

(43) 공개일자 2022년05월13일

- 전체 청구항 수 : 총 5 항

(54) 발명의 명칭 세포의 유전체에서 표적 핵산을 변형시키는 방법

신규한 PAM 서열을 이용하는 세포의 유전체에서 표적 핵산을 변형시키는 방법 및 이에 의해 유전체의 표적 핵산이 변형된 세포를 제공한다. 이에 의하면, 기존에 표적으로 하지 못한 위치를 유전체 편집의 표적으로 하여 유전체 편집을 수행할 수 있고, 이에 의해 유전체 편집의 활용 범위를 확대시킬 수 있다.

(52) CPC특허분류

C12N 15/113 (2013.01)

C12N 15/907 (2013.01)

C12N 9/22 (2013.01)

C12N 2310/20 (2017.05)

- 이 발명을 지원한 국가연구개발사업
- | | |
|-------------|----------------------------------|
| 과제고유번호 | 1711068059 |
| 과제번호 | 2017R1A2B3004198 |
| 부처명 | 과학기술정보통신부 |
| 과제관리(전문)기관명 | 한국연구재단 |
| 연구사업명 | 중견연구자지원사업 |
| 연구과제명 | 크리스퍼 유전자가위의 활성화에 영향을 미치는 인자 규명 및 |
- 대량산출(high-throughput) 방법을 이용한 유전학 연구 기초 기술 개발
- | | |
|---------|-------------------------|
| 기 여 율 | 30/100 |
| 과제수행기관명 | 연세대학교 산학협력단 |
| 연구기간 | 2017.03.01 ~ 2020.02.28 |
- 이 발명을 지원한 국가연구개발사업
- | | |
|-------------|------------------------------------|
| 과제고유번호 | 1711072824 |
| 과제번호 | 2017M3A9B4062403 |
| 부처명 | 과학기술정보통신부 |
| 과제관리(전문)기관명 | 한국연구재단 |
| 연구사업명 | 바이오.의료기술개발(R&D) |
| 연구과제명 | 생체 내 유전자 교정을 통한 근육 및 안 질환 치료 기술 개발 |
| 기 여 율 | 30/100 |
| 과제수행기관명 | 연세대학교 산학협력단 |
| 연구기간 | 2017.06.30 ~ 2019.12.29 |
- 이 발명을 지원한 국가연구개발사업
- | | |
|-------------|---|
| 과제고유번호 | 1465025389 |
| 과제번호 | HI17C0676000018 |
| 부처명 | 보건복지부 |
| 과제관리(전문)기관명 | 한국보건산업진흥원 |
| 연구사업명 | 세계선도의생명과학자육성사업 |
| 연구과제명 | 효율적인 생체 내 유전자 수술 방법 개발을 통한 유전성 간질환 치료법 발굴 |
| 기 여 율 | 30/100 |
| 과제수행기관명 | 연세대학교 산학협력단 |
| 연구기간 | 2017.04.10 ~ 2019.12.31 |
- 이 발명을 지원한 국가연구개발사업
- | | |
|-------------|---------------------------------|
| 과제고유번호 | 1465025972 |
| 과제번호 | HI16C1012020018 |
| 부처명 | 보건복지부 |
| 과제관리(전문)기관명 | 한국보건산업진흥원 |
| 연구사업명 | 임상연구인프라조성(R&D) |
| 연구과제명 | 유전성 대사성 뇌백질 질환 치료용 유전자가위 전달법 개발 |
| 기 여 율 | 5/100 |
| 과제수행기관명 | 연세대학교 산학협력단 |
| 연구기간 | 2016.05.17 ~ 2020.12.31 |
- 이 발명을 지원한 국가연구개발사업
- | | |
|-------------|-------------------------|
| 과제고유번호 | 1711074213 |
| 과제번호 | 2018R1A5A2025079 |
| 부처명 | 과학기술정보통신부 |
| 과제관리(전문)기관명 | 한국연구재단 |
| 연구사업명 | 기초연구사업- 선도연구센터지원사업 |
| 연구과제명 | 만성난치질환 시스템의학 연구센터 |
| 기 여 율 | 5/100 |
| 과제수행기관명 | 연세대학교 산학협력단 |
| 연구기간 | 2017.06.01 ~ 2025.02.28 |
-

명세서

청구범위

청구항 1

분리된 세포의 유전체에서 표적 핵산을 변형시키는 방법으로서, 상기 방법은

표적 핵산을 포함한 세포, Cas(clustered regularly interspaced palindromic repeats-associated protein: Cas) 뉴클레아제를 암호화하는 폴리뉴클레오티드, 및 가이드 RNA를 인큐베이션하는 단계로서,

상기 Cas 뉴클레아제는 SpCas9-NG이고,

상기 표적 핵산은 프로토스페이스어-인접 모티프(protospacer adjacent motif: PAM), 및 상기 가이드 RNA에 상보적인 표적 서열을 포함하고,

상기 PAM은 5'-NA-3'의 뉴클레오티드 서열이고,

상기 표적 핵산은 상기 Cas 뉴클레아제 및 상기 가이드 RNA의 복합체에 의해 인식될 수 있고,

상기 Cas 뉴클레아제 및 가이드 RNA의 복합체는 표적 핵산을 서열-특이적으로 변형시키는 것이고,

상기 방법은 시험관 내(in vitro) 또는 생체 외(ex vivo)에서 수행되는 것인 방법.

청구항 2

청구항 1에 있어서, 상기 가이드 RNA는 상기 표적 핵산에서 PAM의 5' 방향 또는 3' 방향으로 연속적인 2 내지 24 뉴클레오티드 서열과 상보적인 폴리뉴클레오티드인 것인 방법.

청구항 3

청구항 1에 있어서, 상기 가이드 RNA의 길이는 17 뉴클레오티드 내지 24 뉴클레오티드인 것인 방법.

청구항 4

청구항 1에 있어서, 상기 변형은 절단, 삽입, 연결, 탈아미노화, 또는 이들의 조합인 것인 방법.

청구항 5

청구항 1의 방법으로 유전체의 표적 핵산이 변형된 생체 외의 세포.

발명의 설명

기술 분야

[0001] 신규한 프로토스페이스어 인접 모티프를 이용한 세포의 유전체 편집 방법에 관한 것이다.

배경 기술

[0002] CRISPR-Cas9 시스템은, 세균 및 고세균에서 발견되는 적응적 면역 시스템으로, 다양한 생물 및 인간 세포를 포함한 다양한 세포의 유전체 편집에 이용되고 있다. 그러나, 표적 DNA 영역에 이용가능한 프로토스페이스어 인접 모티프(protospacer adjacent motif: PAM)이 없거나, 불충분한 온-타겟(on-target) 활성, 또는 오프-타겟(off-target) 효과 등으로 인해, CRISPR-Cas9의 적용이 흔히 제한된다. 가장 흔히 사용되는 CRISPR 뉴클레아제인, 스트렙토코커스 피오케네스 Cas9(*Streptococcus pyogenes* Cas9: SpCas9)의 PAM 서열은 5'-NGG-3'이다. 포유동물 세포에서 효율적인 활성을 나타내는 다른 CRISPR 뉴클레아제는 NGG 이외에 다른 PAM도 인식한다. PAM 서열이 없는 서열을 표적으로 하는 CRISPR 뉴클레아제 변이체가 개발되었지만, 표적이 되지 않는 서열들이 여전히 있다.

[0003] 최근, NGG가 아닌 PAM을 갖는 서열을 표적화할 수 있는 xCas9 및 SpCas9-NG가 개발되었다(Hu, J.H. et al., Nature vol.556, pp.57-63 (2018)). xCas9은 SpCas9에 비해 더 낮은 오프-타겟 활성을 나타내었다. xCas9 및 SpCas9-NG의 활성에 대해 겨우 각각 20 및 69 내재적인 인간 유전체 서열로만 연구되었을 뿐이고, 표적 서열에

대한 광범위한 조사는 이루어진 바가 없다.

- [0004] 따라서, CRISPR-Cas9 시스템의 적용 범위를 넓히기 위해, xCas9 및 SpCas9-NG의 PAM 호환성 및 표적 서열에 대한 활성을 조사하여, NGG가 아닌 PAM 서열을 발굴할 필요가 있다.

발명의 내용

해결하려는 과제

- [0005] 신규한 PAM 서열을 이용하는 세포의 유전체에서 표적 핵산을 변형시키는 방법을 제공한다.
- [0006] 신규한 PAM 서열을 이용하는 세포의 유전체에서 표적 핵산을 변형시키는 방법에 의해 유전체의 표적 핵산이 변형된 세포를 제공한다.

과제의 해결 수단

- [0007] 일 양상은 세포의 유전체에서 표적 핵산을 변형시키는 방법으로서, 표적 핵산, Cas(clustered regularly interspaced palindromic repeats-associated: CRISPR associated) 뉴클레아제 또는 이의 변이체, 및 가이드 RNA를 인큐베이션하는 단계를 포함하는 방법을 제공한다.
- [0008] 상기 세포는 생체 내 세포 또는 분리된 세포일 수 있다. 상기 세포가 생체 내 세포인 경우, 상기 방법은 생체 내(in vivo)에서 수행될 수 있다. 상기 세포가 분리된 세포인 경우, 상기 방법은 생체 외(ex vivo) 또는 시험관 내(in vitro)에서 수행될 수 있다.
- [0009] 상기 세포는 체세포, 생식세포, 줄기세포, 암세포, 또는 세포주일 수 있다. 상기 세포는 암세포, 줄기세포, 혈관내피 세포, 백혈구, 면역 세포, 상피 세포, 생식 세포, 섬유아세포, 근육 세포, 골수 세포, 표피 세포, 골아세포, 및 신경세포로 이루어진 군으로부터 선택될 수 있다. 상기 세포는 인간을 포함한 동물, 식물, 세균, 또는 균류의 세포일 수 있다.
- [0010] 상기 유전체(genome)은 생명체 또는 세포가 갖는 모든 유전정보를 말한다.
- [0011] 상기 표적 핵산은 변형시키고자 하는 핵산을 의미한다.
- [0012] 세포의 유전체에서 표적 핵산을 변형시키는 방법은 유전체 편집(genome editing)으로도 불릴 수 있다.
- [0013] 상기 방법은 표적 핵산을 포함한 세포, Cas(clustered regularly interspaced palindromic repeats-associated: CRISPR associated) 뉴클레아제 또는 이의 변이체를 암호화하는 폴리뉴클레오티드, 및 가이드 RNA를 인큐베이션하는 단계를 포함한다.
- [0014] 상기 표적 핵산은 프로토스페이서-인접 모티프(protospacer adjacent motif: PAM), 및 상기 가이드 RNA에 상보적인 표적 서열을 포함할 수 있다.
- [0015] 상기 프로토스페이서-인접 모티프(PAM)는 표적 핵산에서 Cas 뉴클레아제가 인식하는 뉴클레오티드 서열일 수 있다. 상기 PAM은 5'-NA-3', 5'-NTG-3', 5'-GTH-3', 5'-VCGD-3', 5'-TCGDG-3', 5'-GGT-3', 5'-BGC-3', 5'-RTG-3', 5'-NYGG-3', 5'-AGCCA-3', 5'-CTGCA-3', 5'-TGTC-3', 5'-GCGCA-3', 5'-GCGCT-3', 5'-GCGTT-3', 5'-AGCCC-3', 5'-CGTCA-3', 5'-GCGTA-3', 5'-AGCCG-3', 5'-GCGTC-3', 5'-AGTCA-3', 5'-AGCAC-3' 및 5'-AGCTC-3'로 이루어진 군으로부터 선택될 수 있다. 상기 뉴클레오티드 서열에서, 'A'는 아데닌(adenine: A)을 의미하고, 'G'는 구아닌(guanine: G)을 의미하고, 'C'는 시토신(cytosine: C)을 의미하고, 및 'T'는 티민(thymine: T)을 의미한다. 'N'은 아데닌(A), 구아닌(G), 시토신(C), 티민(T), 우라실(uracil: U), 불명, 또는 기타의 핵산을 의미한다. 'H'는 구아닌(G)이 아닌 염기를 의미하고, 아데닌(A), 시토신(C), 티민(T), 또는 우라실(U)일 수 있다. 'V'는 티민(T) 및 우라실(U)이 아닌 염기를 의미하고, 아데닌(A), 구아닌(G), 또는 시토신(C)일 수 있다. 'D'는 시토신(C)이 아닌 염기를 의미하고, 아데닌(A), 구아닌(G), 티민(T), 또는 우라실(U)일 수 있다. 'B'는 아데닌(A)이 아닌 염기를 의미하고, 구아닌(G), 시토신(C), 티민(T), 또는 우라실(U)일 수 있다. 'R'은 퓨린(purine)을 의미하고, 아데닌(A) 또는 구아닌(G)일 수 있다. 'Y'는 피리미딘(pyrimidine)을 의미하고, 시토신(C), 티민(T), 또는 우라실(U)일 수 있다. 상기 5'-NA-3'는 5'-AA-3', 5'-GA-3', 5'-CA-3', 또는 5'-TA-3'이다. 5'-NTG-3'는 5'-ATG-3', 5'-GTG-3', 5'-CTG-3', 또는 5'-TTG-3'이다. 5'-GTH-3'는 5'-GTA-3', 5'-GTC-3', 또는 5'-GTT-3'이다. 5'-VCGD-3'는 5'-ACGA-3', 5'-ACGG-3', 5'-ACGT-3', 5'-GCGA-3', 5'-GCGG-3', 5'-GCGT-3', 5'-CCGA-3', 5'-CCGG-3', 또는 5'-CCGT-3'이다. 5'-TCGDG-3'는 5'-TCGAG-3', 5'-

TCGGG-3', 또는 5'-TCGTG-3'이다. 5'-BGC-3'는 5'-GGC-3', 5'-CGC-3', 또는 5'-TGC-3'이다. 5'-RTG-3'는 5'-ATG-3' 또는 5'-GTG-3'이다. 5'-NYGG-3'는 5'-ACGG-3', 5'-ATGG-3', 5'-GCGG-3', 5'-GTGG-3', 5'-CCGG-3', 5'-CTGG-3', 5'-TCGG-3', 또는 5'-TTGG-3'이다. 상기 PAM은 표적 서열의 3'-말단 또는 5'-말단에 바로 인접 또는 연속한 것일 수 있다.

[0016] 상기 Cas 뉴클레아제는 핵산 이중가닥을 절단하는 엔도뉴클레아제(endonuclease)일 수 있다. 상기 Cas 뉴클레아제는 RNA-가이드 DNA 엔도뉴클레아제(RNA guided DNA endonuclease)일 수 있다. 상기 Cas 뉴클레아제는 스트렙토코커스 속(*Streptococcus* sp.), 캄필로박터 속(*Campylobacter* sp.), 레지오넬라 속(*Legionella* sp.), 네이세리아 속(*Neisseria* sp.), 파스테우렐라 속(*Pasteurella* sp.), 프란시셀라(*Francisella* sp.), 및 프레보텔라 속(*Prevotella* sp.)으로 이루어진 군으로부터 선택된 세균으로부터 유래된 뉴클레아제일 수 있다. 상기 Cas 뉴클레아제는 스트렙토코커스 피오게네스(*Streptococcus pyogenes*), 스트렙토코커스 써모필러스(*Streptococcus thermophilus*), 스트렙토코커스 아우레우스(*Streptococcus aureus*), 캄필로박터 제주니(*Campylobacter jejuni*), 레지오넬라 뉴모필라(*Legionella pneumophila*), 네이세리아 메닝기티디스(*Neisseria meningitidis*), 파스테우렐라 물토시다(*Pasteurella multocida*), 프란시셀라 노비시다(*Francisella novicida*), 및 프레보텔라 디시엔스(*Prevotella disiens*)로 이루어진 군으로부터 선택된 세균으로부터 유래된 뉴클레아제일 수 있다. 상기 Cas 뉴클레아제는 Cas9, Cpf1, C2c1, C2c2, C2c3, Cas3, Cas5, Cas7, Cas8, 또는 Cas10일 수 있다. 상기 Cas9은 예를 들어, 스트렙토코커스 피오게네스 유래 Cas9(*Streptococcus pyogenes* Cas9: SpCas9)이다.

[0017] 상기 Cas 뉴클레아제의 변이체는 Cas 뉴클레아제의 기능을 보유한 변이체이다. 상기 Cas 뉴클레아제의 변이체는 xCas9, SpCas9-NG, Cas9 틸대기효소(Cas9 nickase: Cas9 nickase), 불활성화 Cas9(Deactivated Cas9: dCas9), 및 불안정화 Cas9(destabilized Cas9: DD-Cas9)로 이루어진 군으로부터 선택될 수 있다.

[0018] 상기 Cas 뉴클레아제 또는 이의 변이체를 암호화하는 폴리뉴클레오티드는 플라스미드 벡터 또는 바이러스 벡터에 함유된 것일 수 있다.

[0019] 상기 Cas 뉴클레아제의 변이체가 SpCas9-NG인 경우, 상기 PAM은 5'-NA-3', 5'-NTG-3', 5'-GTH-3', 5'-VCGD-3', 및 5'-TCGDG-3'로 이루어진 군으로부터 선택된 뉴클레오티드 서열로 이루어진 폴리뉴클레오티드일 수 있다. 상기 Cas 뉴클레아제가 스트렙토코커스 피오게네스 유래 Cas9(SpCas9)인 경우, 상기 PAM은 5'-GGT-3', 5'-BGC-3', 5'-RTG-3', 5'-NYGG-3', 5'-AGCA-3', 5'-CTGA-3', 5'-TGTA-3', 5'-GCGA-3', 5'-GCGT-3', 5'-GCGTT-3', 5'-AGCC-3', 5'-CGTA-3', 5'-GCGTA-3', 5'-AGCG-3', 5'-GCGTC-3', 5'-AGTA-3', 5'-AGAC-3' 및 5'-AGTC-3'로 이루어진 군으로부터 선택된 뉴클레오티드 서열로 이루어진 폴리뉴클레오티드일 수 있다.

[0020] 용어 "가이드(guide) RNA"는 유전체 편집을 통해 세포내에서 표적 핵산을 인식하여 표적 핵산을 절단, 삽입, 또는 연결시키는 폴리뉴클레오티드를 말한다. 상기 가이드 RNA는 표적 핵산 중 표적 서열에 상보적인 서열을 포함할 수 있다. 상기 가이드 RNA는 상기 표적 핵산에서 PAM의 5' 방향으로 또는 3' 방향으로 연속적인 2 내지 24 뉴클레오티드(이하, 'nt'라 함)의 뉴클레오티드 서열과 상보적인 폴리뉴클레오티드일 수 있다. 상기 가이드 RNA의 길이는 17 nt, 18 nt, 19 nt, 20 nt, 21 nt, 22 nt, 23 nt, 또는 24 nt일 수 있다. 상기 가이드 RNA는 단일-사슬 가이드 RNA(single-chain guide RNA: sgRNA)일 수 있다. 상기 sgRNA는 표적 핵산 서열에 특이적인 crRNA(CRISPR RNA) 및/또는 Cas 뉴클레아제와 복합체를 형성하는 tracrRNA를 포함할 수 있다. 상기 가이드 RNA는 플라스미드 벡터 또는 바이러스 벡터에 함유된 것일 수 있다.

[0021] 상기 인큐베이션은 표적 핵산을 포함하는 세포에 Cas 뉴클레아제 또는 이의 변이체를 암호화하는 폴리뉴클레오티드 및 가이드 RNA를 도입하는 것일 수 있다. 상기 도입은 통합(integration), 형질전환(transformation), 형질도입(transduction), 형질감염(transfection), 또는 이들의 조합을 포함한다. 상기 도입은 일시적(transient) 또는 안정한(stable) 도입일 수 있다.

[0022] 상기 표적 핵산은 상기 Cas 뉴클레아제 또는 이의 변이체, 및 상기 가이드 RNA의 복합체에 의해 인식될 수 있다.

[0023] 상기 Cas 뉴클레아제 또는 이의 변이체, 및 가이드 RNA의 복합체는 표적 핵산을 서열-특이적으로 변형될 수 있다. 상기 변형은 삽입, 절단, 삽입, 연결, 탈아미노화, 또는 이들의 조합일 수 있다. 상기 절단은 유전체 DNA의 이중가닥의 절단일 수 있다. 상기 절단은 평활 말단(blunt end) 또는 점착 말단(sticky end)일 수 있다. 상기 변형은 표적 핵산의 절단 및 절단 부위에 외래 폴리뉴클레오티드의 삽입일 수 있다. 외래 폴리뉴클레오티드를 유전체의 절단 부위에 삽입하는 것은 상동성 의존적 방법에 의한 것일 수 있다. 상기 상동성 의존적 방법은 상동 재조합(homologous recombination) 또는 상동성-지향적 수선(homology-directed repair: HDR)일 수 있다.

[0024] 상기 방법은 시험관 내(in vitro), 생체 외(ex vivo), 또는 생체 내(in vivo)에서 수행될 수 있다.

[0025] 다른 양상은 일 양상의 세포의 유전체에서 표적 핵산을 변형시키는 방법으로 유전체의 표적 핵산이 변형된 세포를 제공한다.

[0026] 상기 세포, 유전체, 표적 핵산, 및 변형은 전술한 바와 같다.

발명의 효과

[0027] 신규한 PAM 서열을 이용하는 세포의 유전체에서 표적 핵산을 변형시키는 방법 및 이에 의해 유전체의 표적 핵산이 변형된 세포에 의하면, 기존에 표적으로 하지 못한 위치를 유전체 편집의 표적으로 하여 유전체 편집을 수행할 수 있고, 이에 의해 유전체 편집의 활용 범위를 확대시킬 수 있다.

도면의 간단한 설명

[0028] 도 1a는 고정된 프로토스페이서를 이용하여 PAM 서열을 결정하기 위한 실험 과정을 나타낸 개략도이고, 도 1b는 PAM 서열 결정을 위한 실험 디자인의 개략도이다.

도 2a 내지 2c는 xCas9 뉴클레아제의 PAM 서열에 따른 인텔 빈도(%)를 나타내 그래프이다.

도 3a 내지 3c는 SpCas9-NG 뉴클레아제의 PAM 서열에 따른 인텔 빈도(%)를 나타내 그래프이다.

도 4a 내지 4c는 SpCas9 뉴클레아제의 PAM 서열에 따른 인텔 빈도(%)를 나타내 그래프이다.

도 5a 내지 5c는 각각 xCas9, SpCas9-NG, 및 SpCas9 뉴클레아제의 PAM 서열에서 첫번째 내지 5번째 위치에 따른 인텔 빈도(%)를 나타내는 히트 맵이다.

도 6a 내지 6c는 각각 xCas9, SpCas9-NG, 및 SpCas9 뉴클레아제의 PAM 서열에 따른 인텔 빈도(%)를 나타내 그래프이다.

도 7a 내지 7c는 PAM 서열에 대해 각각 xCas9과 SpCas9의 평균 인텔 빈도(%), SpCas9-NG와 SpCas9의 평균 인텔 빈도(%), 및 xCas9과 SpCas9-NG의 평균 인텔 빈도(%)를 나타내 그래프이다.

발명을 실시하기 위한 구체적인 내용

[0029] 이하 실시예를 통하여 보다 상세하게 설명한다. 그러나, 이들 실시예는 하나 이상의 구체예를 예시적으로 설명하기 위한 것으로 본 발명의 범위가 이들 실시예에 한정되는 것은 아니다.

[0030] 실시예 1. xCas9, SpCas9-NG, 및 SpCas9 PAM 서열의 확인

[0031] 1. 가이드 RNA, PAM 서열, 및 표적 서열을 함유한 플라스미드 라이브러리의 준비

[0032] 우선, Cas 뉴클레아제에 의해 인식되는 PAM 서열 확인을 위해, Twist Bioscience Co.에 의뢰하여 올리고뉴클레오타이드 라이브러리를 제작하였다.

[0033] 각 올리고뉴클레오타이드는, 5' 말단으로부터 19- 또는 20-nt의 sgRNA, BsmBI 제한효소 부위, 바코드 1(20-nt 서열), 제2 BsmBI 제한효소 부위, 바코드 2(15-nt 서열), 및 PAM을 함유하는 30-nt 표적 서열을 포함하도록 디자인되었다. 특히, 올리고뉴클레오타이드는 79 종의 상이한 5-뉴클레오타이드 PAM 서열(=NNNAT 카테고리에서 64 + 16 in the AGGNN 카테고리에서 16 - 중복 AGGAT)을 포함하고, 80 종의 GN₁₉ 단일 가이드 RNA(single guide RNA: sgRNA) 서열을 포함하므로, 6,320 (=79 x 80)종의 표적 서열을 표적으로 하는 플라스미드 라이브러리이다(도 1b).

[0034] sgRNA와 표적 서열을 함유하는 플라스미드 라이브러리는 2단계 클로닝 방법을 이용하여, 올리고뉴클레오타이드 풀(pool)의 PCR 증폭 과정에서 가이드 RNA와 표적 서열간의 언커플링(uncoupling)을 방지하였다.

[0035] 제1 단계는 가이드 RNA 및 표적 서열을 포함하는 초기 플라스미드 라이브러리 제작이다. BsmBI 제한효소(NEB)를 사용하여 Lenti-gRNA-Puro 플라스미드(Addgene, #84752)를 선형화시키고, 선형화된 벡터에 PCR-증폭된 올리고뉴클레오타이드 풀(표적 서열)을 라이게이션시켰다. 반응물을 대장균에 형질전환시키고, 선별된 콜로니로부터 플라스미드를 분리하였다. 올리고뉴클레오타이드 풀 증폭에 사용된 프라이머는 다음과 같다.

[0036] 정방향 프라이머:

- [0037] 5'-TTGAAAGTATTTTCGATTTCTTGGCTTTATATATCTTGTGGAAGGACGAAACACC-3' (서열번호 1)
- [0038] 역방향 프라이머:
- [0039] 5'-GAGTAAGCTGACCGCTGAAGTACAAGTGGTAGAGTAGAGATCTAGTTACGCCAAGCT-3' (서열번호 2)
- [0040] 제2 단계는 sgRNA 스캐폴드의 삽입이다. 제1 단계에서 제작된 플라스미드 라이브러리는 BsmBI 제한효소(NEB)로 절단하고 아가로스 전기영동 후 겔에서 핵산 단편을 정제하였다. sgRNA 스캐폴드를 함유한 인서트 단편을 합성하고 TOPO 벡터(T-두가닥(blunt) 벡터, Solgent)에 클로닝하였다. 인서트 단편의 서열을 하기에 나타내었고, 폴리 T 서열을 포함한 sgRNA는 굵은 글씨로 표시하고, BsmBI 제한효소 부위는 밑줄로 표시하였다.
- [0041] 5'-CGTCTCTGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAG
- [0042] TGGCACCGAGTCGGTGTCTTTTTGGGAGACG-3' (서열번호 3)
- [0043] 인서트 단편을 함유한 TOPO 벡터를 BsmBI 제한효소(NEB)로 절단하여 83-nt 인서트 단편을 분리하였다. 절단된 플라스미드 라이브러리에 인서트 단편을 라이게이션시키고, 대장균에 형질전환시킨 후, 선별된 콜로니로부터 플라스미드 라이브러리를 분리하였다.
- [0044] **2. 세포 라이브러리의 준비**
- [0045] 우선, 렌티바이러스 라이브러리 생산을 위해, 인간 배아 신장 세포인 HEK293T 세포(ATCC)를 준비하였다. 실시예 1.1에서 준비된 플라스미드, psPAX2, 및 pMD2.G를 혼합하고, 리포펙타민 2000 (Invitrogen)을 사용하여 HEK293T 세포에 형질감염시켰다. 형질감염 후 12 시간에 신선한 배지를 세포에 가하고, 형질 감염 후 36 시간에 바이러스를 함유한 상층액을 수득하였다. 수득된 상층액은 Millex-HV 0.45 μ m 저-단백질 결합 멤브레인(Millipore)으로 여과하고, 분액은 사용시까지 -80℃에 보관하였다. 바이러스 수율은 Lenti-X p24 Rapid Titer Kit(Clontech)로 측정하여 검증하였다. 바이러스 역가 산출을 위해, 순차 희석된 바이러스 분액을 8 μ g/ml의 폴리브렌(polybrene)의 존재에서 HEK293T 세포에 형질도입하고, 2 μ g/ml 퓨로마이신 또는 20 μ g/ml 블라스티시딘(blasticidin) S (InvivoGen)의 존재에서 배양하여 산출하였다.
- [0046] 준비된 렌티바이러스 라이브러리의 형질도입을 위해, HEK293T 세포를 배양 접시에 밤새 배양하였다. 감염다중도(multiplicity of infection: MOI) 0.3의 렌티바이러스 라이브러리를 8 μ g/ml의 폴리브렌의 존재에서 HEK293T 세포에 형질도입하고, 세포를 밤새 배양하였다. 세포를 2 μ g/ml 퓨로마이신의 존재에서 배양하여 형질도입되지 않은 세포를 제거하고, 1.2×10^7 세포의 양으로 세포 라이브러리를 유지하였다.
- [0047] **3. 세포 라이브러리에 Cas9의 전달**
- [0048] 1.2×10^7 세포의 양의 세포 라이브러리를 준비하고, xCas9-, SpCas9-NG, 및 SpCas9-암호화하는 바이러스를 8 μ g/ml의 폴리브렌의 존재에서 세포에 형질도입시켰다. 형질도입은 MOI 5로 수행하였고, 20 μ g/ml 블라스티시딘 S의 존재 하에서 세포를 선별하였다.
- [0049] **4. 인델 빈도의 측정**
- [0050] 실시예 1.3에서 준비된 세포에서 유전체 중 인델(삽입/결실, insertion/deletion: indel)의 빈도를 측정하기 위해, 고심도(deep) 시퀀싱 및 인델 빈도의 분석을 수행하였다.
- [0051] 고심도 시퀀싱을 위해, Wizard Genomic DNA purification kit(Promega)를 사용하여 세포로부터 유전체 DNA를 분리하였다. 삽입된 표적 사열은 고성능 실험을 위해 2X Taq PCR Smart mix(Solgent)를 사용하여 PCR 증폭하였다. 첫번째 PCR에서, 각 세포 라이브러리에 대해 총 240 μ g의 유전체 DNA를 이용하여 라이브러리에 대해 1000x 이상의 커버리지가 되도록 하였다(10^6 개의 세포 당 약 10 μ g의 유전체 DNA). 각 반응 당 2.5 μ g의 유전체 DNA를 1차 PCR하고, 반응산물을 모두 풀링(pooling)한 후 정제하였다. 정제된 산물 중 50 ng의 시료를 Illumina 어댑터 및 바코드 서열을 함유한 프라이머를 이용하여 2차 PCR 증폭하였다. 증폭된 산물은 전기영동 후 정제하고, HiSeq 또는 MiniSeq(Illumina)를 이용하여 분석하였다.
- [0052] 실험에 사용된 프라이머는 다음과 같다.
- [0053] 1차 PCR 반응용 프라이머
- [0054] 정방향 프라이머:

[0055] 5'-ACACTCTTTCCCTACACGACGCTCTTCCGATCTCTTGAAGTGGCACCAGTTCG-3' (정방향 프라이머, 서열번호 4)

[0056] 5'-ACACTCTTTCCCTACACGACGCTCTTCCGATCTCTTGAAGTGGCACCAGTTCG-3' (서열번호 5)

[0057] 5'-ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGCTTGAAGTGGCACCAGTTCG-3' (서열번호 6)

[0058] 역방향 프라이머:

[0059] 5'-GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTAAAGTCGAGTAAGCTGACCGCTGAAG-3' (서열번호 7)

[0060] 5'-GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTATTAAGTCGAGTAAGCTGACCGCTGAAG-3' (서열번호 8)

[0061] 5'-GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTATTAAGTCGAGTAAGCTGACCGCTGAAG-3' (서열번호 9)

[0062] 2차 PCR 반응용 프라이머

[0063] 정방향 프라이머:

[0064] 5'-AATGATACGGCGACCGAGATCTCAC (index) ACACTCTTTCCCTACACGAC-3' (서열번호 10)

[0065] 역방향 프라이머:

[0066] 5'-CAAGCAGAAGACGGCATACGAGAT (index) GTGACTGGAGTTCAGACGTGT-3' (서열번호 11)

[0067] 인텔 빈도의 분석을 위해 Python scripts 프로그램을 변형하여 고심도 시퀀싱 데이터를 분석하였다. 15-nt의 바코드와 그의 상류에 위치한 4-nt의 총 19-nt의 서열을 이용하여 각 가이드 RNA와 표적 서열의 쌍을 식별하였다. 인텔이 예상된 절단 부위(즉, 그 부위의 가운데에 위치한 8-nt 영역)에 위치해 있을 경우, 인텔은 xCas9-, SpCas9-NG-, 또는 SpCas9-유도된 돌연변이라고 간주하였다. 어레이 합성 및 PCR 증폭의 결과인 배경 인텔 빈도를 제거하기 위해, 총 리드(read), 인텔 리드, 및 Cas9을 도입하지 않은 경우의 인텔 빈도를 산출하고, 하기 수식에 따라 인텔 빈도(%)를 산출하였다.

$$\text{인텔 빈도(}\%) = \frac{\text{인텔 리드} - (\text{총 리드의 수} \times \text{배경 인텔 빈도})}{\text{총 리드} - (\text{총 리드의 수} \times \text{배경 인텔 빈도})} \times 100$$

[0068] (수학식 1)

[0069] 고심도 시퀀싱 데이터는 NCBI Sequence Read Archive(SRA; www.ncbi.nlm.nih.gov/sra/)에 Accession No. SRP158724에 업로드하였다.

[0070] 5. PAM 서열의 결정

[0071] 실시예 1.4에 따라 수득된 고심도 시퀀싱 데이터에서, xCas9, SpCas9-NG, 또는 SpCas9 뉴클레아제에서 높은 인텔 빈도를 갖는 PAM 서열을 선별하였다.

[0072] 사람 세포에서 형질 도입 후 6일까지 5% 이상의 평균 인텔 빈도를 갖는 PAM 서열을 선별하고, xCas9, SpCas9-NG, 및 SpCas9 뉴클레아제와 PAM 서열에 따른 인텔 빈도를 분석하였다.

[0073] xCas9 뉴클레아제와 PAM 서열에 따른 인텔 빈도(%)는 도 2a 내지 도 2c에 나타내고, SpCas9-NG 뉴클레아제와 PAM 서열에 따른 인텔 빈도(%)는 도 3a 내지 도 3c에 나타내고, SpCas9 뉴클레아제와 PAM 서열에 따른 인텔 빈도(%)는 도 4a 내지 도 4c에 나타내었다(ANOVA 분석 후 Tukey 사후 검정; *: P<0.05, **: P<0.01, ***: P<0.001; AGGTC에 대해, ##: P<0.01, ###: P<0.001; AGGGC에 대해, §: P<0.001, §§§: P<0.001; AGGAC에 대해 ‡: P<0.05, ‡‡: P<0.01, ‡‡‡: P<0.001). 각 그룹의 서열 이름에서, 가변 뉴클레오티드는 굵은색 검정 글씨로 표시하고, 불변 뉴클레오티드는 회색 글씨로 표시하였다. 박스 플롯에서, 위, 중간, 및 아래의 선은 각각 25번째, 50번째, 및 75번째 백분위를 표시하고, 수염(whisker)은 10번째 및 90번째 백분위를 나타내고, 이상점(outlier)은 점(dot)으로 나타내었다. 개수(N)은 PAM 서열 당 67 내지 79개의 표적 서열이다.

[0074] 도 2a 내지 도 4c에 나타난 바와 같이, PAM 서열은 몇개의 그룹으로 나뉘었다. xCas9, SpCas9-NG, 및 SpCas9 뉴클레아제에서, NGG PAM 서열은 일반적으로 가장 높은 인텔 빈도를 나타내었다.

[0075] xCas9 뉴클레아제에 대해서는, 인텔 빈도가 높은 순서대로, AGGCA, AGGCB (즉, AGGC(C, G, 또는 T)), 또는 NGGNN의 PAM 서열을 갖는 표적 서열에서 높은 인텔 빈도를 유의하게 유도하였다(도 2a 내지 2c). 이외에도 통계적으로 유의한 차이가 있는 6개의 서브세트를 확인하였다. 또한, NGGNN의 PAM 서열의 첫번째, 4번째, 및 5번째 뉴클레오티드 모두 xCas9-유도성 인텔 빈도에 영향이 있었고, 4번째 뉴클레오티드가 C인 경우 가장 높은 활성을

나타내었다(도 2b 및 2c). 4번째 뉴클레오티드가 C인 경우, 5번째 뉴클레오티드가 A 일 때 가장 높은 xCas9 활성이 관찰되었고, ii) 첫번째 위치에서 가장 높은 xCas9 활성을 유도한 뉴클레오티드는 A보다는 S(즉, G 또는 C)였다. 따라서, 이 그룹에서 가장 강력한 xCas9 PAM 서열은 SGGCA일 것으로 예상되었다. xCas9에 대한 대표적인 비-NGG PAM은 NGW (즉, NG (A 또는 T)였고, BGC (또는 더 광범위하게 NGC) 및 SAG (또는 더 광범위하게 NAG)도 PAM으로서 기능을 하였다.

[0076] SpCas9-NG에서, NGT는 비-NGG PAM 서열로서, NGG PAM 서열의 서브세트인 AGGCN PAM에 비해 높은 인텔 빈도를 나타내었다(도 3a). 또한, NGGNN(NGGDN으로 약칭) PAM을 갖는 표적 서열이 AGGCN PAM에 비해 유의하게 높은 인텔 빈도를 유도하였다. NGGDN PAM 서열의 세트 중에서, SpCas9-NG 유도성 인텔 빈도에서 통계적으로 유의한 4개의 서브세트가 확인되었다. NGGNN의 PAM 서열의 첫번째, 4번째, 및 5번째 뉴클레오티드 중 4번째 뉴클레오티드만이 SpCas9-NG-유도성 인텔 빈도에 영향이 있었고, 4번째 뉴클레오티드가 D(C 제외)인 경우 높은 활성을 나타내었다(도 3b 및 3c). SpCas9-NG에 대해 강력한 비-NGG PAM은 NGT였고, NGA, NGC, NGN, NAG, NAH, NTG, GTH 및 VCG도 PAM으로서 기능을 하였다.

[0077] SpCas9 뉴클레아제는, xCas9 및 SpCas9-NG과 달리, NGGDN PAM 서열에 따른 순차적인 인텔 빈도는 나타나지 않았다. 분석된 NGGDN PAM 서열 모두 유사한 SpCas9-유도성 인텔 빈도를 나타내었다(도 4c). NGGNN의 PAM 서열의 첫번째, 4번째, 및 5번째 뉴클레오티드 중 4번째 뉴클레오티드만이 SpCas9-유도성 인텔 빈도에 영향이 있었고, 4번째 뉴클레오티드가 H(G 제외)인 경우 높은 활성을 나타내었다(도 4b 및 4c). SpCas9에 대해 강력한 비-NGG PAM은 NAG 및 NGA였고, GGT, BGC 및 RTG도 PAM으로서 기능을 하였다.

[0078] 6. 광범위의 PAM 서열 및 프로토스페이서를 이용한 PAM 분석

[0079] 실시예 1.5는 80개의 고정된 동일 프로토스페이스 서열을 이용하여 79 종의 PAM 서열을 분석하였다. 더욱 광범위하게 PAM 서열을 분석하기 위해, 실시예 1.1에 기재된 올리고뉴클레오티드 라이브러리에서 약 13,200개가 준비되었다.

[0080] 실시예 1.1 내지 1.4에 기재된 바와 같이, 올리고뉴클레오티드 라이브러리로부터 세포 라이브러리를 준비하고, 세포 라이브러리에 Cas9의 전달한 후, 인텔 빈도를 측정하였다. xCas9, SpCas9-NG, 또는 SpCas9 뉴클레아제에서, PAM 서열에서 첫번째 내지 5번째 뉴클레오티드에 따른 인텔 빈도를 산출하였다. 분석에서 제외된 PAM 서열은 회색으로 표시하였고, PAM 서열 당 표적 서열의 수는 3 내지 306개(평균 33개)였다.

[0081] xCas9, SpCas9-NG, 및 SpCas9 뉴클레아제의 PAM 서열에서 첫번째 내지 5번째 위치에 따른 인텔 빈도(%)를 나타내는 히트맵(heat map)을 각각 도 5a 내지 5c에 나타내고; xCas9, SpCas9-NG, 및 SpCas9 뉴클레아제의 PAM 서열에 따른 인텔 빈도(%)를 각각 도 6a 내지 6c에 나타내고; 및 PAM 서열에 대해 각각 xCas9과 SpCas9의 평균 인텔 빈도(%), SpCas9-NG와 SpCas9의 평균 인텔 빈도(%), 및 xCas9과 SpCas9-NG의 평균 인텔 빈도(%)를 나타내는 그래프를 도 7a 내지 7c에 나타내었다. 도 7a 내지 7c에서, 빨간색 점선은 $y=x$ 을 나타내고, x 축과 y 축의 눈금은 0%에서 10%까지는 직선으로 나타내고, 10%에서 100%까지는 지수로 나타내었다.

[0082] 592종(=37종 3 뉴클레오티드 PAM 서열 $\times 4^2$)의 5-뉴클레오티드 PAM 서열은 실시예 1.5의 분석에서 제시된 가능한 모든 종류의 PAM을 포함한다. 592종의 PAM 서열 중 7종은 표적 서열의 무작위적 선택 및 낮은 시퀀싱 리드 카운트 때문에 분석에서 제외하였다. 남은 585종의 5-뉴클레오티드 PAM 서열에서, PAM 서열 당 표적 서열의 평균은 33종이었다.

[0083] 도 5a 및 6a에 나타난 바와 같이, xCas9 뉴클레아제의 경우, 평균 인텔 빈도가 10%를 넘는 상위 95종의 5-뉴클레오티드 PAM 서열 중 93종이 NGG, NGW, SAG 및 BGC PAM 군에 포함되었다. 나머지 2종의 PAM 서열은 AGCCA 및 CAACG였다. 이러한 결과는 세균성 세포에서 AGC 및 CAA가 xCas9 뉴클레아제에 대한 PAM으로 기능할 수 있다는 기존 결과와 일치한다(Hu, J.H. et al, Nature 556, 57-63 (2018)). 가장 높은 평균 인텔 빈도를 유도한 PAM 서열은 GGGCA, CGGCA, AGGCA 및 TGGCA이었고, 각각 평균 39%, 34%, 33%, 32%의 인텔 빈도를 유도하여, (C 또는 G)GGCA가 가장 강력한 PAM 서열임을 제시한다. 한편, 상위 95종의 5-뉴클레오티드 PAM 서열에서 xCas9에 대한 비-NGG PAM 서열은 NGW, SAG 및 BGC이었고, 대부분 4번째 위치의 뉴클레오티드가 C이었고, NGGC, NGWC, SAGC 및 BGCC는 각각 NGGD, NGWD, SAGD 및 BGCD에 비해 더 높은 평균 인텔 빈도를 유도하였다. 이외에, AGCCA(12%), CAACG(11%), GAACA(9.7%)를 포함하여 5% 이상의 인텔 빈도를 유도하는 PAM 서열이 확인되었다. 또한, xCas9 뉴클레아제에 대해, 5'-NGD-3', 5'-SAG-3', 및 5'-BGCC-3'의 PAM 서열을 확인하였다.

[0084] 도 5b 및 6b에 나타난 바와 같이, SpCas9-NG 뉴클레아제의 경우, 평균 인텔 빈도가 10%를 넘는 상위 455종의 5-뉴클레오티드 PAM 서열은 모두 NGN, NAN, NTG, GTH 및 VCG PAM 군에 포함되었다. xCas9과 달리, 4번째 뉴클레

오티드가 C인 경우 D에 비해 평균 SpCas9-NG-유도성 인텔 빈도가 낮았다(평균 인텔 빈도; NGND vs. NGNC, 32% 대 28%; NAND vs. NANC, 17% 대 13%; NTGD 대 NTGC, 14% 대 8.9%; GTHD 대 GTHC, 10% 대 10%; VCGD 대 VCGC, 7.5% 대 4.2%). TCGDG 서열(평균 인텔 빈도 8.4%)도 PAM으로서 기능하는 것으로 확인되었다. 또한, SpCas9-NG에 대해, 5'-NA-3', 5'-NTG-3', 5'-GTH-3', 5'-VCGD-3', 및 5'-TCGDG-3'의 신규한 PAM 서열을 확인하였다.

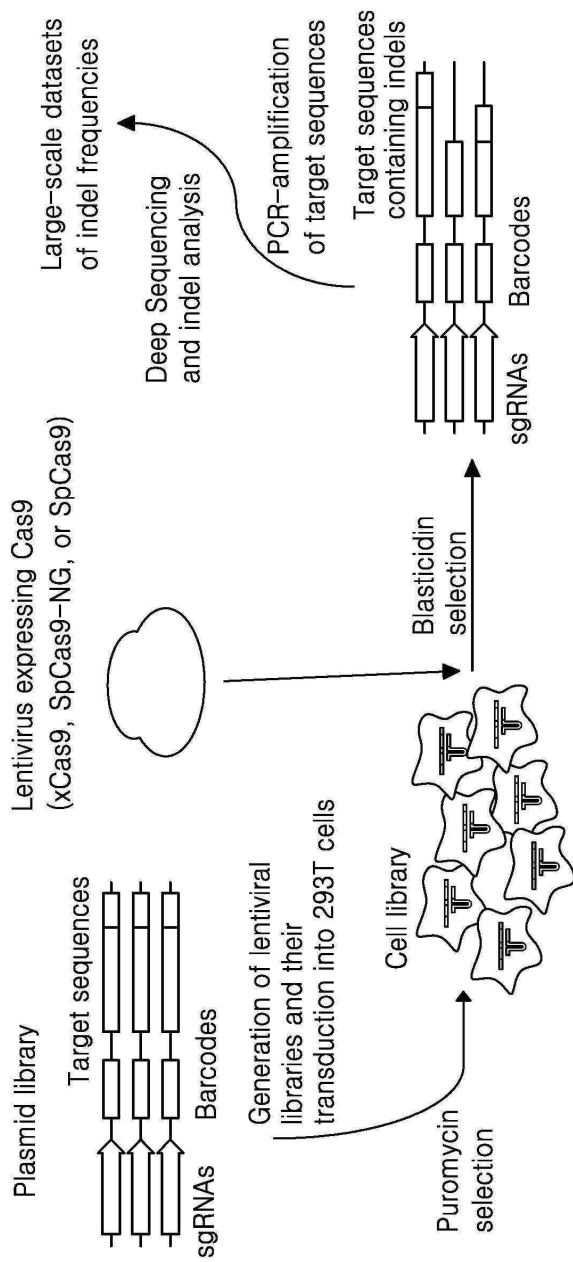
[0085] 도 5c 및 6c에 나타난 바와 같이, SpCas9 뉴클레아제의 경우, 평균 인텔 빈도가 10%를 넘는 상위 243종의 5-뉴클레오티드 PAM 서열 중 241종은 NGG, NAG, NGA, GGT, BGC, RTG 및 NYGG PAM 군에 포함되었다. 상위 243종에서 남은 2종의 PAM 서열은 AGCCA와 CTGCA였고, 각각 16 %와 10 %의 삽입 빈도를 유도하였다. NGG 카테고리의 PAM이 상위 1위에서 64위의 순서였고, 나머지 PAM과 명확히 구별되어, NGG가 SpCas9에 대해 가장 유효한 PAM 서열임을 확인하였다. xCas9 및 SpCas9-NG와 달리, SpCas9은 RTG 및 NYGG를 제외하고, 4번째 위치에서 특정 뉴클레오티드에 대한 선호도가 나타나지 않았다. 또한, SpCas9은 5% 초과인 SpCas9-유도성 인텔 빈도를 유도하는 PAM 서열로 AGCCA(16%), CTGCA(10%), TGTC(8.0%), GCGCA(7.5%), 및 AGCTC(5.1%) 등의 14종이 확인되었다. 또한, SpCas9에 대해, 5'-GGT-3', 5'-BGC-3', 5'-RTG-3', 5'-NYGG-3', 5'-AGCCA-3', 5'-CTGCA-3', 5'-TGTC-3', 5'-GCGCA-3', 5'-GCGCT-3', 5'-GCGTT-3', 5'-AGCCC-3', 5'-CGTCA-3', 5'-GCGTA-3', 5'-AGCCG-3', 5'-GCGTC-3', 5'-AGTCA-3', 5'-AGCAC-3', 및 5'-AGCTC-3'의 신규한 PAM 서열을 확인하였다.

[0086] 도 7a 내지 7c에 나타난 바와 같이, SpCas9은 5'-NGG-3' PAM 서열에 가장 높은 활성을 나타내었고, xCas9은 SpCas9에 비해 5'-GGTC-3' 및 5'-HGT-3' PAM 서열에 더 높은 활성을 가졌다. SpCas9-NG는 SpCas9에 비해 5'-NYGG-3' 및 5'-VCGG-3'를 제외한 대부분의 비-NGG PAM에 더 높은 활성을 가졌다. xCas9과 비교하여, SpCas9-NG는 5'-NGGC-3'를 제외하고 모든 분석된 PAM 서열에 더 높은 활성을 나타내었다.

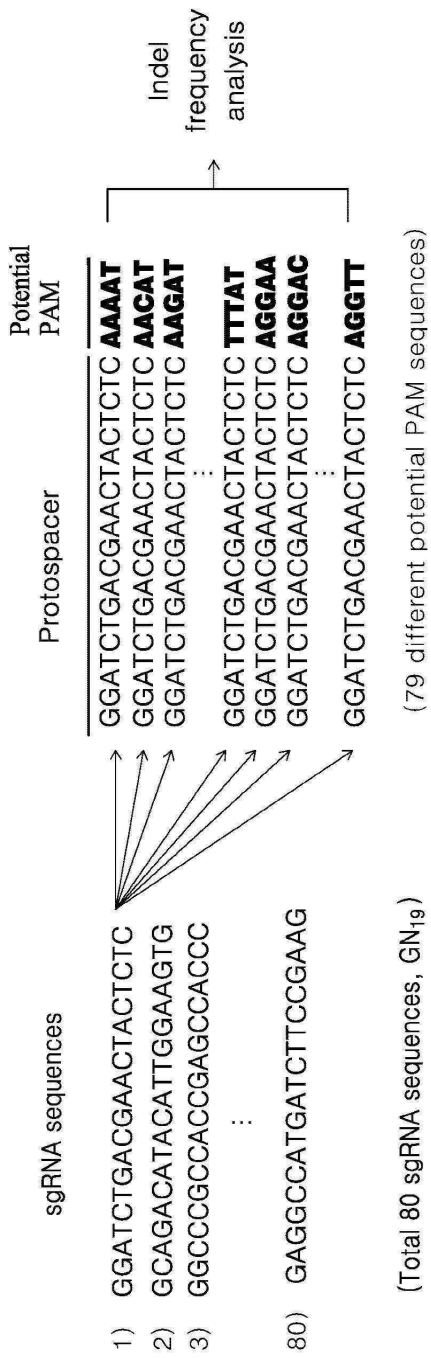
[0087] 따라서, 신규한 PAM 서열을 이용하여 인간 유전체에서 표적가능한 위치를 발굴함하고, PAM 서열에 따른 뉴클레아제의 인텔 유도 활성을 예측함으로써, 유전체 편집의 활용 가능성을 증가시킬 수 있었다.

도면

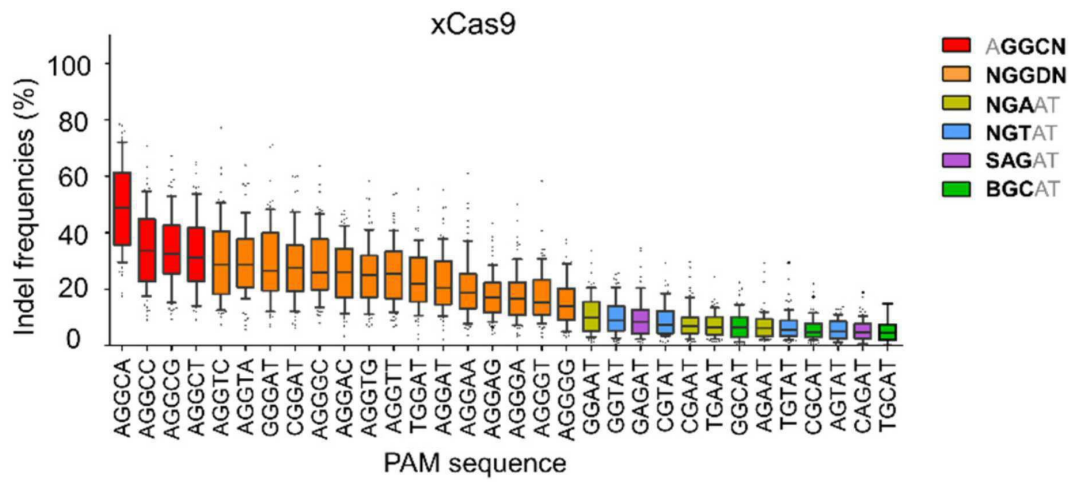
도면1a



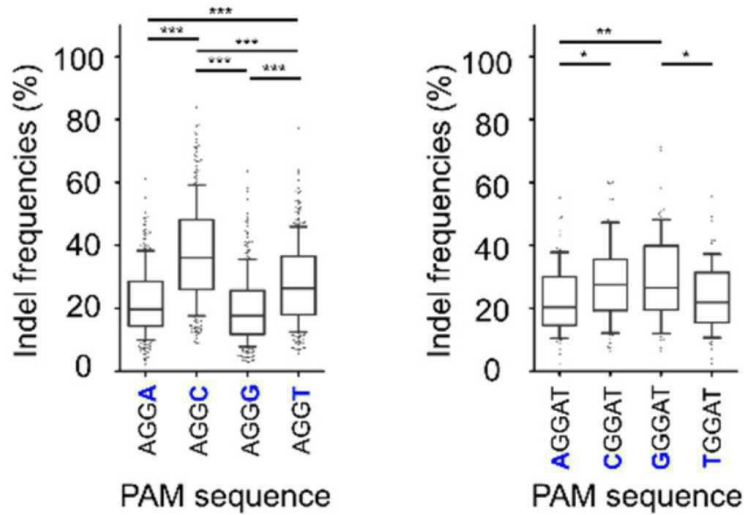
도면1b



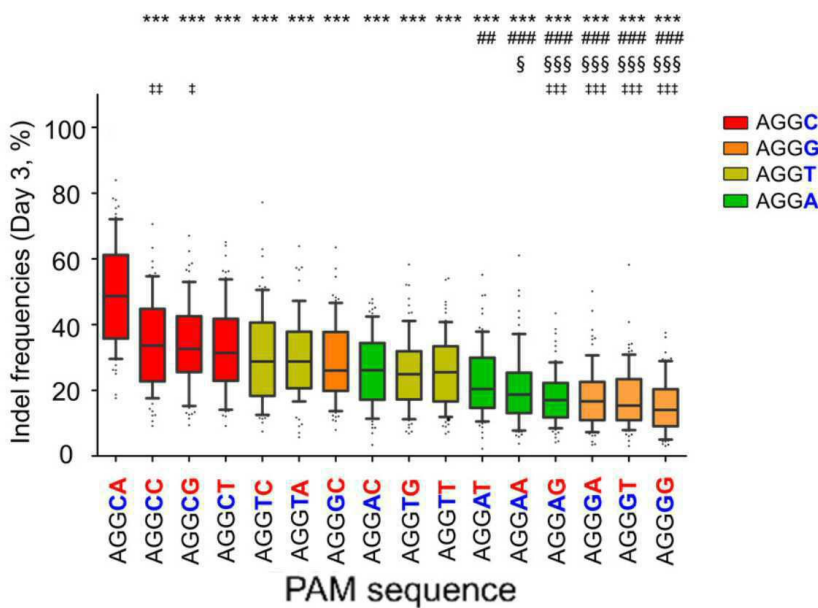
도면2a



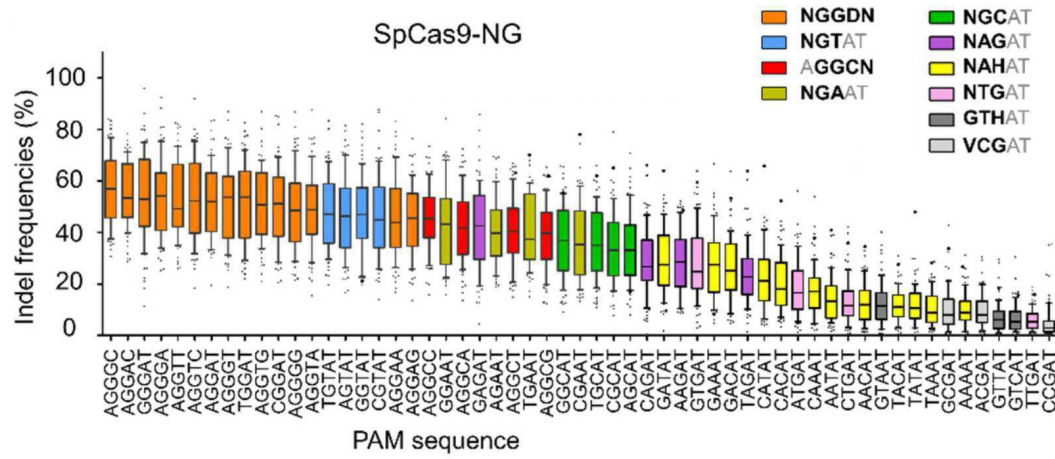
도면2b



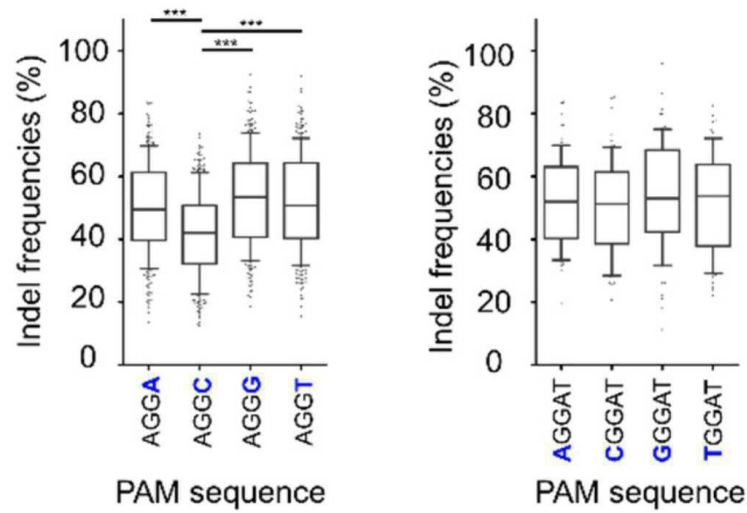
도면2c



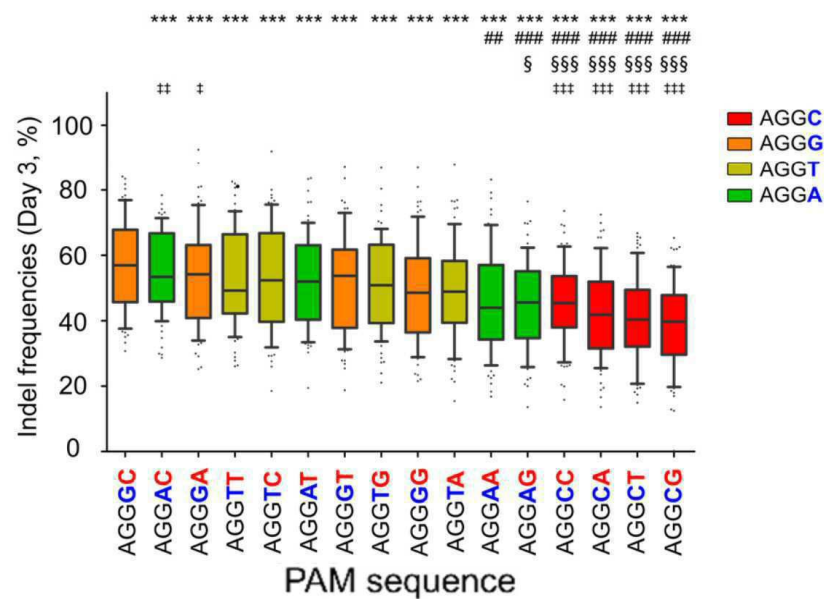
도면3a



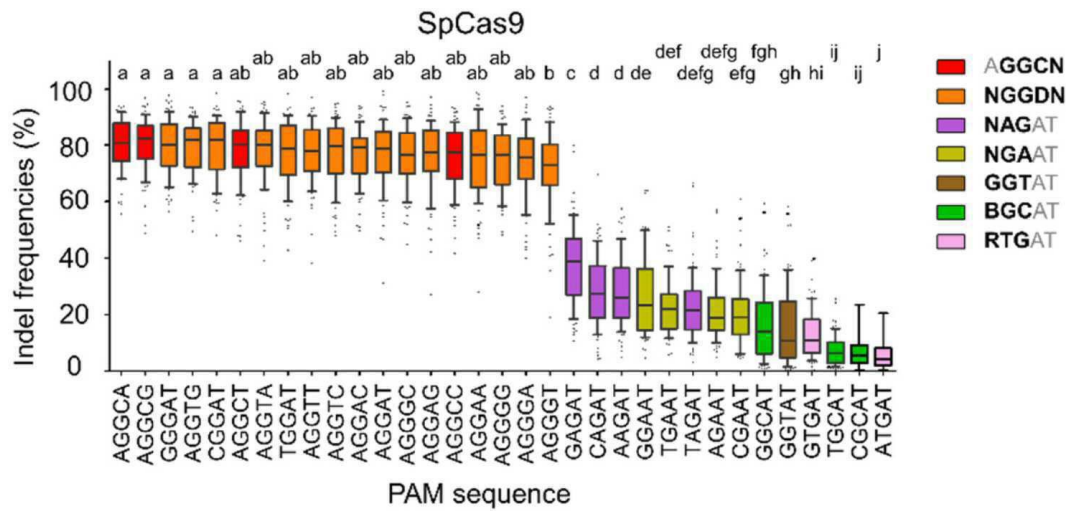
도면3b



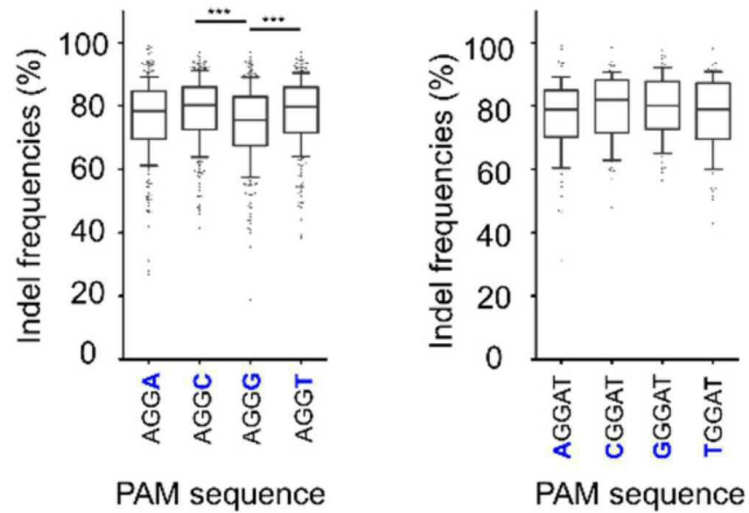
도면3c



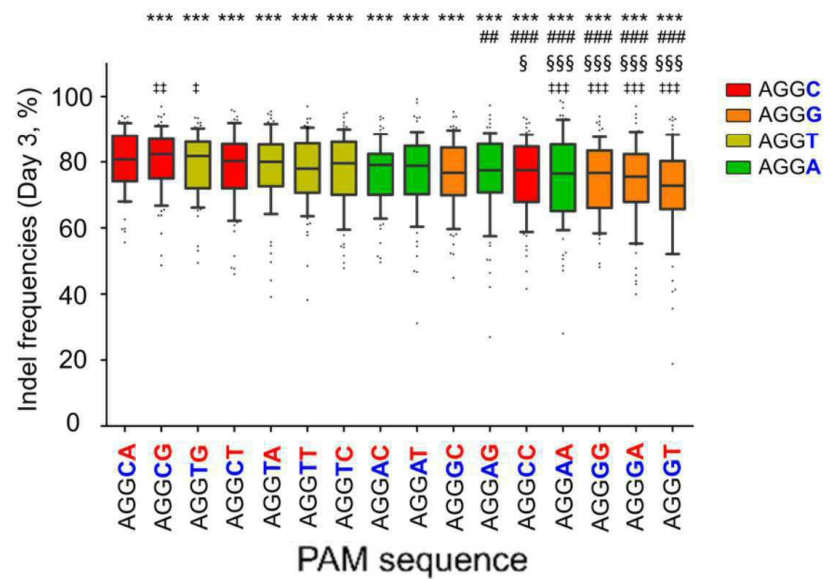
도면4a



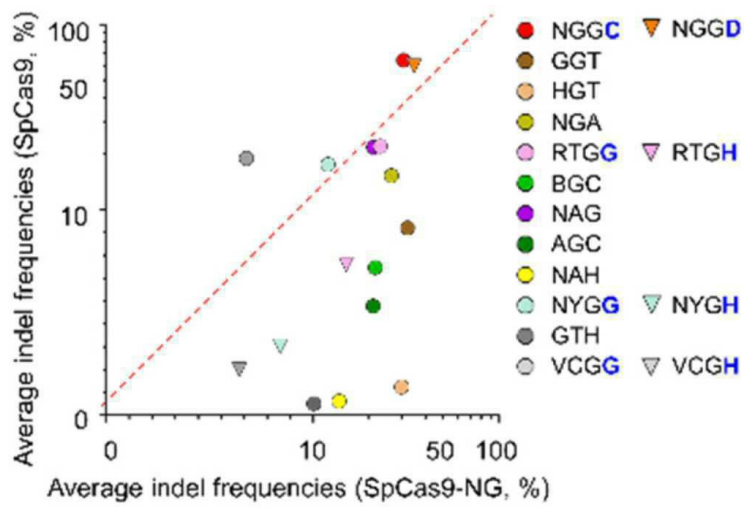
도면4b



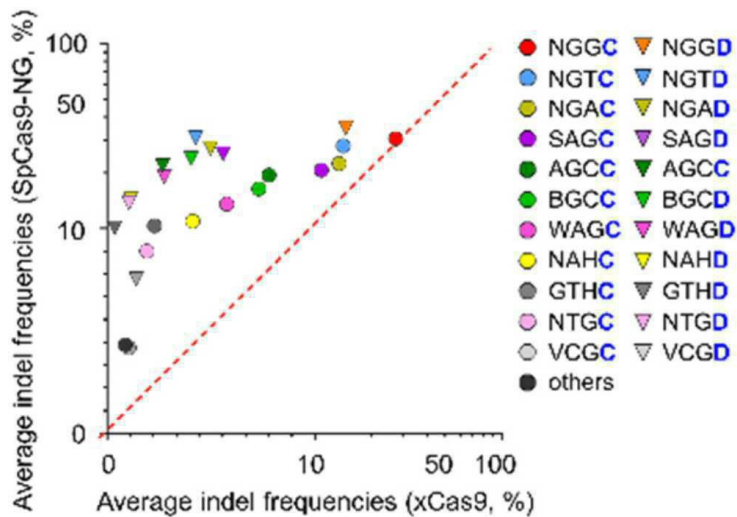
도면4c



도면7b



도면7c



서열 목록

- <110> Industry-Academic Cooperation Foundation, Yonsei University
- <120> Method for modifying a target nucleic acid in the genome of a cell
- <130> PN144218KR
- <150> KR 10-2018-0103979
- <151> 2019-08-31
- <150> KR 10-2019-0000391
- <151> 2019-01-02
- <160> 11
- <170> KoPatent In 3.0
- <210> 1

<211> 55
 <212> DNA
 <213> Artificial Sequence
 <220><223> Forward primer for Oligonucleotide pool amplification
 <400> 1
 ttgaaagtat ttcgatttct tggctttata tatcttgtgg aaaggacgaa acacc 55

<210> 2
 <211> 57
 <212> DNA
 <213> Artificial Sequence
 <220><223> Reverse primer for Oligonucleotide pool amplification
 <400> 2
 gagtaagctg accgctgaag tacaagtggt agagtagaga tctagttacg ccaagct 57

<210> 3
 <211> 97
 <212> DNA
 <213> Artificial Sequence
 <220><223> Insert fragment containing the sgRNA scaffold
 <400> 3
 cgtctctgtt ttagagctag aaatagcaag ttaaaataag gctagtccgt tatcaacttg 60
 aaaaagtggc accgagtcgg tgcttttttg ggagacg 97

<210> 4
 <211> 56
 <212> DNA
 <213> Artificial Sequence
 <220><223> Forward primer for 1st PCR reaction
 <400> 4
 acactctttc cctacacgac gctcttccga tctcttgaaa aagtggcacc gagtcg 56

<210> 5
 <211> 57
 <212> DNA
 <213> Artificial Sequence
 <220><223> Forward primer for 1st PCR reaction

<400> 5
acactctttc cctacacgac gctcttccga tcttcttgaa aaagtggcac cgagtcg 57

<210> 6
<211> 58
<212> DNA
<213> Artificial Sequence

<220><
223> Forward primer for 1st PCR reaction

<400> 6
acactctttc cctacacgac gctcttccga tctcgcttga aaaagtggca ccgagtcg 58

<210> 7
<211> 61
<212> DNA
<213> Artificial Sequence

<220><223> Reverse primer for 1st PCR reaction

<400> 7
gtgactggag ttcagacgtg tgctcttccg atctttaagt cgagtaagct gaccgctgaa 60
g 61

<210> 8
<211> 62
<212> DNA
<213> Artificial Sequence

<220><
223> Reverse primer for 1st PCR reaction

<400> 8
gtgactggag ttcagacgtg tgctcttccg atctattaag tcgagtaagc tgaccgctga 60
ag 62

<210> 9
<211> 63
<212> DNA
<213> Artificial Sequence

<220><223> Reverse primer for 1st PCR reaction

<400> 9
gtgactggag ttcagacgtg tgctcttccg atcttattaa gtcgagtaag ctgaccgctg 60

aag

63

<210> 10

<211> 49

<212> DNA

<213> Artificial Sequence

<220><223> Forward primer for 2nd PCR reaction

<400> 10

aatgatacgg cgaccaccga gatctacaca cactctttcc ctacacgac

49

<210> 11

<211> 45

<212> DNA

<213> Artificial Sequence

<220><223> Reverse primer for 2nd PCR reaction

<400> 11

caagcagaag acggcatacg agatgtgact ggagttcaga cgtgt

45