문헌

- [0007] (특허문헌 0001) 1. 미국 공개 특허 20160153006

비특허문헌

- [0008] (비특허문헌 0001) 1. FEBS J. 2014, 281(7):1717-1725. Dual sgRNAs facilitate CRISPR-Cas9-mediated mouse genome targeting.
- (비특허문헌 0002) 2. Cell. 2013, 152(5):1173-1183. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression.

발명의 내용

해결하려는 과제

- [0009] 본 출원에 의해 개시되는 기술의 일 과제는, 전술한 종래 기술의 문제점을 해결하기 위하여, 중첩된 가이드핵산을 이용하여 타겟 시퀀스를 조작하기 위한 조성물을 제공하는 것이다.
- [0010] 본 출원에 의해 개시되는 기술의 다른 과제는, 전술한 종래 기술의 문제점을 해결하기 위하여, 중첩된 가이드핵산을 이용하여 표적 핵산에 특정 핵산 서열을 삽입하여 상기 표적 핵산을 조작하는 방법을 제공하는 것이다.

과제의 해결 수단

- [0012] 전술한 본 출원에 의해 개시되는 기술의 과제를 달성하기 위하여, 본 출원에 의해 개시되는 기술의 일 양태에 따르면, 표적 핵산에 특정 핵산 서열의 삽입 효율 증가를 위한 조성물 및 방법에 관한 것이다.
- [0013] 전술한 본 발명의 과제를 해결하기 위하여, 본 출원에 의해 개시되는 기술의 일 양태에 의하면, 중첩된 가이드핵산을 이용하여 특정 핵산 서열의 삽입 효율을 증가시킬 수 있는 조성물이 제공된다.
- [0014] 상기 중첩된 가이드핵산을 이용하여 특정 핵산 서열의 삽입 효율을 증가시킬 수 있는 조성물은, Cas 단백질 또는 상기 Cas 단백질을 인코딩하는 제1 핵산; 적어도 제1 영역 및 제2 영역을 포함하는 제2 핵산; 및 적어도 제3 영역을 포함하는 제3 핵산;을 포함할 수 있다. 이 때, 상기 타겟 시퀀스를 조작하기 위하여, 상기 제2 핵산 및 상기 제3 핵산은 상기 Cas 단백질과 상호작용을 할 수 있다. 또한, 상기 제1 영역 및 제2 영역은 상기 타겟 시퀀스의 적어도 제1 부분에 상보적이고, 상기 제3 영역은 상기 타겟 시퀀스의 적어도 제2 부분에 상보적일 수 있으며, 상기 제2 영역 및 상기 제3 영역은 서로 공통 서열을 가지고 있거나 또는 서로 상보적인 서열을 가질 수 있다.
- [0016] 전술한 본 출원에 의해 개시되는 기술의 다른 과제를 해결하기 위하여, 본 출원에 의해 개시되는 기술의 다른 일 양태에 의하면, 중첩된 가이드핵산을 포함하는 표적 핵산에 특정 핵산 서열을 삽입하는 방법이 제공된다.

[0017] 상기 표적 핵산에 특정 핵산 서열을 삽입하는 방법은, i) Cas 단백질 또는 상기 Cas 단백질을 인코딩하는 제1 핵산, ii) 적어도 제1 영역 및 제2 영역을 포함하는 제2 핵산, iii) 적어도 제3 영역을 포함하는 제3 핵산, 및 iv) 상기 표적 핵산의 특정 영역에 삽입되기 위한 상기 특정 핵산 서열을 포함하는 조성물을 세포에 도입시키는 단계를 포함한다. 이 때, 상기 제2 영역 및 상기 제3 영역은 서로 공통 서열을 가지고 있거나 또는 서로 상보적인 서열을 가질 수 있으며, 상기 특정 핵산 서열은 상기 표적 핵산의 영역 중 상기 제2 핵산 및/또는 상기 제3 핵산에 대해 상보적인 영역의 적어도 일부에 중첩되도록 삽입될 수 있다.

발명의 효과

[0019] 본 출원에 의해 개시되는 기술에 따르면 다음과 같은 효과가 발생된다.

[0020] 중첩된 가이드핵산을 사용하여 표적 핵산에 특정 핵산 서열을 삽입하기 위한 조성물 및 방법을 제공할 수 있게 된다. 특히, 본 출원에 의해 개시되는 바와 같이 중첩된 가이드핵산을 사용하게 되면 종래의 기술에 비하여 특정 핵산 서열이 삽입되는 효율이 엄청나게 증가하게 된다.

도면의 간단한 설명

[0022] 도 1은 중첩된 가이드핵산의 예시들을 도시화한 도면이다.

도 2는 Rosa26 유전자를 표적유전자 자리로 한 sgRNAs의 설계를 나타낸 도면이다.

도 3은 Rosa26 유전자를 표적유전자 자리로 한 ssODN의 설계를 나타낸 도면이다.

도 4는 Rosa26 유전자를 표적유전자 자리로 한 배아의 핵인 비율을 나타낸 도면이다.

도 5는 Rosa26 유전자를 표적유전자 자리로 한 마우스 새끼의 유전형 분석결과를 나타낸 도면이다.

도 6은 Rosa26 유전자를 표적유전자 자리로 한 배반포의 유전형 분석결과를 나타낸 도면이다.

도 7은 Rosa26 유전자를 표적유전자 자리로 한 마우스 새끼의 염기서열 분석결과를 나타낸 도면이다.

도 8은 Upf1 유전자를 표적유전자 자리로 한 sgRNAs의 설계를 나타낸 도면이다.

도 9는 Upf1 유전자를 표적유전자 자리로 한 ssODN의 설계를 나타낸 도면이다.

도 10은 Upf1 유전자를 표적유전자 자리로 한 배아의 핵인 비율을 나타낸 도면이다.

도 11은 Upf1 유전자를 표적유전자 자리로 한 sgRNAs의 삭제서열의 분포를 나타낸 도면이다.

도 12는 Upf1 유전자를 표적유전자 자리로 한 배반포의 유전형 분석결과를 나타낸 도면이다.

도 13은 Srebf1 유전자를 표적유전자 자리로 한 sgRNAs의 설계를 나타낸 도면이다.

도 14는 Srebf1 유전자를 표적유전자 자리로 한 ssODN의 설계를 나타낸 도면이다.

도 15는 Srebf1 유전자를 표적유전자 자리로 한 마우스 새끼의 유전형 분석결과를 나타낸 도면이다.

도 16는 Tert 유전자를 표적유전자 자리로 한 sgRNAs의 설계를 나타낸 도면이다.

도 17은 Tert 유전자를 표적유전자 자리로 한 ssODN의 설계를 나타낸 도면이다.

도 18은 Tert 유전자를 표적유전자 자리로 한 마우스 새끼의 유전형 분석결과를 나타낸 도면이다.

도 19는 Morc2a 유전자를 표적유전자 자리로 한 sgRNAs의 설계를 나타낸 도면이다.

도 20은 Morc2a 유전자를 표적유전자 자리로 한 ssODN의 설계를 나타낸 도면이다.

도 21은 Morc2a 유전자를 표적유전자 자리로 한 DSB 패턴 분석결과를 나타낸 도면이다.

도 22는 Morc2a 유전자를 표적유전자 자리로 한 배아의 핵인 비율을 나타낸 도면이다.

도 23은 Morc2a 유전자를 표적유전자 자리로 한 배반포의 유전형 분석결과를 나타낸 도면이다.

도 24는 Morc2a 유전자를 표적유전자 자리로 한 sgRNAs의 삭제서열의 분포를 나타낸 도면이다.

도 25는 Morc2a 유전자를 표적유전자 자리로 한 핵인 효율 분석결과를 나타낸 도면이다.

도 26은 Adora2b 유전자를 표적유전자 자리로 한 sgRNAs와 ssODN의 설계를 나타낸 도면이다.

도 27은 Adora2b 유전자를 표적유전자 자리로 한 마우스 새끼의 유전형 분석결과를 나타낸 도면이다.

도 28는 Adora2b 유전자를 표적유전자 자리로 한 녹인 효율 분석결과를 나타낸 도면이다.

발명을 실시하기 위한 구체적인 내용

- [0023] 본 출원에 의해 개시되는 기술의 일 태양은 특정 핵산 서열의 삽입 효율을 증가시킬 수 있는 조성물에 관한 것이다.
- [0024] 구체적으로, 인위적으로 조작된 중첩된 가이드핵산을 사용한 표적 핵산에 특정 핵산 서열을 삽입하기 위한 조성물에 관한 것이다.
- [0026] 본 출원에 의해 개시되는 기술에서 사용되는 용어 또는 실시는 달리 정의되지 않는 한, 당업계의 기술 내에 있는 면역학, 생화학, 분자 생물학, 미생물학, 세포 생물학, 유전체학 및 재조합 DNA의 통상의 기술 등을 사용하는 당업자에 의해 통상적으로 이해되는 것과 동일한 의미를 가진다.
- [0028] 본 명세서에서 사용되는 용어 "핵산"은 뉴클레오타이드로 구성된 유기물을 의미하며, DNA(deoxyribonucleic acid), RNA(ribonucleic acid), DNA/RNA 혼합형태를 포함한다.
- [0029] 본 명세서에서 사용되는 용어 "뉴클레오타이드"는 인산, 염기, 당으로 구성된 물질을 총칭하며, 상기 당은 5탄당을 포함할 수 있으며, 상기 5탄당은 리보오스(ribose)와 디옥시리보오스(deoxyribose)를 포함할 수 있다. 상기 뉴클레오타이드는 단일가닥(single strand) 또는 이중가닥(double strand)형태로 존재하는 디옥시리보뉴클레오타이드 또는 리보뉴클레오타이드도 포함할 수 있다. 상기 뉴클레오타이드는 임의의 길이의 뉴클레오타이드, 디옥시리보뉴클레오타이드 또는 리보뉴클레오타이드 중 어느 하나, 또는 그것들의 유사체의 중합체 형태도 포함할 수 있다. 또한, 상기 뉴클레오타이드는 임의의 3차원 구조를 가질 수 있으며, 기지의 또는 미지의 임의의 기능을 수행할 수 있다. 상기 뉴클레오타이드의 구조에 대한 변형이 존재한다면, 중합체의 조립 전에 또는 후에 부여될 수 있다.
- [0030] 다음은 뉴클레오타이드의 비제한적인 예들이다. 유전자, 유전자 단편, 염색체 단편, 유전자 또는 유전자 단편의 코딩 영역, 유전자 또는 유전자 단편의 비코딩 영역, 엑손, 인트론, 전령 RNA(mRNA), 운반 RNA(tRNA), 리보솜 RNA(rRNA), 짧은 간섭 RNA(siRNA), 짧은 헤어핀 RNA(shRNA), 마이크로 RNA(miRNA), 리보자임, cDNA, 재조합 폴리뉴클레오타이드, 플라스미드, 벡터, 임의의 서열이 분리된 DNA, 임의의 서열이 분리된 RNA, 핵산 프로브, 프라이머, 하나 이상이 변형된 뉴클레오타이드(예를 들어, 메틸화 뉴클레오타이드 및 뉴클레오타이드 유사체) 등을 포함할 수 있다.
- [0031] 용어 뉴클레오타이드, 뉴클레오티드, 폴리뉴클레오타이드, 폴리뉴클레오티드, 올리고뉴클레오타이드, 올리고뉴클레오티드는 상호교환 가능하게 사용된다.
- [0032] 본 명세서에서 사용되는 용어 "단백질"은 아미노산(amino acid)으로 구성된 유기 물질을 총칭하며, 상기 단백질은 구조 단백질 및 생물학적 활성 단백질 등을 포함할 수 있으나 이에 한정하지 않는다. 상기 구조 단백질은 콜라겐, 케라틴 등으로 일반적으로 가늘고 긴 실 모양의 사슬로 구성되어 있고, 상기 생물학적 활성 단백질은 효소와 같이 생물의 체내에서 일어나는 화학반응을 촉매하거나, 호르몬처럼 신체의 다른 부위로 신호를 전달하거나, 체내의 다른 부위로 물질을 수송하거나, 항원처럼 외부의 침입으로부터 몸을 방어하는 역할 등을 할 수 있다. 상기 단백질은 임의의 3차원 또는 4차원 구조를 가질 수 있으나, 이에 한정하지 않는다. 또한, 상기 단백질은 기지의 또는 미지의 임의의 기능을 수행할 수 있다. 상기 단백질의 구조에 대한 변형이 존재한다면, 중합체의 조립 전에 또는 후에 부여될 수 있다.
- [0033] 상기 단백질은 아미노산의 펩타이드 결합, 이온결합, 공유결합, 비공유결합에 의한 소수성결합, 수소결합, 반데르발스 힘, 정전기적 인력, 이황화(-S-S-)결합 등으로 단백질 구조의 변형이 발생 할 수 있으나, 이에 제한하지는 않는다.
- [0034] 상기 단백질은 임의의 크기의 아미노산 또는 그것들의 유사체의 중합체 형태도 포함할 수 있으며, 이 때 아미노산의 종류와 아미노산 형태의 변형은 한정하지 않는다.
- [0035] 용어 단백질, 효소는 상호교환 가능하게 사용된다.
- [0037] 본 명세서에서 사용되는 용어 "상보성(complementarity)"은 통상적인 왓슨-크릭(Watson-Crick) 염기 쌍형성 또

는 기타 비통상적인 유형에 의한 다른 핵산 서열과 수소 결합(들)을 형성할 수 있는 핵산의 능력을 지칭한다. 상기 핵산서열은 아데닌(adenine, A), 사이토신(cytosine, C), 구아닌(guamine, G), 티민(thymine, T), 우라실(uracil, U)의 염기서열들을 의미한다. 핵산의 종류에 따라 DNA일 때는 아데닌은 다른 가닥의 티민과, 구아닌은 다른 가닥의 사이토신과 각각 수소결합으로 결합한다. RNA일 때는 아데닌이 우라실과 수소결합을 한다. 용어 '상보적'은 상보성을 포함하는 상태 또는 특징을 의미한다.

[0038] 본 명세서에서 사용되는 용어 "상보적인 서열"은 상기 상보성을 가질 수 있는 서열인 것을 의미한다.

[0039] 본 명세서에서, 상보성을 가지는 핵산, 핵산서열, 뉴클레오타이드 등은 수소 결합을 형성할 수 있는 핵산 분자 내의 잔기의 백분율을 의미한다(예를 들어, 10개 중 5, 6, 7, 8, 9, 10개는 50%, 60%, 70%, 80%, 90% 및 100% 상보성을 의미함). 또한, 본 명세서에서 특별한 언급이 없는 한, 상보성의 백분율은 50% 이상, 60% 이상, 70% 이상, 80% 이상, 90% 이상, 95% 이상, 99% 이상, 또는 100%의 서열 상보성을 갖는 모든 뉴클레오타이드 서열을 포함한다.

[0040] 예를 들어, '제1 가이드핵산과 제2 가이드핵산은 서로 상보적인 서열을 가진다'는 말의 의미는 제1 가이드핵산의 염기서열과 제2 가이드핵산의 핵산서열이 서로 수소 결합을 형성할 수 있되, 위에서 언급한 바와 같이 전체 10개의 염기서열 중 5, 6, 7, 8, 9, 10개는 50%, 60%, 70%, 80%, 90% 및 100%의 상보적인 결합을 할 수 있는 서열을 의미한다.

[0041] 본 명세서에서 사용되는 용어 "공통 서열"은 다음과 같이 해석되어야 한다. 예를 들어, '제1 가이드핵산과 제2 가이드핵산은 서로 공통 서열을 가진다'는 말의 의미는 상기 제1 가이드핵산의 핵산서열과 상기 제2 가이드핵산의 핵산서열이 부분적으로 또는 전체적으로 같은 핵산서열을 가지고 있음을 의미한다. 이 때, 상기 제1 가이드핵산의 핵산서열이 10개이고, 상기 제2 가이드핵산의 핵산서열이 10개일 경우 염기서열 중 1, 2, 3, 4, 5, 6, 7, 8, 9, 10개의 염기서열이 같을 수 있는 경우 모두를 포함할 수 있다. 용어 공통 서열, 동일한 서열은 상호교환 가능하게 사용된다.

[0042] 본 명세서에서 사용되는 용어 "중첩"은 둘 이상의 핵산, 핵산 서열이 서로 동일한 핵산 서열을 부분적으로 공유하고 있거나 혹은 둘 이상의 핵산의 적어도 일부 영역들이 서로 상보적인 서열을 가지는 것을 의미한다. 이 때, 중첩된 정도(%)는 예를 들어, 염기 서열 10개 중 1, 2, 3, 4, 5, 6, 7, 8, 9, 10개는 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% 및 100%를 포함할 수 있다.

[0044] 조작용 조성물-유전자 가위 시스템

[0045] 본 출원에 의해 개시되는 기술은 중첩된 가이드핵산을 이용한 표적 핵산에 특정 핵산 서열을 삽입하기 위한 조성물로서, 가이드핵산과 유전자가위 단백질을 포함할 수 있다.

[0047] **1. 가이드핵산(guide nucleic acid)**

[0048] 상기 "가이드핵산"은 표적 핵산을 인식할 수 있는 핵산을 의미하며, Cas 단백질과 결합하여 Cas 단백질을 표적 핵산으로 인도할 수 있는 기능을 가진 것을 지칭한다.

[0049] 상기 가이드핵산은 DNA, RNA 또는 DNA/RNA 혼합의 형태를 포함할 수 있으나, 바람직하게는 가이드 RNA일 수 있다.

[0050] 상기 가이드 RNA는 crRNA(CRISPR RNA) 또는/및 tracrRNA(trans-activating crRNA)를 포함할 수 있다. 상기 crRNA와 상기 tracrRNA의 특정 부위가 융합된 형태를 sgrRNA(single-chain guide RNA)로 지칭할 수 있다.

[0051] 상기 가이드 RNA는 5 내지 150개, 5 내지 100개, 5 내지 50개, 5 내지 40, 또는 5 내지 30개의 핵산서열을 가질 수 있다.

[0052] 상기 가이드핵산은 하나 이상의 도메인을 포함할 수 있다.

[0053] 이 때, 상기 도메인은 가이드 도메인, 제 1 상보적 도메인, 연결 도메인, 제 2 상보적 도메인, 근위 도메인, 꼬리 도메인 등 일 수 있으나, 이에 제한하지 않는다. 또한, 동일한 도메인을 반복적으로 포함하거나, 서로 다른 도메인을 포함할 수 있다.

[0054] 상기 도메인의 개수는 1, 2, 3, 4, 5, 6 또는 그 이상의 가이드 도메인 또는/및 제 1 상보적 도메인 또는/및 연결 도메인 또는/및 제 2 상보적 도메인 또는/및 근위 도메인 또는/및 꼬리 도메인을 포함할 수 있다.

[0055] 선택적으로 상기 가이드 도메인 또는/및 상기 제 1 상보적 도메인 또는/및 상기 연결 도메인 또는/및 상기 제 2

상보적 도메인 또는/및 상기 근위 도메인 또는/및 상기 꼬리 도메인의 염기서열의 일부 또는 전부는 화학적 변형을 포함할 수 있다. 상기 화학적 변형은 methylation, acetylation, phosphorylation, phosphorothioate linkage, locked nucleic acid(LNA), 2'-O-methyl 3'phosphorothioate(MS) 또는 2'-O-methyl 3'thioPACE(MS P)일 수 있으나, 이에 제한되지 않는다.

[0056] 상기 가이드핵산 및 상기 도메인의 길이는 서로 같거나 다를 수 있다.

[0057] 상기 "가이드 도메인"은 표적 유전자 또는 핵산 상에서 표적 서열과 상보적인 결합을 할 수 있는 상보적인 가이드 서열이 포함된 도메인으로, 표적 유전자 또는 핵산과의 특이적인 상호작용을 위한 역할을 할 수 있다.

[0058] 상기 "제 1 상보적 도메인"은 제 2 상보적 도메인과 상보적 핵산서열을 포함하는 핵산 서열로, 제 2 상보적 도메인과 이중가닥을 형성할 수 있을 정도로 상보성을 가질 수 있다.

[0059] 상기 "연결 도메인"은 두 개 이상의 도메인을 연결하는 핵산 서열로, 연결 도메인은 동일한 또는 서로 다른 두 개 이상의 도메인을 연결할 수 있다. 상기 연결 도메인은 두 개 이상의 도메인과 공유결합 또는 비공유결합을 할 수 있고, 또는 두 개 이상의 도메인을 공유적 또는 비공유적으로 연결할 수 있다.

[0060] 상기 "제 2 상보적 도메인"은 상기 제 1 상보적 도메인과 상보적 핵산서열을 포함하는 핵산서열로 제 1 상보적 도메인과 이중가닥을 형성할 수 있을 정도로 상보성을 가질 수 있다.

[0061] 상기 "근위 도메인"은 제 2 상보적 도메인에 근접하게 위치하는 핵산서열일 수 있다.

[0062] 상기 "꼬리 도메인"은 가이드핵산의 양 말단 중 어느 하나 이상의 말단에 위치하는 핵산서열일 수 있다.

[0064] 2. 유전자가위 단백질(genome editing protein)

[0065] 상기 "유전자가위 단백질"은 표적 핵산 또는 유전자를 절단(cleavage)할 수 있는 기능을 수행하는 단백질을 지칭하며, 상기 유전자가위 단백질은 효소를 포함할 수 있으며, 상기 효소는 뉴클레아제, 프로테아제 또는 제한효소 등을 포함할 수 있으나, 이에 제한하지 않는다.

[0066] 상기 유전자가위 단백질은 자연 상태에 존재하는 또는 존재하지 않는 인위적으로 생성된 효소 또는 단백질의 형태 또는 그것들의 일부가 변형된 형태를 포함할 수 있다.

[0067] 상기 효소는 CRISPR 효소일 수 있다.

[0068] 상기 CRISPR 효소는 CRISPR-Cas 시스템을 구성하는 단백질로서, 가이드핵산과 결합하여 복합체를 형성하여 표적 서열을 절단 또는 위치를 변형시키는 기능을 수행할 수 있다.

[0069] 상기 CRISPR 효소는 효소의 활성을 변경시켜 불완전 또는 부분 활성을 가지도록 변형한 형태를 포함할 수 있다.

[0070] 상기 CRISPR 효소는 Cas1, Cas1B, Cas2, Cas3, Cas4, Cas5, Cas6, Cas7, Cas8, Cas9(Csn1 및 Csx12로도 알려짐), Cas10, Csy1, Csy2, Csy3, Cse1, Cse2, Csc1, Csc2, Csa5, Csn2, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Csb1, Csb2, Csb3, Csx17, Csx14, Csx10, Csx16, CsaX, Csx3, Csx1, Csx15, Csf1, Csf2, Csf3, Csf4, Cpf1 등의 자연상태로 존재하는 것 또는 인위적으로 변형된 형태를 포함할 수 있으나, 이에 제한하지는 않는다.

[0071] 상기 CRISPR 효소는 CRISPR 시스템의 종류에 따라 크게 6가지의 Type으로 나눌 수 있다. Type I은 Cas1, Cas2, Cas3, Cas8 등으로 구성될 수 있고; Type II는 Cas1, Cas2, Cas2/Cas4, Cas9 등으로 구성될 수 있고; Type III는 Cas5, Cas6, Cmr1, Cmr3, Cmr4, Cmr5, Cas5/Cmr1, Cam2, Cmr5 등으로 구성될 수 있고; Type IV는 Csf1, Csf2, Csf3 등으로 구성될 수 있고; Type V는 Cas1, Cas2, Cas1, Cas2, Cas4, Cpf1, Cpf2 등으로 구성될 수 있고; Type VI는 C2C2, Cas1, Cas2 등으로 구성될 수 있으나, 이에 한정하지 않는다.

[0072] 상기 CRISPR 효소의 Type 중 Type II의 Cas9과 Type V의 Cpf1이 일반적으로 가장 많이 사용된다.

[0073] 상기 Cas9은 스트렙토코커스 피오게네스(*Streptococcus pyogenes*), 스트렙토코커스 써모필러스(*Streptococcus thermophilus*), 스트렙토코커스 속(*Streptococcus sp.*), 황색포도상구균(*Staphylococcus aureus*), 노카르디옉시스 다손빌레이(*Nocardiaopsis dassonvillei*), 스트렙토마이세스 프리스티네스피랄리스(*Streptomyces pristinaespiralis*), 스트렙토마이세스 비리도크로모게네스(*Streptomyces viridochromogenes*), 스트렙토마이세스 비리도크로모게네스(*Streptomyces viridochromogenes*), 스트렙토스포랑기움 로세움(*Streptosporangium roseum*), 스트렙토스포랑기움 로세움(*Streptosporangium roseum*), 알리사이클로바클루스 아시도칼다리우스(*Alicyclobacillus acidocaldarius*), 바실러스 슈도마이코이데스(*Bacillus pseudomycooides*), 바실러스 셀레니티

레두센스(*Bacillus selenitireducens*), 엑시구오박테리움 시비리쿰(*Exiguobacterium sibiricum*), 락토바실러스 델브루에키(*Lactobacillus delbrueckii*), 락토바실러스 살리바리우스(*Lactobacillus salivarius*), 마이크로실라 마리나(*Microscilla marina*), 부르크홀데리아레스 박테리움(*Burkholderiales bacterium*), 폴라로모나스 나프탈레니보란스(*Polaromonas naphthalenivorans*), 폴라로모나스 속(*Polaromonas* sp.), 크로코스파에라 와트소니(*Crocospaera watsonii*), 시아노테세 속(*Cyanothecae* sp.), 마이크로시스티스 아에루기노사(*Microcystis aeruginosa*), 시네코코커스 속(*Synechococcus* sp.), 아세토할로비움 아라바티쿰(*Acetohalobium arabaticum*), 암모니펙스 데젠시(*Ammonifex degensii*), 칼디셀룰로시럽토 베시(*Caldicelulosiruptor bescii*), 칸디다투스 데술포루디스(*Candidatus Desulforudis*), 클로스트리듐 보툴리눔(*Clostridium botulinum*), 클로스트리듐 디피실레(*Clostridium difficile*), 피네골디아 마그나(*Fingoldia magna*), 나트라나에로비우스 썬모필러스(*Natraneroobius thermophilus*), 펠로토마쿨럼 썬모프로피오니쿰(*Pelotomaculum thermopropionicum*), 아시디티오바실러스 칼두스(*Acidithiobacillus caldus*), 아시디티오바실러스 페로옥시단스(*Acidithiobacillus ferrooxidans*), 알로크로마티움 비노숨(*Allochroamatium vinosum*), 마리노박터 속(*Marinobacter* sp.), 니트로소코커스 할로필러스(*Nitrosococcus halophilus*), 니트로소코커스 와트소니(*Nitrosococcus watsoni*), 슈도알테로모나스 할로플란크티스(*Pseudoalteromonas haloplanktis*), 크테도노박테르 라세미페르(*Ktedonobacter racemifer*), 메타노할로비움 에베스티가툼(*Methanohalobium evestigatum*), 아나베나 바리아빌리스(*Anabaena variabilis*), 노둘라리아 스푸미게나(*Nodularia spumigena*), 노스톡 속(*Nostoc* sp.), 아르트로스피라 맥시마(*Arthrospira maxima*), 아르트로스피라 플라텐시스(*Arthrospira platensis*), 아르트로스피라 속(*Arthrospira* sp.), 링비아속(*Lyngbya* sp.), 마이크로콜레우스 크토노플라스테스(*Microcoleus chthonoplastes*), 오실라토리아 속(*Oscillatoria* sp.), 페트로토가 모빌리스(*Petrotoga mobilis*), 썬모시포 아프리카누스(*Thermosiphon africanus*) 또는 아카리오클로리스 마리나(*Acaryochloris marina*) 등 다양한 미생물 유래의 Cas9일 수 있다.

[0074] 또한, 상기 Cpf1은 *Streptococcus*, *Campylobacter*, *Nitratifactor*, *Staphylococcus*, *Parvibaculum*, *Roseburia*, *Neisseria*, *Gluconacetobacter*, *Azospirillum*, *Sphaerochaeta*, *Lactobacillus*, *Eubacterium*, *Corynebacter*, *Carnobacterium*, *Rhodobacter*, *Listeria*, *Paludibacter*, *Clostridium*, *Lachnospiraceae*, *Clostridiaridium*, *Leptotrichia*, *Francisella*, *Legionella*, *Alicyclobacillus*, *Methanomethyophilus*, *Porphyromonas*, *Prevotella*, *Bacteroidetes*, *Helcococcus*, *Letospira*, *Desulfovibrio*, *Desulfonatronum*, *Opitutaceae*, *Tuberibacillus*, *Bacillus*, *Brevibacillus*, *Methylobacterium* 또는 *Acidaminococcus* 유래의 Cpf1일 수 있다.

[0075] 상기 Cas9은 HNH 도메인, RuvC 도메인, PAM-상호작용(PI)도메인으로 구성될 수 있으며, 구체적인 Cas9의 구조적 특성은 Hiroshi Nishimasu et al. (2014) Cell 156:935-949를 참고할 수 있다.

[0076] 상기 Cpf1은 RuvC 도메인, Nuc 도메인, PAM-상호작용(PI)도메인으로 구성될 수 있으며, 구체적인 Cpf1의 구조적 특성은 Takashi Yamano et al. (2016) Cell 165:949-962를 참고할 수 있다.

[0077] 상기 CRISPR 효소는 가이드핵산과 상호작용하여 복합체를 형성하여 표적 핵산 또는 유전자의 표적 서열을 절단 또는 위치를 변형시키는 기능을 수행할 수 있다. 이 때, 그 기능은 상기 CRISPR 효소가 직접적으로 인지하는 PAM(protospacer adjacent modif)을 포함해야한다. 즉, 상기 PAM의 염기서열은 CRISPR 효소의 종에 따라 달라진다. 예를 들어, 일반적으로 많이 사용되는, *Streptococcus pyogenes* 에서 유래된 Cas9 단백질의 경우 5'-NGG-3'가 PAM 염기서열이 된다. 예를 들어, *Staphylococcus aureus* 에서 유래된 Cas9 단백질은 5'-NNGRRT-3'를 PMA 염기서열로 가진다. 결과적으로 상기 표적 핵산 또는 유전자의 표적 서열을 절단 또는 위치의 변형을 수행할 때, CRISPR 효소의 종의 선택에 따라 필수적으로 PAM의 염기서열이 바뀌며, 그것들에 의해 표적 서열을 절단 또는 위치의 변형이 결정된다.

[0079] 3. 가이드핵산 및 유전자가위 단백질을 포함하는 조성물

[0080] <제1 실시예>

[0081] 본 출원에 의해 개시되는 타겟 시퀀스를 조작하기 위한 조성물은, 표적 핵산에 대한 조작 효율을 높일 수 있는 것으로, 제1 가이드 핵산 또는 이를 인코딩하고 있는 핵산, 제2 가이드 핵산 또는 이를 인코딩하고 있는 핵산 및 유전자가위 단백질 또는 이를 인코딩하고 있는 핵산을 포함할 수 있다. 이 때, 상기 제1 가이드 핵산과 상기 제2 가이드 핵산은 완전히 동일하지는 않다. 즉, 상기 타겟 시퀀스를 조작하기 위한 조성물은, 적어도 둘 이상의 서로 다른 가이드 핵산을 포함할 수 있다.

[0082] 상기 조성물은 서로 동일한 서열을 가지는 복수의 제1 가이드 핵산을 포함할 수 있다. 또한, 상기 조성물은 서

로 동일한 서열을 가지는 복수의 제2 가이드 핵산을 포함할 수 있다. 또한, 상기 조성물은 복수의 유전자 가위 단백질을 포함할 수 있다.

- [0084] 상기 제1 가이드 핵산의 서열 및 상기 제2 가이드 핵산의 서열은 서로 상이할 수 있다. 다만, 본 출원에 의해 개시되는 조성물에 있어서, 상기 제1 가이드 핵산의 서열의 적어도 일부 및 상기 제2 가이드 핵산의 서열의 적어도 일부는 서로 연관되어 있다.
- [0085] 예를 들어, 상기 제1 가이드 핵산 및 상기 제2 가이드 핵산은 서로 동일한 유전자를 표적화할 수 있다.
- [0086] 다른 예를 들어, 상기 제1 가이드 핵산 및 상기 제2 가이드 핵산은 서로 동일한 엑손을 표적화할 수 있다.
- [0087] 또 다른 예를 들어, 상기 제1 가이드 핵산 및 상기 제2 가이드 핵산은 서로 동일한 인트론을 표적화할 수 있다.
- [0088] 본 출원에 의해 개시되는 상기 제1 가이드 핵산 및 상기 제2 가이드 핵산은 서로 중첩될 수 있다.
- [0089] 본 명세서에 있어서, 상기 "중첩"이라는 용어는, 둘 이상의 핵산이 서로 동일한 핵산 서열을 부분적으로 공유하고 있거나 혹은 둘 이상의 핵산의 적어도 일부 영역들이 서로 상보적인 서열을 가지는 것을 의미한다.
- [0090] 예를 들어, 상기 제1 가이드 핵산이 제1 영역 및 제2 영역을 가지고 있고, 상기 제2 가이드 핵산이 제3 영역 및 제4 영역을 가지고 있을 때, 상기 제2 영역의 서열과 상기 제3 영역의 서열이 서로 동일한 서열인 경우, 상기 제1 가이드 핵산과 상기 제2 가이드 핵산은 서로 중첩되어 있다고 할 수 있다.
- [0091] 다른 예를 들어, 상기 제1 가이드 핵산이 제1 영역 및 제2 영역을 가지고 있고, 상기 제2 가이드 핵산이 제3 영역 및 제4 영역을 가지고 있을 때, 상기 제2 영역의 서열과 상기 제3 영역의 서열이 서로 상보적인 서열인 경우, 상기 제1 가이드 핵산과 상기 제2 가이드 핵산은 서로 중첩되어 있다고 할 수 있다.
- [0092] 이하에서, 특별한 언급이 없는 한, 둘 이상의 핵산이 서로 중첩되어 있다고 기재하는 경우, 상기 둘 이상의 핵산이 서로 공통된 염기 서열을 가지는 경우와 서로 상보적인 서열을 가지는 경우를 모두 지칭하는 것으로 한다.
- [0093] 전술한 바와 같이, 상기 제1 가이드 핵산 및 상기 제2 가이드 핵산은 각각 가이드 도메인, 제 1 상보적 도메인, 연결 도메인, 제 2 상보적 도메인, 근위 도메인, 꼬리 도메인을 가질 수 있다.
- [0094] 본 출원에 의해 개시되는 상기 제1 가이드 핵산 및 상기 제2 가이드 핵산에 있어서, 특히, 상기 제1 가이드 핵산의 가이드 도메인 및 상기 제2 가이드 핵산의 가이드 도메인이 서로 중첩되어 있을 수 있다.
- [0096] 상기 조성물에 포함된 유전자가위 단백질은 모두 동일한 종의 CRISPR 효소일 수 있다.
- [0097] 본 출원에 의해 개시되는 상기 조성물에 포함되어 있는 상기 유전자가위 단백질이 모두 동일한 종의 CRISPR 효소(예를 들어, SpCas9)인 경우, 상기 제1 가이드 핵산의 제1 상보적 도메인 및 제2 상보적 도메인은 각각 상기 제2 가이드 핵산의 제1 상보적 도메인 및 제2 상보적 도메인과 서로 동일할 수 있다.
- [0098] 이하에서는, 설명의 편의를 위하여, 둘 이상의 가이드 핵산들이 서로 중첩되어 있다고 기재하면, 특별한 언급이 없는 한, 둘 이상의 가이드 핵산들 각각의 가이드 도메인이 서로 중첩되어 있는 것을 의미하는 것으로 한다.
- [0100] 상기 제1 가이드 핵산 및 상기 제2 가이드 핵산이 서로 중첩된 중첩 길이는 2 - 16 bp일 수 있다.
- [0101] 상기 제1 가이드 핵산 및 상기 제2 가이드 핵산은 서로 10~80% 정도 중첩될 수 있다.
- [0102] 상기 제1 가이드 핵산을 인코딩하고 있는 핵산, 상기 제2 가이드 핵산을 인코딩하고 있는 핵산 및 상기 유전자 가위 단백질을 인코딩하고 있는 핵산은 동일한 벡터에 패키징되어 있을 수 있다.
- [0103] 또는, 상기 제1 가이드 핵산을 인코딩하고 있는 핵산, 상기 제2 가이드 핵산을 인코딩하고 있는 핵산 및 상기 유전자가위 단백질을 인코딩하고 있는 핵산 중 둘은 동일한 벡터에 패키징되어 있고, 나머지 하나는 다른 벡터에 패키징되어 있을 수 있다.
- [0104] 또는, 상기 제1 가이드핵산, 상기 제2 가이드핵산 및 상기 유전자가위 단백질을 인코딩하고 있는 핵산은 동일한 벡터에 패키징 되거나 또는 서로 다른 벡터에 패키징 될 수 있다.
- [0105] 또는, 상기 제1 가이드핵산, 상기 제2 가이드핵산을 인코딩하고 있는 핵산 및 상기 유전자가위 단백질을 인코딩하고 있는 핵산은 동일한 벡터에 패키징 되거나 또는 서로 다른 벡터에 패키징 될 수 있다.
- [0106] 또는, 상기 제1 가이드핵산을 인코딩하고 있는 핵산, 상기 제2 가이드핵산 및 상기 유전자가위 단백질을 인코딩하고 있는 핵산은 동일한 벡터에 패키징 되거나 또는 서로 다른 벡터에 패키징 될 수 있다.

- [0108] 또는, 상기 제1 가이드핵산, 상기 제2 가이드핵산 및 상기 유전자가위 단백질은 서로 결합하여(또는 상호작용하여) 복합체를 형성한 상태로 상기 조성물에 포함되어 있을 수 있다. 이 때, 상기 복합체는 제1 가이드핵산 또는/및 상기 제2 가이드핵산과 유전자가위 단백질의 상호작용을 통해 형성된 복합체를 의미하며, 핵산-단백질 복합체(예를 들어, ribonucleoprotein;RNP)는 상기 제1 가이드핵산 또는/및 상기 제2 가이드핵산과 유전자가위 단백질의 상호작용을 통해 형성된 복합체를 포함할 수 있다.
- [0109] 이 때, 상기 유전자 가위 단백질 1분자에 대하여, 상기 제1 가이드핵산 1분자 및 상기 제2 가이드핵산 1분자가 모두 결합(또는 상호작용)하여 상기 복합체를 형성하는 것은 아니고, 상기 유전자가위 단백질 1분자와 상기 제1 가이드핵산 1분자가 하나의 복합체(제1 가이드핵산-단백질 복합체)를 형성하고, 상기 유전자가위 단백질 1분자와 상기 제2 가이드핵산 1분자가 다른 하나의 복합체(제2 가이드핵산-단백질 복합체)를 형성하는 것이 일반적이다.
- [0110] 이하에서, 특별한 언급이 없는 한, 복합체라고 기재하는 경우, 가이드핵산-단백질 복합체를 포함하는 상기 핵산-단백질 복합체를 복합체로 칭하는 것으로 한다.
- [0111] 상기 제1 가이드핵산 및 상기 유전자가위 단백질은 서로 결합하여 복합체를 형성한 상태로 상기 제2 가이드핵산을 인코딩하고 있는 핵산과 상기 조성물에 포함되어 있을 수 있다.
- [0112] 또는, 상기 제2 가이드핵산 및 상기 유전자가위 단백질은 서로 결합하여 복합체를 형성한 상태로 상기 제1 가이드핵산을 인코딩하고 있는 핵산과 상기 조성물에 포함되어 있을 수 있다.
- [0113] 상기 조성물은 제3 가이드 핵산을 더 포함할 수 있다. 상기 제3 가이드 핵산은 상기 제1 가이드 핵산 또는 상기 제2 가이드 핵산과 서로 중첩되어 있을 수 있다. 보다 구체적으로, 상기 제3 가이드 핵산의 가이드 도메인이 상기 제1 가이드 핵산의 상기 가이드 도메인 또는 상기 제2 가이드 핵산의 상기 가이드 도메인과 서로 중첩되어 있을 수 있다.
- [0114] 상기 조성물은, 상기 제1 내지 제3 가이드 핵산을 제외한 추가적인 가이드 핵산을 더 포함할 수도 있다.
- [0116] 본 출원에 의해 개시되는 조성물에 포함되는 둘 이상의 가이드 핵산들이 서로 어떠한 형태로 중첩될 수 있는지에 대하여 도 1을 참조하여 이하에서 간략하게 설명하기로 한다.
- [0117] 도 1에는 제1 영역(①) 및 제2 영역(②)을 포함하고 있는 제1 가이드핵산, 제3 영역(③) 및 제4 영역(④)을 포함하고 있는 제2 가이드핵산, 제1 PAM 서열, 제2 PAM 서열, 제1 부분(㉠), 제2 부분(㉡) 및 표적 핵산이 도시되어 있다.
- [0118] 도 1의 (a)에는 상기 제1 가이드핵산 및 상기 제2 가이드 핵산이, 타겟 유전자에 대한 이중 가닥의 절단(cleavage) 또는 특정 서열의 유전자의 삽입(insertion)을 유도하기 위하여, 타겟 유전자의 이중 가닥 중 동일한 가닥에 대해 상보적으로 결합하는 경우가 도시되어 있다. 한편, 도 1의 (b)에는 상기 제1 가이드 핵산 및 상기 제2 가이드 핵산이, 타겟 유전자에 대한 이중 가닥의 절단(cleavage) 또는 특정 서열의 유전자의 삽입(insertion)을 유도하기 위하여, 타겟 유전자의 이중 가닥의 서로 다른 가닥에 대해 상보적으로 결합하는 경우가 도시되어 있다.
- [0119] 도 1의 (a)와 같이, 상기 제1 가이드핵산과 상기 제2 가이드핵산이 서로 동일한 가닥(same strand)에 상보적으로 결합하도록 상기 제1 및 제2 가이드핵산이 설계된 경우, 상기 제1 가이드핵산의 상기 제2 영역 또는/및 상기 제2 가이드핵산의 상기 제3 영역이 서로 공통 서열을 가질 수 있다. 이 때, 상기 제1 영역 및 상기 제2 영역이 상기 표적 핵산의 제1 부분(㉠)에 상보적이고, 상기 제2 가이드핵산의 상기 제3 영역 또는/및 상기 제4 영역이 상기 표적 핵산의 상기 제2 부분(㉡)에 상보적일 수 있다.
- [0120] 한편, 도 1의 (b)와 같이, 상기 제1 가이드핵산과 상기 제2 가이드핵산이 서로 다른 가닥(different strand)에 상보적으로 결합하도록 상기 제1 및 제2 가이드핵산이 설계된 경우, 상기 제1 가이드핵산의 상기 제2 영역이 상기 제2 가이드핵산의 상기 제3 영역과 서로 상보적인 서열을 가질 수 있다. 이 때, 상기 제1 영역 및 상기 제2 영역이 상기 표적 핵산의 제1 부분(㉠)에 상보적이고, 상기 제2 가이드핵산의 상기 제3 영역 또는/및 상기 제4 영역이 상기 표적 핵산의 상기 제2 부분(㉡)에 상보적일 수 있다.
- [0122] 진술한 바와 같이, 도 1의 (a)에 도시된 바와 같이, 상기 제1 가이드핵산의 적어도 일부분(예를 들어, 제2 영역)과 상기 제2 가이드핵산의 적어도 일부분(예를 들어, 제3 영역)이 서로 공통된 서열을 가지는 경우와 같이 상기 제1 가이드핵산과 상기 제2 가이드핵산이 중첩될 수 있으며, 도 1의 (b)에 도시된 바와 같이, 상기 제1 가이드핵산의 적어도 일부분(예를 들어, 제2 영역)과 상기 제2 가이드핵산의 적어도 일부분(예를 들어, 제3 영

역)이 서로 상보적 서열을 가지는 경우와 같이 상기 제1 가이드핵산과 상기 제2 가이드 핵산이 중첩될 수도 있다.

[0123] 상기 제1 가이드핵산과 상기 제2 가이드핵산이 서로 중첩된 길이(예를 들어, 상기 제2 영역의 길이 또는 상기 제3 영역의 길이)는 조작하고자 하는 타겟 유전자의 시퀀스에 따라 서로 달라질 수 있으며, 최대 16bp를 넘지 않는 것이 바람직하며, 최소 2bp 보다는 긴 것이 바람직하다.

[0124] 이하에서, 설명하겠지만, 상기 제1 및 제2 가이드핵산들이 서로 중첩된 길이는 타겟 유전자의 시퀀스에 존재하는 PAM 서열들 간의 거리(예를 들어, 제1 PAM 서열의 위치와 제2 PAM 서열의 위치 간의 거리)에 의존적일 수 있다.

[0125] 예를 들어, 도 1의 (a)에 도시한 바와 같이, 상기 제1 및 제2 가이드핵산이 동일한 핵산에 대해 상보적으로 결합하는 경우, 상기 중첩된 길이는 상기 PAM 서열들 간의 거리가 멀어지면 멀어질수록 짧아질 수 있다.

[0126] 다른 예를 들어, 도 1의 (b)에 도시한 바와 같이, 상기 제1 및 제2 가이드핵산이 서로 다른 핵산에 대해 상보적으로 결합하는 경우, 상기 중첩된 길이는 상기 PAM 서열들 간의 거리가 멀어지면 멀어질수록 짧아질 수 있다.

[0128] 한편, 본 출원에 의해 개시되는 타겟 시퀀스를 조작하기 위한 조성물은, 둘 이상의 가이드핵산들을 포함하기만 하면되는 것으로 반드시 두 종류의 가이드핵산만 포함하고 있어야 하는 것은 아니다. 다시 말해, 상기 조성물은, 제1 가이드핵산 및 제2 가이드핵산 이외에도 제3 가이드핵산을 더 포함할 수 있다. 다만, 상기 제3 가이드핵산은 상기 제1 가이드핵산 및/또는 상기 제2 가이드핵산과 서로 중첩되어 있을 수 있다. 예를 들어, 상기 제3 가이드핵산과 상기 제1 가이드핵산(또는 상기 제2 가이드핵산)은 도 1을 참조하여 설명한 상기 제1 및 제2 가이드핵산들 간의 관계를 가질 수 있다.

[0129] 상기 가이드핵산은 제3 가이드핵산에 더하여 추가적인 가이드핵산을 가질 수 있음은, 자세히 설명하지 않아도 당업자에게 자명한 사항이라 할 것이다.

[0131] <제2 실시예>

[0132] 본 출원에 의해 개시되는 타겟 시퀀스를 조작하기 위한 조성물은, 표적 핵산 또는 유전자에 대한 조작 효율을 높일 수 있는 것으로, 제1 가이드 핵산 또는 이를 인코딩하고 있는 핵산, 제2 가이드 핵산 또는 이를 인코딩하고 있는 핵산, 유전자 가위 단백질 또는 이를 인코딩하고 있는 핵산 및 특정 핵산 서열을 포함할 수 있다.

[0133] 제1 실시예에서와 유사하게, 상기 제1 가이드 핵산의 서열 및 상기 제2 가이드 핵산의 서열은 서로 상이할 수 있다. 다만, 본 출원에 의해 개시되는 조성물에 있어서, 상기 제1 가이드 핵산의 서열의 적어도 일부 및 상기 제2 가이드 핵산의 서열의 적어도 일부는 서로 연관되어 있으며, 이에 대한 자세한 설명은 생략하기로 한다.

[0134] 상기 제1 가이드 핵산 및 상기 제2 가이드 핵산은, 제1 실시예에서 설명한 바와 같이, 서로 중첩되어 있을 수 있으며, 이에 대한 자세한 설명은 생략하기로 한다.

[0135] 상기 제1 가이드핵산 또는/및 상기 유전자가위 단백질이 복합체를 형성하여 표적 핵산 또는 유전자의 이중가닥 또는 단일가닥을 절단하고, 손상된 핵산 또는 유전자를 수선하기 위해 상동 재조합 수리(homology directed repairing, HDR)가 일어날 때, 상기 특정 핵산 서열이 사용될 수 있다.

[0136] 일반적으로, 상동 재조합 수리(HDR)는 유전자 가닥에 손상이 생겼을 때, 이중 가닥의 상보적인 특징으로 이용하여 수리되는 것으로, 자연상태에서는, 상기 손상된유전자는 상동 재조합 수리(HDR)에 의해 손상되기 전의 유전자 서열 그대로 복원된다. 그러나, 외부에서 인위적으로 특정 핵산 서열을 넣어주면 상기 상동 재조합 수리에 의해 유전자가 복원되는 과정에서 상기 특정 핵산 서열 그 자체가 혹은 상기 특정 핵산 서열과 동일한 시퀀스를 가지는 폴리뉴클레오타이드가 상기 유전자에 삽입될 수 있다.

[0137] 예를 들어, 상기 삽입하고자 하는 특정 핵산 서열의 양 말단에 상기 유전자에 상보적으로 결합할 수 있는 호몰로지 암(homology arm)이 구비되어 있는 경우, 상기 상동 재조합 수리(HDR) 과정에 의해서 손상된 유전자가 복원될 때, 상기 특정 핵산 서열이 유전자에 삽입되는 빈도가 더 높아진다는 것이 이미 널리 알려진 연구 결과이다.

[0138] 앞으로, 특별한 언급이 없으면, 특정 핵산 서열 그 자체를 '도너(donor)'라고 칭할 것이나, 특별한 언급이 있다면, 특정 핵산 서열 및 그 양 말단에 결합되어 있는 호몰로지 암(homology arm)까지 포함하는 의미로 '도너(donor)'라고 칭하도록 한다.

- [0140] 상기 특정 핵산서열(예를 들어, 도너 시퀀스)은, 다음의 예사와 같은 형태로 상기 조성물에 포함되어 있을 수 있으나, 이에 한정하지는 않는다.
- [0141] 예를 들어, 상기 제1 가이드핵산, 상기 제2 가이드핵산, 상기 유전자가위 단백질 및 상기 특정 핵산서열은 각각 핵산서열 상태로 하나 이상이 동일한 벡터에 존재할 수 있다. 또는, 상이한 벡터에 존재할 수 있다.
- [0142] 예를 들어, 상기 제1 가이드핵산, 상기 제2 가이드핵산 및 상기 유전자가위 단백질은 서로 결합하여 복합체를 형성한 상태로 상기 특정 핵산서열과 상기 조성물에 포함되어 있을 수 있다.
- [0143] 예를 들어, 상기 제1 가이드핵산 및 상기 유전자가위 단백질은 서로 결합하여 복합체를 형성한 상태로 상기 제2 가이드핵산을 인코딩하고 있는 핵산, 상기 특정 핵산서열과 상기 조성물에 포함되어 있을 수 있다.
- [0144] 예를 들어, 상기 제2 가이드핵산 및 상기 유전자가위 단백질은 서로 결합하여 복합체를 형성한 상태로 상기 제1 가이드핵산을 인코딩하고 있는 핵산, 상기 특정 핵산서열과 상기 조성물에 포함되어 있을 수 있다.
- [0146] **4. 가이드핵산 및 유전자가위 단백질을 포함하는 조성물의 이용**
- [0147] 본 출원에 의해 개시되는 기술에 의하면, 전술한 상기 타겟 시퀀스를 조작하기 위한 조성물의 이용예가 제공될 수 있다.
- [0148] 상기 조성물은 대상(target)을 조작하는데 사용될 수 있다.
- [0149] 상기 조작은 녹아웃(Knock-out), 녹인(Knock-in), 녹다운(Knock-down) 중 선택되는 어느 하나일 수 있다.
- [0150] 상기 대상은 표적 핵산 또는 유전자를 포함하는 유기체일 수 있다.
- [0151] 상기 유기체는 세포, 조직, 식물, 동물 또는 인간일 수 있다.
- [0152] 상기 세포는 원핵세포 또는 진핵세포일 수 있다.
- [0153] 일 예로, 상기 조성물은 유전자를 조작하는 데 사용될 수 있다.
- [0154] 구체예로, 상기 유전자는 Rosa26 유전자, Upf1 유전자, Srebf1 유전자, Tert 유전자, Morc2a 유전자, Adora2b 유전자 각각의 야생형 또는 인위적으로 변형된 형태를 포함할 수 있다.
- [0155] 상기 조작은 표적 핵산 또는 유전자를 절단, 변형 및 교정을 포함할 수 있다. 또한, 상기 조작은 상기 표적 핵산 또는 유전자에 특정 핵산 서열을 삽입하는 것을 포함할 수 있다.
- [0156] 상기 특정 핵산 서열이 상기 표적 핵산 또는 유전자에 삽입되면, 녹인이 발생할 수 있다.
- [0157] 상기 조작은 상기 표적 핵산 또는 상기 유전자의 핵산 서열에 대한 하나 이상의 뉴클레오타이드의 삽입, 치환, 결실 및 외부 유래의 하나 이상의 뉴클레오타이드의 삽입 중 하나 이상을 포함할 수 있다.
- [0158] 상기 표적 핵산 또는 유전자의 교정은 상기 표적 핵산 또는 유전자의 핵산 서열 중 일부에 있는 잘못된 시퀀스를 올바른 시퀀스로 바꾸는 것을 의미한다.
- [0159] 상기 잘못된 시퀀스의 예는, 일부 서열의 결실, 일부 서열의 추가, 일부 서열의 중복 등이 있을 수 있다.
- [0160] 일부 서열이 결실되어 있는 경우, 결실된 서열을 상기 표적 핵산 또는 유전자에 삽입함으로써 상기 표적 핵산 또는 유전자가 교정될 수 있다.
- [0162] **5. 타겟 시퀀스 조작 방법**
- [0163] 본 출원에 의해 개시되는 기술의 다른 양태에 의하면, 표적 핵산에 특정 핵산 서열을 삽입하는 방법이 제공된다. 상기 방법에 의하면, 종래의 방법에 비하여, 삽입하고자 하는 특정 핵산 서열을 표적 핵산에 삽입시킬 수 있는 효율(빈도, frequency)를 현저하게 높일 수 있게 된다.
- [0164] 상기 방법에는 전술한 타겟 시퀀스를 조작하기 위한 조성물이 사용될 수 있다.
- [0166] 상기 표적 핵산에 특정 핵산 서열을 삽입하는 방법은, i) Cas 단백질 또는 상기 Cas 단백질을 인코딩하는 제1 핵산, ii) 적어도 제1 영역 및 제2 영역을 포함하는 제2 핵산, iii) 적어도 제3 영역을 포함하는 제3 핵산, 및 iv) 상기 표적 핵산의 특정 영역에 삽입되기 위한 상기 특정 핵산 서열을 포함하는 조성물을 세포에 도입시키는 단계를 포함할 수 있다. 이 때, 상기 제2 영역 및 상기 제3 영역은 서로 공통 서열을 가지고 있거나 또는 서로 상보적인 서열을 가질 수 있다. 또한, 상기 특정 핵산 서열은 상기 표적 핵산에 포함된 서열들 중 상기 제2 핵

산 및/또는 상기 제3 핵산과 중첩되어 있는 영역의 적어도 일부에 중첩되는 위치에 삽입될 수 있다.

- [0168] 상기 표적 핵산에 특정 핵산 서열을 삽입하는 방법은, 상기 조성물이 상기 세포로 보다 잘 도입될 수 있도록 상기 세포를 전처리하는 단계를 더 포함할 수 있다.
- [0169] 예를 들어, 상기 세포가 식물 세포인 경우, 상기 조성물의 적어도 일부가 상기 식물 세포로 보다 더 잘 도입되도록 하기 위하여, 상기 식물 세포의 세포벽을 제거하기 위한 전처리 단계가 수행될 수 있다.
- [0171] 상기 조성물을 세포에 도입하면, 표적 핵산 또는 유전자가 조작될 수 있다. 특히, 상기 표적 핵산 또는 유전자가 낙인될 수 있다.
- [0172] 상기 조성물을 세포에 도입하면, 상기 제2 핵산의 제2 영역과 상기 제3 핵산의 제3 영역의 일부 또는 전체가 상기 제1 핵산과 결합하기 위해서 서로 경쟁을 하기 때문에, 낙인 효율이 더 높아질 수도 있다.
- [0173] 또한, 상기 조성물이 세포에 도입되면, 상기 제2 핵산과 상기 제3 핵산이 상기 특정 핵산 서열이 삽입되는 부위에 작은 서열 결실(small sequence deletion)을 일으켜, 낙인 효율이 더 높아질 수도 있다.
- [0175] 상기 표적 핵산 또는 유전자의 염기서열을 데이터베이스(예를 들어, NCBI 등)를 사용하여 유전자의 염기서열 데이터를 수집하여, 상기 염기서열 중 PAM 서열(예를 들어, 5'-NGG-3')이 있는 부위를 찾고, 그 PAM 서열로부터 ~20bp의 염기서열을 선정하여 적어도 상기 제2 핵산 및 상기 제3 핵산들의 중첩된 다양한 형태를 설계할 수 있다.
- [0176] 상기 제2 핵산, 상기 제3 핵산이 중첩된 가이드핵산 형태를 가질 때, 상기 중첩된 가이드핵산, 상기 PAM 서열은 도 1의 예시와 같이 설계할 수 있으나, 이에 제한되는 것은 아니다. 하기에서는, 중첩된 가이드핵산과 PAM 서열들 간의 관계에 대해서 설명할 것이며, 중첩된 가이드핵산의 형태에 대한 자세한 설명은 이전에 설명하였으므로 생략하기로 한다.
- [0177] 도 1에는 제1 영역(①) 및 제2 영역(②)을 포함하고 있는 제1 가이드핵산, 제3 영역(③) 및 제4 영역(④)을 포함하고 있는 제2 가이드핵산, 제1 PAM 서열, 제2 PAM 서열, 제1 부분(㉑), 제2 부분(㉒) 및 표적 핵산이 도시되어 있다.
- [0178] 예를 들어, 도 1의 (a)와 같이, 상기 제1 가이드핵산의 상기 제2 영역과 상기 제2 가이드핵산의 상기 제3 영역이 서로 공통 서열을 가질 때, 상기 제1 가이드핵산 및 상기 제2 가이드핵산과 상보적 서열을 가질 수 있는 표적 가닥(예를 들어, 도 1의 (a)에 도시된 표적 핵산의 이중 가닥 중 상측에 도시된 가닥)과 다른 가닥(different strand)(예를 들어, 도 1의 (a)에 도시된 표적 핵산의 이중 가닥 중 하측에 도시된 가닥)에 상기 제1 PAM서열 및 상기 제2 PAM서열들이 설계될 수 있다. 이 때, 상기 제1 PAM 서열로부터 20bp의 염기서열을 가지는 제1 가이드핵산 및 상기 제2 PAM 서열로부터 20bp의 염기서열을 가지는 제 2 가이드핵산을 설계할 수 있다.
- [0179] 다른 예를 들어, 도 1의 (b)와 같이, 상기 제1 가이드핵산이 상기 제1 부분에 상보적이고, 상기 제2 가이드핵산이 상기 제2 부분에 상보적일 때, 상기 제1 PAM서열 및 상기 제2 PAM서열은 표적 핵산의 서로 다른 가닥에 설계될 수 있다. 예를 들어, 제1 PAM서열은 도 1의 (b)에 도시된 표적 핵산의 이중 가닥 중 하측에 도시된 가닥에 존재하고, 제2 PAM서열은 도 1의 (b)에 도시된 표적 핵산의 이중 가닥 중 상측에 도시된 가닥에 존재할 수 있다. 이 때, 상기 제1 PAM 서열로부터 20bp의 염기서열을 가지는 제1 가이드핵산 및 상기 제2 PAM 서열로부터 20bp의 염기서열을 가지는 제 2 가이드핵산을 설계할 수 있다.
- [0180] 상기 중첩된 가이드 핵산의 길이(예를 들어, 상기 제2 영역 또는 상기 제3 영역)는 상기 제1 PAM서열과 상기 제2 PAM서열의 거리가 가까워 질수록 길어질 수 있다.
- [0182] 상기 표적 핵산의 표적 서열은 상기 제2 핵산 또는/및 상기 제3 핵산 각각의 가이드 도메인에 포함되는 가이드 서열과 상보성을 가진다. 상기 표적 서열은 표적 핵산 또는 유전자에 따라, 즉 유전자 조작을 하는 대상에 따라 달라질 수 있는 염기서열로, 다양하게 설계될 수 있다.
- [0183] 또한, 상기 표적 서열은 CRISPR 효소, 즉 Cas9 또는 Cpf1이 인식할 수 있는 PAM 서열에 근접한 위치에 위치한 염기서열일 수 있다. 상기 표적 서열은 다음의 예시와 같이 설계될 수 있으나, 이에 제한하지 않는다.
- [0184] 본 출원에 의해 개시되는 타겟 시퀀스를 조작하는 조성물을 이용할 때, 상기 표적 서열 하나에 대해서, 둘 이상의 가이드핵산들을 포함하여
- [0185] 예를 들어, 상기 표적 서열은 CRISPR 효소가 인식할 수 있는 PAM 서열의 5' 말단 또는/및 3' 말단에 인접하여

위치하는 연속하는 5 내지 50개의 염기서열일 수 있다.

[0186] 예를 들어, CRISPR 효소가 SpCas9인 경우, 상기 표적 서열은 5'-NGG-3', 5'-NAG-3' 또는/및 5'-NGA-3' (N= A, T, G 또는 C; 또는 A, U, G 또는 C) 서열의 5' 말단 또는/및 3' 말단에 인접하여 위치하는 연속하는 16 내지 25 개의 염기서열일 수 있다.

[0187] 예를 들어, CRISPR 효소가 SpCas9 단백질인 경우, 상기 PAM 서열은 5'-NGG-3' (N은 A, T, G, 또는 C임)이고, 상기 절단되는 염기서열 부위(표적 부위)는 표적 유전자 내의 5'-NGG-3' 서열의 5' 말단 및/또는 3' 말단에 인접하여 위치하는 연속하는 1bp 내지 25bp, 예컨대 17bp 내지 23bp 또는 21bp 내지 23bp의 염기서열 부위일 수 있다.

[0189] **[발명의 실시를 위한 형태]**

[0190] 이하, 실시예를 통하여 본 발명을 더욱 상세히 설명하고자 한다.

[0191] 이들 실시예는 오로지 본 발명을 보다 구체적으로 설명하기 위한 것으로서, 본 발명의 범위가 이들 실시예에 의해 제한되지 않는다는 것은 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에 있어 자명할 것이다.

[0193] **실험방법**

[0194] **1. 동물**

[0195] C57BL/6 마우스는 코아텍(Koatech, Pyeongtaek, Korea)에서 구입하였다. 모든 마우스는 개별적으로 통풍이 되는 케이지에서 유지하였고, 물과 먹이는 자유식으로 공급하였다. 본 연구는 서울대의 동물실험윤리위원회에 의해 승인 받았고, 승인 받은 가이드라인에 따라 수행하였다.

[0197] **2. Cas9, sgRNA 및 ssODN의 준비**

[0198] Cas9 mRNA는 툴젠(Toolgen Inc, Seoul, Korea)에서 구입하였다. 각 유전자를 위한 sgRNA는 Chopchop을 이용하여 설계하였고, PCR을 증폭한 후에 in vitro RNA 합성 키트(Thermo Fisher Scientific, Waltham, Ma, USA)를 사용하여 합성하였다. ssODNs은 상업적 서비스(IDT, San Jose, CA, USA)를 통해 설계하고 합성하였다. ssODN의 3' 말단은 ssODN의 안정성을 향상시키기 위해 포스포로티오에이트(phosphorothioate)로 변형시켰다.

[0200] **3. 미량주입(microinjection)**

[0201] C57BL/6 암컷 마우스에 임신 말 혈청성 성선 자극 호르몬(pregnant mare's serum gonadotropin, Prospeg, US A)과 인간 융모성 성선 자극 호르몬(human chorionic gonadotropin, Prospeg)을 주입하여 과잉 배란되도록 한 후 배아를 수집하였다. 1~2시간 배양 후, 전핵(pronuclei)을 가진 생존한 배아는 선택하였고, 현미조작장치(Eppendorf, Hamburg, Germany)를 이용하여 미량주입을 실시하였다. 간단하게 설명하면, 50ng/uL Cas9 mRNAs와 20ng/uL sgRNAs 및 20ng/uL ssODNs를 섞은 혼합물을 배아에 미량주입 하였다. 배아는 2-세포 단계(2-cell stage)까지 배양한 후 가임신 암컷 마우스로 옮기거나, 또는 유전형 분석을 위해 배반포(blastocyst) 단계가 될 때까지 배양하였다.

[0203] **4. 유전형 분석(Genotyping) 및 염기서열분석(sequencing)**

[0204] 상실배아(morula) 또는 배반포 상태의 배아 및 마우스 새끼의 꼬리로부터 DNA를 추출하였다. 단일 배아는 피펫(mouth pipet)을 이용하여 20uL DW가 들어있는 150uL 튜브로 옮기고, 3번의 동결/융해과정을 수행하고 95℃에서 15분 동안 변성(denaturation)시킨 후에 PCR을 위한 템플릿으로 사용하였다. 꼬리로부터 DNA 추출은 상업용 추출키트(Intron, Korea)를 이용하여 실시하였다. 염기서열분석은 일반적인 TA 클로닝과 생거 시퀀싱(Sanger sequencing, Cosmogenetech, Korea)을 이용하여 수행하였다.

[0206] **5. 세포배양 및 형질주입(tranfection)**

[0207] NIH3T3 (ATCC® CRL-1658™) 세포주는 high glucose (WelGene, South Korea)와 10% fetal calf serum (WelGene) 및 1X penicillin/streptomycin (WelGene)을 포함하는 Dulbecco's Modified Eagle Medium (DMEM)에서 배양했다. 형질주입을 위해, ssODNs(5ug)을 포함하거나 포함하지 않는 Cas9 단백질(ToolGen, South Korea) 4ug과 sgRNAs(각 1ug)와 ssODNs(5ug)를 포함 또는 포함하지 않도록 섞은 혼합물을 5분 동안 실온에서 배양하였다. 그 후에, 2×10^5 개 세포에 10uL 전기충공용 팁을 사용하는 네온 전기충공장치(Neon electroporator, Thermo Fischer Scientific)를 이용하여 상기 Cas9-sgRNA 혼합물을 세포 내로 전달하였다. 72시간 후에, 형질

주입된 세포를 수거하였고, 게놈 DNA 정제키트(GeneA11)를 이용하여 상기 세포의 게놈 DNA를 추출하였다.

6. 타겟 딥 시퀀싱(targeted deep sequencing)

형질주입된 세포로부터 추출한 게놈 DNA에서 On-target 영역을 Phusion polymerase taq(New England Biolabs)을 사용하여 PCR 증폭하였다. 그 다음, Mi-seq(Illumina)을 이용하여 상기 PCR 증폭물로 페어드-엔드 딥 시퀀싱(paired-end deep sequencing)을 수행하였다. 딥 시퀀싱 결과는 온라인 Cas-Analyzer(www.rgenome.net)를 이용하여 분석하였다. CRISPR/Cas9 반응에 의한 돌연변이 유무 확인을 위하여 PAM 서열로부터 3bp 업스트림에서 인델을 확인하였다. HDR 분석을 위하여, 상기 확인된 서열에서 다른 ssODNs로 인한 특정한 도너 서열이 발견된 녹인(KI)의 빈도를 관찰하였다.

7. 통계분석

통계분석을 위해, Graphpad Prism (Graphpad software, La Jolla, CA, USA)를 사용하여 t-test와 chi-square (X^2 analysis)를 수행하였다. X^2 analysis는 배아와 새끼에서 겹쳐진 sgRNA와 겹쳐지지 않는 sgRNA 사이에 KI 효율의 p 값을 계산하는데 사용하였고, t-test는 NIH3T3 세포주를 이용한 실험에서 KI 효율의 p 값을 계산하는데 사용하였다. $p < 0.05$ 일 때, 현저한 차이가 난다고 규정하였다.

실험예 1. Rosa26 유전자를 표적으로 하는 서열의 삽입 효율 확인

2개의 중첩된 sgRNAs를 이용하여 Rosa26 유전자 자리로 재조합효소 매개 카세트 변화(loxP and loxP2272) 서열(recombinase mediated cassette change (loxP and loxP2272) sequence)을 삽입하였다. 이를 위해, 마우스 Rosa26 유전자 자리를 표적으로 하는 PAM 서열을 포함하는 16bps를 공유하는 2개의 sgRNA(중첩된 sgRNAs)와 Cas9-mRNA, 34nt loxP를 포함하는 ssODN, 42nt 다중 클로닝 사이트(MCS), 34nt loxP2272 및 각 말단에 45nt 상동염기서열(homology arm)로 구성된 표적 위치는 Rosa26 유전자 위치의 인트론 1 영역을 사용하였다(도 2, 도 3). PCR 기반 유전형 분석을 위해 미량주입한 배반포(blastocyst) 단계의 배아로부터 게놈 DNAs(gDNAs)를 분리하여 총 67개의 배아 중에 17개의 배아가 KI된 서열을 가지고 있었고, 미량주입 한 배아 중 태어난 마우스 8마리 중 3마리의 마우스에서 KI된 서열을 가지고 있음을 확인하였다(도 4). 유전형 분석은 마우스 새끼에서 특정한 프라이머(WT 385bp, KI 495bp)로 PCR을 통해 확인하였다(도 5). 배반포의 유전형 분석은 특정 프라이머를 이용하여 PCR을 통해 확인하였다(도 6). 또한, 마우스 새끼의 염기서열 분석 결과로 팔호는 전체 실험한 클론에 대한 KI 클론의 수와 빈도를 나타낸다(도 7). 상기 결과들로 겹쳐지지 않는 sgRNAs 보다 겹쳐진 sgRNAs를 사용하였을 때, KI의 효율이 더 높음을 확인할 수 있었다.

실험예 2. Upf1 유전자를 표적으로 하는 서열의 삽입 효율 확인

넌센스 전사체 1(nonsense transcripts 1)의 조절자를 인코딩하는 Upf1 유전자를 표적 유전자로 선택하여 실험을 수행하였다. Upf1 유전자 자리의 경우, 인트론 1 영역 또는 인트론 2 영역을 표적하는 적어도 6bps가 겹치는 /또는 겹치지 않는 3개의 중첩된 sgRNAs와 loxP를 포함하는 ssODN을 설계하였다(도 8, 도 9). 배아에서 KI의 비율은 PCR을 통해 확인하였다(도 10). 겹치지 않는 sgRNA와 비교하였을 때, 겹치는 sgRNA를 사용하였을 때 KI 효율의 높음이 나타났다. NIH-3T3 세포에 동일한 sgRNAs를 형질주입하여 딥 시퀀싱(deep-sequencing)을 이용해 삭제된 패턴을 분석하였다(도 11). 겹쳐진 sgRNAs를 사용할 경우 짧은 삭제서열의 빈도가 높음을 확인하였다. 배반포의 유전형 분석은 특정 프라이머를 이용하여 PCR을 통해 확인하였다(도 12). NIH-3T3 세포에 Cas9 단백질, sgRNAs 및 ssODN을 형질주입하고 3일 후, 딥시퀀싱을 수행하였다. 상기 결과들로 겹쳐지지 않는 sgRNAs 보다 겹쳐진 sgRNAs를 사용하였을 때, KI의 효율이 더 높음을 확인할 수 있었다.

실험예 3. Srebf1 유전자를 표적으로 하는 서열의 삽입 효율 확인

Sterol regulatory element-binding protein 1을 인코딩하는 Srebf1 유전자를 표적 유전자로 선택하여 실험을 수행하였다. Srebf1 유전자 자리의 경우, 겹쳐지지 않는 2개의 sgRNAs를 이용하여 엑손 1에 loxP 서열을 삽입하였고, 각각 5bps가 겹쳐진 2개의 sgRNA를 엑손 4로 loxP 서열을 삽입하였다(도 13). ssODN은 상동염기서열, 인공서열 및 loxP 서열로 구성하였다(도 14). 마우스 꼬리로부터 수득한 gDNA를 PCR을 통한 유전형을 분석하였다(도 15). 상기 결과들은 전반적으로 겹쳐진 sgRNAs를 사용하였을 때 KI의 효율이 더 높음을 나타낸다.

실험예 4. Tert 유전자를 표적으로 하는 서열의 삽입 효율 확인

Telomerase transcriptase를 인코딩하는 Tert 유전자를 표적 유전자로 선택하여 실험을 수행하였다. Tert 유전자 자리의 경우, 각각 14bps 및 21bps가 겹쳐진 2개의 sgRNAs를 이용하여 각각 인트론 1 및 인트론 15에 loxP

서열을 삽입하였다(도 16). ssODN은 상동염기서열(homology arm)과 loxP로 구성하였다(도 17). 마우스 꼬리로부터 수득한 gDNA를 PCR을 통하여 유전형 분석을 하였다(도 18). 이러한 결과들은 겹쳐진 sgRNAs에서 작은 사이즈를 가지는 외부 유전자의 ssODN 매개의 KI 효율이 더 높음을 나타낸다.

[0227] **실험예 5. Morc2a 유전자를 표적으로 하는 서열의 삽입 효율 확인**

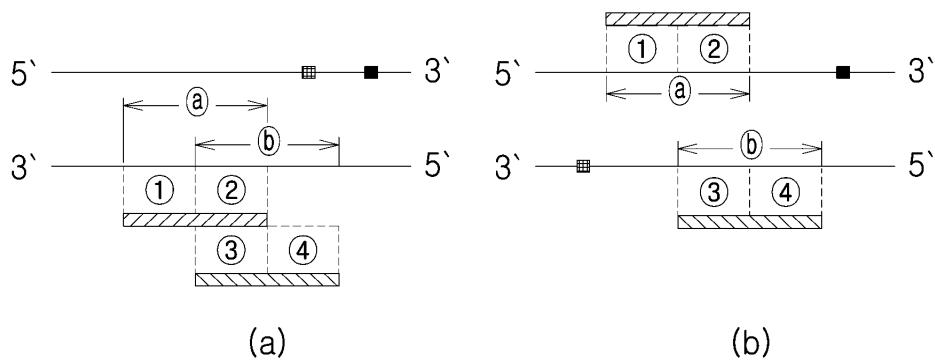
[0228] Microrhynchia 2A를 인코딩하는 Morc2a 유전자를 표적 유전자로 선택하여 실험을 수행하였다. Morc2a 유전자 자리의 경우, 엑손 6에 점돌연변이(point mutation)를 도입하였다(도 19). 상기 실험을 위해, 60bp 상동염기서열(homology arm)과 표적 유전자의 엑손 6에 260번째 사이토신을 티민으로 교체(cC260T)하는 단일 염기 다형성(Single nucleotide polymorphism, SNP)을 야기하는 숨은 돌연변이(silent mutation)를 포함하는 ssODN을 설계하였다(도 20). 상기 숨은 돌연변이는 표적 유전자가 단백질로 번역될 때 세린이 류신으로 아미노산 치환(pS87L)이 되도록 한다. DSB 패턴을 분석하여, sgRNAs의 결합 위치와 단일 염기 다형성(SNP)을 위한 표적 위치를 확인하였다(도 21). KI 분석을 위해 PCR을 이용한 유전형 분석 결과, 숨은 돌연변이를 야기하는 9개 이상의 점돌연변이가 삽입되었다. loxP의 KI 실험과 같은 방법을 이용하여, 겹쳐진 또는 겹쳐지지 않은 2개의 sgRNAs, Cas9-mRNA 및 ssODN을 1 세포기 접합체(one-cell stage zygotes)에 미량주입 하였고, 배아를 수집하여 PCR을 이용한 유전형 분석을 통해 KI 분석을 수행하였다(도 22). 배반포의 유전형 분석도 PCR을 이용하여 수행하였다(도 23). 또한, 겹쳐진 또는 겹쳐지지 않는 sgRNAs를 형질주입 한 NIH-3T3 세포로부터 gDNA를 수집하여 분석하였다. 그 결과, 타겟 딥 시퀀싱을 통해 겹쳐지지 않는 sgRNAs에 비해 겹쳐진 sgRNA에 의해 더 많은 빈도의 짧은 삭제 패턴을 확인하였다(도 24). NIH-3T3 세포에 Cas9 단백질, sgRNAs 및 ssODN을 형질주입하고 3일 후, 딥시퀀싱을 수행하였다. 또한, NIH-3T3 세포에 Cas9 단백질, sgRNAs 및 ssODN을 형질주입하고 3일 후, 딥시퀀싱을 수행하여, KI 효율을 분석하였다(도 25). 상기 결과들은 겹쳐진 sgRNAs를 사용하였을 때, KI 효율이 증가하는 것을 보여준다.

[0230] **실험예 6. Adora2b 유전자를 표적으로 하는 서열의 삽입 효율 확인**

[0231] Adenosine A2b receptor를 인코딩하는 Adora2b 유전자를 표적 유전자로 선택하여 실험을 수행하였다. Adora2b 유전자 자리의 경우, 3개의 겹쳐진 sgRNAs 및 2개의 겹쳐지지 않는 sgRNAs를 엑손 2에 점돌연변이를 도입하였다. 상기 실험을 위해, 50nt 상동염기서열과 표적 유전자의 엑손 2에 cA890G의 단일 염기 다형성을 야기하는 숨은 돌연변이를 포함하는 ssODN을 설계하였다(도 26). 겹쳐진 sgRNAs를 주입한 그룹에서 12마리의 새끼가 태어났고, 겹쳐지지 않는 sgRNAs를 주입한 그룹에서는 13마리의 새끼가 태어났다. 마우스 꼬리로부터 수득한 gDNA를 PCR을 통하여 유전형 분석을 하였다(도 27). 또한, 특정 프라이머로 딥시퀀싱을 수행하여 유전형 분석을 수행하였다. 겹쳐지지 않는 sgRNAs에 비해 겹쳐진 sgRNAs가 adora1 유전자 자리에 염기 치환(Ca890G) 유도 KI 효율이 현저하게 높은 것을 확인하였다(도 28). 이러한 결과들은 겹쳐진 sgRNAs가 마우스 접합체에 염기 치환 유도의 ssODN 매개 KI 효율을 향상시킴을 보여준다.

도면

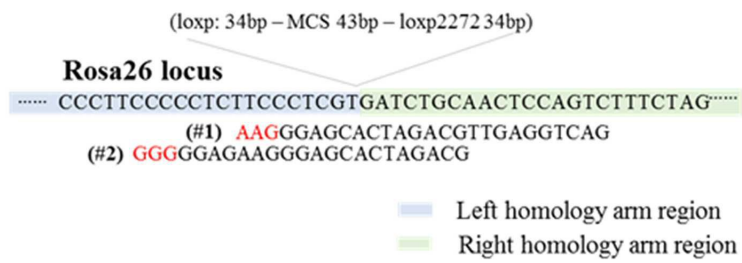
도면1



- ① : 제1 영역
- ② : 제2 영역
- ③ : 제3 영역
- ④ : 제4 영역
- a : 제1 부위
- b : 제2 부위

- ▨ : 제1 가이드핵산
- ▩ : 제2 가이드핵산
- ▬ : 표적핵산
- ▧ : 제1 PAM서열
- : 제2 PAM서열

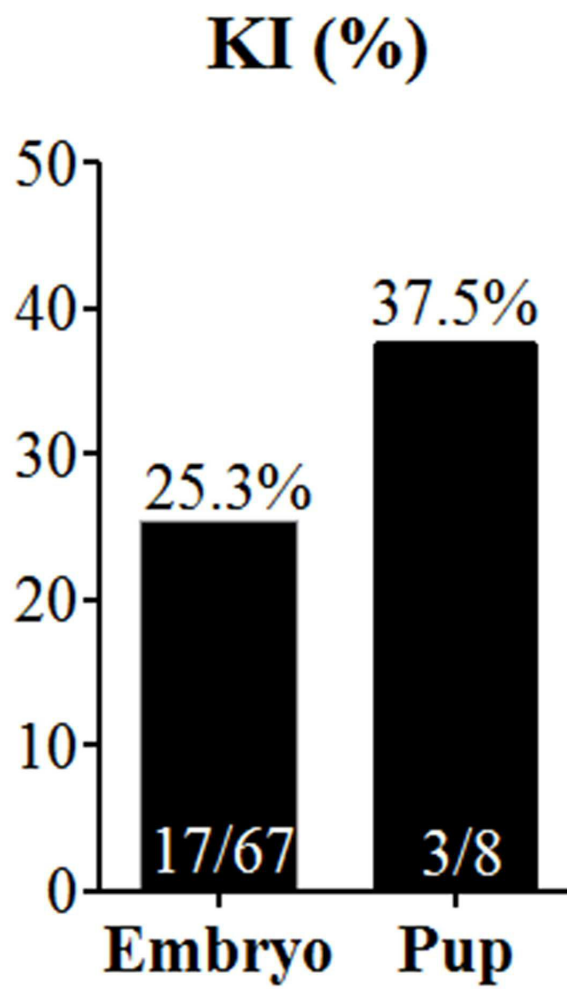
도면2



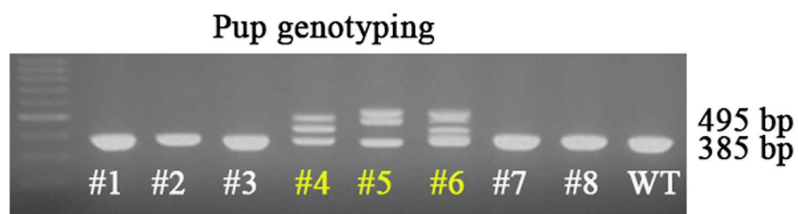
Sequence of ssODN (200 nt)

GGACCGCCCTGGGCCTGGGAGAATCCCTTCCCCCTCTTCCCCTCGT	: homology arm (45 nt)
ATAACTTCGTATAAATGTATGCTATACGAAGTTAT	: loxp (34 nt)
TTCGAATTCCTGCAGTCGACGGTACCGGGGCCCGGGATCCAT	: Multiple cloning site (42 nt)
ATAACTTCGTATAGGATACCTTTATACGAAGTTAT	: loxp (34 nt)
GATCTGCAACTCCAGTCTTTCTAGAAGATGGGGGGGAGTCTTCTG	: homology arm (45 nt), 3' end phosphorothioate taq

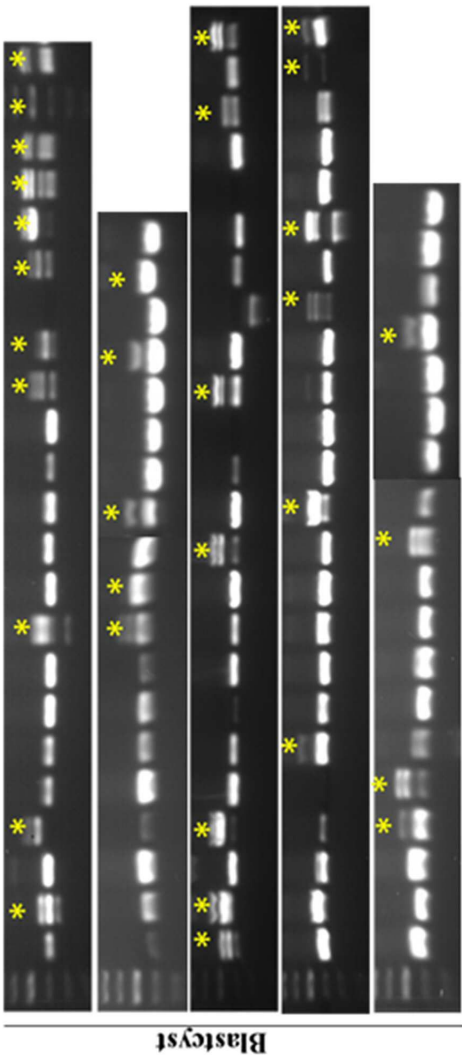
도면4



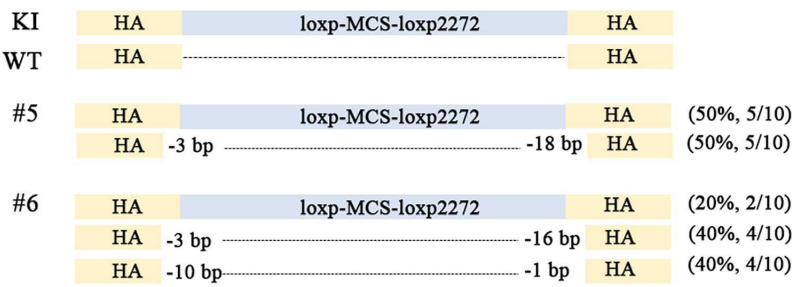
도면5



도면6



도면7



Sequence of ssODN in Upfl1 intron 1 (200 nt)

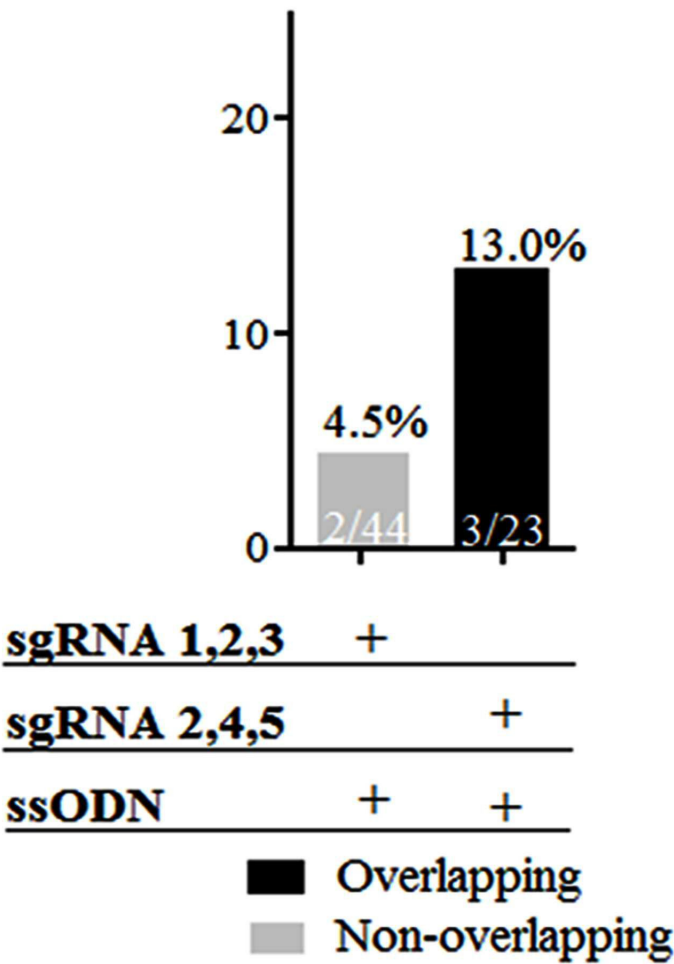
GCTACTTCTAGGCTCCTGAAGAACCCGGTGGCTAAGTGTGGGTGGA : homology arm (48 nt)
TGCTGACTTAGTGTCTCTATAGCTGAAG : homology arm with mutation in PAM sequence
ATAACTTCGTATAGCATACATTATACGAAGTTAT : loxP (34 nt)
ITCCTGGGGTGATGATTAAAGCTGTGCATCTCAGGCTAGTCAGG : homology arm with mutation in PAM sequence
TCTCTGTGGTGGGACTTCAGTCAGACATCCTATGCTGTGACTTTGAG : homology arm (48 nt) with 3' end phosphorothioate tag

Sequence of ssODN in Upfl1 intron 2 (200 nt)

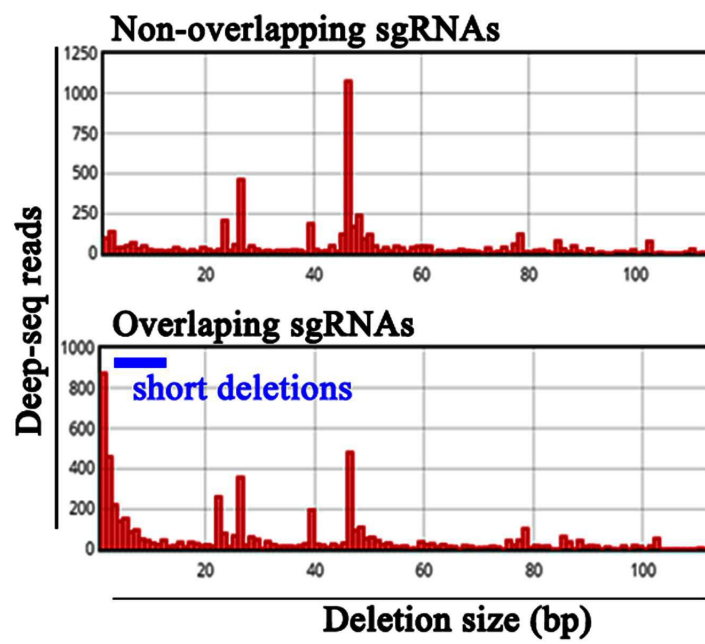
ACCCACACACACCCAGCAAGGCCAGAGCTTGTCTTTACTAGGGTTTTGTAGCCATCTG : homology arm (60 nt)
CGCCGCACTA : homology arm with mutation in PAM sequence (10 nt)
ATAACTTCGTATAGCATACATTATACGAAGTTAT : loxP (34 nt)
CAGGCTGTTTGGCCTCTCGACTCCACACCTATTGCTTTCC : homology arm with mutation in PAM sequence (42 nt)
GTTATAATGGCTCTCTCCTGGTTGACAAGTGTTTCGAGGATAAAACCAGAGC : homology arm (53 nt) with 3' end phosphorothioate tag

도면10

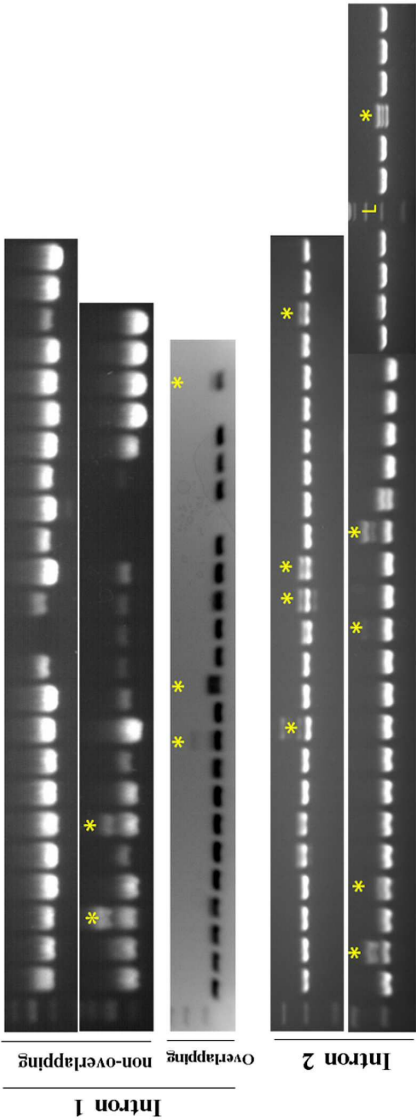
Upf1 (KI, Embryo)

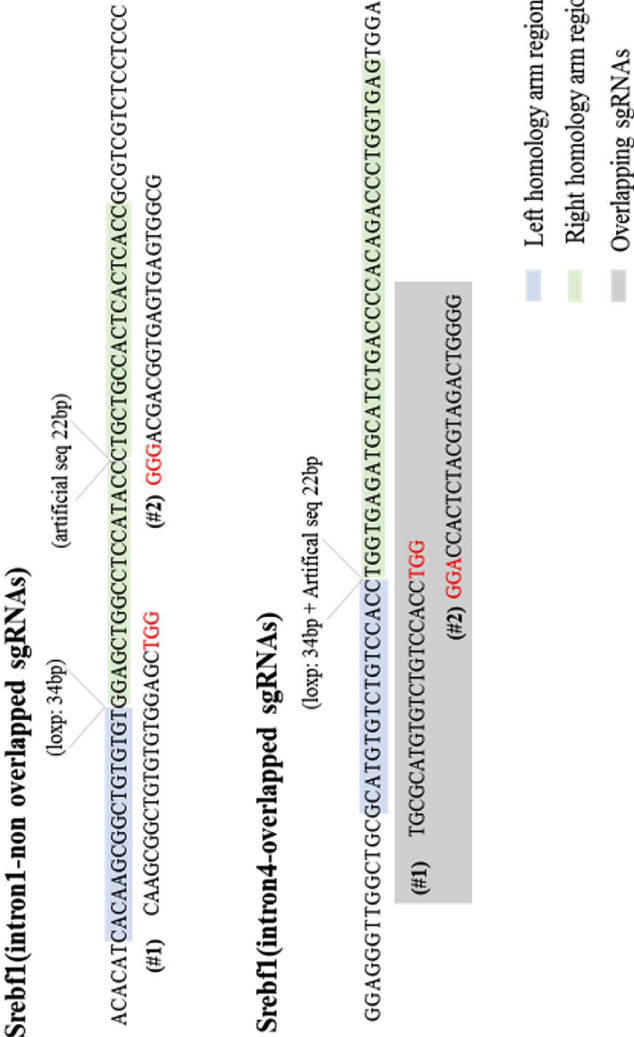


도면11



도면12





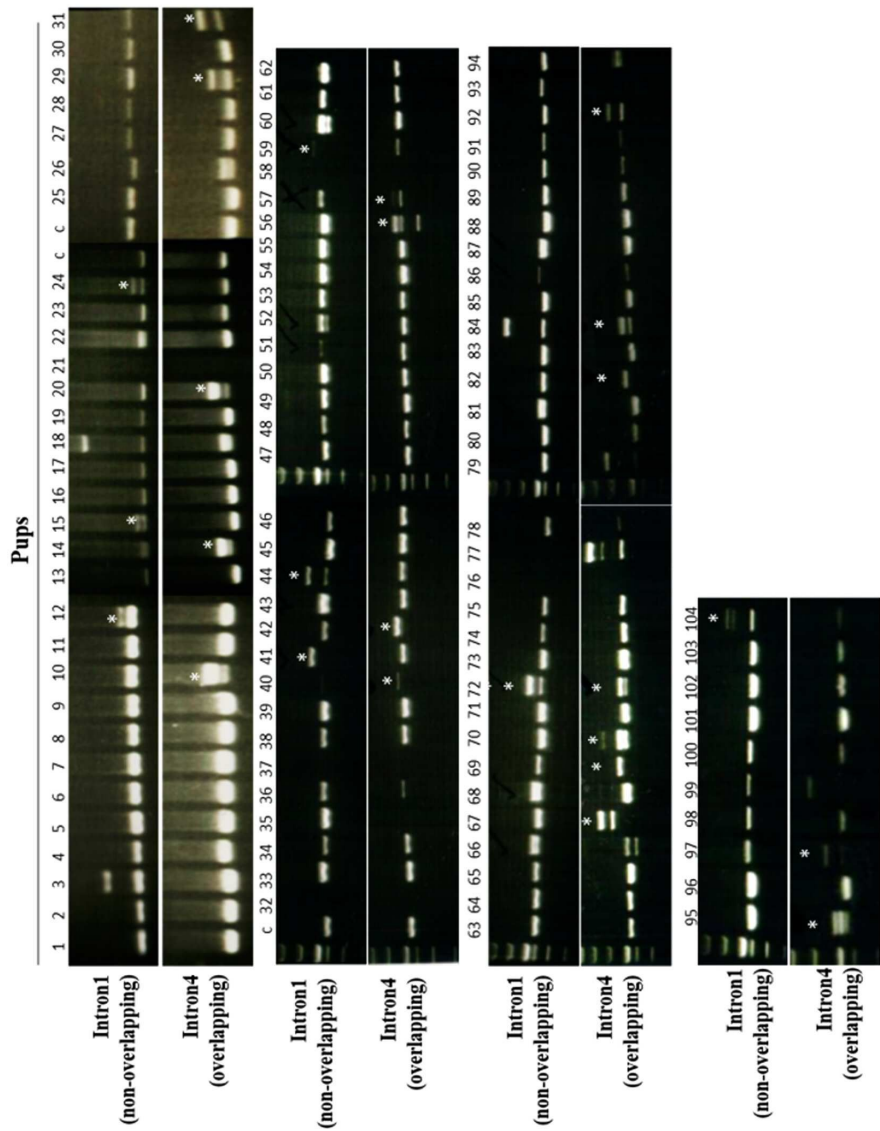
Sequence of ssODN in Srebf1 intron 1 (106 nt)

AGCGGCTGTGTGT : homology arm (13 nt)
ATAACTTCGTATAATGTATGCTATACGAAGTTAT : loxP (34 nt)
GGAGCTGGCCTCCATACC : homology arm (18 nt)
CGTCGCCGTCCAGCTCGACCAG : artificial sequence to screen knockin alleles (22nt)
CTGCTGCCACTCACTCACC : homology arm (19nt)

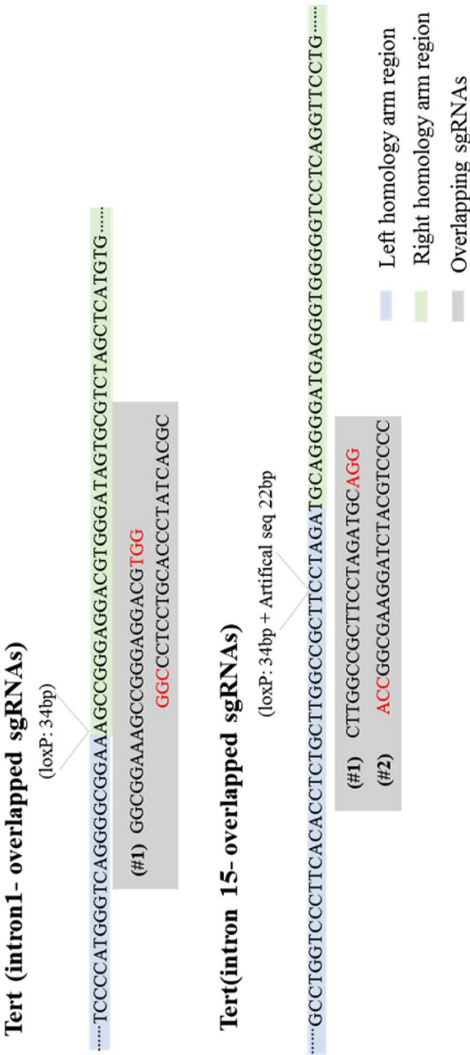
Sequence of ssODN in Srebf1 intron 4 (119 nt)

GGTTGGCTGCGCATGTGTCTGTCCACC : homology arm (27 nt)
ATAACTTCGTATAATGTATGCTATACGAAGTTATTGGCACGAACTCCAGCAGGACCATG : loxP and artificial sequence (59 t)
TGAGATGCACTCTGACCCACAGACCCCTGGTGAG : homology arm (33 nt)

도면15



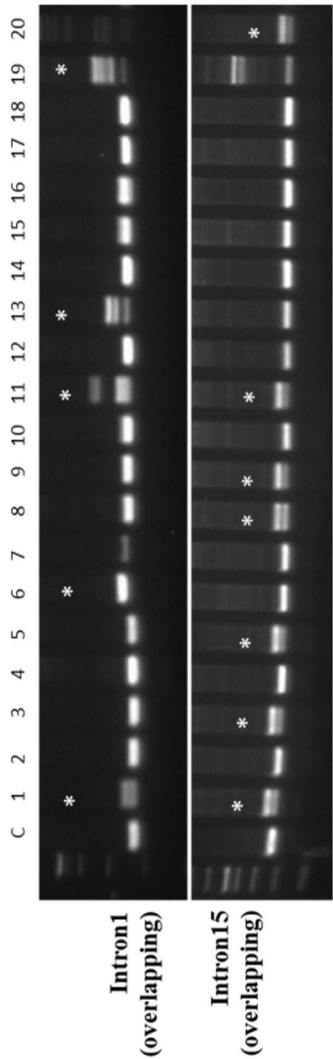
도면16



Sequence of ssODN in Tert intron 1 (194 nt)
GGGGCTCACAGCCCTCCACCTGCCGACCTTTCCTTCCACCAAGGTGGGCTCCAGGCGGGATCCCCATGGGTCAAGGGCGGAA : homology arm (80 nt)
ATAACTTCGTATAATGTATGCTATACGAAGTTAT : loxP (34 nt)
AGCCGGGAGGACGTGGGATAGTGGGTCTAGCTCATGTGTCAAGACCCCTTCTCCTTACCAGGTGTATCCCTGAAAGA : homology arm (79 nt)

Sequence of ssODN in Tert intron 4 (192 nt)
AGCCGGGAGGACGTGGGATAGTGGGTCTAGCTCATGTGTCAAGACCCCTTCTCCTTACCAGGTGTATCCCTGAAAGA : homology arm (79 nt)
TAACTTCGTATAGCATACATTATACGAAGTTATA : loxP sequence (34 nt)
TGCAGGGATGAGGGTGGGGTCTTCAGGTTCTCTGTCTTGTCTTGTTCAGATTGTCAAGCTTGAGGTCTTTTGTGTG : homology arm (79 nt)

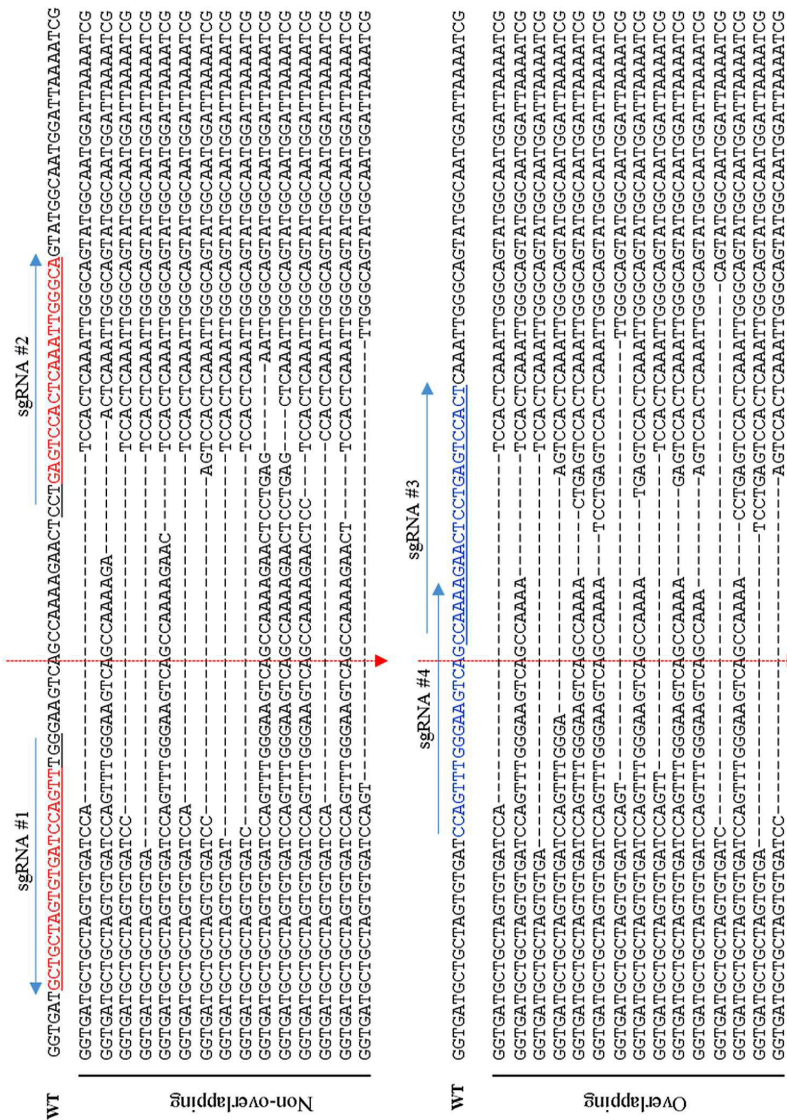
도면18



Sequence of ssODN (165 nt)

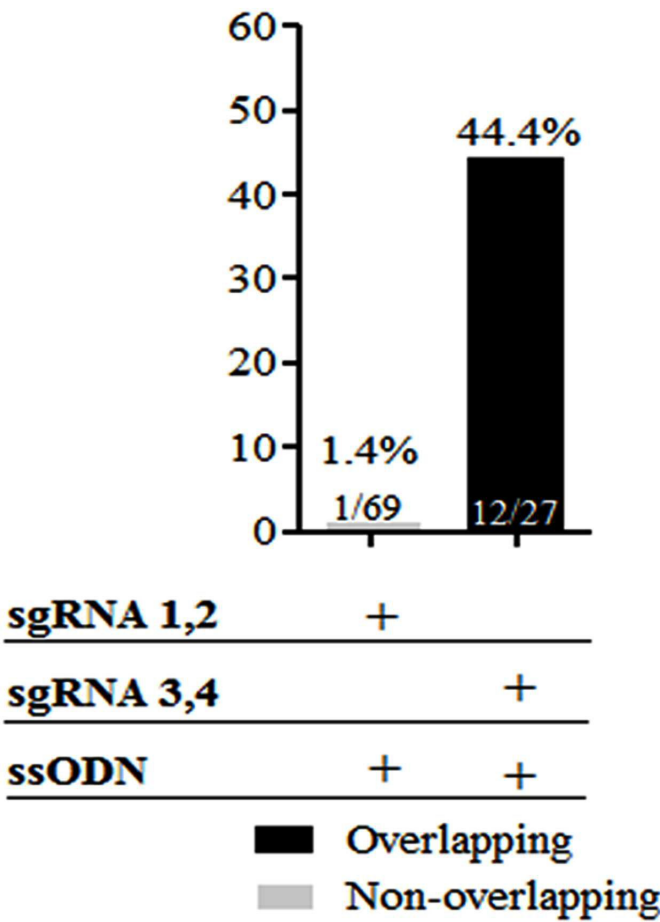
TGAGAAAGTAATGCTATTTTTTCCCTCTTGTTAACCATTGAGGTGAT : homology arm (50 nt)
GCTGCTAGTGTGATACAAATTTGGGAAAT(D)AGCGAAGAGACCCCTGA TCGACTCAAATTGGGCA : insert with silent mutation (65 nt)
GTATGGCAATGGATTAAATTCGTACGTATATTGGAAGCCTAGCATAGGCG : homology arm (50 nt)
with 3' end phosphorothioate tag

도면21

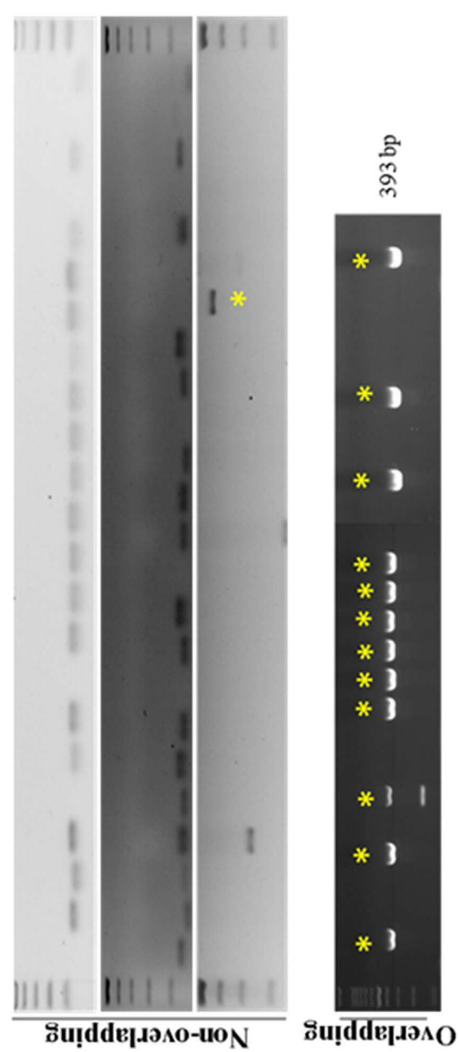


도면22

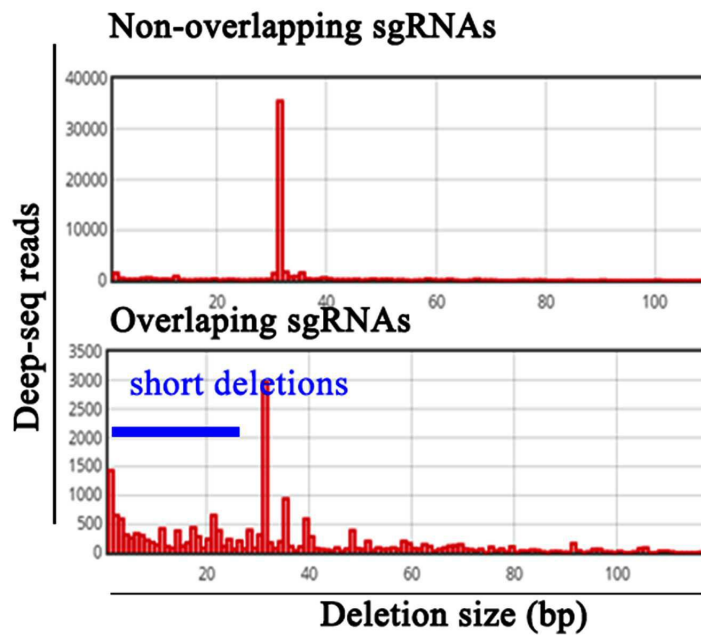
Morc2a (KI, Embryo)



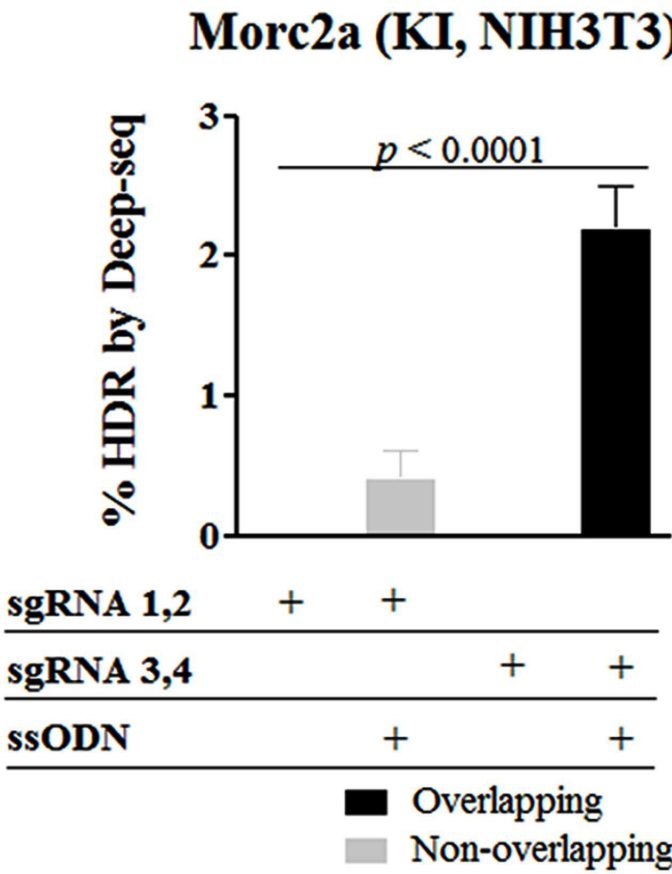
도면23



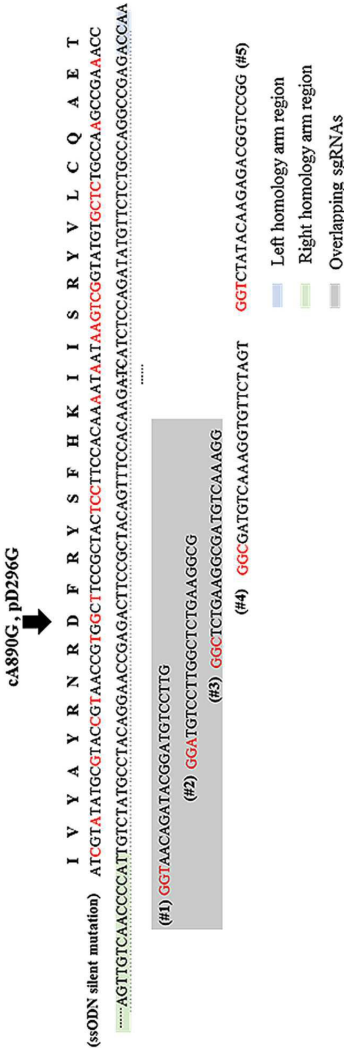
도면24



도면25



도면26



도면27

